MATH1019 LINEAR ALGEBRA AND STATISTICS FOR ENGINEERS
Lab Test (a)
Semester 2, 2017 solutions

Marks: /20

Declaration: *I hereby undertake not to discuss or divulge the content or format of the test paper with any other person until all tests have been written, and declare that I have no prior knowledge of the contents of the test paper.*

I unconditionally accept any action that may be taken should Curtin University consider that an infringement of the statute No. 10 of the Student Disciplinary Statute has occurred.

Name: **ID**: **Signature**: **Date**: / / 2017

Instructions: Open a new *Word* document and save it in your *I*-drive directory as *Your Name Your ID*, for example: *Joe Smith 2456892.docx*. Your test solutions will be stored in this document. Some questions will require you to copy commands/output from *R* and paste it into this document. Once you have completed the test, save this document and **email it to your tutor**. Return this test paper with your name, student ID, signature and date to your tutor before you leave the room.

Question 1

Load the data set **Loblolly** into *R* using the command **data("Loblolly")**. Answer the following questions with reference to the variable **height**. **Use R for all steps**.

a) Obtain the five number summary. Identify the five numbers as min, max, etc. [3 marks]

```
> fivenum(Loblolly$height)
[1]  3.460 10.455 34.000 51.395 64.100
```

1 mark

The numbers correspond to minimum, Q1, Median, Q2, Maximum. Note that the summary command can also be used and may produce slightly different quartiles. Full marks should still be awarded.

2 marks

b) What is the inter-quartile range?
IQR=Q3-Q1 = 40.94

1 mark

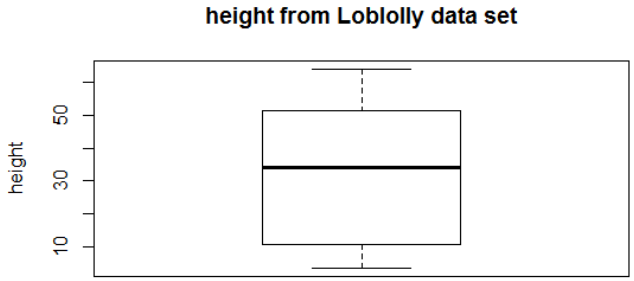
```
> 51.395-10.455
[1] 40.94
```

[1 mark]

c) Obtain a box plot for the variable with appropriate labels and paste it into your document. Are there any outliers? [3 marks]

```
> boxplot(Loblolly$height, main="height from Loblolly data set", ylab="hei
ght")
```

1 mark



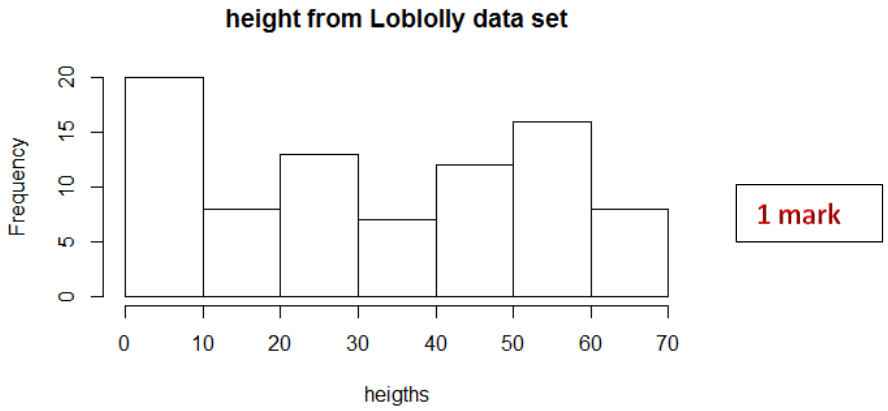
1 mark

There are no outliers. 1 mark

d) Obtain a histogram for the variable with appropriate labels and paste it into your document. [2 marks]

```
> hist(Loblolly$height, main="height from Loblolly data set", xlab="heigth
s")
```

1 mark



1 mark

e) Find the 90% confidence interval for the variable assuming that σ is known and is equal to s. Paste your *R* commands into the document. [3 marks]

```
> sd(Loblolly$height)
[1] 20.6736
```

Assuming σ=s=20.6736

Find the sample size n, one of two ways:

```
> dim(Loblolly)
[1] 84  3
> length(Loblolly$height)
[1] 84
```

So, n=84.

90% CI: $\bar{x} \pm z_{0.05} \frac{\sigma}{\sqrt{n}}$

```
> mean(Loblolly$height)
[1] 32.3644
```

$z_{0.05}$=1.644854

1 mark

Two ways of doing this:

```
> qnorm(0.05,lower.tail = FALSE)
[1] 1.644854
```

or

```
> qnorm(0.95)
[1] 1.644854
```

This can be done using the following statements:

```
> mean(Loblolly$height)-qnorm(0.05,lower.tail = FALSE)*sd(Loblolly$height)
/sqrt(length(Loblolly$height))
[1] 28.65415
> mean(Loblolly$height)+qnorm(0.05,lower.tail = FALSE)*sd(Loblolly$height)
/sqrt(length(Loblolly$height))
[1] 36.07466
```

So the 90% CI is (28.65415, 36.07466) 2 marks

f) True or False: "The probability that the mean lies in the 90% confidence interval you calculated above is 0.9." [1 mark]

False. 1 mark

The confidence level refers to the probability that the method will give the correct answer.

Question 2

If X is a Normally distributed random variable with $\mu = 25$ and $\sigma = 6$, calculate the following using *R*. Paste *R* commands into your document:

a) $P(18 < X < 27)$ [2 marks]

```
> pnorm(27,mean=25,sd=6)-pnorm(18,mean=25,sd=6)
[1] 0.5088862
```

2 marks

b) Find k such that $P(X < k) = 0.7352$. [2 marks]

```
> qnorm(0.7352,mean=25,sd=6)
[1] 28.7717
```

2 marks

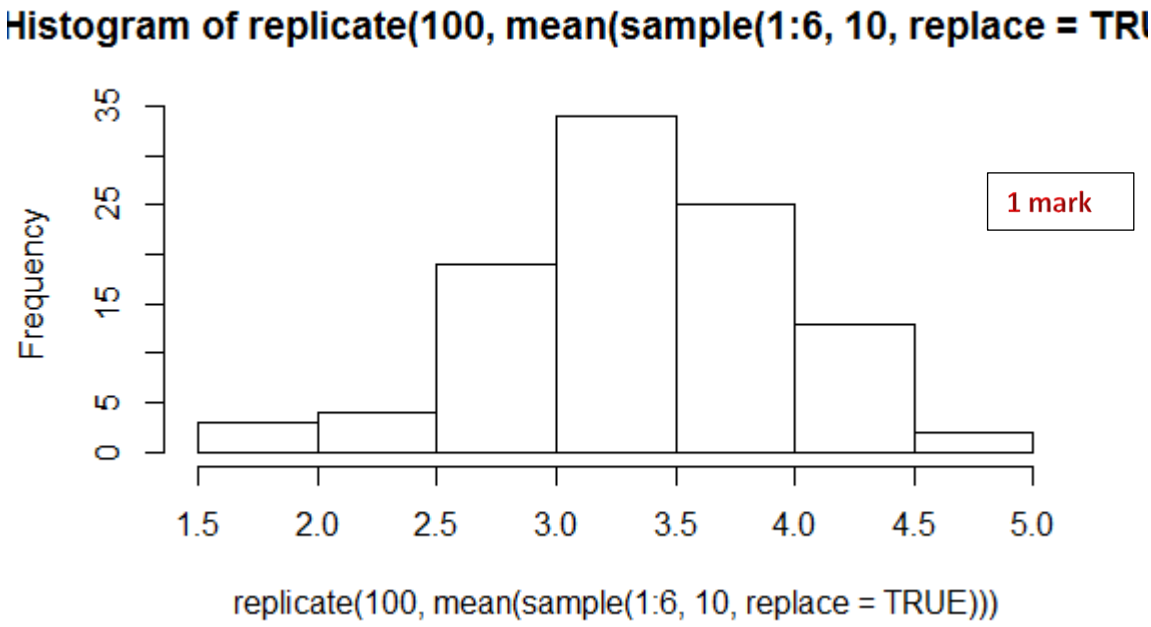
Therefore, k = 28.7717

Question 3

Generate 100 means for samples of size 10 from the digits 1 to 6. Plot your results using a histogram. What do you notice? Paste the histogram into your document. [3 marks]

```
> hist(replicate(100,mean(sample(1:6,10, replace=TRUE))))
```

1 mark



1 mark

The histogram is roughly symmetric. Approximately normal.

1 mark