Statistical Inference Part 2:Toothgrowth Data

Paul Nakonechny

October 24, 2014

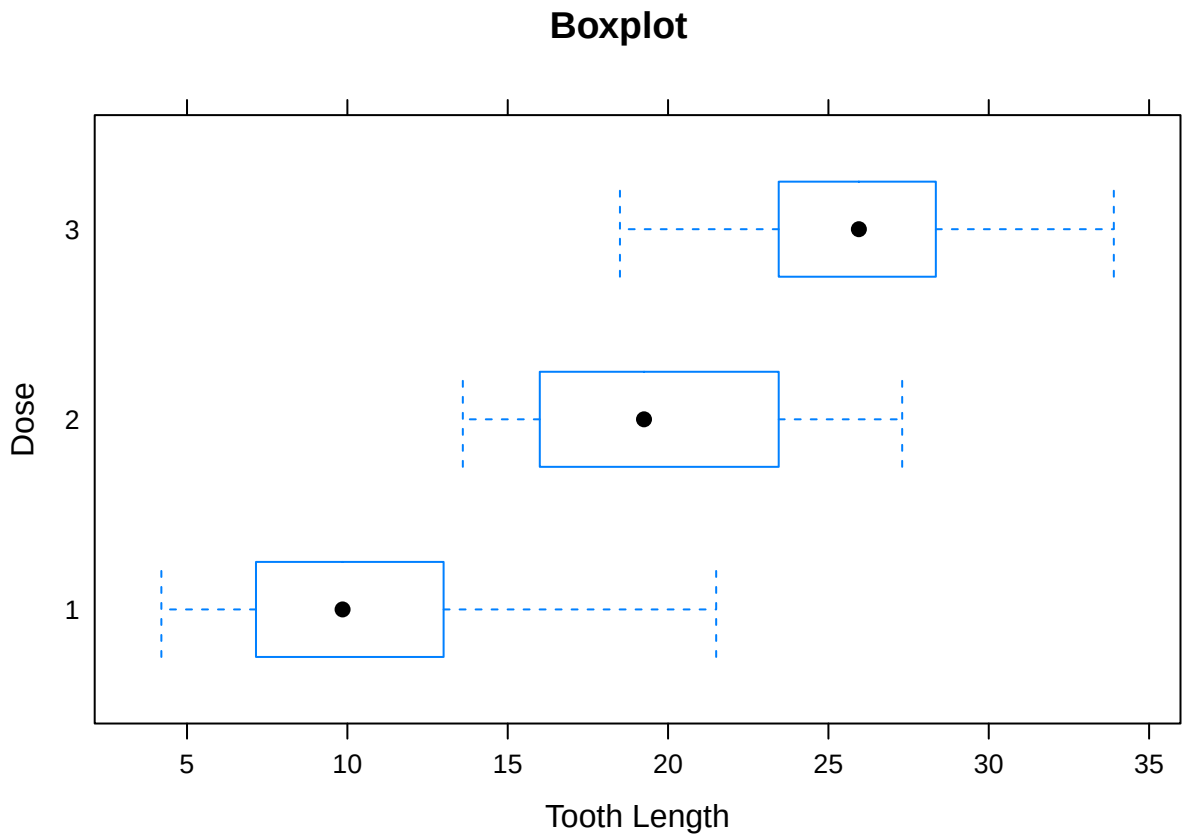
This file constitutes the second part of the Statistical Inference Coures Project

```
#Load ToothGrowth data and any relevant packages
data <- ToothGrowth
attach(data)
library(lattice)
library (latticeExtra)
```

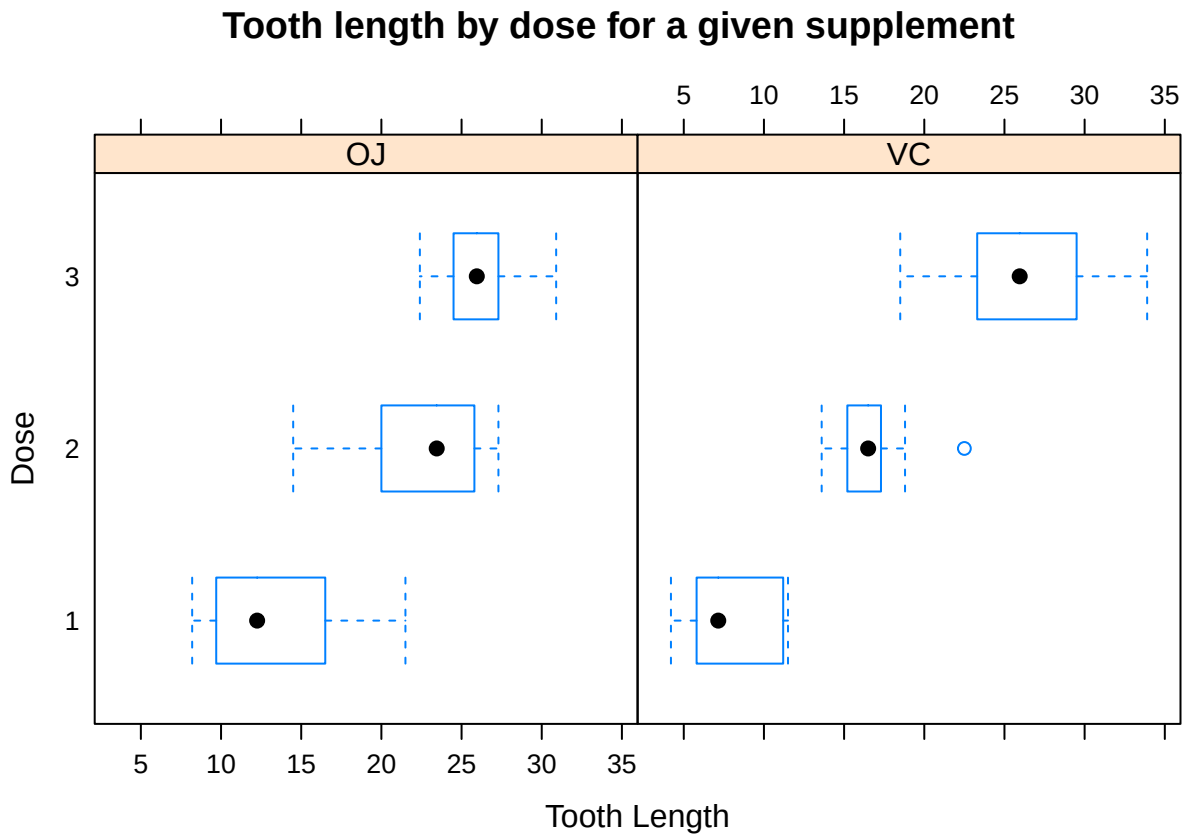
1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
## Loading required package: RColorBrewer
```

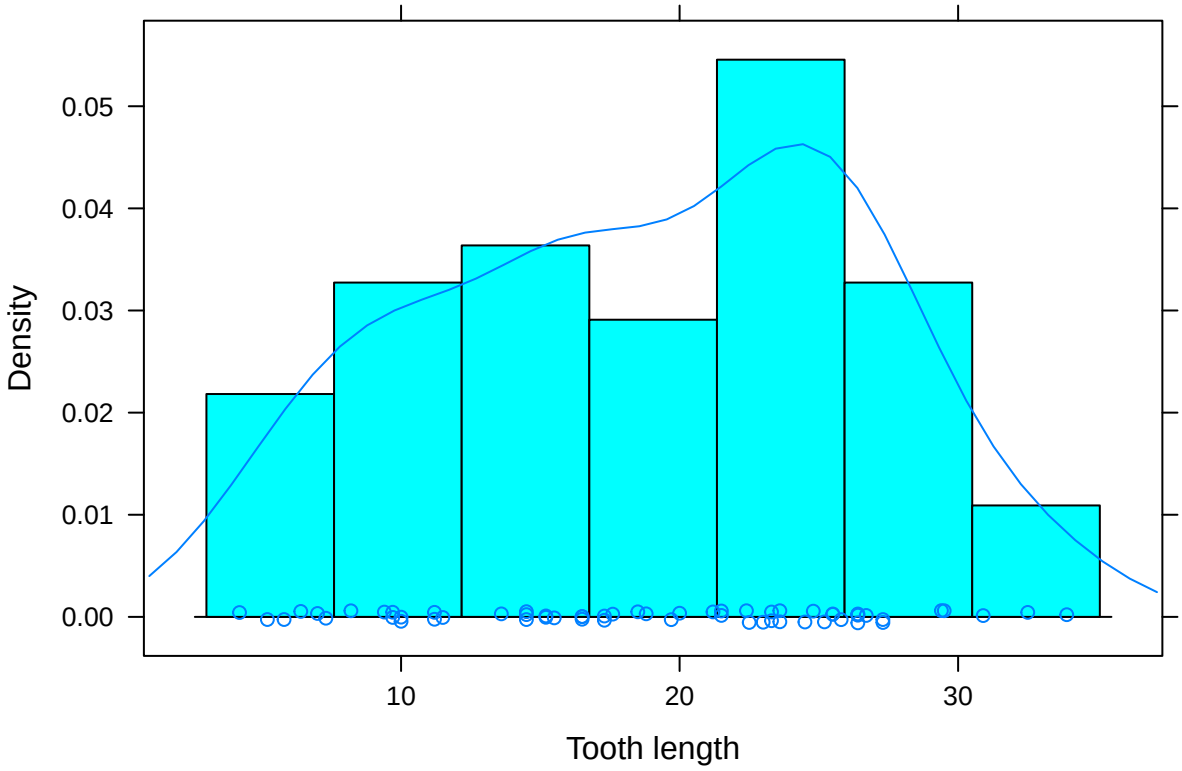
```
#Exploratory analysis
#Boxplots

bwplot(dose ~ len, xlab = "Tooth Length", ylab = "Dose",
       main = "Boxplot")
```



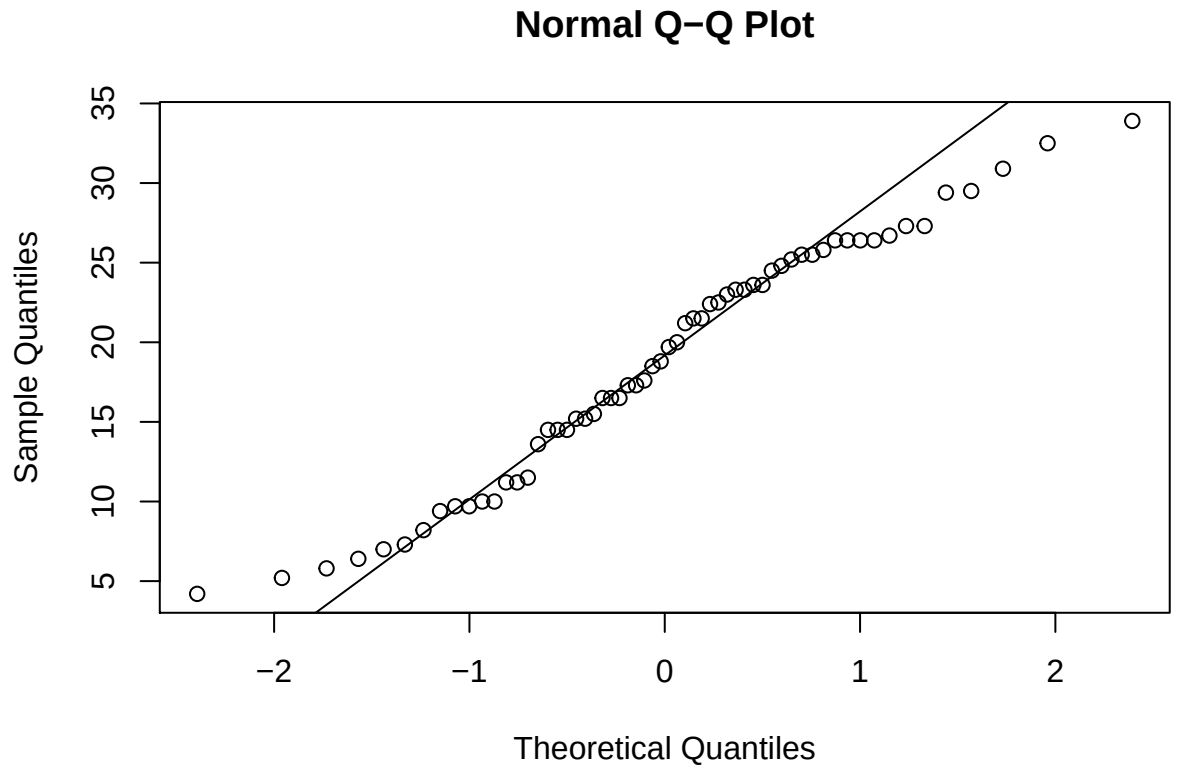
```
bwplot(dose ~ len | supp, xlab = "Tooth Length", ylab = "Dose",
       main = "Tooth length by dose for a given supplement")
```




```
#Histogram with density overlay
histogram(len, type = "density", xlab = "Tooth length") + densityplot(len)
```




```
qqnorm(len)
qqline(len, probs=c(0.25,0.75))
```



```r
#We will assume a normal distribution for the purposes of this analysis.
```

```r
#Data summary
#Quantiles
summary(data)
```

2. Provide a basic summary of the data.

```
##       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
```

```r
#Variance
datavarsupp <- c(var(len[supp == "OJ"]), var(len[supp == "VC"]))

datavardose <- c(var(len[dose == 0.5]), var(len[dose == 1.0]), var(len[dose == 2.0]))
datavarsupp
```

```
## [1] 43.63344 68.32723
```

```r
datavardose
```

```
## [1] 20.24787 19.49608 14.24421
```

For the analysis comparing differences by supplement, we will not assume equal variances.

For the analysis comparing differences between 0.5 doses and 1.0 doses, we will assume equal variances.

```r
#Hypothesis tests
#T-test comparing tooth lenth and supplement
tsupp <- t.test(len~supp, paired = FALSE, var.equal = FALSE, data = ToothGrowth)
tsupp
```

3. Use confidence intervals and hypothesis tests to compare tooth growth by supp and dose.

```
##
##  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
```

```
#T-test comparing tooth length under 0.5 and 1.0 doses
tdose.0.5.1 <- t.test(len[dose == 0.5], len[dose == 1.0], paired = FALSE,
tdose.0.5.1
```

```
##
##  Two Sample t-test
##
## data:  len[dose == 0.5] and len[dose == 1]
## t = -6.4766, df = 38, p-value = 1.266e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983748  -6.276252
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

```
#T-test comparing tooth length under 0.5 and 2.0 doses
tdose.0.5.2 <- t.test(len[dose == 0.5], len[dose == 2.0], paired = FALSE,
                      var.equal = FALSE)
tdose.0.5.2
```

```
##
##  Welch Two Sample t-test
##
## data:  len[dose == 0.5] and len[dose == 2]
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

```
#T-test comparing tooth length under 1.0 and 2.0 doses
tdose.1.2 <- t.test(len[dose == 1.0], len[dose == 2.0], paired = FALSE,
                      var.equal = FALSE)
tdose.1.2
```

```
##
##  Welch Two Sample t-test
##
## data:  len[dose == 1] and len[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
```

4. State your conclusions and the assumptions needed for your conclusions. In conclusion, we fail to reject the null hypothesis when comparing tooth length by the type of supplement.

We reject the null hypothesis when comparing tooth length by dosage levels in all cases.

Assumptions: The sample is normally distributed.

For the analysis comparing differences by supplement, I did not assume equal variances.

For the analysis comparing differences between 0.5 doses and 1.0 doses, I assumed equal variances.