Tooth Growth dataset

Kobra

15/08/2019

Tooth growth data analysis

Synopsis

We are going to analyze the tooth growth data and investigate whether different doses of vitamin C have effect on the tooth growth. In this dataset, the response 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).

1. Basic exploratory data analysis

Loading required packages if necessary:

```
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)
```

Loading ToothGrowth data sets:

```
mydata <- ToothGrowth
```

```
str (mydata)
```

```
## '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 (mydata)
```

```
##    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
```

Converting dose levels to fctor variable:

```
mydata$dose <- as.factor (mydata$dose)
```

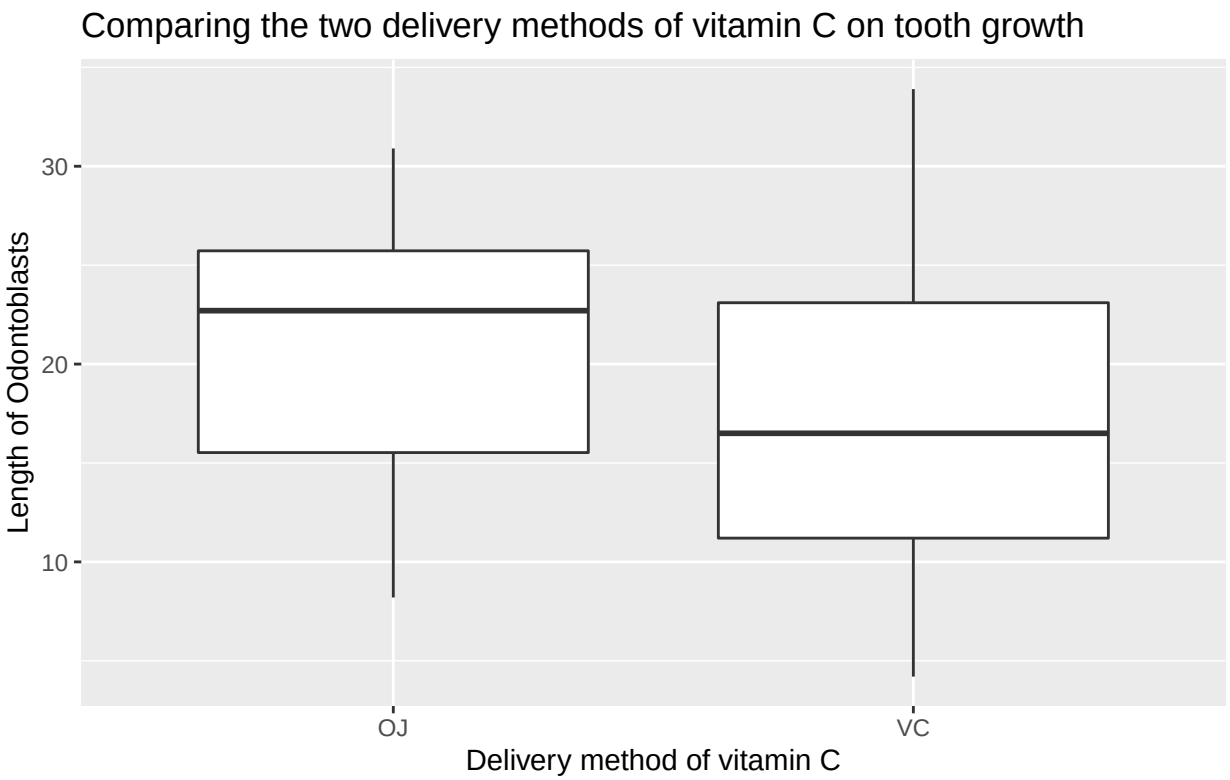
Providing a summary of data file:

```
summary (mydata)
```

```
##       len        supp     dose
##  Min.   : 4.20   OJ:30   0.5:20
##  1st Qu.:13.07   VC:30   1  :20
##  Median :19.25           2  :20
##  Mean   :18.81
##  3rd Qu.:25.27
##  Max.   :33.90
```

Comparing the two methods of delivery:

```
g1 <- ggplot (data= mydata, aes (x=supp, y= len))
g1 <- g1 + geom_boxplot()
g1 <- g1 + labs (x= "Delivery method of vitamin C", y= "Length of Odontoblasts", title = "Comparing the
g1
```



The effect of delivery of vitamin C via orange juice (OJ) versus ascorbic acid (VC) on the length of odontoblas
The results show no significant difference between the two groups.(p = 0.06)

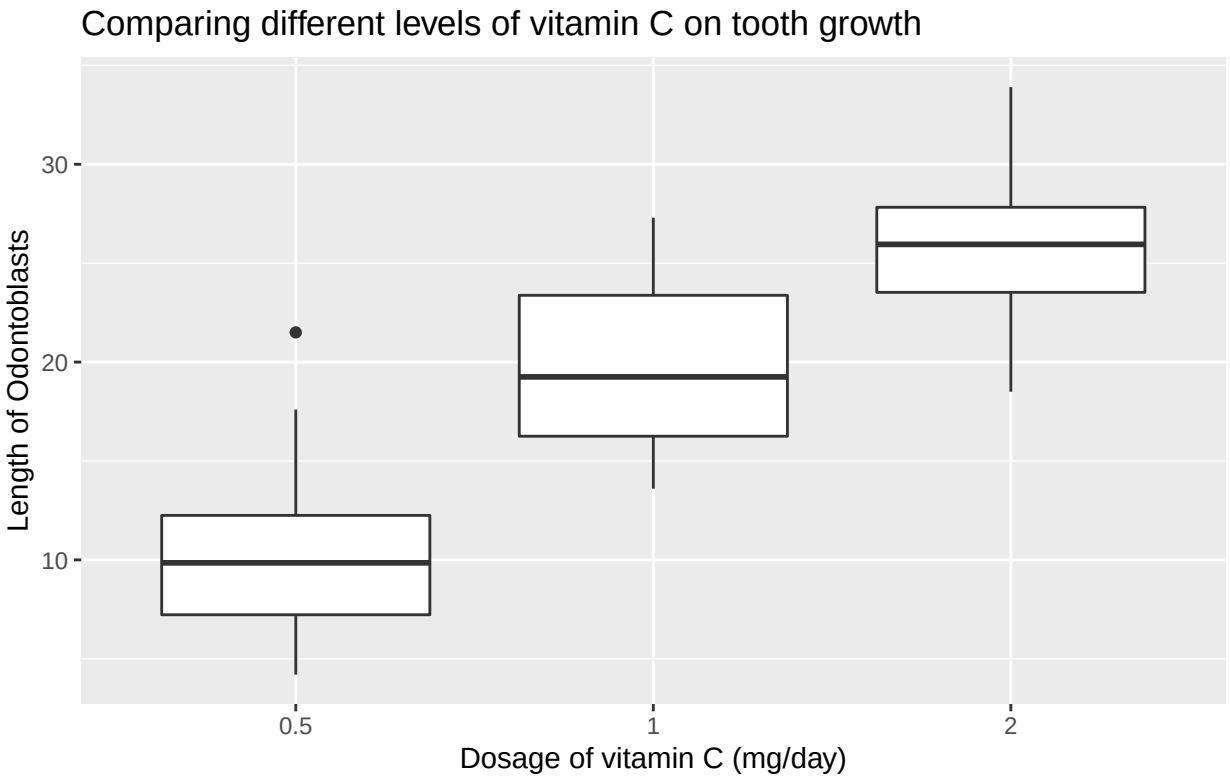
Hypothesis testing: Is there any significant differece between the two method of delivery?

```
t.test(len ~ supp, data = mydata)$p.value
```

```
## [1] 0.06063451
```

Comparing the three different doses of vitamin C:

```
g2 <- ggplot (data= mydata, aes (x=dose, y= len))
g2 <- g2 + geom_boxplot()
g2 <- g2 + labs (x= "Dosage of vitamin C (mg/day)", y= "Length of Odontoblasts", title = "Comparing dif
g2
```



The effect of three different dosage of vitamin C (0.5, 1 and 2 mg/day) on the length of odontoblasts. The results suggests the all three groups are significantly different from each other. (pvlaue <0.05).

Hypothesis testing:

Is there any significant differece between the dose groups?

```
with (mydata, t.test (len [dose == 0.5], len [dose == 1]))$p.value
```

```
## [1] 1.268301e-07
```

```
with (mydata, t.test (len [dose == 0.5], len [dose == 2]))$p.value
```

```
## [1] 4.397525e-14
```

```
with (mydata, t.test (len [dose == 1], len [dose == 2]))$p.value
```

```
## [1] 1.90643e-05
```

#Comparing the mean of lenghts for different doses via each method of delivery:

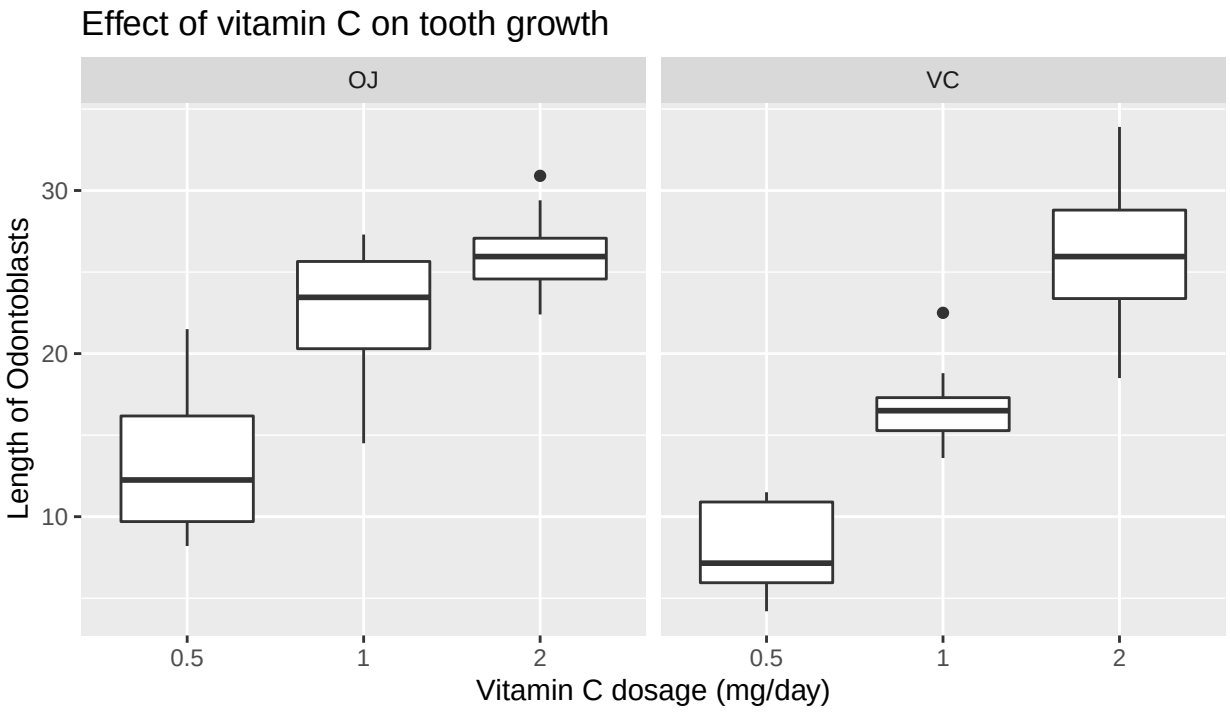
```
grouped_data <- mydata %>% group_by(supp, dose)
averages <- grouped_data %>% summarise(mean = mean (len))
averages
```

```
## # A tibble: 6 x 3
## # Groups:   supp [2]
```

```
##   supp  dose   mean
##   <fct> <fct> <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
```

Graphing lenghts based on dose and delivery method::

```
g3 <- ggplot (data= mydata, aes (x=dose, y= len))
g3 <- g3 + geom_boxplot()+ facet_grid(.~ supp)
g3 <- g3 + labs (x= "Vitamin C dosage (mg/day)", y= "Length of Odontoblasts", title = "Effect of vitami
g3
```



The effect of three dose levels of vitamin C (0.5, 1, and 2 mg/day) on the length of odontoblasts by one of two delivery methods, orange juice (OJ) or ascorbic acid (VC).
The results indicate significantly higher odontoblasts lenght for OJ group at lower dosage (0.5 and 1 mg/day) (pvalue < 0.05) while there was no significant difference between the two groups at 2 mg/day (p= 0.96).

Hypothesis testing:

Is there any significant differnce between the two methos at each level of vitamin C dosage?

```
mydata_0.5 <- subset(mydata, mydata$dose==.5)
mydata_1 <- subset(mydata, mydata$dose==1)
mydata_2 <- subset(mydata, mydata$dose==2)

t_0.5 <- t.test(len~supp, data=mydata_0.5)$p.value
t_1 <- t.test(len~supp, data=mydata_1)$p.value
```

```r
t_2 <- t.test(len~supp, data=mydata_2)$p.value

t_summary <- c ("dose= 0.5"= t_0.5, "dose= 1"= t_1, "dose= 2"= t_2)
t_summary
```

```
##   dose= 0.5     dose= 1     dose= 2
## 0.006358607 0.001038376 0.963851589
```

Conclusion

The results suggest that the deliery method itself will not make any significant differnce on the tooth growth. However, higher dose levels, depsite the different delivery method, will significantly increase the tooth growth. By comparing the lenghts at each dose level, we observed that the lenghts are significantly higher at 0.5 and 1 mg/day for orange juice groupe while no sognificant difference was observed at the dose of 2 mg/day between orange juice and ascorbic acid groups.

Since we were implenting three t tests for each hypothesis testing, we would need to correct the family wise error uremain vaild.