Лекции

по курсу *”Математическая Статистика”*

Ибрагимов Д. Н.

Contents

Источники {-}

- Ивченко Г. И., Медведев Ю. И. “Математическая статистика”, изд. “Высшая школа”, 1984
- Кибзун А. И., Наумов А. В., Горяинова Е. Р. “Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами”, изд “ФИЗМАТЛИТ”, 2013
- Панков А. Р., Платонов Е. Н. “Практикум по математической статистике”, изд. “МАИ”, 2006

1 Многомерное нормальное распределение

Замечание

Вектор $X = (X_1, \dots, X_n)^T$ называется ***случайным***, если $X_1, \dots, X_n$ — *случайные величины (далее **с.в**)*, определенные на одном вероятностном пространстве.

Через $M[X] = m_X$ обозначим ***вектор математического ожидания***:

$$M[X] = m_X = \begin{pmatrix} M[X_1] \\ \vdots \\ M[X_n] \end{pmatrix}$$

Через K_x обозначим *ковариационную матрицу* с.в X:

$$K_X = \begin{pmatrix} \text{cov}(X_1, X_1) & \dots & \text{cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \text{cov}(X_n, X_1) & \dots & \text{cov}(X_n, X_n) \end{pmatrix}$$

1.1 Лемма 1

Пусть $K_X \in \mathbb{R}^{n \times n}$ — ковариационная матрица с.в X. Тогда:

1. $K_X \geqslant 0$, т.е. $\forall x \in \mathbb{R}^n \backslash \{0\}, x^T K_X x \geqslant 0$;
2. $K_X^T = K_X$

1.2 Определение 1

Случайный вектор $X = (X_1, \dots, X_n)^T$ называется ***невырожденным нормальным вектором***:

$$X \sim N(m_X, K_X)$$

если совместная плотность вероятности имеет вид:

$$f_X(x) = ((2\pi)^n \det K_X)^{\frac{-1}{2}} \exp\{\frac{-1}{2}(x - m_X)^T K_X^{-1}(x - m_X)\}$$

где $m_X \in \mathbb{R}^n, K_X \in \mathbb{R}^{n \times n}, K_X > 0, K_x^T = K_X$

1.3 Лемма 2

Пусть X — невырожденный нормальный вектор с параметрами m_X и K_X.
Тогда $M[X] = m_X$, а K_X — корвариационная матрица X.
Рассмотрим основные свойства многомерного нормального распределения.

1.4 Лемма 3

Пусть $X \sim N(m_X, K_X), A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.
Тогда:

$$\begin{aligned} Y = AX + b \sim N(m_Y, K_Y), \\ m_Y = Am_X + b, \\ K_Y = AK_XA^T. \end{aligned}$$

1.5 Лемма 4

Пусть $X \sim N(m_X, K_X)$.
Тогда компоненты вектора X ***независимы*** тогда и только тогда, когда они *некоррелированы.*

note

Доказательство данных утверждений при помощи аппарата функций распределения и плотности довольно сложно. Поэтому рассмотрим аппарат характеристических функций.

1.6 Определение 2

Пусть $X = (X_1, \dots, X_n)^T -$ случайный вектор.
Тогда ***характеристической функцией*** называется:

$$\psi_X(\lambda) = M[e^{i\lambda^T X}] = \int\limits_{\mathbb{R}^n} e^{i\lambda^T X} dF_X(x)$$

Замечание

Характеристическая функция определена для любого случайного вектора или с.в.
Если с.в **дискретная**, то:

$$\psi_X(\lambda) = \sum_{k=1}^{\infty} e^{i\lambda X_k} p_k$$

Если с.в **абсолютно непрерывная**, то

$$\psi_X(\lambda) = \int\limits_{\mathbb{R}} e^{i\lambda X} f_X(x) dx$$

В этом случае $\psi_X(\lambda)$ является ***преобразованием Фурье*** f_X.
Поскольку преобразование Фурье взаимно однозначно, а f_X однозначно определяет распределение, то характеристическая функция $\psi_X(x)$ также однозначно определяет распределение с.в X.

Причем:

$$f_X(x) = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}} e^{-i\lambda^T X} \psi_X(x) d\lambda$$

1.7 Лемма 5

Пусть X — случайный вектор, $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n$.
Тогда:

1. для $Y = AX + b$
$$\psi_Y(\lambda) = e^{i\lambda^T b} \psi_X(A^T \lambda)$$
2. компоненты вектора X ***независимы*** тогда и только тогда, когда
$$\psi_Y(\lambda) = \prod_{k=1}^{n} \psi_{X_k}(\lambda_k)$$

Доказательство

1. $\psi_Y(\lambda) = M[e^{i\lambda^T Y}] = M[e^{i\lambda^T AX} e^{i\lambda^T b}] = e^{i\lambda^T b} M[e^{i(A^T\lambda)^T X}] = e^{i\lambda^T b} \psi_X(A^T \lambda)$
2. $\psi_X(\lambda) = \int\limits_{\mathbb{R}} \dots \int\limits_{\mathbb{R}} e^{i(\lambda_1 x_1 + \dots + \lambda_n x_n)} f_X(x_1, \dots, x_n) dx_1 \cdot \dots \cdot dx_n = \{\text{н/з}\} = \int\limits_{\mathbb{R}} e^{i\lambda_1 x_1} \cdot \dots \cdot e^{\lambda_n x_n} \cdot$ $f_{X_1}(x) \cdot \dots \cdot f_{X_n} dx_1 \cdot \dots \cdot dx_n = \int\limits_{\mathbb{R}} e^{i\lambda_1 x_1} f_{x_1}(x_1) dx_1 \cdot \dots \cdot \int\limits_{\mathbb{R}} e^{i\lambda_n x_n} f_{x_n}(x_n) dx_n = \prod_{k=1}^{n} \psi_{X_k}(\lambda_k)$

■

Замечание

При помощи характеристической функции можно дать другое определение нормального распределения. В том числе для вырожденного K_X.

1.8 Определение 3

Случайный вектор X называется ***нормальным***: $X \sim N(m_X, K_X)$, если:

$$\psi_X(\lambda) = \exp\{i\lambda^T m_X - \frac{1}{2}\lambda^T K_X \lambda\}$$

1.8.1 Доказательство леммы 3

В силу Леммы 5, п.1

$$\psi_Y(\lambda) = e^{i\lambda^T b} \psi_X(A^T \lambda) = e^{i\lambda^T b} \exp\{i\lambda^T A m_x - \frac{1}{2}\lambda^T A K_X A^T \lambda\} =$$
$$= \exp\{i\lambda^T (A m_x + b) - \frac{1}{2}\lambda^T (A K_x A^T) \lambda\}$$

■

1.8.2 Доказательство леммы 4

Пусть $X_i, \dots, X_n$ *попарно некоррелированы.* Тогда $cov(X_i, X_j) = 0, i \neq 0$, т.е. :

$$K_x = diag(\sigma^2_{X_1}, \dots, \sigma^2_{X_n}) = \\ = \begin{pmatrix} \sigma^2_{X_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma^2_{X_n} \end{pmatrix}$$

$\psi_X(\lambda) = \exp\{i\lambda_1 m_{X_1} + \dots + i\lambda_n m_{X_n} - \frac{1}{2}\lambda^T K_X \lambda\} = \exp\{i\lambda_1 m_{X_1} + \dots + i\lambda_n m_{X_n} - \frac{1}{2}(\lambda_1^2\sigma^2_{X_1} + \dots + \lambda_n^2\sigma^2_{X_n})\} = \prod\limits_{k=1}^{n} \exp\{i\lambda_n m_{X_n} - \frac{1}{2}\lambda_k^2\sigma^2_{K_n}\} = \prod\limits_{k=1}^{n} \psi_{X_k}(\lambda_k)$.

Откуда с учетом Леммы 5, п.1 $X_1, \dots, X_n$ — н/з.

Пусть $X_1, \dots, X_n$ — н/з. Тогда $X_1, \dots, X_n$ *попарно некоррелированы.* ■

Замечание

Поскольку K_X — невырожденная, симметричная и положительноопределенная, то существует $S \in \mathbb{R}^{n \times n}$ — *ортогональная* (т. е. $S^T = S^{-1}$) такая, что:

$$S^T K_X S = \Lambda = diag(\lambda_1, \dots, \lambda_n)$$

где $\lambda_i > 0, i = \overline{1, n}$

Определим матрицу $\Lambda^{-\frac{1}{2}} = diag(\lambda_1^{-\frac{1}{2}}, \dots, \lambda_n^{-\frac{1}{2}})$.

Рассмотрим вектор

$$Y = \Lambda^{-\frac{1}{2}} S^T (X - m_X)$$

Тогда $A = \Lambda^{-\frac{1}{2}} S^T, b = -\Lambda^{-\frac{1}{2}} S^T m_X$.

В силу Леммы 3:

$$m_Y = A m_X + b = \Lambda^{-\frac{1}{2}} S^T - \Lambda^{-\frac{1}{2}} S^T m_X = 0,$$
$$K_Y = A K_X A^T = \Lambda^{-\frac{1}{2}} S^T K_X S \Lambda^{-\frac{1}{2}} = I,$$

т. е. $Y \sim N(0, I)$.

При помощи невырожденного линейного преобразования с.в. X может быть преобразован в стандартный нормальный вектор.

Верно и обратное:

$$X = m_X + S\Lambda^{\frac{1}{2}} Y,$$

откуда следует Лемма 2

2 Теорема о нормальной корреляции

2.1 Определение 1

Условным математическим ожиданием абсолютно непрерывного случайного вектора X относительно абсолютно непрерывного случайного вектора Y называется:

$$M[X \mid Y] = \int\limits_{-\infty}^{+\infty} x f_{X|Y}(x \mid Y) dx,$$

где $f_{X|Y}(x \mid Y) = \frac{f_z(x,Y)}{f_y(Y)}, z = \begin{pmatrix} X \\ Y \end{pmatrix}$

2.2 Основные свойства условного М. О.

2.2.1 Свойство 1

$\boxed{M[C \mid Y] = C}$

Доказательство

$M[C \mid Y] = \int\limits_{-\infty}^{+\infty} C f_{X|Y}(x \mid Y) dx = \frac{C \int\limits_{-\infty}^{+\infty} f_z(x,Y) dx}{f_Y(Y)} = C \frac{f_Y(Y)}{f_Y(Y)} = C.$ ■

2.2.2 Свойство 2

$\boxed{M[X\phi(Y)|Y] = \phi(Y) M[X \mid Y]}$

Доказательство

$M[\phi(Y)X \mid Y] = \int\limits_{-\infty}^{+\infty} \phi(Y) X f_{X|Y}(x \mid Y) dx = \phi(Y) \int\limits_{-\infty}^{+\infty} x f_{X|Y}(x \mid Y) dx =$
$= \phi(Y) M[X \mid Y].$ ■

2.2.3 Свойство 3

$\boxed{M[\alpha X_1 + \beta X_2 \mid Y] = \alpha M[X_1 \mid Y] + \beta M[X_2 \mid Y]}$

2.2.4 Свойство 4

Пусть X, Y – *независимые.* Тогда $\boxed{M[X \mid Y] = M[X]}$

Доказательство

$M[X \mid Y] = \int\limits_{-\infty}^{+\infty} x f_{X|Y}(x \mid Y) dx = \int\limits_{-\infty}^{+\infty} x \frac{f_z(x,Y)}{f_Y(Y)} dx = \int\limits_{-\infty} +\infty x \frac{f_X(x) f_Y(Y)}{f_Y(Y)} dx = M[X].$ ■

2.2.5 Свойство 5

$\boxed{M[M[X \mid Y]] = M[X]}$ ***(формула повторного М. О.)***

Доказательство

$M[M[X \mid Y]] = \int\limits_{-\infty}^{+\infty} M[X \mid Y] f_Y(y) dy = \int\limits_{-\infty}^{+\infty} f_Y(y) \int\limits_{-\infty}^{+\infty} x f_{X|Y}(x \mid y) dx dy = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} x \frac{f_Y(y) f_z(x,y)}{f_Y(y)} dy dx = \int\limits_{-\infty}^{+\infty} x \int\limits_{-\infty}^{+\infty} f_z(x,y) dy dx = \int\limits_{-\infty}^{+\infty} x f_X(x) dx = M[X]$. ■

2.3 Лемма 1

Пусть X, Y — случайные векторы с конечными вторыми моментами. Тогда:
$M[(X - \hat{X})\phi(Y)^T] = 0$
где $\hat{X} = M[X \mid Y]$

Доказательство

$M[(X - \hat{X})\phi(Y)^T] = M[X\phi(Y)^T] - M[M[X \mid Y]\phi(Y)^T] =$
= по Свойству 2 $= M[X\phi(Y)^T] - M[M[X\phi(Y)^T \mid Y]] =$ по Свойству 5 $= M[X\phi(Y)^T] - M[X\phi(Y)^T] = 0$. ■

Замечание

Если рассмотреть евклидово пространство $\mathbb{L}_2(\Omega)$ со скалярным произведением:

$$(X, Y) = M[X \cdot Y]$$

то *условное* М. О. — ***оператор ортогонального проектирования*** X на подпространство, порождаемое Y.

2.4 Определение 2

Оценкой X ***по наблюдениям*** Y называется любая измеримая функция $\phi(Y)$.

2.5 Определение 3

Оценка $\hat{X}$ ***называется с.к.-оптимальной оценкой*** X, если для любой другой оценки $\tilde{X}$ верно

$$M[|\tilde{X} - \hat{X}|^2] \leqslant M[|X - \tilde{X}|^2]$$

2.6 Теорема 1

$M[X \mid Y]$ — ***с.к.-оптимальная оценка*** X ***по наблюдениям*** Y.

Доказательство

$M[|X-\tilde{X}|^2] = M[|X-\hat{X}+\hat{X}-\tilde{X}|^2] = M[|X-\hat{X}|^2]+2M[(X-\hat{X})^T(\hat{X}-\tilde{X})]+M[|\hat{X}-\tilde{X}|^2] \overset{*}{=}$

Поскольку по определению $\tilde{X}-\hat{X} = \phi(Y)$, то в силу Леммы 2.1 $M[(X-\hat{X})^T(\tilde{X}-\hat{X})] = 0$.

$\overset{*}{=} M[|X-\hat{X}|^2] + M[|\hat{X}-\tilde{X}|^2] \geqslant M[|X-\hat{X}|^2]$. ■

2.7 Теорема 2 (О нормальной корреляции)

Пусть

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} m_X \\ m_Y \end{pmatrix}, \begin{pmatrix} K_X & K_{XY} \\ K_{XY}^T & K_Y \end{pmatrix}\right)$$

Тогда

1. $Law(X \mid Y) = N(\mu(Y), \Delta)$,

где

$$\mu(Y) = M[X \mid Y] = m_X + K_{XY}K_Y^{-1}(Y - m_Y)$$

$$\Delta = K_X - K_{XY}K_Y^{-1}K_{YX}$$

2. $M[|X-\mu(Y)|^2] = tr(K_X - K_{XY}K_Y^{-1}K_{YX})$

Доказательство

Рассмотрим линейное преобразование Y:

$$\mu(Y) = m_X + K_{XY}K_Y^{-1}(Y - m_Y)$$

В силу Леммы 1.3

$X-\mu(Y) = (I - K_{XY}K_Y^{-1})\begin{pmatrix} X \\ Y \end{pmatrix} - m_X + K_{XY}K_Y^{-1}m_Y \sim N(\mu, K)$

$\mu = (I - K_{XY}K_Y^{-1})\begin{pmatrix} m_X \\ m_Y \end{pmatrix} - m_X + K_{XY}K_Y^{-1}m_Y = 0$

$K = (I - K_{XY}K_Y^{-1})\begin{pmatrix} K_X & K_{XY} \\ K_{XY}^T & K_Y \end{pmatrix}\begin{pmatrix} I \\ -(K_{XY}K_Y^{-1})^T \end{pmatrix} =$

$= (K_X - K_{XY}K_Y^{-1}K_{XY}^T \quad K_{XY} - K_{XY}K_Y^{-1}K_Y)\begin{pmatrix} I \\ K_Y^{-1}K_{XY}^T \end{pmatrix} =$

$= K_X - K_{XY}K_Y^{-1}K_{XY}^{-1} = \Delta$

$cov(X-\mu(Y), Y) = cov(X, Y) - cov(\mu(Y), Y) = cov(X, Y) - cov(K_{XY}K_Y^{-1}Y + m_X - K_{XY}K_Y^{-1}m_Y, Y) = cov(X, Y) - K_{XY}K_Y^{-1}cov(Y, Y) = K_{XY} - K_{XY}K_Y^{-2}K_Y = 0$

т.е. $X-\mu(Y)$ и Y **некорреливаны**.

Тогда в силу Леммы 1.5, п.2 $X - \mu(Y)$ и Y **независимы**. Построим характеристическую функцию условного распределения X относительно Y:

$\psi_{X|Y}(\lambda \mid Y) = \int\limits_{\mathbb{R}^n} e^{i\lambda^T X} f_{X|Y}(x \mid Y) dx = M[e^{i\lambda^T X} \mid Y] = M[e^{i\lambda^T (X-\mu(Y))} e^{i\lambda^T \mu(Y)} \mid Y] \overset{*}{=}$

в силу Леммы 1.2 и *независимости* $X - \mu(Y)$ и Y

$\overset{*}{=} M[e^{i\lambda^T (X-\mu(Y))} \mid Y] \cdot M[e^{i\lambda^T \mu(Y)} \mid Y] = M[e^{i\lambda^T (X-\mu(Y))}] e^{i\lambda^T \mu(Y)} = \psi_{X-\mu(Y)}(\lambda) e^{i\lambda^T \mu(Y)} = \exp\{-\frac{1}{2}\lambda^T \Delta \lambda\} \cdot \exp\{i\lambda^T \mu(Y) - \frac{1}{2}\lambda^T \Delta \lambda\}$

т.е. Условное распределение ***нормальное***:

$$X(Y \sim N(\mu(Y), \Delta)$$

Вычислим с.к. ошибку:

$M[|X - \mu(Y)|^2] = M[\Delta X_1^2 + \Delta X_2^2 + \dots + \Delta X_n^2] = \sum\limits_{k=1}^{n} M[\Delta X_k^2] = \sum\limits_{k=1}^{n} D[\Delta X_k] = \sum\limits_{k=1}^{n} \Delta_{kk} = tr\Delta$. ■

Замечание

1. Из Теоремы о нормальной корреляции следует, что в *гауссовском* случае с.к.-оптимальная оценка является ***линейной***.
2. Если X и Y — *независимы*, то с.к.-оптимальная оценка — m_X.
3. С.к.-оптимальная оценка ***несмещенная***, т.к. $M[X - \mu(Y)] = 0$.

3 Виды сходимости последовательностей случайных величин

3.1 Определение 1

Говорят, что $\{X_n\}_{n=1}^{\infty}$ образует ***последовательность случайных величин***, если $\forall N \in \mathbb{N}$ X_n определены на одном вероятностном пространстве.

3.2 Определение 2

Говорят, что последовательность с.в. $\{X_n\}_{n\in\mathbb{N}}$ ***сходится по вероятности*** к с.в. X, если $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} P(|X_n - X| \leqslant \varepsilon) = 1$$

ИЛИ

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$$

3.3 Определение 3

Говорят, что последовательность с.в. $\{X_n\}_{n\in\mathbb{N}}$ ***сходится почти наверное*** к с.в. X, если

$$P(\{\omega : X_n(\omega) \overset{n\to\infty}{\not\longrightarrow} X(\omega)\}) = 0$$

ИЛИ

$$P(\{\omega : X_n(\omega) \overset{n\to\infty}{\longrightarrow} X(\omega)\}) = 1$$

3.4 Определение 4

Говорят, что последовательность с.в. $\{X\}_{n\in\mathbb{N}}$ ***сходится в среднем квадратическом*** к с.в. X, если

$$M[|X_n - X|^2] \overset{n\to\infty}{\longrightarrow} 0$$

3.5 Пример 1

Пусть $\{X_n\}_{n\in\mathbb{N}}$ — случайная последовательность.

$$X \sim \begin{pmatrix} 0 & n \\ 1 - \frac{1}{n^2} & \frac{1}{n^2} \end{pmatrix}$$

$P(\{\omega : \lim\limits_{n\to\infty} X_n(\omega) \neq 0\}) = P(\{\omega : \forall N \in \mathbb{N} \exists n \geqslant N : X_n(\omega) = n\}) = P(\prod\limits_{N=1}^{\infty} \sum\limits_{n=N}^{\infty} \{\omega : X_n(\omega) = n\}) \overset{*}{=}$

1. $\sum\limits_{n=N+1}^{\infty}\{\omega : X_n(\omega) = n\} \subset \sum\limits_{n=N}^{\infty}\{\omega : X_n(\omega) = n\}$

2. $P(\sum\limits_{n=N}^{\infty}\{\omega : X_n(\omega) = n\} \leqslant \sum\limits_{n=N}^{\infty} P(\{\omega : X_n(\omega) = n\}) = \sum\limits_{n=N}^{\infty}\frac{1}{n^2} \xrightarrow{n\to\infty} 0$, т.к. $\sum\limits_{n=1}^{\infty}\frac{1}{n^2} < \infty$

Тогда в силу *аксиомы непрерывности* $\overset{*}{=} 0$, т.е. $X_n \xrightarrow{\text{п.н.}} 0$

3.6 Пример 2

Рассмотрим ту же последовательность. Выберем $\varepsilon > 0$

$$P(|X_n - 0| \leqslant \varepsilon) = \begin{cases} 1, & \varepsilon \geqslant n \\ 1 - \frac{1}{n^2}, & \varepsilon \in (0; n) \end{cases}$$

Тогда $X_n \xrightarrow{P} 0$

3.7 Пример 3

Рассмотрим ту же последовательность:

$M[|X_n - 0|^2] = M[X_n^2] = 0^2(1 - \frac{1}{n^2}) + n^2\frac{1}{n^2} = 1 \not\xrightarrow{n\to\infty} 0$

Тогда $X_n \not\xrightarrow{\text{с.к.}} 0$

3.8 Пример 4

Пусть $f_{nk} : [0; 1] \longrightarrow \{0; 1\}, n \in \mathbb{N}, k = \overline{1, n}$,

$$f_{n,k} = \begin{cases} 0, & t \notin [\frac{k-1}{n}; \frac{k}{n}], \\ 1, & t \in [\frac{k-1}{n}; \frac{k}{n}]. \end{cases}$$

Пусть $X \sim R(0; 1)$. Рассмотрим последовательность с.в. $X_{nk} = f_{nk}(X)$. $\forall\omega \in \Omega X(\omega) \in [0; 1]$. Тогда $\forall n \in \mathbb{N} \exists k = \overline{1, n}$ такое, что $X(\omega) \in [\frac{k-1}{n}; \frac{k}{n}]$. Т.е. если $\varepsilon = \frac{1}{2}$, то $\forall n \in \mathbb{N}$ найдется $k = \overline{1, n}$ такой, что

$$|f_{nk}(X(\omega)) - 0| > \varepsilon$$

Тогда $X_{nk}(\omega) \not\xrightarrow{n\to\infty} 0$, т.е.

$$\{\omega : \lim_{n,k\to\infty} X_{nk}(\omega) = 0\} = \emptyset$$

$$X_{nk} \not\xrightarrow{\text{п. н.}} 0$$

При этом $\forall\varepsilon > 0$

$$R(|X_{nk}-0|>\varepsilon)=\begin{cases}0, & \varepsilon\geqslant 1,\\ P(X\in[\frac{k-1}{n}, \quad \frac{k}{n}]),\varepsilon\in(0;1)\end{cases}=\begin{cases}0, & \varepsilon\geqslant 1,\\ \frac{1}{n}, & \varepsilon\in(0;1)\end{cases}\xrightarrow{n\to\infty}0$$

$$X_{nk}\xrightarrow{P}0$$

$M[|X_{nk}-0|^2]=M[X_{nk}]=M[f_{nk}(X)]=\int\limits_0^1 f_{nk}(x)f_x(x)dx=\int\limits_0^1 f_{nk}(x)dx=\int\limits_{\frac{k-1}{n}}^{\frac{k}{n}} 1dx=\frac{1}{n}\xrightarrow{n\to\infty}0$

$$X_{nk}\xrightarrow{\text{с.к}}0$$

3.9 Пример 5

Рассмотрим последовательность с.в. $Y_{n_1k}=nK_{n_1k}$ Тогда $Y_{n_1k}\xrightarrow{\text{п.н.}}0$. $\forall\varepsilon>0$.
$P(|Y_{n_1k}-0|>\varepsilon)=\begin{cases}0, & \varepsilon\geqslant n,\\ P(X\in[\frac{k-1}{n};\frac{k}{n}]), & \varepsilon\in(0;n)\end{cases}=\begin{cases}0, & \varepsilon\geqslant n,\\ \frac{1}{n}, & \varepsilon\in(0;n)\end{cases}\xrightarrow{nto\infty}0$

$$Y_{n_1k}\xrightarrow{P}0$$

$M[|Y_{n_1k}-0|^2]=M[n^2X_{nk}^2]=n^2M[X_{nk}]=n\xrightarrow{P}\infty$

$$Y_{n_1k}\not\xrightarrow{\text{с.к}}0$$

Замечание

Согласно определению для исследования на сходимость нужно знать совместное распределение с.в. X_n и X, а для случая *сходимости почти наверное* совместное распределение всей последовательности $\{X_n\}_{n\in\mathbb{N}}$ и X. Поэтому исследование на сходимость иначе, чем к детерминированной константе, довольно проблематично.

3.10 Лемма 1

Пусть $X_n\xrightarrow{\text{п.н.}}X$. Тогда $X_n\xrightarrow{P}X$.

Доказательство

$0=P(\{\omega:\lim\limits_{n\to\infty}X_n(\omega)\neq X(\omega)\})=P(\{\omega:\exists\varepsilon>0\forall N\in\mathbb{N}\exists n\geqslant N:|X_n(\omega)-X(\omega)>\varepsilon\})=P(\sum\limits_{\varepsilon>0)}\prod\limits_{N=1}^{\infty}\sum\limits_{n=N}^{\infty}\{\omega:|X_n(\omega)-X(\omega)|>\varepsilon\})\geqslant P(\prod\limits_{N=1}^{\infty}\sum\limits_{n=N}^{\infty}\{\omega:|X_n(\omega)-X(\omega)|>\varepsilon'\})$, $\forall\varepsilon'>0$

Тогда $0=P(\prod\limits_{N=1}^{\infty}\sum\limits_{n=N}^{\infty}\{\omega:|X_n(\omega)-X(\omega)|>\varepsilon\})$,

$$\sum_{n=N+1}^{\infty}\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) \subset \sum_{n=N}^{\infty}\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$$

В силу *аксиомы непрерывности*:

$0 = \lim\limits_{N\to\infty} P(\sum\limits_{n=N}^{\infty}\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) \geqslant \lim\limits_{N\to\infty} P(\{\omega : |X_N(\omega) - X(\omega)| > \varepsilon\})$

$$X_N \xrightarrow{P} X$$

■

3.11 Лемма 2 (Неравенство Маркова)

Пусть $P(X \geqslant 0) = 1$, $M[X] < \infty$. Тогда $\forall \varepsilon > 0$

$$P(X > \varepsilon) = \frac{M[X]}{\varepsilon}$$

Доказательство

$M[X] = \int\limits_0^{+\infty} x dF_X(x) \geqslant \int\limits_{\varepsilon}^{+\infty} x dF_X(x) \geqslant \varepsilon \int\limits_{\varepsilon}^{+\infty} dF_X(X) = \varepsilon P(X > \varepsilon)$. ■

3.12 Следствие 1

Пусть $M[X^{-k}] < \infty$. Тогда $\forall \varepsilon > 0$:

$$P(|X| > \varepsilon) \leqslant \frac{M[|X|^k|]}{\varepsilon^k}$$

Доказательство

$P(|X| > \varepsilon) = P(|X|^k > \varepsilon^k) \leqslant \frac{M[|X|^k]}{\varepsilon^k}$. ■

3.13 Следствие 2 (Неравенство Чебышёва)

Пусть $M[X^2] < \infty$. Тогда

$$P(|X - M[X]| > \varepsilon) \leqslant \frac{D[X]}{\varepsilon^2}$$

Доказательство

$P(|X - M[X]| > \varepsilon) \leqslant \frac{M[|X - M[X]|^2]}{\varepsilon^2} = \frac{D[X]}{\varepsilon^2}$. ■

3.14 Лемма 3

Пусть $X_n \xrightarrow{\text{с.к.}} X$. Тогда $X_n \xrightarrow{P} X$.

Доказательство

$P(|X_n - X| > \varepsilon) \leqslant \frac{M[|X_n - X|^2]}{\varepsilon^2} \xrightarrow{n\to\infty} 0.$ ■

3.15 Теорема 1 (Бореля — Кантелли)

Пусть $A_1, \dots, A_n \subset \Omega, B = \prod\limits_{N=1}^{\infty} \sum\limits_{n=N}^{\infty} A_n$. Тогда

1. Если $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, то $P(B) = 0$;

2. Если $A_1, \dots, A_n$ **независимы в совокупности** и $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, то $P(B) = 1$.

Доказательство

1. $P(B) = P(\prod\limits_{N=1}^{\infty} \sum\limits_{n=N}^{\infty} A_n) \stackrel{*}{=}$

т.к. $\sum\limits_{n=N+1}^{\infty} A_n \subset \sum\limits_{n=N}^{\infty} A_n$, то *по аксиоме непрерывности* $\stackrel{*}{=} \lim\limits_{N\to\infty} P(\sum\limits_{n=N}^{\infty} A_n) \leqslant \lim\limits_{N\to\infty} \sum\limits_{n=N}^{\infty} P(A_n) \xrightarrow{N\to\infty} 0$, т.к. $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$

2. $P(B) = \dots = \lim\limits_{N\to\infty} P(\sum\limits_{n=N}^{\infty} A_n) = \lim\limits_{N\to\infty} (1 - P(\prod\limits_{n=N}^{\infty} \overline{A_n})) = 1 - \lim\limits_{N\to\infty} P(\prod\limits_{M=N}^{\infty} \prod\limits_{n=N}^{M} \overline{A_n}) \stackrel{*}{=}$

т.к. $\prod\limits_{n=N}^{M+1} \overline{A_n} \subset \prod\limits_{n=N}^{M} \overline{A_n}$, то по *аксиоме непрерывности*

$\stackrel{*}{=} 1 - \lim\limits_{N\to\infty} \lim\limits_{M\to\infty} P(\prod\limits_{n=N}^{M} \overline{A_n}) = 1 - \lim\limits_{N\to\infty} \lim\limits_{M\to\infty} \prod\limits_{n=N}^{M} (1 - P(A_n)) = 1 - \lim\limits_{N\to\infty} \lim\limits_{M\to\infty} \prod\limits_{n=N}^{M} e^{\ln(1-P(A_n))} = 1 - \lim\limits_{N\to\infty} \lim\limits_{M\to\infty} e^{\sum n=N^M \ln(1-P(A_n))} \stackrel{*}{\geqslant}$

т.к. $\ln(1-t) = -t - \frac{t^2}{2} - \frac{t^3}{3} - \frac{t^4}{4} - \dots < -t$, то

$\stackrel{*}{\geqslant} 1 - \lim\limits_{N\to\infty} \lim\limits_{M\to\infty} e^{-\sum\limits_{n=N}^{M} P(A_n)} = 1 - \lim\limits_{n\to\infty} 0 = 1.$ ■

3.16 Лемма 4

Пусть $X_n \xrightarrow{P} X$,

$$\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty$$

Тогда $X_n \xrightarrow{\text{п.н.}} X$

Доказательство

В силу Теоремы 3.1 $\forall\varepsilon > 0$

$$P(\prod_{N=1}^{\infty}\sum_{n=N}^{\infty}\{\omega : |X_n - X| > \varepsilon\}) = 0$$

Тогда

$P(\{\omega : \lim\limits_{n\to\infty} X_n X_n(\omega) \neq X(\omega)\}) = P(\{\omega : \exists\varepsilon > 0, \forall N \in \mathbb{N}\exists n \geqslant N : |X_n(\omega) - X(\omega)| > \varepsilon\}) = P(\{\omega : \exists M \in \mathbb{N}, \forall N \in \mathbb{N}\exists n \geqslant N : |X_n(\omega) - X(\omega)| > \frac{1}{M}\}) = P(\sum\limits_{M=1}^{\infty}\prod\limits_{N=1}^{\infty}\sum\limits_{n=N}^{\infty}\{\omega : |X_n - X| > \frac{1}{M}\}) \leqslant \sum\limits_{M=1}^{\infty} P(\prod\limits_{N=1}^{\infty}\sum n = N^{\infty}\{\omega : |X_n - X| > \frac{1}{M}\}) = 0.$ ■

Замечание

1.

$X_n \xrightarrow{\text{с.к.}} X \implies X_n \xrightarrow{P} X$

ИЛИ

$X_n \xrightarrow{\text{п.н.}} X \implies X_n \xrightarrow{P} X$

2. В силу *теоремы Рисса* (функциональный анализ) если $X_n \xrightarrow{P} X$, то существует *подпоследовательность* $\{X_{n_k}\}_{k\in\mathbb{N}} : X_{n_k} \xrightarrow[k\to\infty]{\text{п.н.}} X$

3. В силу *теоремы о мажорирующей сходимости*, если $X_n \xrightarrow{P} X$ и $\exists Y$ – с.в.: $|X_n| \leqslant Y, M[Y^2] < \infty$, то $X_n \xrightarrow{\text{с.к.}} X$

4. Также из функционального анализа известно, что операция предела (по мере, почти наверное, в средне квадратическом) замкнута относительно линейных операций и непрерывных преобразований.

4 Закон больших чисел

4.1 Определение 1

Выборкой объема n будем называть с.в. $Z_n = (X_1, \dots, X_n)^T$, где $X_1, \dots, X_n$ – *независимые* с.в.

Через $F_k(x)$ обозначим *функцию распределения* k-го элемента выборки.

Если $F_k = F_1, k = \overline{2, n}$, то выборка называется ***однородной***.

4.2 Определение 2

Выборочным средним $\overline{X_n}$ выборки Z_n называется $\overline{X_n} = \frac{1}{n} \sum\limits_{k=1}^{n} X_k$

4.3 Теорема 1 (Закон Больших Чисел Чебышёва)

Пусть Z_n – однородная выборка, $M[X_k^2] < \infty$.

Тогда $\overline{X_n} \xrightarrow{\text{с.к.}} m_X, \overline{X_n} \xrightarrow{P} m_X$

Доказательство

$M[|\overline{X_n} - m_X|^2] = M[|\overline{X_n} - M[\overline{X_n}]|^2|^2] = D[\overline{X_n}] = \frac{1}{n^2} D[\sum\limits_{k=1}^{n} X_k] = \frac{nD[X_1]}{n^2} \overset{n \to \infty}{\longrightarrow} 0$

Т.е. по определению $\overline{X_n} \xrightarrow{\text{с.к.}} m_X$.

С учетом Леммы 3.3 $\overline{X_n} \xrightarrow{P} m_X$. ■

4.4 Теорема 2 (Закон Больших Чисел Колмогорова)

Пусть Z_n – *однородная* выборка, $M[X_k] = m_X < \infty$.

Тогда $\overline{X_n} \xrightarrow{\text{п.н.}} m_X$

Замечание

Т.о. для *однородной* выборки $\overline{X_n}$ сходится **почти наверное** и **по вероятности** к m_X, если оно существует, и **в среднем квадратичном**, если существует дисперсия.

4.5 Теорема 3

Пусть Z_n – *неоднородная* выборка, $M[X_k] = m_X < \infty, D[X_k] = D_k \leqslant D_{max} < \infty$, где $k \in \mathbb{N}$

Тогда $\overline{X_n} \xrightarrow{\text{с.к.}} m_X, \overline{X_n} \xrightarrow{P} m_X$.

Доказательство

$M[|\overline{X_n} - m_X|^2] = M[|\overline{X_n} - M[\overline{X_n}]|^2] = D[\overline{X_n}] = \frac{1}{n^2}\sum\limits_{k=1}^{n} D_k \leqslant \frac{nD_{max}}{n^2} = \frac{D_{max}}{n} \xrightarrow{n\to\infty} 0$

Т.о. $\overline{X_n} \xrightarrow{\text{с.к.}} m_X$. Тогда в силу Леммы 3.3 $\overline{X_n} \xrightarrow{P} m_X$. ■

4.6 Следствие 1

Пусть Z_n — *неоднородная* выборка, $D[X_k] = D_k \leqslant D_{max} < \infty, k \in \mathbb{N}$.
Тогда $\overline{X_n} - \frac{1}{n}\sum\limits_{k=1}^{n} m_{X_k} \xrightarrow{\text{с.к.}} 0, \overline{X_k} - \frac{1}{n}\sum\limits_{k=1}^{n} m_{X_k} \xrightarrow{P} 0.$

Доказательство

$\overline{X_n} - \frac{1}{n}\sum\limits_{k=1}^{n} m_{X_k} = \frac{1}{n}\sum\limits_{k=1}^{n}(X_k - m_{X_k}) = \frac{1}{n}\sum\limits_{k=1}^{n} Y_k,$
где $M[Y_k] = 0, D[Y_k] = D[X_k] =\leqslant D_{max} < \infty$
Тогда $\overline{Y_k}$ удовлетворяет условиям Теоремы 4.3. ■

4.7 Теорема 4

Пусть Z_n — *неоднородная* выборка, $M[X_k] = m_X < \infty, D[X_k] = D_k < \infty,$

$$\sum_{k=1}^{n} \frac{D[X_k]}{k^2} < \infty$$

Тогда

$$\overline{X_n} \xrightarrow{\text{п.н.}} m_X$$

Замечание

Условие Теоремы 4 более мягкое, чем условие Теоремы 3. Пусть $D[X_k] \leqslant D_{max}, k \in \mathbb{N}$. Тогда

$$\sum_{k=1}^{n} \frac{D[X_k]}{k^2} \leqslant \sum_{k=1}^{n} \frac{D_{max}}{k^2} = \frac{\pi^2 D_{max}}{\sigma} < \infty$$

4.8 Следствие 2

Пусть $N(A)$ — число появления события A в серии из N *независимых* опытов. Тогда

$$\frac{N(A)}{N} \xrightarrow{\text{п.н.}} P(A), \frac{N(A)}{N} \xrightarrow{\text{с.к.}} P(A)$$

Доказательство

По условию $N(A) \sim Bi(N; P(A))$. Тогда $\exists X_1, \dots, X_n \sim Be(P(A))$ — независимые с.в.
При этом $M[X_1] = P(A), D[X_1] = P(A)(1 - P(A)) \leqslant \frac{1}{4}$.
Тогда в силу Теоремы 4

$$\frac{N(A)}{N} = \frac{1}{N} \sum_{k=1}^{N} X_k \xrightarrow{\text{с.к.}} P(A)$$

в силу Теоремы 2

$$\frac{N(A)}{N} = \frac{1}{N} \sum_{k=1}^{N} X_k \xrightarrow{\text{п.н.}} P(A)$$

∎

5 Центральная предельная теорема (ЦПТ)

Замечание

Сходимости ***с.к.***, ***п.н.*** и P в общем случае исследования предполагают либо знание совместного распределения элементов последовательности, либо наличие точкой функциональной зависимости от $\omega \in \Omega$.

Как правило, в теории вероятностей это неизвестно, а с.в. описываются при помощи их распределений, а не как функции. При этом если у двух величин совпадают распределения, то это вовсе не значит, что они равны.

Поэтому довольно важным является вид сходимости *по распределению*, т.е. в смысле **“описательного инструмента” с.в.

5.1 Определение 1

Говорят, что последовательность с.в. $\{X_n\}_{n\in\mathbb{N}}$ ***сходится по распределению*** к с.в. X, если

$F_{X_n}(x) \overset{n\to\infty}{\longrightarrow} F_X(x), \forall x -$ точки непрерывности $F_X(x)$.

5.2 Лемма 1

Пусть $X_n \overset{P}{\longrightarrow} X$. Тогда $X_n \overset{d}{\longrightarrow} X$

5.3 Лемма 2

Пусть $X_n \overset{d}{\longrightarrow} X, Y_n \overset{P}{\longrightarrow} C$

Тогда $X_n + Y_n \overset{d}{\longrightarrow} C$

Доказательство

Пусть $C = 0$, $x_0 -$ точка **непрерывности** $F_X(x)$. Тогда $\forall \varepsilon > 0$

$F_{X_n+Y_n}(x_0) = P(X_n+Y_n \leqslant x_0) = \underbrace{P(\{X_n + Y_n \leqslant x_0\}\{|Y_n| > \varepsilon\})}_{p_1} + \underbrace{P(\{X_n + Y_n \leqslant x_0\}\{|Y_n| \leqslant \varepsilon\})}_{p_2}$

$$0 \leqslant p_1 \leqslant P(|Y_n > \varepsilon) \overset{n\to\infty}{\longrightarrow} 0$$

$p_2 = P(\{X_n + Y_n \leqslant x_0\}\{-\varepsilon \leqslant Y_n \leqslant \varepsilon\}) \overset{*}{\leqslant}$

$$\begin{cases} X_n + Y_n \leqslant x_0 \\ -\varepsilon \leqslant Y_n \leqslant \varepsilon \end{cases} \implies \begin{cases} X_n + Y_n \leqslant x_0 \\ -\varepsilon \leqslant -Y_n \leqslant \varepsilon \end{cases} \implies X_n \leqslant x_0 + \varepsilon$$

$\overset{*}{\leqslant} P(X_n \leqslant x_0 + \varepsilon) = F_{X_n}(x_0 + \varepsilon)$

$p_2 \geqslant P(\{\varepsilon + X_n \leqslant x_0\}\{-\varepsilon \leqslant Y_n \leqslant \varepsilon\}) \overset{*}{\geqslant}$

$$P(AB) = P(A) - P(A \backslash B) \geqslant P(A) - P(\overline{B})$$

$\overset{*}{\geqslant} P(\varepsilon + X_n \leqslant x_0) - P(|Y_n > \varepsilon) = F_{X_n}(x_0 - \varepsilon) - P(|Y_n| > \varepsilon)$
Т.о. $F_{X_n}(x_0 - \varepsilon) - P(|Y_n| > \varepsilon) \leqslant F_{X_n + Y_n}(x_0) \leqslant F_{X_n}(x_0 + \varepsilon) + p_1$
Выберем $\varepsilon > 0$ так, чтобы $(x_0 - \varepsilon; x_0 + \varepsilon)$ было областью непрерывности $F_X(x)$
Тогда $\lim\limits_{n\to\infty} F_{X_n}(x_0 \pm \varepsilon) = F_X(x_0 \pm \varepsilon)$

$$F_{X_n}(x_0 - \varepsilon) \leqslant \underline{\lim}_{n\to\infty} F_{X_n+Y_n}(x_0) \leqslant \overline{\lim}_{n\to\infty} F_{X_n+Y_n}(x_0) \leqslant F_X(x_0 + \varepsilon)$$

Возьмем предел по $\varepsilon \to 0$. В силу непрерывности $F_X(x)$ в x_0 $\lim\limits_{\varepsilon\to 0} F_X(x_0 \pm \varepsilon) = F_X(x_0)$
Откуда $F_X(x_0) = \lim\limits_{\varepsilon\to 0} F_{X_n+Y_n}(x_0)$, т.е.

$$X_n + Y_n \xrightarrow{d} X$$

Пусть $C \neq 0$. Тогда $Y_n - C = \tilde{Y} \xrightarrow{P} 0$,
$X_n + C = \tilde{X_n} \xrightarrow{d} X + C = \tilde{X}$
Получаем $X_n + Y_n = \tilde{X_n} + \tilde{Y_n} \xrightarrow{d} \tilde{X} = X + C$. ■

5.4 Доказательство леммы 5.1

$X_n = (X_n - X) + X$, где $X \xrightarrow{d} X$, $X_n - X \xrightarrow{P} 0$. В силу Леммы 5.2 $X_n = (X_n - X) + X \xrightarrow{d} X + 0 = X$. ■

Замечание

Из *теории преобразования Фурье* следует, что $X_n \xrightarrow{d} X$ тогда и только тогда, когда $\psi_{X_n}(\lambda) \xrightarrow{n\to\infty} \psi_X(\lambda), \lambda \in \mathbb{R}$

5.5 Определение 2

Последовательность с.в. $\{X_n\}_{n\in\mathbb{N}}$ называется ***асимптотически нормальной***, если
$X_n \xrightarrow{d} X$, где $X \sim N(m; \sigma^2)$

5.6 Теорема 1 (Центральная предельная)

Пусть $\{X_n\}_{n\in\mathbb{N}}$ — последовательность *независимых* и *одинаково распределенных* с.в., причем

$$M[X_1] = m_X, D[X_1] = \sigma_X^2$$

Тогда

$$S_n \frac{\sum\limits_{k=1}^{n} X_n - nm_X}{\sigma_X \sqrt{n}} \xrightarrow{d} N(0; 1)$$

Доказательство

Обозначим через $Y_k = \frac{X_k - m_X}{\sigma_X \sqrt{n}}$

Тогда $\sum\limits_{k=1}^{n} Y_k$, где $Y_1, \dots, Y_n$ — *независимые* с.в.

В силу Леммы 1.5

$$\psi_{S_n}(\lambda) = \psi_Y(\lambda, \lambda, \dots, \lambda) = \prod_{k=1}^{n} \psi_{Y_k}(\lambda) = \psi_{Y_1}^n(\lambda) = \psi^n(\frac{\lambda}{\sigma_X \sqrt{n}})$$

где $\psi(\lambda)$ — **характеристическая функция** $X_k - m_X$.

$\psi(0) = M[e^{i0(X_k - m_X)}] = M[1] = 1$

$\psi'(0) = M[i(X_k - m_X)e^{i0(X_k - m_X)}] = M[X_k - m_X]i = 0$

$\psi''(0) = -M[(X_k - m_X)^2 e^{i0(X_k - m_X)}] = -M[(X_k - m_X)^2] = -\sigma_X^2$

Тогда согласно *формуле Тейлора*

$\psi(\lambda) = \psi(0) + \psi'(0)\lambda + \psi''(0)\frac{\lambda^2}{2} + 0(\lambda^2) = 1 - \frac{\sigma_X^2}{2}\lambda^2 + 0(\lambda^2)$

Рассмотрим $\ln \psi_{S_n}(\lambda) = n \ln \psi(\frac{\lambda}{\sigma_X \sqrt{n}}) = n \ln(1 - \frac{\lambda^2}{2n} + 0(\frac{\lambda^2}{\sigma_X^2 n})) = n(-\frac{\lambda^2}{2n} + 0(\frac{\lambda^2}{\sigma_X^2 n}) + 0(-\frac{\lambda^2}{2n} + 0(\frac{\lambda^2}{\sigma_X^2 n}))) = n(-\frac{\lambda^2}{2n} + 0(\frac{\lambda^2}{\sigma_X^2 n})) = -\frac{\lambda^2}{2n} + \frac{0(\frac{\lambda^2}{\sigma_X^2 n})}{\frac{\lambda^2}{\sigma_X^2 n}} \cdot \frac{\lambda^2}{2} \xrightarrow{n \to \infty} -\frac{\lambda^2}{2}$,

$$\psi_{S_n}(\lambda) \xrightarrow{n \to \infty} e^{-\frac{\lambda^2}{2}}$$

где $\psi_Y(\lambda) = e^{-\frac{\lambda^2}{2}}$ по определению является *характеристической функцией $Y \sim N(0;1)$

Тогда $S_n \xrightarrow{d} Y \sim N(0;1)$. ■

5.7 Следствие 1 (Теорема Муавра - Лапласа)

Пусть $X_n \sim Bi(n;p)$

Тогда

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0;1)$$

Доказательство

Т.к. $X_n \sim Bi(n;p)$, то существуют **независимые** $\tilde{X}_1, \dots, \tilde{X}_n \sim Be(p)$ такие, что

$$X_n = \sum_{k=1}^{n} \tilde{X}_k, M[\tilde{X}_k] = p = m_X, D[\tilde{X}_k] = p(1-p) = \sigma_X^2$$

Тогда в силу Теоремы 1:

$\frac{\sum\limits_{k=1}^{n} \tilde{X}_k - n m_X}{\sigma_X \sqrt{n}} \xrightarrow{d} N(0;1)$. ■

5.8 Пример 1

Вычислить вероятность того, что при $n = 1000$ подбрасываниях монета упадет *"орлом"* от 400 до 600 раз.

Пусть X — число выпавших *"орлов"*. Тогда $X \sim Bi(1000; \frac{1}{2})$. По *формуле Бернулли*Ж

$$P(X \in [400; 600]) = \sum_{k=400}^{600} C_{1000}^k \frac{1}{2^{1000}}$$

Оценим данную величину с помощью *ЦПТ*.

В силу Теоремы 1 и Следствия

$$\frac{X - n\frac{1}{2}}{\sqrt{n\frac{1}{2}(1-\frac{1}{2})}} \overset{n\to\infty}{\longrightarrow} N(0;1)$$

Тогда

$P(400 \leqslant X \leqslant 600) = P(\frac{400-500}{\sqrt{250}} \leqslant \frac{X-nm_X}{\sigma_X\sqrt{n}} \leqslant \frac{600-500}{\sqrt{250}}) \approx \Phi_0(\frac{100}{5\sqrt{10}}) - \Phi_0(-\frac{100}{5\sqrt{10}}) = 2\Phi_0(2\sqrt{10}) \approx 1.$

5.9 Теорема 2 (Ляпунова)

Пусть $\{X_n\}_{n\in\mathbb{N}}$ — последовательность *независимых* с.в., $M[X_n] = m_{X_n}, D[X_n] = \sigma^2_{X_n}, M[|X_n - m_{X_n}|^3] = C_n^3 < \infty$

При этом $\frac{(\sum\limits_{k=1}^{n} C_k^3)^{\frac{1}{3}}}{(\sum\limits_{k=1}^{n} \sigma^2_{X_n})^{\frac{1}{2}}} \overset{n\to\infty}{\longrightarrow} 0$ ***(Условие Ляпунова)***

Тогда $\frac{(\sum\limits_{k=1}^{n} X_k - M[\sum\limits_{k=1}^{n} X_k]}{\sqrt{D[\sum\limits_{k=1}^{n} X_k]}} \xrightarrow[n\to\infty]{d} \sim N(0;1)$

Замечание

Для *аппроксимации* точности использования ЦПТ используется **неравенство Берри-Эссеена**

5.10 Теорема 3 (Неравенство Берри-Эссеена)

Пусть $\{X_n\}_{n\in\mathbb{N}}$ — последовательность *независимых* и *одинаково распределенных* с.в.

$$M[X_n] = m_X, D[X_n] = \sigma^2_X, M[|X_n - m_X|^3] = \rho < \infty$$

Тогда $\forall x \in \mathbb{R}$ и $n \in \mathbb{N}$

$$|P(\frac{\sum\limits_{k=1}^{n} X_k - nm_X}{\sigma_X\sqrt{n}}) - (\frac{1}{2} + \Phi_0(x))| \leqslant \frac{C_0\rho}{\sigma_X^3\sqrt{n}}$$

Замечание

Точное значение константы C **неизвестно**.
По текущим данным *(2010 г.)* $C_0 \leqslant 0.4784$

5.11 Пример

Оценим точность решения в предыдущем примере: $\sigma_X^2 = \frac{1}{4}, n = 1000, m_X = \frac{1}{2}, X_n \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

$\rho = M[|X_n - m_X|^3] = |0 - \frac{1}{2}|^3 \cdot \frac{1}{2} + |1 - \frac{1}{2}|^3 \cdot \frac{1}{2} = \frac{1}{8}$

Тогда погрешность составит для двухстороннего нер-ва:

$\frac{2C_0\rho}{\sigma_X^3\sqrt{n}} = \frac{2 \cdot 0.4784 \cdot \frac{1}{8}}{\frac{1}{64} \cdot \sqrt{1000}} \approx 0.03$

6 Выборка и ее характеристики

6.1 Определение 1

Выборкой называется $Z_n = (X_1, \dots, X_n)^T$ *независимый вектор* с.в. Если все $X_1, \dots, X_n$ — *одинаково распределены*, а $F(x)$ — *функция распределения*, то говорят, что Z_n — ***однородная*** выбрка, порожденная распределением $F(x)$

6.2 Определение 2

Реализацией выборки $Z_n \in \mathbb{R}^n$ называется *неслучайный* вектор $z_n = Z_n(\omega)$, состоящий из реализаций элементов выборки $X_k, k = \overline{1, n}$.

6.3 Определение 3

Множество S всех возможных реализаций выборки Z_n называют ***выборочным пространством***

Замечание

Обычно распределение, порождающее выбрку, известно неточно.

$$F_X = F_X(x; \theta)$$

Задача состоит в построении оценки θ по элементам выборки.

6.4 Определение 4

С.в. $\phi(Z_n)$, где $\phi : S \to \mathbb{R}$ — *измерима*, называется ***статистикой***.

6.5 Определение 5

k-ой порядковой статистикой называется k-е по величине значение элемента выборки $Z_n = (X_1, \dots, X_n)^T$ и обозначается $X^{(k)}$

Замечание

$X^{(k)}$ является функцией от всей выборки, т.к. при различных $\omega \in \Omega X^{(k)}$ будет совпадать по значению с разными X_i.

6.6 Определение 6

Набор порядковых статистик $X^{(1)}, \dots, X^{(n)}$ называется ***вариационным рядом***.

6.7 Определение 7

$X^{(1)} = \min_{k=\overline{1,n}} X_k, X^{(n)} = \max_{k=\overline{1,n}} X_k$.

6.8 Лемма 1

Пусть однородная выборка Z_n порождена распределением $F(x)$. Тогда функция распределения $X^{(k)}$ имеет вид:

$$F_{(k)}(x) = P(X^{(k)} \leqslant x) = \sum_{i=k}^{n} C_n^i (F(x))^i (1 - F(x))^{n-i}$$

Доказательство

Рассмотрим с.в. Y, равную числу элементов выборки, не превосxодящих x. Тогда $Y \sim Bi(n; F(x))$.

$F_{(k)}(x) = P(X^{(k)} \leqslant x) = P(Y \geqslant k) = \sum\limits_{i=k}^{n} C_n^i (F(x))^i (1 - F(x))^{n-i}$. ■

6.9 Следствие 1

$F_{(1)}(x) = 1 - (1 - F(x))^n$,
$F_{(n)}(x) = (F_{(x)})^n$.

6.10 Определение 8

Выборочной квантилью уровня $\alpha \in (0; 1)$ называется *порядковая статистика* $X^{([nd]+1)}$

6.11 Теорема 1 (Мостеллера)

Пусть X — *абсолютно непрерывная* с.в., x_α — *точка гладкости* $f_X(x), f_X(x_\alpha) > 0$
Тогда $(X^{([nd]+1)} - x_\alpha)\sqrt{\frac{n f_X^2(x_\alpha)}{p(1-p)}} \xrightarrow{d} N(0; 1)$

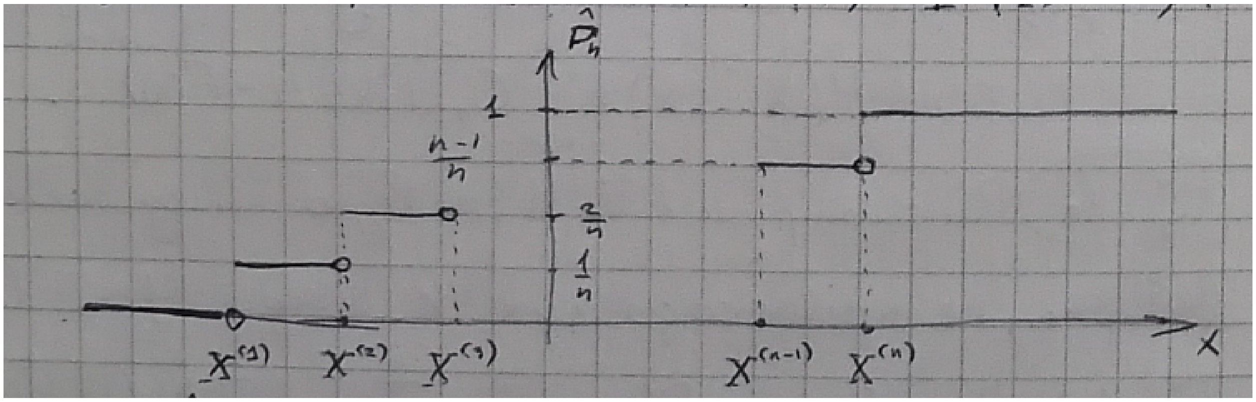
6.12 Определение 9

Выборочной функцией распределения называется статистика $\hat{F}_n(x)$:

$$\hat{F}_n(x) = \begin{cases} \frac{1}{n}\max\{k = \overline{1, n} : X^{(k)} \leqslant x\}, & x \geqslant X^{(1)}, \\ 0, & x < X^{(1)} \end{cases}$$

Замечание

Фактически $\hat{F}_n(x)$ — **частота события** $\{X \leqslant x\}$, которая исользуется для оценки вероятности $F(x) = P(X \leqslant x)$.



6.13 Свойства $\hat{F}_n(x)$

1. $\boxed{n \cdot \hat{F}_n(x) \sim Bi(n; F(x))}$

2. $\boxed{M[\hat{F}_n(x)] = F(x)}$

3. $\boxed{\sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)| \xrightarrow{\text{п.н.}} 0}$

   (Теорема Гливенко - Кантелли)

4. $\boxed{[(\hat{F}_n(x) - F(x))^2] = \frac{F(x)(1-F(x))}{n} \leqslant \frac{1}{4n}}$

5. $\boxed{|\hat{F}_n(x) - F(x)| \xrightarrow{\text{с.к.}} 0}$

6. $\boxed{\frac{\hat{F}_n(x) - F(x)}{\sqrt{F(x)(1-F(x))}} \sqrt{n} \xrightarrow{d} N(0; 1)}$

   (Следует из теоремы Муавра - Лапласа)

Замечание

$\hat{F}_n(x)$ При увеличении n *равномерно приблежается* к $F(x)$, при этом точность приближения можно оценить при помощи свойств **4** и **6**.

Гистрограмма. На основе реализации вариационного ряда построим разбиение $\mathbb{R}$ $-\infty = t_0 < t_1 < t_2 < \dots < t_l < t_{l+1} = +\infty$,

$t_1 \leqslant x^{(1)}, t_l > x(n)$.

Как правило, длина интервалов разбиения выбирается *одинаковой*:

$$h_k = t_{k+1} - t_k = \frac{t_l - t_1}{l - 1}, k = \overline{1, l-1}$$

Вычислим частоту попадания элементов выборки в k-й интервал:

$\hat{p}_k = \frac{n_k}{n}$, где n_k — число элементов выборки, попавших в $[t_k; t_{k+1}), k = \overline{0, l}$

Заметим, что $\hat{p}_0 = \hat{p}_l = 0$.

6.14 Определние 10

Гистограммой называется функция:

$$\hat{f}_n(x) = \begin{cases} 0, & x \in (t_0; t_1) \cup [t_l; t_{l+1}), \\ \hat{p}_k, & x \in (t_k; t_{k+1}), k = \overline{1, l-1} \end{cases}$$

Замечание

Если плотность вероятности $f_X(x)$ *непрерывна* и *ограничена*, а число разрядо гистограммы l_n *удовлетворяет условию*: $l_n \longrightarrow +\infty$, $\frac{n}{l_n} \longrightarrow +\infty$,

то

$$\hat{f}_n(x) \xrightarrow{P} f_X(x)$$

Т.е. гистограмма является статистической аппроксимацией функции плотности вероятности.

Выборочные моменты

6.15 Определние 1

Выборочным начальным и ***центральным моментами*** называется соответственно статистики:

$\hat{\nu_r}(n) = \frac{1}{n} \sum\limits_{k=1}^{n} X_k^r$

И

$\hat{\mu_r}(n) = \frac{1}{n} \sum\limits_{k=1}^{n} (X_k - \hat{\nu}_1(n))^r$

6.16 Определение 2

Выборочным средним и ***выборочой дисперсией*** называются соответственно статистики:

$\overline{X_n} = \hat{\nu_1}(n) = \frac{1}{n} \sum\limits_{k=1}^{n} X_k$

И

$\hat{d_X}(n) = \hat{S}^2(n) = \hat{\mu_2}(n) = \frac{1}{n} \sum\limits_{k=1}^{n} (X_k - \overline{X})^2$

6.17 Определение 3

Пусть $Z_n = (X_1, \dots, X_n)^T$ и $V_n = (Y_1, \dots, Y_n)^T$ – выборки, порожденные распределениями F_X и F_Y соответственно. Тогда ***выборочным коэффициентом корреляции*** называется:

$$r_{\hat{XY}} = \frac{\sum\limits_{k=1}^{n}(X_k - \overline{X_n})(Y_k - \overline{Y_k})}{n\sqrt{\hat{d_X} \cdot \hat{d_Y}}}$$

6.18 Свойства выборочных моментов

1. $\boxed{M[\hat{\nu_r}(n)] = \nu_r, r \in \mathbb{N}}$

Доказательство

$M[\hat{\nu_r}(n)] = \frac{1}{n}\sum\limits_{k=1}^{n} M[X_k^r] = \frac{1}{n}n\nu_r$. ■

2. $\boxed{\text{Если } M[X^r] < \infty, \text{то } \hat{\nu_r}(n) \xrightarrow{\text{п.н.}} \nu_r}$

Доказательство

3. $\boxed{\text{Если } M[X^r] < \infty, \text{то } \hat{\mu_r}(n) \xrightarrow{\text{п.н.}} \mu_r}$

Доказательство

$\hat{\mu_r}(n) = \frac{1}{n}\sum\limits_{k=1}^{n}(X_k - \overline{X_n})^r = \frac{1}{n}\sum\limits_{k=1}^{n}\sum\limits_{i=0}^{r} C_r^i X_k^i(-\hat{X_n})^{r-i} = \sum\limits_{i=0}^{r} C_r^i(-\hat{X_n})^{r-i}(\frac{1}{n}\sum\limits_{k=1}^{n} X_k^i) = \sum\limits_{i=0}^{r} C_r^i(-\hat{X_n})^{r-i}\hat{\nu_i}(n) \underset{\text{св-во (2)}}{\xrightarrow{\text{п.н.}}} \sum\limits_{i=0}^{r} C_r^i(-\nu_1)^{r-i}\hat{\nu_i} = \dots = \mu_i$. ■

4. $\boxed{D[\overline{X_n}] = \frac{1}{n}D[X]}$

Доказательство

С учетом *независимости* $X_1, \dots, X_n$

$D[X\overline{X_n}] = D[\frac{1}{n}\sum\limits_{k=1}^{n} X_k] = \frac{1}{n^2}\sum\limits_{k=1}^{n} D[X_k] = \frac{nD[X]}{n^2}$. ■

5. $\boxed{M[\hat{d_X}(n)] = \frac{n-1}{n}D[X]}$

Доказательство

$M[\hat{d_X}(n)] = \frac{1}{n}\sum\limits_{k=1} nM[(X_k - \overline{X_n})^2] = M[(X_1 - \overline{X_n})^2] = D[X_1 - \overline{X_n})^2] = D[\frac{n-1}{n}X_1 - \sum\limits_{k=2} n\frac{1}{n}\overline{X_k}] \overset{\text{н / з}}{=} \frac{n-1}{n}D[X_1] + \frac{1}{n^2}\sum\limits_{k=2}^{n} D[X_k] = D[X]\frac{n^2-2n+1+n-1}{n^2} = D[X] \cdot \frac{n-1}{n}$. ■

6. $\boxed{\frac{\overline{X_n} - m_X}{\sigma_X}\sqrt{n} \xrightarrow{d} N(0;1)}$

Доказательство

Т.к. $X_1, \dots, X_n$ *независимые*, $M[X_k] = m_k, D[X_k] = \sigma_X^2, k \in \mathbb{N}$, то в силу Теоремы 5.1:

$$\frac{\sum\limits_{k=1}^{n} X_k - n m_X}{\sigma_X \sqrt{n}} = \frac{\frac{1}{n}\sum\limits_{k=1}^{n} X_k - m_X}{\frac{\sigma_X}{\sqrt{n}}} \xrightarrow{d} N(0;1)$$

■

7. $\boxed{\frac{\hat{d_X}(n) - \sigma_X^2}{\sqrt{\mu_4 - \mu_2}} \sqrt{n} \xrightarrow{d} N(0;1)}$