Part A Revision Questions

Q1. Theory for Inference Procedures (3 Marks)

Answer the three short questions. Each correct answer will be awarded 1 mark.

i. What is a $p-$value?

ii. Briefly describe how $p-$value is used in hypothesis testing

iii. What is meant by a Type I error?

iv. What is meant by a Type II error?

Q2. Normal Distribution (6 Marks)

Assume that the diameter of a critical component is normally distributed with a Mean of 50mm and a Standard Deviation of 2mm. You are required to estimate the approximate probability of the following measurements occurring on an individual component.

i. (2 Marks) Between 50 and 51.2mm

ii. (2 Marks) Less than 48.5 mm

iii. (2 Marks) Between 48.2 and 51.6 mm

Use the normal tables to determine the probabilities for the above exercises. You are required to show all of your workings.

Q3. Dixon Q Test For Outliers (4 Marks)

The typing speeds for one group of 12 Engineering students were recorded both at the beginning of year 1 of their studies. The results (in words per minute) are given below:

121	146	150	149	142	170	153
137	161	156	165	137	178	159

Use the Dixon Q-test to determine if the lowest value (121) is an outlier. You may assume a significance level of 5%.

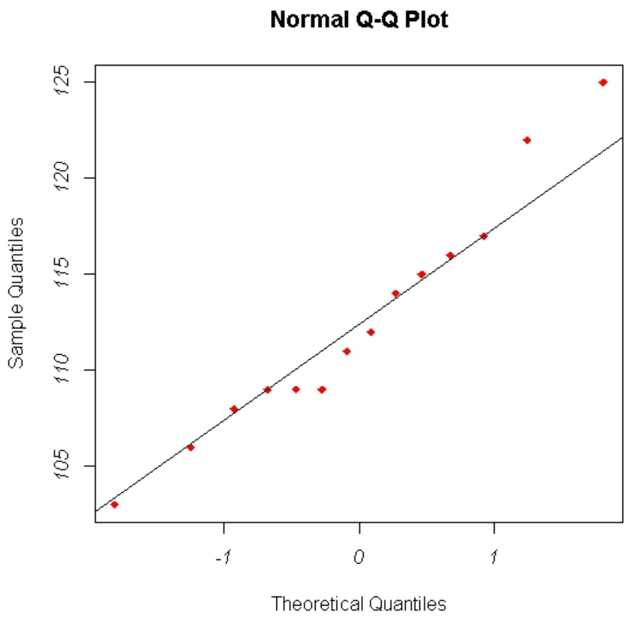
i. (1 Mark) Formally state the null hypothesis and the alternative hypothesis.

ii. (1 Mark) Compute the Test Statistic.

iii. (2 Mark) By comparing the Test Statistic to the appropriate Critical Value, state your conclusion for this test.

Q4. Testing Normality (2 Marks)

A graphical procedure was carried out to assess whether or not this assumption of normality is valid for data set `Y`. Consider the Q-Q plot in the figure below.



i. (1 Mark) Provide a brief description on how to interpret this plot.

ii. (1 Mark) What is your conclusion for this procedure? Justify your answer.

Q5. Testing Normality (3 Marks)

Consider the following inference procedure performed on data set X.

```
> shapiro.test(X)

Shapiro-Wilk normality test

data:  X
W = 0.9619, p-value = 0.6671
```

i. (1 Mark) Describe what is the purpose of this procedure.

ii. (1 Mark) What is the null and alternative hypothesis?

iii. (1 Mark) Write the conclusion that follows from it.

Q6. Testing For Outliers (6 Marks)

(i) (3 Marks) Provide a brief description for three tests from the family of Grubb's Outliers Tests. Include in your description a statement of the null and alternative hypothesis for each test, any required assumptions and the limitations of these tests.

(ii) (3 Marks) Showing your working, use the Dixon Q Test to test the hypothesis that the maximum value of the following data set is an outlier.

$$19, 22, 23, 24, 25, 26, 29, 38$$

Q7. Testing for Outliers (3 Marks)

The following statistical procedure is based on this dataset.

6.98	8.49	7.97	6.64
8.80	8.48	5.94	6.94
6.89	7.47	7.32	4.01

```
> grubbs.test(x, two.sided=T)

Grubbs test for one outlier

data:  x
G = 2.4093, U = 0.4243, p-value = 0.05069
alternative hypothesis: lowest value 4.01 is an outlier
```

i. (1 Mark) Describe what is the purpose of this procedure. State the null and alternative hypothesis.

ii. (1 Mark) Write the conclusion that follows from it.

iii. (1 Mark) State any relevant assumptions for this procedure.

Q8. Numeric Transformation of Data

(i) Describe the purpose of transformations

(ii) Describe the process of transformations

(iii) Describe the purpose of Tukey's Ladder (referencing direction and relative strength)

(iv) Give an example of a transformation for various types of skewed data (use Tukey's Ladder, with an example for both directions)

(v) Describe the limitations of transformations

Q9. - Control Limits

Exam Paper Formulas for Control Limits

- Process Mean

$$\bar{\bar{x}} \pm 3\frac{\bar{s}}{c_4\sqrt{n}}$$

- Process Standard Deviation

$$\bar{s} \pm 3\frac{c_5\bar{s}}{c_4}$$

- Process Range

$$\left[\bar{R}D_3, \bar{R}D_4\right]$$

A normally distributed quality characteristic is monitored through the use of control charts. These charts have the following parameters. All charts are in control.

	LCL	Centre Line	UCL
$\bar{X}$-Chart	542	550	558
R-Chart	0	8.236	16.504

(i.) (2 marks) What sample size is being used for this analysis?

(ii.) (2 marks) Estimate the mean of the standard deviations $\bar{s}$ for this process.

(iii.) (2 marks) Compute the control limits for the process standard deviation chart (i.e. the s-chart).

Q10. Nelson Rules for Control Charts

The **Nelson Rules** are a set of eight decision rules for detecting "out-of-control" or non-random conditions on control charts. These rules are applied to a control chart on which the magnitude of some variable is plotted against time. The rules are based on the mean value and the standard deviation of the samples.

(i) (4×3 Marks) Discuss any four of these rules, and how they would be used to detect "out of control" processes. Support your answer with sketch.

In your answer, you may make reference to the following properties of the Normal Distribution. Consider the random variable X distributed as

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

where μ is the mean and σ^2 is the variance of an random variable X.

- $\Pr(\mu - 1\sigma \leq X \leq \mu + 1\sigma) = 0.6827$
- $\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9545$
- $\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0.9973$