STATISTICS II

Bachelor's degrees in Economics, Finance and Management

2nd year/2nd Semester
2025/2026

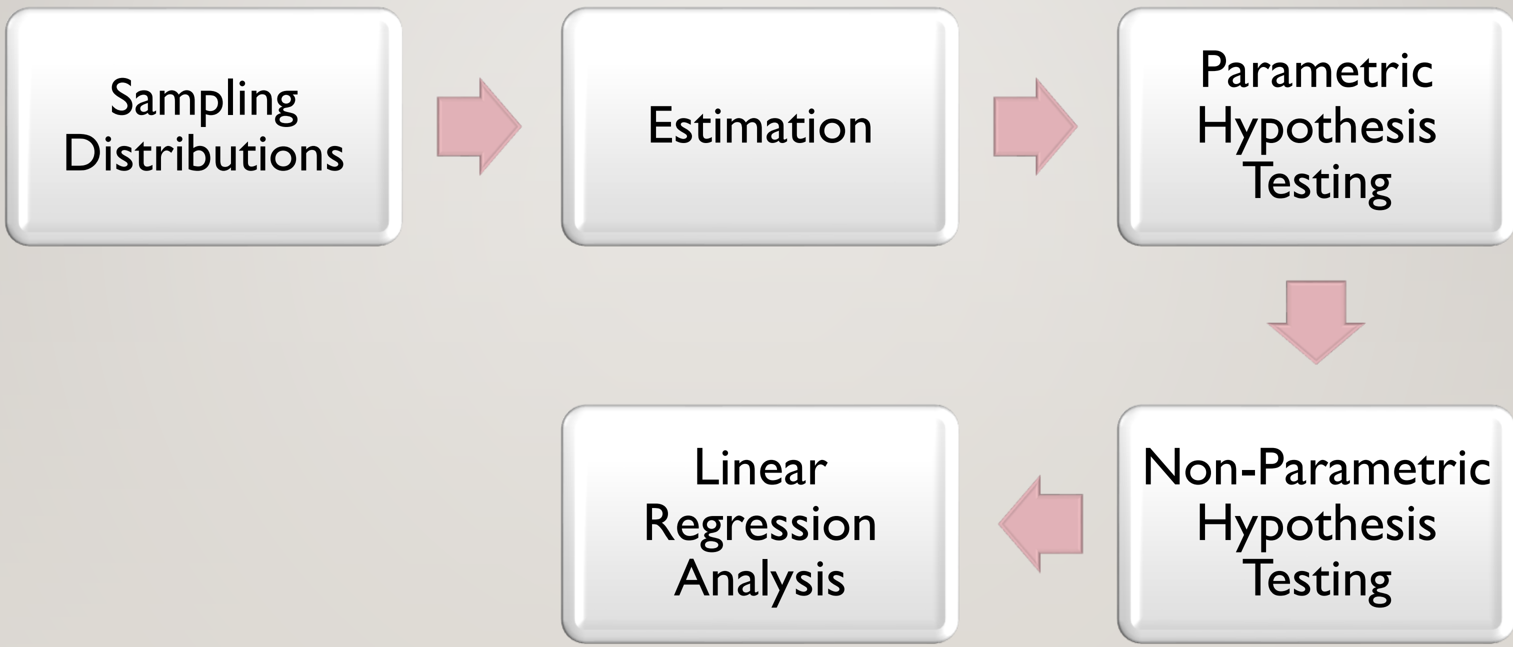
CONTACT

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


Sampling Distributions
Estimation
Parametric Hypothesis Testing
Non-Parametric Hypothesis Testing
Linear Regression Analysis

PRATICAL CLASS 11

Exercises 15.3, 15.9, 14.4, 14.12,14.17,14.20

EXERCISE 15.3

15.3 Given the following analysis of variance table, compute mean squares for between groups and within groups. Compute the *F* ratio and test the hypothesis that the group means are equal.

Source of Variation	Sum of Squares	Degrees of Freedom
Between groups	1,000	2
Within groups	743	15
Total	1,743	17

Newbold et al (2013)

EXERCISE 15.3: SOLUTION

Answer:

One-Way ANOVA

We are given the following analysis of variance table:

Source of Variation	Sum of Squares	Degrees of Freedom
Between groups	1000	2
Within groups	743	15
Total	1743	17

One-Way ANOVA

Step 1: Hypotheses

$$H_0 : \mu_1 = \mu_2 = \mu_3$$
$$H_1 : \text{At least one group mean is different}$$

EXERCISE 15.3: SOLUTION

✅ Answer:

Step 2: Test statistic

The test statistic for a one-way ANOVA is

$$F = \frac{MSB}{MSW}$$

where

$$MSB = \frac{SSB}{df_B}, \qquad MSW = \frac{SSW}{df_W}$$

Compute the mean squares:

$$MSB = \frac{1000}{2} = 500$$

$$MSW = \frac{743}{15} \approx 49.53$$

Thus, the observed value of the test statistic is:

$$F_{\text{obs}} = \frac{500}{49.53} \approx 10.10$$

Under H_0,

$$F \sim F(2, 15)$$

EXERCISE 15.3: SOLUTION

✅ Answer:

Step 3: Decision rule

For a significance level α, reject H_0 if:

$$F_{\text{obs}} > F_{1-\alpha}(2, 15)$$

Step 4: Conclusion

Since

$$F_{\text{obs}} \approx 10.10$$

and this value exceeds the critical value $F_{1-\alpha}(2, 15)$ for standard significance levels (e.g., $\alpha = 0.05$), we **reject** H_0.

Final conclusion

> There is sufficient evidence to conclude that **not all group means are equal**.

EXERCISE 15.9

15.9 Samples of four salespeople from each of four regions were asked to predict percentage increases in sales volume for their territories in the next 12 months. The predictions are shown in the accompanying table.

West	Midwest	South	East
6.8	7.2	4.2	9.0
4.2	6.6	4.8	8.0
5.4	5.8	5.8	7.2
5.0	7.0	4.6	7.6

a. Prepare the analysis of variance table.
b. Test the null hypothesis that the four population mean sales growth predictions are equal.

Newbold et al (2013)

EXERCISE 15.9 A): SOLUTION

✅ Answer:

a) Analysis of Variance Table

Step 1: Sample means

$$\bar{x}_{West} = 5.35$$
$$\bar{x}_{Midwest} = 6.65$$
$$\bar{x}_{South} = 4.85$$
$$\bar{x}_{East} = 7.95$$

Overall mean:

$$\bar{x}. = 6.20$$

Step 2: Sum of Squares

- Between groups:

$$SSB = \sum n(\bar{x}_j - \bar{x}.)^2 = 22.08$$

- Within groups:

$$SSW = 9.26$$

- Total:

$$SST = SSB + SSW = 31.34$$

EXERCISE 15.9 A): SOLUTION

✅ Answer:

ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Between groups	22.08	3	7.36
Within groups	9.26	12	0.77
Total	31.34	15	—

EXERCISE 15.9 B): SOLUTION

✅ Answer:

b) Hypothesis Test

Step 1: Hypotheses

$$H_0 : \mu_{West} = \mu_{Midwest} = \mu_{South} = \mu_{East}$$
$$H_1 : \text{At least one mean is different}$$

Step 2: Test statistic

$$F = \frac{MSB}{MSW} = \frac{7.36}{0.77} \approx 9.56$$

Under H_0,

$$F \sim F(3, 12)$$

EXERCISE 15.9 B): SOLUTION

✅ Answer:

Step 3: Decision rule

At significance level α, reject H_0 if

$$F_{\text{obs}} > F_{1-\alpha}(3, 12)$$

Step 4: Conclusion

Since

$$F_{\text{obs}} \approx 9.56$$

is greater than the critical value $F_{1-\alpha}(3, 12)$ for common significance levels (e.g., $\alpha = 0.05$), we **reject the null hypothesis**.

Final Conclusion

> There is sufficient statistical evidence to conclude that **the mean sales growth predictions are not the same across the four regions**.

EXERCISE 14.4

14.4 Production records indicate that in normal operation for a certain electronic component, 93% have no faults, 5% have one fault, and 2% have more than one fault. For a random sample of 500 of these components from a week's output, 458 were found to have no faults; 30, to have one fault; and 12, to have more than one fault. Test, at the 5% level, the null hypothesis that the quality of the output from this week conforms to the usual pattern.

Newbold et al (2013)

EXERCISE 14.4: SOLUTION

✅ Answer:

Chi-Square Goodness-of-Fit Test

Production records indicate that, under normal operation, the distribution of faults in an electronic component is:

- No faults: 93%
- One fault: 5%
- More than one fault: 2%

A random sample of **500 components** from a week's output produced the following results:

Number of faults	Observed frequency
No faults	458
One fault	30
More than one fault	12
Total	500

We test whether the observed data are consistent with the historical fault distribution.

EXERCISE 14.4: SOLUTION

✅ Answer:

Step 1: Hypotheses

$$H_0 : \text{The distribution of faults follows the stated percentages}$$
$$H_1 : \text{The distribution of faults does not follow the stated percentages}$$

Step 2: Expected frequencies

Expected frequencies are obtained by multiplying the total sample size by the given proportions.

Number of faults	Expected frequency
No faults	$0.93 \times 500 = 465$
One fault	$0.05 \times 500 = 25$
More than one fault	$0.02 \times 500 = 10$

EXERCISE 14.4: SOLUTION

✅ Answer:

Step 3: Test statistic

The chi-square test statistic is

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Compute each contribution:

$$\frac{(458 - 465)^2}{465} = \frac{49}{465} \approx 0.105$$
$$\frac{(30 - 25)^2}{25} = \frac{25}{25} = 1.000$$
$$\frac{(12 - 10)^2}{10} = \frac{4}{10} = 0.400$$

$$\chi^2_{\text{obs}} \approx 0.105 + 1.000 + 0.400 = 1.505$$

Step 4: Degrees of freedom

$$df = k - 1 = 3 - 1 = 2$$

Under H_0,

$$\chi^2 \sim \chi^2(2)$$

EXERCISE 14.4: SOLUTION

✅ Answer:

Step 5: Decision rule

At the 5% significance level ($\alpha = 0.05$), reject H_0 if

$$\chi^2_{\text{obs}} > \chi^2_{0.95}(2) = 5.99$$

Step 6: Conclusion

Since

$$\chi^2_{\text{obs}} \approx 1.51 < 5.99,$$

we **do not reject the null hypothesis**.

Final Conclusion

> At the 5% significance level, there is **no statistical evidence** to suggest that the quality of the electronic components differs from the normal operating standard.

EXERCISE 14.12

14.12 In a period of 100 minutes there were a total of 190 arrivals at a highway toll booth. The accompanying table shows the frequency of arrivals per minute over this period. Test the null hypothesis that the population distribution is Poisson.

Number of arrivals in minutes	0	1	2	3	4 or more
Observed frequency	10	26	35	24	5

Newbold et al (2013)

EXERCISE 14.12: SOLUTION

Answer:

Step 1: Hypotheses

- **H_0**: The number of arrivals per minute follows a Poisson distribution
- **H_1**: It does not follow a Poisson distribution

Step 2: Estimate the Poisson parameter

Total arrivals = 190

Total time = 100 minutes

$$\hat{\lambda} = \frac{190}{100} = 1.9$$

So we test:

$$X \sim \text{Poisson}(1.9)$$

EXERCISE 14.12: SOLUTION

Answer:

Step 3: Compute expected probabilities

We need:

$$P(X = k) = e^{-1.9}\frac{1.9^k}{k!}$$

First compute (approx):

$$e^{-1.9} \approx 0.1496$$

Probabilities

- $P(0) = 0.1496$
- $P(1) = 0.1496 \cdot 1.9 = 0.2842$
- $P(2) = 0.1496 \cdot \frac{1.9^2}{2} = 0.2700$
- $P(3) = 0.1496 \cdot \frac{1.9^3}{6} = 0.1710$

Now:

$$P(\geq 4) = 1 - (P0 + P1 + P2 + P3)$$

$$P(\geq 4) = 1 - (0.1496 + 0.2842 + 0.2700 + 0.1710) = 0.1252$$

EXERCISE 14.12: SOLUTION

✅ Answer:

✔ Step 4: Expected frequencies (n = 100)

Category	Observed (O)	Probability	Expected (E)
0	10	0.1496	14.96
1	26	0.2842	28.42
2	35	0.2700	27.00
3	24	0.1710	17.10
≥4	5	0.1252	12.52

✔ Step 5: Chi-square statistic

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Compute each term:

- 0: $\frac{(10-14.96)^2}{14.96} = 1.64$
- 1: $\frac{(26-28.42)^2}{28.42} = 0.21$
- 2: $\frac{(35-27)^2}{27} = 2.37$
- 3: $\frac{(24-17.10)^2}{17.10} = 2.78$
- ≥4: $\frac{(5-12.52)^2}{12.52} = 4.52$

Total:

$$\chi^2 \approx 11.52$$

EXERCISE 14.12: SOLUTION

Answer:

Step 6: Degrees of freedom

Number of classes = 5

Estimated parameters = 1 (λ)

$$df = 5 - 1 - 1 = 3$$

Step 7: Critical value

At **10% level**:

$$\chi^2_{0.90,3} = 6.251$$

Step 8: Decision

- Test statistic = 11.52
- Critical value = 6.251

$$11.52 > 6.251$$

👉 **Reject H_0**

Final conclusion (exam style)

At the 10% significance level, there is sufficient evidence to conclude that the number of arrivals per minute does not follow a Poisson distribution.

EXERCISE 14.17

14.17 The U.S. Department of Agriculture (USDA) Center for Nutrition Policy and Promotion (CNPP) uses the Healthy Eating Index to monitor the diet quality of the U.S. population, particularly how well it conforms to dietary guidance. The HEI–2005 measures how well the population follows the recommendations of the 2005 *Dietary Guidelines for Americans* (Guenther et al. 2007). Data collected on a random sample of individuals who participated in two extended interviews and medical examinations are contained in the data file **HEI Cost Data Variable Subset**, where the first interview is identified by daycode = 1 and data for the second interview are identified by daycode = 2. One variable in the HEI–2005 study is a participant's activity level, coded as 1 = sedentary, 2 = active, and 3 = very active. In Chapter 1, we constructed bar charts of participants' activity level by gender for data collected on the first interview. Determine if there is an association between activity level and gender.

Activity Level	Male	Female
Sedentary	957	1226
Active	340	417
Very active	842	678

Newbold et al (2013)

EXERCISE 14.17: SOLUTION

✅ Answer:

Chi-Square Test of Independence

We want to determine whether **activity level** is associated with **gender**.

Observed Frequencies

Activity Level	Male	Female	Total
Sedentary	957	1226	2183
Active	340	417	757
Very active	842	678	1520
Total	2139	2321	4460

EXERCISE 14.17: SOLUTION

Answer:

Step 1: Hypotheses

$$H_0 : \text{Activity level and gender are independent}$$
$$H_1 : \text{Activity level and gender are associated}$$

Step 2: Test statistic

The test statistic for a chi-square test of independence is

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O denotes observed frequencies and E denotes expected frequencies.

Expected Frequencies

$$E_{ij} = \frac{(\text{row total})(\text{column total})}{\text{grand total}}$$

Activity Level	Male	Female
Sedentary	1047.1	1135.9
Active	363.1	393.9
Very active	728.8	791.2

EXERCISE 14.17: SOLUTION

Answer:

Chi-Square Contributions

$$\chi^2 \approx \frac{(957 - 1047.1)^2}{1047.1} + \frac{(1226 - 1135.9)^2}{1135.9} + \cdots$$

Summing all six cells:

$$\chi^2_{\text{obs}} \approx 63.7$$

Step 3: Degrees of freedom

$$df = (r-1)(c-1) = (3-1)(2-1) = 2$$

Under H_0,

$$\chi^2 \sim \chi^2(2)$$

EXERCISE 14.17: SOLUTION

✅ Answer:

Step 4: Decision rule

At significance level α, reject H_0 if

$$\chi^2_{\text{obs}} > \chi^2_{1-\alpha}(2)$$

Step 5: Conclusion

Since

$$\chi^2_{\text{obs}} \approx 63.7$$

is far greater than the critical value $\chi^2_{1-\alpha}(2)$ for common significance levels (e.g., $\alpha = 0.05$), we **reject the null hypothesis**.

Final Conclusion

> There is **strong statistical evidence** of an association between **activity level and gender** in this population.

EXERCISE 14.20

14.20 How do customers first hear about a new product? A random sample of 200 users of a new product was surveyed to determine the answer to this question. Other demographic data such as age were also collected. The respondents included 50 people under the age of 21 and 90 people between the ages of 21 and 35; the remainder was over 35 years of age. Of those under 21, 60% heard about the product from a friend, and the remainder saw an advertisement in the local paper. One-third of the people in the age category from 21 to 35 saw the advertisement in the local paper. The other two-thirds heard about it from a friend. Of those over 35, only 30% heard about it from a friend, while the remainder saw the local newspaper advertisement. Set up the contingency table for the variables age and method of learning about the product. Is there an association between the consumer's age and the method by which the customer heard about the new product?

Newbold et al (2013)

EXERCISE 14.20: SOLUTION

Answer:

Step 1: Construct the contingency table

Given information

- Total sample size: 200
- Under 21: 50
- Age 21–35: 90
- Over 35: $200 - 50 - 90 = 60$

Under 21

- 60% heard from a friend → $0.60 \times 50 = 30$
- 40% saw an advertisement → 20

Age 21–35

- One-third saw an advertisement → 30
- Two-thirds heard from a friend → 60

Over 35

- 30% heard from a friend → 18
- 70% saw an advertisement → 42

EXERCISE 14.20: SOLUTION

✅ Answer:

Observed Frequencies

Age Group	Friend	Advertisement	Total
Under 21	30	20	50
21–35	60	30	90
Over 35	18	42	60
Total	108	92	200

Step 2: Hypotheses

$$H_0 : \text{Age and method of learning about the product are independent}$$
$$H_1 : \text{Age and method of learning about the product are associated}$$

EXERCISE 14.20: SOLUTION

✅ Answer:

Step 3: Test statistic

The chi-square test statistic is

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where the expected frequencies are computed as

$$E_{ij} = \frac{(\text{row total})(\text{column total})}{\text{grand total}}$$

Expected Frequencies

Age Group	Friend	Advertisement
Under 21	27.0	23.0
21–35	48.6	41.4
Over 35	32.4	27.6

Chi-square value

$$\chi^2_{\text{obs}} \approx 19.9$$

EXERCISE 14.20: SOLUTION

Answer:

Step 4: Degrees of freedom

$$df = (r-1)(c-1) = (3-1)(2-1) = 2$$

Under H_0,

$$\chi^2 \sim \chi^2(2)$$

Step 5: Decision rule

At significance level α, reject H_0 if

$$\chi^2_{\text{obs}} > \chi^2_{1-\alpha}(2)$$

EXERCISE 14.20: SOLUTION

✅ Answer:

Step 6: Conclusion

Since

$$\chi^2_{\text{obs}} \approx 19.9$$

is much larger than the critical value $\chi^2_{1-\alpha}(2)$ for common significance levels (e.g., $\alpha = 0.05$), we **reject the null hypothesis**.

Final Conclusion

> There is **strong statistical evidence** of an association between **consumer age** and the **method by which customers first hear about the new product**.

THANKS!

Questions?