Statistical Inference: Data Analysis

Dan Klemfuss

1/8/2021

Table of Contents

1. Overview:

This report is an analysis of the 'ToothGrowth' data set in R, which captures the lenght of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods: orange juice or ascorbic acid (a form of vitamin C).

The objective is to create a hypothesis of the data set, whether there is a significant difference between the two delivery methods by dosage.

2. Data Analysis: Tooth Growth

2.1 Summary of Data

Begin by loading the built-in ToothGrowth dataset:

```
# Load ToothGrowth data set:
data(ToothGrowth)
# Summarize ToothGrowth data set (first 5 rows):
head(ToothGrowth)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

```
# Show dimensions of the data set:
dim(ToothGrowth)
```

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

```
# Show unique values of the fields:
unique(ToothGrowth$supp)
```

```
## [1] VC OJ
## Levels: OJ VC
```

```
unique(ToothGrowth$dose)
```

```
## [1] 0.5 1.0 2.0
```

The data set has 60 entries, with 3 fields.The 'supp' field has 2 factors, and the 'dose' field has 3 values.

2.2 Hypothesis Testing

The test for the data will assess the hypothesis that the mean of the tooth Growth for the delivery methods and dosage is equal. The alternate hypothesis is that the tooth Growth for the delivery methods and dosage is not equal. To perform this assessment, a two-sample t-test will be used. From this, the test statistic, p-value, and 95% confidence interval in the mean values will be computed.

Performing a two-sample T-test:

```
dat <- ToothGrowth
output <- list()
# Iterate through all values of dosage:
for (i in unique(dat$dose)){
  index <- as.character(i)
  oj.len <- (ToothGrowth %>% filter(supp=="OJ" & dose==i))$len # Orange Juice
  vc.len <- (ToothGrowth %>% filter(supp=="VC" & dose==i))$len # Absorbic Acid
  # Run t-test:
  test <- t.test(oj.len, vc.len, paired=F)
  output[[index]] <- data.frame(
    oj.mean = test$estimate[1],
    vc.mean = test$estimate[2],
    lower = test$conf[1],
    upper = test$conf[2]
  )
}

t <- as.data.frame(do.call(rbind, output))
```

Results: Dose = 0.5 mg/day:

For the dose of 0.5 mg/day, the average tooth growth rate for orange juice was 13.23 and for absorbic acid was 7.98. The confidence interval is 1.7190573 to 8.7809427, which does not span zero. Therefore we reject the null hypothesis that the delivery method for the dosage of 0.5 mg/day is equal.

Results: Dose = 1.0 mg/day:

For the dose of 1.0 mg/day, the average tooth growth rate for orange juice was 22.7 and for absorbic acid was 16.77. The confidence interval is 2.8021482 to 9.0578518, which does not span zero. Therefore we reject the null hypothesis that the delivery method for the dosage of 1.0 mg/day is equal.

Results: Dose = 1.5 mg/day:

For the dose of 1.5 mg/day, the average tooth growth rate for orange juice was 26.06 and for absorbic acid was 26.14. The confidence interval is -3.7980705 to 3.6380705, which does span zero. Therefore we fail to reject the null hypothesis that the delivery method for the dosage of 1.5 mg/day is equal.

3. Conclusions

Based on the results, there is a significant difference in tooth growth lengths when using Orange juice and Absorbic acid for doses of 0.5 and 1.0 mg/day. There is no significant difference in the mean tooth growth length between Orange juice and Absorbic acid for doses of 1.5 mg/day.