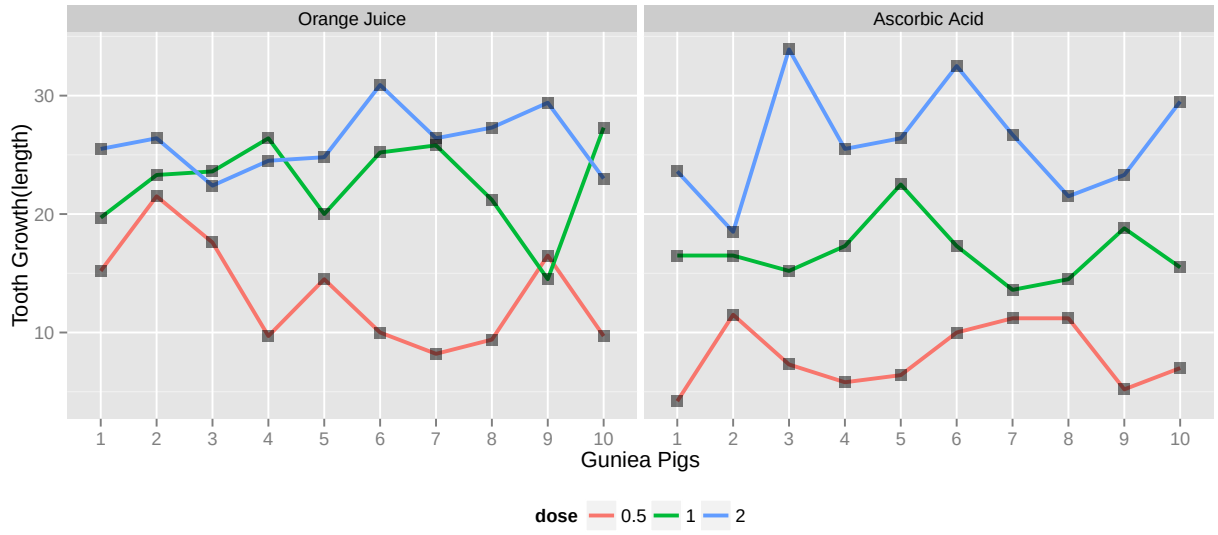
Toothgrowth Analysis

Nagesh Madhwal

October 20, 2015

The Effect of Vitamin C on Tooth Growth in Guinea Pigs

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 - "OJ", or ascorbic acid - "VC") 10 Guinea Pigs (subjects), 6 observations per subject(2 delivery methods x 3 doses). We are measuring Tooth Growth in terms of length increase against 6 treatments as combination of 2 delivery medium & 3 dose combinations. In the report wherever ever we mention tooth growth we are talking about length increase.



We have all 60 observations from the sampling distribution on the plot, Guniea Pigs on the x-axis, tooth growth on y, the plotted lines represent the three dose levels & the plots split for Orange juice & ascorbic acid. This assumes that the data is sequentially arranged so observation nos 1, 11, 21, 31, 41 & 51 are for the same Guinea Pig. Summary data for average tooth growth for each dosage type by the delivery method.

Delivery Method	Dose Level	Avg ToothGrowth
Orange Juice	0.5	13.23
Orange Juice	1.0	22.70
Orange Juice	2.0	26.06
Ascorbic Acid	0.5	7.98
Ascorbic Acid	1.0	16.77
Ascorbic Acid	2.0	26.14

The comparisons that can be done on this data are:
1) Overall effectiveness of the two Delivery mediums. We can check whether the difference in means is due to sample variation or the pvalue is low enough for us to say that one medium is more effective than the other.
2) We can compare the three dose levels for the two mediums with each other(0.5 of OJ with 0.5 of VC & so on), to see if we can establish clear superiority.The methodology for all three comparisions being the same we will compare dose levels of 2.0 between the two delivery mediums.

3) Within a single medium whether the difference in growth level can truly be attributed to the dose level. Here we will test for dose levels of 1.0 & 2.0 for OJ. We will perform the Students t test for the above three cases. The first step in that is to check the conditions.

Conditions Check for t Test

1) We are comparing the means of dependent samples. The observations have been taken on the same 10 pigs for the 6 sets of dose & delivery medium combinations. The samples are paired as we can correlated the same medium/dose combination to the same guniea pig.

2) Population following the normal distribution is an assumption we will take though we have no visibility to the distribution or its skews. To check this we would need to check the information from Source : "C. I. Bliss (1952) The Statistics of Bioassay. Academic Press"" & References: "McNeil, D. R. (1977) Interactive Data Analysis. New York: Wiley".

3) While we can look at the sample variances, that will not be an indicator of the population variances & so we will test values for both the assumptions of variance being equal & unequal.

4) The sample we are dealing with is a random sample & so truly representative of the population, I have not done this check and am taking that as an assumption. Again we could check the source & references for possible information on this aspect.

Case 1 - Overall Effectiveness of Delivery Methods

We will perform the t test first with var.equal = FALSE i.e. the two population variances are not equal. This may not be the case as ultimately the data is on the same set of guniea pigs for the same medium of delivery for similar doses of Vitamin C but since we dont know we are checking for both situations. In all three cases we will work with a significance level of 5%.
Our Null hypothesis H0 is that both the delivery methods are equal. The differences of the means mu = 0. Alternate Hypothesis Ha is that mu > 0 and Orange Juice as a delivery medium on average delivers higher length of Tooth Growth compared to ascorbic acid.

```
##
##  Paired t-test
##
## data:  OJ and VC
## t = 3.3026, df = 29, p-value = 0.001275
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  1.796409      Inf
## sample estimates:
## mean of the differences
##                     3.7
```

Our t value is large & the p value is extremely small comapred to the significance level of 5%. Based on this information & the assumptions we have made about the Population & samples we can reject the null hypothesis and accept the alternate.
On performing the t test with var.equal = TRUE which in my view would be closer to the actual case, we see that the values obtained are identical.For the remaining two Cases I am going to show the var.equal = TRUE analysis in this report.

CASE 2: Effectiveness of delivery medium with same level of dose.

We are setting up a two way hypothes.Signifcance level at 5%.
Null Hypothesis, H0: At a Vitamin C dose level of 2mg there is no difference in average tooth growth based on the delivery medium.
Alternate Hypothesis, Ha : At a Vitamin C dose level of 2mg average tooth growth can be different based on which delivery medium we choose.
Results assuming equal variance.

```
##
##  Paired t-test
##
## data:  VC2mg and OJ2mg
## t = 0.042592, df = 9, p-value = 0.967
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.168976  4.328976
## sample estimates:
## mean of the differences
##                    0.08
```

The results are again idential for equal & unequal variance. In this case we reject the alternate hypothesis as the pvalue is very large. Our null hypothesis which states that there is no difference in average tooth growth based on the delivery medium with 2mg dose of Vitamin C holds.

So we have an interesting situation where over all we can see that the delivery medium makes a difference to tooth growth but for a dose level of 2mg of Vitamin C the delivery medium does not make any difference.

CASE 3: Dose effectiveness within same medium.

Medium: Orange Juice. Dose Levels Vitamin C: 1mg & 2mg. Signifcance level at 5%.
Null Hypothesis, H0: There is no difference in tooth growth by increasing the Vitamin C dose from 1mg to 2mg while using Orange Juice as the delivery medium.
Alternate Hypothesis, Ha: Increasing the dose of Vitamin C from 1mg to 2mg with Orange Juice as the delivery medium increases the growth in tooth length.
One sided hypothesis set for mu > 0

```
##
##  Paired t-test
##
## data:  OJ2mg and OJ1mg
## t = 1.9435, df = 9, p-value = 0.04192
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.1908169       Inf
## sample estimates:
## mean of the differences
##                    3.36
```

The null hypothesis is rejected but the result is very close to our signifiacne level of 5%.

Appendix 1

References: 1) Coursera Statistical Inference Course Notes.
2) Stack Overflow, Quick R, Cran Documentation for code clarifications

Appendix 2

R Code

1) For the plot of tooth growth by dosage

```
K <- ToothGrowth
K$GunieaPigs <- rep(1:10, 6)
K$dose <- as.factor(K$dose)
levels(K$supp) <- c("Orange Juice", "Ascorbic Acid")
library(ggplot2)
g <- ggplot(K, aes(x = factor(GunieaPigs), y = len, group = factor(dose)), )
g <- g + geom_line(size = 1, aes(colour = dose)) + facet_grid(.~supp)
g <- g + geom_point(size =5, pch = 15, fill = "brown", alpha = .5)
g + theme(legend.position = "bottom") + labs(x = "Guniea Pigs", y = "Tooth Growth(length)")
```

2) Summary Data

```
library(dplyr)
TG <- summarise(group_by(ToothGrowth, supp, dose), mean(len))
levels(TG$supp) <- c("Orange Juice", "Ascorbic Acid")
colnames(TG) <- c("Delivery Method", "Dose Level", "Avg ToothGrowth")
library(knitr)
kable(TG)
```

3) For t.tests

```
VC <- ToothGrowth$len[1:30]
OJ <- ToothGrowth$len[31:60]
t.test(OJ, VC, alternative = "greater",paired = TRUE, var.equal = FALSE)
```

```
VC2mg <- ToothGrowth$len[21:30]
OJ2mg <- ToothGrowth$len[51:60]
t.test(VC2mg, OJ2mg, paired = TRUE, var.equal = TRUE)
```

```
OJ1mg <- ToothGrowth$len[41:50]
OJ2mg <- ToothGrowth$len[51:60]
t.test(OJ2mg, OJ1mg, alternative = "greater", paired = TRUE, var.equal = TRUE)
```