Statistical Inference Project, Part 2

S. Smolenski

Introduction

This project explores the ToothGrowth R dataset, which describes: "The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice or ascorbic acid (a form of vitamin C and coded as VC)."

Exploratory Analysis

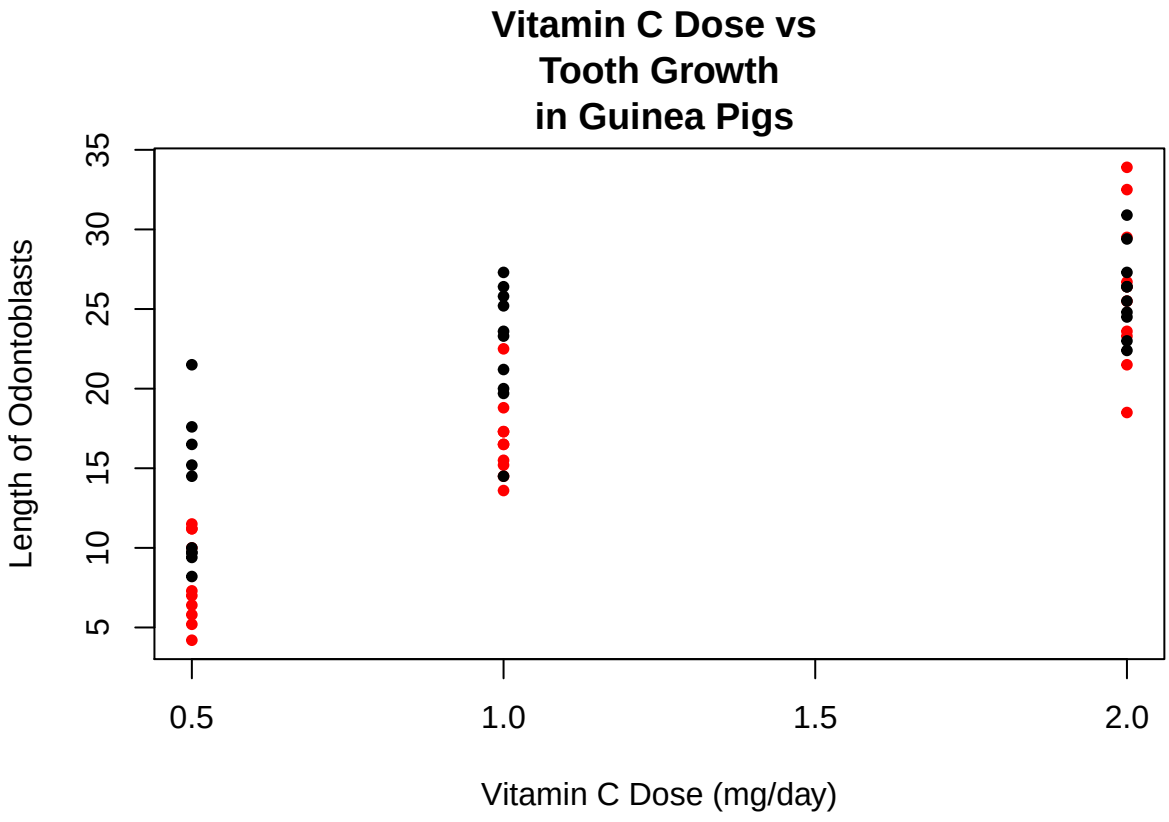
Loading the dataset:

```
data(ToothGrowth)
str(ToothGrowth)
```

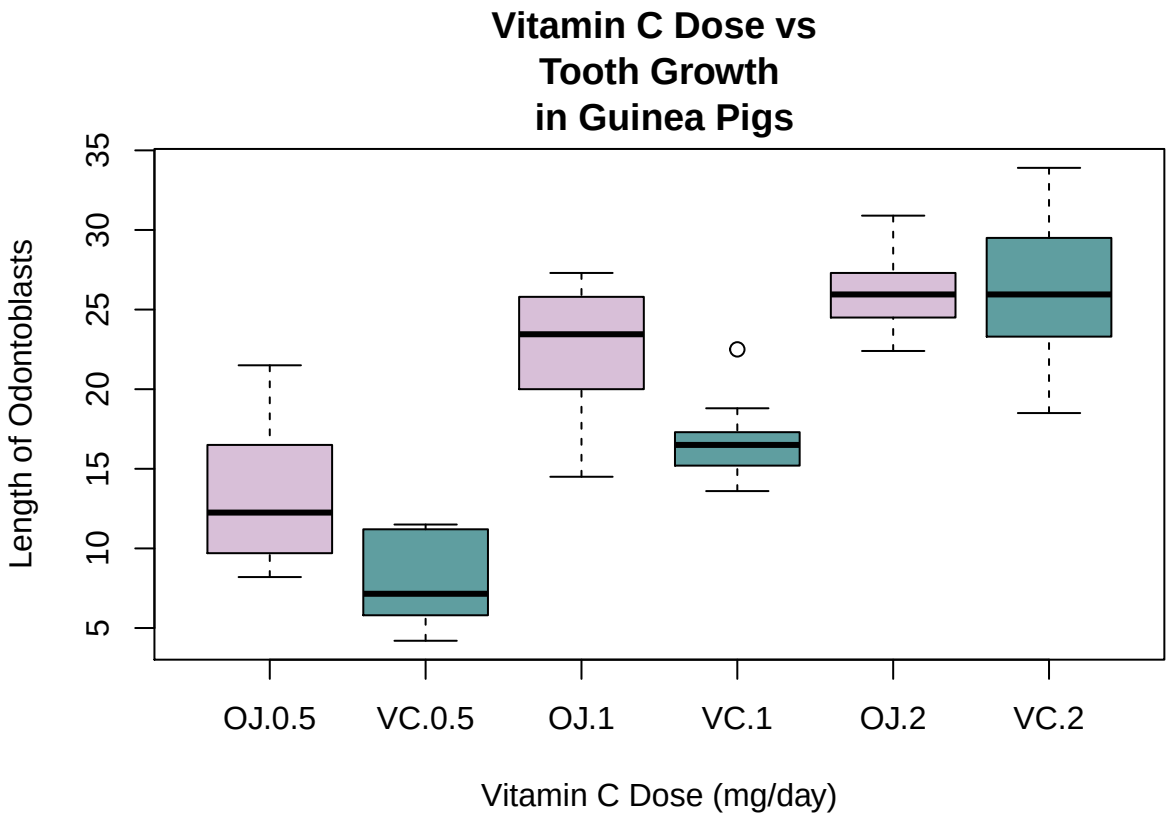
```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Plotting Vitamin C dose versus growth length. Colors are used to separate delivery methods.

```
with(ToothGrowth, plot(dose,len,col=supp, pch=20, main="Vitamin C Dose vs \n Tooth Growth \n in Guinea
```



```r
with(ToothGrowth, boxplot(len~supp*dose,col=c("thistle","cadetblue"), main="Vitamin C Dose vs \n Tooth G
```



Summary

The data seem to imply that tooth growth is correlated with Vitamin C dose, though the relationship between delivery method is less clear. For .5 and 1.0 mg/day, the OJ supplement wins out over the VC supplement, but at 2.0 mg/day, they seem roughly equivalent.

Confidence Interval Testing

Let us more closely examine the relationship between supplement type and tooth growth, since this is the more ambiguous relationship.

We perform a two-sided t test comparing the growth produced by OJ and VC, under the null hypothesis that there is no distinction between the two supplement types.

```r
OJ=ToothGrowth$len[ToothGrowth$supp=="OJ"]
VC=ToothGrowth$len[ToothGrowth$supp=="VC"]
t.test(OJ, VC, conf.level=.975)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ and VC
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 97.5 percent confidence interval:
##  -0.7510928  8.1510928
## sample estimates:
## mean of x mean of y
##  20.66333  16.96333
```

Conclusions

Under a two-sided, 95% confidence level, we obtain a P-value of .06. In other words, we do not reject the null hypothesis. We must conclude from this test that there is no significant difference between the two supplement types.