

SULIT



**KEMENTERIAN PENDIDIKAN TINGGI
JABATAN PENDIDIKAN POLITEKNIK DAN KOLEJ KOMUNITI**

**BAHAGIAN PEPERIKSAAN DAN PENILAIAN
JABATAN PENDIDIKAN POLITEKNIK DAN KOLEJ KOMUNITI
KEMENTERIAN PENDIDIKAN TINGGI**

JABATAN PERDAGANGAN

PEPERIKSAAN AKHIR

SESI II : 2024/2025

DPA10253 : STATISTICS

TARIKH : 20 MEI 2025

MASA : 2.30 PETANG – 4.30 PETANG (2 JAM)

Kertas ini mengandungi **TIGA BELAS (13)** halaman bercetak.

Struktur (4 soalan)

Dokumen sokongan yang disertakan : Formula

JANGAN BUKA KERTAS SOALAN INI SEHINGGA DIARAHKAN

(CLO yang tertera hanya sebagai rujukan)

SULIT

INSTRUCTION:

This section consists of **FOUR (4)** subjective questions. Answer **ALL** the questions.

ARAHAN:

*Bahagian ini terdiri daripada **EMPAT (4)** soalan subjektif. Jawab **SEMUA** soalan.*

QUESTION 1

- CLO1 a) List **FIVE (5)** names that contributed in the history of statistics. [5 marks]
- CLO1 b) Explain **TWO (2)** roles of statistics in daily life. [5 marks]
- CLO1 c) A grocery store tracks the prices of three essential items (rice, sugar, and milk) over two years. The prices are as follows:

Table 1/*Jadual 1*

Item/ <i>Perkara</i>	Price in 2023 (Base Year)	Price in 2024 (Current Year)
Rice/ <i>Beras</i>	RM 20	RM 25
Sugar/ <i>Gula</i>	RM 15	RM 18
Milk/ <i>Susu</i>	RM 10	RM 12

Using the simple index number formula, calculate the price index for each item and the overall price index for 2024, by considering 2023 as the base year. Show all steps clearly and interpret the result.

[15 marks]

SOALAN 1

- CLO1 a) Senaraikan **LIMA (5)** nama tokoh yang telah menyumbang dalam sejarah statistik.
- [5 Markah]
- CLO1 b) Terangkan **DUA (2)** peranan statistik dalam kehidupan harian.
- [5 Markah]
- CLO1 c) Sebuah kedai runcit menjejaki harga tiga barangan keperluan (beras, gula dan susu) dalam tempoh dua tahun. Harga adalah seperti berikut.

Jadual 1

Perkara	Harga pada tahun 2023 (Tahun Asas)	Harga pada tahun 2024 (Tahun Semasa)
Beras	RM 20	RM 25
Gula	RM 15	RM 18
Susu	RM 10	RM 12

Menggunakan formula nombor indeks mudah, hitung indeks harga untuk setiap item dan indeks harga keseluruhan untuk 2024, dengan mempertimbangkan 2023 sebagai tahun asas. Tunjukkan semua langkah dengan jelas dan tafsirkan hasil dapatannya.

[15 Markah]

QUESTION 2

CLO1

a) The following dataset represents the weights (in kg) of 12 individuals:

48, 55, 52, 60, 58, 49, 67, 62, 50, 70, 53, 59

i. Write the dataset in ascending order

[1 mark]

ii. Calculate the 25th percentiles

[3 marks]

iii. Calculate the 50th percentiles

[3 marks]

iv. Calculate 75th percentiles of the dataset.

[3 marks]

v. Using the calculated percentiles, construct a boxplot by labelling the minimum, maximum, median, Q1, and Q3. Calculate interquartile range, lower bound and upper bound to show if any outliers are present.

[10 marks]

CLO1

b) The following data representing the scores of 8 students in a test:

45, 50, 55, 60, 65, 70, 75, 80

Analyze the dispersion of the dataset by calculating:

i. The range.

[1 mark]

ii. The variance.

[2 marks]

iii. The standard deviation

[2 marks]

SOALAN 2

CLO1

a) Set data berikut mewakili berat (dalam kg) bagi 12 individu:
48, 55, 52, 60, 58, 49, 67, 62, 50, 70, 53, 59

i. Tulis set data dalam susunan menaik

[1 Markah]

ii. Kira persentil ke-25

[3 Markah]

iii. Kira persentil ke-50

[3 Markah]

iv. Kira persentil ke-75 daripada set data

[3 Markah]

v. Menggunakan persentil yang dikira, bina plot kotak dengan melabelkan minimum, maksimum, median, $Q1$ dan $Q3$ dengan jelas. Kira julat antara kuartil, batas bawah dan batas atas untuk menunjukkan jika terdapat sebarang outlier.

[10 Markah]

CLO1

b) Data berikut yang mewakili markah 8 pelajar dalam satu ujian:
45, 50, 55, 60, 65, 70, 75, 80

Analisis penyebaran set data dengan mengira:

i. Julat

[1 Markah]

ii. Varians

[2 Markah]

iii. Sisihan piawai

[2 Markah]

QUESTION 3

CLO1

- a) At a restaurant, Ahmad is allowed to choose one item each for the main course, dessert and beverage from menu below.

Table 2

Main course	White rice	Tomato rice	Fried rice	
Dessert	Ice cream	Chocolate cake	Pudding	
Beverage	Tea	Coffee	Soft drink	Fruit juice

Based on the table above, explain:

- i. How many different selections are possible?

[2 marks]

- ii. What is the probability that Ahmad chooses chocolate cake as the dessert?

[4 marks]

- iii. What is the probability that Ahmad includes tomato rice and ice cream in his selection?

[4 marks]

CLO1

- b) A school conducted a survey among 150 students about their preference for two sports, which are football and basketball. The results are as follows:

- 60 students like both football and basketball.
- 40 students like only football.

- 30 students like only basketball.
- 20 students do not like either sport.

i. Prepare a contingency table showing the distribution of preferences.

[4 marks]

ii. Calculate the probability that a randomly chosen student likes basketball.

[2 marks]

iii. Calculate the probability that a student likes football given that they also like basketball.

[2 marks]

iv. Calculate the probability that a randomly chosen student likes at least one of the two sports

[2 marks]

CLO1

c) Calculate how many three-digit numbers can be formed from the digits 1,2,3,4,5 and 6 if:

i. repetition is allowed

[2 Marks]

ii. repetition is not allowed

[3 marks]

SOALAN 3

CLO1

- a) Di sebuah restoran, Ahmad dibenarkan memilih satu item setiap satu untuk hidangan utama, pencuci mulut dan minuman daripada menu di bawah.

Jadual 2

<i>Menu Utama</i>	<i>Nasi Putih</i>	<i>Nasi Tomato</i>	<i>Nasi Goreng</i>	
<i>Pencuci Mulut</i>	<i>Aiskrim</i>	<i>Kek Coklat</i>	<i>Puding</i>	
<i>Minuman</i>	<i>Teh</i>	<i>Kopi</i>	<i>Minuman Ringan</i>	<i>Jus buah</i>

Berdasarkan jadual di atas, jelaskan:

- i. Berapa banyak pilihan berbeza yang kemungkinan dipilih?

[2 Markah]

- ii. Apakah kebarangkalian Ahmad memilih kek coklat sebagai pencuci mulut?

[4 Markah]

- iii. Apakah kebarangkalian Ahmad memasukkan nasi tomato dan aiskrim dalam pilihannya?

[4 Markah]

CLO1

- b) Sebuah sekolah menjalankan tinjauan di kalangan 150 pelajar tentang keutamaan mereka untuk dua sukan iaitu bola sepak dan bola keranjang. Hasilnya adalah seperti berikut:

- 60 pelajar suka bola sepak dan bola keranjang.
- 40 pelajar hanya suka bola sepak.

- 30 pelajar hanya suka bola keranjang.
- 20 pelajar tidak suka kedua-dua sukan.

i. Sediakan jadual luar jangka yang menunjukkan taburan keutamaan

[4 Markah]

ii. Kira kebarangkalian bahawa pelajar yang dipilih secara rawak menyukai bola keranjang.

[2 Markah]

iii. Kira kebarangkalian pelajar suka bola sepak memandangkan mereka juga suka bola keranjang.

[2 Markah]

iv. Kira kebarangkalian bahawa pelajar yang dipilih secara rawak menyukai sekurang-kurangnya satu daripada dua sukan

[2 Markah]

CLO1

c) Kira berapa banyak nombor tiga digit yang boleh dibentuk daripada digit 1,2,3,4,5 dan 6 jika:

i. pengulangan dibenarkan

[2 Markah]

ii. pengulangan tidak dibenarkan

[3 Markah]

QUESTION 4

CLO1

- a) A researcher wants to estimate the average monthly expenditure on school supplies among 10,000 high school students in a city. Instead of surveying all students, the researcher decides to apply random sampling and use sampling distribution concepts.

- i. Explain what is random sampling

[2 marks]

- ii. Explain what is sampling distribution and how sampling distribution helps in making inferences about the population.

[3marks]

CLO1

- b) A survey was conducted to estimate the average monthly expenditure on transportation for university students. A random sample of 50 students was selected, and their mean expenditure was found to be RM75 with a standard deviation of RM15.

- i. Calculate the 95% confidence interval for the true mean monthly transportation expenditure of university students. Show all your calculations including standard error, margin error and confidence interval.

[10 marks]

CLO1

c) Inferential statistics is the branch of statistics that deals with generalizing population parameters using sample data. Inferential statistics uses a random sample of data taken from a population to describe and make inferences about the population. Two main methods used in inferential statistics are parameter estimation and hypothesis testing.

i) Write what do you understand with the term “hypothesis testing”

[2 marks]

ii) Examine the different types of hypothesis testing used in statistical analysis.

1) Z-test

[2 marks]

2) T-test

[2 marks]

3) Chi-square test

[2 marks]

4) ANOVA

[2 marks]

SOALAN 4

CLO1

- a) *Seorang penyelidik ingin menganggarkan purata perbelanjaan bulanan untuk bekalan sekolah di kalangan 10,000 pelajar sekolah menengah di sebuah bandar. Daripada meninjau semua pelajar, penyelidik memutuskan untuk menggunakan persampelan rawak dan menggunakan konsep taburan persampelan.*

i. *Terangkan apa itu persampelan rawak*

[2 Markah]

ii. *Terangkan apa itu taburan persampelan dan bagaimana taburan persampelan membantu dalam membuat kesimpulan tentang populasi.*

[3 Markah]

CLO1

- b) *Satu tinjauan telah dijalankan untuk menganggarkan purata perbelanjaan bulanan untuk pengangkutan untuk pelajar universiti. Sampel rawak 50 pelajar telah dipilih, dan purata perbelanjaan mereka didapati RM75 dengan sisihan piawai RM15*

i. *Kira selang keyakinan 95% untuk purata perbelanjaan pengangkutan bulanan sebenar pelajar universiti. Tunjukkan semua pengiraan anda termasuk ralat standard, ralat margin dan selang keyakinan*

[10 Markah]

CLO1

c) *Statistik inferensial ialah cabang statistik yang berkaitan dengan menyamaratakan parameter populasi menggunakan data sampel. Statistik inferensial menggunakan sampel rawak data yang diambil daripada populasi untuk menerangkan dan membuat kesimpulan tentang populasi. Dua kaedah utama yang digunakan dalam statistik inferensial ialah anggaran parameter dan ujian hipotesis.*

i) *Tulis apa yang anda fahami dengan "ujian hipotesis"*

[2 Markah]

ii) *Periksa pelbagai jenis ujian hipotesis yang digunakan dalam analisis statistik.*

1) *Ujian Z*

[2 Markah]

2) *Ujian T*

[2 Markah]

3) *Ujian Chi-kuasa dua*

[2 Markah]

4) *Ujian ANOVA*

[2 Markah]

SOALAN TAMAT

FORMULA STATISTICS		
Sturges Rule	$k = 1 + 3.3 \log_{10} n$	
Range	$R = \text{Highest value} - \text{Lowest value}$	
Class Width	$c = \frac{\text{Range}}{k}$	
Mean	$\bar{x} = \frac{\sum x}{n}$	$\bar{x} = \frac{\sum fx}{\sum f}$
Mode	$\hat{x} = L_{Mo} + \left[\frac{d_1}{d_1 + d_2} \right] C$	
Median	$\tilde{x} = L_{Med} + \left[\frac{\frac{N}{2} - F}{f_m} \right] C$	
Quartile	$Q_k = L_{Q_k} + \left[\frac{\frac{kN}{4} - F}{f_{Q_k}} \right] C \quad ; k = 1, 2, 3$	
Percentile	$P_k = L_{P_k} + \left[\frac{\frac{kN}{100} - F}{f_{p_k}} \right] C \quad ; k = 1, 2, 3, \dots, 99$	
Mean Deviation	$E = \frac{\sum x - \bar{x} }{n}$	$E = \frac{\sum (x - \bar{x} f)}{\sum f}$
Variance	$s^2 = \frac{\sum (x - \bar{x})^2}{n}$	$s^2 = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n}$
	$s^2 = \frac{\sum [(x - \bar{x})^2 f]}{\sum f}$	$s^2 = \frac{\sum fx^2}{\sum f} - \left[\frac{\sum fx}{\sum f} \right]^2$
Standard Deviation	$s = \sqrt{\text{variance}}$	
Lower Outlier Boundary	$\hat{x} = \bar{x} - 3(\bar{x} - \tilde{x})$	
Coefficient of Skewness	$PCS\ 1 = \frac{\bar{x} - \hat{x}}{s}$	$PCS\ 2 = \frac{3(\bar{x} - \tilde{x})}{s}$
Coefficient of Variation	$cv = \frac{s}{\bar{x}} \times 100$	

FORMULA STATISTICS		
Pearson Correlation Coefficient	$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$	
Spearman's Rank Correlation Coefficient	$\rho = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$	
Slope of The Regression Line	$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$	
y-intercept of The Regression Line	$a = \frac{\sum y}{n} - b \frac{\sum x}{n}$	
Simple Linear Regression Line	$y = a + bx$	
Probability	$P(A) = \frac{n(A)}{n(S)}$	$P(A \cup B) = P(A) + P(B)$
	$P(A B) = \frac{P(B A) \cdot P(A)}{P(B)}$	$P(B A) = \frac{P(A \cap B)}{P(A)}$
	$P(A \cap B) = P(A) \cdot P(B)$	$P(A \cap B) = P(A) \cdot P(B A)$
	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$	
Confidence Interval	$CI = \bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	
Z-score	$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$	
t-statistic	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$	