Econometrics II

Written examination

Tuesday June 04, 2019, 15:00 - 20:00

Examiner: Andreas Rosenblad, Department of Statistics, Stockholm University

Instructions

Allowed tools:

- Pocket calculator
- Text book: Wooldridge, J.M. *Introductory Econometrics: A Modern Approach.* Cengage Learning, Boston.
- Text book: Montgomery, D.C., Jennings, C.L., and Kulachi, M. *Introduction to Time Series Analysis and Forecasting.* John Wiley, New Jersey.
- Notes written in the text books are allowed.

Note that no formula sheet is provided.

Passing rate: 50% of overall total, which is 100 points. For detailed grading criteria, see the course description.

The maximum number of points for each problem is given in the right margin. If not indicated otherwise, to obtain the maximum number of points on each problem, detailed and clear solutions are required. Answers may be given in English or Swedish.

Solutions to the exam questions will be uploaded to Athena on the afternoon of May 02, 2019. The corrected exams will be available at the student office of the Department of Statistics three weeks after the date of the exam.

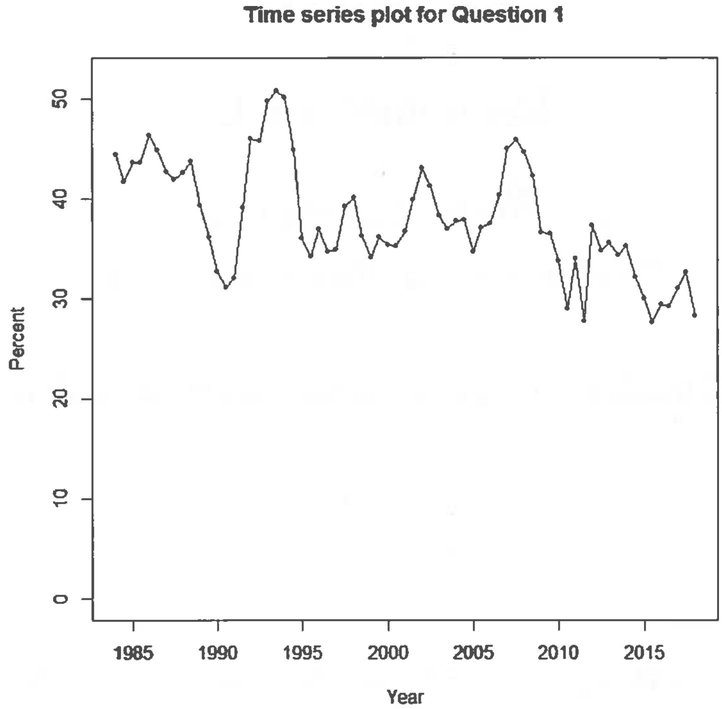

Figure 1: Percent of individuals that would vote for the Swedish Social Democratic Party according to Statistics Sweden's Party Preference Survey (PSU) 1984-2018

Question 1 (22 points)

Statistics Sweden has performed the Party Preference Survey (*Partisympatiundersökningen*, PSU), giving the election results for Sweden "if an election were to be held today", twice a year (May and November) since 1984. Figure 1 gives a time series plot of the $N = 69$ values of y_t, the percent of individuals that would vote for the Swedish Social Democratic Party at each time point, from May 1984 to May 2018.

(a) Using the `auto.arima()` function in the R package `forecast` to estimate an $ARIMA(p,d,q)$ model for y_t suggests that y_t is an $ARIMA(3,1,1)$ process with $\phi_1 = 0.817$, $\phi_2 = 0.101$, $\phi_3 = -0.312$, and $\theta_1 = 0.872$. Write this $ARIMA$ process as an $ARMA(p,q)$ process in equation form, using $\delta = 0$. (8)

(b) We are interested in estimating this time series using simple exponential smoothing (SES) with discount factor $\lambda = 0.8$. The first three values of the time series (i.e. November 1984, May 1985, and November 1985) are $y_1 = y_{1984Nov} = 44.4$, $y_2 = y_{1985May} = 41.7$, and $y_3 = y_{1985Nov} = 43.7$. Using the initial value $\tilde{y}_0 = y_1$, calculate the smoothed (i.e. local mean) values $\tilde{y}_1$, $\tilde{y}_2$, and $\tilde{y}_3$. (8)

(c) In the Swedish General Election (*Riksdagsvalet*) September 9, 2018 (*SGE2018*), the Swedish Social Democratic Party received 28.3% of the votes. Denote this value as $y_{SGE2018} = 28.3$. Utilizing the simple exponential smoothing (SES) method with discount factor $\lambda = 0.8$ applied to the PSU data, we can use the one-step-ahead forecast of y_T made in May 2018 as the one-step-ahead forecasted value of $y_{SGE2018}$, denoted $\hat{y}_{SGE2018}(2018May)$. This results in $\hat{y}_{SGE2018}(2018May) = 29.1$. Calculate the mean absolute percent forecast error (MAPE) of the one-step-ahead forecast $\hat{y}_{SGE2018}(2018May)$. (6)

Question 2 (34 points)

A stationary times series process y_t is given by the equation

$$y_t = 50 + 0.75y_{t-1} - 0.15y_{t-2} + \varepsilon_t$$

where ε_t is a white noise process.

(a) What is this time series process called? (4)

(b) Express this time series process in terms of the backshift operator B. (6)

(c) Show or argue that y_t is a stationary process. (6)

(d) Calculate $E(y_t)$. (6)

(e) Assuming that $\sigma_\varepsilon^2 = \sigma^2 = 1$, $\gamma_y(1) = \gamma(1) = 1.5$, and $\gamma_y(2) = \gamma(2) = 1.75$, calculate the value of $\gamma_y(0)$. (6)

(f) If ε_t is a Gaussian white noise process with $E(\varepsilon_t) = 0$, the current observation of the process is $y_{100} = 60$, and the previous observation was 55, which value would be expected for the next observation? (6)

Question 3 (8 points)

A stationary times series process y_t is given by the equation

$$y_t = \varepsilon_t - 1.5\varepsilon_{t-1} + 0.6\varepsilon_{t-2} +$$

where ε_t is a white noise process.

(a) What is this time series process called? (4)

(b) Show or argue that y_t is a stationary process. (4)

Question 4 (36 points)

Give the correct answer for the following multiple-choice questions. No motivation is needed.

(a) What is the estimator obtained through regression on quasi-demeaned data called? (3)

A. The fixed effects estimator.

B. The mixed effects estimator.

C. The pooled OLS estimator.

D. The random effects estimator.

(b) For a white noise error term ε_t, what is the time series process y_t given by (3)

$$y_t = \varepsilon_t + 0.2\varepsilon_{t-1} - 0.4\varepsilon_{t-2} + 0.6\varepsilon_{t-3}$$

called?

A. An $ARIMA(0,0,3)$ process.
B. An $MA(4)$ process.
C. An $ARMA(3,0)$ process.
D. An $ARIMA(3,0,0)$ process..

(c) What is the equation obtained by differencing each variable in a single cross-sectional equation over time called? (3)

A. The difference-in-differences (DiD) equation.
B. The difference effects equation.
C. The first-differenced equation.
D. The difference-stationary equation.

(d) Which of the following cases describes a situation were two time series are said to be cointegrated? (3)

A. Both series are I(1) but a linear combination of them is I(0).
B. Both series are I(0) but a linear combination of them is I(1).
C. Both series are I(1) and a linear combination of them is I(I).
D. Both series are I(0) and a linear combination of them is I(0).

(e) For time series data on two variables y and z, where y_t and z_t are dated contemporaneously, what is the model (3)

$$y_t = \beta_0 + \beta_1 z_t + u_t, \quad t = 1, 2, ..., n$$

called?

A. An autoregressive conditional heteroskedasticity (ARCH) model.
B. A static model.
C. A finite distributed lag model.
D. An infinite distributed lag model.

(f) Which of the following statements identifies an advantage of using first differencing on time series data? (3)

A. First differencing eliminates the possibility of spurious regression.
B. First differencing eliminates most of the multicollinearity.
C. First differencing eliminates most of the serial correlation.
D. First differencing eliminates most of the heteroskedasticity.

(g) For a given significance level α, what is the correct decision if the observed value of the Durbin-Watson test statistic d lies between the lower and upper critical values? (3)

A. The null hypothesis of no serial correlation is rejected.
B. The null hypothesis of no serial correlation is not rejected.
C. The null hypothesis of serial correlation is rejected.
D. The test is inconclusive.

(h) Which of the following statements about dynamically complete models is true? (3)

A. There is scope of adding more lags to the model to better forecast the dependent variable.
B. The problem of serial correlation does not exist in dynamically complete models.
C. All econometric models are dynamically complete.
D. Sequential endogeneity is implied by dynamic completeness.

(i) Which of the following statements describes a problem with judgment forecasts? (3)

A. They give too little attention to recent events.
B. They are often too pessimistic.
C. The are known for overestimating the variability.
D. They give too much attention to recent events.

(j) If T periods of data have been used to fit a model with p parameters and e_t is the residual from the model-fitting process in period t, what is the value calculated by the formula (3)

$$\ln\left(\frac{1}{T}\sum_{t=1}^{T} e_t^2\right) + \frac{2p}{T}$$

called?

A. The logarithmic mean squared error (LMSE).
B. The log-likelihood (LL) function.
C. The Akaike Information Criteria.
D. The mean logarithmic squared error.

(k) Which of the following four statements is true? (3)

A. The forecast horizon is also called the forecast lead time.
B. The forecast interval is also called the forecast horizon.
C. The prediction interval is also called the forecast horizon.
D. A confidence inteval for a time series is sometimes called a forecast horizon.

(l) Which of the following four models may be written as (3)

$$y_t = \left(1 - \theta_2 B^2 - \theta_4 B^4\right)\varepsilon_t$$

where B is the backshift operator and ε_t is a white noise error term?

A. $y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \theta_3\varepsilon_{t-3} - \theta_4\varepsilon_{t-4}$
B. $y_t = \mu - \theta_2\varepsilon_{t-2} - \theta_4\varepsilon_{t-4}$
C. $y_t = \mu + \varepsilon_t - \theta_2\varepsilon_{t-2} - \theta_4\varepsilon_{t-4}$
D. $y_t = \varepsilon_t - \theta_2\varepsilon_{t-2} - \theta_4\varepsilon_{t-4}$

Department of Statistics

Stockholms universitet

Correction sheet

Date: 190604

Room: Brunnsvikssalen

Exam: Econometrics II

Course: Econometrics

Anonymous code: 0030-EDO

☑ I authorise the anonymous posting of my exam, in whole or in part, on the department homepage as a sample student answer.

NOTE! ALSO WRITE ON THE BACK OF THE ANSWER SHEET

Mark answered questions

1	2	3	4	5	6	7	8	9	Total number of pages
✓	✓	✓	✓						3
Teacher's notes 14	34	8	36						

Points	Grade	Teacher's sign.
92	A	AR

SU, DEPARTMENT OF STATISTICS

Room: Brunnsvikssalen **Anonymous code:** 0030-EDO **Sheet number:** 7

Question 1

a) ARMA(3,1) ⟶ $Y_t = 0.817 Y_{t-1} + 0.101 Y_{t-2} - 0.312 Y_{t-3} + \varepsilon_t + 0.872 \varepsilon_{t-1}$ /0

b) $\lambda = 0.8$ $\tilde{Y}_0 = Y_1$

$Y_1 = 44.4$ $\tilde{Y}_1 = \lambda Y_1 + (1-\lambda)\tilde{Y}_0 = 0.8 \times 44.4 + 0.2 \times 44.4 = 44.4$

$Y_2 = 41.7$ $\tilde{Y}_2 = \lambda Y_2 + (1-\lambda)\tilde{Y}_1 = 0.8 \times 41.7 + 0.2 \times 44.4 = 42.24$

$Y_3 = 43.7$ $\tilde{Y}_3 = \lambda Y_3 + (1-\lambda)\tilde{Y}_2 = 0.8 \times 43.7 + 0.2 \times 42.24 = 43.408$ /8

c) Absolute %error ⟶ $= (e_t(1)/Y_t)100$

$Y_{SUE2018} = Y_1 = 28.3$

$\hat{Y}_{SUE2018} = \hat{Y}_1 = 29.1$

$e_1(1) = Y_1 - \hat{Y}_1 = -0.8$ $|e| = 0.8$

Absolute (%) error $= (0.8/28.3)100 \approx 2.8269$

$MAPE = \frac{1}{n}\sum |re_t(1)|$ ⟶ eftersom vi enbart har ett värde (n=1) blir $MAPE = |re_t(1)| = 2.8269$ /6

SU, DEPARTMENT OF STATISTICS

Room: Brunnsvikssalen **Anonymous code:** 0030-EDO **Sheet number:** 2

Question 2

a) Second-order Autoregressive process, AR(2) /4

b) $(1 - 0.75B + 0.15B^2) Y_t = 50 + \varepsilon_t$ /6

c) En AR(2)-modell är stationär om följande gäller

$\phi_1 + \phi_2 < 1 \rightarrow 0.75 + (-0.15) = 0.6 \rightarrow 0.6 < 1$ ✓

$\phi_2 - \phi_1 < 1 \rightarrow -0.15 - 0.75 = -0.9 \rightarrow -0.9 < 1$ ✓

$|\phi_2| < 1 \rightarrow 0.15 < 1$ ✓

} stationär! /6

d) $E(Y_t) = \delta + \phi_1 E(Y_{t-1}) + \phi_2 E(Y_{t-2}) + 0 \leftarrow E(\varepsilon_t) = 0$

$\mu = \delta + \phi_1 \mu + \phi_2 \mu$

$\rightarrow \mu = \frac{\delta}{1 - \phi_1 - \phi_2} = \frac{50}{1 - 0.75 + 0.15} = \frac{50}{0.4} = 125 \quad E(Y_t) = 125$ /6

e) $\gamma(0) = \phi_1 \gamma(1) + \phi_2 \gamma(2) + \sigma^2$

$= 0.75 \times 1.5 - 0.15 \times 1.75 + 1 = 1.8625 = \gamma_0$ /6

f) $Y_{99} = 55$

$Y_{100} = 60$

$Y_{101} = 50 + 0.75 Y_{100} - 0.15 Y_{99} + 0$

$= 50 + 0.75 \times 60 - 0.15 \times 55 = 86.75$ /6

Σ 34

SU, DEPARTMENT OF STATISTICS

Room: Brunnsviksalen Anonymous code: 0030-EDO Sheet number: 3

Question 3

a) The second-order moving average process, MA(2) /4

b) A finite order moving average (MA) process is always stationary /4

Question 4

a) D /3

b) A /3

c) C /3

d) A /3

e) B /3

f) C /3

g) D /3

h) B /3

i) D /3

j) C /3

k) A /3

l) D /3

Σ 36