STAT 587 - Homework 1

Martin Simonson

January 25, 2019

1.) SAT Mathematics Scores

A person claims that the current SAT Mathematics (SATM) test scores are over-estimates of the ability of typical high school seniors because not all students take the test. They further claim that if the test became required for all seniors, the mean score would be no more than 450.

(a)

- H_0: The mean SATM score = 450;
- H_a: The mean SATM score $\neq$ 450

(b)

```
test.stat<-((460.56-450)/(98.65/sqrt(500)))
t<-round(test.stat,3)
t
```

```
## [1] 2.394
```

- The test statistic is 2.394 with 499 degrees of freedom

(c)

```
p<-round((pt(-abs(t),df=499)),3)
p
```

```
## [1] 0.009
```

- The p-value from the t-table for a two-tailed test with a t-statistic of 2.394 with 499 degrees of freedom is bounded between .02 and .001, and the exact p-value as calculated by R is 0.009.

(d)

- Given the sample data we can conclude that there is *strong* evidence to reject the null hypothesis and thus data supports the alternative hypothesis that the mean SATM test scores are not equal to 450.

2.) Something Fishy in the Cheese Factory

Cheese admins suspect that their milk supplier is watering down milk to increase profits. The freezing point of natural milk is $-0.545°C$. Adding water will raise the freezing temperature. The freezing temperature of 15 lots of suppliers milk is measured. The average of the 15 measurements is $\bar{X} = -0.538°C$ with standard deviation of $s = 0.008°C$. Is this sufficient statistical evidence that the supplier is adding water to the milk?

(a)

- H_0: The freezing point of the supplier's milk = $-0.545°C$;
- H_a: The freezing point of the supplier's milk $\neq -0.545°C$

(b)

```
test.stat<-((-.538- -.545)/(0.008/sqrt(15)))
t<-round(test.stat,3)
t
```

```
## [1] 3.389
```

- The test statistic is 3.389 with 14 degrees of freedom

(c)

```
p<-round((pt(-abs(t),df=14)),3)
p
```

```
## [1] 0.002
```

- The p-value from the t-table for a two-tailed one sample t-test with a t-statistic of 3.389 with 14 degrees of freedom is bounded between 0.01 and and 0.002, and the exact p-value as calculated by R is 0.002.

(d)

- Given this sample data we can conclude that there is *strong* evidence to reject the null hypothesis that the supplier's milk freezes at the same temperature as natural milk and supports the alternative hypothesis that the supplier's milk is watered down and freezes at a temperature not equal to $-0.545°C$.

3.) VW Gas Mileage

The VW's on-board computer was reading a better mile per gallon (MPG) than advertised for that model. Therefore, the computer's MPG and driver-recorded MPG recorded for each fill-up in 2016. The data provided in *mpg.txt* shows the differences between the computer's MPG calculations and the driver's MPG calculations.

(a)

```
data<-read.csv("Data/mpg..csv")
round(mean(data$difference),3)
```

```
## [1] 2.73
```

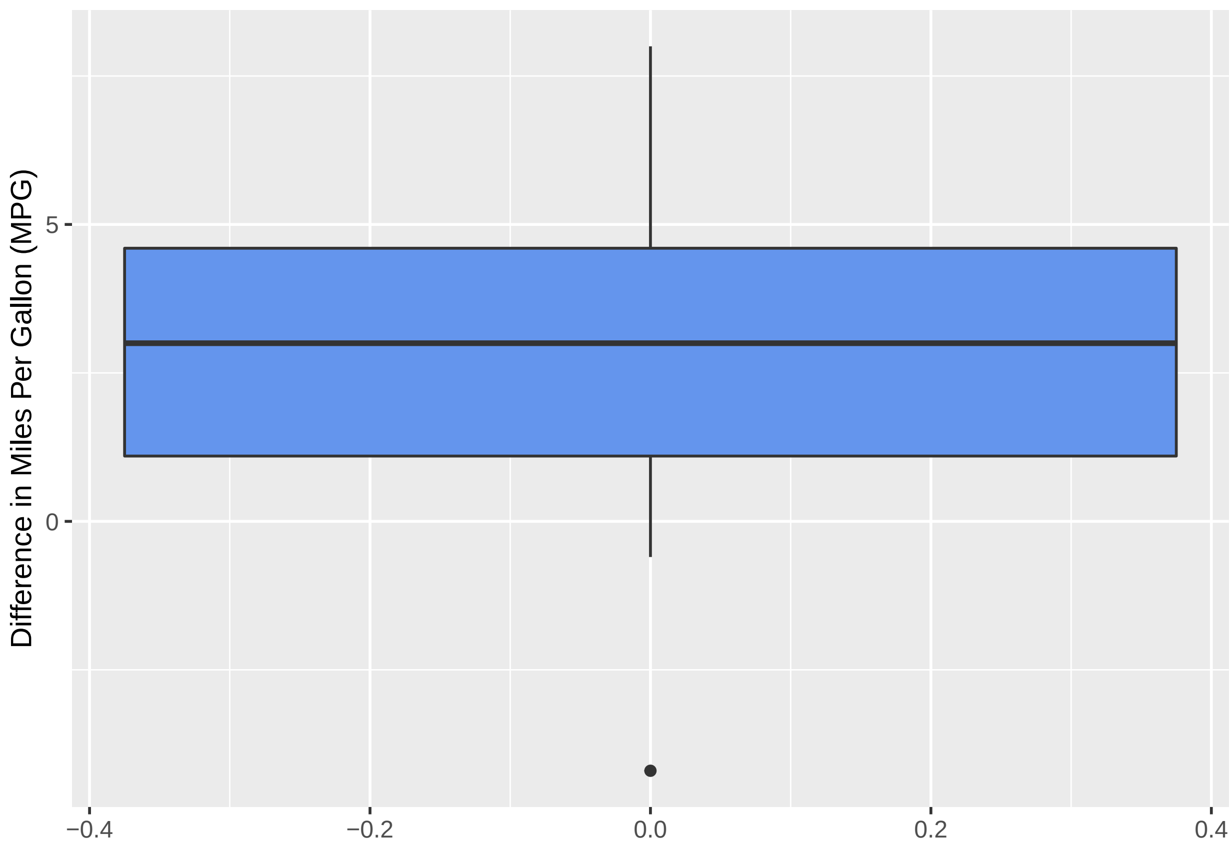
```
round(sd(data$difference),3)
```

```
## [1] 2.802
```

- The sample mean of the difference is 2.73 miles per gallon with a standard deviation of 2.802 miles per gallon.

(b)

```
library(ggplot2)
ggplot(data=data,aes(y=difference))+
       geom_boxplot(fill="cornflowerblue")+
       ylab("Difference in Miles Per Gallon (MPG)")
```



(c)

- H_0: The difference in MPG $= 0$;
- H_a: The difference in MPG $\neq 0$

(d)

```
t.test(data$difference, mu=0)
```

```
##
##  One Sample t-test
##
## data:  data$difference
## t = 4.358, df = 19, p-value = 0.0003386
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.418847 4.041153
## sample estimates:
## mean of x
##      2.73
```

(e)

- We can conclude from the two-tailed one-sample t-test that there is very strong evidence to reject the null hypothesis that there is no difference between the MPG provided by the VW onboard computer and the driver's calculation, and that the data shows very strong evidence supporting the alternative hypothesis that the true mean difference in MPG values between the VW computer and the driver's calculation is not equal to 0.