Случайные векторы. Совместное распределение

октябрь–декабрь 2023 г.

Курс лекций «Теория вероятностей»

Оглавление

- Случайные векторы
- Совместное распределение дискретных случайных величин
- Совместное распределение непрерывных случайных величин
- Независимые случайные величины
- Ковариация, коэффициент корреляции

Случайные векторы

В пространстве $\mathbb{R}^k$ так же, как и на вещественной прямой, определяется борелевская сигма-алгебра, как наименьшая сигма-алгебра, содержащая все открытые множества пространства $\mathbb{R}^k$ (или все замкнутые множества, или все n-мерные прямоугольники и т. п.).

Для определенности полагаем, что вектор в $\mathbb{R}^k$ — вектор-столбец.

Транспонирование матрицы обозначается символом «штрих»: A'. В частности, $x = [x_1, \dots, x_k]' \in \mathbb{R}^k$ — вектор-столбец.

Пусть заданы случайные величины X_1, ..., X_k.

Определение
Упорядоченный набор $X = [X_1, \dots, X_k]'$ называется k-мерным случайным вектором.

Случайные векторы: двумерный случай

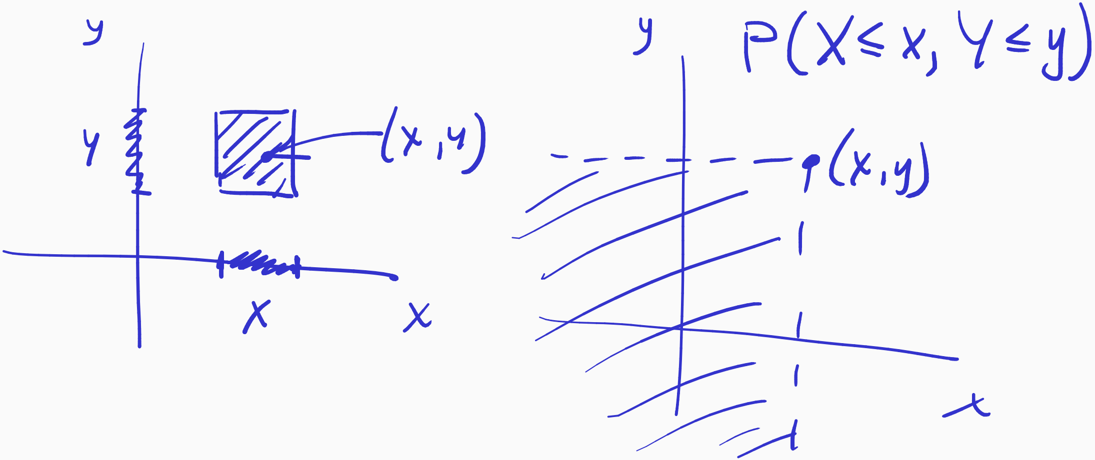
Для простоты ограничимся пока случаем $k = 2$ и будем рассматривать двумерные случайные векторы $U = [X, Y]'$.

Двумерный случайный вектор можно интерпретировать как точку на координатной плоскости. По аналогии с одномерным случаем возникает

Определение
Распределением случайного вектора U называется семейство вероятностей $\{P(U \in A), A \subset \mathbb{R}^2\}$, где A — измеримое подмножество двумерной плоскости, т. е. $A \in \mathcal{B}(\mathbb{R}^2)$.

Как и в одномерном случае рассмотрим отдельно совместное дискретное и непрерывное распределение.

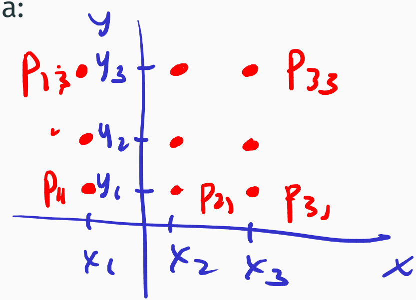
Для заметок



Дискретное совместное распределение

Пусть компоненты X и Y случайного вектора U дискретные и принимают значения $x_1, \ldots, x_m, \ldots$ и $y_1, \ldots, y_n, \ldots$ соответственно. Тогда совместное распределение — набор чисел $p_{ij} = P(X = x_i, Y = y_j) = P(\{X = x_i\} \cap \{Y = y_j\})$. При этом выполняются следующие свойства:

1. $p_{ij} \geq 0,\ i = 1, \ldots, m, \ldots,\ j = 1, \ldots, n, \ldots;$
2. $\sum\limits_i \sum\limits_j p_{ij} = 1.$



Дискретное совместное распределение: таблица

Подобно одномерному случаю эти числа можно объединить в таблицу:

$P(X=x_1, Y=y_1)$

$X \setminus Y$	y_1	…	y_n	…	
x_1	p_{11}	…	p_{1n}	…	$p_{1\bullet} = \sum_j p_{1j}$
⋮	⋮	⋱	⋮	⋱	⋮
x_m	p_{m1}	…	p_{mn}	…	$p_{m\bullet} = \sum_j p_{mj}$
⋮	⋮	⋱	⋮	⋱	⋮
	$p_{\bullet 1} = \sum_i p_{i1}$	…	$p_{\bullet n} = \sum_i p_{in}$	…	

$X = x_1,\ Y = y_j$

$j = 1, \dots n \dots$

распределение X, если не обращать внимания на Y.

Здесь $p_{i\bullet}$, $i = 1, \dots, m \dots$ и $p_{\bullet j}$, $j = 1, \dots, n, \dots$ — маржинальные распределения.

$P(X=x_1) = P(\{X=x_1, Y=y_1\} \cup \{X=x_1, Y=y_2\} \cup \dots) = \sum_j P(X=x_1, Y=y_j) = \sum_j p_{1j} = p_{1\bullet}$

Дискретное совместное распределение: маргинальные распределения

Возникает естественный вопрос: можно ли, зная маргинальные распределения, восстановить совместное распределение. Оказывается, в общем случае ответ отрицательный. Иными словами, можно построить два двумерных дискретных случайных вектора, у которых маргинальные распределения совпадают, а совместные распределения разные.

Упражнение
Постройте такие векторы.

$p_{i\cdot}$ – m штук $p_{\cdot j}$ – n штук

p_{ij} – неизвестные $\Rightarrow$ $n \cdot m$ штук.

$n+m$ ур-й

Для заметок

Пример:

−1 1 − Вратарь прыгает в левый или в правый угол.

−1 1 − Футболист бьёт пенальти в левый или правый угол.

1) Сит. оба независ. друг от друга выбирают 1 или −1

В \ Ф	−1	1	
−1	1/4	1/4	1/2
1	1/4	1/4	1/2
	1/2	1/2	

2) Вратарь знает, куда ударит фут.

В \ Ф	−1	1	
−1	1/2	0	1/2
1	0	1/2	1/2

3) Футболист знает, куда прыгнет В

В \ Ф	−1	1	
−1	0	1/2	1/2
1	1/2	0	1/2

Дискретное совместное распределение: пример

Пример

Уровень образования и доход: X — уровень образования: 1 — среднее, 2 — бакалавр, 3 — магистр, 4 — PhD; Y — ежемесячный доход: 1 — 500-, 2 — 500–1000, 3 — 1000–1500, 4 — 1500–2000, 5 — 2000+.

$EX \approx 2.4$

$X \setminus Y$	1	2	3	4	5	
1	0.1	0.06	0.04	0.02	0.01	0.23
2	0.08	0.09	0.07	0.03	0.05	0.32
3	0.04	0.1	0.08	0.06	0.06	0.34
4	0.01	0.02	0.04	0.02	0.02	0.11
	0.23	0.27	0.23	0.13	0.14	

Числа в ячейках есть вероятности соответствующих совместных событий. Например, число 0.06 в ячейке (1, 2) есть $P(X = 1, Y = 2) = 0.06$. Это же число можно интерпретировать и так: во всей рассматриваемой популяции 6% индивидуумов имеют среднее образование и доход от 500 до 1000.

Дискретное совместное распределение: математическое ожидание

Зная совместное распределение величин X и Y и имея функцию $g(x, y)$, можно сосчитать математическое ожидание $g(X, Y)$:

$$E(g(X, Y)) = \sum_i \sum_j g(x_i, y_j) p_{ij}.$$

марг. расп.

При этом если функция зависит только от x, то получится

$$E(g(X)) = \sum_i \sum_j g(x_i) p_{ij} = \sum_i g(x_i) \sum_j p_{ij} = \sum_i g(x_i) p_{i\bullet},$$

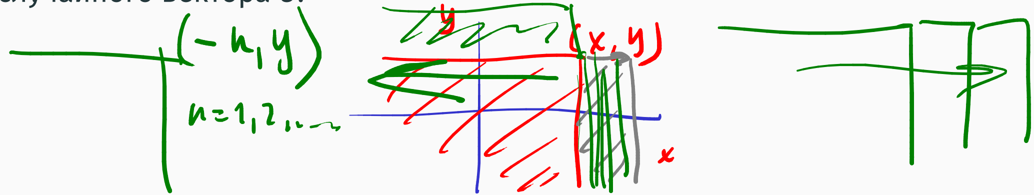
что совпадает с предыдущим определением. Для функции $g(y)$ аналогично.

Функция распределения совместного вектора

Как и раньше, определяется функция распределения случайного вектора $U = [X, Y]'$.

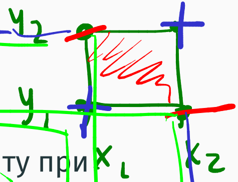
Определение

Функция $F_U(x, y) = F(u) = P(X \leqslant x, Y \leqslant y)$ называется функцией распределения случайного вектора U.



Функция распределения совместного вектора: свойства

В двумерном случае функция распределения обладает примерно теми же свойствами, что и в одномерном случае, а именно:



1. $0 \leqslant F(x, y) \leqslant 1$
2. функция $F(x, y)$ не убывает по каждому аргументу
3. функция $F(x, y)$ непрерывна справа по каждому аргументу при фиксированном другом аргументе
4. $\lim_{x \to -\infty} F(x, y) = 0$, $\lim_{y \to -\infty} F(x, y) = 0$, $\lim_{x \to +\infty, y \to +\infty} F(x, y) = 1$
5. пусть $x_1 \leqslant x_2$, $y_1 \leqslant y_2$, тогда $\boxed{F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)} \geqslant 0$

Действительно, левая часть неравенства (5) это вероятность того, что двумерный вектор попадает в прямоугольник, левая нижняя вершина которого (x_1, y_1), правая верхняя — (x_2, y_2). Как и в одномерном случае, по функции распределения однозначно восстанавливается распределение случайного вектора.

Функция распределения совместного вектора: производная

Предположим, что функция распределения $F(x, y)$ имеет непрерывную вторую смешанную производную

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Непрерывное совместное распределение: плотность

Определение

Функция $f(x,y) = \cancel{\frac{\partial^2 F(x,y)}{\partial x \partial y}}$ называется плотностью распределения вектора $U = [X,Y]'$.

Можно показать, что

$$\boxed{P(U \in A) = \iint_A f(x,y)dxdy} \qquad (1)$$

для любого (измеримого) множества A в $\mathbb{R}^2$.

Упражнение
Докажите, что (1) справедливо для $A = [a,b] \times [c,d]$.

Непрерывное совместное распределение: свойства плотности

Как и в одномерном случае, плотность обладает двумя свойствами:

1. $f(x, y) \geqslant 0$
2. $\iint_{\mathbb{R}^2} f(x, y) dx dy = 1$

и любая (измеримая) функция с этими свойствами может быть реализована как плотность распределения некоторого случайного вектора.

Распределение однозначно восстанавливается по плотности равенством (1).

Непрерывное совместное распределение: маржинальные распределения

Как найти индивидуальные (маржинальные) распределения случайных величин X, Y, если известна плотность $f(x, y)$ случайного вектора $U = [X, Y]'$?
Имеем

$$F_X(x) = P(X \leqslant x) = P(X \leqslant x, -\infty < Y < +\infty) = \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} f(u, v) dv \right) du.$$

Отсюда следует, что $f_X(x) = \int_{-\infty}^{+\infty} f(x, v) dv$.

Итак, зная совместное распределение, можно найти индивидуальные распределения.

Обратное, как и в дискретном случае, неверно, т. е. можно найти два двумерных вектора с одинаковыми распределениями компонентов, но с разными совместными распределениями.

Непрерывное совместное распределение: математическое ожидание

Зная совместное распределение величин X, Y и функцию $g(x, y)$, можно сосчитать математическое ожидание $g(X, Y)$:

$$E(g(X, Y)) = \iint_{\mathbb{R}^2} g(x, y)f(x, y)dxdy.$$

При этом если функция зависит только от x, то получится

$$E(g(X)) = \iint_{\mathbb{R}^2} g(x)f(x, y)dxdy = \int_{-\infty}^{+\infty} g(x)\left(\int_{-\infty}^{+\infty} f(x, y)dy\right)dx = \int_{-\infty}^{+\infty} g(x)f_X(x)dx$$

что совпадает с предыдущим определением. Для функции $g(y)$ аналогично.

Непрерывное совместное распределение: общий случай

В общем случае функция распределения вектора $X = [X_1, \dots, X_k]'$ определяется так:

$$F_X(x) = P(X_1 \leqslant x_1, \dots, X_k \leqslant x_k), \quad x = [x_1, \dots, x_k]' \in \mathbb{R}^k,$$

а плотность распределения (если, конечно, она существует) так:

$$f_X(x) = \frac{\partial^k F_X(x)}{\partial x_1 \dots \partial x_k}.$$

Независимые случайные величины

Пусть заданы случайные величины $X_1, \ldots, X_k$.

Определение
Случайные величины $X_1, \ldots, X_k$ называются независимыми, если

$$P(X_1 \in A_1, \ldots, X_k \in A_k) = \underline{P(X_1 \in A_1) \cdot \ldots \cdot P(X_k \in A_k)}$$

для любых (измеримых) подмножеств $A_j \subset \mathbb{R}$, $j = 1, \ldots, k$.

Ранее, определяя независимость системы событий, мы требовали выполнения соответствующих равенств «независимости» для любой подсистемы. В данном случае достаточно именно приведенного определения.

$, X_2 \in \mathbb{R}, \ldots X_{k-1} \in \mathbb{R}$ $\quad = 1$

Упражнение
Покажите, что верно равенство $P(X_1 \in A_1, X_k \in A_k) = P(X_1 \in A_1) \cdot P(X_k \in A_k)$. $\cdot P(X_2 \in \mathbb{R}) \cdot \ldots \cdot P(X_{k-1} \in \mathbb{R})$

Независимые случайные величины: совместное распределение

Из определения следует, что для вектора с независимыми компонентами выполнено

$$P(X_1 \leq x_1, \ldots X_k \leq x_k)$$

$$\boxed{F(x) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_k}(x_k),} \quad f(x) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_k}(x_k),$$

если каждое маржинальное одномерное распределение имеет плотность. Таким образом, если компоненты вектора X независимы, то по одномерным маржинальным распределениям можно восстановить распределение вектора X.

Аналогично для дискретного двумерного вектора независимость означает выполнение равенств

$$p_{ij} = p_{i\bullet} \cdot p_{\bullet j}. = P(X = x_i) \cdot P(Y = y_j)$$

$$P(X = x_i, Y = y_j)$$

Независимые случайные величины: матожидание $g(X)h(Y)$

Для независимых X, Y и функции вида $\boxed{g(x)h(y)}$ математическое ожидание $E(g(X)h(Y))$ распадается в произведение отдельных средних

$$E(g(X)h(Y)) = \sum_i \sum_j g(x_i)h(y_j)p_{i\bullet}p_{\bullet j} = \left(\sum_i g(x_i)p_{i\bullet}\right)\left(\sum_j h(y_j)p_{\bullet j}\right) = E(g(X))E(h(Y)).$$

Аналогично для непрерывного двумерного вектора:

$$E(g(X)h(Y)) = \iint_{\mathbb{R}^2} g(x)h(y)f_X(x)f_Y(y)dxdy = \left(\int_{-\infty}^{+\infty} g(x)f_X(x)dx\right)\left(\int_{-\infty}^{+\infty} h(y)f_Y(y)dy\right)$$
$$= E(g(X))E(h(Y)).$$

$X \cdot Y$

Независимые случайные векторы

Определение
Случайные векторы (вообще говоря, различных размерностей) $X^{(1)}$, ..., $X^{(m)}$ называются независимыми, если

$$P(X^{(1)} \in \underset{\subset \mathbb{R}^{k_1}}{A_1}, \ldots, X^{(m)} \in \underset{\subset \mathbb{R}^{k_m}}{A_m}) = P(X^{(1)} \in A_1) \cdot \ldots \cdot P(X^{(m)} \in A_m)$$

для любых борелевских множеств A_1, ..., A_m в соответствующих пространствах.

Распределение суммы двух независимых случайных величин

Пусть X, Y — непрерывные независимые случайные величины, $f_X(x)$, $f_Y(y)$ — их плотности. Обозначим $Z = X + Y$, $U = [X, Y]'$, $f_U(x, y) = f_X(x) f_Y(y)$. Тогда

$$F_Z(z) = P(X + Y \leqslant z) = P(U \in G_z) = \iint_{G_z} f_X(x) f_Y(y)\, dx\, dy,$$

" $f(x, y)$

где $G_z = \{(x, y) :\ x + y \leqslant z\}$. Имеем:

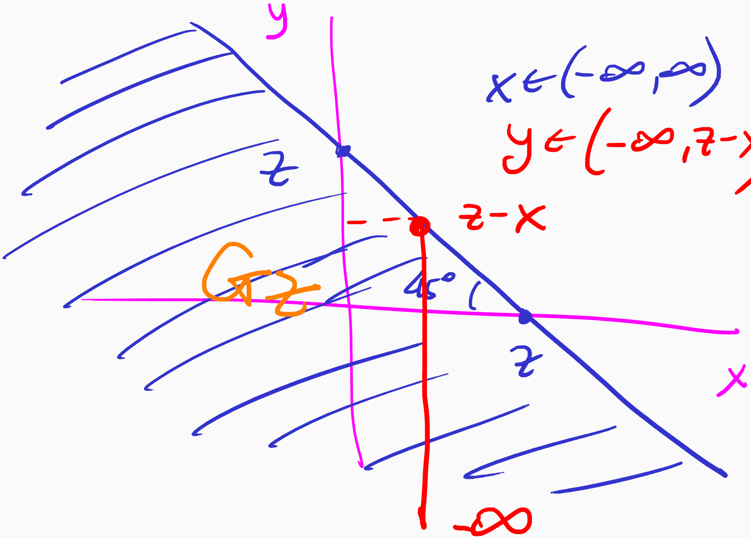
$$\iint_{G_z} f_X(x) f_Y(y)\, dx\, dy = \int_{-\infty}^{+\infty} f_X(x) \left(\int_{-\infty}^{z-x} f_Y(y)\, dy \right) dx = \int_{-\infty}^{+\infty} f_X(x) F_Y(z - x)\, dx.$$

Предполагая возможность дифференцирования под знаком интеграла, получаем:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z - x)\, dx = \int_{-\infty}^{+\infty} f_X(z - y) f_Y(y)\, dy$$

— свертка плотностей $f_X(x)$, $f_Y(y)$.

Для заметок



$x \in (-\infty, \infty)$

$y \in (-\infty, z-x)$

$Z = X + Y$

$\{Z \le z\} =$

$= \{X + Y \le z\}$

$= \{X + Y \in G_z\}$

$\{x + y \le z\}$

Распределение суммы двух независимых случайных величин: пример

Если X, Y — независимые показательные случайные величины с параметром $\lambda > 0$, т. е.

$$f_X(x) = f_Y(y) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

то при $z \geqslant 0$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx = \lambda^2 \int_0^z e^{-\lambda x} e^{-\lambda(z-x)}\,dx = \lambda^2 z e^{-\lambda z}.$$

Также можно проверить, что сумма двух независимых равномерно распределённых на $[0, 1]$ случайных величин имеет «треугольную» плотность.

Ковариация

Числовой характеристикой, связанной с независимостью двух случайных величин, является ковариация.

Определение
Число $\mathrm{Cov}(X, Y) = \sigma_{XY} = E[(X - E(X))(Y - E(Y))] = E[(X - m_X)(Y - m_Y)]$ называется ковариацией случайных величин X, Y.

Легко устанавливается равенство $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$.

Ковариация независимых случайных величин

Неформально, содержательно знак ковариации говорит о положительной или отрицательной связи между величинами, т. е. положительная ковариация показывает, что «в среднем» величины изменяются «в одну сторону», а отрицательная — что в разные.

Предложение
Если случайные величины X и Y независимы, то $E(XY) = E(X)E(Y)$ и $\operatorname{Cov}(X, Y) = 0$.

Доказательство.
Для случаев дискретного и непрерывного совместного распределения X и Y равенство $E(XY) = E(X)E(Y)$ в случае независимости X и Y доказано выше. В общем случае оно тоже справедливо. $\square$

Ковариация: свойства

1. Симметричность: $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
2. Билинейность: $\text{Cov}(aX + bY, Z) = a\,\text{Cov}(X, Z) + b\,\text{Cov}(Y, Z)$.
3. Если случайные величины X, Y независимы, то $\text{Cov}(X, Y) = 0$.

Таким образом, из независимости следует некоррелированность. Обратное, вообще говоря, неверно.

Пример

Пусть $X \sim N(0, 1)$, $Y = X^2$. Тогда X, Y зависимы (даже функционально связаны), но

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(X \cdot X^2) - E(X)E(X^2) = E(X^3) - E(X)E(X^2) = 0 - 0 \cdot 1 = 0.$$

Ковариация суммы

Из определения ковариации непосредственно следует равенство $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y)$ и его обобщение

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j).$$

Действительно,

$$\begin{aligned}\operatorname{Var}(X+Y) = \operatorname{Cov}(X+Y, X+Y) = \operatorname{Cov}(X,X) + \operatorname{Cov}(X,Y) + \operatorname{Cov}(Y,X) + \operatorname{Cov}(Y,Y)\\ = \operatorname{Var}(X) + \operatorname{Var}(X) + 2\operatorname{Cov}(X,Y).\end{aligned}$$

В частности, если величины X_i независимы, то

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i).$$

Ковариация: индикатор зависимости

На первый взгляд, ковариацию (точнее, ее выборочный аналог) можно рассматривать как приближенный индикатор независимости: большая ковариация — есть зависимость, маленькая — нет зависимости. Однако для ковариации в силу ее размерности понятие «большая–маленькая» теряет смысл. Поэтому рассматривается модификация ковариации — *коэффициент корреляции*.

Ковариация: неравенство Коши – Буняковского

Предложение (неравенство Коши – Буняковского)
Для любых случайных величин ξ, η, имеющих вторые моменты, выполнено неравенство $(E(\xi\eta))^2 \leqslant E(\xi^2)E(\eta^2)$.

Доказательство.
При любом x квадратичная функция $f(x) = E((x\xi + \eta)^2) = x^2E(\xi^2) + 2xE(\xi\eta) + E(\eta^2)$ неотрицательна, поэтому ее дискриминант $D = 4((E(\xi\eta))^2 - E(\xi^2)E(\eta^2)) \leqslant 0$. □

Следствие
Для любых случайных величин X, Y выполнено неравенство $(\mathrm{Cov}(X, Y))^2 \leqslant \mathrm{Var}(X)\,\mathrm{Var}(Y)$.

Доказательство.
Обозначим $\xi = X - E(X)$, $\eta = Y - E(Y)$ и применим неравенство Коши – Буняковского. □

Коэффициент корреляции

Определение
Число

$$\rho(X, Y) = \rho_{XY} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

называется коэффициентом корреляции случайных величин X, Y.

Коэффициент корреляции: свойства

1. $-1 \leqslant \rho(X, Y) \leqslant 1$ для любых X, Y (следует из неравенства Коши – Буняковского)
2. $\rho(X, Y) = 0$ тогда и только тогда, когда X и Y некоррелированы. В частности, если X, Y независимы, то $\rho(X, Y) = 0$. Обратное неверно
3. Если $\rho(X, Y) = 1$, то $Y = aX + b$, $a > 0$; если $\rho(X, Y) = -1$, то $Y = aX + b$, $a < 0$

Коэффициент корреляции: комментарий

Коэффициент корреляции представляет собой меру линейной связи между двумя случайными величинами. Если $\rho(X, Y) = \pm 1$, то величины X и Y связаны друг с другом точно. При $0 < |\rho(X, Y)| < 1$ точной линейной связи нет. Если $\rho(X, Y) = 0$, то можно считать, что линейной связи между величинами нет совсем (при этом наилучший линейный прогноз Y, зная X, не зависит от X).

Ненулевой коэффициент корреляции между величинами X и Y говорит только о наличии статистической связи между ними, и ничего не говорит о наличии причинно-следственной связи.

Литература

[1] Sh. Ross. *A First Course in Probability*. 9th ed. London: Pearson, 2014.

[2] Б. В. Гнеденко. *Курс теории вероятностей*. М.: Издательство УРСС, 2005.

[3] С. В. Головань, П. К. Катышев, А. А. Пересецкий. *Сборник задач по курсу теории вероятностей*. М.: Издательский дом «Дело» РАНХиГС, 2021.

Для заметок