Study of ToothGrowth data

Ariane

Saturday, September 20, 2014

The ToothGrowth dataset gives 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). There are 60 observations in the data frame, and the 3 variables are:

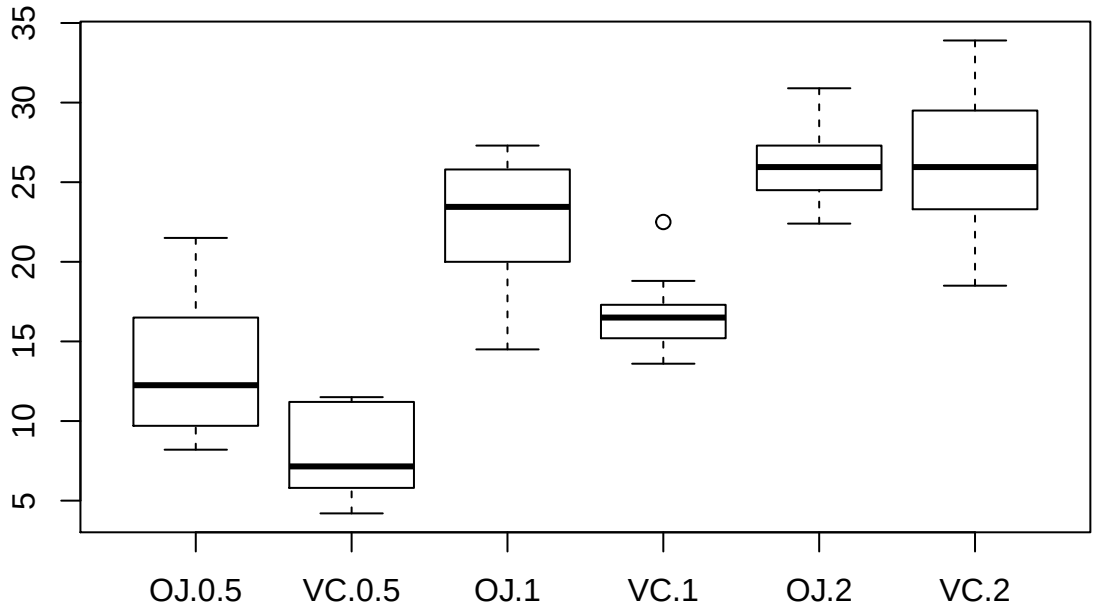
- len: the tooth length
- supp: the supplement type (VC or OJ)
- dose: the dose in milligrams (0.5, 1 or 2)

We want to study whether the supplement type or the dose given to the guinea pigs have an influence on the growth of the teeth.

We read the data.

```
library(datasets)
data(ToothGrowth)
ToothGrowth$supp <- as.factor(ToothGrowth$supp)
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
```

We do a box-and-whisker plot of the data



The length of the teeth seems to be higher when the dose of vitamin increases, and higher with the OJ supplement than with the VC supplement.

First, we consider the influence of the supplement type. The null hypothesis H_0 is "the mean of the difference of tooth growth between OJ and VC is null". We perform a Student's t-Test with confidence levels of 99.7 and 99.8 %.

```
data1 <- ToothGrowth$len[1:30] # VC
data2 <- ToothGrowth$len[31:60] # OJ
test1 <- t.test(data2, data1, paired = TRUE, conf.level=0.997)
test2 <- t.test(data2, data1, paired = TRUE, conf.level=0.998)
```

The confidence interval for the first test is [0.0708 ; 7.3292] and the confidence interval for the second test is [-0.1049 ; 7.5049]. 0 is not included in the first interval, therefore we can reject the null hypothesis with a 99.7 % confidence level. However, 0 is included in the second interval. The attained significance level for this test is comprised between 0.2 and 0.3 %.

We now consider the influence of the dose level. The null hypothesis H_0 are "the mean of the difference of tooth growth between 0.5 and 1 mg is null" and "the mean of the difference of tooth growth between 1 and 2 mg is null". For the first null hypothesis, we perform a Student's t-Test with confidence levels of 99.99985 and 99.9999 %.

```
data1 <- ToothGrowth$len[c(1:10,31:40)] # 0.5
data2 <- ToothGrowth$len[c(11:20,41:50)] # 1
test1 <- t.test(data2, data1, paired = TRUE, conf.level=0.9999985)
test2 <- t.test(data2, data1, paired = TRUE, conf.level=0.999999)
```

The confidence interval for the first test is [0.1323 ; 18.1277] and the confidence interval for the second test is [-0.134 ; 18.394]. 0 is not included in the first interval, therefore we can reject the null hypothesis with a 99.99985 % confidence level. However, 0 is included in the second interval. The attained significance level for this test is comprised between 0.0001 and 0.00015 %.

For the second null hypothesis, we perform a Student's t-Test with confidence levels of 99.98 and 99.985 %.

```
data1 <- ToothGrowth$len[c(11:20,41:50)] # 1
data2 <- ToothGrowth$len[c(21:30,51:60)] # 2
test1 <- t.test(data2, data1, paired = TRUE, conf.level=0.9998)
test2 <- t.test(data2, data1, paired = TRUE, conf.level=0.99985)
```

The confidence interval for the first test is [0.0204 ; 12.7096] and the confidence interval for the second test is [-0.1555 ; 12.8855]. 0 is not included in the first interval, therefore we can reject the null hypothesis with a 99.99985 % confidence level. However, 0 is included in the second interval. The attained significance level for this test is comprised between 0.015 and 0.02 %.

From this analysis, we can conclude that the length of the teeth is higher with the OJ supplement than with the VC supplement. The length of the teeth is also higher when the dose of supplement increases. However, the confidence level is smaller when we test the difference between 1 and 2 mg than when we test the difference between 0.5 and 1 mg. It would be interesting to test a higher dose and check if we can still reject the null hypothesis with an acceptable error rate.