ISCTE Business School
Instituto Universitário de Lisboa

Management Degree
Statistics II

Hypothesis Testing Problems

1. The Eating Place Restaurant has been recording mean daily sales of € 800 for the past six years with a standard deviation of € 64. Daily sales are normally distributed. A new advertising program was started ten weeks ago and the mean daily sales for a random sample of nine days have been € 830. The owner has reason to believe that there is really no actual effect of the new advertising program on mean daily sales. In other words, he believes that the average of € 830 for nine days was due to chance alone. Your objective is to evaluate whether these results are significant or if the higher sales could indeed be due just to chance. Use α=0.05.

Answer: Since the value of Z falls below the critical value (that is, 1.41 < 1.645), we do not reject H_0. We have not been able to show from our data that the new program has a significant effect on daily sales at the 0.05 significance level.

2. A meat market manager is evaluating a supplier's claim that he has changed his procedures and trims his steaks so that there is less fat. The manager had determined that, over the past six months, for each 100 grs of beef purchased, he needed to remove 16.4 grs of fat. The results were normally distributed with a standard deviation of 2.4 grs. To test the supplier's claim, the manager now obtains a sample of 36 (100-grs. groups), giving a mean of 15.64 grs of fat. Was this significantly lower, at the 0.05 significance level, than the original 16.4 grs? What would happen if the manager wanted to make a decision with less risk, for example, with a significance level of 0.01?

Answer: At the 5 percent significance level we reject the null hypothesis, consider the alternative hypothesis as true and conclude that the supplier's new procedure results in less fat. At the 1% significance level the decision should be not reject the null hypothesis.

3. A tire manufacturer has been producing passenger-car radial tires with an average life, μ, of 50,000 miles with a normal distribution and a standard deviation, σ, of 4,000 miles. The company has been experimenting a modified synthetic additive, and a sample of 16 tires of the modified design gave a mean life, $\bar{x}$, of 52,000 miles. Is it likely that this sample could have come from the original population? Based on the results, should the company decide to carry on the new modification and put it into production or not? Let α=0.025.

Answer: We reject the null hypothesis at the 0.025 significance level and conclude that the new tire might be better; the company should carry on the new modification.

4. The owner of Kate Cake Company stated that the average number of buns sold daily was 1,500. A worker in the store wants to test the accuracy of the boss's statement. A random sample of 36 days showed that the average daily sales were 1,450 buns. Using a level of significance of α=0.01 and assuming σ=120, what should the worker conclude?

Answer: Since the value of the test statistic is -2.5, which is between ±2.576, the null hypothesis should be not rejected at the 0.01 level of significance, so the owner might be right.

5. Mr. Drinkwater of the D-T Liquor Shoppe thinks that his business sells an average of 17 pints of Border Ale daily. His partner, Mr. Taylor, thinks this estimate is too optimistic. A random sample of 36 days showed a mean of 15 pints and a sample standard deviation of 4 pints. Test the accuracy of Mr. Drinkwater's statement at the 0.05 level of significance.

Answer: Since the value of the test statistic is less than -1.64, it's necessary to reject Mr. Drinkwater's claim at the 0.05 level of significance. Mr Taylor might be right.

6. A journal article claims that the average height of female adults in Biglandia is 64 inches. To test this claim, a sociologist took a random sample of 16 Biglandia women and found that the mean was 62.9 inches and the standard deviation was 2.5 inches. Can it be concluded that the article is correct at the 0.05 level of significance?

Answer: Since the value of the test statistic falls between ±2.131, there is no reason to reject the article's statement at the 0.05 level of significance.

7. An automatic process produces 96 percent of good parts and 4 percent of defective ones. The manufacturer is satisfied with this since new machinery would be very expensive and the defective parts can be easily sorted out. He does, however, want to be able to detect if something changes in the process that would increase this percentage of defectives. One day he takes a sample of 400 items and finds 22 defective parts. At the 5 percent significance level, should the process be stopped because of a significant difference?

Answer: At the 5 percent significance level, H_0 is rejected if the value of the test statistic is greater than 1.645. Since this value is 1.53, we do not reject H_0 and assume that the sample is from the original process. If the value of the test statistic was greater than 1.645 we would probably stop the process and look for the possible cause.

8. The manager of the Big-Wig Executive Hair Stylists, Hugo Bald, has advertised that 90 percent of the firm's customers are satisfied with the company's services. Polly Tician, a consumer activist, feels that this is an exaggerated statement that might require legal action. In a random sample of 150 of the company's clients, 132 said they were satisfied. What should be concluded if a test were conducted at the 0.05 level of significance?

Answer: Since the value of the test statistic is greater than -1.64, there's no sufficient reason to doubt Hugo's claim. Polly should look for another cause.

9. The market research section of a national distributor wants to determine whether college students can distinguish their cola from other brands just by its taste. An experiment was designed in which 200 randomly chosen college students were each given four small glasses of cola, only one of which contained their favourite brand. After tasting all four drinks, 64 correctly identified their brand.

a) Are the results significantly greater than would be expected by chance? Use $\alpha=0.025$.

b) The market research manager of the cola manufacturer asks: "For what significance levels α would the decision be changed?".

Answers: a) This test supports the conclusion that college students, in general, have some ability to identify their favourite brand of cola by taste alone. It suggests that more than 25% of all college students have the ability of identify their favourite brand in competition with the three other brands. This finding is significant at the 0.025 level, so it indicates that the college students can distinguish their favourite brand by taste alone. b) The p-value for this test is 0.011, so the null hypothesis should not be rejected for any significance level lower than 0.011.

10. The advertising manager of a large television station has used 10 minutes per day per household as the average rate that people spend watching television ads. He wants to know whether people are spending less time watching ads than previously. This could adversely affect the station's major source of revenues, which is from spot ads. A recent monitor by time meters of 100 randomly selected households showed a mean of 9.5 minutes per day. The population standard deviation is 4.0 minutes. Let α=0.025.

a) Identify (in words) the question and the type I and the type II errors.

b) Compute the critical limit, C_1, for a lower-tailed test of the hypothesis H_0: μ = 10, H_1: μ < 10.

c) Compute the probability of type II error if the true mean rate were 9.5 minutes per day. (Notice that is one value in the range of the alternative hypothesis of μ < 10).

d) Develop a power curve for this test (H_0: μ = 10, H_1: μ < 10), that is, graph the probabilities of rejecting H_0 for mean times of spot ads ranging from 8.0 to 10.0 minutes.

e) Suppose that the manager has not tested H0: μ=10 minutes but proposes to take a sample that will satisfy the following conditions:

1. He believes that people watch spot ads an average of 10 minutes per day. He sets the α-risk at 0.025 of rejecting this null hypothesis if, in fact, it is true (that is, α=0.025 when H_0: μ = 10, therefore Z_α=-1.96);

2. If people actually watched only 9.0 minutes per day, he would set the β-risk at only 0.10 of incorrectly not rejecting H_0: μ = 10 (that is, β=0.10 when μ = 9, therefore Z_β=1.28).

Find the sample size required to test this null and alternative hypothesis.

Answers: a) A type I error would be to reject the null hypothesis when it is true and to conclude that people watch fewer ads, although they really watch 10 minutes or more of ads per day. A type II error would be to not reject the null hypothesis when it is false and to conclude that people still watch al least 10 minutes per day of ads, although they really watch less.

b) C_1=9.2 minutes per day.

c) The probability of the type II error, β, is 0.7734 when the true mean is 9.5.

d) The graph shows that the probability of not rejecting H_0 increases toward 1-α as the value for μ approaches μ_0=10.0. Values for μ>10 are of less interest, and the probability of not rejecting for those values approaches 1.

e) Decision rule: take a sample of 168 television viewers and find the sample mean advertising-viewing time per day. If $\bar{x}$ > 9.4 minutes per day, do not reject the hypothesis of a population mean of 10 minutes per day. Otherwise reject the hypothesis and conclude that the spot ad viewing averages less than 10 minutes per day.

11. Two work methods are being compared for assembly of sports equipment. For a random sample of 64 employees working by the traditional method, the mean production rate was 68.8 units. This population standard deviation is 5.2 units. For a random sample of 80 employees who used a new method, the mean production rate was 70.5 units. The standard deviation for the new method is 5.6 units. If the new method yields a statistically higher rate, it will become the standard procedure. Test the hypothesis that the new method results in a statistically higher production mean at α=0.025.

Answer: We cannot reject H_0 of no difference, because the value of the test statistic is 1.89 < 1.96. Therefore, there is no statistical difference in production performance for the new method versus the traditional method.

12. A class of 28 students preparing to take a licensing examination is divided into two groups. The first group of 16 took a 30-hour course, whereas the second group of 12 took a 24-hour course. After the courses were completed, a test was given with the following results, where $\bar{X}$ represents the sample mean test score (it is known that the populations for the two groups are normally distributed with equal variances):

30-Hour Course	24-Hour Course
n_1=16	n_2=12
$\bar{x}_1$ =76	$\bar{x}_2$ =72
s'_1=7	s'_2=9

We want to answer the question, "Is the difference between the mean test scores significant?" Let α=0.05.

Answer: Since the calculated value of the test statistic is 1.324, lower that the critical value of t, which is 2.056, we do not reject the null hypothesis that stated there was no difference. The evidence has not been sufficient to convince us that a significant difference exists between the two population course means, at the 5% level of significance.

13. A distributor of computer products has major outlets in several cities, including Gaia and Setúbal. The company president is considering entering into an exclusive dealership arrangement with a manufacturer of scanners. He wants to know whether potential customers in Gaia and Setúbal have the same preference for the Scannerspeed II. Random samples of customers reveal the data shown in following table. Test the hypothesis that there is no difference in the proportion preferring Scannerspeed II in Gaia and Setúbal. Use a 2% level of significance.

Customer preference	Gaia	Setúbal
Scannerspeed II	62	46
Other brands	47	38
No preference	111	96
Total sample	*220*	*180*

Answer: We cannot reject H_0 of no difference, because the value of the test statistic is 0.61 < 2.05. The evidence suggests that the customer preference for the Scannerspeed II is consistent (no difference) in Gaia and Setúbal for an α=0.02.

14. A claim has been made that there is a greater proportion of women automobile owners in Lisbon than in Porto. Proceeding to test this claim, we have available a random sample of 144 automobile owners from Lisbon and 225 owners from Porto. Our data show that 81 of the 225 automobile owners in Porto are women and 64 of the 144 car owners in Lisbon are women. Can it be concluded that the claim is correct at the 10% level of significance?

Answer: Since the value of the test statistic, 1.62, is greater than the critical value 1.282, we reject the null hypothesis and conclude that the proportion of women car owners in greater in Lisbon than in Porto.

15. A mail advertiser receives its mail brochures packaged in cartons. The quantities can be reasonably measured by the carton weights. The following are the weights in pounds of a sample of five cartons from a large lot shipped by the regular supplier.

Carton No.	1	2	3	4	5
Regular Supplier	1497.7	1493.7	1453.7	1443.8	1438.5

Specifications call for a mean of 1470.0 pounds with a variance of no more than 4% of the mean weight. Excessive variation is a concern because of the potential spillage loss (split cartons from overfills) and potential under- or overfilling of work orders.

a) Is this shipment within the contract for weight variance? Let α=0.05.

b) The mail advertiser has decided to test its regular supplier package weights against those of another supplier. The weights in pounds of a sample of five cartons packed by the new supplier are displayed below:

Carton No.	1	2	3	4	5
New Supplier	1467.9	1487.4	1456.8	1461.1	1470.5

Both suppliers are tested against the contract specifications. Are the cartons shipped by the new supplier less variable in weight than those shipped by the regular supplier? Let α=0.05.

c) A final requirement is for the supplier to be on target for the required mean weight. The variability could be quite small, but if the mean were excessive, the mail advertiser would be charged for overweight cartons! Test the carton weights for equal means using α=0.05.

Answer: a) Reject H_0 since the calculated value of the test statistic $53.93 > \chi^2_{0.05,4} = 9.488$. The sample suggests that the weight variance is beyond the contract specifications for α=0.05.

b) The calculated F is 5.74<6.39. New supplier's cartons are not statistically less variable than those of regular supplier at the α=0.05. There is no sufficient reason to switch suppliers based on the results of this test.

c) We cannot reject the hypothesis of equal means because the |calculated t| < 1.86. Both means appear to be below the target weight, 1470 pounds, and so are acceptable. We will other criteria to decide on the choice of a supplier.

16. A supervisor of a cosmetic packaging operation wants to estimate the increase in production for workers who are given work methods training. She selects a random sample of ten workers. The worker's output rate is measured first as he or she begins a new job and then two weeks after the end of the training. These measures are presented in the following table.

Worker	After training	Before training
1	51	15
2	63	21
3	55	20
4	68	36
5	48	12
6	38	9
7	54	17
8	73	42
9	49	26
10	57	18

a) Test the hypothesis that the average output has increased by less than or equal to 30 units per period against the alternative that the increase is greater than 30 units per period. Assume that the distribution of differences is normal and α=0.05.

b) Gain additional information by setting a 90% confidence interval estimate for the sample mean increase. Interpret the result.

Answer: a) Reject H0 since the calculated t is 2.327 > 1.83. This suggests an average increase of more than 30 units for post training over pre training productivity at the α=0.05 level.

b) The true mean improvement is estimated to be within 31 to 37 units of production.

17. For purpose of advertising, a market researcher would like to determine if a laundry soap is better than another with respect to brightness of the wash. The researcher takes swatches of 12 different types of cloth, each uniformly soiled, cuts each piece in half, and by random choice watches one part in laundry soap A and the other in soap B. The brightness of the resulting washes is then measured with a special meter. The resulting paired observations appear in the following table. Is there a significant difference in the brightness induced by the two laundry soaps? Use a level of significance of 0.05.

Swatch	1	2	3	4	5	6	7	8	9	10	11	12
X_A	5	4	4	1	6	3	4	5	5	3	6	4
X_B	1	3	4	2	4	3	2	3	6	5	3	2

Answer: Not reject H_0 because the value of the test statistic is 1.915 < 2.20 (p-value = 0.082). These data evidence no statistical difference in brightness at the level of significance of 0.05 for the fabrics washed in these two laundry soaps.

18. A building contractor wants to select the best possible plywood sheeting. A key indicator of quality is the weight of the sheets after the wood-preserving treatment is applied. The data of the plywood sheets are randomly selected for the assignment of treatments and are shown in the table. Test whether the mean weights of sheets from wood-preserving processes 1, 2 and 3 are equal. Use α=0.05 and assume that the three samples are drawn from normal populations, all with the same σ^2.

Weight of Plywood Sheets from Different Wood-Preservative Processes

Surface Stain (1)	Pressure Sealer (2)	Liquid Injection Treatment (3)
65	64	68
68	65	70
68	66	70
70	68	74
70	72	74
71	74	76

Answer: Since the calculated F=2.68 is less than $F_{0.05,2,15}$ =3.68 (p-value = 0,110), we conclude the sample data was drawn from three wood-preserving processes populations with no significant mean differences in plywood weights.

Assumptions: Equality of variances: Do not reject (p-value=0.194); Population normality: Group 1 – Not reject (p-value=0.405); Group 2 – Not reject (p-value=0.429)¸ Group 3 – Not reject (p-value=0.458).

19. A national retail chain department store has arranged to give merchandise-pricing training to its employees. Three different groups of sales employees are selected at random to experience the training. The three types of training are (1) hands-on training under the supervision of an experienced pricing manager, (2) video instruction in merchandise pricing, and (3) training by reading a pricing manual and guide. Upon completion of their instruction, the individuals are given a practical test in pricing. The highest possible score is 400 points. Due to illness, vacation and so on, the test produced unequal test groups of 7, 5 and 6 persons (see table). Is there a statistical difference in the population mean scores under the three methods? Use α=0.01 and assume that the three groups are drawn from normal populations, all with the same σ^2.

Group

1 Hands-on	2 Video Instruction	3 Pricing Manual
396	383	368
397	380	369
398	387	366
400	385	366
398	378	364
395		368
383		

Answer: The calculated F of 75.04 is substantially larger than the $F_{0.01,2,15}$ =6.36. So it is significant at p<0.0001. We conclude that the pricing test scores are statistically different for these three population groups. The sample means indicate that performance scores were ranked high to low in order of hands-on experience, video instruction and pricing manual.

Assumptions: Equality of variances: Do not reject (p-value=0.417); Population normality: Group 1 – Reject (p-value=0.007); Group 2 – Not reject (p-value=0.884)¸ Group 3 – Not reject (p-value=0.566).

20. The Portuguese Tax Authority wants to know if there are significant differences between the average VAT (measured in monetary units) paid by the 'Gourmet' restaurants of three cities: Lisbon, Oporto and Faro. With this objective, a random sample of restaurants was collected in each city and the following results were obtained:

Table 1 - Descriptives

Values of VAT (m.u.)

Cities	N	Mean	Std. Deviation	Std. Error
Lisbon	5	91,0000	24,59675	11,00000
Oporto	5	77,0000	27,06474	12,10372
Faro	5	110,0000	40,62019	18,16590
Total	15	92,6667	32,39635	8,36470

Table 2 - ANOVA

	Sum of Squares	df	Mean Square	F
Between Groups	2743,333	(a)	(d)	(f)
Within Groups	11950,000	(b)	(e)	
Total	14693,333	(c)		

a) Complete Table **2** and help the Tax Authority to answer its objective. Identify the hypotheses being tested and the rejection region. Use α=0,05.

b) Identify the hypotheses being tested in Tables **3** and **4** and the decisions to be made. Use α=0,05.

c) In your opinion are all the assumptions needed to apply the previous tests verified? Justify your answer.

Table 3 – Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
1,032	2	12	,386

Table 4 - Tests of Normality

Cities	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Lisbon	,273	5	,200*	,871	5	,271
Oporto	,147	5	,200*	,987	5	,966
Faro	,197	5	,200*	,934	5	,627

a. Lilliefors Significance Correction

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

Answers: a) (a) = 2, (b) = 12, (c) = 14, (d) = 1371.667, (e) = 995.833, (f) = 1.377; Decision: not reject H_0.

b) Table 3: test for equality of variances of VAT (m.u.) paid by the 3 population groups of restaurants (Lisbon, Oporto and Faro; Decision: not reject H_0.

Table 4: for each of the population groups (Lisbon, Oporto and Faro) there are two tests (Kolmogorv-Smirnov and Shapiro-Wilk) on normality of the variable VAT paid (m.u.). Considering the results of Shapiro-Wilk test, more appropriate for small samples dimensions, the decisions are:

- For the population group of Lisbon restaurants: do not reject H_0;
- For the population group of Oporto restaurants: do not reject H_0;
- For the population group of Faro restaurants: do not reject H_0..

c) Yes. (2nd test 4th June 2012)

21. The Portuguese Tax Authority believes that the average VAT to be paid by the end of this year by 'Gourmet' restaurants will be greater than 120 m.u.. However, some tax experts are convinced the true average VAT value will be at most 120 m.u. by the end of the year. A random sample of the population under analysis was collected and the following results were obtained:

$$\sum_{i=1}^{36} X_i = 4320\, u.m. \qquad \sum_{i=1}^{36}(X_i-\bar{X})^2 = 1296\, u.m.^2$$

a) Define the population being studied.

b) Is the Portuguese Tax Authority right? Justify your answer with the appropriate statistical test using a significance level of 5%.

c) In the previous question, which is the probability of making an incorrect decision if the real average value of the VAT paid by the end of the year turns out to be 123 m.u.?

d) Calculate the power of the previous test decision if the real average value of the VAT paid by the end of the year turns out to be 121,645 m.u.?

Answers: a) X: anual value (m.u.) of VAT paid by "Gourmet" restaurants. b) PTA might not be right. c) 0.0869. d) 0.50.

(2nd test 4th June 2012)

22. A car selling company wants to decide if preferences for three new models (A, B and C) are different. A random sample of consumer preferences was collected and the results were as follows:

Car Model	Number of preferences
A	370
B	440
C	390

a) Identify the appropriate statistical test to answer the company's objective.

b) Define the hypotheses being tested and the decision and conclusion to be made for a value of α=0.05.

c) Are the necessary conditions to apply the statistical test verified? Why?

d) Which is the probability (Sig-value) associated to the value of the test? To answer this question use the following table relative to the Qui-square distribution.

Prob.	0,960	0,961	0,962	0,963	0,964	0,965	0,966	0,967	0,968	0,969	0,970
2 d.f.	6,44	6,5	6,54	6,59	6,65	6,70	6,76	6,82	6,88	6,95	7,01
3 d.f.	8,31	8,37	8,42	8,48	8,54	8,61	8,67	8,74	8,80	8,87	8,95

Answers: a) Chi-square test for the goodness of fit of a uniform distribution.
b) Yes, the conditions to apply the test are verified. c) Reject H_0. d) P-value (Sig.) = 0.039.

(2^{nd} test 4^{th} June 2012)

23. A survey was conducted with the main objective of analysing the intention of recently graduate students to follow a master course. The following results were obtained from the SPSS analysis.

Analysis A) Interviewees were asked about the possibility of following a master course within 2 or 3 years and the answers were given in a scale from "5 = not probable at all" to "1 = absolutely sure".

Ranks

	Sex	N	Mean Rank	Sum of Ranks
Possibility of following a master degree	0 Females	102	119,06	12144,00
	1 Males	119	104,09	12387,00
	Total	221		

Test Statistics[a]

	Possibility of following a master degree
Mann-Whitney U	5247,000
Wilcoxon W	12387,000
Z	-1,846
Asymp. Sig. (2-tailed)	,065

a. Grouping Variable: Sex

a) Identify the null and alternative hypotheses in the previous test.

b) Do you consider the previous test adequate to this type of data? If not, please point out an alternative test.

c) Which is the decision to be made for an α = 0,01?

Analysis B) Variable Age (in years) was also analysed for those who admitted some possibility of following a master course in the next 2/3 years.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Age	86	25,37	7,64	,82

One-Sample Test

	Test Value = 25					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Age	,451	85	,653	,37	-1,27	2,01

d) Which are the assumptions for the previous test? Are these assumptions verified?

e) What can one conclude about the age of those who intend to follow a master course in the next 2/3 years?

f) Construct a 95% confidence interval for the mean population age of the variable being studied and interpret the results.

(1st Exam 29h May 2007)

Answers:

a) H_0 : the distribution of the variable "Possibility of following a master degree" is equal for the two population groups of graduate students, females and males.
H_1 the distribution of the variable "Possibility of following a master degree" is different for the two population groups of graduate students, females and males.

b) Yes, it is an ordinal (qualitative) variable and we want to compare the distribution of two population groups using two independent samples.

c) Not reject H_0.

d) $H_0 : \mu = 25$
$H_1 : \mu \neq 25$
The t-test for one population mean assumes that the population distribution is normal; however, we do not need to test this assumption because the t-test statistic is valid when the sample dimension is larger than 30, even if the population distribution is not normal.

e) Do not reject H_0, that is, do not reject that the population mean age of those graduates who admitted some possibility of following a master course, is 25 years.

f) $\left]IC_{0,95}\right[_{\mu} = \left]23,76 \quad ; \quad 26,98\right[$

24. In a study about breast cancer a sample of 1207 women aged between 22 e 88 years was analysed and the following variables were measured: "Number of lymph nodes", "Survival (months)" and "Size of tumor (cms)".

a) Calculate the values of (a), (b) e (c).

b) Define the null and alternative hypothesis for the three tests presented in this analysis and, for each, identify the decisions to be made (α=0,05).

c) Which are the assumptions for the application of ANOVA? Are they verified? If not, present adequate alternative test(s).

d) Which of the multiple comparison tests presented is the most appropriate for this case? Why?

e) Identify, briefly, the main conclusions from all the results presented here. (1st Exam 29h May 2007)

Descriptives

Pathologic Tumor Size (cm)

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Less than 40	121	2,23	1,13	**(a)**	2,02	2,43	,26	7,00
40 - 54 years	394	1,85	1,02	,05	1,75	1,95	,10	5,50
55 - 69 years	394	1,63	,94	,05	1,53	1,72	,15	6,50
70 or more	212	1,43	,81	,06	1,32	1,54	,30	5,20
Total	1121	1,73	1,00	,03	1,68	1,79	,10	7,00

Test of Homogeneity of Variances

Pathologic Tumor Size (cm)

Levene Statistic	df1	df2	Sig.
5,712	3	1117	,001

ANOVA

Pathologic Tumor Size (cm)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	59,016	**(b)**	**(c)**	20,893	,000
Within Groups	1051,722	1117	,942		
Total	1110,738	1120			

Multiple Comparisons

Dependent Variable: Pathologic Tumor Size (cm)

	(I) Age group	(J) Age group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Scheffe	Less than 40	Less than 40					
		40 - 54 years	,37440*	,10085	,003	,0920	,6568
		55 - 69 years	,60136*	,10085	,000	,3190	,8837
		70 or more	,79581*	,11056	,000	,4863	1,1053
	40 - 54 years	Less than 40	-,37440*	,10085	,003	-,6568	-,0920
		40 - 54 years					
		55 - 69 years	,22695*	,06913	,013	,0334	,4205
		70 or more	,42141*	,08265	,000	,1900	,6528
	55 - 69 years	Less than 40	-,60136*	,10085	,000	-,8837	-,3190
		40 - 54 years	-,22695*	,06913	,013	-,4205	-,0334
		55 - 69 years					
		70 or more	,19445	,08265	,137	-,0369	,4259
	70 or more	Less than 40	-,79581*	,11056	,000	-1,1053	-,4863
		40 - 54 years	-,42141*	,08265	,000	-,6528	-,1900
		55 - 69 years	-,19445	,08265	,137	-,4259	,0369
		70 or more					
Dunnett C	Less than 40	Less than 40					
		40 - 54 years	,37440*	,11531		,0746	,6742
		55 - 69 years	,60136*	,11348		,3062	,8965
		70 or more	,79581*	,11722		,4908	1,1008
	40 - 54 years	Less than 40	-,37440*	,11531		-,6742	-,0746
		40 - 54 years					
		55 - 69 years	,22695*	,06999		,0464	,4075
		70 or more	,42141*	,07591		,2252	,6176
	55 - 69 years	Less than 40	-,60136*	,11348		-,8965	-,3062
		40 - 54 years	-,22695*	,06999		-,4075	-,0464
		55 - 69 years					
		70 or more	,19445*	,07310		,0054	,3835
	70 or more	Less than 40	-,79581*	,11722		-1,1008	-,4908
		40 - 54 years	-,42141*	,07591		-,6176	-,2252
		55 - 69 years	-,19445*	,07310		-,3835	-,0054
		70 or more					

*. The mean difference is significant at the .05 level.

Answers:

a) (a) = 0,103; (b) = 3; (c) = 19,672

b) – Levene test for equality of 4 population variances: reject H_0, so, at least two population groups have different variances for the variable "Tumor size";
- ANOVA (test for equality of 4 population means): reject H_0, so, at least two population groups have different means for the variable "Tumor size";
- Dunnett C multiple comparison test for equality of population means two by two: reject H_0 for all comparisons of thwo population means, so there are significant differences on mean tumor sizes for the 4 women age groups; from the sample results we can conclude that the mean size decreases with age.

c) The assumption of equal variances for the 4 population groups is violated since the Sig. value of the Levene test is lower than the 5% significance level.

d) Dunnett C test for the equality of population means, two by two, since the assumption of equal variances for the 4 population groups is being violated.

25. Company XX imports two different types (A and B) of instruments to measure blood pressure with a mean duration of 500 and 400 hours, respectively. It is known that the duration of these instruments follows a normal distribution with a dispersion of 81 hours (equal for both type A and type B instruments). The last order company XX received had no identification but the manager believes that the package contains Type A instruments. In order to identify if the package is type A, the manager decides to take a sample of 9 instruments and to classify the package according to the results of an hypothesis statistical test using a significance level of 0,05.

a) For this specific problem, identify the type of errors that can be made by Company XX manager and calculate their probabilities.

b) If the mean duration of the 9 tested instruments is 436 hours, which should be the decision made by the manager?

c) If the significance level is 0,01, what is result for β ? (1st Exam 29 th May 2007)

Answers:

a) Error type I: classify the order as being type B when in fact is type A
Maximum probability of error type I = 0,05
Error type II: classify the order as being type A when in fact is type B
Probability of error type II when μ_A=400 $\Rightarrow \beta(\mu_A=400) = 0{,}0197$

b) Reject H_0, that the order is Type A.

c) $\beta(\mu_A=400) = 0{,}0838$.

26. The manager of *SEMPRE FRIO, SA* is evaluating the results of an advertising campaign to introduce the new product *"Fever for caramel"*. The manager believes that this campaign wasn't successful because, in his opinion, the percentage of consumers of "*Fever for caramel*" continues to be at most 55%. However, the marketing manager argues that, according to a market survey that questioned 400 individuals, only 152 stated that they never bought this ice cream.

a) Which of these two directors seems to be right? ($\alpha = 0{,}05$)

b) If the marketing director is right and if the percentage of consumers of *"Fever for caramel"* is, in fact, 60%, what is the probability of making a wrong decision?

Answers: a) The marketing manager seems to be right; b) 0. (2nd Exam 30 th June 2006)

27. For data from a normal population and a random sample of dimension 16, many hypothesis tests were applied. If $\alpha = 0{,}01$, which values for the corrected sample variance correspond to the critical area of the following test:

$$H_0 : \sigma^2 \geq 81$$
$$H_A : \sigma^2 < 81$$

Suppose now that the sample variance is 60; which decision should be made?

Answers: The interval $[0\,;28.242]$; Not reject H_0. (2nd Exam 30 th June 2006)

28. The owner of Bar Oceano, very popular among the students of Atlas University, suspects that his female clients tend to spend less than male clients. Challenged by the professor of Statistics of this University, he wants to find an answer to this challenge. This professor, which usually seats at a very discreet place at the bar, decides to register the spending of 17 clients (randomly selected) with the objective of performing an hypothesis test. The results in euros are shown below.

Female clients	5,12	1,50	1,75	3,72	2,12	3,15	6,10	3,42	4,00
Male clients	5,80	6,50	2,80	5,15	1,30	1,40	1,80	6,40	

Without making any assumptions about these populations, help the owner to perform the adequate test to this data. Use a significance level of 5% and find the solution to the question (remember that the owner doesn't know anything about Statistics!).

Answer: Mann-Whitney test for equality of population distributions of female and male clients spending; the decision is not reject the null hypothesis, so there is no reason to believe that female clients spend less than male clients.

29. In a study carried out with the objective of analyzing Company AX workers satisfaction, several results were obtained, using a continuous scale from -3 = completely unsatisfied to 3 = completely satisfied:

Group Statistics

	Sex	N	Mean	Std. Deviation	Std. Error Mean
Remuneration	Male	44	-,110	1,037	,156
	Female	44	,118	,979	,148

Independent Samples Test

		Statistics								
		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
Dependent variables	Assumptions	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Remuneration	Equal variances assumed	,616	,435	**(a)**	86	,292	-,228	,215	-,655	,200
	Equal variances not assumed			-1,059	85,7	,292	-,228	,215	-,655	,200

a) Identify the variables under study and interpret the results in the first table.
b) Calculate the value of (a) and interpret its meaning.
c) Identify the null and alternative hypothesis of all the tests previously presented.
d) Which are the decisions to be taken with an α = 0,05? Why?
e) Which are the necessary assumptions to apply the previous tests? Are they verified? Why?

(1st Exam 31 st May 2008)

Answers:

b) (a) = −1.059 c) Levene test for equality of 2 population variances (male and females) for variable Remuneration; t-test for equality of means for variable Remuneration for male and female populations;
c) For Levene test do not reject H_0, so, the two population groups have equal variances for the variable "Remuneration"; for the t-test do not reject H_0, so, the two population groups have equal means for the variable "Remuneration"; e) both tests assume independent samples and normal population; there is no need to test for population normal distributions since both samples have dimensions larger than 30.

30. The human resources manager of Company AX says that, annually, at least 20% of female workers apply for maternity leave. From the 44 female sampled workers, 4 had maternity leave during the last year. To evaluate if the human resources manager is right, you should apply a statistical hypothesis test.

a) Define the null and alternative hypothesis and any necessary assumptions and conditions to apply this test.
b) Identify the decision to be made (α=0,05).
c) Identify, for this specific problem, the type of error that can be made.
d) Calculate the probability of making the previous type of error if, in fact, the human resources manager is mistaken and the value of the parameter is 10%.

(1st Exam 31 st May 2008)

Answers: a) $H_0 : p \geq 0.20$

$H_1 : p < 0.20$

b) Value of the test statistic = - 1.81 < - 1.645. Reject H_0, so the human resources manager does not seem to be right.

c) Error type I: reject the hypothesis that the population proportion of female workers that annually apply for maternity leave is at least 20% and, in fact, this hypothesis is true. d) 0.05.

31. In the previous study, workers were asked to give their opinion about the "Work environment" (5 point agreement scale (Likert): 1=Disagree completely, 2=Disagree, 3=Neither disagree nor agree, 4=Agree, 5=Agree completely). The following results were obtained:

Ranks

	Working area	N	Mean Rank
Work environment in this. company is, in general, healthy and pleasant	Traffic	46	42,50
	Handling	22	40,05
	Other areas	18	50,28
	Total	86	

Test Statistics(a,b)

	Work environment in this company is, in general, healthy and pleasant.
Chi-Square	2,023
df	2
Asymp. Sig.	,364

a Kruskal Wallis Test
b Grouping Variable: Working area

a) Define the null and alternative hypothesis for the previous test.
b) Identify the decision to be made (α=0,05) and the relevant conclusions.
c) From all the statistical tests you have learned, which can be considered as an alternative to the previous one? For this example, which one should be preferred? Why?

(1st Exam 31 st May 2008)

Answers:

a) H_0 : the distribution of the variable "Work environment in this company is, in general, healthy and pleasant" is equal for the three population groups of workers from the 3 areas.
H_1 at least two population groups of workers have different distributions for the variable "Work environment in this company is, in general, healthy and pleasant".
b) Do not reject H_0 that the distribution of the variable "Work environment in this company is, in general, healthy and pleasant" is equal for the three population groups of workers from the 3 areas.
c) ANOVA test for equality of 3 population means. Since the variable "Work environment in this company is, in general, healthy and pleasant" is ordinal, Kruskal-Wallis test should be preferred because it does not need the assumption of populations normality and 2 of the sample dimensions are lower than 31.

32. One survey of U.S. consumers conducted by the Wall Street Journal and NBC News asked the question: “In general, how would you rate the level of service that American businesses provide?” The distribution of responses to this question was as follows:

Excellent	8%
Pretty good	47%
Only fair	34%
Poor	11%

Suppose that you want to find out whether the results of this consumer survey apply to customers of supermarkets in your city. To do so you interview 207 randomly selected consumers as they leave supermarkets in various parts of the city. You ask the customers how they would rate the level of service at that supermarket. The observed responses from this study are given in the following table:

Response	Frequency
Excellent	21
Pretty good	109
Only fair	62
Poor	15

Determine whether the observed frequencies of responses from this survey are the same as the frequencies that would be expected based on the national survey. Use α=0.05.

Answer: The calculated value of chi-square is 6.25 versus the critical value of 7.815. Because the calculated chi-square is not greater than the critical chi-square value, we do not reject the null hypothesis. Thus, the data gathered in the sample of 207 supermarket shoppers indicates that the population distribution of responses of supermarket shoppers in your city is not significantly different from the distribution of responses to the US survey.

33. Quite often in the business world, random arrivals are Poisson distributed. This distribution is characterized by an average arrival rate, λ, per some interval. Suppose that a teller supervisor believes that the distribution of random arrivals at a local bank is Poisson and sets out to test this hypothesis by gathering information. The following data represent a distribution of frequency of arrivals during one-minute intervals at the bank. Use α=0.05 to test these data in an effort to determine whether they are Poisson distributed.

Number of Arrivals	Observed Frequencies
0	7
1	18
2	25
3	17
4	12
≥5	5

Answer: The calculated value of the test statistic (1.74) is not greater than the critical chi-square value of 9.488, so the supervisor’s decision should be not reject the null hypothesis. That is, there is no evidence to reject the hypothesis that the population distribution of bank arrivals per minute follows a Poisson distribution.

34. Is the type of beverage ordered with lunch at a restaurant independent of the age group of the consumer? A random poll of 309 lunch consumers is taken, resulting in the following contingency table of observed values. Use α=0.01 to determine whether the two variables are independent.

Age group	Coffee/Tea	Soft Drink	Other (Milk, etc.)
21-34	26	95	18
35-55	41	40	20
>55	24	13	32

Answer: The calculated value of the test statistic, 59.41, is greater than the critical value, 13.277 (p-value=0.000), so the null hypothesis is rejected. The two variables are not independent in the population from where we took this sample. The type of beverage that a consumer orders with lunch is related to age. Examination of the sample categories reveals that younger people tend to prefer soft drinks, and older people prefer other types of beverages.

35. An appliance manufacturer obtained the data shown in the table below in a preliminary phone survey designed to evaluate its service nationwide.

Concept of Service	Location		
	East	Midwest	West
Poor	14	16	10
Good	16	14	30

a) What is the expected number of consumers in the West who have a good concept of the firm's service (assuming independence)?

b) What is the number of degrees of freedom for testing the null hypothesis H_0: Location and concept of service are independent?

c) Suppose that the value for the usual statistics that is computed for this type of problem is 6.53. Do you reject the hypothesis given in part b) at the 5% level of significance? Why or why not?

Answer: a) 24 b) 2 c) Yes, reject H_0 because the calculated test statistic value, 6.53, is greater than 5.991, the critical value (p-value=0.038).

36. Do construction workers who purchase lunch from street vendors spend less per meal than construction workers who go to restaurants for lunch? To test this question, a researcher selected two random samples of construction workers, one group that purchases lunch from street vendors and one group that purchases lunch from restaurants. Use the following data and a Mann-Whitney U test to determine whether street vendor lunches are significantly cheaper than restaurant lunches. Let α=0.01.

Vendor	Vendor	Restaurant	Restaurant
$2.75	4.01	$4.10	2.70
3.29	3.68	4.75	3.65
4.53	3.15	3.95	5.11
3.61	2.97	3.50	4.80
3.10	4.05	4.25	6.25
4.29	3.60	4.98	3.89
2.25		5.75	4.80
2.97		4.10	5.50

Answer: As the calculated test statistic value is greater than the critical value, we reject the null hypothesis, thus the population distribution of vendor lunches cost is different from population distribution of the cost of restaurant lunches.

37. Two potential suppliers of street lighting equipment – supplier A and supplier B – have included data on life tests of their streetlights (in months) in bids supplied to a city manager. The manager wishes to test the equality of the two population distributions using rank sums. Are the distributions of lifetime of the streetlights equal?

Supplier	Lifetime of Streetlights (months)
A	35, 66, 58, 46, 42, 40, 49, 59, 32, 58, 68, 41, 29, 75, 30, 53, 60, 63, 58, 56
B	48, 47, 37, 47, 50, 34, 71, 36, 56, 45, 43, 28, 47, 40, 33, 51, 42, 53, 34, 29

Answer: As the calculated Z value (1.85) is less than the critical value (1.96), we don't reject the hypothesis of equal population distributions at $\alpha=0.05$. The sample difference could be due to chance.

38. Suppose that a researcher is interested in determining whether the number of physicians in an office produces significant differences in the number of office patients seen by each physician per day. She takes a random sample of physicians from practices in which

1) there are only two partners,
2) there are three or more partners, or
3) the office is a health maintenance organization (HMO).

The following table shows the data obtained. Let $\alpha=0.01$.

Two Partners	Three or More Partners	HMO
13	24	26
15	16	22
20	19	31
18	22	27
23	25	28
	14	33
	17	

Answer: Because the value of the test statistic 9.56 is larger than the critical value, the researcher rejects the null hypothesis. The population number of patients seen in the office by a physician is not the same in these three sizes of offices. Examination of the samples values in each group reveals that physicians in two-partner offices see fewer patients per physician in the office, and HMO physicians see more patients in the office.

39. Universal Build is a building insurance company operating at national level and its marketing manager believes that the ongoing strategy will have a significant effect on the mean value of net results that will reach at least 4500 monetary units (m.u.) per insurance policy. However, the other company managers do not agree with this statement, considering it too optimistic and not reliable. Assume that the net results per insurance policy follow a normal distribution with a standard deviation of 625 m.u.

a) Is the marketing manager right? Answer this question by performing an adequate hypothesis test with 5% significance level and by knowing that a random sample of 25 contracts shows a mean net result of 4320 m.u. per policy.
b) For this specific problem, identify the type of error that can be made with the decision taken in question a).
c) Which values of the significance level would lead to a different decision from the one made in question a)?
d) If the real value of the mean net results per policy is 4345,625 m.u., which is the probability of making a wrong decision in question a)?
e) The company managers consider the probability obtained in d) too high and want to reduce this probability to 0.10, keeping the 5% significance level. Which should be the sample dimension to make sure these two probabilities are verified?

(1st Exam 1st June 2013)

Answer: a) $H_0 : \mu \geq 4500$
$H_1 : \mu < 4500$ $t = -1{,}44$; Decision: do not reject H_o
b) Error Type II: Do not reject the real mean net result per policy is at least 4500 m.u. when in fact is lower
c) p-value = 0,0749; values of $\alpha \geq 0{,}0749$ would lead to a different decision d) 0.6591 e) $n \geq 141$

40. Attitudes toward a U.S: Treasury Department proposal for a change in the existing tax rate brackets have been assessed for three groups: tax accountants, university accounting faculty, and IRS tax auditors, which scores are presented in the following table. A score of 100 means perfect agreement, whereas 0 means complete disfavour. Is there a real difference among the groups with respect to attitude toward the proposed change? Let α=0.05.

Tax Accountants	Accounting Professors	IRS Tax Auditors
18	93	39
33	21	96
36	27	93
60	33	90
72	48	84
63	62	87
54	51	75
45	32	78
42	66	78
30	69	81
24	71	79
41	68	61

Answer: Since the critical value is 5.99, we reject the hypothesis of agreement on this issue. The distribution average (rank) opinion is not the same for these groups. By the rank-sums, the IRS tax auditors favour the proposal to a higher degree (have lower ranks) than either of the other groups. More testing is required to determine whether the tax-accountant and accounting-professor groups differ in their opinion of the proposed change.

41. A market research company has surveyed a random sample of 206 consumers to know their use and buying behavior on perfumes and eaux-de-toilette. The following results were obtained from a statistical analysis of the collected data.

a) Calculate the values (a), (b), (c) e (d) drawn from Tables 1 and 2 and describe their meaning.

b) Identify the null and alternative hypotheses being tested and the decision to be made for a value of α=0,05.

Table 1: Do you use any perfume at present? * Sex Crosstabulation

			Sex		Total
			Male	Female	
Do you use any perfume at present?	No	Count	11	7	18
		Expected Count	**(a)**	10,7	18,0
		% within Do you use any perfume at present?	61,1%	38,9%	100,0%
		% within Sex	13,3%	**(c)**	8,7%
		% of Total	5,3%	3,4%	8,7%
	Yes	Count	72	116	188
		Expected Count	75,7	112,3	188,0
		% within Do you use any perfume at present?	**(b)**	61,7%	100,0%
		% within Sex	86,7%	94,3%	91,3%
		% of Total	35,0%	56,3%	91,3%
Total		Count	83	123	206
		Expected Count	83,0	123,0	206,0
		% within Do you use any perfume at present?	40,3%	59,7%	100,0%
		% within Sex	100,0%	100,0%	100,0%
		% of Total	40,3%	59,7%	100,0%

Table 2: Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	**(d)**[a]	1	,059

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 7,25.

(1st Exam 1st June 2013)

Answer: a) (a) = 7.3 (b) = 38,3% (c) = 5,7% (d) = 3,4575

b) H_0 : In the population, the use of a perfume at present is independent from sex

H_1 : In the population, the use of a perfume at present is related to categories of sex

Decision: do not reject H_0 because the p-value is bigger than α.

42. TDHIF Portugal sells a high definition TV brand, leader in its market segment. Lately, some clients have been complaining that some TV sets are defective. The Portuguese manager has informed the producer, but the last argued that the percentage of defective sets is at most 1%, so there is no reason to worry. The Portuguese manager wants to test this statement, by counting the number of defective sets in the next 100 to be received from the producer.

a) Identify the hypotheses being tested and the errors that can be made in the context of this specific case.

b) Knowing that from the next 100 received TV sets only 2 are defective, do you consider the producer is right? Use a significance level of 1%.

c) From the 100 received TV sets how many should be defective to lead to a decision opposite to the one made in the previous question?

d) What is the probability of making a wrong decision in question b) if the true percentage of defective TV sets is 2%? (1st Exam 7th June 2014)

Answer: a) $H_0 : p \leq 0.01$

$H_1 : p > 0.01$

b) Decision: do not reject H_0 because the value of the test statistic belongs to the acceptance region; the producer might be right

c) ≈ 4 (3.31)

d) 0.8264

43. In a recent study about the use of public transports, a random sample of 840 individuals aged 15 or more years and resident in the area of Lisbon were asked to evaluate the global service (continuous scale from 1 = no satisfaction at all to 10 = completely satisfied) offered by the different types of transports (bus, train and metro); the following results were obtained.

a) Identify the analysis carried out in the previous tables and its objective.

b) Calculate the values of **(a)**, **(b)** and **(c)** missing from Table 4.

c) Identify the hypotheses tested in Tables 1 and 2 and the decisions to be made in the context of this specific case. Use α=0.05.

d) Identify the hypotheses tested in Table 4 and the decision to be made in the context of this specific case. Use α=0.05. Justify your answer by considering the results of Table 5 and pointing out a possible value for **(d)**.

e) Which are the necessary conditions to apply the previous test? Are they verified? Why?

f) Which of the tests presented in Table 5 is the most adequate for this specific case? Interpret its results.

(1st Exam 7th June 2014)

Table 3: Descriptives

Overall evaluation of the service	N	Mean	Std. Deviation	Std. Error
Bus	400	6,39	1,609	,080
Train	163	6,68	1,578	,124
Metro	277	7,27	1,335	,080
Total	840	6,73	1,565	,054

Table 1: Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Overall evaluation of the service	,167	840	,000	,933	840	,000

a. Lilliefors Significance Correction

Table 2: Test of Homogeneity of Variances

Overall evaluation of the service

Levene Statistic	df1	df2	Sig.
4,313	2	837	,014

Table 4: ANOVA

Overall evaluation of the service	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	127,219	**(a)**	63,610	**(c)**	**(d)**
Within Groups	1928,580	837	**(b)**		
Total	2055,799	839			

Table 5: Multiple Comparisons

Dependent Variable: Overall evaluation of the service

	(I) Type of transport	(J) Type of transport	Mean Difference (I-J)	Std. Error	Sig.
Scheffe	Bus	Comboio	-.293	.141	.115
		Metro	-.880*	.119	.000
	Train	Autocarro	.293	.141	.115
		Metro	-.586*	.150	.001
	Metro	Autocarro	.880*	.119	.000
		Comboio	.586*	.150	.001
Dunnett C	Bus	Comboio	-,293	,147	
		Metro	-,880*	,114	
	Train	Autocarro	,293	,147	
		Metro	-,586*	,147	
	Metro	Autocarro	,880*	,114	
		Comboio	,586*	,147	

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

Answer: a) Oneway ANOVA with test for homogeneity of variances, test of normality and multiple comparison tests

b) (a) = 2, (b) = 2.304, (c) = 27.61

c) K-S test of normality: Reject H_0; Levene test of equality of variances: Reject H_0

d) (d) ≈ 0; test for equality of 3 population means: Reject H_0, so at least two population groups have different means. f) Dunnett C

44. In a recent market research study about public transports, a random sample of 473 users aged 15 or more years and resident in the area of Lisbon were asked to evaluate the quality of service and price (continuous scale from 1 = no satisfaction at all to 10 = completely satisfied) offered by the different types of transports and the following results were obtained:

Table 1

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Satisfaction with: Quality of service	Male	220	4,8767	,89208	,06014
	Female	253	4,1663	1,02098	,06419
Satisfaction with: Price	Male	220	3,4229	1,07769	,07266
	Female	253	3,2165	1,19506	,07513

Table 2

		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
Satisfaction with: Quality of service	Equal variances assumed	3,405	,066	8,000	471	,000	,71032	,08879
	Equal variances not assumed			8,075	470,988	,000	,71032	,08796
Satisfaction with: Price	Equal variances assumed	,783	,377	1,960	471	,051	,20632	,10527
	Equal variances not assumed			1,974	470,367	,049	,20632	,10452

a) Identify the variables under analysis in the previous tables.

b) It is possible to conclude that men are more satisfied with price of transports than women? Identify the hypotheses being tested and the decisions to be made in the context of this specific case. Use α=0.05.

c) And the satisfaction with quality of service? Would it be equal for both gender groups? Use α=0.10.

d) Which are the necessary assumptions to apply the tests in the previous questions? Are they verified? Justify your answer and formulate the hypothesis needed to be tested to verify these assumptions.

(1st Exam 6th June 2015)

Answers: a) X_1 : Degree of satisfaction with quality of service (continuous scale from 1 = no satisfaction at all to 10 = completely satisfied);
X_2: Degree of satisfaction with price (continuous scale from 1 = no satisfaction at all to 10 = completely satisfied);
Gender (qualitative nominal variable with 2 categories: M=Male and F=Female)

b) $H_0 : \mu_{2M} \leq \mu_{2F}$
$H_1 : \mu_{2M} > \mu_{2F}$

Decision: reject H_0 because [Sig/2=0,051/2] < [α=0,05]; i.e., it is possible that, on average, males have a higher level of satisfaction with price than females

c) $H_0 : \mu_{1M} = \mu_{1F}$
$H_1 : \mu_{1M} \neq \mu_{1F}$

Decision: reject H_0 because [Sig=0,000] < [α=0,10]; i.e., it is possible that, on average, males and females have a different level of satisfaction with quality of public transport services

d) The previous hypothesis tests assume that the two variables (X_1 and X_2) follow a normal distribution, both for the male and the female population groups; no results about normality tests are presented but, even if the population groups do not follow a normal distribution, the previous tests are valid because both sample dimensions are bigger than 30 ($n_1 = 220$ and $n_2 = 253$)

45. A sample of 445 workers from different companies of a certain industrial sector was surveyed and the following outputs were obtained:

a) Is it possible to conclude that workers monthly income is related to gender, in this specific activity sector? Identify the hypotheses being tested, the test statistic and the decision to be made at a 5% significance level.

b) Calculate the values of (a), (b), (c) and (d) missing from Table 1 and justify your answer.

Table 1: Gender * monthly income Crosstabulation

			Monthly income			Total
			< 1200 €	1200 € - 2400 €	2401 € - 4000 €	
Gender	Male	Count	109	55	44	208
		Expected Count	119,7	**(a)**	**(b)**	208
	Female	Count	147	61	29	237
		Expected Count	136,3	**(c)**	**(d)**	237
Total		Count	256	116	73	445
		% of Total	57,5%	26,1%	16,4%	100,0%

Table 2: Chi-Square Test

	Value	df	Asymp. Sig.
Pearson Chi-Square	7,174[a]	2	,028
N of Valid Cases	445		

a. 0 cells (0, 0%) have expected count less than 5. The minimum expected count is 34,1.

(1st Exam 6th June 2015)

Answer: a) X: Gender of workers from a certain industrial sector (Qualitative nominal variable with 2 categories: 1=Male; 2=Female);

Y: Monthly income of workers from a certain industrial sector (Qualitative ordinal with 3 categories: 1 = <1200€; 2 = 1200€ - 2400€; 3 = 2401€ - 4000€);

H_0 : in the population , workers monthly income is independent from workers gender

H_1 : in the population , workers monthly income is related to workers gender

Test statistic: $\sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \cap \chi^2_{(c_1-1)\times(c_2-1)}, \quad with \quad i = 1,2 \; e \; j = 1,2,3, \quad c_1 = 2, \quad c_2 = 3$

Decisão: Decision: reject because [Sig/2=0,028] < [α=0,05]; i.e., it is possible to conclude that, in the population of workers of that industrial sector, monthly income is related to gender; from the sample results it is possible to conclude that males are more associated with higher levels of monthly income while females associate more with lower categories of monthly income

b) (a) = 54,2 (b) = 34,1 (c) = 61,8 (d) = 38,9

46. The manager of the national newspaper *Moon & Sun* has introduced some changes in the newspaper format and design and he is now convinced that these changes have made the market share of the newspaper increase to, at least, 50%. To evaluate if he is right, a sample of 120 readers was surveyed about their newspaper preferences and the following sample result was obtained:

$$\sum_{i=1}^{120} X_i = 57$$

a) Define properly the population under study.
b) What can you conclude about the statement of the newspaper manager? Use a significance level of 5%.
c) What is the probability a correct decision was made in the previous question? To answer this question, consider as true the minimum value of the parameter defined under H_0.
d) Suppose now a second sample of the same dimension (n=120) was collected and the following sample result was obtained:

$$\sum_{i=1}^{120} X_i = 42$$

d_1) Which is the decision to be made using the same significance level? Justify your answer.

d_2) Which is the p-value (*Sig.*) associated to the value of test statistic of question d_1)?

(1st Exam 6th June 2015)

Answers: a) X: number of readers, in 1, that will prefer Moon & Sun newspaper

$X \cap B(p)$ ou $X \cap b(n=1; p)$

b) $H_0 : p \geq 0,50$

$H_1 : p < 0.50$

Value of the test statistic = – 0,548 > –1.645; the decision is not Reject H_0, so the newspaper manager might be right

c) P [not reject H_0 | H_0 is true] =
1 – P [reject H_0 | H_0 is true] =
1 – P [test statistic value ≤ –1.645 | p = 0,50] = 1 – 0,05 = 0,95

d_1) With $\bar{x}^* = 0{,}35$, the value of the test statistic becomes –3,29, and the decision is now to Reject H_0, i.e., the newspaper manager is no longer right

d_2) $\alpha \leq 0{,}0005$

47. A survey was carried out with the objective of understanding the opinions and attitudes of the youth about nature. A sample of 202 young people from different regions was asked about the importance given to "Contact with nature" using a continuous scale from 1=No importance at all to 10=Extremely important. One of the objectives of this study was to investigate if the mean importance was different according to different levels of education. Some of the results are shown in the following tables.

Table 1 - Descriptives
Contact with Nature

	N	Mean	Std. Deviation	Std. Error
Primary	31	4.35	1.47	.26
Secondary	82	4.99	1.45	.16
High	89	5.69	1.37	.14
Total	202	5.20	1.49	.10

Table 2- ANOVA
Contact with Nature

	Sum of Squares	Df	Mean Squares	F	Sig.
Between Groups	47.842	2	**(a)**	**(b)**	.000
Within Groups	401.581	199	2.018		
Total	449.423	201			

a) Define the hypotheses being tested according to the objectives of the study.
b) Calculate the values of **(a)** and **(b)**.
c) Which are the necessary assumptions to apply the test in the previous questions?
d) Assuming the assumptions of the previous analysis are verified, what can be concluded from all the presented results? Consider α=0.05.
e) Taking into account the previous conclusion, which additional analysis would you suggest to be made? Justify your answer.

(1st Exam 20th June 2015)

Answers : a) X_1 : Degree of importance given to Contact with Nature (continuous scale from 1= No importance at all to 10=extremely important) by youth with primary education

X_2: Degree of importance given to Contact with Nature (continuous scale from 1= No importance at all to 10=extremely important) by youth with secondary education

X_3: Degree of importance given to Contact with Nature (continuous scale from 1= No importance at all to 10=extremely important) by youth with high education

$H_0 : \mu_1 = \mu_2 = \mu_3$

$H_1 : \exists\ \mu_i \neq \mu_j$ with $i,j = 1,2,3$ and $i \neq j$

b) (a) = 23,921 (b) = 11,854 = F

c) The hypothesis test in table 2 assumes the three population groups (youth with primary, secondary and high education) follow a normal distribution for the importance given to Contact with Nature; assumes the population variances of the three groups for the variable Contact with Nature are equal; and that the samples drawn from the three populations are independent.

d) Decision: reject H_0 because [Sig=0,000] < [α=0,05]; i.e., at least two of the three population groups give different mean importance to Contact with Nature

e) Yes, Scheffé multiple comparison test in case the assumption of homogeneity of variances is confirmed.

48. According to demographic statistics, the population distribution of youth by region of residence is the following: North = 35%, Centre = 20%, Lisbon = 35%, Algarve = 10%. Consider the following results:

Table A: Frequencies
Region of Residence

	Observed N	Expected N	Residual
North	82	70.7	11.3
Centre	36	**(a)**	**(b)**
Lisbon	73	70.7	2.3
Algarve	11	**(c)**	**(d)**
Total	**202**		

Table B: Test Statistics

	Region of Residence
Chi-Square	**6.550**[a]
Df	**3**

a. **(e)** cells (**(f)** %) have expected frequencies less than 5. The minimum expected cell frequency is 20.2.

a) Can this sample of 202 youth be considered representative of the population in terms of region of residence? Define the hypotheses being tested, the test statistic and the decision to be made for α=0.05.
b) Calculate the values of **(a), (b), (c), (d), (e)** and **(f)** from Tables A and B, and justify your answer.
b) Are the necessary conditions for the application of this test verified? Justify.

(1st Exam 20th June 2015)

Answers: a) X: Region of residence of youth (Qualitative nominal variable with 4 categories: 1=North; 2=Centre; 3=Lisbon; 4=Algarve);

H_0 : in the population the distribution of youth by region of residence is the following:

x_i	1	2	3	4
$f(x_i)$	0,35	0,20	0,35	0,10

H_1 : in the population the distribution of youth by region of residence is different from f(x)

Test statistic: $\sum_i \frac{(O_i - E_i)^2}{E_i} \cap \chi^2_{(c-k-1)}, \quad with \quad i = 1,2,3,4, \quad c = 4, \quad k = 0$

Decision: Value of the test statistic is 6,55 < 7,81; the decision is not reject H_0, i.e. this sample of 202 youth might be considered representative of the population in terms of regional distribution.

b) (a) = 40,4 (b) = –4,4 (c) = 20,2 (d) = –9,2 (e) = (f) = 0

c) Yes.

49. A credit institution states that the average number of days needed for approval of a personal loan is, at most, 15. The consumers association is doubtful about this number and decided to get a random sample of 225 consumers that asked for personal loans. The sample results show that consumers wait, on average, 18 days to get the answer from the institution. Assume the variable under study follows a normal distribution with standard deviation equal to 15 days.

a) Define properly the population under study.
b) Is the consumers association right? Use a significance level of 1%.
c) What is the probability of making an error in the decision made in the previous question?
d) What is the probability of making an error type II if, in reality, consumers have to wait, on average, 17 days for approval of a personal loan? Comment the result.

‘ (1st Exam 20th June 2015)

Answers: a) X: number of days waiting for a personal loan approval, $X \cap n(\mu, \sigma=15)$

b) $H_0 : \mu \leq 15$

$H_1 : \mu > 15$

Value of the test statistic = 3 > 2,326, the decision is Reject H_0, so the consumers association might be right

c) P [reject H_0 | H_0 is true] $\leq$ 0,01

d) P [not reject H_0 | $\mu_1 = 17$] = 0,6233