

Statistical__Inference__Part2

emilliman5

October 25, 2014

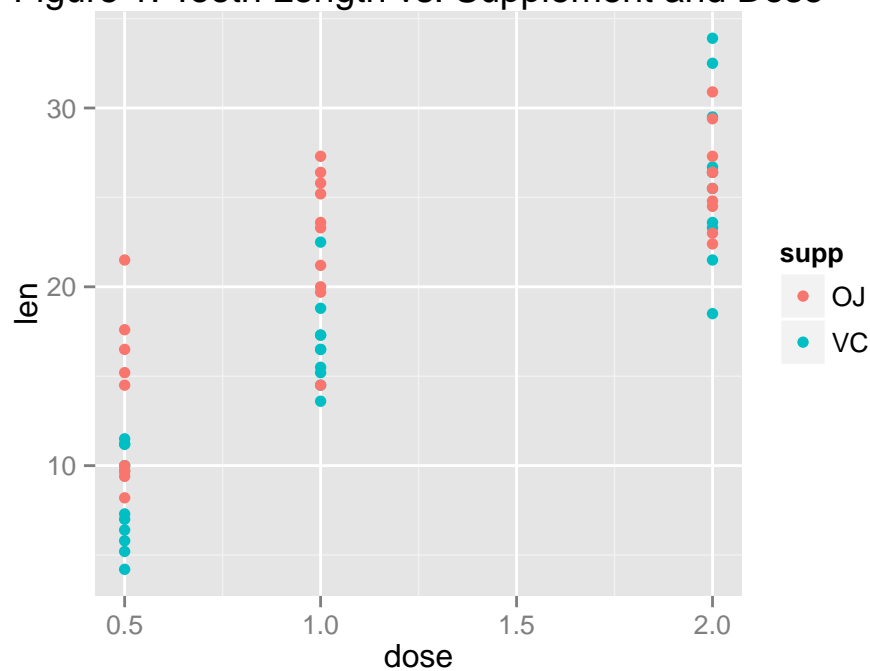
```
## KernSmooth 2.23 loaded
## Copyright M. P. Wand 1997-2009
##
## Attaching package: 'gplots'
##
## The following object is masked from 'package:stats':
##
##     lowess
```

1. Provide a basic summary of the data.

Table 1 displays the toothgrowth data's mean, sd, and quantiles. Figure 1 displays the toothgrowth data graphically separated by dose and colored by supplement.

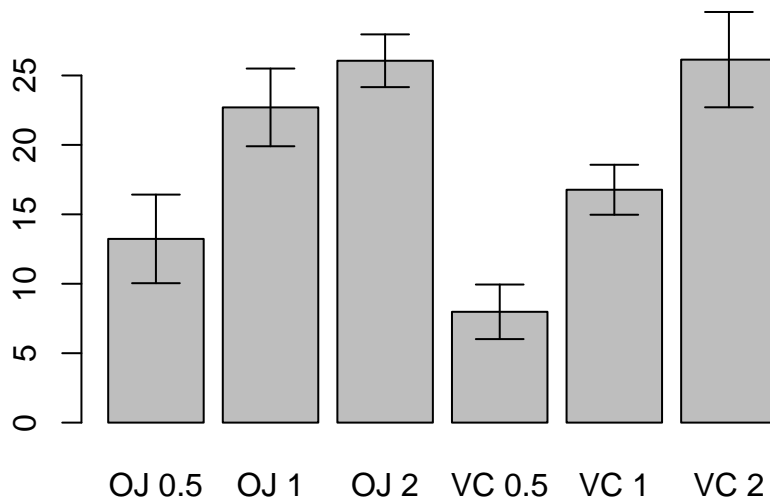
##	Dose	Supp	Mean.Length	S.D	0%	25%	50%	75%	100%
## 1	0.5	OJ	13.23	4.460	8.2	9.70	12.25	16.18	21.5
## 2	1.0	OJ	22.70	3.911	14.5	20.30	23.45	25.65	27.3
## 3	2.0	OJ	26.06	2.655	22.4	24.57	25.95	27.08	30.9
## 4	0.5	VC	7.98	2.747	4.2	5.95	7.15	10.90	11.5
## 5	1.0	VC	16.77	2.515	13.6	15.27	16.50	17.30	22.5
## 6	2.0	VC	26.14	4.798	18.5	23.38	25.95	28.80	33.9

Figure 1: Tooth Length vs. Supplement and Dose



2. Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose. (Use the techniques from class even if there's other approaches worth considering)

Figure 2: Average Length \pm 95%CI



```
pairwise.t.test(ToothGrowth$len, g = paste(ToothGrowth$supp, ToothGrowth$dose), paired = FALSE, p.adjust

##
## Pairwise comparisons using t tests with pooled SD
##
## data: ToothGrowth$len and paste(ToothGrowth$supp, ToothGrowth$dose)
##
##      OJ 0.5  OJ 1    OJ 2    VC 0.5  VC 1
## OJ 1  3.2e-07 -        -        -        -
## OJ 2  1.4e-10 0.04335 -        -        -
## VC 0.5 0.00209 2.0e-12 1.3e-15 -        -
## VC 1   0.03365 0.00059 4.8e-07 1.5e-06 -
## VC 2   1.2e-10 0.03878 0.96089 1.1e-15 4.0e-07
##
## P value adjustment method: none
```

Figure 2 displays the average toothgrowth length for each supplement and dose and its respective 95% CIs. Table 2 shows all pairwise pvalues between doses and supplements. From the table (and figure), we can see that there are some doses and supplements that in which there are more extreme observations made than is expected by chance (p-values ≤ 0.5).

3. State your conclusions and the assumptions needed for your conclusions.

From the data above we can see that increasing the dose of regardless of the supplement seems to increase tooth length greater than is expected by chance. However, the data also suggests that the lower doses of OJ are more effective than than the lower doses of VC but high doses of VC and OJ are statistically indistinguishable. One caveat to this analysis is a lack of mulitple hypothesis testing. The more pair-wise comapraisons that we make the greater the channce we will obtain a large t-statistic by chance (in fact w would expect this to occur for 1 in 20 comparisons)

Appendix: Code

Load data:

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

```
table<-aggregate(ToothGrowth$len,by = list(ToothGrowth$dose, ToothGrowth$supp), mean)
table<-cbind(table,aggregate(ToothGrowth$len,by = list(ToothGrowth$dose, ToothGrowth$supp), sd)[,3],agg
colnames(table)<-c("Dose", "Supp", "Mean.Length", "S.D", "0%", "25%", "50%", "75%", "100%")
table
```

```
qplot(dose, len, data=ToothGrowth, color=supp, main = "Figure 1: Tooth Length vs. Supplement and Dose")
```

```
TGmean<-melt(tapply(ToothGrowth$len,INDEX = list(ToothGrowth$dose, ToothGrowth$supp), FUN=mean))
ul<-melt(tapply(ToothGrowth$len,INDEX = list(ToothGrowth$dose, ToothGrowth$supp), function(x){ t.test(x)
ll<-melt(tapply(ToothGrowth$len,INDEX = list(ToothGrowth$dose, ToothGrowth$supp), function(x){ t.test(x)

barplot2(TGmean$value, plot.ci=TRUE, ci.l=ll$value, ci.u=ul$value, names.arg = paste(TGmean$Var2, TGmean$
pairwise.t.test(ToothGrowth$len, g = paste(ToothGrowth$supp,ToothGrowth$dose), paired = FALSE, p.adjust
```