
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2005/2006

April/May 2006

MST 565 – Linear Models
[Model Linear]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer all four [4] questions.

[Arahan: Jawab semua empat [4] soalan].

.../2-

1. (a) Explain the following terms:

- (i) Less than full rank model
- (ii) Method of least squares
- (iii) Estimable functions
- (iv) Testable hypothesis

Give examples to illustrate your answers.

[80 marks]

(b) Show that the matrix \mathbf{A} below is positive definite, and find a matrix \mathbf{P} such that $\mathbf{P}'\mathbf{P} = \mathbf{A}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

[20 marks]

1. (a) Terangkan sebutan-sebutan berikut:

- (i) Model pangkat kurang penuh
- (ii) Kaedah kuasa dua terkecil
- (iii) Fungsi teranggar
- (iv) Hipotesis teruji

Berikan contoh-contoh untuk mengilustrasi jawapan anda.

[80 markah]

(b) Tunjukkan bahawa matriks \mathbf{A} di bawah adalah tentu positif, dan cari suatu matriks \mathbf{P} sehingga demikian $\mathbf{P}'\mathbf{P} = \mathbf{A}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

[20 markah]

...3/-

2. (a) Let $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ be a random vector with $\boldsymbol{\mu} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Assume that $\sigma_{21} = 3$, $\sigma_1^2 = 4$, and $\sigma_2^2 = 5$. Find $\text{var } \mathbf{y}$.

[20 marks]

- (b) Let

$$\mathbf{A} = \begin{bmatrix} 6 & 1 \\ 4 & 2 \end{bmatrix}.$$

Find $E[\mathbf{y}'\mathbf{A}\mathbf{y}]$.

[20 marks]

- (c) Verify your answer to part (b) by expanding $\mathbf{y}'\mathbf{A}\mathbf{y}$ and taking the expectation of this quadratic form.

[20 marks]

- (d) Consider

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } \mathbf{X}'\mathbf{y} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

for a particular linear model.

- (i) Show that $|\mathbf{X}'\mathbf{X}| = 0$ and hence that $\mathbf{X}'\mathbf{X}$ is singular.
- (ii) Show that $\text{rank}(\mathbf{X}'\mathbf{X}) = 1$.
- (iii) Show that the system of normal equations is consistent.

[40 marks]

...4/-

2. (a) Katakan $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ sebagai suatu vektor rawak dengan $\boldsymbol{\mu} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Andaikan bahawa $\sigma_{21} = 3$, $\sigma_1^2 = 4$, dan $\sigma_2^2 = 5$. Cari var \mathbf{y} .

[20 markah]

- (b) Katakan

$$\mathbf{A} = \begin{bmatrix} 6 & 1 \\ 4 & 2 \end{bmatrix}.$$

Cari $E[\mathbf{y}'\mathbf{A}\mathbf{y}]$.

[20 markah]

- (c) Tentusahkan jawapan anda kepada bahagian (b) dengan mengembangkan $\mathbf{y}'\mathbf{A}\mathbf{y}$ dan mengambil jangkaan bentuk kuadratik ini.

[20 markah]

- (d) Pertimbangkan

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ dan } \mathbf{X}'\mathbf{y} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

bagi sesuatu model linear.

- (i) Tunjukkan bahawa $|\mathbf{X}'\mathbf{X}| = 0$ dan seterusnya tunjukkan bahawa $\mathbf{X}'\mathbf{X}$ adalah singular.
- (ii) Tunjukkan bahawa pangkat($\mathbf{X}'\mathbf{X}$) = 1.
- (iii) Tunjukkan bahawa sistem persamaan normal tersebut adalah konsisten.

[40 markah]

3. (a) The data in Table 3.1 relate grams plant dry weight Y to percent soil organic matter X_1 and kilograms of supplemental soil nitrogen added per 1,000 square metres X_2 .

Table 3.1. Plant Dry Weight Data.

Y	X_1	X_2
78.5	7	2.6
74.3	1	2.9
104.3	11	5.6
87.6	11	3.1
95.9	7	5.2
109.2	11	5.5
102.7	3	7.1

- (i) Propose a suitable model for the data in Table 3.1.
- (ii) What is the X matrix for such a model?
- (iii) Find $X'X$, $(X'X)^{-1}$ and $X'Y$.
- (iv) Find the least squares estimator for β .
- (v) Write the regression equation.
- (vi) Find s^2 and var b_1 .
- (vii) Find and interpret the 95% confidence interval on β_1 .

[50 marks]

- (b) Box, Hunter and Hunter (1978) describe a two-factor experiment in which all combinations of three poisons and four antidotes are considered. Four replicates (laboratory animals) are randomly allocated to each treatment combination. The survival times are shown in Table 3.2.

.../6-

Table 3.2. Box, Hunter and Hunter (1978) Data.

Poison	Antidote			
	A	B	C	D
1	0.31	0.82	0.43	0.45
	0.45	1.10	0.45	0.71
	0.46	0.88	0.63	0.66
	0.43	0.72	0.76	0.62
2	0.36	0.92	0.44	0.56
	0.29	0.61	0.35	1.02
	0.40	0.49	0.31	0.71
	0.23	1.24	0.40	0.38
3	0.22	0.30	0.23	0.30
	0.21	0.37	0.25	0.35
	0.18	0.38	0.24	0.31
	0.23	0.29	0.22	0.33

Derive an ANOVA table for these data and interpret the findings.

[50 marks]

3. (a) Data dalam Jadual 3.1 mengait Y berat kering tumbuhan (gram) dengan X_1 bahan organik tanah dan X_2 nitrogen tanah tambahan (kilogram) seseribu m^2 .

Jadual 3.1. Data Berat Kering Tumbuhan.

Y	X_1	X_2
78.5	7	2.6
74.3	1	2.9
104.3	11	5.6
87.6	11	3.1
95.9	7	5.2
109.2	11	5.5
102.7	3	7.1

.../7-

- (i) Cadangkan suatu model yang sesuai bagi data dalam Jadual 3.1.
- (ii) Apakah matriks X bagi model seperti ini?
- (iii) Cari $X'X$, $(X'X)^{-1}$ dan $X'Y$.
- (iv) Cari penganggar kuasa dua terkecil bagi β .
- (v) Tulis persamaan regresi.
- (vi) Cari s^2 dan var b_1 .
- (vii) Cari dan tafsir selang keyakinan 95% bagi β_1 .

[50 markah]

- (b) Box, Hunter dan Hunter (1978) menghurai suatu ujikaji dua faktor yang mana semua gabungan bagi tiga racun dan empat ubat penawar dipertimbangkan. Empat ukuran (haiwan makmal) diumpukkan secara rawak kepada setiap gabungan rawatan. Masa hayat ditunjukkan dalam Jadual 3.2.

Jadual 3.2. Data Box, Hunter dan Hunter (1978).

		Antidote		
Poison	A	B	C	D
1	0.31	0.82	0.43	0.45
	0.45	1.10	0.45	0.71
	0.46	0.88	0.63	0.66
	0.43	0.72	0.76	0.62
2	0.36	0.92	0.44	0.56
	0.29	0.61	0.35	1.02
	0.40	0.49	0.31	0.71
	0.23	1.24	0.40	0.38
3	0.22	0.30	0.23	0.30
	0.21	0.37	0.25	0.35
	0.18	0.38	0.24	0.31
	0.23	0.29	0.22	0.33

Dapatkan jadual ANOVA yang lengkap bagi data ini dan tafsirkan hasilnya.

[50 markah]

.../8-

4. (a) Table 4.1 shows plasma inorganic phosphate levels (mg/dl) one hour after a standard glucose tolerance test for obese subjects, with or without hyperinsulinemia, and controls (data from Jones, 1987).

Table 4.1. Plasma Phosphate Levels in Obese and Control Subjects.

Hyperinsulinemic Obese	Non-hyperinsulinemic Obese	Controls
2.3	3.0	3.0
4.1	4.1	2.6
4.2	3.9	3.1
4.0	3.1	2.2
4.6	3.3	2.1
4.6	2.9	2.4
3.8	3.3	2.8
5.2	3.9	3.4
3.1		2.9
3.7		2.6
3.8		3.1
		3.2

- (i) Assuming that the one-way classification model with fixed effects is appropriate, use these data to test the hypothesis that there are no mean differences among the three groups. What is your conclusion?
- (ii) Obtain a 95% confidence interval for the difference in means between the two obese groups and comment on the interval.

[50 marks]

- (b) The data displayed in Table 4.2 was collected during a study of a psychiatric screening questionnaire, the General Health Questionnaire or GHQ (see Goldberg, 1972), designed to help identify possible psychiatric 'caseness'. Interest lies in assessing whether GHQ score is predictive of 'caseness' and whether the sex (0 = Female, 1 = Male) of a subject plays a role in this prediction.

.../9-

Table 4.2. Psychiatric Data.

GHQscore	Sex	Case	No Case
0	0	4	80
1	0	4	29
2	0	8	15
3	0	6	3
4	0	4	2
5	0	6	1
6	0	3	1
7	0	2	0
8	0	3	0
9	0	2	0
10	0	1	0
0	1	1	36
1	1	2	25
2	1	2	8
3	1	1	4
4	1	3	1
5	1	3	1
6	1	2	1
7	1	4	2
8	1	3	1
9	1	2	0
10	1	2	0

Perform an appropriate analysis on the psychiatric data. What are your conclusions?

[50 marks]

4. (a) *Jadual 4.1 menunjukkan paras fosfat inorganik plasma (mg/dl) sejam selepas suatu ujian toleransi glukos piawai bagi subjek-subjek sangat gemuk dengan atau tanpa hiperinsulinemik dan subjek kawal (data daripada Jones, 1972).*

.../10-

Jadual 4.1. Paras Fosfat Plasma dalam Subjek Sangat Gemuk dan Subjek Kawal.

<i>Subjek Sangat Gemuk Dengan Hiperinsulinemik</i>	<i>Subjek Sangat Gemuk Tanpa Hiperinsulinemik</i>	<i>Subjek Kawal</i>
2.3	3.0	3.0
4.1	4.1	2.6
4.2	3.9	3.1
4.0	3.1	2.2
4.6	3.3	2.1
4.6	2.9	2.4
3.8	3.3	2.8
5.2	3.9	3.4
3.1		2.9
3.7		2.6
3.8		3.1
		3.2

- (i) Andaikan bahawa model pengelasan sehalu dengan kesan tetap adalah sesuai. Gunakan data ini untuk menguji hipotesis bahawa tidak wujud perbezaan min antara ketiga-tiga kumpulan. Apakah kesimpulan anda?
- (ii) Bina selang keyakinan 95% bagi perbezaan min antara kedua-dua kumpulan sangat gemuk dan komen tentang selang yang diperolehi.

[50 markah]

- (b) Data yang dipaparkan dalam Jadual 4.2 dikutip semasa suatu kajian maklumbalas penskrinan psikiatrik, Maklumbalas Kesihatan Am (the General Health Questionnaire or GHQ (lihat Goldberg, 1972)), direka untuk membantu mengenalpasti kes psikiatrik yang mungkin. Penyelidik berminat menilai sama ada skor GHQ dapat meramal kes psikiatrik dan sama ada jantina (0 = Perempuan, 1 = Lelaki) seorang subjek memainkan peranan dalam ramalan ini.

.../11-

Jadual 4.2. Data Psikiatrik.

<i>SkorGHQ</i>	<i>Jantina</i>	<i>Kes</i>	<i>Tiada Kes</i>
0	0	4	80
1	0	4	29
2	0	8	15
3	0	6	3
4	0	4	2
5	0	6	1
6	0	3	1
7	0	2	0
8	0	3	0
9	0	2	0
10	0	1	0
0	1	1	36
1	1	2	25
2	1	2	8
3	1	1	4
4	1	3	1
5	1	3	1
6	1	2	1
7	1	4	2
8	1	3	1
9	1	2	0
10	1	2	0

Laksanakan analisis yang sesuai bagi data psikiatrik. Apakah kesimpulan anda?

[50 markah]

- 000 O 000 -