UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2008/2009

April/May 2009

MAT 161 – Elementary Statistics
[Statistik Permulaan]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **all ten** [10] questions.

[Arahan: *Jawab **semua sepuluh** [10] soalan.]*

1. The following frequency distribution gives the age distribution of drivers who were at fault in auto accidents during a one-week period in Georgetown.

Age (X years)	Frequency
18 – 20	2
21 – 30	5
31 – 40	15
41 – 50	20
51 – 60	10
61 – 70	2

(i) Find the mean and standard deviation of the age of the drivers.
(ii) What percentage of the drivers are between 35 and 45 years old?
(iii) Use Chebyshev's Theorem to obtain an interval of the ages of at least 75% of the drivers.

[10 marks]

2. A consumer agency randomly selected 1700 flights for two major airlines, A and B. The following table gives the two-way classification of these flights based on airline and arrival times. Note that "less than 30 minutes late" includes flight that arrived early or on time.

	Less Than 30 minutes Late	30 minutes to 1 hour Late	More Than 1 hour Late
Airline A	429	390	92
Airline B	393	316	80

(a) If one flight is selected at random from these 1700 flights, what is the probability that it is
(i) more than 1 hour late or on airline A?
(ii) on airline B or arrived on time?
(iii) on airline A given that it is more than 1 hour late?

(b) Are the events "Airline A" and "more than 1 hour late" independent? Are they mutually exclusive? Explain why or why not?

[10 marks]

1. *Taburan kekerapan berikut memberikan taburan umur pemandu-pemandu yang didapati bersalah dalam kemalangan kenderaan sepanjang suatu tempoh 1-minggu di Georgetown.*

Umur (X tahun)	*Kekerapan*
18 – 20	*2*
21 – 30	*5*
31 – 40	*15*
41 – 50	*20*
51 – 60	*10*
61 – 70	*2*

(i) Cari min dan sisihan piawai bagi umur pemandu-pemandu ini.
(ii) Berapakah peratusan pemandu yang berumu antara 35 hingga 45 tahun?
(iii) Gunakan Teorem Chebyshev untuk mendapatkan selang/julat umur bagi sekurang-kurangnya 75% daripada pemandu-pemandu ini.

[10 markah]

2. *Sebuah agensi pengguna memilih secara rawak 1700 penerbangan bagi dua buah syarikat penerbangan utama, A dan B. Jadual berikut memberikan pengelasan dua hala bagi penerbangan ini berdasarkan tempoh lewat ketibaan. Perhatikan bahawa "lewat kurang daripada 30 minit" termasuklah penerbangan yang sampai tepat pada masa atau lebih awal.*

	Lewat kurang dariapda 30 minit	*Lewat antara 30 minit hingga 1 jam*	*Lewat lebih daripada 1 jam*
Syarikat A	*429*	*390*	*92*
Syarikat B	*393*	*316*	*80*

(a) Sekiranya suatu penerbangan dipilih secara rawak daripada 1700 penerbangan ini, apakah kebarangkalian bahawa penerbangan ini
(i) lewat lebih daripada satu jam atau dari syarikat penerbangan A?
(ii) dari syarikat penerbangan B atau sampai tepat pada masanya?
(iii) dari syarikat penerbangan A jika diketahui bahawa ia lewat lebih daripada 1 jam?

(b) Adakah perisitiwa "syarikat A" dan "lewat lebih daripada 1 jam" tak bersandar? Adakah kedua-duanya saling ekslusif? Terangkan mengapa?

[10 markah]

3. A small car rental company has 5 cars that it rents out at RM100 per car per day. The daily demand for its cars follows a Poisson distribution with mean 5.
   (i) Calculate the probability that the company cannot meet the demand for its cars on a particular day.
   (ii) What is the probability that demand for its cars can be met on at least 5 days of a week?
   (iii) Write the probability distribution for the company's daily income from car rental.
   (iv) Calculate the expected daily income received from the car rental.

[10 marks]

4. The management of a fast food restaurant does not want its customers to wait in line for service for too long. Suppose that the waiting times for customers have a normal distribution with a mean of 6 minutes and a standard deviation of 2 minutes.
   (i) What is the probability that a randomly selected customer will have to wait for less than 3 minutes?
   (ii) Find the probability that the mean waiting time for a sample of 15 customers is within 1 minute of the population mean.
   (iii) Suppose that the management wants to promote its business by guaranteeing a maximum waiting time for its customers. If a customer is not served within that period, the customer will receive a 50% discount on the bill. The management wants to limit this discount to at most 5% of the customers. What should the maximum guaranteed waiting time be?

[10 marks]

5. Beta Airlines claims that at most, only 15% of its flights arrive more than 10 minutes late. The result from a recent survey shows that 10 out of 50 of its flights arrive more than 10 minutes late.
   (i) At $\alpha = 0.05$, test whether there is enough evidence to reject the airline's claim.
   (ii) The delay in arrival time is known to be normally distributed with mean 8 minutes and standard deviation 2 minutes. If the airline wishes to ensure that the probability is about 0.90 that its sample estimate of the mean delay time lies within 0.2 minutes of the population mean, how many flights should it include in its survey?

[10 marks]

3. *Sebuah syarikat kecil sewa kereta mempunyai 5 buah kereta yang disewakan pada kadar RM100 per kereta sehari. Permintaan harian bagi kereta-kereta ini bertaburan Poisson dengan min 5.*

   (i) Hitung kebarangkalian bahawa syarikat ini tidak dapat memenuhi permintaan kereta pada sesuatu hari tertentu.

   (ii) Apakah kebarangkalian bahawa permintaan kereta dapat dipenuhi pada sekurang-kurangnya 5 hari dalam sesuatu minggu?

   (iii) Tuliskan taburan kebarangkalian bagi pendapatan harian syarikat sewa kereta ini.

   (iv) Hitung jangkaan pendapatan harian dari sewaan kereta.

[10 markah]

4. *Pihak pengurusan sebuah restoran makanan segera tidak mahu pelanggannya menunggu terlalu lama dalam baris menunggu. Andaikan masa tunggu bagi pelanggan-pelanggan bertaburan normal dengan min 6 minit dan sisihan piawan 2 minit.*

   (i) Apakah kebarangkalian bahawa seorang pelanggan yang dipilih secara rawak perlu menunggu kurang dari 3 minit?

   (ii) Cari kebarangkalian bahawa min masa tunggu bagi suatu sampel 15 pelanggan adalah sekitar 1 minit dari min populasi.

   (iii) Andaikan pihak pengurusan ingin mempromosi perniagaannya dengan menjamin suatu masa tunggu maksimum kepada pelanggannya. Jika seorang pelanggan tidak dilayan dalam tempoh berkenaan, pelanggan itu diberi diskaun 50% harga. Pihak pengurusan ingin menghadkan diskaun ini kepada tidak lebih dari 5% pelanggannya. Apakah tempoh masa tunggu maksimum ini?

[10 markah]

5. *Beta Airlines mendakwa bahawa tidak lebih dari 15% daripada penerbangannya tiba lewat melebihi 10 minit. Hasil suatu tinjauan terkini menunjukkan bahawa 10 dari 50 penerbangannya tiba lewat melebihi 10 minit.*

   (i) Uji sama ada wujud cukup bukti untuk menolak dakwaan syarikat penerbangan ini, pada aras keertiant $\alpha = 0.05$.

   (ii) Tempoh kelewatan dalam masa ketibaan diketahui bertaburan normal dengan min 8 minit dan sisihan piawai 2 minit. Jika syarikat penerbangan ingin memastikan bahawa anggaran sampel bagi min tempoh lewat berada sekitar 0.2 minit daripada min populasi dengan kebarangkalian 0.90, berapakah bilangan penerbangan yang perlu dipilih dalam tinjauan ini?

[10 markah]

6. It is known that the random variable $X \sim N(\mu, 3.24)$. Suppose that we would like to test the null hypothesis $\mu = -5.5$ against the alternative hypothesis $\mu \neq -5.5$ with probability of a Type I error of 10%. A random sample of 11 observations of X is taken and the sample mean, $\overline{X}$ determined.

(i) Find the acceptance and rejection regions in terms of $\overline{X}$.

(ii) For the alternative case of $\mu = -7.0$, find the probability of a Type II error and the power of the test.

[10 marks]

7. A consumers' association tests the durability of car tyres by running them on a machine until their tread depth reaches a prescribed minimum. 150 tyres of brand A and 120 tyres of brand B were tested by the association. The equivalent distance, measured in thousands of kilometres, for both brands are summarised as follows:

Brand A: $\sum_{i=i}^{150}(x-30) = 0$, $\sum_{i=1}^{150}(x-30)^2 = 10051$

Brand B: $\sum_{i=i}^{120}(y-36) = 0$, $\sum_{i=i}^{120}(y-36)^2 = 10473$

(i) Test at the 5% significance level whether there is sufficient evidence that $\sigma_A^{\,2} = \sigma_B^{\,2}$?

(ii) Is there a difference between the durability of brand A car tyres and brand B car tyres? Test at the 5% significance level.

[10 marks]

8. Mr. A believes his lucky coin is fair. To test his belief, he tosses the coin and counts the number of tails between successive heads. If the coin is fair, then P(r tails) = $\left(\frac{1}{2}\right)^{r+1}$ for r = 1, 2, 3, ...Mr. A's results are as follows:

Number of tails, r	0	1	2	3	4	≥5
Frequency, f	25	18	9	6	2	0

Determine whether there is significant evidence, at the 5% level, that the coin is biased.

[10 marks]

6. *Diketahui bahawa pembolehubah rawak* $X \sim N(\mu, 3.24)$. *Andaikan bahawa kita ingin menguji hipotesis nol* $\mu = -5.5$ *melawan hipotesis alternatif* $\mu \neq -5.5$ *dengan kebarangkalian Ralat Jenis I bernilai 10%.. Suatu sampel rawak sebanyak 11 cerapan X diambil dan min sampel* $\overline{X}$ *ditentukan.*

(i) Cari rantau penerimaan dan rantau penolakan dalam sebutan $\overline{X}$.

(ii) Untuk kes alternatif $\mu = -7.0$, *cari kebarangkalian Ralat Jenis II dan kuasa ujian.*

[10 markah]

7. *Suatu kesatuan pengguna menguji ketahanan tayar kereta dengan mengujinya pada sebuah mesin sehingga bunga tayar haus ke suatu tahap minimum. 150 tayar berjenama A dan 120 tayar berjenama B diuji oleh kesatuan ini. Jarak setara, diukur dalam ribuan kilometer, bagi kedua-dua jenama diringkaskan seperti berikut:*

$$\text{Jenama } A: \sum_{i=i}^{150}(x-30) = 0 \quad , \quad \sum_{i=1}^{150}(x-30)^2 = 10051$$

$$\text{Jenama } B: \sum_{i=i}^{120}(y-36) = 0 \quad , \quad \sum_{i=i}^{120}(y-36)^2 = 10473$$

(i) Uji sama ada wujud bukti yang cukup bahawa $\sigma_A^{\,2} = \sigma_B^{\,2}$ *pada aras keertian 5%.*

(ii) Wujudkah perbezaan ketahanan tayar kedua-dua jenama ini? Uji pada aras keertian 5%.

[10 markah]

8. *En. A percaya bahawa duit syiling beliau adalah saksama. Untuk menguji kepercayaan ini, dia melambung syiling itu dan mengira bilangan kali "bunga" timbul antara dua "kepala" berturutan. Sekiranya, syiling itu saksama, maka P(r bunga)* $= \left(\frac{1}{2}\right)^{r+1}$ *untuk* $r = 1, 2, 3, \ldots$ *Hasil lambungan En. A adalah seperti berikut::*

Bilangan bunga, r	*0*	*1*	*2*	*3*	*4*	≥ 5
Kekerapan, f	*25*	*18*	*9*	*6*	*2*	*0*

Tentukan sama ada wujud bukti bahawa syiling itu tak saksama, pada aras keertian 5 %.

[10 markah]

9. During an interview for a clerical position, two applicants, A and B were asked to type 15 randomly selected documents. The times taken by each applicant were recorded and ranked. Sum of ranks for applicant A was 218, whilst that of applicant B was 268. Based on these results, can you conclude that there is a difference in the median times taken by the two applicants to type the documents. Use a significance level of 5%.

[10 marks]

10. Texas Instrument (TI) has produced a new version of its popular graphing calculator, and wishes to test its efficiency against the older version. For this purpose, ten third year mathematics students were chosen randomly. Each student was given a statistical problem and were asked to solve it using both versions of the graphing calculator, and the times taken were recorded. The data is as below.

Student	1	2	3	4	5	6	7	8	9	10
Old version of the graphing calculator	25	20	22	30	27	24	27	26	29	28
New version of the graphing calculator	17	24	23	21	20	26	21	24	29	26

(i) If TI wishes to determine if the times taken by the students to solve the problem using the old version of the graphing calculator is shorter than the new version, state the appropriate non-parametric test.

(ii) Using the test you have identified in (i) can you conclude that the new version of the graphing calculator is more efficient? Use a significance level of 5%.

[10 marks]

9. *Semasa sesi temuduga untuk jawatan kerani, dua orang pemohon, A dan B telah diminta menaip 15 keping surat yang telah dipilih secara rawak. Masa yang diambil oleh kedua-dua pemohon telah dicatat dan dipangkatkan. Hasil tambah pangkat bagi pemohon A adalah 218 sementara untuk pemohon B pula ialah 268. Berdasarkan hasil ini, bolehkah anda membuat kesimpulan bahawa terdapat perbezaan dari segi median masa yang diambil oleh kedua-dua pemohon ini untuk menaip dokumen? Gunakan aras keertian 5%.*

[10 markah]

10. *Texas Instrument (TI) telah menghasilkan kalkulator grafik versi baru dan ingin menguji kebaikan versi baru berbanding yang lama. Untuk tujuan ini, sepuluh pelajar matematik USM tahun tiga telah dipilih secara rambang. Setiap seorang pelajar telah diberi suatu masalah statistik dan diminta menyelesaikannya dengan mengguna kalkulator grafik kedua-dua versi, dan masa yang diambil dicatatkan. Datanya ditunjukkan di bawah.*

Pelajar	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Kalkulator grafik versi lama	*25*	*20*	*22*	*30*	*27*	*24*	*27*	*26*	*29*	*28*
Kalkulator grafik versi baru	*17*	*24*	*23*	*21*	*20*	*26*	*21*	*24*	*29*	*26*

(i) Sekiranya, TI ingin menentukan sama ada masa yang diambil pelajar untuk menyelesaikan masalah dengan mengguna kalkulator grafik versi lama adalah lebih singkat berbanding versi baru, nyatakan ujian tak berparameter yang bersesuaian.

(ii) Dengan mengguna ujian yang telah anda kenal pasti dalam (i), bolehkah anda membuat kesimpulan bahawa kalkulator versi baru adalah lebih baik? Gunakan aras keertian 5%.

[10 markah]

APPENDIX/FORMULA

$\bar{x} = \frac{\sum xf}{\sum f}$ $s^2 = \frac{\sum (x^2 f) - \frac{(\sum xf)^2}{\sum f}}{\sum f - 1}$	$S_p{}^2 = \frac{(n_x - 1)s_x{}^2 + (n_y - 1)s_y{}^2}{n_x + n_y - 2}$ $\bar{p} = \frac{X + Y}{n_x + n_y}$
Confidence Intervals: $\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$ $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$ *to* $\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$	$(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_x{}^2}{n_x} + \frac{\sigma_y{}^2}{n_y}}$ $(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{s_x{}^2}{n_x} + \frac{s_y{}^2}{n_y}}$ $(\bar{X} - \bar{Y}) \pm t_{\alpha/2}\sqrt{S_p{}^2\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}$ $(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$
Test Statistics: $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$	$T = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n_d}}}$ $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

Test Statistics:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{S_p^2\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

$$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\frac{p_x(1-p_x)}{n_x} + \frac{p_y(1-p_y)}{n_y}}}$$

$$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

$$F = \frac{s_1^2}{s_2^2}$$

$$\chi^2 = \sum \frac{(O-E)^2}{E}, \quad E = np$$

Nonparametric Statistics:

Wilcoxon Signed-rank: $W = \sum R^+$, $W = \sum R^-$

$$Z = \frac{T - \mu_W}{\sigma_W}, \quad \mu_W = \frac{n(n+1)}{4}, \quad \sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Wilcoxon Rank Sum Test: $U = R - \frac{n(n+1)}{2}$

$$Z = \frac{T - \mu_T}{\sigma_T},$$

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2}, \quad \sigma_T = \sqrt{\frac{n_1 n_2 (n+1)(n_1 + n_2 + 1)}{12}}$$