Statistical Inference: Tooth Growth

Mitch Murphy

4/17/2018

Tooth Growth

Summary of Data

Now let us take a look at the affect of two different supplements on tooth growth. First, let's get an high-level overview of the data:

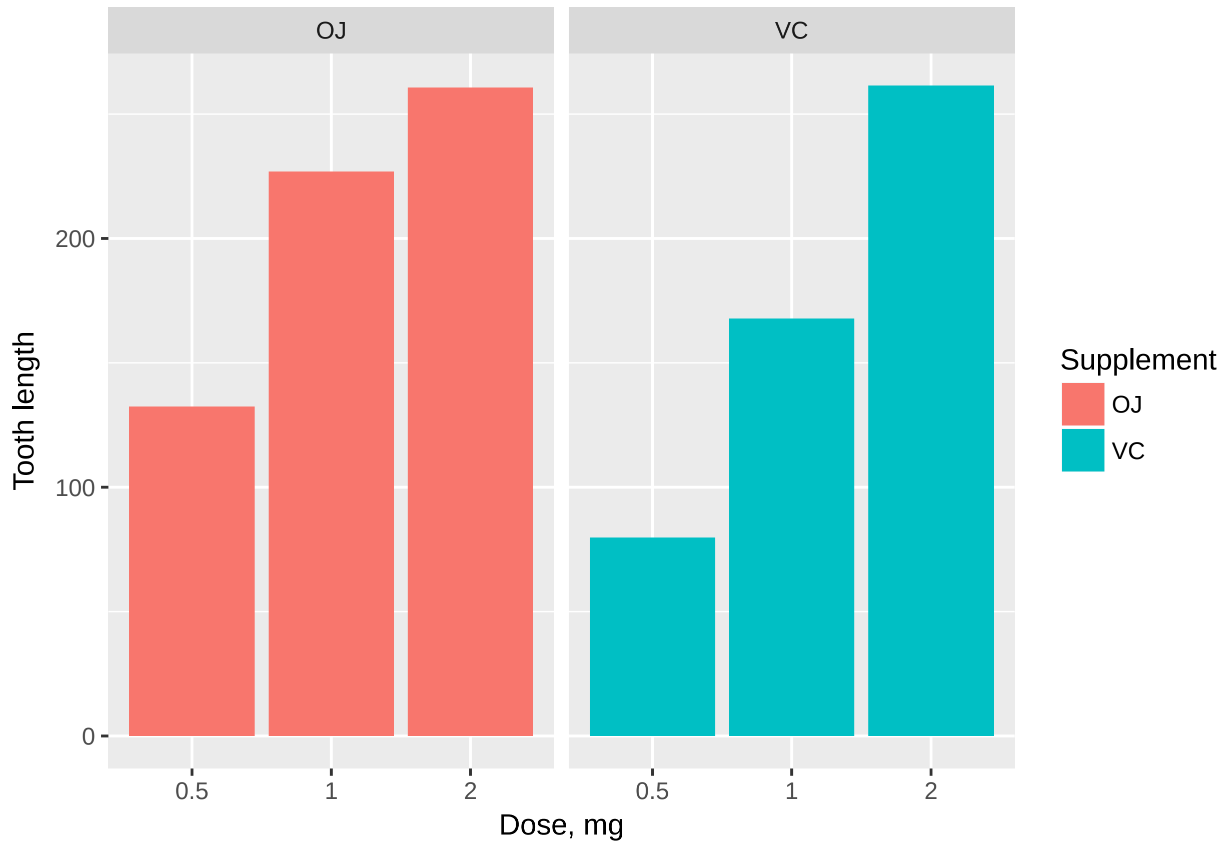
```
tg <- ToothGrowth
tg$dose <- as.factor(tg$dose)
str(tg)
```

```
## '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: Factor w/ 3 levels "0.5","1","2": 1 1 1 1 1 1 1 1 1 1 ...
```

```
summary(tg)
```

```
##       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
```

```
ggplot(data=tg, aes(x=dose, y=len, fill=supp)) + geom_bar(stat="identity",) + facet_grid(. ~ supp) +
  xlab("Dose, mg") + ylab("Tooth length") + guides(fill=guide_legend(title="Supplement"))
```



Confidence Intervals

Now let us construct 95% confidence intervals across the different doses.

```
ci <- t.test(len ~ supp, data = tg)
pval <- ci$p.value
ci
```

```
##
##  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
```

Since zero is contained in this interval, and because p-val$\geq \alpha = 0.05$ we cannot reject the null hypothesis. Let's further explore this data by creating dosage-groups, as so:

```
tg.dose1 <- subset(tg, dose %in% c(0.5,1.0))
tg_dose2 <- subset(tg, dose %in% c(0.5,2.0))
tg_dose3 <- subset(tg, dose %in% c(1.0,2.0))
```

```r
# construct the CI for dose of 0.5 and 1.0
t.test(len ~ dose, data = tg.dose1)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## 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 in group 0.5   mean in group 1
##            10.605            19.735
```

```r
#construct the CI for a dose of 0.5 and 2.0
t.test(len ~ dose, data = tg_dose2)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## 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 in group 0.5   mean in group 2
##            10.605            26.100
```

```r
# construct the CI for a dose of 1.0 and 2.0
t.test(len ~ dose, data = tg_dose3)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## 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 in group 1 mean in group 2
##          19.735          26.100
```

Because zero is not contained in any of the above confidence intervals and the p-value $\leq \alpha$ we can conclude that the dose of the supplement effects tooth length; specifically as the dose increases so does the tooth length. We have already concluded that the supplement type is not significant (first CI above).

Assumptions

When constructing confidence intervals (as we did above), we assumed that the variance of the 2 groups was different.