

Statistical Inference Course Project 2

Marta Regolo

Part 2: Basic Inferential Data Analysis

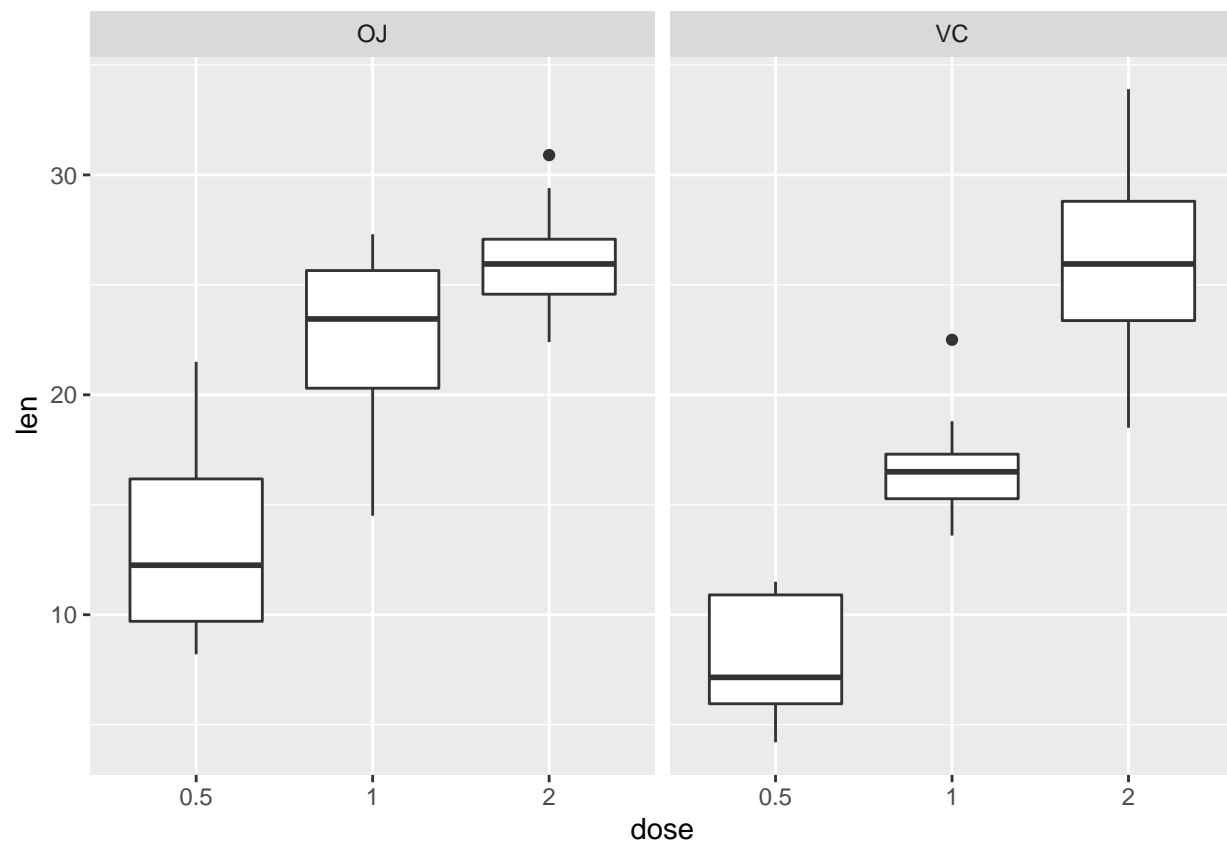
In the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

First, we load the ToothGrowth data and we perform some basic exploratory data analyses. 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).

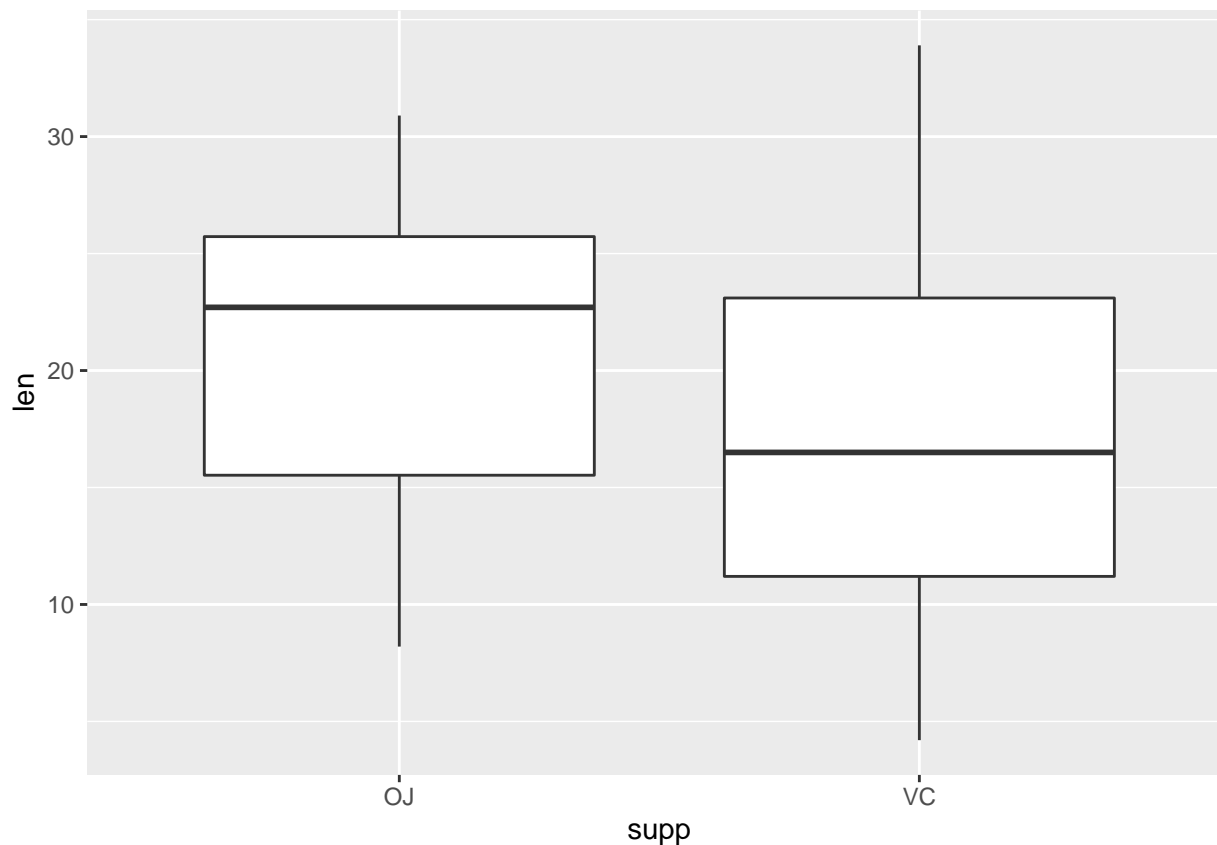
```
data(ToothGrowth)
```

```
summary(ToothGrowth)
```

```
##      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
```



Secondly, we use hypothesis tests to compare growth by supp.



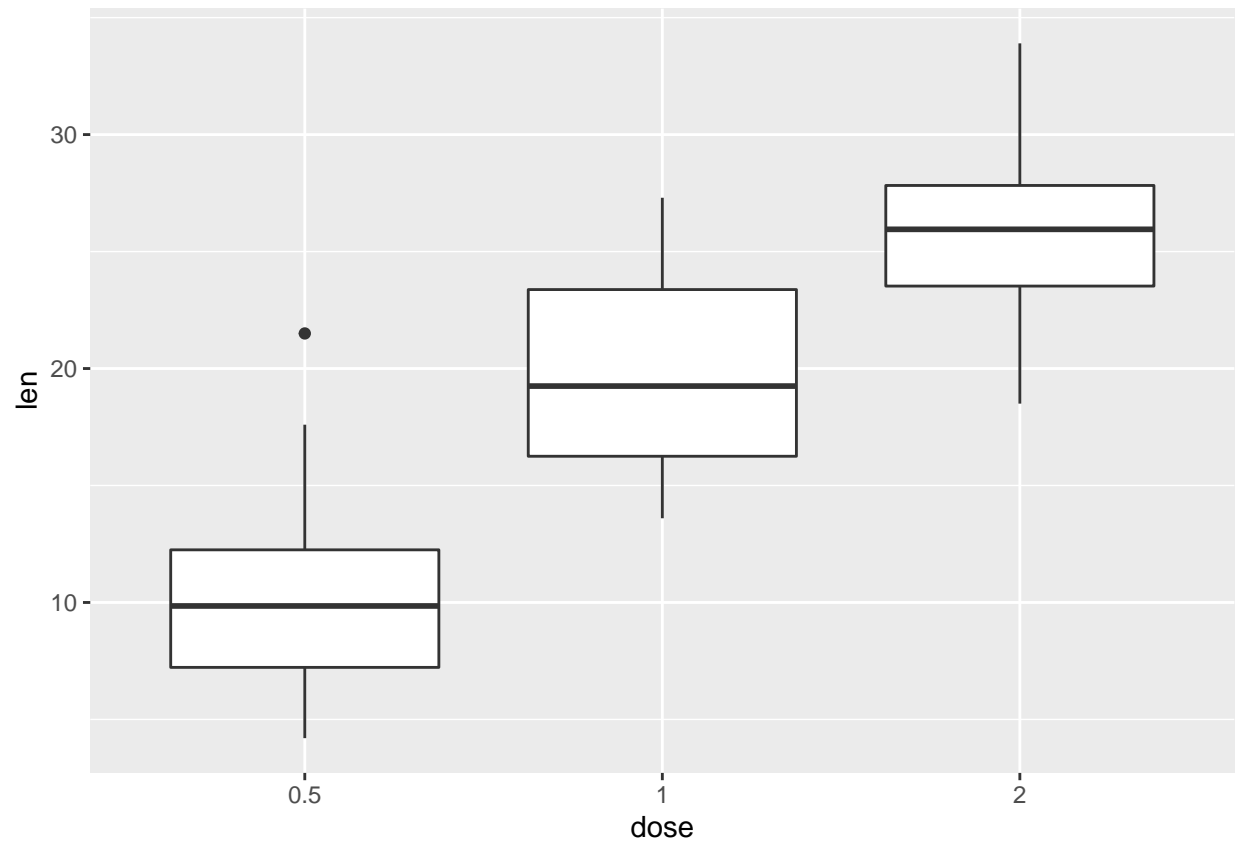
We perform a t-test for the two independent group of observations.

```
t.test(ToothGrowth[ToothGrowth$supp=="VC"],$len,ToothGrowth[ToothGrowth$supp=="OJ"],$len,paired=FALSE,v

##
##  Welch Two Sample t-test
##
## data:  ToothGrowth[ToothGrowth$supp == "VC", ]$len and ToothGrowth[ToothGrowth$supp == "OJ", ]$len
## 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 of x mean of y
##  16.96333  20.66333
```

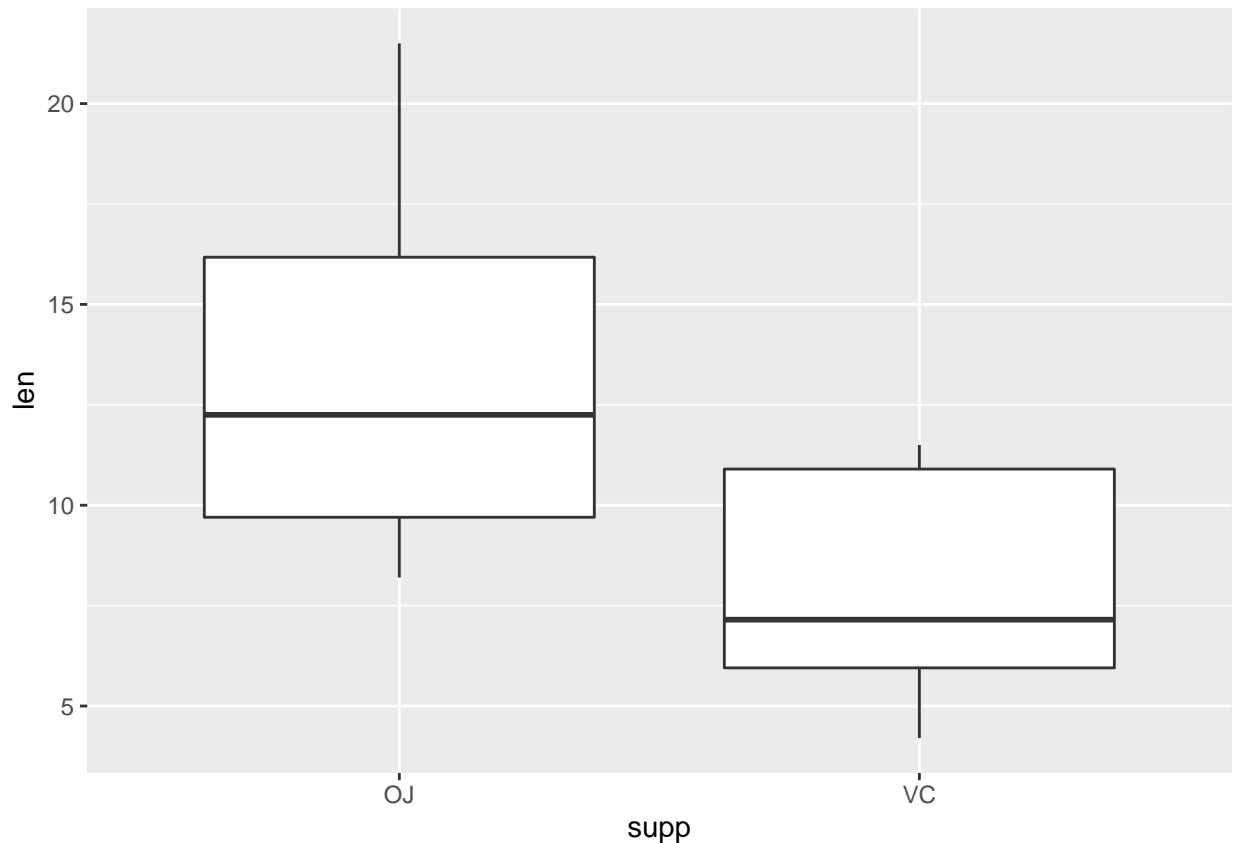
We can see that the p-value is slightly above 0.05, so we are in the acceptance region and we can't reject the null hypothesis. So we accept that the difference in means due to two delivery methods is equal to zero.

Next, we analyze the difference due to the three dose levels of vitamin C (0.5, 1, and 2 mg/day).



We can see that there is a significant difference between the three dose levels.

So, we try to see if there is a significant difference in means due to the delivery method within each dose level.



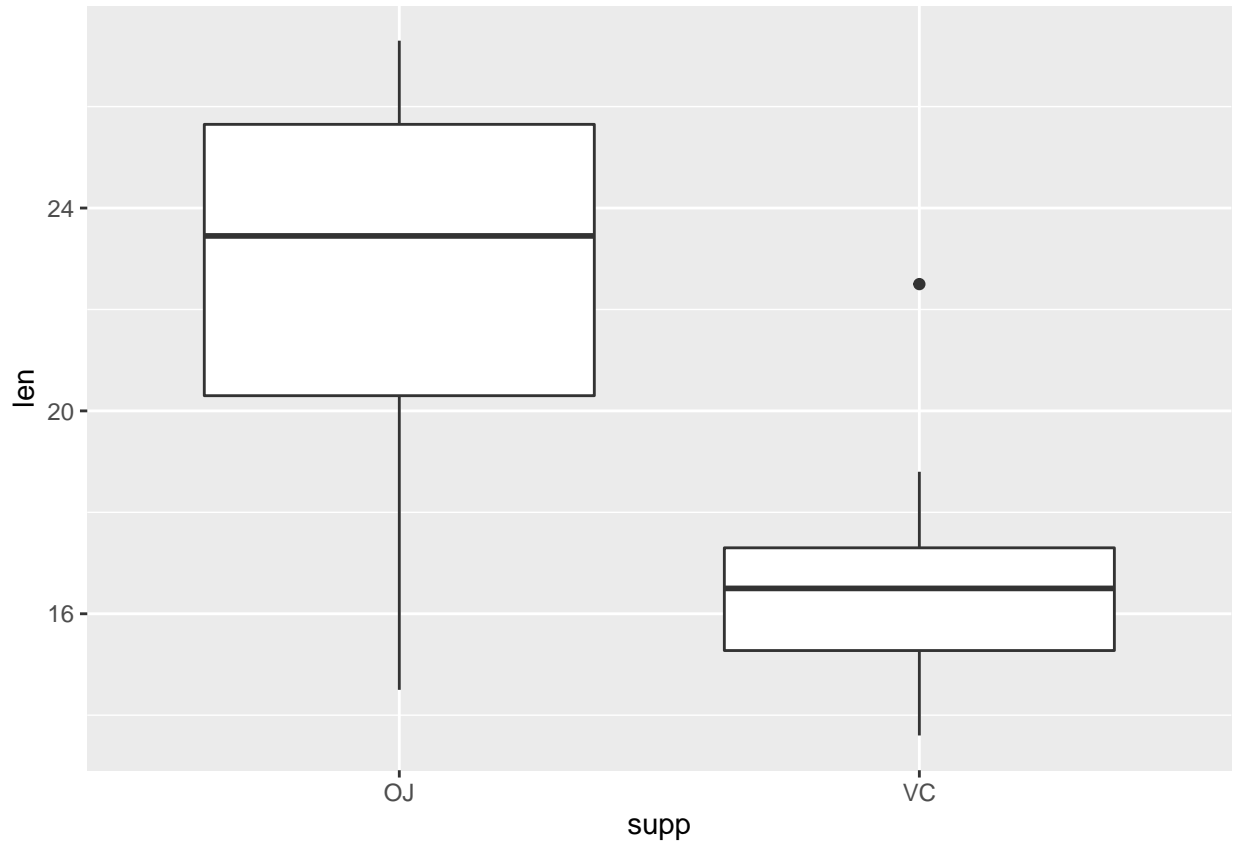
We perform a t-test for the two independent group of observations.

```
t.test(dose05[dose05$supp=="OJ"],$len,dose05[dose05$supp=="VC"],$len,paired=FALSE,var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose05[dose05$supp == "OJ", ]$len and dose05[dose05$supp == "VC", ]$len
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean of x mean of y
##    13.23    7.98
```

We fail to reject the null hypothesis, so the difference in mean is equal to 0.

Dose Level = 1.0



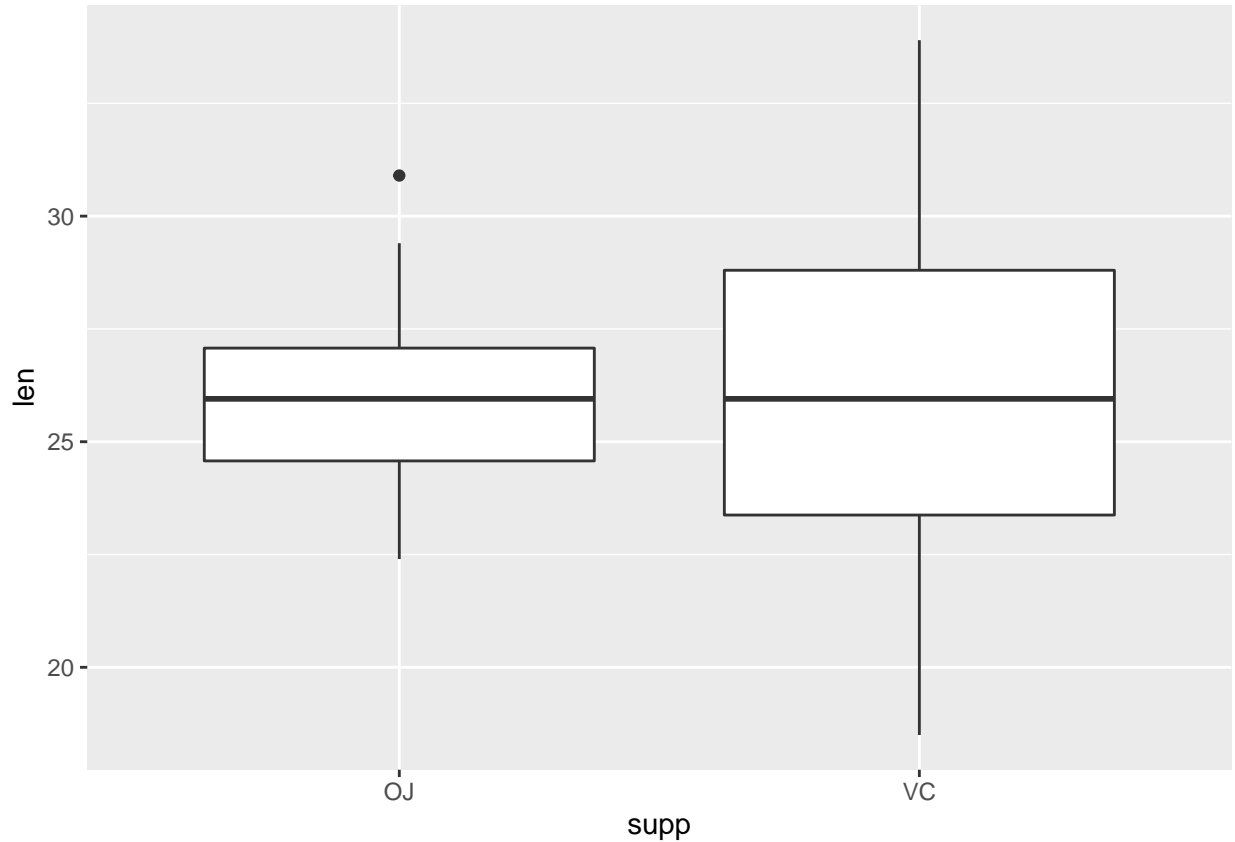
We perform a t-test for the two independent group of observations.

```
t.test(dose1[dose1$supp=="OJ"],dose1[dose1$supp=="VC"],paired=FALSE,var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  dose1[dose1$supp == "OJ", ]$len and dose1[dose1$supp == "VC", ]$len
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean of x mean of y
##    22.70    16.77
```

We reject the null hypothesis, so the difference in mean is different from 0.

Dose Level = 2



We perform a t-test for the two independent group of observations.

```
t.test(dose2[dose2$supp=="OJ"],]$len,dose2[dose2$supp=="VC"],]$len,paired=FALSE,var.equal=FALSE)
```

```
##
## Welch Two Sample t-test
##
## data: dose2[dose2$supp == "OJ", ]$len and dose2[dose2$supp == "VC", ]$len
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.79807 3.63807
## sample estimates:
## mean of x mean of y
## 26.06 26.14
```

We fail to reject the null hypothesis, so the difference in mean is equal to 0.

Conclusions

Overall, we can conclude that the methods used to deliver the vitamin C don't affect the growth of odotoblasts. There's only one case where we can see a significant difference in means between the two methods, and it's when the dose level of vitamin C is 1 mg/day.