IST 772 Homework 2

Due October 19, 2021 at 8:00AM EDT

Homework 2 by Nora Lin: I produced the material below with no assistance.

Excercise 1 p.35:

```
#Flipping a fair coin nine times:
#1 trial of 9 fair coin tosses in each trial
table(rbinom(n=1, size=9, prob=0.5))
```

```
##
## 7
## 1
```

```
#100,000 trails of 9 fair coin tosses in each trial
table(rbinom(n=100000, size=9, prob=0.5))
```

```
##
##     0     1     2     3     4     5     6     7     8     9
##   163  1725  7127 16327 24534 24832 16297  7017  1781   197
```

Excercise 2 p.35:

```
barplot(table( rbinom(n=100000, size=9, prob=0.5)))
```

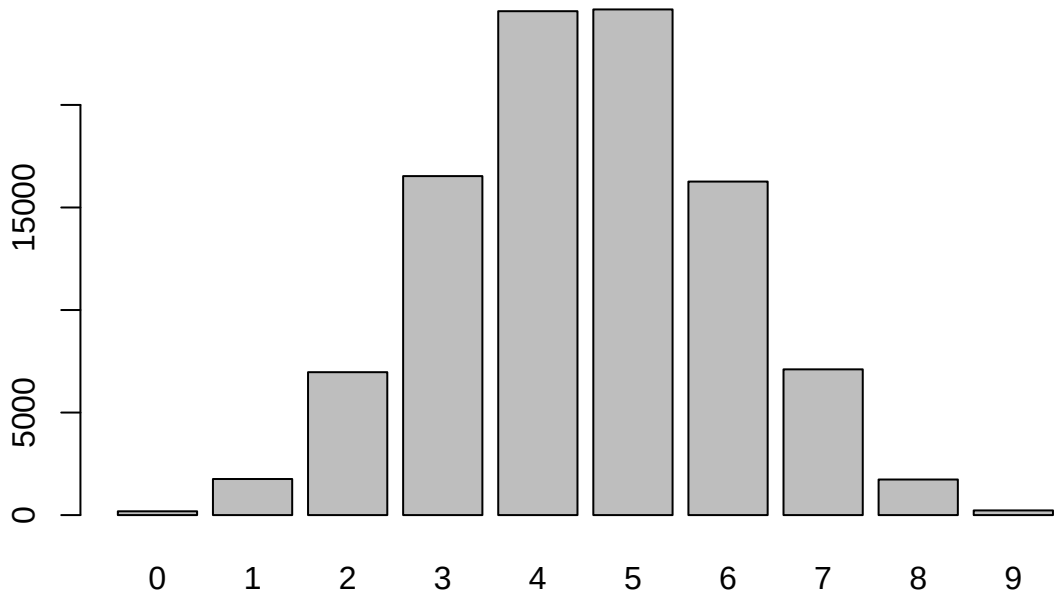

Figure 1: Barplot of n=100,000 trials of nine flips each

```
table( rbinom(n=100000, size=9, prob=0.5))/100000
```

```
##
##       0       1       2       3       4       5       6       7       8       9
## 0.00208 0.01758 0.07170 0.16469 0.24521 0.24733 0.16286 0.06892 0.01741 0.00222
```

```
barplot(table( rbinom(n=100000, size=9, prob=0.5))/100000)
```

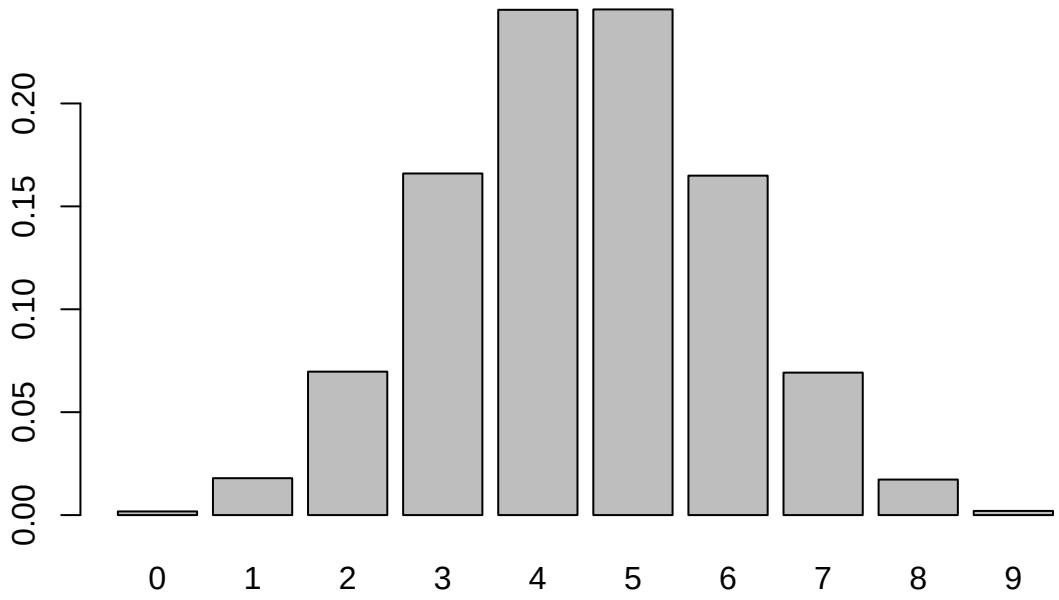

Figure 2:Bar plot of probabilities of each trial for n=100,000 trials in a binomial distribution

Analysis:

Figure 1 signifies a bar plot of 9 fair coins being flipped 100,000 times. The x-axis shows the number of times Heads was recorded. The y-axis shows the raw count for the corresponding number of Heads. In 100,000 trials of flipping a fair coin 9 times, there were 1,856 counts of 1 Head and 9 Tails.

Figure 2 signifies a bar plot of the probabilities of each trial for n=100,000 trails in a binomial distribution. The x-axix is the same as Figure 1. The Y-axis is the probability of each trial. In 100,000 trials of flipping a fair coin 9 times, the probability of getting three heads is 0.16292 or %16.29.

Figure 1 and 2 are related in that figure 1 is the raw counts and figure 2 is the probabilities of those raw counts.

The shape of both figures are normally distributed and centered around 4.5. This makes sense because of central tendency.

Excercise 6 p.36:

```
stat_test <- matrix(c(33,17,50,47,3,50,80,20,100),ncol=3,byrow=TRUE)
colnames(stat_test)<- c("Pass","Fail","Row totals")
rownames(stat_test)<- c("High school students","College students","Column totals")
stat_test <- as.table(stat_test)
stat_test
```

```
##                       Pass Fail Row totals
```

```
## High school students   33   17         50
## College students       47    3         50
## Column totals          80   20        100
```

Table 1.1: Two-by-Two Contingency Table with statistics test labels, data, and marginal totals

```
stat_testProbs <- stat_test/100
stat_testProbs
```

```
##                      Pass Fail Row totals
## High school students 0.33 0.17       0.50
## College students     0.47 0.03       0.50
## Column totals        0.80 0.20       1.00
```

Table 1.2: Two-by-Two Contingency Table with frequencies converted to probabilities.

If one were to only focus on high school students, the probability that a student would pass the test would be 0.33 ot 33%.

Excercise 7 p.36:

```
UKhomes <- matrix(c(2,69,71,93933,5996,99929,93935,6065,100000),ncol=3,byrow=TRUE)
#UKhomes
colnames(UKhomes) <- c("Pass","Fail","Row total")
rownames(UKhomes) <- c("Defaulting mortgage","Not defaulting mortgage", "Column Total")
UKhomes  <-as.table(UKhomes)
UKhomes
```

```
##                           Pass   Fail Row total
## Defaulting mortgage          2     69        71
## Not defaulting mortgage  93933   5996     99929
## Column Total             93935   6065    100000
```

Table 2.1: Two-by-Two Contingency Table with frequencies converted to probabilities

```
UKhomesProbs <- UKhomes/100000
UKhomesProbs
```

```
##                             Pass    Fail Row total
## Defaulting mortgage      0.00002 0.00069   0.00071
## Not defaulting mortgage 0.93933 0.05996   0.99929
## Column Total            0.93935 0.06065   1.00000
```

Table 2.2: Two-by-Two Contingency Table with UK home repossession labels, data, and marginal totals

0.93935 or 93.94% of homes both passed the Barclays Bank test and do not have their homes repossessed.

Excercise 8 p.36:

```
norm_UKhomesProb <- matrix(c(UKhomesProbs[1,2],(UKhomesProbs[1,2]/UKhomesProbs[3,2]),
          UKhomesProbs[2,2],(UKhomesProbs[2,2]/UKhomesProbs[3,2]),
          UKhomesProbs[3,2],(UKhomesProbs[3,2]/UKhomesProbs[3,2])),
          ncol=2,byrow=TRUE)
#norm_UKhomesProb
colnames(norm_UKhomesProb) <- c("Raw Fail Prob", "Norm Fail Prob")
rownames(norm_UKhomesProb) <-c("Defaulting mortgage","Not defaulting mortgage","Marginal")
norm_UKhomesProb
```

```
##                          Raw Fail Prob Norm Fail Prob
## Defaulting mortgage            0.00069     0.01137675
## Not defaulting mortgage        0.05996     0.98862325
## Marginal                       0.06065     1.00000000
```

Table 2.3: Excerpt of Table Showing UK home repossessions with Normalized Probabilities There is a 0.01137675 or 0.11% probability that this new customer who failed the Barclays Bank screening test will actually default on his or her mortgage.

End of Homework Assignment #2