Course6-2

Sangwon Han

2018-12-18

Overview

In the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and explore data analyses

```
data(ToothGrowth)
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 ...
```

```
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
```

```
unique(ToothGrowth$dose)
```

```
## [1] 0.5 1.0 2.0
```

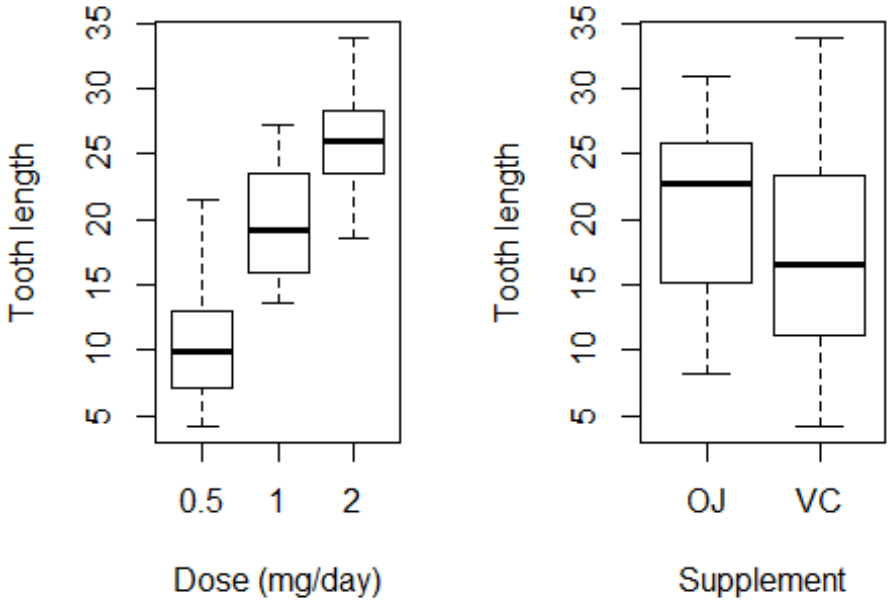
2. Provide a basic summary of the data.

```
summary(ToothGrowth)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

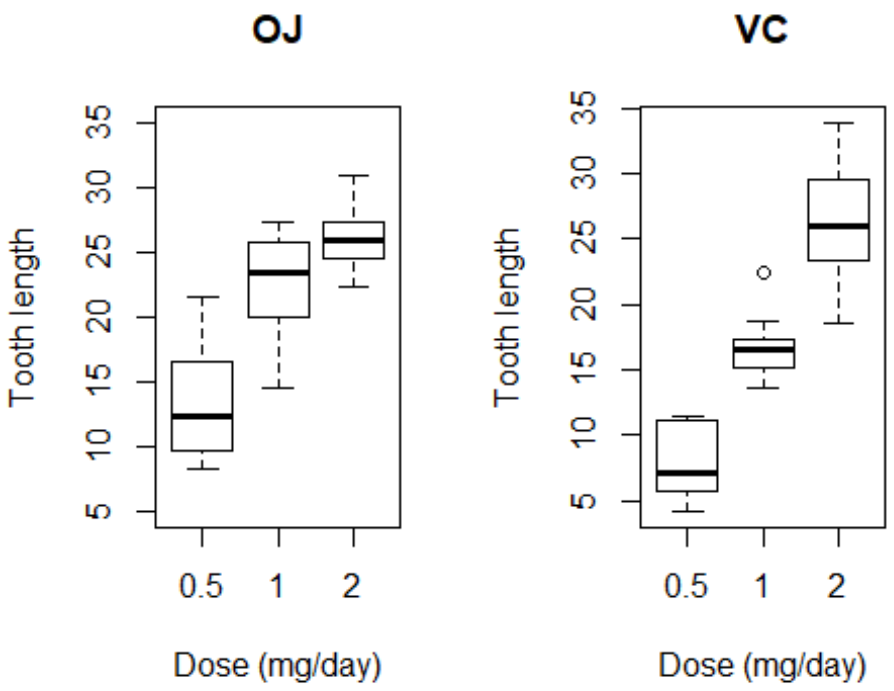
Plot box and whisker diagram

```
par(mfrow = c(1,2))
plot(len ~ as.factor(dose), data = ToothGrowth, xlab = "Dose (mg/day)", ylab
= "Tooth length")
plot(len ~ supp, data = ToothGrowth, xlab = "Supplement", ylab = "Tooth lengt
h")
```



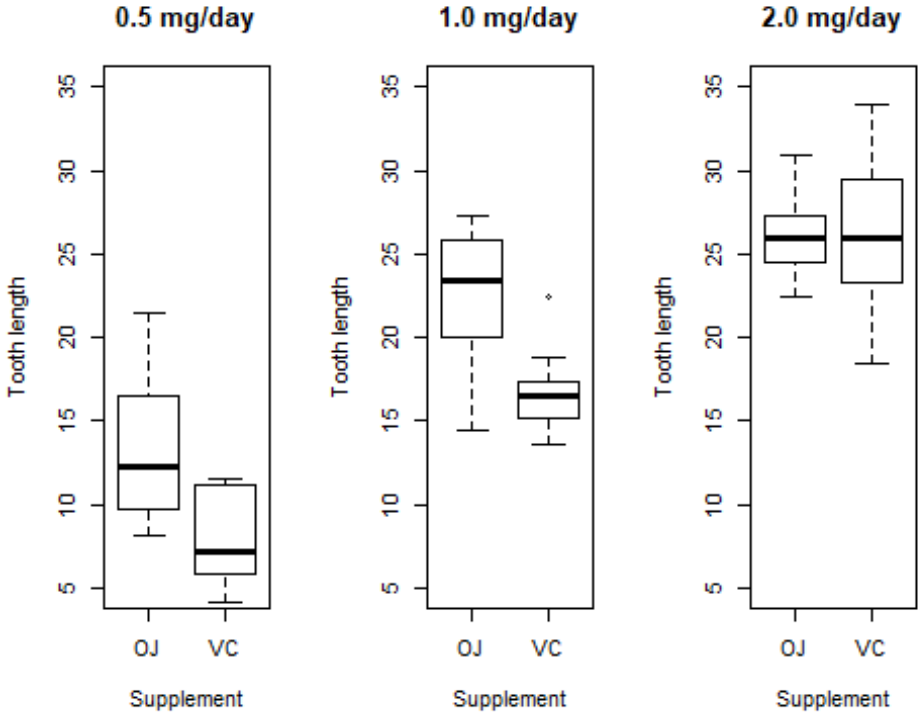
Plot box and whisker diagram according to supplement type subgroups

```
par(mfrow = c(1,2))
plot(len[ToothGrowth$supp == "OJ"] ~ as.factor(dose[ToothGrowth$supp == "OJ
"]), data = ToothGrowth, xlab = "Dose (mg/day)", ylab = "Tooth length", main
= "OJ", ylim = c(5,35))
plot(len[ToothGrowth$supp == "VC"] ~ as.factor(dose[ToothGrowth$supp == "OJ
"]), data = ToothGrowth, xlab = "Dose (mg/day)", ylab = "Tooth length", main
= "VC")
```



Plot box and whisker diagram according to dose subgroups

```
par(mfrow = c(1,3))
plot(len[ToothGrowth$dose == 0.5] ~ supp[ToothGrowth$dose == 0.5], data = Too
thGrowth, xlab = "Supplement", ylab = "Tooth length", main = "0.5 mg/day", yl
im = c(5, 35))
plot(len[ToothGrowth$dose == 1.0] ~ supp[ToothGrowth$dose == 1.0], data = Too
thGrowth, xlab = "Supplement", ylab = "Tooth length", main = "1.0 mg/day", yl
im = c(5, 35))
plot(len[ToothGrowth$dose == 2.0] ~ supp[ToothGrowth$dose == 2.0], data = Too
thGrowth, xlab = "Supplement", ylab = "Tooth length", main = "2.0 mg/day", yl
im = c(5, 35))
```



3. Hypothesis tests to compare tooth growth by supp and dose.

3-1. Supplement

See whether the test meets equal variance assumption

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

```
##
##  F test to compare two variances
##
## data:  len by supp
## F = 0.6386, num df = 29, denom df = 29, p-value = 0.2331
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3039488 1.3416857
## sample estimates:
## ratio of variances
##          0.6385951
```

No significant difference in variance are shown.

Perform t.test function to compare two supplement groups

```
t.test(len~supp, data = ToothGrowth, paired = FALSE, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1670064  7.5670064
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

3-2. Dose

See whether the test meets equal variance assumption

```
bartlett.test(len~as.factor(dose), data = ToothGrowth)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  len by as.factor(dose)
## Bartlett's K-squared = 0.66547, df = 2, p-value = 0.717
```

No significant difference in variance are shown.

Perform aov function to compare three dose groups

```
ANOVA <- aov(len~as.factor(dose), data = ToothGrowth)
summary(ANOVA)
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)
## as.factor(dose)  2   2426    1213   67.42 9.53e-16 ***
## Residuals       57   1026      18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Then, which group shows significant difference in tooth length?

```
TukeyHSD(ANOVA)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = len ~ as.factor(dose), data = ToothGrowth)
##
## $`as.factor(dose)`
##         diff       lwr       upr    p adj
## 1-0.5  9.130  5.901805 12.358195 0.00e+00
## 2-0.5 15.495 12.266805 18.723195 0.00e+00
## 2-1    6.365  3.136805  9.593195 4.25e-05
```

4. conclusions

As shown above, we can conclude

1. Tooth length does not differ according to supplement delivery type.
2. Tooth length does differ among the three dose levels of vitamin C.