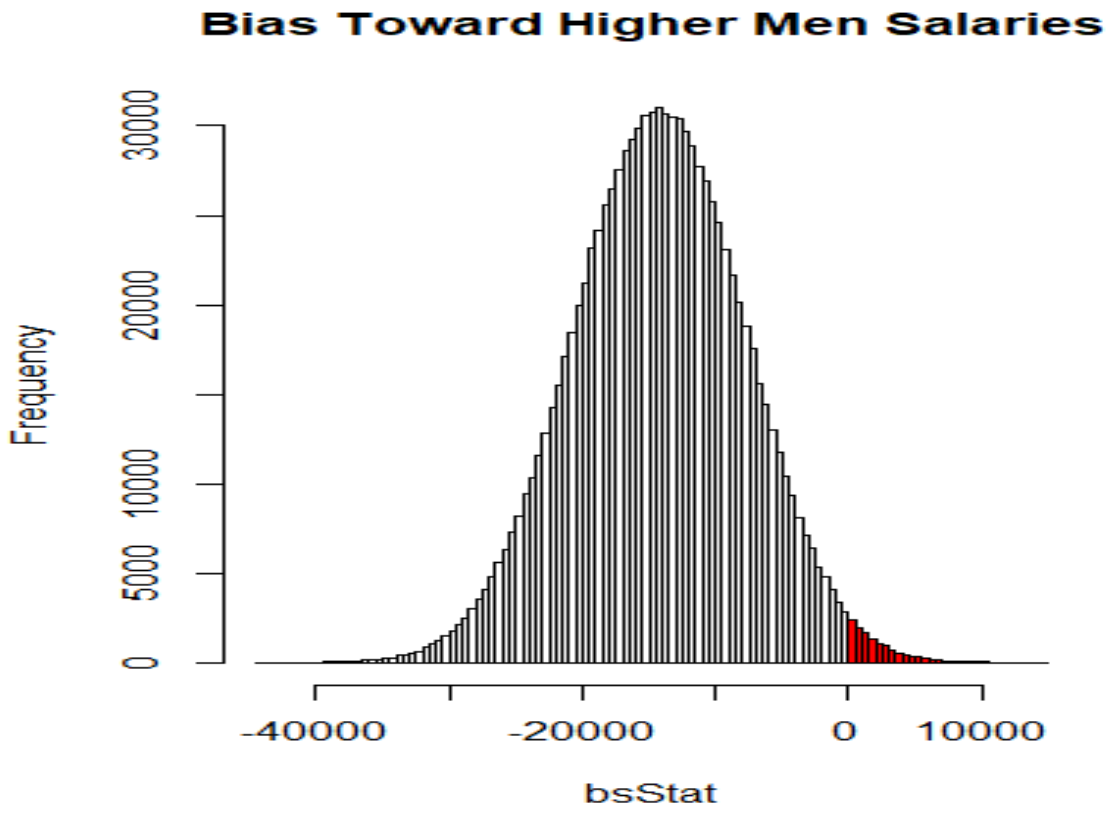
Charles Hesketh

Math 536

Gender Discrimination-HW1

In order to verify that there was no gender bias in the company's pay between men and women I preformed two different tests. One was a standard t-test that evaluated the company's pay with some nice assumptions included, but assumed the variance was not the same. This suggest that the likelihood for bias toward higher Men salaries was 1.427%.

Furthermore, I preformed another analysis using bootstrapping that allowed for no assumptions when running the analysis. This allows us to run a "simulation" using the data we have to provide a more robust analysis. This led to a suggest likelihood bias of 1.2675%. In the graph below you can see the distribution of the difference in average salary between females and males with the red portion of the graph representing 1.2675%, the portion of the distribution that creates the likelihood for bias.



Code Used in Analysis

```
library(matlib)
library(tidyr)
#read File
salarys= read.csv(file = 'C:/Users/User/Desktop/School/Math 536/HW/HW1/HW1P1.csv')
#seperate classes
males=salarys['Males']
females=salarys['Females']
#make males more user friendly for later
males=as.matrix(males)
#Drop rows with no data in Females
femalesKeep=females[complete.cases(females), ]
#Theoretical t.test for values
t.test(femalesKeep,males,alternative="less",var.equal = F)
#create vectors to be used in loop
femBS = rep(0,1000000)
maleBS = rep(0,1000000)
bsStat = rep(0,1000000)
#bootstraping
for(i in 1:1000000){
  #creating populations to use in bootstraping statistic
  femBS[i]=mean(sample(femalesKeep,104,replace=T))
  maleBS[i]=mean(sample(males,115,replace=T))
  #creating bootstraping statistic
  bsStat[i]=femBS[i]-maleBS[i]
}
#Generate P-values for boostraping using >0 as the cut off for difference
p.val.bs = length(bsStat[bsStat>0])/length(bsStat)
p.val.bs
```

```
#create histogram
h=hist(bsStat,breaks=100,plot=F)
#create coloring within histogram to be referenced later in report
cuts=cut(h$breaks,c(-Inf,-0.000001,Inf))
plot(h,main="Bias Toward Higher Men Salaries",col=c("white","red")[cuts])
```