ProjectPart2

Govs

12/20/2014

Coursera Statistical Inference December 2014

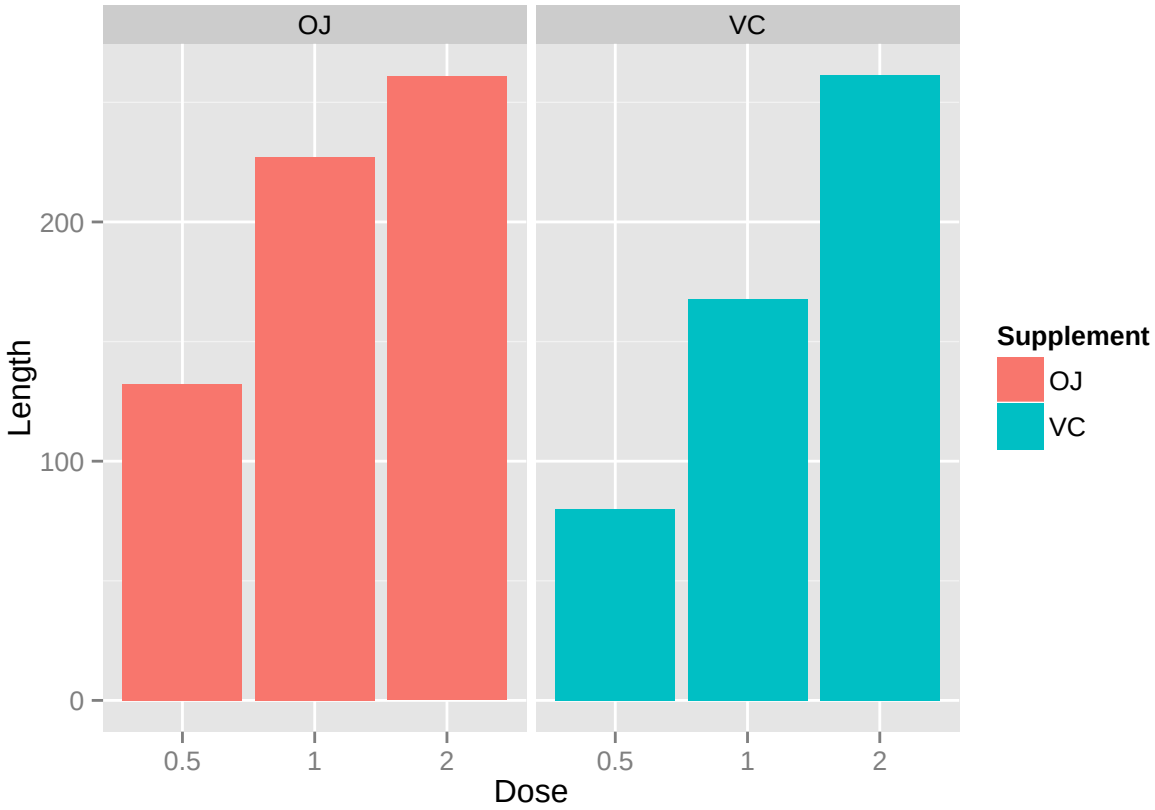
The ToothGrowth dataset has observations about the length of teeth of guinea pigs when Vitamin C is delivered by two methods orange juice and ascorbic acid.

Exploring the Data

```
library(datasets)
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 ...
```

Lets see these observations with graphs:



We can observe that the tooth length increases when the dose is bigger.

Summary

Now here are some tables with summary data of the observations:

Mean of length:

```
with(ToothGrowth, tapply(len, list(supp, dose), mean))
```

```
##      0.5     1     2
## OJ 13.23 22.70 26.06
## VC  7.98 16.77 26.14
```

Median of length:

```
with(ToothGrowth, tapply(len, list(supp, dose), median))
```

```
##      0.5     1     2
## OJ 12.25 23.45 25.95
## VC  7.15 16.50 25.95
```

Variance of length:

```
with(ToothGrowth, tapply(len, list(supp, dose), var))
```

```
##       0.5         1         2
## OJ 19.889 15.295556  7.049333
## VC  7.544  6.326778 23.018222
```

Hypothesis Tests

Now we are going to perform some hypothesis test to see the effect of supplement type and dose level in tooth growth.

First the supplement, we are testing the null hypothesis that the supplement type doesn't affect the tooth growth.

```
t.test(len ~ supp, data = ToothGrowth, paired = FALSE, var.equal = FALSE,conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

The p-value of the test is greater than 0.05 (alpha of 5%), so we cannot reject the hypothesis, we conclude that the supplement type has no effect in the tooth growth.

Now lets see what happen with dosage levels, the null hypothesis is that increasing the dose doesn't affect the tooth growth. We are comparing only three cases, whith dose 0.5vs1,0.5vs2 and 1vs2.

```
t.test(ToothGrowth$len[ToothGrowth$dose==0.5], ToothGrowth$len[ToothGrowth$dose==1], paired = FALSE, va
```

```
## [1] 1.268301e-07
```

```
t.test(ToothGrowth$len[ToothGrowth$dose==0.5], ToothGrowth$len[ToothGrowth$dose==2], paired = FALSE, va
```

```
## [1] 4.397525e-14
```

```
t.test(ToothGrowth$len[ToothGrowth$dose==1], ToothGrowth$len[ToothGrowth$dose==2], paired = FALSE, var.
```

```
##
##  Welch Two Sample t-test
##
## data:  ToothGrowth$len[ToothGrowth$dose == 1] and ToothGrowth$len[ToothGrowth$dose == 2]
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

As you can see the p-value in the three cases is less than 0.05 (alpha of 5%) so we can reject the null hypothesis,and conclude that the change in doses actually has an effect in tooth growth.

Our assumptions were,not equal variances and unpaired in all cases, we have a table of variances and also asume (when in doubt) that they were unequal, about the pairing actually the data of distinct doses and supp have distinct length.