UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2009/2010

November 2009

IUK 108 – Statistics With Computer Applications
[Statistik Dengan Aplikasi Komputer]

Duration: 3 hours
[Masa: 3 jam]

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

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

Instructions: Answer **FOUR (4)** out of five questions. You may answer the question either in Bahasa Malaysia or in English.

*[**Arahan**: Jawab **EMPAT (4)** daripada lima soalan. Anda dibenarkan menjawab soalan sama ada [untuk KBI] dalam Bahasa Malaysia atau Bahasa Inggeris.]*

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].

…2/-

1. Determine whether each statement is true or false. If the statement is false, explain why.

   (a) Histogram can be drawn using vertical and horizontal bars.

   (b) The type of graph used to represent data is determined by the type of data collected and by the researcher's purpose.

   (c) Data collected over a period of time can be graphed by using a pie graph.

   (d) The mean cannot be found for grouped data when there is an open class.

   (e) The range and midrange are both measures of variation.

   (f) Classical probability uses a frequency distribution to compute probabilities.

   (g) A binomial experiment has a fixed number of trials.

   (h) The z value corresponding to a number below the mean is always negative.

   (i) When one is conducting the t test, the population must be approximately normally distributed.

   (j) The null hypothesis for testing between two means is $\mu_1 - \mu_2 = 0$

   (25 marks)

2. Select the best answer.

   (a) What graph should be used to show the relationship between the parts and the whole?

      (i) Histogram.
      (ii) Pie graph.
      (iii) Pareto chart.

   (b) The complement of guessing 5 correct answers on a 5-question true/false exam is

      (i) Guessing 5 incorrect answers.
      (ii) Guessing at least 1 incorrect answer.
      (iii) Guessing no incorrect answers.

(c) What is the sum of the probabilities of all outcomes in a probability distribution?

(i) It cannot be determined
(ii) 1
(iii) between 0 and 1

(d) Which is not a property of the standard normal distribution?

(i) It's symmetric about the mean.
(ii) It's uniform.
(iii) it's bell-shaped.

(e) Which distribution is used to test the difference between two variances?

(i) Chi-square distribution.
(ii) F test.
(iii) t-test.

(f) The probability that an event happens is 41. What is the probability that the event won't happen?

(i) 0
(ii) 59
(iii) 1

(g) The number of absences per year that a worker has is an example of what type of data?

(i) Continuous.
(ii) Discrete.
(iii) Ordinal.

(h) When a distribution is bell-Shaped, approximately what percentage of data values will fall within 1 standard deviation of the mean?

(i) 50%
(ii) 68%
(iii) 99%

(i) What are the boundaries for 8.6-8.8?

(i) 8-9
(ii) 8.55-8.85
(iii) 8.5-8.9

(25 marks)

3. Complete the statements with the best answers.

(a) A measure obtained from sample data is called a (n) …….

(b) A positive square root of the variance is called the ……

(c) When two events cannot occur at the same time, they are said to be ……

(d) Random variable values are determined by ……..

(e) When one is using the standard normal distribution, $P(Z < 0)$ = ……. .

(f) When the sum of the lowest data and the highest data value is divided by 2, the measure is called …….

(g) The mean for a binomial variable can be found by using the formula ……

(h) Rejecting the null hypothesis when it is true is called a(n) ……..

(i) Kurskal-Wallis uses the …… distribution.

(25 marks)

4. The number of absences and the final grade of seven students selected from mathematics class are given in the following Table.

Student	Number of absences X	Final grade y %
A	8	78
B	5	90
C	12	58
D	9	74
E	15	43
F	2	86
G	6	82

(a) Compute the value of the correlation coefficient.

(b) Find the equation of the regression line (Y dependent variable).

(c) Coefficient of determination.

(25 marks)

5. A medical researcher wishes to see whether the variance of the heart rates (in beats per minute) of smokers is different from the variance of heart rates of people who do not smoke. Two samples are selected and the data are as shown. Using $\alpha = 0.05$, is there enough evidence to support the claim?

Smokers	Nonsmokers
$n_1 = 26$	$n_2 = 18$
$S_1^2 = 36$	$S_{21}^2 = 0$

(25 marks)

1. *Tentukan sama ada setiap kenyataan berikut betul atau salah.*

   (a) Histogram boleh dilukis menggunakan turus menegak atau melintang.

   (b) Jenis graf yang mewakili data ditentukan oleh jenis data yang dikumpul serta tujuan pengumpulannya.

   (c) Data yang dikumpul dalam satu tempoh masa boleh dilakar menggunakan carta pie.

   (d) Purata tidak dapat diperolehi untuk data berkumpulan di mana terdapat kelas terbuka.

   (e) Julat dan titik tengah julat dikira menggunakan variasi.

   (f) Kebarangkalian klasik menggunakan taburan frekuensi untuk mengira kebarangkalian.

   (g) Ujikaji binomial mempunyai bilangan percubaan yang khusus.

   (h) Nilai z yang sepadan dengan nombor yang lebih rendah dari purata sentiasa negatif.

   (i) Apabila menguji ujian t, populasi mesti tertabur secara normal.

   (j) Hipotesis nul untuk menguji dua purata adalah $\mu_1 - \mu_2 = 0$

   (25 markah)

2. *Pilih jawapan yang terbaik.*

   (a) Graf mana yang patut digunakan untuk menunjukkan hubungan antara bahagian dan seluruh?

   (i) Histogram.
   (ii) Carta pie.
   (iii) Carta pareto.

   (b) Pelengkap untuk meneka 5 jawapan betul bagi 5 soalan betul/salah ialah,

   (i) meneka 5 jawapan betul.
   (ii) meneka sekurang-kurangnya 1 jawapan salah.
   (iii) meneka tiada jawapan betul.

(c) Apakah jumlah kebarangkalian bagi semua kesudahan dalam taburan kebarangkalian

(i) Tidak dapat ditentukan.
(ii) 1.
(iii) Antara 0 dan 1.

(d) Yang manakah bukan sifat taburan normal piawai?

(i) Simetri berdekatan purata.
(ii) Ia seragam.
(iii) Berbentuk loceng.

(e) Taburan manakah yang digunakan untuk menguji perbezaan antara 2 variasi?

(i) Taburan kuasa dua-Chi.
(ii) Ujian F.
(iii) Ujian t.

(f) Kebarangkalian suatu peristiwa berlaku adalah 41. Apakah kebarangkalian peristiwa tersebut tidak berlaku?

(i) 0
(ii) 59
(iii) 1

(g) Apakah jenis data bagi bilangan cuti untuk pekerja setiap tahun?

(i) Selanjar.
(ii) Diskrit.
(iii) Ordinal.

(h) Apabila taburan berbentuk loceng, lebih kurang berapakah peratus nilai data yang akan berkurang 1 sisihan piawai daripada purata?

(i) 50%
(ii) 68%
(iii) 99%

(i) *Apakah batas bagi 8.6 – 8.8?*

(i) *8-9*
(ii) *8.55-8.85*
(iii) *8.5-8.9*

(25 markah)

3. *Isi tempat kosong dengan jawapan yang tepat.*

(a) *Suatu ukuran yang diperolehi dari sampel data dipanggil a(n) ……*

(b) *Punca kuasa positif bagi variasi dipanggil ……*

(c) *Apabila 2 peristiwa berlaku pada masa yang sama, ia dipanggil ……..*

(d) *Nilai pemboleh ubah rawak ditentukan oleh ……*

(e) *Apabila menggunakan taburan normal piawai,* $P(Z < 0) = \ldots\ldots$.

(f) *Apabila jumlah nilai data terendah dan data tertinggi dibahagi 2, kiraan tersebut dipanggil ……*

(g) *Purata bagi pemboleh ubah binomial boleh diperolehi dengan menggunakan rumus ……*

(h) *Menolak hipotesis nul apabila ia benar dipanggil a (n) ……*

(i) *Kurskal-Wallis menggunakan taburan ……*

(25 markah)

4. *Bilangan hari cuti dan markah akhir bagi tujuh pelajar terpilih dari kelas matematik adalah seperti dalam Jadual di bawah,*

Pelajar	*Bilangan hari, X*	*Markah akhir y %*
A	*8*	*78*
B	*5*	*90*
C	*12*	*58*
D	*9*	*74*
E	*15*	*43*
F	*2*	*86*
G	*6*	*82*

(a) Kirakan nilai pekali korelasi.

(b) Cari persamaan garis regresi (Y pemboleh ubah bersandar).

(c) Pekali penentuan.

(25 markah)

5. *Seorang pengkaji perubatan berharap untuk menilai sama ada variasi kadar degupan jantung (dalam bilangan degupan per minit) seorang perokok berbeza dengan kadar degupan jantung orang yang tidak merokok. Dua sampel dipilih dan data seperti di bawah. Dengan menggunakan* $\alpha = 0.05$*, adakah terdapat bukti yang cukup untuk menyokong penyataan tersebut?*

Perokok	*Bukan perokok*
$n_1 = 26$	$n_2 = 18$
$S_1^2 = 36$	$S_{21}^2 = 0$

(25 markah)

- oooOooo -