
UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2005/2006

November 2005

MAT 161E – Elementary Statistics
[Statistik Permulaan]

Duration : 3 hours
[Masa : 3 jam]

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

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

Instructions: Answer **all four** [4] questions.

Arahan: Jawab **semua empat** soalan].

...2/-

1. (a) Electric fuses, nominally rated at 30A, are tested by passing a gradually increasing current through them and recording the current (x amperes) at which they blow. The result of the test on a sample of 125 fuses are as follows:

Current (x amperes)	Number of fuses, f
25 – 28	6
28 – 29	12
29 – 30	27
30 – 31	30
31 – 32	18
32 – 33	14
33 – 34	9
34 – 35	4
35 – 40	5

$$\sum xf = 3861.5 \quad \sum x^2 f = 119905.25$$

- (i) Draw a histogram to represent these data. Describe the shape of the distribution based on the histogram.
- (ii) Calculate the median current, mean current and the standard deviation of current.
- (iii) A measure of the skewness of a distribution is given by

$$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

Calculate the value of this measure for the above data.

- (iv) Use Chebyshev's theorem to determine a current interval at which at least 75% of the fuses blow.

(Give your answers correct to 3 decimal places.)

- (b) A college is sending a delegation of six students for a sports meet. The students are to be selected from six badminton players, six tennis players and five squash players. No student plays more than one game. The delegation is to consist of at least one but not more than three players from each game. Find the number of ways in which the delegation can be selected. Show the full details of your working.
- (c) Suppose a certain ophthalmic trait is associated with eye color. A sample of 150 randomly selected individuals are studied with results as follows:

		Eye Color			Total
		Blue	Brown	Other	
Trait	Yes	35	15	10	60
	No	10	55	25	90
	Total	45	70	35	150

...3/-

Let events A be that the individual has blue eyes,
 B be that the individual has brown eyes,
 C be that the individual has the trait.

An individual is selected from the sample. Calculate:

(i) $P(A \cup \bar{C})$ (ii) $P(\bar{A} | C)$

Given that 30% of the blue-eyed individuals wear glasses, as do 20% of the brown-eyed and 10% of the other-color eyed.

- (ii) What is the probability that an individual selected at random wears glasses?
 (iii) An individual is selected at random and found to wear glasses. What is the probability that this individual has blue eyes?

[100 marks]

1. (a) Fius elektrik yang dikadarkan pada 30A diuji dengan mengalirkan arus elektrik yang semakin meningkat menerusinya dan kadar arus elektrik pada mana ia 'meletup' (x amperes) direkodkan. Hasil ujian yang dijalankan terhadap suatu sampel 125 fius adalah seperti yang berikut:

Arus elektrik(x amperes)	Bilangan fius, f
25 – 28	6
28 – 29	12
29 – 30	27
30 – 31	30
31 – 32	18
32 – 33	14
33 – 34	9
34 – 35	4
35 – 40	5

$$\sum xf = 3861.5 \quad \sum x^2 f = 119905.25$$

- (i) Lukis suatu histogram bagi mewakili data di atas. Perihalkan bentuk taburan data berdasarkan histogram tersebut.
 (ii) Hitung median arus elektrik, min arus elektrik dan sisihan piawai arus elektrik.
 (iii) Suatu sukatan kepencongan taburan diberikan oleh rumus:

$$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

Hitung nilai sukatan ini bagi data di atas.

- (iv) Guna teorem Chebyshev bagi menentukan suatu selang arus elektrik pada mana sekurang-kurangnya 75% daripada fius meletup.

(Betulkan jawapan anda dalam 3 tempat perpuluhan.)

...4/-

- (b) Sebuah kolej akan menghantar suatu delegasi enam orang pelajar untuk suatu perjumpaan sukan. Pelajar-pelajar akan dipilih daripada kalangan enam orang pemain badminton, enam orang pemain tenis dan lima orang pemain skuasy. Tiada pelajar yang bermain lebih daripada satu jenis sukan. Delegasi tersebut harus terdiri daripada sekurang-kurangnya seorang pemain tetapi tidak lebih daripada tiga orang pemain daripada setiap jenis sukan. Cari bilangan cara untuk membentuk delegasi tersebut. Tunjukkan kerja kira anda dengan sepenuhnya.
- (c) Andaikan sejenis pembawakan optomologi dikaitkan dengan warna mata. Suatu sampel 150 individu yang dipilih secara rawak dikaji dengan hasil seperti yang berikut:

		Warna mata			Jumlah
		Biru	Perang	Lain	
Pembawakan	Ya	35	15	10	60
	Tidak	10	55	25	90
	Jumlah	45	70	35	150

Katakan A ialah peristiwa bahawa seorang individu mempunyai mata biru,
 B ialah peristiwa bahawa seorang individu mempunyai mata perang,
 C ialah peristiwa bahawa seorang individu mempunyai mata warna lain.

Seorang individu dipilih secara rawak daripada sampel tersebut. Hitung:

(i) $P(A \cup \bar{C})$ (ii) $P(\bar{A} | C)$

Diberikan bahawa 30% daripada kalangan individu bermata biru memakai cermin mata, seperti juga 20% daripada kalangan bermata perang dan 10% daripada kalangan yang bermata warna lain.

- (iii) Berapakah kebarangkalian bahawa seorang individu yang dipilih secara rawak memakai cermin mata?
- (iv) Seorang individu yang dipilih secara rawak didapati memakai cermin mata. Berapakah kebarangkalian bahawa individu tersebut mempunyai mata biru?

[100 markah]

2. (a) The probability density function of a continuous random variable X is given by

$$f(x) = \begin{cases} ax, & 1 < x \leq 3 \\ c(4-x), & 3 \leq x < 4 \\ 0, & \text{otherwise} \end{cases}$$

...5/-

- (i) Show that $a = \frac{2}{11}$ and $c = \frac{6}{11}$.
- (ii) Determine $E(X)$ and $\text{Var}(X)$
- (iii) Find $P(X > 2 | X < 4)$
- (b) A book has 500 pages and 50 of them do not contain any typing errors. Assume that the number of typing errors on a randomly selected page has a Poisson distribution.
- (i) Estimate the probability that a randomly selected page do not contain any typing errors.
- (ii) Show that on the average, there are 2.3 typing errors on a page.
- (iii) Calculate the probability that a randomly selected page has at least one typing error.
- (iv) 100 pages are randomly selected from the book. Approximate the probability that at most, 5 of the pages do not contain any typing errors.
- (c) The scores of an examination are normally distributed with mean 72 and standard deviation 14.
- (i) Students who obtain scores of more than 90 are to receive a special commendation. What proportion of students will receive this special commendation?
- (ii) Find the median score of those who will receive this special commendation.
- (iii) A sample of 15 students are randomly selected and their scores noted. What is the probability that their total score is less than 975?

[100 marks]

2. (a) Fungsi taburan kebarangkalian bagi suatu pembolehubah selanjar X diberikan oleh:

$$f(x) = \begin{cases} ax, & 1 < x \leq 3 \\ c(4-x), & 3 \leq x < 4 \\ 0, & \text{di tempat lain} \end{cases}$$

- (i) Tunjukkan bahawa $a = \frac{2}{11}$ dan $c = \frac{6}{11}$.
- (ii) Tentukan $E(X)$ dan $\text{Var}(X)$
- (iii) Cari $P(X > 2 | X < 4)$

...6/-

- (b) *Sebuah buku mempunyai 500 mukasurat dan 50 daripadanya tidak mengandungi sebarang kesilapan taip. Andaikan bilangan kesilapan taip pada sehelai mukasurat yang dipilih secara rawak mempunyai taburan Poisson.*
- (i) *Anggarkan kebarangkalian bahawa sehelai mukasurat yang dipilih secara rawak tidak mengandungi sebarang kesilapan taip.*
 - (ii) *Tunjukkan bahawa pada puratanya, terdapat 2.3 kesilapan taip pada sehelai mukasurat.*
 - (iii) *Hitung kebarangkalian bahawa sehelai mukasurat yang dipilih secara rawak mengandungi sekurang-kurangnya satu kesilapan taip.*
 - (iv) *100 helai mukasurat dipilih secara rawak daripada buku tersebut. Dapatkan kebarangkalian hampiran bahawa sebanyak-banyaknya, 5 daripada mukasurat tersebut tidak mengandungi sebarang kesilapan taip.*
- (c) *Skor bagi suatu peperiksaan tertabur secara normal dengan min 72 dan sisihan piawai 14.*
- (i) *Pelajar yang memperoleh skor melebihi 90 akan mendapat penghargaan khas. Berapakah kadaran pelajar yang akan mendapat penghargaan khas tersebut?*
 - (ii) *Cari skor median bagi pelajar-pelajar yang akan mendapat penghargaan khas tersebut.*
 - (iii) *Suatu sampel 15 orang pelajar dipilih secara rawak dan skor mereka dicatatkan. Berapakah kebarangkalian bahawa jumlah skor mereka kurang daripada 975?*

[100 markah]

3. (a) A sugar company packs sugar in 1.5-kg bags. The amount of sugar per bag (X kg) varies according to a normal distribution with standard deviation 0.01 kg. The consumer association would like to test the hypothesis $H_0 : \mu = 1.5$ versus the alternative $H_1 : \mu \neq 1.5$ to ensure that the amount of sugar in a bag is as stated on the bag. A random sample taken by the association from the population of 1.5-kg bags of sugar yields the following data (in X kg):

Sample: 1.45, 1.55, 1.52, 1.52, 1.47, 1.55, 1.53, 1.50

- (i) State the acceptance and rejection regions in terms of the sample mean, \bar{X} , if the probability of a Type I error is set at 5%.
- (ii) Calculate the probability of committing a Type II error if $H_1 : \mu = 1.55$.
- (iii) What is the power of the test when $\mu = 1.55$?

...7/-

- (b) The manager of store A claims that its customer's waiting time is shorter than at store B . Random samples of customer's waiting time at both stores result in the following summarized data:

Store A	Store B
$n = 25$	$n = 20$
$\sum_{i=1}^{25} (x_i - 8) = 5$	$\sum_{i=1}^{20} (y_i - 10) = 3.5$
$\sum_{i=1}^{25} (x_i - 8)^2 = 151$	$\sum_{i=1}^{20} (y_i - 10)^2 = 135$

- (i) Estimate μ_x and σ_x^2 , the mean and variance of the waiting time of all customers at Store A .
- Estimate μ_y and σ_y^2 , the mean and variance of the waiting time of all customers at Store B .
- (ii) Assuming a common population variance, i.e. $\sigma_x^2 = \sigma_y^2 = \sigma^2$, estimate σ^2 .
- (iii) State any other assumptions that are necessary.
- (iv) Perform a test to determine whether there is evidence, at the 10% significance level, that the above data support the claim made by the manager of store A .
- (c) A food processor produces large quantities of jars of jam. In each batch, the gross weight of a jar (X grams) is known to be normally distributed with standard deviation 7.5 grams. (The gross weight is the weight of the jar plus the weight of the jam.). The gross weights, in x grams, of a random sample from a particular batch were:

514	485	501	486	502	502
496	509	491	497	501	484
506	486	498	490	484	501
494	501	506	490	487	488
507	496	505	498	499	498

$$\sum_{i=1}^{30} x_i = 14,902, \quad \sum_{i=1}^{30} x_i^2 = 7,404,248$$

...8/-

- (i) Estimate the proportion of this batch with gross weight over 500 grams.
Calculate an approximate 95% confidence interval for this proportion.
What do you understand by a 95% confidence interval in this context?
Explain.
- (ii) Calculate a 90% confidence interval for the mean gross weight of this batch.
- (iii) Determine the minimum number of jars that you need to sample if you would like to be 95% confident that an estimate of the mean weight of a jar is within 2 grams of the true mean.

[100 marks]

3. (a) Sebuah syarikat gula memekkan gula dalam kampak-kampak 1.5-kg. Amaun gula dalam setiap kampak (X kg) berubah-ubah mengikut taburan normal dengan sisihan piawai 0.01 kg. Persatuan pengguna ingin menguji hipotesis $H_0 : \mu = 1.5$ berlawanan alternatif $H_1 : \mu \neq 1.5$ untuk memastikan bahawa amaun gula dalam satu kampak adalah seperti yang dinyatakan pada kampak tersebut. Suatu sampel rawak yang diambil oleh persatuan tersebut daripada populasi kampak gula 1.5-kg menghasilkan data yang berikut (dalam X kg):

Sampel: 1.45, 1.55, 1.52, 1.52, 1.47, 1.55, 1.53, 1.50

- (i) Nyatakan kawasan penerimaan dan kawasan penolakan dalam sebutan min sampel, \bar{X} , jika kebarangkalian ralat Jenis I ditetapkan pada 5%.
- (ii) Hitung kebarangkalian melakukan ralat Jenis II jika $H_1 : \mu = 1.55$.
- (iii) Berapakah kuasa ujian apabila $\mu = 1.55$?
- (b) Pengurus kedai A mendakwa bahawa masa menunggu bagi pelanggan di kedainya adalah lebih pendek daripada masa menunggu bagi pelanggan di kedai B. Sampel-sampel rawak masa menunggu bagi pelanggan di kedua-dua kedai menghasilkan data yang diringkaskan seperti berikut:

Kedai A	Kedai B
$n = 25$	$n = 20$
$\sum_{i=1}^{25} (x_i - 8) = 5$	$\sum_{i=1}^{20} (y_i - 10) = 3.5$
$\sum_{i=1}^{25} (x_i - 8)^2 = 151$	$\sum_{i=1}^{20} (y_i - 10)^2 = 135$

- (i) Anggarkan μ_x dan σ_x^2 , min dan varians masa menunggu bagi pelanggan di kedai A.
Anggarkan μ_y dan σ_y^2 , min dan varians masa menunggu bagi pelanggan di kedai B.

...9/-

- (ii) Dengan andaian bahawa varians populasi adalah sepunya, iaitu $\sigma_x^2 = \sigma_y^2 = \sigma^2$, anggarkan σ^2 .
- (iii) Nyatakan sebarang andaian lain yang perlu.
- (iv) Jalankan suatu ujian untuk menentukan sama ada wujud bukti, pada aras keertian 10%, bahawa data di atas menyokong dakwaan yang dibuat oleh pengurus kedai A.
- (c) Sebuah pemproses makanan menghasilkan balang-balang jam dalam kuantiti yang besar. Dalam setiap kumpulan hasil, berat kasar satu balang (X grams) diketahui tertabur secara normal dengan sisihan piawai 7.5 grams. (Berat kasar ialah berat balang berserta berat jam.) Berat kasar, dalam x grams, bagi satu sampel rawak daripada suatu kumpulan hasil ialah:

514	485	501	486	502	502
496	509	491	497	501	484
506	486	498	490	484	501
494	501	506	490	487	488
507	496	505	498	499	498

$$\sum_{i=1}^{30} x_i = 14,902, \quad \sum_{i=1}^{30} x_i^2 = 7,404,248$$

- (i) Anggarkan kadaran yang mempunyai berat kasar melebihi 500 gram bagi kumpulan tersebut.
- Dapatkan selang keyakinan 95% hampiran bagi kadaran tersebut.
- Apakah yang anda faham tentang selang keyakinan 95% dalam konteks masalah di atas? Jelaskan.
- (ii) Hitung suatu selang keyakinan 90% bagi min berat kasar kumpulan tersebut..
- (iii) Tentukan bilangan balang yang minimum yang perlu anda sampelkan jika anda ingin 95% yakin bahawa suatu anggaran min berat sebuah balang adalah dalam sekitar 2 gram daripada min yang sebenar.

[100 markah]

4. (a) State the differences between
- (i) a statistic and a parameter
 - (ii) a binomial distribution and a multinomial distribution
 - (iii) a quantitative variable and a qualitative variable

...10/-

- (b) A new insect spray is applied to 100 samples with 5 insects each. The following table shows the number of samples with a specified number of insects still alive after 10 minutes..

number of insects alive, x	0	1	2	3	4	5
number of samples, f	17	37	29	14	2	1

- (i) Find the mean number of insects still alive per sample and hence, find the estimate of p , the probability that an insect is still alive after 10 minutes.
- (ii) Test, at the 5% significance level, whether the above data follows a binomial distribution.
- (c) A dietitian conducted a study to compare the caloric content (x) and the fat content (y) in 18 of the most popular fast-food items. The results of the study is shown below:

x	120	200	220	230	270	290	310	340	360
y	7	13	11	12	10	8	26	28	8

x	370	420	440	450	460	540	550	640	740
y	36	20	20	22	22	55	25	40	20

$$\sum x = 6950 \quad \sum y = 383 \quad \sum x^2 = 3,126,300 \quad \sum y^2 = 10,885 \quad \sum xy = 168,490$$

- (i) Calculate the correlation coefficient, r . Give your comments on the value obtained.
- (ii) Find the equation of the line of best fit.
- (iii) Explain the meaning of the y -intercept and the slope of the equation.
- (iv) Given that for each value of X , the measured value of Y contains a random error which is normally distributed with mean zero and standard deviation 15. Calculate the probability that, when $X = 250$, the measured value of Y exceeds 15.

[100 marks]

4. (a) Nyatakan perbezaan-perbezaan antara

- (i) *suatu statistik dan suatu parameter.*
- (ii) *suatu taburan binomial dan suatu taburan multinomial*
- (iii) *suatu pembolehubah kuantitatif dan suatu pembolehubah kualitatif*

- (b) *Suatu penyembur serangga yang baru digunakan terhadap 100 sampel yang mengandungi 5 ekor serangga setiap satu. Jadual yang berikut menunjukkan bilangan sampel dengan suatu bilangan serangga yang masih hidup selepas 10 minit.*

...11/-

bilangan serangga yang hidup, x	0	1	2	3	4	5
bilangan sampel, f	17	37	29	14	2	1

- (i) Cari min bilangan serangga yang masih hidup dalam setiap sampel dan seterusnya, cari anggaran bagi p , kebarangkalian bahawa seekor serangga masih hidup selepas 10 minit.
- (ii) Uji pada aras keertian 5% sama ada data di atas tertabur secara binomial.
- (c) Seorang pakar pemakanan menjalankan suatu kajian untuk membandingkan kandungan kalori (x) dan kandungan lemak (y) dalam 18 jenis makanan segera yang paling digemari. Hasil kajiannya adalah seperti yang berikut:

x	120	200	220	230	270	290	310	340	360
y	7	13	11	12	10	8	26	28	8

x	370	420	440	450	460	540	550	640	740
y	36	20	20	22	22	55	25	40	20

$$\sum x = 6950 \quad \sum y = 383 \quad \sum x^2 = 3,126,300 \quad \sum y^2 = 10,885 \quad \sum xy = 168,490$$

- (i) Hitung pekali korelasi linear, r . Beri komen anda mengenai nilai yang diperoleh.
- (ii) Cari persamaan bagi garis penyuaiian terbaik.
- (iii) Terangkan makna pintasan- y dan kecerunan persamaan tersebut.
- (iv) Diberikan bahawa bagi setiap nilai X , nilai Y yang disukat mempunyai ralat rawak yang tertabur secara normal dengan min sifar dan varians 15. Hitung kebarangkalian bahawa apabila $X = 250$, nilai Y yang disukat melebihi 15.

[100 markah]

...12/-

FORMULA

$\bar{x} = \frac{\sum x}{n}$ $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$	$S_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$ $\bar{p} = \frac{X + Y}{n_x + n_y}$
--	---

Confidence Interval

$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ $\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$(\bar{X} - \bar{Y}) \pm z_{\alpha/2} \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$ $(\bar{X} - \bar{Y}) \pm t_{\alpha/2} \sqrt{S_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}$ $(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$
---	--

Test Statistics

$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ $T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$ $T = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n_d}}}$ $T = \frac{b - \beta_1}{s_b}$	$T = r \sqrt{\frac{n-2}{1-r^2}}$ $Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$ $T = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{S_p^2 \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$	$Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\frac{p_x(1-p_x)}{n_x} + \frac{p_y(1-p_y)}{n_y}}}$ $Z = \frac{(\hat{p}_x - \hat{p}_y) - (p_x - p_y)}{\sqrt{\bar{p}(1-\bar{p}) \left(\frac{1}{n_x} + \frac{1}{n_y} \right)}}$ $\chi^2 = \sum \frac{(O-E)^2}{E}, \quad E = np$
---	---	---

Regression and Correlation

$S_{XY} = \sum xy - \frac{(\sum x)(\sum y)}{n}$ $s_e = \sqrt{\frac{S_{YY} - bS_{XY}}{n-2}} ; \quad s_b = \frac{s_e}{\sqrt{S_{XX}}} ; \quad r = \frac{S_{XY}}{\sqrt{S_{XX} \cdot S_{YY}}}$

- 000 O 000 -