

# Comparing means across two factors and six treatments

*James C. Dembowski*

*August 13, 2016*

## Overview

For this project, means for two factors and selected treatments are compared. The data come from a 1952 study by C. I. Bliss, comparing the effect of vitamin C on odontoblast growth in 60 guinea pigs, using two different delivery methods, orange juice and an ascorbic acid solution, and three different doses, 0.5 mg, 1.0 mg, and 2.0 mg.

## Assumptions

We make some assumptions for our analysis. First, we assume that Vitamin C affects odontoblast growth in a statistically significant manner, since we have no control group. Next, we assume that the samples of guinea pigs in the experiment are independent of one another. We also assume that odontoblast growth is normally distributed within the respective populations. Finally, we assume the populations from which our treatments come have equal variances. We verify the last two assumptions are reasonable using the Shapiro-Wilk test and Bartlett's test, respectively. (A standard F-test was performed to check for equal variances for comparing delivery methods, since there are only two samples.)

It is worth noting that t-tests are not ideal for these analyses, as multiple comparisons increase the probability of a Type I error. Per project instructions, we ignore this potential issue and use multiple t-tests without correction.

## Data Summary (full data set)

The basic statistics for the entire data set are listed in Table 1. (Units are microns.)

Table 1: Mean, Median, Standard Deviation

Mean	Median	SD
18.813	19.25	7.649

## Comparing Factor Means

### Delivery Method

The guinea pigs are given vitamin C by two different delivery methods. Half of the animals receive it as orange juice (OJ), and half the animals receive it as an ascorbic acid solution (VC).

Table 2: Mean, Median, Standard Deviation, by Delivery Method

supp	Mean	Median	SD
Ascorbic Acid	16.963	16.5	8.266
Orange Juice	20.663	22.7	6.606

The most noticeable difference for the delivery methods is in the medians. The Shapiro-Wilk test indicates the Orange Juice sample is *not* from a normal population, but the t-test is reasonably robust to the assumption of normality with equal sample sizes, so we still use it. The alternative hypothesis of the t-test is that the true difference in means is not equal to 0. The test yields a p-value of 0.06039, so the difference between the means is not statistically significant at the 95-percent confidence level.

### Dosage level

For each delivery method, the guinea pigs are divided into three groups, and each group receives a different dosage of vitamin C, 0.5 mg, 1 mg, or 2 mg. The dosage, without consideration to delivery method, is examined.

Table 3: Mean, Median, Standard Deviation, by Dosage Level

Dose	Mean	Median	SD
0.5 mg	10.605	9.85	4.500
1.0 mg	19.735	19.25	4.415
2.0 mg	26.100	25.95	3.774

T-tests are performed for each of the three possible pairs, 0.5 mg & 1 mg, 0.5 mg & 2 mg, and 1 mg & 2 mg, with an alternative hypothesis that the difference in means is not equal to 0. The p-values are  $1.266 \times 10^{-7}$ ,  $2.838 \times 10^{-14}$ , and  $1.811 \times 10^{-5}$ , respectively, all statistically significant. The null hypotheses are rejected.

### Comparing Treatment Means

The individual treatments are also considered. The statistics for each of them are included in Table 4.

Table 4: Mean, Median, Standard Deviation, by Treatment

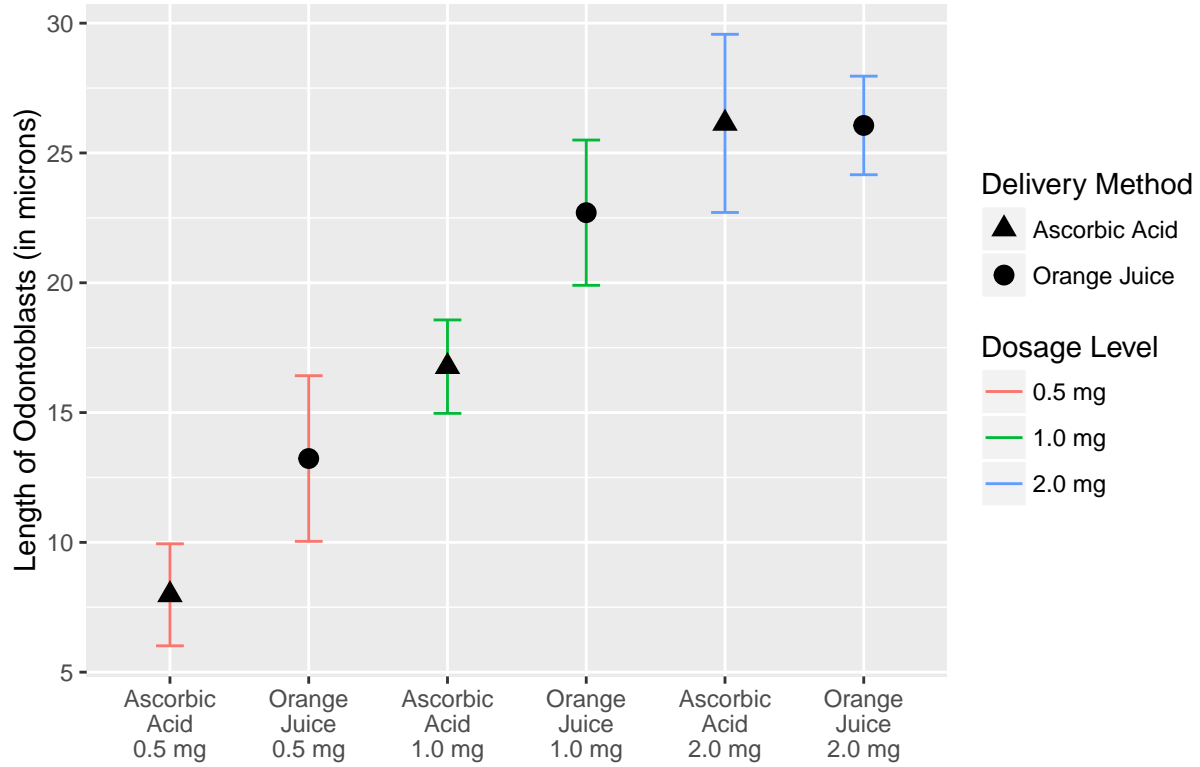
Treatment	Mean	Median	SD
Ascorbic Acid 0.5 mg	7.98	7.15	2.747
Ascorbic Acid 1.0 mg	16.77	16.50	2.515
Ascorbic Acid 2.0 mg	26.14	25.95	4.798
Orange Juice 0.5 mg	13.23	12.25	4.460
Orange Juice 1.0 mg	22.70	23.45	3.911
Orange Juice 2.0 mg	26.06	25.95	2.655

Confidence intervals are calculated for each treatment and are shown graphically in Figure 1.

T-tests are performed for all treatment pairs within each factor. Means are compared for different dosages for guinea pigs that all have vitamin C delivered with the same method. Additionally, means are compared between different delivery methods at each dosage level. Comparisons are not made when two treatments have no common level for either factor, e.g., orange juice at 1 mg is not compared to ascorbic acid at 2 mg. For each test, the alternative hypothesis is that the difference in means is not equal to 0.

The null hypothesis is rejected for every comparison, except the comparison between orange juice and ascorbic acid at a dosage of 2 mg. P-values for each test are provided in Table 5.

Figure 1: 95% Confidence Intervals for Mean by Treatment



The only surprising result is from the comparison between Orange Juice 1.0 and 2.0 mg treatments. The t-test indicates a statistically significant difference in means, but the 95-percent confidence intervals of the means overlap. This highlights the fact that even though non-overlapping confidence intervals indicate statistically significant differences, *the converse is not true*.

Table 5: P-values of t-tests Between Treatments

Treatment 1	Treatment 2	Difference in Means	p-value
Ascorbic Acid 0.5 mg	Ascorbic Acid 1.0 mg	9.47	0.000083576
Ascorbic Acid 0.5 mg	Ascorbic Acid 2.0 mg	12.83	0.000000340
Ascorbic Acid 0.5 mg	Orange Juice 0.5 mg	5.25	0.005303661
Ascorbic Acid 1.0 mg	Ascorbic Acid 2.0 mg	3.36	0.037362796
Ascorbic Acid 1.0 mg	Orange Juice 1.0 mg	5.93	0.000780726
Ascorbic Acid 2.0 mg	Orange Juice 2.0 mg	0.08	0.963709779
Orange Juice 0.5 mg	Orange Juice 1.0 mg	8.79	0.000000649
Orange Juice 0.5 mg	Orange Juice 2.0 mg	18.16	0.000000005
Orange Juice 1.0 mg	Orange Juice 2.0 mg	9.37	0.000033976

## Conclusion

There is a wide range of effects for the odontoblast growth in the guinea pig subjects that receive vitamin C. It is clear that different delivery methods and dosages tend to provide different mean results. It is especially interesting to see that, within dosages, the delivery method is important at two of the three dosage levels, despite there being no statistically significant difference between delivery methods when dosage is not considered.

## Appendix

### R Code

```
library( datasets)
library( dplyr)
library( ggplot2)
library( tidyr)
library( knitr)

teeth = tbl_df( ToothGrowth)

# Clean up formats of delivery method and dosage for nicer tables
teeth$supp = as.character( teeth$supp)
teeth$supp[ teeth$supp == "VC"] = "Ascorbic Acid"
teeth$supp[ teeth$supp == "OJ"] = "Orange Juice "
teeth$supp = as.factor( teeth$supp)
teeth$dose[ teeth$dose == 0.5] = "0.5 mg" # transforms dose into character
teeth$dose[ teeth$dose == "1"] = "1.0 mg"
teeth$dose[ teeth$dose == "2"] = "2.0 mg"
teeth$dose = as.factor( teeth$dose)

summaryOverall = teeth %>%
summarize( Mean = mean( len), Median = median( len), SD = sd( len))
kable( summaryOverall, digits = 3, caption = "Mean, Median, Standard Deviation")

# Table of mean, median, sd, by delivery method
summarySupp = teeth %>%
group_by( supp) %>%
summarize( Mean = mean( len), Median = median( len), SD = sd( len))
kable( summarySupp, digits = 3, caption = "Mean, Median, Standard Deviation, by
      Delivery Method")

# T- test to compare means of delivery methods
# Normality assumption reasonable?
shapiro.test( teeth$len[ teeth$supp == "Orange Juice "]) # p-value == 0.02359!!!!
shapiro.test( teeth$len[ teeth$supp == "Ascorbic Acid"])
# Equal variances?
var.test( teeth$len[ teeth$supp == "Orange Juice "], teeth$len[ teeth$supp == "Ascorbic Acid"])

t.test( len ~ supp, data = teeth, var.equal = TRUE)

# Table of mean, median, sd, by dosage level
summaryDose = teeth %>%
group_by( dose) %>%
summarize( Mean = mean( len), Median = median( len), SD = sd( len)) %>%
rename( Dose = dose)
kable( summaryDose, digits = 3, caption = "Mean, Median, Standard Deviation, by
      Dosage Level")

# T-tests to compare means of dosage levels.
# First, is the normality assumption reasonable?
shapiro.test( teeth$len[ teeth$dose == "0.5 mg"])
```

```

shapiro.test( teeth$len[ teeth$dose == "1.0 mg"])
shapiro.test( teeth$len[ teeth$dose == "2.0 mg"])
# Next, is the equal variance assumption reasonable?
listOfSamples = list( teeth$len[ teeth$dose == "0.5 mg"],
  teeth$len[ teeth$dose == "1.0 mg"], teeth$len[ teeth$dose == "2.0 mg"])
bartlett.test( listOfSamples)

# Assumptions are reasonable
t.test( len ~ dose, data = teeth[ teeth$dose != "2.0 mg",], var.equal = TRUE)
t.test( len ~ dose, data = teeth[ teeth$dose != "1.0 mg",], var.equal = TRUE)
t.test( len ~ dose, data = teeth[ teeth$dose != "0.5 mg",], var.equal = TRUE)

# Table of mean, median, sd, by treatment
summaryTreatments = teeth %>%
group_by( supp, dose) %>%
summarize( Mean = mean( len), Median = median( len), SD = sd( len)) %>%
unite( Treatment, supp:dose, sep = " ")
kable( summaryTreatments, digits = 3, caption = "Mean, Median, Standard Deviation, by
  Treatment")

# Create data frame appropriate for plotting confidence intervals of treatment
# means
treatmentMeanEstimates = teeth %>%
group_by( supp, dose) %>%
summarize( Mean = mean( len), SE = sd( len) / sqrt( 10)) %>%
mutate( Lower = Mean - qt( .975, 9) * SE, Upper = Mean + qt( .975, 9) * SE) %>%
unite( Treatment, supp:dose, sep = " ", remove = FALSE)
treatmentMeanEstimates$Treatment = factor( treatmentMeanEstimates$Treatment,
  levels = treatmentMeanEstimates$Treatment[ c( 1, 4, 2, 5, 3, 6)])
treatmentMeanEstimates$dose = factor( treatmentMeanEstimates$dose)

# Create chart of 95% confidence intervals of treatment means
treatmentMeanCI = ggplot( treatmentMeanEstimates, aes( x = Treatment, y = Mean,
  color = dose))
treatmentMeanCI = treatmentMeanCI + geom_errorbar( aes( x = Treatment,
  ymin = Lower, ymax = Upper), width = 0.2)
treatmentMeanCI = treatmentMeanCI + scale_shape_manual( values = c( 17, 19))
treatmentMeanCI = treatmentMeanCI + geom_point( size = 3, color = "black",
  aes( shape = supp))
treatmentMeanCI = treatmentMeanCI +
  labs( color = "Dosage Level", shape = "Delivery Method")
treatmentMeanCI = treatmentMeanCI +
  ggtitle( "Figure 1: 95% Confidence Intervals for Mean by Treatment")
treatmentMeanCI = treatmentMeanCI + scale_x_discrete( labels = c(
  "Ascorbic\nAcid\n0.5 mg", "Orange\nJuice\n0.5 mg", "Ascorbic\nAcid\n1.0 mg",
  "Orange\nJuice\n1.0 mg", "Ascorbic\nAcid\n2.0 mg", "Orange\nJuice\n2.0 mg"))
treatmentMeanCI = treatmentMeanCI +
  scale_shape_discrete( labels = c( "Orange Juice", "Ascorbic Acid"))
treatmentMeanCI = treatmentMeanCI +
  scale_color_discrete( labels = c( "0.5 mg", "1.0 mg", "2.0 mg"))
treatmentMeanCI = treatmentMeanCI + ylab( "Length of Odontoblasts (in microns)")
treatmentMeanCI = treatmentMeanCI + xlab( "")
print( treatmentMeanCI)

```

```

# T-tests to compare treatment means for selected pairs
# First, is the normality assumption reasonable?
shapiro.test( teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "0.5 mg"])
shapiro.test( teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "1.0 mg"])
shapiro.test( teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "2.0 mg"])
shapiro.test( teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "0.5 mg"])
shapiro.test( teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "1.0 mg"])
shapiro.test( teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "2.0 mg"])
# Next, is the equal variance assumption reasonable?
listOfSamples2 = list( teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "0.5 mg"],
  teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "1.0 mg"],
  teeth$len[ teeth$supp == "Ascorbic Acid" & teeth$dose == "2.0 mg"],
  teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "0.5 mg"],
  teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "1.0 mg"],
  teeth$len[ teeth$supp == "Orange Juice " & teeth$dose == "2.0 mg"])
bartlett.test( listOfSamples2)

# Assumptions are reasonable
results = list( length = 9) # results is a list of results for t.tests for all
  # treatment combinations of interest
results[[ 1]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Orange Juice " & teeth$dose != "2.0 mg",], var.equal = TRUE)
results[[ 2]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Orange Juice " & teeth$dose != "1.0 mg",], var.equal = TRUE)
results[[ 3]] = t.test( len ~ supp, data = teeth[ teeth$dose == "0.5 mg",], var.equal = TRUE)
results[[ 4]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Orange Juice " & teeth$dose != "0.5 mg",], var.equal = TRUE)
results[[ 5]] = t.test( len ~ supp, data = teeth[ teeth$dose == "1.0 mg",], var.equal = TRUE)
results[[ 6]] = t.test( len ~ supp, data = teeth[ teeth$dose == "2.0 mg",], var.equal = TRUE)
results[[ 7]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Ascorbic Acid" & teeth$dose != "2.0 mg",], var.equal = TRUE)
results[[ 8]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Ascorbic Acid" & teeth$dose != "1.0 mg",], var.equal = TRUE)
results[[ 9]] = t.test( len ~ dose,
  data = teeth[ teeth$supp == "Ascorbic Acid" & teeth$dose != "0.5 mg",], var.equal = TRUE)

pValue = vector( length = 9)
meanDiff = vector( length = 9)
for (i in 1:9) {
  pValue[ i] = results[[ i]]$p.value
  meanDiff[ i] = unname( abs( diff( results[[ i]]$estimate)))
}
treatmentTTest = data.frame(
  summaryTreatments$Treatment[ c( 1, 1, 1, 2, 2, 3, 4, 4, 5)],
  summaryTreatments$Treatment[ c( 2, 3, 4, 3, 5, 6, 5, 6, 6)],
  meanDiff, pValue)
treatmentTTest = tbl_df( treatmentTTest)
colnames( treatmentTTest) = c( "Treatment 1", "Treatment 2", "Difference in Means", "p-value")
kable( treatmentTTest, digits = 7, caption = "P-values of t-tests Between
  Treatments")

```