

Статистическое изучение взаимосвязи.

- 1) Виды, формы связей в статистике.
- 2) Причинность, регрессия, корреляция.
- 3) Результативный и качественный признаки.

Основная задача статистики – обнаружить связь между явлениями, её вид и дать количественную характеристику этой связи. Наиболее же общая задача – это прогнозирование и регулирование социально-экономических явлений на основе полученных представлений о связях между явлениями.

1. Виды, формы связей в статистике.

Для описания причинно-следственной связи между явлениями и процессами используется деление статистических признаков, отражающих отдельные стороны взаимосвязанных явлений, на **факторные и результативные**.

Факторные признаки - те, которые влияют на изменение других процессов.

Результативными являются признаки, изменяющиеся под воздействием факторных.

В статистике связи классифицируются по степени их тесноты:

- 1) Функциональная связь - связь, при которой определённому значению факторного признака соответствует одно и только одно значение результативного (например, зависимость площади круга от радиуса: $S=\pi \cdot r^2$).
- 2) Статистическая связь – когда одному и тому же значению факторного признака может соответствовать сколько угодно различных значений результативного признака. Частный случай – корреляционная связь, когда с изменением значения факторного признака закономерно изменяется среднее значение результативного признака.

По направлению корреляционные связи делятся на:

- 1) прямую связь** – такая связь, при которой с изменением значений факторного признака в одну сторону, в ту же сторону меняется и результативный признак.
- 2) обратную связь** – такая связь, при которой с увеличением (уменьшением) факторного признака происходит уменьшение (увеличение) результативного признака.

По форме связи бывают:

- 1) линейные (прямолинейные)
- 2) нелинейные (криволинейные) связи.

Линейные связи выражаются уравнением прямой, а нелинейные – уравнением параболы, гиперболы, степенной и т. п.

По количеству взаимодействующих факторов связи делятся на парную (однофакторную) и множественную (многофакторную) связи. При парной связи на результативный признак действует один факторный, при множественной – несколько факторных признаков.

2. Причинность, регрессия, корреляция.

Исследование существующих связей между явлениями - важнейшая задача общей теории статистики. В процессе статистического исследования зависимостей вскрываются причинно-следственные отношения между явлениями, что позволяет выявлять факторы (признаки), оказывающие существенное влияние на вариацию изучаемых явлений и процессов.

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. Причина всегда должна предшествовать следствию, однако не каждое предшествующее событие следует считать причиной, а последующее - следствием.

Особенностью причинно-следственных связей в социально-экономических явлениях является их транзитивность, т.е. причина X и следствие Y связаны соотношением:

$$X \Rightarrow X' \Rightarrow X'' \Rightarrow Y$$

Например, при использовании показателей международной методологии расчетов фактором валовой прибыли (Y) считается валовое накопление основных и оборотных фондов (X), но при этом допускаются такие факторы, как валовой выпуск (X'), оплата труда (X'') и т. д.



Для количественного описания взаимосвязей между экономическими переменными в статистике используют методы регрессии и корреляции.

Регрессия - статистическая зависимость среднего значения случайной величины (y) от значений другой случайной величины (x) или нескольких случайных величин; введена Фрэнсисом Гальтоном (1886).

При регрессионной зависимости одному и тому же значению (x) могут соответствовать различные значения величины (y). Если при каждом

значении $x = x_i$ наблюдается n_i значений y_{i1}, \dots, y_{in_i} , то зависимость среднего арифметического

$$\langle y \rangle = (y_{i1} + \dots + y_{in_i}) / n_i$$

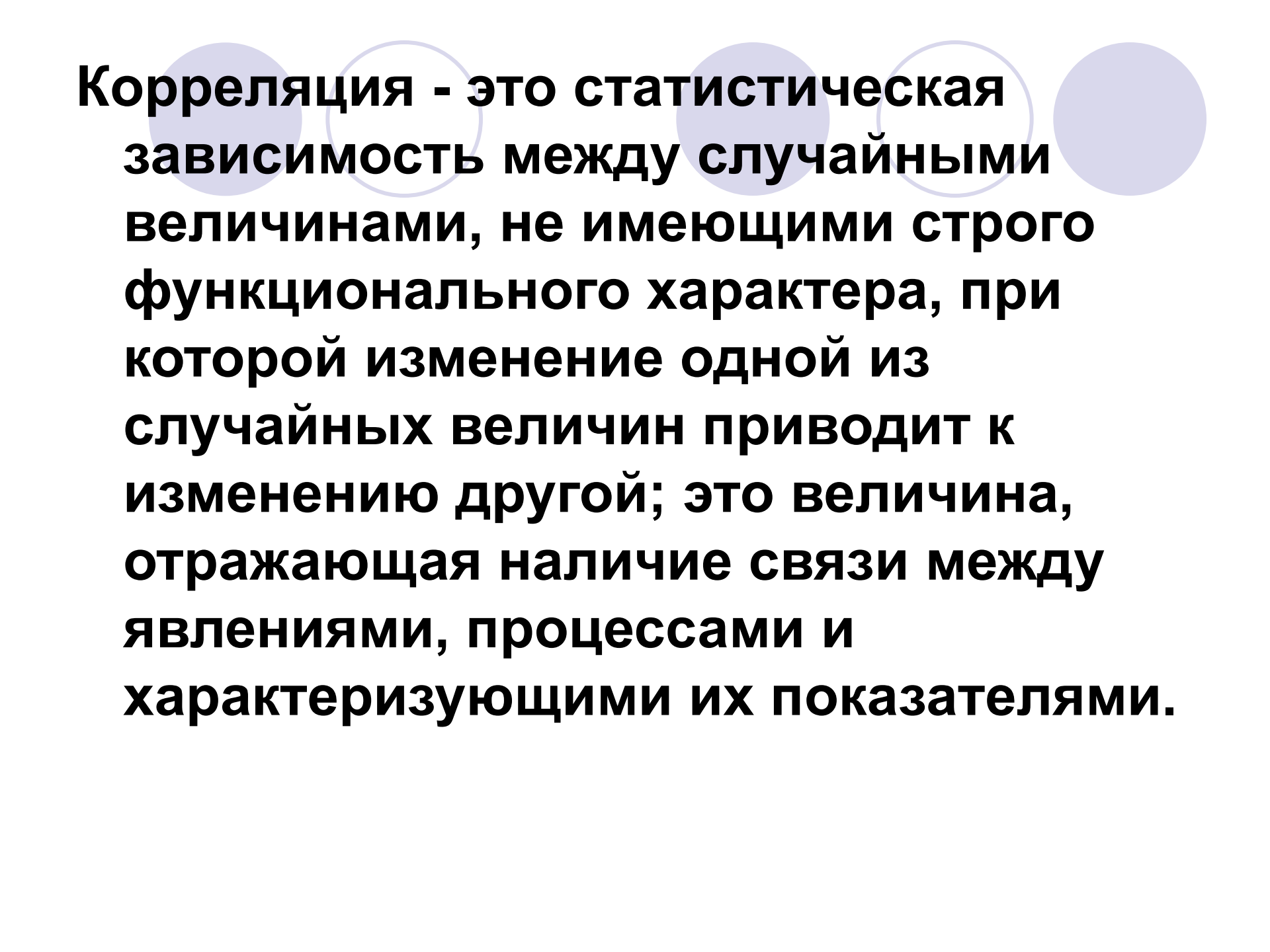
от $x_i, i = 1, \dots, N$ и является средней регрессией.

По направлению связи различают:

- 1) прямую регрессию (положительную), возникающую при условии, если с увеличением или уменьшением независимой величины значения зависимой также соответственно увеличиваются или уменьшаются;**
- 2) обратную (отрицательную) регрессию, появляющуюся при условии, что с увеличением или уменьшением независимой величины зависимая соответственно уменьшается или увеличивается.**



Классическим примером средней регрессии служит зависимость среднего роста детей от роста родителей, а также зависимость средних диаметров деревьев от их высот, зависимость среднего роста человека от его веса и т.п.



Корреляция - это статистическая зависимость между случайными величинами, не имеющими строго функционального характера, при которой изменение одной из случайных величин приводит к изменению другой; это величина, отражающая наличие связи между явлениями, процессами и характеризующими их показателями.

В статистике различаются следующие варианты зависимостей:

- 1) парная корреляция - связь между двумя признаками (результативным и факторным или двумя факторными);**
- 2) частная корреляция - зависимость между результативным и одним факторным признаками при фиксированном значении других факторных признаков;**
- 3) множественная корреляция - зависимость результативного и двух или более факторных признаков, включенных в исследование.**

Показатели корреляции.

- 1) Ковариация. Важной характеристикой совместного распределения двух случайных величин является ковариация (или корреляционный момент). Ковариация является совместным центральным моментом второго порядка.

$$\text{COV}_{XY} = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

где E – математическое ожидание (среднее значение случайной величины).

2) Линейный коэффициент корреляции (коэффициент корреляции Пирсона).

$$r_{XY} = \frac{\text{COV}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}$$

где $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$, $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$ - среднее значение
выборки.

3) Коэффициент Спирмена - метод, который используется с целью статистического изучения связи между явлениями.

Практический расчет коэффициента ранговой корреляции Спирмена включает следующие этапы:

- 1. Сопоставить каждому из признаков их порядковый номер (ранг) по возрастанию (или убыванию).**
- 2. Определить разности рангов каждой пары сопоставляемых значений.**
- 3. Возвести в квадрат каждую разность и суммировать полученные результаты.**

4. Вычислить коэффициент корреляции рангов по формуле:

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

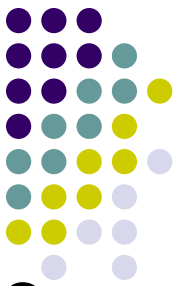
где $\sum d^2$ - сумма квадратов разностей рангов, а n - число парных наблюдений.

3. Результативный и качественный признаки.



Признак – это показатель, характеризующий некоторое свойство объекта совокупности, рассматриваемый, как случайная величина; это качественная особенность единицы совокупности.

Качественный признак относится к признакам, которые определяются характером выражения.

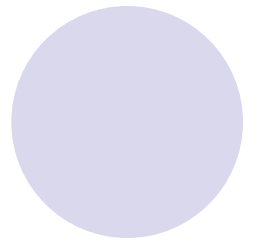
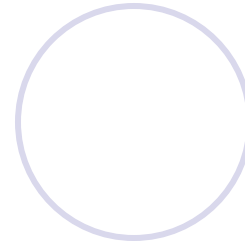
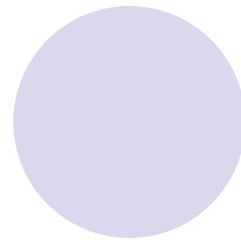
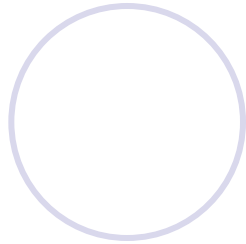
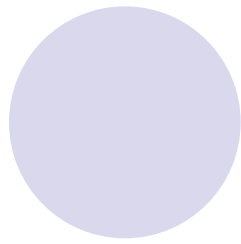


Качественным (атрибутивным) называют такой признак, отдельные значения которого выражаются в виде понятий, наименований (профессия, национальность, вид, свойство и т.д.). Они не имеют числового выражения и представляют собой смысловые понятия.

Признаки, характеризующие статистическую совокупность, взаимосвязаны между собой, поэтому следует различать факторные (признаки-факторы) и результативные признаки.



- ***Факторные признаки*** – это независимые признаки, оказывающие влияние на другие, связанные с ними признаки (например, квалификация, стаж работы и т.д.)
- ***Результативные признаки (признаки следствия)*** – это зависимые признаки, которые изменяются под влиянием факторных признаков, т.е. признаки, которые зависят от вариации других признаков (например, стоимостной объём выпуска продукции является результативным признаком, величина которого зависит от факторных признаков - численности работников и производительности труда).



Спасибо за внимание!)