Laboratory Exercise Week 9

Brian Tipton - STAT-380-001

10/18/23

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. Medical research has shown that repeated interval for extensions beyond 20 degrees increases the risk of wrist and hand injuries. Each of 24 students at Cornell University used a proposed new computer mouse design, and while using the mouse, each student's wrist extension was recorded.

```
wrist <- data.frame(ID = 1:24,
          extension = c(27, 28, 24, 26, 27, 25, 25, 24, 24, 24, 25, 28, 22, 25, 24, 28, 27, 26, 31, 25,
```

i) Compute the mean and standard deviation of the wrist extensions data above. Describe the sample using these summaries.

ii) Create a histogram and QQ-plot of the wrist extensions. Is the normality assumption reasonable?

iii) Use the data to estimate the mean wrist extension for people using this new mouse design using a 97% confidence interval. Provide an interpretation of this interval.

iv) Use the data to test the hypothesis that the mean wrist extension for people using this new design is greater than 20 degrees. Use alpha = 0.01 (1%) level of significance.

Code chunk

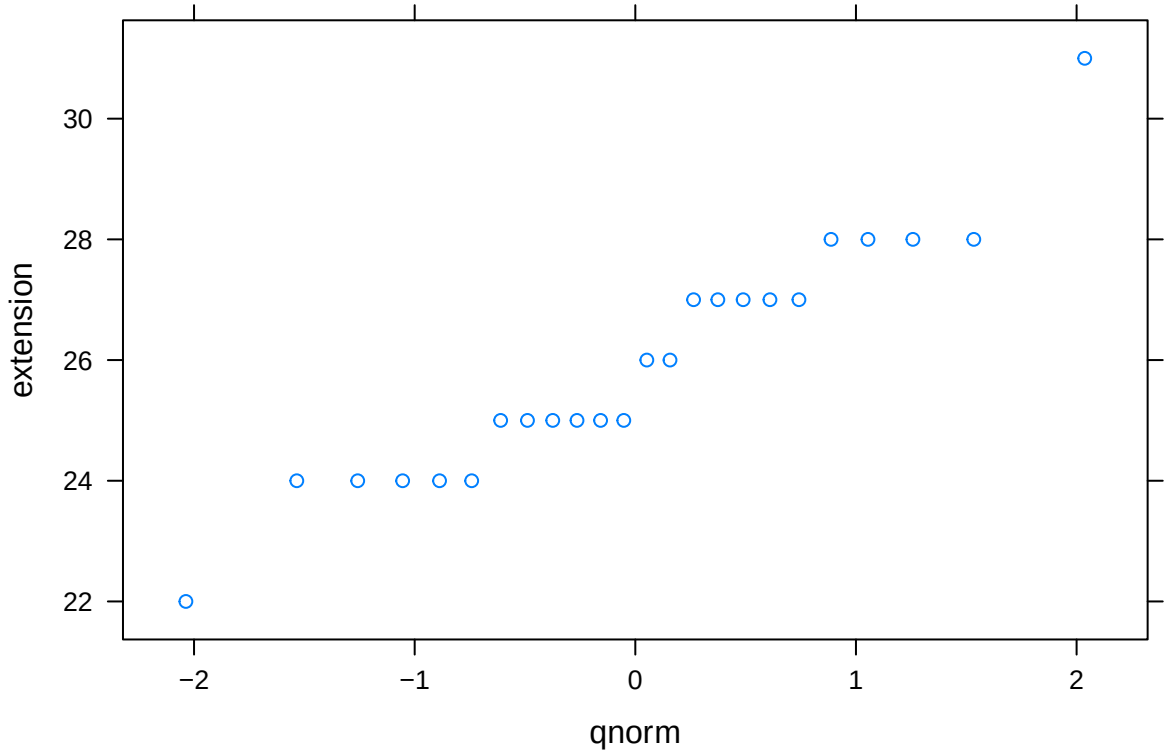
```
# start your code
# i) Compute the mean and standard deviation of the wrist extensions data above.
# Describe the sample using these summaries.
favstats( ~ extension, data = wrist)
```

```
##  min    Q1 median Q3 max     mean       sd  n missing
##   22 24.75   25.5 27  31 25.91667 1.954185 24       0
```

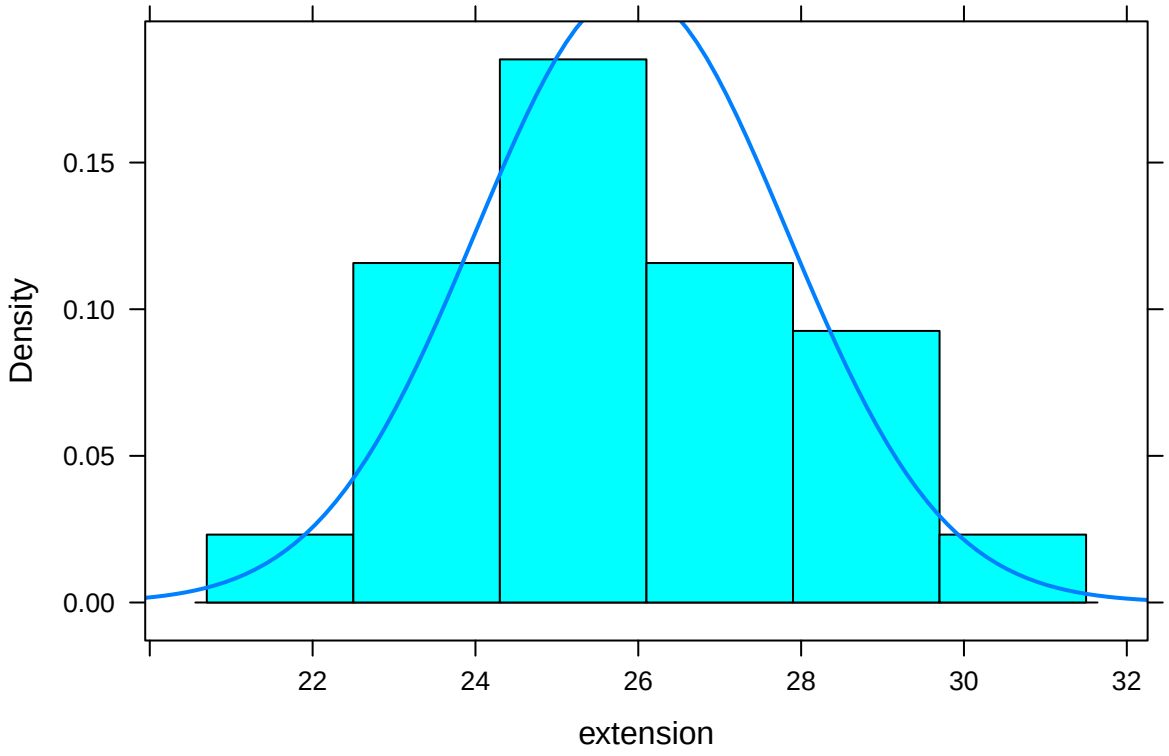
```
cat("The sample mean is approximately 25.92 while the standard deviation is 1.95.\nThis standard deviati
```

```
## The sample mean is approximately 25.92 while the standard deviation is 1.95.
## This standard deviation shows that the values in the Wrist dataset are fairly close
## to the mean.
```

```
# ii) Create a histogram and QQ-plot of the wrist extensions.
# Is the normality assumption reasonable?
qqmath(~ extension, data = wrist)
```




```
histogram(~ extension, data = wrist, fit = "normal")
```




```
cat("Yes the normality fit is reasonable it fits on the curve well.")
```

```
## Yes the normality fit is reasonable it fits on the curve well.
```

```
# iii) Use the data to estimate the mean wrist extension for people
# using this new mouse design using a 97\% confidence interval.
# Provide an interpretation of this interval.
t.test(~ extension, data = wrist, conf.level=0.97)
```

```
##
##  One Sample t-test
##
## data:  extension
## t = 64.971, df = 23, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 97 percent confidence interval:
##  24.99393 26.83941
## sample estimates:
## mean of x
##  25.91667
```

```
cat("The range is 26.83941 - 24.99393 = 1.8455 and the mean is 25.91667.")
```

```
## The range is 26.83941 - 24.99393 = 1.8455 and the mean is 25.91667.
```

```
# iv) Use the data to test the hypothesis that the mean wrist extension for people using
# this new design is greater than 20 degrees.
# Use alpha = 0.01 (1\%) level of significance.
```

```r
t.test(~ extension, data = wrist, alternative = 'greater', mu = 20, conf.level = 0.99)
```

```
##
##  One Sample t-test
##
## data:  extension
## t = 14.833, df = 23, p-value = 1.447e-13
## alternative hypothesis: true mean is greater than 20
## 99 percent confidence interval:
##  24.91948      Inf
## sample estimates:
## mean of x
##  25.91667
```

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

2. Recall the Going Wireless data first mentioned Week 2 of this class. The article Going Wireless (AARP Bulletin, June 2009) reported the estimated percentage of house- holds with only wireless phone service (no land line) for the 50 U.S. states and the District of Columbia. In the accompanying data table, each state was also classified into one of three geographical regions—West (W), Middle states (M), and East (E). Consider only the variable `Wireless` in this data.

```r
wireless.data <- read.csv("https://goo.gl/72BKSf", header = TRUE)
str(wireless.data)
```

```
## 'data.frame':    51 obs. of  3 variables:
##  $ Wireless: num  13.9 11.7 18.9 22.6 9 16.7 5.6 5.7 20 16.8 ...
##  $ Region  : chr  "M" "W" "W" "M" ...
##  $ State   : chr  "AL" "AK" "AZ" "AR" ...
```

i) Compute the mean and standard deviation of the wireless data above. Describe the sample using these summaries.

ii) Create a histogram and QQ-plot of the wireless data. Is the normality assumption reasonable?

iii) Use the data to estimate the mean wireless percentage per state using a 90% confidence interval. Provide an interpretation of this interval.

iv) Use the data to test the hypothesis that the mean wireless percentage per state is less than 17. Use alpha = 0.05 (5%) level of significance.

Code chunk

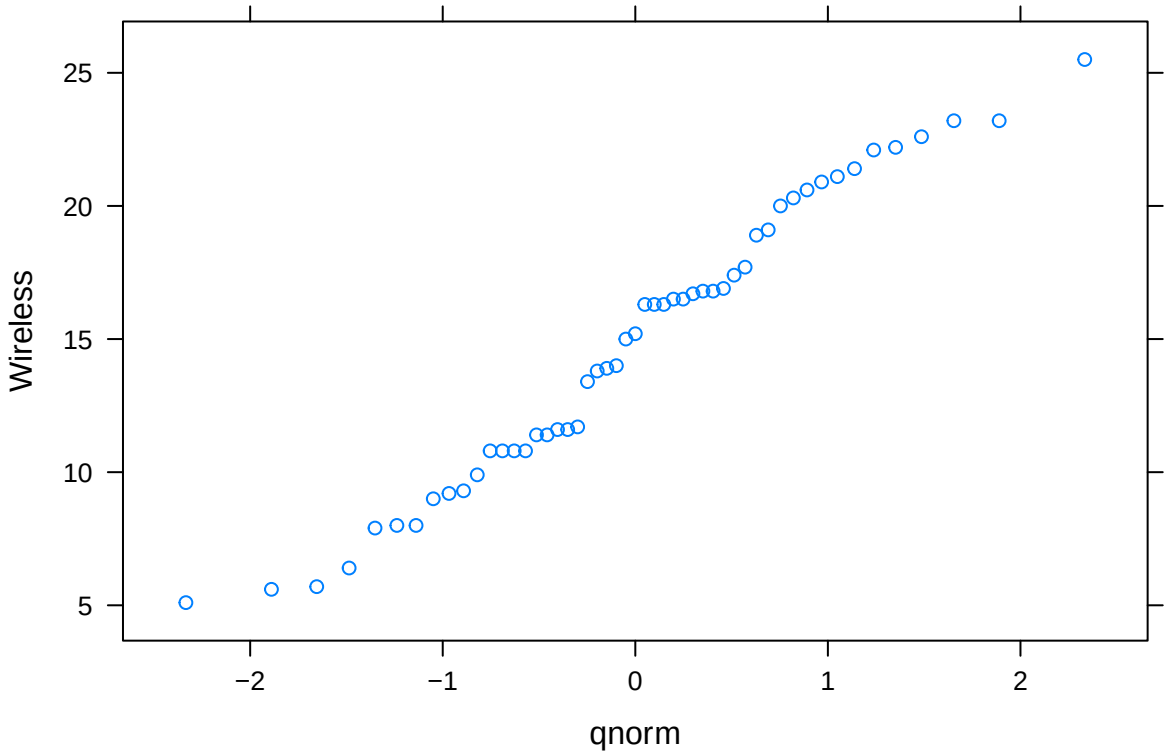
```r
# start your code

# i) Compute the mean and standard deviation of the wireless data above.
# Describe the sample using these summaries.
favstats( ~ Wireless, data = wireless.data)
```

```
##  min   Q1 median Q3  max     mean       sd  n missing
##  5.1 10.8   15.2 19 25.5 14.81569 5.344469 51       0
```

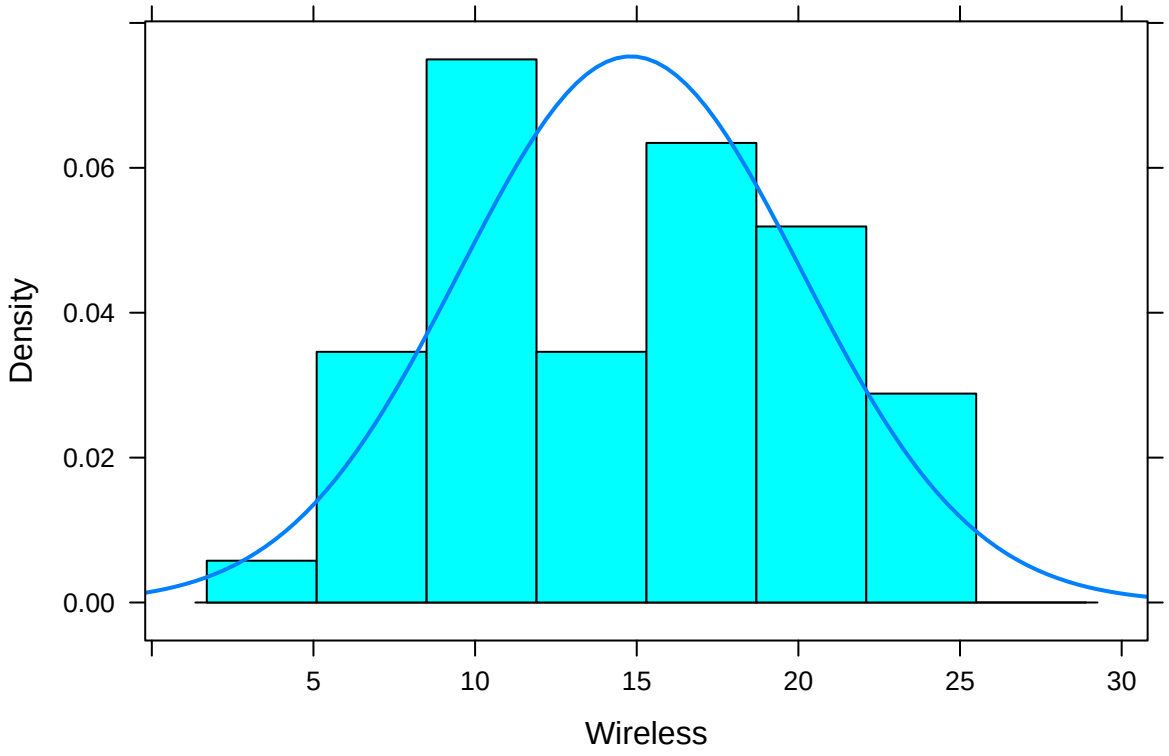
```r
cat("With the mean being approximately 14.82 and a standard deviation of 5.34 shows that\nthere is a de
```

```
## With the mean being approximately 14.82 and a standard deviation of 5.34 shows that
## there is a decent amount of variablility in the Wireless dataset.
```

```r
# ii) Create a histogram and QQ-plot of the wireless data.
# Is the normality assumption reasonable?
qqmath(~ Wireless, data = wireless.data)
```




```r
histogram(~ Wireless, data = wireless.data, fit = "normal")
```




```
cat("No the normality assumption is not reasonable, it has two peaks.")
```

```
## No the normality assumption is not reasonable, it has two peaks.
```

```
# iii) Use the data to estimate the mean wireless percentage per state
# using a 90% confidence interval.
# Provide an interpretation of this interval.
t.test(~ Wireless, data = wireless.data, conf.level = 0.90)
```

```
##
##  One Sample t-test
##
## data:  Wireless
## t = 19.797, df = 50, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 90 percent confidence interval:
##  13.56148 16.06989
## sample estimates:
## mean of x
##  14.81569
```

```
cat("The mean is 14.81569 and the range is 16.06989 - 13.56148 = 2.5084")
```

```
## The mean is 14.81569 and the range is 16.06989 - 13.56148 = 2.5084
```

```
# iv) Use the data to test the hypothesis that the mean wireless percentage per state
# is less than 17.
# Use alpha = 0.05 (5%) level of significance.
```

```r
t.test(~ Wireless, data = wireless.data, alternative = 'less', mu = 17, conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  Wireless
## t = -2.9187, df = 50, p-value = 0.002628
## alternative hypothesis: true mean is less than 17
## 95 percent confidence interval:
##      -Inf 16.06989
## sample estimates:
## mean of x
##  14.81569
```

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