Inference Assignment 1b

Moritz Schneider

Wednesday, December 17, 2014

Assignment

Now in the second portion of the class, we're going to analyze the Tooth Growth data in the R data sets package.

1. Load the ToothGrowth data and perform some basic exploratory data analyses

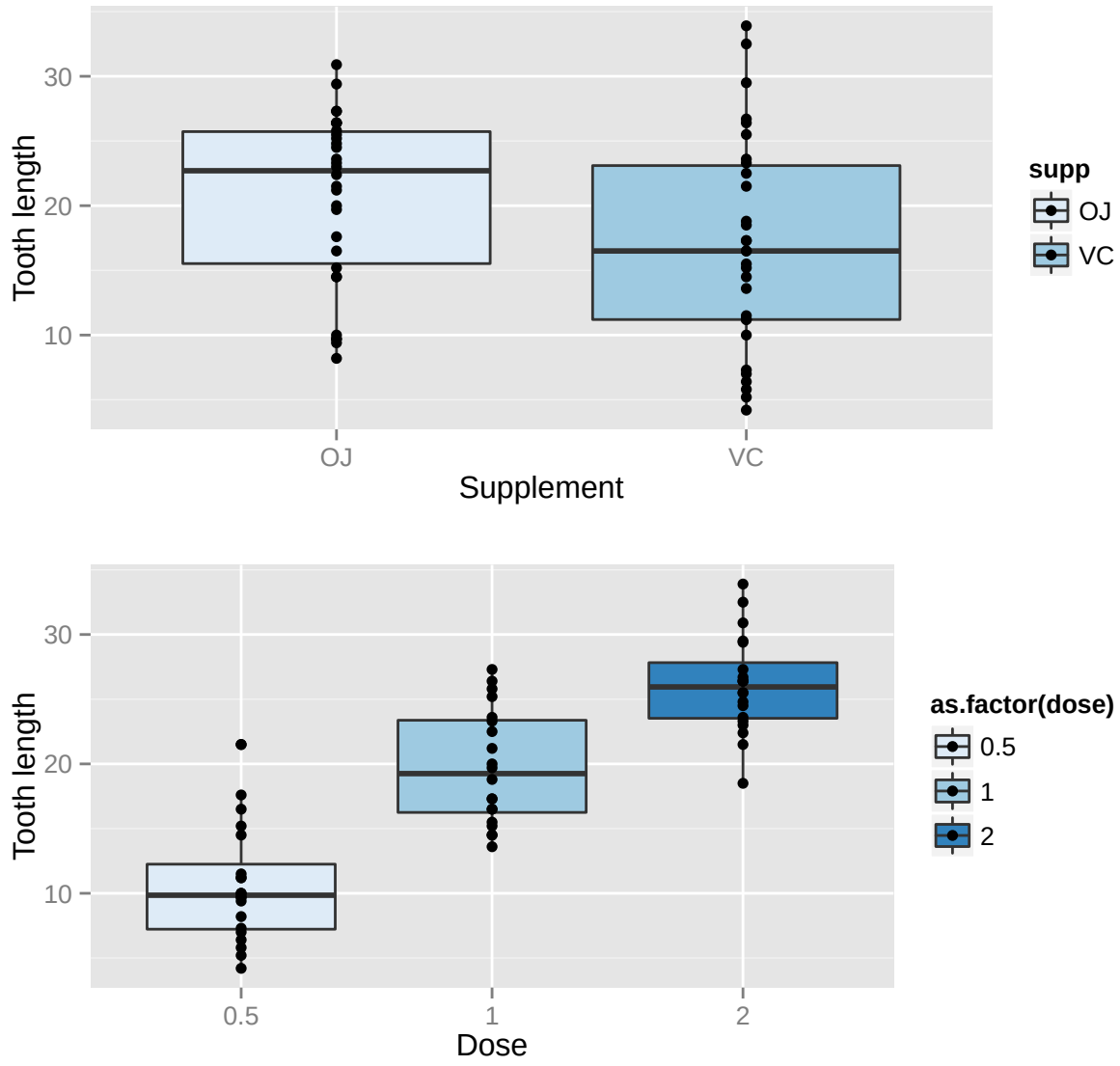
2. Provide a basic summary of the data.

3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose. (Only use

4. State your conclusions and the assumptions needed for your conclusions.

Part 1

For exploratory data analysis, let's take look at some box plots:



In the first plot we can see, that the supplement type, OJ and VC, has an effect on tooth length as OJ supplement seems to lead to increased tooth length. A pattern arising from different dose levels is very obvious in the second box plot, where the boxes resolve tooth length by dose levels. Higher doses seem to lead to increased tooth length.

Part 2

In order to summarize the data in a meaningful way we have to identify the subgroups of the data. There are two types of supplements and each has 10 samples for three different doses. This makes for a total of 6 groups, for which the means and standard deviations of the tooth length can be found in the following table:

	VC	OJ
Dose = 0.5	7.98 +- 2.75	13.23 +- 4.46
Dose = 1	16.77 +- 2.52	22.7 +- 3.91
Dose = 2	26.14 +- 4.8	26.06 +- 2.66

Part 3

Using only the techniques we learned in the lessons, in the following we will perform multiple t-tests to compare tooth growth by supplement and dose. To do this, we will compare the 6 subgroups introduced in part 2. In total, we will perform 8 t-test, so we have to account for multiple testing and will use the method by Bonferroni. This means for a test result to be considered significant the p-vale will have to be less than 0.05/7, i.e approx 0.00714, therefore we use a 99.286% confidence interval. First we will compare the supplements OJ and VC for each dose:

	P-value	Lower Bound	Upper Bound
Dose = 0.5	0.0063586	-10.406039	-0.0939613
Dose = 1	0.0010384	-10.489682	-1.3703184
Dose = 2	0.9638516	-5.373596	5.5335965

The table tells us, that even using the conservative method by Bonferroni to account for multiple testing, there is still a significant difference in tooth length for the lowest dose, as supplement VC leass to less tooth growth than supplement OJ. However, it is noteworthy the upper bound of our adjusted confidence interval is very close to zero. A dose of 1 results in a bigger difference between VC and OJ as can be seen by the lower p-vale and also a decreased upper bound (i.e. less close to zero). Surprisingly, when a dose of 2 is used the supplement type does not seem to have any effect on tooth growth at all, as the p-vale is close to 1 and our confidence interval is almost symmetric round zero.

Next we will take a look at how dose levels affect tooth growth for each supplement.

```
OJ0.5_1 <- t.test(ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5,]$len,
                  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1,]$len,
                  conf.level = 1-0.00714)
OJ1_2 <- t.test(ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1,]$len,
                  ToothGrowth[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2,]$len,
                  conf.level = 1-0.00714)
VC0.5_1 <- t.test(ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 0.5,]$len,
                  ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 1,]$len,
                  conf.level = 1-0.00714)
VC1_2 <- t.test(ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 1,]$len,
                  ToothGrowth[ToothGrowth$supp == "VC" & ToothGrowth$dose == 2,]$len,
                  conf.level = 1-0.00714)
```

```r
output <- rbind(c(OJ0.5_1$p.value, OJ0.5_1$conf.int),
                c(OJ1_2$p.value, OJ1_2$conf.int),
                c(VC0.5_1$p.value, VC0.5_1$conf.int),
                c(VC1_2$p.value, VC1_2$conf.int))
colnames(output) <- c("P-value", "Lower Bound", "Upper Bound")
rownames(output) <- c("OJ: 0.5 vs 1", "OJ: 1 vs 2", "VC: 0.5 vs 1", "VC: 1 vs 2")
kable(output)
```

	P-value	Lower Bound	Upper Bound
OJ: 0.5 vs 1	0.0000878	-15.172941	-3.767059
OJ: 1 vs 2	0.0391951	-7.974036	1.254036
VC: 0.5 vs 1	0.0000007	-12.366524	-5.213476
VC: 1 vs 2	0.0000916	-14.786709	-3.953291

The table tells us, that there a significant increase in tooth length is to be expected when a) increasing dose from 0.5 to 1 using OJ, b) increasing dose from 0.5 to 1 using VC and c) increasing dose from 1 to 2 using VC. In all three cases the p-values are well below our adjusted significance level with upper bounds far from zero. The difference between dose 1 and 2 using OJ would be considered significant using an unadjusted significance level, however, the Bonferroni correction tells us that a p-value of 0.039 is not to be considered significant, accordingly, the adjusted confidence interval includes zero.

Part 4

In conclusion, the comparison between OJ and VC has shown that using low doses of 0.5 or 1 results to a significant difference in tooth growth between the two supplement types, suggesting to use OJ when trying to maximize tooth growth. When using a dose of 2 the supplement type does not seem to make any difference even if we had not adjusted the significance level to account for multiple test.

The comparison between doses has shown, that both supplements lead to significantly longer teeth when increasing dose from 0.5 to 1. However, only VC shows additional significant gain in tooth length from further increasing dose from 1 to 2, for OJ the increase in tooth length from dose 1 to 2 is to small to be considered significant when accounting for multiple tests using Bonferroni's method.