Laboratory Exercise Week 6

Brian Tipton - STAT380-001

9/26/23

Directions:

- Write your R code inside the code chunks after each question.
- Write your answer comments after the # sign.
- To generate the word document output, click the button `Knit` and wait for the word document to appear.
- RStudio will prompt you (only once) to install the `knitr` package.
- Submit your completed laboratory exercise using Blackboard's Turnitin feature. Your Turnitin upload link is found on your Blackboard Course shell under the Laboratory folder.

For this exercise, you will need to use the package `mosaic` to find numerical and graphical summaries.

```r
# install package if necessary
if (!require(mosaic)) install.packages(`mosaic`)
# load the package in R
library(mosaic) # load the package mosaic to use its functions
```

1. A national survey in 2012 showed that 75% of middle and high school teachers believe their students use Wikipedia for typical research statements. Let X equal the number of teachers in a random sample of size n = 40 who believe their students use Wikipedia.

    i) How is X distributed?

    ii) Use `plotDist()` to draw the probability distribution of X.

    iii) Compute the probability that X is equal to 30.

    iv) Compute the probability that X is at most 30. What is the difference between parts (iii) and (iv).

    v) Compute the probability that X is anywhere from 30 to 35.

Code chunk
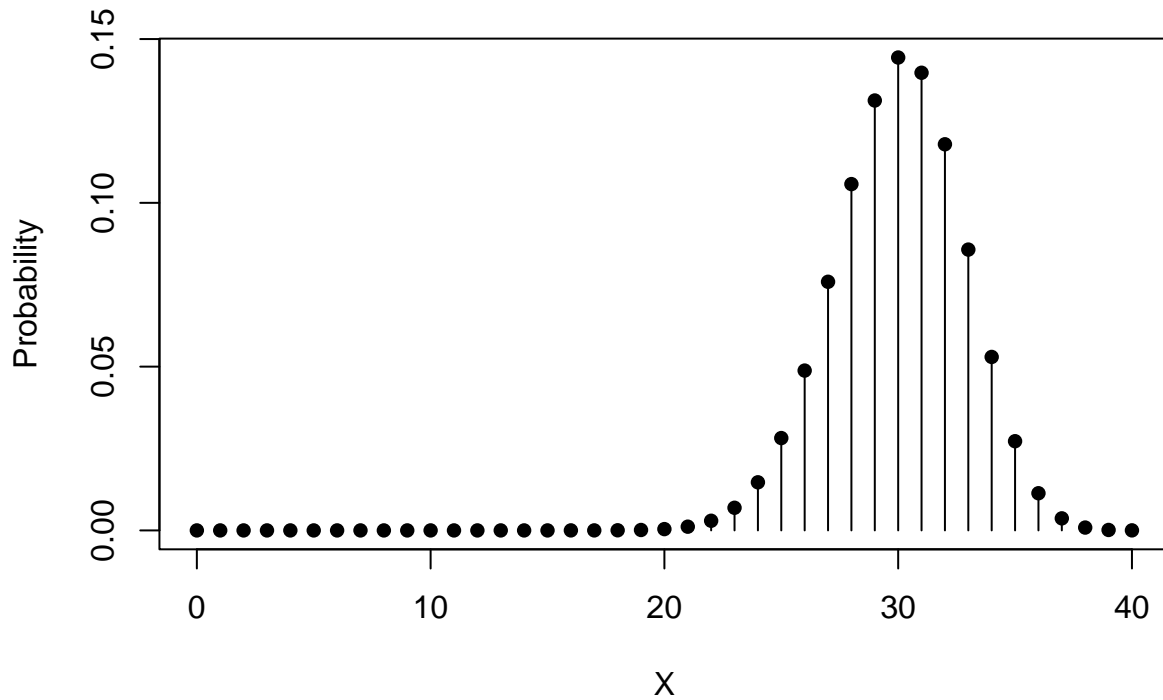
```r
#i) How is X distributed?
cat("X is distributed as Binomial(n=40,p=0.75)Binomial(n=40,p=0.75)")
```

```
## X is distributed as Binomial(n=40,p=0.75)Binomial(n=40,p=0.75)
```

```r
#ii) Use `plotDist()` to draw the probability distribution of X.
n <- 40
p <- 0.75
x <- 0:n
prob <- dbinom(x, size=n, prob=p)

plot(x, prob, type="h", xlab="X", ylab="Probability", main="Probability Distribution of X")
points(x, prob, pch=16)
```

Probability Distribution of X


```r
#iii) Compute the probability that X is equal to 30.
dbinom(30, size=40, prob=0.75)
```

```
## [1] 0.1443643
```

```r
#iv) Compute the probability that X is at most 30. What is the difference between parts (iii) and (iv).
pbinom(30, size=40, prob=0.75)
```

```
## [1] 0.5604603
```

```r
#v) Compute the probability that X is anywhere from 30 to 35.
pbinom(35, size=40, prob=0.75) - pbinom(29, size=40, prob=0.75)
```

```
## [1] 0.5678618
```

2. The size of the left upper chamber of the heart is one measure of cardiovascular health. When the upper left chamber is enlarged, the risk of heart problems is increased. The paper ("Left atrial size increases with body mass index in children")[https://www.ncbi.nlm.nih.gov/pubmed/19147240] (International Journal of Cardiology [2009]) described a study in which the left atrial size was measured for a large number of children age 5 to 15 years. Based on this data, the authors concluded that for healthy children, the left atrial diameter was approximately normally distributed with a mean of 26.4 mm and a standard deviation of 4.2 mm.
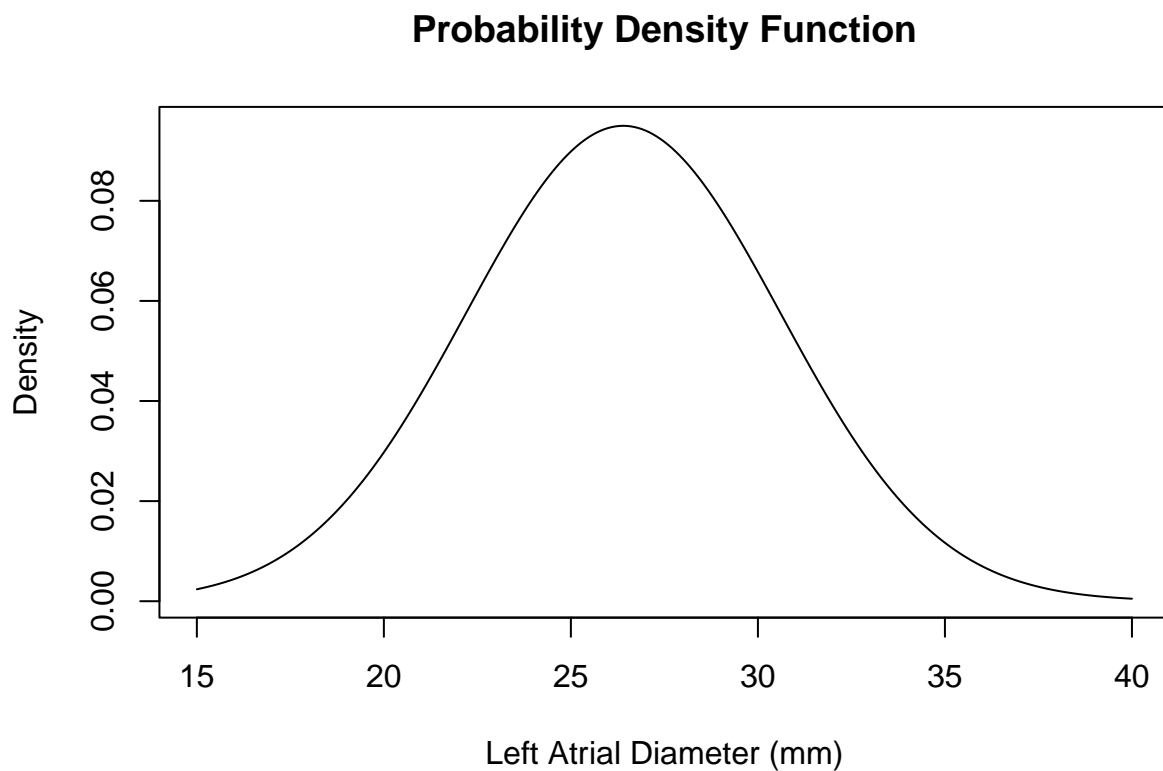
   i) Use `plotDist` to draw the probability density function of the left atrial diameter.

   ii) Generate 50 sample measurements of the left atrial diameters and plot the measurements using a histogram.

iii) Approximately what proportion of healthy children have left atrial diameters less than 23.5 mm?

iv) Approximately what proportion of healthy children have left atrial diameters greater than 32.5 mm?

v) Approximately what proportion of healthy children have left atrial diameters between 24 and 29 mm?

vi) For healthy children, what is the value for which only about 25% have larger left atrial diameter?

Code chunk

```
# start your code
#i) Use `plotDist` to draw the probability density function of the left atrial diameter.
mean <- 26.4
sd <- 4.2
x <- seq(15, 40, by=0.1)
y <- dnorm(x, mean=mean, sd=sd)

plot(x, y, type="l", xlab="Left Atrial Diameter (mm)", ylab="Density", main="Probability Density Functi
```

Probability Density Function


```
# ii) Generate 50 sample measurements of the left atrial diameters and plot the measurements using a hi
samples <- rnorm(50, mean=mean, sd=sd)
hist(samples, breaks=10, xlab="Left Atrial Diameter (mm)", main="Histogram of Sample Measurements")
```

Histogram of Sample Measurements
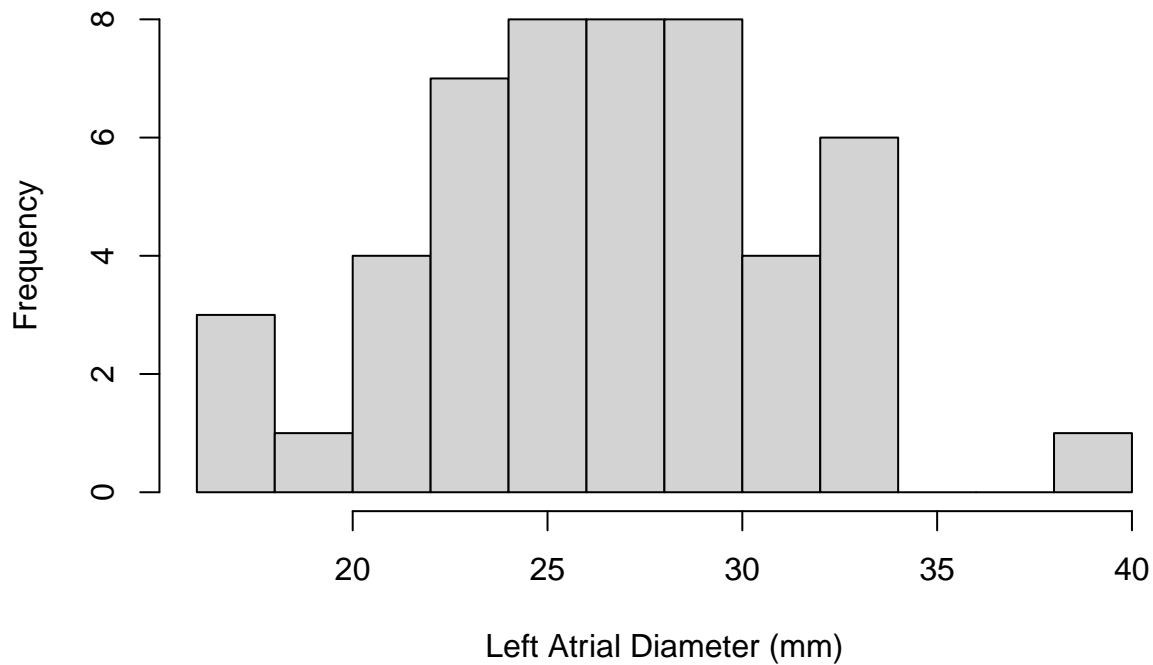

```r
#iii) Approximately what proportion of healthy children have left atrial diameters less than 23.5 mm?
pnorm(23.5, mean=mean, sd=sd)
```

```
## [1] 0.2449474
```

```r
#iv) Approximately what proportion of healthy children have left atrial diameters greater than 32.5 mm?
1 - pnorm(32.5, mean=mean, sd=sd)
```

```
## [1] 0.07319785
```

```r
# v) Approximately what proportion of healthy children have left atrial diameters between 24 and 29 mm?
pnorm(29, mean=mean, sd=sd) - pnorm(24, mean=mean, sd=sd)
```

```
## [1] 0.4482029
```

```r
#vi) For healthy children, what is the value for which only about 25% have larger left atrial diameter?
qnorm(0.75, mean=mean, sd=sd)
```

```
## [1] 29.23286
```

```r
# last R code line
```