

Для сгруппированных данных вводятся следующие понятия:

1) Групповая дисперсия

$$\sigma_i^2 = \frac{\sum_{j=1}^{n_i} (x_j - \bar{x}_i)^2}{n_i},$$

где n_i – кол-во вариантов в i -ой группе (частота);

\bar{x}_i – выборочное среднее i -ой группы;

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j.$$

2) Средняя из групповых дисперсий

$$\bar{\sigma}^2 = \frac{\sum_{i=1}^k \sigma_i^2 * n_i}{n},$$

где k – количество групп;

$$n = \sum_{i=1}^k n_i \text{ – объем выборки;}$$

3) Межгрупповая дисперсия

$$\delta_{\text{межгр}}^2 = \frac{\sum_{i=1}^k (\bar{x}_i - \bar{x})^2 * n_i}{n},$$

где \bar{x} – средняя величина;

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j.$$

4) Общая дисперсия

$$\sigma_{\text{общ}}^2 = \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n}$$

Теорема о сложении дисперсий

$$\sigma_{\text{общ}}^2 = \bar{\sigma}^2 + \delta_{\text{межгр}}^2$$

Эмпирический коэффициент детерминации

$$\eta^2 = \frac{\delta_{\text{межгр}}^2}{\sigma_{\text{общ}}^2}$$

характеризует качество группировки.