Statistic inference report - Part 2

Tom

Sunday, June 14, 2015

Overview

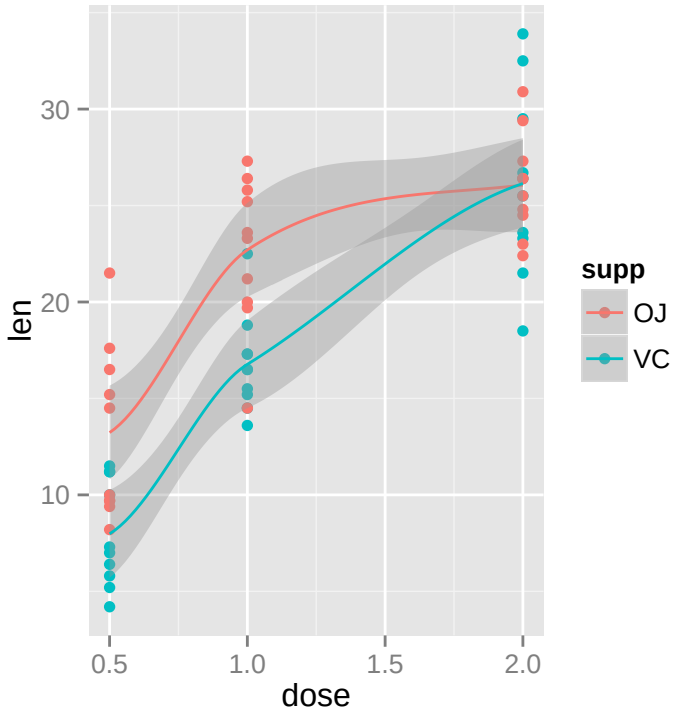
In this report, we are going to explore the following dataset ToothGrowth.

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).

Basic summary of the data

First, we are going to make a plot to provide a basic summary of the data.

```
require(ggplot2)
data(ToothGrowth)
qplot(dose, len, data=ToothGrowth, color=supp) + geom_smooth(data=ToothGrowth)
```



The following matrix present the number of sample by dose and delivery methods. Each different set of dose and delivery methods contains 10 sample.

```r
table(ToothGrowth$dose, ToothGrowth$supp)
```

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

Study

We are going to study if difference delivery methods can result in different tooth length under different dose. We conduct the study using hypothesis testing. $(1-\alpha) = 95\%$ interval is adopted.

Assumption

Compensate for Multiple testing

Because we are conducting three tests, and each of them is similiar to each other, the $\alpha_{fwer} = \alpha/3 = 1.667\%$ is used.

Hypothesis testing

In this study, we does not assume 2 groups has the same variance, and as the tests for difference supplements and dose do not apply to the same subject, the sample is not paired.

Dose level=0.5

Hypothesis test result:

```r
dose05 <- subset(ToothGrowth, dose==0.5)
t.test(len~supp, paired=FALSE, var.equal=FALSE, data=dose05)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

For 95% confidence interval, the test statistic= $3.1697 > t_{(0.99167,14.969)} = 2.69463$. The null hypothesis is rejected.

Dose level=1.0

```
dose10 <- subset(ToothGrowth, dose==1.0)
t.test(len~supp, paired=FALSE, var.equal=FALSE, data=dose10)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

For 95% confidence interval, the test statistic= 4.0328 $> t_{(0.99167,15.358)} = 2.686181$. The null hypothesis is rejected.

Dose level=2.0

```
dose20 <- subset(ToothGrowth, dose==2.0)
t.test(len~supp, paired=FALSE, var.equal=FALSE, data=dose20)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

For 95% confidence interval, the test statistic= -0.046136 $> t_{(0.00833,14.04)} = -2.716925$. The null hypothesis should not be rejected.

Conclusion

In this study, we find out that for low dose level, the delivery method matters, using orange juice as supplement helps the growth of teeth and is more effective than using ascorbic acid. For dose level as high as 2.0, the effect disappears.