

## R Homework Number 8

Fisher Ankney

November 8th, 2018

Statistics 5193

\*Note: this document was created using R Markdown.

### Question 1a.

```
help(source)
setwd("~/Documents/data_science/r_stat_5193/scripts")
source("Day8Script.R")
roster
ls()
```

```
##      Firstname  Lastname Math Science English score grade
## 6      Cheryl   Cushing  512      85      28  0.35      C
## 1        John    Davis   502      95      25  0.56      B
## 9        Joel   England  573      89      27  0.70      B
## 4        David    Jones   358      82      15 -1.16      F
## 8         Greg     Knox   625      95      30  1.34      A
## 5      Janice Markhammer  495      75      20 -0.63      D
## 3 Bullwinkle     Moose   412      80      18 -0.86      D
## 10       Mary    Rayburn   522      86      18 -0.18      C
## 2      Angela   Williams   600      99      22  0.92      A
## 7      Reuven    Ytzrhak   410      80      15 -1.05      F
```

```
## [1] "col"      "English"  "First"    "Firstname" "Last"
## [6] "Lastname" "Math"     "mu"       "mvn.data"  "n"
## [11] "name"     "rho"      "roster"   "Science"   "score"
## [16] "Sigma"    "Student"  "sums"     "t.data"    "x"
## [21] "y"        "z"
```

### Question 1b.

```
score
## [1] 0.56 0.92 -0.86 -1.16 -0.63 0.35 -1.05 1.34 0.70 -0.18

roster$score
## [1] 0.35 0.56 0.70 -1.16 1.34 -0.63 -0.86 -0.18 0.92 -1.05

"roster <- roster[order(Lastname,Firstname),]"

## [1] "roster <- roster[order(Lastname,Firstname),]"
```

### Question 1c.

```
y <- quantile(score, c(.75,.25))

roster$`Easy Grade`[roster$score > y[1]] <- "A"
roster$`Easy Grade`[roster$score < y[1] & roster$score >= y[2]] <- "B"
roster$`Easy Grade`[roster$score < y[2]] <- "C"

roster
```

##	Firstname	Lastname	Math	Science	English	score	grade	Easy Grade
## 6	Cheryl	Cushing	512	85	28	0.35	C	B
## 1	John	Davis	502	95	25	0.56	B	B
## 9	Joel	England	573	89	27	0.70	B	A
## 4	David	Jones	358	82	15	-1.16	F	C
## 8	Greg	Knox	625	95	30	1.34	A	A
## 5	Janice	Markhammer	495	75	20	-0.63	D	B
## 3	Bullwinkle	Moose	412	80	18	-0.86	D	C
## 10	Mary	Rayburn	522	86	18	-0.18	C	B
## 2	Angela	Williams	600	99	22	0.92	A	A
## 7	Reuven	Ytzrhak	410	80	15	-1.05	F	C





### Question 1f.

```
full_name <- paste(roster_sort$Firstname, roster_sort$Lastname)
row.names(roster_sort) <- full_name
```

```
roster_sort_drop <- roster_sort[,c(-1,-2)]
roster_sort_drop
```

##	Math	Science	English	score	grade	Easy	Grade	MathGrade
## Greg Knox	625	95	30	1.34	A		A	A
## Angela Williams	600	99	22	0.92	A		A	A
## Joel England	573	89	27	0.70	B		A	A
## Mary Rayburn	522	86	18	-0.18	C		B	B
## Cheryl Cushing	512	85	28	0.35	C		B	B
## John Davis	502	95	25	0.56	B		B	B
## Janice Markhammer	495	75	20	-0.63	D		B	B
## Bullwinkle Moose	412	80	18	-0.86	D		C	C
## Reuven Ytzrhak	410	80	15	-1.05	F		C	C
## David Jones	358	82	15	-1.16	F		C	C

## Question 2.

```
t_data <- rt(1000,10)
```

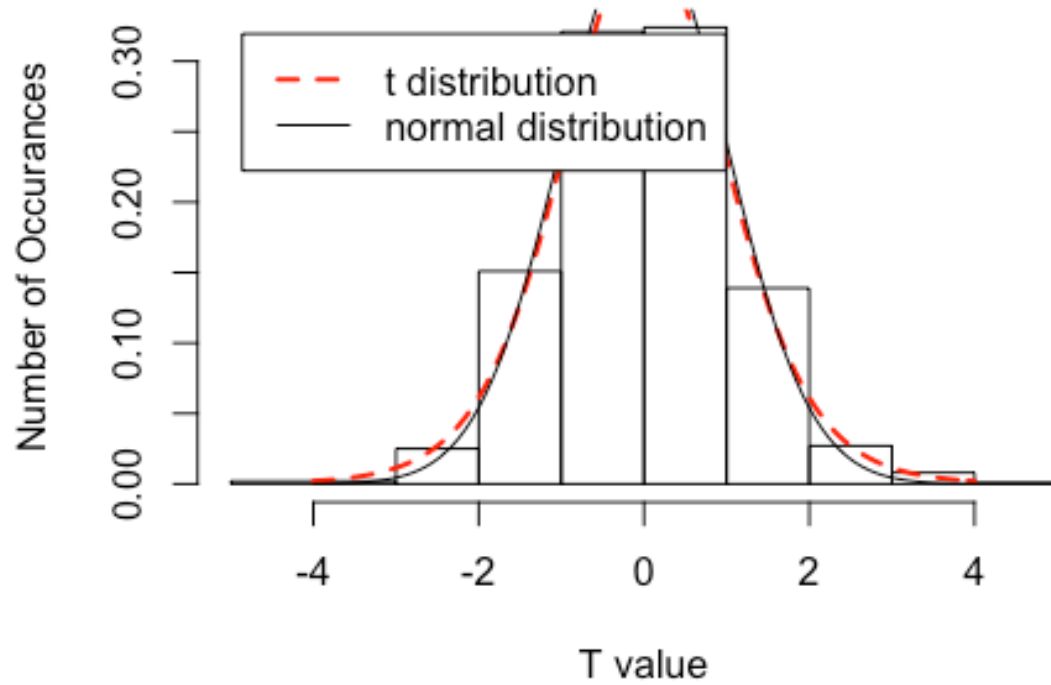
```
hist(t_data,  
     main = "Normal Approximation",  
     ylab = "Number of Occurances",  
     xlab = "T value",  
     probability = T)
```

```
curve(dt(x, df = 10),  
      xlim = c(-4,4),  
      lwd = 2,  
      lty = 2,  
      col = 'red',  
      add = T)
```

```
curve(dnorm(x),  
      xlim = c(-4,4),  
      add = T)
```

```
legend(x="topleft",  
      legend = c("t distribution", "normal distribution"),  
      lty = c(2,1), col = c('red','black'),lwd=c(2,1),  
      inset = 0.05)
```

## Normal Approximation



Question 3ab.

```
hist(rbinom(10000, 100, 0.05),  
     probability = T,
```

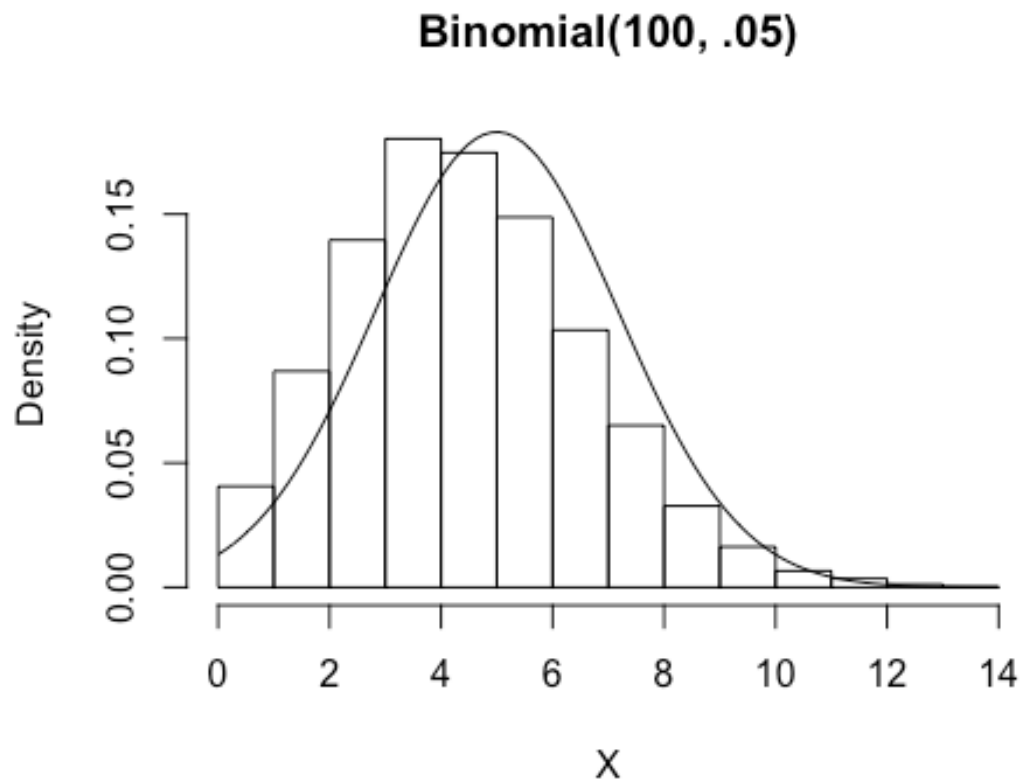
```

    main = 'Binomial(100, .05)',
    xlab = 'X'
  )

mu <- 100 * 0.05
sigma <- sqrt((100*0.05)*0.95)

curve(dnorm(x, mu, sigma), add = T)

```



### Question 3c.

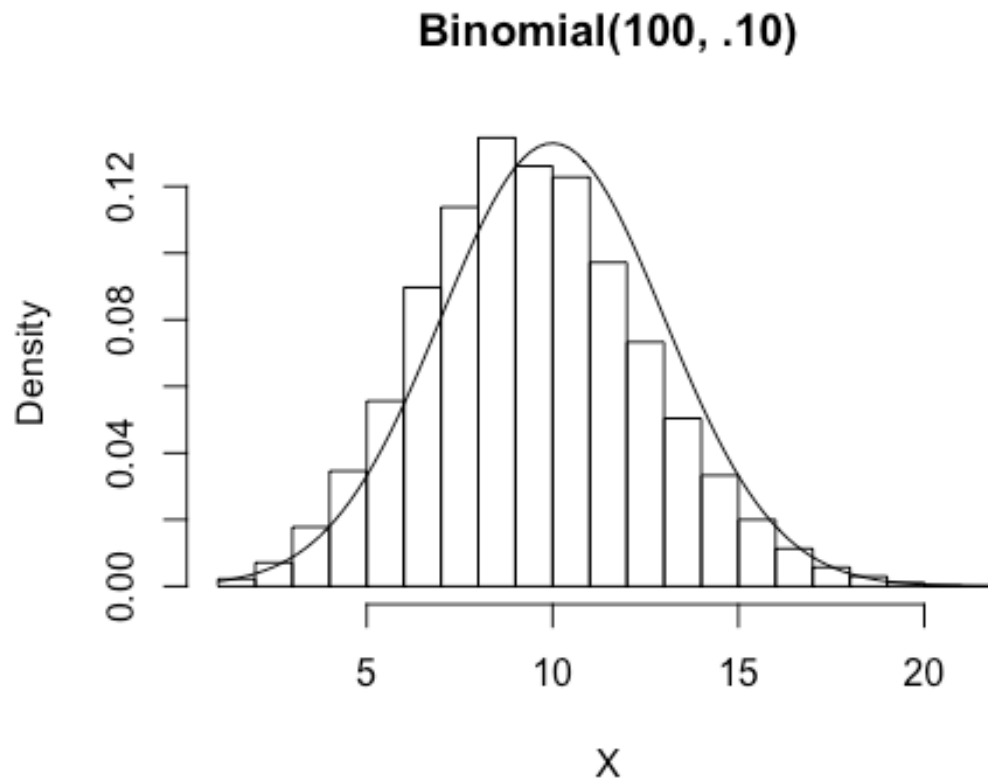
```

hist(rbinom(10000, 100, 0.10),
     probability = T,
     main = 'Binomial(100, .10)',
     xlab = 'X'
  )

```



```
mu <- 100 * 0.10  
sigma <- sqrt((100*0.10)*0.90)  
  
curve(dnorm(x, mu, sigma), add = T)
```



### Question 3d.

Yes, this seems reasonable based off the above plots.