ISEG, Univeridade de Lisboa
Econometrics EXAM
January 5, 2023

TOTAL TIME 2 HOURS

Instructions

The exam has 2 parts: **Section A** and **B**, each with a **50% weight** in the final grade. Your final grade has to be 9.5 to pass this exam (irrespective of the mark in each section). During this exam, you are allowed to **only** use the formulae sheet provided and a calculator (but not tablets or phones). Lecture notes or books are **NOT** allowed.

SECTION A

Total marks 20 (50% weight in the final grade).

Problems 1 and 2 below can be skipped by those who had a mark for the mid-term exam of at least 6 (our of 20). If you wish a higher mark than the one for the mid-term exam then solve Problems 1 and 2 below (if you do that, then the mark for the mid-term exam is not taken into account in the final grade).

Problem 1. Total 3 marks

Answer the following three multiple choice questions. Do not give a justification of your answer. There is only one correct answer for each question: (a), (b) or (c). Each correct answer receives a mark of 1. Each incorrect answer receives a penalty of -0.25.

I. Consider the following equation:

$$y_i = \beta_0 + \beta_1 x_i + u_i. \tag{1}$$

Suppose there is a true linear model where in (1) there is an additional important regressor z_i correlated with x_i. However, you estimate the simple linear regression model (1).

(a) There is no effect on the OLS estimator $\hat{\beta}_1$ from (1), $\hat{\beta}_1$ is unbiased.

(b) The OLS estimator $\hat{\beta}_1$ from (1) will be biased.

(c) The usual OLS standard error of the OLS estimator $\hat{\beta}_1$ remains valid.

Answer: b

II. Under the Gauss-Markov conditions, OLS estimators can be shown to be BLUE. The word "linear" in this acronym refers to:

(a) We are estimating a model linear in the variables.

(b) We are estimating a model linear in the parameters.

(c) We are estimating a model linear in the variables and parameters.

Answer: b

III. How do we use the p-value to reach our conclusion about a test?

(a) A small p-value is evidence against the null hypothesis because one would reject the null hypothesis even at small significance levels.

(b) A large p-value is evidence against the null hypothesis because one would reject the null hypothesis even at large significance levels.

(c) A small p-value is evidence against the alternative hypothesis because one would reject the null hypothesis even at small significance levels.

Answer: a

Problem 2. Total 17 marks

Consider the standard wage equation:

$$\log(wage) = \beta_0 + \beta_1\, educ + \beta_2\, exper + \beta_3\, tenure + u \tag{2}$$

where *educ*, *exper* and *tenure* are years of education, experience and tenure with the current employer respectively.

The estimation results are given in the following Stata output:

Source	SS	df	MS
Model	25.6953242	3	8.56510806
Residual	139.960959	931	.150334005
Total	165.656283	934	.177362188

Number of obs	=	935
F(3, 931)	=	56.97
Prob > F	=	0.0000
R-squared	=	0.1551
Adj R-squared	=	0.1524
Root MSE	=	.38773

lwage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
educ	.0748638	.0065124	11.50	0.000	.062083	.0876446
exper	.0153285	.0033696	4.55	0.000	.0087156	.0219413
tenure	.0133748	.0025872	5.17	0.000	.0082974	.0184522
_cons	5.496696	.1105282	49.73	0.000	5.279782	5.713609

(a) Interpret the estimate of the β_2 coefficient.

(1 Mark)

Answer: One additional year of experience increases the salary by: 1.53 percent.

(b) Are *educ*, *exper* and *tenure* individually significant at the 5% level? Use the Stata output above to justify your answer. State the null and alternative hypotheses.

Short answer: They are all statistically sinficant at 5% because the p-value is smaller than 5%.

(3 Marks)

(c) Are *educ*, *exper* and *tenure* jointly significant at the 5% level? Use the Stata output above to justify your answer. State the null and alternative hypotheses. Calculate the relevant statistic using the Stata output above.

(4 Marks)

Short answer: Yes because the p-value associated with the F-statistic is smaller than 5%. Use the formula for the F-statistic based on the R^2 with the R^2 for the restricted regression equal to 0.

(d) How much in the variation of $\log(wage)$ is explained by the model with *educ*, *exper* and *tenure* as regressors? Using the Stata output above show how you arrive at the number on which your answer is based.

(2 Marks)

Short answer: 15.51 percent (this is the R^2). It can be calculated by using from the Stata output SSR and SST.

(e) State the null hypothesis that another year of workforce experience has the same effect on $\log(wage)$ as another year of tenure with the current employer.

(1 Mark)

Answer: $H_0 : \beta_2 = \beta_3$

(f) Write a new regression equation that allows you to directly test using the t-test the null hypothesis stated in question (e).

(2 Marks)

Answer: Denote $\theta = \beta_2 - \beta_3$ and replace for either β_2 or β_3 in (2) to obtain the new regression which automatically delivers $\hat{\theta}$ and its standard error.

(g) Using a 5% significance level and the Stata outputs (above and below), test the null hypothesis that *sibs* (number of siblings) and *age* (in years) have no effect on wage (after controlling for education, experience and tenure).

(4 Marks)

Answer: calculate $F = \frac{(139.96-138.79)/2}{138.79/(935-6)}$ and compare it with the 5% critical value from the table of the F distribution with 2 and 929 degrees of freedom.

Source	SS	df	MS			
				Number of obs	=	935
				F(5, 929)	=	35.96
Model	26.8611795	5	5.37223589	Prob > F	=	0.0000
Residual	138.795104	929	.149402695	R-squared	=	0.1622
				Adj R-squared	=	0.1576
Total	165.656283	934	.177362188	Root MSE	=	.38653

lwage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
educ	.0683854	.0069057	9.90	0.000	.0548328	.0819381
exper	.0114801	.0039255	2.92	0.004	.0037763	.019184
tenure	.0124307	.0026148	4.75	0.000	.0072992	.0175622
sibs	-.0121724	.0056602	-2.15	0.032	-.0232807	-.0010642
age	.0086558	.0049387	1.75	0.080	-.0010366	.0183481
_cons	5.384746	.1630053	33.03	0.000	5.064844	5.704647

SECTION B

Total 20 marks (50% weight in the final grade)

Solve Problems 3, 4 and 5 below.

Problem 3. Total 3 marks

Answer the following multiple choice questions. Do not give a justification of your answer. There is only one correct answer. Each correct answer receives a mark of 1. Each incorrect answer receives a penalty of -0.25.

I. A covariance-stationary time series y_t is:

(a) also weakly dependent.

(b) also weakly dependent if $Cov(y_t, y_{t+h}) = constant$ as $h \to \infty$.

(c) also weakly dependent if $Cov(y_t, y_{t-h}) = 0$ as $h \to \infty$.

Answer:c

II. Consider the time series process: $y_t = v_t + \delta v_{t-1}$, $t = 1, 2, \ldots$, where v_t is a white noise.

(a) y_t is a MA(2) process and is covariance-stationary only if $|\delta| < 1$.

(b) y_t is an AR(1) process and is covariance-stationary only if $|\delta| < 1$

(c) y_t is a MA(1) process and is covariance-stationary.

Answer:c

III. In cross-sectional data, the errors in the equation models are usually heteroskedastic which means that:

(a) $E(u|x_1, \ldots, x_k)$ varies with regressors.

(b) $Var(u|x_1, \ldots, x_k)$ varies with regressors.

(c) $Cov(u_i, u_j|x_1, \ldots, x_k)$ varies with the regressors.

Answer:b

Problem 4. Total 7 marks

A master student is interested to understand the factors triggering the salaries of chief executive officers (CEOs). She therefore conducts a survey where the CEOs from the financial services sectors are randomly selected. She collects information about their salaries, the return on equity (*roe*, in percentage form) and the *sales* of their companies. She is considering estimating the following equation:

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2\, roe + \beta_3\, investfund + \\ + \beta_4\, insurance + \beta_5\, banking + u \quad (3)$$

where *investfund*, *insurance* and *banking* are dummy variables indicating the investment funds, insurance and banking sectors. The omitted sector is accounting.

The estimated equation is:

$$\log(salary) = \underset{(0.30)}{4.50} + \underset{(0.03)}{0.31} \log(sales) + \underset{(0.002)}{0.01}\, roe + \underset{(0.081)}{0.32}\, investfund +$$

$$+ \underset{(0.082)}{0.18}\, insurance + \underset{(0.008)}{0.1}\, banking + u$$

(a) By what percentage is *salary* predicted to increase if *roe* increases by 50 points? Does *roe* have a large effect on *salary*?

(1 Mark)

Answer: The proportionate effect on salary is 0.01(50)=0.5. To obtain the percentage effect we multiply this by 100: 50%. Thus, a 50% (ceteris paribus) increase in *roe* is predicted to increase salary by 50%. This is a large effect of *roe* on *salary*.

(b) What is the approximate percentage difference in estimated salary between the investment funds and accounting sectors, holding *sales* and *roe* fixed?

(1 Mark)

Answer: The CEO salary in investment funds is predicted to increase by 18% relative to the accounting sector, holding the other factor fixed. We have:

$$\log(salary_{investfund}) - \log(salary_{accounting}) = 0.32$$

(c) What is the approximate percentage difference in estimated salary between investment funds and insurance sector?

(3 Marks) Answer: We have:

$$\log(salary_{investfund}) - \log(salary_{accounting}) = 0.32$$
$$\log(salary_{insurance}) - \log(salary_{accounting}) = 0.18$$
$$\log(salary_{banking}) - \log(salary_{accounting}) = 0.1$$

Then:

$$\log(salary_{investfund}) - \log(salary_{insurance}) = 0.32 - 0.18$$

(d) Write an equation that would allow you to test directly whether the difference mentioned in question (c) above is statistically significant.

(2 Marks) Answer:

$$\begin{aligned}\log(salary) &= \beta_0 + \beta_1 \log(sales) + \beta_2\, roe + \beta_3\, investfund + \\ &+ \beta_4\, accounting + \beta_5\, banking + u\end{aligned}$$

Problem 5. Total 10 marks

A financial analyst wants to understand if information observable to the US stock market prior to quarter t helps predict the return during quarter t. One simple way to check this is to estimate the model:

$$return_t = \alpha + \rho_1\, return_{t-1} + \rho_2\, return_{t-2} + v_t, \quad (4)$$

where v_t is a white noise process with $E(v_t) = 0$ and $Var(v_t) = \sigma_v = constant$.

(i) What is the null hypothesis that needs to be tested in order to check that information prior to quarter t has no impact on the return during quarter t?

(1 Mark)

Answer: $H_0 : \rho_1 = \rho_2 = 0$

(ii) What is the name of equation (4)?

(1 Mark)

Answer:

$AR(2)$

(iii) What are the three necessary requirements needed to conclude that $return$ is covariance stationary?

(2 Marks)

Answer:

$E(return_t) = const., Var(return_t) = const., Cov(return_t, return_{t-h}) = \gamma_h = const.$ (can depend on h but not on t)

(iv) Assuming $return$ is covariance-stationary derive $Var(return_t)$.

(2 Marks)

Answer:

Use the fact that $Var(return_t) = Var(return_{t-h}) = \sigma_r^2$ for any h. Simply apply the formula for the variance of sums of random variables and take into account the properties of v_t. The final answers also depends on $Cov(return_{t-1}, return_{t-2}) = \gamma_1 = const.$

(v) The analyst knows that in many financial studies the variance of returns depends on past returns. The analyst decides to check this for his US data using the Breusch-Pagan test. Write the equation (also known as artificial regression) and the null/alternative hypothesis.

(2 Marks)

Answer:

See slides. You are expected to use the notation in this exercise when steps are described.

(vi) How should the analyst test for seasonality in equation (4)? Write the equation and the null/alternative hypothesis.

(2 Marks)

Answer:

Create 3 dummy variables for 3 of the quarters and add them to AR(2) model. The dummy for the fourth quarter is not included to avoid the perfect multicollinearity trap. The null hypothesis refers to the joint nullity of the coefficients associated with the 3 dummy variables.

END OF EXAM