UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2005/2006

November 2005

MGM 564 – Applied Econometrics
[Ekonometrik Gunaan]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of **THIRTEEN** pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **TIGA BELAS** muka surat yang bercetak sebelum anda memulakan peperiksaan ini].*

Instructions: Answer **all four** [4] questions.

*[Arahan : Jawab **semua empat** soalan].*

...2/-

1. (a) It is known that for a given simple regression $Y = b_0 + b_1 X + \varepsilon$ the least squares estimate for b_1 is given by

$$\hat{b}_1 = \frac{N\Sigma XY - \Sigma X \Sigma Y}{N\Sigma X^2 - (\Sigma X)^2} = \frac{\Sigma xy}{\Sigma x^2}$$

The coefficient of determination, R^2 for the simple regression above is given by

$$R^2 = \frac{ESS}{TSS} = \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2} = \frac{\Sigma \hat{y}_i^2}{\Sigma y_i^2}$$

where *ESS* is the explained sum of squares, *TSS* is the total sum of squares and N is the number of observations. Show that the coefficient of determination can also be written as

$$R^2 = \hat{b}_1 \frac{\Sigma xy}{\Sigma y^2} \quad .$$

[15 marks]

(b) The following table gives data on consumption (in thousand, RM) and yearly income (in thousand, RM) for 10 randomly selected secretaries at different government departments at Putrajaya.

Individual	1	2	3	4	5	6	7	8	9	10
Consumption, C	10	5	3	4	2	10	15	20	10	20
Income, I	20	10	5	6	3	15	20	30	15	35

Given that:

$$\Sigma I^2 = 3545, \quad \Sigma IC = 2195, \quad \Sigma ic = 620.90, \quad \Sigma i^2 = 1016.90$$

Assuming a simple linear relationship between I and C

$$C = \beta_1 + \beta_2 I + \varepsilon$$

(i) What is the theoretical expectation (priori) for the sign of β_1 and β_2 respectively? Discuss.

[5 marks]

(ii) Obtain the ordinary least squares estimators for the regression coefficients β_1 and β_2 and discuss briefly the interpretation of the coefficients. Do these coefficients satisfies *prior* theoretical expectation in (i) above?

[20 marks]

...3/-

(iii) Using t-test at 5% significant level, test the hypothesis $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$. Construct an ANOVA table and calculate the F-statistics. Use your result to acertain your answer for the hypothesis above.

[30 marks]

(iv) Calculate the coefficient of determination and adjusted coefficient of determination for the model above and interpret the values.

[15 marks]

(v) It is known that a newly employed secretary at a certain department has a (starting) yearly income of RM25500. Calculate the estimated consumption for the secretary and construct a 95% confidence interval for the estimate.

[15 marks]

1. (a) *Diketahui bahawa bagi suatu regresi mudah yang diwakili oleh* $Y = b_0 + b_1 X + \varepsilon$ *penganggar kuasa dua terkecil bagi* b_1 *diberikan oleh rumus di bawah*

$$\hat{b}_1 = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} = \frac{\sum xy}{\sum x^2}$$

Pekali penentuan, R^2 *untuk regresi mudah di atas diberikan oleh rumus*

$$R^2 = \frac{ESS}{TSS} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2}$$

yang ESS adalah hasil tambah kuasa dua regresi, TSS adalah jumlah hasil tambah kuasa dua dan N adalah bilangan cerapan. Tunjukkan bahawa pekali penentuan juga boleh dirumuskan sebagai

$$R^2 = \hat{b}_1 \frac{\sum xy}{\sum y^2}.$$

[15 markah]

(b) *Jadual di bawah menunjukkan perbelanjaan (dalam ribu, RM) dan pendapatan tahunan (dalam ribu, RM) bagi 10 kerani yang dipilih secara rawak di jabatan kerajaan yang berbeza di Putrajaya.*

Individu	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Perbelanjaan, C	*10*	*5*	*3*	*4*	*2*	*10*	*15*	*20*	*10*	*20*
Pendapatan, I	*20*	*10*	*5*	*6*	*3*	*15*	*20*	*30*	*15*	*35*

...4/-

Diberikan

$$\sum I^2 = 3545, \quad \sum IC = 2195, \quad \sum ic = 620.90, \quad \sum i^2 = 1016.90$$

Dengan mengandaikan hubungan linear mudah di antara I dan C, iaitu

$$C = \beta_1 + \beta_2 I + \varepsilon$$

(i) *Apakah jangkaan berteori bagi tanda untuk β_1 dan β_2 masing-masing? Bincangkan.*

[5 markah]

(ii) *Dapatkan anggaran kuasa dua terkecil bagi koefisien-koefisien regressi β_1 dan β_2 dan tafsirkannya. Adakah koefisien-koefisien regresi ini memenuhi jangkaan priori?*

[20 markah]

(iii) *Menggunakan ujian-t pada aras keertian 5%, ujikan hipotesis $H_0 : \beta_2 = 0$ terhadap $H_1 : \beta_2 \neq 0$. Bina jadual ANOVA dan hitung statistik-F. Gunakan keputusan anda untuk memastikan jawaban anda bagi hipotesis di atas.*

[30 markah]

(iv) *Hitung pekali penentuan dan pekali penentuan terlaras bagi model di atas dan tafsirkan.*

[15 markah]

(v) *Adalah diketahui bahawa seorang kerani baru di suatu jabatan mempunyai pendapatan tahunan (permulaan) sebanyak RM25500. Hitungkan nilai anggaran perbelanjaan kerani tersebut serta bina selang keyakinan 95% bagi anggaran tersebut.*

[15 markah]

2. For a given multiple regression model

$$Y_i = \beta_1 + \beta_2 X_{1i} + \beta_3 X_{2i} + \ldots + \beta_{k+1} X_{ki} + \varepsilon_i$$

answer the following questions:

(a) Explain briefly the nature or source of the following problems

(i) multicollinearity
(ii) autocorrelation

[10 marks]

...5/-

(b) Discuss the consequences of multicollinearity on parameter estimation.
[10 marks]

(c) Explain how the 2 methods below are used to detect the problem of multicollinearity and autocorrelation respectively
(i) Variance Inflation Factor (VIF)
(ii) Durbin-Watson *d* Test
[25 marks]

(d) Table below shows the import and Gross National Product of a particular country between 1981 and 2000.
(i) Calculate Durbin-Watson *d* statistics and determine if autocorrelation is presence in the data. Show that the approximate first order autocorrelation is 0.53.
(ii) Construct a new table for the corrected values of import, I^* and Gross National Product, G^*. Based on the corrected values the new regression model is given by

$$I^* = -1394 + 0.296\,G^*$$

Investigate if autocorrelation remain present in the new model. Justify your answer.
[40 marks]

Year	Import (I)	GNP (G)	Estimated Import, $\hat{I}$	Residual, $(I-\hat{I})$
1981	3748	21777	3626	122
1982	4010	22418	3805	205
1983	3711	22308	3774	-63
1984	4004	23319	4057	-53
1985	4151	24180	4298	-147
1986	4569	24893	4497	72
1987	4582	25310	4613	-31
1988	4697	25799	4750	-53
1989	4753	25886	4774	-21
1990	5062	26868	5049	13
1991	5669	28134	5403	266
1992	5628	29091	5670	-42
1993	5736	29450	5771	-35
1994	5946	30705	6121	-175
1995	6501	32372	6587	-86
1996	6549	33152	6805	-256
1997	6705	33764	6976	-271
1998	7104	34411	7157	-53
1999	7609	35429	7442	167
2000	8100	36200	7657	443

...6/-

(e) In order to study the dynamic among a few economic variables, an Adjusted Dickey Fuller (ADF) test has been carried to investigate stationarity of the series. The table below shows the results of the ADF at level and after first differencing. The symbol * indicates the test is significant at 5% level. Based on the results discuss stationarity for each of the series.

Series	At Level		First Differenced	
	No trend	With Trend	No Trend	With Trend
LogM1	-1.127	-1.956	-9.813*	-9.838*
LogM2	-1.750	-0.095	-7.641*	-7.919*
LogM3	-3.063*	0.514	-6.316*	-7.254*
LogIPI	-1.763	-1.851	-15.575*	-15.698*
LogCPI	-1.666	-0.819	-7.124*	-7.335*
LogR3M	-1.446	-2.269	-7.770*	-7.831*

[15 marks]

2. *Dalam konteks sebuah regressi berganda*

$$Y_i = \beta_1 + \beta_2 X_{1i} + \beta_3 X_{2i} + \ldots + \beta_{k+1} X_{ki} + \varepsilon_i$$

jawab soalan-soalan berikut:

(a) Huraikan secara ringkas beberapa punca masalah-masalah berikut

(i) multikolinearan
(ii) autokorelasi

[10 markah]

(b) Bincangkan kesan multikolinearan terhadap penanggaran

[10 markah]

(c) Terangkan bagaimana 2 kaedah di bawah, masing-masing dapat mengatasi masalah multikolinearan dan autokorelasi

(i) Variance Inflation Factor (VIF)
(ii) Ujian Durbin-Watson d

[25 markah]

(d) Jadual di bawah menunjukkan impot dan Keluaran Negara Kasar (KNK) bagi suatu negara tertentu di antara tahun 1981 dan 2000.

(i) Hitung statistik Durbin-Watson d dan tentukan sekiranya autokorelasi wujud dalam data. Tunjukkan bahawa nilai anggaran autokorelasi peringkat pertama adalah 0.53.

...7/-

(ii) *Bina jadual baru untuk nilai-nilai selepas pembetulan bagi impot, I* dan Keluaran Negara Kasar, G*. Berdasarkan nilai-nilai selepas pembetulan model regresi terbaru diberikan sebagai*

$$I^* = -1394 + 0.296\,G^*$$

Jalankan kajian sekiranya autokorelasi masih wujud dalam model terbaru ini. Sahkan jawaban anda.

[40 markah]

Tahun	*Impot (I)*	*KNK (G)*	*Impot teranggar,* $\hat{I}$	*Ralat,* $(I - \hat{I})$
1981	*3748*	*21777*	*3626*	*122*
1982	*4010*	*22418*	*3805*	*205*
1983	*3711*	*22308*	*3774*	*-63*
1984	*4004*	*23319*	*4057*	*-53*
1985	*4151*	*24180*	*4298*	*-147*
1986	*4569*	*24893*	*4497*	*72*
1987	*4582*	*25310*	*4613*	*-31*
1988	*4697*	*25799*	*4750*	*-53*
1989	*4753*	*25886*	*4774*	*-21*
1990	*5062*	*26868*	*5049*	*13*
1991	*5669*	*28134*	*5403*	*266*
1992	*5628*	*29091*	*5670*	*-42*
1993	*5736*	*29450*	*5771*	*-35*
1994	*5946*	*30705*	*6121*	*-175*
1995	*6501*	*32372*	*6587*	*-86*
1996	*6549*	*33152*	*6805*	*-256*
1997	*6705*	*33764*	*6976*	*-271*
1998	*7104*	*34411*	*7157*	*-53*
1999	*7609*	*35429*	*7442*	*167*
2000	*8100*	*36200*	*7657*	*443*

(e) *Bagi mengkaji dinamik di antara beberapa pembolehubah ekonomi, ujian Dikey-Fuller terlaras (ADF) telah dijalankan untuk memeriksa kepegunan siri. Jadual di bawah menunjukkan keputusan ujian ADF pada aras dan selepas pembezaan peringkat pertama. Simbol * menerangkan ujian adalah siknifikan pada aras 5%. Berdasarkan keputusan dalam jadual bincangkan kepegunan bagi setiap siri.*

...8/-

Siri	*Aras*		*Pembezaan pertama*	
	Tiada tren	*Bertren*	*Tiada Tren*	*Bertren*
LogM1	*-1.127*	*-1.956*	*-9.813**	*-9.838**
LogM2	*-1.750*	*-0.095*	*-7.641**	*-7.919**
LogM3	*-3.063**	*0.514*	*-6.316**	*-7.254**
LogIPI	*-1.763*	*-1.851*	*-15.575**	*-15.698**
LogCPI	*-1.666*	*-0.819*	*-7.124**	*-7.335**
LogR3M	*-1.446*	*-2.269*	*-7.770**	*-7.831**

[15 markah]

3. (a) Prove that when heteroscedasticity in the form $E(\varepsilon_i^2) = \sigma^2 X_{1i}^2$ is present in the data, the variance of $\hat{\beta}_2^h$ (heteroscedasticity) is given by

$$Var(\hat{\beta}_2^h) = \sigma^2 \frac{\sum x_i^2 X_i^2}{(\sum x_i^2)^2}$$

[10 marks]

(b) The table below shows daily selling of banana (Y, in RM million), total population (X_1, in thousand) and Gross National Product (X_2, in RM million) for 10 particular areas in Malaysia.

Area	1	2	3	4	5	6	7	8	9	10
Y	0.2	102.0	2.9	41.5	4.8	3.7	0.8	4.2	1.8	45.0
X_1	2.3	180.2	7.8	140.1	15.2	21.3	5.7	86.8	14.9	543.1
X_2	0.8	743.2	14.5	110.4	20.2	15.6	3.0	15.3	5.7	115.2
ε	-1.869	-1.995	-1.181	17.808	-0.419	-1.312	-1.740	-4.636	-1.615	-3.040

Given that:

$\sum Y_i = 206.9$ $\sum X_{1i} = 1017.5$ $\sum X_{2i} = 1043.9$

$\sum Y_i^2 = 14217.95$ $\sum X_{1i}^2 = 355597.26$ $\sum X_{2i}^2 = 578943.31$

$\sum X_{1i}X_{2i} = 214141.13$ $\sum X_{1i}Y_i = 49204.84$ $\sum X_{2i}Y_i = 85845.81$

(i) Give reasons to explain why it is mostly likely there are problems of heteroscedasticity in the data given above.

[5 marks]

...9/-

(ii) Using matrix approach, estimate the consumption function $Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \varepsilon$ and test the significance of each partial regression coefficient at 5% significant level. Estimated value of σ^2 is given as 52.0. The matrix $(\mathbf{X^T X})^{-1}$ is given in the Appendix as M1.

[30 marks]

(iii) Based on the result in (ii) verify that the residual for area 4 and 10 are 17.808 and –3.040 respectively as shown in the table. Using Spearman's rank correlation test, determine at 5% significant level whether heteroscedasticity is present in the data above

[20 marks]

(iv) It was found out that heteroscedasticity in the data takes the form of $E(\varepsilon_i^2) = \sigma^2 X_{1i}^2$. Transform the original data and re-estimate the consumption function. Compare the results with that obtained in (ii) above. The matrix $(\mathbf{X^T X})^{-1}$ is given in the Appendix as M2.

[35 marks]

3. *(a) Buktikan bahawa sekiranya heterokedastisiti berbentuk* $E(\varepsilon_i^2) = \sigma^2 X_{1i}^2$ *wujud dalam data, maka variance bagi* $\hat{\beta}_2^h$ *(heteroskedastisiti) diberikan oleh rumus*

$$Var(\hat{\beta}_2^h) = \sigma^2 \frac{\sum x_i^2 X_i^2}{(\sum x_i^2)^2}$$

[10 markah]

(b) Jadual di bawah menunjukkan penjualan pisang harian (Y, dalam juta RM), jumlah penduduk (X_1*, dalam ribu) dan Keluaran Negara Kasar (*X_2*, dalam juta RM) bagi 10 kawasan tertentu di Malaysia.*

Kawasan	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Y	*0.2*	*102.0*	*2.9*	*41.5*	*4.8*	*3.7*	*0.8*	*4.2*	*1.8*	*45.0*
X_1	*2.3*	*180.2*	*7.8*	*140.1*	*15.2*	*21.3*	*5.7*	*86.8*	*14.9*	*543.1*
X_2	*0.8*	*743.2*	*14.5*	*110.4*	*20.2*	*15.6*	*3.0*	*15.3*	*5.7*	*115.2*
ε	*-1.869*	*-1.995*	*-1.181*	*17.808*	*-0.419*	*-1.312*	*-1.740*	*-4.636*	*-1.615*	*-3.040*

Diberikan:

$\sum Y_i = 206.9$ $\quad \sum X_{1i} = 1017.5$ $\quad \sum X_{2i} = 1043.9$

$\sum Y_i^2 = 14217.95$ $\quad \sum X_{1i}^2 = 355597.26$ $\quad \sum X_{2i}^2 = 578943.31$

$\sum X_{1i} X_{2i} = 214141.13$ $\quad \sum X_{1i} Y_i = 49204.84$ $\quad \sum X_{2i} Y_i = 85845.81$

...10/-

(i) *Berikan sebab untuk menerangkan mengapa berkemungkinan besar terdapatnya masalah heteroskedastisiti dalam data di atas.*

[5 markah]

(ii) *Menggunakan kaedah matriks, anggarkan fungsi perbelanjaan* $Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \varepsilon$ *dan uji kesignifikanan setiap koefisien regresi separa pada aras keertian 5%. Nilai teranggar bagi* σ^2 *adalah 52.0. Matriks* $\left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}$ *diberikan dalam Lampiran sebagai M1.*

[30 markah]

(iii) *Berdasarkan keputusan di bahagian (ii), sahkan bahawa nilai ralat bagi kawasan 4 dan 10 masing masing adalah 17.808 dan –3.040 seperti dalam jadual di atas. Dengan menggunakan ujian pekali korelasi pangkat Spearman, tentukan pada aras keertian 5% sama ada terdapat heteroskedastisiti dalam data yang diberi.*

[20 markah]

(iv) *Telah didapati bahawa heteroskedastisiti wujud dalam data dan berbentuk* $E\left(\varepsilon_i^2\right) = \sigma^2 X_{1i}^2$. *Jelmakan data asal dan anggarkan semula fungsi perbelanjaan. Bandingkan keputusan baru dengan jawaban dalam (ii) di atas. Matriks* $\left(\mathbf{X}^{\mathbf{T}}\mathbf{X}\right)^{-1}$ *diberikan dalam Lampiran sebagai M2.*

[35 markah]

4. (a) Rewrite each of the models below using the backward operator B and state the form of ARMA(p,q)

(i) $X_t = \mu(1 + \phi_1 + \phi_2) - \phi_1 X_{t-1} - \phi_2 X_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$

(ii) $X_t = X_{t-1} + \phi_1(X_{t-1} - X_{t-2}) + \phi_2(X_{t-2} - X_{t-3}) + \varepsilon_t + \theta_1 \varepsilon_{t-1}$

(iii) $X_t = \phi_1 X_{t-1} + (\phi_2 - \phi_1) X_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_3 \varepsilon_{t-3}$

[15 marks]

(b) Given a Moving Average of order 2 (MA(2)) model:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Obtain the general form of autocorrelation function for the models above. If $\theta_1 = 1.5$, $\theta_2 = -0.6$, find the autocorrelation for lag k = 1, 2, 3, … . Comment on the values obtained.

[20 marks]

...11/-

(c) Given an ARMA(2,1) model:

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

ε_t is a white noise process with mean 0 and variance σ_ε^2.

(i) Obtain formulae for the one-step-ahead forecast and show that the m-step-ahead forecast is given by

$$\hat{Y}_t(m)=\mu\left(1-\phi_1-\phi_2\right)+\phi_1 Y_t+\phi_2 Y_{t-1}-\theta_1\varepsilon_t \quad \text{for } m=1$$

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

$$\hat{Y}_t(m)=\mu\left(1-\phi_1-\phi_2\right)+\phi_1\hat{Y}_t(m-1)+\phi_2\hat{Y}_t(m-2) \quad \text{for } m\geq 3$$

If estimated values for the coefficients based on time series of 300 observations are $\hat{\phi}_1=1.7$, $\hat{\phi}_2=-0.72$, $\hat{\theta}_1=-0.6$, $\hat{\mu}=200$, $s_\varepsilon^2=4$ with $Y_{299}=180$, $Y_{300}=185$ and $\hat{\varepsilon}_{300}=5$ obtain value of $\hat{Y}_{300}(m)$ for m = 1, 2, ..., 5. Construct a 95% confidence interval for Y_{301} and Y_{302}. Comment on the 5 forecast values obtained above. What is the likely value for $\hat{Y}_{300}(m\to\infty)$? Give an explanation.

[45 marks]

(ii) It is now observed $Y_{301}=210$, calculate the updated values for $Y_{302}\ldots Y_{305}$. Compare these new forecasts with those calculated in (i) and discuss.

[20 marks]

4. *(a) Tulis semula setiap model di bawah dengan menggunakan symbol pengoperasi kebelakang B dan kemudian nyatakan bentuk ARMA(p,q).*

(i) $X_t=\mu\left(1+\phi_1+\phi_2\right)-\phi_1 X_{t-1}-\phi_2 X_{t-2}+\varepsilon_t-\theta_1\varepsilon_{t-1}$

(ii) $X_t=X_{t-1}+\phi_1\left(X_{t-1}-X_{t-2}\right)+\phi_2\left(X_{t-2}-X_{t-3}\right)+\varepsilon_t+\theta_1\varepsilon_{t-1}$

(iii) $X_t=\phi_1 X_{t-1}+\left(\phi_2-\phi_1\right)X_{t-2}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\theta_3\varepsilon_{t-3}$

[15 markah]

(b) Diberikan model Purata Bergerak peringkat kedua (PB(2)):

$$X_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}$$

Dapatkan ungkapan umum fungsi autokorelasi bagi model di atas. Jika $\theta_1=1.5$, $\theta_2=-0.6$, *cari autokorelasi bagi susulan k = 1, 2, 3, ... Bincangkan nilai-nilai yang diperoleh.*

[20 markah]

...12/-

(c) *Diberikan model ARMA(2,1) seperti berikut:*

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

(i) *Dapatkan rumus telahan satu langkah kehadapan dan tunjukkan bahawa telahan m-langkah kehadapan diberikan oleh:*

$$\hat{Y}_t(m)=\mu(1-\phi_1-\phi_2)+\phi_1 Y_t+\phi_2 Y_{t-1}-\theta_1\varepsilon_t \quad \text{untuk } m=1$$

$$\hat{Y}_t(m)=\mu(1-\phi_1-\phi_2)+\phi_1\hat{Y}_t(m-1)+\phi_2 Y_t \quad \text{untuk } m=2$$

$$\hat{Y}_t(m)=\mu(1-\phi_1-\phi_2)+\phi_1\hat{Y}_t(m-1)+\phi_2\hat{Y}_t(m-2) \quad \text{untuk } m\geq 3$$

Sekiranya nilai anggaran bagi koefisien-koefisien berdasarkan satu siri masa terdiri daripada 300 cerapan adalah $\hat{\phi}_1=1.7$, $\hat{\phi}_2=-0.72$, $\hat{\theta}_1=-0.6$, $\hat{\mu}=200$, $s_\varepsilon^2=4$ *dengan* $Y_{299}=180$, $Y_{300}=185$ *dan* $\hat{\varepsilon}_{300}=5$, *dapatkan nilai* $\hat{Y}_{300}(m)$ *bagi m = 1, 2, ..., 5. Bina selang keyakinan 95% bagi* Y_{301} *dan* Y_{302}. *Berikan komen terhadap 5 nilai telahan yang diperoleh. Apakah nilai berkemungkinan bagi* $\hat{Y}_{300}(m\to\infty)$. *Beri penjelasan.*

[45 markah]

(ii) *Diketahui pula* $Y_{301}=210$, *hitung nilai ramaln kemaskini bagi* $Y_{302},\ldots Y_{305}$. *Bandingkan nilai ramalan terkini dengan nilai dihitung di bahagian (i) dan beri penjelasan.*

[20 markah]

...13/-

APPENDIX/LAMPIRAN

$$M1 = \begin{bmatrix} 0.1497930 & -0.0003421 & -0.0001435 \\ -0.0003421 & 0.0000044 & -0.0000010 \\ -0.0001435 & -0.0000010 & 0.0000024 \end{bmatrix}$$

$$M2 = \begin{bmatrix} 0.28167 & -0.80587 & -0.10076 \\ -0.80587 & 6.66587 & 0.16835 \\ -0.10076 & 0.16835 & 0.08099 \end{bmatrix}$$

For $Y_i = \beta_1 + \beta_2 X_{1i} + \beta_3 X_{2i} + \varepsilon_i$

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}_1 - \hat{\beta}_3 \bar{X}_2$$

$$\hat{\beta}_2 = \frac{(\sum y_i x_{1i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{1i} x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2} \qquad \hat{\beta}_3 = \frac{(\sum y_i x_{2i})(\sum x_{1i}^2) - (\sum y_i x_{1i})(\sum x_{1i} x_{2i})}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2}$$

$$Var(\hat{\beta}_1) = \left[\frac{1}{N} + \frac{\bar{X}_1^2 \sum x_{2i}^2 + \bar{X}_2^2 \sum x_{1i}^2 - 2\bar{X}_1 \bar{X}_2 \sum x_{1i} x_{2i}}{\sum x_{1i}^2 \sum x_{2i}^2 - (\sum x_{1i} x_{2i})^2} \right] \sigma^2$$

$$Var(\hat{\beta}_2) = \frac{\sum x_{2i}^2}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2} \sigma^2 \qquad Var(\hat{\beta}_3) = \frac{\sum x_{1i}^2}{(\sum x_{1i}^2)(\sum x_{2i}^2) - (\sum x_{1i} x_{2i})^2} \sigma^2$$

$$Cov(\hat{\beta}_2, \hat{\beta}_3) = \frac{-r_{12}\sigma^2}{(1 - r_{12}^2)\sqrt{\sum x_{1i}^2}\sqrt{\sum x_{2i}^2}}$$

where $\hat{\sigma}^2 = \frac{\sum \hat{\varepsilon}_i^2}{N-3}$, $\sum \hat{\varepsilon}_i^2 = \sum y_1^2 - \hat{\beta}_2 \sum y_i x_{1i} - \hat{\beta}_3 \sum y_i x_{2i}$

$$(\mathbf{X}^{\mathbf{T}}\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathbf{T}}\mathbf{Y} \quad \Rightarrow \quad Cov(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}$$

$$R^2 == \frac{ESS}{TSS} = 1 - \frac{\sum \hat{\varepsilon}_i^2}{\sum y_i^2} = \frac{\hat{\boldsymbol{\beta}}^{\mathbf{T}}\mathbf{X}^{\mathbf{T}}\mathbf{Y} - N\bar{Y}^2}{\mathbf{Y}^{\mathbf{T}}\mathbf{Y} - N\bar{Y}^2}$$

Spearman Rank Test: $r_s = 1 - 6\left[\frac{\sum d_i^2}{N(N^2-1)}\right] \quad \Rightarrow \quad t = \frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}}$

- ooo O ooo -