Name: ______________________________

Matricola: ______________________ email: ______________________

ECONOMETRICS - 25-06-2021 - Time: 2 h 30'

1. Say if the following statements are unambiguously true (True), unambiguously false (False) or impossible to classify the way they are stated (Not necessarily). Write the motivations to your answers **only** in the space provided. A "Not necessarily" answer with no adequate motivation will be considered wrong.

   (a) A square projection matrix is invertible.
   True ◯ False ◯ Not necessarily ◯

   (b) If $\lim_{n\to\infty} Pr(X_n \leq x) = Pr(X \leq x)$, then $X_n \xrightarrow{d} X$.
   True ◯ False ◯ Not necessarily ◯

   (c) Consistent estimators are unbiased.
   True ◯ False ◯ Not necessarily ◯

   (d) If a linear model contains an intercept, the average of OLS residuals is 0.
   True ◯ False ◯ Not necessarily ◯

   (e) Suppose you have a linear model $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. If $V(y_i|x_i) = \exp(x_i)$, then the model is heteroskedastic.
   True ◯ False ◯ Not necessarily ◯

2. Suppose you have two continuous variables y_i and x_i, for which the following holds:

$$E(y_i|x_i) = \beta_1 x_i + \beta_2 x_i^2.$$

   (a) Calculate the unconditional expectation of y_i, given that x_i is a Gaussian rv with mean 1 and variance 1:

$$x_i \sim N(1,1) \Rightarrow E(y_i) = \_\_\_\_\_\_\_\_\_\_$$

   (b) Now suppose you observe $n = 400$ iid realisations of y_i and x_i and the following descriptive statistics:

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 = 2 \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^3 = 4.5 \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^4 = 12$$

$$\frac{1}{n}\sum_{i=1}^{n} y_i^2 = 1.6 \qquad \frac{1}{n}\sum_{i=1}^{n} x_i y_i = 1 \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^2 y_i = 1.5.$$

   Calculate the OLS estimates of the conditional mean parameters.

$$\hat{\beta}_1 = \_\_\_\_\_\_\_\_\_\_ \qquad \hat{\beta}_2 = \_\_\_\_\_\_\_\_\_\_$$

   (c) Calculate the sum of squared residuals and an estimate of $V(y_i|x_i)$

$$SSR = \_\_\_\_\_\_\_\_\_\_ \qquad \hat{\sigma}^2 = \_\_\_\_\_\_\_\_\_\_$$

   (d) Test the hypothesis that $E(y_i|x_i)$ is linear in x_i:

   Test type: ________ Distribution: ______ Test statistic: ______
   Decision: ○ Reject ○ Don't reject

   (e) Test the hypothesis that $E(y_i|x_i) = 0$:

   Test type: ________ Distribution: ______ Test statistic: ______
   Decision: ○ Reject ○ Don't reject

3. The following variables were observed for 106 Italian provinces:

Symbol	Description	Mean	Median	S.D.	Min	Max
A_i	Road accidents in 2019	1609	1065	2099	159	15401
Y_i	Value added per capita	23976	23793	6861	13454	50120
B_i	Number of buses	935.3	559	1140	107	8859
K_i	Number of motorbikes/scooters	64422	44345	73057	9506	516700
V_i	Total vehicles (thousands)	489.5	339.1	491.4	89.85	3541
N_i	Population (thousands)	559	383.5	609.4	84.03	4264
W_i	Area in sq Km	2772	2476	1607	212.5	7398

The following data transformations were performed:

$a_i = \log(A_i/N_i)$ $\quad y_i = \log(Y_i)$ $\quad k_i = \log(K_i/V_i)$ $\quad b_i = \log(B_i/V_i)$ $\quad n_i = \log(N_i)$ $\quad d_i = \log(N_i/W_i)$

and the following model was estimated by OLS

$$a_i = \beta_0 + \beta_1 y_i + \beta_2 k_i + \beta_3 b_i + \beta_4 n_i + \beta_5 d_i + \varepsilon_i \tag{1}$$

Answer the following questions, using the numerical estimates that you find in table 1:

OLS, using observations 1–106
Heteroskedasticity-robust standard errors, variant HC1

	Coefficient	Std. Error	t-ratio	p-value
β_0	-7.9090	0.7937	-9.9649	0.0000
β_1	0.6100	0.0707	8.6333	0.0000
β_2	0.6120	0.0592	10.3412	0.0000
β_3	-0.0807	0.0396	-2.0365	0.0443
β_4	-0.0250	0.0389	-0.6429	0.5218
β_5	0.0034	0.0393	0.0866	0.9312

Mean dependent var	0.966645	S.D. dependent var	0.351059
Sum squared resid	3.663232	S.E. of regression	0.191396
R^2	0.716916	Adjusted R^2	0.702762
$F(5, 100)$	53.52241	P-value(F)	9.32e–27

White's test for heteroskedasticity: Test statistic: LM = 20.536 (p-value = 0.424881)

RESET test for specification: Test statistic: $F(2, 98) = 0.079681$ (p-value = 0.923471)

Table 1: Model for road accidents per capita

(a) Comment on the sign, magnitude and statistical significance of β_1, and provide an interpretation:

(b) Comment on the sign, magnitude and statistical significance of β_2 and β_3, and provide interpretation:

(c) Comment on the sign, magnitude and statistical significance of β_4 and β_5, and provide an interpretation:

(d) Comment on the result of the diagnostic tests.