URJC – GADE BILINGÜE - CORPORATE STATISTICS II
May 2016 Exam
(model X)

SURNAMES:		**NAME**		
DNI: Group:		B:PROBLEM SOLVING (60% weight)	POINTS	STUDENT
A: MULTIPLE CHOICE (40% weight)		Exercise 1	2.5	
		Exercise 2	2.5	
RIGHT (+ 1/10)		Exercise 3	2.5	
WRONG (-0,2/10)	-	Exercise 4	2.5	
MC GRADE out of 10		PS GRADE	out of 10	
40 %		60%		
			FINAL GRADE	

EXAM DURATION: 110 MINUTES

The exam has two sections:

- Multiple Choice section (pages 2-4): 10 questions, weighting 40% of the final grade. **ONLY THOSE ANSWERS MARKED IN THE MASK WILL BE CONSIDERED.** At the end of this section there is space to carry out calculations if needed.

- A correct answer adds 1 point. A wrong answer substracts 0.2 points. A question not answered adds 0 points. **A minimum grade of 4 points in this section is required for assessing section B.**

- Problem Solving section (pages 5-12): 4 exercises weighting 60% of the total grade. **A minimum grade of 5 points is required to pass**.

- The final grade will be the result of adding 40% of the MC and 60% of the PS. **A minimum final grade of 5 points is required to pass**.

MULTIPLE CHOICE ANSWERS

	1	2	3	4	5	6	7	8	9	10
A		X		X						
B	X									
C								X	X	
D			X		X	X	X			X

SECTION A: TEST

1.- Let X be a random variable and a large s.r.s. (simple random sample) obtained from that. Choose the right answer (where CLT stands for Central Limit Theory):

a) The CLT guarantees that any observation in the sample follows a normal distribution
b) The CLT guarantees that the addition of the elements in the sample follows a normal distribution
c) The CLT guarantees that the difference of two elements in the sample follows a normal distribution
d) The three previous statements are right

2.- Choose which one of the coming situations would match the technique of simple random sampling:

a) To choose randomly n individuals from a population of N previously numerically coded in a list
b) An interviewer located in a street surveying people at random
c) A lecturer handing out a survey to the conference attendees chosen at random
d) A survey run among the visitors of a website

3.- In Inference it is very common to say that parameters are unknown. Why?

a) Because they are random variables
b) Because they are statistics
c) Because they are point estimators
d) Because most of the time all elements in the population are not available

4.- Choose the right answer regarding the ideal situation in a hypothesis testing:

a) To reject the null hypothesis H_0 because we control the probability of making type I error
b) To reject the null hypothesis H_0 because we control the probability of making type II error
c) To reject the alternative hypothesis H_1 because we control the probability of making type I error
d) The three previous statements are false

5.- Why do we affirm that the maximum likelihood method (ML) is a good method of obtaining point estimators?

a) Because it gives unbiased estimators in any s.r.s.
b) Because it gives efficient estimators in any s.r.s.
c) Because it gives estimators behaving as normal variables in any s.r.s.
d) Because in large s.r.s it gives estimators being unbiased, efficient and behaving as normal variables

6.- Choose the right answer regarding the p-value in a hypothesis test:

a) It depends on the distribution of the test statistic
b) It depends on the sample
c) It depends on the test statistic
d) The three previous statements are true

7.- Let a B(m;p) where m is large (bigger than 30) and p small (lower than 0.1). Choose the right answer regarding the distribution than can be approximated:

a) t_{n-1}
b) $F_{n-1,m-1}$
c) $N(mp; \sqrt{mp(1-p)}$
d) $P(\lambda = mp)$

8.- Let a random variable N(μ;σ) and a s.r.s. with size n obtained from that, aimed at estimating μ. Now suppose we have the following point estimator:

$$\widehat{\mu_1} = \frac{x_1 - 2x_2}{3}$$

Choose the right answer regarding the distribution followed by the point estimator:

a) $N(\mu;\sigma)$
b) $N(-\frac{\mu}{3};\frac{5}{9}\sigma)$
c) $N(-\frac{\mu}{3};\frac{\sigma}{3}\sqrt{5})$
d) $N(\mu;\frac{\sigma}{\sqrt{n}})$

9.- Choose the right answer in relation with the Shapiro-Wilks test:

a) It is a parametric test connected to the normal distribution
b) It is a goodness of fit test for sample sizes bigger than 100
c) It is a one sided test to the left
d) It is a test suited for discrete variables

10.- Choose the right answer regarding the General Method of Moments (GMM) point estimator $\hat{p}_{GMM}$ of p in a B(m;p):

a) $\hat{p}_{GMM} = \bar{x}$
b) $\hat{p}_{GMM} = \bar{x} - m$
c) $\hat{p}_{GMM} = m\bar{x}$
d) $\hat{p}_{GMM} = \frac{\bar{x}}{m}$

REMEMBER TO PASS YOUR ANSWERS TO THE MASK

SPACE FOR YOUR CALCULATIONS FOR THE MULTIPLE CHOICE SECTION. IF NEEDED.

PROBLEM SOLVING SECTION

Exercise 1 (2,5 points)

A private association against default has studied the time measured in days that city councils delay payments to their suppliers, variable (X). Such variable is assumed to follow a N(μ;σ). With that purpose a s.r.s of 100 city councils has been selected. These are the results:

$$\sum x_i = 8481{,}8 \qquad \sum x_i^2 = 727940{,}77$$

We wish to run a Pearson's Chi square goodness of fit test to match that sample with a N(μ;σ) distribution. To achieve that you must take into account that the sample has been divided in 5 classes with the same probability, apart from the information provided in the next table:

Classes	Oi	Pi	Ei	$\frac{(O_i - E_i)^2}{E_i}$
	24			
	15			
	19			
	21			
	21			
Total:	100			

Where Oi is referred to as the observed frequency in each class, Pi is the probability assigned by the normal distribution and Ei is the expected frequency under such distribution.

Answer:

a) Complete the table. (1 point)

b) Run the corresponding test, with α = 5%, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value. (1.5 points)

Exercise 2 (2.5 points)

In the context of the previous exercise, the association assures that city councils delay payments to their suppliers more than 80 days on average. Based on the s.r.s. of 100 city councils used in exercise 1 answer:

Elaborate the appropriate test under the point of view of the association at a 5% significant level, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve by the p-value.

Exercise 3 (2.5 points)

Under the same scenario, the association wishes to verify if the standard deviation of the variable X is equal to 10. Then based on the s.r.s. employed in exercise 1:

Run the proper test at a 5% significant level, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve by the p-value.

Exercise 4 (2.5 points)

By another side, the association wishes to estimate the proportion of city councils delaying payments to their suppliers more than 90 days. Taking advantage on the s.r.s. obtained in exercise 1, they realized that 30 city councils out of 100 incurred in this practice.

Answer:

a) Elaborate a confidence interval estimation at a 99% level of confidence for the corresponding population proportion, including the pivot statistic and the statistics defining the lower and upper limits of the interval. A manager of the association affirms that 40% of the city councils fall under this category. Discuss. (1.5 points)

b) What will be the sample size required in order to have a sampling error of 3% keeping the level of confidence unchanged? (1 point)