

Tutorial 12!

Hong Xiang Yue

23/05/2019

Part A

Question 2

$$y_t = \alpha + \beta t + u_t$$

- ▶ Linear trend

$$u_t = u_{t-1} + e_t$$

- ▶ Error term is a unit root process (Random Walk)

What is $\Delta y = y_t - y_{t-1}$?

Question 3a,b

An Estimator

- ▶ Function of the data which calculates an estimate of a population parameter
- ▶ Is a random variable because calculated estimates will change depending on the sample
- ▶ $\hat{\beta}$ is an estimator of β

An Unbiased Estimator

- ▶ Estimator whose expected value is equal to the true population parameter
- ▶ It is on average “correct”, not biased
- ▶ If assumptions E.1, E.2 and E.3 hold, then $E[\hat{\beta}] = \beta$

Question 3c

The Best Linear Unbiased Estimator (BLUE)

- ▶ Out of all possible linear unbiased estimators, it has the smallest possible variance
- ▶ Most efficient or precise
- ▶ OLS is BLUE if E.1, E.2, E.3 and E.4 hold

Question 3d

A Consistent Estimator

- ▶ As the sample size increases, the estimator will become more and more precise
- ▶ In the limit the estimator will be equal to the parameter it is trying to estimate
- ▶ As $n \rightarrow \infty$ we say that $\bar{X} \xrightarrow{p} \mu$
- ▶ This also applies to OLS when assumptions E.1 and E.2 hold and $E[u_i | x_{i,1}, x_{i,2}, \dots, x_{i,k}] = 0$
- ▶ The last assumption is a less strict version of E.3: $E[\mathbf{u} | \mathbf{X}] = \mathbf{0}$

Question 3e

An Asymptotically Normal Estimator

- ▶ As the sample size increases, the estimator will have an approximately normal distribution
- ▶ The sample mean is asymptotically normal regardless of the underlying distribution for the data
- ▶ $\bar{X} \overset{asy}{\sim} N(\mu, \sigma^2/n)$
- ▶ $X \sim N(\mu, \sigma^2), X \sim F(p, q), X \sim \Gamma(\alpha, \beta)$
- ▶ This result is also known as the central limit theorem
- ▶ http://www.onlinestatbook.com/stat_sim/sampling_dist/

Part B

Question 1

Assume that y_t is stationary, following:

$$y_t = c + \phi_1 y_{t-1} + e_t, \quad e_t \sim i.i.d.(0, \sigma^2)$$

Show that:

1. $E(y_t) = \mu = \frac{c}{1-\phi_1}$
2. $Var(y_t) = \gamma_0 = \frac{\sigma^2}{1-\phi_1^2}$
3. $y_t - \mu = \phi_1(y_{t-1} - \mu) + e_t$
4. $Cov(y_t, y_{t-1}) = \gamma_1 = \phi_1 \frac{\sigma^2}{1-\phi_1^2}$
5. $Corr(y_t, y_{t-1}) = \rho_1 = \phi_1$

Question 2: Efficient Markets Hypothesis

Returns in the past should not be able to predict returns in this period.

$$E[r_t | r_{t-1}, r_{t-2}, \dots] = E[r_t]$$

A researcher estimates an AR(1) model on weekly returns on the NYSE composite index:

$$\hat{r}_t = \underset{(0.096)}{0.086} - \underset{(0.038)}{0.059} r_{t-1}$$

$$n = 689, R^2 = 0.0035, \bar{R}^2 = 0.0020$$

1. Testing for the EMH

- ▶ What would the null hypothesis be if the EMH holds?
- ▶ What is the conclusion of this test?

Question 2: Efficient Markets Hypothesis

2. The model could be potentially misspecified and that returns in this week, depend on returns more than a week ago
 - ▶ If the model was correctly specified, what type of process would u_t follow? What are its properties?
 - ▶ If the researcher suspected that returns from 3 and 4 weeks ago could add explanatory power, what would he be expected to see in u_t ?

Question 2: Efficient Markets Hypothesis

3. The researcher is also investigating the behaviour of the squared residuals because he is not sure that the variance given past information is constant. He runs the following regression

$$\hat{u}_t^2 = \underset{(0.43)}{4.66} - \underset{(0.201)}{1.104} \times r_{t-1}$$

$$n = 689, R^2 = 0.0042$$

- ▶ What problem is the researcher worried about?
- ▶ Define the problem and conduct a formal test which makes use of the regression's goodness-of-fit, what is the conclusion?
- ▶ What advice would you give the researcher based on the conclusion of the test?

Question 3a

The researcher is considering two models for a stationary time series $\{y_t\}$, an AR(2) and an ARDL(2,1) model:

$$\hat{y}_t = 1.28 - \frac{0.31}{(0.09)} \times y_{t-1} - \frac{0.39}{(0.08)} \times y_{t-2}$$

$$n = 200, \bar{R}^2 = 0.55, SSR = 475, AIC = 1.08, BIC = 1.09$$

$$\hat{y}_t = 1.30 - \frac{0.42}{(0.44)} \times y_{t-1} - \frac{0.37}{(0.08)} \times y_{t-2} - \frac{2.64}{(0.46)} \times x_{t-1}$$

$$n = 200, \bar{R}^2 = 0.71, SSR = 462, AIC = 1.04, BIC = 1.11$$

Which model do you prefer? Why?

Question 3b

The researcher also estimated a model of the form

$$y_t = c + \beta_0 D_t + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + \beta_2 D_t x_{t-1} + u_t$$

$$D_t = \begin{cases} 1 & \text{for } t = 1, 2, \dots, 100 \\ 0 & \text{for } t = 101, 102, \dots, 200 \end{cases}$$

The estimated output is:

$$\hat{y}_t = \underset{(0.44)}{1.32} + \underset{(0.04)}{0.30} D_t - \underset{(0.07)}{0.39} y_{t-1} - \underset{(0.06)}{0.31} y_{t-2} + \underset{(0.41)}{2.15} x_{t-1} + \underset{(0.07)}{0.15} D_t x_{t-1}$$

$$SSR = 450, \bar{R}^2 = 0.69, n = 200$$

Question 3b

1. What is $\hat{E}[y_t|y_{t-1}, y_{t-2}, x_{t-1}]$ for different values of t ?
2. What are the immediate and long run effects for a one unit increase in x on y , before and after $t = 100$
3. Test the null hypothesis that there is no structural break on either the slope or the intercept for x_{t-1}

Question 3c

Carefully describe the steps to conduct a Breusch-Godfrey test for autocorrelation up to the second order in the error term of the ARDL(2,1) model

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \beta_1 x_{t-1} + u_t$$