UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2007/2008

October/November 2007

EUM 222 – PROBABILITY AND APPLIED STATISTICS
[Kebarangkalian Dan Statistik Gunaan]

Duration: 3 hours
[Masa: 3 jam]

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

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

This paper contains SIX questions. THREE questions in Section A and THREE questions in Section B.

[Kertas soalan ini mengandungi ENAM soalan. TIGA soalan dalam Bahagian A dan TIGA *soalan dalam Bahagian B].*

Instructions: Answer **FIVE (5)** questions. Answer **TWO (2)** questions from Section A, **TWO (2)** questions from Section B and **ONE (1)** question from any Section.

*[Jawab **LIMA** soalan: Jawab **DUA** soalan dari Bahagian A, **DUA** soalan dari Bahagian B dan **SATU** soalan dari mana-mana bahagian].*

Use $\alpha = 0.05$, if α value is not given in the question.

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

Answer to any question must start on a new page.

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

You may answer a question either in bahasa Malaysia or in English.

[Anda dibenarkan menjawab soalan sama ada dalam bahasa Malaysia atau bahasa Inggeris].

BAHAGIAN A
SECTION A

1. (a) Andaikan bahawa X dan Y adalah pembolehubah rawak selanjar dengan fungsi ketumpatan kebarangkalian bivariat $f(x,y)$. Jika X dan Y adalah tak bersandar, tunjukkan bahawa $\rho_{XY} = 0$.

Assume that X and Y are continuous random variables with bivariate probability density functions $f(x,y)$. Show that if X and Y are independent then $\rho_{XY} = 0$.

(5 marks)

(b) Min masa pengeringan bagi sejenis cat ialah 12 minit. Satu bahan campuran baru telah ditambahkan dalam cat untuk melihat sama ada campuran baru ini akan mengurangkan masa pengeringan. Seratus sampel telah dicat dan min masa pengeringannya ialah 11.0 minit. Andaikan bahawa varians populasi masa pengeringan ialah 2 minit. Hipothesis nol bagi ujian ini ialah $H_0 : \mu = 12$ melawan hipotesis alternatif $H_1 : \mu < 12$. Hipotesis nol akan ditolak jika $\overline{X} \leq 11.7$.

The mean drying time of a certain paint is 12 minutes. A new additive is added to see whether it reduces drying time. One hundred samples were painted and the mean drying time is 11.0 minutes. Assume that the population variance is 2 minutes. The null hypothesis $H_0 : \mu = 12$ was tested against the alternative hypothesis $H_1 : \mu < 12$. The null hypothesis will be rejected if $\overline{X} \leq 11.7$.

(i) Dapatkan nilai bagi α.

Find the value of α.

...3/-

(ii) Jika min sebenar masa pegeringan bagi cat baru ialah 11.5 minit, dapatkan nilai bagi β.

If the true mean drying time of the new paint is 11.5 minutes, find the value for β.

(iii) Apakah kuasa bagi ujian ini?

What is the power of this test?

(10 marks)

(c) Andaikan X ialah jumlah pengecutan (dalam %) bagi suatu gentian yang dipilih secara rawak apabila dipanaskan pada 100°C. Y mewakili pengecutan tambahan (dalam %) apabila gentian dipanaskan pada 110°C. Fungsi ketumpatan kebarangkalian bivariate bagi X dan Y diberikan oleh

Let X denote the amount of shrinkage (in %) of a randomly chosen fiber when heated at 100°C. Let Y represent the additional shrinkage (in %) when the fiber is heated to 110°C. The joint probability density function of X and Y is given by

$$f(x,y)=\begin{cases}\dfrac{48xy^2}{49} & 3<x<4 \text{ and } 0.5<y<1\\ 0 & \text{otherwise}\end{cases}$$

(i) Cari $P(X<3.25 \quad \text{dan} \quad Y>0.8)$

Find $P(X<3.25 \quad \text{and} \quad Y>0.8)$

(ii) Adakah X dan Y tak bersandar?

Are X and Y independent?

(iii) Carilah kekuatan hubungan antara X dan Y.

Find the strength of the relationship between X and Y.

(10 marks)

2. Dua jenis tayar daripada dua jenama terkenal yang berlainan diuji untuk menentukan sama ada terdapat sebarang perbezaan dalam jangkaan masahayat bunga tayar. Data (dalam ribuan kilometer) diberikan di bawah:

Two types of radial tyres from two different major brands were tested to determine whether there were any differences in the expected tread life. The data (in thousands of kilometers) are given below:

Jenama A Brand A					*Jenama B* Brand B				
50	54	52	47	61	57	61	47	52	53
56	51	51	48	56	57	56	53	67	58
53	43	58	52	48	62	56	56	62	57

(i) Adakah data bagi Jenama A dan bagi Jenama B bertabur secara taburan normal?

Are the data for Brand A and for Brand B normally distributed?

(8 marks)

(ii) Lakukan pengujian hipotesis untuk menentukan bahawa varians sebenar bagi masahayat bunga tayar adalah sama untuk kedua-dua jenama tayar.

Test the hypothesis that the true variances of tread life are equal for the two brands of tyres.

(8 marks)

...5/-

(iii) Binakan selang keyakinan 95% bagi perbezaan sebenar min masahayat bunya tayar.

Construct a 95% confidence interval for the true difference in mean tread life.

(9 marks)

3. (a) Di antara faktor utama yang menyebabkan komputer tak berfungsi adalah kegagalan perisian, kegagalan perkakasan, ralat operator dan bebanan lebihan sistem. Seorang penyelidik mendakwa bahawa 10% daripada penyebab komputer tak berfungsi adalah berpunca daripada kegagalan perisian, 5% disebabkan oleh kegagalan perkakasan, 25% oleh kegagalan operator, 40% oleh bebanan lebihan sistem dan selebihnya disebabkan oleh beberapa punca lain. Satu kajian telah dijalankan dan didapati bahawa 13 kegagalan disebabkan oleh kegagalan perisian, 10 disebabkan kegagalan perkakasan, 42 disebabkan oleh ralat operator dan 65 disebabkan oleh bebanan lebihan sistem dan 20 disebabkan oleh beberapa punca lain. Pada pendapat anda, adakah dakwaan penyelidik itu sah?

Among the major factors that can cause computers to malfunction are software failure, hardware failure, operator error and system overloading. A researcher claims that 10% of malfunction is due to software failure, 5% to hardware failure, 25% to operator error, 40% to system overloading and the rest to other causes. A study was conducted and it was found that 13 malfunctions are due to software failure, 10 to hardware failure, 42 to operator error, 65 to system overloading and 20 due to other causes. Do you think that the researcher's claim is valid?

(9 marks)

(b) *Suatu pembolehubah rawak X mempunyai fungsi ketumpatan kebarangkalian yang diberikan oleh*

A random variable X has probability density function given by

$$f(x;\beta)=\left(\frac{1}{\beta^2}\right)xe^{-\frac{x}{\beta}} \quad \text{for } x>0 \text{ and } \beta>0$$

(i) Dapatkan penganggar kebolehjadian maksimum bagi β berdasarkan sampel rawak bersaiz n.

Find the maximum likelihood estimator for β based on a random sample of size n.

(ii) Jika $E(X)=2\beta$, adakah penganggar kebolehjadian maksimum yang diberikan oleh (i) diatas saksama?

If $E(X)=2\beta$, is the maximum likelihood estimator given in (i) above unbiased?

(9 marks)

(c) Di suatu persimpangan tertentu di Butterworth, purata kelajuan bagi sampel 39 buah kereta adalah 26.5 km/jam dengan sisihan piawai 2.37 km/jam. Purata kelajuan bagi sampel 50 motorsikal adalah 35.0 km/jam dengan sisihan piawai 3.66 km/jam. Dapatkan selang keyakinan 95% bagi perbezaan antara min kelajuan motorsikal dan kereta.

At a particular intersection in Butterworth, the average speed for a sample of 39 cars was 26.5 km/h with a standard deviation of 2.37 km/h. The average speed for a sample of 50 motorcycles was 35.0 km/h with a standard deviation of 3.66 km/h. Find a 95% confidence interval for the difference between the mean speeds of motorcycles and cars.

(7 marks)

...7/-

BAHAGIAN B
SECTION B

4. Peratus kesuburan bagi mani sejenis binatang selepas disimpan telah disukat pada pelbagai gabungan kepekatan tiga bahan yang digunakan untuk menambahkan peluang kesuburannya. Data seperti berikut ;

The percent fertility of certain type of animal semen, after storage, was measured at various combinations of concentrations of three materials used to increase chances of fertility. The data are as follows;

Y (% Kesuburan) *(Fertility in %)*	X_1 (% Berat) *(Weight in %)*	X_2 (% Berat) *(Weight in %)*	X_3 (% Berat) *(Weight in %)*
25.5	1.74	5.30	10.80
31.2	6.32	5.42	9.40
25.9	6.22	8.41	7.20
38.4	10.52	4.63	8.50
18.4	1.19	11.60	9.40
26.7	1.22	5.85	9.90
26.4	4.10	6.62	8.00
25.9	6.32	8.72	9.10
32.0	4.08	4.42	8.70
25.2	4.15	7.60	9.20
39.7	10.15	4.83	9.40
35.7	1.72	3.12	7.60
26.5	1.70	5.30	8.20

(a) Anggarkan model regresi linear berganda untuk data yang diberikan.

Estimate the multiple linear regression model for the given data.

(b) Bina selang keyakinan 95% bagi min sambutan yang diramalkan bila X_1 = 3%, X_2 = 8% dan X_3 = 9%.

Construct a 95% confidence interval for the mean response when X_1 *= 3%,* X_2 *= 8% and* X_3 *= 9%.*

Note :

$$(X^TX)^{-1} = \begin{pmatrix} 8.0640 & -0.0826 & -0.0942 & -0.7905 \\ -0.0826 & 0.0085 & 0.0017 & 0.0037 \\ -0.0942 & 0.0017 & 0.0166 & -0.0021 \\ -0.7905 & 0.0037 & 0.0021 & 0.0886 \end{pmatrix}$$

(25 marks)

5. Empat mesin berlainan dipertimbangkan dalam kerja pemasangan sejenis hasil keluaran. Adalah ditetapkan bahawa enam petugas operasi berlainan digunakan dalam ujikaji rekabentuk blok rawak lengkap untuk membandingkan mesin-mesin tersebut. Operasi mesin-mesin ini memerlukan amaun kekuatan fizikal lengan dan diketahui bahawa terdapat perbezaan di antara petugas-petugas ini dalam kepantasan mereka menggunakan mesin-mesin tersebut. Pengukuran asas bagi mengukur kepantasan menggunakan mesin ialah masa dalam saat sehingga kerja selesai. Data diberikan dalam jadual di bawah;

Four different machines are being considered for the assembling of a particular product. It is decided that 6 different operators are to be used in a randomized block experiment to compare the machines. The machines are assigned in a random order to each operator. The operation of the machines requires physical dexterity and it is the speed with which they operate the machines. The amount of time in seconds was recorded for assembling the product;

Petugas *(Operator)*

Mesin ***Machine***	**1**	**2**	**3**	**4**	**5**	**6**	**Jumlah** ***Total***
1	42.5	39.3	39.6	39.9	42.9	43.6	247.8
2	39.8	40.1	40.5	42.3	42.5	43.1	248.3
3	40.2	40.5	41.3	43.4	44.9	45.1	255.4
4	41.3	42.2	43.5	44.2	45.9	42.3	259.4
Jumlah *Total*	163.8	162.1	164.9	169.8	176.2	174.1	1010.9

(a) Uji hipotesis nol bahawa mesin-mesin itu beroperasi pada kadar min kelajuan yang sama.

Test the null hypothesis that the machines perform at the same mean rate of speed.

(b) Uji hipotesis nol bahawa petugas-petugas itu bekerja pada kadar min kepantasan yang sama.

Test the null hypothesis that the operators perform at the same mean rate of physical dexterity.

(25 marks)

6. Telah diakui bahawa seorang pelajar universiti mampu meningkatkan jumlah markah peperiksaannya dalam satu mata pelajaran tertentu sebanyak sekurang-kurangnya 50 markah jika dia diberikan contoh-contoh soalan peperiksaan lebih awal. Untuk menguji pengakuan ini, 20 orang pelajar dibahagikan kepada 10 pasangan supaya setiap pasangan padanan mempunyai purata markah yang hampir dalam masa 3 tahun pertama di universiti. Contoh-contoh soalan dan jawapan diberikan secara rawak kepada seorang ahli setiap pasangan 1 minggu lebih awal daripada peperiksaan. Jumlah markah peperiksaan berikut direkodkan;

It is claimed that a university student can increase his score in the major field area of the graduate record examination by at least 50 points if he is provided with sample questions in advance. To test this claim, 20 students are divided into 10 pairs such that each matched pair has almost the same overall quality point average for their first 3 years in university. Sample questions and answers are provided at random to one member of each pair 1 week prior to the examination. The following examination scores were recorded;

Pasangan
pair

	1	2	3	4	5	6	7	8	9	10
Dengan contoh soalan *(with sample questions)*	531	621	663	579	451	660	591	719	543	575
Tanpa contoh soalan *(without sample questions)*	509	540	688	502	424	683	568	748	530	524

Uji hipotesis nol bahawa sampel-sampel soalan meningkatkan jumlah markah sebanyak 50 melawan hipotesis alternatif bagi kaedah statistik tak berparameter, bahawa peningkatan itu kurang daripada 50 markah.

Test the null hypothesis that sample questions increase the scores by 50 points against the alternative hypothesis for non-parameter statistic, that the increase is less than 50 points.

(25 marks)

oooOOooo