iscte BUSINESS SCHOOL

Bachelor's Degrees in

Management and Finance and Accounting

Statistics 2

Hypothesis Testing.

Exercises with SPSS outputs. Part B

(Exercises after the midterm 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
(Comparison of two means, independent samples)

We can use this problem to introduce the test to the comparison of two variances (Levene's Test). Establishment of hypotheses and decision making on the equality of variances.

Group Statistics

	Store	N	Mean	Std. Deviation	Std. Error Mean
Time to service	Store 1	10	18,00	10,328	3,266
	Store 2	15	23,00	7,746	2,000

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Time to service	Equal variances assumed	,270	,608	-1,384	23	,180	-5,000	3,612	-12,471	2,471
	Equal variances not assumed			-1,306	15,605	,211	-5,000	3,830	-13,135	3,135

a) Is it to be admitted that the variability of the time until service is different in the two Citizen's Stores? Establish the hypotheses under test and proceed to the decision making assuming α=0.05.

b) Is Pedro right about the average waiting time for service?

I. Define the hypotheses under test. Which test statistic should you choose in this case?

II. Assuming a reference significance level of 5%, determine the Critical and Acceptance regions of the test.

III. What conclusions do you draw? Use the results of a) and the information contained in the tables above.

c) What is the threshold value for the significance level that would lead to a decision contrary to the one you made?

Problem. 2

In a study on sunglasses consumers, information was collected about the opinion of respondents about the importance of 5 factors potentially influencing the purchase of sunglasses: Advertising and marketing, Quality, Ergonomics, Price and Style. These aspects are valued on a continuous scale from 0 to 10.

The aim is to know the extent to which consumers of RB (a brand of sunglasses) differ from non-RB consumers in relation to the above-mentioned aspects. To this end, statistical analyses were carried out, the outputs of which are below.

Group Statistics

	Do you prefer RB?	N	Mean	Std. Deviation	Std. Error Mean
Advertising and marketing	0 No	309	5.0830	1.74907	.09950
	1 Yes	152	5.4052	1.58354	.12844
Quality	0 No	309	7.9036	.84953	.04833
	1 Yes	152	8.0458	.86776	.07038
Ergonomics	0 No	309	7.1128	.91648	.05214
	1 Yes	152	6.8435	1.01101	.08200
Price	0 No	309	6.4642	1.10639	.06294
	1 Yes	152	6.4790	1.08286	.08783
Style	0 No	309	8.1986	.83607	.04756
	1 Yes	152	8.0813	.89823	.07286

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Advertising and marketing	Equal variances assumed	2.65	0.10	-1.92	459	0.056	-0.32	0.17
	Equal variances not assumed			-1.98	329	0.048	-0.32	0.16
Quality	Equal variances assumed	0.00	0.97	-1.68	459	0.094	-0.14	0.08
	Equal variances not assumed			-1.67	295	0.097	-0.14	0.09
Ergonomics	Equal variances assumed	0.90	0.34	2.87	459	0.004	0.27	0.09
	Equal variances not assumed			2.77	276	0.006	0.27	0.10
Price	Equal variances assumed	0.01	0.94	-0.14	459	0.892	-0.01	0.11

	Equal variances not assumed			-0.14	306	0.891	-0.01	0.11
Style	Equal variances assumed	2.47	0.12	1.38	459	0.168	0.12	0.08
	Equal variances not assumed			1.35	282	0.179	0.12	0.09

1. For the Advertising and Marketing variable

a) Indicate the objective of the analysis performed and the **main** hypotheses being tested.

b) What are the assumptions and conditions of application underlying the analysis carried out? What can you conclude about your verification?

c) What are the main conclusions you draw from the analysis carried out? Justify.

2. Characterize RB spectacle consumers, in view of what significantly distinguishes them from non-RB consumers, among the 5 aspects indicated.

Problem 3: A nutritionist is convinced that the new diet he is prescribing to his patients is effective in treating obesity, causing weight loss after 4 weeks and, unlike other diets, reducing the patient's state of anxiety.

Adapted from Exercício 19, p.172, Reis, E. et al (2020) Exercícios de Estatística Aplicada Vol.2
(Comparison of two means, paired samples)

Question: Is it to be admitted that the average weight of patients is lower after 4 weeks of dieting?

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Weight (Kg) before the diet	92,7000	10	7,22726	2,28546
	Weight (Kg) after 4 weeks diet	88,7000	10	6,12917	1,93821

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Weight (Kg) before diet & Weight (Kg) after 4 weeks diet	10	,971	,000

Paired Samples Test

		Paired Differences					t	Df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Weight (Kg) before the diet - Weight (Kg) after 4 weeks diet	4,00000	1,94365	,61464	2,60960	5,39040	6,508	9	,000

Problem. 4:

A group of Sports Science students (n = 20) is selected from the population to investigate whether an innovative 12-week training program improves their performance in the long jump. To test whether this training improves performance, students are tested on their performance in the long jump before starting the training program and after finishing the training program.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Jump 1	2,4815	20	0,16135	0,03608
	Jump 2	2,5155	20	0,15982	0,03574

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Jump 1 – Jump 2	30	,935	,000

Paired Samples Test

		Paired Differences					t	Df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Jump 1 – Jump 2	-0,03400	0,03185	,00712	-0,04891	-0,01909	-4,773	19	,000

Can we conclude that there was an improvement in jump distance after the innovative 12-week training program? Use $\alpha = 0.05$.

Problem. 5

A surprising number of young adults (between the ages of 19 and 25) still live in their parents' house. A random sample conducted by a government institute included 685 men and 921 women in this age group. The survey revealed that 438 of the men and 371 of the women lived with their fathers.

Is this good proof that different proportions of young men and women live with their parents? Use $\alpha = 0.05$.

Independent-Samples Proportions Group Statistics

Gender		Successes	Trials	Proportion	Asymptotic Standard Error
Living with parents = Yes	= Male	438	685	0,639	0,018
	= Female	371	921	0,403	0,016

Independent-Samples Proportions Confidence Intervals

Interval Type		Difference in Proportions	Asymptotic Standard Error	95% Confidence Interval of the Difference	
				Lower	Upper
Living with parents = Yes	Agresti-Caffo	0,237	0,024	0,188	0,284
	Newcombe	0,237	0,024	0,188	0,284

Independent-Samples Proportions Tests

Test Type		Difference in Proportions	Asymptotic Standard Error	Z	Significance	
					One-Sided p	Two-Sided p
Living with parents = Yes	Wald H0	0,237	0,024	9,379	< 0.001	< 0.001

Problem. 6:

A trade association claims that its customers use their credit cards to pay for planned purchases. However, a banking institution counters and says that the credit card is used to make impulse purchases.

To analyze this situation, a random sample was built with 39 consumers who make impulse purchases and 33 consumers who make planned purchases.

To understand whether there are differences between the group of consumers who make impulse purchases and those who make planned purchases, the following hypothesis test was carried out:

Independent-Samples Proportions Group Statistics

Type of purchase		Successes	Trials	Proportion	Asymptotic Standard Error
Credit card payment	= Impulse buying	30	39	0,769	0,067
	= Planned purchase	18	33	0,545	0,087

Independent-Samples Proportions Confidence Intervals

Type of purchase		Difference in Proportions	Asymptotic Standard Error	95% Confidence Interval of the Difference	
				Lower	Upper
Credit card payment	Agresti-Caffo	0,224	0,110	0,002	0,424
	Newcombe	0,224	0,110	0,005	0,419

Independent-Samples Proportions Tests

Test Type		Difference in Proportions	Asymptotic Standard Error	Z	Significance	
					One-Sided p	Two-Sided p
Credit card payment	Wald H0	0,224	0,110	2,007	0,022	0,045

a) Identify the hypothesis test presented.

b) Is it possible to say that there is a difference in credit card use between planned purchase and impulse purchase? Use $\alpha = 0.01$

c) Would you make the same decision if the maximum error accepted in the analysis was 5%?

Problem 7:

An automaker (Not That Random Auto) needs to ascertain if the average lifetime of 4 brands of tires is significantly different, to choose the best supplier for its cars.

To compare the 4 brands, Not That Random sampled a few tires from each brand and measured their lifetime (in thousands of kilometers), and produced the results below.

Adapted from Exercício 22, p.174, Reis, E. et al (2020) Exercícios de Estatística Aplicada Vol.2 (ANOVA)

a) Use a significance level of 5% to test whether there are significant differences in the average lifespans of the four tire brands.

b) Which brands are significantly different from each other?

c) What do you conclude about the assumptions of equality of variances between populations and normality of populations?

Descriptives

Tire life time (thousands of Km)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Brand A	6	28,9167	2,57714	1,05211	26,2121	31,6212	25,00	32,00
Brand B	7	27,4286	2,82000	1,06586	24,8205	30,0366	24,00	32,00
Brand C	5	29,8000	1,15109	,51478	28,3707	31,2293	28,00	31,00
Brand D	6	23,9167	2,33274	,95234	21,4686	26,3647	21,00	27,00
Total	24	27,4167	3,15425	,64386	26,0847	28,7486	21,00	32,00

ANOVA

Tire life time (thousands of Km)

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	115,402	3	38,467	6,783	,002
Within Groups	113,431	20	5,672		
Total	**228,833**	**23**			

Multiple Comparisons

Dependent Variable: Tire life time (thousands of Km)

	(I) Tyre brand	(J) Tire brand	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Scheffe	Brand A	Brand B	1,48810	1,32494	,740	-2,5514	5,5276
		Brand C	-,88333	1,44207	,944	-5,2799	3,5132
		Brand D	5.00000*	1,37496	,016	,8080	9,1920
	Brand B	Brand A	-1,48810	1,32494	,740	-5,5276	2,5514
		Brand C	-2,37143	1,39446	,429	-6,6229	1,8800
		Brand D	3,51190	1,32494	,104	-,5276	7,5514
	Brand C	Brand A	,88333	1,44207	,944	-3,5132	5,2799
		Brand B	2,37143	1,39446	,429	-1,8800	6,6229
		Brand D	5.88333*	1,44207	,006	1,4868	10,2799
	Brand D	Brand A	-5.00000*	1,37496	,016	-9,1920	-,8080
		Brand B	-3,51190	1,32494	,104	-7,5514	,5276
		Brand C	-5.88333*	1,44207	,006	-10,2799	-1,4868

*. The mean difference is significant at the 0.05 level.

Test of Homogeneity of Variances

Tire life time (thousands of Km)

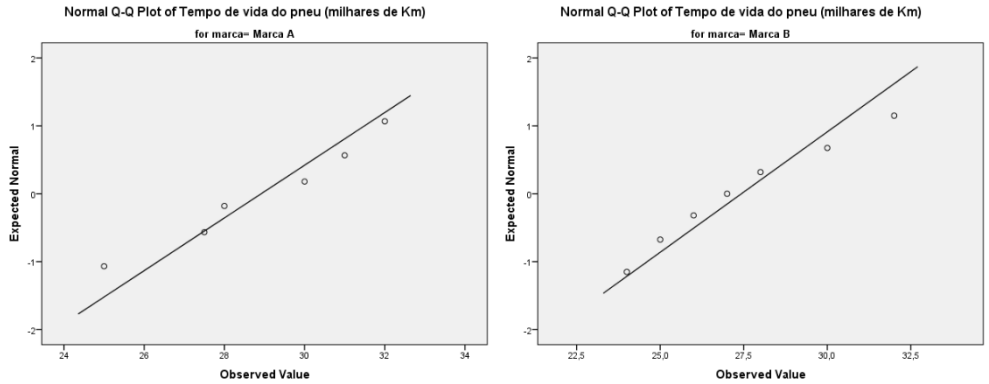
Levene Statistic	df1	df2	Sig.
1,512	3	20	,242

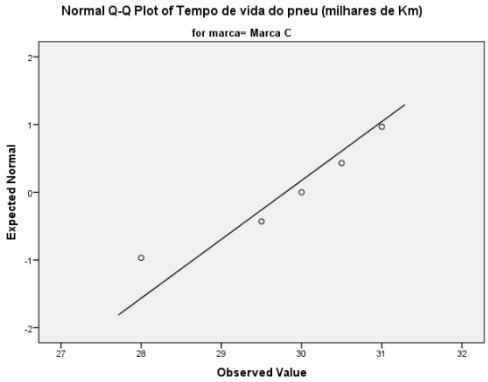
Tests of Normality

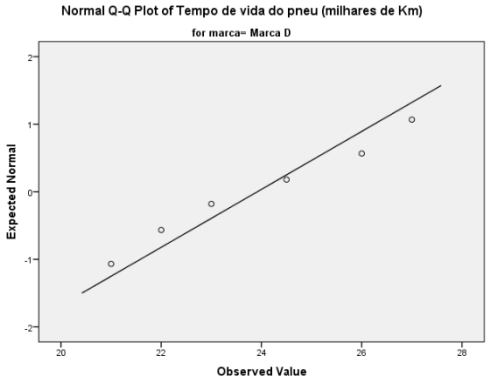
	Tyre brand	Kolmogorov-Smirnova			Shapiro-Wilk		
		Statistic	Df	Sig.	**Statistic**	**Df**	**Sig.**
Tire life time (thousands of Km)	Brand A	,163	6	,200*	**,965**	**6**	**,855**
	Brand B	,134	7	,200*	**,964**	**7**	**,853**
	Brand C	,197	5	,200*	**,943**	**5**	**,685**
	Brand D	,153	6	,200*	**,961**	**6**	**,826**

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction




Normal Q-Q Plot of Tempo de vida do pneu (milhares de Km)
for marca= Marca C
Expected Normal
Observed Value
2
1
0
-1
-2
27
28
29
30
31
32


Normal Q-Q Plot of Tempo de vida do pneu (milhares de Km)
for marca= Marca D
Expected Normal
Observed Value
2
1
0
-1
-2
20
22
24
26
28

Problem. 8

1. A data analyst performed a set of analyses on the variable *Monthly salary* for different job category, having obtained the following tables:

Table 1. Descriptives

Salary (u.m.)

	N	Mean	Std. Deviation	Std. Error
Vendor	363	10,2025	,24586	,01290
Office Employee	27	10,3375	,07001	,01347
Manager	84	11,0296	,26865	,02931
Total	474	10,3568	,39733	,01825

Table 2. ANOVA

Salary (u.m.)

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	46,674	**(A)**	**(B)**	392,564	,000
Within Groups	28,000	471	,059		
Total	74,675	473			

Table 3. Multiple Comparisons

Dependent Variable: Salary (u.m.)

Scheffe

(I)job cat.	(J) job cat.	Mean Difference (I-J)	Std. Error	Sig.
Vendor	Office Employee	-,13492*	,04864	,022
	Manager	-,82709*	,02952	,000
Office Employee	Vendor	,13492*	,04864	,022
	Manager	-,69217*	,05394	,000
Manager	Vendor	,82709*	,02952	,000
	Office Employee	,69217*	,05394	,000

*. The mean difference is significant at the 0.05 level.

a) What is the **primary** Purpose of the analyses carried out?

b) Calculate the values of **(A)** and **(B)** in Table 2.

c) Identify the hypotheses tested in Table 2 and the decision to be made in the context of the problem. Use $\alpha=0.05$. Complete the analysis by also considering the results in Table 3.

2. Subsequently, the analyst was confronted with the information from supplementary analyses that can be found in the tables below.

Table 4. Tests of Normality

		Shapiro-Wilka		
	Job category	Statistic	Df	Sig.
Salary (u.m.)	Vendor	,052	363	,021
	Office Employee	,262	27	,000
	Manager	,057	84	,200*

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 5. Test of Homogeneity of Variances

Salary (u.m.)

Levene Statistic	df1	df2	Sig.
13,168	2	471	,000

d) What are the (generic) hypotheses underlying the tests presented in Table 4? What conclusions do you draw? Use α=0.05.

e) What are the hypotheses underlying the test presented in Table 5? What conclusions do you draw? Use α=0.05.

f) What do the results of the previous paragraphs say about the conditions necessary to carry out the test presented in 1?

Problem. 9

A researcher from the Centre for Pedagogical Innovation is studying how the type of teaching practiced in universities affects the participation of students in class debates. To this end, this researcher counted how many times each student contributed orally in a total period of 24 hours of classes. The students attended traditional education or distance learning.

Ranks

Type of education		N	Mean Rank	Sum of Ranks
Number of Participations	Distance learning	12	8,54	102,50
	Face-to-face teaching	11	15,77	173,50
	Total	23		

Test Statistics[a]

	Age (years)
Mann-Whitney U	24,500
Wilcoxon W	102,500
Z	-2,600
Asymp. Sig. (2-tailed)	0,009
Exact Sig. (2-tailed)	0,009
Exact Sig. (1-tailed)	0,005

a. Grouping Variable: Type of Teaching

a) Identify the null and alternative hypotheses of the test presented.

H_0: The distribution of the number of class participations is the same for the two population groups of students attending distance learning and face-to-face learning

H_1: The distribution of the number of class participations is different for the two population groups of students attending distance learning and face-to-face learning

b) Do you consider the test appropriate for the proposed objective? If you do not agree with the option that was made, indicate alternative(s), justifying it.

The test applied is the non-parametric Mann-Whitney test. This test only requires that the two samples be random and independent and the variable under study be quantitative or qualitative ordinal.

In this case, the conditions are met. The variable under study is quantitative, discrete, (number of participations), and it is intended to compare the behavior of two population groups from two independent samples (students who attend distance learning and students who attend face-to-face education). We assume that the samples are random.

Thus, we can say that the test is adequate.

Note: This test is an alternative test to the t-test for two independent samples when the assumption of normality is violated. In this case, we have no information about the assumption of normality.

c) Is the number of interventions in class different between students who attend traditional education and students who attend distance learning? Use a significance level of 5%.

The output presents the asymptotic significance and the exact significance.

The significance to be interpreted depends on the sample size. As a rule, we can use:

- $n1 + n2 \leq 30$ → exact significance
- $n1 + n2 > 30$ → asymptotic significance

In this case, we use exact significance.

Exact Sig. (2-tailed) = 0.009, $\alpha = 0.05$ and Sig < α → Reject H0

Interpretation:

For the significance level of 5% and the samples used, there is statistical evidence to suggest that the number of interventions in class is different between students who attend traditional education and students who attend distance learning.

Problem. 10

A researcher is studying whether certain antidepressant medications may have the positive side effect of decreasing neurological pain in individuals with back pain when given in doses lower than those prescribed for depression.

The researcher identifies 3 well-known antidepressant drugs that may have this positive side effect and names them drug A, drug B and drug C. With a group of 60 individuals with a similar level of back pain and randomly assign them one of three groups - drug A, drug B or drug C - and prescribe the drug in question for a period of 4 weeks. At the end of the 4-week period, the researcher asks participants to rate their back pain on a scale of 1 to 10, where 10 indicates the maximum level of pain.

The researcher intends to compare the levels of pain felt by the different groups at the end of the drug treatment period.

Ranks

Drug type		N	Mean Rank
Pain classification	Drug A	20	35,33
	Drug B	20	34,83
	Drug C	20	21,35
	Total	60	

Test Statistics[a,b]

	Pain classification
Kruskal-Wallis H	8,520
Df	2
Asymp. Sig.	0,014

a. Kruskal Wallis Test

b. Grouping Variable: Type of Drug

What conclusion do you draw from the results presented? Use a significance level of 5%.

Problem. 11: Chi-square test of independence

In a study on working conditions in a large company, the following analyses were carried out, the outputs of which are below.

Table 1. Crosstab

			Is there flexibility in schedules?		*Total*
			Yes	No	
Company Area	Area1	Count	119	33	152
		% within Area	78,3%	21,7%	100,0%
		% within Is there time flexibility?	34,8%	29,7%	33,6%
	Area 2	Count	55	13	68
		% within Area	80,9%	19,1%	100,0%
		% within Is there time flexibility?	16,1%	11,7%	15,0%
	Area 3	Count	34	45	79
		% within Area	43,0%	57,0%	100,0%
		% within Is there time flexibility?	9,9%	40,5%	17,4%
	Area 4	Count	89	11	100
		% within Area	89,0%	11,0%	100,0%
		% within Is there time flexibility?	26,0%	9,9%	22,1%
	Area 5	Count	45	9	54
		% within Area	83,3%	16,7%	100,0%
		% within Is there time flexibility?	13,2%	8,1%	11,9%
Total		Count	342	111	453
		% within Ward Method	75,5%	24,5%	100,0%
		% within Is there time flexibility?	100,0%	100,0%	100,0%

Table 2. Chi-Square Tests

	Value	*Df*	*Asymp. Sig. (2-sided)*
Pearson Chi-Square	58.35(a)	4	,000
Likelihood Ratio	53,11	4	,000
Linear-by-Linear Association	0,557	1	,455
N of Valid Cases	453		

a 0 cells (,0%) have expected count less than 5. The minimum expected count is 13,23.

a) What was intended to be tested in the analysis carried out? What statistical test was carried out, and what are the hypotheses being tested?

It is intended to test whether there is a significant relationship between 2 qualitative variables: area of the company (Area 1, Area 2, Area 3, Area 4, and Area 5) and the existence of flexible working hours with categories Yes and No.

Independence test of two qualitative variables: company area and existence of flexible hours

X – Company area

Y – Existence of flexible working hours

Hypotheses to be tested:

H_0: X and Y are independent variables

H_1: X and Y are related variables (not independent)

b) What are the conditions of applicability underlying the analysis carried out? Are they all verified?

Conditions of applicability:

- a maximum of 20% of the expected frequencies (E_{ij}) less than 5;

- all expected frequencies (E_{ij}) greater than one

Both conditions are met. All expected values are greater than 5. ($E_{ij} \geq 13{,}23$).

c) Interpret the values marked in grey in Table 1.

Value = 19.1%

In the group of workers belonging to area 2 (68 workers), 19.1% (13 workers) indicated that they did not have flexible working hours.

(13/68) * 100 = 19.1%

Value = 26.0%

In the group of workers who indicated that they had flexible working hours (342 workers), 26.0% (89 workers) belonged to area 4.

(89 / 342) * 100 = 26.0%

d) What is the <u>expected</u> number of workers in area 1 who claim to have flexible hours?

Expected value: E11 = 152*342 / 453 = 114.76

e) What conclusions do you draw, considering a significance level of 5%?

Test Statistic Value: 58.35 and Sig < 0.001, $\alpha = 0.05$

Since Sig (p-value) < α → Decision: Reject H0

Interpretation: The null hypothesis of independence between the two variables should be rejected. In other words, there is a significant relationship between the work area and the existence of flexible hours.

From the sample data it is possible to draw the following conclusions: Workers who belong to areas 1, 2, 4 and 5 stated, for the most part, that they have flexible working hours; Workers belonging to

area 3 have a more balanced distribution with regard to the existence of flexible working hours, with a slightly higher percentage of workers without flexible working hours (57.0%).

The workers who reported having flexible working hours belonged mainly to area 1 (34.8%) and area 4 (26.0%), while the distribution of workers without flexible working hours was different, with emphasis on area 3 (40.5% of the workers), followed by area 1 (29.7%).

Problem. 12:

In a study on newspaper reading habits, the analysis was carried out, the outputs are below.

Preferred Weekly * Sex Crosstabulation

			Sex		
			0 female	1 male	Total
Preferred weekly	Expresso	Count	13	36	49
		Expected Count	16,3	32,7	49,0
		% within Preferred Weekly	26,5%	73,5%	100,0%
		% within Sex	39,4%	54,5%	49,5%
	Semanário	Count	5	20	25
		Expected Count	8,3	16,7	25,0
		% within Preferred Weekly	20,0%	80,0%	100,0%
		% within Sex	15,2%	30,3%	25,3%
	Sol	Count	15	10	25
		Expected Count	8,3	16,7	25,0
		% within Preferred Weekly	60,0%	40,0%	100,0%
		% within Sex	45,5%	15,2%	25,3%
Total		Count	33	66	99
		Expected Count	33,0	66,0	99,0
		% within Preferred Weekly	33,3%	66,7%	100,0%
		% within Sex	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11,020A	2	,004
Likelihood Ratio	10,663	2	,005
Linear-by-Linear Association	6,533	1	,011
N of Valid Cases	99		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,33.

a) What are the conditions of applicability underlying the analysis carried out? Are they all verified?

Conditions of applicability:

- a maximum of 20% of the expected frequencies (E_{ij}) less than 5;
- all expected frequencies (E_{ij}) greater than one

Both conditions are met. All expected values are greater than 5. I.e., $E_{ij} \geq 8.23$.

b) Is it to be admitted that the preference for the various weeklies is influenced by the gender of the readers? (Considering a significance level of 5%)?

X – Favorite weekly newspaper

Y - Sex

Hypotheses to be tested:

H_0: X and Y are independent variables

H_1: X and Y are related variables (not independent)

Test Statistic Value: 11.02 and Sig = 0.004 , $\alpha = 0.05$

Since Sig < α → Decision: Reject H0

Interpretation: The null hypothesis of independence between the two variables is rejected. In other words, there is a relationship between the preferred weekly newspaper and sex.

c) If you had to do the calculations by hand, explain clearly and succinctly the choice of type of test (Critical area/Decision rule) you performed.

The critical region is of the unilateral type on the right.

The test statistic is built based on the difference between the observed frequencies and the expected frequencies. The greater this difference, the greater the value of the test statistic, which means that the null hypothesis should be rejected. I.e., high values of the test statistic should be rejected.

To the extent that the chi-square distribution is a positive distribution, the critical region is on the right-hand side.

For $\alpha = 0.05$: $P\left[\chi^2_{(2)} \leq a\right] = 0{,}95$ a = 5.99

OR: [5.99; +∞[and RA: [0; 5.99[

Problem 13: Chi-square test for adherence (Goodness-of-fit) (previous years exam)

A market study was carried out with the aim of knowing the purchase profile of Portuguese consumers regarding sunglasses, and ISCTE Statistics 2 students were asked to analyze the database of the results of the consumer survey. From this analysis, the following results were obtained, among others:

How much did they cost

	Observed N	Expected N	Residual
Less than 100€	72	153,3	-81,3
100€ - 200€	123	153,3	-30,3
200€ - 300€	360	153,3	206,7
300€ or more	58	153,3	-95,3
Total	613		

Test Statistics

	How much did they cost
Chi-Square	**(a)**
Df	3
Asymp. Sig.	0,000

the. **(b)** cells **(c)** have expected frequencies less than 5. The minimum expected cell frequency is **(d)**.

a) Identify the variables under analysis.

X – Price, in euros, of sunglasses (Ordinal qualitative variable)

b) Calculate the missing values **(a), (b), (c),** and **(d)** in the output tables.

(a) Test Statistic Value

$X2 = [(-81.3)^2 + (-30.3)^2 + 206.72 + (-95.3)^2] / 153.3 = 59376.11 / 153.3 = 387.32$

(b) 0
(c) 0%
(d) 153,3

c) Identify the hypothesis test performed and its objective. Formulate hypotheses and draw conclusions (use α=0.05).

Uniform Distribution Goodness-of-fit Test (Identical Preferences)

Objective: It is intended to find out if consumer preferences are identical when choosing glasses based on price or if, on the contrary, there are differences in consumer preferences when choosing the target price segment

Hypotheses to be tested:

H_0: X follows a uniform distribution

H_1: X does not follow a uniform distribution

Or

H0: $X \sim U_{(4)}$

H1: $X \nsim U_{(4)}$

Sig < 0.001 and α = 0.05 and Sig < α → Decision: Reject H0

Or

$P[\chi^2_{(3)} \leq a] = 0{,}95$ a = 7.81 OR: [7.81; +∞[

ET = 387.32 ∈ RC → Decision: Reject H_0.

Interpretation: Given the level of significance and the sample considered, we can conclude that consumer preferences for the four price categories of sunglasses are not identical.

(d) What type is the critical region of the hypothesis test under (c)? Why? Justify your answer adequately.

The critical region is of the unilateral type on the right.

The test statistic is built based on the difference between the observed frequencies and the expected frequencies. The greater this difference, the greater the value of the test statistic, which means that the null hypothesis should be rejected. I.e., high values of the test statistic should be rejected.

To the extent that the chi-square distribution is positive, the critical region is on the right-hand side.