

Duration: 2 h 10 min

EN A

Name \_\_\_\_\_ N<sup>o</sup> \_\_\_\_\_

*Space for marks*

**Multiple choice questions:** only one option is correct  
each right answer scores **1**  
each wrong answer scores **-0.25**

**True an false questions:** each right answer scores **0.25** points  
each wrong answer scores **-0.25** points

**Derive and justify all your answers.**

If you need extra space, you can use page 8 of the paper, showing clearly which question are you answering.  
The only page you can detach from the staple is the last one, which contains the regression output and can be used as draft paper.

## PART II

1. For each statement, indicate if it is true (T) or false (F). (0.5)

	T	F
If, in a statistic test, the $p$ -value is 0.1, than we reject $H_0$ at the 5% significance level.	<input type="checkbox"/>	<input type="checkbox"/>
The model $y_t = \alpha_1 e^{\alpha_2 x_{2t}}$ can not be studied within the framework of linear regression, as it is not linear.	<input type="checkbox"/>	<input type="checkbox"/>

2. The monthly return on two different investment funds, A and B, are assumed to follow a normal distributions, with equal variances but different means. From the observation of the past 3 years (36 months) a mean monthly return of 0.034 and 0.012, respectively for funds A and B, and a corrected standard deviation of 0.16 and 0.041, respectively for funds A and B, were observed.

(a) The manager of investment fund B states that the mean monthly returns of fund A are zero. Would you agree with this Statement Your based on 95% confidence interval. (1.5)

(b) The manager of investment fund A states that the mean monthly returns of fund A are greater than those of fund B. Would you agree with this statement? Justify your opinion based based on an appropriate test of size 5%. (1.5)

3. The scores (from 0 to 20) in the Statistics course of 173 students were analysed, with the following results (2.0)

	lower than 10	bewteen 10 and 15	higher than 15
<b>male</b>	25	41	11
<b>female</b>	35	47	14

Is there evidence to conclude, at the 5% level, that the Statistics scores depend on gender?

4. In order to study the construction cost of a building in the center of Lisbon, the manager of a construction company proposed the following model, for which he collected data form 37 buildings in the city center:

$$lcost_i = \beta_0 + \beta_1 lfarea_i + \beta_2 ltotarea_i + \beta_3 feheight_i + \beta_4 type_i + u_i, \quad i = 1, \dots, n$$

where the variables are as follows:

- *lcost*: logarithm of the cost of construction per square meter (€);
- *lfarea*: logarithm of the average floor area ( $m^2$ );
- *ltotarea*: logarithm of the total area of the building ( $m^2$ );
- *fheight*: average storey height ( $m$ );
- *type*: dummy variable that takes the value 1 if the building structure is steel, and 0 if it is reinforced concrete.

Taking into account the results of the estimation of the above model, presented in the output, answer the following questions.

- (a) Interpret the estimates of parameters  $\beta_1$ ,  $\beta_2$  and  $\beta_4$ , and comment on the statistical significance of regressors *ltotarea<sub>i</sub>*, *fheight<sub>i</sub>* and *type<sub>i</sub>*. (1.5)
- (b) Interpret the coefficient of adjustment and test, at a 5% level, if global significance of the model. (1.5)
- (c) A civil engineer states that an increase of 5% in the building total area leads, *ceteris paributs*, to an increase of at least 4% in the construction cost. Whould you agree with the engineer? Answer based on an appropriate test at the 5% level. (1.5)

## PART I

1. For each statement, indicate if it is true (T) or false (F). (0.5)

	T	F
The estimator $T$ for $\theta$ is unbiased if $Var(T) = Var(\theta)$ .	<input type="checkbox"/>	<input type="checkbox"/>
A confidence interval for the mean of a normal population, obtained through the usual procedure using the Pivotal, is symmetric with respect to the sample average.	<input type="checkbox"/>	<input type="checkbox"/>

2. The weights of individuals of a certain country, in  $Kg$ , are assumed to follow a normal distribution with mean  $\mu = 66$  and standard deviation of  $\sigma = 17$ . Consider a sample of 10 individuals.

- (a) What is the probability that the heaviest person in the sample weights less than 70  $Kg$ . (1.5)
- (b) What is the probability that the average weight in the samples is higher than 70? (1.5)

3. In a random sample of 100 students of ISEG, 35 stated that training in statistics is the most important component of their graduation courses.

- (a) Obtain the Method of Moments estimator and estimate for the proportion of students who think the same. (1.5)
- (b) Build a 95% confidence interval for the proportion of students who believe that training in statistics is the most important component in their graduation courses. Is it plausible to state that the proportion of students that do believe statistics is the most important course is 50%? (2.0)

- (c) Given a 96% confidence interval for the proportion of students who believe that training in statistics is the most important component in their graduation courses, which of the following quantities is guaranteed to be in that confidence interval? (1.0)

 $\mu$ 

0

 $\bar{x}$ 

0.96

- (d) The sample is composed of 54 men, of which 24 answered that statistics was the most important course during graduation. Of the women, 11 answered the same. Build a confidence interval at 95% for the difference in the proportion of men and women who believe statistics is the most important course in graduation. What can you conclude? (2.0)

Continuing question...

