Statistical Inference Project Part 2

Filipp Trubin

December 05, 2021

Github repo : Statistical Inference

Load the Dataset

```
library(datasets)
head(ToothGrowth)
```

```
##    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(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 ...
```

Summary of Dataset

```
ToothGrowth$dose <- as.factor(ToothGrowth$dose)
summary(ToothGrowth)
```

```
##       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
```

Exploratory Analysis

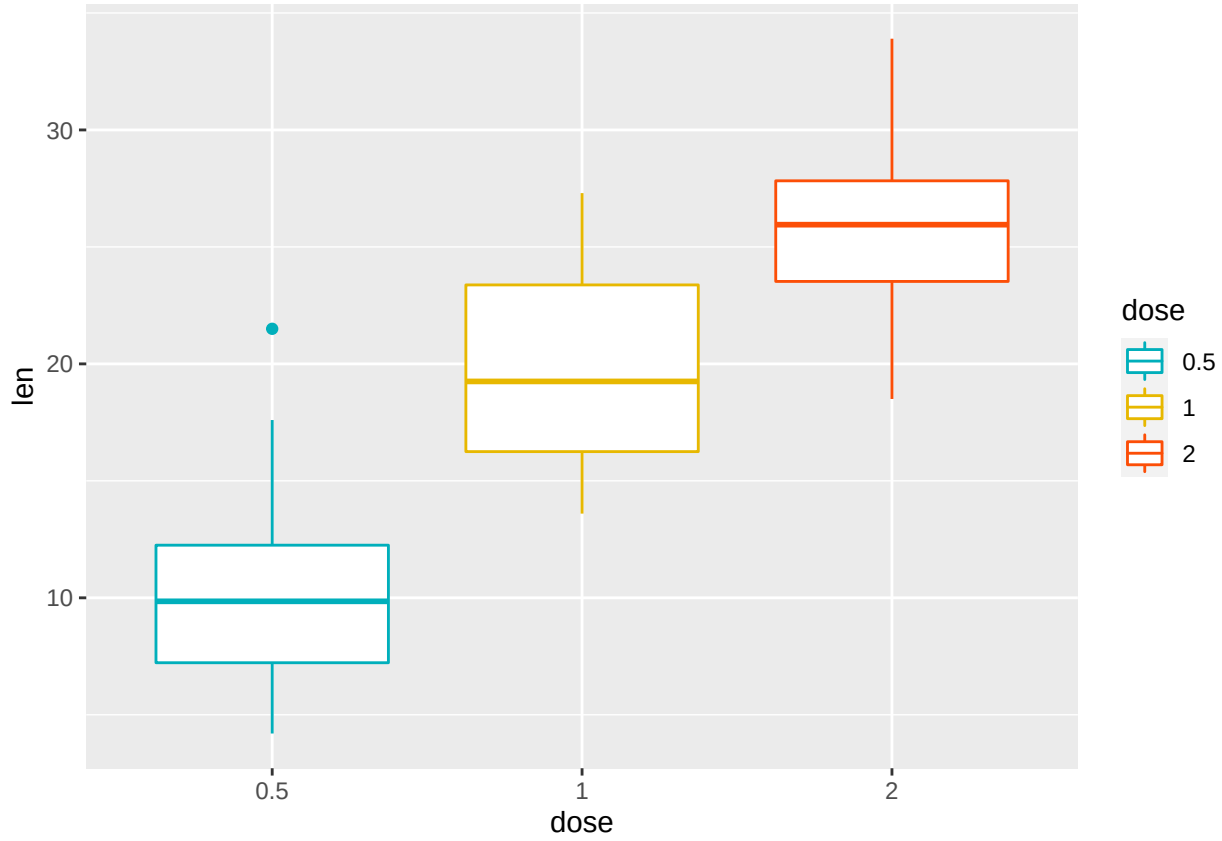
Splitting cases between different dose levels and supplement methods

```
table(ToothGrowth$dose, ToothGrowth$supp)
```

```
##
##      OJ VC
##   0.5 10 10
##   1   10 10
##   2   10 10
```

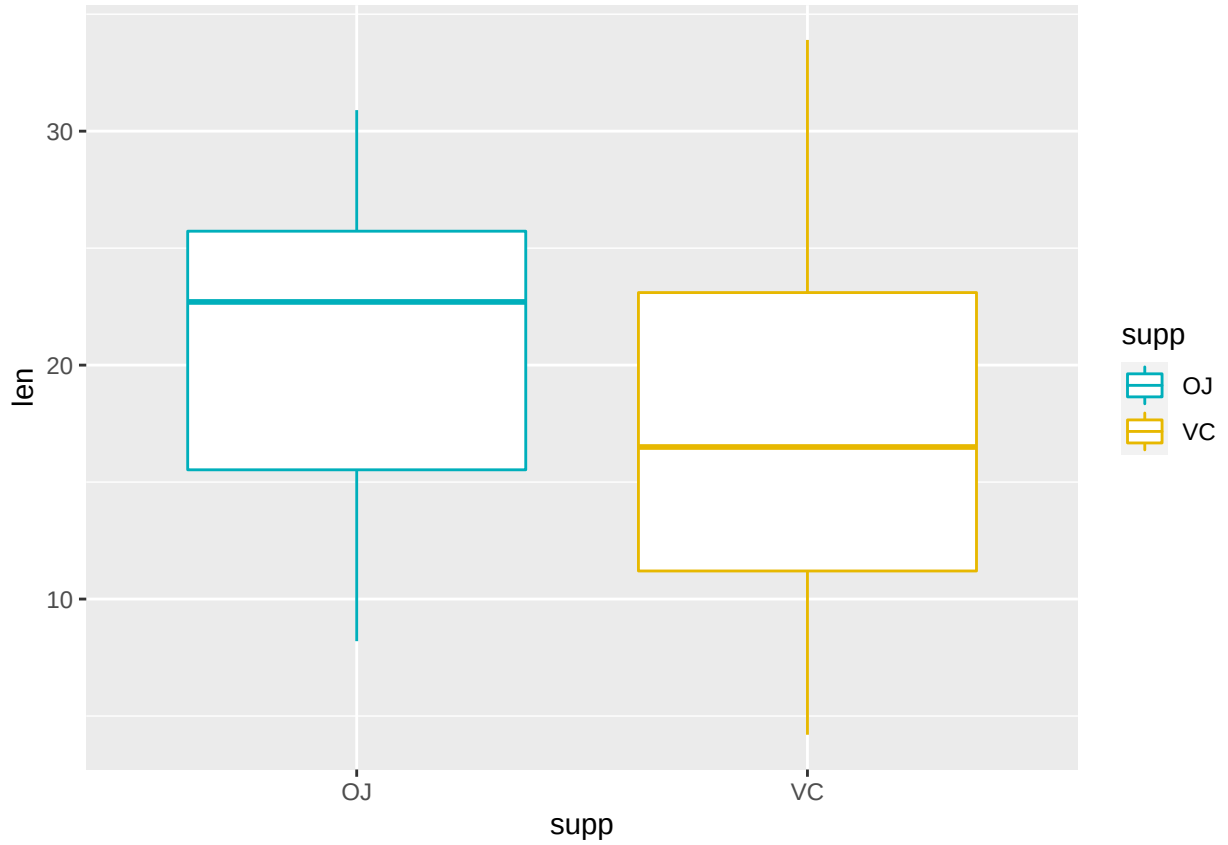
Plot tooth growth as a function of dose

```
library(ggplot2)
ggplot(aes(x=dose, y=len), data=ToothGrowth) + geom_boxplot(aes(color = dose)) + scale_color_manual(valu
```



Plot tooth growth as a function of supplement type

```
library(ggplot2)
ggplot(aes(x=supp, y=len), data=ToothGrowth) + geom_boxplot(aes(color = supp)) +
  scale_color_manual(values = c("#00AFBB", "#E7B800"))
```



##Comparing tooth growth by supp and dose through hypothesis testing

```
t.test(len ~ supp, data = ToothGrowth)
```

```
##
##  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 between group OJ and group VC 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
```

The p-value is 0.06.It means we can not decline the null hypothesis that the different supplement types have no effect on tooth length.

Analyzing the effect of dose on tooth growth:

3 sub-groups as per dose level pairs:

```
doses_0.5_1.0 <- subset (ToothGrowth, dose %in% c(0.5, 1.0))
doses_0.5_2.0 <- subset (ToothGrowth, dose %in% c(0.5, 2.0))
doses_1.0_2.0 <- subset (ToothGrowth, dose %in% c(1.0, 2.0))
```

assuming unequal variances between the two groups

```
t.test(len ~ dose, data = doses_0.5_1.0)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means between group 0.5 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

```
t.test(len ~ dose, data = doses_1.0_2.0)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

For 3 pairs of dose levels, the P value is less than 0.05. Which means we can reject the null hypothesis.

##Conclusions

###The type of supplement has no effect on the growth of teeth. Increased tooth growth depends on the dose increasing.