

Statistical Inference on Tooth Length in Guinea Pigs

A study examined the length of teeth of 10 guinea pigs administered Vitamin C at three dose levels (0.5, 1, and 2 mg) with each delivered via orange juice (OJ) or ascorbic acid (VC).

First, load the data and explore it a little bit.

```
library(datasets)
tooth <- ToothGrowth
dim(tooth)

## [1] 60  3

head(tooth)

##      len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5

str(tooth)

## '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 ...

#pairs(tooth)
boxplot(tooth$len ~ tooth$dose, main = "Length of Teeth by Dose", xlab = "Dose (mg)", ylab = "Tooth Length")

boxplot(tooth$len ~ tooth$supp, main = "Length of Teeth by Delivery", xlab = "Delivery Method", ylab = "Tooth Length")
```



From the boxplots, it appears there are some correlations between tooth length and dose and the delivery method, but a hypothesis test using a t-test should be run to see if these are statistically significant.

```
OJ_0.5 <- tooth[tooth$supp == "OJ" & tooth$dose == 0.5,]
OJ_1 <- tooth[tooth$supp == "OJ" & tooth$dose == 1,]
OJ_2 <- tooth[tooth$supp == "OJ" & tooth$dose == 2,]
VC_0.5 <- tooth[tooth$supp == "VC" & tooth$dose == 0.5,]
VC_1 <- tooth[tooth$supp == "VC" & tooth$dose == 1,]
VC_2 <- tooth[tooth$supp == "VC" & tooth$dose == 2,]
```

Delivery method

First, the delivery method is examined at each dose level. A two sample t-test is run. The null hypothesis (H_0) is that the delivery method does not have an effect on the mean tooth length (the means are equal). The alternative hypothesis (H_a) is that the delivery method increases or decreases tooth length (the means are not equal).

```
t supp1 <- t.test(OJ_0.5$len, VC_0.5$len, paired = F, var.equal = F)
t supp2 <- t.test(OJ_1$len, VC_1$len, paired = F, var.equal = F)
t supp3 <- t.test(OJ_2$len, VC_2$len, paired = F, var.equal = F)
```

The t-value between a OJ and VC at a 0.5 mg dose is 3.17. The t-value between OJ and VC at 1 mg is 4.03. The t-value between a OJ and VC at 2 mg is 8.03. For the two lower doses, the difference in mean tooth length is statistically significant, but not for the highest dose of 2 mg.

The confidence intervals are [1.72, 8.78], [2.8, 9.06], and [6.86, 11.72]. The two lower dose confidence intervals do not contain 0.

Since our null hypothesis is that all three means are equal, we fail to reject the null hypothesis since at the highest dose, there is no statistically significant difference between OJ and VC as a delivery method. However, if the null hypothesis is revised to separate the three dose levels, the delivery method would be statistically significant for dose levels of 0.5 mg and 1 mg, but not for 2 mg.

Dose level

The dose level is examined for OJ and VC separately using a two sample t-test. The null hypothesis (H_0) is that the mean tooth length for the three doses are equal. The alternative hypothesis (H_a) is that the mean tooth length for the three doses are not equal.

```
tdoseOJ1 <- t.test(OJ_0.5$len, OJ_1$len, paired = F, var.equal = F)
tdoseOJ2 <- t.test(OJ_0.5$len, OJ_2$len, paired = F, var.equal = F)
tdoseOJ3 <- t.test(OJ_1$len, OJ_2$len, paired = F, var.equal = F)
tdoseVC1 <- t.test(VC_0.5$len, VC_1$len, paired = F, var.equal = F)
tdoseVC2 <- t.test(VC_0.5$len, VC_2$len, paired = F, var.equal = F)
tdoseVC3 <- t.test(VC_1$len, VC_2$len, paired = F, var.equal = F)
```

The t-values between the 0.5 and 1 mg, 0.5 and 2 mg, and 1 and 2 mg for OJ and VC are: -5.05, -7.82, -2.25, -7.46, -10.39, -5.47. Each of the confidence intervals do not contain 0, so we reject the null hypothesis. Therefore, the dose levels are statistically significant and have an effect on the tooth length.