A basic use of statistical inference

John Lee

22 November 2015

Overview

This report (written using knitr) demonstrates knowledge of basic statistical inference through the use of the t-test to compare treatment groups in the ToothGrowth dataset in the R computing environment.

Results

A quick exploration of the ToothGrowth dataset reveals the three variables len, supp and dose to be a continuous variable, a factor with two levels, and a factor with 3 levels respectively:

```
summary(ToothGrowth$len)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    4.20   13.08   19.25   18.81   25.28   33.90
```

```
summary(ToothGrowth$supp)
```

```
## OJ VC
## 30 30
```

```
unique(ToothGrowth$dose)
```

```
## [1] 0.5 1.0 2.0
```

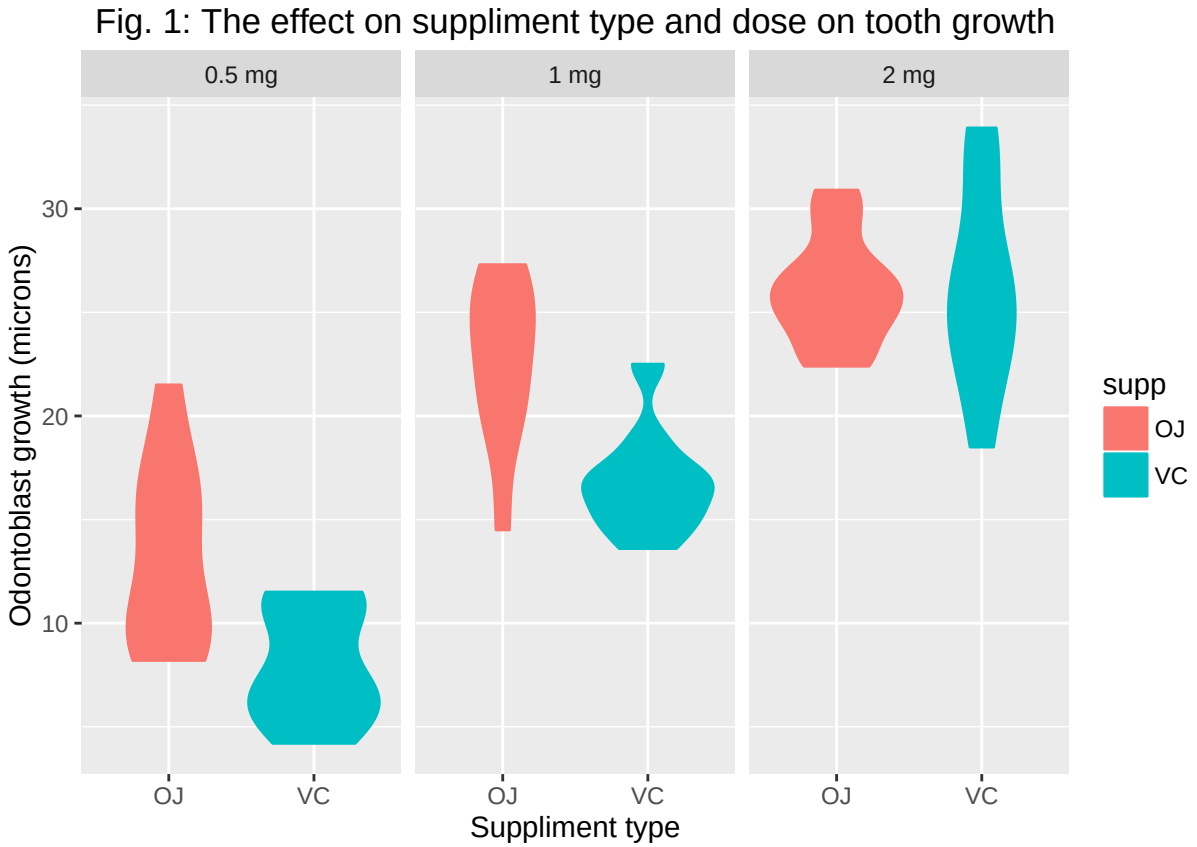
```
ToothGrowth <- mutate(ToothGrowth, dose= factor(paste(dose , "mg")))
```

The help file for this dataset informs us that the supp variable refers to the delivery method of Vitamin C. It is either orange juice (OJ) or ascorbic acid (VC). The dose variable is in units of milligrams. The length is of odontoblasts and is reported in microns (Crampton 1947). A plot of these factor level combinations and their effect on the resulting length of the odontoblast growth is depicted in Figure 1.

It would appear that the delivery method (supp) does not have a substantial effect on tooth growth whereas the dose of Vitamin C does. To assess this using a more formal statistical analysis we will use t-tests discussed in this lecture series to test the effect of the dose and supp variables. Using this testing method makes a number of assumptions and may have some problems.

These assumptions and problems include an assumption about the underlying distribution of the data and an assumption that there is no interaction between the two factors. The t-test assumes that the data is drawn from a normal distribution, an assumption that does not appear to be badly violated by the different groups (as assessed by the violin plot in Figure 1). An interaction between the two factors would confound our results. An interaction is not obvious from looking at Figure 1. When we make multiple comparisons between groups like this we increase the probability of obtaining a false positive (Type I error) but we will ignore that for the purposes of this report.

```
ggplot(ToothGrowth, aes(supp, len, fill = supp,colour = supp)) +
    geom_violin() + facet_grid(. ~ dose)+
    ggtitle("Fig. 1: The effect on suppliment type and dose on tooth growth")+
    xlab("Suppliment type")+
    ylab("Odontoblast growth (microns)")
```



A test for whether the method of delivery has an effect on odontoblast growth is reported below. The 95% confidence interval overlaps with zero indicating that we cannot reject the null hypothesis that the delivery method of Vitamin C has no effect on tooth growth.

```
    t.test(data = ToothGrowth, len ~ supp)
```

```
##
##  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
```

To assess the effect of Vitamin C dose we will test whether a difference in tooth growth is observed between the lowest and highest dose levels:

```r
dose_subset <-   filter(ToothGrowth, dose %in% c("0.5 mg", "2 mg"))
t.test(data = dose_subset, len ~ dose)
```

```
##
##  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 mg   mean in group 2 mg
##               10.605               26.100
```

This test indicates a significant difference between the two dose levels. The 95% confidence interval does not overlap with zero and indicates that with 2 mg supplimentation of Vitamin C vs. 0.5 mg supplimentation that guinea pig odontoblasts will grow an extra 12.8 to 18.1 microns.

Bibliography

Crampton, E W. 1947. "The growth of the odontoblasts of the incisor tooth as a criterion of the vitamin C intake of the guinea pig." *The Journal of Nutrition* 33 (5): 491–504.