Tooth Growth Analysis

MiRo

22 Jul 2015

Excercise

We're going to analyze ToothGrowth data in the R datasets package. * Load the ToothGrowth data and perform some basic exploratory data analyses * Provide a basic summary of the data. * 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) * State your conclusions and the assumptions needed for your conclusions.

Data description

The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, and 2 mg) with each of two delivery methods (orange juice or ascorbic acid).

Loading and preparing data

```r
data("ToothGrowth")
summary(ToothGrowth)
```

```
##       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 discovery

Plotting ToothGrowth data to have a quick overview.

```r
gplot <- ggplot(ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
  geom_violin(scale = "area") +
  facet_grid(. ~ supp) +
  xlab("Dose in miligrams") +
  ylab("Tooth length") +
  guides(fill=guide_legend(title="Supplement type")) +
  scale_fill_brewer(palette="Set1") +
  theme_bw()
gplot
```
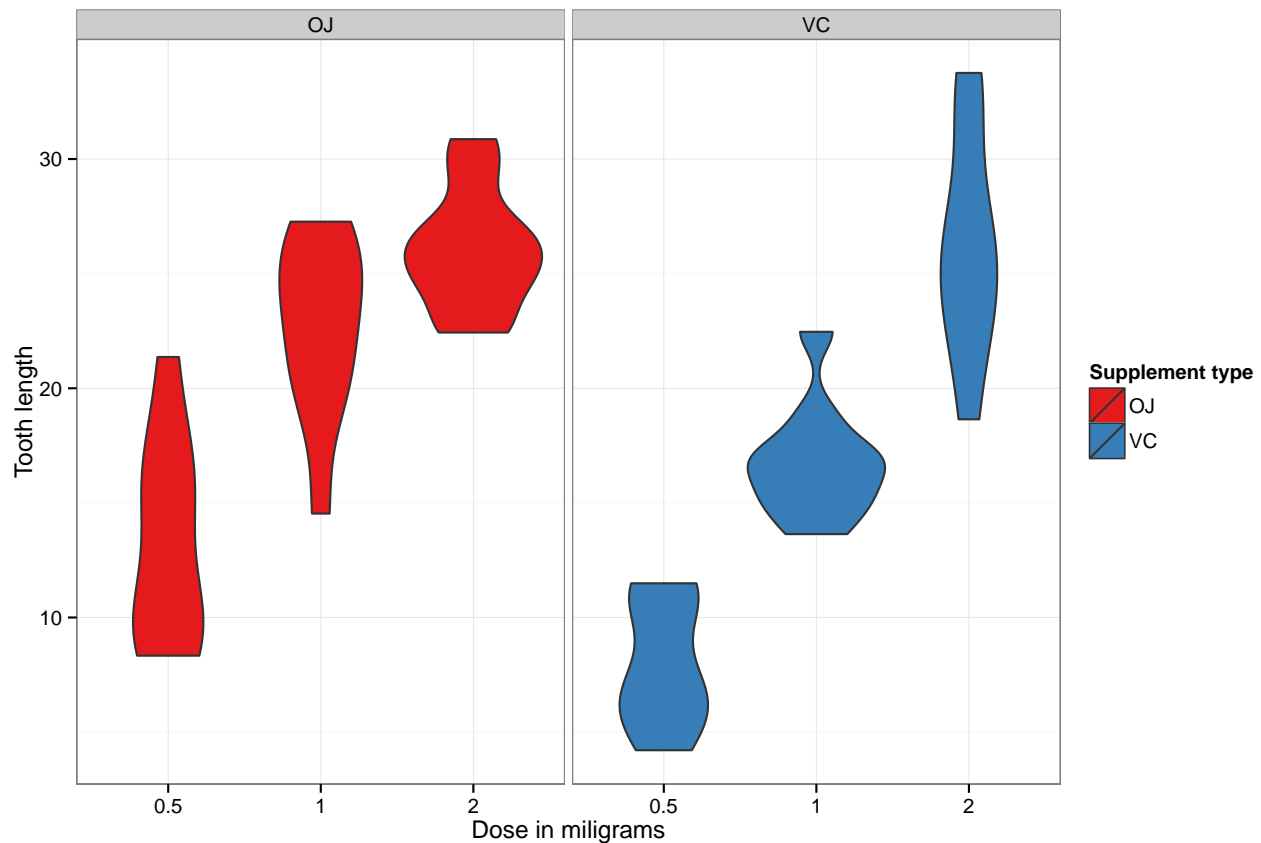
Since intervals are not intersect, we could say that there are singificant difference in tooth lenth depends on dosage. We could even build a linear model to discribe this dependence, but not in this time.

Testing delivery methods

Now our aim is to test how tooth lenght depends on dosage and delivery method. Is there any significant difference between these methods? For every dosage we will take a t-test to compare.

```
meansDT <- ToothGrowth  %>% group_by(supp, dose) %>%
  summarise(means=mean(len)) %>% select(means, supp, dose)
gplot <- ggplot(ToothGrowth, aes(x=len, fill=supp)) +
  geom_density(kernel="gaussian") +
  facet_grid(. ~ dose) +
  geom_vline(aes(xintercept=meansDT$means), meansDT[meansDT$supp=="VC",],
             colour="red", linetype = "longdash", binwidth=0.2) +
  geom_vline(aes(xintercept=meansDT$means), meansDT[meansDT$supp=="OJ",],
             colour="blue", linetype = "longdash", binwidth=0.2) +
  xlab("Driver") +
  ylab("Freq") +
  guides(fill=guide_legend(title="Supplement type")) +
  scale_fill_brewer(palette="Set1") +
  theme_bw()
gplot
```
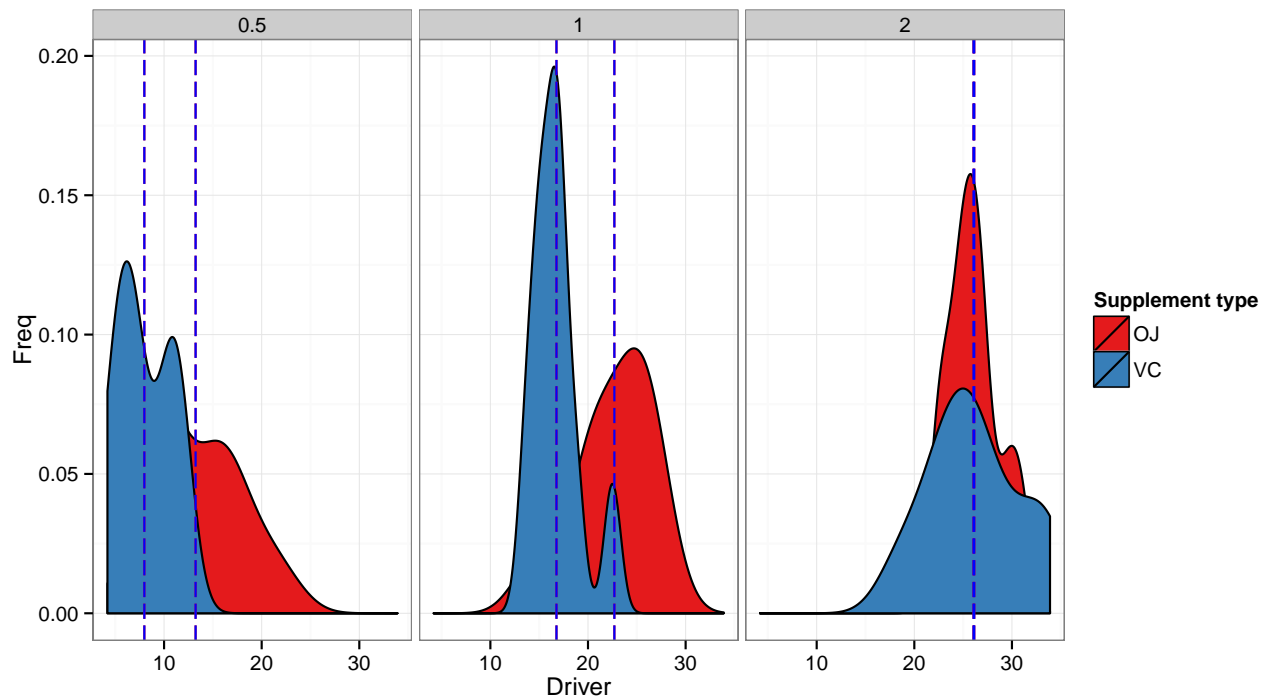
```
alpha <- 0.05
for (selDose in unique(ToothGrowth$dose)) {
  vcSet <- ToothGrowth %>% filter(supp=="VC", dose == selDose)
  ojSet <- ToothGrowth %>% filter(supp=="OJ", dose == selDose)
  t <- t.test(vcSet$len, ojSet$len, alternative = "two.sided",
              conf.level = (1-alpha))
  print(paste("p-Value for dosage ", selDose, " equals to ",
              round(t$p.value,2),".", sep=""))
}
```

```
## [1] "p-Value for dosage 0.5 equals to 0.01."
## [1] "p-Value for dosage 1 equals to 0."
## [1] "p-Value for dosage 2 equals to 0.96."
```

For dosages 0.5 and 1 there are 95% significant difference in resulting tooth lenth.