Statistical Inference Course Project 2: Playing with Tooth Growth Data

Josep Anton Mir Tutusaus

Monday, June 15, 2015

1. Overview: summary of the data

```
library(datasets)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
library(dplyr)
```

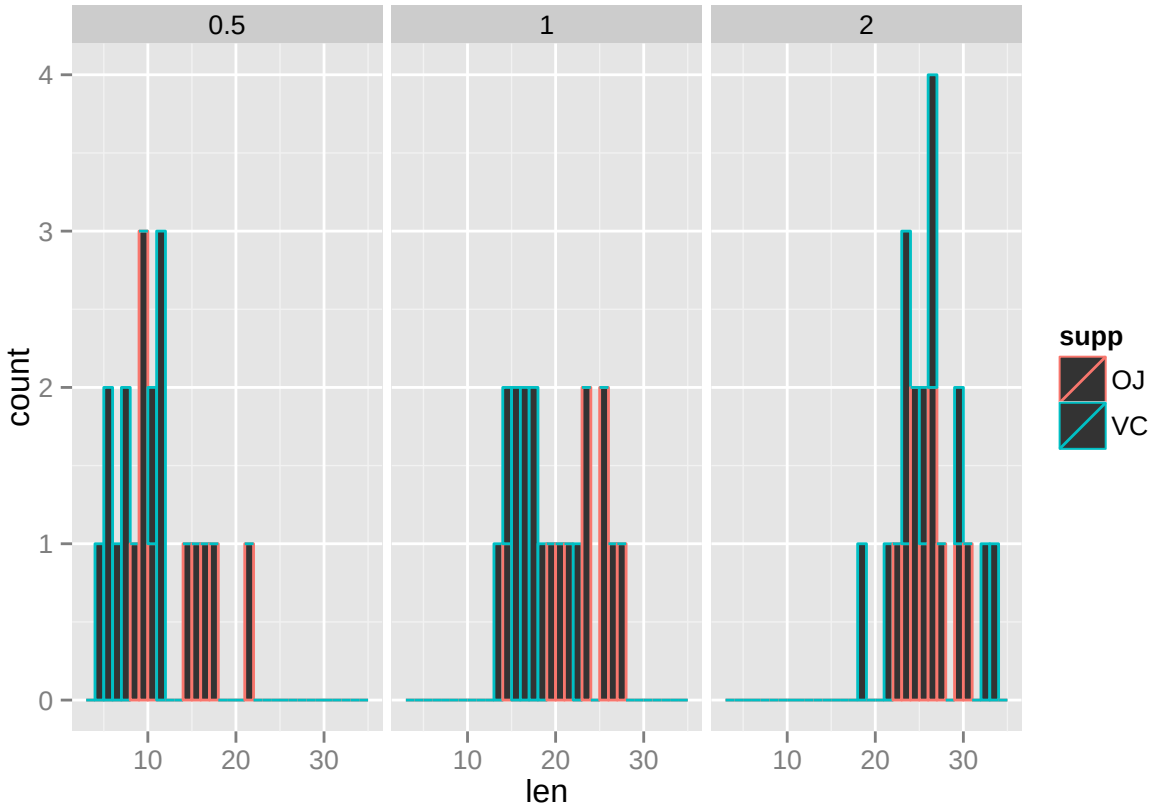
```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

Here there is a small description of the data set, as stated in the R Documentation:

> ### The Effect of Vitamin C on Tooth Growth in Guinea Pigs
>
> **Description**
>
> 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).
>
> **Usage**
>
> ToothGrowth
>
> **Format**
>
> A data frame with 60 observations on 3 variables.
>
> [,1] len numeric Tooth length [,2] supp factor Supplement type (VC or OJ). [,3] dose numeric Dose in milligrams.

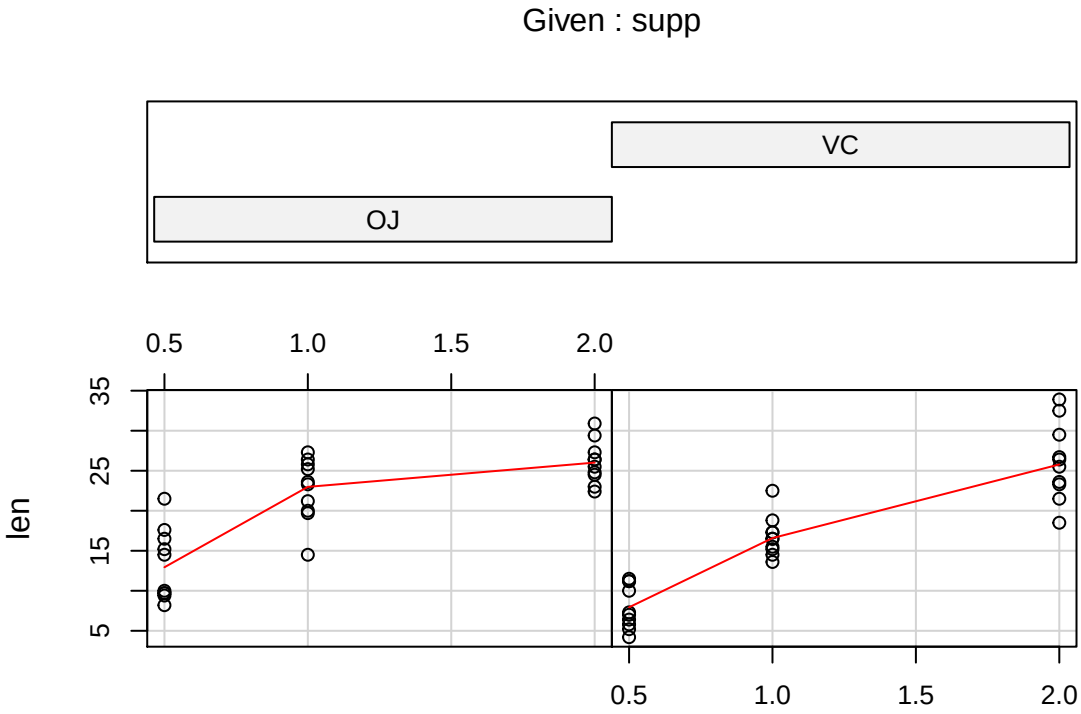
We can have a full view of the data in this plot:

```
qplot(len, data=ToothGrowth, binwidth=1, facets=.~dose, colour=supp)
```



And a better understanding of the realtion between length and dose, and between length and supplement type in this plot:

```
coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth, xlab = "ToothGrowth data: length vs
```



ToothGrowth data: length vs dose, given type of supplement

We can see in the figures that the length of teeth grows as the vitamin C dose rises. It may have better results for orange juice in the 0.5 and 1.0 mg dose, but it may be equal between OJ and VC within the 2.0 mg dose. Is that true?

2. Results and discussion

Here I'm going to present hypothesis tests to compare tooth growth by supp and dose.

2.1. Tooth growth vs. dose

```
TG <- split(ToothGrowth, ToothGrowth$supp)
TGOJ <- as.data.frame(TG[1]); TGVC <- as.data.frame(TG[2])
```

```
TGOJdose <- split(TGOJ, factor(TGOJ$OJ.dose))
aTGOJdose <- as.data.frame(TGOJdose[1]); bTGOJdose <- as.data.frame(TGOJdose[2]); cTGOJdose <- as.data.
```

2.1.1. Orange juice (OJ)

*Are 0.5 mg and 1.0 mg treatments statistically different?

```
t.test(aTGOJdose$X0.5.OJ.len, bTGOJdose$X1.OJ.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  aTGOJdose$X0.5.OJ.len and bTGOJdose$X1.OJ.len
## t = -5.0486, df = 18, p-value = 8.358e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -13.410814  -5.529186
## sample estimates:
## mean of x mean of y
##     13.23     22.70
```

Yes ,they are! 1.0 mg treatment produces larger teeth.

*Are 1.0 mg and 2.0 mg treatments statistically different?

```
t.test(bTGOJdose$X1.OJ.len, cTGOJdose$X2.OJ.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  bTGOJdose$X1.OJ.len and cTGOJdose$X2.OJ.len
## t = -2.2478, df = 18, p-value = 0.03736
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -6.5005017 -0.2194983
## sample estimates:
## mean of x mean of y
##     22.70     26.06
```

Yes ,they are! 2.0 mg treatment produces larger teeth.

Logically, 2.0 mg treatment produce larger theeth than 0.5 mg treatment.

```
TGVCdose <- split(TGVC, factor(TGVC$VC.dose))
aTGVCdose <- as.data.frame(TGVCdose[1]); bTGVCdose <- as.data.frame(TGVCdose[2]); cTGVCdose <- as.data.
```

2.1.2. Ascorbic acid (VC)

*Are 0.5 mg and 1.0 mg treatments statistically different?

```
t.test(aTGVCdose$X0.5.VC.len, bTGVCdose$X1.VC.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  aTGVCdose$X0.5.VC.len and bTGVCdose$X1.VC.len
## t = -7.4634, df = 18, p-value = 6.492e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.264346  -6.315654
## sample estimates:
## mean of x mean of y
##      7.98     16.77
```

Yes ,they are! 1.0 mg treatment produces larger teeth.

*Are 1.0 mg and 2.0 mg treatments statistically different?

```
t.test(bTGVCdose$X1.VC.len, cTGVCdose$X2.VC.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  bTGVCdose$X1.VC.len and cTGVCdose$X2.VC.len
## t = -5.4698, df = 18, p-value = 3.398e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -12.96896  -5.77104
## sample estimates:
## mean of x mean of y
##     16.77     26.14
```

Yes ,they are! 2.0 mg treatment produces larger teeth.

Logically, 2.0 mg treatment produce larger theeth than 0.5 mg treatment.

2.1. Tooth growth vs. supplement

For the sake of the study, let's assume that our goal is to produce larger teeth. The 2.0 mg treatment showed to be the best choice, so now it is important how to supply these 2.0 mg, by orange juice (OJ) or by ascorbic acid (VC).

*Are OJ and VC supplements statistically different, in reagards to 2.0 mg dose?

```
t.test(cTGOJdose$X2.OJ.len, cTGVCdose$X2.VC.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  cTGOJdose$X2.OJ.len and cTGVCdose$X2.VC.len
## t = -0.0461, df = 18, p-value = 0.9637
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.722999  3.562999
## sample estimates:
## mean of x mean of y
##     26.06     26.14
```

No, they are not! It is unimportant what supplement we choose to supply the vitamin C.

*Are OJ and VC supplements statistically different, in reagards to 1.0 mg dose?

```
t.test(bTGOJdose$X1.OJ.len, bTGVCdose$X1.VC.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
```

```
## data:  bTGOJdose$X1.OJ.len and bTGVCdose$X1.VC.len
## t = 4.0328, df = 18, p-value = 0.0007807
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.840692 9.019308
## sample estimates:
## mean of x mean of y
##     22.70     16.77
```

Yes, they are! Oranje juice produces larger teeth.

*Are OJ and VC supplements statistically different, in reagards to 0.5 mg dose?

```
t.test(aTGOJdose$X0.5.OJ.len, aTGVCdose$X0.5.VC.len, var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  aTGOJdose$X0.5.OJ.len and aTGVCdose$X0.5.VC.len
## t = 3.1697, df = 18, p-value = 0.005304
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.770262 8.729738
## sample estimates:
## mean of x mean of y
##     13.23      7.98
```

Yes, they are! Oranje juice produces larger teeth.

3. Conclusions and assumptions

I'm assuming that:

In regards to the experimental design: - There are 60 guinea pigs, one of each receives a specific treatment. - There are 6 groups of 10 guinea pigs: 0.5 mg + OJ, 1 mg + OJ, 2 mg + OJ, 0.5 mg + VC, 1 mg + VC, 2 mg + VC.

In regards to statistical assumptions: - All events are independent. - All events have the same variance. It is reasonable, as the groups contain the same number of guinea pigs.

In light of these assumptions, we conclude that 2.0 mg doses produce larger teeth than 0.5 and 1.0 mg doses. There's only signifficant difference between the two supplements (OJ and VC) in the 0.5 and the 1.0 mg treatments.