Laboratory Exercise Week 7

Brian Tipton

10/4/23

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.

```
# 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. Consider the population of CEO salaries in the lesson this week.
   i) Select a random sample of 40 CEO's and compute the sample mean salary and sample standard deviation salary.
   ii) Use the `do()` function to compute the sample mean of 100 randomly selected samples. Plot the histogram of these 100 sample means.
   iii) Compute the average and standard deviation of these 100 sample means. Are the values what you expect from the properties of the sampling distribution.

Code chunk

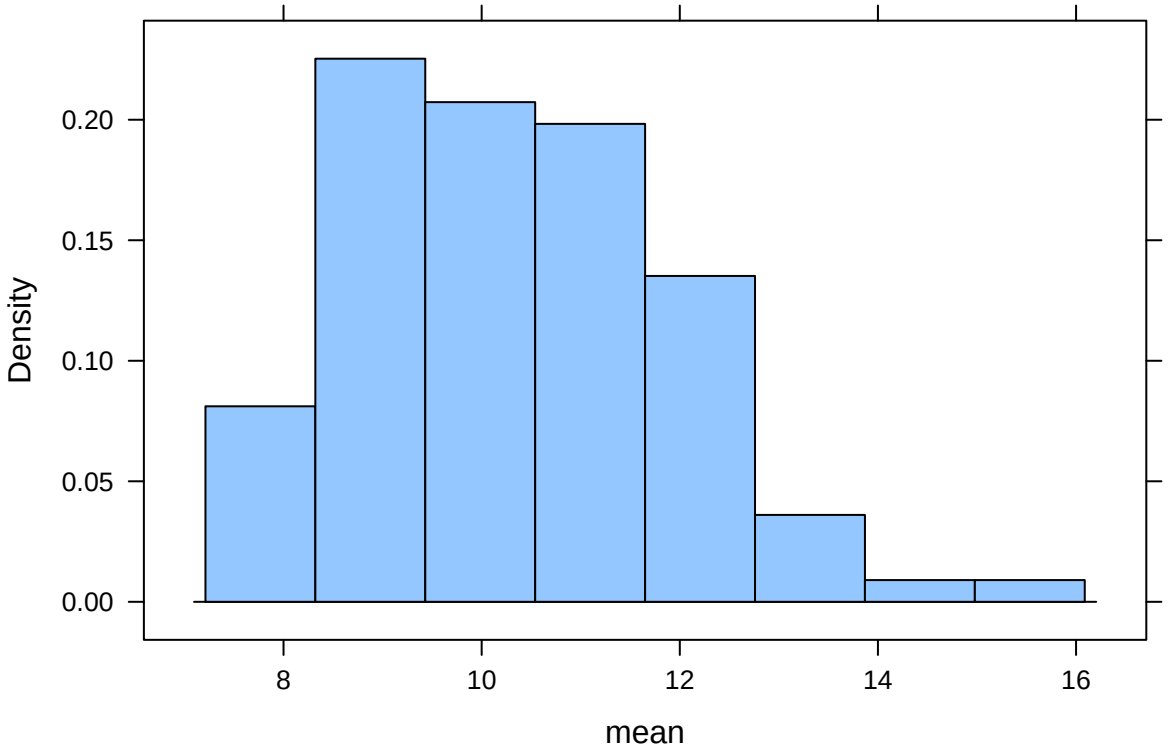
```
# start your code
library(mosaic) # load the package mosaic to use its functions
ceo_salary <-
  read.csv("https://www.siue.edu/~jpailde/CEO_Salary_2012.csv")

#i) Select a random sample of 40 CEO's and compute the sample mean salary and sample standard deviation
sample_40 <- sample_n(ceo_salary, size =  40)
# Original Mean
mean( ~ Annual.pay, data = sample_40)
```

```
## [1] 10.16408
```

```
# Original SD
sdOrig = (sd( ~ Annual.pay, data = sample_40))
sdOrig
```

```
## [1] 8.64106
```

```
# ii) Use the `do()` function to compute the sample mean of 100 randomly selected samples.  Plot the hi
samples <- do (100) * mean( ~ Annual.pay, data = sample_n(ceo_salary, size = 40))
histogram( ~ mean, data = samples)
```




```
#iii) Compute the average and standard deviation of these 100 sample means. Are the values what you exp
meanSamples <- mean( ~ mean, data = samples)
sdSamples <- sd( ~ mean, data = samples)
# Yes, it is as expected for the sd, this is very close to what is exepected. as we calculated by the o
# for the mean too. this is , this is close to what one would expect. as the mean of the initial sample

# last R code line
```

2. Rockwell hardness of pins of a certain type is known to have a mean value of 50 and standard deviation of 1.2.

   i) If the distribution is normal, what is the probability that the sample mean hardness for a random sample of 9 pins is at least 51?

   ii) Without assuming population normality, what is the (approximate) probability that the sample mean hardness for a random sample of 40 pins is at least 51?

Code chunk

```
# start your code
# i) If the distribution is normal, what is the probability that the sample mean hardness for a random
1 - pnorm(q = 51,
          mean = 50,
          sd = 1.2 / sqrt(9))
```

```
## [1] 0.006209665
```

```
# ii) Without assuming population normality, what is the  (approximate) probability that the sample mea
1 - pnorm(q = 51,
          mean = 50,
          sd = 1.2 / sqrt(40))
```

```
## [1] 6.804011e-08
```

3. Suppose that a random sample of size 64 is to be selected from a population with mean 40 and standard deviation 5.

   i) What are the mean and standard deviation of the sampling distribution of the sample mean?

   ii) What is the approximate probability that the sample mean will be within 0.5 of the population mean?

Code chunk

```
# start your code

# i) What are the mean and standard deviation of the sampling distribution of the sample mean?
sample_man = 40
sample_sd = 5 / sqrt(64)

# ii) What is the approximate probability that the sample mean will be within 0.5 of the population mea
pnorm(q = 40.5, mean = sample_man, sd = sample_sd) - pnorm(q = 39.5, mean =
                                                             sample_man, sd = sample_sd)
```

```
## [1] 0.5762892
```

```
# last R code line
```