Лекция 3. Условные распределения и условные математические ожидания. Достаточные статистики

Буре В.М., Грауэр Л.В.

ШАД

Санкт-Петербург, 2013

Содержание

Условные математические ожидания простых случайных величин

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы две простые случайные величины:

$$\xi(\omega) = \sum_{i=1}^{n} a_i I\{\omega \in A_i\},$$

где $a_k \neq a_l$, $A_k \cap A_l = \emptyset$, $k \neq l$, $\bigcup_{k=1}^{n} A_k = \Omega$, $A_k = \xi^{-1}(a_k)$ и

$$\eta(\omega) = \sum_{j=1}^{m} b_j I\{\omega \in B_j\},$$

где $b_k \neq b_l$, $B_k \cap B_l = \emptyset$, $k \neq l$, $\bigcup_{k=1}^{m} B_k = \Omega$, $B_k = \eta^{-1}(b_k)$.

Так как справедливы представления:

$$A_i = \bigcup_{j=1}^{m} (A_i \cap B_j) \quad \text{и} \quad B_j = \bigcup_{i=1}^{n} (A_i \cap B_j),$$

то можно записать случайные величины $\xi(\omega)$ и $\eta(\omega)$ в следующем виде:

$$\xi(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i I\{\omega \in A_i \cap B_j\},$$

$$\eta(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_j I\{\omega \in A_i \cap B_j\}.$$

Совместное распределение простых случайных величин $\xi(\omega)$ и $\eta(\omega)$ представимо в виде таблицы:

$\xi \setminus \eta$	b_1	b_2	$\ldots$	b_m	Σ
a_1	p_{11}	p_{12}	$\ldots$	p_{1m}	$p_{1\cdot}$
a_2	p_{21}	p_{22}	$\ldots$	p_{2m}	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
a_n	p_{n1}	p_{n2}	$\ldots$	p_{nm}	$p_{n\cdot}$
Σ	$p_{\cdot 1}$	$p_{\cdot 2}$	$\ldots$	$p_{\cdot m}$	1

где

$$p_{ij} = P\{A_i \cap B_j\} = P\{\xi^{-1}(a_i) \cap \eta^{-1}(b_j)\} = P\{\xi = a_i, \eta = b_j\},$$

$$p_{i\cdot} = \sum_{j=1}^{m} p_{ij} = \sum_{j=1}^{m} P(A_i \cap B_j) = P\left\{\bigcup_{j=1}^{m} (A_i \cap B_j)\right\} = P(A_i) =$$

$$= P\{\xi = a_i\},$$

$$p_{\cdot j} = \sum_{i=1}^{n} p_{ij} = \sum_{i=1}^{n} P(A_i \cap B_j) = P\left\{\bigcup_{i=1}^{n} (A_i \cap B_j)\right\} = P(B_j) =$$
$$= P\{\eta = b_j\},$$

$$\sum_{i=1}^{n} p_{i\cdot} = \sum_{j=1}^{m} p_{\cdot j} = 1.$$

Случайные величины $\xi(\omega)$ и $\eta(\omega)$ независимы тогда и только тогда, когда для любых i, j выполняется равенство:

$$p_{ij} = p_{i\cdot} p_{\cdot j}.$$

Далее будем предполагать, что $p_{i\cdot} > 0$ для любого i и $p_{\cdot j} > 0$ для любого j.

Рассмотрим строку с номером i таблицы совместного распределения простых случайных величин $\xi(\omega)$ и $\eta(\omega)$. Рассмотрим отношения:

$$\frac{p_{i1}}{p_{i\cdot}}, \quad \frac{p_{i2}}{p_{i\cdot}}, \quad \ldots, \quad \frac{p_{im}}{p_{i\cdot}}.$$

Все эти числа являются неотрицательными, и в сумме дают единицу. Рассмотрим условную вероятность:

$$\frac{p_{ij}}{p_{i\cdot}} = \frac{P\{A_i \cap B_j\}}{P\{A_i\}} = P\{B_j/A_i\} = P\{\eta = b_j/\xi = a_i\}.$$

Набор вероятностей $P\{\eta = b_1/\xi = a_i\}$, $P\{\eta = b_2/\xi = a_i\}$, …, $P\{\eta = b_m/\xi = a_i\}$ — условное распределение случайной величины $\eta(\omega)$ при условии, что $\xi = a_i$.

Рассмотрим столбец с номером j таблицы совместного распределения простых случайных величин $\xi(\omega)$ и $\eta(\omega)$. Рассмотрим отношения:

$$\frac{p_{1j}}{p_{\cdot j}}, \quad \frac{p_{2j}}{p_{\cdot j}}, \quad \ldots, \quad \frac{p_{nj}}{p_{\cdot j}}.$$

Все эти числа являются неотрицательными, и в сумме дают единицу. Рассмотрим условную вероятность:

$$\frac{p_{ij}}{p_{\cdot j}} = \frac{P\{A_i \cap B_j\}}{P\{B_j\}} = P\{A_i/B_j\} = P\{\xi = a_i/\eta = b_j\}.$$

Набор вероятностей $P\{\xi = a_1/\eta = b_j\}$, $P\{\xi = a_2/\eta = b_j\}$, $\ldots$, $P\{\xi = a_n/\eta = b_j\}$ — условное распределение случайной величины $\xi(\omega)$ при условии, что $\eta = b_j$.

Определение 1

Условное математическое ожидание случайной величины $\eta(\omega)$ при условии, что $\xi = a_i$, определяется равенством:

$$E(\eta/\xi = a_i) = \sum_{j=1}^{m} b_j P\{\eta = b_j/\xi = a_i\}. \tag{1}$$

Аналогично можно определить условное математическое ожидание случайной величины $\xi(\omega)$ при условии, что $\eta = b_j$:

$$E(\xi/\eta = b_j) = \sum_{i=1}^{n} a_i P\{\xi = a_i/\eta = b_j\}. \tag{2}$$

Определение 2

Условным математическим ожиданием случайной величины $\eta(\omega)$ относительно случайной величины ξ называется случайная величина $E(\eta/\xi)$, распределение которой определяется следующим образом:

$E(\eta/\xi)$	$E(\eta/\xi = a_1)$	$E(\eta/\xi = a_2)$	$\dots$	$E(\eta/\xi = a_n)$
P	$P\{\xi = a_1\}$	$P\{\xi = a_2\}$	$\dots$	$P\{\xi = a_n\}$

Случайная величина $E(\eta/\xi)$ — математическое ожидание простой случайной величины $\eta(\omega)$ относительно простой случайной величины $\xi(\omega)$.

Возможен и другой способ определения случайной величины $E(\eta/\xi)$.

Определение 3

Пусть

$$h(a_i) = E(\eta/\xi = a_i),$$

где функция $h(x)$ *задана на множестве* $\{a_1, \ldots, a_n\}$. *Тогда случайная величина* $h(\xi(\omega)) = E(\eta/\xi)$ *— условное математическое ожидание случайной величины* η *относительно случайной величины* ξ.

Аналогично определяется условное математическое ожидание $E(\xi/\eta)$ случайной величины ξ относительно случайной величины η.

Введенные таким образом определения 2 и 3 эквивалентны.

Случайная величина $E(\eta/\xi)$ определяется единственным образом с точностью до множества меры нуль, то есть, если $\zeta \overset{\text{п.н.}}{=} E(\eta/\xi)$, то случайная величина ζ также может рассматриваться как условное математическое ожидание случайной величины η относительно случайной величины ξ.

Сформулируем свойства условного математического ожидания.

- Если случайные величины ξ и η взаимно независимы, то справедливо равенство:

$$E(\eta/\xi) = E\eta.$$

Доказательство
Доказательство следует из равенства (1). Если случайные величины ξ и η взаимно независимы, то

$$P\{\eta = b_j/\xi = a_i\} = P\{\eta = b_j\}.$$

- Пусть случайная величина $\eta \overset{\text{п.н.}}{=} c$, $c \in \mathbb{R}$, тогда

$$E(\eta/\xi) \overset{\text{п.н.}}{=} c.$$

Доказательство

Как отмечалось ранее

$$E(\eta/\xi = a_i) = \sum_{j=1}^{m} b_j P\{\eta = b_j/\xi = a_i\}.$$

Если $b_j \neq c$, то

$$P\{\eta = b_j/\xi = a_i\} = 0.$$

Если $b_j = c$, то нетрудно заметить, что

$$P\{\eta = c/\xi = a_i\} = \frac{P\{\eta^{-1}(c) \cap \xi^{-1}(a_i)\}}{P\{\xi^{-1}(a_i)\}} = 1.$$

Действительно,

$$\xi^{-1}(a_i) = \xi^{-1}(a_i) \cap \eta^{-1}(c) + \xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}.$$

Следовательно,

$$P\{\xi^{-1}(a_i)\} = P\{\xi^{-1}(a_i) \cap \eta^{-1}(c)\} + P\{\xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}\}.$$

Очевидно, что $P\{\xi^{-1}(a_i) \cap \overline{\eta^{-1}(c)}\} = 0$. Таким образом, $P\{\eta = c/\xi = a_i\} = 1$.

- Справедливо равенство:

$$EE(\eta/\xi) = E\eta.$$

Доказательство

Как легко заметить,

$$EE(\eta/\xi) = \sum_{i=1}^{n} E\{\eta/\xi = a_i\}P\{\xi = a_i\} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} b_j P\{\eta = b_j/\xi = a_i\}P\{\xi = a_i\} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} b_j P\{\eta = b_j, \xi = a_i\} = \sum_{j=1}^{m} b_j P\{\eta = b_j\} = E\eta.$$

- Пусть $\varphi(x)$ — борелевская функция, тогда имеет место равенство:

$$E\{\eta\varphi(\xi)/\xi\} = \varphi(\xi)E(\eta/\xi).$$

Доказательство
Воспользуемся определением 1:

$$E\{\eta\varphi(\xi)/\xi = a_i\} = \sum_{j=1}^{m} \varphi(a_i) b_j P\{\eta = b_j/\xi = a_i\} =$$

$$= \varphi(a_i)E(\eta/\xi = a_i) = g(a_i),$$

где функция $g(x)$ задана на множестве $\{a_1, a_2, \ldots, a_n\}$. По определению 3 справедливо равенство:

$$E\{\eta\varphi(\xi)/\xi\} = g(\xi) = \varphi(\xi)E(\eta/\xi).$$

Замечание 1

В теории вероятностей все случайные величины определяются с точностью до множества меры нуль, поэтому доказываемое равенство можно записать следующим образом:

$$E\{\eta\varphi(\xi)/\xi\} \overset{\text{п.н.}}{=} \varphi(\xi)E(\eta/\xi).$$

- Пусть $\varphi(x)$ — борелевская функция. Для случайных величин ξ и η справедливо неравенство:

$$E(\eta - E(\eta/\xi))^2 \leqslant E(\eta - \varphi(\xi))^2, \qquad (3)$$

причем неравенство выполняется как равенство только, когда $\varphi(\xi) \overset{\text{п.н.}}{=} E(\eta/\xi)$.

Доказательство

Преобразуем правую часть неравенства (3). К выражению в скобках прибавим и вычтем величину $E(\eta/\xi)$, раскроем скобки, тогда

$$E(\eta - \varphi(\xi))^2 = E\{(\eta - E(\eta/\xi)) + (E(\eta/\xi) - \varphi(\xi))\}^2 =$$
$$= E(\eta - E(\eta/\xi))^2 + E(E(\eta/\xi) - \varphi(\xi))^2 +$$
$$+ 2E\{(\eta - E(\eta/\xi))(E(\eta/\xi) - \varphi(\xi))\}.$$

Рассмотрим последнее слагаемое, воспользуемся третьим свойством условного математического ожидания:

$$E\left(E\left\{(\eta - E(\eta/\xi))(E(\eta/\xi) - \varphi(\xi))/\xi\right\}\right) =$$
$$= E\left[(E(\eta/\xi) - \varphi(\xi))E\left\{(\eta - E(\eta/\xi))/\xi\right\}\right].$$

Нетрудно проверить, что условное математическое ожидание обладает свойством линейности, следовательно,

$$E\{(\eta - E(\eta/\xi))/\xi\} = E(\eta/\xi) - E\{E(\eta/\xi))/\xi\} = E(\eta/\xi) -$$
$$- E(\eta/\xi)E\{1/\xi\} = E(\eta/\xi) - E(\eta/\xi) \overset{\text{п.н.}}{=} 0.$$

Таким образом, последнее слагаемое равно нулю. Итак, получили равенство:

$$E(\eta - \varphi(\xi))^2 = E(\eta - E(\eta/\xi))^2 + E(E(\eta/\xi) - \varphi(\xi))^2. \tag{4}$$

Так как все слагаемые положительные, то очевидно, что имеет место неравенство:

$$E(\eta - \varphi(\xi))^2 \geqslant E(\eta - E(\eta/\xi))^2.$$

Из (4) следует, что неравенство выполняется как равенство только, когда $\varphi(\xi) \overset{\text{п.н.}}{=} E(\eta/\xi)$.

Замечание 2

Доказанное свойство можно записать в виде:

$$E(\eta/\xi) = \arg\inf_{\varphi(\cdot)} E(\eta - \varphi(\xi))^2, \quad E(\varphi^2(\xi)) \in \mathbb{R}.$$

Следовательно, случайная величина $E(\eta/\xi)$ дает наилучший в среднеквадратическом смысле прогноз случайной величины η по случайной величине ξ. Ранее была рассмотрена задача о наилучшем линейном прогнозе.

Условные математические ожидания произвольных случайных величин

Определение 4

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы две случайные величины $\xi(\omega)$, $\eta(\omega)$. Назовем функцию $P(B/x)$, где $B \in \mathcal{B}(\mathbb{R})$, $x \in \mathbb{R}$, условным вероятностным распределением случайной величины η при условии $\xi = x$, если выполнены следующие условия:

1. *При любом фиксированном $x \in \mathbb{R}$ функция $P(\cdot/x)$ как функция первого аргумента является вероятностной мерой на борелевской прямой $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*
2. *При любом фиксированном $B \in \mathcal{B}(\mathbb{R})$ функция $P(B/\cdot)$ как функция второго аргумента является борелевской функцией (т. е. $P(B/\cdot)\colon \mathbb{R} \to \mathbb{R}$ и полный прообраз любого борелевского множества есть борелевское множество).*
3. *Для любых борелевских множеств $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R})$ справедливо равенство:*

$$P\{\xi \in A, \eta \in B\} = \int_A P(B/x)dP_\xi. \tag{5}$$

Замечание 3

Положим в третьем пункте определения $A = \mathbb{R}$*. Тогда событие* $\xi \in \mathbb{R}$ *является достоверным:*

$$P\{\xi^{-1}(A) \cap \eta^{-1}(B)\} = P\{\xi^{-1}(\mathbb{R}) \cap \eta^{-1}(B)\} =$$
$$= P\{\Omega \cap \eta^{-1}(B)\} = P\{\eta \in B\}.$$

Следовательно, имеет место равенство:

$$P\{\eta \in B\} = \int\limits_{\mathbb{R}} P(B/x) dP_{\xi}.$$

Данное равенство является аналогом формулы полной вероятности.

Замечание 4

Определение 4 *также справедливо в векторном случае:*
$\xi = (\xi_1, \ldots, \xi_n)^T$, $\eta = (\eta_1, \ldots, \eta_m)^T$, $x \in \mathbb{R}^n$, $B \in \mathcal{B}(\mathbb{R}^m)$.

Определение 5

Пусть при любом $x \in \mathbb{R}$ *существует интеграл Лебега по вероятностной мере:* $\int_{\mathbb{R}} yP(dy/x)$. *Условным математическим ожиданием случайной величины* η *при условии* $\xi = x$ *назовем величину*

$$E(\eta/\xi = x) = \int_{\mathbb{R}} yP(dy/x),$$

где $P(B/x)$ — *условное распределение* η *при условии* $\xi = x$.

Замечание 5

Сформулированное определение полностью согласуется с определением математического ожидания, если вспомнить теорему о замене переменной под знаком интеграла Лебега.

Замечание 6

Пусть $g\colon \mathbb{R} \to \mathbb{R}$ — борелевская функция. Тогда условным математическим ожиданием случайной величины $g(\eta)$ при условии $\xi = x$ называется

$$E(g(\eta)/\xi = x) = \int\limits_{\mathbb{R}} g(y)P(dy/x),$$

если интеграл существует.

Если существует $E(\eta/\xi = x)$, то с учетом определения 4, а также с учетом определения интеграла Лебега получаем, что $E(\eta/\xi = x) = h(x)$ — борелевская функция.

Определение 6

Пусть для любого $x \in \mathbb{R}$ существует $E(\eta/\xi = x) = h(x)$. Назовем случайную величину $h(\xi) = E(\eta/\xi)$ условным математическим ожиданием случайной величины η относительно случайной величины ξ.

Во всех определениях требовалось выполнение условий для любых $x \in \mathbb{R}$, но достаточно требовать выполнение всех условий почти для всех $x \in \mathbb{R}$ относительно меры P_ξ.
Если условное математическое ожидание $E(\eta/\xi = x)$ существует и конечно почти при всех $x \in \mathbb{R}$ относительно меры P_ξ, и если существует и конечно математическое ожидание $E(\eta)$, то, как нетрудно показать, все свойства условного математического ожидания $E(\eta/\xi)$, рассмотренные в предыдущем параграфе, полностью переносятся на общий случай.

В дальнейшем, пусть при любом фиксированном $x \in \mathbb{R}$ условное распределение $P(B/x)$ имеет плотность $f_{\eta/\xi}(y/x)$ относительно меры Лебега, т. е. справедливо равенство:

$$P(B/x) = \int_B f_{\eta/\xi}(y/x)dy$$

для любого $B \in \mathcal{B}(\mathbb{R})$, $f_{\eta/\xi}(y/x) \geqslant 0$ для любых $y \in \mathbb{R}$, $x \in \mathbb{R}$. Тогда справедливо следующее представление:

$$P\{\xi \in A, \eta \in B\} = \int\limits_A P(B/x)P_\xi(dx) = \int\limits_A \int\limits_B f_{\eta/\xi}(y/x)dyP_\xi(dx). \qquad (6)$$

Кроме того, будем предполагать, что существует плотность распределения $f_\xi(x)$ относительно меры Лебега, т. е.

$$P_\xi(A) = \int\limits_A f_\xi(x)dx$$

для любого $A \in \mathcal{B}(\mathbb{R})$. Тогда равенство (6) можно переписать следующим образом:

$$P\{\xi \in A, \eta \in B\} = \int\limits_A \int\limits_B f_{\eta/\xi}(y/x) f_\xi(x)dydx. \qquad (7)$$

Будем называть плотность $f_{\eta/\xi}(y/x)$ условной плотностью распределения случайной величины η при условии $\xi = x$.
Из равенства (7) следует, что

$$P\{(\xi, \eta) \in A \times B\} = \int\limits_{A \times B} f_{\eta/\xi}(y/x) f_\xi(x)dxdy$$

для любых $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R})$. Тогда для любого $D \in \mathcal{B}(\mathbb{R}^2)$ справедливо

$$P\{(\xi, \eta) \in D\} = \int\limits_D f_{\eta/\xi}(y/x) f_\xi(x)dxdy.$$

Таким образом, доказана следующая теорема.

Теорема 1

Пусть $f_{\eta/\xi}(y/x)$ — условная плотность распределения случайной величины η при условии $\xi = x$. Пусть $f_\xi(x)$ — плотность распределения случайной величины ξ. Тогда совместная плотность распределения $f_{\xi,\eta}(x,y)$ случайных величин ξ, η может быть представлена в виде:

$$f_{\xi,\eta}(x,y) = f_{\eta/\xi}(y/x) f_\xi(x).$$

Аналогично, $f_{\xi,\eta}(x,y) = f_{\xi/\eta}(x/y) f_\eta(y)$.

Справедлива и обратная теорема.

Теорема 2

Пусть $f_{\xi,\eta}(x,y)$ *— совместная плотность распределения случайных величин* ξ, η. *Тогда имеют место равенства:*

$$f_{\xi}(x) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy, \tag{8}$$

$$f_{\eta}(y) = \int\limits_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx, \tag{9}$$

$$f_{\eta/\xi}(y/x) = \frac{f_{\xi,\eta}(x,y)}{f_{\xi}(x)}, \quad f_{\xi/\eta}(x/y) = \frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}.$$

Доказательство

Равенства (8) и (9) были доказаны ранее. Справедливо равенство

$$P\{\xi \in A, \eta \in B\} = \int_A \int_B f_{\xi,\eta}(x,y)dydx.$$

В правой его части выражение под знаком интеграла умножим и разделим на величину $f_\xi(x)$ и перейдем к повторному интегрированию:

$$P\{\xi \in A, \eta \in B\} = \int\limits_A \int\limits_B f_{\xi,\eta}(x,y)\frac{f_\xi(x)}{f_\xi(x)}dydx =$$

$$= \int\limits_A \left(\int\limits_B \frac{f_{\xi,\eta}(x,y)}{f_\xi(x)}dy \right) f_\xi(x)dx.$$

Выражение в круглых скобках — функция, зависящая от x и от B. В силу определения 4 эта функция — условное распределение случайной величины η при условии $\xi = x$. Таким образом, $f_{\xi,\eta}(x,y)/f_\xi(x)$ — плотность меры $P(B/x)$ относительно меры Лебега.

Так как плотность $f_\xi(x)$ может принимать нулевое значение, то всегда ли корректно делить на $f_\xi(x)$? Пусть $N = \{x : f_\xi(x) = 0\}$, тогда

$$P_\xi(N) = \int\limits_N f_\xi(x)dx = 0.$$

Таким образом, функция $f_{\eta/\xi}(y/x)$ не определена на множестве меры нуль, но интегрирование производится по мере P_ξ, следовательно, функцию можно доопределить на этом множестве произвольным образом, что не окажет никакого влияния на правую часть рассматриваемого в теореме выражения.

Достаточные статистики

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ.
Неизвестные параметры составляют вектор
$\theta = (\theta_1, \ldots, \theta_m) \in \Theta \subset \mathbb{R}^m$. Функция распределения генеральной совокупности ξ имеет вид $F_\xi(\cdot/\theta)$.
Будем рассматривать некоторую статистику

$$S(X_{[n]}) = \left(S_1(X_{[n]}), \ldots, S_k(X_{[n]})\right).$$

Обозначим через $P(B/s)$ условное распределение выборки $X_{[n]}$ при условии, что $S(X_{[n]}) = s$.

Распределение $P(B/s)$, где $B \in \mathcal{B}(\mathbb{R}^n)$, называется *условным распределением* выборки $X_{[n]}$ при условии, что $S(X_{[n]}) = s$, если выполнены условия:

1. При любом фиксированном значении статистики s распределение $P(\cdot/s)$ является вероятностной мерой в пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.
2. При любом фиксированном B функция $P(B/\cdot)$ как функция второго аргумента является измеримой и представляет собой условную вероятность $P\{X_{[n]} \in B/S = s\}$.

Определение 7

Статистика $S(X_{[n]})$ называется достаточной, если условное распределение выборки $X_{[n]}$, при условии, что статистика $S(X_{[n]})$ принимает значение s, не зависит от неизвестного параметра θ.

Если статистика является досточной статистикой, то можно говорить о том, что вся информация о параметре θ содержится в значении статистики $S(X_{[n]})$, и оставшийся «разброс» элементов выборки уже не зависит от θ.

Пример 1

Пусть генеральная совокупность ξ подчиняется распределению Пуассона:

$$\xi \sim \frac{\lambda^k}{k!}e^{-\lambda}, \quad \lambda > 0, \quad k \in \mathbb{Z}_+, \quad X_i \in \mathbb{Z}_+.$$

Тогда достаточная статистика имеет вид:

$$S(X_{[n]}) = \sum_{i=1}^{n} X_i.$$

Действительно, пусть $x = (x_1, \ldots, x_n)$, для всех $x_k \in \mathbb{Z}_+$, тогда

$$P\{X_{[n]} = x / S(X_{[n]}) = S(x) = s\} =$$
$$= \frac{P\{X_{[n]} = x, S(X_{[n]}) = S(x) = s\}}{P\{S(X_{[n]}) = S(x) = s\}} = \frac{P\{X_{[n]} = x\}}{P\{S(X_{[n]}) = S(x) = s\}}.$$

Предполагается, что события совместны.

Полученное представление всегда верно для любых дискретных совокупностей.
Рассмотрим числитель:

$$P\{X_{[n]} = x\} = P\{X_1 = x_1, \ldots, X_n = x_n\} = \\ = \frac{\lambda^{\sum x_i}}{x_1! \ldots x_n!} e^{-\lambda n} = \frac{\lambda^s}{x_1! \ldots x_n!} e^{-\lambda n}.$$

Рассмотрим знаменатель:

$$P\{S(X_{[n]}) = S(x) = s\} = \sum_{y: S(y)=s} P\{X_{[n]} = y\} = \\ = \frac{\lambda^s n^s}{s!} e^{-\lambda n} \sum_{y: y_1+\ldots+y_n=s} \frac{s!}{y_1! \ldots y_n!} \cdot \frac{1}{n^s} = \frac{\lambda^s n^s}{s!} e^{-\lambda n}.$$

Тогда условная вероятность имеет следующий вид:

$$\frac{P\{X_{[n]} = x\}}{P\{S(X_{[n]}) = S(x) = s\}} = \frac{s!}{x_1! \ldots x_n!} \frac{1}{n^s}.$$

И, как легко видеть, представляет собой полиномиальное распределение, в котором n исходов, вероятность каждого исхода равна $1/n$ (исходы равновероятны), s — количество испытаний, x_1 — количество исходов с номером 1, ..., x_n — количество исходов с номером n. Вероятность того, что в последовательности таких s испытаний произойдет x_1 исходов с номером 1, ..., x_n исходов с номером n, такова:

$$P\{x_1, \ldots, x_n\} = \frac{s!}{x_1! \ldots x_n!} \frac{1}{n^s}.$$

Замечание 7

Если статистика $S(X_{[n]})$ является достаточной, то взаимно однозначная функция от $S(X_{[n]})$ также достаточная статистика. Следовательно, возвращаясь к примеру 1, получаем:

$$\tilde{S}(X_{[n]}) = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X},$$

также является достаточной статистикой.

Теорема 3 (Неймана-Фишера)

Пусть имеется выборка $X_{[n]}$ *из генеральной совокупности* ξ *с функцией распределения* F_ξ. *Пусть* $\theta \in \Theta$ *— вектор неизвестных параметров распределения. Для того, чтобы статистика* $S(X_{[n]})$ *была достаточной необходимо и достаточно, чтобы функция правдоподобия* $L(X_{[n]}, \theta)$ *была представима в виде:*

$$L(X_{[n]}, \theta) = G(S(X_{[n]}), \theta) H(X_{[n]}),$$

где G, H *— измеримые функции, причем функция* G *зависит от выборки только через значения статистики* $S(X_{[n]})$, *а функция* H *не зависит от неизвестного параметра* θ.

Доказательство

Теорему Неймана-Фишера докажем только для двух случаев.

Рассмотрим «дискретный» случай.

Докажем достаточность. Рассмотрим условную вероятность:

$$P\{X_{[n]} = x/S(X_{[n]}) = s\} =$$
$$= \frac{P\{X_{[n]} = x\}}{\sum\limits_{y:S(y)=s} P\{X_{[n]} = y\}} = \frac{P_\xi(x_1)\dots P_\xi(x_n)}{\sum\limits_{y:S(y)=s} P_\xi(y_1)\dots P_\xi(y_n)} =$$
$$= \frac{L(x;\theta)}{\sum\limits_{y:S(y)=s} L(y,\theta)} = \frac{G(s,\theta)H(x)}{\sum\limits_{y:S(y)=s} G(s,\theta)H(y)} = \frac{H(x)}{\sum\limits_{y:S(y)=s} H(y)}.$$

Как видим, условная вероятность не зависит от θ.

Теперь докажем необходимость. Пусть условная вероятность не зависит от неизвестных параметров. Обозначим через $H(x)$ вероятность $P\{X_{[n]} = x/S(X_{[n]}) = S(x) = s\}$, и, как легко заметить,

$$L(x,\theta) = P\{X_{[n]} = x\} = P\{S(X_{[n]}) = s\}H(x).$$

Вероятность

$$P\{S(X_{[n]}) = S(x) = s\} = G(s,\theta)$$

зависит от неизвестных параметров x значения статистики, и справедливо равенство:

$$L(x,\theta) = G(s,\theta)H(x).$$

Рассмотрим «гладкий» случай.
Будем считать, что существует плотность распределения $f_\xi(\cdot)$, борелевские функции $S(x)$ дифференцируемы для всех $x \in \mathbb{R}^n$, $S(x) \in \mathbb{R}^k$.
Будем предполагать, что мы можем дополнить статистику S статистикой $T(x) \in \mathbb{R}^{n-k}$, при этом

$$(s, t) = (S(x), T(x)),$$

где $(x \in \mathbb{R}^n$. Будем считать, что x взаимно однозначно соответствует (s, t). Пусть (s, t) диффиренцируемо по x, и, наоборот, обратное отображение диффиренцируемо по s и t.
Плотность выборки запишем в виде:

$$f_{X_{[n]}}(x) = L(x, \theta) = \prod_{i=1}^{n} f_\xi(x_i, \theta).$$

Рассмотрим, что происходит с плотностью при взаимно однозначном отображении:

$$f_{S,T}(s,t) = f_{X_{[n]}}(x(s,t)) \left| \det\left(\frac{\partial x}{\partial(s,t)}\right) \right|.$$

Докажем достаточность (факторизация выполнена). Пусть

$$f_{X_{[n]}}(x) = G(S(x),\theta)H(x),$$

тогда

$$f_{S,T}(s,t) = G(s,\theta)H(x(s,t)) \left| \det\left(\frac{\partial x}{\partial(s,t)}\right) \right|.$$

Следовательно,

$$f_{T/S}(t/s) = \frac{f_{S,T}(s,t)}{f_S(s)} = \frac{H(x(s,t)) \left| \det\left(\frac{\partial x}{\partial(s,t)}\right) \right|}{\int\limits_{\mathbb{R}^{n-m}} H(x(s,t)) \left| \det\left(\frac{\partial x}{\partial(s,t)}\right) \right| dt}.$$

Как видим, $f_{T/S}(t/s)$ не зависит от θ, но тогда вероятность

$$P\{X_{[n]} \in B/S(X_{[n]}) = s\} = \int_{\mathbb{R}^{n-m}} I\{x(s,t) \in B\} f_{T/S}(t/s)dt$$

не зависит от θ.
Необходимость следует из того факта, что $f_{T/S}(t/s)$ не зависит от θ, но тогда справедливо равенство:

$$f_{X_{[n]}}(x) = G(S(x),\theta)H(x),$$

где $H(x) = f_{T/S}(t(x)/s(x)) \left|\det\left(\frac{\partial(s,t)}{\partial x}\right)\right|$, $G(S(x),\theta) = f_S(s(x))$.
Теорема доказана.

Пример 2

Предположим, что генеральная совокупность ξ *подчиняется нормальному распределению:* $\xi \sim N(a, \sigma^2)$, $\theta = \left(a, \sigma^2\right)^T \in \mathbb{R} \times \mathbb{R}_+$, *тогда*

$$L(X, \theta) = \prod_{i=1}^{n} f_\xi(X_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(X_i - a)^2}{2\sigma^2}} =$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{-\frac{1}{2\sigma^2} \sum\limits_{i=1}^{n} (X_i - a)^2} = G(S(X_{[n]}), \theta),$$

и функция $H(x) = 1$. *В качестве достаточной статистики* $S(X_{[n]})$ *можно выбрать:*

$$\begin{pmatrix} \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i^2 \end{pmatrix} = \begin{pmatrix} s_1(x) \\ s_2(x) \end{pmatrix} = S(x).$$

Сделаем взаимно однозначное преобразование:

$$\begin{cases} \tilde{S}_2 = a_2^* - (a_1^*)^2 = \frac{1}{n}\sum\limits_{i=1}^{n} X_i^2 - (\frac{1}{n}\sum\limits_{i=1}^{n} X_i)^2, \\ \tilde{S}_1 = \frac{1}{n}\sum\limits_{i=1}^{n} X_i. \end{cases}$$

Следовательно, $(\bar{X}, s^2)$ *- достаточная статистика, и* $(\bar{X}, \tilde{s}^2)$ *также достаточная статистика.*

Пример 3

Пусть случайная величина ξ подчиняется экспоненциальному распределению с параметром λ с плотностью:

$$f_\xi(y) = \begin{cases} \lambda e^{-\lambda y}, & y > 0; \\ 0, & y \leq 0. \end{cases}$$

Тогда

$$f_{X_{[n]}}(x) = L(x, \theta) = \begin{cases} \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}, & \text{если для } \forall\ i\colon x_i > 0; \\ 0, & \text{если } \exists\ i\colon x_i \leq 0. \end{cases}$$

Из теоремы 3 следует, что $\bar{X}$ — достаточная статистика, здесь $H(x) = 1$.

Теорема 4 (Колмогорова-Рао-Блекуэлла)

Пусть имеется выборка $X_{[n]}$ из генеральной совокупности ξ. Пусть $\theta \in \Theta \subset \mathbb{R}$ — неизвестный параметр распределения генеральной совокупности, $\hat{\theta}(X_{[n]})$ — несмещенная оценка параметра θ, $S(X_{[n]})$ — достаточная статистика. Введем следующее обозначение

$$\hat{\theta}_S = E(\hat{\theta}/S(X_{[n]})).$$

Справедливы следующие утверждения:

1. *$\hat{\theta}_S$ — несмещенная оценка параметра θ.*
2. *$\hat{\theta}_S = g(S(X_{[n]}))$, т.е. $\hat{\theta}_S$ зависит от выборки $X_{[n]}$ только через достаточную статистику S.*
3. *$D\hat{\theta}_S \leqslant D\hat{\theta}$, при этом равенство возможно только, если оценки $\hat{\theta}_S$ и $\hat{\theta}$ почти наверное совпадают.*

Доказательство
Заметим, что $\hat{\theta}_s = \int\limits_{R^n} \hat{\theta}(x)P(dx/S = s)$ не зависит от θ и измеримо по s. Таким образом, $\hat{\theta}_s$ действительно является оценкой. Очевидно, что $\hat{\theta}_S = E(\hat{\theta}/S(X_{[n]}))$ зависит от выборки только через статистику S.
По свойству условных математических ожиданий получаем, что $E\hat{\theta}_s = EE(\hat{\theta}/S) = E\hat{\theta} = \theta$, что говорит о несмещенности оценки $\hat{\theta}_s$.
Воспользуемся свойствами условных математических ожиданий:

$$D\hat{\theta} = E(\hat{\theta} - \theta \pm \hat{\theta}_s)^2 = E((\hat{\theta} - \hat{\theta}_s) + (\hat{\theta}_s - \theta))^2 =$$
$$= E(\hat{\theta} - \hat{\theta}_s)^2 + E(\hat{\theta}_s - \theta)^2 + 2E\{(\hat{\theta} - \hat{\theta}_s)(\hat{\theta}_s - \theta)\}.$$

Рассмотрим отдельно третье слагаемое:

$$E\{(\hat{\theta} - \hat{\theta}_s)(\hat{\theta}_s - \theta)\} = E\{(\hat{\theta}_s - \theta)E[(\theta - \hat{\theta}_s)/S]\} = 0.$$

Получаем

$$D\hat{\theta} = E(\hat{\theta} - \hat{\theta}_s)^2 + D\hat{\theta}_s,$$

откуда следует, что

$$D\hat{\theta} \geq D\hat{\theta}_s.$$

Причем, $D\hat{\theta} = D\hat{\theta}_s$ равносильно тому, что $E(\hat{\theta} - \hat{\theta}_s)^2 = 0$ или $\hat{\theta}$ почти наверное совпадает с $\hat{\theta}_s$.

Замечание 8

Теорема 4 переносится на многомерный случай.

Рассмотрим еще одно свойство достаточной статистики.

Лемма 1

Если достаточная статистика существует, то оценка максимального правдоподобия неизвестного параметра является функцией от достаточной статистики.

Доказательство

Пусть $\theta \in \Theta \subset \mathbb{R}^k$. Запишем функцию правдоподобия:

$$L(X_{[n]}, \theta) = G(S(X_{[n]}), \theta)H(X_{[n]}).$$

Для оценки максимального правдоподобия справедливы равенства:

$$\hat{\theta}(X_{[n]}) = \arg\max_{\theta \in \Theta} L(X_{[n]}, \theta) = \arg\max_{\theta \in \Theta} G(S(X_{[n]}), \theta) = \hat{\theta}(S(X_{[n]})),$$

что доказывает справедливость леммы.