

***ПЛАНИРОВАНИЕ
ЭКСПЕРИМЕНТА ДЛЯ
ПРИМЕНЕНИЯ
КОРРЕЛЯЦИОННОГО
АНАЛИЗА***
Лекция 3

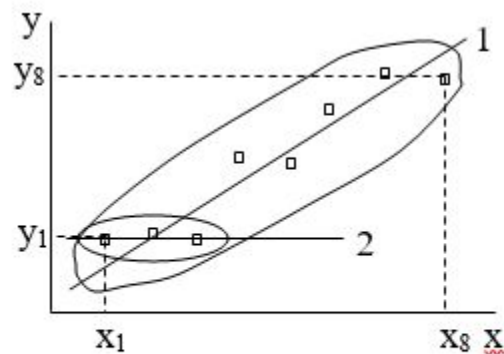
Корреляционный анализ - один из наиболее простых методов математической статистики, позволяющий качественно предсказывать изменения y при изменяющихся значениях x_i (устанавливать связь между этими случайными величинами).

Если каждому значению x_i соответствует всегда строго определенное значение y , то считают, что между этими величинами существует **функциональная связь**, то есть зависимость ϕ ($y = \phi(x_1, x_2, \dots, x_i, \dots, x_k) + \varepsilon$) является **функциональной**.

При наличии функциональной зависимости и знании о ней можно точно предсказывать величину y , задавая конкретное значение x_i .

В большинстве случаев, задавая конкретное значение x_i , можно предсказать лишь тенденцию изменения y .

Эта тенденция обнаруживается лишь при достаточно большом числе m_j различных значений (уровней) изменяемого фактора x_i , а при малых величинах m_j данная тенденция может не наблюдаться (рис.)



Влияние числа значений x (m) на тенденцию изменения y :

1 - тенденция изменения y при $m = 8$; 2 - тенденция изменения y при $m = 3$

Корреляционная связь имеет два крайних предельных случая:
функциональная связь (самая тесная зависимость y от x_i) и **полное отсутствие связи** (влияния x_i на y).

Наличие между y и x_i корреляционной или функциональной связи устанавливается только в результате проведения **корреляционного анализа**.

При корреляционном анализе отражают следующие выводы в форме слов:

- **наличие зависимости** между y и x_j ("есть" или "нет" и др.);

- **характер зависимости** ("функциональная" или "корреляционная") и ее тип ("линейная", "нелинейная", "экспоненциальная", "параболическая", "синусоидальная" и др.);

- **знак связи**: "положительная" - если с увеличением величины значений x_j растет величина y ;
"отрицательная" - если с уменьшением величины значений x_j снижается величина y ;

- **теснота (сила)** корреляционной связи ("очень тесная", "тесная", "не очень тесная", "ярко выраженная", "выраженная", "слабо выраженная" и др.).

Анализ поля корреляции (визуальный анализ)

Поле корреляции называют рисунок (график), выполненный на плоскости в системе двух прямоугольных координат y и x , на котором приведены точки с координатами y_v и x_v (V - номер уровня фактора x от 1 до m).

Анализ поля корреляции проводится визуально.

Анализ выборочного коэффициента корреляции

Корреляция между двумя случайными величинами (y и x)

Присвоим каждой точке на поле корреляции свой номер i (такой же номер будет и у взаимосвязанной пары координат этой точки). Обозначим через N общее число точек с координатами y_i и x_i (количество парных наблюдений в выборке).

Тогда выборочный коэффициент парной корреляции можно рассчитать по формуле

$$r_{yx} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{(N - 1)S_x S_y}$$

Выборочный коэффициент парной корреляции имеет следующие свойства:

$$|r_{yx}| \leq 1$$

Величина r_{yx} не изменяется при изменении начала отсчета величин, а также масштаба координатных осей y и x .

В величине r_{yx} одновременно заложена доля случайности и нелинейности связи между y и x .

По величине и знаку r_{yx} можно сделать большинство выводов корреляционного анализа

Выводы корреляционного анализа в зависимости от значения r_{yx}

<i>Выводы корреляционного анализа</i>	<i>Значения r_{yx}</i>
Наличие зависимости между y_i и x_z : Есть Нет	$0 < r_{yx} \leq 1$ или $r_{yx} = 0$ (при наличии доказательств анализом поля корреляции); $r_{yx} = 1$ (при наличии доказательств анализом поля корреляции)
Характер и тип зависимости: "функциональная линейная" "корреляционная линейная"	$ r_{yx} = 1$ $0 < r_{yx} < 1$
Знак связи: "положительный" "отрицательный"	$r_{yx} > 0$ $r_{yx} < 0$
Теснота (сила) линейной корреляционной связи	Определяется близостью к единице модуля r_{yx}

Оценка тесноты линейной связи (шкала Чаддока)

Значение $ r_{yx} $	0-0,1	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-0,99	1
Наличие функциональной линейной связи	нет	слабая	умеренная	заметная	высокая	очень высокая	полная

Составление планов эксперимента с учетом возможности проведения корреляционного анализа

Корреляционный анализ не предъявляет повышенные требования к планированию эксперимента. Обязательным единственным условием является выполнение соотношения $m_j > 2$.

Для проведения корреляционного анализа желательно, чтобы план эксперимента предусматривал:

- 1) широкую область изменения значений факторов x_j ;
- 2) большое число m_j значений (уровней) факторов x_j , при этом разница между уровнями должна быть больше абсолютной погрешности их измерения;
- 3) повторные опыты для каждого значения факторов x_j ;
- 4) большое общее число измерений (N).