

Analyzing the R data set ToothGrow

David C Latshaw II

Overview

The ToothGrow data set in R contains the response data of 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). In the analysis we will:

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

Simulations

First we load the data set, plotting tools, and the package Rmisc to make some of the statistical analysis more compact and easier to handle.

```
library(datasets)
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.1
```

```
library(Rmisc)
```

```
## Warning: package 'Rmisc' was built under R version 3.2.2
```

```
## Loading required package: lattice
```

```
## Warning: package 'lattice' was built under R version 3.2.1
```

```
## Loading required package: plyr
```

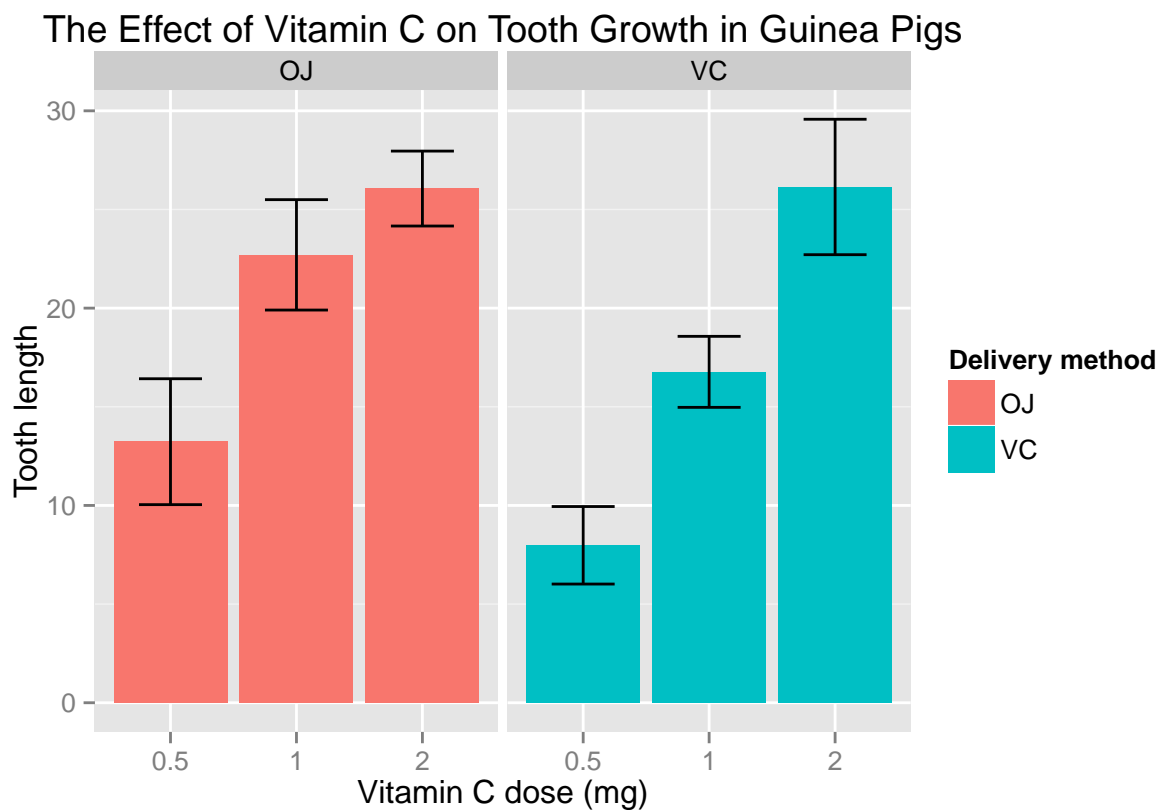
```
## Warning: package 'plyr' was built under R version 3.2.1
```

Next we create a statistical summary of the data using summarySE from the Rmisc package. The next piece is making the variable “dose” a factor so it can be plotted properly in ggplot. Finally we make a dual bar chart of the mean for each observation and include error bars for the 95% confidence intervals.

```

tgsummary <- summarySE(ToothGrowth, measurevar="len", groupvars=c("supp", "dose"))
tgsummary2 <- tgsummary
tgsummary2$dose <- factor(tgsummary2$dose)
ggplot(data=tgsummary2, aes(x=as.factor(dose), y=len, fill=supp)) +
  geom_bar(stat="identity",) +
  facet_grid(. ~ supp) +
  xlab("Vitamin C dose (mg)") +
  ylab("Tooth length") +
  guides(fill=guide_legend(title="Delivery method")) +
  ggtitle("The Effect of Vitamin C on Tooth Growth in Guinea Pigs") +
  geom_errorbar(aes(ymin=len-ci, ymax=len+ci),
    width=.5,
    position=position_dodge(.9))

```



Assumptions - The population is normally distributed and the real population variance is the same for different supp and dose.

Constant Delivery Method

Orange Juice At first glance, increasing the vitamin C dose increases tooth length. this is certainly the case comparing 0.5mg and 1mg doses, but it is relatively unclear if the difference in growth between 1mg and 2mg doses is significant because the 95% confidence intervals overlap. To determine if the difference is significant we calculate the p-value for these observations:

```
data <- subset(ToothGrowth, supp == 'OJ' & dose %in% c(1, 2))
res <- t.test(len ~ dose, data=data)
res$p.value
```

```
## [1] 0.03919514
```

The p-value for these observations is found to be 0.039. A p-value < 0.05 indicates strong evidence against the null hypothesis. In this case it is close but still below the threshold so we can assume this difference is significant. Using 2mg doses with orange juice as a delivery method is the most effective at promoting guinea pig tooth growth.

Ascorbic Acid Similar to the orange juice, increasing the Vitamin C dose increases tooth length. In this case none of the 95% confidence intervals overlap so each increase is statistically significant. Using 2mg doses with ascorbic acid as a delivery method is the most effective at promoting guinea pig tooth growth.

Constant Dose

0.5mg and 1mg Comparing 0.5mg doses for each delivery method and 1mg doses for each delivery method we can see that both doses do not have overlapping 95% confidence intervals. This tells us that at 0.5mg and 1mg doses orange juice is a more effective delivery method.

2mg 2mg doses for orange juice and ascorbic acid have overlapping 95% confidence intervals. Again we will look at the p-value to determine if the difference is significant.

```
data <- subset(ToothGrowth, dose == 2)
res <- t.test(len ~ supp, data=data)
res$p.value
```

```
## [1] 0.9638516
```

The p-value is 0.96 which is far above the threshold of 0.05 indicating weak evidence against the null hypothesis, so the difference is not statistically significant. Thus there is no apparent difference between the delivery methods at 2mg dosing.

Conclusions

In order to maximize tooth growth in guinea pigs, if that is the goal, dosing 2mg of vitamin C in either orange juice or ascorbic acid will be most effective. If a lower dose of 0.5mg or 1mg needs to be used then the most effective delivery method is orange juice.