**Lab 2**

**1.1 Binomial Distribution**

**Questions**

1. Read Concrete\_Data.csv

> concrete\_data <- read.csv(“Concrete\_Data.csv”)

1. Use dim() to determine the dimensions of the concrete data (the number of rows and columns)

> dim(concrete\_data)

[1] 1030 9

1. Use head() and tail() to view the first few and last few rows, respectively, of the concrete data set

> head(concrete\_data)

...  
> tail(concrete\_data)

...

1. Produce a Five-Number Summary of the comprehensive strength of concrete

> summary(concrete\_data$Concrete\_comprehensive\_strength)

Min. 1st Qu. Median Mean 3rd Qu. Max.

2.33 23.71 34.45 35.82 46.13 82.60

1. Plot a histogram of the comprehensive strengths of concretes. Add an appropriate title and - and -axis labels. Note, comprehensive strength is measured in MPa

> hist(concrete\_data$Concrete\_compressive\_strength, main="Comprehensive Strengths of Concretes", xlab="Comprehensive Strength (MPa)", ylab="Number of Concretes")

...

1. Produce a boxplot of comprehensive strengths of concretes. Add an appropriate title and -axis label

> boxplot(concrete\_data$Concrete\_compressive\_strength, horizontal = TRUE, main="Comprehensive Strengths of Concretes", xlab="Comprehensive Strength (MPa)")

...

**Lab 3**

**1.1 Binomial Distribution**

**Questions**

1. Simulate tossing a coin 1000 times. Are the results what you would expect?

> tosses <- rbinom(n=1000, size=1, prob=0.5)

1. Suppose that items are to be inspected from one production line and items are to be inspected from another production line. Let The probability of a defective from line 1 and The probability of a defective from line 2. Let be a Binomial Random Variable with parameters and . Let be a Binomial Random Variable with parameters and . A variable of interest is , which is the total number of defective items observed in both production lines. Let . Use simulation to see how the distribution of will behave. Useful information could be obtained by looking at the histogram of s generated and also considering the sample mean and the sample variance. In your simulation use the following random variables and : is Binomial with and ; and Y is Binomial with and .

> x <- rbinom(n=1000, size=7, prob=0.2)

> y <- rbinom(n=1000, size=8, prob=0.6)

> w <- x + y

**1.3 Normal Distribution**

**Questions**

If is a Normally distributed random variable with = 20 and = 5, calculate the following:

> pnorm(q=15, mean=20, sd=5)

[1] 0.1586553

pnorm(q=23, mean=20, sd=5) - pnorm(q=14, mean=20, sd=5)

[1] 0.6106772

1. Find the value of such that

qnorm(p=0.9345, 20, sd=5)

[1] 27.55085

**2017 Semester 2 Lab Quiz**

**Questions**

Question 1

Load the dataset Loblolly into R by executing the command data(“Loblolly”). Answer the following questions in reference to the variable height.

1. Obtain the Five Number Summary.

> fivenum(Loblolly$height)

[1] 3.460 10.455 34.000 51.395 61.100

The numbers correspond to minimum, Q1, median, Q3 and maximum

1. Determine the inter-quartile range.

> 51.395 – 10.455

[1] 40.94

1. Obtain a box plot for the variable and check if there are any outliers

> boxplot(Loblolly$height, horizontal=TRUE, main="Loblolly Height Attribute", xlab="Height")

...

There are no outliers

1. Obtain a histogram for the variable.

> hist(Loblolly$height, main="Loblolly Height Attribute", xlab="Height")

...

There are no outliers

1. Find the 90% confidence interval for the variable assuming that is known and is equal to .

> x\_bar <- mean(Loblolly$height)

> s <- sd(Loblolly$height)

> z <- qnorm(p=0.95)

> n <- length(Loblolly$height)

> lower\_bound <- x\_bar - z \* (s / sqrt(n))

> lower\_bound

[1] 28.65415

> upper\_bound <- x\_bar + z \* (s / sqrt(n))

> upper\_bound

[1] 36.07466

Therefore, the 90% CI is (28.65415, 36.07466)

1. True or false: “The probability that the mean lies in the 90% CI is 0.9.”

False

Question 2

If is a normally distributed random variable with = 25 and = 6, calculate the following using R:

> pnorm(q=27, mean=25, sd=6) - pnorm(q=18, mean=25, sd=6)

[1] 0.5088862

Therefore, = 0.5088862

1. Find the value of k such that = 0.7352

> qnorm(p=0.7352, mean=25, sd=6)

[1] 28.7717

Therefore, = 28.7717

Question 3

Generate 100 means for samples of size 10 from the digits 1 to 6. Plot your results using a histogram. Note your observations

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

...

The histogram is roughly symmetric and approximately normal.