Statistical Inference Course Project

Part 2: Basic inferential data analysis

Omer Kara

Friday, January 15, 2016

Source code for this entire report can be found here.

Summary of the Data and Exploratory Analysis

In the second part of the project, we analyze the `ToothGrowth` data in the R datasets package. The data is set of 60 observations, and 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).

```
data <- ToothGrowth
data$dose <- as.factor(data$dose)
str(data) ## Summary of data frame.
```

```
## '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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

```
rbind(head(data, 5), tail(data, 5)) ## First 5 and last 5 rows of data
```

```
##     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
## 56 30.9   OJ    2
## 57 26.4   OJ    2
## 58 27.3   OJ    2
## 59 29.4   OJ    2
## 60 23.0   OJ    2
```

```
table(data$supp, data$dose) ## Summary of levels of Delivery Methods and Dose.
```

```
##
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

```r
summary(data) ## Summary of data.
```

```
##      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
```

The following four figures shows the tooth length of guniea pigs by supplement type and dose. Note that Figure 1 is by supplement type, Figure 2 is by dose, and Figure 3 and 4 are by supplement and dose type.

Tooth Length of Guniea Pigs by Supplement Type and Dose
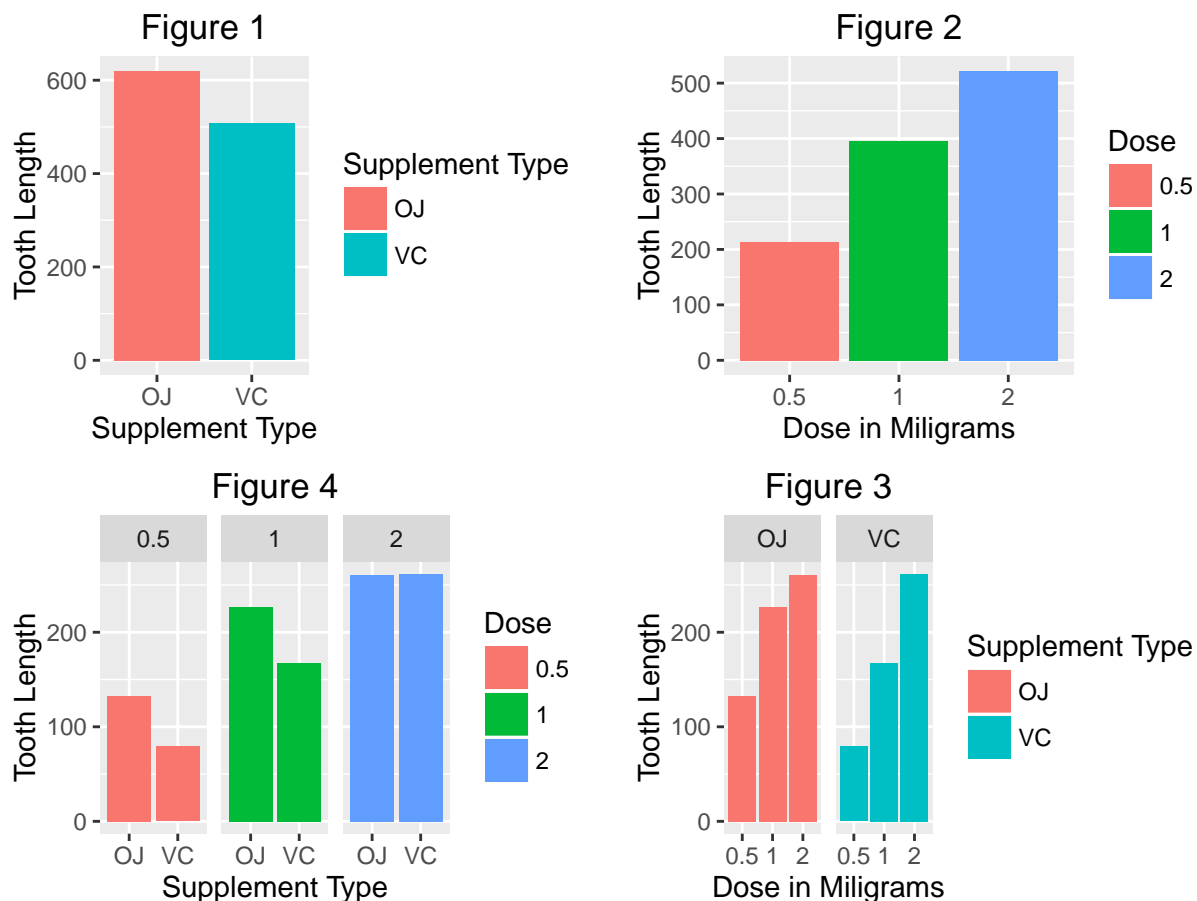
Confidence Interval and Hypothesis Testing

In order to understand Vitamin C's affect on tooth growth, we will conduct the following confidence interval testing scenarios:
- By supplement type.
- By dose.
- Dose by aach each supplement type.
- Supplement type by each dosage.

For each of the comparisons, we will subset data frame appropriately and utilize the `t.test` R function to determine each scenarios confidence interval, p-value and hypothesis test results.[1^] As the theory suggests that a formal two tailed hypothesis test with null equal to 0 is equavalent to confidence interval testing, the inference will be given via hypothesis test results.

The general results from the tables below are;

- Supplement type in general does not affect tooth length.
- Dose level in general affects tooth length. It seems that the more vitamin C, the more tooth length.
- Does level by each supplement type affects tooth lenght. It seems that the more vitamin C, the more tooth length.
- Supplement type in 1 and 0.5 miligrams dose levels affects the tooth length but does not have any effect in 2 miligrams does level.

Table 1: Comparison of Supplement Type

Comparison	CI.95	P.Value	Hypothesis.Test.Result
sVC.sOJ	(-7.571, 0.171)	0.06	Fail to Reject Null: Statistically there is no difference

Table 2: Comparison of Dose

Comparison	CI.95	P.Value	Hypothesis.Test.Result
d2.d1	(3.7335, 8.9965)	0	Reject Null: Statistically there is a difference.
d1.d0.5	(6.2762, 11.9838)	0	Reject Null: Statistically there is a difference.
d2.d0.5	(12.8338, 18.1562)	0	Reject Null: Statistically there is a difference.

Table 3: Comparison of Dose by Each Supplement Type

Supplement.Type	Comparison	CI.95	P.Value	Hypothesis.Test.Result
VC	d2.d1	(5.6857, 13.0543)	0.00	Reject Null: Statistically there is a difference.
VC	d1.d0.5	(6.3143, 11.2657)	0.00	Reject Null: Statistically there is a difference.
VC	d2.d0.5	(14.4185, 21.9015)	0.00	Reject Null: Statistically there is a difference.
OJ	d2.d1	(0.1886, 6.5314)	0.04	Reject Null: Statistically there is a difference.
OJ	d1.d0.5	(5.5244, 13.4156)	0.00	Reject Null: Statistically there is a difference.
OJ	d2.d0.5	(9.3248, 16.3352)	0.00	Reject Null: Statistically there is a difference.

Table 4: Comparison of Supplement Type by Each Dose

Dose	Comparison	CI.95	P.Value	Hypothesis.Test.Result
d2	sVC.sOJ	(-3.6381, 3.7981)	0.96	Fail to Reject Null: Statistically there is no difference
d1	sVC.sOJ	(-9.0579, -2.8021)	0.00	Reject Null: Statistically there is a difference.
d0.5	sVC.sOJ	(-8.7809, -1.7191)	0.01	Reject Null: Statistically there is a difference.

Conclusions

1. As vitamin c dose increases alone, the tooth length increases as well.
2. Supplement type in general does not affect tooth growth.
3. The supplement type with a 0.5 and 1.0 dose size affect the tooth length.
4. When the dose size reached 2.0 milligrams, there is no difference between supplment types.

Assumptions

1. The samples are independent.
2. The confidence intervals are assumed to not be paired, i.e. we are not comparing two different supplement types from individual guinea pig.
3. Since the figures clearly suggest the spread (variance) is not equaly among the supplement type or dosage, unqual variance is assumed in all test.
4. The distribution is approximately normal.