

ВЗАИМОСВЯЗЬ
РЕЗУЛЬТАТОВ
ИЗМЕРЕНИЯ.
МЕТОДЫ ВЫЧИСЛЕНИЯ
КОЭФФИЦИЕНТОВ
ВЗАИМОСВЯЗИ

ЦЕЛИ:

1. Изучить основные положения корреляционного анализа.
2. Изучить условия выбора коэффициента корреляции.
3. Изучить процедуру оценки статистической достоверности коэффициента корреляции.

1. Виды взаимосвязи.
2. Задачи корреляционного анализа.
3. Коэффициенты корреляции.
Основные свойства
коэффициента корреляции.
4. Оценка достоверности
коэффициента корреляции.

Виды взаимосвязи

1. Функциональная

каждому значению одного показателя соответствует одно, строго определенное значение другого показателя (пр-р: $S=V \cdot t$)

2. Статистическая каждому значению одного показателя может соответствовать несколько значений другого показателя

Корреляционная взаимосвязь - среднее значение одного показателя зависит от значения другого показателя
(пр-р: **связь между ростом и весом**)

Корреляционный анализ

Статистический метод, который используется для исследования взаимосвязей, называется *корреляционным анализом*

Основные задачи

определение:

- 1) формы взаимосвязи
- 2) направленности взаимосвязи
- 3) тесноты взаимосвязи

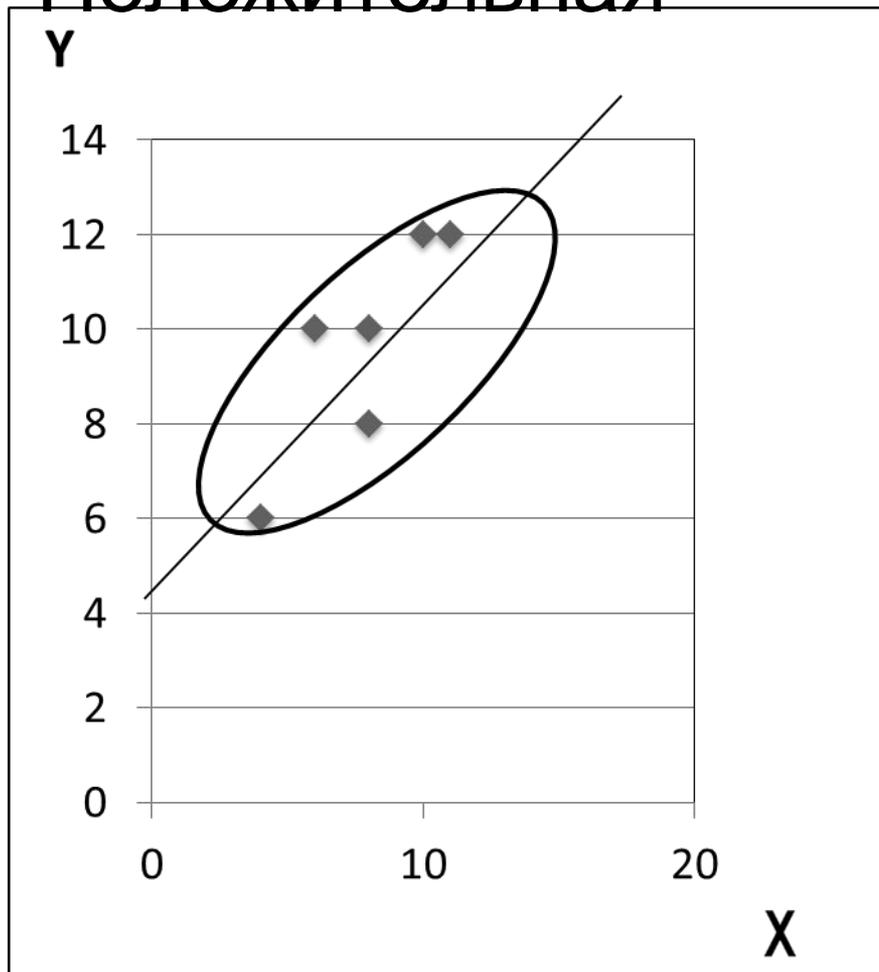
Определение формы

| N | X | Y |
|----------|----------|----------|
| 1 | 4 | 6 |
| 2 | 6 | 10 |
| 3 | 11 | 12 |
| 4 | 10 | 12 |
| 5 | 8 | 10 |
| 6 | 8 | 8 |

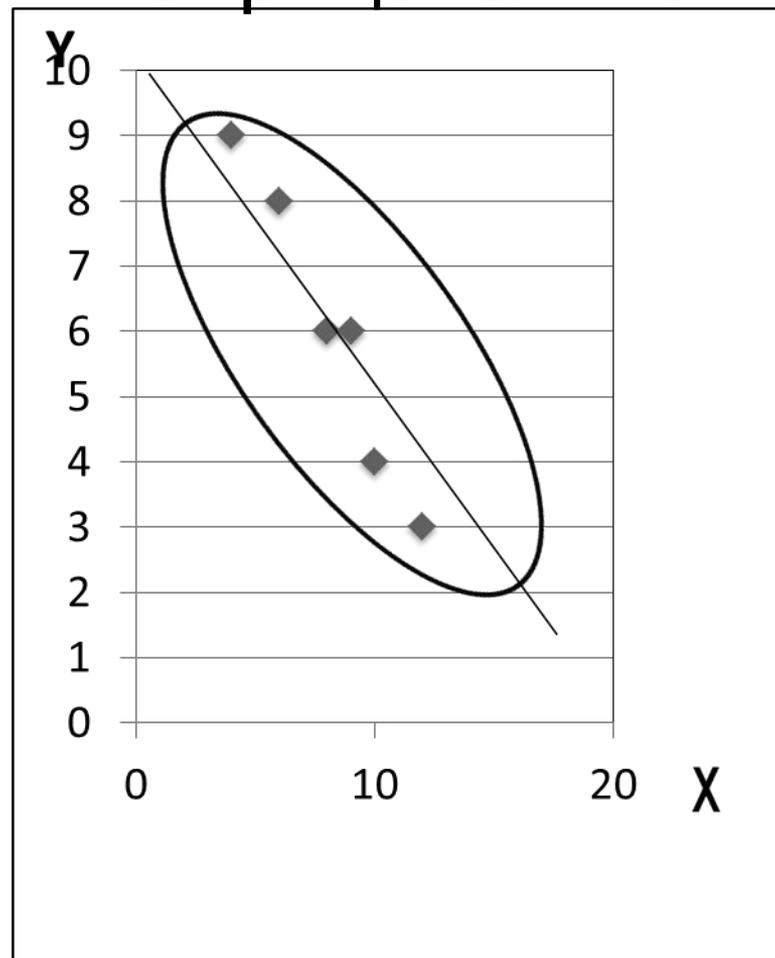


Определение направленности

Положительная

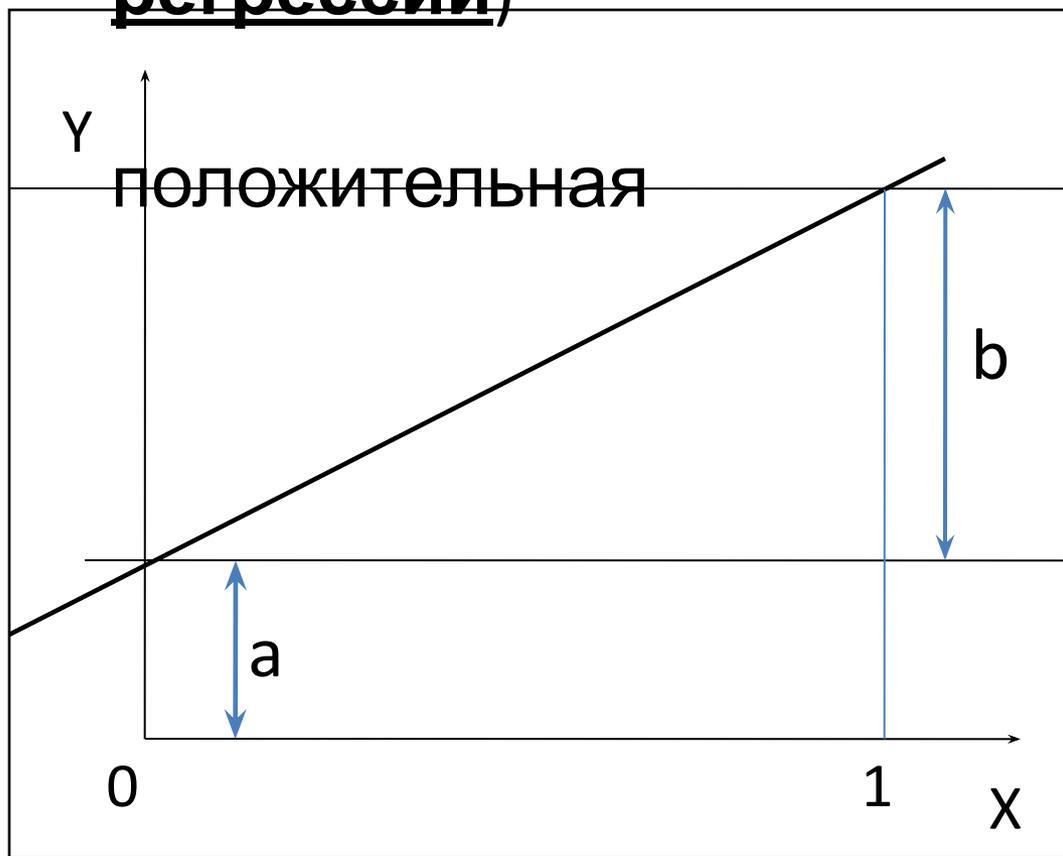


Отрицательная



Определение направленности

$Y=a+b \cdot x$ – уравнение прямой (уравнение регрессии)



$b > 0$ –

$b < 0$ – отрицательная

Определение тесноты

- визуально (по корреляционному полю)
- количественная мера (по коэффициенту корреляции, r)

Свойства КК

1. $-1 \leq r \leq 1$
2. $r < 0$ – взаимосвязь отрицательная
3. $r > 0$ – взаимосвязь положительная

| | | | | | | |
|-------------|---------|---------------|---------------|--------------|--------------------|-------------|
| r | -1 1 | 0,7-0,99 | 0,5-0,69 | 0,2-0,49 | 0,01-0,19 | 0 |
| взаимосвязь | Функц. | Сильная стат. | Средняя стат. | Слабая стат. | Очень слабая стат. | отсутствует |

Коэффициенты корреляции (КК)

Условия выбора КК

1. В какой шкале измерялся изучаемый показатель?
2. Можно ли считать ряд измерений выборкой, имеющей нормальный закон распределения?
3. Как много измерений этого показателя сделано, т.е. какой объем выборки?

Коэффициенты корреляции (КК)

1. Бравэ – Пирсона (r_{xy})

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y} ; r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

(коэффициент детерминации: $D = r^2 \cdot 100\%$)

Пр-р: $r=0,8$, то $D=0,64 \cdot 100\%=64\%$, т.е. взаимосвязь результатов измерений определяется на 64% , 36% - определяются другими причинами

Коэффициенты корреляции (КК)

2. Ранговый коэффициент Спирмена

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Пр-р:

d_x – ранг занятого места

d_y – ранг числа стартов

$$\rho = 1 - (6 \cdot 25,5) / 10 \cdot (10^2 - 1)$$

| № | X | Y | d_x | d_y | $d_x - d_y = d$ | d^2 |
|----|----|----|-------|-------|-----------------|-------------------|
| 1 | 1 | 9 | 1 | 2,5 | -1,5 | 2,25 |
| 2 | 2 | 10 | 2 | 1 | 1 | 1 |
| 3 | 3 | 8 | 3 | 4 | -1 | 1 |
| 4 | 4 | 7 | 4 | 5 | -1 | 1 |
| 5 | 5 | 9 | 5 | 2,5 | 2,5 | 6,25 |
| 6 | 6 | 4 | 6 | 7,5 | -1,5 | 2,25 |
| 7 | 7 | 4 | 7 | 7,5 | -0,5 | 0,25 |
| 8 | 8 | 3 | 8 | 9,5 | -1,5 | 2,25 |
| 9 | 9 | 5 | 9 | 6 | 3 | 9 |
| 10 | 10 | 3 | 10 | 9,5 | 0,5 | 0,25 |
| | | | | | | $\sum d^2 = 25,5$ |

Коэффициенты корреляции (КК)

3. Тетрахорический коэффициент

сопряженности

$$T_4 = \frac{|AD - BC| - 0,5n}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

Пр-р: $n=18$,

x, y – выборки

A – кол-во: $x=1, y=1$

B – кол-во: $x=0, y=1$

C – кол-во: $x=1, y=0$

D – кол-во: $x=0, y=0$

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|
| x | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| y | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |

$$A=7, B=4, C=5, D=2$$

$$A+B=11; C+D=7; A+C=12; B+D=6$$

$$T_4 = \frac{|7 \cdot 2 - 4 \cdot 5| - 0,5 \cdot 18}{\sqrt{11 \cdot 7 \cdot 12 \cdot 6}} = -0,040$$

Коэффициенты корреляции (КК)

4. Частный (порциальный)

$$r_{xyz} = \frac{r_{xy} - r_{xz} \cdot r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

5. Множественный

$$R_{xyz} = \sqrt{\frac{r_{xy}^2 + r_{xz}^2 - 2r_{xy} \cdot r_{xz} \cdot r_{yz}}{1 - r_{yz}^2}}$$

Оценка достоверности КК

Пр-р: Исследовали взаимосвязь (X,Y); n=31
нашли $r_{\text{набл}}=0,26$.

Существует ли взаимосвязь между
показателями?

($H_0: r_{\text{ген}}=0$) и ($H_1: |r_{\text{ген}}|>0$)

$r_{\text{крит}}=0,349$ (из таблицы, при $k=n-1=31-1=30$, $\alpha=0,05$)

0 $r_{\text{крит}}$ 1

Т.к. $|r_{\text{набл}}| < r_{\text{крит}}$ КК статистически

недостоверен взаимосвязь $r_{\text{ген}}=0$