

Measures of location

x_1, x_2, \dots, x_n denotes a data sample of n elements.

The mean

$$\bar{x} = \frac{1}{n} \sum_{x=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Trimmed mean

To get k trimmed mean:

1. Sort a data sample x_1, x_2, \dots, x_n .
2. Remove $\lfloor k \cdot n \rfloor$ numbers from each end of the sorted data sample.
3. Compute a mean of the resulting sample using formula ().

Winsorized mean

To get k winsorized mean:

1. Sort a data sample x_1, x_2, \dots, x_n .
2. Replace $\lfloor k \cdot n \rfloor$ numbers from each end of the sorted data sample with the nearest remaining observation.
3. Compute a mean of the resulting sample using formula ().

The mode

The most frequent value in the data sample.

Quantiles (type 2)

The q -quantile of the data sample is a value x_q that divides a sorted data sample into two parts: The first part contains at least $q \cdot 100\%$ of the data, the second part contains at least $(1 - q) \cdot 100\%$ of the data.

It can be computed using the following formula:

$$n \cdot q = \begin{cases} \text{whole number } c \Rightarrow x_q = \frac{x_{(c)} + x_{(c+1)}}{2} \\ \text{non-whole number} \Rightarrow c = \lceil n \cdot q \rceil \Rightarrow x_q = x_{(c)} \end{cases}$$

Quartiles

First quartile, is 0.25-quantile.

Second quartile, **the median**, is 0.5-quantile.

Third quartile, is 0.75-quantile.

Measures of variability

Variance

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

Standard deviation

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

Range

$$r = \max \{x_1, x_2, \dots, x_n\} - \min \{x_1, x_2, \dots, x_n\}$$

Interquartile range

$$\text{IQR} = x_{0.75} - x_{0.25}$$

Median absolute deviation (MAD)

$$\text{MAD} = \text{median} (|x_i - x_{0.5}|)$$

Measures of shape

Skewness

$$b_1 = \frac{1}{n} \sum_{x=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3$$

Kurtosis

$$b_2 = \frac{1}{n} \sum_{x=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 - 3$$