

Бенин Владимир Евгеньевич | Дир. кафедр мат. статистики.
двух массовых.

Условные мат. ожидания и условные вероятности.

1. $E X$

возникает как $\min_a E(X-a)^2 = \min_a E(X^2 - 2aX + a^2) =$

$$= \min_a (EX^2 - 2aEX + a^2) = \left\{ \begin{array}{l} a^* = EX \\ \text{к вопр. 3-х} \\ \text{элементов } a \end{array} \right\} = EX^2 - 2(EX)^2 + (EX)^2 = EX^2 - (EX)^2 = DX.$$

2. $\left(\frac{+}{-}\right) \min_a E|X-a| = ? \quad (a = \text{med } X, E|X - \text{med } X|)$

$$P(X \geq \text{med } X) \geq \frac{1}{2}, \quad P(X \leq \text{med } X) \geq \frac{1}{2}$$

Раскрыть 4 модуля
и рассмотреть разность
 $E|X-a| - E|X - \text{med } X|.$

3. $\int X$ с.в. не наблюдаемая, но она зависит от наблюдаемой с.в. Y , причем со с.в. Y можно сделать $\rightarrow$ линейное преобразование. Тогда:

$\min_f E(X - f(Y))^2 = E(X - E(X|Y))^2$, т.е. указанный min достигается на условном мат. ожидании $E(X|Y)$, которое найдем найдем при помощи преобразов с.в. Y на X .

4. Понятие условной вероятности и условного E удобно технически:

$$\begin{aligned} \int, \text{ например, } X, Y - \text{н.с.в.}; \text{ найдем } F_{X+Y}(x) &\equiv P(X+Y < x) = \left\{ \begin{array}{l} \text{по ф-ле} \\ \text{полной} \\ \text{вероятности} \end{array} \right\} = \\ &= E P(X+Y < x | Y) = \int P(X+Y < x | Y=y) dF_Y(y) = \left\{ \begin{array}{l} \text{по ф-ле} \\ \text{условной} \\ \text{вероятности} \end{array} \right\} = \\ &= \int P(X < x-y | Y=y) dF_Y(y) = \left\{ \begin{array}{l} X, Y - \text{н.с.в.} \end{array} \right\} = \int P(X < x-y) dF_Y(y) = \\ &= \int P(X < x-y) dF_Y(y) = \int F_X(x-y) dF_Y(y). \end{aligned}$$

Задача $\left(\frac{+}{-}\right)$

$\int f(x) \geq 0, g(x) \geq 0$ - плотности, т.е. $\int f(x) dx = 1, \int g(x) dx = 1$, тогда свертка или плотность суммы н.с.в. $f * g(x) = \int f(x-y)g(y) dy$.

Доказать, что последний интеграл $\exists$ и конечен.

Продемонстрируем полноту понятия условного м.о. на примере с суммой

] $k_1, k_2, \dots$ - н.о.с.в., имеющие смысл страховых требований одноклассных клиентов.

] N - н.о.т. $k_1, k_2, \dots$ с.в.; и $N = \{0, 1, 2, \dots\}$ и имеющая смысл кол-ва клиентов, обратившихся в страх. фирму за опред. время. Определим:

$$S_N = \begin{cases} k_1 + k_2 + \dots + k_N, & N \geq 1 \\ 0, & N = 0 \end{cases}, \text{ которая имеет смысл страховых требований к страхов. фирме за данный период.}$$

Найдем среднее значение страх. требований или ES_N .

$$ES_N = \left\{ \begin{array}{l} \text{по ф.не} \\ \text{полной вероятности} \end{array} \right\} = EE(S_N | N) = \sum_{n=0}^{\infty} E(S_N | N=n) \cdot P(N=n) =$$

$$= \left\{ \begin{array}{l} \text{по с.в.у} \\ \text{условного и.о.} \end{array} \right\} = \sum_{n=0}^{\infty} E(S_N | N=n) \cdot P(N=n) = \left\{ S_n, N \text{ - н.с.в.} \right\} = \sum_{n=0}^{\infty} ES_n \cdot P(N=n) =$$

$$= 0 + \sum_{n=1}^{\infty} E(k_1 + \dots + k_n) \cdot P(N=n) = \sum_{n=1}^{\infty} n \cdot EK_1 \cdot P(N=n) = EK_1 \sum_{n=1}^{\infty} n \cdot P(N=n) = EK_1 \cdot EN.$$

Задача ($\frac{4}{5}$).

$DS_N = ?$

(если правильно получили: то где $N: P(N=n) = 1$ получим $DS_N = n \cdot DK_1$).

6. Опишем еще одну ситуацию, в которой условные вероятности возникают естественно:

] $(\Omega, \mathcal{A}, P)$ - в.р. п-во, где Ω - непустое мн-во, $\mathcal{A}$ - σ -алгебра событий,

P - вероятность на $(\Omega, \mathcal{A})$.

$$\forall B \in \mathcal{A} \quad P(B) > 0, \quad \forall A \in \mathcal{A} : P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \forall A \in \mathcal{A}.$$

$$A \cap B \subseteq B \Rightarrow 0 \leq P(A \cap B) \leq P(B) < \infty.$$

$$B = \{\omega : X(\omega) = x\}$$

] события $B_1, \dots, B_n \in \mathcal{A}$ образуют разбиение Ω , т.е. $\Omega = \bigcup_{i=1}^n B_i, B_i \cap B_j = \emptyset$ $i \neq j, n \geq 1$

Примем $P(B_i) > 0, i = 1, \dots, n$.

Для фиксированного события $A \in \mathcal{A}, X = 1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$

$$Y \text{ - с.в.}, Y = \sum_{i=1}^n v_i \cdot 1_{B_i}(\omega), \quad v_i \in \mathbb{R}^d, i = 1, \dots, n.$$

Попытаемся выбрать коэффич. $v_1, \dots, v_n$: с.в. Y была как можно "ближе" к X :

$$\rho(X, Y) = \sqrt{E(X-Y)^2} = \sqrt{\int (X(\omega) - Y(\omega))^2 dP}$$

Сл-ва след. теорема:

Теорема 1

min угранного более раскраске $p(x, y)$ ^{госмиласм на} $b_1, \dots, b_n$; $b_i^* = \frac{P(A \cap B_i)}{P(B_i)} = P(A|B_i), i = \overline{1, n}$.

Док-во:

$$\begin{aligned} \int_{\Omega} (x(\omega) - y(\omega))^2 dP &= \int_{\Omega = \bigcup_{i=1}^n B_i} (x(\omega) - y(\omega))^2 dP = \int_{\bigcup_{i=1}^n B_i} (x(\omega) - y(\omega))^2 dP = \\ &= \sum_{i=1}^n \int_{B_i} (x(\omega) - y(\omega))^2 dP = \left\{ x = 1_A(\omega) \right. \\ &\quad \left. y \in B_i \Rightarrow y = b_i \right\} = \sum_{i=1}^n \int_{B_i} (1_A(\omega) - b_i)^2 dP = \\ &= \sum_{i=1}^n \int_{B_i} (1_A(\omega) - 2b_i 1_A(\omega) + b_i^2) dP = \sum_{i=1}^n \left(\int_{B_i} 1_A dP - 2b_i \int_{B_i} 1_A dP + b_i^2 \int_{B_i} dP \right) = \\ &= \left\{ \int_{\Omega} dP = \int_{\Omega} 1_{\Omega} dP = P(\Omega) \right\} = \sum_{i=1}^n (P(A \cap B_i) - 2b_i P(A \cap B_i) + b_i^2 P(B_i)) \geq \left\{ \begin{aligned} & b_i^* = \min, \\ & b_i^* = \frac{P(A \cap B_i)}{P(B_i)} \end{aligned} \right\} = \\ &\geq \sum_{i=1}^n (P(A \cap B_i) - 2b_i^* P(A \cap B_i) + b_i^{*2} P(B_i)) = \int_{\Omega} (x(\omega) - y^*(\omega))^2 dP, \forall b_i \in \mathbb{R}^1, i = \overline{1, n} \end{aligned}$$

$$y^*(\omega) = \sum_{i=1}^n b_i^* 1_{B_i}(\omega)$$

$$\int 1_A(\omega); B_i = \Omega.$$

Задача (*)

$$(A \Delta B) \Delta C \stackrel{?}{=} A \Delta (B \Delta C).$$

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$$

$$1_{A \Delta B}(\omega) = (1_A(\omega) - 1_B(\omega))^2 = |1_A(\omega) - 1_B(\omega)|$$

$$1_{(A \Delta B) \Delta C}(\omega) = \dots$$

$$1_{A \Delta (B \Delta C)}(\omega)$$

$$\begin{aligned} A \Delta B = C \Delta D &\Rightarrow A = ? \\ (A \Delta B) \Delta B &= (C \Delta D) \Delta B \\ A \Delta (B \Delta B) &= (C \Delta D) \Delta B \\ A &= (C \Delta D) \Delta B \end{aligned}$$

Определим теперь удобное мат. ожидание. Рассмотрим сначала сл. случай.

$$A, B, P(B > 0), P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Предположим, что н.в. сл. исходов конечно, т.е. $\Omega = \{\omega_1, \dots, \omega_m\}, m \geq 1$

$\Omega = \omega(\Omega)$ - мн-во всех исходов, $|\Omega| = 2^n$,

и $p(\omega_1), \dots, p(\omega_n) \exists$; $p(\omega_i) \geq 0$, $\forall i$, $p(\omega_1) + \dots + p(\omega_n) = 1$.

При этом $P(A) = \sum_{i: \omega_i \in A} p(\omega_i)$, $\forall A \in \mathcal{A}$.

$\exists$ на Ω задана действ. ф-ция: $X = X(\omega): \Omega \rightarrow \mathbb{R}^d$

Обозн $x_1, \dots, x_k$, $k \leq m$ - разн. зн-ия X .

$X: \begin{matrix} \omega_1 & \dots & \omega_k \\ p_1 & \dots & p_k \end{matrix}$, где $p_i \equiv P(X = x_i)$, $i = \overline{1, k}$

Определим мн-во $A_i = \{\omega: X(\omega) = x_i\} = X^{-1}(\{x_i\})$, $i = \overline{1, k}$

$A_i \cap A_j = \emptyset$, $i \neq j$; $\bigcup_{i=1}^k A_i = \Omega$

Если A_i, A_j - непересекающ, то $\exists$ хотя бы одна точка принадл. зн-е x_i, x_j , принадлеж.

При этом можно ф-цию или св $X(\omega) = \sum_{i=1}^k x_i \mathbb{1}_{A_i}(\omega)$.

$\exists X(\omega_0) = x_j \Rightarrow \omega_0 \in A_j$

$Y = Y(\omega)$ - сл. зн-е, заданной на Ω , т.е. $\Omega \rightarrow \mathcal{Y}$. $\forall$ мн-во.

Обозн $y_1, \dots, y_\ell$ - разн. зн-ия, сл. зн-е Y , $\ell \leq m$.

$Y: \begin{matrix} y_1 & \dots & y_\ell \\ q_1 & \dots & q_\ell \end{matrix}$ - распр-е Y ; $q_i = P(Y = y_i)$, $i = \overline{1, \ell}$

$B_i = \{\omega: Y(\omega) = y_i\} = Y^{-1}(\{y_i\})$, $i = \overline{1, \ell}$

$Y(\omega) = \sum_{i=1}^{\ell} y_i \mathbb{1}_{B_i}(\omega)$ - верно, т.к. мы не знаем природы y_i и можем их упоминать на число.

$\forall B \subseteq \Omega$, $P(B) > 0$

Рассмотрим набор чисел: $\frac{p(\omega_1) \cdot \mathbb{1}_B(\omega_1)}{P(B)}$, $\dots$, $\frac{p(\omega_m) \cdot \mathbb{1}_B(\omega_m)}{P(B)}$ (4).

Тогда ясно, что эти числа неотрицательные и в сумме равны 1, т.е.

$$\sum_{i=1}^m \frac{p(\omega_i) \cdot \mathbb{1}_B(\omega_i)}{P(B)} = \frac{1}{P(B)} \left(\sum_{i: \omega_i \in B} p(\omega_i) \cdot 1 + \sum_{i: \omega_i \notin B} p(\omega_i) \cdot 0 \right) =$$

$$= \frac{1}{P(B)} \cdot P(B) = 1$$

$E X$ - обн-е мат. ожид. сл. в. X .

$$E X = \sum_{i=1}^m X(\omega_i) \cdot p(\omega_i) = \sum_{i=1}^m x_i \cdot p_i$$

$$\left(\sum_{i=1}^m X(\omega_i) p(\omega_i) = \sum_{j=1}^k \sum_{\omega_i \in A_j} x(\omega_i) p(\omega_i) \right) - \text{предложен с индексом, проверим}$$

Здесь в этом выражении 1-ое слагаемое и заменим $p(\omega_i)$ на плотность, указанной выше (*), получим определение условного мат. ожидания.

Оп. 1. Условн. мат. ожид. с.в. X от-но мн-ва (события) $B \in \Sigma$, $P(B) > 0$.
Числ-ое выраж. вида: $E(X|B) = \sum_{i=1}^m X(\omega_i) \cdot \frac{p(\omega_i) \cdot \mathbb{1}_B(\omega_i)}{P(B)}$

2. Условн. мат. ожид. с.в. X от-но с.л.а $\mathcal{Y}$ числ-ое с.в. $E(X|Y)$, которое явл-ся ф-цией от Y , т.е.
 $E(X|Y) = h(Y) = h(Y(\omega)) \in \mathbb{R}^d$, $\omega \in \Omega$.

Приведем на мн-вах B_j она имеет вид: $\omega \in B_j \Rightarrow h(Y(\omega)) = h(y_j) =$
 $= E(X|Y=y_j) \equiv E(X|B_j) = \sum_{i=1}^m X(\omega_i) \cdot \frac{p(\omega_i) \cdot \mathbb{1}_{B_j}(\omega_i)}{P(B_j)}$, $j = \overline{1, \ell}$, $P(B_j) > 0$
 $B_j = \{\omega : Y(\omega) = y_j\}$

Теорема 2

Сформулируем след. с.в-ва условного мат. ожид.:

1. $E(\mathbb{1}_A|B) = P(A|B) = \frac{P(A \cap B)}{P(B)}$
 $E(\mathbb{1}_A|Y=y_j) = P(A|Y=y_j)$, $j = \overline{1, \ell}$
2. $E(cX|Y=y_j) = c E(X|Y=y_j)$
3. $E(X_1 + X_2|Y=y_j) = E(X_1|Y=y_j) + E(X_2|Y=y_j)$
4. $E(X|Y=y_j) = \sum_{i=1}^k x_i P(X=x_i|Y=y_j)$, $j = \overline{1, \ell}$

До-во:

$$1. E(\mathbb{1}_A|B) = \left\{ \text{Оп. 1} \right\} = \sum_{i=1}^m \mathbb{1}_A(\omega_i) \cdot \frac{p(\omega_i) \cdot \mathbb{1}_B(\omega_i)}{P(B)} = \sum_{i=1}^m \mathbb{1}_{A \cap B}(\omega_i) \cdot \frac{p(\omega_i)}{P(B)} = \frac{1}{P(B)} \sum_{i=1}^m \mathbb{1}_{A \cap B}(\omega_i) p(\omega_i) =$$

$$= \frac{1}{P(B)} \left(\sum_{i: \omega_i \in A \cap B} p(\omega_i) \cdot 1 + \sum_{i: \omega_i \notin A \cap B} p(\omega_i) \cdot 0 \right) = \frac{1}{P(B)} \cdot P(A \cap B) = P(A|B).$$

$$2. E(cX|Y=y_j) = \left\{ \text{Оп. 1} \right\} = \sum_{i=1}^m c X(\omega_i) \cdot \frac{p(\omega_i) \cdot \mathbb{1}_{B_j}(\omega_i)}{P(B_j)} = c \sum_{i=1}^m X(\omega_i) \cdot \frac{p(\omega_i) \cdot \mathbb{1}_{B_j}(\omega_i)}{P(B_j)} =$$

$$= c E(X|Y=y_j)$$

3. доказать similarly вместо $X(\omega_i)$ $X_1(\omega_i) + X_2(\omega_i)$

$$4. E(X|Y=y_j) = \left\{ X(\omega) = \sum_{i=1}^k x_i \cdot \mathbb{1}_{A_i}(\omega) \right\} = E\left(\sum_{i=1}^k x_i \mathbb{1}_{A_i} | Y=y_j\right) = \left\{ (3) \right\} = \sum_{i=1}^k E(x_i \mathbb{1}_{A_i} | Y=y_j) =$$

$$= \left\{ (2) \right\} = \sum_{i=1}^k x_i E(\mathbb{1}_{A_i} | Y=y_j) = \sum_{i=1}^k x_i P(A_i | Y=y_j) = \sum_{i=1}^k x_i P(X=x_i | Y=y_j), \quad j = \overline{1, \ell}$$

Теорема 3

1. X, Y - н.с.в. $\Rightarrow E(X|Y) = EX$
2. $EX = E E(X|Y)$

Доказ.

$$1. E(X|Y=y_j) = \left\{ \text{математическое ожидание} \right\} = \sum_{i=1}^k x_i P(X=x_i | Y=y_j) = \left\{ k, Y - \text{н.с.р.} \right\} = \left\{ P(X|Y) = P(X) \right\} =$$

$$= \sum_{i=1}^k x_i \cdot P(X=x_i) = EX, \quad \forall j \in \overline{1, \ell}$$

$$2. E(E(X|Y)) = \sum_{j=1}^{\ell} E(X|Y=y_j) \cdot P(Y=y_j) = \left\{ \text{св. б.о.} \right\} = \sum_{j=1}^{\ell} \left(\sum_{i=1}^k x_i \cdot P(X=x_i | Y=y_j) \cdot P(Y=y_j) \right)$$

$$= \sum_{j=1}^{\ell} \left(\sum_{i=1}^k x_i \cdot \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)} \right) \cdot P(Y=y_j) = \sum_{i=1}^k x_i \cdot \sum_{j=1}^{\ell} P(X=x_i, Y=y_j) = \sum_{i=1}^k x_i \cdot \sum_{j=1}^{\ell} P(X=x_i, B_j) =$$

$$= \left\{ B_1 \cap B_j = \emptyset \right\} = \sum_{i=1}^k x_i \cdot P(X=x_i, \bigcup_{j=1}^{\ell} B_j) = \left\{ \bigcup_{j=1}^{\ell} B_j = \Omega \right\} = \sum_{i=1}^k x_i P(X=x_i, \Omega) =$$

$$= P(A \cap \bigcup_{j=1}^{\ell} B_j) =$$

$$= P(\bigcup_{j=1}^{\ell} (A \cap B_j))$$

$$= \sum_{i=1}^k x_i P(X=x_i) = EX \quad \blacksquare$$

Теорема 4

$$1. E(g(X)|Y=y_j) = \sum_{i=1}^k g(x_i) \cdot P(X=x_i | Y=y_j).$$

$$2. E(g(Y)|Y=y_j) = g(y_j), \quad j \in \overline{1, \ell}$$

$$3. E(Xg(Y)|Y=y_j) = g(y_j) \cdot E(X|Y=y_j), \quad j \in \overline{1, \ell}$$

$$4. E(\psi(X, Y)|Y=y_j) = E(\psi(X, y_j)|Y=y_j), \quad j \in \overline{1, \ell}$$

Доказ.

$$1. E(g(X)|Y=y_j) = \left\{ \text{оп. 2} \right\} = \sum_{i=1}^m g(X(\omega_i)) \cdot \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} = \left\{ \text{бегущие значения } A_i \right\} =$$

$$= \sum_{i=1}^k \sum_{\omega_i \in A_i} g(X(\omega_i)) \cdot \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} = \left\{ \omega_i \in A_i \right\} = \sum_{i=1}^k \sum_{\omega_i \in A_i} g(x_i) \cdot \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} =$$

$$= \sum_{i=1}^k g(x_i) \cdot \sum_{\omega_i \in A_i} \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} = \sum_{i=1}^k g(x_i) \cdot \frac{P(A_i \cap B_j)}{P(B_j)} = \sum_{i=1}^k g(x_i) \cdot P(X=x_i | Y=y_j).$$

$$2. E(g(Y)|Y=y_j) = \left\{ \text{оп. 1} \right\} = E(g(Y)|B_j) = \sum_{i=1}^m g(Y(\omega_i)) \cdot \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} =$$

$$= \frac{1}{P(B_j)} \cdot \sum_{i: \omega_i \in B_j} g(Y(\omega_i)) \cdot \mathbb{I}_{B_j}(\omega_i) \cdot P(\omega_i) = \frac{1}{P(B_j)} \cdot \sum_{i: \omega_i \in B_j} g(y_j) \cdot P(\omega_i) =$$

$$= \frac{g(y_j)}{P(B_j)} \cdot \sum_{i: \omega_i \in B_j} P(\omega_i) = \frac{g(y_j)}{P(B_j)} \cdot P(B_j) = g(y_j).$$

$$3. E(Xg(Y)|Y=y_j) = \left\{ \text{оп. 1} \right\} = \sum_{i=1}^m X(\omega_i) \cdot g(Y(\omega_i)) \cdot \frac{P(\omega_i) \cdot \mathbb{I}_{B_j}(\omega_i)}{P(B_j)} =$$

$$= \frac{1}{P(B_j)} \sum_{i: \omega_i \in B_j} X(\omega_i) \cdot g(Y(\omega_i)) \cdot P(\omega_i) = g(y_j) \sum_{i: \omega_i \in B_j} X(\omega_i) \cdot P(\omega_i) \cdot \frac{1}{P(B_j)} \left\{ \text{оп. 1} \right\} =$$

$$= g(y_j) E(X|Y=y_j)$$

4. доказать самим

Теорема 5

$$\min_f E(X - f(Y))^2 = E(X - E(X|Y))^2$$

(X, Y) - заданы с.в.; и в Y нет
неблагоприятных, хотим применить X

Доказ.

$$\begin{aligned} E(X - f(Y))^2 &= E\left((X - E(X|Y)) + (E(X|Y) - f(Y))\right)^2 = \\ &= E(X - E(X|Y))^2 + 2E(X - E(X|Y))(E(X|Y) - f(Y)) + E(E(X|Y) - f(Y))^2 = \end{aligned}$$

Дополн., что второе произведение равно 0. Имеем:

$$E(X - E(X|Y)) \cdot (E(X|Y) - f(Y)) = \begin{cases} \text{поф. 3(2)} \\ E(Z - E(Z|Y)) \end{cases} = E(E(X - E(X|Y)) \cdot (E(X|Y) - f(Y)) | Y) =$$

Точ. 4

$$\begin{aligned} &= E\left((E(X|Y) - f(Y)) E((X - E(X|Y)) | Y)\right) = E((E(X|Y) - f(Y)) (E(X|Y) - E(E(X|Y) | Y))) = \\ &= \begin{cases} \text{поф. 4} \\ \text{пункт 2} \\ E(E(X|Y) | Y) = E(X|Y) \end{cases} = E(E(X|Y) - f(Y)) (E(X|Y) - E(X|Y)) = 0 \end{aligned}$$

Сл-но из последнего рав-ва получаем вид:

$$E(X - f(Y))^2 = E(X - E(X|Y))^2 + E(E(X|Y) - f(Y))^2 \geq E(X - E(X|Y))^2, \forall f.$$

Сл-но $\min_f E(X - f(Y))^2 = E(X - E(X|Y))^2$ ■

Замечание

Выше был рассмотрен частный случай, когда мн-во Ω конечно. Далее
будет рассмотрен общий случай произвольного Ω и будет дано определение условного мат. ожидания.
При этом оно будет удовлетворять тем св-вам: m2 - m4.
Доказ-во теоремы 5 использует только св-ва условного мат. ожидания, поэтому м.б.
сформулировано в общем случае.

Общий случай: $(\Omega, \mathcal{A}, P)$ - в.в. п-во, Ω - не обязательно конечно
 $\mathcal{A}$ - σ -алгебра на которой определена P .

Для $B \in \mathcal{A}$, $P(B) > 0$, $P(A|B) = \frac{P(A \cap B)}{P(B)}$, $\forall A \in \mathcal{A}$

$\exists \varphi(\omega): \Omega \rightarrow \mathbb{R}$ - сл. элемент, принимающий значения в $\mathbb{R}$ и мн-ве $\mathbb{R}$, и

$\exists \varphi(\omega) = \{y_1, y_2, \dots\}$
дискр. зн-ия

Обозн $B_j = \{\omega: \varphi(\omega) = y_j\} = \varphi^{-1}(\{y_j\})$, $j = 1, 2, \dots$

Тогда эти м.в. не пересекаются $B_i B_j = \emptyset$, $\bigcup_{i \in I} B_i = \Omega$.

Усл. вероят. $A \in \mathcal{A}$ от-но с.м. $\mathcal{Y}$ на-м. ф-ции от $\mathcal{Y}$: $P(A|\mathcal{Y}) = g(\mathcal{Y})$ и

$$u \in B_j \Rightarrow g(\mathcal{Y}(u)) = g(y_j) \equiv P(A|\mathcal{Y}=y_j) \equiv P(A|B_j) = \frac{P(A \cap B_j)}{P(B_j)}, \quad j=1,2,\dots$$

Введем интегр. соотношение, которому удовл. эта условная вероятность.

Обозн $\mathcal{A}_{\mathcal{Y}}$ - σ -алгебра, порожденная $\mathcal{Y}$; $\mathcal{A}_{\mathcal{Y}} = \{ \mathcal{Y}^{-1}(C) : C \subseteq \mathcal{Y} \} = \{ \Lambda : \Lambda \subseteq \bigcup_{i \in I} B_i, I \subseteq \mathbb{N}, \}$

$$\forall \Lambda \in \mathcal{A}_{\mathcal{Y}} \text{ имеем } P(A \cap \Lambda) = P(A \cap (\bigcup_{i \in I} B_i)) = P(\bigcup_{i \in I} (A \cap B_i)) = \begin{cases} P(B_i \cap B_j = \emptyset) \\ i \neq j \end{cases} = \\ = \sum_{i \in I} P(A \cap B_i) = \sum_{i \in I} P(A|B_i) \cdot P(B_i) = \sum_{i \in I} P(A|\mathcal{Y}=y_i) P(B_i)$$

$$P(A|\mathcal{Y}) \Rightarrow \int_{\Lambda} P(A|\mathcal{Y}) dP = \int_{\Lambda} P(A \cap \Lambda) dP, \quad \forall A \in \mathcal{A}, \forall \Lambda \in \mathcal{A}_{\mathcal{Y}} \quad (*)$$

$$= \sum_{i \in I} \int_{B_i} P(A|\mathcal{Y}) dP = \sum_{i \in I} P(A|\mathcal{Y}=y_i) \cdot \int_{B_i} dP = \sum_{i \in I} P(A|\mathcal{Y}=y_i) P(B_i)$$

$E(X|\mathcal{Y}) = ?$

Определим теперь условное мат. ожидание действ. дискр. с.в. $K: \Omega \rightarrow \mathbb{R}^d$ ($x_1, x_2, \dots$ - значения с.в. K) от-но дискр. с.м. $\mathcal{Y}$, описанного выше.

Условные м.о. с.в. K от-но с.м. $\mathcal{Y}$ на-м. ф-ции от $\mathcal{Y}$

$$E(K|\mathcal{Y}) = h(\mathcal{Y}): u \in B_i \Rightarrow h(\mathcal{Y}(u)) = h(y_i) = E(K|\mathcal{Y}=y_i) = \sum_{j=1}^d x_j \cdot P(K=x_j|B_i)$$

$$\text{Из этого рав-ва} \Rightarrow E(K|\mathcal{Y} \neq y_i) \cdot P(\mathcal{Y} \neq y_i) = \sum_{j=1}^d x_j P(K=x_j, \mathcal{Y} \neq y_i) \quad (**), \quad i=1,2,\dots$$

Введем интегр. соотношение, которому удовл. усл. м.о.

Имеем $\forall \Lambda \in \mathcal{A}_{\mathcal{Y}} \Rightarrow \Lambda = \bigcup_{i \in I} B_i$. Суммируя (**) по всем $i \in I$ получим:

$$\sum_{i \in I} E(K|\mathcal{Y}=y_i) P(\mathcal{Y}=y_i) = \sum_{i \in I} \sum_{j=1}^d x_j \cdot P(K=x_j, \mathcal{Y}=y_i) \Rightarrow$$

$$\sum_{j=1}^d x_j \sum_{i \in I} P(K=x_j, \mathcal{Y}=y_i) = \sum_{i \in I} E(K|\mathcal{Y}=y_i) \cdot P(\mathcal{Y}=y_i)$$

Левую часть можно записать в виде:

$$\sum_{j=1}^d x_j P(K=x_j, \Lambda) = \sum_{i \in I} E(K|\mathcal{Y}=y_i) \cdot P(\mathcal{Y}=y_i)$$

Учитывая отн. $E(K|\mathcal{Y})$ это рав-во можно записать в интегральном виде:

$$\int_{\Lambda} x dP = \int_{\Lambda} E(K|\mathcal{Y}) dP, \quad \forall \Lambda \in \mathcal{A}_{\mathcal{Y}}, \forall A \in \mathcal{A} \quad (***)$$

Левая часть имеет такой вид поскольку: $\int_{\Omega} \mathbb{1}_C dP = \sum_i c_i \mathbb{1}_{A_i}(\omega) \Rightarrow \int_{\Omega} \mathbb{1}_C dP = \sum_i c_i P(A_i \cap C)$

Для дискретного случая условное м.о. удовл. соотношению (***)

Докажем общее сф-е, в котором постулируем (*) и (**).

$\exists (\Omega, \mathcal{A}, P) - \forall \mathcal{B} \text{ п.в.}$

с по-
мощью
анализа $\mathcal{B} \subseteq \mathcal{A}$

Оп. 2 $\mathcal{B}$ $\mathcal{K}$ -интерпретируемая с.в., т.е. $E(X) < \infty$. Тогда условная м.о. с.в. X относительно $\mathcal{B}$ -подсигмы $\mathcal{B}$ на-и действ. ф-ция $\omega \in \Omega$ удовл. след. усл:

1. $E(X | \mathcal{B}) - \mathcal{B}$ -интерпретация $(h(\omega), h^{-1}(B) \in \mathcal{B})$
2. $\int_{\mathcal{A}} X dP = \int_{\mathcal{A}} E(X | \mathcal{B}) dP, \forall \mathcal{A} \in \mathcal{B}$

Таким ф-ция $E(X | \mathcal{B})$ $\exists$ и определена п.в., это следует из т. Радона-Никодима, поскольку если на измер. п.в. $(\Omega, \mathcal{B})$ рассмотреть ф-цию м.в.

$\rho(\mathcal{A}) = \int_{\mathcal{A}} X dP, \mathcal{A} \in \mathcal{B}$, то из интерпретируемости с.в. X следует, что эта ф-ция

четко-аддитивна и абсол. непр-на отно P ($\rho \ll P$); т.е. по т. Радона-Никодима $\exists \mathcal{B}$ ц.м. ф-ция $f(\omega): \rho(\mathcal{A}) = \int_{\mathcal{A}} f dP, \forall \mathcal{A} \in \mathcal{B}$ т.е. $\int_{\mathcal{A}} f dP = \int_{\mathcal{A}} X dP, \forall \mathcal{A} \in \mathcal{B}$, т.е. $f(\omega)$ удовл. определению, т.е. можно взять $E(X | \mathcal{B}) = f(\omega)$ и почти всюду!

Оп. 3 $A \in \mathcal{A}$, тогда условной вероятн. соб. A отно $\mathcal{B}$ на-и ф-ция от $\omega \in \Omega$:

$P(A | \mathcal{B}): P(A | \mathcal{B}) = E(1_A | \mathcal{B})$, т.е.

1. $P(A | \mathcal{B}) - \mathcal{B}$ -интерпретация
2. $\int_{\mathcal{A}} 1_A dP \equiv P(A \cap \mathcal{A}) = \int_{\mathcal{A}} P(A | \mathcal{B}) dP, \forall A \in \mathcal{A}; \forall \mathcal{A} \in \mathcal{B}$

Определим теперь усл. м.о. и усл. вероятность отно произвольного с.в. $Y = Y(\omega)$. При этом на $\mathcal{Y}$ задана σ -алгебра $\mathcal{B}$, т.е. $(\mathcal{Y}, \mathcal{B})$ -измер. п.в. и с.в. $Y: \Omega \rightarrow \mathcal{Y}$ измерим, т.е. $\mathcal{Y}, \mathcal{A}_Y \equiv \{Y^{-1}(B), \forall B \in \mathcal{B}\} \in \mathcal{A}$, при этом.

$\mathcal{A}_Y$ - σ -подсигма $\mathcal{A}$

Оп. 4 усл. м.о. интерпр. с.в. X отно с.в. Y на-и ф-ция от ω :

$E(X | Y): E(X | Y) \equiv E(X | \mathcal{A}_Y)$

2. усл. вероятностью соб. $A \in \mathcal{A}$ отно с.в. Y на-и ф-ция от ω $P(A | Y): P(A | Y) \equiv P(A | \mathcal{A}_Y)$

Лемма

$\exists Y: (\Omega, \mathcal{A}) \rightarrow (\mathcal{Y}, \mathcal{B})$ - измеримой с.в. и с.в. $\xi = \xi(\omega): \Omega \rightarrow \mathbb{R}^d$

интерпретация отно $\mathcal{A}_Y$, тогда $\exists h: \mathcal{Y} \rightarrow \mathbb{R}^d$:

1. $\xi(\omega) = h(Y(\omega))$ (Шварц)

$$h^{-1}(B) \in \mathcal{A}, \forall B \in \mathbb{R}^d$$

$$\mathcal{A}_Y = \{Y^{-1}(B), B \in \mathcal{B}\} \subseteq \mathcal{A}$$

Задача $\left(\frac{4}{2}\right)$ доказать последнюю лемму.

Замечание: учитывая эту лемму и то, что $E(X|Y)$ и $P(A|Y)$ получаем, что $\exists h: E(X|Y) = h(Y)$

При этом $E(X|Y=y) = h(y)$

Аналогично для $P(A|Y) = g(Y) \quad \exists g(Y) \quad \text{и} \quad P(A|Y=y) = g(y)$.

Замечание: Заметим, что если X, Y - дискретны, то $E(X|Y)$ и $P(A|Y)$ были определены в явном виде выше.

При этом они удовл. соответственно соотношениям (4*) и (4).

Эти ф-ва означают, что они удовл. последнему опр. (в дискр. случае и ф-ция непрерывна), т.к. в качестве б-аппр-а берем мн-во всех подмн-в. поэтому в силу! в дискр. случае усл. мат. ожидания и усл. вероят. опр-ны так, как это было сделано при получении соотн. (4) и (4*).

Св-ва условных мат. ожиданий

$(\Omega, \mathcal{A}, P)$.

$\mathcal{C} \subseteq \mathcal{A}$

$\xi = \xi(\omega) : (\Omega, \mathcal{A}) \rightarrow (R^d, B^d)$ - $\mathcal{C}$ -ц.ф-ия $\Leftrightarrow \xi^{-1}(B) \in \mathcal{C}, \forall B \in B^d$

1. $E(X|\mathcal{C})$ - $\mathcal{C}$ -ц.ф-ия

2. $\int_X dP = \int E(X|\mathcal{C}) dP, \forall X \in \mathcal{C}$

1. $P(A|\mathcal{C})$ - $\mathcal{C}$ -ц.ф-ия, $A \in \mathcal{A}$

2. $P(A \cap L) = \int P(A|\mathcal{C}) dP, \forall L \in \mathcal{C}, \forall A \in \mathcal{A}$

(N1) $\int_L \xi dP = 0, \forall L \in \mathcal{C}, \xi - \mathcal{C}$ -ц.ф-ия $\stackrel{?}{\Rightarrow} \xi = 0$ н.в. (Введем $A^+ = \{\xi \geq 0\}, P(A^+) = 0$)

2/2 (1) $\int_L \xi dP \geq 0, \forall L \in \mathcal{C}, \xi - \mathcal{C}$ -ц.ф-ия $\stackrel{?}{\Rightarrow} \xi \geq 0$ н.в. (1) $\xi \geq 0$ н.в.

$A^+ = \{\omega : \xi(\omega) > 0\}$ - необходимо показать $P(A^+) = P(A^+) = 0$.

$A^+ = \bigcup_{n=1}^{\infty} \{\omega : \xi(\omega) \geq \frac{1}{n}\} \Rightarrow P(A^+) \leq \sum_{n=1}^{\infty} P(\xi \geq \frac{1}{n})$

$\square \quad A_n^+ = \{\omega : \xi(\omega) \geq \frac{1}{n}\} \Rightarrow 0 = \int_{A_n^+} \xi dP \geq \frac{1}{n} \int_{A_n^+} dP = \frac{1}{n} P(A_n^+) \geq 0 \Rightarrow$

$\Rightarrow P(A_n^+) = 0 \quad \forall n$

$A^- = \{\omega : \xi(\omega) < 0\} = \bigcup_{n=1}^{\infty} \{\omega : \xi(\omega) \leq -\frac{1}{n}\} \Rightarrow P(A^-) = 0$

(N2) $\mathcal{C} = \{\emptyset, \Omega\}, \mathcal{C} = \{\emptyset, A, A^c, \Omega\}$ описано все $\mathcal{C}$ -ц.ф-ии св. ξ .

(2) $\mathcal{C} = \mathcal{X}(\Omega)$ - мн-во всех подмн-в $\Omega, \xi - \mathcal{C}$ -ц.ф-ия $\Rightarrow \xi = ?$ - $\forall \omega \in \Omega$
 $\Rightarrow \xi(\omega) = \text{const}, \forall \omega \in \Omega$

Доказано (от противного)

$$\exists f_0(u) \neq c \Rightarrow \exists u_1 \neq u_2 : f_0(u_1) \neq f_0(u_2)$$

$$\begin{array}{c} B_1 \quad B_2 \\ \text{---} \text{---} \text{---} \\ f_0(u_1) \quad f_0(u_2) \end{array} \quad \exists B_1, B_2 \in \mathcal{B}, B_1 \cap B_2 = \emptyset, B_1 \ni f_0(u_1), B_2 \ni f_0(u_2).$$

$$f_0^{-1}(B_1) \notin \mathcal{C}, \text{ т.к. } f_0^{-1}(B_1) \neq \emptyset$$

$$\{\emptyset, \Omega, A, A^c\} \quad f_0(u) = \begin{cases} c_1, & u \in A \\ c_2, & u \in A^c \end{cases} \quad (c_1 = c_2 \text{ не исключено})$$



$$\mathcal{C} = \{\emptyset, \Omega, A_1, A_2, \dots\} \quad \text{ка каждый из них есть разбиение}$$

$$\textcircled{V3} \quad \mathcal{C}_1 = \{\emptyset, \Omega\}, \mathcal{C}_2 = \{\emptyset, A, A^c, \Omega\}, \mathcal{C}_3 = \mathcal{C}(\Omega), E(X|\mathcal{C}) = ?$$

$$\textcircled{3} \quad \text{---} P(A|\mathcal{C}) = ? \quad P(A)$$

$$I \quad E(X|\mathcal{C}_1) = c$$

$$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{C}) dP = c \cdot P(\Omega), \quad \mathcal{L} = \{\emptyset, \Omega\}; \quad \mathcal{L} = \emptyset: \quad 0 = 0$$

$$\mathcal{L} = \Omega: \quad c = EX$$

$$\Rightarrow E(X|\mathcal{C}_1) = EX$$

$$\textcircled{V} \quad \text{---} E(X|\mathcal{C}_2) = \begin{cases} c_1, & u \in A \\ c_2, & u \in A^c \end{cases}$$

$$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{C}) dP, \quad \mathcal{L} = \{\emptyset, A, A^c, \Omega\} \quad (\text{минимум})$$

$$\int_{\Omega} X dP = c_1 \cdot P(A) \Rightarrow c_1 = \frac{1}{P(A)} \int_{\Omega} X dP$$

$$\int_{A^c} X dP = c_2 \cdot P(A^c) \Rightarrow c_2 = \frac{1}{P(A^c)} \int_{A^c} X dP$$

$$P(A) \neq 0, 1$$

$$P(A) \neq 0 \Rightarrow E(X|\mathcal{C}) = \begin{cases} \frac{1}{P(A)} \int_{\Omega} X dP, & u \in A \\ \frac{1}{P(A^c)} \int_{A^c} X dP, & u \in A^c \end{cases}$$

$$\text{Если } P(A) = 0, \text{ то } E(X|\mathcal{C}) = \begin{cases} c_1 = \text{const}, & u \in A \\ \int_{A^c} X dP = EX, & u \in A^c \end{cases}$$

$$\text{Если } P(A) = 1, \text{ то } \dots$$

$$\textcircled{V} \quad E(X|\mathcal{C}_3)$$

$$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{C}) dP; \quad E(X|\mathcal{C}) = X \text{ н.в.}, \text{ т.е. } \int_{\Omega} (X - E(X|\mathcal{C})) dP = 0, \quad \forall \mathcal{L} \in \mathcal{C}$$

$$P(A|\mathcal{C}) = \mathbb{1}_A(u)$$

или наоборот

(N4) $X = C \Rightarrow E(X|\mathcal{O}) = C$ н.в.

$X \geq 0 \Rightarrow E(X|\mathcal{O}) \geq 0$ н.в.

(4) $P(A \cup B|\mathcal{O}) \stackrel{?}{=} P(A|\mathcal{O}) + P(B|\mathcal{O}) - P(A \cap B|\mathcal{O})$ н.в.

$X = C \Rightarrow \int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{O}) dP \Rightarrow C P(\Omega) = \int_{\Omega} E(X|\mathcal{O}) dP$
 $\int_{\Omega} (C - E(X|\mathcal{O})) dP = 0 \quad \forall \Omega \in \mathcal{O} \Rightarrow$ в силу л. $C = E(X|\mathcal{O})$ н.в.

$0 \leq \int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{O}) dP$

Олимпиада: $E(X|\mathcal{O}) \geq 0 \Rightarrow X \geq 0$ н.в.

$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{O}) dP > 0 \quad \forall \Omega \in \mathcal{O}$

не обязательно
 $\int_{\Omega} X dP \geq 0, \forall \Omega \in \mathcal{O}$

(N5) $E(\alpha X + \beta Y|\mathcal{O}) \stackrel{?}{=} \alpha E(X|\mathcal{O}) + \beta E(Y|\mathcal{O})$ н.в.

(5) $X \geq Y \Rightarrow E(X|\mathcal{O}) \geq E(Y|\mathcal{O})$ н.в.
 св. в. монотонности

$\int_{\Omega} (\alpha X + \beta Y) dP = \int_{\Omega} E(\alpha X + \beta Y|\mathcal{O}) dP$

$\left\{ \begin{array}{l} \text{линейность} \\ \text{интеграла} \end{array} \right\} \Rightarrow \alpha \int_{\Omega} X dP + \beta \int_{\Omega} Y dP = \alpha \int_{\Omega} E(X|\mathcal{O}) dP + \beta \int_{\Omega} E(Y|\mathcal{O}) dP = \int_{\Omega} \{ \alpha E(X|\mathcal{O}) + \beta E(Y|\mathcal{O}) \} dP$

$= \int_{\Omega} (\alpha E(X|\mathcal{O}) + \beta E(Y|\mathcal{O})) dP \Rightarrow \int_{\Omega} (E(\alpha X + \beta Y|\mathcal{O}) - (\alpha E(X|\mathcal{O}) + \beta E(Y|\mathcal{O}))) dP = 0$
 $\forall \Omega \in \mathcal{O}$

В силу л. о-н. загара $\Rightarrow$

(N6) $E X = E E(X|\mathcal{O})$

(6) $P(A) \stackrel{?}{=} E P(A|\mathcal{O})$

$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{O}) dP, \quad \forall \Omega \in \mathcal{O}$

$\Omega := \Omega \in \mathcal{O} \Rightarrow E X = E(E(X|\mathcal{O}))$

(N7) $\mathcal{O}_1 \subseteq \mathcal{O}_2 \Rightarrow E(X|\mathcal{O}_1) = E(E(X|\mathcal{O}_2)|\mathcal{O}_1)$

(7) $P(A|\mathcal{O}_1), E(P(A|\mathcal{O}_2)|\mathcal{O}_1)$

$\exists Y = E(X|\mathcal{O}_2) \quad \forall \Omega \in \mathcal{O}_1$

$\int_{\Omega} Y dP = \int_{\Omega} E(Y|\mathcal{O}_1) dP$

$\int_{\Omega} X dP = \int_{\Omega} E(X|\mathcal{O}_1) dP, \quad \forall \Omega \in \mathcal{O}_1$

18) $\mathcal{A}$ и $\mathcal{B}$ - независимы $\Rightarrow E(X|\mathcal{A}) = EX$ п.8

(8) Ω и A - независимы? $P(A|\Omega) = P(A)$

$$\int_{\Omega} E(X|\mathcal{O}) dP = \int_{\Omega} X dP = E \mathbb{1}_\Omega \cdot X = \left\{ \begin{array}{l} \mathbb{1}_\Omega \cdot X \\ \text{невозможна} \\ P(\mathbb{1}_\Omega \in A, X \in B) = P(\Omega \cap X^{-1}(B)) \end{array} \right\} =$$

$$= E \mathbb{1}_\Omega EX = \int_{\Omega} EX dP, \quad \forall \Omega \in \mathcal{O} \Rightarrow \{X \in A\} \Rightarrow E(X|\mathcal{O}) = EX \text{ н.в.}$$

29) k - целое число $\Rightarrow \mathbb{E}(xy|ae) = k \mathbb{E}(y|ae)$.

(9) $A \in \mathcal{A} \Rightarrow P(A \cap B | \mathcal{A}) = 1_A \cdot P(B | \mathcal{A})$

$$\exists X \subseteq I_A \quad A \in \mathcal{A}$$

$$\int_{\Omega} E(XY | \mathcal{A}) dP = \int_{\Omega} X \cdot Y dP = \int_{\Omega} 1_A Y dP = \int_{\Omega} Y dP = \left\{ A \cap \Omega \in \mathcal{A} \right\} =$$

$$z = \int_{A \cap A} E(Y|X) dP = \int_A 1_A E(Y|X) dP = \int_A 1_A \{ \Rightarrow E(1_A \cdot Y|X) = 1_A E(Y|X) \}$$

$$\textcircled{4} \quad \mathbb{E}(\psi(x, y) | Y=y) \stackrel{?}{=} \mathbb{E}(\psi(x, y) | Y=y)$$

Статистические компьютеры

On 3 стр. структурной пар-се н-ка ($\mathbb{X}, \mathbb{F}, \{P_\theta, \theta \in \Theta\}$).

где $\mathcal{L}$ - некоторое м.в., $\mathcal{F}$ - σ -алгебра его подм.в., $P_\theta, \theta \in \Theta$ семейство вероятностей, зависящих от θ , и (4)

Обычно мк-во $\mathcal{X}$ интерпретируется как пространство наблюдений, при этом часто $\mathcal{X} \subseteq \mathbb{R}^n$, $n \geq 1$, т.е. есть сл. эл-т $X = X(\omega)$, заданный на измеримом пр-ве $(\Omega, \mathcal{A})$ и принимающий значения в пр-ве $\mathcal{X}$, при этом его распределение принадлежит сем-ву $\{P_\theta, \theta \in \Theta\}$, т.е. $P(X \in A) = P_\theta(A)$, $\forall A \in \mathcal{F}, \theta \in \Theta$ и основная задача статистики состоит в том, чтобы по конкретному наблюдению $X=x$ "оценить" или сделать вывод о распределении P_θ , при этом мы считаем, что это распределение однозначно определено параметром θ и поэтому основная задача статистики "оценить" θ .

Примеры

4. $\square$ один раз подлог. Динам. сп. в. с вер. успеха P (неизвестной), $m \in \mathbb{X} \sim \text{Bi}(n, \theta) \Leftrightarrow P_\theta(k) = C_n^k \theta^k (1-\theta)^{n-k}$, $k=0,1,\dots,n$, $m \in \mathbb{X}$ здесь $\mathbb{X} = \{0,1,\dots,n\}$, $\mathbb{F} = \mathbb{X}(\mathbb{X})$ - мн-во всех подмн-в $\mathbb{X}$ $\mathbb{Q} = (0,1)$ (интервал)

$$\forall A \in \mathcal{F}, P_\theta(A) = P_\theta(X \in A) = \sum_{k \in A} P_\theta(k) = \int_A p_\theta(x) d\mathcal{D}(x)$$

$\mathcal{D}$ - считающая мера
 $\mathcal{D}(A) \equiv |A|$ - число точек в A
 такая мера каждой $x \in \{0, 1, \dots, n\}$ задает вес 1

2.] наблюдая один раз норм. сл. в. с неизв. пар-ми, т.е. $X \sim N(\mu, \sigma^2) \in \mathcal{H}$
 $\mathcal{X} = \mathbb{R}^1, \mathcal{F} = \mathcal{B}^1, \theta = (\mu, \sigma^2) \in (\mathcal{H}) = \mathbb{R}^1 \times [0, \infty)$ (если $\sigma^2 = 0$, то N вырождена в точку μ)
 $\forall A \in \mathcal{F}, P_\theta(A) = P_\theta(X \in A) = \frac{1}{\sigma} \int_A \varphi\left(\frac{x-\mu}{\sigma}\right) dx; \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt$

Опр. 4 стат. структура $(\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in (\mathcal{H})\})$, $\{P_\theta, \theta \in (\mathcal{H})\}$ или сем-во веф. на $\mathcal{X}$ доминированы, если $\exists$ σ -конечная мера $\mathcal{D}(\cdot)$ на $\mathcal{F}$ не зависящая от θ : выполнено одно из двух экв. условий:

1. $\forall \theta \in (\mathcal{H}), P_\theta \ll \mathcal{D} \Leftrightarrow \forall A \in \mathcal{F} \mathcal{D}(A) = 0 \Rightarrow P_\theta(A) = 0, \forall \theta \in (\mathcal{H})$
 (А не обязательно, нулевая)
2. $\exists$ узн. функ. $p_\theta(x): P_\theta(A) = \int p_\theta(x) d\mathcal{D}(x) \quad \forall A \in \mathcal{F}, p_\theta(x) \geq 0$ - плотность,
 или тогда $p_\theta(x) = \frac{dP_\theta}{d\mathcal{D}}(x)$ - производная Радона-Никодина.

Эквивалентность этих двух условий следует из л. Радона-Никодина

Полной критерий доминир. стат. структура дает след. лемма.

Лемма $(+ \mathcal{D})^*$ (из лем. на эквивалент.)

Для того, чтобы сем-во $\{P_\theta, \theta \in (\mathcal{H})\}$ было доминировано $\Leftrightarrow \exists$ не более чем счетно подсем-во $(\mathcal{H})' \subseteq (\mathcal{H}) : \forall A \in \mathcal{F} P_\theta(A) = 0 \quad \forall \theta \in (\mathcal{H})'$
 $\Leftrightarrow P_\theta(A) = 0, \forall \theta \in (\mathcal{H})$

Следствие (из леммы) • если сем-во $(\mathcal{H})$ конечно или счетно, то стат. структура доминирована;

• если сем-во $\mathcal{X}$ конечно или счетно, то структура также доминирована, поскольку в этом случае в качестве домин. меры можно взять счет. меру $\mathcal{D}$.

Пример (недоминир. структура)

] $X_1, X_2, \dots$ - н.о.б. наблюдение с плотностью $p_\theta(x)$,] N - с.в., принимающая значения N и независимая от $X_1, X_2, \dots$

] $X = (X_1, \dots, X_N)$ - наблюдение
 $\mathcal{X} = \mathbb{R}^\infty, \mathcal{F} = \mathcal{B}^\infty, \forall A \in \mathcal{B}^\infty: P_\theta(A) = P_\theta(X \in A) = P_\theta((X_1, \dots, X_N) \in A) = \mathbb{E} P_\theta((X_1, \dots, X_N) \in A | N) =$
 $= \sum_{n=1}^{\infty} P_\theta((X_1, \dots, X_n) \in A | N=n) \cdot P(N=n) = \sum_{n=1}^{\infty} P_\theta((X_1, \dots, X_n) \in A) \cdot P(N=n) =$

$$= \sum_{n=1}^{\infty} \int_A p_\theta(x_1) \dots p_\theta(x_n) dx_1 \dots dx_n \cdot P(N=n) =$$

(*) задана со (*).

Используя последнюю лемму докажем след. теорему.

Теорема 6

Для того, чтобы стат. стр. $\{\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\}\}$ была доминирова-
на на измеримом п-ве $(\mathcal{X}, \mathcal{F}) \ni P^*$ задаваемом при вымириваемом
распределении, обладающем след. св-вами:

1. $P^*(A) = \sum_{\theta \in \Theta} C_\theta P_\theta(A)$, $\Theta' \subseteq \Theta$, $C_\theta > 0$, $\forall \theta \in \Theta'$, $\sum_{\theta \in \Theta'} C_\theta = 1$, Θ' - не более чем счетно.
строго вычитая или прибавляя
2. $P^* \ll \mu$ относительно μ мерн, доминирующей исх. стат. структуры.
3. $\forall A \in \mathcal{F} \quad P^*(A) = 0 \Leftrightarrow P_\theta(A) = 0, \forall \theta \in \Theta$

Доказ-во:

Согласно последней леммы, исходная стат. структура доминируема $\Leftrightarrow$
 $\exists$ не более чем счетное подмнож-во $\Theta' \subseteq \Theta$ со св-вом, описанным в лемме:

т.е. $\forall A \in \mathcal{F}: P_\theta(A) = 0, \forall \theta \in \Theta'$.

Взяв Θ' , построим $P^*(A)$ так, как в 1 пункте теоремы.

$\exists$ мера μ - доминирующей стат. структуры, т.е. $\forall A \in \mathcal{F}: \mu(A) = 0 \Rightarrow P_\theta(A) = 0$ для $\forall \theta \in \Theta \Rightarrow$
 $\Rightarrow P_\theta(A) = 0 \quad \forall \theta \in \Theta' \Rightarrow P^*(A) = 0 \Rightarrow P^* \ll \mu$.

При этом распр. P^* домин. исх. стат. стр., т.к. $P^*(A) = 0 \Rightarrow \sum_{\theta \in \Theta'} C_\theta P_\theta(A) = 0$

и $C_\theta > 0 \Rightarrow P_\theta(A) = 0 \quad \forall \theta \in \Theta' \Rightarrow$ по лемме $P_\theta(A) = 0 \quad \forall \theta \in \Theta \Rightarrow P_\theta \ll P^*$.

Таким образом 3 точки доказаны. ■

Опр. 5. статистикой $T = T(X)$, заданной на $(\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\})$ наз-м измеримое
отображ-е $(\mathcal{X}, \mathcal{F})$ в изм. п-во $(\mathcal{Y}, \mathcal{H})$ которое не зависит от θ .

Понятие статистики описывает допустимые преобразования наблюд. X , т.е. величину
 X можно рассматривать преобр. $T(X)$, которые измеримы и не зависят от θ , т.к.
 θ неизвестно.

Опр. 6

1. стат. структурами сровн. n повторным наблюдением наз-м стат. структура

$(\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\})^n$, которая имеет вид: $(\mathcal{X}^n, \mathcal{F}^n, \{P_\theta^n, \theta \in \Theta\})$, где

$\mathcal{X}^n, \mathcal{F}^n, P_\theta^n$ - прямые произведения мн-ва $\mathcal{X}, \mathcal{F}$ и P_θ ($P_\theta^n(A_1 \times \dots \times A_n) = P_\theta(A_1) \cdot \dots \cdot P_\theta(A_n)$)

Прямое произв. стат. структур описывает наблюд. вида $X \in (X_1, \dots, X_n)$ - н.о.р. наблюд.

и $X_i \sim (\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\})$

2. $\forall X = (x_1, \dots, x_n) \in \mathcal{X}^n$ выбор. распр-е на $(\mathcal{X}, \mathcal{F})$ наз-м распр-е вида:

$$P_n(A, x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{A_i}(x_i), \quad \forall A \in \mathcal{F}$$

Таким образом, если исходн. наблюдение имеет вид $(X_1, \dots, X_n)$ - н.о.р.к., то

$$P_n(A, X) = P_n(A, X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_A(X_i), \quad \forall A \in \mathcal{F}$$

Если $n: n \rightarrow \infty$, то в силу ЗБЧ $P_n(A, k) \approx E_\theta \mathbb{1}_A(k_i) = P_\theta(k_i \in A) = P_\theta(A)$, $\forall A \in \mathcal{F}$.
 т.е. при $n \rightarrow \infty$ эмпирическое распределение приближается к истинному распределению или является оценкой неизвестного распределения P_θ , в этом и состоит его смысл. И используя θ в статистике часто применяют выборочный принцип.
 $\square$ $X = (X_1, \dots, X_n)$ - наблюдение, н.о.р.; хотим оценить $\psi(P_\theta)$. Поскольку $P_n \xrightarrow{n \rightarrow \infty} P_\theta$, то естественно в качестве оценки взять $\psi(P_n)$. Тогда при опред. условиях регулярности на ψ действительно будет $\psi(P_n) \xrightarrow{n \rightarrow \infty} \psi(P_\theta)$

Пример

1. Эмпир. моменты:

$$X = (X_1, \dots, X_n)$$

$$E_\theta X_i^r, r > 0$$

В этом случае имеем:

$$= \frac{1}{n} \sum_{i=1}^n X_i^r \approx E_\theta X_i^r$$

$$\psi(P_\theta) = E_\theta X_i^r = \int x^r dP_\theta(x) \approx \psi(P_n) = \int x^r dP_n(x) =$$

2. Эмпир. ф.р.

$$\square X = (X_1, \dots, X_n) \text{ - наблюд.; } P_\theta(k_i < x)$$

$$\psi(P_\theta) = P_\theta(k_i < x) = E_\theta \mathbb{1}_{(-\infty, x)}(X_i) = \int_{-\infty}^x dP_\theta(t) \approx \psi(P_n) = \int_{-\infty}^x dP_n(t) = \int \mathbb{1}_{(-\infty, x)}(t) dP_n(t) =$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x)}(X_i) \text{ - эмпирич. ф.р. } \xrightarrow{ЗБЧ} E_\theta \mathbb{1}_{(-\infty, x)}(X_i) = P_\theta(k_i < x), n \rightarrow \infty$$

3. Эмпир. хар. ф-ции

$$X = (X_1, \dots, X_n) \text{ - наблюд.}$$

$$f_{X_i}(t) = E_\theta e^{itX_i}$$

$$\psi(P_\theta) = E_\theta e^{itX_i} = \int e^{itx} dP_\theta(x) \approx \psi(P_n) = \int e^{itx} dP_n(x) = \frac{1}{n} \sum_{i=1}^n e^{itX_i} \text{ - эмпирич. хар. ф-ция}$$

$$\xrightarrow{ЗБЧ} E_\theta e^{itX_i}, n \rightarrow \infty$$

Достаточные статистики

Понятие дост. статистики служит для того, чтобы сократить исходные данные без потери информации о θ .

$\square X = (X_1, \dots, X_n)$, $n = 10^4$ и $\square \square$ дост. стат. $T(X) = \sum_{i=1}^n X_i$, то без потери информации о θ можно перейти от многомерного вектора к одномомерной статистике.

Надеемся сначала общ. опр. дост. σ -подгруппы.

Опр. 7 $\mathcal{D} \subseteq \mathcal{F}$ в $(\mathcal{X}, \mathcal{F}, P_\theta, \theta \in \Theta)$ назовем достаточной, если $\forall A \in \mathcal{F} \exists \tilde{P}_\theta(A|\mathcal{D}) =$
 $= P_\theta(X \in A|\mathcal{D})$, которое не зависит от θ .

Задача (*)

доказать, что полн. опр. $\sim$ и интегр. стат. $Z = E_\theta(Z|\mathcal{D})$ не зависит от θ .

$$\Leftarrow Z = \mathbb{1}_A$$

$$\rightarrow P_\theta(A|\mathcal{D}) = \text{вероятность } Z = \mathbb{1}_A, \text{ берется}$$

Оп. 8 статистика $T = T(X)$ на $(\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\})$ и действ. в $(\mathcal{Y}, \mathcal{H})$ называется достаточной, если достаточна σ -подалгебра вида $T^{-1}(\mathcal{H})$

Физически $\mathcal{D}$ означает, что она содержит всю информацию о θ , поэтому если мы ее расширим, т.е. $\mathcal{D}' : \mathcal{D}' \supseteq \mathcal{D}$, то если $\mathcal{D}$ -достаточна, то $\mathcal{D}'$ -достаточна.

Задача (*)

доказать из оп. 7, что если $\mathcal{D}$ -достаточна и $\mathcal{D}_1 \supseteq \mathcal{D}$, то $\mathcal{D}_1$ - σ -подалгебра, что она достаточна.

Пример

$\square$ канон. имеют вид: $K = (K_1, \dots, K_n)$ - к.о.п.п., $K_i \sim \text{Pois}(\theta)$, $\theta > 0$, $i = \overline{1, n}$.
Докажем, что $T(K) = \sum_{i=1}^n K_i$ - дост. статистика, т.е. необходимо доказать, что

$$P_\theta(K \in A | T=t) = \sum_{x: x \in A} P_\theta(K=x | T=t) = \Phi \text{ не зависит от } \theta.$$

Из этой флос. следует, что достаточно показать, что в.ф.-сти вида

$$P_\theta(K=x | T=t) \text{ не зависит от } \theta \text{ для } \forall t.$$

$$\text{Имеем: } P_\theta(K=x | T=t) = \begin{cases} x = (x_1, \dots, x_n) \\ x_i = 0, 1, 2, \dots \\ t = 0, 1, 2, \dots \end{cases} = \frac{P_\theta(K=x, T=t)}{P_\theta(T=t)} = \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{P_\theta(K=x_1, K_2=x_2, \dots, K_n=x_n, x_1 + \dots + x_n = t)}{P_\theta(T=t)} \end{cases}$$

$$= \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{P_\theta(K_1=x_1, \dots, K_n=x_n)}{P_\theta(T=t)}, \text{ при этом условии совпадают } P_\theta(K_i=x_i, \sum K_i=t) = P_\theta(K_i=x_i) \end{cases} = \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{P_\theta(K_1=x_1) \dots P_\theta(K_n=x_n)}{P_\theta(T=t)}, x_1 + \dots + x_n = t \end{cases}$$

$$= \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{P_\theta(K_1=x_1) \dots P_\theta(K_n=x_n)}{P_\theta(T=t)}, x_1 + \dots + x_n = t \end{cases}, \begin{cases} K_i \sim \text{Pois}(\theta), \forall i \\ T = \sum_{i=1}^n K_i \sim \text{Pois}(n\theta) \end{cases} =$$

$$= \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{e^{-\theta} \frac{\theta^{x_1}}{x_1!} \dots e^{-\theta} \frac{\theta^{x_n}}{x_n!}}{e^{-n\theta} \frac{(n\theta)^t}{t!}}, x_1 + \dots + x_n = t \end{cases} = \begin{cases} 0, x_1 + \dots + x_n \neq t \\ \frac{t!}{n^t \cdot x_1! \dots x_n!}, x_1 + \dots + x_n = t \end{cases}$$

не зависит от θ ; т.е. статистика T достаточна.

На эвристическом уровне докажем, что при переходе от $K \rightarrow T(K)$ информации о θ не теряется.

Т.к. $\forall t, \forall A \in \mathcal{F}$ $P_\theta(K \in A | T=t)$ - из оп. 7 доказ. эти усл. в.ф.-сти не

зависят от θ , то можно смоделировать сл. эл.-т $Z(t) = Z_t$, распр.-е которого совпадает с указанной усл. в.ф.-стью, т.е. $P(Z_t \in A) = P_\theta(K \in A | T=t)$.

При этом ясно, что Z_t не содержит информации о θ .

Предположим теперь, что мы хотим смоделировать дост. статистику $T(X)$: $\forall t$ не зависит от $Z(t)$. Тогда, как показывает след. теорема, составной сл. эл-т $Z_{T(X)}$ имеет такое же распр.е, как и X , т.е. вся информация о θ содержится в $Z_{T(X)}$, и поскольку в Z_t информации нет, то она все в дост. ст-ке $T(X)$.

Теорема 7

Распр.е $Z_{T(X)}$ $\stackrel{d}{=} X$, т.е. $\forall A \in \mathcal{F} \quad P_\theta(Z_{T(X)} \in A) = P_\theta(X \in A), \forall \theta \in \Theta$.

Док-во:

$$\begin{aligned} P_\theta(Z_{T(X)} \in A) &= E_\theta P(Z_{T(X)} \in A | T) = \int P_\theta(Z_{T(X)} \in A | T=t) dP_{\theta T}(t) \quad \text{распр.е статистики } T \\ &= \int P_\theta(Z_t \in A | T=t) dP_{\theta T}(t) = \int_{\substack{Z_t \text{ не зависит} \\ \text{от } T, \forall t}} P_\theta(Z_t \in A) \cdot P_{\theta T}(t) = \\ &= \int_{\substack{\text{по интегр.} \\ Z_t: P_\theta(Z_t \in A) \cdot P_\theta(X \in A | T=t)}} P_\theta(X \in A | T=t) dP_{\theta T}(t) = E_\theta P_\theta(X \in A | T) = \\ &= P_\theta(X \in A), \forall A \in \mathcal{F} \quad \square \end{aligned}$$

При практическом нахождении дост. статистики непосредственно примен. опр. дост. статистики неудобно, т.к. оно требует нахождения дост. ст. $T(X)$ и затем использование условных вероятностей.

Докажем критерий факторизации, позволяющий избежать эти трудности. Рассмотрим сначала дискр. случай, т.е. $\mathcal{X} = \{x_1, x_2, \dots\}$, $\mathcal{T} = \mathcal{X}(\mathcal{X})$ - мн-во всех подмножеств $\mathcal{X}$. Тогда из сл. эл-т $P_\theta(x) = P_\theta(X=x)$, $\theta \in \Theta$, $x \in \mathcal{X}$.

Теорема 8 (критерий факторизации)

$\mathcal{X}$ - дискретно, тогда $T=T(X)$ дост. ст-ка $\Leftrightarrow \exists g(t, \theta) \geq 0, h(x) \geq 0$:

$$P_\theta(x) = P_\theta(X=x) = g(T(x), \theta) \cdot h(x), \quad \forall \theta \in \Theta$$

Заметим, что в последнем рав-ве g зависит от x только через $T(x)$, а h не зависит от θ .

Доказ.

① $T(X)$ - геом. см. $\Rightarrow P_\theta(X=x | T=t)$ не зависит от θ и t .

$$\begin{aligned} \text{Учтем } P_\theta(X=x) &= P_\theta(X=x) \cdot \begin{cases} \text{если } T(X)=t \\ \{\omega: X(\omega)=x\} \subseteq \{\omega: T(X(\omega))=t\} \end{cases} = P_\theta(X=x, T=t) = \\ &= P_\theta(X=x | T=t) \cdot P_\theta(T=t) = \begin{cases} \text{не зависит от } \theta, \text{ так как } P_\theta(T=t) = g(t, \theta) \\ \text{от } X, t. \text{ Если } P_\theta(X=x | T=t) = h(X, t), \end{cases} \\ &= g(t, \theta) \cdot h(X, t) = \begin{cases} t=T(X) \\ h(X) = h(X, T(X)) \end{cases} = g(T(X), \theta) \cdot h(X) = g(T(X), \theta) \cdot h(X). \end{aligned}$$

② $\exists$ такое $P_\theta(X) = g(T(X), \theta) \cdot h(X)$. Покажем, что ф-ция g

$P_\theta(X=x | T=t)$ не зависит от θ .

$$P_\theta(X=x | T=t) = \frac{P_\theta(X=x, T=t)}{P_\theta(T=t)} = \begin{cases} 0, & T(X) \neq t \\ \frac{P_\theta(X=x)}{P_\theta(T=t)}, & T(X) = t \end{cases}$$

$$= \begin{cases} 0, & T(X) \neq t \\ \frac{P_\theta(X=x)}{\sum_{y: T(y)=t} P_\theta(X=y)}, & T(X) = t \end{cases} = \begin{cases} 0, & T(X) \neq t \\ \frac{P_\theta(X=x) \cdot g \cdot h}{\sum_{y: T(y)=t} g(T(y), \theta) \cdot h(y)}, & T(X) = t \end{cases}$$

$$= \begin{cases} 0, & T(X) \neq t \\ \frac{g(t, \theta) \cdot h(X)}{\sum_{y: T(y)=t} g(t, \theta) \cdot h(y)}, & T(X) = t \end{cases} = \begin{cases} 0, & T(X) \neq t \\ \frac{h(X)}{\sum_{y: T(y)=t} h(y)}, & T(X) = t \end{cases}$$

- не зависит от θ

Пример

$X = (X_1, \dots, X_n)$ - н.о.н.

X_i

-1	2	7
$\frac{\theta}{11}$	$\frac{2\theta}{3}$	$1 - \frac{25\theta}{33}$

Найдем г.см. $T(X)$

$$P_\theta(X=x) = \begin{cases} X = (X_1, \dots, X_n) \\ X_i = -1, 2, 7 \end{cases} \Rightarrow P_\theta(X_1=x_1, \dots, X_n=x_n) = \prod_{i=1}^n P_\theta(X_i=x_i) = \left(\frac{\theta}{11}\right)^{x_1-2} \left(\frac{2\theta}{3}\right)^{x_2-2} \left(1 - \frac{25\theta}{33}\right)^{x_3-7}$$

$$= \left(\frac{\theta}{11}\right)^{\frac{(x_1-2)(x_1-7)}{24}} \left(\frac{2\theta}{3}\right)^{\frac{(x_2-2)(x_2-7)}{15}} \left(1 - \frac{25\theta}{33}\right)^{\frac{(x_3-7)(x_3-2)}{40}} \dots \left(\frac{\theta}{11}\right)^{\frac{(x_n-2)(x_n-7)}{24}} \left(\frac{2\theta}{3}\right)^{\frac{(x_n-2)(x_n-7)}{15}} \left(1 - \frac{25\theta}{33}\right)^{\frac{(x_n-7)(x_n-2)}{40}}$$

$$= \left(\frac{\theta}{11}\right)^{\frac{1}{24} \sum_{i=1}^n (x_i^2 - 9x_i + 14)} \left(\frac{2\theta}{3}\right)^{\frac{1}{15} \sum_{i=1}^n (x_i^2 - 6x_i - 7)} \left(1 - \frac{25\theta}{33}\right)^{\frac{1}{40} \sum_{i=1}^n (x_i^2 - x_i - 2)} = c(\theta) \cdot \theta \sum (x_i^2 \cdot \frac{1}{24} - \frac{1}{15}) + x_i(\frac{1}{15} - \frac{1}{40})$$

$$= \left(1 - \frac{25\theta}{33}\right)^{\frac{1}{40} \sum (x_i^2 - x_i)} = c(\theta) \cdot \theta \sum (x_i^2 - x_i) \cdot \left(1 - \frac{25\theta}{33}\right)^{\frac{1}{40} \sum (x_i^2 - x_i)} = g(T(X), \theta) \cdot h(X)$$

$$T(X) = \sum_{i=1}^n (x_i^2 - x_i) \text{ или } T(X) = \sum_{i=1}^n \left(x_i - \frac{1}{2}\right)^2$$

Напомним, что σ -подалгебра $\mathcal{C} \subseteq \mathcal{F}$ ($\mathcal{X}, \mathcal{F}, \{P_\theta, \theta \in \Theta\}$) называется достаточной, если $\exists P_\theta(A|\mathcal{C}), A \in \mathcal{F}$, который не зависит от θ .

Докажем критерий достаточности σ -подалгебры

Напомним, что мы рассматриваем доминировавшую стат. структуру.

При этом $\exists$ привнесённое вероят. распределение $P^*(A) = \sum_{\theta \in \Theta} C_\theta P_\theta(A)$, $\Theta' \subseteq \Theta$ - не более чем счётное подмножество Θ .

$$\sum C_\theta = 1, C_\theta \geq 0$$

При этом $P_\theta \ll P^* \Rightarrow \exists \text{ f.u.d. } P_\theta^*(x) = \frac{dP_\theta}{dP^*} - \exists$ непрерывная плотность, производная Радо-Никодэма

Теорема 9

Для того, чтобы σ -подалгебра $\mathcal{C} \subseteq \mathcal{F}$ была достаточной $\Leftrightarrow$

$\forall \theta \in \Theta \exists P_\theta^*(x) = \frac{dP_\theta}{dP^*}(x)$, которая была бы $\mathcal{C}$ измерима.

До-во.

$\Rightarrow$ $\mathcal{C}$ - достаточна $\Rightarrow \exists \text{ f.u.d. } P_\theta^*(x) = \frac{dP_\theta}{dP^*}(x) : P_\theta(A) = \int_A P^* dP^*, \forall A \in \mathcal{F}$
т.е. $\mathcal{C}$ -дост., то $\exists P_\theta(A|\mathcal{C})$ - не зависящие от θ , и тогда $P(A|\mathcal{C})$

По от. условной вероят. имеем $P_\theta(A \cap B) = \int_B P_\theta(A|\mathcal{C}) dP_\theta, \forall B \in \mathcal{C}$,

$$\Rightarrow \int_B P(A|\mathcal{C}) dP_\theta, \forall A \in \mathcal{F}, \forall B \in \mathcal{C} \quad (*)$$

Умножая эти равенства на $C_\theta, \theta \in \Theta'$ и суммируя по θ с учетом того, что P^* получим (докажем сами) $P^*(A \cap B) = \int_B P(A|\mathcal{C}) dP^*, \forall A \in \mathcal{F}, \forall B \in \mathcal{C}$

Это рав-во означает, что можно взять

$$P^*(A|\mathcal{C}) = P(A|\mathcal{C}) \text{ н.в. } \forall A \in \mathcal{F}.$$

$$\sum_{\theta \in \Theta'} \int_B P(A|\mathcal{C}) dP_\theta = \int_B P(A|\mathcal{C}) d\sum_{\theta \in \Theta'} C_\theta P_\theta = \int_B P(A|\mathcal{C}) dP^* = \int_B P(A|\mathcal{C}) dP^*$$

Покажем в рав-ве (*) $B = \mathcal{X}$ и учитывая это рав-во получим:

$$P_\theta(A) = \int P(A|\mathcal{C}) dP_\theta = \begin{cases} P(A|\mathcal{C}) = P^*(A|\mathcal{C}) \end{cases} = \int P^*(A|\mathcal{C}) dP_\theta \stackrel{!}{=} \begin{cases} P_\theta \ll P^*, \forall \theta \\ P_\theta, P^*, (x, \mathcal{C}) \end{cases} \quad \Leftrightarrow$$

По теореме Радо-Никодэма $\exists$ непрерывная плотность $P_\theta^*(x) = \frac{dP_\theta}{dP^*}(x)$

$$\Leftrightarrow \int P^*(A|\mathcal{C}) P_\theta^* dP^* = E^*(P^*(A|\mathcal{C}) P_\theta^*) = E^* E_{P_\theta^*} P_\theta^* E^*(1_A|\mathcal{C}) = \begin{cases} P_\theta^* - \mathcal{C}\text{-измерима} \end{cases} =$$

$$= E^* E^*(1_A P_\theta^*|\mathcal{C}) = E^*(1_A P_\theta^*) = \int_A P_\theta^* dP^*$$

Сравнивая левую и правую часть этого рав-ва получаем, что $P_\theta^*(x)$ - плотность, $\mathcal{C}$ измерима.

$$\Leftrightarrow \text{т.е. } \exists \mathcal{C} \text{ измеримая } P_\theta^*(x) = \frac{dP_\theta}{dP^*}(x)$$

Докажем, что $P_\theta(A|\mathcal{C})$ не зависит от θ

$$\text{По Ниссену } P_\theta(A \cap B) = E_\theta 1_A 1_B = \int_{\mathcal{C} \cap A} 1_B dP_\theta = \begin{cases} P_\theta \ll P^* \end{cases} = \int 1_A 1_B P_\theta^* dP^* =$$

$$\left\{ \begin{array}{l} \text{но ор.} \\ \text{и.о.} \end{array} \right\} = \int E^*(1_A 1_B P_\theta^* | \mathcal{O}) dP^* = \left\{ \begin{array}{l} B \in \mathcal{O} \Rightarrow \\ 1_B - \mathcal{O}\text{-цикл} \\ P_\theta^* - \mathcal{O}\text{-цикл} \end{array} \right\} = \int 1_B P_\theta^* E^*(1_A | \mathcal{O}) dP^* =$$

$$= \int P^*(A | \mathcal{O}) P_\theta^* dP^* = \int P^*(A | \mathcal{O}) dP_\theta, \quad \forall A \in \mathcal{F}, B \in \mathcal{O}, \theta \in \Theta$$

Сравнивая левую и правую части этого рав-ва получаем

$P_\theta(A | \mathcal{O}) = P^*(A | \mathcal{O})$ п.в., т.к. правая часть не зависит от θ , то не зависит и левая ■

Из этой теоремы непосредственно получаем след. следствие:

□ $\mathcal{O} \subseteq \mathcal{F}$ - дост. σ -подалгебра и $\mathcal{O} \subseteq \mathcal{O}' \subseteq \mathcal{F}$, тогда $\mathcal{O}'$ - достаточна

Если $\mathcal{O}$ достаточна, то $P_\theta^* - \mathcal{O}$ -циклы и т.к. $\mathcal{O} \subseteq \mathcal{O}'$ $P_\theta^* - \mathcal{O}'$ -циклы и сл-но $\mathcal{O}'$ - достаточна

(Задача (*) структура доминируема и и.о.р.)

Докажем теперь с помощью посл. теоремы критерий факторизации, позволяющий на практике строить дост. статистики.

Теорема 10 (критерий факторизации)

Для того, чтобы статистика $T = T(X)$, заданная на доминируемой структуре $\mathcal{O}$ была достаточной $\Leftrightarrow \exists h(x) \geq 0, g(t, \theta) \geq 0$;

$$P_\theta(x) = \frac{dP_\theta(x)}{d\nu} \cdot t = g(T(x), \theta) \cdot h(x), \quad \forall \theta \in \Theta$$

Док-во:

$T; (X, \mathcal{F}) \rightarrow (Y, \mathcal{H})$ достаточна, если $T^{-1}(\mathcal{H}) = \{T^{-1}(H), H \in \mathcal{H}\} \subseteq \mathcal{F}$

достаточна

По предыдущей теореме для этого $\Leftrightarrow P_\theta^*(x) = \frac{dP_\theta}{dP^*}(x) = T^{-1}(\mathcal{H})$ - циклы

но сформулированной выше лемме означает, что $\exists g(t, \theta) \geq 0$ - циклы:

$$P_\theta^*(x) = g(T(x), \theta), \text{ но } P_\theta(x) = \frac{dP_\theta(x)}{d\nu} = \left\{ P^* \ll \nu \right\} = \frac{dP_\theta}{dP^*}(x) \cdot \frac{dP^*}{d\nu}(x) = \left\{ h(x) \equiv \frac{dP^*}{d\nu}(x) \right\} =$$

$$= P_\theta^*(x) \cdot h(x) = g(T(x), \theta) \cdot h(x).$$

Следствие: □ $T(X)$ - дост. статистика на доминируемой структуре и $f(\cdot)$ - взаимно однозн. отображ., тогда $f(T(X))$ - дост. статистика

Док-во: следует из рав-ва м. 10: $P_\theta(x) = \{ \text{м. 10} \} = g(T(x), \theta) \cdot h(x) = \{ \exists f^{-1} \} =$

$$= g(f^{-1}(f(T(x))), \theta) \cdot h(x) = \left\{ \begin{array}{l} \bar{g}(f(T(x)), \theta) = \\ = g(f^{-1}(f(T(x))), \theta) \end{array} \right\} = \bar{g}(f(T(x)), \theta) \cdot h(x) \Rightarrow \text{м. 10. } f(T(X)) \text{ - дост.}$$

Пример (*)

□ $X = (X_1, \dots, X_n)$ - н.о.р.н., $X_i \sim R[0, \theta], \theta > 0$.

$$P_\theta(x) = \{ X = (x_1, \dots, x_n) \} = P_\theta(x_1) \cdot \dots \cdot P_\theta(x_n) = \left\{ \begin{array}{l} X_i \sim R[0, \theta] \\ P_\theta(x_i) = \begin{cases} \frac{1}{\theta}, & x_i \in [0, \theta] \\ 0, & \text{иначе} \end{cases} \end{array} \right\} = \frac{1}{\theta^n} \prod_{i=1}^n 1_{\{x_i \in [0, \theta]\}} = \frac{1}{\theta^n} \cdot 1_{\{x_n \in [0, \theta]\}}$$

Решения и стратегии

2.05.2014

Основная задача статистики - это выбор оптимальной стратегии или опт. поведения.

Для формализации этого рассмотрим один из возможных подходов.

Оп. 9 задача стат. решения описывается заданием стат. структуры $(\mathcal{X}, \mathcal{I}, \{P_\theta, \theta \in \Theta\})$ и цм. п-ва $(\Delta, \mathcal{U})$, при этом стратегии

$S(x, \mathcal{D}) \in \{0, 1\}$ называем переконную вероятность, заданную на $\mathcal{X} \times \mathcal{U}$:

1. $\forall x \in \mathcal{X} \quad S(x, \cdot)$ вероятность на $\mathcal{U}$.

2. $\forall \mathcal{D} \in \mathcal{U} \quad S(x, \mathcal{D})$ функция по x .

Обычно цм. п-во $(\Delta, \mathcal{U})$ интерпретируется как п-во решений, т.е.

$\Delta = \{\delta\}$ интерпр. как п-во решений, которые статистик хочет сделать о параметре θ , при этом при наблюдении $X=x$ решение статистик принимает согласно распр-ию $S(x, \cdot)$ заданному на $\mathcal{U}$.

Если стратегия вырождена в т. $\delta(x)$, т.е. $S(x, \mathcal{D}) = \mathbb{I}_{\mathcal{D}}(\delta(x)) \quad \forall \mathcal{D} \in \mathcal{U}, \delta(x) \in \Delta$ то в этом случае функция $\delta(x)$ называем решающей функцией и такая вырожд. стратегия идентифицируется с этой решающей функцией и называем переконизированной или вырожденной. Такая стратегия состоит в том, что статистик по наблюд. $X=x$ принимает решение $\delta(x) \in \Delta$.

Примеры

1. $\Delta = \Theta$, это интерпр. как оценивание θ .

если в этом случае ограниченные не рандомизированные стратегии, указанного выше типа, то решающая функция $\delta(x)$ называем оценкой θ .

2. Θ разбито на: $\Theta = \bigcup_{i=1}^k \Theta_i, \quad \Theta_i \cap \Theta_j = \emptyset, \quad i \neq j, \quad k \geq 2, \quad \Delta = \{\delta_1, \dots, \delta_k\}$

δ_i решение означающее в том, что $\theta \in \Theta_i, \quad i=1, \dots, k$.

$\mathcal{U} = \mathcal{X}(\Delta)$ - м-во всех подм-в Δ , при этом $\forall$ стратегия $S(x, \cdot)$ имеет вид:

$S(x, \cdot) = \begin{pmatrix} \delta_1 & \dots & \delta_k \\ \varphi_1(x) & \dots & \varphi_k(x) \end{pmatrix}$ ^{задачи вероят.} где $\varphi_i(x) \in [0, 1], \quad \forall i=1, \dots, k$, и цм. п-во, и $\sum_{i=1}^k \varphi_i(x) = 1$

3. Из партии, состоящей из n отдельных изделий для контроля качества берет n изделий, X - число дефективных; $0 \leq X \leq n$, построим модель, позволяющую судить о том, партия дефектна или годна.

$\mathcal{X} = \{0, 1, \dots, n\}, \quad \mathcal{I} = \mathcal{X}(\mathcal{X}), \quad P_\theta = \{B(n, \theta), \quad \theta \in \Theta = (0, 1), \quad \theta$ - вер. степ. дефективных

$\Delta = \{\delta_0, \delta_1\}, \quad \delta_0$ - партия принята, δ_1 - не принимается,

$\mathcal{U} = \mathcal{X}(\Delta)$

Обычно на практике поступают след. образом: назначают некоторый порог $m \in \{0, 1, \dots, n\}$ и считают, что если $X \geq m$, то δ_1 , иначе δ_0 .

Это соответствует стратегии вида: $S(x, \cdot): \begin{matrix} \delta_1 & \delta_2 \\ \uparrow_{1,0,1} & \uparrow_{1,0,1} \end{matrix} \begin{matrix} \delta_1 & \delta_2 \\ \uparrow_{1,0,1} & \uparrow_{1,0,1} \end{matrix}$

Основная задача статистики упорядочить стратегии S и согласно этому упорядочению выбрать оптимальную.

Рассмотрим один из видов подходов, основанный на ф-ции потерь

- Опр 10
1. Ф-ция потерь $L(\theta, \delta)$ назв. $\forall$ неотр. числ. ф-ция $L: (\Pi \times \Delta) \rightarrow [0, +\infty)$ при этом величины $L(\theta, \delta)$ интерпретируются как потери статистика в случае если он принимает решение δ , а истинное θ - параметр θ
 2. Средними потерями $W_S(\theta, x)$ соответ. данной стратегии S и ф-ции потерь L назв. число вида $W_S(\theta, x) = \int L(\theta, \delta) dS(x, \delta)$
 3. Ф-ция риска $R_S(\theta)$ соответ. стратегии S и ф-ции потерь L назв. величины вида $R_S(\theta) = E_{\theta} W_S(\theta, k) = \int W_S(\theta, x) dP_{\theta}(x)$
 4. Ф-ция риска $R_S(\theta)$ позволяет частично упорядочить лев. все стратегии S , т.е. одна стратегия лучше другой, или не хуже ее, если при всех θ ее ф-ция риска не больше, т.е. $R_{S_1} \leq R_{S_2} (=)$
 $R_{S_2}(\theta) \leq R_{S_1}(\theta), \forall \theta \in \Pi$

Ясно, что не все стратегии могут быть упорядочены с помощью ф-ции риска, т.е. $\exists$ несоразмерные стратегии и этот дефект исправляется с помощью байесовского подхода, описанного в книжке.

Байесовский подход: $r_S^{\theta} = \int R_S(\theta) dQ(\theta)$ (здесь Q - упереднение)

Примеры.

1. $\Delta = (\Pi) \in \mathbb{R}^1$ - оценивание θ ,
 $L(\theta, \delta) = c(\theta) (\theta - \delta)^2, c(\theta) > 0$

Рассмотрим только вариант стратегии вида: $S(x, \delta) = \int_{\Delta} (\delta(x))$, $\forall \delta \in \Delta$,
 Имеем $W_S(\theta, x) = \int L(\theta, \delta) dS(x, \delta) = L(\theta, \delta(x)) = c(\theta) (\delta(x) - \theta)^2$ $\delta(x)$ - оценка θ

$R_S(\theta) = E_{\theta} W_S(\theta, x) = c(\theta) E_{\theta} (\delta(x) - \theta)^2$ - ср. квадрат отклонения

Если $L(\theta, \delta) = c(\theta) |\theta - \delta|, c(\theta) \geq 0 \Rightarrow R_S(\theta) = c(\theta) \cdot E_{\theta} |\delta(x) - \theta|$ здесь оценка

$L(\theta, \delta) = \begin{cases} 0, & |\theta - \delta| \leq \lambda \\ a, & |\theta - \delta| > \lambda \end{cases}, a \geq 0, \lambda > 0.$ $\Rightarrow R_S(\theta) = a P_{\theta}(|\delta(x) - \theta| > \lambda)$

2. $\Pi = (\Pi_1) \cup (\Pi_2), (\Pi_1) \cap (\Pi_2) = \emptyset$ - проверка гипотез

$\Delta = \{\delta_1, \delta_2\}, \delta_1 = \{\text{решение } \theta \in (\Pi_1)\}, \delta_2 = \{\text{решение } \theta \in (\Pi_2)\}$ оценивание
 $S(x, \cdot): 1 - \varphi(x), \varphi(x), 0 \leq \varphi(x) \leq 1$

$W_S(\theta, x) = \int_{\Delta} L(\theta, \delta) dS(x, \delta) = L(\theta, \delta_1) \cdot (1 - \varphi(x)) + L(\theta, \delta_2) \cdot \varphi(x) =$
 $L(\theta, \delta) = \begin{cases} 0, & \delta = \delta_1, \theta \in (\Pi_1) \\ 0, & \delta = \delta_2, \theta \in (\Pi_1) \\ a, & \delta = \delta_1, \theta \in (\Pi_2) \\ 0, & \delta = \delta_2, \theta \in (\Pi_2) \end{cases}$
 $\begin{cases} b \varphi(x), & \theta \in (\Pi_1) \\ a(1 - \varphi(x)), & \theta \in (\Pi_2) \end{cases}$

$$R_s(\theta) = E_\theta W_s(\theta, x) = \begin{cases} E_\theta \varphi(x), & \theta \in \Pi_1 \\ E_\theta (1 - \varphi(x)), & \theta \in \Pi_2 \end{cases}$$

$\theta \in \Pi_1$ верно выбрано действие
 $\theta \in \Pi_2$ ошибка 1-го рода
 мы хотим минимизировать все ошибки

$$E_0 \varphi \leq 1$$

Докажем общую теорему, ^{доказ} которая основана на нел/к Уильямса:

$$E g(x) \geq g(E(x)),$$

если $g(x)$ выпукла вниз.

Теорема 11

$\Delta \in \mathbb{R}^k$ - ограниченный отрезок и где $\forall \theta \in \Pi$, $L(\theta, \delta)$ выпукла вниз по δ , тогда для $\forall$ рандомизированной стратегии S $\exists$ не рандомизированная стратегия $\tilde{S}$ риска которой для $\forall \theta$ не больше, риска S .

Замечание эта теорема показывает, что в ее условиях можно ограничиться только не рандом. стратегиями при минимизации функции риска.

Доказ.

$S(x, \cdot)$ - $\forall$ рандомизированная стратегия. Рассмотрим не рандом. стратегию $\delta^*(x) = \int \delta dS(x, \delta) \in \Delta$, т.е. $S^*(x, \delta) = 1_\Delta(\delta^*(x)) \forall \delta \in \Delta$.

$$\text{Далее } W_s(\theta, x) = \int_\Delta L(\theta, \delta) dS(x, \delta) \geq \begin{cases} L(\theta, \delta) \vee \text{ по } \delta \\ \text{выпукла вниз} \\ \text{и-но не-во не-есть} \end{cases} \geq L(\theta, \int_\Delta \delta dS(x, \delta)) =$$

$$= L(\theta, \delta^*(x)) = W_{S^*}(\theta, x), \forall \theta, x$$

Идентифицируя это не-во по x : $R_s(\theta) = E_\theta W_s(\theta, x) \geq E_\theta W_{S^*}(\theta, x) = R_{S^*}(\theta), \forall \theta \in \Pi$, т.е. для $\forall$ рандом. стр. S $\exists$ не-ранг. S^* , которая ее не хуже.

Задача (4)

$\Delta \in \mathbb{R}^k$, $k \geq 2$ Какие условия на Δ для этой теоремы?

$$\delta^* = \int_\Delta \delta dS(\theta, \delta), \text{ - многомерное не-во Уильямса, } \Delta - \text{ выпуклое отрезок.}$$

$\in \Delta$, т.к. S - верное распределение.

Теория оценивания

Δ на основании наблюдений $k \times x$ будем оценивать зт-е некоторой ф-ции $g(\theta): \theta \rightarrow \Gamma$.

Естественно взять $\Delta = \Gamma$

Рассмотрим случай $\Gamma \subseteq \mathbb{R}^1$

Ф-ция потерь $L(\theta, \delta): \Pi \times \Delta \rightarrow [0, +\infty)$ характеризует "близость" $g'(\theta)$ к δ .

Обычно рассматривают ф-ции потерь вида $L(\theta, \delta) = (\delta - g(\theta))^2$.

$$L(\theta, \delta) = c(\theta) (\delta - g(\theta))^2, c(\theta) \geq 0; L(\theta, \delta) = |\delta - g(\theta)|; L(\theta, \delta) = \begin{cases} (\delta - g(\theta))^2, & |\delta - g(\theta)| \leq \varepsilon \\ 2\varepsilon|\delta - g(\theta)| - \varepsilon^2, & |\delta - g(\theta)| > \varepsilon \end{cases}$$

Все эти ф-ции потерь при фикс. θ выпуклы вниз по δ , значит к ним применима пол. теорема.

$\forall S(x, \cdot) \ni$ неандом. $S^*(x, \cdot) : S^*(x, \delta) = \mathbb{1}_\delta(\delta(x)) \quad \forall \delta \in \mathcal{U}$ и не риск
 $R_S(\theta) \geq R_{S^*}(\theta) \quad \forall \theta \in \Theta$.

Потому для выпуклых вып. функций потерь можно ограничиться неандом, стратегиями, и эти неанд. стратегии можно идентифицировать с функцией $\delta(x)$.

Опр 11 1. Оценкой функции $g(\theta)$ на-ся цифровая функция вида $\delta = \delta(x) : \mathcal{X} \rightarrow \Gamma$,
 которые не зависят от θ .
 2. Риском оценки δ на-ся функция вида $R(\theta, \delta) \stackrel{\text{def}}{=} E_\theta L(\theta, \delta(x))$.

Основная задача теории оценивания — найти такую оценку, которая для
 данной функции потерь минимизируется для по всем θ функцию риска.

Покажем теперь как естественно возникает квадратичная функция потерь вида:
 $L(\theta, \delta) = c(\theta)(\delta - g(\theta))^2$

] дана некая функция потерь $L(\theta, \delta)$, которая достаточно гладкая по 2-му
 аргументу. Раскладывая по ф-ле Тейлора по 2-му аргументу с центром в $g(\theta)$
 получим $L(\theta, \delta) = L(\theta, g(\theta)) + (\delta - g(\theta)) L'_\delta(\theta, g(\theta)) + \frac{1}{2}(\delta - g(\theta))^2 L''_\delta(\theta, g(\theta)) + R$.

Т.к. $L(\theta, \delta)$ характеризует "дистанцию" между $g(\theta)$ и δ , то естественно предпо-
 ложить, что $L(\theta, g(\theta)) = 0 \quad \forall \theta \in \Theta$.

По определению $L(\theta, \delta) \geq 0 \quad \forall \theta, \delta$, и $L(\theta, g(\theta)) = 0$, тогда

$L(\theta, \delta)$ по 2-му аргументу в $g(\theta)$ при фикс. θ имеет минимум. В силу
 необход. условия экстремума $L'_\delta(\theta, g(\theta)) = 0 \quad \forall \theta \in \Theta$, т.е. в указанном выше
 разложении первые 2 слагаемых обращаются в 0.

Обозн $c(\theta) = \frac{1}{2} L''_\delta(\theta, g(\theta)) \geq 0$

Потому, отбрасывая остаточный член R , получаем $L(\theta, \delta) = c(\theta)(\delta - g(\theta))^2$

Опр 12 1. Среднеквадр. отклонением оценки $\delta(x)$ или среднекв. риском на-ся
 $R(\theta, \delta) = E_\theta(\delta(x) - g(\theta))^2$

2. Оценка $\delta(x)$ на-ся несмещенной, если она удовл. $E_\theta \delta(x) = g(\theta) \quad \forall \theta \in \Theta$.
 Для несмещенной оценки функция риска $R(\theta, \delta) = D_\theta \delta(x)$

Выше было отмечено, что основная задача теории оценивания — минимизи-
 ровать функцию риска $R(\theta, \delta)$ по $\delta \quad \forall \theta \in \Theta$, т.е. найти δ^* не зависящий от θ .
 $R(\theta, \delta) \geq R(\theta, \delta^*) \quad \forall \theta \in \Theta$. В такой постановке задача не имеет решения.
 $\forall$ фикс. θ_0 рассмотрим $\delta_{\theta_0}(x) = g(\theta_0) \quad \forall x \in \mathcal{X}$, тогда для нее

$$R(\theta, \delta_{\theta_0}) = E_{\theta_0}(\delta_{\theta_0}(x) - g(\theta))^2 = 0 \leq R(\theta, \delta_0). \text{ Меняя } \theta_0 \text{ получаем, что}$$

наилучшая оценка зависит от θ .

Естественный путь убедиться эту трудность — узить класс рассматриваемых
 оценок δ . Например, рассмотрим класс только несмещ. оценок

Однако несмещ. оценки $\ni$ не всегда.

Пример (несмещ. оценка не $\ni$).

$X \sim B_1(n, \theta), \quad \theta \in \Theta = (0, 1)$

$g(\theta) = \frac{1}{\theta}$, докажем, что не-я оценки для $g(\theta)$ не $\ni$

$$E_\theta \delta(x) = g(\theta) \quad \forall \theta \in \Theta \Leftrightarrow \sum_{k=1}^n \delta(x) C_n^k \theta^k (1-\theta)^{n-k} = \frac{1}{\theta} \quad \forall \theta \in (0, 1)$$

Если $\theta \rightarrow 0 \Rightarrow \Sigma \rightarrow \delta(0)$, $\frac{1}{\theta} \rightarrow +\infty$. Поэтому это тогда-то невозможно.

Теорема 12 (Рао - Блэкмана - Хопфмана)

Пусть функция потерь $L(\theta, \delta)$ при $\theta \in \Theta$ выпукла вниз по δ . Тогда $\delta(x)$:

$E_{\theta}[\delta(x)] < \infty$ функции вида $h(t) = E_{\theta}(\delta(x) | T=t)$, где T -дост. статистика, обладает след. св-вами:

1. $h(t)$ не зависит от θ .
2. $\delta(x)$ - несмещ. $\Rightarrow h(T)$ - несм. оценка.
3. $R(\theta, \delta) \geq R(\theta, h(T))$, $\forall \theta \in \Theta$.

Док-во.

1. Т.к. T -дост. статистика, то в силу задачи, приведенной после опр-ия дост. статистики, получаем, $E_{\theta}(\delta(x) | T)$ не зависит от $\theta \Rightarrow h(T)$ не зависит от θ .
2. $E_{\theta} h(T) = E_{\theta} (E_{\theta}(\delta(x) | T)) = E_{\theta} \delta(x) = g(\theta)$ $\forall \theta \in \Theta$
3. $R(\theta, \delta) = E_{\theta} L(\theta, \delta(x)) = E_{\theta} E_{\theta} (L(\theta, \delta(x)) | T) \geq \left\{ \begin{array}{l} \text{неф-во Йенсена} \\ g \cup \Rightarrow E(g | T) \geq g E(1 | T) \end{array} \right\} \geq$
 $\geq E_{\theta} L(\theta, E_{\theta}(\delta(x) | T)) = E_{\theta} L(\theta, h(T)) = R(\theta, h(T))$ $\forall \theta \in \Theta$ ■

Нахождение оценки $h(T)$ по оценке δ най-ше операцией проектирования оценки δ на дост. статистику T , а сама оценка $h(T)$ най-ше проекцией оценки δ .

Задача (*)

Пусть $x = (x_1, \dots, x_n)$ - н.о.б.н. и $E_{\theta} x_1 = g(\theta)$, $T = \sum_{i=1}^n x_i$ - дост. ст.

Тогда $\delta(x) = x_1$ - несм. оценка глв $g(\theta)$.

Найти $h(T) = E_{\theta}(x_1 | T)$ - ?

$$E_{\theta}(x_1 | T) \stackrel{?}{=} E(x_2 | T) = \dots = E(x_n | T) = \frac{1}{n} E_{\theta}(x_1 + \dots + x_n | T) = \frac{1}{n} E_{\theta}(T | T) = \frac{T}{n}$$

$$h(T) = \frac{T}{n}$$

$$D_{\theta} x_1 > D_{\theta} \frac{T}{n} = \frac{n D_{\theta} x_1}{n^2} = \frac{D_{\theta} x_1}{n}$$