

Ошибки выборки

Формирование выборок и подготовка данных

Ошибки выборки



Случайные ошибки уравновешивают друг друга, так как не имеют ярко выраженной тенденции к завышению или занижению показателей, рассчитанных по выборке.

Систематические тенденциозны, направлены в одну сторону и связаны с нарушением отбора объектов в выборку.

Ошибки регистрации связаны с методами отбора элементов в выборку. Их можно попытаться устраниить на этапе отбора данных.

Ошибки репрезентативности устранить невозможно. Они возникают из-за того, что выборка не полностью воспроизводит генеральную совокупность.

Ошибка выборки

N — объём генеральной совокупности

n — объём выборки

$$X_n \in L(\xi)$$

Среднее по выборке

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Среднее по генеральной совокупности

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i$$

Ошибка выборки

$$\varepsilon_{\bar{x}} = |\mu - \bar{x}|$$

Предельная ошибка выборки

Предельная ошибка — максимально возможное расхождение средних значений выборки и генеральной совокупности с заданной вероятностью.

$$P\left\{|\bar{X} - \mu| < \Delta_{\bar{X}}\right\} = p$$

$$\Delta_{\bar{x}} = t \sqrt{\frac{s^2 \left(1 - \frac{n}{N}\right)}{n}}$$

N — объём генеральной совокупности

n — объём выборки

s² — несмешённая дисперсия

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

t — коэффициент доверия Стьюдента

Доверительный интервал

Доверительный интервал — интервал, в который попадает неизвестный параметр с заданной вероятностью.

$$\bar{X} - \Delta_{\bar{x}} < \mu < \bar{X} + \Delta_{\bar{x}}$$