Statistics Inference Course Puroject_ToothGrowth Data Analysis (Part2)

Author: Yiwei Yan Date: January 23, 2015

1. Overview and Data Summary

In this study, we are going to analyze the ToothGrowth data in the R datasets package. We utilize data from the dataset named "ToothGrowth", which is about the 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 or ascorbic acid). Specifically, "mtcars" is a data frame with 60 observations on 3 variables (len, supp factor, and dose numeric). Before start our analysis, Let's open this dataset and review it at first:

```
data(ToothGrowth)
summary(ToothGrowth)
```

```
##      len            supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

2. Assumption

- supplements have a treatment effect and no other fconfounding factors;
- Samples are unpaired that with unequal variances;
- Guinea pigs are essentially identical.

3. Data Exploration

First of all, we plot the data to observe a number of correlations. We can see that as the dosage increases, the length of the tooth increases. The OJ delivery method yields a greater length than the VC (approximately 10mm) for smaller dosages but the difference is negligable by a 2mg dosage.

```
library(ggplot2)
dose <- ToothGrowth$dose
supp <- ToothGrowth$supp
len <- ToothGrowth$len
par(mfrow = c(1,2))
ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose)))+
  geom_boxplot() +
  guides(fill=FALSE) +
  facet_grid(. ~ supp)+
  theme_classic()
```
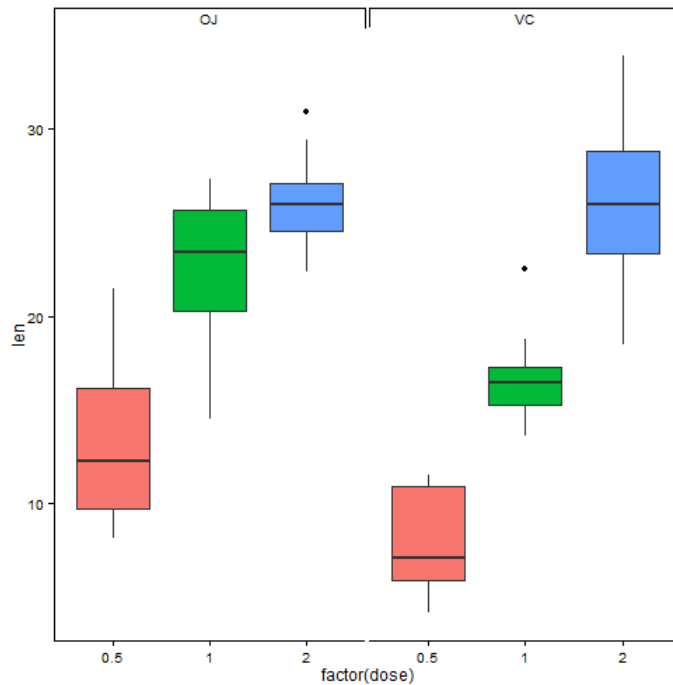
```
ggplot(ToothGrowth, aes(x = factor(supp), y = len, fill = factor(supp)))+
  geom_boxplot() +
  guides(fill=FALSE) +
  facet_grid(. ~ dose)+
  theme_classic()
```
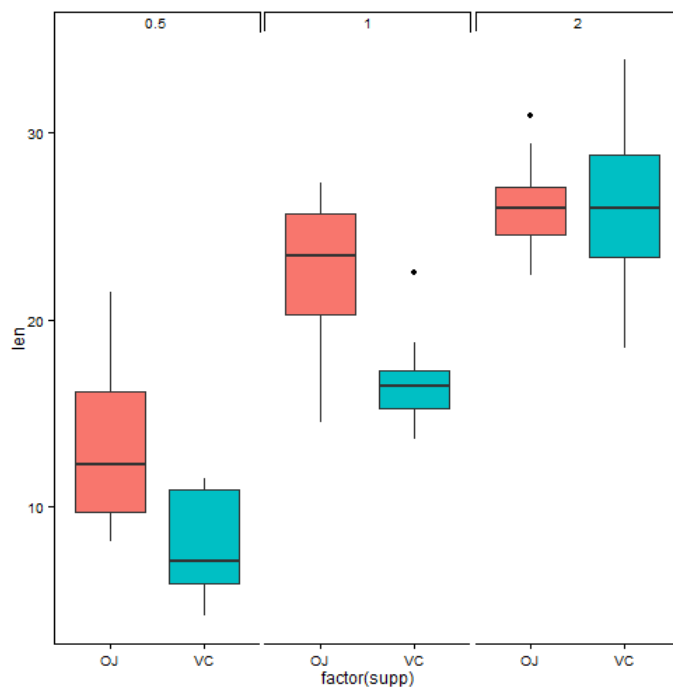

4. Statistical Inference

Then, we calculate the mean and sd of each dosage and supplement, split data by dosages, and conduct a t test between supplements in this section. The tooth growth is compared by supplement for each dosage under the null hypothesis that each supplement has the same effect at a certain dosage on the tooth. Firstly, dosages 1.0 and 1.5 have significant p-values of 0.00636 and 0.00104 respectively indicating that the difference in mean values between the supplements is significant. Dosage 1.0 has a confidence interval of 1.719-8.781 and dosage 2.0 has a confidence interval of 2.802 to 9.058. Secondly, dosage 3.0 has a very high p-value of 0.964 and a confidence interval below zero -3.798 to 3.638. This indicates that there is no significance between the supplements at this dosage. This is also intuitive from the boxplot.

```
Dmeans <- aggregate(ToothGrowth$len, by = list(ToothGrowth$dose), FUN = mean)
Dsds <- aggregate(ToothGrowth$len, by = list(ToothGrowth$dose), FUN = sd)
Smeans <- aggregate(ToothGrowth$len, by = list(ToothGrowth$supp), FUN = mean)
Ssds <- aggregate(ToothGrowth$len, by = list(ToothGrowth$supp), FUN = sd)

d0.5 <- subset(ToothGrowth, dose == 0.5)
d1.0 <- subset(ToothGrowth, dose == 1.0)
d2.0 <- subset(ToothGrowth, dose == 2.0)

test0.5 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d0.5)
test0.5$p.value; test0.5$conf
```

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

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

```
test1.0 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d1.0)
test1.0$p.value; test1.0$conf
```

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

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

```
test2.0 <- t.test(len ~ supp, paired = FALSE, var.equal = FALSE, data = d2.0)
test2.0$p.value; test2.0$conf
```

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

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

5. Comclusion

To sum up, as shown in our analysis results, we can conclude some important results: 1) The supplements orange juice and ascorbic acid have different effects on tooth length for lower dosages of VC based on the t-test. 2) Orange juice outputs a longer tooth for 0.5 and 1.0mg dosages; 3) There is no change in tooth length at a dosage of 3.0mg.