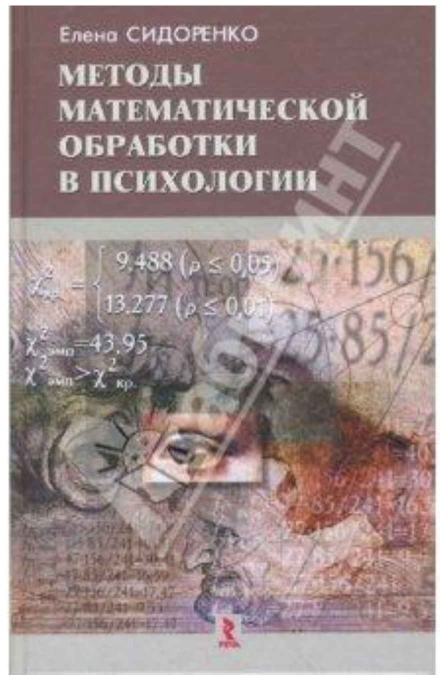
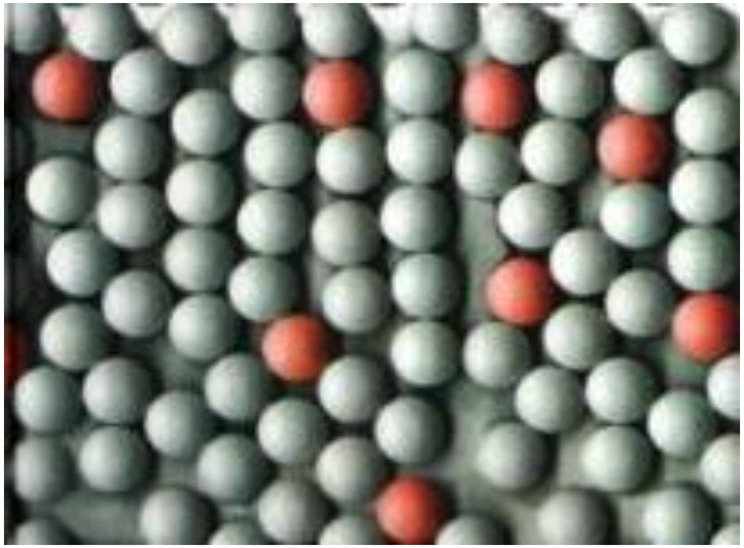
Статистические методы выявления корреляционной связи



Корреляционная связь

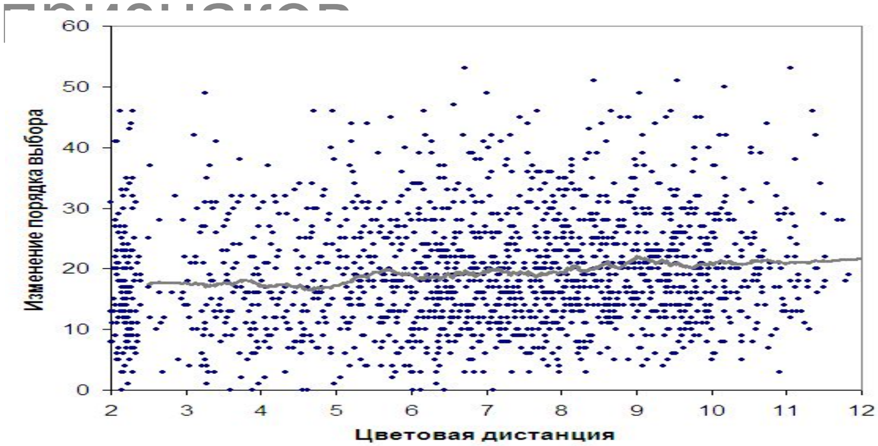
Используется для анализа массовых социально-экономических явлений, когда связь между признаками появляется лишь в среднем, в массе случаев.



Корреляционное поле

При изучении зависимости между двумя случайными величинами используют **поле корреляции.**

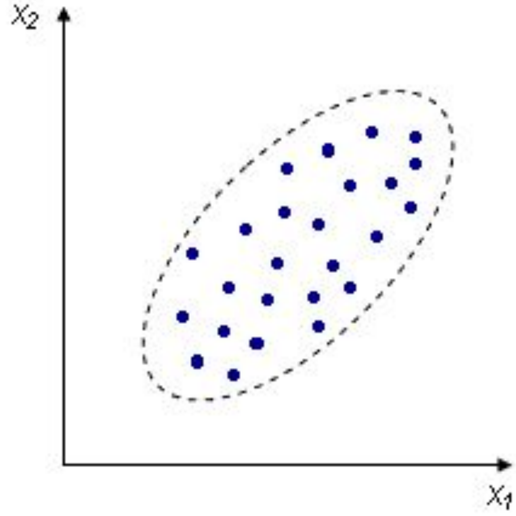
Поле корреляции представляет собой набор точек на плоскости с координатами, соответствующими значениям двух признаков.



Варианты распределения точек на плоскости

Положительная корреляция:

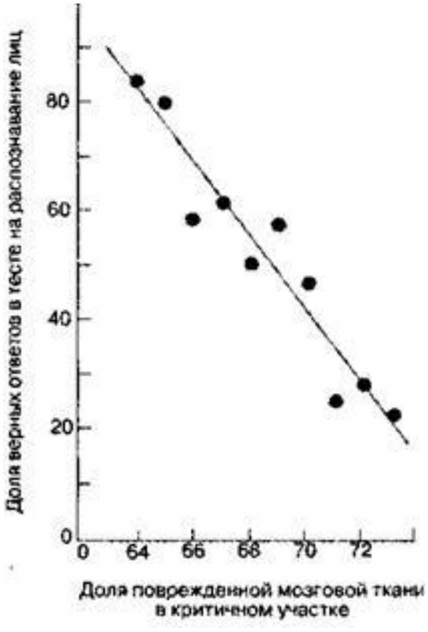
Основная масса точек укладывается на эллипсе, главная диагональ которой образует острый угол с осью .



Варианты распределения точек на плоскости

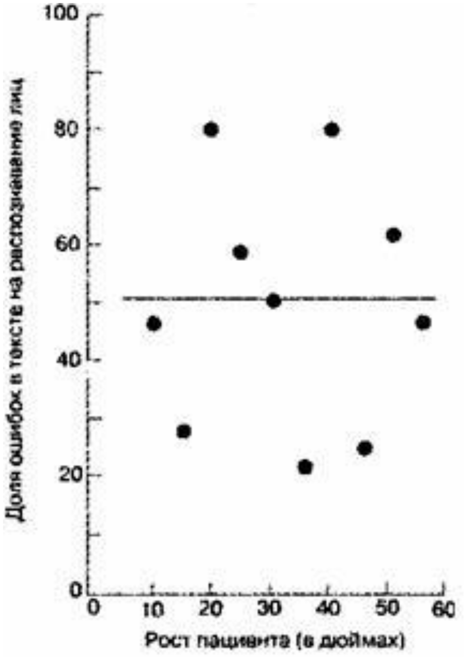
Отрицательная корреляция:

Основная масса точек укладывается на эллипсе, главная диагональ которой образует тупой угол с осью .



Варианты распределения точек на плоскости

Отсутствие корреляционной зависимости: Равномерное распределение точек в пространстве.

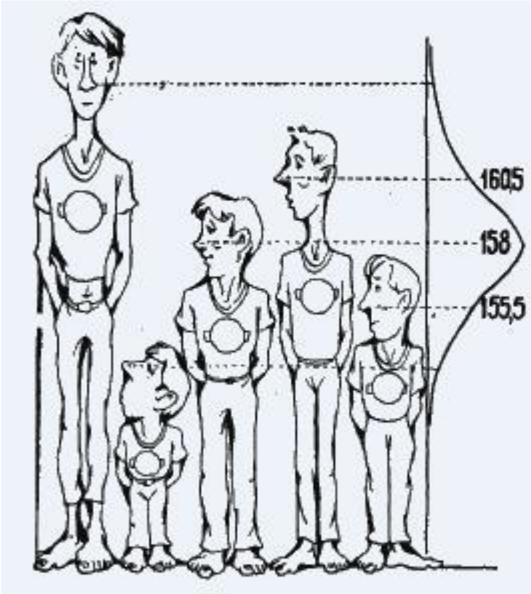


Выборочный коэффициент корреляции

Характеристики тесноты линейной связи между количественными признаками в выборке.

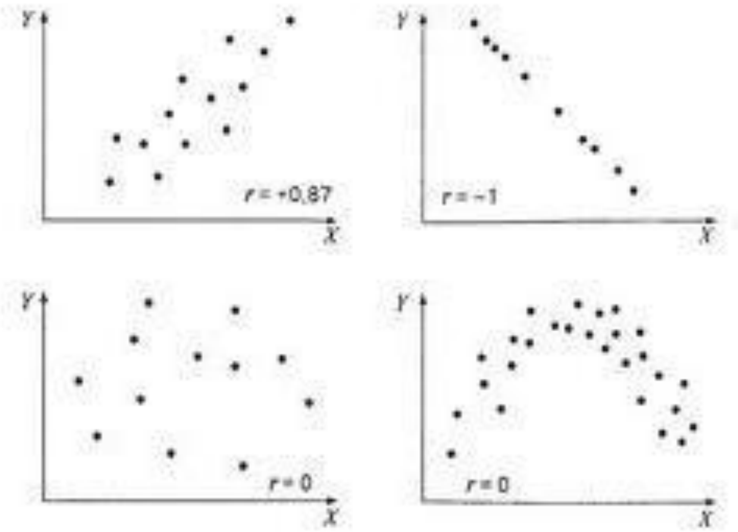
Выборочный коэффициент корреляции

Коэффициент корреляции не имеет размерности, его можно сопоставлять для разных статистических рядов.



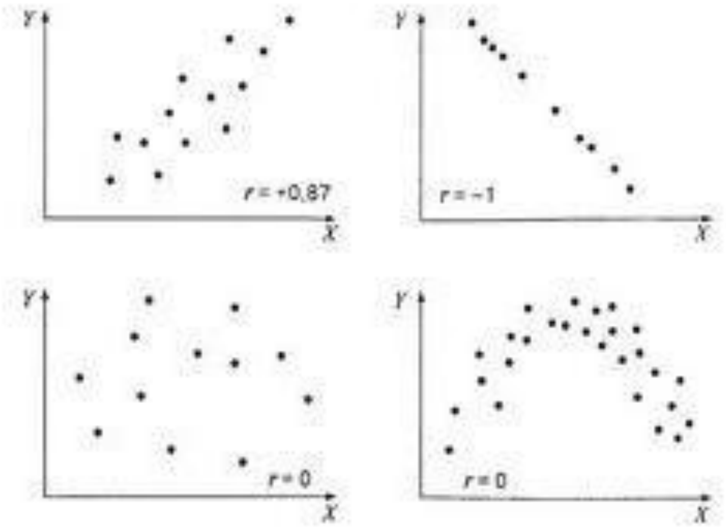
Выборочный коэффициент корреляции

Положительный знак коэффициента корреляции указывает на **положительную корреляцию**, т.е. все данные наблюдения лежат на прямой с положительным углом наклона



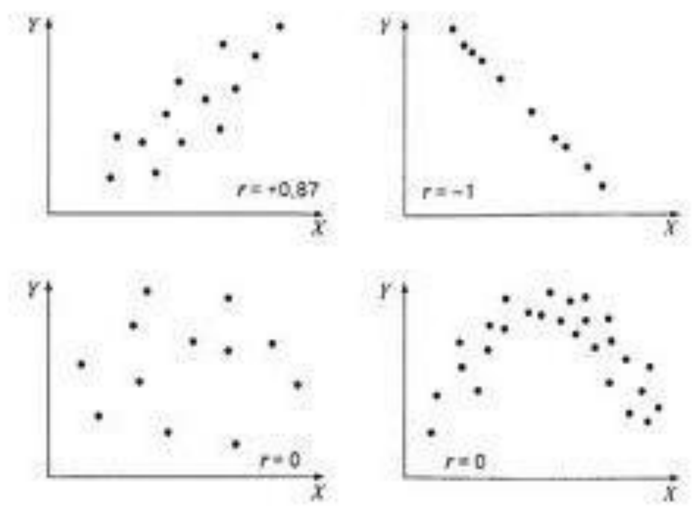
Выборочный коэффициент корреляции

Отрицательный знак коэффициента корреляции указывает на **отрицательную корреляцию**.



Выборочный коэффициент корреляции

Чем ближе значение к единице, тем теснее, а приближение к нулю означает ослабление линейной зависимости между переменными.



Выборочный коэффициент корреляции

$$r_{xy}^{*} = \frac{\alpha_{1,1}^{*}(x,y) - \alpha_{1}^{*}(x) \cdot \alpha_{1}^{*}(y)}{\sqrt{\mu_{2}^{*}(x)} \cdot \sqrt{\mu_{2}^{*}(y)}}.$$

Выборочный коэффициент корреляции

$$r_{xy}^* = \frac{\alpha_{1,1}^*(x,y) - \alpha_1^*(x)\cdot\alpha_1^*(y)}{\sqrt{\mu_2^*(x)}\cdot\sqrt{\mu_2^*(y)}}.$$

$$\alpha_1^*(x) = \frac{\sum_{i=1}^{k} n_i x_i}{n},\ \alpha_2^*(x) = \frac{\sum_{i=1}^{k} n_i x_i^2}{n},\ \mu_2^*(x) = \alpha_2^*(x) - \left[\alpha_1^*(x)\right]^2,$$

$$\alpha_1^*(y) = \frac{\sum_{i=1}^{k} n_i y_i}{n},\ \alpha_2^*(y) = \frac{\sum_{i=1}^{k} n_i y_i^2}{n},\ \mu_2^*(y) = \alpha_2^*(y) - \left[\alpha_1^*(y)\right]^2.$$

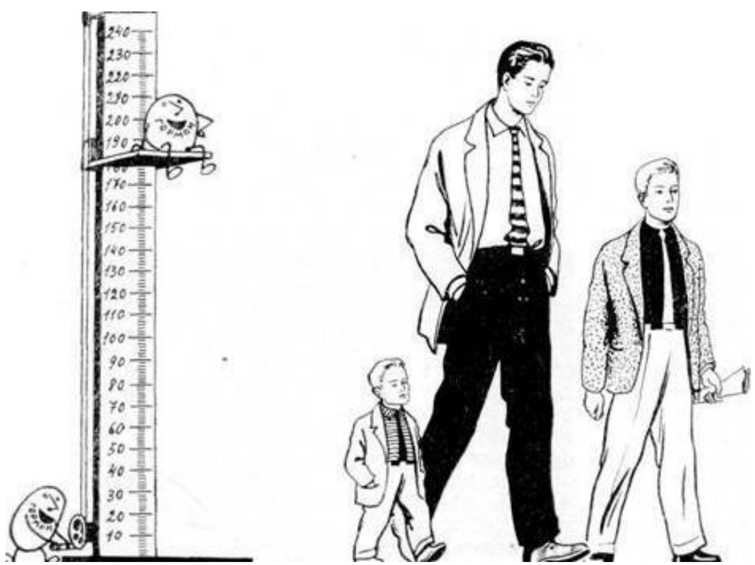
Свойства выборочный коэффициента корреляции

1. $r_{xy} = r_{yx}$.
2. $-1 \le r_{xy}^{*} \le 1$.
3. Если $\left| r_{xy}^{*} \right| = 1$ тогда и только тогда, когда между значениями X, Y имеется линейная зависимость.
4. Если r_{xy}^{*} =0, то между X, Y отсутствует линейная корреляционная связь, но возможно наличие между ними другого типа связи.
5. Если r_{xy}^{*} >0, то увеличение признака X в среднем приводит к увеличению признака Y. Если r_{xy}^{*} <0, то с увеличением X в среднем признак Y уменьшается.

Лабораторная работа: выборочный коэффициент корреляции

Случайные величинь X - рост человека,
Y - вес человека,
Z - размер обуви.

1. Изобразить систему случайных величин.
2. Найти выборочный коэффициент корреляции

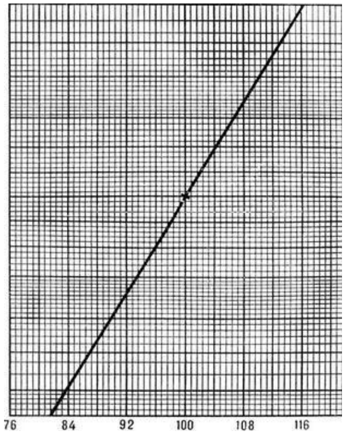


Указания по выполнению задания лабораторной работы

Если элементы системы не повторяются и выборка не большая, то удобно пользоваться таблицей

N	x_i	y_i	x_i^2	y_i^2	$x_i \cdot y_i$
1					
…					
n					
$\sum_{i=1}^{n}$	$\sum_{i=1}^{n} x_i$	$\sum_{i=1}^{n} y_i$	$\sum_{i=1}^{n} x_i^2$	$\sum_{i=1}^{n} y_i^2$	$\sum_{i=1}^{n} x_i \cdot y_i$
$\frac{\sum_{i=1}^{n}}{n}$	$\frac{\sum_{i=1}^{n} x_i}{n} = \alpha_1^*(x)$	$\frac{\sum_{i=1}^{n} y_i}{n} = \alpha_1^*(y)$	$\frac{\sum_{i=1}^{n} x_i^2}{n} = \alpha_2^*(x)$	$\frac{\sum_{i=1}^{n} y_i^2}{n} = \alpha_2^*(y)$	$\frac{\sum_{i=1}^{n} x_i \cdot y_i}{n} = \alpha_{1,1}^*(x,y)$

Лабораторная работа: выборочный коэффициент корреляции



Вариант 1

Для двумерной случайной величины (X,Y)

1. Изобразить систему случайных величин.
2. Найти выборочный коэффициент корреляции.

Лабораторная работа: выборочный коэффициент корреляции

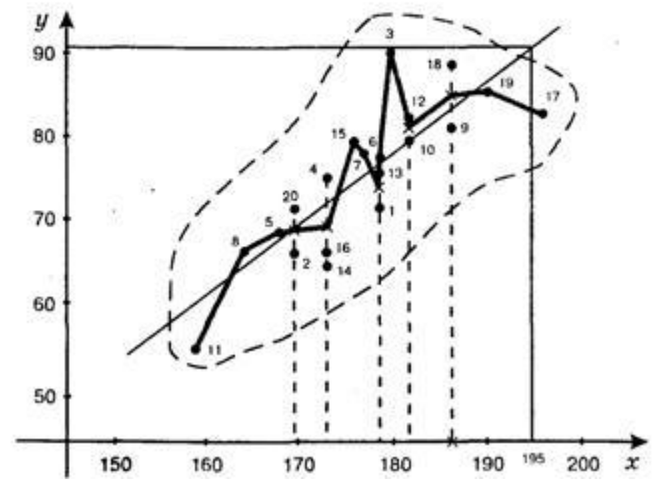
i	x_i	y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	3,9	2,3	–25,29	639,58	–10,11	102,21	255,68
2	0,7	0,3	–28,49	811,68	–12,11	146,65	345,01
3	0,5	0,3	–28,69	823,12	–12,11	146,65	347,44
4	2,4	0,8	–26,79	717,70	–11,61	134,79	311,03
5	5,3	1,2	–23,89	570,73	–11,21	125,66	267,81
6	7,7	2	–21,49	461,82	–10,41	108,37	223,71
7	11	2,4	–18,19	330,88	–10,01	100,20	182,08
8	16,6	4,1	–12,59	158,51	–8,31	69,06	104,62
9	21,2	5,7	–7,99	63,84	–6,71	45,02	53,61
10	27,9	7,8	–1,29	1,66	–4,61	21,25	5,95
11	37,4	11,7	8,21	67,40	–0,71	0,50	–5,83
12	46,6	14,7	17,41	303,11	2,29	5,24	39,87
13	65,5	21,7	36,31	1318,42	9,29	86,30	337,32
14	73,7	26,4	44,51	1981,14	13,99	195,72	622,69
15	117,5	84,7	88,31	7798,66	72,29	5225,84	6383,93
Всего	$\Sigma = 437{,}9$ $\bar{x} = 29{,}19$	$\Sigma = 186{,}1$ $\bar{y} = 12{,}41$		$\Sigma =$ 16 048,25		$\Sigma =$ 6513,49	$\Sigma =$ 9474,93

Вариант 2

Для двумерной случайной величины (X, Z)

1. Изобразить систему случайных величин.
2. Найти выборочный коэффициент корреляции.

Лабораторная работа: выборочный коэффициент корреляции



Вариант 3

Для двумерной случайной величины (Y, Z)

1. Изобразить систему случайных величин.
2. Найти выборочный коэффициент корреляции.

Статистические методы выявления корреляционной связи