

# Project #4: Fitting probabilistic models

Stephen, Amber, Josh

2021-10-25

---

```
#library(nimble)
```

## Problem #2 (14 points)

Here, the goal is to provide a probabilistic model for some of the data available here:

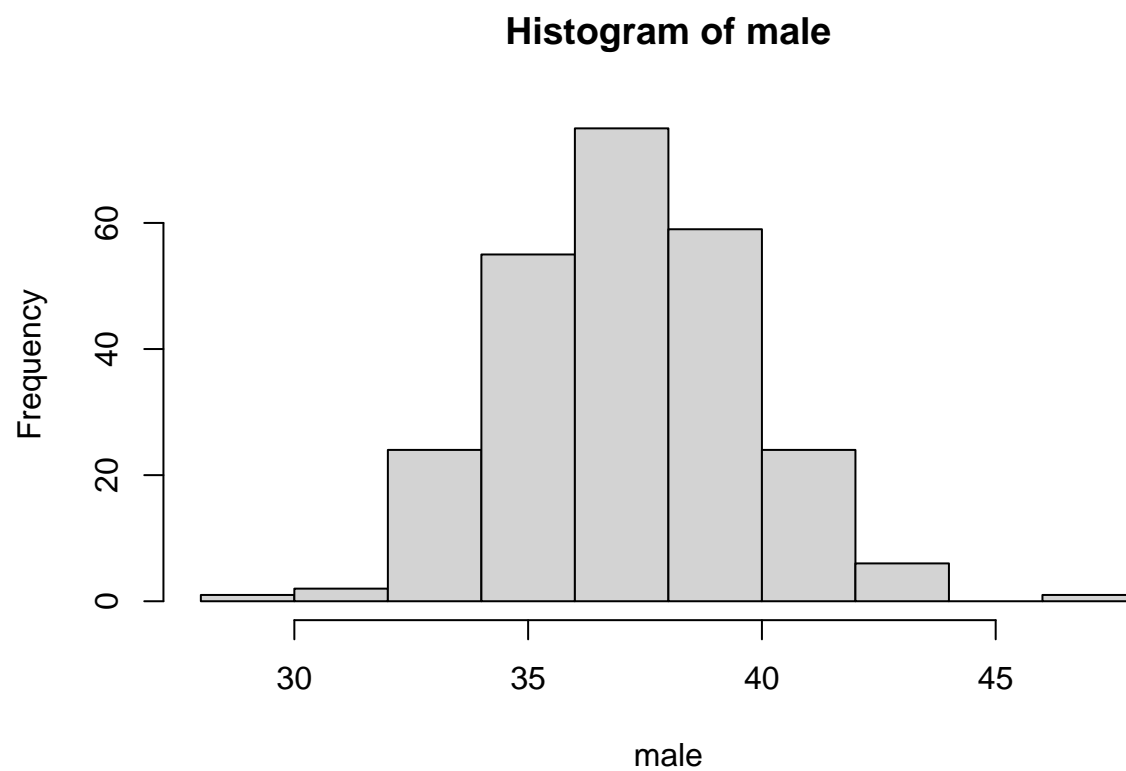
[Data set associated with the textbook](#)

First, download the data set and import it into R. Do not forget that the documentation for the data set is also available under the above link.

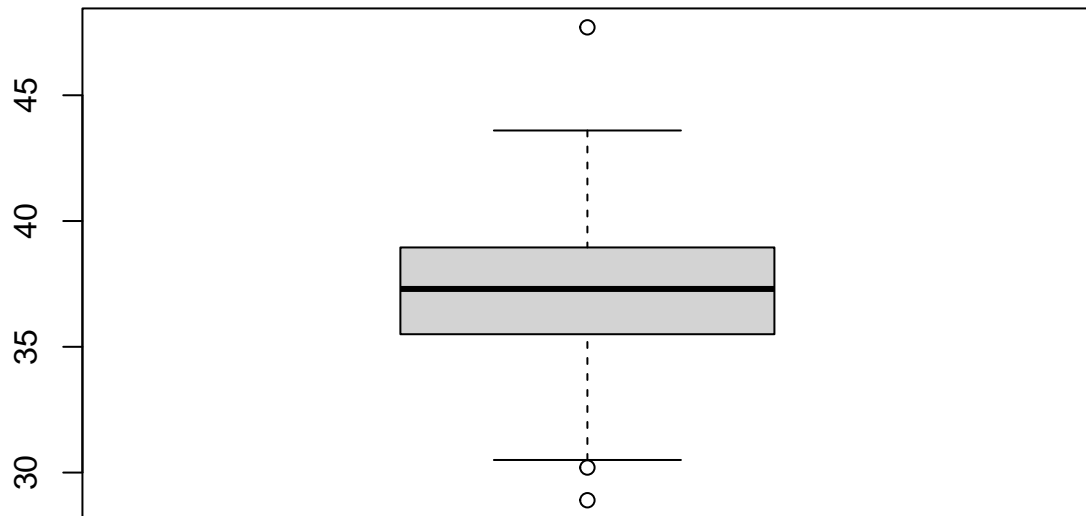
```
getwd()
## [1] "/Users/kariambervasquez/Downloads"
body = read.csv("bdims.csv")
```

**(2 points)** Focus on the measurements of the male respondents' calves' maximum girth in centimeters. Does this set of measurements have any outliers? Yes, this has outliers

```
male = (body$cal_gi[body$sex==1]) # A numerical vector, respondent's calf maximum girth in centimeters
hist(male)
```

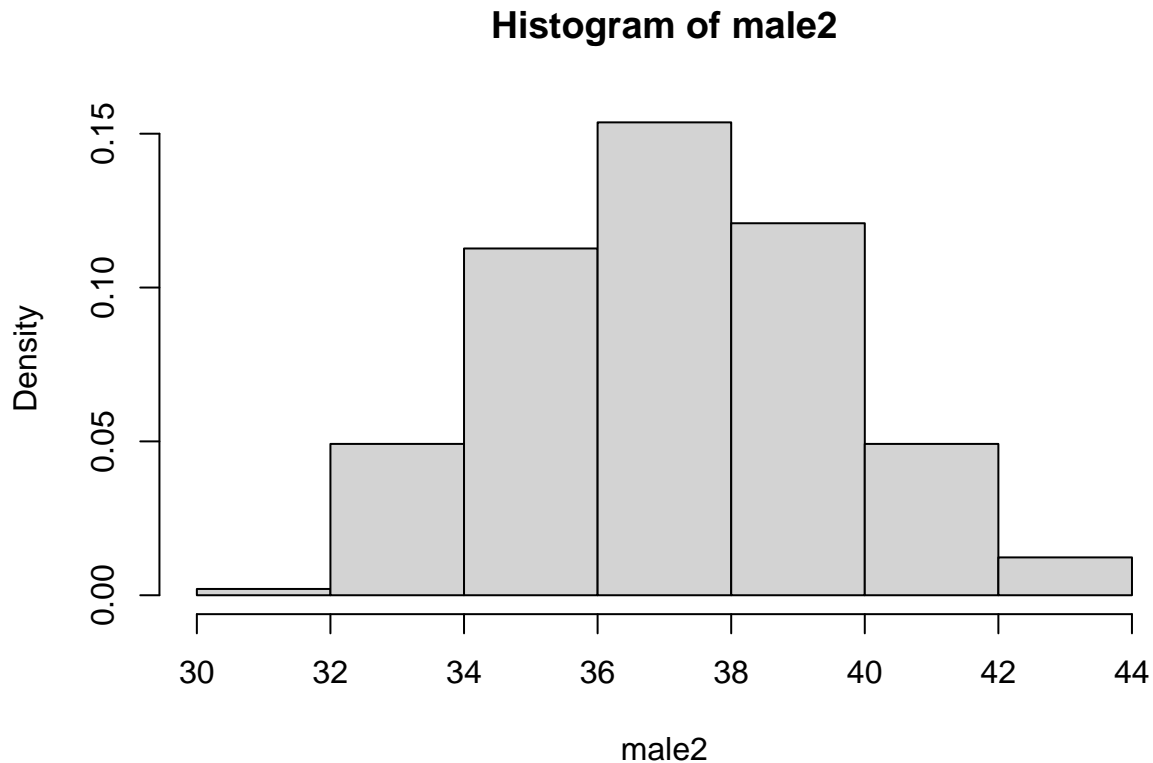


```
boxplot(male)
outliers = boxplot(male)$out #identifies outliers
```



(4 points) If there are outliers, discard them from your data set. Visualize the remaining data points appropriately.

```
male[which(male%in%outliers)] # find the indeces of the outliers
## [1] 30.2 47.7 28.9
male2= male[-which(male%in%outliers)] #returns male except the three
#now should be anything but outliers
hist(male2, prob = TRUE)
```



**(2 points)** Propose a **named parametric distribution** to fit to your data and justify your choice. We chose a Normal Distribution because the histogram shows a symmetry that a normal curve would fit well with

**(2 points)** Using the data, propose a point estimate for any parameters in your model. We propose a point estimate of 37 because it lies in the middle of the distribution

**(4 points)** Superimpose the appropriate graph for your model onto the appropriate graph of the data to convince your reader that your model is valid.

```
hist(male2, prob = TRUE)
mean = mean(male2)
var = var(male2)
sd = sd(male2)
curve(dnorm(x,mean,sd),from = min(male2), to = max(male2) , add = TRUE)
```

