UNIVERSITI KUALA LUMPUR
MALAYSIA ITALY DESIGN INSTITUTE

FINAL EXAMINATION

MARCH 2026 SEMESTER

COURSE CODE	:	PPB20203
COURSE TITLE	:	APPLIED STATISTICS
PROGRAMME NAME	:	BACHELOR OF ENGINEERING TECHNOLOGY WITH HONOURS
DATE	:	25 JUNE 2026
TIME	:	2:00PM - 5:00PM
DURATION	:	3 HOURS

INSTRUCTIONS TO CANDIDATES

1. Please read the instructions given in the question paper CAREFULLY.
2. This question paper is printed on both sides of the paper.
3. This question paper consist of TWO sections.
4. Answer ALL questions for Section A.
5. Section B consist of four questions. Answer THREE (3) questions only.
6. Please write your answer on the answer booklet provided.
7. Please answer all questions in English only.
8. Please answer MCQ/EMQ questions using OMR sheet. ☐ *Tick if applicable*
9. Refer to the attached Formula/ Appendies. ☐ *Tick if applicable*

THERE ARE 12 PAGES OF QUESTIONS INCLUDING THIS PAGE

SECTION A (Total: 40 marks)

Answer ALL questions.

Please use the answer booklet provided.

Question 1

The table below shows the number of food delivery orders made by young adults in a week.

Table : Weekly Food Delivery Orders

Orders per week	Frequency
0 – 2	6
3 – 5	12
6 – 8	14
9 – 11	7
12 – 14	4

(a) Estimate the mean and median number of food delivery orders per week.

(5 marks)

(b) Estimate the standard deviation of the number of food delivery orders per week.

(3 marks)

(c) Describe the shape of distribution of the data.

(2 marks)

Question 2

The time (in minutes) taken by employees to complete a data entry task at a corporate training session follows a normal distribution with a mean of 48 minutes and a standard deviation of 6 minutes.

(a) Estimate the probability that a randomly selected employee completes the task between 45 and 51 minutes.

(3 marks)

(b) The company plans to award a fastest performer certificate to the top 10% fastest employees. Estimate the maximum time an employee must take to qualify.

(3 marks)

(c) In a session with 50 employees, estimate the number of employees who would complete the task in more than 55 minutes.

(4 marks)

Question 3

The production manager at a snack packaging company wants to check whether the average weight of their chocolate bars is equal to the labeled weight of 50 grams. The process standard deviation is known to be 2 grams. A random sample of 40 chocolate bars shows a sample mean weight of 49.3 grams.

(a) Construct the null and alternative hypotheses.

(2 marks)

(b) Examine whether there is sufficient evidence to conclude that the mean weight of the chocolate bars differs from the 50 grams at the 5% level of significance.

(5 marks)

(c) Interpret the results and provide a conclusion to the quality control supervisor.

(3 marks)

Question 4

The data in the table below were collected from 7 online retail stores.

(a) Estimate the slope, b and intercept, a of the regression line. Include the regression equation in your solution.

(5 marks)

(b) Interpret the slope of the regression line based on your answer in part (a).

(2 marks)

(c) Use the regression equation to estimate the daily sales revenue for an online store that conducts 6 hours of promotion per day.

(3 marks)

SECTION B (Total: 60 marks)

Answer THREE (3) questions only.

Please use the answer booklet provided.

Question 1

Answer the following questions:

(a) A workplace wellness committee recommends that employees should spend no more than an average of 100 minutes per day on screen time during working hours. The daily screen time is assumed to be normally distributed with a standard deviation of 20 minutes. A study of 50 employees found that the average screen time was 104 minutes per day.

i. Construct the null and alternative hypotheses for testing whether employees screening time is greater than 100 minutes.

(2 marks)

ii. At the 5% level of significance, test whether there is sufficient evidence to conclude that the average daily screen time is greater than 100 minutes. Show all the working steps.

(6 marks)

iii. Interpret the result of the hypothesis test and state a clear conclusion for the workplace wellness committee.

(2 marks)

(b) A manufacturing company has introduced a new machine maintenance strategy to improve equipment operating time. To evaluate its effectiveness, five machines using the new strategy and four machines using the existing strategy were selected at random. The operating time (in years) of each machine before major breakdown is recorded as follows:

New Strategy	3.2	4.9	2.1	5.4	1.8
Old Strategy	2.4	1.1	3.0	2.6	

At the 0.01 level of significance, test whether the new maintenance strategy results in a longer average operating time than the existing strategy.

i. Construct the null and alternative hypotheses.

(2 marks)

ii. Find the critical value and test statistic. Show all steps clearly.

(6 marks)

iii. Based on your answer in (ii), is there sufficient evidence to conclude that the new maintenance strategy results in a longer average operating time than the existing strategy?

(2 marks)

Question 2

Answer the following questions:

(a) A popular coffee chain wants to determine whether the average time taken to serve a cup of coffee meets the target of 3 minutes. The goal is to test whether the mean service time is different from the target. The standard deviation of service times is known to be 0.5 minutes. A random sample of 36 customers recorded an average service time of 3.4 minutes.

i. Construct the null and alternative hypotheses for this study.

(2 marks)

ii. At the 0.05 level of significance, test whether there is evidence that the mean service time is different from 3 minutes. Show all your calculations.

(6 marks)

iii. Interpret the results and provide a clear conclusion to the coffee shop manager.

(2 marks)

(b) A gaming company wants to compare the average number of levels completed in a new mobile game by players using four different in-game boosters:

- Booster A
- Booster B
- Booster C
- No booster (Control)

The ANOVA summary table below is partially filled:

ANOVA

	Sum of Squares	df	Mean Square	F
Between Groups	***A***	3	32.500	***D***
Within Groups	52.000	***B***	***C***	
Total	120.000	19		

i. Find the values of A, B C and D.

(4 marks)

ii. Write the null and alternative hypotheses for this study.

(2 marks)

iii. At the 0.05 significance level, test if there is a difference in average levels completed among the four booster groups by using F-test.

(4 marks)

Question 3

Answer the following questions:

(a) A gaming accessory company wants to determine whether the average battery life of a new wireless gaming mouse is less than 12 hours. A random sample of 8 mice is tested, and the battery life (in hours) is recorded as follows:

11.5, 12.8, 13.2, 11.0, 12.1, 12.5, 11.8, 12.3

i. Construct the null and alternative hypotheses to test whether the mean battery life is less than 12 hours.

(2 marks)

ii. Find the sample mean and sample standard deviation for the battery life data.

(3 marks)

iii. Solve for the t-statistic using the given sample data.

(2 marks)

iv. Find the critical t-value at the 0.05 significance level (df = 7) and interpret the null hypothesis decision.

(3 marks)

(b) A video game company wants to see if gaming platform preference is associated with player age group. A survey of 60 players recorded the following preferences:

Age Group	PC	Console	Mobile	Total
Teenager	12	8	10	30
Adult	6	12	6	24
Senior	2	1	3	6
Total	20	21	19	60

Chi-Square Test Output

Pearson Chi-Square	5.142
Degrees of Freedom	4
p-value	0.274
Critical Value (0.05)	9.488

i. Compute the expected counts for the Teenager row only.

(4 marks)

ii. Construct the null and alternative hypotheses for this study.

(2 marks)

iii. Using the Chi-Square test output, test at the 0.05 significance level whether platform preference is independent of age group. Provide your decision and interpretation.

(4 marks)

Question 4

A sports scientist wants to study the relationship between running speed (km/h) and average heart rate (beats per minute, bpm) in recreational runners. The following data were collected during treadmill tests:

Speed (km/h)	6	8	10	12	14	16	18	20	22
Heart Rate (bpm)	110	118	125	132	140	148	155	162	170

The output of the data is shown in Table Model Summary and Coefficients.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.992[a]	***P***	.980	3.120

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	102	1.50		68.0	.000
	Speed (km/h)	3.10	0.15	0.985	20.7	.000

(a) Find the dependent and independent variables.

(2 marks)

(b) Find the value of the correlation coefficient in Table Model Summary, and then interpret the value obtained.

(3 marks)

(c) Solve for the coefficient of determination, ***P*** and interpret the value obtained.

(4 marks)

(d) Estimate the values of intercept, ***a*** and slope, ***b*** from Table Coefficients. Hence, determine the linear regression equation of running speed and heart rate.

(4 marks)

(e) Explain the meaning of intercept,a and slope,b obtained in (d).

(3 marks)

(f) Predict the average heart rate if a runner runs at 15 km/h. Is the estimation reliable? reliable? Explain.

(4 marks)

END OF EXAMINATION PAPER

Attachment for SECTION A- QUESTION 4

The data in the table below were collected from 7 online retail stores. x represents the number of hours the online store is actively promoted each day, and y represent the average daily sales revenue (in RM).

Table: Promotion Hours and Daily Sales Revenue

x (promotion hours/day)	3	5	7	4	9	6	8
y (daily revenue in RM)	450	720	980	600	1300	850	1120

FORMULA

Mean	$\bar{x} = \frac{\sum x}{n}\ or\ \frac{\sum fm}{\sum f}$
Median	$\tilde{x} = L_m + i\left(\frac{\frac{n}{2} - f}{f_{md}}\right), ML = \frac{n+1}{2}$
Variance	$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}\quad or\quad \frac{\sum m^2 f - \frac{(\sum fm)^2}{\sum f}}{\sum f - 1}$
Standard Deviation	$s = \sqrt{Variance}$
Skewness	$Skew = \frac{Mean - Mode}{Standard\ Deviation} @ \frac{3(Mean - Median)}{Standard\ Deviation}$
Coefficient of Variation	$CV = \frac{Standard\ Deviation}{Mean} \times 100\%$

Probability Distribution

Binomial Distribution	$P(X) = \frac{n!}{(n-X)!\,X!} \cdot p^X \cdot q^{n-X}$
Poisson Distribution	$P(X) = \frac{e^{-\lambda}\lambda^X}{X!}, \lambda = np$
Normal Distribution	$z_{score} = \frac{x - \mu}{\sigma}$

Hypothesis Testing

For one sample test for μ :	$z_{test} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$	$t_{test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim df = n - 1$
For two sample test for μ :		
$t_{test}\ dependent$	$t = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}\quad where$ $\bar{D} = \frac{\sum D}{n}\quad S_D = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n-1}}$	
$t_{test}\ independent$	$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}\quad \sim df = n - 1$	

For more than two samples means, μ :	$F\ test = \frac{MSB}{MSW}\ where;$ $MSB = \frac{SSB}{k-1}\ and\ MSW = \frac{SSW}{n-k}$ $SSB = \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \frac{{T_3}^2}{n_3} + \frac{{T_4}^2}{n_4} + \cdots\right) - \frac{(\sum x)^2}{n}$ $SSW = \sum x^2 - \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \frac{{T_3}^2}{n_3} + \frac{{T_4}^2}{n_4} + \cdots\right)$ $df_n = k-1\ and\ df_d = n-k$
Chi-square test :	$x^2 = \sum \frac{(O-E)^2}{E}\ \ where$ $E = \frac{(Row\ total)(Column\ total)}{sample\ size}$ $df = (R-1)(C-1)\ \ or\ df = k-1 =$

Regression and Correlation

$y = a + bx$	$t = r\sqrt{\frac{n-2}{1-r^2}} \sim df = n-2$
$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$	$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$
$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$	