UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Year 2003/2004

February/March 2004

MSG 367 – ANALISIS SIRI MASA
(Time Series Analysis)

Time : 3 hours
Masa : 3 Jam

Sila pastikan bahawa kertas peperiksaan ini mengandungi **EMPAT [4]** soalan di dalam **TIGA BELAS [13]** halaman yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab **SEMUA** soalan.

*Please ensure that this examination paper consists of **FOUR [4]** questions on **THIRTEEN [13]** printed pages before you begin the examination.*

*Answer **ALL** questions.*

1. (a) Beri definisi bagi fungsi autokorelasi (fak) dan fungsi autokorelasi separa (faks).
[15 markah]

(b) Terangkan langkah-langkah utama kaedah Box-Jenkins bagi membina model untuk satu set siri masa yang tidak mempamirkan variasi bermusim dan juga tren dalam min. Terutamanya, cadangkan bagaimana fak dan faks bagi sample boleh digunakan untuk menentukan peringkat sesuatu model.
[35 markah]

(c) Beri definasi benar-benar pegun dan pegun peringkat kedua. Tunjukkan bahawa sebarang proses Purata Bergerak adalah pegun peringkat kedua dan tuliskan syarat, yang melibatkan punca-punca suatu persamaan, untuk satu proses Autoregresi pegun peringkat kedua.
[25 markah]

(d) Tulis semula model-model berikut menggunakan pengoperasi anjak kebelakang B dan nyatakan bentuk ARKPB(p,d,q)

(i) $Y_t = \mu(1-\phi_1-\phi_2)+\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\varepsilon_t+\theta_1\varepsilon_{t-1}$
(ii) $Y_t = Y_{t-1}+\phi_1(Y_{t-1}-Y_{t-2})+\phi_2(Y_{t-2}-Y_{t-3})+\varepsilon_t$
(iii) $Y_t = (1+\phi_1)Y_{t-1}+(\phi_2-\phi_1)Y_{t-2}-\phi_2 Y_{t-3}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}$
(iv) $Y_t = -\phi_1 Y_{t-1}+\varepsilon_t-(\theta_2-\theta_1)\varepsilon_{t-1}-\theta_1\theta_2\varepsilon_{t-2}$
(v) $Y_t = 2Y_{t-1}-Y_{t-2}+\varepsilon_t+\theta_1\varepsilon_{t-1}$ [25 markah]

1. (a) Give the definition of the autocorrelation function (acf) and partial autocorrelation function (pacf)
[15 marks]

(b) Explain the major steps in the Box-Jenkins approach to fitting an ARIMA model to a set of time series data which exhibits neither seasonal variation nor a trend

in the mean. Indicate particularly how the sample acf and sample pacf may be used in determining the order of the model.
[35 marks]

(c) Define strict stationarity and second order stationarity. Show that any Moving Average process is second order stationarity and write down a condition, involving the roots of an equation, for an Autoregressive process to be second order stationarity.
[25 marks]

(d) Rewrite each of the models below using the backward operator B and state the form of ARIMA(p,d,q)

(i) $Y_t = \mu(1-\phi_1-\phi_2)+\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\varepsilon_t+\theta_1\varepsilon_{t-1}$
(ii) $Y_t = Y_{t-1}+\phi_1(Y_{t-1}-Y_{t-2})+\phi_2(Y_{t-2}-Y_{t-3})+\varepsilon_t$
(iii) $Y_t = (1+\phi_1)Y_{t-1}+(\phi_2-\phi_1)Y_{t-2}-\phi_2 Y_{t-3}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}$
(iv) $Y_t = -\phi_1 Y_{t-1}+\varepsilon_t-(\theta_2-\theta_1)\varepsilon_{t-1}-\theta_1\theta_2\varepsilon_{t-2}$
(v) $Y_t = 2Y_{t-1}-Y_{t-2}+\varepsilon_t+\theta_1\varepsilon_{t-1}$ *[25 marks]*

2. (a) Suatu jujukan cerapan $\{Y_1, Y_2, \ldots, Y_N\}$ mempunyai varians sample $\hat{\gamma}_0 = 14.5$ dan nilai autokovarians sample susulan-1 $\hat{\gamma}_1 = 5$. Tunjukkan bahawa terdapat lebih daripada satu proses purata bergerak yang boleh disuaikan terhadap cerapan tersebut, tetapi buktikan bahawa hanya satu daripada proses tersebut tersongsangkan.

[20 markah]

(b) Dapatkan suatu rumus am fungsi autokovarians, fungsi autokorelasi dan fungsi autokorelasi separa bagi proses ARPB(2,2) seperti diberi di bawah

$$\left(1-\phi_1 B-\phi_2 B^2\right)Y_t = \left(1-\theta_1 B-\theta_2 B^2\right)\varepsilon_t$$

Pertimbangkan kes khas berikut apabila $\phi_1 = 0.4$, $\phi_2 = 0$, $\theta_1 = -0.4, \theta_2 = 0.45$ dan N = 484. Hitung autokorelasi untuk susulan k = 1, 2, 3, 4 dan autokorelasi separa untuk susulan $k = 1$ dan 2. Komen corak yang diperoleh. Adakah fungsi autokorelasi dan fungsi autokorelasi separa mencadangkan satu model ARPB(p,q)? [Diberi nilai fungsi autokorelasi separa susulan-3, 4 dan 5 masing-masing 0.26, -0.22 dan 0.10].

[40 markah]

(c) Berdasarkan pada kes khas model ARPB(1,2) di bahagian (b):

$$(1-0.4B)Y_t = \left(1+0.4B-0.45B^2\right)\varepsilon_t$$

seorang pelajar yang kurang berpengalaman ingin menyuaikan satu model yang lebih hemat kikir (parsimoni), iaitu pelajar tersebut mempunyai pilihan di antara lima model berikut:

(i) AR(1) $\quad (1-0.4B)Y_t = \varepsilon_t$

(ii) AR(2) $\quad \left(1-0.8B+0.77B^2\right)Y_t = \varepsilon_t$

(iii) MA(1) $\quad Y_t = (1+0.9B)\varepsilon_t$

(iv) MA(2) $\quad Y_t = \left(1+0.4B-0.45B^2\right)\varepsilon_t$

(v) ARMA(1,1) $\quad (1-0.4B)Y_t = (1+0.9B)\varepsilon_t$

Oleh sebab anda seorang yang mahir dalam analisis siri masa, pelajar tersebut ingin mendapatkan nasihat daripada kamu. Terangkan kepada pelajar tersebut alasan kesesuaian ataupun ketidaksesuaian setiap model di atas.

[40 markah]

2. (a) A sequence of observation $\{Y_1, Y_2, \ldots, Y_N\}$ has sample variance $\hat{\gamma}_0 = 14.5$ and sample lag-1 autocovariance $\hat{\gamma}_1 = 5$. Show that there is more than one first order moving average process which can be fitted to these data, but verify that only one of the fitted process is invertible.

[20 marks]

(b) *Find a general formula for autocovariance, autocorrelation and partial autocorrelation functions for an ARMA(2,2) process as given below*

$$\left(1-\phi_1 B-\phi_2 B^2\right)Y_t=\left(1-\theta_1 B-\theta_2 B^2\right)\varepsilon_t$$

Consider the following special case when $\phi_1 = 0.4$, $\phi_2 = 0$, $\theta_1 = -0.4, \theta_2 = 0.45$ *and* $N = 484$. *Calculate the autocorrelation for k = 1, 2, 3 and 4 and partial autocorrelation for k = 1 and 2. Comment on the pattern observed. Does the acf and pacf suggest an ARMA(p,q) model?. [Given the values of pacf at lag 3, 4 and 5 are 0.26 and –0.22 and 0.10 respectively].*

[40 marks]

(c) *Based on the special case of ARMA(1,2) model in (b)*

$$(1-0.4B)Y_t=\left(1+0.4B-0.45B^2\right)\varepsilon_t$$

an inexperienced student intends to fit a more parsimonious model to the data, that is the student has a choice of the following five models

(ii)	*AR(1)*	$(1-0.4B)Y_t=\varepsilon_t$
(ii)	*AR(2)*	$\left(1-0.8B+0.77B^2\right)Y_t=\varepsilon_t$
(iii)	*MA(1)*	$Y_t=(1+0.9B)\varepsilon_t$
(iv)	*MA(2)*	$Y_t=\left(1+0.4B-0.45B^2\right)\varepsilon_t$
(v)	*ARMA(1,1)*	$(1-0.4B)Y_t=(1+0.9B)\varepsilon_t$

As you are an expert in time series analysis, the student is seeking advice from you. Explain to the student the reason for suitability or insuitability of each of the model above.

[40 marks]

3. (a) Terbitkan persamaan Yule-Walker bagi fungsi autokovarians untuk proses autoregresi peringkat kedua

$$Y_t=\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\varepsilon_t$$

dengan $\varepsilon_t \sim \mathrm{N}\left(0,\sigma_\varepsilon^2\right)$.

Tunjukkan bahawa persamaan Yule-Walker boleh ditulis dalam bentuk berikut

$$\begin{bmatrix}\gamma_1 & \gamma_2 & 1\\ \gamma_0 & \gamma_1 & 0\\ \gamma_1 & \gamma_0 & 0\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\\ \sigma_\varepsilon^2\end{bmatrix}=\begin{bmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\end{bmatrix}$$

Suatu siri masa $\{Y_1, Y_2, \ldots Y_{1000}\}$ telah dicerap and memberikan fungsi autokovarians seperti berikut: $\hat{\gamma}_0 = 2.5, \hat{\gamma}_1 = 2.0, \hat{\gamma}_2 = 1.0$. Seorang pelajar mengambil keputusan untuk menyuaikan cerapan tersebut dengan model autoregresi peringkat kedua (AR(2)). Guna persamaan di atas untuk mendapatkan anggaran kaedah momen bagi koefisien-koefisien model autoregresi dan varians hingar putih.

[30 markah]

(b) Menggunakan pakej statistik model AR(2) telah disuaikan terhadap siri masa di atas dan menghasilkan anggaran pembolehubah-pembolehubah seperti di bawah. Bincang serba ringkas perbezaan di antara anggaran yang diperoleh dalam (a) dengan yang dihasilkan oleh pakej. Bincangkan kesesuaian model yang dipilih.

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-1.4854	0.0266	-55.75	0.000
AR 2	-0.5400	0.0266	-20.27	0.000

Number of observations: 1000
Residuals: SS = 4188.77
MS = 4.20 DF = 998

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	41.9	59.3	75.9	85.3
DF	10	22	34	46
P-Value	0.000	0.000	0.000	0.000

[20 markah]

(c) Analisis residual bagi ralat telah dilakukan dan acf serta pacf ditunjukkan dalam Lampiran 1. Pelajar tersebut telah melakukan beberapa analisis terhadap ralat dan telah mencuba untuk menyuaikan model baru terhadap siri masa tersebut. Semua langkah yang telah diambil oleh pelajar tersebut ditunjukkan dalam Lampiran 2. Cadangkan sebab-sebab bagi setiap langkah yang telah diambil oleh pelajar tersebut.

[50 markah]

3. *(a) Derive the Yule-Walker equation for the autocovariance function of the second order autoregression:*

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

where $\varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2)$.

Show that the Yule-Walker equations may be written in the form

$$\begin{bmatrix} \gamma_1 & \gamma_2 & 1 \\ \gamma_0 & \gamma_1 & 0 \\ \gamma_1 & \gamma_0 & 0 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \sigma_\varepsilon^2 \end{bmatrix} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix}$$

A time series $\{Y_1, Y_2, \ldots Y_{1000}\}$ *is observed, yielding the sample autocovariance function as follows:* $\hat{\gamma}_0 = 2.5, \hat{\gamma}_1 = 2.0, \hat{\gamma}_2 = 1.0$. *A student has decided to fit a second order autoregression to the series. Use the equation above to find method of moments estimators for the parameters of the autoregression and the variance of white noise.*

[30 marks]

(b) Using a statistical package an AR(2) has been fitted to the series above and produce estimate of parameters as below. Briefly discuss the difference between the estimates obtained in (a) and that produced by the package. Discuss the suitability of the model.

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-1.4854	0.0266	-55.75	0.000
AR 2	-0.5400	0.0266	-20.27	0.000

Number of observations: 1000
Residuals: SS = 4188.77
MS = .4.20 DF = 998

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	41.9	59.3	75.9	85.3
DF	10	22	34	46
P-Value	0.000	0.000	0.000	0.000

[20 marks]

(c) A residual analysis was carried out and the acf and pacf of the residuals are shown in Appendix 1. The students has carried out a few analysis on the residuals and has also tried to refit a new model to the series. All steps taken by the students are as shown in Appendix 2. Suggest the reasons on each of the step taken by the student.

[50 marks]

4. (a) Diberi model ARPB(2,1)

$$\left(1-\phi_1 B-\phi_2 B^2\right)\left(Y_t-\mu\right)=\left(1-\theta_1 B\right)\varepsilon_t$$

$\{\varepsilon_t\}$ adalah proses hingar putih yang mempunyai min 0 dan varians σ_ε^2.

Dapatkan suatu rumus am untuk ramalan satu-langkah-kehadapan $\hat{Y}_t(1)$ dan tunjukkan bahawa ramalan m-langkah-kehadapan diberi oleh rumus:

$$\hat{Y}_t(m)=\mu\left(1-\phi_1-\phi_2\right)+\phi_1\hat{Y}_t(m-1)+\phi_2\hat{Y}_t(m-2) \quad for\, m\geq 2$$

[20 marks]

(b) Jika nilai-nilai anggaran bagi koefisien-koefisien yang berdasarkan siri masa sebanyak 400 cerapan adalah $\hat{\phi}_1=1.5$, $\hat{\phi}_2=-0.6$, $\theta_1=0.4$, $\hat{\mu}=100$, $s_\varepsilon^2=4$ dengan $Y_{399}=115$, $Y_{400}=122$ dan $\varepsilon_{400}=8$, dapatkan nilai $\hat{Y}_{400}(m)$ bagi m = 1, 2, ..., 6. Bina selang keyakinan 95% bagi Y_{101}, ..., Y_{104}. Komen 6 nilai ramalan yang diperoleh. Apakah nilai berkemungkinan bagi $\hat{Y}_{400}(m\rightarrow\infty)$. Beri penjelasan.

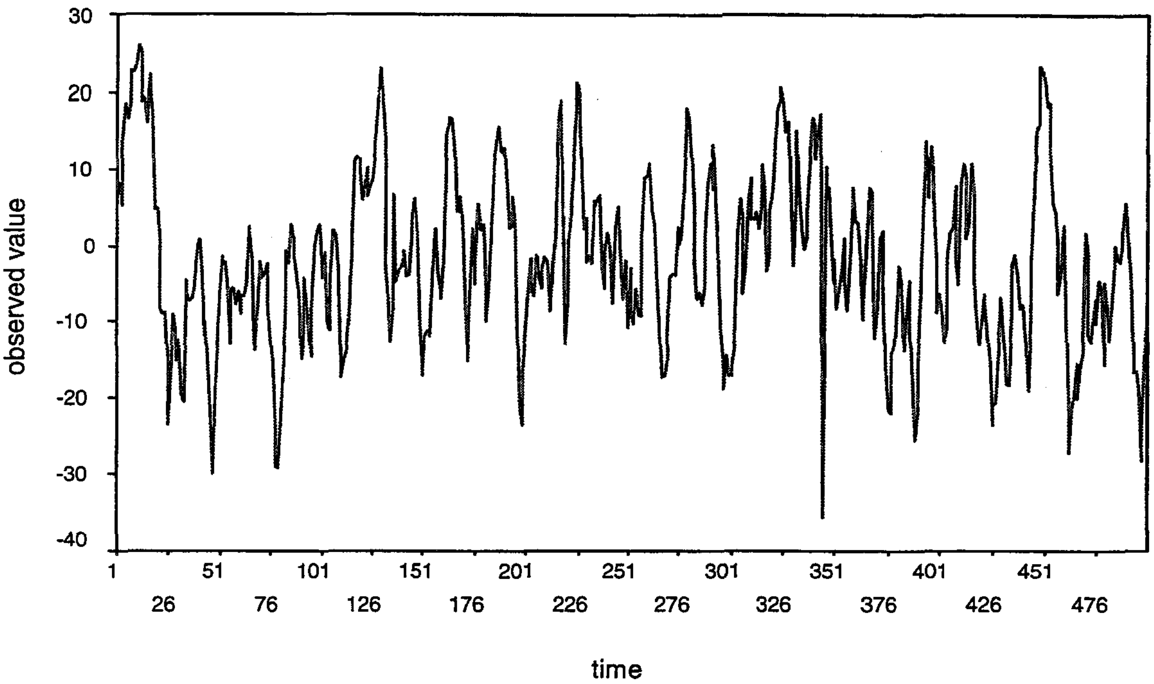
[40 markah]

(c) Diketahui pula $Y_{401}=91$, hitung nilai ramalan kemaskini bagi $Y_{102},\ldots Y_{106}$. Bandingkan nilai ramalan terkini dengan nilai di bahagian (b) dan beri penjelasan.

(d) (i) Bagi model bermusim di bawah, nyatakan bentuk SARIMA(p,d,q)(P,D,Q) Cari dan plot fungsi autokorelasi bagi model tersebut untuk k = 1, 2, ..., 11, 12, 13, ..., 23, 24, 25, 26. Komen secara ringkas terhadap corak yang diperoleh

$$Y_t=\left(1+0.8B\right)\left(1-0.5B^{12}\right)\varepsilon_t$$

(ii) Suatu siri masa sepanjang 500 cerapan seperti ditunjukkan pada graf di bawah telah disuaikan dengan model ARPB(p,q), dengan dan tanpa data terpencil 'additional' pada t = 345. Tuliskan persamaan lengkap bagi model intervensi ARPB(p,q) dan bincangkan kesesuaian model ini berbanding model tulen ARPB(p,q).
[Rujuk komputer output dalam Lampiran 3]



[20 markah]

4. (a) *Given an ARMA(2,1) model*

$$\left(1-\phi_1 B-\phi_2 B^2\right)\left(Y_t-\mu\right)=\left(1-\theta_1 B\right)\varepsilon_t$$

$\{\varepsilon_t\}$ is a white noise process with mean 0 and variance σ_ε^2
Obtain a general formula for the one-step-ahead forecast $\hat{Y}_t(1)$ and show that the m-step-ahead forecast is given by

$$\hat{Y}_t(m)=\mu\left(1-\phi_1-\phi_2\right)+\phi_1\hat{Y}_t(m-1)+\phi_2\hat{Y}_t(m-2) \quad for\ m\geq 2$$

[20 marks]

(b) *If estimated values for the coefficients based on time series of 400 observations are $\hat{\phi}_1=1.5$, $\hat{\phi}_2=-0.6$, $\theta_1=0.4$, $\hat{\mu}=100$, $s_\varepsilon^2=4$ with $Y_{399}=115$, $Y_{400}=122$ and $\varepsilon_{400}=8$, obtain value of $\hat{Y}_{400}(m)$ for m = 1, 2, ..., 6. Construct a 95% confidence interval for Y_{101}, ..., Y_{104}. Comment on the 6 forecast values obtained above. What is the likely value for $\hat{Y}_{400}(m\rightarrow\infty)$? Give explanation.*

[40 marks]

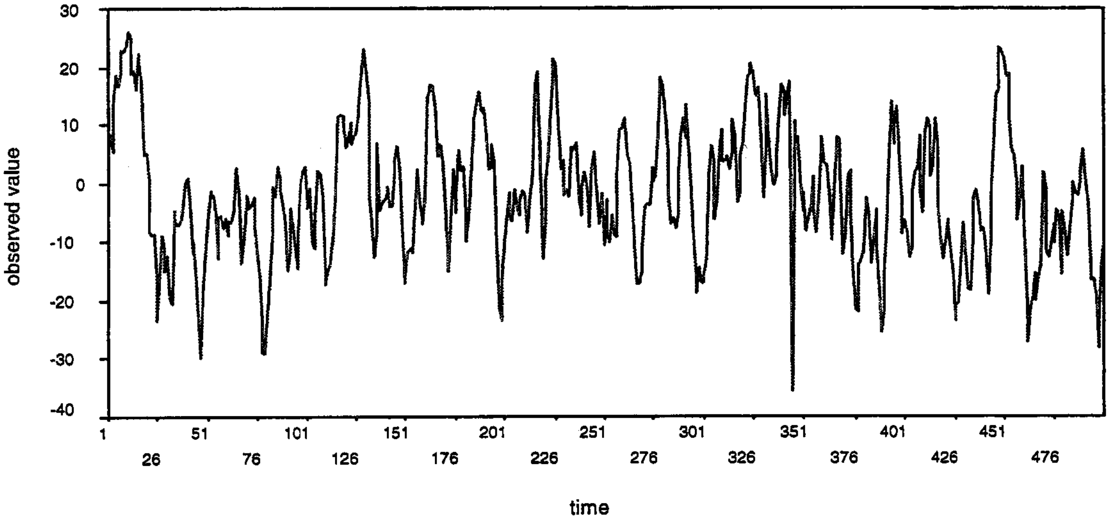
(c) *It is now observed $Y_{401}=91$, calculate the updated values for $Y_{102},\ldots Y_{106}$. Compare these new forecasts with those calculated in (ii) and discuss.*

[20 marks]

(d) (i) *For the seasonal model below, state the SARIMA(p,d,q)(P,D,Q) form. Find and plot the acf for the model for k = 1, 2, ..., 11, 12, 13, ..., 23, 24, 25, 26. Comment briefly on the pattern observed*

$$Y_t = (1+0.8B)(1-0.5B^{12})\varepsilon_t$$

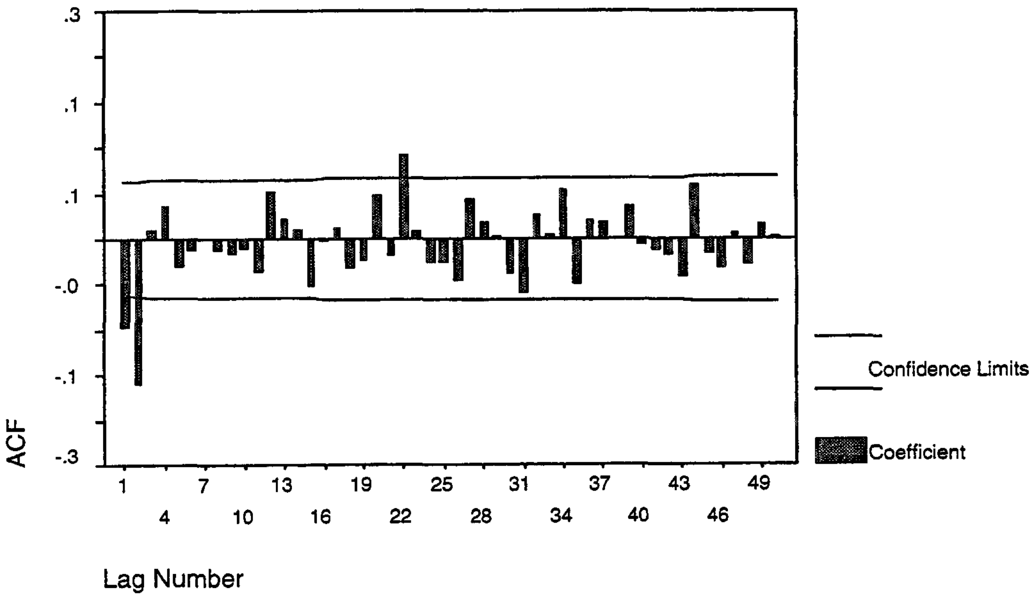
(ii) *A time series of length 500 as shown in the graph below has been fitted to an ARMA(p,q) process, with an without an additional outlier at t = 345. Write down the complete intervention ARMA(p,q) model and discuss the suitability of the model over the pure ARMA(p,q) model. [Refer computer output in Appendix 3].*



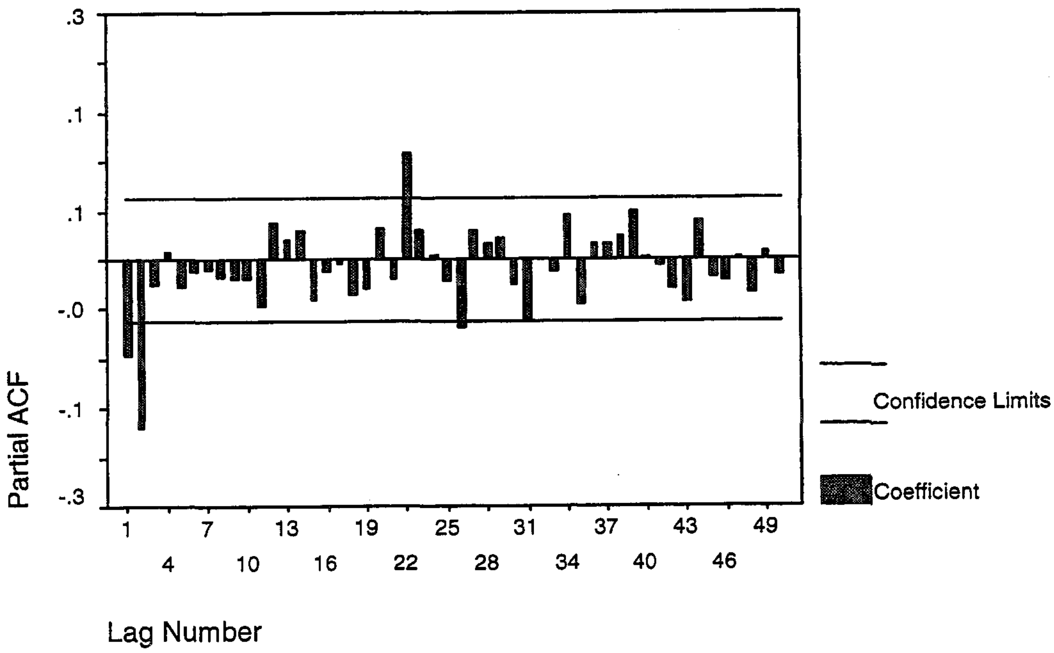
[20 marks]

Appendix 1/Lampiran 1

ACF of residuals from AR(2) model



PACF of residuals from AR(2) model



Appendix 2/Lampiran 2

First step: residual analysis

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-0.0305	0.2008	-0.15	0.879
MA 1	0.0837	0.1982	0.42	0.673
MA 2	0.1615	0.0403	4.01	0.000

Number of observations: 1000
Residuals: SS = 4029.89 (backforecasts excluded)
MS = 4.04 DF = 997

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6.9	24.9	41.2	50.4
DF	9	21	33	45
P-Value	0.649	0.250	0.154	0.268

Second step: residual analysis

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0.1133	0.0313	3.62	0.000
MA 2	0.1578	0.0313	5.05	0.000

Number of observations: 1000
Residuals: SS = 4030.03 (backforecasts excluded)
MS = 4.04 DF = 998

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6.9	24.9	41.0	50.3
DF	10	22	34	46
P-Value	0.735	0.304	0.189	0.308

Appendix 2(continued)/Lampiran 2(sambungan)

Third step: remodeling series

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-1.3892	0.1550	-8.96	0.000
AR 2	-0.4422	0.1450	-3.05	0.002
MA 1	0.2086	0.1609	1.30	0.195
MA 2	0.1055	0.1035	1.02	0.308

Number of observations: 1000
Residuals: SS = 4022.49 (backforecasts excluded)
MS = 4.04 DF = 996

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	5.7	22.7	39.6	49.0
DF	8	20	32	44
P-Value	0.679	0.303	0.167	0.280

Fourth step: remodeling series

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-1.2298	0.0509	-24.16	0.000
AR 2	-0.2932	0.0502	-5.85	0.000
MA 1	0.3720	0.0495	7.51	0.000

Number of observations: 1000
Residuals: SS = 4025.57 (backforecasts excluded)
MS = 4.04 DF = 997

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6.6	23.8	41.5	50.6
DF	9	21	33	45
P-Value	0.680	0.304	0.146	0.263

Appendix 3/Lampiran 3

Initial values:
AR1 .7772 , MA1 -.6806

FINAL PARAMETERS:

Number of residuals	500
Standard error	6.1736
Log likelihood	-1619.2418
AIC	3242.4835
SBC	3250.9127

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	498	19028.574	38.1144

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	.7818	.0329	23.6982	.0000
MA1	-.1862	.0520	-3.5758	.0003

Covariance Matrix:

	AR1	MA1
AR1	.00108	.00092
MA1	.00092	.00271

Correlation Matrix:

	AR1	MA1
AR1	1.0000	.5356
MA1	.5356	1.0000

Initial values:
AR1 .7767, MA1 -.8606, intervention -49.0514

FINAL PARAMETERS:

Number of residuals	500
Standard error	5.1417
Log likelihood	-1527.53
AIC	3061.07
SBC	3073.71

Analysis of Variance:

	DF	Adj. Sum of Squares	Residual Variance
Residuals	497	13185.51	26.4371

Variables in the Model:

	B	SEB	T-RATIO	APPROX. PROB.
AR1	.7805	.0307	25.411	.0000
MA1	-.4449	.0441	-10.083	.0000
intervention	-50.915	3.033	-16.783	.0000

Covariance Matrix:

	AR1	MA1
AR1	.00094	.00056
MA1	.00056	.00194

Correlation Matrix:

	AR1	MA1
AR1	1.000	.4161
MA1	.4161	1.0000

	intervention
intervention	9.2029

-ooo000ooo-