Assessing the To othGrowth Dataset with R

Amol Suraj Mishra

Abstract

In this rep ort we analyze the To othGrowth dataset. We statistically compare the effectiveness of tooth growth in Guinea pigs for each combination of dose level of Vitamin C and delivery method. According to our findings, 0.5 and 1.0 mg doses of Orange Juice promote distinct average to oth growth when compared to Vitamin C. With a 2.0 mg dose, there is a high probability that there is no difference b etween delivery methods.

ToothGrowth Dataset

```
data("ToothGrowth")
str(ToothGrowth)
```

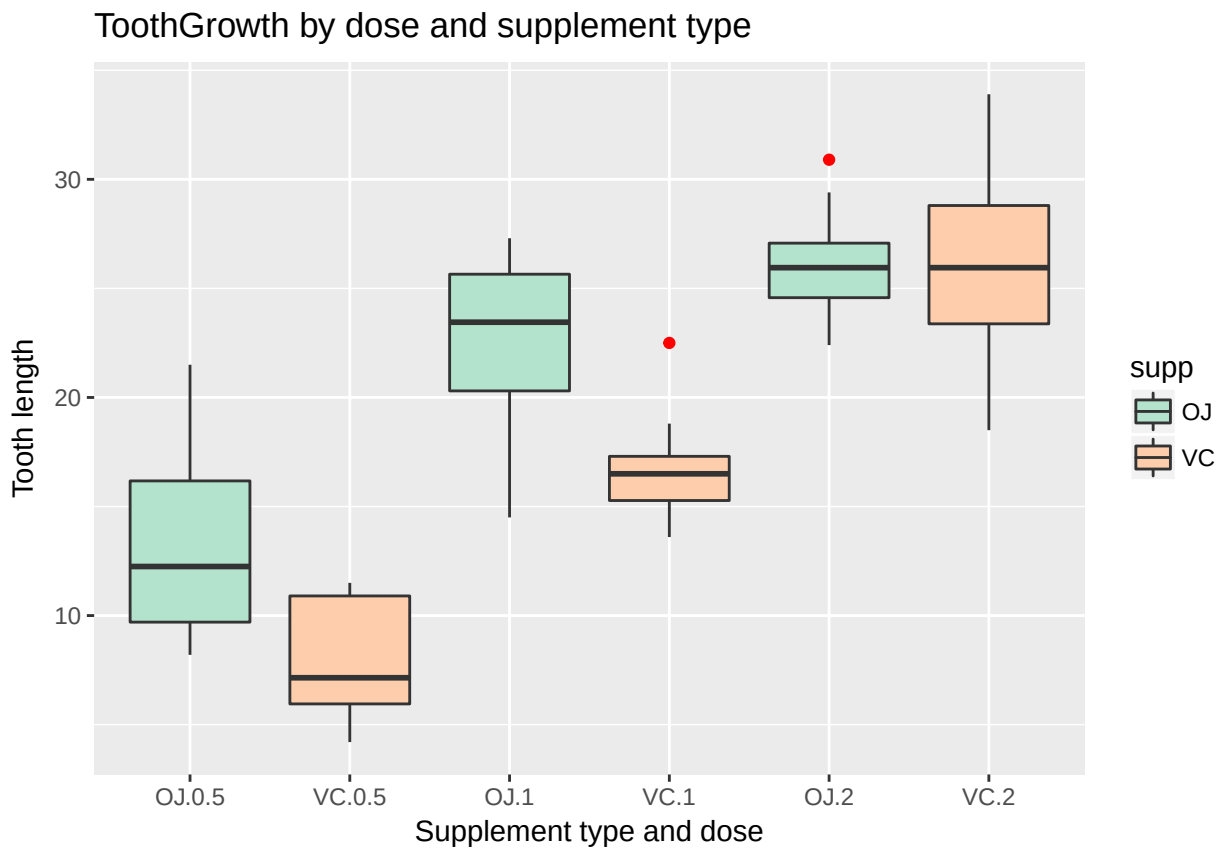
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Let us now see an orange juice vs ascorbic acid plot!

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
g <- ggplot(data = ToothGrowth, aes(x = interaction(supp, dose), y = len, fill = supp))
g <- g + geom_boxplot(outlier.colour = "red") + scale_fill_brewer(palette = "Pastel2")
g <- g + labs(title="ToothGrowth by dose and supplement type",
x="Supplement type and dose", y="Tooth length");
g
```



Data Processing

In this dataset, rows 1 to 10 corresp ond to the observations from Guinea pigs 1 to 10. The eleventh row contains another observation of Guinea pig numb er 1, and so on. In order to tidy this dataset, we are going to fit it into 6 columns and 10 rows. Each row corresp onds to one sp ecific Guinea pig, and each column represents a supplement (VC or OJ) and its asso ciated dose (0.5, 1.0 or 2.0).

```
library(reshape2)
```

```
## Warning: package 'reshape2' was built under R version 3.3.2
```

```
ToothGrowth$guinea.pig <- rep( c(1: 10), 6)
print(tidy <- dcast(ToothGrowth, guinea.pig ~ supp + dose, value.var = "len"))
```

```
##    guinea.pig OJ_0.5 OJ_1 OJ_2 VC_0.5 VC_1 VC_2
## 1           1   15.2 19.7 25.5    4.2 16.5 23.6
## 2           2   21.5 23.3 26.4   11.5 16.5 18.5
## 3           3   17.6 23.6 22.4    7.3 15.2 33.9
## 4           4    9.7 26.4 24.5    5.8 17.3 25.5
## 5           5   14.5 20.0 24.8    6.4 22.5 26.4
## 6           6   10.0 25.2 30.9   10.0 17.3 32.5
## 7           7    8.2 25.8 26.4   11.2 13.6 26.7
## 8           8    9.4 21.2 27.3   11.2 14.5 21.5
```

```
## 9           9   16.5 14.5 29.4    5.2 18.8 23.3
## 10         10    9.7 27.3 23.0    7.0 15.5 29.5
```

Data Analysis

In this section, we are going to p erform several tests to compare tooth growth of Guinea pigs by supplement (OJ, VC) and dose (0.5, 1.0, and 2.0 milligrams). For every comparation, we will consider the following hypotheses: * H_0: For a given dose level of Vitamin C, there is no difference in average tooth growth when we compare the delivery method (Orange Juice or Ascorbic Acid). * H_a: For a given dose level of Vitamin C, there is an actual difference in average tooth growth when we compare the delivery method (Orange Juice or Ascorbic Acid).

Performing a t test for each thing. 0.5, 1.0 and 2.0 milligrams we get.

1. 0.5 milligrams

```
tres_0.5 <- t.test(tidy$OJ_0.5, tidy$VC_0.5, paired = FALSE, var.equal = FALSE)
tres_0.5$p.value
```

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

As p value is very small. We reject the null hypothesis! Thus OJ and VC dont promote the same average tooth growth when the dosage is 0.5 mg.

2. 1.0 milligrams

```
tres_1 <- t.test(tidy$OJ_1, tidy$VC_1, paired = FALSE, var.equal = FALSE)
tres_1$p.value
```

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

As p value is very small again. We reject the null hypothesis! Thus OJ and VC dont promote the same average tooth growth when the dosage is 1.0 mg.

3. 2.0 milligrams

```
tres_2 <- t.test(tidy$OJ_2, tidy$VC_2, paired = FALSE, var.equal = FALSE)
tres_2$p.value
```

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

As p value is big enough here. We reject the alternate hypothesis! Thus OJ and VC equally promote the same average tooth growth when the dosage is 1.0 mg.