Guinea Pig Tooth Growth Experiment

Juan Bosco Mendoza Vega

September 18 2016

Summary

We'll conduct a mean comparison of the effect of dose and supply method on theeth growth of guinea pigs. Means will be compared using t test 95% confidence interval and a significance level of 0.05.

Exploratory statistics

Structure of the **ToothGrowth** data set.

```
'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 ...
```

There are three variables and 60 observations. According to the documentarion called by `?ToothGrowth`, our variables are **len** (Tooth length, numeric); **supp** (Supplement type, factor with two levels. OJ: Orange Juice, VC: Ascorbic acid); and **dose** (Dose in miligrams per day, numeric).

There appear to be differences in Tooth growth by Supplement and Dose (Appendix A), but we need to test if it is statistically significant.

Comparison of means

By Supplement type

Orange juice vs Ascorbic acid

```
    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
```

By Dose in miligrams per day

0.5 miligram per day vs 1 miligram per day

```
    Welch Two Sample t-test

data:  dose_0.5 and dose_1
t = -6.4766, df = 37.986, p-value = 6.342e-08
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -6.753323
sample estimates:
mean of x mean of y
   10.605    19.735
```

0.5 miligram per day vs 2 miligram per day

```
    Two Sample t-test

data:  dose_0.5 and dose_2
t = -11.799, df = 38, p-value = 1.419e-14
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -13.28093
sample estimates:
mean of x mean of y
   10.605    26.100
```

1 miligram per day vs 2 miligram per day

```
    Welch Two Sample t-test

data:  dose_1 and dose_2
t = -4.9005, df = 37.101, p-value = 9.532e-06
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf -4.17387
sample estimates:
mean of x mean of y
   19.735    26.100
```

Conclusion

There's a diference in tooth growth by

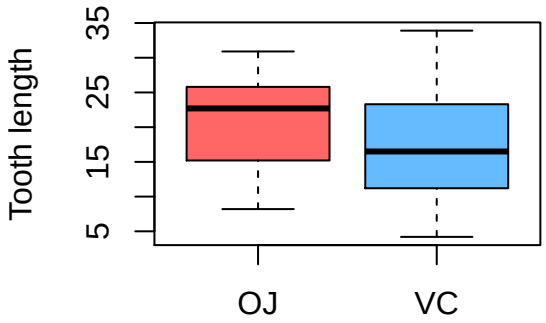
There's no difference by supp.

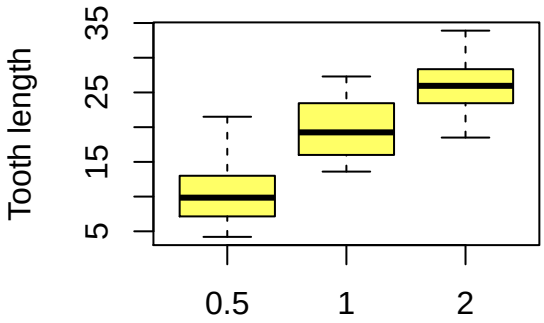
There's a difference by dose. Lower doses correspond to lower growth.

No difference by supp at a dose equal to 2.

Appendix

A. Visualization of the groups





B. Summary of 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
```

C. Test of normality of Tooth length

The ¨*Shapiro-Wilk Normality Test* is used, The null hypothesis for this test is that the given data is normally distributed. Results indicate we can't reject this hypotesis, so **we assume the distribution is normal**.

```
shapiro.test(ToothGrowth$len)
```

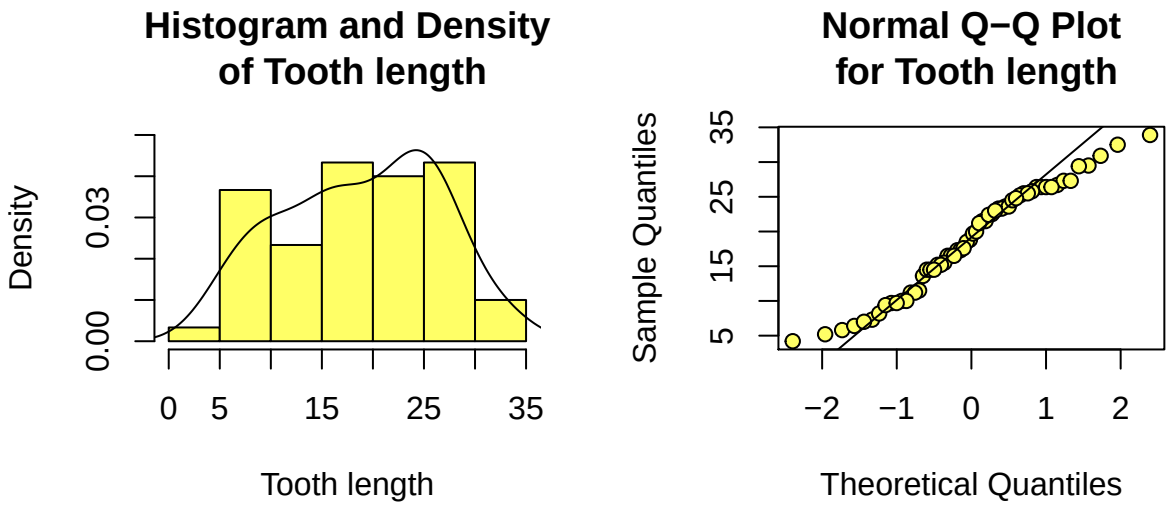
```
    Shapiro-Wilk normality test

data:  ToothGrowth$len
W = 0.96743, p-value = 0.1091
```

A histogram with an overlaid density curve of this variable and a Normal Q-Q plot are generated to support the assumtion its distribution is normal.

```r
par(mfrow = c(1, 2))
hist(ToothGrowth$len, probability = T, ylim = c(0, 0.05),
    main = "Histogram and Density\n of Tooth length",
    xlab = "Tooth length",
    col = "#FFFF66"
    )
lines(density(ToothGrowth$len))

qqnorm(y = ToothGrowth$len,
    main = "Normal Q-Q Plot\n for Tooth length",
    pch = 21, bg = c("#FFFF66"))
qqline(y = ToothGrowth$len)
```



D. Test if variance is equal across groups

The *Levene's test of homogenity of variance is used.* Groups are: *Supplement type*; *Dose in miligrams per day*; and *Supplement type* and *Dose in miligrams per day.* Results indicate groups **don't have equal variance**.

```r
library(car)
leveneTest(len ~ supp, data = ToothGrowth)
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  1.2136 0.2752
      58
```

```r
leveneTest(len ~ as.factor(dose), data = ToothGrowth)
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  2  0.6457 0.5281
      57
```

```r
leveneTest(len ~ supp * as.factor(dose), data = ToothGrowth)
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  5  1.7086 0.1484
      54
```