URJC – GADE BILINGÜE - CORPORATE STATISTICS II

June 2017 Exam

(model X)

SURNAMES:		**NAME**		
DNI: Group:		B:PROBLEM SOLVING (60% weight)	POINTS	STUDENT
A: MULTIPLE CHOICE (40% weight)		Exercise 1	5	
		Exercise 2	3	
RIGHT (+ 1/10)		Exercise 3	2	
WRONG (-0.2/10)	-			
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 3 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 (mark with X)

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

SECTION A: TEST

1.- The weight in kilograms of a hake (a kind of fish) is a random variable following a N(1.5;0.3). Under these conditions, what is the approximate probability for a given hake weighting between 0.9 and 2.1 kilograms?

a) 0,95
b) 1
c) 0,05
d) 0

2.- Let X and Y be two normal variables. Having taken two independent s.r.s. (simple random sample) from them, the quotient between their respective variances $\frac{\sigma_x^2}{\sigma_y^2}$ is estimated through the interval $[0.2; 0.7]$ for a level of confidence of 95%. Choose the right choice regarding what that means at that level of confidence:

a) The variance of X is larger that the variance of Y
b) The variance of X is smaller than the variance of Y
c) There are not significative differences between the variances of X and Y
d) The three previous statements are false

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

$$\hat{\mu} = \frac{3x_1 - x_2}{2n}$$

Choose the right option regarding this point estimator:

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

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

a) It is $\frac{2\sigma^2}{n^2}$

b) It is $\frac{2\sigma^2}{n}$

c) It is $\frac{5\sigma^2}{2n^2}$

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 ($\bar{x}$ is the sample mean and $s_1$is the bias corrected sample standard deviation):

$$\frac{\bar{x} - \mu}{\frac{s_1}{\sqrt{n}}} \sim t_{n-1}$$

It is used in order to build a confidence interval estimation of the population mean μ when the population variance is unknown. Choose the right choice regarding the distribution followed necessarily by that random variable:

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

6.- Consider a random variable following a $B(1;p)$ and a s.r.s. of size n obtained from it $(x_1, x_2, \dots x_n)$. If n is large enough and p is small (smaller than 0.1), choose the right option regarding the approximate probability distribution followed by the addition of the elements in the sample $\sum x_i$:

a) $N(\mu;\sigma)$
b) t_{n-1}
c) $P(\lambda)$
d) $B(1;p)$

7.- Choose which one of the following sampling procedures correspond to a s.r.s.:

a) In a wedding the bride taking five numbered tickets out of n from a bag with the purpose of delivering some gifts; the n guests having received their respective tickets previously

b) A student surveying randomly other students at the entrance to the URJC in Vicálvaro

c) A researcher sending a questionnaire through internet for being answered by the first n people accessing the website where it has been located

d) An actor in a theatre selecting n people among the attendees to the performance

8.- Consider a random variable X with variance σ^2 and a s.r.s. of size n taken from it in order to obtain a point estimator of that parameter. Which one of the following point estimators is unbiased?

a) $\frac{\sum x_i^2}{n}$

b) $\frac{\sum x_i^2}{n-1}$

c) $\frac{\sum (x_i-\bar{x})^2}{n}$

d) $\frac{\sum (x_i-\bar{x})^2}{n-1}$

9.- Let X be a random variable following an $F_{m,n}$. Then choose the right option:

a) It is used in inference in order to compare population means from normal variables

b) It is defined, not only for positive abcises, but also for negative ones

c) The inverse of X, $\frac{1}{X}$, will follow an $F_{n,m}$

d) If $m = 3$ then $F_{3,n}$ will follow a t_n

10.- Which of the following procedure corresponds to a non parametric hypothesis test?

a) A Kolmogorov-Smirnov goodness of fit test
b) A maximum likelihood point estimator
c) A general method of moments point estimator
d) The three previous statements are false

REMEMBER TO PASS YOUR ANSWERS TO THE MASK

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

PROBLEM SOLVING SECTION

Exercise 1 (5 points)

a) A neighbourhood community in Madrid complains because of the level of traffic noise they have to bear in their homes. The president of that community assures that in week days from 10:00 till 14:00 the noise (variable X) is equal to 70 dB (decibels) on average. With the purpose of showing that hypothesis, the president has taken a s.r.s. of 40 measurements with these results:

$$\sum x_i = 2802.5 \qquad \sum x_i^2 = 196358.23$$

Answer:

Elaborate the corresponding hypothesis test for the average level of noise born by the neighbourhood community in a week day from 10:00 till 14:00 (variable X), under the president's point of view, 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) The community has also studied the noise produced in week days in the afternoon from 16:00 till 20:00 (variable Y). With that purpose the president has obtained another s.r.s. of 40 measurements with these results:

$$\sum y_i = 2647.14 \qquad \sum y_i^2 = 175228.11$$

The president asserts that on average the difference between the noise produced in the morning (variable X) and the noise produced in the afternoon (variable Y) is higher than 3 dB. The president wants to run a test to show that hypothesis.

Answer:

Elaborate the corresponding hypothesis test for the difference between the respective means of both variables, under the president's point of view, with α = 1%, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value. (2 points)

c) Discuss the possible errors incurred in sections a) and b) coming from the decisions taken, apart from the respective probabilities involved and the ways to reduce them. (1 point)

Exercise 2 (3 points)

By another side the president wishes to estimate the proportion of neighbours considering the traffic noise they bear to be high. Aimed at that she has taken a s.r.s. of 100 neighbours, 65 of them having answered yes when the question was made.

Answer:

a) Elaborate a 95% confidence interval for the proportion of neighbours feeling that traffic noise they suffer is high, including the pivot statistic and the statistics defining the lower and upper limits of the interval. Consider the maximum uncertainty case. (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 $p_0 = 50\%$ if a level of significance of $\alpha = 5\%$ were used. (0.5 points)

c) Give (without justifying) the maximum likelihood point estimator of the population proportion of section a) and justify if it verifies the properties of unbiasness and consistency. (1 point)

Exercise 3 (2 points)

At the same time, the president wishes to know if the noise is perceived in a uniform way in the different floors of the building regarding their height or not. Then she has spread the 65 neighbours complaining of high noise of the s.r.s. at the previous exercise in 5 classes depending on the floor in which each neighbour lives, having obtained the following results:

Neighbours with high noise	Oi
First floor	25
Second floor	15
Third floor	12
Fourth floor	8
Fifth floor	5

Where Oi stands for the observed frequency of a given class (the number of neighbours complaining at certain floor).

Answer:

Run a Pearson's Chi square goodness of fit test of those data to a discrete uniform distribution, with $\alpha = 1\%$, formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the p-value.

DRAFT