



Кафедра Маркетинга
и менеджмента (ММ)

СТАТИСТИКА I (теория статистики)

14-18

Статистическое изучение взаимосвязей явлений

По характеру зависимости между изменением признака:

- функциональные;
- корреляционные.

По направлению влияния признака – фактора на признак – результат:

- прямые;
- обратные.

По форме выражения связи:

- прямолинейные (уравнение прямой);
- криволинейные (уравнение кривой).

По степени тесноты связи:

- тесные;
- слабые.

По числу факторов:

- однофакторные;
- многофакторные.

В установлении взаимосвязи между признаками общественных явлений определяющая роль принадлежит теоретическому анализу. Статистика использует много приемов установления и измерения связей. Среди которых:

- метод группировок;
- метод параллельных рядов;
- индексный метод;
- балансовый метод;
- корреляционный метод.

Применение метода группировок для изучения взаимосвязей осуществляется в три этапа:

- совокупность распределяется на группы по значению факторного признака;
- в каждой группе определяется среднее значение результативного фактора;
- рассматривается, как изменяется эта средняя от группы к группе.

Метод наименьших квадратов (МНК) - метод оценки параметров модели на основании экспериментальных данных, содержащих случайные ошибки.

В основе метода лежат следующие рассуждения:

- при замене точного (неизвестного) параметра модели приблизительным значением необходимо минимизировать разницу между экспериментальными данными и теоретическими (вычисленными при помощи предложенной модели). Это позволяет рассчитать параметры модели с помощью МНК с минимальной погрешностью.

Мерой разницы в методе наименьших квадратов служит сумма квадратов отклонений действительных (экспериментальных) значений от теоретических.

Выбираются такие значения параметров модели, при которых сумма квадратов разностей будет наименьшей:

$$\sum_i (Y_i - y_i)^2 = \min$$

- Y – теоретическое значение измеряемой величины,
- y – экспериментальное.

Дисперсия - мера рассеивания случайных величин, измеряемая квадратом отклонения от среднего значения

Остаточная дисперсия (дисперсия остаточных величин):

$$\sigma_{\text{ост}}^2 = \frac{\sum \sigma_i^2 * f_i}{\sum f_i - m} \quad \text{или} \quad \sigma_{y-\hat{y}_x}^2 = \frac{\sum (y - \bar{y}_x)^2}{n}$$

- σ_i^2 - групповые дисперсии;
- $\sum f_i$ - общее число единиц наблюдения;
- n - численность выборки.

Ранговые коэффициенты корреляции

- Коэффициент корреляции рангов Спирмена основан на рассмотрении разности рангов значений факторного и результативного признаков:

$$\rho = 1 - \frac{6 \times \sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

- d^2 - квадраты разности рангов
- n - число наблюдений (число пар рангов)
- 6 - число «шесть» (даже в некоторых учебниках это число ошибочно заменено символом стандартного отклонения).

Коэффициент Спирмена принимает любые значения в интервале $[-1;1]$

- **Ранговый коэффициент корреляции Кендалла** также может использоваться для измерения взаимосвязи между качественными и количественными признаками, характеризующими однородные объекты и ранжированные по одному принципу.

$$\tau = \frac{2S}{n(n-1)}$$

- **n** - число наблюдений;
- **S** - сумма разностей между числом последовательностей и числом инверсий по второму признаку.

Линейный коэффициент корреляции (r) – служит для определения степени тесноты парной линейной зависимости.

- Для расчета можно использовать, например, следующие формулы:

$$r_{xy} = \frac{\sum_{i=1}^n [(x_i - \bar{x}) \cdot (y_i - \bar{y})]}{(n-1) \cdot \sigma_x \cdot \sigma_y}$$

- x_i и y_i - сравниваемые количественные признаки
- n – число сравниваемых наблюдений

$$r_{xy} = \frac{n \sum (x_i \cdot y_i) - \sum x_i \cdot \sum y_i}{\sqrt{(n \sum x_i^2 - (\sum x_i)^2) \cdot (n \sum y_i^2 - (\sum y_i)^2)}}$$

- σ_x и σ_y – стандартные отклонения в сопоставляемых рядах.

Ряды динамики — это ряды статистических показателей, характеризующих развитие явлений природы и общества во времени

Ряды динамики содержат два вида показателей.

- **Показатели времени** (годы, кварталы, месяцы и др.) или моменты времени (на начало года, на начало каждого месяца и т.п.).
- **Показатели уровней ряда.** Показатели уровней рядов динамики могут быть выражены абсолютными величинами (производство продукта в тоннах или рублях), относительными величинами (удельный вес городского населения в %) и средними величинами (средняя заработная плата работников отрасли по годам и т. п.).