

Лекции  
по курсу *”Математическая Статистика”*  
Ибрагимов Д. Н.

# Contents

<b>Источники</b>	<b>3</b>
<b>1 Многомерное нормальное распределение</b>	<b>4</b>
Замечание	4
1.1 Лемма 1	4
1.2 Определение 1	4
1.3 Лемма 2	4
1.4 Лемма 3	5
1.5 Лемма 4	5
note	5
1.6 Определение 2	5
Замечание	5
1.7 Лемма 5	6
Доказательство	6
Замечание	6
1.8 Определение 3	6
1.8.1 Доказательство леммы 3	7
1.8.2 Доказательство леммы 4	7
Замечание	7
<b>2 Теорема о нормальной корреляции</b>	<b>9</b>
2.1 Определение 1	9
2.2 Основные свойства условного М. О.	9
2.2.1 Свойство 1	9
2.2.2 Свойство 2	9
2.2.3 Свойство 3	9
2.2.4 Свойство 4	9
2.2.5 Свойство 5	9
2.3 Лемма 1	10
Доказательство	10
Замечание	10
2.4 Определение 2	10
2.5 Определение 3	10
2.6 Теорема 1	10
Доказательство	11
2.7 Теорема 2 (О нормальной корреляции)	11
Доказательство	11
Замечание	12
<b>3 Виды сходимости последовательностей случайных величин</b>	<b>13</b>
3.1 Определение 1	13
3.2 Определение 2	13
3.3 Определение 3	13

---

3.4	Определение 4 . . . . .	13
-----	-------------------------	----

## 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.  $K_x \geq 0$ , т.е.  $\forall x \in \mathbb{R}^n \setminus \{0\}, x^T K_x x \geq 0$ ;
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:

$$f_x(x) = ((2\pi)^n \det K_x)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x - m_k)^T K_x^{-1}(x - m_x)\right\}$$

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  $K_x$  — ковариационная 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:

$$\begin{aligned} Y = AX + b &\sim N(m_y, K_y), \\ m_y &= Am_x + b, \\ K_y &= AK_x A^T. \end{aligned}$$

## 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:

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

## Замечание

*Характеристическая функция* определена для любого случайного вектора или с.в.

Если с.в. **дискретная**, то:

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

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

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

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:

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

### 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.  $\psi_X(\lambda) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} e^{i(\lambda_1 x_1 + \dots + \lambda_n x_n)} f_X(x_1, \dots, x_n) dx_1 \cdot \dots \cdot dx_n = \{ \mathbf{n}/3 \} = \int_{\mathbb{R}} e^{i\lambda_1 x_1} \cdot \dots \cdot e^{i\lambda_n x_n} \cdot f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n) dx_1 \cdot \dots \cdot dx_n = \int_{\mathbb{R}} e^{i\lambda_1 x_1} f_{X_1}(x_1) dx_1 \cdot \dots \cdot \int_{\mathbb{R}} e^{i\lambda_n x_n} f_{X_n}(x_n) dx_n = \prod_{k=1}^n \psi_{X_k}(\lambda_k)$

■

#### 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 [Лемма 5](#) (пункт 1)

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

■

## 1.8.2 Доказательство леммы 4

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

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

$$\begin{aligned}\psi_X(\lambda) &= \exp\{i\lambda_1 m_{X_1} + \dots + i\lambda_n m_{X_n} - \frac{1}{2} \lambda^T K_X \lambda\} = \exp\{i\lambda_1 m_{X_1} + \dots + i\lambda_n m_{X_n} - \\ &= \frac{1}{2} (\lambda_1^2 \sigma_{X_1}^2 + \dots + \lambda_n^2 \sigma_{X_n}^2)\} = \prod_{k=1}^n \exp\{i\lambda_k m_{X_k} - \frac{1}{2} \lambda_k^2 \sigma_{X_k}^2\} = \prod_{k=1}^n \psi_{X_k}(\lambda_k).\end{aligned}$$

Откуда с учетом [Лемма 5](#) (пункт 2)  $X_1, \dots, X_n$  — 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:

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

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

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

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 [Лемма 3](#):

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

$$K_Y = A K_X A^T = \Lambda^{-\frac{1}{2}} S^T K_X S \Lambda^{-\frac{1}{2}} = I,$$

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:*

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

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

### 2.2 Основные свойства условного М. О.

#### 2.2.1 Свойство 1

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

Доказательство

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

#### 2.2.2 Свойство 2

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

Доказательство

$$\begin{aligned} M[\phi(Y) X | Y] &= \int_{-\infty}^{+\infty} \phi(Y) X f_{X|Y}(x | Y) dx = \phi(Y) \int_{-\infty}^{+\infty} x f_{X|Y}(x | Y) dx = \\ &= \phi(Y) M[X | Y]. \blacksquare \end{aligned}$$

#### 2.2.3 Свойство 3

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

#### 2.2.4 Свойство 4

Пусть  $X, Y$  — независимые. Тогда  $\boxed{M[X | Y] = M[X]}$

Доказательство

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

#### 2.2.5 Свойство 5

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

**Доказательство**

$$\begin{aligned}
M[M[X \mid Y]] &= \int_{-\infty}^{+\infty} M[X \mid Y] f_Y(y) dy = \int_{-\infty}^{+\infty} f_Y(y) \int_{-\infty}^{+\infty} x f_{X|Y}(x \mid y) dx dy = \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \frac{f_Y(y) f_Z(x, y)}{f_Y(y)} dy dx = \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_Z(x, y) dy dx = \int_{-\infty}^{+\infty} x f_X(x) dx = M[X]. \blacksquare
\end{aligned}$$

**2.3 Лемма 1**

Пусть  $X, Y$  — случайные векторы с конечными вторыми моментами. Тогда:

$$M[(X - \hat{X})\phi(Y)] = 0$$

где  $\hat{X} = M[X \mid Y]$

**Доказательство**

$$\begin{aligned}
M[(X - \hat{X})\phi(Y)^T] &= M[X\phi(Y)^T] - M[M[X \mid Y]\phi(Y)^T] = \\
&= \text{по Свойство 2} = M[X\phi(Y)^T] - M[M[X\phi(Y)^T \mid Y]] = \text{по Свойство 5} = M[X\phi(Y)^T] - \\
&M[X\phi(Y)^T] = 0. \blacksquare
\end{aligned}$$

**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

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

**2.6 Теорема 1**

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

**Доказательство**

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

Поскольку по определению  $\tilde{X} - \hat{X} = \phi(Y)$ , то в силу [Лемма 1](#)  $M[(X - \hat{X})^T(\tilde{X} - \hat{X})] = 0$ .  
 $\stackrel{*}{=} M[|X - \hat{X}|^2] + M[|\hat{X} - \tilde{X}|^2] \geq M[|X - \hat{X}|^2]$ . ■

## 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

$$1. \text{Law}(X | Y) = N(\mu(Y), \Delta),$$

где

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

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

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

**Доказательство**

Рассмотрим линейное преобразование  $Y$ :

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

В силу [Лемма 1](#) (пункт 3)

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

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

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

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

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

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

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

Тогда в силу [Лемма 1](#) (пункт 2)  $X - \mu(Y)$  и  $Y$  независимы. Построим характеристическую функцию условного распределения  $X$  относительно  $Y$ :

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

$$\stackrel{*}{=} M[e^{i\lambda^T (X - \mu(Y))} | Y] \cdot M[e^{i\lambda^T \mu(Y)} | Y] = M[e^{i\lambda^T (X - \mu(Y))}] e^{i\lambda^T \mu(Y)} = \psi_{X - \mu(Y)}(\lambda) e^{i\lambda^T \mu(Y)} =$$

$$\exp\left\{-\frac{1}{2}\lambda^T \Delta \lambda\right\} \cdot \exp\left\{i\lambda^T \mu(Y) - \frac{1}{2}\lambda^T \Delta \lambda\right\}$$

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

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

Вычислим с.к. ошибку:

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

$$tr \Delta. \blacksquare$$

### Замечание

1. Из **Теорема 2 (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  $\forall \varepsilon > 0$ :

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

ИЛИ

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

#### 3.3 Определение 3

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

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

ИЛИ

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

#### 3.4 Определение 4

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

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