Analysis of ToothGroth database

koji

2022/1/14

load ToothGrowth data and shows the summary

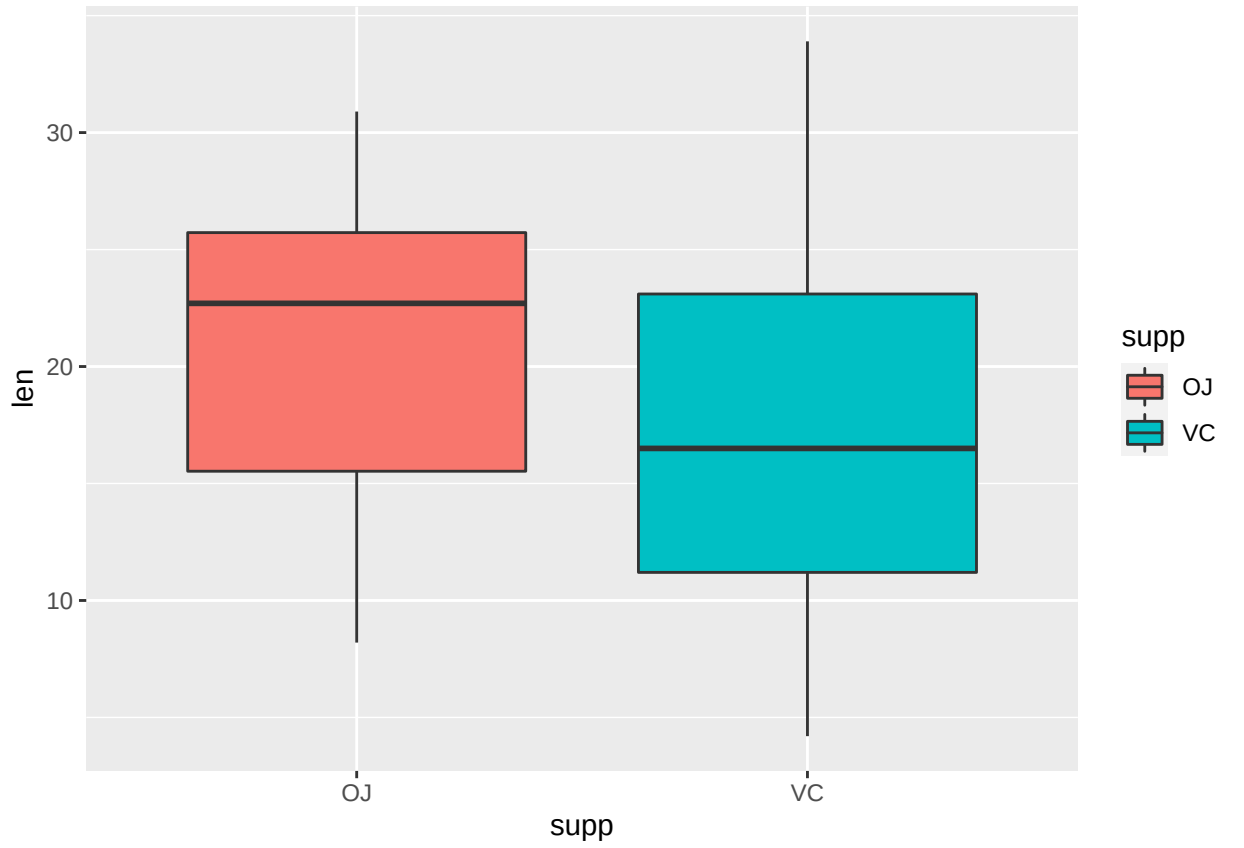
```
data("ToothGrowth")
head(ToothGrowth); summary(ToothGrowth); str(ToothGrowth)
```

```
##    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
```

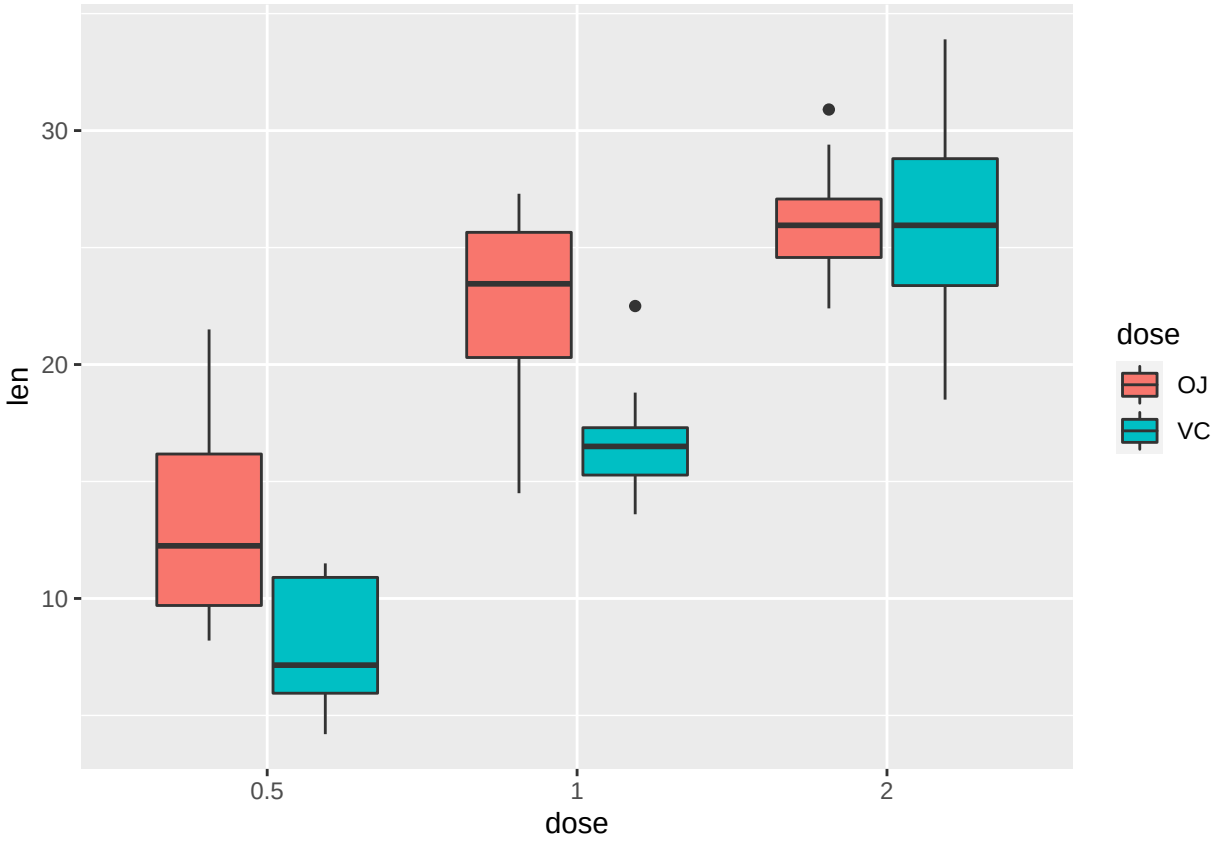
```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

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

```
p1 <- ggplot(ToothGrowth, aes(supp, len, fill = supp)) + geom_boxplot()
p2 <- ggplot(ToothGrowth, aes(factor(dose), len, fill = dose)) + geom_boxplot(aes(fill = supp)) + labs(
p1
```




```
p2
```



t-test in between supplements

```
OJ_tooth <- ToothGrowth$len[ToothGrowth$supp == "OJ"]
VC_tooth <- ToothGrowth$len[ToothGrowth$supp == "VC"]
t.test(OJ_tooth, VC_tooth)
```

```
##
##  Welch Two Sample t-test
##
## data:  OJ_tooth and VC_tooth
## 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
```

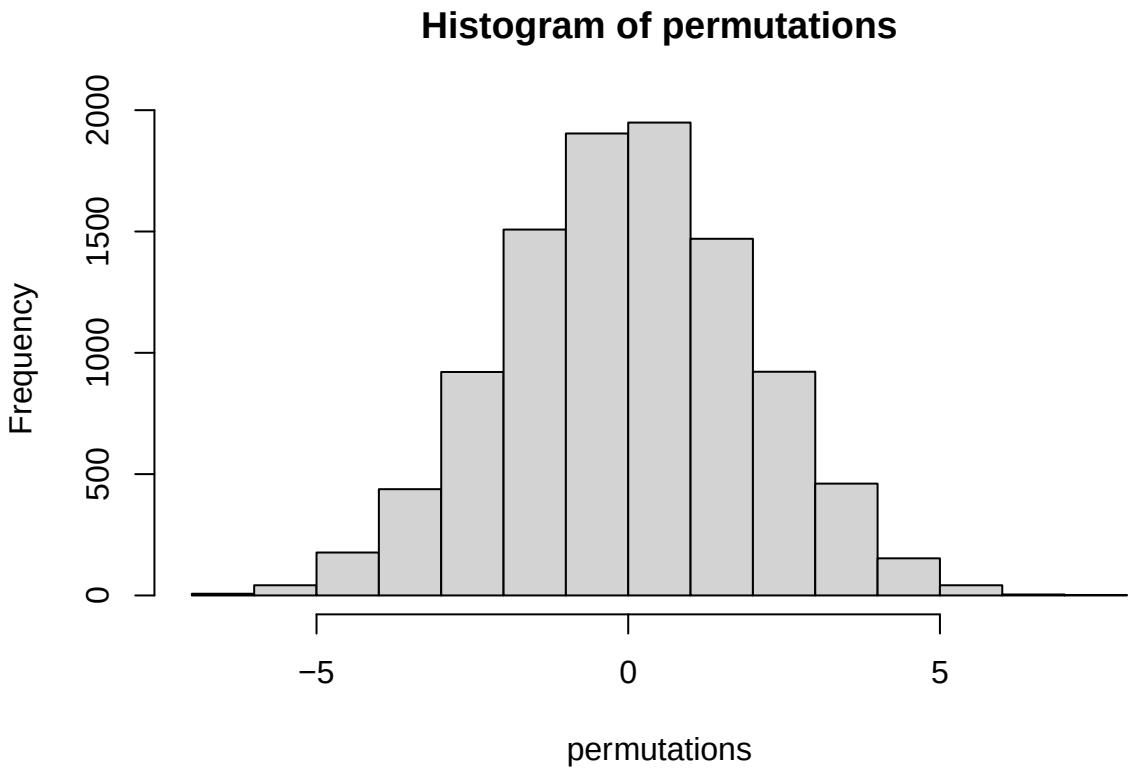
Since the P-value is greater than 0.05, we can not reject Null hypothesis with 95% confidence.

Then how about in permutation test?

Permutation test

```
group <- as.character(ToothGrowth$supp)
testStat <- function(w,g) mean(w[g == "OJ"]) - mean(w[g == "VC"])
observedStat <- testStat(ToothGrowth$len, group)
permutations <- sapply(1: 10000, function(i) testStat(ToothGrowth$len,sample(group)))
```

```
hist(permutations)
```




```
observedStat
```

```
## [1] 3.7
```

```
mean(permutations > observedStat)
```

```
## [1] 0.0299
```

Since this estimate of the P-value is very low and we can reject the NULL with 95% confidence interval

comparison by dosage

we want to reject the probability of rejecting under null to be 5%, thus we reject if our test statistic is larger that qt(.975,37)=2.026 or smaller than qt(0.25,37)=-0.681

```
t.test(ToothGrowth$len[ToothGrowth$dose == 0.5], ToothGrowth$len[ToothGrowth$dose == 1.0])
```

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

```
t.test(ToothGrowth$len[ToothGrowth$dose == 0.5], ToothGrowth$len[ToothGrowth$dose == 2.0])
```

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

```
t.test(ToothGrowth$len[ToothGrowth$dose == 1.0], ToothGrowth$len[ToothGrowth$dose == 2.0])
```

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

As we can see the each t-statistic is quite smaller than qt(0.25,df) = -0.681, thus we reject null and there are no significant differences between dose.

Conclusion

We conducted t-test and permutation test for tooth growth by supplement or dosage with 95% confidence. From the result of t-test, there was no difference in between OJ and VC. But there was significant difference in permutation test in contrast of t-test. In terms of the difference by dosage, there were significant differences for all possible combinations of dosage level.