

**FINAL EXAMINATION****BACHELOR OF ACCOUNTANCY (HONOURS)****COURSE : STATISTICS FOR ACCOUNTING****COURSE CODE : STA2173****DURATION : 3 HOURS****INSTRUCTIONS TO CANDIDATES:**

1. This question paper consists of **FOUR (4)** questions.
2. Answer ALL questions in the Answer Booklet provided.
3. Please check to make sure that this examination pack consists of:
 - i. The Question Paper
 - ii. an Answer Booklet
 - iii. Appendix 1(1) and 1(2)
4. Do not bring any material into the examination hall. Electronic calculator is allowed.
5. Please write your answer using permanent ink.

MYKAD/ PASSPORT NO. : _____**ID. NO. :** _____**LECTURER :** _____**SECTION :** _____

DO NOT OPEN THIS QUESTION PAPER UNTIL YOU ARE TOLD TO DO

This question paper consists of 10 printed pages including the front page

QUESTION 1

- a. What is meant by a base year in index numbers? (1 mark)
- b. The table below shows the prices and quantities for commodities A, B and C in 2022, 2023 and 2024.

Table 1

Commodity	Quantity (weight)	Price per unit (RM)		
		2022	2023	2024
A	250	4.50	3.40	3.00
B	400	8.30	9.10	10.00
C	550	10.20	8.50	9.80

- i. Calculate the simple relative price index for each item for 2023 using 2022 as the base year. (3 marks)
- ii. Use the information in (i) to calculate the average relative price index for 2023. (3 marks)
- iii. Determine the weighted aggregate price index for 2024 using 2023 as the base year. (3 marks)
- c. Steels Industry produces electrical components which require three main raw materials. The following table shows the price and quantity required for each material. Consider 2021 as the base year.

Table 2

Raw Material	Price per kg (RM)		Quantity (kg)	
	2021	2023	2021	2023
D	4.00	5.50	4000	3500
E	6.00	b	2750	3000
F	5.00	4.30	2000	3500

- i. Given the Laspeyres' price index for the year 2023 is 117.29, find **b** value. (4 marks)
- ii. Calculate the Paasche's quantity index for the year 2023 and explain its meaning. (4 marks)

- d. A university cafeteria is analyzing the daily number of vegetarian meals sold over the past month (30 days). The summary statistics are given in Table 3 below.

Table 3	
<i>Daily Vegetarian Meals Sold</i>	
Mean	60.3
Standard Error	2.05
Median	54
Mode	50
Standard Deviation	11.2
Sample Variance	125.4
Kurtosis	0.62
Skewness	1.28
Range	42
Minimum	41
Maximum	83
Sum	1809
Count	30

- i. Before conducting further analysis, the analyst wants to verify whether the data distribution meets the assumption of normality.

Explain why assessing normality is important in this context.

(2 marks)

- ii. The analyst claims that the distribution of daily number of vegetarian meals sold is approximately normal. Using the skewness and kurtosis values in Table 3, assess whether this claim is reasonable. Justify your answer.

(2 marks)

- iii. If a histogram is constructed for the dataset, describe the likely shape of the distribution.

(2 marks)

- iv. Based on the skewness value, state the direction of skewness of the dataset.

(1 mark)

(Total: 25 marks)

QUESTION 2

- a. A university wants to assess whether a new online learning support system has an effect on students' average weekly study hours. Data were collected from a random sample of 20 students after using the system for one semester. The results are shown in Table 4.

Table 4

Weekly Study Hours	n	Mean	Std. Deviation
	20	14.25	G

- i. Given that the standard error of the mean is 0.47, calculate the standard deviation, **G**, of the weekly study hours. (2 marks)
- ii. State the null and alternative hypotheses to test whether the average weekly study time is 15 hours. (2 marks)
- iii. Calculate the 95% confidence interval for the average weekly study hours. Explain the values obtained. (Given $t_{0.025,19} = \pm 2.093$). (5 marks)
- iv. Based on the confidence interval in (iii), can we conclude that the average weekly study hours is 15 hours? (3 marks)

- b. A university is evaluating whether two different teaching strategies (Strategy A and Strategy B) lead to different average final exam scores in an introductory statistics course. Independent random samples of students were taught using each strategy over the same semester. The summary output of a two-sample t-test assuming equal variances ($\alpha = 0.05$) is given in Table 5 below.

Table 5
t-Test: Two-Sample Assuming Equal Variances

	<i>Strategy A</i>	<i>Strategy B</i>
Mean	72.4	78.1
Variance	16.5	15.3
Observations	12	12
Pooled Variance	15.90	
Hypothesized Mean Difference	0	
df	22	
t Stat	-3.503	
P(T<=t) one-tail	0.0010	
t Critical one-tail	1.717	
P(T<=t) two-tail	0.0020	
t Critical two-tail	2.074	

- i. State a reason why the chosen statistical test is suitable for this study.
(2 marks)
- ii. Using the critical value approach, determine if there is a significant difference in the mean final exam scores between the two teaching strategies. Use a significance level of $\alpha = 0.05$.
(3 marks)

- c. A sports science researcher wants to investigate whether a new warm-up routine has an effect on the average 100-meter sprint time for athletes. The sprint times (in seconds) of 14 athletes were recorded before and after implementing the warm-up. The summary statistics for the paired sample data at 1% significance level are provided in Table 6.

Table 6
t-Test: Paired Two Sample for Means

	<i>Before</i>	<i>After</i>
Mean	12.480	12.014
Variance	0.322	0.281
Observations	14	14
Pearson Correlation	0.936	
Hypothesized Mean Difference	0	
df	13	
t Stat	<i>H</i>	
P(T<=t) one-tail	0.002	
t Critical one-tail	1.771	
P(T<=t) two-tail	0.004	
t Critical two-tail	2.160	

- i. Prove that the degree of freedom is 13. (2 marks)
- ii. Calculate the value of ***H***.
(Given the standard deviation of the differences, $s_d = 0.410$). (3 marks)
- iii. Using the *p*-value approach, can we conclude that there is a significant difference in the average sprint time before and after the warm-up routine at 1% significance level? (3 marks)

(Total: 25 marks)

QUESTION 3

- a. A nutrition researcher wants to compare the variability of daily calorie intake between participants following two different diet programmes, Diet A and Diet B. Independent random samples of daily calorie intake (in kcal) were collected from 15 participants in each programme. A two-sample F-test for variances at 5% significance level was performed and the summary is given in Table 7 below.

Table 7		
F-Test Two-Sample for Variances		
	Diet A	Diet B
Mean	2150	2095
Variance	<i>J</i>	400
Observations	15	15
df	14	14
F	3.000	
P(F<=f) one-tail	0.004	
F Critical one-tail	2.168	

- i. State one purpose of using an F-test for variances. (1 mark)
- ii. Calculate the value of *J*. (1 mark)
- iii. Prove that the degree of freedom is 14. (1 mark)
- iv. Using the critical value approach, determine if there is a significant difference in the variability of daily calorie intake between Diet A and Diet B at 5% significance level. (3 marks)

- b. A sports scientist wants to determine whether three different training programmes (Programme A, Programme B, Programme C) lead to different average improvements in athletes' vertical jump height (in cm). Each programme was assigned to a group of 12 athletes, and the improvement in jump height after 8 weeks of training was measured. A one-way ANOVA is conducted at a significance level of 5% and the results are shown in Table 8 and Table 9.

Table 8

Anova: Single Factor
SUMMARY

<i>Training Programme</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Programme A	12	480	40.0	15.2
Programme B	12	522	43.5	18.4
Programme C	12	558	46.5	14.7

Table 9

ANOVA						
<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	<i>K</i>	2	<i>M</i>	<i>O</i>	0.021	3.285
Within Groups	<i>L</i>	33	<i>N</i>			
Total	612.0	35				

- Using the Sum of Square Between (SSB) formula, find *K*. Then, determine the values of *L*, *M*, *N* and *O*.
(11 marks)
- State the null and alternative hypotheses for the ANOVA test.
(2 marks)
- Based on the critical value in Table 9, determine whether there is a significant difference in the mean improvements among the three training programmes at $\alpha = 0.05$.
(3 marks)
- State the decision rule for not rejecting the null hypothesis using the *p*-value approach.
(1 mark)
- If only two training programmes were compared, which statistical test would be more appropriate?
(2 marks)

(Total: 25 marks)

QUESTION 4

- a. A city cultural office wants to know whether residents show equal preference for four weekend leisure activities: Parks, Cafes, Museums and Cinema. They surveyed 120 residents and recorded their favourite weekend activity. The data were analyzed using a chi-square test at $\alpha = 0.10$, and the results were summarized in Table 10 and Table 11.

Table 10

Activity	Observed N	Expected N	Residual
Parks	P	30	10
Cafes	35	30	Q
Museums	20	30	-10
Cinema	R	30	-5

Table 11

Pearson Chi-Square	T
p-value	0.040
Critical value	6.291

- i. Determine the values of **P**, **Q** and **R**.
(3 marks)
- ii. Using the Pearson Chi-Square formula, find **T**.
(3 marks)
- iii. State the null and alternative hypotheses for the study.
(2 marks)
- iv. Using the *p*-value in Table 11, determine if there is a significant difference in preference for weekend leisure activities at 10% significance level.
(3 marks)

- b. A transportation analyst is studying whether commute mode is associated with punctuality among office workers. Workers are classified by their primary commute mode (Car, Public Transit, Motorcycle) and whether they typically arrive On-time or Late. A chi-square test was conducted at $\alpha = 0.10$ and the results are summarized in Table 12 and Table 13.

Table 12

			Punctuality	
			On-time	Late
Commute Mode	Car	Observed	18	12
		Expected	U	9.5
	Public Transit	Observed	15	5
		Expected	13.67	6.33
	Motorcycle	Observed	8	2
		Expected	6.83	X

Table 13

Pearson Chi-Square	W
p-value	0.367
Critical value	4.605

- Compute the values of **U** and **X**.
(4 marks)
- Compute the test statistic, **W**.
(7 marks)
- Using the *p*-value in Table 13, determine if there is a significant association between commute mode and punctuality at the 10% significance level.
(3 marks)

(Total: 25 marks)

(TOTAL: 100 MARKS)

END OF QUESTION PAPER

APPENDIX 1(1)

Index Numbers

$$\text{i. Laspeyres Price Index} = \frac{\sum p_t q_0}{\sum p_0 q_0} \times 100$$

$$\text{ii. Paasche Price Index} = \frac{\sum p_t q_t}{\sum p_0 q_t} \times 100$$

$$\text{iii. Simple Aggregate Price Index} = \frac{\sum p_t}{\sum p_0} \times 100$$

$$\text{iv. Weighted Aggregate price Index} = \frac{\sum p_t w}{\sum p_0 w} \times 100$$

where

p_0 = price of the base year

p_t = price of the current year

q_0 = quantity of the base year

q_t = quantity of the current year

w = weights

Confidence Intervals

- i. Mean μ , for large samples

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

- ii. Mean μ , for small samples

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \quad df = n - 1$$

- iii. Test statistic for two population mean, paired observation

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

- iv. Difference in means, $\mu_1 - \mu_2$, for large and independent samples

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{or} \quad (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- v. Difference in means, $\mu_1 - \mu_2$, for small and independent samples

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad df = n_1 + n_2 - 2$$

- vi. Difference in means, $\mu_1 - \mu_2 = \mu_d$ for paired samples

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}} \quad df = n - 1$$

APPENDIX 1(2)

Analysis of Variance (ANOVA)

- i. Degrees of freedom for the numerator = $k - 1$
- ii. Degrees of freedom for the denominator = $n - k$
- iii. Total sum of squares:

$$SST = \sum x^2 - \frac{(\sum x)^2}{n}$$

- iv. Between-samples sum of squares:

$$SSB = \left(\frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + \frac{T_3^2}{n_3} + \dots \right) - \frac{(\sum x)^2}{n}$$

- v. Within-samples sum of squares = $SST - SSB$

- vi. Mean square between samples:

$$MSB = \frac{SSB}{(k - 1)}$$

- vii. Mean square within samples:

$$MSW = \frac{SSW}{(n - k)}$$

- viii. Test statistic for a one-way ANOVA test:

$$F = \frac{MSB}{MSW}$$

Chi-square

- i. Test statistic for Chi-square:

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

- ii. Expected value = $\frac{\text{row sum} \times \text{column sum}}{\text{grand total}}$