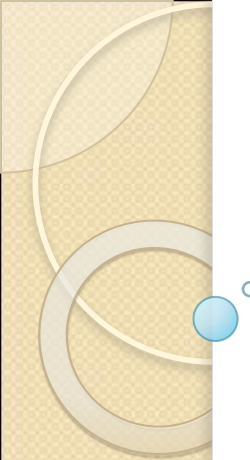




СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ ВЗАИМОСВЯЗИ СОЦИАЛЬНО- ЭКОНОМИЧЕСКИХ ЯВЛЕНИЙ



Причинно-следственные отношения – это связь явлений и процессов, при которой изменение одного из них – причины – ведет к изменению другого – следствия.

Причина – это совокупность условий, обстоятельств, действие которых приводит к появлению следствия.

Признаки, обуславливающие изменения других, связанных с ними признаков, называются **факторными (X)**.

Признаки, изменяющиеся под действием факторных признаков, называются **результативными (Y)**.

В статистике различают **функциональную связь** и **статистическую (стохастическую) связь**.

Функциональной называют такую связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака.

$$y = f(x)$$

Если причинная зависимость проявляется не в каждом отдельном случае, а в общем, среднем при большом числе наблюдений, то такая зависимость называется **стохастической**. Частным случаем стохастической является **корреляционная связь**.

Виды статистической взаимосвязи

По аналитическому
выражению

линейная связь

нелинейная связь

По направлению
связи

прямая
зависимость

обратная
зависимость

По степени тесноты
связи

слабая

умеренная

сильная

По направлению связи бывают:

– **прямыми (положительными)**, с увеличением или уменьшением значений факторного признака происходит увеличение или уменьшение значений результативного признака.

– **обратными (отрицательными)**, значения результативного признака изменяются в противоположном направлении по сравнению с изменением факторного признака.

По аналитическому выражению выделяют связи **прямолинейные** (или просто линейные) и **нелинейные** (криволинейные).

Если статистическая связь между явлениями приближенно выражена:

- 1) уравнением прямой линии, то ее называют **линейной связью**;
- 2) уравнением какой-либо кривой линии (параболы, гиперболы: степенной, показательной, экспоненциальной и т.д.), то такую связь называют **нелинейной** или **криволинейной**.

Для выявления наличия связи, ее характера и направления в статистике используются различные методы.

1. Метод приведения параллельных данных.
2. Графический метод (с помощью поля корреляции).
3. Корреляция.
4. Регрессия.

Метод приведения параллельных данных основан на сопоставлении двух или нескольких рядов статистических величин. Такое сопоставление позволяет установить наличие связи и получить представление о ее характере.

Номер студента	Балл в сессию,	Количество пропущенных семинаров, раз		
1	5	1		
2	3	8		
3	4	3		
4	4	3		
5	3	8		
6	2	10		
7	5	2		
8	4	4		
9	5	2		
10	3	6		

Номер студента	Балл в сессию, у	Количество пропущенных семинаров, раз, х	Приведенные параллельные данные	
			х	у
1	5	1	1	5
2	3	8	2	5
3	4	3	2	5
4	4	3	3	4
5	3	8	4	4
6	2	10	5	4
7	5	2	6	3
8	4	4	8	3
9	5	2	8	3
10	3	6	10	2

Взаимосвязь двух признаков изображается графически с помощью ***поля корреляции***.

Поле корреляции – это поле точек, на котором каждая точка соответствует единице совокупности; ее координаты определяются значениями признаков ***X*** и ***Y***.

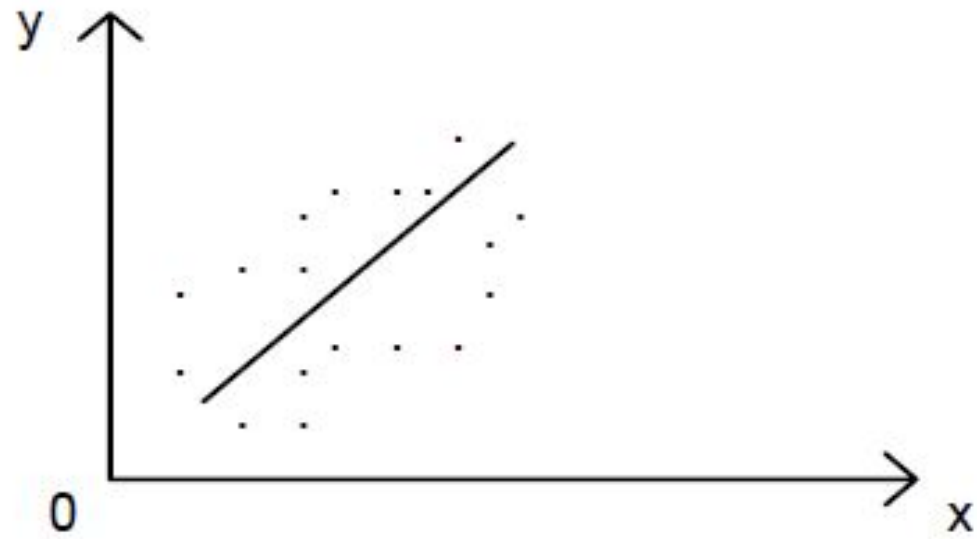


Рис. График корреляционного поля

Корреляционный метод имеет своей задачей количественное определение тесноты связи между двумя признаками (при парной связи) и между результативным и множеством факторных признаков (при многофакторной связи).

Теснота связи количественно выражается величиной коэффициентов корреляции.

В статистике принято различать следующие виды зависимостей:

1. Парная корреляция - связь между двумя признаками (результативным и факторным).

2. Множественная корреляция - влияние нескольких факторов на результативный признак.

Корреляционный анализ

Корреляционный анализ начинается с расчета линейных (парных) коэффициентов корреляции.

$$r = \frac{\sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2}}$$

При изучении совокупностей малого объема ($n \leq 30$) пользуются следующей формулой расчета линейного коэффициента корреляции:

$$r = \frac{n \cdot \sum (x \cdot y) - \sum x \cdot \sum y}{\sqrt{[n \cdot \sum x^2 - (\sum x)^2] \cdot [n \cdot \sum y^2 - (\sum y)^2]}}$$

Количественные критерии оценки тесноты связи

Величина коэффициента корреляции	Характер связи
До $\pm 0,3$	практически отсутствует
$\pm 0,3$ - $\pm 0,5$	слабая
$\pm 0,5$ - $\pm 0,7$	умеренная
$\pm 0,7$ - $\pm 1,0$	сильная

Оценка линейного коэффициента корреляции

Значение линейного коэффициента связи	Характер связи	Интерпретация связи
$r = 0$	Отсутствует	-
$0 < r < 1$	Прямая	С увеличением x увеличивается y
$-1 < r < 0$	Обратная	С увеличением x уменьшается y
$r = 1$	Функциональная	Каждому значению факторного признака строго соответствует одно значение результативного признака.

Корреляционный анализ

Множественный коэффициент корреляции.

$$R_{y/x_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

где r – парные коэффициенты корреляции между признаками.

Множественный коэффициент корреляции изменяется в пределах от 0 до 1 и по определению положителен: $0 \leq R \leq 1$.
Приближение коэффициента к единице свидетельствует о сильной зависимости между признаками.

Регрессионный метод заключается в определении аналитического выражения связи, в котором изменение одной величины (называемой зависимой или результативным признаком) обусловлено влиянием одной или нескольких независимых величин (факторов), а множество всех прочих факторов, также оказывающих влияние на зависимую величину, принимается за постоянные и средние значения.

Аналитически связь между ними описывается уравнениями:

прямой $\bar{y}_x = a_0 + a_1x$

параболы $\bar{y}_x = a_0 + a_1x + a_2x^2$

гиперболы $\bar{y}_x = a_0 + a_1 \cdot \frac{1}{x}$

Оценка параметров уравнения регрессии a_0 и a_1 осуществляется *методом наименьших квадратов* (МНК) .

$$\begin{cases} na_0 + a_1 \sum x = \sum y \\ a_0 \sum x + a_1 \sum x^2 = \sum xy \end{cases}$$

В уравнениях регрессии параметр a_0 показывает усредненное влияние на результативный признак неучтенных (невыделенных для исследования) факторов; параметр a_1 - коэффициент регрессии, показывает, насколько изменяется в среднем значение результативного признака при изменении факторного на единицу его собственного измерения.

Если связь между признаками y и x нелинейная и описывается **уравнением параболы второго порядка**,

$$y_x = a_0 + a_1x + a_2x^2.$$

В данном случае задача сводится к определению неизвестных параметров: a_0, a_1, a_2 . Параметры находят по МНК, и система уравнений имеет вид:

$$\begin{cases} na_0 + a_1 \sum x + a_2 \sum x^2 = \sum y, \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 = \sum xy, \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 = \sum yx^2. \end{cases}$$

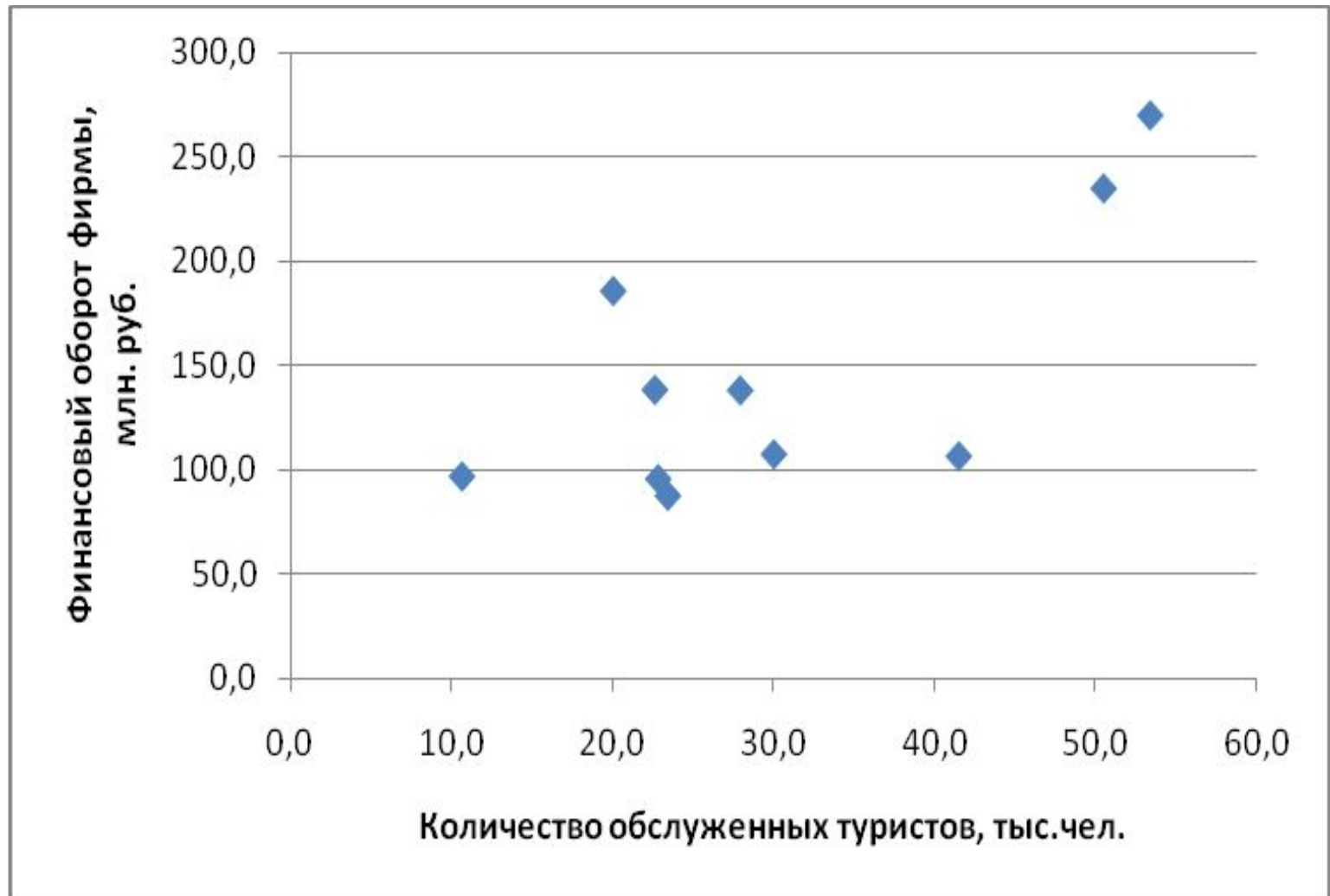
Показатели финансового оборота и количества обслуженных клиентов у туристических фирм за 2010 год

№ п/п	Финансовый оборот (млн. руб.)	Количество обслуженных туристов (тыс. чел.)
1	270,0	53,4
2	235,0	50,5
3	186,0	20,0
4	138,6	22,6
5	138,4	27,9
6	107,8	30,0
7	106,9	41,5
8	97,2	10,6
9	96,0	22,8
10	88,0	23,4

Показатели финансового оборота и количества обслуженных клиентов у туристических фирм за 2010 год

№ п/п	Финансовый оборот (млн. руб.) y_i	Количество обслуженных туристов (тыс. чел.) x_i
1	270,0	53,4
2	235,0	50,5
3	186,0	20,0
4	138,6	22,6
5	138,4	27,9
6	107,8	30,0
7	106,9	41,5
8	97,2	10,6
9	96,0	22,8
10	88,0	23,4

Поле корреляции, характеризующее финансовый оборот туристических фирм и количество обслуженных ими клиентов



№ п/п	Финансовый оборот (млн. руб.) y_i	Количество обслуженных туристов (тыс. чел.) x_i	x_i^2			
1	270,0	53,4	2851,56			
2	235,0	50,5	2550,25			
3	186,0	20,0	400,00			
4	138,6	22,6	510,76			
5	138,4	27,9	778,41			
6	107,8	30,0	900,00			
7	106,9	41,5	1722,25			
8	97,2	10,6	112,36			
9	96,0	22,8	519,84			
10	88,0	23,4	547,56			
	1463,86	302,7	10892,99			

№ п/п	Финансовый оборот (млн. руб.) y_i	Количество обслуженных туристов (тыс. чел.) x_i	x_i^2	$x_i \cdot y_i$		
1	270,0	53,4	2851,56	14418,00		
2	235,0	50,5	2550,25	11867,50		
3	186,0	20,0	400,00	3720,00		
4	138,6	22,6	510,76	3132,36		
5	138,4	27,9	778,41	3861,36		
6	107,8	30,0	900,00	3234,00		
7	106,9	41,5	1722,25	4435,52		
8	97,2	10,6	112,36	1030,32		
9	96,0	22,8	519,84	2188,34		
10	88,0	23,4	547,56	2059,20		
	1463,86	302,7	10892,99	49946,60		

№ п/п	Финансовый оборот (млн. руб.) y_i	Количество обслуженных туристов (тыс. чел.) x_i	x_i^2	$x_i \cdot y_i$	\bar{y}_x	
1	270,0	53,4	2851,56	14418,00	221,72	
2	235,0	50,5	2550,25	11867,50	212,27	
3	186,0	20,0	400,00	3720,00	112,94	
4	138,6	22,6	510,76	3132,36	121,40	
5	138,4	27,9	778,41	3861,36	138,67	
6	107,8	30,0	900,00	3234,00	145,51	
7	106,9	41,5	1722,25	4435,52	182,96	
8	97,2	10,6	112,36	1030,32	82,32	
9	96,0	22,8	519,84	2188,34	122,05	
10	88,0	23,4	547,56	2059,20	124,01	
	1463,86	302,7	10892,99	49946,60	1463,84	

№ п/п	Финансовый оборот (млн. руб.) y_i	Количество обслуженных туристов (тыс. чел.) x_i	x_i^2	$x_i \cdot y_i$	\bar{y}_x	y_i^2
1	270,0	53,4	2851,56	14418,00	221,72	72900
2	235,0	50,5	2550,25	11867,50	212,27	55225
3	186,0	20,0	400,00	3720,00	112,94	34596
4	138,6	22,6	510,76	3132,36	121,40	19209,96
5	138,4	27,9	778,41	3861,36	138,67	19154,56
6	107,8	30,0	900,00	3234,00	145,51	11620,84
7	106,9	41,5	1722,25	4435,52	182,96	11423,33
8	97,2	10,6	112,36	1030,32	82,32	9447,84
9	96,0	22,8	519,84	2188,34	122,05	9212,16
10	88,0	23,4	547,56	2059,20	124,01	7744
	1463,86	302,7	10892,99	49946,60	1463,84	250533,69