Статистическое изучение взаимосвязей финансово-экономических явлений

Для большинства статистических исследований финансовой сферы важно выявить существующие взаимосвязи между финансовыми явлениями и процессами.

Почти все наблюдаемые явления экономической жизни общества следствие действия определенных факторов.

Например, получаемая предприятием прибыль связана с показателями: численностью работников, объемом основных производственных фондов и т.п.

Между общественными и экономическими явлениями существует два основных типа связи — **функциональная** и **статистическая** (называемая также стохастической, или вероятностной).

Кроме того, выделяют ***корреляционную*** связь, которая является частным случаем статистической связи.

Независимыми, или *факторными,* называют признаки, которые вызывают изменения других, связанных с ними, признаков.

Признаки, изменение которых под воздействием определенных факторов требуется проследить, называют ***зависимыми,*** или *результативными*.

При *функциональной* связи изменение независимых переменных приводит к получению точно определенных значений зависимой переменной. Например, если обозначить через *X* независимую переменную, а через *Y* – зависимую, связь $Y=X^3+5$ будет функциональной, так как каждому значению *X* соответствует точно определенное значение *Y* (при $X=0$ значение $Y=5$, при $X=3$ значение $Y=14$ и т.д.), причем это значение не обязательно должно быть единственным. Так, функциональная зависимость вида

$$Y = \sqrt{X} + 5$$

позволит получить не одно, а два значения *Y* (например, при $X=1$ значения $Y = 4$ и 6).

Наиболее часто функциональные связи проявляются при изучении физических явлений, например в механике функциональной является зависимость расстояния, пройденного объектом, от скорости его движения и т.п.

В сфере финансов и в экономике в целом функциональные зависимости также наблюдаются довольно часто – это плата за кредит, начисляемая на основе установленной процентной ставки;

показатель доходности ценной бумаги, находящийся в функциональной зависимости от курса ценной бумаги;

показатели рентабельности, фондоемкости и фондоотдачи, функционально зависящие от объема продукции и стоимости основных фондов и т.д.

При *статистической* связи каждому значению независимой переменной *X* соответствует множество значений зависимой переменной *Y*, причем неизвестно заранее, какое именно.

Например, прибыль коммерческого банка связана с размером уставного капитала. Но нельзя вычислить точную величину прибыли при заданном значении уставного капитала, так как она зависит еще и от множества других факторов, среди которых имеются и случайные, действие которых приводит к статистической зависимости.

Таким образом, статистическая связь отличается от функциональной наличием действия на зависимую переменную большого числа факторов, как *выявленных* с целью описания зависимости в математической форме, так и *случайных*, действие которых трудно учесть при построении модели или же учитывать нецелесообразно ввиду их слабого влияния на зависимую переменную.

***Корреляционной* является статистическая связь между признаками, при которой изменение значений независимой переменной *X* приводит к закономерному *изменению математического ожидания* случайной величины *Y*.**

Предположим, что независимая переменная *X* приняла значение *x1*, тогда зависимая переменная *Y* примет множество значений $\{y_{11}, y_{12}, y_{13}, y_{14} \ldots\}$

с условным математическим ожиданием $M(Y|\mathrm{X} = \mathrm{x1})$

$$X = x_1 \qquad \Rightarrow \qquad Y_1 \in \{y_{11}, y_{12}, y_{13}, y_{14} \ldots\} \to M(Y|\mathrm{X} = \mathrm{x1}).$$

$$\text{При } X = x_2 \quad \text{получим} \quad Y_2 \in \{y_{21}, y_{22}, y_{23}, y_{24} \ldots\} \to M(Y|\mathrm{X} = \mathrm{x2})$$

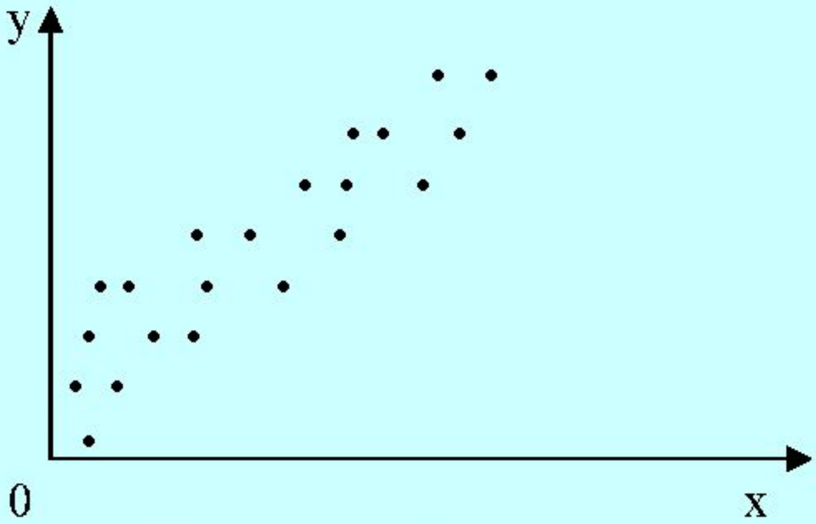
$$\text{при } X = x_3 \qquad \Rightarrow Y_3 \in \{y_{31}, y_{32}, y_{33}, y_{34} \ldots\} \to M(Y|\mathrm{X} = \mathrm{x3})$$

Если имеются закономерности в изменении условных математических ожиданий при изменении значений *X*, рассматриваемая связь между *X* и *Y* будет корреляционной.

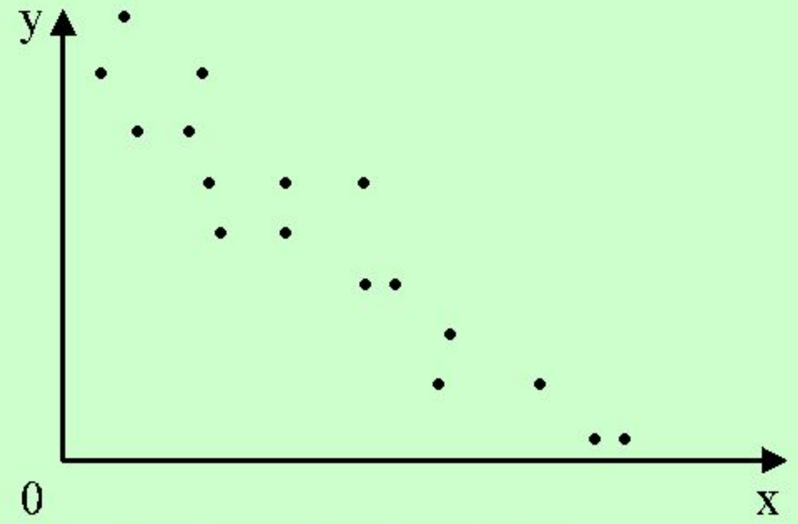
Корреляционная связь, как и функциональная, может быть прямой (положительной) или обратной (отрицательной).

Прямая и обратная зависимости характеризуют направление связи между признаками, которую можно проиллюстрировать графически с помощью **поля корреляции**.

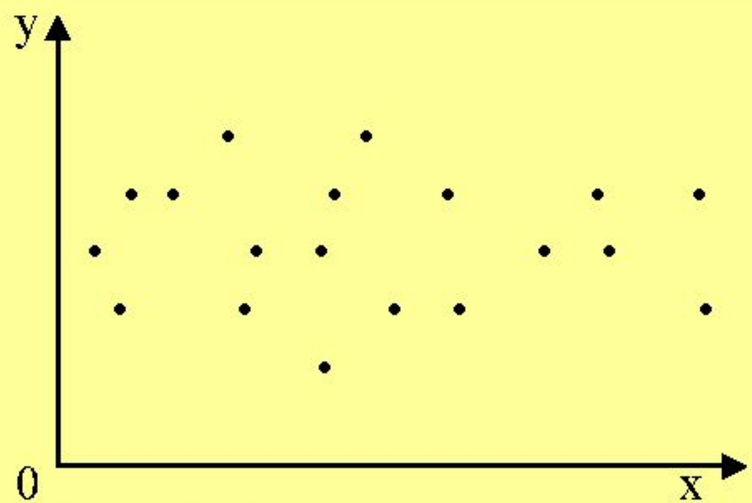
При его построении в прямоугольной системе координат на оси абцисс располагают значения независимой переменной *x*, а на оси ординат – зависимой *y*. Пересечение координат обозначают точками, которые символизируют наблюдения. По форме рассеяния точек на корреляционном поле судят о форме и тесноте связи.



Прямая (положительная) связь



Обратная (отрицательная) связь



Отсутствие связи

Корреляционный анализ

	Y	X_1	X_2	X_3
1	15,3	0,34	7,9	29,5
2	15,2	0,02	7,8	36,0
3	14,5	0,33	7,1	30,5
4	14,1	0,13	7,3	35,6
5	13,8	0,28	6,9	20,7
6	10,2	0,11	5,9	30,4
7	7,8	0,16	3,8	30,1
8	7,4	0,08	3,4	33,9
9	7,3	0,18	3,2	35,4
10	5,5	0,03	2,7	25,2
11	4,8	0,14	2,9	23,7
12	4,4	0,48	2,1	25,1
13	3,7	0,91	2,0	27,4
14	2,1	0,17	0,9	27,3
15	2,0	0,13	0,8	32,1
16	1,8	0,21	0,6	11,4
17	1,6	0,18	0,4	30,9
18	1,4	0,15	0,3	22,2
19	0,8	0,37	0,4	26,1
20	0,7	0,58	0,5	9,3

С помощью методов корреляционного анализа исследуем зависимость показателя прибыли предприятия *(Y)* от следующих факторов:

- затрат на рекламу (*X1)*,
- материальных затрат *(X2),*
- объема основных фондов *(X3).*

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

Парный коэффициент корреляции представляет собой меру *линейной* зависимости между двумя переменными на фоне действия остальных переменных, входящих в модель.

$$r_{yx} = \frac{\overline{yx} - \overline{y} \times \overline{x}}{\sigma_y \times \sigma_x}$$ где

$\overline{x}$ – среднее арифметическое значение *x*;

$\overline{y}$ – среднее арифметическое значение *y*;

$\overline{yx}$ – среднее арифметическое значение из произведений *y* и *x*;

σ_y – среднеквадратическое отклонение признака *y*;

σ_x – среднеквадратическое отклонение признака *x*.

Парный коэффициент корреляции изменяется
в пределах от -1 до +1.

Абсолютное значение, равное единице, свидетельствует о том, что связь функциональная:
-1 – обратная (отрицательная), +1 – прямая (положительная).

Нулевое значение коэффициента указывает на отсутствие линейной связи между признаками.

*Значение коэффициента корреляции (по модулю)**	*Качественная характеристика силы связи*
до 0,3	Практически отсутствует (слабая)
0,3–0,7	Средняя
0,7–0,9	Высокая
0,9–0,99	Весьма высокая

	Y	X_1	X_2	X_3
Сумма	134,4	38,64	66,9	542,8
Среднее значение	**6,72**	**1,932**	**3,345**	**27,14**
Дисперсия	28,317	57,017	7,811	52,2
Среднее квадратическое отклонение	**5,321**	**7,551**	**2,795**	**7,225**

yx_1	yx_2	yx_3	x_1x_2	x_1x_3	x_2x_3
Сумма 543,773	730,580	3996,7	280,338	1117,66	1990,87
Среднее значение 27,189	36,529	199,835	14,017	55,883	99,544

Определим коэффициенты корреляции:

$$r_{yx_1} = \frac{\overline{yx_1} - \overline{y} \times \overline{x_1}}{\sigma_y \times \sigma_{x_1}} = \frac{27{,}189 - 6{,}72 \cdot 1{,}932}{5{,}321 \cdot 7{,}551} = 0{,}354$$

$$r_{yx_2} = \frac{\overline{yx_2} - \overline{y} \times \overline{x_2}}{\sigma_y \times \sigma_{x_2}} = \frac{36{,}529 - 6{,}72 \cdot 3{,}345}{5{,}321 \cdot 2{,}795} = 0{,}945$$

$$r_{yx_3} = \frac{\overline{yx_3} - \overline{y} \times \overline{x_3}}{\sigma_y \times \sigma_{x_3}} = \frac{199{,}835 - 6{,}72 \cdot 27{,}14}{5{,}321 \cdot 7{,}225} = 0{,}454$$

.

$$r_{x_1x_2} = \frac{\overline{x_1x_2} - \overline{x_1} \times \overline{x_2}}{\sigma_{x_1} \times \sigma_{x_2}} = \frac{14,017 - 1,932 \cdot 3,345}{7,551 \cdot 2,795} = 0,358$$

$$r_{x_1x_3} = \frac{\overline{x_1x_3} - \overline{x_1} \times \overline{x_3}}{\sigma_{x_1} \times \sigma_{x_3}} = \frac{55,883 - 1,932 \cdot 27,14}{7,551 \cdot 7,225} = 0,063$$

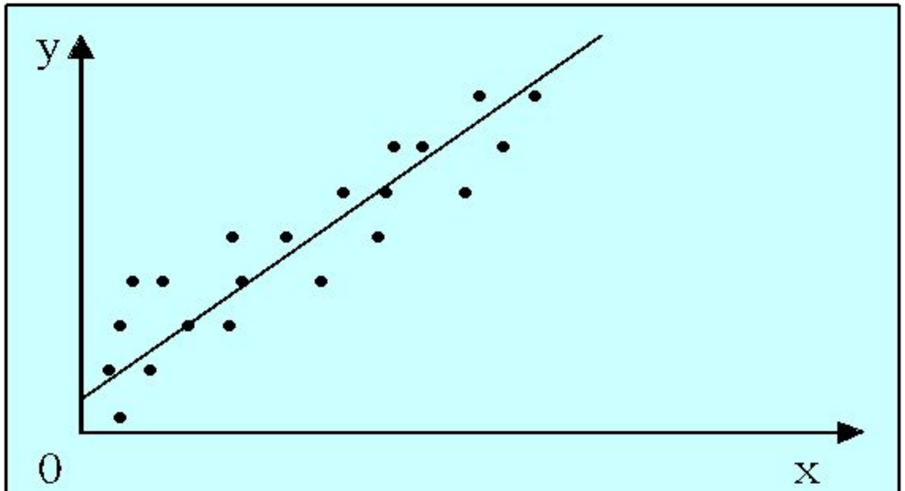
$$r_{x_2x_3} = \frac{\overline{x_2x_3} - \overline{x_2} \times \overline{x_3}}{\sigma_{x_2} \times \sigma_{x_3}} = \frac{99,544 - 3,345 \cdot 27,14}{2,795 \cdot 7,225} = 0,434$$

Корреляционная матрица будет иметь вид:

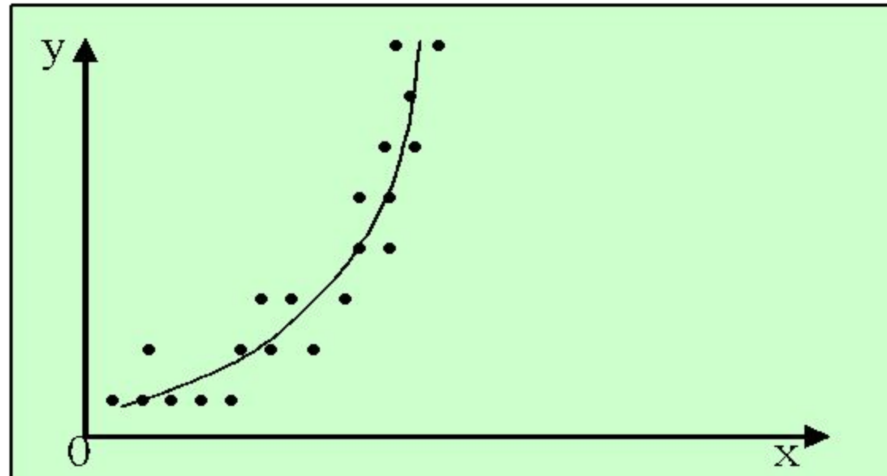
$$R = \begin{array}{c} \\ y \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{cccc} y & x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 1 & 0{,}354 & 0{,}945 & 0{,}454 \\ 0{,}354 & 1 & 0{,}358 & 0{,}063 \\ 0{,}945 & 0{,}358 & 1 & 0{,}434 \\ 0{,}454 & 0{,}063 & 0{,}434 & 1 \end{pmatrix} \end{array}$$

Наибольшее влияние на результативный признак оказывает второй фактор.

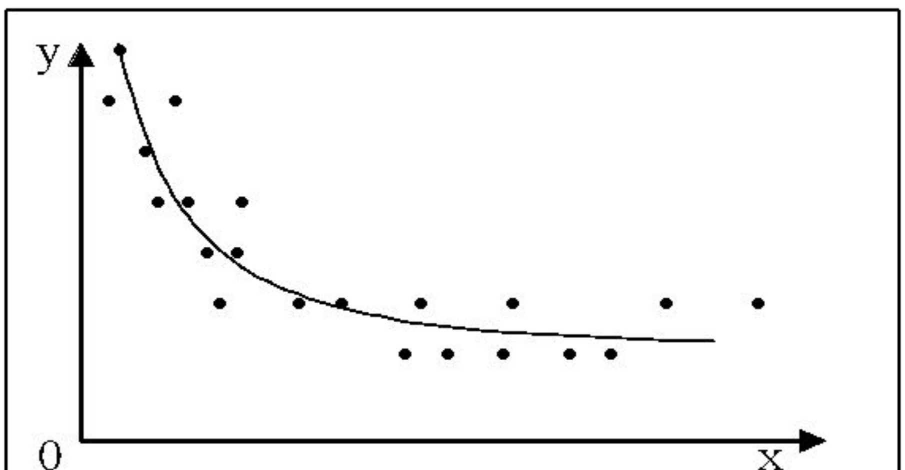
При стат. связи точки фактических наблюдений группируются возле некоторой линии или кривой, называемыми **линиями регрессии**, а описывающие их аналитические выражения – **уравнениями регрессии**.



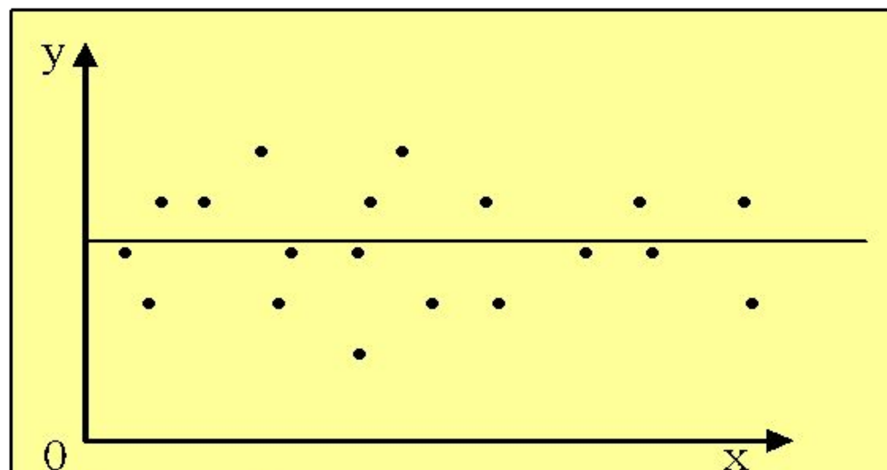
Линейная связь



Параболическая связь



Гиперболическая связь



Нет связи*

Зная уравнение регрессии, можно для любых значений *X*, подставляя их в уравнение, приближенно оценить значение зависимой переменной *Y*.

Точность такой оценки будет тем выше, чем теснее группируются точки фактических наблюдений относительно линии регрессии, т.е. точность модели регрессии определяется тем, насколько тесной является взаимозависимость признаков *X* и *Y*.

При построении парной регрессии (с одной факторной переменной) обычно используются следующие функции:

линейная $\bar{y}_x = a_0 + a_1 x$

степенная $\bar{y}_x = a_0 x^{a_1}$

показательная $\bar{y}_x = a_0 \cdot a_1^{\;x}$

параболическая $\bar{y}_x = a_0 + a_1 x \; + a_2 x^2$

гиперболическая $\bar{y}_x = a_0 + \dfrac{a_1}{x_1}$

логарифмическая $\bar{y}_x = a_0 + a_1 \lg x$

Построение парного линейного уравнения

Если имеется только один факторный признак, строится парная регрессия, выражающаяся уравнением ***прямой:***

$$\boxed{\overline{y}_x = a_0 + a_1 x}$$

Коэффициент регрессии *a1* показывает, на какую величину в среднем изменится результативный признак *Y*, если переменную *X* увеличить на единицу ее собственного измерения.

Свободный член уравнения *a**0*** характеризует усредненное влияние неучтенных в модели факторов (определяет начальные условия развития).

Параметры уравнения получают путем решения следующей системы нормальных уравнений:

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

Рассчитанные по этому уравнению значения $\overline{y}_x$ называются *теоретическими (выравненными) значениями* ***y***.

№ п/п	Баланс. прибыль y	Мат. затраты x	xy	x^2	$\overline{y}_x$
1	15,3	7,9	120,87	62,41	15,3
2	15,2	7,8	118,56	60,84	15,2
3	14,5	7,1	102,95	50,41	13,8
4	14,1	7,3	102,93	53,29	14,2
5	13,8	6,9	95,22	47,61	13,4
6	10,2	5,9	60,18	34,81	11,6
7	7,8	3,8	29,64	14,44	7,6
8	7,4	3,4	25,16	11,56	6,8
9	7,3	3,2	23,36	10,24	6,4
10	5,5	2,7	14,85	7,29	5,5
11	4,8	2,9	13,92	8,41	5,9
12	4,4	2,1	9,24	4,41	4,4
13	3,7	2,0	7,4	4	4,2
14	2,1	0,9	1,89	0,81	2,1
15	2,0	0,8	1,6	0,64	1,9
16	1,8	0,6	1,08	0,36	1,5
17	1,6	0,4	0,64	0,16	1,1
18	1,4	0,3	0,42	0,09	1,0
19	0,8	0,4	0,32	0,16	1,1
20	0,7	0,5	0,35	0,25	1,3

Решим систему уравнений:

$$\begin{cases} 20a_0 + 66{,}9a_1 = 134{,}4 \\ 66{,}9a_0 + 372{,}19a_1 = 730{,}58 \end{cases}$$

Получим параметры уравнения прямой:
a0 = 0,386 и *a1* =1,893.

Таким образом, регрессионная модель имеет вид:

$$\overline{y}_x = 0{,}386 + 1{,}893x$$