RStudio HW

Rick Guha

February 24, 2023

## Read in the data

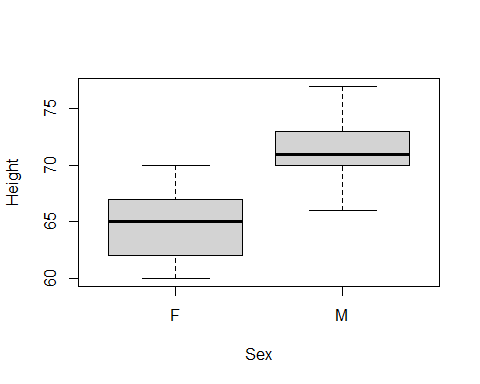
First, we will load the data.

# Activate the Stat2Data package (must be installed first - one-time step)  
library (Stat2Data)  
  
# load the Day1Survey data set  
data (Day1Survey)  
  
head (Day1Survey)

## Section Class Sex Distance Height Handedness Coins WhiteString BlackString  
## 1 1 Senior F 400 62 Right 1.12 42 6  
## 2 1 \* F 450 61 Left 29.00 45 5  
## 3 1 Freshman F 3000 61 Right 1.50 22 4  
## 4 1 Freshman M 100 72 Right 0.07 40 4  
## 5 1 N/A F 2000 69 Right 0.12 48 7  
## 6 1 Senior M 500 73 Right 8.00 30 8  
## Reading TV Pulse Texting  
## 1 80 3 71 3  
## 2 100 10 78 100  
## 3 100 4 80 2  
## 4 50 25 63 200  
## 5 200 5 63 100  
## 6 100 0 56 1

Then make a plot.

# Three different ways to do the same thing  
boxplot (Height ~ Sex, data=Day1Survey)



The plot above suggests that there is some difference between the groups.

Here are the summary statistics by group:

library (mosaic)  
favstats(Height ~ Sex, data=Day1Survey)

## Sex min Q1 median Q3 max mean sd n missing  
## 1 F 60 62 65 67 70 64.64706 3.371594 17 0  
## 2 M 66 70 71 73 77 71.57692 2.386178 26 0

Let’s do a two sample t-test to determine if there is actually a statistically significant difference between the heights of men and women:

t.test (Height ~ Sex, data=Day1Survey, var.equal=T)

##   
## Two Sample t-test  
##   
## data: Height by Sex  
## t = -7.9007, df = 41, p-value = 9.092e-10  
## alternative hypothesis: true difference in means between group F and group M is not equal to 0  
## 95 percent confidence interval:  
## -8.701235 -5.158494  
## sample estimates:  
## mean in group F mean in group M   
## 64.64706 71.57692

The p-value is 9.092e-10, so we reject the null hypothesis and conclude that the mean mean height difference is statistically different between men and women.

We are 95% confident that women, on average, are between -8.7 inches and -5.15 inches shorter than men.