Worksheet2

Peyton Hall

01/24/2025

Load Necessary Libraries

```
library(readxl)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

$H_0 : \mu = 80$ vs $H_a : \mu < 80$ In Class Coding One Sample T Test

```
Reading <- read_excel("~/Desktop/STAT 301/Week 2/Reading.xlsx")
t.test(Reading$read, mu = 80, alternative = "less")
```

```
##
##  One Sample t-test
##
## data:  Reading$read
## t = -3.87, df = 29, p-value = 0.0002844
## alternative hypothesis: true mean is less than 80
## 95 percent confidence interval:
##      -Inf 72.50205
## sample estimates:
## mean of x
##  66.63333
```

t = -3.87, p-value = 0.0003 Reject H0; There is evidence to support that the mean reading score from Prof. White's students is significantly lower than 80

$H_0 : \mu = 0.5$ vs $H_a : \mu > 0.5$ Worksheet 2 Question 2

```r
Mehr_Song_and_Spelke <- read_excel("~/Desktop/STAT 301/Week 2/Mehr Song and Spelke.xlsx")

M1 <- Mehr_Song_and_Spelke %>%
  filter(exp1 == 1)

t.test(M1$Baseline_Proportion_Gaze_to_Singer, mu = 0.5, alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  M1$Baseline_Proportion_Gaze_to_Singer
## t = 0.67438, df = 31, p-value = 0.2525
## alternative hypothesis: true mean is greater than 0.5
## 95 percent confidence interval:
##  0.4680553       Inf
## sample estimates:
## mean of x
## 0.5210967
```

t = 0.67; p-value = 0.25 Fail to reject H0; There is no evidence to support that the average proportion of infants looking for the singer was higher than 0.5.

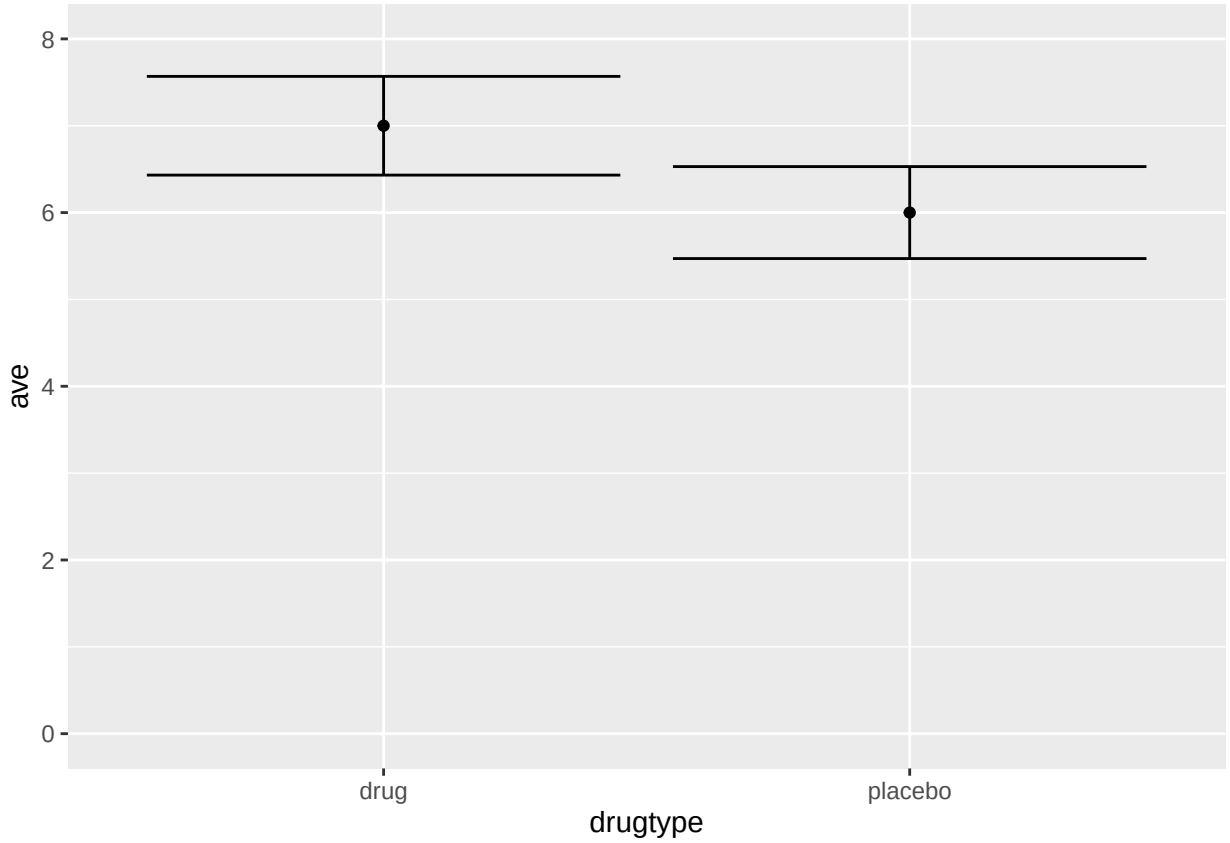
$H_0 : \mu(drug) = \mu(placebo)$ vs $H_a : \mu(drug) > \mu(placebo)$ In Class Coding Slide 11

```r
drugtype <- rep(c("drug", "placebo"), c(10,12))

birthwt <- c(6.9, 7.6, 7.3, 7.6, 6.8, 7.2, 8.0, 5.5, 5.8, 7.3,
             6.4, 6.7, 5.4, 7, 5.3, 6.6, 5.8, 5.7, 6.2, 7.1, 5.6, 4.2)
babydf <- data.frame(drugtype, birthwt)
babydf
```

```
##    drugtype birthwt
## 1      drug     6.9
## 2      drug     7.6
## 3      drug     7.3
## 4      drug     7.6
## 5      drug     6.8
## 6      drug     7.2
## 7      drug     8.0
## 8      drug     5.5
## 9      drug     5.8
## 10     drug     7.3
## 11  placebo     6.4
## 12  placebo     6.7
## 13  placebo     5.4
## 14  placebo     7.0
## 15  placebo     5.3
## 16  placebo     6.6
## 17  placebo     5.8
## 18  placebo     5.7
## 19  placebo     6.2
## 20  placebo     7.1
## 21  placebo     5.6
## 22  placebo     4.2
```

```
library(ggplot2)
newdf <- babydf %>%
  group_by(drugtype) %>%
  summarize(ave = mean(birthwt), ste = sd(birthwt)/sqrt(length(birthwt)),
            tstar = qt(1-0.05/2, length(birthwt) - 1))
ggplot(data = newdf, aes(x = drugtype, y = ave) )+ geom_point() + geom_errorbar(aes(ymin = ave - tstar
```



$H_0 : \mu(drug) = \mu(placebo)$ vs $H_a : \mu(drug) > \mu(placebo)$ In Class Coding Slide 11 T Test

```
# syntax implies  (drug) > (placebo)
t.test(birthwt~drugtype, data = babydf, alternative = "greater")
```

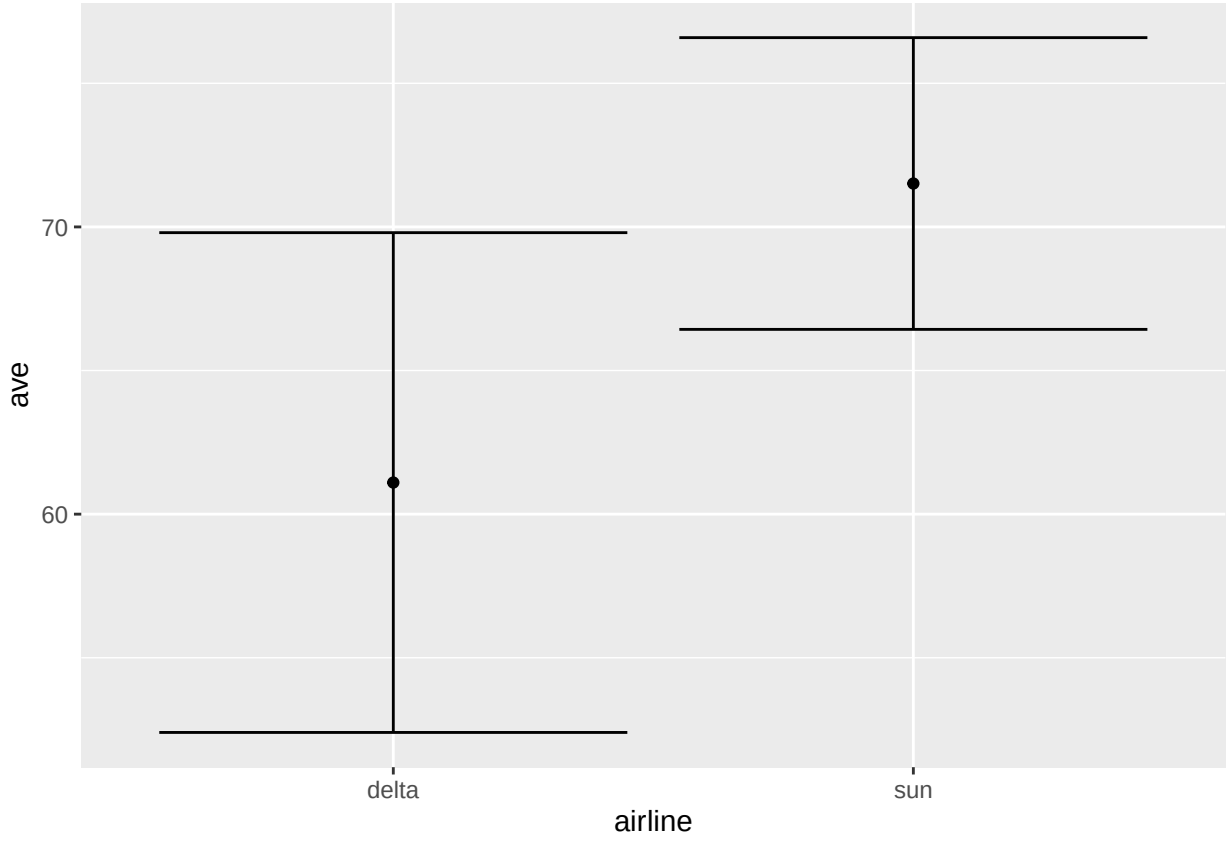
```
##
##  Welch Two Sample t-test
##
## data:  birthwt by drugtype
## t = 2.8749, df = 19.594, p-value = 0.004747
## alternative hypothesis: true difference in means between group drug and group placebo is greater thar
## 95 percent confidence interval:
##  0.3994771       Inf
## sample estimates:
##    mean in group drug mean in group placebo
##                     7                     6
```

t = 2.87; p-value = 0.0047 Reject H0; there is evidence to support that the average baby weight from the drug group is significantly higher than the average baby weight from the placebo group

```r
airline <- rep(c("delta", "sun"), each = 9)

fuel <- c(82.1, 68.4, 52.7, 58.8, 52, 71.4, 60, 59.6, 44.9,
          60.8, 68.3, 77.2, 76.4, 71.4, 63, 81.1, 72.1, 73.3)
airdf <- data.frame(airline, fuel)
```

```r
airdf %>%
  group_by(airline) %>%
  summarize(ave = mean(fuel), ste = sd(fuel)/sqrt(length(fuel)), tstar = qt(1-0.05/2, length(fuel)-1)) %
  ggplot(aes(x=airline, y= ave)) + geom_point() + geom_errorbar(aes(ymin=ave - tstar*ste, ymax=ave+tstar
```




```r
t.test(fuel~airline, data=airdf, alternative = "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  fuel by airline
## t = -2.3831, df = 12.885, p-value = 0.03327
## alternative hypothesis: true difference in means between group delta and group sun is not equal to 0
## 95 percent confidence interval:
##  -19.8577998  -0.9644224
## sample estimates:
## mean in group delta   mean in group sun
##            61.10000            71.51111
```

Reject H_0 There is evidence to support that the ave. fuel efficiency differs sig. between 2 airlines