Simulation

NithaDuff

7/23/2020

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.2
```

ToothGrowth data analysis

Here we are analysing the data reported for the ToothGrowth of a sample of 60 guinea pigs with varying dosage and 2 alternative supplements. The data looks like the following table.

```
data("ToothGrowth")
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
```

```
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
```

```
len <- ToothGrowth$len
```

As seen above the data frame has 3 columns each containing the length(of tooth), supplement group, dosage given. There are 2 supplements and 3 variant dosages of each supplement. This is equally divided between 60 guinea pigs.

```
t.test(len ~ I(relevel(supp,2)), paired = FALSE, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by I(relevel(supp, 2))
## 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:
##  -7.5710156  0.1710156
```

```
## sample estimates:
## mean in group VC mean in group OJ
##         16.96333         20.66333
```

from the above data, we can analyse the group means, the confidence interval of the entire data and so on.

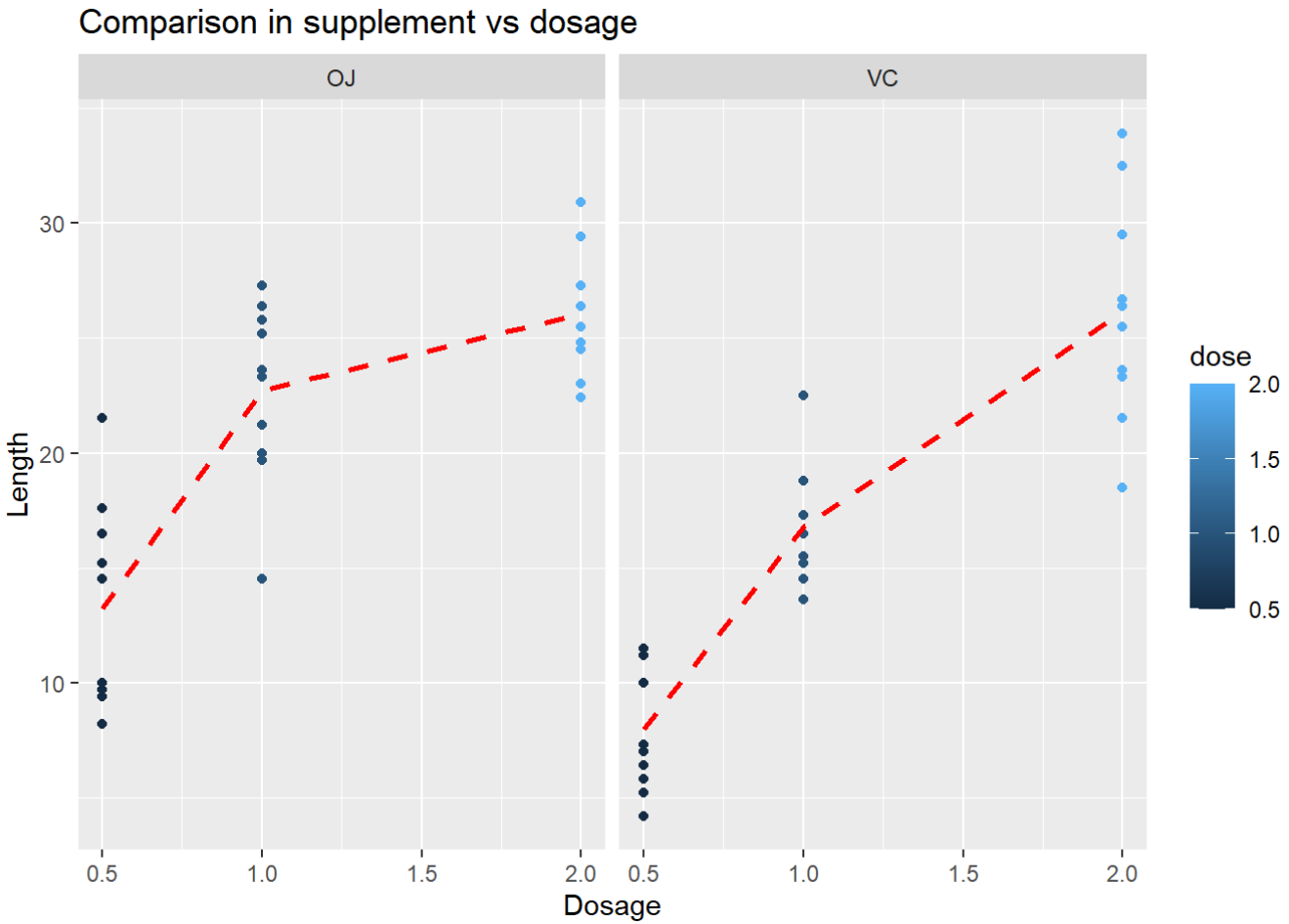
ToothGrowth data Visualisation

```
tooth_mns <- ToothGrowth %>% group_by(supp,dose)  %>% summarise(mean = mean(len)) %>% print()
```

```
## `summarise()` regrouping output by 'supp' (override with `.groups` argument)
```

```
## # A tibble: 6 x 3
## # Groups:   supp [2]
##   supp   dose  mean
##   <fct> <dbl> <dbl>
## 1 OJ      0.5 13.2
## 2 OJ      1   22.7
## 3 OJ      2   26.1
## 4 VC      0.5  7.98
## 5 VC      1   16.8
## 6 VC      2   26.1
```

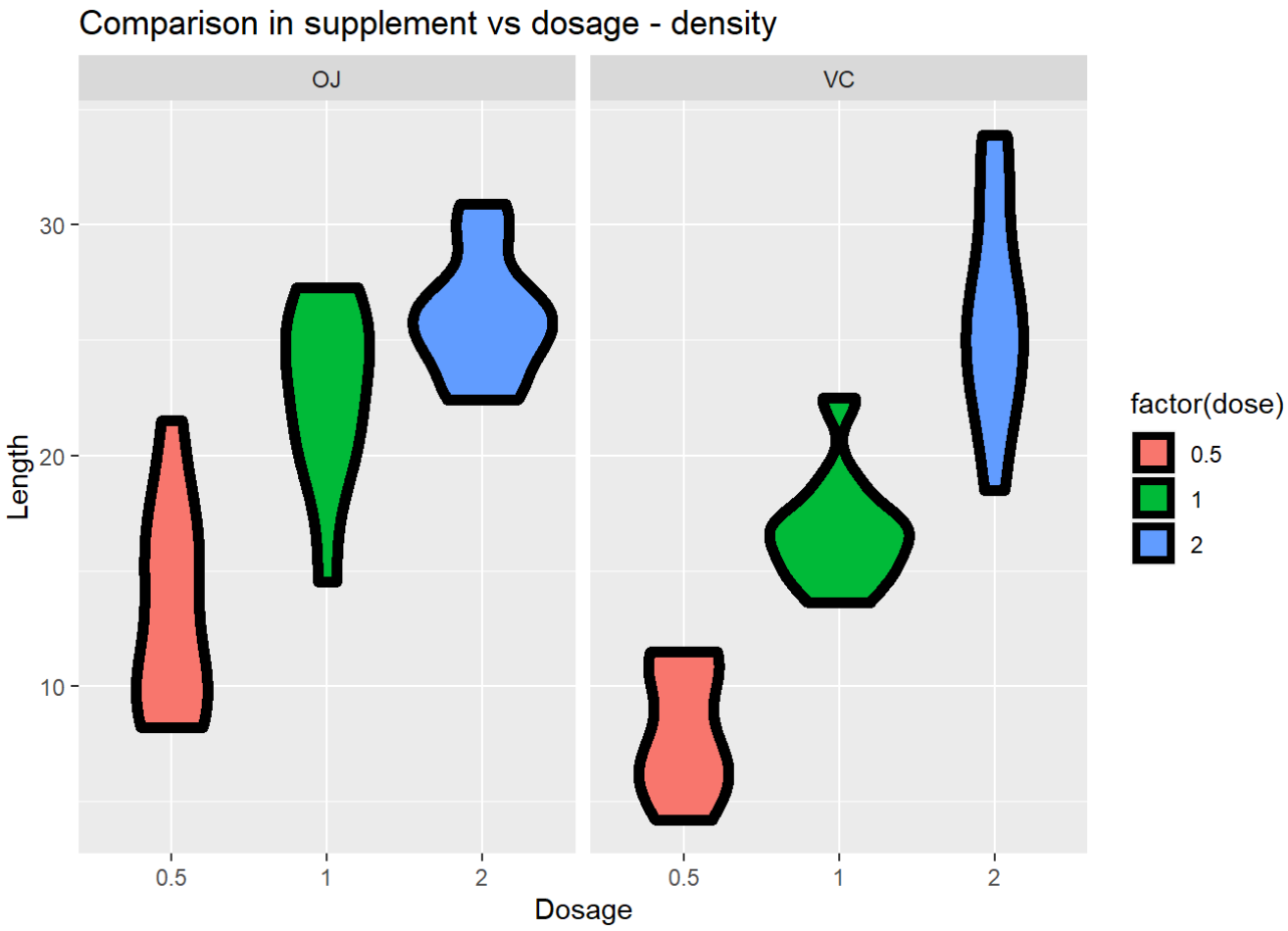
```
g <- ggplot(ToothGrowth, aes(x = dose, y = len, colour = dose, group = supp))
g <- g + geom_point()
g <- g + geom_line(tooth_mns, mapping = aes(x = dose,y = mean),linetype = 2, size = 1, color = "red")
g <- g + facet_grid(. ~ supp) + labs(x = "Dosage", y = "Length" , title = "Comparison in supplement vs dosage")
print(g)
```



The plot here shows the comparison between the 2 supplement types with respect to the level of dosage. From the plot we can infer that in either case of supplement the dosage with value 2.0 has a higher mean. Though the difference in means between 1.0 and 2.0 in case of supplement Orange /juice is very less. Hence the increase in dosage has not made much impact. Whereas in case of supplement VC, the dosage levels show a significant improvement in length.

```
g <- ggplot(ToothGrowth, aes(x = factor(dose), y = len, fill = factor(dose)))
g <- g + geom_violin(col = "black", size = 2)
g <- g + facet_grid(. ~ supp) + labs(x = "Dosage", y = "Length" , title = "Comparison in supplement vs dosage - d
```

```
ensity")
print(g)
```



The violin plots shows that the net effect dosage of VC with value 2.0 has a higher limit. Yet he results of supplement Orange Juice has a comparatively steady outcome. with varying dosages. ## Data Analysis The dosage comparisions in data w.r.t dosage of 0.5 and 1 in case of either supplements.

```
t.test(len[ToothGrowth$dose==0.5],len[ToothGrowth$dose==1], paired = FALSE, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len[ToothGrowth$dose == 0.5] and len[ToothGrowth$dose == 1]
## 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
```

Similarly the statistics of dosages 0.5 and 2.

```
t.test(len[ToothGrowth$dose==0.5],len[ToothGrowth$dose==2], paired = FALSE, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len[ToothGrowth$dose == 0.5] and len[ToothGrowth$dose == 2]
## 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
```

The statistics of dosages 1 and 2.

```
t.test(len[ToothGrowth$dose==1],len[ToothGrowth$dose==2], paired = FALSE, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
```

```
##
## data:  len[ToothGrowth$dose == 1] and len[ToothGrowth$dose == 2]
## 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
```

Conclusion

We can conclude that the Vitamin C for Guinea pigs can be more extracted through orange juice than when compared to ascorbic acid even at an average dosage. Whereas, ascorbic acid has to given in comparatively higher dosages in order to obtain right results.