

# Project2

Mohamed Elashri

March 22, 2020

## Statistical Inference Course Project 2

### Introduction

Load the ToothGrowth data and perform some basic exploratory data analyses

Load the ToothGrowth data and perform some basic exploratory data analyses

```
# libraries
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

library(datasets)
library(gridExtra)
library(GGally)

# The Effect of Vitamin C on Tooth Growth in Guinea Pigs
data(ToothGrowth)
toothGrowth <- ToothGrowth
toothGrowth$dose <- as.factor(toothGrowth$dose) # factor conversion
```

### basic summary of the data

```
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 ...
## $ dose: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 ...

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

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

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

Now we do Analysis of Variance ANOVA test

```
anova <- aov(len ~ supp * dose, data=toothGrowth)
summary(anova)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## supp          1  205.4    205.4   15.572 0.000231 ***
## dose          2 2426.4   1213.2   92.000 < 2e-16 ***
## supp:dose      2  108.3     54.2    4.107 0.021860 *
## Residuals     54  712.1     13.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results here support that there is a notable interaction between the length (len) and dosage (dose) with F values is 15.572 and with  $p < 0.01$ . Also there is weak interaction between the combination of supplement type (supp) and dosage (dose) compared to the length (len) ,  $F = 4.107$  ,  $p < 0.05$ .

## Use confidence intervals

```
confint(anova)
```

```
##              2.5 %    97.5 %
## (Intercept) 10.9276907 15.532309
## suppVC      -8.5059571 -1.994043
## dose1        6.2140429 12.725957
## dose2        9.5740429 16.085957
## suppVC:dose1 -5.2846186  3.924619
## suppVC:dose2  0.7253814  9.934619
```

```
print(model.tables(anova, "means"), digits=3)
```

```
## Tables of means
## Grand mean
##
## 18.81333
##
##   supp
##   supp
##   OJ   VC
## 20.66 16.96
##
##   dose
##   dose
##   0.5   1   2
## 10.60 19.73 26.10
##
##   supp:dose
##   dose
```

```
## supp 0.5    1      2
##    OJ 13.23 22.70 26.06
##    VC  7.98 16.77 26.14
```

### **conclusions**

from our observation, We conclude that a higher dose is the main factor to increase the Tooth Growth using VC but OJ has better results at lower doses.