DOE
DESIGN OF EXPERIMENT
Introduction to Statistics

S. Tassani
Universitat Pompeu Fabra
e-mail:
	simone.tassani@upf.edu
Framework of lesson 0:

• Introduction to normal distributions
  – Concept of “average/mean”
  – Concept of “Standard Deviation” and “Variance”
Normal Distribution

• A distribution is called Normal when assumes the typical bell shape (Gaussian distribution)

• The center of the bell distribution is its Mean (or Average)

• All the *parametric statistic* is based on the assumption of Normal Distribution
  – A parametric test (like t-test) applied to a not-normal distribution gives wrong results.
Normal Distribution: Example

• The average human being has one testicle and one breast

• The number of testicles and breasts is not normally distributed
  – It is wrong to compute the average.
Normal Distribution: Average and Standard Deviation
Normal Distribution: Average and Standard Deviation

• **Arithmetic mean (or mean or average).** In statistics, the term average refers to any of the measures of central tendency. The arithmetic mean is defined as the sum of the numerical values of each and every observation divided by the total number of observations.

  \[ \mu = \frac{1}{n} \sum_{i=1}^{n} a_i \]
Normal Distribution: Average and Standard Deviation

• The **standard deviation** (SD, also represented by the Greek letter sigma σ or s) is a measure that quantify the amount of variation or dispersion of a set of data values. A standard deviation close to 0 indicates that the data points tend to be very close to the mean of the set.

• It is defined as the root mean square of the variance
Normal Distribution: Variance

- **Variance** measures how far a set of numbers are spread out. A variance of zero indicates that all the values are identical.

\[
\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2
\]

Whole population (n very high)

\[
\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2
\]

Sample taken from a bigger Population

Degrees of Freedom
Example

• The heights (at the shoulders) are: 600mm, 470mm, 170mm, 430mm and 300mm.
• Find out the Mean, the Variance, and the Standard Deviation.
• First step is to find the Mean:
  \[ \text{Mean} = \frac{600 + 470 + 170 + 430 + 300}{5} = \frac{1970}{5} = 394\text{mm} \]
Example

• Now we calculate each dog's difference from the Mean:

• To calculate the **Variance**, take each difference, square it, and then average the result:

\[
\sigma^2 = \frac{206^2 + 76^2 + (-224)^2 + 36^2 + (-94)^2}{5} = 21704 \text{ mm}^2
\]
Example

- And the Standard Deviation is just the square root of Variance:
  - \( \sigma = \sqrt{21704} = 147\, mm \)

- So, using the Standard Deviation we have a "standard" way of knowing what is normal, and what is extra large or extra small
Questions

Thank you