

Statistical Inference Project - Part 2

1. Data Summary

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).

Format

A data frame with 60 observations on 3 variables. [,1] len numeric Tooth length [,2] supp factor Supplement type (VC or OJ). [,3] dose numeric Dose in milligrams.

Source

C. I. Bliss (1952) The Statistics of Bioassay. Academic Press.

References

McNeil, D. R. (1977) Interactive Data Analysis. New York: Wiley.

2. Exploration Analysis

```
library(datasets)
library(ggplot2)

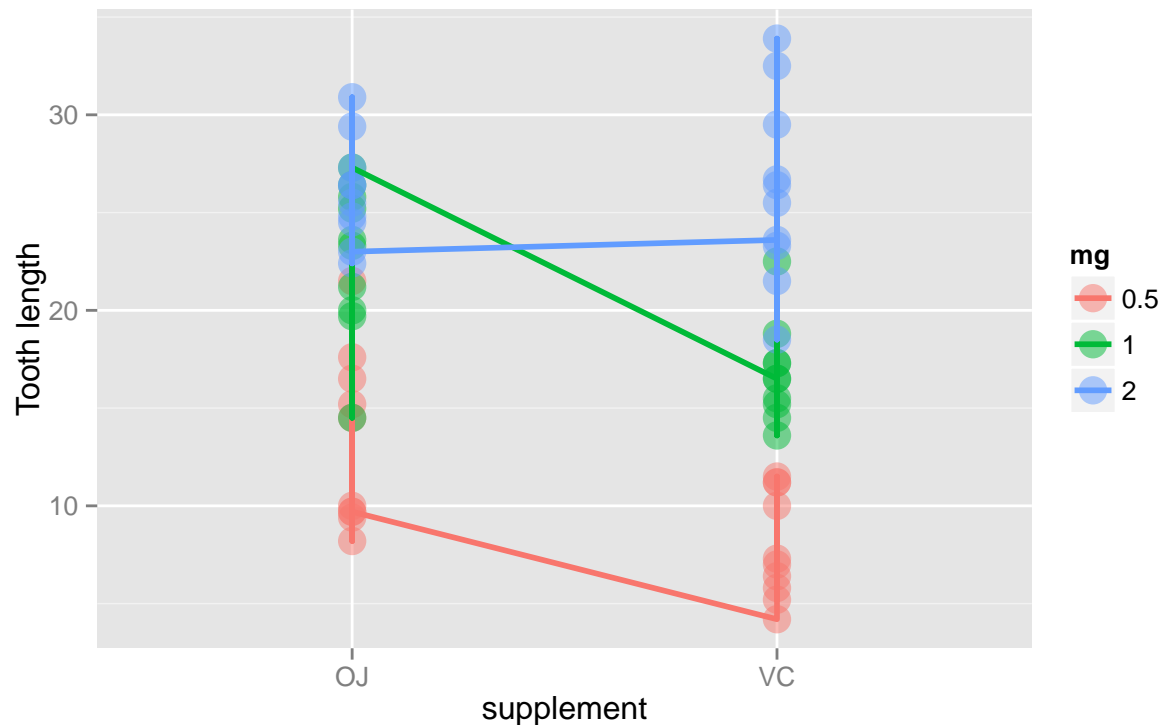
data(ToothGrowth)
head(ToothGrowth)
```

```
##      len supp dose
## 1   4.2   VC  0.5
## 2  11.5   VC  0.5
## 3   7.3   VC  0.5
## 4   5.8   VC  0.5
## 5   6.4   VC  0.5
## 6  10.0   VC  0.5
```

As the graph shows the result for the same dose are correlated. It is normal to expect that different subjects under the same dose and treatment would experience similar effects. In order to obtain the right results we compare both “supplements” “pairing” the results by the dose.

```
ggplot(ToothGrowth, aes(x = supp, y = len, colour = factor(dose), group = factor(dose))) +
  geom_point(size = 5, alpha = .5) +
  geom_line(size = 1) +
  labs(colour = 'mg', x = 'supplement', y = 'Tooth length') +
  ggtitle("ToothGrowth data: length vs dose, given type of supplement")
```

ToothGrowth data: length vs dose, given type of supplement



3. Confidence Interval

Paired observations are often analyzed using the t interval by taking differences. In R the `t.test` function performs one and two sample t -tests on vectors of data. I use this function with the option `paired = TRUE` for the reason explained previously.

```
t.test(len ~ I(relevel(supp, 2)), paired = TRUE, data = ToothGrowth)
```

```
##
## Paired t-test
##
## data: len by I(relevel(supp, 2))
## t = -3.3026, df = 29, p-value = 0.00255
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -5.991341 -1.408659
## sample estimates:
## mean of the differences
## -3.7
```

4. Conclusions

The t -statistic is $t = -3.3026$ and the 95 percent confidence interval: $-5.991341 - 1.408659$. Because the t -statistic is within the confidence interval, this indicates that with a 95 percent confidence, both supplements have the same effect in the teeth growth.