ToothGrowth.Rmd

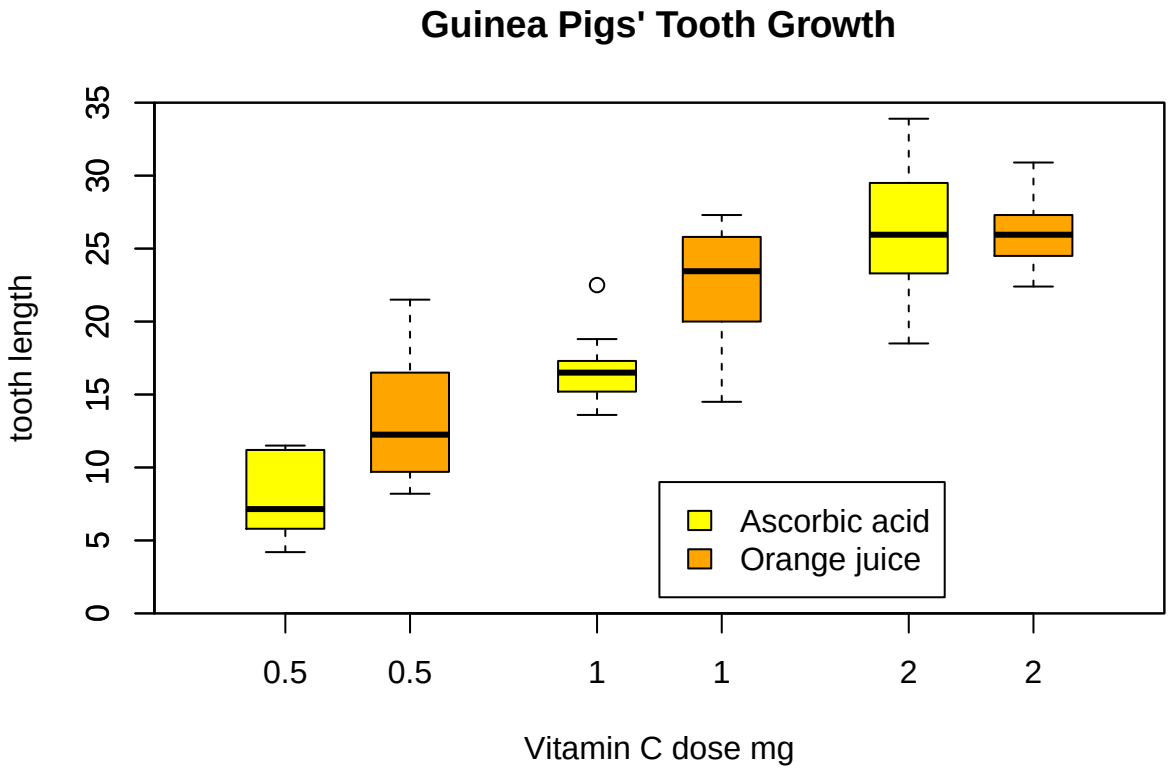
Jonas Wiorek

Sunday, April 26, 2015

This report summarizes the analyses of the ToothGrowth data. The first step in the analyses is to make some basic exploratry data analyses.

```
##       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
```

The data explores the effect of Vitamin C on tooth growth in Guinea Pigs. 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).



The boxplot above indicates that orange juice may be the most promising delivery method regarding tooth growth at least for small dose levels.

Next step is to compare tooth growth by delivery methods and dose levels by means of confidence intervals and hypothesis tests.

We split the data in six groups of delivery methods and dose levels. That is three groups with orange juice delivery method and dose levels of 0.5, 1 respectively 2 mg of Vitamin C, and three groups with ascorbic acid delivery method and the three dose levels.

First we compare groups with the same dose level and different delivery methods. The null hypothesis says that there is no difference in the mean tooth growth for the two delivery methods given the same dose. We set the probablity of Type I error, alpha, to 0.05, i.e. the probability to reject the null hyothesis when, in fact, it is correct.

We compare independent groups under unequal variances with confidence t intervals. When in doubt, assume unequal variances. Under enequal variances the 95% interval is approximated with yhat - xhat +/-tdf*sqrt(Sx^2/nx + Sy^2/ny) where tdf is calculated with degress of freedom df = (Sx^2 /nx + Sy^2 /ny)^2 / ((Sx^2 /nx)^2/(nx-1) + (Sy^2 /ny)^2/(ny-1))

First we compare the two delivery methods for a 0.5 mg dose level of Vitamin C. We calculate the 95% confidence interval for the null hypothesis according to the expression above for two groups under unequal variances and also by using the t.test function in R.

```
oj05 <- ToothGrowth[ToothGrowth$supp == 'OJ' & ToothGrowth$dose == 0.5,]
aa05 <- ToothGrowth[ToothGrowth$supp == 'VC' & ToothGrowth$dose == 0.5,]

noj05 <- length(oj05$len)
muoj05 <- mean(oj05$len)
sdoj05 <- sd(oj05$len)
naa05 <- length(aa05$len)
muaa05 <- mean(aa05$len)
sdaa05 <- sd(aa05$len)

se05 <- sqrt(sdoj05^2/noj05 + sdaa05^2/naa05)
dfn05 <- (sdoj05^2/noj05 + sdaa05^2/naa05)^2
dfd05 <- (sdoj05^2/noj05)^2/(noj05-1) + (sdaa05^2/naa05)^2/(naa05-1)
df05 <- dfn05/dfd05
muoj05 - muaa05 + c(-1,1)*qt(0.975,df05)*se05
```

```
## [1] 1.719057 8.780943
```

```
t.test(oj05$len,aa05$len, paired=FALSE, var.equal = FALSE)$conf
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

0 (no difference between the means) is outside the 95% confidence interval for 0.5 mg and therefore we reject the null hypothesis.

We do the same thing for 1 and 2 mg dose levels of Vitamin C.

```
muoj10 - muaa10 + c(-1,1)*qt(0.975,df10)*se10
```

```
## [1] 2.802148 9.057852
```

```
t.test(oj10$len,aa10$len, paired=FALSE, var.equal = FALSE)$conf
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

```
muoj20 - muaa20 + c(-1,1)*qt(0.975,df20)*se20
```

```
## [1] -3.79807  3.63807
```

```
t.test(oj20$len,aa20$len, paired=FALSE, var.equal = FALSE)$conf
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

0 is outside the 95% confidence interval also for 1 mg and therefore we reject the null hypothesis. However, for 2 mg dose level of Vitamin C 0 is inside the confidence interval. Thus, we fail to reject the null hypothesis for 2 mg.

We conclude that, for small doses of Vitamin C, the delivery method matters. Orange juice is more effective on tooth growth than ascorbic acid. For large doses, however, we couldn't conclude with confidence that the delivery method plays a difference. A theory is that the delivery method matters as long as the dose is low, but above a certain level the Guinea pigs become saturated with Vitamin C regardless of delivery method and the tooth length does no longer increase with increased dose levels.

```
## [1] 0.1885575 6.5314425
## attr(,"conf.level")
## [1] 0.95
```

```
## [1]  5.524366 13.415634
## attr(,"conf.level")
## [1] 0.95
```

```
## [1]  5.685733 13.054267
## attr(,"conf.level")
## [1] 0.95
```

```
## [1]  6.314288 11.265712
## attr(,"conf.level")
## [1] 0.95
```

Finnaly, we also draw some conclusions by testing independent groups under unequal variances with confidence t intervals of the same delivery methods but different dose levels. For both delivery methods increasing from 0.5 to 1 mg and increasing from 1 to 2 mg 0 is outside the confidence interval. Therefore, we reject the null hypothesis. Thus, for the same delivery method, the conclusion is that the dose level matters.