Chapter 8: The Analysis of Variance (ANOVA) Comparisons of Several Population Means

1. ANOVA model
2. ANOVA table and F-test
3. Pairwise comparisons of means (ANOVA setting)
4. Checking ANOVA Assumptions
5. Kruskal-Wallis Test (nonparametric test)
6. Transformations for One Way ANOVA
7. Power for a One-Way Model (Overall F-test)

Examples:

1. Rice Example: One-Way ANOVA
2. Poppies Example: Transformations in the One-Way Model
3. Power in a One-Way

Rice Example: (See "**Rice: One Way ANOVA**" R example)
Goal is to compare effects of 4 acids on the growth of rice seedlings.
t = 4 trts = 4 acids: control, acetic, proponic, butyric
n_T = 20 dishes (with shoots) randomly assigned to trts (n = 5 dishes/acid).

Let y_{ij} = dry wt. after seven days for j^th dish of the i^th treatment (acid),
i=1,..,t=4, j=1,...,n_i=n=5

Let: $\bar{y}_{1.}, \bar{y}_{2.}, \bar{y}_{3.}, \bar{y}_{4.}$ be the sample means
and $\mu_1, \mu_2, \mu_3, \mu_4$ be the population means (unknown)

We could use two-sample t-tests to compare the groups, but there are some problems with this idea:

1. In order to test all possible pairs, we would need to run 6 tests. Inconvenient!
2. We get many different estimated pooled variances.
3. Multiple testing problem. We will discuss this in Chapter 9.

1. ANOVA Model

Two equivalent model statements for One-Way ANOVA

1. **Means Model (No Intercept)** : $y_{ij} = \mu_i + \varepsilon_{ij}$

   μ_i = population group/trt means

   $H_0 \; : \mu_1 = \mu_2 = \mu_3 = \mu_4$

   H_A : one or more difference exists

2. **Effects Model (Default)** : $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$

   μ=reference or intercept, α_i = group/trt effects (NOT Type I error!)

   $H_0 \; : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

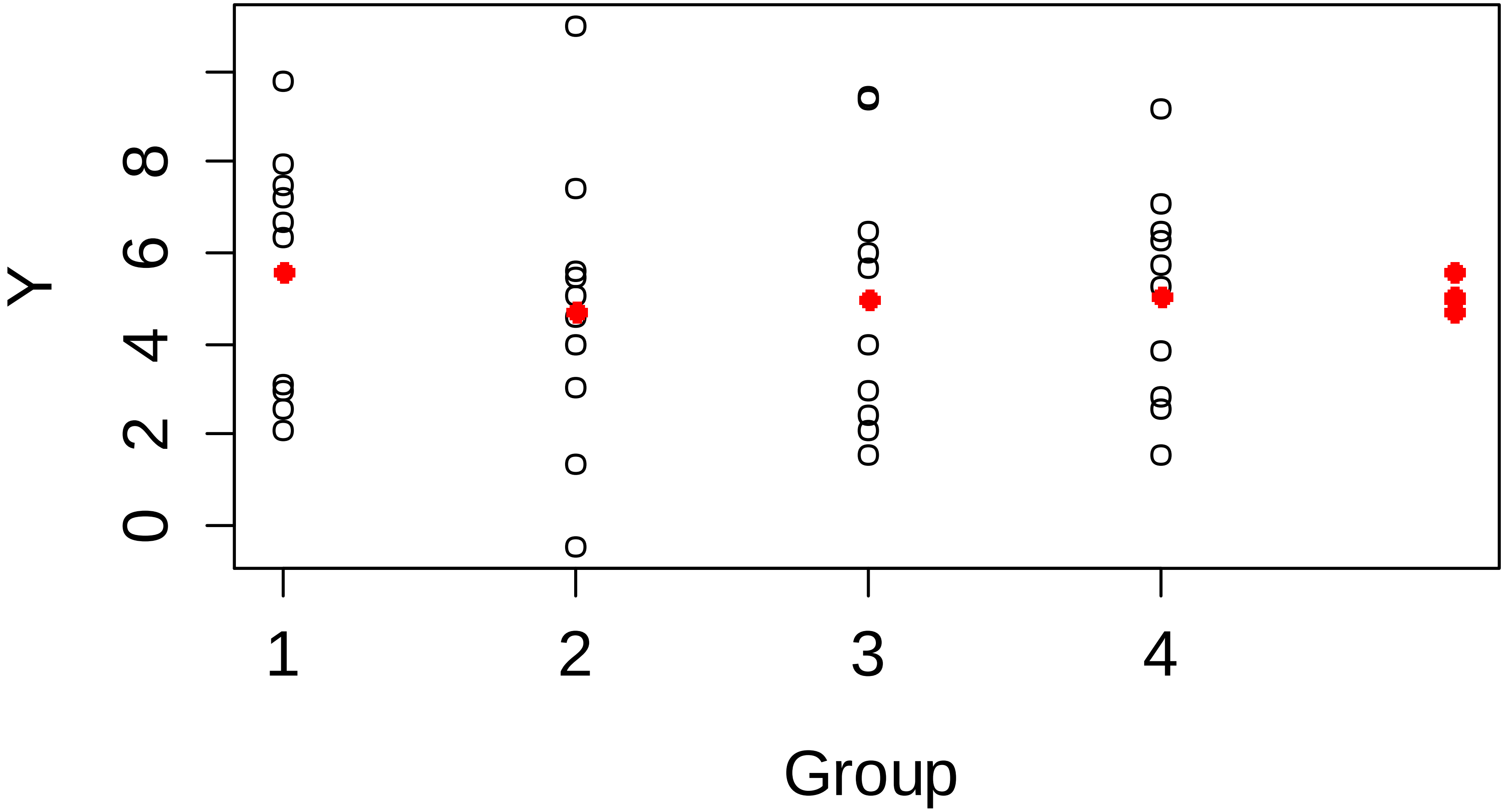
   H_A : one or more α_i is not zero

We will use the Analysis of Variance (ANOVA) F-test to test the equality of means. This test has a NON-directional alternative.

The Idea of ANOVA

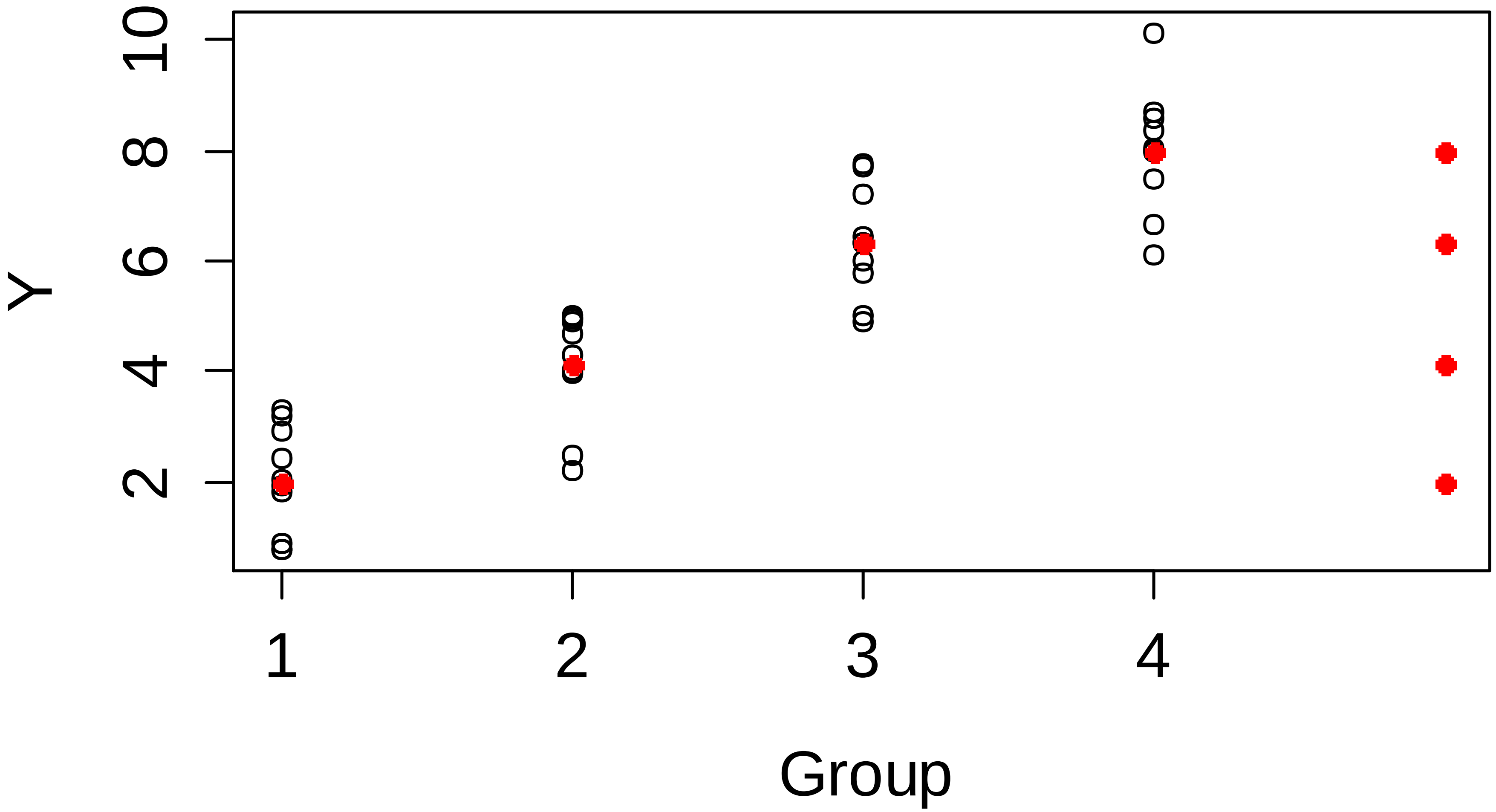
Example of True H_0 ($\mu_1=\mu_2=\mu_3=\mu_4$)

Variation between group means is not large compared to variation within groups.



Example of False H_0

Variation between group means is large compared to variation within groups.



The Idea of ANOVA

If H_0 is true (using either notation), then the model may be written:

$$y_{ij} = \mu + \varepsilon_{ij} \quad \text{(where } \mu \text{ is the common mean)}$$

We test H_0 by forming an F-ratio based on <u>two</u> estimates of variance:

$$F = \frac{s_B^2}{s_W^2} = \frac{MS\mathrm{Trt}}{MS\mathrm{Resid}}, \quad \text{where}$$

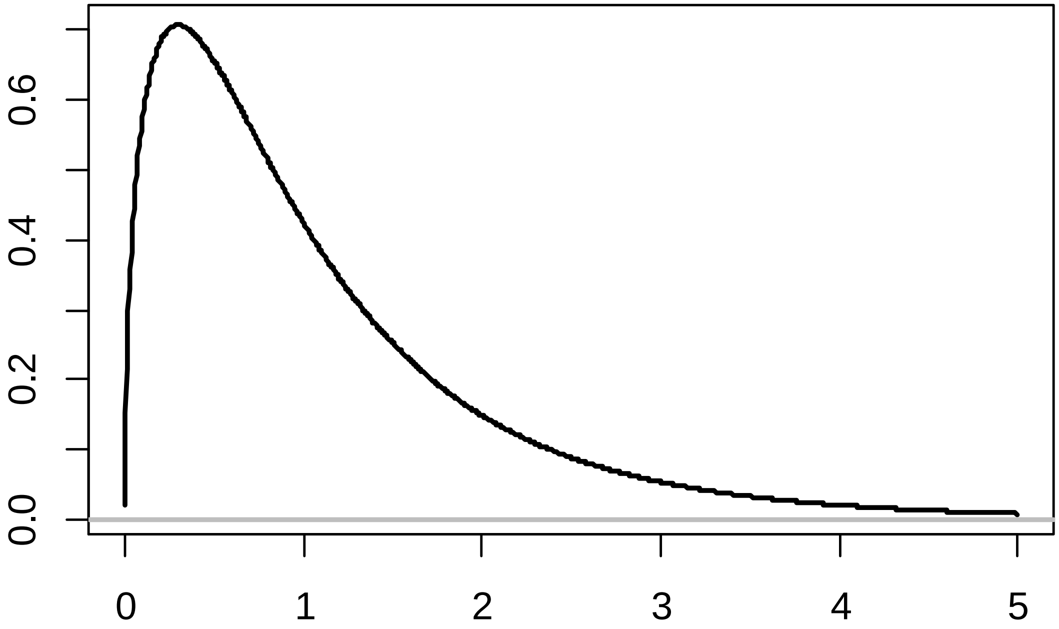
s_W^2 is an estimate of σ^2 formed by pooling sample variances.

s_B^2 is an estimate of σ^2 formed using only only sample means.

$$s_W^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_t - 1)s_t^2}{n_T - t} = \frac{SS\mathrm{Resid}}{n_T - t} = MS\mathrm{Resid}$$

$$s_B^2 = \frac{\sum_{i=1}^{t} n_i(\bar{y}_{i.} - \bar{y}_{..})^2}{t - 1} = \frac{SS\mathrm{Trt}}{t - 1} = MS\mathrm{Trt}$$

- s_W^2 is a pooled estimate of the variance σ^2 that is valid whether H_0 is true or not, because it is based on differences <u>within</u> treatment groups.
- s_B^2 is an estimate of $\sigma^2 + [\Sigma n_i(\mu_i-\mu)^2]/(t-1)$ where $\mu = (\Sigma n_i\mu_i)/n_T$. Hence, when H_0 is true, s_B^2 is a valid estimate of σ^2. If H_0 is false, then s_B^2 will tend to be <u>too big</u> (because it is "contaminated" by mean differences).
- If H_0 is true, the F-statistic will be distributed with the F-distribution with df_1=t-1 and $df_2=n_T$-t, where t is # of trts, n_T is the <u>total</u> sample size.



df1=3, df2=16

$F_{0.05, 3, 16} = 3.24$

2. The ANOVA table and F-test

We can partition the sums of squares:

$$\sum_{ij}(y_{ij}-\bar{y}_{..})^2 \quad = \quad \sum_{i} n_i(\bar{y}_{i.}-\bar{y}_{..})^2 \; + \; \sum_{ij}(y_{ij}-\bar{y}_{i.})^2$$

SSTotal SS Trt SS Resid

(corrected for the mean)

Source	SS	df	MS = SS/df	F-test
Trt	*SSTrt*	t-1	s_B^2 = *SSTrt/(t-1)*	*F=MSTrt/MSResid*
Resid	*SSResid*	n_T-t	s_W^2 = *SSResid/(*n_T*-t)*	
Total	*SSTotal*	n_T-1		

ANOVA F-test for the Equality of Means

Assumptions: Random sample, independent observations, normally distributed residuals, equality of variances.

$H_0 \ : \mu_1 = \mu_2 = \cdots = \mu_t$

$H_A : \text{one or more difference exists}$

Test Statistic: $F = \dfrac{MSTrt}{MS\,\mathrm{Re}\,sid}$

Reject $\mathbf{H_0}$ if $F > F_{\alpha,\, ndf,\, ddf}$

where df1=ndf=t-1 (df Trt) and df2=ddf=n_T-t (df Resid)

P-value is calculated as the area to the right of F (test statistic).

NOTE: MSResid = MSError = MSWithin = $s_W^2 = \hat{\sigma}^2$

MSTrt = MSBetween

Rice Example: One Way ANOVA

- In R, use lm() or aov(). See “**Rice: One Way ANOVA**” example.
- **NOTE:** Be sure “trt” variable is a factor! Check using str().
- In Rcmdr, choose Statistics -> Means -> One-Way ANOVA.

```
> OneWayFit <- lm(weight ~ trt, data = rice)
> anova(OneWayFit)
            Df Sum Sq Mean Sq F value   Pr(>F)
trt          3 1.2199  0.4066   103.5 1.08e-10 ***
Residuals   16 0.0628  0.0039
```

Reject H_0 because p-value < 0.0001
Can also compare test statistic to table value and get the same conclusion: Reject H_0 because calculated F= 103.5 > $F_{0.05, 3, 16}$ =3.24.
Based on the P-value for the F-test we conclude that there is extremely strong evidence indicating differences among the treatment means.

But which means are different?

The linear model **lm()** function can be used to fit a broad class of linear models. This includes one-way ANOVA (this chapter) and simple linear regression (CH11) and many others. Again, be sure trt is defined as a factor! This can be done using str().

The analysis of variance **aov()** function is a “wrapper” for the lm() function.

From ?aov:

- The main difference between lm() and aov() is the way print, summary, etc handle the fit. aov() is expressed in the traditional language of ANOVA rather than that of linear models.
- aov() is designed for balanced designs, and the results can be hard to interpret without balance. Beware that missing values in the response will likely lose the balance.

Note: In STAT512, we will use lm() and Anova() function from the car package!

A note about ANOVA vs t-test when there are t=2 groups:

An ANOVA with t=2 groups is equivalent to running a two-sample t-test assuming equal variances.

The p-values testing H0: $\mu_1=\mu_2$ will be identical.

The test statistics are related as follows $F = t^2$.

Note that F-test is non-directional where as the t-test can easily accommodate a one-sided alternative.

3. Pairwise Comparison of Means (in ANOVA setting)

We will discuss this topic more in Chapter 9. Here we just look at one method: the **Least Significant Difference (LSD) method**. These are **unadjusted** comparisons, meaning **not adjusted for multiple testing**.

Assumptions: Random sample, independent observations, normally distributed residuals, equal variances.

$H_0 : \mu_i = \mu_j \quad 1 \le \text{i, j} \le \text{t (any two)}$

$H_A : \mu_i \neq \mu_j$

Test Statistic:

$$t = \frac{\bar{y}_{i.} - \bar{y}_{j.}}{s_W \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}$$

df = n_T-t = dfResid

Reject H0 if $|t| > t_{\alpha/2}$

NOTE: $s_W = \hat{\sigma} = \sqrt{MS\,\mathrm{Re}\,sid}$

Confidence Interval for the difference between means

Assumptions: Random sample, independent observations, normally distributed residuals, equal variances

$$(\bar{y}_i - \bar{y}_j) \pm t_{\alpha/2} s_W \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

If n$_i$=n$_j$=n (group sizes are equal) the ME reduces to:

$$ME = LSD = t_{\alpha/2} s_W \sqrt{\frac{2}{n}}$$

df = n$_T$-t = df Resid

Recall: $s_W = \hat{\sigma} = \sqrt{MS\text{Resid}}$

Note that we will Reject H$_0$ (μ_i-μ_j=0), whenever $|\bar{y}_i - \bar{y}_j| > ME$ (because 0 will not be contained in the CI).

So, when group sizes are equal, we will Reject H0 for any pair (i,j) where $|\bar{y}_i - \bar{y}_j| > LSD = ME$.

Rice Example: Pairwise comparisons

t = 4 treatment groups. All $n_i = n = 5$

df = dfResid = n_T - t = 20 - 4 =16. $t_{\alpha/2}$=2.120 for α=0.05

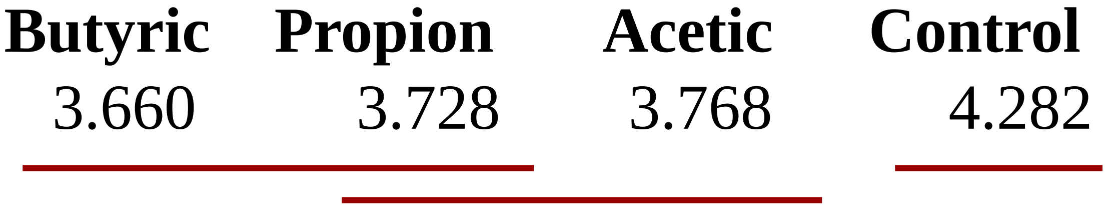
Two treatment means are significantly different if:

$$\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > t_{\alpha/2} s_W \sqrt{\frac{2}{n}} = (2.120)\sqrt{0.00393}\sqrt{\frac{2}{5}} = 0.0840$$

$$95\%\text{ME} = \text{LSD}_{0.05} = 0.0840$$

The result is sometimes reported in a display.

Order the means first:



Butyric 3.660 **Propion** 3.728 **Acetic** 3.768 **Control** 4.282

Then connect any two means that are not significantly different with a line (or group with the same letter or number).

Rice Example: Example CI

$\text{A 95\% CI for } \mu_{\text{Control}} - \mu_{\text{Acetic}}:$

$$\bar{y}_i - \bar{y}_j \quad \pm \quad t_{\alpha/2}\, s_W \sqrt{\frac{2}{n}}$$

$$(4.282 - 3.768) \;\pm\; 0.0840$$

$$(0.430,\;\; 0.598)$$

Notes on Pairwise Comparison of Means in the One-Way ANOVA setting

1. In R, we will use the emmeans() function from the emmeans package to get p-values for all pairwise comparisons . See “**Rice One Way ANOVA**” example.
2. For now, we will use the unadjusted pairwise comparisons. But in practice Tukey adjustment is often used to correct for multiple testing (see CH9 notes).
3. Computer programs and publications often put letter or number superscripts next to members of the same LSD group (rather than underlines). In R, we can use the CLD() function to do this.
4. If there are unequal sample sizes for groups, then statistical significance is based on magnitude of difference and sample size. For this reason, there is no fixed LSD (=ME) value when sample sizes are unequal. Resulting p-values, CIs, etc are fine. But “lines”/CLD display should be interpreted with caution.

Rice Example: simple summary statistics vs emmeans()

```
> SumStats
      trt n  mean          sd         se
1  acetic 5 3.768 0.06140033 0.02745906
2 butyric 5 3.660 0.06442049 0.02880972
3 control 5 4.282 0.06058052 0.02709243
4 propion 5 3.728 0.06418723 0.02870540

> library(emmeans)
> LMFit <- lm(weight ~ trt, data = rice)
> emmeans(LMFit, pairwise ~ trt, adjust = "none")
$emmeans
 trt     emmean         SE df lower.CL upper.CL
 acetic   3.768 0.02802677 16 3.708586 3.827414
 butyric  3.660 0.02802677 16 3.600586 3.719414
 control  4.282 0.02802677 16 4.222586 4.341414
 propion  3.728 0.02802677 16 3.668586 3.787414
```

“Simple summary” statistics vs emmeans()

For one-way ANOVA the simple means and emmeans will be the same, even if the sample sizes are not balanced across groups.

However, there will be a difference between the “simple” SE and the model based SE:
“Simple” SE for group i $= {}^{s_i}/_{\sqrt{n_i}}$
Model based SE for group i = ${}^{s_W}/_{\sqrt{n_i}}$

The difference is that the simple SE allows standard deviation to be estimated separately for each group.
While the model based SE uses a pooled estimate (sw). This makes sense because the ANOVA model assumes equal variance.

Note that if sample sizes are equal, the model based SE will be the same for all groups. This is true for the Rice example.

Rice Example: Pairwise Comparisons using emmeans()

```
> library(emmeans)
> LMFit <- lm(weight ~ trt, data = rice)
> emout <- emmeans(LMFit, pairwise ~ trt, adjust = "none")
> emout
 $contrasts
 contrast          estimate         SE df t.ratio p.value
 acetic - butyric     0.108 0.03963584 16   2.725  0.0150
 acetic - control    -0.514 0.03963584 16 -12.968  <.0001
 acetic - propion     0.040 0.03963584 16   1.009  0.3279
 butyric - control   -0.622 0.03963584 16 -15.693  <.0001
 butyric - propion   -0.068 0.03963584 16  -1.716  0.1055
 control - propion    0.554 0.03963584 16  13.977  <.0001

> CLD(emout$emmeans, adjust = "none")
 trt     emmean         SE df lower.CL upper.CL .group
 butyric  3.660 0.02802677 16 3.600586 3.719414  1
 propion  3.728 0.02802677 16 3.668586 3.787414  12
 acetic   3.768 0.02802677 16 3.708586 3.827414   2
 control  4.282 0.02802677 16 4.222586 4.341414    3
```

4. Checking the ANOVA Assumptions

1. Random sample, independent observations

2. Residuals are normally distributed: $\varepsilon_{ij} \sim N(0, \sigma^2)$
   QQ plot of residuals
   Test of normality for residuals

3. Equality (Homogeneity) of variances: $\text{Var}(\varepsilon_{ij}) = \sigma^2$
   Plot of residuals vs fitted (predicted) values
   Levene’s test

We will be using the **residuals** to check assumptions 2 and 3.
Residuals are calculated as $e_{ij} = y_{ij} - \bar{y}_{i.}$
Fitted/predicted values $\widehat{\mu}_i = \overline{y_i}$ are the sample means.

Note: Both diagnostic plots (and more) can be generated in R by applying the plot() function to a lm or aov object.

Checking for Normality of Residuals:

1. A Q-Q plot of the residuals is a useful graphical tool for checking normality. The Q-Q plot is generally more useful than a histogram of the residuals for this purpose.
2. Tests of normality can be used (ex: Shapiro-Wilks test).

Important Note: When you check the data for normality, check the residuals, not the observations themselves. The errors (i.e., deviations from group means) are assumed to follow a common normal distribution but the observations themselves may come from *different* normal distributions. So, when the means are very different, the combined data will often look non-normal, even when the residuals are close to normal.

Checking for Equality of Variance (Homogeneity of Variance):

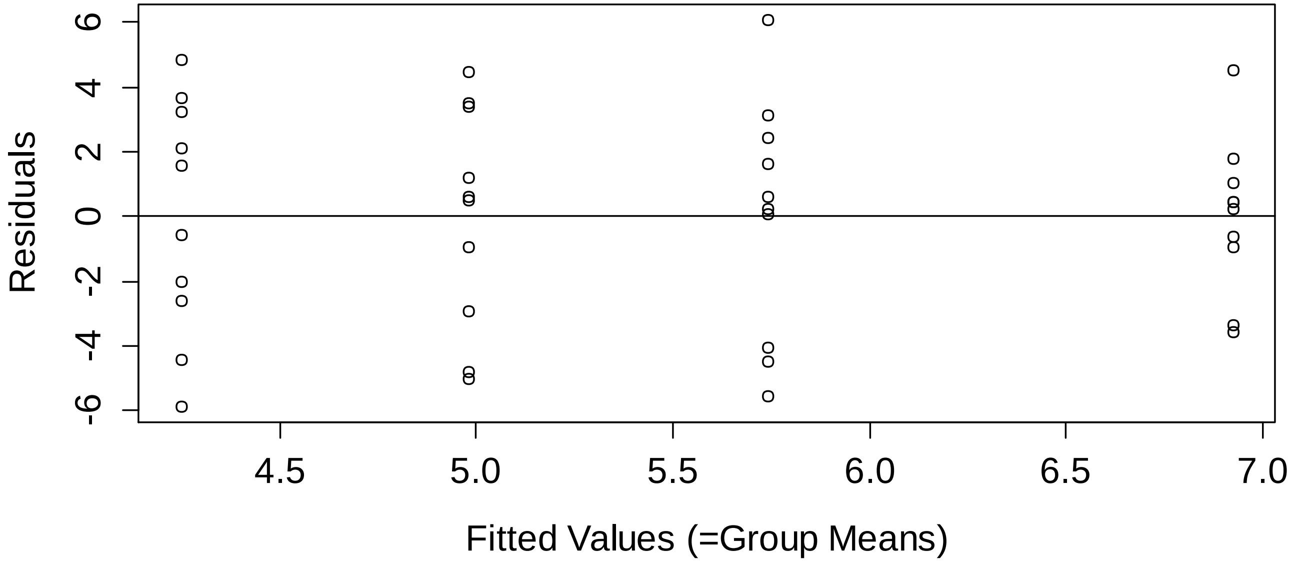
Check plot of residuals versus predicted values (or fitted values). Plots are usually more helpful than the formal tests.

The predicted/fitted value in the i^{th} group is: $\bar{y}_{i.} = \hat{\mu}_i$

The residual for the i,j obs. $= y_{ij} - \bar{y}_{i.}$

Primarily interested in checking for equal "scatter" (representative of equal variance), but sometimes we can also detect "skew" and/or outliers in the residual diagnostic plot.

Megaphone shape is common when assumption of equal variances is NOT met. See slide 25.



Levene’s Test for homogeneity of variances

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_t^2$$

$$H_A : \text{one or more is different}$$

Levene’s test is a generally satisfactory test (even in the absence of normality) and can be done in R.

In R, use leveneTest() from the car package.
In Rcmdr, choose Statistics -> Variances -> Levene’s test.
See “**One Way ANOVA**” example.

Test is based on ANOVA on the absolute differences from the group means: $Z_{ij} = \left| y_{ij} - \bar{y}_{i.} \right|$

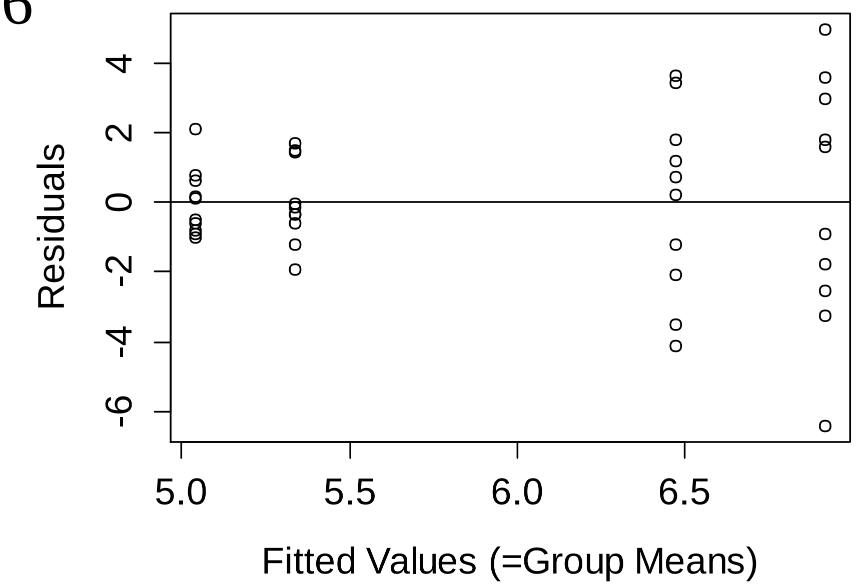
The default in R is to use absolute deviations from the median (center=“median”) sometimes called Brown-Forsythe method, this method is considered to be more robust.
To get the classic Levene’s test, use center=“mean”.

Comments about ANOVA assumptions

1. The ANOVA F-test is "robust" to the assumptions of normality and equal variances. This is especially true when group sizes are equal.
2. Pairwise comparisons of means is not accurate if the variances are not close to equal.

Example: s_w = sqrt(MSResid) = 1.66

```
Group  n    mean       sd
     1 10 5.041818 0.9711052
     2 10 5.334920 1.2085462
     3 10 6.470208 2.6851316
     4 10 6.918262 3.5500434
```



The "pooled" standard deviation overestimates the standard deviation for groups 1 and 2, and underestimates the standard deviation for groups 3 and 4.

5. The Kruskal-Wallis Test

The Kruskal-Wallis test is a non-parametric alternative to the one-way ANOVA F-test.
When the normality of the errors is in doubt, a rank based test is a reasonable alternative. All observations are sorted by size, ranks assigned (n_T for largest, 1 for smallest). The sum of the ranks for the i^{th} group is T_i.

H_0 : the t groups have identical distibutions

H_A : not all of the distribution are the same

$$\text{T.S.}: \quad H = \frac{12}{n_T(n_T+1)}\left(\sum_i \frac{T_i^2}{n_i}\right) - 3(n_T+1)$$

$R.R$: Reject if H exceeds χ^2_α with df = t -1

(See O & L for modification to the fomula when there are ties.)

Notes on Kruskal-Wallis Test

1. The Kruskal-Wallis test is an extension of the Wilcoxon Rank Sum test for more than two groups.
2. In R, use kruskal.test().

   In Rcmdr, choose Statistics -> Nonparametric tests -> Kruskal-Wallis test.
3. It assumes identical shaped distributions (which implies homogeneous variances) with possible location differences, but it is often used even when the variances are not equal, and seems to perform acceptably.
4. Often interpreted as a test for differences in medians.
5. Kruskal-Wallis test is less powerful than the ANOVA F-test if ANOVA assumptions are met.
6. Pairwise comparisons can be done using Dunn's test. Multiple testing adjustments can be applied.

6. Transformations for One Way ANOVA

If ANOVA assumptions are not satisfied on the original scale (Y), then you can consider using a transformed response variable.

Transforming the response makes the model harder to interpret, so we don't want to do it unless it's really necessary.

The literature may have examples of transformations that are common in a particular field. **But the way to decide what transformation is appropriate is by fitting the model with the transformed response and checking the diagnostic plots.**

See "**Transformations for One Way ANOVA**"

Common transformations include:

- **Square root:** `YT=sqrt(Y) or YT=(Y)^0.5` (in R)

Example: Count data (calls to switchboard, insects in trap, etc)

- **Power:** `YT=1/Y` or `YT=Y^2` (in R)

- **Log:** `YT =log(Y) or YT=log10(Y)` (in R)

Examples: Chemical concentrations or hormone levels

Important Note: Watch out for y=0 values which will be undefined after log transformation! If this is not properly accounted for, then these values will be treated as "missing". A simple, common solution is to add a small positive constant before log transformation. You might use y_T=log(y+1) when the y's are 0 to 20, and use y_T=log(y+0.01) when the y's are 0 to 0.25. These are just suggestions!

- The use of transformations like YT=Y^(0.75) is rare. Transformations like YT= Y^(0.63) are seldom used.

The **Box-Cox** approach is a systematic method for choosing a transformation. This approach should only be used for y > 0!

The general form of the Box-Cox transformation is:

$$g(y_i) = (y_i^{\lambda} - 1)/\lambda$$

where λ is a constant to be determined from the data.

If λ = 1, then no transformation is needed.
If λ = 2, then model Y^2 (instead of Y).
If λ = -1, then model Y^{-1} =1/Y (instead of Y).
If λ = 0.5, then model $Y^{0.5}$ = sqrt(Y) (instead of Y).
If λ = 0, them model log(Y) (instead of Y)

We will use the boxcox() function from the MASS package to create a Box-Cox plot to choose λ.

Interpretation after transformation

The means of the transformed variables are not the same as the means of the original variables. However, if the means are significantly different based on the analysis in the transformed scale, it is reasonable to conclude that the means in the original scale are also significantly different.

What should a researcher report when it was necessary to do the ANOVA using a transformed scale?

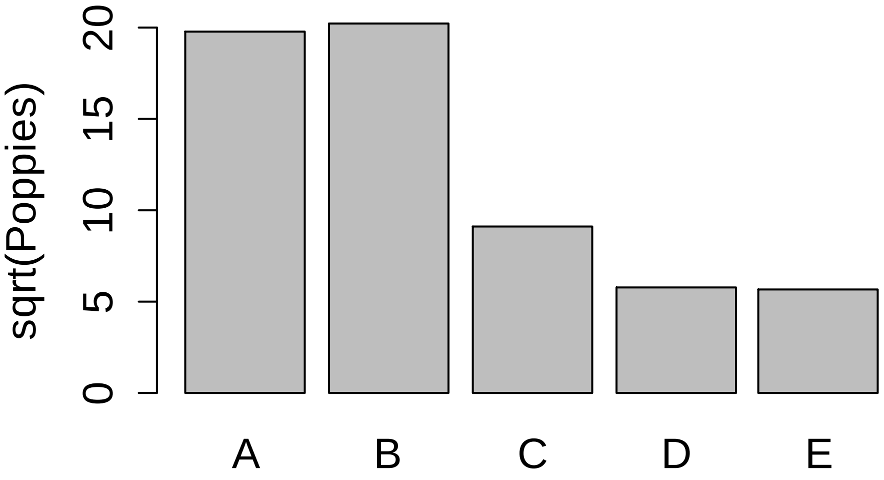
Three options:

1. Present the means on the transformed scale.
2. Present the means on the original scale, but note that analysis was done on transformed data.
3. Back transform to the original scale.

Option 1: Present the means (or graphs) and the comparison of means in the transformed scale. This is a reasonable option when the scale is common in that field of study (e.g. log in chemistry, sqrt in radiological sciences.)

Example: Poppy plants in oats (see R example).

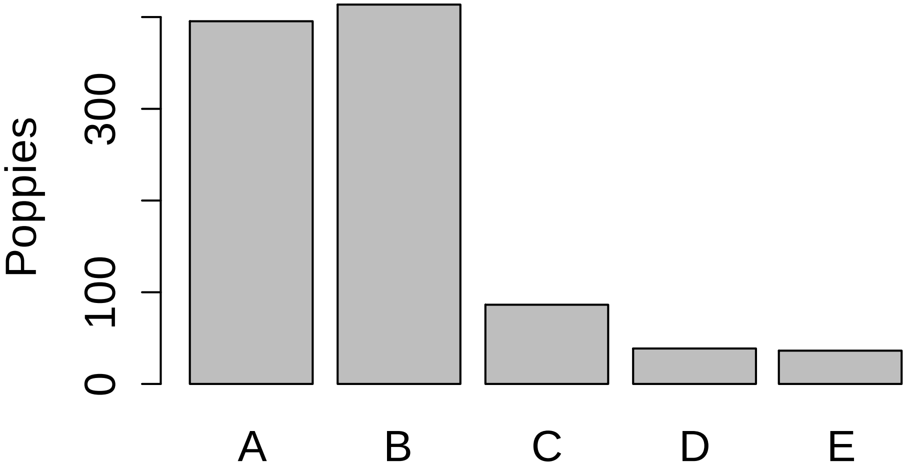
Trt	Orig Scale	**Sqrt Scale**	Back Transformed
A	394.75	19.83	393.09
B	413.00	20.23	409.07
C	86.75	9.08	82.47
D	37.75	5.75	33.11
E	35.25	5.64	31.89



Option 2: Present the means (or graphs) in the original scale, but with the comparison of means based on the transformed scale. Describe what you have done in your methods section and in a footnote to the table or graph of means.

A LSD value based on the transformed scale cannot be used on means based on the original scale. Also, means that are farther apart will not necessarily be more significantly different than means that are closer together.

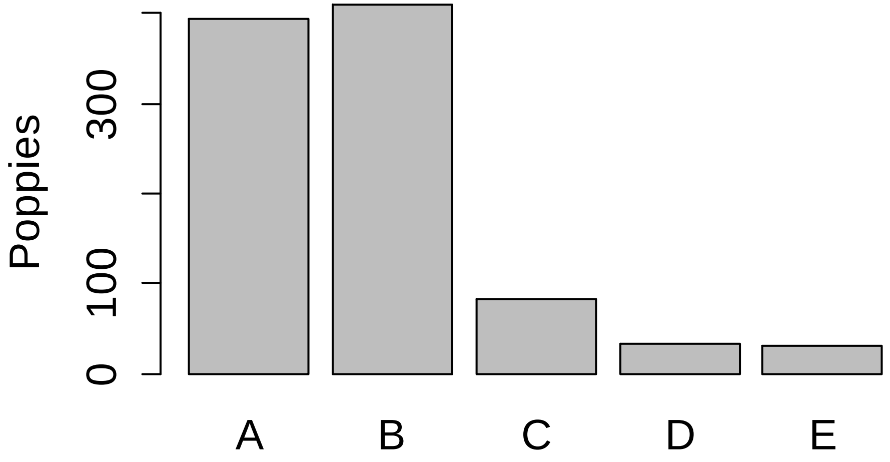
Trt	**Orig Scale**	Sqrt Scale	Back Transformed
A	394.75	19.83	393.09
B	413.00	20.23	409.07
C	86.75	9.08	82.47
D	37.75	5.75	33.11
E	35.25	5.64	31.89



Option 3: Compute the means using the transformed variable, but "back-transform" the means before presentation. To back-transform a square root transformed variable, square the mean.

Backtransformed means can be substantially lower than the means in the original scale, particularly when the transformation is log.

Trt	Orig Scale	Sqrt Scale	**Back Transformed**
A	394.75	19.83	393.09
B	413.00	20.23	409.07
C	86.75	9.08	82.47
D	37.75	5.75	33.11
E	35.25	5.64	31.89



Comments about Interpretation after log Transformation:

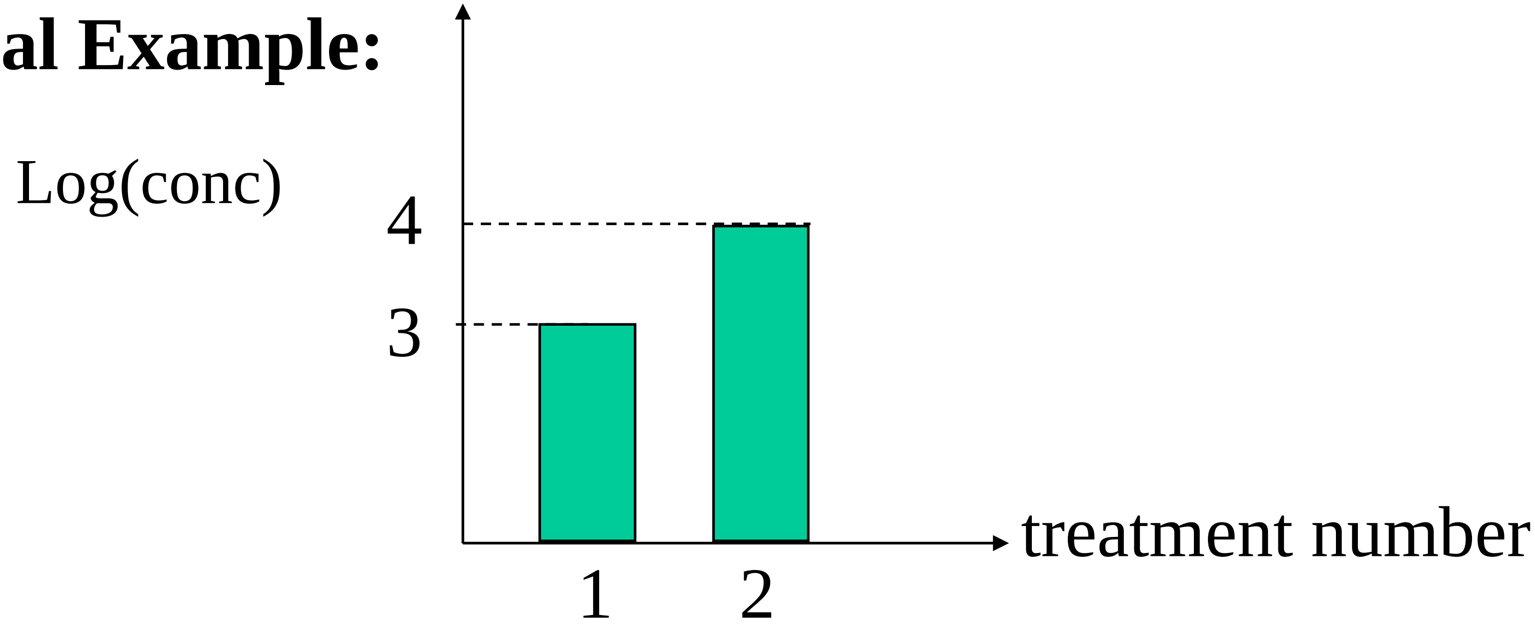
There can be interpretational advantages to doing an analysis in the log scale. Because log(x/y) = log(x) - log(y), differences between means in the log analysis can be interpreted as ratios in the original scale.

In bioinformatics, it is common to use a log2 transformation to satisfy ANOVA assumptions, but also because the differences can be interpreted as log2 fold change (FC) values.

log2(2/1) = +1, log2(½) = -1

When using a log transformation, it is common to see people back transform the difference to a "ratio" scale after analysis.

A Lognormal Example:



$$\text{Est. trt diff (log scale)} = \bar{x}_1 - \bar{x}_2 = 3 - 4 = -1$$

$$\text{Ratio of trt means (orig. scale)} = \frac{\text{trt 1 mean}}{\text{trt 2 mean}} = e^{\mu_1 - \mu_2}$$

$$e^{\bar{x}_1 - \bar{x}_2} = e^{3-4} = e^{-1} = 1/2.718 = 0.368 \approx 37\ \%$$

Trt1 concentration is 37% Trt2 concentration!

A C.I. for this % is given by:

$$\left(e^{LCL}, e^{UCL}\right)$$

(where (LCL,UCL) is a C.I. for the difference in the means of the log data)

7. Sample Size and Power in the ANOVA F-test

As with all of the other sample size calculations, we need:

1) A conjecture about the within-group standard deviation σ.
2) Identification of the true alternative that we want to detect: conjectures for $\mu_1, \mu_2, \ldots, \mu_t$.

Power is a function of the "noncentrality" parameter for the F-distribution, given the alternative:

$$\lambda = \frac{n\sum_{i=1}^{t}\left(\mu_i - \bar{\mu}_{.}\right)^2}{\sigma^2}$$

Power increases as sample size and the differences among the true means increase, and decreases as the error standard deviation increases.

Power can be computed in R using the **pf** function:

```
fcrit=qf(0.95, dfn, dfd)
power = 1 - pf(fcrit, dfn, dfd, lambda)
```

Power can be computed using `power.anova.test()`
See: "**Power for the One-Way Model**" Example

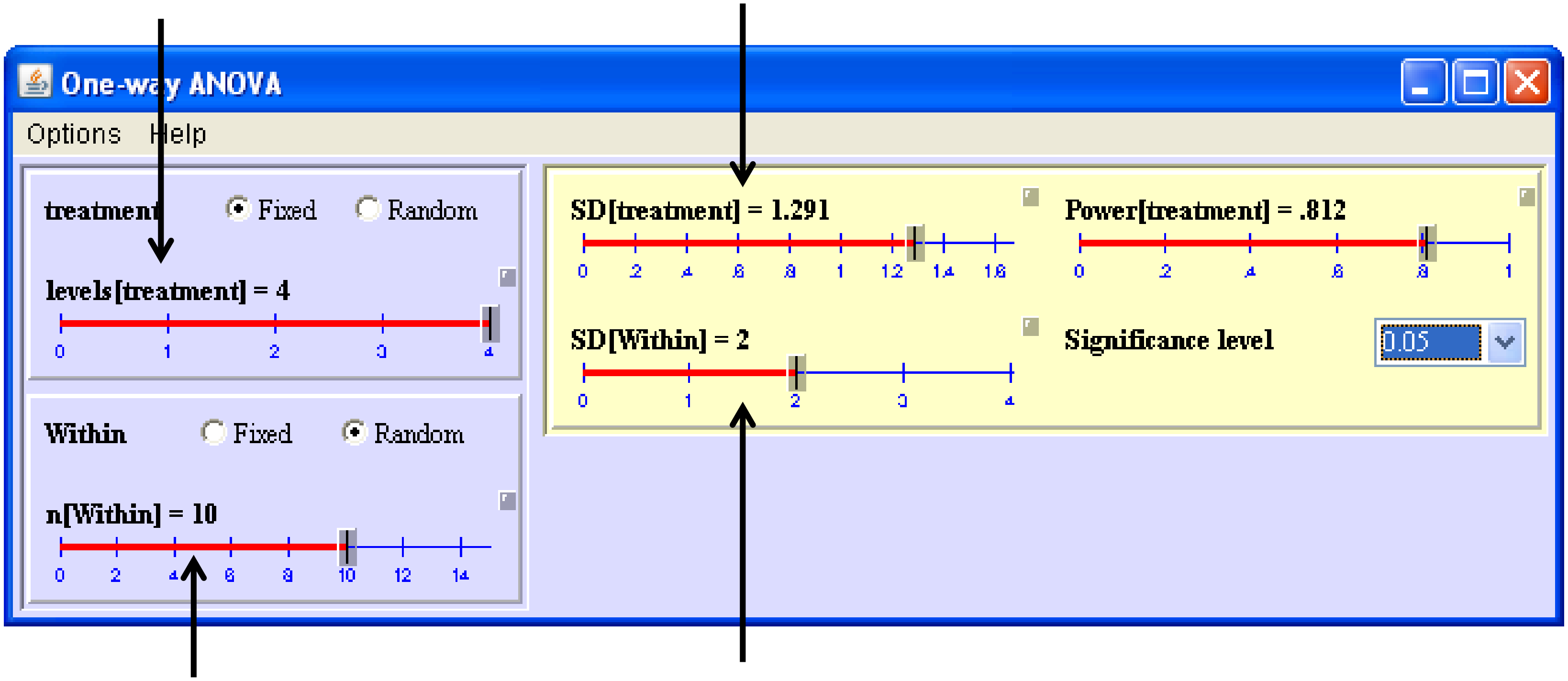
Power for ANOVA F-test using Lenth

http://homepage.stat.uiowa.edu/~rlenth/Power/

* Choose Balanced ANOVA, then Built-in models = "One Way ANOVA", Study the Power of = "F-tests". (We will discuss the Differences/Contrasts option in the Ch9 notes.)

t = # trts

sd of trt means = sd between



One-way ANOVA

Options Help

treatment ● Fixed ○ Random

levels[treatment] = 4

Within ○ Fixed ● Random

n[Within] = 10

SD[treatment] = 1.291

Power[treatment] = .812

SD[Within] = 2

Significance level 0.05

n = n per group

sd within group = sd within