UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2016/2017 Academic Session

June 2017

MST 565 - MODEL LINEAR
[LINEAR MODELS]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **FOUR** (4) questions.

*[**Arahan:** Jawab **semua empat** (4) soalan.]*

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai.]

Question 1

Explain the following terms:

(a) Symmetric matrix

(b) Idempotent matrix

(c) Less than full rank model

(d) Estimable functions

(e) Testable hypothesis

Give examples to illustrate your answers.

[15 marks]

Soalan 1

Jelaskan sebutan-sebutan berikut:

(a) Matriks simetri

(b) Matriks Idempoten

(c) Model pangkat kurang penuh

(d) Fungsi teranggar

(e) Hipotesis teruji

Berikan contoh untuk gambarkan jawapan anda.

[15 markah]

Question 2

(a) Let $\mathbf{y} = (y_1, y_2, y_3)'$ be a random vector with mean vector and covariance matrix

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix}$$

(i) Which variables are independent?

(ii) Let $z_1 = y_1 + 2y_2 + y_3$ and $z_2 = 3y_1 + y_2 - 2y_3$. Find $E(\mathbf{z})$ and $\text{cov}(\mathbf{z})$, where $\mathbf{z} = (z_1, z_2)'$.

(iii) Define $\mathbf{w} = (w_1, w_2, w_3)'$ as follows:

$$w_1 = 2y_1 - y_2 + y_3$$
$$w_2 = y_1 + 2y_2 - 3y_3$$
$$w_3 = y_1 + y_2 + 2y_3$$

Then, by using $\mathbf{z}$ as defined in (ii), find $\text{COV}(\mathbf{z}, \mathbf{w})$.

[13 marks]

(b) Let the random vector $\mathbf{v}$ be $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given as

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{pmatrix}.$$

If $\mathbf{v}$ is partitioned as $\mathbf{v} = (y_1, y_2, x_1, x_2)'$,

(i) what is the partitioned form of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$?

(ii) find the covariance matrix of partial correlations, $\boldsymbol{\Sigma}_{y \cdot x}$ and the matrix of partial correlation, $\mathbf{P}_{y \cdot x}$.

[12 marks]

(c) If $\mathbf{y}$ is $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

find a symmetric matrix $\mathbf{A}$ such that $\mathbf{y}'\mathbf{A}\mathbf{y}$ is $\chi^2(3, \lambda)$. Then find the value of the noncentrality parameter, λ .

[5 marks]

Soalan 2

(a) *Biar* $\mathbf{y} = (y_1, y_2, y_3)'$ *sebagai vektor rawak dengan vektor min and matriks kovarians*

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 3 & 10 \end{pmatrix}.$$

(i) *Pemboleh ubah yang mana adalah tidak bersandar?*

(ii) *Biar* $z_1 = y_1 + 2y_2 + y_3$ *dan* $z_2 = 3y_1 + y_2 - 2y_3$. *Cari* $E(\mathbf{z})$ *dan* $\text{cov}(\mathbf{z})$ *yang mana* $\mathbf{z} = (z_1, z_2)'$.

(iii) *Tafsirkan* $\mathbf{w} = (w_1, w_2, w_3)'$ *seperti berikut:*

$$w_1 = 2y_1 - y_2 + y_3$$
$$w_2 = y_1 + 2y_2 - 3y_3$$
$$w_3 = y_1 + y_2 + 2y_3$$

Kemudian, dengan menggunakan $\mathbf{z}$ *seperti dalam (ii), cari* $\text{cov}(\mathbf{z}, \mathbf{w})$.

[13 markah]

(b) *Biar vektor rawak* $\mathbf{v}$ *sebagai* $N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *yang mana* $\boldsymbol{\mu}$ *dan* $\boldsymbol{\Sigma}$ *diberikan sebagai*

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 5 \\ -2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -3 & 7 \end{pmatrix}.$$

Jika $\mathbf{v}$ *terpetak sebagai* $\mathbf{v} = (y_1, y_2, x_1, x_2)'$,

(i) *apakah bentuk terpetak bagi* $\boldsymbol{\mu}$ *dan* $\boldsymbol{\Sigma}$ *?*

(ii) *cari matriks kovarians bagi korelasi separa,* $\boldsymbol{\Sigma}_{y \cdot x}$ *dan matriks korelasi separa* $\mathbf{P}_{y \cdot x}$.

[12 markah]

(c) *Jika* $\mathbf{y}$ *adalah* $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *yang mana*

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

cari matrik $\mathbf{A}$ *yang simetri supaya* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *adalah* $\chi^2(3, \lambda)$. *Kemudian cari nilai parameter tak berpusat,* λ.

[5 markah]

Question 3

(a) The data in **Table 3.1** relate grams of plant dry weight y to percent of 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 (gram)	x_1 (%)	x_2 (kg)
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 $\mathbf{X}$ matrix for such a model?

(iii) Find $\mathbf{X}'\mathbf{X}$, $(\mathbf{X}'\mathbf{X})^{-1}$ and $\mathbf{X}'\mathbf{y}$.

(iv) Find the least squares estimator for $\boldsymbol{\beta}$.

(v) Write the fitted regression equation.

[20 marks]

(b) In an effort to obtain the maximum yield in a chemical reaction, the value of the following variables were chosen by the experimenter:

x_1 = temperature (^0C)
x_2 = concentration of a reagent (%)
x_3 = time of reaction (hours)

with one response variable was observed:

y = percent of unchanged starting material

The data are in **Table 3.2**.

Table 3.2. Chemical Reaction Data

y	x_1	x_2	x_3
41.5	162	23	3
33.8	162	23	8
27.7	162	30	5
21.7	162	30	8
19.9	172	25	5
15.0	172	25	8
12.2	172	30	5
4.3	172	30	8
19.3	167	27.5	6.5
6.4	177	27.5	6.5
37.6	157	27.5	6.5
18.0	167	32.5	6.5
26.3	167	22.5	6.5
9.9	167	27.5	9.5
25.0	167	27.5	3.5
14.1	177	20	6.5
15.2	177	20	6.5
15.9	160	34	7.5
19.6	160	34	7.5

(i) Find $\text{cov}\left(\hat{\boldsymbol{\beta}}\right)$.

(ii) Discuss the differences between R^2 and R_a^2 (adjusted R^2) of the fitted model.

(iii) Consider the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$, test the hypothesis $H_0 : 2\beta_1 = 2\beta_2 = \beta_3$ by expressing H_0 in the form $H_0 : \mathbf{C\beta} = \mathbf{0}$ versus $H_1 : \mathbf{C\beta} \neq \mathbf{0}$.

[20 marks]

Soalan 3

(a) *Data dalam **Jadual 3.1** menghuraikan y gram berat kering tanaman terhadap x_1 peratus bahan organik tanah dan x_2 kilogram tanah tambahan bernitrogen per seribu m^2 .*

Jadual 3.1. Data Berat Kering Tumbuhan

y (gram)	x_1 (%)	x_2 (kg)
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) *Cadangkan suatu model yang sesuai untuk data dalam **Jadual 3.1**.*

(ii) *Apakah matriks $\mathbf{X}$ bagi model tersebut?*

(iii) *Cari $\mathbf{X}'\mathbf{X}$, $\left(\mathbf{X}'\mathbf{X}\right)^{-1}$ and $\mathbf{X}'\mathbf{y}$.*

(iv) *Cari penganggar kuasa dua terkecil bagi $\boldsymbol{\beta}$.*

(v) *Tulis persamaan regresi terpadankan tersebut.*

[20 markah]

(b) *Dalam usaha untuk mendapatkan hasil yang maksimum dalam suatu tindakbalas kimia, nilai pemboleh ubah berikut telah dipilih oleh seorang pengkaji:*

x_1 *= suhu (*^0C *)*

x_2 *= kepekatan reagen (%)*

x_3 *= masa tindakbalas (jam)*

dengan satu pemboleh ubah respon :

y *= peratusan bahan asal yang tidak berubah*

*Data adalah dalam **Jadual 3.2**.*

Jadual 3.2. Data Tindakbalas Kimia

y_1	y_2	x_1	x_2	x_3
41.5	*45.9*	*162*	*23*	*3*
33.8	*53.3*	*162*	*23*	*8*
27.7	*57.5*	*162*	*30*	*5*
21.7	*58.8*	*162*	*30*	*8*
19.9	*60.6*	*172*	*25*	*5*
15.0	*58.8*	*172*	*25*	*8*
12.2	*58.6*	*172*	*30*	*5*
4.3	*52.4*	*172*	*30*	*8*
19.3	*56.9*	*167*	*27.5*	*6.5*
6.4	*55.4*	*177*	*27.5*	*6.5*
37.6	*46.9*	*157*	*27.5*	*6.5*
18.0	*57.3*	*167*	*32.5*	*6.5*
26.3	*55.0*	*167*	*22.5*	*6.5*
9.9	*58.9*	*167*	*27.5*	*9.5*
25.0	*50.3*	*167*	*27.5*	*3.5*
14.1	*61.1*	*177*	*20*	*6.5*
15.2	*62.9*	*177*	*20*	*6.5*
15.9	*60.0*	*160*	*34*	*7.5*
19.6	*60.6*	*160*	*34*	*7.5*

(i) *Cari cov* $\left(\hat{\boldsymbol{\beta}}\right)$.

(ii) *Bincangkan perbezaan antara* R^2 *dan* R_a^2 *(* R^2 *terlaras) untuk model yang disuaikan.*

(iii) *Pertimbangkan model* $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$, *uji hipotesis* $H_0 : 2\beta_1 = 2\beta_2 = \beta_3$ *dalam bentuk* $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ *lawan* $H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0}$

[20 markah]

Question 4

Table 4.1 shows plasma inorganic phosphate levels (mg/dl) one hour after a standard gluscose 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	3.1
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

(a) 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?

[7 marks]

(b) Obtain a 95% confidence interval for the difference in means between the two obese groups and comment on the interval.

[8 marks]

Soalan 4

Jadual 4.1 *menunjukkan tahap fosfat tak organik plasma (mg/dl) sejam selepas suatu ujian piawai bagi toleransi glukosa terhadap subjek-subjek yang mempunyai berat badan berlebihan dengan atau tanpa hiperinsulinemik dan subjek kawalan (data daripada Jones, 1987).*

Jadual 4.1. Tahap Fosfat dalam Plasma Subjek yang mempunyai Berat Badan Berlebihan dan Subjek Kawalan.

Hyperinsulinemic Obes	*Tanpa- hyperinsulinemic Obes*	*Kawalan*
2.3	*3.0*	*3.1*
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

(a) Andaikan bahawa pengkelasan sehala dengan kesan tetap adalah sesuai, gunakan data tersebut untuk menguji hipotesis bahawa tidak wujud perbezaan min antara ketiga-tiga kumpulan. Apakah kesimpulan anda?

[7 markah]

(b) Dapatkan selang keyakinan 95% bagi perbezaan min antara dua kumpulan subjek yang mempunyai berat berlebihan dan berikan komen tentang selang keyakinan tersebut.

[8 markah]

- ooo 00 ooo -