Statistical Inference Course Project Part 2 - Basic Inferential Data Analysis

Dushan Yovetich

Overview

In this exercise, we analyze the effect of Vitamin C on Tooth Growth in Guinea Pigs. A total of 60 guinea pigs were used in the study. Each animal received one of three doses of Vitamin C (0.5, 1, 2 mg/day) by one of two delivery methods (Orange Juice or Ascorbic Acid, a form of Vitamic C). The animals were then observed for tooth growth.

Tooth Growth Data

The study data is available in the data sets provided in R. The data is comprised of 60 observations with 3 variables for each observation. The variables are length of odontoblasts (cells responsible for tooth growth), dosage level of vitamin C in miligrams per day, and delivery method (OJ = orange juice and VC = ascorbic acid).

```
# load data
data("ToothGrowth")
dim(ToothGrowth)
```

```
## [1] 60  3
```

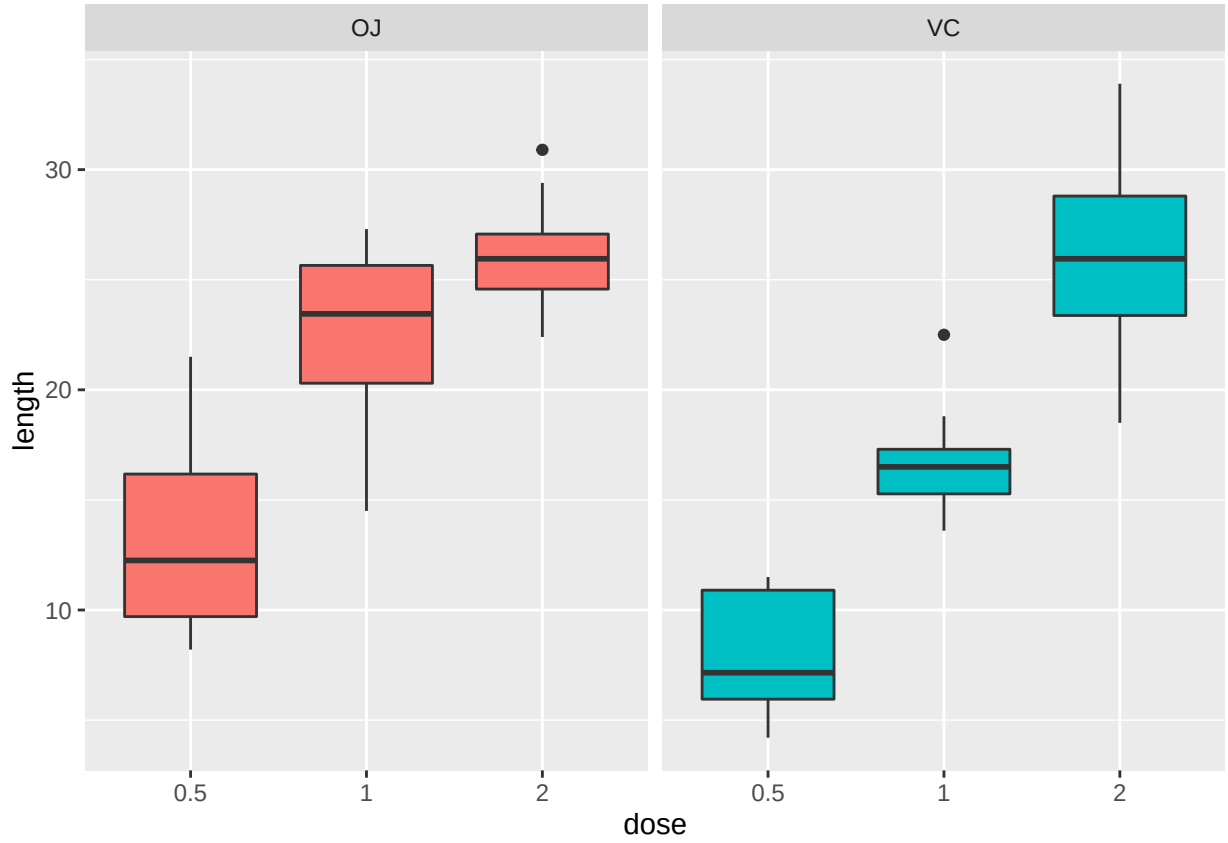
```
# show summary of data
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

Analysis of Tooth Growth Data

```
library(ggplot2)
library(reshape2)
library(dplyr)
```

```
# plot data
ggplot(ToothGrowth,aes(factor(dose),len))+
      geom_boxplot(aes(fill = supp))+ labs(x = "dose",y = "length")+
      facet_grid(.~supp) + theme(legend.position = "none")
```



The data appears to show a positive correlation between dosage and tooth growth. As the dosage of vitamin C is increased, the guinea pigs experienced higher tooth growth. Orange Juice on average appears to have a greater impact in dosages of 0.5 and 1 mg/day versus Ascorbic Acid albeit with more significant variability. However, when the dosage of Vitamin C is set at 2 mg/day, the two delivery methods on average saw similar tooth growth with significant variability in the case of Ascorbic Acid. The following summary shows how close the averages in the data are for the two delivery methods at 2 mg/day.

```
# average length by dose and delivery method
dcast(ToothGrowth,dose~supp,mean,value.var = "len")
```

```
##   dose    OJ    VC
## 1  0.5 13.23  7.98
## 2  1.0 22.70 16.77
## 3  2.0 26.06 26.14
```

Hypothesis Testing

To evaluate the effect of both the dosage level and delivery method to tooth growth, we will use hypothesis testing to look at the relationship

Tooth Growth By Delivery Method

Our first step is testing whether delivery method (Orange Juice versus Ascorbic Acid) has any effect on tooth growth. Ho = 0 is tested against Ha <> 0. We assume a 95% confidence level and use a t-test.

```
test <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth)
```

The t-test provides a p-value of 0.0606345. This exceeds 5% threshold for accepting the alternative hypothesis (Ha <> 0). Further, the value zero is within our calculated confidence interval of -0.1710156 - 7.5710156. As a result, we fail to reject the hypothesis, Ho = 0.

Tooth Growth by Dosage Level

Our second step is testing whether dosage levels have any effect on tooth growth. Ho = 0 is tested against Ha <> 0. Again, we assume a 95% confidence level and use a t-test.

```
# subsets of dosage levels for comparison
set1 <- filter(ToothGrowth, dose %in% c(0.5,1))
set2 <- filter(ToothGrowth, dose %in% c(1,2))
set3 <- filter(ToothGrowth, dose %in% c(0.5,2))

# t-test of each set of dosage levels
test1 <- t.test(len~dose,paired = FALSE, var.equal = FALSE, data = set1)
test2 <- t.test(len~dose,paired = FALSE, var.equal = FALSE, data = set2)
test3 <- t.test(len~dose,paired = FALSE, var.equal = FALSE, data = set3)

# summary of tests
data.frame("p.value" = c(test1$p.value,test2$p.value,test3$p.value),
           "lower.conf" = c(test1$conf.int[1],test2$conf.int[1],test3$conf.int[1]),
           "upper.conf" = c(test1$conf.int[2],test2$conf.int[2],test3$conf.int[2]))
```

```
##        p.value lower.conf upper.conf
## 1 1.268301e-07 -11.983781  -6.276219
## 2 1.906430e-05  -8.996481  -3.733519
## 3 4.397525e-14 -18.156167 -12.833833
```

For each of the t-tests performed, our p-value is below the 5% threshold for accepting the alternative hypothesis (Ha <> 0). Further, the value zero is not included in any of the confidence intervals. As a result, we reject the null hypothesis (Ho = 0).

Tooth Growth by Dosage Level and Delivery Method

Our final step is testing whether delivery method has any effect for each dosage level. Ho = 0 is tested against Ha <> 0. We assume a 95% confidence level and use a t-test.

```
# subsets of data by dosage level
set1 <- filter(ToothGrowth, dose==0.5)
set2 <- filter(ToothGrowth, dose==1)
set3 <- filter(ToothGrowth, dose==2)

# t-test of each set of dosage levels
test1 <- t.test(len~supp,paired = FALSE, var.equal = FALSE, data = set1)
test2 <- t.test(len~supp,paired = FALSE, var.equal = FALSE, data = set2)
test3 <- t.test(len~supp,paired = FALSE, var.equal = FALSE, data = set3)

# summary of tests
```

```
data.frame("p.value" = c(test1$p.value,test2$p.value,test3$p.value),
           "lower.conf" = c(test1$conf.int[1],test2$conf.int[1],test3$conf.int[1]),
           "upper.conf" = c(test1$conf.int[2],test2$conf.int[2],test3$conf.int[2]))
```

```
##       p.value lower.conf upper.conf
## 1 0.006358607   1.719057   8.780943
## 2 0.001038376   2.802148   9.057852
## 3 0.963851589  -3.798070   3.638070
```

For dosage level 0.5 and 1 Mg/day, our p-value is below the 5% threshold for accepting the alternative hypothesis (Ha <> 0). Further, the value zero is not included in the confidence intervals. As a result, we reject the null hypothesis (Ho = 0) in those cases. For the dosage level 2 Mg/day, our p-value far exceeds our 5% threshold for accepting the alternative hypothesis. Further, our confidence interval in this case included the value zero. We fail to reject the null hypothesis in the case of dosages of 2 Mg/day.

Conclusion

In conclusion, we note the following:

- The delivery method (Orange Juice versus Ascorbic Acid) by itself has no overall effect on tooth growth.
- Dosage levels have an effect on tooth growth. Tooth growth increases as the dosage of Vitamin C is increased.
- For dosage levels below 2 Mg/day, Orange Juice is a more effective delivery method than Ascorbic Acid.