1. Frequency distributions

- **Population** is the set of all objects under consideration. A **sample** is a subset of the population.
- **Frequency** distribution specifies for every element in a sample the number of times it appears in the sample.
- **Relative frequency distribution** specifies the percentage of each element in a sample.

2. Mean, median, and the mode

- **Mean** The mean of \( x_1, \ldots, x_n \) is

\[
\frac{x_1 + x_2 + \cdots + x_n}{n}.
\]

- **Median** Let \( x_1 \leq x_2 \leq \cdots \leq x_n \) be the elements in sample.
  - If \( n \) is odd then the median is the element in the middle, i.e. \( x_{(n+1)/2} \).
  - If \( n \) is even then the median is the average of the two elements in the middle, i.e. \( \frac{x_{n/2} + x_{n/2+1}}{2} \).

- **Mode** Mode is the element in a sample that appears most frequently.

3. Standard deviation of a sample

Let \( x_1, x_2, \ldots, x_n \) be the elements in a sample.

- Find the mean, \( \bar{x} = \frac{x_1 + \cdots + x_n}{n} \).
- For each \( x_i \) find \( (x_i - \bar{x})^2 \) and then find the variance

\[
var = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1}
\]
• The standard deviation of the sample is

\[ s = \sqrt{\text{var}} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n - 1}}. \]

• Data points in the interval \([\bar{x} - s, \bar{x} + s]\) lie within one standard deviation of the mean, \([\bar{x} - 2s, \bar{x} + 2s]\) lie within two.