Module 4 R Practice

Angel Waters

2022-06-17

Cats Body Weights

To understand if there are any differences between the body weights of male cats and female cats, data was analyzed from the R Package MASS to understand statistical differences between the two data sets.

The data was subsetted into two groups based on their sex: male cats and female cats.

```
#cats #unquote to view the data
female <- cats %>% filter(Sex=="F")
head(female)
```

```
##   Sex Bwt Hwt
## 1   F 2.0 7.0
## 2   F 2.0 7.4
## 3   F 2.0 9.5
## 4   F 2.1 7.2
## 5   F 2.1 7.3
## 6   F 2.1 7.6
```

```
summary(female)
```

```
##  Sex         Bwt             Hwt
##  F:47   Min.   :2.00   Min.   : 6.300
##  M: 0   1st Qu.:2.15   1st Qu.: 8.350
##         Median :2.30   Median : 9.100
##         Mean   :2.36   Mean   : 9.202
##         3rd Qu.:2.50   3rd Qu.:10.100
##         Max.   :3.00   Max.   :13.000
```

```
male <- cats %>% filter(Sex=="M")
head(male)
```

```
##   Sex Bwt  Hwt
## 1   M 2.0  6.5
## 2   M 2.0  6.5
## 3   M 2.1 10.1
## 4   M 2.2  7.2
## 5   M 2.2  7.6
## 6   M 2.2  7.9
```

```
summary(male)
```

```
##  Sex         Bwt             Hwt       
##  F: 0   Min.   :2.0   Min.   : 6.50  
##  M:97   1st Qu.:2.5   1st Qu.: 9.40  
##         Median :2.9   Median :11.40  
##         Mean   :2.9   Mean   :11.32  
##         3rd Qu.:3.2   3rd Qu.:12.80  
##         Max.   :3.9   Max.   :20.50
```

Identify the Null Hypothesis, Alternative Hypothesis, and claim. The claim is that the average body weight of male cats is not equivalent to the average body weight of female cats, and therefore significantly different. This will be analyzed at an $alpha = 0.05$.

$$H_0:\mu_1=\mu_2\\H_1:\mu_1\neq\mu_2$$

Calculate the necessary sample statistics for the the t test.

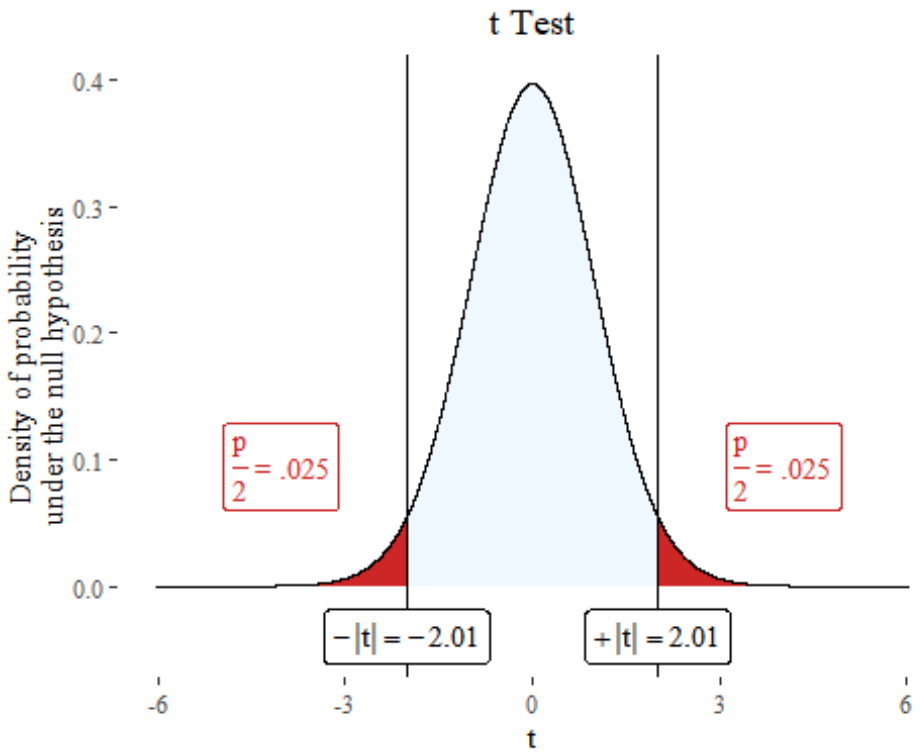
```
nm <- nrow(male)
nf <- nrow(female)
xbarm <- mean(male$Bwt)
xbarf <- mean(female$Bwt)
sm <- sd(male$Bwt)
sf <- sd(female$Bwt)
alpha <- 0.05
dfm <- nm-1
dff <- nf-1
df <- min(dfm, dff) #df to be used
```

This is a two tailed test because it is trying to disprove equivalence with non-equivalence, and not focus on greater than or lesser than results. The critical values were calculated and plotted.

```
halpha <- alpha/2
cv <- qt(p=1-halpha, lower.tail = TRUE, df=df)
cv
```

```
## [1] 2.012896
```

```
#draw curve set to default as a two tailed t test
plotttest(cv, df=df)
```



t test was used to calculate the significance between the two sample averages.

```
##Manual calculation of the t statistic
t <- ((xbarm-xbarf)-(0))/sqrt(((sm^2)/nm)+((sf^2)/nf))
t

## [1] 8.709488

##using the R tool to calculate the t statistic
tsummary <- t.test(female$Bwt, male$Bwt, alternative="two.sided", var.equal =
FALSE, mu=0, conf.level = 0.95)
tsummary

##
##  Welch Two Sample t-test
##
## data:  female$Bwt and male$Bwt
## t = -8.7095, df = 136.84, p-value = 8.831e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.6631268 -0.4177242
## sample estimates:
## mean of x mean of y
##  2.359574  2.900000
```

Based on the summary statistics, the decision around the hypothesis test is:

```
## [1] "Reject the Null Hypothesis"
```

Conclusion

Based on the analysis presented above the following results were summarized about the differences in body weight between male and female cats:

```
## There is enough evidence to support the claim: The average body weight of
male cats is not equivalent to the average body weight of female cats.
```

Meditation Effects on Sleep Quality

A meditation seminar was given to 10 students to see if the workshop improved sleep quality for these students. A paired student's t test was used to analyze if the difference in sleep quality before and after the meditaion workshop was significant and therefore the quality of sleep for the student's improved.

```
before <- c(4.6,7.8,9.1,5.6,6.9,8.5,5.3,7.1,3.2,4.4)
after <- c(6.6,7.7,9.0,6.2,7.8,8.3,5.9,6.5,5.8,4.9)
```

Identify the Null Hypothesis, Alternative Hypothesis, and claim. The claim is that the sleep quality of the students improved after they took the meditation workshop. This will be analyzed at an $alpha = 0.05$.

$$H_0:\mu_D=0\\H_1:\mu_D<0$$

The initial step for gathering the necessary sample statistics is to calculate the individual differences for each pair of points.

```
##    before after    D   D2
## 1     4.6   6.6 -2.0 4.00
## 2     7.8   7.7  0.1 0.01
## 3     9.1   9.0  0.1 0.01
## 4     5.6   6.2 -0.6 0.36
## 5     6.9   7.8 -0.9 0.81
## 6     8.5   8.3  0.2 0.04
## 7     5.3   5.9 -0.6 0.36
## 8     7.1   6.5  0.6 0.36
## 9     3.2   5.8 -2.6 6.76
## 10    4.4   4.9 -0.5 0.25
```

This is a left tailed t test which means the critical value will be calculated via the one-tailed method.

```
n <- nrow(Meditation)
alpha <- 0.05
df <- n-1
cv <- qt(p=1-alpha, df=df, lower.tail = FALSE )
```

The summary statistics were then calculated for the t test.

```
dbar <- mean(Meditation$D)
sumd <- sum(Meditation$D)
sumd2 <- sum(Meditation$D2)
sD <- sqrt(((n*sumd2)-(sumd)^2)/(n*(n-1)))
t <- dbar/(sD/sqrt(n))
```

Based on the t statistic, the following decision was made about the sleep scores:

```
## [1] "Reject the Null Hypothesis"
```

Conclusion

Based on the analysis above, the conclusion can be made around the claim:

```
## There is enough evidence to support the claim: Meditation increases the
average sleep score
```

The method for analyzing the meditation dataset was the two-sample paired t test. This was identified as the appropriate mode for analysis because the data was a before and after study. Each point collected was dependent on a matched point, the student's sleep quality analyzed before the meditation class and the same student's sleep quality analyzed after the meditation class. The aim of the study is to look at the difference therefore making the samples matched-pairs.