Tooth Growth Statistical Inference

Background

We investigate the effective Vitamin C on the tooth grow in Guinea pigs, using R's `ToothGrowth` dataset in the `datasets` package, which records the lengths of odontoblasts (cells responsible for tooth growth) in 60 Guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice (coded as `OJ`) or ascorbic acid (a form of vitamin C and coded as `VC`).

We set out to uncover the statistical correlation between the type/dose of supplements, and the tooth growth in Guinea pigs.

For this analysis, we will use a standard level of significance, $\alpha = 0.05$, and will be making inferences based on the p-value for simplicity.

Data Summary

We begin with a summary of the dataset. There are 60 observations in total.

```
data("ToothGrowth")
str(ToothGrowth)
```

```
## '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: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
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
```

There are two types of supplements (`OJ` and `VC`), and three different doses (0.5, 1.0, and 2.0). There are 10 observations for each supplement-dose combination. The dataset is balanced albeit small in size.

```
with(ToothGrowth, table(dose, supp))
```

```
##      supp
## dose  OJ VC
##   0.5 10 10
##   1   10 10
##   2   10 10
```

General Comparison

We begin with a comparison between the two groups of Guinea pigs receiving different types of supplement. The observations are of 60 individual animals, the observations are hence unpaired. We also do not have sufficient evidence to believe the variance to be equal between groups.

We begin with a two-tailed T-test (i.e., null hypothesis: the type of supplement is not associated with tooth growth).

```r
t.test(len ~ supp,
       paired = FALSE,
       var.equal = FALSE,
       data = ToothGrowth)
```

```
##
##  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
```

We are unable to reject the null hypothesis that different types of supplements are associated with equal outcome.

However, if we switch to a one-tailed test, thereby doubling the rejection region in either tail, our observation becomes significant. In this case, our alternative hypothesis is that group `OJ` correlates with increased tooth growth. We reject the null hypothesis in favour of our one-tailed alternative hypothesis.

```r
t.test(len ~ supp,
       paired = FALSE,
       var.equal = FALSE,
       data = ToothGrowth,
       alternative = "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.03032
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.4682687       Inf
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Effect of Dosage, Controlling for Type of Supplement

We begin by filtering the dataset by the type of supplement.

```r
tg_oj <- filter(ToothGrowth, supp == "OJ")
tg_vc <- filter(ToothGrowth, supp == "VC")
```

Since each group receives 3 different doses, 0.5, 1.0, and 2.0mg/day, we should conduct three individual two-sample T-tests (with unequal variance) to fully understand the correlation between dosage and tooth growth.

This is equivalent to the `pairwise.t.test()` function, if we do not adjust for P-values, and use non-pooled variances.

```r
# Pair-wise T-test Function
pwtt_tg <- function(data_arg, alt_arg) {
  with(data_arg,
       pairwise.t.test(len,
                       dose,
                       p.adjust.method = "none",
                       alternative = alt_arg,
                       pool.sd = FALSE
       ))
}

pwtt_tg(tg_vc, "two.sided")
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  len and dose
##
##   0.5     1
## 1 6.8e-07 -
## 2 4.7e-08 9.2e-05
##
## P value adjustment method: none
```

```r
pwtt_tg(tg_vc, "greater")
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  len and dose
##
##   0.5     1
## 1 3.4e-07 -
## 2 2.3e-08 4.6e-05
##
## P value adjustment method: none
```

Hence, for the `VC` group, increased dosage is associated with increased tooth growth. The same is observed in the `OJ` group.

```r
pwtt_tg(tg_oj, "two.sided")
```

```
##
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  len and dose
##
##   0.5     1
## 1 8.8e-05 -
## 2 1.3e-06 0.039
##
## P value adjustment method: none
```

```r
pwtt_tg(tg_oj, "greater")
```

```
##
```

```
##  Pairwise comparisons using t tests with non-pooled SD
##
## data:  len and dose
##
##   0.5     1
## 1 4.4e-05 -
## 2 6.6e-07 0.02
##
## P value adjustment method: none
```

Effect of Type of Supplement Controlling for Dosage

We slice the dataset based on the dosage.

```
tg_half <- filter(ToothGrowth, dose == 0.5)
tg_one <- filter(ToothGrowth, dose == 1)
tg_two <- filter(ToothGrowth, dose == 2)

ttest_dose <- function(data_arg, alt_arg){
  t.test(len ~ supp,
         paired = FALSE,
         var.equal = FALSE,
         data = data_arg,
         alternative = alt_arg
  )
}
```

For 0.5mg/day:

```
ttest_dose(tg_half, "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

```
ttest_dose(tg_half, "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 3.1697, df = 14.969, p-value = 0.003179
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.34604     Inf
## sample estimates:
## mean in group OJ mean in group VC
##            13.23             7.98
```

For 1.0mg/day:

```
ttest_dose(tg_one, "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

```
ttest_dose(tg_one, "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 4.0328, df = 15.358, p-value = 0.0005192
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  3.356158      Inf
## sample estimates:
## mean in group OJ mean in group VC
##            22.70            16.77
```

For 2.0mg/day:

```
ttest_dose(tg_two, "two.sided")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.79807  3.63807
## sample estimates:
## mean in group OJ mean in group VC
##            26.06            26.14
```

```
ttest_dose(tg_two, "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = -0.046136, df = 14.04, p-value = 0.5181
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -3.1335     Inf
## sample estimates:
```

```
## mean in group OJ mean in group VC
##            26.06            26.14
```

Conclusion

In general, there is some evidence that Vitamin C delivered through orange juice is associated with more tooth growth than through absorbic acid.

On the other hand, there is strong evidence showing that increased dosage correlates with increased tooth growth, for both `OJ` and `VC` groups.

If we control for doses, at 0.5 and 1.0mg/day, `OJ` is linked to increased tooth growth relative to `VC`. Nevertheless, at 2.0mg/day, there appears to be no difference between the two types of supplement.