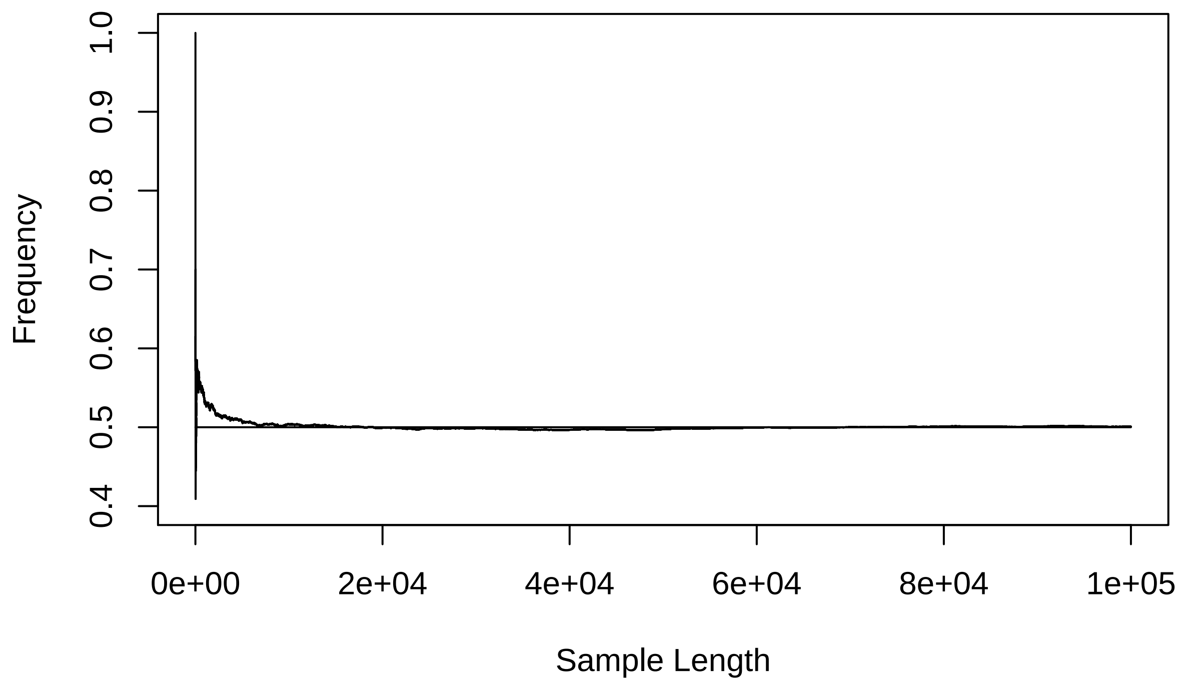
Analyze Trajectories of Multiple Coin Tossing Experiment Assignment 1 - John Navarro

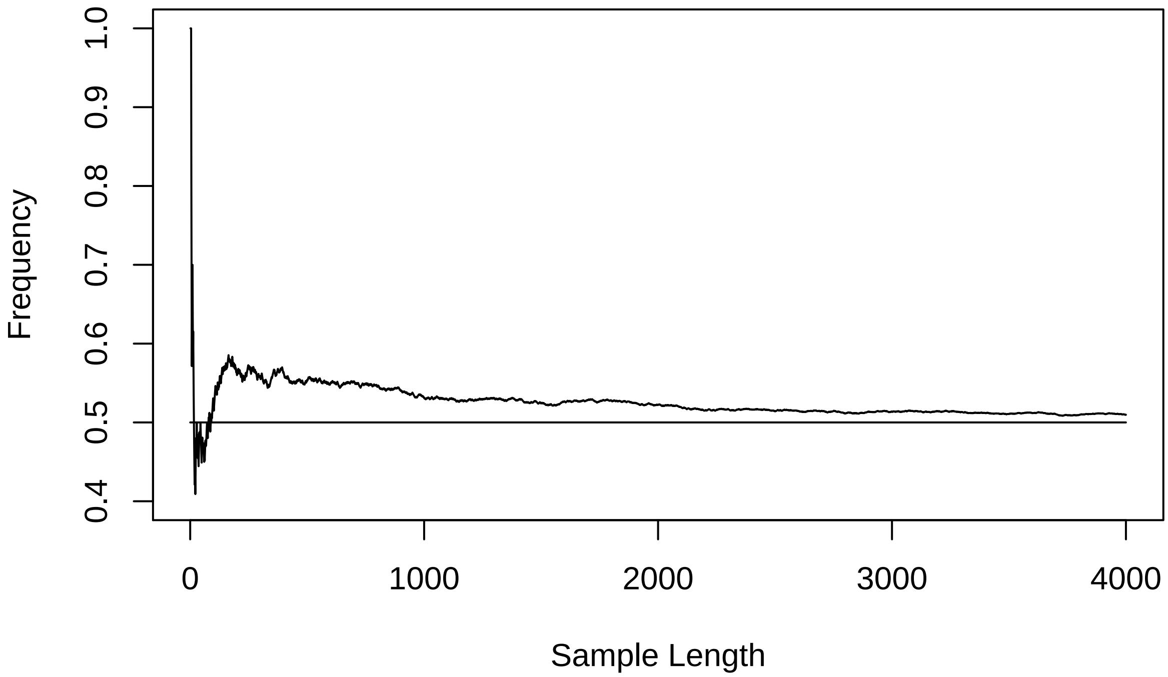
1. Convergence of probability of tail to 0.5

What does this say about fairness of the coin? If the probability of tail converges to 0.5, then it means that the coin is fair

```
##sets the "random" pattern of flips to a specific seed, that way it can be random
##and replicated
set.seed(12345)
##asign value of 100000 to variable nFlips
nFlips <- 100000
##generate a sample of outcomes either 0 or 1, 100000 times, with replacing set to
##true
Flips <- sample(0:1,nFlips,repl=T)
##store the cumulative sum of Flips into variable Trajectory
Trajectory<-cumsum(Flips)
##store the value of Trajectorydivided by the number of flips into freq
freq <- Trajectory/(1:nFlips)
##creates a plot of length of freq vs freq. y axis set from 0.4 to 1, labels the
##axis as Frequency and Sample length
plot(1:length(freq),freq, ylim=c(.4,1),type="l",ylab="Frequency",xlab="Sample Length")
##draws a horizontal line at 0.5
lines(c(0,nFlips),c(.5,.5))
```




```r
##creats a chart of points up to 4000 and their corresponding frequencies, sets y
##axis from .4 to 1, labels axis
plot(1:4000,freq[1:4000], ylim=c(.4,1),type="l",ylab="Frequency",xlab="Sample Length")
##draws a horizontal line at 0.5
lines(c(0,4000),c(.5,.5))
```



The data is noisy when the sample length is low. As the sample length increases, the Frequency approaches 0.5, as is shown by the difference in the two plots. The first plot has 100,000 data points, while the second plot has only 4,000.

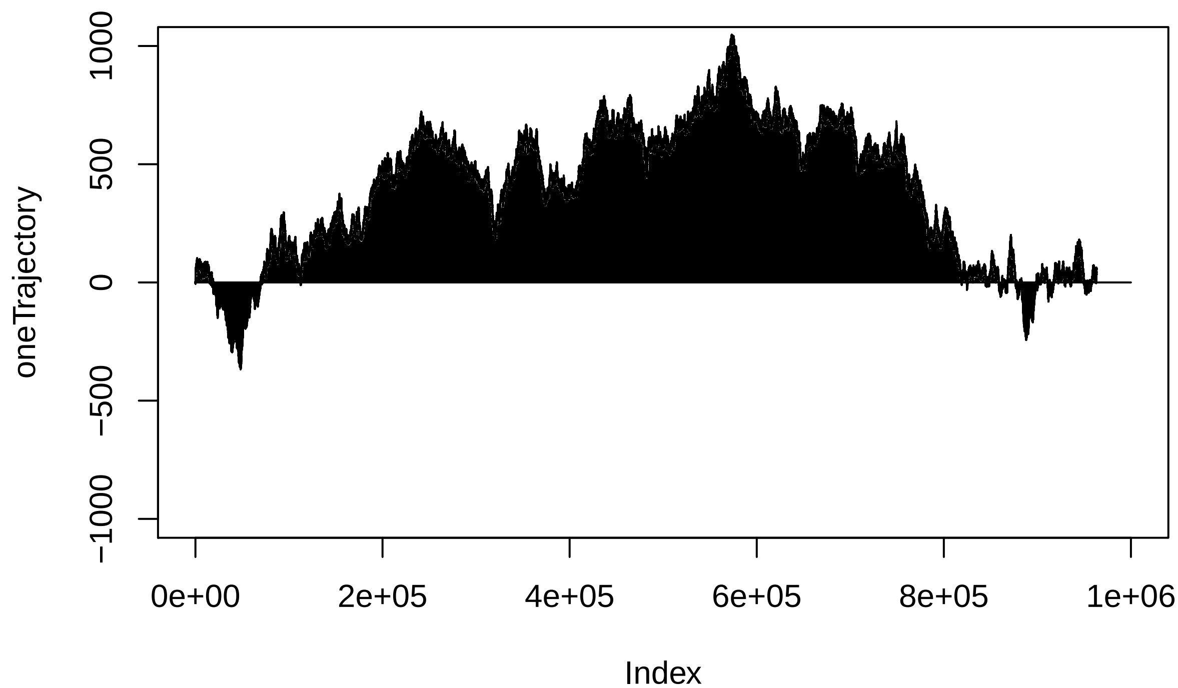
2.1 One trajectory

```
##sets nFlips to 1 million
nFlips<-1000000;
##sets the seed to specific value
set.seed(12345)
##generates sample of values 0 or 1, 1000000 times, with replacing values,
##subtracts 0.5 from each value and multiplies all values by two. This will give us
##-1s and +1s
Flips<-(sample(0:1,nFlips,repl=T)-.5)*2
```

```
##Alternative way to simulate variable Flips using rbinom() instead of sample()
nFlips<-1000000;
set.seed(12345)
Flips<-(rbinom(nFlips,1,.5)-.5)*2
```

How much do you expect the trajectory of wealth to deviate from zero? I expect that the trajectory can deviate far from zero, but will eventually approach zero as the sample size increases. *How long do you expect it to stay on one side above or below zero?* After 4,000 flips

```r
## Assign the cumulative sum of Flips to variable oneTrajectory
oneTrajectory<-cumsum(Flips)
## Plot line graph of one Trajectory vs # of Flips, sets Y axis from -1000 to +1000
plot(oneTrajectory, ylim=c(-1000,1000),type="l")
## Draws a horizontal line at 0
lines(c(0,nFlips),c(0,0))
```



How do the observations match your prior expectations? Even after 1,000,000 flips, in some cases the Trajectory is still not approaching zero.

2.2 Multiple trajectories

What do you expect the probabilities of the following events to be?

Using a binomial distribution calculator, with inputs 0.5, and 500. I found the p(x<253) - p(x<248) = 0.58845-.41154 so, P(Nh-Nt <5) = 0.17691

Again using a binomial distribution calculator, I found the p(x>=263) = 0.13177, p(x<238)= 0.13177, summing the two values gives, P(Nh-Nt > 25) = 0.26354

```r
## Creates a 2000 row by 500 column matrix, using the simulated Flips data, and
##uses apply to run the cumulative sum within each row. Then transpose the matrix
Trajectories2000by500<-t(apply(matrix(Flips,ncol=500),1,cumsum))
##checks the dimensions of the matrix
dim(Trajectories2000by500)
```

```
## [1] 2000  500
```

```
## Sums the number of  trajectories that have an absolute value less than 5, then
##divides by 2000 to find the probability
(probability.less.than.5<-sum(abs(Trajectories2000by500[,500])<5)/2000)
```

```
## [1] 0.18
```

```
## Sums the number of trajectories with an absolute value greater than or equal to
##25, then divides 2000 to find the probability
(probability.greater.than.25<-sum(abs(Trajectories2000by500[,500])>=25)/2000)
```

```
## [1] 0.2515
```

Interpret the results. How did they correspond to your intuition?

Using this strategy of 2000 trajectories of 500 flips, 17.75% of the time we end up less than 1% away from zero. Similarly, 27.35% of the time, the trajectories end up more than 5% away from zero.

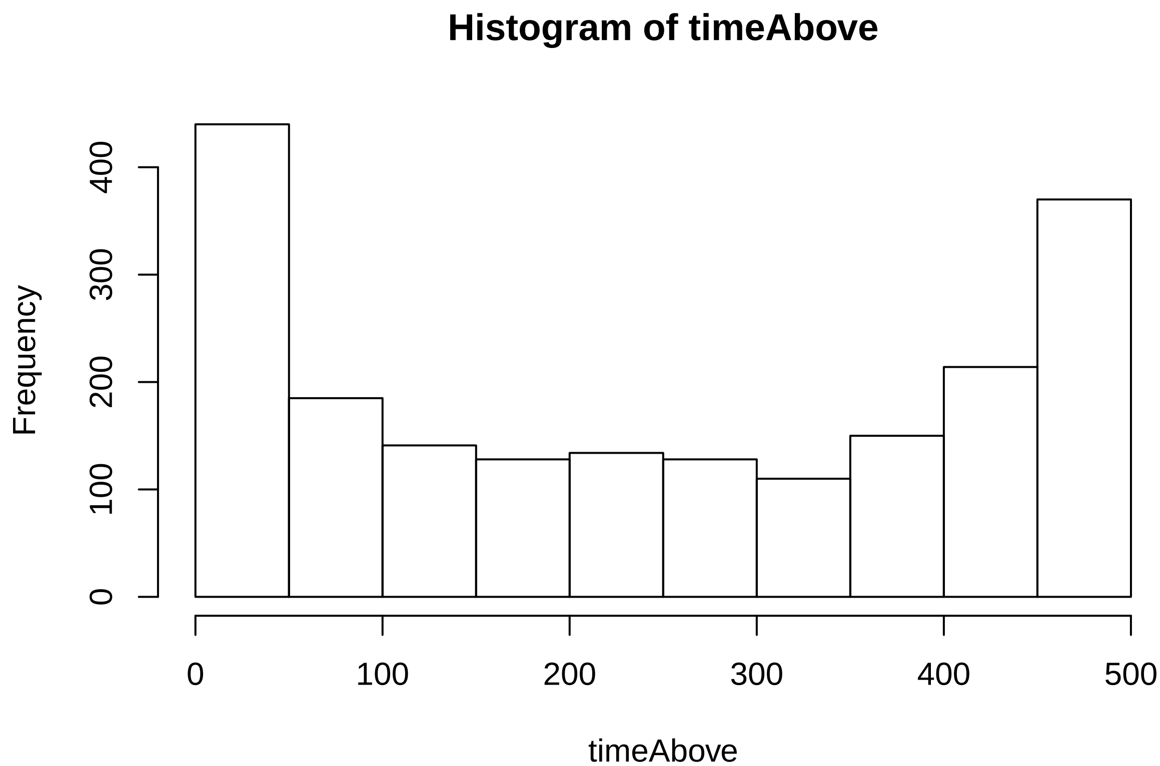
The results are close to the estimations I made using the binomial calculator, I assume that as the number of random walk samples increases above 2000, then values given will approach the ones from the binomial distribution.

2.3 Time on one side

How long do you expect trajectory of random walk to spend on one side from zero, below or above? (Answer before doing calculations)

I expect the trajectory of random walk to be on one side of zero, the majority ofthe time.

```
##Uses the apply function to sum each column in Trajectories2000by500, gives 500 ##sums
timeAbove<-apply(Trajectories2000by500,1,function(z) sum(z>0))
##Creates a histogram of these sums (timeAbove)
hist(timeAbove)
```



The results seem to match my intuition that the majority of the time it will be above/below zero.

According the the histogram, There is a large sample close to 0, which are the cases that spend little time above zero and most of the time below zero. Another large sample is at 500, which means that the trajectories would spend all their time(500 flips) above zero.

This is a demonstration of the gambler's fallacy. That if you have a string of outcomes, that a string of the opposite outcome is just around the corner.