lab_two

Brooke Wheeler

1/19/2022

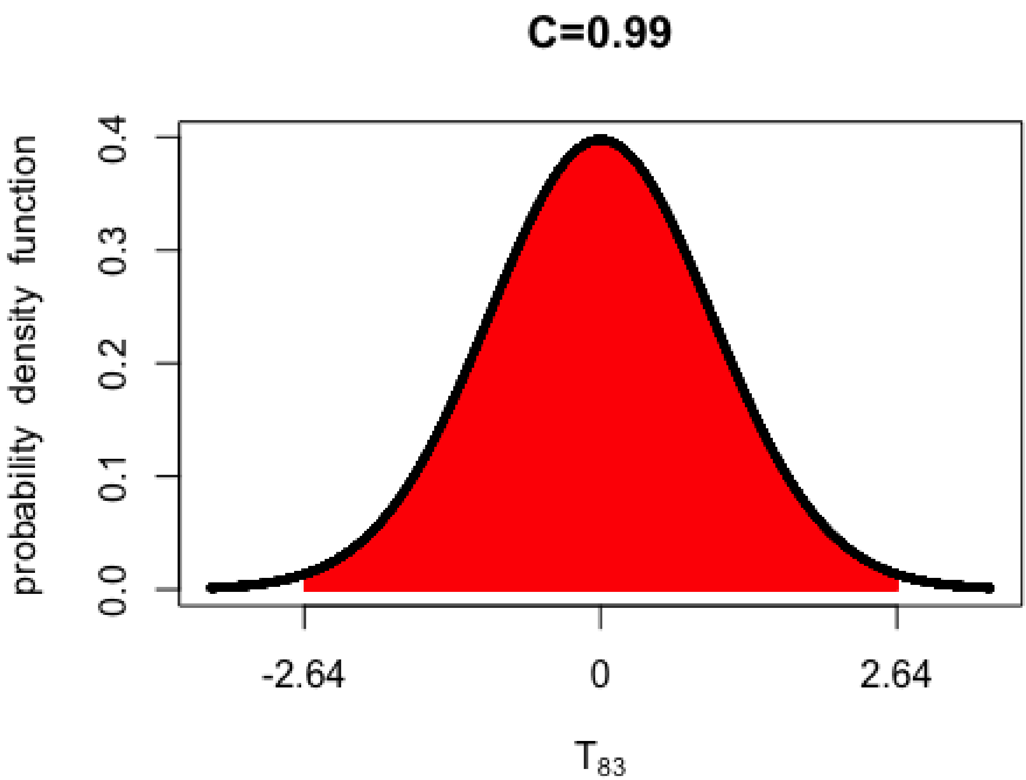
```
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ ggplot2 3.3.5     ✓ purrr   0.3.4
## ✓ tibble  3.1.6     ✓ dplyr   1.0.7
## ✓ tidyr   1.1.4     ✓ stringr 1.4.0
## ✓ readr   2.1.1     ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Part One: Critical Values and confidence intervals

1.Find the following critical values and plot the distribution (with shaded area) to accurately represent the confidence level.

```
library(fastGraph)
shadeDist(c(qt(.005,83), qt(.995,83)), "dt", 83, lower.tail = FALSE, main =
"C=0.99")
```



2.

Use the qt() function to find the exact critical values shown in #1. Report them as t(p;df)=q.

```
qt(.005,83)

## [1] -2.636369

qt(.995, 83)

## [1] 2.636369

# t(.005;83)=-2.64   &   t(.995;83)=2.6
```

Part Two: T tests and F tests

3. save data set as heights

```
getwd()

## [1] "/Users/brookewheeler/Desktop/Regression/Labs"

heights <-read.table("../Data/fballvsbball.csv", header = TRUE, sep = ",")
attach(heights)
```

4. Check to see which t test will be appropriate.

```
var.test(HtFt,HtBk)

##
##  F test to compare two variances
##
## data:  HtFt and HtBk
## F = 1.3575, num df = 44, denom df = 39, p-value = 0.3343
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.7273814 2.5040352
## sample estimates:
## ratio of variances
##            1.35745

# Since the p-value is 0.3343 we fail to reject the null hypothesis that the
variances are equal. Since 1 falls into the interval we can't conclude there
is a statistically sig. difference between the variances. This means that we
will use a test test that assumes the variances of the two samples are equal.
```

5. Based on your conclusions from #4, use R to run the appropriate t test. Carefully think about what your hypotheses should be. Type ?t.test to find out more about the arguments for t.test(). Paste your results below.

```
# Null hypothesis: means are equal Ft =Bk
# Alt. Hypothesis: Ft < BK

t.test(x= heights$HtFt, y= heights$HtBk, c("less"), m= 0, var.equal = TRUE)

##
##  Two Sample t-test
##
## data:  heights$HtFt and heights$HtBk
## t = -3.6841, df = 83, p-value = 0.0002039
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##        -Inf -0.1504837
## sample estimates:
## mean of x mean of y
##  6.178889  6.453250

# p-value=.0002039 , reject the null hypothesis
```

6. Conclusions

We reject the null hypothesis that the mean heights of both basketball and football players are equal at levels of .0002. Meaning that there is statistically significant evidence that football players are not the same height as basketball players.

7. 99% confidence interval

```
t.test(x= heights$HtFt, y= heights$HtBk, m= 0, var.equal = TRUE, conf.level = 
0.99)
```

```
## 
##  Two Sample t-test
## 
## data:  heights$HtFt and heights$HtBk
## t = -3.6841, df = 83, p-value = 0.0004078
## alternative hypothesis: true difference in means is not equal to 0
## 99 percent confidence interval:
##  -0.47069549 -0.07802674
## sample estimates:
## mean of x mean of y 
##  6.178889  6.453250
```

8. 99% confidence interval manually

```
mean(heights$HtFt)
```

```
## [1] 6.178889
```

```
mean(heights$HtBk, na.rm = TRUE)
```

```
## [1] 6.45325
```

```
sd(heights$HtFt)
```

```
## [1] 0.3660987
```

```
sd(heights$HtBk, na.rm = TRUE)
```

```
## [1] 0.3142218
```

```
# critical t values from above -2.64, 2.64
```

Below is a screenshot of my formulas by hand.

The 99% confidence interval is (-.467,-.075)

1 = Ft
2 = Bu

$\bar{x}_1 = 6.179 \quad s_1 = .366 \quad n_1 = 45$

$\bar{x}_2 = 6.45 \quad s_2 = .314 \quad n_2 = 40$

$$\underline{(\bar{x}_1 - x_2) \pm t\left(\tfrac{\alpha}{2}; n_1 + n_2 - 2\right)\sqrt{sp^2\left(\tfrac{1}{n_1} + \tfrac{1}{n_2}\right)}}$$

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$$

$$Sp^2 = \frac{(45-1)(.366^2) + (40-1)(.314^2)}{(45-1) + (40-1)}$$

$$Sp^2 = \frac{5.89 + 3.85}{83}$$

$$Sp^2 = .117$$

$$(6.179 - 6.45) \pm 2.64\sqrt{.117\left(\tfrac{1}{45} + \tfrac{1}{40}\right)}$$

$$-.2.71 \pm 2.64\sqrt{.0055}$$

$$-2.71 \pm .196$$

$$(-2.514, -2.906)$$

$$\cancel{(-.075, -.467)}$$

$$(-.467, -.075)$$