

Study of ToothGrowth data with R

Roman Preobrazhensky

The data represents 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).

Our goal here is to make some conclusions about effectiveness of different methods represented in a data set.

Exploratory analysis

First, lets load the dataset and look at some summary.

```
library(datasets)
data("ToothGrowth")
str(ToothGrowth)
```

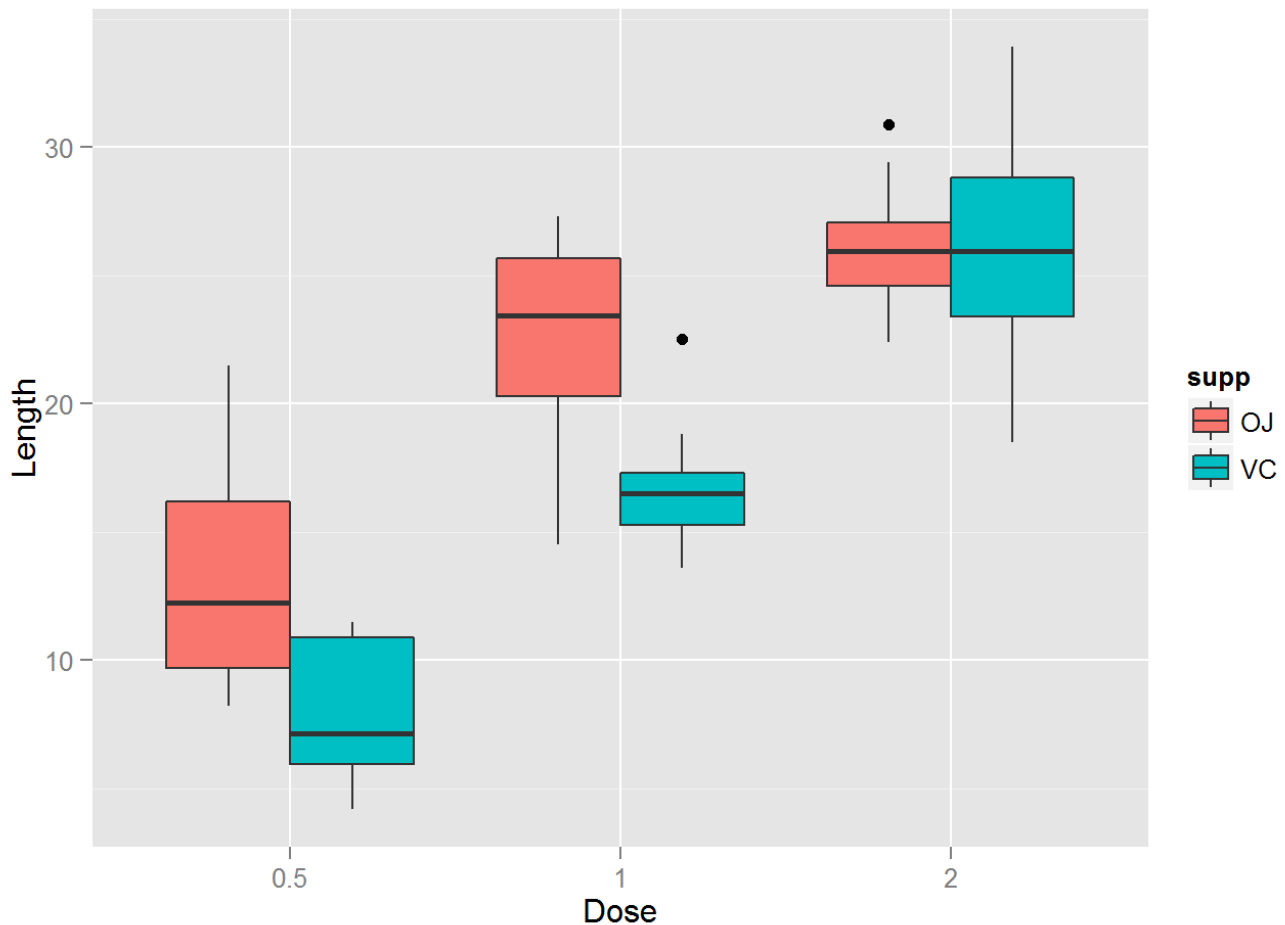
```
## 'data.frame':   60 obs. of  3 variables:
## $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
## $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

For the purposes of exploratory analysis we will create a boxplot for each delivery method and dose.

```
library(ggplot2)
ggplot() + geom_boxplot(data=ToothGrowth, aes(factor(dose), len, fill=supp)) + xlab("Dose") + ylab("Length")
```



From this boxplot it is clear, that length of odontoblasts increasing with an increase of dose for each of delivery methods. Although it is more obvious in case of VC method (neither of whiskers seems to reach means of another groups), let's assume that such growth of effectiveness is obvious for OJ method either. Also it appears that these groups we've plotted are closer by dose rather than by method.

Statistical analysis

In order to perform some statistical testing, we will assume that each distribution of `len` for each `supp` and each `dose` is approximately normal. Also we will assume that these distributions have unequal variances and are not paired. First, let's check if OJ-delivered vitamins made greater effects on length regardless of the dose:

```
VC <- ToothGrowth[which(ToothGrowth$supp=='VC'), c('dose', 'len')]
OJ <- ToothGrowth[which(ToothGrowth$supp=='OJ'), c('dose', 'len')]
t.test(OJ$len, VC$len, paired = FALSE, var.equal = FALSE)$conf.int
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

95% confidence interval contains zero, and hence we cannot reject the null-hypothesis. It appears that in average these methods of vitamin delivery do not show statistically relevant difference.

Let's now compare methods dose-wise. First we'll take a look at 0.5mg dosage:

```
VC0.5 <- VC$len[which(VC$dose==0.5)]
OJ0.5 <- OJ$len[which(OJ$dose==0.5)]
t.test(OJ0.5, VC0.5, paired = FALSE, var.equal = FALSE)$conf.int
```

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

This time we have 95% interval lying above zero, and we can say that we're 95% confident that for 0.5mg dose delivery with orange juice is more efficient than with an ascorbic acid.

Now consider 1mg dosage:

```
VC1 <- VC$len[which(VC$dose==1)]
OJ1 <- OJ$len[which(OJ$dose==1)]
t.test(OJ1, VC1, paired = FALSE, var.equal = FALSE)$conf.int
```

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

And again, we have orange juice method of delivery more efficient than ascorbic acid.

Finally, let's take a look at 2mg dose:

```
VC2 <- VC$len[which(VC$dose==2)]
OJ2 <- OJ$len[which(OJ$dose==2)]
t.test(OJ2, VC2, paired = FALSE, var.equal = FALSE)$conf.int
```

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

This time confidence interval is almost symmetric about zero, which means that we should assume null-hypothesis and say that there are no statistically proven differences between efficiency of delivery methods for 2mg dose.

Conclusions

We can conclude now, based on the assumptions of normally distributed groups of data that we were comparing and their unequal variance and unpairness, that for small and medium dosages (0.5mg and 1mg respectively) vitamins delivery with orange juice is more efficient than with an ascorbic acid. However, for large dose (2mg) there are no statistical differences between these two methods.