0044

The Open University of Sri Lanka
Faculty of Engineering Technology
Department of Mathematics and Philosophy of Engineering

Study Programme	:	Bachelor of Industrial Studies Honours
Name of the Examination	:	Final Examination
Course Code and Title	:	**MHZ4357 Applied Statistics**
Academic Year	:	2024/25
Date	:	21st February 2026
Time	:	0930-1230h
Duration	:	3 hours

General instructions

REGISTRATION NO:

.....

- ***Part A is Compulsory***
  - *Provide short answers in the given space.*
  - *Do not need to show any workings.*
- *Answer **five (05)** questions only out of seven (07) questions from **Part B**.*
  - *Provide answers in separate sheets (answer booklet) which will be provided in the examination.*
  - *Show all your workings.*
- *The number of pages in the paper is **Eight (08)**.*
- *All the symbols are in standard notation unless they are defined.*
- *All the relevant statistical tables and equations are provided.*
- *You are allowed to use a non-programmable calculator during the examination.*
- ***Attach Part A to the Answer script of Part B.***
- *Answer for each question should commence from a new page.*
- *This is a Closed Book Test (CBT).*

Part A

Provide short answers in the given space. Do not need to show any workings or work. Please attach this part to your answer script.

1. State whether the following variables are discrete or continuous:

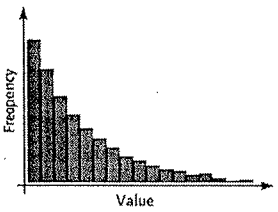
   (i) Number of machine breakdowns per month:.....

   (ii) Time taken to complete a production cycle :.....

2. What happens to the standard deviation if all observations in a dataset are multiplied by 5?
   Answer:.....

3. State the shape of the distribution represented by the following histogram.

   Answer: .....



4. If X ~ Bin(50, 0.3), find mean of X.

   Answer:.....

5. If A and B are mutually exclusive events with P(A)=0.4 and P(B)=0.5, find P(A ∪ B).

   Answer: .....

6. If the probability of selecting a defective fruit is 0.2, the probability of selecting a non-defective fruit is,

   Answer: .....

7. What is the meaning of the level of significance in hypothesis testing?

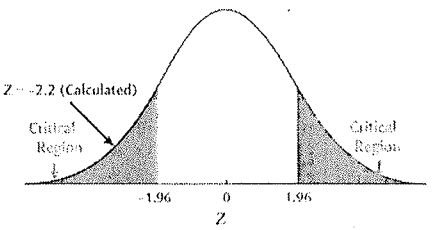
   Answer: .....

8. Find the missing value k: Mean of 5, 7, k, 9, 11 is 8.

   Answer: .....

9. State the decision that can be made according to the position of calculated and critical Z values as displayed in the graph when doing hypothesis testing.

   Answer: .....



10. In hypothesis testing, the probability of failing to reject H_0 when it is false is called,

    Answer: .....

Part B

Q1.

(a) Consider the following extract from a crop quality inspection log maintained at Green Fields Agriculture (Pvt) Ltd.

Crop Quality Inspection Log

lot No	Crop Type	Variety / Colour	Number of Defects	Defect Severity
P100	Rice	White	3	Minor
P101	Maize	Yellow	2	Minor
P102	Tomato	Red	6	Major
P103	Beans	Green	0	None
P104	Potato	Brown	4	Moderate

(i) Determine the type of data and scale of the measurement associated with each of the following variables in the above quality inspection log.

(12 marks)

	Type of data	*Scale of measurement*
Crop Type	.....	.....
Variety/Color	.....	.....
Number of Defects	.....	.....

(ii) Suppose you are the quality control officer at Green Fields Agriculture (Pvt) Ltd. Based on the defect severity levels recorded, propose a simple coding scheme to quantify defect severity for further statistical analysis.

(05 marks)

Defect Severity	*Suggested Code*
None	.....
Minor	.....
Moderate	.....
Major	.....

(iii) Then identify the scale of measurement associated with the coded variable in part (ii) and suggest the best statistical summary for central tendency to use on this variable.

(08 marks)

(b) A large-scale agricultural farm aims to analyze fertilizer waste generated in three cultivation zones (Zone A, Zone B, and Zone C) during one growing week.

The farm uses two types of fertilizers:

- Organic Fertilizer (Eco-friendly Compost)
- Chemical Fertilizer (Synthetic Nitrogen Mix)

At the end of the week, supervisors recorded the fertilizer waste (in kilograms).

Zone	Organic Fertilizer Waste (Kg)	Chemical Fertilizer Waste (Kg)
A	12	24
B	15	18
C	10	30

The farm management plans to reduce chemical fertilizer waste by 25% to minimize soil and water pollution.

(i) Draw a suitable graph to compare waste across zones.

(15 marks)

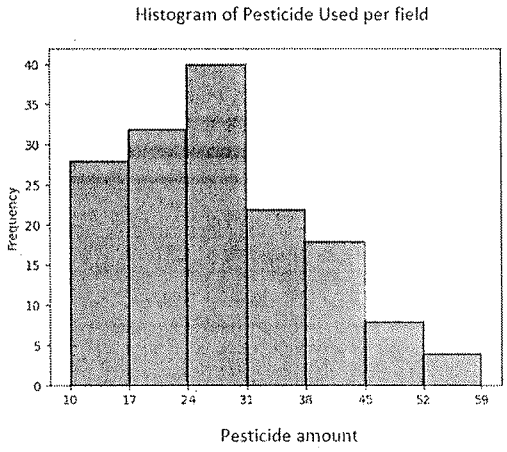
(ii) Explain the differences in waste production among each zone.

(10 marks)

(iii) Identify which zone must reduce the most chemical fertilizer waste.

(15 marks)

(c) An agricultural plantation monitored the amount of pesticide used per field section during one week. A sample of 152 field sections was recorded and summarized in a histogram.



(i) Comment about the shape of the distribution of the data.

(10 marks)

(ii) Draw the frequency polygon corresponding to the histogram.

(15 marks)

(iii) Determine the Total Number of Field Sections That Used Between 24 and 45 Liters. **(10 marks)**

Q2.

At Green Harvest Farms, inspectors recorded the number of pest damage spots per 100 plants from 21 different crop plots during a weekly field inspection. Each value represents the number of pest-damaged plants per plot. The stem represents the tens digit, and the leaf represents the units digit.

Stem	*Leaf*
0	1 1 2 2 3 4
1	0 1 1 1 2 5 5
2	3 3 4
3	5 6
4	8
5	9
6	9

(a) Find the Central Tendency measures (Mean, Median, Mode) of the given data.

(25 marks)

(b) The farm management wants to divide the plots into **four equal performance groups**: Best performing plots (lowest pest damage), Lower-middle group, Upper-middle group, Worst performing plots (highest pest damage)

Find:

i. First Quartile (Q_1)

ii. Second Quartile (Median, Q_2)

iii. Third Quartile (Q_3)

(30 marks)

(c) Obtain the inter quartile range of the number of pest damage spots on a plant by using the results obtained in part (b). Interpret what this tells about the spread of pest damage levels across plots.

(10 marks)

(d) Construct a box and whisker plot to represent the given data, clearly indicating the presence of any outliers. **(25 marks)**

(e) State, giving a reason for which measure of central tendency, you would recommend summarizing this data. **(10 marks)**

Q3. A researcher studying pest infestation in a maize field defines the random variable X = number of infected plants observed in a selected plot.
The probability distribution of X is given by:

$$P(X = x) = c(x + 2), \qquad x = 0,1,2,3$$

where c is a constant.

(a) Find the value of c such that the function becomes a valid probability distribution.

(15 marks)

(b) Find the probability that:
(i) At most 2 plants are infected
(ii) Exactly 3 plants are infected

(20 marks)

(c) Find the expected number of infected plants per plot.

(25 marks)

(d) Find the variance of the number of infected plants per plot.

(25 marks)

(e) Find the probability that the number of infected plants lies between 1 and 3.

(15 marks)

Q4.

(a) Discuss the characteristics of Geometric distribution and the characteristics of Binomial distribution. **(20 marks)**

(b) A seed quality inspection unit tests seeds to identify damaged seeds before packaging. Previous studies show that the probability that a seed is damaged is 0.03.

(i) Find the probability that the first damaged seed is found at the 4th seed inspected. **(20 marks)**

(ii) Find the probability that more than 2 seeds must be inspected before finding the first damaged seed. **(20 marks)**

(iii) Suppose the same inspection process is applied to a batch of 25 seeds. Find the probability that 12% of the seeds in the batch are damaged.

(40 marks)

Q5.

(a) A research farm studies the percentage shrinkage of soil moisture content after the first drying cycle to evaluate soil quality for paddy cultivation. A random sample of ten soil plots was tested, and the shrinkage percentages were recorded as follows:

Shrinkage (%) readings	2.1	1.8	2.4	2.0	1.9	2.3	2.2	1.7	2.0	2.1

(i) Determine the best estimate for the population mean shrinkage.

(25 marks)

(ii) Obtain the best estimate for population variance. **(25 marks)**

(b) An agricultural research institute evaluates the average tomato yield per plant (measured in kilograms) under a new organic fertilizer treatment. A random sample of 120 tomato plants was selected from the experimental field. The findings revealed that the sample mean yield ($\bar{x}$) was 85 kg and the Sample standard deviation (s) was 4 kg.

(i) Find the standard error of the estimation of population mean.

(15 marks)

(ii) Obtain the interval estimate for the population mean at 5% significant level. **(30 marks)**

(iii) Interpret your result. **(05 marks)**

Q6. A fertilizer company claims that a new organic fertilizer formula reduces the average grain weight of harvested paddy without affecting crop quality, making storage and transport more efficient. The traditional method produces paddy with an average grain weight of 1200 g per sample batch, with a population standard deviation known to be 60 g. To test this claim, a random sample of 150 paddy batches treated with the new fertilizer is examined. The average grain weight is found to be 1185 g.

(a) State the null hypotheses and alternative hypotheses for testing the claim.

(20 marks)

(b) Find the test statistics associated with this hypothesis testing.

(30 marks)

(c) At the 1% significance level, determine whether there is sufficient evidence to support the claim that the new fertilizer reduces average grain weight. **(30 marks)**

(d) State your conclusion regarding the hypothesis and comment on the claim. **(20 marks)**

Q7.

An agricultural scientist investigates the relationship between fertilizer application rate X (kg/ha) and crop yield Y (quintals per hectare) for a maize cultivation experiment. The following experimental data were obtained:

Fertilizer (kg/ha)	**120**	**130**	**140**	**150**	**160**	**170**
Yield (q/ha)	68	72	75	78	80	83

You may use the following summary statistics:

$$\sum X = 870, \sum Y = 456, \sum X^2 = 126{,}700, \sum Y^2 = 34{,}722, \sum XY = 66{,}840$$

(a) Draw a scatter diagram and comment on the form and strength of the relationship between fertilizer application and crop yield. **(10 marks)**

(b) Calculate the correlation coefficient for the above data and interpret it. **(20 marks)**

(c) Determine the least squares estimates of the regression coefficients β_0 and β_1 for the model: $Y = \beta_0 + \beta_1 X$ **(30 marks)**

(d) Write down the estimated regression equation and interpret the slope coefficient in an agricultural context. **(10 marks)**

(e)

(i) Calculate the predicted crop yield when fertilizer application is 155 kg/ha. **(10 marks)**

(ii) If the observed yield at 155 kg/ha is 77 quintals per hectare, calculate the residual and comment on its sign. **(10 marks)**

(f) Explain why the fitted regression model may not be reliable for predicting crop yield when fertilizer application is 220 kg/ha. **(10 marks)**

Important Formulars:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

Binomial distribution

$$Pr(X = r) = \frac{n!}{(r!(n-r)!)} p^r q^{(n-r)}$$

Geometric distribution

$$Pr(X = r) = pq^{(r-1)}$$

Poison Distribution

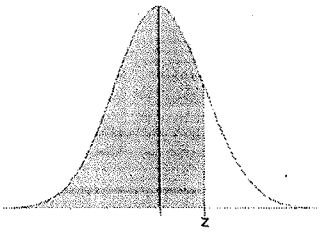
$$Pr(X = r) = e^{(-\pi)} \frac{\mu^r}{r!}$$

If random variable X, follows normal distribution ($X \sim N(\mu, \sigma^2)$) Then $Z = \frac{(X-\mu)}{\sigma}$ follows standard normal distribution where $Z \sim (1, 0)$.

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

$$\beta_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

Standard Normal Cumulative Probability Table



Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998