The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Statistical Inference Course Project 2

Xinchun Tian

Exploratory Data Analysis

Using some simple *R* tools to do some initial exploratory data anlysis. 1. Show part of the data in a table (measurements 25-35)

	len	supp	dose
25	26.4	VC	2.0
26	32.5	VC	2.0
27	26.7	VC	2.0
28	21.5	VC	2.0
29	23.3	VC	2.0
30	29.5	VC	2.0
31	15.2	OJ	0.5
32	21.5	OJ	0.5
33	17.6	OJ	0.5
34	9.7	OJ	0.5
35	14.5	OJ	0.5

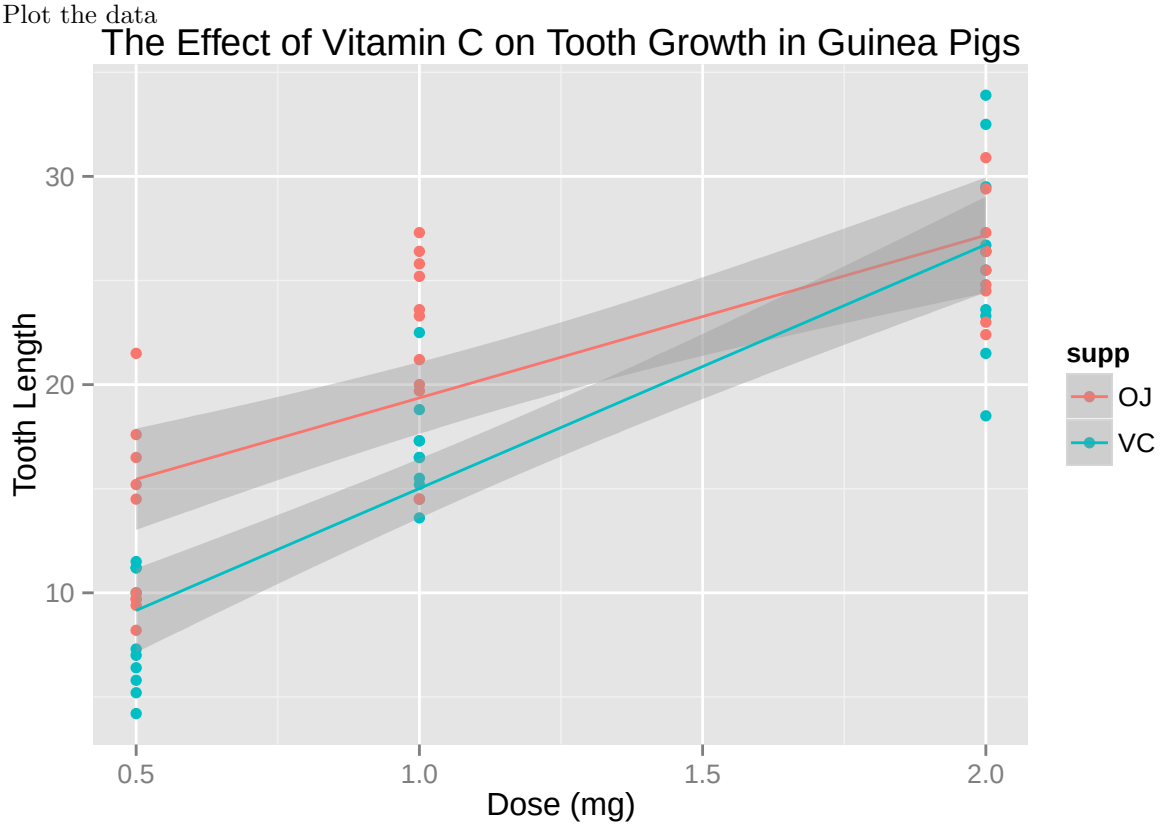
2. Show the summary of the data using *summary*

```
##       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
```

3. Show the summary of the data using *str*

```
## '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 ...
```

4. Plot the data



From the above exploratory data analysis, we can have obtain the following observations about the ToothGrowth dataset.
1. There are 60 observations, 3 variables,
2. Two factors: VC and OJ.
According to the help page of *R*, the summary of the ToothGrowth dataset can be "*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).*"

Two-group Confidence Intervals and Hypothesis Tests

Compare VC vs OJ effects at different dose levels

There are three dose levels (0.5, 1.0 and 2.0 mg).
1. dose is 0.5 mg

```
ToothGrowth_0 <- subset(ToothGrowth, dose==0.5)
```

The variance of len in the group of VC, 7.544, is different from the variance of len in the group of OJ, 19.889, therefore the T-test will use the option of "va.equal=FALSE".

```
interval <- rbind(t.test(len~I(relevel(supp,"OJ")), data=ToothGrowth_0,
                         var.equal=FALSE, paired=FALSE)$conf)
```

The 95% confidence interval is [1.719, 8.781], which means "OJ" is more effective than "VC" on teeth growth at 0.5 mg dose level.

2. dose is 1.0 mg

```
ToothGrowth_1 <- subset(ToothGrowth, dose==1.0)
```

The variance of len in the group of VC, 6.327, is different from the variance of len in the group of OJ, 15.296, therefore the T-test will use the option of "va.equal=FALSE".

```
interval <- rbind(t.test(len~supp, data=ToothGrowth_1,
                         var.equal=FALSE, paired=FALSE)$conf)
```

The 95% confidence interval is [2.802, 9.058], which means "OJ" is more effective than "VC" on teeth growth at 1.0 mg dose level.

3. dose is 2.0 mg

```
ToothGrowth_2 <- subset(ToothGrowth, dose==2.0)
```

The variance of len in the group of VC, 23.018, is different from the variance of len in the group of OJ, 7.049, therefore the T-test will use the option of "va.equal=FALSE".

```
interval <- rbind(t.test(len~supp, data=ToothGrowth_2,
                         var.equal=FALSE, paired=FALSE)$conf)
```

The 95% confidence interval is [-3.798, 3.638], which means "OJ" is not different from "VC" on teeth growth at 2.0 mg dose level.

Compare dose effects (0.5 and 2.0 mg) at VC or OJ

There are three dose levels (0.5, 1.0 and 2.0 mg).
1. supp is VC

```
ToothGrowth_0 <- subset(ToothGrowth, supp=="VC" & (dose==0.5 | dose==2.0))
```

The variance of len at 0.5 mg dose, 7.544, is different from the variance of len at 2.0 mg dose, 23.018, therefore the T-test will use the option of "va.equal=FALSE".

```
interval <- rbind(t.test(len~I(relevel(factor(dose),2)), data=ToothGrowth_0,
                         var.equal=FALSE, paired=FALSE)$conf)
```

The 95% confidence interval is [14.418, 21.902], which means 2.0 mg dose is more effective than 0.5 mg dose on teeth growth in the group of VC.

2. supp is OJ

```
ToothGrowth_1 <- subset(ToothGrowth, supp=="OJ" & (dose==0.5 | dose==2.0))
```

The variance of len at 0.5 mg dose, 19.889, is different from the variance of len at 2.0 mg dose, 7.049, therefore the T-test will use the option of "va.equal=FALSE".

```
interval <- rbind(t.test(len~I(relevel(factor(dose),2)), data=ToothGrowth_1,
                         var.equal=FALSE, paired=FALSE)$conf)
```

The 95% confidence interval is [9.325, 16.335], which means 2.0 mg dose is more effective than 0.5 mg dose on teeth growth in the group of OJ.

Similarly we can compare the dose level of 0.5 vs 1.0, 1.0 vs 2.0 within each OJ and VC group by repeating the above t-test procedures.

The assumption in the above analysis is **the samples are iid and normally distributed**.