Comparing Tooth Growth By Supplement & Dosage

Author: DRC

Overview

This study compares tooth growth by supplement and dosage. The data comes from the R dataset ToothGrowth.

Data Exploration

The dataset contains 60 observations, each with a respective tooth growth length ("len"), supplement ("supp") and dosage ("dose").

```
library(ggplot2)
data(ToothGrowth)

print(str(ToothGrowth))
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
## NULL
```

```
library(dplyr)
print(ToothGrowth %>% distinct(supp))
```

```
##   supp
## 1   VC
## 2   OJ
```

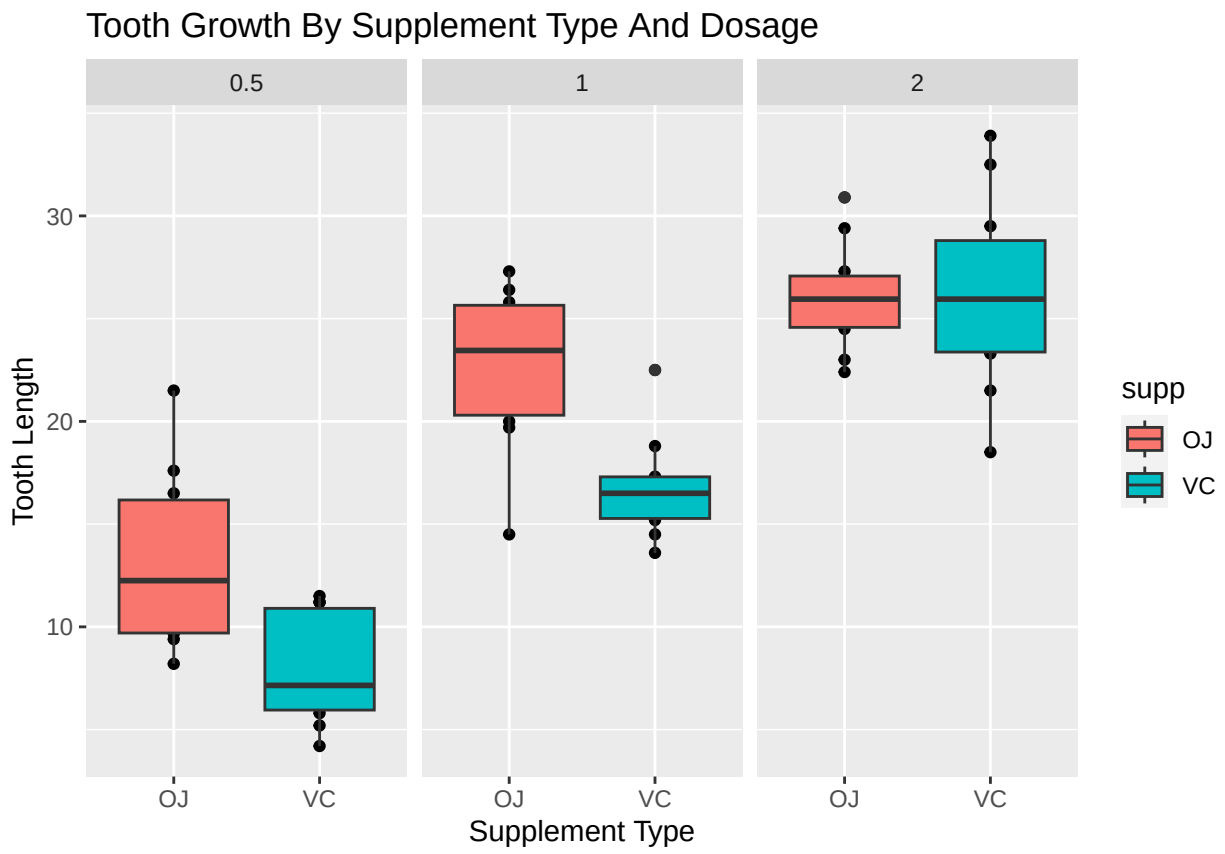
```
print(ToothGrowth %>% distinct(dose))
```

```
##   dose
## 1  0.5
## 2  1.0
## 3  2.0
```

Box plots were created to visualize the distribution of the data by supplement type and by dosage.

```
qplot(supp,
      len,
      data = ToothGrowth,
      facets = ~dose,
      main = "Tooth Growth By Supplement Type And Dosage",
      xlab = "Supplement Type",
      ylab = "Tooth Length") +
  geom_boxplot(aes(fill = supp))
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
```



Hypothesis Testing

1. Did dosage subgroups have longer teeth as dosage increased? Yes. To determine an answer, the ToothGrowth dataset was grouped by dosage amount. Then hypothesis testing was conducted comparing the "half dosage" data subset to the "full dosage" data subset, as well the "full dosage" subgroup was compared to the "double dosage" subgroup.

```
half_dose <- ToothGrowth$len[ToothGrowth$dose == 0.5]
one_dose <- ToothGrowth$len[ToothGrowth$dose == 1]
two_dose <- ToothGrowth$len[ToothGrowth$dose == 2]

print(t.test(one_dose,
             half_dose,
             alternative = "greater",
             paired = FALSE,
             var.equal = FALSE,
             conf.level = 0.95)$p.value)
```

```
## [1] 6.341504e-08
```

```
print(t.test(two_dose,
             one_dose,
             alternative = "greater",
             paired = FALSE,
             var.equal = FALSE,
             conf.level = 0.95)$p.value)
```

```
## [1] 9.532148e-06
```

Assuming that the variances across dosage subgroups are not equal, the null hypotheses were there was no difference in the tooth lengths between subgroups as the dosage increased. The alternative hypothesis was there was an increase in length as the dosage increased. Due to the low p-values of both t-tests, we reject the null hypotheses in favor of the alternative. P-values well outside of our confidence interval suggest it is extremely unlikely that the respective populations these two samples represent have normal distributions centered around the same mean. In other words, we may conclude that as dosage increases, so does tooth length.

2. Did supplement type subgroups have different lengths of teeth? Yes. To determine an answer, the ToothGrowth dataset was grouped by supplement type. Then hypothesis testing was conducted comparing the "OJ" and "VC" subgroups.

```
oj_sub <- ToothGrowth$len[ToothGrowth$supp == 'OJ']
vc_sub <- ToothGrowth$len[ToothGrowth$supp == 'VC']

t.test(oj_sub,
       vc_sub,
       alternative = "greater",
       paired = FALSE,
       var.equal = FALSE,
       conf.level = 0.95)$p.value
```

```
## [1] 0.03031725
```

Assuming that the variances across supplement type subgroups are not equal, the null hypotheses was there was no difference in the tooth lengths between subgroups. The alternative hypothesis was there was a "non-zero difference" in length as the dosage increased. Due to the low p-values of the t-test, we reject the null hypotheses in favor of the alternative. The p-values is outside of our confidence interval. A near 3% p-value suggests it is unlikely that the respective populations these two samples represent have normal distributions centered around the same mean. In other words, we may conclude that tooth length changes across supplement subgroups. Specifically, the "OJ" supplement subset seems to have higher tooth lengths than the "VC" group.

3. Did supplement type subgroups have different tooth lengths with double dosage? No. To determine an answer, the ToothGrowth dataset was grouped by supplement type and dosage. Then hypothesis testing was conducted comparing the "OJ with double dosage" and "VC with double dosage" subgroups.

```
oj_two <- ToothGrowth$len[ToothGrowth$supp == 'OJ' & ToothGrowth$dose == 2]
vc_two <- ToothGrowth$len[ToothGrowth$supp == 'VC' & ToothGrowth$dose == 2]
```

```r
t.test(oj_two,
       vc_two,
       alternative = "greater",
       paired = FALSE,
       var.equal = FALSE,
       conf.level = 0.95)$p.value
```

```
## [1] 0.5180742
```

Assuming that the variances across supplement type subgroups are not equal, the null hypotheses was there was no difference in the tooth lengths between subgroups. The alternative hypothesis was there was a "non-zero difference" in length across the subgroups. Due to the relatively high p-values of the t-test, we fail toreject the null hypotheses in favor of the alternative. The p-values is well within of our confidence interval. A near 52% p-value suggests it is very that the respective populations these two samples represent have normal distributions centered around the same mean. In other words, we may conclude that tooth length does not change across these two supplement subgroups. Specifically, the "OJ" supplement subset does not have higher tooth lengths than the "VC" group when we only observe subject with double dosage.