12. 9. Repeated measures analysis

Sometimes researchers take observations repeatedly on the **same experimental units**, for instance in measuring plant height **over time**; yields of perennial crops over seasons; animal growth; population dynamics over years etc. These observations are not replications and are called repeated measurements. Furthermore, because these observations are taken from the same experimental units, they are **not independent**. It is important to note in this context that the important assumption is not the independence of *measurements* but the independence of *errors*. Indeed, the split plot model implicitly assumes a correlation between measurements.

Time series or multivariate methods can be used to analyze these designs. However, time series methods are more appropriate to analyze long series of data, with more than 20 repeated measures. Time series experiments with large number of replications are more frequent in analyses of stock data or weather data than in agricultural experiments.

The split-plot principle can be applied to experiments where successive observations are made on the same experimental unit over a period of time. For example, a fertilizer trial or variety trial with a perennial crop like alfalfa might be harvested several times. Other examples can be repeated picking of fruit from the same trees or repeated sampling of soil plots over time for nutrient content. The plots to which the treatments are assigned be called main plots, and the several harvests can be called subplots. A subplot in this case, however, differs from the usual subplot in that it consists of data taken from the entire main plot rather than from a designated portion of the main plot, as is the case with the usual split-plot.

The dependency among the observations is used to adjust the degree of freedoms in the analysis of variance to give an approximate test for the repeated measurements. The approximate ANOVA for a CRD is shown below and is followed by an example of RCBD.

12. 9. 1. ANOVA of the repeated measures analysis

Approximate ANOVA of repeated measurement analysis

Source	df	SS	MS	**Conservative df**
Between Experimental Units				
Treatment (A)	a-1	SSA	SSA/(a-1)	
Replic.* Trt (Error A)	a (n-1)	SS(MPE)	SS(MPE)/a(n-1)	
Within Experimental Units				
Response in time (B)	b-1	SSB	SSB /(b-1)	**1**
Response by Trt. (A*B)	(b-1)(a-1)	SSAB	SSAB/(b-1)(a-1)	**a-1**
Error B	a(b-1)(n-1)	SS(SPE)	SS(SPE)/a (b-1) (n-1)	**a(n-1)**

The analysis looks like a split-plot analysis, except conservative degrees of freedoms are used in the *F*-tests for the repeated responses and the interactions. There are no unusual problems in analyzing data on a main plot (A) basis for a single replication or for the totals over several dates of observation. However, *F* values arising from testing the effects of successive observations (B) and the interaction of main plots treatments with successive observations (A*B) may not be distributed as *F*, and too many significant effects may result.

Note that what distinguishes repeated-measures data from any other multivariate data is not so much the existence of the repeated measurements but the desire to examine changes in the measurements taken on each subject.

The split-plot model specifies that pairs of observations within the same main unit are equally correlated. With repeated-measures data, pairs of observations on the same unit are not necessarily equally correlated. Measures close in time can be more highly correlated than measurements far apart in time. Since the unequal correlation between the repeated measurements is ignored in a split-plot analysis, the tests derived in this manner may not be valid.

A **conservative approach** recommended by many statisticians is to require larger F values for significance of B and A*B effects. It is suggested that degrees of freedom of response in time B be used to divide degrees of freedom for B, A*B, and Error B, and to select tabular *F* values on the basis of the resulting degrees of freedom (previous Table, right column).

The uncorrected degrees of freedom are appropriate for independent replications within main plots. The corrected ones (right column) are appropriate for totally dependent replications. Total dependence would be equivalent to have all responses represented by 1 response, and explains why the corrected df is one. This is the worst possible scenario, and the dependence in a real experiment will probably be somewhere between these two extremes.

12. 9. 2. Example of a repeated measurements experiment

The following experiment represents the yields (tons/acre) of 4 alfalfa cultivars tested with 5 replications in a CRD, with repeated measurements (4 cuts of yields, Sept. 10/74; June 25/75; Aug. 5/75; and Sept. 16/75). The data is analyzed as a split-plot design for repeated measurements, in a CRD.

SAS program

```
data rem_mes;
input rep A_var B_cut yield @@;
cards;
1 1 1  2.80191  1 1 2  3.73092  1 1 3  3.09856  1 1 4  2.50965
1 2 1  2.76212  1 2 2  5.40530  1 2 3  3.82431  1 2 4  2.72992
1 3 1  2.29151  1 3 2  3.81140  1 3 3  2.92575  1 3 4  2.39863
1 4 1  2.56631  1 4 2  4.96070  1 4 3  2.81734  1 4 4  2.05752
2 1 1  2.96602  2 1 2  4.43545  2 1 3  3.10607  2 1 4  2.57299
```

```
2 2 1  3.09636  2 2 2  3.90683  2 2 3  3.26229  2 2 4  2.58614
2 3 1  2.54027  2 3 2  3.82716  2 3 3  2.86727  2 3 4  2.16287
2 4 1  2.31630  2 4 2  3.96629  2 4 3  2.91461  2 4 4  2.15764
3 1 1  2.43232  3 1 2  4.32311  3 1 3  2.81030  3 1 4  2.07966
3 2 1  3.09917  3 2 2  4.08859  3 2 3  3.13148  3 2 4  2.60316
3 3 1  2.41199  3 3 2  4.08317  3 3 3  3.03906  3 3 4  2.07076
3 4 1  2.65834  3 4 2  3.71856  3 4 3  2.92922  3 4 4  2.15684
4 1 1  2.93509  4 1 2  3.99711  4 1 3  2.77971  4 1 4  2.44033
4 2 1  2.65256  4 2 2  5.42879  4 2 3  2.70891  4 2 4  2.30163
4 3 1  2.30420  4 3 2  3.27852  4 3 3  2.72711  4 3 4  2.04933
4 4 1  2.47877  4 4 2  3.92048  4 4 3  3.06191  4 4 4  2.35822
5 1 1  2.42277  5 1 2  3.85657  5 1 3  3.24914  5 1 4  2.34131
5 2 1  2.63666  5 2 2  3.77458  5 2 3  3.09734  5 2 4  2.30082
5 3 1  2.36941  5 3 2  3.44835  5 3 3  2.50562  5 3 4  2.08980
5 4 1  2.23595  5 4 2  4.02985  5 4 3  2.85279  5 4 4  1.85736
;
proc glm;
  class rep A_var B_cut;
  model yield= A_var rep*A_var B_cut A_var*B_cut;
  test h=A_var e=rep*A_var;
  means A_var / lsd e=rep*A_var;
  means B_cut/ lsd;
run; quit;
```

Split Plot for a CRD

Class Level Information

Class	Levels	Values
REP	5	1 2 3 4 5
A_VAR	4	1 2 3 4
B_CUT	4	1 2 3 4

Number of observations in data set = 80

Dependent Variable: **YIELD**

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	31	42.884847	1.383382	14.46	0.0001
Error	48	4.592605	**0.095679**		
Corrected Total	79	47.477452			

R-Square	C.V.	Root MSE	YIELD Mean
0.903268	10.33208	0.3093	2.9938

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A_VAR	3	2.840753	0.946918	~~9.90~~	~~0.0001~~
REP*A_VAR	16	2.048682	**0.128043**	1.34	0.2141
B_CUT	3	37.447690	12.482563	**130.46**	0.0001
A_VAR*B_CUT	9	0.547721	0.060858	**0.64**	0.7606

Tests of Hypotheses using the Type III MS for REP*A_VAR as an error term

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A_VAR	3	2.8407531	0.9469177	**7.40**	0.0025

The adjusted degrees of freedom are

B_CUT	1	37.447690	37.447690	130.46	$F_{1,16}$ **
A_VAR*B_CUT	3	0.547721	0.182574	0.64	$F_{3,16}$ NS
Error	16	4.592605	0.287038		

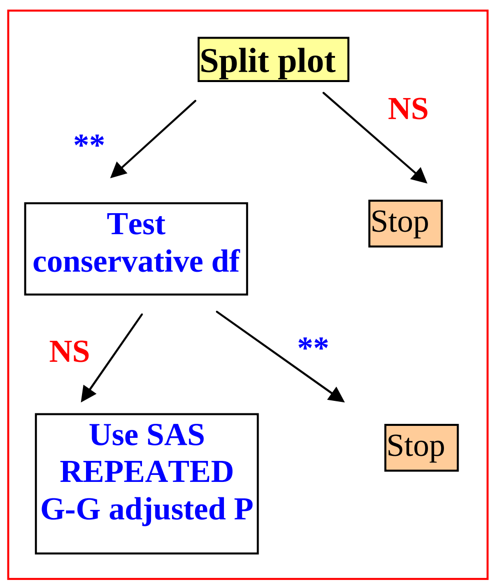
The *F* test for cuttings is significant using the most conservative df. We can conclude that there are significant differences among cuttings. Since no interactions were detected comparisons among the means of the main effects are appropriate.

```
T tests (LSD) for variable: YIELD
Alpha= 0.05  df= 16  MSE= 0.128043
Critical Value of T= 2.12, Least Significant Difference= 0.2399

Means with the same letter are not significantly different.
T Grouping              Mean     N  A_VAR
       A              3.2698     20  2
B      A              3.0444     20  1
B      C              2.9008     20  4
       C              2.7601     20  3

T tests (LSD) for variable: YIELD
Alpha= 0.05  df= 48  MSE= 0.095679
Critical Value of T= 2.01. Least Significant Difference= 0.1967

Means with the same letter are not significantly different.
T Grouping     Mean       N  B_CUT
A            4.09959     20  2
B            2.98544     20  3
C            2.59890     20  1
D            2.29123     20  4
```



Decision tree

If the original split plot in time is NS -> Stop. Conclusion NS

If it is significant, test not the conservative df

If the conservative df is significant -> Stop. Conclusion **

If the original split plot is ** and the conservative df is NS -> used the REPEATED statement: see next section.

12. 9. 3. Univariate Repeated-measures analysis using the REPEATED statement

In the previous example there was no conflict because the adjusted and unadjusted df produce the same conclusions. However, in some cases adjusted and unadjusted df can produce different conclusions. A more precise calculation can be obtained using SAS REPEATED statement that is set up specifically for this situation

Repeated-measures is a special case of multivariate analysis of variance (MANOVA). In the present example the four cuttings are treated as the four response variables and the blocks and varieties are the classification variables. The basic syntax of this statement is

REPEATED *factor* / *options*;

(there are more complex forms of the statement). Here *factor* is the name of the within-subjects factor and must not have been previously defined in the data set.

The previous data set reorganized for the REPEATED statement is:

```
data repeated;
input rep var cut1 cut2 cut3 cut4 @@;
cards;
1 1 2.80191  3.73092  3.09856  2.50965
1 2 2.76212  5.40530  3.82431  2.72992
1 3 2.29151  3.81140  2.92575  2.39863
1 4 2.56631  4.96070  2.81734  2.05752

2 1 2.96602  4.43545  3.10607  2.57299
2 2 3.09636  3.90683  3.26229  2.58614
2 3 2.54027  3.82716  2.86727  2.16287
2 4 2.31630  3.96629  2.91461  2.15764

3 1 2.43232  4.32311  2.81030  2.07966
3 2 3.09917  4.08859  3.13148  2.60316
3 3 2.41199  4.08317  3.03906  2.07076
3 4 2.65834  3.71856  2.92922  2.15684

4 1 2.93509  3.99711  2.77971  2.44033
4 2 2.65256  5.42879  2.70891  2.30163
4 3 2.30420  3.27852  2.72711  2.04933
4 4 2.47877  3.92048  3.06191  2.35822

5 1 2.42277  3.85657  3.24914  2.34131
5 2 2.63666  3.77458  3.09734  2.30082
5 3 2.36941  3.44835  2.50562  2.08980
5 4 2.23595  4.02985  2.85279  1.85736
;
proc anova;
  class var;
  model cut1 cut2 cut3 cut4= var/ nouni;
  repeated time /nom;
run; quit;
```

Note that the 4 cuttings are now to the left of the model statement as four dependent variables. The option "/nouni" prevents SAS to perform and print one univariate analysis for each cutting. After the statement REPEATED, *time* is the name of the within-subjects factor and was not previously defined in the data set. The option "/**nom**" indicates SAS to perform a univariate ANOVA and **no**t a **M**ANOVA. In general the univariate tests are more capable of detecting existing differences than their multivariate counterparts.

Same CRD as a Repeated Measure Analysis

```
Class Level Information
Class     Levels    Values
VAR            4    1 2 3 4

Number of observations in data set = 20

Repeated Measures Analysis of Variance
Repeated Measures Level Information

Dependent Variable       CUT1      CUT2      CUT3      CUT4
    Level of TIME          1         2         3         4



Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source              DF      Anova SS  Mean Square  F Value    Pr > F
VAR                  3      2.840753     0.946918     7.40    0.0025
Error               16      2.048682     0.128043


Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source: TIME
                                                  Adj  Pr > F
DF      SS      MS      F Value    Pr > F    G - G     H - F
3      37.45   12.48     130.46    0.0001    0.0001    0.0001

Source: TIME*VAR
                                                  Adj  Pr > F
DF      SS      MS      F Value   Pr > F     G - G     H - F
9      0.55    0.06       0.64     0.7606    0.6585    0.6948

Source: Error(TIME)
DF      SS      MS
48     4.59    0.096

                  Greenhouse-Geisser Epsilon = 0.4984
                         Huynh-Feldt Epsilon = 0.6413
```

Note that this analysis produces the same F values than the previous split plot analysis but "**correct**" probability values for TIME (previous B_cut) and TIME*VAR (previous B_cut*A_var)

The two additional columns G-G and H-F present two adjusted probabilities that attempt to correct for the unequal correlations among pairs of repeated measures. Both of the tests are carried out by estimating a quantity known as “Epsilon” and then multiplying the numerator and denominator degrees of freedom by this estimate before determining the significance level for the test labeled Adj Pr> F in the output. This is a less drastic reduction of the degrees of freedom than the conservative approach used in the previous split plot that divides the degrees of freedom of B, B*A and Error by the degrees of freedom of B. However, these probabilities are more conservative than the split plot with unadjusted degrees of freedom.

The *F* tests will be correct if the assumptions of the analysis are correct. These assumptions (sphericity tests) can be tested by adding the PRINTE option after “/ nom” (SAS for Linear Models 3rd Ed. p274). The test labeled “Applied to Orthogonal Components” is the one that is important in determining whether the univariate *F* tests for the within-subjects effects are valid. Somewhat simplified, what is being tested is the assumptions that the variances and correlations are the same among the various dependent variables.

The tests labeled G-G (Greenhouse-Geisser) and H-F (Huynh-Feldt) present two adjusted probabilities that attempt to correct for the unequal correlations among pairs of repeated measures. The G-G adjustment is the more conservative one.

12. 9. 3. 1. Mean comparison (Cody & Smith 1991, Chapter 8))

The following keywords can be added after the *factor name* to define single-degree-of freedom contrasts for that factor specified in the REPEATED statement. Since the number of contrasts generated is always one less than the number of levels of the factor, SAS provides some control over which contrast is omitted from the analysis.

The keyword CONTRAST (n) with the REPEATED statement (REPEATED factor_name CONTRAST(n)) generates contrasts between levels of the factor and a reference level.

REPEATED time CONTRAST(1);

Where (n) is a number from 1 to k, with k being the number of levels of the repeated factor. CONTRAST(1) compares the first level of the factor with each of the other levels. It the first level were a control value, for example, this statement would compare the control to each of the other levels. If we want al of the pairwise contrasts we need to write K-1 repeated statements.

The keyword HELMERT generates contrasts between each level of the factor and the mean of the subsequent levels. The keyword MEANS (n) generates contrasts between levels of the factor and the mean of all other levels of the factor. The level specified by (n) is the eliminated contrast.

Which to use, Split Plot or REPEATED?

There is no simple answer as to which of these methods is most appropriate. If there is a clear trend to the data then the assumption of independence of errors is violated and the split plot method with unadjusted df is not appropriate. If responses are significant even when the conservative adjusted df are used, those conclusions are valid.

For the cases where different conclusions are obtained with the adjusted and unadjusted de, one might think that the REPEATED statement would solve all of the problems. It does solve the problem of correlation, but the interpretation of the results requires a number of assumptions or tests that are beyond the scope of this course. You can use the REPEATED statement, but *only* after making sure you understand and test all the assumptions of the model.

Missing values: In cases were many repeated measurements observations are missing, the split-plot approach might be the only way of analyzing a repeated-measures design, because the REPEATED statement ignores cases where **any** of the repeated measures is missing.