Surnames:	*Name:*

Weights: 40% for section A and 60% for section B

SECTION A: MULTIPLE CHOICE
Maximum grade: 10 points
A correct answer adds 1 point
A wrong answer substracts 0.2 points
Minimum grade for assessing SECTION B: 4 points
Those answers not marked in the mask will not be considered

1.- The simple random sampling technique commonly used in Inference involves the density function of the sample $f(x_1, x_2 \ldots x_n)$ being expressed as $f(x_1) \cdot f(x_2) \cdots f(x_n)$. This is due to:
a) The sampling is carried out with replacement
b) The sampling is carried out without replacement
c) The homogeneity of the population
d) The randomness of the sampling

2.- The monthly income earned by an engineer two years after her graduation is a random variable following a N(1800;200). Choose the right option regarding the approximate probability for a given engineer earning less than 1400 € two years after her graduation:

a) 0
b) 0.025
c) 0.05
d) 0.1

3.- Let X be a random variable with mean μ and variance σ^2. Then a s.r.s. of size n has been obtained with the following point estimator of μ being under consideration:

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

Choose the right option regarding the variance of that point estimator:

a) $\frac{3\sigma^2}{n}$

b) $\frac{3\sigma^2}{n^2}$

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

d) $\frac{5\sigma^2}{n^2}$

4.- Under simple random sampling the sample mean $\bar{x}$ is an unbiased estimator of the population mean μ (choose the right option):

a) Only when the sample size n is large enough
b) Only when the variable under study follows a normal distribution
c) In any case, regardless of the sample size n and the distribution followed by the variable under study
d) The statement is false provided that the sample mean is a biased estimator of the population mean

5.- An appropriate tool to collect data from a population is:

a) A parameter
b) An estimate
c) A random variable
d) A survey

6.- Under simple random sampling, when we need to estimate the population variance σ^2 of a variable two point estimators are considered: the sample variance $S^2 = \frac{\sum(x_i-\bar{x})^2}{n}$ and the bias-corrected sample variance $S_1^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$. In which cases do they produce similar inferences?

a) Only when the variable under study follows a normal distribution
b) Only when the sample size n is large enough
c) In any case, regardless of the sample size n and the distribution followed by the variable under study
d) The statement is false, they never produce similar inferences

7.- Choose the right option in a hypothesis test:

a) For a given sample to reduce α involves to reduce β
b) For a given sample to reduce α involves to increase β
c) For a given sample to increase α involves to increase β
d) The level of α does not affect the level of β

8.- Choose the right option regarding the Student's t_n distribution:

a) It approximates a N(0;1) distribution when the degrees of freedom n are large enough
b) The distribution can only take positive values
c) It is used for doing inference over the variance
d) The mean of the variable is 1

9.- Choose the right option regarding the situation arising from a hypothesis contrast when deciding which hypothesis to establish as null and which one as alternative:

a) It doesn't matter which hypothesis to establish as null and which one as alternative
b) It matters which hypothesis to establish as null and which one as alternative provided that it will determine the value of the test statistic in the sample
c) It matters which hypothesis to establish as null and which one as alternative provided that it will affect the errors (type I and type II) that might be incurred
d) The three previous statements are false

10.- Consider a confidence interval estimation for the population mean μ of a variable following a normal distribution. Why cannot a 100% level of confidence be applied?

a) Because the interval would cover all the real line, hence becoming useless
b) Because the interval would be formed by a single point, hence becoming useless
c) Because the precision of the estimation would be maximum
d) Because the variable follows a normal distribution

(model x) MARK YOUR ANSWER WITH AN X

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

<u>SECTION B: Exercise</u>
Maximum grade: 10 points
Minimum grade to pass: 5 points

Exercise 1 (4 Points)

A car manufacturer assures that its X model carbon dioxide emissions are below 50 gCO_2/km. In order to show that hypothesis the manufacturer has obtained a s.r.s. of 20 cars and measured the emissions. Assuming this variable follows a N(μ;σ), the results coming from the sample have been:

$$\sum x_i = 985.4917 \qquad \sum x_i^2 = 48603.0033$$

a) Run the corresponding test at a 5% significance level formulating the hypothesis, the test statistic, the critical region and the decision made. Also solve using the *p-value*.

b) Redo question a) but now applying a 1% level of significance.

c) Considering the agency wishes a 99% level of confidence, which decision should be made? Discuss.

Exercise 2 (4 points)

In the context of the previous exercise answer:

a) Obtain the Generalized Method of Moments (GMM) estimator of both, μ and σ^2, in the normal distribution regarding the carbon dioxide emissions.

b) Elaborate the 95% confidence interval estimation for the standard deviation of the variable under study. Interprete.

c) By another side the car manufacturer wants to compare the emissions of this X model with those released by a Y model, a car with similar characteristics but produced by another manufacturer. The company has data from a s.r.s. of 20 cars of this Y model emissions (this variable is assumed to be normal and this sample is independent from that taken from the X model). These are the data:

$$\sum y_i = 983.0547 \qquad \sum y_i^2 = 48383.5506$$

With that purpose elaborate the 95% confidence interval estimation for the difference between the two population means, interpreting the results. Take into account that population variances are unknown but equal.

Exercise 3 (2 Points)

Run a Shapiro-Wilks goodness of fit test with $\alpha = 5\%$ in order to check if the s.r.s. of model X cars' emissions obtained in exercise 1 matches a $N(\mu;\sigma)$. Formulate the hypothesis, the test statistic, the critical region and the decision made. Also solve using the *p-value*. The corresponding complete ordered sample (of exercise 1) is:

46,7594	47,0860	47,2640	47,8676	48,5466	48,5885	48,7814	48,7932	48,8256	48,9176
49,0541	49,0990	49,5663	49,6425	49,8268	50,3365	51,1258	51,7097	51,7897	51,9115