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 $X = (X_1, \ldots, X_n)^T$ называется ***случайным***, если $X_1, \ldots, X_n$ — *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:

$$K_x = \begin{pmatrix} \mathrm{cov}(X_1, X_1) & \ldots & \mathrm{cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \mathrm{cov}(X_n, X_1) & \ldots & \mathrm{cov}(X_n, X_n) \end{pmatrix}$$

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

$$X \sim N(m_x, K_x)$$

если совместная плотность вероятности имеет вид:

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

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 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 Лемма 5 *(пункт 1)*

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

■

1.8.2 Доказательство леммы 4

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

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

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

Откуда с учетом Лемма 5 *(пункт 2)* $X_1, \ldots, X_n$ — н/з.

Пусть $X_1, \ldots, X_n$ — н/з. Тогда $X_1, \ldots, X_n$ *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:

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

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

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

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:

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

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:

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

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

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

<u>Доказательство</u>

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

2.2.2 Свойство 2

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

<u>Доказательство</u>

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

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

<u>Доказательство</u>

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

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:
$M[(X - \hat{X})\phi(Y)] = 0$
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 Свойство 2 $= M[X\phi(Y)^T] - M[M[X\phi(Y)^T \mid Y]] =$ по Свойство 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 Лемма 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 Лемма 1 *(пункт 3)*

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

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

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

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

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

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

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

Тогда в силу Лемма 1 *(пункт 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 Лемма 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 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 $\forall\varepsilon > 0$:

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

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