Chapter 13 : The Meadowfoam experiment, a CRD factorial

Linear Models – $Y_{ij} = \mu + \tau_{1i} + \tau_{2j} + \tau_1\tau_{2ij} + \varepsilon_{ijk}$ This is a completely randomized design with a 2 by 6 factorial treatment arrangement (two-way ANOVA). In this case the treatments are fixed. The EMS below is for the Meadowfoam example where treatment "TIME" is time (t=2 levels) and "INTENSITY" is light levels (l=6 levels). Within each combination of time by light there are n = 2 replicates.

```
1          ********************************************************;
2          *** The effect of light on Meadowfoam flowering.      ***;
3          *** Results of an experiment where the effedt of six  ***;
4          *** levels of light intensity and the timing of the   ***;
5          *** light treatment was investigated.                 ***;
6          ********************************************************;
7
8          dm'log;clear;output;clear';
9          options nodate nocenter nonumber ps=512 ls=99 nolabel;
10         ODS HTML style=minimal rs=none
10       ! body='C:\Geaghan\Current\EXST3201\Fall2005\SAS\Meadowfoam03.html' ;
NOTE: Writing HTML Body file: C:\Geaghan\Current\EXST3201\Fall2005\SAS\Meadowfoam03.html
11
12         Title1 'Chapter 9 : The effect of light on Meadowfoam flowering';
13         filename input1 'C:\Geaghan\Current\EXST3201\Datasets\ASCII\case0901.csv';
14
15         data Meadowfoam; infile input1 missover DSD dlm="," firstobs=2;
16            input FLOWERS TIME INTENSITY;
17              label Flowers = 'Average number of flowers per plant'
18                    Time = 'Early and Late'
19                    Intensity = 'Level of light intensity';
20            IntensityAgain = Intensity;
21            TimeName = 'Early'; if time eq 1 then Timename = 'Late';
22         datalines;
NOTE: The infile INPUT1 is:
      File Name=C:\Geaghan\Current\EXST3201\Datasets\ASCII\case0901.csv,
      RECFM=V,LRECL=256
NOTE: 24 records were read from the infile INPUT1.
      The minimum record length was 8.
      The maximum record length was 24.
NOTE: The data set WORK.MEADOWFOAM has 24 observations and 5 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds
23         run;
```

The preferred SAS procedure for Analysis of Variance is PROC MIXED. On some occasions we may want to see the sums of squares and will use PROC GLM. PROC GLM can also produce the Expected Mean Squares. The first two examples of the Meadowfoam analysis are done as a PROC GLM as if the effects are random and PROC GLM with fixed effects.

```
25          options ps=512 ls=111;
26          PROC GLM DATA=Meadowfoam; class Intensity TimeName;
27               Title2 'Two-way analysis of variance (CRD factorial) using PROC GLM';
28               Title3 'Analysis done as if the treatments were random';
29             MODEL Flowers = Intensity | TimeName;
30             random Intensity | TimeName;
31             test h=Intensity TimeName e=Intensity * TimeName;
32          RUN;
NOTE: TYPE I EMS not available without the E1 option.
34          options ps=512 ls=111;
NOTE: The PROCEDURE GLM printed pages 1-4.
NOTE: PROCEDURE GLM used (Total process time):
      real time           0.18 seconds
      cpu time            0.05 seconds
```

```
Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC GLM
Analysis done as if the treatments were random

The GLM Procedure

Class Level Information
Class          Levels    Values
INTENSITY           6    150 300 450 600 750 900
TimeName            2    Early Late

Number of Observations Read          24
Number of Observations Used          24

Dependent Variable: FLOWERS             Sum of
Source                      DF          Squares     Mean Square    F Value    Pr > F
Model                       11     3682.011162      334.728287       6.12    0.0020
Error                       12      655.925105       54.660425
Corrected Total             23     4337.936267

R-Square     Coeff Var      Root MSE    FLOWERS Mean
0.848793      13.16993      7.393269        56.13750

Source                      DF       Type I SS     Mean Square    F Value    Pr > F
INTENSITY                    5     2683.513751      536.702750       9.82    0.0006
TimeName                     1      886.950342      886.950342      16.23    0.0017
INTENSITY*TimeName           5      111.547069       22.309414       0.41    0.8342

Source                      DF     Type III SS     Mean Square    F Value    Pr > F
INTENSITY                    5     2683.513751      536.702750       9.82    0.0006
TimeName                     1      886.950342      886.950342      16.23    0.0017
INTENSITY*TimeName           5      111.547069       22.309414       0.41    0.8342

Source                  Type III Expected Mean Square
INTENSITY               Var(Error) + 2 Var(INTENSITY*TimeName) + 4 Var(INTENSITY)
TimeName                Var(Error) + 2 Var(INTENSITY*TimeName) + 12 Var(TimeName)
INTENSITY*TimeName      Var(Error) + 2 Var(INTENSITY*TimeName)
Error                   Var(Error)       (this line was not created by SAS)
```

The last term is the variance component for the residual error term "Var(Error)". This term is omitted by PROC GLM.

```
Tests of Hypotheses Using the Type III MS for INTENSITY*TimeName as an Error Term
Source                      DF     Type III SS     Mean Square    F Value    Pr > F
INTENSITY                    5     2683.513751      536.702750      24.06    0.0017
TimeName                     1      886.950342      886.950342      39.76    0.0015
```

Source	d.f.	Expected Mean Squares (Random)
Treatment 1	$\tilde{t}_1 - 1$	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + \tilde{n}t_2\sigma^2_{\tau 1}$
Treatment 2	$\tilde{t}_2 - 1$	$\sigma^2 + n\sigma^2_{\tau 1\tau 2} + nt_1\sigma^2_{\tau 2}$
Interaction	$(\tilde{t}_1 - 1)(\tilde{t}_2 - 1)$	$\sigma^2 + n\sigma^2_{\tau 1\tau 2}$
Error	$\tilde{t}_1\tilde{t}_2 (n - 1)$	σ^2

```
35         PROC GLM DATA=Meadowfoam; class Intensity TimeName;
36              Title2 'Two-way analysis of variance (CRD factorial) using PROC GLM';
37              Title3 'Analysis done as if the treatments were fixed';
38              Title4 'Linear trend and Lack of Fit calculated as a joint contrast';
39            MODEL Flowers = Intensity | TimeName;
40            contrast 'linear'    intensity -5  -3  -1   1   3   5;
41            contrast 'quadratic' intensity  5  -1  -4  -4  -1   5;
42            contrast 'cubic'     intensity -5   7   4  -4  -7   5;
43            contrast 'quartic'   intensity  1  -3   2   2  -3   1;
44            contrast 'quintic'   intensity -1   5 -10  10  -5   1;
45            contrast 'LOF'        intensity  5  -1  -4  -4  -1   5,
46                                 intensity -5   7   4  -4  -7   5,
47                                 intensity  1  -3   2   2  -3   1,
48                                 intensity -1   5 -10  10  -5   1;
49         RUN;
```

```
NOTE: The PROCEDURE GLM printed pages 5-6.
NOTE: PROCEDURE GLM used (Total process time):
      real time           0.12 seconds
      cpu time            0.05 seconds

Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC GLM
Analysis done as if the treatments were fixed
Linear trend and Lack of Fit calculated as a joint contrast

The GLM Procedure

Class Level Information
Class          Levels    Values
INTENSITY           6    150 300 450 600 750 900
TimeName            2    Early Late

Number of Observations Read          24
Number of Observations Used          24

Dependent Variable: FLOWERS
                                    Sum of
Source                      DF         Squares     Mean Square    F Value    Pr > F
Model                       11     3682.011162      334.728287       6.12    0.0020
Error                       12      655.925105       54.660425
Corrected Total             23     4337.936267

R-Square     Coeff Var      Root MSE    FLOWERS Mean
0.848793      13.16993      7.393269        56.13750

Source                      DF       Type I SS     Mean Square    F Value    Pr > F
INTENSITY                    5     2683.513751      536.702750       9.82    0.0006
TimeName                     1      886.950342      886.950342      16.23    0.0017
INTENSITY*TimeName           5      111.547069       22.309414       0.41    0.8342

Source                      DF     Type III SS     Mean Square    F Value    Pr > F
INTENSITY                    5     2683.513751      536.702750       9.82    0.0006
TimeName                     1      886.950342      886.950342      16.23    0.0017
INTENSITY*TimeName           5      111.547069       22.309414       0.41    0.8342

Contrast                    DF     Contrast SS     Mean Square    F Value    Pr > F
linear                       1     2579.750045     2579.750045      47.20    <.0001
quadratic                    1       63.527416       63.527416       1.16    0.3022
cubic                        1       11.781117       11.781117       0.22    0.6508
quartic                      1        9.315100        9.315100       0.17    0.6870
quintic                      1       19.140073       19.140073       0.35    0.5650
LOF                          4      103.763706       25.940926       0.47    0.7538
```

Source	d.f.	Expected Mean Squares (Random)
Treatment 1	$\tilde{t_1} - 1$	$\sigma^2 + Q_{\tau 1}$
Treatment 2	$\tilde{t_2} - 1$	$\sigma^2 + Q_{\tau 2}$
Interaction	$(\tilde{t_1} - 1)(\tilde{t_2} - 1)$	$\sigma^2 + Q_{\tau 1 \tau 2}$
Error	$\tilde{t_1}\tilde{t_2}(n - 1)$	σ^2

```
51         PROC GLM DATA=Meadowfoam; class Intensity TimeName;
52              Title2 'AnCova with PROC GLM';
53              Title3 'This ANCOVA was previously done with PROC REG';
54              Title4 'Linear trend fitted with quantitative variable';
55              Title5 'Note request of solution to get regression coefficients';
56           MODEL Flowers = intensityagain Intensity TimeName intensityagain*TimeName Intensity*TimeName;
57         RUN;
58
NOTE: The PROCEDURE GLM printed pages 7-8.
NOTE: PROCEDURE GLM used (Total process time):
      real time           0.13 seconds
      cpu time            0.04 seconds
```

```
Chapter 9 : The effect of light on Meadowfoam flowering
AnCova with PROC GLM
This ANCOVA was previously done with PROC REG
Linear trend fitted with quantitative variable
Note request of solution to get regression coefficients

The GLM Procedure

Class Level Information
Class          Levels    Values
INTENSITY           6    150 300 450 600 750 900
TimeName            2    Early Late

Number of Observations Read          24
Number of Observations Used          24

                                    Sum of
Source                      DF         Squares     Mean Square    F Value    Pr > F
Model                       11     3682.011162      334.728287       6.12    0.0020
Error                       12      655.925105       54.660425
Corrected Total             23     4337.936267

R-Square     Coeff Var      Root MSE    FLOWERS Mean
0.848793      13.16993      7.393269        56.13750

Source                      DF       Type I SS     Mean Square    F Value    Pr > F
IntensityAgain               1     2579.750045     2579.750045      47.20    <.0001
INTENSITY                    4      103.763706       25.940926       0.47    0.7538
TimeName                     1      886.950342      886.950342      16.23    0.0017
IntensityAg*TimeName         1        0.576035        0.576035       0.01    0.9199
INTENSITY*TimeName           4      110.971034       27.742759       0.51    0.7314

Source                      DF     Type III SS     Mean Square    F Value    Pr > F
IntensityAgain               0       0.0000000               .          .         .
INTENSITY                    4     103.7637056      25.9409264       0.47    0.7538
TimeName                     1     153.2160135     153.2160135       2.80    0.1199
IntensityAg*TimeName         0       0.0000000               .          .         .
INTENSITY*TimeName           4     110.9710344      27.7427586       0.51    0.7314


59         PROC mixed DATA=Meadowfoam cl covtest; class Intensity TimeName;
60              Title2 'Two-way analysis of variance (CRD factorial) using PROC MIXED';
61              Title3 'Analysis done as if the treatments were random';
62             MODEL Flowers = / outp=resids;
63             Random Intensity | TimeName;
64         RUN;
NOTE: Convergence criteria met.
NOTE: Estimated G matrix is not positive definite.
NOTE: The data set WORK.RESIDS has 24 observations and 12 variables.
NOTE: The PROCEDURE MIXED printed page 9.
NOTE: PROCEDURE MIXED used (Total process time):
      real time           0.26 seconds
      cpu time            0.14 seconds


Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC MIXED
Analysis done as if the treatments were random

The Mixed Procedure

Model Information
Data Set                     WORK.MEADOWFOAM
Dependent Variable           FLOWERS
Covariance Structure         Variance Components
Estimation Method            REML
Residual Variance Method     Profile
Fixed Effects SE Method      Model-Based
Degrees of Freedom Method    Containment
```

```
Class Level Information
Class        Levels    Values
INTENSITY         6    150 300 450 600 750 900
TimeName          2    Early Late

Dimensions
Covariance Parameters             4
Columns in X                      1
Columns in Z                     20
Subjects                          1
Max Obs Per Subject              24

Number of Observations
Number of Observations Read              24
Number of Observations Used              24
Number of Observations Not Used           0

 Iteration History
Iteration    Evaluations    -2 Res Log Like       Criterion

        0              1       188.96140138
        1              2       171.43387902      0.00002298
        2              1       171.43235144      0.00000005
        3              1       171.43234832      0.00000000
                  Convergence criteria met.

Covariance Parameter Estimates
                               Standard        Z
Cov Parm               Estimate       Error     Value        Pr Z     Alpha       Lower       Upper
INTENSITY                122.89     84.9482      1.45      0.0740      0.05     44.8694      946.36
TimeName                 70.1500      104.54      0.67      0.2511      0.05     13.3129      149317
INTENSITY*TimeName             0           .         .           .         .           .           .
Residual                 45.1455     15.4848      2.92      0.0018      0.05     25.4206      101.46

 Fit Statistics
-2 Res Log Likelihood           171.4
AIC (smaller is better)         177.4
AICC (smaller is better)        178.7
BIC (smaller is better)         176.8


66         PROC mixed DATA=Meadowfoam cl covtest; class Intensity TimeName;
67              Title2 'Two-way analysis of variance (CRD factorial) using PROC MIXED';
68              Title3 'Analysis done as if the treatments were fixed';
69            MODEL Flowers = Intensity | TimeName / outp=resids;
70            contrast 'linear'    intensity -5  -3  -1   1   3   5;
71            contrast 'quadratic' intensity  5  -1  -4  -4  -1   5;
72            contrast 'cubic'     intensity -5   7   4  -4  -7   5;
73            contrast 'quartic'   intensity  1  -3   2   2  -3   1;
74            contrast 'quintic'   intensity -1   5 -10  10  -5   1;
75            contrast 'LOF'       intensity  5  -1  -4  -4  -1   5,
76                                 intensity -5   7   4  -4  -7   5,
77                                 intensity  1  -3   2   2  -3   1,
78                                 intensity -1   5 -10  10  -5   1;
79            lsmeans Intensity | TimeName / pdiff adjust=tukey cl;
80           ods output diffs=ppp lsmeans=mmm;
81           ods listing exclude diffs lsmeans;
82         run;
NOTE: The data set WORK.MMM has 20 observations and 11 variables.
NOTE: The data set WORK.PPP has 82 observations and 17 variables.
NOTE: The data set WORK.RESIDS has 24 observations and 12 variables.
NOTE: The PROCEDURE MIXED printed page 10.
NOTE: PROCEDURE MIXED used (Total process time):
      real time           0.91 seconds
      cpu time            0.83 seconds
83         %include 'C:\Geaghan\Current\EXST3201\Fall2005\SAS\pdmix800.sas';
756        %pdmix800(ppp,mmm,alpha=.05,sort=yes);
757        run;
759        quit;
```

Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC MIXED
Analysis done as if the treatments were fixed

The Mixed Procedure

Model Information

Data Set	WORK.MEADOWFOAM
Dependent Variable	FLOWERS
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
INTENSITY	6	150 300 450 600 750 900
TimeName	2	Early Late

Dimensions

Covariance Parameters	1
Columns in X	21
Columns in Z	0
Subjects	1
Max Obs Per Subject	24

Number of Observations

Number of Observations Read	24
Number of Observations Used	24
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
Residual	54.6604	22.3150	2.45	0.0072	0.05	28.1071	148.95

Fit Statistics

-2 Res Log Likelihood	90.4
AIC (smaller is better)	92.4
AICC (smaller is better)	92.8
BIC (smaller is better)	92.9

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
INTENSITY	5	12	9.82	0.0006
TimeName	1	12	16.23	0.0017
INTENSITY*TimeName	5	12	0.41	0.8342

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
linear	1	12	47.20	<.0001
quadratic	1	12	1.16	0.3022
cubic	1	12	0.22	0.6508
quartic	1	12	0.17	0.6870
quintic	1	12	0.35	0.5650
LOF	4	12	0.47	0.7538

```
Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC MIXED
Analysis done as if the treatments were fixed

Effect=INTENSITY ADJUSTMENT=Tukey(P<.05) bygroup=1
```

Obs	INTENSITY	Time Name	Estimate	StdErr	Alpha	Lower	Upper	MSGROUP
1	150		73.2750	3.6966	0.05	65.2207	81.3293	A
2	300		64.1500	3.6966	0.05	56.0957	72.2043	AB
3	450		59.9000	3.6966	0.05	51.8457	67.9543	ABC
4	600		50.0500	3.6966	0.05	41.9957	58.1043	BC
5	750		45.5250	3.6966	0.05	37.4707	53.5793	C
6	900		43.9250	3.6966	0.05	35.8707	51.9793	C

```
Effect=TimeName ADJUSTMENT=Tukey(P<.05) bygroup=2
```

Obs	INTENSITY	Time Name	Estimate	StdErr	Alpha	Lower	Upper	MSGROUP
7	_	Early	62.2167	2.1343	0.05	57.5665	66.8668	A
8	_	Late	50.0583	2.1343	0.05	45.4082	54.7085	B

```
Effect=INTENSITY*TimeName ADJUSTMENT=Tukey(P<.05) bygroup=3
```

Obs	INTENSITY	Time Name	Estimate	StdErr	Alpha	Lower	Upper	MSGROUP
9	150	Early	76.7000	5.2278	0.05	65.3095	88.0905	A
10	300	Early	73.5500	5.2278	0.05	62.1595	84.9405	A
11	150	Late	69.8500	5.2278	0.05	58.4595	81.2405	AB
12	450	Early	64.0500	5.2278	0.05	52.6595	75.4405	ABC
13	600	Early	57.5500	5.2278	0.05	46.1595	68.9405	ABC
14	450	Late	55.7500	5.2278	0.05	44.3595	67.1405	ABC
15	300	Late	54.7500	5.2278	0.05	43.3595	66.1405	ABC
16	750	Early	52.9500	5.2278	0.05	41.5595	64.3405	ABC
17	900	Early	48.5000	5.2278	0.05	37.1095	59.8905	ABC
18	600	Late	42.5500	5.2278	0.05	31.1595	53.9405	BC
19	900	Late	39.3500	5.2278	0.05	27.9595	50.7405	C
20	750	Late	38.1000	5.2278	0.05	26.7095	49.4905	C

```
82         proc univariate data=resids plot normal; var resid;
83              Title4 'Univariate analysis of residuals';
84         run;
NOTE: The PROCEDURE UNIVARIATE printed page 11.
NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.11 seconds
      cpu time            0.02 seconds


Chapter 9 : The effect of light on Meadowfoam flowering
Two-way analysis of variance (CRD factorial) using PROC MIXED
Analysis done as if the treatments were fixed
Univariate analysis of residuals


Tests for Normality
Test                  --Statistic---    -----p Value------
Shapiro-Wilk          W     0.913026    Pr < W      0.0410
Kolmogorov-Smirnov    D     0.112016    Pr > D     >0.1500
Cramer-von Mises      W-Sq  0.085101    Pr > W-Sq   0.1743
Anderson-Darling      A-Sq  0.613184    Pr > A-Sq   0.0985


   Stem Leaf  Boxplot                      Normal Probability Plot
      6 28046                  5     |           7+                                  ** *++*   *
      4 154                    3  +-----+         |                                ***++++
      2 61                     2  |     |         |                              **+++
      0 51                     2  *--+--*         |                           +**+
     -0 15                     2  |     |         |                        +**
     -2 16                     2  |     |         |                    +++**
     -4 451                    3  +-----+         |               ++++***
     -6 64082                  5     |          -7+      *   *++* **
        ----+----+----+----+                        +----+----+----+----+----+----+----+----+----+----+----+
                                                       -2        -1         0        +1        +2
```

```
765        data meadowfoam; set meadowfoam;
766           if TimeName eq 'Early' then IE = intensity; else IL = .;
767           if TimeName eq 'Late' then IL = intensity; else IL = .;
768        run;
NOTE: There were 24 observations read from the data set WORK.MEADOWFOAM.
NOTE: The data set WORK.MEADOWFOAM has 24 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.02 seconds
769
770
771        GOPTIONS DEVICE=CGMLT97L ctitle=black ctext=black ftext='TimesRoman' ftitle='TimesRoman';
772
773        GOPTIONS GSFNAME=OUT2; FILENAME OUT2
       'C:\Geaghan\Current\EXST3201\Fall2005\SAS\MeadowFoam01.CGM';
774        proc gplot data=meadowfoam;
775            Title1 'Effect of light on Meadowfoam flowering';
776            plot Flowers*IL=1 Flowers*IE=2 / overlay HAXIS=AXIS1 VAXIS=AXIS2;
777               AXIS1 LABEL=('Light intensity treatment') MINOR=(N=5) order = 100 to 900 by 100;
778               AXIS2 LABEL=('Flower production') MINOR=(N=3) ORDER = 30 TO 80 BY 10;
779            symbol1 color=red  L=1 i=stdm1tjp mode=include;
780            symbol2 color=blue L=1 i=stdm1tjp mode=include;
781        run;
NOTE: 12 observation(s) contained a MISSING value for the FLOWERS * IL request.
NOTE: 12 observation(s) contained a MISSING value for the FLOWERS * IE request.
NOTE:  11 RECORDS WRITTEN TO C:\Geaghan\Current\EXST3201\Fall2005\SAS\MeadowFoam01.CGM
ERROR: Unable to open C:\Geaghan\Current\EXST3201\Fall2005\SAS\MeadowFoam01.CGM because file is in use.
ERROR: Unable to open graphics device I/O.
ERROR: Unable to initialize graphics device.
ERROR: Driver SASGDGIF will not load.
781      !      quit;
```