

Basic Inferential Data Analysis on R Dataset–ToothGrowth

by Mingzhu Ye

Overview

We're going to load the “ToothGrowth” data and perform some basic exploratory data analysis, provide a basic summary of the data, use confidence intervals and hypothesis tests to compare tooth growth by supp and dose. To figure out details about “ToothGrowth” please check the ‘?ToothGrowth’ in R Help Documentation.

Data Analysis

We will do Data Processing, Data Exploratory Analysis, Regression Models for “ToothGrowth” for this section.

Data Processing

This is the head of “ToothGrowth”, there're 3 variables in the data frame.

```
##      len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
```

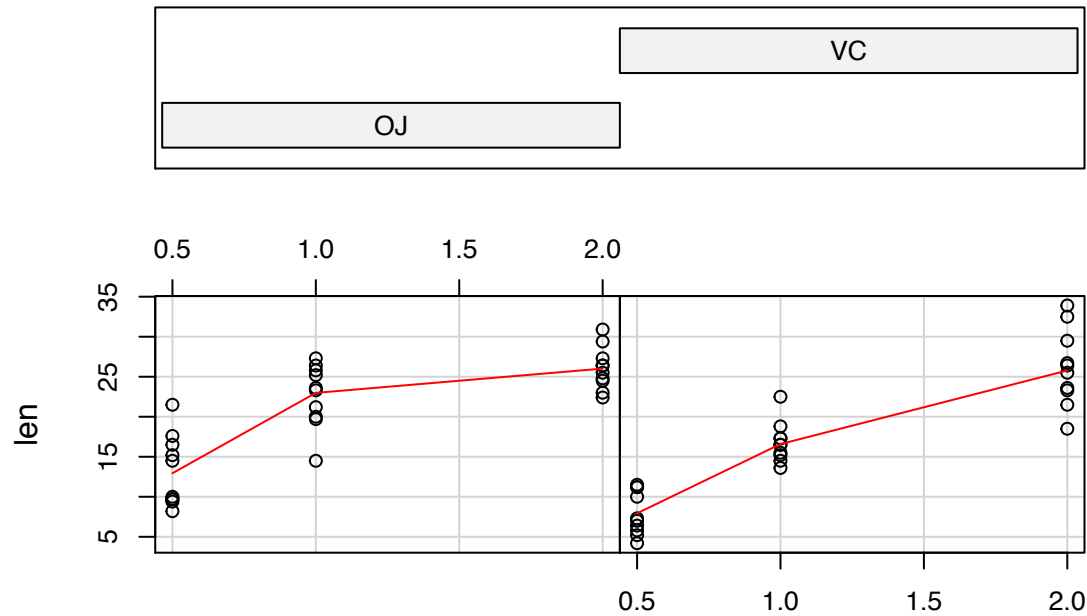
This is the summary of “ToothGrowth”. Specifically, ‘supp’ is an abbreviation for supplement which is divided into two groups, OJ & VC.

```
##      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
```

Data Exploratory Analysis

```
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

As observed in the coplot, tooth length has positive correlation with dose when supp is either OJ or VC.

Regression Models for “ToothGrowth”

```
##
## Call:
## lm(formula = len ~ dose + supp, data = ToothGrowth)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.600 -3.700  0.373  2.116  8.800
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.2725     1.2824   7.231 1.31e-09 ***
## dose          9.7636     0.8768  11.135 6.31e-16 ***
## suppVC       -3.7000     1.0936  -3.383  0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.236 on 57 degrees of freedom
## Multiple R-squared:  0.7038, Adjusted R-squared:  0.6934
## F-statistic: 67.72 on 2 and 57 DF,  p-value: 8.716e-16
```

As shown in the regression model, average length for guinea pigs is 9.7725 without any types of supplement(VC or OJ). Coefficients of dose is 9.7636. Coefficients of supp is -3.7000.

Relevant Confidence Intervals and Tests

Split data based on dosage of 0.5, 1.0, 2.0

```
d_0.5 <- subset(ToothGrowth, dose == 0.5)
d_1.0 <- subset(ToothGrowth, dose == 1.0)
d_2.0 <- subset(ToothGrowth, dose == 2.0)
```

T-test between len and supp under dose = 0.5

```
test0.5 <- t.test (len ~ supp, paired = FALSE, var.equal = FALSE, data = d_0.5)
test0.5$p.value
```

```
## [1] 0.006358607
```

T-test between len and supp under dose = 1.0

```
test1.0 <- t.test (len ~ supp, paired = FALSE, var.equal = FALSE, data = d_1.0)
test1.0$p.value
```

```
## [1] 0.001038376
```

T-test between len and supp under dose = 2.0

```
test2.0 <- t.test (len ~ supp, paired = FALSE, var.equal = FALSE, data = d_2.0)
test2.0$p.value
```

```
## [1] 0.9638516
```

T-test Observation:

T-test shows dose of 0.5 and 1.0 have relatively low P-value(0.006358607, 0.001038376), while dose of 2.0 has pretty high P-value(0.9638516). For P-values > 0.05, we fail to reject the null hypothesis, there are no differences in supplements between different dosage. Difference in mean values between the supplements is not significant. On the contrary, likewise.

Assumption

1. Supplements(OJ or VC) do have effects on the length of Tooth Growth.
2. The guinea pigs sample is not diverse, and sample size is not big enough.
3. Samples are unpaired and variances are unequal.

Conclusion

Supplements of OJ or VC have significantly different effects on pig's tooth length growth for lower(0.5, 1.0) dosages of supplements according to the T-test. However, there is no significant difference in tooth length growth when dosage is high(2.0). OJ yields a longer tooth growth under the dosage of 0.5 or 1.0.