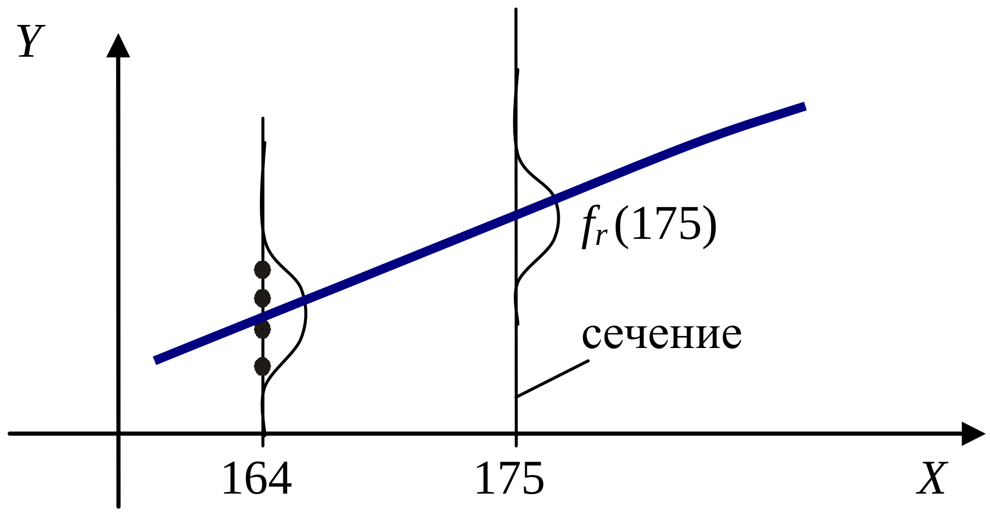
Условное математическое ожидание (продолжение)

Условное математическое ожидание (продолжение)

Определение. Величина $E\left(Y\middle|x\right)$ как функция от x называется **регрессией** Y по x (или x на Y).

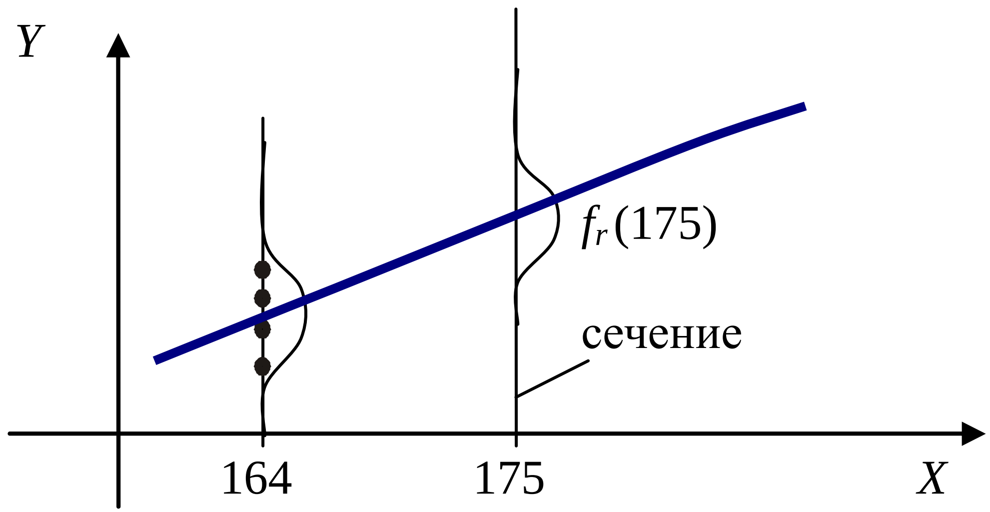
$$\boxed{y = f_r(x)}$$

Термин "регрессия" – предложен Ф. Гальтоном (конец19 века).



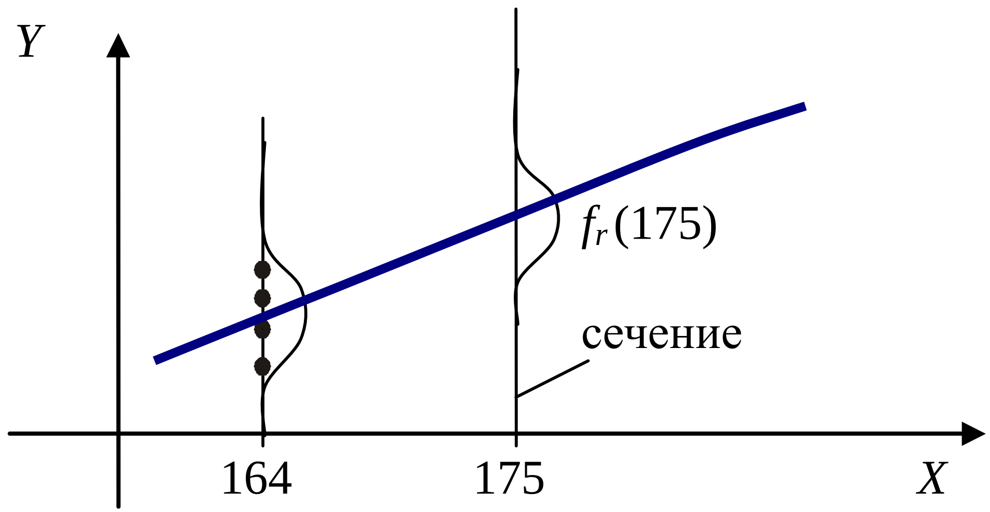
X рост отца, Y - рост сына.

Термин "регрессия" – предложен Ф. Гальтоном (конец19 века).



X рост отца, Y - рост сына.
Вывод: дети родителей с высоким или низким ростом обычно не наследуют такой рост ("регрессия к среднему значению").

Термин "регрессия" – предложен Ф. Гальтоном (конец19 века).



X рост отца, Y - рост сына.
Вывод: дети родителей с высоким или низким ростом обычно не наследуют такой рост ("регрессия к среднему значению").
К.Пирсон предложил использовать термин в общем смысле.

Свойство оптимальности функции регрессии: минимизирует условное математическое ожидание квадрата отклонения от Y при $X = x$:

$$E[(Y - f_r(x))^2] \to \min,$$

Свойство оптимальности функции регрессии: минимизирует условное математическое ожидание квадрата отклонения от Y при $X = x$:

$$E[(Y - f_r(x))^2] \to \min,$$

где минимум берется по всевозможным функциям.

Свойство оптимальности функции регрессии: минимизирует условное математическое ожидание квадрата отклонения от Y при $X = x$:

$$E[(Y - f_r(x))^2] \to \min,$$

где минимум берется по всевозможным функциям.

$$\frac{\partial}{\partial f} E[(Y - f)^2] = -2E[Y - f] = 0 \;\Rightarrow\; f = E[Y].$$

Свойство оптимальности функции регрессии: минимизирует условное математическое ожидание квадрата отклонения от Y при $X = x$:

$$E[(Y - f_r(x))^2] \to \min,$$

где минимум берется по всевозможным функциям.

$$\frac{\partial}{\partial f} E[(Y - f\,)^2] = -2E[Y - f\,] = 0 \;\Rightarrow\; f = E[Y].$$

Функция регрессии может использоваться для прогноза Y в зависимости от x:

$$\boxed{\hat{y} = f_r(x)},$$

т.к. такой прогноз обладает минимальной погрешностью.

Свойство оптимальности функции регрессии: минимизирует условное математическое ожидание квадрата отклонения от Y при $X = x$:

$$E[(Y - f_r(x))^2] \to \min,$$

где минимум берется по всевозможным функциям.

$$\frac{\partial}{\partial f} E[(Y - f)^2] = -2E[Y - f] = 0 \;\Rightarrow\; f = E[Y].$$

Функция регрессии может использоваться для прогноза Y в зависимости от x:

$$\boxed{\hat{y} = f_r(x)},$$

т.к. такой прогноз обладает минимальной погрешностью.

Регрессионный анализ - построение и исследование регрессионных моделей по выборке наблюдений.

Пример. Дискретная двумерная случайная величина задана таблицей распределения: Найти условное математическое ожидание $E(Y \mid x_1 = 1)$.

	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,3	0,1	0,03	0,07

Пример. Дискретная двумерная случайная величина задана таблицей распределения: Найти условное математическое ожидание $E(Y \mid x_1 = 1)$.

	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,3	0,1	0,03	0,07

$$P(x_1) = \sum_{j=1}^{2} P(x_1, y_j) = 0.45;$$

Пример. Дискретная двумерная случайная величина задана таблицей распределения: Найти условное математическое ожидание $E(Y \mid x_1 = 1)$.

	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,3	0,1	0,03	0,07

$$P(x_1) = \sum_{j=1}^{2} P(x_1, y_j) = 0.45;$$

$$P(y_1 \mid x_1) = \frac{P(x_1, y_1)}{P(x_1)} = \frac{1}{3}; \quad P(y_2 \mid x_1) = \frac{P(x_1, y_2)}{P(x_1)} = \frac{2}{3};$$

Пример. Дискретная двумерная случайная величина задана таблицей распределения: Найти условное математическое ожидание $E(Y \mid x_1 = 1)$.

	$x_1 = 1$	$x_2 = 3$	$x_3 = 4$	$x_4 = 8$
$y_1 = 3$	0,15	0,06	0,25	0,04
$y_2 = 6$	0,3	0,1	0,03	0,07

$$P(x_1) = \sum_{j=1}^{2} P(x_1, y_j) = 0.45;$$

$$P(y_1 \mid x_1) = \frac{P(x_1, y_1)}{P(x_1)} = \frac{1}{3}; \quad P(y_2 \mid x_1) = \frac{P(x_1, y_2)}{P(x_1)} = \frac{2}{3};$$

$$E(Y \mid x_1 = 1) = \sum_{j=1} y_j P(Y_j \mid x_1) = 3 \cdot \frac{1}{3} + 6 \cdot \frac{2}{3} = 5.$$

Глава 7. Корреляция случайных величин

Глава 7. Корреляция случайных величин

Определение. Коэффициентом линейной корреляции между случайными величинами X, Y, дисперсии которых существуют и не равны нулю, называется величина

$$\boxed{\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}},$$

Глава 7. Корреляция случайных величин

Определение. Коэффициентом линейной корреляции между случайными величинами X, Y, дисперсии которых существуют и не равны нулю, называется величина

$$\boxed{\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}},$$

где

$$\boxed{\sigma_{XY} = \operatorname{cov}(X, Y) \overset{\text{def}}{=} E(X - EX)(Y - EY)}$$

- коэффициент ковариации между X и Y,

Глава 7. Корреляция случайных величин

Определение. Коэффициентом линейной корреляции между случайными величинами X, Y, дисперсии которых существуют и не равны нулю, называется величина

$$\boxed{\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}},$$

где

$$\boxed{\sigma_{XY} = \operatorname{cov}(X, Y) \overset{\text{def}}{=} E(X - EX)(Y - EY)}$$

- коэффициент ковариации между X и Y, σ_X и σ_Y - стандартное уклонение величин X и Y ($\sigma_X^2 = DX$, $\sigma_Y^2 = DY$).

Другая форма определения:

$$\sigma_{XY} = E(XY) - E(EX \cdot Y) - E(EY \cdot X) + EX \cdot EY =$$

$$= \boxed{E(XY) - EX \cdot EY}.$$

Другая форма определения:

$$\sigma_{XY} = E(XY) - E(EX \cdot Y) - E(EY \cdot X) + EX \cdot EY =$$

$$= \boxed{E(XY) - EX \cdot EY}.$$

Определение. Если $\sigma_{XY} = 0$, то X и Y называют некоррелированными.

Свойства коэффициента корреляции.

1. $-1 \leq \rho_{XY} \leq +1.$

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}$, $Y^* = \dfrac{Y - EY}{\sigma_Y}$.

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}$, $Y^* = \dfrac{Y - EY}{\sigma_Y}$.

Тогда $EX^* = 0$, $EY^* = 0$, $\sigma^2_{X^*} = 1$, $\sigma^2_{Y^*} = 1$,

Свойства коэффициента корреляции.

1. $-1 \leq \rho_{XY} \leq +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}$, $Y^* = \dfrac{Y - EY}{\sigma_Y}$.

Тогда $EX^* = 0,\ EY^* = 0,\ \sigma^2_{X^*} = 1,\ \sigma^2_{Y^*} = 1,$

$$\sigma_{X^*Y^*} = E(X^*Y^*) - 0 \cdot 0 = E\left(\frac{X - EX}{\sigma_X} \cdot \frac{Y - EY}{\sigma_Y}\right) = \rho_{XY}.$$

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}, \; Y^* = \dfrac{Y - EY}{\sigma_Y}.$

Тогда $EX^* = 0, \; EY^* = 0, \; \sigma^2_{X^*} = 1, \; \sigma^2_{Y^*} = 1,$

$$\sigma_{X^*Y^*} = E(X^*Y^*) - 0 \cdot 0 = E\left(\frac{X - EX}{\sigma_X} \cdot \frac{Y - EY}{\sigma_Y}\right) = \rho_{XY}.$$

$$D\left(X^* + Y^*\right) = E\left[X^* + Y^* - E\left(X^* + Y^*\right)\right]^2 =$$

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}, \; Y^* = \dfrac{Y - EY}{\sigma_Y}.$

Тогда $EX^* = 0, \; EY^* = 0, \; \sigma^2_{X^*} = 1, \; \sigma^2_{Y^*} = 1,$

$$\sigma_{X^*Y^*} = E(X^*Y^*) - 0 \cdot 0 = E\left(\frac{X - EX}{\sigma_X} \cdot \frac{Y - EY}{\sigma_Y} \right) = \rho_{XY}.$$

$$D\left(X^* + Y^*\right) = E\left[X^* + Y^* - E\left(X^* + Y^*\right) \right]^2 =$$

$$= E\left[(X^* - EX^*) + (Y^* - EY^* \right]^2 =$$

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}$, $Y^* = \dfrac{Y - EY}{\sigma_Y}$.

Тогда $EX^* = 0,\ EY^* = 0,\ \sigma^2_{X^*} = 1,\ \sigma^2_{Y^*} = 1,$

$$\sigma_{X^*Y^*} = E(X^*Y^*) - 0 \cdot 0 = E\left(\frac{X - EX}{\sigma_X} \cdot \frac{Y - EY}{\sigma_Y}\right) = \rho_{XY}.$$

$$D\left(X^* + Y^*\right) = E\left[X^* + Y^* - E\left(X^* + Y^*\right)\right]^2 =$$

$$= E\left[(X^* - EX^*) + (Y^* - EY^*\right]^2 =$$

$$= E\left[\left(X^* - EX^*\right)^2 + 2\left(X^* - EX^*\right)\left(Y^* - EY^*\right) + \left(Y^* - EY^*\right)^2\right] =$$

Свойства коэффициента корреляции.

1. $-1 \le \rho_{XY} \le +1.$

Доказательство. Обозначим $X^* = \dfrac{X - EX}{\sigma_X}$, $Y^* = \dfrac{Y - EY}{\sigma_Y}$.

Тогда $EX^* = 0,\ EY^* = 0,\ \sigma^2_{X^*} = 1,\ \sigma^2_{Y^*} = 1,$

$$\sigma_{X^*Y^*} = E(X^*Y^*) - 0 \cdot 0 = E\left(\frac{X - EX}{\sigma_X} \cdot \frac{Y - EY}{\sigma_Y}\right) = \rho_{XY}.$$

$$D\left(X^* + Y^*\right) = E\left[X^* + Y^* - E\left(X^* + Y^*\right)\right]^2 =$$

$$= E\left[(X^* - EX^*) + (Y^* - EY^*\right]^2 =$$

$$= E\left[\left(X^* - EX^*\right)^2 + 2\left(X^* - EX^*\right)\left(Y^* - EY^*\right) + \left(Y^* - EY^*\right)^2\right] =$$

$$= \underbrace{\sigma^2_{X^*}}_{1} + 2\underbrace{\sigma_{X^*Y^*}}_{\rho_{XY}} + \underbrace{\sigma^2_{Y^*}}_{1} = 2(1 + \rho_{XY}).$$

Аналогично

$$D\left(X^{*}-Y^{*}\right)=2(1-\rho_{XY}).$$

Аналогично

$$D\left(X^* - Y^*\right) = 2(1 - \rho_{XY}).$$

Так как

$$D\left(X^* \pm Y^*\right) \geq 0, \text{ то } 2\left(1 \pm \rho_{XY}\right) \geq 0 \;\Rightarrow\; -1 \leq \rho_{XY} \leq 1.$$

Аналогично

$$D\left(X^*-Y^*\right)=2(1-\rho_{XY}).$$

Так как

$$D\left(X^*\pm Y^*\right)\geq 0,\text{ то }2\left(1\pm\rho_{XY}\right)\geq 0\;\Rightarrow\;-1\leq\rho_{XY}\leq 1.$$

Замечание.

а) Из доказательства следует, что для произвольных зависимых случайных величин X и Y выполняется:

$$\boxed{D(X\pm Y)=DX+DY\pm 2\cdot\sigma_{XY}}.$$

Аналогично

$$D\left(X^* - Y^*\right) = 2(1 - \rho_{XY}).$$

Так как

$$D\left(X^* \pm Y^*\right) \ge 0, \text{ то } 2\left(1 \pm \rho_{XY}\right) \ge 0 \;\Rightarrow\; -1 \le \rho_{XY} \le 1.$$

Замечание.

а) Из доказательства следует, что для произвольных зависимых случайных величин X и Y выполняется:

$$\boxed{D(X \pm Y) = DX + DY \pm 2 \cdot \sigma_{XY}}.$$

б) Из определения коэффициента ковариации $\Rightarrow$ для зависимых X и Y, $\boxed{E(X \cdot Y) = EX \cdot EY + \sigma_{XY}}$.

Аналогично

$$D\left(X^* - Y^*\right) = 2(1 - \rho_{XY}).$$

Так как

$$D\left(X^* \pm Y^*\right) \geq 0, \text{ то } 2\left(1 \pm \rho_{XY}\right) \geq 0 \;\Rightarrow\; -1 \leq \rho_{XY} \leq 1.$$

Замечание.

а) Из доказательства следует, что для произвольных зависимых случайных величин X и Y выполняется:

$$\boxed{D(X \pm Y) = DX + DY \pm 2 \cdot \sigma_{XY}}.$$

б) Из определения коэффициента ковариации $\Rightarrow$ для зависимых X и Y, $\boxed{E(X \cdot Y) = EX \cdot EY + \sigma_{XY}}$.

в) Из $|\rho_{XY}| \leq 1$ следует, что $\boxed{\left|\sigma_{XY}\right| \leq \sigma_X \cdot \sigma_Y}$.

2. Пусть $Y = a + bX$ (строгая линейная зависимость).
Тогда $|\rho_{XY}| = 1$.

2. Пусть $Y = a + bX$ (строгая линейная зависимость).
Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость). Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость). Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

$$= E[\left(X - EX\right)\cdot(a + bX - a - bEX)] =$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость). Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

$$= E[\left(X - EX\right)\cdot(a + bX - a - bEX)] =$$

$$E[(X - EX)\cdot b(X - EX)] = b\sigma_X^2 .$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость).
Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

$$= E[\left(X - EX\right)\cdot(a + bX - a - bEX)] =$$

$$E[(X - EX)\cdot b(X - EX)] = b\sigma_X^2.$$

$$\sigma_Y^2 = D[a + bX] = b^2 D[X] = b^2\sigma_X^2 \Rightarrow$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость).
Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

$$= E[\left(X - EX\right)\cdot(a + bX - a - bEX)] =$$

$$E[(X - EX)\cdot b(X - EX)] = b\sigma_X^2.$$

$$\sigma_Y^2 = D[a + bX] = b^2 D[X] = b^2\sigma_X^2 \Rightarrow$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{b\sigma_X^2}{\sigma_X\sqrt{b^2\cdot\sigma_X^2}} = \frac{b\sigma_X^2}{\sigma_X\,|b|\cdot\sigma_X} = \frac{b}{|b|} \Rightarrow$$

2. Пусть $Y = a + bX$ (строгая линейная зависимость). Тогда $|\rho_{XY}| = 1$.

Доказательство.

$$\sigma_{XY} = E[(X - EX)\cdot(Y - EY)] =$$

$$= E[(X - EX)\cdot(a + bX - E[a + bX])] =$$

$$= E[\left(X - EX\right)\cdot(a + bX - a - bEX)] =$$

$$E[(X - EX)\cdot b(X - EX)] = b\sigma_X^2.$$

$$\sigma_Y^2 = D[a + bX] = b^2 D[X] = b^2\sigma_X^2 \Rightarrow$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{b\sigma_X^2}{\sigma_X\sqrt{b^2\cdot\sigma_X^2}} = \frac{b\sigma_X^2}{\sigma_X\,|b|\cdot\sigma_X} = \frac{b}{|b|} \Rightarrow$$

$$\rho_{XY} = \begin{cases} 1, & \text{если } b > 0 \\ -1, & \text{если } b < 0 \end{cases}.$$

Таким образом, величина коэффициента корреляции меняется от –1 в случае строгой линейной отрицательной связи до +1 в случае строгой линейной положительной связи.

Таким образом, величина коэффициента корреляции меняется от –1 в случае строгой линейной отрицательной связи до +1 в случае строгой линейной положительной связи.

3. Если X и Y независимы, то они некоррелированы.

Таким образом, величина коэффициента корреляции меняется от –1 в случае строгой линейной отрицательной связи до +1 в случае строгой линейной положительной связи.

3. Если X и Y независимы, то они некоррелированы.

Доказательство.

$$\sigma_{XY} = E[XY] - EX \cdot EY =$$
$$= EX \cdot EY - EX \cdot EY = 0.$$

Замечание. Некоррелированность – необходимое, но не достаточное условие независимости случайных величин.

Замечание. Некоррелированность – необходимое, но не достаточное условие независимости случайных величин.

Пусть, например, имеются зависимые случайные величины $X \sim U(-1,1)$ и $Y = X^2$.

Ковариация $\sigma_{XY} = E(XY) - EX \cdot EY = \underbrace{EX^3}_{0} - \underbrace{EX}_{0} \cdot EX^2 = 0$

Замечание. Некоррелированность – необходимое, но не достаточное условие независимости случайных величин.

Пусть, например, имеются зависимые случайные величины $X \sim U(-1,1)$ и $Y = X^2$.

Ковариация $\sigma_{XY} = E(XY) - EX \cdot EY = \underbrace{EX^3}_{0} - \underbrace{EX}_{0} \cdot EX^2 = 0$

Только для нормально распределенных случайных величин из некоррелированности следует их независимость.

4. Коэффициент корреляции является безразмерной величиной; не зависит от выбора единиц измерения обеих переменных, а также от их порядка ($\rho_{XY} = \rho_{YX}$).

4. Коэффициент корреляции является безразмерной величиной; не зависит от выбора единиц измерения обеих переменных, а также от их порядка ($\rho_{XY} = \rho_{YX}$).

5. Пусть $Y \equiv X$, тогда

$$\rho_{XX} = 1$$

(так как $X = 1 \cdot X + 0$ - линейная функция).

4. Коэффициент корреляции является безразмерной величиной; не зависит от выбора единиц измерения обеих переменных, а также от их порядка ($\rho_{XY} = \rho_{YX}$).

5. Пусть $Y \equiv X$, тогда

$$\rho_{XX} = 1$$

(так как $X = 1 \cdot X + 0$ - линейная функция).

Кроме того,

$$\sigma_{XX} = E(X - EX)(X - EX) = DX .$$

Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

$|\rho_{XY}| < 0{,}25$ — линейная связь отсутствует;

Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

$|\rho_{XY}| < 0{,}25$ — линейная связь отсутствует;

$0{,}25 \leq | \rho_{XY} | < 0{,}5$ — имеется слабая линейная связь;

Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

$|\rho_{XY}| < 0,25$ — линейная связь отсутствует;

$0,25 \le |\rho_{XY}| < 0,5$ — имеется слабая линейная связь;

$0,5 \le |\rho_{XY}| < 0,75$ — имеется умеренная линейная связь;

Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

$|\rho_{XY}| < 0,25$ — линейная связь отсутствует;

$0,25 \leq |\rho_{XY}| < 0,5$ — имеется слабая линейная связь;

$0,5 \leq |\rho_{XY}| < 0,75$ — имеется умеренная линейная связь;

$|\rho_{XY}| \geq 0,75$ — имеется сильная линейная связь.

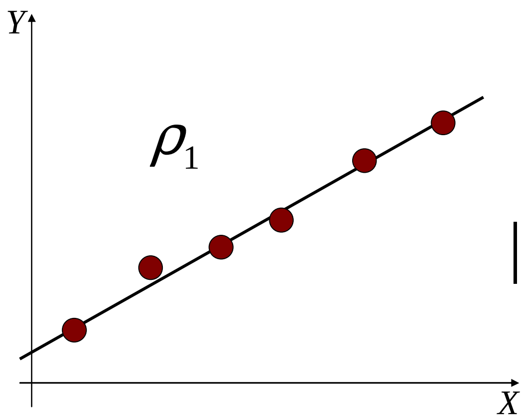
Для оценки степени линейной корреляционной зависимости можно воспользоваться следующей приближенной оценкой:

$|\rho_{XY}| < 0,25$ — линейная связь отсутствует;

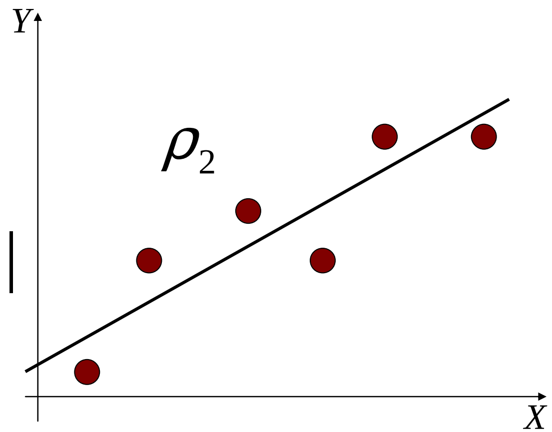
$0,25 \le |\rho_{XY}| < 0,5$ — имеется слабая линейная связь;

$0,5 \le |\rho_{XY}| < 0,75$ — имеется умеренная линейная связь;

$|\rho_{XY}| \ge 0,75$ — имеется сильная линейная связь.

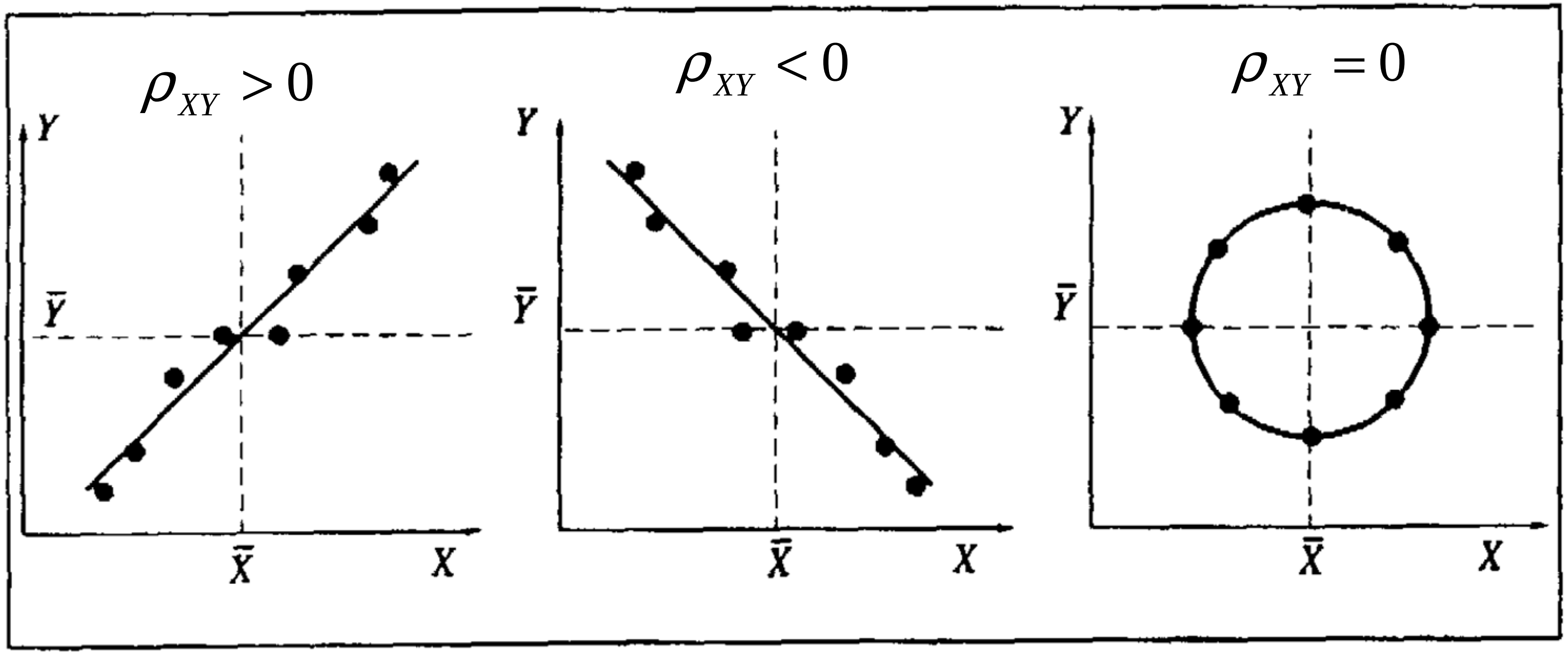


$|\rho_1| > |\rho_2|$

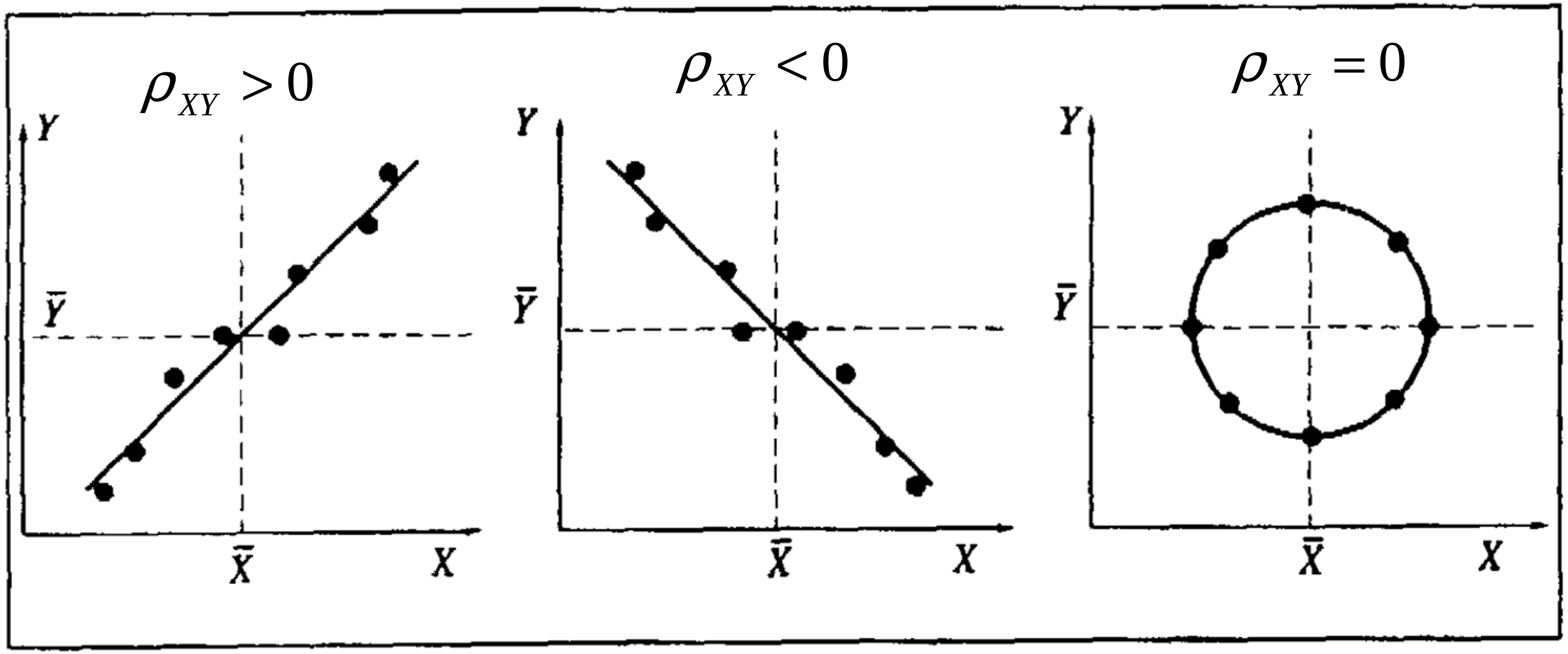


$|\rho|$-мера концентрации распределения около прямой линии

Типы зависимостей и коэффициент корреляции



Типы зависимостей и коэффициент корреляции

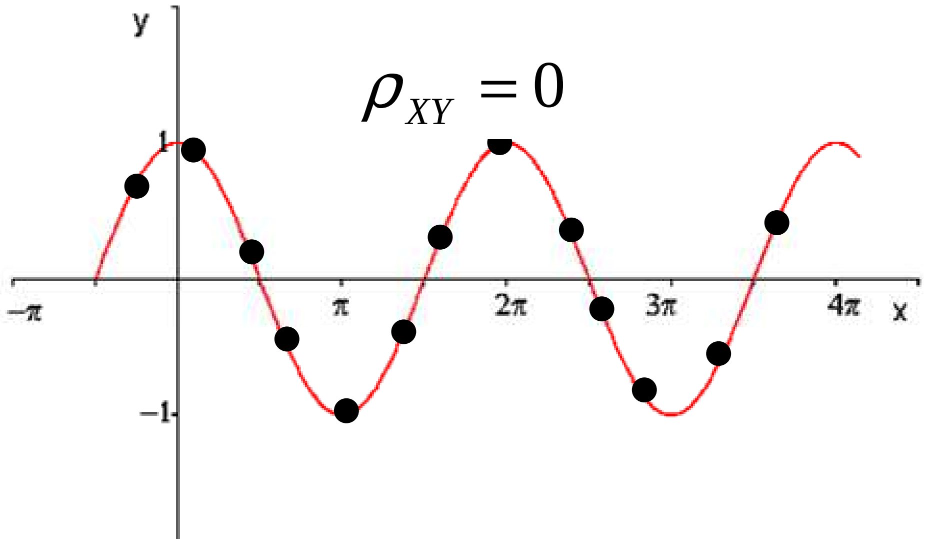


Замечание 1. Высокое значение коэффициента линейной корреляции не обязательно означает наличие причинно-следственной связи между переменными.

Замечание 2. Коэффициент корреляции ρ_{XY} показывает тесноту именно линейной связи между величинами.

Замечание 2. Коэффициент корреляции ρ_{XY} показывает тесноту именно линейной связи между величинами.

В случае строгой нелинейной зависимости между X и Y (например, $Y = \cos X$) коэффициент корреляции может быть равен нулю.



Корреляционная и ковариационная матрицы

Пусть имеется несколько случайных величин $X_1, X_2, \ldots, X_n$. Тогда можно найти коэффициенты корреляции

$$\rho_{i,j} \overset{\text{def}}{=} \rho_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1, \ldots, n)$.

Корреляционная и ковариационная матрицы

Пусть имеется несколько случайных величин $X_1, X_2, \ldots, X_n$. Тогда можно найти коэффициенты корреляции

$$\rho_{i,j} \overset{\text{def}}{=} \rho_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1, \ldots, n)$.

Запишем эти коэффициенты с помощью корреляционной матрицы

$$\mathrm{P} = (\rho_{i,j}).$$

Корреляционная и ковариационная матрицы

Пусть имеется несколько случайных величин $X_1, X_2, \ldots, X_n$. Тогда можно найти коэффициенты корреляции

$$\rho_{i,j} \overset{\text{def}}{=} \rho_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1,\ldots,n)$.

Запишем эти коэффициенты с помощью корреляционной матрицы

$$\mathrm{P} = (\rho_{i,j}).$$

- Матрица является симметричной относительно главной диагонали $(\rho_{i,j} = \rho_{j,i})$;

Корреляционная и ковариационная матрицы

Пусть имеется несколько случайных величин $X_1, X_2, \ldots, X_n$. Тогда можно найти коэффициенты корреляции

$$\rho_{i,j} \overset{\text{def}}{=} \rho_{X_i X_j}$$

для каждой пары X_i, X_j $(i, j = 1, \ldots, n)$.

Запишем эти коэффициенты с помощью корреляционной матрицы

$$\mathrm{P} = (\rho_{i,j}).$$

- Матрица является симметричной относительно главной диагонали $(\rho_{i,j} = \rho_{j,i})$;
- все диагональные элементы равны единице $(\rho_{i,i} = 1)$.

Корреляционная и ковариационная матрицы

Пусть имеется несколько случайных величин $X_1, X_2, \ldots, X_n$. Тогда можно найти коэффициенты корреляции

$$\rho_{i,j} \overset{\text{def}}{=} \rho_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1, \ldots, n)$.

Запишем эти коэффициенты с помощью корреляционной матрицы

$$\mathrm{P} = (\rho_{i,j}).$$

- Матрица является симметричной относительно главной диагонали ($\rho_{i,j} = \rho_{j,i}$);
- все диагональные элементы равны единице ($\rho_{i,i} = 1$).

Пример (n=3):

$$\mathrm{P} = \begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} \\ \rho_{2,1} & 1 & \rho_{2,3} \\ \rho_{3,1} & \rho_{3,2} & 1 \end{pmatrix}$$

Аналогично положим

$$\sigma_{i,j} \overset{\text{def}}{=} \sigma_{X_i X_j}$$

для каждой пары $X_i,\ X_j$ $(i, j = 1,\ldots,n)$.

Аналогично положим

$$\sigma_{i,j} \overset{\text{def}}{=} \sigma_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1,\dots,n)$.

Из этих коэффициентов составим ковариационную матрицу

$$\Sigma = (\sigma_{i,j}).$$

Аналогично положим

$$\sigma_{i,j} \overset{\text{def}}{=} \sigma_{X_i X_j}$$

для каждой пары $X_i,\ X_j$ $(i, j = 1,\ldots,n)$.

Из этих коэффициентов составим ковариационную матрицу

$$\Sigma = (\sigma_{i,j}).$$

- Матрица является симметричной относительно главной диагонали ($\sigma_{i,j} = \sigma_{j,i}$);

Аналогично положим

$$\sigma_{i,j} \overset{\text{def}}{=} \sigma_{X_i X_j}$$

для каждой пары $X_i,\ X_j\ (i, j = 1,\dots,n)$.

Из этих коэффициентов составим ковариационную матрицу

$$\Sigma = (\sigma_{i,j}).$$

- Матрица является симметричной относительно главной диагонали $(\sigma_{i,j} = \sigma_{j,i})$;

- Диагональные элементы равны дисперсиям:

$$\sigma_{i,i} = DX_i,\ i = 1,\dots,n.$$

Глава 8. Предельные теоремы

Глава 8. Предельные теоремы

Пусть имеется последовательность случайных величин $X_1, X_2, ...$ Будем считать, что эти величины независимы и подчиняются одному и тому же распределению.

Например, $X_i \in \{0,1\}$ - исход эксперимента в схеме Бернулли.

Глава 8. Предельные теоремы

Пусть имеется последовательность случайных величин $X_1, X_2, \ldots$. Будем считать, что эти величины независимы и подчиняются одному и тому же распределению.

Например, $X_i \in \{0,1\}$ - исход эксперимента в схеме Бернулли.

Как ведут себя характеристики последовательности с ростом числа ее элементов?

Глава 8. Предельные теоремы

Пусть имеется последовательность случайных величин $X_1, X_2, \ldots$. Будем считать, что эти величины независимы и подчиняются одному и тому же распределению.

Например, $X_i \in \{0,1\}$ - исход эксперимента в схеме Бернулли.

Как ведут себя характеристики последовательности с ростом числа ее элементов?

Предельные теоремы

Законы больших чисел

Центральная предельная теорема (несколько форм)

Замечание. Независимым испытаниям эквивалентны две схемы:

- проведение n раз одного и того же испытания;
- проведение по одному независимому испытанию над n копиями одного и того же.

Аналогия: 100 раз монету подбрасывает 1 человек или 100 человек подбрасывают по одной одинаковой монете.

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

$$\boxed{X_n \xrightarrow{p} X}$$

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

$$\boxed{X_n \xrightarrow{p} X}$$

Теорема 1 (ЗБЧ в форме Чебышева). Пусть случайные величины $X_1, X_2, \ldots$ независимы и одинаково распределены, причем $EX_1^2 < \infty$.

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

$$\boxed{X_n \xrightarrow{p} X}$$

Теорема 1 (ЗБЧ в форме Чебышева). Пусть случайные величины $X_1, X_2, \ldots$ независимы и одинаково распределены, причем $EX_1^2 < \infty$. Обозначим $a = EX_1$, $S_n = \sum_{i=1}^{n} X_i$.

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

$$\boxed{X_n \xrightarrow{p} X}$$

Теорема 1 (ЗБЧ в форме Чебышева). Пусть случайные величины $X_1, X_2, \ldots$ независимы и одинаково распределены, причем $EX_1^2 < \infty$. Обозначим $a = EX_1$, $S_n = \sum_{i=1}^{n} X_i$. Тогда при $n \to \infty$

$$\frac{S_n}{n} \xrightarrow{p} a.$$

Определение. Последовательность $\{X_n\}$ сходится по вероятности к случайной величине X, если для любого $\varepsilon > 0$ вероятность события

$$P(|X_n - X| \geq \varepsilon)$$

стремится к нулю при $n \to \infty$.

$$\boxed{X_n \xrightarrow{p} X}$$

Теорема 1 (ЗБЧ в форме Чебышева). Пусть случайные величины $X_1, X_2, \ldots$ независимы и одинаково распределены, причем $EX_1^2 < \infty$. Обозначим $a = EX_1$, $S_n = \sum_{i=1}^{n} X_i$. Тогда при $n \to \infty$

$$\frac{S_n}{n} \xrightarrow{p} a.$$

При увеличении числа испытаний среднее сходится по вероятности к математическому ожиданию.

Доказательство.

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{na}{n} = a.$$

Доказательство.

$$E\left(\frac{S_n}{n}\right)=\frac{1}{n}\sum_{i=1}^{n}EX_i=\frac{na}{n}=a.$$

Пусть $\sigma^2=DX_1$. По неравенству Чебышева,

$$P\left(\left|\frac{S_n}{n}-a\right|\geq\varepsilon\right)\leq\frac{D\left(S_n/n\right)}{\varepsilon^2}=\frac{1}{n^2}\frac{D\left(S_n\right)}{\varepsilon^2}=$$

Доказательство.

$$E\left(\frac{S_n}{n}\right)=\frac{1}{n}\sum_{i=1}^{n}EX_i=\frac{na}{n}=a.$$

Пусть $\sigma^2=DX_1$. По неравенству Чебышева,

$$P\left(\left|\frac{S_n}{n}-a\right|\geq\varepsilon\right)\leq\frac{D\left(S_n/n\right)}{\varepsilon^2}=\frac{1}{n^2}\frac{D\left(S_n\right)}{\varepsilon^2}=$$

$$=\frac{1}{n^2}\frac{D\left(\sum_{i=1}^{n}X_i\right)}{\varepsilon^2}=\frac{1}{n^2}\frac{n\sigma^2}{\varepsilon^2}=\frac{\sigma^2}{n\varepsilon^2}\xrightarrow[n\to\infty]{}0.$$

Доказательство.

$$E\left(\frac{S_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{na}{n} = a.$$

Пусть $\sigma^2 = DX_1$. По неравенству Чебышева,

$$P\left(\left|\frac{S_n}{n} - a\right| \geq \varepsilon\right) \leq \frac{D\left(S_n/n\right)}{\varepsilon^2} = \frac{1}{n^2}\frac{D\left(S_n\right)}{\varepsilon^2} =$$

$$= \frac{1}{n^2}\frac{D\left(\sum_{i=1}^{n} X_i\right)}{\varepsilon^2} = \frac{1}{n^2}\frac{n\sigma^2}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0.$$

Значит, по определению сходимости,

$$\frac{S_n}{n} \xrightarrow{p} a \quad \text{при } n \to \infty.$$

Практический вывод из ЗБЧ: неизвестное значение математического ожидания случайной величины можно заменить ее средним значением, полученным по достаточно большому числу опытов.

Закон больших чисел в форме Бернулли

Теорема 2 (Бернулли). Пусть m - число успехов в n испытаниях схемы Бернулли, p - вероятность успеха в одном испытании. Тогда

$$\boxed{\frac{m}{n} \xrightarrow{p} p}$$

при $n \to \infty$.

Закон больших чисел в форме Бернулли

Теорема 2 (Бернулли). Пусть m - число успехов в n испытаниях схемы Бернулли, p - вероятность успеха в одном испытании. Тогда

$$\boxed{\frac{m}{n} \xrightarrow{p} p}$$

при $n \to \infty$.

Доказательство. Пусть X_i - число успехов в i-м испытании, тогда X_i подчиняется распределению Бернулли, все $X_1, ..., X_n$ независимы и $EX_1 = p$.

Закон больших чисел в форме Бернулли

Теорема 2 (Бернулли). Пусть m - число успехов в n испытаниях схемы Бернулли, p - вероятность успеха в одном испытании. Тогда

$$\boxed{\frac{m}{n} \xrightarrow{p} p}$$

при $n \to \infty$.

Доказательство. Пусть X_i - число успехов в i-м испытании, тогда X_i подчиняется распределению Бернулли, все $X_1, \ldots, X_n$ независимы и $EX_1 = p$. Положим $m = X_1 + \ldots + X_n$. Из теоремы 1 $\Rightarrow$ $\dfrac{m}{n} \xrightarrow{p} p$.

Закон больших чисел в форме Бернулли

Теорема 2 (Бернулли). Пусть m - число успехов в n испытаниях схемы Бернулли, p - вероятность успеха в одном испытании. Тогда

$$\boxed{\frac{m}{n} \xrightarrow{p} p}$$

при $n \to \infty$.

Доказательство. Пусть X_i - число успехов в i-м испытании, тогда X_i подчиняется распределению Бернулли, все $X_1, \ldots, X_n$ независимы и $EX_1 = p$. Положим $m = X_1 + \ldots + X_n$. Из теоремы 1 $\Rightarrow$ $\dfrac{m}{n} \xrightarrow{p} p$.

Частота наступления события при увеличении числа испытаний сходится к вероятности этого события.

Замечание. Согласно неравенству Чебышева,

$$P\left(\left|\frac{m}{n}-p\right|\geq\varepsilon\right)\leq\frac{1}{n}\frac{D\left(X_i\right)}{\varepsilon^2}=$$

Замечание. Согласно неравенству Чебышева,

$$P\left(\left|\frac{m}{n}-p\right|\geq\varepsilon\right)\leq\frac{1}{n}\frac{D\left(X_i\right)}{\varepsilon^2}=$$

$$=\frac{1}{n}\frac{p(1-p)}{\varepsilon^2}=\frac{pq}{n\varepsilon^2},$$

где $q=1-p$.

Замечание. Согласно неравенству Чебышева,

$$P\left(\left|\frac{m}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{n}\frac{D\left(X_i\right)}{\varepsilon^2} =$$

$$= \frac{1}{n}\frac{p(1-p)}{\varepsilon^2} = \frac{pq}{n\varepsilon^2},$$

где $q = 1 - p$.

То есть получили оценку отклонения частоты от вероятности события.

Пример. С целью установления доли брака было проверено 1000 единиц продукции. Какова вероятность того, что установленная этой выборкой доля брака по абсолютной величине будет отличаться от доли брака по всей партии не более чем на $0,01$, если известно, что в среднем на каждые $10\,000$ изделий приходится 500 бракованных?

Пример. С целью установления доли брака было проверено 1000 единиц продукции. Какова вероятность того, что установленная этой выборкой доля брака по абсолютной величине будет отличаться от доли брака по всей партии не более чем на $0,01$, если известно, что в среднем на каждые 10000 изделий приходится 500 бракованных?

Решение. $n = 1000$, $p = 500/10\ 000 = 0,05$; $q = 1 - p = 0,95$; $\varepsilon = 0,01$.

Пример. С целью установления доли брака было проверено 1000 единиц продукции. Какова вероятность того, что установленная этой выборкой доля брака по абсолютной величине будет отличаться от доли брака по всей партии не более чем на $0,01$, если известно, что в среднем на каждые $10\,000$ изделий приходится 500 бракованных?

Решение. $n = 1000$, $p = 500/10\ 000 = 0,05$; $q = 1 - p = 0,95$; $\varepsilon = 0,01$. Из неравенства Чебышева $\Rightarrow$

$$P(|m/n - p| < 0,01) > 1 - pq/(n\varepsilon^2) =$$

Пример. С целью установления доли брака было проверено 1000 единиц продукции. Какова вероятность того, что установленная этой выборкой доля брака по абсолютной величине будет отличаться от доли брака по всей партии не более чем на $0{,}01$, если известно, что в среднем на каждые $10\,000$ изделий приходится 500 бракованных?

Решение. $n = 1000$, $p = 500/10\,000 = 0{,}05$; $q = 1 - p = 0{,}95$; $\varepsilon = 0{,}01$. Из неравенства Чебышева $\Rightarrow$

$$P(|m/n - p| < 0{,}01) > 1 - pq/(n\varepsilon^2) =$$

$$= 1 - 0{,}05 \cdot 0{,}95 / \left(1000 \cdot 0{,}0001\right) = 0{,}527.$$