Basic inferential data analysis

1. ToothGrowth data - basic exploratory data analysis

Loading and subsetting

```
library(datasets)
attach(ToothGrowth)
OJ <- split(split(len,f = as.factor(supp))[[1]],f = as.factor(dose[1:30]))
VC <- split(split(len,f = as.factor(supp))[[2]],f = as.factor(dose[31:60]))
```

Creating boxplot

```
par(mfrow = c(1,2))
boxplot(OJ,main="Orange Juice",xlab="dose (mg)",ylab="Tooth length", col="orange", ylim = c(0,35))
boxplot(VC,main="Vitamin C",xlab="dose (mg)",ylab="Tooth length", col="gray", ylim = c(0,35))
```

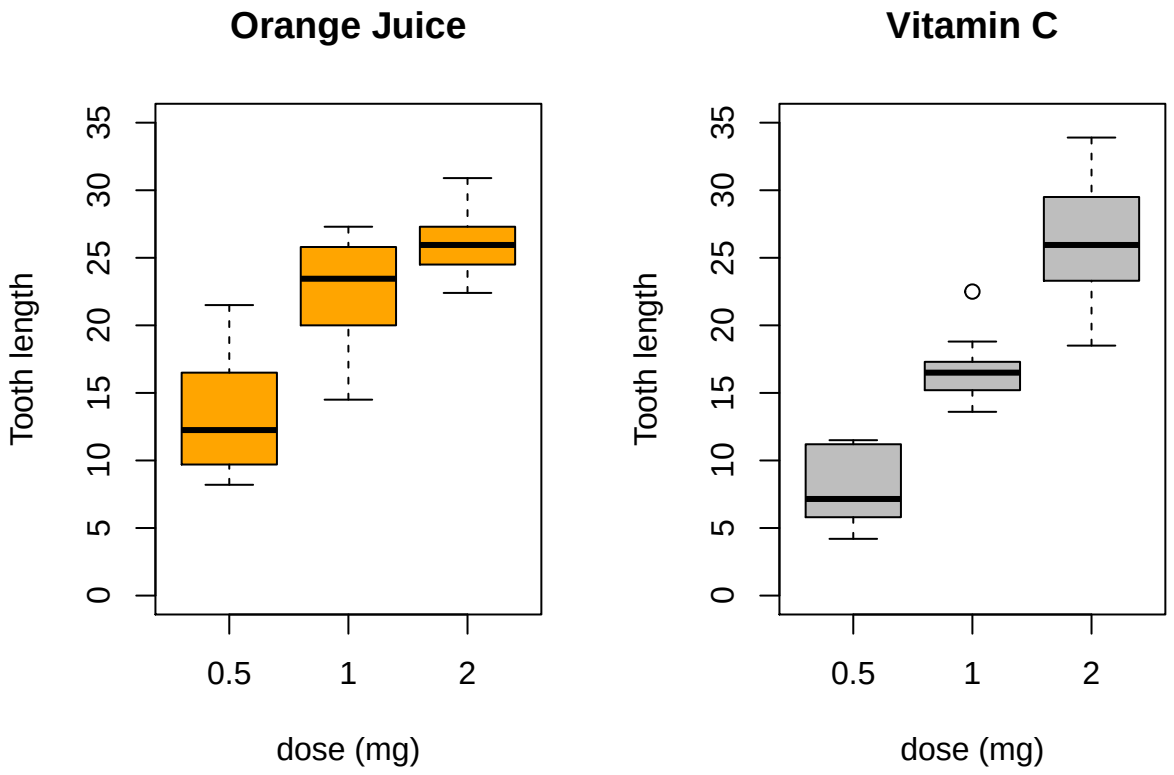

Figure 1. Boxplot showing the effect of vitamin C on tooth growth (in form of Orange Juice or Vitamin C [ascorbic acid])

2. Basic summary of the data

Applying summary function

Tooth growth quantiles of animals supplied with 0.5, 1 or 2 mg vitamin C equivalent Orange Juice

```
x <- sapply(OJ,summary)
colnames(x) <- c("0.5mg","1mg","2mg")
x
```

```
##         0.5mg   1mg   2mg
## Min.     8.20 14.50 22.40
## 1st Qu.  9.70 20.30 24.58
## Median  12.25 23.45 25.95
## Mean    13.23 22.70 26.06
## 3rd Qu. 16.18 25.65 27.08
## Max.    21.50 27.30 30.90
```

Tooth growth quantiles of animals supplied with 0.5, 1 or 2 mg Vitamin C (ascorbic acid)

```
x <- sapply(VC,summary)
colnames(x) <- c("0.5mg","1mg","2mg")
x
```

```
##         0.5mg   1mg   2mg
## Min.     4.20 13.60 18.50
## 1st Qu.  5.95 15.27 23.38
## Median   7.15 16.50 25.95
## Mean     7.98 16.77 26.14
## 3rd Qu. 10.90 17.30 28.80
## Max.    11.50 22.50 33.90
```

3. Reporting t confidence intervals

T confidence intervals for Orange Juice treatment

```
x <- matrix(nrow = 3,ncol =3)
for(i in 1:3){x[i,c(1,3)] <- t.test(OJ[[i]])$conf.int[1:2]}
x[,2] <- sapply(OJ,mean)
rownames(x) <- c("0.5mg","1mg","2mg")
colnames(x) <- c("lower","mean","upper")
x
```

```
##          lower  mean    upper
## 0.5mg 10.03972 13.23 16.42028
## 1mg   19.90227 22.70 25.49773
## 2mg   24.16069 26.06 27.95931
```

T confidence intervals for vitamin C treatment

```
x <- matrix(nrow = 3,ncol =3)
for(i in 1:3){x[i,c(1,3)] <- t.test(VC[[i]])$conf.int[1:2]}
x[,2] <- sapply(VC,mean)
rownames(x) <- c("0.5mg","1mg","2mg")
colnames(x) <- c("lower","mean","upper")
x
```

```
##           lower  mean     upper
## 0.5mg  6.015176  7.98  9.944824
## 1mg   14.970657 16.77 18.569343
## 2mg   22.707910 26.14 29.572090
```

T-tests p-values within Orange Juice treated for different doses

```
x <- vector(length = 3)

x[1] <- t.test(OJ[[1]],OJ[[2]])$p.value
x[2] <- t.test(OJ[[1]],OJ[[3]])$p.value
x[3] <- t.test(OJ[[2]],OJ[[3]])$p.value
names(x) <- c("0.5mg-1mg","0.5mg-2mg","1mg-2mg")
x
```

```
##    0.5mg-1mg    0.5mg-2mg      1mg-2mg
## 8.784919e-05 1.323784e-06 3.919514e-02
```

T-tests p-values within Vitamin C treated for different doses

```
x <- vector(length = 3)

x[1] <- t.test(VC[[1]],VC[[2]])$p.value
x[2] <- t.test(VC[[1]],VC[[3]])$p.value
x[3] <- t.test(VC[[2]],VC[[3]])$p.value
names(x) <- c("0.5mg-1mg","0.5mg-2mg","1mg-2mg")
x
```

```
##    0.5mg-1mg    0.5mg-2mg      1mg-2mg
## 6.811018e-07 4.681577e-08 9.155603e-05
```

T-tests p-values between Orange Juice vs. vitamin C treated for given dose

```
x <- vector(length = 3)

x[1] <- t.test(VC[[1]],OJ[[1]])$p.value
x[2] <- t.test(VC[[2]],OJ[[2]])$p.value
x[3] <- t.test(VC[[3]],OJ[[3]])$p.value
names(x) <- c("0.5mg","1mg","2mg")
x
```

```
##       0.5mg         1mg         2mg
## 0.006358607 0.001038376 0.963851589
```

4. Conclusions

- Increasing amounts of Orange Juice or Vitamin C supplements resulted in increased tooth growth
- 1 mg supplement significantly increased tooth growth compared to 0.5 mg supplement
- 1 mg Orange Juice was more effective than 1 mg Vitamin C
- 2 mg supplements were the most effective (Orange Juice and Vitamin C had equal effect)

Assumptions:

- t-test assumes that the variables are normally distributed
- multiple tests were applied without any test correction