

Э. ЛЕМАН

**ПРОВЕРКА
СТАТИСТИЧЕСКИХ
ГИПОТЕЗ**

Э. ЛЕМАН

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Перевод с английского
Ю. В. ПРОХОРОВА

Издание второе, исправленное



МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1979

Леман Э. Проверка статистических гипотез. — М.: Наука. Главная редакция физико-математической литературы, 1979, 408 с.

Один из основных разделов математической статистики — теория проверки статистических гипотез — исчерпывающим образом изложен в книге Э. Лемана, известного американского специалиста.

Статистические критерии приводятся вместе с указанием как тех областей, где их применение вполне оправдано, так и тех областей, где применение требует осторожности. Большое внимание уделено построению критериев, в том или ином смысле наилучших. Ценность книги увеличивается большим количеством примеров из разнообразных областей (техники, биологии, медицины и др.), удачно подобранными задачами и обширным списком аннотированной литературы.

Книга может быть полезна физикам, инженерам и другим специалистам, интересующимся сознательным и критическим применением мощного аппарата математической статистики.

Э. Леман

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

М., 1979 г., 408 стр. с илл.

Редактор *А. В. Прохоров*

Техн. редактор *Л. В. Лихачева*. Корректор *А. Л. Ипатова*

ИБ № 11460

Слано в набор 16.01.79. Подписано к печати 08.06.79. Бумага 60×90^{1/16}, тип. № 1. Литературная гарнитура. Высокая печать. Условн. печ. л. 25,5. Уч.-изд л. 28,58. Тираж 11 000 экз. Заказ № 25. Цена книги 2 р. 30 к.

Издательство «Наука»

Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени Ленинградская типография № 2 имени Евгении Соколовой «Союзполиграфпрома» при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли, 198052, Ленинград, Л 52, Измайловский проспект, 29

ОГЛАВЛЕНИЕ

Предисловие	6
Глава 1. Общая проблема решения	9
1. Статистические выводы и статистические решения	9
2. Точная постановка проблемы решения	10
3. Рандомизация. Выбор эксперимента	14
4. Оптимальные процедуры	16
5. Инвариантность и несмещенность	18
6. Байесовские и минимаксные процедуры	21
7. Метод максимума правдоподобия	23
8. Полные классы	25
9. Достаточные статистики	26
10. Задачи	31
11. Литературные ссылки	36
Глава 2. Вероятностные основы	40
1. Вероятность и мера	40
2. Интегрирование	44
3. Статистики и подполя	47
4. Условное математическое ожидание и условная вероятность	49
5. Условные распределения вероятностей	54
6. Характеристика достаточности	59
7. Экспоненциальные семейства	62
8. Задачи	66
9. Литературные ссылки	71
Глава 3. Равномерно наиболее мощные критерии	73
1. Постановка проблемы	73
2. Фундаментальная лемма Неймана — Пирсона	77
3. Распределения с монотонным отношением правдоподобия	82
4. Сравнение экспериментов	90
5. Доверительные границы	93
6. Обобщение фундаментальной леммы	99
7. Двусторонние гипотезы	104
8. Наименее благоприятные распределения	107
9. Проверка гипотез о среднем и дисперсии в нормальной совокупности	111
10. Последовательный критерий отношений вероятностей	114
11. Мощность и средний размер выборки для последовательного критерия отношений вероятностей	118
12. Оптимальное свойство последовательных критериев отношений вероятностей	122
13. Задачи	128
14. Литературные ссылки	138

Глава 4. Несмещенность: теория и первые применения	143
1. Несмещенность при проверке гипотез	143
2. Однопараметрические экспоненциальные семейства	144
3. Подобие и полнота	148
4. РНМ несмещенные критерии для экспоненциальных семейств со многими параметрами	153
5. Сравнение двух пуассоновских или биномиальных совокупностей	159
6. Проверка независимости в 2×2 таблицах	162
7. Критерий знаков	166
8. Задачи	169
9. Литературные ссылки	177
Глава 5. Несмещенность: применения к нормальным распределениям; доверительные интервалы	180
1. Статистики, не зависящие от достаточной статистики	180
2. Проверка гипотез о параметрах нормального распределения	183
3. Сравнение средних и дисперсий двух нормальных распределений	189
4. Доверительные интервалы и семейства критериев	194
5. Несмещенные доверительные множества	197
6. Регрессия	202
7. Критерии, основанные на перестановках	204
8. Наиболее мощные критерии перестановок	206
9. Рандомизация как основа статистических выводов	211
10. Критерии перестановок и рандомизация	214
11. Проверка независимости в двумерном нормальном распределении	219
12. Задачи	222
13. Литературные ссылки	235
Глава 6. Инвариантность	237
1. Симметрия и инвариантность	237
2. Максимальные инварианты	240
3. Наиболее мощные инвариантные критерии	243
4. Выборочный контроль по количественному признаку	247
5. Почти инвариантность	251
6. Несмещенность и инвариантность	255
7. Ранговые критерии	260
8. Задача сравнения двух выборок	264
9. Гипотеза симметрии	270
10. Инвариантные доверительные множества	273
11. Доверительные границы для функций распределения	276
12. Задачи	278
13. Литературные ссылки	293
Глава 7. Линейные гипотезы	293
1. Каноническая форма	298
2. Линейные гипотезы и метод наименьших квадратов	303
3. Критерии однородности	306
4. Двойная классификация: одно наблюдение в клетке	311
5. Двойная классификация: m наблюдений в клетке	314
6. Регрессия	318
7. Модель II: одинарная классификация	323
8. Классификации по подчиненности («гнездовые» классификации)	327
9. Многомерная линейная гипотеза	331
10. Редукция с учетом инвариантности	334
11. Применения	338
12. χ^2 -критерий: простая гипотеза и неограниченные альтернативы	342
13. χ^2 -критерии и критерии отношения правдоподобия	345

14. Задачи	351
15. Литературные ссылки	363
Глава 8. Принцип минимакса	367
1. Критерии с гарантированной мощностью	367
2. Примеры	371
3. Максимальные критерии и инвариантность	375
4. Теорема Ханта — Стейна	378
5. Наиболее строгие критерии	382
6. Задачи	384
7. Литературные ссылки	389
Дополнение	391
1. Отношения эквивалентности; группы	391
2. Сходимость распределений	392
3. Доминированные семейства распределений	396
4. Теорема о слабой компактности	398
Именной указатель	401
Предметный указатель	403

ПРЕДИСЛОВИЕ

Математическая теория проверки гипотез, в которой критерии появляются как решения точно поставленных оптимальных проблем, была создана Ю. Нейманом и Э. Пирсоном в 30-х годах и с тех пор получила значительное развитие. Цель настоящей книги состоит в том, чтобы дать систематическое изложение этой теории и тесно связанной с ней теории доверительных множеств вместе с их главными приложениями. В число последних входят стандартные задачи (как для одной, так и для двух выборок), касающиеся нормального, биномиального и пуассоновского распределений; некоторые аспекты дисперсионного анализа и анализа регрессии (линейные гипотезы); некоторые многомерные проблемы и последовательный анализ. Книга знакомит также с непараметрическими критериями, хотя здесь теоретический подход не развит полностью. Большая область методологии, а именно класс методов, основанных на рассмотрении больших выборок, в частности критерий χ^2 и критерий отношения правдоподобия, по существу, не нашла отражения. Подход к ним и применяемый математический аппарат столь отличны от используемого в этой книге, что адекватное изложение потребовало бы отдельного тома. Краткие указания на теорию этих критериев содержатся в конце главы 7.

В настоящее время теория проверки гипотез подвергается важным изменениям по меньшей мере в двух направлениях. Одно из них возникло на почве осознания того факта, что стандартная формулировка чрезмерно упрощает задачу. Как следствие, теория была пересмотрена с точки зрения статистических решающих функций, введенных Вальдом. Хотя эти исследования и проливают новый свет на классическую теорию, они в целом подтверждают ее открытия. Я сохранил формулировку Неймана — Пирсона в основной части книги, но включил обсуждение понятий общей теории решений в главу 1 с тем, чтобы дать более широкую основу для оправдания некоторых результатов. Эти понятия служат также базисом для развития теорий проверки гипотез и доверительных интервалов.

Значительно более важным является тот факт, что многие из рассматриваемых задач, которые обычно формулируют в терминах проверки гипотез, в действительности оказываются множественными проблемами решения. В этих последних отклонение основной гипотезы приводит к необходимости выбора одного из нескольких решений. Развитие подходящих процедур в этих задачах представляет собой в настоящее время одну из важнейших проблем статистики. Однако поскольку значительная часть работ в этом направлении имеет пробный характер, я предпочел указывать традиционные критерии даже в тех случаях, где большинство применений создает впечатление необходимости более тонких процедур; в этих случаях я ограничивался лишь предостережением о том, что подход справедлив в таких-то пределах. Действительно, кажется правдоподобным, что упомянутые критерии, по причине своей простоты, останутся полезными и тогда, когда будет закончена более полная теория множественных решений.

Естественную математическую основу систематического изложения способов проверки гипотез образует теория меры в абстрактных пространствах. Поскольку вводные курсы по теории действительного переменного или теории меры часто ограничиваются одной лишь мерой Лебега, в разделах 1 и 2 главы 2 приведены краткие формулировки, относящиеся к общей теории. В самом деле, большая часть книги может быть прочитана без знания теории меры, если символ $\int p(x) d\mu(x)$ интерпретировать или как $\int p(x) dx$, или как $\Sigma p(x)$ и если не обращать внимания на теоретико-множественную природу некоторых доказательств. Формально знание статистики не предполагается — все необходимые понятия вводятся с самого начала. С другой стороны, так как читатель обычно уже имеет некоторый опыт обращения со статистическими методами, то применения каждого метода описываются в общих терминах — конкретные примеры с числовыми данными не приводятся. Их можно найти в большинстве стандартных учебников.

В конце каждой главы помещены задачи, многие из которых сопровождаются указаниями на путь решения. Здесь встречаются упражнения, и дополнительные примеры, и задачи, вводящие в новые темы. В конце каждой главы имеется также аннотированный список литературы, послужившей источником как идей, так и отдельных результатов. Примечания не ставят целью привести основные результаты цитированных статей, а лишь указывают значение каждой из них для соответствующей главы. Составляя эти списки, я не стремился к полноте, а пытался дать полезный «путеводитель» по литературе.

Набросок этой книги появился в 1949 году в форме записок лекций, составленных Колином Блитом (Colin Blyth) в летнем семестре в Калифорнийском университете. С тех пор я излагал те или иные части материала в Колумбийском, Принстонском и Стэнфордском университетах, и несколько раз — в Калифорнийском университете. Мне очень помогли замечания студентов, и я сожалею, что не могу поблагодарить их каждого в отдельности. На различных этапах работы я получил много полезных предложений от В. Готски (W. Gautschi), А. Хейланда (A. Høyland), Л. Д. Сэвиджа (L. J. Savage) и особенно от К. Стрибел (C. Striebel); критическое чтение ею предпоследнего варианта рукописи привело ко многим усовершенствованиям. Также я хотел бы отметить с благодарностью ту пользу, которую мне принесли многие продолжительные беседы с Чарлзом Стейном (Charles Stein).

Приятно отметить поддержку этой работы Бюро Морских исследований; без нее, вероятно, книга не была бы написана. Наконец, я хотел бы поблагодарить Дж. Рубалкава (J. Rubalcava), которая печатала и перепечатывала различные черновые варианты рукописи с неослабевающим терпением, точностью и скоростью.

Беркли, Калифорния,
июнь 1959

Э. Л. Леман

ГЛАВА 1

ОБЩАЯ ПРОБЛЕМА РЕШЕНИЯ

1. Статистические выводы и статистические решения

«Сырьем» для статистического исследования служит совокупность результатов наблюдений; эти результаты представляют собой значения случайных величин X , распределение которых P_θ хотя бы частично неизвестно. О параметре θ предполагается известным лишь то, что это один из элементов некоторого множества Ω (*пространства параметров*). Статистические выводы используют материал наблюдений для получения информации относительно распределения X (или значения параметра θ , от которого оно зависит). Чтобы прийти к более точной формулировке задачи, мы рассмотрим цели статистических выводов.

Необходимость статистического анализа возникает из того факта, что распределение X , а следовательно, и некоторые черты ситуации, лежащей в основе математической модели, неизвестны. Следствием этого недостаточного знания является неопределенность в выборе наилучшего поведения. Чтобы формализовать понятие «поведения», допустим, что необходимо выбрать одно из нескольких действий. Результаты наблюдений, давая информацию относительно исходного распределения, тем самым указывают путь к наилучшему решению. Задача состоит в определении правила, ставящего в соответствие каждому результату наблюдений решение, которое должно быть принято. С точки зрения математики, такое правило есть функция δ , которая каждому возможному значению x случайных величин приписывает решение $d = \delta(x)$, т. е. функция, областью определения которой является множество значений X и областью значений — множество возможных решений.

Чтобы понять, как следует выбирать δ , необходимо сравнить последствия использования различных правил. С этой целью предположим, что принятие решения d при условии, что распределение X равно P_θ , приводит к *потере*, которая может быть выражена неотрицательным числом $L(\theta, d)$. Тогда средний убыток, возникающий от применения δ в длинном ряду повторений

эксперимента, приближенно равен математическому ожиданию $E_\theta [L(\theta, \delta(X))]$, вычисленному в предположении, что распределением X является P_θ . Это математическое ожидание, которое зависит от решающего правила δ и распределения P_θ , называется *функцией риска* для δ и будет обозначаться $R(\theta, \delta)$. Принимая решение по результатам наблюдений, мы заменяем первоначальную задачу выбора решения d при функции потерь $L(\theta, d)$ задачей выбора δ при функции потерь $R(\theta, \delta)$ *).

Приведенные выше рассуждения наводят на мысль о том, что целью статистики является выбор решающей функции, минимизирующей возникающий риск. Как будет видно впоследствии, такое определение цели статистики недостаточно точно, чтобы быть вполне осмысленным; его надлежащая интерпретация в действительности оказывается одной из основных проблем теории.

2. Точная постановка проблемы решения

Методы, требующиеся для решения какой-либо отдельной статистической задачи, существенно зависят от трех элементов, которые ее определяют: класса $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$, которому, по предположению, принадлежит распределение X ; структуры пространства D возможных решений d ; формы функции потерь L . Чтобы прийти к конкретным результатам, необходимо, следовательно, сделать определенные предположения относительно этих элементов. С другой стороны, если рассматривать теорию, как нечто большее, чем собрание изолированных результатов, предположения должны быть достаточно широкими, с тем чтобы или быть часто применимыми или выделять классы задач, допускающих единообразный подход.

Рассмотрим сначала класс $\mathcal{P}$. Точные числовые предположения относительно вероятностей или вероятностных распределений обычно не могут быть оправданы. Однако часто можно допустить, что некоторые события имеют равные вероятности, а некоторые другие статистически независимы. Предположения другого типа касаются относительной величины некоторых инфинитезимальных вероятностей, например вероятностей наступления событий во временном или пространственном интервале при длине интервала, стремящейся к нулю. Приводимые ниже классы распределений построены на основании предположений только такого рода и, следовательно, могут встретиться в большом числе случаев.

*) Иногда принимают во внимание не только математическое ожидание потерь, связанных с решающим правилом, но и другие его качества.

Биномиальное распределение $b(p, n)$ с

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x=0, \dots, n; \quad 0 \leq p \leq 1, \quad (1)$$

есть распределение числа успехов в n независимых испытаниях, в каждом из которых вероятность успеха равна p .

Распределение Пуассона $P(\tau)$ с

$$P(X=x) = \frac{\tau^x}{x!} e^{-\tau}, \quad x=0, 1, \dots; \quad 0 < \tau, \quad (2)$$

это — распределение числа событий, происходящих в фиксированном временном или пространственном интервале, при условии, что вероятность наступления более чем одного события в короткий промежуток времени является величиной более высокого порядка малости, чем вероятность наступления одного события, и что числа событий в непересекающихся интервалах статистически независимы. При этих предположениях процесс появления событий называется пуассоновским процессом*).

Нормальное распределение $N(\xi, \sigma^2)$ с плотностью вероятности

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(x-\xi)^2\right], \quad -\infty < x, \quad \xi < \infty; \quad 0 < \sigma, \quad (3)$$

при весьма общих предположениях, которые уточняются центральной предельной теоремой, является приближенным распределением суммы большого числа независимых случайных величин, каждая из которых мала по сравнению со всей суммой.

Мы рассмотрим теперь структуру пространства решений D . Большое разнообразие возможностей указывается следующими примерами.

Пример 1. Пусть $X_1, \dots, X_n$ — выборка из совокупности с одним из распределений (1)–(3), т. е. величины X независимы и одинаково распределены по одному из этих законов. Пусть θ обозначает p, τ или пару (ξ, σ) соответственно, и пусть $\gamma = \gamma(\theta)$ — действительная функция θ .

(I) Если хотя бы решить, превосходит γ некоторое определенное значение γ_0 или нет, то выбирают одно из решений: $d_0: \gamma > \gamma_0$ или $d_1: \gamma \leq \gamma_0$. В отдельных приложениях эти решения могут соответствовать принятию или отклонению партии товаров или опытного образца самолета (по его пригодности к испытательным полетам), или нового способа лечения, предложенного в качестве лучшего по сравнению с принятым, и т. д. Функция потерь зависит от области применений. Типично, однако, что потеря равна нулю, если выбрано правильное решение, в то время как при неправильном решении потери $L(\gamma, d_0)$ и $L(\gamma, d_1)$ являются возрастающими функциями $|\gamma - \gamma_0|$.

(II) В силу сложности пространства решений значительно более трудной проблемой является отыскание числовой оценки для γ .

* Эти процессы рассматриваются, например, в книгах: В. Феллер, Введение в теорию вероятностей и ее приложения, т. I, Мир, 1967, Дж. Дуб, Вероятностные процессы, ИЛ, 1956.

Здесь решением d , которое принимает статистик, служит действительное число — оценка γ , и функцией потерь может быть, например, $L(\gamma, d) = v(\gamma)\omega(|d - \gamma|)$, где ω — строго возрастающая функция ошибки $|d - \gamma|$.

(III) Промежуточное положение занимает случай выбора между тремя альтернативными решениями — d_0 : $\gamma < \gamma_0$, d_1 : $\gamma > \gamma_1$, d_2 : $\gamma_0 \leq \gamma \leq \gamma_1$, например, принятием нового способа лечения, отклонением его или рекомендацией дальнейшего его изучения.

Различия, иллюстрируемые этим примером, лежат в основе одного из главных способов классификации статистических методов. Задачи с двумя решениями, такие, как (I), формулируются обычно в терминах *проверки гипотезы*, которая должна быть принята или отвергнута (см. главу 3). Мы будем иметь дело, в значительной степени, с теорией именно этого класса задач. Другой важной ветвью статистики является теория *точечных оценок*, занимающаяся такими проблемами, как (II). Исследование *процедур со многими решениями*, таких, как (III), начато только в последние годы.

Пример 2. Предположим, что даны выборки X_{ij} , $j = 1, \dots, n_i$, из нормальных совокупностей $N(\xi_i, \sigma^2)$, $i = 1, \dots, s$.

(I) Рассмотрим сначала случай $s = 2$. Поставим вопрос о том, имеется ли существенное различие между этими двумя совокупностями. Вопрос имеет ту же самую структуру, что и задача (III) в предыдущем примере. Здесь производится выбор между тремя решениями — d_0 : $|\xi_2 - \xi_1| \leq \Delta$, d_1 : $\xi_2 > \xi_1 + \Delta$, d_2 : $\xi_2 < \xi_1 - \Delta$, где Δ — наперед заданное число. Аналогичная проблема, включающая $s + 1$ возможное решение, возникает в общем случае s совокупностей. Здесь нужно произвести выбор между решениями d_0 : $\max |\xi_i - \xi_j| \leq \Delta$ (совокупности не отличаются существенно одна от другой) и d_k : $\max |\xi_i - \xi_j| > \Delta$ и ξ_k — наибольшее из средних.

(II) Близкой к рассмотренной является задача о расположении совокупностей в порядке возрастания соответствующих средних значений.

(III) Пусть задан некоторый стандарт ξ . Задача состоит в том, чтобы определить, превосходят ли этот стандарт какие-либо из средних значений (и какие именно).

Пример 3. Рассмотрим два распределения, для определенности — два распределения Пуассона $P(\tau_1)$, $P(\tau_2)$. Предположим, что τ_1 меньше τ_2 , но в остальных параметрах τ неизвестны. Пусть случайные величины $z_1, \dots, z_n$ распределены независимо, каждая в соответствии с законом $P(\tau_1)$ или $P(\tau_2)$. Требуется каждую из величин z отнести к одному из двух классов, соответственно теоретическому распределению. Здесь функция потерь могла бы равняться числу ошибок в классификации, умноженному на некоторую функцию от τ_1 и τ_2 . Степень сложности, которой может достигать эта задача, а также возникающие трудности, как математические, так и концептуальные, иллюстрируются усилиями антропологов разделить человечество на некоторое число однородных рас с помощью изучения частот различных групп крови и других генетических характеристик.

Все рассмотренные до сих пор проблемы могли бы быть названы *проблемами действия*. В каждой из них предполагалось, что если θ известно, то лишь одно решение d будет правильным, т. е. при любом данном θ существует единственное d , для которого $L(\theta, d) = 0$. Однако не все статистические задачи ставятся так отчетливо. Часто вопрос заключается в удобном сводном

описании данных или в указании содержащейся в них информации относительно неизвестного параметра или распределения. Эта информация используется затем в различных обстоятельствах, но она не является единственной основой какого-либо определенного решения. В таких случаях в статистической проблеме ударение делается на выводах, а не на решении, хотя формально проблема может рассматриваться как проблема решения: для этого результат вывода должен интерпретироваться как решение, которое следует принять. Важный класс подобных задач составляют оценки интервалами *); для иллюстрации ниже приводится пример.

Пример 4. Пусть $X = (X_1, \dots, X_n)$ является выборкой из нормальной совокупности $N(\xi, \sigma^2)$, и пусть решение состоит в выборе интервала $[L, L]$ и утверждении, что он содержит ξ . Предположим далее, что процедура решения допускает только интервалы $[\underline{L}(X), \bar{L}(X)]$, математическое ожидание длины которых при всех ξ и σ не превосходит $k\sigma$, где k — некоторая заранее заданная константа. Функцию потерь можно было бы положить равной нулю, если решение правильное, а в остальных случаях сделать зависящей от относительного расположения интервала и истинного значения ξ . В таком случае данному распределению $N(\xi, \sigma^2)$ соответствует несколько правильных решений.

Нам остается обсудить выбор функции потерь. Вероятно, что из трех элементов, определяющих статистическую проблему, этот уточняется с наибольшим трудом. Даже в простейшем случае, когда все потери в конечном счете сводятся к денежным, трудно рассчитывать на то, что мы сумеем оценить все как немедленные, так и более отдаленные следствия некоторого действия. Часто можно упростить положение, принимая во внимание только некоторые черты функции потерь. В качестве иллюстрации возьмем пример 1 (I). Пусть $L(\theta, d_0) = a$ для $\gamma(\theta) \leq \gamma_0$ и $L(\theta, d_1) = b$ для $\gamma(\theta) > \gamma_0$. Тогда функция риска равна

$$R(\theta, \delta) = \begin{cases} aP_\theta\{\delta(X) = d_0\}, & \text{если } \gamma \leq \gamma_0, \\ bP_\theta\{\delta(X) = d_1\}, & \text{если } \gamma > \gamma_0, \end{cases} \quad (4)$$

и, таким образом, зависит только от двух вероятностей ошибок, веса при которых могут быть подобраны с учетом сравнительного значения этих ошибок. Аналогично в примере 3 можно ограничиться числом ошибок в классификации.

К сожалению, подобное естественное упрощение не всегда возможно. Отсутствие точных сведений приводит к необходимости выбирать функцию потерь некоторым условным образом. При этом математическая простота рассматривается как важное обстоятельство. Так, в проблемах точечной оценки, подобных рассмотренной в примере 1 (II), оценивая действительную

*) Более обычную формулировку в терминах доверительных интервалов см. в главе 3 раздел 5 и в главе 5 разделы 4 и 5.

функцию $\gamma = \gamma(\theta)$, обычно используют квадрат ошибки или, несколько более общим образом, выражение

$$L(\theta, d) = v(\theta)(d - \gamma)^2. \quad (5)$$

Будучи весьма простым математически, оно может восприниматься как приближение к истинной функции потерь в предположении, что при каждом фиксированном θ функция $L(\theta, d)$ дважды дифференцируема по d , $L(\theta, \gamma(\theta)) = 0$ для всех θ , и что ошибка невелика.

Часто случается, что в рамках одной задачи возможны различные типы потерь, для которых трудно подыскать общую единицу измерения. Вернемся еще раз к примеру 1 (I) и допустим, что γ_0 представляет собою значение γ , получающееся при использовании в некоторой ситуации (например, в медицине, сельском хозяйстве или промышленности) стандартного способа действий. Задача состоит в том, чтобы сравнить какой-то новый способ (с неизвестным γ) с указанным стандартным способом. Отклонение нового метода, когда он превосходит старый, или принятие его в противоположном случае влекут за собою, очевидно, весьма различные последствия. Иногда в таких случаях бывает удобно оперировать по отдельности с несколькими компонентами, скажем $L_1, L_2, \dots, L_r$. Предположим, в частности, что $r = 2$ и что L_1 соответствует более существенной стороне дела. Тогда можно ограничить эту компоненту риска, т. е. наложить условие вида

$$EL_1(\theta, \delta(X)) \leq \alpha \quad (6)$$

и при этом минимизировать другую компоненту. Эту процедуру иллюстрирует пример 4. Длина интервала $[\underline{L}, \bar{L}]$ (измеренная в σ -единицах) является одной компонентой функции потерь. Другую образуют потери, возникающие тогда, когда этот интервал не покрывает истинное значение ξ .

3. Рандомизация. Выбор эксперимента

Данное выше описание решающего правила в некоторых отношениях остается слишком узким. Мы предполагали, что каждому возможному значению случайных величин соответствует ровно одно решение. Вместо этого удобно допустить выбор одного из нескольких решений в соответствии с заданными вероятностями или, более общим образом, в соответствии с некоторым распределением вероятностей на пространстве решений. Это распределение зависит, конечно, от наблюдаемого значения x . Один из способов описания такой рандомизированной процедуры состоит в том, что задаются нерандомизированная процедура, зависящая от X , и случайная величина Y , значения которой

лежат в пространстве решений и условное распределение которой при данном x не зависит от θ .

Хотя интуитивно может показаться, что такая дополнительная рандомизация не имеет значения, это допущение большей свободы выбора не может нанести вреда. Если интуитивные сомнения будут правильны, то окажется, что оптимальные процедуры всегда имеют нерандомизированный характер. В действительности введение рандомизированных процедур создает важное математическое упрощение: расширение класса функций риска делает его выпуклым. К тому же, существуют проблемы, в которых некоторые характеристики функции риска (например, ее максимум) могут быть лучше использованы, если допустить и рандомизированные процедуры.

Другое предположение, которое молчаливо до сих пор принималось, состоит в том, что был выбран определенный план эксперимента. Поэтому известно, какие именно наблюдения будут произведены. Однако статистические рассуждения, связанные с планированием эксперимента, несколько не менее важны, чем касающиеся анализа этого эксперимента. В частности, еще до того, как предпринято исследование, необходимо решить, сколько наблюдений должно быть произведено, учитывая, что потери от неправильных решений не должны быть чрезмерными. Часто оказывается, что требуемый объем выборки зависит от неизвестного распределения и, следовательно, не может быть фиксирован заранее. Вместо этого объем определяется в зависимости от наблюдений, и решение о продолжении эксперимента принимается *последовательно* на каждом его шаге на основе имеющихся к этому моменту результатов наблюдений.

Пример 5. Допустим, что по выборке $X_1, \dots, X_n$ из нормального распределения $N(\xi, \sigma^2)$ желают оценить среднее значение ξ . Здесь функция риска для оценки (например, математическое ожидание квадрата ее отклонения от ξ) зависит от σ . При больших σ выборка содержит мало информации в том смысле, что два распределения $N(\xi_1, \sigma^2)$ и $N(\xi_2, \sigma^2)$ с фиксированной разностью $\xi_2 - \xi_1$ становятся неразличимыми при $\sigma \rightarrow \infty$; вследствие этого риск стремится к бесконечности. Наоборот, риск стремится к нулю при $\sigma \rightarrow 0$, так как в этом случае среднее почти точно известно. Таким образом, число наблюдений, необходимое для поддержания риска на заданном уровне, неизвестно. Однако после того как произведено некоторое количество наблюдений, возможно оценить σ^2 и, следовательно, определить число требуемых дополнительно наблюдений.

Пример 6. Пусть относительно последовательности испытаний с постоянной вероятностью успеха p желают выяснить, какое из двух неравенств $p \leq 1/2$ или $p > 1/2$ верно. Как правило, мы будем быстро приходить к решению, если p близко к 0 или 1, когда практически все наблюдения приводят к одному и тому же результату. В то же время для промежуточных значений p потребуются выборка большего объема. Это различие может частично уравновешиваться тем обстоятельством, что для промежуточных значений потери, возникающие из-за неправильного решения, предположительно менее серьезны, чем для значений, близких к краям.

Пример 7. Возможность последовательного определения объема выборки важна не только потому, что распределения P_θ имеют большую или меньшую информационную ценность, но и потому, что это же самое верно для результатов наблюдений. Рассмотрим, например, выборку из равномерного на интервале $(\theta - 1/2, \theta + 1/2)$ распределения*), и пусть задача состоит в оценке θ . Здесь нет различия в количестве информации, доставляемой различными распределениями P_θ . Однако по выборке $X_1, X_2, \dots, X_n$ можно почти точно определить θ , если $\max |X_j - X_i|$ достаточно близок к единице; если же $\max |X_j - X_i|$ близок к нулю, то вся выборка дает информацию несколько не большую, чем одно наблюдение. Снова необходимый объем выборки должен определяться последовательно.

Определение надлежащего размера выборки представляет собою, за исключением простейших ситуаций, лишь одну из сторон задачи планирования эксперимента. Вообще говоря, мы должны решить не только вопрос о числе наблюдений, но и то, какого они должны быть типа. Формально все эти вопросы можно включить в общую проблему решения, как она описана в начале этого раздела. С этой целью X интерпретируется как множество всех наблюдаемых случайных величин; вводятся понятия решения о продолжении и решения об остановке эксперимента на различных этапах; определяется (в случае продолжения), какого типа величину следует наблюдать на следующем шаге; наконец, стоимость наблюдений включается в функцию потерь. Однако, несмотря на эту формальную возможность, отыскание оптимального плана эксперимента в конкретных ситуациях представляет собою обычно задачу более трудную, чем отыскание оптимального решающего правила для данного эксперимента. Эта задача решена лишь в небольшом числе случаев. Здесь в первую очередь мы будем заниматься проблемой в том виде, в котором она предстает перед нами после того, как эксперимент выбран, и только в нескольких случаях мы постараемся сравнить различные планы эксперимента.

4. Оптимальные процедуры

В конце первого раздела было сказано, что цель статистической теории — отыскание такой решающей функции δ , которая минимизирует функцию риска

$$R(\theta, \delta) = E_\theta[L(\theta, \delta(X))]. \quad (7)$$

К сожалению, минимизирующее δ зависит, вообще говоря, от неизвестного значения θ . Возьмем, например, некоторое специальное решение d_0 и рассмотрим решающую процедуру $\delta(x) \equiv d_0$, в соответствии с которой решение d_0 принимается независимо от исхода эксперимента. Предположим, что d_0 является правиль-

*) Это распределение определено в задаче 1 в конце главы.

ным решением для некоторого θ_0 , т. е. $L(\theta_0, d_0) = 0$. Тогда δ минимизирует риск для θ_0 , поскольку $R(\theta_0, \delta) = 0$, но, вероятно, ценной значительного риска для других значений θ .

При отсутствии решающей функции, которая минимизирует риск для всех значений θ , проблема остается математически неопределенной, так как неясно, что подразумевается под наилучшей процедурой. Хотя представляется невозможным дать определение оптимальности, которое было бы подходящим во всех случаях, приводимые ниже два подхода часто оказываются удовлетворительными.

Отсутствие оптимального решающего правила связано с тем, что процедура уделяет слишком большое внимание отдельному

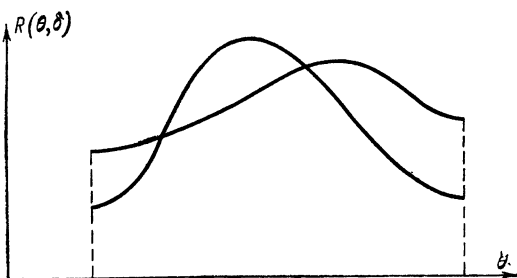


Рис. 1.

значению параметра, пренебрегая многими другими. Это приводит к мысли ограничиться решающими процедурами, в какой-то степени беспристрастными; возможно, что внутри такого класса найдутся процедуры с равномерно наименьшим риском. Два ограничения подобного рода — инвариантность и несмещенность, будут обсуждены в следующем разделе.

Вместо сужения класса процедур возможен и другой подход к проблеме. Рассмотрим функции риска, соответствующие двум различным решающим правилам δ_1 и δ_2 . Если $R(\theta, \delta_1) < R(\theta, \delta_2)$ при всех θ , то ясно, что δ_1 предпочтительнее δ_2 , так как δ_1 приводит к меньшему риску, каково бы ни было истинное значение θ . Однако положение неясно, когда графики этих функций пересекаются, как на рис. 1. Необходим принцип, который в подобных случаях устанавливает, какая из двух функций риска предпочтительнее, т. е. который упорядочивает множество функций риска. Оптимальной будет та процедура, функция риска которой является наилучшей в смысле этого порядка.

Некоторые критерии, которые были предложены для упорядочения функций риска, будут рассмотрены в разделе 6.

Слабым пунктом теории оптимальных процедур, как она намечена выше, является ее зависимость от не относящихся к

существо дела обстоятельств: ограничения класса процедур, принципа упорядочения и от сведений о функции потерь и распределениях наблюдаемых случайных величин. Эти сведения в приложениях часто бывают или недоступны, или ненадежны. Подобные трудности могут вызвать сомнение в ценности оптимальной теории, покоящейся на столь шатком основании. Но они в принципе не отличаются от трудностей, возникающих в любых применениях математики к действительности. Математические формулировки всегда включают упрощение и аппроксимацию; поэтому на решения, полученные с их использованием, нельзя полагаться без дополнительной проверки. В нашем случае проверка состоит из всеобъемлющей оценки действия процедуры, даваемой теорией, и в исследовании ее чувствительности к отклонениям от предположений, при которых она была рекомендована.

Отмеченные трудности могут быть частично преодолены рассмотрением одной и той же проблемы в разных постановках. Если различные критерии оптимальности приводят к одному и тому же решению, то оно будет наилучшим с нескольких точек зрения, и, следовательно, правдоподобно, что оно будет удовлетворительным вообще. В противоположном случае этот метод указывает слабые и сильные стороны различных решений и тем самым может привести к некоторой компромиссной процедуре. Аналогично, чувствительность процедуры к отклонениям от предположений, при которых она была выведена, может быть испытана, например, отбрасыванием одного из предположений и сравнением процедуры, полученной в более широкой модели, с первоначальной.

5. Инвариантность и несмещенность *)

Естественное определение беспристрастности появляется в ситуациях, симметричных по отношению к интересующим нас значениям параметра. *От процедуры требуется в этом случае симметричность по отношению к этим значениям параметра.*

Пример 8. Предположим, что сравниваются два способа действий, каждый из которых применяется n раз. Соответствующие результаты наблюдений $X_{11}, \dots, X_{1n}$ и $X_{21}, \dots, X_{2n}$ представляют собою выборки из $N(\xi_1, \sigma^2)$ и $N(\xi_2, \sigma^2)$ соответственно. Допускаются три решения — d_0 : $|\xi_2 - \xi_1| \leq \Delta$, d_1 : $\xi_2 > \xi_1 + \Delta$, d_2 : $\xi_2 < \xi_1 - \Delta$, и принятие решения d_j при условии, что истинно d_i , приводит к потере w_{ij} . Если способы действий сравниваются только в терминах ξ и никакие посторонние соображения не принимаются во внимание, то потери являются симметричными по отношению к этим способам, так что $w_{01} = w_{02}$, $w_{10} = w_{20}$, $w_{12} = w_{21}$. Предположим теперь, что индексы 1 и 2, отмечающие

*) Понятия, обсуждаемые здесь в рамках общей теории решений, будут изучены в более специализированной форме в последующих главах. Поэтому настоящий раздел может быть опущен при первом чтении.

способы, меняют местами и соответственно меняют индексы у X , ξ и решений d_1 и d_2 . Это изменяет смысл символов, но формальная проблема решения, в силу ее симметрии, остается прежней. Поэтому естественно требовать соответствующей симметрии от процедуры δ , т. е. требовать, чтобы $\delta(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}) = d_0, d_1$ или d_2 одновременно с $\delta(x_{21}, \dots, x_{2n}, x_{11}, \dots, x_{1n})$. Если бы это условие не было выполнено, то решение того, какая из выборок имеет большее среднее, зависело бы от совершенно несущественного обстоятельства — нумерации выборок. Аналогичные замечания применимы и к другим типам симметрии, встречающимся в этой задаче.

Пример 9. Рассмотрим выборку $X_1, \dots, X_n$ из распределения с плотностью $\sigma^{-1}f[(x - \xi)/\sigma]$. Пусть проблема состоит в оценке параметра сдвига ξ , скажем математического ожидания X_k при функции потерь $(d - \xi)^2/\sigma^2$, т. е. равной квадрату ошибки, выраженной в σ -единицах. Предположим, что первоначально результаты наблюдений выражены в футах, и пусть $X'_i = aX_i$, где $a = 12$, — соответствующие значения в дюймах. В преобразованной задаче плотность равна $\sigma'^{-1}f[(x' - \xi')/\sigma']$, где $\xi' = a\xi$, $\sigma' = a\sigma$. Так как $(d' - \xi')^2/\sigma'^2 = (d - \xi)^2/\sigma^2$, то формально проблема не меняется. Та же самая процедура, которая используется для первоначальных наблюдений, пригодна после преобразования. Для $\xi' = a\xi$, где ξ' измерено в дюймах, она дает оценку $\delta(aX_1, \dots, aX_n)$. Переводя последнее выражение в футы, мы видим, что если потребовать независимость результата от выбора единицы измерения, то δ должно удовлетворять условию масштабной инвариантности:

$$\delta(aX_1, \dots, aX_n)/a = \delta(X_1, \dots, X_n).$$

Общим математическим выражением симметрии является инвариантность относительно подходящей группы преобразований. Мы скажем, что группа G преобразований g пространства выборок оставляет статистическую проблему решения инвариантной, если она удовлетворяет следующим условиям:

(I) Она оставляет инвариантным семейство распределений $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$, т. е. для любого возможного распределения P_θ случайных величин X распределение gX , скажем $P_{\theta'}$, также принадлежит $\mathcal{P}$. Предполагается, что соотношение $\theta' = \bar{g}\theta$ отображает Ω на Ω и притом взаимно однозначным образом *).

(II) Для каждого $g \in G$ существует отображение $g^* = h(g)$ пространства решений D на себя, определяемое гомоморфизмом h (т. е. h удовлетворяет условию $h(g_1g_2) = h(g_1)h(g_2)$) и такое, что функция потерь не меняется при этом отображении, т. е.

$$L(\bar{g}\theta, g^*d) = L(\theta, d).$$

При этих предположениях преобразованная задача в терминах $X' = gX$, $\theta' = \bar{g}\theta$ и $d' = g^*d$ формально идентична первоначальной задаче в терминах X , θ и d . Решающая процедура δ первоначальной задачи остается пригодной и после преобразования. Интерпретируя преобразование как замену координатной системы и, следовательно, замену названий элементов, мы выбрали бы, наблюдая x' , решение, которое в новой системе записывается

*) Термин «на Ω » используется для указания того, что $\bar{g}\Omega = \Omega$, т. е. что для любого θ' в Ω существует θ с $\bar{g}\theta = \theta'$.

как $\delta(x')$, а в старой — как $g^{*-1}\delta(x')$. Если принимаемое решение не должно зависеть от выбора системы координат, то последнее решение должно совпадать с первоначальным δ , т. е. процедура должна удовлетворять условию *инвариантности*:

$$\delta(gx) = g^*\delta(x) \quad \text{для всех } x \in X, g \in G. \quad (8)$$

Рассмотрения, основанные на инвариантности, применимы лишь тогда, когда проблема обнаруживает некоторую симметрию. Альтернативная формализация понятия беспристрастности, применяемая к проблемам другого типа, дается приводимым ниже условием несмещенности. Предположим задачу такой, что для каждого θ существует единственное правильное решение и что каждое решение является правильным для некоторого θ . Допустим далее, что $L(\theta_1, d) = L(\theta_2, d)$ для всех d , если правильные решения для θ_1 и θ_2 совпадают. Тогда функция потерь $L(\theta, d')$ зависит только от принятого решения, скажем d' , и правильного решения d . Потеря может быть обозначена поэтому $L(d, d')$ и эта последняя функция указывает, как далеко отстоит d' от d . При перечисленных предположениях решающая функция δ называется несмещенной, если при всех θ и d'

$$E_{\theta}L(d', \delta(X)) \geq E_{\theta}L(d, \delta(X)),$$

где индекс θ указывает распределение, по отношению к которому берутся математические ожидания, и d обозначает решение, правильное для θ . Таким образом, δ является несмещенным, если в среднем $\delta(X)$ подходит ближе к правильному решению, чем к любому из неправильных. Распространяя приведенное определение на любые проблемы решения, δ называют *несмещенным* правилом, если для всех θ и θ'

$$E_{\theta}L(\theta', \delta(X)) \geq E_{\theta}L(\theta, \delta(X)). \quad (9)$$

Пример 10. Предположим, что в задаче оценки действительного параметра θ с помощью доверительных интервалов (пример 4) потеря равна 0 или 1 в зависимости от того, покрывает или не покрывает интервал $[\underline{L}, \bar{L}]$ истинное значение θ . Тогда система интервалов $[\underline{L}(X), \bar{L}(X)]$ является несмещенной, если вероятность покрытия истинного значения не меньше вероятности покрытия любого другого значения.

Пример 11. Рассмотрим задачу с двумя решениями такую, как в примере 1 (I). Пусть ω_0 и ω_1 обозначают множества значений θ , для которых правильными решениями являются d_0 и d_1 соответственно. Предположим, что потеря равна 0, если выбирается правильное решение, а в противном случае определяется равенствами: $L(\theta, d_0) = a$ для $\theta \in \omega_1$ и $L(\theta, d_1) = b$ для $\theta \in \omega_0$. Тогда

$$E_{\theta}L(\theta', \delta(X)) = \begin{cases} aP_{\theta}\{\delta(X) = d_0\}, & \text{если } \theta' \in \omega_1, \\ bP_{\theta}\{\delta(X) = d_1\}, & \text{если } \theta' \in \omega_0, \end{cases}$$

так что (9) приводится к неравенству

$$aP_{\theta}\{\delta(X) = d_0\} \geq bP_{\theta}\{\delta(X) = d_1\} \quad \text{для } \theta \in \omega_0$$

и к обратному неравенству для $\theta \in \omega_1$. Так как $P_\theta \{\delta(X) = d_0\} + P_\theta \{\delta(X) = d_1\} = 1$, то условие несмещенности (9) превращается в неравенства

$$\begin{aligned} P_\theta \{\delta(X) = d_1\} &\leq \frac{a}{a+b} \quad \text{для } \theta \in \omega_0, \\ P_\theta \{\delta(X) = d_1\} &\geq \frac{a}{a+b} \quad \text{для } \theta \in \omega_1. \end{aligned} \quad (10)$$

Пример 12. В задаче оценки действительной функции $\gamma(\theta)$ при потере, равной квадрату ошибки, условие несмещенности записывается в виде

$$E_\theta [\delta(X) - \gamma(\theta')]^2 \geq E_\theta [\delta(X) - \gamma(\theta)]^2 \quad \text{для всех } \theta, \theta'.$$

Добавляя и вычитая внутри скобок в обеих частях неравенства $h(\theta) = E_\theta \delta(X)$, мы преобразуем его к виду

$$[h(\theta) - \gamma(\theta')]^2 \geq [h(\theta) - \gamma(\theta)]^2 \quad \text{для всех } \theta, \theta'.$$

Если $h(\theta)$ является одним из значений функции γ , то последнее соотношение выполняется тогда и только тогда, когда

$$E_\theta \delta(X) = \gamma(\theta). \quad (11)$$

В теории точечных оценок (11) обычно принимается в качестве определения несмещенности. Кроме весьма патологических случаев, равенство (11) необходимо и достаточно для того, чтобы δ удовлетворяло (9) (см. задачу 2 в конце главы).

6. Байесовские и минимаксные процедуры

Мы вернемся теперь к обсуждению способов упорядочения решающих процедур и их функций риска. Одно из таких упорядочений получается в предположении, что в повторных экспериментах сам параметр является случайной величиной Θ . Если предположить для простоты, что его распределение имеет плотность $\rho(\theta)$, то полная средняя потеря от применения решающей процедуры δ равна

$$r(\rho, \delta) = \int E_\theta L(\theta, \delta(X)) \rho(\theta) d\theta = \int R(\theta, \delta) \rho(\theta) d\theta \quad (12)$$

и, чем меньше $r(\rho, \delta)$, тем лучше δ . Оптимальной является процедура, минимизирующая $r(\rho, \delta)$. Она называется *байесовским решением* данной проблемы, соответствующим априорной плотности ρ . Минимум $r(\rho, \delta)$ называется *байесовским риском* для ρ .

К сожалению, чтобы использовать этот принцип, надо предположить не только то, что θ — случайная величина, но и то, что распределение ее известно. Последняя предпосылка на практике обычно не оправдывается. Правая часть (12) может быть иначе истолкована как взвешенное среднее рисков. В частности, для $\rho(\theta) = 1$ это есть площадь под графиком функции риска. При такой интерпретации выбор весовой функции ρ показывает

степень значимости, которую экспериментатор приписывает различным значениям θ .

При отсутствии априорной информации относительно θ имеет смысл рассматривать максимум функции риска как ее наиболее важную характеристику. Тогда из двух функций риска предпочтительнее та, которая имеет меньший максимум. Оптимальными процедурами являются те, которые минимизируют максимальный риск, т. е. обладают *минимаксным* свойством. Так как максимум указывает самые тяжелые (в среднем) потери, которые могут возникнуть при использовании данной процедуры, то минимаксное решение, по сравнению с другими, дает самую надежную защиту от больших потерь. Рис. 2 показывает, что

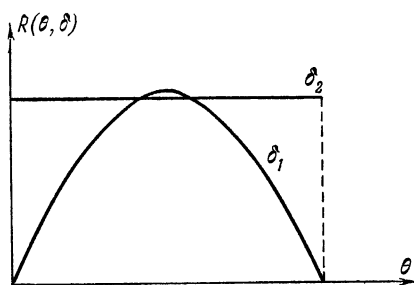


Рис. 2.

иногда такой принцип может оказаться неблагоприятным: в большинстве случаев мы предпочли бы правило δ_1 , хотя его функция риска имеет больший максимум.

Вероятно, что наиболее типичен случай, промежуточный между двумя только что описанными. С одной стороны, мы могли встречаться в прошлом с таким же или сходным экспериментом и потому можем иметь некоторое представление

об ожидаемых значениях θ . С другой стороны, эта информация не является ни достаточно точной, ни достаточно надежной, чтобы оправдать байесовский подход. В этой обстановке представляется желательным использовать доступную информацию, не опираясь на нее, однако, настолько, чтобы в случае ее неточности столкнуться с опасностью катастрофически высоких потерь. Для достижения этой цели можно задать верхнюю границу риска и рассматривать лишь такие процедуры δ , для которых

$$R(\theta, \delta) \leq C \quad \text{для всех } \theta \quad (13)$$

(здесь константа C должна быть больше максимального риска C_0 минимаксной процедуры, так как в противном случае не найдется процедуры, удовлетворяющей (13)). Гарантировав, таким образом, что риск ни при каких обстоятельствах не выйдет из-под контроля, экспериментатор может без опасений использовать свои сведения о ситуации. Последние могут быть следствием как теоретических соображений, так и прошлого опыта. Экспериментатор может следовать своим влечениям и угадывать распределение p параметра θ . Это приводит к выбору процедуры δ (*ограниченного байесовского решения*), которая минимизирует риск

(12) для этого априорного распределения (при условии (13)). Чем более мы уверены в выборе ρ , тем большим можно выбрать C . При этом, управляя большими значениями риска в случае неудачного предположения о ρ , мы уменьшаем риск при хорошей догадке.

Вместо непосредственного задания упорядочения, мы можем постулировать требования, которым оно должно удовлетворять.

Были исследованы различные системы таких условий *). Это привело к общему заключению, что этим требованиям удовлетворяют лишь упорядочения, основанные на значении байесовского риска по отношению к некоторому априорному распределению θ .

7. Метод максимума правдоподобия

Другим подходом, основанным на соображениях, отличных от использованных в предыдущих разделах, является метод максимума правдоподобия. В большом числе проблем он приводит к рациональным процедурам и до сих пор играет доминирующую роль при отыскании новых критериев или оценок. Предположим на время, что X может принимать только конечное или счетное число значений $x_1, x_2, \dots$ с вероятностями $P_\theta(x) = P_\theta\{X = x\}$. Допустим, что мы хотим определить истинное значение θ по наблюдаемому x . Естественно посмотреть для каждого возможного θ , сколь вероятным было бы x , если бы θ было истинным значением параметра. Чем выше эта вероятность, тем более мы склоняемся к мнению, что рассматриваемое θ является истинным; т. е., иными словами, тем более правдоподобным представляется нам это значение. В соответствии со сказанным, выражение $P_\theta(x)$, рассматриваемое при фиксированном x как функция от θ , называется *функцией правдоподобия* для θ . Чтобы подчеркнуть это изменение точки зрения, обозначим функцию правдоподобия $L_x(\theta)$. Предположим теперь, что мы сталкиваемся с проблемой действия, включающей конечное или счетное число возможных решений, и что эта проблема сформулирована не в терминах потерь, как обычно, а в терминах выигрышей: выигрыш равен нулю, если принято неправильное решение, и равен $a(\theta) > 0$, если принято правильное решение и θ — истинное значение параметра. Тогда представляется естественным снабдить значения функции правдоподобия $L_x(\theta)$ весами, равными соответствующим выигрышам, и определить значение θ , максимизирующее произведение $a(\theta)L_x(\theta)$, а затем принять

*) См., например, Savage, The Foundations of Statistics, New York, Wiley, 1954, и раздел 4.3 книги: Blackwell, Girshick, Theory of Games and Statistical Decisions, New York, Wiley, 1954. (Русский перевод: Блэквелл и Гиршик, Теория игр и статистических решений, ИЛ, 1958.)

такое решение, которое было бы правильным, если бы найденное θ было верным*). Сделанные замечания применимы и к случаю непрерывных распределений, когда $P_\theta(x)$ обозначает плотность вероятности. Приведенные мотивировки неприменимы непосредственно к задаче оценки непрерывного параметра, однако этот случай можно рассматривать как предельный.

В задачах точечной оценки $a(\theta)$ предполагают обычно не зависящим от θ . Тем самым в качестве оценки θ используют значение, максимизирующее функцию правдоподобия $L_x(\theta)$, т. е., как говорят, *оценку наибольшего правдоподобия*. Другой интересный случай — проблемы с двумя решениями (ср. пример 1 (I)). Пусть ω_0 и ω_1 обозначают множества значений θ , для которых правильными являются решения d_0 и d_1 соответственно, и пусть $a(\theta) = a_0$ при $\theta \in \omega_0$ и $a(\theta) = a_1$ при $\theta \in \omega_1$. Мы принимаем решение d_0 или d_1 в зависимости от того, будет ли $a_1 \sup_{\theta \in \omega_1} L_x(\theta)$

меньше или больше $a_0 \sup_{\theta \in \omega_0} L_x(\theta)$, т. е. будет ли

$$\frac{\sup_{\theta \in \omega_0} L_x(\theta)}{\sup_{\theta \in \omega_1} L_x(\theta)} > \frac{a_1}{a_0} \quad \text{или} \quad \frac{\sup_{\theta \in \omega_1} L_x(\theta)}{\sup_{\theta \in \omega_0} L_x(\theta)} < \frac{a_1}{a_0}. \quad (14)$$

Эта процедура приводит к *критерию отношения правдоподобия**)*.

Хотя принцип максимального правдоподобия не основывается на каких-либо ясно выраженных соображениях оптимальности, он применяется весьма успешно и приводит к удовлетворительным процедурам во многих конкретных задачах. Было показано также, что в широком классе задач метод максимального правдоподобия обладает свойством асимптотической оптимальности при неограниченно растущем объеме выборки***). С другой стороны, существуют примеры, в которых метод максимального правдоподобия более чем бесполезен. Действительно, в этих примерах он столь плох, что можно найти лучший ответ, вовсе не прибегая к наблюдениям (см. задачу 18 в главе 6).

*) Вариант такого подхода был предложен в работе Lindley, Statistical inference, J. Roy. Stat. Soc., Ser. B, vol. XI (1953), 30—76.

**) Данное определение слегка отличается от обычного, где в знаменателе левой части (14) $\sup$ берется на множестве $\omega_0 \cup \omega_1$. Оба определения совпадают, когда левая часть (14) не превосходит 1. Процедуры совпадают, следовательно, если $a_1 < a_0$.

***) По этому вопросу см., например, Wald, Tests of statistical hypotheses concerning several parameters when the number observations is large, Trans. Am. Math. Soc. 54 (1943), 426—482, и Le Cam, On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. Univ. Calif. Publ. Statistics, 1 (1953), 277—330. См. также библиографию по этому вопросу в Дополнении к книге: Г. Крамер, Математические методы статистики, 2-е изд., М., Мир, 1975.

8. Полные классы

Ни один из описанных до сих пор подходов не представляется вполне надежным в том смысле, что рекомендуемая им процедура не является с необходимостью удовлетворительной. Имеются задачи, в которых можно указать процедуру δ_0 с равномерно наименьшим риском среди всех несмещенных или инвариантных процедур; в то же время существует процедура δ_1 , не обладающая этими свойствами беспристрастности, но лучшая δ_0 (см. задачи 14 и 16). Как было показано раньше, минимаксные процедуры также могут оказаться нежелательными, а успех в применении байесовских или ограниченных байесовских решений связан с наличием априорной информации, которая, если и доступна, то, как правило, не бывает очень надежной. Действительно, кажется вероятным, что при отсутствии надежной априорной информации ни один из принципов, приводящих к единственному ответу, не может быть вполне удовлетворительным.

Это выдвигает возможность, хотя бы в качестве первого шага, не настаивать на единственном решении, но исследовать, сколь сильно может быть редуцирована проблема без потери, относящейся к делу информации. Мы уже видели, что решающая процедура δ может быть отклонена по той причине, что существует процедура δ' , *доминирующая* ее в том смысле, что

$$\begin{aligned} R(\theta, \delta') &\leq R(\theta, \delta) \quad \text{для всех } \theta, \\ R(\theta, \delta') &< R(\theta, \delta) \quad \text{для некоторого } \theta. \end{aligned} \tag{15}$$

В подобном случае δ будет называться *недопустимой*. С другой стороны, δ будет называться *допустимой*, если не существует доминирующей ее δ' . Класс $\mathcal{C}$ решающих процедур называют *полным*, если для любого δ , не принадлежащего $\mathcal{C}$, существует доминирующее его δ' в $\mathcal{C}$. Полный класс называют *минимальным*, если никакой его подкласс не является полным. Если минимальный полный класс существует (что типично), то он совпадает с множеством всех допустимых процедур.

Удобно ввести также следующий вариант понятия полного класса. Класс $\mathcal{C}$ называют *существенно полным*, если для каждой процедуры δ найдется процедура δ' из $\mathcal{C}$, для которой при всех θ $R(\theta, \delta') \leq R(\theta, \delta)$. Очевидно, что каждый полный класс является существенно полным. Действительно, оба определения расходятся только в обращении с эквивалентными решающими правилами, т. е. с решающими правилами, имеющими идентичные функции риска. Если δ принадлежит минимальному полному классу $\mathcal{C}$, то любое эквивалентное решающее правило также входит в $\mathcal{C}$. С другой стороны, минимальный существенно

полный класс должен содержать не более одного представителя от каждого множества эквивалентных между собою процедур.

Минимальный, существенно полный класс соответствует максимальной в некотором смысле редукции проблемы решения. С другой стороны, нет никаких причин рассматривать какую-либо из отклоненных (т. е. не входящих в этот класс) процедур. Для каждой из них найдется такая же или лучшая процедура в $\mathcal{C}$. В то же время этот класс невозможно уменьшить, так как, если даны две процедуры из $\mathcal{C}$, то иногда одна из них будет лучше, чем другая, и без дополнительной информации нельзя сказать, которая из них предпочтительнее.

Первоначальной заботой статистики было точное определение процедур (или классов процедур) в различных специальных задачах. Наиболее интенсивно изучались проблемы оценки и проблемы выбора между двумя решениями (проверка гипотез), теория которых составляет предмет настоящей книги. Однако некоторые заключения возможны без такой специализации. В частности, при весьма общих условиях были доказаны два результата, касающиеся структуры полных классов и минимаксных процедур*):

(I) Совокупность всех байесовских решений и их пределов образует полный класс.

(II) Минимаксные процедуры являются байесовскими решениями по отношению к *наименее благоприятному* априорному распределению, т. е. такому, которое максимизирует байесовский риск. Минимаксный риск равен этому байесовскому риску. Более общим образом, если не существует наименее благоприятного априорного распределения, а имеется лишь последовательность, для которой байесовские риски стремятся к максимуму, то минимаксные процедуры являются пределами соответствующей последовательности байесовских решений.

9. Достаточные статистики

Мы видели в предыдущем разделе, что минимальный полный класс соответствует максимальной возможной редукции проблемы решения, осуществляемой без потери информации. Часто возможно произвести сокращение (редукцию) данных, применимое одновременно ко всем проблемам, касающимся заданного класса $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ распределений случайной величины X . Эта редукция состоит в отбрасывании той части данных, кото-

*) Точные формулировки и доказательства даны в книге: Wald, Statistical Decision Functions, New York, Wiley, 1950. (Русский перевод: В а л ь д А., Статистические решающие функции, Сб. Позиционные игры, М., 1967, 300—522.)

рая не содержит информации относительно неизвестного распределения P_θ и, следовательно, бесполезна для любой проблемы решения, касающейся θ .

Пример 13. Производятся независимые испытания с постоянной неизвестной вероятностью p успеха. Пусть X_i равно 1, если i -е испытание закончилось успехом, и равно 0 в противном случае. Выборка $(X_1, \dots, X_n)$ показывает число успехов и номера испытаний, в которых они произошли. Последняя информация ничего не проясняет в вопросе о значении p . Коль скоро общее число успехов $t = \sum X_i$ дано, то каждое из $\binom{n}{t}$ возможных расположений

этих успехов имеет одну и ту же вероятность относительно к тому, каково p . Отсюда следует, что, зная $\sum X_i$, но не зная ни отдельных X_i , ни p , возможно с помощью таблицы случайных чисел построить множество случайных величин $X'_1, \dots, X'_n$, совместное распределение которых такое же, как и у $X_1, \dots, X_n$. Следовательно, информация, содержащаяся в X_i , такая же, как в $\sum X_i$ и таблице случайных чисел.

Пример 14. Пусть $X_1, \dots, X_n$ независимы и распределены нормально с математическим ожиданием 0 и дисперсией σ^2 . Тогда условное распределение выборочной точки на каждой из сфер $\sum X_i^2 = \text{const}$ равномерно, каково бы ни было σ^2 . Зная $\sum X_i^2$, можно построить эквивалентную выборку $X'_1, \dots, X'_n$, если располагать механизмом, который наудачу наносит точки на сферу.

Вообще, статистика T называется *достаточной* для семейства $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ (или достаточной для θ , если из контекста ясно, о каком Ω идет речь), когда условное распределение X при условии $T = t$ не зависит от θ . При слабых ограничениях*) отсюда следует, так же как и в указанных двух примерах, что нет необходимости использовать сами первоначальные результаты наблюдений X . Если мы наблюдаем лишь T (вместо X), то класс доступных решающих процедур не становится уже. Пусть для каждого значения t статистики T X_t обозначает случайную величину, распределение которой совпадает с условным распределением X при $T = t$. Такая величина может быть, по крайней мере теоретически, построена с помощью надлежащего случайного механизма. Если сначала наблюдение дает значение t для T , а затем значение x' для X_t , то случайная величина X' , получаемая в результате такой двукратной процедуры, имеет то же распределение, что и X . Таким образом, для любой заданной процедуры, основанной на X , возможно построить эквивалентную ей, основанную на X' . Эта последняя может рассматриваться как рандомизированная процедура, использующая только T . Следовательно, если разрешена рандомизация (а мы всюду будем предполагать, что это так и есть), то ограничение достаточными статистиками не приводит к потере общности.

*) Они связаны с трудностями, касающимися понятия условной вероятности. По этому поводу см. разделы 3—5 главы 2.

Необходимость вычислять условное распределение X при данном t с тем, чтобы установить — является ли статистика T достаточной или нет, оказывается неудобной. Простой способ проверки дает следующий *критерий факторизации*.

Рассмотрим сначала случай дискретного X , и пусть $P_\theta(x) = P_\theta\{X=x\}$. Тогда необходимое и достаточное условие для того, чтобы статистика T была достаточной для θ , состоит в существовании факторизации

$$P_\theta(x) = g_\theta[T(x)] h(x), \quad (16)$$

где первый множитель может зависеть от θ , а от x зависит только через $T(x)$, в то время как второй множитель не зависит от θ .

Предположим, что (16) выполняется, и пусть $T(x) = t$. Тогда $P_\theta\{T=t\} = \sum P_\theta(x')$, где сумма берется по x' с $T(x') = t$, и условная вероятность

$$P_\theta\{X=x | T=t\} = P_\theta(x)/P_\theta\{T=t\} = h(x)/\sum h(x')$$

не зависит от θ . Обратно, если это условное распределение не зависит от θ и равно, скажем, $k(x, t)$, то $P_\theta(x) = P_\theta\{T=t\} \times k(x, t)$, так что (16) выполняется.

Пример 15. Пусть $X_1, \dots, X_n$ независимы и одинаково распределены по закону Пуассона (2). Тогда

$$P_\tau(x_1, \dots, x_n) = \frac{\tau^{\sum x_i} e^{-n\tau}}{\prod_{j=1}^n x_j!},$$

откуда следует, что $\sum X_i$ является достаточной статистикой для τ .

Рассмотрим случай, когда распределение X непрерывно и имеет плотность $p_\theta^X(x)$. Пусть X и T — векторы, скажем $X = (X_1, \dots, X_n)$ и $T = (T_1, \dots, T_r)$. Допустим, что на пространстве выборок существует функция $Y = (Y_1, \dots, Y_{n-r})$ такая, что преобразование

$$(x_1, \dots, x_n) \leftrightarrow (T_1(x), \dots, T_r(x), Y_1(x), \dots, Y_{n-r}(x)) \quad (17)$$

взаимно однозначно в соответствующей области и что совместная плотность T и Y существует и связана с плотностью X обычной формулой *)

$$p_\theta^X(x) = p_\theta^{T,Y}(T(x), Y(x)) \cdot |J|, \quad (18)$$

*) Условия регулярности, при которых верно (18), даны Tukey, A smooth invertibility theorem, Ann. Math. Stat. 29 (1958), 581—584; см. также Lehmann and Scheffé, On the problem of similar regions, Proc. Nat. Acad. Sci. 33 (1947), 382—386.

где J — якобиан $(T_1, \dots, T_r, Y_1, \dots, Y_{n-r})$ по отношению к $(x_1, \dots, x_n)$. Так, в примере 14 $T = \sqrt{\sum X_i^2}$ и $Y_1, \dots, Y_{n-1}$ могут быть определены как полярные координаты выборочной точки. По совместной плотности $p_{\theta}^{T, Y}(t, y)$ T и Y мы находим условную плотность Y при данном $T = t$:

$$p_{\theta}^{Y|t}(y) = \frac{p_{\theta}^{T, Y}(t, y)}{\int p_{\theta}^{T, Y}(t, y') dy'}, \quad (19)$$

предполагая, что знаменатель отличен от нуля.

Так как в условном распределении при данном t меняются только величины Y , то T достаточна для θ , если условное распределение Y при данном t не зависит от θ . Предположим, что T удовлетворяет (19). Тогда, аналогично дискретному случаю, статистика T будет достаточной тогда и только тогда, когда плотность допускает факторизацию вида

$$p_{\theta}^X(x) = g_{\theta}[T(x)] h(x) \quad (20)$$

(см. задачу 19). Ниже приводятся два примера, иллюстрирующие применение критерия в этом случае. В обоих примерах существование функции Y , удовлетворяющей (17) — (19), предполагается, но не доказывается. Как мы увидим позже (раздел 6 главы 2), это предположение не является необходимым для справедливости критерия факторизации.

Пример 16. Пусть $X_1, \dots, X_n$ независимы и имеют нормальную плотность распределения

$$p_{\xi, \sigma}(x) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\xi}{\sigma^2} \sum x_i - \frac{n}{2\sigma^2} \xi^2\right).$$

Критерий факторизации показывает, что статистика $(\sum X_i, \sum X_i^2)$ достаточна для (ξ, σ) .

Пример 17. Пусть $X_1, \dots, X_n$ независимы и имеют равномерное распределение $R(0, \theta)$ на отрезке $(0, \theta)$. Тогда $p_{\theta}(x) = \theta^{-n} u(\max x_i, \theta)$, где $u(a, b) = 1$ при $a \leq b$ и $u(a, b) = 0$ при $a > b$. Следовательно, статистика $\max X_i$ достаточна для θ .

Другой критерий достаточности устанавливает прямую связь между этим понятием и некоторыми из основных понятий теории статистических решений. Как в байесовском подходе предположим, что неизвестный параметр θ является случайной величиной Θ с априорным распределением. Для простоты примем, что оно имеет плотность $\rho(\theta)$. Тогда, если T достаточна, то условное распределение Θ при данном $X = x$ зависит только от $T(x)$. Обратно, если $\rho(\theta) \neq 0$ для всех θ и если условное распределение Θ при данном x зависит только от $T(x)$, то T достаточна для θ .

Действительно, при сделанных предположениях совместная плотность X и Θ равна $p_{\theta}(x)\rho(\theta)$. Если T достаточна, то из (20) вытекает, что условная плотность Θ при данном x зависит только от $T(x)$. Обратно, предположим, что для некоторого априорного распределения с $\rho(\theta) \neq 0$ при всех θ условное распределение Θ при данном x зависит только от $T(x)$. Тогда

$$\frac{p_{\theta}(x)\rho(\theta)}{\int p_{\theta'}(x)\rho(\theta')d\theta'} = f_{\theta}[T(x)],$$

откуда для $p_{\theta}(x)$ получаем выражение, показывающее достаточность T .

Каждое байесовское решение зависит только от условного распределения Θ при данном x (см. задачу 8) и, следовательно, от $T(x)$. Так как обычно байесовские решения и их пределы образуют существенно полный класс, то это же верно и в применении к решающим процедурам, основанным на T . К этому заключению мы пришли также (и более непосредственно) в начале настоящего раздела.

Ограничиваясь достаточными статистиками, мы сокращаем данные. Желательно это сокращение сделать максимальным. Чтобы проиллюстрировать различные возможности, рассмотрим еще раз пример 13. Каково бы ни было целое m , меньшее n , пара (T_1, T_2) , где $T_1 = \sum_{i=1}^m X_i$ и $T_2 = \sum_{i=m+1}^n X_i$, образует достаточную статистику, так как условное распределение $X_1, \dots, X_n$ при условии $T_1 = t_1, T_2 = t_2$ не зависит от p . По той же причине и полная выборка $(X_1, \dots, X_n)$ также является достаточной статистикой. Однако $T = \sum_{i=1}^n X_i$ приводит к большему сокращению, чем

любая из вышеприведенных статистик и чем вообще любая, которая может быть построена. Достаточная статистика называется *минимальной*, если данные не допускают дальнейшей редукции без потери свойства достаточности. Для биномиального распределения, например, минимальной достаточной статистикой будет $\sum_{i=1}^n X_i$ (задача 17). Этим иллюстрируется тот факт, что достаточная статистика, найденная с помощью критерия факторизации, в конкретных примерах часто оказывается минимальной *).

*) Способы построения минимальных достаточных статистик (некоторые авторы называют их необходимыми и достаточными статистиками) даны в работах: Lehman and Scheffé, Completeness, similar regions and unbiased estimation, Sankhyā, 10 (1950), 305—340 и Bahadur, Sufficiency and statistical decision functions, Ann. Math. Stat. 25 (1954), 423—462. См. также

10. Задачи

К разделу 2

1. Нижеследующие распределения получаются на основе предположений, сходных с использованными в формулах (1)–(3).

(I) Независимые испытания с постоянной вероятностью успеха p проводятся до тех пор, пока не наступит заданное число m успехов. Если число необходимых испытаний равно $X + m$, то X имеет отрицательно-биномиальное распределение

$$P\{X = x\} = \binom{m+x-1}{x} p^m (1-p)^x, \quad x = 0, 1, 2, \dots$$

(II) Дана последовательность случайных событий такая, что число событий, наступающих в любом промежутке времени длины τ , имеет пуассоновское распределение $P(\lambda\tau)$, и числа событий в неперекрывающихся промежутках времени независимы. Тогда «время ожидания» T , т.е. время от начального момента, скажем $t = 0$, до наступления первого события, имеет экспоненциальную плотность вероятности

$$p(t) = \lambda e^{-\lambda t}, \quad t \geq 0.$$

(Пусть T_1 обозначает промежуток времени от начала отсчета до наступления 1-го события, а T_i ($i \geq 2$) обозначает промежуток времени между $(i-1)$ -м и i -м событиями. Тогда верно, хотя и труднее доказывается, следующее утверждение: $T_1, T_2, \dots$ независимы и распределены одинаково *).)

(III) Пусть точка X выбирается «наудачу» в интервале (a, b) , т.е., иными словами, вероятность попадания X в какой-либо подынтервал (a, b) зависит только от его длины (а не от положения). Тогда X имеет равномерное или прямоугольное распределение $R(a, b)$ с плотностью

$$p(x) = \frac{1}{(b-a)}, \quad a < x < b.$$

[К (II). Если $t > 0$, то $T > t$ в том и только том случае, когда в промежутке $(0, t)$ не происходит ни одного события.]

К разделу 5

2. Несмещенность точечных оценок. Предположим, что пространство параметров Ω связно; γ — непрерывная действительная функция на Ω , не обращающаяся в константу ни на каком открытом подмножестве Ω , математическое ожидание $h(\theta) = E_{\theta} \delta(X)$ непрерывно как функция от θ при любой оценке $\delta(X)$ величины $\gamma(\theta)$. Тогда условие (11) необходимо и достаточно для того, чтобы оценка $\delta(X)$ была несмещенной при функции потерь, равной квадрату ошибки.

[Несмещенность влечет неравенство $\gamma^2(\theta') - \gamma^2(\theta) \geq 2h(\theta)[\gamma(\theta') - \gamma(\theta)]$ при всех θ и θ' . Если θ не является ни относительным минимумом, ни относительным максимумом γ , то существуют точки θ' , сколько угодно близкие к θ , для одних из которых $\gamma(\theta) + \gamma(\theta') \geq 2h(\theta)$, а для других $\gamma(\theta) + \gamma(\theta') \leq 2h(\theta)$. Следовательно, $\gamma(\theta) = h(\theta)$. Из сделанных вначале предположений вытекает, что равенство выполняется и для точек экстремума γ]

Дынкин, О достаточных и необходимых статистиках для семейства распределений вероятностей, ДАН 75 (1950), 161–164.

*) Доказательство см. в книге: Д.ж. Дуб, Вероятностные процессы, ИЛ, 1956, стр. 362.

3. *Медианная несмещенность.* (I). Число m называется медианой случайной величины Y , если $P\{Y \geq m\} \geq \frac{1}{2}$ и $P\{Y \leq m\} \geq \frac{1}{2}$. Для любых действительных a_1, a_2 , для которых $m \leq a_1 \leq a_2$ или $m \geq a_1 \geq a_2$, имеем $E|Y - a_1| \leq E|Y - a_2|$.

(II) Пусть $\delta(X)$ — оценка $\gamma(\theta)$, и пусть $m^-(\theta)$ и $m^+(\theta)$ означают точную нижнюю и точную верхнюю грани медиан $\delta(X)$. Предположим, что m^- и m^+ — непрерывные функции θ .

Пусть Ω связно, а $\gamma(\theta)$ непрерывна и отлична от константы на каждом открытом подмножестве Ω . Тогда оценка $\delta(X)$ является несмещенной по отношению к функции потерь $L(\theta, d) = |\gamma(\theta) - d|$ тогда и только тогда, когда при каждом θ $\gamma(\theta)$ есть медиана $\delta(X)$. Оценки с этим свойством называют медианно-несмещенными.

4. *Отсутствие несмещенных процедур.* Рассмотрим проблему решения, в которой для каждого θ существует единственное правильное решение d . Предположим, что

$$L(\theta, d') = h(\theta) V(d, d') \quad \text{для } \theta \in \omega_d,$$

где ω_d — множество θ , для которых правильно решение d . Тогда, если h принимает по меньшей мере два различных значения на каждом ω_d , то функция риска для любой несмещенной процедуры тождественно равна нулю, т. е., как правило, несмещенные процедуры не существуют. Пусть, например, $X_1, \dots, X_n$ независимы и одинаково распределены с плотностью $(1/a)f_n(x - \xi)/a$ и $\theta = (\xi, a)$. Тогда для ξ не существует оценки, несмещенной по отношению к функции потерь $(\xi - d)^2/a^2$.

5. Пусть $\mathcal{G}$ — класс процедур, замкнутый относительно преобразований группы G (в том смысле, что из $\delta \in \mathcal{G}$ вытекает $g^*\delta g^{-1} \in \mathcal{G}$ для всех $g \in G$). Если существует единственная процедура δ_0 , которая минимизирует риск равномерно в классе $\mathcal{G}$, то δ_0 инвариантна. Если δ_0 единственна с точностью до множеств меры нуль, то она почти инвариантна, т. е. для каждого g удовлетворяет соотношению $\delta(gx) = g^*\delta(x)$ при всех x , кроме, быть может, множества N_g нулевой меры.

6. *Соотношение несмещенности и инвариантности.* (I) Если δ_0 есть единственная (с точностью до множеств нулевой меры) несмещенная процедура с равномерно наименьшим риском, то она почти инвариантна.

(II) Если $\bar{G}$ транзитивна, а G^* коммутативна, и если среди всех инвариантных (почти инвариантных) процедур существует процедура δ_0 , равномерно минимизирующая риск, то она является несмещенной.

[(I) следует из предыдущей задачи и того факта, что несмещенность δ влечет за собою несмещенность $g^*\delta g^{-1}$. (II) По самому определению транзитивности для любых θ, θ' существует такое $\bar{g}$, что $\theta' = \bar{g}\theta$. Следовательно, для любых θ, θ'

$$E_{\theta'} L(\theta', \delta_0(X)) = E_{\theta} L(\bar{g}\theta, \delta_0(X)) = E_{\theta} L(\theta, g^{*-1}\delta_0(X)).$$

В силу коммутативности G процедура $g^{*-1}\delta_0$ инвариантна, так что

$$R(\theta, g^{*-1}\delta_0) \geq R(\theta, \delta_0) = E_{\theta} L(\theta, \delta_0(X)).$$

7. *Противоречащий пример.* Заключение (II) задачи 6 не обязано выполняться, если нарушены предположения относительно G^* и $\bar{G}$. Рассмотрим, например, задачу об оценке среднего ξ в нормальном распределении $N(\xi, \sigma^2)$ при функции потерь $(\xi - d)^2/\sigma^2$. Эта задача инвариантна относительно групп G_1 : $gx = x + b$, $-\infty < b < \infty$, и G_2 : $gx = ax + b$, $0 < a < \infty$, $-\infty < b < \infty$. Наилучшей инвариантной оценкой в обоих случаях служит $\bar{X}$. Однако не существует оценки, не смещенной по отношению к указанной функции потерь.

К разделу 6

8. Структура байесовских решений. (I) Пусть Θ — случайная величина с плотностью $\rho(\theta)$, не наблюдаемая непосредственно. Пусть плотность X при $\Theta = \theta$ равна $p_\theta(x)$. Тогда, если решение $\delta(x)$ при каждом x выбирается так, чтобы минимизировать

$$\int L(\theta, \delta(x)) \pi(\theta | x) d\theta, \quad \text{где } \pi(\theta | x) = \rho(\theta) p_\theta(x) / \int \rho(\theta') p_{\theta'}(x) d\theta'$$

— условная (апостериорная) плотность Θ при данном x , то δ — байесовское решение.

(II) Рассмотрим проблему с двумя решениями и с функцией потерь такой же, как в примере 11. Тогда байесовское решение состоит в выборе d_0 , если

$$aP\{\Theta \in \omega_1 | x\} < bP\{\Theta \in \omega_0 | x\},$$

и выборе d_1 , если верно обратное неравенство. В случае равенства можно выбрать любое решение.

(III) В задаче точечной оценки действительной функции $g(\theta)$ при функции потерь $L(\theta, d) = (g(\theta) - d)^2$ байесовским решением служит $\delta(x) = E[g(\theta) | x]$. Если изменить функцию потерь, полагая $L(\theta, d) = |g(\theta) - d|$, то байесовским решением будет любая из медиан условного распределения $g(\theta)$ при данном x .

(I) Байесовский риск может быть записан в форме

$$\int \left[\int L(\theta, \delta(x)) \pi(\theta | x) d\theta \right] p(x) dx,$$

где

$$p(x) = \int \rho(\theta') p_{\theta'}(x) d\theta'.$$

(II) Условное математическое ожидание $\int L(\theta, d_0) \pi(\theta | x) d\theta$ приводится к $aP\{\Theta \in \omega_1 | x\}$ (и аналогично для d_1).]

9. (I) Рассмотрим пример, в котором рандомизация уменьшает максимум риска: известно, что монета или стандартна (ГР), или имеет герб на обеих сторонах (ГГ). Одну из этих возможностей необходимо выбрать после одного бросания. Потеря равна 1 при неправильном решении и 0 при правильном. Пусть при появлении решетки принимается решение ГР, в противном случае решение принимается «наудачу»; ГР с вероятностью ρ , ГГ с вероятностью $1 - \rho$.

Тогда максимум риска минимален при $\rho = 1/3$.

10. Несмещенность и минимакс. Пусть $\Omega = \Omega_0 \cup \Omega_1$, где Ω_0 и Ω_1 не пересекаются. Рассмотрим задачу с двумя решениями и функцией потерь $L(\theta, d_i) = a_i$ при $\theta \in \Omega_j$ ($j \neq i$) и $L(\theta, d_i) = 0$ при $\theta \in \Omega_i$ ($i = 0, 1$).

(I) Каждая минимаксная процедура является несмещенной.

(II) В предположении, что $P_\theta(A)$ непрерывна по θ при каждом A , и множества Ω_0 и Ω_1 имеют хотя бы одну общую граничную точку, верно утверждение, обратное к (I).

(I) Условие несмещенности в рассматриваемом случае равносильно неравенству $\sup R_\delta(\theta) \leq a_0 a_1 / (a_0 + a_1)$. Это неравенство выполняется для любой минимаксной процедуры, что видно из сравнения с процедурой $\delta(x) = d_0$ или $\delta(x) = d_1$ с вероятностями $a_1/(a_0 + a_1)$ и $a_0/(a_0 + a_1)$ соответственно.

(II) Если θ_0 — общая для Ω_0 и Ω_1 граничная точка, то в силу непрерывности функции риска для любой несмещенной процедуры $R_\delta(\theta_0) = a_0 a_1 / (a_0 + a_1)$. Следовательно, $\sup R_\delta(\theta) = a_0 a_1 / (a_0 + a_1)$.]

11. Инвариантность и минимакс. Пусть проблема инвариантна относительно групп G , $\bar{G}$ и G^* преобразований пространств $\mathcal{X}$, Ω и D соответственно.

Рандомизированная процедура Y_x называется инвариантной, если при всех x и g условное распределение Y_x при данном x совпадает с распределением $g^{*-1}Y_{gx}$.

(I) Рассмотрим решающую процедуру, инвариантную относительно конечной группы $G = \{g_1, \dots, g_N\}$. Если минимаксные процедуры существуют, то среди них найдется инвариантная.

(II) Как показывает нижеследующий пример, для бесконечных групп это утверждение может не выполняться. Пусть пространство параметров Ω состоит из всех элементов θ свободной группы с двумя образующими, т. е. формальных произведений $\pi_1 \dots \pi_n$ ($n = 0, 1, 2, \dots$), где каждое π_i совпадает с одним из элементов a, a^{-1}, b, b^{-1} и где члены вида $aa^{-1}, a^{-1}a, bb^{-1}$ и $b^{-1}b$ выбрасываются. Пустое произведение ($n = 0$) обозначается e . Выборочная точка X получается умножением θ справа на один из четырех элементов a, a^{-1}, b, b^{-1} , выбираемых с вероятностью $1/4$ каждый (и с сокращениями, если они возможны: например, когда случайный множитель равен π_n^{-1}). Проблема оценки θ при функции потерь $L(\theta, d) = 0$, при $d = \theta$ и $L(\theta, d) = 1$ в других случаях, инвариантна относительно умножения X, θ и d слева на произвольную последовательность $\pi_{-m} \dots \pi_{-2}\pi_{-1}$ ($m = 0, 1, \dots$). Инвариантная процедура, которая минимизирует максимум риска, имеет функцию риска $R(\theta, \delta) = 3/4$. Существует, однако, не инвариантная процедура с максимальным риском $1/4$.

[(I) Если Y_x — минимаксная процедура (может быть рандомизированная), то инвариантная минимаксная процедура Y'_x определяется соотношением

$$P(Y'_x = d) = \sum_{i=1}^N P(Y_{g_i x} = g_i^* d) / N.$$

(II) Наилучшая процедура состоит в том, что за оценку θ принимается или $\pi_1 \dots \pi_{k-1}$, если наблюдение дало $\pi_1 \dots \pi_k$ ($k \geq 1$), или с вероятностью $1/4$ один из четырех элементов a, a^{-1}, b, b^{-1} , если наблюдение дало e . Оценка будет правильной, за исключением случая, когда последний элемент X сократился, т. е. будет правильной с вероятностью $\geq 3/4$.]

К разделу 7

12. (I). Пусть X имеет плотность вероятности $p_\theta(x)$, где θ принимает одно из значений $\theta_1, \dots, \theta_n$, так что задача состоит в выборе одного из n решений: $d_1 = \theta_1, \dots, d_n = \theta_n$ с выигрышем $a(\theta_i)$, если $d_i = \theta_i$ и 0 в других случаях. Если θ — случайная величина, принимающая каждое из n значений с вероятностью $1/n$, то байесовское решение (которое максимизирует средний выигрыш) совпадает с оценкой максимального правдоподобия.

(II) Пусть X имеет плотность $p_\theta(x)$ ($0 \leq \theta \leq 1$). Тогда оценка максимального правдоподобия совпадает с модой (точкой максимума) апостериорной плотности Θ при данном x , где априорное распределение Θ равномерно в интервале $(0, 1)$.

13. (I) Пусть $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$. Рассмотрим проблему выбора между решениями $\omega_0: \xi < 0$ и $\omega_1: \xi \geq 0$. Обозначим $\bar{x} = \sum x_i / n$ и $C = (a_1/a_0)^{2/n}$. Процедура максимального правдоподобия приводит к выбору d_0 или d_1 в соответствии с тем, меньше или больше k будет выражение

$$\frac{\sqrt{n}\bar{x}}{\sqrt{\sum (x_i - \bar{x})^2}},$$

где $k = -\sqrt{C-1}$ при $C > 1$ и $k = \sqrt{(1-C)/C}$ при $C < 1$.

(II) В задаче выбора между $\omega_0: \sigma < \sigma_0$ и $\omega_1: \sigma \geq \sigma_0$ процедура максимального правдоподобия приводит к решению d_0 или d_1 в зависимости от того, меньше или больше k будет

$$\sum \frac{(x_i - \bar{x})^2}{n\sigma_0^2},$$

где k — наименьший корень уравнения $Cx = e^{x-1}$ при $C > 1$ и наибольший корень уравнения $x = Ce^{x-1}$ при $C < 1$ (C определяется как в (I)).

К разделу 8

14. *Допустимость несмещенных процедур.* (I) Если в предположениях задачи 10 среди несмещенных процедур имеется процедура с равномерно наименьшим риском, то она допустима.

(II) Нижеследующий пример показывает, что в общем случае процедура с равномерно наименьшим риском не обязана быть допустимой. Пусть X имеет усеченное в нуле пуассоновское распределение, т. е. $P_\theta\{X=x\} = \theta^x e^{-\theta} / [x!(1-e^{-\theta})]$ при $x=1, 2, \dots$. При функции потерь $L(\theta, d) = (d-\theta)^2$ для $\gamma(\theta) = e^{-\theta}$ существует единственная несмещенная оценка, но она не является допустимой.

[(II) Единственная несмещенная оценка $\delta_0(x) = (-1)^{x+1}$ доминируется оценкой $\delta_1(x) = 0$ при x четном и $\delta_1(x) = 1$ при x нечетном.]

15. *Допустимость инвариантных процедур.* Если проблема решения инвариантна относительно конечной группы и если среди инвариантных процедур имеется процедура δ_0 , равномерно минимизирующая риск, то δ_0 допустима.

[Это вытекает из тождества $R(\theta, \delta) = R(\bar{g}\theta, g^*\delta g^{-1})$ и замечания к задаче 11 (I).]

16. (I) Пусть X принимает значения $\theta - 1$ и $\theta + 1$ с вероятностью $1/2$ каждое. Проблема оценки θ при функции потерь $L(\theta, d) = \min(|\theta - d|, 1)$ остается инвариантной при преобразованиях $gX = X + c$, $g\theta = \theta + c$, $g^*d = d + c$. Оценки, принимающие значения $X - 1$ и $X + 1$ с вероятностями p и q (не зависящими от X), равномерно минимизируют риск в классе всех инвариантных оценок.

(II) Утверждение задачи 15 не обязано выполняться для бесконечной G . Это можно установить сравнением наилучших инвариантных оценок п. (I) с оценкой $\delta_1(x) = X + 1$ при $X < 0$ и $\delta_1(x) = X - 1$ при $X \geq 0$.

К разделу 9

17. Рассмотрим последовательность n независимых испытаний с постоянной вероятностью p успеха и положим X_i равным 1 или 0 в зависимости от того, окончилось ли i -е испытание успехом или нет. Тогда $\sum_{i=1}^n X_i$ является минимальной достаточной статистикой.

[Пусть $T = \sum X_i$. Предположим, что $U = f(T)$ достаточна и что $f(k_i) = \dots = f(k_r) = u$. Тогда $P\{T=t|U=u\}$ зависит от p .]

18. (I) Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения $R(0, \theta)$, $0 < \theta < \infty$, и $T = \max(X_1, \dots, X_n)$. Доказать, что T достаточна, двумя способами: один раз, используя непосредственно определение достаточности, а другой — применяя критерий факторизации и допуская существование стигистик Y_i , удовлетворяющих условиям (17) — (19).

(II) Пусть $X_1, \dots, X_n$ — выборка из экспоненциального распределения с плотностью $ae^{-a(x-b)}$ при значениях $x \geq b$ ($0 < a < \infty$, $-\infty < b < \infty$). Применить критерий факторизации для доказательства того, что статистика

$\left(\min (X_1, \dots, X_n), \sum_{i=1}^n X_i \right)$ достаточна для a, b , допуская существование статистики Y_i , удовлетворяющих (17)–(19).

19. Статистика T , удовлетворяющая условиям (17)–(19), достаточна тогда и только тогда, когда для нее выполняется (20).

11. Литературные ссылки

Некоторые из основных понятий статистической теории были введены в первой четверти XIX в. Лапласом в его фундаментальном труде «*Théorie Analytique des Probabilités*» (1812) и Гауссом в его работах по методу наименьших квадратов. Функции риска и потерь упоминались ими при обсуждении проблемы точечной оценки, для которой Гауссом было введено понятие несмещенности. Детальное изложение этой работы и последовавших за ней работ XIX в. дано в книге Чубера (1891).

В конце столетия работы К. Пирсона открыли период интенсивного развития статистических методов. Две области: оценка параметров и проверка гипотез, были особенно тщательно исследованы Р. А. Фишером, Ю. Нейманом и другими. Работа Фишера отражена в его книгах (1925, 1935) и в собрании его избранных статей (1950). Многие из принципиальных идей Неймана собраны в опубликованном цикле лекций (1938). Общее введение в современные приемы оценки параметров и проверки гипотез дается в книгах Уилкса (1944), Крамера (1946), Кендалла (1946), Ван-дер-Вардена (1957) и в более теоретической работе Шметтерера (1956).

Формальное объединение теорий оценки параметров и проверки гипотез, охватывающее также и ряд других проблем, было дано Вальдом. Ему принадлежит единая исчерпывающая формулировка задач общей теории решающих процедур. Полное изложение этой теории, близкой по своим идеям к созданной Дж. фон Нейманом теории игр, имеется в книге Вальда (1950). Часть статей Вальда, посвященных как отдельным аспектам общей теории решений, так и другим разделам статистики, помещена в отдельном томе (1958). Общая теория служит предметом двух недавних книг: Блекуэлла и Гиршика (1954) и Сэвиджа (1954)*).

Блекуэлл и Гиршик (Blackwell D. and Girschick M.)

(1954) *Theory of Games and Statistical Decisions*, New York, John Wiley & Sons. (Русский перевод: Блекуэлл Д. и Гиршик М. А., Теория игр и статистических решений, ИЛ, 1958.)

*) См. также Де Гроот, Оптимальные статистические решения, пер. с англ., М., Мир, 1974. — *Прим. ред.*

Браун (Brown George)

(1947) On small sample estimation, *Ann. Math. Stat.*, vol. 18, pp. 582—585. [Определение медианной несмещенности.]

Вальд (Wald Abraham)

(1939) Contributions to the theory of statistical estimation and testing hypotheses, *Ann. Math. Stat.*, vol. 10, 299—326. [Общая формулировка статистических проблем, содержащая задачи оценки параметров и задачу проверки гипотез в качестве специальных случаев. Обсуждение байесовских и минимаксных процедур.]

(1947) An essentially complete class of admissible decision functions, *Ann. Math. Stat.*, vol. 18, 549—555. [Определение и характеристика полных классов решающих процедур для общих проблем решения. Идеи этой и предшествующей работ были развиты позднее в ряде работ, завершившихся книгой Вальда, упомянутой выше.]

(1950) *Statistical Decision Functions*, New York, John Wiley & Sons. (Русский перевод: Вальд А., Статистические решающие функции. — Сб. Позиционные игры, М., Наука, 1967, стр. 300—522.)

(1958) *Selected Papers in Statistics and Probability* by Abraham Wald, Stanford University Press.

Ван-дер-Варден (van der Waerden B. L.)

(1957) *Mathematische Statistik*, Berlin, Springer-Verlag. (Русский перевод: Б. Л. Ван-дер-Варден, Математическая статистика, ИЛ, 1960.)

Вольфовиц (Wolfowitz I.)

(1951) On ϵ -complete classes of decision functions, *Ann. Math. Stat.*, vol. 22, 461—465.

(1953) The method of maximum likelihood and the Wald theory of decision functions, *Indag. Math.*, vol. 15, 114—119.

Кендалл (Kendall M. G.)

(1946) *The Advanced Theory of Statistics*, vol. 2, London. (Русский перевод 3-го издания, написанного совместно с А. Стюартом: М. Дж. Кендалл, А. Стюарт, Статистические выводы и связи, М., Наука, 1973.)

Колмогоров А. Н.

(1942) Определение центра рассеивания и меры точности по ограниченному числу наблюдений, *ИАН СССР, серия матем.*, 6, 3—32. [Определение достаточности в терминах распределений для параметров.]

Крамер (Cramér H.)

(1946) *Mathematical Methods of Statistics*, Princeton University Press. (Русский перевод: Гаральд Крамер, Математические методы статистики, ИЛ, 1948; 2-е изд., М., Мир, 1975).

Лаплас (Laplace P. S.)

(1812) *Théorie Analytique des Probabilités*, Paris.

Леман (Lehmann E. L.)

(1947) On families of admissible tests, *Ann. Math. Stat.*, vol. 18, 97—104. [Вводится понятие полных классов в связи со специальными статистическими проблемами.]

(1950) Some principles of the theory of hypothesis testing, *Ann. Math. Stat.*, vol. 21, 1—26.

(1951) A general concept of unbiasedness, *Ann. Math. Stat.*, vol. 22, 587—597. [Определение (8); Задачи 2—7 и 14.]

Ле Кам (Le Cam)

(1953) On some asymptotic properties of maximum likelihood estimates and related Bayes estimates, *Univ. Calif. Publ. Statistics*, vol. 1, 277—329, Univ.

Calif. Press, Berkeley and Los Angeles. [Строгое и весьма общее изложение теории оценок наибольшего правдоподобия для случая выборок большого объема с обзором предшествующей литературы по этому вопросу.]

Нейман (Neyman J.)

(1935) Sur un theorema concernente le cosidette statistiche sufficienti, Giorn. Ist. Ital. Att., vol. 6, 320—334. [Получена теорема о факторизации в виде (20).]

(1938) L'estimation statistique traitée comme un problème classique de probabilité, Actualites sci. et ind., № 739, 25—57. [Выдвигается точка зрения, что статистика в первую очередь интересуется тем, каково должно быть поведение в условиях неопределенности, а не определением неизвестных значений параметров, т. е. индуктивным поведением, а не индуктивными выводами.]

(1938) Lectures and Conferences on Mathematical Statistics and Probability, Washington, Graduate School. U. S. Dept. Agriculture, 1st. ed., 1938, 2nd ed., 1952.

(1962) Two breakthroughs in the theory of statistical decision making, Int. Stat. Rev., 30, № 1, 11—27.

(1968) Вводный курс теории вероятностей и математической статистики, пер. с англ., М., Наука.

Нейман и Пирсон (Neyman J. and Pearson E. S.)

(1928) On the use and interpretation of certain test criteria for purposes of statistical inference, Biometrika, vol. 20A, 175—240, 263—294. [Предложено использовать отношение правдоподобия для получения рациональных критериев, даны применения ко многим конкретным проблемам.]

(1933) On the testing of statistical hypotheses in relation to probability a priori, Proc. Cambridge Phil. Soc., vol. 29, 492—510. [В связи с задачей проверки гипотез предлагается «взвешивание» различных возможных ошибочных решений и использование принципа минимакса.]

Пейсаков (Peisakoff M.)

(1951) Transformation of parameters, Princeton. Неопубликованная диссертация. [Теория инвариантности Ханта—Стейна распространена на более общие классы проблем решения; задача 11 (II). Далее теория была обобщена в работах: J. Kiefer, Invariance, minimax sequential estimation, and continuous time processes, Ann. Math. Stat., vol 28 (1957), 573—601, и H. Kudo, On minmax invariant estimates of the transformations parameter, Nat. Sci. Rept. Ochanomizu Univ., Tokyo, vol. 6 (1955), 31—73.]

Питмен (Pitman E. I. G.)

(1939) Location and scale parameters, Biometrika, vol. 30, 391—421.

(1939) Tests of hypotheses concerning location and scale parameters, Biometrika, vol. 31, 200—215.

[В этих работах ограничение инвариантными процедурами вводится в задачи оценки (или проверки гипотез) параметров сдвига и масштабных параметров.]

Сэвидж (Savage L. I.)

(1954) The Foundations of Statistics, New York, John Wiley & Sons.

Уилкс (Wilks S. S.)

(1944) Mathematical Statistics, Princeton University Press. (Русский перевод последнего издания: С. Уилкс, Математическая статистика, М., Наука, 1967.)

Фишер (Fisher R. A.)

(1920) A mathematical examination of the methods of determining the accuracy of an observation by the mean error and by the mean square error, Monthly Notices Roy. Astron. Soc., vol. 80, 758—770.

(1921) On the mathematical foundation of theoretical statistics, Phil. Trans. Roy. Soc., London, Ser. A, vol. 222, 309—368.

(1922) On the mathematical foundations of theoretical statistics, Phil. Trans. Roy. Soc. London, Ser. A, vol. 222, 309—368. [Развита теория точечных оценок на основе метода максимального правдоподобия.]

(1925) Theory of statistical estimation, Proc. Cambridge Phil. Soc., vol. 22, 700—725.

[В этих статьях развита идея достаточности главным образом в связи с теорией точечных оценок. Теорема факторизации дана формально в более слабой форме, но по существу эквивалентна (20).]

(1925) Statistical Methods for Research Workers, Edinburgh, Oliver and Boyd, 1 st ed., 1925; 11 th ed., 1950. (Русский перевод: Р. Фишер, Статистические методы для исследователей, М., Госстатиздат, 1958.)

(1935) The Design of Experiments, Edinburgh, Oliver and Boyd.

(1950) Contributions to Mathematical Statistics, New York, John Wiley & Sons.

Хант и Стейн (Hunt G. and Stein C.)

(1946) Most stringent tests of statistical hypotheses. [В этой работе, оставшейся, к сожалению, неопубликованной, развита общая теория инвариантности для проверки гипотез.]

Ходжес и Леман (Hodges J. L., Jr. and Lehmann E. L.)

(1952) The use of previous experience in reaching statistical decisions, Ann. Math. Stat., vol. 23, 396—407. [Теория ограниченных байесовских решений.]

Хотеллинг (Hotelling H.)

(1936) Relations between two sets of variates, Biometrika, vol. 28, 321—377. [Одна из ранних работ, отчетливо использующих соображения инвариантности.]

Чубер (Czuber E.)

(1891) Theorie der Beobachtungsfehler, Leipzig, B. G. Teubner.

Шметтерер (Schmetterer L.)

(1956) Einführung in die mathematische Statistik, Wien, Springer-Verlag. (Русский перевод 2-го издания: Л. Шметтерер, Введение в математическую статистику, М., Наука, 1976.)

Эджворт (Edgeworth F. Y.)

(1908—1909) On the probable errors of frequency constants, I. Roy. Stat. Soc., vol. 71, 381—397, 499—512, 651—678; vol. 72, 81—90. [Следуя Лапласу и Гауссу, автор исследует оценки «методом обратных вероятностей», т. е. в предположении равномерного априорного распределения параметра. Поскольку оценки определяются так, что они максимизируют моду апостериорного распределения, то они являются оценками максимального правдоподобия.]

ГЛАВА 2

ВЕРОЯТНОСТНЫЕ ОСНОВЫ

1. Вероятность и мера

Математическую основу теории статистических решений составляет теория вероятностей, которая, в свою очередь, опирается на теорию меры и интегрирования. Настоящий и следующий разделы имеют целью определить некоторые важные понятия этих теорий, установить обозначения и сформулировать без доказательства некоторые из основных результатов. В оставшейся части главы более детально разбирается ряд специальных вопросов.

Теория вероятностей имеет дело с ситуациями, допускающими различные исходы. В абстрактной модели совокупность возможных исходов описывается множеством точек некоторого пространства $\mathcal{X}$. Поскольку изучаемые события состоят из этих исходов, они представляются подмножествами $\mathcal{X}$. Объединение двух множеств A_1, A_2 будет обозначаться $A_1 \cup A_2$, их пересечение — $A_1 \cap A_2$, дополнение к A будет обозначаться $\bar{A} = \mathcal{X} - A$. Пустое множество обозначим 0. Вероятность $P(A)$ события A — действительное число, лежащее между нулем и единицей; в частности,

$$P(0) = 0 \quad \text{и} \quad P(\mathcal{X}) = 1. \quad (1)$$

Вероятность обладает свойством *счетной аддитивности*

$$P(UA_i) = \sum P(A_i), \quad \text{если} \quad A_i \cap A_j = 0 \quad \text{при} \quad i \neq j. \quad (2)$$

К сожалению, оказывается, что функции множеств, с которыми мы будем встречаться, не могут быть, как правило, разумным образом определены для всех подмножеств $\mathcal{X}$, если сохранить требование (2). Невозможно, например, дать разумное определение «площади» для всех подмножеств единичного квадрата на плоскости.

Множества, для которых вероятностная функция определена, будут называться «измеримыми». Область определения P должна включать наряду с каждым множеством A его дополнение и

вместе с каждым счетным классом событий — их объединение. По (1) она содержит также $\mathcal{X}$. Класс подмножеств, содержащий $\mathcal{X}$ и замкнутый относительно операций перехода к дополнению и счетного объединения, называется σ -полем. Такой класс автоматически замкнут относительно счетных пересечений.

Отправной точкой любых вероятностных исследований являются, стало быть, пространство $\mathcal{X}$ возможных исходов и σ -поле $\mathcal{A}$ подмножеств $\mathcal{X}$, представляющее события, для которых будет определена вероятность. Такая пара $(\mathcal{X}, \mathcal{A})$ называется *измеримым пространством*, а элементы $\mathcal{A}$ — *измеримыми множествами*. *Мерой* называется определенная на $\mathcal{A}$, неотрицательная, счетно аддитивная (не обязательно конечная) функция множеств μ , для которой $\mu(0) = 0$. Если ее значение для $\mathcal{X}$ равно 1, то она называется *вероятностной мерой*. Более общим образом, если $\mu(\mathcal{X}) < \infty$, то μ , по определению, *конечна*; если найдется последовательность множеств $A_1, A_2, \dots$ из $\mathcal{A}$ (их можно взять попарно непересекающимися), для которой $\bigcup A_i = \mathcal{X}$ и $\mu(A_i) < \infty$, $i = 1, 2, \dots$, то μ , по определению, σ -*конечна*. В нижеследующих примерах указаны важные специальные случаи.

Пример 1. Пусть $\mathcal{X}$ — n -мерное евклидово пространство и $\mathcal{A}$ — наименьшее σ -поле, содержащее все «прямоугольники»

$$R = \{(x_1, \dots, x_n): a_i < x_i \leq b_i, \quad i = 1, \dots, n\}^*.$$

Элементы $\mathcal{A}$ называются *борелевскими множествами* в $\mathcal{X}$. На $\mathcal{A}$ можно определить единственную меру μ , значение которой для любого прямоугольника R совпадает с его объемом

$$\mu(R) = \prod_{i=1}^n (b_i - a_i).$$

Мера μ может быть *пополнена* присоединением к $\mathcal{A}$ всех подмножеств меры нуль. Область определения μ расширяется, таким образом, до σ -поля $\mathcal{A}'$, элементы которого называются *измеримыми по Лебегу* множествами. Термин *мера Лебега* будет использован для μ как в случае, когда она рассматривается на борелевских множествах, так и в случае, когда она рассматривается на множествах, измеримых по Лебегу.

В этом примере объем может быть заменен любой неотрицательной функцией множеств ν , которая определена и счетно аддитивна на классе всех прямоугольников R . Как и прежде, существует единственная мера μ на $(\mathcal{X}, \mathcal{A})$, которая совпадает с ν для всех R . Эта мера также может быть пополнена; однако возникающее σ -поле зависит от μ и не обязано совпадать с σ -полем $\mathcal{A}'$, введенным выше.

* Если $\pi(x)$ — утверждение, касающееся некоторых объектов x , то $\{x: \pi(x)\}$ обозначает множество всех тех x , для которых $\pi(x)$ верно.

Пример 2. Предположим, что $\mathcal{X}$ счетно, и пусть $\mathcal{A}$ — класс всех подмножеств $\mathcal{X}$. Для любого A положим $\mu(A)$ равным числу его элементов, если оно конечно; в противном случае положим меру равной $+\infty$. Эту меру иногда называют *считающей мерой*.

В приложениях вероятности на $(\mathcal{X}, \mathcal{A})$ отражают обстановку случайного эксперимента или наблюдений, возможные исходы которых описываются точками $x \in \mathcal{X}$. Обозначим результаты этих наблюдений, которые могут быть, например, действительными числами или векторами, буквой X . Пусть вероятность того, что X попадает во множество A , равна $P\{X \in A\} = P(A)$. В таком контексте вероятность $P(A)$ будет иногда обозначаться $P^X(A)$, а вероятностная мера P — знаком P^X . Мы будем говорить об X , как о *случайной величине**) в пространстве $(\mathcal{X}, \mathcal{A})$ и о вероятностной мере P или P^X — как о *распределении вероятностей* X . Математически, таким образом, случайная величина есть не что иное, как «носитель» своего распределения. Пусть $\pi(x)$ — какое-либо утверждение относительно точек x , и A — множество тех точек, для которых $\pi(x)$ истинно. Тогда вероятность $P^X(A)$ мы будем иногда записывать в виде $P\{\pi(X)\}$.

Пусть X — действительная случайная величина с распределением вероятностей P^X , определенном на борелевских множествах числовой прямой. Функция точки a , определяемая для любого a равенством $F(a) = P\{X \leq a\}$, называется *функцией распределения* X . Функция F неубывающая, непрерывная справа, и $F(-\infty) = 0$, $F(+\infty) = 1$. Если F — любая функция с этими свойствами, то равенство $P\{a < X \leq b\} = F(b) - F(a)$ определяет меру на интервалах. Как и в примере 1, эта мера единственным образом продолжается до меры на борелевских множествах. Таким образом, между распределениями вероятностей и функциями распределения устанавливается взаимно однозначное соответствие. Это остается справедливым и для распределений в n -мерном евклидовом пространстве, где функция распределения определяется соотношением $F(a_1, \dots, a_n) = P\{X_1 \leq a_1, \dots, X_n \leq a_n\}$. По распределению X можно найти распределение любой функции от X . Пусть T — функция от результатов наблюдений, принимающая значения в некотором пространстве $\mathcal{T}$. Эта функция порождает в $\mathcal{T}$ σ -поле $\mathcal{B}'$, состоящее из множеств B , полные прообразы которых

$$A = T^{-1}(B) = \{x : x \in \mathcal{X}, T(x) \in B\}$$

входят в $\mathcal{A}$. Значения, принимаемые $T(X)$, опять возникают как исходы случайного эксперимента, так что $T = T(X)$ является

*) Это определение отличается от даваемого в большинстве учебников теории вероятностей, где случайная величина есть измеримая функция, отображающая первоначальное пространство в пространство $(\mathcal{X}, \mathcal{A})$.

случайной величиной в пространстве $(\mathcal{T}, \mathcal{B}')$. Так как $X \in T^{-1}(B)$ тогда и только тогда, когда $T(X) \in B$, то распределение вероятностей T в $(\mathcal{T}, \mathcal{B}')$ задается равенством

$$P^T(B) = P\{T \in B\} = P\{X \in T^{-1}(B)\} = P^X(T^{-1}(B)). \quad (3)$$

Часто случается, что в $\mathcal{T}$ заранее задано σ -поле $\mathcal{B}$ его подмножеств, такое, что вероятность события $T \in \mathcal{B}$ должна быть определена для тех и только тех B , которые $\in \mathcal{B}$. Отсюда с необходимостью вытекает, что $T^{-1}(B) \in \mathcal{A}$ для всех $B \in \mathcal{B}$. Отображение (функция, преобразование) T пространства $(X, \mathcal{A})$ в $(\mathcal{T}, \mathcal{B})$ называется в этом случае *измеримым*. Другим следствием является удобное порой использование вероятностных утверждений только применительно ко множествам, входящим в $\mathcal{B}$, даже если существует $B \notin \mathcal{B}$, для которых $T^{-1}(B) \in \mathcal{A}$ и вероятность которых, следовательно, могла бы быть определена.

В приложениях «сырьем» для исследования служит совокупность результатов наблюдений, это единственно доступные данные. Результаты наблюдений представляются случайной величиной X , так что все другие случайные величины, которые могут встретиться, являются функциями от X . Пространство $(\mathcal{X}, \mathcal{A})$, то, в котором задано X , называется *выборочным пространством*, а каждое измеримое отображение T пространства $(\mathcal{X}, \mathcal{A})$ в $(\mathcal{T}, \mathcal{B})$ — *статистикой*. Распределение T определяется для всех $B \in \mathcal{B}$ с помощью (3). При таком подходе статистика T описывается заданием как функции T , так и σ -поля $\mathcal{B}$. Мы условимся считать, однако, что каждый раз функция T принимает значения в евклидовом пространстве, в качестве σ -поля $\mathcal{B}$ измеримых множеств будет выбираться класс борелевских множеств (если только явно не сказано противное). При таком соглашении указывать $\mathcal{B}$, в том числе и в обозначениях, нет никакой необходимости.

Различие между статистикой и случайной величиной, как они введены выше, оказывается незначительным. Термин «статистика» используется лишь с тем, чтобы подчеркнуть, что эта величина является функцией «основных» результатов наблюдений. Все статистики в какой-либо данной задаче представляют собой функции на выборочном пространстве $(\mathcal{X}, \mathcal{A})$. С другой стороны, каждая статистика есть случайная величина, так как она имеет распределение в $(\mathcal{T}, \mathcal{B})$, и она будет именоваться слу-

*) Термин «в» употребляется для указания того, что область значений $T(\mathcal{X})$ функции T лежит в $\mathcal{T}$, если $T(\mathcal{X}) = \mathcal{T}$, то говорят, что T отображает $\mathcal{X}$ на $\mathcal{T}$.

чайной величиной в обстоятельствах, когда ее происхождение несущественно. Какой термин употреблять — зависит, следовательно, от точки зрения и до известной степени произвольно *).

2. Интегрирование

В соответствии с соглашением предыдущего раздела, действительная функция f , определенная на $(\mathcal{X}, \mathcal{A})$, измерима, если $f^{-1}(B) \in \mathcal{A}$ для любого борелевского множества B на прямой. Такая функция называется *простой*, если она принимает лишь конечное число значений. Пусть μ — мера в $(\mathcal{X}, \mathcal{A})$ и f — простая функция, принимающая различные значения $a_1, \dots, a_m$ на множествах $A_1, \dots, A_m$, принадлежащих $\mathcal{A}$ (так как f — измерима). Если $\mu(A_i) < \infty$ при $a_i \neq 0$, то интеграл от f относительно μ определяется равенством

$$\int f d\mu = \sum a_i \mu(A_i). \quad (4)$$

Какова бы ни была неотрицательная измеримая функция f , существует неубывающая последовательность простых функций f_n , сходящаяся к f . Тогда интеграл от f определяется формулой

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu. \quad (5)$$

Можно показать, что это определение не зависит от специального выбора последовательности f_n . Для любой измеримой функции ее положительная и отрицательная части

$$f^+(x) = \max[f(x), 0] \text{ и } f^-(x) = \max[-f(x), 0] \quad (6)$$

также измеримы, и

$$f(x) = f^+(x) - f^-(x).$$

Если интегралы от f^+ и f^- оба конечны, то f называется *интегрируемой*, и ее интеграл полагается равным

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

Если же один из двух интегралов конечен, а другой бесконечен, то интегралу от f приписывается соответствующее бесконечное значение.

*) Данное определение статистики близко к определению случайной величины в теории вероятностей. В теоретико-множественном отношении разницы между этими двумя понятиями нет. Различие, сделанное здесь, соответствует сложившемуся способу употребления терминов в статистической литературе. Подробнее об этих подходах к определению случайной величины см. Гнеденко Б. В., Колмогоров А. Н., Предельные распределения для сумм независимых случайных величин, М., Гостехиздат, 1949 или дополнение к русскому переводу книги: Крамер Г., Математические методы статистики, М., Мир, 1975. — Прим. ред.

Пример 3. Пусть $\mathcal{X}$ — замкнутый интервал $[a, b]$, $\mathcal{A}$ — класс борелевских или измеримых по Лебегу множеств и μ — мера Лебега. Тогда интеграл от f по μ записывается как $\int_a^b f(x) dx$ и называется интегралом Лебега от f .

Этот интеграл обобщает интеграл Римана в том смысле, что если последний существует, то и первый существует и имеет то же самое значение.

Пример 4. Пусть $\mathcal{X}$ — счетное множество, состоящее из точек $x_1, x_2, \dots$; $\mathcal{A}$ — класс всех подмножеств $\mathcal{X}$ и μ приписывает точке x_i меру b_i . Тогда f интегрируема, если ряд $\sum f(x_i) b_i$ сходится абсолютно, и интеграл $\int f d\mu$ равен этой сумме.

Пусть P^X — распределение вероятностей случайной величины X и T — действительная статистика. Если $T(x)$ интегрируема, то ее математическое ожидание определяется формулой

$$E(T) = \int T(x) dP^X(x). \quad (7)$$

Как будет показано в лемме 2 раздела 3 (см. ниже), интегрирование можно также проводить в t -пространстве по отношению к распределению T , даваемому (3), так что

$$E(T) = \int t dP^T(t). \quad (8)$$

Приведенному выше определению интеграла соответствует следующая основная теорема о сходимости.

Теорема 1. Пусть f_n — последовательность измеримых функций и пусть $f_n(x) \rightarrow f(x)$ при всех x . Тогда

$$\int f_n d\mu \rightarrow \int f d\mu,$$

если выполнено одно из следующих условий:

(I) (Теорема Лебега о монотонной сходимости) f_n — неотрицательны и последовательность не убывает.

(II) (Теорема Лебега об ограниченной сходимости) — существует интегрируемая функция g такая, что $|f_n(x)| \leq g(x)$ при всех n и x .

Для любого множества $A \in \mathcal{A}$ обозначим I_A его индикатор:

$$I_A(x) = 1 \text{ или } 0, \text{ когда } x \in A \text{ или } x \notin A, \quad (9)$$

и положим

$$\int_A f d\mu = \int f I_A d\mu. \quad (10)$$

Если μ — мера и f — неотрицательная измеримая функция на $(\mathcal{X}, \mathcal{A})$, то

$$\nu(A) = \int_A f d\mu \quad (11)$$

также является мерой на $(\mathcal{X}, \mathcal{A})$. Тот факт, что (11) выполняется для всех $A \in \mathcal{A}$, выражается записью

$$d\nu = f d\mu \quad \text{или} \quad f = \frac{d\nu}{d\mu}. \quad (12)$$

Пусть μ и ν — две данные σ -конечные меры на $(\mathcal{X}, \mathcal{A})$. Если существует функция f , удовлетворяющая (12), то она этим равенством определяется с точностью до множеств меры нуль, так как из равенства $\int_A f d\mu = \int_A g d\mu$ при всех $A \in \mathcal{A}$ вытекает, что $f = g$ μ -почти всюду *). Такая f называется *производной Радона — Никодима* меры ν по мере μ и, в частном случае, когда ν есть вероятностная мера, *плотностью вероятности* ν по отношению к μ .

Вопрос существования функции f , удовлетворяющей (12), решается в терминах, даваемых следующим определением. Мера ν называется *абсолютно непрерывной* по отношению к μ , если

$$\text{из } \mu(A) = 0 \text{ следует } \nu(A) = 0.$$

Теорема 2. (Радон — Никодим). Если μ и ν — σ -конечные меры на $(\mathcal{X}, \mathcal{A})$, то измеримая функция f , удовлетворяющая (12), существует в том и только том случае, когда ν абсолютно непрерывна по отношению к μ .

Прямым (или *декартовым*) *произведением* $A \times B$ двух множеств A, B называется множество всех упорядоченных пар (x, y) с $x \in A, y \in B$. Пусть $(\mathcal{X}, \mathcal{A})$ и $(\mathcal{Y}, \mathcal{B})$ — два измеримых пространства, и пусть $\mathcal{A} \times \mathcal{B}$ обозначает наименьшее σ -поле, содержащее все множества вида $A \times B$, где $A \in \mathcal{A}$ и $B \in \mathcal{B}$. Если μ и ν — две σ -конечные меры на $(\mathcal{X}, \mathcal{A})$ и $(\mathcal{Y}, \mathcal{B})$ соответственно, то существует единственная мера $\lambda = \mu \times \nu$ на $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B})$, называемая *произведением* μ и ν , такая, что для любых $A \in \mathcal{A}, B \in \mathcal{B}$

$$\lambda(A \times B) = \mu(A) \nu(B). \quad (13)$$

Пример 5. Пусть $\mathcal{X}$ и $\mathcal{Y}$ — евклидовы пространства размерности m и n соответственно, $\mathcal{A}$ и $\mathcal{B}$ — σ -поля борелевских множеств в этих пространствах. Тогда $\mathcal{X} \times \mathcal{Y}$ является $(m + n)$ -мерным евклидовым пространством, а $\mathcal{A} \times \mathcal{B}$ — классом его борелевских множеств.

Пример 6. Пусть $Z = (X, Y)$ — случайная величина на $(\mathcal{X} \times \mathcal{Y}, \mathcal{A} \times \mathcal{B})$. Предположим, что случайные величины X и Y имеют распределения P^X и P^Y на $(\mathcal{X}, \mathcal{A})$ и $(\mathcal{Y}, \mathcal{B})$. Тогда X и Y называют *независимыми*, если распределение P^Z величины Z является произведением $P^X \times P^Y$.

*) Об утверждении, которое выполняется для всех точек x , исключая, быть может, множество μ меры 0, говорят, что оно выполняется μ -почти всюду (п. в.); или выполняется $(\mathcal{A}, \mu)$ — почти всюду, если желательно указать σ -поле, на котором μ определена.

В терминах этих понятий сведение двойного интеграла к повторному дается следующей теоремой.

Теорема 3. (Фубини.) Пусть μ и ν — σ -конечные меры на $(\mathcal{X}, \mathcal{A})$ и $(\mathcal{Y}, \mathcal{B})$ соответственно, и пусть $\lambda = \mu \times \nu$. Если $f(x, y)$ интегрируема относительно λ , то

(I) для почти всех (y) фиксированных y функция $f(x, y)$ интегрируема относительно μ ;

(II) функция $\int f(x, y) d\mu(x)$ интегрируема относительно ν , и

$$\int f(x, y) d\lambda(x, y) = \int \left[\int f(x, y) d\mu(x) \right] d\nu(y). \quad (14)$$

3. Статистики и подполя

В соответствии с определением раздела 1, статистика есть измеримое отображение T выборочного пространства $(\mathcal{X}, \mathcal{A})$ в измеримое пространство $(\mathcal{T}, \mathcal{B})$. Это отображение индуцирует в первоначальном выборочном пространстве подполе*)

$$\mathcal{A}_0 = T^{-1}(\mathcal{B}) = \{T^{-1}(B) : B \in \mathcal{B}\}. \quad (15)$$

Так как $T^{-1}[T(A)]$ содержит A , но не обязано с ним совпадать, то подполе $\mathcal{A}_0$ не обязано совпадать с $\mathcal{A}$ и, следовательно, может оказаться собственным подполем $\mathcal{A}$. С другой стороны, предположим на время, что $\mathcal{T} = T(\mathcal{X})$, т. е. что T является отображением на $\mathcal{T}$ (а не в $\mathcal{T}$). Тогда

$$T[T^{-1}(B)] = B \quad \text{для всех } B \in \mathcal{B}, \quad (16)$$

так что формула $A_0 = T^{-1}(B)$ устанавливает взаимно однозначное соответствие между классами $\mathcal{A}_0$ и $\mathcal{B}$. Это соответствие изоморфно, т. е. сохраняет теоретико-множественные операции пересечения, объединения и перехода к дополнению. Поэтому для большинства целей несущественно, оперировать ли с пространством $(\mathcal{X}, \mathcal{A}_0)$, или с пространством $(\mathcal{T}, \mathcal{B})$. Эти пространства порождают два эквивалентных класса событий, два класса измеримых функций, возможных решающих процедур и т. п. Если T является отображением в $\mathcal{T}$, то указанное взаимно однозначное соответствие применяется к классу $\mathcal{B}'$ входящих в $\mathcal{B}$ подмножеств $\mathcal{T}' = T(\mathcal{X})$ (а не непосредственно к $\mathcal{B}$). Однако любое множество $B \in \mathcal{B}$ эквивалентно $B' = B \cap \mathcal{T}'$ в том смысле, что любая мера на $(\mathcal{X}, \mathcal{A})$ принимает на B' и B одинаковые значения. Следовательно, рассматриваемые как классы событий $\mathcal{A}_0$ и $\mathcal{B}$ продолжают быть эквивалентными с той лишь

*) Мы будем употреблять этот термин вместо более длинного «под- σ -поле».

разницей, что $\mathcal{B}$ содержит несколько (эквивалентных) представлений одного и того же события.

Пусть, например, $\mathcal{X}$ — действительная прямая, $\mathcal{A}$ — класс борелевских множеств и $T(x) = x^2$. Пусть $\mathcal{T}$ означает или положительную полупрямую, или всю прямую, а $\mathcal{B}$ — класс борелевских подмножеств $\mathcal{T}$. Тогда $\mathcal{A}_0$ является классом всех борелевских множеств, которые симметричны по отношению к началу координат. Рассматривая, скажем, действительные измеримые функции и оперируя в $\mathcal{T}$ -пространстве, мы ограничились бы измеримыми функциями от x^2 . Вместо этого мы могли бы остаться в первоначальном пространстве, и ограничение сводилось бы к обращению только с четными измеримыми функциями x . Эквивалентность подходов очевидна. Какое представление выбрать — это вопрос удобства.

Нижеследующая лемма показывает, что соответствие между множествами $A_0 = T^{-1}(B) \in \mathcal{A}_0$ и $B \in \mathcal{B}$ продолжается в аналогичное соответствие между измеримыми функциями на $(\mathcal{X}, \mathcal{A}_0)$ и $(\mathcal{T}, \mathcal{B})$.

Лемма 1. Пусть статистика T , отображающая $(\mathcal{X}, \mathcal{A})$ в $(\mathcal{T}, \mathcal{B})$, индуцирует подполе $\mathcal{A}_0$. Тогда действительная $\mathcal{A}$ -измеримая функция f является $\mathcal{A}_0$ -измеримой тогда и только тогда, когда существует такая $\mathcal{B}$ -измеримая функция g , что

$$f(x) = g[T(x)]$$

при всех x .

Доказательство. Предположим сначала, что такая функция g существует. Тогда множество

$$\{x : f(x) < r\} = T^{-1}(\{t : g(t) < r\})$$

лежит в $\mathcal{A}_0$ и f $\mathcal{A}_0$ -измерима. Обратно, если f $\mathcal{A}_0$ -измерима, то множества

$$A_{in} = \left\{ x : \frac{i}{2^n} < f(x) \leq \frac{i+1}{2^n} \right\}, \quad i = 0, \pm 1, \pm 2, \dots$$

при каждом фиксированном n попарно не пересекаются, принадлежат $\mathcal{A}_0$ и в сумме дают все $\mathcal{X}$. Далее, существуют $B_{in} \in \mathcal{B}$ такие, что $A_{in} = T^{-1}(B_{in})$. Пусть

$$B_{in}^* = B_{in} \cap \bigcap_{j \neq i} \overline{B_{jn}}.$$

Так как при $i \neq j$ A_{in} и A_{jn} не пересекаются, то множество $T^{-1}(B_{in} \cap B_{jn})$, а вместе с ним и множество $T^{-1}(B_{in} \cap \tilde{B}_{jn}^*)$ пусто. Следовательно, при фиксированном n множества B_{in}^* попарно не пересекаются и удовлетворяют соотношению $A_{in} = T^{-1}(B_{in}^*)$. Полагая

$$f_n(x) = \frac{i}{2^n} \quad \text{для} \quad x \in A_{in}, \quad i = 0, \pm 1, \pm 2, \dots,$$

мы можем написать

$$f_n(x) = g_n[T(x)],$$

где

$$g_n(t) = \begin{cases} \frac{i}{2^n} & \text{для } t \in B_{in}^*, \quad i = 0, \pm 1, \pm 2, \dots \\ 0 & \text{в других случаях.} \end{cases}$$

Так как функции g_n $\mathcal{B}$ -измеримы, то множество B , на котором $g_n(t)$ сходятся к конечному пределу, входит в $\mathcal{B}$. Пусть $R = T(\mathcal{X})$ — область значений T . Тогда для $t \in R$

$$\lim g_n[T(x)] = \lim f_n(x) = f(x)$$

для всех $x \in \mathcal{X}$, так что R содержится в B . Следовательно, функция g , равная $\lim g_n(t)$ для $t \in B$ и 0 в других случаях, является искомой.

Соотношение между интегралами функций f и g дается следующей леммой.

Лемма 2. Пусть T — измеримое отображение пространства $(\mathcal{X}, \mathcal{A})$ в $(\mathcal{T}, \mathcal{B})$, μ — σ -конечная мера на $(\mathcal{X}, \mathcal{A})$ и g — действительная измеримая функция от t . Если μ^* — мера, определенная на $(\mathcal{T}, \mathcal{B})$ равенством

$$\mu^*(B) = \mu[T^{-1}(B)] \quad \text{для всех } B \in \mathcal{B}, \quad (17)$$

то для любого $B \in \mathcal{B}$

$$\int_{T^{-1}(B)} g[T(x)] d\mu(x) = \int_B g(t) d\mu^*(t) \quad (18)$$

в том смысле, что если один из этих интегралов существует, то существует и другой и их значения совпадают.

Доказательство. Не ограничивая общности, допустим, что B совпадает со всем пространством $\mathcal{T}$. Если g является индикатором некоторого множества $B_0 \in \mathcal{B}$, то утверждение леммы верно, так как правая и левая части (18) приводятся к $\mu[T^{-1}(B_0)]$ и $\mu^*(B_0)$ соответственно, а эти выражения равны, по определению μ^* . Отсюда (18) переносится последовательно на все простые функции, на все неотрицательные измеримые функции и, наконец, на все интегрируемые функции.

4. Условное математическое ожидание и условная вероятность

Если две статистики индуцируют одно и то же подполе $\mathcal{A}_0$, то они эквивалентны в том смысле, что они приводят к эквивалентным классам измеримых событий. Эта эквивалентность существенна в рассуждениях, касающихся условной вероятности.

Так, если X распределено нормально с нулевым средним значением, то статистики $|X|$, X^2 , e^{-X^2} и т. д. доставляют нам одну и ту же информацию. Из равенств $|X| = t$, $X^2 = t^2$, $e^{-X^2} = e^{-t^2}$ следует, что $X = \pm t$ и любое разумное определение условной вероятности должно приписать каждому из этих значений вероятность $1/2$. В данное ниже общее определение условной вероятности существенным образом входит именно $\mathcal{A}_0$, а не пространство $\mathcal{T}$ значений статистики T . Однако если рассматривать только $\mathcal{A}_0$, то понятие проигрывает в наглядности, и разрыв между элементарным и общим определениями без всякой необходимости увеличивается. По этой причине часто оказывается более удобным иметь дело с конкретными представлениями статистик, включающими определенный выбор пространства значений $(\mathcal{T}, \mathcal{B})$.

Пусть P — распределение вероятностей на $(\mathcal{X}, \mathcal{A})$, T — статистика с пространством значений $(\mathcal{T}, \mathcal{B})$ и $\mathcal{A}_0$ — порожденное ею подполе. Рассмотрим неотрицательную функцию f , которая $(\mathcal{A}, P)$ интегрируема, т. е. $\mathcal{A}$ -измерима и P -интегрируема. Тогда $\int_A f dP$ определен для всех $A \in \mathcal{A}$ и тем более для всех $A_0 \in \mathcal{A}_0$. Из теоремы Радона — Никодима (теорема 2) вытекает, что существует функция f_0 , интегрируемая $(\mathcal{A}_0, P)$, для которой

$$\int_{A_0} f dP = \int_{A_0} f_0 dP \quad \text{при всех } A_0 \in \mathcal{A}_0, \quad (19)$$

и что f_0 $(\mathcal{A}_0, P)$ -единственна. По лемме 1 f_0 зависит от x только через $T(x)$. В примере со случайной величиной X , распределенной нормально с нулевым средним, и $T = X^2$ функция f_0 определяется тем, что (19) должно выполняться для всех множеств A_0 , симметричных относительно начала координат, так что

$$f_0(x) = \frac{1}{2} [f(x) + f(-x)].$$

Функция f_0 , участвующая в (19), определяется двумя свойствами:

(I) Ее среднее значение по мере P на любом множестве A_0 то же, что и у f .

(II) Она зависит от x только через $T(x)$ и, стало быть, постоянна на множествах D_x , на которых постоянна T .

Интуиция подсказывает способ построения такой функции: $f_0(x)$ следует определить, как условное P -среднее на множестве D_x . Мы заменили бы тогда однократный процесс усреднения, т. е. интегрирования f , представляемый левой частью, двукратным процессом усреднения (повторным интегрированием). Эта конструкция может быть осуществлена, когда X — дискретная

величина, или в регулярном случае, рассмотренном в разделе 9 главы 1. В этих случаях $f_0(x)$ оказывается условным математическим ожиданием $f(X)$ при данном $T(x)$. В общем случае неясно, как можно определить условное математическое ожидание непосредственно. Так как оно, однако, должно обладать свойствами (I) и (II) и так как f_0 определяется $(\mathcal{A}_0, P)$ однозначно (через (19)), то мы принимаем в качестве общего определения *условного математического ожидания* $E[f(X)|T(x)]$ функцию $f_0(x)$ из равенства (19). Если $f_0(x) = g[T(x)]$, то можно записать

$$E[f(x)|T=t] = g(t),$$

т. е. $E[f(x)|t]$ является $\mathcal{B}$ -измеримой функцией, определенной с точностью до $(\mathcal{B}, P^T)$ -эквивалентности. Если в указанном в лемме 2 соотношении между интегралами принять $\mu = P^X$, то $\mu^* = P^T$, и мы видим, что функция g может быть выражена непосредственно через f с помощью соотношения

$$\int_{T^{-1}(B)} f(x) dP^X(x) = \int_B g(t) dP^T(t) \quad \text{для всех } B \in \mathcal{B}, \quad (20)$$

эквивалентного (19).

До сих пор функция f предполагалась неотрицательной. В общем случае условное математическое ожидание f определяется как разность:

$$E[f(X)|t] = E[f^+(X)|t] - E[f^-(X)|t].$$

Пример 7. Пусть $X_1, \dots, X_n$ независимы и имеют одну и ту же непрерывную функцию распределения. Положим

$$T(x_1, \dots, x_n) = (x^{(1)}, \dots, x^{(n)}),$$

где $x^{(1)} \leq \dots \leq x^{(n)}$ обозначают x_k , расположенные в порядке возрастания. Не ограничивая общности, мы можем рассматривать только точки, где $x^{(1)} < \dots < x^{(n)}$, так как вероятность совпадения каких-либо координат равна нулю. Тогда $\mathcal{B}$ будет множеством точек с попарно различными координатами, $\mathcal{T}$ — множеством точек со строго возрастающими координатами, $\mathcal{A}$ и $\mathcal{B}$ — классами борелевских подмножеств $\mathcal{B}$ и $\mathcal{T}$. При обратном отображении T^{-1} множество, состоящее из одной точки $a = (a_1, \dots, a_n)$, переходит во множество из $n!$ точек $(a_{i_1}, \dots, a_{i_n})$, получающихся из a всевозможными перестановками координат. Отсюда следует, что $\mathcal{A}_0$ является классом всех множеств, симметричных в том смысле, что если A_0 содержит $x = (x_1, \dots, x_n)$, то оно содержит и все точки $(x_{i_1}, \dots, x_{i_n})$.

Для любой интегрируемой функции f положим

$$f_0(x) = \frac{1}{n!} \sum f(x_{i_1}, \dots, x_{i_n}),$$

где суммирование производится по всем $n!$ перестановкам $(x_1, \dots, x_n)$. Тогда f_0 будет $\mathcal{A}$ -измеримой, так как она симметрична относительно n своих аргу-

ментов. Итак,

$$\int_{A_0} f(x_1, \dots, x_n) dP(x_1) \dots dP(x_n) = \int_{A_0} f(x_{i_1}, \dots, x_{i_n}) dP(x_1) \dots dP(x_n),$$

так что f_0 удовлетворяет (19), откуда следует, что $f_0(x)$ представляет собой условное математическое ожидание $f(X)$ при данном $T(x)$.

Условное математическое ожидание $f(X)$ при заданной статистике $T(x)$ может быть вычислено и без предположения, что X_k независимы и одинаково распределены. Предположим, что X имеет плотность $h(x)$ по отношению к мере μ (такой, как мера Лебега), которая симметрична, т. е. которая при любом $A \in \mathcal{A}$ и любой перестановке $(i_1, \dots, i_n)$ приписывает всем множествам $\{x: (x_{i_1}, \dots, x_{i_n}) \in A\}$ одну и ту же меру. Пусть

$$f_0(x_1, \dots, x_n) = \frac{\sum f(x_{i_1}, \dots, x_{i_n}) h(x_{i_1}, \dots, x_{i_n})}{\sum h(x_{i_1}, \dots, x_{i_n})},$$

причем суммирование в этой формуле (так же, как и в дальнейших) распространяется на все $n!$ перестановок $(i_1, \dots, i_n)$. Функция f_0 симметрична по n своим аргументам и, следовательно, $\mathcal{A}_0$ -измерима. Каково бы ни было симметричное множество A_0 , интеграл

$$\int_{A_0} f_0(x_1, \dots, x_n) h(x_{i_1}, \dots, x_{i_n}) d\mu(x_1, \dots, x_n)$$

имеет одно и то же значение при любой перестановке $(x_{i_1}, \dots, x_{i_n})$, и потому

$$\begin{aligned} \int_{A_0} f_0(x_1, \dots, x_n) h(x_1, \dots, x_n) d\mu(x_1, \dots, x_n) &= \\ &= \int_{A_0} f_0(x_1, \dots, x_n) \frac{1}{n!} \sum h(x_{i_1}, \dots, x_{i_n}) d\mu(x_1, \dots, x_n) = \\ &= \int_{A_0} f(x_1, \dots, x_n) h(x_1, \dots, x_n) d\mu(x_1, \dots, x_n). \end{aligned}$$

Отсюда мы видим, что $f_0(x) = E[f(X) | T(x)]$.

Статистике $T(x) = (x^{(1)}, \dots, x^{(n)})$ (координаты которой называются *рядковыми статистиками*) эквивалентна статистика $U(x) = (\sum x_i, \sum x_i^2, \dots, \sum x_i^n)$. Это немедленно вытекает из того факта, что равенства $T(x^0) = i^0$ и $U(x^0) = u^0$ влекут равенство $T^{-1}(\{i^0\}) = U^{-1}(\{u^0\}) = S$, где $\{i^0\}$ и $\{u^0\}$ обозначают множества, содержащие по одной точке (i^0 или u^0), и где S — множество точек $x = (x_1, \dots, x_n)$, получаемых из $x^0 = (x_1^0, \dots, x_n^0)$ всевозможными перестановками координат. Очевидно, что $T^{-1}(\{i^0\}) = S$. Чтобы установить аналогичное соотношение для U^{-1} , введем

$$V(x) = \left(\sum_i x_i, \sum_{i < j} x_i x_j, \sum_{i < j < k} x_i x_j x_k, \dots, x_1 x_2 \dots x_n \right),$$

так что компоненты $V(x)$ являются элементарными симметрическими функ-

циями $v_1 = \sum x_i, \dots, v_n = x_1 \dots x_n$ от n аргументов $x_1, \dots, x_n$. Тогда

$$(x - x_1) \dots (x - x_n) = x^n - v_1 x^{n-1} + v_2 x^{n-2} - \dots + (-1)^n v_n.$$

Следовательно, из $V(x^0) = v^0 = (v_1^0, \dots, v_n^0)$ вытекает, что $V^{-1}(\{v^0\}) = S$; в этом случае и $U^{-1}(\{u^0\}) = S$, как это видно из соотношений (так называемых тождеств Ньютона)*

$$u_k - v_1 u_{k-1} + v_2 u_{k-2} - \dots + (-1)^{k-1} v_{k-1} u_1 + (-1)^k k v_k = 0, \quad 1 \leq k \leq n.$$

Как легко вывести из данного выше определения, условное математическое ожидание обладает большинством обычных свойств математических ожиданий. Из неединственности этого определения следует, что эти свойства выполняются только $(\mathcal{B}, P^T)$ -почти всюду. Формулировки даны в следующей лемме.

Лемма 3. Если T — некоторая статистика и функции $f, g, \dots (\mathcal{A}, P)$ -интегрируемы, то $(\mathcal{B}, P^T)$ -почти всюду

- (I) $E[af(X) + bg(X)|t] = aE[f(X)|t] + bE[g(X)|t]$;
- (II) $E[h(T)f(X)|t] = h(t)E[f(X)|t]$;
- (III) $a \leq f(x) \leq b$ влечет $a \leq E[f(X)|t] \leq b$ $(\mathcal{A}, P)$ -почти всюду;
- (IV) $|f_n| \leq g, f_n(x) \rightarrow f(x)$ влечет

$$E[f_n(X)|t] \rightarrow E[f(X)|t] \quad (\mathcal{A}, P)\text{-почти всюду.}$$

Другой полезный результат получается из (20) в случае, когда B равно всему пространству $\mathcal{T}$.

Лемма 4. Если $E[f(X)] < \infty$ и $g(t) = E[f(X)|t]$, то

$$Ef(X) = Eg(T), \quad (21)$$

т. е. математическое ожидание может быть вычислено, как среднее значение условного математического ожидания.

Так как $P\{X \in A\} = E[I_A(X)]$, где I^A — индикатор множества A , то условную вероятность A при данном $T = t$ естественно определить равенством

$$P(A|t) = E[I_A(X)|t]. \quad (22)$$

Принимая во внимание (20), мы можем записать уравнение, определяющее $P(A|t)$, в виде

$$\begin{aligned} P^X(A \cap T^{-1}(B)) &= \int_{A \cap T^{-1}(B)} dP^X(x) = \\ &= \int_B P(A|t) dP^T(t) \quad \text{для всех } B \in \mathcal{B}. \end{aligned} \quad (23)$$

Из леммы 3 немедленно вытекает, что с соответствующими оговорками, касающимися множеств нулевой меры, для $P(A|t)$

* Доказательство этих соотношений см., например, в книге: Dickson, New First Course in the Theory of Equations, New York, John Wiley & Sons, 1939, глава X.

имеют место обычные для вероятностей соотношения. Результаты сформулированы в следующей лемме.

Лемма 5. Если T — некоторая статистика с пространством значений $(\mathcal{T}, \mathcal{B})$ и $A, B, A_1, A_2, \dots$ — множества, принадлежащие $\mathcal{A}$, то $(\mathcal{B}, P_T)$ -почти всюду:

$$(I) 0 \leq P(A|t) \leq 1;$$

(II) если множества $A_1, A_2, \dots$ попарно не пересекаются, то $P(\cup A_i|t) = \sum P(A_i|t)$;

$$(III) A \subset B \text{ влечет } P(A|t) \leq P(B|t).$$

В соответствии с определением (22), условная вероятность $P(A|t)$ должна рассматриваться при каждом фиксированном A как $\mathcal{B}$ -измеримая функция t . Мы видим контраст между этим и элементарным определением; в последнем t считают фиксированным и рассматривают $P(A|t)$ при меняющемся A как функцию множеств на $\mathcal{A}$. Лемма 5 может привести к мысли, что интерпретация $P(A|t)$ при фиксированном t , как распределения вероятностей на $\mathcal{A}$, сохранится и в общем случае. Однако равенство $P(A_1 \cup A_2|t) = P(A_1|t) + P(A_2|t)$, например, может нарушаться на множестве нулевой меры, а это множество может меняться в зависимости от A_1 и A_2 . Объединение всех исключительных нулевых множеств *) не обязано иметь меру нуль.

Для важного класса случаев эту трудность можно преодолеть за счет неединственности выбора значений функции $P(A|t)$, которая при каждом фиксированном t определяется с точностью до множеств нулевой меры по t . Так как различные варианты выбора эквивалентны, то достаточно найти такой специальный вариант для каждого A , чтобы при фиксированном t получить распределение вероятностей на $\mathcal{A}$. Эта возможность иллюстрируется примером 7, в котором можно принять, что условное распределение при данном $T(x) = t$ приписывает вероятность $\frac{1}{n!}$ каждой из $n!$ точек, для которых $T(x) = t$. В более общей обстановке вопрос о существовании подобных условных распределений будет изучен в следующем разделе.

5. Условные распределения вероятностей **)

Мы исследуем теперь вопрос о существовании условного распределения вероятностей при условии, выполняющемся в большинстве статистических приложений, а именно, что $\mathcal{B}$ — боре-

*) Этот термин употребляется взамен более длинного выражения «множество меры нуль».

**) Этот раздел может быть опущен при первом чтении. Основное применение имеющихся здесь результатов — в доказательстве леммы 8 (II) раздела 7, которая в свою очередь используется только в доказательстве теоремы 3 главы 4.

левское множество в евклидовом пространстве. Для краткости мы будем говорить, что $\mathcal{X}$ — евклидово и будем предполагать, если явно не оговорено противное, что $\mathcal{A}$ — класс борелевских подмножеств $\mathcal{X}$.

Теорема 4. *Если $\mathcal{X}$ — евклидово, то существует вариант выбора $P(A|t)$ такой, что при каждом t $P(A|t)$ будет вероятностной мерой на $\mathcal{A}$.*

Доказательство. Полагая вероятность любого борелевского множества, лежащего вне $\mathcal{X}$, равной нулю, мы можем распространить данную меру на класс всех борелевских множеств и можем, следовательно, не ограничивая общности, предположить, что $\mathcal{X}$ совпадает со всем пространством. Для простоты мы рассмотрим только одномерный случай. Выберем некоторый вариант вероятностной функции и при каждом x положим $F(x, t) = P((-\infty, x]|t)$. Пусть $r_1, r_2, \dots$ — рациональные числа, занумерованные в каком-либо порядке. Тогда неравенство $r_i < r_j$ влечет неравенство $F(r_i, t) \leq F(r_j, t)$ для всех t , кроме входящих в некоторое нулевое множество M_{ij} . Следовательно, $F(x, t)$ при каждом t , не входящем в нулевое множество $N' = \bigcup M_{ij}$, возрастает на множестве рациональных значений x . Аналогично, из леммы 5 вытекает, что для всех t , не принадлежащих некоторому нулевому множеству N'' , при $n \rightarrow \infty$ справедливы равенства: $\lim F(r_i + \frac{1}{n}, t) = F(r_i, t)$ для $i = 1, 2, \dots$; $\lim F(n, t) = 1$ и $\lim F(-n, t) = 0$. Поэтому для всех t , не входящих в нулевое множество $N' \cup N''$, $F(x, t)$, рассматриваемая как функция x (x — рационально), надлежаще нормирована на бесконечности, монотонна и непрерывна справа. Обозначим для t , не входящих в $N' \cup N''$, $F^*(x, t)$, единственную непрерывную справа по x функцию, которая совпадает с $F(x, t)$ для всех рациональных x . Тогда $F^*(x, t)$ будет функцией распределения и потому определяет некоторую вероятностную меру $P^*(A|t)$ на $\mathcal{A}$. Мы покажем теперь, что $P^*(A|t)$ является условной вероятностью A при данном t , для чего установим, что при каждом фиксированном A эта функция $\mathcal{B}$ -измерима по t и удовлетворяет (23). Последние утверждения мы получим из равенства

$$P^*(A|t) = P(A|t) \quad (\mathcal{B}, P^T)\text{-почти всюду}$$

при любом фиксированном $A \in \mathcal{A}$. По определению P^* это равенство верно, если A имеет вид $(-\infty, x]$, где x — рационально. Оно выполняется и для интервалов $(a, b] = (-\infty, b] - (-\infty, a]$, концы которых рациональны, так как P^* — мера, а для P верна лемма 5 (II). Следовательно, доказываемое равенство выполняется для всех элементов поля $\mathcal{F}$, образованного конечными суммами интервалов $(a_i, b_i]$ с рациональными концами. Класс множеств, для которых наше равенство выполняется, оказы-

вается монотонным (см. задачу 1) и поэтому содержит σ -поле, порожденное $\mathcal{F}$, т. е. $\mathcal{A}$. Мера $P^*(A|t)$ на $\mathcal{A}$, как она введена выше, определена для t , не входящих в $N' \cup N''$. Но поскольку ни измеримость функции, ни величина ее интеграла не зависят от значений на множестве меры нуль, то для $t \in N' \cup N''$ мы можем выбрать произвольные вероятностные меры на $\mathcal{A}$ и тем завершим построение.

Пусть X — случайный вектор с распределением P^X и T — статистика, заданная на $(\mathcal{X}, \mathcal{A})$. Пусть $P^{X|t}$ обозначает любой из вариантов семейства условных распределений $P(A|t)$ на $\mathcal{A}$, существование которых утверждается теоремой 4. Связь с понятием условного математического ожидания указывается следующей теоремой.

Теорема 5. Если X — случайный вектор и $E|f(X)| < \infty$, то $(\mathcal{B}, P^T)$ -почти всюду

$$E[f(X)|t] = \int f(x) dP^{X|t}(x). \quad (24)$$

Доказательство. Соотношение (24) выполняется, если f — индикатор любого множества $A \in \mathcal{A}$. Из леммы 3 следует, что оно верно для любой простой функции и потому для любой интегрируемой функции.

Вариант условного математического ожидания $E[f(X)|t]$, представляемый правой частью (24), обладает для каждого t обычными свойствами математических ожиданий ((I), (III) и (IV) в лемме 3). Ранее мы должны были исключать множества нулевой меры, зависящие от рассматривавшихся функций $f, g, \dots$. При условиях теоремы 4 аналогичное усиление допускает утверждение (II) леммы 3. Теперь мы в состоянии показать, что оно выполняется, за исключением, быть может, нулевого множества N , не зависящего от h . Для наших целей достаточно провести доказательство для случая, когда пространство значений статистики также евклидово *).

Теорема 6. Если T — статистика, у которой область определения $(\mathcal{X}, \mathcal{A})$ и пространство значений $(\mathcal{T}, \mathcal{B})$ — евклидовы, то найдется вариант $P^{X|t}$ условных вероятностных распределений и нулевое множество N такие, что условное математическое ожидание, вычисленное по формуле

$$E[f(X)|t] = \int f(x) dP^{X|t}(x)$$

для всех $t \notin N$, удовлетворяет соотношению

$$E[a(T)f(X)|t] = a(t)E[f(X)|t]. \quad (25)$$

*) Доказательство утверждения без этого ограничения см. в теореме A раздела 26 книги: L o é v e, Probability Theory, New York, D. Van Nostrand Co, 1955. (Русский перевод: М. Л о э в, Теория вероятностей, ИЛ, 1962.)

Доказательство. Для простоты допустим, что T — действительная функция (это по существу не ограничивает общности). Для каждого t обозначим $P^{X|t}(A)$ распределение вероятностей на $\mathcal{A}$, существование которого гарантируется теоремой 4. Для $B \in \mathcal{B}$ индикатор $I_B(t)$ $\mathcal{B}$ -измерим и

$$\int_{B'} I_B(t) dP^T(t) = P^T(B' \cap B) = P^X(T^{-1}B' \cap T^{-1}B).$$

Таким образом, в силу (20),

$$I_B(t) = P^{X|t}(T^{-1}B) \quad P^T\text{-почти всюду.}$$

Обозначим B_n , $n = 1, 2, \dots$, интервалы в $\mathcal{T}$ с рациональными концами. Тогда существует P -нулевое множество $N = \bigcup N_n$ такое, что для $t \notin N$

$$I_{B_n}(t) = P^{X|t}(T^{-1}B_n)$$

при всех n . При фиксированном $t \notin N$ две функции множества, $P^{X|t}(T^{-1}B)$ и $I_B(t)$, являются распределениями вероятностей на $\mathcal{B}$; второе из них приписывает множеству вероятность 1 или 0 в зависимости от того, содержит ли это множество точку t или нет. Так как значения этих функций совпадают на рациональных интервалах B_n , то они совпадают при всех $B \in \mathcal{B}$. В частности, для $t \notin N$ множество, состоящее из единственной точки t , входит в $\mathcal{B}$, и если обозначить

$$A^{(t)} = \{x: T(x) = t\},$$

то при всех $t \notin N$

$$P^{X|t}(A^{(t)}) = 1. \quad (26)$$

Таким образом, для $t \notin N$

$$\begin{aligned} \int a[T(x)] f(x) dP^{X|t}(x) &= \\ &= \int_{A^{(t)}} a[T(x)] f(x) dP^{X|t}(x) = a(t) \int f(x) dP^{X|t}(x), \end{aligned}$$

что и требовалось доказать.

Как следствие теоремы 6 получаем, что для $t \notin N$ $E[a(T)|t] = a(t)$ и, в частности, $P(T \in B|t)$ равна 1 или 0 при $t \in B$ или $t \notin B$.

Условные распределения $P^{X|t}$ отличаются от введенных в элементарном случае (раздел 9 главы 1) тем, что определены на $(\mathcal{B}, \mathcal{A})$, а не на $A^{(t)}$ и σ -поле $\mathcal{A}^{(t)}$ его борелевских подмножеств. Однако (26) влечет при всех $t \notin N$

$$P^{X|t}(A) = P^{X|t}(A \cap A^{(t)}).$$

Вычисления с условными вероятностями и математическими ожиданиями можно производить, следовательно, заменив для $t \notin N$ $P^{X|t}$ на распределение $\bar{P}^{X|t}$, которое определено на $(A^{(t)}, \mathcal{A}^{(t)})$ и приписывает каждому подмножеству $A^{(t)}$ ту же вероятность, что и $P^{X|t}$.

Теорема 6 устанавливает при $t \notin N$ существование условных распределений вероятностей $\bar{P}^{X|t}$, определенных на $(A^{(t)}, \mathcal{A}^{(t)})$, для которых по лемме 4

$$E[f(X)] = \int_{\mathcal{T}-N} \left[\int_{A^{(t)}} f(x) d\bar{P}^{X|t}(x) \right] dP^T(t) \quad (27)$$

при любой интегрируемой функции f . Обратно, рассмотрим любое семейство распределений, удовлетворяющих (27), и эксперимент, в котором сначала наблюдается T , а при $T = t$ — случайная величина с распределением $\bar{P}^{X|t}$. Результатом этой двухступенной процедуры будет случайная точка в $(\mathcal{X}, \mathcal{A})$, распределенная по тому же закону, что и X . Таким образом, $\bar{P}^{X|t}$ удовлетворяет этому «функциональному» определению условной вероятности.

Если $(\mathcal{X}, \mathcal{A})$ является произведением $(\mathcal{T} \times \mathcal{Y}, \mathcal{B} \times \mathcal{C})$, то $A^{(t)}$ будет произведением $\mathcal{Y}$ на множество, состоящее из одной точки t . Для $t \notin N$ условное распределение $\bar{P}^{X|t}$ индуцирует распределение на $(\mathcal{Y}, \mathcal{C})$, которое по аналогии с элементарным случаем будет обозначаться $P^{Y|t}$. В этом случае определение можно распространить на все $\mathcal{T}$, принимая, например, что при $t \in N$ $P^{Y|t}$ приписывает вероятность 1 некоторой выбранной нами точке y_0 . При таком определении из (27) выводим

$$Ef(T, Y) = \int_{\mathcal{T}} \left[\int_{\mathcal{Y}} f(t, y) dP^{Y|t}(y) \right] dP^T(t). \quad (28)$$

В качестве приложения мы докажем следующую лемму, которая затем будет использована в разделе 7.

Лемма 6. Пусть $(\mathcal{T}, \mathcal{B})$ и $(\mathcal{Y}, \mathcal{C})$ — евклидовы пространства, и пусть $P_0^{T, Y}$ — некоторое распределение на произведении $(\mathcal{X}, \mathcal{A}) = (\mathcal{T} \times \mathcal{Y}, \mathcal{B} \times \mathcal{C})$. Предположим, что другое распределение P_1 на $(\mathcal{X}, \mathcal{A})$ таково, что

$$dP_1(t, y) = a(y) b(t) dP_0(t, y)$$

с $a(y) > 0$ при всех y . Тогда при P_1 частное распределение T и вариант условных распределений Y при данном t даются формулами:

$$dP_1^T(t) = b(t) \left[\int a(y) dP_0^{Y|t}(y) \right] dP_0^T(t)$$

и

$$dP_1^{Y|t}(y) = \frac{a(y) dP_0^{Y|t}(y)}{\int_{\mathcal{Y}} a(y') dP_0^{Y|t}(y')}.$$

Доказательство. Первое утверждение леммы следует из равенств

$$\begin{aligned} P_1\{T \in B\} &= E_1[I_B(T)] = E_0[I_B(T) a(Y) b(T)] = \\ &= \int_B b(t) \left[\int_{\mathcal{Y}} a(y) dP_0^{Y|t}(y) \right] dP_0^T(t). \end{aligned}$$

Чтобы проверить второе утверждение, достаточно показать, что для любой интегрируемой f математическое ожидание $E_1 f(Y, T)$ удовлетворяет (28). Но это получается сразу. Заметим, что знаменатель у $dP_1^{Y|t}$ положителен, так как $a(y) > 0$ при всех y .

6. Характеристика достаточности

Мы можем теперь обобщить определение достаточности, данное в разделе 9 главы 1. Если $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ — любое семейство распределений, определенных на одном и том же выборочном пространстве $(\mathcal{X}, \mathcal{A})$, то статистика T называется *достаточной* для $\mathcal{P}$ (или для θ), если для каждого A из $\mathcal{A}$ можно выбрать вариант условной вероятностной функции $P_\theta(A|t)$, не зависящий от θ . В качестве примера возьмем случай, когда $X_1, \dots, X_n$ независимы и имеют общую функцию распределения $F_\theta, \theta \in \Omega$. Из примера 7 следует, что совокупность $T(X) = (X^{(1)}, \dots, X^{(n)})$ порядковых статистик достаточна для θ .

Теорема 7. Если $\mathcal{X}$ — евклидово и статистика T достаточна для $\mathcal{P}$, то найдется вариант условных распределений вероятностей $P_\theta(A|t)$, не зависящий от θ , и такой, что для каждого фиксированного t $P(A|t)$ — вероятностная мера на $\mathcal{A}$.

Доказательство. Справедливость утверждения можно усмотреть из доказательства теоремы 4. По определению достаточности мы можем при каждом рациональном r выбрать функцию $F(r, t)$ не зависящей от θ . Получаемое этим путем условное распределение также не будет зависеть от θ .

В главе 1 понятие достаточности разъяснялось указанием на то, что в некотором смысле достаточная статистика содержит всю доступную информацию. Теорема 7 показывает, что это замечание сохраняет силу в общем случае евклидовых выборочных пространств. Исходя из достаточной статистики T , мы мо-

жем с помощью случайного механизма построить случайный вектор X' , имеющий то же самое распределение, что и первоначальный вектор X . Другое обобщение прежнего результата, не связанное с ограничением евклидовыми выборочными пространствами, дано в задаче 11.

Критерий факторизации для достаточных статистик, выведенный в главе 1, может быть распространен на любое *доминированное* семейство распределений, т. е. на любое семейство $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ распределений, обладающих плотностями p_θ по отношению к σ -конечной мере на $(\mathcal{X}, \mathcal{A})$. Доказательство этого утверждения основано на существовании распределения вероятностей $\lambda = \sum c_i P_{\theta_i}$ (см. теорему 2 в дополнении), *эквивалентного* $\mathcal{P}$ в том смысле, что для любого $A \in \mathcal{A}$

$$\begin{aligned} \lambda(A) = 0 & \text{ тогда и только тогда,} \\ \text{когда } P_\theta(A) = 0 & \text{ при всех } \theta \in \Omega. \end{aligned} \quad (29)$$

Теорема 8. Пусть $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ — доминированное семейство распределений вероятностей на $(\mathcal{X}, \mathcal{A})$ и пусть $\lambda = \sum c_i P_{\theta_i}$ удовлетворяет (29). Тогда статистика T с пространством значений $(\mathcal{T}, \mathcal{B})$ достаточна для $\mathcal{P}$ в том и только том случае, когда существует неотрицательная $\mathcal{B}$ -измеримая функция $g_\theta(t)$ такая, что

$$dP_\theta(x) = g_\theta[T(x)] d\lambda(x) \quad (30)$$

для всех $\theta \in \Omega$.

Доказательство. Предположим, что T достаточна для θ и пусть $\mathcal{A}_0$ — подполе, индуцированное T . Тогда для всех $\theta \in \Omega$, $A_0 \in \mathcal{A}_0$ и $A \in \mathcal{A}$

$$\int_{A_0} P(A|T(x)) dP_\theta(x) = P_\theta(A \cap A_0),$$

и так как $\lambda = \sum c_i P_{\theta_i}$, то

$$\int_{A_0} P(A|T(x)) d\lambda(x) = \lambda(A \cap A_0),$$

т. е. $P(A|T(x))$ является условной вероятностной функцией также и для λ . Пусть $g_\theta(T(x))$ обозначает производную Радона — Никодима $dP_\theta(x)/d\lambda(x)$ для $(\mathcal{A}_0, \lambda)$. Для доказательства (30) необходимо установить, что $g_\theta(T(x))$ является также производной P_θ для $(\mathcal{A}, \lambda)$. Последнее следует из приводимых ниже выкладок. В качестве начального шага мы полагаем в первом

из приведенных выше соотношений A_0 равным $\mathcal{E}$:

$$\begin{aligned} P_\theta(A) &= \int P(A|T(x)) dP_\theta(x) = \int E_\lambda[I_A(x)|T(x)] dP_\theta(x) = \\ &= \int E_\lambda[I_A(x)|T(x)] g_\theta(T(x)) d\lambda(x) = \\ &= \int E_\lambda[g_\theta(T(x)) I_A(x)|T(x)] d\lambda(x) = \\ &= \int_A g_\theta(T(x)) I_A(x) d\lambda(x) = \int_A g_\theta(T(x)) d\lambda(x). \end{aligned}$$

Здесь второе равенство использует установленное в самом начале доказательства утверждение о том, что $P(A|T(x))$ — условное распределение для λ ; третье равенство выполняется, так как подынтегральная функция $\mathcal{A}_0$ -измерима и $dP_\theta = g_\theta d\lambda$ для $(\mathcal{A}_0, \lambda)$; четвертое равенство — приложение леммы 3 (II); пятое использует определение условных математических ожиданий.

Предположим теперь обратное, т. е. что (30) имеет место. Мы докажем, что условная вероятностная функция $P_\lambda(A|t)$ является в то же время условной вероятностной функцией для всех $P \in \mathcal{P}$. Пусть $g_\theta(T(x)) = dP_\theta(x)/d\lambda(x)$ на $\mathcal{A}$. Для фиксированных A и θ введем меру ν на $\mathcal{A}$, определив ее равенством $d\nu = I_A dP_\theta$. Тогда на $\mathcal{A}_0$ $dv(x)/dP_\theta(x) = E_\theta[I_A(X)|T(x)]$ и поэтому

$$dv(x)/d\lambda(x) = P_\theta[A|T(x)] g_\theta(T(x)) \text{ на } \mathcal{A}_0.$$

С другой стороны, $dv(x)/d\lambda(x) = I_A(x) g_\theta(T(x))$ на $\mathcal{A}$, и, следовательно,

$$\frac{dv(x)}{d\lambda(x)} = E_\lambda[I_A(X) g_\theta(T(X))|T(x)] = P_\lambda[A|T(x)] g_\theta(T(x)) \text{ на } \mathcal{A}_0.$$

Поэтому $(\mathcal{A}_0, \lambda)$ -почти всюду (а стало быть, и $(\mathcal{A}_0, P_\theta)$ -почти всюду) $P_\lambda(A|T(x)) g_\theta(T(x)) = P_\theta(A|T(x)) g_\theta(T(x))$. Так как $(\mathcal{A}_0, P_\theta)$ -почти всюду $g_\theta(T(x)) \neq 0$, то это равенство показывает, что $P_\theta(A|T(x)) = P_\lambda(A|T(x))$ $(\mathcal{A}_0, P_\theta)$ -почти всюду, т. е. что $P_\lambda(A|T(x))$ является вариантом $P_\theta(A|T(x))$.

Вместо вышеприведенной формулировки, в которой явно участвует распределение λ , иногда бывает удобна другая, данная непосредственно в терминах доминирующей меры μ .

Следствие 1. (Теорема факторизации.) Пусть распределения P_θ семейства $\mathcal{P}$ имеют плотности $p_\theta = dP_\theta/d\mu$ по отношению к σ -конечной мере μ . Статистика T достаточна для $\mathcal{P}$ тогда и только тогда, когда существуют неотрицательная $\mathcal{B}$ -измеримая функция g_θ на $\mathcal{T}$ и неотрицательная $\mathcal{A}$ -измеримая функция h на $\mathcal{X}$ такие, что

$$p_\theta(x) = g_\theta[T(x)] h(x) \quad (\mathcal{A}, \mu)\text{-почти всюду.} \quad (31)$$

Доказательство. Пусть $\lambda = \sum c_i P_{\theta_i}$ удовлетворяет (29). Тогда, если T достаточна, то (31) следует из (30) с $h = d\lambda/d\mu$. Обратно, если выполняется (31), то

$$d\lambda(x) = \sum c_i g_{\theta_i}[T(x)] h(x) d\mu(x) = k[T(x)] h(x) d\mu(x)$$

и потому $dP_{\theta}(x) = g_{\theta}^*(T(x)) d\lambda(x)$, где $g_{\theta}^*(t) = g_{\theta}(t)/k(t)$ при $k(t) > 0$ и может быть определено произвольным образом при $k(t) = 0$.

7. Экспоненциальные семейства

Важным примером семейств распределений, допускающих редукцию с помощью достаточных статистик, являются *экспоненциальные семейства*, плотность которых по отношению к некоторой σ -конечной мере μ на евклидовом выборочном пространстве $(\mathcal{X}, \mathcal{A})$ имеет вид

$$p_{\theta}(x) = C(\theta) \exp \left[\sum_{j=1}^r Q_j(\theta) T_j(x) \right] h(x). \quad (32)$$

Распределение выборки $X = (X_1, \dots, X_n)$, взятой из биномиальной, пуассоновской или нормальной совокупности, относится к этому классу. В биномиальном случае, например, плотность по отношению к мере, которая для каждого целого числа равна единице, имеет вид

$$\binom{n}{x} p^x (1-p)^{n-x} = (1-p)^n \exp \left[x \log \left(\frac{p}{1-p} \right) \right] \binom{n}{x}.$$

Пример 8. Если величины $Y_1, \dots, Y_n$ независимы и каждая из них имеет плотность (по отношению к мере Лебега), равную

$$p_{\sigma}(y) = \frac{y^{\frac{f}{2}-1} \exp[-y/(2\sigma^2)]}{(2\sigma^2)^{\frac{f}{2}} \Gamma(f/2)}, \quad y > 0, \quad (33)$$

то совместное распределение Y_k образует экспоненциальное семейство. При $\sigma = 1$ (33) является плотностью χ^2 -распределения с f степенями свободы; в частности, при целом f последняя плотность совпадает с плотностью суммы

$\sum_{j=1}^f X_j^2$, где X_j образуют выборку из нормального распределения $N(0, 1)$.

Пример 9. Рассмотрим n независимых испытаний, каждое из которых с вероятностями $p_1, \dots, p_s$ заканчивается одним из исходов $E_1, \dots, E_s$. Пусть X_{ij} равно 1, если i -е испытание закончилось E_j -м исходом и равно 0 в других случаях. Тогда совместное распределение X_K имеет вид

$$P\{X_{11} = x_{11}, \dots, X_{ns} = x_{ns}\} = p_1^{\sum x_{1i}} p_2^{\sum x_{2i}} \dots p_s^{\sum x_{si}},$$

где каждое x_{ij} равно 0 или 1 и $\sum_j x_{ij} = 1$. Это семейство снова экспоненциаль-

ное с $T_j(x) = \sum_{i=1}^n x_{ij}$ ($j = 1, \dots, s-1$). Статистики T_j имеют совместное полиномиальное распределение

$$P\{T_1=t_1, \dots, T_{s-1}=t_{s-1}\} = \frac{n!}{t_1! \dots t_{s-1}! (n-t_1-\dots-t_{s-1})!} \times \\ \times p_1^{t_1} \dots p_{s-1}^{t_{s-1}} (1-p_1-\dots-p_{s-1})^{n-t_1-\dots-t_{s-1}}. \quad (34)$$

Если $X_1, \dots, X_n$ — выборка из распределения с плотностью (32), то совместное распределение X_i ($i = 1, \dots, n$) образует экспоненциальное семейство с достаточными статистиками

$\sum_{i=1}^n T_j(X_i)$ ($j = 1, \dots, k$). Таким образом, сколь бы велик ни был объем выборки, для $(X_1, \dots, X_n)$ существует k -мерная достаточная статистика. Обратно, пусть $X_1, \dots, X_n$ является выборкой из распределения с некоторой плотностью $p_\theta(x)$ и пусть множество, на котором эта плотность положительна, не зависит от θ . Тогда при условиях регулярности, позволяющих осмысленно применить понятие размерности, можно и из существования k -мерной достаточной статистики с $k < n$ вывести, что плотности $p_\theta(x)$ образуют экспоненциальное семейство *).

Используя более естественную параметризацию и включая множитель $h(x)$ в μ , мы придадим экспоненциальному семейству форму $dP_\theta(x) = p_\theta(x) d\mu(x)$, где

$$p_\theta(x) = C(\theta) \exp \left[\sum_{j=1}^k \theta_j T_j(x) \right]. \quad (35)$$

Правая часть (35), если ее интеграл конечен, может быть надлежащим выбором $C(\theta)$ превращена в плотность вероятности. Множество Ω параметрических точек $\theta = (\theta_1, \dots, \theta_k)$, для которых это имеет место, называется *естественным параметрическим пространством* экспоненциального семейства (35).

Оптимальные критерии для проверки различных гипотез, касающихся θ_j , находятся в главе 4. Мы установим сейчас некоторые свойства экспоненциальных семейств, необходимые для этой цели.

Л е м м а 7. *Естественное параметрическое пространство экспоненциального семейства выпукло.*

*) Доказательство и формулировку условий регулярности см. в работе: Koopman, On distributions admitting a sufficient statistic, Trans. Am. Math. Soc., 39 (1936), 399—409. Результат обсуждается также у Darmois, Sur les lois de probabilité à estimation exhaustive, Compt. Rend. Acad. Sci., Paris, 260 (1935), 1265—1266, и у Pitman, Sufficient Statistics and intrinsic accuracy, Proc. Cambridge Phil. Soc. 32 (1936), 567—579.

Доказательство. Пусть $(\theta_1, \dots, \theta_k)$ и $(\theta'_1, \dots, \theta'_k)$ — две параметрические точки, для которых интеграл от (35) конечен. Тогда по равенству Гельдера для любого $0 < \alpha < 1$

$$\begin{aligned} \int \exp \left[\sum [\alpha \theta_i + (1 - \alpha) \theta'_i] T_i(x) \right] d\mu(x) &\leq \\ &\leq \left[\int \exp \left[\sum \theta_i T_i(x) \right] d\mu(x) \right]^\alpha \left[\int \exp \left[\sum \theta'_i T_i(x) \right] d\mu(x) \right]^{1-\alpha} < \infty. \end{aligned}$$

Если выпуклое множество Ω лежит в линейном пространстве размерности меньшей k , то (35) может быть переписано в форме, включающей менее k компонент T . Поэтому мы, не ограничивая общности, предположим Ω k -мерным.

Из теоремы о факторизации вытекает, что статистика $T(x) = (T_1(x), \dots, T_k(x))$ достаточна для $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$.

Лемма 8. Пусть X имеет распределение из экспоненциального семейства

$$dP_{\theta, \vartheta}^X(x) = C(\theta, \vartheta) \exp \left[\sum_{i=1}^r \theta_i U_i(x) + \sum_{j=1}^s \vartheta_j T_j(x) \right] d\mu(x).$$

Тогда существуют меры λ_θ и вероятностные меры ν_t на s и r -мерном евклидовом пространствах соответственно, такие, что

(I) распределение $T = (T_1, \dots, T_s)$ принадлежит экспоненциальному семейству

$$dP_{\theta, \vartheta}^T(t) = C(\theta, \vartheta) \exp \left[\sum_{j=1}^s \vartheta_j t_j \right] d\lambda_\theta(t), \quad (36)$$

(II) условное распределение $U = (U_1, \dots, U_r)$ при данном $T = t$ принадлежит экспоненциальному семейству вида

$$dP_{\theta}^{U|T}(u) = C_t(\theta) \exp \left[\sum_{i=1}^r \theta_i u_i \right] d\nu_t(u), \quad (37)$$

т. е. в частности не зависит от ϑ .

Доказательство. Пусть (θ^0, ϑ^0) — точка естественного параметрического пространства и $\mu^* = P_{\theta^0, \vartheta^0}^X$. Тогда

$$\begin{aligned} dP_{\theta, \vartheta}^X(x) &= \\ &= \frac{C(\theta, \vartheta)}{C(\theta^0, \vartheta^0)} \exp \left[\sum_{i=1}^r (\theta_i - \theta_i^0) U_i(x) + \sum_{j=1}^s (\vartheta_j - \vartheta_j^0) T_j(x) \right] d\mu^*(x), \end{aligned}$$

и результат следует из леммы 6 с

$$d\lambda_\theta(t) = \exp \left(\sum \vartheta_i^0 t_i \right) \left[\int \exp \left[\sum_{i=1}^r (\theta_i - \theta_i^0) u_i \right] dP_{\theta^0, \vartheta^0}^{U|T}(u) \right] dP_{\theta^0, \vartheta^0}^T(t)$$

и

$$dv_t(u) = dP_{\theta^0, \theta^0}^{U|t}(u).$$

Теорема 9. Пусть φ — ограниченная измеримая функция на $(\mathcal{X}, \mathcal{A})$. Тогда
(I) интеграл

$$\int \varphi(x) \exp \left[\sum_{j=1}^k \theta_j T_j(x) \right] d\mu(x), \quad (38)$$

рассматриваемый как функция комплексных переменных $\theta_j = \xi_j + i\eta_j$ ($j = 1, \dots, k$), является аналитической по каждой из этих переменных в области R параметрических точек, для которых $(\xi_1, \dots, \xi_k)$ является внутренней точкой естественного параметрического пространства Ω ;

(II) производные любого порядка от интеграла (38) по θ могут вычисляться дифференцированием под знаком интеграла.

Доказательство. Если $|\varphi| \leq M$, то

$$|\varphi(x) \exp[\sum \theta_j T_j(x)]| \leq M \exp[\sum \xi_j T_j(x)],$$

так что интеграл (38) существует и конечен для всех точек $(\xi_1, \dots, \xi_k)$ из Ω . Пусть $(\xi_1^0, \dots, \xi_k^0)$ — некоторая фиксированная внутренняя точка Ω . Рассмотрим одну из величин, например θ_1 . Выделяя в множителе $\varphi(x) \exp[(\xi_2^0 + i\eta_2^0) T_2(x) + \dots + (\xi_k^0 + i\eta_k^0) T_k(x)]$ его действительную и мнимую части и разлагая каждую из последних на положительную и отрицательную части, включая затем надлежащие множители в меру μ , мы представим интеграл (38) в виде

$$\begin{aligned} & \int \exp[\theta_1 T_1(x)] d\mu_1(x) - \int \exp[\theta_1 T_1(x)] d\mu_2(x) + \\ & + i \int \exp[\theta_1 T_1(x)] d\mu_3(x) - i \int \exp[\theta_1 T_1(x)] d\mu_4(x). \end{aligned}$$

Достаточно доказать требуемый результат для интегралов вида

$$\psi(\theta_1) = \int \exp[\theta_1 T_1(x)] d\mu(x).$$

Так как $(\xi_1^0, \dots, \xi_k^0)$ — внутренняя точка Ω , то существует такое $\delta > 0$, что $\psi(\theta_1)$ конечна для всех θ_1 с $|\xi_1 - \xi_1^0| \leq \delta$. Рассмотрим разностное отношение

$$\frac{\psi(\theta_1) - \psi(\theta_1^j)}{\theta_1 - \theta_1^j} = \int \frac{\exp[\theta_1 T_1(x)] - \exp[\theta_1^j T_1(x)]}{\theta_1 - \theta_1^j} d\mu(x).$$

Подынтегральное выражение может быть записано в виде

$$\exp[\theta_1^0 T_1(x)] \left[\frac{\exp[(\theta_1 - \theta_1^0) T_1(x)] - 1}{\theta_1 - \theta_1^0} \right].$$

Применяя ко второму множителю неравенство

$$\left| \frac{\exp(az) - 1}{z} \right| \leq \frac{\exp(\delta |a|)}{\delta} \quad \text{при } |z| \leq \delta,$$

мы видим, что подынтегральное выражение по абсолютной величине не превосходит

$$\frac{1}{\delta} |\exp(\theta_1^0 T_1 + \delta |T_1|)| \leq \frac{1}{\delta} |\exp[(\theta_1^0 + \delta) T_1] + \exp[(\theta_1^0 - \delta) T_1]|,$$

если $|\theta_1 - \theta_1^0| \leq \delta$. Так как правая часть интегрируема, то из теоремы Лебега об ограниченной сходимости (теорема 1 (II)) вытекает, что для любой последовательности точек $\theta_1^{(n)}$, сходящейся к θ_1^0 , разностное отношение ψ стремится к

$$\int T_1(x) \exp[\theta_1^0 T_1(x)] d\mu(x).$$

Этим завершается доказательство (I) и доказывается (II) для случая первой производной. Доказательство для высших производных проводится по индукции и вполне аналогично тому, что сделано выше.

8. Задачи

К разделу 1

1. *Монотонные классы.* Класс $\mathcal{F}$ подмножеств данного пространства называется *полем*, если он содержит все пространство, замкнут относительно дополнений и конечных пересечений. Класс $\mathcal{M}$ называется *монотонным*, если объединение (пересечение) любой возрастающей (убывающей) последовательности множеств из $\mathcal{M}$ снова входит в $\mathcal{M}$. Наименьший монотонный класс $\mathcal{M}_0$, содержащий данное поле $\mathcal{F}$, совпадает с наименьшим σ -полем $\mathcal{A}$, содержащим $\mathcal{F}$.

[Сначала докажете, что $\mathcal{M}_0$ — поле. Чтобы показать, например, что $A \cap B \in \mathcal{M}_0$, когда A и B принадлежат $\mathcal{M}_0$, рассмотрим для фиксированного $A \in \mathcal{F}$ класс $\mathcal{M}_A$ всех $B \in \mathcal{M}_0$, для которых $A \cap B \in \mathcal{M}_0$. Тогда $\mathcal{M}_A$ — монотонный класс, содержащий $\mathcal{F}$, и потому $\mathcal{M}_A = \mathcal{M}_0$. Таким образом, $A \cap B \in \mathcal{M}_0$ при всех B . Эти рассуждения могут быть теперь повторены с произвольным фиксированным $B \in \mathcal{M}_0$ и классом $\mathcal{M}_B$ тех множеств $A \in \mathcal{M}_0$, для которых $A \cap B \in \mathcal{M}_0$. Так как $\mathcal{M}_0$ — поле и к тому же монотонно, то оно является σ -полем, содержащим $\mathcal{F}$, а следовательно, и $\mathcal{A}$. Но любое σ -поле является монотонным классом, так что $\mathcal{M}_0$ содержится в $\mathcal{A}$.]

К разделу 2

2. Производные Радона — Никодима. (I) Если λ и μ — σ -конечные меры на $(\mathcal{X}, \mathcal{A})$ и μ абсолютно непрерывна по отношению к λ , то

$$\int f d\mu = \int f \frac{d\mu}{d\lambda} d\lambda$$

для любой μ -интегрируемой функции f .

(II) Если λ , μ и ν — σ -конечные меры на $(\mathcal{X}, \mathcal{A})$ такие, что ν абсолютно непрерывна по μ , а μ — по λ , то

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda} \quad \lambda\text{-почти всюду.}$$

(III) Если μ и ν — σ -конечные меры, которые эквивалентны в том смысле, что каждая абсолютно непрерывна по другой, то

$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu} \right)^{-1} \quad \mu, \nu\text{-почти всюду.}$$

(IV) Если μ_k , $k = 1, 2, \dots$, и μ — σ -конечные меры на $(\mathcal{X}, \mathcal{A})$ такие, что $\sum_{k=1}^{\infty} \mu_k(A) = \mu(A)$ при всех $A \in \mathcal{A}$, и если μ_k абсолютно непрерывны по отношению к σ -конечной мере λ , то и μ абсолютно непрерывна по λ и

$$\frac{d \sum_{k=1}^n \mu_k}{d\lambda} = \sum_{k=1}^n \frac{d\mu_k}{d\lambda}, \quad \lim_{n \rightarrow \infty} \frac{d \sum_{k=1}^n \mu_k}{d\lambda} = \frac{d\mu}{d\lambda} \quad \lambda\text{-почти всюду.}$$

[(I) Рассматриваемое соотношение выполняется в случае, когда f — индикатор некоторого множества. Отсюда переходим к простым f и затем к интегрируемым.]

(II) Применить (I) с $f = d\nu/d\mu$.]

К разделу 3

3. Пусть $(\mathcal{X}, \mathcal{A})$ — измеримое пространство и $\mathcal{A}_0$ — σ -поле, содержащееся в $\mathcal{A}$. Для любой функции T определим σ -поле $\mathcal{B}$ как совокупность тех множеств B , для которых $T^{-1}(B) \in \mathcal{A}_0$. Может не существовать функция T такая, что $T^{-1}(\mathcal{B}) = \mathcal{A}_0$.

[Примером может служить любое $\mathcal{A}_0$, содержащее все одноточечные множества.]

К разделу 4

4. (I) Пусть $\mathcal{P}$ — семейство распределений $X = (X_1, \dots, X_n)$ таких, что

$$P\{(X_i, X_{i+1}, \dots, X_n, X_1, \dots, X_{i-1}) \in A\} = P\{(X_1, \dots, X_n) \in A\}$$

для всех борелевских множеств A и всех $i = 1, \dots, n$. Для любой выборочной точки $(x_1, \dots, x_n)$ определим $(y_1, \dots, y_n) = (x_i, x_{i+1}, \dots, x_n, x_1, \dots, x_{i-1})$, где $x_i = x^{(i)} = \min(x_1, \dots, x_n)$. Тогда условное математическое ожидание $f(x)$ при данном $Y = y$ равно

$$f_0(y_1, \dots, y_n) = \frac{1}{n} [f(y_1, \dots, y_n) + f(y_2, \dots, y_n, y_1) + \dots + f(y_n, y_1, \dots, y_{n-1})].$$

(II) Пусть $G = \{g_1, \dots, g_r\}$ — некоторая группа перестановок координат $x_1, \dots, x_n$ точек x n -мерного пространства. Обозначим gx точку, получающуюся применением g к координатам x . Пусть $\mathcal{P}$ — некоторое семейство распределений P вектора $X = (X_1, \dots, X_n)$. Допустим, что для P имеет место

$$P\{gX \in A\} = P\{X \in A\} \quad \text{для всех } g \in G. \quad (39)$$

Для любой точки x выберем какую-либо одну из k точек $g_k x$, $k = 1, \dots, r$ (например, точку с наименьшей первой координатой, если она этим требованием определяется однозначно, или, в противном случае, точку с наименьшей второй координатой и т. д.). Полученную функцию обозначим $T(x) = t$. Тогда

$$E[f(X) | t] = \frac{1}{r} \sum_{k=1}^r f(g_k t).$$

(III) Допустим, что в (II) распределение P не удовлетворяет условию инвариантности (39), а определяется формулой

$$dP(x) = h(x) d\mu(x),$$

где μ инвариантна в том смысле, что $\mu\{x: gx \in A\} = \mu(A)$. Тогда

$$E[f(X) | t] = \frac{\sum_{k=1}^r f(g_k t) h(g_k t)}{\sum_{k=1}^r h(g_k t)}.$$

К разделу 5

5. Докажите теорему 4 для случая n -мерного выборочного пространства [Условие возрастания функции распределения заменяется на

$$P\{x_1 < X_1 \leq x'_1, \dots, x_n < X_n \leq x'_n\} \geq 0.$$

Условие непрерывности справа принимает вид

$$\lim_{m \rightarrow \infty} F\left(x_1 + \frac{1}{m}, \dots, x_n + \frac{1}{m}\right) = F(x_1, \dots, x_n).$$

6. Пусть $\mathcal{X} = \mathcal{Y} \times \mathcal{T}$ и P_0, P_1 — два распределения, для которых

$$dP_0(y, t) = f(y) g(t) d\mu(y) dv(t), \quad dP_1(y, t) = h(y, t) d\mu(y) dv(t),$$

где $h(y, t)/f(y)g(t) < \infty$. Тогда при P_1 плотность Y по отношению к μ равна

$$p_1^Y(y) = f(y) E_0 \left[\frac{h(y, T)}{f(y) g(T)} \mid Y = y \right].$$

$$\left[p_1^Y(y) = \int_{\mathcal{T}} h(y, t) dv(t) = f(y) \int_{\mathcal{T}} \frac{h(y, t)}{f(y) g(t)} g(t) dv(t). \right]$$

К разделу 6

7. Симметричные распределения. (I) Пусть $\mathcal{P}$ — некоторое семейство распределений $X = (X_1, \dots, X_n)$, симметричных в том смысле, что

$$P\{(X_{i_1}, \dots, X_{i_n}) \in A\} = P\{(X_1, \dots, X_n) \in A\}$$

для всех борелевских множеств A и всех перестановок $(i_1, \dots, i_n)$ чисел $(1, \dots, n)$. Тогда статистика T примера 7 достаточна для $\mathcal{P}$, и формула для условного математического ожидания $E[f(X)|T(x)]$, данная в первой части примера, остается справедливой.

(II) Статистика Y задачи 4 достаточна.

(III) Пусть $X_1, \dots, X_n$ независимы и имеют одно и то же непрерывное распределение $P \in \mathcal{P}$. Предположим, что распределения класса $\mathcal{P}$ симметричны по отношению к началу координат. Пусть $V_i = |X_i|$ и $W_i = V_i^{(i)}$. Тогда статистика $(W_1, \dots, W_n)$ достаточна для $\mathcal{P}^*$.

8. Достаточность отношений правдоподобия. Пусть P_0, P_1 — два распределения с плотностями p_0, p_1 . Тогда статистика $T(x) = p_1(x)/p_0(x)$ достаточна для $\mathcal{P} = \{P_0, P_1\}$.

[Это следует из критерия факторизации и равенств

$$p_1 = T \cdot p_0, \quad p_0 = 1 \cdot p_0.]$$

9. Попарная достаточность. Статистика T называется попарно достаточной для $\mathcal{P}$, если она достаточна для каждой пары распределений из $\mathcal{P}$.

(I) Если $\mathcal{P}$ счетно и T — попарно достаточна для $\mathcal{P}$, то T достаточна для $\mathcal{P}$.

(II) Если $\mathcal{P}$ — доминированное семейство и T попарно достаточна для $\mathcal{P}$, то T достаточна для $\mathcal{P}$.

[(I) Пусть $\mathcal{P} = \{P_0, P_1, \dots\}$ и $\mathcal{A}_0$ — достаточное подполе, индуцированное T . Пусть $\lambda = \sum c_j P_j$ ($c_j > 0$) эквивалентно $\mathcal{P}$. При каждом $j = 1, 2, \dots$ вероятностная мера λ_j , пропорциональная $(c_0/n)P_0 + c_j P_j$, эквивалентна $\{P_0, P_j\}$. В силу попарной достаточности производная

$$f_j = dP_0 / [(c_0/n)dP_0 + c_j dP_j]$$

является $\mathcal{A}_0$ -измеримой. Пусть $S_j = \{x: f_j(x) = 0\}$ и $S = \bigcup_{j=1}^n S_j$. Тогда $S \in \mathcal{A}_0$, $P_0(S) = 0$ и на $\mathcal{X} - S$ производная $dP_0/d \sum_{j=1}^n c_j P_j$ равна $\sum_{j=1}^n 1/f_j$, т. е.

$\mathcal{A}_0$ -измеримой функции. Из задачи 2 вытекает, что тогда $\mathcal{A}_0$ -измерима и

$$\frac{dP_0}{d\lambda} = \frac{dP_0}{d \sum_{j=1}^n c_j P_j} \frac{d \sum_{j=1}^n c_j P_j}{d\lambda}.$$

(II) Пусть $\lambda = \sum_{j=1}^{\infty} c_j P_{\theta_j}$ эквивалентно $\mathcal{P}$. Из попарной достаточности T следует, что при любом θ_0 $dP_{\theta_0}/(dP_{\theta_0} + d\lambda)$ и $dP_{\theta_0}/d\lambda$ являются измеримыми функциями от T .]

10. Если статистика T достаточна для $\mathcal{P}$, то, какова бы ни была $(\mathcal{A}, P_\theta)$ -интегрируемая при всех $\theta \in \Omega$ функция f , найдется вариант условной функции $E_\theta[f(X)|t]$, не зависящий от θ .

[Если $\mathcal{X}$ — евклидово, то утверждение следует из теорем 5 и 7. В общем случае, если f неотрицательна, то существует неубывающая последовательность неотрицательных простых функций f_n , сходящаяся к f . По лемме 3 (II) условное математическое ожидание простой функции может быть выбрано не зависящим от θ . Теперь желаемый результат вытекает из леммы 3 (IV).]

*) Обозначения заимствованы из примера 7. — Прим. пер.

11. В проблеме решения с конечным числом возможных решений класс процедур, зависящих только от достаточной статистики T , является существенно полным *).

[Для евклидовых выборочных пространств это следует из теоремы 4 без каких бы то ни было ограничений на пространство решений. Для рассматриваемого случая обозначим решающую процедуру $\delta(x) = (\delta^{(1)}(x), \dots, \delta^{(m)}(x))$, где $\delta^{(i)}(x)$ — вероятность, с которой принимают решение d_i , если наблюдают x . Если T достаточна и $\eta^{(i)}(t) = E[\delta^{(i)}(X) | t]$, то процедуры δ и η имеют одну и ту же функцию риска.]

К разделу 7

12. Пусть X_i ($i = 1, \dots, s$) независимы и распределены по закону Пуассона $P(\lambda_i)$. Пусть $T_0 = \sum X_j$, $T_i = X_i$, $\lambda = \sum \lambda_j$. Тогда T_0 имеют распределение Пуассона $P(\lambda)$, и условное распределение $T_1, \dots, T_{s-1}$ при $T_0 = t_0$ является полиномиальным распределением (34) с $n = t_0$ и $p_i = \lambda_i/\lambda$. [Используйте прямой подсчет.]

13. *Испытания на продолжительность жизни.* Пусть $X_1, \dots, X_n$ независимы и имеют распределение с показательной плотностью $(2\theta)^{-1}e^{-x/2\theta}$ при $x \geq 0$. Обозначим X_k , расположенные в порядке возрастания, через $Y_1 \leq Y_2 \leq \dots \leq Y_n$. Допустим, что в процессе наблюдений мы обнаруживаем сначала Y_1 , затем Y_2 и т. д. до Y_r . Подобный процесс может встретиться при испытаниях на продолжительность жизни: каждая из величин X может означать продолжительность «жизни» электронной лампы, когда одновременно испытывается n ламп. Другой пример: радиоактивный распад вещества, где n — число атомов, а наблюдения продолжаются до появления r -й α -частицы.

(I) Совместные распределения $Y_1, \dots, Y_r$ образуют экспоненциальное семейство с плотностью

$$\frac{1}{(2\theta)^r} \frac{n!}{(n-r)!} \exp \left[- \frac{\sum_{i=1}^r y_i + (n-r)y_r}{2\theta} \right], \quad \theta \leq y_1 \leq \dots \leq y_r.$$

(II) Величина $\left[\sum_{i=1}^r Y_i + (n-r)Y_r \right] / \theta$ имеет χ^2 -распределение с $2r$ степенями свободы.

(III) Пусть $Y_1, Y_2, \dots$ обозначают промежутки времени до наступления первого, второго и т. д. событий в процессе Пуассона с параметром $1/2\theta'$ (см. задачу 1 в главе 1). Тогда величины $Z_1 = Y_1/\theta'$, $Z_2 = (Y_2 - Y_1)/\theta'$, $Z_3 = (Y_3 - Y_2)/\theta'$ независимы и имеют каждая распределение χ^2 с двумя степенями свободы. Совместные плотности $Y_1, \dots, Y_r$ образуют экспоненциальное семейство с плотностью

$$\frac{1}{(2\theta')^r} \exp \left(- \frac{y_r}{2\theta'} \right), \quad 0 \leq y_1 \leq \dots \leq y_r.$$

Величина Y_r/θ' имеет распределение χ^2 с $2r$ степенями свободы.

(IV) Та же модель пригодна для описания испытаний на продолжительность жизни, если число n ламп поддерживается постоянным, для чего перегоревшие лампы заменяются новыми. Здесь Y_1 обозначает момент времени, когда перегорает первая лампа, Y_2 — момент, когда перегорает вторая лампа и т. д. При этом предполагается, что фиксирован некоторый начальный момент времени, от которого идет отсчет.

*) О более общем результате см. Bahadur, A characterization of sufficiency, Ann. Math. Stat. 26 (1955), 286—293 и Elfving, Sufficiency and completeness, Ann. Acad. Sci. Fennicae, Ser. A, 135, 1952.

[(II) Случайные величины $Z_i = (n - i + 1)(Y_i - Y_{i-1})/\theta$ ($i = 1, \dots, r$) независимы и каждая из них имеет распределение χ^2 с двумя степенями свободы и $\left[\sum_{i=1}^r Y_i + (n - r) Y_r \right] / \theta = \sum_{i=1}^r Z_i$.]

14. Математические ожидания и ковариации статистик T_j в экспоненциальном семействе (35) даются формулами

$$E[T_j(X)] = -\partial \log C(\theta) / \partial \theta_j \quad (j = 1, \dots, k),$$

$$E[T_i(X) T_j(X)] - [ET_i(X) ET_j(X)] = -\frac{\partial^2 \log C(\theta)}{\partial \theta_i \partial \theta_j} \quad (i, j = 1, \dots, k).$$

15. Пусть Ω означает естественное параметрическое пространство экспоненциального семейства (35), и пусть $\Omega'_{\theta_1}, \dots, \theta_r$ — естественное параметрическое пространство для семейства условных распределений при условии $T_{r+1} = t_{r+1}, \dots, T_k = t_k$ ($t_{r+1}, \dots, t_k$ — фиксированные числа).

(I) Тогда $\Omega'_{\theta_1}, \dots, \theta_r$ содержит проекцию $\Omega_{\theta_1}, \dots, \theta_r$ пространства Ω на $\theta_1, \dots, \theta_r$.

(II) Семейство плотностей

$$p_{\theta_1, \theta_2}(x, y) = C(\theta_1, \theta_2) \exp(\theta_1 x + \theta_2 y - xy), \quad x, y > 0,$$

даст пример, в котором $\Omega_{\theta_1}, \dots, \theta_r$ является собственным подмножеством $\Omega'_{\theta_1}, \dots, \theta_r$.

9. Литературные ссылки

Теория меры и интегрирования в абстрактных пространствах излагается во многих книгах, в том числе в книгах Халмоша (1950), Лоева (1955) и Сакса (1937)*).

Определение и основные свойства условных вероятностей и математических ожиданий были даны Колмогоровым (1933). Более детальное изложение, содержащее большое число дополнительных результатов, можно найти в книгах Дуба (1953) и Лоева (1955)**).

Названия этих книг, а также более специальные ссылки на литературу к разделам 3, 6 и 7 даются ниже.

Бахадур (Bahadur R. R.)

(1954) Sufficiency and statistical decision functions, Ann. Math. Stat., vol. 25, 423—462. [Подробное изложение абстрактной теории достаточных статистик, включая теорему факторизации, структурную теорему для минимальных достаточных статистик и обсуждение достаточности для случая последовательных экспериментов.]

(1955) Statistics and subfields, Ann. Math. Stat., vol. 26, 490—497.

*) См. также Данфорд, Шварц, Линейные операторы. Общая теория, пер. с англ., М., ИЛ, 1962; Колмогоров А. Н., Фомин С. В., Элементы теории функций и функционального анализа, М., Наука, 1968. — Прим. ред.

**) См. также Невё, Математические основы теории вероятностей, пер. с франц. М., Мир, 1969; Феллер В., Введение в теорию вероятностей и ее приложения, том 2, пер. с англ., М., Мир, 1967. — Прим. ред.

Бахадур и Леман (Bahadur R. R. and Lehman E. L.)

(1955) Two comments on «sufficiency and statistical decision functions», Ann. Math. Stat., vol. 26, 139—142. [Задача 3]

Дуб (Doob J. L.)

(1953) Stochastic Processes, New York, John Wiley & Sons. (Русский перевод: Дж. Л. Дуб, Вероятностные процессы, ИЛ, 1956.)

Колмогоров

(1933) Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin, J. Springer, или русский вариант: А. Н. Колмогоров, Основные понятия теории вероятностей, ОНТИ, М., 1936; 2-е изд., М., Наука, 1974.

Линник Ю. В.

(1966) Статистические задачи с мешающими параметрами, М., Наука.

Лоэв (Loève M.)

(1955) Probability Theory, New York D. Van Nostrand Co. (Русский перевод второго (1960 г.) издания: М. Лоэв, Теория вероятностей, ИЛ, 1962.)

Сакс (Saks S.)

(1937) Theory of the Integral, New York, G. E. Stechert and Co. (Русский перевод: С. Сакс, Теория интеграла, ИЛ, 1949.)

Халмош (Halmos P. R.)

(1950) Measure theory, New York, D. Van Nostrand Co. (Русский перевод:

П. Халмош, Теория меры, ИЛ, 1953.)

Халмош и Сэвидж (Halmos P. R. and Savage L. J.)

(1949) Application of the Radon — Nikodym theorem to the theory of sufficient statistics, Ann. Math. Stat., vol 20, 225—241. [Первая абстрактная формулировка понятия достаточной статистики; теорема факторизации. Задача 9].

Эпштейн и Собел (Epstein B. and Sobel M.)

(1954) Some theorems relevant to life testing from an exponential distribution, Ann. Math. Stat., vol. 25, 373—381. [Задача 13.]

ГЛАВА 3

РАВНОМЕРНО НАИБОЛЕЕ МОЩНЫЕ КРИТЕРИИ

1. Постановка проблемы

Мы начинаем теперь изучение статистических проблем, теория которых была разработана в наибольшей степени, а именно проблем проверки гипотез. Как видно из самого названия, при этом хотят выяснить, верна или неверна какая-либо данная гипотеза. Выбор здесь производится только между двумя решениями: принятием или отклонением гипотезы. Решающая процедура в таких проблемах называется *критерием* гипотезы, о которой идет речь.

Решение должно быть основано на значениях некоторой случайной величины X , о которой известно, что ее распределение принадлежит данному классу $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$. Мы предположим, что если бы θ было известно, то можно было бы сказать — верна гипотеза или нет. Соответственно, распределения класса $\mathcal{P}$ могут быть разделены на две группы: для одной из групп гипотеза верна, а для другой — нет. Эти группы будут обозначаться буквами H и K , а соответствующие им подмножества Ω — буквами Ω_H и Ω_K , так что $H \cup K = \mathcal{P}$ и $\Omega_H \cup \Omega_K = \Omega$. Математически гипотеза эквивалентна утверждению, что P_θ принадлежит H . Поэтому удобно отождествить гипотезу и это утверждение и обозначать ее также буквой H . Аналогично мы называем распределения из K альтернативными к H , так что K есть *класс альтернатив*.

Обозначим решения, состоящие в принятии или отклонении H , буквами d_0 и d_1 соответственно. Нерандомизированная процедура проверки гипотезы приписывает каждому возможному значению x случайной величины X одно из двух решений. Тем самым выборочное пространство делится на два взаимно дополнительных множества S_0 и S_1 . Если X попадает в S_0 , гипотеза принимается, в других случаях она отвергается. Множество S_0 называется областью принятия гипотезы, а множество S_1 — областью отклонения или *критической областью*.

В процессе проверки можно или прийти к правильному решению, или совершить одну из двух ошибок: отвергнуть гипотезу, когда она верна (ошибка первого рода), или принять гипотезу, когда она неверна (ошибка второго рода). Последствия этих ошибок часто оказываются совершенно различными. Например, если проверяется наличие некоторого заболевания, то неправильное заключение о необходимости лечения может создать пациенту неудобства. С другой стороны, неудача в попытке обнаружить имеющееся заболевание может привести к смерти пациента.

Желательно провести проверку таким образом, чтобы свести к минимуму вероятности обоих типов ошибок. К сожалению, когда число испытаний задано, мы не можем управлять обеими вероятностями ошибок одновременно. Обычно задают границу для вероятности отклонения H , когда она верна, и при этом условии стремятся минимизировать вероятность другой ошибки. Иными словами, выбирают число α между 0 и 1 (называемое *уровнем значимости*) и налагают условие

$$P_{\theta} \{ \delta(X) = d_1 \} = P_{\theta} \{ X \in S_1 \} \leq \alpha \quad \text{для всех } \theta \in \Omega_H. \quad (1)$$

При этом условии желательно сделать минимальной $P_{\theta} \{ \delta(X) = d_0 \}$ для $\theta \in \Omega_K$ или, что то же самое, сделать максимальной

$$P_{\theta} \{ \delta(X) = d_1 \} = P_{\theta} \{ X \in S_1 \} \quad \text{для всех } \theta \in \Omega_K. \quad (2)$$

Хотя обычно из (1) вытекает, что

$$\sup_{\Omega_H} P_{\theta} \{ X \in S_1 \} = \alpha, \quad (3)$$

удобно для левой части (3) ввести специальное название: ее называют *размером* критерия или критической области.

Принимая условие (1), мы ограничиваемся, следовательно, критериями, размер которых не превосходит заданного уровня значимости. Вероятность (2), рассматриваемая как функция θ , $\theta \in \Omega_K$, называется *мощностью* критерия при альтернативе θ . Рассматриваемая как функция θ , $\theta \in \Omega$, вероятность (2) называется *функцией мощности* критерия и обозначается $\beta(\theta)$.

Обычно выбор уровня значимости α до некоторой степени произволен, поскольку в большинстве ситуаций нет точной границы для «разрешенной» вероятности ошибки первого рода. Стало обычным выбирать для α одно из стандартных значений, таких как 0,005, 0,01, 0,05. Эта стандартизация имеет некоторые преимущества, так как она позволяет сократить объем таблиц, используемых при проведении различных испытаний. Никакой другой специальной причины для выбора именно этих значений нет. Действительно, выбирая уровень значимости, необходимо

также обращать внимание на мощность критерия при альтернативных гипотезах. Если мощность оказывается слишком малой, то может оказаться желательным использование значений α , превосходящих обычные, например 0,1 или 0,2 *).

Другое обстоятельство, которое часто влияет на выбор уровня значимости — это наше отношение к гипотезе до проведения эксперимента. Если мы твердо верим в истинность гипотезы, то потребуются убедительные свидетельства против нее для того, чтобы мы отказались от своей уверенности; соответственно уровень значимости будет выбран весьма низким (низкий уровень значимости приводит к тому, что гипотеза отвергается при таких комбинациях результатов наблюдений, полная вероятность которых при нашей гипотезе мала, так что появление этих результатов крайне неправдоподобно при справедливости H).

В приложениях мы обычно сталкиваемся с семействами вложенных друг в друга критических областей. В этих случаях полезно определять не только то, что гипотеза принимается или отвергается с данным уровнем значимости, но и указывать для каждого x наименьший уровень значимости $\hat{\alpha} = \hat{\alpha}(x)$, *критический уровень*, при котором гипотеза отвергается для данного результата наблюдения. Это число дает представление о том, сколь сильно данные наблюдений противоречат гипотезе (или ее подкрепляют), и делают возможным для других прийти к заключению, соответствующему выбранному ими уровню значимости (см. задачу 7 и в главе 4 задачу 2).

Рассмотрим теперь структуру рандомизированного критерия. При любом x такой критерий приводит к выбору между двумя решениями — принятием или отклонением гипотезы, и этот выбор осуществляется с зависящими от x вероятностями, которые будут обозначаться $\varphi(x)$ и $1 - \varphi(x)$ соответственно. Если X принимает значение x , то производится случайный эксперимент с двумя возможными исходами: R и $\bar{R}$, вероятности которых равны $\varphi(x)$ и $1 - \varphi(x)$. Если эксперимент заканчивается исходом R , то гипотеза отвергается, а в противном случае принимается. Рандомизированный критерий, таким образом, полностью характеризуется функцией φ — *критической функцией*, $0 \leq \varphi(x) \leq 1$ при всех x . Если φ принимает только значения 0 и 1, то мы возвращаемся к случаю нерандомизированного критерия. Множество точек, в которых $\varphi(x) = 1$, совпадает в этом случае с критической областью, так что для нерандомизированного критерия φ является индикатором (характеристической функцией) критической области.

*) Практически полезный метод выбора α в зависимости от мощности критерия предложен в статье: Lehmann, Significance level and power, Ann. Math. Stat. 29 (1958), 1167—1176.

Если X распределена по закону P_θ и используется критическая функция φ , то вероятность отвергнуть гипотезу равна

$$E_\theta \varphi(X) = \int \varphi(x) dP_\theta(x),$$

т. е. равна условной вероятности отклонения гипотезы при данном x , проинтегрированной по распределению X . Задача состоит в выборе такого φ , чтобы сделать мощность

$$\beta_\varphi(\theta) = E_\theta \varphi(X) \quad \text{для всех } \theta \in \Omega_K \quad (4)$$

максимальной при условии

$$E_\theta \varphi(X) \leq \alpha \quad \text{для всех } \theta \in \Omega_H. \quad (5)$$

Но здесь возникает та же самая трудность, с которой мы уже встречались в общих рассуждениях главы 1. Как правило, критерий, максимизирующий мощность при какой-либо определенной альтернативе класса K , зависит от этой альтернативы. Поэтому необходимы дополнительные соображения о том, что следует понимать под оптимальной процедурой. Имеется одно важное исключение: если класс K содержит ровно одно распределение, т. е. мы имеем дело с одной альтернативной гипотезой, то проблема целиком определяется соотношениями (4) и (5), и с математической точки зрения сводится к максимизации некоторого интеграла при некоторых определенных ограничениях. Теория для этого случая и ее статистические приложения составляют основное содержание настоящей главы. В некоторых случаях может, конечно, оказаться, что один и тот же критерий максимизирует мощность для всех альтернатив из K (даже если их много). Примеры таких *равномерно наиболее мощных* (РНМ) критериев будут указаны в разделах 3 и 7.

В данной выше формулировке задачу можно рассматривать, как специальный случай общей проблемы решения с двумя типами потерь. Соответственно двум типам ошибок мы вводим две компоненты функции потерь

$$\begin{cases} L_1(\theta, d_1) = 1 \quad \text{или} \quad 0 \quad \text{при} \quad \theta \in \Omega_H \quad \text{или} \quad \theta \in \Omega_K \\ L_1(\theta, d_0) = 0 \quad \text{для всех } \theta \end{cases}$$

и

$$\begin{cases} L_2(\theta, d_0) = 0 \quad \text{или} \quad 1 \quad \text{при} \quad \theta \in \Omega_H \quad \text{или} \quad \theta \in \Omega_K \\ L_2(\theta, d_1) = 0 \quad \text{для всех } \theta. \end{cases}$$

При таком определении проблема проверки гипотезы, как она сформулирована выше, равносильна минимизации $EL_2(\theta, \delta(x))$ при ограничении $EL_1(\theta, \delta(X)) \leq \alpha$.

Формально введенные функции потерь L_1 и L_2 , вообще говоря, не будут совпадать с истинными потерями. Потеря, воз-

никающая при ошибочном принятии гипотезы, может не быть одной и той же для всех альтернатив: чем больше отличается альтернатива от основной гипотезы, тем более серьезны последствия такой ошибки. Как было указано ранее, мы намеренно уклонились от большей детализации, требуемой этой критикой. Предпочтительнее основывать теорию на простых и интуитивно привлекательных понятиях ошибок первого и второго рода. Впоследствии мы увидим, что по крайней мере некоторые из результатов останутся верными и в более сложной обстановке.

2. Фундаментальная лемма Неймана — Пирсона

Класс распределений будет называться *простым*, если он содержит ровно одно распределение. В противном случае он будет называться *сложным*. Проблема проверки гипотезы полностью описывается соотношениями (4) и (5), если класс K — простой. Если класс H также простой, то проблема может быть решена точно. Пусть гипотезе H соответствует распределение P_0 , а гипотезе K — распределение P_1 . Допустим на время, что эти распределения дискретны, и обозначим $P_i\{X=x\} = P_i(x)$ для $i = 0, 1$. Если ограничиться нерандомизированными критериями, то оптимальный критерий соответствует критической области S , для которой

$$\sum_{x \in S} P_0(x) \leq \alpha \quad (6)$$

и

$$\sum_{x \in S} P_1(x) = \max.$$

Нетрудно понять, какие именно точки должны быть включены в S . Выбранные точки должны иметь суммарную вероятность, не превосходящую α , в одном случае, и вероятность возможно большую — в другом. Подобные задачи часто возникают и в иной обстановке. Так, покупатель с ограниченными средствами, желающий приобрести как можно больше необходимых ему предметов, будет руководствоваться их *количеством на доллар*. Лицо, желающее как можно быстрее преодолеть некоторое расстояние, должно выбирать скорейший способ передвижения, т. е. такой, при котором число *миль в час* оказывается максимальным. Аналогично и в нашей проблеме, наиболее ценными являются точки x с большими значениями:

$$r(x) = \frac{P_1(x)}{P_0(x)}.$$

Точки, следовательно, упорядочиваются соответственно величине этого отношения и затем в S включают максимальное число их, согласующееся с требованием (6). Формально это означает, что

S есть множество тех точек x , для которых $r(x) > c$, где c определяется условием

$$P_0\{X \in S\} = \sum_{x: r(x) < c} P_0(x) = \alpha.$$

Здесь возникает некоторая трудность. Может случиться, что, включив очередную точку, мы еще не достигнем уровня α , а включив следующую — превзойдем его. Точное значение α при этом или совсем не может быть получено, или же достигается присоединением к S какой-либо из следующих по порядку точек. Эта трудность исчезает при переходе к рандомизированным критериям. С помощью рандомизации можно «расщепить» очередную точку, взяв в S такую ее «часть», чтобы получить суммарную вероятность, в точности равную α , не нарушая принятого порядка точек. Эти рассуждения формализуются в нижеследующей теореме (фундаментальной лемме Неймана — Пирсона).

Теорема 1. Пусть P_0 и P_1 — распределения вероятностей, обладающие плотностями p_0 и p_1 соответственно, по отношению к некоторой мере μ *).

(I) **Существование.** Для проверки H : p_0 при конкурирующей гипотезе K : p_1 найдется критерий φ и константа k такие, что

$$E_0\varphi(X) = \alpha \quad (7)$$

и

$$\varphi(x) = \begin{cases} 1, & \text{когда } p_1(x) > kp_0(x), \\ 0, & \text{когда } p_1(x) < kp_0(x). \end{cases} \quad (8)$$

(II) **Достаточное условие для критерия наибольшей мощности.** Если критерий удовлетворяет требованиям (7) и (8) при некотором k , то он является наиболее мощным критерием уровня α для проверки распределения p_0 при конкурирующем p_1 .

(III) **Необходимое условие для критерия наибольшей мощности.** Если φ — наиболее мощный критерий уровня α для проверки распределения p_0 при конкурирующем p_1 , то при некотором k он удовлетворяет (8) почти всюду по мере μ . Он также удовлетворяет (7), кроме случая, когда существует критерий размера $< \alpha$ и мощности 1.

Доказательство. Для $\alpha = 0$ и $\alpha = 1$ теорема очевидным образом верна: достаточно принять в (8) $k = +\infty$, полагая произведение $0 \cdot \infty$ равным 0. Поэтому мы предположим в дальнейшем, что $0 < \alpha < 1$.

*) Это предложение не ограничивает общности, так как всегда можно взять

$$\mu = P_0 + P_1.$$

(I) Обозначим $\alpha(c) = P_0\{p_1(X)\} > cp_0(X)\}$. При вычислении P_0 -вероятностей достаточно рассматривать множества, где $p_0(x) > 0$. Поэтому $\alpha(c)$ равно вероятности того, что случайная величина $p_1(X)/p_0(X)$ превосходит c . Таким образом, $1 - \alpha(c)$ является функцией распределения и $\alpha(c)$ не возрастает и непрерывна справа,

$$\alpha(c - 0) - \alpha(c) = P_0\{p_1(X)/p_0(X) = c\}, \quad \alpha(-\infty) = 1, \quad \alpha(\infty) = 0.$$

При данном α , $0 < \alpha < 1$, определим c_0 из соотношений $\alpha(c_0) \leq \alpha \leq \alpha(c_0 - 0)$ и введем критерий φ , полагая

$$\varphi(x) = \begin{cases} 1, & \text{когда } p_1(x) > c_0 p_0(x), \\ \frac{\alpha - \alpha(c_0)}{\alpha(c_0 - 0) - \alpha(c_0)}, & \text{когда } p_1(x) = c_0 p_0(x), \\ 0, & \text{когда } p_1(x) < c_0 p_0(x). \end{cases}$$

Здесь второе равенство имеет смысл всегда, кроме случая $\alpha(c_0) = \alpha(c_0 - 0)$, но в этом последнем случае $P_0\{p_1(X) = c_0 p_0(X)\} = 0$, следовательно φ определена почти всюду. Размер φ равен

$$E_0 \varphi(X) = P_0\left\{\frac{p_1(X)}{p_0(X)} > c_0\right\} + \frac{\alpha - \alpha(c_0)}{\alpha(c_0 - 0) - \alpha(c_0)} P_0\left\{\frac{p_1(X)}{p_0(X)} = c_0\right\} = \alpha.$$

Таким образом, величина c_0 , участвующая в формулировке теоремы, может быть выбрана равной k .

Интересно отметить, что величина c_0 определяется, по существу, единственным способом. Единственное исключение — это случай, когда $\alpha(c) = \alpha$ для целого интервала значений c . Если (c', c'') — такой интервал, причем

$$C = \left\{x: p_0(x) > 0 \text{ и } c' < \frac{p_1(x)}{p_0(x)} < c''\right\},$$

то

$$P_0(C) = \alpha(c') - \alpha(c'' - 0) = 0 \text{ и } \mu(C) = 0, \quad P_1(C) = 0.$$

То есть множества, соответствующие различным значениям c , отличаются друг от друга точками, в совокупности имеющими нулевую вероятность при обоих распределениях, и потому вообще могут быть исключены из выборочного пространства.

(II) Предположим, что критерий φ удовлетворяет (7) и (8). Пусть φ^* — какой-либо другой критерий с $E_0 \varphi^*(X) \leq \alpha$. Обозначим S^+ и S^- множества точек выборочного пространства, в которых $\varphi(x) - \varphi^*(x) > 0$ и $\varphi(x) - \varphi^*(x) < 0$ соответственно. Если $x \in S^+$, то $\varphi(x) > 0$ и $p_1(x) > kp_0(x)$. Аналогично $p_1(x) \leq kp_0(x)$ для всех $x \in S^-$. Следовательно,

$$\int (\varphi - \varphi^*) (p_1 - kp_0) d\mu = \int_{S^+ \cup S^-} (\varphi - \varphi^*) (p_1 - kp_0) d\mu \geq 0,$$

и для разности мощностей φ и φ^* получаем

$$\int (\varphi - \varphi^*) p_1 d\mu \geq k \int (\varphi - \varphi^*) p_0 d\mu \geq 0,$$

что и требовалось доказать.

(III) Предположим, что φ^* является наиболее мощным критерием уровня α для проверки гипотезы p_0 при альтернативе p_1 , и φ удовлетворяет условиям (7) и (8). Пусть S обозначает пересечение множества $S^+ \cup S^-$, на котором $\varphi \neq \varphi^*$, и множества $\{x: p_1(x) \neq kp_0(x)\}$. Допустим, что $\mu(S) > 0$. Так как произведение $(\varphi - \varphi^*)(p_1 - kp_0)$ положительно на S , то

$$\int_{S^+ \cup S^-} (\varphi - \varphi^*)(p_1 - kp_0) d\mu = \int_S (\varphi - \varphi^*)(p_1 - kp_0) d\mu > 0,$$

т. е. φ оказывается при альтернативе p_1 более мощным, чем φ^* . Мы приходим к противоречию. Следовательно, $\mu(S) = 0$, что и требовалось доказать.

Если бы φ^* имел размер $< \alpha$ и мощность < 1 , то было бы возможно включить в критическую область дополнительные точки (или «части» точек). Этим мы увеличили бы или мощность до 1, или размер критерия до α . Таким образом, или $E_{0\varphi^*}(X) = \alpha$, или $E_{1\varphi^*}(X) = 1$.

Доказательство части (III) показывает, что наиболее мощный критерий определяется соотношениями (7) и (8) однозначно всюду, кроме, быть может, точек, для которых $p_1(x) = kp_0(x)$. На этом множестве φ можно определить произвольным образом, лишь бы размер критерия был равен α . Действительно, мы показали, что на этом пограничном множестве φ можно выбрать постоянным. В тривиальном случае, когда существует критерий мощности 1, константа k в (8) равна 0, и мы будем принимать H во всех точках, где $p_1(x) = kp_0(x)$, хотя, быть может, размер критерия окажется меньше α .

Из этих замечаний вытекает, что наиболее мощный критерий определяется однозначно из соотношений (7) и (8) (с точностью до множеств нулевой меры) каждый раз, когда множество $p_1(x) = kp_0(x)$ имеет μ -меру, равную 0. Очевидно, что этот единственный критерий будет нерандомизированным. Вообще, мы видели, что рандомизация может потребоваться только на пограничном множестве и только в случае необходимости сделать размер критерия в точности равным α . На практике предпочитают в таком случае заменить уровень значимости таким, который не требует рандомизации. В случае, когда существует критерий мощности 1, соотношения (7) и (8) по-прежнему определяют наиболее мощный критерий, но он может не быть единственным, поскольку может существовать другой наиболее мощ-

ный критерий, удовлетворяющий (7) и (8) при некотором $\alpha' < \alpha$.

Следствие 1. Пусть β обозначает мощность наиболее мощного критерия уровня α для проверки гипотезы P_0 при конкурирующей P_1 . Тогда $\alpha < \beta$, за исключением случая $P_0 = P_1$.

Доказательство. Так как критерий уровня α , определяемый формулой $\varphi(x) \equiv \alpha$, имеет мощность α , то ясно, что $\alpha \leq \beta$. Если $\alpha = \beta < 1$, то критерий $\varphi(x) \equiv \alpha$ оказывается наиболее мощным, и по теореме 1 (III) он должен удовлетворять (8).

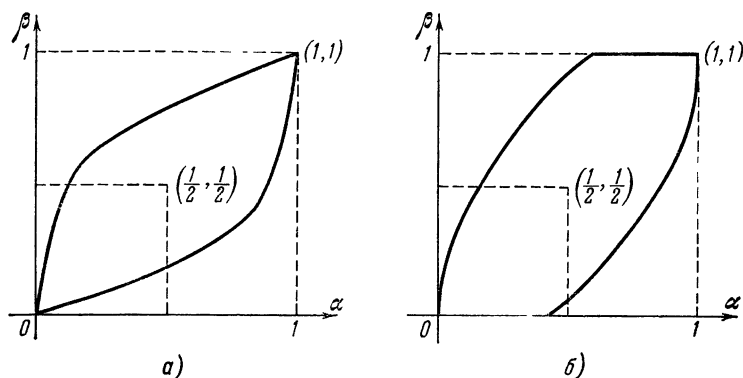


Рис. 3.

Тогда $p_0(x) = p_1(x)$ почти всюду по мере μ , следовательно $P_0 = P_1$.

Альтернативный метод доказательства теоремы настоящего раздела основан на следующем геометрическом представлении проблемы проверки простой гипотезы при простой альтернативе. Пусть N обозначает множество всех точек (α, β) , обладающих свойством: существует критерий φ , для которого

$$\alpha = E_0\varphi(X), \quad \beta = E_1\varphi(X).$$

Это множество выпукло, содержит точки $(0, 0)$ и $(1, 1)$ и центрально симметрично по отношению к точке $(\frac{1}{2}, \frac{1}{2})$ (т. е. вместе с любой точкой (α, β) это множество содержит также точку $(1 - \alpha, 1 - \beta)$) (рис. 3, а). Кроме того, множество замкнуто. [Это следует из теоремы о слабой компактности множества критических функций, см. теорему 3 в добавлении; аргументация совпадает с используемой в доказательстве теоремы 5(I).]

При каждом $0 < \alpha_0 < 1$ критерии уровня α_0 изображаются точками, абсциссы которых $\leq \alpha_0$. Наиболее мощный из этих критериев (а он существует, так как N замкнуто) соответствует точке на верхней границе N , имеющей абсциссу α_0 . Эта точка будет

единственной, соответствующей наиболее мощному критерию уровня α_0 , за исключением случая, когда в N существует точка $(\alpha, 1)$ с $\alpha < \alpha_0$ (см. рис. 3, б).

Рассмотрим в качестве примера геометрическое доказательство следствия 1. Предположим, что для некоторого $0 < \alpha_0 < 1$ мощность наиболее мощного критерия уровня α_0 равна α_0 . Тогда из выпуклости N следует, что включение $(\alpha, \beta) \in N$ влечет $\beta \leq \alpha$, а из симметрии N вытекает, что N совпадает с отрезком, соединяющим точки $(0, 0)$ и $(1, 1)$. Это означает, что $\int \varphi p_0 d\mu = \int \varphi p_1 d\mu$ для всех φ и потому $p_0 = p_1$ μ -почти всюду, что и требовалось доказать.

Опирающееся на эти идеи доказательство теоремы 1 содержится в более общем доказательстве теоремы 5.

3. Распределения с монотонным отношением правдоподобия

Случай, когда и проверяемая гипотеза, и класс альтернатив являются простыми, представляет главным образом теоретический интерес, так как в прикладных задачах встречаются, как правило, параметрические семейства распределений, зависящие от одного или нескольких непрерывных параметров. В простейшей из подобных ситуаций распределения зависят от одного действительного параметра и проверяемая гипотеза является односторонней, скажем $H: \theta \leq \theta_0$. Наиболее мощный критерий проверки H при альтернативе $\theta_1 > \theta_0$ зависит, вообще говоря, от θ_1 и потому не будет РНМ. Однако при некоторых дополнительных требованиях РНМ критерий существует. Говорят, что зависящее от действительного параметра семейство плотностей имеет *монотонное отношение правдоподобия*, если для любых θ и θ' распределения P_θ и $P_{\theta'}$ различны и существует функция $T(x)$ такая, что отношение $p_{\theta'}(x)/p_\theta(x)$ является неубывающей функцией от $T(x)$.

Теорема 2. Пусть θ — действительный параметр и случайная величина X имеет плотность вероятности $p_\theta(x)$ с монотонным отношением $T(x)$ отношением правдоподобия. Тогда

(I) Для проверки $H: \theta \leq \theta_0$ при конкурирующей гипотезе $K: \theta > \theta_0$ существует РНМ критерий

$$\varphi(x) = \begin{cases} 1, & \text{когда } T(x) > C, \\ \gamma, & \text{когда } T(x) = C, \\ 0, & \text{когда } T(x) < C, \end{cases} \quad (9)$$

где C и γ выбираются так, чтобы выполнялось равенство

$$E_{\theta_0} \varphi(x) = \alpha. \quad (10)$$

(II) *Функция мощности*

$$\beta(\theta) = E_{\theta}\varphi(X)$$

этого критерия строго возрастает во всех точках, где $\beta(\theta) < 1$.

(III) Для всех θ' критерий, определенный равенствами (9) и (10), является РНМ для проверки $H': \theta \leq \theta'$ при конкурирующей $K': \theta > \theta'$ и при уровне $\alpha' = \beta(\theta')$.

(IV) Для любого $\theta < \theta_0$ наш критерий минимизирует $\beta(\theta)$ (вероятность ошибки первого рода) в классе всех критериев, удовлетворяющих (10).

Доказательство. (I) и (II). Рассмотрим сначала гипотезу $H: \theta = \theta_0$ и некоторую простую альтернативу $\theta_1 > \theta_0$. Применяя фундаментальную лемму, мы находим, что наиболее мощный критерий отклоняет гипотезу, когда

$$\frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} > C,$$

или, эквивалентным образом, когда

$$T(x) > C^*).$$

Из теоремы 1 (I) следует, что существуют такие постоянные C и γ , что равенства (9) и (10) выполняются. Если $\theta' < \theta''$, то по теореме 1 (II) соответствующий критерий является наиболее мощным для проверки $P_{\theta'}$ при конкурирующей $P_{\theta''}$ и при уровне $\alpha' = \beta(\theta')$. Утверждение (II) теоремы вытекает теперь из следствия 1. Так как $\beta(\theta)$ не убывает, то критерий удовлетворяет неравенствам

$$E_{\theta}\varphi(X) \leq \alpha \quad \text{для} \quad \theta \leq \theta_0. \quad (11)$$

Класс критериев, для которых верно (11), содержится в классе критериев с $E_{\theta_0}\varphi(X) \leq \alpha$. Так как данный критерий максимизирует $\beta(\theta_1)$ в этом более широком классе, он тем более максимизирует $\beta(\theta_1)$ при условии (11). Так как он не зависит от специального выбора альтернативы $\theta_1 > \theta_0$, то он является РНМ при альтернативе K .

(III) доказывается аналогичными рассуждениями.

(IV) следует из того факта, что критерий, минимизирующий мощность в задаче проверки простой гипотезы при простой альтернативе, получается из фундаментальной леммы (теорема 1) если в ней все неравенства заменить на обратные.

Очевидными изменениями проведенных рассуждений мы получим решение двойственной проблемы $H: \theta \geq \theta_0$, $K: \theta < \theta_0$.

*) Здесь и в аналогичных рассуждениях C обозначает неопределенную постоянную. (C можно понимать как сокращение для const. — Прим. пер.)

Ниже будут даны некоторые примеры семейств с монотонным отношением правдоподобия и соответствующих РНМ критериев. Однако главные применения теорема 2 получит позже, когда в качестве подобных семейств будут выступать или множества условных распределений, взятых по отношению к некоторой достаточной статистике (главы 4 и 5), или распределения максимального инварианта (главы 6 и 7).

Пример 1. Из партии, содержащей N деталей, наудачу берут n деталей и каждую из них проверяют. Если полное число дефектных изделий в партии равно D , то число X дефектных изделий в выборке имеет *гипергеометрическое* распределение

$$P\{X=x\} = P_D(x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}, \quad x=0, 1, \dots, D.$$

Интерпретируя $P_D(x)$ как плотность по отношению к мере μ , которая для каждого множества на прямой равна числу содержащихся в нем точек с целыми неотрицательными координатами, и замечая, что

$$\frac{P_{D+1}(x)}{P_D(x)} = \frac{D+1}{N-D} \frac{N-D-n+x}{D+1-x},$$

мы заключаем, что рассматриваемые распределения имеют монотонное отношение правдоподобия с $T(x) = x$. Следовательно, существует РНМ критерий проверки гипотезы $H: D \leq D_0$ при альтернативе $K: D > D_0$. Этот критерий отвергает H , когда X чрезмерно велико. Аналогичный критерий можно построить для проверки $H': D \geq D_0$.

Важный класс семейств распределений, для которых выполняются предпосылки теоремы 2, образует однопараметрические экспоненциальные семейства.

Следствие 2. Пусть θ — действительный параметр и пусть X имеет распределение с плотностью (по отношению к некоторой мере μ)

$$p_\theta(x) = C(\theta) e^{Q(\theta) T(x)} h(x), \quad (12)$$

где Q — строго монотонная функция. Тогда существует РНМ критерий для проверки $H: \theta \leq \theta_0$ при альтернативе $K: \theta > \theta_0$. Если Q возрастает, то

$$\varphi(x) = \begin{cases} 1 & \text{при } T(x) > C, \\ \gamma & \text{при } T(x) = C, \\ 0 & \text{при } T(x) < C, \end{cases}$$

где C и γ определяются равенством $E_{\theta_0} \varphi(X) = \alpha$. Если Q убывает, то эти неравенства заменяются противоположными.

Если X имеет дискретное распределение, то правая часть (12) будет обозначаться $P_\theta(x)$. Как и в примере 1, $p_\theta(x)$ может рассматриваться как плотность по отношению к мере μ , которая

для любого множества на прямой равна числу возможных значений X , входящих в него.

Пример 2. Биномиальное распределение $b(p, n)$ с

$$P_p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

удовлетворяет (12) с $T(x) = x$, $\theta = p$, $Q(p) = \log[p/(1-p)]$. Проблема проверки гипотезы $H: p \geq p_0$ возникает, например, в обстановке примера 1, если предположить, что процесс производства подвергается статистическому контролю и при этом результаты обследования выбранных изделий независимы с вероятностью встречи дефектного изделия, равной p . Число X дефектных изделий в выборке объема n является в этом случае достаточной статистикой для совместного распределения случайных величин X_i ($i = 1, \dots, n$), где X_i обозначает число дефектных изделий при i -м обследовании (т.е. 0 или 1). Очевидно, X имеет биномиальное распределение $b(p, n)$. Следовательно, существует РНМ критерий проверки H , который отвергает эту гипотезу, если X слишком мало.

Другой план выбора, который иногда применяется в «биномиальной» ситуации, — это *обратный биномиальный выбор*. Здесь испытания заканчиваются только по достижении определенного числа m успехов (например, выздоровлений при использовании нового способа лечения). Пусть Y_i обозначает число испытаний между $(i-1)$ -м и i -м успехами. Тогда вероятность того, что $Y_i = y$, равна pq^y , $y = 0, 1, \dots$. Поэтому совместное распределение величин $Y_1, \dots, Y_m$ равно

$$P_p(y_1, \dots, y_m) = p^m q^{\sum y_i}, \quad y_k = 0, 1, \dots; \quad k = 1, \dots, m.$$

Это семейство оказывается экспоненциальным с $T(y) = \sum y_i$ и $Q(p) = \log(1-p)$. Так как $Q(p)$ — убывающая функция p , то РНМ критерий проверки гипотезы $H: p \leq p_0$ отвергает ее, когда T слишком мало. Этот вывод совпадает с тем, чего естественно было бы ожидать, поскольку наступление m успехов при числе испытаний, лишь немногим превосходящем m , указывает на большое значение p . Статистика T , равная разности между числом испытаний до m -го успеха включительно и числом m , имеет отрицательно-биномиальное распределение (глава 1, задача 1 (I)):

$$P(t) = \binom{m+t-1}{m-1} p^m q^t, \quad t = 0, 1, \dots$$

Пример 3. Если $X_1, \dots, X_n$ независимы и распределены по закону Пуассона с $E(X_i) = \lambda$, то их совместное распределение равно

$$P_\lambda(x_1, \dots, x_n) = \frac{\lambda^{x_1 + \dots + x_n}}{x_1! \dots x_n!} e^{-n\lambda}.$$

Мы снова получаем экспоненциальное семейство с $T(x) = \sum x_i$ и $Q(\lambda) = \log \lambda$. Односторонние гипотезы о значении λ могут возникать, например, если λ обозначает плотность бактерий, а X_k — количество бактерий в пробе k ; или если X_k обозначают число α -частиц, выделяемых в равные и непрерывающиеся промежутки времени радиоактивным источником, и т.д. РНМ критерий для гипотезы $\lambda \leq \lambda_0$ отвергает ее, если $\sum X_i$ чрезмерно велика. Здесь статистика $\sum X_i$, на которой основан критерий, сама имеет распределение Пуассона с параметром $n\lambda$.

Вместо того чтобы наблюдать радиоактивное вещество заданное время или подсчитывать число бактерий на определенном участке стеклянной пластинки, можно использовать также метод «обратной выборки»^{*)}. В этом

^{*)} В оригинале «inverse sampling». — *Прим. пер.*

случае эксперимент продолжается (или интервал наблюдения расширяется) до появления m -го из интересующих нас объектов. Результатом наблюдения является система чисел $T_1, \dots, T_m$, равных соответственно времени до первого появления, времени от первого появления до второго и т. д. Если наблюдаемый процесс — процесс Пуассона и число появлений во временном или пространственном промежутке длины τ имеет распределение

$$P(x) = \frac{(\lambda \tau)^x}{x!} e^{-\lambda \tau}, \quad x = 0, 1, \dots,$$

то величины $T_1, \dots, T_m$ независимы и каждая имеет показательное распределение с плотностью $\lambda e^{-\lambda t}$ для $t \geq 0$ (см. задачу 1 (II) в главе 1). Совместные плотности

$$p_\lambda(t_1, \dots, t_m) = \lambda^m \exp\left(-\lambda \sum_{i=1}^m t_i\right), \quad t_1, \dots, t_m \geq 0,$$

образуют экспоненциальное семейство с $T(t_1, \dots, t_m) = \sum t_i$ и $Q(\lambda) = -\lambda$. РНМ критерий проверки $H: \lambda \leq \lambda_0$ отвергает ее, когда $T = \sum T_i$ слишком мало. Так как $2\lambda T_i$ имеет плотность $1/2 e^{-u/2}$ при $u \geq 0$, т. е. такую же, как χ^2 -распределение с двумя степенями свободы, то $2\lambda T$ имеет χ^2 -распределение с $2m$ степенями свободы. Поэтому границу критической области можно определить по таблицам χ^2 -распределения.

В формулировке проблемы проверки гипотез, данной в начале этой главы, потери, возникающие от принятия неправильного решения, выражались в терминах ошибок первого и второго рода. Чтобы прийти к более детальному описанию задачи проверки $H: \theta \leq \theta_0$ при альтернативе $\theta > \theta_0$, можно поставить ее как проблему с решениями d_0 (принятие H) и d_1 (отклонение H) и функцией потерь $L(\theta, d_i) = L_i(\theta)$. Обычно $L_0(\theta) = 0$ при $\theta \leq \theta_0$ и строго возрастает при $\theta \geq \theta_0$; $L_1(\theta) = 0$ при $\theta \geq \theta_0$ и строго убывает при $\theta \leq \theta_0$. Разность L_1 и L_0 удовлетворяет неравенству

$$L_1(\theta) - L_0(\theta) \geq 0, \quad \text{если } \theta \leq \theta_0. \quad (13)$$

Теорема 3. (I) В предположениях теоремы 2 семейство критериев, определяемых формулами (9) и (10) с $0 \leq \alpha \leq 1$, является существенно полным, если допустить, что функции потерь удовлетворяют (13).

(II) Это семейство является минимальным существенно полным, если множество точек x , для которых $p_\theta(x) > 0$, не зависит от θ .

Доказательство. (I) Для любого критерия φ функция риска равна

$$\begin{aligned} R(\theta, \varphi) &= \int p_\theta(x) \{\varphi(x) L_1(\theta) + [1 - \varphi(x)] L_0(\theta)\} d\mu(x) = \\ &= \int p_\theta(x) \{L_0(\theta) + [L_1(\theta) - L_0(\theta)] \varphi(x)\} d\mu(x) \end{aligned}$$

и, следовательно, разность двух функций риска равна

$$R(\theta, \varphi') - R(\theta, \varphi) = [L_1(\theta) - L_0(\theta)] \int (\varphi' - \varphi) p_\theta d\mu.$$

Это выражение ≤ 0 для всех θ , если

$$\beta_{\varphi'}(\theta) - \beta_\varphi(\theta) = \int (\varphi' - \varphi) p_\theta d\mu \equiv 0 \quad \text{при} \quad \theta \equiv \theta_0.$$

Для заданного критерия φ положим $E_{\theta_0} \varphi(X) = \alpha$. Из теоремы 2 (I) следует, что существует удовлетворяющий (9) и (10) РНМ критерий φ' уровня α для проверки гипотезы $\theta = \theta_0$ при альтернативе $\theta > \theta_0$. По теореме 2 (IV) φ' также минимизирует мощность при $\theta < \theta_0$. Поэтому $R(\theta, \varphi') \leq R(\theta, \varphi)$ при всех θ , что и требовалось доказать.

(II) Пусть φ_α и $\varphi_{\alpha'}$ имеют размеры α и α' соответственно и, кроме того, являются РНМ для проверки гипотезы $\theta = \theta_0$ при альтернативе $\theta > \theta_0$. Тогда $\beta_{\varphi_\alpha}(\theta) < \beta_{\varphi_{\alpha'}}(\theta)$ при всех $\theta > \theta_0$, исключая случай $\beta_{\varphi_\alpha}(\theta) = 1$. Рассматривая задачу проверки той же гипотезы $\theta = \theta_0$ при альтернативе $\theta < \theta_0$, заключаем, что это неравенство выполняется и при всех $\theta < \theta_0$, исключая случай $\beta_{\varphi_{\alpha'}}(\theta) = 0$. Так как, в силу наших предположений, исключительные случаи невозможны, то $R(\theta, \varphi') \leq R(\theta, \varphi)$ при $\theta \equiv \theta_0$. Следовательно, каждая из двух функций риска лучше другой для подходящих значений θ .

Теперь мы видим, что полученный ранее класс РНМ критериев, соответствующих различным величинам уровня значимости, является существенно полным классом в значительно более общей проблеме решения, в которой функции потерь должны удовлетворять некоторым широким качественным условиям. С этой точки зрения фиксирование уровня значимости можно рассматривать как простой способ выбора какой-либо отдельной процедуры из существенно полного класса.

Свойство монотонности отношения правдоподобия порождает весьма сильное упорядочение семейства распределений. В дальнейшем мы будем использовать также следующее несколько более слабое определение. Семейство одномерных функций распределения F_θ будет называться *стохастически возрастающим* (этот же термин будет применяться и к случайным величинам с такими функциями распределения), если эти распределения различны и из $\theta < \theta'$ вытекает $F_\theta(x) \geq F_{\theta'}(x)$ для всех x . Тогда, если X и X' имеют распределения F_θ и $F_{\theta'}$, соответственно, то при всех x $P\{X > x\} \leq P\{X' > x\}$, так что X' имеет тенденцию принимать большие значения, чем X . В этом случае мы скажем, что случайная величина X' *стохастически больше* X . Это отношение

между величинами поясняется следующей леммой, характеризующей стохастическое упорядочение двух распределений.

Лемма 1. Пусть F_0 и F_1 — две функции распределения на прямой. Неравенство $F_1(x) \leq F_0(x)$ выполняется при всех x в том и только том случае, когда существуют две неубывающие функции f_0 и f_1 и случайная величина V такие, что (а) $f_0(v) \leq f_1(v)$ при всех v и (б) распределения $f_0(V)$ и $f_1(V)$ совпадают с F_0 и F_1 соответственно.

Доказательство. Допустим сначала, что требуемые f_0 , f_1 и V существуют. Тогда

$$F_1(x) = P\{f_1(V) \leq x\} \leq P\{f_0(V) \leq x\} = F_0(x)$$

при всех x . Обратно, пусть $F_1(x) \leq F_0(x)$ при всех x и пусть $f_i(y) = \inf\{x : F_i(x - 0) \leq y \leq F_i(x)\}$, $i = 0, 1$. Эти функции не убывают и для них выполняются неравенства (мы для краткости опускаем индекс i):

$$f[F(x)] \leq x \text{ и } F[f(y)] \geq y \text{ для всех } x \text{ или } y.$$

Поэтому $y \leq F(x_0)$ влечет $f(y) \leq f[F(x_0)] \leq x_0$ и, обратно, $f(y) \leq x_0$ влечет $F[f(y)] \leq F(x_0)$ и $y \leq F(x_0)$, так что неравенства $f(y) \leq x_0$ и $y \leq F(x_0)$ — эквивалентны. Пусть V распределено равномерно на $(0, 1)$. Тогда $P\{f_i(V) \leq x\} = P\{V \leq F_i(x)\} = F_i(x)$. Из того, что при всех x $F_1(x) \leq F_0(x)$ вытекает, что при всех y $f_0(y) \leq f_1(y)$, чем доказательство завершается.

Одним из простейших примеров стохастически упорядоченного семейства является семейство с параметром сдвига, т. е. семейство, для которого

$$F_\theta(x) = F(x - \theta).$$

Покажем, что это семейство стохастически возрастает. Пусть X — случайная величина с функцией распределения $F(x)$. Тогда из $\theta < \theta'$ вытекает

$$F(x - \theta) = P\{X \leq x - \theta\} \geq P\{X \leq x - \theta'\} = F(x - \theta'),$$

что и требовалось доказать.

Другим примером могут служить семейства с монотонным отношением правдоподобия. Это видно из приводимой ниже леммы, в которой указываются некоторые существенные черты таких семейств.

Лемма 2*). Пусть $p_\theta(x)$ — семейство одномерных плотностей с отношением правдоподобия, монотонным относительно x .

*) Это — специальный случай теоремы Карлина, связывающей число перемен знака у $E_\theta \psi(x)$ и у $\psi(x)$ для плотностей типа Пойа. См. Karlin, Pólya type distributions II, Ann. Math. Stat. 28 (1957), 281—308.

(I) Если ψ — неубывающая функция x , то $E_{\theta}\psi(x)$ — неубывающая функция θ ; если $X_1, \dots, X_n$ независимы и имеют одну и ту же плотность p_{θ} , а ψ' — функция от $x_1, \dots, x_n$, неубывающая по каждому аргументу в отдельности, то $E_{\theta}\psi'(X_1, \dots, X_n)$ является неубывающей функцией θ .

(II) Для любых $\theta < \theta'$ функции распределения X , соответствующие значениям θ и θ' параметра, удовлетворяют неравенству

$$F_{\theta'}(x) \leq F_{\theta}(x) \text{ при всех } x.$$

(III) Пусть ψ — функция, меняющая знак ровно один раз. Более точно, пусть существует значение x_0 такое, что $\psi(x) \leq 0$ при $x < x_0$ и $\psi(x) \geq 0$ при $x \geq x_0$. Тогда существует θ_0 , для которого $E_{\theta}\psi(X) \leq 0$ при $\theta < \theta_0$ и $E_{\theta}\psi(X) \geq 0$ при $\theta > \theta_0$; при этом исключается случай, когда $E_{\theta}\psi(X)$ положительно (или отрицательно) при всех θ .

Доказательство. (I) Пусть $\theta < \theta'$. Обозначим A и B множества, на которых соответственно $p_{\theta'}(x) < p_{\theta}(x)$ и $p_{\theta'}(x) > p_{\theta}(x)$. Если $a = \sup_A \psi(x)$ и $b = \inf_B \psi(x)$, то $b - a \geq 0$ и

$$\begin{aligned} \int \psi(p_{\theta'} - p_{\theta}) d\mu &\geq a \int_A (p_{\theta'} - p_{\theta}) d\mu + b \int_B (p_{\theta'} - p_{\theta}) d\mu = \\ &= (b - a) \int_B (p_{\theta'} - p_{\theta}) d\mu \geq 0, \end{aligned}$$

что доказывает первое утверждение. Для $n > 1$ применяется индукция.

(II) Положим $\psi(x) = 1$ при $x > x_0$ и $\psi(x) = 0$ в других случаях. Тогда утверждение (II) следует из (I).

(III) Мы покажем сначала, что при любых $\theta' < \theta''$ из $E_{\theta'}\psi(X) > 0$ вытекает $E_{\theta''}\psi(X) > 0$. Если $p_{\theta''}(x_0)/p_{\theta'}(x_0) = \infty$, то $p_{\theta'}(x) = 0$ при $x \geq x_0$ и, следовательно, $E_{\theta'}\psi(X) \leq 0$. Допустим поэтому, что $p_{\theta''}(x_0)/p_{\theta'}(x_0) = c < \infty$. Тогда $\psi(x) \geq 0$ на множестве $S = \{x: p_{\theta'}(x) = 0 \text{ и } p_{\theta''}(x) > 0\}$ и

$$E_{\theta''}\psi(X) \geq \int_S \psi \frac{p_{\theta''}}{p_{\theta'}} p_{\theta'} d\mu \geq \int_{-\infty}^{x_0-} c\psi p_{\theta'} d\mu + \int_{x_0}^{\infty} c\psi p_{\theta'} d\mu = cE_{\theta'}\psi(X) \geq 0.$$

Выбирая $\theta_0 = \inf\{\theta: E_{\theta}\psi(X) > 0\}$, мы приходим к нужному результату.

Часть (II) леммы показывает, что каждое семейство распределений с монотонным по x отношением правдоподобия является стохастически возрастающим. Обратное утверждение неверно.

Так, для распределений Коши

$$\frac{1}{\pi} \frac{1}{1 + (x - \theta)^2}$$

отношение правдоподобия не будет монотонным, хотя семейство стохастически возрастает (так как θ — параметр сдвига). Условия, при выполнении которых семейство с параметром сдвига обладает монотонным отношением правдоподобия, даны в главе 8 (пример 1).

4. Сравнение экспериментов *)

Предположим, что для проверки простой гипотезы H при простой альтернативе K возможно провести различные эксперименты. Допустим, что результатом одного эксперимента является случайная величина X , которая имеет плотность $f(g)$ при гипотезе H (K); результатом другого эксперимента пусть будет X' с плотностью $f'(g')$. Обозначим $\beta(\alpha)$ и $\beta'(\alpha)$ мощности наиболее мощных критериев, основанных на X и X' соответственно. В общем случае соотношение между $\beta(\alpha)$ и $\beta'(\alpha)$ зависит от α . Однако, если $\beta'(\alpha) \leq \beta(\alpha)$ при всех α , то X или эксперимент $f(g)$ будет называться *лучше информирующим*, чем X' . Рассмотрим, например, экспоненциальное семейство плотностей $p_\theta(x)$ (см. (12)) и примем

$$f = f' = p_{\theta_0}, \quad g = p_{\theta_1}, \quad g' = p_{\theta_2}, \quad \text{где } \theta_0 < \theta_1 < \theta_2.$$

Тогда, в силу теоремы 2, эксперимент (f, g) лучше информирует, чем (f', g') .

Простое достаточное условие **) того, чтобы X лучше информировало, чем X' , состоит в том, что существуют функция $h(x, u)$ и случайная величина U , не зависящая от X и имеющая известное распределение, такие, что плотность $Y = h(X, U)$ равна f' или g' , если плотность X равна f или g . Так же как и в теории достаточных статистик, это утверждение вытекает из возможности построить, исходя из X (с помощью U) случайную величину Y , эквивалентную X' . Можно рассуждать также следующим образом. Пусть $\phi(x')$ — наиболее мощный критерий уровня α для проверки f' при альтернативе g' . Положим $\psi(x) = E\phi[h(x, U)]$. Тогда $E\psi(X) = E\phi(X')$ как при H , так и при K . Критерий $\psi(x)$ имеет уровень α и мощность $\beta'(\alpha)$ и, следовательно, $\beta(\alpha) \geq \beta'(\alpha)$.

*) Этот раздел представляет отступление и может быть пропущен.

**) Доказательство того, что это условие также необходимо, см. Blackwell, Comparison of experiments, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, Univ. Calif. Press, 1951.

В случаях, когда указанная функция h существует, мы будем говорить, что эксперимент (f, g) *достаточен* для (f', g') . Тогда, если $X_1, \dots, X_n$ и $X'_1, \dots, X'_n$ — суть выборки из распределений, соответствующих X и X' , то первая из этих выборок лучше информирует, чем вторая. Она также лучше информирует, чем $(Z_1, \dots, Z_n)$, где каждое из Z_i с некоторой вероятностью равно X_i или X'_i .

Пример 4. Каждый элемент совокупности обладает или не обладает одним из двух признаков — A и B . Необходимо проверить гипотезу о независимости этих признаков. Вероятности $p = P(A)$ и $\pi = P(B)$ (т. е. доли объектов, обладающих признаками A и B) предполагаются известными. Это знание может проистекать оттого, что ранее каждый из признаков был исследован по отдельности. Вероятности четырех возможных комбинаций $AB, A\tilde{B}, \tilde{A}B, \tilde{A}\tilde{B}$, как вычисленные при гипотезе независимости, так и вычисленные при альтернативе, что $P(AB)$ равно некоторому заданному ρ , — приведены в следующей таблице:

	При H		При K	
	B	$\tilde{B}$	B	$\tilde{B}$
A	$p\pi$	$p(1 - \pi)$	ρ	$p - \rho$
$\tilde{A}$	$(1 - p)\pi$	$(1 - p)(1 - \pi)$	$\pi - \rho$	$1 - p - \pi + \rho$

Допустим, что экспериментальный материал доставляется выборкой объема s . Эта выборка может быть взята, например, наудачу из числа элементов совокупности, которые обладают признаком A . Для каждого элемента выборки отмечают, обладает он признаком B или нет, т. е., следовательно, имеют дело с выборкой из биномиального распределения с вероятностями

$$H: P(B|A) = \pi \quad \text{и} \quad K: P(B|A) = \frac{\rho}{p}.$$

Вместо A можно было бы взять какой-нибудь из остальных признаков. При этом каждый раз получалась бы выборка из биномиального распределения. Соответствующие вероятности приведены в таблице:

Выбранная совокупность	Вероятность	H	K
A	$P(B A)$	π	ρ/p
B	$P(A B)$	p	ρ/π
$\tilde{B}$	$P(A \tilde{B})$	p	$(p - \rho)/(1 - \pi)$
$\tilde{A}$	$P(B \tilde{A})$	π	$(\pi - \rho)/(1 - p)$

Не ограничивая общности, допустим, что обозначения классов A , $\bar{A}$, B и $\bar{B}$ выбраны так, что $p \leq \pi \leq 1/2$. Мы покажем теперь, что из четырех перечисленных экспериментов первый (выбор из A) является лучше всего информирующим и, более того, достаточным для каждого из оставшихся остальных.

Сравним выбор из A и из B . Пусть X и X' принимают значения 0 и 1, и пусть вероятность равенства единице дается соответственно первый и второй строками таблицы. Пусть U не зависит от X и распределено равномерно на $(0, 1)$, и пусть $Y = h(X, U) = 1$ при $X = 1$ и $U \leq p/\pi$, $Y = 0$ в других случаях. Тогда $P\{Y = 1\}$ равна p при H и равна p/π при K , так что Y имеет то же распределение, что и X' . Этим доказано, что X достаточно для X' и, следовательно, является лучше информирующим. Для сравнения выбора из A и выбора из B положим $Y = 1$ при $X = 0$ и $U \leq p/1 - \pi$, $Y = 0$ в других случаях. Тогда вероятности события $Y = 1$ совпадают с указанными в третьей строке таблицы. Наконец, вероятности события $Y = 1$ совпадают с указанными в последней строке таблицы, если положить $Y = 1$ при $X = 1$ и $U \leq (\pi - p)/(1 - p)$ и при $X = 0$ и $U > (1 - \pi - p)/(1 - p)$.

Из общих замечаний, предшествующих разбираемому примеру, вытекает, что экспериментальный материал, если он должен доставляться выборкой объема s , необходимо брать из категории A , т. е. самой редкой из всех категорий. Этот способ предпочтительнее извлечения выборки объема s из всей совокупности, так как последняя процедура эквивалентна выбору каждого элемента или из A (с вероятностью p), или из $\bar{A}$ (с вероятностью $1 - p$).

Установленное соотношение между различными экспериментами не зависит не только от α , но и от p . Далее, если выборка взята из категории A , то по следствию 2 существует РНМ критерий H при односторонней альтернативе положительной связи: $P(B|A) > \pi$, т. е. $p > \pi p$, в соответствии с которой вероятности AB и $\bar{A}\bar{B}$ больше, а вероятности $A\bar{B}$ и $\bar{A}B$ меньше, чем при гипотезе независимости. Этот критерий обладает наибольшей мощностью, которой можно достичь, проверяя гипотезу независимости по выборке объема s .

Пример 5. В процессе Пуассона число событий, наступающих в интервале времени длины v , имеет распределение Пуассона $P(\lambda v)$. Проблема проверки гипотезы о том, что значение параметра равно λ_0 при альтернативном значении λ_1 , возникает и при изучении пространственного распределения частиц в объеме v . Покажем, что эксперимент информирует тем лучше, чем длиннее промежутки наблюдения v . Пусть $v < w$. Обозначим X и Y число событий в промежутках $(t, t + v)$ и $(t + v, t + w)$. Тогда величины X и Y независимы и имеют каждая распределение Пуассона; $Z = X + Y$ является достаточной статистикой для λ . Таким образом, каждый критерий, основанный на X , может быть заменен равносильным критерием, основанным на Z , т. е. Z лучше информирует, чем X . То, что Z действительно информирует существенно лучше, чем X (в очевидном значении этого высказывания), можно понять из следующего: единственный наиболее мощный критерий проверки λ_0 при альтернативе λ_1 зависит от $X + Y$ и, следовательно, не может быть заменен критерием, основанным только на X .

Иногда бывает невозможно определить число наступивших событий, но можно установить, имело ли место хотя бы одно из них или нет. В бактериологии, например, бактериальная культура смешивается с некоторым количеством воды, из которой берется затем несколько проб фиксированного объема и каждая испытывается на наличие или отсутствие в ней бактерий, т. е. в каждой из p проб может наступить или не наступить некоторое событие. В результате мы приходим к биномиальной схеме с вероятностью успеха (т. е. появления по крайней мере одной бактерии)

$$p = 1 - e^{-\lambda v}.$$

Так как очень большие или очень маленькие объемы почти достоверно приво-

дят к успеху или неудаче, то можно подозревать, что для проверки λ_0 при альтернативе λ_1 промежуточные значения v лучше информируют, чем крайние. Оказывается, однако, что эксперименты $(\lambda_0 v, \lambda_1 v)$ и $(\lambda_0 w, \lambda_1 w)$ не сравнимы ни при каких значениях v и w (см. задачу 14)*).

5. Доверительные границы

Теорию РНМ односторонних критериев можно применить к отысканию нижней и верхней границ для действительного параметра θ . Задача отыскания нижней границы возникает, например, когда θ есть величина разрушающего усилия для нового сплава; отыскание верхней границы необходимо, когда θ — токсичность медикамента или вероятность нежелательного события. Рассуждения для случая верхней и нижней границ сходны. Поэтому достаточно рассмотреть только случай нижней границы θ .

Так как $\underline{\theta} = \underline{\theta}(X)$ должно быть функцией результатов наблюдения, то невозможно требовать, чтобы $\underline{\theta}$ достоверным образом было меньше θ . Можно лишь требовать, чтобы вероятность этого события была близка к единице. Выберем число $1 - \alpha$ — *доверительный уровень* — и ограничимся границами $\underline{\theta}$, для которых

$$P_{\theta} \{ \underline{\theta}(X) \leq \theta \} \geq 1 - \alpha \quad \text{для всех } \theta. \quad (14)$$

Функция $\underline{\theta}$ называется нижней *доверительной границей* для θ , соответствующей доверительному уровню $1 - \alpha$; точная нижняя грань левой части (14), которая практически обычно равна $1 - \alpha$, называется *доверительным коэффициентом* $\underline{\theta}$.

Границу $\underline{\theta}$ следует выбирать так, чтобы она удовлетворяла условию (14) и недооценка θ была возможно меньше. Можно потребовать, например, чтобы вероятность попадания $\underline{\theta}$ левее любого θ была минимальна. Функция $\underline{\theta}$, для которой

$$P_{\theta} \{ \underline{\theta}(X) \leq \theta' \} = \min \quad (15)$$

при всех $\theta' < \theta$, и для которой имеет место (14), называется равномерно *наиболее точной* нижней доверительной границей для θ , соответствующей доверительному уровню $1 - \alpha$.

Пусть $L(\theta, \underline{\theta})$ обозначает меру потерь, возникающих от недооценки θ , так что при каждом фиксированном θ функция $L(\theta, \underline{\theta})$ определена и неотрицательна для $\underline{\theta} < \theta$, и не возрастает по второму аргументу. Тогда желательно минимизировать

$$E_{\theta} L(\theta, \underline{\theta}) \quad (16)$$

*) Вопрос о выборе v в этой и подобных ситуациях обсуждается у Hodges, The choice of inspection stringency in acceptance sampling by attributes, Univ. Calif. Publ. Statistics, 1 (1949), 1—14.

при условии (14). Можно показать, что равномерно наиболее точная нижняя доверительная граница $\underline{\theta}$ минимизирует (16) при условии (14), какова бы ни была функция потерь L (см. задачу 17).

Отыскание равномерно наиболее точных нижних доверительных границ облегчается введением более общего понятия, которое будет рассмотрено детальнее в главе 5. Семейство подмножеств $S(x)$ пространства параметров Ω является, по определению, семейством *доверительных множеств*, соответствующих доверительному уровню $1 - \alpha$, если

$$P_{\theta} \{ \theta \in S(X) \} \geq 1 - \alpha \quad \text{для всех } \theta \in \Omega, \quad (17)$$

т. е. если случайное множество $S(X)$ покрывает истинную параметрическую точку с вероятностью $\geq 1 - \alpha$. Нижняя доверительная граница соответствует специальному случаю, когда $S(x)$ есть односторонний интервал

$$S(x) = \{ \theta: \underline{\theta}(x) \leq \theta < \infty \}.$$

Теорема 4. (I) Рассмотрим для каждого $\theta_0 \in \Omega$ какой-либо критерий уровня α для проверки гипотезы $H(\theta_0): \theta = \theta_0$. Обозначим $A(\theta_0)$ область принятия гипотезы. При каждом x определим множество $S(x)$ значений параметра, полагая

$$S(x) = \{ \theta: x \in A(\theta), \theta \in \Omega \}.$$

Тогда $S(x)$ является семейством доверительных множеств для θ с доверительным уровнем $1 - \alpha$.

(II) Если $A(\theta_0)$ является РНМ критерием уровня α для проверки $H(\theta_0)$ при альтернативе $K(\theta_0)$, то $S(x)$ минимизирует вероятность

$$P_{\theta} \{ \theta' \in S(X) \} \quad \text{для всех } \theta \in K(\theta')$$

в классе всех семейств доверительных множеств для θ с доверительным уровнем $1 - \alpha$.

Доказательство. (I) По определению $S(x)$

$$\theta \in S(x) \quad \text{тогда и только тогда, когда } x \in A(\theta), \quad (18)$$

и, следовательно,

$$P_{\theta} \{ \theta \in S(X) \} = P_{\theta} \{ X \in A(\theta) \} \geq 1 - \alpha.$$

(II) Если $S^*(x)$ — любое другое семейство доверительных множеств уровня $1 - \alpha$ и если $A^*(\theta) = \{ x: \theta \in S^*(x) \}$, то

$$P_{\theta} \{ X \in A^*(\theta) \} = P_{\theta} \{ \theta \in S^*(X) \} \geq 1 - \alpha,$$

так что $A^*(\theta_0)$ является областью принятия для критерия $H(\theta_0)$ с уровнем α . Из допущений, касающихся $A(\theta_0)$, вытекает, что

при любом $\theta \in K(\theta_0)$

$$P_{\theta} \{X \in A^*(\theta_0)\} \geq P_{\theta} \{X \in A(\theta_0)\}$$

и что, следовательно,

$$P_{\theta} \{\theta_0 \in S^*(X)\} \geq P_{\theta} \{\theta_0 \in S(X)\},$$

что и требовалось доказать.

Соотношение эквивалентности (18) описывает структуру доверительных множеств $S(x)$; $S(x)$ есть совокупность тех значений параметра, для которых принимается $H(\theta)$ в случае, когда наблюдается x . Семейство доверительных множеств связывается, таким образом, с критериями различных гипотез $H(\theta)$, показывая как те значения, при которых гипотеза принимается ($\theta \in S(x)$), так и те, при которых она отвергается ($\theta \in \bar{S}(x)$).

Следствие 3. Допустим, что семейство плотностей $p_{\theta}(x)$, $\theta \in \Omega$, имеет отношение правдоподобия, монотонное относительно $T(x)$, и что функция распределения $F_{\theta}(t)$ статистики $T = T(X)$ непрерывна по t при каждом фиксированном θ .

(I) При каждом значении $1 - \alpha$ доверительного уровня существует равномерно наиболее точная доверительная граница θ для θ .

(II) Если x обозначает наблюдаемое значение X и $t = T(x)$ и если уравнение

$$F_{\theta}(t) = 1 - \alpha \quad (19)$$

имеет решение $\theta = \hat{\theta} \in \Omega$, то это решение единственно и $\theta(x) = \hat{\theta}$.

Доказательство. (I) Для каждого θ_0 найдется константа $C(\theta_0)$ такая, что $P_{\theta_0} \{T > C(\theta_0)\} = \alpha$ и по теореме 2 множество $T > C(\theta_0)$ можно рассматривать как критическую область РНМ критерия уровня α для проверки гипотезы $\theta = \theta_0$ при альтернативе $\theta > \theta_0$. По следствию 1 мощность этого критерия при любой альтернативе $\theta_1 > \theta_0$ превосходит α . Следовательно, $C(\theta_0) < C(\theta_1)$ и функция C — строго возрастающая. Обозначим $A(\theta_0)$ соответствующую область принятия: $T \leq C(\theta_0)$ и определим $S(x)$ формулой (18). Из монотонности функции C вытекает, что $S(x)$ состоит из тех $\theta \in \Omega$, для которых $\underline{\theta} \leq \theta$, где

$$\underline{\theta} = \inf \{\theta: T(x) \leq C(\theta)\}.$$

По теореме 4 множества $\{\theta: \underline{\theta}(x) \leq \theta\}$ образуют семейство доверительных множеств с уровнем $1 - \alpha$, и это семейство минимизирует $P\{\underline{\theta} \leq \theta'\}$ при всех $\theta \in K(\theta')$, т. е. при всех $\theta > \theta'$. Этим доказано, что $\underline{\theta}$ — равномерно наиболее точная доверительная граница для θ .

(II) Из следствия 1 вытекает, что $F_{\theta}(t)$ является строго убывающей функцией θ в каждой точке t , в которой $0 < F_{\theta}(t) < 1$.

Следовательно, (19) имеет не более одного решения. Предположим теперь, что наблюдаемое значение T равно t и что уравнение $F_{\theta}(t) = 1 - \alpha$ имеет решение $\hat{\theta} \in \Omega$. Тогда $F_{\hat{\theta}}(t) = 1 - \alpha$ и, по определению функции C , $C(\hat{\theta}) = t$. Неравенство $t \leq C(\theta)$ эквивалентно неравенству $C(\hat{\theta}) \leq C(\theta)$, а следовательно, $\hat{\theta} \leq \theta$. Таким образом, $\theta = \hat{\theta}$, что и требовалось доказать.

При тех же предположениях соответствующая верхняя доверительная граница с коэффициентом доверия $1 - \alpha$ является решением $\bar{\theta}$ уравнения $P_{\theta}\{T \geq t\} = 1 - \alpha$ или уравнения $F_{\theta}(t) = \alpha$.

Пример 6. Чтобы определить верхнюю границу степени радиоактивности λ радиоактивного вещества, его наблюдают до момента регистрации счетчиком Гейгера m -й частицы. Совместная плотность вероятности для промежутков T_i ($i = 1, \dots, m$) между $(i-1)$ -й и i -й регистрациями равна

$$p(t_1, \dots, t_m) = \lambda^m e^{-\lambda \sum t_i}, \quad t_1, \dots, t_m \geq 0.$$

Пусть $T = \sum T_i$ обозначает полное время наблюдения. Тогда $2\lambda T$ имеет χ^2 -распределение с $2m$ степенями свободы и, как показано в примере 3, область принятия для наиболее мощного критерия проверки $H(\lambda_0)$: $\lambda = \lambda_0$ при альтернативе $\lambda < \lambda_0$ имеет вид $2\lambda_0 T \leq C$, где C определяется уравнением

$$\int_0^C \chi_{2m}^2 = 1 - \alpha.$$

Множество $S(t_1, \dots, t_m)$, определенное формулой (18), оказывается совокупностью тех λ , для которых $\lambda \leq C/2T$ и, как следует из теоремы 4, $\bar{\lambda} = C/2T$ является равномерно наиболее точной верхней оценкой для λ . Это заключение можно вывести также из следствия 3.

Если X или T имеют дискретное распределение, то следствие 3 не может быть непосредственно применено, так как функция распределения $F_{\theta}(t)$ не будет непрерывной, и для большинства значений θ_0 оптимальный критерий для $H: \theta = \theta_0$ будет рандомизированным. Однако каждый рандомизированный критерий, основанный на X , может быть представлен, как нерандомизированный критерий, зависящий от X и случайной величины U , распределенной равномерно на $(0,1)$. При этом X и U независимы. При данной критической функции ϕ определим критическую область R равенством

$$R = \{(x, u): u \leq \phi(x)\}.$$

Тогда, каково бы ни было распределение X

$$P\{(X, U) \in R\} = P\{U \leq \phi(X)\} = E\phi(X),$$

так что R имеет ту же самую функцию мощности, что и ϕ , и эти два критерия эквивалентны. Пара (X, U) имеет особенно про-

стое представление, когда X целочисленно. В этом случае статистика

$$T = X + U$$

эквивалентна паре (X, U) , так как с вероятностью 1

$$X = [T], \quad U = T - [T],$$

где $[T]$ обозначает наибольшее целое $\leq T$. Распределение T непрерывно и доверительные границы можно строить, отправляясь от этой статистики.

Пример 7. Допустим, что мы интересуемся верхней границей для биномиальной вероятности p , которая может быть, например, вероятностью того, что в дозе вакцины полиомиелита, произведенной в соответствии с определенными правилами, содержится хотя бы один живой вирус. Пусть $X_1, \dots, X_n$ обозначают исходы n испытаний, X_i принимает значения 1 и 0 с вероятностями p и q соответственно. Пусть $X = \sum X_i$. Тогда $T = X + U$ имеет плотность

$$\binom{n}{[t]} p^{[t]} q^{n-[t]}, \quad 0 \leq t < n+1.$$

Эти плотности удовлетворяют условиям следствия 3, и верхняя доверительная граница $\bar{p}$ является решением (если оно существует) уравнения

$$P_p \{T < t\} = \alpha,$$

где t — наблюдаемое значение T . Решение существует для всех $\alpha \leq t \leq n + \alpha$. Для $n + \alpha < t$ гипотеза $H(p_0): p = p_0$ при альтернативе $p < p_0$ принимается при любом p_0 , и следовательно, $\bar{p} = 1$. Для $t < \alpha$ $H(p_0)$ отклоняется при всех p_0 и доверительное множество $S(t)$ оказывается пустым. Рассмотрим теперь множества $S^*(t)$, которые равны $S(t)$ при $t \geq \alpha$ и состоят из одной точки $p = 0$ при $t < \alpha$. Они также образуют семейства доверительных множеств с уровнем $1 - \alpha$, так как при всех p

$$P_p \{p \in S^*(T)\} \geq P_p \{p \in S(T)\} = 1 - \alpha.$$

С другой стороны, $P_p \{p' \in S^*(T)\} = P_p \{p' \in S(T)\}$ при всех $p' > 0$, и, следовательно,

$$P_p \{p' \in S^*(T)\} = P_p \{p' \in S(T)\} \quad \text{при всех } p' > p.$$

Таким образом, семейство множеств $S^*(t)$ минимизирует вероятность покрытия p' при всех $p' > p$ в классе семейств с уровнем $1 - \alpha$. Соответствующая доверительная граница

$$\bar{p}^*(t) = \bar{p}(t) \quad \text{при } t \geq \alpha \quad \text{и} \quad \bar{p}^*(t) = 0 \quad \text{при } t < \alpha$$

является равномерно наиболее точной верхней доверительной границей для p с уровнем $1 - \alpha$.

На практике, с целью избежать рандомизации и установить границу, не зависящую от дополнительной величины U , величину T обычно заменяют величиной $X + 1 = [T] + 1$. Так как $\bar{p}^*(t)$ — неубывающая функция t , то получаемая на этом пути верхняя доверительная граница $\bar{p}^*([t] + 1)$ оказывается несколько завышенной; в качестве компенсации она приводит к большей вероятности попадания в область правее истинного p .

Пусть $\underline{\theta}$ и $\bar{\theta}$ обозначают нижнюю и верхнюю границы для θ с доверительными коэффициентами $1 - \alpha_1$ и $1 - \alpha_2$, и предположим,

что $\underline{\theta}(x) < \bar{\theta}(x)$ при всех x . Так, например, будет в условиях следствия 3, если $\alpha_1 + \alpha_2 < 1$. Интервалы $(\underline{\theta}, \bar{\theta})$ являются в этом случае *доверительными интервалами* для θ с доверительным коэффициентом $1 - \alpha_1 - \alpha_2$, т. е. они содержат истинное значение параметра с вероятностью $1 - \alpha_1 - \alpha_2$, так как

$$P_{\theta} \{ \underline{\theta} \leq \theta \leq \bar{\theta} \} = 1 - \alpha_1 - \alpha_2 \text{ при всех } \theta.$$

Если границы $\underline{\theta}$ и $\bar{\theta}$ равномерно наиболее точные, то они минимизируют, при данном уровне, математические ожидания $E_{\theta} L_1(\theta, \underline{\theta})$ и $E_{\theta} L_2(\theta, \bar{\theta})$ при любой L_1 , не возрастающей по $\underline{\theta}$ при $\underline{\theta} < \theta$ и равной 0 при $\underline{\theta} \geq \theta$, и при любой L_2 , не убывающей по $\bar{\theta}$ при $\bar{\theta} > \theta$ и равной 0 при $\bar{\theta} \leq \theta$. Полагая

$$L(\theta; \underline{\theta}, \bar{\theta}) = L_1(\theta, \underline{\theta}) + L_2(\theta, \bar{\theta}),$$

мы видим, что интервалы $(\underline{\theta}, \bar{\theta})$ минимизируют $E_{\theta} L(\theta; \underline{\theta}, \bar{\theta})$ при условиях

$$P_{\theta} \{ \underline{\theta} > \theta \} \leq \alpha_1, \quad P_{\theta} \{ \bar{\theta} < \theta \} \leq \alpha_2.$$

Примером подобной функции потерь может служить

$$L(\theta; \underline{\theta}, \bar{\theta}) = \begin{cases} \bar{\theta} - \underline{\theta}, & \text{если } \underline{\theta} \leq \theta \leq \bar{\theta}, \\ \bar{\theta} - \theta, & \text{если } \theta < \underline{\theta}, \\ \theta - \underline{\theta}, & \text{если } \bar{\theta} < \theta, \end{cases}$$

что представляется естественной мерой точности интервалов. Отметим, что действительная длина $\bar{\theta} - \underline{\theta}$ не была бы столь же осмысленной мерой, так как короткие интервалы, лежащие далеко от истинного θ , не дают никакой выгоды.

Важный предельный случай соответствует уровням $\alpha_1 = \alpha_2 = 1/2$. В предположениях следствия 3 и при условии, что область, где плотность положительна, не зависит от θ , так что критерии мощности невозможны при $\alpha < 1$, верхняя и нижняя доверительные границы $\bar{\theta}$ и $\underline{\theta}$ совпадают. Их общее значение удовлетворяет соотношению

$$P_{\theta} \{ \underline{\theta} \leq \theta \} = P_{\theta} \{ \theta \leq \bar{\theta} \} = \frac{1}{2},$$

и оценка $\underline{\theta}$ для θ в равной степени как недооценивает θ , так и переоценивает его. Оценки с этим свойством называются *медианно несмещенными* (о соотношении этого понятия несмещенности с другими см. задачу 3 в главе 1). Из вышеприведенных резуль-

татов вытекает, что среди всех медианно несмещенных оценок $\underline{\theta}$ минимизирует $EL(\theta, \underline{\theta})$ при любой функции потерь, которая для фиксированного θ имеет нулевой минимум при $\theta = \underline{\theta}$ и не убывает по мере удаления θ от $\underline{\theta}$ вправо или влево. Полагая, в частности, $L(\theta, \underline{\theta}) = 0$ при $|\theta - \underline{\theta}| \leq \Delta$ и $L(\theta, \underline{\theta}) = 1$ в других случаях, мы заключаем, что в классе всех медианно несмещенных оценок $\underline{\theta}$ минимизирует вероятность отклонения от θ большего заданной величины. Более общим образом $\underline{\theta}$ максимизирует вероятность

$$P_{\theta} \{-\Delta_1 \leq \theta - \underline{\theta} \leq \Delta_2\}$$

при любых $\Delta_1, \Delta_2 \geq 0$.

6. Обобщение фундаментальной леммы

Нижеследующее утверждение представляет собой полезное обобщение теоремы 1 на случай нескольких ограничивающих условий.

Теорема 5. Пусть $f_1, \dots, f_{m+1}$ — действительные функции, определенные на евклидовом пространстве $\mathcal{X}$ и интегрируемые относительно некоторой меры μ . Предположим, что при заданных константах $c_1, \dots, c_m$ существуют критические функции φ , для которых

$$\int \varphi f_i d\mu = c_i, \quad i = 1, \dots, m. \quad (20)$$

Обозначим $\mathcal{C}$ класс всех таких критических функций.

(I) В классе $\mathcal{C}$ существует элемент, максимизирующий

$$\int \varphi f_{m+1} d\mu,$$

(II) Для того чтобы элемент из $\mathcal{C}$ максимизировал

$$\int \varphi f_{m+1} d\mu,$$

достаточно существования таких констант $k_1, \dots, k_m$, что

$$\begin{aligned} \varphi(x) &= 1 \quad \text{при} \quad f_{m+1}(x) > \sum_{i=1}^m k_i f_i(x), \\ \varphi(x) &= 0 \quad \text{при} \quad f_{m+1}(x) < \sum_{i=1}^m k_i f_i(x). \end{aligned} \quad (21)$$

(III) Если элемент класса $\mathcal{C}$ удовлетворяет (21) с константами $k_1, \dots, k_m \geq 0$, то он максимизирует

$$\int \varphi f_{m+1} d\mu$$

в классе всех критических функций, для которых

$$\int \varphi f_i d\mu \leq c_i, \quad i = 1, \dots, m. \quad (22)$$

(IV) Множество M точек m -мерного пространства с координатами

$$\left(\int \varphi f_1 d\mu, \dots, \int \varphi f_m d\mu \right),$$

где φ — критическая функция, является выпуклым и замкнутым. Если $(c_1, \dots, c_m)$ — внутренняя точка*) M , то существуют константы $k_1, \dots, k_m$ и критерий φ , удовлетворяющие (20) и (21), и необходимое условие для того, чтобы элемент из $\mathcal{C}$ максимизировал

$$\int \varphi f_{m+1} d\mu,$$

состоит в том, что (21) выполняется μ -почти всюду.

Здесь термин «внутренняя точка M » в утверждении (IV) может означать точку, внутреннюю для M как относительно всего m -мерного пространства, так и относительно наименьшего линейного подпространства (размерности $\leq m$), содержащего M . Теорема справедлива при обеих интерпретациях, но является более сильной при последней, при которой она и будет доказана.

Отметим также, что вполне аналогичный результат верен в задаче минимизации $\int \varphi f_{m+1} d\mu$.

Доказательство. (I) Пусть $\{\varphi_n\}$ — такая последовательность функций из $\mathcal{C}$, для которой $\int \varphi_n f_{m+1} d\mu$ стремится к $\sup_{\varphi} \int \varphi f_{m+1} d\mu$. По теореме о слабой компактности для критических функций (теорема 3 в Дополнении) существуют подпоследовательность $\{\varphi_{n_i}\}$ и критическая функция φ такие, что

$$\int \varphi_{n_i} f_k d\mu \rightarrow \int \varphi f_k d\mu \quad \text{для } k = 1, \dots, m+1.$$

Отсюда следует, что $\varphi \in \mathcal{C}$ и максимизирует интеграл от $f_{m+1} d\mu$ в классе $\mathcal{C}$.

(II) и (III) доказываются точно так же, как часть (II) в теореме 1.

(IV) Замкнутость M вытекает из теоремы о слабой компактности, а его выпуклость из того факта, что, каковы бы ни были критические функции φ_1 и φ_2 , их комбинация $\alpha\varphi_1 + (1-\alpha)\varphi_2$ при любом $0 \leq \alpha \leq 1$ также является критической функцией.

*) Случай, когда это предположение не выполняется, рассмотрен в книге: Dantzig and Wald, On the fundamental lemma of Neyman and Pearson, Ann. Math. Stat. 22 (1951), 87—93.

Обозначим N множество точек $(m+1)$ -мерного пространства с координатами

$$\left(\int \phi f_1 d\mu, \dots, \int \phi f_{m+1} d\mu \right),$$

где ϕ пробегает класс всех критических функций. Тогда, по тем же причинам, что и выше, N будет выпуклым и замкнутым. Условимся обозначать координаты точек M и N символами $(u_1, \dots, u_m)$ и $(u_1, \dots, u_{m+1})$ соответственно. Точки N , первые m координат которых равны $c_1, \dots, c_m$, образуют замкнутый интервал $[c^*, c^{**}]$.

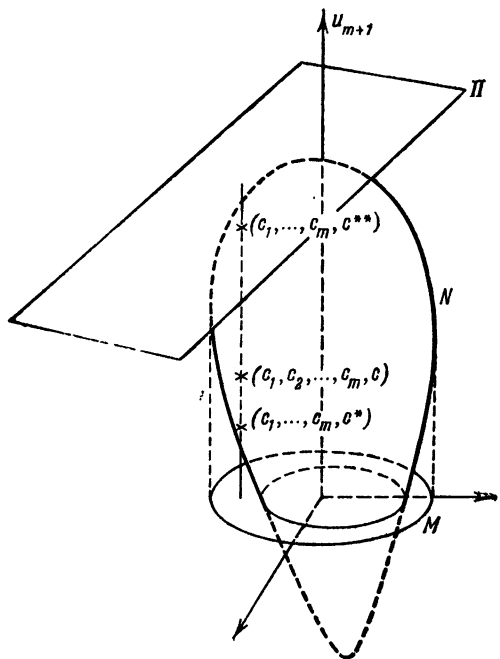


Рис. 4.

Предположим сначала, что $c^* < c^{**}$ (рис. 4). Так как $(c_1, \dots, c_m, c^{**})$ является граничной точкой N , то существует гиперплоскость Π , проходящая через эту точку и такая, что все точки N лежат или на Π , или под Π . Пусть уравнение Π имеет вид

$$\sum_{i=1}^{m+1} k_i u_i = \sum_{i=1}^m k_i c_i + k_{m+1} c^{**}.$$

Так как $(c_1, \dots, c_m)$ — внутренняя точка M , то коэффициент $k_{m+1} \neq 0$. Чтобы показать, это, выберем c , $c^* < c < c^{**}$. Точка

$(c_1, \dots, c_m, c)$ — внутренняя точка N . Поэтому существует сфера с центром в этой точке, целиком лежащая в N , и следовательно, расположенная под Π . Таким образом, мы видим, что точка $(c_1, \dots, c_m, c)$ не лежит на Π и, стало быть, $k_{m+1} \neq 0$. Мы можем положить $k_{m+1} = -1$, и тогда для каждой точки из N

$$u_{m+1} - \sum_{i=1}^m k_i u_i \leq c_{m+1}^* - \sum_{i=1}^m k_i c_i.$$

Иными словами, для любой критической функции φ_m

$$\int \varphi \left(f_{m+1} - \sum_{i=1}^m k_i f_i \right) d\mu \leq \int \varphi^* \left(f_{m+1} - \sum_{i=1}^m k_i f_i \right) d\mu,$$

где φ^* — критерий, соответствующий точке $(c_1, \dots, c_m, c^*)$. Таким образом, критическая функция φ^* максимизирует правую часть последнего неравенства. Так как интеграл, стоящий в этой части, получает максимальное значение при $\varphi = 1$ в точках, где подынтегральное выражение положительно, а при $\varphi = 0$ в других точках, то φ^* μ -почти всюду должна удовлетворять (21).

Рассмотрим случай $c^* = c^{**}$. Пусть $(c'_1, \dots, c'_m)$ — точка из M , отличная от $(c_1, \dots, c_m)$. Мы покажем сейчас, что существует ровно одно действительное число c' такое, что $(c'_1, \dots, c'_m, c')$ лежит в N . Допустим противное, т. е. что две точки $(c'_1, \dots, c'_m, c')$ и $(c'_1, \dots, c'_m, \bar{c}')$ лежат в N . Возьмем любую точку $(c''_1, \dots, c''_m, \bar{c}'')$ из N так, что точка $(c_1, \dots, c_m)$ лежит внутри прямолинейного отрезка, соединяющего $(c'_1, \dots, c'_m)$ и $(c''_1, \dots, c''_m)$. Такая точка существует, так как $(c_1, \dots, c_m)$ — внутренняя точка M . Тогда выпуклое множество, «натянутое» на три точки

$$(c'_1, \dots, c'_m, c'), \quad (c'_1, \dots, c'_m, \bar{c}') \quad \text{и} \quad (c''_1, \dots, c''_m, c''),$$

содержится в N и содержит точки $(c_1, \dots, c_m, \underline{c})$ и $(c_1, \dots, c_m, \bar{c})$ с $\underline{c} < \bar{c}$, что приводит к противоречию. Так как N выпукло, содержит начало и имеет самое большое одну точку на каждой вертикали $u_1 = c'_1, \dots, u_m = c'_m$, то N содержится в гиперплоскости, проходящей через начало и не параллельной оси u_{m+1} . Поэтому

$$\int \varphi f_{m+1} d\mu = \sum_{i=1}^m k_i \int \varphi f_i d\mu$$

для всех φ . Это может быть лишь в тривиальном случае

$$f_{m+1} = \sum_{i=1}^m k_i f_i \quad \mu\text{-почти всюду}$$

и, следовательно, (21) выполняется автоматически.

Следствие 4. Пусть $p_1, \dots, p_m, p_{m+1}$ являются плотностями по отношению к мере μ и пусть $0 < \alpha < 1$. Тогда существует критерий φ такой, что $E_i \varphi(X) = \alpha$ ($i = 1, \dots, m$) и $E_{m+1} \varphi(X) > \alpha$, исключая тот случай, когда $p_{m+1} = \sum_{i=1}^m k_i p_i$ μ -почти всюду.

Доказательство. Мы используем индукцию по m . При $m = 1$ утверждение сводится к следствию 1. Допустим, что оно доказано для любой системы из m плотностей, и рассмотрим случай $(m + 1)$ плотности $p_1, \dots, p_{m+1}$. Если $p_1, \dots, p_m$ линейно зависимы, то число p_i может быть сокращено, и теорема верна по предположению индукции. Примем теперь, что $p_1, \dots, p_m$ линейно независимы. Тогда для каждого $j = 1, \dots, m$ существуют, в силу предположения индукции, критерии φ_j и φ'_j такие что $E_i \varphi_j(X) = E_i \varphi'_j(X) = \alpha$ при всех $i = 1, \dots, j-1, j+1, \dots, m$ и $E_j \varphi_j(X) < \alpha < E_j \varphi'_j(X)$. Отсюда следует, что точка m -мерного пространства, все координаты которой равны α , является внутренней точкой M , так что применимо утверждение (IV) теоремы 5. Критерий $\varphi(x) \equiv \alpha$ таков, что $E_i \varphi(X) = \alpha$ при $i = 1, \dots, m$. Если среди всех критериев, удовлетворяющих наложенным ограничениям, этот критерий является наиболее мощным, он должен удовлетворять (21). Так как $0 < \alpha < 1$, то из сказанного следует

$$p_{m+1} = \sum_{i=1}^m k_i p_i \quad \mu\text{-почти всюду,}$$

что и требовалось доказать.

Наиболее полезными в теоремах 1 и 5 являются части (II), дающие достаточные условия для того, чтобы критическая функция максимизировала некоторый интеграл при определенных дополнительных ограничениях. Эти результаты могут быть весьма просто получены (как показано ниже) методом неопределенных множителей.

Лемма 3. Пусть $F_1, \dots, F_{m+1}$ — действительные функции, определенные на пространстве U . Рассмотрим проблему максимизации $F_{m+1}(u)$ при ограничениях $F_i(u) = c_i$ ($i = 1, \dots, m$). Для того чтобы точка u^0 , удовлетворяющая этим ограничениям, была решением данной проблемы, достаточно, чтобы она максимизировала выражение

$$F_{m+1}(u) - \sum_{i=1}^m k_i F_i(u)$$

хотя бы при одном выборе $k_1, \dots, k_m$.

Применяя эту лемму, максимизацию производят обычно при произвольных k и затем определяют эти константы так, чтобы выполнялись наложенные ограничения.

Доказательство. Пусть u — любая точка, удовлетворяющая наложенным ограничениям. Тогда

$$F_{m+1}(u) - \sum_{i=1}^m k_i F_i(u) \leq F_{m+1}(u^0) - \sum_{i=1}^m k_i F_i(u^0),$$

и, следовательно, $F_{m+1}(u) \leq F_{m+1}(u^0)$.

Рассмотрим, в качестве иллюстрации, проблему, разобранный в теореме 5. Пусть U будет пространством всех критических функций φ и пусть $F_i(\varphi) = \int \varphi f_i d\mu$. Тогда достаточное условие для того, чтобы φ максимизировала $F_{m+1}(\varphi)$ при ограничениях $F_i(\varphi) = c_i$, состоит в том, что φ максимизирует $F_{m+1}(\varphi) - \sum k_i F_i(\varphi) = \int (f_{m+1} - \sum k_i f_i) \varphi d\mu$. Последнее же выражение делается максимальным, если принять $\varphi(x) = 1$ или 0 при $f_{m+1}(x) >$ или $< \sum k_i f_i(x)$.

7. Двусторонние гипотезы

РНМ критерии существуют не только для односторонних, но также и для некоторых двусторонних гипотез типа

$$H: \theta \leq \theta_1 \quad \text{или} \quad \theta \geq \theta_2 \quad (\theta_1 < \theta_2). \quad (23)$$

Проблемы подобного рода могут возникать, например, если мы желаем определить, удовлетворяются ли заданные требования относительно пропорции ингредиента в каком-либо медикаменте. Другой пример — определение того, является ли измерительный инструмент, скажем линейка, правильно размеченным. Мы выдвигаем гипотезу, что θ не лежит в заданных пределах. Здесь ошибка первого рода состоит в объявлении, что θ удовлетворительно, в то время, как оно таковым не является. На практике решение принять H сопровождается обычно указанием, считать ли $\theta \leq \theta_1$ или $\theta \geq \theta_2$. Последствия гипотезы H бывают, обычно, весьма весомы, и поэтому принятие H в любом случае должно сопровождаться более тщательным исследованием. Так, если при испытании измерительного прибора перед его выпуском обнаруживают, что он не сбалансирован надлежащим образом, то потребуются дополнительная работа по доведению этого прибора до необходимого стандарта. В научном исследовании неравенства $\theta \leq \theta_1$ и $\theta \geq \theta_2$ могут противоречить некоторым высказанным ранее предположениям, и может потребоваться более сложная теория или дополнительные эксперименты. В подобных ситуациях допустимы только две основные возможности: или действовать так, как если бы $\theta_1 < \theta < \theta_2$, или производить дальнейшее исследование. Сведение задачи к проверке гипотезы H

может оказаться весьма подходящим. В настоящем разделе мы докажем для экспоненциальных семейств существование РНМ критерия для H .

Теорема 6. (I) В однопараметрическом экспоненциальном семействе существует РНМ критерий гипотезы $H: \theta \leq \theta_1$ или $\theta \geq \theta_2$ ($\theta_1 < \theta_2$) при альтернативе $K: \theta_1 < \theta < \theta_2$. Этот критерий определяется равенствами

$$\varphi(x) = \begin{cases} 1, & \text{если } C_1 < T(x) < C_2 \ (C_1 < C_2), \\ \gamma_i, & \text{если } T(x) = C_i, \ i = 1, 2, \\ 0, & \text{если } T(x) < C_1 \text{ или } > C_2, \end{cases} \quad (24)$$

где C_i и γ_i определяются из условий

$$E_{\theta_0} \varphi(X) = E_{\theta_0} \psi(X) = \alpha. \quad (25)$$

(II) Этот критерий минимизирует $E_{\theta} \varphi(X)$ при условии (25) при всех $\theta < \theta_1$ и $\theta > \theta_2$.

(III) При $0 < \alpha < 1$ функция мощности этого критерия имеет максимум в точке θ_0 , лежащей между θ_1 и θ_2 , и строго убывает при удалении θ от θ_0 вправо или влево. При этом исключается тот случай, когда существуют два числа t_1 и t_2 , для которых

$$P_{\theta} \{T(X) = t_1\} + P_{\theta} \{T(X) = t_2\} = 1 \quad \text{при всех } \theta.$$

Доказательство. (I) Мы можем ограничиться случаем, когда $T = T(X)$ — достаточная статистика, распределение которой в соответствии с леммой 7 главы 2 равно

$$dP_{\theta}(t) = C(\theta) e^{Q(\theta)t} dv(t),$$

где $Q(\theta)$ предполагается строго возрастающей. Пусть $\theta_1 < \theta' < \theta_2$. Рассмотрим сначала проблему максимизации $E_{\theta} \psi(T)$ при условии (25), если в нем $\varphi(x) = \psi[T(x)]$. Обозначим M совокупность всех точек $(E_{\theta_1} \psi(T), E_{\theta_2} \psi(T))$, получающуюся, когда ψ пробегает класс всех критических функций. Тогда точка (α, α) будет внутренней точкой M . Это вытекает из следующего соображения. По следствию 1 M содержит точки (α, u_1) и (α, u_2) с $u_1 < \alpha < u_2$, а потому содержит все точки (u, u) с $0 < u < 1$. По теореме 5 (утверждение (IV)) существуют константы k_1, k_2 и критерий $\psi_0(t)$ такие, что $\varphi_0(x) = \psi_0[T(x)]$ удовлетворяют (25) и что $\psi_0(t) = 1$, когда

$$k_1 C(\theta_1) e^{Q(\theta_1)t} + k_2 C(\theta_2) e^{Q(\theta_2)t} < C(\theta') e^{Q(\theta')t},$$

т. е. когда

$$a_1 e^{b_1 t} + a_2 e^{b_2 t} < 1 \quad (b_1 < 0 < b_2),$$

и $\psi_0(t) = 0$, когда левая часть последнего неравенства больше единицы.

Здесь оба числа α не могут одновременно быть отрицательными, так как тогда бы критерий всегда отклонял гипотезу. Если одно $\alpha \leq 0$, а другое $\alpha > 0$, то левая часть строго монотонна, и критерий имеет односторонний характер, рассмотренный в следствии 2. Поэтому его функция мощности строго монотонна и не может удовлетворять (25). Следовательно, оба числа α положительны и критерий удовлетворяет (24). По теореме 5 (III) он максимизирует $E_{\theta'}\psi(T)$ и при более слабых ограничениях $E_{\theta_i}\psi(T) \leq \alpha$ ($i = 1, 2$). Для завершения доказательства того, что этот критерий проверки H РНМ, необходимо установить, что он удовлетворяет неравенствам $E_{\theta}\psi(T) \leq \alpha$ при $\theta \leq \theta_1$ и при $\theta \geq \theta_2$. Это выводится из (II) посредством сравнения с критерием $\psi(t) \equiv \alpha$.

(II) Пусть $\theta' < \theta_1$. Для минимизации $E_{\theta'}\phi(X)$ при условии (25) используем теорему 5 (IV). Нетрудно видеть, что интересующему нас критерию соответствует критическая область вида

$$a_1 e^{b_1 t} + a_2 e^{b_2 t} < 1 \quad (b_1 < 0 < b_2).$$

Таким образом, он совпадает с критерием, найденным в (I). По теореме 5 (IV) первое и третье из условий (24) являются необходимыми и потому в предположении $P\{T = C_i\} = 0$ оптимальный критерий единствен.

(III) Не ограничивая общности, допустим, что $Q(\theta) = \theta$. Из (I) и непрерывности $\beta(\theta) = E_{\theta}\phi(X)$ следует, что или $\beta(\theta)$ удовлетворяет (III), или существуют три точки $\theta' < \theta'' < \theta'''$, для которых $\beta(\theta') = \beta(\theta'') = \beta(\theta''')$. Общее значение c этих величин лежит между 0 и 1, так как предположение $\beta(\theta') = 0$ (или 1) влекло бы $\phi(t) = 0$ (или 1) ν -почти всюду. Но эта возможность исключается равенствами (25). Как мы видели при доказательстве пункта (I), рассматриваемый критерий максимизирует $E_{\theta}\phi(X)$ при условиях $E_{\theta'}\phi(X) = E_{\theta'''}\phi(X) = c$, каково бы ни было $\theta' < \theta < \theta'''$. Возможность равенства $E_{\theta''}\phi(X) = c$ исключается следствием 4 во всех случаях, кроме того, когда $p_{\theta''} = k_1 p_{\theta'} + k_2 p_{\theta'''}$ ν -почти всюду. По сделанным в формулировке пункта (III) предположениям, из последнего равенства вытекало бы существование трех точек t_1, t_2, t_3 , для которых

$$1 = k_1 \frac{p_{\theta'}(t_i)}{p_{\theta''}(t_i)} + k_2 \frac{p_{\theta'''}(t_i)}{p_{\theta''}(t_i)}, \quad i = 1, 2, 3.$$

Но это невозможно, так как функция $[k_1 p_{\theta'}(t) + k_2 p_{\theta'''}(t)]/p_{\theta''}(t)$ выпукла.

На практике, чтобы определить C_i и γ_i , начинают с некоторых пробных значений C_1^* , γ_1^* , затем отыскивают такие C_2^* , γ_2^* , что $\beta^*(\theta_1) = \alpha$, и вычисляют $\beta^*(\theta_2)$. Последнее, как правило, оказывается или слишком большим, или слишком маленьким.

При выборе следующих пробных значений полезно заметить следующее: если $\beta^*(\theta_2) < \alpha$, то истинная область приемки лежит справа от выбранной, т. е. или $C_1 > C_1^*$, или $C_1 = C_1^*$ и тогда $\gamma_1 < \gamma_1^*$. Если $\beta^*(\theta_2) > \alpha$, то справедливо обратное утверждение. Сказанное вытекает из леммы 2, примененной к $T(x)$. Любой критерий φ^* , удовлетворяющий соотношениям (24) и $\beta^*(\theta_1) = \alpha$, должен располагаться или справа, или слева от критерия φ , удовлетворяющего (24) и (25). В соответствии с тем, находится ли φ слева или справа от φ^* функция $\psi(t) = \varphi^*(t) - \varphi(t)$ является монотонно возрастающей или монотонно убывающей, и по лемме $\beta^*(\theta_2) \geq \alpha$ или $\beta^*(\theta_2) \leq \alpha$.

Хотя в экспоненциальном семействе существует РНМ критерий проверки гипотезы: $\theta \leq \theta_1$ или $\theta \geq \theta_2$, РНМ критерий для двойственной гипотезы $H: \theta_1 \leq \theta \leq \theta_2$ или для гипотезы $\theta = \theta_0$ может не существовать (см. задачу 26). Но, как будет показано в главе 4, существуют РНМ несмещенные критерии для этих гипотез.

8. Наименее благоприятные распределения

Как следует из теоремы 1, всегда существует наиболее мощный критерий проверки простой гипотезы при простой альтернативе. Возьмем более общий случай. Пусть в евклидовом выборочном пространстве даны мера μ и функции f_θ , $\theta \in \omega$, и g , являющиеся плотностями по отношению к этой мере. В задаче проверки $H: f_\theta$, $\theta \in \omega$ при простой альтернативе $K: g$ существует наиболее мощный критерий уровня α . Это можно установить, исходя из теоремы о слабой компактности для критических функций (теорема 3 Дополнения); ср. теорему 5 (I).

В теореме 1 для случая простых гипотез указывается также способ построения наиболее мощного критерия. Мы распространим теперь эту теорему на сложные гипотезы, придерживаясь направления теоремы 5 и используя метод неопределенных множителей. Однако результат здесь оказывается заметно менее точным. Он оставляет, по существу, открытым вопрос о выборе множителей, совокупность которых принимает теперь форму произвольного распределения. В конкретных задачах этот выбор сопряжен со значительными трудностями.

С другой точки зрения, наш подход к задаче можно рассматривать, как сведение сложных гипотез к простым. Это достигается введением «взвешенных средних» распределений из H . Сложная гипотеза H заменяется при этом простой гипотезой H_λ , при которой плотность X определяется формулой

$$h_\lambda(x) = \int_{\omega} f_\theta(x) d\lambda(\theta),$$

где λ — распределение вероятностей на ω . Проблему отыскания надлежащего λ часто можно упростить с помощью следующих соображений. Так как H не дает информации о точном значении θ и так как H_λ должно быть эквивалентно H (с точки зрения альтернативы g), то естественно требовать, чтобы знание λ было как можно «менее полезным». Уточним эту мысль. Допустим, что θ имеет распределение λ . Тогда максимальная мощность β_λ , которая может быть достигнута при альтернативе g , соответствует наиболее мощному критерию ϕ_λ проверки H_λ при альтернативе g . Назовем распределение λ *наименее благоприятным* (при уровне α), если при всех λ' имеет место неравенство $\beta_\lambda \leq \beta_{\lambda'}$.

Теорема 7. Пусть на ω задано σ -поле такое, что плотности $f_\theta(x)$ измеримы по совокупности переменных θ и x . Предположим, что существует определенное на этом σ -поле распределение вероятностей λ со следующим свойством: наиболее мощный критерий ϕ_λ уровня α для проверки H_λ при альтернативе g имеет размер $\leq \alpha$ по отношению к первоначальной гипотезе H . Тогда:

(I) Критерий ϕ_λ является наиболее мощным для H при альтернативе g .

(II) Если ϕ_λ — единственный наиболее мощный критерий уровня α для H_λ при альтернативе g , то он будет единственным наиболее мощным критерием для H при альтернативе g .

(III) Распределение λ является наименее благоприятным.

Доказательство. Заметим сначала, что h_λ также является плотностью по отношению к μ . Действительно, по теореме Фубини (теорема 3 главы 2)

$$\int h_\lambda(x) d\mu(x) = \int_{\omega} d\lambda(\theta) \int f_\theta(x) d\mu(x) = \int_{\omega} d\lambda(\theta) = 1.$$

Предположим, что ϕ_λ — критерий уровня α для H , и пусть ϕ^* — любой другой критерий уровня α . Так как при всех $\theta \in \omega$ $E_\theta \phi^*(x) \leq \alpha$, то мы имеем

$$\int \phi^*(x) h_\lambda(x) d\mu(x) = \int_{\omega} E_\theta \phi^*(X) d\lambda(\theta) \leq \alpha.$$

Следовательно, ϕ^* — критерий уровня α для H_λ , и потому его мощность не превосходит мощности ϕ_λ . Этим доказаны (I) и (II). Пусть теперь λ' — любое распределение. Тогда ϕ_λ является критерием уровня α для $H_{\lambda'}$ и, следовательно, его мощность при альтернативе g не превосходит мощности наиболее мощного критерия, а эта последняя по определению равна $\beta_{\lambda'}$.

Условиям доказанной теоремы можно придать несколько иной вид. Заметим, что φ_λ может удовлетворять условиям

$$\int_{\omega} E_{\theta} \varphi_{\lambda}(X) d\lambda(\theta) = \alpha \quad \text{и} \quad E_{\theta} \varphi_{\lambda}(X) \leq \alpha \quad \text{для всех } \theta$$

только тогда, когда множество всех θ , для которых $E_{\theta} \varphi_{\lambda}(X) = \alpha$ имеет λ -меру единица.

Следствие 5. Пусть λ — распределение вероятностей на ω и ω' — подмножество ω , для которого $\lambda(\omega') = 1$. Пусть φ_λ — такой критерий, что

$$\varphi_{\lambda}(x) = \begin{cases} 1, & \text{если } g(x) > k \int f_{\theta}(x) d\lambda(\theta), \\ 0, & \text{если } g(x) < k \int f_{\theta}(x) d\lambda(\theta). \end{cases} \quad (26)$$

Тогда в предположении, что

$$E_{\theta'} \varphi_{\lambda}(X) = \sup_{\theta \in \omega} E_{\theta} \varphi_{\lambda}(X) = \alpha \quad \text{для всех } \theta' \in \omega', \quad (27)$$

φ_λ является наиболее мощным критерием для H при альтернативе g .

Из теоремы 7 легко получаются теоремы 2 и 6. Множество ω' , на котором сосредоточено наименее благоприятное распределение λ , в первом из этих примеров состоит из единственной точки θ_0 , а во втором — из двух точек θ_1 и θ_2 . Этого результата естественно было ожидать, так как в обоих случаях в H имеются распределения, представляющиеся ближайшими к K . Ниже приводится другой пример, когда наименее благоприятное распределение сосредоточено в одной точке.

Пример 8. Пусть X — характеристика, измеряющая качество изделий некоторого производства. Это может быть, например, величина разрывающего усилия для образцов материала, или длительность горения, или яркость электроламп. Изделие считается удовлетворительным, если X превосходит заданную величину u . Требуется проверить гипотезу $H: p \geq p_0$, где

$$p = P\{X \leq u\}$$

— вероятность получить дефектное изделие. Пусть $X_1, \dots, X_n$ — значения рассматриваемой характеристики в выборке объема n , так что X_k независимы и имеют одну и ту же функцию распределения, о которой никаких сведений не имеется. Каждое распределение на прямой может быть охарактеризовано заданием вероятности p и условных распределений P_- и P_+ величины X при условиях $X \leq u$ и $X > u$ соответственно. Если распределения P_- и P_+ имеют плотности p_- и p_+ , например, по отношению к мере $\mu = P_- + P_+$, то совместная плотность $X_1, \dots, X_n$ в выборочной точке $x_1, \dots, x_n$, для которой

$$x_{i_1}, \dots, x_{i_m} \leq u < x_{j_1}, \dots, x_{j_{n-m}},$$

будет равна

$$p^m (1-p)^{n-m} p_-(x_{i_1}) \dots p_-(x_{i_m}) p_+(x_{j_1}) \dots p_+(x_{j_{n-m}}).$$

Фиксируем теперь определенную альтернативу к H , например (p_1, P_-, P_+) с $p_1 < p_0$. При этом следовало бы ожидать, что наименее благоприятное распределение на H приписывает вероятность 1 распределению (p_0, P_-, P_+) , так как последнее представляется ближайшим к выбранной альтернативе. При таком выборе λ критерий (26) превращается в

$$\varphi_\lambda(x) = 1 \quad \text{или} \quad 0 \quad \text{при} \quad \left(\frac{p_1}{p_0}\right)^m \left(\frac{q_1}{q_0}\right)^{n-m} > \text{или} < C$$

(т. е. при $m < \text{или} > C$ соответственно). Критерий отвергает, следовательно, гипотезу, если число дефектных изделий M в выборке слишком мало. Точнее, он достоверно отвергает гипотезу при $M < C$ и отвергает ее с вероятностью γ при $M = C$, C и γ определяются равенством

$$P\{M < C\} + \gamma P\{M = C\} = \alpha \quad \text{при} \quad p = p_0. \quad (28)$$

Распределение числа M является биномиальным распределением $b(p, n)$ и не зависит от P_+ и P_- . Как следствие этого, функция мощности критерия зависит только от p и является убывающей функцией p , так что при H она принимает свое наибольшее значение для $p = p_0$. Этим доказано, что λ — наименее благоприятно, а критерий φ_λ — наиболее мощный. Так как критерий не зависит от специального выбора альтернативы, то он оказывается РНМ.

Выраженная в терминах переменных $Z_i = X_i - u$ статистика M , на которой основан критерий, равна числу этих переменных ≤ 0 . Сам критерий становится при этом *критерием знаков* (см. раздел 7 главы 4). Он может служить примером *непараметрического критерия*, так как он выведен без предположений такого рода, как предположения нормальности, равномерности, принадлежности классу законов Пуассона распределения X . Предположения такого рода оставляют неизвестными лишь некоторые параметры.

С очевидными видоизменениями вышеприведенные рассуждения можно применить к случаю, когда изделие признается удовлетворительным по признаку $u < X < v$. Такой случай может встретиться, например, если X — длина металлического изделия или пропорция ингредиента в некоторой смеси, для которых заданы некоторые допустимые пределы. Рассуждения применимы и в ситуациях, где X — вектор. Предположим, что изделие считается удовлетворительным только в том случае, когда X попадает в некоторое множество S , например, если размеры детали или пропорции ингредиентов лежат в заданных пределах. Вероятность дефектного изделия будет при этом равна

$$p = P\{X \in \tilde{S}\},$$

а P_- и P_+ будут обозначать условные распределения X при условиях $X \in S$ и $X \in \tilde{S}$ соответственно. Как и прежде, здесь существует РНМ критерий для H : $p \geq p_0$, который отвергает H , если число дефектных изделий M в выборке достаточно мало. Граница критической области определяется из (28).

Распределение λ , удовлетворяющее условиям теоремы 7, существует в большинстве обычных статистических задач и, в частности, при следующих предположениях *). Пусть выборочное пространство евклидово, ω — борелевское множество в s -мерном евклидовом пространстве, $f_\theta(x)$ — непрерывно по θ при почти всех x .

*) См. Lehmann, On the existence of least favorable distributions, Ann. Math. Stat. 23 (1952), 408—416.

Пусть

$$\lim_{n \rightarrow \infty} \int_S f_{\theta_n}(x) d\mu(x) = 0$$

для любого ограниченного множества S в выборочном пространстве и любой последовательности векторов θ_n , расстояние которых от начала неограниченно возрастает. Тогда, каково бы ни было g , существует удовлетворяющее условиям теоремы 7 распределение λ .

Отсюда следует (подобно тому, как следствия 1 и 4 вытекают из теорем 1 и 5), что при выполнении вышеприведенных условий и при $0 < \alpha < 1$ существует критерий проверки $H_0: f_{\theta}, \theta \in \omega$ при альтернативе g мощности $\beta > \alpha$. При этом исключается случай, когда при некотором λ $g = \int f_{\theta} d\lambda(\theta)$. Последняя возможность осуществляется, например, если f_{θ} и g являются нормальными плотностями $N(\theta, \sigma_0^2)$ и $N(\theta, \sigma_1^2)$, соответственно с $\sigma_0^2 < \sigma_1^2$ (см. стр. 112—113).

9. Проверка гипотез о среднем и дисперсии в нормальной совокупности

Проблемы проверки гипотез о среднем ξ и дисперсии σ^2 нормального распределения особенно важны, так как имеют широкую область применения. Здесь и в аналогичных проблемах в дальнейшем мы предполагаем, что параметр, который не оценивается, — неизвестен; но это не будет указываться при задании гипотезы. Мы будем писать, например, $\sigma \leq \sigma_0$ вместо более точного утверждения $\sigma \leq \sigma_0, -\infty < \xi < \infty$. Стандартные (основанные на отношении правдоподобия) критерии гипотез $\sigma \leq \sigma_0$ и $\xi \leq \xi_0$ задаются критическими областями:

$$\sum (x_i - \bar{x})^2 \geq C \quad (29)$$

и

$$\frac{\sqrt{n} (\bar{x} - \xi_0)}{\sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}} \geq C. \quad (30)$$

Критерии для гипотез $\sigma \geq \sigma_0$ и $\xi \geq \xi_0$ получаются заменой знаков неравенств в (29) и (30) на противоположные. Как будет показано в последующих главах, эти четыре критерия являются РНМ как в классе всех несмещенных, так и в классе всех инвариантных критериев. Однако при обычных уровнях значимости только первый из них является РНМ.

Пусть $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$. Рассмотрим сначала гипотезы $H_1: \sigma \geq \sigma_0$ и $H_2: \sigma \leq \sigma_0$ и простую альтернативу

K : $\xi = \xi_1$, $\sigma = \sigma_1$. Кажется целесообразным предположить, что наименее благоприятное распределение λ на (ξ, σ) -плоскости сконцентрировано на прямой $\sigma = \sigma_0$. Так как пара $Y = \Sigma X_i/n = \bar{X}$ и $U = \Sigma (X_i - \bar{X})^2$ образует достаточную статистику для параметров (ξ, σ) , то можно ограничиться этими величинами. Их совместная плотность при гипотезе H_λ равна

$$C_0 u^{\frac{1}{2}(n-3)} \exp\left(-\frac{u}{2\sigma_0^2}\right) \int \exp\left[-\frac{n}{2\sigma_0^2}(y - \xi)^2\right] d\lambda(\xi),$$

а при K равна

$$C_1 u^{\frac{1}{2}(n-3)} \exp\left(-\frac{u}{2\sigma_1^2}\right) \exp\left[-\frac{n}{2\sigma_1^2}(y - \xi_1)^2\right].$$

Как видно, выбор λ влияет только на распределение Y . Наименее благоприятное λ должно, следовательно, обладать тем свойством, что плотность Y при H_λ

$$\int \frac{\sqrt{n}}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{n}{2\sigma_0^2}(y - \xi)^2\right] d\lambda(\xi)$$

сближается возможно сильнее с альтернативной плотностью

$$\frac{\sqrt{n}}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{n}{2\sigma_1^2}(y - \xi_1)^2\right].$$

Начиная с этого места, гипотезы H_1 и H_2 следует рассматривать по отдельности. В первом случае $\sigma_1 < \sigma_0$. Надлежащим выбором λ среднее значение Y может быть сделано равным ξ_1 , однако дисперсия если и изменится, то только возрастет по сравнению с первоначальным значением σ_{20} . Это наводит на мысль что наименее благоприятное распределение приписывает вероятность 1 точке $\xi = \xi_1$, так как при этом распределение Y оказывается нормальным и при H , и при K , причем среднее будет одно и то же в обоих случаях, а расхождение между дисперсиями будет минимальным. Для H_2 (при которой $\sigma_0 < \sigma_1$) положение будет несколько иным. Если наименее благоприятное распределение имеет плотность, скажем, λ' , то плотность Y при H_λ будет

$$\int_{-\infty}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{n}{2\sigma_0^2}(y - \xi)^2\right] \lambda'(\xi) d\xi.$$

Это — плотность суммы двух независимых случайных величин, одна из которых распределена по $N(0, \sigma_0^2/n)$, а другая имеет плотность $\lambda'(\xi)$. Если в качестве λ выбрать $N(\xi_1, (\sigma_1^2 - \sigma_0^2)/n)$, то

распределение Y при H_λ (как, впрочем, и при K) равно $N(\xi_1, \sigma_1^2/n)$.

Мы используем теперь следствие 5 с λ , предложенным выше. При H_1 более удобно оперировать с первоначальными величинами, чем с Y и U . Из (26) получаем, что $\varphi(x) = 1$ при

$$\frac{(2\pi\sigma_1^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma_1^2} \sum (x_i - \xi_1)^2 \right]}{(2\pi\sigma_0^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma_0^2} \sum (x_i - \xi_1)^2 \right]} > C,$$

т. е. при

$$\sum (x_i - \xi_1)^2 \leq C. \quad (31)$$

Чтобы оправдать выбор λ , мы должны показать, что

$$P \left\{ \sum (x_i - \xi_1)^2 \leq C \mid \xi, \sigma \right\}$$

в полуплоскости $\sigma \geq \sigma_0$ достигает максимального значения в точке $\xi = \xi_1, \sigma = \sigma_0$. Для любого фиксированного σ это выражение равно вероятности того, что случайная точка, координаты которой независимы и имеют распределение $N(\xi, \sigma^2)$, попадает внутрь сферы фиксированного радиуса. Это вероятность максимальна, когда центр сферы совпадает с центром распределения, т. е. когда $\xi = \xi_1$. Ее значение при этом равно

$$P \left\{ \sum \left(\frac{X_i - \xi_1}{\sigma} \right)^2 \leq \frac{C}{\sigma^2} \mid \xi_1, \sigma \right\} = P \left\{ \sum V_i^2 \leq \frac{C}{\sigma^2} \right\},$$

где $V_1, \dots, V_n$ независимы и распределены по $N(0, 1)$. Правая часть является убывающей функцией σ и потому максимальна при $\sigma = \sigma_0$.

В случае H_2 применение следствия 5 к достаточной статистике (Y, U) показывает, что $\varphi(y, u) = 1$ при

$$\frac{C_1 u^{\frac{1}{2}(n-3)} \exp \left(-\frac{u}{2\sigma_1^2} \right) \exp \left[-\frac{n}{2\sigma_1^2} (y - \xi_1)^2 \right]}{C_0 u^{\frac{1}{2}(n-3)} \exp \left(-\frac{u}{2\sigma_0^2} \right) \int \exp \left[-\frac{n}{2\sigma_0^2} (y - \xi)^2 \right] \lambda'(\xi) d\xi} = \\ = K \exp \left[-\frac{u}{2} \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right) \right] \geq C,$$

т. е. при

$$u = \sum (x_i - \bar{x})^2 \geq C. \quad (32)$$

Так как распределение $\sum (X_i - \bar{X})^2 / \sigma^2$ не зависит ни от ξ , ни от σ , то вероятность $P \{ \sum (X_i - \bar{X})^2 \geq C \mid \xi, \sigma \}$ не зависит от ξ и возрастает вместе с σ , так что условия следствия 5 выполняются. Критерий (32), будучи независимым от ξ_1 и σ_1 , является РНМ

для проверки гипотезы $\sigma \leq \sigma_0$ при альтернативе $\sigma > \sigma_0$. Как видно, он совпадает с критерием отношения правдоподобия (29). С другой стороны, наиболее мощный критерий (31) для проверки гипотезы $\sigma \geq \sigma_0$ при альтернативе $\sigma < \sigma_0$ не зависит от значения среднего при альтернативной гипотезе, т. е. от ξ_1 .

До сих пор неявно предполагалось, что $n > 1$. При $n = 1$ и гипотезе H_1 наши рассуждения применимы без изменений и приводят к (31) с $n = 1$. При гипотезе H_2 статистика U отпадает, и Y совпадает с X . Используя то же λ , что и раньше, мы видим, что X имеет одно и то же распределение как при H_λ , так и при K , и критерий ϕ_λ принимает вид $\phi_\lambda(x) \equiv \alpha$. Последний критерий удовлетворяет условиям следствия 5 и потому оказывается наиболее мощным критерием в рассматриваемой задаче. Мы приходим к выводу, что единственное наблюдение не имеет значения при проверке гипотезы H_2 (что интуитивно очевидно), но что оно может быть использовано для проверки H_1 , если класс альтернатив достаточно узок.

Соответствующее исследование для гипотезы $\xi \leq \xi_0$ более сложно. Можно показать*), что критерий Стьюдента (30) является наиболее мощным при уровне $\alpha \geq 1/2$ и любой альтернативе $\xi_1 > \xi_0$, σ_1 . Иными словами, этот критерий РНМ при $\alpha \geq 1/2$. С другой стороны, при $\alpha < 1/2$ наиболее мощный критерий для H отклоняет гипотезу при $\sum (x_i - a)^2 \leq b$, где a и b зависят от альтернативы (ξ_1, σ_1) и от α . Таким образом, для уровней значимости, представляющих действительный интерес, РНМ критерий для H не существует. Случай $\xi > \xi_0$ не порождает новых задач, так как преобразованием $Y_i = \xi_0 - (X_i - \xi_0)$ он сводится к уже рассмотренному.

10. Последовательный критерий отношений вероятностей

В соответствии с фундаментальной леммой Неймана — Пирсона, наилучшая процедура проверки простой гипотезы H : плотность X равна p_0 , при простой альтернативе: плотность X равна p_1 , принимает или отклоняет H в зависимости от того, больше или меньше отношение

$$\frac{p_{1n}}{p_{0n}} = \frac{p_1(x_1) \dots p_1(x_n)}{p_0(x_1) \dots p_0(x_n)}$$

некоторой надлежаще выбранной константы C . Однако, если не фиксировать объем выборки заранее, но сделать его зависящим от хода наблюдений, то возможны дальнейшие улучшения. Наилучшая, в некотором смысле, процедура доставляется следующим *последовательным критерием отношений вероятностей*.

*) См. Lehmann and Stein, Most powerful tests of composite hypotheses. I. Normal distributions, Ann. Math. Stat. 19 (1948), 495—516.

Пусть $A_0 < A_1$ — две заданные константы. Предположим, что наблюдения проводятся все время, пока отношение p_{1n}/p_{0n} удовлетворяет неравенствам

$$A_0 < \frac{p_{1n}}{p_{0n}} < A_1. \quad (33)$$

Гипотеза H принимается или отвергается при первом нарушении (33). Эти решения соответствуют неравенствам $p_{1n}/p_{0n} \leq A_0$ и $p_{1n}/p_{0n} \geq A_1$. Качество подобной процедуры измеряется обычно вероятностями α_0 и α_1 отклонения H при $p = p_0$ и принятия H при $p = p_1$ и, кроме того, средним числом $E_i(N)$ наблюдений при $p = p_i (i = 0, 1)$.

Теорема 8. Среди всех критериев (последовательных или нет), для которых

$$P_1 \text{ (принятия } H) \leq \alpha_1, \quad P_0 \text{ (отклонения } H) \leq \alpha_0$$

и $E_0(N)$, $E_1(N)$ конечны, последовательный критерий отношений правдоподобия с вероятностями ошибок α_0 и α_1 минимизирует как $E_0(N)$, так и $E_1(N)$.

В частности, последовательный критерий отношений вероятностей требует в среднем меньше наблюдений, чем критерий с фиксированным объемом выборки и такими же вероятностями ошибок. Мы отложим доказательство теоремы до раздела 12. В настоящем и следующем разделах будут указаны основные свойства последовательного критерия отношений вероятностей.

Точное определение границ A_0 и A_1 , соответствующих заданным α_0 и α_1 , сопряжено с большими трудностями. Поэтому полезны приводимые ниже неравенства. Обозначим R_n часть n -мерного пространства, определяемую неравенствами

$$A_0 < \frac{p_{1k}}{p_{0k}} < A_1 \quad \text{при } k = 1, \dots, n-1 \quad \text{и} \quad A_1 \leq \frac{p_{1n}}{p_{0n}}.$$

Это множество тех точек $(x_1, \dots, x_n)$, для которых процедура заканчивается на $N = n$ -м шаге отклонением H . Тогда

$$\alpha_0 = \sum_{n=1}^{\infty} \int_{R_n} p_{0n} \leq \frac{1}{A_1} \sum_{n=1}^{\infty} \int_{R_n} p_{1n} = \frac{1 - \alpha_1}{A_1}.$$

Аналогично, если S_n обозначает часть n -мерного пространства, для которой $N = n$ и H принимается, то

$$1 - \alpha_0 = \sum_{n=1}^{\infty} \int_{S_n} p_{0n} \geq \frac{\alpha_1}{A_0}.$$

При этом мы молчаливо предполагали, что

$$\sum_{n=1}^{\infty} p_i \{N = n\} = \sum_{n=1}^{\infty} \int_{R_n \cup S_n} p_{in} = 1 \quad \text{для } i = 0, 1,$$

т. е. что вероятность неограниченного продолжения процедуры равна 0. Доказательство этого факта см. в задачах (34) и (35). Неравенства

$$A_0 \geq \frac{\alpha_1}{1 - \alpha_0}, \quad A_1 \leq \frac{1 - \alpha_1}{\alpha_0} \quad (34)$$

наводят на мысль об аппроксимации границ A_0 и A_1 , соответствующих заданным α_0 и α_1 , величинами

$$A'_0 = \frac{\alpha_1}{1 - \alpha_0}, \quad A'_1 = \frac{1 - \alpha_1}{\alpha_0}.$$

В силу (34) вероятности ошибок в этой приближенной процедуре удовлетворяют неравенствам

$$\frac{\alpha'_1}{1 - \alpha'_0} \leq A'_0 = \frac{\alpha_1}{1 - \alpha_0} \quad \text{и} \quad \frac{1 - \alpha'_1}{\alpha'_0} \geq A'_1 = \frac{1 - \alpha_1}{\alpha_0},$$

откуда

$$\alpha'_0 \leq \frac{\alpha_0}{1 - \alpha_1} \quad \text{и} \quad \alpha'_1 \leq \frac{\alpha_1}{1 - \alpha_0}.$$

Если, как это типично бывает, α_0 и α_1 имеют порядок от 0,01 до 0,1, то превышение α'_i над α_i ($i = 1, 0$) оказывается пренебрежимо малым, т. е. с большой точностью можно принять, что ошибки обоих типов ограничены сверху заданными α_0 и α_1 . Последнее заключение подкрепляется тем фактом, что $\alpha'_0 + \alpha'_1 \leq \alpha_0 + \alpha_1$. Это можно видеть, складывая неравенства $\alpha'_1(1 - \alpha'_0) \leq \alpha_1(1 - \alpha'_0)$ и $\alpha'_0(1 - \alpha'_1) \leq \alpha_0(1 - \alpha'_1)$.

Единственный серьезный риск, связанный с употреблением приближенных границ, состоит, следовательно, в том, что α'_0 и α'_1 могут оказаться намного меньше заданных значений, что приведет к существенному увеличению числа необходимых наблюдений. Однако есть причины надеяться, что это увеличение будет умеренным. Действительно, положим

$$z_i = \log \left[\frac{p_1(x_i)}{p_0(x_i)} \right]. \quad (35)$$

Тогда (33) превращается в

$$\log A_0 < \sum_{i=1}^n z_i < \log A_1,$$

и если гипотеза H отклоняется, то величины z удовлетворяют неравенствам

$$z_1 + \dots + z_{n-1} < \log A_1 \leq z_1 + \dots + z_n.$$

Аппроксимация состоит в замене $z_1 + \dots + z_n$ на $\log A_1$. Ошиб-

ка, как правило, будет умеренной, так как после $(n - 1)$ -го наблюдения $\sum z_i$ все еще остается $< A_1$, и погрешности не накапливаются, а обуславливаются единственным наблюдением. Аналогичные замечания применимы к другой границе.

Пример 9. Рассмотрим последовательность биномиальных испытаний с постоянной вероятностью успеха p . Пусть задача состоит в проверке $p = p_0$ при альтернативе $p = p_1$ ($p_0 < p_1$). Тогда

$$\frac{p_{1n}}{p_{0n}} = \frac{p_1^{\sum x_i} (1 - p_1)^{n - \sum x_i}}{p_0^{\sum x_i} (1 - p_0)^{n - \sum x_i}} = \left(\frac{p_1 q_0}{p_0 q_1} \right)^{\sum x_i} \left(\frac{q_1}{q_0} \right)^n.$$

В случае, когда отношение $\log(p_1 p_0^{-1}) / \log(q_0 q_1^{-1})$ рационально, известны точные формулы *) для вероятностей ошибок и среднего размера выборки, что позволяет оценить результат замены точных A_0 и A_1 на приближенные A'_0 и A'_1 . Для иллюстрации **) предположим $p_0 = 0,05$, $p_1 = 0,17$, $\alpha_0 = 0,05$, $\alpha_1 = 0,1$. Тогда $\alpha'_0 = 0,031$, $\alpha'_1 = 0,099$ и средние размеры выборки для приближенной процедуры равны $E'_0(N) = 31,4$, $E'_1(N) = 30,0$. Имеется и другой план, найденный методом последовательных приближений, для которого $\alpha_0^* = 0,046$, $\alpha_1^* = 0,097$, $E_0^*(N) = 30,5$, $E_1^*(N) = 26,1$. С другой стороны, процедура с фиксированным объемом выборки и вероятностями ошибок 0,05 и 0,10 требует 57 наблюдений.

Для определенности мы предположили в определении последовательного критерия отношений вероятностей, что наблюдения продолжаются только то время, пока отношение вероятностей лежит строго между A_0 и A_1 . Но рассуждения применимы, в равной мере, и к правилу, по которому наблюдения продолжаются, пока $A_0 < p_{1n}/p_{0n} < A_1$ заканчиваются соответствующими решениями, когда впервые $p_{1n}/p_{0n} < A_0$ или $> A_1$, а на границах решения принимаются с некоторыми вероятностями. Термин «последовательный критерий отношений вероятностей» будет применяться и к этой более общей процедуре. Все эти процедуры эквивалентны, когда отношение $p_1(X)/p_0(X)$ имеет непрерывное распределение. Однако в случае дискретно распределенных отношений вероятностей целесообразно сохранить возможность рандомизации на границе, с тем чтобы получать в точности заданные вероятности ошибок. Если допустить рандомизацию так-

*) Girshick, Contributions to the theory of sequential analysis, II, III, Ann. Math. Stat. 17 (1946), 282—298 и Pólya, Exact formulas in the sequential analysis of attributes, Univ. Calif. Publ. Mathematics, New Series 29 (1958), 938—939.

**) Взято из книги: Robinson, A note on exact sequential analysis, Univ. Calif. Publ. Mathematics, New Series, 1 (1948), 241—246.

же и при решении вопроса: начинать наблюдения, или принять решение без наблюдений, то, как можно показать, достижимы любые наперед заданные вероятности ошибок *).

11. Мощность и средний размер выборки для последовательного критерия отношений вероятностей

Предыдущий раздел в какой-то степени вводит в заблуждение тем, что в нем рассматривается проверка простой гипотезы при простой альтернативе. Но эта постановка задачи интересна, главным образом, теми следствиями, которые получаются из нее применительно к более реалистической ситуации семейства распределений, зависящих от непрерывного параметра. К сожалению, утверждение о равномерно наибольшей мощности, справедливое для критерия отношений вероятностей с фиксированным объемом выборки в случае семейств с монотонным отношением правдоподобия (теорема 2), не распространяется на последовательный критерий. Более точно, рассмотрим последовательный критерий отношений вероятностей для проверки гипотезы $H: \theta_0$ при альтернативе $K: \theta_1$ и пусть $\beta(\theta) = P_\theta$ (отклонения H) — его функция мощности. Тогда, если θ_2 — некоторая другая альтернатива, то последовательный критерий отношений вероятностей для проверки θ_0 при этой альтернативе, соответствующий вероятностям ошибок α_0 и α_1 , не совпадает, вообще говоря, с первоначальным критерием, который поэтому не минимизирует $E_{\theta_2}(N)$. Действительно, представляется правдоподобным, что с общей точки зрения последовательный критерий отношений вероятностей не приводит к наилучшей последовательной процедуре для семейств с непрерывным параметром, хотя, обычно он лучше, чем соответствующий критерий с фиксированным объемом выборки.

Допустим, что плотность зависит от действительного параметра θ и что проверяется гипотеза $\theta \leq \theta_0$. При этом обычно не интересуются мощностью критерия при альтернативах θ , близких к θ_0 , но желают контролировать вероятность обнаружения альтернатив, достаточно удаленных от θ_0 . Критерий, следовательно, должен удовлетворять условиям

$$\begin{aligned} \beta(\theta) &\leq \alpha \quad \text{при} \quad \theta \leq \theta_0 \\ \beta(\theta) &\geq \beta \quad \text{при} \quad \theta \geq \theta_1 \end{aligned} \quad (\theta_0 < \theta_1), \quad (36)$$

что будет иметь место, в частности, когда

$$\beta(\theta_0) = \alpha, \quad \beta(\theta_1) = \beta$$

*) Этот результат содержится в не опубликованной статье: Stein, Existence of sequential probability ratio tests. См. также резюме Wijsman, On the existence of Wald's sequential test, Ann. Math. Stat. 29 (1958), 938—939.

и $\beta(\theta)$ — неубывающая функция θ . Последовательный критерий отношений вероятностей для проверки θ_0 при альтернативе θ_1 и вероятностях ошибок $\alpha_0 = \alpha$, $\alpha_1 = 1 - \beta$ — решает поставленную задачу, если предположить, что его функция мощности — неубывающая.

Лемма 4. Пусть $X_1, X_2, \dots$ независимы и одинаково распределены с плотностью $p_\theta(x)$. Допустим, что семейство $p_\theta(x)$ имеет отношение правдоподобия, монотонное относительно $T(x)$. Тогда любой последовательный критерий отношений вероятностей для проверки θ_0 при альтернативе θ_1 ($\theta_0 < \theta_1$) имеет неубывающую функцию мощности.

Доказательство. Пусть $Z_i = \log[p_{\theta_1}(X_i)/p_{\theta_0}(X_i)] = h(T_i)$ где h не убывает, и пусть $\theta < \theta'$. По лемме 2 для функции распределения $F_\theta(t)$ величины T_i имеем $F_{\theta'}(t) \leq F_\theta(t)$ при всех t . Следовательно, по лемме 1, существуют случайные величины V_i и функции f и f' ($f(v) \leq f'(v)$ при всех v) такие, что распределения величин $f(V_i)$ и $f'(V_i)$ совпадают с F_θ и $F_{\theta'}$ соответственно. Рассматриваемый последовательный критерий допускает следующее геометрическое представление на $(n, \sum_{i=1}^n h(t_i))$ -плоскости. Наблюдения продолжаются до тех пор, пока выборочная точка остается внутри полосы, образованной параллельными прямыми

$$\sum_{i=1}^n h(t_i) = \log A_j, \quad j = 0, 1.$$

Гипотеза отвергается, если путь, образованный точками

$$(1, h(t_1)), (2, h(t_1) + h(t_2)), \dots, (N, h(t_1) + \dots + h(t_N)),$$

выходит из полосы через верхнюю границу. Вероятность этого события равна, следовательно, вероятности отклонения гипотезы. Для ее подсчета при θ нужно заменить T_i на $f(V_i)$, а при θ' — T_i на $f'(V_i)$. Так как $f(V_i) \leq f'(V_i)$ при всех i , то путь, порожденный величинами $f'(V_i)$, заведомо покидает полосу через верхнюю границу, если это происходит с путем, порожденным $f(V_i)$. Следовательно, $\beta(\theta) \leq \beta(\theta')$, что и требовалось доказать.

Таким образом, в случае монотонного отношения правдоподобия последовательный критерий отношений вероятностей с вероятностями ошибок $\alpha_0 = \alpha$ и $\alpha_1 = 1 - \beta$ удовлетворяет (36). Из оптимального свойства, сформулированного в разделе 10, вытекает, что среди всех критериев, удовлетворяющих (36), последовательный критерий отношений вероятностей минимизирует средний объем выборки при $\theta = \theta_0$ и при $\theta = \theta_1$. Но теперь нас интересует $E_\theta(N)$ при всех θ . Обычно функция $E_\theta(N)$ имеет максимум между θ_0 и θ_1 и убывает, когда точка θ удаляется от точки

максимума в любом направлении. Часто оказывается, что максимум меньше n_0 — наименьшего фиксированного размера выборки, при котором существует критерий, удовлетворяющий (36). Но это не всегда так. В частности, в обстановке примера 9 при $p_0 = 0,4$; $p_1 = 0,6$; $\alpha_0 = \alpha_1 = 0,005$ фиксированный объем выборки n_0 равен 160, и хотя $E_p(N)$ для большинства значений p оказывается меньшим этой границы, при $p = 1/2$ $E_p(N)$ равно 170. Важная проблема отыскания критерия, который минимизировал бы $\sup E_\theta(N)$ при условии (36), все еще не решена.

Точное определение функции мощности $\beta(\theta)$ и среднего объема выборки $E_\theta(N)$ для последовательного критерия отношений вероятностей в общем случае представляет собой крайне трудную задачу. Однако, если уравнение

$$E_\theta \{ [p_{\theta_1}(X)/p_{\theta_0}(X)]^h \} = 1 \quad (37)$$

имеет ненулевое решение $h = h(\theta)$ (что верно при не слишком жестких ограничениях), то возможно дать простые приближенные формулы (см. задачу 38). В этом случае

$$p_\theta^*(x) = \left[\frac{p_{\theta_1}(x)}{p_{\theta_0}(x)} \right]^{h_\theta} p_\theta(x)$$

также является плотностью вероятности. Допустим теперь, что $h > 0$ (другой случай разбирается аналогично), и рассмотрим последовательный критерий отношений правдоподобия с границами A_0^h , A_1^h для проверки p_θ при альтернативе p_θ^* . При такой процедуре наблюдения продолжаются до тех пор, пока

$$A_0^h < \frac{p_\theta^*(x_1)}{p_\theta(x_1)} \dots \frac{p_\theta^*(x_n)}{p_\theta(x_n)} < A_1^h.$$

Если $\alpha_0^*(1 - \alpha_1^*)$ обозначает вероятность отклонения гипотезы при истинном распределении $p_\theta(p_\theta^*)$, то по (34) можно определить приближенные выражения для границ

$$A_0^h \sim \frac{\alpha_1^*}{1 - \alpha_0^*}, \quad A_1^h \sim \frac{1 - \alpha_1^*}{\alpha_0^*}.$$

Однако рассматриваемый критерий в точности совпадает с последовательным критерием отношений вероятностей с вероятностями ошибок $\alpha_0 = \alpha$, $\alpha_1 = 1 - \beta$ для проверки θ_0 при альтернативе θ_1 . Следовательно, α_0^* и $\beta(\theta)$ — вероятности отклонения гипотезы этими критериями при истинном распределении p_θ — должны быть равны. Разрешая вышеприведенные соотно-

шения относительно α_0^* , мы находим

$$\beta(\theta) \sim \frac{1 - A_0^h}{A_1^h - A_0^h}. \quad (38)$$

Приближенное значение для $E_\theta(N)$ может быть получено из тождества Вальда

$$E_\theta(Z_1 + \dots + Z_N) = E_\theta(N) E_\theta(Z), \quad (39)$$

которое, во всяком случае, выполняется тогда, когда величины независимы и одинаково распределены, и процедура такова, что средний объем выборки конечен. Доказательство этого соотношения смотри в задаче 37. Если величины $Z_i (i = 1, 2, \dots)$ определены равенством (35), и процедура соответствует последовательному критерию отношений вероятностей, то $Z_1 + \dots + Z_N$ может быть аппроксимирована выражением $\log A_1$, когда H отклоняется, и выражением $\log A_0$, когда H принимается. Поэтому в предположении $E_\theta(Z) \neq 0$, из (39) вытекает

$$E_\theta(N) \sim \frac{\beta(\theta) \log A_1 + [1 - \beta(\theta)] \log A_0}{E_\theta(Z)}. \quad (40)$$

Пример 10. В биномиальной схеме примера 9 уравнение (37) принимает вид

$$p \left(\frac{p_1}{p_0} \right)^h + q \left(\frac{q_1}{q_0} \right)^h = 1. \quad (41)$$

Левая часть, как функция h , является выпуклой и равна единице при $h = 0$. Отсюда видно, что, за исключением случая $p = \log(q_0/q_1)/\log(p_1q_0/p_0q_1)$ (когда левая часть минимальна при $h = 0$), уравнение (41) имеет ненулевое решение. Соотношения (38) и (41) задают приближенную функцию мощности в параметрическом виде. Ее можно вычислять, придавая h различные значения и определяя из (38) и (41) соответствующие значения p и β (для $h = 0$ β определяется по непрерывности). Ниже приведено сравнение точных и приближенных значений $\beta(p)$ и $E_p(N)$; числовые данные взяты из примера 9: $p_0 = 0,05$, $p = 0,099$, $p_1 = 0,17^*$)

$\beta(p_0)$	$\beta(p)$	$\beta(p_1)$	$E_{p_0}(N)$	$E_p(N)$	$E_{p_1}(N)$	
0,05 0,031	0,44 0,409	0,90 0,901	30 31,4	39 46,8	15 30,0	Приближенно Точно

*) Взято из уже цитированной книги Робинсона, содержащей много других примеров.

12. Оптимальное свойство последовательных критериев отношений вероятностей *)

Основная часть доказательства теоремы 8 состоит в решении следующей вспомогательной задачи. Пусть проверяется гипотеза H : плотность X есть p при альтернативе, что она есть p_1 . Пусть потери, возникающие от неправильного отклонения (принятия) H равны $w_0(w_1)$ и пусть стоимость каждого наблюдения равна c . Риск последовательной процедуры (т. е. средние потери плюс средняя стоимость наблюдений) при истинной плотности p_i равен

$$\alpha_i w_i + c E_i(N),$$

где

$$\alpha_0 = P_0 \text{ (отклонения } H), \quad \alpha_1 = P_1 \text{ (принятия } H)$$

— вероятности ошибок. Допустим, что номер i плотности также случаен и принимает значения 0 и 1 с вероятностями π и $1 - \pi$ соответственно. Тогда полный средний риск процедуры δ равен

$$r(\pi, \delta) = \pi [\alpha_0 w_0 + c E_0(N)] + (1 - \pi) [\alpha_1 w_1 + c E_1(N)]. \quad (42)$$

Мы определим байесовскую процедуру для этой задачи, т. е. процедуру, которая минимизирует (42). Здесь интерпретация (42) как байесовского риска помогает лучше понять доказательство и приводит к некоторым задачам, имеющим самостоятельный интерес. Однако, с точки зрения теоремы 8, введение величин w , c и π является просто математическим приемом, и задача состоит в минимизации формального выражения (42).

Байесовское решение содержит две постоянные $\pi' \leq \pi''$, которые однозначно выражаются через w_0 , w_1 и c с помощью уравнений (44) и (45) (см. ниже), и которые не зависят от π . Нам достаточно будет ограничиться случаем $0 < \pi' < \pi'' < 1$, предполагая также, что априорная вероятность π удовлетворяет неравенству $\pi' \leq \pi \leq \pi''$.

Лемма 5. Пусть π' и π'' удовлетворяют уравнениям (44). Если $0 < \pi' < \pi'' < 1$, тогда при всех $\pi' \leq \pi < \pi''$ байесовский риск (42) минимизируется любым последовательным критерием отношений вероятностей с границами

$$A_0 = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi''}{\pi''}, \quad A_1 = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi'}{\pi'}. \quad (43)$$

Доказательство. (1) Мы начнем с исследования того, стоит ли вообще производить наблюдения, в результате чего риск будет не меньше c , или лучше принять решение немедленно

*) В этом разделе рассматривается специальный вопрос, на который ссылок в дальнейшем не будет.

но. Пусть $\delta_0(\delta_1)$ — процедура, которая отвергает (принимает) H без наблюдений, так что

$$r(\pi, \delta_0) = \pi w_0 \quad (r(\pi, \delta_1) = (1 - \pi) w_1).$$

Пусть

$$\rho(\pi) = \inf_{\delta \in \mathcal{S}} r(\pi, \delta),$$

где $\mathcal{S}$ обозначает класс всех процедур, требующих хотя бы одного наблюдения. Тогда для любого $0 < \lambda < 1$ и любых π_0, π_1

$$\begin{aligned} \rho[\lambda\pi_0 + (1 - \lambda)\pi_1] &= \inf_{\delta \in \mathcal{S}} [\lambda r(\pi_0, \delta) + (1 - \lambda)r(\pi_1, \delta)] \geq \\ &\geq \lambda\rho(\pi_0) + (1 - \lambda)\rho(\pi_1). \end{aligned}$$

Следовательно, ρ — вогнутая функция, и так как ее значения

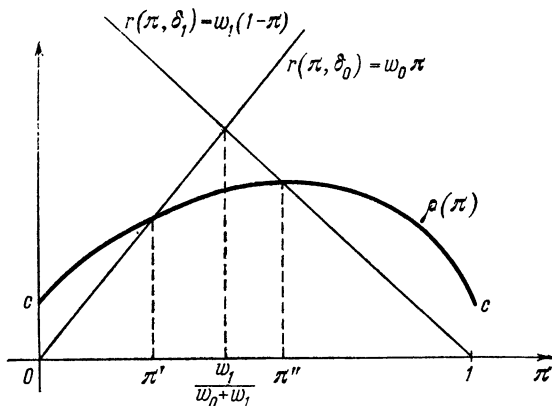


Рис. 5.

ограничены снизу нулем, то она непрерывна на интервале $(0, 1)^*$). Если

$$\rho\left(\frac{w_1}{w_0 + w_1}\right) < \frac{w_0 w_1}{w_0 + w_1},$$

то мы определим π' и π'' из уравнений

$$r(\pi', \delta_0) = \rho(\pi') \quad \text{и} \quad r(\pi'', \delta_1) = \rho(\pi'') \quad (44)$$

(см. рис. 5). В противном случае положим

$$\pi' = \pi'' = \frac{w_1}{w_0 + w_1}. \quad (45)$$

В рассматриваемом нами случае $(0 < \pi' < \pi'' < 1)$ δ_0 минимизирует (42) тогда и только тогда, когда $\pi \leq \pi'$, а δ_1

*) См., например, раздел 3.18 книги: Hardy, Littlewood, Pólya, Inequalities, Cambridge Univ. Press, 1934.

минимизирует (42) тогда и только тогда, когда $\pi \geq \pi''$. Отсюда вытекает единственный оптимальный способ поведения на первом шаге при $\pi \neq \pi'$, π'' : если $\pi < \pi'$ ($\pi > \pi''$), то наблюдения не производятся, а гипотеза H отвергается (принимается); если $\pi' < \pi < \pi''$, то производится первое наблюдение X_1 .

(2) Доказательство завершается теперь по индукции. Пусть $\pi' < \pi < \pi''$ и произведено n наблюдений с исходами $X_1 = x_1, \dots, X_n = x_n$. Мы сталкиваемся с возможностями: не брать ни одного дополнительного наблюдения, отклоняя или принимая H (с потерями ω_0 и ω_1 в случае неправильного решения), или произвести наблюдение X_{n+1} . Положение весьма сходно с проанализированным в пункте 1. Допустимо неограниченное накопление наблюдений. Тот факт, что уже израсходовано n единиц, не влияет на задачу, так как раз уже мы понесли эти потери, то никакие будущие события не могут их уменьшить. Таким образом, процедура совпадает с прежней: наблюдения не производятся, если вероятность того, что H верна, меньше π' или больше π'' , в то время как при $\pi' < \pi < \pi''$ производится дополнительное наблюдение X_{n+1} .

Наличие результатов $x_1, \dots, x_n$ уже произведенных наблюдений изменит только одну сторону ситуации. Вероятность того, что гипотеза H верна, будет теперь равна не π , а

$$\pi(x_1, \dots, x_n) = \frac{\pi p_{0n}}{\pi p_{0n} + (1 - \pi) p_{1n}},$$

т. е. условной (апостериорной) вероятности H при $X_1 = x_1, \dots, X_n = x_n$. Процедура, следовательно, требует продолжения наблюдений только в случае $\pi' < \pi(x_1, \dots, x_n) < \pi''$, т. е. в случае

$$A_0 = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi''}{\pi''} < \frac{p_{1n}}{p_{0n}} < \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi'}{\pi'} = A_1.$$

Гипотеза H принимается, если при первом нарушении этих неравенств $\frac{p_{1n}}{p_{0n}} < A_0$, и отвергается, если это отношение больше A_1 .

(3) Как показано в части (1) доказательства, на первом шаге процедуры решение однозначно определяется (δ_0 при $\pi < \pi'$, δ_1 при $\pi > \pi''$, произвести наблюдения при $\pi' < \pi < \pi''$). При $\pi = \pi'$ процедура δ_0 снова минимизирует (42), но перестает однозначно определяться этим свойством, т. е. существует процедура $\delta \in \mathcal{S}$, для которой $r(\pi', \delta) = \rho(\pi')$. Чтобы принадлежать к $\mathcal{S}$, эта процедура должна требовать по меньшей мере одного наблюдения. Коль скоро значение X_1 стало известным, то, как следует из части (2), наилучшая процедура в $\mathcal{S}$ получается, если наблюдения продолжать при

$$\pi' < \pi(x_1, \dots, x_n) < \pi''.$$

На первом шаге, следовательно, является несущественным, будут ли при пограничном значении эксперименты продолжены, или мы примем вышеуказанное решение. То же самое верно и для последующих шагов. Этим доказано, в частности, что при $\pi' \leq \pi \leq \pi''$ процедура, состоящая в проведении первого наблюдения и в следовании в дальнейшем последовательному критерию отношений вероятностей с границами (43), является байесовской.

Связь между вспомогательной и основной задачами устанавливается следующей леммой.

Лемма 6. Каковы бы ни были $0 < \pi'_0 < \pi''_0 < 1$, найдутся числа $0 < \omega < 1$ и $c > 0$ такие, что байесовское решение вспомогательной задачи с $\omega_0 = 1 - \omega$, $\omega_1 = \omega$, с и априорной вероятностью π ($\pi'_0 < \pi < \pi''_0$) дается последовательным критерием отношений вероятностей с границами

$$A_0 = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi''_0}{\pi''_0}, \quad A_1 = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi'_0}{\pi'_0}.$$

*Доказательство *).* (1) По лемме 5 величины π' и π'' — суть функции от ω и c и потому достаточно найти такие ω и c , что $\pi'(\omega, c) = \pi'_0$, $\pi''(\omega, c) = \pi''_0$. Обозначим (при фиксированном ω) $\pi'(c) = \pi'(\omega, c)$ и $\pi''(c) = \pi''(\omega, c)$. Если c_0 — наименьшее c , при котором $\pi'(c_0) = \pi''(c_0)$, то при $0 < c < c_0$ величины $\pi'(c)$ и $\pi''(c)$ определяются из уравнений

$$(1 - \omega)\pi' = \rho(\pi', c), \quad (1 - \pi'')\omega = \rho(\pi'', c),$$

где $\rho(\pi, c)$ обозначает то, что ранее мы обозначали $\rho(\pi)$. Функция $\rho(\pi', c)$, рассматриваемая как функция c при фиксированном π' , обладает следующими свойствами: (I) она непрерывна, что вытекает из ее вогнутости; (II) она строго возрастает, так как при любом $\delta \in \mathcal{E}$ риск строго возрастает вместе с c и так как минимальный риск $\rho(\pi', c)$ достигается на процедуре $\delta \in \mathcal{E}$; (III) при c , стремящемся к нулю, $\rho(\pi', c)$ и $\rho(\pi'', c)$ также стремятся к нулю. Последнее утверждение вытекает из того, что, рассматривая критерии с фиксированным объемом выборки n и беря n достаточно большим, можно сделать вероятности ошибок сколь угодно малыми.

Из этих свойств функции ρ вытекает, что при $0 < c < c_0$ функция $\pi'(\pi'')$ непрерывна, строго возрастает (убывает) и что $\pi'(c) \rightarrow 0$, $\pi''(c) \rightarrow 1$ при $c \rightarrow 0$. С другой стороны, при $c \rightarrow c_0$ $\pi''(c) - \pi'(c) \rightarrow 0$, так что обе величины стремятся к решению $\pi' = \pi'' = \omega$ уравнения $\pi'(1 - \omega) = (1 - \pi'')\omega$. Из

*) Это доказательство было сообщено мне Л. Ле-Камом.

перечисленных свойств следует, что при фиксированном ω

$$\lambda(c) = \frac{\pi'(c)}{1 - \pi'(c)} \cdot \frac{1 - \pi''(c)}{\pi''(c)}$$

является непрерывной, строго возрастающей функцией, которая изменяется от 0 до 1 при изменении c от 0 до $c_0 = c_0(\omega)$.

(2) Пусть

$$\lambda(\omega, c) = \frac{\pi'(\omega, c)}{1 - \pi'(\omega, c)} \cdot \frac{1 - \pi''(\omega, c)}{\pi''(\omega, c)}, \quad \gamma(\omega, c) = \frac{\pi''(\omega, c)}{1 - \pi''(\omega, c)}.$$

Вместо того, чтобы оперировать с переменными π' и π'' , более удобно иметь дело с λ и γ . Докажем, что существуют ω и c такие, что

$$\lambda(\omega, c) = \frac{\pi'_0}{1 - \pi'_0} \cdot \frac{1 - \pi''_0}{\pi''_0} = \lambda_0, \quad \gamma(\omega, c) = \frac{\pi''_0}{1 - \pi''_0} = \gamma_0.$$

Согласно части (1) доказательства, для любого ω существует единственное $c = c(\omega)$ такое, что $\lambda(\omega, c) = \lambda_0$. Ниже будет показано, что функция $\gamma(\omega) = \gamma[\omega, c(\omega)]$ осуществляет взаимно однозначное отображение интервала $0 < \omega < 1$ на $0 < \gamma < \infty$ и, следовательно, найдется единственное значение ω , для которого $\gamma(\omega) = \gamma_0$. Этим доказательство леммы будет завершено.

(3) По лемме 5 для вспомогательной задачи с константами $\omega, c = c(\omega)$ и $\pi = \pi'[\omega, c(\omega)]$ существует байесовское решение δ' , которое является последовательным критерием отношений вероятностей с границами

$$A'_0 = \frac{\pi'[\omega, c(\omega)]}{1 - \pi'[\omega, c(\omega)]} \cdot \frac{1 - \pi''[\omega, c(\omega)]}{\pi''[\omega, c(\omega)]} = \lambda[\omega, c(\omega)] = \lambda_0, \quad A'_1 = 1.$$

Пусть δ'' обозначает соответствующее решение при константах $\omega, c = c(\omega)$ и $\pi = \pi''[\omega, c(\omega)]$, так что для δ'' границы равны

$$A''_0 = 1, \quad A''_1 = \frac{\pi''[\omega, c(\omega)]}{1 - \pi''[\omega, c(\omega)]} \cdot \frac{1 - \pi'[\omega, c(\omega)]}{\pi'[\omega, c(\omega)]} = \frac{1}{\lambda_0}.$$

Тогда вероятности ошибок и средние размеры выборок, т. е. $\alpha'_0, \alpha'_1, E'_0(N), E'_1(N)$ для δ' и $\alpha''_0, \alpha''_1, E''_0(N), E''_1(N)$ для δ'' зависят от ω и c только через λ_0 , но не через γ . Поэтому при фиксированном λ_0 они будут фиксированными числами. Байесовские риски для $\pi = \pi'[\omega, c(\omega)]$ и $\pi = \pi''[\omega, c(\omega)]$ равны

$$\rho(\pi') = r(\pi', \delta') \quad \text{и} \quad \rho(\pi'') = r(\pi'', \delta''),$$

и из (44) следует, что

$$r(\pi', \delta_0) = r(\pi', \delta') \quad \text{и} \quad r(\pi'', \delta_1) = r(\pi'', \delta'').$$

Эти уравнения могут быть более подробно записаны в форме

$$\pi'(1 - \omega) = \pi'[\alpha'_0(1 - \omega) + cE'_0(N)] + (1 - \pi')[\alpha'_1\omega + cE'_1(N)]$$

и

$$(1 - \pi'')w = \pi''[\alpha_0''(1 - w) + cE_0''(N)] + (1 - \pi'')[\alpha_1''w + cE_1''(N)].$$

Если сюда подставить $\lambda_0\gamma$ вместо $\pi'/(1 - \pi')$ и γ вместо $\pi''/(1 - \pi'')$, а затем исключить c , то мы придем к одному уравнению, связывающему γ и w :

$$\begin{aligned} \{\lambda_0\gamma(1 - \alpha_0') - w[\lambda_0\gamma(1 - \alpha_0') + \alpha_1']\} \{ \gamma E_0''(N) + E_1''(N) \} = \\ = \{ -\gamma\alpha_0'' + w[(1 - \alpha_1'') + \gamma\alpha_0''] \} \cdot \{ \lambda_0\gamma E_0'(N) + E_1'(N) \}. \end{aligned}$$

Оно линейно относительно w и при каждом $\gamma > 0$ имеет решение $0 < w < 1$. Как функция γ оно является квадратным трехчленом, причем коэффициент при квадрате и свободный член имеют при $0 < w < 1$ противоположные знаки. Следовательно, в этом случае существует единственное положительное решение γ , которое и устанавливает желаемое взаимно однозначное соответствие между γ и w .

Чтобы закончить доказательство теоремы 8, рассмотрим теперь какой-либо последовательный критерий отношений вероятностей с $A_0 < 1 < A_1$ и любой константой $0 < \pi < 1$. Пусть

$$\pi' = \frac{\pi}{A_1(1 - \pi) + \pi}, \quad \pi'' = \frac{\pi}{A_0(1 - \pi) + \pi}.$$

Эти значения удовлетворяют (43) и $0 < \pi' < \pi < \pi'' < 1$. По лемме 6 найдутся, следовательно, константы $0 < w < 1$ и $c > 0$ такие, что данный критерий оказывается байесовским решением вспомогательной задачи, в которой истинная плотность равна p_0 с вероятностью π , потери равны $w_0 = 1 - w$ и $w_1 = w$ и стоимость одного наблюдения — c . Пусть $\alpha_0, \alpha_1, E_0(N), E_1(N)$ будут вероятностями ошибок и средними объемами выборок для данного критерия. Рассмотрим любую конкурирующую процедуру δ^* с вероятностями ошибок $\alpha_i^* \leq \alpha_i$ и размерами выборок $E_i^*(N) < \infty$ ($i = 0, 1$). Так как данный критерий минимизирует байесовский риск, то для него

$$\begin{aligned} \pi[(1 - w)\alpha_0 + cE_0(N)] + (1 - \pi)[w\alpha_1 + cE_1(N)] \leq \\ \leq \pi[(1 - w)\alpha_0^* + cE_0^*(N)] + (1 - \pi)[w\alpha_1^* + cE_1^*(N)] \end{aligned}$$

и, следовательно,

$$\pi E_0(N) + (1 - \pi)E_1(N) \leq \pi E_0^*(N) + (1 - \pi)E_1^*(N).$$

Справедливость этого неравенства при всех $0 < \pi < 1$ влечет

$$E_0(N) \leq E_0^*(N) \quad \text{и} \quad E_1(N) \leq E_1^*(N),$$

что и требовалось доказать.

13. Задачи

К разделу 2

1. РНМ критерий для $R(0, \theta)$. Пусть $X = (X_1, \dots, X_n)$ — выборка из распределения, равномерного на $(0, \theta)$.

(I). Для проверки $H: \theta \leq \theta_0$ при альтернативе $K: \theta > \theta_0$ РНМ уровня α будет любой критерий, для которого $E_{\theta_0} \phi(X) = \alpha$, $E_{\theta} \phi(X) \leq \alpha$ при $\theta \leq \theta_0$ и $\phi(x) = 1$, когда $\max(x_1, \dots, x_n) > \theta_0$.

(II) Для проверки $H: \theta = \theta_0$ при альтернативе $K: \theta \neq \theta_0$ существует единственный РНМ критерий. Для него $\phi(x) = 1$ при $\max(x_1, \dots, x_n) > \theta_0$

или $\max(x_1, \dots, x_n) \leq \theta_0 \sqrt[n]{\alpha}$ и $\phi(x) = 0$ в других случаях.

[(II) Определите РНМ критерий для проверки $\theta = \theta_0$ при альтернативе $\theta < \theta_0$ и соедините этот результат с (I).]

2. РНМ критерий для экспоненциальных семейств. Пусть $X_1, \dots, X_n$ — выборка из экспоненциального распределения с плотностью $ae^{-a(x-b)}$, $x \geq b$.

(I) Определите РНМ критерий для проверки $H: b = b_0$ при альтернативе $K: b \neq b_0$, если a известно.

(II) Определите РНМ критерий для проверки $H: a = a_0, b = b_0$ при альтернативах $a > a_0, b < b_0$. Объясните, с чем связано (весьма неожиданное) существование в этом случае РНМ критерия для двухпараметрической проблемы.

[(I) Случайные величины $Y_i = e^{-aX_i}$ образуют выборку из равномерного на $(0, e^{-ab})$ распределения.]

3. Если выборочное пространство $\mathcal{X}$ евклидово и P_0 и P_1 имеют плотности по отношению к лебеговской мере, то существует нерандомизированный наиболее мощный критерий проверки P_0 при альтернативе P_1 при любом уровне значимости α^* .

[Это вытекает из теоремы 1 и следующей леммы**): пусть $f \geq 0$

и $\int_A f(x) dx = a$; тогда при любом $0 \leq b \leq a$ существует подмножество B

множества A такое, что $\int_B f(x) dx = b$.]

4. *Вполне информирующие статистики.* Статистика T называется *вполне информирующей*, если в любой проблеме решения процедуры, основанные только на T , образуют существенно полный класс. Если семейство $\mathcal{P}$ доминировано и статистика T вполне информирующая, то T достаточна.

[Рассмотрим любую пару распределений $P_0, P_1 \in \mathcal{P}$ с плотностями p_0, p_1 , и пусть $g_i = p_i/(p_0 + p_1)$. Предположим, что T — вполне информирующая статистика. Обозначим $\mathcal{A}_0$ подполе, индуцированное T . Тогда $\mathcal{A}_0$ содержит подполе, порожденное парой (g_0, g_1) , так как оно содержит каждую критическую область, которая является (единственной) наиболее мощной для проверки P_0 при альтернативе P_1 (или P_1 при альтернативе P_0) для какого-либо уровня α .

*) О более общих результатах, касающихся возможности обойтись без рандомизированных процедур, см. Dvoretzky, Wald and Wolfowitz, Elimination of randomization in certain statistical decision procedures and zero-sum two-person games, Ann. Math. Stat. 22 (1951), 1—21.

**) Доказательство этой теоремы см. в книге: П. Халмош, Теория меры, ИЛ, 1953, 171. Лемма является частным случаем теоремы Ляпунова, см. «О вполне аддитивных вектор-функциях», Изд. АН, сер. матем., 4 (1940), 465—478.

Следовательно, T достаточна для каждой пары распределений и потому (см. задачу 9 в главе 2) достаточна для $\mathcal{P}$.]

К разделу 3

5. Пусть X — число успехов в n независимых испытаниях с вероятностью успеха p . Пусть $\phi(x)$ — РНМ критерий (9) для проверки $p \leq p_0$ при альтернативе $p > p_0$ с уровнем значимости α .

(I) Для $n = 6$, $p_0 = 0,25$ и уровней $\alpha = 0,05; 0,1; 0,2$ определите C и γ и найдите мощность критерия при альтернативах $p_1 = 0,3; 0,4; 0,5; 0,6; 0,7$.

(II) Допустим, что при $p = 0,2$ и $\alpha = 0,05$ желательно построить критерий мощности $\beta \geq 0,9$ при альтернативе $p_1 = 0,4$. Определить необходимый объем выборки: (а) используя таблицы биномиального распределения, (б) используя нормальное приближение*).

(III) Используйте нормальное приближение для определения необходимого объема выборки при $\alpha = 0,05; \beta = 0,9; p_0 = 0,01; p_1 = 0,02$.

6. (I) Допустим, что существует смешанная вторая производная $\partial^2 \log p_\theta(x) / \partial \theta \partial x$. Плотности $p_\theta(x)$ имеют монотонное относительно x отношение правдоподобия тогда и только тогда, когда эта производная ≥ 0 при всех θ и x .

(II) Равносильное условие состоит в том, что

$$p_\theta(x) \frac{\partial^2 p_\theta(x)}{\partial \theta \partial x} \geq \frac{\partial p_\theta(x)}{\partial \theta} \frac{\partial p_\theta(x)}{\partial x} \quad \text{при всех } \theta \text{ и } x.$$

7. Пусть плотности p_θ случайной величины X имеют монотонное относительно $T(x)$ отношение правдоподобия. Рассмотрим проблему проверки $H: \theta \leq \theta_0$ при альтернативе $\theta > \theta_0$. Если распределение T непрерывно, то критический уровень $\hat{\alpha}$ определяется равенством $\hat{\alpha} = P_{\theta_0}\{T \geq t\}$, где t — наблюдаемое значение T . Это верно и без предположения непрерывности, если определить $\hat{\alpha}$ для рандомизированного критерия, как наименьший уровень значимости, при котором гипотеза отвергается с вероятностью 1.

8. Пусть $X_1, \dots, X_n$ независимы и имеют плотность $(2\theta)^{-1} e^{-x/2\theta}$, $x \geq 0$, и пусть $Y_1 \leq \dots \leq Y_n$ обозначают X_i ($i = 1, 2, \dots, n$), расположенные в порядке возрастания. Допустим, что в процессе наблюдений мы обнаруживаем сначала Y_1 , затем Y_2 , потом Y_3 и т. д. до Y_r . На основе $Y_1, \dots, Y_r$ необходимо проверить гипотезу $H: \theta > \theta_0 = 1000$ с уровнем значимости $\alpha = 0,05$ при альтернативе $\theta < \theta_0$.

(I) Определите критическую область при $r = 4$ и найдите мощность критерия при альтернативе $\theta_1 = 500$.

(II) Найдите значение r , необходимое для достижения мощности $\beta \geq 0,95$ при этой альтернативе.

[В задаче 13 главы 2 установлено, что величина $\left[\sum_{i=1}^r Y_i + (n-r) Y_r \right] / \theta$

имеет χ^2 -распределение с $2r$ степенями свободы.]

9. В процессе Пуассона, наблюдаемом в промежутке времени длины t , число появившихся событий имеет распределение Пуассона $P(\lambda t)$. В другой схеме процесс наблюдается до момента появления r -го события; при этом время наблюдения T является случайной величиной, и $2\lambda T$ имеет распределение χ^2 с $2r$ степенями свободы. Применяя для проверки гипотезы $H: \lambda \leq \lambda_0$ с уровнем α любую из этих схем, мы можем, выбирая t и r достаточно боль-

*) Обсуждение другого удобного метода, применимого к этой и близким проблемам, см. в работе: Mosteller and Tukey, The uses and usefulness of binomial probability paper, J. Am. Stat. Assoc. 44 (1949), 174—212.

шими, достичь любой наперед заданной мощности β по отношению к альтернативе λ_1 .

(I) Отношение необходимого времени наблюдения для первой схемы к необходимому в среднем времени для второй равно $\lambda t/r$.

(II) Определите, для каких значений λ какая схема предпочтительнее, если $\lambda_0 = 1$, $\lambda_1 = 2$, $\alpha = 0,05$, $\beta = 0,9$.

10. *Обобщение леммы 2.* Пусть P_0 и P_1 — два распределения с плотностями p_0 и p_1 такими, что отношение $p_1(x)/p_0(x)$ является неубывающей функцией от действительной статистики $T(x)$.

(I) Если плотность T при P_i равна p'_i , то $p'_1(t)/p'_0(t)$ — неубывающая функция t .

(II) $E_0\psi(T) \leq E_1\psi(T)$ для любой неубывающей функции ψ .

(III) Если $p_1(x)/p_0(x)$ — строго возрастающая функция $t = T(x)$, то такова же $p'_1(t)/p'_0(t)$ и $E_0\psi(T) < E_1\psi(T)$, исключая случаи, когда $\psi[T(x)]$ постоянна ($P_0 + P_1$)-почти всюду или когда $E_0\psi(T) = E_1\psi(T) = \pm\infty$.

(IV) Для любых двух различных распределений с плотностями p_0 , p_1

$$-\infty \leq E_0 \log \left[\frac{p_1(X)}{p_0(X)} \right] < E_1 \log \left[\frac{p_1(X)}{p_0(X)} \right] \leq \infty.$$

[(I) Предположим, не ограничивая общности, что $p_1(x)/p_0(x) = T(x)$. Тогда для любой интегрируемой φ

$$\int \varphi(t) p'_1(t) d\nu(t) = \int \varphi[T(x)] T(x) p_0(x) d\mu(x) = \int \varphi(t) t p'_0(t) d\nu(t),$$

и, следовательно, $p'_1(t)/p'_0(t) = t$ почти всюду.

(IV) Возможность $E_0 \log [p_1(X)/p_0(X)] = \infty$ исключается, так как в силу выпуклости логарифмической функции,

$$E_0 \log \left[\frac{p_1(X)}{p_0(X)} \right] \leq \log E_0 \left[\frac{p_1(X)}{p_0(X)} \right] = 0.$$

То же верно и для E_1 . Строгое неравенство вытекает из (III) с $T(x) = p_1(x)/p_0(x)$.

11. Пусть F_0 и F_1 — две функции распределения на действительной прямой такие, что при всех x $F_1(x) \leq F_0(x)$. Тогда $E_0\psi(X) \leq E_1\psi(X)$ для любой неубывающей функции ψ .

К разделу 4

12. Если эксперимент (f, g) лучше информирует, чем (f', g') , то (g, f) лучше информирует, чем (g', f') .

13. *Условия, обеспечивающие сравнимость.* (I) Пусть X и X' — две случайные величины, принимающие значения 1 и 0. Предположим, что $P\{X=1\} = p_0$, $P\{X'=1\} = p'_0$ или $P\{X=1\} = p_1$, $P\{X'=1\} = p'_1$. Допустим, не ограничивая общности, что $p_0 < p'_0$, $p_0 < p_1$, $p'_0 < p'_1$ (этого можно достичь перестановкой случайных величин X и X' и значений 0 и 1 у одной или обеих случайных величин). Тогда X лучше информирует, чем X' в том и только в том случае, когда $(1 - p_1)(1 - p'_0) \leq (1 - p_0)(1 - p'_1)$.

(II) Пусть U_0 и U_1 независимы и распределены равномерно на отрезке $(0, 1)$, и пусть $Y = 1$, если $X = 1$, $U_1 \leq \gamma_1$ или если $X = 0$, $U_0 \leq \gamma_0$ и $Y = 0$ в других случаях. В предположениях пункта (I) существуют $0 \leq \gamma_0, \gamma_1 \leq 1$, для которых $P\{Y=1\} = p'_i$ при $P\{X=1\} = p_i$ ($i = 0, 1$) в предположении, что $(1 - p_1)(1 - p'_0) \leq (1 - p_0)(1 - p'_1)$. Это неравенство, являющееся, как

мы видим, достаточным для того, чтобы выборка $X_1, \dots, X_n$ из X лучше информировала, чем выборка $X'_1, \dots, X'_n$ из X' , оказывается и необходимым. Аналогично условие $p'_0 p_1 \leq p_0 p'_1$ необходимо и достаточно для того, чтобы выборка из X' лучше информировала, чем выборка из X .

[(I) Мощность $\beta(\alpha)$ основанного на X наиболее мощного критерия уровня α для проверки p_0 при альтернативе p_1 равна $\alpha p_1 / p_0$ для $\alpha \leq p_0$ и $p_1 + q_1 q_0^{-1} (\alpha - p_0)$ для $p_0 \leq \alpha$. Сравнивая графики $\beta(\alpha)$ и $\beta'(\alpha)$, мы приходим к искомому результату.

(II) Вторая часть пункта (II) вытекает из сравнения при α , близких к 1, мощностей $\beta_n(\alpha)$ и $\beta'_n(\alpha)$ наиболее мощных критериев уровня α , основанных на ΣX_i и $\Sigma X'_i$ соответственно. Двойственное условие получается из задачи 12.]

1. Для случая 2×2 таблицы, описанной в примере 4 при сделанном там предположении $p \leq p \leq 1/2$, выборка из B лучше информирует, чем выборка из $\bar{A}$. С другой стороны, выборки из B и $\bar{B}$ несравнимы.

[Необходимое и достаточное условие сравнимости дано в предыдущей задаче.]

15. В обстановке примера 5 производится n биномиальных испытаний с вероятностью успеха $p = 1 - e^{-\lambda v}$. Необходимо проверить $\lambda = \lambda_0$ при альтернативе $\lambda = \lambda_1$. Эксперименты, отвечающие различным значениям v , несравнимы.

К разделу 5

16. (I) Для $n = 5, 10$ и $1 - \alpha = 0,95$ начертите верхние доверительные пределы $\bar{p}$ и $\bar{p}^*$ примера 7 как функции от $t = x + u$.

(II) Для тех же значений n и $\alpha_1 = \alpha_2 = 0,05$ начертите нижний и верхний доверительные пределы $\underline{p}$ и $\bar{p}$.

17. *Доверительные границы с минимальным риском.* Допустим, что функция $L(\theta, \underline{\theta})$ неотрицательна, не возрастает по второму аргументу при $\underline{\theta} < \theta$ и равна 0 для $\underline{\theta} \geq \theta$. Если $\underline{\theta}$ и $\underline{\theta}^*$ — две нижние доверительные границы для θ такие, что

$$P_{\theta} \{ \underline{\theta} \leq \theta' \} \leq P_{\theta} \{ \underline{\theta}^* \leq \theta' \} \quad \text{при всех } \theta' \leq \theta,$$

то

$$E_{\theta} L(\theta, \underline{\theta}) \leq E_{\theta} L(\theta, \underline{\theta}^*).$$

[Определите функции распределения F и F^* равенствами $F(u) = P_{\theta} \{ \underline{\theta} \leq u \} / P_{\theta} \{ \underline{\theta}^* \leq u \}$; $F^*(u) = P_{\theta} \{ \underline{\theta}^* \leq u \} / P_{\theta} \{ \underline{\theta}^* \leq \theta \}$ при $u < \theta$ и $F(u) = F^*(u) = 1$ при $u \geq \theta$. Тогда при всех u $F(u) \leq F^*(u)$ и из задачи 11 вытекает, что

$$E_{\theta} [L(\theta, \underline{\theta})] \leq P_{\theta} \{ \underline{\theta}^* = \theta \} \int L(\theta, u) dF(u) \leq \\ \leq P_{\theta} \{ \underline{\theta}^* \leq \theta \} \int L(\theta, u) dF^*(u) = E_{\theta} [L(\theta, \underline{\theta}^*)].]$$

К разделу 6

18. Обозначим $\beta(\theta)$ функцию мощности РНМ критерия следствия 2. Если функция Q в (12) дифференцируема, то $\beta'(\theta) > 0$ для всех θ , для которых $Q'(\theta) > 0$.

[Чтобы показать, что $\beta'(\theta_0) > 0$, рассмотрите проблему максимизации производной $\beta'(\theta_0)$ (или, что то же самое, величины $E_{\theta_0} T[(X) \varphi(X)]$ при условии $E_{\theta_0} \varphi(X) = \alpha$.]

19. *Оптимальная процедура отбора.* Каждый элемент некоторой совокупности измеряется n раз. Пусть $(X_1, \dots, X_n) = X$ — результаты измерений, Это

могут быть, например, итоги n тестов проверки способностей, которые проводятся для отбора кандидатов на обучение по некоторой программе. Результат будущего измерения Y (такого, как итог проверки по завершении программы) весьма интересен, но не доступен. Совместное распределение X и Y предполагается данным.

(I) Желательно отобрать заданную пропорцию α кандидатов таким образом, чтобы сделать максимальным математическое ожидание Y для этой группы. Этого можно достигнуть, отбирая тех кандидатов, для которых $E(Y|x) \geq C$, где C определяется условием: кандидат отбирается с вероятностью α . При $E(Y|x) = C$ может оказаться необходимой рандомизация (чтобы получить вероятность, в точности равную α).

(II) При иной постановке, когда требуется максимизировать вероятность того, что Y для отобранных кандидатов превзойдет заданный уровень y_0 , следует выбирать тех, для кого условная вероятность $P\{Y \geq y_0|x\}$ достаточно велика.

[(I) Пусть $\varphi(x)$ обозначает вероятность отбора кандидата с характеристикой x . Задача состоит в максимизации

$$\int \left[\int y p^{Y|x}(y) \varphi(x) dy \right] p^X(x) dx$$

при условии

$$\int \varphi(x) p^X(x) dx = \alpha.$$

20. Нижеследующий пример показывает, что следствие 4 не распространяется на счетные (бесконечные) семейства распределений. Пусть p_n — равномерная плотность на $[0, 1 + 1/n]$ и p_0 — равномерная плотность на $[0, 1]$.

(I) Тогда p_0 линейно независимо от $(p_1, p_2, \dots)$, т. е. не существует таких констант $c_1, c_2, \dots$, что $p_0 = \sum c_n p_n$.

(II) Не существует критерия φ такого, что

$$\int \varphi p_n = \alpha \text{ при } n = 1, 2, \dots, \text{ но } \int \varphi p_0 > \alpha.$$

21. Пусть $F_1, \dots, F_{m+1}$ — действительные функции, определенные на пространстве U . Чтобы u_0 максимизировало F_{m+1} при условиях $F_i(u) \leq c_i$ ($i = 1, \dots, m$), достаточно, чтобы оно максимизировало $F_{m+1}(u) - \sum k_i F_i(u)$ при некоторых константах $k_i \geq 0$, причем для i с $k_i > 0$ должно быть $F_i(u_0) = c_i$.

К разделу 7

22. Для случайной величины X с биномиальным распределением $b(p, n)$ определите константы C_i, γ_i ($i = 1, 2$) в РНМ критерии (24) для проверки гипотезы $H: p \leq 0,2$ или $p \geq 0,7$, когда $\alpha = 0,1$ и $n = 15$. Найдите мощность критерия относительно альтернативы $p = 0,4$.

23. Классы типа Поля. Семейство распределений с плотностями $p_\theta(x)$, которые непрерывны по действительным аргументам θ и x , относится к типу Поля, если при всех $x_1 < \dots < x_n$ и $\theta_1 < \dots < \theta_n$

$$\Delta_n = \begin{vmatrix} p_{\theta_1}(x_1) & \dots & p_{\theta_1}(x_n) \\ \vdots & & \vdots \\ p_{\theta_n}(x_1) & \dots & p_{\theta_n}(x_n) \end{vmatrix} \geq 0 \text{ для } n = 1, 2, \dots \quad (46)$$

Оно относится к типу Поля в строгом смысле, если в (46) имеют место строгие неравенства. При $n = 1$ из (46) выводим, что $p_\theta(x) \geq 0$, при $n = 2$, — что $p_\theta(x)$ имеет монотонное отношение правдоподобия. Экспоненциальные семейства (12) с $T(x) = x$ и $Q(\theta) = \theta$ относятся к типу Поля в строгом смысле.

[Положительность определителя $|e^{\theta_i x_j}|$, $i, j = 1, \dots, n$, может быть доказана по индукции. Разделим i -й столбец на $e^{\theta_1 x_i}$, $i = 1, \dots, n$, вычтем в полученном определителе $(n-1)$ -й столбец из n -го, $(n-2)$ -й из $(n-1)$ -го, $\dots$, 1-й из 2-го. Разложим полученный определитель по первой строке. Тогда мы увидим, что Δ_n имеет тот же самый знак, что и

$$\Delta'_n = |e^{\eta_i x_j} - e^{\eta_i x_{j-1}}|, \quad i, j = 2, \dots, n,$$

где $\eta_i = \theta_i - \theta_1$. Если разложить этот определитель по элементам первого столбца, то получится сумма вида

$$a_2 (e^{\eta_2 x_2} - e^{\eta_2 x_1}) + \dots + a_n (e^{\eta_n x_2} - e^{\eta_n x_1}) = \\ = h(x_2) - h(x_1) = (x_2 - x_1) h'(y_2)$$

где $x_1 \leq y_2 \leq x$. Переписав $h'(y_2)$ в форме детерминанта, у которого все столбцы, кроме первого, совпадают со столбцами Δ'_n , и поступая аналогично с последующими столбцами, мы придем к определителю $|e^{\eta_i y_j}|$, $i, j = 2, \dots, n$, который положителен по предположению индукции.]

24. *Тип Поля 3.* Пусть θ и x означают действительные переменные, и допустим, что плотности $p_\theta(x)$ таковы, что $p_{\theta'}(x)/p_\theta(x)$ строго возрастает по x при $\theta < \theta'$. Тогда приводимые ниже два условия эквивалентны.

(а) Для $\theta_1 < \theta_2 < \theta_3$ и $k_1, k_2, k_3 > 0$ положим

$$g(x) = k_1 p_{\theta_1}(x) - k_2 p_{\theta_2}(x) + k_3 p_{\theta_3}(x).$$

Если $g(x_1) = g(x_3) = 0$, то функция g положительна вне интервала (x_1, x_3) и отрицательна внутри его.

(б) Определитель Δ_3 , задаваемый (46), положителен при всех $\theta_1 < \theta_2 < \theta_3$, $x_1 < x_2 < x_3$. (Из (а) следует, что уравнение $g(x) = 0$ имеет не более двух решений.)

[Что (б) влечет (а) можно установить, рассматривая при $x_1 < x_2 < x_3$ определитель

$$\begin{vmatrix} g(x_1) & g(x_2) & g(x_3) \\ p_{\theta_2}(x_1) & p_{\theta_2}(x_2) & p_{\theta_2}(x_3) \\ p_{\theta_3}(x_1) & p_{\theta_3}(x_2) & p_{\theta_3}(x_3) \end{vmatrix}.$$

Обратно, предположим, что (а) выполняется. Монотонность отношений правдоподобия влечет, что ранг Δ_3 не менее двух, так что существуют константы k_1, k_2, k_3 , для которых $g(x_1) = g(x_3) = 0$. Положительность чисел k вытекает из монотонности отношений правдоподобия.]

25. *Обобщение теоремы 6.* Заключение теоремы 6 остается верным, если плотности p_θ достаточной статистики T (которую, не ограничивая общности, мы примем равной X) удовлетворяют следующим условиям: (а) $p_\theta(x)$ непрерывна по x при каждом θ ; (б) $p_{\theta'}(x)/p_\theta(x)$ строго возрастает по x при $\theta < \theta'$ и определитель Δ_3 из формулы (46) положителен при всех $\theta_1 < \theta_2 < \theta_3$ и $x_1 < x_2 < x_3$.

[При доказательстве теоремы 6 были использованы два свойства экспоненциальных семейств, а именно непрерывность по x и свойство из пункта (а) предыдущей задачи.]

26. Для проверки гипотезы $H': \theta_1 \leq \theta \leq \theta_2$ ($\theta_1 \leq \theta_2$) при альтернативах $\theta < \theta_1$ или $\theta > \theta_2$ или гипотезы $\theta = \theta_0$ при альтернативах $\theta \neq \theta_0$ в предположении, что семейство экспоненциально (или, в общем случае, удовлетворяет условиям задачи 25), не существует РНМ критерия.

Это вытекает из рассмотрения РНМ критериев для односторонних гипотез $H_1: \theta \geq \theta_1$ и $H_2: \theta \leq \theta_2$.]

К разделу 8

27. Пусть случайные величины X_i ($i = 1, \dots, s$) независимы и распределены по закону Пуассона $P(\lambda_i)$. Для проверки гипотезы $H: \Sigma \lambda_j \leq a$ (например, гипотезы о том, что общая радиоактивность нескольких образцов радиоактивного вещества не превосходит a) существует РНМ критерий, отвергающий ее при $\Sigma X_j > C$.

[Разложите совместное распределение X_i в произведение частного распределения ΣX_j (распределение Пуассона со средним $\Sigma \lambda_j$) и условного распределения величин $Y_i = X_i / \Sigma X_j$ при данной сумме ΣX_j (полиномиальное с вероятностями $p_i = \lambda_i / \Sigma \lambda_j$); дальнейшие рассуждения те же, что и в примере 8.]

28. *Доверительные границы для медианы* Пусть $X_1, \dots, X_n$ — выборка из распределения с непрерывной функцией распределения F . Обозначим ξ медиану F , если она единственна, или, в общем случае, $\xi = \inf\{\xi' : F(\xi') = 1/2\}$.

(I) Пусть X_i расположены в порядке возрастания $X^{(1)} < \dots < X^{(n)}$. Равномерно наиболее точная нижняя доверительная граница для ξ определяется соотношениями $\underline{\xi} = X^{(k)}$ с вероятностью ρ , $\underline{\xi} = X^{(k+1)}$ — с вероятностью $(1 - \rho)$, где k и ρ находятся из уравнения

$$\rho \sum_{j=k}^n \binom{n}{j} \frac{1}{2^n} + (1 - \rho) \sum_{j=k+1}^n \binom{n}{j} \frac{1}{2^n} = 1 - \alpha.$$

(II) Эта граница имеет доверительный коэффициент $1 - \alpha$ для любой медианы F .

(III) Найдите наиболее точную нижнюю доверительную границу для 100ρ -процентной точки ξ распределения F , определяемой равенством $\xi = \inf\{\xi' : F(\xi') = \rho\}$.

[При фиксированном ξ_0 проблема проверки $H: \xi = \xi_0$ при альтернативе $K: \xi > \xi_0$ эквивалентна проверке $H': p = 1/2$ при альтернативе $K': p < 1/2$.]

29. *Противоречащий пример.* Как правило, при меняющемся α наиболее мощные критерии уровня α для проверки гипотезы H при простой альтернативе устроены так, что соответствующие критические области R_α вложены одна в другую $R_\alpha \subset R_{\alpha'}$ при любом $\alpha < \alpha'$. Это соотношение всегда выполняется, если гипотеза H простая, но, как показывает приводимый ниже пример, этого может не быть при сложной H .

	1	2	3	4
P_0	$\frac{2}{13}$	$\frac{4}{13}$	$\frac{3}{13}$	$\frac{4}{13}$
P_1	$\frac{4}{13}$	$\frac{2}{13}$	$\frac{1}{13}$	$\frac{6}{13}$
Q	$\frac{4}{13}$	$\frac{3}{13}$	$\frac{2}{13}$	$\frac{4}{13}$

равно P_0 или P_1 при альтернативе Q и при уровне $\alpha = 5/13$, отклоняет гипотезу, когда $X = 1$ или $X = 3$; если же уровень $\alpha = 6/13$, то при $X = 1$ или $X = 2$.

30. Пусть X и Y обозначают число успехов в двух сериях по n биномиальных испытаний в каждом и с вероятностями успеха p_1 и p_2 соответственно.

(I) При уровне $\alpha < 1/2$ наиболее мощный критерий проверки гипотезы $H: p_2 \leq p_1$ при альтернативе (p'_1, p'_2) с $p'_1 < p'_2$ и $p'_1 + p'_2 = 1$ отклоняет гипотезу всегда, когда $Y - X > C$, и с вероятностью γ , когда $Y - X = C$.

(II) Этот критерий не является РНМ при альтернативе $p_1 < p_2$.

[(I) Возьмите в качестве априорного распределения на H распределение λ , приписывающее вероятность 1 точке $p_1 = p_2 = 1/2$. Наиболее мощный критерий при альтернативе (p'_1, p'_2) совпадает тогда с предложенным выше. Можно показать, что λ является наименее благоприятным. Рассмотрим вероятность $\beta(p_1, p_2)$ отклонения гипотезы для $p_1 = p_2 = p$. По симметрии

$$2\beta(p, p) = P\{|Y - X| > C\} + \gamma P\{|Y - X| = C\}.$$

Пусть $X_i(Y_i)$ равно 1 или 0 в зависимости от того, закончилось i -е испытание первой (второй) серии успехом или нет. Тогда $Y - X = \sum_{i=1}^n (Y_i - X_i)$, и тот факт, что $2\beta(p, p)$ достигает максимума при $p = 1/2$, может быть доказан индукцией по n .

(II) Так как $\beta(p, p) < \alpha$ при $p \neq 1/2$, то мощность $\beta(p_1, p_2)$ будет $< \alpha$ для альтернатив $p_1 < p_2$, достаточно близких к отрезку $p_1 = p_2$. То, что предложенный критерий не является РНМ, следует теперь из сравнения с $\varphi(x, y) \equiv \alpha$.]

31. Достаточные статистики при «мешающих» параметрах.

(I) Статистика T называется достаточной для θ при наличии «мешающего» параметра η , если параметрическое пространство является прямым произведением множеств возможных значений θ и η и если выполняются следующие два условия: (а) условное распределение при $T = t$ зависит только от η ; (б) частное распределение T зависит только от θ . Если эти условия выполнены, то существует зависящий только от T РНМ критерий проверки сложной гипотезы $H: \theta = \theta_0$ при сложной альтернативе $\theta = \theta_1$.

(II) Часть (I) указывает новое доказательство того, что критерий примера 8 является РНМ.

[Пусть $\varphi_0(t)$ — наиболее мощный критерий уровня α для проверки θ_0 при альтернативе θ_1 , зависящий только от t . Пусть $\varphi(x)$ — любой критерий уровня α и $\psi(t) = E\eta_1[\varphi(X)|t]$. Так как $E_{\theta_1}\psi(T) = E_{\theta_1}\eta_1\varphi(X)$, то ψ — критерий уровня α для H , и потому его мощность, а вместе с тем и мощность φ не превосходит мощности φ_0 .]

К разделу 9

32. Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — две независимые выборки из $N(\xi, 1)$ и $N(\eta, 1)$. Рассмотрим гипотезу $H: \eta \leq \xi$ при альтернативе $K: \eta > \xi$. Тогда существует РНМ критерий, отвергающий гипотезу, если разность $\bar{Y} - \bar{X}$ слишком велика.

[Пусть $\xi_1 < \eta_1$ — специальная альтернатива. Тогда распределение, приписывающее вероятность 1 точке $\eta = \xi = (m\xi_1 + n\eta_1)/(m+n)$, является наименее благоприятным.]

33. Пусть $X, \dots, X_m, Y_1, \dots, Y_n$ независимы и распределены нормально со средними ξ и η и дисперсиями σ^2 и τ^2 соответственно. Рассмотрим гипотезу $H: \tau \leq \sigma$ при альтернативе $K: \sigma < \tau$.

(I) Если ξ и η известны, то существует РНМ критерий с критической областью $\Sigma(Y_1 - \eta)^2 / \Sigma(X_i - \xi)^2 \geq C$.

II) РНМ критерий не существует при неизвестных ξ и η .

К разделу 10

34. Распределение объема последовательной выборки. Пусть X_i ($i = 1, 2, \dots$) независимы и одинаково распределены, и пусть Z_i определены по (35). Обозначим N число наблюдений до принятия окончательного решения в последовательном критерии отношений вероятностей (33) с $A_0 < 1 < A_1$. Если

истинное распределение X таково, что $P\{Z=0\} < 1$, то найдутся $0 < \delta < 1$ и $C > 0$, для которых $P\{N \geq n\} \leq C\delta^n$.

[Обозначим $c = \log A_1 - \log A_0$ и допустим сначала, что $P\{|Z| \leq c\} = p < 1$. Из события $N \geq n$ вытекает, что

$$\log A_0 < z_1, \quad z_1 + z_2, \quad \dots, \quad z_1 + \dots + z_{n-1} < \log A_1,$$

и, следовательно, что величины $|z_1|, |z_2|, \dots, |z_{n-1}|$ все $\leq c$. Тогда $P\{N \geq n\} \leq p^{n-1} = p^{-1} \cdot p^n$. Если $P\{|Z| \leq c\} = 1$, то существует такое r , что

$$P\{|Z_1 + \dots + Z_r| \leq c\} = p < 1,$$

откуда следует, что $P\{N \geq rm\} \leq p^{m-1}$, и потому

$$P\{N \geq n\} \leq p^{\lfloor n/r \rfloor - 1} \leq p^{n/r-2} = p^{-2} (p^{1/2})^n.]$$

35. Моменты объема последовательной выборки. В предположениях предыдущей задачи (I) $E(N) < \infty$ и (II) при некотором $t > 0$ $E(e^{tN}) < \infty$, так что $E(N^k) < \infty$ при всех $k = 1, 2, \dots$. Кроме того, (III) $P\{Z=0\} < 1$, если истинное распределение X совпадает с P_0 или P_1 и $P_0 \neq P_1$.

$$[(I) \quad E(N) = \sum_{n=1}^{\infty} nP\{N=n\} = \sum_{n=1}^{\infty} P\{N \geq n\} \leq C \sum_{n=1}^{\infty} \delta^n < \infty.$$

$$(II) \quad E(e^{tN}) \leq \sum e^{tn} P\{N \geq n\} \leq C \sum (\delta e^t)^n < \infty \text{ в предположении } e^t < \delta^{-1}.]$$

К разделу 11

36. Мощность последовательного критерия отношений вероятностей в биномиальном случае. В примерах 9 и 10 последовательный критерий отношений вероятностей применялся к проверке гипотезы: вероятность успеха равна p_0 при альтернативе p_1 . Положим $p_1 = q_0$ и допустим, что $\log A_0 / \log(q_0 p_0^{-1}) = -a$ и $\log A_1 / \log(q_0 p_0^{-1}) = b$, где a и b — положительные целые числа.

(I) Тогда неравенства (34) превращаются в равенства, и приближения (38) и (40) становятся точными формулами.

(II) Мощность критерия равна

$$\beta(p) = \frac{q^a p^b - p^{a+b}}{q^{1+b} - p^{a+b}} \quad \text{для } p \neq \frac{1}{2},$$

$$\beta\left(\frac{1}{2}\right) = \frac{a}{(a+b)} \quad (\text{по непрерывности}).$$

(III) Правило остановки испытаний совпадает с правилом окончания игры в задаче о разорении, где два игрока с капиталами a и b проводят последовательность игр с единичной ставкой с вероятностями выигрыша p и q соответственно до разорения одного из них *).

[Испытания продолжаются до тех пор, пока $-a < 2 \sum x_i - n < b$, и (I) и (II) следуют из того факта, что средний член этого неравенства равен 0 при $n = 0$ и при каждом испытании увеличивается или уменьшается на 1.]

*) Альтернативный вывод формулы для $\beta(p)$ в этом случае см., например, в разделе 2 главы 14 книги: Feller, An Introduction to Probability Theory and Its Applications I, New York, John Wiley & Sons, 2nd ed., 1957. (Русский перевод: Феллер В., Введение в теорию вероятностей и ее приложения, том I, М., Мир, 1967.)

37. *Тожество Вальда.* Если $Z_1, Z_2, \dots$ независимы, одинаково распределены и $E|Z| < \infty$, и число наблюдений определяется в соответствии с последовательным правилом, для которого $E(N) < \infty$, то

$$E(Z_1 + \dots + Z_N) = E(N) E(Z). \quad (47)$$

[Левая часть равна

$$\begin{aligned} \sum_{n=1}^{\infty} \left[P\{N=n\} \sum_{i=1}^n E(Z_i | N=n) \right] &= \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} P\{N=n\} E(Z_i | N=n) = \\ &= \sum_{i=1}^{\infty} P\{N \geq i\} E(Z_i | N \geq i). \end{aligned}$$

Так как событие $N \geq i$ зависит только от $Z_1, \dots, Z_{i-1}$ и не зависит от Z_i и так как $\sum_{i=1}^{\infty} P\{N \geq i\} = E(N)$, то мы приходим к требуемому соотношению.

Чтобы оправдать перестановку членов бесконечного ряда, заменим всюду Z_i на $|Z_i|$. Это показывает, что

$$E|Z_1 + \dots + Z_N| \leq E(|Z_1| + \dots + |Z_N|) = E|Z| \cdot E(N) < \infty,$$

откуда вытекает нужная абсолютная сходимость.]

38. (I) Пусть Z — такая случайная величина, что (а) $E(Z) \neq 0$, (б) $\psi(h) = E(e^{hz})$ существует при всех действительных h , (в) $P\{e^Z < 1 - \delta\}$ и $P\{e^Z > 1 + \delta\}$ положительны при некотором $\delta > 0$. Тогда существует одно и только одно решение $h \neq 0$ уравнения $\psi(h) = 1$.

(II) Это приводит к достаточным условиям существования ненулевого решения у уравнения (37).

[(I) Функция ψ выпукла, так как $\psi''(h) = E(Z^2 e^{hz}) > 0$; кроме того, $\psi(h) \rightarrow \infty$ при $h \rightarrow \pm\infty$. Следовательно, ψ имеет минимум h_0 , в котором $\psi'(h_0) = E(Z e^{h_0 z}) = 0$, так что по (а) $h_0 \neq 0$. Раз $\psi(0) = 1$, то существует единственное $h_1 \neq 0$, для которого $\psi(h_1) = 1$.

(II) Определяя Z с помощью (35), мы можем записать (37) в виде $E(e^{hz}) = 1$.]

39. Следующий пример показывает, что мощность критерия иногда может быть увеличена за счет случайности объема выборки (даже когда рандомизация не связана с результатами наблюдений *). Пусть $X_1, \dots, X_n$ независимы и распределены одинаково по $N(\theta, 1)$. Рассмотрим проблему проверки гипотезы $H: \theta = 0$ при альтернативе $K: \theta = \theta_1 > 0$.

(I) Мощность наиболее мощного критерия, как функция от объема выборки n , не обязана быть вогнутой.

(II) В частности, при $\alpha = 0,005$, $\theta_1 = 1/2$ большая мощность получается, если взять 2 или 16 наблюдений (то и другое с вероятностью 1/2), чем если взять выборку фиксированного объема 9.

(III) Дальнейшего увеличения мощности можно достичь, если допустить два возможных уровня значимости α_1 и α_2 для двух объемов выборки и требовать только, чтобы средний уровень значимости был равен $\alpha = 0,005$. Примеры: (а) с вероятностью 1/2 берут $n_1 = 2$ и строят критерий с уровнем $\alpha_1 = 0,001$ или берут $n_2 = 16$ и критерий с уровнем $\alpha_2 = 0,009$; (б) с вероятностью 1/2 берут $n_1 = 0$ или $n_2 = 18$ при соответствующих уровнях $\alpha_1 = 0$, $\alpha_2 = 0,01$.

*) Этот и подобные примеры обсуждались Крускалом в семинаре, происходившем в Колумбийском университете в 1954 г. Недавно более детальное исследование этого явления было произведено в работе: Cohen, On mixed single sample experiments, Ann. Math. Stat. 29 (1958), 947—971.

14. Литературные ссылки

Методы проверки гипотез развивались постепенно и на первых этапах сводились к весьма расплывчатым утверждениям о значимости или незначимости совокупности результатов наблюдений. Отдельные приложения (самые ранние восходят, вероятно, к Лапласу (1773)) встречаются в 19 веке в работах, например, Гаварэ (1840), Лексиса (1875, 1877) и Эджворта (1885). Систематическое использование проверки гипотез началось с работ Карла Пирсона; особенно следует отметить статью о χ^2 -критерии (1900).

Первыми авторами, указавшими, что рациональный выбор критерия должен основываться не только на проверяемой гипотезе, но и на допускаемых альтернативах, были Нейман и Пирсон (1928). Они ввели понятие ошибок первого и второго рода и мотивировали использование отношения правдоподобия как основы общего метода построения критериев. Логическое завершение этих исследований дано в работе Неймана и Пирсона (1933), где была развита теория РНМ критериев *).

Вероятно, самый ранний пример доверительных интервалов встречается у Лапласа (1812), который показал, как можно обратить (приближенное) вероятностное утверждение о степени расхождения наблюдаемой частоты и биномиальной вероятности p , с тем чтобы найти подходящий интервал для p . Другие примеры можно найти в работах Гаусса (1816), Фурье (1826) и Лексиса (1875). Однако, хотя во всех этих случаях сделанные утверждения формально правильны, авторы стремились рассматривать параметр как переменную, которая с заданной вероятностью попадает в фиксированный доверительный интервал. Правильная интерпретация была в первый раз отмечена, по-видимому, у Е. Б. Вилсона (1927). Примерно в это же время два точных доверительных утверждения были высказаны Уоркингом и Хотеллингом (1929), и Хотеллингом (1931).

Общий метод отыскания точных доверительных границ для действительного параметра в непрерывном распределении был предложен Фишером (1930), который впоследствии отказался от этой интерпретации своей работы. (Библиография, касающаяся фишеровской концепции доверительных вероятностей, в терминах которых формулируется теория, приведена у Тьюки

*) Иной подход к проверке гипотез, основанный на использовании априорных вероятностей, развит в книге: Jeffreys, *Theory of Probability*, Oxford, Clarendon Press, 2nd ed., 1948. Некоторые вопросы связи между двумя теориями обсуждаются в работах: Lindley, *A statistical paradox*, *Biometrika* 44 (1957), 187—192 и Bartlett, *A comment on D. V. Lindley's statistical paradox*, *Biometrika* 44 (1957), 523—534.

(1957).) Примерно в то же время*) Нейманом была разработана общая теория доверительных утверждений**). Им же была показана их тесная связь с теорией проверки гипотез. Детальное изложение результатов, которое лежит в основе приведенной здесь трактовки материала, было опубликовано Нейманом в статьях 1937 и 1938 годов.

Бирнбаум и Чепмэн (Birnbauм Z. W. and Chapman D. G.)

(1950) On optimum selections from multinormal populations, *Ann. Math. Stat.*, vol. 21, 443—447. [Задача 19.]

Блекуэлл (Blackwell D.)

(1951) Comparison of experiments, *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Univ. Calif. Press, 93—102.

(1953) Equivalent comparisons of experiments, *Ann. Math. Stat.*, vol. 24, 265—272. [Теория, пример 4 и задачи раздела 4.]

Вальд (Wald A.)

(1947) *Sequential Analysis*, New York, John Wiley & Sons. (Русский перевод: А. Вальд, Последовательный анализ, Физматгиз, 1960.) [Теория и применение последовательных критериев отношения вероятностей.]

Вальд и Вольфовиц (Wald A. and Wolfowitz J.)

(1948) Optimum character of the sequential probability ratio test, *Ann. Math. Stat.*, vol. 19, 326—339.

(1950) Bayes solutions of sequential decision problems, *Ann. Math. Stat.*, vol. 21, 82—99 [В этих работах доказываются оптимальные свойства последовательного критерия отношения вероятностей, данные в разделе 12. В этой связи см также работу Эрроу, Блекуэлла и Гиршика.]

Вилсон (Wilson E. B.)

(1927) Probable inference, the law of succession, and statistical inference, *J. Am. Stat. Assoc.*, vol. 22, 209—212.

Вольфовиц (Wolfowitz J.)

(1947) The efficiency of sequential estimates and Wald's equation for sequential processes. *Ann. Math. Stat.*, vol. 18, 215—230. [Доказательство тождества Вальда (45), данного в задаче 37.]

Гаварэ (Gavarret J.)

(1840) *Principes généraux de statistique médicale*, Paris.

Гаусс (Gauss C. F.)

(1816) Bestimmung der Genauigkeit der Beobachtungen, *Z. Astronomie und verwandte Wissenschaften*, vol. 1. (Перепечатано в собрании сочинений Гаусса, т. 4, стр. 109—119.)

Гренандер (Grenander U.)

(1950) Stochastic processes and statistical inference, *Arkiv för Matematik*, vol. 1, 195—277. [Применение фундаментальной леммы к проблемам в стохастических процессах.]

Данциг и Вальд (Dantzig B. and Wald A.)

(1951) On the fundamental lemma of Neyman — Pearson, *Ann. Math. Stat.*, vol. 22, 87—83. [Даны необходимые условия (включая условия теоремы 5)]

*) Ср. Neyman, Fiducial argument and the theory of confidence intervals, *Biometrika* 32 (1941), 128—150.

**) См. также Neyman J., Silver jubilee of my dispute with Fisher, *J. Oper Res. Soc. of Japan* 3, 145—154. — *Прим. пер.*

для того, чтобы критическая функция, которая максимизирует некоторый интеграл при нескольких ограничениях интегрального типа, удовлетворяла (21).]

Дворецкий, Кифер и Вольфовиц (Dvoretzky A., Kiefer J. and Wolfowitz J.)

(1953) Sequential decision problems for processes with continuous time parameter. Testing hypotheses, *Ann. Math. Stat.*, vol. 24, 254—264. [Оптимальное свойство последовательного критерия отношений вероятностей распространяется на непрерывно наблюдаемые случайные процессы.]

Карлин (Karlin S.)

(1955) Decision theory for Pólya type distributions. Case of two actions, I., *Proc. Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, Berkeley, Univ. Calif. Press, 115—129.

(1957) Pólya type distributions, II, *Ann. Math. Stat.*, vol. 28, 281—308. [Свойства распределений типа Пуассона, включая задачи 23—25.]

Карлин и Рубин (Karlin S. and Rubin H.)

(1956) The theory of decision procedures for distributions with monotone likelihood ratio, *Ann. Mat. Stat.*, vol. 27, 272—299. [Общая теория семейств с монотонным отношением правдоподобия, включая теорему 3.]

Кендалл и Стьюарт (Kendall M. G., Stuart A.)

(1973) Статистические выводы и связи, пер. с англ., М., Наука.

Лаплас (Laplace P. S.)

(1773) Mémoire sur l'inclinaison moyenne des orbites des comètes, *Mem. acad. roy. sci. Paris*, vol. VII (1776), 503—524.

(1812) *Théorie Analytique des Probabilités*, Paris. [Третье издание 1820 г. перепечатано в виде седьмого тома собраний сочинений Лапласа.]

Лексис (Lexis W.)

(1875) *Einleitung in die Theorie der Bevölkerungsstatistik*, Strassburg.

(1877) *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*, Freiburg.

Леман (Lehmann E. L.)

(1955) Ordered families of distributions, *Ann. Math. Stat.*, vol. 26, 399—419. [Леммы 1, 2 и 5.]

Леман и Стейн (Lehmann E. L. and Stein C.)

(1948) Most powerful tests of composite hypotheses, *Ann. Math. Stat.*, vol. 19, 495—516. [Теорема 7 и ее применения.]

Нейман (Neuman J.)

(1937) Outline of a theory of statistical estimation based on the classical theory of probability, *Phil. Trans. Roy. Soc.*, vol. 236, 333—380. [Развита теория оптимальных доверительных множеств. Задача сводится к определению оптимальных критериев в соответствующих классах гипотез.]

(1938) L'estimation statistique traitée comme un problème classique de probabilité, *Actualités sci. et ind.*, vol. 739, 25—57.

(1952) Lectures and Conferences on Mathematical Statistics, Washington, Graduate School, U. S. Dept. Agriculture, 2nd ed., 43—46. [Описание разнообразных подходов к проблеме проверки гипотез.]

(1959) Optimal asymptotic tests of composite, *Probability and Stat.: The H. Gramér volume*, Upsala, 213—234.

Нейман и Пирсон (Neuman J. and Pearson E. S.)

(1928) On the use and interpretation of certain test criteria, *Biometrika*, vol. 20A, 175—240, 263—294.

(1933) On the problem of the most efficient tests of statistical hypotheses, Phil. Trans. Roy. Soc., Ser. A, vol. 231, 289—337. [Основная работа по теории проверки гипотез. Проблема сформулирована в терминах ошибок двух родов; изложена основа теории, включая фундаментальную лемму. Применения, включая задачу 1.]

(1936) Contributions to the theory of testing statistical hypotheses 1. Unbiased critical regions of type A and type A_1 , Stat. Res. Mem., vol. 1, 1—37. [Обобщение фундаментальной леммы на случай нескольких ограничивающих условий.]

(1936) Sufficient statistics and unformly most powerful tests of statistical hypotheses, Stat. Res. Mem., vol. 1, 113—137. [Задача 2 (II).]

Пирсон (Pearson K.)

(1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Phil. Mag., Ser. 5, vol. 50, 157—172.

Стейн (Stein C. M.)

(1946) A note on cumulative sums, Ann. Math. Stat., vol. 17, 489—499. [Задачи 34 и 35.]

(1951) A property of some tests of composite hypotheses, Ann. Math. Stat., vol. 22, 475—476. [Задача 29.]

Томпсон (Thompson W. R.)

(1936) On confidence ranges for the median and other expectation distributions for populations of unknown distribution form, Ann. Math. Stat., vol. 7, 122—128. [Задача 28.]

Тьюки (Tukey W.)

(1957) Some examples with fiducial relevance, Ann. Math. Stat., vol. 28, 687—695.

Уоркинг и Хотеллинг (Working H. and Hotelling H.)

(1929) Applications of the theory of error to the interpretation of trends, J. Am. Stat. Assoc., Suppl., vol. 24, 73—85.

Фишер (Fisher R. A.)

(1930) Inverse probability, Proc. Cambridge Phil. Soc., vol. 26, 528—535.

Фрезер (Fraser D. A. S.)

(1953) Non-parametric theory: Scale and location parameters, Canad. J. Math., vol. 6, 46—68. [Пример 8.]

(1956) Sufficient statistics with nuisance parameters, Ann. Math. Stat., vol. 27, 838—842. [Задача 31.]

Фурье (Fourier J. B. J.)

(1826) Recherches statistiques sur la ville de Paris et le département de la Seine, vol. 3.

Хотеллинг (Hotelling H.)

(1931) The generalization of Student's ratio, Ann. Math. Stat., vol. 2, 360—378.

Чернов и Шеффэ (Chernoff H. and Scheffé H.)

(1952) A generalization of the Neyman — Pearson fundamental lemma, Ann. Math. Stat., vol. 23, 213—225.

Ширяев А. Н.

(1961) Статистический последовательный анализ, М., Наука.

Эджворт (Edgeworth F. Y.)

(1885) Methods of statistics, Jubilee volume of the Stat. Soc., London, E. Stanford.

Эпштейн и Собел (Epstein B. and Sobel M.)

(1953) Life testing. J. Am. Stat. Assoc., vol. 48, 486—502. [Задача 8.]

Эрроу, Блекуэлл и Гиршик (Arrow R. J., Blackwell D. and Girschick M. A.)

(1949) Bayes and minimax solutions of sequential decision problems, Econometrica, vol. 17, 213—244.

ГЛАВА 4

НЕСМЕЩЕННОСТЬ: ТЕОРИЯ И ПЕРВЫЕ ПРИМЕНЕНИЯ

1. Несмещенность при проверке гипотез

Простое условие, которое можно требовать от критериев гипотезы $H: \theta \in \Omega_H$ при сложной альтернативе $K: \theta \in \Omega_K$, состоит в том, что ни при одной альтернативе из K вероятность отклонения гипотезы не должна быть меньше размера критерия. Если это условие нарушено, то найдутся альтернативы, при которых принятие гипотезы более вероятно, чем в некоторых случаях, когда она верна. Критерий φ , для которого выполняется указанное условие, т. е. для функции мощности $\beta_\varphi(\theta) = E_\theta \varphi(X)$ которого верны неравенства

$$\begin{aligned} \beta_\varphi(\theta) &\leq \alpha, & \text{если } \theta \in \Omega_H, \\ \beta_\varphi(\theta) &\geq \alpha, & \text{если } \theta \in \Omega_K, \end{aligned} \quad (1)$$

называется *несмещенным*. При надлежащем выборе функции потерь это определение оказывается, как было показано в главе 1, частным случаем общего определения несмещенности. РНМ критерий, если он существует, оказывается несмещенным, так как его мощность не может быть меньше мощности критерия $\varphi(x) \equiv \alpha$.

В широком классе проблем, где РНМ критерии не существуют, тем не менее существуют РНМ несмещенные критерии. В частности, сюда включаются некоторые гипотезы вида $\theta \leq \theta_0$ или $\theta = \theta_0$ в предположении, что распределения результатов наблюдений зависят кроме θ также и от других параметров.

Если $\beta_\varphi(\theta)$ — непрерывная функция θ , то несмещенность влечет

$$\beta_\varphi(\theta) = \alpha \quad \text{для всех } \theta \in \omega, \quad (2)$$

где ω — общая граница Ω_H и Ω_K , т. е. множество точек θ предельных как для Ω_H , так и для Ω_K . Критерии, удовлетворяющие последнему условию, называются *подобными на границе* (H и

К). Так как более удобно оперировать с (2), чем с (1), то нижеследующая лемма играет важную роль в отыскании РНМ несмещенных критериев.

Л е м м а 1. Если распределения P_θ таковы, что функция мощности любого критерия непрерывна, и если φ_0 является РНМ среди всех критериев, удовлетворяющих (2), и имеет уровень α как критерий H , то φ_0 — РНМ несмещенный критерий.

Доказательство. Класс критериев, удовлетворяющих (2), содержит в себе класс несмещенных критериев и, следовательно, φ_0 равномерно не менее мощен, чем любой несмещенный критерий. С другой стороны, критерий φ_0 является несмещенным, так как он равномерно не менее мощен, чем $\varphi(x) \equiv \alpha$.

2. Однопараметрические экспоненциальные семейства

Пусть θ — действительный параметр и $X = (X_1, \dots, X_n)$ — случайный вектор, имеющий по отношению к некоторой мере μ плотность

$$p_\theta(x) = C(\theta) e^{\theta T(x)} h(x).$$

Как было показано в главе 3, РНМ критерий существует в случаях, когда гипотеза H и класс альтернатив K определяются условиями (I) $H: \theta \leq \theta_0$, $K: \theta > \theta_0$ (следствие 2) или (II) $H: \theta \leq \theta_1$ или $\theta \geq \theta_2$ ($\theta_1 < \theta_2$), $K: \theta_1 < \theta < \theta_2$ (теорема 6), и не существует в случае (III) $H: \theta_1 \leq \theta \leq \theta_2$, $K: \theta < \theta_1$ или $\theta > \theta_2$. Мы покажем теперь, что в случае (III) существует РНМ несмещенный критерий, задаваемый равенствами

$$\varphi(x) = \begin{cases} 1 & \text{при } T(x) < C_1 \text{ или } > C_2, \\ \gamma_i & \text{при } T(x) = C_i \quad i = 1, 2, \\ 0 & \text{при } C_1 < T(x) < C_2, \end{cases} \quad (3)$$

где постоянные C и γ находятся из уравнений

$$E_{\theta_0} \varphi(X) = E_{\theta_2} \varphi(X) = \alpha. \quad (4)$$

Функция мощности непрерывна по теореме 9 главы 2, так что лемма 1 применима. Множество ω состоит из двух точек θ_1 и θ_2 , поэтому мы рассмотрим сначала задачу максимизации $E_{\theta'} \varphi(x)$ для θ' , лежащего вне интервала $[\theta_1, \theta_2]$, и с учетом ограничений (4). Если задачу сформулировать в терминах $1 - \varphi(x)$, то из части (II) теоремы 6 главы 3 вытекает, что ее решение дается формулами (3) и (4). Следовательно, построенный критерий является РНМ в классе критериев, подчиненных (4), и по лемме 1 — в классе всех несмещенных критериев. Как видно из части (III) цитированной теоремы, мощность этого критерия имеет минимум в точке, лежащей между θ_1 и θ_2 , и строго возрастает при удалении от точки минимума влево или вправо,

Проблема проверки $H: \theta = \theta_0$ при альтернативе $\theta \neq \theta_0$ близка к уже разобранной. Для нее также существует РНМ несмещенный критерий, определяемый по (3); константы в этом случае находятся из уравнений

$$E_{\theta_0}[\varphi(X)] = \alpha \quad (5)$$

и

$$E_{\theta_0}[T(X)\varphi(X)] = E_{\theta_0}[T(X)]\alpha. \quad (6)$$

Чтобы доказать это, обозначим θ' какую-либо специальную альтернативу и возьмем достаточную статистику T , распределение которой в соответствии с леммой 7 главы 2 имеет вид

$$dP_{\theta}(t) = C(\theta) e^{\theta t} dv(t).$$

Несмещенность критерия $\psi(t)$ влечет (5) с $\varphi(x) = \psi[T(x)]$; отсюда же вытекает наличие у функции мощности $\beta(\theta) = E_{\theta}[\psi(T)]$ минимума при $\theta = \theta_0$. По теореме 9 главы 2 функция $\beta(\theta)$ дифференцируема и ее производная может быть вычислена дифференцированием $E_{\theta}\psi[T]$ под знаком математического ожидания, так что для всех критериев $\psi(t)$

$$\beta'(\theta) = E_{\theta}[T\psi(T)] + \frac{C'(\theta)}{C(\theta)} E_{\theta}[\psi(T)].$$

Для $\psi(t) \equiv \alpha$ последнее уравнение превращается в

$$0 = E_{\theta}(T) + \frac{C'(\theta)}{C(\theta)}.$$

Подстановка в выражение для $\beta'(\theta)$ дает

$$\beta'(\theta) = E_{\theta}[T\psi(T)] - E_{\theta}(T) E_{\theta}[\psi(T)],$$

и, следовательно, несмещенность, в дополнение к (5), влечет и (6).

Пусть M обозначает множество точек $(E_{\theta_0}[\psi(T)], E_{\theta_0}[T\psi(T)])$, где ψ пробегает совокупность всех критических функций. Тогда M выпукло и содержит все точки $(u, uE_{\theta_0}(T))$ с $0 < u < 1$. M также содержит точки (α, u_2) с $u_2 > \alpha E_{\theta_0}(T)$. Это следует из того факта, что существуют критерии с $E_{\theta_0}[\psi(T)] = \alpha$ и $\beta'(\theta_0) > 0$ (см. задачу 18 главы 3). Так как по сходным причинам M содержит точки (α, u_1) с $u_1 < \alpha E_{\theta_0}(T)$, то точка $(\alpha, \alpha E_{\theta_0}(T))$ является внутренней для M . Поэтому по теореме 5 (IV) главы 3 существуют константы k_1, k_2 и критерий $\psi(t)$, удовлетворяющие (5) и (6) с $\varphi(x) = \psi[T(x)]$ и такие, что $\psi(t) = 1$ при

$$C(\theta_0)(k_1 + k_2 t) e^{\theta_0 t} < C(\theta') e^{\theta' t},$$

т. е. при

$$a_1 + a_2 t < e^{bt}.$$

Эта область представляет собою или полупрямую, или внешнюю часть некоторого интервала. По теореме 2 (II) главы 3 односторонний критерий имеет строго монотонную функцию мощности и потому не может удовлетворять (6). Таким образом, $\psi(t)$ равно 1 при $t < C_1$ или $t > C_2$, и наиболее мощный критерий, подчиненный (5) и (6), дается формулой (3). Этот критерий несмещенный, как можно видеть, сравнивая его с $\varphi(x) \equiv \alpha$. Он также является РНМ несмещенным, так как класс критериев, удовлетворяющих (5) и (6), включает класс несмещенных критериев.

Определение критерия можно упростить, если при $\theta = \theta_0$ распределение T симметрично относительно некоторой точки a , так что $P_{\theta_0}\{T < a - u\} = P_{\theta_0}\{T > a + u\}$ при всех действительных u . Любой критерий, симметричный относительно a и удовлетворяющий (5), непременно удовлетворяет и (6), поскольку $E_{\theta_0}[T\psi(T)] = E_{\theta_0}[(T - a)\psi(T)] + aE_{\theta_0}\psi(T) = a\alpha = E_{\theta_0}(T)\alpha$. Константы C и γ находятся поэтому из уравнений

$$P_{\theta_0}\{T < C_1\} + \gamma_1 P_{\theta_0}\{T = C_1\} = \frac{\alpha}{2}, \quad C_2 = 2a - C_1, \quad \gamma_2 = \gamma_1.$$

Полученные выше критерии гипотез $\theta_1 \leq \theta \leq \theta_2$ и $\theta = \theta_0$ являются *строго несмещенными* в том смысле, что их мощность больше α при всех альтернативах θ . Для первого из этих критериев, определенного формулами (3) и (4), строгая несмещенность является немедленным следствием теоремы 6 (III) главы 3. Действительно, последняя устанавливает, что мощность рассматриваемого критерия имеет минимум в точке θ_0 между θ_1 и θ_2 и строго возрастает при удалении θ от θ_0 в любом направлении. Второй из этих критериев, определяемый формулами (3), (5) и (6), имеет непрерывную функцию мощности с минимумом размера α при $\theta = \theta_0$. Таким образом, существуют $\theta_1 < \theta_0 < \theta_2$ такие, что $\beta(\theta_1) = \beta(\theta_2) = c$, где $\alpha \leq c < 1$. Рассматриваемый критерий совпадает, следовательно, с РНМ несмещенным критерием уровня c для проверки гипотезы $\theta_1 \leq \theta \leq \theta_2$, и его мощность строго возрастает при удалении от θ_0 в любом направлении. Утверждение доказано.

Пример 1. Пусть X — число успехов в n биномиальных испытаниях с вероятностью успеха p . Теория, которая должна быть проверена, приписывает параметру p значение p_0 , так что ей соответствует гипотеза $H: p = p_0$. Отвергая гипотезу H , обычно стремятся указать, превосходит ли p значение p_0 или нет. Раз заключение $p \neq p_0$ все равно приводит к необходимости дальнейших исследований, то на первом шаге разумно решить, согласуются ли данные с гипотезой $p = p_0$ или нет. Поэтому формулировка задачи, как задачи проверки вышеуказанной гипотезы, вполне приемлема.

РНМ несмещенный критерий проверки H дается формулой (3) с $T(X) = X$. Уравнение (5) превращается в

$$\sum_{x=C_1+1}^{C_2-1} \binom{n}{x} p_0^x q_0^{n-x} + \sum_{i=1}^2 (1 - \gamma_i) \binom{n}{C_i} p_0^{C_i} q_0^{n-C_i} = 1 - \alpha,$$

и левая часть может быть подсчитана с помощью таблиц вероятностей отдельных значений и функции распределения X . Соотношение (6) с помощью тождества

$$x \binom{n}{x} p_0^x q_0^{n-x} = n p_0 \binom{n-1}{x-1} p_0^{x-1} q_0^{(n-1)-(x-1)}$$

приводится к виду

$$\begin{aligned} \sum_{x=C_1+1}^{C_2-1} \binom{n-1}{x-1} p_0^{x-1} q_0^{(n-1)-(x-1)} + \\ + \sum_{i=1}^2 (1 - \gamma_i) \binom{n-1}{C_i-1} p_0^{C_i-1} q_0^{(n-1)-(C_i-1)} = 1 - \alpha. \end{aligned}$$

Левая часть снова может быть подсчитана по таблицам биномиального распределения.

При возрастании n распределение величины $(X - np_0)/\sqrt{np_0q_0}$ стремится к нормальному $N(0, 1)$. Поэтому при не слишком малых объемах выборки и p_0 , не слишком близком к 0 или 1, распределение X приближенно симметрично относительно точки np_0 . В этом случае много проще критерий, которому соответствуют «равные хвосты» распределения, т. е. для которого C и γ определяются уравнениями

$$\begin{aligned} \sum_{x=0}^{C_1-1} \binom{n}{x} p_0^x q_0^{n-x} + \gamma_1 \binom{n}{C_1} p_0^{C_1} q_0^{n-C_1} = \\ = \gamma_2 \binom{n}{C_2} p_0^{C_2} q_0^{n-C_2} + \sum_{x=C_2+1}^n \binom{n}{x} p_0^x q_0^{n-x} = \frac{\alpha}{2}. \end{aligned}$$

Этот критерий является приближенно несмещенным и хорошо аппроксимирует несмещенный критерий. Ясно, что если n велико, то константы можно определять прямо по таблицам нормального распределения.

Пример 2. Пусть $X = (X_1, \dots, X_n)$ — выборка из нормального распределения со средним 0 и дисперсией σ^2 , так что плотность X равна

$$\left(\frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \left(-\frac{1}{2\sigma^2} \sum x_i^2 \right).$$

Статистика $T(x) = \sum x_i^2$ достаточна для σ^2 и имеет плотность $(1/\sigma^2) f_n(y/\sigma^2)$, где

$$f_n(y) = \frac{1}{2^{n/2} \Gamma(n/2)} y^{(n/2)-1} e^{-(y/2)} \quad (y > 0)$$

— плотность χ^2 -распределения с n степенями свободы. При меняющемся σ эти распределения образуют экспоненциальное семейство, которое встречается также в задаче об испытаниях на продолжительность жизни (см. задачу 13

главы 2) и в задачах, касающихся нормальных распределений с неизвестными средним и дисперсией (раздел 3 главы 5). Область принятия у РНМ несмещенного критерия для гипотезы $H: \sigma = \sigma_0$ состоит из точек, для которых

$$C_1 \leq \frac{\sum x_i^2}{\sigma_0^2} \leq C_2,$$

где

$$\int_{C_1}^{C_2} f_n(y) dy = 1 - \alpha,$$

$$\int_{C_1}^{C_2} y f_n(y) dy = \frac{(1 - \alpha) E_{\sigma_0}(\sum X_i^2)}{\sigma_0^2} = n(1 - \alpha).$$

Для определения констант из таблиц χ^2 -распределения удобно использовать тождество

$$y f_n(y) = n f_{n+2}(y)$$

и переписать второе уравнение в форме

$$\int_{C_1}^{C_2} f_{n+2}(y) dy = 1 - \alpha$$

или же можно проинтегрировать $\int_{C_1}^{C_2} y f_n(y) dy$ по частям и привести второе уравнение к форме

$$C_1^{n/2} e^{-C_1/2} = C_2^{n/2} e^{-C_2/2}.$$

Хорошую аппроксимацию к несмещенному критерию, исключая случаи малых n или σ_0 , близких к 0 или ∞ , дает критерий с «равными хвостами», для которого

$$\int_0^{C_1} f_n(y) dy = \int_{C_2}^{\infty} f_n(y) dy = \frac{\alpha}{2}.$$

Это вытекает из того, что при надлежащей нормировке распределение T стремится к нормальному, и, следовательно, T распределена асимптотически симметрично.

3. Подобие и полнота

Во многих важных задачах подлежащие проверке гипотезы касаются одного действительного параметра, в то время как распределение результатов наблюдений зависит дополнительно от некоторых «мешающих» параметров. В широком классе таких задач существует РНМ несмещенный критерий, который можно отыскать с помощью леммы 1. Но для этого требуется характеристика критериев ϕ , для которых

$$E_{\theta} \phi(X) = \alpha$$

при всех распределениях X , принадлежащих заданному семейству $\mathcal{P}^X = \{P_\theta, \theta \in \omega\}$. Такие критерии называются *подобными* по отношению к $\mathcal{P}^X$ или ω по следующей причине. Если критерий ϕ является нерандомизированным и имеет критическую область S , то S «подобна выборочному пространству» $\mathcal{X}$ в том смысле, что обе вероятности $P_\theta\{X \in S\}$ и $P_\theta\{X \in \mathcal{X}\}$ не зависят от $\theta \in \omega$.

Пусть T — достаточная статистика для $\mathcal{P}^X$ и пусть $\mathcal{P}^T$ означает семейство $\{P_\theta^T, \theta \in \omega\}$ распределений T , где θ пробегает ω . Тогда любой критерий, для которого

$$E[\phi(X) | t] = \alpha \quad \mathcal{P}^T\text{-почти всюду}^*), \quad (7)$$

оказывается подобным по отношению к $\mathcal{P}^X$, так как

$$E_\theta[\phi(X)] = E_\theta\{E[\phi(X) | T]\} = \alpha \quad \text{для всех } \theta \in \omega.$$

О критериях, удовлетворяющих (7), говорят, что они имеют *неймановскую структуру* относительно T . Характерное обстоятельство состоит в том, что вероятность отклонения гипотезы равна α на каждой из поверхностей $T = t$. Так как распределение на каждой из этих поверхностей не зависит от θ для $\theta \in \omega$, то условие (7) по существу сводит проблему к проверке простой гипотезы при каждом значении t . Среди критериев, обладающих неймановской структурой, часто без труда находится наиболее мощный, для чего проблема оптимизации решается на каждой поверхности отдельно. Найденный критерий оказывается наиболее мощным среди всех подобных критериев в предположении, что каждый подобный критерий обладает неймановской структурой. Условия, при которых верно это последнее утверждение, мы сформулируем в терминах следующего определения.

Семейство $\mathcal{P}$ распределений вероятностей P называется *полным*, если для любой измеримой функции f

$$E_P[f(X)] = 0 \quad \text{для всех } P \in \mathcal{P} \quad (8)$$

влечет

$$f(x) = 0 \quad \mathcal{P}\text{-почти всюду.} \quad (9)$$

В приложениях роль $\mathcal{P}$ будет играть семейство распределений достаточной статистики.

Пример 3. Рассмотрим n независимых испытаний с вероятностью p успеха, и пусть X_i равно 1 или 0, в соответствии с тем, закончилось i -е испытание успехом или нет. Тогда $T = X_1 + \dots + X_n$ является достаточной статистикой для p , и семейство возможных распределений равно $\mathcal{P}^T = \{b(p, n), 0 \leq p \leq 1\}$. Для этого семейства из (8) вытекает, что

$$\sum_{t=0}^n f(t) \binom{n}{t} p^t = 0 \quad \text{при всех } 0 < p < \infty,$$

*) Говорят, что утверждение выполняется $\mathcal{P}$ -почти всюду, если оно не выполняется лишь на множестве N с $P(N) = 0$ для всех $P \in \mathcal{P}$.

где $p = p/(1-p)$. Левая часть есть многочлен от p , все коэффициенты которого должны быть равны нулю. Следовательно, $f(t) = 0$ для $t = 0, \dots, n$, и биномиальное семейство распределений полно (или, что то же самое, T -полно).

Пример 4. Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения $R(0, \theta)$, $0 < \theta < \infty$. Тогда $T = \max(X_1, \dots, X_n)$ является достаточной статистикой для θ , и (8) превращается в

$$\int f(t) dP_{\theta}^T(t) = n\theta^{-n} \int_0^{\theta} f(t) t^{n-1} dt = 0 \quad \text{при всех } \theta.$$

Пусть $f(t) = f^+(t) - f^-(t)$, где $f^+(t)$ и $f^-(t)$ обозначают положительную и отрицательную части $f(t)$ соответственно.

Тогда

$$v^+(A) = \int_A f^+(t) t^{n-1} dt \quad \text{и} \quad v^-(A) = \int_A f^-(t) t^{n-1} dt$$

являются мерами на борелевских множествах $(0, \infty)$, которые совпадают для всех интервалов, а следовательно, и для всех A . Отсюда вытекает, что $f^+(t) = f^-(t)$, за исключением, быть может, множества нулевой меры Лебега. Поэтому $f(t) = 0$ $\mathcal{P}^T$ -почти всюду.

Пример 5. Пусть $X_1, \dots, X_m; Y_1, \dots, Y_n$ независимы и распределены нормально по законам $N(\xi, \sigma^2)$ и $N(\xi, \tau^2)$ соответственно. Тогда совместная плотность этих случайных величин равна

$$C(\xi, \sigma, \tau) \exp \left(-\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\xi}{\sigma^2} \sum x_i - \frac{1}{2\tau^2} \sum y_j^2 + \frac{\xi}{\tau^2} \sum y_j \right).$$

Статистика $T = (\sum X_i, \sum X_i^2, \sum Y_j, \sum Y_j^2)$ является достаточной. Она неполна, так как $E(\sum Y_j/n - \sum X_i/m)$ равно тождественно нулю. Однако если Y_j ($j = 1, \dots, n$) имеет среднее $E(Y) = \eta$, которое меняется независимо от ξ , то множество возможных значений параметров $\theta_1 = -1/2\sigma^2$, $\theta_2 = \xi/\sigma^2$, $\theta_3 = -1/2\tau^2$, $\theta_4 = \eta/\tau^2$ содержит четырехмерный прямоугольник, и потому, как следует из теоремы 1 (см. ниже), семейство $\mathcal{P}^T$ полно.

В широком классе случаев (включающем пример 3) полнота семейств распределений может быть выведена из следующей теоремы.

Теорема 1. Пусть X — случайный вектор с распределением вероятностей

$$dP_{\theta}(x) = C(\theta) \exp \left[\sum_{j=1}^k \theta_j T_j(x) \right] d\mu(x)$$

и пусть $\mathcal{P}^T$ — семейство распределений статистики $T = (T_1(X), \dots, T_k(X))$ при θ , пробегающем множество ω . Тогда, если ω содержит k -мерный интервал, то $\mathcal{P}^T$ полно.

Доказательство. Производя, если необходимо, сдвиг в пространстве параметров, мы можем предположить, не ограничивая общности, что ω содержит прямоугольник

$$I = \{(\theta_1, \dots, \theta_k): -a \leq \theta_j \leq a, \quad j = 1, \dots, k\}.$$

Пусть функция $f(t) \equiv f^+(t) - f^-(t)$ такова, что

$$E_\theta f(T) = 0 \quad \text{при всех } \theta \in \omega.$$

Обозначим ν меру, индуцированную μ в пространстве значений T . Тогда при всех $\theta \in I$

$$\int e^{\Sigma \theta_j t_j} f^+(t) d\nu(t) = \int e^{\Sigma \theta_j t_j} f^-(t) d\nu(t)$$

и, следовательно, в частности,

$$\int f^+(t) d\nu(t) = \int f^-(t) d\nu(t).$$

Разделив f на надлежащую константу, мы можем принять, что общее значение последних интегралов равно 1, так что

$$dP^+(t) = f^+(t) d\nu(t) \quad \text{и} \quad dP^-(t) = f^-(t) d\nu(t)$$

являются вероятностными мерами, для которых

$$\int e^{\Sigma \theta_j t_j} dP^+(t) = \int e^{\Sigma \theta_j t_j} dP^-(t)$$

при всех $\theta \in I$. Рассмотрим теперь эти интегралы, как функции комплексных переменных $\theta_j = \xi_j + i\eta_j$, $j = 1, \dots, k$. При любых фиксированных $\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_k$, действительные части которых лежат строго внутри промежутка от $-a$ до $+a$, эти интегралы по теореме 9 главы 2 являются аналитическими функциями θ_j в полосе R_j : $-a < \xi_j < a$, $-\infty < \eta_j < \infty$ комплексной плоскости. При фиксированных действительных $\theta_2, \dots, \theta_k$, лежащих между $-a$ и $+a$, равенство интегралов имеет место в интервале $\{(\xi_1, \eta_1): -a < \xi_1 < a, \eta_1 = 0\}$ и потому верно в полосе R_1 , в которой интегралы аналитичны. По индукции равенство может быть распространено на m -мерную комплексную область $\{(\theta_1, \dots, \theta_k): (\xi_j, \eta_j) \in R_j, j = 1, \dots, k\}$. Отсюда, в частности, вытекает, что при всех действительных $(\eta_1, \dots, \eta_k)$

$$\int e^{i \Sigma \eta_j t_j} dP^+(t) = \int e^{i \Sigma \eta_j t_j} dP^-(t).$$

Последние интегралы представляют собой характеристические функции распределений P^+ и P^- соответственно, и по теореме единственности для характеристических функций*), распределения P^+ и P^- должны совпадать. Из их определения следует, что $f^+(t) = f^-(t)$ ν -почти всюду, и потому $f(t) = 0$ $\mathcal{P}^T$ -почти всюду, что и требовалось доказать.

*) См., например, раздел 10.6 книги: Г. Крамер, Математические методы статистики, Мир, 1975.

Пример 6. Пусть $X_1, \dots, X_N$ независимы и одинаково распределены с функцией распределения $F \in \mathcal{F}$, где $\mathcal{F}$ — семейство всех непрерывных распределений. Как было показано в разделе 6 главы 2, совокупность порядковых статистик $T(X) = (X^{(1)}, \dots, X^{(N)})$ достаточна для $\mathcal{F}$. Мы докажем теперь полноту распределений T . Так как статистика $T'(X) = (\Sigma X_i, \Sigma X_i^2, \dots, \Sigma X_i^N)$ эквивалентна $T(X)$ в том смысле, что обе индуцируют одно и то же подполе в выборочном пространстве, то $T'(X)$ также достаточна. Кроме того, $T'(X)$ полна тогда и только тогда, когда полна статистика $T(X)$. Чтобы доказать полноту $T'(X)$, а тем самым и полноту $T(X)$, рассмотрим семейство плотностей

$$f(x) = C(\theta_1, \dots, \theta_N) \exp(-x^{2N} + \theta_1 x + \dots + \theta_N x^N),$$

где C — нормирующая константа. Эти плотности определены при всех значениях аргументов θ , так как интегралы от экспоненты конечны. Соответствующие распределения принадлежат $\mathcal{F}$. Плотность для выборки объема N равна

$$C^N \exp(-\Sigma x_j^{2N} + \theta_1 \Sigma x_j + \dots + \theta_N \Sigma x_j^N),$$

и эти плотности образуют экспоненциальное семейство $\mathcal{F}_0$. По теореме 1 $T'(X)$ полно для $\mathcal{F}_0$, а следовательно, и для $\mathcal{F}$, что и требовалось доказать. (Иное доказательство см. в задачах 12 и 13).

Этот же метод доказательства позволяет установить более общий результат. Пусть X_{ij} , $j = 1, \dots, N_i$; $i = 1, \dots, c$, независимы и имеют непрерывные функции распределения F_i , и пусть $X_i^{(1)} < \dots < X_i^{(N_i)}$ обозначают N_i наблюдений $X_{i1}, \dots, X_{iN_i}$, расположенных в возрастающем порядке. Тогда система порядковых статистик

$$(X_1^{(1)}, \dots, X_1^{(N_1)}, \dots; X_c^{(1)}, \dots, X_c^{(N_c)})$$

достаточна и полна для семейства распределений, получаемого, когда $F_1, \dots, F_c$ независимо пробегает класс $\mathcal{F}$. Здесь полнота доказывается рассмотрением подсемейства $\mathcal{F}_0$ семейства $\mathcal{F}$, в котором распределения F_i имеют плотности вида

$$f_i(x) = C_i(\theta_{i1}, \dots, \theta_{iN_i}) \exp(-x^{2N_i} + \theta_{i1}x + \dots + \theta_{iN_i}x^{N_i}).$$

Для настоящих целей достаточным было бы и более слабое понятие ограниченной полноты: семейство $\mathcal{P}$ называется *ограниченно полным*, если (8) влечет (9) для всех ограниченных функций f . Если $\mathcal{P}$ полно, то оно тем более ограничено полно.

Теорема 2. Пусть X — случайная величина с распределением $P \in \mathcal{P}$ и пусть T — достаточная статистика для $\mathcal{P}$. Тогда, чтобы все подобные критерии имели неймановскую структуру относительно T , необходимо и достаточно, чтобы семейство $\mathcal{P}^T$ распределений T было ограничено полным.

Доказательство. Предположим, что $\mathcal{P}^T$ ограничено полно и $\phi(X)$ — критерий, подобный относительно $\mathcal{P}$. Тогда

$$E[\phi(X) - \alpha] = 0 \quad \text{при всех } P \in \mathcal{P},$$

и, следовательно, если $\psi(t)$ обозначает условное математическое ожидание разности $\phi(X) - \alpha$ при данном t , то

$$E\psi(T) = 0 \quad \text{при всех } P^T \in \mathcal{P}^T,$$

Так как по лемме 3 главы 2 можно считать, что $\psi(t)$ ограничена, то из ограниченной полноты $\mathcal{P}^T$ вытекает $\psi(t) = 0$ и $E[\varphi(X) | t] = \alpha$ $\mathcal{P}^T$ -почти всюду, что и требовалось доказать.

Обратно, допустим, что $\mathcal{P}^T$ не является ограниченно полным. Тогда найдется функция f такая, что $|f(t)| \leq M$ при некотором M , $Ef(T) = 0$ при всех $P^T \in \mathcal{P}^T$ и $f(T) \neq 0$ с положительной вероятностью при некотором $P^T \in \mathcal{P}^T$. Положим $\varphi(t) = cf(T) + \alpha$, где $c = \min(\alpha, 1 - \alpha)/M$. Тогда φ — критическая функция, так как $0 \leq \varphi(t) \leq 1$, и она определяет подобный критерий, поскольку $E\varphi(t) = \alpha$ для всех $P^T \in \mathcal{P}^T$. Но φ не имеет неймановской структуры, так как $\varphi(T) \neq \alpha$ с положительной вероятностью по крайней мере для одного распределения класса $\mathcal{P}^T$.

4. РНМ несмещенные критерии

для экспоненциальных семейств со многими параметрами

Важный класс составляют гипотезы относительно действительного параметра в экспоненциальном семействе, при условии, что остальные параметры остаются неуточненными, «мешающими» параметрами. Во многих подобных случаях РНМ несмещенные критерии существуют и могут быть построены с помощью теории предыдущего раздела.

Пусть X распределено по закону

$$dP_{\theta, \vartheta}^X(x) = C(\theta, \vartheta) \exp \left[\theta U(x) + \sum_{i=1}^k \vartheta_i T_i(x) \right] d\mu(x), \quad (\theta, \vartheta) \in \Omega, \quad (10)$$

и пусть $\vartheta = (\vartheta_1, \dots, \vartheta_k)$ и $T = (T_1, \dots, T_k)$. Мы рассмотрим проблему проверки следующих гипотез H_j при альтернативах K_j , $j = 1, \dots, 4$:

$$H_1: \theta \leq \theta_0$$

$$K_1: \theta > \theta_0$$

$$H_2: \theta \leq \theta_1 \quad \text{или} \quad \theta \geq \theta_2$$

$$K_2: \theta_1 < \theta < \theta_2$$

$$H_3: \theta_1 \leq \theta \leq \theta_2$$

$$K_3: \theta < \theta_1 \quad \text{или} \quad \theta > \theta_2$$

$$H_4: \theta = \theta_0$$

$$K_4: \theta \neq \theta_0.$$

Мы предположим, что пространство параметров Ω выпукло и что его размерность равна $k + 1$, т. е., что оно не содержится ни в каком линейном пространстве размерности $< k + 1$. В частности, это имеет место, если Ω — естественное пространство параметров экспоненциального семейства. Мы предположим также, что в Ω найдутся как точки, для которых θ меньше, так и точки, для которых θ больше θ_0 , θ_1 и θ_2 соответственно.

Можно ограничиться только достаточными статистиками (U, T) , которые имеют совместное распределение

$$dP_{\theta, \vartheta}^{U, T}(u, t) = C(\theta, \vartheta) \exp \left[\theta u + \sum_{i=1}^k \vartheta_i t_i \right] d\nu(u, t), \quad (\theta, \vartheta) \in \Omega. \quad (11)$$

При данном $T = t$ U остается единственной переменной, и по лемме 8 главы 2 условное распределение U при данном t входит в экспоненциальное семейство

$$dP_{\theta}^{U|t}(u) = C_t(\theta) \exp(\theta u) dv_t(u).$$

Для условных распределений по следствию 2 главы 3 существует РНМ критерий для проверки H_1 с критической функцией φ_1 такой, что

$$\varphi_1(u, t) = \begin{cases} 1 & \text{при } u > C_0(t), \\ \gamma_0(t) & \text{при } u = C_0(t), \\ 0 & \text{при } u < C_0(t), \end{cases} \quad (12)$$

где константы C_0 и γ_0 определяются из уравнения

$$E_{\theta_0}[\varphi_1(U, T) | t] = \alpha \quad \text{при всех } t. \quad (13)$$

Для проверки H_2 в семействе условных распределений можно по теореме 6 главы 3 построить РНМ критерий с критической функцией

$$\varphi_2(u, t) = \begin{cases} 1 & \text{при } C_1(t) < u < C_2(t), \\ \gamma_i(t) & \text{при } u = C_i(t), \quad i = 1, 2, \\ 0 & \text{при } u < C_1(t) \text{ или } > C_2(t), \end{cases} \quad (14)$$

где константы C и γ определяются из уравнения

$$E_{\theta_1}[\varphi_2(U, T) | t] = E_{\theta_2}[\varphi_2(U, T) | t] = \alpha. \quad (15)$$

Рассмотрим теперь критерий φ_3 такой, что

$$\varphi(u, t) = \begin{cases} 1 & \text{при } u < C_1(t) \text{ если } > C_2(t), \\ \gamma_i(t) & \text{при } u = C_i(t), \quad i = 1, 2, \\ 0 & \text{при } C_1(t) < u < C_2(t), \end{cases} \quad (16)$$

где C и γ определяются из уравнения

$$E_{\theta_1}[\varphi_3(U, T) | t] = E_{\theta_2}[\varphi_3(U, T) | t] = \alpha. \quad (17)$$

При данном $T = t$ этот критерий в соответствии с разделом 2 настоящей главы будет РНМ несмещенным для проверки H_3 и РНМ среди всех критериев, удовлетворяющих (17).

Пусть, наконец, φ_4 — критическая функция, удовлетворяющая (16) с константами C и γ , которые находятся из уравнений

$$E_{\theta_0}[\varphi_4(U, T) | t] = \alpha \quad (18)$$

и

$$E_{\theta_0}[U \varphi_4(U, T) | t] = \alpha E_{\theta_0}[U | t]. \quad (19)$$

При данном $T = t$ из результатов раздела 2 снова выводим, что φ_4 является РНМ несмещенным критерием для проверки H_4 и РНМ в классе всех критериев, для которых выполняются (18) и (19).

До сих пор критические функции φ_j рассматривались как условные при $T = t$. Интерпретируя их теперь как критерии, основанные на U и T и предназначенные для гипотез относительно распределений X (или совместного распределения U и T), т. е. возвращаясь к первоначальной точке зрения, мы приходим к следующей важной теореме.

Теорема 3. *Определим критические функции: φ_1 по (12) и (13); φ_2 по (14) и (15); φ_3 по (16) и (17), φ_4 по (16), (18) и (19). Эти функции приводят к РНМ несмещенным критериям уровня α для проверки гипотез $H_1, \dots, H_4$ соответственно, если совместное распределение U и T дается формулой (11).*

Доказательство. Статистика T достаточна для θ при любом фиксированном θ и, следовательно, T достаточно для каждого

$$\omega_j = \{(\theta, \vartheta): (\theta, \vartheta) \in \Omega, \theta = \theta_j\}, \quad j = 0, 1, 2.$$

По лемме 8 главы 2 соответствующее семейство распределений T получается из формулы

$$dP_{\theta_j, \vartheta}^T(t) = C(\theta_j, \vartheta) \exp\left(\sum_{i=1}^k \vartheta_i t_i\right) dv_{\vartheta}(t), \quad (\theta_j, \vartheta) \in \omega_j; \\ j = 0, 1, 2.$$

Так как, по предположению, Ω выпукло, имеет размерность $k+1$ и содержит точки, лежащие по обе стороны от $\theta = \theta_j$, то ω_j выпукло и имеет размерность k . Поэтому ω_j содержит k -мерный прямоугольник. По теореме 1 семейство

$$\mathcal{P}_j^T = \{P_{\theta_j, \vartheta}^T: (\theta, \vartheta) \in \omega_j\}$$

полно. Из подобия критерия φ на ω_j вытекает

$$E_{\theta_j}[\varphi(U, T) | t] = \alpha.$$

(1) Рассмотрим сначала H_1 . По теореме 6 главы 2 для экспоненциального семейства мощность любого критерия непрерывна. Следовательно, достаточно доказать, что φ_1 является РНМ критерием в классе всех критериев, подобных на ω_0 (лемма 1), и тем самым в классе критериев, удовлетворяющих (13). С другой стороны, безусловная мощность критерия φ при альтернативе (θ, ϑ) равна

$$E_{\theta, \vartheta}[\varphi(U, T)] = \int \left[\int \varphi(u, t) dP_{\theta}^{U|t}(u) \right] dP_{\theta, \vartheta}^T(t). \quad (20)$$

Мы сделаем максимальной безусловную мощность, максимизируя условную мощность при каждом отдельном t (т. е. максимизируя выражение, стоящее в квадратных скобках). Так как φ_1 делает максимальной условную мощность при любой альтернативе $\theta > \theta_0$ и при ограничении (13), то мы получаем требуемый результат.

(2) Доказательства для H_2 и H_3 совершенно аналогичны. По лемме 1 достаточно доказать, что φ_2 и φ_3 являются РНМ среди всех критериев, подобных на ω_1 и ω_2 , и, следовательно, среди критериев, удовлетворяющих (15). При каждом t φ_2 и φ_3 максимизируют условную мощность в соответствующих задачах при указанном ограничении. Поэтому они максимизируют и безусловную мощность.

(3) Несмещенность критерия для H_4 влечет подобие на ω_0 и

$$\frac{\partial}{\partial \theta} [E_{\theta, \phi} \Phi(U, T)] = 0 \text{ на } \omega_0.$$

В левой части этого соотношения дифференцирование можно перенести под знак математического ожидания, и с помощью тех же вычислений, что ранее привели к (6), мы находим, что это уравнение равносильно следующему:

$$E_{\theta, \phi} [U \Phi(U, T) - \alpha U] = 0 \text{ на } \omega_0.$$

Поэтому, так как $\mathcal{P}_0^T$ полно, несмещенность влечет (18) и (19). Как и в уже разобранных случаях, критерий (который удовлетворяет (16)) является РНМ среди всех критериев, удовлетворяющих этим двум условиям. Так же, как при доказательстве леммы 1, из сравнения с критерием $\phi(u, t) \equiv \alpha$ вытекает, что предложенный критерий — РНМ несмещенный.

(4) Функции $\varphi_1, \dots, \varphi_4$ определены при каждом фиксированном t как функции от u . Для завершения доказательства необходимо установить, что эти функции измеримы по совокупности переменных u и t , так что математическое ожидание (20) существует. Мы докажем здесь это утверждение для φ_1 ; доказательство для остальных случаев намечено в задачах 14 и 15. Чтобы показать измеримость φ_1 , мы выведем сначала измеримость по t $C_0(t)$ и $\gamma_0(t)$, определенных в (12) и (13). Опуская индекс 0 и обозначая

$$F_t(u) = P_{\theta_0} \{U \leq u | t\}$$

условную функцию распределения U при $T = t$ и $\theta = \theta_0$, мы можем переписать (13) в форме

$$F_t(C) - \gamma[F_t(C) - F_t(C - 0)] = 1 - \alpha.$$

Здесь $C = C(t)$ таково, что $F_t(C - 0) \leq 1 - \alpha \leq F_t(C)$ и, следовательно,

$$C(t) = F_t^{-1}(1 - \alpha),$$

где $F_t^{-1}(y) = \inf \{u: F_t(u) \geq y\}$. Мы видим, что $C(t)$ и $\gamma(t)$ будут измеримы, если $F_t(u)$ и $F_t(u-0)$ измеримы по совокупности переменных u и t и $F_t^{-1}(1-\alpha)$ измерима по t .

При каждом фиксированном u функция $F_t(u)$ измерима по t и при каждом фиксированном t она является функцией распределения, а потому монотонна и непрерывна справа. Из второго свойства вытекает, что $F_t(u) \geq c$ в том и только том случае, когда для каждого n существует рациональное r такое, что

$$u \leq r < u + 1/n, \quad F_t(r) \geq c.$$

Следовательно, обозначая $r_1, r_2, \dots$, занумерованные каким-либо образом рациональные числа, имеем

$$\{(u, t): F_t(u) \geq c\} = \bigcap_n \bigcup_i \left\{ (u, t): 0 \leq r_i - u < \frac{1}{n}, \quad F_t(r_i) \geq c \right\}.$$

Отсюда видно, что $F_t(u)$ измерима по совокупности переменных u и t . Для $F_t(u-0)$ доказательство совершенно аналогично. Так как $F_t^{-1}(y) \leq u$ равносильно $F_t(u) \geq y$, то $F_t^{-1}(y)$ t -измерима при любом фиксированном y . Доказательство закончено.

Критерий φ_1 доказанной теоремы остается РНМ несмещенным, если заменить Ω множеством $\Omega' = \Omega \cap \{(\theta, \vartheta): \theta > \theta_0\}$, т. е. если проверять гипотезу H' : $\theta = \theta_0$ при альтернативе $\theta > \theta_0$. Предположение о том, что Ω содержит точки с $\theta < \theta_0$, было в действительности использовано только для доказательства наличия внутри ω_0 k -мерного прямоугольника. Но это утверждение остается верным, если Ω заменить на Ω' .

Оставшаяся часть этой главы, равно как и вся следующая глава, посвящена в основном применениям предыдущей теоремы к различным статистическим задачам. Этот путь — наиболее быстрый путь доказательства того, что соответствующие критерии являются РНМ несмещенными. Однако имеется и другой, более элементарный, вариант подхода к проблеме. Именно, доказательство теоремы 3 совершенно элементарно во всех пунктах, за исключением следующих: (I) того факта, что условные распределения U при данном $T = t$ образуют экспоненциальное семейство, (II) что семейство распределений T полно, (III) что производная от $E_{\theta, \varphi}(U, T)$ существует и может быть вычислена дифференцированием под знаком математического ожидания, (IV) что функции $\varphi_1, \dots, \varphi_4$ — измеримы. Вместо проверки свойств (I) — (IV) в общем случае, как было сделано в выше приведенном доказательстве, можно проверять эти свойства непосредственно в каждой отдельной задаче, что порой очень легко.

С помощью преобразования параметров теорема 3 может быть распространена на гипотезы относительно параметров типа

$$\theta^* = a_0 \theta + \sum_{i=1}^k a_i \theta_i, \quad a_0 \neq 0.$$

Преобразование формально дается нижеследующей леммой, доказательство которой очевидно.

Л е м м а 2. *Экспоненциальное семейство распределений (10) может быть представлено в виде*

$$dP_{\theta, \Phi}^x(x) = K(\theta^*, \Phi) \exp[\theta^* U^*(x) + \sum \theta_i T_i^*(x)] d\mu(x),$$

где

$$U^* = \frac{U}{a_0}, \quad T_i^* = T_i - \frac{a_i}{a_0} U.$$

Применение теоремы 3 к распределениям в той форме, как они даны в этой лемме, приводит к РНМ несмещенным критериям для гипотезы H_1^* : $\theta^* \leq \theta_0$ и определенных, аналогично предыдущему, гипотез H_2^* , H_3^* , H_4^* .

Проверяя одну из гипотез H_j , мы часто интересуемся мощностью $\beta(\theta', \Phi)$ критерия φ_j по отношению к какой-либо альтернативе θ' . Как подчеркнуто в обозначении и видно из (20), эта мощность зависит, обычно, от неизвестных «мешающих» параметров Φ . С другой стороны, мощность условного критерия при данном $T = t$

$$\beta(\theta' | t) = F_{\theta'}[\varphi(U, T) | t]$$

не зависит от Φ и имеет, следовательно, известное значение.

Величину $\beta(\theta' | t)$ можно интерпретировать двумя способами. (I) Это есть вероятность отклонения H при $T = t$. Если значение T уже известно и равно t , то возникает представление, что, по крайней мере в некоторых проблемах, указанная вероятность дает более подходящее к ситуации выражение мощности, чем $\beta(\theta', \Phi)$. Последнее получается осреднением $\beta(\theta' | t)$ по значениям t , не имеющим отношения к рассматриваемому случаю. Этот аргумент наталкивается на трудности, так как могут налагаться еще и дополнительные условия, и неясно, когда же следует остановиться.

(II) Более отчетливая интерпретация состоит в том, что $\beta(\theta' | t)$ рассматривается как оценка $\beta(\theta', \Phi)$. Так как

$$E_{\theta', \Phi}[\beta(\theta' | T)] = \beta(\theta', \Phi),$$

то эта оценка является несмещенной в смысле главы 1 (равенство 11). Далее, из теории несмещенных оценок и полноты экспоненциального семейства следует, что в классе всех несмещен-

ных оценок мощности $\beta(\theta', \theta)$ предлагаемая имеет наименьшую дисперсию *).

Безотносительно к способу интерпретации $\beta(\theta'|t)$ имеет по сравнению с безусловной мощностью тот недостаток, что ее величина становится известной после проведения наблюдений. Следовательно, она непригодна для планирования эксперимента, в частности, для определения размера выборки, если это необходимо сделать заранее, до проведения эксперимента. С другой стороны, заданная мощность β при альтернативе $\theta = \theta'$ достигается последовательным накоплением наблюдений до тех пор, пока условная мощность $\beta(\theta'|t)$ не станет $\geq \beta$.

5. Сравнение двух пуассоновских или биномиальных совокупностей

Часто возникает необходимость сравнения двух способов воздействия на изучаемые объекты или какого-либо способа воздействия с контрольной ситуацией, где он не применяется. Если результатом наблюдений является число успехов в последовательности испытаний каждого способа, например число пациентов, излеченных от некоторой болезни, проблема сводится к проверке равенства двух биномиальных вероятностей. При сравнении степени радиоактивности двух веществ исходными были бы распределения Пуассона, а проблема состояла бы в проверке тождественности этих распределений.

Желая проверить, дает ли предлагаемый способ выигрыш по сравнению с контрольными испытаниями, где он не применяется, мы формулируем одностороннюю гипотезу. Обозначая ξ_2 и ξ_1 значения параметра при применении воздействия и в контрольных опытах, мы представим класс альтернатив в виде $K: \xi_2 > \xi_1$. Основная гипотеза имеет вид $\xi_2 = \xi_1$, если априори известно, что предлагаемый способ или не вносит изменений, или дает преимущество, и вид $\xi_2 \leq \xi_1$, если допускается возможность ухудшения. Так как при обеих формулировках критерии оказываются одними и теми же, то более безопасная вторая формулировка часто оказывается и более предпочтительной.

Односторонние гипотезы оказываются уместными и в тех случаях сравнения новых способов со стандартными, когда первые представляют интерес исключительно лишь при условии улучшения результатов. С другой стороны, если два способа находятся в одинаковом положении, то проверяют гипотезу $\xi_2 = \xi_1$ при двусторонней альтернативе $\xi_2 \neq \xi_1$. Здесь формулировка представляется довольно искусственной, так как в случае

*) См. теорему 5.1 в работе: Lehmann and Scheffé, Completeness, similar regions and unbiased estimates, Sankhya, 10 (1950), 305—340.

отклонения гипотезы мы несомненно стремимся понять, какой же способ лучше *). Однако двусторонние критерии находят важное применение при построении доверительных границ для меры превосходства одного способа над другим.

Чтобы применить теорему 3 к задаче сравнения, необходимо представить соответствующие распределения в экспоненциальной форме с $\theta = f(\xi_1, \xi_2)$, например, $\theta = \xi_2 - \xi_1$ или ξ_2/ξ_1 , словом, в такой форме, чтобы интересующая нас гипотеза была эквивалентна встречающимся в теореме 3. В настоящем разделе будут рассмотрены пуассоновское и биномиальное распределения. Случаем нормальных распределений мы займемся в главе 5.

Начнем с распределений Пуассона. Пусть X и Y независимы и имеют распределения $P(\lambda)$ и $P(\mu)$, соответственно, так что их совместное распределение может быть представлено в форме

$$P\{X=x, Y=y\} = \frac{e^{-(\lambda+\mu)}}{x!y!} \exp\left[y \log \frac{\mu}{\lambda} + (x+y) \log \lambda\right].$$

Согласно теореме 3 существует РНМ несмещенные критерии для четырех гипотез H_1, H_2, H_3, H_4 , касающихся параметра $\theta = \log(\mu/\lambda)$, или, что то же самое, отношения $\rho = \mu/\lambda$. Сюда; в частности, включаются гипотезы $\mu \leq \lambda$, (или $\mu = \lambda$) при альтернативе $\mu > \lambda$, и $\mu = \lambda$ при альтернативе $\mu \neq \lambda$. Сравнивая распределение (X, Y) с (10), мы видим, что здесь $U = Y$ и $T = X + Y$, и по теореме 3, критерии применяются условным образом в целых точках отрезка прямой $X + Y = t$, лежащего в первом квадранте x, y -плоскости. Условное распределение Y при $X + Y = t$ равно (см. задачу 12 главы 2)

$$P\{Y=y | X+Y=t\} = \binom{t}{y} \left(\frac{\mu}{\lambda+\mu}\right)^y \left(\frac{\lambda}{\lambda+\mu}\right)^{t-y},$$

$$y=0, 1, \dots, t,$$

т. е. является биномиальным распределением, соответствующим t испытаниям с вероятностью успеха $p = \mu/(\lambda + \mu)$. Первоначальные гипотезы сводятся этим приемом к аналогичным гипотезам относительно параметра p биномиального распределения. Например, гипотеза $H: \mu \leq a\lambda$ превращается в $H: p \leq a/(a+1)$, которую отвергают при слишком больших Y . Критическое значение зависит, конечно, не только от a , но и от t . Его можно определить по таблицам биномиального распределения, а при больших t — (приближенно) по таблицам нормального закона.

Во многих применениях отношение $\rho = \mu/\lambda$ оказывается разумной мерой расхождения между пуассоновскими совокупностями

*) Сравнение двух способов как проблема с тремя решениями рассматривается в работах: V a h a d u r, A property of the t -statistic, Sankhyā, 12 (1952), 79—88 и L e h m a n n, A theory of some multiple decision procedures, Ann. Math. Stat. 28 (1957), 1—25, 547—572.

ми, поскольку параметры λ и μ указывают на скорость (в пространстве или во времени), с которой появляются события в сравниваемых процессах Пуассона. Можно было бы надеяться, что мощность указанных выше критериев зависит только от этого отношения, но это в действительности не так. Более того, при каждом фиксированном ρ , соответствующем альтернативе к проверяемой гипотезе, мощность $\beta(\lambda, \mu) = \beta(\lambda, \rho\lambda)$ является возрастающей функцией λ и стремится к 1 при $\lambda \rightarrow \infty$ и к α при $\lambda \rightarrow 0$. Чтобы показать это, рассмотрим мощность $\beta(\rho|t)$ условного критерия при данном t . Эта мощность есть возрастающая функция t , поскольку она представляет собой мощность оптимального критерия, построенного по t биномиальным испытаниям. Статистика T имеет распределение Пуассона с параметром $\lambda(1 + \rho)$, и ее распределения при меняющемся λ образуют экспоненциальное семейство. Из леммы 2 главы 3 следует, что безусловная мощность $E'[\beta(\rho|T)]$ является возрастающей функцией λ . При $\lambda \rightarrow 0$ или $\lambda \rightarrow \infty$ T , по вероятности, стремится к 0 или ∞ , а мощность относительно фиксированной альтернативы ρ — к α или 1.

Приведенные выше критерии применимы и к выборкам $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ из двух распределений Пуассона. Статистики $X = \sum_{i=1}^m X_i$ и $Y = \sum_{j=1}^n Y_j$ будут при этом достаточными для λ и μ и будут распределены по закону Пуассона с параметрами $m\lambda$ и $n\mu$ соответственно. При планировании эксперимента естественно желание взять $m = n$ столь большим, чтобы критерий гипотезы, скажем $H: \rho \leq \rho_0$, имел при фиксированной альтернативе ρ_1 мощность, не меньшую заданного β . Однако замечания, сделанные по поводу функции мощности при $n = 1$, остаются справедливыми при любом n , и потому высказанное пожелание неосуществимо при фиксированном n , сколь бы велико оно ни было. Это можно увидеть и непосредственно, поскольку при $\lambda \rightarrow 0$ вероятность события $X = Y = 0$ стремится к 1 как при $\rho = \rho_0$, так и при $\rho = \rho_1$. Поэтому мощность любого критерия уровня α при альтернативе $\rho = \rho_1$ и при меняющемся λ не может быть «удалена» от α . Эту трудность можно преодолеть только при последовательной процедуре наблюдения. Мы можем, например, выбрать t_0 столь большим, чтобы критерий гипотезы $\rho \leq \rho_0/(1 + \rho_0)$, построенный по t_0 биномиальным испытаниям, имел мощность $\geq \beta$ при альтернативе $\rho = \rho_1/(1 + \rho_1)$. Наблюдая пары $(X_1, Y_1), (X_2, Y_2), \dots$ до тех пор, пока не станет $\Sigma(X_i + Y_i) \geq t_0$, мы получим критерий, мощность которого $\geq \beta$ при всех альтернативах с $\rho \geq \rho_1$ *).

*) Обсуждение этой и альтернативной процедур для достижения той же самой цели дается в работе: Birnbaum, Statistical methods for Poisson processes and exponential populations, J. Am. Stat. Assoc. 49, 254—266.

Сравнение биномиальных вероятностей совершенно аналогично. Пусть X и Y — независимые случайные величины с совместным распределением

$$P\{X=x, Y=y\} = \binom{m}{x} p_1^x q_1^{m-x} \binom{n}{y} p_2^y q_2^{n-y} = \\ = \binom{m}{x} \binom{n}{y} q_1^m q_2^n \exp \left[y \left(\log \frac{p_2}{q_2} - \log \frac{p_1}{q_1} \right) + (x+y) \log \frac{p_1}{q_1} \right].$$

Мы можем проверить каждую из четырех гипотез H_1, H_2, H_3, H_4 относительно параметра $\theta = \log \left(\frac{p_2}{q_2} / \frac{p_1}{q_1} \right)$ или, что эквивалентно, относительно отношения $\rho = \frac{p_2}{q_2} / \frac{p_1}{q_1}$. Сюда включаются, в частности, проблемы проверки $H'_1: p_2 \leq p_1$ при альтернативе $p_2 > p_1$ и $H'_4: p_2 = p_1$ при альтернативе $p_2 \neq p_1$. Как и в случае распределений Пуассона, $U=Y$ и $T=X+Y$, и критерий выражается в терминах условного распределения Y на отрезке прямой $X+Y=t$. Это распределение равно

$$P\{Y=y | X+Y=t\} = C_t(\rho) \binom{m}{t-y} \binom{n}{y} \rho^y, \quad y=0, 1, \dots, t, \quad (21)$$

где

$$C_t(\rho) = 1 / \sum_{y'=0}^t \binom{m}{t-y'} \binom{n}{y'} \rho^{y'}.$$

В частном случае гипотез H'_1 и H'_4 пограничное значение θ_0 (см. формулы (13), (18), (19)) равно 0, а соответствующее значение ρ равно 1. Условное распределение превращается в гипергеометрическое распределение *)

$$P\{Y=y | X+Y=t\} = \frac{\binom{m}{t-y} \binom{n}{y}}{\binom{m+n}{t}}, \quad y=0, \dots, t.$$

6. Проверка независимости в 2×2 таблицах

Задача о том, независимо ли распределение некоторых характеристик A и B в данной совокупности, обсуждалась в разделе 4 главы 2 (пример 4). При этом предполагалось, что частные вероятности $p(A)$ и $p(B)$ известны. Мы видели, что лучше всего информирующей среди выборок объема s является та, ко-

*) Таблицы, облегчающие применение критериев в этом случае, даются в книгах: Mainland, Herrera and Sutcliffe, Tables for Use with Binomial Samples, New York, Department of Medical Statistics, N. Y. Univ. College of Medicine, 1956, и Armsen, Tables for significance tests of 2×2 contingency tables, Biometrika 42 (1955), 494—511.

торая взята из самой редкой из четырех разновидностей $A, \tilde{A}, B, \tilde{B}$. Пусть это будет, например, A . Тогда задача сводится к проверке гипотезы $H: p = p(B)$ относительно биномиального распределения $b(p, s)$.

В более типичной ситуации, когда $p(A)$ и $p(B)$ неизвестны, выборка только из одной категории, скажем A , не может служить основой для различения гипотезы и ее альтернатив. Это следует из того факта, что число элементов выборки, обладающих свойством B , имеет биномиальное распределение с вероятностью успеха $p(B|A)$, совершенно неизвестной как в случае, когда гипотеза верна, так и в противоположном случае. Гипотеза может быть, однако, проверена, если выборка берется из двух категорий: A и $\tilde{A}$ (или B и $\tilde{B}$). Возьмем, например, последний случай. Если объемы выборок равны m и n , то числа элементов этих выборок, обладающих свойством A , будут независимыми случайными величинами с биномиальными распределениями $b(p_1, m)$ и $b(p_2, n)$ соответственно. Здесь $p_1 = P(A|B)$ и $p_2 = P(A|\tilde{B})$. Гипотеза независимости двух характеристик: $p(A|B) = p(A)$ эквивалентна гипотезе $p_1 = p_2$ и задача сводится, таким образом, к разобранной в предыдущем разделе.

Вместо извлечения выборок из двух категорий часто бывает удобнее взять случайную выборку из совокупности, рассматриваемой как целое. Результаты подобной выборки могут быть сведены в следующую 2×2 таблицу сопряженности признаков. В ее клетках указано количество объектов разных категорий:

	A	$\tilde{A}$	
B	X	X'	M
$\tilde{B}$	Y	Y'	N
	T	T'	s

Величины X, X', Y и Y' имеют совместное полиномиальное распределение

$$\begin{aligned}
 P\{X=x, X'=x', Y=y, Y'=y'\} &= \\
 &= \frac{s!}{x! x'! y! y'!} p_{AB}^x p_{\tilde{A}B}^{x'} p_{A\tilde{B}}^y p_{\tilde{A}\tilde{B}}^{y'} = \\
 &= \frac{s!}{x! x'! y! y'!} p_{\tilde{A}\tilde{B}}^s \exp \left[x \log \frac{p_{AB}}{p_{\tilde{A}\tilde{B}}} + x' \log \frac{p_{\tilde{A}B}}{p_{\tilde{A}\tilde{B}}} + y \log \frac{p_{A\tilde{B}}}{p_{\tilde{A}\tilde{B}}} \right].
 \end{aligned}$$

Лемма 2 и теорема 3 применимы к любому параметру вида

$$\theta^* = a_0 \log \frac{p_{AB}}{p_{\tilde{A}\tilde{B}}} + a_1 \log \frac{p_{\tilde{A}B}}{p_{\tilde{A}\tilde{B}}} + a_2 \log \frac{p_{A\tilde{B}}}{p_{\tilde{A}\tilde{B}}}.$$

Полагая $a_1 = a_2 = 1$, $a_0 = -1$, $\Delta = e^{\theta^*} = (p_{\tilde{A}B} p_{A\tilde{B}}) / (p_{AB} p_{\tilde{A}\tilde{B}})$ и обозначая $p_A = p_{AB} + p_{A\tilde{B}}$ и $p_B = p_{AB} + p_{\tilde{A}B}$ вероятности категорий A и B в рассматриваемой совокупности, мы имеем

$$p_{AB} = p_A p_B + \frac{1 - \Delta}{\Delta} p_{\tilde{A}B} p_{A\tilde{B}},$$

$$p_{\tilde{A}B} = p_{\tilde{A}} p_B - \frac{1 - \Delta}{\Delta} p_{\tilde{A}B} p_{A\tilde{B}},$$

$$p_{A\tilde{B}} = p_A p_{\tilde{B}} - \frac{1 - \Delta}{\Delta} p_{\tilde{A}B} p_{A\tilde{B}},$$

$$p_{\tilde{A}\tilde{B}} = p_{\tilde{A}} p_{\tilde{B}} + \frac{1 - \Delta}{\Delta} p_{\tilde{A}B} p_{A\tilde{B}}.$$

Независимость A и B эквивалентна равенству $\Delta = 1$, а неравенства $\Delta < 1$ и $\Delta > 1$ указывают соответственно на положительную и отрицательную зависимость *).

Критерий гипотезы независимости, или любой другой гипотезы относительно Δ , описывается в терминах условного распределения X при данных $X + X' = m$, $X + Y = t$. Вместо того чтобы непосредственно вычислять это распределение, рассмотрим сначала условное распределение при условии $X + X' = m$ (которое влечет $Y + Y' = s - m = n$). Это распределение, как легко видеть, равно

$$P\{X = x, Y = y | X + X' = m\} =$$

$$= \binom{m}{x} \binom{n}{y} \left(\frac{p_{AB}}{p_B} \right)^x \left(\frac{p_{\tilde{A}B}}{p_B} \right)^{m-x} \left(\frac{p_{A\tilde{B}}}{p_{\tilde{B}}} \right)^y \left(\frac{p_{\tilde{A}\tilde{B}}}{p_{\tilde{B}}} \right)^{n-y},$$

т. е. представляет собой совместное распределение двух независимых биномиальных величин с параметрами m , $p_1 = p_{AB}/p_B$, и n , $p_2 = p_{A\tilde{B}}/p_{\tilde{B}}$ соответственно. Действительно, это ясно и без вычислений, поскольку теперь мы имеем дело с выборками фиксированных объемов m и n из категорий B и $\tilde{B}$, вероятность A в которых равна p_1 и p_2 . Если наложить дополнительное условие $X + Y = t$, то условное распределение X при двух условиях

*) Δ эквивалентна мере зависимости признаков $Q = (1 - \Delta)/(1 + \Delta)$, предложенной Юлом. Эта и подобные меры обсуждаются в статье: Goodman and Kruskal, Measures of association for error classifications, J. Am. Stat. Assoc. 49 (1954), 732—764. См. также М. Дж. Кендалл, А. Стьюарт, Статистические выводы и связи, М., 1973, гл. 33. — Прим. пер.

$X + X' = m$ и $X + Y = t$ будет таким же, как у X при данном $X + Y = t$ в случае двух независимых биномиальных величин, рассмотренных в предыдущем параграфе. Стало быть, оно равно

$$P\{X = x | X + X' = m, X + Y = t\} = C_t(\rho) \binom{m}{x} \binom{n}{t-x} \rho^{t-x}, \quad x = 0, \dots, t$$

(мы заменили в (21) y на x ; следует заметить, что выбор именно X для построения критериев совершенно произволен; с таким же успехом мы могли бы выбрать Y). Для параметра ρ имеем выражение

$$\rho = \frac{p_2/p_1}{q_2/q_1} = \frac{p_{\tilde{A}B} p_{A\tilde{B}}}{p_{AB} p_{\tilde{A}\tilde{B}}} = \Delta.$$

Из проведенных рассуждений мы видим, что условный критерий при данных $X + X' = m$ и $X + Y = t$ для проверки любой гипотезы, относящейся к Δ , идентичен с условным критерием при данном $X + Y = t$ для той же самой гипотезы относительно $\rho = \Delta$ в обстановке предыдущего раздела (в котором $X + X' = m$ было дано априори). В частности, условный критерий для проверки независимости $\Delta = 1$ (критерий Фишера — Ирвина) совпадает с критерием проверки равенства двух биномиальных параметров ρ и потому может быть выражен в терминах гипергеометрического распределения.

В начале раздела уже отмечалось, что гипотеза независимости может быть проверена на основании выборок, взятых различными способами. Выборки фиксированного объема можно взять из A и $\tilde{A}$ или из B и $\tilde{B}$, или, наконец, из совокупности как целого. Какой из этих планов более эффективен, зависит от стоимости выбора из различных категорий и из всей совокупности, а также от стоимости необходимой классификации отобранных индивидуумов в соответствии с рассматриваемыми характеристиками. Предположим, однако, на время, что мы пренебрегаем этой стороной дела и желаем сравнить планы эксперимента в терминах мощностей отвечающих им критериев при одной и той же альтернативе. Тогда при возрастании объема выборки s до бесконечности верны следующие асимптотические утверждения *):

(1) Если выборки объемов m и n ($m + n = s$) извлекаются из B и $\tilde{B}$ или A и $\tilde{A}$, то наилучшими значениями m и n будут $m = n = s/2$.

*) Справедливость этих результатов была предположена Берксоном и доказана Нейманом в курсе лекций о χ^2 -критерии.

(II) Выборки равного объема $s/2$ лучше извлекать из B и $\bar{B}$, чем из A и $\bar{A}$, если $\left| p_B - \frac{1}{2} \right| > \left| p_A - \frac{1}{2} \right|$.

(III) Случайная выборка из всей совокупности хуже, чем выборки равного объема из A и $\bar{A}$ или из B и $\bar{B}$.

Мы не будем доказывать здесь эти утверждения. Отметим лишь, что их можно получить, используя нормальное приближение к биномиальным распределениям X и Y , и замечая, что при случайном выборе из всей совокупности M/s и N/s стремятся по вероятности к p_B и $p_{\bar{B}}$ соответственно.

7. Критерий знаков

Чтобы определить, какой из двух товаров предпочитает потребитель, опрашивают n лиц. Результат отмечается знаком плюс, когда предпочитают товар A , и знаком минус в другом случае. Полное число плюсов Y имеет тогда биномиальное распределение $b(p, n)$. Рассмотрим проблему проверки гипотезы $p = 1/2$ об отсутствии разницы в спросе при альтернативе $p \neq 1/2$ (как и раньше, мы отвлекаемся от того факта, что при отклонении гипотезы необходимо решить, какой из товаров предпочитают). Соответствующий критерий, двусторонний критерий знаков, отклоняет гипотезу, если значение $\left| Y - \frac{1}{2} n \right|$ слишком велико. Это — РНМ несмещенный критерий (раздел 2).

Допустим, что потребители могут воздержаться от ответа. Обозначим p_+ , p_- и p_0 вероятности предпочесть A , предпочесть B и воздержаться соответственно, тогда числа X , Y и Z потребителей, для которых осуществляются указанные возможности, имеют совместное полиномиальное распределение с вероятностями

$$\frac{n!}{x! y! z!} p_+^x p_-^y p_0^z \quad (x + y + z = n), \quad (22)$$

и гипотеза, которую надлежит проверить, имеет вид $H: p_+ = p_-$. Распределение (22) можно записать в форме

$$\frac{n!}{x! y! z!} \left(\frac{p_+}{1 - p_0 - p_+} \right)^y \left(\frac{p_0}{1 - p_0 - p_+} \right)^z (1 - p_0 - p_+)^n,$$

из которой видно, что это — экспоненциальное семейство с $U = Y$, $T = Z$, $\theta = \log [p_+ / (1 - p_0 - p_+)]$, $\phi = \log [p_0 / (1 - p_0 - p_+)]$. Гипотеза H , как легко видеть, эквивалентна гипотезе $\theta = 0$. Поэтому существует РНМ несмещенный критерий для H . Мы придем к этому критерию, если сначала фиксируем z , а затем определим наилучший несмещенный условный критерий для H при $Z = z$. Так как условное распределение Y при данном z

является биномиальным $b(p, n - z)$ с $p = p_+ / (p_+ + p_-)$, то проблема сводится к проверке гипотезы $p = 1/2$ в биномиальном распределении с $n - z$ испытаниями, а в этом случае критическая область имеет вид $\left| Y - \frac{1}{2}(n - z) \right| > C(z)$. Мы получаем, следовательно, РНМ критерий, отбрасывая случаи, когда предположение не было высказано («ничьи»), и применяя критерий знаков к оставшимся данным.

Мощность критерия сильно зависит от p_0 , определяющего распределение величины Z . Можно ожидать, что при больших p_0 число $n - z$ испытаний в условном биномиальном распределении будет мало, и поэтому мощность критерия также будет невелика. Правда, в этом случае есть и преимущество — большое значение p_0 указывает, безотносительно к величине p_+ / p_- , что потребители в целом весьма безразличны к товарам, о которых идет речь.

Вышеприведенный условный критерий знаков применим в любой ситуации, где наблюдаются n независимых испытаний, каждое из которых заканчивается успехом (+), неудачей (—) или ничьей. В качестве альтернативного способа обращения с ничьими иногда предлагают приписать каждой ничьей наудачу (с вероятностями $1/2$) знак плюс или минус. После такого «расщепления» ничьих полное число плюсов становится биномиальной случайной величиной с распределением $b(\pi, n)$, где $\pi = p_+ + \frac{1}{2} p_0$. Гипотеза H заменяется гипотезой $\pi = 1/2$ и отклоняется, когда $\left| Y' - \frac{1}{2}n \right| > C$. C определяется тем, что вероятность отклонения должна быть равна α при $\pi = 1/2$. Этот критерий можно рассматривать как рандомизированный критерий, основанный на X , Y и Z . Как критерий для H , он не смещен, поскольку $p_+ = p_-$ или $p_+ \neq p_-$ при $\pi = 1/2$ или $\pi \neq 1/2$. Так как критерий включает рандомизацию не только на границе критической области, то он менее мощен, чем РНМ несмещенный критерий для этой ситуации. Таким образом, случайное «расщепление» ничьих приводит к потере мощности.

Можно полагать, что это замечание поможет ответить на следующий вопрос: что лучше делать при определении вкусов потребителей — позволять им воздержаться или настаивать на выражении ими определенного мнения. Однако здесь неприменимо предположение о полной случайности «расщепления» ничьих. Даже в случаях, когда индивидуум полностью не уверен в том, что именно он предпочитает, обычно имеется небольшая склонность к одной из двух возможностей. Она в большинстве случаев обнаруживается при вынужденном решении. Но эта сторона дела уравнивается большей изменчивостью вынужденных

решений по сравнению с добровольными. Какой из этих двух факторов главный, зависит от степени предпочтения.

Часто вопрос о предпочтении возникает, когда сравниваются стандартный продукт и некоторая его модификация или новый продукт. Если каждому опрашиваемому необходимо выразить определенное мнение, то гипотеза формулируется обычно как односторонняя: $p_+ \leq p_-$, где $+$ означает, что предпочитают модификацию. Однако, если можно воздерживаться от ответа, то проверяемой гипотезой будет не $p_+ \leq p_-$, а $p_+ \leq p_0 + p_-$, поскольку, как правило, модификация интересна, если ее действительно предпочитают. Как было показано в примере 8 главы 3, односторонний критерий знаков, который отвергает гипотезу при слишком большом числе плюсов, является РНМ для этой задачи.

В некоторых исследованиях опрашиваемое лицо должно не только выразить свое предпочтение, но и дать более детальную оценку, например указать число из некоторой шкалы оценок. В зависимости от ситуации, гипотеза может принимать одну из двух форм. Можно интересоваться тем, имеется ли различие в отношении потребителей к двум видам товаров. Формально соответствующая гипотеза утверждает, что распределение оценок $X_1, \dots, X_n$, где X_i выражает степень предпочтения i -м лицом модифицированного продукта, симметрично относительно начала. Эта задача, в которой РНМ несмещенный критерий без дополнительных предположений может не существовать, будет рассмотрена в разделе 9 главы 6.

При ином подходе гипотеза имеет прежний вид $H: p_+ = p_-$. Так как $p_- = P\{X < 0\}$ и $p_+ = P\{X > 0\}$, то

$$H: P\{X > 0\} = P\{X < 0\}.$$

Здесь симметрия величин X более не предполагается, даже в случае $P\{X < 0\} = P\{X > 0\}$. Если не делать никаких предположений о распределении X , кроме того, что множество его возможных значений известно, то критерий знаков, основанный на числе положительных и отрицательных X_i , продолжает быть РНМ несмещенным.

Для доказательства заметим, что любое распределение X полностью определяется заданием вероятностей

$$p_- = P\{X < 0\}, \quad p_+ = P\{X > 0\}, \quad p_0 = P\{X = 0\}$$

и условных распределений F_- и F_+ величины X при условиях $X < 0$ и $X > 0$ соответственно. Фиксируем любые два распределения F'_-, F'_+ и обозначим $\mathcal{F}_0$ семейство всех распределений с $F_- = F'_-, F_+ = F'_+$ и произвольными p_-, p_+, p_0 . Каждый критерий для H , который не смещен в первоначальном семействе рас-

пределений $\mathcal{F}$, где F_- и F_+ неизвестны, остается несмещенным и в более узком семействе $\mathcal{F}_0$. Ниже мы покажем, что существует РНМ несмещенный критерий φ_0 для H в классе $\mathcal{F}_0$. Оказывается, что φ_0 является также несмещенным критерием для проверки H в $\mathcal{F}$ и не зависит от F'_- , F'_+ . Пусть φ — любой другой несмещенный критерий для гипотезы H в $\mathcal{F}$. Рассмотрим какую-либо фиксированную альтернативу. Не ограничивая общности, можно предположить, что она входит в $\mathcal{F}_0$. Так как критерий φ не смещен в $\mathcal{F}$, он не смещен и для гипотезы $p_+ = p_-$, проверяемой в $\mathcal{F}_0$. Следовательно, мощность φ_0 при выбранной альтернативе не меньше мощности φ . Поэтому φ_0 — РНМ несмещенный критерий.

Чтобы построить РНМ несмещенный критерий для H в $\mathcal{F}_0$, обозначим f'_- и f'_+ плотности распределений F_- и F_+ по отношению к некоторой мере μ . Совместная плотность величин X в точке $(x_1, \dots, x_n)$ с

$$x_{i_1}, \dots, x_{i_r} < 0 = x_{j_1} = \dots = x_{j_s} < x_{k_1}, \dots, x_{k_m}$$

равна

$$p_-^{r-s} p_+^s f'_-(x_{i_1}) \dots f'_-(x_{i_r}) f'_+(x_{k_1}) \dots f'_+(x_{k_m}).$$

Совокупность статистик (r, s, m) достаточна для (p_-, p_0, p_+) и ее распределение дается формулой (22) с $x = r$, $y = m$, $z = s$. Мы видим, что, как и раньше, критерий знаков является РНМ несмещенным.

8. Задачи

К разделу 1

1. *Допустимость.* Любой РНМ несмещенный критерий φ_0 является допустимым, т. е. не существует другого критерия φ_1 , который был бы не менее мощен, чем φ_0 при всех альтернативах, и более мощен при некоторой из них.

[Если критерий φ несмещенный и φ' равномерно не хуже (в смысле мощности), чем φ , то φ' также несмещенный].

2. *Критические уровни.* Рассмотрим семейство критериев для H : $\theta = \theta_0$ (или $\theta \leq \theta_0$) уровня α , с критическими областями S_α такими, что (а) $P_{\theta_0}\{X \in S_\alpha\} = \alpha$ для всех $0 < \alpha < 1$ и (б) $S_{\alpha_0} = \bigcap_{\alpha > \alpha_0} S_\alpha$ для всех $0 < \alpha_0 < 1$ (отсюда вытекает, в частности, что $S_\alpha \subset S_{\alpha'}$ при $\alpha < \alpha'$).

(I) Тогда критический уровень равен $\hat{\alpha} = \hat{\alpha}(x) = \inf\{\alpha: x \in S_\alpha\}$.

(II) При $\theta = \theta_0$ распределение $\hat{\alpha}$ равномерно на $(0, 1)$.

(III) Если критерии S_α -несмещенные, то распределение $\hat{\alpha}$ при любой альтернативе θ удовлетворяет неравенству

$$P_\theta\{\hat{\alpha} \leq \alpha\} \geq P_{\theta_0}\{\hat{\alpha} \leq \alpha\} = \alpha,$$

т. е. как бы «сдвинуто» к началу координат.

Если критические числа определяются из нескольких независимых экспериментов, то в сочетании с (II) и (III) можно получить безусловный критерий*) для рассматриваемой гипотезы $\{\hat{\alpha} \leq \alpha$ тогда и только тогда, когда $x \in S_\alpha$, следовательно, $P_\theta\{\hat{\alpha} \leq \alpha\} = P_\theta\{X \in S_\alpha\} = \beta_\alpha(\theta)$, что равно α при $\theta = \theta_0$ и $\geq \alpha$, если θ — альтернатива H_1].

К разделу 2

3. Пусть X имеет биномиальное распределение $b(p, n)$. Рассмотрим гипотезу $H: p = p_0$ при уровне значимости α . Определите границы РНМ несмещенного критерия для $n = 10$; $\alpha = 0,1$; $p_0 = 0,2$ и $\alpha = 0,05$; $p_0 = 0,4$. В каждом случае постройте функции мощности несмещенного критерия и критерия с равными «хвостами».

4. Пусть X имеет распределение Пуассона $P(\tau)$. Рассмотрим гипотезу $H: \tau = \tau_0$. Тогда условие (6) сводится к

$$\sum_{x=C_1+1}^{C_2-1} \frac{\tau_0^{x-1}}{(x-1)!} e^{-\tau_0} + \sum_{i=1}^2 (1 - \gamma_i) \frac{\tau_0^{C_i-1}}{(C_i-1)!} e^{-\tau_0} = 1 - \alpha.$$

5. Пусть T_n/θ имеет χ^2 -распределение с n степенями свободы. Найдите, сколь большим должно быть n , чтобы мощность РНМ несмещенного критерия для проверки $H: \theta = 1$ с уровнем значимости $\alpha = 0,05$ была $\geq 0,9$ как при $\theta \geq 2$, так и при $\theta \leq 1/2$. Сколь велико должно быть n , если не требовать несмещенности критерия?

6. Пусть X и Y независимы и распределение каждой величины входит в однопараметрическое экспоненциальное семейство, так что совместное распределение дается формулой

$$dP_{\theta_1, \theta_2}(x, y) = C(\theta_1) e^{\theta_1 T(x)} d\mu(x) K(\theta_2) e^{\theta_2 U(y)} d\nu(y).$$

Тогда для проверки $H: \theta_1 = a, \theta_2 = b$ при альтернативе $\theta_1 \neq a$ или $\theta_2 \neq b$ не существует РНМ несмещенного критерия.

[Наиболее мощные несмещенные критерии при альтернативах $\theta_1 = a, \theta_2 \neq b$ и $\theta_1 \neq a, \theta_2 = b$ имеют соответственно области принятия $K_1 < U(y) < K_2$ и $C_1 < T(x) < C_2$. Эти критерии остаются несмещенными и при более широком классе альтернатив $K: \theta_1 \neq a$ или $\theta_2 \neq b$, или и то и другое.]

7. Допустим, что распределение (X, Y) входит в экспоненциальное семейство

$$dP_{\theta_1, \theta_2}(x, y) = C(\theta_1, \theta_2) e^{\theta_1 x + \theta_2 y} d\mu(x, y).$$

Единственным несмещенным критерием для проверки $H: \theta_1 \leq a, \theta_2 \leq b$ при альтернативах $K: \theta_1 > a$ или $\theta_2 > b$, или и то и другое, является $\varphi(x, y) = \alpha$.

[Возьмем $a = b = 0$. Пусть $\beta(\theta_1, \theta_2)$ — мощность какого-либо критерия уровня α . Несмещенность влечет $\beta(0, \theta_2) = \alpha$ для $\theta_2 < 0$ и, следовательно, для всех θ_2 , так как $\beta(0, \theta_2)$ — аналитическая функция θ_2 . При фиксированном $\theta_2 > 0$ $\beta(\theta_1, \theta_2)$, рассматриваемая как функция θ_1 , имеет минимум при $\theta_1 = 0$, так что производная $\partial\beta(\theta_1, \theta_2)/\partial\theta_1$ равна нулю при $\theta_1 = 0$ и всех положительных θ_2 , а следовательно, и при всех θ_2 . Рассматривая положительные и отрицательные значения θ_2 и используя тот факт, что частные производные всех порядков от $\beta(\theta_1, \theta_2)$ по θ_1 аналитичны, мы находим, что при каждом фиксированном θ_2 эти производные все обращаются в нуль в точке $\theta_1 = 0$.

*) Обсуждение нескольких таких критериев см. в работах: Wallis, Combining probabilities from independent significance tests, *Econometrica* vol. 10 (1942), 229—248 и Birnbaum, Combining independent tests of significance, *J. Am. Stat. Assoc.*, vol. 49 (1954), 559—574.

Поэтому функция β должна быть постоянной. В силу полноты (X, Y) из $\beta((\theta_1, \theta_2) \equiv \alpha$ вытекает $\varphi(x, y) \equiv \alpha$].

8. Рассмотрим задачу проверки гипотезы $H: \theta = \theta_0$ в однопараметрическом экспоненциальном семействе раздела 2. Обозначим $\mathcal{C}$ класс всех критериев, удовлетворяющих (3) и (5) при некоторых $-\infty \leq C_1 \leq C_2 \leq \infty$ и $0 \leq \gamma_1, \gamma_2 \leq 1$.

(I) Класс $\mathcal{C}$ полон в том смысле, что, каков бы ни был критерий φ_0 для H уровня α , найдется $\varphi \in \mathcal{C}$ такой, что он равномерно не менее мощен, чем φ_0 .

(II) Если $\varphi_1, \varphi_2 \in \mathcal{C}$, то ни один из этих критериев не может быть равномерно более мощным, чем другой.

(III) Рассмотрим задачу с двумя решениями d_0 и d_1 , соответствующими принятию и отклонению H , и с функциями потерь $L(\theta, d_i) = L_i(\theta)$, $i = 0, 1$. В предположении, что $L_1(\theta) < L_0(\theta)$ при всех $\theta \neq \theta_0$, $\mathcal{C}$ является существенно полным классом.

(IV) Распространите результат части (III) на гипотезу $H': \theta_1 \leq \theta \leq \theta_2$.

[(I) Обозначим ρ производную функции мощности β_{φ_0} критерия φ_0 в точке $\theta = \theta_0$. Тогда найдется критерий $\varphi \in \mathcal{C}$ такой, что $\beta'_\varphi(\theta_0) = \rho$, и φ является РНМ среди всех критериев с этим свойством.

(II) См. конец раздела 7 главы 3.

(III) См. доказательство теоремы 3 в главе 3.]

К разделу 3

9. Пусть $X_1, \dots, X_n$ — выборка из (I) нормального распределения $N(a, \sigma^2)$ с фиксированным a и $0 < \sigma < \infty$; (II) равномерного распределения $R\left(\theta - \frac{1}{2}, \theta + \frac{1}{2}\right)$, $-\infty < \theta < \infty$; (III) равномерного распределения $R(\theta_1, \theta_2)$, $-\infty < \theta_1 < \theta_2 < \infty$. В этих трех семействах распределений следующие статистики являются достаточными: (I) $T = (\Sigma X_i, \Sigma X_i^2)$, (II) и (III) $T = (\min(X_1, \dots, X_n), \max(X_1, \dots, X_n))$. Семейство распределений T полно в случае (III), но в случаях (I) и (II) не является ни полным, ни даже ограниченно полным.

[(I) Распределение отношения $\Sigma X_i / \sqrt{\Sigma X_i^2}$ не зависит от σ .]

10. Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из $N(\xi, \sigma^2)$ и $N(\xi, \tau^2)$. В примере 5 было показано, что $T = (\Sigma X_i, \Sigma Y_j, \Sigma X_i^2, \Sigma Y_j^2)$ не является полной. T не будет и ограниченно полной.

[Введите $f(i)$, равное 1 или -1 в зависимости от того, положительно $\bar{y} - \bar{x}$ или нет.]

11. *Противоречащий пример.* Пусть случайная величина X принимает значения $-1, 0, 1, 2, \dots$ с вероятностями

$$P_\theta\{X = -1\} = \theta; \quad P_\theta\{X = x\} = (1 - \theta)^2 \theta^x, \quad x = 0, 1, \dots$$

Тогда $\mathcal{P} = \{P_\theta, 0 < \theta < 1\}$ не будет полным, но будет ограниченно полным.

12. Пусть $\mathcal{P} = \{P\}$ — семейство распределений с тем свойством, что для любых $P, Q \in \mathcal{P}$ найдется $0 < p < 1$ такое, что $pP + (1 - p)Q \in \mathcal{P}$. Предположим, что $h(x_1, \dots, x_n)$ — симметричная функция, для которой

$$\int h(x_1, \dots, x_n) dP(x_1) \dots dP(x_n) = 0 \quad \text{при всех } P \in \mathcal{P}. \quad (23)$$

Тогда

$$\int h(x_1, \dots, x_n) dP_1(x_1) \dots dP_n(x_n) = 0 \quad \text{при всех } P_1, \dots, P_n \in \mathcal{P}. \quad (24)$$

[(1) Если $P_1, \dots, P_k \in \mathcal{P}$, то существуют вероятности $p_1, \dots, p_k$, в сумме дающие единицу и такие, что $(p_1 P_1 + \dots + p_k P_k) / (p_1 + \dots + p_k) \in \mathcal{P}$ при всех $i = 1, \dots, k$.

(2) Для любых целых $1 \leq i_1 < i_2 < \dots < i_k \leq n$ обозначим $\alpha(i_1, \dots, i_k)$ множество всех n -наборов $(j_1, \dots, j_n)$ таких, что (а) каждая компонента совпадает с одним из $i_1, \dots, i_k$ и (б) каждое из этих целых встречается хотя бы раз среди $(j_1, \dots, j_n)$. Если

$$I(j_1, \dots, j_n) = p_{j_1} \dots p_{j_n} \int h(x_1, \dots, x_n) dP_{j_1}(x_1) \dots dP_{j_n}(x_n),$$

то в силу (23)

$$\sum_{\alpha(i_1, \dots, i_k)} I(j_1, \dots, j_n) = 0$$

для всех $(i_1, \dots, i_k)$ с $k \leq n$.

Это утверждение доказывается индукцией по k . При $k = 1$ оно следует непосредственно из (23). Чтобы доказать, например, что $\sum_{\alpha(1, 2)} I(j_1, \dots, j_n) = 0$, обозначим $P = (p_1 P_1 + p_2 P_2) / (p_1 + p_2)$ элемент $\mathcal{P}$, существование которого гарантируется (1). Тогда

$$\begin{aligned} 0 &= \int h(x_1, \dots, x_n) dP(x_1) \dots dP(x_n) (p_1 + p_2)^n = \\ &= \sum_{\alpha(1, 2)} I(j_1, \dots, j_n) + \sum_{\alpha(1) \cup \alpha(2)} I(j_1, \dots, j_n), \end{aligned}$$

и, так как второй член в правой части равен нулю, мы получаем желаемый результат.

(3) Из (2) при $k = n$ следует, что $\sum I(j_1, \dots, j_n) = 0$, где суммирование распространяется на все перестановки $(j_1, \dots, j_n)$ чисел $(1, \dots, n)$. Так как I симметрично по своим аргументам, отсюда вытекает $I(1, \dots, n) = 0$, что и требовалось доказать.]

13. *Продолжение.* Пусть $\mathcal{C}$ — класс всех равномерных распределений на конечном интервале и $\mathcal{P}$ — класс конечных выпуклых комбинаций распределений из $\mathcal{C}$. Если $X_1, \dots, X_n$ независимы и распределены по закону $P \in \mathcal{P}$, то совокупность порядковых статистик $T = (X^{(1)}, \dots, X^{(n)})$ достаточна для $\mathcal{P}$, и семейство $\mathcal{P}^T$ распределений T -полно.

[Достаточность T видна из примера 7 главы 2. Для доказательства полноты рассмотрите уравнение $E_P h(T) = 0$ для всех $P \in \mathcal{P}$ и примените результаты предыдущей задачи.]

14. *Измеримость критериев теоремы 3.* Функция Φ_3 , определяемая (16) и (17), измерима по совокупности переменных u и t .

[Обозначим $C_1 = v$, $C_2 = w$. Определяющие уравнения для v , w , γ_1 , γ_2 имеют вид

$$F_t(v^-) + [1 - F_t(w)] + \gamma_1 [F_t(v) - F_t(v^-)] + \gamma_2 [F_t(w) - F_t(w^-)] = \alpha \quad (25)$$

и

$$G_t(v^-) + [1 - G_t(w)] + \gamma_1 [G_t(v) - G_t(v^-)] + \gamma_2 [G_t(w) - G_t(w^-)] = \alpha, \quad (26)$$

где

$$F_t(u) = \int_{-\infty}^u C_t(\theta_1) e^{\theta_1 y} d\nu_t(y), \quad G_t(u) = \int_{-\infty}^u C_t(\theta_2) e^{\theta_2 y} d\nu_t(y) \quad (27)$$

— условные функции распределения U при данном t , когда $\theta = \theta_1$ и $\theta = \theta_2$ соответственно.

(1) Для каждого $0 \leq y \leq \alpha$ положим $v(y, t) = F_t^{-1}(y)$ и $w(y, t) = F_t^{-1}(1 - \alpha + y)$, где обратная функция определена как при доказательстве теоремы 3. Выберем такие $\gamma_1(y, t)$ и $\gamma_2(y, t)$, что для $v = v(y, t)$ и $w = w(y, t)$

$$\begin{aligned} F_t(v^-) + \gamma_1[F_t(v) - F_t(v^-)] &= y, \\ 1 - F_t(w) + \gamma_2[F_t(w) - F_t(w^-)] &= \alpha - y. \end{aligned}$$

(2) Обозначим $H(y, t)$ левую часть (26) с $v = v(y, t)$ и т. д. Тогда $H(0, t) > \alpha$ и $H(\alpha, t) < \alpha$. Это следует по теореме 2 главы 3 из того, что $v(0, t) = -\infty$, $w(\alpha, t) = \infty$ (т. е. что условные критерии при $y = 0$ и $y = \alpha$ являются односторонними) и что левая часть (26) при любом y есть функция мощности условного критерия.

(3) При фиксированном t функции

$$H_1(y, t) = G_t(v^-) + \gamma_1[G_t(v) - G_t(v^-)]$$

и

$$H_2(y, t) = 1 - G_t(w) + \gamma_2[G_t(w) - G_t(w^-)]$$

суть непрерывные функции y . Действительно, как видно из (27), разрывы и горизонтальные участки функций F_t и G_t совпадают $\mathcal{P}^T$ -почти всюду, а отсюда следует высказанное утверждение.

(4) Функция $H(y, t)$ измерима по совокупности переменных y и t . Это можно вывести из непрерывности H с помощью рассуждений, сходных с использованными в тексте при доказательстве измеримости $F_t(u)$. Положим

$$y(t) = \inf \{y: H(y, t) < \alpha\},$$

и пусть $v(t) = v[y(t), t]$ и т. д. Тогда (25) и (26) выполняются при всех t . Измеримость $v(t)$, $w(t)$, $\gamma_1(t)$ и $\gamma_2(t)$, введенных таким путем, будет следовать из измеримости по t функций $y(t)$ и $F_t^{-1}[y(t)]$. Последнее же видно из выполняющихся при любом действительном c соотношений

$$\{t: y(t) < c\} = \bigcup_{r < c} \{t: H(r, t) < \alpha\},$$

где r пробегает все рациональные числа, и

$$\{t: F_t^{-1}[y(t)] \leq c\} = \{t: y(t) - F_t(c) \leq 0\}.$$

15. *Продолжение.* Функция φ_* , определенная (16), (18) и (19), измерима по совокупности переменных u и t .

[Доказательство, в общем сходное с намеченным в предыдущей задаче, опирается, в частности, на измеримость по z и t интеграла

$$g(z, t) = \int_{-\infty}^{z-} u dF_t(u).$$

Этот интеграл абсолютно сходится для всех t , поскольку F_t принадлежит экспоненциальному семейству. При любом $z < \infty$ $g(z, t) = \lim g_n(z, t)$, где

$$g_n(z, t) = \sum_{j=1}^{\infty} \left(z - \frac{j}{2^n} \right) \left[F_t \left(z - \frac{j-1}{2^n} - 0 \right) - F_t \left(z - \frac{j}{2^n} - 0 \right) \right],$$

и измеримость g следует из измеримости функций g_n . Неравенства, соответствующие найденным в части 2 предыдущей задачи, можно вывести из свойства условных односторонних критериев, сформулированного в задаче 18 главы 3.]

16. РНМ несмещенные критерии для гипотез $H_1, \dots, H_4$, указанные в теореме 3, являются единственными в классе критериев, основанных на U и T .

К разделу 5

17. Пусть X и Y независимы и имеют распределения Пуассона $P(\lambda)$ и $P(\mu)$. Найдите мощность РНМ несмещенного критерия для гипотезы $H: \mu \leq \lambda$ при альтернативах $\lambda = 0,1, \mu = 0,2; \lambda = 1, \mu = 2; \lambda = 10, \mu = 20; \lambda = 0,1; \mu = 0,4$. Уровень значимости $\alpha = 0,1$.

[Так как $T = X + Y$ имеет распределение Пуассона $P(\lambda + \mu)$, то мощность равна

$$\beta = \sum_{t=0}^{\infty} \beta(t) \frac{(\lambda + \mu)^t}{t!} e^{-(\lambda + \mu)},$$

где $\beta(t)$ — мощность условного критерия при данном t и рассматриваемой альтернативе.]

18. *Последовательное сравнение двух биномиальных схем.* Рассмотрим две серии биномиальных испытаний с вероятностями успеха p_1 и p_2 соответственно, и пусть $\rho = (p_2/q_2) : (p_1/q_1)$.

(I) Допустим $\alpha < \beta$. Никакой критерий с фиксированными числами m и n испытаний для проверки $H: \rho = \rho_0$ не может иметь мощность $\geq \beta$ при всех альтернативах с $\rho = \rho_1$.

(II) Ниже указывается простая последовательная процедура, приводящая к желаемому результату. Допустим, что испытания в обеих сериях идут параллельно, парами. Рассмотрим пары, в которых одно испытание закончилось успехом, а другое неудачей. Если эксперимент продолжается до появления N -й такой пары, то число пар, в которых успешное испытание относится к первой серии, имеет биномиальное распределение $b(\pi, N)$ с $\pi = p_1 q_2 / (p_1 q_2 + p_2 q_1) = 1/(1 + \rho)$. Взяв N достаточно большим, можно получить критерий сколь угодно высокой мощности при альтернативе ρ_1 .

(III) Пары испытаний, которыми мы ограничиваемся в (II), подчиняются биномиальной схеме с вероятностью успеха π . Отличная от указанной в (II) процедура проверки $H: \pi = \pi_0$ (или $\pi \leq \pi_0$) состоит в применении последовательного критерия отношений вероятностей, основанного на отмеченных парах и соответствующего проверке $\pi = \pi_0$ при альтернативе $\pi = \pi_1$.

К разделу 6

19. *Серии.* Рассмотрим последовательность N зависимых испытаний, и пусть X_i равно 1 или 0 в зависимости от того, заканчивается i -е испытание успехом или неудачей. Допустим, что последовательность обладает *марковским* свойством *)

$$P\{X_i = 1 \mid x_1, \dots, x_{i-1}\} = P\{X_i = 1 \mid x_{i-1}\}$$

и свойством *стационарности*, в соответствии с которым $P\{X_i = 1\}$ и $P\{X_i = 1 \mid x_{i-1}\}$ не зависят от i . В этом случае распределение X_i полностью задается вероятностями

$$p_1 = P\{X_i = 1 \mid X_{i-1} = 1\} \quad \text{и} \quad p_0 = P\{X_i = 1 \mid X_{i-1} = 0\}$$

*) Современное обсуждение статистических проблем для более сложных цепей Маркова см. в работах: Anderson and Goodman, Statistical inference about Markov chains, Ann. Math. Stat. 28 (1957), 89—110 и Goodman, Simplified runs tests and likelihood ratio tests for Markoff chains, Biometrika 45 (1958), 181—197. См. также Billingsley P., Statistical methods in Markov Chains, Ann. Math. Stat. 32, № 1 (1961).

и начальными вероятностями

$$\pi_1 = P\{X_1 = 1\} \quad \text{и} \quad \pi_0 = 1 - \pi_1 = P\{X_1 = 0\}.$$

(I) Стационарность влечет

$$\pi_1 = \frac{p_0}{(p_0 + q_1)}, \quad \pi_0 = \frac{q_1}{(p_0 + q_1)}.$$

(II) Совокупность последовательных исходов $x_i, \dots, x_{i+j}$ образует *серию* нулей, если $x_i = x_{i+1} = \dots = x_{i+j} = 0$ и $x_{i-1} = 1$ (или $i = 1$) и $x_{i+j+1} = 1$ (или $i + j = N$). Аналогично определяется серия единиц. Вероятность какой-либо отдельной последовательности исходов $(x_1, \dots, x_N)$ равна

$$\frac{1}{p_0 + q_1} p_0^v p_1^{n-v} q_1^u q_0^{m-u},$$

где m и n обозначают число нулей и единиц, а u и v — число серий нулей и единиц в этой последовательности.

20. *Продолжение.* Для проверки гипотезы независимости X_i H : $p_0 = p_1$ при альтернативе K : $p_0 < p_1$ применим критерий числа серий, который отвергает H , когда полное число серий $R = U + V$ меньше константы $C(m)$, зависящей от числа m нулей в последовательности. При $R = C(m)$ гипотеза отвергается с вероятностью $\gamma(m)$. C и γ определяются уравнением

$$P_H\{R < C(m) | m\} + \gamma(m) P_H\{R = C(m) | m\} = \alpha.$$

(I) При любой альтернативе из K наиболее мощный подобный критерий (который не менее мощен, чем наиболее мощный несмещенный критерий) совпадает с критерием числа серий, отвергающим H при $R < C(m)$. Только правило дополнения до α условной вероятности отклонения гипотезы при данном m зависит от выбора альтернативы.

(II) Критерий числа серий является несмещенным при альтернативах из K .

(III) Условное распределение R при данном m в случае, когда H истинна, равно *)

$$P\{R = 2r\} = \frac{r \binom{m-1}{r-1} \binom{n-1}{r-1}}{\binom{m+n}{m}};$$

$$P\{R = 2r + 1\} = \frac{\binom{m-1}{r-1} \binom{n-1}{r} + \binom{m-1}{r} \binom{n-1}{r-1}}{\binom{m+n}{m}}.$$

[(I) Из несмещенности вытекает, что условная вероятность отклонения гипотезы при данном m равна α при всех m . Наиболее мощный условный критерий уровня α отвергает H для тех выборочных последовательностей, для которых $\Delta(u, v) = (p_0/p_1)^v (q_1/q_0)^u$ слишком велико. Так как $p_0 < p_1$ и $q_1 < q_0$ и поскольку $|v - u|$ может принимать только значения 0 и 1, то

$$\Delta(1, 1) > \Delta(1, 2), \quad \Delta(2, 1) > \Delta(2, 2) > \Delta(2, 3), \quad \Delta(3, 2) > \dots$$

*) Это распределение табулировано у Swed and Eisenhart, Tables for testing randomness of grouping in a sequence of alternatives, Ann. Math. Stat. 14 (1943), 66—87. Дополнительные сведения о критерии числа серий см. в работе: Wolfowitz, On the theory of runs with some applications to quality control, Ann. Math. Stat. 14 (1943), 280—288.

Таким образом, только соотношение между $\Delta(i, i+1)$ и $\Delta(i+1, i)$ зависит от специального выбора альтернативы, что приводит к желаемому результату.

(II) Несмещенность условного критерия при каждом m можно установить, записав его мощность в виде

$$\beta(p_0, p_1 | m) = (1 - \gamma) P\{R < C(m) | m\} + \gamma P\{R \leq C(m) | m\}$$

и замечая, в соответствии с (I), что критические области $R < C(m)$ и $R \leq C(m) + 1$ являются РНМ при отвечающих им условных уровнях значимости.

(III) Если H истинна, то условная вероятность при данном m любого набора из m нулей и n единиц равна $1/\binom{m+n}{m}$. n единиц можно разбить на r групп $\binom{n-1}{r-1}$ способами и m нулей на $r+1$ группу $\binom{m-1}{r}$ способами. Поэтому условная вероятность получить $r+1$ серию нулей и r серий единиц равна $\binom{m-1}{r} \binom{n-1}{r-1} / \binom{m+n}{m}$. Чтобы закончить доказательство, остается заметить, что полное число серий равно $2r+1$ в том и только том случае, когда имеется или $r+1$ серия нулей и r серий единиц или r серий нулей и $r+1$ серия единиц.]

21. *Критерий суммы рангов.* Пусть $Y_1, \dots, Y_N$ независимы и распределены по биномиальным законам $b(p_i, n_i)$, $i = 1, \dots, N$, где

$$p_i = 1/[1 + e^{-(\alpha + \beta x_i)}].$$

Такая модель часто встречается в биологических испытаниях. Тогда x_i обозначает дозу (или функцию от дозы, например логарифм) некоторого вещества, данную n_i подопытным животным, а Y_i — число тех из них, которые реагируют на это вещество определенным образом. Здесь x_i известны, а α и β — неизвестные параметры.

(I) Совместные распределения Y_i образуют экспоненциальное семейство, и РНМ несмещенные критерии существуют для четырех гипотез теоремы 3 как в случае α , так и в случае β .

(II) Предположим, в частности, что $x_i = \Delta_i$, где Δ известно, и что $n_i = 1$ при всех i . Обозначим n число успехов в N испытаниях и допустим, что успех наблюдался в $s_1, s_2, \dots, s_n$ испытаниях, $s_1 < s_2 < \dots < s_n$. Тогда РНМ несмещенный критерий для проверки $H: \beta = 0$ при альтернативе $\beta > 0$ применяется условно при фиксированном n и отклоняет гипотезу, если сумма рангов $\sum_{i=1}^n s_i$ слишком велика*).

(III) Пусть $Y_1, \dots, Y_M$ и $Z_1, \dots, Z_N$ — результаты двух независимых экспериментов описанного в начале задачи типа. Они могут соответствовать, например, двум различным испытываемым веществам. Если Y_i имеет распределение $b(p_i, m_i)$, а Z_j — распределение $b(\pi_j, n_j)$, где

$$p_i = \frac{1}{[1 + e^{-(\alpha + \beta u_i)}]}, \quad \pi_j = \frac{1}{[1 + e^{-(\gamma + \delta v_j)}]},$$

то РНМ несмещенные критерии существуют для всех четырех гипотез относительно $\gamma - \alpha$ и $\delta - \beta$.

*) Таблицы этого критерия, который формально эквивалентен критерию Вилкоксона (раздел 8 главы 6) см. у Fix and Hodges, Significance probabilities of the Wilcoxon test, Ann. Math. Stat. 26 (1955), 301—312.

9. Литературные ссылки

Барра (Barra J. R.)

(1974) Основные понятия математической статистики, пер. с франц. М., Мир, 1974, гл. 2. [Соотношение между понятиями подобия, полноты, достаточности и независимости.]

Бартлетт (Bartlett M.)

(1937) Properties of sufficiency and statistical tests, Proc. Roy. Soc. London, Ser. A, vol. 160, 268—282. [Указывается, что точные (т. е. подобные) критерии могут быть получены комбинированием условных критериев, определяемых различными значениями достаточных статистик. Даны примеры применений.]

Вальд (Wald A.)

(1947) Sequential Analysis, New York, John Wiley & Sons, Section 6.3. (Русский перевод: А. Вальд. Последовательный анализ. Физматгиз, Москва, 1960.) [Задача 18 (III).]

Гиршик, Мостеллер и Сэвидж (Girshick M. A., Mosteller F. and Savage L. J.)

(1946) Unbiased estimates for certain binomial sampling problems with applications, Ann. Math. Stat., vol. 17, 13—23. [Задача 11.]

Гхош (Ghosh M. N.)

(1948) On the problem of similar regions, Sankhyā, vol. 8, 329—338. [Теорема 1.]

Дэвид (David F. N.)

(1947) A power function for tests of randomness in a sequence, Biometrika, vol. 34, 335—339. [Рассматривается критерий числа серий в связи с моделью задачи 19.]

Ирвин (Irwin J. O.)

(1935) Tests of significance for differences between percentages based on small numbers, Metron, vol. 12, 83—94. [Предложен условный критерий для гипотезы независимости в 2×2 таблицах, который был предложен также Фишером (ср. Yates, Contingency tables involving small numbers and the χ^2 -test, J. Roy. Stat. Soc., Suppl., vol. 1 (1934), 217—235.)]

Кендалл и Стьюарт (Kendall M. G., Stuart A.)

(1973) Статистические выводы и связи, пер. с англ., М., Наука.

Крускал (Kruskal W. H.)

(1957) Historical notes on the Wilcoxon unpaired two-sample test, J. Am. Stat. Assoc., vol. 52, 356—360. [Дана история критерия суммы рангов задачи 21.]

Леман (Lehmann E. L.)

(1947) On families of admissible tests, Ann. Math. Stat., vol. 18, 97—104. [Задача 8.]

(1950) Some principles of the theory of testing hypotheses, Ann. Math. Stat., vol. 21, 1—26. [Лемма 1.]

(1952) Testing multiparameter hypotheses, Ann. Math. Stat., vol. 23, 541—552. [Задача 7.]

Леман и Шеффé (Lehmann E. L. and Scheffé H.)

(1950, 1955) Completeness, similar regions and unbiased estimation, Sankhyā, vol. 10, 305—340, vol. 15, 219—236. [Вводится понятие полноты. Теорема 3 и ее применение.]

Нанди (Nandi H. K.)

(1951) On type B_1 and type B regions, *Sankhyā*, vol. 11, 13—22. [Один из случаев теоремы 3 при некоторых предположениях регулярности.]

Нейман (Neyman J.)

(1935) Sur la vérification des hypothèses statistiques composées, *Bull. Soc. Math. France*, vol. 63, 246—266. [Теория локально несмещенных и локально наиболее мощных критериев для сложных гипотез.]

(1941) On a statistical problem arising in routine analyses and in sampling inspection of mass distributions, *Ann. Math. Stat.*, vol. 12, 46—76.

Нейман и Пирсон (Neyman J. and Pearson E. S.)

(1933) On the problem of the most efficient tests of statistical hypotheses, *Phil. Trans. Roy. Soc., Ser. A*, vol. 231, 289—337. [Введено понятие подобия и изложен метод для нахождения всей совокупности подобных областей.]

(1936, 1938) Contributions to the theory of testing statistical hypotheses, *Stat. Res. Mem.*, vol. I, 1—37; vol. II, 25—57. [Определяется несмещенность и устанавливаются как локальные, так и РНМ несмещенные критерии для некоторых классов простых гипотез.]

Путтер (Putter J.)

(1955) The treatment of ties in some nonparametric tests, *Ann. Math. Stat.*, vol. 26, 368—386. [Обсуждается способ обращения с ничьими в плане применения критерия знаков.]

Пшиборовский и Виленский (Przyborowski J. and Wilenski H.)

(1939) Homogeneity of results in testing samples from Poisson series, *Biometrika*, vol. 31, 313—323. [Получен РНМ подобный критерий для проверки равенства параметров двух пуассоновских распределений.]

Свердруп (Sverdrup E.)

(1953) Similarity, unbiasedness, minimaxibility and admissibility of statistical test procedures, *Skand. Aktuar. Tidskrift*, vol. 36, 64—86. [Теорема 1, результаты типа теоремы 3. Применения, включая 2×2 таблицу.]

Токер (Tocher K. D.)

(1950) Extension of Neyman — Pearson theory of tests to discontinuous variates, *Biometrika*, vol. 37, 130—144. [Доказано оптимальное свойство критерия Фишера — Ирвина, данного в разделе 6.]

Уолш (Walsh E.)

(1949) Some significance tests for the median which are valid under very general conditions, *Ann. Math. Stat.*, vol. 20, 64—81. [Содержит результат, связанный с задачей 12.]

Феллер (Feller W.)

(1936) Note on regions similar to the sample space, *Stat. Res. Mem.*, vol. II, 117—125. [Получен результат, который влечет полноту порядковых статистик.]

Фишер (Fisher R. A.)

(1934) *Statistical Methods for Research Workers*, Edinburgh, Oliver and Boyd, 5-е и последующие издания. Section 21.02. (Русский перевод: Р. Фишер, *Статистические методы для исследователей*, М., Госстатиздат, 1958.) [Предложен условный критерий для гипотезы независимости в 2×2 таблицах.]

Фрезер (Fraser D. A. S.)

(1953) Completeness of order statistics, *Canad. J. Math.*, vol. 6, 42—45. [Задача 12 и 13.]

Холдейн и Смит (Haldane J. B. S. and Smith C. A. B.)

(1948) A simple exact test for birth—order effect, *Ann. Eugenics*, vol. 14, 117—124. [Предложен критерий суммы рангов в той же ситуации, что в задаче 21.]

Хоул (Hoel P. G.)

(1945) Testing the homogeneity of Poisson frequencies, *Ann. Math. Stat.*, vol. 16, 362—368. [Первый пример раздела 5.]

(1948) On the uniqueness of similar regions, *Ann. Math. Stat.*, vol. 19, 66—71. [Теорема 1 при некоторых предположениях регулярности.]

Шеффé (Scheffé H.)

(1943) On a measure problem arising in the theory of non-parametric tests, *Ann. Math. Stat.*, vol. 14, 227—233. [Доказывается полнота порядковых статистик.]

ГЛАВА 5

НЕСМЕЩЕННОСТЬ: ПРИМЕНЕНИЯ К НОРМАЛЬНЫМ РАСПРЕДЕЛЕНИЯМ; ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

1. Статистики, не зависящие от достаточной статистики

Общий вид РНМ несмещенных критериев для гипотез $H_1: \theta \leq \theta_0$ и $H_4: \theta = \theta_0$ в экспоненциальном семействе

$$dP_{\theta, \phi}(x) = C(\theta, \phi) \exp[\theta U(x) + \sum \phi_i T_i(x)] d\mu(x) \quad (1)$$

был указан в теореме 3 предыдущей главы. Однако он оказывается неудобным для применений к нормальному и некоторым другим семействам непрерывных распределений, с которыми нам придется иметь дело в настоящей главе. В этих применениях упомянутым критериям можно придать более удобную форму. Они уже не будут выступать в форме условных критериев, построенных по U при данном t , а будут выражены в терминах одной статистики.

Эта редукция связана с существованием статистики $V = h(U, T)$, которая не зависит от T при $\theta = \theta_0$, и при каждом фиксированном t монотонна по U (для H_1) или линейна по U (для H_4). Критическая функция ϕ_1 для проверки H_1 тогда имеет вид

$$\phi(v) = \begin{cases} 1 & \text{при } v > C_0, \\ \gamma_0 & \text{при } v = C_0, \\ 0 & \text{при } v < C_0, \end{cases} \quad (2)$$

где C_0 и γ_0 уже не зависят от t и определяются из уравнений

$$E_{\theta_0} \phi_1(V) = \alpha. \quad (3)$$

Аналогично, критерий ϕ_4 для H_4 принимает форму

$$\phi(v) = \begin{cases} 1 & \text{при } v < C_1 \text{ или } v > C_2, \\ \gamma_i & \text{при } v = C_i, i = 1, 2, \\ 0 & \text{при } C_1 < v < C_2, \end{cases} \quad (4)$$

где C и γ определяются из уравнений

$$E_{\theta_0}[\varphi_4(V)] = \alpha \quad (5)$$

и

$$E_{\theta_0}[V\varphi_4(V)] = \alpha E_{\theta_0}(V). \quad (6)$$

Для соответствующей редукции при гипотезах $H_2: \theta \leq \theta_1$ или $\theta \geq \theta_2$ и $H_3: \theta_1 \leq \theta \leq \theta_2$ требуется монотонность V по U при каждом фиксированном t и независимость V от T при $\theta = \theta_1$ и $\theta = \theta_2$. Критерий φ_3 в этом случае дается формулой (4), а константы C и γ находятся из уравнения

$$E_{\theta_1}\varphi_3(V) = E_{\theta_2}\varphi_3(V) = \alpha. \quad (7)$$

Критерий для H_2 , как и раньше, имеет критическую функцию

$$\varphi_2(v; \alpha) = 1 - \varphi_3(v; 1 - \alpha).$$

Высказанные утверждения суммируются в следующей теореме.

Теорема 1. *Предположим, что распределение X задается формулой (1) и что $V = h(U, T)$ не зависит от T при $\theta = \theta_0$. Тогда φ_1 является РНМ несмещенным критерием для проверки H_1 , если функция h возрастает по u при каждом t ; φ_4 является РНМ несмещенным критерием для H_4 , если*

$$h(u, t) = a(t)u + b(t), \quad \text{где } a(t) > 0.$$

Критерии φ_2 и φ_3 суть несмещенные РНМ критерии для H_2 и H_3 , если V не зависит от T при $\theta = \theta_1$ и $\theta = \theta_2$, и если h возрастает по u для каждого t .

Доказательство. Критерий для H_1 , определенный формулами (12) и (13) главы 4, эквивалентен критерию (2) с константами, вычисляемыми из соотношения

$$P_{\theta_0}\{V > C_0(t) | t\} + \gamma_0(t) P_{\theta_0}\{V = C_0(t) | t\} = \alpha.$$

По предположению, V не зависит от T при $\theta = \theta_0$, следовательно, C_0 и γ_0 также не зависят от t . Этим завершается доказательство для H_1 . Для H_2 и H_3 оно вполне аналогично.

Критерий гипотезы H_4 , данный в разделе 4 главы 4, эквивалентен критерию (4) с константами C_i и γ_i , определяющимися из уравнений $E_{\theta_0}[\varphi_4(V, t) | t] = \alpha$ и

$$E_{\theta_0}\left[\varphi_4(V, t) \frac{V - b(t)}{a(t)} \middle| t\right] = \alpha E_{\theta_0}\left[\frac{V - b(t)}{a(t)} \middle| t\right].$$

Последнее сводится к $E_{\theta_0}[V\varphi_4(V, t) | t] = \alpha E_{\theta_0}[V | t]$. Так как при $\theta = \theta_0$ V не зависит от T , то тем же свойством обладают величины C_i и γ_i , что и требовалось доказать.

В специальных случаях применения теоремы 1 независимость V и T может быть установлена стандартными приемами:

преобразованием переменных, с помощью характеристических функций, или геометрическим методом. Другой подход, особенно полезный при отыскании надлежащей статистики V , указывается следующей теоремой.

Теорема 2*). Пусть $\mathcal{P} = \{P_\theta, \theta \in \omega\}$ — семейство возможных распределений X , T — достаточная статистика для $\mathcal{P}$ и семейство $\mathcal{P}^T$ распределений T ограничено полно. Тогда статистика V не зависит от T при всех θ в том и только в том случае, когда распределение V не зависит от θ .

Доказательство. Распределение V не зависит от θ тогда и только тогда, когда для любой критической функции ϕ математическое ожидание $E_\theta \phi(V)$ не зависит от θ . Необходимое и достаточное условие для справедливости последнего состоит в том, что условное математическое ожидание $E[\phi(V) | t]$ постоянно $\mathcal{P}^T$ -почти всюду (см. теорему 2 главы 4), что эквивалентно независимости V и T .

Следствие 1. Обозначим $\mathcal{P}$ экспоненциальное семейство, получающееся из (1) при фиксированном θ . Тогда статистика V не зависит от T при всех θ в том и только том случае, когда распределение V не зависит от θ .

Доказательство. Из теоремы 1 главы 4 следует, что $\mathcal{P}^T$ полно, а потому и ограничено полно. Поэтому можно применить предыдущую теорему.

Пример 1. Пусть $X_1, \dots, X_n$ независимы и распределены нормально со средним ξ и дисперсией σ^2 . Предположим сначала, что σ^2 фиксировано и равно σ_0^2 . Тогда условия следствия 1 выполняются с $T = \bar{X} = \Sigma X_i/n$ и θ , пропорциональным ξ . Пусть f — любая функция, для которой

$$f(x_1 + c, \dots, x_n + c) = f(x_1, \dots, x_n) \quad \text{при всех действительных } c.$$

Если

$$V = f(X_1, \dots, X_n),$$

то тогда $V = f(X_1 - \xi, \dots, X_n - \xi)$. Так как случайные величины $X_1 - \xi$ имеют распределение $N(0, \sigma_0^2)$, не содержащее параметра ξ , то распределение V не зависит от ξ . Из следствия 1 вытекает, что любая такая статистика, и в частности $V = \Sigma (X_i - \bar{X})^2$, не зависит от $\bar{X}$. Это верно при всех σ .

Допустим теперь, что ξ фиксировано и равно ξ_0 . Тогда следствие 1 применимо с $T = \Sigma (X_i - \xi_0)^2$ и $\theta = -1/2\sigma^2$. Пусть f — любая функция, для которой

$$f(cx_1, \dots, cx_n) = f(x_1, \dots, x_n) \quad \text{при всех } c > 0,$$

и пусть

$$V = f(X_1 - \xi_0, \dots, X_n - \xi_0).$$

Тогда V не изменится, если каждую из величин $X_i - \xi_0$ заменить на $(X_i - \xi_0)/\sigma$. Эти величины имеют нормальное распределение со средним значе-

*) В части, касающейся необходимости (см. статью Басу в Sankhyā. 15, 1955, 377—380), теорема 2 сформулирована неточно. Точная формулировка необходимого условия независимости статистики от достаточной статистики имеется в работе Басу в Sankhyā 20, 1958, 223—226. — Прим. ред.

нием нуль и единичной дисперсией. Поэтому распределение V не зависит от σ . Мы видим, что все подобные статистики, и в частности

$$\frac{(\bar{X} - \xi_0)}{\sqrt{\sum (X_i - \bar{X})^2}} \quad \text{и} \quad \frac{(\bar{X} - \xi_0)}{\sqrt{\sum (X_i - \xi_0)^2}},$$

не зависят от $\sum (X_i - \xi_0)^2$. Однако это имеет место не при всех ξ , но лишь при $\xi = \xi_0$.

Пример 2. Пусть U_1/σ_1^2 и U_2/σ_2^2 независимы и имеют распределения χ^2 с f_1 и f_2 степенями свободы соответственно. Предположим, что $\sigma_2^2/\sigma_1^2 = a$. Тогда совместная плотность величин U имеет вид

$$C u_1^{\frac{1}{2} f_1 - 1} u_2^{\frac{1}{2} f_2 - 1} \exp \left[-\frac{1}{2\sigma_2^2} (a u_1 + u_2) \right],$$

так что следствие 1 применимо с $T = aU_1 + U_2$ и $\phi = -1/2\sigma_2^2$. Так как распределение отношения

$$V = \frac{U_2}{U_1} = a \frac{U_2/\sigma_2^2}{U_1/\sigma_1^2}$$

не зависит от σ_2 , то V не зависит от $aU_1 + U_2$. Для специального случая $\sigma_2 = \sigma_1$ это доказывает независимость U_2/U_1 и $U_1 + U_2$.

Пример 3. Пусть $(X_1, \dots, X_n)$ и $(Y_1, \dots, Y_n)$ — выборки из нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ соответственно. Тогда статистика $T = (\bar{X}, \Sigma X_i^2, \bar{Y}, \Sigma Y_i^2)$ достаточна для $(\xi, \sigma^2, \eta, \tau^2)$ и семейство распределений T полно. Так как

$$V = \frac{\Sigma (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma (X_i - \bar{X})^2 \Sigma (Y_i - \bar{Y})^2}}$$

не меняется при замене X_i и Y_i на $(X_i - \xi)/\sigma$ и $(Y_i - \eta)/\tau$, то распределение V не зависит ни от одного из четырех параметров, и теорема 2 показывает, что V не зависит от T .

2. Проверка гипотез о параметрах нормального распределения

Четыре гипотезы $\sigma \leq \sigma_0$, $\sigma \geq \sigma_0$, $\xi \leq \xi_0$, $\xi \geq \xi_0$ о среднем и дисперсии нормального распределения рассматривались в разделе 9 главы 3. Там отмечалось, что при обычных уровнях значимости РНМ критерий существует только для первой гипотезы. Мы покажем теперь, что стандартные (основанные на отношении правдоподобия) критерии являются РНМ несмещенными для всех четырех гипотез, а также и для некоторых двусторонних гипотез.

При меняющихся ξ и σ плотности

$$(2\pi\sigma^2)^{-n/2} \exp \left(-\frac{n\xi^2}{2\sigma^2} \right) \exp \left(-\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\xi}{\sigma^2} \sum x_i \right) \quad (8)$$

выборки $X_1, \dots, X_n$ из $N(\xi, \sigma^2)$ образуют двухпараметрическое семейство, которое совпадает с (1) при

$$\theta = -\frac{1}{2\sigma^2}, \quad \vartheta = \frac{n\xi}{\sigma^2}, \quad U(x) = \sum x_i^2, \quad T(x) = \bar{x} = \frac{\sum x_i}{n}.$$

Поэтому по теореме 3 главы 4 существует РНМ несмещенный критерий для гипотезы $\theta \geq \theta_0$, которая при $\theta = -1/2\sigma_0^2$ эквивалентна $H: \sigma \geq \sigma_0$. Критическая область для этого случая может быть получена из формулы (12) главы 4 при замене неравенств на обратные, так как теперь гипотеза имеет вид $\theta \geq \theta_0$. В рассматриваемом случае имеем

$$\sum x_i^2 \leq C_0(\bar{x}),$$

где

$$P_{\sigma} \{ \sum X_i^2 \leq C_0(\bar{x}) \mid \bar{x} \} = \alpha.$$

Если это записать в форме

$$\sum x_i^2 - n\bar{x}^2 \leq C'_0(\bar{x}),$$

то из независимости $\sum X_i^2 - n\bar{X}^2 = \sum (X_i - \bar{X})^2$ и $\bar{X}$ (см. пример 1) будет следовать, что $C'_0(\bar{x})$ не зависит от $\bar{x}$. Поэтому критерий отвергает гипотезу, когда $\sum (x_i - \bar{x})^2 \leq C'_0$, т. е. когда

$$\frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \leq C_0, \quad (9)$$

где C_0 определяется равенством $P_{\sigma_0} \{ \sum (X_i - \bar{X})^2 / \sigma_0^2 \leq C_0 \} = \alpha$. Поскольку $\sum (X_i - \bar{X})^2 / \sigma_0^2$ имеет χ^2 -распределение с $n-1$ степенью свободы, то уравнение для C_0 приводится к виду

$$\int_0^{C_0} \chi_{n-1}^2(y) dy = \alpha, \quad (10)$$

где χ_{n-1}^2 обозначает плотность χ^2 -распределения с $n-1$ степенью свободы.

К этому результату можно прийти и с помощью теоремы 1. В качестве статистики $V = h(U, T)$, фигурирующей в условиях теоремы, т. е. не зависящей от $\bar{X}$ при $\sigma = \sigma_0$ и при всех ξ , можно взять

$$V = \sum (X_i - \bar{X})^2 = U - nT^2.$$

Эта статистика, действительно, не зависит от $\bar{X}$ при всех ξ и σ^2 . Так как $h(u, t)$ является возрастающей функцией u при каждом t , то критическая область для РНМ несмещенного критерия должна иметь вид $V \leq C'_0$.

Этот вывод показывает также, что РНМ несмещенная критическая область для $H: \sigma \leq \sigma_1$ или $\sigma \geq \sigma_2$ имеет вид

$$C_1 < \sum (x_i - \bar{x})^2 < C_2, \quad (11)$$

где константы C определяются из уравнений

$$\int_{c_1/\sigma_1^2}^{c_2/\sigma_1^2} \chi_{n-1}^2(y) dy = \int_{c_1/\sigma_2^2}^{c_2/\sigma_2^2} \chi_{n-1}^2(y) dy = \alpha. \quad (12)$$

Так как $h(u, t)$ линейна по u , мы видим, что РНМ несмещенному критерию гипотезы $H: \sigma = \sigma_0$ соответствует область принятия

$$C'_1 < \frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} < C'_2 \quad (13)$$

с константами, вычисляемыми из соотношений

$$\int_{c'_1}^{c'_2} \chi_{n-1}^2(y) dy = \frac{1}{n-1} \int_{c'_1}^{c'_2} y \chi_{n-1}^2(y) dy = 1 - \alpha. \quad (14)$$

Это — критерий, найденный в примере 2 главы 4 при замене $\sum x_i^2$ на $\sum (x_i - \bar{x})^2$ и числа степеней свободы n на $n - 1$, что можно было предвидеть. Теорема 1 показывает, что РНМ несмещенный критерий для этой (и аналогичных) гипотез зависит только от V . Так как распределение V не зависит от ξ и при меняющемся σ порождает экспоненциальное семейство, то задачи сводятся к аналогичным задачам для однопараметрического экспоненциального семейства, т. е. к уже решенным.

Мощность вышеприведенных критериев может быть точно выражена в терминах χ^2 -распределения. Например, в случае одностороннего критерия (9) она равна

$$\beta(\sigma) = P_\sigma \left\{ \frac{\sum (X_i - \bar{X})^2}{\sigma^2} \leq \frac{C_0 \sigma_0^2}{\sigma^2} \right\} = \int_0^{C_0 \sigma_0^2 / \sigma^2} \chi_{n-1}^2(y) dy.$$

Этот же самый метод может быть применен к проверке гипотез $\xi \leq \xi_0$ (при альтернативе $\xi > \xi_0$) и $\xi = \xi_0$ (при альтернативе $\xi \neq \xi_0$). Переходя, если необходимо, к величинам $X_i - \xi_0$, мы можем, не ограничивая общности, предположить $\xi_0 = 0$. Приведем (8) к виду (1), полагая

$$\theta = \frac{\eta \xi}{\sigma^2}, \quad \vartheta = -\frac{1}{2\sigma^2}, \quad U(x) = \bar{x}, \quad T(x) = \sum x_i^2.$$

Теорема 3 главы 4 показывает, что для гипотез $\theta \leq 0$ и $\theta = 0$, эквивалентных $\xi \leq 0$ и $\xi = 0$, существуют РНМ несмещенные критерии. Поскольку при $\xi = 0$

$$V = \frac{\bar{X}}{\sqrt{\sum (X_i - \bar{X})^2}} = \frac{U}{\sqrt{T - nU^2}}$$

не зависит от $T = \sum X_i^2$ (см. пример 1), то по теореме 1 РНМ несмещенная критическая область для $H: \xi \leq 0$ имеет вид $V \geq C'_0$ или, что то же самое, вид

$$t(x) \geq C_0, \quad (15)$$

где

$$t(x) = \frac{\sqrt{n} \bar{x}}{\sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}}. \quad (16)$$

Перейдем к гипотезе $H': \xi = 0$. Обозначим $W = \bar{X}/\sqrt{\sum X_i^2}$. При $\xi = 0$ эта величина не зависит от $\sum X_i^2$ и к тому же линейна относительно $U = \bar{X}$. Распределение W при $\xi = 0$ симметрично относительно точки 0. Следовательно, условия (4), (5), (6) с заменой V на W выполняются для критической области вида $|w| \geq C'$ с $P_{\xi=0}\{|W| \geq C'\} = \alpha$. Так как

$$t(x) = \frac{\sqrt{(n-1)n} W(x)}{\sqrt{1 - nW^2(x)}},$$

то абсолютная величина $t(x)$ является возрастающей функцией от $|W(x)|$. Поэтому критическая область принимает вид

$$|t(x)| \geq C. \quad (17)$$

Из (16) мы видим, что $t(X)$ есть отношение двух независимых случайных величин $\sqrt{n} \bar{X}/\sigma$ и $\sqrt{\sum (X_i - \bar{X})^2/(n-1)\sigma^2}$. Знаменатель распределен как квадратный корень из χ^2 -величины с $n-1$ степенями свободы, деленной на $n-1$. Распределение числителя при $\xi = 0$ нормально $N(0, 1)$. Распределение подобного отношения называется распределением *Стьюдента* с $n-1$ -й степенью свободы и имеет плотность

$$t_{n-1}(y) = \frac{1}{\sqrt{\pi(n-1)}} \frac{\Gamma\left(\frac{1}{2}n\right)}{\Gamma\left[\frac{1}{2}(n-1)\right]} \frac{1}{\left(1 + \frac{y^2}{n-1}\right)^{\frac{1}{2}n}}. \quad (18)$$

Распределение симметрично относительно 0, и константы C_0 и C для одно- и двустороннего критериев находятся из уравнений

$$\begin{aligned} \int_{C_0}^{\infty} t_{n-1}(y) dy &= \alpha, \\ \int_C^{\infty} t_{n-1}(y) dy &= \frac{\alpha}{2}. \end{aligned} \quad (19)$$

При $\xi \neq 0$ $t(X)$ имеет так называемое *нецентральное t -распределение* (оно выведено в задаче 3). Некоторые свойства функции мощности одно- и двустороннего t -критерия даны в задачах 1, 2 и 4. Отметим, что распределение $t(X)$, а вместе с тем и мощность указанных выше критериев, зависит только от «параметра нецентральности» $\delta = \sqrt{n} \xi/\sigma$. Это видно, например, из формулы для плотности, приводимой в задаче 3, но может быть установлено и следующим непосредственным путем. Допустим, что $\xi'/\sigma' = \xi/\sigma \neq 0$ и обозначим c общее значение отношений ξ'/ξ и σ'/σ , $c \neq 0$. Если $X'_i = cX_i$ и X_i распределены по закону $N(\xi, \sigma^2)$, то X'_i имеют распределение $N(\xi', \sigma'^2)$. Итак, $t(X) = t(X')$ и, следовательно, $t(X')$ имеет такое же распределение, как и $t(X)$, что и требовалось доказать.

Пусть ξ_1 — некоторая альтернатива к $\xi = 0$. Тогда мощность $\beta(\xi_1, \sigma) = f(\delta)$ зависит от σ . При $\sigma \rightarrow \infty$ и $\delta \rightarrow 0$

$$\beta(\xi_1, \sigma) \rightarrow f(0) = \beta(0, \sigma) = \alpha,$$

так как по теореме 6 главы 2 функция f непрерывна. Следовательно, каков бы ни был объем выборки, вероятность обнаружить ложность гипотезы при $\xi \geq \xi_1 > 0$ не может быть сделана $\geq \beta > \alpha$ при всех σ . Это не удивительно, так как распределения $N(0, \sigma^2)$ и $N(\xi_1, \sigma^2)$ становятся практически неразличимыми, если σ слишком велико. Чтобы получить процедуру с гарантированной мощностью при $\xi \geq \xi_1$, объем выборки следует сделать зависящим от σ . Этого можно достигнуть применением последовательной процедуры с правилом остановки, опирающимся на оценки σ , но невозможно при выборках фиксированного объема (см. задачи 15 и 17).

Критерии для более общих гипотез $\xi \leq \xi_0$ и $\xi = \xi_0$ получаются из найденных выше заменой переменных на $X_i - \xi_0$. Критические области для этих гипотез, как и раньше, даются формулами (15), (17) и (19), но на этот раз с

$$t(x) = \frac{\sqrt{n}(\bar{x} - \xi_0)}{\sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}}.$$

Из записи распределений (8) в форме экспоненциального семейства с $\theta = n\xi/\sigma^2$ видно, что существует РНМ несмещенный критерий для гипотезы $a \leq \xi/\sigma^2 \leq b$, но метод неприменим к более интересной гипотезе $a \leq \xi \leq b^*$; он неприменим и к аналогичной гипотезе, когда среднее выражено в единицах стандартного отклонения: $a \leq \xi/\sigma \leq b$. Последняя гипотеза будет обсуждаться в главе 6.

Данные выше РНМ несмещенные критерии для среднего и дисперсии в определенном отношении заметно различаются. Если случайные величины $X_1, \dots, X_n$ образуют выборку из любого распределения с конечной дисперсией и нулевым средним и если объем выборки достаточно велик, то статистика (16) распределена приближенно по закону $N(0, 1)$. Это следует из центральной предельной теоремы (в соответствии с которой $\sqrt{n} \bar{X}/\sigma$ имеет предельное распределение $N(0, 1)$), того факта, что $\sum (X_i - \bar{X})^2/(n-1)\sigma^2$ стремится по вероятности к единице, и из одной теоремы о сходимости, принадлежащей Крамеру **). Поэтому размер t -критерия будет, по крайней мере при больших выборках, приближенно равен заданному уровню значимости, даже если основное распределение отклоняется от нормального ***).

С другой стороны, предельное распределение для $\sum (X_i - \bar{X})^2/\sqrt{n}\sigma^2$ уже будет зависящим от исходного распределения величин X , точнее будет связано с величиной четвертого момента $E(X_i^4)$. Так как $\sum (X_i - \bar{X})^2$ не зависит от среднего значения X_i , то, не ограничивая общности, мы можем принять $E(X_i) = 0$. Тогда $\sqrt{n} \bar{X}$ имеет предельное распределение $N(0, \sigma^2)$, а $n\bar{X}^2/\sqrt{n}$ стремится по вероятности к нулю. Поэтому величина $[\sum (X_i - \bar{X})^2 - n\sigma^2]/\sqrt{n}$ имеет то же самое предельное распределение, что и $[\sum X_i^2 - n\sigma^2]/\sqrt{n}$, а именно, по цент-

*) Эта проблема обсуждается в разделе 3 работы: Hodges and Lehmann, Testing the approximate validity of statistical hypotheses, J. Roy. Stat. Soc., ser. B. 16 (1954), 261—268.

**) Формулировку и доказательство этой теоремы см. в книге: Крамер, Математические методы статистики, Мир, 1975, 281.

***) Более детальное исследование поведения t -критерия для распределений, отличающихся от нормального, было предпринято Gaуen, The distribution of Student's t in random samples of any size from non-normal universes, Biometrika 36 (1949), 353—369 и Geary, The distribution of Student's ratio for non-normal samples, Suppl. J. Roy. Stat. Soc. 3 (1936), 178—184. В частности, показано, что приближение к предельному поведению в двустороннем случае более быстрое, чем в одностороннем, и что в действительности односторонний критерий при малых выборках очень чувствителен к отклонениям от нормальности. См. также Tukey, Some elementary problems of importance to small sample practice, Human Biology 20 (1948), 205—214 и обзорную статью Wallace, Asymptotic approximations so distributions, Ann. Math. Stat. 29 (1958), 635—654.

ральной предельной теореме, — нормальное распределение $N(0, \tau^2)$, где τ^2 — дисперсия величин X_i^2 . Отсюда вытекает, что размер критериев для дисперсии (9) и (11) может быть далек от заданных уровней значимости, даже и при больших объемах выборки, если только исходное распределение отклоняется от нормального.

3. Сравнение средних и дисперсий двух нормальных распределений

Проблема сравнения параметров двух нормальных распределений возникает при сравнении двух методов обработки, сравнении двух продуктов и т. д., при условиях, подобных обсуждавшимся в начале раздела 5 главы 4. Мы рассмотрим сначала задачу сравнения двух дисперсий. Эта задача встречается при сравнении степени изменчивости анализов, сделанных в двух лабораториях или двумя различными методами. Специально мы остановимся на гипотезах $H: \tau^2/\sigma^2 \leq \Delta_0$ и $H': \tau^2/\sigma^2 = \Delta_0$.

Пусть $X = (X_1, \dots, X_m)$ и $Y = (Y_1, \dots, Y_n)$ — выборки из нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ с совместной плотностью

$$C(\xi, \eta, \sigma, \tau) \exp \left(-\frac{1}{2\sigma^2} \sum x_i^2 - \frac{1}{2\tau^2} \sum y_j^2 + \frac{m\xi}{\sigma^2} \bar{x} + \frac{n\eta}{\tau^2} \bar{y} \right).$$

Это есть экспоненциальное семейство с четырьмя параметрами

$$\theta = -\frac{1}{2\tau^2}, \quad \theta_1 = -\frac{1}{2\sigma^2}, \quad \theta_2 = \frac{n\eta}{\tau^2}, \quad \theta_3 = \frac{m\xi}{\sigma^2}$$

и достаточными статистиками

$$U = \sum Y_j^2, \quad T_1 = \sum X_i^2, \quad T_2 = \bar{Y}, \quad T_3 = \bar{X}.$$

Эквивалентное выражение можно получить (см. лемму 2 главы 4) в терминах параметров

$$\theta^* = \left(-\frac{1}{2\tau^2} \right) + \left(\frac{1}{2\Delta_0\sigma^2} \right), \quad \theta_i^* = \theta_i \quad (i = 1, 2, 3)$$

и статистик

$$U^* = \sum Y_j^2, \quad T_1^* = \sum X_i^2 + \left(\frac{1}{\Delta_0} \right) \sum Y_j^2, \quad T_2^* = \bar{Y}, \quad T_3^* = \bar{X}.$$

Гипотезам $\theta^* \leq 0$ и $\theta^* = 0$, эквивалентным H и H' соответственно, по теореме 3 главы 4 отвечают РНМ несмещенные критерии.

При $\tau^2 = \Delta_0\sigma^2$ распределение статистики

$$V = \frac{\sum (Y_j - \bar{Y})^2 / \Delta_0}{\sum (X_i - \bar{X})^2} = \frac{\sum (Y_j - \bar{Y})^2 / \tau^2}{\sum (X_i - \bar{X})^2 / \sigma^2}$$

не зависит от σ , ξ , η и, как вытекает из следствия 1, величина V не зависит от (T_1^*, T_2^*, T_3^*) . Поэтому РНМ несмещенный критерий для H дается формулами (2) и (3), так что критическая область имеет вид

$$\frac{\sum (Y_j - \bar{Y})^2 / \Delta_0 (n-1)}{\sum (X_i - \bar{X})^2 / (m-1)} \geq C_0. \quad (20)$$

При $\tau^2 = \Delta_0 \sigma^2$ статистика в левой части (20) является отношением двух независимых χ^2 -величин $\sum (Y_j - \bar{Y})^2 / \tau^2$ и $\sum (X_i - \bar{X})^2 / \sigma^2$, причем каждая разделена на соответствующее число степеней свободы. Распределение такого отношения есть F -распределение с $n-1$ и $m-1$ степенями свободы и с плотностью

$$F_{n-1, m-1}(y) = \frac{\Gamma\left[\frac{1}{2}(m+n-2)\right]}{\Gamma\left[\frac{1}{2}(m-1)\right] \Gamma\left[\frac{1}{2}(n-1)\right]} \left(\frac{n-1}{m-1}\right)^{\frac{1}{2}(n-1)} \times \\ \times \frac{y^{\frac{1}{2}(n-1)-1}}{\left(1 + \frac{n-1}{m-1} y\right)^{\frac{1}{2}(m+n-2)}}. \quad (21)$$

Константа C_0 в (20) находится из уравнения

$$\int_{C_0}^{\infty} F_{n-1, m-1}(y) dy = \alpha. \quad (22)$$

Чтобы применить теорему 1 к H' , положим

$$W = \frac{\sum (Y_j - \bar{Y})^2 / \Delta_0}{\sum (X_i - \bar{X})^2 + (1/\Delta_0) \sum (Y_j - \bar{Y})^2}$$

При $\tau^2 = \Delta_0 \sigma^2$ эта статистика также не зависит от $T^* = (T_1^*, T_2^*, T_3^*)$. Кроме того, она линейна по U^* . Следовательно, РНМ несмещенная область принятия H' имеет вид

$$C_1 \leq W \leq C_2, \quad (23)$$

где константы определяются из уравнений (5) и (6), в которых V заменено на W . Разделив числитель и знаменатель W на σ^2 , мы видим, что при $\tau^2 = \Delta_0 \sigma^2$ статистика W может быть представлена в виде отношения $W_1 / (W_1 + W_2)$, где W_1 и W_2 — независимые χ^2 -величины с $n-1$ и $m-1$ степенями свободы, соответственно. Эквивалентным образом, $W = Y / (1 + Y)$, где $Y = W_1 / W_2$ и $(m-1)Y / (n-1)$ имеет распределение $F_{n-1, m-1}$.

Распределение W есть *бета-распределение* *) с плотностью

$$B_{\frac{1}{2}(n-1), \frac{1}{2}(m-1)}(w) = \\ = \frac{\Gamma\left[\frac{1}{2}(m+n-2)\right]}{\Gamma\left[\frac{1}{2}(m-1)\right]\Gamma\left[\frac{1}{2}(n-1)\right]} w^{\frac{1}{2}(n-3)} (1-w)^{\frac{1}{2}(m-3)}, \quad 0 < w < 1. \quad (24)$$

Условия (5) и (6) с помощью соотношений

$$E(W) = \frac{n-1}{m+n-2}$$

и

$$w B_{\frac{1}{2}(n-1), \frac{1}{2}(m-1)}(w) = \frac{n-1}{m+n-2} B_{\frac{1}{2}(n+1), \frac{1}{2}(m-1)}(w)$$

могут быть превращены в

$$\int_{C_1}^{C_2} B_{\frac{1}{2}(n-1), \frac{1}{2}(m-1)}(w) dw = \int_{C_1}^{C_2} B_{\frac{1}{2}(n+1), \frac{1}{2}(m-1)}(w) dw = 1 - \alpha. \quad (25)$$

Определение статистики V показывает, что ее распределение зависит только от отношения τ^2/σ^2 , а следовательно, и распределение W зависит только от отношения τ^2/σ^2 . Мощность критериев (20) и (23) будет функцией только от $\Delta = \tau^2/\sigma^2$. Точное выражение для мощности можно получить в терминах F -распределения. Например, в первом случае

$$\beta(\Delta) = P \left\{ \frac{\sum (Y_i - \bar{Y})^2 / \tau^2 (n-1)}{\sum (X_i - \bar{X})^2 / \sigma^2 (m-1)} \geq \frac{C_0 \Delta_0}{\Delta} \right\} = \int_{C_0 \Delta_0 / \Delta}^{\infty} F_{n-1, m-1}(y) dy.$$

Гипотеза равенства средних ξ , η в двух нормальных распределениях с неизвестными дисперсиями σ^2 и τ^2 (так называемая *проблема Берэнса — Фишера*), не поддается указанным здесь методам (см. пример 5 главы 4; о возможном подходе к задаче см. раздел 6 главы 6). Мы рассмотрим поэтому лишь простейший случай, в котором дисперсии предполагаются равными. Тогда совместная плотность X_i и Y_i равна

$$C(\xi, \eta, \sigma) \exp \left[-\frac{1}{2\sigma^2} \left(\sum x_i^2 + \sum y_i^2 \right) + \frac{\xi}{\sigma^2} \sum x_i + \frac{\eta}{\sigma^2} \sum y_i \right]. \quad (26)$$

*) Отношение $W = Y/(1+Y)$ показывает, что распределение F и бета-распределение эквивалентны. Преимущество последнего состоит в наличии обширных таблиц функции распределения, *Tables of the Incomplete Beta Function*, Cambridge Univ. Press., 1932, составленных К. Пирсоном. (См. русский перевод: Пирсон К., Таблицы неполной бета-функции, М., 1974, а также Болшев Л. Н., Смирнов Н. В., Таблицы математической статистики, М., 1965. — Прим. ред.)

Это есть экспоненциальное семейство с параметрами

$$\theta = \frac{\eta}{\sigma^2}, \quad \theta_1 = \frac{\xi}{\sigma^2}, \quad \theta_2 = -\frac{1}{2\sigma^2}$$

и достаточными статистиками

$$U = \sum Y_j, \quad T_1 = \sum X_i, \quad T_2 = \sum X_i^2 + \sum Y_j^2.$$

Для целей проверки гипотез

$$H: \eta - \xi \leq 0 \quad \text{и} \quad H': \eta - \xi = 0$$

удобнее представить плотности в экспоненциальной форме с параметрами

$$\theta^* = \frac{\eta - \xi}{\left(\frac{1}{m} + \frac{1}{n}\right)\sigma^2}, \quad \theta_1^* = \frac{m\xi + n\eta}{(m+n)\sigma^2}, \quad \theta_2^* = \theta_2$$

и достаточными статистиками

$$U^* = \bar{Y} - \bar{X}, \quad T_1^* = m\bar{X} + n\bar{Y}, \quad T_2^* = \sum X_i^2 + \sum Y_j^2.$$

Возможность такого перехода видна из тождества

$$m\xi\bar{x} + n\eta\bar{y} = \frac{(\bar{y} - \bar{x})(\eta - \xi)}{\frac{1}{m} + \frac{1}{n}} + \frac{(m\bar{x} + n\bar{y})(m\xi + n\eta)}{m+n}.$$

Как вытекает из теоремы 3 главы 4, существуют РНМ несмещенные критерии для проверки гипотез $\theta^* \leq 0$ и $\theta^* = 0$, а потому — и для гипотез H и H' .

При $\eta = \xi$ распределение статистики

$$V = \frac{\bar{Y} - \bar{X}}{\sqrt{\sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2}} = \frac{U^*}{\sqrt{T_2^* - \frac{1}{m+n} T_1^{*2} - \frac{mn}{m+n} U^{*2}}}$$

не зависит от ξ и σ , что можно видеть, заменяя X_i на $(X_i - \xi)/\sigma$ и Y_j на $(Y_j - \xi)/\sigma$. Поэтому V не зависит от (T_1^*, T_2^*) . Критическая область для РНМ несмещенного критерия H имеет вид $V \geq \geq C'_0$ или

$$t(X, Y) \geq C_0, \quad (27)$$

где

$$t(X, Y) = \frac{(\bar{Y} - \bar{X})/\sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{[\sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2]/(m+n-2)}}. \quad (28)$$

Статистика $t(X, Y)$ является отношением двух независимых случайных величин

$$\frac{\bar{Y} - \bar{X}}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right) \sigma^2}} \quad \text{и} \quad \sqrt{\frac{\sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2}{(m+n-2) \sigma^2}}.$$

Числитель этого отношения распределен нормально со средним $(\eta - \xi) / \sqrt{m^{-1} + n^{-1}} \sigma$ и единичной дисперсией; знаменатель распределен как квадратный корень из χ^2 -величины $(m+n-2)$ степенями свободы, поделенной на $(m+n-2)$. Поэтому $t(X, Y)$ имеет нецентральное t -распределение с $(m+n-2)$ степенями свободы и параметром нецентральности

$$\delta = \frac{\eta - \xi}{\sqrt{\frac{1}{m} + \frac{1}{n}} \sigma}.$$

В частности, при $\eta - \xi = 0$ распределение $t(X, Y)$ превращается в t -распределение Стьюдента и константа C_0 определяется из уравнения

$$\int_C^\infty t_{m+n-2}(y) dy = \alpha. \quad (29)$$

Как и прежде, при проверке H' предположения теоремы 1 не выполняются для самой статистики V , а имеют место для статистики

$$W = \frac{\bar{Y} - \bar{X}}{\sqrt{\sum X_i^2 + \sum Y_j^2 - \frac{(\sum X_i + \sum Y_j)^2}{m+n}}},$$

связанной с V соотношением

$$V = \frac{W}{\sqrt{1 - \frac{mn}{m+n} W^2}}.$$

При $\eta = \xi$ W , будучи функцией от V , также не зависит от (T_1^*, T_2^*) ; к тому же, она является линейной функцией от U^* с коэффициентами, зависящими только от T^* . Заметим, что распределение W симметрично относительно 0 при $\eta = \xi$. Отсюда вытекает, так же как и при выводе формулы (17) для одной выборки, что РНМ несмещенный критерий H' отвергает ее при слишком больших $|W|$, или, иначе, при

$$|t(X, Y)| > C. \quad (30)$$

Константа C определяется из уравнения

$$\int_C^{\infty} t_{m+n-2}(y) dy = \frac{\alpha}{2}. \quad (31)$$

Мощность критериев (27) и (30) зависит только от $(\eta - \xi)/\sigma$ и может быть выражена в терминах нецентрального t -распределения. Ее свойства аналогичны свойствам мощности t -критерия для одной выборки (задачи 1, 2 и 4).

Так же как и в случае одной выборки, критерии, основанные на t -статистике (28), нечувствительны к отклонениям от нормальности, но это утверждение неверно для критериев, основанных на F -отношении (20)*. Доказательство использует те же рассуждения, что и в случае одной выборки. К устойчивости t -критерия мы вернемся еще раз в разделе 8, где будет обсуждаться видоизмеренный критерий, размер которого не зависит от исходного распределения.

4. Доверительные интервалы и семейства критериев

Доверительные границы для параметра θ , соответствующие доверительному уровню $1 - \alpha$, были определены в разделе 5 главы 3. При этом рассматривался случай, когда распределение X зависит только от θ . При наличии «мешающих параметров» ϕ определение нижней доверительной границы $\underline{\theta}$ имеет вид

$$P_{\theta, \phi} \{ \underline{\theta}(X) \leq \theta \} \geq 1 - \alpha \quad \text{при всех } \theta, \phi. \quad (32)$$

Аналогично этому, доверительные интервалы для θ с доверительным уровнем $1 - \alpha$ определяются как множество случайных интервалов с концами $\underline{\theta}(X)$, $\bar{\theta}(X)$ такими, что

$$P_{\theta, \phi} \{ \underline{\theta}(X) \leq \theta \leq \bar{\theta}(X) \} = 1 - \alpha \quad \text{при всех } \theta, \phi. \quad (33)$$

Точная нижняя грань по (θ, ϕ) левых частей (32) и (33) называется *коэффициентом доверия* этих утверждений.

Как уже отмечалось в главе 3, доверительные утверждения допускают двойственную интерпретацию. Прежде всего они дают границы для неизвестного параметра θ и тем самым решение проблемы оценки θ . Хотя утверждение $\underline{\theta} \leq \theta \leq \bar{\theta}$ не является столь точным, как точечная оценка, оно имеет то преимущество,

*) Критерии для двух и большего числа дисперсий, лишенные этого недостатка, обсуждаются в работах: Box, Non-normality and tests of variances, *Biometrika* 40 (1953), 318—335 и Box and Anderson, Permutation theory in the derivation of robust criteria and the study of departures from assumptions, *J. Roy. Stat. Soc., ser. B.* 17 (1955), 1—34.

что вероятность его правильности гарантированным образом не может быть сделана меньше $1 - \alpha$. Аналогично нижнюю доверительную границу можно представлять себе как оценку $\underline{\theta}$, которая переоценивает параметр с вероятностью $\leq \alpha$. В частности, при $\alpha = 1/2$ граница $\underline{\theta}$, для которой

$$P_{\theta, \vartheta} \{ \underline{\theta} \leq \theta \} = P_{\theta, \vartheta} \{ \underline{\theta} \geq \theta \} = \frac{1}{2},$$

одинаково часто недооценивает и переоценивает истинное значение и поэтому называется *медианно-несмещенной* (относительно связи этого свойства с общим понятием несмещенности см. задачу 3 главы 1).

С другой стороны, как было показано в главе 3, доверительные утверждения можно рассматривать как эквивалент семейства критериев. Нижеследующие рассуждения в значительной степени повторяют сказанное об этом соотношении в главе 3, правда, здесь мы ограничиваемся двусторонними гипотезами. Для каждого θ_0 обозначим $A(\theta_0)$ область принятия критерия уровня α для гипотезы $H(\theta_0): \theta = \theta_0$ (на время мы предположим, что этот критерий — нерандомизированный). Если

$$S(x) = \{ \theta: x \in A(\theta) \},$$

то

$$\theta \in S(x) \quad \text{эквивалентно} \quad x \in A(\theta) \quad (34)$$

и, следовательно,

$$P_{\theta, \vartheta} \{ \theta \in S(X) \} \geq 1 - \alpha \quad \text{при всех } \theta, \vartheta. \quad (35)$$

Таким образом, любое семейство областей принятия с уровнем α приводит через соотношение (34) к семейству доверительных множеств с доверительным уровнем $1 - \alpha$.

Обратно, возьмем любой класс доверительных множеств $S(x)$, удовлетворяющих (35). Пусть

$$A(\theta) = \{ x: \theta \in S(x) \}. \quad (36)$$

Тогда множества $A(\theta_0)$ являются областями принятия с уровнем α для гипотез $H(\theta_0): \theta = \theta_0$, а доверительное множество $S(x)$ показывает при каждом θ_0 , следует ли при данном наблюдаемом x принимать или отвергать гипотезу $\theta = \theta_0$ (с уровнем α).

Эти же рассуждения применимы, если множества $A(\theta_0)$ являются областями принятия для гипотез $\theta \leq \theta_0$. Ниже мы покажем, что, как правило (хотя и не всегда), одно- и двусторонние критерии приводят соответственно к односторонним доверительным границам и к доверительным интервалам.

Пример 4. Доверительные интервалы для среднего значения ξ нормального распределения с неизвестной дисперсией могут быть получены из

областей принятия $A(\xi_0)$ гипотез $H: \xi = \xi_0$. Последние даются формулой

$$\frac{|\sqrt{n}(\bar{x} - \xi_0)|}{\sqrt{\Sigma(x_i - \bar{x})^2/(n-1)}} \leq C,$$

где C определяется через t -распределение условием, чтобы при $\xi = \xi_0$ вероятность вышеприведенного неравенства была равна $1 - \alpha$ (см. формулы (17) и (19) раздела 2). Множество $S(x)$ состоит из тех чисел, которые удовлетворяют приведенному неравенству при $\xi = \xi_0$, т. е. заполняют собой интервал

$$\bar{x} - \frac{C}{\sqrt{n}} \sqrt{\frac{1}{n-1} \Sigma(x_i - \bar{x})^2} \leq \xi \leq \bar{x} + \frac{C}{\sqrt{n}} \sqrt{\frac{1}{n-1} \Sigma(x_i - \bar{x})^2}. \quad (37)$$

Мы получаем доверительные интервалы для ξ с коэффициентом доверия $1 - \alpha$.

Длина интервала (37) пропорциональна $\sqrt{\Sigma(x_i - \bar{x})^2}$, и математическое ожидание его длины пропорционально σ . Следовательно, при больших σ эти интервалы дают мало информации относительно неизвестного ξ . Это следует и из того факта (который порождал аналогичные трудности в задаче проверки гипотез), что нормальные распределения $N(\xi_0, \sigma^2)$ и $N(\xi_1, \sigma^2)$ с фиксированной разностью средних становятся неразличимыми при σ , стремящемся к бесконечности. Чтобы получить доверительные интервалы для ξ , длина которых не стремится к бесконечности вместе с σ , необходимо число наблюдений определять последовательным образом, согласованным с величиной σ . Последовательная процедура, приводящая к доверительным интервалам заданной длины, указана в задачах 15 и 16.

Однако и эта последовательная процедура в действительности не позволяет избавиться от трудности. Получив возможность контролировать длину интервала, мы теряем контроль над числом наблюдений. При $\sigma \rightarrow \infty$ число наблюдений, необходимое для получения доверительного интервала ограниченной длины, также стремится к бесконечности. В практических задачах мы часто имеем представление о порядке величины σ . Взяв выборку фиксированного объема или последовательную выборку, мы должны найти компромиссное решение, учитывающее желаемый доверительный уровень $1 - \alpha$, точность, определяемую длиной l интервала, и число наблюдений n , которое нам позволяют наши средства. При таком подходе две из трех величин $1 - \alpha$, l и n фиксированы, а третья является случайной величиной, распределение которой зависит от σ , так что она менее контролируема, чем другие. Если фиксировать $1 - \alpha$, то выбор между последовательной схемой и схемой с фиксированным объемом выборки определяется тем, что важнее контролировать, l или n .

Чтобы получить нижние доверительные границы для ξ , рассмотрим области принятия

$$\frac{\sqrt{n}(\bar{x} - \xi_0)}{\sqrt{\Sigma(x_i - \bar{x})^2/(n-1)}} \leq C_0$$

для проверки $\xi \leq \xi_0$ при альтернативе $\xi > \xi_0$. Множества $S(x)$ являются односторонними интервалами

$$\bar{x} - \frac{C_0}{\sqrt{n}} \sqrt{\frac{1}{n-1} \Sigma(x_i - \bar{x})^2} \leq \xi,$$

левый конец которых и дает искомую нижнюю границу $\underline{\xi}$. Если $\alpha = 1/2$, то константа C_0 равна 0. Соответствующая доверительная граница $\underline{\xi} = \bar{X}$ яв-

ляется медианно несмещенной оценкой ξ и в классе всех таких оценок равномерно минимизирует

$$P\{-\Delta_1 \leq \xi - \underline{\xi} \leq \Delta_2\} \text{ при всех } \Delta_1, \Delta_2 \geq 0$$

(доказательство см. на стр. 99).

5. Несмещенные доверительные множества

Доверительные множества можно рассматривать как семейство критериев для гипотез $\theta \in H(\theta')$ при альтернативах $\theta \in K(\theta')$, когда θ' меняется. Если коэффициент доверия равен $1 - \alpha$, то это означает, что все критерии имеют уровень α и

$$P_{\theta, \theta'}\{\theta' \in S(x)\} \geq 1 - \alpha \text{ при всех } \theta \in H(\theta') \text{ и всех } \theta'. \quad (38)$$

В случае, когда гипотеза $H(\theta')$ имеет вид $\theta = \theta'$ и $S(X)$ является интервалом $[\underline{\theta}(X), \bar{\theta}(X)]$, это согласуется с (33). В случае односторонней гипотезы $H(\theta') : \theta \leq \theta'$ и $S(X) = \{\theta : \underline{\theta}(X) \leq \theta\}$ условие (38) принимает форму $P_{\theta, \theta'}\{\underline{\theta}(X) \leq \theta'\} \geq 1 - \alpha$ для всех $\theta' \geq \theta$, эквивалентную (32). При такой интерпретации доверительных множеств вероятности

$$P_{\theta, \theta'}\{\theta' \in S(x)\}, \quad \theta \in K(\theta') \quad (39)$$

суть вероятности ошибочного принятия $H(\theta')$ (ошибки второго рода), и чем они меньше, тем лучше сами критерии.

Рассмотрим вопрос о качестве оценок. Формула (39) указывает вероятности покрытия «ложного» значения θ' . При заданной вероятности покрытия правильного значения доверительные множества дают тем больше информации, чем реже накрывают они ложные значения параметра. В этом смысле вероятности (39) могут служить мерой точности доверительных множеств. Обоснование (39) с точки зрения функций потерь для одностороннего случая было дано в разделе 5 главы 3.

При наличии мешающих параметров РНМ критерии обычно не существуют, соответственно не существуют и равномерно наиболее точные доверительные множества, т. е. такие, которые минимизируют (39) при всех θ' с $\theta \in K(\theta')$ при всех θ . Это наводит на мысль ограничиться доверительными множествами, в некотором смысле несмещенными. По аналогии с соответствующим определением для критериев, мы называем семейство доверительных множеств с доверительным уровнем $1 - \alpha$ *несмещенным*, если

$$P_{\theta, \theta'}\{\theta' \in S(x)\} \leq 1 - \alpha \text{ для всех } \theta' \text{ с } \theta \in K(\theta') \text{ при всех } \theta \text{ и } \theta', \quad (40)$$

так что вероятность покрытия этих «ложных» значений не превосходит доверительного уровня.

В отмеченных выше одно- и двустороннем случаях условие (40) приводится к следующим:

$$P_{\theta, \vartheta} \{ \underline{\theta} \leq \theta' \} \leq 1 - \alpha \text{ при всех } \theta' < \theta \text{ и всех } \vartheta,$$

и

$$P_{\theta, \vartheta} \{ \theta' \leq \bar{\theta} \} \leq 1 - \alpha \text{ при всех } \theta' \neq \theta \text{ и всех } \vartheta.$$

При таком определении несмещенности, несмещенные семейства критериев порождают несмещенные доверительные множества и обратно. Семейство доверительных множеств называется равномерно наиболее точным несмещенным семейством с доверительным уровнем $1 - \alpha$, если оно минимизирует вероятности

$$P_{\theta, \vartheta} \{ \theta' \in S(x) \} \text{ при всех } \theta' \text{ с } \theta \in K(\theta') \text{ для всех } \vartheta \text{ и } \theta$$

при ограничениях (38) и (40). На основе РНМ несмещенных критериев этой и предыдущей глав получаются равномерно наиболее точные несмещенные доверительные множества. В частности, это относится и к доверительным интервалам, построенным в предыдущем разделе. Ниже приводятся еще некоторые примеры.

Пример 5. Если $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$, то РНМ несмещенный критерий для гипотезы $\sigma = \sigma_0$ имеет область принятия (13)

$$C'_1 \leq \Sigma (x_i - \bar{x})^2 / \sigma_0^2 \leq C'_2,$$

где C'_1 и C'_2 определяются из (14). Следовательно, наиболее точным несмещенным доверительным интервалом для σ^2 является

$$\frac{1}{C'_2} \Sigma (x_i - \bar{x})^2 \leq \sigma^2 \leq \frac{1}{C'_1} \Sigma (x_i - \bar{x})^2.$$

Аналогично этому из (9) и (10) получается наиболее точная несмещенная верхняя доверительная граница для σ^2 :

$$\sigma^2 \leq \frac{1}{C_0} \Sigma (x_i - \bar{x})^2,$$

где

$$\int_{C_0}^{\infty} \chi_{n-1}^2(y) dy = 1 - \alpha.$$

Соответствующая нижняя доверительная граница является равномерно наиболее точной (без ограничения несмещенности). Это следует из раздела 10 главы 3.

Пример 6. Доверительные интервалы для разности $\Delta = \eta - \xi$ средних в двух нормальных совокупностях с одинаковой дисперсией получаются из критериев гипотезы $\eta - \xi = \Delta_0$. Если $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из $N(\xi, \sigma^2)$ и $N(\eta, \sigma^2)$ соответственно, и если $Y'_j = Y_j - \Delta_0$, $\eta' = \eta - \Delta_0$, то в терминах величин X_i и Y'_j гипотеза принимает вид $\eta' - \xi = 0$. Из (28) и

(30) мы видим, что РНМ несмещенная область принятия дается неравенством

$$\frac{|\bar{y} - \bar{x} - \Delta_0| / \sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{[\Sigma(x_i - \bar{x})^2 + \Sigma(y_j - \bar{y})^2] / (m + n - 2)}} \leq C,$$

где C находится из (31). Наиболее точные несмещенные доверительные интервалы для $\eta - \xi$ суть

$$(\bar{y} - \bar{x}) - CS \leq \eta - \xi \leq (\bar{y} - \bar{x}) + CS, \quad (41)$$

где

$$S^2 = \left(\frac{1}{m} + \frac{1}{n} \right) \frac{\Sigma(x_i - \bar{x})^2 + \Sigma(y_j - \bar{y})^2}{m + n - 2}.$$

Односторонние интервалы получаются аналогичным образом.

Пример 7. Если $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$, то наиболее точные несмещенные доверительные интервалы для $\Delta = \eta - \xi$ находятся с помощью (23) и имеют вид

$$\frac{1 - C_2}{C_2} \frac{\Sigma(y_j - \bar{y})^2}{\Sigma(x_i - \bar{x})^2} \leq \frac{\tau^2}{\sigma^2} \leq \frac{1 - C_1}{C_1} \frac{\Sigma(y_j - \bar{y})^2}{\Sigma(x_i - \bar{x})^2}, \quad (42)$$

где C_1 и C_2 вычисляются по (25)*. При $m = n$ выражение для интервалов упрощается:

$$\frac{1}{k} \frac{\Sigma(y_j - \bar{y})^2}{\Sigma(x_i - \bar{x})^2} \leq \frac{\tau^2}{\sigma^2} \leq k \frac{\Sigma(y_j - \bar{y})^2}{\Sigma(x_i - \bar{x})^2}, \quad (43)$$

где k определяется по таблицам F -распределения. Наиболее точная несмещенная нижняя доверительная граница для отношения дисперсий равна

$$\underline{\Delta} = \frac{1}{C_0} \frac{\Sigma(y_j - \bar{y})^2 / (n - 1)}{\Sigma(x_i - \bar{x})^2 / (m - 1)} \leq \frac{\tau^2}{\sigma^2}, \quad (44)$$

где C_0 дается формулой (22). Если в (22) положить $\alpha = 1/2$, то нижняя доверительная граница $\underline{\Delta}$ превращается в медианно-несмещенную оценку для τ^2/σ^2 . В классе всех таких оценок она минимизирует

$$P \left\{ -\Delta_1 \leq \frac{\tau^2}{\sigma^2} - \underline{\Delta} \leq \Delta_2 \right\} \quad \text{при всех } \Delta_1, \Delta_2 \geq 0.$$

(Доказательство см. на стр. 99.)

До сих пор мы предполагали, что критерии, на основе которых строятся доверительные множества, — нерандомизированные. Видоизменения, которые необходимо внести, когда это условие не выполнено, обсуждались в главе 3. Рандомизированные критерии при этом интерпретируются как нерандомизированные, но зависящие от X и вспомогательной переменной V , равномерно распределенной в единичном интервале. Если X — целочисленная величина (как в биномиальной и пуассоновской схемах),

*) Сравнение этих границ с полученными из критерия равных «хвостов» дается в работе: Schéffé, On the ratio of the variances of two normal populations, Ann. Math. Stat. 13 (1942), 371—388.

то эти критерии могут быть описаны в терминах непрерывной случайной величины $X + V$. Этим путем мы можем, например, получить наиболее точные несмещенные доверительные интервалы для биномиальной вероятности p , отправляясь от РНМ несмещенного критерия для гипотезы $H: p = p_0$ (см. пример 1 главы 4). Априори неясно, что соответствующие доверительные множества для p будут интервалами. Однако это вытекает из следующей леммы.

Лемма 1. Пусть X — действительная случайная величина, плотности которой $p_\theta(X)$ имеют отношение правдоподобия, монотонное относительно x . Предположим, что РНМ несмещенные критерии для гипотез $H(\theta_0): \theta = \theta_0$ существуют и имеют области принятия

$$C_1(\theta_0) \leq x \leq C_2(\theta_0).$$

Пусть, кроме того, эти критерии являются строго несмещенными. Тогда функции $C_i(\theta)$ строго возрастают по θ , и наиболее точные несмещенные доверительные интервалы для θ имеют вид

$$C_2^{-1}(x) \leq \theta \leq C_1^{-1}(x).$$

Доказательство. Пусть $\theta_0 < \theta_1$ и пусть $\beta_0(\theta)$ и $\beta_1(\theta)$ обозначают функции мощности указанных в формулировке критериев φ_0 и φ_1 для проверки гипотез $\theta = \theta_0$ и $\theta = \theta_1$. Из строгой несмещенности этих критериев вытекает, что

$$E_{\theta_0}[\varphi_1(X) - \varphi_0(X)] = \beta_1(\theta_0) - \alpha > 0 > \alpha - \beta_0(\theta_1) = E_{\theta_1}[\varphi_1(X) - \varphi_0(X)].$$

Таким образом, ни один из интервалов $[C_1(\theta_i), C_2(\theta_i)]$ ($i = 0, 1$) не содержит другого, и, как видно из леммы 2 (II) главы 3, $C_i(\theta_0) < C_i(\theta_1)$ для $i = 1, 2$. Следовательно, функции C_i имеют обратные функции, и неравенства, определяющие область принятия $H(\theta)$, эквивалентны $C_2^{-1}(x) \leq \theta \leq C_1^{-1}(x)$, что и требовалось доказать.

Сказанное иллюстрирует рис. 6. По границам $x = C_1(\theta)$ и $x = C_2(\theta)$ области принятия $A(\theta)$ мы находим при каждом фиксированном x доверительное множество $S(x)$ — это интервал тех значений θ , для которых $C_1(\theta) \leq x \leq C_2(\theta)$.

В соответствии с разделом 2 главы 4 условия леммы выполняются, в частности, для однопараметрического экспоненциального семейства, в предположении, что критерии нерандомизированы. В случае распределений биномиального или пуассоновского типа, когда семейство экспоненциально, но X — целочисленная величина (так что требуется рандомизация), доверительные интервалы могут быть получены применением леммы не к X , а к $X + V$. Здесь V не зависит от X и распределена равномерно на $(0, 1)$.

В лемме 1 предполагается, что распределение X зависит только от θ . Рассмотрим теперь экспоненциальное семейство (1), в котором дополнительно к θ имеются и мешающие параметры. РНМ несмещенные критерии гипотез $\theta = \theta_0$ выступают в этом случае в форме условных критериев при данном $T = t$. Как следствие доверительные интервалы также получаются условным об-

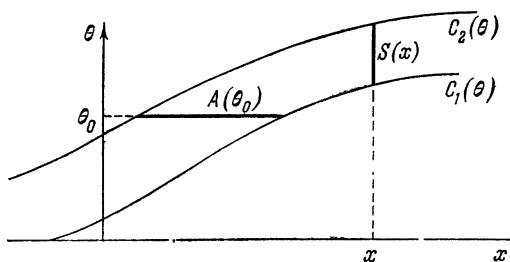


Рис. 6.

разом. Если условные распределения непрерывны, то области принятия имеют вид

$$C_1(\theta; t) \leq u \leq C_2(\theta; t),$$

где, по лемме 1, функции C_i являются возрастающими при каждом t . Поэтому доверительные интервалы суть

$$C_2^{-1}(u; t) \leq \theta \leq C_1^{-1}(u; t).$$

Если условные распределения дискретны, то непрерывность достигается введением дополнительной случайной величины, распределенной равномерно.

Пример 8. Пусть X и Y независимы и имеют распределения Пуассона со средними значениями λ и μ . Обозначим $\rho = \mu/\lambda$. Условное распределение Y при условии $X + Y = t$ является биномиальным $b(p, t)$ с

$$p = \frac{\rho}{1 + \rho}.$$

РНМ несмещенный критерий $\varphi(y, t)$ для гипотезы $\rho = \rho_0$ определяется при каждом t как РНМ несмещенный условный критерий для гипотезы $p = \rho_0/(1 + \rho_0)$. Если

$$\underline{p}(t) \leq p \leq \bar{p}(t)$$

суть соответствующие наиболее точные несмещенные доверительные интервалы для p при данном t , то наиболее точные несмещенные доверительные интервалы для μ/λ имеют вид

$$\frac{\underline{p}(t)}{[1 - \underline{p}(t)]} \leq \frac{\mu}{\lambda} \leq \frac{\bar{p}(t)}{[1 - \bar{p}(t)]}.$$

Биномиальные критерии, по которым определяются границы $\underline{p}(t)$ и $\bar{p}(t)$, рассматривались в примере 1 главы 4.

6. Регрессия

Изучение связи между двумя величинами X и Y может производиться на основе выборки и регистрации n пар измерений $(X_1, Y_1), \dots, (X_n, Y_n)$ (см. раздел 11 и задачу 11 в главе 6). Но часто бывает возможно регулировать одну из этих величин, например возраст обследуемого лица, или температуру, при которой проводится эксперимент, или силу применяемого лечебного препарата. Наблюдения $Y_1, \dots, Y_n$ величины Y получают при этом для наперед заданных значений $x_1, \dots, x_n$ величины x . Предположим, что при фиксированном x распределение Y нормально с постоянной дисперсией σ^2 и средним, которое зависит от x (его называют *регрессией Y на x*). Предположим, что эта зависимость линейна

$$E[Y | x] = \alpha + \beta x.$$

Полагая $v_i = (x_i - \bar{x}) / \sqrt{\Sigma(x_j - \bar{x})^2}$ и $\gamma + \delta v_i = \alpha + \beta x_i$, так что $\Sigma v_i = 0$, $\Sigma v_i^2 = 1$ и

$$\alpha = \gamma - \delta \frac{\bar{x}}{\sqrt{\Sigma(x_j - \bar{x})^2}}, \quad \beta = \frac{\delta}{\sqrt{\Sigma(x_j - \bar{x})^2}},$$

мы представим совместную плотность $Y_1, \dots, Y_n$ в форме

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \Sigma (y_i - \gamma - \delta v_i)^2 \right].$$

Эти плотности образуют экспоненциальное семейство (1) с

$$U = \Sigma v_i Y_i, \quad T_1 = \Sigma Y_i^2, \quad T_2 = \Sigma Y_i;$$

$$\theta = \frac{\delta}{\sigma^2}, \quad \theta_1 = -\frac{1}{2\sigma^2}, \quad \theta_2 = \frac{\gamma}{\sigma^2}.$$

Из такого представления вытекает существование РНМ несмещенных критериев для гипотез типа $a\gamma + b\delta = c$, где a, b, c — заданные постоянные, а следовательно, и существование наиболее точных несмещенных доверительных интервалов для параметра $\rho = a\gamma + b\delta$.

Чтобы найти для этих интервалов точное выражение, заметим, что РНМ несмещенный критерий для $H: \rho = \rho_0$ имеет область принятия

$$\frac{|b \Sigma v_i Y_i + a \bar{Y} - \rho_0| / \sqrt{(a^2/n) + b^2}}{\sqrt{[\Sigma(Y_i - \bar{Y})^2 - (\Sigma v_i Y_i)^2] / (n-2)}} \leq C, \quad (45)$$

где

$$\int_{-C}^C t_{n-1}(y) dy = 1 - \alpha$$

(см. задачу 20 и раздел 6 главы 7). Соответствующие доверительные интервалы для ρ имеют центры $b \sum v_i Y_i + a \bar{Y}$ и длину

$$L = 2C \sqrt{[(a^2/n) + b^2][\sum (Y_i - \bar{Y})^2 - (\sum v_i Y_i)^2]/(n-2)}.$$

Из преобразования, указанного в задаче 20, вытекает, что величина $[\sum (Y_i - \bar{Y})^2 - (\sum v_i Y_i)^2]/\sigma^2$ имеет χ^2 -распределение с $n-2$ степенями свободы, и поэтому математическое ожидание длины доверительных интервалов равно

$$E(L) = 2C_n \sigma \sqrt{\left(\frac{a^2}{n}\right) + b^2}.$$

Как правило, a и b являются функциями x . Если они находятся в распоряжении экспериментатора, и, следовательно, есть некоторая свобода в выборе a и b , то средняя длина L минимизируется одновременно с суммой $(a^2/n) + b^2$. Но в действительности средняя длина может и не быть хорошей характеристикой точности доверительных интервалов, так как короткие интервалы желательны, если они покрывают истинное значение параметра, и только при этом условии. Однако тот же вывод справедлив и по отношению к другим критериям качества, таким как математическое ожидание от $(\bar{\rho} - \rho)^2 + (\rho - \underline{\rho})^2$ или, более общим образом, от $f_1(|\bar{\rho} - \rho|) + f_2(|\rho - \underline{\rho}|)$, где f_1 и f_2 — возрастающие функции своих аргументов (см. задачу 20). Далее, тот же самый выбор a и b минимизирует вероятность покрытия интервалом любого ложного значения параметра. Мы будем поэтому рассматривать $(a^2/n) + b^2$ как обратную величину к мере точности интервалов.

Пример 9. Доверительные интервалы для наклона $\beta = \delta/\sqrt{\sum (x_i - \bar{x})^2}$ получаются из указанных выше при $a = 0$ и $b = 1/\sqrt{\sum (x_i - \bar{x})^2}$. Здесь точность возрастает вместе с $\sum (x_i - \bar{x})^2$ и, если точки должны быть выбраны в отрезке $[C_0, C_1]$, то точность становится максимальной при помещении половины всех x_i в каждый конец. Однако с практической точки зрения такой план часто не очень хорош, так как не позволяет проверить линейность регрессии.

Пример 10. Другой интересный параметр — математическое ожидание Y в точке $x = x_0$, т. е. $\alpha + \beta x_0$. Так как

$$\alpha + \beta x_0 = \frac{\gamma + \delta(x_0 - \bar{x})}{\sqrt{\sum (x_i - \bar{x})^2}},$$

то a и b имеют значения $a = 1$, $b = (x_0 - \bar{x})/\sqrt{\sum (x_i - \bar{x})^2}$. Максимум точности достигается при наименьшем $|\bar{x} - x_0|$, и если невозможно принять $\bar{x} = x_0$, то следует минимизировать $\sum (x_i - \bar{x})^2$.

Пример 11. Часто представляет интерес оценка точки x , в которой $\alpha + \beta x$ имеет заданное значение. Например, мы можем искать точку

$x = -\alpha/\beta$, в которой $E(Y | x) = 0$ или, что то же самое, искать значение $v = (x - \bar{x})/\sqrt{\sum (x_i - \bar{x})^2}$, при котором $\gamma + \delta v = 0$. Наиболее точные несмещенные доверительные множества для решения $-\gamma/\delta$ последнего уравнения можно получить из РНМ несмещенных критериев для гипотез $-\gamma/\delta = v_0$. Области принятия для этих гипотез даются (45) с $a = 1$, $b = v_0$ и $\rho_0 = 0$. Соответствующие доверительные множества для v содержат все точки, для которых

$$v^2 [C^2 S^2 - (\sum v_i Y_i)^2] - 2v\bar{Y} (\sum v_i Y_i) - \frac{1}{n} (C^2 S^2 - n\bar{Y}^2) \geq 0,$$

где $S^2 = [\sum (Y_i - \bar{Y})^2 - (\sum v_i Y_i)^2]/(n-2)$. Если квадратный трехчлен имеет корни $\underline{v}$ и $\bar{v}$, то доверительные утверждения принимают вид

$$\underline{v} \leq v \leq \bar{v} \quad \text{при} \quad |\sum v_i Y_i|/S > C$$

и

$$v \leq \underline{v} \quad \text{или} \quad v \geq \bar{v} \quad \text{при} \quad |\sum v_i Y_i|/S < C.$$

Удивительная на первый взгляд возможность получить в качестве доверительного множества внешность интервала, в действительности весьма естественна в этой задаче. Когда прямая $y = \gamma + \delta v$ почти параллельна оси v , ее пересечение с этой осью происходит в точке с большой по абсолютной величине абсциссой. Но знак абсциссы может измениться при совсем малом изменении угла наклона. Возможно также, что дискриминант квадратного трехчлена отрицателен

$$n\bar{Y}^2 + (\sum v_i Y_i)^2 < C^2 S^2.$$

При этом соответствующее квадратное уравнение не имеет вещественных решений. Из выписанного условия видно, что старший коэффициент квадратного трехчлена положителен, так что доверительным множеством в этом случае является вся прямая. То обстоятельство, что доверительные множества могут быть неограниченными, привело к соглашению, чтобы их использовать, исключая этот случай. Однако такое соглашение изменит вероятности, с которыми эти множества покрывают истинное значение параметра, а тем самым и величину коэффициента доверия *).

7. Критерии, основанные на перестановках

Допустим, что некоторый способ обработки сравнивается с контрольной ситуацией, где он не применяется. Если результаты измерений $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ представляют собой выборки из нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \sigma^2)$, то, как было показано в разделе 3, односторонний t -критерий является РНМ несмещенным критерием для проверки $H: \eta = \xi$ при альтернативе $\eta - \xi = \Delta > 0$. Мы рассмотрим теперь эту задачу без предположения о нормальности распределений. Вместо этого мы допустим, что величины X и Y выбраны из распределений с плот-

*) Метод, учитывающий этот эффект, был развит Нейманом, и на основе этого Э. Фикс были вычислены таблицы. Об этом см. в работе: Bennett, On the performance characteristic of certain methods of determining confidence limits, Sankhyā. 18 (1957), 1—12.

ностями $f(x)$ и $f(y - \Delta)$, где f — почти всюду непрерывна, а в остальном неизвестна. В этой непараметрической формулировке совместная плотность величин X и Y равна

$$f(x_1) \dots f(x_m) f(y_1 - \Delta) \dots f(y_n - \Delta), \quad f \in \mathcal{F}, \quad (46)$$

где $\mathcal{F}$ — семейство всех плотностей вероятности, непрерывных почти всюду.

Если изменчивость в совокупности, откуда берется выборка, велика, то чувствительность эксперимента часто может быть увеличена разбиением этой совокупности на более однородные подгруппы, отправляясь от таких характеристик, как, например, возраст или пол. Затем из i -й подгруппы отбирается N_i ($i = 1, \dots, c$) элементов, m_i из которых предназначается для контроля, а $n_i = N_i - m_i$ — для обработки по предлагаемому способу. Если результаты, полученные при наблюдениях над i -й подгруппой такой *расслоенной выборки*, обозначить

$$(X_{i1}, \dots, X_{im_i}; Y_{i1}, \dots, Y_{in_i}) = (Z_{i1}, \dots, Z_{iN_i}),$$

то плотность $Z = (Z_{11}, \dots, Z_{cN_c})$ примет вид

$$p_{\Delta}(z) = \prod_{i=1}^c [f_i(x_{i1}) \dots f_i(x_{im_i}) f_i(y_{i1} - \Delta) \dots f_i(y_{in_i} - \Delta)]. \quad (47)$$

Несмещенность некоторого критерия φ для проверки гипотезы $\Delta = 0$ при альтернативе $\Delta > 0$ влечет при всех $f_1, \dots, f_c$ равенство

$$\int \varphi(z) p_0(z) dz = \alpha \quad (dz = dz_{11} \dots dz_{cN_c}). \quad (48)$$

Теорема 3. Пусть $\mathcal{F}$ обозначает семейство всех вероятностных плотностей f , которые непрерывны почти всюду. Равенство (48) выполняется при всех $f_1, \dots, f_c \in \mathcal{F}$ тогда и только тогда, когда

$$\frac{1}{N_1! \dots N_c!} \sum_{z' \in S(z)} \varphi(z') = \alpha \text{ почти всюду,} \quad (49)$$

где $S(z)$ — множество точек, получающихся из z перестановкой при каждом $i = 1, \dots, c$ координат z_{ij} ($j = 1, \dots, N_i$) внутри i -й подгруппы всеми возможными $N_1! \dots N_c!$ способами.

Доказательство. Установим результат для $c = 1$. Отметим, что совокупность порядковых статистик $T(Z) = (Z^{(1)}, \dots, Z^{(N)})$ является полной достаточной статистикой для $\mathcal{F}$ (пример 6 главы 4). Поэтому условие, необходимое и достаточное для (48), имеет вид

$$E[\varphi(Z) | T(z)] = \alpha \text{ почти всюду.} \quad (50)$$

Множество $S(z)$ в рассматриваемом случае ($c = 1$) состоит из $N!$ точек, получаемых из z всевозможными перестановками координат, так что $S(z) = \{z' : T(z') = T(z)\}$. Из раздела 4 главы 2 следует, что условное распределение Z при данном $T(z)$ приписывает вероятность $1/N!$ каждой из $N!$ точек $S(z)$, т. е. (50) эквивалентно

$$\frac{1}{N!} \sum_{z' \in S(z)} \varphi(z') = \alpha \text{ почти всюду,} \quad (51)$$

что и требовалось доказать. Общий случай доказывается совершенно аналогично и помещен в упражнения (задача 21).

Критерии, удовлетворяющие (49), называются критериями *перестановок*. Это определение обобщается в задаче 31.

8. Наиболее мощные критерии перестановок

Пусть проверяется гипотеза $H: \Delta = 0$ об отсутствии эффекта предлагаемого способа обработки, причем основой служит расслоченная выборка с плотностью (47). В предыдущем разделе было показано, что в этом случае несмещенность влечет (49). Мы определим сейчас критерий, который удовлетворяет (49) и максимизирует мощность при любой фиксированной альтернативе (47), и, более того, при любой альтернативе с фиксированной плотностью $h(z)$.

Мощность критерия φ при альтернативе h равна

$$\int \varphi(z) h(z) dz = \int E[\varphi(Z) | t] dP^T(t).$$

Пусть $t = T(z) = (z^{(1)}, \dots, z^{(N)})$, так что $S(z) = S(t)$. Как мы видели в примере 7 и задаче 4 главы 2, условное математическое ожидание $\varphi(Z)$ при данном $T(Z) = t$ равно

$$\psi(t) = \frac{\sum_{z \in S(t)} \varphi(z) h(z)}{\sum_{z \in S(t)} h(z)}.$$

Чтобы максимизировать мощность φ при условии (49), необходимо максимизировать $\psi(t)$ при каждом t при этом же условии. Проблема сводится, таким образом, к определению функции φ , которая, подчиняясь условию

$$\sum_{z \in S(t)} \varphi(z) \frac{1}{N_1! \dots N_c!} = \alpha,$$

максимизирует

$$\sum_{z \in S(t)} \varphi(z) \frac{h(z)}{\sum_{z' \in S(t)} h(z')}.$$

По фундаментальной лемме Неймана — Пирсона к этому можно прийти, отклоняя гипотезу H для тех z из $S(t)$, для которых отношение

$$\frac{h(z) N_1! \dots N_c!}{\sum_{z' \in S(t)} h(z')}$$

слишком велико. Поэтому наиболее мощный критерий определяется критической функцией

$$\varphi(z) = \begin{cases} 1 & \text{при } h(z) > C[T(z)], \\ \gamma & \text{при } h(z) = C[T(z)], \\ 0 & \text{при } h(z) < C[T(z)]. \end{cases} \quad (52)$$

Для применения критерия $N_1! \dots N_c!$ точек каждого множества $S(z)$ упорядочивают в соответствии с величиной плотности h . Гипотеза отвергается для k больших значений и с вероятностью γ для $(k+1)$ -го, где k и γ определяются уравнением

$$k + \gamma = \alpha N_1! \dots N_c!.$$

Рассмотрим, в частности, альтернативы (47). Мы видим, что наиболее мощный критерий перестановок зависит от Δ и f_i и потому не будет РНМ критерием.

Особый интерес представляет класс нормальных альтернатив с одинаковой дисперсией

$$f_i = N(\xi_i, \sigma^2).$$

Наиболее мощный критерий при этих альтернативах оказывается не зависящим от ξ_i , σ^2 и Δ . Этот критерий уместен и в том случае, когда предполагается приближенная нормальность, но это предположение не является слишком надежным. В этом случае желательно поддерживать размер критерия на уровне α независимо к форме плотностей f_i и сделать критерий несмещенным при всех альтернативах (47). Однако внутри класса критериев, удовлетворяющих этим широким условиям, естественно произвести отбор с той целью, чтобы максимизировать мощность при таких альтернативах, которые мы вправе ожидать, т. е. при нормальных альтернативах.

Выбирая f_i , как указано выше, запишем (47) в форме

$$h(z) = (\sqrt{2\pi}\sigma)^{-N} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^c \left(\sum_{j=1}^{m_i} (z_{ij} - \xi_i)^2 + \sum_{j=m_i+1}^{N_i} (z_{ij} - \xi_i - \Delta)^2 \right) \right]. \quad (53)$$

Так как множитель $\exp \left[- \sum_i \sum_{j=1}^{N_i} (z_{ij} - \xi_i)^2 / 2\sigma^2 \right]$ сохраняет на $S(t)$ постоянное значение, то критерий (52) отвергает H , когда $\exp \left(\Delta \sum_i \sum_{j=m_i+1}^{N_i} z_{ij} \right) > C[T(z)]$, т. е. когда

$$\sum_{i=1}^c \sum_{j=1}^{n_i} y_{ij} = \sum_{i=1}^c \sum_{j=m_i+1}^{N_i} z_{ij} > C[T(z)]. \quad (54)$$

Из $N_1! \dots N_c!$ значений, которые статистика, входящая в (54), принимает на $S(t)$, различных будет только $\binom{N_1}{n_1} \dots \binom{N_c}{n_c}$, поскольку статистика принимает одинаковые значения в любых двух точках z' и z'' , для которых $(z'_{i1}, \dots, z'_{im_i})$ и $(z''_{i1}, \dots, z''_{im_i})$ получаются друг из друга перестановкой ($i = 1, \dots, c$). Поэтому достаточно рассматривать эти различные значения и отклонять H для k' наибольших, а для $(k' + 1)$ -го по величине — отклонять с вероятностью γ' , где

$$k' + \gamma' = \alpha \binom{N_1}{n_1} \dots \binom{N_c}{n_c}.$$

Критерий (54) является наиболее мощным при рассматриваемых нормальных альтернативах в классе всех несмещенных критериев уровня α при проверке гипотезы $H: \Delta = 0$ в первоначальном семействе (47) с $f_1, \dots, f_c \in \mathcal{F}$. Чтобы завершить доказательство этого утверждения, остается проверить несмещенность критерия при альтернативах (47). Мы докажем большее, а именно, что этот критерий является несмещенным при всех альтернативах, для которых X_{ij} ($j = 1, \dots, m_i$), Y_{ik} ($k = 1, \dots, n_i$) независимы и имеют функции распределения F_i, G_i соответственно, причем Y_{ik} стохастически больше чем X_{ij} , т. е. $G_i(z) \leq F_i(z)$ при всех z . Это утверждение вытекает из следующей леммы.

Лемма 2. Пусть $X_1, \dots, X_m; Y_1, \dots, Y_n$ — выборки из непрерывных распределений F, G , и пусть $\varphi(x_1, \dots, x_m; y_1, \dots, y_n)$ — критическая функция такая, что (а) ее математическое ожидание равно α , если $G = F$, и (б) $y_i \leq y'_i$ при $i = 1, \dots, n$ влечет

$$\varphi(x_1, \dots, x_m; y_1, \dots, y_n) \leq \varphi(x_1, \dots, x_m; y'_1, \dots, y'_n).$$

Тогда математическое ожидание $\beta = \beta(F, G)$ функции φ будет $\geq \alpha$ для всех пар распределений, для которых Y стохастически больше чем X . Это математическое ожидание $\leq \alpha$, если X стохастически больше чем Y .

Доказательство. По лемме 1 главы 3 существуют функции f, g и независимые случайные величины $V_1, \dots, V_{m+n}$ такие, что распределения $f(V_i)$ и $g(V_i)$ совпадают с F и G соответственно, и что $f(z) \leq g(z)$ при всех z . Тогда

$$E\varphi[f(V_1), \dots, f(V_m); f(V_{m+1}), \dots, f(V_{m+n})] = \alpha$$

и

$$E\varphi[f(V_1), \dots, f(V_m); g(V_{m+1}), \dots, g(V_{m+n})] = \beta.$$

Так как при всех $(v_1, \dots, v_{m+n})$

$$\begin{aligned} \varphi[f(v_1), \dots, f(v_m); f(v_{m+1}), \dots, f(v_{m+n})] &\leq \\ &\leq \varphi[f(v_1), \dots, f(v_m); g(v_{m+1}), \dots, g(v_{m+n})] \end{aligned}$$

то такое же неравенство верно и для математических ожиданий обеих частей, т. е. $\alpha \leq \beta$.

Доказательство для случая, когда X стохастически больше чем Y , совершенно аналогично. Лемма обобщается и на случай c векторов $(X_{i1}, \dots, X_{im_i}; Y_{i1}, \dots, Y_{in_i})$ с распределениями (F_i, G_i) . Пусть математическое ожидание функции φ равно α при $F_i = G_i$ и φ не убывает по каждому y_{ij} , когда все остальные переменные фиксированы. Как и раньше, математическое ожидание φ будет $\geq \alpha$, если случайные величины с распределениями G_i стохастически больше величин с распределениями F_i .

Применяя лемму к критерию перестановок (54), достаточно рассмотреть случай $c = 1$, поскольку для общего случая рассуждения совершенно аналогичны. Так как при $F = G$ вероятность отклонения гипотезы критерием (54) равна α , остается проверить, что критическая функция φ удовлетворяет (б). Отме-

тим, что $\varphi = 1$, если $\sum_{i=m+1}^{m+n} z_i$ больше достаточно многих сумм

$\sum_{i=m+1}^{m+n} z_{j_i}$, т. е. если достаточное число разностей

$$\sum_{i=m+1}^{m+n} z_i - \sum_{i=m+1}^{m+n} z_{j_i}$$

положительно. Для отдельной перестановки $(j_1, \dots, j_{m+n})$

$$\sum_{i=m+1}^{m+n} z_i - \sum_{i=m+1}^{m+n} z_{j_i} = \sum_{i=1}^p z_{s_i} - \sum_{i=1}^p z_{r_i},$$

где $r_1 < \dots < r_p$ обозначают те из целых $j_{m+1}, \dots, j_{m+n}$, которые $\leq m$, а $s_1 < \dots < s_p$ — те из целых $m+1, \dots, m+n$, которые не входят во множество $(j_{m+1}, \dots, j_{m+n})$. Если разность $\sum z_{s_i} - \sum z_{r_i}$ положительна и $y_i \leq y'_i$, т. е. $z_i \leq z'_i$, при $i = m+1, \dots, m+n$, то разность $\sum z'_{s_i} - \sum z_{r_i}$ также положительна и, следовательно, φ удовлетворяет (б).

Такие же рассуждения показывают, что вероятность отклонения гипотезы $\leq \alpha$, если плотность рассматриваемых величин дается формулой (47) с $\Delta \leq 0$. Поэтому критерий остается пригодным при замене гипотезы $\Delta = 0$ на $\Delta \leq 0$.

Использование критерия перестановок (54) в точной форме возможно только при небольших объемах выборок N_i ; в противном случае, вычислительные трудности становятся непреодолимыми. При $c = 1$, например, для определения критической точки $C[T(z)]$ требуется отыскать множества индексов $(j_1, \dots, j_n)$,

дающих k наибольших по величине значений сумм $\sum_{i=1}^n z_{j_i}$.

Здесь k — наибольшее целое, не превосходящее $\alpha \binom{N}{n}$. Если $\alpha = 0,05$, то при $m = n = 5$ будет $k = 12$, но уже при $m = n = 10$ величина k превосходит 9000. Однако для случая больших выборок можно указать весьма удобную аппроксимацию. Умножая обе части неравенства

$$\sum y_i > C[T(z)]$$

на $[(1/m) + (1/n)]$ и вычитая затем $(\sum x_i + \sum y_i)/m$, мы для $c = 1$ приведем критическую область к виду $\bar{y} - \bar{x} > C[T(z)]$

или $W = (\bar{y} - \bar{x}) / \sqrt{\sum_{i=1}^N (z_i - \bar{z})^2} > C[T(z)]$, так как знаменатель у W постоянен на $S(z)$ и потому зависит только от $T(z)$. Последнее неравенство, как было показано в конце раздела 3, эквивалентно следующему

$$\frac{(\bar{y} - \bar{x}) / \sqrt{\frac{1}{m} + \frac{1}{n}}}{\sqrt{[\sum (x_i - \bar{x})^2 + \sum (y_j - \bar{y})^2] / (m + n - 2)}} > C[T(z)]. \quad (55)$$

Критическая область имеет такой же вид, как у t -критерия, только константа C_0 из (27) заменяется случайной величиной. Можно показать, что при справедливости гипотезы, т. е. когда величины Z независимы и одинаково распределены, и при условиях (а) $E|Z|^3 < \infty$ (б) m/n не приближается ни к нулю, ни к бесконечности при m и n , стремящихся к бесконечности, разность между случайной критической точкой $C[T(Z)]$ и C_0 стремится к нулю по вероятности*). В пределе, следовательно, критерий пе-

*) Соответствующие предельные теоремы и ссылки на исходные работы Двассы, Хёффдинга, Нётер, Вальда и Вольфовица даны в разделе 6 главы 6 книги: Fraser, Nonparametric Methods in Statistics, N. Y. Wiley, 1957. О более точных аппроксимациях критерия перестановок см. Box and Andersen, Permutation theory in the derivation of robust criteria and the study of departures from assumption, J. Roy. Stat. Soc., Ser. B, XVII (1955), pp. 1—34.

рестановок становится эквивалентным t -критерию (27) — (29). Таким образом, *при больших объемах выборки критерий рестановок может быть аппроксимирован стандартным t -критерием*. В точности аналогичный результат справедлив и при $c > 1$. Соответствующий t -критерий указан в задаче 7 главы 7.

9. Рандомизация как основа статистических выводов

Проблема проверки действительности какого-либо способа обработки рассматривалась в разделе 3 (в предположении, что результаты измерения как при обработке, так и в контрольных испытаниях, т. е. $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$, образуют выборки из нормального распределения) и в разделах 7 и 8 (без предположения нормальности).

Допустим, что для проведения эксперимента из рассматриваемой совокупности отбирается наудачу $m + n$ элементов (пациентов, растений или образцов материала и т. д.). Из них n подвергаются обработке, а оставшиеся m служат для контроля. Затем измеряется характеристика, на которую, возможно, влияет обработка, что приводит к результатам $X_1, \dots, X_m; Y_1, \dots, Y_n$.

Для определенности предположим, что обработка состоит в инъекции лекарства и что имеется $m + n$ ампул, предназначенных для $m + n$ пациентов. Результат i -го измерения может рассматриваться как сумма двух компонент. Одна, скажем U_i , связана с i -м пациентом. Другая — V_i с i -й ампулой и с обстоятельствами, при которых она использована и при которых измерен соответствующий результат. Мы предположим, что случайные величины U_i и V_i независимы, причем распределение V_i нормально и равно $N(\eta, \sigma^2)$, если ампула содержит исследуемое лекарство, или $N(\xi, \sigma^2)$, если ампула употребляется для контроля. Принимая дополнительно, что U_i образуют выборку из $N(\mu, \sigma_1^2)$, мы видим, что величины X и Y независимы и распределены нормально с общей дисперсией $\sigma^2 + \sigma_1^2$ и средними

$$E(X) = \mu + \xi, \quad E(Y) = \mu + \eta.$$

С точностью до обозначений их совместное распределение задается формулой (26), и гипотеза $\eta = \xi$ может быть проверена с помощью стандартного t -критерия.

К сожалению, в реальных условиях часто невозможно добиться того, чтобы пациенты или другие экспериментальные единицы образовывали случайную выборку из соответствующей генеральной совокупности: мы рассматриваем или совокупность пациентов, находящихся в данном госпитале в данный момент, или совокупность лиц, добровольно согласившихся на эксперимент и т. п. В этом случае величины U следовало бы рассматривать, как неизвестные константы, так как они не являются

результатом какой-либо определенной выборочной процедуры. Это предположение уместно и в другом контексте. Допустим, что экспериментальными единицами являются машины некоторого цеха или поля некоторой фермы. Если эксперимент имеет целью отыскать метод, наилучший именно для этого цеха или этой фермы, то только указанные единицы имеют отношение к задаче. В этом случае повторение эксперимента состояло бы в сравнении двух способов обращения с одними и теми же машинами или полями, а не с новыми, наудачу выбранными из более обширной совокупности. В этом случае сами экспериментальные единицы постоянны, а следовательно, постоянны и величины u .

При высказанных допущениях совместная плотность $m + n$ результатов измерений равна

$$\frac{1}{(\sqrt{2\pi} \sigma)^{m+n}} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^m (x_i - u_i - \xi)^2 + \sum_{j=1}^n (y_j - u_{m+j} - \eta)^2 \right) \right].$$

Так как величины u совершенно произвольны, то ясно, что невозможно провести различие между гипотезой H : $\eta = \xi$ и альтернативами K : $\eta > \xi$. Действительно, каждое распределение из K принадлежит также H и обратно. Поэтому наиболее мощный критерий уровня α для проверки H при любой простой альтернативной гипотезе относительно величин ξ , η , σ и u_i отвергает H с вероятностью α независимо от исхода наблюдений.

Данные, которые могут служить базисом для проверки эффективности способа обработки, могут быть получены с помощью весьма важного приема — *рандомизации*. Предположим, что $N = m + n$ пациентам прописаны N ампул, причем наудачу, т. е. таким образом, что каждое из $N!$ возможных соответствий имеет одну и ту же вероятность $1/N!$. Тогда при каком-либо данном соответствии результаты N измерений будут распределены нормально с дисперсией σ^2 и средними $\xi + u_{i_i}$ ($i = 1, \dots, m$) и $\eta + u_{i_j}$ ($i = m + 1, \dots, m + n$). Совместная плотность величин

$$(Z_1, \dots, Z_N) = (X_1, \dots, X_m; Y_1, \dots, Y_n)$$

равна, следовательно,

$$\frac{1}{N!} \sum_{(j_1, \dots, j_N)} \frac{1}{(\sqrt{2\pi} \sigma)^N} \times \\ \times \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^m (x_i - u_{i_{j_i}} - \xi)^2 + \sum_{i=1}^n (y_i - u_{i_{m+j_i}} - \eta)^2 \right) \right], \quad (56)$$

где внешнее суммирование производится по всем $N!$ перестановкам $(j_1, \dots, j_N)$ чисел $(1, \dots, N)$. При гипотезе $\eta = \xi$ эта плот-

ность имеет вид

$$\frac{1}{N!} \sum_{(j_1, \dots, j_N)} \frac{1}{(\sqrt{2\pi} \sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^N (z_i - \xi_{j_i})^2 \right], \quad (57)$$

где $\xi_{j_i} = u_{j_i} + \xi = u_{j_i} + \eta$.

Без рандомизации появление множества y_k , больших по сравнению с x_k , могло бы быть объяснено целиком в терминах компонент u_i , связанных с отдельными пациентами. Однако если связь между u_i и y_k случайна, то эффект этих компонент будет в среднем таков же, как и компонент, приписанных x_k . Следовательно, заметное превосходство второй выборки становится весьма неправдоподобным в условиях проверяемой гипотезы и потому должно быть приписано эффективности способа обработки.

Метод сопоставления «наудачу» способов обработки с экспериментальными единицами позволяет построить критерий уровня α для проверки гипотезы $\eta = \xi$, мощность которого больше α при всех альтернативах $\eta - \xi > 0$. Действительная мощность этого критерия зависит, однако, не только от альтернативного значения $\eta - \xi$, которое измеряет эффективность способа обработки, но также и от эффектов u_i отдельных единиц. В частности, если изменчивость величин u_i чрезмерно велика, то она может «подавить» эффект обработки (так же, как могло бы быть при возрастании дисперсии σ^2). В соответствии с этим, мощность критерия была бы мала и не позволила бы обнаруживать альтернативы $\eta - \xi$.

В таких случаях чувствительность эксперимента может быть повышена приемом, в точности аналогичным способу «расслоенной выборки», обсуждавшемуся в разделе 7. В рассматриваемом случае это означает замену описанного выше процесса полной рандомизации более ограниченной процедурой рандомизации. Экспериментальный материал разбивается на подгруппы, более однородные, чем материал в целом, так что внутри каждой группы разности между величинами u малы. В опытах с животными, например, этого часто можно достичь разделением по пометам. Рандомизация применяется только внутри каждой из групп. Если i -я группа содержит N_i единиц, то n_i из них выбираются для обработки, а остающиеся $m_i = N_i - n_i$ служат для контроля ($\sum N_i = N$, $\sum m_i = m$, $\sum n_i = n$).

Примером такого подхода может служить метод *сходных пар*. В нем экспериментальные единицы делятся на пары. Единицы каждой пары стремятся подобрать возможно более похожими друг на друга во всех существенных для эксперимента отношениях, так что внутри каждой пары разность значений делается возможно меньшей. Предположим, что материал разбит на n таких пар, и обозначим соответствующие единицам эффекты

(величины U предыдущих рассуждений) $U_1, U'_1; \dots; U_n, U'_n$. Пусть первый и второй члены каждой пары или подвергаются обработке, или служат для контроля, соответственно, и пусть результаты наблюдений за i -й парой равны X_i и Y_i . Если подбор пар вполне удачен, что может быть, например, когда один и тот же пациент дважды используется при исследовании снотворного средства, то $U'_i = U_i$ при всех i и плотность величин X и Y равна

$$\frac{1}{(\sqrt{2\pi}\sigma)^{2n}} \exp\left[-\frac{1}{2\sigma^2}\left[\sum (x_i - \xi - u_i)^2 + \sum (y_i - \eta - u_i)^2\right]\right]. \quad (58)$$

РНМ несмещенный критерий для проверки $H: \eta = \xi$ при альтернативе $\eta > \xi$ определяется в этом случае в терминах разностей $W_i = Y_i - X_i$; его критическая область имеет вид

$$\frac{\sqrt{n} \bar{w}}{\sqrt{\frac{1}{n-1} \sum (w_i - \bar{w})^2}} > C \quad (59)$$

(см. задачу 25).

Однако обычно не желают доверять предположению $u'_i = u_i$ даже после объединения в пары, и снова становится необходимой рандомизация. Так как, по-видимому, объединение в пары делает изменчивость u внутри пары значительно меньшей, чем по всей совокупности, то рандомизация проводится только внутри каждой пары. Одна из единиц каждой пары выбирается с вероятностью $1/2$ для обработки, в то время как остальные служат для контроля. Плотность величин X и Y в этом случае равна

$$\frac{1}{2^n} \frac{1}{(\sqrt{2\pi}\sigma)^{2n}} \prod_{i=1}^n \left\{ \exp\left[-\frac{1}{2\sigma^2} [(x_i - \xi - u_i)^2 + (y_i - \eta - u_i')^2]\right] + \right. \\ \left. + \exp\left[-\frac{1}{2\sigma^2} [(x_i - \xi - u_i')^2 + (y_i - \eta - u_i)^2]\right] \right\}. \quad (60)$$

При гипотезе $\eta = \xi$ это превращается в

$$\frac{1}{2^n} \sum \frac{1}{(\sqrt{2\pi}\sigma)^{2n}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^2 (z_{ij} - \zeta'_{ij})^2\right], \quad (61)$$

где

$$z_{i1} = x_i, \quad z_{i2} = y_i, \quad \zeta_{i1} = \xi + u_i, \quad \zeta_{i2} = \eta + u'_i \quad (i = 1, \dots, n).$$

Внешнее суммирование распространяется на 2^n точек $\zeta' = (\zeta'_{11}, \dots, \zeta'_{n2})$, для которых $(\zeta'_{11}, \zeta'_{12})$ равно или (ζ_{11}, ζ_{12}) или (ζ_{12}, ζ_{11}) .

10. Критерии перестановок и рандомизация

В предыдущем разделе было показано, что рандомизация дает базис для проверки гипотезы $\eta = \xi$ об отсутствии эффекта обработки без каких-либо предположений об экспериментальных

единицах. В настоящем разделе будет получен специальный критерий для этой проблемы. В случае, когда экспериментальные единицы рассматриваются как постоянные, совместная плотность результатов наблюдений дается формулой (56) (при полной рандомизации) или формулой (60) (при использовании сходных пар). Более общим образом, пусть экспериментальный материал разбит на c подгрупп, рандомизация применяется в каждой из подгрупп и пусть результаты наблюдений в i -й подгруппе равны

$$(Z_{i1}, \dots, Z_{iN_i}) = (X_{i1}, \dots, X_{im_i}; Y_{i1}, \dots, Y_{in_i}).$$

Для любой точки $u = (u_{11}, \dots, u_{cN_c})$ обозначим, как и раньше, знаком $S(u)$ множество из $N_1! \dots N_c!$ точек, получаемых из u всевозможными $N_1! \dots N_c!$ перестановками координат внутри каждой подгруппы. Тогда совместная плотность величин Z равна

$$\frac{1}{N_1! \dots N_c!} \sum_{u' \in S(u)} \frac{1}{(\sqrt{2\pi} \sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^c \left(\sum_{j=1}^{m_i} (z_{ij} - \xi - u'_{ij})^2 + \sum_{j=m_i+1}^{N_i} (z_{ij} - \eta - u'_{ij})^2 \right) \right] \quad (62)$$

и при гипотезе отсутствия эффекта обработки

$$p_{\sigma, \xi}(z) = \frac{1}{N_1! \dots N_c!} \sum_{\xi' \in S(\xi)} \frac{1}{(\sqrt{2\pi} \sigma)^N} \times \\ \times \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^c \sum_{j=1}^{N_i} (z_{ij} - \xi'_{ij})^2 \right]. \quad (63)$$

Может случиться, что не все координаты u или ξ различны. Тогда, если некоторые из точек $S(u)$ или $S(\xi)$ совпадают, то каждая точка должна считаться с учетом ее кратности. Более точно, если $N_1! \dots N_c!$ рассматриваемых перестановок $N_1 + \dots + N_c$ координат обозначить g_k , $k = 1, \dots, N_1! \dots N_c!$, то в качестве $S(\xi)$ можно взять упорядоченное множество точек $g_k \xi$, $k = 1, \dots, N_1! \dots N_c!$. Тогда соотношение (63), например, принимает вид

$$p_{\sigma, \xi}(z) = \frac{1}{N_1! \dots N_c!} \sum_{k=1}^{N_1! \dots N_c!} \frac{1}{(\sqrt{2\pi} \sigma)^N} \exp \left(-\frac{1}{2\sigma^2} |z - g_k \xi|^2 \right),$$

где $|u^2|$ обозначает $\sum_{i=1}^c \sum_{j=1}^{N_i} u_{ij}^2$.

Теорема 4. Для того чтобы критическая функция φ удовлетворяла неравенству

$$\int \varphi(z) p_{\sigma^2}(z) dz \leq \alpha \quad (dz = dz_{11} \dots dz_{cN_c}) \quad (64)$$

при всех $\sigma > 0$ и всех векторах ξ , необходимо и достаточно выполнение условия

$$\frac{1}{N_1! \dots N_c!} \sum_{z' \in S(z)} \varphi(z') \leq \alpha \text{ почти всюду.} \quad (65)$$

Доказательство основано на следующей лемме.

Лемма 3. Пусть A — множество в N мерном пространстве, мера Лебега $\mu(A)$ которого положительна. Тогда для любого $\varepsilon > 0$ найдутся такие действительные числа $\sigma > 0$ и $\xi_1, \dots, \xi_N$, что

$$P\{(X_1, \dots, X_n) \in A\} \geq 1 - \varepsilon,$$

где X_i — независимые нормально распределенные случайные величины со средними $E(X_i) = \xi_i$ и дисперсиями $\sigma_{X_i}^2 = \sigma^2$.

Доказательство. Не ограничивая общности, допустим, что $\mu(A) < \infty$. Для любого $\eta > 0$ найдется такой квадрат Q , что

$$\mu(Q \cap \tilde{A}) \leq \eta \mu(Q).$$

Это следует из того, что почти в каждой своей точке A имеет метрическую плотность 1^*) (или из более элементарного факта — возможности аппроксимировать измеримое множество по мере объединения непересекающихся квадратов). Пусть a таково, что

$$\frac{1}{\sqrt{2\pi}} \int_{-a}^a \exp\left(-\frac{t^2}{2}\right) dt = \left(1 - \frac{\varepsilon}{2}\right)^{1/N},$$

и пусть

$$\eta = \frac{\varepsilon}{2} \left(\frac{\sqrt{2\pi}}{2a}\right)^N.$$

Если $(\xi_1, \dots, \xi_N)$ — центр Q , и если $\sigma = b/a = (1/2a)[\mu(Q)]^{1/N}$, где $2b$ — длина стороны Q , то

$$\begin{aligned} & \frac{1}{(\sqrt{2\pi}\sigma)^N} \int_{\tilde{A} \cap \tilde{Q}} \exp\left[-\frac{1}{2\sigma^2} \sum (x_i - \xi_i)^2\right] dx_1 \dots dx_N \leq \\ & \leq \frac{1}{(\sqrt{2\pi}\sigma)^N} \int_{\tilde{Q}} \exp\left[-\frac{1}{2\sigma^2} \sum (x_i - \xi_i)^2\right] dx_1 \dots dx_N = \\ & = 1 - \left[\frac{1}{\sqrt{2\pi}} \int_{-a}^a \exp\left(-\frac{t^2}{2}\right) dt \right]^N = \frac{\varepsilon}{2}. \end{aligned}$$

*) См., например, Hobson, Theory of Functions of a Real Variable, vol. 1, Cambridge Univ. Press, 3rd ed., 1927, 194.

С другой стороны,

$$\frac{1}{(\sqrt{2\pi}\sigma)^N} \int_{\hat{A} \cap Q} \exp \left[-\frac{1}{2\sigma^2} \sum (x_i - \xi_i)^2 \right] dx_1 \dots dx_N \leqslant \frac{1}{(\sqrt{2\pi}\sigma)^N} \mu(\hat{A} \cap Q) < \frac{\varepsilon}{2}.$$

Складывая почленно оба неравенства, приходим к искомому результату.

Доказательство теоремы. Пусть φ — любая критическая функция и пусть

$$\Psi(z) = \frac{1}{i_1 \dots N_c!} \sum_{z' \in S(z)} \varphi(z').$$

Если (65) не выполняется, то существует такое $\eta > 0$, что на множестве положительной меры $\Psi(z) > \alpha + \eta$. По лемме существуют $\sigma > 0$ и $\xi = (\xi_{11}, \dots, \xi_{cN_c})$ такие, что $P\{Z \in A\} > 1 - \eta$, где $Z_{11}, \dots, Z_{cN_c}$ независимы и распределены нормально с общей дисперсией σ^2 и средними $E(Z_{ij}) = \xi_{ij}$. Мы имеем

$$\begin{aligned} \int \varphi(z) p_{\sigma, \xi}(z) dz &= \int \Psi(z) p_{\sigma, \xi}(z) dz \geqslant \\ &\geqslant \int_A \Psi(z) \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \sum \sum (z_{ij} - \xi_{ij})^2 \right] dz > \\ &> (\alpha + \eta)(1 - \eta), \end{aligned} \quad (66)$$

что больше α , так как $\alpha + \eta < 1$. Этим доказано, что (64) влечет (65). Обратное утверждение следует из первого равенства в (66).

Следствие 2. Пусть H — класс плотностей

$$\{p_{\sigma, \xi}(z): \sigma > 0, \quad -\infty < \xi_{ij} < \infty\}.$$

Полным семейством критериев для проверки H с уровнем α может служить класс $\mathcal{C}$ критериев, для которых

$$\frac{1}{i_1 \dots N_c!} \sum_{z' \in S(z)} \varphi(z') = \alpha \text{ почти всюду.} \quad (67)$$

Доказательство. Следствие утверждает, что для любого данного критерия φ_0 с уровнем α найдется элемент φ в $\mathcal{C}$, который равномерно не менее мощен, чем φ_0 . По предыдущей теореме среднее значение φ_0 на каждом множестве $S(z)$ не превосходит α . На множествах, где это неравенство является строгим, мы можем увеличить φ_0 таким образом, чтобы получить критическую функцию φ , удовлетворяющую (67) и такую, что $\varphi_0(z) \leqslant \varphi(z)$ для любых z . При всех альтернативах мощность φ не меньше мощности φ_0 , т. е. φ — искомый критерий. Точная кон-

струкция критерия ϕ , показывающая его измеримость, приведена в задаче 28. Из следствия 2 видно, что в нормальном случае модель рандомизации (62) приводит в точности к тому же классу критериев, который был найден ранее в несколько иной обстановке. Из раздела 8 можно вывести, что наиболее мощный критерий уровня α для проверки (63) при простой альтернативе (62) дается формулой (52) с $h(z)$, равной плотности (62). Если $\eta - \xi = \Delta$, то критическая область этого критерия имеет вид

$$\sum_{u' \in S(u)} \exp \left[\frac{1}{\sigma^2} \sum_{i=1}^c \left(\sum_{j=1}^{N_i} z_{ij} u'_{ij} + \Delta \sum_{j=m_i+1}^{N_i} (z_{ij} - u'_{ij}) \right) \right] > C[T(z)], \quad (68)$$

так как обе суммы $\sum \sum z_{ij}$ и $\sum \sum z_{ij}^2$ постоянны на $S(z)$ и потому зависят только от $T(z)$. Мы видим, что критерий зависит от Δ и индивидуальных эффектов u_{ij} , так что РНМ критерий не существует.

Среди альтернатив (62) имеется один подкласс, занимающий центральное положение и представляющий особенный интерес. Это класс альтернатив, выделяемый предположением, что индивидуальные эффекты u_i образуют выборку из нормального распределения. Хотя трудно ожидать, что это предположение выполняется точно, во многих случаях можно с основанием предполагать его приближенное выполнение. Соответствующий подкласс альтернатив дается вероятностными плотностями

$$\frac{1}{(\sqrt{2\pi} \sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^c \left(\sum_{j=1}^{m_i} (z_{ij} - u_i - \xi)^2 + \sum_{j=m_i+1}^{N_i} (z_{ij} - u_i - \eta)^2 \right) \right]. \quad (69)$$

Эти альтернативы привлекательны и с несколько иной точки зрения. Процедура случайного (внутри каждой подгруппы) сопоставления способов обработки с экспериментальными единицами уместна, как мы видели, тогда, когда изменчивость u внутри каждой группы мала. Эта процедура используется при уверенности, что последнее условие выполняется, что приводит к предположению, может быть приближенному, о постоянстве $u_{ij} = u_i$. Такой случай может рассматриваться как предельный для нормальных распределений с дисперсией, стремящейся к нулю. Для последних плотность задается формулой (69).

Так как альтернативы (69) совпадают с альтернативами (53) раздела 8 с $u_i - \xi = \xi_i$, $u_i - \eta = \xi_i - \Delta$, то критерий перестановок (54) оказывается наиболее мощным критерием для проверки

гипотезы $\eta = \xi$ в нормальной модели рандомизации (62) при альтернативах (69) с $\eta - \xi > 0$. Критерий сохраняет это свойство в значительно более общей обстановке, в которой ни нормальность, ни выборочные свойства величин U не участвуют. Пусть совместная плотность рассматриваемых случайных величин равна

$$\sum_{u' \in S(u)} \prod_{i=1}^c \left[\prod_{j=1}^{m_i} f_i(z_{ij} - u'_{ij} - \xi) \prod_{j=m_i+1}^{N_i} f_i(z_{ij} - u'_{ij} - \eta) \right],$$

где f_i почти всюду непрерывны, а в остальном произвольны*). При гипотезе $H: \eta = \xi$ эта плотность симметрична по переменным $(z_{i1}, \dots, z_{iN_i})$ i -й подгруппы при каждом i , так что критерий перестановок (49) имеет вероятность отклонения, равную α при всех распределениях из H . По следствию 2 эти критерии перестановок образуют полный класс, что и требовалось установить.

11. Проверка независимости в двумерном нормальном распределении

До сих пор методы настоящей главы применялись, главным образом, к проблеме двух выборок. Мы используем теперь два рассматривавшихся подхода — нормальную модель раздела 3 и непараметрическую модель раздела 7 — к проверке гипотезы независимости в двумерном нормальном распределении.

Плотность вероятности для выборки $(X_1, Y_1), \dots, (X_n, Y_n)$ из двумерного нормального распределения равна

$$\frac{1}{(2\pi\sigma\tau\sqrt{1-\rho^2})^n} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{1}{\sigma^2} \sum (x_i - \xi)^2 - \frac{2\rho}{\sigma\tau} \sum (x_i - \xi)(y_i - \eta) + \frac{1}{\tau^2} \sum (y_i - \eta)^2 \right) \right]. \quad (70)$$

Здесь (ξ, σ^2) и (η, τ^2) являются средними и дисперсиями X и Y соответственно, а ρ — коэффициент корреляции между X и Y . Гипотезы $\rho \leq \rho_0$ и $\rho = \rho_0$ могут быть исследованы методами настоящей главы, и такое исследование будет предпринято в главе 6. Здесь же мы рассмотрим только гипотезу $\rho = 0$, т. е. гипотезу независимости X и Y , и соответствующую одностороннюю гипотезу $\rho \leq 0$.

Семейство плотностей (70) имеет экспоненциальную форму (1) с

$$U = \sum X_i Y_i, \quad T_1 = \sum X_i^2, \quad T_2 = \sum Y_i^2, \quad T_3 = \sum X_i, \quad T_4 = \sum Y_i$$

*) В действительности требуется лишь, чтобы $f_1, \dots, f_c \in \mathcal{F}$, где $\mathcal{F}$ — любое семейство, содержащее все нормальные распределения.

и

$$\theta = \frac{\rho}{\sigma\tau(1-\rho^2)}, \quad \vartheta_1 = \frac{-1}{2\sigma^2(1-\rho^2)}, \quad \vartheta_2 = \frac{-1}{2\tau^2(1-\rho^2)},$$

$$\vartheta_3 = \frac{1}{1-\rho^2} \left(\frac{\xi}{\sigma^2} - \frac{\eta\rho}{\sigma\tau} \right), \quad \vartheta_4 = \frac{1}{1-\rho^2} \left(\frac{\eta}{\tau^2} - \frac{\xi\rho}{\sigma\tau} \right).$$

Гипотеза H : $\rho \leq 0$ и гипотеза $\theta \leq 0$ — эквивалентны. Так как выборочный коэффициент корреляции

$$R = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}$$

не меняется при замене X_i и Y_i на $(X_i - \xi)/\sigma$ и $(Y_i - \eta)/\tau$, то распределение R не зависит от ξ, η, σ и τ , а зависит только от ρ . Следовательно, при $\theta = 0$ оно не зависит от $\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4$. По теореме 2 при $\theta = 0$ R не зависит от (T_1, T_2, T_3, T_4) . Из теоремы 1 выводим, что РНМ несмещенный критерий отклоняет H при

$$R \geq C_0 \quad (71)$$

или, что то же самое, при

$$\frac{R}{\sqrt{(1-R^2)/(n-2)}} > K_0. \quad (72)$$

Статистика R линейна по U , и ее распределение при $\rho = 0$ симметрично относительно 0. Поэтому РНМ несмещенный критерий гипотезы $\rho = 0$ при альтернативе $\rho \neq 0$ отвергает гипотезу, если

$$\frac{|R|}{\sqrt{(1-R^2)/(n-2)}} > K_1. \quad (73)$$

Так как величина $\sqrt{n-2} R / \sqrt{1-R^2}$ имеет при $\rho = 0$ t -распределение с $(n-2)$ степенями свободы (см. задачу 32), то константы K_0 и K_1 находятся из уравнений

$$\int_{K_0}^{\infty} t_{n-2}(y) dy = \alpha \quad \text{и} \quad \int_{K_1}^{\infty} t_{n-2}(y) dy = \frac{\alpha}{2}. \quad (74)$$

Раз распределение R зависит только от ρ , то и функция мощности критерия имеет это свойство.

Рассмотрим теперь задачу без предположения нормальности — в непараметрической постановке. Для любого двумерного распределения (X, Y) обозначим Y_x случайную величину, распределение которой совпадает с условным распределением Y при данном x . Мы скажем, что между X и Y имеется *положительная зависимость*, если при любых $x < x'$ величина $Y_{x'}$ стохастически больше Y_x . Грубо говоря, бóльшие значения Y соответствуют

большим значениям X : в этом интуитивный смысл положительной зависимости. Примером может служить любое двумерное нормальное распределение с $\rho > 0$ (см. задачу 36).

Возьмем теперь общее двумерное распределение, имеющее плотность относительно меры Лебега, и рассмотрим гипотезу независимости при альтернативе положительной зависимости. Если Φ — какой-либо несмещенный критерий, то при независимости X и Y вероятность отклонения гипотезы равна α , т. е.

$$\int \Phi(x_1, \dots, x_n; y_1, \dots, y_n) f_1(x_1) \dots f_1(x_n) f_2(y_1) \dots f_2(y_n) dx dy = \alpha$$

для всех плотностей f_1 и f_2 . По теореме 3 отсюда находим, что

$$\frac{1}{(n!)^2} \sum \Phi(x_{i_1}, \dots, x_{i_n}; y_{j_1}, \dots, y_{j_n}) = \alpha.$$

Здесь суммирование распространяется на все $(n!)^2$ точек множества $S(x, y)$, которое получается из фиксированной точки (x, y) с $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$ всевозможными перестановками x_k и y_k в отдельности.

В классе всех критериев, удовлетворяющих этому условию, наиболее мощный по отношению к нормальным альтернативам (70) с $\rho > 0$ выделяется тем, что отклоняет гипотезу при k' наибольших значениях функции (70) на каждом множестве $S(x, y)$; здесь $k'/(n!)^2 = \alpha$. Так как все суммы: $\sum x_i^2$, $\sum y_i^2$, $\sum x_i$, $\sum y_i$ постоянны на $S(x, y)$, то можно сказать, что критерий отвергает гипотезу для k' наибольших значений $\sum x_i y_i$ в каждом $S(x, y)$.

Из $(n!)^2$ значений, которые статистика $\sum X_i Y_i$ принимает на $S(x, y)$, различных будет только $n!$. В самом деле, значение статистики не изменяется, если величины X и Y подвергаются одной и той же перестановке. Поэтому критерию может быть придана более простая форма: гипотеза отвергается для k наибольших значений суммы $\sum x^{(i)} y_{j_i}$ на каждом множестве $S(x, y)$; здесь предполагается, что $x^{(1)} < \dots < x^{(n)}$ и $k/n! = \alpha$. Можно показать, что этот критерий не смещен по отношению ко всем альтернативам с положительной зависимостью (см. задачу 41 в главе 6).

Сравним критерий перестановок со стандартным нормальным критерием, основанным на выборочном коэффициенте корреляции R . Пусть $T(X, Y)$ обозначает множество упорядоченных X и Y

$$T(X, Y) = (X^{(1)}, \dots, X^{(n)}; Y^{(1)}, \dots, Y^{(n)}).$$

Критическая область для критерия перестановок имеет вид

$$\sum X_i Y_i > C[T(X, Y)]$$

или, что то же, вид

$$R > K [T(X, Y)].$$

Оказывается*), что разность между $K [T(X, Y)]$ и граничной точкой C_0 соответствующего нормального критерия (71) стремится к нулю, и оба критерия становятся эквивалентными при n , стремящемся к бесконечности. Достаточные условия для этого состоят в том, что

$$\sigma_X^2 > 0, \quad \sigma_Y^2 > 0$$

и

$$E(|X|^3) < \infty, \quad E(|Y|^3) < \infty.$$

Следовательно, при больших n стандартный нормальный критерий (71) служит приближением к критерию перестановок, который непрактичен при сколько-нибудь значительном объеме выборки.

12. Задачи

К разделу 2

1. Пусть $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$. Мощность t -критерия Стьюдента является возрастающей функцией от ξ/σ в одностороннем случае

$$H: \xi \leq 0, \quad K: \xi > 0,$$

и от $|\xi|/\sigma$ в двустороннем случае

$$H: \xi = 0, \quad K: \xi \neq 0.$$

[Если

$$S = \sqrt{\frac{1}{n-1} \sum (X_i - \bar{X})^2},$$

то мощность в двустороннем случае равна

$$1 - P \left\{ -\frac{CS}{\sigma} - \frac{\sqrt{n}\xi}{\sigma} \leq \frac{\sqrt{n}(\bar{X} - \xi)}{\sigma} \leq \frac{CS}{\sigma} - \frac{\sqrt{n}\xi}{\sigma} \right\},$$

и желаемый результат вытекает из того факта, что он имеет место условно при каждом фиксированном значении S/σ]

2. В обстановке предыдущей задачи не существует критерия уровня α для проверки $H: \xi = 0$, который при всех σ имел бы мощность $\geq \beta > \alpha$ при альтернативах (ξ, σ) с $\xi = \xi_1 > 0$.

[Пусть $\beta(\xi_1, \sigma)$ — мощность некоторого критерия уровня α для проверки H , и пусть $\beta(\sigma)$ — мощность наиболее мощного критерия для проверки $\xi = 0$ при альтернативе $\xi = \xi_1$, когда σ известно. Тогда $\beta_\sigma(\xi_1, \sigma) \leq \inf_{\sigma} \beta(\sigma) = \alpha$.]

3. (I) Пусть Z и V независимы и распределены по законам $N(0, 1)$ и χ^2_f с f степенями свободы соответственно. Тогда отношение $Z/\sqrt{V/f}$ имеет не-

*) Доказательство см. в книге Фрезера, указанной в сноске в разделе 8.

центральное t -распределение с f степенями свободы и параметром нецентральности δ и плотность этого отношения равна *)

$$p_{\delta}(t) = \frac{1}{2^{\frac{1}{2}(f+1)} \Gamma\left(\frac{1}{2}f\right) \sqrt{\pi f}} \times \\ \times \int_0^{\infty} y^{\frac{1}{2}(f-1)} \exp\left(-\frac{1}{2}y\right) \exp\left[-\frac{1}{2}\left(t\sqrt{\frac{y}{f}} - \delta\right)^2\right] dy \quad (75)$$

или, иначе,

$$p_{\delta}(t) = \frac{1}{2^{\frac{1}{2}(f-1)} \Gamma\left(\frac{1}{2}f\right) \sqrt{\pi f}} \exp\left(-\frac{1}{2}\frac{f\delta^2}{f+t^2}\right) \times \\ \times \left(\frac{f}{f+t^2}\right)^{\frac{1}{2}(f+1)} \int_0^{\infty} v^f \exp\left[-\frac{1}{2}\left(v - \frac{\delta t}{\sqrt{f+t^2}}\right)^2\right] dv. \quad (76)$$

Другая форма получается подстановкой $w = t\sqrt{y}/\sqrt{f}$ в выражении (75)

(II) Если $X_1, \dots, X_n$ независимы и распределены по закону $N(\xi, \sigma^2)$, то $\sqrt{n}\bar{X}/\sqrt{\sum (X_i - \bar{X})^2/(n-1)}$ имеет нецентральное t -распределение с $(n-1)$ -й степенью свободы и параметром нецентральности $\delta = \sqrt{n}\xi/\sigma$.

[(I) Первое выражение получается из совместной плотности Z и V переходом к переменным $t = z/\sqrt{v/f}$ и v .]

4. Пусть $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$. Обозначим $\beta(\xi/\sigma)$ мощность одностороннего t -критерия для $H: \xi \leq 0$ при альтернативе ξ/σ , а $\beta^*(\xi/\sigma)$ — мощность соответствующего критерия при известном σ . Определить $\beta(\xi/\sigma)$ для $n = 5, 10, 15$, $\alpha = 0,05$, $\xi/\sigma = 0,7; 0,8; 0,9; 1,0; 1,1; 1,2$ и в каждом случае сравнить с $\beta^*(\xi/\sigma)$. Сделайте то же самое для двустороннего случая.

5. Пусть $Z_1, \dots, Z_n$ независимы и распределены нормально с общей дисперсией σ^2 и средними

$$E(Z_i) = \xi_i \quad (i = 1, \dots, s), \\ E(Z_i) = 0 \quad (i = s+1, \dots, n).$$

Существуют РНМ несмещенные критерии для проверки $\xi_1 \leq \xi_1^0$ и $\xi_1 = \xi_1^0$. Их критические области

$$\frac{Z_1 - \xi_1^0}{\sqrt{\sum_{i=s+1}^n Z_i^2/(n-s)}} > C_0$$

*) Функция распределения, так же как и плотность вероятности этого распределения, табулирована у Resnikoff and Lieberman, Tables of the Non-central t -distribution, Stanford Univ. Press, 1957. См. также: Mergington and Pearson, An approximation to the distribution of non-central t , Biometrika 45 (1958), 484—491. (См. другие ссылки в книге: Болышев Л. Н., Смирнов Н. В., Таблицы математической статистики, М., 1965, стр. 98. — Прим. ред.)

и

$$\frac{|Z_1 - \xi_1^0|}{\sqrt{\sum_{i=s+1}^n Z_i^2 / (n-s)}} > C.$$

При $\xi_1 = \xi_1^0$ статистика имеет t -распределение с $n-s$ степенями свободы.

6. Пусть $X_1, \dots, X_n$ независимы и распределены нормально с общей дисперсией σ^2 и средними $\xi_1, \dots, \xi_n$. Пусть $Z_i = \sum_{j=1}^n a_{ij} X_j$ — ортогональное преобразование (т. е. $\sum_{i=1}^n a_{ij} a_{ik} = 1$ или 0 при $j = k$ или $j \neq k$). Тогда величины Z распределены нормально с общей дисперсией σ^2 и средними $\xi_i = \sum a_{ij} \xi_j$.

[Плотность Z получается из плотности X подстановкой $x_i = \sum b_{ij} z_j$, где (b_{ij}) — матрица, обратная к (a_{ij}) и умножением на якобиан, который здесь равен 1.]

7. Если $X_1, \dots, X_n$ — выборка из $N(\xi, \sigma^2)$, то РНМ несмещенный критерий для $\xi \leq 0$ и $\xi = 0$ может быть получен из результатов задач 5 и 6 ортогональным преобразованием X в Z , где $Z_1 = \sqrt{n} \bar{X}$.

[Тогда

$$\sum_{i=2}^n Z_i^2 = \sum_{i=1}^n Z_i^2 - Z_1^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2 = \sum_{i=1}^n (X_i - \bar{X})^2.]$$

8. Пусть $X_1, X_2, \dots$ — последовательность независимых случайных величин, распределенных по закону $N(\xi, \sigma^2)$, и пусть

$$Y_n = [nX_{n+1} - (X_1 + \dots + X_n)] / \sqrt{n(n+1)}.$$

(I) Величины $Y_1, Y_2, \dots$ независимы и распределены по закону $N(0, \sigma^2)$.

(II) С помощью величин Y гипотеза $\sigma = \sigma_0$ при альтернативе $\sigma = \sigma_1$ может быть проверена последовательным критерием отношений вероятностей.

К разделу 3

9. Пусть $X_1, \dots, X_n$ и $Y_1, \dots, Y_n$ — независимые выборки из $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ соответственно. Определите объем выборки, необходимый для того, чтобы получить мощность $\geq \beta$ при альтернативах $\tau/\sigma > \Delta$, если $\alpha = 0,05$, $\beta = 0,9$, $\Delta = 1,5$; 2; 3 и проверяемая гипотеза есть $H: \Delta \leq 1$.

10. При $m = n$ область принятия (23) может быть записана в форме

$$\max (S_Y^2 / \Delta_0 S_X^2, \Delta_0 S_X^2 / S_Y^2) \leq (1 - C) / C,$$

где $S_X^2 = \sum (X_i - \bar{X})^2$, $S_Y^2 = \sum (Y_i - \bar{Y})^2$ и где C определяется из

$$\int_0^C B_{n-1, n-1}(w) dw = \frac{\alpha}{2}.$$

11. Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$. РНМ несмещенный критерий для гипотезы $\eta - \xi = 0$ может быть получен

из результатов задач 5 и 6 ортогональным преобразованием $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ в $(Z_1, \dots, Z_{m+n})$ таким, что

$$Z_1 = (\bar{Y} - \bar{X}) / \sqrt{(1/m) + (1/n)}, \quad Z_2 = (\sum X_i + \sum Y_j) / \sqrt{m+n}.$$

12. *Экспоненциальные плотности.* Пусть $X_1, \dots, X_n$ — выборка из распределения с экспоненциальной плотностью $a^{-1}e^{-(x-b)/a}$ при $x \geq b$.

(I) Для проверки $a = 1$ существует РНМ несмещенный критерий с областью принятия

$$C_1 \leq 2 \sum [x_i - \min(x_1, \dots, x_n)] \leq C_2.$$

При $a = 1$ соответствующая статистика имеет χ^2 -распределение с $2n - 2$ степенями свободы, а C_1 и C_2 определяются из уравнений

$$\int_{C_1}^{C_2} \chi_{2n-2}^2(y) dy = \int_{C_1}^{C_2} \chi_{2n}^2(y) dy = 1 - \alpha.$$

(II) Для проверки $b = 0$ существует РНМ несмещенный критерий с областью принятия

$$0 \leq \frac{n \cdot \min(x_1, \dots, x_n)}{\sum [x_i - \min(x_1, \dots, x_n)]} \leq C.$$

При $b = 0$ соответствующая статистика имеет плотность

$$p(u) = \frac{(n-1)}{(1+u)^n}, \quad u \geq 0.$$

[Эти распределения при меняющемся b не образуют экспоненциального семейства, и поэтому теорема 3 непосредственно неприменима.]

(I) Можно ограничиться упорядоченными наблюдениями $X^{(1)} < \dots < X^{(n)}$, поскольку они достаточны для a и b , и перейти к новым переменным $Z_1 = nX^{(1)}$, $Z_i = (n-i+1)[X^{(i)} - X^{(i-1)}]$, $i = 2, \dots, n$ (так же, как в задаче 13 главы 2). При $a = 1$ Z_1 является полной достаточной статистикой для b , и критерий находится путем решения условной задачи при данном z_1 .

Так как $\sum_{i=2}^n Z_i$ не зависит от Z_1 , то условный РНМ несмещенный критерий

имеет при каждом z_1 область принятия $C_1 \leq \sum_{i=2}^n Z_i \leq C_2$, что приводит к нужному результату.

(II) При $b = 0$ $\sum_{i=1}^n Z_i$ является полной достаточной статистикой для a

и критерий находится решением условной задачи при данной $\sum_{i=1}^n z_i$. Затем

используется тот факт, что при $b = 0$ $Z_1 / \sum_{i=1}^n Z_i$ не зависит от $\sum_{i=1}^n Z_i$,

и рассуждения завершаются, как в доказательстве теоремы 1.]

13. Распространите результаты предыдущей задачи на случай, рассмотренный в задаче 8 главы 3, когда наблюдения продолжаются только до тех пор, пока не появятся $X^{(1)}, \dots, X^{(r)}$.

К разделу 4

14. На основе выборки $X = (X_1, \dots, X_n)$ фиксированного объема из $N(\xi, \sigma^2)$ нельзя построить доверительные интервалы для ξ с положительным коэффициентом доверия и ограниченной длины.

[Рассмотрите любое семейство доверительных интервалов $\delta(X) \pm L/2$ постоянной длины L . Пусть $\xi_1, \dots, \xi_{2N}$ таковы, что $|\xi_i - \xi_j| < L$ при $i \neq j$. Тогда множества $S_i = \{x: |\delta(x) - \xi_i| \leq L/2\}$ ($i = 1, \dots, 2N$) попарно не пересекаются. Существует $\sigma_0 > 0$ такое, что

$$|P_{\xi_i, \sigma} \{X \in S_i\} - P_{\xi_1, \sigma} \{X \in S_1\}| \leq 1/2N \text{ при } \sigma > \sigma_0,$$

что можно увидеть, переходя к новым переменным $Y_i = (X_i - \xi_i)/\sigma$ и применяя леммы 2 и 4 Дополнения. Так как $\min_i P_{\xi_i, \sigma} \{X \in S_i\} \leq 1/2N$, то при $\sigma > \sigma_0$ $\min_i P_{\xi_i, \sigma} \{X \in S_i\} \leq 1/N$ и, следовательно,

$$\inf_{\xi, \sigma} P_{\xi, \sigma} \left\{ |\delta(X) - \xi| \leq \frac{L}{2} \right\} \leq \frac{1}{N}.$$

Доверительный коэффициент, связанный с системой интервалов $\delta(X) \pm L/2$, равен, следовательно, нулю. Тем более это должно быть верно для любого множества доверительных интервалов длины $\leq L$.

15. *Двухступенная процедура, предложенная Стейном.*

(I) Если mS^2/σ^2 имеет χ^2 -распределение с m степенями свободы и если условное распределение Y при данном $S = s$ равно $N(0, \sigma^2/S^2)$, то Y имеет t -распределение Стьюдента с m степенями свободы.

(II) Пусть $X_1, X_2, \dots$ независимы и распределены по закону $N(\xi, \sigma^2)$.

Пусть $\bar{X}_0 = \sum_{i=1}^{n_0} X_i/n_0$, $S^2 = \sum_{i=1}^{n_0} (X_i - \bar{X}_0)^2/(n_0 - 1)$, и пусть $a_1 = \dots = a_{n_0} = a$, $a_{n_0+1} = \dots = a_n = b$ и $n \geq n_0$ — измеримые функции от S . Тогда

$$Y = \frac{\sum_{i=1}^n a_i (X_i - \xi)}{\sqrt{S^2 \sum_{i=1}^n a_i^2}}$$

имеет распределение Стьюдента с $n_0 - 1$ степенями свободы.

(III) Рассмотрим двухступенную схему выборки Π_1 , в которой S^2 вычисляется, исходя из первоначальной выборки объема n_0 , а затем производится $n - n_0$ дополнительных наблюдений. Объем второй выборки таков, что

$$n = \max \left\{ n_0 + 1, \left[\frac{S^2}{c} \right] + 1 \right\},$$

где c — заданная константа и $[y]$ обозначает наибольшее целое $\leq y$. Тогда существуют числа $a_1, \dots, a_n$ такие, что

$$a_1 = \dots = a_{n_0}, \quad a_{n_0+1} = \dots = a_n, \quad \sum_{i=1}^n a_i = 1, \quad \sum_{i=1}^n a_i^2 = c/S^2.$$

Из (II) следует, что $\sum_{i=1}^n a_i (X_i - \xi)/\sqrt{c}$ имеет t -распределение Стьюдента с $n_0 - 1$ степенями свободы.

(IV) Ниже описывается схема выборки Π_2 , при которой вторая выборка может быть и пустой, т. е. не содержащей ни одного элемента. По отношению

к применениям, которые будут сделаны в задачах 16 и 17, P_2 несколько более эффективна, чем P_1 . Пусть n_0 , S^2 и c определены, как раньше, пусть

$$n = \max \left\{ n_0, \left[\frac{S^2}{c} \right] + 1 \right\},$$

$a_i = 1/n$ ($i = 1, \dots, n$) и $\bar{X} = \sum_{i=1}^n a_i X_i$. Тогда $\sqrt{n} (\bar{X} - \xi)/S$ снова имеет t -распределение с $n_0 - 1$ степенью свободы.

[(II) При данном $S = s$ величины a , b и n постоянны,

$$\sum_{i=1}^{n_0} a_i (X_i - \xi) = n_0 a (\bar{X}_0 - \xi)$$

распределена по закону $N(0, n_0 \cdot a^2 \cdot \sigma^2)$ и, следовательно, числитель Y распределен нормально с нулевым средним и дисперсией $\sigma^2 \sum_{i=1}^n a_i^2$. Теперь требуемый результат получается из (I).]

16. *Доверительные интервалы фиксированной длины для среднего значения нормального распределения.* (I) Рассмотрим двухступенную процедуру P_1 , определенную в п. (III) предыдущей задачи. Допустим, что при данных $L > 0$ и $0 < \gamma < 1$ число c находится из уравнения

$$\int_{-L/2\sqrt{c}}^{L/2\sqrt{c}} t_{n_0-1}(y) dy = \gamma,$$

где t_{n_0-1} обозначает плотность t -распределения с $(n_0 - 1)$ степенью свободы.

Тогда интервалы $\sum_{i=1}^n a_i X_i \pm L/2$ являются доверительными интервалами для ξ длины L и с коэффициентом доверия γ .

(II) Пусть c определено как в п. (I), а выборочная процедура P_2 такова же, как в п. (IV) предыдущей задачи. Интервалы $\bar{X} \pm L/2$ являются доверительными интервалами длины L для ξ с коэффициентом доверия $\geq \gamma$ и в то же время число необходимых наблюдений несколько меньше, чем в P_1 .

[(I) Вероятность того, что интервал покрывает ξ , равно

$$P_{\xi, \sigma} \left\{ -\frac{L}{2\sqrt{c}} \leq \frac{\sum_{i=1}^n a_i (X_i - \xi)}{\sqrt{c}} \leq \frac{L}{2\sqrt{c}} \right\} = \gamma.$$

(II) Вероятность того, что интервал покрывает ξ , равна

$$P_{\xi, \sigma} \left\{ \frac{\sqrt{n} |\bar{X} - \xi|}{S} \leq \frac{\sqrt{n} L}{2S} \right\} \geq P_{\xi, \sigma} \left\{ \frac{\sqrt{n} |\bar{X} - \xi|}{S} \leq \frac{L}{2\sqrt{c}} \right\} = \gamma.]$$

17. *Двухступенные t -критерии с мощностью, не зависящей от σ .* (I) Возьмем процедуру P_1 с некоторым c и определим C посредством равенства

$$\int_C^\infty t_{n_0-1}(y) dy = \alpha.$$

Тогда критическая область $\left(\sum_{i=1}^n a_i X_i - \xi_0\right)/\sqrt{c} > C$ определяет критерий уровня α для $H: \xi \leq \xi_0$ со строго возрастающей функцией мощности $\beta_c(\xi)$, зависящей только от ξ .

(II) При любой альтернативе ξ_1 и любых $\alpha < \beta < 1$ число c может быть выбрано так, что $\beta_c(\xi_1) = \beta$.

(III) Критерий с критической областью $\sqrt{n}(\bar{X} - \xi_0)/S > C$, основанный на P_2 и с тем же значением c , что и в п. (I), является критерием уровня α для H , причем равномерно более мощным, чем критерий, данный в (I).

(IV) Распространите утверждения пп. (I)–(III) на случай проверки гипотезы $\xi = \xi_0$ при альтернативе $\xi \neq \xi_0$.

[(I) (II) Мощность критерия равна

$$\beta_c(\xi) = \int_{C - \frac{\xi - \xi_0}{\sqrt{c}}}^{\infty} t_{n_0-1}(y) dy.$$

(III) Это следует из неравенства $\sqrt{n}|\xi - \xi_0|/S \geq |\xi - \xi_0|/\sqrt{c}$.]

К разделу 5

18. Пусть $X_1, \dots, X_n$ распределены, как в задаче 12. Тогда наиболее точные несмещенные доверительные интервалы для масштабного параметра a суть

$$\frac{2}{C_2} \sum [x_i - \min(x_1, \dots, x_n)] \leq a \leq \frac{2}{C_1} \sum [x_i - \min(x_1, \dots, x_n)].$$

19. В нижеследующих ситуациях существуют наиболее точные несмещенные доверительные интервалы:

(I) для параметра $p_1 q_2 / p_2 q_1$, если X и Y независимы и имеют биномиальные распределения $b(p_1, m)$ и $b(p_2, n)$.

(II) для параметра Δ раздела 6 главы 4 — в 2×2 таблице.

К разделу 6

20. (I) В предположениях, сделанных в начале раздела 6, РНМ критерий для $H: \rho = \rho_0$ дается формулой (45).

(II) Пусть $(\underline{\rho}, \bar{\rho})$ — соответствующие наиболее точные несмещенные доверительные интервалы для $\rho = a\gamma + b\delta$, где $\underline{\rho} = \underline{\rho}(a, b)$, $\bar{\rho} = \bar{\rho}(a, b)$. Тогда, если f_1 и f_2 — возрастающие функции, то математическое ожидание от $f_1(|\bar{\rho} - \rho|) + f_2(|\underline{\rho} - \rho|)$ является возрастающей функцией от $a^2/n + b^2$.

[(I) Произведите любое ортогональное преобразование $y_1, \dots, y_n$ в новые переменные $z_1, \dots, z_n$ такие, что

$$z_1 = \sum_i [bv_i + (a/n)] y_i / \sqrt{(a^2/n) + b^2}, \quad z_2 = \sum_i (av_i - b) \cdot y_i / \sqrt{a^2 + nb^2},$$

и затем примените результаты задач 5 и 6.

(II) Если $a_1^2/n + b_1^2 < a_2^2/n + b_2^2$, то случайная величина $|\bar{\rho}(a_2, b_2) - \rho|$ стохастически больше, чем $|\bar{\rho}(a_1, b_1) - \rho|$. Аналогично и для $\underline{\rho}$.]

К разделу 7

21. Докажите теорему 3 для произвольных значений c .

К разделу 8

22. Пусть $c = 1$, $m = n = 4$, $\alpha = 0,1$ и упорядоченные по возрастанию координаты $z^{(1)}, \dots, z^{(N)}$ точки z суть 1,97; 2,19; 2,61; 2,79; 2,88; 3,02; 3,28; 3,41. Определите точки множества $S(z)$, входящие в критическую область (54).

23. Доверительные интервалы для сдвига. Пусть $X_1, \dots, X_m; Y_1, \dots, Y_n$ независимы и имеют непрерывные функции распределения $F(x)$ и $G(y) = F(y - \Delta)$ соответственно. Из критерия перестановок для гипотезы $H(\Delta_0)$: $\Delta = \Delta_0$, без всяких дальнейших предположений относительно F , можно получить доверительные интервалы для Δ . Именно, возьмем точку $(z_1, \dots, z_{m+n}) = (x_1, \dots, x_m, y_1 - \Delta, \dots, y_n - \Delta)$ и $\binom{m+n}{m}$ перестановок $i_1 < \dots < i_m; i_{m+1} < \dots < i_{m+n}$ целых чисел $1, \dots, m+n$. Допустим, что гипотеза $H(\Delta)$ принимается для k из этих перестановок, которые приводят к наименьшим значениям

$$\left| \sum_{j=m+1}^{m+n} \frac{z_{i_j}}{n} - \sum_{j=1}^m \frac{z_{i_j}}{m} \right|,$$

причем $k = (1 - \alpha) \binom{m+n}{m}$. Тогда совокупность значений Δ , для которых $H(\Delta)$ принимается, образует интервал, и эти интервалы являются доверительными интервалами для Δ с доверительным уровнем $1 - \alpha$.

[Некоторая точка принадлежит критической области для $H(\Delta)$, если величину

$$\left| \frac{\sum (y_i - \Delta)}{n} - \frac{\sum x_i}{m} \right| = |\bar{y} - \bar{x} - \Delta|$$

превосходят по меньшей мере $\binom{m+n}{m} - k$ из величин $|\bar{y}' - \bar{x}' - \gamma\Delta|$, где $(x'_1, \dots, x'_m, y'_1, \dots, y'_n)$ получается перестановкой из $(x_1, \dots, x_m, y_1, \dots, y_n)$, величина γ определяется этой перестановкой и $|\gamma| \leq 1$. Требуемый результат вытекает теперь из следующего замечания. Если

$$(\bar{y} - \bar{x} - \Delta)^2 \leq (\bar{y}' - \bar{x}' - \gamma\Delta)^2$$

или, более общим образом, если $(a - \Delta)^2 \leq (b - \gamma\Delta)^2$ при некоторых a, b и $\Delta = \Delta_0, \Delta = \Delta_1$, то это же неравенство выполняется для любого Δ между Δ_0 и Δ_1]

К разделу 9

24. Рассмотрим эксперимент со сходными парами для проверки эффекта обработки. Допустим, что наблюдаются только разности $Z_i = Y_i - X_i$. Предположим, что величины Z образуют выборку из неизвестного непрерывного распределения, которое при гипотезе отсутствия эффекта обработки является симметричным относительно начала координат. При альтернативных гипотезах оно симметрично относительно точки $\xi > 0$. Найдите критерий, который в классе всех несмещенных критериев максимизирует мощность при альтернативах: величины Z образуют выборку из $N(\xi, \sigma^2)$ с $\xi > 0$.

[При указанной гипотезе совокупность статистик $\left(\sum_{i=1}^n Z_i^2, \dots, \sum_{i=1}^n Z_i^{2n} \right)$ будет достаточной; так же, как в теореме 3, можно показать, что она является полной. Дальнейшие рассуждения следуют плану раздела 8.]

25. (I) Если $X_1, \dots, X_n; Y_1, \dots, Y_n$ — независимые нормально распределенные величины с общей дисперсией σ^2 и средними $E(X_i) = \xi_i, E(Y_i) = \xi_i + \Delta$, то РНМ несмещенный критерий для $\Delta = 0$ при альтернативе $\Delta > 0$ дается формулой (59).

(II) Найдите наиболее точные несмещенные доверительные интервалы для Δ .

[(I) Структура задачи становится ясной, если произвести ортогональное преобразование $X'_i = (Y_i - X_i)/\sqrt{2}, Y' = (X_i + Y_i)/\sqrt{2}$.]

26. *Сравнение двух планов эксперимента.* В предположениях, сделанных в начале раздела 9, можно провести следующее сравнение метода полной рандомизации и метода сходных пар. Эффекты отдельных единиц и экспериментальные эффекты: U_i и V_i независимы и распределены нормально с дисперсиями σ_i^2, σ^2 и средними $E(U_i) = \mu$ и $E(V_i)$, равным ξ или η (в зависимости от того, соответствует ли V_i контрольной или обрабатываемой единице). При полной рандомизации результатами наблюдений будет $X_i = U_i + V_i$ ($i = 1, \dots, n$) у контрольных единиц и $Y_i = U_{n+i} + V_{n+i}$ ($i = 1, \dots, n$) у обрабатываемых единиц, причем $E(X_i) = \mu + \xi, E(Y_i) = \mu + \eta$. В методе сходных пар, если сходство предположить идеальным, величины X будут такими же, как и раньше, но $Y_i = U_i + V_{n+i}$. РНМ несмещенный критерий определяется формулой (27) в случае полной рандомизации и формулой (59) в методе сходных пар. Распределение соответствующей статистики при альтернативе $\Delta = \eta - \xi$ будет нецентральным t -распределением с параметром нецентральности $\sqrt{n} \Delta / \sqrt{2(\sigma^2 + \sigma_1^2)}$ и $(2n - 2)$ степенями свободы в первом случае, с параметром нецентральности $\sqrt{n} \Delta / \sqrt{2} \sigma$ и $(n - 1)$ степенью свободы во втором. Таким образом, недостатком метода сходных пар является меньшее число степеней свободы, а преимущество — больший параметр нецентральности. Для $\alpha = 0,05$ и $\Delta = 2$ сравните мощность двух методов как функцию от n при $\sigma_1 = 1, \sigma = 2$ и при $\sigma_1 = 2, \sigma = 1$.

27. *Продолжение.* Другой способ сравнения двух планов эксперимента — изучение средней длины наиболее точных несмещенных доверительных интервалов для $\Delta = \eta - \xi$ в каждом случае. Проведите это сравнение при меняющемся n и коэффициенте доверия $1 - \alpha = 0,95$ при $\sigma_1 = 1, \sigma = 2$ и при $\sigma_1 = 2, \sigma = 1$.

К разделу 10

28. Пусть критическая функция удовлетворяет (65), но не (67), и пусть $\alpha < 1/2$. Нижеследующая конструкция приводит к измеримой критической функции ϕ такой, что для нее выполняется (67) и $\phi_0(z) \leq \phi(z)$ при всех z . Определим последовательности функций $\phi_1, \phi_2, \dots$ и $\psi_0, \psi_1, \dots$ по индукции с помощью соотношений

$$\psi_m(z) = \sum_{z' \in S(z)} \phi_m(z') / N_1! \dots N_c!, \quad m = 0, 1, \dots$$

и

$$\phi_m(z) = \begin{cases} \phi_{m-1}(z) + [\alpha - \psi_{m-1}(z)], & \text{если как } \phi_{m-1}(z), \text{ так и } \psi_{m-1}(z) \\ \phi_{m-1}(z) & \text{меньше } \alpha, \text{ в других случаях.} \end{cases}$$

Функция $\phi_m(z) = \lim \phi_m(z)$ является школой.

[Функции ϕ_m образуют неубывающую последовательность и лежат между 0 и 1. Далее, как можно показать по индукции, $0 \geq \alpha - \phi_m(z) \leq (1 - \gamma)^m [\alpha - \psi_0(z)]$, где $\gamma = 1/N_1! \dots N_c!]$

29. Рассмотрим проблему проверки $H: \eta = \xi$ в семействе плотностей (62), когда известно, что $\sigma > c > 0$ и что точка $(\xi_{11}, \dots, \xi_{cNc})$ (см. (63)) лежит в ограниченной области R , содержащей прямоугольник; при этом c и R считаются известными. Теорема 4 уже неприменима. Однако несмещенность некоторого критерия Φ для проверки H влечет (67) и тем самым редуцирует проблему к критериям перестановок.

[Несмещенность влечет $\int \Phi(z) p_{\sigma, \xi}(z) dz = \alpha$ и, следовательно,

$$\alpha = \int \Phi(z) p_{\sigma, \xi}(z) dz = \int \Phi(z) \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp \left[-\frac{1}{2\sigma^2} \sum \sum (z_{ij} - \xi_{ij})^2 \right]$$

при всех $\sigma > c$ и ξ из R . Желаемый результат вытекает из полноты последнего семейства]

30. Обобщим теорему 4 на другие планы эксперимента. Пусть $Z = (Z_1, \dots, Z_N)$ и пусть $G = \{g_1, \dots, g_r\}$ — группа перестановок N координат, или, более общим образом, группа ортогональных преобразований N -мерного пространства. Если

$$p_{\sigma, \xi}(z) = \frac{1}{r} \sum_{k=1}^r \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp \left(-\frac{1}{2\sigma^2} |z - g_k \xi|^2 \right), \quad (77)$$

где $|z|^2 = \sum z_i^2$, то выполнение при всех $\sigma > 0$ и всех ξ неравенства $\int \Phi(z) p_{\sigma, \xi}(z) dz \leq \alpha$ влечет

$$\frac{1}{r} \sum_{z' \in S(z)} \Phi(z') \leq \alpha \text{ почти всюду,} \quad (78)$$

где $S(z)$ — множество точек N -мерного пространства, получаемых из точки z применением к ней всех преобразований g_k , $k = 1, \dots, r$.

31. Обобщение следствия 2. Пусть H — класс плотностей (77) с $\sigma > 0$ и $-\infty < \xi_i < \infty$ ($i = 1, \dots, N$). Полное семейство критериев для H с уровнем значимости α образуют все критерии перестановок, для которых

$$\frac{1}{r} \sum_{z' \in S(z)} \Phi(z') = \alpha \text{ почти всюду.} \quad (79)$$

К разделу 11

32. (I) Если совместное распределение X и Y представляет собой двумерное нормальное распределение (70), то условное распределение Y при данном x также нормально с дисперсией $\tau^2(1 - \rho^2)$ и средним $\eta + (\rho\tau/\sigma)(x - \xi)$.

(II) Пусть $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения, и пусть R — выборочный коэффициент корреляции и $\rho = 0$. Тогда условное распределение $\sqrt{n-2}R/\sqrt{1-R^2}$ при данных $x_1, \dots, x_n$ является t -распределением Стьюдента с $n-2$ степенями свободы (в предположении, что $\sum (x_i - \bar{x})^2 > 0$). Таково же, следовательно, и безусловное распределение этой статистики.

(III) Плотность вероятности для R равна

$$p(r) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left[\frac{1}{2}(n-1)\right]}{\Gamma\left[\frac{1}{2}(n-2)\right]} (1-r^2)^{\frac{1}{2}n-2}. \quad (80)$$

[(II) Если $v_i = (x_i - \bar{x}) / \sqrt{(\sum x_j - \bar{x})^2}$ так, что $\sum v_i = 0$ и $\sum v_i^2 = 1$, то исследуемую статистику можно представить в форме

$$\frac{\sum v_i Y_i}{\sqrt{\sum [Y_i^2 - n\bar{Y}^2 - (\sum v_i Y_i)^2 / (n-2)]}}.$$

Так как ее распределение зависит только от ρ , то мы можем принять $\eta = 0$, $\tau = 1$. Желаемый результат получается из задачи 6, если применить ортогональное преобразование $(Y_1, \dots, Y_n)$ в $(Z_1, \dots, Z_n)$ такое, что $Z_1 = \sqrt{n} \bar{Y}$, $Z_2 = \sum v_i Y_i$.]

33. (I) Пусть $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения (70). Пусть

$$S_1^2 = \sum (X_i - \bar{X})^2, \quad S_2^2 = \sum (Y_i - \bar{Y})^2, \quad S_{12} = \sum (X_i - \bar{X})(Y_i - \bar{Y}).$$

Тогда существует РНМ несмещенный критерий для проверки гипотезы $\tau/\sigma = \Delta$. Соответствующая область принятия:

$$\frac{|\Delta^2 S_1^2 - S_2^2|}{\sqrt{(\Delta^2 S_1^2 + S_2^2)^2 - 4\Delta^2 S_{12}^2}} \leq C,$$

а плотность соответствующей статистики при истинности гипотезы определяется по (80).

(II) В предположении $\tau = \sigma$ существует РНМ несмещенный критерий для проверки $\eta = \xi$ с областью принятия

$$|\bar{Y} - \bar{X}| / \sqrt{S_1^2 + S_2^2 - S_{12}} \leq C.$$

При умножении на надлежащую постоянную соответствующая статистика при $\xi = \eta$ имеет t -распределение Стьюдента с $n-1$ степенью свободы (без предположения $\tau = \sigma$ эта гипотеза является специальным случаем гипотезы, рассмотренной в примере 11 главы 7).

[(I) Преобразование $U = \Delta X + Y$, $V = X - (1/\Delta)Y$ сводит задачу к проверке гипотезы о равенстве нулю коэффициента корреляции в двумерном нормальном распределении.

(II) Перейдите к новым переменным $V_i = Y_i - X_i$, $U_i = Y_i + X_i$.]

34. Пусть $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения (70) и пусть

$$S_1^2 = \sum (X_i - \bar{X})^2, S_{12} = \sum (X_i - \bar{X})(Y_i - \bar{Y}), \quad S_2^2 = \sum (Y_i - \bar{Y})^2.$$

(I) Тогда статистики (S_1^2, S_{12}, S_2^2) и $(\bar{X}, \bar{Y})$ независимы и совместное распределение у первых такое же, как у $\left(\sum_{i=1}^{n-1} X_i'^2, \sum_{i=1}^{n-1} X_i' Y_i', \sum_{i=1}^{n-1} Y_i'^2 \right)$, где (X_i', Y_i') , $i = 1, \dots, n-1$ — выборка из нормального распределения (70) с $\xi = \eta = 0$.

(II) Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_m$ — две выборки из $N(0, 1)$. Тогда совместная плотность для $S_1^2 = \sum X_i^2$, $S_{12} = \sum X_i Y_i$, $S_2^2 = \sum Y_i^2$ равна

$$\frac{1}{4\pi\Gamma(m-1)} (s_1^2 s_2^2 - s_{12}^2)^{1/2(m-3)} \exp \left[-\frac{1}{2} (s_1^2 + s_2^2) \right]$$

при $s_{12}^2 \leq s_1^2 s_2^2$ и нулю в других случаях.

(III) Совместная плотность статистик (S_1^2, S_{12}, S_2^2) п. (I) равна

$$\frac{(s_1^2 s_2^2 - s_{12}^2)^{1/2} (n-4)}{4\pi \Gamma(n-2) (\sigma \tau \sqrt{1-\rho^2})^{n-1}} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{s_1^2}{\sigma^2} - \frac{2\rho s_{12}}{\sigma\tau} + \frac{s_2^2}{\tau^2} \right) \right] \quad (81)$$

при $s_{12}^2 \leq s_1^2 s_2^2$ и нулю в других случаях.

[(I) Произведите ортогональное преобразование $X_1, \dots, X_n$ в $X'_1, \dots, X'_n$ такое, что $X'_n = \sqrt{n} \bar{X}$, и примените это же ортогональное преобразование к $Y_1, \dots, Y_n$. Тогда

$$Y'_n = \sqrt{n} \bar{Y},$$

$$\sum_{i=1}^{n-1} X_i'^2 = \sum_{i=1}^n (X_i - \bar{X})^2, \quad \sum_{i=1}^{n-1} X'_i Y'_i = \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}),$$

$$\sum_{i=1}^{n-1} Y_i'^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

Пары случайных величин $(X'_1, Y'_1), \dots, (X'_n, Y'_n)$ независимы и имеют каждая двумерное нормальное распределение с теми же дисперсиями и корреляцией, что и (X, Y) и со средними $E(X'_i) = E(Y'_i) = 0$ при $i = 1, \dots, n-1$.

(II) Рассмотрим сначала совместное распределение $S_{12} = \sum x_i Y_i$ и $S_2^2 = \sum Y_i^2$ при данных $x_1, \dots, x_m$. Положим $Z_1 = S_{12} / \sqrt{\sum x_i^2}$ и произведем ортогональное преобразование $Y_1, \dots, Y_m$ в $Z_1, \dots, Z_m$. Тогда $S_2^2 = \sum_{i=1}^m Z_i^2$.

и случайные величины Z_1 и $\sum_{i=2}^m Z_i^2 = S_2^2 - Z_1^2$ будут независимы и распределены по законам $N(0, 1)$ и χ_{m-1}^2 соответственно. Отсюда совместная условная плотность величин $S_{12} = s Z_1$ и S_2^2 получается простой заменой переменных. Так как это условное распределение зависит от x только через s_1^2 , то совместная плотность S_1^2, S_{12}, S_2^2 находится умножением указанной условной плотности на частную плотность величины S_1^2 , которая равна χ_m^2 . Доказательство завершается применением тождества

$$\Gamma \left[\frac{1}{2} (m-1) \right] \Gamma \left(\frac{1}{2} m \right) = \frac{\sqrt{\pi} \Gamma(m-1)}{2^{m-2}}.$$

(III) Если $(X', Y') = (X'_1, Y'_1, \dots, X'_m, Y'_m)$ — выборка из двумерного нормального распределения с $\xi = \eta = 0$, то статистика $T = \left(\sum X_i'^2, \sum X_i' Y_i', \sum Y_i'^2 \right)$ достаточна для $\theta = (\sigma, \rho, \tau)$ и плотность T находится по данной в п. (II) для $\theta_0 = (1, 0, 1)$ с помощью тождества (см. задачу (10 (I) в главе 3)

$$p_\theta^T(t) = p_{\theta_0}^T(t) \left[\frac{p_{\theta_0}^{X', Y'}(x', y')}{p_{\theta_0}^{X', Y'}(x', y')} \right].$$

Требуемый результат получается теперь из п. (I) при $m = n-1$.

35. Если $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения, то плотность вероятности выборочного коэффициента корреляции R равна *)

$$\rho_R(r) = \frac{2^{n-3}}{\pi (n-3)!} (1-\rho^2)^{1/2(n-1)} (1-r^2)^{1/2(n-4)} \sum_{k=0}^{\infty} \Gamma^2\left[\frac{1}{2}(n+k-1)\right] \frac{(2\rho r)^k}{k!} \quad (82)$$

или, что то же самое,

$$\rho_R(r) = \frac{n-2}{\pi} (1-\rho^2)^{1/2(n-1)} (1-r^2)^{1/2(n-4)} \int_0^1 \frac{t^{n-2}}{(1-\rho r t)^{n-1}} \frac{1}{\sqrt{1-t^2}} dt. \quad (83)$$

Другая форма получается при замене переменных $t = (1-v)/(1-\rho r v)$ в интеграле правой части (83). Интеграл принимает вид

$$\frac{1}{(1-\rho r)^{1/2(2n-3)}} \int_0^1 \frac{(1-v)^{n-2}}{\sqrt{2v}} \left[1 - \frac{1}{2} v(1+\rho r)\right]^{-1/2} dv. \quad (84)$$

Разлагая последний множитель по степеням v , получаем для плотности выражение

$$\begin{aligned} \frac{n-2}{\sqrt{2\pi}} \frac{\Gamma(n-1)}{\Gamma\left(n-\frac{1}{2}\right)} (1-\rho^2)^{1/2(n-1)} (1-r^2)^{1/2(n-4)} (1-\rho r)^{-n+3/2} \times \\ \times F\left(\frac{1}{2}; \frac{1}{2}; n-\frac{1}{2}; \frac{1+\rho r}{2}\right), \end{aligned} \quad (85)$$

где

$$F(a, b, c, x) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)}{\Gamma(a)} \frac{\Gamma(b+j)}{\Gamma(b)} \frac{\Gamma(c)}{\Gamma(c+j)} \frac{x^j}{j!}$$

— гипергеометрическая функция.

[Чтобы получить первое выражение, произведите преобразование статистики (S_1^2, S_2^2, S_{12}) с плотностью (81) в (S_1^2, S_2^2, R) и разложите множитель $\{\rho s_{12}/(1-\rho^2)\sigma\tau\} = \exp\{\rho r s_1 s_2/(1-\rho^2)\sigma\tau\}$ в степенной ряд. Получающийся ряд можно проинтегрировать почленно по переменным s_1^2 и s_2^2 . Эквивалентность со вторым выражением усматривается из разложения множителя $(1-\rho r t)^{-(n-1)}$ под интегралом (83) и почленного интегрирования образующегося ряда.]

36. Если X и Y имеют двумерное нормальное распределение с коэффициентом корреляции $\rho > 0$, то они положительно зависимы в смысле раздела 11.

[Условное распределение Y при данном x нормально со средним $\eta + \rho\sigma\sigma^{-1}(x - \xi)$ и дисперсией $\tau^2(1-\rho^2)$. При добавлении к случайной величине с таким распределением положительного числа $\rho\sigma\sigma^{-1}(x' - x)$ она превращается в случайную величину с распределением, равным условному распределению Y при данном $x' > x$.]

*) Эта плотность и соответствующая функция распределения табулированы у David, Tables of the Correlation Coefficient, Cambridge Univ. Press, 1938.

13. Литературные ссылки

Басу (Basu D.)

(1955) On statistics independent of a complete sufficient statistic, *Sankhyā*, vol. 15, 377—380. [Теорема 2.]

(1958) On statistics independent of a complete sufficient statistics, *Sankhyā*, vol. 20, 223—226. [Исправленный вариант теоремы 2.]

Гиршик (Girshick M. A.)

(1946) Contribution to the theory of sequential analysis I, *Ann. Math. Stat.*, vol. 17, 123—143. [Задача 8.]

Кендалл и Стюарт (Kendall M. G., Stuart A.)

(1973) Статистические выводы и связи, М., Наука.

Леман (Lehmann E. L.)

(1947) On optimum tests of composite hypothesis with one constraint, *Ann. Math. Stat.*, vol. 18, 473—494. [Определяются наилучшие подобные области для некоторых задач, включая задачу 12.]

Леман и Стейн (Lehmann E. L. and Stein C.)

(1949) On the theory of some non — parametric hypotheses, *Ann. Math. Stat.*, vol. 20, 28—45. [Изложена теория оптимальных критериев перестановок.]

Морган (Morgan W. A.)

(1939) A test for the significance of the difference between the two variances in a sample from a normal bivariate population, *Biometrika*, vol. 31, 13—19. [Задача 33(1).]

Нейман (Neyman J.)

(1938) On statistics the distribution of which is independent of the parameters involved in the original probability law of the observed variables, *Stat. Res. Mem.*, vol. II, 58—59. [Теорема 2 при некоторых предположениях регулярности.]

Паулсон (Paulson E.)

(1941) On certain likelihood ratio tests associated with the exponential distribution, *Ann. Math. Stat.*, vol. 12, 301—306. [Рассматривается мощность критериев задачи 12.]

Питмен (Pitman E. I. G.)

(1937/38) Significance tests which may be applied to samples from any populations, I. *Roy. Stat. Suppl.*, vol. 4, 119—130, 225—232, and *Biometrika*, vol. 29, 322—335. [Изложена теория рандомизированных критериев с большим числом применений.]

(1939) A note on normal correlation, *Biometrika*, vol. 31, 9—12. [Задача 33(1).]

Стейн (Stein C.)

(1945) A two-sample test for a linear hypotheses whose power is independent of the variance, *Ann. Math. Stat.*, vol. 16, 243—258. [Задачи 15—17.]

Стьюдент (Student (Gosset W. S.))

(1908) On the probable error of a mean, *Biometrika*, vol. 6, 1—25. [Получено распределение t -статистики при X_i , распределенных по закону $N(0, \sigma^2)$. Строгое доказательство было дано в работе: R. A. Fisher, Note on Dr. Burnside's recent paper on error of observation, *Proc. Camb. Phil. Soc.*, vol. 21 (1923), 655—658.]

Сюй (Hsu C. T.)

(1940) On samples from a normal bivariate population, *Ann. Math. Stat.*, vol. 11, 410—426. [Задача 33 (II)]

Фишер (Fisher R. A.)

(1915) Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biometrika*, vol. 10, 507—521. [Получено распределение выборочного коэффициента корреляции для двумерного нормального распределения.]

(1931) Properties of the [Hh] functions, *Brit. Assoc. Math. Tables*, vol. 1 (3rd ed., 1951, XXVIII—XXXVII). [Вывод нецентрального t -распределения.]

(1935) *The Design of Experiments*, Edinburgh, Oliver and Boyd. [Содержит основные идеи, касающиеся критерия перестановок. В частности, отмечается, каким образом может быть использована рандомизация, и предлагается связанный с перестановками вариант t -критерия, как не требующий предположения нормальности.]

Хельмерт (Helmert F. R.)

(1876) Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers direkter Beobachtungen gleicher Genauigkeit, *Astron. Nachrichten*, vol. 88, № 2096—97, 113—132. [Получено распределение $\sum (X_i - \bar{X})^2$, где величины X независимы и распределены по нормальному закону.]

(1950) Some two — sample tests, *Ann. Math. Stat.*, vol. 21, 601—606. [Задачи 15—17 распространяются на случай сравнения двух средних.]

ГЛАВА 6

ИНВАРИАНТНОСТЬ

1. Симметрия и инвариантность

Многие статистические задачи обладают симметрией, приводящей к естественным ограничениям на статистические процедуры, которые следует использовать. Предположим, например, что $X_1, \dots, X_n$ — независимые случайные величины с плотностями распределения вероятностей $p_{\theta_1}(x_1), \dots, p_{\theta_n}(x_n)$ соответственно. При проверке гипотезы $H: \theta_1 = \dots = \theta_n$ при альтернативе, состоящей в том, что не все θ равны, естественно ограничиться рассмотрением только критериев, симметричных относительно значений $x_1, \dots, x_n$, поскольку в противном случае принятие или отклонение той или иной гипотезы зависело бы (что совершенно не относится к делу) от нумерации этих переменных.

В качестве другого примера рассмотрим круговую мишень с центром в точке O , на которой отмечены следы от попаданий выстрелов. Предположим, что выстрелы совершаются независимо и рассеяние их попаданий характеризуется двумерным нормальным распределением с центром в точке O . В задаче проверки круговой симметрии этого распределения относительно точки O представляется естественным ограничиться рассмотрением критериев, также обладающих подобной симметрией. В противном случае принятие или отклонение гипотез зависело бы от выбора необходимой для описания критерия двумерной системы координат (скажем, декартовой), которая при сделанных предположениях совершенно произвольна и не имеет никакого отношения к рассматриваемой проблеме.

Математическим выражением симметрии является инвариантность относительно подходящей группы преобразований. В первом из рассмотренных выше двух примеров такой группой является группа всех перестановок величин $x_1, \dots, x_n$, поскольку функция от n переменных симметрична тогда и только тогда, когда она остается инвариантной при всевозможных перестановках этих переменных. Во втором примере круговая симметрия

относительно точки O эквивалентна инвариантности относительно вращений около этой точки.

Вообще, пусть случайная величина X имеет распределение вероятностей P_θ , $\theta \in \Omega$ и пусть g — преобразование выборочного пространства $\mathcal{X}$. Все такие преобразования, рассматриваемые в связи с инвариантностью, будут предполагаться взаимно однозначными отображениями $\mathcal{X}$ на себя. Обозначим через gX случайную величину, принимающую значения gx при $X = x$, и предположим, что, когда распределение X есть P_θ , $\theta \in \Omega$, распределение величины gX равно $P_{\bar{g}\theta}$ с $\bar{g}\theta$, также принадлежащим Ω . Элемент $\bar{g}\theta \in \Omega$, связанный с θ указанным образом, будет обозначаться $\bar{g}\theta$, так что

$$P_\theta \{gX \in A\} = P_{\bar{g}\theta} \{X \in A\}. \quad (1)$$

Индекс θ в левой части равенства (1) указывает на распределение случайной величины X , а не gX . Равенство (1) может быть переписано в виде $P_\theta(g^{-1}A) = P_{\bar{g}\theta}(A)$, так что

$$P_{\bar{g}\theta}(gA) = P_\theta(A). \quad (2)$$

Параметрическое множество Ω остается инвариантным относительно преобразования g (или сохраняется при g), если $\bar{g}\theta \in \Omega$ для всех $\theta \in \Omega$ и если, дополнительно, для каждого $\theta' \in \Omega$ существует $\theta \in \Omega$ такое, что $\bar{g}\theta = \theta'$. Эти два условия можно выразить в форме равенства

$$\bar{g}\Omega = \Omega. \quad (3)$$

Определенное таким образом преобразование $\bar{g}$ множества Ω на себя взаимно однозначно, если только распределения P_θ , соответствующие различным значениям θ , различны. В самом деле, пусть $\bar{g}\theta_1 = \bar{g}\theta_2$. Тогда $P_{\bar{g}\theta_1}(gA) = P_{\bar{g}\theta_2}(gA)$ и поэтому $P_{\theta_1}(A) = P_{\theta_2}(A)$ для всех A , так что $\theta_1 = \theta_2$.

Л е м м а 1. Пусть g и g' — два преобразования, сохраняющие Ω . Тогда преобразования $g'g$ и g^{-1} , определенные формулами

$$(g'g)x = g'(gx) \quad \text{и} \quad g(g^{-1}x) = x \quad \text{при всех } x \in \mathcal{X},$$

также сохраняют Ω и удовлетворяют соотношениям

$$\overline{g'g} = \bar{g}'\bar{g} \quad \text{и} \quad (\bar{g}^{-1}) = (\bar{g})^{-1}. \quad (4)$$

Доказательство. Если распределение случайной величины X есть P_θ , так что распределение gX равно $P_{\bar{g}\theta}$, то распределением величины $g'gX = g'(gX)$ является $P_{\bar{g}'\bar{g}\theta}$. Этим устанавливается первое из соотношений (4); доказательство второго проводится аналогично.

Мы будем говорить, что *задача проверки гипотезы* $H: \theta \in \Omega_H$ *при альтернативе* $K: \theta \in \Omega_K$ *остаётся инвариантной* относительно преобразования g , если $\bar{g}$ сохраняет пространства Ω_H и Ω_K , т. е. если в дополнение к (3) выполнено соотношение

$$\bar{g}\Omega_H = \Omega_H. \quad (5)$$

Пусть $\mathcal{G}$ — класс преобразований, удовлетворяющий этим двум условиям, и пусть G будет наименьший класс преобразований, содержащий $\mathcal{G}$ и такой, что из $g, g' \in G$ следует, что $g'g$ и g^{-1} принадлежит G . Тогда G есть группа преобразований, каждое из которых, согласно лемме 1, сохраняет Ω и Ω_H . Поэтому любой класс преобразований $\mathcal{G}$, оставляющий проблему инвариантной, может быть расширен до группы G . Из леммы 1 также следует, что класс индуцированных преобразований $\bar{g}$ образует группу, скажем, $\bar{G}$. Соотношения (4) выражают тот факт, что $\bar{G}$ гомоморфно группе G .

При наличии симметрии выборочного и параметрических пространств, выраженной группами G и $\bar{G}$, естественно ограничиться рассмотрением только симметричных критериев, т. е. критериев, удовлетворяющих соотношению

$$\varphi(gx) = \varphi(x) \quad \text{для всех } x \in X \text{ и } g \in G. \quad (6)$$

Критерий φ , удовлетворяющий (6), называется *инвариантным относительно группы* G . Ограничения инвариантными критериями есть частный случай применения принципа инвариантности, сформулированного в разделе 5 главы 1, где, равно как и в приведенных ранее примерах, было указано, что преобразование g может быть интерпретировано как замена системы координат. С этой точки зрения, критерий будет инвариантным, если он не зависит от выбора частной системы координат, служащей для выражения результатов наблюдений.

Чтобы проблема оставалась инвариантной, преобразование g необходимо должно сохранять и класс $\mathcal{A}$ измеримых множеств, на которых определены распределения P_θ . Это означает, что каждое множество $A \in \mathcal{A}$ преобразуется в множество же из $\mathcal{A}$ и является образом такого же множества, т. е., что gA и $g^{-1}A$ оба принадлежат $\mathcal{A}$. Каждое преобразование, удовлетворяющее этим условиям, называется *биизмеримым*. Поскольку группа вместе с каждым элементом g содержит также и g^{-1} , то все ее элементы автоматически биизмеримы, если каждый из них измерим. Если g' и g биизмеримы, то таковы же $g'g$ и g^{-1} . Элементы группы G , порожденной классом преобразований $\mathcal{G}$ указанным выше способом, будут все биизмеримыми, если только таковыми являются преобразования из $\mathcal{G}$.

2. Максимальные инварианты

Если задача проверки гипотез остается инвариантной относительно некоторой группы преобразований, то применение *принципа инвариантности* приводит к рассмотрению инвариантных критериев. Чтобы выбрать среди них наилучший, удобно сначала охарактеризовать совокупность всех инвариантных критериев.

Две точки x_1, x_2 будут рассматриваться как эквивалентные относительно группы G ,

$$x_1 \sim x_2 \pmod{G},$$

если существует преобразование $g \in G$, для которого $x_2 = gx_1$. Определенное так отношение действительно является отношением эквивалентности, поскольку G есть группа, и поэтому множества эквивалентных точек, *траекторий* G , образуют разбиение выборочного пространства (см. Дополнение, раздел 1). Точка x пробегает траекторию, когда к ней применяются все преобразования g из G ; это означает, что траектория, содержащая точку, состоит из всех точек gx с $g \in G$. Из определения инвариантности следует, что каждая функция инвариантна тогда и только тогда, когда она постоянна на каждой траектории.

Функция T называется *максимальным инвариантом*, если она инвариантна и если выполняется условие:

из $T(x_1) = T(x_2)$ следует, что $x_2 = gx_1$, для некоторого $g \in G$, (7)

т. е. если она постоянна на каждой траектории и на различных траекториях принимает различные значения. Все максимальные инварианты эквивалентны в том смысле, что множества точек, где они постоянны, совпадают.

Теорема 1. Пусть $T(x)$ — максимальный инвариант относительно группы G . Тогда необходимое и достаточное условие для инвариантности функции φ состоит в том, чтобы φ зависело от x только через $T(x)$, т. е. чтобы существовала функция h такая, что $\varphi(x) = h(T(x))$ для всех x .

Доказательство. Если $\varphi(x) = h(T(x))$ для всех x , тогда $\varphi(gx) = h(T(gx)) = h(T(x)) = \varphi(x)$, т. е. φ инвариантна. С другой стороны, если φ инвариантна и если $T(x_1) = T(x_2)$, тогда $x_2 = gx_1$ для некоторой g , и поэтому $\varphi(x_2) = \varphi(x_1)$.

Пример 1. (I) Пусть $x = (x_1, \dots, x_n)$, и пусть G — группа сдвигов $gx = (x_1 + c, \dots, x_n + c)$, $-\infty < c < \infty$.

Тогда множество разностей $y = (x_1 - x_n, \dots, x_{n-1} - x_n)$ инвариантно относительно G . Чтобы показать, что функция y в то же время есть и максимальный инвариант, предположим, что $x'_i - x'_n = x_i - x_n$ для $i = 1, \dots, n-1$. Полагая $x'_n - x_n = c$, видим, что $x'_i = x_i + c$ для всех i , что и нужно было показать. Функция y является, конечно, только одним из примеров представления максимального инварианта. В качестве других можно указать на функ-

цию $(x_1 - x_2, x_2 - x_3, \dots, x_{n-1} - x_n)$ или $(x_1 - \bar{x}, \dots, x_n - \bar{x})$. В частном случае, когда $n = 1$, не существует нетривиальных инвариантных функций. Все пространство состоит из одной траектории, так что любые две точки преобразованием из G могут быть переведены одна в другую. В таком случае группа преобразований называется *транзитивной*. Инвариантными функциями могут быть только константы $\varphi(x) \equiv c$.

(II) Если G есть группа преобразований

$$gx = (cx_1, \dots, cx_n), \quad c \neq 0,$$

то специальную роль играет начало координат. Однако в статистических приложениях множество точек, отличных от начала координат, имеет меру единицы, поэтому можно ограничиться рассмотрением только этой части выборочного пространства и тогда множество отношений $x_1/x_n, \dots, x_{n-1}/x_n$ будет максимальным инвариантом. Без этого ограничения две точки x, x' эквивалентны относительно максимального инвариантного разбиения, если среди их координат имеется одно и то же число нулевых координат (при условии, что таковые вообще имеются), если они стоят на одном и том же месте и если для каждого двух ненулевых координат x_i, x'_i отношения x_j/x_i и x'_j/x'_i равны.

(III) Пусть $x \neq (x_1, \dots, x_n)$ и пусть G — группа всех ортогональных преобразований $x' = Gx$ n -мерного пространства. Тогда функция $\sum x_i^2$ является максимальным инвариантом, т. е. две точки x и x^* могут быть переведены одна в другую ортогональным преобразованием в том и только том случае, когда они находятся на одинаковых расстояниях от начала координат. Доказательство этого получается немедленно, если ограничиться рассмотрением плоскости, содержащей точки x, x^* и начало координат.

Пример 2. Пусть $x = (x_1, \dots, x_n)$ и G — множество из $n!$ перестановок координат x . Тогда множество упорядоченных координат (*порядковых статистик*) $x^{(1)} \leq \dots \leq x^{(n)}$ является максимальным инвариантом. Перестановка координат x_i не изменяет, очевидно, множества значений координат и поэтому не изменяет величин $x^{(i)}$. С другой стороны, две точки с одними и теми же значениями порядковых статистик могут быть получены одна из другой перестановкой координат.

Пример 3. Пусть G — совокупность преобразований $x'_i = f(x_i)$, $i = 1, \dots, n$, таких, что f непрерывна и строго возрастает, и предположим, что мы рассматриваем только точки, у которых все координаты различны. Если координаты рассматривать как n точек на действительной прямой, то указанное преобразование сохраняет их порядок. Обратно, если $x_1, \dots, x_n$ и $x'_1, \dots, x'_n$ — две точки с одним и тем же порядком, скажем $x_{i_1} < \dots < x_{i_n}$ и $x'_{i_1} < \dots < x'_{i_n}$, то существует преобразование f , удовлетворяющее сформулированным условиям и такое, что для всех i $x'_i = f(x_i)$. Например, такую функцию можно определить, положив $f(x) = x + (x'_{i_1} - x_{i_1})$ для $x \leq x_{i_1}$, $f(x) = x + (x'_{i_n} - x_{i_n})$ для $x \geq x_{i_n}$ и сделав ее линейной между значениями x_{i_k} и $x_{i_{k+1}}$, $k = 1, \dots, n-1$. Формальным выражением для максимального инварианта в этом случае служит множество *рангов* $(r_1, \dots, r_n)$ вектора $(x_1, \dots, x_n)$. Здесь ранг r_i величины x_i определяется посредством соотношения

$$x_i = x^{(r_i)},$$

так что r_i есть число координат $x \leq x_i$. В частности, $r_i = 1$, если x_i есть самая наименьшая из координат, $r_i = 2$, если x_i — вторая наименьшая координата, и т. д.

Часто оказывается удобным получать максимальные инварианты в несколько этапов, на каждом шагу находя их для подгрупп группы G . Чтобы проиллюстрировать этот процесс и трудности, которые могут встретиться в его применении, положим $x = (x_1, \dots, x_n)$ и допустим, что все координаты x_i различны и группа задается преобразованием

$$gx = (ax_1 + b, \dots, ax_n + b), \quad a \neq 0, \quad -\infty < b < \infty.$$

Применяя сначала подгруппу сдвигов $x'_i = x_i + b$, найдем, что максимальный инвариант есть функция $y = (y_1, \dots, y_{n-1})$ с $y_i = x_i - x_n$. Другая подгруппа состоит в умножении на скалярный множитель $x''_i = ax_i$. Это преобразование индуцирует соответствующее умножение на множитель для y_k : $y''_i = ay_i$, и максимальным инвариантом относительно такого преобразования пространства y_k является функция $z = (z_1, \dots, z_{n-2})$ с $z_i = y_i/y_{n-1}$. Откуда, в терминах x_k получим, что функции $z_i = (x_i - x_n)/(x_{n-1} - x_n)$ являются максимальным инвариантом относительно группы G .

Предположим сейчас, что этот процесс производится в другом порядке. Применяя сначала подгруппу $x''_i = ax_i$, получаем максимальный инвариант $u = (u_1, \dots, u_n)$ с $u_i = x_i/x_n$. Однако преобразования $x'_i = x_i + b$ не индуцируют преобразования в u -пространстве, поскольку $(x_i + b)/(x_n + b)$ не является функцией от x_i/x_n .

Вообще, пусть группа преобразований G порождена двумя подгруппами D и E в том смысле, что она является наименьшей группой, содержащей D и E . Тогда G состоит из множества произведений $e_m d_m \dots e_1 d_1$, $m = 1, 2, \dots$, $d_i \in D$, $e_i \in E$ ($i = 1, \dots, m$)*). Следующая теорема показывает, что всякий раз, когда процесс нахождения максимальных инвариантов в несколько этапов вообще можно провести, он всегда приводит к максимальным инвариантам относительно группы G .

Теорема 2. Пусть G — группа преобразований и пусть D и E — подгруппы, порождающие G . Предположим, что $y = s(x)$ — максимальный инвариант относительно группы D и для каждого $e \in E$ выполняется условие

$$s(x_1) = s(x_2) \text{ влечет } s(ex_1) = s(ex_2). \quad (8)$$

Если $z = t(y)$ есть максимальный инвариант относительно группы E^* , состоящей из элементов e^* таких, что $e^*y = s(ex)$, когда $y = s(x)$, то $z = t[s(x)]$ является максимальным инвариантом относительно группы G .

*) См. раздел I в Дополнении,

Доказательство. Чтобы показать, что функция $t[s(x)]$ инвариантна, возьмем $x' = gx$, $g = e_m d_m \dots e_1 d_1$. Тогда

$$\begin{aligned} t[s(x')] &= t[s(e_m d_m \dots e_1 d_1 x)] = t[e_m^* s(d_m \dots e_1 d_1 x)] = \\ &= t[s(e_{m-1} d_{m-1} \dots e_1 d_1 x)], \end{aligned}$$

и по индукции последнее выражение можно свести к $t[s(x)]$. Теперь, чтобы показать, что функция $t[s(x)]$ в действительности является максимальным инвариантом, предположим, что $t[s(x')] = t[s(x)]$. Полагая $y' = s(x')$, $y = s(x)$, получим $t(y') = t(y)$ и поскольку $t(y)$ — максимальный инвариант относительно E^* , то существует e^* такое, что $y' = e^* y$. Тогда $s(x') = e^* s(x) = s(ex)$ и, в силу того, что функция $s(x)$ — максимальный инвариант относительно D , существует $d \in D$ такое, что $x' = dex$. Но так как de есть элемент G , то это и завершает доказательство.

3. Наиболее мощные инвариантные критерии

Класс всех инвариантных функций может быть получен как совокупность функций от максимального инварианта $T(x)$. Поэтому, в частности, класс всех инвариантных критериев совпадает со множеством критериев, зависящих только от максимальной инвариантной статистики $T(x)$. Это замечание, будучи справедливым во всех обычных ситуациях, требует некоторых ограничений на класс измеримых множеств в T -пространстве, которые будут обсуждены в конце настоящего раздела; в приводимых ниже примерах они выполняются.

Пример 4. Пусть $X = (X_1, \dots, X_n)$. Предположим, что плотность случайной величины X при гипотезе H_i ($i = 0, 1$) равна $f_i(x_1 - \theta, \dots, x_n - \theta)$, где θ изменяется от $-\infty$ до ∞ . Задача проверки гипотезы H_0 при альтернативе H_1 инвариантна относительно группы (G) преобразований

$$gx = (x_1 + c, \dots, x_n + c), \quad -\infty < c < \infty,$$

которая в параметрическом пространстве индуцирует преобразования

$$\bar{g}\theta = \theta + c.$$

Согласно примеру 1 максимальным инвариантом относительно группы G является функция $Y = (X_1 - X_n, \dots, X_{n-1} - X_n)$. Распределение Y не зависит от θ и при гипотезе H_i имеет плотность

$$\int_{-\infty}^{\infty} f_i(y_1 + z, \dots, y_{n-1} + z, z) dz.$$

Если рассматривать случайную величину Y , то проблема проверки гипотезы H_0 при альтернативе H_1 превращается в задачу проверки простой гипотезы при простой альтернативе. Наиболее мощный критерий тогда не зависит от θ

и поэтому является РНМ среди всех инвариантных критериев. Согласно лемме Неймана — Пирсона гипотеза H_0 отвергается, если

$$\frac{\int_{-\infty}^{\infty} f_1(y_1 + z, \dots, y_{n-1} + z, z) dz}{\int_{-\infty}^{\infty} f_0(y_1 + z, \dots, y_{n-1} + z, z) dz} = \frac{\int_{-\infty}^{\infty} f_1(x_1 + u, \dots, x_n + u) du}{\int_{-\infty}^{\infty} f_0(x_1 + u, \dots, x_n + u) du} > C.$$

Пример 5. Если $X_1, \dots, X_n$ — выборка из нормального распределения $N(\xi, \sigma^2)$, то гипотеза $H: \sigma \geq \sigma_0$ остается инвариантной при преобразовании $X'_i = X_i + c$, $-\infty < c < \infty$. В терминах достаточных статистик $Y = \bar{X}$, $S^2 = \sum (X_i - \bar{X})^2$ это преобразование переходит в преобразование $Y' = Y + c$, $(S^2)' = S^2$, и максимальным инвариантом является S^2 . Поэтому класс инвариантных критериев есть класс критериев, зависящих от S^2 . Как следует из теоремы 2 главы 3, существует РНМ инвариантный критерий, отклоняющий гипотезу H_0 , когда $\sum (X_i - \bar{X})^2 \leq C$. Этот критерий совпадает с РНМ несмещенным критерием (9) из главы 5.

Пример 6. Если $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из двух нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$, то множество достаточных статистик состоит из $T_1 = \bar{X}$, $T_2 = \bar{Y}$, $T_3 = \sqrt{\sum (X_i - \bar{X})^2}$, $T_4 = \sqrt{\sum (Y_i - \bar{Y})^2}$. Задача проверки гипотезы $H: \tau^2/\sigma^2 \leq \Delta_0$ остается инвариантной при преобразованиях $T'_1 = T_1 + c_1$, $T'_2 = T_2 + c_2$, $T'_3 = T_3$, $T'_4 = T_4$, $-\infty < c_1, c_2 < \infty$, а также относительно умножения на одно и то же число всех четырех величин. Максимальным инвариантом относительно первой группы преобразований является функция (T_3, T_4) . В пространстве значений этой функции группа умножений на числа индуцирует преобразования $T''_3 = cT_3$, $T''_4 = cT_4$, $0 < c$, относительно которых максимальным инвариантом будет отношение T_4/T_3 . Статистика $Z = T_4^2/(n-1) : T_3^2/(m-1)$, поделенная на $\Delta = \tau^2/\sigma^2$, имеет распределение F с плотностью, задаваемой формулой (21) главы 5, так что плотность распределения вероятностей величины Z равна

$$\frac{C(\Delta) z^{1/2(n-3)}}{\left(\Delta + \frac{n-1}{m-1} z\right)^{1/2(m+n-2)}}, \quad z > 0.$$

Это семейство плотностей, зависящее от параметра Δ , является семейством с монотонным отношением правдоподобия; таким образом, среди всех критериев для проверки гипотезы H , основанных на статистике Z , а следовательно, и среди всех инвариантных критериев, существует РНМ с критической областью $Z > C$. Этот критерий совпадает с РНМ несмещенным критерием (20) главы 5.

Пример 7. В методе *парных сравнений*, применяемом для выяснения эффективности какого-либо способа обработки, экспериментальный материал состоит из n пар субъектов. Из каждой пары случайным образом выделяется для обработки один субъект, тогда как второй служит для контроля. Пусть X_i равно 1 или 0 в зависимости от того, говорит ли результат эксперимента с i -й парой в пользу субъекта, подвергнувшегося обработке или в пользу контрольного; пусть $p_i = P\{X_i = 1\}$. Рассмотрим гипотезу H об отсутствии эффекта обработки, т. е. о том, что $p_i = 1/2$, $i = 1, \dots, n$, при альтернативе, состоящей в том, что для всех i $p_i > 1/2$.

Эта задача остается инвариантной относительно всех перестановок n величин $X_1, \dots, X_n$, и максимальным инвариантом относительно этой группы является функция, равная суммарному числу успехов $X = X_1 + \dots + X_n$. Распределение X задается формулой

$$P\{X = k\} = q_1 \dots q_n \sum \frac{p_{i_1}}{q_{i_1}} \dots \frac{p_{i_k}}{q_{i_k}},$$

где $q_i = 1 - p_i$ и где суммирование распространено по всем $\binom{n}{k}$ наборам индексов $i_1 < \dots < i_k$. Наиболее мощный инвариантный критерий при альтернативной гипотезе $(p'_1, \dots, p'_n)$ отвергает H , когда

$$f(k) = \frac{1}{\binom{n}{k}} \sum \frac{p'_{i_1}}{q'_{i_1}} \dots \frac{p'_{i_k}}{q'_{i_k}} > C.$$

Чтобы показать, что $f(k)$ есть возрастающая функция от k , заметим, что $a_i = p'_i/q'_i > 1$ и что

$$\sum_j \sum a_j a_{i_1} \dots a_{i_k} = (k+1) \sum a_{i_1} \dots a_{i_{k+1}}$$

и

$$\sum_j \sum a_{i_1} \dots a_{i_k} = (n-k) \sum a_{i_1} \dots a_{i_k}.$$

Здесь в обоих равенствах второе суммирование в левых частях распространено по всем индексам $i_1 < \dots < i_k$, из которых ни один не равен j , а суммирование в правых частях производится соответственно по всем $i_1 < \dots < i_k$ и $i_1 < \dots < i_{k+1}$ без всяких ограничений. Тогда

$$\begin{aligned} f(k+1) &= \frac{1}{\binom{n}{k+1}} \sum a_{i_1} \dots a_{i_{k+1}} = \\ &= \frac{1}{(n-k) \binom{n}{k}} \sum_j \sum a_j a_{i_1} \dots a_{i_k} > \frac{1}{\binom{n}{k}} \sum a_{i_1} \dots a_{i_k} = f(k), \end{aligned}$$

что и надо было показать. Поэтому независимо от выбранной альтернативы критерий отвергает гипотезу H , когда $k > C$, и, следовательно, является РНМ инвариантным критерием.

Если i -му сравнению приписать знак плюс при X_i , равном 1, и минус при X_i равном 0, то, как легко видеть, этот критерий превращается в критерий знаков (см. пример 8 главы 3 и раздел 7 главы 4).

Достаточные статистики упрощают задачу посредством редукиции выборочного пространства, причем этот процесс не изменяет параметрического пространства. С другой стороны, инвариантность путем редуцирования данных к максимально инвариантной статистике T (распределение которой может зависеть только от некоторой функции от параметра) обычно сокращает также и параметрическое пространство. Точные утверждения содержатся в следующей теореме.

Теорема 3. Если функция $T(x)^1$ инвариантна относительно группы преобразований G и если $v(\theta)$ является максимальным инвариантом относительно индуцированной группы $\bar{G}$, то распределение $T(X)$ зависит только от $v(\theta)$.

Доказательство. Пусть $v(\theta_1) = v(\theta_2)$. Тогда $\theta_2 = \bar{g}\theta_1$, и, следовательно,

$$\begin{aligned} P_{\theta_1} \{T(X) \in B\} &= P_{\theta_1} \{T(gX) \in B\} = \\ &= P_{\bar{g}\theta_1} \{T(X) \in B\} = P_{\theta_2} \{T(X) \in B\}. \end{aligned}$$

Этот результат можно перефразировать, сказав, что принцип инвариантности идентифицирует все параметрические точки, эквивалентные относительно преобразований $\bar{G}$.

В приложениях, например в примерах 5 и 6, максимальные инварианты $T(x)$ и $\delta = v(\theta)$ относительно G и $\bar{G}$, часто оказываются действительными функциями, а семейство вероятностных плотностей $p_\delta(t)$ величины T имеет монотонное отношение правдоподобия. Тогда при проверке гипотезы $H: \delta \leq \delta_0$ среди всех критериев, зависящих только от T , существует РНМ, который, следовательно, является РНМ инвариантным критерием. При этом гипотеза H отвергается, если $t \geq C$, где

$$\int_C^\infty p_{\delta_0}(t) dt = \alpha. \quad (9)$$

Рассмотрим теперь эту проблему как задачу с двумя решениями d_0 и d_1 , соответствующими принятию или отклонению гипотезы H , и с функцией потерь $L(\theta, d_i) = L_i(\theta)$. Предположим, что $L_i(\theta)$ зависит только от параметра δ , скажем, пусть $L_i(\theta) = L'_i(\delta)$ и пусть

$$L'_1(\delta) - L'_0(\delta) \geq 0, \quad \text{когда} \quad \delta \geq \delta_0. \quad (10)$$

В таком случае, из теоремы 3 главы 3 следует, что семейству критических областей $t \geq C(\alpha)$, при изменении α от 0 до 1 соответствует семейство решающих процедур полное относительно класса процедур, зависящих только от t , т. е. полное семейство инвариантных процедур. При этом, как и раньше, выбор частных значений уровня значимости α можно рассматривать как удобный способ выделения критериев из этого семейства.

В начале данного раздела было установлено, что класс всех инвариантных критериев совпадает с классом критериев, основанных на статистике $T = T(X)$, являющейся максимальным инвариантом. Однако указанием некоторой функции статистика полностью не определяется, поскольку требуется также указать класс $\mathcal{B}$ измеримых множеств. Если в рассматриваемом случае $\mathcal{B}$ есть класс всех множеств B , для которых $T^{-1}(B) \in \mathcal{A}$, то

класс всех инвариантных критериев совпадает с классом критериев, основанных на статистике T . Действительно, пусть $\varphi(x) = \psi[T(x)]$, φ — $\mathcal{A}$ -измерима и C — борелевское множество на числовой прямой. Тогда $\varphi^{-1}(C) = T^{-1}[\psi^{-1}(C)] \in \mathcal{A}$ и, следовательно, $\psi^{-1}(C) \in \mathcal{B}$, так что ψ $\mathcal{B}$ -измерима и $\varphi(x) = \psi[T(x)]$ является критерием, основанным на статистике T .

В большинстве приложений $T(x)$ является измеримой функцией со значениями в евклидовом пространстве. Поэтому удобно в качестве системы $\mathcal{B}$ множеств взять здесь класс всех борелевских множеств. Тогда если $\varphi(x) = \psi[T(x)]$ — произвольная измеримая функция, зависящая только от $T(x)$, то неясно, будет ли $\psi(t)$ необходимо $\mathcal{B}$ -измеримой. Факт измеримости можно установить, если предположить, что $\mathcal{X}$ также евклидово пространство, $\mathcal{A}$ — класс борелевских множеств и что область T — борелевское множество. Мы сейчас установим этот факт при одном дополнительном предположении (которое в приложениях обычно очевидным образом выполняется и которое поэтому в дальнейшем не будет проверяться в каждом отдельном случае), что существует измеримая по Борелю векторная функция $Y(x)$ такая, что $[T(x), Y(x)]$ отображает пространство $\mathcal{X}$ в борелевское подмножество пространства $\mathcal{T} \times \mathcal{Y}$, что это отображение взаимно однозначно и что обратное отображение также является измеримым по Борелю. Тогда для каждой измеримой функции φ от x существует измеримая функция φ' от (t, y) такая, что $\varphi(x) = \varphi'[T(x), Y(x)]$. Если φ зависит только от $T(x)$, то функция φ' зависит только от t , т. е. $\varphi'(t, y) = \psi(t)$, где ψ есть измеримая функция от t *). Так, в примере 1 (I), где $x = (x_1, \dots, x_n)$ и $T(x) = (x_1 - x_n, \dots, x_{n-1} - x_n)$, в качестве функции $Y(x)$ можно взять $Y(x) = x_n$.

4. Выборочный контроль по количественному признаку

Из партии некоторых изделий производится выборка с целью решения вопроса о том, является ли качество этой партии приемлемым. В простейшем случае каждое изделие классифицируется непосредственно как годное или негодное (контроль по *качественному признаку*), и решение выносится на основе суммарного числа дефектных изделий. В более общих случаях качество изделия характеризуется случайной величиной Y (*контроль по количественному признаку*) и изделие считается годным, если Y превышает заданную константу u . Тогда вероятность того, что изделие дефектно, равна

$$p = P\{Y \leq u\},$$

и задача состоит в том, чтобы проверить гипотезу $H: p \geq p_0$.

*) Последнее утверждение немедленно следует, например, из теоремы 2 § 34 книги П. Халмоша, Теория меры, ИЛ, 1953.

Как было показано в примере 8 главы 3, конкретные значения Y нельзя использовать для проверки этой гипотезы, если ничего не известно о распределении Y . В отсутствие подобной информации решение, как и раньше, будет приниматься просто по числу дефектных изделий в выборке. Мы рассмотрим сейчас эту задачу в предположении, что величины $Y_1, \dots, Y_n$ образуют выборку из нормального распределения $N(\eta, \sigma^2)$.

Тогда

$$p = \int_{-\infty}^u \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(y - \eta)^2\right] dy = \Phi\left(\frac{u - \eta}{\sigma}\right),$$

где

$$\Phi(y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt$$

обозначает функцию распределения стандартного нормального распределения, и гипотеза H эквивалентна тому, что $(u - \eta)/\sigma \geq \Phi^{-1}(p_0)$. В терминах переменных $X_i = Y_i - u$, имеющих среднее $\xi = \eta - u$ и дисперсию σ^2 , исходная задача сводится к проверке гипотезы

$$H: \xi/\sigma \leq \theta_0,$$

с $\theta_0 = -\Phi^{-1}(p_0)$. Необходимость проверки этой гипотезы, которая для случая $\theta_0 = 0$ была рассмотрена в разделе 2 главы 5, возникает также и в других случаях. К подобной задаче приходят, например, когда интересуются средним ξ нормального распределения, выраженным в единицах стандартного отклонения, а не в единицах фиксированной шкалы.

Для проверки гипотезы H можно ограничиться рассмотрением пары величин $\bar{X}$, $S = \sqrt{\sum (X_i - \bar{X})^2}$, поскольку они образуют множество достаточных статистик для (ξ, σ) . Эти случайные величины независимы, распределение $\bar{X}$ совпадает с нормальным $N(\xi, \sigma^2/n)$, а распределение S/σ есть χ_{n-1} -распределение. Умножение $\bar{X}$ и S на общий множитель $c > 0$ преобразует параметры в $\xi' = c\xi$, $\sigma' = c\sigma$, так что отношение ξ/σ , а следовательно, и задача проверки гипотезы H остаются инвариантными. Относительно этих преобразований максимальным инвариантом является функция $\bar{x}/s$ или

$$t = \frac{\sqrt{n} \bar{x}}{s/\sqrt{n-1}},$$

распределение которой зависит только от максимального инварианта в параметрическом пространстве $\theta = \xi/\sigma$ (см. раздел 2 главы 5). Тем самым, инвариантные критерии — это критерии,

которые зависят только от t . Остается в этом классе найти наиболее мощный критерий для проверки гипотезы $H: \theta \leq \theta_0$.

Плотность вероятности t равна (задача 3, главы 5)

$$P_\delta(t) = \\ = C \int_0^\infty \exp \left[-\frac{1}{2} \left(t \sqrt{\frac{w}{n-1}} - \delta \right)^2 \right] w^{\frac{1}{2}(n-2)} \exp \left(-\frac{1}{2} w \right) dw,$$

где $\delta = \sqrt{n}\theta$ — параметр нецентральности. Покажем, что семейство этих распределений имеет монотонное отношение правдоподобия. Чтобы установить, что отношение

$$r(t) = \frac{\int_0^\infty \exp \left[-\frac{1}{2} \left(t \sqrt{\frac{w}{n-1}} - \delta_1 \right)^2 \right] w^{\frac{1}{2}(n-2)} \exp \left(-\frac{1}{2} w \right) dw}{\int_0^\infty \exp \left[-\frac{1}{2} \left(t \sqrt{\frac{w}{n-1}} - \delta_0 \right)^2 \right] w^{\frac{1}{2}(n-2)} \exp \left(-\frac{1}{2} w \right) dw}$$

является возрастающей функцией от t для $\delta_0 < \delta_1$, предположим сначала, что $t < 0$ и положим $v = -t \sqrt{w/(n-1)}$. Тогда рассматриваемое отношение пропорционально

$$\frac{\int_0^\infty f(v) \exp [-(\delta_1 - \delta_0)v - (n-1)v^2/2t^2] dv}{\int_0^\infty f(v) \exp [-(n-1)v^2/2t^2] dv} = \\ = \int \exp [-(\delta_1 - \delta_0)v] g_{t^2}(v) dv,$$

где

$$f(v) = \exp(-\delta_0 v) v^{n-1} \exp \left(-\frac{v^2}{2} \right)$$

и

$$g_{t^2}(v) = \frac{f(v) \exp [-(n-1)v^2/2t^2]}{\int_0^\infty f(z) \exp [-(n-1)z^2/2t^2] dz}.$$

Поскольку семейство плотностей $g_{t^2}(v)$ имеет монотонное отношение правдоподобия, то интеграл от $\exp [-(\delta_1 - \delta_0)v]$ по этой плотности является убывающей функцией от t^2 (задача 10 из главы 3) и, следовательно, возрастающей функцией от t для $t < 0$. Аналогично рассмотрев преобразование $v = t \sqrt{w/(n-1)}$, найдем, что $r(t)$ есть возрастающая функция от t при $t > 0$. Отсюда по непрерывности следует, что функция $r(t)$ возрастает с ростом t для всех t .

Тем самым для проверки гипотезы $H: \xi/\sigma \leq \theta_0$ существует РНМ инвариантный критерий, который ее отвергает, когда $t > C$, где константа C определяется из условия (9). В терминах первоначальных величин Y_i область отклонения гипотезы $H: \rho \geq \rho_0$ для РНМ инвариантного критерия принимает вид

$$\frac{\sqrt{n}(\bar{y} - u)}{\sqrt{\sum (y_i - \bar{y})^2 / (n-1)}} > C. \quad (11)$$

Если эту задачу рассматривать как задачу с двумя решениями и функциями потерь $L_0(\rho)$, $L_1(\rho)$ (потери от принятия или отклонения гипотезы $\rho \geq \rho_0$), зависящими только от ρ и удовлетворяющими условию, аналогичному (10), то класс критериев (11) образует полное семейство инвариантных процедур, когда C изменяется в пределах от $-\infty$ до ∞ .

Рассмотрим далее сравнение двух партий изделий, основанное на выборках $X_1, \dots, X_m$; $Y_1, \dots, Y_n$ из распределений $N(\xi, \sigma^2)$ и $N(\eta, \sigma^2)$ соответственно. Если

$$\rho = \Phi\left(\frac{u - \xi}{\sigma}\right), \quad \pi = \Phi\left(\frac{u - \eta}{\sigma}\right),$$

то возникает задача проверки гипотезы $\rho \leq \pi$ или, что то же, проверки гипотезы

$$H: \eta \leq \xi.$$

Статистики $\bar{X}$, $\bar{Y}$, $S = \sqrt{\sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2}$ образуют семейство достаточных статистик для ξ , η , σ . Рассматриваемая задача остается инвариантной при добавлении произвольной общей константы к $\bar{X}$, $\bar{Y}$. Отсюда следует, что $\bar{Y} - \bar{X}$ и S являются максимальными инвариантами. Эти функции остаются также инвариантными и при умножении величин $\bar{X}$, $\bar{Y}$ и S , а следовательно, и величин $\bar{Y} - \bar{X}$ и S на общую положительную константу, что редуцирует исходные данные к максимальному инварианту $(\bar{Y} - \bar{X})/S$. Поскольку величина

$$t = \frac{(\bar{y} - \bar{x}) / \sqrt{\frac{1}{m} + \frac{1}{n}}}{s / \sqrt{m + n - 2}}$$

имеет нецентральное t -распределение с параметром нецентральности $\delta = \frac{\sqrt{mn}(\eta - \xi)}{\sqrt{m + n}\sigma}$, то РНМ инвариантный критерий отклоняет гипотезу $H: \eta - \xi \leq 0$, когда $t > C$. Этот критерий совпадает с РНМ несмещенным критерием (27) из раздела 3 главы 5.

Аналогично, соответствующий двусторонний критерий (30) главы 5, с критической областью $|t| \geq C$ является РНМ инвариантным критерием для проверки гипотезы $p = \pi$ при альтернативе $p \neq \pi$. (Задача 10.)

5. Почти инвариантность

Пусть G — группа преобразований, оставляющая семейство $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ распределений X инвариантным. Критерий φ будет называться *эквивалентным инвариантному критерию*, если существует инвариантный критерий ψ такой, что $\varphi(x) = \psi(x)$ для всех x , за исключением множества N $\mathcal{P}$ -меры нуль; φ называется *почти инвариантным относительно группы G* , если

$$\varphi(gx) = \varphi(x) \quad \text{для всех } x \in \mathcal{X} - N_g, \quad g \in G, \quad (12)$$

где исключительное множество N_g меры нуль, возможно, зависит от g . Это определение будет использоваться при изучении связи между инвариантностью, с одной стороны, несмещенностью и некоторыми другими оптимальными свойствами, с другой. Поэтому важно знать, является ли РНМ инвариантный критерий также и РНМ среди почти инвариантных критериев. Оказывается, что в предположениях, которые точно сформулированы ниже, в теореме 4, и которые выполняются во всех обычных применениях, это действительно так.

Если критерий φ эквивалентен некоторому инвариантному критерию, то $\varphi(gx) = \varphi(x)$ для всех $x \notin N \cup g^{-1}N$. Поскольку $P_\theta(g^{-1}N) = P_{g\theta}(N) = 0$, то отсюда следует, что критерий φ почти инвариантен. Следующая теорема дает условия, при которых, наоборот, каждый почти инвариантный критерий эквивалентен некоторому инвариантному критерию.

Теорема 4. Пусть G — группа преобразований $\mathcal{X}$ и пусть $\mathcal{A}$ и $\mathcal{B}$ — σ -поля подмножеств $\mathcal{X}$ и G такие, что для каждого $A \in \mathcal{A}$ множество пар (x, g) , для которых $gx \in A$, $\mathcal{A} \times \mathcal{B}$ -измеримо. Предположим далее, что существует σ -конечная мера ν на G такая, что из $\nu(B) = 0$ следует, что $\nu(Bg) = 0$ для всех $g \in G$. Тогда каждая измеримая почти инвариантная относительно G функция («почти» относительно некоторой σ -конечной меры μ) эквивалентна некоторой инвариантной функции.

Доказательство. В силу сделанных предположений об измеримости, функция $\varphi(gx)$, рассматриваемая как функция двух переменных x и g , является $\mathcal{A} \times \mathcal{B}$ -измеримой. Следовательно, $\varphi(gx) - \varphi(x)$ — также $\mathcal{A} \times \mathcal{B}$ -измерима и таковым же является множество S точек (x, g) , где $\varphi(gx) \neq \varphi(x)$. Если критерий φ почти инвариантен, то каждое сечение множества S с фиксированным g имеет μ -меру нуль. По теореме Фубини (тео-

рема 3 главы 2) существует поэтому множество N μ -меры нуль такое, что для всех $x \in \mathcal{X} - N$

$$\varphi(gx) = \varphi(x) \text{ } \nu\text{-почти всюду.}$$

Без ограничения общности можно считать, что $\nu(G) = 1$. Пусть A — множество точек x , для которых

$$\int \varphi(g'x) d\nu(g') = \varphi(gx) \text{ } \nu\text{-почти всюду.}$$

Если

$$f(x, g) = \left| \int \varphi(g'x) d\nu(g') - \varphi(gx) \right|,$$

то A есть множество точек x , для которых

$$\int f(x, g) d\nu(g) = 0.$$

Поскольку этот интеграл является измеримой функцией от x , то множество A измеримо. Пусть

$$\psi(x) = \begin{cases} \int \varphi(gx) d\nu(g), & \text{если } x \in A, \\ 0, & \text{если } x \notin A. \end{cases}$$

Тогда ψ измерима и $\psi(x) = \varphi(x)$ для $x \notin N$, поскольку из равенства $\varphi(gx) = \varphi(x)$ почти всюду относительно меры ν следует, что $\int \varphi(g'x) d\nu(g') = \varphi(x)$ и что $x \in A$. Чтобы показать теперь, что ψ инвариантна, достаточно доказать, что множество A — инвариантно. Для каждой точки $x \in A$ функция $\varphi(gx)$, рассматриваемая как функция от g , постоянна, за исключением подмножества меры нуль $N_x \in G$. Тогда $\varphi(ghx)$ имеет то же самое значение и для всех $g \notin N_x h^{-1}$; последние множества согласно предположению имеют ν -меру нуль. Таким образом, $hx \in A$, чем и заканчивается доказательство.

Следствие 1. Предположим, что задача проверки гипотезы $H: \theta \in \omega$ при альтернативе $K: \theta \in \Omega \setminus \omega$ остается инвариантной относительно G , и пусть выполнены предположения теоремы 4. Тогда если φ_0 РНМ инвариантный критерий, то он будет также РНМ и в классе почти инвариантных критериев.

Доказательство. Если критерий φ почти инвариантен, то согласно теореме 4 он эквивалентен некоторому инвариантному критерию ψ . Критерии φ и ψ имеют одну и ту же функцию мощности и, следовательно, φ_0 равномерно не менее мощен, чем φ .

В приложениях семейство $\mathcal{P}$ обычно является доминированным, а μ — любая σ -конечная мера, эквивалентная $\mathcal{P}$ (которая существует согласно теореме 2 из Дополнения). Если критерий φ почти инвариантен для семейства $\mathcal{P}$, то тогда φ почти инва-

риантен и по отношению к мере μ и, следовательно, эквивалентен некоторому инвариантному критерию. Как правило, выборочное пространство $\mathcal{X}$ является n -мерным евклидовым пространством, $\mathcal{A}$ — классом всех борелевских множеств, а элементы группы G представимы в форме $y = f(x, \tau)$, где τ изменяется на множестве положительной меры в m -мерном пространстве, а векторная функция f , как функция $m + n$ переменных, измерима по Борелю. Если в качестве $\mathcal{B}$ взять класс всех борелевских множеств в m -мерном пространстве, то условия измеримости теоремы 4 выполняются.

То требование, что для всех $g \in G$ и $B \in \mathcal{B}$

$$\text{из } \nu(B) = 0 \text{ вытекает, что } \nu(Bg) = 0, \quad (13)$$

выполняется, в частности, когда

$$\nu(Bg) = \nu(B) \text{ для всех } g \in G, \quad B \in \mathcal{B}. \quad (14)$$

Существование такой *правоинвариантной меры* для большого класса групп гарантируется теорией мер Хаара. С другой стороны, обычно бывает нетрудно проверить условие (13) и непосредственно.

Пример 8. Пусть G — группа всех невырожденных преобразований n -мерного пространства. В фиксированной системе координат элементы группы G представляются невырожденными матрицами $A = (a_{ij})$, $A' = (a'_{ij})$, ... порядка $n \times n$ с групповой операцией, задаваемой матричным произведением. В качестве σ -поля $\mathcal{B}$ можно взять класс борелевских множеств в пространстве n^2 элементов матриц, а в качестве меры ν — обычную меру Лебега на $\mathcal{B}$. Рассмотрим теперь множество S матриц с $\nu(S) = 0$, и множество S^* матриц $A'A$ с $A' \in S$ и фиксированным A . Если $a = \max |a_{ij}|$, $C' = A'A$ и $C'' = A''A$, то из неравенства $|a''_{ij} - a'_{ij}| \leq \varepsilon$ для всех i, j следует, что $|c''_{ij} - c'_{ij}| \leq na\varepsilon$. Поскольку каждое множество имеет ν -меру нуль тогда и только тогда, когда оно может быть покрыто соединением прямоугольников с суммарной мерой, не превосходящей заданного $\varepsilon > 0$, то отсюда следует, что $\nu(S^*) = 0$, а это и следовало показать.

В предыдущих главах сравнение различных критериев производилось только в терминах их функций мощности (возможно взвешенных в зависимости от степени значимости соответствующих потерь).

Поскольку ограничение инвариантными критериями представляет собой отход от этой точки зрения, то интересно выяснить вопрос о том, что же дает применение инвариантности к функциям мощности. Каждый критерий, инвариантный или почти инвариантный относительно G , имеет функцию мощности, которая остается инвариантной относительно группы $\bar{G}$, индуцированной группой G в параметрическом пространстве.

Обратное, вообще говоря, неверно; пусть случайные величины X_1, X_2, X_3 независимы и нормально распределены со средним

ξ и дисперсией σ^2 . Рассмотрим гипотезу $\sigma \geq \sigma_0$. Критерий, который отвергает эту гипотезу, если

$$|X_2 - X_1| > k, \quad \text{когда } \bar{X} < 0,$$

$$|X_3 - X_2| > k, \quad \text{когда } \bar{X} \geq 0,$$

не инвариантен относительно группы G преобразований $X'_i = X_i + c$, однако его функция мощности инвариантна относительно группы $\bar{G}$.

Два свойства — почти инвариантность критерия φ и инвариантность его функции мощности — становятся эквивалентными, если еще до применения принципа инвариантности задача сведена к достаточным статистикам, распределения которых образуют ограниченно полное семейство.

Лемма 2. Пусть семейство $\mathcal{P}^T = \{P_\theta^T, \theta \in \Omega\}$ распределений T ограничено полно и пусть задача проверки гипотезы $H: \theta \in \Omega_H$ остается инвариантной при группе G преобразований T при всех θ . Тогда, для того чтобы функция мощности критерия $\psi(t)$ была инвариантной относительно группы $\bar{G}$, действующей в пространстве Ω , необходимо и достаточно, чтобы функция $\psi(t)$ была почти инвариантной при преобразованиях из G .

Доказательство. Для всех $\theta \in \Omega$ мы имеем $E_{\bar{g}\theta}\psi(T) = E_\theta\psi(gT)$. Если ψ — почти инвариантно, то $E_\theta\psi(T) = E_\theta\psi(gT)$ и, следовательно, $E_{\bar{g}\theta}\psi(T) = E_\theta\psi(T)$, так что функция мощности критерия ψ является инвариантной. Обратно, если $E_\theta\psi(T) = E_{\bar{g}\theta}\psi(T)$, то $E_\theta\psi(T) = E_\theta\psi(gT)$, и из ограниченной полноты семейства $\mathcal{P}^T$ следует, что $\psi(gt) = \psi(t)$ почти всюду относительно $\mathcal{P}^T$. Используя эту лемму, можно показать, что РНМ инвариантный критерий обладает также следующим оптимальным свойством.

Теорема 5. Пусть выполнены предположения леммы 2 и пусть функция $v(\theta)$ — максимальный инвариант относительно $\bar{G}$. Предположим, что среди критериев для гипотезы H , основанных на достаточной статистике T , существует РНМ почти инвариантный критерий, скажем $\psi_0(t)$. Тогда $\psi_0(t)$ является РНМ в классе всех критериев, основанных на первоначальных наблюдениях, функции мощности которых зависят только от $v(\theta)$.

Доказательство. Пусть $\varphi(x)$ — некоторый такой критерий и пусть $\psi(t) = E[\varphi(x) | t]$. Функция мощности критерия $\psi(t)$, будучи идентична функции мощности $\varphi(x)$, зависит тогда только от $v(\theta)$ и, следовательно, инвариантна относительно $\bar{G}$. Из леммы 2 следует, что критерий $\psi(t)$ почти инвариантен при G , и что $\psi_0(t)$ имеет мощность, равномерно не меньшую, чем $\psi(t)$, и, следовательно, чем $\varphi(x)$.

Пример 9. В задаче проверки гипотезы $\tau^2 \leq \sigma^2$ относительно дисперсий двух нормальных распределений величины $(\bar{X}, \bar{Y}, S_X^2, S_Y^2)$ образуют пол-

ное множество достаточных статистик. Как было показано в примере 6, при подходящим образом подобранной группе преобразований существует РНМ инвариантный критерий, который отвергает рассматриваемую гипотезу, если $S_Y^2/S_X^2 > C_0$. Поскольку в настоящем случае из почти инвариантности этого критерия относительно G вытекает, что он эквивалентен некоторому инвариантному (задача 13), то применима теорема 5 с $v(\theta) = \Delta = \tau^2/\sigma^2$. Этот критерий является, следовательно, РНМ среди всех критериев, функции мощности которых зависят только от Δ .

6. Несмещенность и инвариантность

Принципы несмещенности и инвариантности дополняют друг друга в том отношении, что один дает удачные результаты там, где применение другого не приносит успеха. Так, например, для сравнения двух биномиальных или пуассоновских распределений имеется РНМ несмещенный критерий, тогда как соображения инвариантности здесь неприменимы. РНМ несмещенный критерий существует также при проверке гипотезы $\sigma = \sigma_0$ при альтернативе $\sigma \neq \sigma_0$ для нормально распределенных наблюдений, тогда как применение принципа инвариантности не дает здесь значительного продвижения. И обратно, существуют РНМ инвариантные критерии для проверки гипотез о специальных значениях одного или нескольких параметров (что будет рассмотрено в гл. 7), тогда как РНМ несмещенных критериев не существует. Имеются также гипотезы (как, например, односторонняя гипотеза $\xi/\sigma < \theta_0$ для одномерного нормального распределения или гипотеза $\rho \leq \rho_0$ для двумерного (задача 11) с $\theta_0, \rho_0 \neq 0$), где РНМ инвариантный критерий существует, тогда как показать существование РНМ несмещенного критерия методами главы 5 не удается, и вопрос до сих пор остается открытым.

С другой стороны, в некоторых задачах с одинаковым успехом могут применяться оба принципа. В число таких проблем входят задачи проверки гипотез Стьюдента $\xi \leq \xi_0$ и $\xi = \xi_0$ относительно среднего нормального распределения, и соответствующие задачи проверки гипотез $\eta - \xi \leq \Delta_0$ и $\eta - \xi = \Delta_0$ для случая двух выборок в предположении равенства дисперсий в обоих выборках. Другим примером может служить задача проверки гипотез $\sigma^2 \geq \sigma_0^2$ и $\tau^2/\sigma^2 \geq \Delta_0$ относительно дисперсии одного или двух нормальных распределений. Еще один такой пример дает проверка гипотезы о независимости $\rho = 0$ в двумерном нормальном распределении (задача 11). Во всех перечисленных примерах обе оптимальные процедуры совпадают. Мы сейчас покажем, что это совпадение не случайно и имеет место в тех случаях, когда РНМ инвариантный критерий является РНМ также и среди всех почти инвариантных критериев и когда РНМ несмещенный критерий единствен. В этом смысле принципы несмещенности и почти инвариантности согласованы.

Теорема 6. Пусть для данной задачи проверки гипотезы существует единственный (с точностью до множеств меры 0) РНМ несмещенный критерий φ^* и, кроме того, существует РНМ критерий, почти инвариантный относительно некоторой группы преобразований G . Тогда последний также является единственным (с точностью до множеств меры 0), и оба критерия совпадают почти всюду.

Доказательство. Если $U(\alpha)$ есть класс несмещенных критериев уровня α и $g \in G$, то $\varphi \in U(\alpha)$ тогда и только тогда, когда $\varphi g \in U(\alpha)^*$. Обозначая функцию мощности критерия φ через $\beta_\varphi(\theta)$, мы тем самым получим

$$\begin{aligned}\beta_{\varphi^*g}(\theta) &= \beta_{\varphi^*}(\bar{g}\theta) = \sup_{\varphi \in U(\alpha)} \beta_\varphi(\bar{g}\theta) = \\ &= \sup_{\varphi \in U(\alpha)} \beta_{\varphi g}(\theta) = \sup_{\varphi g \in U(\alpha)} \beta_{\varphi g}(\theta) = \beta_{\varphi^*}(\theta).\end{aligned}$$

Следовательно, φ^* и φ^*g имеют одну и ту же функцию мощности, и в силу предположенной единственности критерий φ^* почти инвариантен. Поэтому, если φ' есть РНМ почти инвариантный критерий, то $\beta_{\varphi'}(\theta) \geq \beta_{\varphi^*}(\theta)$ для всех θ . С другой стороны, φ' — несмещенный критерий, как это следует из сравнения его с инвариантным критерием $\varphi(x) \equiv \alpha$; поэтому $\beta_{\varphi'}(\theta) \leq \beta_{\varphi^*}(\theta)$ для всех θ . Так как критерии φ' и φ^* имеют одну и ту же функцию мощности, то они совпадают почти всюду в силу единственности φ^* , что и требовалось доказать.

Эта теорема дает возможность сделать определенные выводы для некоторых критериев, рассмотренных в главе 5. В теореме 3 главы 4 существование РНМ несмещенных критериев было установлено для одно- и двусторонних гипотез относительно значений параметра θ в экспоненциальном семействе распределений (10) главы 4. Для этого семейства статистики (U, T) достаточны и полны и поэтому в терминах этих статистик РНМ несмещенный критерий единствен. Удобные явные выражения для некоторых из критериев, рассмотренных в главе 5, можно было бы получить, заметив, что в тех случаях, когда РНМ почти инвариантный критерий существует, он, согласно теореме 6, необходимо будет также и РНМ несмещенным. Этим доказывается, например, что критерии, построенные в примерах 5 и 6 настоящей главы, являются РНМ несмещенными.

Принципы несмещенности и инвариантности могут быть применены дополнительно один к другому в случаях, где каждый из них в отдельности не приводит к полному решению, а их совместное применение приводит к нему. В качестве примера рассмот-

*) φg обозначает критическую функцию, равную $\varphi(gx)$ для каждого значения x .

рим выборку $X_1, \dots, X_n$ из $N(\xi, \sigma^2)$ и гипотезу $H: \xi/\sigma = \theta_0 \neq 0$ при двусторонней альтернативе: $\xi/\sigma \neq \theta_0$. Здесь достаточность и инвариантность сводят задачу к рассмотрению статистики $t = \sqrt{n} \bar{x} / \sqrt{\sum (x_i - \bar{x})^2 / (n-1)}$. Эта статистика имеет нецентральное t -распределение с параметром нецентральности $\sigma = \sqrt{n} \xi / \sigma$ и $n-1$ степенью свободы. Можно показать, что это семейство, зависящее от параметра δ , относится к строгому типу Пойа и, следовательно, в частности, к 3-му типу *). Как и в задаче 25 главы 3, можно показать, что среди всех критериев, основанных на статистике t , существует РНМ несмещенный критерий с областью принятия H вида $C_1 \leq t \leq C_2$, где C_1, C_2 определяются из условия

$$P_{\delta_0} \{C_1 \leq t \leq C_2\} = 1 - \alpha \quad \text{и} \quad \left. \frac{\partial P_{\delta} \{C_1 \leq t \leq C_2\}}{\partial \delta} \right|_{\delta=\delta_0} = 0.$$

В терминах исходных наблюдений построенный критерий будет РНМ среди всех несмещенных и инвариантных критериев, но вопрос о том, будет ли этот критерий РНМ несмещенным без ограничений инвариантности, остается открытым.

Другим случаем, в котором совместное применение инвариантности и несмещенности дает, по-видимому, многообещающий подход, является так называемая *проблема Беренка* — Фишера **). Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из двух нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ соответственно. Проблема состоит в том, чтобы проверить гипотезу $H: \eta \leq \xi$ (или $\eta = \xi$) без предположения равенства дисперсий σ^2 и τ^2 . Множество достаточных статистик для $(\xi, \eta, \sigma, \tau)$ есть в этом случае $(\bar{X}, \bar{Y}, S_X^2, S_Y^2)$, где $S_X^2 = \sum (X_i - \bar{X})^2$ и $S_Y^2 = \sum (Y_j - \bar{Y})^2$. Добавление одной и той же постоянной к $\bar{X}$ и $\bar{Y}$ редуцирует задачу к статистикам $\bar{Y} - \bar{X}, S_X^2, S_Y^2$, а умножение всех переменных на одну и ту же положительную константу приводит к рассмотрению статистик

$$\frac{(\bar{Y} - \bar{X})}{\sqrt{S_X^2 + S_Y^2}} \quad \text{и} \quad \frac{S_Y^2}{S_X^2}.$$

*) Karlin, Decision theory for Pólya type distributions. Case of two actions, I, Proc. Third Berkeley Symposium on Mathematical Statistics and Probability I, Berkley, Univ. Calif. Press, 115—129.

**) Решение проблемы освещается, с одной стороны, в книге Ю. В. Линника, Статистические задачи с мешающими параметрами, М., Наука, 1966, с другой стороны, в работах Пагуровой, О сравнении средних значений в двух нормальных выборках, Теор. вер. и ее примен. 13 (1968), № 3, 561—568 и Критерий сравнения средних значений по двум нормальным выборкам, вып. 5, М., Изд. ВЦ АН СССР, 1968. См. также Большеев Л. Н., Смирнов Н. В., Таблицы математической статистики, М., Наука, 1965. — Прим. ред

Можно ожидать, что каждая разумная критическая область представляется в виде

$$\frac{\bar{Y} - \bar{X}}{\sqrt{S_X^2 + S_Y^2}} \geq g\left(\frac{S_Y^2}{S_X^2}\right) \quad (15)$$

с определенным образом подобранной функцией g . Если от этого критерия потребовать так же несмещенности, то вероятность множества (15) должна быть равна α , когда $\eta = \xi$, для всех значений τ/σ . Но, существует ли функция g с таким свойством, до сих пор не выяснено. Однако имеется приближенное решение, которое протабулировано *) и которое дает для практических целей удовлетворительные результаты.

Каждый РНМ несмещенный критерий обладает важным свойством допустимости (задача 1 главы 4), так что в этом случае не может существовать другого критерия равномерно не менее мощного, а для некоторых альтернатив даже более мощного, чем заданный. Аналогичное свойство для РНМ инвариантных критериев может и не выполняться, что показывает следующий пример.

Пример 10 **). Пусть случайные векторы (X_{11}, X_{12}) и (X_{21}, X_{22}) распределены нормально с нулевыми средними и матрицами ковариаций

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \text{ и } \begin{pmatrix} \Delta\sigma_1^2 & \Delta\rho\sigma_1\sigma_2 \\ \Delta\rho\sigma_1\sigma_2 & \Delta\sigma_2^2 \end{pmatrix}.$$

Предположим, что эти матрицы не вырождены, т. е. что $|\rho| \neq 1$; в других же отношениях σ_1 , σ_2 , ρ и Δ остаются неизвестными. Задача различения гипотез $\Delta = 1$ и $\Delta > 1$ инвариантна относительно группы G всех невырожденных преобразований

$$\begin{aligned} X'_{i1} &= a_{11}X_{i1} + a_{12}X_{i2}, \\ X'_i &= a_{21}X_{i1} + a_{22}X_{i2} \end{aligned} \quad (i = 1, 2).$$

Поскольку вероятность того, что $X_{11}X_{22} = X_{12}X_{21}$, равна нулю, матрицы второго порядка (X_{ij}) с вероятностью единица не вырождены, и поэтому выборочное пространство можно считать состоящим только из невырожденных матриц. Для любых двух выборочных точек $Z = (X_{ij})$ и $Z' = (X'_{ij})$ существует невырожденное линейное преобразование A такое, что $Z' = AZ$. Поэтому относительно G не существует инвариантных критериев, за исключением тривиаль-

*) Welch, The generalization of Student's problem when several different population variances are involved, *Biometrika* 34 (1947), 28—35; Aspin, Tables for use in comparisons whose accuracy involves two variances, *Biometrika* 36 (1949), 290—296. См. также Chernoff, Asymptotic studentization in testing of hypotheses, *Ann. Math. Stat.* 20 (1949), 268—278 и Wallace, Asymptotic approximations to distributions, *Ann. Math. Stat.* 29 (1958), 635—654.

**) Этот пример сообщен мне проф. Стейном.

ного критерия $\varphi \equiv \alpha$ уровня α , мощность которого $\beta(\Delta) \equiv \alpha$. С другой стороны, X_{11} и X_{21} независимы и распределены нормально $N(0, \sigma_1^2)$ и $N(0, \Delta\sigma_1^2)$. Основываясь на этих наблюдениях, можно построить РНМ критерий для различения гипотез $\Delta = 1$ и $\Delta > 1$ с критической областью $X_{21}^2/X_{11}^2 > C$ (задача 33 главы 3). Функция мощности этого критерия строго возрастает по Δ и, следовательно, больше α для всех $\Delta > 1$.

Тем самым свойство допустимости оптимальных инвариантных критериев заведомо не выполняется автоматически и его надо устанавливать в каждом отдельном случае. Пусть функция $\delta = v(\theta)$ — максимальный инвариант относительно группы преобразований $\bar{G}$. Предположим, для определенности, что проверке подлежит гипотеза $\delta \leq \delta_0$. Чтобы проверить допустимость критерия φ_0 с уровнем значимости α , достаточно показать для некоторого подмножества альтернатив Ω' , что для любого критерия φ уровня α из условия $E_\theta \varphi(X) \geq E_\theta \varphi_0(X)$ для всех $\theta \in \Omega'$ вытекает равенство $E_\theta \varphi(X) = E_\theta \varphi_0(X)$ для всех θ . Типичные доказательства допустимости можно разделить на три категории, соответственно тому, как они устанавливают указанное свойство:

(а) локально, т. е. для всех θ , удовлетворяющих условию $\delta_0 < v(\theta) < \delta_1$ для некоторого $\delta_1 > \delta_0$; (b) для всех достаточно далеких альтернатив, т. е. для всех альтернатив, удовлетворяющих условию $v(\theta) > \delta_2$ для некоторого $\delta_2 > \delta_0$; (с) для всех альтернатив с любым заданным расстоянием δ , т. е. для альтернатив, удовлетворяющих условию $v(\theta) = \delta$. Доказательства типа (а) или (b) не совсем удовлетворительны, поскольку они не исключают возможности существования критерия с мощностью, лучшей для всех практически важных альтернатив и худшей только тогда, когда оба критерия имеют мощность, очень близкую к 1, или когда альтернативы так близки к гипотезам, что значения их функций мощности становятся несущественными.

В качестве примера рассмотрим РНМ несмещенный критерий φ_1 (из теоремы 3 главы 4) для проверки гипотезы $H: \theta \leq \theta_0$ при альтернативе $\theta > \theta_0$ и при наличии мешающих параметров θ . Чтобы показать, что этот критерий локально допустим, предположим, что φ — некоторый другой критерий для H , имеющий уровень значимости α . Если $E_{\theta_0} \varphi(X) < \alpha$ для некоторого θ , тогда по непрерывности найдется такое $\theta_1 > \theta_0$, что для всех $\theta_0 < \theta < \theta_1$ $E_\theta \varphi(X) < \alpha < E_{\theta_0} \varphi_1(X)$. Отсюда следует, что локально мощность критерия φ не больше, чем мощность критерия φ_1 . С другой стороны, если $E_{\theta_0} \varphi(X) = \alpha$ для всех θ , то $E_\theta \varphi(X) \leq E_{\theta_0} \varphi_1(X)$ для всех $\theta > \theta_0$ и всех θ , поскольку при доказательстве теоремы 3 было показано, что критерий φ_1 является РНМ среди всех критериев, подобных на границе. Эта аргументация, однако, не исключает возможности существования критерия, который смещен вблизи H , но который в то же

время является равномерно более мощным, чем φ_0 при всех альтернативах, находящихся от H на расстоянии, большем некоторого числа.

Свойство допустимости относительно удаленных альтернатив было установлено в задачах проверки некоторых гипотез о параметрах экспоненциальных семейств *) распределений, а также для альтернатив на любом расстоянии — в некоторых задачах **) о параметре сдвига. К числу последних относится задача проверки гипотезы $\xi/\sigma \leq \theta_0$ при альтернативе $\xi/\sigma = \theta$ для нормального распределения.

7. Ранговые критерии ***)

Одной из основных задач статистики является задача проверки по двум выборкам гипотезы о равенстве их распределений. Типичным примером служит задача сравнения некоторого способа обработки со стандартным (контрольным), где проверяется гипотеза об отсутствии преимуществ этого способа при альтернативе его более благотворного действия. Этот пример рассматривался в главах 4 и 5 в предположениях нормальности; соответствующий критерий основывался на t -статистике Стьюдента. Было также показано, что в тех случаях, когда предположение о приближенной нормальности допустимо, но не может быть проверено, приходят к замене t -критерия его аналогом (основанным на перестановках), который, в свою очередь, может быть аппроксимирован первоначальным t -критерием.

Ниже мы будем рассматривать эту проблему, отказавшись по крайней мере на время от всяких предположений относительно даже приближенной формы рассматриваемых распределений, предполагая только, что они непрерывны. Результаты наблюдений в таком случае образуют две выборки $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ из распределений с непрерывными функциями распределения F и G . Задача состоит в том, чтобы проверить гипотезу

$$H_1: G = F.$$

Если предположить, что эффект обработки аддитивен, то альтернативы имеют вид $G(y) = F(y - \Delta)$. Мы рассмотрим здесь более общий случай, в котором эффект обработки может зависеть

*) Birnbaum, Characterizations of complete classes of tests of some multiparameter hypotheses with applications to likelihood ratio test, Ann. Math. Stat. 26 (1955), 21—36 и Stein, The admissibility of Hotelling's T^2 -test, Ann. Math. Stat. 27 (1956), 616—623.

**) Lehmann and Stein, The admissibility of certain invariant statistical tests involving a translation parameter, Ann. Math. Stat. 24 (1953), 473—479.

***). См. подробнее Гаек, Шидак, Теория ранговых критериев, пер. с англ., М., Наука, 1971. — Прим. ред.

и от значения y (так что Δ есть неотрицательная функция от y). Тогда гипотеза H_1 будет проверяться при односторонней альтернативе, что величины Y стохастически больше величин X , т. е. при альтернативе

$$K_1: G(z) \leq F(z) \text{ для всех } z \text{ и } G \neq F.$$

Альтернативный эксперимент, который проводится для выяснения эффективности обработки, состоит в сравнении N пар предметов, группирующихся по парам так, чтобы исключить расхождения, не являющиеся следствием обработки. Один член каждой пары выбирается случайным образом для обработки, тогда как другой служит для контроля. Если предположения (см. раздел 4 главы 5) о нормальности распределений отбрасываются и пары предметов рассматриваются как один элемент выборки, то предполагается, что распределения независимых наблюдений $(X_1, Y_1), \dots, (X_N, Y_N)$ описываются двумерной непрерывной функцией распределения F . Гипотеза отсутствия эффекта обработки эквивалентна тогда предположению, что функция распределения F симметрична относительно линии $y = x$

$$H_2: F(x, y) = F(y, x).$$

Другая основная проблема, которая возникает во многих случаях, касается зависимости или независимости двух случайных величин. В частности, если $(X_1, Y_1), \dots, (X_N, Y_N)$ — выборка из двумерного распределения F , то может представить интерес проверка гипотезы

$$H_3: F(x, y) = G_1(x) H_2(y),$$

т. е. гипотезы, что X и Y независимы; для случая нормального распределения эта задача рассматривалась в разделе 9 главы 5. В качестве интересной альтернативы можно рассмотреть, к примеру, гипотезу о том, что X и Y положительно зависимы (см. раздел 11 главы 5). Альтернативная формулировка возникает, когда значения x не являются случайными, а выбираются заранее, до эксперимента. Пусть выбраны значения $x_1 < \dots < x_N$ и пусть F_i обозначает распределение Y при заданном x_i ; предполагается, что Y_i независимы и их функции распределения $F_1, \dots, F_N$ непрерывны. Гипотеза независимости Y от x тогда заключается в том, что

$$H_4: F_1 = \dots = F_N,$$

в то время как альтернативная гипотеза положительной зависимости Y_i состоит в стохастическом их возрастании с ростом i .

В этих и других подобных задачах инвариантность редуцирует исходные данные так, что истинные значения наблюдений становятся несущественными и сохраняются только некоторые отношения порядка между различными группами переменных.

Тем не менее оказывается возможным таким образом проверить исследуемые гипотезы, а получаемые этим способом критерии часто имеют мощность, близкую к мощности стандартных нормальных критериев. Мы сейчас проведем эту редукцию для четырех рассмотренных выше проблем.

В случае двух выборок задача проверки гипотезы H_1 при альтернативе K_1 остается инвариантной относительно группы G всех преобразований

$$x'_i = f(x_i), \quad y'_j = f(y_j) \quad (i = 1, \dots, m; j = 1, \dots, n)$$

таких, что функции f непрерывны и строго возрастают. Это утверждение следует из того факта, что указанные преобразования сохраняют непрерывность распределений и свойства величин оставаться одинаково распределенными или одной быть стохастически больше другой. Как видно из примера 3 (в других обозначениях), максимальным инвариантом относительно группы G является множество рангов

$$(R'; S') = (R'_1, \dots, R'_m; S'_1, \dots, S'_n)$$

величин $X_1, \dots, X_m; Y_1, \dots, Y_n$, рассматриваемых как элементы одной выборки. Поскольку распределение $(R'_1, \dots, R'_m; S'_1, \dots, S'_n)$ симметрично относительно первых m и последних n переменных для всех распределений F и G , то множеством достаточных статистик для (R', S') является множество рангов величин X и Y без учета индексов у X и Y . Эти статистики можно представить в виде упорядоченных рангов величин X и Y :

$$R_1 < \dots < R_m \quad \text{и} \quad S_1 < \dots < S_n.$$

А поскольку каждое из этих множеств определяется другим, то достаточно рассматривать только одну из этих статистик. Тем самым каждый инвариантный критерий является *ранговым критерием*, т. е. зависит только от рангов наблюдений, например, от $(S_1, \dots, S_n)$.

Чтобы получить подобную же редукцию для гипотезы H_2 , сделаем сначала преобразование $Z_i = Y_i - X_i$, $W_i = X_i + Y_i$. Пары величин (Z_i, W_i) снова образуют выборку из непрерывного двумерного распределения. Тогда проверяемая гипотеза состоит в том, что это распределение симметрично относительно оси w , в то время как при альтернативе это распределение сдвинуто в положительном направлении оси z . Задача проверки этих гипотез не изменится, если все w подвергнуть одному и тому же взаимно однозначному преобразованию $w'_i = g(w_i)$, где g имеет самое большее конечное число точек разрыва. Максимальным инвариантом относительно этой группы [см. задачу 2 (II)] будет функция $(Z_1, \dots, Z_N)$.

Случайные величины Z_i имеют непрерывную одномерную функцию распределения D , и гипотеза ее симметрии относительно начала координат

$$H'_2: D(z) + D(-z) = 1 \quad \text{для всех } z$$

проверяется при альтернативе, что распределение сдвинуто в положительном направлении оси z . Эта задача инвариантна относительно группы G всех преобразований

$$z'_i = f(z_i) \quad (i = 1, \dots, N)$$

таких, что f непрерывна, нечетна и строго возрастает. Пусть $z_{i_1}, \dots, z_{i_m} < 0 < z_{j_1}, \dots, z_{j_n}$, где $i_1 < \dots < i_m$ и $j_1 < \dots < j_n$. Обозначим $s'_1, \dots, s'_n$ ранги величин $z_{i_1}, \dots, z_{i_n}$ среди абсолютных значений $|z_1|, \dots, |z_N|$ и $r'_1, \dots, r'_m$ — ранги величин $|z_{i_1}|, \dots, |z_{i_m}|$ среди $|z_1|, \dots, |z_N|$. Преобразование f сохраняет знак каждого наблюдения и, следовательно, в частности, числа m и n . Поскольку f непрерывная, строго возрастающая функция от $|z|$, она оставляет инвариантным порядок абсолютных величин и, тем самым, ранги r'_i и s'_j . Покажем, что эти функции являются максимальными инвариантами. Пусть $(z_1, \dots, z_N)$ и $(z'_1, \dots, z'_N)$ — два множества точек с $m' = m$, $n' = n$ и теми же самыми r'_i и s'_j . Существует непрерывная, строго возрастающая функция, определенная на положительной оси, такая, что $|z'_i| = f(|z_i|)$ и $f(0) = 0$. Если f определить для отрицательных z посредством соотношения $f(-z) = -f(z)$, то такая функция будет принадлежать G , и тогда $z'_i = f(z_i)$ для всех i , что и надо было показать. Как и в предыдущей задаче, достаточность позволяет совершить дальнейшую редукцию данных к упорядоченным рангам $r_1 < \dots < r_m$ и $s_1 < \dots < s_n$. Эта редукция сохраняет информацию о рангах абсолютных значений независимо от того, являются ли результаты наблюдений положительными или отрицательными, но она не сохраняет информации о том, с каким значением, положительным или отрицательным, связываются ранги.

Ситуация очень похожа и в случае гипотез H_3 и H_4 . Задача проверки независимости в двумерном распределении при альтернативной гипотезе о положительной зависимости не изменяется, если X_i и Y_i подвергнуть преобразованиям $X'_i = f(X_i)$, $Y'_i = g(Y_i)$ таким, что f и g непрерывны и строго возрастают. Относительно этого преобразования максимальными инвариантами являются ранги $(R'_1, \dots, R'_N)$ величин $(X_1, \dots, X_N)$ среди X и ранги $(S'_1, \dots, S'_N)$ величин $(Y_1, \dots, Y_N)$ среди Y . Совместное распределение величин $(R'_i, S'_i), \dots, (R'_N, S'_N)$ симметрично относительно этих N пар, каковы бы ни были распределения (X, Y) . Отсюда следует, что достаточной статистикой является $(S_1, \dots, S_N)$, где

$(1, S_1), \dots, (N, S_N)$ — перестановка величин $(R'_1, S'_1), \dots, (R'_N, S'_N)$ и где, следовательно, S_i есть ранг величины Y_i , связанной с i -м наименьшим значением X .

Гипотеза H_4 о том, что $Y_1, \dots, Y_n$ одинаково распределены, проверяется при альтернативе K_4 , состоящей в том, что Y_i стохастически возрастают с ростом i . Эта проблема инвариантна относительно группы преобразований $y'_i = f(y_i)$, где функция f непрерывна и строго убывает. Максимальным инвариантом здесь является множество рангов $S_1, \dots, S_N$ величин $Y_1, \dots, Y_N$.

Некоторые инвариантные критерии для проверки гипотез H_1 и H_2 будут рассмотрены в следующих двух разделах. Соответствующие результаты, относящиеся к гипотезам H_3 и H_4 , даны в задачах 39—41.

8. Задача сравнения двух выборок

Принцип инвариантности для случая двух выборок редуцирует задачу проверки гипотезы $H: G = F$ при односторонней альтернативе K , состоящей в том, что величины Y стохастически больше величин X , к изучению рангов $S_1 < \dots < S_n$ величин Y . Задание величин S_i эквивалентно указанию для составной выборки из $N = m + n$ элементов наименьшего значения, следующего за ним по величине значения и т. д., независимо от того, являются ими x или y . Поскольку в каждом множестве из N результатов наблюдений y_k занимают n положений и так как, в предположении $G = F$, все $\binom{N}{n}$ возможных выборов n положений y_k равновероятны, то совместное распределение S_i при гипотезе H равно

$$P\{S_1 = s_1, \dots, S_n = s_n\} = \frac{1}{\binom{N}{n}}, \quad (16)$$

для каждого множества $1 \leq s_1 < s_2 < \dots < s_n \leq N$. Поэтому для любого рангового критерия с уровнем $\alpha = k / \binom{N}{n}$ критическая область состоит в точности из k точек $(s_1, \dots, s_n)$.

В задаче проверки гипотезы H при альтернативе K не существует РНМ рангового критерия и, следовательно, не существует РНМ инвариантного критерия. Это следует, например, из рассмотрения двух стандартных критериев, применяемых в этой задаче, поскольку каждый из них является наиболее мощным среди всех ранговых критериев для случая некоторых альтернатив. Критерии, о которых идет речь, имеют критическую область вида

$$h(s_1) + \dots + h(s_n) > C. \quad (17)$$

Один из них, *двухвыборочный критерий Вилкоксона* *), получается, если в (17) положить $h(s) = s$, так что гипотеза H отвергается, когда сумма рангов величин y достаточно велика. Ниже мы покажем, что для достаточно малых Δ этот критерий является наиболее мощным при альтернативе: F есть логистическое распределение $F(x) = 1/(1 + e^{-x})$, а $G(y) = F(y - \Delta)$. Другой критерий, *критерий Фишера — Ийтса*, получим, если в (17) положить $h(s) = E(V^{(s)})$, где $V^{(1)} < \dots < V^{(N)}$ — упорядоченная выборка объема N из стандартного нормального распределения **). Этот критерий является наиболее мощным при проверке гипотез о том, что распределения F и G нормальны с общей дисперсией и средними ξ и $\eta = \xi + \Delta$, когда Δ достаточно мало.

Чтобы доказать, что эти критерии обладают указанными свойствами, необходимо знать распределение $(S_1, \dots, S_n)$ при каждой из гипотез. Если распределения F и G имеют плотность f и g такие, что f всегда положительна, когда положительно g , то совместное распределение величин S_i задается формулой

$$P\{S_1 = s_1, \dots, S_n = s_n\} = \frac{E \left[\frac{g(V^{(s_1)})}{f(V^{(s_1)})} \dots \frac{g(V^{(s_n)})}{f(V^{(s_n)})} \right]}{\binom{N}{n}}, \quad (18)$$

где $V^{(1)} < \dots < V^{(N)}$ — упорядоченная выборка объема N из распределения F (см. задачу 22). Рассмотрим, в частности, альтернативу сдвига

$$g(y) = f(y - \Delta)$$

и обратимся к задаче максимизации функции мощности для малых значений Δ . Предположим, что плотность f дифференцируема и что вероятность (18), являющаяся функцией от Δ , можно дифференцировать под знаком математического ожидания.

*) По поводу таблиц для этого критерия см. задачу 21 в главе 4. См. также по поводу ранговых критериев и соответствующих таблиц Болев Л. Н., Смирнов Н. В., Таблицы математической статистики, М., Наука, 1965.

**) Таблицы, необходимые для работы с порядковыми статистиками, для случая нормального распределения, даны в «Biometrika Tables for Statisticians», 1, Cambridge Univ. Press, 1954, Table 28 (с 3 десятичными знаками для $N \leq 20$ и 2 знаками для $N \leq 50$) и в работе Teichrow, Tables of expected values of order statistics and products of order statistics..., Ann. Math. Stat. 27 (1956), 410—426 (с 10 десятичными знаками для $N \leq 20$). См. также Д. Тейкроу, Таблицы первых моментов порядковых статистик в выборках из совокупностей с нормальным распределением. — Сб.: Сархан и Гринберг, Введение в теорию порядковых статистик, пер. с англ., М., Статистика, 1970.

Производная вероятности (18) в точке $\Delta = 0$ равна

$$\begin{aligned} \frac{\partial}{\partial \Delta} P_{\Delta} \{S_1 = s_1, \dots, S_n = s_n\} |_{\Delta=0} = \\ = \frac{-E \left[\frac{f'(V^{(s_1)})}{f(V^{(s_1)})} + \dots + \frac{f'(V^{(s_n)})}{f(V^{(s_n)})} \right]}{\binom{N}{n}}. \end{aligned}$$

Поскольку при гипотезе H вероятность каждого набора рангов дается формулой (16), то из расширенной формы леммы Неймана — Пирсона (теорема 5 главы 3) следует, что производная функции мощности в точке $\Delta = 0$ максимизируется, если критическая область имеет следующий вид:

$$-\sum_{i=1}^n E \left[\frac{f'(V^{(s_i)})}{f(V^{(s_i)})} \right] > C. \quad (19)$$

Этот же критерий максимизирует и самую мощность для достаточно малых Δ . Чтобы это показать, обозначим s ранговую точку $(s_1, \dots, s_n)$ и $s^{(j)}$ ранговую точку, на которой достигается j -е наибольшее значение левой стороны неравенства (19). Если $\alpha = \frac{k}{\binom{N}{n}}$, то мощность этого критерия равна

$$\beta(\Delta) = \sum_{j=1}^k P_{\Delta}(s^{(j)}) = \sum_{j=1}^k \left[\frac{1}{\binom{N}{n}} + \Delta \frac{\partial}{\partial \Delta} P_{\Delta}(s^{(j)}) |_{\Delta=0} + \dots \right].$$

Поскольку существует только конечное число точек s , то для каждого j существует число $\Delta_j > 0$ такое, что точка $s^{(j)}$ также дает j -е наибольшее значение величине $P_{\Delta}(s)$ для всех $\Delta < \Delta_j$. Когда Δ меньше наименьшего из чисел Δ_j , $j = 1, \dots, \binom{N}{n}$, то этот критерий также максимизирует $\beta(\Delta)$.

Если $f(x)$ — плотность нормального распределения $N(\xi, \sigma^2)$, то

$$-\frac{f'(x)}{f(x)} = -\frac{d}{dx} \log f(x) = \frac{x - \xi}{\sigma^2},$$

и левая часть равенства (19) превращается в

$$\sum E \left(\frac{V^{(s_i)} - \xi}{\sigma^2} \right) = \frac{1}{\sigma} \sum E(W^{(s_i)}),$$

где $W^{(1)} < \dots < W^{(N)}$ есть упорядоченная выборка из $N(0, 1)$. Критерием, максимизирующим мощность при этих альтернати-

вах (для достаточно малых Δ), является, следовательно, критерий Фишера — Иэйтса.

В случае логистического распределения

$$F(x) = \frac{1}{(1 + e^{-x})}, \quad f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$$

и, следовательно,

$$-\frac{f'(x)}{f(x)} = 2F(x) - 1.$$

Поэтому локально наиболее мощный ранговый критерий тем самым отвергает гипотезу, когда $\sum E[F(V^{(s_i)})] > C$. Если V имеет распределение F и $0 \leq y \leq 1$, то

$$P\{F(V) \leq y\} = P\{V \leq F^{-1}(y)\} = F[F^{-1}(y)] = y,$$

так что случайная величина $U = F(V)$ равномерно распределена на $(0,1)$ *). Критическая область может быть записана в виде $\sum E(U^{(s_i)}) > C$, где $U^{(1)} < \dots < U^{(N)}$ — упорядоченная выборка объема N из равномерного распределения $R(0,1)$. Поскольку $E(U^{(s_i)}) = s_i/(N+1)$, то этот критерий есть не что иное, как критерий Вилкоксона.

И критерий Фишера — Иэйтса, и критерий Вилкоксона являются несмещенными при односторонних альтернативах K . В самом деле, пусть ϕ — критическая функция любого критерия, определенного формулой (17) с неубывающей функцией h . Тогда ϕ не убывает по y , и вероятность отвергнуть гипотезу равна α для всех $F = G$. Из леммы 2 главы 5 следует, что этот критерий является несмещенным для всех альтернатив из K .

Из свойств несмещенности этих критериев вытекает, что наиболее мощные инвариантные критерии в двух рассмотренных примерах являются также наиболее мощными при соответствующих альтернативах и среди всех инвариантных и несмещенных критериев. Следовательно, РНМ критерий может не существовать даже, если, помимо ограничения инвариантности, наложить еще условие несмещенности. Аналогично и применение одного только принципа несмещенности не приводит к решению, что мы уже видели при обсуждении критериев, основанных на перестановках в разделе 8 главы 5. Отказ же от этих принципов (от одного или от обоих сразу) оставляет проблему не только без решения, но даже и без формулировки. Одна из возможных формулировок (строгость) будет рассматриваться в гл. 8. Однако отыскание наиболее строгого критерия в задаче двух выборов остается открытой проблемой.

*) Это преобразование, переводящее случайную величину с непрерывной функцией распределения F в равномерно распределенную, известно под названием *вероятностного интегрального преобразования*.

Хотя оптимальные свойства не установлены ни для одного из двух выборочных критериев, оба упомянутых выше критерия представляются довольно удовлетворительными с практической точки зрения (точно так же, как и другой такой критерий, критерий Ван-дер-Вардена *) с критической областью (17), в которой $h(s) = \Phi^{-1}(s/N + 1)$, где Φ — функция распределения стандартного нормального распределения). Даже тогда, когда F и G нормальны с общей дисперсией, эти критерии имеют мощность, близкую к мощности t -критерия.

Чтобы получить численное сравнение этих критериев, предположим, что обе выборки имеют одинаковый объем, и рассмотрим отношение n^*/n числа наблюдений, требуемых двумя критериями для получения одной и той же мощности β при одних и тех же альтернативах. Пусть $m = n$ и $m^* = n^* = g(n)$ — объемы выборки, требуемые одним из ранговых критериев и t -критерием соответственно, и предположим (что верно для рассматриваемых критериев), что отношение n^*/n стремится к пределу e при $n \rightarrow \infty$, не зависящему от α и β . Тогда e называется *асимптотической эффективностью* рангового критерия по отношению к t -критерию. Поэтому, если, в частности, $e = 1/2$, то ранговый критерий требует примерно в два раза больше наблюдений для достижения той же мощности, что дает t -критерий.

В частном случае критерия Вилкоксона **) e оказывается равным $3/\pi \sim 0,95$, если F и G — нормальные распределения с одинаковой дисперсией. В случае, когда F и G не предполагаются нормальными, но отличаются только сдвигом, эффективность e зависит от вида распределений. Эта эффективность всегда $\geq 0,864$, но может превысить 1 и даже быть бесконечной. Для критерия Фишера — Иэйтса сравнение оказывается еще более благоприятным. Его асимптотическая эффективность по отношению к t -критерию всегда ≥ 1 , когда F и G отличаются только сдвигом, и равна 1 в частном случае нормального распределения F . Тот же результат справедлив и для критерия Ван-дер-Вар-

*) Таблицы, упрощающие применение этого критерия, даны Ван-дер-Варденом и Нивергельтом: van der Waerden B. L., Nievergelt E., Tables for Comparing Two Samples by X-Test and Sign Test, Berlin, Springer Verlag, 1956. О свойствах критерия Ван-дер-Вардена см. также в книге: Б. Л. Ван-дер-Варден, Математическая статистика, пер. с нем., М., ИЛ, 1960.

**) Эти и ряд других результатов об эффективности приведены, например, в работах: Hodges and Lehmann, The efficiency of some nonparametric competitors of the t -test, Ann. Math. Stat. 27 (1956), 324—335; Chernoff and Savage, Asymptotic normality and efficiency of certain nonparametric test statistics, Ann. Math. Stat. 29 (1958), 972—994; van der Waerden, Order tests for the two-sample problem and their powers, Koninkl. Ned. Acad. Wetenschap., Proc., ser. A. 55 (1952), 435—458 and 56 (1953), 303—316; Ван-дер-Варден Б. Л., Математическая статистика, пер. с нем., М., ИЛ, 1960.

дена, который асимптотически эквивалентен критерию Фишера — Иэйтса.

Упомянутые результаты не зависят от предположения о равенстве объема выборок; они остаются также справедливыми, если m/n и m^*/n^* стремятся к общему пределу ρ , $0 < \rho < \infty$, когда $n \rightarrow \infty$. Асимптотические результаты хорошо согласуются с теми, которые найдены для выборок малого объема по крайней мере в случае нормального распределения F .

Для проверки гипотезы $G = F$ при двусторонней альтернативе, состоящей в том, что величины Y стохастически меньше или больше величин X , можно применить двусторонний вариант рассмотренных выше критериев. В частности, если $m = n$, то (17) подсказывает, что критическую область здесь можно взять в виде

$$|\sum h(s_j) - \sum h(r_i)| > C.$$

Теория этих вопросов, однако, здесь находится на менее удовлетворительном уровне, чем для случая односторонних гипотез. Так, например, для двустороннего критерия Вилкоксона с $h(k) = k$ и других подобных критериев даже не известно, являются ли они несмещенными при рассматриваемых двусторонних альтернативах или будут допустимыми в классе всех ранговых критериев. С другой стороны, относительная асимптотическая эффективность их та же, что и для случая односторонних альтернатив.

Гипотеза $G = F$ для двух выборок может также проверяться при общей альтернативе $G \neq F$. Эта задача возникает, скажем, при решении вопроса о том, можно ли два продукта, два множества данных и т. п. объединить вместе, когда ничего не известно о характере их распределений. Поскольку на альтернативные распределения не налагается ограничений, проблема остается инвариантной при всех преобразованиях

$$x'_i = f(x_i), \quad y'_j = f(y_j), \quad i = 1, \dots, m; \quad j = 1, \dots, n,$$

таких, что функция f имеет только конечное число разрывов. Не существует критериев, инвариантных относительно этой группы, за исключением только критерия $\phi(x, y) \equiv \alpha$.

Этот критерий, однако, не является допустимым, поскольку существуют критерии для H , строго несмещенные при всех альтернативах $G \neq F$ (задача 34). Наиболее часто к рассматриваемой задаче применяется критерий Смирнова, состоящий в следующем. Пусть эмпирические функции распределения двух выборок определены по формулам

$$S_{x_1, \dots, x_m}(z) = \frac{a}{m}, \quad S_{y_1, \dots, y_n}(z) = \frac{b}{n},$$

где a и b — числа x и y , меньших или равных z , соответственно. Тогда гипотеза $H: G = F$ отвергается*), если

$$\sup_z |S_{x_1, \dots, x_m}(z) - S_{y_1, \dots, y_n}(z)| > C.$$

9. Гипотеза симметрии

В разделе 7 было показано, что проверка гипотезы об отсутствии эффекта обработки методом парных сравнений сводится с помощью принципа инвариантности к проверке гипотезы

$$H'_2: D(z) + D(-z) = 1 \text{ для всех } z,$$

т. е. к проверке того, что распределение D разностей $Z_i = Y_i - X_i (i = 1, \dots, N)$ симметрично относительно начала координат. Распределение D может быть описано тройкой (ρ, F, G) , где

$$\begin{aligned} \rho &= P\{Z \leq 0\}, \quad F(z) = P\{|Z| \leq z | Z < 0\}, \\ G(z) &= P\{Z \leq z | Z > 0\}. \end{aligned}$$

В этих терминах гипотеза симметрии распределения относительно начала координат будет эквивалентна гипотезе

$$H: \rho = \frac{1}{2}, \quad G = F.$$

Как было показано, инвариантность и достаточность редуцируют данные к рангам $S_1 < \dots < S_n$ положительных величин Z в ряду абсолютных значений $|Z_1|, \dots, |Z_n|$. Вероятность, что $S_1 = s_1, \dots, S_n = s_n$, равна вероятности этого события при условии, что n задано, помноженной на вероятность того, что число положительных наблюдений равно n . Следовательно,

$$\begin{aligned} P\{S_1 = s_1, \dots, S_n = s_n\} &= \\ &= \binom{N}{n} (1 - \rho)^n \rho^{N-n} P_{F, G}\{S_1 = s_1, \dots, S_n = s_n | n\}, \end{aligned}$$

где последний сомножитель дается формулой (18). Если справедлива гипотеза H , то эта вероятность равна

$$P\{S_1 = s_1, \dots, S_n = s_n\} = \frac{1}{2^N},$$

*) Обзор, посвященный теории этого и других сходных критериев, содержащий также ссылки на относящиеся к ним таблицы, дал Darling, The Kolmogorov — Smirnov, Cramer — von Mises tests, Ann. Math. Stat. 28 (1957), 823—838. Детальное изучение распределения соответствующей статистики при гипотезе равенства провел Hodges, The significance probability of the Smirnov two-sample test, Arkiv Mat. 3 (1957), 469—486. См. дальнейшие ссылки в Дополнении к книге: Г. Крамер, Математические методы статистики, 2-е изд., М., Мир, 1975.

для каждой из $\sum_{n=0}^N \binom{N}{n} = 2^N$ групп $(s_1, \dots, s_n)$, удовлетворяю-

щих условию $1 \leq s_1 < \dots < s_n \leq N$. Поэтому каждый ранговый критерий уровня $\alpha = k/2^N$ отвергает гипотезу в области, состоящей в точности из k таких точек $(s_1, \dots, s_n)$.

Альтернатива K о благоприятности эффекта применяемого способа обработки характеризуется тем, что случайная величина Z стохастически больше, чем некоторая случайная величина, симметрично распределенная около 0. Представляется естественным искать область, где гипотеза H отбрасывается, в виде $h(s_1) + \dots + h(s_n) > C$, здесь n не будет постоянным, как и в задачах о двух выборках, но зависит от результатов наблюдений. Частными случаями являются *одновыборочный критерий Вилкоксона* с $h(s) = s$ и аналог критерия Фишера — Иэйтса с $h(s) = E(W^{(s)})$, где $W^{(1)} < \dots < W^{(N)}$ — упорядоченные значения величин $|V_1|, \dots, |V_N|$, причем V_i образуют выборку из распределения $N(0, 1)$. Тем самым W здесь образуют упорядоченную выборку объема N из распределения с плотностью, равной $\sqrt{2/\pi} e^{-w^2/2}$ для $w \geq 0$.

Как и в задаче о двух выборках, можно показать, что каждый из этих критериев является наиболее мощным (среди всех инвариантных критериев) для некоторых альтернатив и что они оба оказываются несмещенными при альтернативах класса K . Их асимптотическая эффективность по сравнению с t -критерием, служащим для проверки того, что среднее значение Z равно нулю, имеет те же самые значения $3/\pi$ и 1 соответственно, как и для аналогичных двувывборочных критериев, если предположить, что случайные величины Z нормально распределены.

В некоторых приложениях, например тогда, когда выводы приходится делать на основе данных, полученных при разных экспериментальных условиях или разными методами, было бы нереально считать величины $Z_1, \dots, Z_N$ одинаково распределенными. Вместо этого предположим, что Z_i все еще остаются независимыми, но их (непрерывные) распределения D_i произвольны. Проверяемая гипотеза состоит в том, что каждое из этих распределений симметрично относительно точки 0.

Эта задача остается инвариантной при всех преобразованиях $z'_i = f_i(z_i)$, $i = 1, \dots, N$, где функции f_i непрерывны, нечетны и строго возрастают. Максимальным инвариантом здесь является число положительных наблюдений и, как это следует из примера 7, существует РНМ инвариантный критерий, а именно *критерий знаков*, отклоняющий проверяемую гипотезу, когда n достаточно велико. Этот критерий отражает тот факт, что размеры самих наблюдений или их абсолютных значений можно всецело

объяснить в терминах «разброса» распределений D_i , так что существенны только знаки величин Z .

Часто случается, что величины Z разумно считать одинаково распределенными, но доверять всецело этому предположению не представляется возможным. В этом случае предпочтительнее было бы использовать информацию, содержащуюся в рангах s_i , но потребовать, чтобы применяемый критерий позволял контролировать вероятность ложного отбрасывания гипотезы даже и тогда, когда сделанное предположение неверно. Как показывает нижеследующая лемма, это требование на самом деле выполняется для каждого (симметричного) рангового критерия. Фактически, лемма не требует даже независимости величин Z ; будет показано, что каждый симметричный ранговый критерий сохраняет установленный уровень значимости, если только внутри каждой пары предмет обработки выбирается случайным образом.

Лемма 3. Пусть $\varphi(z_1, \dots, z_N)$ — симметричная функция от N переменных такая, что

$$E_D \varphi(Z_1, \dots, Z_N) = \alpha, \quad (20)$$

когда величины Z образуют выборку из некоторого непрерывного распределения D , которое симметрично относительно начала координат, тогда

$$E \varphi(Z_1, \dots, Z_N) = \alpha, \quad (21)$$

если совместное распределение Z не меняется при 2^N преобразованиях $Z'_1 = \pm Z_1, \dots, Z'_N = \pm Z_N$.

Доказательство. Из условия (20) вытекает, что

$$\sum_{(j_1, \dots, j_N)} \sum \varphi(\pm z_{i_1}, \dots, \pm z_{i_N}) / (2^N \cdot N!) = \alpha \text{ почти всюду,} \quad (22)$$

где внешнее суммирование распределено на все $N!$ перестановок $(j_1, \dots, j_N)$, а внутреннее — на все 2^N возможных выборов знаков $+$ и $-$. Этот факт устанавливается в точности так же, как и в теореме 3 главы 5. Если к тому же функция φ симметрична, то из (22) следует, что

$$\sum \varphi \frac{(\pm z_1, \dots, \pm z_N)}{2^N} = \alpha. \quad (23)$$

Предположим, что распределение Z инвариантно относительно указанных выше 2^N преобразований. Тогда условная вероятность каждой комбинации знаков величин $Z_1, \dots, Z_N$ при заданных $|Z_1|, \dots, |Z_N|$ равна $1/2^N$. Следовательно, (23) эквивалентно тому, что

$$E[\varphi(Z_1, \dots, Z_N) | |Z_1|, \dots, |Z_N|] = \alpha \text{ почти всюду,} \quad (24)$$

откуда следует (21), что и надо было показать.

10. Инвариантные доверительные множества

Доверительные множества для параметра θ при наличии мешающих параметров ϕ рассматривались в главе 5 (разделы 4 и 5) в предположении, что параметр θ — действительный. Соответствие между областью $A(\theta_0)$, где принимается гипотеза $H(\theta_0)$: $\theta = \theta_0$, и доверительным множеством $S(x)$ для θ задается формулами (34) и (35) главы 5. Однако оно не зависит от этого предположения и выполнено независимо от того, является ли θ действительным числом, вектором или служит только для обозначения неизвестной функции распределения (в последнем случае доверительные интервалы становятся доверительными границами для функции распределения). Указанное соответствие, которое можно выразить в виде соотношения

$$\theta \in S(x) \text{ тогда и только тогда, когда } x \in A(\theta), \quad (25)$$

служило основой для получения равномерно наиболее точных и равномерно наиболее точных несмещенных доверительных множеств. В настоящем разделе это соотношение будет использовано для получения равномерно наиболее точных инвариантных доверительных множеств.

Мы начнем с определения инвариантности для доверительных множеств. Пусть G — группа преобразований переменной X , сохраняющая семейство распределений $\{P_{\theta, \phi}(\theta, \phi) \in \Omega\}$, а G — индуцированная группа преобразований на Ω . Если $\bar{g}(\theta, \phi) = (\theta', \phi')$, то мы будем предполагать, что θ' зависит только от $\bar{g}$ и θ , но не от ϕ , так что $\bar{g}$ индуцирует преобразование в пространстве значений параметра θ . Чтобы не усложнять обозначений, будем писать $\theta' = \bar{g}\theta$. Для каждого преобразования $g \in G$ обозначим через g^* преобразование множеств S в пространстве параметра θ , определенное формулой

$$g^*S = \{\bar{g}\theta: \theta \in S\}, \quad (26)$$

так что g^*S есть множество, полученное применением преобразования $\bar{g}$ к каждой точке $\theta \in S$. Каждая доверительная процедура, задаваемая классом доверительных множеств $S(x)$, будет называться *инвариантной относительно группы преобразований G* , если

$$g^*S(x) = S(gx) \text{ для всех } x \in \mathcal{X}, \quad g \in G. \quad (27)$$

Это определение является частным случаем применения общего принципа инвариантности, рассмотренного в главе 1. Если преобразование g интерпретировать как замену системы координат, то (27) означает, что доверительные множества не зависят от координатной системы, в которой выражены исходные данные. Утверждение, что преобразованный параметр $\bar{g}\theta$ лежит в $S(gx)$, эквивалентно тому, что $\theta \in g^{*-1}S(gx)$, а это эквивалентно

первоначальному утверждению, что $\theta \in S(x)$, если только выполнено условие (27).

Пример 11. Пусть X, Y — независимые нормально распределенные величины со средними ξ, η и единичной дисперсией, и пусть G — группа всех движений плоскости, которая порождается всеми сдвигами и ортогональными преобразованиями. Здесь $\bar{g} = g$ для всех $g \in G$. Инвариантный класс доверительных множеств образуют, например, все круги радиуса $\sqrt{C}$ с центром в (x, y) :

$$S(x, y) = \{(\xi, \eta): (x - \xi)^2 + (y - \eta)^2 \leq C\}.$$

Множество $g^*S(x, y)$ — это множество всех точек $g(\xi, \eta)$ с $(\xi, \eta) \in S(x, y)$ и, следовательно, оно получается, если $S(x, y)$ подвергнуть движению g . В результате получится окружность радиуса $\sqrt{C}$ с центром $g(x, y)$; тем самым условие (27) удовлетворяется.

В соответствии с определением, данным в главах 3 и 5, класс доверительных множеств для θ будет называться *равномерно наиболее точным инвариантным классом* с доверительным уровнем $1 - \alpha$, если среди всех инвариантных классов множеств с этим уровнем он минимизирует вероятность

$$P_{\theta, \theta'}\{\theta' \in S(x)\} \quad \text{для всех } \theta' \neq \theta.$$

Чтобы из семейства РНМ инвариантных критериев получить доверительные множества с этим свойством, изучим соотношения между инвариантностью доверительных множеств и связанных с ними критериев.

Предположим, что для каждого θ_0 существует группа преобразований G_{θ_0} , которая сохраняет инвариантной задачу проверки гипотезы $H(\theta_0): \theta = \theta_0$. Пусть G — группа преобразований, порожденная совокупностью групп G_{θ} .

Лемма 4. (I) Пусть $S(x)$ — любой класс доверительных множеств, который инвариантен относительно группы преобразований G , и пусть $A(\theta) = \{x: \theta \in S(x)\}$; тогда область принятия $A(\theta)$ инвариантна при G_{θ} для каждого θ .

(II) Если, кроме того, для каждого θ_0 область принятия $A(\theta_0)$ является РНМ инвариантной для гипотезы $H(\theta_0)$ уровня α , то класс доверительных множеств $S(x)$ будет равномерно наиболее точным среди всех доверительных множеств доверительного уровня $1 - \alpha$.

Доказательство. (I) Рассмотрим некоторое фиксированное θ и пусть $g \in G_{\theta}$. Тогда

$$\begin{aligned} gA(\theta) &= \{gx: \theta \in S(x)\} = \{x: \theta \in S(g^{-1}x)\} = \\ &= \{x: \theta \in g^{*-1}S(x)\} = \{x: \bar{g}\theta \in S(x)\} = \\ &= \{x: \theta \in S(x)\} = A(\theta). \end{aligned}$$

Здесь третье равенство выполнено, так как $S(x)$ инвариантно, а пятое — поскольку $g \in G_{\theta}$ и, следовательно, $\bar{g}\theta = \theta$.

(II) Если $S'(x)$ есть некоторый другой инвариантный класс доверительных множеств с заданным уровнем, то соответствующие области $A'(\theta)$, согласно (I), определяют инвариантные критерии для гипотез $H(\theta)$. Отсюда следует, что эти критерии являются равномерно не более мощными, чем критерии с областями принятия $A(\theta)$. Следовательно,

$$P_{\theta, \theta} \{\theta' \in S(x)\} \leq P_{\theta, \theta} \{\theta' \in S'(x)\} \quad \text{для всех } \theta' \neq \theta,$$

что и надо было показать.

Из этой леммы непосредственно следует, что если для каждой гипотезы $H(\theta)$ (инвариантной относительно G_θ) найдены РНМ инвариантные области $A(\theta)$ и если доверительные множества $S(x) = \{\theta : x \in A(\theta)\}$ инвариантны относительно G , то эти множества являются равномерно наиболее точными инвариантными доверительными множествами.

Пример 12. В предположениях, сделанных в примере 11, задача проверки гипотезы $\xi = \xi_0, \eta = \eta_0$ инвариантна относительно группы G_{ξ_0, η_0} ортогональных преобразований около точки (ξ_0, η_0) :

$$\begin{aligned} X' - \xi_0 &= a_{11}(X - \xi_0) + a_{12}(Y - \eta_0), \\ Y' - \eta_0 &= a_{21}(X - \xi_0) + a_{22}(Y - \eta_0), \end{aligned}$$

где матрица a_{ij} ортогональна. Относительно этой группы преобразований существует РНМ инвариантный критерий с областью принятия гипотезы (задача 8 главы 7)

$$(X - \xi_0)^2 + (Y - \eta_0)^2 \leq C.$$

Пусть G_0 — наименьшая группа, содержащая группы $G_{\xi, \eta}$ для всех ξ, η . Поскольку эта группа является подгруппой группы G примера 11 (эти две группы в действительности совпадают, но сейчас это несущественно), то доверительные множества $(X - \xi)^2 + (Y - \eta)^2 \leq C$ инвариантны при G_0 и, следовательно, являются среди инвариантных равномерно наиболее точными.

Пример 13. Пусть $X_1, \dots, X_n$ — независимые, нормально распределенные величины со средним ξ и дисперсией σ^2 . Построение доверительных интервалов для ξ основано на решении задачи о проверке гипотезы $H(\xi_0)$: $\xi = \xi_0$, которая инвариантна относительно группы G_{ξ_0} преобразований $X'_i = a(X_i - \xi_0) + \xi_0$ ($a \neq 0$). РНМ инвариантный критерий для $H(\xi_0)$ имеет область принятия

$$\frac{\sqrt{(n-1)n} |\bar{X} - \xi_0|}{\sqrt{\sum (X_i - \bar{X})^2}} \leq C,$$

которой соответствуют доверительные интервалы

$$\bar{X} - \frac{C}{\sqrt{n(n-1)}} \sqrt{\sum (X_i - \bar{X})^2} \leq \xi \leq \bar{X} + \frac{C}{\sqrt{n(n-1)}} \sqrt{\sum (X_i - \bar{X})^2}. \quad (28)$$

Группа G в настоящем случае состоит из всех преобразований g : $X'_i = aX_i + b$ ($a \neq 0$), которая для параметра ξ индуцирует преобразование $\bar{g}$: $\xi' = a\xi + b$. Применение соответствующего преобразования g^* к интервалу (28) преобразует его во множество точек $a\xi + b$, где ξ удовлетворяет (28), т. е.

в интервал, начальная и конечная точки которого имеют вид

$$a\bar{X} + b - |a|C \sqrt{\frac{\sum (X_i - \bar{X})^2}{n(n-1)}}$$

и

$$a\bar{X} + b + |a|C \sqrt{\frac{\sum (X_i - \bar{X})^2}{n(n-1)}}.$$

Поскольку интервал с этими концами совпадает с интервалом, полученным заменой в (28) X_i на $aX_i + b$, то доверительные интервалы (28) инвариантны при G_0 и, следовательно, являются среди инвариантных равномерно наиболее точными.

11. Доверительные границы для функций распределения

Предположим, что $X = (X_1, \dots, X_n)$ есть выборка из совокупности с неизвестной непрерывной функцией распределения F ; пусть надо определить нижние и верхние границы L_x и M_x так, чтобы с заданной вероятностью $1 - \alpha$ неравенства

$$L_x(u) \leq F(u) \leq M_x(u), \quad \text{где } u \text{ любое,}$$

выполнялись для всех непрерывных функций распределения F . Эта задача инвариантна относительно группы G , состоящей из преобразований

$$X'_i = g(X_i), \quad i = 1, \dots, n,$$

где g — любая непрерывная строго возрастающая функция. Индуцированное преобразование в пространстве параметров задается формулой $\bar{g}F = F(g^{-1})$.

Если $S(x)$ есть множество непрерывных функций распределения

$$S(x) = \{F : L_x(u) \leq F(u) \leq M_x(u) \text{ для всех } u\},$$

тогда

$$\begin{aligned} g^*S(x) &= \{\bar{g}F : L_x(u) \leq F(u) \leq M_x(u) \text{ для всех } u\} = \\ &= \{F : L_x[g^{-1}(u)] \leq F(u) \leq M_x[g^{-1}(u)] \text{ для всех } u\}. \end{aligned}$$

Для инвариантных процедур это множество должно совпадать с множеством

$$\begin{aligned} S(gx) &= \{F : L_g(x_1), \dots, g(x_n)(u) \leq F(u) \leq \\ &\leq M_g(x_1), \dots, g(x_n)(u) \text{ для всех } u\}. \end{aligned}$$

Поэтому принцип инвариантности приводит к условию

$$L_g(x_1), \dots, g(x_n)[g(u)] = L_x(u); \quad M_g(x_1), \dots, g(x_n)[g(u)] = M_x(x)$$

для всех x и u .

Чтобы охарактеризовать совокупность инвариантных процедур, рассмотрим *эмпирическую* функцию распределения T_x , задаваемую формулой

$$T_x(u) = \frac{i}{n} \quad \text{для} \quad x^{(i)} \leq u < x^{(i+1)}, \quad i = 0, \dots, n,$$

где $x^{(1)} < \dots < x^{(n)}$ — упорядоченная выборка и где $x^{(0)} = -\infty$, $x^{(n+1)} = \infty$. Тогда необходимое и достаточное условие для того, чтобы L и M удовлетворяли указанному выше условию инвариантности, состоит в существовании чисел $a_0, \dots, a_n$; $a'_0, \dots, a'_n$ таких, что

$$L_x(u) = a_i, \quad M_x(u) = a'_i \quad \text{для} \quad x^{(i)} < u < x^{(i+1)}.$$

Достаточность этого условия очевидна. Для доказательства необходимости обозначим u и u' две точки, удовлетворяющие условию $x^{(i)} < u < u' < x^{(i+1)}$. Для заданных $y_1, \dots, y_n$ и v с $y^{(i)} < v < y^{(i+1)}$ существуют $g, g' \in G$ такие, что

$$g(y^{(i)}) = g'(y^{(i)}) = x^{(i)}, \quad g(v) = u, \quad g'(v) = u'.$$

Если L_x, M_x инвариантны, то $L_x(u') = L_y(v)$, $L_x(u) = L_y(v)$ и, следовательно, $L_x(u') = L_x(u)$; аналогично этому $M_x(u') = M_x(u)$, что и требовалось доказать. Это показывает, что L_x и M_x будут ступенчатыми функциями, точки разрыва которых совпадают с точками разрыва T_x .

Поскольку каждые две непрерывные строго возрастающие функции распределения могут быть преобразованы друг в друга посредством преобразования $\bar{g}$, то вероятность того, что они будут накрыты инвариантной доверительной полосой (см. задачу 48), одна и та же. Предположим теперь, что F непрерывна, но не обязательно строго возрастает. Если I есть некоторый интервал постоянства $\bar{F}$, то наблюдения не попадают в I , так что I является интервалом постоянства так же и для эмпирической функции распределения. Отсюда следует, что вероятность доверительной полосе накрыть F не зависит от наличия интервала I и, следовательно, имеет одно и то же значение для всех непрерывных функций распределения F .

Для любых чисел a_i, a'_i определим Δ_i, Δ'_i посредством соотношений

$$a_i = \frac{i}{n} - \Delta_i, \quad a'_i = \frac{i}{n} + \Delta'_i.$$

Из сказанного выше видно, что любые числа $\Delta_0, \dots, \Delta_n$; $\Delta'_0, \dots, \Delta'_n$ определяют доверительную полосу для F , которая инвариантна, и, следовательно, имеет постоянную вероятность накрыть распределение F . С помощью этих доверительных областей может быть получен *критерий согласия* для гипотезы о том,

что неизвестное F равно некоторому F_0 : $F = F_0$. Гипотеза принимается, если F_0 лежит целиком внутри полосы, т. е. если

$$-\Delta_i < F_0(u) - T_x(u) < \Delta'_i \quad \text{для всех} \quad x^{(i)} < u < x^{(i+1)} \\ \text{и всех} \quad i = 1, \dots, n.$$

Внутри этого класса не существует РНМ критерия и наиболее удобен выбор $\Delta_i = \Delta'_i = \Delta$ для всех i . Область, в которой для этого критерия, называемого *критерием Колмогорова* *), принимается гипотеза $F = F_0$, может быть тогда записана в виде

$$\sup_{-\infty < u < \infty} |F_0(u) - T_x(u)| < \Delta.$$

Этот критерий является предельным случаем двувыворочного критерия Смирнова, когда число наблюдений во второй выборке стремится к бесконечности.

12. Задачи

К разделу 1

1. Пусть G — группа измеримых преобразований пространства $(\mathcal{X}, \mathcal{A})$, оставляющая семейство $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ инвариантным, и пусть $T(x)$ — измеримое отображение в пространство $(\mathcal{T}, \mathcal{B})$. Предположим, что из $T(x_1) = T(x_2)$ вытекает, что $T(gx_1) = T(gx_2)$ для всех $g \in G$ так, что G индуцирует группу G^* на $\mathcal{T}$ с помощью соотношений $g^*T(x) = T(gx)$. Предположим далее, что индуцируемое преобразование g^* $\mathcal{B}$ -измеримо. Тогда группа G^* оставляет семейство распределений $\mathcal{P}^T = \{P_\theta^T, \theta \in \Omega\}$ на T инвариантным.

К разделу 2

2. (I) Пусть $\mathcal{X}$ — совокупность точек $x = (x_1, \dots, x_n)$, все координаты которых отличны от нуля, и пусть G — группа преобразований $x'_i = cx_i$, $c > 0$. Тогда максимальным инвариантом относительно группы G является $(\operatorname{sgn} x_n, x_1/x_n, \dots, x_{n-1}/x_n)$, где $\operatorname{sgn} x$ равно 1 или -1 , когда x положительно или отрицательно.

(II) Пусть $\mathcal{X}$ — пространство точек $x = (x_1, \dots, x_n)$, все координаты которых различны, и пусть G — группа всех преобразований $x'_i = f(x_i)$, $i = 1, \dots, n$, таких, что f есть взаимно однозначное отображение действительной прямой на себя с самое большее конечным числом точек разрыва. Тогда группа G транзитивна на $\mathcal{X}$.

*) Обзор, посвященный теории этого и других связанных с ним критериев (включая и критерии согласия для того случая, когда рассматриваемые гипотезы относятся не к одному-единственному распределению, а к целому параметрическому семейству), дан в работе: Darling, The Kolmogorov — Smirnov, Cramér — von Mises tests, Ann. Math. Stat. 28 (1957), 823—838. Эта статья содержит также ссылки на необходимые таблицы. Некоторые вопросы, относящиеся к односторонним критериям, обсуждаются в статье: Chapman, A comparative study of several one-sided goodness-of-fit tests, Ann. Math. Stat. 29 (1958), 655—674. См. дальнейшие ссылки в Дополнении к книге Г. Крамера, Математические методы статистики, 2-е изд., М., Мир, 1975.

[(II) Пусть $x = (x_1, \dots, x_n)$ и $x' = (x'_1, \dots, x'_n)$ — две точки из $\mathcal{X}$. Пусть $I_1, \dots, I_n$ — множество непересекающихся полуинтервалов, которые (вместе с их концевыми точками) покрывают действительную прямую и такие, что $x_i \in I_j$. Обозначим $I'_1, \dots, I'_n$ соответствующие множества интервалов для $x'_1, \dots, x'_n$. Тогда существует преобразование f , которое отображает каждое I_j непрерывным образом в I'_j , точки x_j — в x'_j , и множество из $n-1$ концевой точки интервалов $I_1, \dots, I_n$ — в множество концевых точек интервалов $I'_1, \dots, I'_n$.]

3. (I) Достаточное условие для того, чтобы (8) было выполнено, состоит в том, что D есть нормальная подгруппа G .

(II) Если G есть группа преобразований $x' = ax + b$, $a \neq 0$, $-\infty < b < \infty$, тогда подгруппа сдвигов $x' = x + b$ является нормальной, подгруппа же преобразований $x' = ax$ таковой не будет. [Подгруппа называется нормальной, если для каждого $d \in D$, $g \in G$, существует $d' \in D$ такой, что $gd = d'g$. Из равенства $s(x_1) = s(x_2)$ следует, что $x_2 = dx_1$, для некоторого $d \in D$ и, следовательно, $ex_2 = edx_1 = d'ex_1$. Результат п. (I) немедленно вытекает из того, что преобразование s инвариантно при D .]

К разделу 3

4. Пусть $f(x, y)$ — совместная плотность величин X, Y . Тогда интеграл
$$h(z) = \int_{-\infty}^{\infty} f(y - z, y) dy$$
 конечен для почти всех z и является плотностью вероятности для $Z = Y - X$.

[Поскольку $P\{Z \leq b\} = \int_{-\infty}^b h(z) dz$ конечна, то функция h конечна почти всюду.]

5. (I) Пусть $X = (X_1, \dots, X_n)$ — случайный вектор с плотностью $(1/\theta^n) f[(x_1 - \xi)/\theta, \dots, (x_n - \xi)/\theta]$, где $-\infty < \xi < \infty$, $0 < \theta$ неизвестны и где f есть четная функция. Задача проверки гипотезы $f = f_0$ при альтернативе $f = f_1$ остается инвариантной относительно преобразований $x'_i = ax_i + b$ ($i = 1, \dots, n$), $a \neq 0$, $-\infty < b < \infty$, и наиболее мощный инвариантный критерий отвергает гипотезу $f = f_0$ в области

$$\int_{-\infty}^{\infty} \int_0^{\infty} v^{n-2} f_1(vx_1 + u, \dots, vx_n + u) dv du > C \int_{-\infty}^{\infty} \int_0^{\infty} v^{n-2} f_0(vx_1 + u, \dots, vx_n + u) dv du.$$

(II) Пусть $X = (X_1, \dots, X_n)$ имеет плотность вероятности

$$f\left(x_1 - \sum_{j=1}^k w_{1j} \beta_j, \dots, x_n - \sum_{j=1}^k w_{nj} \beta_j\right),$$

где $k < n$, величины w заданы, а величины β неизвестны. Желательно проверить гипотезу $f = f_0$ при альтернативной гипотезе $f = f_1$. Эта задача остается инвариантной при преобразованиях $x'_i = x_i + \sum_{j=1}^k w_{ij} \gamma_j$, $-\infty < \gamma_1, \dots$

..., $\gamma_k < \infty$, и наиболее мощный инвариантный критерий характеризуется критической областью

$$\frac{\int \dots \int f_1 \left(x_1 - \sum w_{1j} \beta_j, \dots, x_n - \sum w_{nj} \beta_j \right) d\beta_1 \dots d\beta_k}{\int \dots \int f_0 \left(x_1 - \sum w_{1j} \beta_j, \dots, x_n - \sum w_{nj} \beta_j \right) d\beta_1 \dots d\beta_k} > C.$$

[Максимальным инвариантом является функция

$$y = \left(x_1 - \sum_{r=n-k+1}^n a_{1r} x_r, x_2 - \sum_{r=n-k+1}^n a_{2r} x_r, \dots, x_{n-k} - \sum_{r=n-k+1}^n a_{n-k, r} x_r \right)$$

с надлежащим образом подобранными константами a_{ir}]

6. Пусть $X_1, \dots, X_m; Y_1, \dots, Y_n$ — выборки из экспоненциальных распределений с плотностями $\sigma^{-1} e^{-(x-\xi)/\sigma}$ для $x \geq \xi$ и $\tau^{-1} e^{-(y-\eta)/\tau}$ для $y \geq \eta$.

(I) В проверке гипотезы $\tau/\sigma \leq \Delta$ при альтернативе $\tau/\sigma > \Delta$ существует РНМ критерий, инвариантный относительно группы преобразований G :

$$X'_i = aX_i + b, \quad Y'_j = aY_j + c, \quad a > 0, \quad -\infty < b, \quad c < \infty.$$

Гипотеза отвергается в области

$$\frac{\sum [y_j - \min(y_1, \dots, y_n)]}{\sum [x_i - \min(x_1, \dots, x_m)]} > C.$$

(II) Этот критерий является также РНМ несмещенным.

(III) Распространить эти результаты на тот случай, когда наблюдаются только r наименьших значений X и S наименьших значений Y .

[(II) См. задачу 12 главы 5.]

7. Если $X_1, \dots, X_n$ и $Y_1, \dots, Y_n$ — выборки из нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ соответственно, то задача проверки гипотезы $\tau^2 = \sigma^2$ при двусторонней альтернативе $\tau^2 \neq \sigma^2$ остается инвариантной относительно группы G , порожденной преобразованиями

$$X'_i = aX_i + b, \quad Y'_i = aY_i + c, \quad a \neq 0 \quad \text{и} \quad X'_i = Y_i, \quad Y'_i = X_i.$$

Существует РНМ инвариантный относительно G критерий, отклоняющий гипотезу в области

$$W = \max \left\{ \frac{\sum (Y_i - \bar{Y})^2}{\sum (X_i - \bar{X})^2}, \frac{\sum (X_i - \bar{X})^2}{\sum (Y_i - \bar{Y})^2} \right\} \geq k.$$

[Отношение плотностей вероятностей величины W , подсчитываемое в предположении $\tau^2/\sigma^2 = \Delta$ и $\tau^2/\sigma^2 = 1$, пропорционально $[(1+w)/(\Delta+w)]^{n-1} + [(1+w)/(1+\Delta w)]^{n-1}$ для $w \geq 1$. Производная этого выражения ≥ 0 для всех Δ .]

К разделу 4

8. (I) Воспользовавшись критерием (11) для проверки гипотезы $H: p \leq p_0$ при альтернативе $K: p > p_0$, определить число необходимых наблюдений для получения мощности β , если этот критерий применяется для проверки гипотезы $p = p_1$, $\alpha = 0,05$, $\beta = 0,9$ для случаев $p_0 = 0,1$, $p_1 = 0,15; 0,20; 0,25$; $p_0 = 0,05$, $p_1 = 0,10; 0,15; 0,20; 0,25$; $p_0 = 0,01$, $p_1 = 0,02; 0,05; 0,10, 0,15; 0,20$.

(II) Сравнить полученный результат с числом наблюдений, требуемых в том случае, когда проверка производится по качественному признаку, т. е. критерий основывается на суммарном числе дефектных изделий, и также со средним числом наблюдений, если для проверки гипотез p_0 и p_1 применяется биномиальный последовательный критерий отношений вероятностей.

9. *Последовательный t -критерий*. Гипотезу $p \geq p_0$, что эквивалентно $\sqrt{n}\xi/\sigma \leq \delta_0$, рассмотренную в разделе 4, можно проверять с помощью следующего последовательного t -критерия. Пусть $\delta_0 < \delta_1$,

$$s_n^2 = \sum_{i=1}^n \left(x_i - \frac{x_1 + \dots + x_n}{n} \right)^2 \quad \text{и} \quad t_{n-1} = \frac{(x_1 + \dots + x_n)/\sqrt{n}}{s_n/\sqrt{n-1}}.$$

Если $p_\delta(t_1, \dots, t_{n-1})$ обозначает совместную плотность величин $t_1, \dots, t_{n-1}$ то наблюдения продолжаются, если

$$A_0 < \frac{p_{\delta_1}(t_1, \dots, t_{n-1})}{p_{\delta_0}(t_1, \dots, t_{n-1})} < A_1,$$

и при первом нарушении одного из этих неравенств гипотеза принимается или отвергается в зависимости от того, $< A_0$ или $> A_1$ вероятностное отношение.

(I) Можно показать*), что эта процедура заканчивается с вероятностью единица. Используя этот факт, показать, что неравенства (34) главы 3 выполнены в настоящем случае.

(II) Эта процедура очень проста и может быть основана на таблицах плотности нецентрального t -распределения, если заметить, что

$$\frac{p_{\delta_1}(t_1, \dots, t_{n-1})}{p_{\delta_0}(t_1, \dots, t_{n-1})} = \frac{p_{\delta_1}(t_{n-1})}{p_{\delta_0}(t_{n-1})},$$

где $p_\delta(t_{n-1})$ есть плотность нецентрального t -распределения, задаваемого формулой (75) главы 5, с $f = n - 1$.

(III) Интересно отметить, что это отношение вероятностей может быть представлено как отношение плотностей первоначальных величин $X_1, \dots, X_n$, усредненных по инвариантной (относительно умножения на число) мере $d\sigma/\sigma$; так что оно равно

$$\frac{\int_0^\infty \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \sum (x_i - \delta_1\sigma)^2 \right] \frac{d\sigma}{\sigma}}{\int_0^\infty \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \sum (x_i - \delta_0\sigma)^2 \right] \frac{d\sigma}{\sigma}}.$$

[(I) Доказательство проводится аналогично доказательству формулы (34) главы 3.

*) David and Kruskal, The Wagr sequential t -tet reaches a decision with probability one, Ann. Math. Stat. 27 (1956), 797—805 и 29 (1958), 936.

(II) *).

(III) Сделать преобразование $v' = v/\delta$ и сравнить с отношением $h\delta$]

10. *Двусторонний t -критерий*. (I) Пусть $X_1, \dots, X_n$ — выборка из $N(\xi, \delta^2)$. В задаче различения гипотез $\xi = 0$ и $\xi \neq 0$ существует РНМ инвариантный критерий (относительно группы преобразований $X'_i = cX_i, c \neq 0$), являющийся двусторонним t -критерием (17) главы 5.

(II) Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из $N(\xi, \delta^2)$ и $N(\eta, \delta^2)$ соответственно. При проверке гипотезы $\eta = \xi$ при альтернативе $\eta \neq \xi$ существует РНМ инвариантный относительно группы преобразований $X'_i = aX_i + b, Y'_j = aY_j + b, a \neq 0$, критерий, а именно двусторонний t -критерий (30) главы 5.

[(I) Достаточность и инвариантность редуцируют задачу к статистике $|t|$, которая в обозначениях раздела 4 имеет плотность вероятности $p_\delta(t) + p_\delta(-t)$ для $t > 0$. Отношение этих плотностей при $\delta = \delta_1$ и $\delta = 0$

пропорционально $\int_0^\infty (e^{\delta_1 v} + e^{-\delta_1 v}) g_{t^2}(v) dv$, что является возрастающей функцией от t^2 и, следовательно, от $|t|$.]

11. *Проверка гипотез о коэффициенте корреляции*. Пусть $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения.

(I) В задаче проверки гипотезы $\rho \leq \rho_0$ при альтернативе $\rho > \rho_0$ существует РНМ критерий, инвариантный относительно группы всех преобразований $X'_i = aX_i + b, Y'_i = cY_i + d$ с $a > 0$ и $c > 0$. Согласно этому критерию гипотеза отвергается, когда выборочный коэффициент корреляции R достаточно велик.

(II) Задача проверки гипотезы $\rho = 0$ при альтернативе $\rho \neq 0$ остается инвариантной еще и при преобразованиях $Y'_i = -Y_i, X'_i = X_i$. Относительно группы, порожденной этим преобразованием и преобразованиями из (I), существует РНМ инвариантный критерий с областью отклонения гипотезы $|R| \geq C$.

[(I) Показать, что плотность вероятности $p_\rho(r)$ для R имеет монотонное отношение правдоподобия, применить условие из главы 3 (задача 6 (I)) к выражению (85) для этой плотности, данному в главе 5. Полагая $t = \rho r - 1$, для второй производной $\partial^2 \log p_\rho(r)/\partial \rho \partial r$ получим выражение

$$\frac{\sum_{i,j=0}^{\infty} c_i c_j t^{i+j-2} [(j-i)^2 (t-1) + (i+j)]}{2 \left[\sum_{i=0}^{\infty} c_i t^i \right]^2}.$$

Для доказательства того, что числитель положителен для всех $t > 0$, заметим, что он больше, чем

$$2 \sum_{i=0}^{\infty} c_i t^{i-2} \sum_{j=i+1}^{\infty} c_j t^j [(j-i)^2 (t-1) + (i+j)].$$

*) В работе Cox, Sequential tests for composite hypotheses, Proc. Camb. Phil. Soc. 48 (1952), 290—299 дано доказательство (II), основанное на тех фактах, что $\sum x_i/n$ и s_n являются достаточными статистиками для (ξ, δ) , что t_{n-1} есть максимальная инвариантная функция от этих достаточных статистик относительно изменения масштаба и что $t_1, \dots, t_{n-1}$ инвариантны при этих преобразованиях. Сходным методом в этой работе разобрано несколько других примеров.

Фиксируя индекс i и употребляя неравенство $c_{i+1} < \frac{1}{2} c_j$, получим, что во внутренней сумме коэффициент при $i!$ будет ≥ 0 .

12. В задаче проверки гипотезы о том, что коэффициент корреляции ρ двумерного нормального распределения будет $\leq \rho_0$, определить мощность, если альтернатива состоит в том, что $\rho = \rho_1$. Уровень значимости α равен 0,05, $\rho_0 = 0,3$, $\rho_1 = 0,5$ и число наблюдений n равно 50; 100; 200.

К разделу 5

13. Почти инвариантность критерия φ относительно группы преобразований G задачи 7 и примера 6 влечет за собой тот факт, что φ эквивалентно некоторому инвариантному критерию.

К разделу 6

14. Рассмотрим задачу проверки гипотез, которая инвариантна относительно группы G преобразований пространства выборок, и пусть $\mathcal{C}$ — класс критериев, замкнутый относительно G , т. е. из условия $\varphi \in \mathcal{C}$ следует $\varphi g \in \mathcal{C}$, где φg — критерий, определенный формулой $\varphi g(x) = \varphi(gx)$. Если в этом классе существует единственный с точностью до множеств меры 0 РНМ критерий φ_0 , тогда φ_0 является почти инвариантным.

15. *Огибающая функция мощности.* Пусть $S(\alpha)$ — класс критериев уровня α для гипотезы H , и пусть $\beta_\alpha^*(\theta)$ — *огибающая функция мощности*, определенная соотношением

$$\beta_\alpha^*(\theta) = \sup_{\varphi \in S(\alpha)} \beta_\varphi(\theta),$$

где β_φ обозначает функцию мощности критерия φ . Если задача проверки гипотезы H инвариантна относительно группы G , то $\beta_\alpha^*(\theta)$ инвариантна относительно преобразований индуцированной группы $\bar{G}$.

16. Обобщением равенства (1) служит соотношение

$$\int_A f(x) dP_\theta(x) = \int_{gA} f(g^{-1}x) dP_{\bar{g}\theta}(x).$$

(II) Если P_{θ_1} абсолютно непрерывна относительно меры P_{θ_0} , тогда $P_{\bar{g}\theta_1}$ абсолютно непрерывна относительно $P_{\bar{g}\theta_0}$ и

$$\frac{dP_{\theta_1}}{dP_{\theta_0}}(x) = \frac{dP_{\bar{g}\theta_1}}{dP_{\bar{g}\theta_0}}(gx) \quad (P_{\theta_0}\text{-почти всюду}).$$

(III) Распределение плотности $dP_{\theta_1}/dP_{\theta_0}(X)$, когда X распределено по закону P_{θ_0} , то же самое, что и распределение величины $dP_{\bar{g}\theta_1}/dP_{\bar{g}\theta_0}(X')$, когда X' распределена по $P_{\bar{g}\theta_0}$.

17. *Инвариантность отношения правдоподобия.* Пусть семейство распределений $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ доминируется мерой μ , $\dot{P}_\theta = dP_\theta/d\mu$, и пусть μg^{-1} — мера, определенная соотношением $\mu g^{-1}(A) = \mu[g^{-1}(A)]$; предположим, что мера μ абсолютно непрерывна по мере μg^{-1} для всех $g \in G$.

(I) Тогда

$$p_\theta(x) = p_{\bar{g}\theta}(gx) \frac{d\mu}{d\mu g^{-1}}(gx) \quad (\mu\text{-почти всюду}).$$

(II) Пусть Ω и ω инвариантны относительно $\bar{G}$ и счетны. Тогда отношение правдоподобия $\sup_{\Omega} p_{\theta}(x) / \sup_{\omega} p_{\theta}(x)$ почти инвариантно относительно G .

(III) Предположим, что $p_{\theta}(x)$ — непрерывная функция по θ для всех x , что Ω сепарабельно и что Ω и ω инвариантны. Тогда отношение правдоподобия почти инвариантно относительно G .

18. *Пример, когда критерий отношения правдоподобия не является допустимым.* Во многих применениях, в которых РНМ инвариантный критерий существует, он совпадает с критерием, основанным на отношении правдоподобия. Однако это не всегда так, что видно из следующего примера. Пусть $P_1, \dots, P_n$ — n равноудаленных точек на окружности $x^2 + y^2 = 4$ и $Q_1, \dots, Q_n$ — на окружности $x^2 + y^2 = 1$. Обозначим начало координат в плоскости x, y через O ; пусть фиксировано α , $0 < \alpha \leq 1/2$, и пусть величина (X, Y) принимает $2n + 1$ значение $P_1, \dots, P_n, Q_1, \dots, Q_n, O$ с вероятностями, задаваемыми следующей таблицей:

	P_i	Q_i	O
H	α/n	$(1-2\alpha/n)$	α
K	p_i/n	0	$(n-1)/n$

где $\sum p_i = 1$. Задача остается инвариантной относительно вращений плоскости на углы $2k\pi/n$ ($k = 0, 1, \dots, n-1$). Область, где гипотеза H отвергается, если пользоваться критерием отношения правдоподобия, состоит из точек $P_1, \dots, P_n$, и его мощность равна $1/n$. С другой стороны, РНМ инвариантный критерий отвергает H , когда $X = Y = 0$, и его мощность равна $(n-1)/n$.

19. Пусть G — группа преобразований пространства $\mathcal{X}$, $\mathcal{A}$ — σ -поле подмножеств $\mathcal{X}$ и μ — мера на $(\mathcal{X}, \mathcal{A})$. Тогда множество $A \in \mathcal{A}$ называется почти инвариантным, если таковой является его характеристическая функция (индикатор).

(I) Совокупность почти инвариантных множеств образует σ -поле $\mathcal{A}_0$, и критическая функция почти инвариантна тогда и только тогда, когда она $\mathcal{A}_0$ -измерима.

(II) Пусть $\mathcal{P} = \{P_{\theta}, \theta \in \Omega\}$ — доминированное семейство вероятностных распределений на $(\mathcal{X}, \mathcal{A})$, и предположим, что $\bar{g}\theta = \theta$ для всех $\bar{g} \in \bar{G}, \theta \in \Omega$. Тогда σ -поле $\mathcal{A}_0$ почти инвариантных множеств достаточно для $\mathcal{P}$.

[Пусть $\lambda = \sum c_i P_{\theta_i}$ эквивалентна $\mathcal{P}$. Тогда

$$\frac{dP_{\theta}}{d\lambda}(gx) = \frac{dP_{\bar{g}^{-1}\theta}}{\sum c_i dP_{\bar{g}^{-1}\theta_i}}(x) = \frac{dP_{\theta}}{d\lambda}(x) \quad (\lambda\text{-почти всюду}),$$

так что плотность $dP_{\theta}/d\lambda$ почти инвариантна и, следовательно, $\mathcal{A}_0$ -измерима.]

К разделу 8

20. *Двувыборочный критерий Вилкоксона.* Пусть $U_{ij} = 1$ или 0 , когда $X_i < Y_j$ или $X_i > Y_j$, и пусть $U = \sum \sum U_{ij}$ — число тех пар X_i, Y_j , для которых $X_i < Y_j$.

(I) Тогда $U = \sum S_i - \frac{1}{2}n(n+1)$, где $S_1 < \dots < S_n$ — ранги величин Y ; таким образом, критерий с критической областью $U > C$ эквивалентен критерию Вилкоксона.

(II) Любое заданное расположение x_i и y_j может быть преобразовано в расположение $x \dots xy \dots y$ посредством конечного числа перестановок соседних элементов. Наименьшее число шагов, за которое заданное расположение можно перевести в указанное, равно $m + n - U$.

21. Математическое ожидание и дисперсия статистики Вилкоксона. Если X и Y — две случайные величины с непрерывными распределениями F и G соответственно, то математическое ожидание и дисперсия статистики Вилкоксона U , определенной в предыдущей задаче, задаются формулами

$$E\left(\frac{U}{mn}\right) = P\{X < Y\} = \int F dG \quad (29)$$

и

$$\begin{aligned} mnD\left(\frac{U}{mn}\right) &= \int F dG + (n-1) \int (1-G)^2 dF + \\ &+ (m-1) \int F^2 dG - (m+n-1) \left(\int F dG\right)^2. \end{aligned} \quad (30)$$

При гипотезе $G = F$ они равны

$$E\left(\frac{U}{mn}\right) = \frac{1}{2}, \quad D\left(\frac{U}{mn}\right) = \frac{(m+n+1)}{12mn}. \quad (31)$$

22. (I) Пусть $Z_1, \dots, Z_n$ — независимые случайные величины с плотностями $f_1, \dots, f_n$ соответственно, и пусть ранг величины Z_i обозначается T_i . Если f есть некоторая плотность вероятности, положительная всякий раз, когда по крайней мере одна из функций f_i положительна, то

$$P\{T_1 = t_1, \dots, T_n = t_n\} = \frac{1}{n!} E \left[\frac{f_1(V^{(t_1)})}{f(V^{(t_1)})} \dots \frac{f_n(V^{(t_n)})}{f(V^{(t_n)})} \right], \quad (32)$$

где $V^{(1)} < \dots < V^{(n)}$ — упорядоченная выборка из распределения с плотностью f .

(II) Если $N = m + n$, $f_1 = \dots = f_m = f$, $f_{m+1} = \dots = f_{m+n} = g$ и $S_1 < \dots < S_n$ обозначают упорядоченные ранги величин $Z_{m+1}, \dots, Z_{m+n}$ среди всех Z , то распределение вероятностей величин $S_1, \dots, S_n$ задается формулой (18).

[(I) Искомая вероятность равна $\int \dots \int f_1(z_1) \dots f_n(z_n) dz \dots dz_n$,

где интегрирование производится по множеству, в котором z_i есть t_i -е наименьшее из чисел z , $i = 1, \dots, n$. При преобразовании $w_{t_i} = z_i$ интеграл

становится равным $\int \dots \int f_1(w_{t_1}) \dots f_n(w_{t_n}) dw_1 \dots dw_n$, где интегрирование производится по множеству $w_1 < \dots < w_n$. Желаемый результат теперь следует из того факта, что плотность вероятности порядковой статистики $V^{(1)} < \dots < V^{(n)}$ равна $N! f(w_1) \dots f(w_n)$ для $w_1 < \dots < w_n$.]

23. (I) Для каждой непрерывной функции распределения F определим $F^{-1}(0) = -\infty$, $F^{-1}(y) = \inf\{x: F(x) = y\}$, для $0 < y < 1$, $F^{-1}(1) = \infty$, если $F(x) < 1$ для всех конечных x , и в противном случае $F^{-1}(y) = \inf\{x: F(x) = y\}$. Тогда $F[F^{-1}(y)] = y$ для всех $0 \leq y \leq 1$, но $F^{-1}[F(y)]$ может быть и меньше y .

(II) Пусть Z имеет непрерывную функцию распределения $G(z) = h[F(z)]$, где F и h — непрерывные функции распределения, последняя из которых

определена на $(0, 1)$. Если $Y = F(Z)$, тогда $P\{Y < y\} = h(y)$ для всех $0 \leq y \leq 1$.

(III) Если Z имеет непрерывную функцию распределения F , тогда $F(Z)$ равномерно распределено на $(0, 1)$.

$$[(II) P\{F(Z) < y\} = P\{Z < F^{-1}(y)\} = F[F^{-1}(y)] = y.]$$

24. Пусть Z_i имеет непрерывную функцию распределения F_i ($i = 1, \dots, N$), и пусть G — группа всех преобразований $Z'_i = f(Z_i)$ таких, что f непрерывна и строго возрастает.

(I) Преобразование, индуцируемое функцией f в пространстве распределений, задается формулой $F'_i = F_i(f^{-1})$.

(II) Две N -элементные группы распределений $(F_1, \dots, F_N)$ и $(F'_1, \dots, F'_N)$ принадлежат к одной и той же траектории, отвечающей преобразованию G тогда и только тогда, когда существуют непрерывные функции распределения $h_1, \dots, h_N$, определенные на $(0, 1)$, и строго возрастающие непрерывные функции F и F' такие, что $F_i = h_i(F)$, $F'_i = h_i(F')$.

$$[(I) P\{f(Z_i) \leq y\} = P\{Z_i \leq f^{-1}(y)\} = F_i[f^{-1}(y)].$$

(II) Если $F_i = h_i(F)$ и F'_i принадлежат одной и той же траектории, т. е. $F'_i = F_i(f^{-1})$, тогда $F'_i = h_i(F')$ с $F' = F(f^{-1})$. Обратно, если $F_i = h_i(F)$ $F'_i = h_i(F')$, тогда $F'_i = F_i(f^{-1})$ с $f = F'^{-1}(F)$.]

25. В предположениях предыдущей задачи, при $F_i = h_i(F)$ распределение рангов $T_1, \dots, T_N$ величин $Z_1, \dots, Z_N$ зависит только от h_i , но не от F . Если функции h_i дифференцируемы, то распределение T_i задается формулой

$$P\{T_1 = t_1, \dots, T_N = t_N\} = \frac{E' [h'_1(U^{(t_1)}) \dots h'_N(U^{(t_N)})]}{N!}, \quad (33)$$

где $U^{(1)} < \dots < U^{(N)}$ есть упорядоченная выборка объема N из равномерного распределения $R(0, 1)$ на $(0, 1)$. [Левая часть равенства (33) есть вероятность того, что i -я из величин $F(Z_1), \dots, F(Z_N)$ является t_i -й наименьшей, $i = 1, \dots, N$. Эта вероятность равна $\int \dots \int h'_1(y_1) \dots h'_N(y_N) dy$, где интегрирование производится по области, в которой y_i есть t_i -е наименьшее значение из y , $i = 1, \dots, N$. Завершается доказательство так же, как и в задаче 22.]

26. *Распределение порядковых статистик.* (I) Если $Z_1, \dots, Z_N$ — выборка из распределения F с плотностью f , то совместная плотность величин $Y_i = Z^{(s_i)}$, $i = 1, \dots, n$, задается формулой

$$\frac{N! f(y_1) \dots f(y_n)}{(s_1 - 1)! (s_2 - s_1 - 1)! \dots (N - s_n)!} [F(y_1)]^{s_1 - 1} [F(y_2) - F(y_1)]^{s_2 - s_1 - 1} \dots \\ \dots [1 - F(y_n)]^{N - s_n} \quad (34)$$

для $y_1 < \dots < y_n$.

(II) Для того частного случая, когда случайные величины Z равномерно распределены на $(0, 1)$, указанная плотность вероятности равна

$$\frac{N!}{(s_1 - 1)! \dots (N - s_n)!} y_1^{s_1 - 1} (y_2 - y_1)^{s_2 - s_1 - 1} \dots (1 - y_n)^{N - s_n}. \quad (35)$$

Для $n = 1$ (35) есть плотность бета-распределения $B_{s, N-s+1}$, которое тем самым оказывается распределением порядковой статистики $Z^{(s)}$, если Z_i имеют распределение $R(0, 1)$.

(III) Пусть распределение величин $Y_1, \dots, Y_n$ задается формулой (35), и пусть величины V_i определены из формул $Y_i = V_{i+1} \dots V_n$, $i = 1, \dots, n$. Тогда совместное распределение величин V_i равно

$$\frac{N!}{(s_1 - 1)! \dots (N - s_n)!} \prod_{i=1}^n v_i^{s_i-1} (1 - v_i)^{s_{i+1} - s_i - 1} \quad (s_{n+1} = N + 1).$$

Таким образом, случайные величины V_i независимы и распределены согласно бета-распределению $B_{s_i, s_{i+1} - s_i}$.

[(I) Если $Y_1 = Z^{(s_1)}, \dots, Y_n = Z^{(s_n)}$ и $Y_{n+1}, \dots, Y_N$ обозначают оставшиеся Z , взятые в первоначальном порядке их индексов, то совместное распределение величин $Y_1, \dots, Y_n$ имеет плотность $N(N-1) \dots (N-n+1) \times \int \dots \int f(y_{n+1}) \dots f(y_N) dy_{n+1} \dots dy_N$, где интегрирование производится по области, в которой $s_1 - 1$ значение y меньше y_1 , $s_2 - s_1 - 1$ значение лежит между y_1 и y_2 , $\dots$, и $N - s_n$ значений $> y_n$. Рассмотрим любое множество, где заданное $s_1 - 1$ значение из y меньше y_1 , $s_2 - s_1 - 1$ значение из них лежит между y_1 и y_2 и т. д. Всего существует $N!/(s_1 - 1)! \dots (N - s_n)!$ таких областей, и интеграл, взятый по каждой из этих областей, принимает одно и то же значение

$$[F(y_1)]^{s_1-1} [F(y_2) - F(y_1)]^{s_2-s_1-1} \dots [1 - F(y_n)]^{N-s_n}.$$

27. (I) Если $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ имеют непрерывные функции распределения F и $G = h(F)$ соответственно, и если функция h дифференцируема, то распределение рангов $S_1 < \dots < S_n$ величин Y задается формулой

$$P\{S_1 = s_1, \dots, S_n = s_n\} = \frac{E[h'(U^{(s_1)}) \dots h'(U^{(s_n)})]}{\binom{m+n}{m}}, \quad (36)$$

где $U^{(1)} < \dots < U^{(m+n)}$ есть упорядоченная выборка из равномерного распределения $R(0, 1)$.

(II) Если, в частности, $G = F^k$, где k — положительное целое число, то формула (36) редуцируется к

$$P\{S_1 = s_1, \dots, S_n = s_n\} = \frac{k^n}{\binom{m+n}{m}} \prod_{j=1}^n \frac{\Gamma(s_j + jk - j)}{\Gamma(s_j)} \cdot \frac{\Gamma(s_j + 1)}{\Gamma(s_j + 1 + jk - j)}. \quad (37)$$

28. Для достаточно малых $\theta > 0$ критерий Вилкоксона максимизирует мощность (среди ранговых критериев) при альтернативах (F, G) с $G = (1 - \theta)F + \theta F^2$.

29. Другое доказательство оптимального свойства критерия Вилкоксона в задаче определения сдвига в логистическом распределении можно получить из предшествующей задачи, если приравнять $F(x - \theta)$ выражению $(1 - \theta)F(x) + \theta F^2(x)$ и пренебречь более высокими степенями θ . Это приводит к дифференциальному уравнению $F - \theta F' = (1 - \theta)F + \theta F^2$, решением которого является логистическое распределение.

30. Пусть $\mathcal{F}_0$ — семейство вероятностных мер на $(\mathcal{X}, \mathcal{A})$, и пусть $\mathcal{C}$ — класс преобразований пространства $\mathcal{X}$. Определим $\mathcal{F}_1$ как класс распределений такой, что $F_1 \in \mathcal{F}_1$, если существуют $F_0 \in \mathcal{F}_0$ и $f \in \mathcal{C}$ такие, что распределение $f(X)$ есть F_1 , когда распределением X является F_0 . Если ϕ — критерий, удовлетворяющий условиям (а) $E_{F_0} \phi(X) = \alpha$ для всех $F_0 \in \mathcal{F}_0$ и (б) $\phi(x) \leq \phi[f(x)]$ для всех x и всех $f \in \mathcal{C}$, тогда ϕ является несмещенным при различии $\mathcal{F}_0$ и $\mathcal{F}_1$.

31. Пусть $X_1, \dots, X_m; Y_1, \dots, Y_n$ — две выборки из одного и того же непрерывного распределения F . Тогда статистика Вилкоксона U , определенная в задаче 20, симметрично распределена около точки $1/2 mn$, даже если $m \neq n$.

32. *Доверительные интервалы для сдвига.* Пусть $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ — выборки из совокупностей с функциями распределения $F(x)$ и $G(y) = F(y - \Delta)$ соответственно. Гипотезу $\Delta = \Delta_0$ можно проверить, применяя двусторонний критерий Вилкоксона к наблюдениям X_i и $Y_j - \Delta_0$.

(I) Доверительным интервалом для Δ является интервал

$$Y^{(1)} - X^{(m)} < \Delta < Y^{(n)} - X^{(1)},$$

когда доверительный коэффициент равен $1 - 2 / \binom{m+n}{m}$, и

$$\min(Y^{(1)} - X^{(m-1)}, Y^{(2)} - X^{(m)}) < \Delta < \max(Y^{(n)} - X^{(2)}, Y^{(n-1)} - X^{(1)}),$$

если этот коэффициент равен $1 - 4 / \binom{m+n}{m}$.

(II) Определить доверительный интервал для Δ , если доверительный коэффициент равен $20/21$, $m = n = 6$, наблюдения x : 0,113; 0,212; 0,249; 0,522; 0,709; 0,788, а y : 0,221; 0,433; 0,724; 0,913; 0,917; 1,58.

33. (I) Пусть X, X' и Y, Y' — независимые выборки объема 2 из непрерывных распределений F и G соответственно. Тогда

$$p = P\{\max(X, X') < \min(Y, Y')\} + P\{\max(Y, Y') < \min(X, X')\} = \frac{1}{3} + \Delta,$$

где $\Delta = \int (F - G)^2 d[(F + G)/2]$.

(II) $\Delta = 0$ тогда и только тогда, когда $F = G$.

[(I) $p = \int (1 - F)^2 dG^2 + \int (1 - G)^2 dF^2$, что после некоторых преобразований дает требуемый результат.

(II) Из условия $\Delta = 0$ следует, что $F(x) = G(x)$, за исключением множества N , имеющего меру нуль для каждого из распределений F и G . Предположим теперь, что $G(x_1) - F(x_1) = \eta > 0$. Тогда существует x_0 такое, что $G(x_0) = F(x_0) + \frac{1}{2} \eta$ и $F(x) < G(x)$ для $x_0 \leq x \leq x_1$. Поскольку $G(x_1) - G(x_0) > 0$, то $\Delta > 0$].

34. *Продолжение.* (I) Для каждого уровня значимости α существует критерий проверки гипотезы $H: G = F$, мощность которого $> \alpha$ для всех непрерывных альтернатив (F, G) с $F \neq G$.

(II) Не существует нерандомизированного несмещенного критерия для проверки гипотезы H при всех альтернативах $G \neq F$ с уровнем значимости $\alpha = 1 / \binom{m+n}{m}$.

[(I) Пусть $X_i, X'_i; Y_i, Y'_i$ ($i = 1, \dots, n$) — независимо распределенные случайные величины, причем X имеют распределение F , Y — распределение G . Пусть далее, $V_i = 1$, если $\max(X_i, X'_i) < \min(Y_i, Y'_i)$ или $\max(Y_i, Y'_i) < \min(X_i, X'_i)$, и $V_i = 0$ в противном случае. Тогда $p = P(V_i = 1) = \frac{1}{3} + \Delta$.

$< \min(X_i, X'_i)$ и $V_i = 0$ в остальных случаях. Тогда $\sum V_i$ имеет биномиальное распределение с вероятностью p , определенной в задаче 33, и задача сводится к проверке гипотезы $p = 1/3$ при альтернативе $p > 1/3$.

(II) Рассмотреть частные альтернативы, для которых $P(X < Y)$ равна 1 или 0.]

К разделу 9

35. (I) Пусть m и n — число отрицательных и положительных наблюдений среди $Z_1, \dots, Z_N$ и пусть $S_1 < \dots < S_n$ обозначают ранги положительных Z в ряду величин $|Z_1|, \dots, |Z_N|$. Рассмотрим $N + \frac{1}{2}N(N-1)$ различных сумм $Z_i + Z_j$ с $i = j$ и $i \neq j$. Ранговая сумма Вилкоксона $\sum S_j$ равна числу тех сумм $(Z_i + Z_j)$, которые положительны.

(II) Если D есть распределение величин Z , то

$$E\left(\sum S_j\right) = \frac{1}{2}N(N+1) - ND(0) - \frac{1}{2}N(N-1) \int D(-z) dD(z).$$

[(I) Пусть K — число положительных сумм. Поскольку $Z_i + Z_j$ положительно тогда и только тогда, когда наибольшее из чисел $|Z_i|$ и $|Z_j|$ положительно, то $K = \sum_{i=1}^N \sum_{j=1}^N U_{ij}$, где $U_{ij} = 1$, если $Z_j > 0$ и $|Z_i| \leq Z_j$, и $U_{ij} = 0$ в остальных случаях.]

36. Пусть $Z_1, \dots, Z_N$ будет выборка из распределения с плотностью $f(z - \theta)$, где $f(z)$ положительна для всех z и симметрична около 0, и пусть m и n определены, как и в предыдущей задаче.

(I) Распределение величин n и S_j задается формулой

$$P\{\text{число положительных } Z \text{ равно } n \text{ и } S_1 = s_1, \dots, S_n = s_n\} = \\ = \frac{1}{2^N} E \left[\frac{f(V^{(r_1)} + \theta) \dots f(V^{(r_m)} + \theta) f(V^{(s_1)} - \theta) \dots f(V^{(s_n)} - \theta)}{f(V^{(1)}) \dots f(V^{(N)})} \right], \quad (38)$$

где $V^{(1)} < \dots < V^{(N)}$ есть упорядоченная выборка из распределения с плотностью $2f(v)$ для $v > 0$ и 0 в остальных случаях.

(II) Максимизирующий производную функции мощности в точке $\theta = 0$ и, следовательно, максимизирующий мощность для достаточно малых $\theta > 0$ ранговый критерий для гипотезы симметрии распределения относительно начала координат отвергает эту гипотезу, когда

$$-E \left[\sum_{j=1}^n \frac{f'(V^{(s_j)})}{f(V^{(s_j)})} \right] > C.$$

(III) В частном случае, когда $f(z)$ есть плотность нормального распределения с нулевым средним, указанная в п. (II) область сводится к области $\sum E(V^{(s_j)}) > C$, где $V^{(1)} < \dots < V^{(N)}$ есть упорядоченная выборка из χ^2 -распределения с одной степенью свободы.

(IV) Определить плотность f таким образом, чтобы одновыборочный критерий Вилкоксона был наиболее мощным при альтернативе $f(z - \theta)$ для достаточно малых положительных θ .

[(I) Применить задачу 22 (I) к нахождению выражения для условной вероятности $P\{S_1 = s_1, \dots, S_n = s_n \text{ при условии, что число положительных } Z \text{ равно } n\}$.]

37. Другое выражение для (38) получится, если распределение Z описывать тройкой (ρ, F, G) . Если тогда $G = h(F)$ и h дифференцируема, то распределение для n и s_j дается формулой

$$\rho^m (1 - \rho)^n E[h'(U^{(s_1)}) \dots h'(U^{(s_n)})], \quad (39)$$

где $U^{(1)} < \dots < U^{(N)}$ — упорядоченная выборка из $R(0, 1)$.

38. *Несмещенные критерии симметрии.* Пусть $Z_1, \dots, Z_N$ — выборка и Φ — ранговый критерий, служащий для проверки гипотезы симметрии относительно начала, такой, что для всех $z_i \leq z'_i$, $i = 1, \dots, n$, $\Phi(z_1, \dots, z_N) \leq \Phi(z'_1, \dots, z'_N)$. Тогда критерий Φ является несмещенным в задаче проверки одно-сторонней альтернативы о том, что Z стохастически больше, чем некоторая случайная величина, симметрично распределенная относительно начала координат.

39. *Гипотеза случайности.* Пусть $Z_1, \dots, Z_N$ — независимые случайные величины с распределениями $F_1, \dots, F_N$ соответственно, и пусть T_i — ранг величины Z_i среди всех Z . Для проверки *гипотезы случайности*: $F_1 = \dots = F_N$ при альтернативе K о *положительном тренде*, т. е. гипотезы, что Z_i стохастически возрастают с ростом i , рассмотрим две области

$$\sum it_i > C \quad (40)$$

и

$$\sum iE(V^{(t_i)}) > C, \quad (41)$$

где $V^{(1)} < \dots < V^{(N)}$ — упорядоченная выборка из стандартного нормального распределения, а t_i — значения, принимаемые величинами T_i .

(I) Второй из этих критериев является наиболее мощным ранговым критерием для нормальных альтернатив $F = N(\gamma + i\delta, \sigma^2)$ при достаточно малых δ .

(II) Найти альтернативы, для которых первый критерий является наиболее мощным ранговым критерием.

(III) Оба критерия являются несмещенными для альтернатив положительного тренда; таким является каждый ранговый критерий Φ , удовлетворяющий условию $\Phi(z_1, \dots, z_N) \leq \Phi(z'_1, \dots, z'_N)$ для каждой двух точек, для которых из $z_i < z_j$, $i < j$, следует, что $z'_i < z'_j$ для всех i и j .

[(III) Применить задачу 30, взяв в качестве $\mathcal{S}$ класс преобразований $z'_i = z_i$, $z'_i = f_i(z_i)$ для $i > 1$, где $z < f_2(z) < \dots < f_N(z)$ и каждая из функций f_N неубывающая. Если $\mathcal{F}_0$ есть класс N -мерных векторов $(F_1, \dots, F_N)$ с $F_1 = \dots = F_N$, тогда $\mathcal{F}_1$ совпадает с классом альтернатив K .]

40. Пусть $U_{ij} = 1$, если $(j-1)(Z_j - Z_i) > 0$ и 0 в остальных случаях.

(I) Статистика $\sum iT_i$ может быть выражена в терминах величин U посредством соотношения

$$\sum_{i=1}^N iT_i = \sum_{i < j} (j-i) U_{ij} + \frac{N(N+1)(N+2)}{6}.$$

(II) Наименьшее число шагов (в смысле задачи 20 (II)), которыми выборка $(z_1, \dots, z_N)$ может быть преобразована в упорядоченную выборку $(z^{(1)}, \dots, z^{(N)})$, равно $[N(N-1)/2] - U$, где $U = \sum_{i < j} U_{ij}$. Тем самым область $U > C$ может быть использована в качестве области, где отвергается гипотеза, рассмотренная в предыдущей задаче.

[(I) Пусть $V_{ij} = 1$ или 0, когда $Z_i \leq Z_j$ или $Z_i > Z_j$.
Тогда

$$T_j = \sum_{i=1}^N V_{ij}$$

и

$$V_{ij} = U_{ij} \quad \text{или} \quad V_{ij} = 1 - U_{ij},$$

если $i < j$ или $i \geq j$. Выражая $\sum_{j=1}^N jT_j = \sum_{j=1}^N j \sum_{i=1}^N V_{ij}$ в терминах U и используя тот факт, что $U_{ij} = U_{ji}$, требуемый результат получаем с помощью простых вычислений.

41. *Гипотеза независимости.* Пусть $(X_1, Y_1), \dots, (X_N, Y_N)$ — выборка из двумерного распределения и $(X^{(1)}, Z_1), \dots, (X^{(N)}, Z_N)$ — та же выборка, расположенная в порядке возрастания X , тем самым величины Z получают некоторую перестановку из величин Y . Пусть R_i — ранг X_i среди X , S_i — ранг Y_i среди Y и T_i — ранг Z_i среди Z . Рассмотрим гипотезу о независимости величин X и Y при альтернативе положительной зависимости.

(I) В терминах величин T эта проблема эквивалентна задаче проверки гипотезы случайности Z при альтернативе положительного тренда.

(II) Критерий (40) эквивалентен критерию, отвергающему гипотезу, когда *ранговый коэффициент корреляции*

$$\frac{\sum (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum (R_i - \bar{R})^2 \sum (S_i - \bar{S})^2}} = \frac{12}{N^3 - N} \sum \left(R_i - \frac{N+1}{2} \right) \left(S_i - \frac{N+1}{2} \right)$$

достаточно велик.

(III) Другое выражение для рангового коэффициента корреляции*) дается формулой

$$1 - \frac{6}{N^3 - N} \sum (S_i - R_i)^2 = 1 - \frac{6}{N^3 - N} \sum (T_i - i)^2.$$

(IV) Критерий $U > C$ задачи 40 (II) эквивалентен критерию, согласно которому гипотеза отвергается, когда t -статистика Кендалла $\sum_{i < j} V_{ij}/N(N-1)$

достаточно велика; здесь V_{ij} равно $+1$ или -1 в зависимости от того, положительно или отрицательно $(Y_j - Y_i)(X_j - X_i)$.

(V) Критерии (II) и (IV) для альтернатив положительной зависимости являются несмещенными**).

К разделу 10

42. В примере 11 семейство множеств $S(x, y)$ является классом инвариантных доверительных множеств тогда и только тогда, когда существует множество $\mathcal{R}$ действительных чисел такое, что

$$S(x, y) = \bigcup_{r \in \mathcal{R}} \{(\xi, \eta): (x - \xi)^2 + (y - \eta)^2 = r^2\}.$$

*) Дальнейший материал, посвященный этим статистикам, см. в работе: Hoeffding, A class of statistics with asymptotically normal distributions, Ann. Math. Stat. 19 (1948), особенно раздел 9, и в книге: Kendall, Rank Correlation Methods, London, Charles Griffin and Co., 2nd ed., 1955. (Русский перевод: М. Кендэл, Ранговые корреляции, пер. с англ., М., Статистика, 1975. — Прим. пер.)

**) Другой тип критерия, в котором играют особенную роль крайние значения наблюдений, был предложен Olmstead'ом и Tukey, A corner test for association, Ann. Math. Stat. 18 (1947), 495—513.

43. Пусть $X_1, \dots, X_n; Y_1, \dots, Y_n$ — выборки из нормальных распределений $N(\xi, \sigma^2)$ и $N(\eta, \tau^2)$ соответственно. Тогда доверительные интегралы (43) из главы 5 для τ^2/σ^2 , которые можно записать в виде

$$\frac{\sum (Y_i - \bar{Y})^2}{k \sum (X_i - \bar{X})^2} \leq \frac{\tau^2}{\sigma^2} \leq \frac{k \sum (Y_i - \bar{Y})^2}{\sum (X_i - \bar{X})^2},$$

являются равномерно наиболее точными среди инвариантных относительно наименьшей группы G , содержащей преобразования $X'_i = aX + b$, $Y'_i = aY + c$ для всех $a \neq 0$, b, c и преобразования $X'_i = dY_i$, $Y'_i = X_i/d$ для всех $d \neq 0$ [см. задачу 6].

44. *Односторонние инвариантные доверительные пределы.* Пусть θ — действительный параметр. Предположим, что для каждого θ_0 задача проверки гипотез $\theta \leq \theta_0$ и $\theta > \theta_0$ (в присутствии мешающих параметров Φ) остается инвариантной относительно группы G_{θ_0} , и пусть $A(\theta_0)$ есть область принятия, соответствующая РНМ инвариантному критерию с уровнем α . Пусть соответствующие доверительные множества $S(x) = \{\theta: x \in A(\theta)\}$ будут односторонними интервалами $S(x) = \{\theta: \theta(x) \leq \theta\}$. Предположим, что они инвариантны при всех преобразованиях G_θ и, следовательно, относительно группы G , порожденной ими. Тогда нижние доверительные границы $\underline{\theta}(X)$ являются равномерно наиболее точными среди инвариантных границ с доверительным уровнем $1 - \alpha$ (в смысле минимизации $P_{\theta, \theta'} \{\underline{\theta}(X) \leq \theta'\}$ для всех $\theta' < \theta$).

45. Пусть $X_1, \dots, X_n$ — независимые случайные величины, распределенные по закону $N(\xi, \sigma^2)$. Верхние доверительные границы $\sigma^2 \leq \sum (X_i - \bar{X})^2 / C_0$ примера 5 главы 5 являются равномерно наиболее точными инвариантными относительно группы $X'_i = X_i + c$, $-\infty < c < \infty$ границами. Они также инвариантны (и, следовательно, являются равномерно наиболее точными инвариантными границами) относительно более широкой группы $X'_i = aX_i + c$, $-\infty < a, c < \infty$.

46. (I) Пусть случайные величины $X_1, \dots, X_n$ независимы и распределены по закону $N(\xi, \sigma^2)$, и пусть $\theta = \xi/\sigma$. Нижние доверительные границы $\underline{\theta}$ для θ с доверительным уровнем $1 - \alpha$, являющиеся равномерно наиболее точными инвариантными границами при преобразованиях $X'_i = aX_i$ имеют вид

$$\theta = C^{-1} \left[\frac{\sqrt{n} \bar{X}}{\sqrt{\frac{\sum (X_i - \bar{X})^2}{(n-1)}}} \right],$$

где функция $C(\theta)$ определяется с помощью таблиц нецентрального t -распределения так, чтобы

$$P_\theta \left\{ \frac{\sqrt{n} \bar{X}}{\sqrt{\frac{\sum (X_i - \bar{X})^2}{(n-1)}}} \leq C(\theta) \right\} = 1 - \alpha.$$

(II) Определить $\underline{\theta}$, когда значения x равны 7,6; 21,2; 15,1; 32,0; 19,7; 25,3; 29,1; 18,4, а доверительный уровень равен $1 - \alpha = 0,95$.

47. (I) Пусть $(X_1, Y_1), \dots, (X_n, Y_n)$ — выборка из двумерного нормального распределения, и пусть

$$\rho = C^{-1} \left\{ \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \right\},$$

где $C(\rho)$ определено так, что

$$P_\rho \left\{ \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \leq C(\rho) \right\} = 1 - \alpha$$

Тогда ρ является нижней доверительной границей для теоретического коэффициента корреляции ρ с доверительным уровнем $1 - \alpha$; эта граница равномерно наиболее точная среди инвариантных относительно группы преобразований

$$X'_i = aX_i + b, \quad Y'_i = cY_i + d \quad \text{с} \quad ac > 0, \quad -\infty < b, \quad d < \infty.$$

(II) Определить ρ , если $1 - \alpha = 0,95$ и наблюдения таковы:

$$(12,9; 0,56), \quad (9,8; 0,92), \quad (13,1; 0,42), \quad (12,5; 1,01), \\ (8,7; 0,63), \quad (10,7; 0,58), \quad (9,3; 0,72), \quad (11,4; 0,64).$$

К разделу 11

48. Если доверительные множества $S(x)$ инвариантны относительно группы G , тогда вероятность $P_\theta \{ \theta \in S(X) \}$ того, что истинное значение параметра будет накрыто этими множествами, инвариантна относительно индуцируемой группы $\bar{G}$.

49. Рассматривается задача построения (двусторонних) доверительных полос для неизвестной непрерывной функции распределения F .

(I) Показать, что эта задача инвариантна как при строго возрастающих, так и при строго убывающих непрерывных преобразованиях $X'_i = f(X_i), i = 1, \dots, n$. Определить максимальный инвариант, соответствующий этой группе.

(II) Показать, что задача не инвариантна при преобразовании

$$X'_i = \begin{cases} X_i, & \text{если} \quad |X_i| \geq 1, \\ X_i - 1, & \text{если} \quad 0 < X_i < 1, \\ X_i + 1, & \text{если} \quad -1 < X_i < 1. \end{cases}$$

[(II) Для этого преобразования g множество $g^*S(x)$ уже не будет полосой.]

13. Литературные ссылки

Соображения инвариантности использовались в специальных классах задач Хотеллингом и Питменом (см. ссылки в главе I). Общая теория инвариантных и почти инвариантных критериев вместе с главными параметрическими приложениями была дана Хантом и Стейном в неопубликованной работе (1946). В их работе инвариантность не рассматривалась как свойство, желательное само по себе, но как средство получения наиболее строгих критериев (см. главу 8). За исключением этого

расхождения в точке зрения, изложение здесь основано на идеях Ханта и Стейна, с которыми я познакомился во время бесед с Чарлзом Стейном в период 1947—1950.

Область непараметрической статистики, в которой многие основные проблемы еще не решены, находится в настоящее время в процессе стремительного развития. Здесь возможно указать только некоторые из сделанных работ, в частности потому, что многие главные результаты относятся к области больших выборок. Более детальное освещение некоторых аспектов непараметрической статистики дано в книгах Фрезера (1957) и Кендалла (1955) и в обзорных статьях Шеффэ (1943), Уилкса (1948), Вольфовица (1949), Морана, Уайтфилда и Даниелса (1950), Кендалла и Сандрума (1953), ван Данцига и Хемельрийка (1954). Обширная библиография приведена у Сэвиджа (1953).

Андерсон (Anderson T. W.)

(1958) *An Introduction to Multivariate Statistical Analysis*, New York, John Wiley and Sons, 99. [Задача 11.] (Русский перевод: Андерсон, Введение в многомерный статистический анализ, М., Физматгиз, 1963.)

Арнольд (Arnold K. J.)

(1951) *Tables to facilitate sequential t -tests*, Appl. Math. Ser. Nat. Bur. Standards (U. S.), vol. 7, V—VIII. [Задача 9 (II).]

Барнард (Barnard G.)

(1950) *The Behrens — Fisher test*, *Biometrika*, vol. 37, 203—207.

Вальд (Wald A.)

(1955) *Testing the difference between the means of two normal populations with unknown standard deviations*, опубликовано посмертно в *Selected Papers in Statistics and Probability* by Abraham Wald, Stanford, Stanford Univ. Press. [Рассматриваются инвариантные критерии для проблемы Беренса — Фишера в случае выборок равного объема.]

ван Данциг и Хемельрийк (van Dantzig D. and Hemelrijk I.)

(1954) *Statistical methods based on few assumptions*, *Bull. Int. Stat. Inst.*, vol. 34, 2nd part, 3—31.

Ван-дер-Варден (van der Waerden B. L.)

(1952, 1953) *Order tests for the two-sample problem and their power*, *Proc. Koninkl. Ned. Akad. Wetenschap.*, vol. 55, 453—458 and vol. 56, 303—316. [Для задачи двух выборок предлагается критерий, основанный на функции, обратной к нормальной функции распределения.]

(1960) *Математическая статистика*, пер. с нем., М., ИЛ.

Вилкоксон (Wilcoxon F.)

(1945) *Individual comparisons by ranking methods*, *Biometrics*, vol. 1, 80—83. [Предлагаются два критерия, носящие имя автора (см. также Дейхлер).]

Вольфовиц (Wolfowitz J.)

(1949) *Non-parametric statistical inference*, *Proc. Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Univ. Calif. Press, 93—113.

(1949) *The power of the classical tests associated with the normal distribution*, *Ann. Math. Stat.*, vol. 20, 540—551. [Доказана лемма 2 для нескольких специальных случаев.]

Гаеки и Шидак (Hájek J., Sidak Z.)

(1967) Теория ранговых критериев, пер. с англ., М., Наука.

Дейхлер (Deuchler G.)

(1914) Über die Methoden der Korrelationsrechnung in der Paedagogik und Psychologie, Z pädag. Psychol., vol. 15, 114—131, 145—159, 229—242. [Кажется, что здесь впервые предлагается для задачи двух выборок процедура, известная как критерий Вилкоксона, который позже был независимо открыт многими авторами. Об истории этого критерия см. W. H. Kruskal, Historical notes on the Wilcoxon unpaired two-sample test, J. Am. Stat. Assoc., vol. 52 (1957), 356—360.]

Кендалл (Kendall M. G.)

(1955) Rank Correlation Methods, London, Charles Griffin and Co., 2nd ed. (Русский перевод: М. Кендэл, Ранговые корреляции, пер. с англ., М., Статистика, 1975.)

Кендалл и Сандрум (Kendall M. G. and Sundrum R. M.)

(1953) Distribution-free methods and order properties, Rev. Int. Stat. Inst., vol. 23, 124—134.

Кендалл и Стьюарт (Kendall M. G., Stuart A.)

(1973) Статистические выводы и связи, пер. с англ., М., Наука.

Крускал (Kruskal W.)

(1954) The monotonicity of the ratio of two non-central t density functions, Ann. Math. Stat., vol. 25, 162—165.

Леман (Lehmann E. L.)

(1950) Some principles of the theory of testing hypotheses, Ann. Math. Stat., vol. 21, 1—26. [Лемма 2; теорема 6; приводится пример Стейна, взятый за образец для задачи 18.]

(1951) Consistency and unbiasedness of certain nonparametric tests, Ann. Math. Stat., vol. 22, 165—179. [Задачи 33, 34.]

(1953) The power of rank tests, Ann. Math. Stat., vol. 24, 28—43. [Соображения инвариантности применяются к непараметрическим задачам.]

Моран, Уайтфилд и Даниелс (Moran P. A. P., Whitfield J. W. and Daniels H. E.)

(1950) Symposium on ranking methods, J. Roy Stat. Soc., Ser. B., vol. 12, 153—191.

Питмен (Pitman E. J. G.)

(1939) Tests of hypotheses concerning location and scale parameters, Biometrika, vol. 31, 200—215. [Вводятся идеи инвариантности и применяются к задачам, похожим на приведенную в примере 4.]

(1949) Lecture notes on nonparametric statistical inference, неопубликовано. [Исследуется понятие относительной асимптотической эффективности. Оно применяется к нескольким примерам, в том числе к критерию Вилкоксона.]

Питчер (Pitcher T. S.)

(1957) Trans. Amer. Math. Soc., 85, 166—173. [О достаточных статистиках в связи с теорией инвариантности.]

Раштон (Rushton S.)

(1950) On a sequential t -test, Biometrika, vol. 37, 326—333.

(1952) On a two-sided sequential t -test, Biometrika, vol. 39, 302—308. [Предлагается последовательный t -критерий задач 9 и 10 и обсуждаются некоторые из его свойств. Имеется ссылка на близкую по теме неопубликованную работу Барнарда, Голдберга и Стейна.]

Сукхатм (Sukhatme P. V.)

(1936) On the analysis of k samples from exponential distributions with expected reference to the problem of random intervals, Stat. Res. Mem., vol. 1, 94—112.

Сэвидж (Savage I. R.)

(1953) Bibliography of nonparametric statistics and related topics, J. Am. Stat. Assoc., vol. 48, 844—906.

Сюй (Hsu P. L.)

(1938) Contributions to the theory of Student's t -test as applied to the problem of two samples, Stat. Res. Mem., vol. II, 1—24. [Показано, что двухвыборочный t -критерий в случае равных и не очень малых объемов выборок является приближенно несмещенным, даже если дисперсии не равны, и что, следовательно, для этого случая t -критерий приближенно решает проблему Беренса — Фишера.]

Уилкс (Wilks, S. S.)

(1948) Order statistics, Bull. Am. Math. Soc., vol. 54, 6—50.

Уолш (Walsh, John E.)

(1949) Some significance tests for the median which are valid under very general conditions, Ann. Math. Stat., vol. 20, 64—81. [Лемма 3; предлагается критерий Вилкоксона в форме, данной в задаче 35. Эквивалентность двух критериев была показана Тьюки в неопубликованном отчете (1949 г.).]

Фишер (Fisher R. A.)

(1956) Statistical Methods and Scientific Inference, Edinburgh and London, Oliver and Boyd. [В главе IV автор формулирует свою точку зрения на проверку гипотез, в частности обсуждается проблема Беренса — Фишера.]

Фишер и Иэйтс (Fisher R. A. and Yates F.)

(1948) Statistical Tables for Biological, Agricultural and Medical Research, London, Oliver and Boyd, 3rd ed. [Во введении к таблицам XX и XXI рассматриваются ранговые критерии, подобные (19).]

Фрезер (Fraser D. A. S.)

(1957) Nonparametric Methods in Statistics, New York, John Wiley and Sons.

Хант и Стейн (Hunt G. and Stein C.)

(1946) Most stringent tests of statistical hypotheses, неопубликовано.

Хемельрийк (Hemelrijk J.)

(1950) A family of parameter-free tests for symmetry with respect to a given point, Proc. Koninkl. Ned. Akad. Wetenschap., vol. 53, 945—955, 1186—1198. [Обсуждается связь гипотезы симметрии с проблемой двух выборок.]

Хёфдинг (Hoeffding W.)

(1951) «Optimum» nonparametric tests, Proc. 2nd Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, Univ. Calif. Press, 83—92. [Выводится общее распределение для рангов, частным случаем которого является формула (18). На этой основе строятся локально оптимальные критерии типа (19). Результаты автора применены к проблеме двух выборок в работе: Milton E. T., Some rank order tests which are most powerful against specific parametric alternatives, Ann. Math. Stat., vol. 23 (1952), 346—366.]

Хопф (Hopf E.)

(1937) Ergodentheorie, Ergeb. Math., vol. 5, № 2. [Доказывается результат, очень сходный с теоремой 4 (стр. 9—10).]

Шеффе (Scheffé H.)

(1943) Statistical inference in the non-parametric case, Ann. Math. Stat., vol. 14, 305—332.

(1943) On solutions of the Behrens—Fisher problem, based on the t -distribution, Ann. Math. Stat., vol. 14, 35—44.

Шметтерер (Schmetterer L.)

(1976) Введение в математическую статистику, пер. с нем., М., Наука, гл. 7, §§ 4, 6.

Эпштейн и Цзао (Epstein B. and Tsao C. K.)

(1953) Some tests based on ordered observations from two exponential populations, Ann. Math. Stat., vol. 24, 458—466.

Юсава (Usawa H.)

(1960) Locally most powerful rank tests for two-sample problems, Ann. Math. Stat., 31, 685—702. [В работе, в частности, обобщаются результаты работы Лемана (1953).]

ГЛАВА 7

ЛИНЕЙНЫЕ ГИПОТЕЗЫ

1. Каноническая форма

Многие задачи проверки гипотез касаются средних значений нормальных распределений и являются специальными случаями следующей *общей одномерной линейной гипотезы*. Пусть $X_1, \dots, X_n$ независимы и распределены нормально со средними значениями $\xi_1, \dots, \xi_n$ и общей дисперсией σ^2 . Известно, что вектор $\underline{\xi}$, образованный из средних значений ξ_i), лежит в данном s -мерном линейном подпространстве Π_Ω ($s < n$), и проверяемая гипотеза H состоит в том, что $\underline{\xi}$ лежит в данном $(s - r)$ -мерном подпространстве Π_ω , входящем в Π_Ω ($r \leq s$).

Пример 1. В задаче о проверке равенства средних в двух выборках из нормальных распределений (рассмотренной в других обозначениях в разделе 3 главы 5) дано, что $\xi_i = \xi$ при $i = 1, \dots, n_1$ и $\xi_i = \eta$ при $i = n_1 + 1, \dots, n_1 + n_2$, и гипотеза, подлежащая проверке: $\eta = \xi$. Пространство Π_Ω натянуто на векторы $(1, \dots, 1, 0, \dots, 0)$ и $(0, \dots, 0, 1, \dots, 1)$, т. е. $s = 2$ и его точки суть

$$(\xi, \dots, \xi, \eta, \dots, \eta) = \xi(1, \dots, 1, 0, \dots, 0) + \eta(0, \dots, 0, 1, \dots, 1).$$

Аналогично, Π_ω состоит из всех векторов $(\xi, \dots, \xi) = \xi(1, \dots, 1)$, и, следовательно, $r = 1$.

Другая гипотеза, которую можно проверять в этой ситуации, $\eta = \xi = 0$. Пространство Π_ω состоит из одной точки — начала координат, $s - r = 0$, т. е. $r = 2$. Более общая гипотеза $\xi = \xi_0$, $\eta = \eta_0$ не будет линейной, так как Π_ω не содержит начала координат. Однако она сводится к предыдущему случаю преобразованием $X'_i = X_i - \xi_0$ ($i = 1, \dots, n_1$), $X'_i = X_i - \eta_0$ ($i = n_1 + 1, \dots, n_1 + n_2$).

Пример 2. Задача о регрессии из раздела 6 главы 5 дает другой пример линейной гипотезы. Чтобы согласовать обозначения с употребляемыми здесь, положим $\xi_i = \alpha + \beta t_i$, где α, β неизвестны, а t_i известны и не все равны между собой. Здесь Π_Ω — пространство векторов $\alpha(1, \dots, 1) + \beta(t_1, \dots, t_n)$, и его размерность равна 2. Проверяемая гипотеза или может быть

*) На протяжении этой главы предполагается, что в n -мерном пространстве фиксирована некоторая система координат. Вектор с компонентами $\xi_1, \dots, \xi_n$ обозначается $\underline{\xi}$, а $n \times 1$ -матрица-столбец с элементами $\xi_1, \dots, \xi_n$ обозначается ξ .

$\alpha = \beta = 0$ ($r = 2$), или может состоять в том, что один из параметров равен нулю ($r = 1$). Более общая гипотеза $\alpha = \alpha_0$, $\beta = \beta_0$ сводится к предыдущей заменой $X'_i = X_i - \alpha_0 - \beta_0 t_i$, так как $E(X'_i) = \alpha' + \beta' t_i$ с $\alpha' = \alpha - \alpha_0$, $\beta' = \beta - \beta_0$.

Полиномиальная регрессия и регрессия при нескольких переменных также охватываются схемой линейной гипотезы. Так, если $\xi_i = \alpha + \beta t_i + \gamma t_i^2$ или, более общим образом, $\xi_i = \alpha + \beta t_i + \gamma u_i$, где t_i и u_i известны, то можно проверять равенство нулю одного или нескольких из коэффициентов регрессии α , β , γ , и с помощью замены $X'_i = X_i - \alpha_0 - \beta_0 t_i - \gamma_0 u_i$ — гипотезы о равенстве этих коэффициентов некоторым заданным значениям, отличным от нуля.

В общем случае гипотезе может быть придан более простой вид ортогональным переходом к новым переменным

$$Y = CX, \quad C = (c_{ij}); \quad i, j = 1, \dots, n, \quad (1)$$

таким, что первые s векторов-строк $c_1, \dots, c_s$ матрицы C порождают Π_Ω , а $c_{s+1}, \dots, c_s$ порождают Π_ω . Тогда $Y_{s+1} = \dots = Y_n = 0$ равносильно тому, что X лежит в Π_Ω , а $Y_1 = \dots = Y_r = Y_{s+1} = \dots = Y_n = 0$ — тому, что X лежит в Π_ω . Пусть $\eta_i = E(Y_i)$, так что $\eta = C\xi$. Так как ξ лежит в Π_Ω по предположению и лежит в Π_ω при гипотезе H , то $\eta_i = 0$ при $i = s+1, \dots, n$ в обоих случаях и $\eta_i = 0$ при $i = 1, \dots, r$ в случае истинности H . Наконец, раз преобразование ортогональное, то случайные величины $Y_1, \dots, Y_n$ снова распределены нормально с общей дисперсией σ^2 , т. е. проблема сводится к следующей канонической форме.

Случайные величины $Y_1, \dots, Y_n$ независимы и распределены нормально с общей дисперсией σ^2 и средними $E(Y_i) = \eta_i$ при $i = 1, \dots, s$ и $E(Y_i) = 0$ при $i = s+1, \dots, n$, так что совместная плотность имеет вид

$$\frac{1}{(\sqrt{2\pi} \sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^s (y_i - \eta_i)^2 + \sum_{i=s+1}^n y_i^2 \right) \right]. \quad (2)$$

Величины η и σ^2 неизвестны и проверяемая гипотеза

$$H: \quad \eta_1 = \dots = \eta_r = 0 \quad (r \leq s < n). \quad (3)$$

Пример 3. Чтобы проиллюстрировать преобразование (1), рассмотрим еще раз модель регрессии $\xi_i = \alpha + \beta t_i$ из примера 2. Мы видели, что Π_Ω натянуто на $(1, \dots, 1)$ и $(t_1, \dots, t_n)$. Если проверяется гипотеза $\beta = 0$, то Π_ω — пространство, порожденное первым из этих векторов. Вектор-строка c_2 лежит в Π_ω и имеет длину 1, следовательно, $c_2 = (1/\sqrt{n}, \dots, 1/\sqrt{n})$. Так как c_1 лежит в Π_Ω , имеет длину 1 и ортогонален к c_2 , то его координаты должны иметь вид $a + bt_i$, $i = 1, \dots, n$, где a и b определяются соотношениями $\sum (a + bt_i) = 0$ и $\sum (a + bt_i)^2 = 1$. Решая эти уравнения, находим $a = -b\bar{t}$, $b = 1/\sqrt{\sum (t_j - \bar{t})^2}$, следовательно, $a + bt_i = (t_i - \bar{t})/\sqrt{\sum (t_j - \bar{t})^2}$

и

$$Y_1 = \frac{\sum X_i (t_i - \bar{t})}{\sqrt{\sum (t_j - \bar{t})^2}} = \frac{\sum (X_i - \bar{X}) (t_i - \bar{t})}{\sqrt{\sum (t_j - \bar{t})^2}}.$$

Остальные векторы-строки C можно взять любыми при условии, что они ортогональны, имеют длину 1 и ортогональны к Π_Q ; их явные выражения нам не нужны.

Если проверяется гипотеза $\alpha = 0$, то Π_Q порождается вектором $(t_1, \dots, t_n)$, так что i -я координата c_2 равна $t_i / \sqrt{\sum t_j^2}$. Координаты c_1 снова имеют вид $a + bt_i$, где a и b определяются теперь из уравнений $\sum (a + bt_i) t_i = 0$ и $\sum (a + bt_i)^2 = 1$. Решая их, находим $b = -a\bar{t} / \sum t_j^2$, $a = \sqrt{\sum t_j^2 / n \sum (t_j - \bar{t})^2}$ и, следовательно,

$$Y_1 = \sqrt{\frac{n \sum t_j^2}{\sum (t_j - \bar{t})^2}} \left(\bar{X} - \frac{\bar{t}}{\sum t_j^2} \sum t_i X_i \right).$$

В случае гипотезы $\alpha = \beta = 0$ Π_Q состоит из начала координат и в качестве c_1, c_2 можно взять два любых ортогональных единичных вектора из Π_Q . Возможен в частности, такой же выбор, как при гипотезе $\beta = 0$. Тогда Y_1 — линейная функция, указанная ранее, а $Y_2 = \sqrt{n} \bar{X}$.

Общая линейная гипотеза, записанная в терминах Y_k , остается инвариантной относительно группе G_1 преобразований $Y'_i = Y_i + c_i$ при $i = r+1, \dots, s$; $Y'_i = Y_i$ при $i = 1, \dots, r$; $s+1, \dots, n$. При этом $Y_1, \dots, Y_r$ и $Y_{s+1}, \dots, Y_n$ — максимальные инварианты. Другая группа, составляющая задачу инвариантной, — группа G_2 всех ортогональных преобразований переменных $Y_1, \dots, Y_r$. Так как величины с промежуточными индексами не затрагиваются, то из примера 1 (III) главы 6 видно, что максимальными инвариантами относительно G_2 будут

$$U = \sum_{i=1}^r Y_i^2, Y_{s+1}, \dots, Y_n. \text{ По соображениям достаточности их}$$

можно свести к U и $V = \sum_{i=s+1}^n Y_i^2$. Наконец, задача остается инвариантной относительно группы G_3 изменений масштаба $Y'_i = cY_i$, $c \neq 0, i = 1, \dots, n$. В пространстве U и V эта группа индуцирует преобразование $U^* = c^2 U$, $V^* = c^2 V$, относительно которого $W = U/V$ является максимальным инвариантом. Таким образом, принцип инвариантности сводит данные к единственной статистике

$$W = \frac{\sum_{i=1}^r Y_i^2}{\sum_{i=s+1}^n Y_i^2}. \quad (4)$$

Каждая из трех групп преобразований $G_i (i = 1, 2, 3)$, осуществляющих вышеуказанную редукцию, индуцирует соответствующую группу $\bar{G}_i$ в пространстве параметров. Группа $\bar{G}_i$ — это группа сдвигов $\eta'_i = \eta_i + c_i$ ($i = r + 1, \dots, s$), $\eta'_i = \eta_i$ ($i = 1, \dots, r$), $\sigma' = \sigma$ с максимальными инвариантами $(\eta_1, \dots, \eta_r, \sigma)$. Так как каждое ортогональное преобразование $Y_1, \dots, Y_r$ индуцирует такое же преобразование величин $\eta_1, \dots, \eta_r$ и оставляет σ^2 неизменным, то максимальным инвариантом относительно $\bar{G}_2$ будет $(\sum_{i=1}^r \eta_i^2, \sigma^2)$. Наконец, элементами $\bar{G}_3$ будут преобразования $\eta'_i = c\eta_i$, $\sigma' = c\sigma$, и поэтому максимальным инвариантом по отношению к этим трем группам преобразований будет

$$\psi^2 = \frac{\sum_{i=1}^r \eta_i^2}{\sigma^2}. \quad (5)$$

Из теоремы 3 главы 6 вытекает, что распределение W зависит только от ψ^2 , так что принцип инвариантности сводит дело к проверке простой гипотезы $H: \psi = 0$. Точное выражение для плотности W имеет вид (см. задачи 2 и 3):

$$p_\psi(w) = e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} c_k \frac{\left(\frac{1}{2}\psi^2\right)^k}{k!} \frac{w^{\frac{1}{2}r-1+k}}{(1+w)^{\frac{1}{2}(r+n-s)+k}}, \quad (6)$$

где

$$c_k = \frac{\Gamma\left[\frac{1}{2}(r+n-s)+k\right]}{\Gamma\left(\frac{1}{2}r+k\right)\Gamma\left[\frac{1}{2}(n-s)\right]}.$$

Для любого ψ_1 отношение $p_{\psi_1}(w)/p_0(w)$ является возрастающей функцией w и из фундаментальной леммы Неймана — Пирсона вытекает, что наиболее мощный инвариантный критерий для проверки $\psi = 0$ при альтернативе $\psi = \psi_1$ отклоняет гипотезу, если W слишком велико, т. е. если

$$W^* = \frac{\sum_{i=1}^r Y_i^2/r}{\sum_{i=s+1}^n Y_i^2/(n-s)} > C. \quad (7)$$

Граничная точка C определяется из условия, что при $\psi = 0$ вероятность отклонения равна α . Так как в этом случае W^* будет отношением двух независимых χ^2 величин, каждая из которых поделена на соответствующее число степеней свободы, то

распределением W служит F -распределение с r и $n - s$ степенями свободы, а C определяется из уравнения

$$\int_C^\infty F_{r, n-s}(y) dy = \alpha. \quad (8)$$

Критерий не зависит от ψ_1 и потому является РНМ в классе всех инвариантных критериев. По теореме 5 главы 6 он будет также РНМ в классе всех критериев, функция мощности которых зависит только от ψ^2 .

Критическую область (7) можно также определить неравенством

$$\frac{\sum_{i=1}^r Y_i^2}{\sum_{i=1}^r Y_i^2 + \sum_{i=s+1}^n Y_i^2} > C'. \quad (9)$$

При $\psi = 0$ левая часть имеет бета-распределение с r и $n - s$ степенями свободы (см. формулу (24) главы 5), так что C' находится из

$$\int_{C'}^1 B_{\frac{1}{2}r, \frac{1}{2}(n-s)}(y) dy = \alpha. \quad (10)$$

При альтернативных значениях ψ левая часть (9) имеет *нецентральное бета-распределение* с параметром нецентральности ψ и с плотностью (задача 3)

$$g_\psi(y) = e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\psi^2\right)^k}{k!} B_{\frac{1}{2}r+k, \frac{1}{2}(n-s)}(y). \quad (11)$$

Мощность критерия при альтернативе ψ равна *)

$$\beta(\psi) = \int_{C'}^1 g_\psi(y) dy.$$

В частном случае, при $r = 1$, критическая область (7) сводится к

$$\frac{|Y_1|}{\sqrt{\sum_{i=s+1}^n Y_i^2 / (n-s)}} > C_0. \quad (12)$$

*) Номограммы для этой мощности даны в работах: Pearson and Hartley, Charts of the power function for analysis of variance tests, derived from the noncentral F distribution, *Biometrika* 38 (1951), 112—130; Fox, Charts of the power of the F -test, *Ann. Math. Stat.* 27 (1956), 484—497. Вычислительная формула для нецентрального бета-распределения рассматривается Ходжесом, On the noncentral beta distribution, *Ann. Math. Stat.* 26 (1955), 648—653.

Это — двусторонний t -критерий, который в соответствии с теорией главы 5 (см., например, задачу 5 этой главы) является РНМ несмещенным. С другой стороны, при $r > 1$ не существует РНМ несмещенного критерия.

2. Линейные гипотезы и метод наименьших квадратов

В приложениях к конкретным проблемам приведение к канонической форме в явном виде проводить неудобно. Статистику W , соответствующую критерию, можно выразить в терминах первоначальных переменных, если заметить, что $\sum_{i=s+1}^n Y_i^2$ является минимальным значением для

$$\sum_{i=1}^n (Y_i - \eta_i)^2 + \sum_{i=s+1}^n Y_i^2 = \sum_{i=1}^n [Y_i - E(Y_i)]^2$$

при неограниченном изменении величин η . Так как преобразование $Y = CX$ ортогонально, а ортогональные преобразования оставляют расстояния неизменными, то

$$\sum_{i=1}^n [Y_i - E(Y_i)]^2 = \sum_{i=1}^n (X_i - \xi_i)^2.$$

Далее, существует взаимно однозначное соответствие между совокупностью всех s -членных наборов $(\eta_1, \dots, \eta_s)$ и совокупностью векторов $\underline{\xi}$ из Π_Ω . Следовательно,

$$\sum_{i=s+1}^n Y_i^2 = \sum_{i=1}^n (X_i - \hat{\xi}_i)^2, \quad (13)$$

где $\hat{\xi}$ — оценки величин ξ при условии Ω , т. е. значения, которые минимизируют $\sum_{i=1}^n (X_i - \xi_i)^2$ при условии, что $\underline{\xi}$ лежит в Π_Ω .

Подобным же образом можно увидеть, что

$$\sum_{i=1}^r Y_i^2 + \sum_{i=s+1}^n Y_i^2 = \sum_{i=1}^n (X_i - \hat{\xi}_i)^2,$$

где величины $\hat{\xi}$ — это значения, минимизирующие $\sum (X_i - \xi_i)^2$ при условии, что $\underline{\xi}$ лежит в Π_ω . Критерий (7) принимает вид

$$W^* = \frac{\left[\sum_{i=1}^n (X_i - \hat{\xi}_i)^2 - \sum_{i=1}^n (X_i - \hat{\xi}_i)^2 \right] / r}{\sum_{i=1}^n (X_i - \hat{\xi}_i)^2 / (n - s)} > C, \quad (14)$$

где C определяется из условия (8). Геометрически векторы $\hat{\xi}$ и $\hat{\xi}$ являются проекциями $\underline{X}$ на Π_Ω и Π_ω , так что треугольник, образованный $\underline{X}$, $\hat{\xi}$ и $\hat{\xi}$ имеет прямой угол при вершине $\hat{\xi}$ (рис. 7). Таким образом, знаменатель и числитель W^* , если отвлечься от множителей $1/(n-s)$ и $1/r$, являются квадратами расстояний

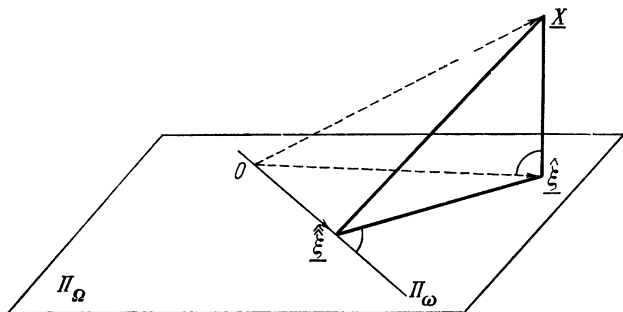


Рис. 7.

между $\underline{X}$ и $\hat{\xi}$ и между $\hat{\xi}$ и $\hat{\xi}$ соответственно. Поэтому альтернативным выражением для W^* будет

$$W^* = \frac{\sum_{i=1}^n (\hat{\xi}_i - \hat{\xi}_i)^2 / r}{\sum_{i=1}^n (X_i - \hat{\xi}_i)^2 / (n-s)}. \quad (15)$$

Желательно также выразить в терминах ξ параметр нецентральности $\psi^2 = \sum_{i=1}^r \eta_i^2 / \sigma^2$. Мы имеем $X = C^{-1}Y$, $\xi = C^{-1}\eta$ и

$$\sum_{i=1}^r Y_i^2 = \sum_{i=1}^n (X_i - \hat{\xi}_i)^2 - \sum_{i=1}^n (X_i - \hat{\xi}_i)^2. \quad (16)$$

Если правую часть (16) обозначить $f(X)$, то $\sum_{i=1}^r \eta_i^2 = f(\xi)$. Некоторым обобщением линейных гипотез являются неоднородные гипотезы типа: вектор $\underline{\xi}$ лежит в гиперплоскости Π'_ω , содержащейся в Π_Ω и не проходящей через начало. Обозначим Π_ω подпространство Π_Ω , проходящее через начало и параллельное Π'_ω . Если ξ^0 некоторая точка Π'_ω , то множество Π'_ω состоит из всех точек вида $\xi = \xi^* + \xi^0$, где ξ^* пробегает Π_ω . Применяя преобразование (1) к Π_ω , вектор средних значений $\underline{\eta}$ для $\xi \in \Pi'_\omega$

можно представить в канонической форме (2) $\eta = C\xi = C\xi^* + C\xi^0$, и совокупность этих векторов описывается уравнениями $\eta_1 = \eta_1^0, \dots, \eta_r = \eta_r^0, \eta_{s+1} = \dots = \eta_n = 0$, где η_i^0 — i -я координата $C\xi^0$. Поэтому в канонической форме неоднородная гипотеза $\xi \in \Pi_\omega$ принимает вид $\eta_i = \eta_i^0$ ($i = 1, \dots, r$). Этот случай приводится к однородному заменой Y_i на $Y_i - \eta_i^0$ и, как следует из (7), РНМ несмещенный критерий имеет критическую область

$$\sum_{i=1}^r (Y_i - \eta_i^0)^2 / r \Big/ \sum_{i=s+1}^n Y_i^2 / (n - s) > C \quad (17)$$

и параметр нецентральности равен $\psi^2 = \sum_{i=1}^r (\eta_i - \eta_i^0)^2 / \sigma^2$.

В приложениях обычно бывает более удобным применять преобразование $X_i - \xi_i^0$ непосредственно к (14) или (15). Из (17) видно, что такое преобразование всегда оставляет знаменатель неизменным. Это можно понять и геометрически, так как рассматриваемое преобразование представляет собой сдвиг пространства параллельно Π_Ω и потому оставляет расстояние $\Sigma (X_i - \xi_i)^2$ от X до Π_Ω неизменным.

Параметр нецентральности может быть вычислен, как и раньше, заменой X на ξ в преобразованном числителе (16).

Некоторые примеры линейной гипотезы, все с $r = 1$, уже обсуждались в главе 5. В дальнейшем два из них рассматриваются с новой точки зрения.

Пример 4. Пусть $X_1, \dots, X_n$ независимы и распределены нормально с общими средним μ и дисперсией σ^2 . Рассмотрим гипотезу $H: \mu = 0$. Здесь Π_Ω — прямая $\xi_1 = \dots = \xi_n$, Π_ω — начало координат, s и r равны 1. Из тождества

$$\sum (X_i - \mu)^2 = \sum (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2 \quad \left(\bar{X} = \frac{\sum X_i}{n} \right)$$

видно, что $\hat{\xi}_i = \bar{X}$, в то время как $\hat{\xi}_i = 0$. Статистика, на которой основан критерий, и ψ^2 равны, следовательно,

$$W = \frac{n\bar{X}^2}{\sum (X_i - \bar{X})^2} \quad \text{и} \quad \psi^2 = \frac{n\mu^2}{\sigma^2}.$$

При истинности гипотезы распределение $(n-1)W$ таково же, как у квадрата величины, имеющей t -распределение Стьюдента с $n-1$ степенями свободы.

Пример 5. В задаче двух выборов, рассмотренной в примере 1, сумма квадратов

$$\sum_{i=1}^{n_1} (X_i - \xi)^2 + \sum_{i=n_1+1}^n (X_i - \eta)^2$$

минимизируется значениями

$$\hat{\xi} = X^{(1)} = \sum_{i=1}^{n_1} \frac{X_i}{n_1}, \quad \hat{\eta} = X^{(2)} = \sum_{i=n_1+1}^n \frac{X_i}{n_2}$$

в то время как при гипотезе $\eta - \xi = 0$

$$\hat{\xi} = \hat{\eta} = \bar{X} = \frac{[n_1 X^{(1)} + n_2 X^{(2)}]}{n}.$$

Числитель статистики (15) равен, следовательно,

$$n_1 (X^{(1)} - \bar{X})^2 + n_2 (X^{(2)} - \bar{X})^2 = \frac{n_1 n_2}{n_1 + n_2} [X^{(2)} - X^{(1)}]^2.$$

Более общая гипотеза $\eta - \xi = \theta_0$ сводится к уже рассмотренной заменой X_i на $X_i - \theta_0$ при $i = n_1 + 1, \dots, n$ и, следовательно, отклоняется при

$$\frac{(X^{(2)} - X^{(1)} - \theta_0)^2 / \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}{\left[\sum_{i=1}^{n_1} (X_i - X^{(1)})^2 + \sum_{i=n_1+1}^n (X_i - X^{(2)})^2 \right] / (n_1 + n_2 - 2)} > C.$$

Параметр нецентральности равен $\psi^2 = (\eta - \xi - \theta_0)^2 / \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma^2$. При истинности гипотезы квадратный корень из соответствующей статистики имеет t -распределение с $n_1 + n_2 - 2$ степенями свободы.

3. Критерии однородности

Полученный в предыдущем разделе РНМ инвариантный критерий для проверки равенства средних значений двух нормальных распределений с общей дисперсией оказывается также РНМ несмещенным (см. раздел 3 главы 5). Однако если число сравниваемых совокупностей больше 2, то РНМ несмещенный критерий не существует и принцип инвариантности приводит к новому результату. Пусть X_{ij} ($j = 1, \dots, n_i; i = 1, \dots, s$) независимы и распределены по законам $N(\mu_i, \sigma^2)$ соответственно. Рассмотрим гипотезу

$$H: \mu_1 = \dots = \mu_s.$$

Она возникает, например, при сравнении нескольких способов обработки, процессов, условий, размещений и т. п. с целью выяснить, влияют ли различия между ними на интересующий нас исход X . Более общим образом, она возникает каждый раз в любой ситуации, связанной с *одинарной классификацией* исходов, т. е. такой, где разбиение на классы производится по значениям одного признака.

Гипотеза H является линейной с $r = s - 1$, с Π_Ω , определяемым соотношениями $\xi_{ij} = \xi_{ik}$ при $j, k = 1, \dots, n; i = 1, \dots, s$

с Π_ω , представляемым прямой, на которой все $n = \sum n_i$ координат ξ_{ij} равны между собой. Мы имеем

$$\sum \sum (X_{ij} - \mu_i)^2 = \sum \sum (X_{ij} - X_{i.})^2 + \sum n_i (X_{i.} - \mu_i)^2,$$

где $X_{i.} = \sum_{j=1}^{n_i} X_{ij}/n_i$ и, следовательно, $\hat{\xi}_{ij} = X_{i.}$ Итак,

$$\sum \sum (X_{ij} - \mu)^2 = \sum \sum (X_{ij} - X_{..})^2 + n (X_{..} - \mu)^2$$

с $X_{..} = \sum \sum X_{ij}/n$, так что $\hat{\xi}_{ij} = X_{..}$. Используя формулу (15) статистики W^* , мы получаем критерий

$$W^* = \frac{\sum n_i (X_{i.} - X_{..})^2 / (s-1)}{\sum \sum (X_{ij} - X_{i.})^2 / (n-s)} > C. \quad (18)$$

Параметр нецентральности равен

$$\psi^2 = \sum n_i (\mu_i - \mu)^2 / \sigma^2$$

с

$$\mu = \frac{\sum n_i \mu_i}{n}.$$

Суммы квадратов как в числителе, так и в знаменателе (18) допускают три близкие между собой интерпретации:

I) как две компоненты разложения полной суммы квадратов отклонений от общего среднего («полной изменчивости»)

$$\sum \sum (X_{ij} - X_{..})^2 = \sum \sum (X_{ij} - X_{i.})^2 + \sum n_i (X_{i.} - X_{..})^2,$$

на а) сумму квадратов отклонений каждой величины от соответствующего группового среднего значения («изменчивость внутри групп») и б) сумму квадратов отклонений групповых средних от общего среднего («изменчивость между группами»);

(II) как основу сравнения двух источников изменчивости (по критерию (18));

(III) как оценки для их математических ожиданий $(n-s)\sigma^2$ и $(s-1)\sigma^2 + \sum n_i (\mu_i - \mu)^2$ (см. задачу 9). Это разбиение полной изменчивости вместе с указанными интерпретациями компонент дает пример *дисперсионного анализа*, который к более сложным проблемам будет применен в последующих разделах.

Мы оставим теперь на время схему линейных гипотез и обратимся к гипотезе равенства дисперсий в случае, когда величины X_{ij} распределены по законам $N(\mu_i, \sigma_i^2)$, $i = 1, \dots, s$. РНМ несмещенный критерий для этой гипотезы был найден в разделе 3 главы 5 в предположении $s = 2$; но он не существует при $s > 2$ (см., например, задачу 6 главы 4). К сожалению,

в этой задаче нет и группы, для которой существовал бы РНМ инвариантный критерий. Тем не менее, чтобы получить критерий, мы используем асимптотический метод, который при достаточно больших n сводит задачу, по сути дела, к проверке равенства s средних.

Удобно сначала редуцировать совокупность результатов наблюдений, заменив ее достаточными статистиками $X_{i.} = \sum_j X_{ij}/n_i$

и $S_i^2 = \sum_j (X_{ij} - X_{i.})^2$, $i = 1, \dots, s$. Гипотеза

$$H: \sigma_1 = \dots = \sigma_s$$

остается инвариантной относительно преобразований $X'_{ij} = X_{ij} + c_i$, которые в пространстве достаточных статистик индуцируют преобразования $S_i'^2 = S_i^2$, $X'_{i.} = X_{i.} + c_i$. Максимальным инвариантом относительно этой группы служит совокупность величин $S_1^2, \dots, S_s^2$. Каждая статистика S_i^2 является суммой квадратов $n_i - 1$ независимых нормально распределенных величин с нулевыми средними и дисперсиями σ_i^2 . Из центральной предельной теоремы вытекает, что при больших n_i величина

$$\sqrt{n_i - 1} \left(\frac{S_i^2}{n_i - 1} - \sigma_i^2 \right)$$

распределена приблизительно по закону $N(0, 2\sigma_i^4)$. Такая аппроксимация для наших целей неудобна, так как неизвестные параметры σ_i входят не только в среднее значение, но и в дисперсию предельного распределения.

Эту трудность можно преодолеть с помощью подходящего преобразования, стабилизирующего дисперсию. Подобные преобразования могут быть получены из следующих соображений *). Пусть T_n — последовательность действительных статистик, такая, что $\sqrt{n}(T_n - \theta)$ имеет предельное распределение $N(0, \tau^2)$. Тогда для любой непрерывно дифференцируемой функции f предельное распределение $\sqrt{n}[f(T_n) - f(\theta)]$ нормально со средним нуль и дисперсией $\tau^2(df/d\theta)^2$. Следовательно, в предположении, что производная от $f(\theta)$ пропорциональна $1/\tau(\theta)$, эта дисперсия не зависит от θ .

Применим сказанное к рассматриваемому случаю, полагая $n = n_i - 1$, $T_n = S_i^2/(n_i - 1)$, $\theta = \sigma_i^2$ и $\tau^2 = 2\theta^2$. В качестве преобразования берем $f(\theta) = \log \theta$. Здесь производная пропорциональна $1/\theta$. Тогда предельное распределение для

*) Доказательство см., например, в книге: Rao, Advanced Statistical Methods in Biometric Research, New York John Wiley & Sons, 1952, Section 5e.

$\sqrt{n_i - 1} \{ \log [S_i^2 / (n_i - 1)] - \log \sigma_i^2 \}$ будет нормальным со средним нуль и дисперсией 2, так что при больших n_i случайные величины $Z_i = \log [S_i^2 / (n_i - 1)]$ распределены приблизительно по законам $N(\xi_i, a_i^2)$ с $\xi_i = \log \sigma_i^2$, $a_i^2 = 2 / (n_i - 1)$.

Теперь проблема сводится к проверке равенства средних значений s независимых случайных величин Z_i , имеющих распределение $N(\xi_i, a_i^2)$, где a_i известны. В частном случае, когда все n_i равны друг другу, дисперсии a_i^2 равны между собой и асимптотическая проблема становится простым вариантом (простым в том смысле, что дисперсии известны) задачи, рассмотренной в начале раздела. Гипотеза $\xi_1 = \dots = \xi_s$ инвариантна относительно добавления одной и той же константы к каждой из величин Z и относительно ортогональных преобразований гиперплоскостей, перпендикулярных прямой линии $Z_1 = \dots = Z_s$. РНМ инвариантной критической областью будет

$$\sum (Z_i - \bar{Z})^2 / a^2 > C, \quad (19)$$

где a^2 — общее значение дисперсий Z_i и C определяется из уравнения

$$\int_C^\infty \chi_{s-1}^2(y) dy = \alpha. \quad (20)$$

В более общем случае неравных a_i проблема сводится к линейной гипотезе с известной дисперсией посредством преобразования $Z'_i = Z_i / a_i$ и РНМ инвариантный относительно подходящей группы линейных преобразований критерий отвергает гипотезу, когда

$$\sum \frac{1}{a_i^2} \left(Z_i - \frac{\sum Z_i / a_j^2}{\sum 1/a_j^2} \right)^2 = \sum \left(\frac{Z_i}{a_i} \right)^2 - \frac{(\sum Z_i / a_j^2)^2}{\sum (1/a_j^2)} > C \quad (21)$$

(см. задачу 10), где C опять определяется равенством (20). Эта критическая область РНМ, инвариантная при проверке гипотезы $\xi_1 = \dots = \xi_s$ относительно предельного распределения, может рассматриваться как обладающая этим свойством асимптотически при проверке первоначальной гипотезы $H: \sigma_1 = \dots = \sigma_s^*$).

Этот же метод может быть применен для проверки однородности нескольких биномиальных или пуассоновских распределений. Подробности указаны в задаче 11.

*) Часто в качестве асимптотического критерия для H используется критерий Бартлетта (см., например, раздел 6а цит. книги Rao), который является, по существу, критерием отношения правдоподобия.

В применениях принципа инвариантности важно быть уверенным в том, что предполагаемая симметрия действительно имеет место. Например, при проверке равенства средних значений $\mu_1, \dots, \mu_s$ нормальных совокупностей все параметрические точки с одним и тем же значением $\psi^2 = \sum n_i (\mu_i - \mu.)^2 / \sigma^2$ отождествляются вследствие принципа инвариантности. Это уместно только в предположении, что такие альтернативы можно рассматривать как одинаково удаленные от проверяемой гипотезы. В частности, должно быть несущественным, как получилось данное значение ψ^2 : за счет нескольких небольших слагаемых или за счет одного большого. Ситуации, где напротив, существенно обнаружить большие индивидуальные отклонения, не обладают требуемой симметрией, и критерий, основанный на (18), не обязан более быть оптимальным.

В подобных ситуациях обычно привлекают процедуры более общие, чем проверка отдельной гипотезы. Сравнивая несколько совокупностей или способов обработки, мы, как правило, желали бы не только определить, одинаковы ли они, но, в случае отклонения гипотезы — также упорядочить их или как-то сгруппировать, или, по крайней мере, выделить из них лучшие. Предположим для простоты, что объемы выборок равны $n_i = n$ при $i = 1, \dots, s$. Естественная процедура, приводящая к группировке значений μ_i , состоит в объявлении μ_i и μ_j различными при $|X_j - X_i| > CS / \sqrt{sn(n-1)}$, где $S^2 = \sum \sum (X_{ki} - X_{k.})^2$ *). Гипотеза H равенства всех средних принимается в этом случае, если

$$\frac{\max |X_j - X_i|}{S / \sqrt{sn(n-1)}} \leq C^{**}. \quad (22)$$

Если H отвергается, то считают, что $\mu_j > \mu_i$ для всех пар (i, j) , для которых $X_j - X_i > CS / \sqrt{sn(n-1)}$. Уровень значимости здесь равен вероятности объявления какой-либо из разностей $\mu_j - \mu_i$ значимой, в то время как в действительности все величины μ равны друг другу. Левая часть (22) называется *стюден-товым размахом* выборочных средних.

Аналогичный подход возможен при сравнении нескольких дисперсий. Предположим снова, что объемы выборок равны и

*) Другие типы процедур со многими решениями для этой и более общих линейных гипотез были предложены в числе других в работах: Duncan, Multiple range and multiple F -tests, *Biometrics* 11 (1955); 1—42; Scheffé, A method for judging all contrasts in the analysis of variance, *Biometrika* 40 (1953), 87—104; Tukey, Comparing individual means in the analysis of variance, *Biometrics* 5 (1949), 99—114 и в одной неопубликованной работе.

**) Таблицы C даны в *Biometrika Tables*, vol. 1, Cambridge Univ. Press, 1954, Table 29, и работе May, Extended and corrected tables of the upper percentage points of the studentized range, *Biometrika* 39 (1952), 192—193.

что классификация осуществляется по принципу: $\sigma_i^2 < \sigma_j^2$, если $S_j^2/S_i^2 > C$; $\sigma_i^2 > \sigma_j^2$, если $S_j^2/S_i^2 < 1/C$ ($C > 1$), и $\sigma_i^2 = \sigma_j^2$, если ни одно из этих неравенств не выполняется. Полная гипотеза $\sigma_1 = \dots = \sigma_s$ принимается в этом случае, когда

$$\max_{i, j} \left(\frac{S_j^2}{S_i^2} \right) = \frac{\max_k S_k^2}{\min_k S_k^2} \leq C^* \quad (23)$$

Критерии, основанные на стьюдентовом размахе и на максимуме F -отношений, не обладают, как кажется, никакими оптимальными свойствами, если их рассматривать как критерии для гипотез $\mu_1 = \dots = \mu_s$ и $\sigma_1 = \dots = \sigma_s$ соответственно. Однако они обладают свойствами оптимальности, если их рассматривать как решение проблемы об упорядочении средних или дисперсий (при допущении «ничьих»)**).

4. Двойная классификация: одно наблюдение в клетке

Гипотеза равенства нескольких средних возникает тогда, когда необходимо сравнить несколько различных способов обработки или процедур, или совокупностей, или проявлений каких-либо факторов. Часто интересуются изучением эффекта более чем одного фактора или эффекта одного фактора при изменении некоторых из условий эксперимента (что играет роль дополнительных факторов). В настоящем разделе мы рассмотрим случай, когда число факторов, влияющих на исход эксперимента, равно двум.

Предположим, что при каждой возможной комбинации интенсивностей обоих факторов производится ровно одно наблюдение. Пусть X_{ij} ($i = 1, \dots, a$; $j = 1, \dots, b$) — результат наблюдения, соответствующий случаю, когда первый фактор находится на i -м, а второй на j -м уровне. Предполагается, что X_{ij} независимы и распределены нормально с постоянной дисперсией σ^2 и (на время) что оба фактора действуют независимо (в этом случае говорят, что они *аддитивны*), так что ξ_{ij} имеет вид $\alpha'_i + \beta'_j$. Полагая $\mu = \alpha'_i + \beta'_j$ и $\alpha_i = \alpha'_i - \alpha'$, $\beta_j = \beta'_j - \beta'$, мы можем написать

$$\xi_{ij} = \mu + \alpha_i + \beta_j, \quad \sum \alpha_i = \sum \beta_j = 0, \quad (24)$$

где величины α , β и μ однозначно определяются из (24):

$$\alpha_i = \xi_{i.} - \xi_{..}, \quad \beta_j = \xi_{.j} - \xi_{..}, \quad \mu = \xi_{..} \quad (25)$$

*) Таблицы C даны в Biometrika Tables (цит. книга), Table 31.

**) Lehmann, A theory of some multiple decision problems, Ann. Math. Stat. 28 (1957), 1—25, 547—572.

***) Замена индекса точкой указывает на то, что величина усреднена по этому индексу.

Рассмотрим гипотезу

$$H: \alpha_1 = \dots = \alpha_a = 0$$

о том, что первый фактор не влияет на исходы испытаний. Эта гипотеза возникает в двух различных обстоятельствах. Интересующим нас фактором, соответствующим нескольким способам обработки, может быть β , в то время как α соответствует классификации по месту, где производятся наблюдения (ферма, лаборатория, город и т. п.). Указанная гипотеза соответствует предположению о том, что эта вспомогательная классификация не влияет на результаты эксперимента и, следовательно, не должна приниматься во внимание. При другом истолковании фактором, представляющим главный интерес, является α . В этом случае формулировка проблемы как задачи проверки гипотезы была бы чрезмерным упрощением действительного положения, так как в случае отклонения H потребовались бы оценки величин α или, по крайней мере, их разбиение на группы больших и малых.

Гипотеза H является линейной гипотезой с $r = a - 1$, $s = = 1 + (a - 1) + (b - 1) = a + b - 1$ и $n - s = (a - 1)(b - 1)$. Оценки параметров в Ω по методу наименьших квадратов могут быть получены из тождества

$$\begin{aligned} \sum \sum (X_{ij} - \xi_{ij})^2 &= \sum \sum (X_{ij} - \mu - \alpha_i - \beta_j)^2 = \\ &= \sum \sum [(X_{ij} - X_{i.} - X_{.j} + X_{..}) + (X_{i.} - X_{..} - \alpha_i) + \\ &\quad + (X_{.j} - X_{..} - \beta_j) + (X_{..} - \mu)]^2 = \\ &= \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2 + b \sum (X_{i.} - X_{..} - \alpha_i)^2 + \\ &\quad + a \sum (X_{.j} - X_{..} - \beta_j)^2 + ab (X_{..} - \mu)^2, \end{aligned}$$

которое выполняется, так как в разложении третьей суммы квадратов попарные произведения исчезают. Мы имеем

$$\hat{\alpha}_i = X_{i.} - X_{..}, \quad \hat{\beta}_j = X_{.j} - X_{..}, \quad \hat{\mu} = X_{..}$$

и

$$\sum \sum (X_{ij} - \hat{\xi}_{ij})^2 = \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2.$$

При гипотезе H имеем, кроме того, $\hat{\beta}_j = X_{.j} - X_{..}$, $\hat{\mu} = X_{..}$ и $\hat{\xi}_{ij} - \hat{\xi}_{ij} = X_{i.} - X_{..}$. Следовательно, наилучший инвариантный критерий отклоняет гипотезу, когда

$$W^* = \frac{b \sum (X_{i.} - X_{..})^2 / (a - 1)}{\sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2 / (a - 1)(b - 1)} > C. \quad (26)$$

Параметр нецентральности, от которого зависит мощность критерия, равен

$$\psi^2 = b \sum (\xi_{i.} - \xi_{..})^2 / \sigma^2 = b \sum \alpha_i^2 / \sigma^2. \quad (27)$$

Эта задача доставляет новый пример применения дисперсионного анализа. Полная изменчивость может быть разбита на три компоненты:

$$\begin{aligned} \sum \sum (X_{ij} - X_{..})^2 &= b \sum (X_{i.} - X_{..})^2 + a \sum (X_{.j} - X_{..})^2 + \\ &+ \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2. \end{aligned}$$

Из них первая описывает изменчивость, обусловленную фактором α , а вторая — фактором β . Последняя компонента в канонической записи раздела 1 имеет вид $\sum_{i=s+1}^n Y_i^2$, т. е. это сумма квадратов тех случайных величин, средние которых равны нулю и при Ω . Так как эта часть (которая после деления на $n - s$ дает оценку σ^2) не может быть связана с действием факторов α и β , то ее часто называют «ошибкой», подчеркивая этим названием, что она связана со случайностью результатов наблюдения, а не с каким-либо расхождением в средних значениях. В действительности разбиение не является столь отчетливым, как можно вывести из сказанного. Каждая компонента, например, относимая к фактору α , содержит некоторые «ошибки», как можно видеть из величины ее математического ожидания

$$E \sum (X_{i.} - X_{..})^2 = (a - 1) \sigma^2 + b \sum \alpha_i^2.$$

Вместо проверки гипотезы о наличии эффекта некоторого фактора можно стремиться оценить размер этого эффекта при разных интенсивностях (уровнях) фактора. Другими параметрами, оценка которых может представлять интерес, являются средние значения исходов (например, размер урожая) $\xi_{1.}, \dots, \xi_{a.}$ при различных значениях фактора. Пусть $\theta_i = \mu + \alpha_i = \xi_{i.}$. Тогда доверительные множества для $(\theta_1, \dots, \theta_a)$ получаются рассмотрением гипотезы $H(\theta^0)$: $\theta_i = \theta_i^0$ ($i = 1, \dots, a$). При проверке гипотезы $\theta_1 = \dots = \theta_a = 0$ оценки по методу наименьших квадратов для ξ_{ij} равны $\hat{\xi}_{ij} = X_{i.} + X_{.j} - X_{..}$ и $\hat{\xi}_{ij} = X_{.j} - X_{..}$. Как и раньше, сумма квадратов в знаменателе равна $\sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2$, а сумма квадратов в числителе теперь равна

$$\sum \sum (\hat{\xi}_{ij} - \hat{\xi}_{ij}^0)^2 = b \sum X_{i.}^2.$$

Общий случай сводится к этому специальному заменой величин X_{ij} на $X_{ij} - \theta_i^0$. Так как $s = a + b - 1$ и $r = a$, то гипотеза

$H(\theta^0)$ отклоняется при

$$\frac{b \sum (X_{i.} - \theta_i^0)^2 / a}{\sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2 / (a-1)(b-1)} > C.$$

Соответствующие доверительные множества для $(\theta_1, \dots, \theta_a)$ являются сферами

$$\sum (\theta_i - X_{i.})^2 \leq \frac{aC \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2}{(a-1)(b-1)b}.$$

Рассматривая доверительные множества для эффектов $\alpha_1, \dots, \alpha_a$, следует иметь в виду, что величины α связаны между собой. Их общая сумма равна нулю и поэтому достаточно ограничиться величинами $\alpha_1, \dots, \alpha_{a-1}$. Однако решение оказывается более легким и более симметричным, если сохранить все α . Критическая область для H : $\alpha_i = \alpha_i^0$ при $i = 1, \dots, a$ ($\sum \alpha_i^0 = 0$) получается из (26) заменой $X'_{ij} = X_{ij} - \alpha_i^0$ и потому имеет вид

$$b \sum (X_{i.} - X_{..} - \alpha_i^0)^2 > \frac{C \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2}{(b-1)}.$$

Соответствующие доверительные множества состоят из точек $(\alpha_1, \dots, \alpha_a)$, для которых $\sum \alpha_i = 0$ и

$$\sum [\alpha_i - (X_{i.} - X_{..})]^2 \leq \frac{C \sum \sum (X_{ij} - X_{i.} - X_{.j} + X_{..})^2}{b(b-1)}.$$

В пространстве $(\alpha_1, \dots, \alpha_a)$ это неравенство определяет шар с центром $(X_{1.} - X_{..}, \dots, X_{a.} - X_{..})$ в гиперплоскости $\sum \alpha_i = 0$, т. е. доверительные множества для величин α состоят из внутренности и поверхности гиперсфер, получающихся в пересечении a -мерных сфер с гиперплоскостью $\sum \alpha_i = 0$.

Как в этом, так и в предшествующем случае обычными методами можно показать, что класс доверительных множеств инвариантен относительно подходящей группы линейных преобразований и что среди инвариантных эти множества являются равномерно наиболее точными.

5. Двойная классификация: m наблюдений в клетке

В предыдущем разделе предполагалось, что эффекты факторов α и β независимы и, следовательно, аддитивны. Однако факторы могут взаимодействовать в том смысле, что эффект одного из них зависит от интенсивности другого. Так, успех учителя зависит, например, от степени подготовленности или возраста студентов, а урожай при различных условиях орошения зависит от типа почвы и сорта выращиваемых растений. Если

отказаться от предположения аддитивности, то средние значения ξ_{ij} величин X_{ij} уже не будут более задаваться соотношениями (24) при Ω , но станут совершенно произвольными. Поэтому для каждой комбинации значений признаков требуется не менее ab наблюдений, так как в противном случае $s = n$. Мы рассмотрим здесь лишь простейший случай, когда при каждой комбинации значений признаков производится одно и то же число наблюдений.

Пусть $X_{ijk} (i = 1, \dots, a; j = 1, \dots, b; k = 1, \dots, m)$ — независимые случайные величины, распределенные нормально с общей дисперсией σ^2 и средними $E(X_{ijk}) = \xi_{ij}$. По аналогии с прежним случаем мы пишем

$$\begin{aligned} \xi_{ij} &= \xi_{..} + (\xi_{i.} - \xi_{..}) + (\xi_{.j} - \xi_{..}) + (\xi_{ij} - \xi_{i.} - \xi_{.j} + \xi_{..}) = \\ &= \mu + \alpha_i + \beta_j + \gamma_{ij} \end{aligned}$$

с $\sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$. Тогда α_i представляет собой среднее значение фактора 1 на уровне i (усреднение производится по b уровням фактора 2; аналогично интерпретируются и величины β_j). Величины γ называются *взаимодействиями*, так как γ_{ij} указывает количество, на которое совместный эффект $\xi_{ij} - \xi_{..}$ факторов 1 и 2 превосходит сумму их индивидуальных эффектов. Рассмотрим снова гипотезу о равенстве нулю всех величин α . Тогда $r = a - 1$, $s = ab$ и $n - s = (m - 1)ab$. Из разложения

$$\sum \sum \sum (X_{ijk} - \xi_{ij})^2 = \sum \sum \sum (X_{ijk} - X_{ij.})^2 + m \sum \sum (X_{ij.} - \xi_{ij})^2$$

и

$$\begin{aligned} \sum \sum (X_{ij.} - \xi_{ij})^2 &= \sum \sum (X_{ij.} - X_{i..} - X_{.j.} + X_{...} - \gamma_{ij})^2 + \\ &+ b \sum (X_{i..} - X_{...} - \alpha_i)^2 + a \sum (X_{.j.} - X_{...} - \beta_j)^2 + ab (X_{...} - \mu)^2 \end{aligned}$$

следует, что

$$\hat{\mu} = \hat{\mu} = \hat{\xi}_{..} = X_{...}, \quad \hat{\alpha}_i = \hat{\xi}_{i.} - \hat{\xi}_{..} = X_{i..} - X_{...},$$

$$\hat{\beta}_j = \hat{\beta}_j = \hat{\xi}_{.j} - \hat{\xi}_{..} = X_{.j.} - X_{...},$$

$$\hat{\gamma}_{ij} = \hat{\gamma}_{ij} = X_{ij.} - X_{i..} - X_{.j.} + X_{...}$$

и что

$$\sum \sum \sum (X_{ijk} - \hat{\xi}_{ij})^2 = \sum \sum \sum (X_{ijk} - X_{ij.})^2,$$

$$\sum \sum \sum (\hat{\xi}_{ij} - \hat{\xi}_{..})^2 = mb \sum (X_{i..} - X_{...})^2.$$

Наиболее мощный инвариантный критерий отвергает гипотезу, когда

$$W^* = \frac{mb \sum (X_{i..} - X_{...})^2 / (a-1)}{\sum \sum \sum (X_{ijk} - X_{ij.})^2 / (m-1) ab} > C. \quad (28)$$

Здесь параметр нецентральности распределения W^* равен

$$\frac{mb \sum (\xi_{i.} - \xi_{..})^2}{\sigma^2} = \frac{mb \sum \alpha_i^2}{\sigma^2}. \quad (29)$$

Интерес представляет и гипотеза о независимости факторов *) H' : $\gamma_{ij} = 0$ при всех i, j . Оценки параметров по методу наименьших квадратов снова легко вычисляются. РНМ инвариантный критерий имеет критическую область (см. задачу 12)

$$W^* = \frac{m \sum \sum (X_{ij.} - X_{i..} - X_{.j.} + X_{...})^2 / (a-1)(b-1)}{\sum \sum \sum (X_{ijk} - X_{ij.})^2 / (m-1) ab} > C. \quad (30)$$

При гипотезе H' статистика W^* имеет F -распределение с $(a-1)(b-1)$ и $(m-1)ab$ степенями свободы. Параметр нецентральности при любых альтернативных значениях величин γ равен

$$\psi^2 = \frac{m \sum \sum \gamma_{ij}^2}{\sigma^2}. \quad (31)$$

Разложение полной изменчивости на компоненты в настоящем случае имеет вид

$$\begin{aligned} \sum \sum \sum (X_{ijk} - X_{...})^2 &= mb \sum (X_{i..} - X_{...})^2 + \\ &+ ma \sum (X_{.j.} - X_{...})^2 + m \sum \sum (X_{ij.} - X_{i..} - X_{.j.} + X_{...})^2 + \\ &+ \sum \sum \sum (X_{ijk} - X_{ij.})^2. \end{aligned}$$

Здесь первые три суммы представляют компоненты изменчивости, связанные с величинами α , β и γ соответственно. Последняя компонента описывает «ошибку». Критерии для гипотезы о том, что все α , все β или все γ равны нулю (первая и третья из них имеют критические области (28) и (30)), получаются сравнением соответствующей суммы квадратов с последней суммой квадратов. Аналогичное разложение возможно и в случае, когда величины γ априори принимаются равными нулю. Тогда третья компонента, ранее связывавшаяся с γ , описывает дополнительный

*) Критерий для H' при некоторых ограниченных альтернативах был предложен для случая одного наблюдения в клетке Tukey, One degree of freedom for non-additivity, Biometrics 5 (1949), 232—242.

эффект ошибок. Само разложение принимает вид

$$\sum \sum \sum (X_{ijk} - X_{...})^2 = mb \sum (X_{i..} - X_{...})^2 + \\ + ma \sum (X_{.j.} - X_{...})^2 + \sum \sum \sum (X_{ijk} - X_{i..} - X_{.j.} + X_{...})^2,$$

где последний член соответствует ошибкам. Гипотеза $H: \alpha_1 = \dots = \alpha_a = 0$ отклоняется при

$$\frac{mc \sum (X_{i..} - X_{...})^2 (a-1)}{\sum \sum \sum (X_{ijk} - X_{i..} - X_{.j.} + X_{...})^2 / (abm - a - b + 1)} > C.$$

Допустим теперь, что предположение об отсутствии взаимодействия, при котором получен критерий, не оправдано. Тогда сумма квадратов в знаменателе имеет нецентральное χ^2 -распределение (вместо обычного) и потому стохастически больше, чем предполагалось (см. задачу 13). Отсюда следует, что действительная вероятность отклонения меньше, чем она была бы при $\sum \sum \gamma_{ij}^2 = 0$. Это показывает, что вероятность ошибки первого рода не превосходит номинального уровня значимости, каковы бы ни были значения γ . Однако мощность убывает при возрастании $\sum \sum \gamma_{ij}^2 / \sigma^2$ и стремится к нулю, когда это отношение стремится к бесконечности.

Дисперсионный анализ и связанные с ним критерии, рассмотренные в этом разделе для случая двух факторов, непосредственно переносятся на случай большего числа факторов (см., например, задачу 14). С другой стороны, если число наблюдений для каждой комбинации значений факторов (для каждой *клетки*) не одно и то же, то задача, хотя и охватывается схемой линейных гипотез, но становится значительно сложнее.

Чрезвычайно важными являются планы эксперимента, в которых появляются только некоторые комбинации значений факторов, поскольку тогда возможно сократить объем экспериментальной работы. Так, например, действия трех независимых факторов с t уровнями каждый можно проанализировать всего с t^2 наблюдениями (вместо t^3 наблюдений, если производить по одному наблюдению для каждой комбинации уровней), если использовать метод латинских квадратов (см. задачу 15).

Рассмотренный здесь класс задач как частный случай содержит задачу двух выборок, обсуждавшуюся в главе 5, так как последняя касается одного фактора с двумя уровнями. Изложенные в этой связи вопросы о возможной неоднородности экспериментального материала и о рандомизации как средстве компенсации этой неоднородности сохраняют свою важность и в настоящей, более сложной ситуации. Если неоднородный материал разделен на несколько однородных групп, то эту классификацию можно трактовать как порожденную одним или несколькими

дополнительными факторами. Выбор этих групп является существенным шагом в определении подходящего плана эксперимента*) (очень простой пример обсуждался в задачах 26 и 27 главы 5).

Как защиту от возможных неоднородностей внутри подгрупп (и других отклонений от сделанных предположений) используют рандомизацию при назначении способов обработки внутри групп. Как и в случае задачи о двух выборках, сам по себе процесс рандомизации (без каких бы то ни было предположений относительно метода отбора экспериментального материала, нормальности, независимости и т. д.) позволяет строить критерии уровня α для различных интересных гипотез. Критерии перестановок в рассматриваемом случае сводятся к вычислению подходящей F -статистики W^* и сравнению этого значения с тем, которое получается применением к наблюдениям перестановок, связанных с процессом рандомизации**). Как и раньше, эти критерии асимптотически эквивалентны соответствующим F -критериям, которыми они, следовательно, могут быть аппроксимированы.

6. Регрессия

Как уже отмечалось в примере 2, гипотезы о специальных значениях коэффициентов регрессии α и β , когда $X_1, \dots, X_n$ — независимые нормально распределенные величины с общей дисперсией σ^2 и средними

$$\xi_i = \alpha + \beta t_i, \quad (32)$$

являются частными случаями общей линейной гипотезы. Гипотезы $H_1: \alpha = \alpha_0$ и $H_2: \beta = \beta_0$ обсуждались в разделе 6 главы 5, где было показано, что для них существуют РНМ несмещенные критерии. Мы рассмотрим теперь гипотезы H_1 и H_2 и гипотезу $H_3: \alpha = \alpha_0, \beta = \beta_0$ с новой точки зрения. В соответствии с общей теорией раздела 1, соответствующие критерии будут РНМ инвариантными относительно надлежащих групп линейных преобразований. Для первых двух гипотез, для которых $r = 1$, на этом пути получается, если учесть рассуждения раздела 6 главы 6, новое доказательство того, что рассматриваемые критерии — РНМ несмещенные.

*) Обсуждение различных планов и условия, при которых они уместны, см., например, в книгах: Kempthorne, The Design and Analysis of Experiments, New York, John Wiley & Sons, 1952 и Cochran and Cox, Experimental Designs, New York, John Wiley & Sons, 2nd ed., 1957. Свойства оптимальности некоторых планов, доказанные Вальдом, Эренфельдом, Кифером и другими, рассматриваются в работе: Kiefer, On the nonrandomized optimality and randomized nonoptimality of symmetrical designs, Ann. Math. Stat. 29 (1958), 675—699.

**) Подробности см. в цит. книге Kempthorne.

Пространство Π_Ω — одно и то же для всех трех гипотез. Оно натянуто на векторы $(1, \dots, 1)$ и $(t_1, \dots, t_n)$ и потому имеет размерность $s = 2$, за исключением случая, когда все t_i равны друг другу. Этот случай мы не будем рассматривать. Оценки для α и β по методу наименьших квадратов при Ω получаются минимизацией $\sum (X_i - \alpha - \beta t_i)^2$. При любом фиксированном β минимум достигается при $\alpha = \bar{X} - \beta \bar{t}$. При таком α сумма квадратов сводится к $\sum [(X_i - \bar{X}) - \beta (t_i - \bar{t})]^2$. Минимизируя последнее выражение относительно β , мы находим

$$\hat{\beta} = \frac{\sum (X_i - \bar{X})(t_i - \bar{t})}{\sum (t_i - \bar{t})^2}, \quad \hat{\alpha} = \bar{X} - \hat{\beta} \bar{t}; \quad (33)$$

кроме того, сумма квадратов в знаменателе для всех трех гипотез равна

$$\sum (X_i - \hat{\alpha} - \hat{\beta} t_i)^2 = \sum (X_i - \bar{X})^2 - \hat{\beta}^2 \sum (t_i - \bar{t})^2.$$

Числитель статистики, по которой строится критерий, равен Y_1^2 при проверке гипотез $\alpha = 0$ и $\beta = 0$ и равен $Y_1^2 + Y_2^2$ при проверке гипотезы $\alpha = \beta = 0$.

При гипотезе $\alpha = 0$, как показано в примере 3, статистика Y_1 равна

$$\left(\frac{\bar{X} - \bar{t} \sum t_i X_i}{\sum t_i^2} \right) \sqrt{\frac{n \sum t_i^2}{\sum (t_i - \bar{t})^2}} = \hat{\alpha} \sqrt{\frac{n \sum (t_i - \bar{t})^2}{\sum t_i^2}}.$$

Так как при этом

$$E(Y_1) = \alpha \sqrt{\frac{n \sum (t_i - \bar{t})^2}{\sum t_i^2}},$$

то гипотеза $\alpha = \alpha_0$ эквивалентна гипотезе $E(Y_1) = \eta_1^0 = \alpha_0 \sqrt{n \sum (t_i - \bar{t})^2 / \sum t_i^2}$, для которой критическая область (17) имеет вид $(n - s)(Y_1 - \eta_1^0)^2 / \sum_{i=s+1}^n Y_i^2 > C_0$ или

$$\frac{|\hat{\alpha} - \alpha_0| \sqrt{n \sum (t_i - \bar{t})^2 / \sum t_i^2}}{\sqrt{\sum (X_i - \hat{\alpha} - \hat{\beta} t_i)^2 / (n - 2)}} > C_0. \quad (34)$$

Как было показано ранее, при гипотезе $\beta = 0$ Y_1 равно

$$\frac{\sum (X_i - \bar{X})(t_i - \bar{t})}{\sqrt{\sum (t_i - \bar{t})^2}} = \hat{\beta} \sqrt{\sum (t_i - \bar{t})^2}.$$

Так как при этом $E(Y_1) = \beta \sqrt{\sum (t_i - \bar{t})^2}$, то гипотеза $\beta = \beta_0$ эквивалентна гипотезе $E(Y_1) = \eta_1^0 = \beta_0 \sqrt{\sum (t_i - \bar{t})^2}$ и, следова-

тельно, критическая область определяется неравенством

$$\frac{|\hat{\beta} - \beta_0| \sqrt{\sum (t_j - \bar{t})^2}}{\sqrt{\sum (X_i - \hat{\alpha} - \hat{\beta} t_i)^2 / (n-2)}} > C_0. \quad (35)$$

В примере 3 было указано, что при гипотезе $\alpha = \beta = 0$

$$Y_1 = \hat{\beta} \sqrt{\sum (t_j - \bar{t})^2}, \quad Y_2 = \sqrt{n} \bar{X} = \sqrt{n} (\hat{\alpha} + \hat{\beta} \bar{t}),$$

так что числитель в формуле (7) равен

$$\frac{(Y_1^2 + Y_2^2)}{2} = \frac{[n(\hat{\alpha} + \hat{\beta} \bar{t})^2 + \hat{\beta}^2 \sum (t_j - \bar{t})^2]}{2}.$$

Более общая гипотеза $\alpha = \alpha_0$, $\beta = \beta_0$ эквивалентна гипотезе $E(Y_1) = \eta_1^0$, $E(Y_2) = \eta_2^0$, где $\eta_1^0 = \beta_0 \sqrt{\sum (t_j - \bar{t})^2}$, $\eta_2^0 = \sqrt{n}(\alpha_0 + \beta_0 \bar{t})$. Поэтому критическая область может быть представлена в виде

$$\frac{[n(\hat{\alpha} - \alpha_0)^2 + 2n\bar{t}(\hat{\alpha} - \alpha_0)(\hat{\beta} - \beta_0) + \sum t_i^2(\hat{\beta} - \beta_0)^2]^{1/2}}{\sum (X_i - \hat{\alpha} - \hat{\beta} t_i)^2 / (n-2)} > C. \quad (36)$$

Соответствующие доверительные множества для (α, β) получаются обращением этих неравенств и заменой α_0 и β_0 на α и β . Получающиеся множества имеют вид эллипсов с центром в $(\hat{\alpha}, \hat{\beta})$.

Простую модель регрессии (32) можно обобщать во многих направлениях; например, средние ξ_i могут быть полиномами от t_i степени выше первой (см. задачу 18) или более сложными функциями, скажем, тригонометрическими многочленами, или функциями нескольких переменных, например t_i , u_i , v_i . Некоторые дальнейшие обобщения будут теперь проиллюстрированы несколькими примерами.

Пример 6. Несколько проблем возникает при наличии более чем одной линии регрессии. Допустим, что случайные величины X_{ij} независимы и распределены нормально с общей дисперсией и средними

$$\xi_{ij} = \alpha_i + \beta_i t_{ij} \quad (j = 1, \dots, n_i; \quad i = 1, \dots, b). \quad (37)$$

Гипотеза о том, что эти линии регрессии имеют один и тот же наклон

$$H: \beta_1 = \dots = \beta_b,$$

может встретиться, например, при проверке равенства нескольких скоростей роста. Пространство параметров Π_Ω имеет размерность $s = 2b$, если предположить, что ни одна из сумм $\sum_j (t_{ij} - \bar{t}_i)^2$ не обращается в нуль. Число ограничений, налагаемых гипотезой, равно $r = b - 1$. Минимальное значение $\sum \sum (X_{ij} - \xi_{ij})^2$ в Ω получается минимизацией при каждом i сумм

$\sum_j (X_{ij} - \alpha_i - \beta_i t_{ij})^2$, так что в силу (33)

$$\hat{\beta}_i = \frac{\sum_j (X_{ij} - X_{i.})(t_{ij} - t_{i.})}{\sum_j (t_{ij} - t_{i.})^2}, \quad \hat{\alpha}_i = X_{i.} - \hat{\beta}_i t_{i.}.$$

При гипотезе H следует минимизировать сумму $\sum \sum (X_{ij} - \alpha_i - \beta_i t_{ij})^2$, что при любом фиксированном β приводит к значению $\alpha_i = X_{i.} - \beta t_{i.}$ и сумму квадратов превращает в $\sum \sum [(X_{ij} - X_{i.}) - \beta (t_{ij} - t_{i.})]^2$. Минимизируя последнее выражение по β , находим

$$\hat{\beta} = \frac{\sum \sum (X_{ij} - X_{i.})(t_{ij} - t_{i.})}{\sum \sum (t_{ij} - t_{i.})^2}; \quad \hat{\alpha}_i = X_{i.} - \hat{\beta} t_{i.}.$$

Так как

$$X_{ij} - \hat{\xi}_{ij} = X_{ij} - \hat{\alpha}_i - \hat{\beta}_i t_{ij} = (X_{ij} - X_{i.}) - \hat{\beta}_i (t_{ij} - t_{i.})$$

и

$$\hat{\xi}_{ij} - \hat{\xi}_{ij} = (\hat{\alpha}_i - \hat{\alpha}_i) + t_{ij} (\hat{\beta}_i - \hat{\beta}) = (\hat{\beta}_i - \hat{\beta}) (t_{ij} - t_{i.}),$$

то критическая область (15) равна

$$\frac{\sum_i (\hat{\beta}_i - \hat{\beta})^2 \sum_j (t_{ij} - t_{i.})^2 / (b - 1)}{\sum \sum [(X_{ij} - X_{i.}) - \hat{\beta}_i (t_{ij} - t_{i.})]^2 / (n - 2b)} > C, \quad (38)$$

где левая часть при гипотезе H имеет F -распределение с $(b - 1)$ и $(n - 2b)$ степенями свободы.

Так как

$$E(\hat{\beta}_i) = \beta_i \quad \text{и} \quad E(\hat{\beta}) = \frac{\sum_i \beta_i \sum_j (t_{ij} - t_{i.})^2}{\sum \sum (t_{ij} - t_{i.})^2},$$

то параметр нецентральности указанного распределения при альтернативных значениях β равен $\psi^2 = \sum_i (\beta_i - \tilde{\beta})^2 \sum_j (t_{ij} - t_{i.})^2 / \sigma^2$, где $\tilde{\beta} = E(\hat{\beta})$. В частном случае, когда n_i и t_{ij} не зависят от i , $\tilde{\beta}$ принимает вид $\tilde{\beta} = \sum \beta_j / b$.

Пример 7. Модель регрессии (37) возникает при сравнении нескольких способов обработки при условии, что экспериментальные единицы считаются постоянными, а соответствующие эффекты u_{ij} (определенные в разделе 10 главы 5) — пропорциональными известным константам t_{ij} . Здесь t_{ij} может измерять степень плодородия и i j -го участка почвы или вес i j -го животного до эксперимента. В этой обстановке часто можно принять, что множитель пропорциональности β_i не зависит от способа обработки. Тогда (37) сводится к

$$\xi_{ij} = \alpha_i + \beta t_{ij}, \quad (39)$$

а гипотеза об отсутствии эффекта обработки принимает вид

$$H: \alpha_1 = \dots = \alpha_b.$$

Пространство Π_Ω совпадает с пространством Π_ω предыдущего примера, так что $s = b + 1$ и

$$\hat{\beta} = \frac{\sum \sum (X_{ij} - X_{i.})(t_{ij} - t_{i.})}{\sum \sum (t_{ij} - t_{i.})^2}; \quad \hat{\alpha}_i = X_{i.} - \hat{\beta} t_{i.}.$$

Минимизация суммы $\sum \sum (X_{ij} - \alpha - \beta t_{ij})^2$ дает

$$\hat{\beta} = \frac{\sum \sum (X_{ij} - X_{..})(t_{ij} - t_{..})}{\sum \sum (t_{ij} - t_{..})^2}, \quad \hat{\alpha} = X_{..} - \hat{\beta} t_{..},$$

где $X_{..} = \sum \sum X_{ij}/n$, $t_{..} = \sum \sum t_{ij}/n$, $n = \sum n_i$. Сумма квадратов в числителе W^* (см. (15)) равна

$$\sum \sum (\hat{\xi}_{ij} - \hat{\xi}_{i.})^2 = \sum \sum [(X_{i.} - X_{..}) + \hat{\beta}(t_{ij} - t_{i.}) - \hat{\beta}(t_{i.} - t_{..})]^2.$$

Следовательно, гипотеза H отвергается при

$$\frac{\sum \sum [(X_{i.} - X_{..}) + \hat{\beta}(t_{ij} - t_{i.}) - \hat{\beta}(t_{i.} - t_{..})]^2 / (b - 1)}{\sum \sum [(X_{ij} - X_{i.}) - \hat{\beta}(t_{ij} - t_{i.})]^2 / (n - b - 1)} > C, \quad (40)$$

где при гипотезе H левая часть имеет F -распределение с $b - 1$ и $n - b - 1$ степенями свободы.

Гипотеза H может быть проверена и без того, чтобы значения t_{ij} считать известными. В этом случае она становится гипотезой об отсутствии эффекта в одинарной классификации, рассмотренной в разделе 3, и соответствующий критерий дается формулой (18). Действительно, поскольку в этом случае индивидуальные эффекты u_{ij} предполагаются постоянными (но неизвестными), то способы обработки приписываются экспериментальным единицам или совершенно случайно, или случайно внутри подгрупп. Соответствующий критерий является рандомизированным; критерий (18) служит приближением к нему.

Пример 7 иллюстрирует важный класс ситуаций, в которых дисперсионный анализ (в нашем случае касающийся одинарной классификации) комбинируется с проблемой регрессии (в нашем случае — линейная регрессия на единственную «сопутствующую переменную» t). Обе части проблемы могут, конечно, оказаться более сложными, чем это было в примере. В общем случае такой комбинированной проблемы можно проверить гипотезы о величине эффектов, подобно тому как это сделано выше. Аналогичный анализ можно провести для коэффициентов регрессии. Разбиение полной изменчивости на компоненты, соответствующие различным способам обработки и различным коэффициентам регрессии, является предметом так называемого *анализа ковариаций*.*)

*) См., например, ссылку на работу Кокрена (1957) на стр. 364 в разделе 15. — *Прим. ред.*

7. Модель II: одинарная классификация

Мы видели, что способ анализа действия одного или нескольких признаков зависит от того, считаются ли экспериментальные единицы постоянными или же образуют случайную выборку из совокупности таких единиц. То же различие возникает и при изучении эффектов самих факторов. В некоторых применениях они постоянны, а в некоторых других являются не наблюдаемыми случайными величинами. Если размеры эффектов все постоянны или все случайны, то мы говорим о *модели I* или *модели II* соответственно; ситуации, где встречаются оба типа, мы обозначаем термином *смешанная модель*. Ясно, что схемой линейных гипотез, как она определена в начале этой главы, охватывается только модель I. В настоящем разделе мы рассмотрим для случая одного фактора (одинарная классификация) модель II. Модель I для этого случая была разобрана в разделе 3.

В качестве иллюстрации рассмотрим какую-либо промышленную продукцию, например сталь, которая производится партиями. Предположим, что из каждой из s партий берется выборка объема n и что результаты наблюдений X_{ij} ($j = 1, \dots, n$; $i = 1, \dots, s$) независимы и распределены нормально с дисперсией σ^2 и средними ξ_i . Если бы фактор, соответствующий индексу i , был постоянен, с одним и тем же эффектом α_i в каждом повторении эксперимента, то мы имели бы

$$\xi_i = \mu + \alpha_i \quad (\sum \alpha_i = 0) \quad \text{и} \quad X_{ij} = \mu + \alpha_i + U_{ij},$$

где U_{ij} независимы и распределены по закону $N(0, \sigma^2)$. Гипотеза об отсутствии эффекта имеет вид $\xi_1 = \dots = \xi_s$, или, что эквивалентно, $\alpha_1 = \dots = \alpha_s = 0$. Однако эффект связан с партиями, из которых при каждом повторении эксперимента берется новая часть; поэтому эффект не остается постоянным. Мы допустим, что эффекты, связанные с партиями, образуют выборку из нормального распределения, и, чтобы подчеркнуть их случайную природу, мы будем писать A_i вместо α_i , так что

$$X_{ij} = \mu + A_i + U_{ij}. \quad (41)$$

Предположение об аддитивности (т. е. об отсутствии взаимодействия) эффектов, связанных с партией, и индивидуальных эффектов влечет в этой модели независимость случайных величин A и U . Если математические ожидания величин A_i включить в μ , то все величины A и U будут независимыми нормально распределенными со средними нуль и дисперсиями, равными σ_A^2 и σ^2 соответственно. Конечно, величины X уже не будут независимыми.

Гипотеза об отсутствии эффекта партий, т. е. гипотеза о том, что A равны нулю и, следовательно, постоянны, принимает форму

$$H: \sigma_A^2 = 0.$$

Это предположение в рассматриваемой ситуации нереалистично, но оно может рассматриваться как предельный случай гипотезы

$$H(\Delta_0): \frac{\sigma_A^2}{\sigma^2} \leq \Delta_0$$

о том, что эффект партий мал по сравнению с изменчивостью материала внутри партии. Этим гипотезам в модели I соответствуют гипотезы $\sum \alpha_i^2 = 0$ и $\sum \alpha_i^2 / \sigma^2 \leq \Delta_0$.

Чтобы получить критерий для $H(\Delta_0)$, удобно начать с того же самого преобразования переменных, которое приводило соответствующую модель I к канонической форме. Каждый набор $(X_{i1}, \dots, X_{in})$ мы подвергаем ортогональному преобразованию

$Y_{ij} = \sum_{k=1}^n c_{jk} X_{ik}$ такому, что $Y_{i1} = \sqrt{n} X_{i1}$. Так как $c_{1k} = 1/\sqrt{n}$ при $k=1, \dots, n$ (см. пример 3), то из предположения ортогональности следует, что $\sum_{k=1}^n c_{jk} = 0$ для $j=2, \dots, n$ и, что, сле-

довательно, $Y_{ij} = \sum_{k=1}^n c_{jk} U_{ik}$ при $j > 1$. Таким образом, вели-

чины Y_{ij} с $j > 1$ независимы и распределены нормально с нулевыми средними значениями и дисперсией σ^2 . Они также независимы от U_i , поскольку $(\sqrt{n} U_i, Y_{i2}, \dots, Y_{in})' = C(U_{i1}, U_{i2}, \dots, U_{in})'$ (штрих указывает операцию транспонирования матрицы). С другой стороны, случайные величины $Y_{i1} = \sqrt{n} X_{i1} = \sqrt{n}(\mu + A_i + U_{i1})$ также независимы и распределены нормально со средними значениями $\sqrt{n}\mu$ и дисперсией $\sigma^2 + n\sigma_A^2$. Если произвести дополнительное ортогональное преобразование вектора $(Y_{i1}, \dots, Y_{is})$ в вектор $(Z_{i1}, \dots, Z_{is})$ такой, что $Z_{i1} = \sqrt{s} Y_{i1}$, то величины Z будут независимыми, нормально распределенными с общей дисперсией $\sigma^2 + n\sigma_A^2$ и средними $E(Z_{i1}) = \sqrt{sn}\mu$ и $E(Z_{ij}) = 0$ при $j > 1$. Полагая, по определению, $Z_{ij} = Y_{ij}$ при $j > 1$, мы найдем, что совместная плотность величин Z равна

$$(2\pi)^{-ns/2} \sigma^{-(n-1)s} (\sigma^2 + n\sigma_A^2)^{-s/2} \exp \left[-\frac{1}{2(\sigma^2 + n\sigma_A^2)} \times \right. \\ \left. \times \left((z_{i1} - \sqrt{sn}\mu)^2 + \sum_{i=2}^s z_{i1}^2 \right) - \frac{1}{2\sigma^2} \sum_{i=1}^s \sum_{j=2}^n z_{ij}^2 \right]. \quad (42)$$

Проблема проверки $H(\Delta_0)$ инвариантна относительно добавления произвольной постоянной к Z_{11} ; по отношению к этому преобразованию совокупность остальных величин Z является максимальным инвариантом. Эта совокупность составлена из двух выборок — объемов $s(n-1)$ и $s-1$ — из нормальных распределений со средними нуль и дисперсиями σ^2 и $\tau^2 = \sigma^2 + n\sigma_A^2$. Гипотеза $H(\Delta_0)$ эквивалентна гипотезе $\tau^2/\sigma^2 \leq 1 + \Delta_0 n$, и проблема сводится к задаче сравнения дисперсий двух нормальных распределений, которая рассматривалась в примере 6 главы 6 (без предположения о том, что средние равны нулю). РНМ инвариантный (относительно умножения всех Z_{ij} на одну и ту же положительную константу) критерий отклоняет гипотезу, когда

$$W^* = \frac{1}{(1 + \Delta_0 n)} \cdot \frac{S_A^2/(s-1)}{S^2/(n-1)s} > C, \quad (43)$$

где

$$S_A^2 = \sum_{i=2}^n Z_{i1}^2 \quad \text{и} \quad S^2 = \sum_{i=1}^s \sum_{j=2}^n Z_{ij}^2 = \sum_{i=1}^s \sum_{j=2}^n Y_{ij}^2.$$

Константа C определяется из уравнения

$$\int_C^\infty F_{s-1, (n-1)s}(y) dy = \alpha.$$

Так как

$$\sum_{j=1}^n Y_{ij}^2 - Y_{i1}^2 = \sum_{j=1}^n U_{ij}^2 - nU_{i1}^2.$$

и

$$\sum_{i=1}^s Z_{i1}^2 - Z_{11}^2 = \sum_{i=1}^s Y_{i1}^2 - sY_{11}^2,$$

то суммы квадратов, стоящие в числителе и знаменателе W^* , выраженные в терминах величины X , принимают вид

$$S_A^2 = n \sum_{i=1}^s (X_{i.} - X_{..})^2 \quad \text{и} \quad S^2 = \sum_{i=1}^s \sum_{j=1}^n (X_{ij} - X_{i.})^2.$$

В частном случае, при $\Delta_0 = 0$, критерий (43) эквивалентен соответствующему критерию (18) для модели I, но они являются решениями разных проблем и имеют разные функции мощности. Сумма квадратов в числителе W^* пропорциональна обычному χ^2 , даже если гипотеза неверна (в то время как в модели I эта сумма имела нецентральное χ^2 -распределение). Мощность критерия (43) при альтернативном значении Δ вычисляется

через F -распределение по формуле

$$\beta(\Delta) = P_{\Delta} \{W^* > C\} = \int_{\frac{1+\Delta sn}{1+\Delta n} C}^{\infty} F_{s-1, (n-1)s}(y) dy.$$

Семейство критериев (43) при меняющемся Δ_0 эквивалентно доверительным утверждениям

$$\underline{\Delta} = \frac{1}{n} \left[\frac{S_A^2/(s-1)}{C'S^2/(n-1)s} - 1 \right] \leq \Delta. \quad (44)$$

Соответствующие верхние доверительные границы для Δ получаются из критериев для гипотез $\Delta \geq \Delta_0$. Области принятия для этих последних имеют вид $W^* \geq C'$, где W^* определяется по

(43), а C' — равенством $\int_{C'}^{\infty} F_{s-1, (n-1)s} = 1 - \alpha$. Соответствующие

доверительные границы имеют вид

$$\Delta \leq \frac{1}{n} \left[\frac{S_A^2/(s-1)}{C'S^2/(n-1)s} - 1 \right] = \bar{\Delta}. \quad (45)$$

Как доверительные множества (44), так и доверительные множества (45) инвариантны относительно группы преобразований, порожденной группами, рассматривавшимися при проверке гипотез, и, следовательно, среди инвариантных являются равномерно наиболее точными.

Когда $\underline{\Delta}$ отрицательно, доверительное множество $(\underline{\Delta}, \infty)$ содержит все возможные значения параметра Δ . При малых Δ это будет происходить с большой вероятностью ($1 - \alpha$ для $\Delta = 0$), чего следовало ожидать, так как $\underline{\Delta}$ должно быть надежной нижней границей для величины, которая равна нулю или очень близка к нему. Более неприятна возможность получить отрицательное значение для $\bar{\Delta}$, так что доверительное множество $(-\infty, \bar{\Delta})$ оказывается пустым. Интерпретация подсказывается тем фактом, что это может быть тогда и только тогда, когда гипотеза $\Delta \geq \Delta_0$ отклоняется при любых положительных значениях Δ_0 . Последнее можно воспринимать как указание на неудачный выбор модели. Правда, следует иметь в виду, что при малых Δ вероятность события $\bar{\Delta} < 0$ близка к α , даже если требуемые предпосылки выполнены; так что изредка это событие может наблюдаться.

Критерии для гипотез $\Delta \leq \Delta_0$ и $\Delta \geq \Delta_0$ являются не только РНМ инвариантными, но и РНМ несмещенными. РНМ несмещенный критерий существует и для гипотезы $\Delta = \Delta_0$ при двусторонней альтернативе $\Delta \neq \Delta_0$. Это вытекает из того факта, что

совместные распределения величин Z образуют экспоненциальное семейство. Доверительные множества, связанные с этими тремя семействами критериев, оказываются равномерно наиболее точными в классе несмещенных (см. задачу 19). То обстоятельство, что в модели II существуют оптимальные несмещенные процедуры, а в соответствующей модели I они отсутствуют, объясняется различной структурой двух гипотез. Гипотеза $\sigma_A^2 = 0$ в модели II налагает одно ограничение, так как касается единственного параметра σ_A^2 . С другой стороны, соответствующая гипотеза $\sum_{i=1}^s \alpha_i^2 = 0$ в модели I касается значений s параметров $\alpha_1, \dots, \alpha_s$. Так как $s - 1$ из них независимы, то число ограничений равно $s - 1$.

8. Классификации по подчиненности («гнездовые» классификации)

Теория предыдущего раздела неприменима даже в столь простой ситуации, как одинарная классификация с неравными численностями различных классов (см. задачу 22). Однако принцип несмещенности может быть распространен на важный случай «гнездовых» (иерархических) классификаций с равными численностями классов. Характер этого обобщения легко понять, если провести его для случая двух факторов. Случай большего числа факторов охватывается по индукции.

Возвращаясь к примеру с партиями какого-либо продукта, допустим, что одной партии сырья достаточно для производства нескольких партий окончательного продукта. Пусть экспериментальный материал состоит из ab партий, причем каждая из a партий сырья идет на изготовление b партий продукта. И пусть из каждой партии взята выборка объема n . Тогда формула (41) принимает вид

$$X_{ijk} = \mu + A_i + B_{ij} + U_{ijk} \quad (46)$$

$$(i = 1, \dots, a; j = 1, \dots, b; k = 1, \dots, n),$$

где A_i обозначает эффект, связанный с i -й партией сырья, B_{ij} — с j -й партией окончательного продукта, полученного из этой партии сырья, U_{ijk} — с k -й единицей, выбранной из последней партии. Все эти величины предполагаются независимыми и распределенными нормально с нулевыми средними значениями и дисперсиями, равными соответственно σ_A^2 , σ_B^2 и σ^2 . Основная часть доказательства по индукции состоит в указании ортогонального перехода к величинам Z_{ijk} , совместная плотность которых,

с точностью до постоянного множителя, равна

$$\exp \left[-\frac{1}{2(\sigma^2 + n\sigma_B^2 + bn\sigma_A^2)} \left((z_{i11} - \sqrt{abn} \mu)^2 + \sum_{i=2}^a z_{i11}^2 \right) - \right. \\ \left. - \frac{1}{2(\sigma^2 + n\sigma_B^2)} \sum_{i=1}^a \sum_{j=2}^b z_{ij1}^2 - \frac{1}{2\sigma^2} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=2}^n z_{ijk}^2 \right]. \quad (47)$$

Прежде всего, при каждом i, j существует ортогональное преобразование величин $(X_{ij1}, \dots, X_{ijn})$ в $(Y_{ij1}, \dots, Y_{ijn})$ такое, что

$$Y_{ij1} = \sqrt{n} X_{ij.} = \sqrt{n} \mu + \sqrt{n} (A_i + B_{ij} + U_{ij}).$$

Как и в случае простой классификации, величины Y_{ijk} с $k > 1$ зависят только от U , независимы между собой и распределены нормально с нулевым средним и дисперсией σ^2 и не зависят от U_{ij} . С другой стороны, величины Y_{ij1} имеют ту же структуру, что величины Y_{ij} в одинарной классификации

$$Y_{ij1} = \mu' + A'_i + U'_{ij},$$

где $\mu' = \sqrt{n} \mu$, $A'_i = \sqrt{n} A_i$, $U'_{ij} = \sqrt{n} (B_{ij} + U_{ij})$ и где дисперсии A'_i и U'_{ij} равны $\sigma_A'^2 = n\sigma_A^2$ и $\sigma'^2 = \sigma^2 + n\sigma_B^2$ соответственно. Эти величины могут быть, следовательно, преобразованы в величины Z_{ij1} , совместная плотность которых задается формулой (42) с заменой Z_{ij} на Z_{ij1} . Полагая $Z_{ijk} = Y_{ijk}$ при $k > 1$, мы видим, что совместная плотность величин Z_{ijk} определяется выражением (47).

На основе формулы (47) можно проверить гипотезы H_1 : $\sigma_A'^2/(\sigma^2 + n\sigma_B^2) \leq \Delta_0$ и H_2 : $\sigma_B'^2/\sigma^2 \leq \Delta_0$, которые устанавливают, что одна или другая из классификаций меньше влияет на исход испытаний. Пусть

$$S_A^2 = \sum_{i=2}^a Z_{i11}^2, \quad S_B^2 = \sum_{i=1}^a \sum_{j=2}^b Z_{ij1}^2, \quad S^2 = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=2}^n Z_{ijk}^2.$$

Чтобы получить критерий для H_1 , кажется соблазнительной попытка исключить S^2 с помощью инвариантности относительно умножения величин Z_{ijk} с $k > 1$ на произвольную постоянную. Однако эти преобразования не оставляют выражение (47) инвариантным, так как они не всегда сохраняют тот факт, что σ^2 является наименьшей из трех дисперсий: σ^2 , $\sigma^2 + n\sigma_B^2$ и $\sigma^2 + n\sigma_B^2 + bn\sigma_A^2$. Вместо этого мы рассмотрим задачу с точки зрения несмещенности. Для любого несмещенного критерия гипотезы H_1 вероятность отклонения равна α каждый раз как $\sigma_A'^2/(\sigma^2 + n\sigma_B^2) = \Delta_0$; в частности, это имеет место, если дисперсии равны σ^2 , τ_0^2

и $(1 + bn\Delta_0)\tau_0^2$, где τ_0^2 — любое фиксированное число и $\sigma^2 < \tau_0^2$. Теми же приемами, что и в главе 4, можно показать, что условная вероятность отклонения гипотезы при данном $S^2 = s^2$ должна быть равна α для почти всех значений s^2 . При фиксированном S^2 совместное распределение оставшихся величин имеет такую же форму, как (42), после удаления Z_{111} . РНМ несмещенный условный критерий при данном $S^2 = s^2$ имеет критическую область

$$W_1^* = \frac{1}{1 + bn\Delta_0} \cdot \frac{S_A^2/(a-1)}{S_B^2/(b-1)a} \geq C_1. \quad (48)$$

Так как S_A^2 и S_B^2 не зависят от S^2 , то константа C_1 определяется тем условием, что при $\sigma_A^2/(\sigma^2 + n\sigma_B^2) = \Delta_0$ статистика W_1^* имеет распределение $F_{a-1, (b-1)a}$, т. е., в частности, не зависит от s . Критерий (48) является, очевидно, несмещенным и, следовательно, РНМ несмещенным.

В применении к H_2 совершенно аналогичные рассуждения показывают, что РНМ несмещенный критерий имеет критическую область

$$W_2^* = \frac{1}{1 + n\Delta_0} \cdot \frac{S_B^2/(b-1)a}{S^2/(n-1)ab} \geq C_2, \quad (49)$$

где C_2 определяется тем условием, что при $\sigma_B^2/\sigma^2 = \Delta_0$ статистика W_2^* имеет распределение $F_{(b-1)a, (n-1)ab}$.

Остается выразить статистики S_A^2 , S_B^2 и S^2 в терминах величин X . Из соответствующих выражений для одинарной классификации получаем

$$S_A^2 = \sum_{i=1}^a Z_{i11}^2 - Z_{111}^2 = b \sum (Y_{i.1} - Y_{..1})^2,$$

$$S_B^2 = \sum_{i=1}^a \left[\sum_{j=1}^b Z_{ij1}^2 - Z_{i11}^2 \right] = \sum \sum (Y_{ij1} - Y_{i.1})^2$$

и

$$\begin{aligned} S^2 &= \sum_{i=1}^a \sum_{j=1}^b \left[\sum_{k=1}^n Y_{ijk}^2 - Y_{ij1}^2 \right] = \sum_i \sum_j \left[\sum_{k=1}^n U_{ijk}^2 - nU_{ij.}^2 \right] = \\ &= \sum_i \sum_j \sum_k (U_{ijk} - U_{ij.})^2. \end{aligned}$$

Следовательно,

$$S_A^2 = bn \sum (X_{i..} - X_{...})^2, \quad S_B^2 = n \sum \sum (X_{ij.} - X_{i..})^2,$$

$$S^2 = \sum \sum \sum (X_{ijk} - X_{ij.})^2. \quad (50)$$

Из выражений для этих статистик в терминах величин Z видно, что их математические ожидания равны $E[S_A^2/(a-1)] = \sigma^2 + n\sigma_B^2 + bn\sigma_A^2$, $E[S_B^2/(b-1)a] = \sigma^2 + n\sigma_B^2$ и $E[S^2/(n-1)ab] = \sigma^2$. Поэтому разложение

$$\sum \sum \sum (X_{ijk} - X_{...})^2 = S_A^2 + S_B^2 + S^2$$

служит основой для анализа дисперсий X_{ijk}

$$\text{Var}(X_{ijk}) = \sigma_A^2 + \sigma_B^2 + \sigma^2,$$

давая оценки для компонент дисперсии σ_A^2 , σ_B^2 , σ^2 , и для некоторых отношений этих компонент.

Двойные классификации по подчиненности могут встречаться в форме смешанных моделей. Предположим, возвращаясь к предыдущим иллюстрациям, что фирма производит продукцию на нескольких заводах. Пусть α_i обозначает эффект i -го завода (эта величина будет постоянной, так как при повторении эксперимента заводы не меняются), B_{ij} — эффект партий, U_{ijk} — индивидуальный эффект. Тогда наблюдения имеют следующую структуру:

$$X_{ijk} = \mu + \alpha_i + B_{ij} + U_{ijk}. \quad (51)$$

Вместо полного приведения величин X к канонической форме в терминах величин Z здесь удобнее осуществить редукцию к величинам Y (так что $Y_{ij1} = \sqrt{n} X_{ij1}$) и первое из двух преобразований, переводящих Y в Z . Обозначим так полученные величины W_{ijk} . Они удовлетворяют соотношениям $W_{i11} = \sqrt{b} Y_{i1}$, $W_{ijk} = Y_{ijk}$ при $k > 1$ и

$$\sum_{i=1}^a (W_{i11} - W_{\cdot 11})^2 = S_A^2, \quad \sum_{i=1}^a \sum_{j=2}^b W_{ij1}^2 = S_B^2, \quad \sum_{i=1}^a \sum_{j=1}^b \sum_{k=2}^n W_{ijk}^2 = S^2,$$

где S_A^2 , S_B^2 и S^2 определяются по (50). С точностью до константы совместная плотность величин W равна

$$\exp \left[-\frac{1}{2(\sigma^2 + n\sigma_B^2)} \left(\sum_{i=1}^a (w_{i11} - \mu - \alpha_i)^2 + \sum_{i=1}^a \sum_{j=2}^b w_{ij1}^2 \right) - \frac{1}{2\sigma^2} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=2}^n w_{ijk}^2 \right]. \quad (52)$$

Отсюда ясно видно отличие задачи проверки гипотез о малости эффектов отдельных заводов

$$H: \alpha_1 = \dots = \alpha_a = 0 \quad \text{или} \quad H': \frac{\sum \alpha_i^2}{(\sigma^2 + n\sigma_B^2)} \leq \Delta_0$$

и задачи проверки аналогичной гипотезы для эффектов партий: $\sigma_B^2/\sigma^2 \leq \Delta_0$. Первые две гипотезы включаются в модель I (линейная гипотеза). Как и прежде, из несмещенности вытекает, что условная вероятность отклонения при $S^2 = s^2$ равна почти всюду величине α . При фиксированном S^2 гипотеза H проверяется как линейная гипотеза и РНМ инвариантный условный критерий при условии $S^2 = s^2$ имеет критическую область (48) с $\Delta_0 = 0$. Константа C_1 снова не зависит от S^2 и найденный критерий является РНМ в классе всех критериев, как несмещенных, так и инвариантных. Критерий с аналогичным свойством существует и для гипотезы H' . Его критическая область имеет вид

$$\frac{S_A^2/(a-1)}{S_B^2/(b-1)a} \geq C',$$

где C' определяется с помощью нецентрального F -распределения (a не из центрального F -распределения как раньше; см. задачу 5).

С другой стороны, гипотеза $\sigma_B^2/\sigma^2 \leq \Delta_0$ относится, по существу, к модели II. Она инвариантна относительно добавления произвольной постоянной к каждой из величин W_{i11} . При этом преобразовании максимальный инвариант образует суммы $\sum_{i=1}^a \sum_{j=2}^b W_{ij1}^2$ и $\sum_{i=1}^a \sum_{j=1}^b \sum_{k=2}^n W_{ijk}^2$. Задача сводится, таким образом, по структуре к чистой модели II с одинарной классификацией. Критерий, как и раньше, определяется по (49). Он является и РНМ инвариантным, и РНМ несмещенным.

Смешанная модель с двумя факторами, которые взаимодействуют, будет рассмотрена в примере 11.

9. Многомерная линейная гипотеза

Одномерная линейная модель раздела 1 возникает при изучении воздействия различных экспериментальных условий (факторов) на единственную характеристику, такую, как размер урожая, вес, продолжительность жизни, давление крови и т. п. Эта характеристика предполагается распределенной нормально со средним, зависящим от исследуемых факторов, и с дисперсией, не зависящей от них. Мы рассмотрим теперь многомерный аналог этой модели, который применим к изучению действия одного или большего числа факторов на несколько характеристик, например, к изучению влияния рациона питания коров на жирность и на количество молока.

Многомерным аналогом действительной нормально распределенной случайной величины является случайный вектор

$(X_1, \dots, X_p)$ с многомерной нормальной плотностью вероятности

$$\frac{\sqrt{|A|}}{(2\pi)^{\frac{1}{2}p}} \exp \left[-\frac{1}{2} \sum \sum a_{ij} (x_i - \xi_i) (x_j - \xi_j) \right], \quad (53)$$

где матрица $A = (a_{ij})$ положительно определена и $|A|$ обозначает ее детерминант. Средние значения и матрица ковариаций X_i задаются формулами

$$E(X_i) = \xi_i, \quad E(X_i - \xi_i)(X_j - \xi_j) = \sigma_{ij}, \quad (\sigma_{ij}) = A^{-1}. \quad (54)$$

Рассмотрим теперь n независимых многомерных нормальных векторов $X_\alpha = (X_{\alpha 1}, \dots, X_{\alpha p})$, $\alpha = 1, \dots, n$, со средними $E(X_{\alpha i}) = \xi_{\alpha i}$ и общей матрицей ковариаций A^{-1} . Как и в одномерном случае, *многомерная линейная гипотеза* формулируется в терминах двух линейных подпространств Π_Ω и Π_ω n -мерного пространства, имеющих размерности $s < n$ и $0 \leq s - r \leq s$ соответственно. Предполагается известным, что при всех $i = 1, \dots, p$ векторы $(\xi_{1i}, \dots, \xi_{ni})$ лежат в Π_Ω ; проверяемая гипотеза утверждает, что они лежат в Π_ω . Эта проблема приводится к канонической форме применением к каждому из p векторов $(X_{1i}, \dots, X_{ni})$ ортогонального преобразования (1). Если

$$X = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ \vdots & & \vdots \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

и если новые величины суть $X_{\alpha i}^*$, то преобразование может быть записано в следующей матричной форме:

$$X^* = CX,$$

где $C = (c_{\alpha\beta})$ — ортогональная матрица.

Чтобы найти совместное распределение величин $X_{\alpha i}^*$, рассмотрим сначала ковариацию любых двух из них, скажем

$X_{\alpha i}^* = \sum_{\gamma=1}^n c_{\alpha\gamma} X_{\gamma i}$ и $X_{\beta j}^* = \sum_{\delta=1}^n c_{\beta\delta} X_{\delta j}$. Используя тот факт, что ковариация величин $X_{\gamma i}$ и $X_{\delta j}$ равна нулю при $\gamma \neq \delta$ и равна σ_{ij} при $\gamma = \delta$, мы получаем

$$\text{Cov}(X_{\alpha i}^*, X_{\beta j}^*) =$$

$$= \sum_{\gamma=1}^n \sum_{\delta=1}^n c_{\alpha\gamma} c_{\beta\delta} \text{Cov}(X_{\gamma i}, X_{\delta j}) = \sigma_{ij} \sum_{\gamma=1}^n c_{\alpha\gamma} c_{\beta\gamma} = \begin{cases} \sigma_{ij}, & \text{если } \alpha = \beta, \\ 0, & \text{если } \alpha \neq \beta. \end{cases}$$

Таким образом, строки матрицы X^* снова являются независимыми многомерными нормальными векторами с общей матри-

цей ковариаций A^{-1} . Как и в одномерном случае, векторы средних значений должны удовлетворять соотношениям

$$\xi_{s+1, i}^* = \dots = \xi_{ni}^* = 0 \quad (i = 1, \dots, p)$$

(равносильным принадлежности Ω) и проверяемая гипотеза принимает вид

$$H: \xi_{1i}^* = \dots = \xi_{ri}^* = 0 \quad (i = 1, \dots, p).$$

Изменим обозначения. Буквами Y , U и Z будут обозначаться соответственно первые r , последующие $s - r$ и последние $m = n - s$ выборочных векторов. Мы приходим тогда к следующей канонической форме. Векторы Y_α , U_β , Z_γ ($\alpha = 1, \dots, r$; $\beta = 1, \dots, s - r$; $\gamma = 1, \dots, m$) независимы и имеют p -мерные нормальные распределения с общей матрицей ковариаций A^{-1} . Дано, что средние величин Z равны нулю, а проверяемая гипотеза H состоит в том, что средние значения величин Y также равны нулю. Если

$$Y = \begin{pmatrix} Y_{11} & \dots & Y_{1p} \\ \vdots & & \vdots \\ Y_{r1} & \dots & Y_{rp} \end{pmatrix} \quad \text{и} \quad Z = \begin{pmatrix} Z_{11} & \dots & Z_{1p} \\ \vdots & & \vdots \\ Z_{m1} & \dots & Z_{mp} \end{pmatrix},$$

то, как можно показать, принципы инвариантности и достаточности позволяют редуцировать данные наблюдений к двум $p \times p$ -матрицам $Y'Y$ и $Z'Z$. Поэтому удобно иметь выражение для этих статистик в терминах первоначальных наблюдений.

Как и в одномерном случае, обозначим $(\hat{\xi}_{1i}, \dots, \hat{\xi}_{ni})$ и $(\hat{\xi}_{1i}, \dots, \hat{\xi}_{ni})$ проекции вектора $(X_{1i}, \dots, X_{ni})$ на Π_Ω и Π_ω . Тогда

$$\sum_{\alpha=1}^n (X_{\alpha i} - \hat{\xi}_{\alpha i})(X_{\alpha j} - \hat{\xi}_{\alpha j})$$

представляет собой скалярное произведение двух векторов, каждый из которых есть разность между некоторым вектором и его проекцией на Π_Ω . Отсюда следует, что это выражение остается инвариантным при ортогональных преобразованиях системы координат, в которой выражены результаты наблюдений. Преобразование

$$C \begin{pmatrix} X_{1i} \\ \vdots \\ X_{ni} \end{pmatrix}$$

можно интерпретировать как запись вектора $(X_{1i}, \dots, X_{ni})$ в новой системе координат, у которой первые s координатных осей лежат в Π_Ω . Проекция преобразованного вектора $(Y_{1i}, \dots, Y_{ri}, U_{1i}, \dots, U_{s-r, i}, Z_{1i}, \dots, Z_{mi})$ на Π_Ω равна $(Y_{1i}, \dots, Y_{ri}, U_{1i}, \dots$

$\dots, U_{s-r, i}, 0, \dots, 0)$, так что разность между этим вектором и его проекцией равна $(0, \dots, 0, Z_{1i}, \dots, Z_{mi})$. Элемент, стоящий в матрице $Z'Z$ на (i, j) -м месте, равен

$$\sum_{\gamma=1}^m Z_{\gamma i} Z_{\gamma j} = \sum_{\alpha=1}^n (X_{\alpha i} - \hat{\xi}_{\alpha i})(X_{\alpha j} - \hat{\xi}_{\alpha j}). \quad (55)$$

Аналогично, проекция преобразованного вектора $(Y_{1i}, \dots, Y_{ri}, U_{1i}, \dots, U_{s-r, i}, Z_{1i}, \dots, Z_{mi})$ на Π_{ω} равна $(0, \dots, 0, U_{1i}, \dots, U_{s-r, i}, Z_{1i}, \dots, Z_{mi})$ и разность между его проекциями на Π_{Ω} и Π_{ω} равна $(Y_{1i}, \dots, Y_{ri}, 0, \dots, 0, 0, \dots, 0)$. Мы видим, что сумма $\sum_{\beta=1}^r Y_{\beta i} Y_{\beta j}$ равна скалярному произведению разностей этих проекций, вычисленных для i -го и j -го векторов. Сравнивая эту сумму с выражением для того же самого скалярного произведения в первоначальной системе координат, мы находим (i, j) -й элемент матрицы $Y'Y$

$$\sum_{\beta=1}^r Y_{\beta i} Y_{\beta j} = \sum_{\alpha=1}^n (\hat{\xi}_{\alpha i} - \hat{\xi}_{\alpha i})(\hat{\xi}_{\alpha j} - \hat{\xi}_{\alpha j}). \quad (56)$$

10. Редукция с учетом инвариантности

Многомерная линейная гипотеза, каноническая форма которой была указана в предыдущем разделе, является инвариантной относительно некоторых групп преобразований. Чтобы найти максимальные инварианты относительно этих групп, нам потребуется, в дополнение к стандартным теоремам о квадратичных формах, также следующая лемма.

Лемма 1. Пусть M — любая $t \times r$ -матрица. Тогда (I) матрица $M'M$ положительно полуопределена, (II) ее ранг равен рангу M , так что, в частности, $M'M$ не будет вырожденной тогда и только тогда, когда $t \geq r$ и ранг M равен r .

Доказательство. (I) Рассмотрим квадратичную форму $Q = u'(M'M)u$. Если $w = Mu$, то

$$Q = w'w \geq 0.$$

(II) Сумма квадратов $w'w$ равна нулю в том и только том случае, когда вектор w равен нулю. Требуемый результат вытекает теперь из того факта, что решения u системы уравнений $Mu = 0$ образуют линейное пространство размерности $r - \rho$, где ρ — ранг M .

Мы рассмотрим теперь три группы, относительно которых инвариантна изучаемая проблема.

Группа G_1 . Добавление произвольной постоянной $d_{\beta i}$ к каждой из величин $U_{\beta i}$ оставляет проблему инвариантной. Этим

исключаются величины U , так как совокупность величин Y и Z образует максимальный инвариант.

Группа G_2 . В процессе приведения проблемы к канонической форме было показано, что ортогональное преобразование

$$Y^* = CY$$

не изменяет ни факта независимости строк матрицы Y , ни матрицы ковариаций этих строк. Средние значения векторов Y^* равны нулю в том и только том случае, когда равны нулю средние значения векторов Y и, следовательно, проблема остается инвариантной при этих преобразованиях.

Матрица $Y'Y$, составленная из скалярных произведений векторов-столбцов матрицы Y , является инвариантной относительно группы G_2 , так как $Y'^*Y^* = Y'C'CY = Y'Y$. Мы покажем, что равенство $Y'Y = Y'^*Y^*$ влечет существование такой ортогональной матрицы C , что $Y^* = CY$. Отсюда будет следовать, что матрица $Y'Y$ является максимальным инвариантом относительно G_2 . Рассмотрим сначала случай $r = p$. Не ограничивая общности, мы можем предположить, что первые p столбцов матрицы Y линейно независимы, так как исключительное множество величин Y , для которых это не выполняется, имеет меру нуль. Из равенства $Y'Y = Y'^*Y^*$ вытекает, что матрица $C = Y^*Y^{-1}$ ортогональна и что $Y^* = CY$, а это и требовалось доказать. Предположим теперь, что $r > p$. По-прежнему мы, не ограничивая общности, предположим, что первые p столбцов матрицы Y линейно независимы. Так как для любых двух p -мерных подпространств r -мерного пространства существует ортогональное преобразование, переводящее одно из них в другое, то можно предположить, что (после надлежащего ортогонального преобразования) p векторов-столбцов матриц Y и Y^* лежат в одном и том же p -мерном пространстве. Тем самым задача сводится к случаю $r = p$. Если, наконец, $r < p$, то можно предположить, что первые r столбцов матрицы V линейно независимы. Обозначая матрицы, образованные первыми r и последними $p - r$ столбцами матрицы Y , буквами Y_1 и Y_2 соответственно, так что

$$Y = (Y_1 Y_2),$$

мы получим $Y_1'^* Y_1^* = Y_1' Y_1$ и, в силу сказанного выше, найдется ортогональная матрица B такая, что $Y_1^* = B Y_1$. Из соотношения $Y_1'^* Y_2^* = Y_1' Y_2$ вытекает, что $Y_2^* = (Y_1'^*)^{-1} Y_1' Y_2 = B Y_2$, чем и завершается доказательство.

Аналогично, проблема остается инвариантной относительно преобразований

$$Z^* = DZ,$$

для которых максимальным инвариантом служит $Z'Z$. Редукция к $Z'Z$ может быть проведена и на основе того факта, что $Z'Z$

вместе с Y и U образует систему достаточных статистик. Как бы то ни было, при группах G_1 и G_2 данные редуцируются к двум матрицам $V = Y'Y$ и $S = Z'Z$.

Группа G_3 . Мы наложим теперь ограничение $m \geq p$ (см. задачу 24), которое обеспечивает наличие числа степеней свободы, достаточного для разумной оценки матрицы ковариаций. Рассмотрим преобразования

$$Y^* = YB, \quad Z^* = ZB,$$

где B — любая невырожденная матрица. Эти преобразования действуют раздельно на каждый из независимых нормальных векторов $(Y_{\beta 1}, \dots, Y_{\beta p})$, $(Z_{\gamma 1}, \dots, Z_{\gamma p})$ и, очевидно, оставляют проблему инвариантной. Индуцированные преобразования матриц $V = Y'Y$ и $S = Z'Z$ имеют вид

$$V^* = B'VB, \quad S^* = B'SB.$$

Так как $|B'(V - \lambda S)B| = |B|^2|V - \lambda S|$, то корни характеристического уравнения

$$|V - \lambda S| = 0 \tag{57}$$

инвариантны относительно этой группы. Покажем, что их совокупность образует максимальный инвариант. Допустим, что уравнения $|V - \lambda S| = 0$ и $|V^* - \lambda S^*| = 0$ имеют одни и те же корни. Опять-таки не ограничивая общности, мы можем предположить, что первые p строк матрицы Z являются линейно независимыми, так что матрицы Z и Z^* имеют ранг p . По лемме 1 матрица S будет положительно определенной и из теорем об одновременном приведении двух квадратичных форм к сумме квадратов *) вытекает, что существует невырожденная матрица B_1 , такая, что

$$B_1'VB_1 = \Lambda, \quad B_1'SB_1 = I,$$

где Λ — диагональная матрица, элементы которой суть корни уравнения (57), а I — единичная матрица. Существует также матрица B_2 , такая что

$$B_2'V^*B_2 = \Lambda, \quad B_2'S^*B_2 = I.$$

Таким образом, $B = B_1B_2^{-1}$ преобразует V в V^* и S в S^* .

Среди корней уравнения (57), образующих максимальный инвариант, некоторые могут быть равны нулю. Действительно, так как эти корни являются диагональными элементами матрицы Λ , то число ненулевых корней равно рангу Λ , а следовательно, ран-

*) См., например, Anderson, An Introduction to Multivariate Statistical Analysis, New York, John Wiley & Sons, 1958, Theorem 3 of Appendix 1. [Русский перевод: Андерсон, Введение в многомерный статистический анализ, Физматгиз, 1963.]

гу $V = B_1'^{-1} \Lambda B_1^{-1}$, который по лемме 1, равен $\min(p, r)$. Если это число больше 1, то РНМ инвариантный критерий не существует. Случай $p = 1$ — это случай одномерной линейной гипотезы, уже изученный в разделе 1. Мы рассмотрим теперь остающуюся возможность $r = 1$.

При $r = 1$ уравнение (57) и эквивалентное ему уравнение

$$|VS^{-1} - \lambda I| = 0$$

имеют только один ненулевой корень. Записывая определитель в форме многочлена по степеням λ , мы видим, что коэффициенты при степенях λ , меньших $p - 1$, должны быть равны нулю и что поэтому уравнение имеет вид

$$(-\lambda)^p + W(-\lambda)^{p-1} = 0,$$

где W — сумма диагональных элементов (след) матрицы VS^{-1} . Пусть S^{ij} обозначает (i, j) -й элемент матрицы S^{-1} и пусть $(Y_1, \dots, Y_p)$ — единственный Y -вектор. Тогда простые вычисления показывают, что

$$W = \sum_{i=1}^p \sum_{j=1}^p S^{ij} Y_i Y_j. \quad (58)$$

Таким образом, необходимое и достаточное условие инвариантности критерия относительно групп G_1 , G_2 и G_3 состоит в том, что он зависит только от W .

Распределение W зависит только от максимального инварианта в пространстве параметров, который равен

$$\psi^2 = \sum_{i=1}^p \sum_{j=1}^p a_{ij} \eta_i \eta_j, \quad (59)$$

где $\eta_i = E(Y_i)$. Плотность вероятности W имеет вид (см. задачи 28—30)

$$p_\psi(w) = e^{-\frac{1}{2}\psi^2} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{2}\psi^2\right)^k}{k!} c_k \frac{w^{\frac{1}{2}p-1+k}}{(1+w)^{\frac{1}{2}(m+1)+k}}. \quad (60)$$

Она совпадает с плотностью (6) статистики для одномерного случая с $r = p$ и $n - s = m + 1 - p$. При любых $\psi_0 < \psi_1$ отношение $p_{\psi_1}(w)/p_{\psi_0}(w)$ является возрастающей функцией от w . Из леммы Неймана — Пирсона вытекает, что наиболее мощный инвариантный критерий для проверки гипотезы $H: \eta_1 = \dots = \eta_p = 0$ отвергает ее при слишком больших W или, что то же самое, при

$$\frac{m+1-p}{p} W > C. \quad (61)$$

Величина mW , которая при $p = 1$ сводится к квадрату t -статистики Стьюдента, по сути дела, является T^2 -статистикой Хотеллинга. Константа C определяется из того условия, что при $\psi = 0$ статистика $(m + 1 - p)W/p$ имеет F -распределение с p и $m + 1 - p$ степенями свободы. Как и в одномерном случае, РНМ инвариантный критерий существует и для более общей гипотезы H' : $\psi^2 \leq \psi_0^2$. Его критическая область имеет вид $W > C'$.

Так как при $\min(p, r) > 1$ РНМ инвариантный критерий не существует, то были предложены различные критерии, основанные на тех или иных функциях от корней уравнения (57), например, на сумме корней, на максимуме или минимуме корней, на произведении $\prod_{i=1}^p (1 + \lambda_i)^{-1}$ (это критерий отношения правдоподобия).

11. Применения

Различные одномерные линейные гипотезы с $r = 1$, фиксирующие, например, среднее значение нормального распределения, разность средних значений двух нормальных совокупностей с равными дисперсиями, наклон линии регрессии и т. п., могут быть распространены теперь на многомерный случай.

Пример 8. Пусть $(X_{\alpha 1}, \dots, X_{\alpha p})$, $\alpha = 1, \dots, n$ — выборка из многомерного нормального распределения со средним $(\xi_1, \dots, \xi_p)$ и матрицей ковариаций A^{-1} , причем и то и другое считается неизвестным. Рассмотрим задачу проверки гипотезы H : $\xi_1 = \dots = \xi_p = 0$. В примере 4 было показано, что

$$\hat{\xi}_{\alpha i} = \sum_{\beta=1}^n \frac{X_{\beta i}}{n} = X_{.i}; \quad \hat{\xi}_{\alpha t} = 0.$$

В силу (55) (i, j) -й элемент матрицы $S = Z'Z$ равен

$$S_{ij} = \sum_{\alpha=1}^n (X_{\alpha i} - X_{.i})(X_{\alpha j} - X_{.j})$$

и по (56)

$$Y_i Y_j = n X_{.i} X_{.j}.$$

Критерий основан на статистике W (см. (58)) и имеет критическую область (61) с $s = 1$ и $m = n - s = n - 1$. Статистика $T^2 = (n - 1)W$ известна под названием T^2 -статистики Хотеллинга.

Пример 9. Пусть $(X_{\alpha 1}^{(1)}, \dots, X_{\alpha p}^{(1)})$, $\alpha = 1, \dots, n_1$ и $(X_{\beta 1}^{(2)}, \dots, X_{\beta p}^{(2)})$, $\beta = 1, \dots, n_2$ — независимые выборки из многомерных нормальных распределений с одной и той же матрицей ковариаций A^{-1} и средними $(\xi_1^{(1)}, \dots, \xi_p^{(1)})$ и $(\xi_1^{(2)}, \dots, \xi_p^{(2)})$. Рассмотрим гипотезу H : $\xi_i^{(1)} = \xi_i^{(2)}$ для $i = 1, \dots, p$. Тогда $s = 2$ и, как следует из примера 5, при всех α и β

$$\hat{\xi}_{\alpha i}^{(1)} = X_{.i}^{(1)}, \quad \hat{\xi}_{\beta i}^{(2)} = X_{.i}^{(2)}$$

и

$$\hat{\xi}_{\alpha i}^{(1)} = \hat{\xi}_{\beta i}^{(2)} = \frac{\sum_{\alpha=1}^{n_1} X_{\alpha i} + \sum_{\beta=1}^{n_2} X_{\beta i}}{(n_1 + n_2)} = \bar{X}_i.$$

Следовательно,

$$S_{ij} = \sum_{\alpha=1}^{n_1} (X_{\alpha i}^{(1)} - X_{\cdot i}^{(1)}) (X_{\alpha j}^{(1)} - X_{\cdot j}^{(1)}) + \sum_{\beta=1}^{n_2} (X_{\beta i}^{(2)} - X_{\cdot i}^{(2)}) (X_{\beta j}^{(2)} - X_{\cdot j}^{(2)})$$

и выражение для $Y_i Y_j$ упрощается:

$$Y_i Y_j = n_1 (X_{\cdot i}^{(1)} - \bar{X}_i) (X_{\cdot j}^{(1)} - \bar{X}_j) + n_2 (X_{\cdot i}^{(2)} - \bar{X}_i) (X_{\cdot j}^{(2)} - \bar{X}_j)^*.$$

Помимо задачи проверки вышеприведенных и аналогичных обобщений одномерных гипотез, критерий (61) применим также и к некоторым проблемам, которые сами по себе не укладываются в схему линейных гипотез, как она описана в разделе 9, но сводятся к ней с помощью соображений инвариантности. Пусть $(X_{\alpha 1}, \dots, X_{\alpha p})$, $\alpha = 1, \dots, n$ — выборка из многомерного нормального распределения со средним $(\xi_1, \dots, \xi_p)$ и матрицей ковариаций A^{-1} . Рассмотрим гипотезу о том, что вектор $(\xi_1, \dots, \xi_p)$ лежит в $(p-r)$ -мерном подпространстве p -мерного пространства. Обычным образом можно преобразовать результаты наблюдений в величины $(Y_{\alpha 1}, \dots, Y_{\alpha r}, Z_{\alpha 1}, \dots, Z_{\alpha l})$, $p = r + l$, образующие выборку из p -мерного нормального распределения со средним $(\eta_1, \dots, \eta_r, \zeta_1, \dots, \zeta_l)$ таким, что проверяемая гипотеза принимает форму $H: \eta_1 = \dots = \eta_r = 0$.

Эта задача остается инвариантной относительно некоторой группы линейных преобразований, для которой максимальным инвариантом является совокупность величин Y . В терминах Y гипотеза сводится к рассмотренной в примере 8. Следовательно, существует РНМ инвариантный критерий гипотезы H , задаваемый формулой (61) с $p = r$ и $m = n - 1$. Перед тем как доказать, что величины Z могут быть отброшены, мы укажем два примера задач этого типа.

Пример 10. Пусть $(X_{\alpha 1}, \dots, X_{\alpha q}, X_{\alpha, q+1}, \dots, X_{\alpha, 2q})$, $\alpha = 1, \dots, n$, — выборка из многомерного нормального распределения. Рассмотрим задачу проверки $H: \xi_{q+i} = \xi_i$ для $i = 1, \dots, q$. Она может возникнуть, например, когда $X_{\alpha 1}, \dots, X_{\alpha q}$ и $X_{\alpha, q+1}, \dots, X_{\alpha, 2q}$ суть результаты q измерений одного и того же объекта в два различных периода после некоторой обработки. В терминах переменных

$$Y_{\alpha i} = X_{\alpha, q+i} - X_{\alpha i}; \quad Z_{\alpha i} = X_{\alpha i} \quad (\alpha = 1, \dots, n; \quad i = 1, \dots, q)$$

гипотеза принимает вид $\eta_i = E(Y_{\alpha i}) = 0$ для $i = 1, \dots, q$ и РНМ инвариантный критерий получается из критерия примера 8, примененного к величинам Y , с заменой p на q .

*) Критерий гипотезы H для случая $p > n_1 + n_2 - 2$ рассматривается в работе: Dempster, A high dimensional two-sample significance test, Ann. Math. Stat. 29 (1958), 995—1010.

Пример 11. Пусть $(X_{\alpha 1}, \dots, X_{\alpha p}), \alpha = 1, \dots, n$, — выборка из p -мерного нормального распределения. Рассмотрим задачу проверки гипотезы $H: \xi_1 = \dots = \xi_p$. В терминах новых переменных $Y_{\alpha i} = X_{\alpha i} - X_{\alpha p} (i = 1, \dots, p-1)$ и $Z_{\alpha} = X_{\alpha p}$ гипотеза снова принимает каноническую форму $\eta_1 = \dots = \eta_{p-1} = 0$ и задача сводится к рассмотренной в примере 8 (с заменой p на $p-1$). Возьмем конкретный пример: пусть на фабрике имеется p машин, производящих некоторую продукцию, качество которой измеряется случайной величиной X . Эксперимент состоит в том, что n рабочих обслуживают каждую из машин; при этом $X_{\alpha i}$ — качество, получаемое, когда рабочий с номером α обслуживает i -ю машину. Если указанные n рабочих выбраны наудачу из большой совокупности, то можно предположить, что векторы $(X_{\alpha 1}, \dots, X_{\alpha p})$ образуют выборку из p -мерного нормального распределения. Из двух факторов, влияющих на результат эксперимента, один (машины) фиксирован, а другой (рабочие) случаен в том смысле, что повторение эксперимента происходило бы с теми же самыми машинами, но другими рабочими. Проверяемая гипотеза состоит в том, что эффект, связанный с неслучайным фактором, отсутствует. В этой смешанной модели критерий совершенно другой, чем в соответствующей модели 1, когда оба эффекта фиксированы (что рассматривалось в разделе 4).

Мы вернемся теперь к общему случаю, когда $(Y_{\alpha 1}, \dots, Y_{\alpha r}, Z_{\alpha 1}, \dots, Z_{\alpha l}), \alpha = 1, \dots, n$ — выборка из p -мерного нормального распределения со средним $(\eta_1, \dots, \eta_r, \xi_1, \dots, \xi_l)$ и когда проверяется гипотеза $\eta_1 = \dots = \eta_r = 0$, проиллюстрированная примерами 10 и 11. Интерпретируя временно совокупность pn величин как множество $p = r + l$ векторов в n -мерном пространстве: $(Y_{1i}, \dots, Y_{ni}), i = 1, \dots, r$, и $(Z_{1j}, \dots, Z_{nj}), j = 1, \dots, l$, рассмотрим ортогональное преобразование n -мерного пространства, которое переводит $(x_1, \dots, x_n)$ в $(x'_1, \dots, x'_n)$, причем так, что $x'_1 = \sqrt{n} \bar{x}$. Допустим, что это преобразование применяется к каждому из p наблюдаемых векторов, и пусть оно переводит $(Y_{1i}, \dots, Y_{ni})$ и $(Z_{1j}, \dots, Z_{nj})$ в $(U_{1i}, \dots, U_{ni})$ и $(V_{1j}, \dots, V_{nj})$ соответственно. Тогда, в частности, $U_{1i} = \sqrt{n} Y_{\cdot i}, V_{1j} = \sqrt{n} Z_{\cdot j}$ и совокупности случайных величин $(U_{\alpha 1}, \dots, U_{\alpha r}, V_{\alpha 1}, \dots, V_{\alpha l}), \alpha = 1, \dots, n$, независимы и имеют $(r + l)$ -мерные нормальные распределения с одной и той же матрицей ковариаций и средними значениями $E(U_{1i}) = \sqrt{n} \eta_i, E(V_{1j}) = \sqrt{n} \xi_j$ и $E(U_{\alpha i}) = E(V_{\alpha j}) = 0$ при $\alpha > 1$.

Полагая

$$U = \begin{pmatrix} U_{21} & \dots & U_{2r} \\ \vdots & & \vdots \\ U_{n1} & \dots & U_{nr} \end{pmatrix} \quad \text{и} \quad V = \begin{pmatrix} V_{21} & \dots & V_{2l} \\ \vdots & & \vdots \\ V_{n1} & \dots & V_{nl} \end{pmatrix},$$

мы видим, что проблема остается инвариантной относительно следующих двух групп преобразований.

Группа G_1 . Добавление произвольной постоянной c_j к каждой из величин $V_{1j}, j = 1, \dots, l$.

Г р у п п а G_2 . Преобразования

$$V^* = UB + VC, \quad U^* = U,$$

где B — любая матрица порядка $r \times l$, а C — любая невырожденная $l \times l$ -матрица.

До применения принципа инвариантности удобно редуцировать данные к достаточным статистикам.

Случайные величины U_{1i} , V_{1j} вместе с матрицами скалярных произведений $U'U$, $U'V$ и $V'V$ образуют систему достаточных статистик для неизвестного вектора средних значений и неизвестной матрицы ковариаций. В соответствии с результатом задачи 1 главы 6, группы G_1 и G_2 оставляют задачу инвариантной и после того, как она сведена к достаточным статистикам. Максимальным инвариантом относительно G_1 является совокупность величин U_{1i} и матриц $U'U$, $U'V$ и $V'V$. Мы докажем теперь, что максимальный инвариант относительно группы, которую G_2 индуцирует на этом множестве статистик, образуют U_{1i} и $U'U$. Этим будет завершено желаемое исключение величин V , а вместе с тем и величин Z .

Для доказательства высказанного утверждения необходимо показать, что для любой данной $(n-1) \times l$ -матрицы V^{**} найдутся матрицы B и C такие, что $V^* = UB + VC$ удовлетворяет соотношениям

$$U'V^* = U'V^{**} \quad \text{и} \quad V'^*V^* = V'^{**}V^{**}.$$

Геометрически эти уравнения устанавливают, что существуют векторы $(V_{2i}^*, \dots, V_{ni}^*)$, $i = 1, \dots, l$, которые лежат в пространстве E , натянутом на векторы-столбцы матриц U и V , и скалярные произведения которых между собой и на векторы-столбцы U имеют заданную величину.

Рассмотрим сначала случай $l = 1$. Если $r + 1 \geq n - 1$, то можно предположить, что V и столбцы U порождают $(n-1)$ -мерное пространство; тогда можно принять, что $V^* = V^{**}$. Если $r + 1 < n - 1$, то можно предположить, что V и столбцы U линейно независимы. Тогда существует такой поворот вокруг натянутого на столбцы U пространства, который переводит V^{**} в вектор, лежащий в E , и этот вектор имеет свойства, требуемые от V^* .

Доказательство завершается повторным применением полученного результата. Мы применяем его сначала к вектору $(V_{21}, \dots, V_{n1})$ и определяем первый столбец B и число c_{11} , добавляя к которому нули, мы построим первый столбец матрицы C . Присоединяя преобразованный вектор $(V_{21}^*, \dots, V_{n1}^*)$ к столбцам U и применяя снова указанный результат к вектору $(V_{22}, \dots, V_{n2})$, мы построим вектор $(V_{22}^*, \dots, V_{n2}^*)$, который лежит в пространстве, натянутом на $(V_{21}, \dots, V_{n1})$, $(V_{22}, \dots, V_{n2})$ и

столбцы U , и который, кроме того, имеет заданные скалярные произведения с $(V_{21}^*, \dots, V_{n1}^*)$, со столбцами U , и заданный скалярный квадрат. Этим вторым шагом мы определяем второй столбец матрицы B и два числа c_{12} , c_{22} , добавляя к которым нули, получим второй столбец C . Продолжая это построение, мы получим «треугольную» матрицу C с нулями под главной диагональю; эта матрица не вырождена. Так как можно предположить, что U , V и V^{**} имеют максимальный ранг, то из леммы 1 и уравнения $V^*V^* = V^{**}V^{**}$ следует, что ранг V^* также максимален. Этим доказательство завершается.

12. χ^2 -критерий: простая гипотеза и неограниченные альтернативы

РНМ инвариантные критерии существуют в довольно узком классе задач, важнейшие из которых, вероятно, связаны с линейными гипотезами. Однако, если число наблюдений велико, то часто можно указать критерии, приближенно обладающие этим свойством. Хотя подробное изложение теории больших выборок не входит в задачи этой книги, здесь приводятся краткие сведения о двух типах критериев подобного рода: критерии χ^2 и критерии отношения правдоподобия. В обоих случаях приближенная оптимальность вытекает из асимптотической эквивалентности рассматриваемой задачи и некоторой задачи о проверке линейной гипотезы. Эта связь будет прослежена в следующем разделе. В качестве подготовительного шага мы обсудим здесь специальный класс задач, связанных с χ^2 -критерием.

Удобно начать со следующего видоизменения модели линейной гипотезы. Пусть $Y = (Y_1, \dots, Y_q)$ имеет многомерное нормальное распределение с плотностью вероятности

$$\frac{\sqrt{|A|}}{(2\pi)^{\frac{1}{2}q}} \exp \left[-\frac{1}{2} \sum_{i=1}^q \sum_{j=1}^q a_{ij} (y_i - \eta_i) (y_j - \eta_j) \right] \quad (62)$$

и с известной матрицей ковариаций A^{-1} . Известно, что вектор средних значений $\eta = (\eta_1, \dots, \eta_q)$ лежит в данном s -мерном линейном пространстве Π_s с $s \leq q$. Проверяемая гипотеза состоит в том, что η лежит в заданном $(s-r)$ -мерном линейном подпространстве Π_r пространства Π_s ($r \leq s$). Проблема инвариантна относительно подходящей группы G линейных преобразований и существует РНМ инвариантный относительно G критерий с критической областью

$$\begin{aligned} \sum \sum a_{ij} (y_i - \hat{\eta}_i) (y_j - \hat{\eta}_j) - \sum \sum a_{ij} (y_i - \hat{\eta}_i) (y_j - \hat{\eta}_j) = \\ = \sum \sum a_{ij} (\hat{\eta}_i - \hat{\eta}_i) (\hat{\eta}_j - \hat{\eta}_j) \leq C. \end{aligned} \quad (63)$$

Здесь $\hat{\eta}$ — точка Π_Ω , ближайшая к выборочной точке y в смысле метрики, порождаемой квадратичной формой $\sum \sum a_{ij} x_i x_j$, т. е. точка, которая минимизирует выражение $\sum \sum a_{ij} (y_i - \eta_i)(y_j - \eta_j)$ при $\eta \in \Pi_\Omega$. Аналогично, $\hat{\eta}$ минимизирует это выражение на Π_ω .

Когда гипотеза верна, левая часть (63) имеет χ^2 -распределение с r степенями свободы, так что C определяется из соотношения

$$\int_C^\infty \chi_r^2(z) dz = \alpha. \quad (64)$$

Если η не принадлежит Π_ω , то вероятность отклонения гипотезы равна *)

$$\int_C^\infty p_\lambda(z) dz, \quad (65)$$

где $p_\lambda(z)$ — нецентральная плотность χ^2 (см. формулу (86) в задаче 2) с r степенями свободы и с параметром нецентральности λ^2 , получаемым заменой величин y_i , $\hat{\eta}_i$, $\hat{\eta}_i$ в (63) их математическими ожиданиями (рассматривая (63) как функцию y , можно сказать, что при этом всюду y заменяется на η). Указанное выражение для мощности остается верным и тогда, когда принятая модель неудовлетворительна, в том смысле, что $E(Y) = \eta$ не лежит в Π_Ω . В случае $\eta \in \Pi_\Omega$ второй член в выражении для λ^2 равен нулю. Доказательство высказанных утверждений получается сведением задачи к проверке линейной гипотезы с помощью надлежащего линейного преобразования (см. задачу 33).

Вернемся к χ^2 -критерию и рассмотрим его в применении к полиномиальной схеме. Возьмем n полиномиальных испытаний с m возможными исходами. Обозначим $p = (p_1, \dots, p_m)$ вероятности этих исходов, и пусть X_i — число испытаний, закончившихся i -м исходом. Тогда распределение $X = (X_1, \dots, X_m)$ равно

$$P(x_1, \dots, x_m) = \frac{n!}{x_1! \dots x_m!} p_1^{x_1} \dots p_m^{x_m} \left(\sum x_i = n, \sum p_i = 1 \right). \quad (66)$$

Простейшая из связанных с χ^2 -критерием задач состоит в проверке гипотезы $H: p = \pi$, где $\pi = (\pi_1, \dots, \pi_m)$ дано, при неограниченной альтернативе $p \neq \pi$. При $n \rightarrow \infty$ мощность критерия, о котором будет идти речь, стремится к единице при любой

*) Таблицы даны у следующих авторов: Patnaik, The noncentral χ^2 and F -distributions and their applications, Biometrika, vol. 36 (1949), 202—232; Fix, Tables of noncentral χ^2 , Univ. Calif. Publ. Statistics, vol. 1 (1949), 15—19; Fix, Hodges and Lehmann, The restricted χ^2 test, in Studies in Probability and Statistics Dedicated to Harald Cramér, Almqvist and Wiksell, Stockholm, 1959.

фиксированной альтернативе *). Для изучения функции мощности такого рода критерия целесообразно рассмотреть последовательность альтернатив $p^{(n)}$, стремящихся к π при $n \rightarrow \infty$. Если скорость сходимости по порядку больше $1/\sqrt{n}$, то даже для наиболее мощного критерия мощность будет стремиться к уровню значимости α . Последовательностями, отражающими наиболее интересные стороны поведения функции мощности и дающими, по всей видимости, полезную аппроксимацию действительной мощности при больших, но конечных n , являются те, для которых $\sqrt{n}(p^{(n)} - \pi)$ имеет ненулевой предел, так что

$$p_i^{(n)} = \pi_i + \frac{\Delta_i}{\sqrt{n}} + R_n, \quad (67)$$

где $\sqrt{n}R_n$ стремится к нулю при $n \rightarrow \infty$.

Пусть

$$Y_i = \frac{(X_i - n\pi_i)}{\sqrt{n}}. \quad (68)$$

Тогда $\sum_{i=1}^m Y_i = 0$ и среднее значение Y_i равно нулю при гипотезе H и стремится к Δ_i при альтернативах (67). Матрица ковариаций для Y имеет элементы

$$\sigma_{ij} = -\pi_i \pi_j \quad \text{при } i \neq j; \quad \sigma_{ii} = \pi_i (1 - \pi_i), \quad (69)$$

если гипотеза H верна. При альтернативах (67) элементы матрицы ковариаций стремятся к этим значениям. При $n \rightarrow \infty$ распределение $Y = (Y_1, \dots, Y_{m-1})$ стремится к многомерному нормальному распределению со средними $E(Y_i) = 0$ при H и $E(Y_i) = \Delta_i$ при альтернативах (67) и с матрицей ковариаций, определяемой по (69) в обоих случаях **). Плотность предельного распределения равна

$$c \exp \left[-\frac{1}{2} \left(\sum_{i=1}^{m-1} \frac{(y_i - \Delta_i)^2}{\pi_i} + \frac{\left(\sum_{j=1}^{m-1} (y_j - \Delta_j) \right)^2}{\pi_m} \right) \right] \quad (70)$$

и проверяемая гипотеза принимает вид $H: \Delta_1 = \dots = \Delta_{m-1} = 0$.

В соответствии с (63) РНМ инвариантный критерий в этой асимптотической модели отвергает гипотезу, когда

$$\sum_{i=1}^{m-1} \frac{y_i^2}{\pi_i} + \frac{1}{\pi_m} \left(\sum_{j=1}^{m-1} y_j \right)^2 > C,$$

*) Последовательность критериев с таким свойством называется *состоятельной*.

**) Доказательство в предположении, что гипотеза H верна, дано, например, в книге: Крамер Г., Математические методы статистики, 2-е изд., М., Мир, 1975 (§ 30.1). С очевидным применением оно переносится на случай, когда H неверна.

т. е. когда

$$n \sum_{i=1}^m \frac{(v_i - \pi_i)^2}{\pi_i} > C, \quad (71)$$

где $v_i = X_i/n$ и где C определяется из (64) с $r = m - 1$. Предельная мощность рассматриваемого критерия при альтернативах (67) вычисляется по формуле (65) с $\lambda^2 = \sum_{i=1}^m \Delta_i^2 / \pi_i$. Это дает приближенное представление о мощности при фиксированном n и фиксированной альтернативе p , если отождествить p с $p^{(n)}$ для этого n . Из (67) мы выводим, что приближенно $\Delta_i = \sqrt{n} (p_i - \pi_i)$, так что параметр нецентральности принимает вид

$$\lambda^2 = n \sum_{i=1}^m \frac{(p_i - \pi_i)^2}{\pi_i}. \quad (72)$$

Пример 12. Предположим, что проверяемая гипотеза состоит в том, что некоторые события (рождения, смерти, происшествия) происходят равномерно в заданном интервале времени (например, день или год). Разобьем этот интервал на m равных частей и обозначим p_i вероятность наступления указанного события в i -м подынтервале. Тогда гипотеза может быть представлена в виде $H: p_i = 1/m, i = 1, \dots, m$, а статистика критерия — в виде

$$mn \sum_{i=1}^m \left(v_i - \frac{1}{m} \right)^2,$$

где v_i — относительная частота событий в i -м подынтервале. Приближенное значение мощности критерия находится из (65) с $r = m - 1$ и $\lambda^2 =$
 $= mn \sum_{i=1}^m [p_i - (1/m)]^2.$

13. χ^2 -критерии и критерии отношения правдоподобия

Сила и слабость χ^2 -критерия предыдущего раздела связаны с тем, что его асимптотическая мощность зависит только от взвешенной суммы квадратов отклонений (72), а не от знаков этих отклонений или их распределения между различными значениями i . Этим создается преимущество, если ничего не известно об альтернативах. Тогда критерий одинаково надежен при всех альтернативах, равно удаленных от гипотезы $H: p = \pi$ в смысле метрики (72). Однако часто бывают известны ожидаемые типы отклонений. В этих случаях критерий можно видоизменить так, чтобы его асимптотическая мощность при интересующих нас альтернативах возросла.

Для вывода видоизмененного критерия допустим, что задан ограниченный класс альтернатив к H

$$K: p \in \mathcal{P}, \quad p \neq \pi,$$

где поверхность $\mathcal{P}$ имеет параметрическое представление

$$p_i = f_i(\theta_1, \dots, \theta_s), \quad i = 1, \dots, m.$$

Пусть

$$\pi_i = f_i(\theta_1^0, \dots, \theta_s^0).$$

Будем считать, что θ_j действительны, частные производные $\partial f_i / \partial \theta_j$ существуют и непрерывны в точке θ^0 и что якобиан матрицы $(\partial f_i / \partial \theta_j)$ имеет в точке θ^0 ранг s . Если $\theta^{(n)}$ — любая последовательность, для которой

$$\sqrt{n}(\theta_j^{(n)} - \theta_j^{(0)}) \rightarrow \delta_j, \quad (73)$$

то предельное распределение случайных величин $(Y_1, \dots, Y_{m-1})$, введенных в предыдущем параграфе, является нормальным со средними значениями

$$E(Y_i) = \Delta_i = \sum_{j=1}^s \delta_j \left. \frac{\partial f_i}{\partial \theta_j} \right|_{\theta^0} \quad (74)$$

и матрицей ковариаций (69). Это устанавливается разложением f_i в ряд в окрестности точки θ^0 и применением (70). Проблема проверки гипотезы H при всех последовательностях альтернатив из K , удовлетворяющих (73), оказывается асимптотически эквивалентной задаче проверки гипотезы

$$\Delta_1 = \dots = \Delta_{m-1} = 0$$

в семействе (70) при альтернативах $\bar{K}: (\Delta_1, \dots, \Delta_{m-1}) \in \Pi_\Omega$, где Π_Ω — линейное пространство, образованное совокупностью всех точек с координатами

$$\Delta_i = \sum_{j=1}^s \delta_j \left. \frac{\partial f_i}{\partial \theta_j} \right|_{\theta^0}. \quad (75)$$

Заметим для последующих целей, что при любом фиксированном n совокупность точек

$$p_i = \pi_i + \frac{\Delta_i}{\sqrt{n}}, \quad i = 1, \dots, m,$$

с Δ_i , удовлетворяющими (75), образует плоскость, касательную к $\mathcal{P}$ в точке π . Эта плоскость будет обозначаться $\bar{\mathcal{P}}$.

Пусть $(\hat{\Delta}_1, \dots, \hat{\Delta}_m)$ — значения, минимизирующие $\sum_{i=1}^m (y_i - \Delta_i)^2 / \pi_i$ при условиях $(\Delta_1, \dots, \Delta_{m-1}) \in \Pi_\Omega$ и $\Delta_m = -(\Delta_1 + \dots$

... + Δ_{m-1}). Тогда, в силу (63), асимптотически РНМ инвариантный критерий отклоняет H (т. е. действует в пользу K) при

$$\frac{\sum_{i=1}^m y_i^2}{\pi_i} - \frac{\sum_{i=1}^m (y_i - \hat{\Delta}_i)^2}{\pi_i} = \frac{\sum_{i=1}^m \hat{\Delta}_i^2}{\pi_i} > C,$$

т. е. при

$$\frac{n \sum_{i=1}^m (v_i - \pi_i)^2}{\pi_i} - \frac{n \sum_{i=1}^m (v_i - \hat{p}_i)^2}{\pi_i} = \frac{n \sum_{i=1}^m (\hat{p}_i - \pi_i)^2}{\pi_i} > C, \quad (76)$$

где $\hat{p}_i$ минимизируют $\sum (v_i - p_i)^2 / \pi_i$ при условии $p \in \bar{\mathcal{P}}$. Константа C определяется из (64) с $r = s$. Критерий, асимптотически эквивалентный этому (вычисление его, как правило, более трудно), получается, если $\hat{p}_i$ определять из условия минимальности при $p \in \mathcal{P}$ (а не при $p \in \bar{\mathcal{P}}$). Асимптотическое выражение для мощности критерия получается из (65). При этом λ^2 находится с помощью замены v_i на p_i (отметим, что $\hat{p}_i$ рассматриваются как функции от v_i).

Пример 13. Допустим, что в примере 12, где проверяется гипотеза равномерности, альтернативой выдвигается циклическое изменение, которое, по крайней мере приближенно, представляется «синусоидальной волной»:

$$p_i = \frac{1}{m} + \rho \int_{(i-1)\frac{2\pi}{m}}^{i\frac{2\pi}{m}} \sin(u - \theta) du, \quad i = 1, \dots, m.$$

Здесь ρ — амплитуды и θ — фаза циклического возмущения. Полагая $\xi = \rho \cos \theta$, $\eta = \rho \sin \theta$, мы получаем

$$p_i = \frac{1}{m} (1 + a_{-i} \xi + b_i \eta),$$

где

$$a_i = 2m \sin \frac{\pi}{m} \sin (2i - 1) \frac{\pi}{m}, \quad b_i = -2m \sin \frac{\pi}{m} \cos (2i - 1) \frac{\pi}{m}.$$

Уравнение для p_i определяет в этом случае плоскость $\mathcal{P}$, которая, следовательно, совпадает с $\bar{\mathcal{P}}$.

Величины $\hat{\xi}$, $\hat{\eta}$, минимизирующие $\sum (v_i - p_i)^2 / \pi_i$ при условии $p \in \mathcal{P}$, равны

$$\hat{\xi} = \frac{\sum a_i v_i}{\sum a_i^2 \pi_i}, \quad \hat{\eta} = \frac{\sum b_i v_i}{\sum b_i^2 \pi_i}$$

с $\pi_i = 1/m$. Пусть $m > 2$. Используя тот факт, что $\sum a_i = \sum b_i = \sum a_i b_i = 0$ и что

$$\sum_{i=1}^m \sin^2 (2i - 1) \frac{\pi}{m} = \sum_{i=1}^m \cos^2 (2i - 1) \frac{\pi}{m} = \frac{m}{2},$$

мы, после некоторых упрощений, можем придать критерию форму

$$2n \left[\sum_{i=1}^m v_i \sin(2i-1) \frac{\pi}{m} \right]^2 + 2n \left[\sum_{i=1}^m v_i \cos(2i-1) \frac{\pi}{m} \right]^2 > C,$$

где число степеней свободы левой части равно $s = 2$. Параметр нецентральности, через который приближенно выражается мощность, равен

$$\lambda^2 = n \left(\xi m \sin \frac{\pi}{m} \right)^2 + n \left(\eta m \sin \frac{\pi}{m} \right)^2 = n \rho^2 m^2 \sin^2 \frac{\pi}{m}.$$

Обсуждавшиеся до сих пор варианты критерия χ^2 предназначались для простых гипотез. Рассмотрим теперь более общую проблему проверки гипотезы $H: p \in \mathcal{T}$ при альтернативах $K: p \in \mathcal{P}, p \notin \mathcal{T}$, где $\mathcal{T} \subset \mathcal{P}$ и где $\mathcal{P}$ и $\mathcal{T}$ имеют параметрические представления

$$\mathcal{P}: p_i = f_i(\theta_1, \dots, \theta_s); \quad \mathcal{T}: p_i = f_i(\theta_1^0, \dots, \theta_r^0, \theta_{r+1}, \dots, \theta_s).$$

Базисом для анализа проблемы при больших объемах выборки служит тот факт, что при больших n можно указать сферу радиуса $\rho/\sqrt{n}$, которая при достаточно больших ρ покрывает истинное значение p с вероятностью, сколь угодно близкой к единице. Поэтому можно ограничиться последовательностями точек $p^{(n)} \in \mathcal{P}$, которые стремятся к некоторой фиксированной точке $\pi \in \mathcal{T}$ со скоростью $1/\sqrt{n}$. Более точно, пусть $\pi_i = f_i(\theta_1^0, \dots, \theta_s^0)$ и пусть $\theta^{(n)}$ — последовательность, удовлетворяющая (73). Тогда случайные величины $(Y_1, \dots, Y_{m-1})$ имеют нормальное предельное распределение с матрицей ковариаций (69) и с вектором средних (74). Пусть Π_Ω имеет тот же смысл, что и раньше, а Π_ω пусть обозначает линейное пространство

$$\Pi_\omega: \Delta_i = \sum_{j=r+1}^s \delta_j \frac{\partial p_i}{\partial \theta_j} \Big|_{\theta^0}.$$

Рассмотрим задачу проверки гипотезы о том, что $p^{(n)}$ — последовательность из H , для которой $\theta^{(n)}$ удовлетворяют (73), причем альтернативами являются все последовательности из K , удовлетворяющие этому же условию. Эта задача асимптотически эквивалентна задаче, обсуждавшейся в начале раздела 12, а именно задаче проверки гипотезы $(\Delta_1, \dots, \Delta_{m-1}) \in \Pi_\omega$ в семействе (70), если дано, что $(\Delta_1, \dots, \Delta_{m-1}) \in \Pi_\Omega$. В силу (63) соответствующая критическая область имеет вид

$$\frac{\sum \frac{(y_i - \hat{\Delta}_i)^2}{\pi_i}}{\sum \frac{(y_i - \hat{\Delta}_i)^2}{\pi_i}} > C,$$

где $\hat{\Delta}_i$ и $\hat{\Delta}_i$ минимизируют $\sum (y_i - \Delta_i)^2 / \pi_i$ при условиях, что $\Delta_m = -(\Delta_1 + \dots + \Delta_{m-1})$ и что $(\Delta_1, \dots, \Delta_{m-1})$ лежит в Π_Ω или в Π_ω соответственно. В терминах первоначальных величин критическая область имеет вид

$$\frac{n \sum (v_i - \hat{p}_i)^2}{\pi_i} - \frac{n \sum (v_i - \hat{p}_i)^2}{\pi_i} > C. \quad (77)$$

Здесь $\hat{p}_i$ и $\hat{p}_i$ минимизируют

$$\frac{\sum (v_i - p_i)^2}{\pi_i} \quad (78)$$

при условии, что p лежит в касательных плоскостях к поверхностям $\mathcal{P}$ и $\mathcal{T}$ соответственно, проведенных через π . Константа C определяется из (64).

Это решение задачи зависит от точки π , которая нам неизвестна. Критерий, асимптотически эквивалентный (77) и не зависящий от π , получается, если заменить $\hat{p}_i$ и $\hat{p}_i$ значениями p_i^* и p_i^{**} , минимизирующими (78) при условии, что p лежит на $\mathcal{P}$ или на $\mathcal{T}$, а не на их касательных плоскостях, и, кроме того, заменить π_i в (77) и (78) подходящими оценками, например v_i . Это приводит к критической области

$$\frac{n \sum (v_i - p_i^{**})^2}{v_i} - \frac{n \sum (v_i - p_i^*)^2}{v_i} = \frac{n \sum (p_i^* - p_i^{**})^2}{v_i} > C, \quad (79)$$

где p_i^{**} и p_i^* минимизируют

$$\sum (v_i - p_i)^2 / v_i \quad (80)$$

при условиях $p \in \mathcal{T}$ и $p \in \mathcal{P}$ соответственно и где C , как и раньше, определяется формулой (64). Приближенное выражение для мощности критерия при фиксированных n и альтернативе p дается формулой (65) с λ^2 , определяемым по (79) с заменой v_i на p_i (и где p_i^* и p_i^{**} рассматриваются как функции от v_i *).

Для случая больших выборок более общий подход, не связанный в отличие от метода χ^2 с полиномиальным распределением, опирается на метод наибольшего правдоподобия. Укажем кратко на основные положения этой теории. Перечисляя основные факты, мы опустим сложно формулируемые условия регулярности, при которых они справедливы **).

*) Доказательство вышеприведенных утверждений и обсуждение критериев, которые асимптотически эквивалентны (76) и которые иногда легче вычисляются, см. в цитированной работе Fix, Hodges and Lehmann.

**) Подробное изложение см. в работе: Wald, Tests of statistical hypotheses concerning several parameters when the number of observations is large, Trans. Ann. Math. Soc. 54 (1943), 426—483.

Пусть $p_\theta(x)$, $\theta = (\theta_1, \dots, \theta_r)$ — некоторое семейство одномерных плотностей вероятности. Рассмотрим на основе большой выборки $X_1, \dots, X_n$ задачу проверки простой гипотезы $H: \theta_i = \theta_i^0$, $i = 1, \dots, r$. Пусть $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_r)$ — оценка максимального правдоподобия для θ , т. е. значение параметрического вектора, максимизирующее $p_\theta(x_1) \dots p_\theta(x_n)$. Тогда асимптотически, при $n \rightarrow \infty$, можно ограничиться оценками $\hat{\theta}_i$, так как они являются «асимптотически достаточными»^{*}). Мощность критериев, о которых пойдет речь, стремится к единице при любой фиксированной альтернативе. Здесь, как и в случае χ^2 -критерия, интерес представляют альтернативы $\theta_i^{(n)}$, для которых

$$\sqrt{n} (\theta_i^{(n)} - \theta_i^0) \rightarrow \Delta_i. \quad (81)$$

Если $Y_i = \sqrt{n} (\hat{\theta}_i - \theta_i^0)$, то предельное распределение величин $Y_1, \dots, Y_r$ является многомерным нормальным распределением (62) с

$$a_{ij} = a_{ij}(\theta^0) = -E \left(\frac{\partial^2 \log p_\theta(X)}{\partial \theta_i \partial \theta_j} \right) \Big|_{\theta = \theta^0} \quad (82)$$

и с $\eta_i = 0$ при гипотезе H и $\eta_i = \Delta_i$ при альтернативах, удовлетворяющих (81).

В силу (63) РНМ инвариантный критерий в асимптотической модели отклоняет гипотезу, когда

$$- \sum_{i=1}^r \sum_{j=1}^r a_{ij} n (\hat{\theta}_i - \theta_i^0) (\hat{\theta}_j - \theta_j^0) > C. \quad (83)$$

При гипотезе H левая часть имеет в пределе χ^2 -распределение с r степенями свободы. При альтернативах (81) предельным распределением является нецентральное χ^2 -распределение с параметром нецентральности

$$\lambda^2 = - \sum_{i=1}^r \sum_{j=1}^r a_{ij} n (\theta_i^{(n)} - \theta_i^0) (\theta_j^{(n)} - \theta_j^0). \quad (84)$$

Приближенное выражение для мощности при какой-либо специальной альтернативе θ дается, следовательно, формулой (65), где λ^2 находится по (84), с заменой $\theta^{(n)}$ на θ .

^{*} Это было установлено Вальдом (цитированная работа); об определении асимптотической достаточности и дальнейших результатах, касающихся этого понятия, см. Le Cam, On the asymptotic theory of estimation and testing hypotheses, Proc. Third Berkeley Symposium on Mathematical and Probability, Univ. Calif. Press, 1956.

Критерий (83) асимптотически эквивалентен критерию отношения правдоподобия, который отклоняет гипотезу при

$$\Lambda_n = \frac{p_\theta(x_1) \dots p_\theta(x_n)}{p_{\theta^0}(x_1) \dots p_{\theta^0}(x_n)} > k. \quad (85)$$

Эту эквивалентность можно установить, разлагая $\sum_{v=1}^n \log p_{\theta^0}(x_v)$

в ряд около $\sum_{v=1}^n \log p_\theta(x_v)$ и используя тот факт, что при $\theta = \hat{\theta}$ частные производные $\partial \sum \log p_\theta(x_v) / \partial \theta_i$ равны нулю. Применение закона больших чисел показывает, что величина $-2 \log \Lambda_n$ отличается от левой части (83) членом, стремящимся к нулю по вероятности при $n \rightarrow \infty$. Отсюда, в частности, следует, что обе статистики имеют одно и то же предельное распределение.

Распространение этого метода на сложные гипотезы аналогично распространению на этот же случай χ^2 -критерия. Пусть $\theta = (\theta_1, \dots, \theta_s)$ и $H: \theta_i = \theta_i^0$ при $i = 1, \dots, r$ ($r < s$). Если ограничиться последовательностями $\theta^{(n)}$, удовлетворяющими (81) при $i = 1, \dots, s$, и произвольными $\theta_{r+1}^0, \dots, \theta_s^0$, то проблема асимптотически превращается в задачу проверки гипотезы $\eta_1 = \dots = \eta_r = 0$ при неограниченной альтернативе $(\eta_1, \dots, \eta_s)$, причем распределения определяются формулой (62) с $a_{ij} = a_{ij}(\theta^0)$ из (82). Тогда $\hat{\eta}_i = Y_i$ при всех i , в то время как $\hat{\eta}_i = 0$ для $i = 1, \dots, r$ и $\hat{\eta}_i = Y_i$ для $i = r+1, \dots, s$. Поэтому РНМ инвариантный критерий имеет вид (83). Коэффициенты $a_{ij} = a_{ij}(\theta^0)$ зависят от $\theta_{r+1}^0, \dots, \theta_s^0$, но, заменяя $a_{ij}(\theta^0)$ на $a_{ij}(\hat{\theta})$, приходим к асимптотически эквивалентному критерию. И снова статистика асимптотически эквивалентна взятому со знаком минус удвоенному логарифму отношения правдоподобия, т. е. критерий асимптотически эквивалентен критерию отношения правдоподобия *).

14. Задачи

К разделу 1

1. Математическое ожидание сумм квадратов. Средние значения числителя и знаменателя статистики W^* , определенной формулой (7), равны

$$E\left(\sum_{i=1}^r \frac{Y_i^2}{r}\right) = \sigma^2 + \frac{1}{r} \sum_{i=1}^r \eta_i^2 \quad \text{и} \quad E\left(\sum_{i=s+1}^n \frac{Y_i^2}{n-s}\right) = \sigma^2.$$

*) Асимптотическая теория критериев отношения правдоподобия была распространена Черновым на более общие типы проблем, включая, в частности, случай ограниченных классов альтернатив. См. Chernoff, On the distribution of the likelihood ratio, Ann. Math. Stat., vol. 25 (1954), 573—578.

2. *Нецентральное χ^2 -распределение.* (I) Если X распределено по закону $N(\psi, 1)$, то плотность вероятности величины $V = X^2$ представляется как

$p_\psi^V(v) = \sum_{k=0}^{\infty} P_k(\psi) f_{2k+1}(v)$, где $P_k(\psi) = (\psi^2/2)^k e^{-\frac{1}{2}\psi^2}/k!$ и где f_{2k+1} является плотностью вероятности χ^2 -величины с $2k+1$ степенями свободы.

(II) Пусть $Y_1, \dots, Y_r$ независимы и нормально распределены с единичной дисперсией и средними $\eta_1, \dots, \eta_r$. Тогда $U = \sum_{i=1}^r Y_i^2$ имеет нецентральное χ^2 -распределение с r степенями свободы и параметром нецентральности $\psi^2 = \sum_{i=1}^r \eta_i^2$; плотность вероятности величины U дается формулой

$$P_\psi^U(u) = \sum_{k=0}^{\infty} P_k(\psi) f_{r+2k}(u). \quad (86)$$

Здесь $P_k(\psi)$ и $f_{r+2k}(u)$ имеют тот же смысл, что и в (I), так что распределение является «смесью» χ^2 -распределений с пуассоновскими весами.

[(I) Это следует из равенства

$$p_\psi^V(v) = \frac{e^{-\frac{1}{2}(\psi^2+v)} (e^{\psi\sqrt{v}} + e^{-\psi\sqrt{v}})}{2\sqrt{2\pi v}},$$

если выражение в скобках разложить в степенной ряд и использовать тот факт, что $\Gamma(2k) = 2^{2k-1} \Gamma(k) \Gamma(k + \frac{1}{2}) / \sqrt{\pi}$.

(II) Перейдем с помощью ортогонального преобразования к величинам $Z_1, \dots, Z_r$ таким, что $Z_i = \sum \eta_i Y_i / \psi$. Тогда величины Z независимы и нормальны с единичной дисперсией и средними $E(Z_1) = \psi$ и $E(Z_i) = 0$ при $i > 1$].

3. *Нецентральные F- и бета-распределения.* Пусть $Y_1, \dots, Y_r; Y_{s+1}, \dots, Y_n$ независимы и нормально распределены с общей дисперсией σ^2 и средними $E(Y_i) = \eta_i$ ($i = 1, \dots, r$); $E(Y_i) = 0$ ($i = s+1, \dots, n$).

(I) Плотность вероятности статистики $W = \sum_{i=1}^r Y_i^2 / \sum_{i=s+1}^n Y_i^2$ дается формулой (6). Величина $(n-s)W/r$, пропорциональная W , имеет *нецентральное F-распределение*.

(II) Статистика $B = \sum_{i=1}^r Y_i^2 / \left(\sum_{i=1}^r Y_i^2 + \sum_{i=s+1}^n Y_i^2 \right)$ имеет *нецентральное бета-распределение* с плотностью

$$\sum_{k=0}^{\infty} P_k(\psi) g_{\frac{1}{2}r+k, \frac{1}{2}(n-s)}(b), \quad (87)$$

где

$$g_{p,q}(b) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} b^{p-1} (1-b)^{q-1}, \quad 0 \leq b \leq 1, \quad (88)$$

является плотностью (центрального) бета-распределения.

4. (I) Если $p_\psi(x)$ — плотность нецентрального χ^2 или нецентрального F-распределения, то отношение $p_{\psi_1}(x)/p_{\psi_0}(x)$ возрастает по x при $\psi_0 < \psi_1$.

(II) В предположениях раздела 1, гипотеза $H': \psi^2 \leq \psi_0^2$ (при данном $\psi_0 > 0$) остается инвариантной относительно преобразований G_i ($i = 1, 2, 3$), которые были использованы для редукции $H: \psi = 0$, и существует РНМ инвариантный критерий с критической областью $W > C'$. Константа C' определяется равенством $P_{\psi_0} \{W > C'\} = \alpha$, где плотность W дается формулой (6).

[(I) Пусть $f(z) = \sum_{k=0}^{\infty} b_k z^k / \sum_{k=0}^{\infty} a_k z^k$, где постоянные $a_k, b_k > 0$ и ряды $\sum a_k z^k$ и $\sum b_k z^k$ сходятся при всех $z > 0$. Допустим, что $b_k/a_k < b_{k+1}/a_{k+1}$ для всех k . Тогда производная

$$f'(z) = \frac{\sum_{k < n} \sum (n-k) (a_k b_n - a_n b_k) z^{k+n-1}}{\left(\sum_{k=0}^{\infty} a_k z^k \right)^2}$$

положительна, так как $(n-k)(a_k b_n - a_n b_k)$ при $k < n$ и, следовательно, f возрастает.]

5. *Наилучшая средняя мощность.* (I) Рассмотрим общую линейную гипотезу в канонической форме, данной формулами (2) и (3) раздела 1, и пусть для любых $\eta_{r+1}, \dots, \eta_s, \sigma$ и ρ $S = S(\eta_{r+1}, \dots, \eta_s, \sigma; \rho)$ обозначает сферу $\left\{ (\eta_1, \dots, \eta_r): \sum_{i=1}^r \eta_i^2 / \sigma^2 = \rho^2 \right\}$. Если $\beta_\varphi(\eta_1, \dots, \eta_s, \sigma)$ обозначает мощность критерия φ для H , то критерий (9) максимизирует среднюю мощность

$$\frac{\int_S \beta_\varphi(\eta_1, \dots, \eta_s, \sigma) dA}{\int_S dA}$$

в классе всех несмещенных (или подобных) критериев для каждого $\eta_{r+1}, \dots, \eta_s, \sigma$ и ρ . Здесь dA обозначает дифференциал площади поверхности сферы.

(II) Результат (I) дает альтернативное доказательство того факта, что критерий (9) является РНМ среди всех критериев, функция мощности которых зависит только от $\sum_{i=1}^r \eta_i^2 / \sigma^2$.

[(I) Если $U = \sum_{i=1}^r Y_i^2$, $V = \sum_{i=s+1}^n Y_i^2$, то из несмещенности (или подобия) следует, что условная вероятность отклонения данных $Y_{r+1}, \dots, Y_s$ и $U + V$ равна α почти всюду. Следовательно, для любых данных $\eta_{r+1}, \dots, \eta_s, \sigma$ и ρ средняя мощность максимизируется при отклонении гипотезы в случае, когда отношение средней плотности к плотности при гипотезе H больше, чем соответствующая постоянная $C(y_{r+1}, \dots, y_s, u + v)$, и, следовательно, когда

$$g(y_1, \dots, y_r; \eta_1, \dots, \eta_r) = \int_S \exp \left(\sum_{i=1}^r \frac{\eta_i y_i}{\sigma^2} \right) dA > C(y_{r+1}, \dots, y_s, u + v).$$

Как будет показано ниже, функция g зависит от $y_1, \dots, y_r$ только через u и,

как функция u , является возрастающей. Так как при проверяемой гипотезе $U/(U+V)$ не зависит от $Y_{r+1}, \dots, Y_s$ и $U+V$, то критерий задается формулой (9).

Показатель в интеграле, определяющем g , может быть записан в виде

$\sum_{i=1}^r \eta_i y_i / \sigma^2 = \rho \sqrt{u} \cos \beta / \sigma$, где β — угол ($0 \leq \beta \leq \pi$) между $(\eta_1, \dots, \eta_r)$ и $(y_1, \dots, y_r)$. В силу симметрии сферы он не меняется, если β заменить углом γ между $(\eta_1, \dots, \eta_r)$ и произвольным вектором. Этим доказывается, что g зависит от u только через u . Обозначим величину g при фиксированных $\eta_1, \dots, \eta_r$ через $h(u)$. Пусть S' — подмножество S , для которого $0 \leq \gamma \leq \pi/2$. Тогда

$$h(u) = \int_{S'} \left[\exp \left(\rho \sqrt{u} \frac{\cos \gamma}{\sigma} \right) + \exp \left(-\rho \sqrt{u} \frac{\cos \gamma}{\sigma} \right) \right] dA,$$

что доказывает желаемый результат.]

К разделу 2

6. В предположениях раздела 1 допустим, что средние значения ξ_i имеют вид

$$\xi_i = \sum_{j=1}^s a_{ij} \beta_j,$$

где константы a_{ij} известны, матрица $A = (a_{ij})$ не вырождена и β_j — неизвестные параметры. Пусть $\theta = \sum_{j=1}^s e_j \beta_j$ — заданная линейная комбинация величин β_j .

(I) Если $\hat{\beta}_j$ обозначает величину β_j , минимизирующую $\sum (X_i - \xi_i)^2$, и если $\hat{\theta} = \sum_{j=1}^s e_j \hat{\beta}_j = \sum_{j=1}^n d_j X_j$, то критическая область для гипотезы $H: \theta = \theta_0$ имеет вид

$$\frac{|\hat{\theta} - \theta_0| / \sqrt{\sum d_i^2}}{\sqrt{\sum (X_i - \hat{\xi}_i)^2 / (n-s)}} > C_0, \quad (89)$$

где левая часть при гипотезе H имеет такое же распределение, как абсолютная величина t -статистики Стьюдента с $n-s$ степенями свободы.

(II) Соответствующие доверительные интервалы для θ имеют вид

$$\hat{\theta} - k \sqrt{\frac{\sum (X_i - \hat{\xi}_i)^2}{(n-s)}} \leq \theta \leq \hat{\theta} + k \sqrt{\frac{\sum (X_i - \hat{\xi}_i)^2}{(n-s)}}, \quad (90)$$

где $k = C_0 \sqrt{\sum d_i^2}$. Эти интервалы являются равномерно наиболее точными среди инвариантных относительно подходящей группы преобразований.

[(I) Рассмотрим сначала гипотезу $\theta = 0$. Допустим, не ограничивая общности, что $\theta = \beta_1$, так как общий случай может быть сведен к этому линейным преобразованием в пространстве β . Если $a_1, \dots, a_s$ — столбцы матрицы A , которые по предположению порождают Π_Ω , то $\underline{\xi} = \beta_1 \underline{a}_1 + \dots + \beta_s \underline{a}_s$, и так как $\underline{\xi}$ лежит в Π_Ω , то $\underline{\xi} = \hat{\beta}_1 \underline{a}_1 + \dots + \hat{\beta}_s \underline{a}_s$. Пространство Π_ω , определяемое гипотезой $\beta_1 = 0$, можно рассматривать как натянутое или на векторы $\underline{a}_2, \dots, \underline{a}_s$, или на векторы-строки $\underline{c}_2, \dots, \underline{c}_s$ матрицы C из (1). В то же вре-

мая $\underline{c}_1$ ортогонально Π_ω . В силу 1 вектор $\underline{X}$ определяется равенством $\underline{X} = \sum_{i=1}^n Y_i \underline{c}_i$, а его проекция $\underline{\hat{x}}$ на Π_Ω равна $\underline{\hat{x}} = \sum_{i=1}^s Y_i \underline{c}_i$. Приравнивая друг другу два выражения для $\underline{\hat{x}}$ и беря скалярное произведение обеих частей полученного уравнения на $\underline{c}_1$, получим $Y_1 = \hat{\beta}_1 \sum_{i=1}^n a_{i1} c_{i1}$, так как векторы $\underline{c}$ образуют ортогональную систему единичных векторов. Мы показали, что Y_1 пропорционально $\hat{\beta}_1$, и так как дисперсия Y_1 равна дисперсии величин X , то $|Y_1| = |\hat{\beta}_1| / \sqrt{\sum d_i^2}$. Способ проверки гипотезы $\beta_1 = 0$ вытекает теперь из (12) и (13). Критерий для гипотезы $\beta_1 = \beta_1^0$ получается после преобразования $X_i^* = X_i - a'_{i1} \beta_1^0$.

(II) Инвариантность интервалов (90) можно доказывать, не ограничивая общности, в предположении, что параметр θ равен β_1 . Тогда в канонической форме раздела 1 $E(Y_1) = \eta_1 = \lambda \beta_1$ с $|\lambda| = \sqrt{\sum d_i^2}$, в то время как $\eta_2, \dots, \eta_s$ не содержит β_1 . Поэтому гипотеза $\beta_1 = \beta_1^0$ эквивалентна гипотезе $\eta_1 = \eta_1^0$ с $\eta_1^0 = \lambda \beta_1^0$. Последняя инвариантна а) относительно добавления произвольных постоянных к $Y_2, \dots, Y_s$; б) относительно преобразований $Y_1^* = -(Y_1 - \eta_1^0) + \eta_1^0$; в) относительно изменения масштаба $Y_i^* = c Y_i$ ($i = 2, \dots, n$), $Y_1^* - \eta_1^{0*} = c(Y_1 - \eta_1^0)$. Доверительные интервалы для $\theta = \beta_1$ являются равномерно наиболее точными среди инвариантных относительно группы, получаемой из а), б) и в) при меняющемся η_1^0 .]

7. Пусть X_{ij} ($j = 1, \dots, m_i$) и Y_{ik} ($k = 1, \dots, n_i$) независимы и распределены нормально с общей дисперсией σ^2 и средними $E(X_{ij}) = \xi_i$ и $E(Y_{ij}) = \xi_i + \Delta$. Тогда РНМ инвариантный критерий для гипотезы $H: \Delta = 0$ дается формулой (89) с $\theta = \Delta$, $\theta_0 = 0$

$$\hat{\theta} = \frac{\sum_i \frac{m_i n_i}{N_i} (Y_{i.} - X_{i.})}{\sum_i \frac{m_i n_i}{N_i}}, \quad \hat{\xi}_i = \frac{\sum_{j=1}^{m_i} X_{ij} + \sum_{k=1}^{n_i} (Y_{ik} - \hat{\theta})}{N_i},$$

где $N_i = m_i + n_i$.

8. Пусть $X_1, \dots, X_n$ независимы и распределены нормально с известной дисперсией σ_0^2 и средними $E(X_i) = \xi_i$. Рассмотрим любую линейную гипотезу с $s \leq n$ (вместо $s < n$, что требуется при неизвестной дисперсии). Проблема остается инвариантной относительно некоторой подгруппы той группы, которая использовалась в случае неизвестной дисперсии. РНМ инвариантный критерий имеет критическую область

$$\sum (X_i - \hat{\xi}_i)^2 - \sum (X_i - \hat{\xi}_i)^2 = \sum (\hat{\xi}_i - \hat{\xi}_i^*)^2 > C \sigma_0^2 \quad (91)$$

где C определяется соотношением

$$\int_C^\infty \chi_r^2(y) dy = \alpha. \quad (92)$$

К разделу 3

9. Если случайные величины X_{ij} ($j = 1, \dots, n_i$; $i = 1, \dots, s$) независимы и распределены по закону $N(\mu_i, \sigma^2)$, то

$$E[\sum n_i (X_{i.} - X_{..})^2] = (s-1)\sigma^2 + \sum n_i (\mu_i - \mu_{..})^2,$$

$$E[\sum \sum (X_{ij} - X_{i.})^2] = (n-s)\sigma^2.$$

10. Пусть $Z_1, \dots, Z_s$ независимы и распределены по законам $N(\xi_i, a_i^2)$, $i = 1, \dots, s$, где a_i — известные постоянные.

(I) При надлежаще выбранной группе линейных преобразований существует РНМ инвариантный критерий для гипотезы $H: \xi_1 = \dots = \xi_s$, с критической областью (21).

(II) Мощность этого критерия равна интегралу от C до ∞ от нецентральной χ^2 -плотности с $s-1$ степенями свободы и параметром нецентральности λ^2 , получаемым подстановкой ξ_i вместо Z_i в левую часть (21).

11. (I) Если X имеет распределение Пуассона с $E(X) = \lambda$, то при больших λ статистика $\sqrt{X}$ распределена приближенно по закону $N(\sqrt{\lambda}, 1/4)$.

(II) Если величина X имеет биномиальное распределение $b(p, n)$, то при больших n величина $\arcsin \sqrt{X/n}$ распределена приближенно по закону $N(\arcsin \sqrt{p}, 1/4n)^*$.

К разделу 5

12. Построенный в соответствии с теорией линейных гипотез критерий проверки гипотезы об отсутствии взаимодействия в двойной классификации при m наблюдениях в клетке дается формулой (30).

13. Обозначим X_λ случайную величину, имеющую нецентральное χ^2 -распределение с f степенями свободы и параметром нецентральности λ^2 . Тогда X_λ стохастически больше, чем X_λ при $\lambda < \lambda'$.

[Достаточно показать, что при Y распределенном по закону $N(0, 1)$, величина $(Y + \lambda')^2$ стохастически больше, чем $(Y + \lambda)^2$. Равносильное утверждение, что при любом $z > 0$

$$P\{|Y + \lambda'| \leq z\} \leq P\{|Y + \lambda| \leq z\},$$

непосредственно вытекает из формулы для нормальной функции плотности. Иное доказательство можно получить, комбинируя задачу 4 с леммой 2 главы 3.

14. Пусть X_{ijk} ($i = 1, \dots, a$; $j = 1, \dots, b$; $k = 1, \dots, m$) независимы и распределены нормально с общей дисперсией σ^2 и средними

$$E(X_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_k \quad (\sum \alpha_i = \sum \beta_j = \sum \gamma_k = 0).$$

Найдите критерий проверки линейной гипотезы $H: \alpha_1 = \dots = \alpha_a = 0$.

15. Предположим, что в трехфакторной ситуации предыдущей задачи $b = c = m$. Тогда гипотеза H может быть проверена на основе m^2 наблюдений, если использовать следующий способ. Для каждой пары уровней (i, j) первых двух факторов производится одно наблюдение, о котором мы будем говорить, что оно находится на пересечении i -й строки и j -го столбца. Если

* Эти преобразования подробно рассматриваются Эйзенхартом в главе 16 книги *Selected Techniques of Statistical Analysis*, New York, McGraw-Hill Book Co., 1947. Некоторые усовершенствования обсуждают Anscombe, *Transformations of Poisson, binomial and negative binomial data*, *Biometrika*, 35 (1948), 246—254 и Freeman and Tukey, *Transformations related to the angular and the square root*, *Ann. Math. Stat.* 21 (1950), 607—611.

уровни третьего фактора выбраны таким образом, что каждый из них встречается один и только один раз в каждой строке и каждом столбце, то мы называем план эксперимента *латинским квадратом*. Результаты m^2 наблюдений обозначаются $X_{ij(k)}$, где третий индекс обозначает уровень третьего фактора, когда первые два находятся на уровнях i и j соответственно. Предполагается, что $E(X_{ij(k)}) = \xi_{ij(k)} = \mu + \alpha_i + \beta_j + \gamma_k$ с $\sum \alpha_i = \sum \beta_j = \sum \gamma_k = 0$.

(I) Параметры вычисляются по величинам ξ из соотношений

$$\xi_{i.(.)} = \mu + \alpha_i, \quad \xi_{.j(.)} = \mu + \beta_j, \quad \xi_{..(k)} = \mu + \gamma_k, \quad \xi_{..(.)} = \mu$$

(суммирование по j при фиксированном i автоматически приводит к суммированию по k).

(II) Оценки параметров по методу наименьших квадратов могут быть найдены из тождества

$$\begin{aligned} \sum_i \sum_j [x_{ij(k)} - \xi_{ij(k)}]^2 &= m \sum [x_{i.(.)} - x_{..(.)} - \alpha_i]^2 + \\ &+ m \sum [x_{.j(.)} - x_{..(.)} - \beta_j]^2 + m \sum [x_{..(k)} - x_{..(.)} - \gamma_k]^2 + \\ &+ m^2 [x_{..(.)} - \mu]^2 + \sum_i \sum_k [x_{ij(k)} - x_{i.(.)} - x_{.j(k)} - x_{..(k)} + 2x_{..(.)}]^2. \end{aligned}$$

(III) При проверке гипотезы $H: \alpha_1 = \dots = \alpha_m = 0$ соответствующая критерию статистика W^* (см. (15)) равна

$$\frac{m \sum [X_{i.(.)} - X_{..(.)}]^2}{\sum \sum [X_{ij(k)} - X_{i.(.)} - X_{.j(k)} - X_{..(k)} + 2X_{..(.)}]^2 / (m-2)}.$$

Число степеней свободы равно $m-1$ для числителя и $(m-1)(m-2)$ для знаменателя; параметр нецентральности равен $\psi^2 = m \sum \alpha_i^2 / \sigma^2$.

К разделу 6

16. Допустим, что в задаче регрессии наблюдаемые значения X_j и Y_j аргумента (независимой переменной) и функции от него отличаются от некоторых истинных значений X'_j и Y'_j ошибками U_j , V_j , которые независимы, распределены нормально с нулевыми средними и дисперсиями σ_U^2 и σ_V^2 . Предполагается, что истинные значения удовлетворяют линейному соотношению $Y'_j = \alpha + \beta X'_j$. Однако контролируемые и, следовательно, постоянными величинами являются не X'_j , а X_j . Обозначая заново X_j через x_j , мы имеем $x_j = X'_j + U_j$, $Y_j = Y'_j + V_j$ и, наконец, $Y_j = \alpha + \beta x_j + W_j$, где $W_j = V_j - \beta U_j$. К проверке гипотез о том, что β или $\alpha + \beta x_0$ имеют некоторые определенные значения, могут быть применены теперь результаты раздела 6.

17. Пусть $X_1, \dots, X_m$; $Y_1, \dots, Y_n$ независимы и распределены нормально с общей дисперсией σ^2 и средними значениями $E(X_i) = \alpha + \beta(u_i - \bar{u})$, $E(Y_j) = \gamma + \delta(v_j - \bar{v})$, где u и v — известные числа. Найдите РНМ инвариантные критерии для гипотез $H: \beta = \delta$ и $H: \alpha = \gamma, \beta = \delta$.

18. Пусть $X_1, \dots, X_n$ независимы и распределены нормально с общей дисперсией σ^2 и средними $\xi_i = \alpha + \beta t_i + \gamma t_i^2$, где t_i известны. Если векторы $(t_1^k, \dots, t_n^k)$, $k = 0, 1, 2$, линейно независимы, то пространство параметров Π_Ω имеет размерность $s = 3$ и оценки $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ по методу наименьших

квадратов являются единственным решением системы уравнений

$$\alpha \sum t_i^k + \beta \sum t_i^{k+1} + \gamma \sum t_i^{k+2} = \sum t_i^k X_i \quad (k=0, 1, 2).$$

Эти решения зависят от величин X_i , и если $\hat{\gamma} = \sum c_i X_i$, то гипотеза $\gamma = 0$ отклоняется при

$$\frac{|\hat{\gamma}|/\sqrt{\sum c_i^2}}{\sqrt{\sum (X_i - \hat{\alpha} - \hat{\beta}t_i - \hat{\gamma}t_i^2)^2/(n-3)}} > C_0.$$

К разделу 7

19. (I) Критерий (43) для гипотезы $H: \Delta \leq \Delta_0$ — РНМ несмещенный.

(II) Найдите РНМ несмещенный критерий для гипотезы $H: \Delta = \Delta_0$ и соответствующие наиболее точные несмещенные доверительные множества для Δ .

20. В модели (41) коэффициент корреляции ρ между двумя наблюдениями X_{ij} , X_{ik} , принадлежащими одному и тому же классу (так называемый коэффициент *внутригрупповой корреляции*), вычисляется по формуле

$$\rho = \sigma_A^2 / (\sigma_A^2 + \sigma^2).$$

К разделу 8

21. Критерии (48) и (49) являются РНМ несмещенными.

22. Пусть X_{ij} определены по (41), но число наблюдений n_i для каждой партии может отличаться от других. Найдите каноническую форму, соответствующую (42), полагая $Y_{i1} = \sqrt{n_i} X_i$. Заметьте, что здесь система достаточных статистик более обширна, чем в случае одинаковых n_i .

23. В рамках модели II общая «гнездовая» классификация с постоянным числом наблюдений в клетках имеет структуру

$$X_{ijk} = \mu + A_i + B_{ij} + C_{ijk} + \dots + U_{ijk} \dots, \quad (93)$$

$$i = 1, \dots, a; \quad j = 1, \dots, b; \quad k = 1, \dots, c; \dots$$

(I) В этом случае возможно приведение к канонической форме, обобщающей (47).

(II) Существуют РНМ несмещенные критерии для гипотез

$$H_A: \frac{\sigma_A^2}{(cd \dots \sigma_B^2 + d \dots \sigma_C^2 + \dots + \sigma^2)} \leq \Delta_0,$$

$$H_B: \frac{\sigma_B^2}{(d \dots \sigma_C^2 + \dots + \sigma^2)} \leq \Delta_0.$$

К разделу 10

24. (I) Если $t < p$, то матрица S , а с нею и матрица S/m , являющаяся несмещенной оценкой для неизвестной матрицы ковариаций исходного p -мерного распределения, будут вырожденными. Если $t \geq p$, то эти матрицы с вероятностью 1 не вырождены.

(II) Если $r + t \leq p$, то лишь один критерий, а именно $\phi(y, u, z) \equiv \alpha$, инвариантен относительно групп G_1 и G_3 раздела 10.

[(II) Величины U устраняются с помощью соображений инвариантности относительно G_1 . $r + t$ строк матриц Y и Z можно предположить линейно независимыми, а любая такая система векторов может быть переведена в любую аналогичную преобразованием группы G_3 .]

25. (I) Если $p < r + m$ и $V = Y'Y$, $S = Z'Z$, то $p \times p$ -матрица $V + S$ с вероятностью 1 не вырождена и корни характеристического уравнения

$$|V - \lambda(V + S)| = 0 \quad (94)$$

образуют максимальный инвариант относительно групп G_1 , G_2 и G_3 .

(II) Среди корней уравнения (94) имеется $p - \min(r, p)$, равных нулю, и $p - \min(m, p)$, равных единице. Других постоянных корней нет, так что число переменных корней, образующих максимальный инвариант, равно $\min(r, p) + \min(m, p) - p$.

[Кратность корня $\lambda = 1$ равна p минус ранг S , т. е. $p - \min(m, p)$. Уравнение (94) не может выполняться ни для какого постоянного $\lambda \neq 0, 1$ для почти всех V, S , так как при любом $\mu \neq 0$ матрица $V + \mu S$ с вероятностью 1 не вырождена.]

26. (I) Пусть A и B — матрицы порядков $k \times m$ и $m \times k$ соответственно. Тогда произведения AB и BA имеют одни и те же ненулевые характеристические числа.

(II) Отсюда получается альтернативное доказательство того, что W (см. (58)) является единственным ненулевым корнем характеристического уравнения (57).

[(I) Если x — ненулевое решение уравнения $ABx = \lambda x$ с $\lambda \neq 0$, то $y = Bx$ будет ненулевым решением для $BAy = \lambda y$.]

27. В случае $r = 1$ статистика W , заданная (58), будет максимальным инвариантом относительно группы, индуцированной группами G_1 и G_3 в области значений статистик Y_i , $U_{\alpha i}$ ($i = 1, \dots, p$; $\alpha = 1, \dots, s - 1$) и $S = Z'Z$.

[Существует такая невырожденная матрица B , что $B'SB = I$ и только первая координата у YB отлична от нуля. Доказательство можно начать с отыскания матрицы B_1 , для которой $B_1'SB_1 = I$, а затем подобрать ортогональную матрицу Q такую, что только первая координата у YB_1Q отлична от нуля.]

28. Пусть $Z_{\alpha i}$ ($\alpha = 1, \dots, m$; $i = 1, \dots, p$) независимы и распределены по закону $N(0, 1)$ и пусть $Q = Q(Y)$ — ортогональная $m \times m$ -матрица, являющаяся функцией от случайной величины Y , не зависящей от величины Z . Определим величины $Z_{\alpha i}^*$ равенством

$$(Z_{1i}^* \dots Z_{mi}^*) = (Z_{1i} \dots Z_{mi}) Q'.$$

Тогда $Z_{\alpha i}^*$ независимы между собой, распределены по закону $N(0, 1)$ и не зависят от Y .

[При каждом y требуемым свойством обладает условное распределение $(Z_{1i} \dots Z_{mi}) Q'(y)$ при условии $Y = y$.]

29. Пусть Z — $m \times p$ -матрица $(Z_{\alpha i})$, где $p \leq m$ и где $Z_{\alpha i}$ независимы и распределены по закону $N(0, 1)$. Обозначим $S = Z'Z$ и S_1 — матрицу, получаемую из S отбрасыванием последней строки и последнего столбца. Тогда отношение определителей $|S|/|S_1|$ имеет χ^2 -распределение с $m - p + 1$ степенями свободы.

[Пусть Q — зависящая от $Z_{11}, \dots, Z_{m1}$ ортогональная матрица, такая, что $(Z_{11} \dots Z_{m1}) Q' = (R 0 \dots 0)$, где $R^2 = \sum_{\alpha=1}^m Z_{\alpha 1}^2$. Тогда

$$S = Z'Q'QZ = \begin{bmatrix} R & 0 & \dots & 0 \\ Z_{12}^* & \dots & \dots & Z_{m2}^* \\ \vdots & \vdots & \ddots & \vdots \\ Z_{1p}^* & \dots & \dots & Z_{mp}^* \end{bmatrix} \begin{bmatrix} R & Z_{12}^* & \dots & Z_{1p}^* \\ 0 & \dots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & Z_{m2}^* & \dots & Z_{mp}^* \end{bmatrix},$$

где $Z_{\alpha i}^*$ получаются из Z преобразованием Q . Первая матрица справа равна произведению

$$\left(\frac{R}{Z_1^*} \middle| \frac{O}{I} \right) \left(\frac{1}{0} \middle| \frac{0}{Z^{*'}} \right),$$

где Z^* — это $(m-1) \times (p-1)$ -матрица с элементами $Z_{\alpha i}^*$ ($\alpha = 2, \dots, m$; $i = 2, \dots, p$), I — это $(p-1) \times (p-1)$ единичная матрица, Z_1^* — вектор-столбец $(Z_{12}^* \dots Z_{1p}^*)'$ и где 0 обозначает строку или столбец, состоящий из нулей. Мы видим, что $|S|$ равен R^2 , умноженному на определитель матрицы $Z^{*'}Z^*$. Теперь, S_1 получается из матрицы Z двумя операциями: отбрасыванием у Z последнего столбца и последующим умножением слева полученной $m \times (p-1)$ -матрицы на транспонированную. Поэтому $|S_1|$ равен R^2 , умноженному на детерминант матрицы, получаемой из $Z^{*'}Z^*$ отбрасыванием последней строки и последнего столбца. Таким образом, мы свели отношение $|S|/(S_1)$ к аналогичному отношению для величин $Z_{\alpha i}^*$, но уже с заменой m на $m-1$ и p на $p-1$. По индукции устанавливаем, что при любом $k < p$ можно заменить m и p на $m-k$ и $p-k$. В частности, можно вычислять $|S|/|S_1|$ в предположении, что m и p заменены на $m-(p-1)$ и $p-(p-1) = 1$. В этом случае Z' есть матрица-строка $(Z_{11} \dots Z_{m-p+1, 1})$, детерминант S равен $|S| = \sum_{\alpha=1}^{m-p+1} Z_{\alpha 1}^2$ и потому имеет распределение χ_{m-p+1}^2 . Наконец, так как S есть (1×1) -матрица, то $|S_1|$ заменяется единицей.]

30. Определенная формулой (58) статистика $W = YS^{-1}Y'$ имеет такое же распределение вероятностей, как отношение двух независимых величин, причем числитель имеет нецентральное χ^2 -распределение с параметрами нецентральности ψ^2 и p степенями свободы, а знаменатель — центральное χ^2 -распределение с $m+1-p$ степенями свободы.

[Распределение W не изменяется, если к $(Y_1, \dots, Y_p)$ и к каждому из m векторов $(Z_{\alpha 1}, \dots, Z_{\alpha p})$ применить одно и то же невырожденное преобразование. Поэтому можно предположить, что эти векторы имеют единичную матрицу ковариаций. Пусть Q — зависящая от Y ортогональная матрица, такая, что $(Y_1 \dots Y_p)Q = (00 \dots T)$, где $T^2 = \sum Y_i^2$. Так как QQ' равно единичной матрице, то

$$W = (YQ)(Q'S^{-1}Q)(Q'Y') = (0 \dots 0T)(Q'S^{-1}Q)(0 \dots 0T)'$$

Следовательно, W равно произведению величины T^2 (которая имеет нецентральное распределение χ^2 с p степенями свободы и параметром нецентральности ψ^2) на элемент, стоящий на пересечении p -й строки и p -го столбца матрицы $Q'S^{-1}Q = (Q'SQ)^{-1} = (Q'Z'ZQ)^{-1}$. В соответствии с задачами 28 и 29 эта матрица не зависит от Y и величина, обратная к рассматриваемому элементу, имеет распределение χ_{m-p+1}^2 .]

31. Пусть $(X_{v1}, \dots, X_{vp})$, $v = 1, \dots, N$, — выборка из p -мерного нормального распределения с неизвестной матрицей ковариаций и средними $E(X_{vi}) = \alpha_i + \beta_i(u_v - u)$, где u — известные постоянные. Гипотеза H : $\beta_1 = \dots = \beta_p = 0$ является многомерной линейной гипотезой с $r = 1$, $s = 2$. Мы имеем $\hat{\alpha}_i = \hat{\alpha}_i = X_{.i}$, $\hat{\beta}_i = \sum_{v=1}^N (u_v - u)(X_{vi} - X_{.i}) / \sum_{v=1}^N (u_v - u)^2$. Статистика критерия W задается формулой (58) с

$$S_{ij} = \sum [X_{vi} - \hat{\alpha}_i - \hat{\beta}_i(u_v - u)][X_{vj} - \hat{\alpha}_j - \hat{\beta}_j(u_v - u)]$$

и

$$Y_i Y_j = \hat{\beta}_i \hat{\beta}_j \sum (u_v - u.)^2.$$

32. Пусть $X = (X_{ai})$, $i = 1, \dots, p$; $a = 1, \dots, N$, — выборка из p -мерного нормального распределения, и пусть $q < p$, $\max(q, p - q) \leq N$. Рассмотрим гипотезу H о том, что $(X_{11}, \dots, X_{1q})$ не зависит от $(X_{1, q+1}, \dots, X_{1p})$, т. е. что ковариации $\sigma_{ij} = E(X_{ai} - \xi_i)(X_{aj} - \xi_j)$ равны нулю при всех $i \leq q$, $j > q$. Проблема проверки H остается инвариантной относительно преобразований $X_{ai}^* = X_{ai} + b_i$ и $X^* = XC$, где C — любая невырожденная $p \times p$ -матрица, имеющая структуру

$$C = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{22} \end{pmatrix},$$

причем C_{11} и C_{22} имеют порядок $q \times q$ и $(p - q) \times (p - q)$ соответственно.

(I) Максимальным инвариантом преобразования, индуцированного в пространстве значений достаточных статистик $X_{\cdot i}$ и матрицы S , которую мы представим в форме $S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$, являются q корней уравнения $|S_{12} S_{22}^{-1} S_{21} - \lambda S_{11}| = 0$.

(II) В случае $q = 1$ максимальным инвариантом будет статистика $R^2 = S_{12} S_{22}^{-1} S_{21} / S_{11}$, являющаяся квадратом *множественного коэффициента корреляции* между X_{11} и $(X_{12}, \dots, X_{1p})$. Распределение R^2 зависит только от ρ^2 — квадрата теоретического коэффициента множественной корреляции, который получается из R^2 заменой элементов матрицы S их математическими ожиданиями σ_{ij} .

(III) Используя то, что плотность вероятности R^2 равна *)

$$\frac{(1 - R^2)^{\frac{1}{2}(N - p - 2)} (R^2)^{\frac{1}{2}(p - 1) - 1} (1 - \rho^2)^{\frac{1}{2}(N - 1)}}{\Gamma\left[\frac{1}{2}(N - 1)\right] \Gamma\left[\frac{1}{2}(N - p)\right]} \sum_{h=0}^{\infty} \frac{(\rho^2)^h (R^2)^h \Gamma^2\left[\frac{1}{2}(N - 1) + h\right]}{h! \Gamma\left[\frac{1}{2}(p - 1) + h\right]}$$

и что гипотеза H при $q = 1$ эквивалентна гипотезе $\rho = 0$, покажите, что РНМ инвариантный критерий отвергает гипотезу при $R^2 > C_0$.

(IV) При $\rho = 0$ статистика $\frac{R^2}{1 - R^2} \cdot \frac{N - p}{p - 1}$ имеет F -распределение с $p - 1$ и $N - p$ степенями свободы.

[(I) Допустим, что уравнения $|S_{12} S_{22}^{-1} S_{21} - \lambda S_{11}| = 0$ и $|S_{12}^* S_{22}^{*-1} S_{21}^* - \lambda S_{11}^*| = 0$ имеют одни и те же корни. Тогда найдутся матрицы B и C такие, что $BS_{11}B' = I = CS_{11}^*C'$ и $BS_{12}S_{22}^{-1}S_{21}B' = CS_{12}^*S_{22}^{*-1}S_{21}^*C' = A$, где A — диагональная матрица, на диагонали которой стоят корни λ . Так как матрицы S_2^{-1} и S_2^{*-1} положительно определены, то существуют невырожденные матрицы E и E^* такие, что $S_2^{-1} = EE'$ и $S_2^{*-1} = E^*E^{*-1}$ (это можно установить, приводя E к диагональной форме с помощью ортогонального преобразования). Тогда

$$(BS_{12}E)(BS_{12}E)' = (CS_{12}^*E^*)(CS_{12}^*E^*)'$$

*) См., например, Anderson, An Introduction to Multivariate Statistical Analysis, New York, John Wiley & Sons, 1958. [Русский перевод см. стр. 363.]

и аргументация, использованная в разделе 10 в связи с G_2 , показывает, что существует ортогональная матрица Q , такая, что $BS_{12}EQ = CS_{12}^*E^*$, откуда $C_{11} = C^{-1}B$ и $C'_{22} = EQE^{*-1}$.]

К разделу 12

33. Проблема проверки гипотезы $\eta \in \Pi_\omega$, когда распределение Y определяется из (62) с $\eta \in \Pi_\Omega$, остается инвариантной относительно подходящей группы линейных преобразований. По отношению к этой группе критерий (63) будет РНМ инвариантным. Вероятность отклонения гипотезы этим критерием находится из (65) и (66) для всех точек $(\eta_1, \dots, \eta_d)$.

[Существует невырожденное линейное преобразование $Z = CY$, для которого $C'A^{-1}C$ совпадает с единичной матрицей; в терминах Z задача сводится к линейной гипотезе с известной дисперсией.]

К разделу 13

34. Допустим, что уравнение касательной плоскости $\bar{\mathcal{P}}$ в точке π имеет вид $p_i = \pi_i(1 + a_{i1}\xi_1 + \dots + a_{is}\xi_s)$ и что векторы $(a_{i1}, \dots, a_{is})$ ортогональны в том смысле, что $\sum a_{ik}a_{il}\pi_i = 0$ при всех $k \neq l$.

(1) Если $(\hat{\xi}_1, \dots, \hat{\xi}_s)$ минимизирует $\sum (v_i - p_i)^2/\pi_i$ при условии $p \in \bar{\mathcal{P}}$, то $\hat{\xi}_j = \sum_i a_{ij}v_i / \sum_i a_{ij}^2\pi_i$.

(11) Статистика (76) критерия для проверки $H: p = \pi$ сводится к

$$n \sum_{j=1}^s \left[\left(\sum_{i=1}^m a_{ij}v_i \right)^2 / \sum_{i=1}^m a_{ij}^2\pi_i \right].$$

35. *Независимость в таблицах сопряженности признаков.* Рассмотрим классификацию n элементов по двум признакам, первый из которых принимает значения $A_1, \dots, A_a$, а второй — $B_1, \dots, B_b$. Пусть n_{ij} — число элементов, обладающих свойствами A_i и B_j . Критерий отношения правдоподобия для проверки гипотезы H о независимости признаков A и B отвергает гипотезу, когда

$$\Lambda = \prod_{i,j} n_{ij}^{n_{ij}} / \prod_i n_i^{bn_i} \cdot \prod_j n_j^{an_j}$$

слишком велико. Здесь $n_i = \sum_j n_{ij}/b$, $n_j = \sum_i n_{ij}/a$. При больших n величина $-2 \log \Lambda$ при гипотезе H распределена приблизительно по закону χ^2 с $(r-1)(s-1)$ степенями свободы.

«Правдоподобие» полиномиальной выборки $x_1, \dots, x_m$ с m классами пропорционально выражению $p_1^{x_1} \dots p_m^{x_m}$, максимальное значение которого равно $(x_1/n)^{x_1} \dots (x_m/n)^{x_m}$. Это можно доказать, например, рассматривая n чисел, среди которых x_i равны p_i/x_i при $i = 1, \dots, m$, и замечая, что их среднее геометрическое не превосходит их среднего арифметического. Применяя найденный результат к полиномиальным схемам с вероятностями p_{ij} ($i = 1, \dots, a$; $j = 1, \dots, b$), образующими Ω , и $p_{ij} = r_i p'_j$ ($\sum p_i = \sum p'_j = 1$), образуящими ω , мы приходим к требуемому выводу.]

15. Литературные ссылки

Настоящая глава служит кратким введением в каждую из трех тем, которые в совокупности охватывают подавляющую часть современных статистических методов: дисперсионный анализ, многомерный анализ, теорию χ^2 -критериев.

Разработка дисперсионного анализа была начата главным образом Р. А. Фишером, значительная часть результатов приведена в его книгах (1925, 1935). Исчерпывающим образом дисперсионный анализ изложен в книгах Кемпторна (1952) и Шеффе (1959). Обзор некоторых аспектов метода имеется в статьях Шеффе (1956) и Кокрена (1957)*).

Многомерный анализ подробно изложен в книгах Андерсона (1958), Рао (1952) и Кендалла (1957)**).

Обзор применений χ^2 -метода дан в статьях Кокрена (1952, 1954***), Ватсона (1959), Ланкастера (1969).

Андерсон (Anderson T. W.)

(1958) An Introduction to Multivariate Analysis, New York, John Wiley and Sons. (Русск. перевод: Андерсон Т., Введение в многомерный статистический анализ, М., Физматгиз, 1963.)

Бартлетт (Bartlett M. S.)

(1947) The use of transformations, Biometrics, vol. 3, 39—52. [Рассматриваются в числе других следующие преобразования: логарифмическое извлечение квадратного корня и арк-синуса.]

Бокс (Box G. E. P.)

(1949) A general distribution theory for a class of likelihood ratio criteria, Biometrika, vol. 36, 317—346.

Вальд (Wald A.)

(1942) On the power function of the analysis of variance test, Ann. Math. Stat. vol. 13, 434—439. [Задача 5. Эта задача имеется также в работе: Hsu, On the power function of the E^2 -test and the T^2 -test, Ann. Math. Stat., vol. 16 (1945), 278—286.]

(1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large, Trans. Am. Math. Soc., vol. 54, 426—482. [Общее асимптотическое распределение и свойства оптимальности критериев отношения правдоподобия и им асимптотически эквивалентных.]

Ватсон (Watson S. S.)

(1959) Some recent results in chi-square goodness-of-fit tests, Biometrics, 15, № 3, 440—468.

Гербах (Herbach L. H.)

(1957) Optimum properties of analysis of variance tests based on model II and some generalizations of model II, Scientific Paper № 6, Engin. Stat. Lab., New York University.

*) См. Рао, Линейные статистические методы и их применение, пер. с англ., М., Наука, 1968. — *Прим. ред.*

**) См. Кендалл М., Стьюарт А., Многомерный статистический анализ и временные ряды, пер. с англ., М., Наука, 1976. — *Прим. ред.*

***) См. также Watson G. S., Some recent results in chi-square goodness-of-fit test, Biometrics 15, № 3, 440—468; Чибисов Д. М., Некоторые критерии типа хи-квадрат для непрерывных распределений, Теория вероятн. и ее примен. 16, № 1, 3—20.

(1959) Properties of model II-type analysis of variance tests, *Ann. Math. Stat.*, vol. 30, 939—959.

Кемпторн (Kempthorne O.)

(1952) *The Design and Analysis of Experiments*, New York, John Wiley and Sons.

Кендалл (Kendall M. G.)

(1957) *A Course in Multivariate Analysis*, London, Charles Griffin.

Кендалл и Стьюарт (Kendall M. G., Stuart A.)

(1973) Статистические выводы и связи, пер. с англ., М., Наука.

(1976) Многомерный статистический анализ и временные ряды, пер. с англ., М., Наука.

Кокрен (Cochran W. G.)

(1952) The χ^2 -test of goodness of fit, *Ann. Math. Stat.*, vol. 2, 315—345.

(1954) Some methods for strengthening the common χ^2 -test, *Biometrics*, vol. 10, 417—451.

(1957) Analysis of covariance: Its nature and uses, *Biometrics*, vol. 13, 261—281.

Колоджейчик (Kolodziejczyk S.)

(1935) An important class of statistical hypotheses, *Biometrika*, vol. 37, 161—190. [Общая линейная одномерная гипотеза рассматривается с точки зрения отношения правдоподобия.]

Ланкастер (Lancaster H. O.)

(1969) *The chi-squared distribution*, N. Y., John Wiley and Sons.

Нейман (Neuman J.)

(1949) Contributions to the theory of the χ^2 -test, *Proc. Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, Univ. Calif. Press, 239—273. [Дана теория χ^2 -критериев с ограниченными альтернативами.]

Нейман и Пирсон (Neuman J. and Pearson E. S.)

(1928) The use and interpretation of certain test criteria for purposes of statistical inference, *Biometrika*, vol. 20, A, 175—240, 263—294. [Предложен критерий отношения правдоподобия как метод, который учитывает оба типа ошибок. Критерий применяется к целому ряду проблем проверки гипотез.]

Ньюман (Newman D.)

(1939) Range in samples from a normal population, *Biometrika*, vol. 31, 20—30. [Рассматривается критерий (22), который приписывается автором «Студенту» (Госсету).]

Пирсон (Pearson K.)

(1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that in can be reasonably supposed to have arisen from random sampling, *Phil. Mag. Ser. 5*, vol. 50, 157—172. [Предложен χ^2 -критерий (71) для проверки простой полиномиальной гипотезы. Дается его предельное распределение при проверяемой гипотезе. Критерий обобщается на случай сложных гипотез, но содержит ошибку в определении числа степеней свободы предельного распределения; правильное решение для общего случая было найдено Фишером (1924а). Применения.]

Рао (Rao C. R.)

(1952) *Advanced Statistical Methods in Biometric Research*, New York, John Wiley and Sons.

(1968) Линейные статистические методы и их применения, пер. с англ., М., Наука.

Симаик (Simaika J. B.)

(1941) An optimum property of two statistical tests, *Biometrika*, vol. 32, 70—80. [Показано, что критерий (58) является РНМ в классе всех критериев, функция мощности которых зависит только от параметра нецентральности (59), и установлено соответствующее свойство для критерия множественной корреляции, данного в задаче 32 (III).]

Сюй (Hsu P. L.)

(1938) Notes on Hotelling's generalized T^2 , *Ann. Math. Stat.*, vol. 9, 231—243. [Получено распределение T^2 -статистики в нецентральном случае. Статистика применена к классу проблем, описанных в разделе II после примера 9. Вывод T^2 -распределения, указанный в задачах 29 и 30, принадлежит Wilks, Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis, *Ann. Math. Stat.*, vol. 28 (1957), 415—423. Этот вывод был также отмечен Стейном (ср. Вейсман, цит. работа, стр. 416) и Боукером (ср. Андерсон (1958), стр. 107).]

(1941) Analysis of variance from the power function stand-point, *Biometrika*, vol. 32, 62—69. [Показано, что критерий (7) является РНМ в классе всех критериев, функция мощности которых зависит только от параметра нецентральности.]

Тэнг (Tang P. C.)

(1938) The power function of the analysis of variance test with tables and illustrations of their use, *Stat. Math. Res. Mem.*, vol. II, 126—149.

Уилкс (Wilks S. S.)

(1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses, *Ann. Math. Stat.*, vol. 9, 60—62. [Выведено асимптотическое распределение отношения правдоподобия в случае справедливости гипотезы.]

Фишер (Fisher R. A.)

(1924a) The conditions under which chi square measures the discrepancy between observation and hypothesis, *J. Roy. Stat. Soc.*, vol. 87, 442—450. [Получено предельное распределение (при проверяемой гипотезе) χ^2 -статистики для случая сложных гипотез. Обсуждается зависимость этого распределения от метода, используемого при оценке параметров.]

(1924b) On a distribution yielding the error functions of several well-known statistics, *Proc. Int. Math. Congress, Toronto*, 805—813. [Обсуждается вопрос использования Z-распределения (которое эквивалентно F-распределению) в дисперсионном анализе (модель I) и регрессионном анализе. Получено распределение выборочного коэффициента множественной корреляции в случае, когда соответствующий теоретический коэффициент равен 0.]

(1925) *Statistical Methods for Research Workers*, Edinburgh, Oliver and Boyd, 1st ed.

(1928) The general sampling distribution of the multiple correlation coefficient, *Proc. Roy. Soc., Ser. A*, vol. 121, 664—673. [Вывод нецентрального χ^2 -и нецентрального бета-распределения и распределения выборочного коэффициента множественной корреляции для произвольных значений теоретического коэффициента множественной корреляции]

(1935) *The Design of Experiments*, Edinburgh, Oliver and Boyd, 1st ed.

Ханти Стейн (Hunt G. and Stein C. M.)

(1946) Most stringent tests of statistical hypotheses Не опубликована. [Докладывается, что критерии (7) и (61) РНМ почти инвариантны и что корни уравнения (57) составляют максимальный инвариант.]

Хартли (Hartley H. O.)

(1950) Maximum F ratio as a short-cut test for heterogeneity of variance, *Biometrika*, vol. 37, 308—312. [Предложен критерий (23).]

Хотеллинг (Hotelling H.)

(1931) The generalization of Student's ratio, *Ann. Math. Stat.*, vol. 2, 360—378. [Предложена статистика (58) в качестве многомерного обобщения t -статистики Стьюдента и получено распределение этой статистики при проверяемой гипотезе.]

Шеффе (Scheffé H.)

(1956) A «mixed model» for the analysis of variance, *Ann. Math. Stat.*, vol. 27, 23—36. [Пример 11.]

(1956) Alternative models for the analysis of variance, *Ann. Math. Stat.*, vol. 27, 251—271.

(1959) Fitting straight lines when one variable is controlled, *J. Am. Stat. Assoc.*, vol. 53, 106—117. [Задача 16.]

(1959) *Analysis of Variance*, New York, John Wiley and Sons, (Русск. перевод: Шеффе, Дисперсионный анализ, М., Физматгиз, 1963).

Эйзенхарт (Eisenhart C.)

(1947) The assumptions underlying the analysis of variance, *Biometrics*, vol. 3, 1—21. [Обсуждается различие между моделями I и II.]

ГЛАВА 8

ПРИНЦИП МИНИМАКСА

1. Критерии с гарантированной мощностью

Обсуждавшиеся до сих пор ограничения — несмещенность и инвариантность — обладают тем недостатком, что они применимы и приводят к оптимальным решениям только в весьма ограниченных классах задач. Поэтому теперь мы обратимся к альтернативному подходу, который потенциально имеет более широкую область применимости. К сожалению, его использование в конкретных задачах приводит к трудностям и может быть успешно проведено главным образом в случаях, когда существуют РНМ инвариантные критерии.

Одна из важных для планирования эксперимента характеристик — число наблюдений, необходимое для того, чтобы результирующая статистическая процедура имела заданную точность или чувствительность. В проблеме проверки гипотез это означает, что вероятности ошибок первого и второго рода не должны превосходить некоторых заданных границ, скажем α и $1 - \beta$, так что рассматриваемые критерии должны удовлетворять условиям

$$\begin{aligned} E_{\theta} \varphi(X) &\leq \alpha \quad \text{при } \theta \in \Omega_H, \\ E_{\theta} \varphi(X) &\geq \beta \quad \text{при } \theta \in \Omega_K. \end{aligned} \quad (1)$$

Если функция мощности $E_{\theta} \varphi(X)$ непрерывна и $\alpha < \beta$, то (1) не может выполняться для множеств Ω_H и Ω_K , которые соприкасаются. Эта математическая трудность связана отчасти с тем фактом, что разделение значений параметра на два класса Ω_H и Ω_K , которым соответствуют два различных решения, часто оказывается не резким. Между значениями, для которых одно или другое решение очевидным образом правильно, могут лежать другие, для которых относительные преимущества и недостатки принятия и отклонения гипотезы приближенно уравниваются.

ся. Соответственно мы предположим, что Ω разбито на сумму трех множеств

$$\Omega = \Omega_H + \Omega_I + \Omega_K,$$

где Ω_I обозначает *зону безразличия*, а Ω_K — класс значений параметра, отличающихся от постулируемых гипотезой так сильно, что ложное принятие H является серьезной ошибкой, вероятность которой не должна превышать $1 - \beta$.

Чтобы понять, как определяется размер выборки в этой ситуации, предположим, что доступные для наблюдения случайные величины образуют последовательность $X_1, X_2, \dots$. Допустим на время, что n фиксировано, и пусть $X = (X_1, \dots, X_n)$. В обычной для ряда применений обстановке (точную формулировку см. в задаче 1) существует критерий φ_n , который максимизирует

$$\inf_{\Omega_K} E_{\theta} \varphi(X) \quad (2)$$

в классе всех критериев уровня α , основанных на X . Обозначим $\beta_n = \inf_{\Omega_K} E_{\theta} \varphi(X)$ и допустим, что при достаточно больших n существует критерий, удовлетворяющий (1)*).

Желаемый размер выборки (т. е. наименьшее n , для которого $\beta_n \geq \beta$) определяется в данном случае методом проб. Для этого требуется уметь при каждом фиксированном n определить критерий, который максимизирует (2) при условии

$$E_{\theta} \varphi(X) \leq \alpha \quad \text{при} \quad \theta \in \Omega_H. \quad (3)$$

Метод отыскания критерия с подобным *максиминным* свойством (т. е. критерия, максимизирующего минимальную мощность на Ω_K) получается обобщением теоремы 7 главы 3. В последующих рассуждениях удобно изменить обозначения: мы назовем теперь ω и ω' подмножества Ω , которые ранее обозначались Ω_H и Ω_K . Пусть $\mathcal{P} = \{P_{\theta}, \theta \in \omega \cup \omega'\}$ — семейство распределений вероятностей на пространстве выборок $(\mathcal{X}, \mathcal{A})$ с плотностями $p_{\theta} = dP_{\theta}/d\mu$ по некоторой σ -конечной мере μ . Допустим, что эти плотности $p_{\theta}(x)$, рассматриваемые как функции пары переменных (x, θ) , измеримы относительно $(\mathcal{A} \times \mathcal{B})$ и $(\mathcal{A} \times \mathcal{B}')$, где $\mathcal{B}$ и $\mathcal{B}'$ — заданные σ -поля в ω и ω' . При этих предположениях и при условиях, указываемых в приводимой ниже теореме, критерий с желательными свойствами совпадает с решением надлежаще подобранной байесовской проблемы.

*) Условия, при которых это справедливо, даны в статье: Berger, On uniformly consistent tests, Ann. Math. Stat. 22 (1951), 289—293.

Теорема 1. Пусть λ и λ' — любые распределения на $\mathcal{B}$ и $\mathcal{B}'$ и пусть $\Phi_{\lambda, \lambda'}$ — наиболее мощный критерий уровня α для проверки

$$h(x) = \int_{\omega} p_{\theta}(x) d\lambda(\theta)$$

при альтернативе

$$h'(x) = \int_{\omega'} p_{\theta}(x) d\lambda'(\theta).$$

Пусть $\beta_{\lambda, \lambda'}$ — мощность этого критерия при альтернативе h' . Если существуют λ и λ' такие, что

$$\begin{aligned} \sup_{\omega} E_{\theta} \Phi_{\lambda, \lambda'}(X) &\leq \alpha, \\ \inf_{\omega'} E_{\theta} \Phi_{\lambda, \lambda'}(X) &= \beta_{\lambda, \lambda'}, \end{aligned} \quad (4)$$

то

(I) $\Phi_{\lambda, \lambda'}$ максимизирует $\inf_{\omega'} E_{\theta} \Phi(X)$ в классе всех критериев уровня α для проверки гипотезы $H: \theta \in \omega$; он является единственным критерием с этим свойством, если единствен наиболее мощный критерий уровня α для проверки h при альтернативе h' .

(II) Пара распределений λ, λ' является наименее благоприятной в том смысле, что для любой другой пары ν, ν'

$$\beta_{\lambda, \lambda'} \leq \beta_{\nu, \nu'}.$$

Доказательство. (I) Пусть Φ^* — любой другой критерий уровня α для проверки H . Тогда он является критерием уровня α и для проверки простой гипотезы: плотность X равна h . Поэтому мощность Φ^* при альтернативе h' не превосходит $\beta_{\lambda, \lambda'}$. Мы имеем

$$\inf_{\omega'} E_{\theta} \Phi^*(X) \leq \int_{\omega'} E_{\theta} \Phi^*(X) d\lambda'(\theta) \leq \beta_{\lambda, \lambda'} = \inf_{\omega'} E_{\theta} \Phi_{\lambda, \lambda'}(X),$$

и последнее неравенство будет строгим, если критерий $\Phi_{\lambda, \lambda'}$ единствен.

(II) Пусть ν, ν' — любые другие распределения на $(\omega, \mathcal{B})$ и $(\omega', \mathcal{B}')$. Обозначим

$$g(x) = \int_{\omega} p_{\theta}(x) d\nu(\theta); \quad g'(x) = \int_{\omega'} p_{\theta}(x) d\nu'(\theta).$$

Но как $\Phi_{\lambda, \lambda'}$, так и $\Phi_{\nu, \nu'}$ являются критериями уровня α для гипотезы: плотность X равна $g(x)$. Поэтому

$$\beta_{\nu, \nu'} \geq \int \Phi_{\lambda, \lambda'}(x) g'(x) d\mu(x) \geq \inf_{\omega'} E_{\theta} \Phi_{\lambda, \lambda'}(X) = \beta_{\lambda, \lambda'}.$$

Следствие 1. Пусть λ, λ' — два распределения вероятностей, а C — такая постоянная, что

$$\Phi_{\lambda, \lambda'}(x) = \begin{cases} 1, & \text{если } \int_{\omega'} p_{\theta}(x) d\lambda'(\theta) > C \int_{\omega} p_{\theta}(x) d\lambda(\theta), \\ \gamma, & \text{если } \int_{\omega'} p_{\theta}(x) d\lambda'(\theta) = C \int_{\omega} p_{\theta}(x) d\lambda(\theta), \\ 0, & \text{если } \int_{\omega'} p_{\theta}(x) d\lambda'(\theta) < C \int_{\omega} p_{\theta}(x) d\lambda(\theta) \end{cases} \quad (5)$$

служит критерием уровня α для гипотезы: плотность X равна $\int_{\omega} p_{\theta}(x) d\lambda(\theta)$, и что

$$\lambda(\omega_0) = \lambda'(\omega'_0) = 1, \quad (6)$$

где

$$\begin{aligned} \omega_0 &= \{\theta: \theta \in \omega \text{ и } E_{\theta} \Phi_{\lambda, \lambda'}(X) = \sup_{\theta' \in \omega} E_{\theta'} \Phi_{\lambda, \lambda'}(X)\}, \\ \omega'_0 &= \{\theta: \theta \in \omega' \text{ и } E_{\theta} \Phi_{\lambda, \lambda'}(X) = \inf_{\theta' \in \omega'} E_{\theta'} \Phi_{\lambda, \lambda'}(X)\}. \end{aligned}$$

Тогда имеет место заключение теоремы 1.

Доказательство. Определим h, h' и $\beta_{\lambda, \lambda'}$ так же, как в теореме 1. Из наших предположений вытекает, во-первых, что $\Phi_{\lambda, \lambda'}$ является наиболее мощным критерием уровня α для проверки h при альтернативе h' ; во-вторых, что

$$\sup_{\omega} E_{\theta} \Phi_{\lambda, \lambda'}(X) = \int_{\omega} E_{\theta} \Phi_{\lambda, \lambda'}(X) d\lambda(\theta) = \alpha$$

и, в-третьих, что

$$\inf_{\omega'} E_{\theta} \Phi_{\lambda, \lambda'}(X) = \int_{\omega'} E_{\theta} \Phi_{\lambda, \lambda'}(X) d\lambda'(\theta) = \beta_{\lambda, \lambda'}.$$

Таким образом, (4) выполняется, а потому применима теорема 1.

Допустим, что множества Ω_H, Ω_I и Ω_K определены в терминах неотрицательной функции d (измеряющей расстояние от θ до H) формулами $\Omega_H = \{\theta: d(\theta) = 0\}$; $\Omega_I = \{\theta: 0 < d(\theta) < \Delta\}$; $\Omega_K = \{\theta: d(\theta) \geq \Delta\}$.

Примем также, что функция мощности любого критерия непрерывна по θ . В пределе, при $\Delta = 0$, зона безразличия отсутствует. В этом случае Ω_K превращается в множество $\{\theta: d(\theta) > 0\}$ и точная нижняя грань $\beta(\theta)$ на Ω_K будет $\leq \alpha$ для любого критерия уровня α . Эта точная нижняя грань максимизируется, следовательно, любым критерием, для которого $\beta(\theta) \geq \alpha$ при

всех $\theta \in \Omega_K$, т. е. любым несмещенным критерием. Мы видим, что несмещенность получается предельным переходом из максиминного свойства. Другая предельная форма (более полезная, так как она обычно приводит к единственному критерию) дается следующим определением: говорят, что критерий φ_0 *локально максимизирует минимум мощности* *), если, каков бы ни был другой критерий φ , найдется Δ_0 такое, что

$$\inf_{\omega_\Delta} \beta_{\varphi_0}(\theta) \geq \inf_{\omega_\Delta} \beta_\varphi(\theta) \text{ при всех } 0 < \Delta < \Delta_0. \quad (7)$$

Здесь ω_Δ обозначает множество тех θ , для которых $d(\theta) \geq \Delta$.

2. Примеры

В главе 3 было установлено, что для семейства плотностей, зависящих от действительного параметра θ таким образом, что при $\theta < \theta'$ отношение $p_{\theta'}(x)/p_\theta(x)$ является монотонной функцией от некоторой действительной статистики, существует РНМ критерий проверки H : $\theta \leq \theta_0$ при альтернативе $\theta > \theta_0$. Указанное предположение, хотя оно и выполняется для однопараметрических экспоненциальных семейств, весьма ограничительно. Поэтому в действительности РНМ критерий для H существует довольно редко. Более общий подход предлагается формулировками предыдущего раздела. Так, если зону безразличия составляют те значения θ , для которых $\theta_0 < \theta < \theta_1$, то задача сводится к максимизации минимума мощности на классе альтернатив ω' : $\theta \geq \theta_1$. Можно ожидать, что при надлежащих ограничениях наименее благоприятные распределения λ и λ' из теоремы 1 приписывают вероятность 1 точкам θ_0 и θ_1 соответственно и что, следовательно, область отклонения максиминного критерия задается неравенством $p_{\theta_1}(x)/p_{\theta_0}(x) > C$. Приводимая ниже лемма дает для этого достаточные условия.

Лемма 1. Пусть $X_1, \dots, X_n$ независимы и одинаково распределены с плотностью $f_\theta(x)$, где θ и x — действительные переменные. Допустим, что при любом $\theta < \theta'$ отношение $f_{\theta'}(x)/f_\theta(x)$ является неубывающей функцией x . Тогда критерий φ уровня α для проверки H , максимизирующей минимум мощности на ω' , определяется равенствами

$$\varphi(x_1, \dots, x_n) = \begin{cases} 1, & \text{если } r(x_1, \dots, x_n) > C, \\ \gamma, & \text{если } r(x_1, \dots, x_n) = C, \\ 0, & \text{если } r(x_1, \dots, x_n) < C, \end{cases} \quad (8)$$

*) Свойство локальной оптимальности вне связи с расстоянием d рассматривается в задаче 4.

где $r(x_1, \dots, x_n) = f_{\theta_1}(x_1) \dots f_{\theta_1}(x_n) / f_{\theta_0}(x_1) \dots f_{\theta_0}(x_n)$ и где C и γ находятся из уравнения

$$E_{\theta_0} \Phi(X_1, \dots, X_n) = \alpha. \quad (9)$$

Доказательство. Функция $\Phi(x_1, \dots, x_n)$ не убывает по каждому из своих аргументов, так что по лемме 2 главы 3 при $\theta < \theta'$

$$E_{\theta} \Phi(X_1, \dots, X_n) \leq E_{\theta'} \Phi(X_1, \dots, X_n).$$

Следовательно, функция мощности для Φ монотонна и Φ — критерий уровня α . Так как $\Phi = \Phi_{\lambda, \lambda'}$, где λ и λ' — распределения, приписывающие вероятность 1 точкам θ_0 и θ_1 , то условие (4) выполнено, что доказывает как желаемый результат, так и то, что распределения λ, λ' — наименее благоприятные.

Пример 1. Пусть θ — параметр сдвига, т. е. $f_{\theta}(x) = g(x - \theta)$. Допустим для простоты, что $g(x) > 0$ при всех x . Мы покажем, что семейство $f_{\theta}(x)$ имеет монотонное относительно x отношение правдоподобия тогда и только тогда, когда функция $-\log g$ является выпуклой. Условие монотонности по x отношения правдоподобия эквивалентно

$$\frac{g(x - \theta')}{g(x - \theta)} \leq \frac{g(x' - \theta')}{g(x' - \theta)} \quad \text{при всех } x < x', \quad \theta < \theta',$$

что в свою очередь эквивалентно

$$\log g(x' - \theta) + \log g(x - \theta') \leq \log g(x - \theta) + \log g(x' - \theta').$$

Поскольку $x - \theta = t(x - \theta') + (1 - t)(x' - \theta)$ и $x' - \theta' = (1 - t)(x - \theta) + t(x' - \theta)$, где $t = (x' - x) / (x' - x + \theta' - \theta)$, то выпуклость $-\log g$ достаточна для выполнения рассматриваемого соотношения. Покажем, что это условие и необходимо. Пусть $a < b$ — любые действительные числа, и пусть $x - \theta' = a$, $x' - \theta = b$, $x' - \theta' = x - \theta$. Тогда $x - \theta = \frac{1}{2}(x' - \theta + x - \theta') = \frac{1}{2}(a + b)$ и условие монотонности отношения правдоподобия влечет неравенство

$$\frac{1}{2} [\log g(a) + \log g(b)] \leq \log g\left[\frac{1}{2}(a + b)\right].$$

Так как функция $-\log g$ измерима, то она выпукла *).

Вышеприведенному условию удовлетворяют, например (кроме нормального распределения, для которого совместные плотности $p_{\theta}(x_1, \dots, x_n)$ образуют экспоненциальное семейство), симметричное экспоненциальное распределение с

$$g(x) = \frac{1}{2} e^{-|x|}$$

и логистическое распределение с функцией распределения

$$G(x) = \frac{1}{1 + e^{-x}},$$

т. е. с плотностью $g(x) = e^{-x} / (1 + e^{-x})^2$.

*) См. Sierpinski, Sur les fonctions convexes mesurables, Fundamenta Mathematicae I (1920), 125—129.

Пример 2. Рассмотрим соответствующую задачу для масштабного параметра. Пусть $f_{\theta}(x) = \theta^{-1} h(x/\theta)$, где h — четная функция. Не ограничивая общности, мы можем иметь дело только с неотрицательными x , так как абсолютные значения $|X_1|, \dots, |X_n|$ образуют систему достаточных статистик для θ . Если $Y_i = \log X_i$ и $\eta = \log \theta$, то плотность Y_i имеет вид $h(e^{y-\eta}) e^{y-\eta}$. Если $h(x) > 0$ при $x \geq 0$, то в соответствии с примером 1 необходимым и достаточным условием, чтобы $f_{\theta'}(x)/f_{\theta}(x)$ было неубывающей функцией x при всех $\theta < \theta'$, является выпуклость функции $-\log [e^y h(e^y)]$ или, что то же самое, выпуклость $-\log h(e^y)$. Помимо нормального и симметричного экспоненциального распределений (для которых совместные плотности образуют экспоненциальные семейства), примером, в котором это свойство выполняется, может служить *распределение Коши* с плотностью

$$h(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$

Так как выпуклость $-\log h(y)$ влечет выпуклость $-\log h(e^y)$, то для четной h из монотонности отношения правдоподобия у семейства $h(x-\theta)$ вытекает его монотонность у $h(x/\theta)$. Когда распределение h нормальное или симметричное экспоненциальное, это свойство $h(x/\theta)$ следует уже из примера 1. Однако монотонность отношения правдоподобия для масштабного параметра не влечет аналогичного свойства для соответствующего семейства с параметром сдвига. Это иллюстрируется примером распределения Коши. Поэтому рассматриваемое условие в применении к параметру сдвига более ограничительно, чем в применении к масштабному параметру.

Основная трудность в применении теоремы 1 к конкретным проблемам состоит в необходимости знать, или по крайней мере уметь правильно угадывать, пару (λ, λ') наименее благоприятных распределений. Руководством для отыскания этих распределений иногда могут служить соображения инвариантности. Если существует группа G преобразований X такая, что индуцированная группа $\bar{G}$ оставляет инвариантными ω и ω' , то задача симметрична по отношению к тем θ , которые могут быть переведены друг в друга преобразованиями группы $\bar{G}$. В этом случае кажется вероятным, что, коль скоро λ и λ' не обнаруживают той же симметрии, они делают задачу статистика более легкой и, следовательно, не будут наименее благоприятными.

Пример 3. В задаче попарного сравнения, рассмотренной в примере 7 главы 6, результаты наблюдений X_i ($i = 1, \dots, n$) представляют собой независимые случайные величины, принимающие значения 1 и 0 с вероятностями p_i и $q_i = 1 - p_i$. Гипотезе H , которую необходимо проверить, соответствует множество ω : $\max p_i \leq 1/2$. Принимаются во внимание только альтернативы, для которых $p_i \geq 1/2$ при всех i . В качестве ω' мы выберем подмножество тех альтернатив, для которых $p_i \geq 1/2 + \delta$. Следовало бы ожидать, что λ приписывает вероятность 1 точке $p_1 = \dots = p_n = 1/2$, а λ' приписывает положительные вероятности только n точками $(p_1, \dots, p_n)$, у которых $n-1$ координат равны $1/2$, а оставшиеся равны $1/2 + \delta$. Следствие симметрии по n переменным представляется вероятным, что λ' должно приписывать каждой из этих точек вероятность $1/n$. При таком определении критерий $\Phi_{\lambda, \lambda'}$

отвергает гипотезу при

$$\sum_{i=1}^n \left(\frac{\frac{1}{2} + \delta}{\frac{1}{2} - \delta} \right)^{x_i} > C.$$

Последнее эквивалентно

$$\sum_{i=1}^n x_i > C,$$

а эта критическая область, как было показано ранее, соответствует РНМ инвариантному критерию для данной задачи. Так как критическая функция $\Phi_{\lambda, \lambda'}(x_1, \dots, x_n)$ не убывает по каждому из своих аргументов, то по лемме 2 главы 3 из неравенств $p_i \leq p'_i$ при всех $i = 1, \dots, n$ следует

$$E_{p_1, \dots, p_n} \Phi_{\lambda, \lambda'}(X_1, \dots, X_n) \leq E_{p'_1, \dots, p'_n} \Phi_{\lambda, \lambda'}(X_1, \dots, X_n),$$

т. е. выполняются условия теоремы 1.

Пример 4. Пусть $X = (X_1, \dots, X_n)$ выборка из $N(\xi, \sigma^2)$. Рассмотрим задачу проверки гипотезы $H: \sigma = \sigma_0$ при альтернативах $\omega': \sigma \leq \sigma_1$ или $\sigma \geq \sigma_2 (\sigma_1 < \sigma_0 < \sigma_2)$. Задача остается инвариантной при преобразованиях $X'_i = X_i + c$, которые индуцируют в пространстве параметров группу $\bar{G}$ преобразований $\xi' = \xi + c$, $\sigma' = \sigma$. Поэтому следовало бы ожидать, что наименее благоприятное распределение λ на прямой $\omega: -\infty < \xi < \infty$, $\sigma = \sigma_0$ инвариантно относительно $\bar{G}$. Отсюда вытекает, что λ приписывает любому интервалу меру, пропорциональную его длине, т. е. λ не может быть распределением вероятностей, и теорема 1 непосредственно не применима. Эту трудность можно преодолеть, аппроксимируя λ последовательностью распределений вероятностей, в разбираемом случае, например, последовательностью нормальных распределений $N(0, k)$, $k = 1, 2, \dots$

Оказывается, что в той специальной задаче, которую мы сейчас рассматриваем, существуют наименее благоприятные распределения λ и λ' , являющиеся настоящими распределениями вероятностей и, следовательно, не обладающие инвариантностью. Эти распределения можно найти путем исследования соответствующей односторонней задачи из раздела 9 главы 3, что и производится ниже. На множестве ω , где меняется только ξ , в качестве распределения λ для ξ принимается нормальное распределение с произвольным средним ξ_1 и дисперсией $(\sigma_2^2 - \sigma_0^2)/n$. Распределение λ' следует сконцентрировать на двух прямых $\sigma = \sigma_1$ и $\sigma = \sigma_2$ (ξ, σ)-плоскости. Положим $\lambda' = p\lambda'_1 + q\lambda'_2$, где λ'_1 нормально со средним ξ_1 и дисперсией $(\sigma_2^2 - \sigma_1^2)/n$, в то время как λ'_2 приписывает вероятность 1 точке (ξ_1, σ_2) . Вычисления, аналогичные проведенным в разделе 9 главы 3, показывают, что область принятия гипотезы определяется неравенством

$$\begin{aligned} & \frac{\frac{p}{\sigma_1^{n-1}\sigma_2} \exp \left[\frac{-1}{2\sigma_1^2} \sum (x_i - \bar{x})^2 - \frac{n}{2\sigma_2^2} (\bar{x} - \xi_1)^2 \right]}{\frac{1}{\sigma_0^{n-1}\sigma_2} \exp \left[\frac{-1}{2\sigma_0^2} \sum (x_i - \bar{x})^2 - \frac{n}{2\sigma_2^2} (\bar{x} - \xi_1)^2 \right]} + \\ & + \frac{\frac{q}{\sigma_2^n} \exp \left[\frac{-1}{2\sigma_2^2} \left[\sum (x_i - \bar{x})^2 + n(\bar{x} - \xi_1)^2 \right] \right]}{\frac{1}{\sigma_0^{n-1}\sigma_2} \exp \left[\frac{-1}{2\sigma_0^2} \sum (x_i - \bar{x})^2 - \frac{n}{2\sigma_2^2} (\bar{x} - \xi_1)^2 \right]} < C, \end{aligned}$$

которое эквивалентно неравенству

$$C_1 \leq \sum (x_i - \bar{x})^2 \leq C_2.$$

Вероятность последнего неравенства не зависит от ξ . Следовательно, C_1 и C_2 могут быть определены таким образом, что вероятность принятия гипотезы будет равна $1 - \alpha$ при $\sigma = \sigma_0$ и будет иметь равные значения при $\sigma = \sigma_1$ и $\sigma = \sigma_2$.

Из раздела 7 главы 3 выводим, что существуют такие p и C , которые приводят к указанным C_1 и C_2 , и что вышеприведенный критерий удовлетворяет условиям следствия с $\omega_0 = \omega$ и с ω'_0 , состоящим из двух прямых линий: $\sigma = \sigma_1$ и $\sigma = \sigma_2$.

3. Максиминные критерии и инвариантность

Когда проблема проверки Ω_H при альтернативе Ω_K остается инвариантной относительно преобразований некоторой группы, представляется разумным ожидать, что существует инвариантная пара наименее благоприятных распределений (или, по меньшей мере, последовательности распределений, которые в пределе, в некотором смысле, являются наименее благоприятными и инвариантными) и что, вместе с тем, существует инвариантный максиминный критерий. Это приводит к возможности обойти трудности подхода, намеченного в предыдущих параграфах. Если бы мы смогли доказать, что для инвариантной проблемы всегда существует инвариантный критерий, который максимизирует минимум мощности на Ω_K , то мы были бы вправе ограничиться только инвариантными критериями. В частности, в этом случае РНМ несмещенный критерий автоматически обладал бы желательным максиминным свойством. Эти догадки оказываются правильными в важном классе задач, хотя они, к сожалению, не переносятся на общий случай. Чтобы найти соответствующие условия, удобно сначала разделить статистический и теоретико-групповой аспекты проблемы посредством следующей леммы.

Лемма 2. Пусть $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ — доминированное семейство распределений на $(\mathcal{X}, \mathcal{A})$ и пусть G — группа преобразований на $(\mathcal{X}, \mathcal{A})$ такая, что индуцированная группа $\bar{G}$ оставляет инвариантными подмножества Ω_H и Ω_K пространства Ω . Допустим, что для любой критической функции ϕ существует (почти) инвариантная критическая функция ψ , для которой

$$\inf_{\bar{G}} E_{\theta\theta\phi}(X) \leq E_\theta\psi(X) \leq \sup_{\bar{G}} E_{\theta\theta\phi}(X) \quad (10)$$

при всех $\theta \in \Omega$. Тогда, если существует критерий ϕ_0 уровня α , максимизирующий $\inf_{\Omega_K} E_{\theta\phi}(X)$, то существует и (почти) инвариантный критерий с этим свойством.

Доказательство. Пусть $\inf_{\Omega_K} E_{\theta\phi_0}(X) = \beta$ и пусть ψ_0 — (почти) инвариантный критерий, такой, что (10) выполняется с

$\varphi = \varphi_0, \psi = \psi_0$. Тогда

$$E_{\theta}\psi_0(X) \leq \sup_{\bar{G}} E_{g\theta}\varphi_0(X) \leq \alpha \quad \text{при всех } \theta \in \Omega_H$$

и

$$E_{\theta}\psi_0(X) \geq \inf_{\bar{G}} E_{g\theta}\varphi_0(X) \geq \beta \quad \text{при всех } \theta \in \Omega_K,$$

что и требовалось доказать.

Чтобы установить условия, при которых существует инвариантный или почти инвариантный критерий ψ , удовлетворяющий (10), рассмотрим сначала простейший случай конечной группы G . Пусть $G = \{g_1, \dots, g_N\}$. Тогда, если определить ψ равенством

$$\psi(x) = \frac{1}{N} \sum_{i=1}^N \varphi(g_i x), \quad (11)$$

то ясно, что ψ будет критической функцией, и притом инвариантной относительно G . Она также удовлетворяет (10), так как $E_{\theta}\varphi(gX) = E_{g\theta}\varphi(X)$ и $E_{\theta}\psi(X)$ является средним нескольких величин, минимальная и максимальная из которых совпадают соответственно с первым и последним членами в (10).

Конечный случай иллюстрируется примером 3. В нем проблема остается инвариантной относительно $n!$ перестановок переменных $(X_1, \dots, X_n)$. Лемма 2 применима и показывает, что существует инвариантный критерий, максимизирующий $\inf_{\Omega_K} E_{\theta}\varphi(X)$. Поэтому, в частности, РНМ инвариантный критерий,

полученный в примере 7 главы 6, обладает требуемым максиминным свойством и, следовательно, является решением проблемы.

Определение (11) приводит к мысли о возможности и в других случаях получить $\psi(x)$ усреднением значений $\varphi(gx)$ по отношению к надлежаще выбранному распределению вероятностей на группе G . Посмотрим, каким условиям должно было бы удовлетворять подобное распределение. Пусть $\mathcal{R}$ σ -поле подмножеств G и ν — распределение вероятностей на $(G, \mathcal{R})$. Отвлекаясь на время от проблем измеримости, определим ψ соотношением

$$\psi(x) = \int \varphi(gx) d\nu(g). \quad (12)$$

Тогда $0 \leq \psi \leq 1$ и (10) имеет место, как можно заключить из теоремы Фубини (теорема 3 главы 2), применяя ее к интегралу от ψ по распределению P_{θ} . Для любого $g_0 \in G$

$$\psi(g_0 x) = \int \varphi(g g_0 x) d\nu(g) = \int \varphi(hx) d\nu^*(h),$$

где $h = gg_0$ и где ν^* — мера, определенная формулой

$$\nu^*(B) = \nu(Bg_0^{-1}) \quad \text{для всех } B \in \mathcal{B},$$

т. е. мера, в которую ν переводится трансформацией $h = gg_0$. Таким образом, ψ будет обладать желаемой инвариантностью, $\psi(g_0x) = \psi(x)$ при всех $g_0 \in G$, если ν — *правоинвариантна*, т. е. если

$$\nu(Bg) = \nu(B) \quad \text{при всех } B \in \mathcal{B}, g \in G. \quad (13)$$

Для справедливости приведенных выше рассуждений нужны следующие предположения измеримости:

(I) При любом $A \in \mathcal{A}$ множество пар (x, g) с $gx \in A$ является $(\mathcal{A} \times \mathcal{B})$ -измеримым (этим обеспечивается измеримость функции ψ , определяемой по (12));

(II) при любых $B \in \mathcal{B}, g \in G$ множество Bg принадлежит $\mathcal{B}$.

Пример 5. Пусть G — конечная группа с элементами $g_1, \dots, g_N$, $\mathcal{B}$ — класс всех подмножеств G и ν — вероятностная мера, приписывающая вероятность $1/N$ каждому из N элементов. В этом случае (13) выполняется и определение (12) сводится к (11).

Пример 6. Рассмотрим группу G ортогональных $(n \times n)$ -матриц Γ и определим групповое произведение $\Gamma_1 \Gamma_2$ как обычное произведение матриц. Каждую матрицу можно рассматривать как точку в n^2 -мерном евклидовом пространстве, координаты которой равны элементам матрицы. Тогда группа порождает некоторое множество в этом пространстве. В качестве σ -поля $\mathcal{B}$ мы возьмем борелевские подмножества G . Чтобы доказать существование на $(G, \mathcal{B})$ правоинвариантной вероятностной меры*), построим случайную ортогональную матрицу, распределение вероятностей которой удовлетворяет (13) и тем самым является искомой мерой. Свяжем с каждой невырожденной матрицей $x = (x_{ij})$ ортогональную матрицу $y = f(x)$, получаемую применением к n векторам-строкам $x_i = (x_{i1}, \dots, x_{in})$ матрицы x следующего процесса ортогонализации Грама — Шмидта. Пусть y_1 — единичный вектор в направлении x_1 , y_2 — единичный вектор в плоскости, натянутой на x_1 и x_2 , ортогональный к y_1 и образующий острый угол с x_2 , и т. д. Пусть $y = (y_{ij})$ — матрица, i -я строка которой равна y_i .

Допустим теперь, что случайные величины X_{ij} ($i, j = 1, \dots, n$) независимы и распределены по закону $N(0, 1)$. Обозначим X случайную матрицу (X_{ij}) и положим $Y = f(X)$. Покажем, что распределение случайной ортогональной матрицы Y удовлетворяет (13). С этой целью рассмотрим любую фиксированную ортогональную матрицу Γ и любое фиксированное множество $B \in \mathcal{B}$. Тогда $P\{Y \in B\Gamma\} = P\{Y\Gamma' \in B\}$, и из определения f видно, что $Y\Gamma' = f(X\Gamma')$. Так как n^2 элементов матрицы $X\Gamma'$ имеют то же самое совместное распределение, что и элементы матрицы X , то и матрицы $f(X\Gamma')$ и $f(X)$ имеют одно и то же распределение, что и требовалось доказать.

Примеры 5 и 6 достаточны для тех применений, которые будут исследованы здесь. Эти примеры являются простыми част-

*) Более детальное изложение вопросов, связанных с этой инвариантной мерой, дано в работе: James, Normal multivariate analysis and the orthogonal group, Ann. Math. Stat. 25 (1954), 40—75.

ными случаями по отношению к общим условиям существования инвариантных вероятностных мер, даваемых теорией меры Хаара *).

4. Теорема Ханта — Стейна

Инвариантные меры существуют (и с точностью до константы) для большого класса групп, но, к сожалению, часто бывают бесконечными и, следовательно, не могут быть взяты в качестве распределений вероятностей. Это явление сходно (и, по существу, связано) с отсутствием пары наименее благоприятных распределений в теореме 1. Трудность обычно можно преодолеть, рассматривая последовательность распределений, которая «в пределе» обладает требуемым свойством. Аналогично, мы обобщим теперь конструкцию φ как среднего по отношению к правоинвариантной мере, рассматривая последовательность распределений на G , которые при больших n приближенно правоинвариантны.

Пусть $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ — семейство распределений на евклидовом пространстве $(\mathcal{X}, \mathcal{A})$, доминированное σ -конечной мерой μ , и пусть G — группа преобразований $(\mathcal{X}, \mathcal{A})$ такая, что индуцированная группа $\bar{G}$ оставляет Ω инвариантным.

Теорема 2 (Ханта — Стейна). Пусть $\mathcal{B}$ — σ -поле подмножеств группы G , так что для любого $A \in \mathcal{A}$ множество пар (x, g) с $gx \in A$ принадлежит $\mathcal{A} \times \mathcal{B}$ и для любых $B \in \mathcal{B}$ и $g \in G$ $Bg \in \mathcal{B}$. Допустим, что существует последовательность распределений вероятностей ν_n на $(G, \mathcal{B})$, которые асимптотически правоинвариантны в том смысле, что для любых $g \in G$, $B \in \mathcal{B}$

$$\lim_{n \rightarrow \infty} |\nu_n(Bg) - \nu_n(B)| = 0. \quad (14)$$

Тогда, какова бы ни была критическая функция φ , существует критическая функция ψ , которая почти инвариантна и удовлетворяет (10).

Доказательство. Пусть

$$\psi_n(x) = \int \varphi(gx) d\nu_n(g).$$

Как и раньше, эта функция измерима и лежит между 0 и 1. По теореме о слабой компактности (теорема 3 в Дополнении) существует подпоследовательность $\{\psi_{n_i}\}$ и измеримая функция ψ ,

*) Мера Хаара рассматривается, например, в книгах: Montgomery and Zippin, Topological Transformation Groups, New York, Interscience Publishers, 1955, главы I, II, и П. Халмош, Теория меры, ИЛ, 1953, главы XI, XII.

лежащая между 0 и 1, для которой

$$\lim_{i \rightarrow \infty} \int \psi_{n_i} p \, d\mu = \int \psi p \, d\mu$$

для всех μ -интегрируемых функций p . В частности,

$$\lim_{i \rightarrow \infty} E_{\theta} \psi_{n_i}(X) = E_{\theta} \psi(X)$$

при всех $\theta \in \Omega$. По теореме Фубини

$$E_{\theta} \psi_{n_i}(X) = \int [E_{\theta} \varphi(gX)] \, d\nu_{n_i}(g) = \int E_{\bar{g}\theta} \varphi(X) \, d\nu_{n_i}(g),$$

так что

$$\inf_{\bar{G}} E_{\bar{g}\theta} \varphi(X) \leq E_{\theta} \psi_{n_i}(X) \leq \sup_{\bar{G}} E_{\bar{g}\theta} \varphi(X)$$

и ψ удовлетворяет (10).

Чтобы доказать почти инвариантность ψ , мы покажем теперь, что для всех x и g

$$\psi_{n_i}(gx) - \psi_{n_i}(x) \rightarrow 0. \quad (15)$$

По теореме об ограниченной сходимости (теорема 1 (II) главы 2) отсюда будет следовать, что

$$\int_A [\psi_{n_i}(gx) - \psi_{n_i}(x)] \, dP_{\theta}(x) \rightarrow 0$$

при всех $\theta \in \Omega$ и $A \in \mathcal{A}$, а следовательно, что $\psi(gx) = \psi(x)$ ($\mathcal{P}$ — почти всюду), что и требовалось доказать.

Фиксируем x и, выбрав целое m , разобьем G на попарно непересекающиеся множества

$$B_k = \left\{ h \in G: a_k < \varphi(hx) \leq a_k + \frac{1}{m} \right\}, \quad k = 0, \dots, m,$$

где $a_k = (k-1)/m$. В частности, B_0 — это множество $\{h \in G: \varphi(hx) = 0\}$. Из определения множеств B_k видно, что

$$\begin{aligned} \sum_{k=0}^m a_k \nu_{n_i}(B_k) &\leq \sum_{k=0}^m \int_{B_k} \varphi(hx) \, d\nu_{n_i}(h) \leq \\ &\leq \sum_{k=0}^m \left(a_k + \frac{1}{m}\right) \nu_{n_i}(B_k) \leq \sum_{k=0}^m a_k \nu_{n_i}(B_k) + \frac{1}{m} \end{aligned}$$

и, аналогично,

$$\left| \sum_{k=0}^m \int_{B_k g^{-1}} \varphi(hgx) \, d\nu_{n_i}(h) - \sum_{k=0}^m a_k \nu_{n_i}(B_k g^{-1}) \right| \leq \frac{1}{m},$$

откуда следует, что

$$|\psi_{n_i}(gx) - \psi_{n_i}(x)| \leq \sum |a_k| \cdot |\nu_{n_i}(B_k g^{-1}) - \nu_{n_i}(B_k)| + \frac{2}{m}.$$

В силу (14) первый член в правой части стремится к нулю при i , стремящемся к бесконечности, и этим доказательство завершается.

Если существует правоинвариантная мера ν на G и последовательность G_n подмножеств G с $G_n \subseteq G_{n+1}$, $\bigcup G_n = G$ и $\nu(G_n) = c_n < \infty$, то естественно пытаться выбрать в качестве вероятностных мер ν_n теоремы 2 меры ν/c_n , рассматриваемые на G_n . В указанном ниже примере это приводит к успеху. С другой стороны, имеются случаи, когда существуют подобные последовательности G_n , но нет инвариантных критериев, удовлетворяющих (10), и нет, следовательно, последовательности ν_n , для которой выполнялось бы (14).

Пример 7. Пусть $x = (x_1, \dots, x_n)$, $\mathcal{A}$ — класс борелевских множеств в n -мерном пространстве и G — группа сдвигов $(x_1 + g, \dots, x_n + g)$, $-\infty < g < \infty$. Элементы G могут быть представлены действительными числами, и групповое произведение оказывается в этом случае обычной суммой $g + g'$. Пусть $\mathcal{B}$ — класс борелевских множеств на прямой. Тогда предположения измеримости, входящие в теорему 2, выполнены. Пусть ν — мера Лебега, которая, очевидно, инвариантна относительно G . Определим ν_n как распределение, равномерное в интервале $I(-n, n) = \{g: -n \leq g \leq n\}$. Тогда для всех $B \in \mathcal{B}$, $g \in G$

$$|\nu_n(B) - \nu_n(Bg)| = \frac{1}{2n} |\nu[B \cap I(-n, n)] - \nu[B \cap I(-n - g, n - g)]| \leq \frac{|g|}{2n},$$

так что (14) выполняется.

Эти рассуждения применимы также и к группе изменений масштаба $(ax_1, \dots, ax_n)$, $0 < a < \infty$, которая может быть превращена в группу сдвигов переходом к логарифмам.

Редукция максиминной проблемы может быть проведена шаг за шагом в предположениях теоремы 2 главы 6. Допустим, что проблема остается инвариантной по отношению к двум группам D и E . Обозначим $y = s(x)$ максимальный инвариант относительно D и буквой E^* — группу, определенную в теореме 2 главы 6 (которую E индуцирует в пространстве y). Если D и E^* удовлетворяют условиям теоремы Ханта — Стейна, то, во-первых, существует максиминный критерий, зависящий только от $y = s(x)$, во-вторых, существует максиминный критерий, зависящий только от максимального инварианта $z = t(y)$ относительно E^* .

Пример 8. Рассмотрим одномерную линейную гипотезу, в канонической форме которой $Y_1, \dots, Y_n$ независимы и распределены по $N(\eta_i, \sigma^2)$, причем дано, что $\eta_{s+1} = \dots = \eta_n = 0$. Проверке подлежит гипотеза $\eta_1 = \dots = \eta_r = 0$. В разделе 1 главы 7 было показано, что эта проблема инвариантна относительно некоторых групп преобразований и что по отношению к этим группам существует РНМ инвариантный критерий. Указанные группы

являются группами ортогональных преобразований или группами сдвигов типа, рассмотренного в примере 7, или группами изменений масштаба. Так как каждая из этих групп удовлетворяет условиям теоремы Ханта — Стейна и поскольку они оставляют инвариантной проблему максимизации минимума мощности на множестве альтернатив

$$\sum_{i=1}^r \frac{\eta_i^2}{\sigma^2} \geq \psi_1^2 \quad (\psi_1 > 0), \quad (16)$$

то РНМ инвариантный критерий главы 7 является решением также и этой максиминной проблемы. Более общим образом критерий, который является РНМ инвариантным (относительно той же самой группы) для проверки

$$\sum_{i=1}^r \frac{\eta_i^2}{\sigma^2} \leq \psi_0^2$$

(задача 4 главы 7), максимизирует минимальную мощность на альтернативах (16) для $\psi_0 < \psi_1$.

Пример 9 (Стейн). Пусть G — группа всех невырожденных линейных преобразований p -мерного пространства. При $p > 1$ она удовлетворяет условиям теоремы 2, как можно увидеть из нижеследующей задачи, которая инвариантна относительно G , для которой РНМ инвариантный критерий не максимизирует минимум мощности. Обобщая пример 10 главы 6, допустим, что $X = (X_1, \dots, X_p)$ и $Y = (Y_1, \dots, Y_p)$ независимы и имеют p -мерное нормальное распределение с нулевыми средними и невырожденными матрицами ковариаций $E(X_i X_j) = \sigma_{ij}$ и $E(Y_i Y_j) = \Delta \sigma_{ij}$. Пусть проверке подлежит гипотеза $H: \Delta \leq \Delta_0$ при альтернативе $\Delta \geq \Delta_1$ ($\Delta_0 < \Delta_1$), причем σ_{ij} предполагаются неизвестными.

Эта проблема остается инвариантной, если оба вектора подвергаются любому (одному и тому же) невырожденному преобразованию, и так как с вероятностью единица эта группа транзитивна на выборочном пространстве, то РНМ инвариантный критерий тривиален: $\phi(x, y) \equiv \alpha$. Следовательно, максиминная мощность при альтернативах $\Delta \geq \Delta_1$, которая может быть достигнута при использовании инвариантных критериев, равна α . С другой стороны, критерий с критической областью $Y_1^2/X_1^2 > C$ имеет строго возрастающую функцию мощности $\beta(\Delta)$, минимум которой на множестве альтернатив $\Delta \geq \Delta_1$ равен $\beta(\Delta_1) > (\beta\Delta_0) = \alpha$.

Интересная особенность теоремы 2 состоит в том, что ее предположения касаются только группы G , а не распределений P_θ . Если эти предположения выполняются для некоторой группы G , то, как можно вывести из (10) (путем, указанным в доказательстве леммы 2), для любой проблемы, инвариантной относительно G и допускающей РНМ инвариантный критерий, этот критерий максимизирует минимум мощности в любом инвариантном классе альтернатив. Наоборот, допустим, что мы установили, что в некоторой отдельной задаче РНМ инвариантный относительно G критерий не максимизирует минимум мощности (как это имело место в группе линейных преобразований примера 9). Тогда предпосылки теоремы 2 не могут быть выполнены. Однако это не исключает той возможности, что

в другой проблеме, инвариантной относительно G , РНМ инвариантный критерий будет максимизировать минимум мощности. Произойдет это или нет — зависит уже не только от группы, но и от семейства распределений.

Рассмотрим, в частности, задачу проверки гипотезы $H: \xi_1 = \dots = \xi_p = 0$ на основе выборки $(X_{\alpha 1}, \dots, X_{\alpha p})$, $\alpha = 1, \dots, n$, из p -мерного нормального распределения со средними $E(X_{\alpha i}) = \xi_i$ и общей матрицей ковариаций $(\sigma_{ij}) = (a_{ij})^{-1}$. Как мы показали в разделе 10 главы 7, эта проблема инвариантна относительно нескольких групп, в число которых входит группа всех невырожденных линейных преобразований p -мерного пространства. Кроме того, было установлено существование РНМ инвариантного критерия. Инвариантным относительно этих групп является класс альтернатив

$$\frac{\sum \sum a_{ij} \xi_i \xi_j}{\sigma^2} \geq \Psi_1^2. \quad (17)$$

Здесь теорема 2 неприменима, и вопрос о том, будет ли РНМ инвариантный критерий максимизировать минимум мощности при альтернативах (17), остается открытым.

5. Наиболее строгие критерии

Одна из практических трудностей, связанных с критериями, максимизирующими минимум мощности на классе Ω_K , состоит в выборе подходящего Ω_K . Если информация, на основе которой можно было бы провести такой выбор, отсутствует и если соображения инвариантности не приводят к естественному определению, то часто разумное определение может быть получено в терминах мощности, достигаемой при различных альтернативах. *Огибающая функция мощности* β_α^* была определена в задаче 15 главы 6 формулой

$$\beta_\alpha^*(\theta) = \sup \beta_\phi(\theta), \quad (18)$$

где β_ϕ обозначает мощность критерия ϕ и где супремум берется по всем критериям проверки H с уровнем α . Таким образом, $\beta_\alpha^*(\theta)$ является максимальной мощностью, которая может быть достигнута при уровне α и альтернативе θ (то, что супремум достигается, вытекает из теоремы 3 Дополнения при очень слабых ограничениях). Если

$$S_\Delta^* = \{\theta : \beta_\alpha^*(\theta) = \Delta\}$$

и $\theta_1 \in S_{\Delta_1}^*$, $\theta_2 \in S_{\Delta_2}^*$ — две альтернативы, то θ_1 можно полагать более близкой, равноудаленной или более удаленной от H , чем θ_2 , в соответствии с тем, будет ли $\Delta_1 <, =$ или $> \Delta_2$.

С идеей измерения расстояния от альтернативы до H в терминах доступной информации мы уже сталкивались ранее. Если, например, $X_1, \dots, X_n$ образуют выборку из $N(\xi, \sigma^2)$, то проблема проверки $H: \xi \leq 0$ может рассматриваться (см. раздел 2 главы 5) с альтернативами ξ , измеренными как в абсолютных единицах, так и в σ -единицах. Последний случай соответствует сделанному выше предположению, так как из соображений инвариантности (задача 15 главы 6) следует, что $\beta_\alpha^*(\xi, \sigma)$ постоянно на линиях $\xi/\sigma = \text{const}$.

Фиксируя значение Δ и принимая в качестве Ω_K класс альтернатив θ , для которых $\beta_\alpha^*(\theta) \geq \Delta$, мы можем определить критерий, который максимизирует минимум мощности на Ω_K . Другая возможность, которая исключает необходимость выбора значения Δ , состоит в том, что для любого критерия φ рассматривается разность $\beta_\alpha^*(\theta) - \beta_\varphi(\theta)$. Эта разность показывает, насколько действительная мощность $\beta_\varphi(\theta)$ меньше, чем максимально достижимая. Критерий, который минимизирует выражение

$$\sup_{\Omega - \omega} [\beta_\alpha^*(\theta) - \beta_\varphi(\theta)], \quad (19)$$

называют *наиболее строгим* *). Таким образом, критерий является наиболее строгим, если он минимизирует максимальную «нехватку» мощности.

Обозначим φ_Δ критерий, который максимизирует минимальную мощность на S_Δ^* и, следовательно, минимизирует максимум разности между $\beta_\alpha^*(\theta)$ и $\beta_\varphi(\theta)$ на S_Δ^* . Если окажется, что φ_Δ не зависит от Δ , то φ_Δ будет наиболее строгим. Это замечание делает возможным применение результатов предыдущего раздела к отысканию наиболее строгих критериев. Допустим, что проблема проверки $H: \theta \in \omega$ при альтернативах $\theta \in \Omega - \omega$ остается инвариантной относительно группы G , что существует РНМ почти инвариантный относительно G критерий φ_0 и что выполняются предположения теоремы 2. Так как $\beta_\alpha^*(\theta)$ и, следовательно, множество S_Δ^* инвариантны относительно G (задача 15 главы 6), то мы заключаем, что φ_0 максимизирует минимум мощности на S_Δ^* при каждом Δ . Стало быть, критерий φ_0 — наиболее строгий.

Как пример применения метода, рассмотрим проблему проверки $H: p_1, \dots, p_n \leq 1/2$ при альтернативе $K: p_i > 1/2$ при всех i , где p_i — вероятность успеха в i -м из последовательности n независимых испытаний. В зависимости от того, заканчивается i -е испытание успехом или неудачей, X_i равно 1 или 0. В этом случае проблема остается инвариантной по отношению к перестановкам величин X_i , и РНМ инвариантный критерий имеет кри-

*) В оригинале — most stringent — термин Вальда, см. Вальд (1942) на стр. 389.

тическую область $\Sigma X_i > C$ (пример 7 главы 6). Из сделанных выше замечаний мы выводим, что этот критерий — наиболее строгий.

Другим примером служит общая одномерная линейная гипотеза. В этом случае из рассуждений примера 8 вытекает, что стандартный критерий проверки $H: \eta_1 = \dots = \eta_r = 0$ или

$H': \sum_{i=1}^r \eta_i^2 / \sigma_2 \leq \psi_0^2$ является наиболее строгим.

Во многих задачах, к которым неприменимы соображения инвариантности, наиболее строгие критерии до сих пор не найдены. Ниже указывается один класс задач, где эти критерии находятся непосредственно. Пусть распределения X образуют однопараметрическое экспоненциальное семейство с плотностями, определяемыми формулой (12) главы 3. Рассмотрим гипотезу $H: \theta = \theta_0$. Тогда, в соответствии с тем, будут ли $\theta > \theta_0$ или $\theta < \theta_0$, огибающая функция мощности $\beta_\alpha^*(\theta)$ является мощностью РНМ одностороннего критерия для проверки H при альтернативах $\theta > \theta_0$ или $\theta < \theta_0$. Предположим, что существует двусторонний критерий φ_0 , задаваемый формулой (3) главы 4 и такой, что

$$\sup_{\theta < \theta_0} [\beta_\alpha^*(\theta) - \beta_{\varphi_0}(\theta)] = \sup_{\theta > \theta_0} [\beta_\alpha^*(\theta) - \beta_{\varphi_0}(\theta)] \quad (20)$$

и что супремум достигается в обеих частях равенства, скажем, в точках $\theta_1 < \theta_2 < \theta_2$. Если $\beta_{\varphi_0}(\theta_i) = \beta_i$, $i = 1, 2$, то применение фундаментальной леммы (теорема 5 (III) главы 3) к точкам $\theta_1, \theta_2, \theta_0$ показывает, что среди всех критериев φ с $\beta_\varphi(\theta_1) \geq \beta_1$ и $\beta_\varphi(\theta_2) \geq \beta_2$ только φ_0 удовлетворяет условию $\beta_\varphi(\theta_0) \leq \alpha$. Для любого другого критерия уровня α имеем: или $\beta_\varphi(\theta_1) < \beta_1$, или $\beta_\varphi(\theta_2) < \beta_2$, т. е. мы видим, что φ_0 является единственным наиболее строгим критерием. Существование критерия, удовлетворяющего (20), может быть выведено с помощью соображений непрерывности (по отношению к изменению констант C_i и γ_i , которые определяют границу критерия (3) главы 4) из того факта, что для РНМ одностороннего критерия при альтернативах $\theta > \theta_0$ правая часть (20) равна нулю, а левая положительна, в то время как для другого одностороннего критерия имеет место обратное соотношение.

6. Задачи

К разделу 1

1. *Существование максиминных критериев.* Пусть $(\mathcal{X}, \mathcal{A})$ — евклидово пространство выборок, и пусть распределения P_θ , $\theta \in \Omega$, доминированы σ -конечной мерой на $(\mathcal{X}, \mathcal{A})$. Для любых взаимно непересекающихся подмножеств $\Omega_n, \Omega_k \in \Omega$ существует критерий уровня α , максимизирующий (2).

[Пусть $\beta = \sup_{\Omega_K} [\inf_{\Omega_K} E_{\theta\Phi}(X)]$, где верхняя грань берется по всем критериям уровня α гипотезы $H: \theta \in \Omega_H$. Пусть Φ_n — последовательность критериев уровня α таких, что величина $\inf_{\Omega_K} E_{\theta\Phi_n}(X)$ стремится к β . Если Φ_{n_i} — подпоследовательность и Φ — критерий (гарантированный теоремой 3 дополнения) такие, что $E_{\theta\Phi_{n_i}}(X)$ стремится к $E_{\theta\Phi}(X)$ для всех $\theta \in \Omega$, то Φ является критерием уровня α и $\inf_{\Omega_K} E_{\theta\Phi}(X) = \beta$.]

2. *Локально наиболее мощные критерии.* Обозначим d меру расхождения между альтернативой θ и данной гипотезой H . Критерий Φ_0 уровня α называется *локально наиболее мощным* (ЛНМ), если для любого другого критерия Φ уровня α существует Δ такое, что

$$\beta_{\Phi_0}(\theta) \geq \beta_{\Phi}(\theta) \quad \text{при всех } \theta \text{ с } 0 < d(\theta) < \Delta. \quad (21)$$

Допустим, что θ действительно и функция мощности каждого критерия непрерывно дифференцируема в точке θ_0 .

(I) Тогда ЛНМ критерий для проверки гипотезы $H: \theta = \theta_0$ при альтернативе $\theta > \theta_0$ существует и определяется тем, что среди всех критериев уровня α для проверки H он максимизирует $\beta'(\theta_0)$.

(II) ЛНМ критерий локально максимизирует минимум мощности при условии, что его функция мощности «удалена» от α для каждого множества альтернатив, которые «удалены» от H .

(III) Пусть $X_1, \dots, X_n$ — выборка из распределения Коши с неизвестным параметром сдвига θ , так что совместная плотность X_k равна $\pi^{-n} \prod_{i=1}^n [1 + (x_i - \theta)^2]^{-1}$. ЛНМ критерий для проверки гипотезы $\theta = 0$ при альтернативе $\theta > 0$, имеющий уровень $\alpha < 1/2$, не является несмещенным и, следовательно, не максимизирует локально минимум мощности.

[(III) Существует M столь большое, что общая точка с $x_i \geq M$ для всех $i = 1, \dots, n$ лежит в области принятия ЛНМ критерия. Следовательно, мощность критерия стремится к нулю при θ , стремящемся к бесконечности.]

3. Критерий Φ_0 называется ЛНМ *несмещенным*, если он является несмещенным и если для любого другого несмещенного критерия Φ уровня α существует Δ такое, что выполняется (21). Предположим, что θ действительно и что функция мощности каждого критерия дважды непрерывно дифференцируема в точке θ_0 . Тогда ЛНМ несмещенный критерий уровня α для $H: \theta = \theta_0$ при альтернативе $\theta \neq \theta_0$ существует и определяется фактически тем, что максимизирует $\beta''(\theta_0)$ среди всех несмещенных критериев уровня α для H .

К разделу 2

4. Пусть распределение X зависит от параметров $(\theta, \Phi) = (\theta_1, \dots, \theta_r, \Phi_1, \dots, \Phi_s)$. Критерий для $H: \theta = \theta^0$ называется *локально строго несмещенным*, если для каждого Φ выполняется: (а) $\beta_{\Phi}(\theta^0, \Phi) = \alpha$, (б) существует θ -окрестность точки θ^0 , в которой $\beta_{\Phi}(\theta, \Phi) > \alpha$ при $\theta \neq \theta^0$.

(I) Допустим, что первая и вторая производные

$$\beta_{\Phi}^i(\Phi) = \frac{\partial}{\partial \theta_i} \beta_{\Phi}(\theta, \Phi) \Big|_{\theta_0} \quad \text{и} \quad \beta_{\Phi}^{ij}(\Phi) = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \beta_{\Phi}(\theta, \Phi) \Big|_{\theta_0}$$

существуют при всех критических функциях Φ и всех θ . Тогда, для того чтобы критерий Φ был локально строго несмещенным, необходимо и достаточно, чтобы $\beta_{\Phi}^i(\Phi) = 0$ при всех i и Φ и матрица $(\beta_{\Phi}^{ij}(\Phi))$ являлась положительно определенной при всех Φ .

(II) Критерий для H называется критерием типа E (типа D , если $s = 0$, так что не существует «мешающих» параметров), если он является локально строго несмещенным и в классе всех критериев с этим свойством максимизирует определитель $|\beta_{\Phi}^{ij}|$. (Этот определитель при перечисленных условиях оказывается равным гауссовской кривизне поверхности мощности в точке θ^0 .) Тогда критерий Φ_0 , заданный формулой (7) главы 7, для проверки общей линейной одномерной гипотезы (3) главы 7 является критерием типа E .

[(II) В соответствии с задачей 5 главы 7 при $\theta = (\eta_1, \dots, \eta_r)$ и $\Phi = (\eta_{r+1}, \dots, \eta_s, \sigma)$ критерий Φ_0 максимизирует поверхностный интеграл

$$\int_S [\beta_{\Phi}(\eta, \sigma^2) - \alpha] dA,$$

где $S = \left\{ (\eta_1, \dots, \eta_r) : \sum_{i=1}^r \eta_i^2 = \rho^2 \sigma^2 \right\}$, в классе всех подобных (и, следовательно, всех локально несмещенных) критериев. Полагая ρ стремящимся к нулю и используя условия

$$\beta_{\Phi}^i(\theta) = 0, \quad \int_S \eta_i \eta_j dA = 0 \quad \text{при } i \neq j, \quad \int_S \eta_i^2 dA = K(\rho\sigma),$$

найдем, что Φ_0 максимизирует $\sum_{i=1}^r \beta_{\Phi}^{ii}(\eta, \sigma^2)$ среди всех локально несмещенных критериев. Для любой положительно определенной матрицы $|\beta_{\Phi}^{ij}| \leq \prod \beta_{\Phi}^{ii}$. Из этого следует, что для любого строго локально несмещенного *) критерия

$$|\beta_{\Phi}^{ij}| \leq \prod \beta_{\Phi}^{ii} \leq \left[\frac{\sum \beta_{\Phi}^{ii}}{r} \right]^r \leq \left[\frac{\sum \beta_{\Phi_0}^{ii}}{r} \right]^r = [\beta_{\Phi_0}^{11}]^r = |\beta_{\Phi_0}^{ij}|.$$

5. Пусть $Z_1, \dots, Z_N$ независимы и имеют одно и то же непрерывное распределение D , о котором известно лишь, что оно симметрично относительно некоторой (неизвестной) точки. Тогда критерий знаков максимизирует минимум мощности, если проверяется $H: D(0) = \frac{1}{2}$ при альтернативах $K:$

$$D(0) \leq q \left(q < \frac{1}{2} \right).$$

[Пара наименее благоприятных распределений приписывает вероятность 1 соответственно распределениям $F \in H$, $G \in K$ с плотностями

$$f(x) = \frac{1-2q}{2(1-q)} \left(\frac{q}{1-q} \right)^{[|x|]}; \quad g(x) = (1-2q) \left(\frac{q}{1-q} \right)^{[|x|]},$$

где для всех x , положительных, отрицательных и равных нулю, $[x]$ обозначает наибольшее целое число $\leq x$.]

6. Пусть $f_{\theta}(x) = \theta g(x) + (1-\theta)h(x)$ с $0 \leq \theta \leq 1$. Тогда $f_{\theta}(x)$ удовлетворяет предположениям леммы 1 при условии, что отношение $g(x)/h(x)$ является неубывающей функцией x .

7. Пусть $x = (x_1, \dots, x_n)$. Обозначим $g_{\theta}(x, \xi)$ семейство плотностей вероятности, зависящих от $\theta = (\theta_1, \dots, \theta_r)$ и действительного параметра ξ и

*) В оригинале «Strictly locally unbiased». Уместно обратить внимание читателя на то, что термин «строгий» (stringent) ранее употреблялся в ином смысле. — Прим. перев.

измеримых по совокупности x и ξ . При каждом θ обозначим $h_\theta(\xi)$ плотность вероятности по отношению к σ -конечной мере ν так, что существует $p_\theta(x) = \int g_\theta(x, \xi) h_\theta(\xi) d\nu(\xi)$. Мы будем говорить, что функция f двух аргументов $u = (u_1, \dots, u_r)$ и $v = (v_1, \dots, v_s)$ не убывает по (u, v) , если $f(u', v)/f(u, v) \leq f(u', v')/f(u, v')$ для всех (u, v) удовлетворяющих $u_i \leq u'_i$, $v_j \leq v'_j$ ($i = 1, \dots, r$; $j = 1, \dots, s$). Тогда $p_\theta(x)$ — неубывающая функция (x, θ) при условии, что произведение $g_\theta(x, \xi) h_\theta(\xi)$ обладает следующими свойствами: (а) не убывает по (x, θ) при каждом фиксированном ξ ; (б) не убывает по (ξ, θ) при каждом фиксированном x ; (в) не убывает по (x, ξ) при каждом фиксированном θ .

[Интерпретируя $g_\theta(x, \xi)$ как условную плотность x при данном ξ и $h_\theta(\xi)$ как априорную плотность ξ , обозначим $\rho(\xi)$ апостериорную плотность ξ при данном x , и пусть $\rho'(\xi)$ определяется аналогично, с заменой θ на θ' . Тогда то, что функция $p_\theta(x)$ является неубывающей по двум своим аргументам, эквивалентно соотношению

$$\int \frac{g_\theta(x', \xi)}{g_\theta(x, \xi)} \rho(\xi) d\nu(\xi) \leq \int \frac{g_{\theta'}(x', \xi)}{g_\theta(x, \xi)} \rho'(\xi) d\nu(\xi).$$

В силу (а), достаточно доказать, что

$$D = \int \left[\frac{g_\theta(x', \xi)}{g_\theta(x, \xi)} \right] [\rho'(\xi) - \rho(\xi)] d\nu(\xi) \geq 0.$$

Пусть $S_- = \{\xi: \rho'(\xi)/\rho(\xi) < 1\}$ и $S_+ = \{\xi: \rho'(\xi)/\rho(\xi) \geq 1\}$. По (II) множество S_- полностью лежит левее S_+ . Из (III) следует существование $a \leq b$ таких, что

$$D = a \int_{S_-} [\rho'(\xi) - \rho(\xi)] d\nu(\xi) + b \int_{S_+} [\rho'(\xi) - \rho(\xi)] d\nu(\xi)$$

и, следовательно, что $D = (b - a) \int_{S_+} [\rho'(\xi) - \rho(\xi)] d\nu(\xi) \geq 0$

8. (I) Пусть X имеет биномиальное распределение $b(p, n)$. Рассмотрим проверку гипотезы $H: p = p_0$ при альтернативах $\Omega_K: p/q \leq \frac{1}{2} p_0/q_0$ или $p/q \geq 2p_0/q_0$. При уровне $\alpha = 0,05$ определить наименьший размер выборки, для которого существует критерий с мощностью $\geq 0,8$ при альтернативах Ω_K , если $p_0 = 0,1; 0,2; 0,3; 0,4; 0,5$.

(II) Пусть $X_1, \dots, X_n$ независимы и распределены по закону $N(\xi, \sigma^2)$ и проверяется гипотеза $\sigma = 1$. Определить наименьший размер выборки, для которого существует критерий с уровнем $\alpha = 0,05$ и мощностью $\geq 0,9$ при альтернативах $\sigma^2 \leq \frac{1}{2}$ и $\sigma^2 \geq 2$. [См. задачу 5 главы 4.]

9. Симметричное экспоненциальное распределение. Пусть $X_1, \dots, X_n$ — выборка из симметричного экспоненциального распределения с плотностью $\frac{1}{2} e^{-|x-\theta|}$. ЛНМ критерием для проверки $\theta \leq 0$ против $\theta > 0$ является критерий знаков.

[Докажем это для случая, когда уровень α представляется в виде

$$\alpha = \sum_{k=0}^m C_n^k / 2^n, \text{ т. е. когда критерий знаков уровня } \alpha \text{ будет нерандомизирован-}$$

ным. Обозначим R_k ($k = 0, \dots, n$) подмножество пространства выборок, в котором k величин X положительны и $n - k$ — отрицательны. Допустим, далее, что $0 \leq k < l < n$, и выберем в R_k и R_l подмножества S_k и S_l соответственно такие, что $P_0(S_k) = P_0(S_l) \neq 0$. Тогда из рассмотрения $P_\theta(S_k)$ и $P_\theta(S_l)$ при малых θ следует существование Δ такого, что $P_\theta(S_k) < P_\theta(S_l)$ для $0 < \theta < \Delta$. Допустим теперь, что область отклонения нерандомизированным критерием гипотезы $\theta = 0$ при конкурирующей гипотезе $\theta > 0$ не совпадает ни с одним из множеств, для которых число положительных X больше константы. Тогда этот критерий может быть превращен в критерий знаков того же самого размера посредством конечного числа шагов, каждый из которых состоит в замене S_k на S_l с $k < l$, и поэтому увеличивает мощность при достаточно малых θ . Для рандомизированных критериев аргументация та же самая с ϕ_k, ϕ_l вместо S_k, S_l].

К разделу 4

10. Пусть $X = (X_1, \dots, X_p)$ и $Y = (Y_1, \dots, Y_p)$ независимы и имеют одно и то же p -мерное нормальное распределение с нулевым средним и матрицами ковариаций $E(X_i X_j) = \sigma_{ij}$ и $E(Y_i Y_j) = \Delta \sigma_{ij}$.

(I) Задача проверки $H: \Delta \leq \Delta_0$ остается инвариантной относительно группы G преобразований $X^* = XA$, $Y^* = YA$, где $A = (a_{ij})$ — любая невырожденная $(p \times p)$ -матрица с $a_{ij} = 0$ при $i > j$, и существует РНМ инвариантный относительно G критерий с областью отклонения $Y_1^2/X_1^2 > C$.

(II) Критерий с областью отклонения $Y_1^2/X_1^2 > C$ максимизирует минимум мощности для проверки $\Delta \leq \Delta_0$ при альтернативе $\Delta \geq \Delta_1$ ($\Delta_0 < \Delta_1$).

[(II) Тот факт, что теорема Ханта — Стейна применима к G , может быть доказан рассмотрением групп G_q преобразований $X'_q = \alpha_1 X_1 + \dots + \alpha_q X_q$ ($X'_i = X_i$ для $i = 1, \dots, q-1, q+1, \dots, p$) последовательно для $q = 1, \dots, p-1$. Здесь $\alpha_q \neq 0$, так как матрица A является невырожденной тогда и только тогда, когда $\alpha_{ii} \neq 0$ для всех i . Групповое произведение $(\gamma_1, \dots, \gamma_q)$ двух таких преобразований $(\alpha_1, \dots, \alpha_q)$ и $(\beta_1, \dots, \beta_q)$ задается соотношениями $\gamma_1 = \alpha_1 \beta_q + \beta_1$, $\gamma_2 = \alpha_2 \beta_q + \beta_2$, $\dots$, $\gamma_{q-1} = \alpha_{q-1} \beta_q + \beta_{q-1}$, $\gamma_q = \alpha_q \beta_q$, которые показывают, что G_q изоморфна группе изменений масштаба (умножение всех компонент на β_q) и сдвигов (сложение с $(\beta_1, \dots, \beta_{q-1}, 0)$). Теперь результат следует из теоремы Ханта — Стейна и примера 7, так как допущения теоремы Ханта — Стейна, за исключением легко проверяемых условий измеримости, связаны лишь с абстрактной структурой $(G, \mathcal{A})$, а не частной реализацией элементов группы G как преобразований некоторого пространства.]

11. Допустим, что проблема проверки гипотезы $\theta \in \Omega_H$ при альтернативе $\theta \in \Omega_K$ остается инвариантной относительно G , что существует РНМ почти инвариантный относительно G критерий ϕ_0 и что выполняются предположения теоремы 2. Тогда ϕ_0 максимизирует $\inf_{\Omega_K} [\omega(\theta) E_{\theta} \phi(X) + u(\theta)]$ при любых ве-

совых функциях $\omega(\theta) \geq 0$, $u(\theta)$, инвариантных относительно $\bar{G}$.

К разделу 5

12. *Существование наиболее строгих критериев.* В предположениях задачи 1 существует наиболее строгий критерий для проверки гипотезы $\theta \in \Omega_H$ при альтернативе $\theta \in \Omega - \Omega_H$.

13. Обозначим $\{\Omega_\Delta\}$ класс взаимно непересекающихся множеств альтернатив такой, что огибающая функция мощности постоянна на каждом Ω_Δ и $\cup \Omega_\Delta = \Omega - \Omega_H$. Пусть φ_Δ максимизирует минимум мощности на Ω_Δ . Если $\varphi_\Delta = \varphi$ не зависит от Δ , то φ является наиболее строгим критерием для проверки $\theta \in \Omega_H$.

14. Пусть $(Z_1, \dots, Z_N) = (X_1, \dots, X_m, Y_1, \dots, Y_n)$ распределены с совместной плотностью (56) главы 5. Рассмотрим проблему проверки $H: \eta = \xi$ при конкурирующей гипотезе, что X_i и Y_j независимы и нормально распределены с общей дисперсией σ^2 и средними $\eta \neq \xi$. Тогда критерий перестановок с областью отклонения $|\bar{Y} - \bar{X}| > C[T(Z)]$ (двусторонний вариант критерия (55) главы 5) является наиболее строгим.

[Применить задачу 13 к случаю, когда каждое из множеств Ω_Δ состоит из двух точек (ξ_1, η_1, σ) и (ξ_2, η_2, σ) таких, что

$$\xi_1 = \xi - \frac{n}{m+n} \delta, \quad \eta_1 = \xi + \frac{m}{m+n} \delta,$$

$$\xi_2 = \xi + \frac{n}{m+n} \delta, \quad \eta_2 = \xi - \frac{m}{m+n} \delta$$

для некоторых ξ и δ .]

7. Литературные ссылки

Идеи и результаты раздела 1, по существу, содержатся в развитой Вальдом для общих проблем решения теории минимакса. Изложение этой теории и некоторых ее применений дано в книге Вальда (1950). Материал разделов 3—5, включая, в частности, лемму 2, теорему 2 и пример 8, составляет основную часть неопубликованной работы Ханта и Стейна (1946).

Вальд (Wald A.)

(1942) On the principles of statistical inference, Notre Dame Math. Lectures № 1, Notre Dame, Ind. [Определение наиболее строгих критериев.]

(1950) Statistical Decision Functions, New York, John Wiley and Sons.

Веслер (Wesler O.)

(1959) Invariance theory and a modified minimax principle, Ann. Math. Stat., 30, 1—20. [Обобщение теоремы Ханта — Стейна.]

Вольфовиц (Wolfowitz J.)

(1949) The power of the classical tests associated with the normal distribution, Ann. Math. Stat., vol. 20, 540—551. [Доказано, что стандартные критерии для одномерной линейной гипотезы и для проверки отсутствия множественной корреляции являются наиболее строгими в классе всех подобных критериев, а также обладают другими чертами оптимальности.]

Изаacson (Isaacson S. L.)

(1951) On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters, Ann. Math. Stat., vol. 22, 217—234. [Вводятся критерии типа D и E .]

Кифер (Kiefer J.)

(1958) On the nonrandomized optimality and randomized nonoptimality of symmetrical designs, Ann. Math. Stat., vol. 29, 675—699. [Задача 4.]

Леман (Lehmann E. L.)

(1947) On families of admissible tests, *Ann. Math. Stat.*, vol. 18, 97—104. [Последний пример раздела 5.]

(1950) Some principles of the theory of testing hypotheses, *Ann. Math. Stat.*, vol. 21, 1—26. [Теорема 1, задача 10.]

(1955) Ordered families of distributions, *Ann. Math. Stat.*, vol. 26, 399—419. [Лемма 1, задачи 2, 3 *), 8.]

Леман и Стейн (Lehmann E. L. and Stein C.)

(1949) On the theory of some nonparametric hypotheses, *Ann. Math. Stat.*, vol. 20, 28—45. [Задача 14.]

Нейман (Neuman J.)

(1935) Sur la vérification des hypothèses statistiques composées, *Bull. Soc. Math. France*, vol. 63, 246—266. [Определяются критерии типа *B*, т. е. критерии, являющиеся при наличии «мешающих» параметров ЛНМ в классе всех локально несмещенных критериев, а также показано, как их получить.]

Нейман и Пирсон (Neuman J. and Pearson E. S.)

(1936, 1938) Contributions to the theory of testing statistical hypotheses, *Stat. Res. Mem.*, vol. I, 1—37; vol. II, 25—57. [Обсуждаются критерии типа *A*, т. е. критерии, являющиеся ЛНМ в классе всех локально несмещенных критериев при отсутствии мешающих параметров.]

Руист (Ruist E.)

(1954) Comparison of tests for nonparametric hypotheses, *Arkiv Math.*, vol. 3, 133—163. [Задача 4.]

Хант и Стейн (Hunt G. and Stein C.)

(1946) Most stringent tests of statistical hypotheses. [Не опубликована.]

Шёнберг (Schoenberg I. J.)

(1951) On Pólya frequency functions, I, *J. Anal. Math.*, vol. 1, 331—374. [Пример 1.]

Шметтерер (Schmetterer L.)

(1976) Введение в математическую статистику, пер. с нем., М., Наука, гл. 3.

*) Задача 3 представляет собой исправленный вариант теоремы 3 этой работы. Я благодарен Р. Блюменталю, обнаружившему ошибку в этой теореме.

ДОПОЛНЕНИЕ

1. Отношения эквивалентности; группы

Отношение $x \sim y$ между точками пространства $\mathcal{X}$ называется *отношением эквивалентности*, если оно рефлексивно, симметрично и транзитивно, т. е. если

- (I) $x \sim x$ для всех $x \in \mathcal{X}$,
- (II) $x \sim y$ влечет $y \sim x$,
- (III) $x \sim y, y \sim z$ влечет $x \sim z$.

Пример 1. Рассмотрим класс статистических решающих процедур как пространство, точками которого являются отдельные процедуры. Тогда отношение $\delta \sim \delta'$, если процедуры δ и δ' имеют одну и ту же функцию риска, есть отношение эквивалентности. В качестве другого примера рассмотрим на действительной прямой все функции с действительными значениями как точки пространства. Тогда соотношение $f \sim g$, если $f(x) = g(x)$ почти всюду, будет отношением эквивалентности.

Пусть дано отношение эквивалентности; обозначим D_x множество точек пространства, эквивалентных точке x . Тогда $D_x = D_y$, если $x \sim y$, и $D_x \cap D_y = \emptyset$ в противном случае. Так как по (I) каждая точка пространства принадлежит, по крайней мере, одному из множеств D_x , то из этого следует, что множества D_x — *классы эквивалентности*, определенные отношением $\sim$, образуют разбиение пространства на части.

Множество G элементов называется *группой*, если выполняются следующие условия.

(I) Определена операция — групповое умножение, которая любым двум элементам $a, b \in G$ ставит в соответствие элемент $c \in G$. Элемент c называется произведением a и b и обозначается ab .

(II) Групповое умножение подчиняется ассоциативному закону

$$(ab)c = a(bc).$$

(III) Существует элемент $e \in G$, называемый *единичным*, такой, что

$$ae = ea = a \text{ для всех } a \in G.$$

(IV) Для каждого элемента $a \in G$ существует ему обратный элемент $a^{-1} \in G$ такой, что

$$aa^{-1} = a^{-1}a = e.$$

Можно показать, что как единичный элемент, так и обратный любому элементу a определяются однозначно.

Пример 2. Множество всех $(n \times n)$ -ортогональных матриц составляет группу, если матричное умножение и обращение матриц принять соответственно за групповое умножение и обратную операцию и если единичную матрицу принять за единичный элемент группы. С тем же самым определением групповой операции класс всех невырожденных $(n \times n)$ -матриц также образует группу. С другой стороны, класс всех $(n \times n)$ -матриц не удовлетворяет условию (IV).

Если элементы G являются отображениями некоторого пространства на себя, причем групповое произведение ba определяется как результат выполнения преобразования с помощью элемента a и последующего применения элемента b , то G называется *группой преобразований*. В данном случае предположение (IV) удовлетворяется автоматически. Для любой группы преобразований, определенной на пространстве $\mathcal{X}$, отношение между точками $\mathcal{X}$:

$x \sim y$, если существует элемент $a \in G$ такой, что $y = ax$, будет отношением эквивалентности.

Тот факт, что удовлетворяются условия эквивалентности (I), (II) и (III), следует из формулировок (III), (IV) и (I) (соответственно) свойств группы.

Пусть $\mathcal{G}$ — любой класс взаимно однозначных отображений пространства и пусть G является классом всех конечных произведений $a_1^{\pm 1} a_2^{\pm 1} \dots a_m^{\pm 1}$ с $a_1, \dots, a_m \in \mathcal{G}$, $m = 1, 2, \dots$, где каждый показатель равен $+1$ или -1 и где среди элементов $a_1, a_2, \dots$, могут быть равные. Тогда легко проверяется, что G — группа и к тому же наименьшая группа, содержащая $\mathcal{G}$.

2. Сходимость распределений

Когда изучаются свойства сходимости функций, часто удобнее рассматривать класс функций как реализацию абстрактного пространства $\mathcal{F}$ точек f , в котором определена сходимость последовательности f_n к пределу f , обозначаемая $f_n \rightarrow f$.

Пример 3. Пусть μ — мера на измеримом пространстве $(\mathcal{X}, \mathcal{A})$.

(I) Обозначим $\mathcal{F}$ класс интегрируемых функций. Тогда f_n сходится к f в среднем, если *)

$$\int |f_n - f| d\mu \rightarrow 0. \quad (1)$$

*) Здесь и в следующих примерах предельная функция f не единственна. Более определенно, если $f_n \rightarrow f$, то $f_n \rightarrow g$ тогда и только тогда, когда $f = g$ μ -почти всюду. Положив $f \sim g$ при $f = g$ μ -почти всюду, мы можем получить однозначность, переходя к классам эквивалентности.

(II) Пусть $\mathcal{F}$ — равномерно ограниченный класс измеримых функций. Последовательность f_n называют сходящейся к f слабо, если

$$\int f_n p \, d\mu \rightarrow \int f p \, d\mu \quad (2)$$

для всех μ -интегрируемых функций p .

(III) Пусть $\mathcal{F}$ является классом измеримых функций. Тогда f_n сходятся к f поточно, если

$$f_n(x) \rightarrow f(x) \quad \mu\text{-почти всюду.} \quad (3)$$

Подмножество $\mathcal{F}_0$ пространства $\mathcal{F}$ плотно в $\mathcal{F}$, если для любого элемента $f \in \mathcal{F}$ существует последовательность элементов из $\mathcal{F}_0$, имеющая f своей предельной точкой. Пространство $\mathcal{F}$ называется *сепарабельным*, если существует счетное плотное подмножество $\mathcal{F}$. Пространство $\mathcal{F}$ такое, что каждая последовательность содержит сходящуюся подпоследовательность, предельная точка которой принадлежит $\mathcal{F}$, называется *компактом**). Пространство $\mathcal{F}$ является *метрическим пространством*, если для каждой пары точек $f, g \in \mathcal{F}$ определено расстояние $d(f, g) \geq 0$ такое, что

(I) $d(f, g) = 0$ тогда и только тогда, когда $f = g$;

(II) $d(f, g) = d(g, f)$;

(III) $d(f, g) + d(g, h) \geq d(f, h)$ для всех f, g, h .

Пространство *псевдометрично*, если (I) заменяется на

$$(I') \quad d(f, f) = 0 \quad \text{для всех } f \in \mathcal{F}.$$

Псевдометрическое пространство может быть превращено в метрическое введением отношения эквивалентности $f \sim g$, если $d(f, g) = 0$. Классы эквивалентности $F, G, \dots$ в этом случае образуют метрическое пространство по отношению к расстоянию $D(F, G) = d(f, g)$, где $f \in F, g \in G$.

В любом псевдометрическом пространстве естественно определить сходимость, полагая $f_n \rightarrow f$, если $d(f_n, f) \rightarrow 0$.

Пример 4. Пространство интегрируемых функций примера 3 (I) становится псевдометрическим пространством, если мы положим

$$d(f, g) = \int |f - g| \, d\mu.$$

Тогда определение сходимости для пространства с такой метрикой совпадает с данным формулой (I).

Пример 5. Обозначим $\mathcal{P}$ семейство распределений вероятностей на $(\mathcal{X}, \mathcal{A})$. Тогда $\mathcal{P}$ — метрическое пространство по отношению к метрике

$$d(P, Q) = \sup_{A \in \mathcal{A}} |P(A) - Q(A)|. \quad (4)$$

*) Термин *компактность* обычно употребляется для другого понятия, которое в случае метрических пространств совпадает с данным здесь. Поэтому для данного здесь понятия иногда употребляют название *секвенциальная компактность*. — Прим. пер.

Лемма 1. Каждое подмножество сепарабельного псевдометрического пространства $\mathcal{F}$ сепарабельно.

Доказательство. Согласно предположению, существует плотное счетное подмножество $\{f_n\}$ пространства $\mathcal{F}$. Пусть

$$S_{m,n} = \{f : d(f, f_n) < 1/m\}$$

и пусть A — любое подмножество $\mathcal{F}$. Выберем один элемент из каждого непустого пересечения $A \cap S_{m,n}$ и обозначим полученное счетное множество элементов A_0 . Если a — любой элемент из A и m — любое положительное число, то существует элемент f_{n_m} такой, что $d(a, f_{n_m}) < 1/m$. Поэтому $a \in S_{m,n_m}$, пересечение $A \cap S_{m,n_m}$ не пусто и существует элемент в A_0 такой, что расстояние от a до этого элемента $< 2/m$. Этим показано, что A_0 плотно в A и что, следовательно, A сепарабельно.

Лемма 2. Последовательность интегрируемых функций f_n сходится к f в среднем тогда и только тогда, когда

$$\int_A f_n d\mu \rightarrow \int_A f d\mu \text{ равномерно для } A \in \mathcal{A}. \quad (5)$$

Доказательство. Очевидно, что (1) влечет (5), так как для всех $A \in \mathcal{A}$

$$\left| \int_A f_n d\mu - \int_A f d\mu \right| \leq \int_A |f_n - f| d\mu.$$

Обратно, пусть выполняется (5). Обозначим A_n и A'_n множества точек x , для которых $f_n(x) > f(x)$ и $f_n(x) < f(x)$ соответственно. Тогда

$$\int |f_n - f| d\mu = \int_{A_n} (f_n - f) d\mu - \int_{A'_n} (f_n - f) d\mu \rightarrow 0.$$

Лемма 3. Последовательность равномерно ограниченных функций f_n сходится к ограниченной функции f слабо тогда и только тогда, когда

$$\int_A f_n d\mu \rightarrow \int_A f d\mu \text{ для всех } A \text{ с } \mu(A) < \infty. \quad (6)$$

Доказательство. Условие (6) следует из слабой сходимости, если в (2) взять в качестве функции P индикатор множества A , который интегрируем, если $\mu(A) < \infty$. Обратно, из (6) следует, что (2) выполняется, если p — любая простая функция $s = \sum a_i I_{A_i}$, где A_i таковы, что $\mu(A_i) < \infty$. Для любой интегрируемой функции p по определению интеграла существует такая простая функция s , для которой $\int |p - s| d\mu < \varepsilon/3M$, где M —

константа, ограничивающая $|f|$. Тогда мы имеем

$$\left| \int (f_n - f) p \, d\mu \right| \leq \\ \leq \left| \int f_n (p - s) \, d\mu \right| + \left| \int f (s - p) \, d\mu \right| + \left| \int (f_n - f) s \, d\mu \right|.$$

Первые два слагаемых правой части неравенства меньше $\varepsilon/3$ (каждое), а третье слагаемое стремится к нулю, когда n стремится к бесконечности. Таким образом, левая часть неравенства $< \varepsilon$ при n достаточно больших, что и требовалось доказать.

Лемма 4 *). Пусть f и f_n , $n = 1, 2, \dots$, — неотрицательные интегрируемые функции и такие, что

$$\int f \, d\mu = \int f_n \, d\mu = 1.$$

Тогда поточечная сходимость f_n к f влечет сходимость f_n к f в среднем.

Доказательство. Если $g_n = f_n - f$, то $g_n \geq -f$ и отрицательная часть $g_n^- = \max(-g_n, 0)$ удовлетворяет неравенству $|g_n^-| \leq f$. Так как $g_n(x) \rightarrow 0$ (μ -почти всюду), то из теоремы 1 (II) главы 2 следует, что $\int g_n^- \, d\mu \rightarrow 0$; $\int g_n^+ \, d\mu$ здесь также стремится к нулю, так как $\int g_n \, d\mu = 0$. Поэтому $\int |g_n| \, d\mu = \int (g_n^+ + g_n^-) \, d\mu \rightarrow 0$, что и требовалось.

Пусть P и P_n , $n = 1, 2, \dots$, — распределения вероятностей на $(\mathcal{X}, \mathcal{A})$ с плотностями p и p_n по отношению к мере μ . Рассмотрим определения сходимости

(а) $p_n \rightarrow p$ (μ -почти всюду);

(б) $\int |p_n - p| \, d\mu \rightarrow 0$;

(в) $\int g p_n \, d\mu \rightarrow \int g p \, d\mu$ для всех ограниченных измеримых g ;

и

(б') $P_n(A) \rightarrow P(A)$ равномерно для всех $A \in \mathcal{A}$;

(в') $P_n(A) \rightarrow P(A)$ для всех $A \in \mathcal{A}$.

Тогда леммы 2 и 4 вместе с незначительно видоизмененной леммой 3 показывают, что (а) влечет (б) и (б) влечет (в); и что (б) эквивалентно (б'), а (в) эквивалентно (в'). Кроме того, можно показать, что ни (а) и (б), ни (б) и (в) не эквивалентны **).

*) Scheffé, A useful convergence theorem for probability distributions, Ann. Math. Stat. 18 (1947), 434—438.

**) Robbins, Convergence of distributions, Ann. Math. Stat. 19 (1948), 72—76.

3. Доминируемые семейства распределений

Пусть $\mathcal{M}$ — семейство мер, определенных на измеримом пространстве $(\mathcal{X}, \mathcal{A})$. Говорят, что $\mathcal{M}$ *доминируемо* σ -конечной мерой μ , определенной на $(\mathcal{X}, \mathcal{A})$, если каждый элемент $\mathcal{M}$ абсолютно непрерывен по отношению к μ . Семейство $\mathcal{M}$ называется *доминируемым*, если существует σ -конечная мера, которой оно доминируемо. Действительно, если $\mathcal{M}$ является доминируемым, то всегда существует конечная доминирующая мера. В самом деле, допустим, что $\mathcal{M}$ доминируемо мерой μ и что $\mathcal{X} = \bigcup A_i$ с $\mu(A_i)$ конечной для всех i . Если множества A_i взять взаимно непересекающимися, то мера ν , определенная как

$$\nu(A) = \sum \mu(A \cap A_i) / 2^i \mu(A_i),$$

также доминирует $\mathcal{M}$ и конечна.

Теорема 1*). Семейство $\mathcal{P}$ вероятностных мер на евклидовом пространстве $(\mathcal{X}, \mathcal{A})$ доминируемо тогда и только тогда, когда оно сепарабельно по отношению к метрике (4) или, что эквивалентно, по отношению к сходимости

$$P_n \rightarrow P, \text{ если } P_n(A) \rightarrow P(A) \text{ равномерно для } A \in \mathcal{A}.$$

Доказательство. Допустим сначала, что $\mathcal{P}$ сепарабельно и что последовательность $\{P_n\}$ плотна в $\mathcal{P}$, и пусть

$$\mu = \sum P_n / 2^n.$$

Тогда $\mu(A) = 0$ влечет $P_n(A) = 0$ для всех n , и следовательно, $P(A) = 0$ для всех $P \in \mathcal{P}$. Обратно, пусть $\mathcal{P}$ доминируемо мерой μ , которая без ограничения общности предполагается конечной. Тогда мы должны показать, что множество интегрируемых функций $dP/d\mu$ сепарабельно по отношению к сходимости, определенной (5), или, согласно лемме 2, по отношению к сходимости в среднем. Но, исходя из утверждения леммы 1, достаточно доказать, что этим свойством обладает класс $\mathcal{F}$ всех μ -интегрируемых функций f . Так как по определению интеграла каждая интегрируемая функция может быть аппроксимирована в среднем простыми функциями, то достаточно доказать наше утверждение для случая, когда $\mathcal{F}$ является классом всех простых интегрируемых функций. Любая простая функция может быть аппроксимирована в среднем простыми функциями с рациональными значениями, так что достаточно показать сепарабельность класса функций $\sum r_i I_{A_i}$, где r рациональны и A — борелевские множества с конечной μ -мерой (так как функции f интегрируемы). Наконец, достаточно взять в качестве $\mathcal{F}$ класс

*) Berger, Remark on separable spaces of probability measures, Ann. Math. Stat. 22 (1951), 119—120.

функций I_A — индикаторов борелевских множеств с конечной мерой. Однако любое такое множество может быть аппроксимировано конечными суммами непересекающихся прямоугольников с рациональными концами. Множество всех таких сумм счетно, поэтому рациональные линейные комбинации соответствующих индикаторов образуют счетное плотное подмножество $\mathcal{F}$.

Анализ доказательства показывает, что «евклидовость» пространства $(\mathcal{X}, \mathcal{A})$ была использована лишь для установления факта существования счетного числа множеств $A_i \in \mathcal{A}$ таких, что любому $A \in \mathcal{A}$ с конечной мерой соответствует подпоследовательность A_{ij} с $\mu(A_{ij}) \rightarrow \mu(A)$. Это свойство выполняется, вообще, для любого σ -поля $\mathcal{A}$, которое обладает *счетным множеством образующих*, т. е. в котором существует счетное число множеств B_i таких, что $\mathcal{A}$ является наименьшим σ -полем, содержащим все B_i *).

Отсюда следует, что теорема 1 выполняется для любого σ -поля с таким свойством. В статистических приложениях такие σ -поля встречаются в последовательном анализе, где выборочное пространство $\mathcal{X}$ есть сумма $\mathcal{X} = \bigcup_i \mathcal{X}_i$ борелевских подмножеств $\mathcal{X}_i$ i -мерного евклидова пространства. В этих проблемах $\mathcal{X}_i$ — множество точек $(x_1, \dots, x_i)$, которые требуют ровно i наблюдений. Если $\mathcal{A}_i$ — σ -поле борелевских подмножеств $\mathcal{X}_i$, то можно взять в качестве $\mathcal{A}$ σ -поле, порожденное $\mathcal{A}_i$, и так как

каждое $\mathcal{A}_i$ обладает счетным множеством образующих, то $\mathcal{A}$ также обладает им.

Если $\mathcal{A}$ не обладает счетным множеством образующих, то может быть сделано более слабое утверждение. Два семейства мер $\mathcal{M}$ и $\mathcal{N}$ называются *эквивалентными*, если равенство $\mu(A) = 0$ для всех $\mu \in \mathcal{M}$ влечет $\nu(A) = 0$ для всех $\nu \in \mathcal{N}$, и наоборот.

Теорема 2).** Семейство $\mathcal{P}$ вероятностных мер доминируемо σ -конечной мерой тогда и только тогда, когда $\mathcal{P}$ имеет счетное эквивалентное подмножество.

Доказательство. Допустим сначала, что $\mathcal{P}$ имеет счетное эквивалентное подмножество $\{P_1, P_2, \dots\}$. Тогда $\mathcal{P}$ доминируемо мерой $\mu = \sum P_n/2^n$. Обратно, пусть $\mathcal{P}$ доминировано σ -конечной мерой μ , про которую, без ограничения общности, можно предположить, что она конечна. Пусть $\mathcal{Q}$ — класс всех вероятностных мер Q вида $\sum c_i P_i$, где $P_i \in \mathcal{P}$, все c_i положительны и $\sum c_i = 1$. Класс $\mathcal{Q}$ также доминируется мерой μ , и мы обозначим q фиксированный вариант плотности $dQ/d\mu$.

*) Доказательство дано, например, в книге: Халм ош П., Теория меры, ИЛ, 1953 (теорема 2 раздела 40).

**) H a l m o s and S a v a g e, Application of the Radon — Nikodym theorem to the theory of sufficient statistics, Ann. Math. Stat. 20 (1948), 225—241.

Докажем эквивалентное теореме утверждение: существует $Q_0 \in \mathcal{Q}$ такое, что из $Q_0(A) = 0$ следует $Q(A) = 0$ для всех $Q \in \mathcal{Q}$.

Рассмотрим класс $\mathcal{V}$ множеств C в $\mathcal{A}$, для которых существует $Q \in \mathcal{Q}$ такое, что $q(x) > 0$ μ -почти всюду на C и $Q(C) > 0$.

Пусть $\mu(C_i)$ стремится к $\sup_{\mathcal{C}} \mu(C)$, пусть $q_i(x) > 0$ почти всюду на C_i ; обозначим C_0 сумму множеств C_i . Тогда $q_0^*(x) = \sum c_i q_i(x)$ совпадает почти всюду с плотностью вероятности $Q_0 = \sum c_i Q_i$, q_0^* положительна почти всюду на C_0 , так что $C_0 \in \mathcal{V}$. Допустим теперь, что $Q_0(A) = 0$. Пусть Q — любой другой элемент $\mathcal{Q}$ и пусть $C = \{x: q(x) > 0\}$. Тогда $Q_0(A \cap C_0) = 0$, поэтому $\mu(A \cap C_0) = 0$ и $Q(A \cap C_0) = 0$. Точно так же и $Q(A \cap \bar{C}_0 \cap \bar{C})$ равно нулю. Наконец, неравенство $Q(A \cap \bar{C}_0 \cap C) > 0$ привело бы к тому, что $\mu(C_0 \cup [A \cap \bar{C}_0 \cap C]) > \mu(C_0)$, а это противоречит равенству $\mu(C_0) = \sup_{\mathcal{C}} \mu(C)$, так как $A \cap \bar{C}_0 \cap C$, а поэтому и $C_0 \cup [A \cap \bar{C}_0 \cap C]$ принадлежат $\mathcal{V}$.

4. Теорема о слабой компактности

Следующая теорема создает основу для доказательства существования наиболее мощных критериев, наиболее строгих критериев и т. д.

Теорема 3*) (Теорема о слабой компактности). Пусть μ — σ -конечная мера в евклидовом пространстве или, более общим образом, в любом измеримом пространстве $(\mathcal{X}, \mathcal{A})$, в котором $\mathcal{A}$ обладает счетным множеством образующих. Тогда множество измеримых функций f с $0 \leq f \leq 1$ является компактом по отношению к слабой сходимости, определяемой условием (2).

Доказательство. Возьмем любую последовательность $\{f_n\}$. Мы должны доказать существование подпоследовательности $\{f_{n_i}\}$ и функции f таких, что

$$\lim \int f_{n_i} p \, d\mu = \int f p \, d\mu$$

при всех интегрируемых p . Если μ^* — конечная мера, эквивалентная μ , то p^* μ^* -интегрируема тогда и только тогда, когда $p = (d\mu^*/d\mu) p^*$ μ -интегрируема, и $\int f p \, d\mu = \int f p^* \, d\mu^*$ при всех f . Поэтому мы можем, без ограничения общности, предположить, что μ конечна. Пусть $\{p_n\}$ — последовательность, плотная в классе функций p по отношению к сходимости в среднем. Су-

*) Banach, Théorie des opérations linéaires, Warszawa, Fundusz Kultury Narodowej, 1932, 131.

существование такой последовательности гарантировано теоремой 1 и следующим за ней замечанием. Если

$$\Phi_n(p) = \int \varphi_n p \, d\mu,$$

то последовательность $\Phi_n(p)$ ограничена при каждом p . Можно выбрать подпоследовательность Φ_{n_k} так, что $\Phi_{n_k}(p_m)$ сходится при каждом p_m . Для этого применяем следующий диагональный процесс. Рассмотрим сначала последовательность чисел $\{\Phi_n(p_1)\}$, которая обладает сходящейся подпоследовательностью $\Phi_{n'_1}(p_1)$, $\Phi_{n'_2}(p_1), \dots$. Следующая последовательность $\Phi_{n'_1}(p_2)$, $\Phi_{n'_2}(p_2), \dots$ содержит сходящуюся подпоследовательность $\Phi_{n''_1}(p_2)$, $\Phi_{n''_2}(p_2), \dots$. Продолжая этот процесс, положим $n_1 = n'_1$, $n_2 = n''_2$, $n_3 = n'''_3, \dots$. Тогда $n_1 < n_2 < \dots$, и последовательность $\{\Phi_{n_i}\}$ сходится при каждом p_m .

Из неравенства

$$\left| \int (\varphi_{n_j} - \varphi_{n_i}) p \, d\mu \right| \leq \left| \int (\varphi_{n_j} - \varphi_{n_i}) p_m \, d\mu \right| + 2 \int |p - p_m| \, d\mu$$

следует, что $\Phi_{n_i}(p)$ сходится при всех p . Обозначим ее предел $\Phi(p)$ и определим функцию множества Φ^* на $\mathcal{A}$, полагая

$$\Phi^*(A) = \Phi(I_A).$$

Тогда Φ^* неотрицательна и ограничена, так как при всех A $\Phi^*(A) \leq \mu(A)$. К тому же она счетно-аддитивна. Действительно, пусть $A = \bigcup A_k$, где A_k не пересекаются. Тогда

$$\Phi^*(A) = \lim \Phi_{n_i}^*(\bigcup A_k)$$

и

$$\left| \int_{\bigcup A_k} \varphi_{n_i} \, d\mu - \sum \Phi^*(A_k) \right| \leq \left| \int_{\bigcup_{k=1}^m} \varphi_{n_i} \, d\mu - \sum_{k=1}^m \Phi^*(A_k) \right| + \left| \int_{\bigcup_{k=m+1}^{\infty} A_k} \varphi_{n_i} \, d\mu - \sum_{k=m+1}^{\infty} \Phi^*(A_k) \right|.$$

Здесь второе слагаемое равно нулю в случае конечных сумм $A = \bigcup_{k=1}^m A_k$, в противном случае оно не превосходит величины $2\mu\left(\bigcup_{k=m+1}^{\infty} A_k\right)$, которая при m достаточно большом может быть сделана произвольно малой. При любом фиксированном m

первое слагаемое стремится к нулю, когда i стремится к бесконечности. Таким образом, Φ^* представляет собой конечную меру на $(\mathcal{X}, \mathcal{A})$. Она, кроме того, абсолютно непрерывна относительно меры μ , так как $\mu(A) = 0$ влечет $\Phi_{n_i}(I_A) = 0$ при всех i и поэтому $\Phi(I_A) = \Phi^*(A) = 0$. Мы можем теперь, применяя теорему Радона — Никодима, получить

$$\Phi^*(A) = \int_A \varphi d\mu \quad \text{для всех } A,$$

где $0 \leq \varphi \leq 1$. Тогда мы имеем

$$\int_A \varphi_{n_i} d\mu \rightarrow \int_A \varphi d\mu \quad \text{для всех } A,$$

и слабая сходимость последовательности φ_{n_i} к φ следует из леммы 3.

ИМЕННОЙ УКАЗАТЕЛЬ

- Андерсен (Andersen S. L.) 194, 210
 Андерсон (Anderson T. W.) 174, 294, 336, 361, 363
 Анскомб (Anscombe F. J.) 356
 Армсен (Armsen F.) 162
 Арнольд (Arnold R. J.) 294
- Банах (Banach S.) 398
 Барнард (Barnard G.) 294, 295
 Барра (Barra J. R.) 177
 Бартлетт (Bartlett M. S.) 138, 177, 363
 Басу (Basu D.) 182, 235
 Бахадур (Bahadur R. R.) 30, 70, 71, 72, 160
 Беннет (Bennett B.) 204
 Бергер (Berger A.) 368, 396
 Беренс (Behrens W. V.) 296
 Берксон (Berkson J.) 165
 Бирнбаум (Birnbau A.) 161, 260
 Бирнбаум (Birnbau Z. W.) 139, 170
 Блекуэл (Blackwell D.) 23, 36, 90, 139, 141
 Блюменталь (Blumenthal R.) 390
 Бокс (Box G. E. P.) 194, 210, 463
 Большев Л. Н. 223, 257, 265
 Боукер (Bowker A. H.) 365
 Браун (Brown G.) 36
- Вальд (Wald A.) 26, 36, 37, 100, 128, 139, 177, 210, 294, 318, 349, 350, 363, 389
 Ван-дер-Варден (van der Waerden B. L.) 36, 37, 268, 294
 Ватсон (Watson G. G.) 363
 Веслер (Wesler O.) 389
 Вийсман (Wijsman K. A.) 118, 365
 Виленский (Wilenski H.) 178
 Вилкоксон (Wilcoxon F.) 176, 294, 296
 Вилсон (Wilson E. B.) 138, 139
 Вольфовиц (Wolfowitz J.) 37, 128, 139, 140, 175, 210, 294, 389
- Гаварэ (Gavarret J.) 138, 139
 Гаек (Hájek J.) 280, 295
 Гаусс (Gauss C. F.) 36, 138, 139
 Гейен (Geyen A. K.) 188
 Гербрах (Herbrach L. H.) 363
 Гери (Geary R. C.) 188
 Гиршик (Girshick M. A.) 23, 36, 117, 139, 141, 177, 235
 Голдберг (Goldberg M.) 295
 Госсет (Gossett W. S.) см. Стьюдент
 Гренандер (Grenander V.) 139, 235, 366
 Гринберг (Greenberg B. J.) 265
 де Гроот (De Groot M.) 36
 Гудман (Goodman L.) 164, 174
 Гхош (Ghosh M. N.) 177
- Даниелс (Daniels H. E.) 294, 295
 Данфорд (Dunford N.) 71
 ван Дантиг (van Dantzig D.) 294
- Дантиг (Dantzig G. B.) 100, 139
 Дарлинг (Darling D.) 270, 278
 Дармуа (Darmois G.) 63
 Дворецкий (Dvoretzky A.) 128, 140
 Демпстер (Dempster A. P.) 339
 Дейхлер (Deuchler G.) 294, 295
 Джеймс (James A. T.) 377
 Джефрис (Jeffreys H.) 138
 Диксон (Dickson L. E.) 53
 Дуб (Doob J. L.) 11, 31, 71, 72
 Дункан (Duncan D. B.) 310
 Дуосс (Dwass M.) 210
 Дынкин Е. Б. 31
 Дэвид (David F. N.) 177, 234
 Дэвид (David H. T.) 281
- Зиппин (Zippin L.) 378
- Изааксон (Isaakson S. L.) 389
 Ирвин (Irwin J. O.) 177
 Йейтс (Yates F.) 296
- Карлин (Karlin S.) 88, 140, 257
 Кемпторн (Kempthorne O.) 318, 363, 364
 Кендалл (Kendall M. G.) 36, 37, 140, 164, 177, 235, 291, 294, 295, 363, 394
 Кифер (Kiefer J.) 38, 140, 318, 389
 Кокрен (Cochran W. G.) 318, 363, 364
 Кокс (Cox D. R.) 282
 Кокс (Cox G.) 318
 Колмогоров А. Н. 37, 44, 71, 72, 270, 278
 Колоджейчик (Kolodziejczyk S.) 364
 Коэн (Cohen L.) 137
 Крамер (Cramér H.) 24, 36, 37, 44, 151, 188, 270, 278
 Крускал (Kruskal W. H.) 137, 164, 177, 281, 295, 344
 Кудо (Kudo H.) 38
 Купмен (Koopman B. O.) 62
- Ланкастер (Lancaster H. O.) 363, 364
 Лаплас (Laplace P. S.) 36, 37, 138, 140
 Ле Кам (Le Cam L.) 24, 37, 125, 350
 Лексис (Lexis W.) 133, 140
 Леман (Lehmann E. L.) 28, 30, 37, 39, 72, 110, 114, 140, 159, 160, 177, 188, 235, 268, 295, 310, 311, 343, 349, 390
 Либерман (Lieberman G.) 223
 Линдли (Lindley D. V.) 24, 138
 Линник Ю. В. 72, 257
 Литлвуд (Littlewood J. E.) 123
 Люэв (Loève M.) 56, 71, 72
 Ляпунов А. М. 128
- Мей (May J. M.) 310
 Мейнланд (Mainland D.) 162
 Мерингтон (Merrington M.) 223

- Мильтон (Milton E. T.) 296
 Монтгомери (Montgomery D.) 378
 Моран (Moran P. A. P.) 294, 295
 Морган (Morgan W. A.) 235
 Мостеллер (Mosteller F.) 128, 177
- Нанди (Nandi H. K.) 178
 Невё (Neveu J.) 71
 Нейман (Neuman J.) 36, 38, 138, 139, 140, 165, 178, 204, 235, 364, 390
 фон Нейман (von Neumann J.) 36
 Нётер (Noether G.) 210
 Нивергельт (Nievergelt E.) 268
 Ньюман (Newmann D.) 364
- Ольмстед (Olmstead P. S.) 291
- Пагурова В. И. 257
 Патнаик (Patnaik P. B.) 343
 Паулсон (Paulson E.) 235
 Пейсакوف (Peisakoff M.) 38
 Пирсон (Pearson K.) 36, 138, 141, 223, 364
 Пирсон (Pearson E. S.) 38, 138, 140, 178, 302, 364, 390
 Питмен (Pitman E. J. G.) 38, 63, 235, 293, 295
 Питчер (Pitcher T. S.) 295
 Поля (Polya G.) 117, 123
 Путтер (Putter J.) 178
 Пшиборовский (Przyborowski J.) 178
- Рао (Rao C. R.) 308, 309, 363, 364
 Раштон (Rushton S.) 295
 Резников (Resnikoff G.) 223
 Роббинс (Robbins H.) 395
 Робинсон (Robinson J.) 117
 Рубин (Rubin H.) 140
 Руист (Ruist E.) 390
- Сакс (Saks S.) 71, 72
 Сандрум (Sundrum R. M.) 294, 295
 Сархан (Sarhan A. E.) 265
 Свед (Swed F.) 175
 Сведруп (Sverdrup E.) 178
 Серпинский (Sierpinski W.) 372
 Симанк (Simaika J. B.) 365
 Смирнов Н. В. 223, 257, 265, 270, 278
 Смит (Smith C. A. B.) 179
 Собел (Sobel M.) 72, 141
 Стейн (Stein C.) 39, 114, 118, 140, 141, 235, 258, 260, 294, 295, 296, 365, 389, 390
 Стюарт (Stuart A.) 37, 140, 164, 177, 235, 295, 363, 364
 Студент (Student) 235, 366
 Сукхатм (Sukhatme P. V.) 296
 Сэвидж (Savage I. R.) 268, 294, 296
 Сэвидж (Savage L. J.) 23, 36, 38, 72, 177, 397
 Сюй (Hsu C. T.) 235
 Сюй (Hsu P. L.) 296, 365
 Сютклифф (Sutcliffe M. I.) 162
- Тейкроу (Teichroew D.) 265
 Токер (Tocher K. D.) 178
- Томпсон (Thompson W. R.) 141
 Тьюки (Tukey J.) 28, 129, 138, 141, 188, 291, 296, 310, 316, 356
 Тэнг (Tang P. G.) 365
- Уайтфилд (Whitfield J. W.) 294, 295
 Уилкс (Wilks S. S.) 36, 38, 294, 296, 365
 Уоллес (Wallace D.) 188, 258
 Уоллис (Wallis W. A.) 170
 Уолш (Walsh J.) 178, 296
 Уоркинг (Working H.) 138, 141
 Уэлч (Welch B. L.) 258
- Феллер (Feller W.) 11, 71, 136, 178
 Фикс (Fix E.) 204, 343, 349
 Фишер (Fisher R. A.) 36, 38, 138, 141, 177, 178, 235, 296, 363, 365
 Фокс (Fox M.) 302
 Фомин С. В. 71
 Фрезер (Fraser D. A. S.) 141, 178, 210, 222, 294, 296
 Фримен (Freeman M. F.) 356
 Фурье (Fourier J. B. J.) 138, 141
- Халмош (Halmos P.) 71, 72, 128, 247, 378, 397
 Хант (Hunt G.) 39, 294, 296, 363, 389, 390
 Харди (Hardy G. H.) 123
 Хартли (Hartley H. O.) 302, 365
 Хельмерт (Helmert F. R.) 236
 Хемелрийк (Hemelrijk J.) 294, 296
 Херера (Herrera L.) 162
 Хёфдинг (Hoeffding W.) 210, 291, 296
 Хобсон (Hobson E. W.) 216
 Ходжес (Hodges J. L.) 39, 93, 188, 268, 270, 302, 343, 349
 Холдейн (Haldane J. B. S.) 179
 Хопф (Hopf E.) 296
 Хотеллинг (Hotelling H.) 39, 138, 141, 293, 366
 Хоул (Hoel P. G.) 179
- Цзао (Tsao C. K.) 297
- Чепмэн (Chapman D. G.) 139, 236, 278
 Чернов (Chernoff H.) 141, 258, 268, 351
 Чубер (Czuber E.) 36, 39
- Шварц (Schwartz J. T.) 71
 Шеффе (Scheffé H.) 28, 30, 141, 159, 177, 179, 198, 294, 297, 310, 363, 365, 395
 Шёнберг (Schoenberg I. J.) 390
 Шидак (Sidak Z.) 260, 296
 Шириев А. Н. 141
 Шметтерер (Schmetterer L.) 36, 39, 297, 390
- Эджворт (Edgeworth F. Y.) 39, 138, 141
 Эйзенхарт (Eisenhart C.) 175, 356, 366
 Эльфинг (Elfvig G.) 70
 Эпстейн (Epstein B.) 72, 141, 297
 Эренфельд (Ehrenfeld S.) 318
 Эрроу (Arrow K. J.) 139, 141
 Эспин (Aspin A.) 258
- Юсава (Usawa H.) 297

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ*)

Абсолютная непрерывность 46
Аддитивность эффектов 311; ~, критерий 316; ~ в модели II 323
Анализ ковариаций 322, 364
Арка синуса преобразование для биномиального распределения 356, 363

Байесовский риск 21

Байесовское решение 21, 33, 34; ~ ограниченное 22; ~ для последовательных проблем с двумя решениями 122

Бартлета критерий для равенства дисперсий 309

Бернса—Фишера проблема 191, 257, 294, 296, 297

Бета-распределение 191; ~, связь с *F*-распределением 191; ~ при исследовании отношений дисперсий 191, 224; ~ как распределение порядковых статистик 286; ~ в задаче проверки линейных гипотез 302; ~ нецентральное 302

Бишмеримое преобразование 239

Биномиальная вероятностная бумага 129

Биномиальное распределение 11; ~, достаточные статистики 26, 36; ~ как экспоненциальное семейство 85; ~, полнота 149; ~ в задаче о сравнении двух распределений Пуассона 160; ~, преобразование арксинуса 356, 363. См. также Полномиальное распределение; Отрицательно-биномиальное распределение

Биномиальные вероятности; ~ односторонний критерий 85, 129, 168, 281; ~, доверительные границы 97, 131; ~, последовательный критерий отношений вероятностей 117, 121, 136, 281; ~, двусторонний критерий 132, 146, 166, 170; ~, сравнение двух 134, 162, 174, 228; ~, сравнение нескольких 309. См. также Таблица сопряженности признаков; Медиана; Парные сравнения; Выборочный контроль; Критерий знаков

Биологические испытания 176

Борелевские множества 41

Вальда тождество 121, 136

Ван-дер-Вардена критерий 268

Взаимодействия (interactions) 315; ~, критерий отсутствия 316; ~, ~ в смешанных моделях 340

Вилкоксона критерий; ~ двухвыборочный 176, 265, 267, 284, 285, 289; ~, относительная асимптотическая эффективность 268; ~ одновыборочный 271, 289; ~, доверительные интервалы 288

Вилкоксона статистика; ~, таблица распределения 176; ~, математическое ожидание и дисперсия 285; ~, симметрия 288
Время ожидания См. Пуассоновский процесс; Экспоненциальное распределение
Выборка 11; ~ расслоенная (*stratified sample*) 205

Выборочное пространство 43; ~ для последовательного анализа 397

Выборочный контроль; ~ по качественному признаку 84, 247, 281; ~ по количественному признаку 109, 247; ~, степень строгости (*stringency*) 93; ~ при сравнении двух партий изделий 166, 250; ~, сравнение двух методов 281

Гарантированная мощность; ~ в последовательных критериях отношений вероятностей 118; ~, достижимая с помощью последовательных процедур 159, 161; ~ при минимальном объеме выборки 368

Гипергеометрическое распределение 84; ~, равномерно наиболее мощный (РНМ) односторонний критерий для проверки среднего 84; ~ в задаче сравнения двух биномиальных вероятностей 162; ~ и проверка независимости в 2×2 таблице 165

«Гнездовая» (nested) классификация 327, 358

Группа 391; ~ конечная 34, 376, 377; ~ свободная 34; ~ перестановок 231; ~ сдвигов 240, 380; ~ транзитивная 241; ~ ортогональная 241, 377, 392

Двойная классификация (two-way classification); ~ с одним наблюдением в клетке 311; ~, доверительные интервалы 314; ~ с несколькими наблюдениями в клетке 314; ~, смешанная модель 330, 340. См. также Взаимодействия; «Гнездовая» классификация

Двумерное нормальное распределение 219; ~, доверительные границы 293; ~, критерий для него 282; ~, критерий независимости 220; ~, критерий отношения дисперсий 232; ~, критерий равенства средних 232; ~, распределение коэффициента корреляции 231, 232, 233; ~, совместное распределение вторых моментов 233. См. также Многомерное нормальное распределение

Двухступенная (two-stage) процедура Стейна 226

Декартово произведение множеств 46

*) Знак ~ заменяет группу слов, выделенных курсивом.

- Дисперсионный анализ** 307, 363; ~ при одинарной классификации 307; ~ при двойной классификации 313, 316; ~, рандомизация 318; ~, различные модели 323, ~ в модели II 330. См. также Линейная гипотеза одномерная
- Доверительные множества** 13, 14, 98, 195; ~, равномерно наиболее точные (*most accurate*) 93, 131, ~, связь с критериями 94, 95, 195; ~ пустые 97, 326; ~, связь с медианой несмещенными оценками 98, 195, 197; ~ равномерно наиболее точные, несмещенные 198; ~, не являющиеся интервалами 204; ~ равномерно наиболее точные, инвариантные 274, 292
- Доверительный коэффициент** (коэффициент доверия) 93, 194
- Доверительный уровень** 93
- Доминируемое (dominated) семейство распределений** 60, 396
- Допустимость (admissibility)** 25; ~ несмещенных и инвариантных процедур 35; ~ РНМ несмещенных критериев 169, 258; ~ РНМ инвариантных критериев 258; ~ критериев t и T^2 260; ~ локальная 259; ~ критериев отношения правдоподобия 260, 284
- Достаточная статистика** 27, 59; ~, критерий факторизации 28, 35, 61; ~, определение в терминах апостериорного распределения 30; ~ минимальная 30, 35; ~, отношение правдоподобия 69; ~ попарная 69; ~, связь с вполне информирующей статистикой 128; ~ при наличии мешающих параметров 135; ~, независимость от нее 182; ~ и инвариантность 282; ~ асимптотически 350
- Достаточный эксперимент** 91
- Евклидово пространство** 55
- Естественное параметрическое пространство экспоненциального семейства** 63
- Зависимость (dependence)**; ~ положительная 92, 164, 220, 234, 261; ~ меры 164
- Задача сравнения двух выборок (two-sample problem), непараметрическая**; ~, критерий перестановок 204, 210; ~, доверительные интервалы для сдвига 229, 288; ~, ранговый критерий при односторонних и двусторонних альтернативах 260, 264, 265; ~, ранговый критерий при общих альтернативах 269, 288
- Задача сравнения двух выборок, параметрическая**. См. Биномиальное распределение; Нормальное распределение, Распределение Пуассона; Экспоненциальное распределение
- Задача сравнения нескольких выборок (k-sample problem)** См. Однородности критерий
- «Иерархическая» (hierarchical) классификация** 327, 358
- Измеримое**; ~ пространство 41; ~ множество 41; отображение 43
- Инвариантная мера** 253, 377, 378; ~ на ортогональной группе 377
- Инвариантность**; ~ решающих процедур 20, 32, 33, 34, 35, 38; ~, связь с несмещенностью 32, 255; ~ связь с принципом минимакса 33, 375, 376; ~ критерия 239; ~ меры 253, 377; ~ функции мощности 253; ~ доверительных множеств 273, 276, 291, 292, 293; ~, связь с достаточностью 282, 284; ~ отношения правдоподобия 283; ~, предостережение против неоправданного употребления 310. См. также Почти инвариантность
- Индикатор** 45
- Интегрируемая функция** 44
- Испытания на продолжительность жизни (life-testing)** 70, 142. См. также Пуассоновский процесс; Экспоненциальное распределение
- Каноническая форма**; ~ для одномерной линейной гипотезы 298, 305; ~ для «гнездовой» классификации в модели II 327, 358; ~ для многомерной линейной гипотезы 333
- Качественные признаки (attributes)**; ~, выборочный контроль по ним 84, 247; ~, попарные сравнения 244, 373
- Кендалла t-статистика** 291
- Классификация «по подчиненности»** См. «Гнездовая» классификация
- Классы эквивалентности** 391
- Колмогорова критерий согласия** 278
- Комбинация критериев** 170
- Компакт** 393
- Компоненты дисперсии** 330
- Коши распределение** 90, 373, 385
- Коэффициент внутригрупповой корреляции (intraclass correlation coefficient)** 358
- Коэффициент корреляции**; ~ в двумерном нормальном распределении 220, 234, 282, 293; ~ ранговый 291; ~ внутригрупповой корреляции 358; ~ множественный 361
- Крамера — Мизеса критерий** 270, 278
- Критерии с неймановской структурой** 149, 152
- Критерий 73**; ~ рандомизированный 75, 96, ~ равномерно наиболее мощный (РНМ) 76; ~ типа A 141, 390; ~ типа A_1 141; ~ несмещенный 143; ~ подобный 143, 149; ~ строго несмещенный 146; ~ типа B 178, 390; ~ типа B_1 178; ~ локально наиболее мощный (ЛНМ) 178, 385; ~ инвариантный 239; ~ почти инвариантный 251; ~ максимальный 368; ~ локально максимальный 371; ~ наиболее строгий 382, 388, ~ типов D и E 386
- Критерий знаков (sign test)** 110; ~ при проверке того, какой из товаров предпочтительнее 166; ~, трактовка «ничьих» 167; ~ при отыскании центра симметрии 168, 386; ~ при попарных сравнениях 245; ~ при проверке гипотезы симметрии 271; ~ при проверке гипотез о симметричном экспоненциальном распределении 387. См. также Биномиальные вероятности; Медиана; Выборочный контроль
- Критерий отношения правдоподобия** 24, 34, 38; ~, инвариантность 283; ~, пример, когда не является допустимым 284; ~, теория, основанная на больших выборках 349, 350, 351, 362
- Критерий перестановок** 204, 206, 231; ~ наиболее мощный, при непараметрических гипотезах 205, 221; ~, аппроксимация стандартными нормальными критериями 211, 222; ~ наиболее мощный, при рандомизации 214; ~, доверительные интервалы, основанные на нем 229; ~, наиболее строгий 389
- Критерий суммы рангов** 176 См. также Вилкоксона критерий
- Критерий T^2** ; ~, допустимость 260; ~ как РНМ инвариантный критерий 338; ~, максимальное свойство как открытый вопрос 382

Критерий χ^2 342, 345, 363, 364
Критическая область 73
Критическая функция 75
Критический уровень 75, 129, 169

Лапласа распределение см. Симметричное экспоненциальное распределение

Латинский квадрат 317, 357

Лебега теорема сходимости 45

Линейная гипотеза многомерная 331; ~, каноническая форма 333; ~, редукция с учетом инвариантности 334, 358; ~, свойство критерия 338, 382; ~, обобщение 339; ~, критерий отношения правдоподобия 338; ~ с известной матрицей ковариаций 342. См. также Многомерное нормальное распределение; Критерий T^2

Линейная гипотеза одномерная 298, ~, каноническая форма 299; ~, оптимальные свойства критерия 302, 353, 380, 384, 386; ~, мощность критерия 385; ~, обобщение 304, 354; ~, доверительные множества, связанные с ней 314, 355; ~, критерий перестановок 318; ~ при известной дисперсии 355. См. также Дисперсионный анализ; Смешанная модель; Модель II; Одноранговая классификация; Регрессия; Двойная классификация

Логарифмическое преобразование 308, 363
Логистическое распределение 265, 267, 287, 372

Локально оптимальные свойства критериев 178, 371, 385, 390

Максимальное правдоподобие 24, 34, 349, 362

Максимальный инвариант 240, ~, отыскание в несколько этапов 242

Максимальный критерий 368, ~ локальный 371, 385; ~, связь с инвариантностью 375; ~, существование 384

Максимум F -отношений 311

Маркова цепь 174

Масштабный параметр (scale parameter); ~, устранение с использованием инвариантности 279; ~, критерий 295, 373

Математическое ожидание 45; ~ условное 51, 53, 56

Медиана; ~, доверительные границы 134, 141; ~, критерий 178, 296

Мера зависимости Юла 164

Мера Лебега 41

Мера Хаара 253, 378

Меры зависимости 164, 291

Метод сходных пар (matched pairs) см. Парные сравнения

Метрическое пространство 393

Минимальная достаточная статистика 30, 35, 71

Минимальный полный класс решающих процедур 25

Многомерная линейная гипотеза см. Линейная гипотеза многомерная

Многомерное нормальное распределение 332; ~, проверка гипотез о векторе средних значений 338; ~, проверка равенства векторов средних значений в двух 338; ~ как предел полиномиальных распределений 344; ~, проверка гипотезы о независимости двух групп компонент 361. См. также Двумерное нормальное распределение

Многомерный анализ 363. См. также Линейная гипотеза многомерная; Множественный коэффициент корреляции

Множественный коэффициент корреляции 361; ~, распределение 361; ~, оптимальный критерий 361, 365, 389

Модель I в дисперсионном анализе 306, 323

Модель II в дисперсионном анализе 323, 327, 358

Монотонное отношение правдоподобия 82; ~ в случае гипергеометрического распределения 84; ~, следствие 88, ~, условие его существования 129; ~ в случае нецентрального t -распределения 249; ~ в случае распределения коэффициента корреляции 282, ~ нецентральных χ^2 и F -распределений 352; ~, критерии, основанные на выборках из распределений 371; ~ в случае семейства распределений с параметром сдвига 372

Монотонный класс множеств 66

Мощность критерия 74

Наиболее строгий (stringent) критерий 267, 382, 388

Наименее благоприятное (favorable) распределение 26, 107, 369

Наименьших квадратов метод 303

Независимость 46; ~, критерий в таблице сопряженности признаков 91, 162, 362; ~ статистики от полной достаточной статистики 182; ~, критерий в двумерном нормальном распределении 218, 282; ~, непараметрическая гипотеза 221, 261, 263; ~, непараметрический критерий 291

Неопределенные множители 103

Непараметрические критерии; ~, обзор 294; ~, библиография 296. См. также Критерий перестановок; Ранговый критерий, Критерий знаков

Несмещенность; ~ решающих процедур 20; ~ критериев 20, 143; ~ доверительных множеств 20, 197; ~ точечных оценок 21, 31, 158; ~ медианная 32, 98; ~, связь с инвариантностью 32, 255; ~, связь с принципом минимакса 33, 371; ~, связь с понятием подобия 144; ~ строгая (strict) 146

Нецентральное; ~ t -распределение 187, 223, 249, 281; ~ F -распределение 302, 331, 343, 352; ~ χ^2 -распределение 343, 352, 356; ~ бета-распределение 302, 352

«Ничь» в критерии знаков 166

Нормальная подгруппа 279

Нормальное распределение 11; ~, достаточные статистики 29

Нормальное распределение, две выборки; ~, проверка равенства средних при равных дисперсиях 135, 191, 210, 236, 250, 305; ~, проверка равенства дисперсий 135, 189, 194, 224, 254, 280, ~, проверка равенства средних при неравных дисперсиях 191, 257; ~, доверительные интервалы для разности средних значений 198; ~, доверительные интервалы для отношения дисперсий 199, 292; ~, критерий перестановок для разности средних значений 210, 218, 389; ~, сравнение полной рандомизации с методом сходных пар 230. См. также Берэнса — Фишера проблема; Однородности критерий; Стьюдента t -распределение

Нормальное распределение, дисперсия; ~, оптимальный критерий 111, 148, 185, 188, 244, 374, 387; ~, доверительные интервалы 198, 292; ~, последовательный критерий 224

Нормальное распределение, среднее; ~, отсутствие РНМ критерия 114; ~, критерий

- рий, основанный на выборке случайного объема 137, ~, оптимальный критерий 186, 224, 248, 256, 282, 305; ~, устойчивость критериев 188; ~, доверительные интервалы 195, 275; ~, медиано несмещенная оценка 197; ~, отсутствие критерия с контролируемой мощностью 222; ~, доверительные интервалы фиксированной длины 227; ~, критерии с контролируемой мощностью 228; ~, критерий и доверительные границы для отношения к стандарту 248, 250, 257, 281, 292; ~, последовательный критерий 281. См. также Попарные сравнения, Критерий перестановок; Стьюдента t -распределение; Симметрия
- Обратный выбор;** ~ для биномиальных испытаний 85; ~ для распределений Пуассона 85
- Огибающая (envelope) функция мощности** 283, 382
- Ограниченная полнота (bounded completeness)** 152; ~ и отсутствие полноты 171. См. также Полнота
- Одинарная классификация** 306; ~ в модели I 323
- Однородности критерий;** ~ для распределений Пуассона 178, 309; ~ для нормальных распределений 336, 309; ~ для биномиальных распределений 309
- Отношения эквивалентности** 391
- Отрицательно-биномиальное распределение** 31, 85
- Оценка медианно-несмещенная** 32; ~, связь с доверительными границами 98, 195; ~, примеры 197, 199
- Оценки см.** Доверительные множества, Максимальное правдоподобие; Оценка медианно-несмещенная, Несмещенность
- Ошибки первого и второго рода** 74
- Параметр сдвига (location parameter);** ~, устранение с использованием инвариантности 243, 279; ~, критерий 295, 372
- Планирование экспериментов** 15, 165, 318, 389. См. также Сравнение экспериментов
- Плотность вероятности** 46; ~, теорема сходности 395
- Подобный критерий** 143, 149; ~, связь с несмещенностью 144; ~, характеристика 152
- Поля распределение** 88, 132, 140
- Показательное распределение** см. Экспоненциальное распределение
- Полиномиальное распределение** 62; ~ как условное распределение 70; ~ в задаче оценки вкусов потребителя 168; ~ и χ^2 -критерий 343; ~, предельное распределение 344; ~, оценки максимального правдоподобия для параметров 362
- Полнота (completeness),** ~ семейства распределений 149; ~ биномиальных распределений 150; равномерных распределений 150, 171, ~ экспоненциальных семейств 150; ~ порядковых статистик 152, 172; ~, связь с ограниченной полнотой 152, 171
- Полный класс;** ~ решающих процедур 25, ~, связь с достаточностью 70; ~ односторонних критериев 86; ~ двусторонних критериев 171
- Положительная связь** см. Зависимость
- Попарная достаточность** 69
- Попарные сравнения (paired comparisons);** ~, теория для нормально распределенных величин и критерии перестановок 213, 230; ~, обобщение 215; ~, сравнение с методом полной рандомизации 230; ~ по качественному признаку 244, 373; ~, ранговые критерии 261, 270
- Пополнение (completion) меры** 41
- Порядковые статистики** 52; ~, эквивалентность суммам степеней 52; ~ как достаточная статистика 69, ~, полнота 152, 172; ~ в критериях перестановок 205; ~ как максимальный инвариант 241; ~, математические ожидания 265, ~, распределение 286
- Последовательные процедуры;** ~, их преимущества 15; ~ с гарантированной мощностью 159; ~ для сравнения двух распределений Пуассона 161. См. также Обратный выбор; Двухступенная процедура Стейна
- Последовательный критерий отношений вероятностей** 114, 117; ~, функция мощности 118, 120, 136; ~, объем выборки 119, 135; ~, оптимальные свойства 122; ~ для сравнения двух биномиальных распределений 174; ~ для сравнения двух дисперсий 224
- Последовательный t -критерий** 281
- Почти всюду** 46, 149
- Почти инвариантность;** ~ решающих процедур 32; ~ критериев 251; ~, связь с инвариантностью 251, 283; ~, связь с инвариантностью функции мощности 254; ~, соотношение между почти инвариантностью и несмещенностью 255; ~ отношения правдоподобия 284; ~ множеств 284; ~ связь с максимальными критериями 375
- Правдоподобия функция** 23
- Преобразование;** ~ интегралов 49, вероятностное интегральное 267; ~, стабилизирующее дисперсию 308; ~ логарифмическое 308, ~ арксинуса 356; ~ квадратного корня 356
- Преобразование с помощью функции распределения (probability integral transformation)** (переводящее произвольную случайную величину в равномерно распределенную) 267
- Приемочный контроль** см. Выборочный контроль
- Принцип минимакса** 21, 26; ~, связь с несмещенностью 33, 371; ~, связь с инвариантностью 33, 375; ~ при проверке гипотез 369. См. также Максимальный критерий
- Проблема с тремя решениями** 160
- Произведение мер** 46
- Простая гипотеза** 77
- Простая функция** 44
- Пространство решений**
- Процедура отбора** 131, 139
- Процедуры со многими решениями (multiple decision procedures)** 12; ~ для расположения по величине средних значений или дисперсий нормальных совокупностей 310, 311. См. также Проблема с тремя решениями
- Прямое произведение множеств** 46
- Прямоугольное распределение** см. Равномерное распределение
- Пуассона распределение** 11; ~, достаточные статистики 23; ~ как распределение суммы пуассоновских величин 70; ~, односторонний критерий для среднего значения 85, 129, ~, гипотезы о сумме параметров 134; ~, сравнение нескольких выборок 178, 309; ~, доверительные интер

- валы для отношения средних значений двух распределений 201; $\sim$, преобразование квадратного корня 356
- Пуассоновский процесс** 11; $\sim$, распределение времени ожидания 31, $\sim$, критерий для масштабного параметра 86, 129, $\sim$, сравнение экспериментов 92, 129; $\sim$, доверительные границы для масштабного параметра 96; $\sim$, сравнение двух процессов 163. См. также Экспоненциальное распределение
- Равномерно наиболее мощный (РНМ) См. Критерий**
- Равномерное распределение** 31; $\sim$, оценка среднего значения 16; $\sim$, достаточные статистики 29, 35, 171; $\sim$, критерий для масштабного параметра 128; $\sim$, полнота 150, 171; $\sim$ критического уровня 169; $\sim$, порядковые статистики 286; $\sim$, проверка равномерности распределения 345, 347
- Радона—Никодима производная** 46, $\sim$, свойства 67
- Радона—Никодима теорема** 46
- Размер критерия** 74
- Ранги** 241; $\sim$ как максимальный инвариант 241, 262; $\sim$, распределение 264, 285, 286, 287, 289, 290
- Ранговый коэффициент корреляции** 291
- Ранговый критерий** 260, 262; $\sim$ в задаче сравнения двух выборок 264, 284, 288; $\sim$ для гипотезы симметрии 270, 290; $\sim$ для гипотезы случайности 290, 291; $\sim$ для гипотезы независимости 291
- Рандомизация** 14; $\sim$ с целью понизить максимум риска 33; $\sim$, возможность отказа 128; $\sim$ как основа для статистических выводов 211; $\sim$, связь с критерием перестановок 212
- Рандомизированные; $\sim$ решающие процедуры** 14; $\sim$ инвариантные процедуры 34; $\sim$ критерии 75, 96; $\sim$ доверительные множества 97
- Распределение** См. распределения: Бета, Биномиальное, Гипергеометрическое, Двумерное нормальное, Коши, Многомерное нормальное, Нормальное, Отрицательно-биномиальное, Пойа, Полиномиальное, Пуассона, Равномерное, Симметричное экспоненциальное, Стюдента, T^2 -распределение, χ^2 -распределение, Экспоненциальное
- Распределение вероятностей случайной величины** 42
- Раслоенная выборка (stratified sample)** 205
- Регрессия; $\sim$, критерий для коэффициентов** 202, 299, 318, 357; $\sim$, доверительные интервалы 203, 320; $\sim$, доверительные множества для абсциссы линии 203; $\sim$ полиномиальная 320, 357, $\sim$, сравнение линий 320, 357; $\sim$ в анализе ковариаций 322; $\sim$ в многомерном распределении 360
- Решающее правило** 9
- Сдвиг (shift); $\sim$, доверительные интервалы, основанные на критериях перестановок** 229; $\sim$, основанные на ранговых критериях 288 ($\sim$, критерии см. Задача сравнения двух выборок, непараметрическая)
- Семейство распределений с параметром сдвига; $\sim$ является стохастически возрастающим** 88, $\sim$, условие для того, чтобы оно имело монотонное отношение правдоподобия 372
- Епарабельное пространство** 393
- Серии (runs); $\sim$ для проверки гипотезы независимости в цепи Маркова** 175; $\sim$, распределение 175 **σ -конечная мера** 41
- σ -поле** 41; $\sim$ со счетным числом образующих 397
- Симметричное экспоненциальное распределение (double exponential)** 372, 387
- Симметрия; $\sim$, связь с инвариантностью** 18, 237; $\sim$, достаточные статистики для распределений, обладающих симметрией 68; $\sim$, гипотеза симметричности относительно начала координат 168, 229, 262, 270, 290, 386
- Сложная гипотеза** 77
- Случайная величина** 42; $\sim$, связь со статистикой 43, 44
- Случайность; $\sim$, гипотеза** 261; $\sim$, критерий 290, 291
- Смешанная модель (mixed model)** 323; $\sim$ для «гнездовой» классификации 330; $\sim$ с двумя факторами 331, 340
- Смирнова критерий в задаче сравнения двух выборок** 269, 278
- Совласия критерий** 277. См. также χ^2 -критерий
- Состоятельные последовательности критериев** 344
- Сравнение экспериментов** 90, 130; $\sim$ для проверки независимости в 2×2 таблицах 91, 162; $\sim$ по выборочному контролю 93, 281; $\sim$ по оценке параметра пуассоновского процесса 129; $\sim$ по проверке вкуса потребителей 167; $\sim$ по анализу регрессии 203; $\sim$ по сопоставлению двух способов обработки 230. См. также Планирование экспериментов; Выборки объем
- Средняя мощность наилучшая** 353
- Статистика** 43, $\sim$, подполе, индуцированное ею 47; $\sim$, эквивалентные представления 49; $\sim$, вполне информирующая (fully informative) 128
- Стационарность** 174
- Стохастический процесс** 139. См. также Пуассоновский процесс
- Стохастическое возрастание** 87, 261
- Строго (strictly) несмещенный критерий** 146
- Стюдента t -критерий; $\sim$ как критерий отношения правдоподобия** 34; $\sim$ как РНМ критерий 114; $\sim$ как РНМ несмещенный критерий 186, 192; $\sim$, устойчивость 188, 194; $\sim$, связь с критерием перестановок 211; $\sim$, мощность 222, 223; $\sim$ как РНМ инвариантный критерий 250; $\sim$, допустимость 260; $\sim$, эффективность по сравнению с ранговыми критериями 268. См. также Последовательный t -критерий
- Стюдента t -распределение** 186; $\sim$ нецентральное 187, 223, 249, 281; $\sim$ в задаче сравнения двух выборок 193; $\sim$ в регрессионном анализе 202; $\sim$ в методе попарных сравнений 214, 230; $\sim$ в процедуре Стейна 227; $\sim$ как условное распределение функции от коэффициента корреляции 231; $\sim$ в линейной гипотезе с одним ограничением 302. См. также Нормальное распределение, среднее значение; Нормальное распределение, две выборки
- Студентов размах (studentized range)** 310
- Сходимость; $\sim$ в среднем** 392; $\sim$ слабая 393; $\sim$ поточечная (pointwise) 393
- Счетная аддитивность** 40
- Считающая (counting) мера** 42
- Таблица сопряженности признаков** 91, 131, 162, 362

Теорема о слабой компактности 398
Теорема факторизации для достаточных статистик 28, 35, 61
Теорема о сходимости; ~ ограниченной 45; ~ монотонной 45; ~ Крамера 188; ~ плотностей 395
Транзитивная группа преобразований 241
T²-распределение 338
Тренд; ~, критерий отсутствия 290, 291

Упорядочение решающих процедур по предпочтению 17, 21
Уровень значимости 74
Условная вероятность 53
Условная мощность; ~, интерпретация 158
Условное математическое ожидание 51; ~, свойства 53
Условное распределение 54
Условный критерий 155, 177
Устойчивость критериев 188, 194

Фишера — Ирвина критерий независимости в 2 × 2 таблицах 165
Фишера — Изйтса критерий 265, 267, 268
F-распределение 190, ~ при построении доверительных интервалов и критериев для отношения дисперсий 190, 199; ~, связь с бета-распределением 191; ~ при проверке линейных гипотез 302; ~ нецентральное 302, 331, 343, 352; ~ в модели II дисперсионного анализа 326, 329; ~, связь с T²-распределением 338; ~, связь с распределением множественного коэффициента корреляции 361
Фубини теорема 47
Фундаментальная лемма Неймана — Пирсона 77; ~ обобщенная 99, 141
Функция мощности 74; ~ одностороннего критерия 82, 131; ~ двустороннего критерия 105; ~ последовательного критерия отношений вероятностей 118, ~, оценка 158, ~ инвариантного критерия 254
Функция потерь 9; ~, ее выбор 13; ~ для одностороннего критерия 86, 246; ~ для доверительных границ 93, 131; ~ для доверительных интервалов 98, 203, ~ для

медианно-несмещенных оценок 99; ~ для двусторонних критериев 171
Функция распределения 42; ~ эмпирическая 269; ~ доверительные границы 276
Функция риска 10

Ханта — Стейна теорема 378, 388
Хотеллинга T²-критерий См. T²-распределение; T²-критерий
χ²-распределение 62; ~, проверка гипотез о масштабном параметре 148, 170; ~ при проверке однородности 309; ~ в линейной гипотезе с известной матрицей ковариаций 342, ~ нецентральное 343, 352, 356; ~ в многомерном анализе 360. См. также Испытания на продолжительность жизни; Нормальное распределение, дисперсия, Пуассоновский процесс; Экспоненциальное распределение

Эквивалентность; ~ распределений 60; ~ семейств мер 397; ~ статистик 49
Экспоненциальное распределение 31; ~, связь с пуассоновским процессом 31, 85, 161; ~, достаточные статистики 35; ~, порядковые, статистики выборки 70; ~, доверительные границы и критерии для параметра сдвига и масштабного параметра 96, 128, 225, 228; ~, задача сравнения двух выборок 280, 297. См. также Испытания на продолжительность жизни
Экспоненциальное семейство распределений 62; ~, естественное параметрическое пространство 63, 71; ~, аналитичность некоторых интегралов 65; ~, моменты достаточных статистик 71; ~, критерии в однопараметрическом семействе 84; ~, полнота 150; ~, критерии в многопараметрическом семействе 153, 170; ~, эквивалентная запись 158; ~, допустимость критериев 260
Эмпирическая (выборочная) функция распределения 269, 277
Эффективность; ~ относительная асимптотическая 268