

Измерение вариаций



Вариация - колеблемость,
изменяемость величины признака у
единиц совокупности.

Вариационными называют ряды
распределения, построенные по
количественному признаку.

Задачи статистического изучения вариации:

- 1) изучение характера и степени вариации признаков у отдельных единиц совокупности;
- 2) определение роли отдельных факторов или их групп в вариации тех или иных признаков совокупности.

Различают вариацию в пространстве и вариацию во времени.

Под вариацией в пространстве понимают колеблемость значений признака у единиц совокупности, представляющих отдельные территории.

Под вариацией во времени подразумевают, изменение значений признака в различные периоды времени.

Частота - относительный показатель частоты, который может быть выражен в долях единицы или процентах и позволяет сопоставлять вариационные ряды с различным числом наблюдений.

$$f_{\text{отн}} = \frac{f_{\text{абс}}}{\sum_{i=1}^n f_{\text{абс}}}$$

Размах вариации (R) представляет собой разность между максимальным и минимальным значениями признака в изучаемой совокупности ($R = X_{\max} - X_{\min}$).

Среднее линейное отклонение представляет собой среднюю арифметическую из абсолютных значений отклонений отдельных вариантов от их средней арифметической

$$\bar{x}_L = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

$$\bar{x}_L = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{\sum x_i}$$

Дисперсия - средняя из квадратов отклонений вариантов значений признака от их средней величины

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Для не сгруппированных данных:

$$\sigma^2 = (x_n - \bar{x})^2$$

Для сгруппированных данных:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 n_i}{\sum n_i}$$

Свойства дисперсии.

1. Дисперсия постоянной величины равна 0.
2. Уменьшение всех значений признака на одну и ту же величину не изменяет величину дисперсии

$$\sigma^2(x - a) = \sigma^2(x)$$

3. Уменьшение всех значений признака в k раз уменьшает дисперсию в k^2 раз

$$\sigma^2(kx) = k^2 \sigma^2(x)$$

4. Средний квадрат отклонений, исчисленный от среднего арифметического, всегда будет меньше среднего квадрата отклонений, исчисляемого от любой другой величины. Величина различия между ними вполне определенная, это квадрат разности между средней и этой условной величиной A

$$\sigma_{\bar{x}}^2 = \sigma^2 + (\bar{x} - A)^2$$

$$\sigma_{\bar{x}}^2 = \sigma^2 + (\bar{x} + A)^2$$

$$\sigma_{\bar{x}}^2 = \frac{\sum (x_i - \bar{x})^2}{\sum x_i} - (\bar{x} - A)^2$$

Средняя величина альтернативного признака

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{1 \times n_1 + 0 \times n_2}{n_1 + n_2}$$

Дисперсия альтернативного признака

$$\begin{aligned} s^2 &= \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} = \frac{(1 - \bar{x})^2 n_1 + (0 - \bar{x})^2 n_2}{n_1 + n_2} = \bar{x}^2 n_1 + \bar{x}^2 n_2 = \\ &= \bar{x} n_1 \bar{x} + \bar{x} n_2 = \bar{x} n = \bar{x} (1 - \bar{x}) \end{aligned}$$

Среднее квадратическое отклонение альтернативного признака

$$\sigma_{\text{альт}} = \sqrt{\overline{\sigma_{\text{альт}}^2}} = \sqrt{\overline{\sigma^2(1 - p)}} = \sqrt{\overline{\sigma^2 p q}}$$

Среднее квадратическое отклонение групповых средних от общей средней

$$\sigma_{\text{групп}}^2 = \frac{\sum (\sigma_{\text{групп}} - \bar{\sigma})^2 \sigma_{\text{групп}}}{\sum \sigma_{\text{групп}}}$$

Дисперсия - средняя из квадратов отклонений вариантов значений признака от их средней величины

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}$$

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}$$

Среднее квадратическое отклонение (σ) представляет собой корень квадратный из дисперсии

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}}$$

Коэффициент осцилляции

$$\sigma_{osc} = \frac{\sigma_{max}}{\sigma_{min}} \times 100\%$$

Относительное линейное отклонение

$$\sigma_{lin} = \frac{\sigma_{max} - \sigma_{min}}{\sigma_{avg}} \times 100\%$$

Коэффициент вариации - наиболее часто применяемый показатель относительной колеблемости, характеризующий однородность совокупности

$$\sigma_v = \frac{\sigma}{\bar{x}} \times 100\%$$

Мода - это наиболее часто встречающееся в совокупности значение признака.

$$x_0 = x_{j_0} + h \frac{x_{j_0} - x_{j_0}}{x_{j_0} - x_{j_0-1} + (x_{j_0} - x_{j_0+1})}$$

Медиана - это значение изучаемого признака, приходящееся на середину ранжированной совокупности

$$x_1 = \frac{x_{j_1} + x_{j_1+1}}{2}, \quad x_1 = 2/ + 1$$

Квартили - это порядковые характеристики, отделяющие четверти ранжированных совокупностей

$$Q_1 = Q_{[n/4]} + h \frac{\frac{1}{4} - Q_{[n/4]-1}}{Q_{[n/4]}}$$

$$Q_3 = Q_{[3n/4]} + h \frac{\frac{3}{4} - Q_{[3n/4]-1}}{Q_{[3n/4]}}$$