Bachelor's Degrees in

Management and

Finance and Accounting

Statistics 2

Hypothesis tests

Exercises with SPSS outputs. Part A

(Exercises up to the mid-term test)

Academic year 2024/2025

Problem 1:

Pedro and João discuss how quickly the same service can be provided in two different Citizen's Bureaux. The former says that in store 1 the service is faster on average than in store 2, but the latter doubts it. In order to clarify this situation, they decided to monitor the service on the same day and in the same time slot. They randomly selected 10 customers in store 1 and 15 customers in store 2, recording the time (in minutes) that elapsed between their arrival and the moment they were attended to.

Adapted from Exercício 17, p.171, Reis, E. et al (2020) Exercícios de Estatística Aplicada Vol.2
(One sample mean test)

Question1: Can you consider the average waiting time at the Citizen's Bureau to be 20 minutes?

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Waiting time (min)	25	21,00	9,014	1,803

One-Sample Test

	Test Value = 20						
			Significance			95% Confidence Interval of the Difference	
	t	df	One-Sided p	Two-Sided p	Mean Difference	Lower	Upper
Waiting time (min)	0,555	24	0,292	0,584	1,000	-2,72	4,72

a) Which test statistic should you choose? What assumptions should you make?

X - waiting time at the Citizen's Bureau, in minutes

Test statistic: $\frac{\bar{X}-\mu_0}{S'/\sqrt{n}} \sim t_{(n-1)}$

Assumptions:

- σ^2 unknown
- assume normal population, because n = 25 < 30

b) For a significance level (α) of 5%, what decision do you make?

H_0: $\mu = 20$
H_1: $\mu \neq 20$

Two-sided test → interpret two-sided p
p-value = 0,584 > α = 0,05 → do not reject H0 (for this significance level and this sample)

Interpretation: It can be assumed that the average waiting time at the Citizen's Bureau is 20 minutes, for α = 5%.

c) What is the limit value for the significance level that would lead to the opposite decision to the one you made?

RR:]-∞, -2,064] U [2,064, +∞[
NRR:]-2,064, 2,064[

To reject H_0, the significance level needs to be 0.584. With this new value for α
t = 0,555 ∈ RR.

Question 2: Can you consider that the average waiting time at Citizen's Bureau 2 is less than 25 minutes?

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Waiting time (min)	15	23,00	7,746	2,000

One-Sample Test

	Test Value = 25						
			Significance			95% Confidence Interval of the Difference	
	t	df	One-Sided p	Two-Sided p	Mean Difference	Lower	Upper
Waiting time (min)	-1,000	14	0,167	0,334	-2,000	-6,29	2,29

a) Define the hypotheses being tested. What assumptions you need to make when carrying out the test?

X - waiting time at the Citizen's Bureau 2, in minutes

H_0: $\mu \geq 25$
H_1: $\mu < 25$

Assumptions:
- - σ^2 unknown
- assume normal population, because n = 25

b) For a significance level of 5%, what decision do you make?

One-sided left test → interpret one-sided p
p-value = 0,167 > α = 0,05 → do not reject H0

Extra:
Interpretation: It can be assumed that the average waiting time at Citizen's Bureau 2 is at least 25 minutes, for α = 5%

Problem 2 (Midterm test 2018/2019)

As part of a market study on purchasing behaviour, the analyst wanted to test the hypothesis that the average age of those who buy clothing online is 38.

The SPSS outputs available are as follows:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	100	43,90	23,726	2,373

One-Sample Test

	Test Value = 38				
			Significance		
	t	df	One-Sided p	Two-Sided p	Mean Difference
Age	2,487	99	0,008	**0,015**	5,900

a) Formulate the hypotheses of the test that allows you to evaluate the hypothesis put forward by the analyst and determine the Critical Region and the Acceptance Region of the test.
Based on the information contained in the Outputs presented and the regions defined, indicate which decision should be taken.
Note: Consider for this purpose α = 0.01.

X – age of person who buy clothes online

H_0: μ = 38
H_1: μ ≠ 38

Two-sided test:

Test statistics (SPSS): $\frac{\bar{X}-\mu_0}{S\prime/\sqrt{n}} \sim t_{(99)}$

RR:]-∞, -1,984] U [1,984, +∞[
NRR:]-1,984, 1,984[

Decision:
$t = 2,487 \in RR$ → reject H_0 (given the significance level and the given sample)

NOTE: Since n>30, we could use the approximate distribution for the Test Statistic: $N(0,1)$, and in this case, we would have
RR:]-∞, -1,96] U [1,96, +∞[
NRR:]-1,96, 1,96[

b) Interpret, in detail and in the context of the hypothesis test under study, the value highlighted in the Output.

p-value = 0,015 (two-sided test) < α = 0,05 → reject H_0

p-value = 0,015 corresponds to the minimum significance level for which H_0 is rejected

Problem 3

In a study carried out to measure the time spent traveling from home to work, the following results were obtained:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
time spent traveling from home to work (min.)	86	25,37	7,640	0,820

One-Sample Test

	Test Value = 25					
			Significance		95% Confidence Interval of the Difference	
	t	df	Two-Sided p	Mean Difference	Lower	Upper
time spent traveling from home to work (min.)	0,451	85	0,653	0,370	-1,27	2,01

a) What are the hypotheses and assumptions of this test? Are any assumptions violated?
X – time spent traveling from home to work, in minutes

$H_0: \mu = 25$
$H_1: \mu \neq 25$

Assumptions:
- Any population distribution
- n = 86>30
- σ^2 unkown

No assumptions were violated.

b) What can you say about the average time spent traveling from home to work?
Two-sided test → interpret two-sided p
p-value = 0,653 > α = 0,05 → do not reject H_0

Interpretation: It can be assumed that the average time spent traveling from home to work is 25 minutes, for > α = 5%.

c) What should be the threshold value for the significance level that would lead to the opposite decision to the one you made?
Two-sided test:
Test Statistic (SPSS): $\frac{\bar{X}-\mu_0}{S'/\sqrt{n}} \sim t_{(85)}$

RR:]-∞, -1,984] U [1,984, +∞[
NRR:]-1,984, 1,984[

To reject H_0, the significance level needs to be 0.653. With this new value for α, t = 0,451 ∈ RR.

Problem 4 (Midterm test 2017/2018)

Data obtained from the PISA2002 study made it possible to develop various indicators. These include the indicator "Love for Mathematics", expressed on a scale of 0 to 100. A certain analysis of the data available in a sub-sample produced the following results:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Love for Math	50	54,06	15,798	2,234

One-Sample Test

	Test Value = 50					
			Significance		95% Confidence Interval of the Difference	
	t	df	Sig (2-tailed)	Mean Difference	Lower	Upper
Love for Math	1,819	49	0,075	4,064	-0,43	8,55

a) What is the purpose of this analysis? State the null and alternative hypotheses being tested, **in the context of the problem**.

The aim of this analysis is to find out whether the average indicator "Love for Mathematics" differs from 50 (midpoint on a scale of 0 to 100).

Com esta análise pretende-se perceber se a média das perceções sobre o gosto pela matemática difere de 50 (ponto intermédio numa escala de 0 a 100)

X – Love for Math, in a scale 0 to 100

H_0: $\mu = 50$
H_1: $\mu \neq 50$

b) What conclusion should be drawn (α=0.05)?
Two-sided test → interpret two-sided p
p-value = 0,075 > α = 0,05 → Do not reject H_0

c) Is there reason to believe that Love for Mathematics is on average above the midpoint of the scale? Justify.
No. By not rejecting H_0, we can admit that, on average, the indicator of liking mathematics do not differ from 50 (the middle point on a scale of 0 to 100).

Problem 5

Elections for the ISCTE-IUL Students' Association will be held soon.

Some A-list students decided to carry out a poll to find out whether 50% of students intend to vote for their list, and obtained the following results:

One-Sample Proportions Tests

		Observed						Significance	
Test Type		Successes	Trials	Proportion	Observed - Test Value[a]	Asymptotic Standard Error	Z	One-Sided p	Two-Sided p
List A: Yes	Score (Continuity Corrected)	53	100	0,530	0,030	0,050	0,500	0,309	0,617

a. Test Value = 0,5

a) Define the population under study, indicating its distribution.
X – number of ISCTE-IUL students, in 1, who will vote for list A in the Student Association elections.

$X \sim Bern(p)$ (Pop. Bernoulli)

b) What is the purpose of this analysis? Indicate the null and alternative hypotheses being tested, in the context of the problem.
The aim of this analysis is to determine whether the proportion of students intending to vote for list A is 0,5.

H_0: p = 0,50
H_1: p ≠ 0,50

c) What conclusion should be drawn (α=0.05)?
Two-sided test → interpret two-sided p

p-value = 0,617 > α = 0,05 → do not reject H_0

Interpretation: It can be assumed that the proportion of students intending to vote for list A is 0,5, for α = 5%.

Problem 6

A team of health professionals is trying to find out if at least 60% of the Portuguese population aged between 60 and 70 has been vaccinated against COVID-19.

To this end, a random sample of 112 people was taken. Some of the results obtained are presented below:

One-Sample Proportions Tests

Test Type		Observed						Significance	
		Successes	Trials	Proportion	Observed - Test Value[a]	Asymptotic Standard Error	Z	One-Sided p	Two-Sided p
Vaccinate: Yes	Score (Continuity Corrected)	58	112	0,518	-0,082	0,0472	-1,678	0,047	0,093

a. Test Value = 0,6

a) Which test statistic should you choose? What assumptions should you make?

X – number of Portuguese aged 60 to 70, out of 1, who are vaccinated against COVID-19

Test Statistic: $\frac{\bar{X}_b - p_0}{\sqrt{\frac{p_0 q_0)}{n}}}$

(Note: as in Statistics I, there was no continuity correction, we decided to consider the test statistic above)

Assumptions:
- population Bernoulli
- n = 112 > 30

b) State the null and alternative hypotheses being tested, in the context of the problem.

H_0: p ≥ 0,60

H_1: $p < 0,60$

c) For a significance level (α) of 5%, what decision do you make?

One-sided left test → interpret one-sided p

p-value = 0,047 < α = 0,05 → reject H_0

d) Is there reason to believe that at least 60% of the Portuguese population aged between 60 and 70 is vaccinated against COVID-19? Justify.

No. The proportion of Portuguese between the ages of 60 and 70 vaccinated against COVID-19 is less than 60%, for α = 5%.