

UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2004/2005

March 2005

MST 565 – LINEAR MODEL
[MODEL LINEAR]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of **ELEVEN [11]** pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **SEBELAS [11]** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

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

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

...2/-

1. (a) Explain the following terms:

- (i) non-central chi-squared distribution
- (ii) non-singular
- (iii) positive semidefinite
- (iv) orthogonal

Give examples to illustrate your answers.

[40 marks]

(b) Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & 5 \\ -3 & 6 & 7 \\ 5 & 7 & -8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 3 & -4 \\ -3 & 0 & 5 \\ 4 & -5 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- (i) Are these matrices symmetric? Explain your answers.
- (ii) Suppose \mathbf{A} and \mathbf{B} are symmetric. Are $\mathbf{A} + \mathbf{B}$ and \mathbf{AB} symmetric?
- (iii) Suppose

$$\mathbf{A} = \begin{bmatrix} 1/9 & 8/9 & -4/9 \\ 4/9 & -4/9 & -7/9 \\ 8/9 & 1/9 & 4/9 \end{bmatrix}.$$

Is \mathbf{A} orthogonal?

- (iv) Let

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \end{bmatrix}$$

where \mathbf{P} is orthogonal. Show that $\mathbf{P}_1^T \mathbf{P}_1$ is idempotent.

[60 marks]

1. (a) Terangkan sebutan-sebutan berikut:

- (i) taburan khi-kuasa tak memusat
- (ii) tak singular
- (iii) semitentu positif
- (iv) berortogon

Berikan contoh-contoh untuk mengilustrasi jawapan anda.

[40 markah]

(b) Pertimbangkan matriks-matriks yang berikut:

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & 5 \\ -3 & 6 & 7 \\ 5 & 7 & -8 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 3 & -4 \\ -3 & 0 & 5 \\ 4 & -5 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- (i) Adakah matriks-matriks ini bersimetri? Terangkan jawapan anda.
 (ii) Katakan \mathbf{A} dan \mathbf{B} adalah bersimetri. Adakah $\mathbf{A} + \mathbf{B}$ dan \mathbf{AB} bersimetri?
 (iii) Katakan

$$\mathbf{A} = \begin{bmatrix} 1/9 & 8/9 & -4/9 \\ 4/9 & -4/9 & -7/9 \\ 8/9 & 1/9 & 4/9 \end{bmatrix}.$$

Adakah \mathbf{A} berortogon?

- (iv) Biarkan

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \end{bmatrix}$$

di mana \mathbf{P} adalah berortogon. Tunjukkan bahawa $\mathbf{P}_1' \mathbf{P}_1$ adalah idempotent.

[60 markah]

2. (a) Let y_1, y_2, y_3 be a random sample from a distribution that has mean 3 and variance 2. Let

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

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

[20 marks]

- (b) Let

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

be a multivariate normal random variable with mean $\mathbf{0}$ and variance-covariance matrix

$$\mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

...4/-

Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

- (i) Write $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ in summation form.
- (ii) What is the distribution of $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$?
- (iii) Are $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ independent?

[40 marks]

(c) Suppose that for a particular linear model,

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}.$$

- (i) Show that $|\mathbf{X}'\mathbf{X}| = 0$ and hence that $(\mathbf{X}'\mathbf{X})$ is singular.
- (ii) Show that $r(\mathbf{X}'\mathbf{X}) = 2$.
- (iii) Show that the system of normal equations is consistent.

[40 marks]

2. (a) Biarkan y_1, y_2, y_3 sebagai sebuah sampel rawak daripada satu taburan yang mempunyai min 3 dan varians 2. Biarkan

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

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

[20 markah]

(b) Biarkan

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

menjadi sebuah pembolehubah rawak normal multivariat dengan min 0 dan matriks varians-kovarians

$$\mathbf{V} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

...5/-

Biarkan

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ dan } \mathbf{B} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

- (i) Tulis $\mathbf{y}'\mathbf{A}\mathbf{y}$ dan $\mathbf{y}'\mathbf{B}\mathbf{y}$ dalam bentuk hasil tambah.
 (ii) Apakah taburan $\mathbf{y}'\mathbf{A}\mathbf{y}$ dan $\mathbf{y}'\mathbf{B}\mathbf{y}$?
 (iii) Adakah $\mathbf{y}'\mathbf{A}\mathbf{y}$ dan $\mathbf{y}'\mathbf{B}\mathbf{y}$ tak bersandar?

[40 markah]

(c) Katakan untuk sebuah model linear tertentu,

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \text{ dan } \mathbf{X}'\mathbf{y} = \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}.$$

- (i) Tunjukkan bahawa $|\mathbf{X}'\mathbf{X}| = 0$ dan seterusnya tunjukkan bahawa $(\mathbf{X}'\mathbf{X})$ adalah singular.
 (ii) Tunjukkan bahawa $r(\mathbf{X}'\mathbf{X}) = 2$.
 (iii) Tunjukkan bahawa system persamaan normal tersebut adalah konsisten.

[40 markah]

3. (a) In a simple linear regression we might want to model the amount of diesel fuel used by trucks as a function only of the time during which their motors are running. It is evident that if $x=0$, the motor does not run at all, then $y=0$, i.e. no fuel is used. The desired model should have intercept 0 and hence should take the form

$$\mathbf{y} = \beta_1 \mathbf{x} + \varepsilon.$$

Consider this model for the data in Table 3.1.

Table 3.1. Fuel Consumption data.

Amount of fuel in gallons (y)	Time motor runs in hours (x)
3	0.6
5	2.0
7	2.1
9	2.0
10	2.4

- (i) What is the X matrix for such a model?
- (ii) Find $X'X$, $(X'X)^{-1}$ and $X'Y$.
- (iii) Find the least squares estimator for β_1 .
- (iv) Find s^2 and var b_1 .
- (v) Find and interpret the 95% confidence interval on β_1 .

Note: To use SAS to fit a model through the origin, the model statement is

MODEL Y + X / XPX I P NOINT;

[50 marks]

- (b) Ozonation as a secondary treatment for affluent following absorption by ferrous chloride was studied for three reactions and three PH levels. These data shown in Table 3.2 are obtained on effluent decline.

Table 3.2. Effluent data.

	PH Level		
Reaction Time (mins.)	7	9	10.5
20	23	16	14
	21	18	13
	22	15	16
40	20	14	12
	22	13	11
	19	12	10
60	21	13	11
	20	12	13
	19	12	12

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

[50 marks]

3. (a) Dalam suatu regresi linear ringkas kita mungkin ingin memodelkan amaun minyak diesel yang digunakan oleh truk sebagai fungsi hanya masa ketika motor-motor mereka sedang beroperasi. Adalah jelas bahawa jika $x=0$, motor tidak akan beroperasi, maka $y=0$, iaitu tiada minyak digunakan. Model yang diinginkan harus mempunyai pintasan 0 dan seharusnya mengambil bentuk

$$y = \beta_1 x + \varepsilon.$$

Pertimbangkan model ini bagi data dalam Jadual 3.1.

Jadual 3.1. Data penggunaan minyak .

Amaun minyak dalam gallon (y)	Masa motor beroperasi dalam jam (x)
3	0.6
5	2.0
7	2.1
9	2.0
10	2.4

- (i) Apakah matriks X bagi model seperti ini?
- (ii) Cari $X'X$, $(X'X)^{-1}$ dan $X'Y$.
- (iii) Cari penganggar kuasa dua terkecil bagi β_1 .
- (iv) Cari s^2 dan $\text{var } b_1$.
- (v) Cari dan tafsir selang keyakinan 95% bagi β_1 .

Nota: Untuk menggunakan SAS bagi penyuaian model melalui asalan, pernyataan model ialah

MODEL Y + X / XPX I P NOINT;

[50 markah]

- (b) Ozonasi sebagai suatu rawatan yang sekunder bagi sisa mengikut penyerapan oleh 'ferrous chloride' telah dikaji bagi tiga waktu reaksi dan tiga aras PH. Data yang ditunjukkan dalam Jadual 3.2 diperoleh bagi kemerosotan sisa.

Jadual 3.2. Data sisa.

Reaction Time (mins.)	PH Level		
	7	9	10.5
20	23	16	14
	21	18	13
	22	15	16
40	20	14	12
	22	13	11
	19	12	10
60	21	13	11
	20	12	13
	19	12	12

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

[50 markah]

4. (a) The data displayed in Table 4.1 was collected to examine the extent to which erythrocyte sedimentation rate (ESR) (i.e., the rate at which red blood cells [erythrocytes] settle out of suspension in blood plasma) is related to two plasma proteins: fibrinogen and γ -globulin, both measured in gm/l. The ESR for a 'healthy' individual should be less than 20 mm/h and, because the absolute value of ESR is relatively unimportant, the response variable used here denotes whether or not this is the case. A response of zero signifies a healthy individual ($ESR < 20$), while a response of unity refers to an unhealthy individual ($ESR \geq 20$). The aim of the analysis for these data is to determine the strength of any relationship between ESR level and the levels of the two plasmas (these data are in the SPSS file, 'ESR.sav', in the diskette provided).

Table 4.1. ESR data.

Fibrinogen	γ -Globulin	ESR
2.52	38	0
2.56	31	0
2.19	33	0
2.18	31	0
3.41	37	0
2.46	36	0
3.22	38	0
2.21	37	0
3.15	39	0
2.60	41	0
2.29	36	0
2.35	29	0
5.06	37	1
3.34	32	1
2.38	37	1
3.15	36	0
3.58	46	1
2.68	34	0
2.60	38	0
2.23	37	0
2.88	30	0
2.65	46	0
2.09	44	1
2.28	36	0
2.67	39	0
2.29	31	0
2.15	31	0
2.54	28	0
3.93	32	1
3.34	30	0
2.99	36	0
3.32	35	0

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

[50 marks]

- (b) A study of sleep deprivation is conducted. Twenty-four subjects are randomly divided into three groups of size eight each. After differing amounts of sleep deprivation, all subjects are asked to perform a task that requires manual dexterity. A score from 0 (poor) to 10 (excellent) is recorded. Table 4.2 displays the data obtained.

Table 4.2. Sleep Deprivation data.

Group I (16 hours deprivation)	Group II (20 hours deprivation)	Group III (24 hours deprivation)
8.95	7.70	5.99
8.04	5.81	6.79
7.72	6.61	6.43
6.21	6.07	5.85
6.48	8.04	5.78
7.81	5.96	7.60
7.50	7.30	5.78
6.90	7.46	6.00

Assuming that the one-way classification model with fixed effects is appropriate, use these data to test $H_0: \tau_1 = \tau_2 = \tau_3$. What is your conclusion?

[50 marks]

4. (a) *Data yang dipaparkan dalam Jadual 4.1 dikutip untuk menyiasat sejauh manakah kadar mendak eritrosit (ESR) (iaitu, kadar di mana sel-sel darah merah [eritrosit] tenggelam dari mendak dalam plasma darah) berkait dengan dua protin plasma: fibrinogen dan globulin- γ , kedua-duanya disukat dalam gm/l. ESR bagi seorang individu yang 'sihat' haruslah kurang daripada 20 mm/h dan, kerana nilai mutlak ESR adalah tidak penting secara relatif, pembolehubah sambutan yang digunakan di sini menandakan sama ada perkara ini benar atau tidak. Sambutan sifar menandakan seorang individu yang sihat ($ESR < 20$), manakala sambutan satu merujuk kepada individu yang tidak sihat ($ESR \geq 20$). Tujuan analisis data ini ialah untuk menentukan kekuatan hubungan antara aras ESR dan aras-aras dua plasma tersebut (data ini adalah di dalam fail SPSS, 'ESR.sav', dalam disket yang diberikan).*

Jadual 4.1. Data ESR.

Fibrinogen	γ -Globulin	ESR
2.52	38	0
2.56	31	0
2.19	33	0
2.18	31	0
3.41	37	0
2.46	36	0
3.22	38	0
2.21	37	0
3.15	39	0
2.60	41	0
2.29	36	0
2.35	29	0
5.06	37	1
3.34	32	1
2.38	37	1
3.15	36	0
3.58	46	1
2.68	34	0
2.60	38	0
2.23	37	0
2.88	30	0
2.65	46	0
2.09	44	1
2.28	36	0
2.67	39	0
2.29	31	0
2.15	31	0
2.54	28	0
3.93	32	1
3.34	30	0
2.99	36	0
3.32	35	0

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

[50 markah]

- (b) Satu kajian perlucutan tidur (sleep deprivation) telah dilaksanakan. Dua puluh empat subjek dibahagi secara rawak ke dalam tiga kumpulan yang bersaiz lapan. Setelah amaran perlucutan tidur didapati berbeza, semua subjek diminta melaksanakan suatu tugas yang memerlukan ketangkasan secara manual (manual dexterity). Skor dari 0 (lemah) hingga 10 (cermelang) dicatatkan. Jadual 4.2 memaparkan data yang diperolehi.

Jadual 4.2. Data perlucutan tidur.

Kumpulan I (perlucutan 16 hours)	Kumpulan II (perlucutan 20 hours)	Kumpulan III (perlucutan 24 hours)
8.95	7.70	5.99
8.04	5.81	6.79
7.72	6.61	6.43
6.21	6.07	5.85
6.48	8.04	5.78
7.81	5.96	7.60
7.50	7.30	5.78
6.90	7.46	6.00

Andaikan bahawa model pengelasan sehala dengan kesan tetap adalah sesuai. Gunakan data ini untuk menguji $H_0: \tau_1 = \tau_2 = \tau_3$. Apakah kesimpulan anda?

[50 markah]

-ooo000ooo-