Analysis of Tooth Growth Data

Matthew Mariano

February 22, 2015

Introduction

The purpose of this report is to analyze the ToothGrowth data. In this analysis I will reform the data to make it workable, make some initial plots and finally apply some confidence intervals. Since this data consists of subjects on different supplements and differnt dosages, analysis may be done on both paired and nonpaired data. So for example I may consider the pair where each subject is on the same dose but the supplement changes.

```
library(datasets)
library(ggplot2)
data(ToothGrowth)
library(reshape2)
```

Look at the data. For conveience in this report I do a head, but in my analysis I opened the whole data set which is rather small.

```
tg=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
```

```
tg$dose = as.factor(tg$dose)
```

Notice that the data is for 10 subjects , their are two supplements (VC and OJ) and there are three dosages (.5, 1.0 and 2.0). I want to compare the effects of different supplements at each dosage. To do this I need to reshape the data. I will do this so that there will be three columns for dose (dose1 , dose2, dose3) and the length associated with each one will be in the respective column.

- id a sequuence of id for each subject
- tgnew = the reformed dataset
- d1 = the data for the supplement OJ
- d2 = the data for the supplement VC

```
id = 1:10
tgnew = cbind(id, tg)
tgnew = dcast(tgnew, id + supp ~ dose, value.var = "len")
names(tgnew)[3:5] = c("dose1", "dose2", "dose3")
tgnew = tgnew[order(tgnew$supp),]
d1=tgnew[tgnew[,c("supp")]=="OJ",]
d2=tgnew[tgnew[,c("supp")]=="VC",]
```

Define some functions for plotting

```
plot1=function(supp)
{
  x=tgnew[,c("supp")]==supp
  d1=tgnew[x,c("dose1","dose2","dose3")]
  # note you need the mode here or else a dataframe is still returned by as.vector
  c1=as.vector(cumsum(d1)[nrow(d1),],mode="numeric")
  p1=barplot(c1, main=supp,
          xlab="dose(mg)",col=c("red","green","blue"))
}
```

Confidence Tests for Paired Data

paired data

```
oj1=d1[,c("dose1")]
vc1=d2[,c("dose1")]

difference <- oj1-vc1
mn <- mean(difference); s <- sd(difference); n =10
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1] 1.263458 9.236542
```

```
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 2.9791, df = 9, p-value = 0.01547
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.263458 9.236542
## sample estimates:
## mean of x
##      5.25
```

```
oj2=d1[,c("dose2")]
vc2=d2[,c("dose2")]

difference <- oj2-vc2
mn <- mean(difference); s <- sd(difference); n =10
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1] 1.951911 9.908089
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 3.3721, df = 9, p-value = 0.008229
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.951911 9.908089
## sample estimates:
## mean of x
##      5.93
```

```r
oj3=d1[,c("dose3")]
vc3=d2[,c("dose3")]

difference <- oj3-vc3
mn <- mean(difference); s <- sd(difference); n =10
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1] -4.328976  4.168976
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = -0.0426, df = 9, p-value = 0.967
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -4.328976  4.168976
## sample estimates:
## mean of x
##     -0.08
```

Confidence Tests Unpaired

```r
message("COMPARISON OF DOSAGE FOR SUPPLEMENT OJ")
```

```
## COMPARISON OF DOSAGE FOR SUPPLEMENT OJ
```

```r
message("first compare dose 2 to dose 1")
```

```
## first compare dose 2 to dose 1
```

```r
difference <- d1$dose2-d1$dose1
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference9.47
```

```r
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1]  4.324616 14.615384
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 4.1635, df = 9, p-value = 0.002435
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   4.324616 14.615384
## sample estimates:
## mean of x
##      9.47
```

```r
message("second compare dose 3 to dose 1")
```

```
## second compare dose 3 to dose 1
```

```r
difference <- d1$dose3-d1$dose1
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference12.83
```

```r
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1]  8.956042 16.703958
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 7.4919, df = 9, p-value = 3.724e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   8.956042 16.703958
## sample estimates:
## mean of x
##     12.83
```

```r
message("second compare dose 3 to dose 2")
```

```
## second compare dose 3 to dose 2
```

```r
difference <- d1$dose3-d1$dose2
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference3.36
```

```r
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1] -0.5509376  7.2709376
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 1.9435, df = 9, p-value = 0.08384
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.5509376  7.2709376
## sample estimates:
## mean of x
##      3.36
```

```r
message("--------------------------------------------------------------")
```

```
## --------------------------------------------------------------
```

```r
message("COMPARISON OF DOSAGE FOR SUPPLEMENT VC")
```

```
## COMPARISON OF DOSAGE FOR SUPPLEMENT VC
```

```r
message("first compare dose 2 to dose 1")
```

```
## first compare dose 2 to dose 1
```

```r
difference <- d2$dose2-d2$dose1
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference8.79
```

```r
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1]  5.549601 12.030399
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 6.1364, df = 9, p-value = 0.0001715
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   5.549601 12.030399
## sample estimates:
## mean of x
##      8.79
```

```r
message("second compare dose 3 to dose 1")
```

```
## second compare dose 3 to dose 1
```

```r
difference <- d2$dose3-d2$dose1
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference18.16
```

```r
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1] 13.9643 22.3557
```

```r
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 9.7912, df = 9, p-value = 4.264e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  13.9643 22.3557
## sample estimates:
## mean of x
##     18.16
```

```r
message("second compare dose 3 to dose 2")
```

```
## second compare dose 3 to dose 2
```

```
difference <- d2$dose3-d2$dose2
mn <- mean(difference); s <- sd(difference); n =10
message("Mean difference",mn)
```

```
## Mean difference9.37
```

```
mn + c(-1, 1) * qt(.975, n-1) * s / sqrt(n)
```

```
## [1]  5.405082 13.334918
```

```
t.test(difference)
```

```
##
##  One Sample t-test
##
## data:  difference
## t = 5.346, df = 9, p-value = 0.0004648
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   5.405082 13.334918
## sample estimates:
## mean of x
##      9.37
```

Conclusion

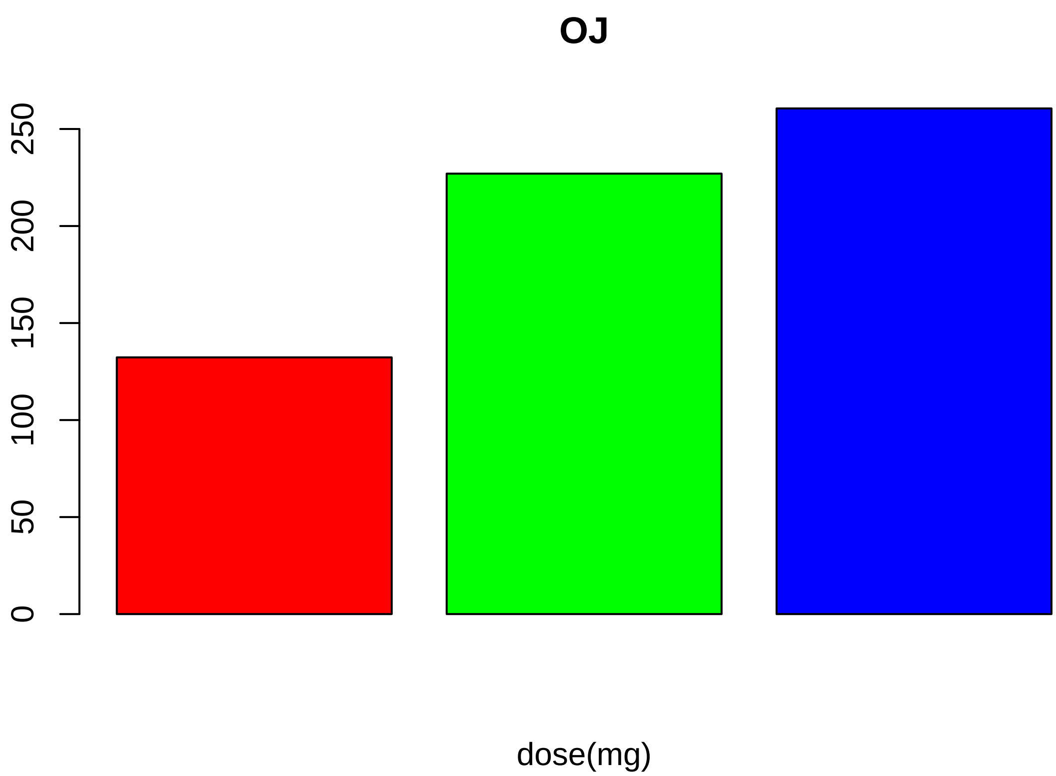
The paired confidence intervals indicate that we should reject the null hypothesis for the first two dosages. To see that consider the mean of the differnce for each case and also consider the p value from the test. The first two tests have large means and small p values. In other words the delivery method, or supplement is important for tooth length when the dosage is .5 to 1.0 mg, however at 2.0 mg the delivery method does not determine the tooth length.

For the unpaired tests, testing was done to see if the dosage for a particular supplement is statistically significant for tooth length. Here we see that all of the mean in difference is large. And also the p value from the test is very small. In this case the null hypothesis must be rejected for both supplements and all doses. In other words the dosage is statistically significant in determining tooth length for either VC or OJ.

Appendix

Figure 1 Tooth Length versus Dosage for Supplement OJ

```
plot1("OJ")
```




```r
# Figure 1 Tooth Length versus Dosage for Supplement OJ
plot1("VC")
```