UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Year 2003/2004

September/October 2003

MGM 564 – APPLIED ECONOMETRICS
(Ekonometrik Gunaan)

Time : 3 hours

Please ensure that this examination paper consists of **TWELVE [12]** printed pages before you begin the examination.

Answer all **FOUR [4]** questions.

Sila pastikan bahawa kertas peperiksaan ini mengandungi **EMPAT [4]** soalan di dalam **DUABELAS [12]** halaman yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab kesemua **EMPAT [4]** soalan.

...2/-

1. a) It is known that the F statistic for goodness of fit test of a regression model is given by

$$F = \frac{ESS/(p-1)}{RSS/(N-p)}$$

where ESS is the explained sum of squares, RSS is the residuals sum of squares, N is the number of observations and p is the number of coefficients in the model. Show that the F statistics can also be written as

$$F = \frac{R^2/(p-1)}{(1-R^2)/(N-p)}$$

where R^2 is the coefficient of determination.

[10 marks]

b) The following table gives data on consumption (in thousand, RM) and yearly income (in thousand, RM) for 20 lecturers at a certain IPTA in the year 1988. The data was collected during holiday between the first and second semester.

Individu	1	2	3	4	5	6	7	8	9	10
Consumption, C	46.8	19.2	27.6	44.7	21.6	22.8	21.6	21.6	23.7	26.4
Income, I	48.9	20.4	25.8	45.9	26.1	23.4	21.9	24.9	28.2	32.4

Individu	11	12	13	14	15	16	17	18	19	20
Consumption, C	46.2	12.3	33.3	7.2	34.5	12.6	20.1	36.3	33.3	14.1
Income, I	55.8	15.3	34.8	8.1	35.4	13.8	22.2	38.7	39.9	17.7

Given that:

$$\sum c_i^2 = 2493.23, \quad \sum i_i^2 = 2920.21, \quad \sum c_i i_i = 2643.05$$

Assuming a simple linear relationship between I and C

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

i) What is the theoretical expectation 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. Are these coefficients satisfies *prior* theoretical expectation?

[25 marks]

...3/-

iii) Test for the overall goodness of fit of the regression model at 5% significant level. Hence, test the hypothesis $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$ at the same significant level.

[20 marks]

iv) What are the information that can be used from the values of coefficient of determination and adjusted coefficient of determination? 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 particular lecturer who is currently on a sabbatical leave has a yearly income of RM52500. Calculate the estimated consumption for the lecturer and construct a 95% confidence interval for the estimate.

[15 marks]

vi) It is also known that at the beginning of the second semester there are a few more lecturers who have given information on their yearly income and consumption. An updated regression model is estimated, this model gives coefficient of determination R^2 of 93.3% and F statistics of 332.31. Determine the number of new lecturers who have just given their information.

[10 marks]

1. a) *Diketahui bahawa statistik F bagi ujian ketepatan padanan model regresi diberikan oleh rumus di bawah*

$$F = \frac{ESS/(p-1)}{RSS/(N-p)}$$

yang ESS adalah hasil tambah kuasa dua regresi, RSS adalah hasil tambah kuasa dua reja, N ialah bilangan cerapan dan p ialah bilangan pekali dalam model regresi. Tunjukkan bahawa statistik F juga boleh dirumuskan sebagai

$$F = \frac{R^2/(p-1)}{(1-R^2)/(N-p)}$$

yang R^2 *adalah pekali penentuan.*

[10 markah]

...4/-

b) *Jadual di bawah menunjukkan perbelanjaan (dalam ribu, RM) dan juga pendapatan tahunan (dalam ribu, RM) bagi 20 pensyarah di sebuah IPTA dalam tahun 1988. Data telah dikumpulkan semasa cuti di antara semester 1 dan 2.*

Individu	1	2	3	4	5	6	7	8	9	10
Consumption, *C*	46.8	19.2	27.6	44.7	21.6	22.8	21.6	21.6	23.7	26.4
Income, *I*	48.9	20.4	25.8	45.9	26.1	23.4	21.9	24.9	28.2	32.4

Individu	11	12	13	14	15	16	17	18	19	20
Consumption, *C*	46.2	12.3	33.3	7.2	34.5	12.6	20.1	36.3	33.3	14.1
Income, *I*	55.8	15.3	34.8	8.1	35.4	13.8	22.2	38.7	39.9	17.7

Diberikan:

$$\sum c_i^2 = 2493.23, \quad \sum i_i^2 = 2920.21, \quad \sum c_i i_i = 2643.05$$

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 regresi β_1 dan β_2 dan tafsirkannya. Adakah koefisien-koefisien regresi ini memenuhi jangkaan a priori?*

[25 markah]

iii) *Ujikan ketepatan padanan keseluruhan model regresi pada aras keertian 5%. Seterusnya ujikan hipotesis $H_0 : \beta_2 = 0$ terhadap $H_1 : \beta_2 \neq 0$ pada aras keertian yang sama.*

[20 markah]

iv) *Apakah maklumat yang boleh digunakan daripada nilai-nilai pekali penentuan dan pekali penentuan terlaras? Hitung pekali penentuan dan pekali penentuan terlaras bagi model di atas dan tafsirkan.*

[15 markah]

v) *Adalah diketahui bahawa seorang pensyarah yang sedang bercuti sabatikal mempunyai pendapatan sebanyak RM52500 setahun. Hitungkan nilai anggaran perbelanjaan pensyarah tersebut serta bina selang keyakinan 95% bagi anggaran tersebut.*

[15 markah]

...5/-

vi) *Adalah diketahui pada permulaan semester kedua beberapa orang pensyarah lagi telah pun memberikan maklumat pendapatan dan perbelanjaan mereka. Model regressi terkini telah dianggarkan dan memberikan nilai pekali penentuan* R^2 *sebanyak 93.3% dan nilai statistik ujian F sebanyak 332.31. Hitungkan bilangan pensyarah yang baru memberikan maklumat mereka.*

[10 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 of the following problems

i) multicollinearity
ii) autocorrelation
iii) heteroscedasticity

[15 marks]

b) Discuss the consequences of multicollinearity on parameter estimation.

[10 marks]

c) Explain 3 methods for detecting each of the problems mentioned in (a) above.

[30 marks]

d) Explain how the 2 methods below are used to solve the problem of heteroscedasticity and autocorrelation respectively

i) assumption that error variance is proportional to X_{ji}^2, $j \in \{1, 2, \ldots, k\}$
ii) Cochran-Orcutt method

[20 marks]

e) Prove that when $k = 1$, and if first order autocorrelation or heteoscedasticity in the form $E(\varepsilon_i^2) = \sigma^2 X_{1i}^2$ is presence in the data, therefore

i) Estimated values of $\hat{\beta}_2^a$ (autocorrelation) and $\hat{\beta}_2^h$ (heteroscedasticity) remain unbiased, that is $E(\hat{\beta}_2^a) = E(\hat{\beta}_2^h) = 0$

...6/-

ii) $$Var\left(\hat{\beta}_2^a\right)=\frac{\sigma^2}{\sum x_i^2}\left(1+2\rho\frac{\sum_{i=1}^{N-1}x_i x_{i+1}}{\sum x_i^2}+2\rho^2\frac{\sum_{i=1}^{N-2}x_i x_{i+2}}{\sum x_i^2}+\ldots+2\rho^{N-1}\frac{x_1 x_N}{\sum x_i^2}\right)$$

Give explanation for a particular situation that produces

$$Var\left(\hat{\beta}_2^a\right)> Var\left(\hat{\beta}_2\right)$$

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

[25 marks]

2. *Dalam konteks sebuah model regresi 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
iii) heteroskedastisiti

[15 markah]

b) *Bincangkan kesan multikolinearan terhadap penganggaran*

[10 markah]

c) *Jelaskan 3 langkah bagi mengesan setiap masalah (a) di atas*

[30 markah]

d) *Terangkan bagaimana 2 kaedah di bawah, masing-masing dapat mengatasi masalah heteroskedastisiti dan autokorelasi*

i) andaian bahawa ralat varians berkadar terus dengan X_{ji}^2, $j\in\{1,2,\ldots,k\}$
ii) kaedah Cochran-Orcutt

[20 markah]

...7/-

e) *Buktikan bahawa apabila* $k=1$, *dan sekiranya autokorelasi peringkat pertama atau heteroskedastisiti berbentuk* $E(\varepsilon_i^2)=\sigma^2 X_{1i}^2$ *wujud dalam data, maka*

i) *nilai anggaran* $\hat{\beta}_2^a$ *(autokorelasi) dan* $\hat{\beta}_2^h$ *(heteroskedastisiti) masih lagi saksama iaitu* $E(\hat{\beta}_2^a)=E(\hat{\beta}_2^h)=0$

ii) $$Var(\hat{\beta}_2^a)=\frac{\sigma^2}{\sum x_i^2}\left(1+2\rho\frac{\sum_{i=1}^{N-1} x_i x_{i+1}}{\sum x_i^2}+2\rho^2\frac{\sum_{i=1}^{N-2} x_i x_{i+2}}{\sum x_i^2}+\ldots+2\rho^{N-1}\frac{x_1 x_N}{\sum x_i^2}\right)$$

Berikan penjelasan situasi yang akan menghasilkan

$$Var(\hat{\beta}_2^a)>Var(\hat{\beta}_2)$$

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

[25 markah]

3. The table below shows consumption on clothing (Y), total consumption (X_1), clothe price (X_2) and residuals from a multiple regression (ε_i) in a particular country starting in 1990 (in million, RM)

Year	90	91	92	93	94	95	96	97	98	99
Y	3.5	4.3	5.0	6.0	7.0	9.0	8.0	10.0	12.0	14.0
X_1	15	20	30	42	50	54	65	72	85	90
X_2	16	13	10	7	7	5	4	3	3.5	2
ε_i	-0.165	0.428	0.117	-0.216	-0.501	1.254	-1.315	-0.241	-0.428	1.066

Given that:

$$\sum Y_i = 78.8 \qquad \sum X_{1i} = 523 \qquad \sum X_{2i} = 70.5$$
$$\sum Y_i^2 = 725.74 \qquad \sum X_{1i}^2 = 33439 \qquad \sum X_{2i}^2 = 689.25$$
$$\sum X_{1i}X_{2i} = 2667.5 \qquad \sum X_{1i}Y_i = 4896.5 \qquad \sum X_{2i}Y_i = 429.9$$

$$\begin{array}{lll} Y & = 12.488 - 0.654X_2 & \\ \text{standard error} & (1.008) \quad (0.121) & \\ t \text{ statistics} & (12.39) \quad (-5.38) & r^2 = 0.748 \end{array}$$

...8/-

a) Give reasons to explain why it is mostly likely there are problems of autocorrelation, heteroscedasticity and multicollinearity in the data given above.

[10 marks]

b) Estimate the consumption function $Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \varepsilon$ and test the significance of each partial correlation coefficient at 5% level.

[30 marks]

c) Based on the residual values given in the table, test the presence of first order autocorrelation using DW statistics at 5% significance level. Use the **adjusted** Durbin-Watson test. Show that $DW = 2.66$. [Value for lower and upper limit is 0.525 and 2.016 respectively, for $N = 10, p = 3$]

[20 marks]

d) By using Spearman's rank correlation test, determine at 5% significant level whether heteroscedaticity is present in the data above.

[20 marks]

e) From the information given and the result in (b), discuss the presence of multicollinearity in the data above. Calculate the variance inflation factor and use the value to confirm your earlier expectation.

[20 marks]

f) Based on the result in (c) show that the estimated value for the autocorrelation coefficient $\rho = -0.33$. Use the estimated value to adjust the original data and re-estimate the consumption function. Test the hypothesis $H_0 : \beta_3 = 0$ against $H_1 : \beta_3 \neq 0$ at 5% significant level.

[40 marks]

g) Based on the latest consumption function calculate the (pair-wise) correlation coefficient between the adjusted independent variables X_1 and X_2. Hence, calculate the variance inflation factor and confirm that multicollinearity is present in the adjusted data. Exclude the insignificant variable and estimate the simple regression model. Prove that autocorrelation does not present in the new model.

[60 marks]

...9/-

3. *Jadual di bawah menunjukkan perbelanjaan untuk pakaian (Y), jumlah perbelanjaan (X_1), harga pakaian (X_2) dan juga ralat daripada regressi berganda (ε_i) di suatu negara mulai tahun 1990 (dalam juta, RM)*

Tahun	*90*	*91*	*92*	*93*	*94*	*95*	*96*	*97*	*98*	*99*
Y	*3.5*	*4.3*	*5.0*	*6.0*	*7.0*	*9.0*	*8.0*	*10.0*	*12.0*	*14.0*
X_1	*15*	*20*	*30*	*42*	*50*	*54*	*65*	*72*	*85*	*90*
X_2	*16*	*13*	*10*	*7*	*7*	*5*	*4*	*3*	*3.5*	*2*
ε_i	*-0.165*	*0.428*	*0.117*	*-0.216*	*-0.501*	*1.254*	*-1.315*	*-0.241*	*-0.428*	*1.066*

Diberikan:

$$\sum Y_i = 78.8 \qquad \sum X_{1i} = 523 \qquad \sum X_{2i} = 70.5$$
$$\sum Y_i^2 = 725.74 \qquad \sum X_{1i}^2 = 33439 \qquad \sum X_{2i}^2 = 689.25$$
$$\sum X_{1i}X_{2i} = 2667.5 \qquad \sum X_{1i}Y_i = 4896.5 \qquad \sum X_{2i}Y_i = 429.9$$

$$\begin{array}{lll} Y & = 12.488 - 0.654X_2 & \\ \text{sisihan piawai} & (1.008) \quad (0.121) & \\ \text{statistik } t & (12.39) \quad (-5.38) & r^2 = 0.748 \end{array}$$

a) *Berikan sebab-sebab untuk menerangkan mengapa berkemungkinan besar terdapatnya masalah autokorelasi, heteroskedastisiti dan multikolinearan dalam data di atas.*

[10 markah]

b) *Anggarkan fungsi perbelanjaan $Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \varepsilon$ dan uji keertian setiap pekali regresi separa pada aras 5%.*

[30 markah]

c) *Menggunakan nilai ralat daripada model regresi berganda yang diberi, uji kewujudan autoregresi peringkat pertama dengan menggunakan statistik DW pada aras keertian 5%. Gunakan ujian Durbin-Watson* ***terubah****. Buktikan bahawa $DW = 2.66$. [Nilai had bawah dan had atas masing-masing 0.525 dan 2.016 bagi $N = 10$, $p = 3$]*

[20 markah]

d) *Dengan menggunakan pekali korelasi pangkat Spearman, uji sama ada terdapat heteroskedastisiti dalam data pada aras keertian 5%.*

[20 markah]

e) *Daripada maklumat yang diberi dan keputusan di bahagian (b), bincangkan sekiranya terdapat multikolinearan dalam data di atas. Hitung nilai faktor inflasi varians dan gunakannya untuk mengesahkan sekiranya jangkaan kamu benar.*

[20 markah]

...10/-

f) Berdasarkan keputusan di bahagian (c) buktikan bahawa nilai anggaran pekali autokorelasi $\rho = -0.33$. Gunakan nilai anggaran yang diperolehi untuk mengubah data-data asal dan kemudian anggarkan semula fungsi perbelanjaan. Uji hipotesis $H_0 : \beta_3 = 0$ lawan $H_1 : \beta_3 \neq 0$ pada aras keertian 5%.

[40 markah]

g) Berdasarkan fungsi perbelanjaan terkini, hitungkan pekali korelasi boleh bilang di antara pembolehubah tak bersandar X_1 dan X_2 yang telah diubahsuai. Seterusnya hitung nilai faktor inflasi varians dan sahkan bahawa multikolinearan wujud dalam data yang telah diubahsuai. Gugurkan pembolehubah yang tidak beerti dan anggarkan model regresi ringkas. Buktikan bahawa autokorelasi tidak wujud dalam model terbaru ini.

[60 markah]

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

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

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

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

[15 marks]

b) Given 2 models:

$$X_t = \phi_1 X_{t-1}+\varepsilon_t \quad \text{dan} \quad X_t = \varepsilon_t-\theta_1\varepsilon_{t-1}$$

Obtain the general form of the partial autocorrelation function for the models above. If $\phi_1 = \theta_1 = 0.9$, find the partial autocorrelation for k = 1, 2, 3. Discuss on the values obtained.

[15 marks]

c) Given an AR(2) model:

$$(1-\phi_1 B-\phi_2 B^2)X_t = \varepsilon_t$$

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

i) Obtain the general form of the autocorrelation function at lag k for the model above.

[10 marks]

...11/-

ii) If it is known that a time series Y_t has a constant mean μ, show that the general form of the autocorrelation function for time series $X_t = Y_t - \mu$ is also equal to (i) above.

[10 marks]

iii) 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

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

If estimated values for the coefficients based on time series of 100 observations are $\hat{\phi}_1 = 1.5$, $\hat{\phi}_2 = -0.6$, $\hat{\mu} = 10$, $s_\varepsilon^2 = 0.4$ with $Y_{99} = 11.5$ and $Y_{100} = 11.2$, obtain value of $\hat{Y}_{100}(m)$ for m = 1, 2, ..., 5. Construct a 95% confidence interval for Y_{101} and Y_{102}. Give comment on the 5 forecast values obtained above. What is the likely value for $\hat{Y}_{100}(m \to \infty)$? Give explanation.

[50 marks]

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

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

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

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

[15 markah]

b) Diberikan 2 model berikut:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t \text{ dan } X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Dapatkan ungkapan umum fungsi autokorelasi separa bagi model-model di atas. Sekiranya $\phi_1 = \theta_1 = 0.9$, *dapatkan autokorelasi separa bagi k = 1, 2, 3. Bincangkan nilai-nilai yang diperolehi.*

[15 markah]

c) Diberikan model AR(2) berikut:

$$(1 - \phi_1 B - \phi_2 B^2) X_t = \varepsilon_t$$

ε_t *ialah proses hingar putih dengan min 0 dan varians* σ_ε^2

...12/-

i) *Dapatkan ungkapan umum fungsi autokorelasi susulan k bagi model di atas.*

[10 markah]

ii) *Sekiranya diketahui siri masa Y_t mempunyai min malar μ, tunjukkan bahawa ungkapan umum fungsi autokorelasi siri masa $X_t = Y_t - \mu$ sama seperti (i) di atas.*

[10 markah]

iii) *Terbitkan ungkapan telahan satu langkah kehadapan $\hat{Y}_t(1)$, dan tunjukkan bahawa telahan m langkah ke hadapan diberi oleh*

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

Sekiranya anggaran kepada parameter model berdasarkan satu siri masa yang terdiri daripada 100 cerapan diberi sebagai $\hat{\phi}_1 = 1.5$, $\hat{\phi}_2 = -0.6$, $\hat{\mu} = 10$, $s_\varepsilon^2 = 0.4$, dengan $Y_{99} = 11.5$ dan $Y_{100} = 11.2$, dapatkan nilai $\hat{Y}_{100}(m)$ bagi m = 1, 2 , ..., 5. Bina selang keyakinan 95% bagi Y_{101} dan Y_{102}. Berikan komen terhadap 5 nilai telahan yang diperolehi. Apakah nilai berkemungkinan bagi $\hat{Y}_{100}(m \to \infty)$? Beri penjelasan.

[50 markah]

- oooOooo -