Hypothesis Testing, Probability and Distributions

Pablo E. GutiÃ©rrez-Fonseca

Fall 2024

Normality, Probability and Significance

- Why did we focus on normality?
  - The **normal distribution** is a key tool for determining the probability of a given value occurring in a population that follows this distribution.
- It allows us to make inferences about a population by calculating how likely it is for data to fall within certain ranges.
- Many statistical tests assume data follows a **normal distribution**, which helps in determining **significance** and making reliable conclusions.

University of Vermont

Hypothesis Testing

- All inferential tests use a formula that calculates a **test statistic**, quantifying the relationship or difference you are testing.

- Independent t-test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

  - Dependent t-test

$$t = \frac{\bar{D}}{\frac{s_D}{\sqrt{n}}}$$

  - One sample z-test

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- F-test (ANOVA)

$$F = \frac{MST}{MSE}$$

- Pearson correlation

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

University of Vermont