Statistical Inference- Project Part 2

Jeff Sheremata

October 25, 2015

This is a statistical inference analysis of the Tooth Growth data set in R. The data in this study examines the 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). Specifically, I will examine if dosage has a statistical impact on tooth length, and if delivery method has a statistically signifficant impact on tooth length.

Load the Data

```
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 ...
```

```
summary(ToothGrowth)
```

```
##       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
```

```
unique(ToothGrowth$dose)
```

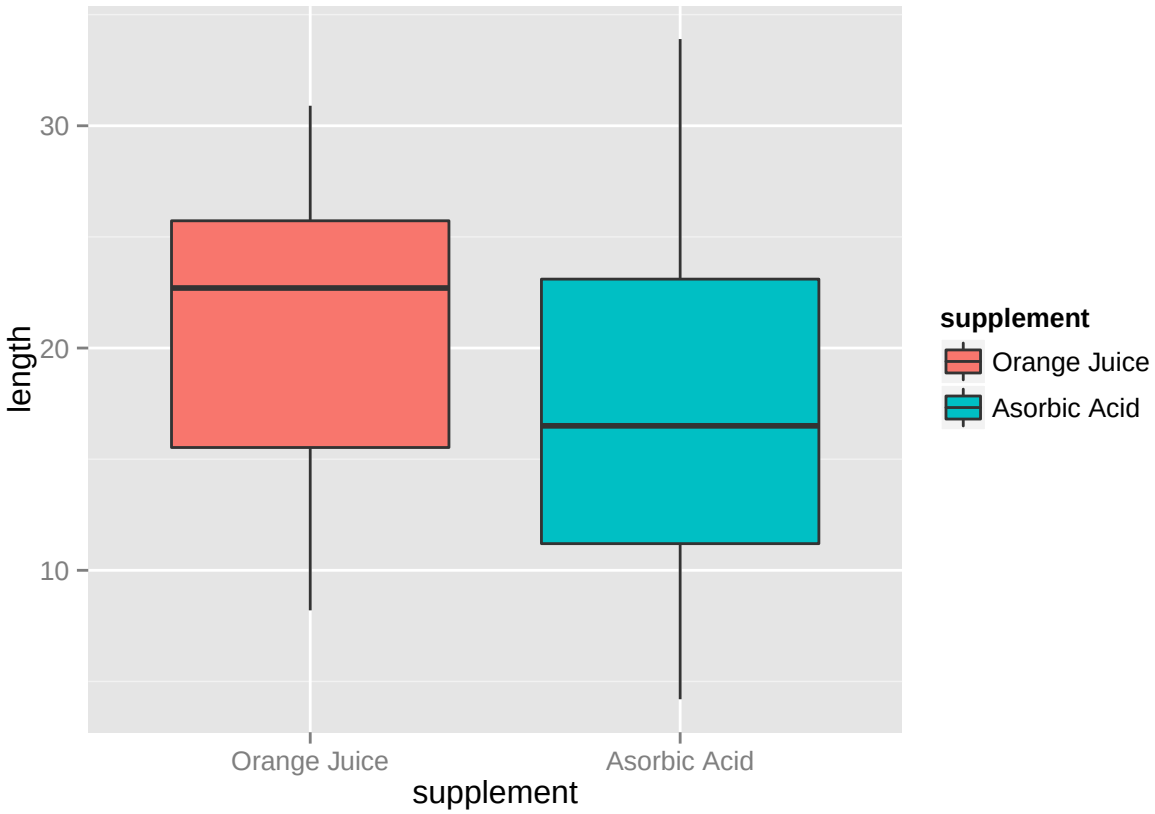
```
## [1] 0.5 1.0 2.0
```

The data frame contains 60 data points and has three variables: len (tooth length), supp (delivery type - either OJ-orange juice, or VC-asorbic acid), and dose - dosage - either 0.5, 1.0, or 2.0 mg.

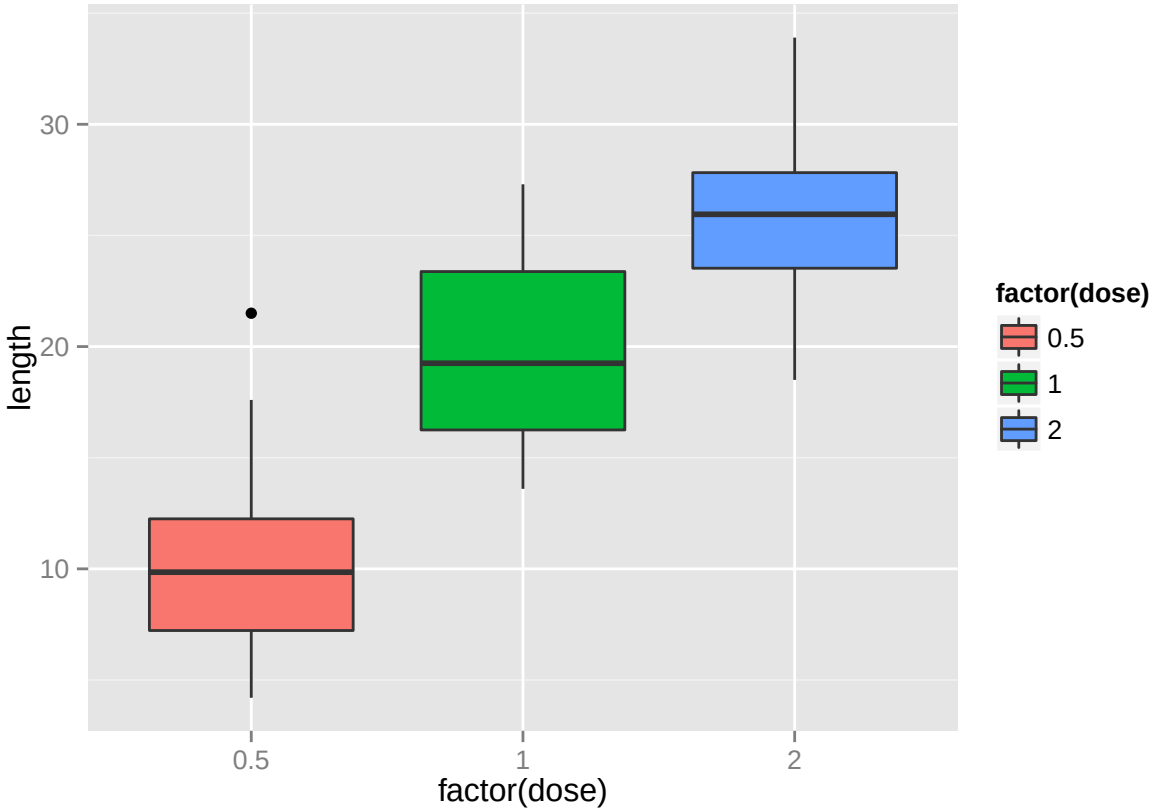
Let's change the name of supp to supplement, VC to Asorbic Acid, and OK to Orange Juice. Improved nomenclature leads to improved and easier analyses.

```
library(plyr)
ToothGrowth$supplement<-revalue(ToothGrowth$supp, c("VC"="Asorbic Acid", "OJ"="Orange Juice"))
ToothGrowth$length<-ToothGrowth$len
```

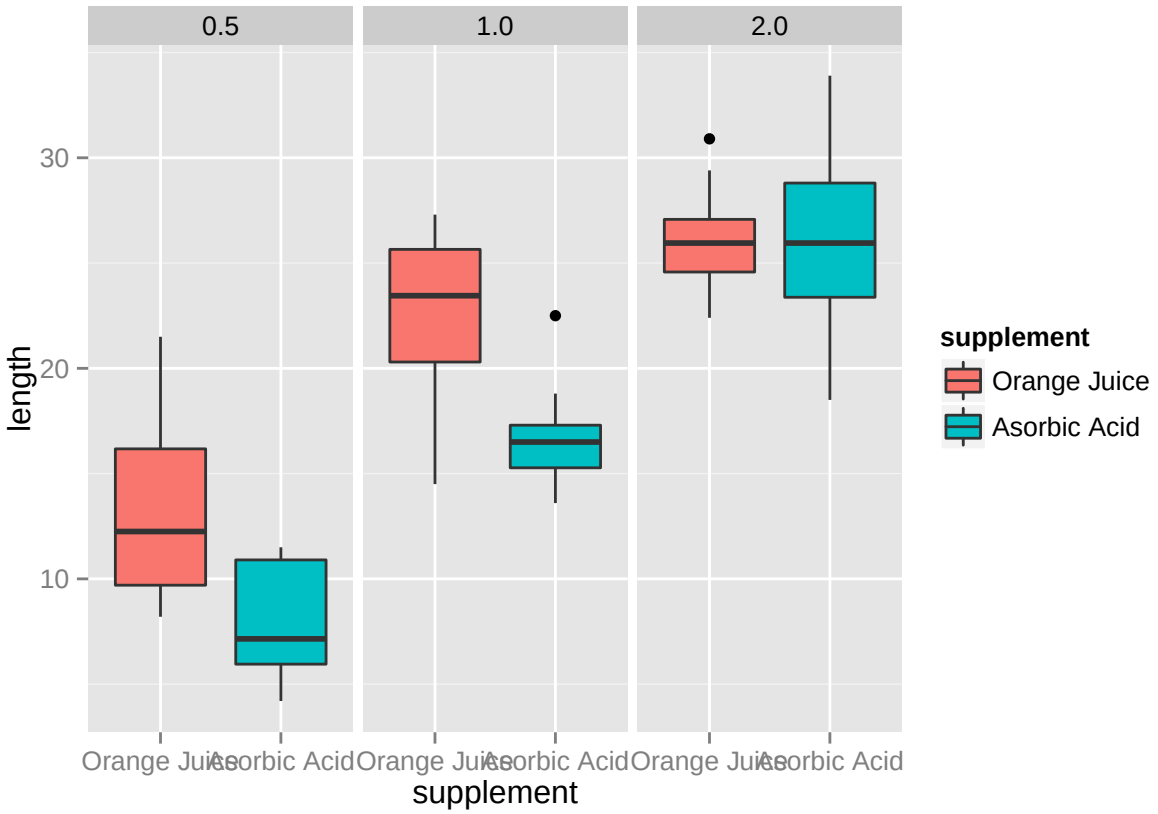
This plot seperates the delivery data into orange juice and asorbic acid classes.



This plot examines tooth length as a function of dose.



This plot presents the delivery method, orange juice and asorbic acid, data at each dose.



From visual inspection, it appears that for asorbic acid, as dosage increases, measured tooth length also increases. For orange juice, as dosage inceases from 0.5 mg to 1.0 mg, tooth length increases. However, it appears that for orange juice as dosage increases from 1.0 mg to 2.0 mg, tooth length does not signifficantly change.

These are just visual inspections of the data. In the next section, I will present the results of t-tests of the data.

Statistical Tests

Part 1-Impact of Delivery Method at Each Dose

In this part, I will subset the data for each dose, and then perform t-tests to examine if delivery method has a statistically signifficant impact on tooth length.

Let's first subset the data into Orange Juice and Asorbic Acid classes

```
OJ<-subset(ToothGrowth,supp=="OJ")
VC<-subset(ToothGrowth,supp=="VC")
```

Impact of Delivery Method at 0.5 mg Dosages

Null Hypthesis-Delivery method does not have an impact at a dosage of 0.5 mg.

```
OJP5=subset(OJ,dose==0.5)
VCP5=subset(VC,dose==0.5)
```

```
tresult<-t.test(OJP5$length,VCP5$length,paired = F, var.equal = F)
tresult
```

```
##
##  Welch Two Sample t-test
##
## data:  OJP5$length and VCP5$length
## 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 of x mean of y
##     13.23      7.98
```

The 95% confidence interval is 1.7190573, 8.7809427 . We can reject the null hypothesis, and conclude that at the 0.5 mg dosage, delivery by orange juice results in tooth lengths that are different than delivery by asorbic acid.

Impact of Delivery Method at 1.0 mg Dosages

Null Hypthesis-Delivery method does not have an impact at a dosage of 0.5 mg.

```
OJ1=subset(OJ,dose==1.0)
VC1=subset(VC,dose==1.0)

tresult<-t.test(OJ1$length,VC1$length,paired = F, var.equal = F)
tresult
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ1$length and VC1$length
## 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 of x mean of y
##     22.70     16.77
```

The 95% confidence interval is 2.8021482, 9.0578518 . We can reject the null hypothesis, and conclude that at the 1.0 mg dosage, delivery by orange juice results in tooth lengths that are different than delivery by asorbic acid.

Impact of Delivery Method at 2.0mg Dosages

Null Hypthesis-Delivery method does not have an impact at a dosage of 2.0 mg.

```
OJ2=subset(OJ,dose==2.0)
VC2=subset(VC,dose==2.0)

tresult<-t.test(OJ2$length,VC2$length,paired = F, var.equal = F)
tresult
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ2$length and VC2$length
## t = -0.0461, 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 of x mean of y
##     26.06     26.14
```

The 95% confidence interval is -3.7980705, 3.6380705 . We accept the null hypothesis, and conclude that at the 2.0 mg dosage, delivery by orange juice results in tooth lengths that are not different than delivery by asorbic acid.

Part 2-Impact of Delivery Method at Each Dose

In this part, t-tests will be constructed to examine the impact of dosage on tooth growth. Delivery results will be lumped, and t-tests between different doses will be constructed.

Increasing Dosage From 0.5mg to 1.0 mg

Null Hypthesis-Increasing dosage from 0.5 mg to 1.0 mg does not have an impact on tooth length

```
TGP5=subset(ToothGrowth,dose==0.5)
TG1=subset(ToothGrowth,dose==1.0)

tresult<-t.test(TGP5$length,TG1$length,paired = F, var.equal = F)
tresult
```

```
##
##  Welch Two Sample t-test
##
## data:  TGP5$length and TG1$length
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

The 95% confidence interval is -11.9837813, -6.2762187 . We can reject the null hypothesis, and conclude that at tooth lengths at the 1.0 mg and 0.5 mg dosages are signifficantly different.

Increasing Dosage From 1.0 mg to 2.0 mg

Null Hypthesis-Increasing dosage from 1.0 mg to 2.0 mg does not have an impact on tooth length

```
TG1=subset(ToothGrowth,dose==1.0)
TG2=subset(ToothGrowth,dose==2.0)

tresult<-t.test(TG1$length,TG2$length,paired = F, var.equal = F)
tresult
```

```
##
##  Welch Two Sample t-test
##
## data:  TG1$length and TG2$length
## 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
```

The 95% confidence interval is -8.9964805, -3.7335195 . We can reject the null hypothesis, and conclude that at tooth lengths at the 1.0 mg and 2.0 mg dosages are signifficantly different.

Assumptions

The following assumptions were made in this statistical analysis:

1. The variances in the populations were different
2. The populations were independent
3. Samples were random
4. All samples were similar
5. Each population follows a normal distribution
6. The data are continuous
7. All subjects were cared for under the same conditions

Conclusions

Based on the statistical analyses in this study, it can be concluded that at 0.5 mg and 1.0 mg Vitamin C dosages, orange juice delivery results in tooth lengths that are signifficantly larger than by asorbic acid delivery. However, at 2.0 Vitamin C mg dosages, there was not statistical evidence that tooth lengths are impacted by orange juice delivery any differently compared to asorbic acid delivery.

The t-test results clearly indicate that in this study, as Vitamin dosage increased from 0.5-1.0 mg, and from 1.0-2.0 mg, tooth lengths also increased. Thus, the data imply that dosage and tooth length are correlated in the dosage range used in this study.