Odontoblasts Growth Exploratory Analysis

Shaun

March 5, 2017

Overview

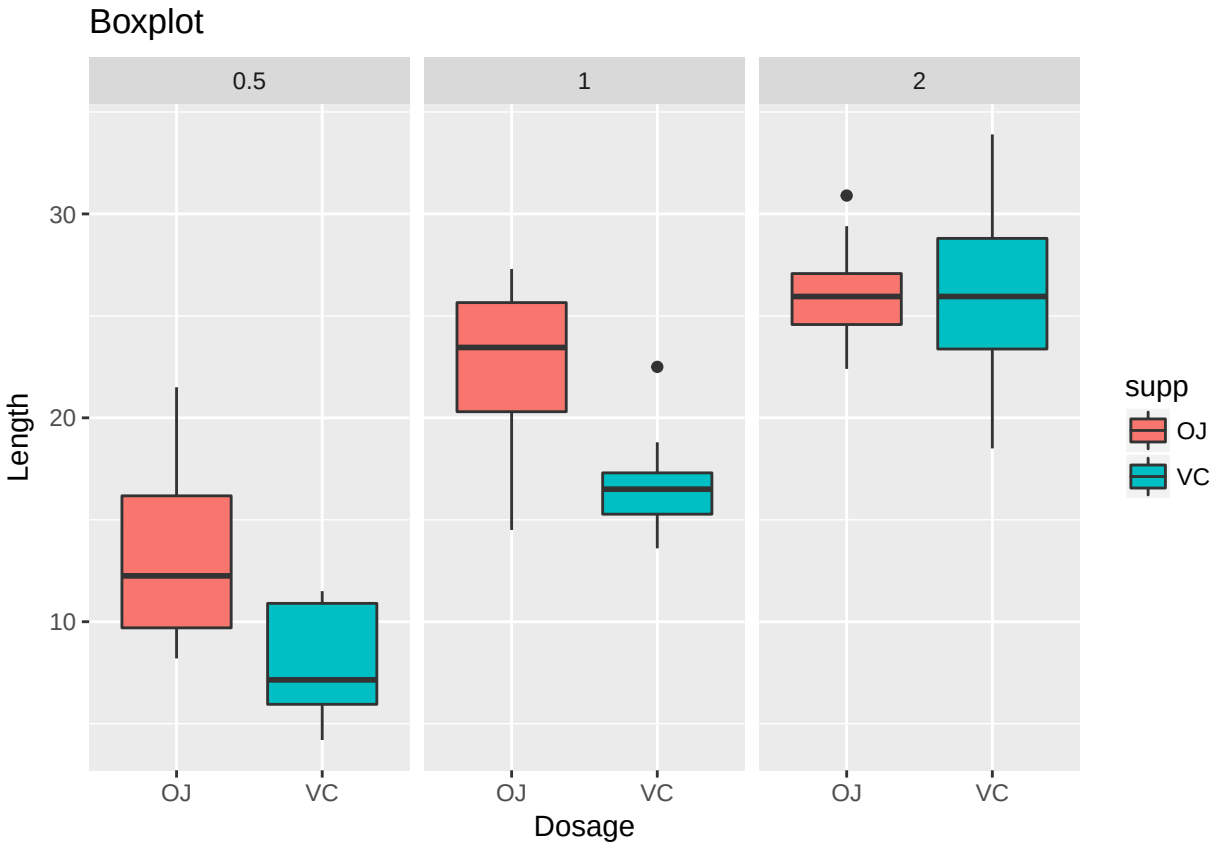
For Part 2 of this assignment, we'll perform an exploratory analysis on the impact that vitamin C has on the length of odontoblasts in guinea pigs via two delivery methods (orange juice and ascorbic acid) at varying dosage levels.

It will be shown that the delivery method does not impact on the length of odontoblast but the dosage level does.

Visual exploration

```
library(ggplot2)
library(gridExtra)
library(datasets)
suppressMessages(library(dplyr))
data("ToothGrowth") #loading the required dataset
ToothGrowth$dose <- as.factor(ToothGrowth$dose) #setting dosage variable to factor

ggplot(ToothGrowth, aes(y = len, x = supp))+
  geom_boxplot(aes(fill = supp))+
  facet_wrap(~ dose)+
  labs(title="Boxplot", x="Dosage", y="Length")
```

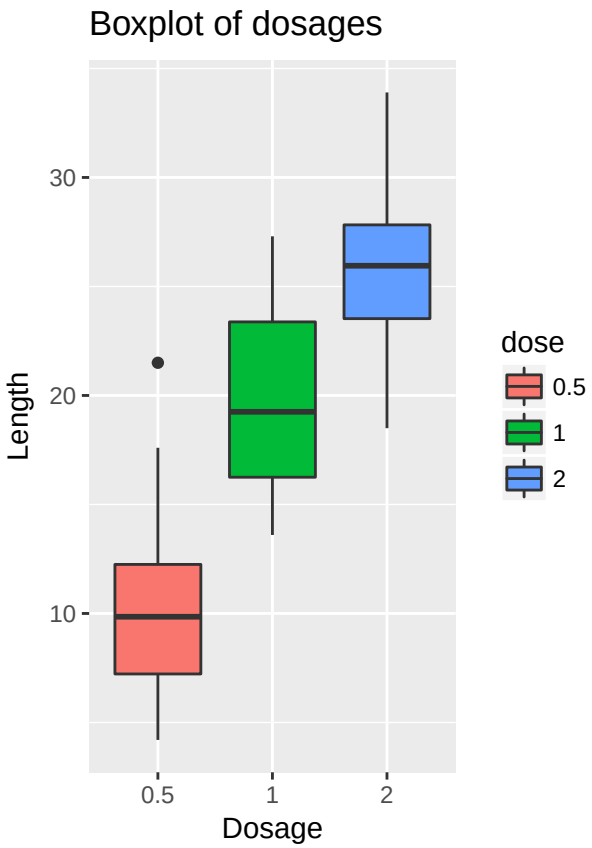


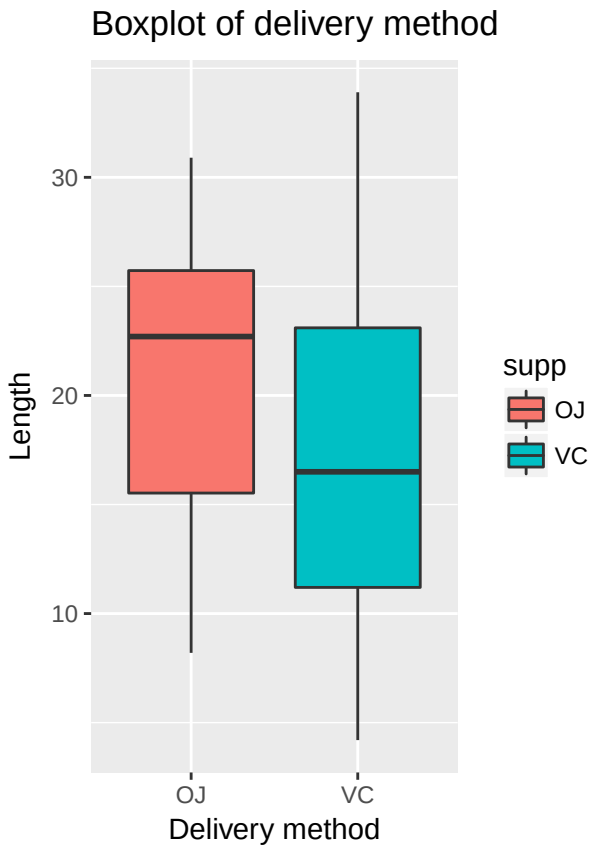
The above graph represents the box-plots of different dosage rates for each delivery method. It appears that there is an increase of legnth in odontoblasts as there is an increase of dosage rates.

```
dosage.graph <- ggplot(ToothGrowth, aes(y = len, x = dose))+
  geom_boxplot(aes(fill=dose))+
  labs(title="Boxplot of dosages", x="Dosage", y="Length")

supp.graph <- ggplot(ToothGrowth, aes(y = len, x = supp))+
  geom_boxplot(aes(fill=supp))+
  labs(title="Boxplot of delivery method", x="Delivery method", y="Length")

grid.arrange(dosage.graph, supp.graph, ncol=2)
```





From the graphs above, it appears that there are distinct groups if you were to seperate them by dosage levels. The seperation does not appear to be as clear between delivery methods.

Numerical analysis

Summary of dose rate:

```
dose.summary <- tapply(ToothGrowth$len, ToothGrowth$dose, summary)
dose.summary
```

```
## $`0.5`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.200   7.225   9.850  10.600  12.250  21.500
##
## $`1`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   13.60   16.25   19.25   19.74   23.38   27.30
##
## $`2`
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   18.50   23.52   25.95   26.10   27.83   33.90
```

Summary of the delivery method:

```
supp.summary <- tapply(ToothGrowth$len, ToothGrowth$supp, summary)
supp.summary
```

```
## $OJ
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    8.20   15.52   22.70   20.66   25.72   30.90
##
## $VC
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    4.20   11.20   16.50   16.96   23.10   33.90
```

```
supp.tresults <- t.test(data = ToothGrowth, len ~ supp, paired = F, var.equal = F)
supp.tresults
```

```
##
##  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
```

From the t-test above, we can see that there is insufficient evidence to reject the null hypothesis that the means between the two groups are equal as the p-value $= 0.0606345 > 0.05$.

```
dose.tresults <- pairwise.t.test( ToothGrowth$len, ToothGrowth$dose, p.adj = "bonf", paired = F, pool.s
dose.tresults
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  ToothGrowth$len and ToothGrowth$dose
##
##   0.5     1
## 1 3.8e-07 -
## 2 1.3e-13 5.7e-05
##
## P value adjustment method: bonferroni
```

From the pairwise t-test for the dosage levels with a bonferroni correction, we can see that the null hypothesis that the means are equal can be rejected with alpha $= 0.05$. Each subgroup is statistically different.

Conclusions and assumptions

We can conclude that dosage rates have a significant impact on the length of odontoblasts for guinea pigs. The following assumptions were made:
* The population are independant
* The populations have unequal variances