

Statistical Inference Project 2

Jack Chesson

Overview

The purpose of this analysis is to explore the ToothGrowth dataset from R datasets which includes possible predictors for the growth of teeth in Guinea Pigs. The analysis will check the statistical significance and create confidence intervals for each possible predictor.

Loading Data and Libraries

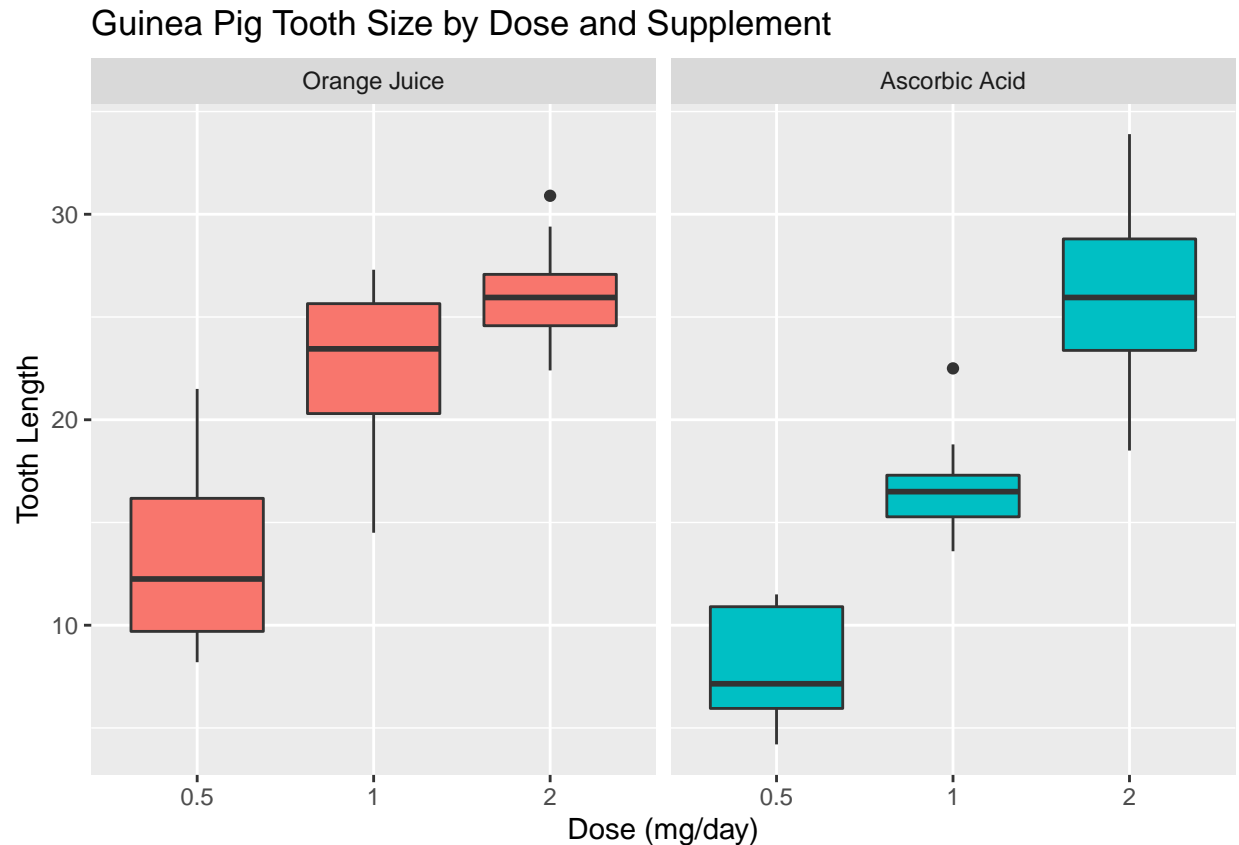
```
library(datasets)
library(data.table)
library(ggplot2)
library(knitr)

toothData <- datasets::ToothGrowth
head(toothData)
```

```
##      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
```

Initial Exploration

```
levels(toothData$supp) = c("Orange Juice", "Ascorbic Acid")
ggplot(toothData, aes(x = factor(dose), y = len)) +
  facet_grid(.~supp) +
  geom_boxplot(aes(fill = supp), show.legend = FALSE)+
  labs(title = "Guinea Pig Tooth Size by Dose and Supplement", x = "Dose (mg/day)", y = "Tooth Length")
```



At a glance, there appears to be a clear positive trend between dose and tooth length. It is unapparent if the supplement taken is significantly affecting the tooth length.

Further Exploration

We will use hypothesis testing using T-Tests in order to determine any significant trends.

Requirements for significance: If the confidence interval is on a set that does not contain zero and if the p-value, at a significance level of 95%, is smaller than 0.05 the null Hypothesis can be rejected.

Hypothesis: The two supplements deliver the same tooth growth according to the data as a whole

```
H1 <- t.test(len~supp, data = toothData)
H1$conf.int
```

```
## [1] -0.1710156 7.5710156
## attr(,"conf.level")
## [1] 0.95
```

```
H1$p.value
```

```
## [1] 0.06063451
```

At a confidence level of 95%, it can be concluded that the hypothesis is not rejected and the two supplements do not deliver a significant difference in tooth growth across the entire set of data.

Hypothesis: The two supplements deliver the same effects on tooth length at the 0.5 mg/day dose

```
H2 <- t.test(len~supp, data = subset(toothData, dose == 0.5))
H2$conf.int
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

```
H2$p.value
```

```
## [1] 0.006358607
```

At a confidence level of 95% it can be concluded that the hypothesis is rejected, and since the confidence interval lies on a positive set the Orange Juice supplement had a significantly larger effect on tooth length at the 0.5 mg dose.

Hypothesis: The two supplements deliver the same effects on tooth length at the 1.0 mg/day dose

```
H3 <- t.test(len~supp, data = subset(toothData, dose == 1.0))
H3$conf.int
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

```
H3$p.value
```

```
## [1] 0.001038376
```

At a confidence level of 95% it can be concluded that the hypothesis is rejected, and since the confidence interval lies on a positive set the Orange Juice supplement had a significantly larger effect on tooth length at the 1.0 mg dose.

Hypothesis: The two supplements deliver the same effects on tooth length at the 2.0 mg/day dose

```
H4 <- t.test(len~supp, data = subset(toothData, dose == 2.0))
H4$conf.int
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

```
H4$p.value
```

```
## [1] 0.9638516
```

At a confidence level of 95%, it can be concluded that the hypothesis is not rejected and the two supplements do not deliver a significant difference in tooth growth at the 2.0 mg dose.

Conclusions

Based on the assumptions that Tooth Length in Guinea Pigs is normally distributed, it can be concluded with 95% confidence that Orange Juice has a significantly larger effect on tooth growth than Ascorbic Acid at the 0.5 mg/day and 1.0 mg/day doses. There is no significant evidence that the two supplements deliver different effects at the 2.0 mg/day dose.

Looking Forward

Based on the initial exploration of data, it appears the effects of Orange Juice on the tooth length diminishes at the 2.0 mg/day dose while the Ascorbic Acid seems to continue a linear upward trend. If ethically reasonable, it would be interesting to explore the effects of each on a dose of 3.0 mg/day or more to see if there is a continuation in both trends.