University of Ruhuna

Bachelor of Science General/Special Degree

Level III (Semester I) Examination – November 2021

Subject: Industrial Mathematics/Applied Mathematics

Course Unit: IMT313β/AMT314β/MSP316β (Applied Statistics II)   Time: Two (02) Hours

Answer all the questions

Calculators are provided.

1.

1.1. The weights of loaves from a certain bakery are normally distributed with mean 500g and standard deviation 20g.

  1.1.1. Determine what percentage of the output would fall below 475g.

(10 marks)

  1.1.2. The bakery produces 1000 loaves daily at a cost of Rs. 40.00 per loaf and can sell all those above 475g at Rs. 50.00 each but not allowed to sell the rest. Calculate the expected daily profit.

(10 marks)

  1.1.3. A random sample of 25 loaves are selected from the production and the sample mean is 490g. Test at the 5% level of significance, whether the population mean weight is less than 500g.

(30 marks)

1.2. A company recently opened two supermarkets in two different areas. The management wants to know if the mean sales per day for these two supermarkets are different. A sample of 10 days for Supermarket A produced the following data on daily sales (in million rupees).

47.56 57.66 51.23 58.29 43.71 49.33 52.35 50.13 47.45 53.86

A sample of 12 days for Supermarket B produced the following data on daily sales (in million rupees).

56.34 63.55 61.64 63.75 54.78 58.19 55.40 59.44 62.33 67.82 56.65 67.90

Assume that the daily sales of the two supermarkets are both normally distributed with equal but unknown standard deviations.

  1.2.1. Construct a 95% confidence interval for the difference between the mean daily sales for these two supermarkets.

(30 marks)

  1.2.2. Test at the 5% level of significance whether the mean daily sales for these two supermarkets are different.
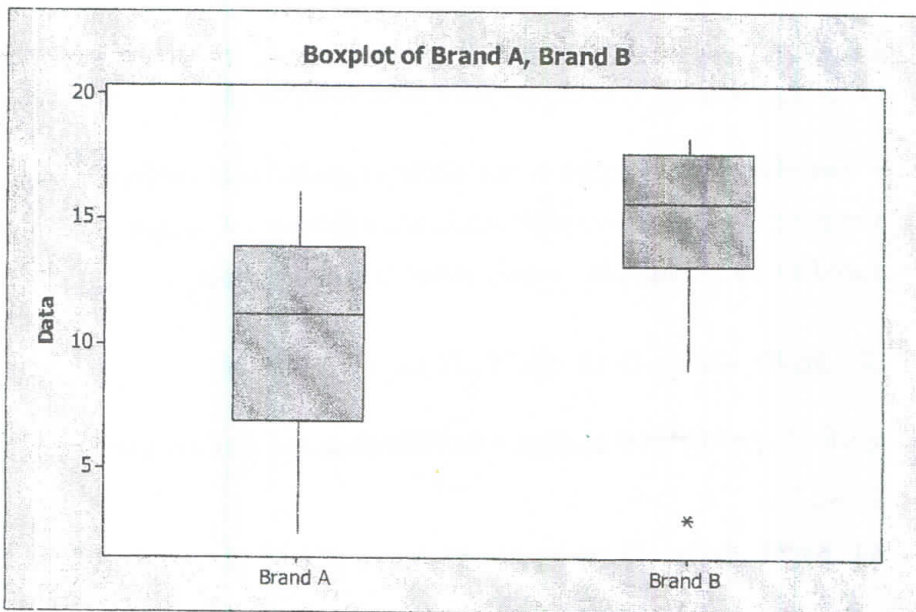
(20 marks)

2.

2.1. Suppose we want to compare the mean lifetime of two kind of 9-volt batteries on the basis of the following lifetime (in hours):

Brand A : 6.9 11.2   14.0   13.2   9.1   13.9   16.1   9.3   2.4   6.4   11.5

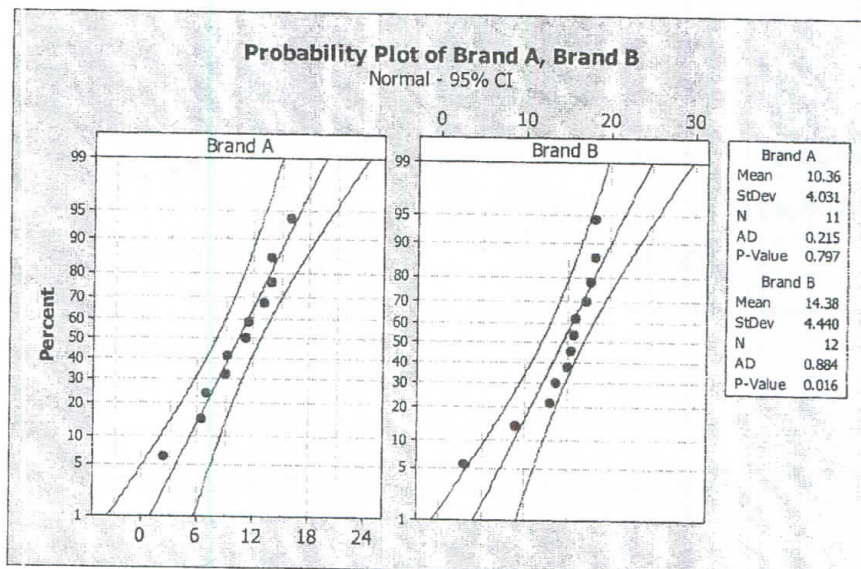Brand B : 15.5 3.0   16.0   15.8   18.2   13.7   18.3   9.0   17.2   17.8   13.0
          15.1

2.1.1. Based on the side-by-side boxplots of the two samples, comment on the shapes of the two distributions.

(05 marks)

Boxplot of Brand A, Brand B

2.1.2. Based on the normal probability plots of two samples, does normality seem reasonable in either distribution?

(05 marks)

Probability Plot of Brand A, Brand B
Normal - 95% CI

Brand A	
Mean	10.36
StDev	4.031
N	11
AD	0.215
P-Value	0.797

Brand B	
Mean	14.38
StDev	4.440
N	12
AD	0.884
P-Value	0.016

2.1.3. Based on your results in parts 2.1.1 and 2.1.2, complete the most suitable parametric or non-parametric test to test the hypothesis that there is no difference in the mean lifetime of the two Brands of batteries at the 5% level of significance.

(40 marks)

2.2. The number of defects in printed circuit boards is hypothesized to follow a Poisson distribution. A random sample of 60 printed boards has been collected, and the following number of defects observed.

Number of defects	Observed frequency
0	32
1	15
2	9
3	4

2.2.1. Show that the sample mean is 0.75.

(10 marks)

2.2.2. Based on the above 60 observations, is a Poisson distribution with a mean 0.75 an appropriate model? Perform a goodness of fit test at the 5% level of significance.

(40 marks)

3.

3.1. The yields of tomato plants grown using different types of fertilizer are given in the table below:

	Yield (kg)				
Fertilizer X	3.5	4.0	3.8	4.1	4.4
Fertilizer Y	4.7	5.0	4.5	5.3	4.6
Fertilizer Z	3.6	3.9	4.2	4.1	4.0

The analysis of variance table is given for the above data with a few missing values:

One-way ANOVA: Fertilizer X, Fertilizer Y, Fertilizer Z

Source of Variation	DF	Sum of Square	Mean Square	F value
Treatment	2	2.4653	b	d
Error	a	1.0920	c	
Total	14			

3.1.1. State the assumptions you need to make in order to carry out an analysis of variance on these data.

(15 marks)

3.1.2. State the null and alternative hypotheses to determine whether the fertilizers produce different yields.

(10 marks)

3.1.3. Find the missing values a, b, c and d.

(10 marks)

3.1.4. What is the F ratio? Should the null hypothesis be rejected at the 5% level of significance?

(10 marks)

3.1.5. State your conclusion.

(05 marks)

3.2. A sample of 500 male registered voters showed that 57% of them voted in the last presidential election. Another sample of 400 female registered voters showed that 55% of them voted in the same election.

3.2.1. Construct a 95% confidence interval for the difference between the proportions of all male and all female registered voters who voted in the last presidential election.

(30 marks)

3.2.2. Test at the 5% significance level whether the proportion of all male voters who voted in the last presidential election is different from that of all female voters.

(20 marks)

4. The following data were obtained in a study of the relationship between the weight and chest size of infants at birth:

Weight in kg (x)	2.76	2.17	5.53	4.31	2.30	3.70
Chest size in cm (y)	29.5	26.3	36.6	30.5	28.3	28.6

4.1. Assuming the linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ with $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma^2$, obtain the least squares estimators for the model parameters β_0 and β_1.

(20 marks)

4.2. Obtain the least squares estimates of β_0 and β_1.

(20 marks)

4.3. Give a brief interpretation of the estimate value of β_1

(05 marks)

4.4. Predict the mean chest size for a weight of an infant of 3.0kg.

(05 marks)

4.5. Can the least-squares line be used to predict the chest size for a weight of 2.0kg? If so, predict the chest size. If not, explain why not.

(05 marks)

4.6. Compute the standard deviation of errors.

(10 marks)

4.7. Test the hypothesis that $\beta_1 = 0$ at the 5% level of significance. Do the results of this test indicate that a linear trend is significant?

(25 marks)

4.8. Calculate the coefficient of determination and interpret the value.

(10 marks)