Московский авиационный институт
(государственный технический университет)

А.Р. Панков Е.Н. Платонов

ПРАКТИКУМ

ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

ИЗДАТЕЛЬСТВО «МАИ»
Москва 2006

519.2 (075)
К 686
УДК: 519.246.2 (075.8)

Панков А.Р., Платонов Е.Н. Практикум по математической статистике: Учебное пособие. — М.: Изд-во МАИ, 2006.

Данное учебное пособие предназначено для практических занятий и самостоятельной работы студентов по математической статистике.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Данное учебное пособие предназначено для методического обеспечения практических занятий и самостоятельной работы студентов в рамках курса “Математическая статистика”, изучаемого на факультете “Прикладной математики и физики” МАИ в объеме 32 часов лекций и 32 часов практических занятий.

Пособие состоит из тринадцати разделов. Первые четыре раздела посвящены более углубленному изучению разделов курса теории вероятностей, имеющих особое значение для математической статистики (гауссовское многомерное распределение, сходимость последовательностей случайных величин, законы больших чисел, центральная предельная теорема). Разделы с пятого по двенадцатый посвящены изучению важнейших понятий и методов собственно математической статистики (выборочный метод, оценки параметров и методы их построения, проверка статистических гипотез, линейный регрессионный анализ). В последнем разделе приведены таблицы, используемые для статистических расчетов. Каждый из разделов содержит три подраздела: в первом содержатся основные определения и утверждения (в виде лемм, теорем и следствий), во втором приведены важные с методической точки зрения примеры, снабженные подробными решениями и комментариями, а в третьем подразделе приведены условия задач, предназначенных для самостоятельной работы студентов. Все задачи снабжены указаниями и ответами.

При подготовке материала пособия авторы пользовались источниками $[5, 7-9]$ (теория вероятностей), $[2, 4-8]$ (математическая статистика), $[1, 3, 4]$ (примеры и задачи), в которых могут быть найдены доказательства основных теоретических положений, дополнительные примеры и задачи.

Пособие ориентировано не только на студентов, обучающихся по специальности “Прикладная математика”, но также на студентов технических университетов, специализирующихся в области теории управления, обработки информации, экономической статистики, социологии.

СПИСОК ОСНОВНЫХ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

СВ — случайная величина или случайный вектор;

СП — случайная последовательность;

$\mathbb{R}^n$ — n-мерное (вещественное) евклидово пространство;

A^* — транспонированная матрица;

A^{-1} — обратная матрица;

I — единичная матрица;

$\mathrm{tr}[A]$ — след матрицы A;

$\det[A]$ — определитель матрицы A;

$A \geqslant 0$ — неотрицательно определенная матрица;

$\exp\{x\} = e^x$ — экспонента;

$\max(x_1, \ldots, x_n)$ — максимум из $x_1, \ldots, x_n$;

$\arg\min\limits_{x \in X} f(x)$ — точка минимума функции $f(x)$ на множестве X;

Ω — пространство элементарных событий (исходов) ω;

$\mathcal{F}$ — σ-алгебра случайных событий (подмножеств Ω);

$\mathbf{P}\{A\}$ — вероятность (вероятностная мера) события A;

$\{\Omega, \mathcal{F}, \mathbf{P}\}$ — основное вероятностное пространство;

$\varnothing$ — невозможное событие;

$F_\xi(x)$ — функция распределения СВ ξ;

$\xi \sim F(x)$ — СВ ξ имеет распределение $F(x)$;

$p_\xi(x)$ — плотность распределения СВ ξ;

$m_\xi = \mathbf{M}\{\xi\}$ — математическое ожидание (среднее) СВ ξ;

$D_\xi = \mathbf{D}\{\xi\}$ — дисперсия СВ ξ;

$\mathbf{cov}\{\xi, \eta\}$ — ковариация СВ ξ и η;

$\overset{\circ}{\xi}$ — центрированная СВ ξ;

$\Pi(\lambda)$ — распределение Пуассона с параметром λ;

$Bi(N; p)$ — биномиальное распределение с параметрами N, p;

$R[a; b]$ — равномерное распределение на отрезке $[a, b]$;

$E(\lambda)$ — экспоненциальное распределение с параметром λ;

$\mathcal{N}(m; D)$ — гауссовское распределение со средним m и дисперсией (ковариационной матрицей) D;

$\Psi_X(\lambda)$ — характеристическая функция n-мерного гауссовского распределения;

χ^2_n, $\mathcal{H}_n$ — распределение хи-квадрат с n степенями свободы;

T_r — распределение Стьюдента с r степенями свободы;

$\Phi(x)$ — интеграл вероятностей (функция Лапласа);

$\xi_n \xrightarrow{\text{p}} \xi$ — сходимость по вероятности;

$\xi_n \xrightarrow{\text{с.к.}} \xi$ — сходимость в среднем квадратическом (с.к.-сходимость);

$\xi_n \xrightarrow{\text{п.н.}} \xi$ — сходимость почти наверное;

$\xi_n \xrightarrow{\text{d}} \xi$ — сходимость по распределению (слабая сходимость);

u_α — квантиль уровня α распределения $\mathcal{N}(0; 1)$;

$k_\alpha(n)$ —- квантиль уровня α распределения $\mathcal{H}_n$;

$t_\alpha(r)$ — квантиль уровня α распределения T_r

§ 1. Гауссовский случайный вектор

1.1. Теоретические положения. Пусть $m_X \in \mathbb{R}^n$ — произвольный вектор, а $K_X \in \mathbb{R}^{n\times n}$ — симметричная неотрицательно определенная матрица ($K_X = K_X{}^\top$, $K_X \geqslant 0$).

Определение 1.1. Случайный вектор $X \in \mathbb{R}^n$ имеет *n-мерное гауссовское распределение* с параметрами $(m_X; K_X)$, если его характеристическая функция $\Psi_X(\lambda)$, $\lambda \in \mathbb{R}^n$ имеет вид

$$\Psi_X(\lambda) = \exp\left\{i\lambda^\top m_X - \frac{1}{2}\lambda^\top K_X\lambda\right\}, \tag{1.1}$$

где i — мнимая единица ($i^2 = -1$).

Обозначение: $X \sim \mathcal{N}(m_X; K_X)$.

Параметры m_X и K_X являются, соответственно, математическим ожиданием и ковариационной матрицей вектора X.

Определение 1.2. Гауссовский вектор X называется *невырожденным*, если матрица K_X — положительно определенная ($K_X > 0$).

Если $K_X > 0$, а $\Delta_X = \det[K_X]$ — определитель матрицы K_X, то X имеет в каждой точке $x \in \mathbb{R}^n$ плотность вероятности следующего вида:

$$p_X(x) = [(2\pi)^n\Delta_X]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-m_X)^\top K_X^{-1}(x-m_X)\right\} \tag{1.2}$$

Гауссовский вектор X имеет следующие основные свойства.

1) Если $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ — неслучайные матричные параметры, $X \sim \mathcal{N}(m_X; K_X)$, а $Y = AX + b$, то $X \sim \mathcal{N}(m_Y; K_Y)$, где

$$m_Y = Am_X + b; \quad K_Y = AK_XA^\top. \tag{1.3}$$

Из (1.3) следует, в частности, что любой подвектор гауссовского вектора также является гауссовским. Например, если X_i — i-ая компонента вектора X, то $X_i \sim \mathcal{N}(m_i; \sigma_i^2)$, где m_i — i-ая компонента m_X, а σ_i^2 — i-й диагональный элемент матрицы K_X.

2) Если $X \sim \mathcal{N}(m_X; K_X)$, причем компоненты $\{X_1, \ldots, X_n\}$ вектора X некоррелированы (т.е. K_X — диагональная матрица), то случайные величины $\{X_1, \ldots, X_n\}$ независимы в совокупности. Наоборот, произвольная совокупность $\{X_1, \ldots, X_n\}$ независимых гауссовских случайных величин образует гауссовский случайный вектор.

Пусть $Z = \{X^\top, Y^\top\}^\top$ — гауссовский вектор. Обозначим $m_X = \mathbf{M}\{X\}$, $m_Y = \mathbf{M}\{Y\}$, $K_X = \mathbf{cov}(X,X)$, $K_Y = \mathbf{cov}(Y,Y)$, $K_{XY} = \mathbf{cov}(X,Y)$, где $\mathbf{M}\{\cdot\}$ — математическое ожидание, а $\mathbf{cov}(\cdot,\cdot)$ — ковариация.

Теорема 1.1. *Пусть $K_Y > 0$. Условное распределение вектора X относительно Y является гауссовским с параметрами $m_{X|Y}$ и $K_{X|Y}$, где*

$$m_{X|Y} = m_X + K_{XY}K_Y^{-1}(Y - m_Y), \tag{1.4}$$

$$K_{X|Y} = K_X - K_{XY}K_Y^{-1}(K_{XY})^\top. \tag{1.5}$$

Случайный вектор $m_{X|Y}$ называется *условным математическим ожиданием* X относительно Y, а неслучайная матрица $K_{X|Y}$ — *условной ковариационной матрицей.*

Предположим, что требуется найти приближенное значение вектора $X \in \mathbb{R}^p$ по наблюдениям $Y \in \mathbb{R}^q$, причем $Z = \{X^\top, Y^\top\}^\top \in \mathbb{R}^n$, $n = p + q$ — гауссовский вектор.

Определение 1.3. *Оценкой для X по наблюдениям Y* будем называть случайный вектор $\widetilde{X} = \varphi(Y)$, где $\varphi(\cdot)$ — произвольная борелевская функция, отображающая $\mathbb{R}^q$ в $\mathbb{R}^p$. Величина

$$J(\varphi) = \mathbf{M}\left\{|X - \widetilde{X}|^2\right\} = \mathbf{M}\{|X - \varphi(Y)|^2\} \tag{1.6}$$

называется *среднеквадратической погрешностью* (с.к.-погрешностью) оценки $\widetilde{X} = \varphi(Y)$.

Определение 1.4. Оценка $\widehat{X} = \widehat{\varphi}(Y)$ называется *с.к.-оптимальной оценкой* для X по наблюдениям Y, если

$$J(\widehat{\varphi}) \leqslant J(\varphi), \quad \forall\varphi \in \mathcal{B},$$

где $\mathcal{B}$ — класс всех борелевских отображений $\mathbb{R}^q \to \mathbb{R}^p$.

Теорема 1.2. *Пусть $K_Y > 0$, тогда*

$$\widehat{X} = \widehat{\varphi}(Y) = m_X + K_{XY}K_Y^{-1}(Y - m_Y), \tag{1.7}$$

$$J(\widehat{\varphi}) = \min_{\varphi\in\mathcal{B}} J(\varphi) = \mathrm{tr}\left[K_X - K_{XY}K_Y^{-1}(K_{XY})^\top\right], \tag{1.8}$$

где $\mathrm{tr}[A]$ *— след матрицы A.*

Теорема 1.2 дает явный вид (1.7) с.к.-оптимальной оценки $\widehat{X}$ в гауссовском случае, причем из (1.4) и (1.7) следует, что $\widehat{X}$ совпадает с условным математическим ожиданием $m_{X|Y}$. При этом ковариационная матрица $K_{\Delta_X} = \mathbf{M}\{\Delta X(\Delta X)^\top\}$ ошибки $\Delta X = \widehat{X} - X$ оценки $\widehat{X}$ совпадает с условной ковариационной матрицей $K_{X|Y}$, которая в гауссовском случае оказывается неслучайной в силу (1.5).

1.2. Примеры.

Пример 1.1. Доказать, что линейное преобразование $Y = AX + b$ гауссовского вектора X также является гауссовским вектором с параметрами, определенными в (1.3).

Решение. Пусть $X \sim \mathcal{N}(m_X; K_X)$, тогда из (1.1) следует, что

$$\Psi_X(\lambda) = \exp\{i\lambda^\top m_X - \tfrac{1}{2}\lambda^\top K_X\lambda\}. \qquad (1.9)$$

Найдем характеристическую функцию $\Psi_Y(\lambda)$:

$$\Psi_Y(\lambda) = \mathbf{M}\{\exp\{\lambda^T Y\}\} = \mathbf{M}\{\exp\{i\lambda^\top(AX+b)\}\} =$$
$$= \exp\{i\lambda^\top b\}\,\mathbf{M}\{\exp\{i\lambda^\top AX\}\} = \exp\{i\lambda^\top b\}\,\Psi_X(\gamma),$$

где $\gamma = A^\top\lambda$. Используя (1.9), из последнего выражения получаем

$$\Psi_Y(\lambda) = \exp\{i\lambda^\top b\}\exp\{i\gamma^\top m_X - \tfrac{1}{2}\gamma^\top K_X\gamma\} =$$
$$= \exp\{i\lambda^\top(Am_X + b) - \tfrac{1}{2}\lambda^\top(AK_XA^\top)\lambda\}.$$

Из полученного выражения для $\Psi_Y(\lambda)$ и определения 1.1 следует:

$$Y \sim \mathcal{N}\left(Am_X + b; AK_XA^\top\right),$$

что согласуется с (1.3). ■

Пример 1.2. Двумерный гауссовский вектор $X = \{X_1, X_2\}^\top$ имеет плотность вероятности

$$p_X(x_1, x_2) = C\exp\{-\tfrac{1}{2}Q(x_1, x_2)\},$$

где $Q(x_1, x_2) = 2x_1^2 + 3x_2^2 + 2x_1x_2 - 8x_1 - 14x_2 + 18$.

Найти закон распределения СВ $\eta = 2X_1 - X_2$ и вычислить C.

Решение. Пусть $x = \{x_1, x_2\}^\top$, $X \sim \mathcal{N}(m_X; K_X)$, тогда из (1.2) получаем, что $Q(x) = (x - m_X)^\top K_X^{-1}(x - m_X)$, причем $K_X > 0$, так как X невырожден по условию.

Очевидно, что $Q(x) \geqslant Q(m_X) = 0$ для любых $x \in \mathbb{R}^2$. Поэтому m_X найдем из условия $m_X = \arg\min\limits_x Q(x)$. Воспользуемся необходимым условием экстремума:

$$\begin{cases} \dfrac{\partial Q(x)}{\partial x_1} = 4x_1 + 2x_2 - 8 = 0, \\ \dfrac{\partial Q(x)}{\partial x_2} = 2x_1 + 6x_2 - 14 = 0. \end{cases} \qquad (1.10)$$

Решая систему уравнений (1.10), находим $m_1 = \mathbf{M}\{X_1\} = 1$; $m_2 = \mathbf{M}\{X_2\} = 2$. Итак, $m_X = \{1; 2\}^\top$.

Теперь найдем K_X^{-1}, оставив в выражении для $Q(x)$ только квадратичные члены:

$$x^\top K_X^{-1}x = 2x_1^2 + 3x_2^2 + 2x_1x_2.$$

Из последнего выражения следует, что $K_X^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$. Таким образом, $K_X = \begin{bmatrix} 0{,}6 & -0{,}2 \\ -0{,}2 & 0{,}4 \end{bmatrix}$. Так как $\Delta_X = \det[K_X] = 0{,}2$, то $C = \left(2\pi\sqrt{\Delta_X}\right)^{-1} = \dfrac{\sqrt{5}}{2\pi}$.

Так как по условию $\eta = AX$, где $A = [2; -1]$, то $\eta \sim \mathcal{N}\left(Am_X; AK_XA^\top\right)$. Используя найденные параметры m_X и K_X распределения X, находим

$$m_\eta = Am_X = [2; -1]\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0;$$

$$D_\eta = AK_XA^\top = [2; -1]\begin{bmatrix} 0{,}6 & -0{,}2 \\ -0{,}2 & 0{,}4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 3{,}6.$$

Итак, $\eta \sim \mathcal{N}(0; 3{,}6)$. ■

Пример 1.3. Пусть $Y = X + \varepsilon$, где $X \sim \mathcal{N}(m_X; D_X)$, $D_X > 0$, $\varepsilon \sim \mathcal{N}(m_\varepsilon; D_\varepsilon)$, причем СВ X и ε — независимы. Найти с.к.-оптимальные оценки для Y по наблюдению X и для X по наблюдению Y.

Решение. По условию СВ X и Y образуют гауссовский вектор. Тогда по теореме 1.2

$$\widehat{Y} = \widehat{\varphi}(X) = m_{Y|X} = m_Y + K_{YX}K_X^{-1}(X - m_X).$$

$$K_{YX} = \mathbf{cov}(Y,X) = \mathbf{M}\left\{\overset{\circ}{Y}\overset{\circ}{X}\right\} = \mathbf{M}\left\{\left(\overset{\circ}{X} + \overset{\circ}{\varepsilon}\right)\overset{\circ}{X}\right\} = \mathbf{D}\{X\} + K_{\varepsilon X} = D_X,$$

так как $K_{\varepsilon X} = \mathbf{cov}(\varepsilon, X) = 0$ по условию.

$$K_X = \mathbf{cov}(X,X) = \mathbf{D}\{X\} = D_X, \quad m_Y = \mathbf{M}\{X+\varepsilon\} = m_X + m_\varepsilon.$$

Таким образом,

$$\widehat{Y} = m_X + m_\varepsilon + D_X D_X^{-1}(X - m_X) = X + m_\varepsilon.$$

Найдем теперь $\widehat{X} = \widehat{\psi}(Y) = m_{X|Y}$. По теореме 1.2

$$\widehat{X} = m_X + K_{XY}K_Y^{-1}(Y - m_Y) = m_X + D_X K_Y^{-1}(Y - m_X - m_\varepsilon).$$

$K_Y = \mathbf{cov}(Y,Y) = \mathbf{D}\{Y\} = \mathbf{D}\{X+\varepsilon\} = \mathbf{D}\{X\} + \mathbf{D}\{\varepsilon\} = D_X + D_\varepsilon$ с учетом того, что X и ε независимы. Итак,

$$\widehat{X} = m_X + \frac{D_X}{D_X + D_\varepsilon}(Y - m_X - m_\varepsilon).$$

Заметим, что дисперсию ошибки $\Delta\widehat{X} = \widehat{X} - X$ оценки можно вычислить по формуле (1.5):

$$\mathbf{D}\left\{\Delta\widehat{X}\right\} = K_{X|Y} = D_X - K_{XY}K_Y^{-1}(K_{XY})^\top =$$
$$= D_X - \frac{D_X^2}{D_X + D_\varepsilon} = D_X\left[1 - \frac{D_X}{D_X + D_\varepsilon}\right] < D_X.$$

Таким образом, информация об X в виде измерения Y позволяет уточнить тривиальную оценку для X: $\widetilde{X} = m_X$, дисперсия ошибки которой $\mathbf{D}\left\{\Delta\widetilde{X}\right\} = D_X$. ■

Пример 1.4. Характеристическая функция вектора $X = \{X_1, X_2\}^\top$ имеет вид $\Psi_X(\lambda) = \exp\left\{-(\lambda_1^2 + \lambda_2^2)\right\}$. Вычислить $\mathbf{P}\left(|X_1 - X_2| \leqslant 6\right)$.

Решение. Из (1.1) следует, что компоненты X_1 и X_2 независимы и распределены по закону $\mathcal{N}(0;2)$. Поэтому $\xi = X_1 - X_2 \sim \mathcal{N}(m_\xi; D_\xi)$, где $m_\xi = \mathbf{M}\{X_1 - X_2\} = 0$, а $D_\xi = \mathbf{D}\{X_1 - X_2\} = D_{X_1} + D_{X_2} = 4$.

Таким образом,

$$\mathbf{P}\left(|X_1 - X_2| \leqslant 6\right) = \mathbf{P}(-6 \leqslant \xi \leqslant 6) = \Phi\left(\frac{6 - m_\xi}{\sigma_\xi}\right) - \Phi\left(\frac{-6 - m_\xi}{\sigma_\xi}\right)$$
$$= \Phi(3) - \Phi(-3) = 0{,}997,$$

где $\sigma_\xi = \sqrt{D_\xi} = 2$, а $\Phi(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$ — функция Лапласа (интеграл вероятностей). ■

1.3. Задачи для самостоятельного решения.

1. Пусть компоненты вектора X — независимые гауссовские величины с параметрами $(0;1)$. Найти преобразование $Y = \varphi(X)$ такое, что $Y \sim \mathcal{N}(m_Y; K_Y)$ для заданных параметров m_Y и K_Y.

Указание. Воспользоваться факторизацией $K_Y = AA^\top$.

2. Случайный вектор $Z = \{X,Y\}^\top$ имеет плотность вероятности $p_Z(x,y) = \dfrac{\sqrt{2}}{\pi}\exp\{-Q(x,y)\}$, где $Q(x,y) = 2x^2 + 1{,}5y^2 - 2xy + 2x + 6y + 18$. Вычислить $\mathbf{P}(X > Y)$.

Ответ. $1 - \Phi\left(-\dfrac{5}{2}\sqrt{\dfrac{2}{3}}\right)$.

3. Известно, что $X \sim \mathcal{N}(m_X; K_X)$, где $m_X = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $K_X = \begin{bmatrix} 2 & 0{,}5 \\ 0{,}5 & 1 \end{bmatrix}$. Найти плотность вероятности вектора X.

Ответ. $p_X(x,y) = \dfrac{1}{\pi\sqrt{7}}\exp\left\{-\frac{2}{7}\left(x^2 + 2y^2 - xy + 4x - 9y + 11\right)\right\}$.

4. В условиях задачи 3 найти характеристическую функцию и плотность вероятности первой компоненты вектора $Y = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix} X + \begin{bmatrix} -2 \\ 2 \end{bmatrix}$.

Ответ. $\Psi_Y(\lambda) = \exp\left\{3i\lambda_1 - 2i\lambda_2 + 7\lambda_1\lambda_2 - 7\lambda_1^2 - 2\lambda_2^2\right\}$; $p_{Y_1}(x) = (28\pi)^{-\frac{1}{2}}\exp\left\{-\dfrac{(x-3)^2}{28}\right\}$.

5. Случайный вектор $Z = \{X,Y\}^\top$ имеет характеристическую функцию $\Psi_Z(\lambda_1,\lambda_2) = \exp\left\{\lambda_1\lambda_2 - \lambda_1^2 - 1{,}5\lambda_2^2\right\}$. Найти $\mathbf{M}\left\{(X-Y)^2\right\}$.

Ответ. 7.

6. Получены два наблюдения $Y_k = 2X + \varepsilon_k$, $k = 1,2$ случайной величины $X \sim \mathcal{N}(1;4)$. Ошибки наблюдения ε_1 и ε_2 не зависят друг от друга и от X и имеют распределение $\mathcal{N}(0;1)$. Найти с.к.-оптимальную оценку $\widehat{X}$ для X по наблюдениям $Y = \{Y_1, Y_2\}^\top$.

Ответ. $\widehat{X} = \dfrac{8(Y_1 + Y_2) + 1}{33}$.

7. Пусть $\{X^\top, Y^\top\}^\top$ — гауссовский вектор, $K_Y > 0$, $\widehat{X}$ — с.к.-оптимальная оценка для X по Y, а $\Delta X = \widehat{X} - X$ — ее ошибка. Доказать, что

а) $\mathbf{M}\{\Delta X\} = 0$;

б) $K_{\Delta X} = K_X - K_{XY}K_Y^{-1}(K_{XY})^\top$;

в) случайные векторы ΔX и $\widehat{X}$ независимы.

§ 2. Сходимость последовательностей случайных величин

2.1. Теоретические положения. Пусть $\{X_1, \ldots, X_n, \ldots\}$ — последовательность произвольных случайных величин (заданных на одном вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbf{P}\}$).

Определение 2.1. $X_n \to X$ *по вероятности*, если для любого $\varepsilon > 0$ $\lim\limits_{n\to\infty} \mathbf{P}\left(|X_n - X| > \varepsilon\right) = 0$.

Определение 2.2. $X_n \to X$ *в среднем квадратическом*, если $\lim\limits_{n\to\infty} \mathbf{M}\{|X_n - X|^2\} = 0$.

Определение 2.3. $X_n \to X$ *почти наверное (с вероятностью 1)*, если $\mathbf{P}\Big(\omega \in \Omega\colon \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)\Big) = 1$.

Указанные виды сходимости будем обозначать, соответственно, $X_n \xrightarrow{\text{p}} X$, $X_n \xrightarrow{\text{с.к.}} X$, $X_n \xrightarrow{\text{п.н.}} X$, $n \to \infty$.

Перечислим некоторые свойства сходящихся последовательностей.

1) Если $X_n \xrightarrow{\text{п.н.}} X$ или $X_n \xrightarrow{\text{с.к.}} X$, $n \to \infty$, то $X_n \xrightarrow{\text{p}} X$, $n \to \infty$.

2) Если $X_n \xrightarrow{\text{п.н.}} X$, $Y_n \xrightarrow{\text{п.н.}} Y$, $n \to \infty$, $a, b = \text{const}$, тогда $aX_n + bY_n \xrightarrow{\text{п.н.}} aX + bY$, $n \to \infty$.

3) Пусть $g(x)$ — произвольная борелевская функция, заданная на прямой $\mathbb{R}^1$, а A — множество точек разрыва функции $g(x)$. Если $X_n \xrightarrow{\text{п.н.}} X$, $n \to \infty$, причем $\mathbf{P}(X \in A) = 0$, то $g(X_n) \xrightarrow{\text{п.н.}} g(X)$, $n \to \infty$.

4) Если $X_n \sim \mathcal{N}(m_n; D_n)$ и $X_n \xrightarrow{\text{с.к.}} X$, $n \to \infty$, то $X \sim \mathcal{N}(m_X; D_X)$, где $m_X = \lim\limits_{n\to\infty} m_n$, $D_X = \lim\limits_{n\to\infty} D_n$, причем пределы существуют и конечны.

5) Свойства 2) и 3) справедливы для сходимости по вероятности, а свойство 2) — для с.к.-сходимости.

Для исследования сходимости последовательностей имеют важное значение следующие вспомогательные утверждения.

Теорема 2.1. *Для любых $\varepsilon > 0$ выполнено неравенство Чебышева*

$$\mathbf{P}\left(|\xi_n| > \varepsilon\right) \leqslant \frac{\mathbf{M}\{\xi_n^2\}}{\varepsilon^2}. \tag{2.1}$$

Теорема 2.2 (Борель-Кантелли). *Пусть $\{A_1, A_2, \ldots, A_n, \ldots\}$ — бесконечная последовательность случайных событий, а $B = \bigcap\limits_{n\geqslant 1} \bigcup\limits_{k\geqslant n} A_k$ — событие, состоящее в том, что произойдет бесконечно много событий $\{A_i\}$. Тогда*

1) если $\sum\limits_{k=1}^{\infty} \mathbf{P}(A_k) < \infty$, то $\mathbf{P}(B) = 0$;

2) если $\{A_1, A_2, \ldots, A_n, \ldots\}$ независимы и $\sum\limits_{k=1}^{\infty} \mathbf{P}(A_k) = \infty$, то $\mathbf{P}(B) = 1$.

2.2. Примеры.

Пример 2.1. Пусть $X_n \xrightarrow{\text{p}} X$, $Y_n \xrightarrow{\text{p}} Y$, $n \to \infty$. Показать, что $X_n + Y_n \xrightarrow{\text{p}} X + Y$, $n \to \infty$.

Решение. Пусть $\varepsilon > 0$, тогда

$$\mathbf{P}\left(|(X_n + Y_n) - (X + Y)| > \varepsilon\right) = \mathbf{P}\left(|(X_n - X) + (Y_n - Y)| > \varepsilon\right) \leqslant$$

$$\leqslant \mathbf{P}\left(|X_n - X| + |Y_n - Y| > \varepsilon\right) \leqslant$$

$$\leqslant \mathbf{P}\left(|X_n - X| > \frac{\varepsilon}{2}\right) + \mathbf{P}\left(|Y_n - Y| > \frac{\varepsilon}{2}\right) \to 0, \quad n \to \infty,$$

так как $\mathbf{P}\left(|X_n - X| > \delta\right) \to 0$, $\mathbf{P}\left(|Y_n - Y| > \delta\right) \to 0$, $n \to \infty$ для любого $\delta > 0$ по условию. ■

Пример 2.2. Пусть $\{X_n, n = 1, 2, \ldots\}$ — последовательность независимых случайных величин, $\mathbf{M}\{X_n\} = 0$, $\mathbf{D}\{X_n\} \leqslant D < \infty$, а $Y_n = n^{-\alpha} \sum\limits_{k=1}^{n} X_k$. Показать, что $Y_n \xrightarrow{\text{p}} 0$, если $\alpha > \frac{1}{2}$.

Решение. $\mathbf{M}\{Y_n\} = n^{-\alpha} \sum\limits_{k=1}^{n} \mathbf{M}\{X_k\} = 0$. Следовательно,

$$\mathbf{M}\{Y_n^2\} = \mathbf{D}\left\{n^{-\alpha} \sum_{k=1}^{n} X_k\right\} = n^{-2\alpha} \sum_{k=1}^{n} \mathbf{D}\{X_k\} \leqslant n^{-2\alpha} n D = \frac{D}{n^{2\alpha-1}}.$$

По условию $\alpha = \frac{1}{2} + \gamma$, где $\gamma > 0$. Отсюда $n^{2\alpha-1} = n^{2\gamma}$.

Из неравенства Чебышева для $\varepsilon > 0$

$$\mathbf{P}\left(|Y_n| > \varepsilon\right) \leqslant \frac{\mathbf{M}\{Y_n^2\}}{\varepsilon^2} \leqslant \frac{D}{n^{2\gamma}\varepsilon^2} \to 0, \quad n \to \infty,$$

так как $\gamma > 0$. Итак, $Y_n \xrightarrow{\text{p}} 0$, $n \to \infty$. ■

Следующие два примера показывают, что из сходимости по вероятности не следуют сходимости почти наверное и в среднем квадратическом.

Пример 2.3. Пусть X_n при каждом $n \geqslant 1$ имеет плотность вероятности $p_{X_n}(x) = \dfrac{1}{\pi}\dfrac{n}{1+(nx)^2}$. Доказать, что $X_n \xrightarrow{\text{p}} 0$, $n \to \infty$, но не сходится к нулю в среднем квадратическом.

Решение. $\mathbf{P}(|X_n| \leqslant \varepsilon) = \dfrac{1}{\pi}\displaystyle\int_{-\varepsilon}^{\varepsilon}\frac{n\,dx}{1+(nx)^2} = \frac{2}{\pi}\int_0^{n\varepsilon}\frac{dy}{1+y^2} = \frac{2}{\pi}\operatorname{arctg}(n\varepsilon) \to 1$, $n \to \infty$, так как при $\varepsilon > 0$ $\operatorname{arctg}(n\varepsilon) \to \operatorname{arctg}(\infty) = \dfrac{\pi}{2}$. Отсюда $\mathbf{P}(|X_n| > \varepsilon) = 1 - \mathbf{P}(|X_n| \leqslant \varepsilon) \to 0$, $n \to \infty$, т.е. $X_n \xrightarrow{\text{p}} 0$, $n \to \infty$.

Однако, $\mathbf{M}\{X_n^2\} = \dfrac{2}{\pi}\displaystyle\int_0^{\infty}\frac{x^2 n\,dx}{1+(nx)^2} = \frac{2}{\pi n^2}\int_0^{\infty}\frac{y^2dy}{1+y^2} = +\infty$, поэтому $\mathbf{M}\{X_n^2\}$ не сходится к нулю при $n \to \infty$, т.е. X_n не сходится к нулю в среднем квадратическом. ■

Пример 2.4. Пусть $\{X_n,\ n = 1, 2, \ldots\}$ — последовательность независимых случайных величин с распределениями $\mathbf{P}(X_n = 0) = 1 - \frac{1}{n}$, $\mathbf{P}(X_n = 1) = \frac{1}{n}$. Доказать, что $X_n \xrightarrow{\text{p}} 0$, $n \to \infty$, но X_n не сходится к нулю почти наверное.

Решение. $\mathbf{M}\{X_n^2\} = 0 \cdot \left(1 - \dfrac{1}{n}\right) + 1 \cdot \dfrac{1}{n} = \dfrac{1}{n} \to 0$, $n \to \infty$. Поэтому $X_n \xrightarrow{\text{с.к.}} 0$, $n \to \infty$ и, следовательно, $X_n \xrightarrow{\text{p}} 0$, $n \to \infty$ (что следует из неравенства Чебышева).

Обозначим $A_k = \{X_k = 1\}$. Если $X_n \xrightarrow{\text{п.н.}} 0$, то из последовательности $\{A_k,\ k = 1, 2, \ldots\}$ может произойти только конечное число событий, т.е. $\mathbf{P}(B) = \mathbf{P}\left(\displaystyle\bigcap_{n\geqslant 1}\bigcup_{k\geqslant n} A_k\right) = 0$. Так как по условию $\mathbf{P}(A_k) = \mathbf{P}(X_k = 1) = \dfrac{1}{k}$, получаем $\displaystyle\sum_{k=1}^{\infty}\mathbf{P}(A_k) = \sum_{k=1}^{\infty}\frac{1}{k} = +\infty$, что с учетом теоремы Бореля-Кантелли и независимости событий $\{A_k\}$ означает: $\mathbf{P}(B) = 1$. Итак, $\{X_n\}$ не сходится к нулю почти наверное, так как $\mathbf{P}(B) > 0$. ■

2.3. Задачи для самостоятельного решения.

1. Доказать, что из $X_n \xrightarrow{\text{p}} X$, $Y_n \xrightarrow{\text{p}} Y$, $n \to \infty$ следует $aX_n + bY_n \xrightarrow{\text{p}} aX + bY$, $n \to \infty$ для любых чисел a, b.

Указание. Воспользоваться результатом примера 2.1.

2. Пусть $\mathbf{P}(|X| < \infty) = 1$, а $X_n = \dfrac{X}{n}$. Доказать, что $X_n \xrightarrow{\text{п.н.}} 0$, $n \to \infty$.

Указание. Воспользоваться определением п.н.-сходимости.

3. Доказать, что из сходимости почти наверное не следует сходимость в среднем квадратическом.

Указание. Рассмотреть случайную величину X_n с распределением $\mathbf{P}(X_n = 0) = 1 - \dfrac{1}{n^2}$, $\mathbf{P}(X_n = 1) = \dfrac{1}{n^2}$.

4. Пусть $\{\xi_n,\ n = 1, 2, \ldots\}$ — последовательность независимых случайных величин, причем ξ_n имеет равномерное распределение на отрезке $[-\sqrt[4]{n}; \sqrt[4]{n}]$. Доказать, что $X_n \xrightarrow{\text{p}} 0$, $n \to \infty$, если $X_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}\xi_k$.

Указание. Вычислить дисперсию $\mathbf{D}\{X_n\}$.

5. Случайные величины $\{\xi_k,\ k = 1, 2, \ldots\}$ независимы, причем $\xi_k \sim \mathcal{N}(a_k; D)$, $D < \infty$, $k = 1, 2, \ldots$. Доказать, что $X_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}(\xi_k - a_k)^2 \xrightarrow{\text{с.к.}} D$, $n \to \infty$.

Указание. Учесть, что $\mathbf{M}\{(\xi - m)^4\} = 3D^2$, если $\xi \sim \mathcal{N}(m; D)$.

6. Доказать, что $X_n \xrightarrow{\text{п.н.}} X$, $n \to \infty$, если числовой ряд $\displaystyle\sum_{k=1}^{\infty}\mathbf{P}(|X_n - X| > \varepsilon)$ сходится для любого $\varepsilon > 0$.

Указание. Воспользоваться теоремой Бореля-Кантелли.

§ 3. Центральная предельная теорема

3.1. Теоретические положения. Пусть $\{X_n,\ n = 1, 2, \ldots\}$ — последовательность случайных величин. Пусть также B — любое борелевское множество действительной оси $\mathbb{R}^1$, а ∂B — его граница.

Определение 3.1. Последовательность $\{X_n,\ n = 1, 2, \ldots\}$ *сходится по распределению* к случайной величине X, если для любого B такого, что $\mathbf{P}(X \in \partial B) = 0$, выполнено

$$\mathbf{P}(X_n \in B) \to \mathbf{P}(X \in B), \quad n \to \infty. \tag{3.1}$$

Сходимость по распределению, которую также называют *слабой сходимостью*, будем обозначать

$$X_n \xrightarrow{\text{d}} X, \quad n \to \infty. \tag{3.2}$$

Пусть $F_n(x)$ и $F_X(x)$ — функции распределения, соответственно, X_n и X.

Теорема 3.1. *Следующие три утверждения эквивалентны:*

1) $X_n \xrightarrow{\mathrm{d}} X$, $n \to \infty$;
2) $F_n(x) \to F_X(x)$, $n \to \infty$ *для любой точки непрерывности функции* $F_X(x)$;
3) $\mathbf{M}\{g(X_n)\} \to \mathbf{M}\{g(X)\}$, $n \to \infty$ *для любой непрерывной и ограниченной на* $\mathbb{R}^1$ *функции* $g(x)$.

Теорема 3.2. *Если* $X_n \xrightarrow{\mathrm{p}} X$, $n \to \infty$, *то* $X_n \xrightarrow{\mathrm{d}} X$, $n \to \infty$.

Таким образом, слабая сходимость случайной последовательности следует из сходимости по вероятности, а, следовательно, из сходимости почти наверное и сходимости в среднем квадратическом.

В одном важном частном случае сходимость по распределению и сходимость по вероятности эквивалентны: если $X_n \xrightarrow{\mathrm{d}} a$, $n \to \infty$, где $a = \mathrm{const}$, то $X_n \xrightarrow{\mathrm{p}} a$, $n \to \infty$.

Для установления факта слабой сходимости обычно используется следующее утверждение.

Теорема 3.3. *Пусть* $\Psi_n(\lambda)$ *и* $\Psi(\lambda)$ — *характеристические функции, соответственно,* X_n *и* X. *Пусть также для любого* $\lambda \in \mathbb{R}^1$

$$\Psi_n(\lambda) \to \Psi(\lambda), \quad n \to \infty.$$

Тогда $X_n \xrightarrow{\mathrm{d}} X$, $n \to \infty$.

Особое значение в статистике имеет случай, когда предельная случайная величина X имеет нормальное распределение.

Определение 3.2. Последовательность $\{X_n, n = 1, 2, \dots\}$ называется *асимптотически нормальной* с параметрами $(m; \sigma^2)$, если

$$X_n \xrightarrow{\mathrm{d}} X, \quad n \to \infty, \text{ где } X \sim \mathcal{N}(m; \sigma^2). \tag{3.3}$$

Из определения 3.2 и теоремы 3.1 следует, что для любого $x \in \mathbb{R}^1$

$$F_n(x) \to \varPhi\left(\frac{x - m}{\sigma}\right), \quad n \to \infty. \tag{3.4}$$

Практическое значение (3.4) состоит в том, что СВ X_n можно считать нормально распределенной с параметрами $(m; \sigma^2)$, если n достаточно велико.

Пусть теперь $\{\xi_k, k = 1, 2, \dots\}$ — последовательность независимых случайных величин с параметрами

$$\mathbf{M}\{\xi_k\} = a; \quad \mathbf{D}\{\xi_k\} = \sigma^2 > 0.$$

Рассмотрим случайную *последовательность сумм* этих величин:

$$X_n = \sum_{k=1}^{n} \xi_k, \quad n = 1, 2, \dots \tag{3.5}$$

Очевидно, $\mathbf{M}\{X_n\} = na$, $\mathbf{D}\{X_n\} = n\sigma^2$.

Введем *стандартизованную сумму*

$$\widetilde{X}_n = \frac{X_n - na}{\sigma\sqrt{n}}, \quad n = 1, 2, \dots \tag{3.6}$$

Нетрудно проверить, что $\mathbf{M}\left\{\widetilde{X}_n\right\} = 0$, $\mathbf{D}\left\{\widetilde{X}_n\right\} = 1$.

Следующее утверждение, называемое *центральной предельной теоремой* (ЦПТ), имеет особое значение для математической статистики.

Теорема 3.4 (ЦПТ для одинаково распределенных слагаемых). *Пусть случайные величины* $\{\xi_k, k = 1, 2, \dots\}$ *одинаково распределены, тогда последовательность* $\{\widetilde{X}_n, n = 1, 2, \dots\}$ *асимптотически нормальна с параметрами* $(0; 1)$.

Следствие 3.1. *Для любых чисел* $a \leqslant b$ *выполнено*

$$\mathbf{P}(a \leqslant \widetilde{X}_n \leqslant b) \to \varPhi(b) - \varPhi(a), \quad n \to \infty. \tag{3.7}$$

Следствие 3.2 (Интегральная теорема Муавра-Лапласа). *Пусть* X_n — *число успехов в серии из* n *испытаний Бернулли, а* p — *вероятность успеха в одном испытании. Тогда при* $n \to \infty$

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0; 1). \tag{3.8}$$

Если слагаемые $\{\xi_k, k = 1, 2, \dots\}$ распределены не одинаково, то утверждение, аналогичное ЦПТ, останется в силе при некоторых дополнительных ограничениях.

Пусть $\mathbf{M}\{\xi_k\} = a_k$, $\mathbf{D}\{\xi_k\} = \sigma_k^2$, $\mathbf{M}\{|\xi_k - a_k|^3\} = c_k^3$. Обозначим $A_n = \mathbf{M}\{X_n\} = \sum_{k=1}^{n} a_k$, $D_n^2 = \mathbf{D}\{X_n\} = \sum_{k=1}^{n} \sigma_k^2$, $C_n^3 = \sum_{k=1}^{n} c_k^3$.

Теорема 3.5 (Ляпунов). *Пусть* A_n, D_n, C_n *конечны при всех* $n \geqslant 1$, *причем* $\frac{C_n}{D_n} \to 0$, $n \to \infty$. *Тогда последовательность* $\{\widetilde{X}_n, n = 1, 2, \dots\}$, *где* $\widetilde{X}_n = \dfrac{X_n - A_n}{D_n}$, *асимптотически нормальна с параметрами* $(0; 1)$.

3.2. Примеры.

Пример 3.1. Доказать интегральную теорему Муавра-Лапласа (следствие 3.2).

Решение. Пусть X_n — число "успехов" в серии из n испытаний Бернулли с вероятностью "успеха" p. Тогда $X_n = \sum_{k=1}^{n} \xi_k$, где ξ_k — результат k-го испытания, т.е. случайная величина с распределением $\mathbf{P}(\xi_k = 1) = p$, $\mathbf{P}(\xi_k = 0) = 0$. По условию случайные величины $\{\xi_k\}$ независимы в совокупности, причем $\mathbf{M}\{\xi_k\} = p$, $\mathbf{D}\{\xi_k\} = p\,(1-p)$. Поэтому $\mathbf{M}\{X_n\} = \sum_{k=1}^{n} \mathbf{M}\{\xi_k\} = np$, а $\mathbf{D}\{X_n\} = \sum_{k=1}^{n} \mathbf{D}\{\xi_k\} = np\,(1-p)$. Поэтому последовательность стандартизованных сумм $\widetilde{X}_n = \dfrac{X_n - np}{\sqrt{np\,(1-p)}}$ асимптотически нормальна по теореме 3.4. ∎

Пример 3.2. Симметричная монета подбрасывается 10000 раз. Используя ЦПТ, оценить вероятность того, что "герб" выпадет более 6000 раз. Указать реальные пределы изменения числа выпавших "гербов".

Решение. Из условия следует, что рассматривается схема из $n = 10000$ испытаний Бернулли с $p = 0{,}5$. Тогда, если X_n — число выпавших "гербов", то $\mathbf{M}\{X_n\} = np = 5000$, $\mathbf{D}\{X_n\} = np\,(1-p) = 2500$. В силу ЦПТ с учетом $n \gg 1$ полагаем, что $\widetilde{X}_n = \dfrac{X_n - np}{\sqrt{np\,(1-p)}} = \dfrac{X_n - 5000}{50} \sim \mathcal{N}(0;1)$.

Отсюда $\mathbf{P}(X_n > 6000) = \mathbf{P}\left(\dfrac{X_n - 5000}{50} > 20\right) = \mathbf{P}(\widetilde{X}_n > 20) = 1 - \varPhi(20) \cong 0$. Таким образом, событие $\{X_n > 6000\}$ представляется совершенно невероятным.

В силу приведенных выше рассуждений можно считать, что $X_n \sim \mathcal{N}(5000; 50^2)$. Известно, что если $X \sim \mathcal{N}(m_X; \sigma_X^2)$, то $\mathbf{P}(m_X - 3\sigma_X \leqslant X \leqslant m_X + 3\sigma_X) = \varPhi(3) - \varPhi(-3) \cong 0{,}997$. Поэтому $\mathbf{P}(5000 - 3\cdot 50 \leqslant X_n \leqslant 5000 + 3\cdot 50) = \mathbf{P}(4850 \leqslant X_n \leqslant 5150) = \varPhi(3) - \varPhi(-3) \cong 0{,}997$. Таким образом, с вероятностью, близкой к 1, число выпавших "гербов" X_n будет лежать в промежутке от 4850 до 5150, т.е. отличаться от ожидаемого числа 5000 не более чем на 150. ∎

Пример 3.3. Пусть $P_n^* = \dfrac{X_n}{n}$ — частота появления события A в схеме из n испытаний Бернулли, где X_n — количество появлений события A, а $p = \mathbf{P}(A)$. Доказать, что последовательность $Y_n = \sqrt{n}\,(P_n^* - p)$ асимптотически нормальна с параметрами $(0; p\,(1-p))$.

Решение. Так как в примере 3.1 показано, что $\mathbf{M}\{X_n\} = np$, $\mathbf{D}\{X_n\} = np\,(1-p)$, то $\mathbf{M}\{P_n^*\} = p$, $\mathbf{D}\{P_n^*\} = \dfrac{p\,(1-p)}{n}$. Поэтому

$$\begin{aligned} F_n(x) &= \mathbf{P}(Y_n \leqslant x) = \mathbf{P}\left(\frac{X_n}{n} - p \leqslant \frac{x}{\sqrt{n}}\right) = \\ &= \mathbf{P}\left(\frac{X_n - np}{\sqrt{np\,(1-p)}} \leqslant \frac{x}{\sqrt{p\,(1-p)}}\right) = \\ &= \mathbf{P}\left(\widetilde{X}_n \leqslant \frac{x}{\sqrt{p\,(1-p)}}\right) \to \varPhi\left(\frac{x}{\sqrt{p\,(1-p)}}\right), \quad n \to \infty. \end{aligned}$$

С учетом утверждения теоремы 3.1 и определения 3.2 заключаем, что $Y_n \xrightarrow{\mathrm{d}} Y$, $n \to \infty$, где $Y \sim \mathcal{N}(0; p\,(1-p))$, т.е. $\{Y_n\}$ асимптотически нормальна с параметрами $(0; p\,(1-p))$.

Заметим, что при $n \gg 1$ полученный результат позволяет считать, что $P_n^* \sim \mathcal{N}\left(p; \frac{p\,(1-p)}{n}\right)$. ∎

Пример 3.4. Пусть $X_n = \sum_{k=1}^{n} \xi_k$, где $\xi_k \sim R[-k; k]$, а случайные величины $\{\xi_k,\ k = 1, 2, \ldots\}$ независимы в совокупности. Доказать, что последовательность $\widetilde{X}_n = \frac{X_n}{D_n}$, где $D_n^2 = \mathbf{D}\{X_n\}$ асимптотически нормальна с параметрами $(0;1)$.

Решение. По условию $a_k = \mathbf{M}\{\xi_k\} = \dfrac{k + (-k)}{2} = 0$; $\sigma_k^2 = \mathbf{D}\{\xi_k\} = \dfrac{(k - (-k))^2}{12} = \dfrac{k^2}{3}$. Вычислим $\mathbf{M}\{|\xi_k|^3\}$: $c_k^3 = \mathbf{M}\{|\xi_k|^3\} = \displaystyle\int_{-k}^{k} |x|^3 \frac{1}{2k} dx = \frac{1}{k}\int_0^k x^3 dx = \frac{k^3}{4}$. Таким образом,

$$A_n = \mathbf{M}\{X_n\} = 0; \quad D_n^2 = \sum_{k=1}^{n} \sigma_k^2 = \frac{1}{3}\sum_{k=1}^{n} k^2 = \frac{n\,(n+1)\,(2n+1)}{18};$$

$$C_n^3 = \sum_{k=1}^{n} c_k^3 = \frac{1}{4}\sum_{k=1}^{n} k^3 = \frac{n^2 (n+1)^2}{16}.$$

Из полученных выражений следует, что

$$\frac{C_n}{D_n} = \alpha \frac{\sqrt[3]{n^4 + 2n^3 + n^2}}{\sqrt[2]{2n^3 + 3n^2 + n}} = O\left(n^{-1/6}\right) \to 0, \quad n \to \infty.$$

Таким образом, условия теоремы 3.5 Ляпунова выполнены. Поэтому $\widetilde{X}_n = \dfrac{X_n - A_n}{D_n} = 3\sqrt{2}\dfrac{X_n}{\sqrt{2n^3+3n^2+n}} \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0;1)$, $n \to \infty$, что и требовалось доказать. ■

3.3. Задачи для самостоятельного решения.

1. Вероятность рождения мальчика равна 0,52, девочки — 0,48. Какова вероятность того, что среди 10000 новорожденных детей девочек будет не меньше, чем мальчиков?

Ответ. $3 \cdot 10^{-5}$.

2. Длина шага человека имеет равномерное распределение на промежутке [0,9; 1,1] (в метрах). Какова вероятность того, что, сделав 100 шагов, он пройдет не менее 105 метров? В каких пределах с вероятностью 0,95 может лежать пройденный путь L?

Ответ. $\mathbf{P}(L > 105) \cong 0$; $\mathbf{P}(98{,}86 \leqslant L \leqslant 101{,}14) \cong 0{,}95$.

3. В результате 100 подбрасываний монеты "герб" выпал 70 раз. Насколько правдоподобным является предположение о симметричности монеты?

Указание. Вычислить $\mathbf{P}(X_n \geqslant 70)$ в предположении, что $p = 0{,}5$.

Ответ. Предположение симметричности неправдоподобно.

4. Случайная величина X_n имеет распределение хи-квадрат с n степенями свободы, т.е. $X_n = \sum\limits_{k=1}^{n} \xi_k^2$, где $\xi_k \sim \mathcal{N}(0;1)$, $k = 1,\ldots,n$ и независимы в совокупности. Доказать, что $X_n \sim \mathcal{N}(n;2n)$, если $n \gg 1$.

Указание. Показать, что $\left\{\dfrac{X_n - n}{\sqrt{2n}}\right\}$ асимптотически нормальна.

5. Производственный процесс состоит из 100 независимых операций, выполняемых одна за другой. Длительность каждой операции τ_k (сек) имеет экспоненциальное распределение. Известно, что $\mathbf{P}(\tau_k \leqslant 1) = 1 - \dfrac{1}{\sqrt{e}}$. Какова вероятность того, что длительность T процесса превысит одну минуту?

Ответ. $1 - \Phi(2) \cong 0{,}023$.

6. Случайная величина X_n имеет распределение Пуассона с параметром n ($X_k \sim \Pi(n)$). Доказать, что последовательность $Y_n = \dfrac{X_n - n}{\sqrt{n}}$ асимптотически нормальна.

Указание. Показать, что $X_n = \sum\limits_{k=1}^{n} \xi_k$, где $\xi_k \sim \Pi(1)$.

7. Случайные величины ξ_k независимы в совокупности и распределены по закону $R[0;1]$. Вычислить $\lim\limits_{n\to\infty} \mathbf{P}\left(\prod\limits_{k=1}^{n} \xi_k \leqslant e^{-n}\right)$.

Ответ. 0,5.

8. Необходимо сложить миллион чисел, округленных до пятого десятичного знака. Считая, что ошибки округления всех чисел независимы в совокупности и распределены равномерно на соответствующем интервале, найти пределы, в которых суммарная ошибка округления δ лежит с вероятностью 0,95.

Ответ. $|\delta| \leqslant 0{,}00566$.

§ 4. Закон больших чисел

4.1. Теоретические положения. Во всех приводимых ниже утверждениях (теоремы 4.1-4.5) предполагается, что $\{X_k, k = 1, 2, \ldots\}$ — последовательность *независимых СВ.*

Определение 4.1. *Выборочным средним* называется СВ

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \quad n = 1, 2, \ldots .$$

Законом больших чисел называется совокупность утверждений о поведении последовательности $\{\overline{X}_n, n = 1, 2, \ldots\}$ выборочных средних при $n \to \infty$.

Будем далее обозначать $\mathbf{M}\{X_k\} = a_k$, $\mathbf{D}\{X_k\} = D_k = \sigma_k^2$.

Теорема 4.1. *Если $\{X_k, k = 1, 2, \ldots\}$ одинаково распределены и $\mathbf{M}\{X_k\} = a$ конечно, то $\overline{X}_n \xrightarrow{\text{п.н.}} a$, $n \to \infty$.*

Теорема 4.1 называется *"Усиленный закон больших чисел А.Н. Колмогорова".*

Теорема 4.2. *Если $\mathbf{M}\{X_k\} = a$, $\mathbf{D}\{X_k\} = D_k < \infty$, причем*

$$\sum_{k=1}^{\infty} \frac{D_k}{k^2} < \infty, \tag{4.1}$$

то $\overline{X}_n \xrightarrow{\text{п.н.}} a$, $n \to \infty$ и $\overline{X}_n \xrightarrow{\text{с.к.}} a$, $n \to \infty$.

Следствие 4.1. *Если $a_k = a$, $D_k \leqslant \overline{D} < \infty$ $\forall k = 1, 2, \ldots$, то утверждение теоремы 4.2 справедливо.*

Пусть теперь $\{a_k\}$ не одинаковы. Обозначим $\overline{a}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} a_k$.

Следствие 4.2. *Если выполнено условие (4.1), то*

$$\overline{X}_n - \overline{a}_n \xrightarrow{\text{п.н.}} 0,\ n \to \infty; \quad \overline{X}_n - \overline{a}_n \xrightarrow{\text{с.к.}} 0,\ n \to \infty.$$

Следующее утверждение показывает, как будет вести себя последовательность выборочных средних, если СВ $\{X_k, k = 1, 2, \ldots\}$ не имеют конечного среднего (например, X_k имеет распределение Коши).

Т е о р е м а 4.3. *Пусть* $\{X_k, k = 1, 2, \ldots\}$ *одинаково распределены, но* $\mathbf{M}\{X_k\}$ *не существует ни при каких* $k = 1, 2, \ldots$. *Тогда последовательность* $\{\overline{X}_n, n = 1, 2, \ldots\}$ *с вероятностью 1 неограничена, т.е.* $\mathbf{P}\left(\sup\limits_{n\geqslant 1}|\overline{X}_n| > C\right) = 1$ *для любого* $C > 0$.

Весьма часто последовательность выборочных средних асимптотически нормальна (см. определение 3.2).

Т е о р е м а 4.4. *Пусть* $\{X_k, k = 1, 2, \ldots\}$ *— последовательность независимых одинаково распределенных СВ, причем* $\mathbf{M}\{X_k^2\} < \infty$, $k = 1, 2, \ldots$. *Тогда*

$$\sqrt{n}(\overline{X}_n - a) \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0; \sigma^2), \quad n \to \infty, \tag{4.2}$$

где $\sigma = \sqrt{\mathbf{D}\{X_k\}}$, $a = \mathbf{M}\{X_k\}$, $k = 1, 2, \ldots$.

Утверждение (4.2) позволяет пользоваться при $n \gg 1$ приближенным соотношением:

$$\overline{X}_n \sim \mathcal{N}\left(a; \frac{\sigma^2}{n}\right). \tag{4.3}$$

Для случая, когда СВ $\{X_k, k = 1, 2, \ldots\}$ распределены неодинаково, аналог теоремы 4.4 можно получить, используя теорему 3.5 Ляпунова.

Т е о р е м а 4.5. *Пусть выполнены условия теоремы 3.5, тогда*

$$\frac{n}{D_n}\left(\overline{X}_n - \overline{a}_n\right) \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0; 1), \quad n \to \infty.$$

Аналогично (4.3), в данном случае можно считать, что

$$\overline{X}_n \sim \mathcal{N}\left(\overline{a}_n; \frac{D_n^2}{n^2}\right). \tag{4.4}$$

4.2. Примеры.

П р и м е р 4.1. Пусть $\{\xi_k, k = 1, 2, \ldots, n\}$ — независимые СВ, причем $\xi_k \sim R[0; 1]$, $k = 1, 2, \ldots$. Показать, что $Y_n = \dfrac{1}{n}\sum\limits_{k=1}^{n}(\xi_k)^2$ при $n \to \infty$ сходится почти наверное к $\frac{1}{3}$.

Р е ш е н и е. По условию $\mathbf{M}\{\xi_k\} = \frac{1}{2}$, $\mathbf{D}\{\xi_k\} = \frac{1}{12}$. Пусть $X_k = \xi_k^2$, тогда $\{X_k, k = 1, 2, \ldots, n\}$ независимы и одинаково распределены, причем $\mathbf{M}\{X_k\} = \mathbf{M}\{(\xi_k)^2\} = \mathbf{D}\{\xi_k\} + (\mathbf{M}\{\xi_k\})^2 = \dfrac{1}{12} + \dfrac{1}{4} = \dfrac{1}{3}$. Таким образом, $Y_n = \overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k \xrightarrow{\text{п.н.}} \mathbf{M}\{X_1\} = \dfrac{1}{3}$ при $n \to \infty$ по теореме 4.1 А.Н. Колмогорова. ■

П р и м е р 4.2. Доказать, что в условиях теоремы 4.2 $\overline{X}_n \xrightarrow{\text{с.к.}} a$, $n \to \infty$.

Р е ш е н и е. $\mathbf{M}\{\overline{X}_n\} = \mathbf{M}\left\{\dfrac{1}{n}\sum\limits_{k=1}^{n} X_k\right\} = \dfrac{1}{n}\sum\limits_{k=1}^{n} a = \dfrac{na}{n} = a$. $\mathbf{D}\{\overline{X}_n\} = \dfrac{1}{n^2}\sum\limits_{k=1}^{n}\mathbf{D}\{X_k\} = \dfrac{1}{n^2}\sum\limits_{k=1}^{n} D_k \to 0$, $n \to \infty$. Действительно, известная лемма Кронекера утверждает, что если числовые последовательности $\{v_k\}$ и $\{u_k\}$ таковы, что $\sum\limits_{k=1}^{\infty} v_k < \infty$, $0 < u_n \uparrow \infty$, $n \to \infty$, то $\dfrac{1}{u_n}\sum\limits_{k=1}^{n} v_k u_k \to 0$, $n \to \infty$. Теперь положим $v_k = \dfrac{D_k}{k^2}$, $u_k = k^2$ и заметим, что $\sum\limits_{k=1}^{\infty} v_k = \sum\limits_{k=1}^{\infty}\dfrac{D_k}{k^2} < \infty$ по условию. Тогда $\dfrac{1}{u_n}\sum\limits_{k=1}^{n} v_k u_k = \dfrac{1}{n^2}\sum\limits_{k=1}^{n} D_k \to 0$, $n \to \infty$.

Таким образом, $\mathbf{M}\{|\overline{X}_n - a|^2\} = \mathbf{D}\{\overline{X}_n\} \to 0$, $n \to \infty$, т.е. $\overline{X}_n \xrightarrow{\text{с.к.}} a$, $n \to \infty$. ■

П р и м е р 4.3. Пусть $P_n^*(A) = \dfrac{n(A)}{n}$ — частота появления случайного события A в серии из n независимых одинаковых опытов, а $n(A)$ — количество опытов, в которых A произошло. Доказать, что $P_n^*(A) \xrightarrow{\text{п.н.}} \mathbf{P}(A)$, $n \to \infty$ и $P_n^*(A) \xrightarrow{\text{с.к.}} \mathbf{P}(A)$, $n \to \infty$, где $\mathbf{P}(A)$ — вероятность события A.

Р е ш е н и е. По условию $n(A) = \sum\limits_{k=1}^{n} X_k$, где $\{X_k, k = 1, 2, \ldots\}$ — независимые СВ с распределением Бернулли, причем $\mathbf{P}(X_k = 1) = \mathbf{P}(A)$. Поэтому $\mathbf{M}\{X_k\} = p$, $\mathbf{D}\{X_k\} = pq$, где $p = \mathbf{P}(A)$, $q = 1 - p$. Таким образом, $P_n^*(A) = \overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k \xrightarrow{\text{п.н.}} p$, $n \to \infty$ по теореме 4.1, так как $p = \mathbf{M}\{X_k\} \in [0; 1]$. Очевидно, что $\mathbf{M}\{P_n^*(A)\} = \dfrac{1}{n}\sum\limits_{k=1}^{n}\mathbf{M}\{X_k\} = p$, а $\mathbf{D}\{P^*_n(A)\} = \frac{1}{n^2}\sum\limits_{k=1}^{n}\mathbf{D}\{X_k\} = \frac{npq}{n^2} = \frac{pq}{n}$. Поэтому $\mathbf{M}\{|P^*_n(A)-p|^2\} = \mathbf{D}\{P^*_n(A)\} = \frac{pq}{n} \to 0$, $n\to\infty$. Последнее означает, что $P^*_n(A) \xrightarrow{\text{с.к.}} p$, $n\to\infty$. ■

Пример 4.4. В условиях примера 4.3 оценить минимальное число опытов n, при котором $\mathbf{P}\left(|P^*_n(A)-\mathbf{P}(A)| \leqslant 0{,}1\right) \geqslant 0{,}95$.

Решение. Из теоремы 4.4 и решения примера 4.3 следует, что $\sqrt{n}\,(P^*_n(A)-p) \xrightarrow{\text{d}} X \sim \mathcal{N}(0;pq)$, $n\to\infty$. Поэтому при $n\gg 1$ можно считать, что $\xi = P^*_n(A)-p \sim \mathcal{N}\left(0;\frac{pq}{n}\right)$. Отсюда $\mathbf{P}(|\xi|\leqslant 0{,}1) = \mathbf{P}(-0{,}1\leqslant\xi\leqslant 0{,}1) = \varPhi\left(\frac{0{,}1\sqrt{n}}{\sqrt{pq}}\right) - \varPhi\left(-\frac{0{,}1\sqrt{n}}{\sqrt{pq}}\right) = 2\varPhi_0\left(\frac{0{,}1\sqrt{n}}{\sqrt{pq}}\right)$, где

$\varPhi_0(z) = \frac{1}{2\pi}\int\limits_0^z e^{-t^2/2}dt$.

Решая неравенство $2\varPhi_0\left(\frac{0{,}1\sqrt{n}}{\sqrt{pq}}\right) \geqslant 0{,}95$, по таблице 13.1 находим: $\frac{0{,}1\sqrt{n}}{\sqrt{pq}} \geqslant 1{,}96$. Отсюда следует: $n\geqslant (19{,}6)^2\cdot pq = 384{,}2\cdot pq$. Таким образом, $n = [384{,}2\cdot pq]+1$.

Если p неизвестна, то можно получить гарантированную оценку $n_{\text{гар}}$ необходимого числа опытов, справедливую для любой вероятности p. Действительно, $\max\limits_{p\in[0;1]} p(1-p) = 0{,}25$ и достигается при $p=q=0{,}5$, поэтому $pq\leqslant 0{,}25$ при любом $p\in[0;1]$. Отсюда $n_{\text{гар}} \geqslant 384{,}2\cdot 0{,}25 = 96{,}05$. Итак, $n_{\text{гар}} = 97$. ■

Пример 4.5. Пусть СВ $X_k \sim E(k)$, $k=1,2,\ldots$ и независимы. Доказать, что $\overline{\xi}_n = \frac{1}{n}\sum\limits_{k=1}^{n}\xi_k$, где $\xi_k = X_k - \frac{1}{k}$ сходится в среднем квадратическом к нулю.

Решение. По условию $\mathbf{M}\{X_k\} = \frac{1}{k}$, $\mathbf{D}\{X_k\} = \frac{1}{k^2}$. Поэтому $\mathbf{M}\{\xi_k\} = \mathbf{M}\left\{X_k - \frac{1}{k}\right\} = 0$, $\mathbf{D}\{\xi_k\} = \mathbf{D}\{X_k\} = \frac{1}{k^2}$, $k=1,2,\ldots$. Отсюда заключаем, что

$$\mathbf{M}\{\overline{\xi}_n\} = \frac{1}{n}\sum_{k=1}^{n}\mathbf{M}\{\xi_k\} = 0;$$

$$\mathbf{D}\{\overline{\xi}_n\} = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{D}\{\xi_k\} = \frac{1}{n^2}\sum_{k=1}^{n}\frac{1}{k^2} \leqslant \frac{C}{n^2} \to 0, \quad n\to\infty,$$

где $C = \sum\limits_{k=1}^{\infty}\frac{1}{k^2} < \infty$. Итак, $\mathbf{M}\{\overline{\xi}^2_n\} = \mathbf{D}\{\overline{\xi}_n\} \to 0$, $n\to\infty$, т.е. $\overline{\xi}_n \xrightarrow{\text{с.к.}} 0$, $n\to\infty$. ■

4.3. Задачи для самостоятельного решения.

1. Пусть X, $\{X_k,\ k = 1, 2, \ldots\}$ — независимые одинаково распределенные СВ, а $\varphi(x)$ — непрерывная функция. Доказать, что если $\mathbf{M}\{\varphi(X)\} = m_\varphi$ конечно, то

$$\frac{1}{n}\sum_{k=1}^{n}\varphi(X_k) \xrightarrow{\text{п.н.}} m_\varphi, \quad n \to \infty.$$

Указание. Воспользоваться теоремой 4.1.

2. СВ $\{X_k,\ k = 1, 2, \ldots\}$ одинаково распределены с функцией распределения $F(x)$. Пусть $\widehat{F}_n(x) = \dfrac{n_x}{n}$, где n_x — количество опытов, в которых произошло событие $\{X_k \leqslant x\}$, $k \leqslant n$. Доказать, что $\widehat{F}_n(x) \xrightarrow{\text{п.н.}} F(x)$, $n \to \infty$.

Указание. Смотри пример 4.3.

3. Пусть $\{X_k,\ k = 1, 2, \ldots\}$ — независимые СВ с распределением $X_k \sim \Pi(\sqrt{k})$. Доказать, что последовательность $\xi_n = \dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - \sqrt{k})$ сходится к нулю почти наверное и в среднем квадратическом при $n \to \infty$.

Указание. Смотри теорему 4.2.

4. Пусть $\{X_k \sim R[0;1],\ k = 1, 2, \ldots\}$ — независимые СВ. В каких пределах с вероятностью 0,997 будет лежать $\overline{X}_n$, если $n = 300$?

Ответ. $[0{,}45; 0{,}55]$.

5. Независимые СВ $\{X_k,\ k = 1, 2, \ldots\}$ распределены одинаково с плотностью вероятности $p(x) = \begin{cases} 2x\exp\{-x^2\}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$

Показать, что $\dfrac{1}{n}\sum\limits_{k=1}^{n} X_k^2 \xrightarrow{\text{п.н.}} 1$, $n \to \infty$.

Указание. Вычислить $\mathbf{M}\{X_k^2\}$.

6. Пусть $\{X_k,\ k = 1, 2, \ldots\}$ — независимые стандартные гауссовские СВ. Доказать, что $\dfrac{1}{n}\sum\limits_{k=1}^{n}|X_k| \xrightarrow{\text{п.н.}} \sqrt{\dfrac{2}{\pi}}$, $n \to \infty$.

Указание. Найти $\mathbf{M}\{|X_k|\}$.

§ 5. Выборка и ее основные характеристики

5.1. Теоретические положения. Пусть X — произвольная случайная величина с функцией распределения $F(x) = \mathbf{P}(X \leqslant x)$, $x \in \mathbb{R}^1$.

Определение 5.1. Совокупность $\{X_k,\ k = 1, \ldots, n\}$ независимых случайных величин, имеющих одинаковые функции распределения

$F_{X_k}(x) = F(x)$, называется *однородной выборкой объема* n, соответствующей функции распределения $F(x)$.

СВ X_k $(k = 1, \ldots, n)$ называется k-м *элементом выборки.*

Из определения 5.1 следует, что выборку можно рассматривать как случайный вектор $Z_n = \{X_1, \ldots, X_n\}^\top$ с независимыми компонентами. Кроме того, СВ $\{X_k,\ k = 1, \ldots, n\}$ — независимые вероятностные "копии" СВ X, поэтому мы также будем говорить, что *выборка* Z_n *порождена СВ* X с распределением $F(x)$.

Определение 5.2. Выборка $\{X_k,\ k = 1, \ldots, n\}$ называется *гауссовской*, если Z_n — n-мерный гауссовский вектор.

Определение 5.3. Выборка Z_n называется *неоднородной*, если законы распределения $F_{X_k}(x)$ ее элементов неодинаковы.

Далее полагается, что выборка Z_n — однородная, если специально не указано обратное.

Из приведенных определений следует, что выборка является математической моделью последовательности одинаковых опытов со случайными исходами, проводимых в неизменных условиях, причем результаты опытов статистически независимы.

Определение 5.4. *Реализацией выборки* Z_n называется неслучайный вектор $z_n = \{x_1, \ldots, x_n\}^\top$, компонентами которого являются реализации соответствующих элементов выборки.

Пусть $z_n = \{x_1, \ldots, x_n\}^\top$ — некоторая реализация выборки Z_n, а $z_{(n)} = \{x_{(1)}, \ldots, x_{(n)}\}^\top$ — вектор, компонентами которого являются упорядоченные по возрастанию числа $(x_1, \ldots, x_n)$, т.е. $x_{(1)} \leqslant x_{(2)} \leqslant \ldots \leqslant x_{(n)}$.

Определение 5.5. СВ $X_{(k)}$, реализацией которой для каждой z_n является число $x_{(k)}$, называется k*-й порядковой статистикой*, $k = 1, \ldots, n$. Случайный вектор $Z_{(n)} = \{X_{(1)}, \ldots, X_{(n)}\}^\top$ называется *вариационным рядом* выборки.

СВ $X_{(1)}$ и $X_{(n)}$ (т.е. крайние элементы вариационного ряда) называются *экстремальными статистиками.*

Определение 5.6. СВ $Y = \varphi(X_1, \ldots, X_n)$, где $\varphi(x_1, \ldots, x_n)$ — произвольная (борелевская) функция на $\mathbb{R}^n$, называется *статистикой.*

Обычно предполагается, что в случае, когда функция распределения $F(x)$, которой соответствует выборка, зависит от некоторого вектора θ неизвестных параметров, $\varphi(x_1, \ldots, x_n)$ не зависит от θ.

Рассмотрим некоторые важнейшие для приложений виды статистик.

Определение 5.7.

1) $\overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$ называется *выборочным средним.*

2) $\overline{S}^2_n = \dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - \overline{X}_n)^2$ называется *выборочной дисперсией.*

3) $\overline{\nu}_r(n) = \dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k)^r$, $r = 1, 2, \ldots$, называется *выборочным начальным моментом* r-го порядка.

4) $\overline{\mu}_r(n) = \dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - \overline{X}_n)^r$, $r = 1, 2, \ldots$, называется *выборочным центральным моментом* r-го порядка.

Заметим, что $\overline{X}_n = \overline{\nu}_1(n)$, а $\overline{S}^2_n = \overline{\mu}_2(n)$.

Пусть распределение $F(x)$ таково, что существуют следующие теоретические моменты любого элемента X_k выборки: $m_X = \mathbf{M}\{X_k\}$, $D_X = \mathbf{D}\{X_k\}$, $\nu_r = \mathbf{M}\{(X_k)^r\}$, $\mu_r = \mathbf{M}\{(X_k - m_X)^r\}$, $r = 2, 3, \ldots$. Тогда справедливо следующее утверждение.

Теорема 5.1. *При неограниченном увеличении объема выборки n выборочные моменты $\overline{\nu}_r(n)$ и $\overline{\mu}_r(n)$, $r = 1, 2, \ldots$ почти наверное сходятся к теоретическим моментам ν_r и μ_r соответственно.*

Следствие 5.1. *Если $m_X = \mathbf{M}\{X_k\}$ существует, то $\overline{X}_n \xrightarrow{\text{п.н.}} m_X$, $n \to \infty$. Если ν_2 существует, то $\overline{S}^2_n \xrightarrow{\text{п.н.}} D_X$, $n \to \infty$.*

При определенных дополнительных условиях выборочные моменты обладают свойством асимптотической нормальности.

Теорема 5.2. *Пусть для некоторого $r \geqslant 1$ существует и конечен момент ν_{2r}. Тогда $\sqrt{n}\,(\overline{\nu}_r(n) - \nu_r) \xrightarrow{\text{d}} \xi \sim \mathcal{N}(0; \nu_{2r} - \nu_r^2)$, $n \to \infty$.*

Следствие 5.2. *Если $D_X < \infty$, то*

$$\sqrt{n}\,(\overline{X}_n - m_X) \xrightarrow{\text{d}} \xi \sim \mathcal{N}(0; D_X), \quad n \to \infty.$$

Если $\nu_4 < \infty$, то $\sqrt{n}\,(\overline{\nu}_2(n) - \nu_2) \xrightarrow{\text{d}} \xi \sim \mathcal{N}(0; \nu_4 - D_X^2)$, $n \to \infty$.

Из приведенных выше утверждений следует, что при $n \gg 1$ выборочные моменты $\overline{\nu}_r(n)$ и $\overline{\mu}_r(n)$ практически не отличаются от своих теоретических значений ν_r и μ_r. Кроме того, можно считать, что $\overline{\nu}_r(n) \sim \mathcal{N}\left(\nu_r; \dfrac{\nu_{2r} - \nu_r^2}{n}\right)$, если $n \gg 1$.

Пусть выборка $\{X_k, k = 1, \ldots, n\}$ порождена СВ X с функцией распределения $F(x)$. Для любого $x \in \mathbb{R}^1$ введем событие $A_X = \{X \leqslant x\}$, тогда $\mathbf{P}(A_X) = F(x)$. Обозначим через $M_n(x)$ случайное число элементов выборки, не превосходящих x.

Определение 5.8. Случайная функция $\widehat{F}_n(x) = \dfrac{M_n(x)}{n}$, $x \in \mathbb{R}^1$, называется *выборочной (эмпирической) функцией распределения* СВ X.

При достаточно больших n функция $\widehat{F}_n(x)$ весьма точно аппроксимирует функцию распределения $F(x)$, которой соответствует выборка, о чем свидетельствуют следующие утверждения.

Теорема 5.3 (Гливенко-Кантелли). *$\widehat{F}_n(x)$ сходится к $F(x)$ почти наверное равномерно по x при $n \to \infty$, т.е.*

$$\sup_{x \in \mathbb{R}^1} |\widehat{F}_n(x) - F(x)| \xrightarrow{\text{п.н.}} 0, \quad n \to \infty.$$

Теорема 5.4. *При любом $x \in \mathbb{R}^1$ последовательность $\{\widehat{F}_n(x), n = 1, 2, \ldots\}$ асимптотически нормальна:*

$$\sqrt{n}\left(\widehat{F}_n(x) - F(x)\right) \xrightarrow{\text{d}} \xi \sim \mathcal{N}\left(0; F(x)(1 - F(x))\right), \quad n \to \infty.$$

Пусть двумерная выборка $\{(X_k, Y_k), k = 1, \ldots, n\}$ порождена случайным вектором $\xi = \{X, Y\}^\top$. Обозначим через $k_{XY} = \mathbf{M}\{(X - m_X)(Y - m_Y)\} = \mathbf{M}\{XY\} - m_X m_Y$ ковариацию случайных величин X и Y.

Определение 5.9. Статистика $\widehat{k}_{XY}(n) = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k Y_k - \overline{X}_n \overline{Y}_n$ называется *выборочной ковариацией* случайных величин X и Y.

Теорема 5.5. *Если СВ X и Y имеют конечные дисперсии, то*

1) $\mathbf{M}\left\{\widehat{k}_{XY}(n)\right\} = \dfrac{n-1}{n} k_{XY}$;

2) $\widehat{k}_{XY}(n) \xrightarrow{\text{п.н.}} k_{XY}$, $n \to \infty$;

3) *если $\mathbf{M}\{|X|^4 + |Y|^4\} < \infty$, то*

$$\sqrt{n}\left(\widehat{k}_{XY}(n) - k_{XY}\right) \xrightarrow{\text{d}} \eta \sim \mathcal{N}\left(0; \mu_{22} - k_{XY}^2\right), \quad n \to \infty,$$

где $\mu_{22} = \mathbf{M}\{(X - m_X)^2 (Y - m_Y)^2\}$.

5.2. Примеры.

Пример 5.1. Пусть выборка Z_n соответствует распределению $F(x)$. Доказать, что $X_{(1)}$ и $X_{(n)}$ имеют функции распределения, соответственно, $F_{(1)}(x) = 1 - (1 - F(x))^n$ и $F_{(n)}(x) = F^n(x)$.

Решение. По определению $X_{(n)} = \max(X_1, \ldots, X_n)$, поэтому $F_{(n)}(x) = \mathbf{P}(X_{(n)} \leqslant x) = \mathbf{P}\left(\{X_1 \leqslant x\} \cdot \{X_2 \leqslant x\} \cdot \ldots \cdot \{X_n \leqslant x\}\right) =$

$= \mathbf{P}\left(\prod\limits_{k=1}^{n}\{X_k \leqslant x\}\right)$. Так как элементы выборки статистически независимы и одинаково распределены, получаем

$$F_{(n)}(x) = \prod_{k=1}^{n} \mathbf{P}(X_k \leqslant x) = \prod_{k=1}^{n} F_{X_k}(x) = F^n(x).$$

Аналогично,

$$F_{(1)}(x) = 1 - \mathbf{P}(X_{(1)} > x) = 1 - \prod_{k=1}^{n} \mathbf{P}(X_k > x) =$$
$$= 1 - \prod_{k=1}^{n}\left(1 - F_{X_k}(x)\right) = 1 - (1 - F(x))^n.$$

■

Пример 5.2. Пусть выборка Z_n порождена СВ X с конечным моментом ν_r. Доказать, что выборочный начальный момент $\overline{\nu}_r(n)$ обладает по отношению к ν_r *свойством несмещенности*, т.е. $\mathbf{M}\{\overline{\nu}_r(n)\} = \nu_r$, и *свойством сильной состоятельности*, т.е. $\overline{\nu}_r(n) \xrightarrow{\text{п.н.}} \nu_r$, $n \to \infty$.

Решение. По условию $\mathbf{M}\{(X_k)^r\} = \mathbf{M}\{X^r\} = \nu_r$. Поэтому $\mathbf{M}\{\overline{\nu}_r(n)\} = \mathbf{M}\left\{\frac{1}{n}\sum\limits_{k=1}^{n}(X_k)^r\right\} = \frac{1}{n}\sum\limits_{k=1}^{n}\mathbf{M}\{(X_k)^r\} = \frac{1}{n}\sum\limits_{k=1}^{n}\nu_r = \nu_r$. Свойство несмещенности доказано.

Обозначим $\xi_k = (X_k)^r$, тогда величины $\{\xi_1, \ldots, \xi_n\}$ независимы, одинаково распределены и $\mathbf{M}\{\xi_k\} = \nu_r$. По усиленному закону больших чисел Колмогорова (теорема 4.1) $\overline{\nu}_r(n) = \frac{1}{n}\sum\limits_{k=1}^{n}\xi_k \xrightarrow{\text{п.н.}} \mathbf{M}\{\xi_1\} = \nu_r$, $n \to \infty$. Свойство сильной состоятельности доказано. ■

Пример 5.3. В условиях примера 5.2 для $r = 2$ показать, что выборочная дисперсия $\overline{S}_n^2$ обладает *свойством асимптотической несмещенности*, т.е. $\mathbf{M}\{\overline{S}_n^2\} \to D_X$, $n \to \infty$, и свойством сильной состоятельности.

Решение. По определению $\overline{S}_n^2 = \frac{1}{n}\sum\limits_{k=1}^{n}(X_k - \overline{X}_n)^2 = \frac{1}{n}\sum\limits_{k=1}^{n}(X_k)^2 - (\overline{X}_n)^2 = \overline{\nu}_2(n) - (\overline{\nu}_1(n))^2$. Из результата примера 5.2 следует, что $\overline{\nu}_2(n) \xrightarrow{\text{п.н.}} \nu_2$, $\overline{\nu}_1(n) \xrightarrow{\text{п.н.}} \nu_1$, $n \to \infty$. Тогда в силу свойства сходимости почти наверное (см. разд. 2.1) заключаем: $\overline{S}_n^2 = \overline{\nu}_2(n) - (\overline{\nu}_1(n))^2 \xrightarrow{\text{п.н.}} \nu_2 - \nu_1^2 = \mathbf{M}\{X^2\} - (\mathbf{M}\{X\})^2 = D_X$, $n \to \infty$. Свойство сильной состоятельности доказано.

Пусть теперь $\xi_k = (X_k - \overline{X}_n)^2$, а $m_X = \mathbf{M}\{X\}$. Тогда $\mathbf{M}\{\xi_k\} = \mathbf{M}\left\{\left((X_k - m_X) - (\overline{X}_n - m_X)\right)^2\right\} = \mathbf{M}\left\{\left(\overset{\circ}{X}_k - \left(\frac{1}{n}\sum\limits_{i=1}^{n}\overset{\circ}{X}_i\right)\right)^2\right\} = \mathbf{M}\left\{\overset{\circ}{X}{}_k^2\right\} - \frac{2}{n}\sum\limits_{i=1}^{n}\mathbf{M}\left\{\overset{\circ}{X}_k\overset{\circ}{X}_i\right\} + \frac{1}{n^2}\sum\limits_{i,j=1}^{n}\mathbf{M}\left\{\overset{\circ}{X}_i\overset{\circ}{X}_j\right\}$. С учетом независимости X_i и X_j при $i \neq j$ получаем $\mathbf{M}\{\xi_k\} = D_X - \frac{2}{n}D_X + \frac{1}{n}D_X = \frac{n-1}{n}D_X$. Поэтому $\mathbf{M}\{\overline{S}_n^2\} = \frac{1}{n}\sum\limits_{k=1}^{n}\mathbf{M}\{\xi_k\} = \frac{n-1}{n}D_X$. Таким образом, $\overline{S}_n^2$ не обладает свойством несмещенности по отношению к дисперсии D_X, так как $\mathbf{M}\{\overline{S}_n^2\} \neq D_X$. Однако, $\lim\limits_{n\to\infty}\mathbf{M}\{\overline{S}_n^2\} = \lim\limits_{n\to\infty}\frac{n-1}{n}D_X = D_X$, т.е. свойство асимптотической несмещенности имеет место. ■

Пример 5.4. Выборка $\{X_k,\ k = 1, \ldots, 175\}$ соответствует распределению $R[-1;1]$. Оценить вероятность того, что $|\overline{\nu}_3(175)| \leqslant \frac{1}{70}$.

Решение. По условию $X_k \sim R[-1;1]$, поэтому $\nu_3 = \mathbf{M}\{X_k^3\} = \int\limits_{-1}^{1} x^3\frac{1}{2}dx = 0$. Так как $n = 175 \gg 1$, то для искомой оценки вероятности можно воспользоваться теоремой 5.2, из которой для $r = 3$ с учетом $\nu_3 = 0$ следует, что $\overline{\nu}_3(175) \sim \mathcal{N}\left(0; \frac{\nu_6}{175}\right)$. Так как $\nu_6 = \frac{1}{2}\int\limits_{-1}^{1}x^6dx = \frac{1}{7}$, то $\overline{\nu}_3(175) \sim \mathcal{N}\left(0; \frac{1}{1225}\right)$. Отсюда $\mathbf{P}\left(|\nu_3(175)| \leqslant \frac{1}{70}\right) \cong \Phi\left(\frac{\sqrt{1225}}{70}\right) - \Phi\left(-\frac{\sqrt{1225}}{70}\right) = \Phi\left(\frac{1}{2}\right) - \Phi\left(-\frac{1}{2}\right) = 2\Phi_0\left(\frac{1}{2}\right) = 0{,}383$. ■

Пример 5.5. Выборка Z_n порождена СВ $X \sim R[0;1]$. Для любого $\varepsilon > 0$ оценить $\mathbf{P}\left(|\widehat{F}_n(x) - x| \leqslant \varepsilon\right)$ при каждом $x \in [0;1]$ и $n \gg 1$.

Решение. Так как $X \sim R[0;1]$, то $F(x) = x$, $x \in [0;1]$. Поэтому $\widehat{F}_n(x) - x = \widehat{F}_n(x) - F(x) \sim \mathcal{N}\left(0; \frac{F(x)(1-F(x))}{n}\right)$ по теореме 5.4. Итак, $\widehat{F}_n(x) - x \sim \mathcal{N}\left(0; \frac{x(1-x)}{n}\right)$ для $x \in [0;1]$ и $n \gg 1$. Отсюда $\mathbf{P}\left(|\widehat{F}_n(x) - x| \leqslant \varepsilon\right) \cong \Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{x(1-x)}}\right) - \Phi\left(-\frac{\varepsilon\sqrt{n}}{\sqrt{x(1-x)}}\right) =$

$= 2\Phi_0\left(\dfrac{\varepsilon\sqrt{n}}{\sqrt{x(1-x)}}\right)$. Так как $x(1-x) \leqslant \frac{1}{4}$, то наихудший результат будет при $x = \frac{1}{2}$. Например, если $\varepsilon = 0{,}1$, $n = 100$, то $\mathbf{P}\left(|\widehat{F}_n(\frac{1}{2}) - \frac{1}{2}| \leqslant 0{,}1\right) \cong 2\Phi_0(2) \cong 0{,}95$. Для сравнения, при $x = 0{,}1$ $\mathbf{P}\left(|\widehat{F}_n(0{,}1) - 0{,}1| \leqslant 0{,}1\right) \cong 2\Phi_0\left(\dfrac{0{,}1\sqrt{100}}{\sqrt{0{,}09}}\right) \cong 2\Phi_0(3{,}3) \cong 0{,}998$. ■

5.3. Задачи для самостоятельного решения.

1. Найти функцию распределения k-ой порядковой статистики $X_{(k)}$, $k = 1, \ldots, n$.

Ответ. $F_{(k)}(x) = \sum\limits_{m=k}^{n} C_n^m F^m(x)(1-F(x))^{n-m}$.

2. Выборка соответствует распределению $R[0;1]$. Вычислить $\mathbf{M}\left\{X_{(n)}\right\}$ и $\mathbf{D}\left\{X_{(n)}\right\}$.

Ответ. $\mathbf{M}\left\{X_{(n)}\right\} = \dfrac{n}{n+1}$; $\mathbf{D}\left\{X_{(n)}\right\} = \dfrac{n}{(n+1)^2(n+2)}$.

3. Выборка соответствует распределению $F(x)$ с конечным моментом ν_r. Доказать, что $\overline{\mu}_r(n) \xrightarrow{\text{п.н.}} \mu_r$, $n \to \infty$.

Указание. Воспользоваться формулой $(a-b)^r = \sum\limits_{m=0}^{r} (-1)^m C_r^m a^m b^{r-m}$ и примером 5.3.

4. Выборка объема $n \gg 1$ соответствует распределению $\mathcal{N}(0;\sigma^2)$. Найти асимптотическое распределение выборочного момента $\overline{\nu}_2(n)$.

Ответ. $\mathcal{N}\left(\sigma^2; \dfrac{2\sigma^4}{n}\right)$.

5. Двумерная выборка объема $n \gg 1$ соответствует распределению $\mathcal{N}(\mu; K)$, где $K = \begin{bmatrix} D_X & k_{XY} \\ k_{YX} & D_Y \end{bmatrix}$. Доказать, что $\sqrt{n}\left(\widehat{k}_{XY}(n) - k_{XY}\right) \xrightarrow{\text{d}} \xi \sim \mathcal{N}\left(0; D_X D_Y + k_{XY}^2\right)$.

Указание. Вычислить μ_{22}, воспользоваться теоремой 5.5.

6. Выборка соответствует распределению $E(\lambda)$, $\lambda > 0$. Найти предел, к которому почти наверное сходится $\overline{\nu}_2(n)$ при $n \to \infty$.

Ответ. $\dfrac{2}{\lambda^2}$.

7. Выборка объема $n \gg 1$ порождена СВ $X \sim E(1)$. Оценить $\mathbf{P}\left(|\widehat{F}_n(1) - F_X(1)| \leqslant \dfrac{1}{\sqrt{n}}\right)$.

Ответ. $2\Phi_0\left(\dfrac{e}{\sqrt{e-1}}\right)$.

§ 6. Точечные оценки и их свойства

6.1. Теоретические положения. Пусть $\theta \in \Theta \subseteq \mathbb{R}^1$ — некоторая детерминированная или случайная величина (параметр), а $Z_n = \{X_k, k =$

$= 1, \ldots, n\}$ — выборка.

Определение 6.1. *Точечной оценкой параметра* θ по выборке Z_n называется любая статистика $\widehat{\theta}_n = \varphi_n(Z_n)$, принимающая значения из множества Θ.

На практике вычисляют реализацию оценки $\widehat{\theta}_n$ (по имеющейся реализации z_n) и принимают ее за приближенное значение параметра θ. Поэтому желательно, чтобы при любом возможном θ величина $\widehat{\theta}_n$ была бы близка к θ.

Определение 6.2. Величина $\Delta\widehat{\theta}_n = \widehat{\theta}_n - \theta$ называется *ошибкой оценки* $\widehat{\theta}_n$.

Определение 6.3. Оценка $\widehat{\theta}_n$ называется *несмещенной*, если $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} = 0$. Если же $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} \neq 0$, но $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} \to 0$, $n \to \infty$, то оценка $\widehat{\theta}_n$ называется *асимптотически несмещенной.*

Определение 6.4. Оценка $\widehat{\theta}_n$ называется *сильно состоятельной*, если $\Delta\widehat{\theta}_n \xrightarrow{\text{п.н.}} 0$, $n \to \infty$ и *состоятельной в среднем квадратическом* (с.к.-состоятельной), если $\Delta\widehat{\theta}_n \xrightarrow{\text{с.к.}} 0$, $n \to \infty$.

Определение 6.5. *Среднеквадратической погрешностью* (с.к.-погрешностью) оценки $\widehat{\theta}_n$ называется величина

$$\Delta_n = \mathbf{M}\left\{|\Delta\widehat{\theta}_n|^2\right\}. \tag{6.1}$$

Теорема 6.1. *Пусть* $\theta \in \mathbb{R}^1$ *и* $\mathbf{M}\left\{|\Delta\widehat{\theta}_n|^2\right\} < \infty$, *тогда*

$$\Delta_n = l_n^2 + d_n, \tag{6.2}$$

где $l_n = \mathbf{M}\left\{\Delta\widehat{\theta}_n\right\}$ *— смещение оценки* $\widehat{\theta}_n$, *а* $d_n = \mathbf{D}\left\{\Delta\widehat{\theta}_n\right\}$ *— дисперсия ее ошибки.*

Определение 6.6. Оценка $\widehat{\theta}_n$ называется *асимптотически нормальной*, если существует детерминированная последовательность $\{C_n, n = 1, 2, \ldots\}$ такая, что $C_n\Delta\widehat{\theta}_n \xrightarrow{\text{d}} \xi \sim \mathcal{N}(0;1)$, $n \to \infty$.

Пусть теперь оценка $\widehat{\theta}_n = \varphi_n(Z_n)$ принадлежит некоторому заданному классу *допустимых оценок*, т.е. $\varphi_n \in \Phi_n$, $n = 1, 2, \ldots$, где Φ_n — фиксированный класс допустимых преобразований выборки Z_n.

Определение 6.7. Оценка $\widehat{\theta}_n = \varphi(Z_n)$ называется *оптимальной в среднем квадратическом* (с.к.-оптимальной) на Φ_n, если

$$\Delta_n = \mathbf{M}\left\{|\Delta\widehat{\theta}_n|^2\right\} \leqslant \mathbf{M}\left\{|\theta_n - \widetilde{\theta}_n|^2\right\}, \quad n = 1, 2, \ldots,$$

где $\widetilde{\theta}_n$ — произвольная допустимая оценка: $\widetilde{\theta}_n = \psi_n(Z_n)$, $\psi_n \in \Phi_n$.

Если $\theta \in \mathbb{R}^m$, где $m \geqslant 2$, то все вышеприведенные определения остаются в силе со следующими уточнениями:

1) в (6.2) величина $l_n^2 = \delta_n^\top \delta_n$, где $\delta_n = \mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} \in \mathbb{R}^m$ — вектор смещения оценки $\widehat{\theta}_n$, а $d_n = \mathrm{tr}[K_n]$, где $K_n = \mathbf{cov}(\Delta\widehat{\theta}_n, \Delta\widehat{\theta}_n)$ — ковариационная матрица ошибки $\Delta\widehat{\theta}_n$, $\mathrm{tr}[A]$ — след матрицы A;

2) в определении 6.6 $\{C_n, n = 1, 2, \ldots\}$ — последовательность неслучайных матриц размера $m \times m$, а предельное распределение $\mathcal{N}(0; I)$ — m-мерное стандартное гауссовское распределение.

6.2. Примеры.

Пример 6.1. Пусть выборка $\{X_k, k = 1, \ldots, n\}$ имеет вид

$$X_k = \theta + \varepsilon_k, \quad k = 1, \ldots, n,$$

θ — неслучайный скалярный параметр, $\{\varepsilon_k, k = 1, \ldots, n\}$ — независимые случайные величины, $\mathbf{M}\{\varepsilon_k\} = 0$, $\mathbf{D}\{\varepsilon_k\} = D_k \leqslant \overline{D} < \infty$ для всех $k \geqslant 1$. Доказать, что выборочное среднее $\overline{X}_n$ является несмещенной и состоятельной оценкой θ.

Решение. По определению 4.1 $\widehat{\theta}_n = \overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^n X_k$, поэтому $\overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^n (\theta + \varepsilon_k) = \theta + \dfrac{1}{n}\sum\limits_{k=1}^n \varepsilon_k = \theta + \overline{\varepsilon}_n$. Отсюда $\Delta\overline{X}_n = \Delta\widehat{\theta}_n = \overline{X}_n - \theta = \overline{\varepsilon}_n$ — ошибка оценки $\widehat{\theta}_n = \overline{X}_n$. Погрешность $\Delta_n = \mathbf{M}\{|\Delta\overline{X}_n|^2\} = \mathbf{M}\{|\overline{\varepsilon}_n|^2\} = \dfrac{1}{n^2}\sum\limits_{k=1}^n \mathbf{D}\{\varepsilon_k\} \leqslant \dfrac{\overline{D}}{n} \to 0$, $n \to \infty$. Поэтому оценка $\overline{X}_n$ с.к.-состоятельна.

Докажем теперь сильную состоятельность оценки $\overline{X}_n$. Так как $\mathbf{M}\{\varepsilon_k\} = a_k = 0$, а $\sum\limits_{k=1}^\infty \dfrac{D_k}{k^2} \leqslant \sum\limits_{k=1}^\infty \dfrac{\overline{D}}{k^2} = \overline{D}\sum\limits_{k=1}^\infty \dfrac{1}{k^2} < \infty$, то $\overline{\varepsilon}_n = \dfrac{1}{n}\sum\limits_{k=1}^n \varepsilon_k \xrightarrow{\text{п.н.}} 0$, $n \to \infty$ по теореме 4.2. Поэтому $\Delta\overline{X}_n = \overline{\varepsilon}_n \xrightarrow{\text{п.н.}} 0$, $n \to \infty$, т.е. $\overline{X}_n \xrightarrow{\text{п.н.}} \theta$, $n \to \infty$, что означает сильную состоятельность $\overline{X}_n$.

Наконец, для любого $n \geqslant 1$ $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} = \mathbf{M}\{\Delta\overline{X}_n\} = \mathbf{M}\{\overline{\varepsilon}_n\} = \dfrac{1}{n}\sum\limits_{k=1}^n \mathbf{M}\{\varepsilon_k\} = 0$, т.е. оценка $\overline{X}_n$ — несмещенная. ■

Пример 6.2. Пусть в условиях примера 6.1 СВ $\{\varepsilon_k k = 1, 2, \ldots\}$ одинаково распределены, причем $\mathbf{M}\{\varepsilon_k\} = 0$, $\mathbf{D}\{\varepsilon_k\} = \sigma^2$, где $\sigma < \infty$. Доказать, что оценка $\widehat{\theta}_n = \overline{X}_n$ асимптотически нормальна.

Решение. Из решения примера 6.1 следует, что $\Delta\widehat{\theta}_n = \Delta\overline{X}_n = \overline{\varepsilon}_n$, причем $\mathbf{M}\{\varepsilon_k\} = 0$, $\mathbf{D}\{\varepsilon_k\} = \sigma^2$. Тогда из теоремы 4.4 следует, что $\sqrt{n}\overline{\varepsilon}_n \xrightarrow{\mathrm{d}} X \sim \mathcal{N}(0;\sigma^2)$, $n \to \infty$. Отсюда $\dfrac{\sqrt{n}}{\sigma}\Delta\overline{X}_n = \dfrac{\sqrt{n}}{\sigma}\overline{\varepsilon}_n \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0;1)$, $n \to \infty$. Таким образом, $C_n\Delta\overline{X}_n \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0;1)$, $n \to \infty$, если $\left\{C_n = \dfrac{\sqrt{n}}{\sigma}, n = 1, 2, \ldots\right\}$. ■

Пример 6.3. Выборка $\{X_k, k = 1, \ldots, n\}$ порождена СВ $X \sim R[0;\theta]$, $\theta > 0$. Доказать, что $\widehat{\theta}_n = X_{(n)}$ — асимптотически несмещенная оценка параметра θ.

Решение. По условию $F(x) = \mathbf{P}(X \leqslant x) = \dfrac{x}{\theta}$, $x \in [0;\theta]$. Из примера 5.1 следует, что $F_{(n)}(x) = \mathbf{P}(X_{(n)} \leqslant x) = F^n(x) = \dfrac{x^n}{\theta^n}$, $x \in [0;\theta]$. Тогда $\mathbf{M}\left\{X_{(n)}\right\} = \int\limits_0^\theta x dF_{(n)}(x) = \int\limits_0^\theta x\dfrac{nx^{n-1}}{\theta^n}dx = \dfrac{n}{\theta^n}\int\limits_0^\theta x^n dx = \dfrac{n}{n+1}\theta$. Отсюда $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} = \mathbf{M}\left\{X_{(n)} - \theta\right\} = \mathbf{M}\left\{X_{(n)}\right\} - \theta = \dfrac{n}{n+1}\theta - \theta = -\dfrac{\theta}{n+1} \to 0$, $n \to \infty$. Итак, при любом $\theta > 0$ $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} < 0$, поэтому смещение $l_n \neq 0$, но $\mathbf{M}\left\{\Delta\widehat{\theta}_n\right\} \to 0$, $n \to \infty$, т.е. $\widehat{\theta}_n = X_{(n)}$ асимптотически несмещенная. ■

Пример 6.4. Выборка $Z_n = \{X_k, k = 1, \ldots, n\}$ соответствует распределению $\mathcal{N}(m;\theta^2)$. Найти величину C, при которой статистика $\varphi(Z_n) = \dfrac{C}{n}\sum\limits_{k=1}^n |X_k - m|$ будет несмещенной и сильно состоятельной оценкой параметра θ.

Решение. Пусть $\xi = |X - m|$, где $X \sim \mathcal{N}(m;\theta^2)$. Тогда $\mathbf{M}\{\xi\} = \mathbf{M}\{|X - m|\} = \dfrac{1}{\sqrt{2\pi}\theta}\int\limits_{-\infty}^{\infty} |x - m|\exp\left\{-\dfrac{(x-m)^2}{2\theta^2}\right\}dx = \sqrt{\dfrac{2}{\pi}}\theta\int\limits_0^\infty y\exp\left\{-\dfrac{y^2}{2}\right\}dy = \sqrt{\dfrac{2}{\pi}}\theta$. Отсюда с учетом $\mathbf{M}\{|X_k - m|\} = \mathbf{M}\{\xi\}$ получим $\mathbf{M}\{\varphi(Z_n) - \theta\} = \dfrac{C}{n}\sum\limits_{k=1}^n \mathbf{M}\{|X_k - m|\} - \theta = \left(C\sqrt{\dfrac{2}{\pi}} - 1\right)\theta$.

Последнее выражение равно нулю при всех θ, только если $C\sqrt{\dfrac{2}{\pi}} - 1 = 0$. Отсюда $C = \sqrt{\dfrac{\pi}{2}}$ есть условие несмещенности оценки $\widehat{\theta}_n = \varphi(Z_n)$.

По усиленному закону больших чисел имеем: $\nu_n = \dfrac{1}{n}\sum\limits_{k=1}^n |X_k - m| \xrightarrow{\text{п.н.}} \mathbf{M}\{\xi\} = \sqrt{\dfrac{2}{\pi}}\theta$, $n \to \infty$. Поэтому $\widehat{\theta}_n = C\nu_n \xrightarrow{\text{п.н.}} C\mathbf{M}\{\xi\} = \theta$, если $C = \sqrt{\dfrac{\pi}{2}}$. Итак, оценка $\widehat{\theta}_n = \dfrac{\sqrt{\pi}}{n\sqrt{2}}\sum\limits_{k=1}^n |X_k - m|$ — несмещенная и сильно состоятельная оценка среднего квадратического отклонения θ. ■

Пример 6.5. Пусть выборка $Z_n = \{X_k, k = 1, \ldots, n\}$ порождена СВ X, причем $\mathbf{M}\{X\} = \theta$, а $\mathbf{D}\{X\} = \sigma^2$ — известная величина. Доказать, что оценка $\widehat{\theta}_n = \overline{X}_n$ параметра θ с.к.-оптимальна на классе всех линейных несмещенных оценок вида $\widetilde{\theta}_n = \sum\limits_{k=1}^n \alpha_k X_k$, где $\{\alpha_k\}$ — некоторые числовые коэффициенты.

Решение. По условию $\widetilde{\theta}_n$ — несмещенная оценка, поэтому $\mathbf{M}\left\{\widetilde{\theta}_n - \theta\right\} = \mathbf{M}\left\{\sum\limits_{k=1}^n \alpha_k X_k - \theta\right\} = \sum\limits_{k=1}^n \alpha_k\theta - \theta = \left(\sum\limits_{k=1}^n \alpha_k - 1\right)\theta$. Таким образом, условие несмещенности $\mathbf{M}\left\{\widetilde{\theta}_n - \theta\right\} = 0$ влечет условие $\sum\limits_{k=1}^n \alpha_k = 1$. Обозначим через Φ_n соответствующий класс оценок. Заметим, что если $\alpha_k = \dfrac{1}{n}$, $k = 1, \ldots, n$, то $\sum\limits_{k=1}^n \alpha_k = 1$, поэтому оценка $\widehat{\theta}_n = \sum\limits_{k=1}^n \alpha_k X_k = \dfrac{1}{n}\sum\limits_{k=1}^n X_k = \overline{X}_n$ принадлежит классу Φ_n.

Найдем с.к.-погрешность произвольной оценки $\widetilde{\theta}_n$ из Φ_n. Так как $\mathbf{M}\left\{\widetilde{\theta}_n - \theta\right\} = 0$ по доказанному выше, то $\Delta_n = \mathbf{M}\left\{|\widetilde{\theta}_n - \theta|^2\right\} = \mathbf{D}\left\{\widetilde{\theta}_n - \theta\right\} = \mathbf{D}\left\{\widetilde{\theta}_n\right\} = \sigma^2\sum\limits_{k=1}^n \alpha_k^2$. Таким образом, коэффициенты $\{\widehat{\alpha}_k, k = 1, \ldots, n\}$, определяющие оптимальную оценку $\widehat{\theta}_n$, удовлетворяют условию

$$\sum_{k=1}^{n} \widehat{\alpha}_k^2 \leqslant \sum_{k=1}^{n} \alpha_k^2 \text{ для любых } \{\alpha_k\} \text{ таких, что } \sum_{k=1}^{n} \alpha_k = 1.$$

Обозначим $e = \{1, \ldots, 1\}^\top$, $\alpha = \{\alpha_1, \ldots, \alpha_n\}^\top$. Из неравенства Коши–Буняковского следует:

$$1 = \left(\sum_{k=1}^{n} \alpha_k\right)^2 = |(e, \alpha)|^2 \leqslant |e|^2 |\alpha|^2,$$

причем равенство достигается только при $\alpha = \lambda e$. Отсюда $|\alpha|^2 = \sum_{k=1}^{n} \alpha_k^2 \geqslant \geqslant \frac{1}{|e|^2} = \frac{1}{n}$. Если теперь положить $\widehat{\alpha}_k = \frac{1}{n}$, $k = 1, \ldots, n$, то $\sum_{k=1}^{n} \widehat{\alpha}_k^2 = = \frac{1}{n} \leqslant \sum_{k=1}^{n} \alpha_k^2$. Итак, $\widehat{\theta}_n = \sum_{k=1}^{n} \widehat{\alpha}_k X_k = \frac{1}{n} \sum_{k=1}^{n} X_k = \overline{X}_n$ — с.к.-оптимальная оценка на классе Φ_n всех линейных несмещенных оценок. Заметим также, что оценка $\widehat{\theta}_n$ — единственная (в силу единственности набора оптимальных коэффициентов $\{\widehat{\alpha}_k = \frac{1}{n}, k = 1, \ldots, n\}$). ■

6.3. Задачи для самостоятельного решения.

1. Доказать теорему 6.1.

Указание. Учесть, что $\mathbf{M}\{X^2\} = D_X + m_X^2$.

2. Пусть θ — неслучайный параметр, а $\widehat{\theta}_n$ — его несмещенная оценка. Показать, что $\Delta_n = \mathbf{D}\left\{\widehat{\theta}_n\right\}$.

3. Выборка $\{X_k, k = 1, \ldots, n\}$ порождена СВ X с известным средним $m = \mathbf{M}\{X\}$ и неизвестной дисперсией $\theta = \mathbf{D}\{X\}$. Доказать, что статистика $S_n^2 = \frac{1}{n} \sum_{k=1}^{n} (X_k - m)^2$ является несмещенной и сильно состоятельной оценкой параметра θ.

4. Выборка $\{X_k, k = 1, \ldots, n\}$ соответствует распределению $R[0; \theta]$, $\theta > 0$. Показать, что $X_{(n)}$ — с.к.-состоятельная оценка параметра θ.

Указание. Вычислить $\mathbf{M}\left\{(X_{(n)})^2\right\}$ и учесть результат примера 6.3.

5. Выборка $\{X_k, k = 1, \ldots, n\}$ соответствует распределению $E(\theta)$, $\theta > > 0$. Доказать, что $\widehat{\theta}_n = \sqrt{\dfrac{2n}{\sum_{k=1}^{n} X_k^2}}$ является сильно состоятельной оценкой параметра θ.

Указание. Найти п.н.-предел $\xi_n = \frac{1}{n} \sum_{k=1}^{n} X_k^2$.

6. Пусть выборка $\{X_k, k = 1, \ldots, n\}$ соответствует нормальному распределению $\mathcal{N}(\theta, \sigma^2)$, где σ — известно. Показать, что статистика $T_n = (\overline{X}_n)^2 - \frac{\sigma^2}{n}$ несмещенно и сильно состоятельно оценивает функцию $g(\theta) = \theta^2$.

Указание. Показать, что $\mathbf{M}\{(\overline{X}_n)^2\} = \theta^2 + \frac{\sigma^2}{n}$.

7. Выборка $\{X_k, k = 1, \ldots, n\}$ порождена СВ $X \sim R[0; \theta]$, $\theta > 0$. Доказать, что $\widehat{\theta}_n = 2\overline{X}_n$ — несмещенная и сильно состоятельная оценка для θ.

8. Пусть выборка $\{X_k, k = 1, \ldots, n\}$ соответствует распределению $Bi(N; p)$, где N — известно. Показать, что статистика $T_n = \frac{\overline{X}_n(N - \overline{X}_n)}{N}$ является асимптотически несмещенной и сильно состоятельной оценкой параметра $\theta = \mathbf{D}\{X_1\}$.

Указание. Вычислить $\mathbf{M}\{\overline{X}_n^2\} = \mathbf{D}\{\overline{X}_n\} + \left(\mathbf{M}\{\overline{X}_n\}\right)^2$ и учесть, что $\mathbf{M}\{\overline{X}_n\} = pN$.

§ 7. Методы построения точечных оценок параметров

7.1. Теоретические положения. Пусть $Z_n = \{X_k, k = = 1, 2, \ldots, n\}$ — выборка, порожденная СВ X, функция распределения которой $F_X(x; \theta)$ известна с точностью до m-мерного вектора $\theta = = \{\theta_1, \ldots, \theta_m\}^\top$ неизвестных неслучайных параметров. Для построения оценок параметров $(\theta_1, \ldots, \theta_m)$ по выборке Z_n можно использовать *метод моментов*, если СВ X имеет конечные начальные моменты ν_r для всех $r \leqslant m$.

Алгоритм метода моментов.

1) Найти аналитические выражения для начальных моментов ν_r:

$$\nu_r(\theta) = \mathbf{M}\{X^r\} = \int_{-\infty}^{\infty} x^r dF_X(x; \theta), \quad r = 1, \ldots, m. \tag{7.1}$$

2) Вычислить соответствующие выборочные начальные моменты

$$\overline{\nu}_r(n) = \frac{1}{n} \sum_{k=1}^{n} (X_k)^r, \quad r = 1, \ldots, m. \tag{7.2}$$

3) Составить систему из m уравнений для переменных $\{\theta_1, \ldots, \theta_m\}^\top$, приравнивая соответствующие теоретические (7.1) и выборочные (7.2) моменты:

$$\nu_r(\theta) = \overline{\nu}_r(n), \quad r = 1, \ldots, m. \tag{7.3}$$

4) Найти решение $\widehat{\theta}_n$ системы уравнений (7.3).

Определение 7.1. Решение $\widehat{\theta}_n$ системы уравнений (7.3) называется *оценкой метода моментов* вектора параметров θ закона распределения $F_X(x;\theta)$, которому соответствует выборка.

Заметим, что при составлении системы уравнений (7.3) можно использовать не только начальные моменты, но также и центральные моменты $\mu_r(\theta)$ и $\overline{\mu}_r(n)$, если это удобно.

Основным достоинством метода моментов является простота его практической реализации.

Важнейшим методом построения точечных оценок вектора θ является *метод максимального правдоподобия* (ММП). Предположим, что $\theta \in \Theta$, где Θ — множество допустимых значений вектора θ. Если СВ X, порождающая выборку, является дискретной, то пусть

$$p(x;\theta) = \mathbf{P}(X = x;\theta), \quad x \in \mathcal{X}, \tag{7.4}$$

где $\mathcal{X}$ — множество всех возможных значений СВ X, а $\mathbf{P}(X = x;\theta)$ — закон распределения дискретной СВ X.

Если же СВ X абсолютно непрерывна, то

$$p(x;\theta) = \frac{dF(x;\theta)}{dx}, \tag{7.5}$$

т.е. является плотностью вероятности СВ X.

Определение 7.2. *Функцией правдоподобия* выборки Z_n называется функция $L_n(\theta; Z_n)$, $\theta \in \Theta$ вида

$$L_n(\theta; Z_n) = \prod_{k=1}^{n} p\,(X_k;\theta). \tag{7.6}$$

Заметим, что для случая (7.5) функция $L_n(\theta; x)$ является *плотностью вероятности* случайного вектора Z_n в точке $x \in \mathbb{R}^n$.

Определение 7.3. Пусть $\widehat{\theta}_n$ — точка глобального максимума функции $L_n(\theta; Z_n)$ на Θ. Статистика $\widehat{\theta}_n$ называется *оценкой максимального правдоподобия* вектора θ (МП-оценкой).

Итак, $\widehat{\theta}_n = \arg\max\limits_{\theta\in\Theta} L_n(\theta; Z_n)$ — МП-оценка.

Обычно в расчетах используют *логарифмическую функцию правдоподобия*

$$\overline{L}_n(\theta) = \ln L_n(\theta; Z_n) = \sum_{k=1}^{n} \ln p\,(X_k;\theta). \tag{7.7}$$

Очевидно, что $\arg\max\limits_{\theta\in\Theta} L_n(\theta; Z_n) = \arg\max\limits_{\theta\in\Theta} \overline{L}_n(\theta)$.

Для построения МП-оценки $\widehat{\theta}_n$ можно использовать *необходимые условия экстремума* функции $\overline{L}_n(\theta)$:

$$\frac{\partial \overline{L}_n(\theta)}{\partial \theta_k} = 0, \quad k = 1, \ldots, m. \tag{7.8}$$

Система уравнений (7.8), решением которой при определенных условиях является оценка $\widehat{\theta}_n$, называется *системой уравнений правдоподобия.*

Следующее утверждение называется *принципом инвариантности* для оценивания по методу максимального правдоподобия.

Теорема 7.1. *Пусть выборка Z_n соответствует распределению $F(x;\theta)$, $\theta \in \Theta$, а функция $g(\theta)$ отображает Θ в некоторый промежуток Δ действительной оси. Тогда если $\widehat{\theta}_n$ — МП-оценка вектора θ, то $g(\widehat{\theta}_n)$ — МП-оценка функции $g(\theta)$.*

При определенных условиях МП-оценка параметра θ обладает замечательными асимптотическими свойствами. Предположим, что θ_0 — истинное значение скалярного параметра θ, Θ — замкнутое ограниченное подмножество $\mathbb{R}^1$, а θ_0 лежит внутри Θ. Пусть также выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$ соответствует распределению с плотностью вероятности $p(x;\theta)$.

Теорема 7.2. *Пусть выполнены следующие условия.*

1) *При каждом $\theta \in \Theta$ $\left|\dfrac{\partial^{(k)} p(x;\theta)}{\partial\theta^{(k)}}\right| \leqslant g_k(x)$, $k = 1, 2, 3$, причем $g_1(x)$ и $g_2(x)$ интегрируемы на $\mathbb{R}^1$, а $\sup\limits_{\theta\in\Theta} \int\limits_{-\infty}^{\infty} g_3(x) p(x;\theta)dx < \infty$.*

2) *При каждом $\theta \in \Theta$ функция*

$$i(\theta) = \int\limits_{-\infty}^{\infty} \left[\frac{\partial \ln p(x;\theta)}{\partial\theta}\right]^2 p(x;\theta)dx$$

конечна и положительна.

Тогда уравнение правдоподобия (7.8) имеет решение $\widehat{\theta}_n$, обладающее следующими свойствами:

а) $\mathbf{M}\left\{\widehat{\theta}_n-\theta_0\right\}\to 0$, $n\to\infty$ *(асимптотическая несмещенность);*

б) $\widehat{\theta}\xrightarrow{\text{п.н.}}\theta_0$, $n\to\infty$ *(сильная состоятельность);*

в) $\sqrt{n\,i(\theta_0)}(\widehat{\theta}_n-\theta_0)\xrightarrow{\text{d}}\xi\sim\mathcal{N}(0;1)$, $n\to\infty$ *(асимптотическая нормальность).*

Утверждения теоремы 7.2 могут быть обобщены на случай многомерного параметра θ.

7.2. Примеры.

Пример 7.1. Выборка Z_n порождена СВ $X\sim R[\theta_1;\theta_2]$, $\theta_1<\theta_2$. Найти оценку вектора $\theta=\{\theta_1,\theta_2\}^\top$ методом моментов.

Решение. Известно, что $\nu_1(\theta)=\mathbf{M}\{X\}=\dfrac{\theta_1+\theta_2}{2}$, а $\mu_2(\theta)=$ $=\mathbf{M}\left\{(X-\nu_1(\theta))^2\right\}=\mathbf{D}\{X\}=\dfrac{(\theta_2-\theta_1)^2}{12}$. Выборочными оценками моментов $\nu_1(\theta)$ и $\mu_2(\theta)$ являются, соответственно, выборочное среднее и выборочная дисперсия (см. раздел 6):

$$\overline{\nu}_1(n)=\overline{X}_n=\frac{1}{n}\sum_{k=1}^{n}X_k,$$

$$\overline{\mu}_2(n)=\overline{S}^2_n=\frac{1}{n}\sum_{k=1}^{n}(X_k-\overline{X}_n)^2.$$

Подставляя найденные теоретические и выборочные моменты в систему уравнений метода моментов (7.3), получаем

$$\begin{cases}\theta_1+\theta_2=2\overline{X}_n,\\ \theta_2-\theta_1=2\sqrt{3}\,\overline{S}_n.\end{cases}$$

Решая полученную систему уравнений относительно θ_1, θ_2, находим окончательный вид оценок:

$$\widehat{\theta}_1=\overline{X}_n-\sqrt{3}\,\overline{S}_n,\quad \widehat{\theta}_2=\overline{X}_n+\sqrt{3}\,\overline{S}_n.$$

■

Пример 7.2. В условиях примера 7.1 найти оценки максимального правдоподобия параметров θ_1 и θ_2.

Решение. По условию $p(x;\theta)=\begin{cases}\dfrac{1}{\theta_2-\theta_1}, & \text{если } x\in[\theta_1,\theta_2];\\ 0, & \text{если } x\notin[\theta_1,\theta_2].\end{cases}$ Отсюда

$$L_n(\theta;Z_n)=\begin{cases}\dfrac{1}{(\theta_2-\theta_1)^n}, & \text{если } X_i\in[\theta_1,\theta_2],\ i=1,\ldots,n;\\ 0, & \text{если } \exists j: X_j\notin[\theta_1,\theta_2].\end{cases}$$

Из полученного выражения следует, что при любых $\theta_1<\theta_2$ $L_n(\theta;Z_n)\leqslant$ $\leqslant\dfrac{1}{(X_{(n)}-X_{(1)})^n}=L_{max}$, где $X_{(1)}=\min(X_1,\ldots,X_n)$, $X_{(n)}=$ $=\max(X_1,\ldots,X_n)$. Отсюда $\widehat{\theta}_1=X_{(1)}$, $\widehat{\theta}_2=X_{(n)}$, так как $L_n(\widehat{\theta}_1,\widehat{\theta}_2;Z_n)=L_{max}$. Заметим, что МП-оценки $\widehat{\theta}_1$, $\widehat{\theta}_2$ не совпадают с оценками метода моментов, построенными в примере 7.1. ■

Пример 7.3. Пусть выборка $Z_n=\{X_k,\ k=1,\ldots,n\}$ соответствует распределению $Bi(N;p)$, где N — известно. Найти МП-оценку параметра θ (с учетом $\theta\in(0;1)$).

Решение. Из условия следует, что $p(x;\theta)=C_N^x\theta^x(1-\theta)^{N-x}$, где $x=$ $=0,1,\ldots,N$, а $C_N^x=\dfrac{N!}{x!(N-x)!}$. Поэтому функция правдоподобия имеет вид

$$L_n(\theta;Z_n)=\prod_{k=1}^{n}p(X_k;\theta)=\prod_{k=1}^{n}C_N^{X_k}\theta^{X_k}(1-\theta)^{N-X_k}. \tag{7.9}$$

Логарифмируя (7.9), найдем логарифмическую функцию правдоподобия:

$$\overline{L}_n(\theta)=\ln L_n(\theta;Z_n)=\sum_{k=1}^{n}\left(\ln C_N^{X_k}+X_k\ln\theta+(N-X_k)\ln(1-\theta)\right)=$$
$$=\sum_{k=1}^{n}\ln C_N^{X_k}+\ln\theta\sum_{k=1}^{n}X_k+\ln(1-\theta)\left(Nn-\sum_{k=1}^{n}X_k\right).$$

Уравнение правдоподобия (7.8) имеет вид

$$\frac{d\overline{L}_n(\theta)}{d\theta}=\frac{1}{\theta}\sum_{k=1}^{n}X_k-\frac{1}{1-\theta}\left(Nn-\sum_{k=1}^{n}X_k\right)=0.$$

Решая полученное уравнение относительно θ, находим $\widehat{\theta}_n = \dfrac{\sum\limits_{k=1}^{n} X_k}{Nn} = \dfrac{\overline{X}_n}{N}$. Оценка $\widehat{\theta}_n$ будет несмещенной, сильно состоятельной и асимптотически нормальной. ■

Пример 7.4. Дана гауссовская выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$, где $X_k \sim \mathcal{N}(\theta_1; \theta_2)$. Найти МП-оценку среднего θ_1 и дисперсии $\theta_2 > 0$.

Решение. По условию для $x \in \mathbb{R}^1$ и $\theta = \{\theta_1, \theta_2\}^\top$ имеем $p(x;\theta) = \dfrac{1}{\sqrt{2\pi\theta_2}}\exp\left\{-\dfrac{(x-\theta_1)^2}{2\theta_2}\right\}$. Поэтому

$$L_n(\theta; Z_n) = \prod_{k=1}^{n} p\,(X_k;\theta) = (2\pi\theta_2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\theta_2}\sum_{k=1}^{n}(X_k-\theta_1)^2\right\}.$$

Отсюда

$$\overline{L}_n(\theta) = \ln L_n(\theta; Z_n) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta_2 - \frac{1}{2\theta_2}\sum_{k=1}^{n}(X_k-\theta_1)^2.$$

Для нахождения максимума функции $\overline{L}_n(\theta)$ по θ воспользуемся уравнениями правдоподобия (7.8):

$$\begin{cases} \dfrac{\partial \overline{L}_n(\theta)}{\partial\theta_1} = \dfrac{1}{\theta_2}\sum\limits_{k=1}^{n}(X_k-\theta_1) = 0 \\ \dfrac{\partial \overline{L}_n(\theta)}{\partial\theta_2} = -\dfrac{n}{2\theta_2} + \dfrac{1}{2\theta_2^2}\sum\limits_{k=1}^{n}(X_k-\theta_1)^2 = 0. \end{cases}$$

Решая полученную систему уравнений относительно θ_1 и θ_2, находим требуемые оценки $\widehat{\theta}_1$ и $\widehat{\theta}_2$:

$$\widehat{\theta}_1 = \frac{1}{n}\sum_{k=1}^{n} X_k = \overline{X}_n; \qquad \widehat{\theta}_2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2 = \overline{S}_n^2.$$

Итак, выборочное среднее $\overline{X}_n$ и выборочная дисперсия $\overline{S}_n^2$ являются МП-оценками, соответственно, математического ожидания θ_1 и дисперсии θ_2 по гауссовской выборке. Из результатов примеров 5.2 и 5.3 следует, что $\widehat{\theta}_1$ — несмещенная и сильно состоятельная оценка θ_1, $\widehat{\theta}_2$ — асимптотически несмещенная и сильно состоятельная оценка для θ_2. Можно показать, что обе оценки асимптотически нормальны. ■

7.3. Задачи для самостоятельного решения.

1. Доказать, что оценки параметров, построенные в примерах 7.1 и 7.3, являются асимптотически несмещенными и сильно состоятельными.

Указание. Использовать асимптотические свойства выборочных моментов (см. разделы 4 и 5).

2. Выборка объема n порождена СВ $X \sim E(\theta)$, $\theta > 0$. Найти МП-оценку параметра θ и доказать ее сильную состоятельность.

Ответ. $\widehat{\theta}_n = \dfrac{1}{\overline{X}_n}$.

3. Выборка объема n соответствует распределению Пуассона $\Pi(\theta)$, $\theta > 0$. Найти МП-оценку для θ, доказать ее несмещенность, сильную состоятельность и асимптотическую нормальность.

Ответ. $\widehat{\theta}_n = \overline{X}_n$.

4. Выборка $\{X_k,\ k = 1, \ldots, n\}$ порождена СВ $X \sim E(\theta_1, \theta_2)$, $\theta_2 > 0$, т.е.

$$p_X(x) = \begin{cases} \theta_2\exp\{-\theta_2(x-\theta_1)\}\,, & \text{если } x \geqslant \theta_1, \\ 0, & \text{если } x < \theta_1. \end{cases}$$

Найти оценки параметров θ_1 и θ_2 методом моментов. Доказать сильную состоятельность полученных оценок.

Ответ. $\widehat{\theta}_1 = \overline{X}_n - \overline{S}_n$; $\widehat{\theta}_2 = \dfrac{1}{\overline{S}_n}$.

5. Выборка Z_n соответствует распределению Рэлея с функцией распределения $F(x;\theta) = 1 - \exp\left\{-\dfrac{x^2}{\theta}\right\}$, $x \geqslant 0$, $\theta > 0$. Найти МП-оценку параметра θ.

Ответ. $\widehat{\theta}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k)^2$.

6. Выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$ объема $n = 2m+1$ (m — натуральное) соответствует распределению Лапласа с плотностью $p(x;\theta) = \dfrac{1}{2}\exp\{-|x-\theta|\}$. Найти МП-оценку параметра θ.

Ответ. $\widehat{\theta}_n = X_{(m+1)}$.

7. В условиях задачи 6 для случая $n = 2m$ показать, что МП-оценкой для θ является любая статистика вида $\widehat{\theta}_n = (1-\lambda)X_{(m)} + \lambda X_{(m+1)}$, $\lambda \in [0;1]$.

8. Пусть $\widehat{\theta}_n$ — МП-оценка параметра θ распределения Бернулли $Bi(1;\theta)$. Показать, что последовательность $\sqrt{n}(\widehat{\theta}_n - \theta)$ асимптотически нормальна с параметрами $(0;\theta(1-\theta))$.

Указание. См. пример 7.3.

9. Выборка Z_n соответствует нормальному распределению с параметрами $(\sqrt{\theta};2)$, $\theta \geqslant 0$. Найти МП-оценку для θ.

Ответ. $\widehat{\theta}_n = (\overline{X}_n)^2$, если $\overline{X}_n \geqslant 0$ и $\widehat{\theta}_n = 0$ в противном случае.

10. Выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$ соответствует логнормальному распределению с параметрами $(\theta_1; \theta_2)$, т.е. $\ln X_k \sim \mathcal{N}(\theta_1; \theta_2)$. Найти МП-оценку параметра $\theta = \mathbf{M}\{X_k\}$, доказать ее сильную состоятельность.

Указание. Показать, что $\theta = \exp\left\{\theta_1 + \frac{\theta_2}{2}\right\}$; воспользоваться принципом инвариантности.

Ответ. $\widehat{\theta}_n = \exp\left\{\overline{Y}_n + \frac{\overline{S}_n^2}{2}\right\}$, где $\overline{Y}_n = \frac{1}{n}\sum\limits_{k=1}^{n} Y_k$, $\overline{S}_n^2 = \frac{1}{n}\sum\limits_{k=1}^{n}(Y_k - \overline{Y}_n)^2$, $Y_k = \ln X_k$, $k = 1, \ldots, n$.

11. Найти МП-оценку параметра $\theta = \mathbf{M}\{X_k\}$ по выборке $\{X_k,\ k = 1, \ldots, n\}$, соответствующей распределению $R[\theta_1, \theta_2]$.

Ответ. $\widehat{\theta}_n = \frac{X_{(1)} + X_{(n)}}{2}$.

§ 8. Эффективность точечных оценок

8.1. Теоретические положения. Для определенности предположим, что выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$ соответствует абсолютно непрерывному распределению $F(x; \theta)$ с плотностью $p(x; \theta)$, где $\theta \in \Theta \subseteq \mathbb{R}^1$, Θ — произвольный промежуток.

Определение 8.1. Распределение $F(x; \theta)$ называется *регулярным*, если выполнены следующие два условия.

R.1) Функция $\sqrt{p(x;\theta)}$ непрерывно дифференцируема по θ на Θ для почти всех x (по мере Лебега).

R.2) Функция

$$i(\theta) = \mathbf{M}_\theta\left\{\left(\frac{\partial \ln p(X;\theta)}{\partial\theta}\right)^2\right\} = \int\limits_{-\infty}^{\infty}\left(\frac{\partial \ln p(x;\theta)}{\partial\theta}\right)^2 p(x;\theta)dx \qquad (8.1)$$

конечна, положительна и непрерывна по θ на Θ.

В формуле (8.1) СВ X имеет плотность распределения $p(x; \theta)$, $\theta \in \Theta$, а $\mathbf{M}_\theta\{\xi\}$ означает усреднение СВ ξ по этому распределению.

Определение 8.2. Функция $i(\theta)$ называется *информационным количеством Фишера одного наблюдения* с распределением $p(x; \theta)$.

Если СВ X, порождающая выборку Z_n, является дискретной, а $\{\xi_1, \ldots, \xi_m\}$ — множество ее допустимых значений, то $\mathbf{P}(X = \xi_k; \theta) = P_k(\theta)$, $\theta \in \Theta$, $p(x; \theta) = \mathbf{P}(X = x; \theta)$, а в формуле (8.1) интеграл

заменяется суммой:

$$i(\theta) = \sum_{k=1}^{m}\left(\frac{\partial \ln p(\xi_k;\theta)}{\partial\theta}\right)^2 P_k(\theta). \qquad (8.2)$$

Пусть $L_n(\theta; Z_n)$ — функция правдоподобия выборки Z_n (см. определение 7.2), а $\overline{L}_n(\theta) = \ln L_n(\theta; Z_n)$ — логарифмическая функция правдоподобия.

Определение 8.3. Функция

$$U_n(\theta; Z_n) = \frac{d\,\overline{L}_n(\theta)}{d\,\theta} \qquad (8.3)$$

называется *вкладом выборки* Z_n.

Определение 8.4. Функция $I_n(\theta)$, определенная на Θ формулой

$$I_n(\theta) = \mathbf{M}_\theta\left\{U_n^2(\theta; Z_n)\right\} = \int\limits_{\mathbb{R}^n} U_n^2(\theta; x) L_n(\theta; x)dx, \qquad (8.4)$$

называется *количеством информации Фишера* о параметре θ, содержащимся в выборке Z_n, соответствующей распределению $p(x; \theta)$, $\theta \in \Theta$.

Теорема 8.1. *Пусть выполнены условия регулярности* R.1) *и* R.2), *тогда* $I_n(\theta) = n\, i(\theta)$, *где* $i(\theta)$ *имеет вид* (8.1) *или* (8.2).

Пусть $\theta_0 \in \Theta$ — неизвестное истинное значение параметра θ распределения $F(x; \theta)$, а $\widehat{\theta}_n$ — произвольная несмещенная оценка для θ_0, построенная по выборке Z_n: $\mathbf{M}\left\{\widehat{\theta}_n - \theta_0\right\} = 0$. Пусть также $\Delta_n = \mathbf{M}\left\{|\widehat{\theta}_n - \theta_0|^2\right\}$ — с.к.-погрешность оценки $\widehat{\theta}_n$.

Теорема 8.2 (неравенство Рао–Крамера). *Пусть выполнены условия регулярности* R.1) *и* R.2), *тогда справедливы следующие утверждения.*

1)

$$\Delta_n \geqslant \frac{1}{I_n(\theta_0)} = \Delta_n^{min}, \qquad (8.5)$$

где Δ_n^{min} *— нижняя граница Рао–Крамера с.к.-погрешности несмещенной оценки* $\widehat{\theta}_n$.

2) *Если в* (8.5) *для некоторой оценки* $\widehat{\theta}_n$ *достигается равенство, то ее можно представить в виде*

$$\widehat{\theta}_n = \theta_0 + a(\theta_0) U_n(\theta_0; Z_n), \qquad (8.6)$$

где $a(\theta)$ *не зависит от* Z_n, *а* $U_n(\theta; Z_n)$ *— вклад выборки* (8.3).

Определение 8.5. Несмещенная оценка $\widehat{\theta}_n$, с.к.-погрешность которой совпадает при всех $n \geqslant 1$ с нижней границей Δ_n^{min}, называется *эффективной по Рао–Крамеру (R-эффективной)*.

Из приведенных определений и утверждений следует:

1) эффективная оценка является с.к.-оптимальной на классе всех несмещенных оценок параметра θ_0;

2) если эффективная оценка существует, то она имеет вид (8.6).

Следующее утверждение поясняет связь между эффективной оценкой и МП-оценкой.

Теорема 8.3. *Пусть в условиях теоремы 8.2 существует эффективная оценка $\widehat{\theta}_n$, тогда она единственна и является МП-оценкой.*

Заметим, что с учетом несмещенности эффективной оценки $\widehat{\theta}_n$ и утверждения теоремы 8.1

$$\Delta_n(\theta_0) = \mathbf{M}\Big\{(\widehat{\theta}_n - \theta_0)^2\Big\} = \mathbf{D}\Big\{\widehat{\theta}_n\Big\} = \frac{1}{n\,i(\theta_0)}, \tag{8.7}$$

где $i(\theta)$ — информация Фишера одного наблюдения (8.1) или (8.2).

Из (8.7) видно, что $\mathbf{D}\Big\{\widehat{\theta}_n\Big\} = O\left(\dfrac{1}{n}\right)$, т.е. убывает с ростом объема выборки со скоростью, пропорциональной $\dfrac{1}{n}$. Кроме того, всякая эффективная оценка с.к.-состоятельна, так как $\Delta_n = \mathbf{D}\Big\{\widehat{\theta}_n\Big\} \to 0$, $n \to \infty$.

Для случая $\theta \in \Theta \subseteq \mathbb{R}^m$, $m > 1$ условия регулярности $R.1$) и $R.2$) принимают следующий вид:

R.1') $\sqrt{p(x;\theta)}$ непрерывно дифференцируема по θ_j, $j = 1, \ldots, m$ на Θ для почти всех x.

R.2') Матрица $I(\theta) = \{I_{ij}(\theta)\}$ с элементами

$$I_{ij}(\theta) = \int\limits_{-\infty}^{\infty} \frac{\partial \ln p(x;\theta)}{\partial \theta_i} \cdot \frac{\partial \ln p(x;\theta)}{\partial \theta_j}\, p(x;\theta)dx \tag{8.8}$$

непрерывна по θ на Θ и положительно определена.

В этом случае неравенство Рао–Крамера (8.5) принимает вид

$$\mathbf{M}\Big\{(\widehat{\theta}_n - \theta_0)(\widehat{\theta}_n - \theta_0)^\top\Big\} \geqslant \frac{I^{-1}(\theta_0)}{n}, \tag{8.9}$$

где θ_0 — истинное значение вектора параметров θ, а $\widehat{\theta}_n$ — его произвольная несмещенная оценка. Знак неравенства в (8.9) имеет следующий

смысл: если матрицы A и B симметричны и неотрицательно определены, то $A \geqslant B$ означает, что $A - B$ неотрицательно определена. Матрица $I(\theta)$ называется *информационной матрицей Фишера*.

Если выборка соответствует регулярному распределению, $\theta \in \Theta \subseteq \mathbb{R}^1$, а для некоторой несмещенной оценки $\widehat{\theta}_n$ выполнено $\dfrac{\mathbf{D}\Big\{\widehat{\theta}_n\Big\}}{\Delta_n^{min}} \to 1$, $n \to \infty$, то $\widehat{\theta}_n$ называется *асимптотически эффективной* оценкой.

8.2. Примеры.

Пример 8.1. Выборка Z_n соответствует распределению $\mathcal{N}(\theta;\sigma^2)$, $\sigma > 0$. Доказать, что выборочное среднее $\overline{X}_n$ является эффективной оценкой математического ожидания θ.

Решение. По условию $p(x;\theta) = \dfrac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\dfrac{(x-\theta)^2}{2\sigma^2}\right\}$, поэтому условие $R.1$) очевидно выполнено. Проверим условие $R.2$). Пусть $X \sim \sim \mathcal{N}(\theta;\sigma^2)$, тогда

$$l(X;\theta) = \ln p(X;\theta) = \ln\left(\frac{1}{\sqrt{2\pi}\sigma}\right) - \frac{(X-\theta)^2}{2\sigma^2}.$$

Отсюда $\varphi(X;\theta) = \dfrac{\partial l(X;\theta)}{\partial \theta} = \dfrac{X-\theta}{\sigma^2}$, и, следовательно,

$$i(\theta) = \mathbf{M}_\theta\left\{\varphi^2(X;\theta)\right\} = \mathbf{M}_\theta\left\{\frac{(X-\theta)^2}{\sigma^4}\right\} = \frac{\sigma^2}{\sigma^4} = \frac{1}{\sigma^2}.$$

Итак, информация Фишера для гауссовского распределения $i(\theta) = \dfrac{1}{\sigma^2}$ удовлетворяет $R.2$) при любом $\sigma \in (0,+\infty)$.

Теперь видно, что нижняя граница в неравенстве (8.5) Рао–Крамера $\Delta_n^{min} = \dfrac{1}{n\,i(\theta)} = \dfrac{\sigma^2}{n}$ и не зависит от θ.

Так как $\overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$ по определению, то $\mathbf{M}\{\overline{X}_n\} = \dfrac{1}{n}\sum\limits_{k=1}^{n}\mathbf{M}\{X_k\} = \dfrac{n\theta}{n} = \theta$, т.е. $\overline{X}_n$ — несмещенная оценка. При этом

$$\Delta_n = \mathbf{D}\{\overline{X}_n\} = \mathbf{D}\left\{\frac{1}{n}\sum_{k=1}^{n} X_k\right\} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Итак, $\Delta_n = \Delta_n^{min}$, поэтому $\overline{X}_n$ — эффективная оценка для θ при любой величине с.к.о. σ. ■

Пример 8.2. Показать, что распределение Бернулли $Bi(1;\theta)$, $\theta \in (0;1)$ является регулярным и найти информацию Фишера $i(\theta)$.

Решение. По условию $p(x;\theta) = \theta^x(1-\theta)^{1-x}$, $x = 0, 1$, а $\theta \in \Theta = (0;1)$. Обозначим $f(x;\theta) = \dfrac{\partial\sqrt{p(x;\theta)}}{\partial\theta} = -\dfrac{\partial p(x;\theta)}{\partial\theta}\dfrac{1}{2\sqrt{p(\theta;x)}}$. Если $x = 0$, то $p(x;\theta) = 1-\theta$, и, следовательно, $f(x;\theta) = -\dfrac{1}{2\sqrt{1-\theta}}$ непрерывна по θ на $(0;1)$. Аналогично, для $x = 1$ $p(x;\theta) = \theta$, т.е. $f(x;\theta) = -\dfrac{1}{2\sqrt{\theta}}$ также непрерывна по θ на Θ. Таким образом, условие регулярности $R.1)$ выполнено.

Теперь найдем $i(\theta)$. Если $X \sim Bi(1;\theta)$, то $p(X;\theta) = \theta^X(1-\theta)^{1-X}$. Отсюда $l(X;\theta) = \ln p(X;\theta) = X\ln\theta + (1-X)\ln(1-\theta)$. Поэтому $\varphi(X;\theta) = \dfrac{\partial l(X;\theta)}{\partial\theta} = \dfrac{X}{1-\theta} - \dfrac{1-X}{\theta} = \dfrac{X-\theta}{\theta(1-\theta)}$. Теперь

$$i(\theta) = \mathbf{M}_\theta\left\{\varphi^2(X;\theta)\right\} = \frac{\mathbf{M}_\theta\left\{(X-\theta)^2\right\}}{\theta^2(1-\theta)^2} = \frac{\mathbf{D}_\theta\{X\}}{\theta^2(1-\theta)^2} = \frac{\theta(1-\theta)}{\theta^2(1-\theta)^2} = \frac{1}{\theta(1-\theta)}.$$

Видно, что $0 < i(\theta) < \infty$ при любом $\theta \in \Theta$ и $i(\theta)$ непрерывна по θ на Θ, т.е. условие $R.2)$ также выполнено. ■

Пример 8.3. Доказать, что частота $\widehat{\theta}_n = P_n^*(A)$ случайного события A является эффективной оценкой вероятности $\theta = \mathbf{P}(A)$ этого события.

Решение. По определению $P_n^*(A) = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$, где $X_k \sim Bi(1;\theta)$ — независимые бернуллиевские СВ (см. пример 4.3). Поэтому $\mathbf{M}\{P_n^*(A)\} = \theta$, а $\mathbf{D}\{P_n^*(A)\} = \dfrac{\mathbf{D}\{X_1\}}{n} = \dfrac{\theta(1-\theta)}{n}$. Из примера 8.2 следует, что количество информации в выборке $Z_n = \{X_k,\ k = 1, \ldots, n\}$ о параметре θ равно $I_n(\theta) = n\, i(\theta) = \dfrac{n}{\theta(1-\theta)}$. Поэтому $\mathbf{D}\left\{\widehat{\theta}_n\right\} = \mathbf{D}\{P_n^*(A)\} = \dfrac{1}{I_n(\theta)}$, т.е. в неравенстве Рао–Крамера достигается нижняя граница. Таким образом, $\widehat{\theta}_n = P_n^*(A)$ эффективно оценивает $\theta = \mathbf{P}(A)$. Применимость теоремы Рао–Крамера в данном случае обосновывается регулярностью распределения $Bi(1;\theta)$ для всех $\theta \in (0;1)$, что было доказано в примере 8.2. ■

Следующий пример показывает, что выборочное среднее отнюдь не всегда является эффективной оценкой математического ожидания.

Пример 8.4. Выборка $\{X_k,\ k = 1, \ldots, n\}$ соответствует распределению Лапласа с параметрами (θ, λ), где $\lambda > 0$, т.е.

$$p(x;\theta) = \frac{1}{2\lambda}\exp\left\{-\frac{|x-\theta|}{\lambda}\right\}, \quad \theta \in \mathbb{R}^1. \tag{8.10}$$

Доказать, что $\overline{X}_n$ является несмещенной, но не эффективной оценкой среднего θ при любом известном λ.

Решение. Можно показать, что в условиях примера неравенство (8.5) выполнено, причем $I_n(\theta) = \dfrac{n}{\lambda^2}$.

Если X имеет распределение (8.10), то $\mathbf{M}\{X\} = \dfrac{1}{2\lambda}\displaystyle\int_{-\infty}^{\infty} x\exp\left\{-\frac{|x-\theta|}{\lambda}\right\}dx = \theta + \frac{\lambda}{2}\int_{-\infty}^{\infty} y\exp\{-|y|\}\,dy = \theta$, поэтому $\mathbf{M}\{\overline{X}_n\} = \mathbf{M}\{X\} = \theta$ (см. решение примера 8.1). Далее $\mathbf{D}\{\overline{X}_n\} = \dfrac{\mathbf{D}\{X\}}{n} = \dfrac{2\lambda^2}{n}$, так как $\mathbf{D}\{X\} = \dfrac{1}{2\lambda}\displaystyle\int_{-\infty}^{\infty}(x-\theta)^2\exp\left\{-\frac{|x-\theta|}{\lambda}\right\}dx = 2\lambda^2$.

Отсюда видно, что $\Delta_n = \mathbf{D}\{\overline{X}_n\} = \dfrac{2\lambda^2}{n} > \dfrac{1}{I_n(\theta)} = \dfrac{\lambda^2}{n} = \Delta_n^{min}$. Таким образом, с.к.-погрешность Δ_n оценки $\overline{X}_n$ параметра θ в два раза больше нижней границы Рао–Крамера Δ_n^{min} при любом объеме выборки n и любом $\theta \in \mathbb{R}^1$. Последнее означает, что $\overline{X}_n$ не может быть эффективной оценкой для θ. Более того, $\overline{X}_n$ не является даже асимптотически эффективной, так как $\dfrac{\Delta_n}{\Delta_n^{min}} \not\to 1$, $n \to \infty$. ■

Приведем пример нерегулярного распределения и рассмотрим точность МП-оценки параметра этого распределения.

Пример 8.5. Показать, что распределение $R[0;\theta]$, $\theta > 0$ нерегулярно. Исследовать поведение с.к.-погрешности МП-оценки параметра θ при $n \to \infty$.

Решение. Зафиксируем любое $x > 0$. По условию

$$p(x;\theta) = \begin{cases} 0, & \text{если } \theta < x, \\ \dfrac{1}{\theta}, & \text{если } \theta \geqslant x. \end{cases}$$

Таким образом, $\sqrt{p(x;\theta)}$ терпит разрыв в точке $\theta = x$ и, естественно, не является непрерывно дифференцируемой при любом $x > 0$. Итак, условие $R.1)$ нарушено.

Пусть $\widehat{\theta}_n$ — МП-оценка параметра θ, тогда $\widehat{\theta}_n = X_{(n)} = \max\{X_1, \ldots, X_n\}$ (см. пример 7.2). В примере 5.1 было показано, что $X_{(n)} \sim F_{(n)}(x) = \dfrac{x^n}{\theta^n}$, если $x \in [0;\theta]$, поэтому

$$p(x;\theta) = \begin{cases} \dfrac{nx^{n-1}}{\theta^n}, & \text{если } x \in [0;\theta], \\ 0, & \text{если } x \notin [0;\theta]. \end{cases}$$

Отсюда немедленно следует, что для любого $\theta \in \Theta$

$$\mathbf{M}\left\{\widehat{\theta}_n\right\} = \frac{n}{\theta^n}\int\limits_0^\theta x^n dx = \frac{n}{n+1}\theta.$$

Поэтому "подправленная" оценка $\widetilde{\theta}_n = \dfrac{n+1}{n}\widehat{\theta}_n$ — несмещенная. Найдем теперь дисперсию оценки $\widetilde{\theta}_n$:

$$\mathbf{M}\left\{(\widetilde{\theta}_n)^2\right\} = \left(\frac{n+1}{n}\right)^2 \mathbf{M}\left\{\left(\widehat{\theta}_n\right)^2\right\} = \left(\frac{n+1}{n}\right)^2 \int\limits_0^\theta x^2 \frac{nx^{n-1}}{\theta^n} dx =$$

$$= \frac{(n+1)^2}{n\theta^n}\int\limits_0^\theta x^{n+1}dx = \frac{(n+1)^2}{n(n+2)}\theta^2.$$

Поэтому $\mathbf{D}\left\{\widetilde{\theta}_n\right\} = \mathbf{M}\left\{(\widetilde{\theta}_n)^2\right\} - \theta^2 = \left[\dfrac{(n+1)^2}{n(n+2)} - 1\right]\theta^2 = \dfrac{\theta^2}{n(n+2)} = O\left(\dfrac{1}{n^2}\right)$.

Итак, мы видим, что для θ найдена несмещенная оценка, с.к.-погрешность которой убывает существенно быстрее, чем $O\left(\dfrac{1}{n}\right)$, что разрешено неравенством Рао–Крамера. Указанный эффект вызван нерегулярностью распределения $R[0;\theta]$ и известен как "сверхэффективность" оценки $\widetilde{\theta}_n$. ■

8.3. Задачи для самостоятельного решения.

1. Выборка соответствует распределению $Bi(N;\theta)$, $\theta \in (0;1)$. Проверить условия регулярности, найти $i(\theta)$ и доказать эффективность МП-оценки параметра θ.

Указание. $\widehat{\theta}_n = \dfrac{\overline{X}_n}{N}$.

Ответ. $i(\theta) = \dfrac{N}{\theta(1-\theta)}$.

2. Показать, что распределение Пуассона $\Pi(\theta)$, $\theta > 0$, регулярно. Найти $i(\theta)$. Доказать эффективность МП-оценки $\widehat{\theta}_n$ параметра θ.

Указание. $\widehat{\theta}_n = \overline{X}_n$.

Ответ. $i(\theta) = \dfrac{1}{\theta}$.

3. Для распределения $\mathcal{N}(\mu;\theta^2)$, $\theta > 0$, где μ — известно, найти информацию Фишера $i(\theta)$.

Ответ. $i(\theta) = \dfrac{2}{\theta^2}$.

4. Проверить регулярность распределения $E(\theta)$, $\theta > 0$, вычислить $I_n(\theta)$. Доказать, что оценка $\widetilde{\theta}_n = \dfrac{n-1}{n}\widehat{\theta}_n$ асимптотически эффективна, если $\widehat{\theta}_n$ — МП-оценка для θ.

Указание. $\widehat{\theta}_n = \dfrac{1}{\overline{X}_n}$.

Ответ. $I_n(\theta) = \dfrac{n}{\theta^2}$; $\mathbf{D}\left\{\widetilde{\theta}_n\right\} = \dfrac{\theta^2}{n-2}$.

5. Показать, что информация $I_n(\theta)$, содержащаяся в выборке Z_n, соответствующей распределению Лапласа (8.10), равна $\dfrac{n}{\lambda^2}$.

6. Сравнить по точности оценку θ_n^* параметра θ распределения $R[0;\theta]$, $\theta > 0$, полученную методом моментов, с оценкой $\widetilde{\theta}_n$, рассмотренной в примере 8.5.

Указание. $\theta_n^* = 2\overline{X}_n$.

Ответ. $\dfrac{\mathbf{D}\{\theta_n^*\}}{\mathbf{D}\left\{\widetilde{\theta}_n\right\}} = \dfrac{n+2}{3}$.

7. Пусть выборка соответствует распределению $\mathcal{N}(\mu;\theta)$, $\theta > 0$, μ — известно. Доказать, что МП-оценка дисперсии θ эффективна.

Указание. $\widehat{\theta}_n = \dfrac{1}{n}\sum\limits_{k=1}^n (X_k - \mu)^2$.

Ответ. $\mathbf{D}\left\{\widehat{\theta}_n\right\} = \dfrac{2\theta^2}{n}$; $i(\theta) = \dfrac{1}{2\theta^2}$.

8. Выборка Z_n соответствует распределению $\mathcal{N}(\theta_1;\theta_2^2)$. Найти информационную матрицу Фишера $I(\theta_1,\theta_2)$.

Ответ. $I(\theta_1,\theta_2) = \begin{bmatrix} \dfrac{1}{\theta_2^2} & 0 \\ 0 & \dfrac{2}{\theta_2^2} \end{bmatrix}$.

9. Для оценок, построенных в задачах 1, 2 и 7, найти их представление (8.7) через вклад выборки.

Ответ. Во всех случаях $a(\theta) = \dfrac{1}{I_n(\theta)}$.

§ 9. Интервальные оценки параметров

9.1. Теоретические положения. Пусть $Z_n = \{X_k,\ k = 1, \ldots, n\}$ — выборка, соответствующая распределению $F(x;\theta)$, где $\theta \in \Theta \subseteq \mathbb{R}^1$ — неизвестный параметр. Выберем некоторое малое положительное число p и предположим, что найдутся статистики $T_1 = T_1(Z_n)$ и $T_2 = T_2(Z_n)$, $T_1 < T_2$, такие, что для любого $\theta \in \Theta$

$$\mathbf{P}(T_1 \leqslant \theta \leqslant T_2) = 1 - p. \tag{9.1}$$

Определение 9.1. Промежуток $[T_1, T_2]$ называется *доверительным интервалом для θ надежности $q = 1 - p$*. Доверительный интервал также называют *интервальной оценкой* параметра θ.

Число $p = 1 - q$ называют *уровнем значимости* и обычно на практике полагают $p = 0{,}05$ или $p = 0{,}01$.

Пусть ξ — СВ, имеющая непрерывную функцию распределения $F_\xi(x)$.

Определение 9.2. Для любого $\alpha \in (0;1)$ число

$$x_\alpha = \min\{x : F_\xi(x) \geqslant \alpha\} \tag{9.2}$$

называется *квантилью уровня* α распределения $F_\xi(x)$.

Из (9.2) следует, что

$$\mathbf{P}(\xi \leqslant x_\alpha) = \alpha, \quad \mathbf{P}(\xi \geqslant x_\alpha) = 1 - \alpha. \tag{9.3}$$

Понятие квантили имеет существенное значение для построения доверительных интервалов и проверки статистических гипотез.

Центральный доверительный интервал. Пусть $G(Z_n;\theta)$ — такая СВ, что ее функция распределения $F_G(x) = \mathbf{P}\left(G(Z_n;\theta) \leqslant x\right)$ не зависит от θ. Пусть также для каждой реализации z_n выборки Z_n числовая функция $G_n(\theta) = G(z_n;\theta)$ непрерывна и строго монотонна по θ на Θ.

Определение 9.3. СВ $G(Z_n;\theta)$ называется *центральной статистикой* для θ.

Пусть задан уровень значимости p и выбраны произвольно $p_1 > 0$ и $p_2 > 0$ такие, что $p = p_1 + p_2$ (например, $p_1 = p_2 = \frac{p}{2}$). Если g_1 и g_2 — квантили распределения $F_G(x)$ уровней, соответственно, p_1 и $1 - p_2$, то для любого $\theta \in \Theta$

$$\mathbf{P}\left(g_1 \leqslant G(Z_n;\theta) \leqslant g_2\right) = 1 - p.$$

Найдем решения t_1 и t_2 уравнений $G(Z_n;\theta) = g_i$, $i = 1, 2$ и положим $T_1 = \min\{t_1, t_2\}$, $T_2 = \max\{t_1, t_2\}$. Тогда

$$\mathbf{P}\left(T_1 \leqslant \theta \leqslant T_2\right) = 1 - p = q,$$

т.е. $[T_1, T_2]$ — доверительный интервал для θ надежности q.

В силу произвола в выборе p_1 и p_2 интервал $[T_1, T_2]$ определен неоднозначно. Если при построении T_1 и T_2 с помощью $G(Z_n;\theta)$ дополнительно предположить, что $p_1 = p_2 = \frac{p}{2}$, то $[T_1, T_2]$ называют *центральным доверительным интервалом.*

В общем случае выбор p_1 и p_2 осуществляется так, чтобы длина интервала $T_2 - T_1$ была минимальной при неизменной надежности q (в этом случае интервальная оценка будет самой точной среди всех оценок надежности q).

Следующее утверждение дает общий способ построения центральной статистики.

Теорема 9.1. *Пусть выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$ соответствует функции распределения $F(x;\theta)$, удовлетворяющей следующим требованиям:*

1) *$F(x;\theta)$ непрерывна по x для любого $\theta \in \Theta$;*
2) *$F(x;\theta)$ непрерывна и монотонна по θ для любого x.*

Тогда $G(Z_n;\theta) = -\sum\limits_{k=1}^{n} \ln F(X_k;\theta)$ является центральной статистикой для $\theta \in \Theta$.

Асимптотический доверительный интервал. При больших объемах выборки ($n \gg 1$) для построения доверительного интервала можно воспользоваться любой асимптотически нормальной оценкой $\widehat{\theta}_n$ параметра θ. Пусть

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0; d(\theta)), \quad n \to \infty, \tag{9.4}$$

где $d(\theta)$ — *асимптотическая дисперсия* оценки $\widehat{\theta}_n$.

Зададим надежность q, уровень значимости $p = 1 - q$ и определим (по таблице 13.2) квантиль u_α уровня $\alpha = 1 - \dfrac{p}{2}$ распределения $\mathcal{N}(0;1)$. Так как функция Лапласа строго монотонна, $\varPhi(u_\alpha) = 1 - \dfrac{p}{2}$. Кроме того, если $\beta = \dfrac{p}{2}$, то $u_\beta = -u_\alpha$.

Если $d(\theta)$ непрерывна по $\theta \in \Theta$, то из (9.4) следует:

$$\mathbf{P}\left(\widehat{\theta}_n - u_\alpha\sqrt{\frac{d(\widehat{\theta}_n)}{n}} \leqslant \theta \leqslant \widehat{\theta}_n + u_\alpha\sqrt{\frac{d(\widehat{\theta}_n)}{n}}\right) \to \varPhi(u_\alpha) - \varPhi(-u_\alpha) = q.$$

Последнее означает, что интервал

$$\widehat{I} = \left[\widehat{\theta}_n - u_\alpha\sqrt{\frac{d(\widehat{\theta}_n)}{n}};\ \widehat{\theta}_n + u_\alpha\sqrt{\frac{d(\widehat{\theta}_n)}{n}}\right], \quad \alpha = 1 - \frac{p}{2} = \frac{1+q}{2}$$

при $n \gg 1$ накрывает оцениваемый параметр θ с вероятностью, близкой к $q = 1 - p$.

Если $\widehat{\theta}_n$ — МП-оценка параметра θ, то в условиях теоремы 7.2 $d(\theta) = \dfrac{1}{i(\theta)}$, где $i(\theta)$ — информация Фишера одного наблюдения. Пусть распределение, определяющее выборку, регулярно (см. определение 8.1), тогда $i(\theta) > 0$, $d(\theta) = \dfrac{1}{i(\theta)}$ непрерывна по θ, причем $\widetilde{d}(\theta) \geqslant d(\theta)$, если $\widetilde{d}(\theta)$ — асимптотическая дисперсия любой другой асимптотически нормальной оценки $\widetilde{\theta}_n$ параметра θ. Поэтому интервал $\widehat{I}$, построенный с использованием МП-оценки $\widehat{\theta}_n$, будет асимптотически наикратчайшим.

Рассмотрим теперь некоторые специальные вероятностные распределения, необходимые для построения доверительных интервалов.

Определение 9.4. Пусть $\{X_k, k = 1, \ldots, n\}$ — выборка, соответствующая распределению $\mathcal{N}(0; 1)$. Тогда СВ

$$\chi_n^2 = \sum_{k=1}^{n} (X_k)^2$$

имеет *χ^2-распределение ("хи-квадрат" распределение) с n степенями свободы*. Обозначение: $\chi_n^2 \sim \mathcal{H}_n$.

СВ χ_n^2 имеет плотность вероятности

$$p_{\chi_n^2}(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)}\, x^{\frac{n-2}{2}} \exp\Bigl\{-\dfrac{x}{2}\Bigr\}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

где $\Gamma(\lambda) = \int\limits_0^\infty t^{\lambda-1} e^{-t} dt$ — гамма функция.

Моментные характеристики: $\mathbf{M}\{\chi_n^2\} = n$, $\mathbf{D}\{\chi_n^2\} = 2n$.

Распределение $\mathcal{H}_n$ асимптотически нормально (по числу степеней свободы n): $\dfrac{\chi_n^2 - n}{\sqrt{2n}} \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0; 1)$, $n \to \infty$.

Определение 9.5. Пусть $X \sim \mathcal{N}(0; 1)$, $Y_n \sim \mathcal{H}_n$, X и Y — независимы. Тогда СВ

$$\tau_n = \frac{X}{\sqrt{\frac{1}{n} Y_n}}$$

имеет *распределение Стьюдента с n степенями свободы*.

Обозначение: $\tau_n \sim \mathcal{T}_n$.

СВ τ_n имеет плотность вероятности

$$p_{\tau_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

Свойства распределения $\mathcal{T}_n$:

1) если $n > 2$, то $\mathbf{M}\{\tau_n\} = 0$, $\mathbf{D}\{\tau_n\} = \dfrac{n}{n-2}$;

2) если $n = 1$, то τ_n имеет *распределение Коши*: $p_{\tau_n}(x) = \dfrac{1}{\pi(1 + x^2)}$;

3) асимптотическая нормальность: $\tau_n \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0; 1)$, $n \to \infty$.

Пусть теперь $Z_n = \{X_k, k = 1, \ldots, n\}$ — выборка, соответствующая распределению $\mathcal{N}(\theta; \sigma^2)$, $\sigma > 0$, $\overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$ — выборочное среднее, $\overline{S}_n^2 = \dfrac{1}{n}\sum\limits_{k=1}^{n} (X_k - \overline{X}_n)^2$ — выборочная дисперсия.

Теорема 9.2. *Статистики $\overline{X}_n$ и $\overline{S}_n^2$ независимы и обладают следующими свойствами:*

1) $\overline{X}_n \sim \mathcal{N}\left(\theta; \dfrac{\sigma^2}{n}\right)$;

2) $g_n = \dfrac{n\,\overline{S}_n^2}{\sigma^2} \sim \mathcal{H}_{n-1}$;

3) $\tau_{n-1} = \dfrac{\sqrt{n-1}(\overline{X}_n - \theta)}{\overline{S}_n} \sim \mathcal{T}_{n-1}$.

Утверждения теоремы 9.2 существенно облегчают построение доверительных интервалов для параметров гауссовского распределения.

9.2. Примеры.

Пример 9.1. Выборка $Z_n = \{X_k, k = 1, \ldots, n\}$ соответствует распределению $\mathcal{N}(\theta; \sigma^2)$; $\sigma^2 > 0$ — известная дисперсия. Построить для θ доверительный интервал надежности $q = 1 - p$.

Решение. Пусть $G(Z_n; \theta) = \dfrac{\sqrt{n}\,(\overline{X}_n - \theta)}{\sigma}$. По теореме 9.2 $G(Z_n; \theta) \sim \mathcal{N}(0; 1)$. При фиксированном $\overline{X}_n$ $G(Z_n; \theta)$ монотонно убывает по θ. Итак, $G(Z_n; \theta)$ — центральная статистика. Пусть $p_1 + p_2 = p$, $p_1 > 0$, $p_2 > 0$. Найдем квантили g_1 и g_2 из соответствующих уравнений $\Phi(g_1) = p_1$

и $\Phi(g_2) = 1 - p_2$. Тогда $\mathbf{P}\left(g_1 \leqslant \dfrac{\sqrt{n}\,(\overline{X}_n - \theta)}{\sigma} \leqslant g_2\right) = q$. Отсюда

$$\mathbf{P}\left(\overline{X}_n - g_2\frac{\sigma}{\sqrt{n}} \leqslant \theta \leqslant \overline{X}_n - g_1\frac{\sigma}{\sqrt{n}}\right) = q. \tag{9.5}$$

Найдем g_1 и g_2 посредством минимизации длины полученного доверительного интервала: $\dfrac{\sigma}{\sqrt{n}}(g_2 - g_1) \to \min$ при условии $\Phi(g_2) - \Phi(g_1) = q$. Для этого рассмотрим функцию Лагранжа

$$\mathcal{L}(g_1, g_2, \lambda) = \frac{\sigma}{\sqrt{n}}(g_2 - g_1) + \lambda\left(\Phi(g_2) - \Phi(g_1) - q\right), \quad \lambda > 0.$$

Найдем стационарные точки функции $\mathcal{L}(g_1, g_2, \lambda)$:

$$\frac{\partial\mathcal{L}(g_1, g_2, \lambda)}{\partial g_1} = -\frac{\sigma}{\sqrt{n}} - \lambda p_G(g_1) = 0,$$

$$\frac{\partial\mathcal{L}(g_1, g_2, \lambda)}{\partial g_2} = \frac{\sigma}{\sqrt{n}} + \lambda p_G(g_2) = 0,$$

где $p_G(x)$ — плотность распределения $\mathcal{N}(0;1)$. Отсюда следует, что $p_G(g_1) = p_G(g_2)$. Так как, $p_G(x) = p_G(-x)$ для всех $x \in \mathbb{R}^1$, то либо $g_1 = g_2$, либо $g_1 = -g_2$. Первый случай не подходит, так как $\Phi(g_2) - \Phi(g_1) = 0 \neq q$. Отсюда заключаем, что $\Phi(g_2) - \Phi(-g_2) = q$. Таким образом, $g_2 = u_\alpha$ — квантиль уровня $\alpha = 1 - \dfrac{p}{2}$, а $g_1 = -u_\alpha$. Подставляя найденные g_1 и g_2 в (9.5), окончательно имеем

$$\mathbf{P}\left(\overline{X}_n - u_\alpha\frac{\sigma}{\sqrt{n}} \leqslant \theta \leqslant \overline{X}_n + u_\alpha\frac{\sigma}{\sqrt{n}}\right) = q. \tag{9.6}$$

Заметим, что из $g_2 = -g_1 = u_\alpha$ следует, что $p_1 = p_2 = \dfrac{p}{2}$. Таким образом, доверительный интервал (9.6) является центральным. ■

Пример 9.2. Дана реализация z_n выборки Z_n объема $n = 9$, порожденной СВ $X \sim \mathcal{N}(\theta;\sigma^2)$: $z_n = \{1{,}23; -1{,}384; -0{,}959; 0{,}731; 0{,}717; -1{,}805; -1{,}186; 0{,}658; -0{,}439\}$. Построить для θ доверительные интервалы надежности $q = 0{,}95$, если а) $\sigma^2 = 1$; б) σ^2 неизвестна.

Решение. а) По условию $p = 1 - q = 0{,}05$, поэтому $\alpha = 1 - \dfrac{p}{2} = 0{,}975$. По таблице 13.2 находим: $u_\alpha = 1{,}96$. По реализации выборки z_n вычисляем реализацию $\overline{x}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} x_k = -0{,}271$ выборочного среднего $\overline{X}_n$. Теперь из (9.6) следует, что искомый интервал $I_1 = \left[\overline{x}_n - u_\alpha\dfrac{\sigma}{\sqrt{n}}; \overline{x}_n + u_\alpha\dfrac{\sigma}{\sqrt{n}}\right]$. Подставляя $\overline{x}_n$, $n = 9$, $\sigma = 1$ и $u_\alpha = 1{,}96$, находим, что $I_1 = [-0{,}924; 0{,}382]$.

б) Теперь дисперсия σ^2 неизвестна. Воспользуемся статистикой $G_n(Z_n;\theta) = \dfrac{\sqrt{n-1}\,(\overline{X}_n - \theta)}{\overline{S}_n}$, которая является центральной. Действительно, по теореме 9.2 $G_n(Z_n;\theta) \sim \mathcal{T}_{n-1}$, а монотонность по θ очевидна. Повторяя практически дословно рассуждения, приведенные в примере 9.1, находим доверительный интервал наименьшей длины:

$$\mathbf{P}\left(\overline{X}_n - t_\alpha\frac{\overline{S}_n}{\sqrt{n-1}} \leqslant \theta \leqslant \overline{X}_n + t_\alpha\frac{\overline{S}_n}{\sqrt{n-1}}\right) = q, \tag{9.7}$$

где t_α — квантиль уровня $\alpha = 0{,}975$ распределения Стьюдента $\mathcal{T}_r$ с $r = n - 1 = 8$ степенями свободы. По таблице 13.4 находим, что $t_\alpha = 2{,}306$.

По реализации z_n вычисляем реализацию $\overline{s}_n^2 = \dfrac{1}{n}\sum\limits_{k=1}^{n}(x_k - \overline{x}_n)^2 = 1{,}115$ выборочной дисерсии $\overline{S}_n^2$. Теперь из (9.7) с учетом $n = 9$, найденного t_α и того, что $\overline{s}_n = 1{,}056$, следует

$$I_2 = \left[\overline{x}_n - t_\alpha\frac{\overline{s}_n}{\sqrt{n-1}} \leqslant \theta \leqslant \overline{x}_n + t_\alpha\frac{\overline{s}_n}{\sqrt{n-1}}\right] = [-1{,}132;\ 0{,}59]$$

Итак, $I_2 = [-1{,}132; 0{,}59]$ — искомый доверительный интервал. Заметим, что выборка, приведенная в условии, в действительности соответствует $\theta_0 = 0$ и $\sigma_0^2 = 1$. Видим, что оба полученных интервала "накрывают" истинное значение θ_0 параметра θ.

Заметим, что I_1 и I_2, конечно, являются лишь *реализациями доверительных интервалов*, соответствующими конкретной реализации z_n выборки Z_n. ■

Пример 9.3. В условиях примера 9.2 построить доверительный интервал нажежности $q = 0{,}95$ для неизвестной дисперсии σ^2.

Решение. Статистика $g_n(\sigma^2) = \dfrac{n\overline{S}_n^2}{\sigma^2}$ является центральной для σ^2, так как $g_n(\sigma^2) \sim \mathcal{H}_{n-1}$ и монотонно убывает по σ^2. Пусть k_α и k_β — квантили χ^2-распределения $\mathcal{H}_{n-1}$ уровней, соответственно, $\alpha = \dfrac{p}{2}$ и $\beta = 1 - \dfrac{p}{2}$. Тогда $\mathbf{P}\left(k_\alpha \leqslant \dfrac{n\overline{S}_n^2}{\sigma^2} \leqslant k_\beta\right) = q$. Отсюда $\mathbf{P}\left(\dfrac{n\overline{S}_n^2}{k_\beta} \leqslant \sigma^2 \leqslant \dfrac{n\overline{S}_n^2}{k_\alpha}\right) = 0{,}95$, если $p = 0{,}05$.

Итак, искомый интервал для σ^2 имеет вид

$$I^2 = \left[\frac{n\overline{S}_n^2}{k_\beta};\ \frac{n\overline{S}_n^2}{k_\alpha}\right].$$

Для $n-1=8$, $\alpha = 0{,}025$, $\beta = 0{,}975$ по таблице 13.3 находим $k_\alpha = 2{,}18$, $k_\beta = 17{,}5$. Реализация интервала I^2 с учетом данных примера 9.2 и того, что $\overline{s}_n^2 = 1{,}115$, имеет вид $\left[\frac{n\overline{s}_n^2}{k_\beta};\ \frac{n\overline{s}_n^2}{k_\alpha}\right] = \left[\frac{8\cdot 1{,}115}{17{,}5};\ \frac{8\cdot 1{,}115}{2{,}18}\right] = [0{,}51; 4{,}09]$.

Для с.к.о. σ соответствующий интервал равен $I = \left[\sqrt{0{,}51}; \sqrt{4{,}09}\right] = [0{,}71;\ 2{,}02]$. Заметим, что истинное значение $\sigma_0 = 1$ накрывается найденным интервалом I. ■

Пример 9.4. Выборка $\{X_k,\ k = 1, \ldots, n\}$, где $n \gg 1$, соответствует распределению $Bi(N;\theta)$, $\theta > 0$. Построить асимптотический доверительный интервал для θ.

Решение. Известно (см. задачу 1 из раздела 8.3), что оценка $\widehat{\theta}_n = \frac{\overline{X}_n}{N}$ эффективна. Так как $\mathbf{M}\{X_k\} = N\theta$, то из центральной предельной теоремы следует: $\sqrt{n}\,(\overline{X}_n - N\theta) \xrightarrow{\mathrm{d}} \xi \sim \mathcal{N}(0; N\theta(1-\theta))$, где $N\theta(1-\theta) = \mathbf{D}\{X_k\}$. Отсюда заключаем, что $\sqrt{n}\,(\widehat{\theta}_n - \theta) \xrightarrow{\mathrm{d}} \eta \sim \mathcal{N}(0; d(\theta))$, где $d(\theta) = \frac{\theta(1-\theta)}{N}$ — асимптотическая дисперсия. Теперь, если u_α — квантиль уровня $\alpha = 1 - \frac{p}{2}$ распределения $\mathcal{N}(0;1)$, то искомый интервал имеет вид

$$\widehat{I}(n) = \left[\widehat{\theta}_n - u_\alpha\sqrt{\frac{\widehat{\theta}_n(1-\widehat{\theta}_n)}{n\,N}};\ \widehat{\theta}_n + u_\alpha\sqrt{\frac{\widehat{\theta}_n(1-\widehat{\theta}_n)}{n\,N}}\right],$$

где $\widehat{\theta}_n = \frac{\overline{X}_n}{N}$. При этом $\mathbf{P}\left(\theta \in \widehat{I}(n)\right) \to q$, $n \to \infty$. ■

9.3. Задачи для самостоятельного решения.

1. Выборка $\{X_1, \ldots, X_n\}$, $n \gg 1$ соответствует распределению Пуассона $\Pi(\theta)$, $\theta > 0$. Построить асимптотический доверительный интервал для θ надежности q.

Ответ. $\left[\overline{X}_n - u_\alpha\sqrt{\frac{\overline{X}_n}{n}};\ \overline{X}_n + u_\alpha\sqrt{\frac{\overline{X}_n}{n}}\right]$, $\alpha = \frac{1+q}{2}$.

2. Выборка $\{X_1, \ldots, X_n\}$ соответствует распределению $\mathcal{N}(\mu;\theta)$, μ — известно. Построить центральный доверительный интервал надежности q для дисперсии θ.

Указание. Показать, что $\frac{\sum_{k=1}^{n}(X_k-\mu)^2}{\theta}$ — центральная статистика с распределением $\mathcal{H}_n$. Воспользоваться примером 9.3.

3. Выборка $\{X_1, \ldots, X_n\}$, $n \gg 1$ соответствует распределению $E\left(\frac{1}{\theta}\right)$, $\theta > 0$. Построить асимптотический доверительный интервал надежности q для параметра θ.

Указание. Воспользоваться МП-оценкой $\widehat{\theta}_n$ для θ.

Ответ. $\left[\left(1 - \frac{u_\alpha}{\sqrt{n}}\right)\overline{X}_n;\ \left(1 + \frac{u_\alpha}{\sqrt{n}}\right)\overline{X}_n\right]$; $\alpha = \frac{1+q}{2}$.

4. По выборке $z_n = \{-0{,}26; -0{,}36; 1{,}83; 0{,}54; -2{,}06\}$, соответствующей распределению $\mathcal{N}(\theta_1;\theta_2^2)$, найти доверительные интервалы надежности $q = 0{,}9$ для θ_1 и θ_2.

Ответ. $[-1{,}41; 1{,}29]$ — для θ_1; $[0{,}94; 3{,}43]$ — для θ_2.

5. Выборка $Z_n = \{X_k,\ k = 1, \ldots, n\}$, $n \gg 1$ соответствует равномерному распределению $R[0;a]$, $a > 0$. Построить асимптотический доверительный интервал надежности q для параметра $\theta = \mathbf{M}\{X_1\}$.

Указание. Использовать оценку $\widehat{\theta}_n = \overline{X}_n$.

Ответ. $\left[\left(1 - \frac{u_\alpha}{\sqrt{3n}}\right)\overline{X}_n;\ \left(1 + \frac{u_\alpha}{\sqrt{3n}}\right)\overline{X}_n\right]$; $\alpha = \frac{1+q}{2}$.

6. Выборка $Z_n = \{X_1, \ldots, X_n\}$, $n \gg 1$ соответствует распределению с плотностью вероятности $p(x;\theta) = \exp\{\theta - x\}$, $x \geqslant \theta$, где $\theta > 0$. Показать, что доверительным интервалом надежности q для θ является $\left[X_{(1)} + \frac{\ln(1-q)}{n};\ X_{(1)}\right]$.

Указание. Показать, что $G(Z_n;\theta) = n(X_{(1)} - \theta)$ — центральная статистика.

7. В условиях задачи 6 построить центральный доверительный интервал надежности q для θ.

Ответ. $\left[X_{(1)} + \frac{1}{n}\ln\left(\frac{1-q}{2}\right);\ X_{(1)} + \frac{1}{n}\ln\left(\frac{1+q}{2}\right)\right]$.

8. Пусть выборка $\{X_1, \ldots, X_n\}$, $n \gg 1$ соответствует распределению $R[0;\theta]$, $\theta > 0$. Показать, что $\left(\frac{X_{(n)}}{\theta}\right)^n$ — центральная статистика, и построить для θ доверительный интервал минимальной длины и надежности q.

Ответ. $\left[X_{(n)};\ \frac{X_{(n)}}{(1-q)^{1/n}}\right]$.

9. Пусть $Z_n^{(1)} = \{X_1, \ldots, X_n\}$ и $Z_n^{(2)} = \{Y_1, \ldots, Y_n\}$ — две независимые выборки, причем $Z_n^{(1)}$ соответствует распределению $\mathcal{N}(\theta_1;\sigma_1^2)$, а $Z_n^{(2)}$ — $\mathcal{N}(\theta_2;\sigma_2^2)$ (σ_1 и σ_2 — известны). Требуется построить доверительный интервал надежности q для параметра $\theta = \theta_1 - \theta_2$.

Указание. Использовать $G(Z_n^{(1)}; Z_n^{(2)}; \theta) = \dfrac{\overline{X}_n - \overline{Y}_n - \theta}{\sigma}$, где $\sigma^2 = \dfrac{\sigma_1^2 + \sigma_2^2}{n}$.

Ответ. $\left[\overline{X}_n - \overline{Y}_n - u_\alpha\sigma;\ \overline{X}_n - \overline{Y}_n + u_\alpha\sigma\right]$, $\alpha = \dfrac{1+q}{2}$.

§ 10. Проверка параметрических гипотез

10.1. Теоретические положения. Пусть СВ X имеет закон распределения, заданный функцией распределения $F(x;\theta)$ или плотностью вероятности $p(x;\theta)$, где θ — некоторый скалярный или векторный параметр.

Определение 10.1. Любое предположение о возможных значениях параметра θ называется *параметрической гипотезой.*

Определение 10.2. Параметрическая гипотеза, состоящая в том, что $\theta = \theta_0$, где θ_0 — фиксированная величина, называется *простой гипотезой.*

Определение 10.3. Параметрическая гипотеза называется *сложной*, если она состоит в том, что $\theta \in \Theta_0$, где Θ_0 — некоторое фиксированное подмножество, принадлежащее множеству Θ возможных значений параметра θ и содержащее более одной точки.

Обычно проверяемая (основная) гипотеза обозначается $H_0 : \theta \in \Theta_0$, а гипотеза, обозначаемая $H_1 : \theta \in \Theta_1$, где $\Theta_1 \in \Theta \setminus \Theta_0$, называется *альтернативой* по отношению к H_0.

Определение 10.4. *Статистическим критерием* называется алгоритм проверки гипотезы H_0 по выборке Z_n, соответствующей распределению $F(x;\theta)$, $\theta \in \Theta$.

Рассмотрим общую структуру статистического критерия. Пусть V_0 — множество всех возможных значений вектора Z_n в предположении, что H_0 — верна. Выберем малое положительное число $p \in (0;1)$ и область $S_p \in V_0$ такую, что

$$P_0(S_p) = \mathbf{P}\left(Z_n \in S_p \mid H_0\text{— верна}\right) = p.$$

Определение 10.5. Число p называется *уровнем значимости* критерия, а множество S_p — *критической областью уровня p.*

Пусть $z_n = \{x_1, \ldots, x_n\}^\top$ — конкретная реализация выборки Z_n. Предположим, что $z_n \in S_p$, тогда гипотеза H_0 *отвергается на уровне значимости p.* Если же $z_n \in \overline{S}_p = V_0 \setminus S_p$, то H_0 — принимается. Область $\overline{S}_p$ обычно называют *доверительной областью.* Очевидно, что $P_0(\overline{S}_p) = 1 - p = q$, где q — *доверительная вероятность* или *надежность* критерия.

Определение 10.6. *Ошибкой первого рода* называется факт отвержения гипотезы H_0 в случае, когда она верна. Принятие гипотезы H_0 при условии, что верна альтернатива H_1, называется *ошибкой второго рода.*

Очевидно, что вероятность ошибки первого рода равна $P_0(S_p) = p$, т.е. совпадает с уровнем значимости критерия.

Вероятность ошибки второго рода имеет вид

$$\beta = P_1(\overline{S}_p) = \mathbf{P}\left(Z_n \in \overline{S}_p \mid H_1\text{— верна}\right).$$

Определение 10.7. Пусть $H_\gamma : \theta = \gamma$. Функция $W(S_p, \gamma) = \mathbf{P}\left(Z_n \in S_p \mid H_\gamma\text{— верна}\right)$ называется *мощностью критерия.*

Если альтернатива H_1 — простая ($\theta = \theta_1$), то вероятность β ошибки второго рода связана с мощностью критерия очевидным соотношением $\beta = 1 - W(S_p, \theta_1)$. Отсюда следует, что "хороший" критерий проверки H_0 должен быть устроен так, чтобы при фиксированном p мощность $W(S_p, \theta_1)$ была бы как можно ближе к единице. Последнее достигается посредством выбора из множества I_p всех возможных критических областей S_p такой области S_p^*, что

$$W(S_p^*, \theta_1) = \max_{S_p \in I_p} W(S_p, \theta_1). \qquad (10.1)$$

Критерий, соответствующий критической области S_p^*, называется *наиболее мощным.* В некоторых частных случаях область S_p^* существует и может быть найдена аналитически (см. далее теорему 10.1).

Обычно на практике критическую область S_p задают неявно с помощью некоторой *статистики критерия* $T(Z_n)$. Пусть Δ_p — область на $\mathbb{R}^1$ такая, что

$$\mathbf{P}\left(T(Z_n) \in \Delta_p \mid H_0\text{— верна}\right) = p.$$

Тогда критическая область S_p определяется так:

$$S_p = \{z : z \in V_0 \text{ и } T(z) \in \Delta_p\}. \qquad (10.2)$$

Как правило, описать явно одномерную область Δ_p существенно проще, чем n-мерную область S_p. Например, $T(z)$ и Δ_p достаточно просто определяются с помощью метода доверительных интервалов (см. пример 10.1).

Рассмотрим общий способ выбора статистики $T(z)$, приводящий к наиболее мощному критерию для проверки простой гипотезы $H_0 : \theta = \theta_0$ против простой альтернативы $H_1 : \theta = \theta_1$.

Пусть Z_n — выборка, соответствующая распределению с плотностью $p(x;\theta)$, где $\theta = \theta_0$ или $\theta = \theta_1$, $\theta_0 \neq \theta_1$, а $z_n = \{x_1, \ldots, x_n\}^\top$ — реализация Z_n. Введем *статистику правдоподобия*:

$$T(z_n) = \frac{\prod\limits_{k=1}^{n} p(x_k;\theta_1)}{\prod\limits_{k=1}^{n} p(x_k;\theta_0)}. \qquad (10.3)$$

Теорема 10.1. *(Нейман-Пирсон). Наиболее мощный критерий для проверки H_0 на уровне значимости p против H_1 существует и задается оптимальной в смысле* (10.1) *критической областью $S_p^* = \{z_n : T(z_n) \geqslant \delta\}$, где $T(z_n)$ имеет вид* (10.3), *а параметр δ определяется из условия*

$$\mathbf{P}\left(T(Z_n) \geqslant \delta \mid H_0 — \textit{верна}\right) = p.$$

Заметим, что в условиях теоремы 10.1 критическая область Δ_p для $T(Z_n)$ имеет простой вид: $\Delta_p = [\delta, +\infty)$.

10.2. Примеры.

Пример 10.1. По выборке $Z_n = \{X_k, k = 1, \ldots, n\}$, соответствующей распределению $\mathcal{N}(\theta;\sigma^2)$, где $\sigma^2 > 0$ — известна, проверить гипотезу H_0: $\theta = \theta_0$ на уровне значимости p против альтернативы H_1: $\theta \neq \theta_0$.

Решение. Для проверки H_0 воспользуемся методом доверительных интервалов. Рассмотрим статистику $T(Z_n) = \overline{X}_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$. Из примера 9.1 следует, что при $\theta = \theta_0$ (т.е. H_0 — верна) и $\alpha = 1 - \dfrac{p}{2}$

$$\mathbf{P}\left(\overline{X}_n - u_\alpha \frac{\sigma}{\sqrt{n}} \leqslant \theta_0 \leqslant \overline{X}_n + u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - p,$$

если u_α — квантиль уровня α распределения $\mathcal{N}(0;1)$. Отсюда

$$\mathbf{P}\left(\overline{X}_n \in \left[\theta_0 - u_\alpha \frac{\sigma}{\sqrt{n}}, \theta_0 + u_\alpha \frac{\sigma}{\sqrt{n}}\right]\right) = 1 - p.$$

Таким образом, критическая область Δ_p для статистики критерия $T(Z_n) = \overline{X}_n$ принимает вид

$$\Delta_p = \left\{x : |x - \theta_0| > u_\alpha \frac{\sigma}{\sqrt{n}}\right\},$$

а доверительная область $\overline{\Delta}_p = \mathbb{R}^1 \setminus \Delta_p = \left\{x : |x - \theta_0| \leqslant u_\alpha \dfrac{\sigma}{\sqrt{n}}\right\}$.

Итак, если $z_n = \{x_1, \ldots, x_n\}^\top$ — реализация выборки Z_n, а $\overline{x}_n = T(z_n) = \dfrac{1}{n}\sum\limits_{k=1}^{n} x_k$ — соответствующая реализация выборочного среднего $\overline{X}_n$ (т.е. статистики критерия), то гипотезу H_0 на уровне значимости p следует отвергнуть, если $\overline{x}_n \in \Delta_p$, т.е. $|\overline{x}_n - \theta_0| > u_\alpha \dfrac{\sigma}{\sqrt{n}}$. Если же $\overline{x}_n \in \overline{\Delta}_p$, то H_0 следует принять. ∎

Пример 10.2. В условиях примера 10.1 вычислить мощность критерия и вероятность ошибки второго рода, если H_1: $\theta = \gamma$, $\gamma \neq \theta_0$.

Решение. По определению 10.7 с учетом (10.2) имеем

$$\begin{aligned}
W(S_p, \gamma) &= \mathbf{P}\left(\overline{X}_n \in \Delta_p \mid H_1 — \text{верна}\right) = \\
&= \mathbf{P}\left(\frac{\sqrt{n}|\overline{X}_n - \theta_0|}{\sigma} > u_\alpha \mid H_1 — \text{верна}\right) = \\
&= 1 - \mathbf{P}\left(\theta_0 - u_\alpha \frac{\sigma}{\sqrt{n}} \leqslant \overline{X}_n \leqslant \theta_0 + u_\alpha \frac{\sigma}{\sqrt{n}} \mid H_1 — \text{верна}\right).
\end{aligned}$$

Если верна альтернатива H_1, то $\overline{X}_n \sim \mathcal{N}\left(\gamma; \dfrac{\sigma^2}{n}\right)$, поэтому

$$\begin{aligned}
W(S_p, \gamma) &= 1 - \left[\Phi\left(\frac{\theta_0 + u_\alpha \frac{\sigma}{\sqrt{n}} - \gamma}{\frac{\sigma}{\sqrt{n}}}\right) - \Phi\left(\frac{\theta_0 - u_\alpha \frac{\sigma}{\sqrt{n}} - \gamma}{\frac{\sigma}{\sqrt{n}}}\right)\right] = \\
&= 1 - \left[\Phi\left(\frac{\sqrt{n}(\theta_0 - \gamma)}{\sigma} + u_\alpha\right) - \Phi\left(\frac{\sqrt{n}(\theta_0 - \gamma)}{\sigma} - u_\alpha\right)\right] = 1 - \varphi_n(\theta_0, \gamma).
\end{aligned}$$

Очевидно, что $W(S_p, \theta_0) = 1 - [\Phi(u_\alpha) - \Phi(-u_\alpha)] = 1 - (1 - p) = p$ — вероятность ошибки первого рода.

По определению 10.7 вероятность ошибки второго рода $\beta = 1 - W(S_p, \gamma) = \varphi_n(\theta_0, \gamma)$.

Сделаем некоторые выводы о зависимости $\varphi_n(\theta_0, \gamma)$ от величины γ и объема выборки n (θ_0 — фиксировано).

1) Если $n = \text{const}$, а $|\theta_0 - \gamma| \to \infty$, то $\varphi_n(\theta_0, \gamma) \to 0$. Поэтому $W(S_p, \theta_0) \to 1$, а $\beta \to 0$. Последнее означает, что при фиксированном объеме выборки n хорошо различаются "далекие" гипотезы H_0 и H_1 (т.е. $|\theta_0 - \gamma| \gg 0$). Если же $\theta_0 \cong \gamma$, то $\beta \cong 1 - W(S_p, \theta_0) = 1 - p$, т.е. близка к единице, так как p мало по условию.

2) Если же $\theta_0 \neq \gamma$, но $n \to \infty$, то $\dfrac{\sqrt{n}|\theta_0 - \gamma|}{\sigma} \to \infty$. Поэтому $\varphi_n(\theta_0, \gamma) \to 0$ при $n \to \infty$, θ_0, γ — фиксированы. Последнее означает, что критерий будет хорошо различать даже "близкие" гипотезы ($\theta_0 \cong \gamma$), если объем выборки n достаточно велик. Указанный факт называется *свойством состоятельности критерия* против любой простой альтернативы H_1.

■

П р и м е р 10.3. По реализации z_n выборки Z_n объема $n = 100$, соответствующей распределению $\mathcal{N}(\theta; 1)$, вычислена реализация выборочного среднего $\overline{x}_n = 0{,}153$. На уровне значимости $p = 0{,}05$ проверить гипотезу H_0: $\theta = 0$ против альтернативы H_1: $\theta = 0{,}5$. Вычислить мощность критерия и вероятность ошибки второго рода β.

Р е ш е н и е. Воспользуемся результатами примеров 10.1 и 10.2. По условию $n = 100$, $\sigma = 1$, $p = 0{,}05$, $\alpha = 1 - \dfrac{p}{2} = 0{,}975$, $u_\alpha = 1{,}96$. Доверительная область $\overline{\Delta}_p$ имеет вид

$$\overline{\Delta}_p = \left[\theta_0 - u_\alpha \frac{\sigma}{\sqrt{n}}; \theta_0 + u_\alpha \frac{\sigma}{\sqrt{n}}\right] = [-0{,}196; 0{,}196],$$

где учтено, что $\theta_0 = 0$ по условию. Так как $\overline{x}_n = 0{,}153 \in \overline{\Delta}_p$, гипотеза H_0 принимается. Заметим, что проводя аналогичные выкладки для гипотезы H_1, мы получили бы доверительную область $\overline{\Delta}_p^{(1)} = [-0{,}196 + 0{,}5; 0{,}196 + 0{,}5] = [0{,}304; 0{,}696]$. Так как $\overline{x}_n \notin \overline{\Delta}_p^{(1)}$, то гипотезу H_1 следует отвергнуть.

Из примера 10.2 следует, что при $\theta_0 = 0$ и $\gamma = 0{,}5$

$$W(S_p, \gamma) = 1 - [\varPhi(5 + 1{,}96) - \varPhi(5 - 1{,}96)] \cong \varPhi(3{,}04) = 0{,}9987.$$

Поэтому вероятность ошибки второго рода весьма мала: $\beta = 1 - W(S_p, \gamma) = 0{,}0013$. ■

П р и м е р 10.4. В условиях примера 10.1 построить наиболее мощный критерий для проверки гипотезы H_0: $\theta = \theta_0$ против альтернативы H_1: $\theta = \gamma$, $\gamma > \theta_0$.

Р е ш е н и е. Воспользуемся теоремой 10.1 Неймана-Пирсона. Статистика критерия (10.3) с учетом гауссовости выборки принимает вид

$$T(z_n) = \frac{\left(\sqrt{2\pi}\sigma\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \gamma)^2\right\}}{\left(\sqrt{2\pi}\sigma\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^n (x_k - \theta_0)^2\right\}} = \\ = \exp\left\{\frac{\sum_{k=1}^n x_k(\gamma - \theta_0)}{\sigma^2} - \frac{n}{2\sigma^2}\left(\gamma^2 - \theta_0^2\right)\right\}.$$

Поэтому неравенство $T(z_n) \geqslant \delta$ эквивалентно $\ln(T(z_n)) \geqslant \ln\delta$, т.е. $\overline{x}_n \geqslant \delta_1$, где $\overline{x}_n = \dfrac{1}{n}\sum_{k=1}^n x_k$, а $\delta_1 = \dfrac{1}{2}(\theta_0 + \gamma) + \dfrac{\sigma^2 \ln\delta}{(\gamma - \theta_0)n}$.

Найдем теперь δ_1 с учетом того, что $\overline{X}_n \sim \mathcal{N}\left(\theta_0; \dfrac{\sigma^2}{n}\right)$, если H_0 — верна. Из теоремы 10.1 следует:

$$p = \mathbf{P}\left(T(Z_n) \geqslant \delta \mid H_0\text{— верна}\right) = \mathbf{P}\left(\overline{X}_n \geqslant \delta_1 \mid H_0\text{— верна}\right) = \\ = 1 - \varPhi\left(\frac{\sqrt{n}(\delta_1 - \theta_0)}{\sigma}\right).$$

Отсюда следует, что $\varPhi\left(\dfrac{\sqrt{n}(\delta_1 - \theta_0)}{\sigma}\right) = 1 - p$, т.е. $\dfrac{\sqrt{n}(\delta_1 - \theta_0)}{\sigma} = u_\alpha$, где u_α — квантиль уровня $\alpha = 1 - p$ распределения $\mathcal{N}(0; 1)$. Таким образом, $\delta_1 = \theta_0 + u_\alpha \dfrac{\sigma}{\sqrt{n}}$. Итак, если реализация выборочного среднего $\overline{x}_n$ удовлетворяет неравенству $\overline{x}_n \geqslant \theta_0 + u_\alpha \dfrac{\sigma}{\sqrt{n}}$, то гипотезу H_0 следует отвергнуть. В заключение заметим, что граница δ_1 зависит от θ_0, но не зависит от конкретного значения γ (учтено лишь, что $\gamma > \theta_0$). ■

10.3. Задачи для самостоятельного решения.

1. Обобщить результат примера 10.1 на случай, когда дисперсия σ^2 неизвестна.

Указание. Смотри примеры 9.2 (б) и 10.1.

Ответ. H_0 принимается, если $|\overline{x}_n - \theta_0| \leqslant t_\alpha \sqrt{\dfrac{\overline{s}_n^2}{n-1}}$, где t_α — квантиль уровня $\alpha = 1 - \dfrac{p}{2}$ распределения Стьюдента с $n-1$ степенью свободы, $\overline{s}_n^2$ — реализация выборочной дисперсии $\overline{S}_n^2$.

2. Выборка Z_n соответствует распределению $\mathcal{N}(\theta_1; \theta_2)$. Проверить на уровне значимости p гипотезу H_0: $\theta_2 = \sigma^2$ против $H_1 : \theta_2 \neq \sigma^2$.

Указание. Смотри пример 9.3.

Ответ. H_0 принимается, если $\overline{s}_n^2 \in \left[k_1 \dfrac{\sigma^2}{n}; k_2 \dfrac{\sigma^2}{n}\right]$, где k_1 и k_2 — квантили распределения $\mathcal{H}_{n-1}$ уровней $\dfrac{p}{2}$ и $1 - \dfrac{p}{2}$ соответственно.

3. Пусть $\{X_k, k = 1, \ldots, n\}$ и $\{Y_m, m = 1, \ldots, n\}$ — независимые выборки, порожденные СВ $X \sim \mathcal{N}(\theta_1; \sigma_1^2)$ и $Y \sim \mathcal{N}(\theta_2; \sigma_2^2)$, σ_1 и σ_2 — известны. На уровне значимости p проверить гипотезу $H_0 : \theta_1 = \theta_2$ против $H_1 : \theta_1 \neq \theta_2$.

Указание. Смотри задачу 9 из раздела 9.3.

Ответ. H_0 принимается, если $|\overline{x}_n - \overline{y}_n| \leqslant u_\alpha \sqrt{\dfrac{\sigma_1^2 + \sigma_2^2}{n}}$, u_α — квантиль распределения $\mathcal{N}(0; 1)$ уровня $\alpha = 1 - \dfrac{p}{2}$; $\overline{x}_n$ и $\overline{y}_n$ — реализации $\overline{X}_n$ и $\overline{Y}_n$.

4. В условиях примера 10.4 найти вероятность β ошибки второго рода.

Ответ. $\beta = \varPhi\left(\sqrt{n}\dfrac{(\theta_0 - \gamma)}{\sigma} + u_\alpha\right)$.

5. В условиях предыдущей задачи определить, при каком минимальном объеме выборки n_0 величина β будет не больше 0,01, если $\theta_0 = 0$, $\gamma = 1$, $\sigma^2 = 1$, $p = 0{,}05$.

Ответ. $n_0 = 9$.

6. В условиях примера 10.4 найти минимальный объем выборки n_0, при котором вероятности ошибок первого и второго рода не больше заданных значений, соответственно, $a > 0$ и $b > 0$.

Ответ. $n_0 = \left\{\dfrac{\sigma^2(u_a + u_b)}{(\theta_0 - \gamma)^2}\right\} + 1$, где $\{\cdot\}$ — целая часть числа, u_a и u_b — квантили распределения $\mathcal{N}(0; 1)$ уровней a и b соответственно.

7. Выборка Z_n соответствует распределению $\mathcal{N}(0; \theta^2)$, $\theta > 0$. Построить наиболее мощный критерий для проверки на уровне значимости p гипотезы $H_0 : \theta = \theta_0$ против альтернативы $H_1 : \theta = \sigma > \theta_0$.

Ответ. H_0 отвергается, если $\sum\limits_{k=1}^{n} x_k^2 \geqslant k_\alpha \theta_0^2$, где k_α — квантиль уровня $\alpha = 1 - p$ распределения $\mathcal{H}_{n-1}$.

§ 11. Проверка непараметрических гипотез

11.1. Теоретические положения. Пусть выборка $Z_n = \{X_k, k = 1, \ldots, n\}$ соответствует некоторому распределению с неизвестной функцией распределения $F(x)$.

Определение 11.1. *Непараметрической гипотезой* называется предположение H_0: $F(x) = F_0(x)$, где $F_0(x)$ — некоторая известная функция распределения.

В непараметричемкой гипотеы H_0 функция $F_0(x)$ может содержать неизвестные параметры.

Пусть $F_0(x)$ задает распределение дискретной СВ X, принимающей значения $\{\xi_1, \ldots, \xi_K\}$ с вероятностями $p_m = \mathbf{P}(X = \xi_m)$, $m = 1, \ldots, K$. Построим по выборке Z_n частоты $\{p_m^*,\ m = 1, \ldots, K\}$ появления соответствующих значений: $p_m^* = \dfrac{n_m}{n}$, где n_m — случайное число элементов выборки, равных ξ_m, $m = 1, \ldots, K$, а n — объем выборки.

Определение 11.2. *Статистикой хи-квадрат* (статистикой К. Пирсона) называется функция от выборки Z_n вида

$$G_n = T(Z_n) = n \sum_{m=1}^{K} \frac{(p_m - p_m^*)^2}{p_m} = \sum_{m=1}^{K} \frac{(n_m - np_m)^2}{np_m}. \tag{11.1}$$

Из (11.1) видно, что G_n является некоторой мерой расхождения между теоретическими вероятностями $\{p_m\}$ и их оценками $\{p_m^*\}$, построенными по результатам наблюдений.

На практике G_n удобно вычислять по формуле, следующей из (11.1):

$$G_n = \sum_{m=1}^{K} \frac{(n_m)^2}{np_m} - n. \tag{11.2}$$

Теорема 11.1. *Если Z_n действительно соответствует распределению $F_0(x)$ (т.е. H_0 — верна), то G_n асимптотически распределена по закону хи-квадрат с $r = K - 1$ степенью свободы, т.е.*

$$G_n \xrightarrow{\mathrm{d}} \chi_r^2 \sim \mathcal{H}_r,\ n \to \infty.$$

Теорема 11.1 является теоретическим обоснованием следующего критерия проверки гипотезы H_0: $F(x) = F_0(x)$ на некотором уровне значимости $p \in (0; 1)$.

1) По реализации z_n выборки Z_n вычисляем соответствующую реализацию g_n статистики G_n из (11.2).

2) По таблице 13.3 находим квантиль $k_\alpha(r)$ уровня $\alpha = 1 - p$ распределения $\chi^2(r)$, где $r = K - 1$.

3) Строим критическую $\Delta_p = (k_\alpha(r), +\infty)$ и доверительную $\overline{\Delta}_p = [0; k_\alpha(r)]$ области.

4) Если $g_n \in \Delta_p$, то H_0 отвергается на уровне значимости p. Если же $g_n \in \overline{\Delta}_p$, то на указанном уровне значимости H_0 принимается.

Описанный алгоритм называется *хи-квадрат критерием (критерием согласия К. Пирсона)* для проверки непараметрической гипотезы H_0. На практике обычно полагают, что $p = 0{,}05$.

Если $F_0(x)$ задает распределение СВ X непрерывного типа, то проверка H_0 по хи-квадрат критерию проводится после предварительной *группировки данных*. Для этого область V_X всех возможных значений СВ X, имеющей распределение $F_0(x)$, разбивается на $K > 1$ непересекающихся интервалов $\{\delta_m : m = 1, \ldots, K\}$: $\bigcup\limits_{m=1}^{K} \delta_m = V_X$, $\delta_m \cap \delta_i = \varnothing$, $m \neq i$. Статистика G_n по-прежнему вычисляется по формуле (11.2), где n_m — число элементов выборки, попавших в промежуток δ_m, а p_m — теоретическая вероятность попадания X (т.е. любого элемента выборки) в δ_m (при условии, что H_0 — верна). Например, если $\delta_m = [a_m; a_{m+1})$, то $p_m = F_0(a_{m+1}) - F_0(a_m)$, $m = 1, \ldots, K$.

Далее проверка H_0 проводится точно так же, как и в случае дискретного распределения $F_0(x)$.

Если $F_0(x) = f(x;\theta)$, где θ — вектор неизвестных параметров размера $(l \times 1)$, то в критерии Пирсона следует предварительно заменить θ на его МП-оценку $\widehat{\theta}_n$. В этом случае предельное распределение статистики G_n будет (при весьма общих условиях) по-прежнему хи-квадрат распределением $\mathcal{H}_r$, но уже с числом степеней свободы $r = K - l - 1$, где l — число неизвестных параметров распределения $F_0(x)$.

На практике критерием Пирсона можно пользоваться, если объем выборки достаточно велик ($n \geqslant 50$), а при проведении группировки данных соблюдено условие $n \cdot p_m \geqslant 5$, $m = 1, \ldots, K$. При выборе числа K интервалов группировки можно воспользоваться *формулой Стерджеса*: $K = 1 + \{3{,}32 \lg n\}$, где $\{a\}$ — целая часть числа a.

Хи-квадрат критерием можно воспользоваться так же для проверки непараметрической *гипотезы H_0 о независимости случайных величин X и Y* дискретного типа.

Пусть СВ X принимает значения $\{a_1, \ldots, a_s\}$ с вероятностями $p_i = \mathbf{P}(X = a_i)$, $i = 1, \ldots, s$, а СВ Y — значения $\{b_1, \ldots, b_t\}$ с вероятностями $q_j = \mathbf{P}(X = b_j)$, $j = 1, \ldots, t$. Предположим, что имеется двумерная выборка Z_n, порожденная случайным вектором $W = \{X, Y\}^\top$. Обозначим через n_{ij} число опытов, в которых W принимал значение $\{a_i, b_j\}$, $i = 1, \ldots, s$; $j = 1, \ldots, t$. Пусть также $N_i = \sum\limits_{j=1}^{t} n_{ij}$, $M_j = \sum\limits_{i=1}^{s} n_{ij}$.

Рассмотрим статистику хи-квадрат, аналогичную (11.1):

$$\widehat{G}_n = \sum_{i=1}^{s} \sum_{j=1}^{t} \frac{(n_{ij} - np_i^* q_j^*)^2}{np_i^* q_j^*}, \tag{11.3}$$

где $p_i^* = \dfrac{N_i}{n}$ и $q_j^* = \dfrac{M_j}{n}$ — частоты появления a_i и b_j соответственно. Если гипотеза H_0 о независимости X и Y верна, то можно показать, что

$$\widehat{G}_n \xrightarrow{\text{d}} \chi_r^2 \sim \mathcal{H}_r, \ n \to \infty,$$

где $r = (s-1)(t-1)$.

Таким образом, H_0 отвергают, если реализация $\widehat{g}_n$ статистики $\widehat{G}_n$ попадает в критическую область $\Delta_p = (k_\alpha(r), +\infty)$, где $\alpha = 1 - p$, p — уровень значимости критерия.

Заметим, что $\widehat{G}_n$ удобнее вычислять по формуле

$$\widehat{G}_n = n \left(\sum_{i=1}^{s} \sum_{j=1}^{t} \frac{(n_{ij})^2}{N_i M_j} - 1 \right). \tag{11.4}$$

Для случая $s = t = 2$, часто встречающегося на практике, формула (11.4) принимает весьма простой вид

$$\widehat{G}_n = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{N_1 N_2 M_1 M_2}. \tag{11.5}$$

11.2. Примеры.

Пример 11.1. Монета подброшена $n = 4000$ раз, в результате чего "герб" выпал 2028 раз. Проверить на уровне значимости $p = 0{,}05$ гипотезу H_0: монета симметрична.

Решение. Свяжем с k-м подбрасыванием монеты случайную величину X_k: $\{X_k = 1\}$ означает, что выпал "герб", $\{X_k = 0\}$ — выпала "решка". Очевидно, что $X_k \sim Bi(1; p_1)$, где $p_1 = \mathbf{P}(X_k = 1)$ — вероятность выпадения "герба". Если гипотеза H_0 справедлива, то $p_1 = \frac{1}{2} = p_2$, где $p_2 = 1 - p_1$ — вероятность выпадения "решки". Отсюда $np_1 = np_2 = 2000$.

По условию $n = 4000$, $n_1 = 2028$, $n_2 = n - n_1 = 1972$. Вычислим реализацию g_n статистики G_n по формуле (11.2):

$$g_n = \frac{n_1^2}{np_1} + \frac{n_2^2}{np_2} - n = \frac{2028^2}{2000} + \frac{1972^2}{2000} - 4000 = 0{,}784.$$

Критическая область имеет вид $\Delta_p = (k_\alpha(r), +\infty)$, где $k_\alpha(r) = k_{0,95}(1) = 3{,}84$ (здесь $\alpha = 1 - p = 0{,}95$, $r = K - 1 = 1$, так как $K = 2$ по условию). Итак, $\Delta_p = (3{,}84, +\infty)$. Так как $g_n = 0{,}784 \notin \Delta_p$, то H_0 принимается на уровне значимости $p = 0{,}05$. ■

Пример 11.2. В таблице 11.1 приведена реализация выборки объема $n = 50$, соответствующей некоторому распределению $F(x)$. На уровне значимости $p = 0{,}05$ проверить гипотезу H_0: $F(x) = R[0; 100]$.

Таблица 11.1

65	48	11	76	74	17	46	85	9	50
37	77	69	74	45	80	12	43	56	35
72	70	80	15	45	31	82	23	74	45
35	9	98	17	77	39	27	72	14	43
55	2	10	19	69	91	58	17	73	23

Решение. Если H_0 верна, то выборка порождена СВ $X \sim R[0; 100]$, множество V_X допустимых значений которой есть $[0; 100]$. Разобьем V_X на $K = 5$ подынтервалов $\{\delta_m, m = 1, \ldots, K\}$ одинаковой длины и проведем группировку выборки, результаты которой приведены в таблице 11.2.

Таблица 11.2

m	1	2	3	4	5
δ_m	$[0, 20)$	$[20, 40)$	$[40, 60)$	$[60, 80)$	$[80, 100]$
n_m	12	8	11	13	6

Так как все подынтервалы δ_m одинаковой длины $L = 20$, а $K = 5$, то $p_m = \mathbf{P}(X \in \delta_m) = \dfrac{20}{100} = 0{,}2$ (при условии, что $X \sim R[0; 100]$). Используя результаты, приведенные в таблице 11.2, вычислим g_n по формуле (11.2) с учетом того, что $np_m = 50 \cdot 0{,}2 = 10$, $m = 1, \ldots, K$:

$$g_n = \sum_{m=1}^{K} \frac{(n_m)^2}{np_m} - n = \left[\frac{144 + 64 + 121 + 169 + 36}{10}\right] - 50 = 3{,}4.$$

Критическая область $\Delta_p = (k_\alpha(r), +\infty) = (k_{0,95}(4), +\infty)$, так как $\alpha = 1 - p = 0{,}95$, $r = K - 1 = 4$ (тестируемое распределение не содержит неизвестных параметров). Квантиль $k_{0,95}(4) = 9{,}49$, поэтому $\Delta_p = (9{,}49, +\infty)$ — критическая область, а $\overline{\Delta}_p = [0, 9{,}49]$ — доверительная область.

Так как $g_n = 3{,}4 \in \overline{\Delta}_p$, то H_0: $X \sim R[0; 100]$ принимается на 5%-ном уровне значимости. ■

Пример 11.3. В условиях примера 11.2 проверить гипотезу H_0: выборка соответствует гауссовскому распределению.

Решение. Если H_0 верна, то $X \sim \mathcal{N}(m_X; D_X)$. Заметим, что гипотетическое распределение содержит $l = 2$ неизвестных параметра, так как точные значения m_X и D_X не заданы. По выборке (таблица 11.1) вычислим соответствующие реализации оценок максимального правдоподобия для m_X и D_X (см. пример 7.4):

$$\widehat{m}_X = \overline{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k = 47{,}9 \quad \text{(выборочное среднее)};$$

$$\widehat{D}_X = \overline{s}_n^2 = \frac{1}{n}\sum_{k=1}^{n} (x_k - \widehat{m}_X)^2 = 698{,}02 \quad \text{(выборочная дисперсия)};$$

$$\widehat{\sigma}_X = \sqrt{\widehat{D}_X} = \overline{s}_n = 26{,}42.$$

В силу гауссовости СВ X множество ее значений $V_X = (-\infty, +\infty)$, поэтому подынтервал δ_1 необходимо расширить до $(-\infty, 20)$, а δ_5 — до $[80, +\infty)$. Пусть $\delta_m = [a_m, a_{m+1})$, тогда оценку $\widehat{p}_m$ вероятности попадания СВ X в δ_m можно вычислить по следующей формуле:

$$\widehat{p}_m = \varPhi\left(\frac{a_{m+1} - \overline{x}_n}{\overline{s}_n}\right) - \varPhi\left(\frac{a_m - \overline{x}_n}{\overline{s}_n}\right) = \varPhi\left(\frac{a_{m+1} - 47{,}9}{26{,}42}\right) - \varPhi\left(\frac{a_m - 47{,}9}{26{,}42}\right),$$

где $\varPhi(x)$ — функция Лапласа, $m = 1, \ldots, K$. Результаты соответствующих вычислений приведены в таблице 11.3.

Таблица 11.3

δ_m	$(-\infty, 20)$	$[20, 40)$	$[40, 60)$	$[60, 80)$	$[80, +\infty)$
$n\widehat{p}_m$	20,25	14,5	10,15	4,04	1,06
n_m	12	8	11	13	6

Вычислим реализацию статистики критерия g_n по формуле (11.2), заменяя p_m на $\widehat{p}_m$:

$$g_n = \sum_{m=1}^{K} \frac{(n_m)^2}{n\widehat{p}_m} - n = \left[\frac{144}{20{,}25} + \frac{64}{14{,}5} + \frac{121}{10{,}15} + \frac{169}{4{,}04} + \frac{36}{1{,}06}\right] - 50 = 49{,}24.$$

Найдем критическую область Δ_p для $p = 0{,}05$. Так как $K = 5$, а $l = 2$, то число степеней свободы $r = K - l - 1 = 2$. Отсюда $\Delta_p = (k_\alpha(r), +\infty) = (k_{0{,}95}(2), +\infty) = (5{,}99, +\infty)$. Видим, что $g_n = 49{,}24 \in \Delta_p$, поэтому H_0 отвергается на уровне значимости $p = 0{,}05$. ■

Пример 11.4. Имеются две группы данных о приеме в ВУЗы: A — абитуриент принят ($\overline{A}$ — не принят), B — абитуриент мужчина ($\overline{B}$ — женщина). Всего абитуриентов $n = 442$, а суммарные данные о сочетании признаков в обследуемой группе приведены в таблице 11.4:

Таблица 11.4

	B	$\overline{B}$
A	97	40
$\overline{A}$	263	42

Проверить на уровне значимости $p = 0{,}05$ гипотезу H_0: признаки A и B независимы.

Решение. С признаком A свяжем СВ X следующим образом: $\{X = 1\}$ — абитуриент принят, $\{X = 0\}$ — не принят. Аналогично, $\{Y = 1\}$ — абитуриент мужчина, $\{Y = 0\}$ — женщина. Таким образом, проверка независимости признаков A и B эквивалентна проверки независимости СВ X и Y.

Из таблицы 11.4 следует, что $n_{11} = 97$, $n_{12} = 40$, $n_{21} = 263$ и $n_{22} = 42$. Отсюда $N_1 = n_{11} + n_{12} = 137$, $N_2 = n_{21} + n_{22} = 305$, $M_1 = 97 + 263 = 360$, $M_2 = 40 + 42 = 82$. В данном примере X и Y принимают только два значения, т.е. $s = t = 2$, поэтому реализацию $\widehat{g}_n$ статистики хи-квадрат $\widehat{G}_n$ можно вычислить по формуле (11.5):

$$\widehat{g}_n = \frac{n(n_{11}n_{22} - n_{12}n_{21})^2}{N_1 N_2 M_1 M_2} = \frac{442(97 \cdot 42 - 40 \cdot 263)^2}{137 \cdot 305 \cdot 360 \cdot 82} = 14{,}89.$$

Если H_0 верна, то можно считать, что $\widehat{g}_n$ — реализация СВ $\chi^2(r)$, где $r = (s-1)(t-1) = 1$. Поэтому для $\alpha = 0{,}95$ критическая область $\Delta_p = (k_\alpha(r), +\infty) = (3{,}84, +\infty)$. Так как $\widehat{g}_n \in \Delta_p$, то H_0 следует отвергнуть. Таким образом, имеющиеся статистические данные свидетельствует в пользу того, что признаки A и B зависимы. ■

11.3. Задачи для самостоятельного решения.

1. Среди $n = 10000$ новорожденных детей мальчиков оказалось 5140. С помощью хи-квадрат критерия на уровне значимости $p = 0{,}05$ проверить гипотезу H_0: вероятности рождения мальчика и девочки одинаковы.

Ответ. H_0 отвергается.

2. Результаты группировки реализаций выборки, порожденной некоторой СВ X, приведены в таблице:

δ_m	$(-\infty; -2)$	$[-2; -1)$	$[-1; 0)$	$[0; 1)$	$[1; 2)$	$[2; +\infty)$
n_m	24	132	351	335	140	18

На уровне значимости $p = 0{,}05$ проверить гипотезу H_0: $X \sim \mathcal{N}(0; 1)$.

Ответ. H_0 принимается.

3. В таблице случайных чисел на 1000 знаков цифры $0, 1, \ldots, 9$ появлялись следующее число раз:

0	1	2	3	4	5	6	7	8	9
90	105	112	97	108	101	93	87	103	104

С помощью критерия Пирсона проверить на 5%-ом уровне значимости гипотезу H_0: цифры появляются равновероятно.

Ответ. H_0 принимается.

4. Среди 2020 семей, имеющих двух детей, 527 имеют двух мальчиков, а 476 — двух девочек. Пусть СВ X — количество мальчиков в двухдетной семье. На уровне значимости $p = 0{,}05$ проверить гипотезу H_0: $X \sim Bi(2; \theta)$.

Ответ. $\widehat{\theta}_n = 0{,}513$; H_0 принимается.

5. В таблице приведены 818 случаев, классифицированных по двум признакам: A — наличию прививки против холеры и B — отсутствию заболевания:

	B	$\overline{B}$
A	276	3
$\overline{A}$	473	66

С помощью хи-квадрат критерия на уровне значимости $p = 0{,}05$ проверить гипотезу H_0: признаки A и B независимы.

Ответ. H_0 отклоняется.

§ 12. Метод наименьших квадратов

12.1. Теоретические положения. Пусть $Z_n = \{X_k, k = 1, \ldots, n\}$ — *неоднородная гауссовская выборка*, элементы которой заданы соотношениями

$$X_k = h^\top(t_k)\theta + \varepsilon_k, \quad k = 1, \ldots, n, \tag{12.1}$$

где $\theta \in \mathbb{R}^p$ — вектор неслучайных неизвестных параметров; $h(t) = \{h_1(t), \ldots, h_p(t)\}^\top$ — вектор известных функций; ε_k — гауссовская СВ с параметрами $(0; \sigma^2)$, $\sigma > 0$; t_k — момент времени, соответствующей наблюдению X_k; n — общее число наблюдений. Будем предполагать также, что СВ $\{\varepsilon_k, k = 1, \ldots, n\}$ независимы в совокупности.

Введем обозначения: H_n — матрица размера $(n \times p)$, k-й строкой которой является $h^\top(t_k)$; $E_n = \{\varepsilon_1, \ldots, \varepsilon_n\}^\top$. Тогда (12.1) можно представит в виде

$$Z_n = H_n\theta + E_n. \tag{12.2}$$

Определение 12.1. Модель (12.2), удовлетворяющая всем сделанным выше предположениям, называется *линейной гауссовской регрессионной моделью* порядка p.

Уравнения (12.1) и (12.2) можно трактовать как математическую модель процесса наблюдения “*полезного сигнала*” $\varphi(t;\theta) = h^\top(t)\theta$ в дискретные моменты времени $\{t_k\}$ со случайными аддитивными *ошибками наблюдения* $\{\varepsilon_k, k = 1, \ldots, n\}$.

Рассмотрим задачу оценивания по выборке Z_n вектора $Y = L\theta$, где L — заданная неслучайная матрица размера $(q \times p)$. Для построения оценки $\widehat{Y}_n$ вектора Y воспользуемся методом максимального правдоподобия. По предположению вектор наблюдений Z_n в модели (12.2) имеет гауссовское распределение:

$$Z_n \sim \mathcal{N}(H_n\theta; \sigma^2 I). \tag{12.3}$$

Теорема 12.1. *Пусть матрица $W_n = H_n^\top H_n$ — невырожденная. Тогда МП-оценка $\widehat{Y}_n$ вектора Y по выборке Z_n имеет вид*

$$\widehat{Y}_n = L\widehat{\theta}_n, \tag{12.4}$$

где $\widehat{\theta}_n$ — оценка вектора θ вида

$$\widehat{\theta}_n = W_n^{-1}H_n^\top Z_n. \tag{12.5}$$

Из доказательства теоремы 12.1 (см. пример 12.1) следует, что $\widehat{\theta}_n$ есть точка минимума квадратичной функции потерь:

$$\widehat{\theta}_n = \arg\min_\theta |Z_n - H_n\theta|^2. \tag{12.6}$$

Определение 12.2. Оценка $\widehat{\theta}_n$ вектора, определяемая из условия (12.6) и имеющая вид (12.5), называется *оценкой метода наименьших квадратов* (МНК-оценкой) вектора θ в модели линейной регрессии (12.2).

Статистические свойства ошибки $\Delta\widehat{Y}_n = \widehat{Y}_n - Y$ оценки $\widehat{Y}_n$ вектора Y приведены в следующей теореме.

Теорема 12.2. *Оценка $\widehat{Y}_n$ и ее ошибка $\Delta\widehat{Y}_n$ обладают следующими свойствами:*

1) $\mathbf{M}\Big\{\widehat{Y}_n\Big\} = Y$, $\mathbf{M}\Big\{\Delta\widehat{Y}_n\Big\} = 0$ — *несмещенность;*

2) $\widehat{Y}_n \sim \mathcal{N}\Big(Y; \widehat{K}_{Y_n}\Big)$, $\Delta\widehat{Y}_n \sim \mathcal{N}\Big(0; \widehat{K}_{Y_n}\Big)$, *где* $\widehat{K}_{Y_n} = \mathbf{cov}\{\Delta\widehat{Y}_n, \Delta\widehat{Y}_n\} = \sigma^2 L W_n^{-1} L^\top$ — *ковариационная матрица ошибки $\Delta\widehat{Y}_n$ оценки $\widehat{Y}_n$;*

3) *оценка $\widehat{Y}_n$ эффективна по Рао-Крамеру.*

Так как при $L = I$ оценка $\widehat{Y}_n$ превращается в $\widehat{\theta}_n$, то справедливо следующее утверждение.

Следствие 12.1. *Оценка $\widehat{\theta}_n$ вида* (12.5) *имеет распределение $\mathcal{N}\Big(\theta; \widehat{K}_n\Big)$, где $\widehat{K}_n = \sigma^2 W_n^{-1}$ — ковариационная матрица ее ошибки $\Delta\widehat{\theta}_n = \widehat{\theta}_n - \theta$. МНК-оценка $\widehat{\theta}_n$ — несмещенная и эффективная по Рао-Крамеру.*

Следствие 12.2. *СВ $\xi = \dfrac{1}{\sigma^2}|Z_n - H_n\widehat{\theta}_n|^2$ имеет хи-квадрат распределение с $r = n - p$ степенями свободы: $\xi \sim \mathcal{H}_{n-p}$.*

Если отказаться от предположения о том, что вектор ошибок E_n — гауссовский, то $\widehat{Y}_n$ теряет свойство эффективности, но остается наилучшей (по с.к.-критерию) среди всех линейных несмещенных оценок.

Определение 12.3. Оценка $\widetilde{Y}_n$ вектора Y по наблюдениям Z_n называется *линейной несмещенной оценкой*, если она имеет вид $\widetilde{Y}_n = A_n Z_n$, где A_n — неслучайная матрица размера $(q \times n)$, причем $\mathbf{M}\Big\{\widetilde{Y}_n\Big\} = Y$.

Теорема 12.3 (Гаусс-Марков). *Пусть $\widehat{K}_{Y_n}$ — ковариационна матрица ошибки оценки $\widehat{Y}_n$ вида* (12.4), (12.5), *а $\widetilde{K}_{Y_n}$ — ковариационная матрица произвольной линейной несмещенной оценки $\widetilde{Y}_n$ вектора Y. Тогда $\widehat{K}_{Y_n} \leqslant \widetilde{K}_{Y_n}$.*

Таким образом, оценка $\widehat{Y}_n$ является *наилучшей линейной несмещенной оценкой* вектора Y в модели линейной регрессии (12.2) (НЛН-оценкой).

Пусть теперь вектор ошибок наблюдений E_n в модели (12.2) имеет нулевое среднее $\mathbf{M}\{E_n\} = 0$ и произвольную невырожденную ковариационную матрицу $\mathbf{cov}\,(E_n, E_n) = V_n > 0$. Соответствующую модель наблюдения обычно называют *обобщенной линейной регрессионной моделью.*

Т е о р е м а 12.4. *Если матрица $W_n = H_n^\top H_n$ и матрица ковариаций V_n ошибок наблюдений невырождены, то НЛН-оценка $\widehat{Y}_n$ вектора Y имеет вид* (12.4)*, где*

$$\widehat{\theta}_n = \left(H_n^\top V_n^{-1} H_n\right)^{-1} H_n^\top V_n^{-1} Z_n. \qquad (12.7)$$

$\widehat{\theta}_n$ является НЛН-оценкой для θ, а ковариационная матрица ее ошибки $\Delta\widehat{\theta}_n = \widehat{\theta}_n - \theta_n$ равна

$$\widehat{K}_n = \left(H_n^\top V_n^{-1} H_n\right)^{-1}. \qquad (12.8)$$

О п р е д е л е н и е 12.4. Оценка (12.7) называется *оценкой обобщенного метода наименьших квадратов* (ОМНК-оценкой).

Заметим, что при дополнительном условии $E_n \sim \mathcal{N}(0; V_n)$ оценка $\widehat{\theta}_n$, определенная в теореме 12.4, является МП-оценкой, эффективной по Рао-Крамеру.

12.2. Примеры.

П р и м е р 12.1. Доказать теорему 12.1.

Р е ш е н и е. Найдем МП-оценку вектора θ. По условию закон распределения наблюдения X_k имеет плотность

$$p_k(x;\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x - h^\top(t_k)\theta)^2}{2\sigma^2}\right\}.$$

Поэтому функция правдоподобия выборки Z_n принимает вид (см. определение 7.2)

$$L_n(\theta; Z_n) = \prod_{k=1}^{n} p_k(X_k;\theta) = (2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(X_k - h^\top(t_k)\theta)^2\right\}.$$

Последнее выражение, используя матричные обозначения в (12.2), можно представить следующим образом:

$$L_n(\theta; Z_n) = (2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}|Z_n - H_n\theta|^2\right\}. \qquad (12.9)$$

Из (12.9) следует, что задача максимизации $L_n(\theta; Z_n)$ по θ эквивалентна минимизации по θ функции $\mathcal{L}_n(\theta) = |Z_n - H_n\theta|^2$. Таким образом, МП-оценку $\widehat{\theta}_n$ можно определить из условия

$$\widehat{\theta}_n = \arg\min_{\theta} \mathcal{L}_n(\theta).$$

Найдем явный вид оценки $\widehat{\theta}_n$.

$$\mathcal{L}_n(\theta) = |Z_n - H_n\theta|^2 = Z_n^\top Z_n - 2\theta^\top H_n^\top Z_n + \theta^\top H_n^\top H_n\theta.$$

Вычислим $\dfrac{\partial\mathcal{L}_n(\theta)}{\partial\theta}$, воспользовавшись следующими правилами матричного дифференцирования: $\dfrac{\partial(\theta^\top A)}{\partial\theta} = A$; $\dfrac{\partial(\theta^\top A\theta)}{\partial\theta} = 2A\theta$, если $A = A^\top$.

Итак, $\dfrac{\partial\mathcal{L}_n(\theta)}{\partial\theta} = -2H_n^\top Z_n + 2H_n^\top H_n\theta$.

Необходимое условие экстремума функции $\mathcal{L}_n(\theta)$ дает нам следующие соотношения:

$$\frac{\partial\mathcal{L}_n(\theta)}{\partial\theta} = 0 \quad \Longrightarrow H_n^\top H_n\theta = H_n^\top Z_n.$$

Так как матрица $W_n = H_n^\top H_n > 0$ по условию, полученная система уравнений имеет единственное решение:

$$\widehat{\theta}_n = W_n^{-1} H_n^\top Z_n.$$

Заметим, что функция $\mathcal{L}_n(\theta)$ строго выпукла по θ, поэтому $\widehat{\theta}_n$ — единственная точка минимума $\mathcal{L}_n(\theta)$, и, следовательно, единственная МП-оценка вектора θ.

Так как $Y = L\theta$, то МП-оценка для Y имеет вид $\widehat{Y} = L\widehat{\theta}_n$, что непосредственно следует из принципа инвариантности для МП-оценок (см. теорему 7.1). ■

П р и м е р 12.2. Пусть в модели наблюдения (12.2) вектор ошибок E_n имеет ковариационную матрицу $K_{E_n} = \sigma^2 I$, где $\sigma > 0$. Показать, что МНК-оценка $\widehat{\theta}_n$ является наилучшей линейной несмещенной оценкой вектора θ.

Р е ш е н и е. Пусть $\widetilde{\theta}_n = A_n Z_n$ — произвольная линейная несмещенная оценка θ. Тогда

$$\mathbf{M}\left\{\widetilde{\theta}_n\right\} = A_n\mathbf{M}\{Z_n\} = A_n(H_n\theta + \mathbf{M}\{E_n\}) = A_n H_n\theta = \theta.$$

Отсюда $(A_nH_n - I)\theta = 0$. Последнее в силу произвольности θ означает, что
$$A_nH_n = I.$$
Заметим, что $\widehat{\theta}_n = A_n^0 Z_n$, где $A_n^0 = W_n^{-1}H_n^\top$, причем
$$\mathbf{M}\left\{\widehat{\theta}_n\right\} = W_n^{-1}H_n^\top H_n\theta + W_n^{-1}H_n\mathbf{M}\{E_n\} = W_n^{-1}W_n\theta = \theta.$$

Таким образом, $\widehat{\theta}_n$ — линейная несмещенная оценка. Отсюда следует, что $A_n^0H_n = I$. Обозначим $T_n = A_n - A_n^0$.
$$T_nH_n = (A_n - A_n^0)H_n = A_nH_n - A_n^0H_n = I - I = 0.$$
Отсюда $T_n\left(A_n^0\right)^\top = T_nH_nW_n^{-1} = 0$.

Найдем ковариационную матрицу $\widetilde{K}_n$ ошибки $\Delta\widetilde{\theta}_n = \widetilde{\theta} - \theta$ оценки $\widetilde{\theta}_n$. Заметим, что $\widetilde{\theta}_n = \theta + A_nE_n$. Отсюда
$$\widetilde{K}_n = \mathbf{M}\left\{\Delta\widetilde{\theta}_n\left(\Delta\widetilde{\theta}_n\right)^\top\right\} = \mathbf{M}\left\{A_nE_n(A_nE_n)^\top\right\} = \\ = A_n\mathbf{M}\{E_nE_n^\top\}A_n^\top = \sigma^2A_nA_n^\top.$$

Учитывая, что $A_n = A_n^0 + T_n$ и $T_n\left(A_n^0\right)^\top = 0$, получаем
$$\widetilde{K}_n = \sigma^2\left(A_n^0 + T_n\right)\left(A_n^0 + T_n\right)^\top = \sigma^2A_n^0\left(A_n^0\right)^\top + T_n\left(A_n^0\right)^\top + \\ + \left(T_n\left(A_n^0\right)^\top\right)^\top + \sigma^2T_n(T_n)^\top = \sigma^2A_n^0\left(A_n^0\right)^\top + \sigma^2T_n(T_n)^\top \geqslant \widehat{K}_n,$$
где $\widehat{K}_n = \sigma^2A_n^0\left(A_n^0\right)^\top = \sigma^2W_n^{-1}$ — ковариационная матрица ошибки $\Delta\widehat{\theta}_n = \widehat{\theta}_n - \theta$ МНК-оценки $\widehat{\theta}_n$, а $\sigma^2T_n(T_n)^\top \geqslant 0$ при любой величине матрицы T_n. Итак, наименьшей ковариационной матрицей ошибки среди всех линейных несмещенных оценок обладает МНК-оценка $\widehat{\theta}_n$. ■

Пример 12.3. Линейная функция $\varphi(t;\theta) = \theta_1 + \theta_2t$ измеряется в дискретные моменты $\{t_k\}$ по схеме
$$X_k = \varphi(t_k;\theta) + \varepsilon_k, \quad k = 1, \ldots, 5.$$
Случайные ошибки измерений $\{\varepsilon_k\}$ — независимые центрированные гауссовские СВ с дисперсией $\sigma^2 = 0{,}01$. Результаты наблюдений $\{x_k\}$ приведены в таблице 12.1.

Таблица 12.1

k	1	2	3	4	5
t_k	0	1	2	3	4
x_k	−1,10	1,15	3,20	4,85	7,10

Найти МНК-оценку $\widehat{\varphi}(t;\theta)$ наблюдаемой функции.

Решение. По условию $\varphi(t;\theta) = h^\top(t)\theta$, где $h^\top(t) = \{1;t\}$. Поэтому $h^\top(t_k) = \{1;t_k\}$. Отсюда
$$W_n = H_n^\top H_n = \begin{bmatrix} n & \sum_{k=1}^n t_k \\ \sum_{k=1}^n t_k & \sum_{k=1}^n t_k^2 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 30 \end{bmatrix}.$$
$$H_n^\top z_n = \begin{bmatrix} \sum_{k=1}^n x_k \\ \sum_{k=1}^n t_kx_k \end{bmatrix} = \begin{bmatrix} 15{,}2 \\ 50{,}5 \end{bmatrix},$$ где $z_n = \{x_1, \ldots, x_n\}^\top$ — реализация выборки Z_n, приведенная в таблице 12.1.

Реализация МНК-оценки вектора $\theta = \{\theta_1;\theta_2\}^\top$ имеет вид
$$\widehat{\theta}_n = W_n^{-1}H_n^\top z_n = \begin{bmatrix} 0{,}6 & -0{,}2 \\ -0{,}2 & 0{,}1 \end{bmatrix} \cdot \begin{bmatrix} 15{,}2 \\ 50{,}5 \end{bmatrix} = \begin{bmatrix} -0{,}98 \\ 2{,}01 \end{bmatrix}.$$

Согласно принципу инвариантности
$$\widehat{\varphi}(t;\theta) = \varphi(t;\widehat{\theta}_n) = h^\top(t)\widehat{\theta}_n = -0{,}98 + 2{,}01t.$$
■

Пример 12.4. В условиях примера 12.3 найти закон распределения оценки $\widehat{\varphi}(t;\theta)$ полезного сигнала и построить доверительный интервал надежности $q = 0{,}95$ для $\varphi(t;\theta)$ в момент $t = 5$.

Решение. Так как $\widehat{\varphi}(t;\theta) = h^\top(t)\widehat{\theta}_n$, то
$$\mathbf{M}\{\widehat{\varphi}(t;\theta)\} = h^\top(t)\mathbf{M}\left\{\widehat{\theta}_n\right\} = h^\top(t)\theta,$$
$$\mathbf{D}\{\widehat{\varphi}(t;\theta)\} = h^\top(t)\widehat{K}_nh(t) = \sigma^2h^\top(t)W_n^{-1}h(t),$$

где $\widehat{K}_n = \sigma^2 W_n^{-1}$ — ковариационная матрица ошибки оценки $\widehat{\theta}_n$ (см. теорему 12.2). Итак,

$$\widehat{\varphi}(t;\theta) \sim \mathcal{N}\left(h^\top(t)\theta;\ \sigma^2 h^\top(t) W_n^{-1} h(t)\right).$$

При этом ошибка оценки $\Delta\widehat{\varphi}(t;\theta) = \widehat{\varphi}(t;\theta) - \varphi(t;\theta)$ имеет распределение $\mathcal{N}\left(0;\ \sigma^2 h^\top(t) W_n^{-1} h(t)\right)$.

В условиях примера 12.3 для $t = 5$ находим

$$\sigma^2 h^\top(t) W_n^{-1} h(t) = 0{,}01 \begin{bmatrix} 1 \\ 5 \end{bmatrix}^\top \cdot \begin{bmatrix} 0{,}6 & -0{,}2 \\ -0{,}2 & 0{,}1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} = 0{,}011.$$

Итак, $\widehat{\varphi}(5;\theta) - \varphi(5;\theta) \sim \mathcal{N}(0; 0{,}011)$. Отсюда

$$\mathbf{P}\left(|\widehat{\varphi}(5;\theta) - \varphi(5;\theta)| \leqslant u_\alpha \sqrt{0{,}011}\right) = 0{,}95,$$

если $u_\alpha = 1{,}96$ — квантиль уровня $\alpha = 0{,}975$ распределения $\mathcal{N}(0;1)$. Таким образом, искомый доверительный интервал имеет вид $\left[\widehat{\varphi}(5;\theta) - 1{,}96\sqrt{0{,}011};\ \widehat{\varphi}(5;\theta) + 1{,}96\sqrt{0{,}011}\right]$. Найдем реализацию этого интервала, используя результаты примера 12.3. Так как $\widehat{\varphi}(5;\theta) = -0{,}98 + 2{,}01 \cdot 5 = 9{,}07$, окончательно получаем интервал $[8{,}86;\ 9{,}28]$, который с высокой надежностью накрывает точное значение $\varphi(5;\theta)$ полезного сигнала в точке $t = 5$. ■

12.3. Задачи для самостоятельного решения.

1. Доказать теорему Гаусса-Маркова (теорема 12.3).

Указание. См. пример 12.2.

2. Вывести выражение (12.7) для НЛН-оценки $\widehat{\theta}_n$ в случае произвольной невырожденной ковариационной матрицы V_n вектора ошибок наблюдения E_n.

Указание. С помощью матрицы Σ_n: $V_n = \Sigma_n(\Sigma_n)^\top$ преобразовать обобщенную регрессию к регрессии с некоррелированными ошибками. Воспользоваться следствием 12.1.

3. Доказать, что ОМНК-оценка $\widehat{\theta}_n$ (12.7) вектора θ в модели обобщенной линейной регрессии является МП-оценкой, если $E_n \sim \mathcal{N}(0; V_n)$, $V_n > 0$.

Указание. Показать, что $\theta_n = \arg\min\limits_\theta (Z_n - H_n\theta)^\top V_n^{-1}(Z_n - H_n\theta)$.

4. Пусть $\widehat{\theta}_n$ — МНК-оценка вектора θ в модели линейной гауссовской регрессии (12.2). Показать, что $\widehat{\sigma}_n^2 = \dfrac{1}{n-p}|Z_n - H_n\widehat{\theta}_n|^2$ является несмещенной оценкой дисперсии σ^2 ошибок наблюдений.

Указание. Воспользоваться следствием 12.2.

5. Модель (12.1) имеет следующий частный вид:

$$X_k = \theta_1 + k\theta_2 + \varepsilon_k, \quad k = 1, \ldots, n.$$

Найти явное аналитической выражение для оценки $\widehat{\theta}_n$ вектора $\theta = \{\theta_1, \theta_2\}^\top$ и доказать ее состоятельность.

Указание. Воспользоваться следствием 12.1. Доказать, что $\widehat{K}_n \to 0$ при $n \to \infty$.

6. Модель (12.1) имеет вид

$$X_k = k\theta + \varepsilon_k, \quad k = 1, \ldots, n,$$

где $\{\varepsilon_k\}$ — независимые гауссовские СВ, $\varepsilon_k \sim \mathcal{N}(0;\sigma^2)$, $\sigma > 0$. Используя МНК-оценку $\widehat{\theta}_n$ параметра θ, построить для него доверительный интервал надежности $q = 0{,}95$.

Ответ. $\left[\widehat{\theta}_n - 1{,}96\,\sigma\sqrt{\psi(n)};\ \widehat{\theta}_n + 1{,}96\,\sigma\sqrt{\psi(n)}\right]$, где $\widehat{\theta}_n = \psi(n)\sum\limits_{k=1}^{n} kX_k$, $\psi(n) = 6(2n^2 + 3n + 1)^{-1}$.

7. В условиях примера 12.3 проверить на уровне значимости $p = 0{,}05$ параметрическую гипотезу H_0: $\theta_2 = 0$.

Указание. Построить доверительный интервал надежности $q = 0{,}95$ для параметра θ_2.

Ответ. H_0 отвергается.

§ 13. Таблицы

Таблица 13.1. Функция $\Phi_0(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2}\, dt.$

x	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	,0039	,0079	,0119	,0159	,0199	,0239	,0279	,0318	,0358
0,1	,0398	,0438	,0477	,0517	,0556	,0596	,0635	,0674	,0714	,0753
0,2	,0792	,0831	,0870	,0909	,0948	,0987	,1025	,1064	,1102	,1140
0,3	,1179	,1217	,1255	,1293	,1330	,1368	,1405	,1443	,1480	,1517
0,4	,1554	,1591	,1627	,1664	,1700	,1736	,1772	,1808	,1843	,1879
0,5	,1914	,1949	,1984	,2019	,2054	,2088	,2122	,2156	,2190	,2224
0,6	,2257	,2290	,2323	,2356	,2389	,2421	,2453	,2485	,2517	,2549
0,7	,2580	,2611	,2642	,2673	,2703	,2733	,2763	,2793	,2823	,2852
0,8	,2881	,2910	,2938	,2967	,2995	,3023	,3051	,3078	,3105	,3132
0,9	,3159	,3185	,3212	,3228	,3263	,3289	,3314	,3339	,3364	,3389
1,0	,3413	,3437	,3461	,3485	,3508	,3531	,3554	,3576	,3599	,3621
1,1	,3643	,3665	,3686	,3707	,3728	,3749	,3769	,3790	,3810	,3829
1,2	,3849	,3868	,3887	,3906	,3925	,3943	,3961	,3979	,3997	,4014
1,3	,4032	,4049	,4065	,4082	,4098	,4114	,4130	,4146	,4162	,4177
1,4	,4192	,4207	,4222	,4236	,4250	,4264	,4278	,4292	,4305	,4318
1,5	,4331	,4344	,4357	,4369	,4382	,4394	,4406	,4417	,4429	,4440
1,6	,4452	,4463	,4473	,4484	,4495	,4505	,4515	,4525	,4535	,4544
1,7	,4554	,4563	,4572	,4581	,4590	,4599	,4608	,4616	,4624	,4632
1,8	,4640	,4648	,4656	,4663	,4671	,4678	,4685	,4692	,4699	,4706
1,9	,4712	,4719	,4725	,4732	,4738	,4744	,4750	,4755	,4761	,4767
2,0	,4772	,4777	,4783	,4788	,4793	,4798	,4803	,4807	,4812	,4816
2,1	,4821	,4825	,4830	,4834	,4838	,4842	,4846	,4850	,4853	,4857
2,2	,4861	,4864	,4867	,4871	,4874	,4877	,4880	,4884	,4887	,4889
2,3	,4892	,4895	,4898	,4901	,4903	,4906	,4908	,4911	,4913	,4915
2,4	,4918	,4920	,4922	,4924	,4926	,4928	,4930	,4932	,4934	,4936
2,5	,4937	,4939	,4941	,4943	,4944	,4946	,4947	,4949	,4950	,4952
2,6	,4953	,4954	,4956	,4957	,4958	,4959	,4960	,4962	,4963	,4964
2,7	,4965	,4966	,4967	,4968	,4969	,4970	,4971	,4972	,4972	,4973
2,8	,4974	,4975	,4976	,4976	,4977	,4978	,4978	,4979	,4980	,4980
2,9	,4981	,4981	,4982	,4983	,4983	,4984	,4984	,4985	,4985	,4986
3,0	,49865									
3,5	,499767									
4,0	,4999683									
4,5	,4999966									
5,0	,49999971									

Таблица 13.2. Квантили u_α уровня α распределения $\mathcal{N}(0;1)$.

α	0,6	0,8	0,9	0,95	0,975	0,99	0,995	0,9995
u_α	0,253	0,842	1,282	1,645	1,960	2,326	2,576	3,291

Таблица 13.3. Квантили k_α уровня α хи-квадрат распределения $\mathcal{H}_r$.

r \ α	0,05	0,1	0,2	0,3	0,5	0,7	0,8	0,9	0,95
1	0,00	0,02	0,06	0,15	0,46	1,07	1,64	2,71	3,84
2	0,10	0,21	0,45	0,71	1,39	2,41	3,22	4,60	5,99
3	0,35	0,58	1,01	1,42	2,37	3,66	4,64	6,25	7,82
4	0,71	1,06	1,65	2,20	3,36	4,88	5,99	7,78	9,49
5	1,15	1,61	2,34	3,00	4,35	6,06	7,29	9,24	11,07
6	1,65	2,20	3,07	3,83	5,35	7,23	8,56	10,64	12,59
7	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07
8	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51
9	3,32	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92
10	3,94	4,86	6,18	7,27	9,34	11,78	13,44	15,99	18,31
11	4,58	5,58	6,99	8,15	10,34	12,90	14,63	17,28	19,68
12	5,23	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,03
13	5,89	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,36
14	6,57	7,79	9,47	10,82	13,34	16,22	18,15	21,06	23,69
15	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,31	25,00
16	7,96	9,31	11,15	12,62	15,34	18,42	20,47	23,54	26,29
17	8,67	10,08	12,00	13,53	16,34	19,51	21,62	24,78	27,60
18	9,39	10,86	12,86	14,44	17,34	20,60	22,76	25,59	28,87
19	10,11	11,65	13,72	15,35	18,34	21,70	23,90	27,20	30,14
20	10,85	12,44	14,58	16,27	19,34	22,80	25,04	28,41	31,41
21	11,59	13,24	15,44	17,18	20,30	23,90	26,17	29,61	32,67
22	12,34	14,04	16,31	18,10	21,30	24,90	27,30	30,81	33,92
23	13,09	14,85	17,19	19,02	22,30	26,00	28,43	32,01	35,17
24	13,85	15,66	18,06	19,94	23,30	27,10	29,55	33,20	36,42
25	14,61	16,47	18,94	20,90	24,30	28,20	30,78	34,38	37,65
26	15,38	17,29	19,82	21,80	25,30	29,20	31,80	35,56	38,89
27	16,15	18,11	20,70	22,70	26,30	30,30	32,91	36,74	40,11
28	16,93	18,94	21,60	23,60	27,30	31,40	34,03	37,92	41,34
29	17,71	19,77	22,50	24,60	28,30	32,50	35,14	39,09	42,56
30	18,49	20,60	23,40	25,50	29,30	33,50	36,25	40,26	43,77

Таблица 13.4. Квантили t_α уровня α распределения Стьюдента T_r.

$r \setminus \alpha$	0,6	0,8	0,9	0,95	0,975	0,99	0,995	0,9995
1	0,325	1,376	3,078	6,314	12,706	31,821	63,657	636,619
2	0,289	1,061	1,886	2,920	4,303	6,965	9,925	31,598
3	0,277	0,978	1,638	2,353	3,182	4,541	5,841	12,941
4	0,271	0,941	1,533	2,132	2,776	3,747	4,604	8,610
5	0,267	0,920	1,476	2,015	2,571	3,365	4,032	6,859
6	0,265	0,906	1,440	1,943	2,447	3,143	3,707	5,959
7	0,263	0,896	1,415	1,895	2,365	2,998	3,499	5,405
8	0,262	0,889	1,397	1,860	2,306	2,896	3,355	5,041
9	0,261	0,883	1,383	1,833	2,262	2,821	3,250	4,781
10	0,260	0,879	1,372	1,812	2,228	2,764	3,169	4,587
11	0,260	0,876	1,363	1,796	2,201	2,718	3,106	4,437
12	0,259	0,873	1,356	1,782	2,179	2,681	3,055	4,318
13	0,259	0,870	1,350	1,771	2,160	2,650	3,012	4,221
14	0,258	0,868	1,345	1,761	2,145	2,624	2,977	4,140
15	0,258	0,866	1,341	1,753	2,131	2,602	2,947	4,073
16	0,258	0,865	1,337	1,746	2,120	2,583	2,921	4,015
17	0,257	0,863	1,333	1,740	2,110	2,567	2,898	3,965
18	0,257	0,862	1,330	1,734	2,101	2,552	2,878	3,922
19	0,257	0,861	1,328	1,729	2,093	2,539	2,861	3,883
20	0,257	0,860	1,325	1,725	2,086	2,528	2,845	3,850
21	0,257	0,859	1,323	1,721	2,080	2,518	2,831	3,819
22	0,256	0,858	1,321	1,717	2,074	2,508	2,819	3,792
23	0,256	0,858	1,319	1,714	2,069	2,500	2,807	3,767
24	0,256	0,857	1,318	1,711	2,064	2,492	2,797	3,745
25	0,256	0,856	1,316	1,708	2,060	2,485	2,787	3,725
26	0,256	0,856	1,315	1,706	2,056	2,479	2,779	3,707
27	0,256	0,855	1,314	1,703	2,052	2,473	2,771	3,690
28	0,256	0,855	1,313	1,701	2,048	2,467	2,763	3,674
29	0,256	0,854	1,311	1,699	2,045	2,462	2,756	3,659
30	0,256	0,854	1,310	1,697	2,042	2,457	2,750	3,646
40	0,255	0,851	1,303	1,684	2,021	2,423	2,704	3,551
60	0,254	0,848	1,296	1,671	2,000	2,390	2,660	3,460
120	0,254	0,845	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,253	0,842	1,282	1,645	1,960	2,326	2,576	3,291

СПИСОК ЛИТЕРАТУРЫ

1. *Болдин М.В. и др.* Методические указания к практическим занятиям по математической статистике. М.: МАИ, 1988.
2. *Боровков А.А.* Математическая статистика. М.: Наука, 1984.
3. *Ватутин В.А. и др.* Теория вероятностей и математическая статистика в задачах. М.: Агар, 2003.
4. *Ивченко Г.И., Медведев Ю.И.* Математическая статистика. М.: Высшая школа, 1984.
5. *Кибзун А.И. и др.* Теория вероятностей и математическая статистика в задачах. М.: Физматлит, 2002.
6. *Кочетков Е.С.* Метод наименьших квадратов. М.: МАИ, 1993.
7. *Пугачев В.С.* Теория вероятностей и математическая статистика. — М.: Наука, 1979.
8. *Севастьянов Б.А.* Курс теории вероятностей и математической статистики. М.: Наука, 1982.
9. *Ширяев А.Н.* Вероятность. — М.: Наука, 1989.