Statistical Inference Course Project - Part 2

Anthony Macey

18 November 2017

The Effect of Vitamin C on Tooth Growth in Guinea Pigs

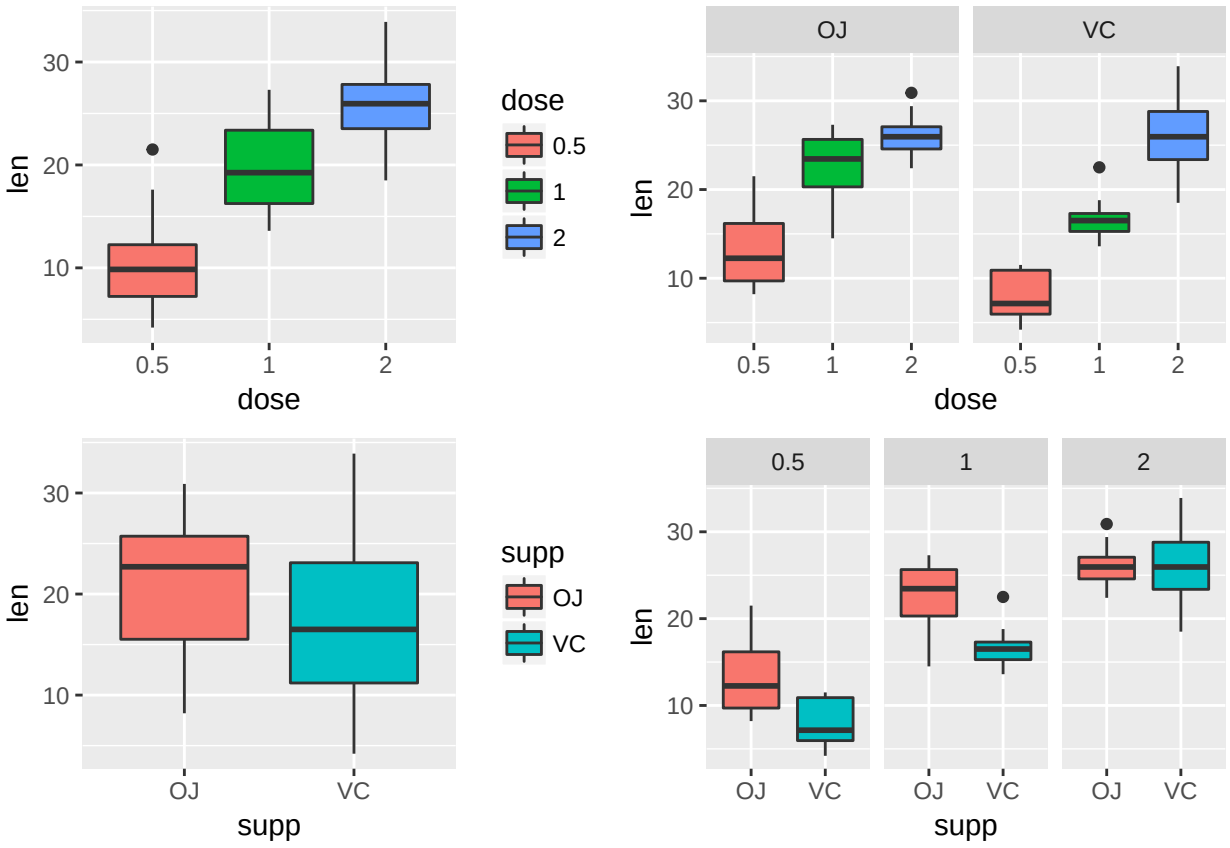
Now in the second portion of the project, we're going to analyze the ToothGrowth data in the R datasets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses
2. Provide a basic summary of the data.
3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use the techniques from class, even if there's other approaches worth considering)
4. State your conclusions and the assumptions needed for your conclusions.

Load the data

Basic analyis of the data

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



Test 1

Using `t.test` to check if there is a difference in the performance of the treatments across all doses.

```
t.test(len ~ supp,
       data = ToothGrowth,
       paired = FALSE,
       var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## 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 in group OJ mean in group VC
##         20.66333         16.96333
```

The p-value is 0.06, and the confidence interval contains zero. This indicates that we can not reject the null hypothesis that the different supplement types have no effect on tooth length.

Test 2

Using `t.test` to check if there is a difference in the performance of the treatments in different dose pairs.

```
t.test(
  len ~ supp,
  data = subset (ToothGrowth, dose %in% c(0.5, 1.0)),
  paired = FALSE,
  var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.0503, df = 36.553, p-value = 0.004239
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.875234 9.304766
## sample estimates:
## mean in group OJ mean in group VC
##           17.965           12.375
```

```
t.test(
  len ~ supp,
  data = subset (ToothGrowth, dose %in% c(0.5, 2.0)) ,
  paired = FALSE,
  var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 0.92163, df = 35.105, p-value = 0.363
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.10849  8.27849
## sample estimates:
## mean in group OJ mean in group VC
##           19.645           17.060
```

```
t.test(
  len ~ supp,
  data = subset (ToothGrowth, dose %in% c(1.0, 2.0)),
  paired = FALSE,
  var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.8397, df = 31.273, p-value = 0.07533
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.3166175  6.1666175
## sample estimates:
## mean in group OJ mean in group VC
```

```
##           24.380           21.455
```

For all three dose level pairs, the p-value is less than 0.05, and the confidence interval does not contain zero. The mean tooth length increases on raising the dose level and this indicates that we can reject the null hypothesis, and establish that increasing the dose level leads to an increase in tooth length.

Conclusion

1. Supplement type has no effect on tooth growth.
2. Increasing the dosage level leads to increased tooth growth.

Assumptions

1. The experiment was done with random assignment of guinea pigs.
2. The guinea pigs, are representative of the entire population of guinea pigs.