
UNIVERSITI SAINS MALAYSIA

Second Semester Examination
2009/2010 Academic Session

April/May 2010

MSG 162 – Applied Statistical Methods
[Kaedah Statistik Gunaan]

Duration : 3 hours
[Masa : 3 jam]

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

[Sila pastikan bahawa kertas peperiksaan ini mengandungi TUJUH 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 was designed to evaluate the effectiveness of various diets combined with various levels of jogging per day on weight loss (in kilogram). Twelve men, each approximately 20 kilograms overweight, are randomly assigned to two diets and three levels of jogging. Each man consumes the same number of calories per day, but the diets differ in their proportion of fat and carbohydrates. After participating in the program for six months, the reduction in weight was recorded.

Jogging (in kilometers)	Diet		Total
	High carbohydrate	High fat	
0	15.5	8.5	45.0
	13.5	7.5	
	29.0	16.0	
1	21.0	13.0	63.0
	18.0	11.0	
	39.0	24.0	
2	24.5	22.0	101.0
	27.5	27.0	
	52.0	49.0	
Total	120.0	89.0	209.0

$$\sum_i \sum_j \sum_k Y_{ijk}^2 = 4,175.5$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.025$.
(iii) Perform a follow-up analysis.
- (b) Three different teaching techniques are to be investigated for teaching elementary calculus.

Technique 1: Computer-assisted instruction (CAI) in conjunction with lectures
Technique 2: Computer-assisted instruction only
Technique 3: Lectures only

Five students were randomly assigned to each of the teaching techniques. A skills test was administered at the end of the learning period and the scores recorded.

Technique	Student					Total
	1	2	3	4	5	
1	90	88	94	67	96	435
2	90	59	88	87	86	410
3	78	70	65	74	43	330
Total	258	217	247	228	225	1,175

$$\sum_i \sum_j Y_{ij}^2 = 95,209$$

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

[100 marks]

1. (a) Suatu kajian dirancang untuk menilai keberkesanan pelbagai diet digabung dengan pelbagai aras larian dalam sehari keatas penurunan berat badan (dalam kilogram). Dua belas orang lelaki, setiapnya mempunyai berat badan berlebihan sebanyak 20 kilogram, diumpukkan secara rawak kepada dua jenis diet dan tiga aras larian. Setiap orang lelaki memakan makanan yang mempunyai bilangan kalori yang sama dalam sehari, tetapi diet tersebut berbeza dalam kadar lemak dan karbohidrat. Selepas mengikuti program tersebut selama enam bulan, pengurangan dalam berat badan dicatat.

Larian (dalam kilometer)	Diet		Jumlah
	Berkarbohidrat tinggi	Berlemak tinggi	
0	15.5	8.5	45.0
	13.5	7.5	
	29.0	16.0	
1	21.0	13.0	63.0
	18.0	11.0	
	39.0	24.0	
2	24.5	22.0	101.0
	27.5	27.0	
	52.0	49.0	
Jumlah	120.0	89.0	209.0

$$\sum_1 \sum_j \sum_k Y_{ijk}^2 = 4,175.5$$

- (i) Nyatakan suatu model dan anggapannya.
(ii) Lakukan suatu analisis yang sesuai pada $\alpha = 0.025$.
(iii) Lakukan suatu tatacara lanjutan.
- (b) Tiga teknik pengajaran yang berbeza dikaji bagi pengajaran kalkulus permulaan.

Teknik 1: Pengajaran berbantuan komputer (CAI) bersamaan dengan syarahan

Teknik 2: Pengajaran berbantuan komputer sahaja

Teknik 3: Syarahan sahaja

Lima orang pelajar diumpukkan secara rawak kepada setiap daripada teknik pengajaran. Suatu ujian kemahiran diberikan pada akhir tempoh pembelajaran dan skornya dicatat.

Teknik	Pelajar					Jumlah
	1	2	3	4	5	
1	90	88	94	67	96	435
2	90	59	88	87	86	410
3	78	70	65	74	43	330
Jumlah	258	217	247	228	225	1,175

$$\sum \sum Y_{ij}^2 = 95,209$$

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

[100 markah]

2. (a) An anesthesiologist studied the effects of acupuncture and codeine on postoperative dental pain in male subjects. The four treatments were as follows:

Placebo – a sugar capsule and two inactive acupuncture points

Codeine only – a codeine capsule and two inactive acupuncture points

Acupuncture only – a sugar capsule and two active acupuncture points

Codeine and acupuncture - a codeine capsule and two active acupuncture points

Subjects for the study were selected according to an initial evaluation of their level of pain tolerance. The subjects in different levels of pain tolerance (level 1: lowest tolerance to level 8: highest tolerance) were then randomly assigned to the treatments. Pain relief scores (the higher the pain relief score, the more effective the treatment) were obtained for all subjects two hours after dental treatment.

Treatment	Level of pain tolerance								Total
	1	2	3	4	5	6	7	8	
Placebo	0.0	0.3	0.2	0.5	0.4	1.1	1.0	1.2	4.7
Codeine only	0.5	0.6	0.8	0.7	1.0	1.9	1.8	1.7	9.0
Acupuncture only	0.6	0.7	0.9	1.0	0.9	1.5	1.7	1.6	8.9
Codeine and acupuncture	1.2	1.3	1.8	2.2	2.5	2.3	2.1	2.4	15.8
Total	2.3	2.9	3.7	4.4	4.8	6.8	6.6	6.9	38.4

$$\sum \sum Y_{ij}^2 = 60.76$$

- (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.
 - (iv) Determine the following effects:
 - a. codeine effect
 - b. acupuncture effect
- (b) In order to study the efficiency of microwave cooking in sterilizing meat, a food scientist takes a random sample of nine beef sausages, and by means of hypodermic needle she inoculates each with the same volume of a nutrient broth containing a heavy suspension of salmonella. She then cooks each sausage for a different length of time in a microwave oven set for a constant temperature. The contents of the sausages are then mixed with an agar solution and poured into petri dishes. The dishes are placed in an incubator. After 18 hours of incubation, the number of salmonella colonies per dish are counted.

Time cooked in microwave oven (in minutes)	Number of salmonella colonies
0	740
2	410
4	210
6	100
8	45
10	25
12	10
14	6
16	4

- (i) State a model and its assumptions.
- (ii) Perform an appropriate analysis at $\alpha = 0.025$.

[100 marks]

2. (a) Seorang pakar bius mengkaji keberkesanan akupunktur dan kodine atas kesakitan gigi selepas pembedahan bagi subjek lelaki. Keempat-empat rawatan adalah seperti berikut:

Placebo – suatu kapsul gula dan dua titik akupunktur yang tak aktif

Kodein sahaja – suatu kapsul kodeine dan dua titik akupunktur yang tak aktif

Akupunktur sahaja – suatu kapsul gula dan dua titik akupunktur yang aktif

Kodein dan akupunktur - suatu kapsul kodein dan dua titik akupunktur yang aktif

Subjek bagi kajian tersebut dipilih mengikut suatu penilaian awal aras ketahanan sakit mereka. Subjek dalam aras ketahanan sakit yang berbeza (aras 1: ketahanan yang paling rendah kepada aras 8: ketahanan yang paling tinggi) diumpukkan secara rawak kepada rawatan-rawatan tersebut. Skor kelegaan sakit (paling tinggi skor kelegaan sakit, lebih berkesan rawatan tersebut) didapati bagi kesemua subjek dua jam selepas rawatan gigi.

Rawatan	Aras ketahanan sakit								Jumlah
	1	2	3	4	5	6	7	8	
Placebo	0.0	0.3	0.2	0.5	0.4	1.1	1.0	1.2	4.7
Kodein sahaja	0.5	0.6	0.8	0.7	1.0	1.9	1.8	1.7	9.0
Akupunktur sahaja	0.6	0.7	0.9	1.0	0.9	1.5	1.7	1.6	8.9
Kodein dan akupunktur	1.2	1.3	1.8	2.2	2.5	2.3	2.1	2.4	15.8
Jumlah	2.3	2.9	3.7	4.4	4.8	6.8	6.6	6.9	38.4

$$\sum \sum Y_{ij}^2 = 60.76$$

- Nyatakan suatu model dan anggapannya.
 - Lakukan suatu analisis yang sesuai pada $\alpha = 0.01$.
 - Lakukan suatu tatacara lanjutan menggunakan ujian julat berganda Duncan.
 - Tentukan keberkesanan berikut:
 - kesan kodein
 - kesan akupunktur
- (b) Dalam mengkaji kecekapan masakan ketuhar dalam mensteril daging, seorang saintis pemakanan mengambil suatu sampel rawak sembilan sosej daging. Dengan menggunakan suatu jarum hipodermik dia menyuntik setiap satu sosej dengan suatu sup bernutrien yang berisipadu sama. Sup tersebut mengandungi suatu bahan yang sarat dengan salmonella. Dia kemudiannya memasak setiap sosej untuk suatu jangka masa di dalam oven ketuhar bagi suatu suhu yang tetap. Kandungan sosej tersebut dicampur dengan cecair agar-agar dan dituang ke dalam bekas petri. Bekas petri tersebut diletakkan di dalam inkubator. Selepas 18 jam di dalam inkubator, bilangan koloni salmonella di dalam setiap bekas dikira.

Jangka masa memasak di dalam ketuhar (dalam minit)	Bilangan koloni salmonella
0	740
2	410
4	210
6	100
8	45
10	25
12	10
14	6
16	4

- Nyatakan suatu model dan anggapannya.
- Lakukan suatu analisis yang sesuai pada $\alpha = 0.025$.

[100 markah]

...6/-

3. (a) An endocrinologist conducts an experiment to study the effect of diet on the activity of an enzyme that consumes fat in the body. These diets are regularly given to groups of rats of an identical breed.

Diet I: Protein-rich food; free feeding

Diet II: Protein-rich food; controlled feeding

Diet III: Carbohydrate-rich food; controlled feeding

Measurements of fatty deposits found on the kidneys and intestines of the rats given each diet were determined by six technicians.

Technician	Diet			Total
	1	2	3	
1	120	96	98	314
2	93	42	92	227
3	95	84	81	260
4	96	86	93	275
5	105	69	75	249
6	96	74	51	221
Total	605	451	490	1,546

$$\sum \sum Y_{ij}^2 = 138,704$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.025$.
(iii) Perform a follow-up analysis.
- (b) A researcher studied the sugar content in carbonated drinks by selecting at random four brands sold in a metropolitan area. The researcher selected eight 360-milliliter cans from retail outlets in the area and measured the sugar content (in milligrams) of each can.

Can	Brand				Total
	A	B	C	D	
1	24.4	10.2	21.3	25.3	81.2
2	22.6	12.1	20.2	25.2	80.1
3	23.8	10.3	20.7	33.6	88.4
4	22.0	10.2	20.8	32.9	85.9
5	24.5	9.9	20.1	27.7	82.2
6	22.3	11.2	18.8	28.5	80.8
7	25.0	12.0	21.1	29.2	87.3
8	24.5	9.5	20.3	29.3	83.6
Total	189.1	85.4	163.3	231.7	669.5

$$\sum \sum Y_{ij}^2 = 15,512.4$$

- (i) State a model and its assumptions.
(ii) Perform an appropriate analysis at $\alpha = 0.01$.
(iii) Perform a follow-up analysis.

[100 marks]

3. (a) Seorang ahli endokrinologi melakukan suatu ujikaji untuk mengkaji keberkesanan diet atas aktiviti suatu enzim yang menyerap lemak ke dalam badan. Diet tersebut diberikan secara berkala kepada beberapa kumpulan tikus yang sama jenis.

Diet I: Makanan berprotin tinggi; pemakanan bebas

Diet II: Makanan berprotin tinggi; pemakanan dikawal

Diet III: Makanan berkabihidrat tinggi; pemakanan dikawal

Ukuran mendapan lemak yang didapati di atas ginjal dan usus tikus-tikus yang diberi setiap diet ditentukan oleh enam orang juruteknik.

Juruteknik	Diet			Jumlah
	1	2	3	
1	120	96	98	314
2	93	42	92	227
3	95	84	81	260
4	96	86	93	275
5	105	69	75	249
6	96	74	51	221
Jumlah	605	451	490	1,546

$$\sum \sum Y_{ij}^2 = 138,704$$

- (i) Nyatakan suatu model dan anggapannya.
(ii) Lakukan suatu analisis yang sesuai pada $\alpha = 0.025$.
(iii) Lakukan suatu tatacara lanjutan.
- (b) Seorang penyelidik mengkaji kandungan gula di dalam minuman berkarbonat dengan memilih secara rawak empat jenama yang dijual di suatu kawasan metropolitan. Penyelidik tersebut memilih lapan tin 360-mililiter daripada pusat jualan kawasan tersebut dan mengukur kandungan gula (dalam miligram) bagi setiap tin.

Tin	Jenama				Jumlah
	A	B	C	D	
1	24.4	10.2	21.3	25.3	81.2
2	22.6	12.1	20.2	25.2	80.1
3	23.8	10.3	20.7	33.6	88.4
4	22.0	10.2	20.8	32.9	85.9
5	24.5	9.9	20.1	27.7	82.2
6	22.3	11.2	18.8	28.5	80.8
7	25.0	12.0	21.1	29.2	87.3
8	24.5	9.5	20.3	29.3	83.6
Jumlah	189.1	85.4	163.3	231.7	669.5

$$\sum \sum Y_{ij}^2 = 15,512.4$$

- (i) Nyatakan suatu model dan anggapannya.
(ii) Lakukan suatu analisis yang sesuai pada $\alpha = 0.01$.
(iii) Lakukan suatu tatacara lanjutan.

[100 markah]

4. Mercury contamination of edible fish poses a health threat. A study was conducted to examine factors that might influence the amount of mercury contamination. Samples of fish were taken from 15 lakes and the mercury concentration in their muscle tissue was measured. Since fish absorb mercury over time (older fish tend to have higher concentrations), standardized mercury levels (in parts per million, ppm) for three year old fish were calculated for each lake. Water samples were also collected from each lake and each sample was analyzed for alkalinity and calcium levels of the water.

Standard mercury (in ppm)	Alkalinity (mg/l)	Calcium (mg/l)
250	71.8	40.5
550	5.0	23.6
160	71.4	55.2
510	6.5	13.8
150	83.7	66.5
190	97.5	45.5
450	7.5	32.0
420	10.5	36.3
530	8.5	12.5
310	55.4	45.9
470	6.3	33.3
250	67.0	48.6
410	28.8	30.2
160	66.5	56.0
430	15.5	35.2

Refer to the appendix output for question 4.

- Perform an appropriate analysis to obtain the model and its assumptions. Use $\alpha = 0.02$.
- Obtain an estimation interval and prediction interval of a lake with alkalinity level of 45 milligram per liter and calcium level of 45 milligram per liter .

[100 marks]

4. Pencemaran raksa dalam ikan yang boleh dimakan memberi suatu ancaman kesihatan. Suatu kajian dilakukan untuk memeriksa faktor yang mungkin mempengaruhi amaun pencemaran raksa. Beberapa sampel ikan diambil daripada 15 buah tasik dan kepekatan raksa di dalam tisu otot diukur. Oleh kerana ikan menyerap raksa mengikut masa (ikan yang lebih tua berkemungkinan mempunyai kepekatan yang lebih tinggi), aras raksa yang diipawai (dalam bahagian sejuta) untuk ikan yang berumur tiga tahun dikira bagi setiap tasik. Sampel air juga diambil daripada setiap tasik dan setiap sampel dianalisis untuk aras alkalin dan kalsium dalam air.

Raksa piawai (dalam ppm)	Aras alkalin (mg/l)	Aras kalsium (mg/l)
250	71.8	40.5
550	5.0	23.6
160	71.4	55.2
510	6.5	13.8
150	83.7	66.5
190	97.5	45.5
450	7.5	32.0
420	10.5	36.3
530	8.5	12.5
310	55.4	45.9
470	6.3	33.3
250	67.0	48.6
410	28.8	30.2
160	66.5	56.0
430	15.5	35.2

Rujuk kepada apendiks output bagi soalan 4.

- Lakukan suatu analisis yang sesuai untuk mendapatkan suatu model dan anggapanannya. Guna $\alpha = 0.02$.
- Dapatkan suatu selang anggaran dan selang ramalan bagi suatu tasik yang mempunyai aras alkalin 45 miligram seliter dan aras kalsium 45 miligram seliter.

[100 markah]

APPENDIX: OUTPUT

Question 4

Predictor	Coefficient	Standard error
Constant	612.08	23.72
Alkalinity	-2.5626	0.3915
Calcium	-4.1710	0.8702

Analysis of Variance

Source of Variation	df	SS
Regression	2	286725
Error	12	9968
Total	14	296693

$$\mathbf{X}'\mathbf{X}^{-1} = \begin{bmatrix} 0.677411 & 0.005385 & -0.021565 \\ 0.005385 & 0.0001845 & -0.0003336 \\ -0.021565 & -0.0003336 & 0.0009116 \end{bmatrix}$$

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

Tukey HSD: $r_{\alpha, a, df}$, a = number of treatments, df = degrees of freedom

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

6. Correlation

$$r = \frac{SS_{XY}}{\sqrt{SS_X SS_Y}}$$

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

7. Regression

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

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

$$SS_{XY} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}$$

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

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

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

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

8. Kruskal-Wallis Test

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

9. Friedman Test

$$T = \frac{12}{ba(b+1)} \sum_{i=1}^k R_i^2 - 3b(b+1)$$

10. Cochran Test

$$T = \frac{a(a-1) \sum_{i=1}^k A_i^2 - (a-1) N^2}{a N - \sum_{j=1}^J B_j^2}$$

11. Spearman Test

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

APPENDIX: TABLES**Duncan Multiple Range Table**

Tukey HSD Table

Spearman Table