Guinea pig's tooth growth analysis

Kirill Panarin

Overview

The response is 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).

ToothGrowth data set contains 3 variables:

- **len**: tooth length, numeric
- **supp**: supplement type (VC or OJ), factor
- **dose**: dose in milligrams, numeric

Exploratory analysis

Let's load ToothGrowth data set and display structure of it:

```
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 ...
```

```
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  
```

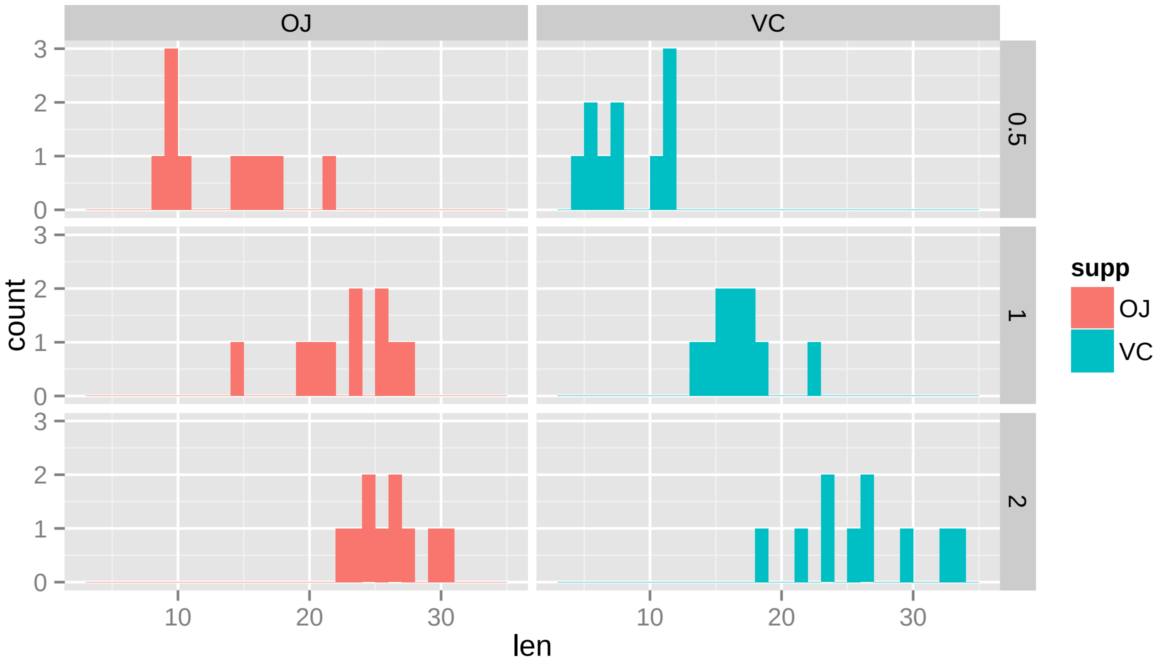
```
head(ToothGrowth)
```

```
##    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
```

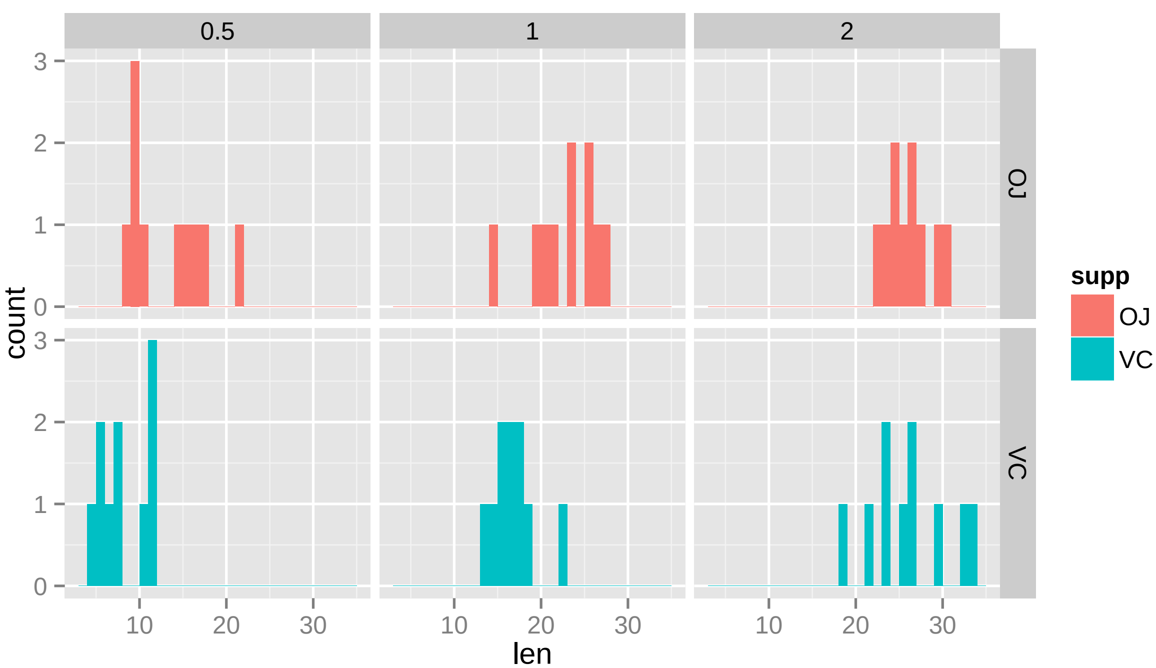
So we have a data set with 60 observations of 3 variables. It will be better to convert `dose` variable from numeric to factor:

```
ToothGrowth$dose <-as.factor(ToothGrowth$dose)
```

Let's plot how `len` is distributed for different `supp` and `dose` in order to highlight some basic features of the data:



Base on the figure above we can assume that as dose increases the chances to get higher tooth length become higher.



Based on the figure above, I can't assume that supplement type somehow affects tooth growth rate.

Tooth growth by supp and dose comparison

Comparison by supp

First, let's apply t-statistics for length by supplement type for each dose:

```
stat_by_supp <- sapply(unique(ToothGrowth$dose), function(d) {
    stat <- t.test(len ~ supp, paired = F, var.equal = F, data = ToothGrowth[ToothGrowth$dose == d,])
    stat$dose <- as.character(d)
    stat$conf.int_min <- stat$conf.int[1]
    stat$conf.int_max <- stat$conf.int[2]
    stat
})
t(stat_by_supp)[, c(10, 1:3, 11, 12)]
```

```
##      dose  statistic  parameter p.value     conf.int_min conf.int_max
## [1,] "0.5" 3.169733   14.96875  0.006358607 1.719057     8.780943
## [2,] "1"   4.03277    15.35767  0.001038376 2.802148     9.057852
## [3,] "2"   -0.0461361 14.03982  0.9638516   -3.79807     3.63807
```

For doses 0.5 and 1 we have high t-statistics and low p-value (< 5%), so we have to reject null hypothesis and assume that supplement type OJ is more effective than VC i.e. **orange juice is more effective than ascorbic acid for small doses**. In contrast, for dose 2 t-statistics is small, p-value is high (>>5%) and 0 is in the confidence interval, so we can assume that **for big doses supplement type doesn't play significant role in tooth growth rate**.

Comparison by dose of Vitamin C in orange juice

In this case we have more analysis permutation than in the case with comparison by supp. There are three combinations of doses. Let's run t-statistics for all three combinations in case when we are dealing with OJ supplement type:

```
dose_combn <- combn(as.character(unique(ToothGrowth$dose)), 2)
statf <- function(st, f, s) {
    dose_pair <- c(f, s)
    stat <- t.test(len ~ dose, paired = F, var.equal = F, data = ToothGrowth[ToothGrowth$dose %in% dose_
    stat$conf.int_min <- stat$conf.int[1]
    stat$conf.int_max <- stat$conf.int[2]
    stat$dose.first <- dose_pair[1]
    stat$dose.second <- dose_pair[2]
    stat
}
stat_by_dose_oj <- mapply(function(f, s) statf("OJ", f, s), dose_combn[1, ], dose_combn[2, ])
colnames(stat_by_dose_oj) <- c()
t(stat_by_dose_oj)[, c(12, 13, 1:3, 10, 11)]
```

```
##      dose.first dose.second statistic parameter p.value      conf.int_min
## [1,] "0.5"      "1"         -5.048635 17.69835  8.784919e-05 -13.41563
## [2,] "0.5"      "2"         -7.817021 14.6678   1.323784e-06 -16.33524
## [3,] "1"        "2"         -2.247761 15.84238  0.03919514   -6.531443
##      conf.int_max
## [1,] -5.524366
```

```
## [2,] -9.324759
## [3,] -0.1885575
```

As we can see for doses as dose increases from 0.5 to 1 and 2 t-statistic is big and p-value is small, and 0 doesn't fall into confidence interval, so we can reject null hypothesis and assume that **as dose increases from 0.5 to 1 tooth growth rate increases too**. In contrast, for the doses pair 1 and 2 t-statistics is not so big, and p-value is less than 5%, and the right boundary of the confidance interval is almost 0, so we can't reject our null hypothesis. **When dose increases from 1 to 2 there is no significant influence on tooth growth rate.**

Comparison by dose of Vitamin C in ascorbic acid

There are three combinations of doses. Let's run t-statistics for all three combinations in case when we are dealing with VC supplement type:

```
stat_by_dose_vc <- mapply(function(f, s) statf("VC", f, s), dose_combn[1, ], dose_combn[2, ])
colnames(stat_by_dose_vc) <- c()
t(stat_by_dose_vc)[, c(12, 13, 1:3, 10, 11)]
```

```
##      dose.first dose.second statistic parameter p.value      conf.int_min
## [1,] "0.5"      "1"         -7.46343  17.86244  6.811018e-07 -11.26571
## [2,] "0.5"      "2"         -10.3878  14.32712  4.681577e-08 -21.90151
## [3,] "1"        "2"         -5.469814 13.59996  9.155603e-05 -13.05427
##      conf.int_max
## [1,] -6.314288
## [2,] -14.41849
## [3,] -5.685733
```

As we can see for all cases t-statistics is big, p-value is small, and confidence interval doesn't contain zero, so we can reject null hypothesis and assume that **when dose increases from 0.5 to and from 1 to 2 the tooth growth rate increases too**.

Conclusions

- Orange juice is more effective than ascorbic acid for small doses.
- For big doses supplement type doesn't play significant role in tooth growth rate.
- In the case of orange juice, as dose increases from 0.5 to 1 tooth growth rate increases too, and when dose increases from 1 to 2 there is no significant influence on tooth growth rate.
- In the case of ascorbic acid, when dose increases from 0.5 to and from 1 to 2 the tooth growth rate increases too.

Base on the conclusions above we can generalize the ide, and state that orange juice is more effective supplement type than ascorbic acid. In addition, we can state that Vitamin C definitely increases tooth growth rate.