ToothGrowth analysis

Robert Jonczy

26.07.2015

Introduction

Exploratory analysis

Load `ToothGroowth` dataset:

```
library(datasets)
data(ToothGrowth)
```

Perform basis data exploratory

```
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 ...
```

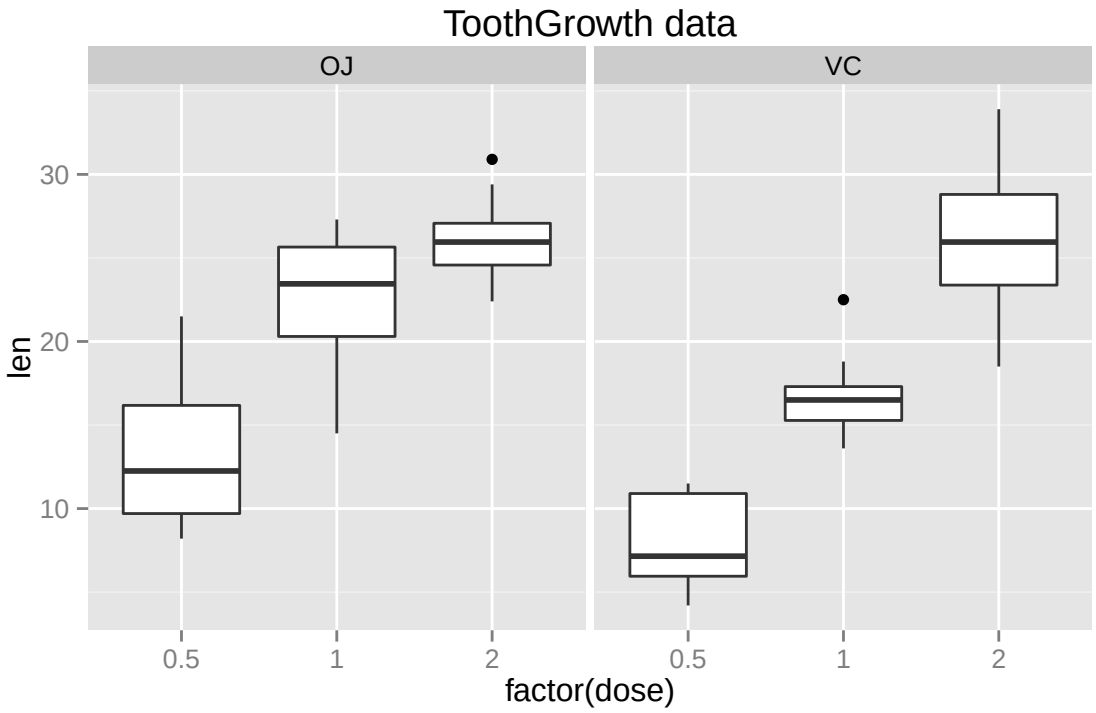
```
head(ToothGrowth)
```

```
##    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
```

```
tail(ToothGrowth)
```

```
##     len supp dose
## 55 24.8   OJ    2
## 56 30.9   OJ    2
## 57 26.4   OJ    2
## 58 27.3   OJ    2
## 59 29.4   OJ    2
## 60 23.0   OJ    2
```

As we can see this dataset contains length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice (OJ) or ascorbic acid (VC).



From the above plot we can see that with increase in dose level (0.5 mg to 2 mg) results in increased tooth growth. For low dose levels (0.5 mg and 1 mg) orange juice (OJ) suggests to have larger effect on tooth growth than ascorbic acid (VC). Below i will perform some statistical analysis of a data.

Data summary

```
library(plyr)
data_summary <- ddply(ToothGrowth , .(dose, supp) , summarize, mean = mean(len), sd = sd(len))
as.factor(data_summary$dose)
```

```
## [1] 0.5 0.5 1   1   2   2
## Levels: 0.5 1 2
```

```
data_summary
```

```
##   dose supp  mean       sd
## 1  0.5   OJ 13.23 4.459709
## 2  0.5   VC  7.98 2.746634
## 3  1.0   OJ 22.70 3.910953
## 4  1.0   VC 16.77 2.515309
## 5  2.0   OJ 26.06 2.655058
## 6  2.0   VC 26.14 4.797731
```

Confidence intervals

affect of 0.5 mg dose

```
tw05 <- subset(ToothGrowth, dose == 0.5)
t.test(len ~ supp, data = tw05, paired = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

Above T-test shows that the confidence interval is above 0, inferring that we reject the null hypothesis and accept that tooth growth is affected significantly by supp when dose is at 0.5. For dose 0.5, the p-value of OJ in comparison to VC is 0.0064. Since it is less than 0.05 (strong presumption against null hypothesis), it means that there is a difference between both methods.

affect of 1.0 mg dose

```
tw10 <- subset(ToothGrowth, dose == 1.0)
t.test(len ~ supp, data = tw10, paired = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

Above T-test shows that the confidence interval is above 0, inferring that we reject the null hypothesis and accept that tooth growth is affected significantly by supp when dose is at 1.0. For dose 1.0, the p-value of OJ in comparison to VC is 0.001. Since it is less than 0.05 (strong presumption against null hypothesis), it means that there is a difference between both methods.

affect of 2.0 mg dose

```r
tw20 <- subset(ToothGrowth, dose == 2.0)
t.test(len ~ supp, data = tw20, paired = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

Above T-test shows that the confidence interval holds the value of 0, inferring that we reject the alternative hypothesis and accept that tooth growth is not affected significantly by supp when dose is at 2.0. For dose 2.0, the p-value of OJ in comparison to VC is 0.064. Since it is greater than 0.05 (low presumption against null hypothesis), it means that there is a no that much of a difference between both methods.

Conclusion

We can accept the hypothesis that the supp OJ gives increased tooth growth as compared to VC. In addition, the boxplot also displays an approximately linear trendpoint where increasing dosage begets increased tooth length as well, hence we can conclude that increased dosage promotes tooth growth.

Assumptions

1. Group's (OJ, VC) variances are not equal,
2. Group's (OJ, VC) are independent of each other, thats why we promote unpaired testing