HW1

Swetha Jayapathy

10/6/2020

R Markdown

Question 1

Exercise 3 a. Cause and effect cannot be inferred from the provided information, there might be other major factors causing it.

 b. Confounding variables which may be responsible are • Work conditions • Their food habits • Wealth • The people they stay with

Exercise 4:

Yes, others should question the claim whether this is as good as a randomized experiment. It is mentioned that the rats are assigned haphazardly which may not be a proper randomized experiment as there is a bias in this. The researcher would have selected the rats which were weak and so was easier to be pulled out. A healthier rat would have easily escaped from being pulled out. Hence, it is better to use a proper randomization technique to come to a conclusion.

Question 2:

Scope of Inference : Trace OSU study selects the volunteers from a group of enrollees who may be either a Faculty, Staff or a Student from the specified locations. Therefore, the selection of the subject is the main problem. We cannot assume that people from one group gets higher priority in selection than the people in another group. For instance, we can't infer that a greater number of Faculties would be selected than the number of Students as they are being randomly selected. There are equal chances of the individuals from each group to get selected. Hence the result of the study would apply to the entire population that is sampled.

Question 3

```r
library(Sleuth3) # Load the Sleuth3 package
library(ggplot2) # and the ggplot2 package
```

3a - Produce side-by-side boxplots for the two groups

```r
qplot(Group, Lifetime, data=ex0211, geom="boxplot")
```
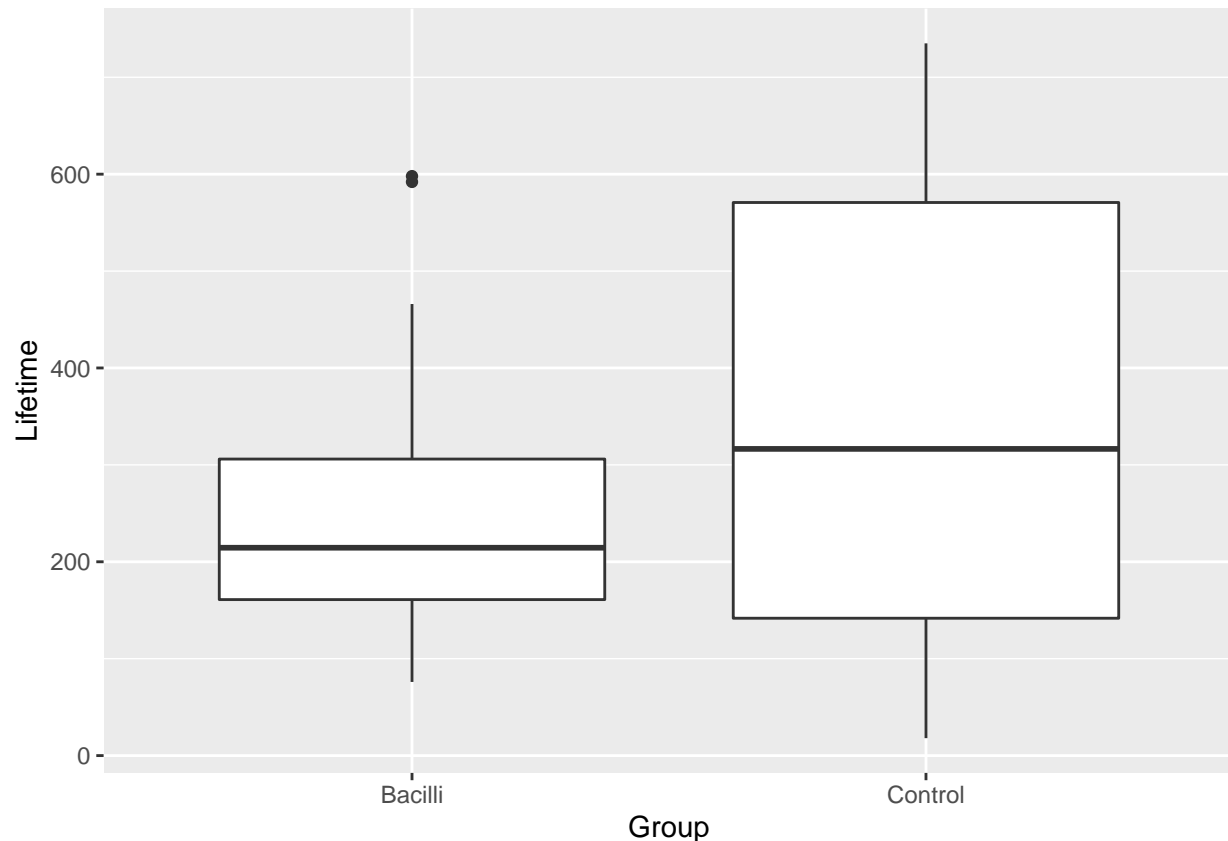
3b Difference in Sample means : = Mean of Control Group – Mean if Bacilli Group

$= 345.23 - 242.53 = 102.7$

Let us consider the Null Hypothesis to be no difference in the lifetimes between the two groups. As we can see the observed test statistics is in the extreme end of the histogram, and seems to be very unusual if H0 is true. This says how inconsistent our null hypothesis is, hence, our H0 is not plausible.

3C

```
#3c - t test with one sided confidence interval
t.test(Lifetime~Group, alternative="less", data=ex0211, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  Lifetime by Group
## t = -3.1411, df = 120, p-value = 0.00106
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf -48.50152
## sample estimates:
## mean in group Bacilli mean in group Control
##              242.5345              345.2344
```

```
# With 2 sided confidence interval
t.test(Lifetime~Group, data=ex0211, var.equal=TRUE)
```

```
## 
##  Two Sample t-test
## 
## data:  Lifetime by Group
## t = -3.1411, df = 120, p-value = 0.00212
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -167.43548  -37.96431
## sample estimates:
## mean in group Bacilli mean in group Control
##              242.5345              345.2344
```

Statistical conclusion :

The data provide convincing evidence that the mean difference in the Lifetime of the Bacilli group and that of Controlled group are not equal.Average Lifetime of Bacilli group is lesser than that of the Controlled group (p value=0.00106).Therefore, the dose of tubercle bacilli does affect the lifetime of guinea pigs. The mean lifetime of Bacilli group is lesser than the mean lifetime of Controlled group by nearly 102 years(95% confidence interval:-167.43 to -37.96).