Statistics II: Quiz II (Example) - Parametric Tests.

1. *(2 points)* Which of the following statements corresponds to the concept of significance level:
   (a) Type I error, i.e., rejecting the null hypothesis when it is actually true;
   (b) Probability of rejecting the null hypothesis when it is false;
   (c) Type II error, i.e., not rejecting the null hypothesis when in reality it is false;
   (d) Probability of not rejecting the null hypothesis when it is false;
   (e) None of the other options is correct.

2. *(2 points)* In a study on satisfaction with public transports in the Lisbon Metropolitan Area, a sample of 473 users was asked about the quality of service (continuous satisfaction scale from 1 = *not at all satisfied* to 10 = *completely satisfied*). If you want to test whether the average level of satisfaction is at least 6, what assumptions should you check:
   (a) Only the assumption that the population from which the sample was taken follows a normal distribution;
   (b) Just the assumption that the sample is random;
   (c) Just the assumption that the population variance is known;
   (d) The assumptions that the sample is random and the population is normally distributed;
   (e) The assumptions that the sample is random, the population is normally distributed, and the population variance is known.

3. *(2 points)* In a study on satisfaction with public transports in the Lisbon Metropolitan Area, a sample of 65 users was asked about the price of tickets/passes (continuous satisfaction scale from 1 = *not at all satisfied* to 10 = *completely satisfied*). What is the test statistic if you want to test that the average level of satisfaction is the same for public transport users with and without a car (the population variances are not known but it is assumed that they will be different)?
   (a) $\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{(n_1-1)S_1'^2+(n_2-1)S_2'^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \overset{o}{\frown} N(0,\,1)$, any n_1 and n_2;
   (b) $\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{(n_1-1)S_1'^2+(n_2-1)S_2'^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \overset{o}{\frown} N(0,\,1)$, if $n_1 > 30$ and $n_2 > 30$;
   (c) $\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{S_1'^2/n_1+S_2'^2/n_2}} \overset{o}{\frown} N(0,\,1)$,if $n_1 > 30$ and $n_2 > 30$;
   (d) $\dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)_0}{\sqrt{\frac{(n_1-1)S_1'^2+(n_2-1)S_2'^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \frown t_{(n_1+n_2-2)}$;
   (e) None of the other options is correct.

4. *(2 points)* Thirty ISCTE students (random sample) were asked about the time they usually take between home and university. The following information was obtained:

$$\sum_{i=1}^{30} x_i = 990, \qquad \sum_{i=1}^{30} (x_i - \bar{x})^2 = 300.$$

A group that has been investigating this subject intends to test whether the average time taken by ISCTE students to go from home to university is less than 35 minutes. Considering a significance level of $\alpha = 0.05$ and that the home-university travel time of ISCTE students follows a $Normal$ distribution, define the null hypothesis (H_0) and the alternative hypothesis (H_1) of this statistical test:

(a) $H_0 : \mu < 35$ *vs* $H_1 : \mu \geq 35$;
(b) $H_0 : \mu \geq 35$ *vs* $H_1 : \mu < 35$;
(c) $H_0 : \mu = 35$ *vs* $H_1 : \mu \neq 35$;
(d) $H_0 : \mu \leq 35$ *vs* $H_1 : \mu > 35$;
(e) None of the other options is correct.

5. *(2 points)* Thirty ISCTE students (random sample) were asked about the time they usually take between home and university. The following information was obtained:

$$\sum_{i=1}^{30} x_i = 990, \qquad \sum_{i=1}^{30} (x_i - \bar{x})^2 = 300.$$

A group that has been investigating this subject, states that the mean and standard deviation of the time taken between home and university by ISCTE students are, respectively, 30 minutes and 10 minutes. Considering a significance level of $\alpha = 0.05$, comment on this group's statement regarding the variability of the home-university travel time of ISCTE students:

(a) The research group is right because $\bar{x} \in RR_{\bar{x}} =]\infty;\ 28,799] \cup [36,578;\ +\infty[$;
(b) The research group is not right because $\bar{x} \notin RR_{\bar{x}} =]\infty;\ 28,799] \cup [36,578;\ +\infty[$;
(c) The research group is right because the test statistic is $3 \in RR = [0;\ 16,047] \cup [45,722;\ +\infty[$;
(d) The research group is not right because the test statistic is $3 \notin RR = [42,557;\ +\infty[$;
(e) None of the other options is correct.

6. *(2 points)* In a study on satisfaction with public transports in the Lisbon Metropolitan Area, a sample of 473 users was asked about satisfaction with the price (continuous satisfaction scale from 1 = $not\,at\,all\,satisfied$ to 10 = $completely\,satisfied$). The following results were obtained to test whether female and male users have the same average degree of satisfaction:

Group Statistics

	Sexo	N	Mean	Std. Deviation	Std. Error Mean
Price satisfaction	Male	220	3,4229	1,07769	0,07266
	Female	253	3,21657	1,19506	0,07513

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
						Significance			
		F	Sig.	t	df	One - Sided p	Two - Sided p	Mean Difference	Std. Error Difference
Price satisfaction	Equal variances assumed	0,783	0,377	1,960	471	0,025	0,051	0,20632	0,10527
	Equal variances not assumed			1,974	470,367	0,025	0,049	0,20632	0,10452

Which of the following statements is correct regarding the decision to be made, considering a significance level of $\alpha = 0.05$:

(a) There is a significant difference between the sample means because $Two-Sided\ p = 0,049$;

(b) There is a significant difference between the population means because $Two-Sided\ p = 0,049$;

(c) There is no significant difference between the population means because $Two-Sided\ p = 0,051$.

(d) There is no significant difference between the population means because $Sig. = 0,377$;

(e) None of the other options is correct.

7. *(2 points)* In a study on satisfaction with public transports in the Lisbon Metropolitan Area, a sample of 473 users was asked about service quality and price (for the two variables, a continuous satisfaction scale of 1 = *not at all satisfied* to 10 = *completely satisfied* was used). If you want to test using SPSS software whether the average level of satisfaction is the same for service quality and price, which of the following procedures should you choose:

(a) Analyze → Compare means and proportions → One sample $t-test$;

(b) Analyze → Compare means and proportions → Independent-samples $t-test$;

(c) Analyze → Compare means and proportions → Paired-samples $t-test$;

(d) Analyze → Compare means and proportions → Paired-samples proportions;

(e) Analyze → Compare means and proportions → Independent-samples proportions.

8. *(2 points)* In a recent interview, the Minister of Health stated that, **the current situation has changed**, and at the moment more than half of the nursing professionals in public hospitals have permanent contracts. To test the Minister's assertion, the Portuguese Nurses Union collected a random sample of 412 nursing professionals from various public hospitals in the country, obtaining the following results.

One-Sample Proportions Test

		Observed						Significance	
Test Type		Successes	Trials	Proportion	Observed - Test Value[a]	Asymptotic Standard Error	Z	One - Sided p	Two - Sided p
Has Permanent contract=Yes	Wald	215	412	0,522	0,022	0,025	0,888	0,187	0,375

a. Teste t Value=0,5

Which of the following statements is correct?

(a) The Minister of Health is not right because $Two-Sided\ p = 0,375$;

(b) The Minister of Health is not right because $One-Sided\ p = 0,187$;

(c) The Minister of Health is right because $Two-Sided\ p = 0,375$;

(d) The Minister of Health is right because $One-Sided\ p = 0,187$;

(e) Nothing can be concluded because the assumptions of the test are not verified.

9. *(4 points)* A study was carried out on the ages of the populations of higher and secondary education teachers. A sample of 800 higher education teachers and 1000 secondary education teachers was collected. The means and variances were then calculated, obtaining the following values:

Group Statistics

Education	N	Mean	Std. Deviation
Higher	800	37,2	21,3448
Secondary	1000	35,9	20,8038

(a) Define the populations under study and formulate, justifying, the opposing hypotheses if you intend to compare the average ages of the two populations. $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

(b) Considering a significance level of $\alpha = 0.05$, find out whether the mean ages of higher and secondary education teacher populations are different (population variances are unknown, but believed to be different). p-value = 0.5569 → Do not Reject H_0

(c) Identify the two types of errors, in the context of this exercise, and say what error you may be committing, with the decision taken in the test conducted on (b).

(d) What is the probability of making a Type II error if in fact the difference of the means is $\mu_1 - \mu_2 = 3$?

$P[\text{ERROR Type II} \mid \mu_1 - \mu_2 = 3] \simeq 0.15$