Part 2 Basic Inferential Data_Analysis

Volha Leusha

In this paper analysis of the ToothGrowth dataset from Rdataset package is presented and basic summary of data is provided. In second part confidence intervals and hypothesis tests are used to compare tooth growth by supp and dose, with subsequent conclusions on results.

Dataset Base Analysis

1. Load dataset from datasets library:
2. Basic summary of data exploratory data analysis:

```
##    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
```

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

```
##       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
```

The dataset is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice or ascorbic acid (a form of vitamin C and coded as VC).

Dataset has 60 observarions of 3 variables:

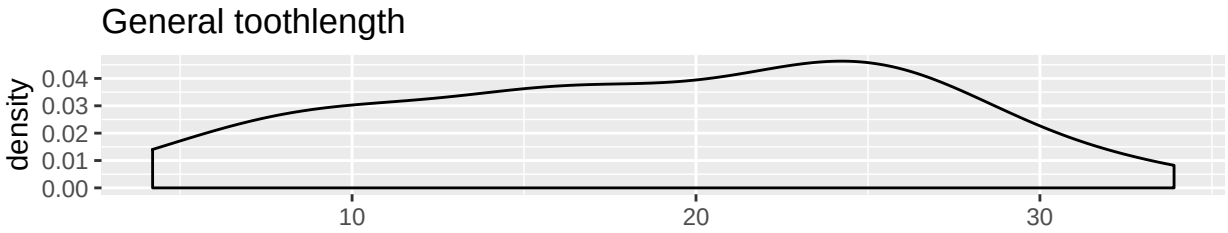
- supp: factor with only two possible supplement types: OJ, VC with equal proportion
- dose: numerical categorical variable that can take only 3 unique values (milligrams/day)

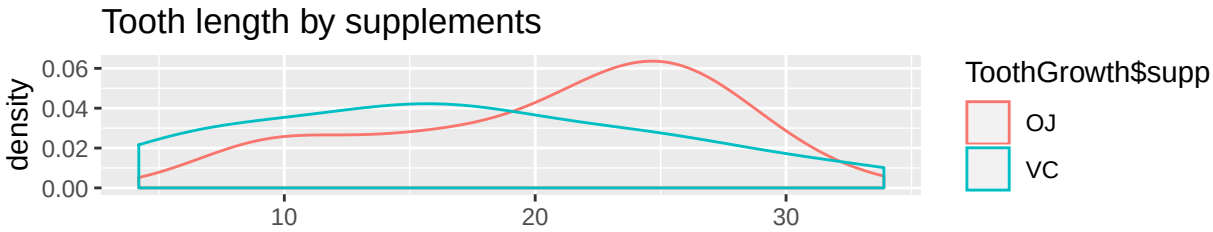
```
## [1] 0.5 1.0 2.0
```

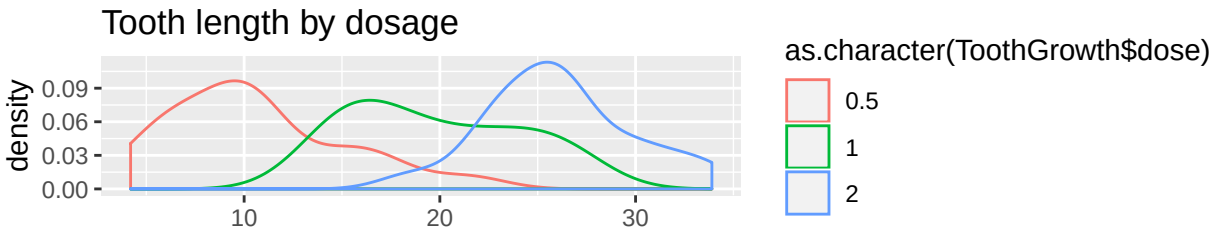
- len: numerical tooth length in range from 4.2 to 33.90 with mean 18.81 and standard deviation

```
## [1] 7.649315
```

len variable has following density distribution:



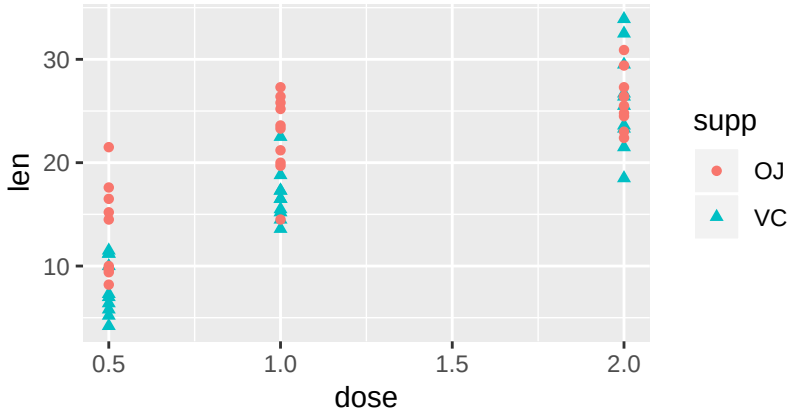




From graphs can be seen that data is skewed in general for all groups.

Tooth growth comparisson by supp and dose

Plot dataset grouped by supplement type and dosage:



From the graph and table it can be seen that there is potential difference in tooth length between supplements, suggesting that pigs on supplement OJ have longer teeth than on VC. Also, the bigger the dossage, the longer are teeth. To prove above assumptions, three t-tests are performed.

1. First test:
   Ho: there is no difference between tooth length with different supplements Ha: tooth length for supplement OJ in higher
   alpha =.05

```
## [1] 1.408659 5.991341
## attr(,"conf.level")
## [1] 0.95
```

Confidence interval is fully above zero (and p-value is less than alpha) consequently null hypothesis can be rejected and it can be assumed that pig with supplement OJ have bigger teeth length.

2. Second test:
   Ho: there is no difference between dose0.5 and 1.0 for tooth length Ha: tooth length for dose 1.0 is higher than for dose 0.5
   alpha= .05

```
##
##  One Sample t-test
##
## data:  diff2
## t = 6.9669, df = 19, p-value = 1.225e-06
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##   6.387121 11.872879
## sample estimates:
## mean of x
##      9.13
```

Confidence interval is fully above zero (and p-value is less than alpha) consequently null hypothesis can be rejected and it can be assumed that Ha is true.

3. Third test:
   Ho: there is no difference between dose 1.0 and 2.0 for tooth length Ha: tooth length for dose 2.0 is higher than for dose 1.0
   alpha= .05

```
##
##  One Sample t-test
##
## data:  diff3
## t = 4.6046, df = 19, p-value = 0.0001934
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.471814 9.258186
## sample estimates:
## mean of x
##     6.365
```

Confidence interval is fully above zero (and p-value is less than alpha) consequently null hypothesis can be rejected and it can be assumed that Ha is true.

Appendix