ASE, Spring Class exercises.

Ron Smith: r.smith@bbk.ac.uk

Spring term 2020

EMS, Birkbeck: GDE & BScs

1. Introduction

Try the class exercises before the class or you will not know what your difficulties are.

2. Week 2, Estimation and testing

Suppose real returns on the S&P500 stock market index, Y_t, in year $t = 1, 2, ..., T$, are determined as $Y_t = \alpha + u_t$where $E(u_t) = 0$, $E(u_t^2) = \sigma^2$, $E(u_t u_{t-i}) = 0$, $i \neq 0$.

Consider an estimator of $\alpha : \widehat{\alpha} = T^{-1} \sum_{t=1}^{T} Y_t$ and an estimator of $\sigma : s = \sqrt{\sum_{t=1}^{T} \widehat{u}_t^2/(T-1)}$; where $\widehat{u}_t = Y_t - \widehat{\alpha}$.

Real returns 1917-2016 had $T = 100$; $\widehat{\alpha} = 2.3$ $s = 19.2$.

(a). Show that $\widehat{\alpha}$ (a) makes $\sum_{t=1}^{T} \widehat{u}_t = 0$ (b) makes $\sum_{t=1}^{T} \widehat{u}_t^2$ as small as possible.

(b). Show $\widehat{\alpha}$ is unbiased. What is its standard error?

(c). How would you forecast Y_{T+1}?

(d). For large T, how would you construct a 95% confidence interval for $\widehat{\alpha}$.

(e). Use the information on the real return on the S&P 1917-2016 test $H_0 : \alpha = 0$ against $H_1 : \alpha \neq 0$.

(f) Explain Type I and Type II errors. How did you deal with these two types of error in your answer to (e)?

3. Week 3 Regression

A dependent variable, Y_t, in year $t = 1, 2, ..., T$, is determined by an exogenous independent variable X_t :

$$Y_i = \alpha + \beta X_t + u_i$$

where $E(u_t) = 0$, $E(u_t^2) = \sigma^2$, $E(u_t u_{t-i}) = 0$, $i \neq 0$. Denote estimators of α and β by $\widehat{\alpha}$ and $\widehat{\beta}$ and denote $\widehat{u}_t = Y_t - \widehat{\alpha} - \widehat{\beta} X_t$. Define $y_t = Y_t - \overline{Y}$, where $\overline{Y} = T^{-1} \sum_{t=1}^{T} Y_t$ and similarly for x_t.

(a). Show that $\widehat{\alpha} = Y - \widehat{\beta X}$ makes $\sum_{t=1}^{T} \widehat{u}_t = 0$.

(b) Show that

$$\widehat{\beta} = \frac{T^{-1} \sum_{t=1}^{T} x_t y_t}{T^{-1} \sum_{t=1}^{T} x_t^2}$$

(i) makes $\sum_{t=1}^{T} x_t \widehat{u}_t = 0$ and (ii) makes $\sum_{t=1}^{T} \widehat{u}_t^2$ as small as possible.

(c). Show $\widehat{\beta}$ is unbiased and derive its standard error.

(d). How would you estimate σ^2?

(e). How would you forecast Y_{T+1}, given a value for X_{t+1}?

(f). For large T, how would you construct a 95% confidence interval for $\widehat{\beta}$.

4. Week 4 Life Expectancy

For a sample of countries, $i = 1, 2, ..., 189$ life expectancy, LE, measured in years, in 2013, was regressed on the (natural to base e) logarithm of per capita income, LPCI, measured in purchasing power parity, PPP, inflation adjusted international dollars. The results (with standard errors in parentheses) are:

$$\begin{array}{rlll} LE_i = & 24.28 & +5.17 & LPCI_i \quad +\widehat{u}_i \\ & (2.57) & (0.28) & \end{array}$$

With coefficient of determination $R^2 = 0.646$ and standard error of regression $SER = 4.8$.

(a) Explain how R^2, SER and are $\widehat{u}_t$ calculated and what they measure.

(b) Which coefficients are significantly different from zero?

(c) If per-capita income increases by 10%, how much would the equation predict that life expectancy would increase?

(d) Equatorial Guinea had a life expectancy of 58 years and per-capita income of \$37,741. What is the life expectancy predicted by this regression and the residual?

(e) When the log of life expectancy was used as the dependent variable, the results were:

$$\begin{array}{lccccc} LLE_i = & 3.58 & +0.07 & LPCI_i & +\widehat{u}_i \\ & (0.038) & (0.004) & & \end{array}$$

With $R^2 = 0.623$ and $SER = 0.073$. Does the fact that the R^2 of this equation is lower indicate that it is a worse equation? How would you interpret the SER?

(f) When the dependent variable was LE but log per-capita income, LPCI was replaced by per-capita income, PCI, $R^2 = 0.373$, $SER = 6.4$. Does the fact that the R^2 of this equation is lower indicate that it is a worse equation?

5. Week 5 Multiple regression and the properties of least squares

(a) What are the assumptions required for the Ordinary Least Squares, OLS, estimator to (i) be unbiased (ii) to have minimum variance among linear unbiased estimators (iii) for the standard errors to be correct..

(b) Explain what heteroskedsticity and serial correlation are. How might you recognise them in a plot of the residuals.

(c) Explain what exogeneity means, give examples of how this assumption might fail.

(d) Suppose the true model was $y_i = \beta x_i + \gamma x_i^2 + e_i$, but you estimated $y_i = bx_i + u_i$, what would be the consequences for your estimate of the coefficient of x_i.

(e) In the model $y_i = \beta x_i + \gamma x_i^2 + e_i$ how would you measure the effect of a change in x?

6. Week 6, Reading week

There will be no lecture or tutorial class, instead there will be computer classes in room 742 to introduce you to statistical programs.

7. Week 7 Matrix form for linear regression

Consider the model

$$y = X\beta + u,$$

where y is an $T \times 1$ vector of observations on a dependent variable; X is a $T \times k$, rank k, matrix of observations on exogenous independent variables; β is a $k \times 1$ vector of unknown parameters; and u is a $T \times 1$ vector of unobserved disturbance terms with $E(u) = 0$, $E(u\ u') = \sigma^2 I$

(a) Suppose $k = 2$

$$X = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ .. & .. \\ 1 & X_T \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_2 \\ .. \\ u_T \end{bmatrix}$$

what are (i) $X'X$, (ii) $X'u$, (iii) $u\ u'$, (iv) $u'u$?

(b) Explain the meaning of the assumptions that X is (i) exogenous (ii) of rank k. What is the consequence of the failure of each of them?

(c) Explain the meaning of the assumption $E(u\ u') = \sigma^2 I$. What is the consequnce of its failure?

(d) Derive the method of moments estimator $\widehat{\beta}$ that makes $X'\widehat{u} = 0$, where $\widehat{u} = y - X\widehat{\beta}$.

(e) Derive the Ordinary Least Squares, OLS, estimator, $\widehat{\beta}$, that minimises $\widehat{u}'\widehat{u}$.

8. Week 8 Properties of Least Squares in Matrix form

In the model for Week 7.

(a) Show that the OLS estimator, $\widehat{\beta}$, is unbiased. What assumptions are required for this?

(b) Derive the variance covariance matrix of $\widehat{\beta}$. What assumptions are required for this?

(c) How would you estimate the standard errors of the individual coefficient estimates $\widehat{\beta}_i$, $i = 1, 2, ..., k$.

(d) Suppose $E(u\ u') = \sigma^2 \Omega$, what is the variance covariance matrix of $\widehat{\beta}$ in this case?

9. Week 9, Estimates

It is believed that an energy demand equation takes the form:

$$q_t = \alpha + \beta y_t + \gamma p_t + \varepsilon_t,$$

where q_t is the logarithm of per capita energy demand in year t; p_t the logarithm of real energy prices; y_t the logarithm of per-capita real GDP; ε_t is a well behaved disturbance term. The following estimates (with standard errors in parentheses) were obtained using data for the period t =1974-1990.

	β	γ	R^2	SER
India	1.006	-0.068	0.38	0.027
	(0.102)	(0.080)		
Indonesia	1.564	-0.488	0.52	0.034
	(0.234)	(0.195)		
Korea	1.074	-0.136	0.54	0.031
	(0.125)	(0.189)		

SER is the standard error of the regression.

(a) How would you interpret β and γ?

(b) Explain what R^2 and SER are and what they tell you. How would you interpret the fact that while Korea has the largest R^2 it does not have the lowest SER?

(c) For Indonesia, test (i) the hypothesis $\beta = 1$ and (ii) the hypothesis $\gamma = 0$.

(d) Interpret the stochastic component of the model. How would you estimate it?

(e) Suppose that you believed that there was a trend increase in energy efficiency in these countries. How would you adjust the model to allow for this.

(f) Why might the exogeniety assumption fail in this model?

10. Week 10, Tests

The regression below uses data on 159 countries in 2010, LE is life expectancy in years; LPPT is the log of physicians per thousant people; LYPC$ is log per-capita income in US dollars at market exchange rates; LYPC is log percapita income at PPP international dollars.

Dependent Variable: LE
Method: Least Squares
Date: 12/19/18 Time: 13:02
Sample: 1 159
Included observations: 159

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	43.44386	5.110221	8.501365	0.0000
LPPT	1.976569	0.371373	5.322324	0.0000
LYPC$	1.844996	0.894236	2.063208	0.0408
LYPC	1.364243	1.207985	1.129354	0.2605

R-squared	0.753444	Mean dependent var	71.30836
Adjusted R-squared	0.748672	S.D. dependent var	8.259025
S.E. of regression	4.140463	Akaike info criterion	5.704328
Sum squared resid	2657.232	Schwarz criterion	5.781533
Log likelihood	-449.4940	Hannan-Quinn criter.	5.735680
F-statistic	157.8872	Durbin-Watson stat	2.043228
Prob(F-statistic)	0.000000		

(a) Are the coefficents of LYPC$ and LYPC individually significant (i) at the 5% level (ii) at the 1% level? A test of their joint significance gave a F(2,155) p value of 0.0000. Explain the difference between the hypotheses in the individual and joint tests.

(b) (i) What hypothesis is tested by the F-statistic in the last 2 rows of the left hand column? What are its degrees of freedom?

(c) The data are ordered by per-capita income. A Chow test for the hypothesis that the coefficients for the first 80, low income, observations were the same as the following 79, high income, observations gave a F(4,151) p value of 0.1509. Are the determinants of LE the same in high and low income countries? Explain the degrees of freedom of the F statistic.

(d) A test that the residuals had skewness of zero and kurtosis 3 gave a p value of 0.000. The F-statistic in a regression of the squared residuals on the regressors gave a p value of 0.0533. A RESET test involving a regression of the residuals on the squared fitted values gave a p value of 0.881. These diagnostic tests are testing for 3 possible failures of the assumptions. What are the possible failures? Is there evidence of any of these failures?

(e) It is argued that it is a mistake to include the two measures of per-capita income since they will be correlated and this will cause multicollinearity. Do you agree?

11. Week 11, Dynamic Models.

Using US data on on the logarithms of company earnings (profits) LE and dividends paid out to shareholder, LD, the following results were obtained

Dependent Variable: LD
Method: Least Squares
Date: 12/19/18 Time: 13:57
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.133988	0.016276	-8.232040	0.0000
LD(-1)	0.651616	0.032320	20.16136	0.0000
LE	0.203134	0.025576	7.942294	0.0000
LE(-1)	0.115419	0.032928	3.505239	0.0006

R-squared	0.997099	Mean dependent var	0.364613
Adjusted R-squared	0.997036	S.D. dependent var	1.578354
S.E. of regression	0.085925	Akaike info criterion	-2.043109
Sum squared resid	1.026254	Schwarz criterion	-1.960232
Log likelihood	150.0823	Hannan-Quinn criter.	-2.009431
F-statistic	15924.80	Durbin-Watson stat	1.805160
Prob(F-statistic)	0.000000		

Dependent Variable: D(LD)
Method: Least Squares
Date: 12/19/18 Time: 13:55
Sample (adjusted): 1872 2014
Included observations: 143 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.133988	0.016276	-8.232040	0.0000
LD(-1)	-0.348384	0.032320	-10.77922	0.0000
D(LE)	0.203134	0.025576	7.942294	0.0000
LE(-1)	0.318554	0.028591	11.14183	0.0000

R-squared	0.526412	Mean dependent var	0.035118
Adjusted R-squared	0.516191	S.D. dependent var	0.123533
S.E. of regression	0.085925	Akaike info criterion	-2.043109
Sum squared resid	1.026254	Schwarz criterion	-1.960232
Log likelihood	150.0823	Hannan-Quinn criter.	-2.009431
F-statistic	51.50140	Durbin-Watson stat	1.805160
Prob(F-statistic)	0.000000		

Denoting LE as, e_t, and LD as, d_t, the equations can be written

$$d_t = \alpha_0 + \alpha_1 d_{t-1} + \beta_0 e_t + \beta_1 e_{t-1} + u_t \quad (11.1)$$

$$\Delta d_t = a_0 + a_1 d_{t-1} + b_0 \Delta e_t + b_1 e_{t-1} + u_t \quad (11.2)$$

(a) (i) What is the relationship between the two equations? (ii) It is argued that the first is better because the R^2 is higher. Do you agree?

(b) Calculate the long-run elasticity of dividends with respect to earnings for each case.

(c) The parameter $\phi = \alpha_1 + \beta_1 + \beta_2 - 1$ in (11.1) is estimated as 0.029831 with a standard error of 0.005738. Test $H_0 : \phi = 0$ and interpret the hypothesis.

(d) The parameter $\theta = -b_2/a_2$ in (11.2) is estimated as 0.914374 with a standard error of 0.012283. Test $H_0 : \theta = 1$ and interpret the hypothesis.

(e) A test for second order serial correlation in (11.2) had a p value of 0.0672. Explain what second order serial correlation is and why it is a problem. Is it a problem in this case?