
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2011/2012 Academic Session

June 2012

MSG 162 – Applied Statistical Methods
[Kaedah Statistik Gunaan]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **all 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].

1. (a) A study has been initiated to investigate the cleaning ability of three laundry detergents. Three different brands of automatic washing machines are to be used in the experiment. Two stacks of soiled laundry, containing an equal amount of soil, are to be washed with each detergent tested in each washer. At the completion of each wash load, the “whiteness” scores, in percentage, were measured with laser equipment.

Brand of washing machine							
Detergent	A		B		C		Total
A	80.9		81.3		81.8		487.8
	80.9		81.7		81.2		
		161.8		163.0		163.0	
B	80.5		79.4		80.0		480.6
	79.8		80.2		80.7		
		160.3		159.6		160.7	
C	79.7		78.8		80.4		478.8
	80.0		79.2		80.7		
		159.7		158.0		161.1	
Total	481.8		480.6		484.8		1,447.2

$$\sum_{i=1} \sum_{j=1} \sum_{k=1} Y_{ijk}^2 = 116,366.880$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.01$.
(iii) Perform a follow-up analysis using Duncan multiple range test.
- (b) Some varieties of nematodes, roundworms that live in the soil and frequently are so small as to be invisible to the naked eye, feed on the roots of lawn grasses and other plants. This pest which is particularly troublesome in warm climates, can be treated by the application of nematicides. To evaluate the effectiveness of the nematicides, eight plots of land were randomly selected and various rates of application were randomly assigned to the plots. Data collected on the percent kill of nematodes are as follows:

Plot	1	2	3	4	5	6	7	8
Rate of application	3	5	2	2	4	3	5	4
Percent kill	69	99	38	30	82	71	97	86

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at 0.01 significance level .

[25 marks]

1. (a) Suatu kajian telah dilakukan untuk menyelidik kebersihan tiga pencuci baju. Tiga jenama mesin pembasuh automatik digunakan dalam ujikaji ini. Dua longgok baju kotor yang mengandungi amaun kekotoran yang sama, dicuci dengan setiap pencuci yang diuji di dalam setiap mesin pembasuh. Di akhir setiap muatan basuhan, skor “putihan”, dalam peratus, diukur dengan alat laser.

Jenama mesin pembasuh

Pencuci	A	B	C	Jumlah
A	80.9	81.3	81.8	487.8
	80.9	81.7	81.2	
	161.8	163.0	163.0	
B	80.5	79.4	80.0	480.6
	79.8	80.2	80.7	
	160.3	159.6	160.7	
C	79.7	78.8	80.4	478.8
	80.0	79.2	80.7	
	159.7	158.0	161.1	
Jumlah	481.8	480.6	484.8	1,447.2

$$\sum_{i=1} \sum_{j=1} \sum_{k=1} Y_{ijk}^2 = 116,366.880$$

- (i) Nyatakan suatu model serta anggapannya.
- (ii) Lakukan suatu analisis yang sepadan pada $\alpha = 0.01$.
- (iii) Lakukan suatu analisis lanjutan menggunakan ujian julat berganda Duncan.
- (b) Beberapa jenis nematod, cacing bulat yang hidup di dalam tanah yang biasanya adalah sangat kecil untuk dilihat dengan mata kasar, memakan akar-akar rumput laman dan tanaman lain. Makhluk perosak ini khususnya membawa masalah kepada iklim bercuaca panas, boleh dirawat menggunakan nematisid. Untuk menilai keberkesanan nematisid, lapan plot tanah dipilih secara rawak dan pelbagai kadar aplikasi diumpukkan secara rawak kepada plot-plot tersebut. Data yang diambil dalam peratusan nematod yang mati adalah seperti berikut:

Plot	1	2	3	4	5	6	7	8
Kadar aplikasi	3	5	2	2	4	3	5	4
Peratusan mati	69	99	38	30	82	71	97	86

- (i) Nyatakan suatu model serta anggapannya.
- (ii) Lakukan suatu analisis yang sepadan pada aras signifikan 0.01.

[25 markah]

2. (a) A study was designed to evaluate the combined effect of diet and exercise on weight loss. Four weight treatment programs were tested for six different trial periods. A group of overweight male individuals were randomly selected and in each period, four individuals were randomly assigned to the different treatment programs. At the completion of the experiment, weight changes were recorded to the nearest kilogram, for all individuals.

Trial	Diet	Exercise	Diet and exercise	Behavior modification	Total
1	10	8	1	18	37
2	5	4	1	12	22
3	4	2	-1	10	15
4	3	2	-2	7	10
5	0	-1	-2	1	-2
6	-3	-3	-3	0	-9
Total	19	12	-6	48	73

$$\sum_{i=1}^6 \sum_{j=1}^4 Y_{ij}^2 = 895$$

- State a model and its assumptions.
 - Perform an appropriate analysis at $\alpha = 0.05$.
 - Perform a follow-up analysis using Tukey HSD test.
 - Determine the following effects on weight loss:
 - combined effect of diet and exercise
 - diet effect
- (b) A study was conducted to determine whether the quantity of dissolved oxygen, a measure of pollution varied from one location to another. Location A and B were above an industrial plant, one near the shore and the other in midstream, location C was adjacent to the industrial water discharge for the plant and location D was slightly downward in midstream. Water samples were randomly selected at each location on five randomly selected days and the quantity of dissolved oxygen was recorded.

Location	Day					Total
	1	2	3	4	5	
A	5.9	16.1	36.3	6.1	16.0	80.4
B	6.3	16.6	37.4	6.4	16.5	83.2
C	4.8	14.3	35.0	4.7	15.1	73.9
D	6.0	16.2	36.3	5.8	15.9	80.2
Total	23.0	63.2	145.0	23.0	63.5	317.7

$$\sum_{i=1}^4 \sum_{j=1}^5 Y_{ij}^2 = 7,537.35$$

- State a model and its assumptions.
- Perform an appropriate analysis at $\alpha = 0.05$.
- Perform a follow-up analysis.

[30 marks]

2. (a) Suatu kajian telah dirancang untuk menilai kesan gabungan diet dan senaman keatas pengurangan berat badan. Empat program rawatan diuji untuk enam tempoh cubaan berlainan. Sekumpulan individu lelaki yang berlebihan berat dipilih secara rawak dan dalam setiap tempoh, empat individu diumpukkan secara rawak kepada program rawatan berlainan. Diakhir ujikaji tersebut, perubahan berat dicatat kepada kilogram terhampir, bagi kesemua individu.

Cubaan	Diet	Senaman	Diet dan senaman	Perubahan makanan	Jumlah
1	10	8	1	18	37
2	5	4	1	12	22
3	4	2	-1	10	15
4	3	2	-2	7	10
5	0	-1	-2	1	-2
6	-3	-3	-3	0	-9
Jumlah	19	12	-6	48	73

$$\sum_{i=1}^6 \sum_{j=1}^4 Y_{ij}^2 = 895$$

- (i) Nyatakan suatu model serta anggapannya.
(ii) Lakukan suatu analisis yang sepadan pada $\alpha = 0.05$.
(iii) Lakukan suatu analisis lanjutan menggunakan ujian Tukey HSD.
(iv) Tentukan kesan berikut atas pengurangan berat badan:
a. kesan gabungan diet dan senaman
b. kesan diet
- (b) Suatu kajian dilakukan untuk menentukan sama ada kuantiti oksigen terlarut, suatu ukuran pencemaran berubah daripada satu lokasi ke lokasi lain. Lokasi A dan B terletak di atas daripada sebuah kilang, satu bersebelahan pantai dan satu lagi di pertengahan air. Lokasi C adalah bersebelahan dengan sisa air industri kilang tersebut manakala lokasi D terletak di bawah sedikit di pertengahan air. Sampel air dipilih secara rawak pada setiap lokasi pada lima hari yang dipilih secara rawak dan kuantiti oksigen terlarut dicatat.

Jumlah	Hari					Jumlah
	1	2	3	4	5	
A	5.9	16.1	36.3	6.1	16.0	80.4
B	6.3	16.6	37.4	6.4	16.5	83.2
C	4.8	14.3	35.0	4.7	15.1	73.9
D	6.0	16.2	36.3	5.8	15.9	80.2
Jumlah	23.0	63.2	145.0	23.0	63.5	317.7

$$\sum_{i=1}^4 \sum_{j=1}^5 Y_{ij}^2 = 7,537.35$$

- (i) Nyatakan suatu model serta anggapannya.
(ii) Lakukan suatu analisis yang sepadan pada $\alpha = 0.05$.
(iii) Lakukan suatu analisis lanjutan.

[30 markah]

...6/-

3. (a) Physicians depend on laboratory test results when managing medical problems such as diabetes or epilepsy. In a uniformity test for glucose tolerance, three laboratories were randomly selected. Each laboratory was sent five identical blood samples from a person who had drunk 50 milligrams of glucose dissolved in water. The laboratory results were as follows.

Blood sample	Laboratory			Total
	A	B	C	
1	120.1	98.3	103.0	321.4
2	110.7	112.1	108.5	331.3
3	108.9	107.7	101.1	317.7
4	104.2	107.9	110.0	322.1
5	100.4	99.2	105.4	305.0
Total	544.3	525.2	528	1,597.5

$$\sum_{i=1} \sum_{j=1} Y_{ij}^2 = 170,598.77$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.10$.
(iii) Perform a follow-up analysis.
- (b) A forest entomologist has isolated three formulations of insecticides and herbicides to control gypsy moths.

A: a concentration of 0.25 μg of insecticide with no herbicide
B: a concentration of 0.25 μg of insecticide with 50 μg of herbicide
C: a concentration of 0.25 μg of insecticide with 100 μg of herbicide

She wants to determine whether any one of them produces significantly greater mortality than the others when applied to adult gypsy moths. Using standard bioassay techniques, she applies a given formulation to the abdomen of each of 100 moths. This procedure is repeated six times for each formulation, with new solution being prepared each time. Mortality is recorded after 24 hours for each formulation trial

Formulation	Trial						Total
	1	2	3	4	5	6	
A	55	82	83	88	89	92	489
B	90	92	90	91	93	51	507
C	98	98	100	97	57	99	549
Total	243	272	273	276	239	242	1545

$$\sum_{i=1} \sum_{j=1} Y_{ij}^2 = 136,629$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.10$.
(ii) Perform a follow-up analysis.

[25 marks]

...7/-

3. (a) Ahli perubatan bergantung kepada keputusan ujian makmal apabila mengurus masalah perubatan seperti penyakit kencing manis atau epilepsi. Dalam suatu ujian keseragaman bagi ketahanan glukos, tiga makmal dipilih secara rawak. Setiap makmal dihantar lima sampel darah yang sama daripada seseorang yang telah minum 50 miligram glukos yang dilarut dalam air. Keputusan makmal adalah seperti berikut.

Sampel darah	Makmal			Jumlah
	A	B	C	
1	120.1	98.3	103.0	321.4
2	110.7	112.1	108.5	331.3
3	108.9	107.7	101.1	317.7
4	104.2	107.9	110.0	322.1
5	100.4	99.2	105.4	305.0
Jumlah	544.3	525.2	528	1,597.5

$$\sum_{i=1} \sum_{j=1} Y_{ij}^2 = 170,598.77$$

- (i) Nyatakan suatu model serta anggapannya.
(ii) Lakukan suatu analisis yang sepadan pada $\alpha = 0.10$.
(iii) Lakukan suatu analisis lanjutan.
- (b) Seorang pakar entomologi perhutanan telah mengasingkan tiga formulasi racun perosak dan herbisid untuk mengawal kupu-kupu gipsi.

A: suatu kepekatan 0.25 μg racun perosak tanpa herbisid
B: suatu kepekatan 0.25 μg racun perosak dengan 50 μg of herbisid
C: suatu kepekatan 0.25 μg racun perosak dengan 100 μg of herbisid

Dia ingin menentukan sama ada salah satu daripadanya menghasilkan mortaliti lebih besar secara signifikan daripada yang lain apabila digunakan kepada kupu-kupu gipsi dewasa. Menggunakan teknik-teknik bioassay piawai, dia meletakkan suatu formulasi ke perut setiap daripada 100 kupu-kupu. Tatacara ini diulangi enam kali bagi setiap rumusan, dengan larutan baru disediakan setiap kali. Mortaliti dicatat selepas 24 jam bagi setiap cubaan rumusan.

Rumusan	Cubaan						Jumlah
	1	2	3	4	5	6	
A	55	82	83	88	89	92	489
B	90	92	90	91	93	51	507
C	98	98	100	97	57	99	549
Jumlah	243	272	273	276	239	242	1545

$$\sum_{i=1} \sum_{j=1} Y_{ij}^2 = 136,629$$

- (i) Nyatakan suatu model serta anggapannya.
(ii) Lakukan suatu analisis yang sepadan pada $\alpha = 0.10$.
(iii) Lakukan suatu analisis lanjutan.

[25 markah]

..8/-

4. A study was conducted to investigate the effects of temperature and lighting intensity on the speed of productivity among piecework employees. Ten piecework employees were randomly selected from an electric company. Each employee is given 100 transistors to solder to the connecting relays under different temperatures and lighting intensities. The time taken (in hours) is recorded for each subject to complete the task.

Time taken	4.5	4.9	4.7	5.7	4.8	5.3	5.1	5.4	5.6	6.4
Temperature, °F	64	64	66	66	68	68	70	70	72	72
Lighting intensity	60	65	60	65	60	65	60	65	60	65

Refer to the appendix output for question 4

- State the model and its assumptions.
- Perform an appropriate analysis to obtain the estimated model. Use $\alpha = 0.05$
- Obtain an estimation interval and prediction interval of the time taken to complete the tasks with temperature of 69°F and lighting intensity of 63.

[20 marks]

4. Suatu kajian dilakukan untuk mengkaji keberkesanan suhu dan kecerahan lampu ke atas keberhasilan antara pekerja cebis demi cebis. Sepuluh pekerja cebis demi cebis dipilih secara rawak daripada suatu syarikat elektrik. Setiap pekerja diberi 100 transistor untuk dikimpal kepada geganti penyambung di bawah suhu dan kecerahan lampu yang berbeza. Masa yang diambil (dalam jam) dicatat bagi setiap subjek untuk menyiapkan tugas tersebut.

Masa yang diambil	4.5	4.9	4.7	5.7	4.8	5.3	5.1	5.4	5.6	6.4
Suhu, °F	64	64	66	66	68	68	70	70	72	72
Kecerahan lampu	60	65	60	65	60	65	60	65	60	65

Rujuk kepada apendiks output bagi soalan 4

- Nyatakan model serta anggapannya.
- Lakukan suatu analisis yang sesuai untuk mendapatkan suatu model yang dianggarkan. Guna $\alpha = 0.05$
- Dapatkan suatu selang anggaran dan selang ramalan bagi masa yang diambil untuk menyiapkan tugas dengan suhu 69°F dan kecerahan lampu 63.

[20 markah]

APPENDIX: OUTPUT

Question 4

Predictor	Coefficient	Standard error
Constant	-11.27	3.157
Temperature	0.1325	0.03217
Lighting	0.12	0.03639

Analysis of Variance

Source of Variation	df	SS
Regression	2	2.3045
Residual Error	7	0.5795
Total	9	2.8840

$$\mathbf{X}'\mathbf{X}^{-1} = \begin{bmatrix} 120.40 & -0.85 & -1.00 \\ -0.85 & 0.0125 & 0.000 \\ -1.00 & 0.000 & 0.016 \end{bmatrix}$$

APPENDIX: FORMULAS

1. Completely Randomized Design

$$SST = \sum_i \sum_j Y_{ij}^2 - \frac{Y^2}{N}$$

$$SSA = \sum_i \frac{Y_{i.}^2}{n_i} - \frac{Y^2}{N}$$

For any contrast : $L = \sum_i c_i \bar{Y}_{i.}$

$$SSL = \frac{\left(\sum_i c_i \bar{Y}_{i.} \right)^2}{\sum_i \frac{c_i^2}{n_i}}$$

2. Completely Randomized Block Design

$$SST = \sum_i \sum_j Y_{ij}^2 - \frac{Y^2}{N}$$

$$SSA = \sum_i \frac{Y_{i.}^2}{b} - \frac{Y^2}{N}$$

$$SSB = \sum_j \frac{Y_{.j}^2}{a} - \frac{Y^2}{N}$$

3. Latin Square Design

$$SST = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{Y^2}{N}$$

$$SSR = \sum_i \frac{Y_{i..}^2}{a} - \frac{Y^2}{N}$$

$$SSC = \sum_j \frac{Y_{.j.}^2}{a} - \frac{Y^2}{N}$$

$$SSA = \sum_k \frac{Y_{..k}^2}{a} - \frac{Y^2}{N}$$

4. Two-way Factorial Design

$$SST = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{Y^2}{N}$$

$$SSA = \sum_i \frac{Y_{i..}^2}{bn} - \frac{Y^2}{N}$$

$$SSB = \sum_j \frac{Y_{.j.}^2}{an} - \frac{Y^2}{N}$$

$$SSE = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{Y_{ij.}^2}{n}$$

5. Distributions:

Duncan: $r_{\alpha, p, df}$, $p = \text{range}$ $df = \text{degrees of freedom}$

Tukey: $\frac{1}{\sqrt{2}} q_{\alpha, a, df}$, $a = \text{number of treatments}$ $df = \text{degrees of freedom}$

Scheffe': $\sqrt{a-1 F_{\alpha, a-1, df}}$, $a = \text{number of treatments}$ $df = \text{degrees of freedom}$

6. Regression

$$b_1 = \frac{SS_{XY}}{SS_X} , \quad b_0 = \bar{Y} - b_1 \bar{X}$$

$$SSE = SS_Y - \frac{[SS_{XY}]^2}{SS_X}$$

$$SS_{XY} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}$$

$$SS_X = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$SS_Y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

$$\text{Var}(b_1) = \frac{\sigma^2}{SS_X}$$

$$\text{Var}(\hat{Y}_h) = \sigma^2 \left[\frac{1}{n} + \frac{(X_h - \bar{X})^2}{SS_X} \right] , \quad \text{Var}(\hat{Y}_h) = \text{MSE } \mathbf{X}_h' (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h$$

7. Correlation

$$r = \frac{SS_{XY}}{\sqrt{SS_{XX} SS_{YY}}}$$

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

8. Kruskal-Wallis Test

$$T = \frac{12}{N(N+1)} \sum_i \frac{R_i^2}{n_i} - 3 \frac{N+1}{2}$$

9. Friedman Test

$$T = \frac{12}{ba(a+1)} \sum_i R_i^2 - 3b \frac{a+1}{2}$$

10. Cochran Test

$$T = \frac{a(a-1) \sum_i A_i^2 - a(a-1) N^2}{a N - \sum_j B_j^2}$$

11. Spearman Test

$$r_s = 1 - \frac{6 \sum_{i=1}^n [R_{X_i} - R_{Y_i}]^2}{n(n^2-1)} = 1 - \frac{6 \sum_i d_i^2}{n(n^2-1)}$$

APPENDIX: TABLES

Duncan Multiple Range Table

Spearman Table

Tukey HSD Table