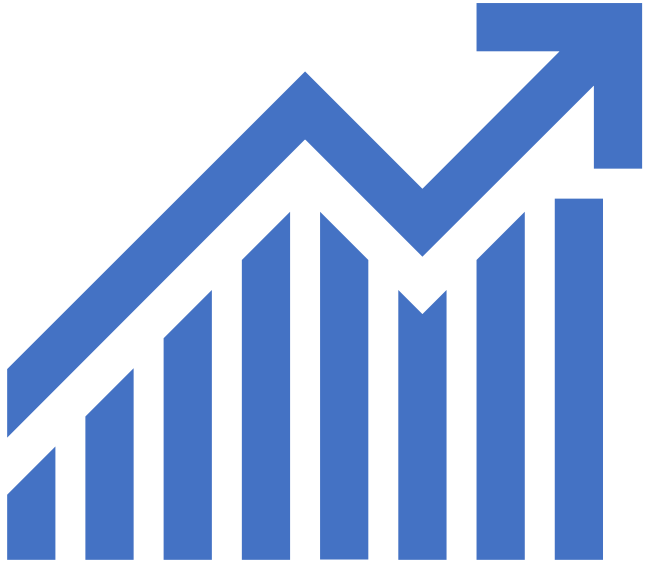
ТЕОРИЯ РИСКА



лекции

Поспелова Ирина Игоревна

Распределения случайных величин

Базовые вероятностные модели для актуарных ситуаций, как правило, являются либо непрерывными, либо дискретными (т.е. не смешанными).

Требуются для подсчета чего-либо дискретного (число страховых случаев, число выплат) либо для определения объемов платежей (отдельные выплаты, сумма выплат).

В моделях не требуется учитывать отрицательные значения.

Набор всех возможных функций распределения слишком велик для понимания. Поэтому при поиске функции распределения для использования в качестве модели случайного явления может быть полезно сузить множество возможных вариантов. Ранее рассматривали классификацию распределений степени тяжести хвоста.

Далее рассмотрим классификацию в зависимости от сложности модели. Сначала для множества непрерывных моделей, а затем дискретных.

Параметры моделей

- Число необходимых для задания модели количественных характеристик (параметров) является индикатором ее сложности.

Простая модель:

- малое число элементов, для которых необходима оценка, определяется более точно.
- модель более устойчива во времени и в разных настройках.
- сглаживание нерегулярных данных.

Сложная модель:

- большое число специфицируемых параметров обеспечивает более точное соответствие действительности.
- позволяет учитывать нерегулярность в данных.

Параметры более простых моделей могут быть оценены более точно, но сама модель может быть слишком грубой.

Следует использовать самую простую модель, адекватно отражающую реальность. «Адекватность» зависит от цели применения модели.

Далее рассмотрим различные категории вероятностных моделей, определяемых функциями распределения.

Параметрические и масштабируемые распределения

Определение 1. Параметрическим распределением называется множество функций распределения, каждая из которых определяется указанием одного или более значений, называемых *параметрами*. Число параметров фиксировано и конечно.

Самое известное параметрическое распределением – нормальное с параметрами μ и $\sigma 2$.

Чем меньше параметров требуется задать, тем проще модель.

Для многих актуарных задач удобно, чтобы распределение не менялось, когда случайная величина умножается на константу. Наиболее распространенное использование этого в моделях – учет эффекта инфляции и изменения денежной единицы.

Определение 2. Параметрическое распределение называется масштабируемым, если случайная величина, подчиняющаяся этому распределению, умноженная на положительную константу, также подчиняется этому распределению.

Пример 1

Покажем, что экспоненциальное распределение является масштабируемым

Функция распределения:

$$F_X(x) = 1 - e^{-\frac{x}{\theta}}.$$

Пусть $Y = cX,\ c > 0$. Тогда

$$F_Y(y) = \mathrm{P}\{Y \le y\} = \mathrm{P}\{cX \le y\} = \mathrm{P}\left\{X \le \frac{y}{c}\right\} = 1 - e^{-\frac{y}{c\theta}}.$$

Получили экспоненциальное распределение с параметром $c\theta$.

Определение 3. Когда случайная величина с неотрицательным носителем, подчиняющаяся масштабируемому распределению, умножается на положительную величину:

1. параметр масштаба умножается на ту же величину.
2. другие параметры не меняются.

Пример 2

Покажем, что гамма-распределение имеет параметр масштаба.

Пусть X подчиняется гамма-распределению, $Y = cX$. Тогда

$$F_Y(y) = \mathrm{P}\left\{X \leq \frac{y}{c}\right\} = \Gamma\left(\alpha; \frac{y}{c\theta}\right)$$

Таким образом, Y имеет гамма распределение с параметрами α и $c\theta$. Поэтому θ - параметр масштаба.

Нередко плотность гамма распределения записывается в виде

$$f(x) = \frac{\beta\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0$$

Мы используем вид функции плотности, в котором параметр масштаба θ, равен $1/\beta$.

Параметр масштаба легко распознать по виду функции или плотности распределения. В частности, в функции распределения x появляется только вместе с параметром θ в виде x/θ.

Семейства параметрических распределений

Несколько более сложными являются параметрические распределения, в которых число параметров конечно, но не фиксировано заранее

Определение 4. Семейство параметрических распределений – это множество параметрических распределений, которые связаны каким-то осмысленным образом.

Некоторые типы семейств параметрических распределений основаны на заданном параметрическом распределении. Другие члены семейства получаются в результате задания одного или более параметров.

Рассмотрим наиболее распространённый тип параметрического семейства.

Пример 3

Покажем, что преобразованное бета семейство является семейством параметрических распределений.

$$f(x) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x/\theta)^{\gamma\tau}}{x(1+(x/\theta)^{\gamma})^{\alpha+\tau}}$$

Преобразованное бета распределение имеет 4 параметра. Каждое распределение в семействе получается при определенных значениях параметров:

при $\gamma = \tau = 1$ получаем распределение Парето:

$$f(x) = \frac{\alpha\theta^{\alpha}}{(\theta+x)^{\alpha+1}}$$

при $\tau = 1$ and $\gamma = \alpha$ паралогистическое:

$$f(x) = \frac{\alpha^2(x/\theta)^{\alpha}}{x(1+(x/\theta)^{\alpha})^{\alpha+1}}$$

Заметим, что число параметров (от 2 до 4) заранее не известно.

Если нужно подобрать подходящее преобразованное бета-распределение, следует искать все 4 параметра по диапазонам возможных значений.

Если требуется использовать семейство преобразованных бета-распределений, рассматривается возможность использовать специальные случаи.

Например, если получено значение $\tau = 1.01$, в первом случае оно и будет использоваться. Во втором можно заметить, что $\tau = 1$ приводит к распределению Бурра, и далее использовать более простой вариант

Семейство бета-распределений

- Преобразованное бета-распределение $f(x) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x/\theta)^{\gamma\tau}}{x(1+(x/\theta)^{\gamma})^{\alpha+\tau}}$
- Преобразованное распределение Парето $f(x) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\theta^{\alpha}\, x^{\tau-1}}{(x+\theta)^{\alpha+\tau}}$
- Распределение Бурра $f(x) = \frac{\alpha\gamma(x/\theta)^{\gamma}}{x(1+(x/\theta)^{\gamma})^{\alpha+1}}$
- Обратное распределение Бурра $f(x) = \frac{\tau\gamma(x/\theta)^{\gamma\tau}}{x(1+(x/\theta)^{\gamma})^{\tau+1}}$
- Распределение Парето $f(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}$
- Обратное распределение Парето $f(x) = \frac{\tau\theta^{\tau-1}}{(x+\theta)^{\tau+1}}$
- Логлогистическое $f(x) = \frac{\gamma(x/\theta)^{\gamma}}{x(1+(x/\theta)^{\gamma})^{2}}$
- Паралогистическое $f(x) = \frac{\alpha^2(x/\theta)^{\alpha}}{x(1+(x/\theta)^{\alpha})^{\alpha+1}}$
- Обратное паралогистическое $f(x) = \frac{\tau^2(x/\theta)^{\tau^2}}{x(1+(x/\theta)^{\tau})^{\tau+1}}$

Конечные смеси распределений

- Одна из причин смешивания состоит в том, что моделируемое явление само по себе является смесью явлений с неизвестными вероятностями. Например, случайно выбранный иск по возмещению оплаты стоматологических услуг может касаться профилактического осмотра, пломбирования, протезирования, хирургических вмешательств. Распределения разных видов помощи могут быть похожими, но иметь разные моды.
- **Определение 5.** Пусть $\{X_1, X_2, \dots, X_k\}$ – множество случайных величин с соответствующими функциями распределений $\{F_{X_1}, F_{X_2}, \dots, F_{X_k}\}$. Случайная величина Y является *k*-точечной смесью случайных величин $\{X_1, X_2, \dots, X_k\}$, если ее функция распределения имеет вид

$$F_Y(y) = a_1 F_{X_1}(y) + a_2 F_{X_2}(y) + \cdots + a_k F_{X_k}(y)$$

  где $a_j > 0$ и $a_1 + a_2 + \cdots + a_k = 1$.
- Это не совместное распределение $\{X_1, X_2, \dots, X_k\}$!
- Распределение Y зависит только от функций распределения $\{X_1, X_2, \dots, X_k\}$.
- Смесь приписывает вероятность a_j тому исходу Y, который реализуется случайной величиной X_j .
- Если есть 20 испытаний для данной случайной величины, то двухточечная смесь дает возможность создать 200 более новых распределений ($\binom{20}{2}$+20).

Пример 4

Для модели общего страхования ответственности можно рассмотреть смесь двух распределений Парето. Эта модель содержит 5 параметров. Однако если 5 параметров кажутся избыточными, можно выбрать следующей вариант:

$$F(x) = 1 - a\left(\frac{\theta_1}{\theta_1 + x}\right)^{\alpha} - (1 - a)\left(\frac{\theta_2}{\theta_2 + x}\right)^{\alpha+2}$$

Параметр формы α в двух распределениях отличается на 2. Второе распределение приписывает большую вероятность меньшим значениям (более легких хвост). Это может быть модель для частот, небольших возмещений, в то время как первое распределение относится к большим, но нечастым убыткам. Такое распределение имеет 4 параметра, ограничивая излишнюю свободу в выборе параметров.

Переменные смеси

- Если число распределений в смеси заранее не известно, то k становится параметром
- **Определение 6.** Смесь распределений с переменным числом компонент имеет функцию распределения следующего вида:

$$F(x) = \sum_{j=1}^{K} a_j F_j\,(x)\,, \qquad \sum_{j=1}^{K} a_j = 1, a_j > 0;\ j = 1, 2, \dots, K, \qquad K = 1, 2, \dots$$

- Такие модели называются полупараметрическими, они сложнее параметрических и проще непараметрических.
- Если все компоненты имеют одно и то же параметрическое распределение (но разные значения параметров), результирующее распределение называется *переменной смесью* соответствующий распределений

Пример 5

- Определим функцию распределения, плотность распределения и уровень риска для переменной смеси экспоненциальных распределений. Смесь экспоненциальных распределений:

$$F(x) = 1 - a_1 e^{-\frac{x}{\theta_1}} - a_2 e^{-\frac{x}{\theta_2}} - \cdots - a_K e^{-\frac{x}{\theta_K}}, \qquad \sum_{j=1}^{K} a_j = 1, a_j, \theta_j > 0;\ j = 1, \ldots, K, K = 1, 2, \ldots$$

- Тогда функция плотности

$$f(x) = a_1 \theta_1^{-1} e^{-\frac{x}{\theta_1}} + a_2 \theta_2^{-1} e^{-\frac{x}{\theta_2}} + \cdots + a_K \theta_K^{-1} e^{-\frac{x}{\theta_K}}$$

$$h(x) = \frac{a_1 \theta_1^{-1} e^{-\frac{x}{\theta_1}} + a_2 \theta_2^{-1} e^{-\frac{x}{\theta_2}} + \cdots + a_K \theta_K^{-1} e^{-\frac{x}{\theta_K}}}{a_1 e^{-\frac{x}{\theta_1}} + a_2 e^{-\frac{x}{\theta_2}} + \cdots + a_K e^{-\frac{x}{\theta_K}}}$$

- Число параметров не фиксировано и даже не ограничено. Например, когда K = 2, есть три параметра $(a_1;\ \theta_1;\ \theta_2)$. Однако, если K = 4, то будет 7 параметров.

Пример 6

- Покажем, как двухточечная смесь гамма распределений может создавать бимодальное распределение.
- Рассмотрим смесь 50–50 двух гамма распределений.
- Пусть одно имеет параметры

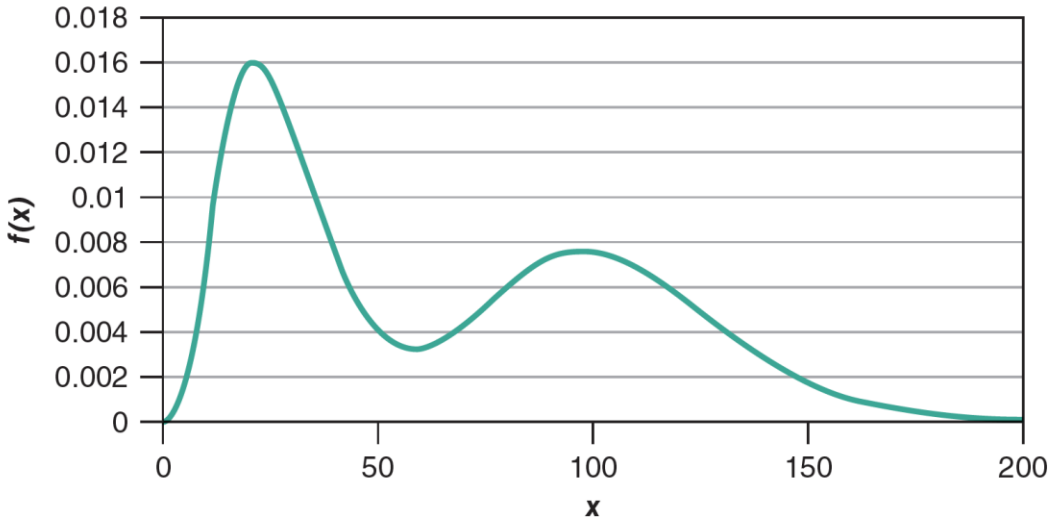
α = 4 и θ = 7 (мода $\theta(\alpha - 1)$ =21) и

другое имеет параметры

α = 15 и θ = 7 (мода 98).

Функция плотности

$$f(x) = 0.5\frac{x^3 e^{-\frac{x}{7}}}{3!\,7^4} + 0.5\frac{x^{14} e^{-x/7}}{14!\,7^{15}}$$



Распределения, зависящие от данных

- Модели 1-5 и многие другие примеры основываются на соответствующем явлении (случайной величине), а не на наблюдения за явлением. Например, в отсутствие каких-либо наблюдений за исками по результатам оказания стоматологической помощи, мы можем предположить логнормальное распределение с параметрами μ = 5 и σ = 1 или, возможно, смесь двух логнормальных распределений. Такая модель может давать плохое описание стоматологических убытков, но это другой вопрос. Существует другой тип модели, который, в отличие от логнормального, требует данных. Эти модели также имеют параметры, но часто называются непараметрическими.
- **Определение 7.** Зависящее от данных распределение является по крайней мере настолько же сложным, как и сами данные или знания, которые его породили. Число «параметров» увеличивается при увеличении числа точек или количества знания.
- Эти модели имеют столько же (или более) параметров, сколько наблюдений в наборе данных.
- **Определение 8.** Эмпирическая модель – это дискретное распределение, основанное на выборке размера n, в которой вероятность $1/n$ приписана каждой точке

Пример 7

- Рассмотрим выборку размера 8, в которой наблюдаются значения 3, 5, 6, 6, 6, 7, 7, 10. Эмпирическая модель имеет функцию вероятности

$$p(x) = \begin{cases} 0.125, & x = 3 \\ 0.125, & x = 5 \\ 0.375, & x = 6 \\ 0.25, & x = 7 \\ 0.125, & x = 10 \end{cases}$$

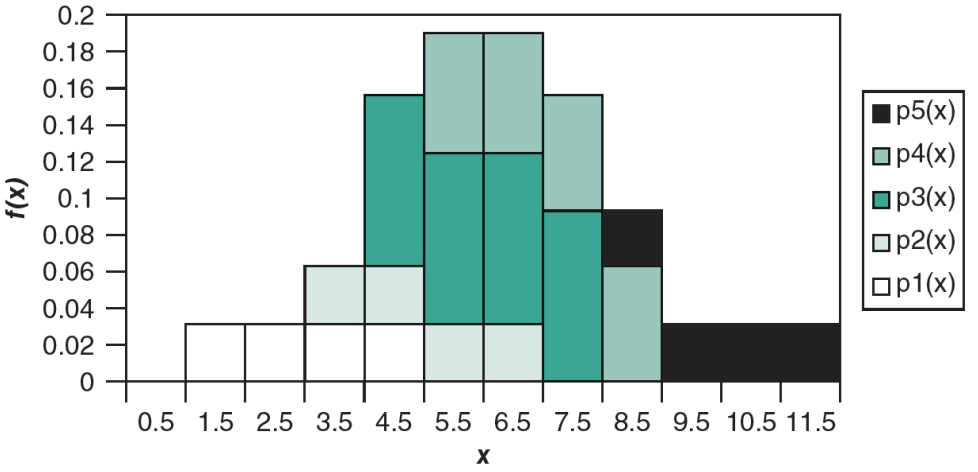
- Заметим, что дискретные модели с конечным носителем выглядят как эмпирические. Модель 3 могла бы быть эмпирической для выборки размера 10, содержащий 50 нулей, 25 единиц, 12 двоек, 8 троек и 5 четверок. Но термин «эмпирическая модель» используется только, когда имеется настоящая выборка.
- Эмпирическое распределение зависит от данных. Каждое значение дает вклад $1/n$ в функцию вероятности, поэтому n параметров – это n наблюдений в наборе данных, которые привели к эмпирическому распределению.
- Другим примером моделей, зависящих от данных, являются аппроксимация плотности распределения с помощью *ядерных функций*. Вместо того, чтобы поставить вероятность $1/n$ каждому данному, непрерывная функция плотности с областью, имеющую площадь $1/n$, заменяет точку данных. Область центрируется так, что модель похожа на данные, но не идеально. Это обеспечивает сглаживание в отличие от эмпирического распределения.

Пример 8

Построим аппроксимацию с помощью ядерных функций по данным Примера 7, используя однородное ядро и ширину полосы 2.
Функция плотности вероятности

$$f(x) = \sum_{j=1}^{5} p_8(x_j) K_j\ (x)$$

$$K_j(x) = \begin{cases} 0, |x\ -\ x_j| > 2, \\ 0.25, |x\ -\ x_j\ | \leq 2 \end{cases}$$



Сумма берется по 5 точкам, в которых исходная модель имеет положительную вероятность.
Например, первое слагаемое в сумме - это функция

$$p_8(x_1)K_1(x)\ = \begin{cases} 0, x\ <\ 1 \\ 0.03125, 1\ \leq\ x\ \leq\ 5 \\ 0, x\ >\ 5 \end{cases}$$

И аппроксимация с помощью ядерных функций, и эмпирическое распределение могут быть представлены в виде смеси распределений. Причина, по которой эти модели рассматриваются отдельно, состоит в том, что число компонент связано с размером выборки, а не и с самим явлением и числом случайных величин

Получение новых распределений

- В актуарных приложения основной интерес представляют распределения с положительным носителем, т.е. $F(0) = 0$. Таким свойством обладают экспоненциальное, гамма, Парето и логнормальное. Для любого распределения можно построить новые в результате преобразований или других математических манипуляций. Рассмотрим такие методы на примерах.
- **1. Умножение на константу**
- Это преобразование эквивалентно применению инфляции равномерно к убыткам всех уровней и соответствует изменению масштаба. Например, если в этом году убытки заданы случайной величиной X, тогда равномерная инфляция 5% показывает, что в следующем году убытки описываются случайной величиной $Y = 1.05X$.
- **Теорема 1.** Пусть X - непрерывная случайная величина с плотностью распределения $f_X(x)$ и функцией распределения $F_X(x)$. Пусть $Y = \theta X$, $\theta > 0$. Тогда $F_Y(y) = F_X\left(\frac{y}{\theta}\right)$,

$$f_Y(y) = \frac{1}{\theta} f_X\left(\frac{y}{\theta}\right)$$

Очевидно, что θ – параметр масштаба

Пример 9

- Пусть X имеет плотность распределения $f(x) = e^{-x}, x > 0.$
- Определим функцию и плотность распределения $Y = \theta X$:

$$F_X(x) = 1 - e^{-x}, \qquad F_Y(y) = 1 - e^{-\frac{y}{\theta}}, \qquad f_Y = \frac{1}{\theta} e^{-y/\theta}$$

Возведение в степень

Теорема 2. Пусть X - непрерывная случайная величина с плотностью $f_X(x)$ и функцией распределения $F_X(x), F_X(0) = 0$. Пусть $Y = X^{1/\tau}$. Тогда

если $\tau > 0$, $F_Y(y) = F_X(y^\tau)$, $f_Y(y) = \tau\, y^{\tau-1}\, f_X(y^\tau)$, $y > 0$

если $\tau < 0$, $F_Y(y) = 1 - F_X(y^\tau)$, $f_Y(y) = -\tau\, y^{\tau-1}\, f_X(y^\tau)$.

Доказательство.

Если $\tau > 0$, $F_Y(y) = P\{X \le y^\tau\} = F_X(y^\tau)$. Если $\tau < 0$, $F_Y(y) = P\{X \ge y^\tau\} = 1 - F_X(y^\tau)$

Так как обычно параметры считаются положительными, то при $\tau < 0$ после замены переменных получаем $F_Y(y) = 1 - F_X(y^{-\tau})$, $f_Y(y) = \tau\, y^{-\tau-1}\, f_X(y^{-\tau})$

Определение 9. При возведении в степень, если $\tau > 0$, полученное распределение называется преобразованным; если $\tau = -1$, то обратным; если $\tau < 0$ (но не равно -1), то обратным преобразованным.

Пример 10

Пусть X имеет экспоненциальное распределение. Определим функцию распределения обратного, преобразованного и обратного преобразованного экспоненциального распределения.

Обратное экспоненциальное распределение без параметра масштаба имеет функцию распределения $F(y) = 1 - \left[1 - e^{-\frac{1}{y}}\right] = e^{-\frac{1}{y}}$. С параметром масштаба $F(y) = e^{-\theta/y}$.

Преобразованное экспоненциальное распределение без параметра масштаба $F(y) = 1 - e^{-y^{\tau}}$.

Если есть параметр масштаба, то $F(y) = 1 - e^{-(y/\theta)^{\tau}}$. Получили известное распределение Вейбулла.

Обратное преобразованное экспоненциальное распределение без параметра масштаба имеет функцию распределения

$$F(y) = 1 - [1 - e^{-y^{-\tau}}] = e^{-y^{-\tau}}$$

Если добавить параметр масштаба, получаем $F(y) = e^{-(\theta/y)^{\tau}}$ - обратное распределение Вейбулла.

Другое базовое распределение имеет плотность распределения $f(x) = x^{\alpha-1}e^{-x}/\Gamma(\alpha)$. Добавление параметра масштаба приводит к гамма-распределению. Оно имеет обратную и преобразованную версию, которые могут быть получены аналогично.

Неполная гамма-функция

Определение 10. Неполная гамма функция с параметром $\alpha > 0$

$$\Gamma(\alpha;\, x) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t}\, dt,$$

где

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt.$$

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1),$$

для натурального n $\Gamma(n) = (n-1)!$

Экспоненцирование

Теорема 3. Пусть X - непрерывная случайная величина с функцией плотности $f_X(x)$ и функцией распределения $F_X(x)$, $f_X(x) > 0 \ \forall \ x$. Пусть $Y = e^X$. Тогда для $y > 0$

$$F_Y(y) = F_X(\ln y), \qquad f_Y(y) = \frac{1}{y} f_X(\ln y)$$

Пример 11.

- Пусть X имеет нормальное распределение со средним μ и дисперсией σ^2. Определим функцию распределения и плотность распределения $Y = e^X$.

$$F_Y(y) = \Phi\left(\frac{\ln y - \mu}{\sigma}\right),$$

$$f_Y(y) = \frac{1}{y\sigma}\,\phi\left(\frac{\ln y - \mu}{\sigma}\right) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right]$$

- Получили логнормальное распределение.

Смешивание

Концепция смешивания может быть расширена со смеси конечного числа случайных переменных на несчетное число.

Далее функция плотности $f_\Lambda(\lambda)$ играет роль дискретных вероятностей a_j *в* k-точечной смеси.

Теорема 4. Пусть X имеет плотность распределения $f_{X|\Lambda}(x|\lambda)$ и функцию распределения $F_{X|\Lambda}(x|\lambda)$, где λ – параметр распределения X. Другие параметры не имеют значения. Пусть λ - реализация случайной величины Λ с плотностью распределения $f_\Lambda(\lambda)$. Тогда безусловная функция плотности случайной величины X

$$f_X(x) = \int f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda,$$

где интеграл берется по всем значениям λ, имеющим положительную вероятность.

Получившееся распределение является смесью распределений

Смешивание

Функция смеси распределений

$$F_X(x) = \int_{-\infty}^{x} \int f_{X|\Lambda}(y|\lambda) f_\Lambda(\lambda) d\lambda \, dy = \int \int_{-\infty}^{x} f_{X|\Lambda}(y|\lambda) f_\Lambda(\lambda) dy \, d\lambda = \int F_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda$$

Моменты смеси распределений:

$$E\left[X^k\right] = E\left[E\left[X^k|\Lambda\right]\right],$$

$$Var[X] = E[Var[X|\Lambda]] + Var[E[X|\Lambda]]$$

Как правило, смесь распределений, имеет тяжелые хвосты, поэтому такой способ является удобным при необходимости учитывать вероятности редких событий.

Если $f_{X|\Lambda}(x|\lambda)$ имеет убывающую функцию уровня риска для всех λ, тогда смесь распределений также имеет убывающую функцию уровня риска.

Пример 12

Пример показывает, как распределение с тяжелым хвостом получается из смеси распределений.

Пусть $X|\Lambda$ имеет экспоненциальное распределение с параметром $1/\Lambda$.

Пусть Λ имеет гамма распределение. Определим безусловное распределение X.

$$f_X(x) = \frac{\theta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda e^{-\lambda x} \lambda^{\alpha-1} e^{-\theta\lambda} d\lambda = \frac{\theta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda(x+\theta)} \lambda^\alpha d\lambda = \frac{\theta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{(x+\theta)^{\alpha+1}} = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}$$

Получили распределение Парето.

Пример 13

Предположим, что для данного $\Theta = \theta$ X имеет нормальное распределение со средним θ и дисперсией υ:

$$f_{X|\Theta}(x|\theta) = \frac{1}{\sqrt{2\pi\upsilon}} \exp\left[-\frac{1}{2\upsilon}(x-\theta)^2\right], -\infty < x < \infty,$$

и пусть Θ само имеет нормальное распределение со средним μ и дисперсией a:

$$f_{\Theta}(\theta) = \frac{1}{\sqrt{2\pi a}} \exp\left[-\frac{1}{2a}(\theta-\mu)^2\right], -\infty < \theta < \infty$$

Функция плотности

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\upsilon}} \exp\left[-\frac{1}{2\upsilon}(x-\theta)^2\right] \frac{1}{\sqrt{2\pi a}} \exp\left[-\frac{1}{2a}(\theta-\mu)^2\right] d\theta = \frac{1}{2\pi\sqrt{\upsilon a}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\upsilon}(x-\theta)^2 - \frac{1}{2a}(\theta-\mu)^2\right] d\theta$$

$$= \frac{\exp\left[-\frac{(x-\mu)^2}{2(a+\upsilon)}\right]}{\sqrt{2\pi(\upsilon+a)}} \int_{-\infty}^{\infty} \sqrt{\frac{a+\upsilon}{2\pi\upsilon a}} \exp\left[-\frac{a+\upsilon}{2a\upsilon}\left(\theta - \frac{ax+\upsilon\mu}{a+\upsilon}\right)^2\right] d\theta = \frac{1}{\sqrt{2\pi(\upsilon+a)}} \exp\left[-\frac{(x-\mu)^2}{2(a+\upsilon)}\right]$$

так как подынтегральное выражение есть плотность нормального распределения. В результате получилось также нормальное распределение со средним μ и дисперсией $a + \upsilon$

Пример 14

Пример показывает, как появляются смеси распределений. В частности, непрерывные смеси часто используются в качестве модели для неопределенного параметра. Точное значение параметра неизвестно, но для описания параметра можно выяснить функцию плотности

При оценке гарантии на автомобили важно учитывать, что километраж зависит от водителя. Но для конкретного водителя километраж может меняться по годам. Предположим, что километраж для случайно выбранного водителя имеет обратное распределение Вейбулла, но изменение параметра масштаба от года к году имеет преобразованное гамма-распределение с тем же значением τ. Определим распределение километража за случайно выбранный год случайно выбранного водителя.

$$f_X(x) = \int_0^\infty \frac{\tau\lambda^\tau}{x^{\tau+1}} e^{-\left(\frac{\lambda}{x}\right)^\tau} \frac{\tau\lambda^{\tau\alpha-1}}{\theta^{\tau\alpha}\Gamma(\alpha)} e^{-\left(\frac{\lambda}{\theta}\right)^\tau} d\lambda = \frac{\tau^2}{\theta^{\tau\alpha}\Gamma(\alpha)x^{\tau+1}} \int_0^\infty \lambda^{\tau+\tau\alpha-1} \exp[-\lambda^\tau (x^{-\tau}+\theta^{-\tau})] \quad d\lambda$$

$$= \{y = \lambda^\tau(x^{-\tau}+\theta^{-\tau})\} = \frac{\tau}{\theta^{\tau\alpha}\Gamma(\alpha)x^{\tau+1}(x^{-\tau}+\theta^{-\tau})^{\alpha+1}} \int_0^\infty y^\alpha e^{-y} dy = \frac{\tau\alpha\theta^\tau x^{\tau\alpha-1}}{(x^\tau+\theta^\tau)^{\alpha+1}}$$

Получили обратное распределение Бурра. Заметим, что это распределение применимо к конкретному водителю. Другой водитель может иметь другой параметр формы τ распределения Вейбулла

Модели уязвимости (Frailty models)

- Важным типом распределения смеси является модель уязвимости (нестабильности). Хотя этот тип смеси изначально использовался для моделирования распределений продолжительности жизни в анализе выживаемости, оказалось, что такой подход также можно рассматривать как полезный способ генерирования новых распределений путем смешивания.
- Введем случайную величину уязвимости $\Lambda > 0$ и определим условный (при заданном $\Lambda = \lambda$) уровень риска X:

$$h_{X|\Lambda}(x|\lambda) = \lambda a(x),$$

где $a(x)$ - заданная функция (зависит от конкретного приложения).

Уязвимость предназначена для количественной оценки неопределенности, связанной со степенью риска, которая

действует мультипликативно. Тогда условная функция выживания $X|\Lambda$ равна

$$S_{X|\Lambda}(x|\lambda) = e^{-\int_0^x h_{X|\Lambda}(t|\lambda)\,dt} = e^{-\lambda A(x)},$$

где $A(x) = \int_0^x a(t)dt.$

Модели уязвимости (Frailty models)

Чтобы найти смесь распределений (т.е. безусловное распределение X), определим производящую функцию моментов случайной величины уязвимости Λ как $M_\Lambda(z) = E[e^{z\Lambda}]$. Тогда

$$S_X(x) = E[e^{-\Lambda A(x)}] = M_\Lambda[-A(x)]$$

и $F_X(x) = 1 - S_X(x)$.

Тип используемой смеси определяется выбором $a(x)$ и, значит, $A(x)$.

Наиболее важный подкласс моделей уязвимости – это класс экспоненциальных смесей

с $a(x) = 1$ и $A(x) = x$, так что $S_{X|\Lambda}(x|\lambda) = e^{-\lambda x}, \; x \geq 0.$

Другая полезная смесь – Вейбулла с $a(x) = \gamma x^{\gamma-1}$ и $A(x) = x^\gamma$.

Для оценки распределения уязвимости требуется выражение для производящей функции моментов $M_\Lambda(z)$.

Наиболее общий выбор – гамма уязвимости, но и другие варианты, например, обратное распределение Гаусса, также используются.

Пример 15

Пусть Λ имеет гамма распределение и пусть $X|\Lambda$ имеет распределение Вейбулла с условной функцией выживания $S_{X|\Lambda}\,(x|\lambda) \;=\; e^{-\lambda x^{\gamma}}$. Определим безусловное распределение X.

- Производящая функция моментов гамма-распределения $M_{\Lambda}(z) \;=\; (1-\theta z)^{-\alpha}$, тогда X имеет функцию выживания

$$S_X(x) \;=\; M_{\Lambda}(-x^{\gamma}) \;=\; (1 \;+\; \theta x^{\gamma})^{-\alpha}$$

- Это распределение Бурра с параметром $\theta^{-1/\gamma}$. Заметим, что при $\gamma \;=\; 1$, получим экспоненциальную смесь, которая является распределением Парето (см. Пример 12)

Склейка (splicing)

Еще один метод получения новых распределений - *склейка*.

Похож на смешивание в том смысле, что два или более отдельных процесса отвечают за порождение убытков.

При смешивании различные процессы действуют на подмножествах совокупности.

При склейке процессы различаются в зависимости от величины потерь (на разных интервалах действуют разные модели).

Определение 11. *k*-компонентное склеенное распределение имеет функцию плотности, которая представима в следующем виде:

$$f_X(x) = \begin{cases} a_1 f_1(x), c_0 < x < c_1 \\ a_2 f_2(x), c_1 < x < c_2 \\ \dots \\ a_k f_k(x), c_{k-1} < x < c_k \end{cases}$$

Для $j = 1, \dots, k$ $a_j > 0$ и $f_j(x)$ является функцией плотности, заданной на (c_{j-1}, c_j), $a_1 + \cdots + a_k = 1$.

Пример 16

Покажем, что Модель 5 является двухкомпонентной склейкой.

Функция плотности

$$f(x) = \begin{cases} 0.01, 0 \leq x < 50 \\ 0.02, 50 \leq x < 75 \end{cases}$$

и склейка получается при $f_1(x) = 0.02, 0 \leq x < 50$, которая является плотностью равномерного распределения на интервале от 0 до 50, и $f_2(x) = 0.04, 50 \leq x < 75$, которая является плотностью равномерного распределения на интервале от 50 до 75. Коэффициенты равны $a_1 = 0.5$ и $a_2 = 0.5$.

Свойства склейки

Наличие функций плотности и коэффициентов гарантирует, что результат является функцией плотности.

Частая причина использования склейки параметрических моделей: поведение хвоста может быть несовместимым с поведением при небольших потерях.

Например, опыт (основанный на знаниях, выходящих за рамки того, что доступно на текущем, возможно, небольшом наборе данных) может указывать на то, что хвост соответствует распределению Парето, но в окрестности моды распределение больше соответствует логнормальному распределению.

Другой вариант. Имеется большой объем данных небольших значений, но ограниченный для больших значений. Тогда можно использовать эмпирическое распределение до определенной точки и параметрическую модель за пределами этого значения.

Свойства склейки

Определение 11 работает, когда точки $c_0, \dots, c_k$ известны заранее.

Другой способ построения склейки состоит в том, чтобы использовать стандартные распределения в диапазоне от c_0 до c_k.

Пусть $g_j(x)$ - j-я функция плотности. Тогда в Определении 11 можно заменить

$f_j\ (x)$ на $g_j(x)/\left[G\left(c_j\ \right) - \ G\left(c_{j-1}\right)\right]$.

Такой способ позволяет сделать точки склейки параметрами, которые также могут быть оценены.

Пример 17

Ни один из способов склейки не гарантирует, что результирующая функция плотности будет непрерывной.

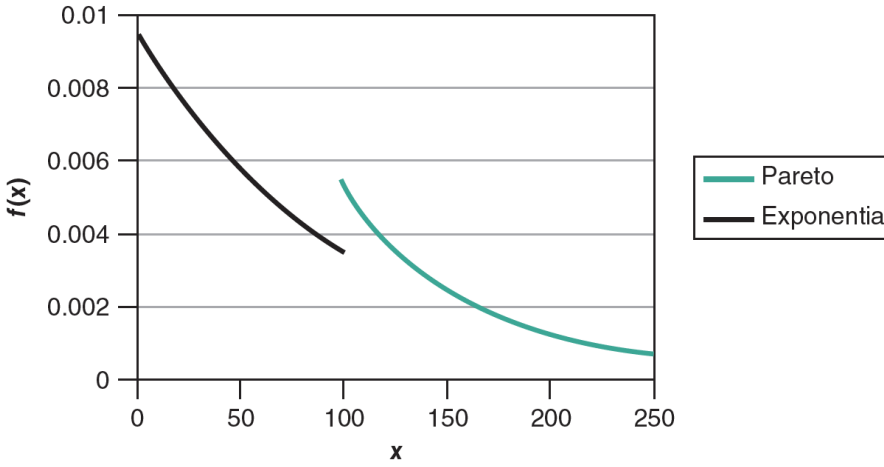
Построим двухкомпонентную склейку, используя экспоненциальное распределение на интервале от 0 до c и распределение Парето на $(c,\ \infty)$ (параметр γ используется вместо θ).

Склейка имеет вид:

$$f_X(x) = \begin{cases} a_1 \dfrac{\theta^{-1} e^{-\frac{x}{\theta}}}{1 - e^{-\frac{c}{\theta}}}, 0 < x < c, \\ a_2 \dfrac{\alpha \gamma^{\alpha} (x+\gamma)^{-\alpha-1}}{\gamma^{\alpha} (c+\gamma)^{-\alpha}}, c < x < \infty \end{cases}$$

Так как интеграл от функции плотности должен быть равен 1, то $a_1 = v, a_2 = 1 - v$. Получаем:

$$f_X(x) = \begin{cases} v \dfrac{\theta^{-1} e^{-\frac{x}{\theta}}}{1 - e^{-\frac{c}{\theta}}}, 0 < x < c, \\ (1-v) \dfrac{\alpha (c+\gamma)^{\alpha}}{(x+\gamma)^{\alpha+1}}, c < x < \infty \end{cases}$$



c = 100, v = 0.6, θ = 100, γ = 200, α = 4

Некоторые распределения и соотношения

Существует много способов организации распределений в группы.

Есть семейства Pearson (12 типов), Burr (12 типов), Stoppa (5 типов) and Dagum (11 типов). Одно и то же распределение может присутствовать в нескольких семействах, показывая разнообразие взаимосвязей.

Семейства, представленные дальше, широко используются в практике актуарного моделирования, так как все они имеют положительный носитель и имеют правую асимметрию.

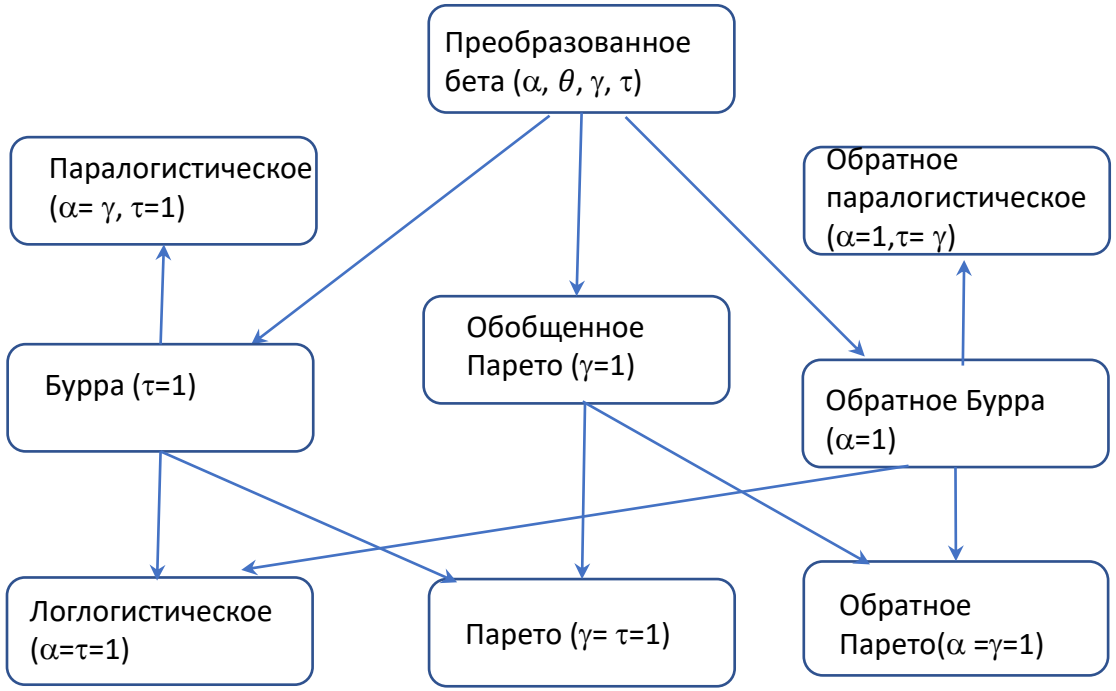
Два параметрических семейства

Многие распределения являются частными случаями других.

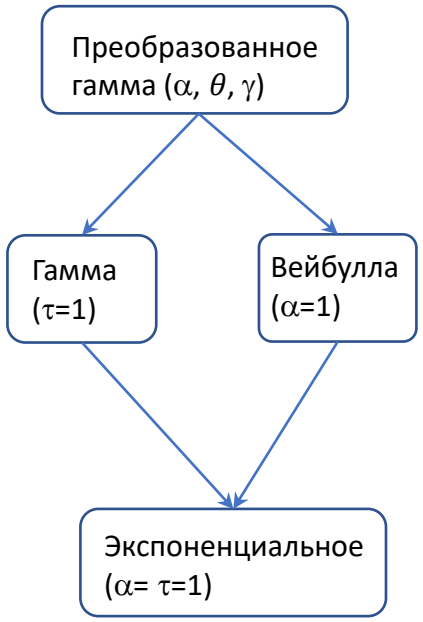
Например, распределение Вейбулла с $\tau = 1$ и произвольным θ является экспоненциальным распределением.

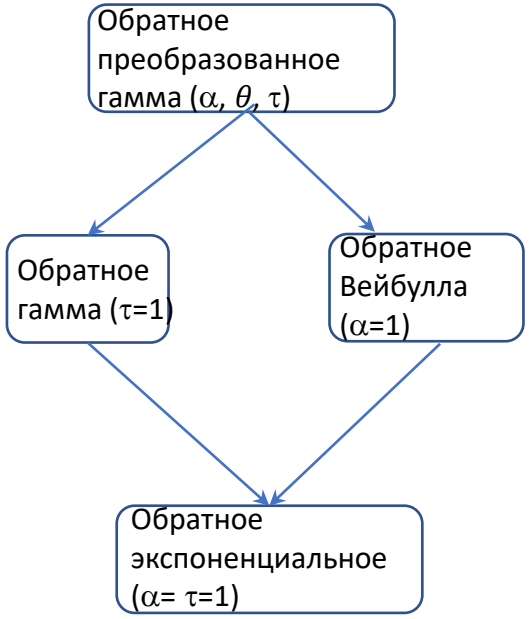
Два параметрических семейства

Преобразованное бета



Преобразованное гамма ($f(x) = \frac{\tau u^{\alpha} e^{-u}}{x\Gamma(\alpha)}, u = \left(\frac{x}{\theta}\right)^{\tau}$)





Линейное экспоненциальное семейство

- Большое параметрическое семейство включающее в себя множество распределений, интересных с актуарной точки зрения, используется в Байесовском анализе и доверительной теории.
- **Определение 12.** Случайная величина X (дискретная или непрерывная) имеет распределение из линейного экспоненциального семейства, если плотность ее распределения может быть параметризована с помощью параметра θ и выражена в виде

$$f(x;\ \theta) = \frac{p(x)e^{r(\theta)x}}{q(\theta)}$$

- Функция $p(x)$ зависит только от x и функция $q(\theta)$ является нормировочной. Также носитель случайной величины не должен зависеть от θ. Параметр $r(\theta\ )$ называется каноническим параметром распределения.
- Такую форму имеют многие стандартные распределения.

Пример 18

Покажем, что нормальное распределение входит в линейное экспоненциальное семейство.

Полагая среднее равным θ

$$f(x,\theta) = (2\pi\upsilon)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\upsilon}\,(x-\theta)^2\right] = (2\pi\upsilon)^{-1/2} \exp\left(-\frac{x^2}{2\upsilon} + \frac{\theta}{\upsilon}\,x - \frac{\theta^2}{2\upsilon}\right)$$

$$= \frac{\left[(2\pi\upsilon)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\upsilon}\right)\right] \exp\left(\frac{\theta}{\upsilon}\,x\right)}{\exp\left(\frac{\theta^2}{2\upsilon}\right)}$$

Получили вид функции из определения 12 при условии

$$p(x) = (2\pi\upsilon)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\upsilon}\right), r(\theta) = \frac{\theta}{\upsilon}\,, q(\theta) = \exp\left(\frac{\theta^2}{2\upsilon}\right)$$

Пример 19

Гамма распределение принадлежит линейном экспоненциальному семейству

$$f(x,\theta) = \theta^{-\alpha}\frac{x^{\alpha-1}e^{-\frac{x}{\theta}}}{\Gamma(\alpha)}$$

при

$$r(\theta) = -\frac{1}{\theta}, q(\theta) = \theta^{\alpha}, p(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}.$$

Математическое ожидание

Как видно из примеров, могут быть другие параметры помимо θ, но они не играют явной роли в последующем анализа линейного экспоненциального семейства.

Найдем математическое ожидание и дисперсию. Для этого сначала заметим, что

$$\ln f(x,\theta) = \ln p(x) + r(\theta)x - \ln q(\theta).$$

Дифференцируя по θ, получим

$$\frac{\partial}{\partial\theta} f(x,\theta) = \left[r'(\theta)x - \frac{q'(\theta)}{q(\theta)} \right] f(x,\theta)$$

Интегрируя по x, получим

$$\int \frac{\partial}{\partial\theta} f(x,\theta)dx = r'(\theta) \int x\, f(x,\theta)dx - \frac{q'(\theta)}{q(\theta)} \int f(x,\theta)\, dx$$

Меняя порядок интегрирования и дифференцирования в левой части, получим

$$\frac{\partial}{\partial\theta} \int f(x,\theta)dx = r'(\theta) \int x\, f(x,\theta)dx - \frac{q'(\theta)}{q(\theta)} \int f(x,\theta)\, dx$$

$$\int f(x,\theta)dx = 1, \qquad \int x\, f(x,\theta)dx = E[X], \qquad E[X] = \mu(\theta) = \frac{q'(\theta)}{r'(\theta)q(\theta)}$$

Дисперсия

Перепишем дифф. уравнение $\frac{\partial}{\partial\theta}f(x,\theta)=\left[r'(\theta)x-\frac{q'(\theta)}{q(\theta)}\right]f(x,\theta)$ в виде $\frac{\partial}{\partial\theta}f(x,\theta)=r'(\theta)[x-\mu(\theta)]f(x,\theta)$

Дифференцируем по θ

$$\frac{\partial^2}{\partial\theta^2}f(x,\theta)=r''(\theta)[x-\mu(\theta)]f(x,\theta)-r'(\theta)\mu'(\theta)f(x,\theta)+[r'(\theta)]^2[x-\mu(\theta)]^2f(x,\theta)$$

Снова интегрируем по x

$$\int\frac{\partial^2}{\partial\theta^2}f(x,\theta)dx=\int r''(\theta)[x-\mu(\theta)]f(x,\theta)dx-\int r'(\theta)\mu'(\theta)f(x,\theta)dx+\int\left[r'(\theta)\right]^2[x-\mu(\theta)]^2f(x,\theta)dx$$

$$\frac{\partial^2}{\partial\theta^2}\int f(x,\theta)dx=r''(\theta)\int[x-\mu(\theta)]f(x,\theta)dx-r'(\theta)\mu'(\theta)\int f(x,\theta)dx+[r'(\theta)]^2\int[x-\mu(\theta)]^2f(x,\theta)dx$$

$$0=-r'(\theta)\mu'(\theta)+\left[r'(\theta)\right]^2\int[x-\mu(\theta)]^2f(x,\theta)dx\quad\Longrightarrow\quad Var[X]=v(\theta)=\frac{\mu'(\theta)}{r'(\theta)}$$

Линейное экспоненциальное семейство включает также дискретные распределения – Пуассона, биномиальное и отрицательное биномиальное.

Дискретные распределения

В страховании счетные распределения могут быть использованы для описания числа убытков или претензий.

Одновременный учет числа исков и их размеров позволяет решать разнообразные задачи, связанные со страховыми выплатами.

Кроме того, модель числа убытков получить довольно легко, и опыт показывает, что обычно используемые распределения действительно позволяют оценить страховые риски и точки зрения порождения убытков.

Функция вероятности p_k обозначает вероятность того, что произойдет ровно k событий (таких как претензии или убытки). Пусть N - случайная величина, представляющая число таких убытков. Тогда

$$p_k = P\{N = k\}, \qquad k = 0, 1, 2, \ldots$$

Производящая функция вероятности дискретной случайной величины N с функцией вероятности p_k равна

$$P(z) = P_N(z) = E\left[z^N\right] = \Sigma_{k=0}^{N} p_k z_k$$

Как и производящая функция моментов, производящая функция вероятности порождает моменты:

$P'(1) = E[N], P''(1) = E[N(N-1)]$.

$$P^m(z) = E\left[\frac{d^m}{dz^m} z^N\right] = E[N(N-1)\cdots(N-m+1)z^{N-m}] = \sum_{k=m}^{\infty} k(k-1)\cdots(k-m+1)z^{k-m} p_k$$

$$P^m(0) = m!\, p_m, \qquad p_m = \frac{P^m(0)}{m!}$$

Распределение Пуассона

Функция вероятности: $p_k = \frac{e^{-\lambda}\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$ Производящая функция вероятности: $P(z) = e^{\lambda(z-1)}, \lambda > 0$

Среднее и дисперсия могут быть вычислены из производящей функции вероятности:

$$E[N] = P'(1) = \lambda$$

$$E[N(N-1)] = P''(1) = \lambda^2$$

$$Var[N] = E[N(N-1)] + E[N] - E^2[N] = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Распределение Пуассона имеет два полезных свойства.

Теорема 5. Пусть $N_1, \ldots, N_n$ независимые случайные величины, имеющие распределения Пуассона с параметрами $\lambda_1, \ldots, \lambda_n$. Тогда $N = N_1 + \cdots + N_n$ имеет распределение Пуассона с параметром $\lambda_1 + \cdots + \lambda_n$.

Доказательство. Для суммы случайных величин, распределенных по закону Пуассона, имеем

$$P_N(z) = \prod_{j=1}^{n} P_{N_j}(z) = \prod_{j=1}^{n} \exp\left[\lambda_j(z-1)\right] = \exp\left[\sum_{j=1}^{n} \lambda_j(z-1)\right] = e^{\lambda(z-1)},$$

где $\lambda = \lambda_1 + \cdots + \lambda_n$. Как и производящая функция моментов, производящая функция вероятности единственна, поэтому N имеет распределение Пуассона с параметром λ.

Распределение Пуассона

Второе свойство полезно на практике при моделировании рисков в страховании.

Пусть число требований в заданный период времени, например, 1 год, имеет распределение Пуассона.

Требования разделены на m различных типов. Например, они могут различаться по размеру (выше и ниже установленного порога). Тогда число требований выше порога также будет иметь распределение Пуассона, но с другим значением параметра.

Предположим, что число заявок на получение сложного медицинского покрытия подчиняется распределению Пуассона. Пусть «типы» требований, относятся к различным медицинским процедурам. Если одна из позиций удаляется из покрытия, то распределение числа требований по пересмотренному договору имеет распределение Пуассона, но с другим параметром.

Распределение Пуассона

Теорема 6. Пусть число событий N - случайная величина, имеющая распределение Пуассона со средним λ. Предположим, что каждое событие может быть отнесено к одному из m типов с вероятностями $p_1, \dots, p_m$ независимо от всех других событий. Тогда число событий $N_1, \dots, N_m$, соответствующих типам событий 1,…,m, взаимно независимые случайные величины, имеющие распределения Пуассона со средними $\lambda p_1, \dots, \lambda p_m$, соответственно.

Доказательство. Для фиксированного $N = n$ условное совместное распределение $(N_1, \dots, N_m)$ является полиномом с параметрами $(n, p_1, \dots, p_m)$. Также для фиксированного $N = n$ условное распределение N_j бином с параметрами (n, p_j). Совместная функция вероятности $(N_1, \dots, N_m)$

$$P\{N_1 = n_1, \dots, N_m = n_m\} = P\{N_1 = n_1, \dots, N_m = n_m | N = n\} \times P\{N = n\} = \frac{n!}{n_1!\, n_2! \cdots n_m!} p_1^{n_1} \cdots p_m^{n_m} \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= \prod_{j=1}^{m} e^{-\lambda p_j} \frac{(\lambda p_j)^{n_j}}{n_j!}$$

где $n = n_1 + n_2 + \cdots + n_m$. Аналогично функция вероятности для N_j равна

$$P\{N_j = n_j\} = \sum_{n=n_j}^{\infty} P\{N_j = n_j | N = n\} P\{N = n\} = \sum_{n=n_j}^{\infty} \binom{n}{n_j} p_j^{n_j} (1 - p_j)^{n-n_j} \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= e^{-\lambda} \frac{(\lambda p_j)^{n_j}}{n_j!} \sum_{n=n_j}^{\infty} \frac{[\lambda(1 - p_j)]^{n-n_j}}{(n - n_j)!} = e^{-\lambda} \frac{(\lambda p_j)^{n_j}}{n_j!} e^{\lambda(1-p_j)} = e^{-\lambda p_j} \frac{(\lambda p_j)^{n_j}}{n_j!}$$

Пример 20

В исследовании медицинского страхования ожидаемое количество претензий по отдельному полису составляет 2.3, число претензий распределено по закону Пуассона.

Рассмотрим возможность исключения одной медицинской процедуры из страхового покрытия в рамках этого полиса. На основании исторических данных установлено, что на эту процедуру приходится примерно 10% претензий. Определим новое распределение.

По теореме 6 распределение количества претензий, ожидаемых в соответствии с пересмотренным страховым полисом после исключения процедуры из страхового покрытия, является пуассоновским со средним 0.9(2.3) = 2.07.

Следует отметить, что это не подразумевает снижения премии на 10%, поскольку распределение суммы претензии может измениться.

Отрицательное биномиальное распределение

Отрицательное биномиальное распределение широко используется в качестве альтернативы Пуассоновскому. Как и распределение Пуассона, оно имеет положительные вероятности для неотрицательных целых чисел. Наличие двух параметров делает его более гибким по форме, чем пуассоновское.

Определение 13. Функция вероятности отрицательного биномиального распределения

$$P\{N = k\} = p_k = \binom{k + r - 1}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1 + \beta}\right)^k, k = 0, 1, 2, \ldots\ r > 0, \beta > 0.$$

Биномиальные коэффициенты равны

$$\binom{x}{k} = \frac{x(x - 1)\cdots(x - k + 1)}{k!}$$

k является целым, x может быть любым действительным. При $x > k$ - 1

$$\binom{x}{k} = \frac{\Gamma(x + 1)}{\Gamma(k + 1)\Gamma(x - k + 1)}$$

что может быть полезным, так как вычисление $\ln\Gamma(x)$ является стандартной функцией в электронных таблицах, языках программирования, мат. пакетах.

Отрицательное биномиальное распределение

Производящая функция вероятностей отрицательного биномиального распределения

$$P(z) = [1 - \beta(z - 1)]^{-r}$$

Отсюда среднее и дисперсия отрицательного биномиального распределения

$$E[N] = r\beta\,, Var[N] = r\beta(1 + \beta).$$

Так как β положительно, то дисперсия больше мат. ожидания. Это соотношение отличается от распределения Пуассона, для которого дисперсия и мат. ожидание равны. Таким образом, если для конкретного набора данных наблюдаемая дисперсия больше наблюдаемого среднего, отрицательное биномиальное распределение может быть более подходящим, чем распределение Пуассона.

При этом отрицательное биномиальное распределение является обобщением распределения Пуассона, по крайней мере, двумя способами: как смесь распределения Пуассона с гамма-распределением и как составное распределение Пуассона с логарифмическим вторичным распределением

Частный случай: геометрическое распределение

- Геометрическое распределение является частным случаем отрицательного биномиального распределения с r = 1. В некотором смысле, его можно считать дискретным аналогом непрерывного экспоненциального распределения. И геометрическое, и экспоненциальное имеют убывающие функции вероятности и, следовательно, свойство отсутствия памяти, которое может быть интерпретировано по-разному. Если экспоненциальное распределение - это распределение продолжительности жизни, то ожидаемая будущая продолжительность жизни постоянна для любого возраста. Если экспоненциальное распределение описывает размер страховых претензий, то отсутствие памяти означает следующее: учитывая, что претензия превышает определенный уровень d, ожидаемая сумма претензии, превышающая d, постоянна и поэтому не зависит от d, т.е. если применяется франшиза в размере d, ожидаемый платеж по каждой претензии останется неизменным, но ожидаемое количество платежей уменьшится.
- Если геометрическое распределение описывает количество претензий, то отсутствие памяти можно интерпретировать следующим образом: учитывая, что существует по меньшей мере m претензий, распределение вероятностей количества претензий, превышающих m, не зависит от m.
- Среди непрерывных распределений экспоненциальное распределение используется для различения субэкспоненциальных распределений с тяжелыми хвостами и распределений с легкими хвостами. Аналогично для частотных распределений, распределения, которые затухают в хвосте медленнее, чем геометрическое распределение, часто считаются имеющими тяжелые хвосты, тогда как распределения, которые затухают быстрее, чем геометрическое, имеют легкие хвосты. Отрицательное биномиальное распределение имеет тяжелый хвост (падает медленнее, чем геометрическое распределение), когда r < 1, и более легкий хвост, чем геометрическое распределение, когда r > 1.

Отрицательное биномиальное распределение как смесь пуассоновских

- Один из способов получить отрицательное биномиальное распределение состоит в смеси пуассоновских. Предположим, что риск имеет распределенное по закону Пуассона число претензий с известным λ. Можно считать λ реализацией случайной величины Λ. Обозначим плотность распределения Λ через $u(\lambda)$, причем Λ может непрерывным или дискретным, функцию распределения обозначим через $U(\lambda)$.
- Причин, по которым λ является реализацией случайной величины, может быть несколько. Во-первых, можно считать, имеется набор неоднородных рисков с параметром риска Λ. Рассмотрим портфель страховых полисов с одинаковой премией, например, группу автолюбителей с одной и той же категорией. Такие категории обычно имеет довольно широкие диапазоны значений, например 0–7,500 миль в год, гараж, проезжает менее 50 в неделю и т.п. Известно, что не все водители в одной категории похожи, даже если они кажутся похожими с точки зрения страховщика и платят одинаковую премию. Параметр λ измеряет ожидаемое число происшествий для заданного водителя. Если λ меняется в зависимости от совокупности водителей, мы можем рассматривать отдельного застрахованного как выборочное значение из совокупности водителей. Для конкретного водителя λ неизвестно, но имеет некоторое распределение $u(\lambda)$. Реальное значение λ не наблюдаемо. Мы можем наблюдать только число происшествий с одним водителем. Таким образом, появляется дополнительная неопределенность относительно параметра.

Смесь пуассоновских распределений

Это то же смешивание распределений, которое было в непрерывном случае. В некоторых случаях это называется неопределенностью параметра. В Байесовском подходе распределение Λ называется априорным, а параметры его распределения называются гиперпараметрами. Роль распределения $u(\cdot)$ очень важна в доверительной теории.

Когда параметр λ неизвестен, вероятность того, что будет ровно k претензий может быть записана как ожидаемое значение с той же вероятностью, но при условии $\Lambda = \lambda$, где мат. ожидание берется по распределению Λ. Согласно формуле полной вероятности,

$$p_k = P\{N = k\} = E[P\{N = k|\Lambda\}] = \int_0^\infty P\{N = k|\Lambda = \lambda\}\, u(\lambda)d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!}\, u(\lambda)\, d\lambda.$$

Теперь предположим, что Λ имеет гамма распределение. Тогда

$$p_k = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!}\frac{\lambda^{\alpha-1}e^{-\left(\frac{\lambda}{\theta}\right)}}{\theta^\alpha\Gamma(\alpha)}\, d\lambda = \frac{1}{k!}\frac{1}{\theta^\alpha\Gamma(\alpha)}\int_0^\infty e^{-\lambda\left(1+\frac{1}{\theta}\right)}\, \lambda^{k+\alpha-1}\, d\lambda = \frac{1}{k!}\frac{1}{\Gamma(\alpha)}\left(\frac{\theta}{1+\theta}\right)^k\left(\frac{1}{1+\theta}\right)^\alpha\Gamma(k+\alpha)$$

$$= \binom{k+\alpha-1}{k}\left(\frac{\theta}{1+\theta}\right)^k\left(\frac{1}{1+\theta}\right)^\alpha$$

Получили отрицательное биномиальное распределение.

Предел распределения Пуассона

Распределение Пуассона является предельным случаем отрицательного биномиального распределения.

Пусть $r \to \infty, \beta \to 0$, но $r\beta = const = \lambda$. Заменив $\beta = \lambda / r$, получим

$$\lim_{r\to\infty} \left[1 - \frac{\lambda}{r}(z-1)\right]^{-r} = \lim_{r\to\infty} \left(\left[1 + \frac{\lambda(z-1)}{-r}\right]^{-\frac{r}{\lambda(z-1)}}\right)^{\lambda(z-1)} = e^{\lambda(z-1)}$$

Получили функцию вероятности распределения Пуассона

Биномиальное распределение

(Положительное) биномиальное распределение обладает свойствами, отличными от свойств распределения Пуассона и отрицательного биномиального, которые делают его практически полезным. Во-первых, дисперсия меньше мат. ожидания. Во-вторых, оно описывает ситуацию, в которой каждый из m рисков является предметом претензии или убытка.

Рассмотрим m независимых и идентичных рисков, каждый с вероятностью q приводящий к убытку. Например, в наборе договоров страхования жизни все индивидуумы находятся в одной и той же группе по уровню смертности, т.е. они могут быть курящие мужчины возраста 35, и срок действия полиса составляет 5 лет. Тогда q - вероятность того, что человек умрет в следующем году. Число претензий от одного лица подчиняется распределению Бернулли: с вероятностью 1 - q равно 0 и вероятностью q равно 1. Производящая функция вероятности числа претензий на индивидуума

$$P(z) = (1 - q)z^0 + qz^1 = 1 + q(z - 1)$$

Если есть m таких независимых индивидуумов, то производящая функция вероятности должна умножаться сама на себя, чтобы получилось общее число убытков от группы из m индивидуумов:

$$P(z) = \big(1 + q(z - 1)\big)^m, \qquad 0 < q < 1:$$

Тогда

$$p_k = P\{N = k\} = \binom{m}{k} q^k (1 - q)^{m-k}, k = 0, 1, \dots, m$$

Биномиальное распределение

Получили функцию вероятности биномиального распределения с параметрами m и q.

Из схемы испытания Бернулли ясно, что может произойти не более m событий. Следовательно, распределение имеет положительные вероятности только для неотрицательных целых чисел до m включительно.

Биномиальное распределение имеет конечный носитель, то есть диапазон значений, для которых существуют положительные вероятности, имеет конечную длину. Это позволяет моделировать ситуации, которых есть верхний предел диапазона возможных значений.

Например, вряд ли возможно, чтобы число претензий, связанных с авариями с участием одного автомобиля в течение года, было больше 12, так как для ремонта автомобиля между авариями требуется время. Если в такой ситуации используется модель со значениями, выходящими за пределы 12, их вероятности должны быть малы, и потому не будут оказывать влияние на любые принимаемые решения.

Среднее и дисперсия биномиального распределения равны

$$E[N] = mq, \qquad Var[N] = mq(1 - q)$$

Класс $(a, b, 0)$

Определение 14. Пусть p_k - функция вероятности дискретного распределения. Распределение входит в класс $(a, b, 0)$, если существуют такие a и b, что

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1}, k = 1, 2, \ldots$$

Эта рекурсия описывает относительный размер последовательных вероятностей в счетных распределениях. Вероятность нулевого значения p_0 может быть получена из рекурсивной формулы, так как сумма вероятностей равна 1. Класс распределений $(a, b, 0)$ является двухпараметрическим.

Пример 21

Покажем, что биномиальное распределение с параметрами m и q принадлежит классу $(a, b, 0)$.

$$p_k = \binom{m}{k} q^k (1-q)^{(m-k)}, k = 0, 1, \ldots, m,$$

и $p_k = 0, k > m$. Перепишем вероятности для $k = 1, 2, \ldots, m$

$$p_k = \frac{m!}{(m-k)!\,k!} q^k (1-q)^{m-k}$$

$$= \frac{m-(k-1)}{k} \frac{q}{1-q} \left(\frac{m!}{(m-(k-1))!\,(k-1)!} q^{k-1} (1-q)^{m-(k-1)} \right)$$

$$= \frac{q}{1-q} \left(-1 + \frac{m+1}{k} \right) p_{k-1}$$

Таким образом, $p_k = \left(a + \frac{b}{k} \right) p_{k-1}$ для $k = 1, 2, \ldots, m$ при $a = -q/(1-q)$ и $b = \frac{(m+1)q}{1-q}$.

Осталось проверить выполнение рекурсии для $k = m+1, m+2, \ldots$ Для $k = m+1$

$$\left(a + \frac{b}{m+1} \right) p_m = \left(-\frac{q}{1-q} + \frac{q}{1-q} \right) p_m = 0 = p_{m+1}.$$

Для $k = m+2, m+3, \ldots$ рекурсия, очевидно, выполняется.

Класс $(a, b, 0)$

Распределение Пуассона и отрицательное биномиальное распределение также принадлежат классу (a, b, 0). Геометрическое распределения является частным случаем отрицательного биномиального, поэтому также принадлежит этому классу.

Рекурсивная формула может быть переписана в виде (при $p_{k-1} > 0$)

$$k \frac{p_k}{p_{k-1}} = ak + b, k = 1, 2, 3, \ldots$$

Выражение слева линейно по k. В таблице видно, что наклон прямой a =0 Для распределения Пуассона, a <0 для биномиального распределения, a >0 для отрицательного биномиального распределения. Это свойство позволяет графически определить, какое из распределений больше подойдет под данные.

Распределение	a	b	p_0
Пуассона	0	λ	$e^{-\lambda}$
Биномиальное	$-\frac{q}{1-q}$	$\frac{(m+1)q}{1-q}$	$(1-q)^m$
Отрицательное биномиальное	$\frac{\beta}{1+\beta}$	$\frac{(r-1)\beta}{1+\beta}$	$(1+\beta)^{-r}$
Геометрическое	$\frac{\beta}{1+\beta}$	0	$(1+\beta)^{-1}$

Пример 22

Сначала необходимо изобразить

$$k\frac{\hat{p}_k}{\hat{p}_{k-1}} = k\frac{n_k}{n_{k-1}}$$

относительно k. Наблюдения должны образовывать почти прямую линию. Значение угла наклона показывает, какую именно модель следует выбрать. Заметим, что этого нельзя делать, если какое-либо n_k =0. Таким образом, эта процедура хуже работает для маленьких объемов наблюдений.

Рассмотрим данные о происшествиях в таблице. Для 9461 полисов автомобильного страхования указано число происшествий в рамках действия полиса. Также в таблице приведены наблюдаемые значения, которые должны быть линейны.

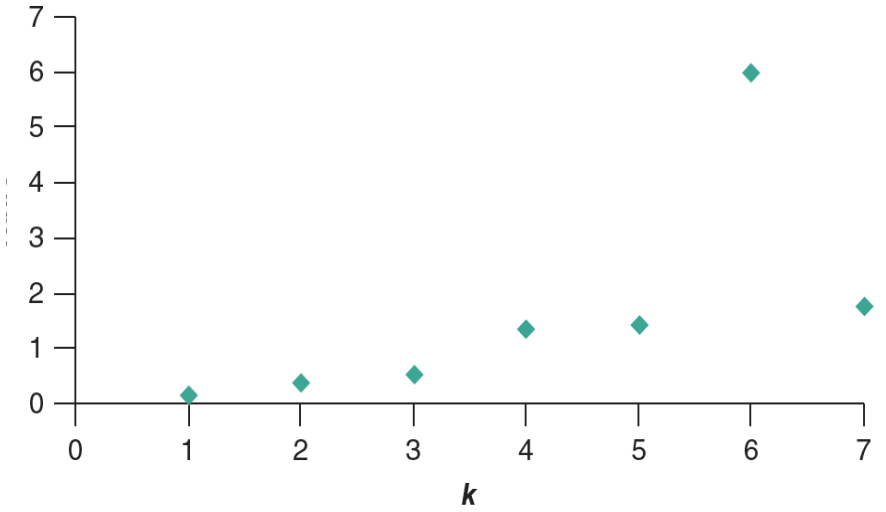
k	n_k	$k\frac{n_k}{n_{k-1}}$
0	7840	
1	1371	0.17
2	239	0.36
3	42	0.53
4	14	1.33
5	4	1.43
6	4	6.00
7	1	1.75
8+	0	
Всего	9461	

Пример 22

На рисунке изображены эти значения в зависимости от k (числа происшествий). Видно, что зависимость примерно линейная, кроме точки $k = 6$. Надежность величин по мере увеличения k уменьшается, поскольку число наблюдений становится небольшим, а вариабельность результатов растет, что иллюстрирует слабость этой процедуры.

Видно, что наклон положительный, что позволяет предположить, что отрицательное биномиальное распределение является подходящей моделью.

Однако, значительно ли наклон отличается от нуля по графику определить нельзя. За счет изменения масштаба вертикальной оси графика можно сделать так, чтобы наклон выглядел более крутым. Графически трудно провести различие между пуассоновским и отрицательным биномиальным распределением. Однако биномиальное распределение, вероятно, можно исключить из рассмотрения. В этом случае желательно подогнать как пуассоновское, так и отрицательное биномиальное распределения и провести более формальный тест для выбора между ними.



Усечение и модификация в нуле

Иногда распределения неадекватно описывают характеристики некоторых наборов данных, встречающихся на практике. Например, хвост отрицательного биномиального распределения может оказаться недостаточно тяжелым или распределения в классе $(a, b, 0)$ не могут отразить форму набора данных в какой-либо другой части распределения.

Рассмотрим проблему плохого соответствия в левом конце распределения.

Для данных о количестве страховых случаев нулевая вероятность - это вероятность того, что претензий нет.

Существуют также ситуации, которые порождают большие вероятности при нулевом значении.

Класс $(a, b, 1)$

Определение 15. Пусть p_k - функция вероятности дискретного распределения. Распределение принадлежит классу $(a, b, 1)$, если существуют константы a и b такие, что $p_k = \left(a + \frac{b}{k}\right) p_{k-1}, k = 2, 3, 4, \ldots$

В отличии от $(a, b, 0)$ этот класс начинает рекурсию с p_1, а не с p_0.

Распределения от $k = 1$ до $k = \infty$ имеет ту же форму, что и класс $(a, b, 0)$ в том смысле, что вероятности те же с точностью до коэффициента пропорциональности, так как $\sum_{k=1}^{\infty} p_k$ можно привести к любому числу из интервала $[0;\ 1]$. Оставшаяся вероятность относится к $k = 0$.

Необходимо различать ситуации, в которых $p_0 = 0$ и $p_0 > 0$. Первый подкласс называется *усеченными* (в нуле) распределениями. Ему принадлежат усеченное распределение Пуассона, усеченное биномиальное и усеченное отрицательное биномиальное.

Второй подкласс называется *модифицированными* в нуле распределениями, так как вероятность отличается от того, что есть в классе $(a, b, 0)$. Эти распределения можно рассматривать как смесь распределения $(a, b, 0)$ и вырожденного распределения с вероятностью, сосредоточенной в нуле. Покажем эту эквивалентность формально. Заметим, что все усеченные в нуле распределения можно рассматривать как модифицированные в нуле распределения с $p_0 = 0$. Будем для кратности использовать аббревиатуры ZT (zero truncated) и ZM (zero modified). Для того, чтобы различать вероятности будем использовать $p_k = P\{N = k\}$. Для усеченных в нуле p_k^T и для модифицированных p_k^M

Пусть $P(z) = \sum_{k=0}^{\infty} p_k z^k$ - производящая функция вероятности для распределения из класса $(a, b, 0)$. Пусть $P^M(z) = \sum_{k=0}^{\infty} p_k^M z^k$ обозначает производящую функцию вероятности соответствующего распределения из класса $(a, b, 1)$, т.е. $p_k^M = cp_k;\ k = 1, 2, 3, .. u\ p_0^M$ – произвольное число. Тогда

$$P^M(z) = p_0^M + \sum_{k=1}^{\infty} p_k^M z^k = p_0^M + c \sum_{k=1}^{\infty} p_k z^k = p_0^M + c(P(z) - p_0).$$

Так как $P^M(1) = P(1) = 1$, то $1 = p_0^M + c(.1 - p_0)$, т.е. $c = \frac{1 - p_0^M}{1 - p_0}$ или $p_0^M = 1 - c(1 - p_0)$

Это соотношение необходимо, чтобы гарантировать, что сумма p_k^M *равна* 1. Тогда имеем

$$P^M(z) = p_0^M + \frac{1 - p_0^M}{1 - p_0}(P(z) - p_0) = \left(1 - \frac{1 - p_0^M}{1 - p_0}\right)1 + \frac{1 - p_0^M}{1 - p_0} P(z)$$

Это средневзвешенное производящих функций вероятности вырожденного распределения и соответствующего $(a, b, 0)$ распределения.

Более того, $p_k^M = \frac{1 - p_0^M}{1 - p_0} p_k,\ \ k = 1, 2, \ldots$

Пусть $P^T(z)$ обозначает производящую функцию вероятности усеченного в нуле распределения, соответствующего $(a, b, 0)$ производящей функции вероятности $P(z)$. Тогда полагая $p_0^M = 0$,

$P^T(z) = \frac{P(z) - p_0}{1 - p_0}$, $p_k^T = \frac{p_k}{1 - p_0}$, $k = 1, 2, \ldots$

$$p_k^M = (1 - p_0^M) p_k^T, \qquad k = 1, 2, \ldots$$

$$P^M(z) = p_0^M(1) + (1 - p_0^M) P^T(z)$$

Пример 4

Рассмотрим отрицательное биномиальное распределение с параметрами $\beta = 0.5$ и $r = 2.5$.

Определим первые четыре вероятности такой случайной величины. Затем определим соответствующие вероятности для усеченного в нуле и модифицированного в нуле (с $p_0^M = 0.6$).

Для отрицательного биномиального распределения $p_0 = (1 + 0.5)^{-2.5} = 0.362887, a = \frac{0.5}{1.5} = \frac{1}{3}; b = \frac{(2.5-1)0.5}{1.5} = \frac{1}{2}$

Рекурсивно получаем

$$p_1 = 0.362887\left(\frac{1}{3} + \frac{1}{2}\frac{1}{1}\right) = 0.302406;\ p_2 = 0.302406\left(\frac{1}{3} + \frac{1}{2}\frac{1}{2}\right) = 0.176404;\ p_3 = 0.176404\left(\frac{1}{3} + \frac{1}{2}\frac{1}{3}\right) = 0.088202.$$

Для усеченного в нуле распределения $p_0^T = 0$ по определению. Рекурсия начинается с $p_1^T = \frac{0.302406}{(1-0.362887)} = 0.474651$. Затем $p_2^T = 0.474651\left(\frac{1}{3} + \frac{1}{2}\frac{1}{2}\right) = 0.276880;\ p_3^T = 0.276880\left(\frac{1}{3} + \frac{1}{2}\frac{1}{3}\right) = 0.138440.$

Если бы все исходные значения были доступны, то усеченные до нуля вероятности могли бы быть получены умножением исходных значений на $1 / (1 - 0.362887) = 1.569580$.

Для модифицированного в нуле $p_0^M = 0.6$ произвольно. $p_1^M = \frac{(1-0.6)0.302406}{1-0.362887} = 0.189860$. Поэтому

$$p_2^M = 0.189860\left(\frac{1}{3} + \frac{1}{2}\frac{1}{2}\right) = 0.110752;\ p_3^M = 0.110752\left(\frac{1}{3} + \frac{1}{2}\frac{1}{3}\right) = 0.055376.$$

В этом случае каждая вероятность отрицательного биномиального распределения умножается на

Частным случаем модифицированного в нуле распределения является **zero-inflated**. Единственная разница состоит в том, что такое распределение требует $p_0^M > p_0$. Показано, что для zero-inflated Пуассоновского распределения дисперсия всегда больше среднего. Это дает альтернативу отрицательному биномиальному распределению когда это свойство требуется. Хотя мы рассматриваем только модифицированные в нуле из класса $(a;\ b;\ 0)$, класс $(a;\ b;\ 1)$ допускает другие распределения. Пространство параметров $(a;\ b)$ может быть расширено, чтобы допустить расширение отрицательного биномиального распределения на случаи $-1 < r < 0$. Для класса $(a;\ b;\ 0)$ требуется $r > 0$. Добавляя дополнительную область к пространству значений выборки, получаем «расширенное» усеченное отрицательное биномиальное распределение (ETNB) с параметрами $\beta > 0, r > -1, r \neq 0$.

Покажем, что рекурсивное соотношение $p_k = p_{k-1}\left(a + \frac{b}{k}\right), k = 2,3, c\ p_0 = 0$ определяет соответствующее распределение. Достаточно показать, что для любого p_1 последовательные значения p_k, полученные рекурсивно, положительны и $\sum_{k=1}^{\infty} p_k < \infty$.

Для ETNB, пространство параметров $a = \frac{\beta}{1+\beta},\ \beta > 0;\ b = \frac{(r-1)\beta}{1+\beta}, r > -1, r \neq 0$.

При $r \to 0$, предельным случаем ETNB является логарифмическое распределение с $p_k^T = \frac{\left[\frac{\beta}{1+\beta}\right]^k}{k\ln(1+\beta)}, k = 1,2,3$.

Производящая функция вероятности логарифмического распределения равна $\frac{P^T(z) = 1 - \ln[1-\beta(z-1)]}{\ln(1+\beta)}$. Модифицированное в нуле логарифмическое распределение получается путем приписывания произвольной вероятности в нуле и уменьшением остальных вероятностей.

Также интересен и другой предельный случай с $-1 < r < 0$ и $\beta \to \infty$. Соответствующее распределение называется распределением Сибуя. Оно имеет производящую функцию вероятности $P(z) = 1 - (1 - z)^{-r}$ и никаких моментов у нее не существует. Распределения, не имеющие моментов не имеют практического интереса при моделировании числа убытков, если только правый хвост впоследствии не модифицируется, так как иначе оказывается, что ожидается бесконечное число претензий

Пример 5

Определим вероятности для ETNB распределения с $r = -0.5$ и $\beta = 1$. Сделаем это для усеченной и модифицированной версий с произвольным $p_0^M = 0.6$

Имеем $a = 1/(1 + 1) = 0.5$ и $b = \frac{(-0.5-1)(1)}{1+1} = -0.75$. Также $p_1^T = -\frac{0.5(1)}{(1+1)^{0.5}-(1+1)} = 0.853553$.

Следующие значения

$$p_2^T = \left(0.5 - \frac{0.75}{2}\right) 0.853553 = 0.106694, p_3^T = \left(0.5 - \frac{0.75}{3}\right) 0.106694 = 0.026674.$$

Для модифицированных вероятностей усеченные вероятности должны быть умножены на 0.4. Тогда получится $p_1^M = 0.341421, p_2^M = 0.042678$ и $p_3^M = 0.010670$.

Естественно возникает вопрос, существует ли «настоящий» представитель ETNB распределения для примера, т.е. такой, для которого рекурсия начиналась бы с p_1, а не с p_2. Тогда p_0 должно удовлетворять $p\_1 = \left(0.5 - \frac{0.75}{1}\right) p_0 = -0.25 p_0$. Т.е. одна из двух вероятностей должны быть отрицательными, поэтому здесь решения не существует. Легко показать, что это происходит для любого $r < 0$.

Других представителей класса $(a, b, 1)$ не существует.