# Statistical Inference Course Project

Part 2: Basic Inferential Data Analysis

*audrey*

*25 October, 2015*

## Introduction

The **ToothGrowth**[1] data set investigates the effect of vitamin C on tooth growth in 10 guinea pigs. The data set contains 60 observations on 3 variables: *len* denotes the tooth length, *supp* represents the vitamin C supplement type [absorbic acid (VC) or orange juice (OJ)], and *dose* is the dose in milligrams. Note: I am inferring that "VC" represents absorbic acid as the dataset description is poor at best. The response is the tooth length growth; we are interested in exploring whether growth is affected by dose or supplement type or both factors. We achieve this through statistical inference; however, it is of interest to complete an exploratory data analysis first in order to drive our inferential statistics. Note that *response* and *tooth length growth* are used interchangeably in this report.

## Exploratory data analysis

We obtain some basic information and summary of the data including the number of observations, variables and quartile values (see Appendix for more data summary statistics).

```
## '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 ...
```

We reshape the data to show the average tooth length readings according to supplement type and dose.

```
##   supp dose   len
## 1   OJ  0.5 13.23
## 2   OJ  1.0 22.70
## 3   OJ  2.0 26.06
## 4   VC  0.5  7.98
## 5   VC  1.0 16.77
## 6   VC  2.0 26.14
```

In this section, we use boxplots to illustrate the relationship between the response and supplement typse by ignoring dose levels (see Fig. 1) and factoring in dose levels (see Fig. 3). We also investigate the relationship of response against dose ignoring supplement type (see Fig. 2).

In Fig. 1, we compare the distributions between mean response readings obtained from the OJ and VC delivery methods ignoring dose. It appears that the supplement type may affect the tooth growth since the mean readings recorded for the OJ are higher than the VC type. However, the boxplot alone does not provide sufficient evidence on whether our observations are statistically significant. In Fig. 2, we plot the relationship between response and supplement dose ignoring delivery method. We observe an increasing trend with increasing dose (see also Fig. 4 in Appendix). Finally, Fig. 3 shows the relationship between response and supplement type grouped by dose; while the mean responses suggest a difference in OJ and VC delivery methods for dose levels of 0.5 and 1 mg, the two methods appear very similar for a 2 mg dose.

[1] Source: C. I. Bliss (1952) The Statistics of Bioassay. Academic Press.

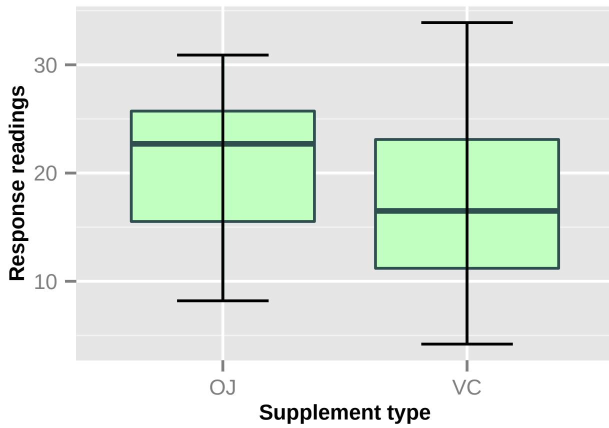


Figure 1: Relationship between response readings and supplement type

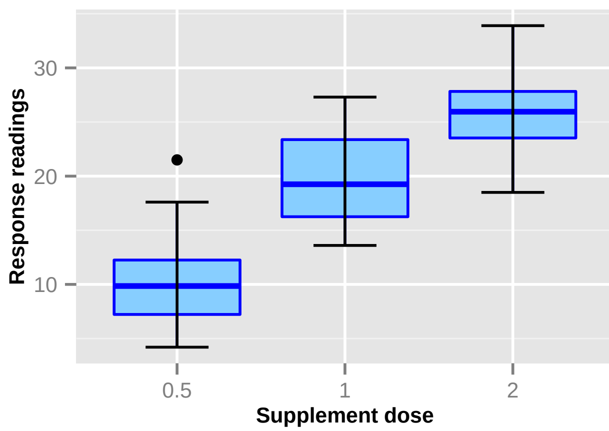


Figure 2: Relationship between response readings and supplement dose

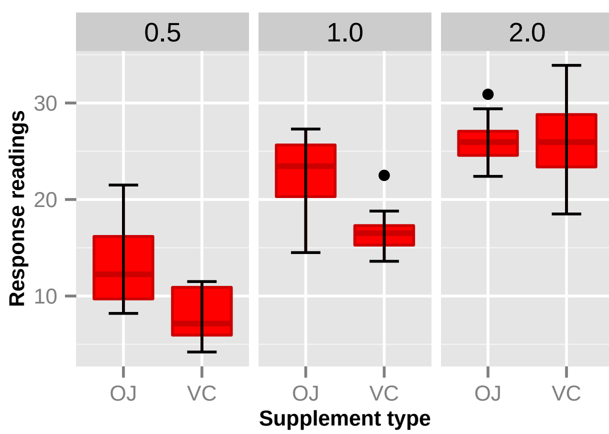


Figure 3: Relationship between response readings and supplement type for each dose

## Inferential statistics

The statistics given above are merely descriptive; the averaged response readings differ depending on supplement type and dose but to show whether the observed differences are reliable, one resorts to inference. We note that we have 60 observations corresponding to 60 different guinea pigs; it follows that we treat the data as **unpaired** in all the tests in this report. Consequently, we use an **independent samples t-test** to test the effectiveness of supplement type and dose on response. Further, we assume unequal variance; the latter leads to an adjusted degree of freedom calculation according to Welch's test. The results are discussed in the main body of the report and the full statistics are included in the Appendix.

*Effectiveness of supplement type on tooth growth*

For the first test, we split the data in two groups according to supplement type and perform an unpaired t-test. The results are shown below; the p-value indicates that there is a probability of 0.061 of obtaining a difference in means of 3.7 or larger **by chance**. That said, in most research, the conventional p-value associated with what is considered statistically significant, is $< 0.05$. We note that our calculated value is marginally close to the conventional cut-off for statistical significance. We now look at the 95% confidence interval output from the t-test. The confidence interval for the difference between the two means contains all the values of $(\mu_{OJ} - \mu_{VC})$, where $\mu_{OJ}$ and $\mu_{VC}$ denote the mean tooth length growth from the OJ and VC supplement types, respectively, which would not be rejected under the null hypothesis of $H_0 : \mu_{OJ} - \mu_{VC} = 0$. Since the calculated interval under the 95% confidence level includes 0, it implies that a mean of 0 is a possibility for the true value of the difference. We can conclude, based on the independent samples t-test, that the difference between the means is *not* significant and thus we fail to reject the null hypothesis.

```
##            dose = all
## t          1.91526827
## ci.lower -0.17101562
## ci.upper  7.57101562
## p.value   0.06063451
```

*Effectiveness of supplement dose on tooth growth*

Referring to Fig. 2, we observe that averaged responses suggest a difference across dose. Observe that we have 3 groups for each dose[2]; we proceed by comparing possible pairs of dose [1-0.5, 2-0.5 and 2-1] with t-tests. As suggested by p-values of $<< 0.05$ as well as confidence intervals entirely above 0, we conclude that the results are statistically significant.

```
##             1 vs .5 mg   2 vs .5 mg    2 vs 1 mg
## t         6.476648e+00 1.179905e+01 4.9004843172
## ci.lower 6.276219e+00 1.283383e+01 3.7335194831
## ci.upper 1.198378e+01 1.815617e+01 8.9964805169
## p.value   1.268301e-07 4.397525e-14 0.0000190643
```

*Effectiveness of supplement dose and type on tooth growth*

We now repeat the first test taking into account supplemtn dose and type. We note that our study ignores the 2 mg dose since the observed sample means show little difference (see Fig. 3). Two t-tests are performed with response data from VC and OJ corresponding to a dose of 0.5 and 1 mg. Again, the t-test results suggest statistical significance.

```
##            dose = 0.5    dose = 1
## t         3.169732784 4.032769634
## ci.lower 1.719057271 2.802148249
## ci.upper 8.780942729 9.057851751
## p.value   0.006358607 0.001038376
```

[2]For the comparison of more than two means, computations like ANOVA are better suited. Since ANOVA is beyond the scope of the class, we compare possible pairs with t-tests, although this might not be completely valid.

## Concluding remarks

We have carried out inferential data analyses on the ToothGrowth data set. Independent samples t-tests were used to test the effectiveness of supplement dose and type on tooth length. As discussed above, we have assumed an unpaired data set and unequal variance. Additionally, we make the following assumptions when using the t-test statistic:

- The data follow the normal probability distribution (see Shapiro-Wilk normality test and Q-Q plot in Appendix).
- The two samples are independent: there is no relationship between the guinea pigs in one sample as compared to the other.

The results are summarised as follows:

- Effectiveness on supplement type on response, t(df = 55.309) = 1.92, p = 0.06: no significant difference was found
- Effectiveness on supplement dose on response:
  - Pair 1 (1 vs 0.5 mg), t(df = 37.986) = 6.47, p = 1.26e-07: 1 mg is associated with *higher* growth than 0.5 mg
  - Pair 2 (2 vs 0.5 mg), t(df = 36.883) = 11.80, p = 4.40e-15: 2 mg is associated with *higher* growth than 0.5 mg
  - Pair 3 (2 vs 1 mg), t(df = 37.101) = 4.90, p = 1.913-05: 2 mg is associated with *higher* growth than 1 mg
- Effectiveness on supplement type (per dose level) on response:
  - For 0.5 mg, t(df = 14.969), p = 0.006: OJ type is associated with *higher* growth than VC
  - For 1 mg, t(df = 15.358), p = 0.001: OJ type is associated with *higher* growth than VC

## Appendix

### Summary statistics

```
##       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  
```

An increase in dose leads to an increase in averaged tooth length readings while OJ as the supplement type leads to higher response at 0.5, 1 mg dose but the effect decreases significantly at 0.2 mg. This trend is most clearly observed graphically in Fig. 4.

### Independent samples t-test results: R output

Here, we show the results as output by R on the various t-tests discussed in the main body of the report.

*Test 1*: effectiveness on supplement type on response

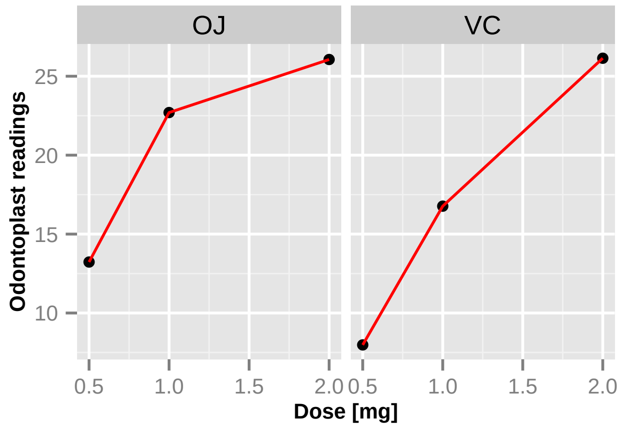


Figure 4: Relationship between tooth length readings and supplement dose grouped by supplement type.

```
##
##  Welch Two Sample t-test
##
## data:  TGOJ$len and TGVC$len
## 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 of x mean of y
##  20.66333  16.96333
```

*Test 2*: effectiveness on 1 and 0.5 mg dose on response:

```
##
##  Welch Two Sample t-test
##
## data:  TG1$len and TG05$len
## t = 6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   6.276219 11.983781
## sample estimates:
## mean of x mean of y
##    19.735    10.605
```

*Test 3*: effectiveness on 2 and 0.5 mg dose on response:

```
##
##  Welch Two Sample t-test
##
## data:  TG2$len and TG05$len
## 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:
##  12.83383 18.15617
```

```
## sample estimates:
## mean of x mean of y
##    26.100    10.605
```

*Test 4*: effectiveness on 2 and 1 mg dose on response:

```
##
##  Welch Two Sample t-test
##
## data:  TG2$len and TG1$len
## t = 4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.733519 8.996481
## sample estimates:
## mean of x mean of y
##    26.100    19.735
```

*Test 5*: effectiveness on supplement type on response with dose = 0.5

```
##
##  Welch Two Sample t-test
##
## data:  TGOJ05$len and TGVC05$len
## 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 of x mean of y
##     13.23      7.98
```

*Test 6*: effectiveness on supplement type on response with dose = 1

```
##
##  Welch Two Sample t-test
##
## data:  TGOJ1$len and TGVC1$len
## 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 of x mean of y
##     22.70     16.77
```

### Shapiro Wilk normality test and Q-Q plot

As indicated in the main body of the report, the t-test requires that the distributions do not departure from normality. The Shapiro-Wilk test uses the null hypothesis to check whether our sample data used to compute the t statistic comes from a normally distributed population. We check here that whether the data corresponding to 1mg dose is normally distributed.

```r
shapiro.test(TG1$len)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  TG1$len
## W = 0.93134, p-value = 0.1639
```

Since the p-value>0.05, the null hypothesis that the population is normally distributed cannot be rejected. In addition, we include a Q-Q plot which verifies that the data is approximately normal.

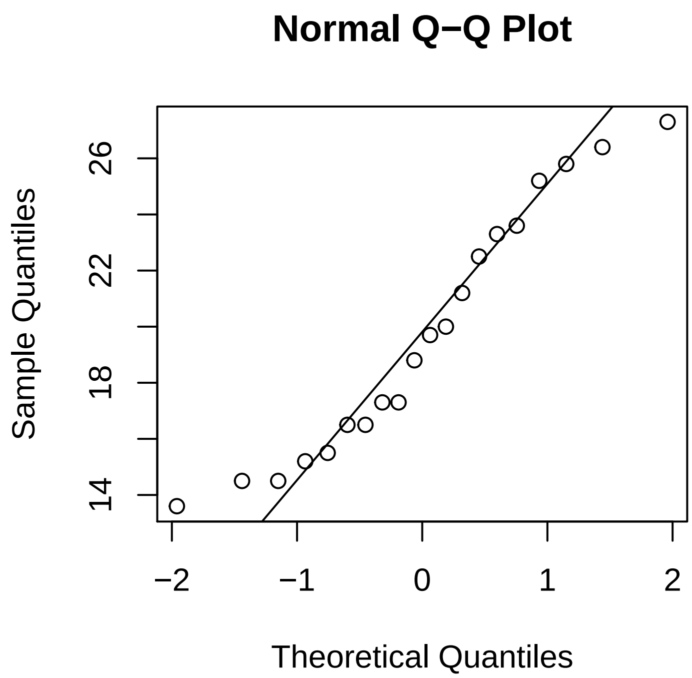


Figure 5: Q-Q plot for response corresponding to 1 mg dose.