104neutrinos

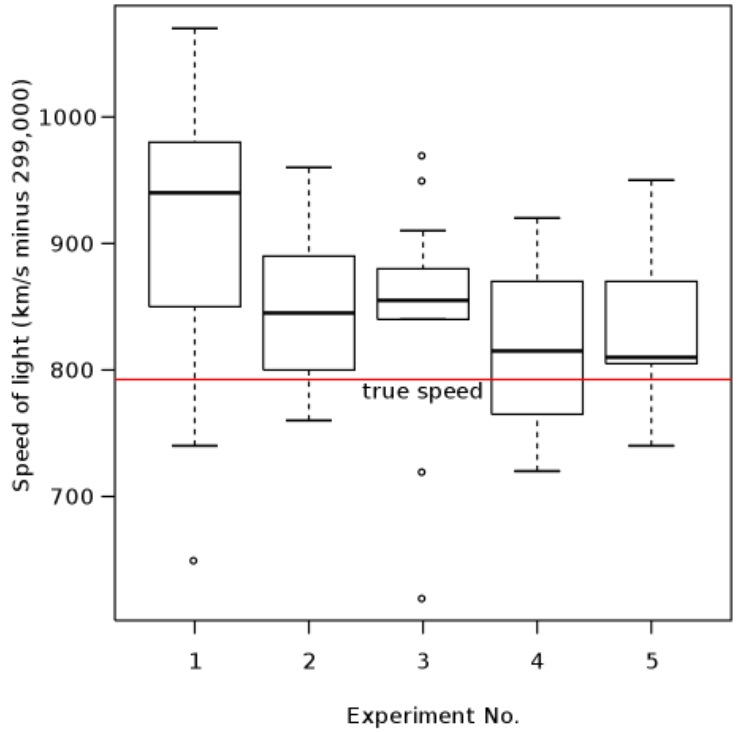
B-MAT-100

Descriptive statistics

- Science that quantitatively describes and analyses features of a collection of data points
- The goal is to summarize a sample, rather than inferring information on a larger population it may represent
- Applications everywhere
  - Industry (reliability, projections…)
  - Social sciences (census, polls…)
  - Economy (market analysis…)

Descriptive statistics

- Given a data set $(x_0, x_1, \ldots, x_n)$, we can describe the observations using measures such as:
  - Central tendency (ex: mean)
  - Dispersion (ex: variance)



Central tendency measures

- Arithmetic mean

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{1}{n}\sum_{i=0}^{n} x_i$$

- Root mean square

$$x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots + x_n^2)} = \sqrt{\frac{1}{n}\sum_{i=0}^{n} x_i^2}$$

- Harmonic mean (if $x_i > 0$)

$$H = \frac{n}{\frac{1}{x_0} + \frac{1}{x_1} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum_{i=0}^{n}\frac{1}{x_i}} = \left(\frac{1}{n}\sum_{i=0}^{n}\frac{1}{x_i}\right)^{-1}$$

Dispersion measures

- Indicate how close are the values to the central tendency
- Variance

$$Var(x) = \frac{1}{n}\sum_{i=0}^{n}(x_i - \bar{x})^2$$
$$= \left(\frac{1}{n}\sum_{i=0}^{n}x_i^2\right) - \bar{x}^2$$
$$= x_{rms}^2 - \bar{x}^2$$

- Standard deviation

$$\sigma = \sqrt{Var(x)}$$

104neutrinos

- Goal: Compute central tendency and dispersion measures for a data set evolving with time
- Inputs: n a h sd
  - n: number of values already in the set
  - a: current arithmetic mean
  - h: current harmonic mean
  - sd: current standard deviation
- When started, the program asks the user to enter a new value

104neutrinos

- If a new value is entered:
  - Display the number of recorded values
  - Display the arithmetic mean
  - Display the root mean square
  - Display the harmonic mean
  - Display the standard deviation
  - Asks the user for another value
- If the keyword ‘END’ is entered, the program stops

Exercise: Arithmetic mean

- Create a function that takes a number of recorded values, an arithmetic mean and a new value as parameters, and returns the updated arithmetic mean.

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{1}{n}\sum_{i=0}^{n} x_i$$

Exercise: Root mean square

- Create a function that takes a number of recorded values, a root mean square, and a new value as parameters, and returns the updated root mean square.

$$x_{rms} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots + x_n^2)} = \sqrt{\frac{1}{n}\sum_{i=0}^{n} x_i^2}$$

Exercise: Harmonic mean

- Create a function that takes a number of recorded values, an harmonic mean, and a new value as parameters, and returns the updated harmonic mean.

$$H = \frac{n}{\frac{1}{x_0} + \frac{1}{x_1} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum_{i=0}^{n} \frac{1}{x_i}} = \left(\frac{1}{n} \sum_{i=0}^{n} \frac{1}{x_i} \right)^{-1}$$

(with $x_i > 0$)

Exercise: Wrapping up

- The root mean square is not passed up as a parameter. Create a function that computes it from the parameters of the 104neutrinos.

- Can you find an efficient way to update the standard deviation ?