URJC – GADE BILINGÜE - CORPORATE STATISTICS II

May 2017 Exam

(model X)

SURNAMES:		**NAME**		
DNI: Group:		B:PROBLEM SOLVING (60% weight)	POINTS	STUDENT
A: MULTIPLE CHOICE (40% weight)		Exercise 1	3	
		Exercise 2	2,5	
RIGHT (+ 1/10)		Exercise 3	2,5	
WRONG (-0.2/10)	-	Exercise 4	2	
MC GRADE out of 10		PS GRADE	out of 10	
40 %		60%		
			FINAL GRADE	

EXAM DURATION: 100 MINUTES

The exam has two sections:

- Section A: Multiple Choice (pages 2-4) with 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 subtracts 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 and for passing the exam.**

- Section B: Problem Solving (pages 5-14) with 4 exercises weighting 60% of the total grade. **A minimum grade of 5 points in this section is required to pass**.

- The final grade will be the result of adding 40% of section A and 60% of section B. **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					
B		X		X				X		
C						X				X
D	X		X				X		X	

SECTION A: TEST

1.- In a Spanish city, the price in euros of a theatre ticket follows a random variable N(26;2). Then what is the approximate probability that the ticket price for a given performance being under 20 euros?

a) 0,95
b) 0,99
c) 0,05
d) 0

2.- The difference between two competitors A and B in their average yearly sales, $\mu_A - \mu_B$, measured in millions of euros is estimated to fall in the interval $[-2; -5]$ for a level of confidence of 95%. Choose the right choice in relation to that information at that level of confidence:

a) A sells more than B
b) B sells more than A
c) There are no significant differences in sales between A and B
d) The three previous statements are false

3.- Let a random variable N(μ;σ) and a s.r.s. (simple random sample) of size n obtained from it in order to estimate μ. Now consider this point estimator:

$$\hat{\mu} = \frac{2x_1 - x_2}{2}$$

Choose the right option regarding this point estimator:

a) It is unbiased
b) It is biased but asymptotically unbiased
c) Its bias is $\frac{1}{2}\mu$
d) Its bias is $-\frac{1}{2}\mu$

4.- In the context of the previous question, choose the right choice regarding the variance of the point estimator:

a) It is $0{,}75\sigma^2$

b) It is $1{,}25\sigma^2$

c) It is $1{,}5\sigma^2$

d) The three previous statements are false

5.- Consider the following statistic coming from a s.r.s. of size n obtained from a random variable N(μ;σ):

$$\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2}$$

Choose the right choice regarding its behaviour as a random variable:

a) χ^2_{n-1}
b) t_{n-1}
c) $N(\mu, \sigma)$
d) *B(m;p)*

6.- Which advantage does a s.r.s. supply over a non probabilistic sample when estimating a parameter by a confidence interval? Choose the right option:

a) Economy, provided that a s.r.s. is always cheaper than a non probabilistic sample
b) Subjectivity, an essential feature when sampling
c) The possibility of measuring the probability of making a mistake
d) The three previous statements are false

7.- When applying the maximum likelihood method for estimating parameters, the likelihood function of a s.r.s. selected from a given random variable is employed. That likelihood function is the product of the density function of each element in the sample. This is due to:

a) The elements in a s.r.s. are dependent in probability
b) The central limit theorem
c) The variable follows a $N(\mu;\sigma)$
d) The three previous statements are false

8.- Among those students attending a given postgraduate program a s.r.s. with n of them has been selected. By one side, their grades obtained in the entry exam $(x_1, x_2, \dots x_n)$ have been taken, assuming a $N(\mu_x;\sigma_x)$ in the corresponding variable X. By another side, their grades obtained in the final exam $(y_1, y_2, \dots y_n)$ over similar contents have been also taken, assuming a $N(\mu_y;\sigma_y)$ in the corresponding variable Y. Choose the right choice regarding the pivot statistic in the confidence interval for $\mu_x - \mu_y$ (where $\bar{d}\ and\ S_{1d}$ are, respectively, the mean and the bias corrected standard deviation at the sample obtained with the corresponding differences between both samples $(d_1 = x_1 - y_1; d_2 = x_2 - y_2; \dots d_n = x_n - y_n)$, meanwhile $s_{1x}^2\ and\ s_{1y}^2$ are referred to as the bias corrected sample variances of X and Y respectively):

a) $\frac{(\bar{x}-\bar{y})-(\mu_x-\mu_y)}{\sqrt{\frac{\sigma_x^2}{n}+\frac{\sigma_y^2}{n}}} \sim N(0;1);$ b) $\frac{\bar{d}-(\mu_x-\mu_y)}{s_{1d}/\sqrt{n}} \sim t_{n-1}$ c) $\frac{(\bar{x}-\bar{y})-(\mu_x-\mu_y)}{\sqrt{\frac{s_{1x}^2}{n}+\frac{s_{1y}^2}{n}}} \sim N(0;1);$

d) The three previous statements are false

9.- Let a s.r.s. of size n = 20 coming from a $B(1;p)$ with $p = 0{,}3$. Choose the right option in relation to the variable arising when adding all the elements in that sample:

a) t_{n-1}

b) $N\left(6;\sqrt{4{,}2}\right)$

c) $P(\lambda = 6)$

d) $B(20;0{,}3)$

10.- In a survey, the section reporting of aspects such as the universe, the sample size, the sample procedure or the sampling error is called:

a) Questionnair
b) Population
c) Technical sheet or methodology
d) Stratum

REMEMBER TO PASS YOUR ANSWERS TO THE MASK

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

PROBLEM SOLVING SECTION

Exercise 1 (3 points)

The technical department of a courier service is carrying out a study in order to measure its efficiency. Aimed at that the economist in charge has collected a s.r.s. of n = 30 services, taking note of the time elapsed in hours since the customer makes a request and he or she receives the shipment. Assuming that the variable follows a N(μ;σ) the results obtained are:

$$\sum x_i = 1197{,}24 \qquad \sum x_i^2 = 47896{,}25$$

Based on that, the economist assures that the average time employed in a service is lower than 48 hours. Answer:

a) Elaborate the corresponding test under the point of view of the economist with α = 5%, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value. (2 points)

b) Draw a graph with the following information: the probabilistic distribution of the sample mean under the null hypothesis considered in question a), indicating α, apart from the probabilistic distribution of the sample mean under the alternative hypothesis considered in question a), indicating β. (1 point)

Exercise 2 (2,5 points)

By another side, the economist wants to test if the dispersion in the time employed in a service, measured by the standard deviation, is exactly equal to 2 hours. Based on the s.r.s. of n = 30 employed in exercise 1, answer:

Run the corresponding test with $\alpha = 5\%$, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value.

Exercise 3 (2,5 points)

At the same time the economist wishes to know the condition in which the shipment arrives to the customer. With that objective she wants to estimate the proportion of deliveries with some flaw. Based on the services surveyed in the s.r.s. of exercise 1, 3 out of the n = 30 shipments had some damage. Answer:

a) Elaborate a 95% confidence interval for the proportion of deliveries with some damage, including the pivot statistic and the statistics defining the lower and upper limits of the interval. (1,5 points)

b) Justify without doing any calculation what would be the decision taken in a two sided hypothesis test over the parameter considered in question a) of this exercise for the value $p_0 = 12\%$ if a level of significance of $\alpha = 5\%$ were used. (1 point)

Exercise 4 (2 points)

Finally the economist is going to test if customer's satisfaction is independent on the weight of the delivery. With this purpose she has considered two kind of customers (satisfied and unsatisfied) and three weight groups (< 200 grs, between 200 and 1000 grs; and >1000 grs). Then the economist has collected a s.r.s. of n = 60 shipments with these results:

	< 200 grs.	200-1000 grs.	> 1000 grs.
Satisfied	14	11	12
Unsatisfied	6	9	8

Run an independence test at a 1% significance level, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value.

DRAFT