
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2011/2012 Academic Session

June 2012

MST 565 – Linear Models
[Model Linear]

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of EIGHT pages of printed materials before you begin the examination.

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

Instructions: Answer **all ten** [10] questions.

[Arahan: Jawab **semua sepuluh** [10] 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].

1. The columns of the following matrix are mutually orthogonal:

$$\mathbf{T} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

- (a) Normalize the columns of $\mathbf{T}$ and denote the resulting matrix by $\mathbf{D}$.
 (b) Show that $\mathbf{D}$ is an orthogonal matrix.

[6 marks]

1. Lajur matriks berikut adalah saling ortogon:

$$\mathbf{T} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix}$$

- (a) Normalkan lajur T dan wakikan matriks yang terhasil sebagai $\mathbf{D}$.
 (b) Tunjukkan $\mathbf{D}$ adalah matriks ortogon.

[6 markah]

2. Consider the system of equations

$$\begin{aligned} 2\mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 &= 1 \\ 4\mathbf{x}_1 + 6\mathbf{x}_2 + \mathbf{x}_3 &= 1 \\ -4\mathbf{x}_2 - 3\mathbf{x}_3 &= 2 \end{aligned}$$

Is this system of equations consistent? Explain your answer.

[6 marks]

2. Pertimbangkan sistem persamaan berikut

$$\begin{aligned} 2\mathbf{x}_1 + \mathbf{x}_2 - \mathbf{x}_3 &= 1 \\ 4\mathbf{x}_1 + 6\mathbf{x}_2 + \mathbf{x}_3 &= 1 \\ -4\mathbf{x}_2 - 3\mathbf{x}_3 &= 2 \end{aligned}$$

Adakah sistem persamaan ini konsisten? Jelaskan jawapan anda.

[6 markah]

3. For the least square estimator $\hat{\beta} = \mathbf{X}'\mathbf{X}^{-1}\mathbf{X}'\mathbf{y}$ it is known that $\Sigma = \text{cov } \hat{\beta} = \sigma^2 \mathbf{X}'\mathbf{X}^{-1}$. We also know that Σ is symmetric and hence may be written as $\mathbf{P}\mathbf{\Lambda}\mathbf{P}'$ where $\mathbf{P}$ is a matrix of eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues.

(a) Explain why Σ^{-1} is also symmetric. What is the rank of Σ^{-1} ?

(b) Explain why $\frac{\hat{\beta} - \beta' \mathbf{X}'\mathbf{X} \hat{\beta}}{\sigma^2} \sim \chi^2_{k+1}$.

[8 marks]

3. Bagi penganggar kuasa dua terkecil $\hat{\beta} = \mathbf{X}'\mathbf{X}^{-1}\mathbf{X}'\mathbf{y}$, diketahui bahawa $\Sigma = \text{cov } \hat{\beta} = \sigma^2 \mathbf{X}'\mathbf{X}^{-1}$. Kita juga tahu bahawa Σ adalah simetri dan dengan itu boleh ditulis seperti $\mathbf{P}\mathbf{\Lambda}\mathbf{P}'$ di mana $\mathbf{P}$ ialah matriks bagi vektor eigen dan $\mathbf{\Lambda}$ adalah matriks pepenjuru bagi nilai eigen.

(a) Terangkan kenapa Σ^{-1} juga adalah simetri. Apakah pangkat Σ^{-1} ?

(b) Terangkan kenapa $\frac{\hat{\beta} - \beta' \mathbf{X}'\mathbf{X} \hat{\beta}}{\sigma^2} \sim \chi^2_{k+1}$.

[8 markah]

4. Let $T_1, \dots, T_n$ and $S_1, \dots, S_n$ be independent random samples from $N(\mu_1, \sigma^2)$ and $N(\mu_2, 2\sigma^2)$ distributions respectively. Let $\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$ and $\bar{S} = \frac{1}{n} \sum_{i=1}^n S_i$.

(a) Find the distribution of $2 \frac{\sum_{i=1}^n (T_i - \bar{T})^2}{\sum_{i=1}^n (S_i - \bar{S})^2}$.

(b) Find the mean and variance of $\frac{1}{\sigma^2} \sum_{i=1}^n (T_i - \bar{T})^2 + \frac{1}{2\sigma^2} \sum_{i=1}^n (S_i - \bar{S})^2$.

[8 marks]

4. Andaikan $T_1, \dots, T_n$ and $S_1, \dots, S_n$ masing-masing adalah sampel rawak tak bersandar daripada taburan $N(\mu_1, \sigma^2)$ dan $N(\mu_2, 2\sigma^2)$. Katakan $\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$ dan $\bar{S} = \frac{1}{n} \sum_{i=1}^n S_i$.

(a) Cari taburan untuk $2 \frac{\sum_{i=1}^n (T_i - \bar{T})^2}{\sum_{i=1}^n (S_i - \bar{S})^2}$.

(b) Cari min dan varians bagi $\frac{1}{\sigma^2} \sum_{i=1}^n (T_i - \bar{T})^2 + \frac{1}{2\sigma^2} \sum_{i=1}^n (S_i - \bar{S})^2$.

[8 markah]

5. Explain the following terms and give examples to illustrate your answers.
- Non-central chi-squared distribution
 - Moment generating function
 - Less and full rank model

[10 marks]

5. Terangkan istilah berikut dan berikan contoh untuk menjelaskan jawapan anda.
- Taburan khi-kuasa dua tak memusat
 - Fungsi penjana momen
 - Model pangkat penuh dan kurang.

[10 markah]

6. Suppose that $\mathbf{X} \sim N_{\mathbf{\mu}, \mathbf{\Sigma}}$ where $\mu_1 = 1, \mu_2 = 0, \mu_3 = -1$, $\sigma_{11} = 8/13, \sigma_{22} = 7/13, \sigma_{33} = 5/13$, $\sigma_{12} = -2/13, \sigma_{13} = 1/13, \sigma_{23} = 3/13$. Let X_i be the i -th element of $\mathbf{X}$ ($i=1,2,3$).
- Find the joint density function of X_1 and X_2 .
 - Find the conditional distribution of X_1 , given X_2 and X_3 .
 - Find the joint density function of $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - 2X_3$.

[12 marks]

6. Katakan $\mathbf{X} \sim N_{\mathbf{\mu}, \mathbf{\Sigma}}$ dengan $\mu_1 = 1, \mu_2 = 0, \mu_3 = -1$, $\sigma_{11} = 8/13, \sigma_{22} = 7/13, \sigma_{33} = 5/13$, $\sigma_{12} = -2/13, \sigma_{13} = 1/13, \sigma_{23} = 3/13$. Andaikan X_i adalah elemen ke- i bagi $\mathbf{X}$ ($i=1,2,3$).
- Cari fungsi ketumpatan tercantum X_1 dan X_2 .
 - Cari taburan bersyarat X_1 , diberi X_2 dan X_3 .
 - Cari fungsi ketumpatan tercantum $Y_1 = X_1 + X_2 + X_3$, $Y_2 = X_1 - 2X_3$.

[12 markah]

7. Let $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\mathbf{X}$ is a $n \times k+1$ full rank matrix of real numbers, $\boldsymbol{\beta}$ is a $k+1 \times 1$ vector of parameters, and $\boldsymbol{\varepsilon}$ is a $n \times 1$ random vector with mean 0 and variance $\sigma^2 \mathbf{I}$. Let $\mathbf{b}$ be the least squares estimator for $\boldsymbol{\beta}$.
- Show that $E \mathbf{y} = \mathbf{X}\boldsymbol{\beta}$
 - Use the rules of expectation to show that $\mathbf{b}$ is unbiased estimator for $\boldsymbol{\beta}$.
 - Use the rules for variance to show that $\text{var } \mathbf{b} = \mathbf{X}\mathbf{X}^{-1} \sigma^2$.

[10 marks]

7. Biar $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ yang mana $\mathbf{X}$ satu matriks pangkat penuh nombor nyata $n \times k+1$, $\boldsymbol{\beta}$ adalah vektor parameter $k+1 \times 1$ dan $\boldsymbol{\varepsilon}$ adalah vektor rawak $n \times 1$ dengan min 0 dan varians $\sigma^2 \mathbf{I}$. Biarkan $\mathbf{b}$ adalah penganggar kuasa dua terkecil $\boldsymbol{\beta}$.
- Tunjukkan $E \mathbf{y} = \mathbf{X}\boldsymbol{\beta}$.
 - Tunjukkan $\mathbf{b}$ adalah penganggar saksama $\boldsymbol{\beta}$ dengan menggunakan petua jangkaan.
 - Tunjukkan $\text{var } \mathbf{b} = \mathbf{X}\mathbf{X}^{-1} \sigma^2$ dengan menggunakan petua untuk varians

[10 markah]

...5/-

8. The following data represents heat evolved in calories per gram of cement (Y) as a function of the amount of two ingredients in the mix: tricalcium aluminate (X_1) and tetracalcium silicate(X_2). The data are shown below.

Observation	X_1	X_2	Y
1	68.5	16.7	174.4
2	45.2	16.8	164.4
3	91.3	18.2	244.2
4	47.8	16.3	154.6
5	46.9	17.3	181.6
6	66.1	18.2	207.5
7	49.5	15.9	152.8
8	52.0	17.2	163.2
9	48.9	16.6	145.4
10	38.4	16.0	137.2
11	87.9	18.3	241.9
12	72.8	17.1	191.1
13	88.4	17.4	232.0
14	42.9	15.8	145.3
15	52.5	17.8	161.1
16	85.7	18.4	209.7
17	41.3	16.5	146.4
18	51.7	16.3	144.0
19	89.6	18.1	232.6
20	82.7	19.1	224.1
21	52.3	16.0	166.5

Based on SAS output given in Figure 1, answer the following questions.

- Test whether calories per gram of cement are related to tricalcium aluminate and tetracalcium silicate.
- Write the estimated regression models and interpret the parameter estimates.
- Interpret the result of R^2 .
- Write a complete SAS program to produce the output in Figure 1.

[14 marks]

The REG Procedure

Dependent Variable: y

Number of Observations Read 21

Number of Observations Used 21

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	24015	12008	99.10	<.0001
Error	18	2180.92741	121.16263		
Corrected Total	20	26196			

Root MSE	11.00739	R-Square	0.9167
Dependent Mean	181.90476	Adj R-Sq	0.9075
Coeff Var	6.05118		

Parameter Estimates				
Variable	DF	Parameter Estimate	Standard Error	t Value
Intercept	1	-68.85707	60.01695	-1.15
x1	1	1.45456	0.21178	6.87
x2	1	9.36550	4.06396	2.30

Figure 1

8. Data berikut mewakili haba yang dikeluarkan dalam kalori bagi setiap gram simen (Y) sebagai suatu fungsi dari dua campuran bancuhan: tricalcium aluminate (X_1) and tetracalcium silicate(X_2). Data tersebut ditunjukkan dibawah.

Cerapan	X_1	X_2	Y
1	68.5	16.7	174.4
2	45.2	16.8	164.4
3	91.3	18.2	244.2
4	47.8	16.3	154.6
5	46.9	17.3	181.6
6	66.1	18.2	207.5
7	49.5	15.9	152.8
8	52.0	17.2	163.2
9	48.9	16.6	145.4
10	38.4	16.0	137.2
11	87.9	18.3	241.9
12	72.8	17.1	191.1
13	88.4	17.4	232.0
14	42.9	15.8	145.3
15	52.5	17.8	161.1
16	85.7	18.4	209.7
17	41.3	16.5	146.4
18	51.7	16.3	144.0
19	89.6	18.1	232.6
20	82.7	19.1	224.1
21	52.3	16.0	166.5

Berdasarkan output SAS pada Rajah 1, jawab soalan-soalan berikut.

- Uji sama ada kalori bagi setiap gram simen adalah berkaitan dengan tricalcium aluminate dan tetracalcium silicate.
- Tuliskan model regresi yang dianggar dan tafsir parameter yang dianggar.
- Tafsir keputusan R^2 .
- Tulis satu program SAS lengkap untuk menghasilkan output dalam Rajah 1.

[14 markah]

The REG Procedure
Dependent Variable: y

Number of Observations Read	21
Number of Observations Used	21

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	24015	12008	99.10	<.0001
Error	18	2180.92741	121.16263		
Corrected Total	20	26196			

Root MSE	11.00739	R-Square	0.9167
Dependent Mean	181.90476	Adj R-Sq	0.9075
Coeff Var	6.05118		

Parameter Estimates				
Variable	DF	Parameter Estimate	Standard Error	t Value
Intercept	1	-68.85707	60.01695	-1.15
x1	1	1.45456	0.21178	6.87
x2	1	9.36550	4.06396	2.30

Rajah 1

9. Refer to the previous question on heat evolved in calories per gram of cement and SAS output in Figure 1, if

$$\mathbf{X}'\mathbf{X}^{-1} = \begin{pmatrix} 29.7289 & 0.0722 & -1.9926 \\ 0.0722 & 0.00037 & -0.0056 \\ -1.9926 & -0.0056 & 0.1363 \end{pmatrix},$$

- Find var $\mathbf{b}$ where $\mathbf{b}$ is the least squares estimator, β .
- Test the effect of tricalcium aluminate and tetracalcium silicate on calories per gram of cement. Interpret your answer.
- Construct and interpret the 99% confidence interval for the parameters.

[12 marks]

9. Merujuk kepada soalan sebelumnya berkaitan haba yang dikeluarkan dalam kalori bagi setiap gram simen dan output SAS pada Rajah 1, jika

$$\mathbf{X}'\mathbf{X}^{-1} = \begin{pmatrix} 29.7289 & 0.0722 & -1.9926 \\ 0.0722 & 0.00037 & -0.0056 \\ -1.9926 & -0.0056 & 0.1363 \end{pmatrix}$$

- Cari var $\mathbf{b}$ di mana $\mathbf{b}$ adalah penganggar kuasa dua terkecil β .
- Uji kesan tricalcium aluminate dan tetracalcium silicate terhadap kalori bagi setiap gram simen. Tafsir keputusan anda.
- Bina dan tafsir selang keyakinan 99% bagi parameter-parameter tersebut.

[12 markah]

...8/-

10. Consider the following analysis of variance model:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n$$

Discuss the estimation procedure for the parameters of the model and also show the procedure for testing the hypothesis $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$.

[14 marks]

10. *Pertimbangkan model analisis varians berikut:*

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n$$

Bincangkan langkah penganggaran parameter-parameter model dan juga tunjukkan langkah bagi pengujian hipotesis $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$.

[14 markah]