Author: Giuseppa Cefalu
Date: 9/30/2020

```
rm(list=ls())

library(dplyr)
library(ggplot2)
library(datasets)
library(gridExtra)
library(knitr)
library(rmarkdown)

data(ToothGrowth)
data <- ToothGrowth

cat("EFFECT OF SUPPLEMENT TYPE AND DOSE IN TOOH GROWTH\n")
```

```
## EFFECT OF SUPPLEMENT TYPE AND DOSE IN TOOH GROWTH
```

```
cat("This is and analysis of the R TeoohGrowth data in R data sets package\n")
```

```
## This is and analysis of the R TeoohGrowth data in R data sets package
```

```
cat("EXPLORATORY DATA ANALYSIS\n\n")
```

```
## EXPLORATORY DATA ANALYSIS
```

```
cat("VARIABLES")
```

```
## VARIABLES
```

```
print(str(data))
```

```
## '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 ...
## NULL
```

```
cat("SUMMARY")
```

```
## SUMMARY
```

```
params <- data %>%
  group_by(supp) %>%
  summarise(Mean = mean(len, na.rm = TRUE), Sd = sd(len, na.rm = TRUE),
            Max = max(len, na.rm = TRUE), Min = mean(len, na.rm = TRUE),
            Median = median(len, na.rm = TRUE),
            Quantile0 = quantile(len, na.rm = TRUE, disp = 0, prob = 0),
            Quantile25 = quantile(len, na.rm = TRUE, disp = 0.25, prob = 0.25),
            Quantile50 = quantile(len, na.ram = TRUE, disp = 0.5, prob = 0.5),
            Quantile75 = quantile(len, na.rm = TRUE, disp = 0.75, prob = 0.75),
            Quantile100 = quantile(len, na.rm = TRUE, disp = 1, prob = 1))
```

```
print(params)
```

```
## # A tibble: 2 x 11
##   supp   Mean    Sd   Max   Min Median Quantile0 Quantile25 Quantile50
##   <fct> <dbl> <dbl> <dbl> <dbl>  <dbl>     <dbl>      <dbl>      <dbl>
## 1 OJ     20.7  6.61  30.9  20.7   22.7       8.2       15.5       22.7
```

```
## 2 VC     17.0  8.27  33.9  17.0   16.5       4.2       11.2       16.5
## # ... with 2 more variables: Quantile75 <dbl>, Quantile100 <dbl>
```

```
cat("\nThere is some variation in the data between the 2 samples. I will assume that
variance is different\n")
```

```
##
## There is some variation in the data between the 2 samples. I will assume that
## variance is different
```

```
params2 <- data %>%
  group_by(dose) %>%
  summarise(Mean = mean(len, na.rm = TRUE), Sd = sd(len, na.rm = TRUE),
            Max = max(len, na.rm = TRUE), Min = mean(len, na.rm = TRUE),
            Median = median(len, na.rm = TRUE),
            Quantile0 = quantile(len, na.rm = TRUE, disp = 0, prob = 0),
            Quantile25 = quantile(len, na.rm = TRUE, disp = 0.25, prob = 0.25),
            Quantile50 = quantile(len, na.ram = TRUE, disp = 0.5, prob = 0.5),
            Quantile75 = quantile(len, na.rm = TRUE, disp = 0.75, prob = 0.75),
            Quantile100 = quantile(len, na.rm = TRUE, disp = 1, prob = 1))
```

```
print(params2)
```

```
## # A tibble: 3 x 11
##    dose  Mean    Sd   Max   Min Median Quantile0 Quantile25 Quantile50
##   <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl>     <dbl>      <dbl>      <dbl>
## 1   0.5  10.6  4.50  21.5  10.6   9.85       4.2       7.22       9.85
## 2   1    19.7  4.42  27.3  19.7  19.2       13.6      16.2       19.2
## 3   2    26.1  3.77  33.9  26.1  26.0       18.5      23.5       26.0
## # ... with 2 more variables: Quantile75 <dbl>, Quantile100 <dbl>
```

```
p <- ggplot(data, aes(x = dose, y = len, fill =  factor(dose))) +
  geom_dotplot(binaxis='y',  dotsize=0.75) + facet_grid(.~supp) +
  stat_summary(fun=mean, geom="point", shape=18,
               size=3, color="black") + ggtitle("Tooth growth versus dose and suplement") +
  theme(plot.title=element_text(face="bold", size=9))

OJ <- filter(data, supp == 'OJ', na.rm = TRUE)
VC <- filter(data, supp == 'VC', na.rm = TRUE)

pdOJ <- ggplot(OJ, aes(x =len, fill = 30)) + theme_bw() +
  geom_histogram(aes(y = ..density..), alpha = 0.7, binwidth = 0.75,  col = "black") +
  ggtitle("OJ Supplement tooth growth distribution") +
  theme(plot.title=element_text(face="bold", size=9))


pdVC <- ggplot(VC, aes(x =len, fill = 30)) + theme_bw() +
  geom_histogram(aes(y = ..density..), alpha = 0.7,  binwidth = 0.75, col = "black") +
  ggtitle("VC Supplement tooth growth dsitribution") +
  theme(plot.title=element_text(face="bold", size=9))

grid.arrange(p, pdOJ, pdVC)
```

```
cat("\nTeeth growth may be affected by dose and type of supplement\n\n")
```

```
##
## Teeth growth may be affected by dose and type of supplement
```

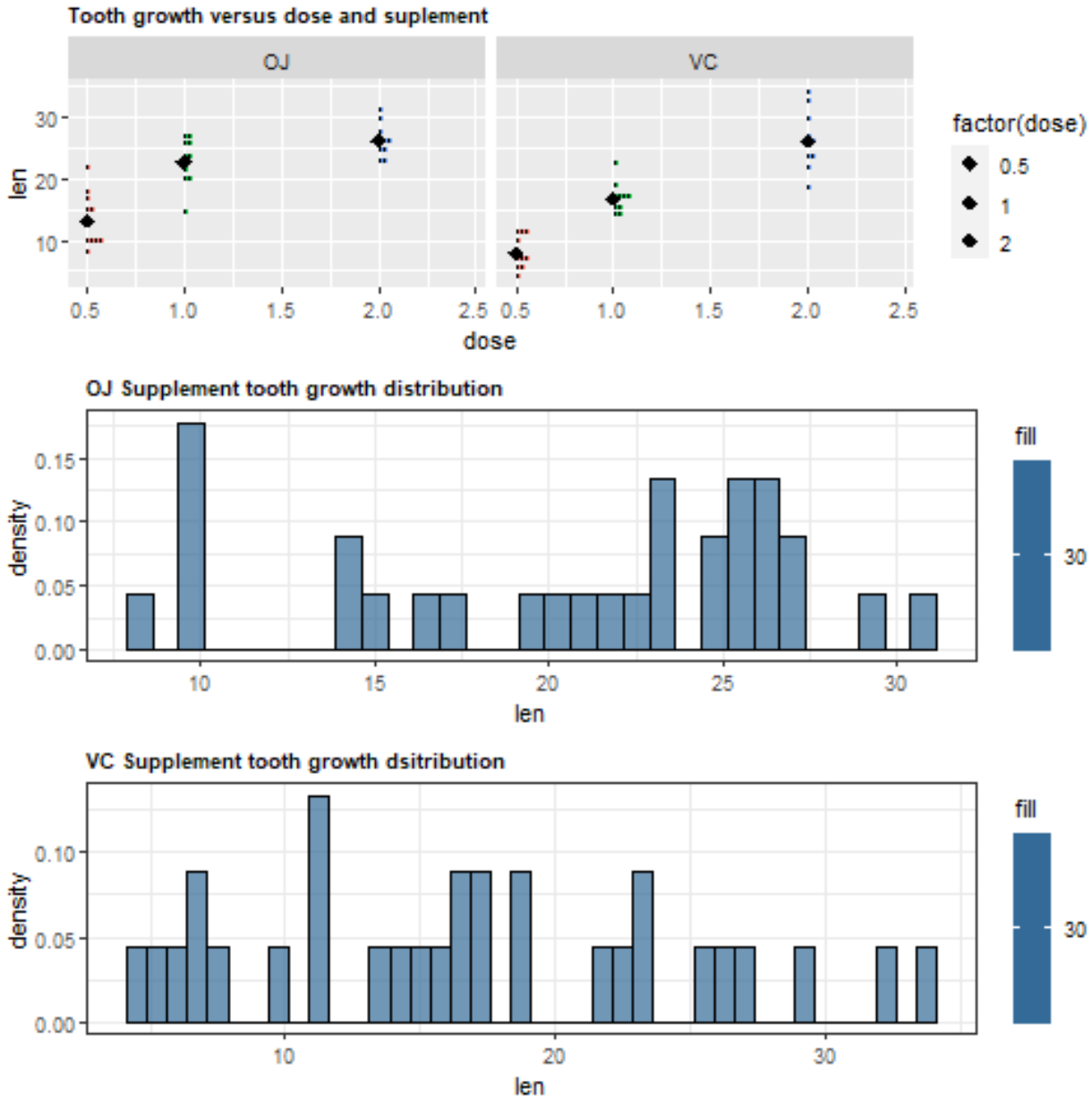

Figure 1: plots of frquencies and distributions

```r
cat("HYPOTHSESIS TESTING\n\n")
```

```
## HYPOTHSESIS TESTING
```

```r
cat("Assumptions:\n")
```

```
## Assumptions:
```

```r
cat("    The variables are (iid) independent and identically distributed.
    tooth growth changes with supplement and dose
    Tooth growth data are normally distributed in the two samples: OJ and CV
    Variance seems to be different both while grouping the data by supp or dose
    I will use a T test with a significance level with alpha=0.05 var.equal = FALSE\n\n")
```

```
##     The variables are (iid) independent and identically distributed.
##     tooth growth changes with supplement and dose
##     Tooth growth data are normally distributed in the two samples: OJ and CV
##     Variance seems to be different both while grouping the data by supp or dose
##     I will use a T test with a significance level with alpha=0.05 var.equal = FALSE
```

```r
cat("\nEFECT OF SUPLEMENT ON TEETH GROWTH\n\n")
```

```
##
## EFECT OF SUPLEMENT ON TEETH GROWTH
```

```r
OJlen <- select(OJ, len)
VClen <- select(VC, len)

print(t.test(OJlen, VClen, alternative = "two.sided", paired = FALSE,
             var.equal = FALSE, conf.level = 0.95))
```

```
##
##  Welch Two Sample t-test
##
## data:  OJlen and VClen
## 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 of x mean of y
##  20.66333  16.96333
```

```r
cat("It does not seem to be any difference in the effect exerted by supplement type
in teeth growth\n")
```

```
## It does not seem to be any difference in the effect exerted by supplement type
## in teeth growth
```

```r
cat("\nEFFECT OF SUPPLEMENT DOSE ON TEETH GROWTH\n\n")
```

```
##
## EFFECT OF SUPPLEMENT DOSE ON TEETH GROWTH
```

```r
data0.5 <- filter(data, dose == 1/2)
data1 <- filter(data, dose == 1)
data2 <- filter(data, dose == 2)
```

```r
Len05 <- select(data0.5, len)
Dose05 <- select(data0.5, dose)

Len1 <- select(data1, len)
Dose1 <- select(data1, dose)

Len2 <- select(data2, len)
Dose2 <- select(data2, dose)


print(t.test(Len05,Len1, alternative = "two.sided", paired = F,
             var.equal = F, conf.level = 0.95))
```

```
##
##  Welch Two Sample t-test
##
## data:  Len05 and Len1
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean of x mean of y
##    10.605    19.735
```

```r
print(t.test(Len1, Len2, alternative = "two.sided", paired = F,
             var.equal = F, conf.level = 0.95))
```

```
##
##  Welch Two Sample t-test
##
## data:  Len1 and Len2
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean of x mean of y
##    19.735    26.100
```

```r
print(t.test(Len05,Len2, alternative = "two.sided", paired = F,
             var.equal = F, conf.level = 0.95))
```

```
##
##  Welch Two Sample t-test
##
## data:  Len05 and Len2
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean of x mean of y
##    10.605    26.100
```

```r
cat("Supplement dose affects teeth growth")
```

```
## Supplement dose affects teeth growth
```

```r
cat("\nCONCLUSSION\n\n")
```

```
##
## CONCLUSSION
```

```r
cat("increasing the supplement's dose benefits tooth growth, but the type of supplement does not have an effect on teeth growth.")
```

```
## increasing the supplement's dose benefits tooth growth, but the type of supplement does not have
## an effect on teeth growth.
```