Statistical Inference Project Part 2

By Aiman D.

Overview

Part 2 of the project will analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.
4. State your conclusions and the assumptions needed for your conclusions.

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.

Source

C. I. Bliss (1952) The Statistics of Bioassay. Academic Press.

References

McNeil, D. R. (1977) Interactive Data Analysis. New York: Wiley.

Data Loading and Analysis

```
> library(datasets)
> data(ToothGrowth)
> 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
```

```
> 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 ...
```

Basic Summary of Data

```
> ToothGrowth$dose <- as.factor(ToothGrowth$dose)
> table(ToothGrowth$supp, ToothGrowth$dose)

      0.5  1  2
  OJ   10 10 10
  VC   10 10 10

> 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
> mean(ToothGrowth$len)
[1] 18.81333
> sd(ToothGrowth$len)
[1] 7.649315
```

Plotting Data

```
1  require(ggplot2)
2  plot <- ggplot(ToothGrowth, aes(x=factor(dose),y=len,fill=factor(dose)))
3  plot + geom_boxplot(notch=F) + facet_grid(.~supp) +
4    scale_x_discrete("Dosage (mg)") +
5    scale_y_continuous("Tooth Length") +
6    ggtitle("Dosage and Supplement Effect on Teeth Growth")
```
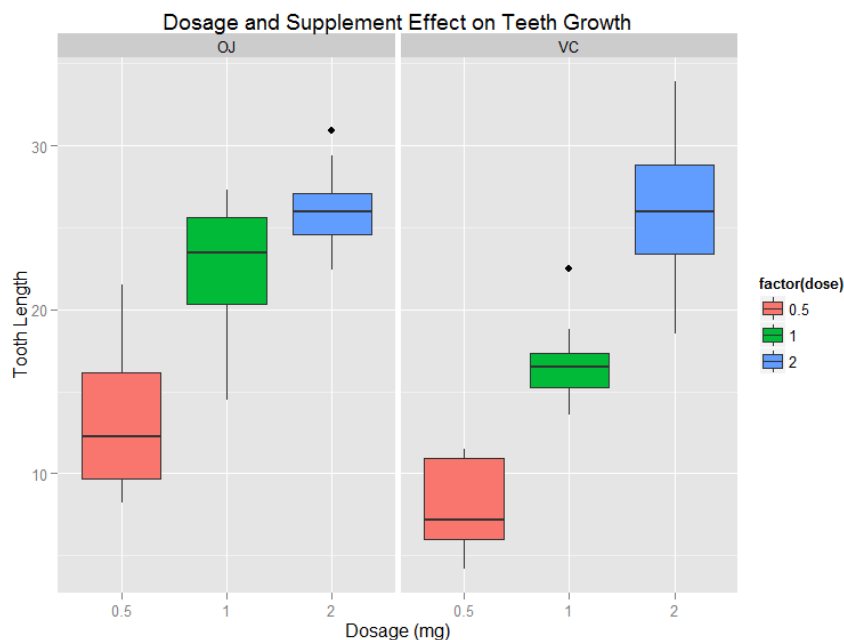
Confidence Intervals and/or hypothesis tests to compare tooth growth by supp and dose

```
1  supp.t1 <- t.test(len~supp, paired=F, var.equal=T, data=ToothGrowth)
2  supp.t2 <- t.test(len~supp, paired=F, var.equal=F, data=ToothGrowth)
3  supp.result <- data.frame("p-value"=c(supp.t1$p.value, supp.t2$p.value),
4                "Low Conf"=c(supp.t1$conf[1],supp.t2$conf[1]),
5                "High Conf"=c(supp.t1$conf[2],supp.t2$conf[2]),
6                row.names=c("Equal Var","UnEqual Var"))
7  supp.result
```

```
                p.value    Low.Conf High.Conf
Equal Var    0.06039337 -0.1670064  7.567006
UnEqual Var  0.06063451 -0.1710156  7.571016
```

Conclusions

Analyzing the data and the plots:

1. The dosage seems to have the biggest effect on the growth if teeth.
2. OJ is better for teeth growth than VC at lower dosages (0.5 mg - 1 mg).
3. OJ and VC at the higher dosage (2 mg) seem to be statistically indifferent but with a slight edge for VC.