**SAS Library  
SAS Code for Some Advanced Experimental Designs**

**This page was adapted from a page titled SAS Code for Some Experiemental Designs  created by Oliver Schabenberger.  We thank Professor Schabenberger for permission to adapt and distribute this page via our web site.**

DISCLAIMER: No guarantee is given for the correctness of the code herein.

Recently, PROC MIXED was added to the palette of SAS/STAT procedures. One of PROC MIXED strengths is the analysis of statistical models with combined random and fixed effects. Many experimental design situations that had a non-optimal solution in the otherwise powerful GLM procedure have now become much simpler. For example tests across whole- and split-plot factors in Split-Plot experiments, Block designs with random block effects etc. There is no room to discuss the common and disparate features of the GLM and MIXED procedures in detail. The interested user is pointed to

* *SAS System for Mixed Models*. Littell, R.C.; Milliken, G.A.; Stroup, W.W.; Wolfinger, R.D., Cary, NC: SAS Institute Inc., 1996. 633pp.
* *SAS/STAT Software:Changes and Enhancements, Release 6.11*, Cary, NC: SAS Institute, Inc. 1995.
* *SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2,* Cary, NC: SAS Institute Inc., 1989. 846 pp.

Do not use the statements below blindly. The statements are rudimentary in that they do not produce all possible tests and contrasts for complex designs. Consult the manuals. This is true in particular for the TEST statement in PROC GLMas well as theLSMEANS, ESTIMATE, andCONTRAST statements in MIXED andGLM.

For a brief comparison ofPROC GLM andPROC MIXED, click [here](http://www.ats.ucla.edu/stat/sas/library/GLMvsMIXED_os.htm).

**Notation**

|  |  |
| --- | --- |
| A,B,C,... | Experimental factors |
| A1,A2,A3,... | Levels 1,2,3 of experimental factor A |
| A\*B | Interaction between factors A and B |
| Y | The experimental response (what you measured) |
| Tx | The treament factor in a non-factorial design |
| X | Supplementary information (ANCOVA) |
| WhP | Whole-plot factor in Split-Plot type design |
| SubP | Sub-plot factor in Split-Plot type design |
| Rep | Replications of a basic design |

1. Completely Randomized Design (CRD)
   1. [CRD without subsampling](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#CRD)
   2. [CRD with subsampling](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#CRDSS)
2. Randomized Complete Block Designs (RCBD)
   1. [RCBD with fixed block effects, without subsampling](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RCBDf)
   2. [RCBD with fixed block effects, with subsampling](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RCBDfss)
   3. [RCBD with random block effects](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RCBDr)
3. Multilocation RCBDs over time
   1. [Different sites at each time point](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RCBDDifferentSites)
   2. [Same site at different time points](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RCBDSameSite) (randomized separately at each time point)
4. Generalized Randomized Complete Block Design (GRBD)
   1. [GRBD with fixed block effects](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#GRBDf)
   2. [GRBD with random block effects](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#GRBDr)
5. Incomplete Block Designs (IBD)
   1. [Incomplete, fixed blocks](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#IBD)
   2. [Intra-block analysis in incomplete random blocks](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#IBDi)
   3. [Inter- and intra-block analysis in incomplete random blocks](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#IBDii)
6. Latin Square Designs (LSD)
   1. [Standard Latin Square](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#LSD)
   2. [Incomplete Latin Square](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#LSD)
   3. [Latin Rectangle](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#LSD)
   4. [Replicated LSD (rows and columns crossed w/ reps)](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RLSD1)
   5. [Replicated LSD (rows nested, columns crossed w/ reps)](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RLSD2)
   6. [Replicated LSD (rows crossed, columns nested w/ reps)](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RLSD3)
   7. [Replicated LSD (rows and columns nested w/ reps)](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#RLSD4)
   8. [Graeco-Latin Square](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#GLSD)
7. Split-Plot Designs (SPD)
   1. [Standard SPD with whole-plot factor in a RCBD, replication fixed](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SPD(RCBD,RCBD)f)
   2. [Standard SPD with whole-plot factor in a RCBD, replication random](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SPD(RCBD,RCBD)r)
   3. [SPD with whole-plot factor in a CRD](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SPD(CRD,RCBD)r)
   4. [SPD with whole-plot factor in a GRBD](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SPD(GRBD,RCBD))
8. Split-Block (Strip-Plot) Designs (SBD)
   1. [Standard SBD, replications fixed](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SBDf)
   2. [Standard SBD, replications random](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SBDr)
   3. [SBDs: same sites at different times](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SBDLocC)
   4. [SBDs: different sites](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SBDLocN)
9. [Split-Split-Plot Design (SSPD)](http://www.ats.ucla.edu/stat/sas/library/SASExpDes_os.htm#SSPD)

**Completely Randomized Design (CRD)**

* **CRD without subsampling**
* proc glm data=yourdata;
* class tx;
* model y = tx;

      run;

* **CRD with subsampling**
* proc glm data=yourdata;
* class rep tx;
* model y = rep(tx) tx;
* test h=tx e=rep(tx);

     run;

Note: SAS would use the remainder term as the error for all tests. For a CRD with subsampling this would be the observational error. But to compare treatments we need the experimental error term, which is REP(Tx). In order to use this as the error term in the E= option of the TEST statement, the term must be part of the model statement. In a CRD without subsampling, the remainder term **is** experimental error, and all tests are correct as they appear on the output.

**Randomized Complete Block Designs (RCBD)**

* **RCBD with fixed block effects, without subsampling**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class block tx;     model y = block tx;   run; | proc mixed data=yourdata;      class block tx;      model y = block tx;   run; |

* **RCBD with fixed block effects, with subsampling**
* proc glm data=yourdata;
* class block tx;
* model y = block tx block\*tx;
* test h=tx e=block\*tx;

     run;

Note: The correct error term for testing treatment effects is the block-treatment interaction. The remainder is again observational error and carries no information about the variabilty of the experimental units.

* **RCBD with random block effects**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class block tx;     model y = block tx;     lsmeans tx /stderr pdiff;     random block;   run; | proc mixed data=yourdata;      class block tx;      model y = tx;      lsmeans tx /stderr pdiff;      random block;   run; |



**Multilocation RCBDs**

* **Different sites at each time point**
* proc glm data=multi;
* class year loc block a b;
* model y = Year|A|B
* block(Loc\*Year)        /\* the block effect \*/
* loc(year)
* loc\*A(year) loc\*B(year)
* loc\*A\*B(year) /ss3 ;
* run;

quit;

* **Same site at different time points**
* proc glm data=multi;
* class year loc block a b;
* model y = Year|Loc|A|B
* block(Loc\*Year)        /\* the block effect \*/
* /ss3 ;
* run;

quit;

**Generalized Randomized Complete Block Design (GRBD)**

* **GRBD with fixed block effects**
* proc glm data=yourdata;
* class block tx;
* model y = block tx block\*tx;
* means tx block\*tx;

      run;

* **GRBD with random block effects**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class block tx;     model y = block tx block\*tx;     random block block\*tx;   run; | proc mixed data=yourdata;      class block tx;      model y = tx;      random block block\*tx;   run; |



**Incomplete Block Designs (IBD)**

* **Incomplete, fixed blocks**
* proc glm data=yourdata;
* class block tx;
* model y = block tx;
* lsmeans tx/stderr pdiff;

    run;

* **Intra-block analysis in incomplete random blocks**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class block tx;     model y = block tx;     lsmeans tx/stderr pdiff;     estimate 'Tx 1' intercep 1 tx 1;     estimate 'Tx1 vs Tx3' intercep 1 Tx 1 0 -1;     random block;   run; | proc mixed data=yourdata;     class block tx;     model y = tx;     lsmeans tx /stderr pdiff;     estimate 'Tx 1' intercept 1 tx 1;     estimate 'Tx1 vs Tx3' intercept 1 Tx 1 0 -1;     random block;   run; |

* **Inter and intra-block analysis in incomplete random blocks**
* proc mixed data=yourdata; /\* assume 6 treatments in 3 blocks of size 4 \*/
* class block tx;
* model y = tx;
* lsmeans tx /stderr pdiff;
* estimate 'Tx 1' intercept 1 tx 1;  /\* intra-analysis \*/
* estimate 'Tx 1' intercept 12 tx 6 |
* blk 1 1 1 /divisor=12; /\* inter-analysis \*/
* estimate 'Tx 1 vs Tx 3' intercept 1 Tx 1 0 -1;
* random block;

     run;

Note: Specifying the second estimate statement in the intra-block GLM analysis above leads to exactly the same estimate, standard error, and test as the first estimate since GLM does not utilize the random effects in calculating standard errors for the estimates.

**Latin Square Designs**

* **Standard Latin Square, Latin Rectangle and Incomplete Latin Squares**
* proc glm data=yourdata;
* class row col tx;
* model y = row col tx;

    run;

* **Replicated LSD (rows and columns crossed w/ reps)**
* proc glm data=yourdata;
* class rep row col tx;
* model y = rep row col tx;

    run;

* **Replicated LSD (rows nested, columns crossed w/ reps)**
* proc glm data=yourdata;
* class rep row col tx;
* model y = rep row(rep) col tx;

    run;

* **Replicated LSD (rows crossed, columns nested w/ reps)**
* proc glm data=yourdata;
* class rep row col tx;
* model y = rep row col(rep) tx;

    run;

* **Replicated LSD (rows and columns nested w/ reps)**
* proc glm data=yourdata;
* class rep row col tx;
* model y = rep row(rep) col(rep) tx;

    run;

* **Graeco Latin Square**
* proc glm data=yourdata;
* class greek row col tx;
* model y = row col greek tx;

    run;

**Split-Plot Designs (SPD)**

* **Standard SPD with whole-plot factor in a RCBD, replication fixed**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class rep whP subP;      model y = rep whP rep\*whP subP subP\*whP;      test h=whP e=rep\*whP;  run; | proc mixed data=yourdata;      class rep whP subP;      model y = rep whP subP subP\*whP/ddfm=satterth;      random rep\*whP;   run; |

* Note: The PROC MIXED statements trick SAS into using the correct error terms without assuming that the REP effect is random. The SATTERTH option in the MODEL statement of MIXED requests calculation of degrees of freedom by the Satterthwaite procedure.
* **Standard SPD with whole-plot factor in a RCBD, replication random**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class rep whP subP;      model y = rep whP rep\*whP subP subP\*whP;      random rep rep\*whOP / test;      test h=whP e=rep\*whP;  run; | proc mixed data=yourdata;      class rep whP subP;      model y = whP subP subP\*whP/ddfm=satterth;      random rep rep\*whP;   run; |

* Note: Standard errors for least square means, estimates, and contrasts in GLM are mostly incorrect. Only a single error term can be specified to override the default. For most comparisons the error term is in fact a linear combination of various variance components. GLM does not compute these combinations or allow the user to use them in the formulation of test statistics. PROC MIXED always chooses the correct error terms and should be preferred. The SATTERTH option in the MODEL statement of MIXED requests calculation of degrees of freedom by the Satterthwaite procedure.
* **SPD with whole-plot factor in a CRD**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class unit A B;     model y = A unit(A) B A\*B;     random unit(A) / test;     lsmeans A / Stderr E=unit(A);   run; | proc mixed data=yourdata;     class unit A B;     model y = A B A\*B / ddfm=satterth;     random unit(A);     lsmeans A B A\*B;   run; |

* Note: GLM can not compute standard errors for B and A\*B means. Only the standard errors for A means are correct. Use PROC MIXED instead. The SATTERTH option in the MODEL statement of MIXED requests calculation of degrees of freedom by the Satterthwaite procedure.
* **SPD with whole-plot factor in a GRBD**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class rep a b c;      model y = a b a\*b rep(a\*b)                c a\*c b\*c a\*b\*c;      test h=h b\*a  e=rep(a\*b);   run; | proc mixed data=yourdata;      class rep a b c;      model y = a b a\*b                s a\*c b\*c a\*b\*c;      random rep(a\*b);   run; |



**Split-Block (Strip-Plot) Designs (SBD)**

* **Standard SBD, replications fixed**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class rep A B;      model y = rep A rep\*A B rep\*B A\*B;      test  h = A e=rep\*A;      test  h = B e=rep\*B;   run; | proc mixed data=yourdata;     class rep A B;     model y = rep A B A\*B /ddfm=satterth ;     random rep\*A rep\*B;   run; |

* **Standard SBD, replications random**
* proc mixed data=yourdata;
* class rep A B;
* model y = A B A\*B /ddfm=satterth ;
* random rep rep\*A rep\*B rep\*A\*B;

     run;

* **SBDs: same sites at different times**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class site rep A B;      model y = Site rep  Site\*rep                A  A\*Site A\*Rep\*Site                B  B\*Site B\*Rep\*Site                A\*B  A\*B\*Site;      test  h = A    e=A\*Rep\*Site;      test  h = B    e=B\*Rep\*Site;      test  h=A\*Site e=A\*Rep\*Site;      test  h=B\*Site e=B\*Rep\*Site;   run; | proc mixed data=yourdata;      class site rep A B;      model y = Site rep rep\*Site                A    A\*Site                B    B\*Site                A\*B  A\*B\*Site;      random rep\*B\*site rep\*A\*site;   run; |

* Note: Reps are crossed with sites (locations).
* **SBDs: different sites**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;      class site rep A B;      model y = Site rep  rep(Site)                A  A\*Site A\*Rep(Site)                B  B\*Site B\*Rep(Site)                A\*B  A\*B\*Site;      test  h = A    e=A\*Rep(Site);      test  h = B    e=B\*Rep(Site);      test  h=A\*Site e=A\*Rep(Site);      test  h=B\*Site e=B\*Rep(Site);   run; | proc mixed data=yourdata;      class site rep A B;      model y = Site rep  rep(Site)                A  A\*Site                B  B\*Site                A\*B  A\*B\*Site;      random rep\*B(site) rep\*A(site);   run; |

* Note: Reps are nested with sites (locations).

**Split-Split-Plot Design (SSPD)**

|  |  |
| --- | --- |
| **GLM** | **MIXED** |
| proc glm data=yourdata;     class rep a b c;     model y = rep     a   rep\*a               b       a\*b rep\*a\*b               c       a\*c b\*c a\*b\*c;     test h=a   e=rep\*a;     test h=b   e=rep\*a\*b;     test h=a\*b e=rep\*a\*b;   run; | proc mixed data=yourdata;     class rep a b c;     model y = rep     a               b       a\*b               c       a\*c b\*c a\*b\*c ;     random rep\*a rep\*a\*b;   run; |