

ESTADÍSTICA II
 EXERCISES CHAPTER 3
 ACADEMIC YEAR 2010/11

1. It is desired to evaluate the effect of the noise when a specific task is being carrying out. For that purpose, 32 people that are able to carry out the same task at the same time are randomly chosen in order to keep such a task in an environment under control and under two different kinds of background noise. From these 32 people 16 are chosen to carry out the task under background noise level 2. The remaining people will carry out it with low background noise level (level 1). The registered time (in minutes) for each individual is:

level 1	14	12	15	15	11	16	17	12	14	13	18	13	18	15	16	11
level 2	20	22	18	18	19	15	18	15	22	18	19	15	21	22	18	16

- a) Choose an appropriate test, specifying clearly the null hypothesis and alternative you are considering.
- b) Calculate the p -value of the test when normal distribution and equal variance hypotheses are considered.
2. It is desired to determine if unemployment in Madrid and Barcelona is different. For that purpose random samples of both cities are taken, each of them composed by 500 people, and the results show that 35 people were unemployed in Madrid and 25 in Barcelona. Is the any reason to think than unemployment frequencies are different in Madrid and Barcelona?
- a) Choose an appropriate test, specifying clearly the null hypothesis and alternative you are considering, as well as the working hypotheses.
- b) Calculate the p -value of the test.
3. Build a confidence interval at $1 - \alpha$ level for the difference of the means of two populations when variance equality can be assumed, $X \sim N(\mu_X, \sigma^2)$, $Y \sim N(\mu_Y, \sigma^2)$, and deal with two simple random samples $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ mutually independent. What is the width of the confidence interval?

Remark: Use as estimator of $\mu_X - \mu_Y$, difference of sample means $\bar{X} - \bar{Y}$, and apply the basic results considered in class to define a statistic that relates $\mu_X - \mu_Y$ to $\bar{X} - \bar{Y}$. The distribution of the statistic must be known.

4. The MATWES¹ method was designed to measure the attitude towards executive women. A high score indicates a negative attitude from the person that is being evaluated, while a low score indicates a positive attitude. It is conjectured that the attitude towards executive women changes in function of gender.

In order to test this hypothesis two independent simple random samples of M.B.A. students were taken, one with 151 men and another with 108 women. In the sample of men the mean of score was 85,8 with standard deviation 19,3. For women the mean of the score was 71,5 with standard deviation 12,2.

- a) What conclusion do you get with $\alpha = 0,01$ significance level? Choose an appropriate test, specifying clearly the null hypothesis and alternative you are considering, as well as the working hypotheses.
- b) Taking into account the results, do you think that the confidence interval at 99% level would contain the 0 value inside?
5. In order to compare the speed level of two computers, A and B, the time spent in carrying out some kind of operations is measured. A sample of 5 operations of this kind was taken and each operation was carried out by both computers. The following data of time (in milliseconds) were obtained:

¹P. Dubno, "Attitudes toward woman executives: A longitudinal approach", *Academy of Management Journal*, 28 (1985), 135-39, see Newbold

A	110	125	141	113	182
B	102	120	135	114	175

Analyze whether there are differences:

- Taking into account that the data are matched (the same operation is carried out by the computers). Obtain the p -value of the test.
 - Consider independent samples.
 - In the second case, when considering independent samples, what may be happening in your opinion? the following plot, in which the data of time of both samples are represented simultaneously, can be helpful:
- It is conjectured that the assets of a company will suffer from more variability in industry with competence in prices than in one with duopoly and tacit collusion. In a study about industry of steam turbine generators², it was seen that in 4 years the variability in price competence with respect to General Electric was 114,09. In the next 7 years, in which there was duopoly and tacit collusion, this variability was 16,08. Assume that the data may be considered independent random samples for two normal populations and test the previous conjecture with 5% significance level.
 - A user of large amount of electric components usually purchases them mainly from two suppliers, A and B. Since the structure of the prices are better, the user will do business only with supplier B whenever the proportion of defective items for B does not exceed the proportion of supplier A (it is less reliable). From two big lots, one of supplier A and the other of supplier B, the user selected randomly 125 items from A and 100 from B; inspected the items and found out 7 defective items of each supplier (7 and 7 for A and B respectively). Under appropriate hypotheses and the information above, is there any reason for not obtaining items from supplier B?
 - A professor in Political Science thinks that the interest of the students in the subject that she teaches could be related to her participation degree in the election process of her country. In order to check such a hypothesis, she wants to compare the mean of the marks of the students that polled in the last elections with respect to those that did not. From a couple of independent samples of 114 student that did vote and 123 that did not, in the first group it was obtained that the mean of the mark was 2,71, with a standard deviation equal to 0,64. In the second group the mean of the mark was 2,79, with a standard deviation equal to 0,56. Choose an appropriate test, specifying clearly the null hypothesis and alternative you are considering, as well as the working hypotheses. Define the critical region at α significance level and find the p -value of the test.
 - It is desired to analyze the whether a regular class attendance has any effect on the final mark in Statistics I. For that purpose independent samples of 8 student that usually attended were taken and on the other hand 8 that did not. The following results were obtained:

Regularly (1)	2.5	1.5	7	4	8.2	6.8	9.5	8
Occasionally (0)	0.75	1.5	3	1	4	4.5	6.5	5

- Propose a two-sided test in order to see whether there are significative differences due to not attending. What conclusion do you obtain at $\alpha = 0,05$ significance level? Provide a bound for the p -value.
- Taking into account the result of the previous section, would your answer change if you propose a one-sided test rather than a two-sided? How does the p -value change?
- Furthermore, the hours that each student spends studying out of class are shown in the following plot (triangle: more than 1 hour per week, circle: less than 1 hour per week).

Taking into account such a plot, do you think that the power of the test would increase if we propose to work with matched samples? In case of an affirmative answer, what could be a criterion for matching?

²B.T.Allen, "Tacit collusion and market sharing: The case of steam turbine generators", *Industrial Organization Review* 4 (1976), 48-57. See Newbold.

10. We are interested in knowing if the mean of salary in two big cities of a country are homogeneous or on the contrary, there are significative differences. For that purpose independent samples of workers (with similar activity and similar level) were taken from both cities and the following data were obtained:

City 1	5.9	6.1	6.3	6.1	6	6.2	5.7	6.3	6.6		
City 2	6.4	6.3	6.5	6.1	5	5.5	4.7	5.1	2.8	1.8	1.5

- a) Propose a one-sided test in order to see if there are significative differences in the salary of the workers in both cities. What is the conclusion at $\alpha = 0,1$ significance level? Provide a bound for the p -value.
- b) In the following plot the data were represented taking into account additional information, gender of workers (triangles are for women)

Taking into account this plot, would you propose a new study?