Brunel University

School of Social Sciences

Mid-Term Exam

EC5501-EC5509

2011-2012

Answer one question.

Each part of the question carries equal weight

1. a) Consider the bivariate regression $\widehat{Y}_i = \alpha + \beta X_i$, where $N = 5$, $\overline{Y} = 8$, $\overline{X} = 4$, $\sum xy = 70$, $\sum x^2 = 40$, $\sum y^2 = 124$.

i) Obtain the regression coefficients α and β.

ii) Calculate the explained and residual sum of squares as well as the correlation coefficient between Y and X (r)

(use the form: $\sum y^2 = \sum \widehat{e}^2 + r^2 \sum y^2$).

iii) Obtain the estimated standard error of the regression.

b) i) Explain the White, Breusch-Pagan and ARCH test statistics for heteroscedasticity.

ii) Consider the classical linear regression model: $\underset{(Nxk)}{\mathbf{X}}\underset{(kx1)}{\mathbf{b}} + \mathbf{e}$. What is the formula for the variance-covariance matrix $\underset{(kxk)}{\mathbf{V}}(\beta)$ of the $(k \times 1)$ least-squares $\underset{(kx1)}{\boldsymbol{\beta}}$ vector under heteroscedasticity?

c) i) Write four factors that affecting housing prices and explain how one can use regression analysis to test for the effect of each of these factors on housing prices.

ii) Consider the following regression:

$$log(\text{wage}) = 0.284 - \underset{(0.007)}{0.092}(\text{educ}) + \underset{(0.0017)}{0.0041}\ (\text{exper}) + \underset{(0.003)}{0.022}(\text{tenure}),\ n = 526,$$

where standard errors are in parentheses. Test the null hypothesis that years of labour market experience have a significant effect on hourly wages.

d) Show that if r is the correlation coefficient between n pairs of variables (X_i,Y_i), then the squared correlation between the n pairs ($aX_i + b$, cY_{i+d}), where $a; b; c$ and d are constants, is also r^2.

2) a) i) Consider the classical linear regression model: $\underset{(Nxk)}{\mathbf{X}}\underset{(kx1)}{\mathbf{b}} + \mathbf{e}$. What is the formula for the $(k \times 1)$ least-squares $\underset{(kx1)}{\boldsymbol{\beta}}$ vector?

ii) In the two variable regression: $Y_i = b_1 + b_2 X_i + e_i$, use your result in i) to derive the least-squares estimates β_1 and β_2.

b) i) Consider the classical linear regression model: $\underset{(Nxk)}{\mathbf{X}}\underset{(kx1)}{\mathbf{b}} + \mathbf{e}$. What is the formula for the variance-covariance matrix $\underset{(kxk)}{\mathbf{V}}(\beta)$ of the $(k \times 1)$ least-squares $\underset{(kx1)}{\boldsymbol{\beta}}$ vector?

ii) In the two variable regression: $Y_i = b_1 + b_2 X_i + e_i$, use your result in i) to derive the $V(\beta_2)$ and the $Cov(\beta_1, \beta_2)$.

c) From a sample of 200 observations the following quantities were calculated:

$$\sum X = 11.34, \ \sum Y = 20.72, \ \sum x^2 = 11.52$$
$$\sum y^2 = 82.81, \ \sum xy = 20.95.$$

Estimate both regression equations: the regression of Y on X, and the regressions of X on Y.

d) i) Write three factors that affecting hourly wages and explain how one can use regression analysis to test for the effect of each of these factors on hourly wages.

ii) Consider the following regression:

$$log(\text{price}) = \underset{}{11.08} - \underset{(0.117)}{0.954} log(\text{nox}) - \underset{(0.043)}{0.134} \log(\text{dist}) + \underset{(0.019)}{0.255}(\text{rooms}) - \underset{(0.006)}{0.052}(\text{stratio}),$$

where standard errors are in parentheses and $n = 506$. Test the null hypothesis that the elasticity of housing prices with respect to air pollution is -1.

3) a) The following sums (in deviation form) were obtained from 10 sets of observations on Y, X_1, and X_2:

$$\sum y^2 = 48.2, \;\; \sum x_1^2 = 2, \sum x_2^2 = 3,$$
$$\sum x_1 x_2 = -1, \sum y x_1 = -1, \;\; \sum y x_2 = 8.$$

The standard error of the regression is: $s^2 = 3.6$. Estimate the regression of Y on X_1 and X_2, including an intercept term, and test the hypothesis that the coefficient of X_2 is zero.

(The 5% critical value for the t distribution with 7 degrees of freedom is 2.365).

b) Explain the Durbin-Watson test and the Box-Pierce statistic for serial correlation.

c) i) Let β_{xy}be the estimated slope coefficient from the regression of X on Y. Write an equation that relates r_{yx} (the sample correlation coefficient), β_{xy} and β_{yx}.

ii) Assume that Y is a linear combination of X. Prove that ρ (the theoretical correlation coefficient) between the two variables is 1.

d) Explain what is: i) the probability of type I error; ii) the 95% confidence interval; iii) the p value;) the probability of type II error; v) the power of a test.