ToothGrowth

Egor Ignatenkov

27.10.2014

Here we'll try to analyze data from ToothGrowth dataset.

First, let's load data and see how it looks.

```
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 ...
```

Let's factorize the dose column:

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

```
summary(ToothGrowth)
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

Let's see how growth depends on supp:

```
g1<-ToothGrowth$len[ToothGrowth$supp == "VC"]
g2<-ToothGrowth$len[ToothGrowth$supp == "OJ"]
difference<-g2-g1
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 3.3026, df = 29, p-value = 0.00255
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.408659 5.991341
## sample estimates:
## mean of x
##       3.7
```

Confidence interval is all above zero, so we can easily suggest that with supp = OJ teeth grow faster.

Let's check on dose:

```
d1<-ToothGrowth$len[ToothGrowth$dose == 0.5]
d2<-ToothGrowth$len[ToothGrowth$dose == 1]
d3<-ToothGrowth$len[ToothGrowth$dose == 2]
diff21<-d2-d1
diff31<-d3-d1
diff32<-d3-d2
```

1 against 0.5:

```
t.test(diff21)
```

```
##
##  One Sample t-test
##
## data:  diff21
## t = 6.9669, df = 19, p-value = 1.225e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   6.387121 11.872879
## sample estimates:
## mean of x
##      9.13
```

Again, confidence interval above zero, so when dose is 1, teeth grow better than with dose = 0.5

2 against 0.5:

```
t.test(diff31)
```

```
##
##  One Sample t-test
```

```
##
## data:  diff31
## t = 11.2915, df = 19, p-value = 7.19e-10
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  12.6228 18.3672
## sample estimates:
## mean of x
##    15.495
```

again, confidence interval above zero.

2 against 1:

```
t.test(diff32)
```

```
##
##  One Sample t-test
##
## data:  diff32
## t = 4.6046, df = 19, p-value = 0.0001934
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.471814 9.258186
## sample estimates:
## mean of x
##     6.365
```

again, confidence interval above zero.

I should've thought how these two factors (dose and supp) may overlap and what I can do to deal with that but I have 7 minutes till deadline, sorry :(.