MATH 21-01 (Introductory Statistics), HW5 (100 points). Due: 09/30/2016 in class.

Not from textbook (60 pts)

In this homework, we will use R to simulate the effect of a coin toss experiment. All the code you need is essentially given to you below. The goal of this assignment is to help you understand how such a simulation in R works and employ the methods we discuss. Suppose, you and your friend play a game of tossing coins. You flip the coin many times. For each head your toss yields, your friend plays you one dollar (+1 dollar for you). For each tail, you pay one dollar to your friend (−1 dollar for you).
You will now simulate the profit/loss of such a game in R.
Sample, a sequence of 2000 coin tosses:

```
N <- 2000;
coinTossResults <- sample(c(-1,1), N, replace = TRUE);
```

- (A - 10 pts) Compute the mean, median, and std deviation of the coinTossResults list. Make N equal to 50000 and repeat. Describe how you would interpret these numbers?
- (B - 5 pts) How would you compute the total profit or loss that you would get out of playing this game with a fixed number of coin tosses N?
- (C - 5 pts) Consider running the command:

```
plot(cumsum(coinTossResults), type = 'l')
```

  Label this plot and include in your homework. What does the resulting plot tell you? How do you interpret the net profit or loss after $N = 2000$ tosses from this plot?

- (D - 15 pts) Simulate the results of several trials of this game (with $N = 2000$ coin tosses each). This can be done using the following code:

```
N <- 2000;
num_trials <- 10;
net_profits <- list();
for(tn in 1:num_trials) {
        coinTossResults   <- cumsum(sample(c(-1,1), N, replace = TRUE))
        net_profits[[tn]] <- coinTossResults[length(coinTossResults)]
}
```

  Describe what this code does, a line at a time. How do you interpret the results of the net_profits list?

- (E - 25 pts) Run the game using num_trials = 5, 100, 1000, 20000. In each case, plot and label a histogram of the net_profits. Plot a normal curve over the histogram to see how close to normal the distribution is in each case. State if or not the distribution is approximately bell shaped for each number of trials. Include the labeled histograms in your output. You may find the following code useful.

```
x <- unlist(net_profits)
h<-hist(x, breaks=40, col="red", xlab="Profit",
   main="Histogram of profit/loss distribution")
xfit<-seq(min(x),max(x),length=length(x))
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="blue", lwd=2)
```

  Please adjust parameters as necessary to make clear plots. What happens as the number of trials increases? What can you say about the expected profit of the game over a large number of trials?

From textbook (40 pts)

Section 3-3: 41,42,43,44
Section 3-4: 5