

---

UNIVERSITI SAINS MALAYSIA

Peperiksaan Kursus Semasa Cuti Panjang  
Sidang Akademik 2006/2007

Jun 2007

**EUM 222 – KEBARANGKALIAN DAN STATISTIK GUNAAN**

Masa: 3 jam

---

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

Kertas soalan ini mengandungi **ENAM** soalan.

Jawab **EMPAT** soalan.

Gunakan  $\alpha = 0.05$  jika nilai  $\alpha$  tidak diberikan dalam soalan.

Mulakan jawapan anda untuk setiap soalan pada muka surat yang baru.

Agihan markah bagi setiap soalan diberikan di sudut sebelah kanan soalan berkenaan.

Jawab semua soalan dalam bahasa Malaysia atau bahasa Inggeris atau kombinasi kedua-duanya.

1. (a) Sebutkan takrifan bagi Ralat Jenis I dan Ralat Jenis II.

*State the definition of Type I and Type II error.*

(5 markah)

- (b) Dua komponen minikomputer boleh diperihalkan oleh fungsi ketumpatan kebarangkalian bercantum seperti yang berikut:

*Two components of a minicomputer can be described by the bivariate probability density function as follows:*

$$f(x, y) = \begin{cases} xe^{-x(1+y)} & x \geq 0, y \geq 0 \\ 0 & \text{selainnya} \end{cases}$$

*X dan Y ialah masa hayat bagi komponen minikomputer itu.*

*X and Y are the lifetimes for the components of the minicomputer.*

- (i) Apakah kebarangkalian bahawa masa hayat *X* bagi komponen pertama melebihi 3?

*What is the probability that the lifetime X for the first component exceeds 3?*

- (ii) Dapatkan fungsi ketumpatan sut bagi *X* dan *Y*. Adakah *X* dan *Y* tak bersandar?

*Determine the marginal probability density function for X and Y. Are X and Y independent?*

- (iii) Apakah kebarangkalian bahawa masa hayat bagi salah satu komponen melebihi 3?

*What is the probability that the lifetimes for at least one component exceeds 3?*

(15 markah)

...3/-

- (c) Andaikan seorang doktor sedang menguji hipotesis bahawa sekurang-kurangnya 30% daripada penduduk mengidap penyakit lelah. Terangkan dengan jelas bagaimana doktor tersebut boleh melakukan ralat jenis pertama dan ralat jenis kedua.

*Assume that a doctor is testing the hypothesis that at least 30% of the population is infected by asthma. Describe clearly how the doctor can make a Type I and Type II error.*

(5 markah)

2. (a) Andaikan satu sampel bersaiz  $n = 9$  dikutip daripada taburan normal dengan  $\sigma = 4$ . Rantau penolakan bagi ujian ini ialah  $\xi = \{\bar{x} : \bar{x} < 28\}$  untuk menguji  $H_0 : \mu = 30$  melawan  $H_1 : \mu < 30$ .

*Assume that a sample of size  $n = 9$  are obtained from the normal distribution with  $\sigma = 4$ . The rejection region is*

*$\xi = \{\bar{x} : \bar{x} < 28\}$  for testing  $H_0 : \mu = 30$  versus  $H_1 : \mu < 30$ .*

- (i) Dapatkan nilai paras keertian bagi ujian ini.  
*Determine the value for the significance level of this test.*
- (ii) Apakah nilai kebarangkalian ralat jenis II jika  $\mu = 25$  ?  
*What is the probability of Type II error if  $\mu = 25$  ?*

(10 markah)

- (b) Seorang jurutera berpendapat bahawa rintangan wayar 1 lebih kuat daripada rintangan wayar 2. Data yang berikut (dalam ohm) telah diperoleh daripada ujian yang telah dilakukan terhadap sampel setiap wayar.

*An engineer believes that the resistance for wire 1 is greater than that of wire 2. The following data (in ohms) was obtained from tests which were conducted on the samples of each wire.*

Wayar Wire	Ohm					
1	0.140	0.138	0.143	0.142	0.144	0.139
2	0.135	0.140	0.142	0.136	0.137	

- (i) Adakah setiap sampel bertabur secara taburan normal?  
*Are the samples for each wire normally distributed?*
- (ii) Jika varians boleh dianggap sebagai sama, adakah pendapat jurutera itu benar?  
*Can you say that the opinion of the engineer is correct? Assume that the variances for both wires are the same.*

(15 markah)

3. (a) Jadual yang di bawah ialah kekerapan bagi suatu pembolehubah rawak.

*The table below is the frequency for a random variable.*

Nilai <i>Value</i>	0	1	2	3	4
Kekerapan <i>Frequency</i>	4	21	10	13	2

Adakah taburan binomial dengan  $n=6$  dan  $p=0.25$  sesuai untuk data ini?

*Can the binomial distribution with  $n=6$  and  $p=0.25$  be fitted to the data?*

(10 markah)

- (b) Suatu ujikaji telah dijalankan untuk membandingkan rintangan elektrik (disukat dalam ohm) bagi perintang yang dibekalkan oleh dua buah pengeluar, A dan B. Enam buah perintang dari setiap pengeluar dipilih secara rawak dan rintangan elektrik disukat. Data yang berikut diperoleh:

*An experiment was conducted to compare the electrical resistance (measured in ohms) for resistors which are produced by two producers, A and B. Six resistors were selected from each producer and the electrical resistance calculated. The following data are obtained:*

A	0.140	0.138	0.143	0.142	0.144	0.137
B	0.135	0.140	0.142	0.136	0.138	0.140

Andaikan sukatan ini adalah daripada populasi normal.

*Assume that the measurements are normally distributed.*

- (i) Adakah varians populasi bagi rintangan elektrik sama untuk kedua-dua pengeluar?

*Are the population variance the same for the electrical resistance produced by the two different producers?*

- (ii) Dapatkan selang keyakinan 95% bagi perbezaan antara dua min perintang ini.

*Determine the 95% confidence interval for the difference in means of the two resistors.*

(15 markah)

4. Seorang jurutera pembuatan telah mengenalpasti empat kaedah yang boleh dilakukan untuk menyiapkan satu tugas. Tanggungjawab jurutera berkenaan ialah untuk memastikan bahawa masa menyiapkan tugas itu dapat dikurangkan. Enam belas orang pekerja telah dipilih dan diumpukkan kepada empat kaedah secara rawak. Masa (dalam minit) melaksanakan tugas diberikan dalam jadual di bawah.

*A manufacturing engineer has established four methods which can be used for completing a task. The responsibility of the engineer is to determine that the task can be completed in the shortest time. Sixteen workers were chosen at random and assigned to the four methods. The time (in minutes) to complete the task is given in the following table:*

Kaedah Method			
1	2	3	4
27	25	29	31
32	28	31	30
29	24	29	33
30	26	30	32

- (a) Adakah terdapat bukti bahawa wujud perbezaan bererti dalam purata masa menyiapkan tugas menggunakan empat kaedah berlainan itu?

*Are there any evidence to suggest that there are significant differences in the mean time to complete the task using the four different methods?*

(10 markah)

- (b) Kaedah manakah yang paling baik?

*Which method is the best?*

(15 markah)

5. Data yang berikut memberikan hubungan antara penggunaan minyak petrol (dalam batu per gelen) bagi kereta berdasarkan beratnya (dalam tan) dan suhu semasa memandu (dalam °F).

*The following data shows the relationship between consumption of petrol (in miles per gallon) for cars based on its weight (in tons) and temperature of the car during driving (in °F).*

Nombor kereta <i>Car number</i>	Batu per gelen <i>Miles per gallon</i>	Berat <i>Weight</i>	Suhu <i>Temperature</i>
1	17.9	1.35	90
2	16.5	1.90	30
3	16.4	1.70	80
4	16.8	1.80	40
5	18.8	1.30	35
6	15.5	2.05	45
7	17.5	1.60	50
8	16.4	1.80	60
9	15.9	1.85	65
10	18.3	1.40	30

- (a) Suaikan satu model regresi linear berbentuk  
*Fit a linear regression model of the form*

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2.$$

- (b) Ujikan kecukupan model.  
*Test for significance of the regression model.*



- (c) Kirakan nilai  $R^2$ .  
*Determine  $R^2$ .*
  
- (d) Cari selang keyakinan 95% bagi  $\beta_1$ .  
*Find the 95% confidence interval for  $\beta_1$ .*

(25 markah)

CATATAN:  $(X^T X)^{-1} = \begin{pmatrix} 6.070769 & -3.02588 & -0.0171888 \\ -3.02588 & 1.738599 & 0.002166306 \\ -0.0171888 & 0.002166306 & 0.0002582903 \end{pmatrix}$

6. (a) Suatu ujikaji telah dilakukan untuk membandingkan jumlah ingatan yang diperlukan untuk menganalisis suatu set data menggunakan dua jenis perisian statistik yang berlainan. Data yang berikut telah diperoleh:

*An experiment was conducted to compare the amount of memory required for the analysis of a dataset using two different statistical packages. The following data are obtained:*

Program	Perisian X Software X	Perisian Y Software Y
1	512 kb	500 kb
2	650	600
3	890	890
4	410	400
5	1050	1025
6	1500	1400
7	600	625
8	750	710

- (i) Ujikan hipotesis bahawa jumlah ingatan yang diperlukan untuk menganalisis set data adalah berbeza apabila menggunakan perisian statistik yang berlainan. Gunakan ujian tak berparameter yang sesuai.

*Test the hypothesis that the amount of memory required for analyzing a dataset are different when using different statistical packages. Use a suitable nonparametric test.*

(10 markah)

- (ii) Nyatakan andaian yang diperlukan untuk menggunakan ujian tak berparameter dalam bahagian (i) di atas.

*State the required assumptions for using the nonparametric test in (i) above.*

(5 markah)

- (b) Suatu kajian telah dilakukan untuk menguji ketakbersandaran antara mutu udara dengan suhu udara. Data yang berikut telah diperoleh daripada rekod bagi 200 hari yang telah dipilih secara rawak dalam jangka masa beberapa tahun yang lampau. Adakah data ini menunjukkan bahawa terdapat hubungan antara mutu udara dengan suhu udara?

*A study was conducted to test whether air quality and air temperature are independent. The following data was obtained for 200 days which are selected at random from previous records of a few years. Does this data show an association between air quality and air temperature?*

Suhu Temperature	Mutu Udara Air quality		
	Buruk Bad	Sederhana Moderate	Baik Good
Rendah Low	1	3	24
Sederhana Moderate	12	28	76
Tinggi High	12	14	30

(10 markah)