UNIVERSITI SAINS MALAYSIA

Peperiksaan Kursus Semasa Cuti Panjang
Sidang Akademik 2010/2011

Jun 2011

MAT 161 – Elementary Statistics
[Statistik Permulaan]

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 **all ten** [10] questions.

*[**Arahan:** Jawab **semua sepuluh** [10] 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. The speeds of 50 randomly selected cars that passed through a busy road in a town were monitored by radar and recorded as in the following table:

Speed of car, X (km/hour)	Frequency, f_i
18 - 24	10
24 - 30	22
30 - 36	8
36 - 42	5
42 - 48	3
48 - 54	2

(i) Calculate the mean speed and its standard deviation.
(ii) Cars that exceeded the limit of 40 km/hour were remanded by the police. What percentage of the cars that passed through the road were remanded by the police?
(iii) Use Chebyshev's Theorem to obtain an interval of the speeds of at least 75% of cars that passed through the road.

[15 marks]

1. Kelajuan 50 buah kereta yang dipilih secara rawak yang melalui sebatang jalan sibuk dalam sebuah bandar dipantau dengan alat radar dan direkodkan seperti dalam jadual yang berikut:

Kelajuan kereta, X (km/jam)	*Kekerapan, f_i*
18 - 24	*10*
24 - 30	*22*
30 - 36	*8*
36 - 42	*5*
42 - 48	*3*
48 - 54	*2*

(i) Hitung purata kelajuan kereta dan sisihan piawainya.
(ii) Kereta yang melebihi had 40 km/jam telah ditahan oleh polis. Berapakah peratusan kereta yang melalui jalan tersebut telah ditahan oleh polis?
(iii) Gunakan Teorem Chebyshev untuk mendapatkan suatu selang kelajuan bagi sekurang-kurangnya 75% daripada kereta-kereta yang melalui jalan tersebut

[15 markah]

2. A box contains three white marbles, three red marbles and a blue marble. Three marbles were randomly taken out of the box, without replacement. Let X represents the number of red marbles that were taken out.

(i) Write the probability distribution of X.
(ii) Find the probability that one white, one red and one blue marble were selected.
(iii) Find the probability that at least one red marble was taken out.

[10 marks]

2. Sebuah kotak mengandungi tiga biji guli putih, tiga biji guli merah dan sebiji guli biru. Tiga biji guli dikeluarkan secara rawak daripada kotak tersebut, tanpa pengembalian. Andaikan X mewakili bilangan guli merah yang dikeluarkan.

(i) Tuliskan taburan kebarangkalian bagi X.
(ii) Dapatkan kebarangkalian bahawa sebiji guli putih, sebiji guli merah dan sebiji guli biru terpilih.
(iii) Dapatkan kebarangkalian bahawa sekurang-kurangnya sebiji guli merah dikeluarkan.

[10 markah]

3. The probability density function of a continuous variable X is given by

$$f(x)=\begin{cases}\dfrac{x+1}{6}, & 1<x<3\\ 0, & \textit{otherwise,}\end{cases}$$

(i) Find E(X) and Var(X).
(ii) The median for a continuous probability distribution is a value m that divides the distribution into two equal areas, i.e.

$$\int_{-\infty}^{m} f(x)\,dx = \int_{m}^{\infty} f(x)\,dx = 0.5.$$

Find the median of $f(x)$.

[10 marks]

3. Taburan kebarangkalian pembolehubah selanjar X diberikan oleh

$$f(x)=\begin{cases}\dfrac{x+1}{6}, & 1<x<3\\ 0, & \textit{di tempat lain,}\end{cases}$$

(i) Dapatkan E(X) dan Var(X).
(ii) Median bagi taburan kebarangkalian selanjar ialah suatu nilai m yang membahagikan taburan tersebut kepada dua bahagian yang sama besar iaitu

$$\int_{-\infty}^{m} f(x)\,dx = \int_{m}^{\infty} f(x)\,dx = 0.5.$$

Dapatkan median f(x).

[10 markah]

4. A process for making glass plates produces small bubbles (defects) scattered at random in the glass, at an average rate of four small bubbles per $10m^2$. Consider glass pieces with dimensions 2.5m x 2.0m.

   (i) Determine the probability that a piece of glass contains at least one small bubble.
   (ii) Suppose ten glass pieces were randomly chosen. What is the probability that five of the pieces contain at least one small bubble each?

   [10 marks]

4. Suatu proses pembuatan kepingan kaca menghasilkan gelembung-gelembung kecil (kecacatan) yang tersebar secara rawak dalam kaca dengan kadar purata empat gelembung kecil pada setiap $10m^2$. Pertimbangkan kepingan-kepingan kaca dengan dimensi 2.5m x 2.0m.

   (i) Tentukan kebarangkalian bahawa sekeping kaca mengandungi sekurang kurangnya satu gelembung kecil.
   (ii) Andaikan sepuluh kepingan kaca dipilih secara rawak. Apakah kebarangkalian bahawa lima daripadanya mengandungi sekurang-kurangnya satu gelembung kecil setiap satu?

   [10 markah]

5. Two companies, *A* and *B*, drill wells in a rural area. Company *A* charges a flat fee of RM3500 to drill a well, regardless of its depth. Company *B* charges a basic fee of RM1000 plus RM12 per foot to drill a well. The depths of wells in this area have a normal distribution with a mean of 250 feet and a standard deviation of 40 feet.

   (i) What is the probability that Company *B* would charge more than Company *A* to drill a well?
   (ii) Find the expected amount charged by Company *B* to drill a well in the area.

   [10 marks]

5. Dua buah syarikat, A dan B, menggali perigi di sebuah kawasan pedalaman. Syarikat A mengenakan bayaran pukal sebanyak RM3500 untuk menggali perigi tanpa mengambil kira kedalamannya. Syarikat B mengenakan bayaran asas sebanyak RM1000 campur RM12 bagi setiap kaki kedalaman perigi yang digali. Kedalaman perigi di kawasan tersebut tertabur secara normal dengan min 250 kaki dan sisihan piawai 40 kaki.

   (i) Apakah kebarangkalian bahawa Syarikat B akan mengenakan bayaran yang lebih tinggi daripada Syarikat A untuk menggali sebuah perigi?
   (ii) Dapatkan bayaran yang dijangka dikenakan oleh Syarikat B untuk menggali sebuah perigi di kawasan tersebut.

   [10 markah]

6. It is claimed that the performance of the first year undergraduate students in Statistics in one university can be studied in their ratios of grades. The ratios of distributions to grade A, B and C are 3:5:2 respectively. A total of 100 students are surveyed and the survey reveals the following result:

Grade A	Grade B	Grade C
34	54	12

(i) If the null hypothesis (3:5:2) is true, what are the expected number of students that score A, B and C respectively in Statistics?
(ii) Complete the hypothesis test using the 0.01 significance level.

[10 marks]

6. Terdapat dakwaan yang menyatakan bahawa prestasi pelajar ijazah tahun pertama dalam statistic boleh dikaji dari nisbah gred mereka. Taburan untuk gred A, B dan C adalah dalam nisbah 3:5:2. Seramai 100 mahasiswa ditinjau dan hasil tinjauan adalah seperti yang berikut.

Gred A	*Gred B*	*Gred C*
34	*54*	*12*

(i) Jika hipotesis nol (3:5:2) adalah benar, berapakah bilangan pelajar yang dijangka mendapat gred A, B dan C?
(ii) Lengkapkan pengujian hipotesis dengan menggunakan aras keertian 0.01.

[10 markah]

7. The diameters of apples from an orchard are normally distributed with a mean of 2.63 inches and a standard deviation of 0.25 inch.

(i) What percentage of the apples from this orchard have diameters less than 2.25 inches?
(ii) A random sample of 100 apples is gathered, and the mean diameter obtained is $\bar{x} = 2.56$ inches. What is the probability that the sample mean will be greater than 2.56 inches?

[10 marks]

7. Diameter epal dari sebuah dusun tertabur secara normal dengan min 2.63 inci dan sisihan piawai 0.25 inci.

(i) Apakah peratusan epal dari dusun ini yang mempunyai diameter kurang daripada 2.25 inci?
(ii) Satu sampel rawak 100 biji epal dikumpul dan sampel tersebut mempunyai min $\bar{x} = 2.56$ *inci. Apakah kebarangkalian bahawa min sampel melebihi 2.56 inci?*

[10 markah]

8. Supermarket A requires that all apples from orchards must meet the store's specifications of a mean diameter of 7.0cm and a standard deviation of no more than 0.2cm. Supermarket A sends a buyer to a potential fruit supplier and selects a random sample of 50 apples from the supplier's orchard. The diameter of each apple is measured and the mean is reported to be 6.95cm with a standard deviation of 0.24cm.

(i) Show if the sample sufficiently meets the specification with regard to the mean diameter. Use a significance level of 0.05.
(ii) Is there any sufficient evidence that the apples do not meet the specification with regard to the standard deviation? Conduct a hypothesis test at the 0.05 significance level.
(iii) Consider results from (i) and (ii). Should supermarket A take the apples from the fruit supplier? Explain.

[15 marks]

8. Pasaraya A mensyaratkan bahawa semua epal dari dusun-dusun harus memenuhi spesifikasi min diameter 7.0 cm dengan sisihan piawai tidak melebihi 0.2 cm. Pasaraya A menghantar pembelinya kepada seorang bakal pembekal buah-buahan dan memilih suatu sampel rawak 50 biji epal dari dusunnya. Diameter setiap biji epal diukur dan minnya dilaporkan 6.95cm dengan sisihan piawai 0.24cm.

(i) Tunjukkan sama ada sampel tersebut memenuhi spesifikasi dengan secukupnya dari segi min diameter. Gunakan aras keertian 0.05.
(ii) Adakah terdapat bukti cukup yang menunjukkan bahawa epal-epal tersebut tidak memenuhi spesifikasi dari segi sisishan piawai? Lakukan ujian hipotesis pada aras keertian 0.05.
(iii) Pertimbangkan keputusan (i) dan (ii). Haruskah pasaraya A mengambil epal daripada pembekal buah-buahan tersebut? Jelaskan.

[15 markah]

9. The table below shows the average daily temperature (in degree Celsius) of 10 days in September for two cities, A and B. The samples are assumed to be normally distributed.

City A	14.0	12.5	11.5	12.2	12.4	12.3	11.8	11.9	13.7	13.2
City B	12.0	12.5	11.6	13.3	13.0	13.0	12.1	12.8	12.2	12.6

City A : $\bar{x} = 12.55$ $s^2 = 0.6827$ $s = 0.826$
City B : $\bar{x} = 12.51$ $s^2 = 0.2832$ $s = 0.532$

(i) Conduct a parametric test at the 5% level of significance to determine if the average temperatures between city A and B are different.
(ii) Suppose that the normality assumption cannot be made. Conduct an appropriate alternative test to determine if the average temperature between city A and B are different. Use a significance level of 0.05

[15 marks]

9. *Jadual di bawah menunjukkan purata suhu harian (dalam darjah celsius) selama 10 hari dalam bulan September untuk dua bandaraya, A dan B. Sampel-sampel tersebut diandaikan tertabur secara normal.*

Bandaraya A	*14.0*	*12.5*	*11.5*	*12.2*	*12.4*	*12.3*	*11.8*	*11.9*	*13.7*	*13.2*
Bandaraya B	*12.0*	*12.5*	*11.6*	*13.3*	*13.0*	*13.0*	*12.1*	*12.8*	*12.2*	*12.6*

$$\textit{Bandar A}: \bar{x} = 12.55 \quad s^2 = 0.6827 \quad s = 0.826$$
$$\textit{Bandar B}: \bar{x} = 12.51 \quad s^2 = 0.2832 \quad s = 0.532$$

(i) Jalankan suatu ujian berparameter pada aras keertian 5%, untuk menentukan sama ada suhu purata bandaraya A dan suhu purata bandaraya B berbeza.

(ii) Katakan andaian kenormalan tidak dapat dibuat. Jalankan suatu ujian alternatif yang sesuai untuk menentukan sama ada suhu purata bandaraya A dan suhu purata bandaraya B berbeza. Guna aras keertian 5%.

[15 markah]

10. It is claimed that the registration rate for a Statistics course by first year undergraduates in university B is higher than that at university A. In order to investigate the claim, two independent samples are taken randomly. The results are summarized in the following table.

University	*n*	No. of registered students in Statistics
A	100	48
B	100	54

(i) Construct a 95% confidence interval for the difference between the proportions of the two populations.

(ii) Conduct a hypothesis test to determine if the registration rates in Statistics at the two universities are the same. Use the $\alpha = 0.05$ significance level.

[15 marks]

10. Terdapat dakwaan bahawa kadar pendaftaran bagi kursus Statistik oleh pelajar ijazah tahun pertama di Universiti B lebih tinggi daripada kadar penyertaan di Universiti A. Untuk menyiasat kebenaran dakwaan ini, dua sampel tak bersandar diambil secara rawak. Hasilnya diringkaskan dalam jadual berikut.

Universiti	*n*	*Bil. pelajar berdaftar kursus Statistik*
A	*100*	*48*
B	*100*	*54*

(i) Bina suatu selang keyakinan 95% bagi perbezaan antara kadar kedua-dua populasi.

(ii) Lakukan ujian hipotesis untuk menentukan sama ada kadar pendaftaran bagi kursus Statistik di kedua-dua universiti adalah sama. Gunakan aras keertian $\alpha = 0.05$.

[15 markah]

APPENDIX/FORMULA

$\bar{x} = \frac{\sum xf}{\sum f}$ $s^2 = \frac{\sum (x^2 f) - \frac{\left(\sum xf\right)^2}{\sum f}}{\sum f - 1}$	$S_p^{\ 2} = \frac{(n_x - 1)s_x^{\ 2} + (n_y - 1)s_y^{\ 2}}{n_x + n_y - 2}$ $\bar{p} = \frac{X + Y}{n_x + n_y}$
Confidence Intervals: $\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ $\overline{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$ $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $\frac{(n-1)s^2}{\chi^2_{\alpha/2}}$ *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: $U = R - \frac{n(n+1)}{2}$

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

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