STATISTICAL METHODS

Master in Industrial Management,

Operations and Sustainability (MIMOS)

2nd year/1st Semester

2025/2026

CONTACT

Professor: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/
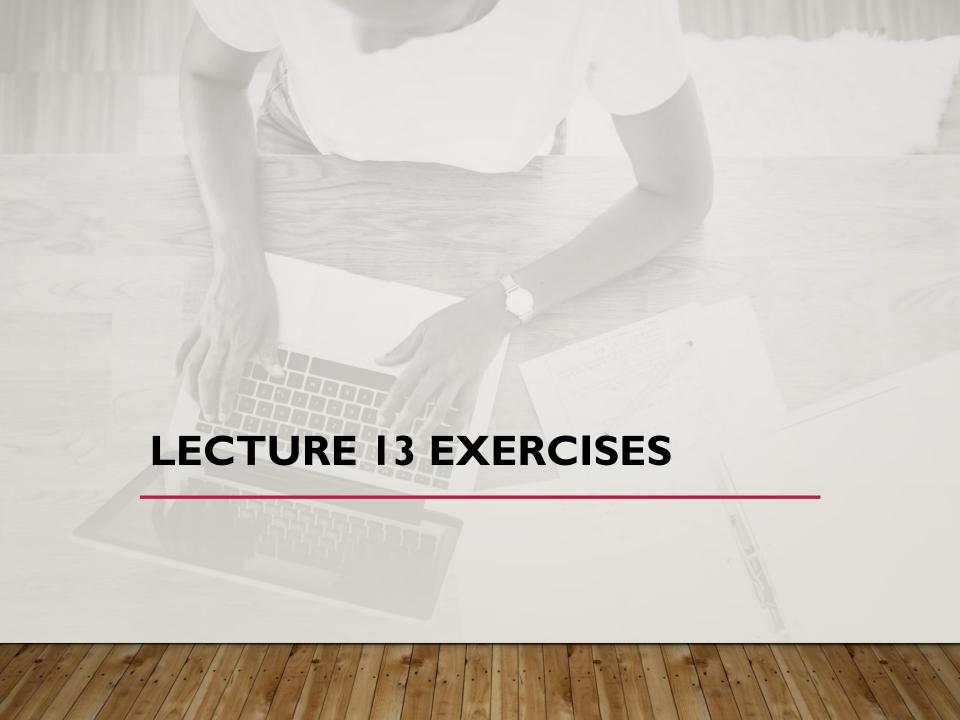
PROGRAM

1 Fundamental Concepts of Statistics

2 Descriptive Data Analysis

3 Introduction to Inferential Analysis

4 Parametric Hypothesis Testing

5 Non-Parametric Hypothesis Testing

6 Linear Regression Analysis

LECTURE 13 EXERCISES

EXERCISE 1

Exercise 1 – Mean vs Median

A dataset contains household incomes in a city. Most values are between 800 and 1200 euros, but a few households earn more than 10,000 euros.

Which measure best represents the "typical" income?

a) Mean
b) Median
c) Range
d) Variance

EXERCISE I: SOLUTION

✅ Answer:

✅ **Exercise 1 – Mean vs Median**

Correct answer: b) Median

Justification:

The mean is strongly affected by extreme values (outliers), such as very high incomes.

The median is more robust and better represents a typical value when outliers are present.

EXERCISE 2

Exercise 2 – Direction of Correlation

A scatter plot shows that as **distance from the city centre increases, house prices tend to decrease.**

Which statement best describes this relationship?

a) Strong positive linear relationship

b) Moderate negative linear relationship

c) No linear relationship

d) Perfect positive linear relationship

EXERCISE 2: SOLUTION

✅ Answer:

✅ **Exercise 2 – Direction of Correlation**

Correct answer: b) Moderate negative linear relationship

Justification:

As one variable increases (distance), the other decreases (price).

This indicates a **negative** relationship, and since it is not perfect, it is **moderate**.

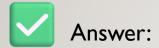# EXERCISE 3

Exercise 3 – Correlation vs Causation

A study finds a positive correlation between **number of cafés in a city** and **number of tourists**.

Which statement is correct?

a) Cafés cause tourism

b) Tourism causes cafés

c) The variables are correlated, but causation is not implied

d) There is no relationship between the variables

EXERCISE 3: SOLUTION

✅ Answer:

✅ **Exercise 3 – Correlation vs Causation**

Correct answer: c) The variables are correlated, but causation is not implied

Justification:

Correlation only measures association.

It does **not** prove that one variable causes the other.

EXERCISE 4

Exercise 4 – Basic Probability (Two Groups)

In a school, **70% of students are undergraduates** and **30% are postgraduate students.**

- The probability that an undergraduate owns a laptop is 0.9
- The probability that a postgraduate owns a laptop is 0.95

What is the probability that a randomly selected student owns a laptop?

a) 0.905
b) 0.915
c) 0.930
d) 0.950

EXERCISE 4: SOLUTION

Answer:

✅ **Exercise 4 – Basic Probability (Law of Total Probability)**

Correct answer: b) 0.915

Justification:

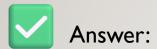$$P(\text{Laptop}) = 0.7(0.9) + 0.3(0.95) = 0.63 + 0.285 = 0.915$$

EXERCISE 5

Exercise 5 – Hypothesis Testing (Intuition)

In hypothesis testing, a **small p-value** indicates:

a) Strong evidence in favour of the null hypothesis

b) Strong evidence against the null hypothesis

c) The null hypothesis is always true

d) The significance level is large

EXERCISE 5: SOLUTION

✅ Answer:

✅ **Exercise 5 – Hypothesis Testing (p-value)**

Correct answer: b) Strong evidence against the null hypothesis

Justification:

A small p-value indicates that the observed result is unlikely under H_0, so it provides evidence **against** the null hypothesis.
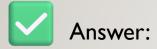
EXERCISE 6

Exercise 6 – Sampling Distribution (Idea)

Which statement about the sampling distribution of the sample mean is correct?

a) It always has the same variance as the population

b) Its standard deviation decreases as sample size increases

c) It depends on the sample mean only

d) It is never approximately normal

EXERCISE 6: SOLUTION

✅ Answer:

✅ **Exercise 6 – Sampling Distribution**

Correct answer: b) Its standard deviation decreases as sample size increases

Justification:

The standard deviation of the sample mean is:

$$\frac{\sigma}{\sqrt{n}}$$

As n increases, the variability decreases.

EXERCISE 7

Exercise 7 – Confidence Interval (Interpretation)

A 95% confidence interval for a population mean is (48, 52).

Which interpretation is correct?

a) 95% of the sample values lie between 48 and 52

b) The probability that the true mean lies in the interval is 0.95

c) We are 95% confident that the true mean lies between 48 and 52

d) The true mean is exactly 50

EXERCISE 7: SOLUTION

✅ Answer:

☑ **Exercise 7 – Confidence Interval Interpretation**

Correct answer: c) We are 95% confident that the true mean lies between 48 and 52

Justification:

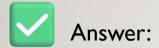This is the correct frequentist interpretation of a confidence interval.

EXERCISE 8

Exercise 8 – Margin of Error (Conceptual)

Which action will **reduce the margin of error** of a confidence interval?

a) Decreasing the sample size

b) Increasing the confidence level

c) Increasing the sample size

d) Increasing the standard deviation

EXERCISE 8: SOLUTION

✅ Answer:

☑ **Exercise 8 – Margin of Error**

Correct answer: c) Increasing the sample size

Justification:

The margin of error is inversely proportional to $\sqrt{n}$.

Larger samples → smaller margin of error.

EXERCISE 9

Exercise 9 – Chi-Square Test (Purpose)

The chi-square goodness-of-fit test is mainly used to:

a) Compare two population means

b) Test if observed frequencies match expected proportions

c) Test correlation between two variables

d) Estimate a confidence interval

EXERCISE 9: SOLUTION

✅ Answer:

☑ **Exercise 9 – Chi-Square Test**

Correct answer: b) Test if observed frequencies match expected proportions

Justification:

The chi-square goodness-of-fit test compares **observed** and **expected** frequencies.

EXERCISE 10

Exercise 10 – t-Test (When to Use)

When should a **t-test for the mean** be used instead of a z-test?

a) When the population variance is known

b) When the sample size is very large

c) When the population variance is unknown and the sample is small

d) When data are categorical

EXERCISE 10: SOLUTION

✅ Answer:

☑ **Exercise 10 – t-Test vs z-Test**

Correct answer: c) When the population variance is unknown and the sample is small

Justification:

The t-test is used when σ is unknown and the sample size is small.

EXERCISE 11

11. Assume that a random variable X follows a Normal distribution with mean $\mu = 20$ and standard deviation $\sigma = 5$. Which of the following statements is correct?

a) The median of X is 15.

b) The mode of X is 25.

c) The median of X is 20.

d) The mean of X is less than the median.

EXERCISE 11: SOLUTION

✅ Answer:

Resolução:

- For a **Normal distribution**, the distribution is **symmetric** around the mean.
- Therefore:

$$\text{Mean} = \text{Median} = \text{Mode} = \mu$$

- Given that $\mu = 20$, the median is also 20.

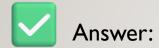**Resposta correta: c) The median of X is 20**

EXERCISE 12

Exercise 12 – Uniform Distribution (Understanding)

If a random variable is uniformly distributed between 0 and 10, which statement is true?

a) Values near 5 are more likely

b) All values between 0 and 10 have the same probability

c) The mean is 10

d) The distribution is skewed

EXERCISE 12: SOLUTION

✅ Answer:

✅ **Exercise 12 – Uniform Distribution**

Correct answer: b) All values between 0 and 10 have the same probability

Justification:

In a uniform distribution, all values in the interval are equally likely.

EXERCISE 13

Exercise 13 – Poisson Distribution (When to Use)

The Poisson distribution is appropriate when modelling:

a) Continuous waiting times

b) The number of events in a fixed interval

c) Proportions in a population

d) Sample means

EXERCISE 13: SOLUTION

✅ Answer:

✅ **Exercise 13 – Poisson Distribution**

Correct answer: b) The number of events in a fixed interval

Justification:

The Poisson distribution models **counts of events** in a fixed time or space interval.

THANKS!

Questions?