Topics 6 & 7

· Testing all assumptions of ANOVA
· Randomized Complete Block Design (RCBD)
  Tukey 1-df Test for Nonadditivity
· Latin Squares
  Unreplicated
  Replicated, with shared columns and rows
  Replicated, independent
· Power Analysis of RCBD

· APPENDIX: Thinking about the Tukey 1-df Test for Nonadditivity

Testing all assumptions of ANOVA

The results of an ANOVA are valid only if the data satisfy the assumptions (i.e. criteria) of the test. The first thing you must always do, therefore, is make sure your data meet the assumptions. There are four:

1. **Errors are independent** · Satisfied through proper randomization

2. **Errors (a.k.a. residuals) are normally distributed** · Verified by testing for normality of residuals

   Until now, we've used the Shapiro-Wilk test to determine whether or not the observations themselves are normally distributed. While sample normality is a direct result of the normal distribution of errors, this approach is untenable for small sample sizes (the common case of treatments with only 3 or 4 replications, for example). Verification of the normal distribution of the residuals for the experiment as a whole (not treatment-by-treatment) is the accepted proxy indicator of the normality of the samples under consideration.

3. **Variances are homogeneous** · Verified using Levene's Test for Homogeneity of Variances

   Treatment was the only effect in a CRD, so Levene's Test was straightforward. For RCBD's, if you intend to make mean comparisons among treatments only (i.e. not among blocks), you must do Levene's Test for treatment, using a one-way ANOVA. *Levene's Test is valid only for one-way ANOVA's.*

4. **Model effects are additive** · Verified using the Tukey 1-df Test for Nonadditivity

   It is necessary to test this assumption only when there is just a single replication per block-treatment combination, thereby leaving you no way to measure directly the error or noise in your experiment. The Tukey Test checks to make sure the block-treatment interaction is not significant and therefore can be used as a good estimator of MSE for use in the ANOVA.
Randomized Complete Block Design (RCBD)

Example 4.1

In this experiment, Figueruelo et al. (1993) took seven soil samples from the Canary Islands and divided each sample into five equal parts. Five amounts of phosphate (0, 50, 100, 150, and 200 ppm, in the form of Na$_2$PO$_4$H) were randomly assigned among each of these groups of five soil samples and mixed in thoroughly. Afterward, the amount of exchangeable calcium (in meq/100 g) was measured for each sample. [DESIGN: RCBD with 1 replication per block*treatment combination.]

```
Data Calcium;
    Input Block Phosphate Calcium; * Every observation now belongs
to an associated block;
Cards;
1   0 3.51 2  0 5.07 3 0 2.97 4 0 2.68 5 0 6.54 6 0 2.38 7 0 2.06
1  50 3.68 2  50 3.94 3 50 2.86 4 50 2.50 5 50 7.25 6 50 3.51 7 50 2.06
1 100 3.62 2 100 3.97 3 100 2.92 4 100 2.47 5 100 7.27 6 100 3.38 7 100 2.61
1 150 3.75 2 150 4.03 3 150 2.93 4 150 2.45 5 150 7.13 6 150 3.40 7 150 2.06
1 200 3.71 2 200 5.14 3 200 3.60 4 200 3.12 5 200 7.96 6 200 3.40 7 200 2.22
;
Proc GLM Data = Calcium;
    Title 'The ANOVA';
    Class Block Phosphate; * We now have TWO class variables;
    Model Calcium = Block Phosphate;
    Output Out = CalciumPR r = Res p = Pred;
        * Generates the residuals (R) and predicted (P) values needed for
          testing normality of residuals as well as for carrying out the
          Tukey 1-df Test for Nonadditivity;

Proc Univariate Normal Data = CalciumPR; * Testing for normality of residuals;
    Var Res;

Proc GLM Data = Calcium; * Levene's Test for Treatment;
    Title 'Levene for Treatment';
    Class Phosphate;
    Model Calcium = Phosphate;
    Means Phosphate / hovtest = Levene;

Proc GLM Data = Calcium; * Levene's Test for Block (not necessary);
    Title 'Levene for Block';
    Class Block;
    Model Calcium = Block;
    Means Block / hovtest = Levene;

Proc GLM Data = CalciumPR;
    Title 'Tukey 1-df Test';
    Class Block Phosphate;
    Model Calcium = Block Phosphate Pred*Pred;
        * Note that 'Pred*Pred,' as a regression variable, is not
          listed as a class variable and appears only in the model;

Proc Plot Data = CalciumPR;
    Plot Res * Pred = Phosphate /vpos = 25;
Proc GPlot Data = CalciumPR;
    Plot Res * Pred = Phosphate;
Run;
Quit;
```
Discussion of the Code

Because the observations are classified according to both Block and Treatment, both variables appear in the **Class** statement. By appearing on the left side of the equal sign in the **Model** statement, “Calcium” is specified as the response variable. The two known, non-error sources of variation in the experiment are Block (large soil sample) and Treatment (added phosphate, in five different amounts), so these variables appear on the right side of the equal sign in the **Model** statement.

The first **Proc GLM** codes for a standard RCBD analysis. While the **Output** statement is not essential to the ANOVA itself, it is very useful for the subsequent testing of assumptions. As you've seen before, the **Output** statement produces a new data set. The name of this set (CalciumPR) is specified after "Out =", and it contains the old variables plus two additional ones: the first one (specified after "p =") includes the predicted values and the second (specified after "r =") includes the residuals. SAS calculates the predicted and residual values based on the linear model you supplied in the **Model** statement above.

In a CRD, predicted values are determined by adding individual treatment effects to the overall mean, such that the predicted value of an observation within Treatment, is simply the mean observation within that treatment ($y_i = \mu + \tau_i$). We've seen this before.

Now, in an RCBD, the predicted value is the sum of the overall mean, the treatment effect, and the block effect (remember, it's an additive model):

$$y_i = \mu + \tau_i + \beta_i$$

The residuals, then, are simply the deviations of the observed values from these predicted values. The normal distribution of these residuals is tested using **Proc Univariate**. The second **Proc GLM** uses Levene's test to evaluate the homogeneity of variances among treatments; the third **Proc GLM** does the same for blocks. Since we have no intention to compare Block means, this latter Levene's test is unnecessary; it is included here just to show you how one would code for it and to emphasize the fact that **Levene's Test is valid only for 1-way ANOVAs**.

The role of the fourth **Proc GLM** is to carry out Tukey's 1-df Test for Nonadditivity. For this test, “Pred*Pred” is a regression variable and, as such, is not included in the Class statement. As with the soybean example last week, this “Pred*Pred” regression variable is testing for a significant quadratic component to the relationship between the observed and predicted values [see Appendix].

A couple versions of **Proc Plot** have also been included so that we may visually check the residuals (reading the tea leaves); each produces a plot of residual versus predicted values.

**Output**

**Normality of Residuals**

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>0.982268</td>
<td>0.8302</td>
</tr>
</tbody>
</table>

We fail to reject $H_0$ (errors are normally distributed).
**Levene's Test for Treatment**

Levene's Test for Homogeneity of Calcium Variance
ANOVA of Squared Deviations from Group Means

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phosphate</td>
<td>4</td>
<td>2.6741</td>
<td>0.6685</td>
<td>0.04</td>
<td>0.9975 NS</td>
</tr>
</tbody>
</table>

We fail to reject H₀ (variances among treatments are homogeneous).

**Levene's Test for Block (again, unnecessary unless you're comparing block means)**

Levene's Test for Homogeneity of Calcium Variance
ANOVA of Squared Deviations from Group Means

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>6</td>
<td>0.3566</td>
<td>0.0594</td>
<td>1.74</td>
<td>0.1492 NS</td>
</tr>
<tr>
<td>Error</td>
<td>28</td>
<td>0.9579</td>
<td>0.0342</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tukey 1-df Test for Nonadditivity**

Dependent Variable: Calcium

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>86.63008776</td>
<td>7.87546252</td>
<td>64.46</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>23</td>
<td>2.80990081</td>
<td>0.12216960</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>34</td>
<td>89.43998857</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.968583      Coeff Var: 9.262867      Root MSE: 0.349528        Calcium Mean: 3.773429

Source: Block                        DF     Type I SS     Mean Square     F Value     Pr > F
| Block | 6   | 84.92026857     | 14.15337810  | 115.85   | <.0001 **|
| Phosphate | 4  | 1.42607429     | 0.35651857   | 2.92     | 0.0434 NS |
| Pred*Pred | 1  | 0.28374490      | 0.28374490   | 2.32     | 0.1411 NS |

We fail to reject H₀ (model effects are additive); therefore, we are justified in using the Block*Treatment interaction as a reliable estimate of the true experimental error (MSE).

**Finally, the ANOVA**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>10</td>
<td>86.34634286</td>
<td>8.63463429</td>
<td>66.99</td>
<td>&lt;.0001 **</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>3.09364571</td>
<td>0.12890190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>34</td>
<td>89.43998857</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.965411      Coeff Var: 9.514666      Root MSE: 0.359029        Calcium Mean: 3.773429

Source: Block                        DF     Type III SS     Mean Square     F Value     Pr > F
| Block | 6   | 84.92026857     | 14.15337810  | 109.80   | <.0001 *** |
| Phosphate | 4  | 1.42607429     | 0.35651857   | 2.77     | 0.0506 NS |

R-Square: 0.965411      Coeff Var: 9.514666      Root MSE: 0.359029        Calcium Mean: 3.773429

Source: Block                        DF     Type III SS     Mean Square     F Value     Pr > F
| Block | 6   | 84.92026857     | 14.15337810  | 109.80   | <.0001 *** |
| Phosphate | 4  | 1.42607429     | 0.35651857   | 2.77     | 0.0506 NS |
Some further comments

- Be careful: SAS does not make a big deal differentiating between the second Tukey Test output and the first real ANOVA table.
- Disregard the results for Block and Treatment in the table for the Tukey Test.

Notice the DF of the model in the actual ANOVA and the Tukey Test. Why does this difference exist?

Plot of residuals (tea leaves), from Proc Plot

Plot of residuals (tea leaves), from Proc GPlot

There is no evidence in these plots of residuals that there are any serious problems with nonadditivity or variance homogeneity. This is verified by the NS Levene and Tukey Tests above.
**Latin Squares**

This first example features an unreplicated Latin Square with four treatments. Since there is only one replication per column-row-treatment combination, the ANOVA uses the interactions of these effects as the error term (MSE). In this case, however, there are now *three* possible two-way interactions:

- Column x Row
- Column x Treatment
- Row x Treatment

It is necessary to test for the significance of each of these, requiring *three separate Proc GLM’s* with three separately-generated Pred-Res datasets.

---

**Example 4.2 Un replicated Latin Square**  
*ST&D pg. 230 [Lab4ex2.sas]*

To assist you in using the following code as a template, the classification and response variables have been labeled generically as possible (e.g. Row, Col, Trtmt, and Response).

```sas
Data LSWheat;
Input Row Col Trtmt $ Response;
Cards;
1 1 C 10.5  2 1 B 11.1  3 1 D 5.8  4 1 A 11.6
1 2 D 7.7  2 2 A 12.0  3 2 C 12.2  4 2 B 12.3
1 3 B 12.0  2 3 C 10.3  3 3 A 11.2  4 3 D 5.9
1 4 A 13.2  2 4 D  7.5  3 4 B 13.7  4 4 C 10.2
;
Proc GLM Data = LSWheat Order = Data; * The ANOVA;
   Class Row Col Trtmt;
   Model Response = Row Col Trtmt;
      * You can put in contrasts, separations, whatever - no new GLM needed;
   Means Trtmt / Tukey;
   Contrast 'D vs. AB&C' Trtmt  1 -3 1 1;
   Output Out = LSWheatPR P = Pred R = Res;
Proc Univariate Normal Data = LSWheatPR; * Testing normality of residuals;
   Var Res;
Proc GLM Data = LSWheat; * Testing treatment homogeneity of variances;
   Class Trtmt;
   Model Response = Trtmt;
   Means Trtmt / hovtest = Levene;
Proc GLM Data = LSWheat; * Testing ColxRow interaction;
   Title 'Interaction ColxRow';
   Class Col Row;
   Model Response = Col Row;
   Output Out = ColRow p = PredCR;
Proc GLM Data = ColRow;
   Title 'Output Tukey ColxRow';
   Class Col Row;
   Model Response = Col Row PredCR*PredCR;
Proc GLM Data = LSWheat; * Testing ColxTrt interaction;
   Title 'Interaction ColxTrtmt';
   Class Col Trtmt;
   Model Response = Col Trtmt;
   Output Out = ColTrt p = PredCT;
Proc GLM Data = ColTrt;
```

---

PLS205 4.6 Lab 4 (Topics 6-7) 6
Title 'Output Tukey ColxTrt';
Class Col Trtmt;
Model Response = Col Trtmt PredCT*PredCT;
Proc GLM Data = LSWheat; * Testing RowxTrt interaction;
Title 'Interaction RowxTrt';
Class Row Trtmt;
Model Response = Row Trtmt;
Output Out = RowTrt p = PredRT;
Proc GLM Data = RowTrt;
Title 'Output Tukey ColxRow';
Class Row Trtmt;
Model Response = Row Trtmt PredRT*PredRT;
Run;
Quit;

Output

Normality of Residuals

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>--Statistic---</th>
<th>-----p Value------</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W 0.989685</td>
<td>Pr &lt; W 0.9991 NS</td>
</tr>
</tbody>
</table>

We fail to reject $H_0$. There is no evidence that the residuals (i.e. errors) are not normally distributed.

Treatment Homogeneity of Variances

Levene's Test for Homogeneity of Yield Variance

ANOVA of Squared Deviations from Group Means

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trtmt</td>
<td>3</td>
<td>0.2174</td>
<td>0.0725</td>
<td>0.13</td>
<td>0.9393 NS</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>6.5985</td>
<td>0.5499</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We fail to reject $H_0$. There is no evidence that the treatment variances are not homogeneous.

Tukey 1-df Tests for Nonadditivity

ColxRow Interaction

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col</td>
<td>3</td>
<td>6.80000000</td>
<td>2.26666667</td>
<td>0.22</td>
<td>0.8779</td>
</tr>
<tr>
<td>Row</td>
<td>3</td>
<td>1.95500000</td>
<td>0.65166667</td>
<td>0.06</td>
<td>0.9774</td>
</tr>
<tr>
<td>PredCR*PredCR</td>
<td>1</td>
<td>0.25494817</td>
<td>0.25494817</td>
<td>0.03</td>
<td>0.8781 NS</td>
</tr>
</tbody>
</table>

ColxTrt Interaction

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Col</td>
<td>3</td>
<td>6.80000000</td>
<td>2.26666667</td>
<td>3.98</td>
<td>0.0525</td>
</tr>
<tr>
<td>Trtmt</td>
<td>3</td>
<td>78.92500000</td>
<td>26.30833333</td>
<td>46.20</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PredCT*PredCT</td>
<td>1</td>
<td>0.11927929</td>
<td>0.11927929</td>
<td>0.21</td>
<td>0.6594 NS</td>
</tr>
</tbody>
</table>
RowxTrt Interaction

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
<td>3</td>
<td>1.95500000</td>
<td>0.65166667</td>
<td>0.55</td>
<td>0.6621</td>
</tr>
<tr>
<td>Trtmt</td>
<td>3</td>
<td>78.92500000</td>
<td>26.30833333</td>
<td>22.21</td>
<td>0.0003</td>
</tr>
<tr>
<td>PredRT*PredRT</td>
<td>1</td>
<td>0.04333230</td>
<td>0.04333230</td>
<td>0.04</td>
<td>0.8531</td>
</tr>
</tbody>
</table>

We fail to reject H0 in all three cases. There is no evidence of nonadditive two-way interactions.

**Finally, the ANOVA**

Dependent Variable: Yield

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
<td>3</td>
<td>1.95500000</td>
<td>0.65166667</td>
<td>1.44</td>
<td>0.3219</td>
</tr>
<tr>
<td>Col</td>
<td>3</td>
<td>6.80000000</td>
<td>2.26666667</td>
<td>5.00</td>
<td>0.0452</td>
</tr>
<tr>
<td>Trtmt</td>
<td>3</td>
<td>78.92500000</td>
<td>26.30833333</td>
<td>58.03</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Notice that Row is NS. Would you include “Row” as a blocking variable the next time you do the experiment? It’s difficult to say because the Row SS is certainly reducing the Error SS (which increases our sensitivity), but it does so at a price, namely a reduction in the error df. To really answer this question, it is necessary to consider the relative efficiencies of the two experiments (LS vs. RCBD) [see Topic 7.7 in your lecture notes].

Note: Since Levene's Test is only valid for 1-way ANOVA's, you would have to perform three separate Levene's Tests here if you were interested in carrying out mean separations for column, row, and treatment. However, since we are comparing treatment means only, we need simply to conduct the Levene's Test for treatment, shown above.

**Example 4.3 Replicated Latin Squares** [Lab4ex3.sas]

In this example, three gasoline additives (TREATMENTS A, B, & C) were tested for combustion efficiency by three drivers (ROWS 1, 2, & 3) using three different tractors (COLUMNS 1, 2, & 3). The variable measured was the amount of CO trapped from the exhaust. The experiment was carried out twice, on two different days.

Data LSGasAdd;
   Input Day Tractor Driver Trtmt $ CO;
Cards;
1 1 1 B 26.0
1 1 2 C 28.7
1 1 3 A 25.3
1 2 1 C 25.0
1 2 2 A 23.6
1 2 3 B 28.4
1 3 1 A 21.3
1 3 2 B 28.5
1 3 3 C 30.1

2 1 1 C 32.4
2 1 2 B 31.7
2 1 3 A 24.9
2 2 1 B 28.7
2 2 2 A 24.3
2 2 3 C 29.3
2 3 1 A 25.8
2 3 2 C 30.5
2 3 3 B 29.2
Important: The above program, as a template, illustrates four different ways to analyze a replicated Latin Square. You would only run ONE of these analyses, depending on which blocking variables are shared between the two replications of the experiment.

Output

Replicated LS sharing both rows and columns

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>7</td>
<td>132.0038889</td>
<td>18.8576984</td>
<td>8.19</td>
<td>0.0018</td>
</tr>
<tr>
<td>Error</td>
<td>10</td>
<td>23.0122222</td>
<td>2.3012222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>17</td>
<td>155.0161111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R-Square</th>
<th>Coeff Var</th>
<th>Root MSE</th>
<th>CO Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.851549</td>
<td>5.530809</td>
<td>1.516978</td>
<td>27.42778</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>22.00055556</td>
<td>22.00055556</td>
<td>9.56</td>
<td>0.0114 *</td>
</tr>
<tr>
<td>Tractor</td>
<td>2</td>
<td>8.01444444</td>
<td>4.00722222</td>
<td>1.74</td>
<td>0.2244</td>
</tr>
<tr>
<td>Driver</td>
<td>2</td>
<td>7.29711111</td>
<td>3.60055556</td>
<td>1.56</td>
<td>0.2563</td>
</tr>
<tr>
<td>Trtmt</td>
<td>2</td>
<td>94.78777778</td>
<td>47.39388889</td>
<td>20.60</td>
<td>0.0003 ***</td>
</tr>
</tbody>
</table>

In this case, our model has 7 df and our error has 10 df; the effect of driver and tractor are each found to be NS.
Replicated independent LS

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>11</td>
<td>152.37944444</td>
<td>13.8526768</td>
<td>31.52</td>
<td>0.0002</td>
</tr>
<tr>
<td>Error</td>
<td>6</td>
<td>2.63666667</td>
<td>0.43944444</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>17</td>
<td>155.01611111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE CO Mean
0.982991 2.416915 0.662906 27.42778

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>22.00055556</td>
<td>22.00055556</td>
<td>50.06</td>
<td>0.0004 ***</td>
</tr>
<tr>
<td>Tractor(Day)</td>
<td>4</td>
<td>9.42222222</td>
<td>2.35555556</td>
<td>5.36</td>
<td>0.0350 *</td>
</tr>
<tr>
<td>Driver(Day)</td>
<td>4</td>
<td>26.16888889</td>
<td>6.54222222</td>
<td>14.89</td>
<td>0.0029 **</td>
</tr>
<tr>
<td>Trtmt</td>
<td>2</td>
<td>94.787777778</td>
<td>47.39388889</td>
<td>107.85</td>
<td>&lt;.0001 ***</td>
</tr>
</tbody>
</table>

In this case, the model df has increased to 11 and the error df has decreased to 6. Now all model effects are found to be significant.

Replicated LS sharing rows but with independent columns

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>9</td>
<td>133.4116667</td>
<td>14.8235185</td>
<td>5.49</td>
<td>0.0126</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>21.60444444</td>
<td>2.70055556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>17</td>
<td>155.01611111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE CO Mean
0.860631 5.991505 1.643337 27.42778

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>22.00055556</td>
<td>22.00055556</td>
<td>8.15</td>
<td>0.0213 *</td>
</tr>
<tr>
<td>Tractor(Day)</td>
<td>4</td>
<td>9.42222222</td>
<td>2.35555556</td>
<td>0.87</td>
<td>0.5207</td>
</tr>
<tr>
<td>Driver</td>
<td>2</td>
<td>7.20111111</td>
<td>3.60055556</td>
<td>1.33</td>
<td>0.3164</td>
</tr>
<tr>
<td>Trtmt</td>
<td>2</td>
<td>94.787777778</td>
<td>47.39388889</td>
<td>17.55</td>
<td>0.0012 **</td>
</tr>
</tbody>
</table>

In this case, the model df has dropped to 9 and the error df has increased to 8. Significance levels have also changed. Similar changes can be seen in the next ANOVA:

Replicated LS sharing columns but with independent rows

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>9</td>
<td>150.9716667</td>
<td>16.7746296</td>
<td>33.18</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>8</td>
<td>4.04444444</td>
<td>0.50555556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>17</td>
<td>155.01611111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE CO Mean
0.973910 2.592351 0.711024 27.42778
Source                      DF     Type III SS     Mean Square    F Value    Pr > F
Day                          1     22.00055556     22.00055556      43.52    0.0002 ***
Tractor                      2      8.01444444      4.00722222       7.93    0.0127 *
Driver(Day)                  4     26.16888889      6.54222222      12.94    0.0014 **
Trtmt                        2     94.78777778     47.39388889      93.75    <.0001 ***

So, very different results depending on how the rows and columns relate to each other across the two replications of the experiment.

**Proc GLM Power (Power Analysis of RCBD)**

**Example 4.4 Proc GLM Power**

```
Data flax;
Do trtmt = 1 to 6;
   Do block = 1 to 4;
      Input oil @@;
      Output;
   End;
End;
Cards;
34.4 35.9 36.0 34.1
33.3 31.9 34.9 37.1
34.4 34.0 34.5 33.1
36.8 36.6 37.0 36.4
36.6 34.9 35.9 37.2
36.4 37.3 37.7 36.7;
Proc Print;
   var Trtmt Block oil;
Proc GLM Data = flax;
   Class Trtmt Block;
   Model Oil = Trtmt Block;
Proc GLMPower Data = flax;
   Class Trtmt Block;
   Model Oil = Trtmt Block;
   Power
      Stddev = 1.161596
      Alpha = .05
      ntotal = 24
      Power = .;
Run;
Quit;
```
Dependent Variable: oil

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>8</td>
<td>35.14000000</td>
<td>4.39250000</td>
<td>3.26</td>
<td>0.0234</td>
</tr>
<tr>
<td>Error</td>
<td>15</td>
<td>20.23958333</td>
<td>1.34930556</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>23</td>
<td>55.37958333</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square       | Coeff Var | Root MSE   | oil Mean |
---------------|-----------|------------|-----------|
0.634530       | 3.267883  | 1.161596   | 35.54583  |

The GLMPOWER Procedure

Fixed Scenario Elements

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>oil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Error Standard Deviation</td>
<td>1.161596</td>
</tr>
<tr>
<td>Total Sample Size</td>
<td>24</td>
</tr>
<tr>
<td>Error Degrees of Freedom</td>
<td>15</td>
</tr>
</tbody>
</table>

Computed Power

<table>
<thead>
<tr>
<th>Index</th>
<th>Source</th>
<th>DF</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>trtmt</td>
<td>5</td>
<td>0.902</td>
</tr>
<tr>
<td>2</td>
<td>block</td>
<td>3</td>
<td>0.173</td>
</tr>
</tbody>
</table>

In order to run the Proc GLMPOWER statement, it is necessary to have the Root MSE value (1.161596), from the ANOVA. So, run an Proc GLM on your original data first, then use the Root MSE from the output as the Stddev in the Proc GLMPOWER statement.

In this case, we see that the RCBD Power is .902 (90.2%), which is well over 80%.
APPENDIX: Thinking about the Tukey 1-df Test for Nonadditivity

How does a regression of the observed data against the squares of its predicted values tell you anything about the existence of nonadditive effects? One way to think about it:

To begin, recall that under our linear model, each observation is characterized as:

\[ y_{ij} = \mu + \beta_j + \tau_i + \varepsilon_{ij} \]

And its predicted value is determined by:

\[ \text{pred}_{ij} = \mu + \beta_j + \tau_j \]

In looking at these two equations, the first thing to notice is the fact that, if we had no error in our experiment (i.e. if \( \varepsilon_{ij} = 0 \)), the observed data would exactly match its predicted values and a correlation plot of the two would yield a perfect line with slope = 1:

Now let's introduce some error. If the errors in the experiment are in fact random and independent (criteria of the ANOVA and something achieved by proper randomization from the outset), then \( \varepsilon_{ij} \) will be a random variable that causes no systematic deviation from this linear relationship, as indicated in the next plot:

![Observed vs. Predicted Values (RCBD, no error)](image_url)
As this plot shows, while random error may decrease the overall strength of correlation, it will not systematically compromise its underlying linear nature.

So far so good. But what happens when you have an interaction (e.g. Block * Treatment) but lack the degrees of freedom necessary to include it in the linear model (e.g. when you have only 1 replication per block*treatment combination)? In this case, the df and the variation assigned to the interaction are relegated to the error term simply because we need a nonzero \( \text{df}_{\text{error}} \) to carry out our F tests. Under such circumstances, you can think of the error term as now containing two separate components:

\[
\varepsilon_{yi} = \varepsilon_{\text{RANDOMij}} + \text{B*T Interaction Effects}
\]

While the first component is random and will not affect the underlying linear correlation seen above, the second component is non-random and will cause systematic deviations from linearity. Indeed, if this interaction component is too large, the observed vs. predicated correlation will become detectably non-linear, thereby violating the ANOVA assumption of random and independent error, not to mention making your F tests much less sensitive.

The plot on the following page illustrates the deviation from linearity that results when significant multiplicative effects (one kind of nonadditive effect) cannot be accommodated by the model. The quadratic component of the trend is unmistakable.
If you buy all of that, then it's a logical jump to state:

If the observed and predicted values obey a linear relationship, then the non-random Interaction Effects buried in the error term are sufficiently small to uphold our assumption of random, independent error.

Seen in this light, our test for unaccounted-for nonadditivity [significant nonadditive (i.e. interaction) effects] becomes a simple test for linear regression. So why regress the observed values with the squares of the predicted values? Because, as we have discussed before when talking about contrasts, to establish the existence of a linear relationship (as opposed to a correlation of a higher power), one must test for (and successfully reject $H_0$ for) the quadratic component.

And that's the idea.