

Inferential analysis of ToothGrowth dataset

Piotr Turski

September 23, 2015

Overview

This document shows analysis of ToothGrowth dataset version 3.2.2 provided by R.

Exploratory data analyses

```
data(ToothGrowth)
str(ToothGrowth)
```

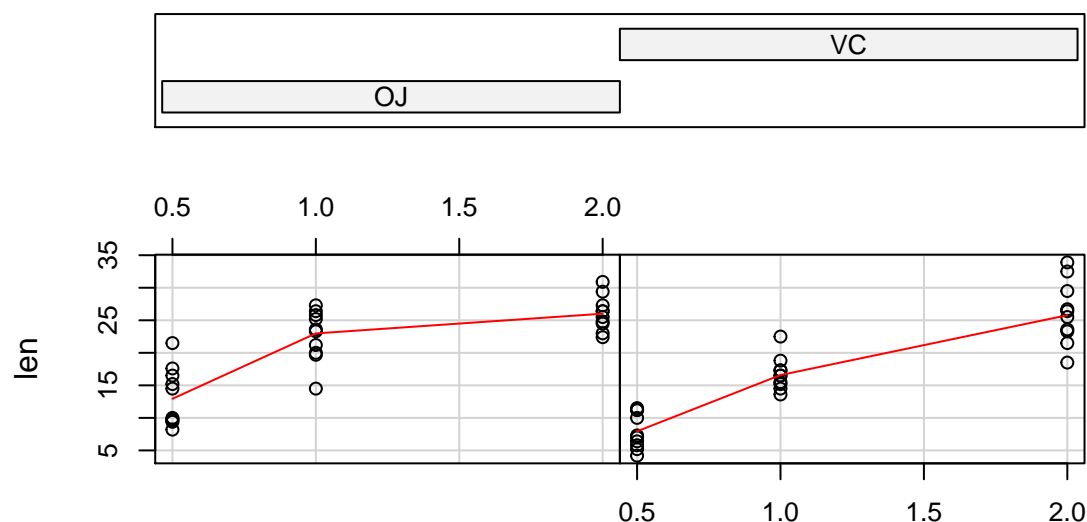
```
## 'data.frame': 60 obs. of 3 variables:
## $ len : num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
## $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
## $ dose: num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

According to `?ToothGrowth`, dataset content: “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)”. Newer version’s (3.3.0) description says that ascorbic acid is coded as VC.

A plot suggested in the documentation helps to understand the data:

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

Given : supp



ToothGrowth data: length vs dose, given type of supplement

Compare tooth growth by supp and dose

Let's do `t.test` to compare: supplement types, dosage, supplement type per each dose

Compare supplement type

```
t.test(len ~ supp, data = ToothGrowth)

##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##      20.66333      16.96333
```

Confidence interval contains 0 so we cannot reject null hypothesis that supplement type doesn't matter.

Compare dosage

```
library(dplyr)
t.test(len ~ dose, data = filter(ToothGrowth, dose != 2))

##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5 mean in group 1
##      10.605      19.735
```

Confidence interval doesn't contain 0 so we reject null hypothesis that dose 0.5mg vs 1mg doesn't matter. Mean length for dose 1mg is greater.

```
t.test(len ~ dose, data = filter(ToothGrowth, dose != 0.5))

##
##  Welch Two Sample t-test
##
## data:  len by dose
```

```
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##      19.735      26.100
```

Similarly, dose 2mg gives larger length mean than dose 1mg.

Compare supplement type per dosage

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 0.5))
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##      13.23      7.98
```

Confidence interval doesn't contain 0 so we reject null hypothesis that supplement type for dose 0.5mg doesn't matter. Mean length for dose 1mg is greater for orange juice.

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 1))
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##      22.70      16.77
```

Confidence interval doesn't contain 0 so we reject null hypothesis that supplement type for dose 1mg doesn't matter. Mean length for dose 1mg is greater for orange juice.

```
t.test(len ~ supp, data = filter(ToothGrowth, dose == 2))
```

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

```
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##           26.06           26.14
```

Confidence interval contains 0 so we cannot reject null hypothesis that supplement type for dose 2mg doesn't matter.

Conclusions

- Tooth length increases with dose.
- For doses 0,5mg and 1mg orange juice causes greater tooth length than ascorbic acid.
- For dose 2mg and irrespective of the dose, supplement type doesn't matter.

Assumptions

- Distribution of means is approximately normal
- Samples are random and independent
- Tested animals are representative for the whole population