The experiment: *One Red clover plant* is grown in each of *25 pots* (one plant per pot). Each *plant (or pot)* is inoculted from a culture of one of *5 different Rhizobium strains*. Plants are incubated for *two weeks* and *each Red clover plant* is tested for *nitrogen content*.

Ask yourself , "What is the dependent variable and what are the treatments?"

        1) What is the experimental unit?

        2) What is the sampling unit?

        3) Are the treatments fixed or random?

1) H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu$

2) H_1: some μ_i is different

3)   a) Assume that the observations are normally distributed about each mean, or that the residuals (i.e. deviations) are normally distributed.

       b) Assume that the observations are independent

       c) Assume that the variances are homogeneous

4) Set the level of type I error. Usually $\alpha = 0.05$

5) Determine the critical value. The test is an ANOVA (F test). There are t=5 treatments (the Rhizobium strains) so the numerator has t-1 = 4 d.f. Each treatment has n=5 observations. The error degrees of freedom will be t(n-1) = 5(4) = 20 d.f. $F_{\alpha=0.05, 4, 20 \text{ d.f.}} = 2.8661$.

6) Obtain data and evaluate.

The raw data for this experiment is given below.

Obs	Treatment 3DOk1	3DOk4	3DOk5	3DOk7	3DOk13
1	19.4	17.0	17.7	20.7	14.3
2	32.6	19.4	24.8	21.0	14.4
3	27.0	9.1	27.9	20.5	11.8
4	32.1	11.9	25.2	18.8	11.6
5	33.0	15.8	24.3	18.6	14.2

The calculations needed for the ANOVA are as follows.

	3DOk1	3DOk4	3DOk5	3DOk7	3DOk13	Sums	Grand Mean
Sum	144.1	73.2	119.9	99.6	66.3	503.1	20.124
Mean	28.8	14.6	24.0	19.9	13.3		Correction Factor
n	5	5	5	5	5	25	10124.3844
Var	33.642	16.943	14.267	1.277	2.038		MSE
CSS	134.568	67.772	57.068	5.108	8.152	272.668	13.6334
							CSS Treatment
Sum²/n	4152.96	1071.65	2875.20	1984.03	879.14	10962.98	838.5976
Mean²*n	4152.96	1071.65	2875.20	1984.03	879.14	10962.98	838.5976

The treatment sum of squares, as developed by Fisher, are converted to a "variance" and tested with an F test against the pooled error variance. In practice, the sum of squares are usually calculated and presented with the degrees of freedom in a table called an ANOVA table.

The uncorrected SS for treatments is $USS_{Treatments} = \dfrac{\sum\limits_{i=1}^{t}\left(\sum\limits_{j=1}^{n} Y_{ij}\right)^2}{n} = n\sum\limits_{i=1}^{t}\left(\dfrac{\sum\limits_{j=1}^{n} Y_{ij}}{n}\right)^2 =$

$\dfrac{144.1^2}{5}+\dfrac{73.2^2}{5}+\dfrac{119.9^2}{5}+\dfrac{99.6^2}{5}+\dfrac{66.3^2}{5} = 4152.96+1071.65+2875.20+1984.03+879.14=10962.98$

The uncorrected SS for the total $SS_{Total} = \sum\limits_{i}\sum\limits_{j} Y_{ij}^2 = 19.4^2+17.0^2+17.7^2+\ldots+14.3^2=11235.65$

The correction factor for both terms is $CF = \dfrac{\left(\sum\limits_{i}^{t}\sum\limits_{j}^{n} Y_{ij}\right)^2}{tn} = \dfrac{503.1^2}{25} = 10124.3844$

The same correction factor is used for both the SSTotal and SSTreatments, so

    a) The corrected SSTotal = 11235.65 - 10124.3844 = 1111.2656

    b) The corrected SSTreatment = 10962.98 - 10124.3844 = 838.5976

The error term (SSError) is calculated by either;

    a) subtracting the Uncorrected SSTotal from the Uncorrected SSTreatment.

    b) subtracting the Corrected SSTotal from the Corrected SSTreatment.

    c) the pooled within group variance $SSError = S_p^2 = \dfrac{SS_1 + SS_2 + SS_3 + SS_4 + SS_5}{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5}$.

        = 1111.2656 - 838.5976 = 11235.65 - 10962.98 = 272.668

The Analysis of Variance done in EXCEL is given below.

ANOVA: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
3DOk1	5	144.1	28.82	33.642
3DOk4	5	73.2	14.64	16.943
3DOk5	5	119.9	23.98	14.267
3DOk7	5	99.6	19.92	1.277
3DOk13	5	66.3	13.26	2.038

ANOVA Table

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	838.5976	4	209.6494	15.3776	0.00000676	2.8661
+Within Groups	272.6680	20	13.6334			
Total	1111.2656	24				

Our ANOVA analyses will be done with PROC MIXED and PROC GLM. There is a PROC ANOVA, but it is a subset of PROC GLM.

Homogeniety of Variance

Your textbook discusses one test by Hartley. It is one of the simplest tests, but not usually the best. To do this test we calculate the largest observed variance divided by the smallest observed variance. This statistics is tested with a special table by Hartley (Appendix Table 5.A in your Freund & Wilson textbook). For our experiment above the variances were 33.642, 16.943, 14.267, 1.277 and 2.038. Calculate ratio of the largest individual treatment variance to the smallest variance. $F_{max}=S^2_{max}/S^2_{min}=^{33.642}/_{1.277}=26.34455756$. The critical values are 3.52 for $\alpha=0.05$ and 4.6 for $\alpha=0.01$. The variances do not appear homogeneous in this case.

A number of other tests are available in SAS, but only for a simple CRD (i. e. a One-way ANOVA). These test are briefly discussed below.

To get all of the tests available in SAS, use the following statement following PROC GLM.

```
 MEANS your_treatment_name / HOVTEST=BARTLETT HOVTEST=BF
     HOVTEST=LEVENE(TYPE=ABS) HOVTEST=LEVENE(TYPE=SQUARE)
     HOVTEST=OBRIEN WELCH;
```

To get all of the tests available in SAS, use the following statement following PROC GLM.

Levene's Test: *This test is basically an ANOVA of the squared deviations* (**TYPE=SQUARE**)*. It can also be done with absolute values* (**TYPE=ABS**)*. This is one of the most popular HOV tests.*

O'Brien's Test: *This test is a modification of Levene's with an additional adjustment for kurtosis.*

Brown and Forsythe's Test: *This test is similar to Levene's, but uses absolute deviations from the* <u>*median*</u> *instead of more ANOVA like means. There is a "nonparametric" ANOVA that employs deviations from the median instead of the usual deviations from the mean used for the normal ANOVA.*

Bartlett's Test for Equality: *This test is similar to Hartley's, but uses a likelihood ratio test instead of an F test. This test can be inaccurate if the data is not normally distributed.*

Welch's ANOVA: *This test is a weighted ANOVA. It is not a test of homogeneity of variance. This ANOVA weights the observations by an inverse function of the variances and is intended to address the problem of non-homogeneous variance and to be use when the variance is not homogeneous.*

```
45        model percent = treatmnt / SS3;
46        MEANS treatmnt / HOVTEST=BARTLETT HOVTEST=BF HOVTEST=LEVENE(TYPE=ABS)
47                         HOVTEST=LEVENE(TYPE=SQUARE) HOVTEST=OBRIEN WELCH;
```
          See SAS OUTPUT

The Homogeniety of Variance (HOV) tests discussed above can be done in SAS (PROC GLM). Note that the last one is NOT an HOV test, it is another type of ANOVA called a weighted ANOVA.

Post-hoc or Post-Anova tests! Once you have found out some treatment(s) are "different", how do you determine which one(s) are different?

If we had done a t-test on the individual pairs of treatments, the test would have been done as

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{MSE \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$. If the difference between $\bar{Y}_1 - \bar{Y}_2$ was large enough, the t

value would have been greater than the $t_{critical}$ and we would conclude that there was a significant

difference between the means. Since we know the value of $t_{critical}$ we could figure out how large a difference is needed for significance for any particular values of MSE, n_1 and n_2. We do this by replacing t with $t_{critical}$ and solving for $\bar{Y}_1 - \bar{Y}_2$.

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{MSE \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}, \text{ so}$$

$$t_{critical} \sqrt{MSE \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \bar{Y}_1 - \bar{Y}_2 \quad \text{or} \quad \bar{Y}_1 - \bar{Y}_2 = t_{critical} S_{\bar{Y}_1 - \bar{Y}_2}$$

This value is the exact width of an interval $\bar{Y}_1 - \bar{Y}_2$ which would give a t-test equal to $t_{critical}$. Any larger values would be "significant" and any smaller values would not. This is called the "Least Significant Difference". $$LSD = t_{critical} S_{\bar{Y}_1 - \bar{Y}_2}$$

This least significant difference calculation can be used to either do pariwise tests on observed differences or to place a confidence interval on observed differences.

The LSD can be done in SAS in one of two ways. The MEANS statement produces a range test (LINES option) or confidence intervals (CLDIFF option), while the LSMEANS statement gives pairwise comparisons.

```
49              means treatmnt / lsd tukey bon scheffe duncan cldiff;
48              means treatmnt / lsd tukey bon scheffe duncan lines;
```
          See SAS OUTPUT

Post-hoc test continued.

The LSD has an α probability of error on each and every test. The whole idea of ANOVA is to give a probability of error that is α for the whole experiment, so, much work in statistics has been dedicated to this problem. Some of the most common and popular alternatives are discussed below. Most of these are also discussed in your textbook.

The LSD is the LEAST conservative of those discussed, meaning it is the one most likely to detect a difference and it is also the one most likely to make a Type I error when it finds a difference. However, since it is unlikely to miss a difference that is real, it is also the most powerful. The probability distribution used to produce the LSD is the t distribution.

Bonferroni's adjustment. Bonferroni pointed out that in doing k tests, eacn at a probability of Ttype I error equal to α, the overall experimentwise probability of Type I error will be NO MORE than $k*\alpha$, where k is the number of tests. Therefore, if we do 7 tests, each at $\alpha=0.05$, the overall rate of error will be NO MORE than =.35, or 35%. So, if we want to do 7 tests and keep an error rate of 5% overall, we can do each individual test at a rate of $\alpha/k = 0.055/7 = 0.007143$. For the 7 tests we have an overall rate of $7*0.007143 = 0.05$. The probability distribution used to produce the LSD is the t distribution.

Duncan's multiple range test. This test is intended to give groupings of means that are not significantly different among themselves. The error rate is for each group, and has sometimes been called a familywise error rate. This is done in a manner similar to Bonferroni, except the calculation used to calculate the error rate is $[1-(1-\alpha)^{r-1}]$ instead of the sum of α. For comparing

two means that are r steps apart, where for adjacent means r=2. Two means separated by 3 other means would have r = 5, and the error rate would be $[1-(1-\alpha)^{r-1}] = [1-(1-0.05)^4] = 0.1855$. The value of a needed to keep an error rate of α is the reverse of this calculation, $[1-(1-0.05)^{1/4}] = 0.0127$.

Tukey's adjustment The Tukey adjustment allows for **all possible pairwise tests**, which is often what an investigator wants to do. Tukey developed his own tables (see Appendix table A.7 in your book, "percentage points of the studentized range). For "t" treatments and a given error degrees of freedom the table will provide 5% and 1% error rates that give an experimentwise rate of Type I error.

Scheffé's adjustment This test is the most conservative. It allows the investigator to do not only all pairwise tests, but **all possible tests**, and still maintain an experimentwise error rate of α. "All possible" tests includes not only all pairwise tests, but comparisons of all possible combinations of treatments with other combinations of treatments (see CONTRASTS below). The calculation is based on a square root of the F distribution, and can be used for range type tests or confidence intervals. The test is more general than the others mentioned, for the special case of pairwise comparisons, the statistic is $\sqrt{(t-1)} * F_{t-1, n(t-1)}$ for a balanced design with t treatments and n observations per treatment.

Place the post-hoc tests above in order from the one most likely to detect a difference (and the one most likely to be wrong) to the one least likely to detect a difference (and the one least likely to be wrong). **LSD is first, followed by Duncan's test, Tukey's and finally Scheffé's**. Dunnett's is a special test that is similar to Tukey's, but for a specific purpose, so it does not fit well in the ranking. The Bonferroni approach produces an upper bound on the error rate, so it is conservative for a given number of tests. It is a useful approach if you want to do a few tests, fewer than allowed by one of the others (e.g. you may want to do just a few and not all possible pairwise). In this case, the Bonferroni may be better.

Dunnett's adjustment - Dunnetts adjustment is a special adjustment intended only for testing **all treatments against one other treatment**, usually a control. For "t-1" treatments to be tested against the control the test gives an experimentwise rate of Type I error.

```
50              means treatmnt / dunnett('3DOk1');
```
         See SAS OUTPUT

All three options SAS output options shown for the LSD above are available for most of the SAS post-hoc tests (no LINES option with Dunnett's). I put all of them in the computer output on the internet, but not in this handout. Sample output is given below for the Tukey and Scheffé options only.

```
51              lsmeans treatmnt / pdiff stderr;
52              lsmeans treatmnt / pdiff stderr adjust=tukey;
```
         See SAS OUTPUT

Contrasts

A calculation similar to the LSD, but extended to more than just 2 means, is called a contrast. Suppose we wish to test the mean of the first two means against the mean of the last 3 means.

1) H_0: $\dfrac{\mu_1 + \mu_2}{2} = \dfrac{\mu_3 + \mu_4 + \mu_5}{3}$ or $\dfrac{\mu_1 + \mu_2}{2} - \dfrac{\mu_3 + \mu_4 + \mu_5}{3} = 0$ or

$$\left(\tfrac{1}{2}\mu_1 + \tfrac{1}{2}\mu_2\right) - \left(\tfrac{1}{3}\mu_3 + \tfrac{1}{3}\mu_4 + \tfrac{1}{3}\mu_5\right) = 0 \text{ or } \tfrac{1}{2}\mu_1 + \tfrac{1}{2}\mu_2 + \left(-\tfrac{1}{3}\right)\mu_3 + \left(-\tfrac{1}{3}\right)\mu_4 + \left(-\tfrac{1}{3}\right)\mu_5 = 0 \text{ or }$$

$$3\mu_1 + 3\mu_2 + (-2)\mu_3 + (-2)\mu_4 + (-2)\mu_5 = 0$$

This expression is what is a "linear model", and the last expression of this linear model is the easiest form for us to work with. We can evaluate the linear model, and if we can find the variance we can test the linear model. Generically, the variance of a linear model is "the sum of the variances", however there are a few other details. As with the transformations discussed earlier in the semester, when we multiply a value by "a" the mean changes by "a", but the variance changes by "a^2". Also, if there are covariances between the observations these must also be included in the variance. For our purposes, since we have assumed independence, there are no covariances.

The linear expression to evaluate is then: $a_1T_1 + a_2T_2 + a_3T_3 + a_4T_4 + ... + a_kT_k$ where the "a" are the coefficients and the "T" are the treatment means (sums can also be used).

The variance is then: $a^2_1Var(T_1) + a^2_2Var(T_2) + a^2_3Var(T_3) + a^2_4Var(T_4) + ... + a^2_kVar(T_k)$

In an ANOVA, the best estimate of the variance is the MSE, and the variance of a treatment mean is MSE/n, where n is the number of observations in that treatment. We can therefore factor out MSE, and in the balanced case (1/n) can also be factored out. The result is $(MSE/n)(a^2_1 + a^2_2 + a^2_3 + a^2_4 + ... + a^2_k)$

If we were to use a t-test to test the linear combination against zero, the t-test would be:

$$\frac{a_1T_1 + a_2T_2 + a_3T_3 + a_4T_4 + ... + a_kT_k}{\sqrt{\dfrac{MSE}{n}\left(a_1^2 + a_2^2 + a_3^2 + a_4^2 + ... + a_k^2\right)}} == \frac{\sum\limits_{i=1}^{k} a_i T_i}{\sqrt{\dfrac{MSE}{n}\sum\limits_{i=1}^{k} a_i^2}}$$

This is the test done with treatment means. If treatment totals are used the equation is modified silghtly to $\sum\limits_{i=1}^{k} a_i T_i \Big/ \sqrt{nMSE\sum\limits_{i=1}^{k} a_i^2}$ and will give the same result.

One final modification. If we calculate our "contrasts" as above without the "MSE" in the denominator, then we calculate $Q = \sum\limits_{i=1}^{k} a_i T_i \Big/ \sqrt{n\sum\limits_{i=1}^{k} a_i^2}$, without the MSE, then all that would remain to complete the t-test is to divide by $\sqrt{MSE}$.

The value called "Q", when divided by $\sqrt{MSE}$ gives a t statistic. If we calculate Q^2 and divide by MSE we get an F statistic. SAS uses F tests. All we need provide SAS is the values of "a", the coefficients, in the correct order, and it will calculate and test the "Contrast" with an F statistic.

Three arbitrary contrasts were done on the sample dataset. The statements used were

```
 ** treatments in order=data ===========> 3DOk1 3DOk4 3DOk5 3DOk7 3DOk13;
    contrast '3 low vrs 2 high' treatmnt   -2    -2    -2     3     3;
    contrast 'odd vrs even'     treatmnt   -1     4    -1    -1    -1;
    contrast '1st vrs 2nd'      treatmnt   -1     1     0     0     0;
```

Note that it is important to know the order of the means. In this case the means are in the order entered in the dataset because an "ORDER=DATA" option was added to the PROC GLM. The default order is alphabetical. The first of the 4 program lines above is just a comment I used to keep the order of the treatments straight. The three contrasts follow the comment. Note that the third contrast is just a pairwise test of two means, so all our pairwise tests could also be done as contrasts. Output from the contrasts statements are given below.

```
Percent fine gravel in surface soils
Simple ANOVA with non orthogonal contrasts

General Linear Models Procedure
```

Dependent Variable: PERCENT   Percent N content

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
3 low vrs 2 high	1	55.5712667	55.5712667	4.08	0.0571
odd vrs even	1	294.4656000	294.4656000	21.60	0.0002
1st vrs 2nd	1	605.2840000	605.2840000	44.40	0.0001

The error rate for each contrast is also α. Therefore, doing multiple contrasts may also inflate the error rate. Generally, there is little concern for this as long as the number of contrasts done does not exceed the degrees of freedom. If many contrasts are done Scheffé's adjustment should be considered.

Evaluating the assumptions for ANOVA.

We have already discussed some techniques for the evaluation of data for homogeneous variance. The assumption of independence is somewhat more difficult to evaluate. Random sampling is the best guarantee of independence and should be used as much as possible.

The third assumption is normality. The observarions are assumed to be normally distributed within each treatment, but how the treatments come together to form the dependent variable Y_{ij} may cause them to look non-normal. The best way to test for normality is to examine the residuals, pooling the normal distribution across the treatments to a common mean of zero. SAS will output the residuals with an output statement, and PROC UNIVARIATE has a number of tools to evaluate normality.

To output residuals from PROC GLM and other procedures use the output statement below as part of the PROC GLM section and then the PROC UNIVARIATE.

```
57      output out=resids residual=e;
60   proc univariate data=clover normal; by treatmnt; var percent;
61      TITLE2 'Univariate analysis of treatment groups';
62   run;
```

          See SAS OUTPUT

The variance / mean plot.

One final diagnostic technique for examining the assumption of homogeneous variance is a plot of the VARIANCE on the MEAN. In biological data it is often the case that when variance is not homogeneous it increases as the variance increases. This pattern will appear as a increasing (or decreasing) trend in the variances when plotted on the mean. The pattern may occur when the data follow a distribution other than the normal, because for many distributions the variance and the mean are NOT independent. Logarithmic or square root transformations can be used to try to address this

problem. Also, some newer statistical techniques (PROC MIXED) allow the fitting of different variances for different treatments.

    See SAS OUTPUT

Other techniques for examining homogeniety of variance

The previous analysis of variance (*Rhizobium strain effect on Red clover nitrogen content*) is redone below with PROC MIXED follows. The PROC MIXED is a potentially very different way of fitting ANOVA models, applying an interative solution to estimate the random effects instead of the least squares estimates from PROC GLM. For this simple case, with only one random error, the results are the same, but will often differ for more complex designs.

```
71    proc mixed data=clover order=data; class treatmnt;
72         TITLE2 'ANOVA with PROC MIXED';
73      model percent = treatmnt / htype=3 DDFM=Satterthwaite;
74    run;
```

```
Nitrogen content of red clover (S&T 1980)
ANOVA with PROC MIXED


The Mixed Procedure
                   Model Information
Data Set                       WORK.CLOVER
Dependent Variable             percent
Covariance Structure           Diagonal
Estimation Method              REML
Residual Variance Method       Profile
Fixed Effects SE Method        Model-Based
Degrees of Freedom Method      Residual


          Class Level Information
Class        Levels    Values
treatmnt        5      3DOk1 3DOk13 3DOk4 3DOk5 3DOk7


            Dimensions
Covariance Parameters              1
Columns in X                       6
Columns in Z                       0
Subjects                           1
Max Obs Per Subject               25
Observations Used                 25
Observations Not Used              0
Total Observations                25


Covariance Parameter Estimates
Cov Parm      Estimate
Residual       13.6334


        Fit Statistics
-2 Res Log Likelihood           117.1
AIC (smaller is better)         119.1
AICC (smaller is better)        119.3
BIC (smaller is better)         120.1


        Type 3 Tests of Fixed Effects
               Num        Den
```

Effect	DF	DF	F Value	Pr > F
treatmnt	**4**	**20**	**15.38**	**<.0001**

One of the options in PROC MIXED allows the fitting of different variances to each treatment. Four our simple case this is done with the GROUP specification on the REPEATED statement.

```
76    proc mixed data=clover order=data; class treatmnt;
77        TITLE2 'ANOVA with PROC MIXED - separate variances';
78      model percent = treatmnt / htype=3 DDFM=Satterthwaite;
79      repeated / group = treatmnt;
80      lsmeans treatmnt / adjust=tukey pdiff;
81    run;
```

Nitrogen content of red clover (S&T 1980)
ANOVA with PROC MIXED - separate variances

The Mixed Procedure
                 Model Information
Data Set   WORK.CLOVER
Dependent Variable   percent
Covariance Structure   Variance Components
Group Effect   treatmnt
Estimation Method   REML
Residual Variance Method   None
Fixed Effects SE Method   Model-Based
Degrees of Freedom Method   Between-Within

            Class Level Information
Class   Levels   Values
treatmnt   5   3DOk1 3DOk13 3DOk4 3DOk5 3DOk7

             Dimensions
Covariance Parameters   **5**
Columns in X   6
Columns in Z   0
Subjects   25
Max Obs Per Subject   1
Observations Used   25
Observations Not Used   0
Total Observations   25

                 Iteration History
Iteration   Evaluations   -2 Res Log Like   Criterion
        0   1   117.05518418
        1   1   104.64497787   0.00000000
                 Convergence criteria met.

      Covariance Parameter Estimates
Cov Parm   Group   Estimate
Residual   treatmnt 3DOk1   33.6420
Residual   treatmnt 3DOk13   2.0380
Residual   treatmnt 3DOk4   16.9430
Residual   treatmnt 3DOk5   14.2670
Residual   treatmnt 3DOk7   1.2770

         Fit Statistics
-2 Res Log Likelihood   104.6

```
AIC (smaller is better)              114.6
AICC (smaller is better)             118.9
BIC (smaller is better)              120.7

  Null Model Likelihood Ratio Test
    DF      Chi-Square        Pr > ChiSq
     4          12.41             0.0145

        Type 3 Tests of Fixed Effects
             Num      Den
Effect        DF       DF     F Value     Pr > F
treatmnt       4       20      25.64      <.0001
```

Once we have fitted the model with separate variances we can compare the version fitted with a pooled variance to the model with separate variances and test for a difference using the "-2 Res Log Likelihood" values. The differences between these values for a "full" and "reduced" model follows a Chi square distribution. The degrees of freedom for this Chi square will be the difference in the number of variance parameters estimated (one for the first model and 5 for the second = 4).

Likelihood ratio test between ANOVA models with common and separate variances.

ANOVA with PROC MIXED	Covariance Parameters	-2 Res Log Likelihood
Common variances	1	117.1
Separate variances	5	104.6
Difference	4	12.5
	Actual P value =	**0.01400**
	Tabular χ^2 d.f.=4, α=0.05	**9.49**
	Tabular χ^2 d.f.=4, α=0.01	**13.28**

Note that for this particular, simple case SAS provides a section titled "Null Model Likelihood Ratio Test". This does exactly the same test we just did, but for many other more complicated models this test will not be provided by SAS.

```
        Null Model Likelihood Ratio Test
            DF      Chi-Square        Pr > ChiSq
             4          12.41             0.0145
```

```
43        proc glm data=clover order=data; class treatmnt;
44            TITLE2 'Simple ANOVA with non orthogonal contrasts';
45          model percent = treatmnt / SS3;
46          MEANS treatmnt / HOVTEST=BARTLETT HOVTEST=BF HOVTEST=LEVENE(TYPE=ABS)
47                       HOVTEST=LEVENE(TYPE=SQUARE) HOVTEST=OBRIEN WELCH;
48          means treatmnt / lsd tukey bon scheffe duncan lines;
49          means treatmnt / lsd tukey bon scheffe duncan cldiff;
50          means treatmnt / dunnett('3DOk1');
51          lsmeans treatmnt / pdiff stderr;
52          lsmeans treatmnt / pdiff stderr adjust=tukey;
53      ** treatments in order=data ==========> 3DOk1 3DOk4 3DOk5 3DOk7 3DOk13;
54          contrast '3 low vrs 2 high' treatmnt  -2    -2    -2     3     3;
55          contrast 'odd vrs even'     treatmnt  -1     4    -1    -1    -1;
56          contrast '1st vrs 2nd'      treatmnt  -1     1     0     0     0;
57          output out=resids residual=e;
58      run;
58   !      quit;
```