

Statistical Inference - The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Jonathan Hill, MPA

Wednesday, April 22, 2015

Load necessary R packages and the ToothGrowth dataset

```
# Loading necessary packages #
library(datasets)
library(lattice)

# Load in the data #
data(ToothGrowth)
tooth.dat <- ToothGrowth

# Change dose to a factor variable #
tooth.dat$dose <- as.factor(tooth.dat$dose)
```

Summary of the Data

```
##      len      supp      dose
##  Min.   : 4.20    OJ:30    0.5:20
## 1st Qu.:13.07    VC:30     1 :20
##  Median :19.25           2 :20
##   Mean  :18.81
## 3rd Qu.:25.27
##   Max.  :33.90
```

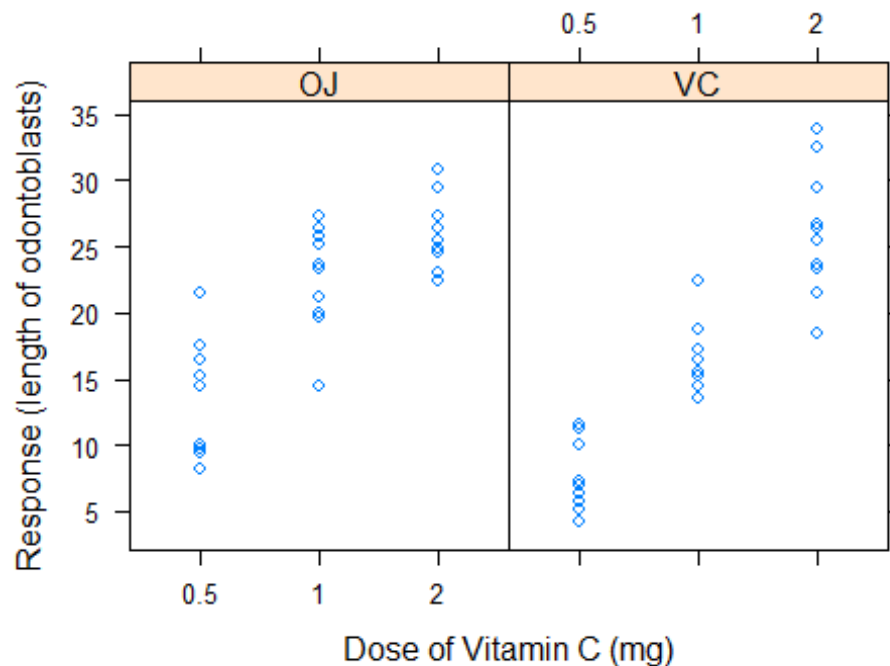
len = numeric variable of tooth length

supp = factor variable of supplement type (ascorbic acid and orange juice)

dose = numeric variable of dosage in milligrams

Exploratory plot of the data

Vitamin C Delivery: Orange Juice (OJ) Vs. Ascorbic Acid



Orange juice seems to amplify tooth growth at lower dosages (0.5 & 1.0 mg) of Vitamin C, and Vitamin C seems to result in more tooth growth regardless of the delivery method.

T-tests: Orange Juice versus Ascorbic Acid by Dose

First, the null and alternative hypotheses for the t-tests are:

Ho: Tooth growth is the same for both delivery methods at the given dosage of Vitamin C

Ha: Tooth growth is probably impacted by the delivery method at the given dosage of Vitamin C

Vitamin C Dose = 0.5 mg

```
tooth.dat.5 <- tooth.dat[tooth.dat$dose==0.5,]
t.test(len~supp,data=tooth.dat.5,paired = FALSE, var.equal = FALSE)

##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##          13.23          7.98
```

We must reject the null at the 95% significance level because the p-value is less than .05. There is evidence that orange juice significantly amplifies guinea pig tooth growth for a 0.5 mg dosage.

Vitamin C Dose = 1.0 mg

```
tooth.dat1 <- tooth.dat[tooth.dat$dose==1.0,]
t.test(len~supp,data=tooth.dat1,paired = FALSE, var.equal = FALSE)

##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##          22.70          16.77
```

We must reject the null at the 95% significance level because the p-value is less than .05. There is evidence that orange juice significantly amplifies guinea pig tooth growth for a 1.0 mg dosage.

Vitamin C Dose = 2.0 mg

```
tooth.dat2 <- tooth.dat[tooth.dat$dose==2.0,]
t.test(len~supp,data=tooth.dat2,paired = FALSE, var.equal = FALSE)

##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##          26.06          26.14
```

We cannot reject the null at the 95% significance level because the p-value is not less than .05. There is no evidence that the delivery method significantly amplifies guinea pig tooth growth for a 2.0 mg dosage.

Conclusion:

The delivery method (orange juice) significantly impacts the growth of guinea pig teeth at lower dosages of Vitamin C (.5 & 1.0 mg), but there is no evidence that it does so at higher dosage levels (2.0 mg).

T-test Assumptions

1. Adequate power

This dataset only has 60 observations and each subset includes only 20 observations. The power of these tests is low, and any results need to be reproduced before they could be considered solid findings.

2. The data are produced from a random sample

There is no mention of the data being produced using proper sampling methodology.

3. The population is normally distributed

```
##  
## Shapiro-Wilk normality test  
##  
## data:  tooth.dat$len  
## W = 0.9674, p-value = 0.1091
```

A Shapiro Wilk Normality test reveals that this assumption has been violated because the p-value is greater than .05.

4. Independent Samples

There is no language indicating that these are independent samples with equal variance; therefore, the options "paired = FALSE" and "equal.var = FALSE" were specified in each of the tests above.