

# Level III: Proc GLM, Proc MIXED, Proc GLIMMIX – an overview – CRD

The goals of this workshop are:

- to compare Proc GLM, Proc MIXED, Proc GLIMMIX using a Completely Randomized Design (CRD) for the example by:
  - showing coding differences
  - showing output differences
- to provide guidelines/explanations as to why and when you would use GLM, MIXED, and GLIMMIX

Proc GLIMMIX, appears to be the “new” kid on the block when it comes to analyzing our data. But believe it or not, GLIMMIX has existed for many years, but never really caught on, until a few years ago. Many of us now are relearning our traditional analyses methods in SAS and converting to GLIMMIX.

There will be several workshops that will concentrate on the use of Proc GLIMMIX. The idea is that we will start with the straightforward experimental designs and increase the complexity to showcase the strengths of GLIMMIX and maybe convince you to make the switch to this more robust SAS procedure. This workshop will use the basic Completely Randomized Design to primarily show coding and output differences among the 3 procedures.

## History of ANOVA analyses in SAS

1966 – SAS is released with Proc ANOVA, which is to be used with:

- balanced data ONLY!
- FIXED effects ONLY!
- NOTE from SAS Online Docs: “**Caution:** If you use PROC ANOVA for analysis of unbalanced data, you must assume responsibility for the validity of the results.

1976 – SAS released Proc GLM

- balanced (Type I SS) and unbalanced (Type III SS)
- RANDOM statement introduced – provides EMS (expected mean squares equations, but you need to do the calculations!)

1992 – Proc MIXED

- RANDOM statement incorporated
- REPEATED statement introduced
- “Normally distributed” data ONLY
- linear effects

1992 – Proc GENMOD

- Non-normal data
- Fixed effects ONLY

xxxx? – Proc NLMIXED

- normal, binomial, Poisson distributions
- nonlinear effects

2005 – Proc GLIMMIX

- Proc MIXED
- Proc NLMIXED
- Non-normal data

## Completely Randomized Design

Our fictitious design has 6 treatments (A, B, C, D, F, G) with 4 observations per treatment. Our Null Hypothesis states that all treatment means are equal, with our Alternate Hypothesis stating that at least 2 means are not equal. We will have a model to reflect this design as:

Outcome variable(Weight) = overall mean + Treatment effect + residual error

```

/*****
/*
/*  Reading data gathered from a Completely Randomized Design conducted across 6
/*  treatments. Variables include treatments and weight collected in hypothetical units*/
/*  This is dummy data created for the purposes of a demonstration and workshop
/*  Created and entered by A.M. Edwards  May 23, 2017
/*
/*
/*****/

```

```

Data crd;
  input ID trmt$ weight;
  datalines;

```

1	A	41
2	A	24
3	A	33
4	A	38
5	B	24
6	B	21
7	B	16
8	B	43
9	C	46
10	C	33
11	C	14
12	C	19
13	D	32
14	D	38
15	D	15
16	D	17
17	F	31
18	F	15
19	F	36
20	F	46

21	G	28
22	G	40
23	G	37
24	G	39

```
;
Run;
```

## Proc GLM – General Linear Model

Proc GLM was the second generation PROCedure developed in SAS to conduct ANOVAs (analysis of variance). This Proc is still used today for situations where you have a FIXED effects model and a balanced design – same number of observations in each treatment group.

```
Proc glm data=crd;
  class trmt;
  model weight = trmt;
  title "Proc GLM Results";
Run;
Quit;
```

**Proc glm** – calls on the GLM Procedure. data=crd – specifies the dataset which you want Proc GLM to use.

**Class** statement – list your classification variables here. Think of these variables are those that tell you which group your observations fall into.

**Model** statement – this should be based on your experimental design. In this case we have a CRD – our dependent variable = independent variable or our fixed effect.

**Title** statement – another great little habit to start. Create a title statement for each procedure you use. This way you will have a title at the top of our output window. You will never guess again as to what that output was about. If you want more titles or subtitles simply type title2 or title 3, etc.... You can also use the Footnote option to add notes to the bottom of our output page.

**Run** statement finishes the Procedure.

**Quit** statement will let SAS know that you do not want to add any more information to the Proc GLM. Proc GLM is one of the few SAS Procedures that will wait for more instructions by running in the background. In order to close it out, you will need to add a Quit.

[View Proc GLM Results](#)

## Proc MIXED

With the increasing use of mixed models – models that include both fixed and random effects, Proc MIXED was developed. Proc MIXED can also account for unbalanced designs. Using the same CRD dataset:

```
Proc mixed data=crd;  
  class trmt;  
  model weight = trmt;  
  title "Proc MIXED Results";  
Run;
```

You should obtain the SAME results with both procedures with a basic CRD design. For most straightforward models, Proc GLM and Proc MIXED should yield the same results.

**Proc mixed** – calls on the MIXED Procedure. data=crd – specifies the dataset which you want Proc MIXED to use.

**Class** statement – list your classification variables here. Think of these variables are those that tell you which group your observations fall into.

**Model** statement – this should be based on your experimental design. In this case we have a CRD – our dependent variable = independent variable or our fixed effect.

**Run** statement finishes the Procedure.

[View Proc MIXED Results](#)

## Proc GLIMMIX

Proc GLIMMIX does it all! ok, almost. For our purposes, Proc GLIMMIX handles the different types of experimental designs that are used in OAC and in the agricultural field.

```
Proc glimmix data=crd;  
  class trmt;  
  model weight = trmt;  
  title "Proc GLIMMIX Results";  
Run;
```

**Proc glimmix** – calls on the GLIMMIX Procedure. data=crd – specifies the dataset which you want Proc GLIMMIX to use.

**Class** statement – list your classification variables here. Think of these variables are those that tell you which group your observations fall into.

**Model** statement – this should be based on your experimental design. In this case we have a CRD – our dependent variable = independent variable or our fixed effect.

**Run** statement finishes the Procedure.

[View Proc GLIMMIX Results](#)

## Randomized Complete Block Design (RCBD)

```
/******  
/*  
/*  Reading data gathered from a Randomized Complete Block Design conducted across  /*  
/*  4 blocks with 6 treatments from Kuehl (2000).  /*  
/*  Entered by A.M. Edwards  May 25, 2017  /*  
/*  
/******
```

```
Data rcdb;  
  input block trmt Nitrogen;  
  datalines;
```

1	1	34.98
1	2	40.89
1	3	42.07
1	4	37.18
1	5	37.99
1	6	34.89
2	1	41.22
2	2	46.69
2	3	49.42
2	4	45.85
2	5	41.99
2	6	50.15
3	1	36.94
3	2	46.65
3	3	52.68
3	4	40.23
3	5	37.61
3	6	44.57
4	1	39.97
4	2	41.9
4	3	42.91
4	4	39.2
4	5	40.45
4	6	43.29

```
;  
Run;
```

```

/* Proc GLM Statements */

Proc glm data=rcbd;
  class block trmt;
  model Nitrogen = block trmt;
  random block;
  title "Proc GLM Results";
Run;
Quit;

/* Proc MIXED Statements with an LSMEANS for treatment differences */

Proc mixed data=rcbd;
  class block trmt;
  model Nitrogen = trmt;
  random block;
  title "Proc MIXED Results";
Run;

/* Proc GLIMMIX Statements with an LSMEANS for treatment differences */

Proc glimmix data=rcbd;
  class block trmt;
  model Nitrogen = trmt;
  random block;
  title "Proc GLIMMIX Results";
Run;

```

The SAS program can be found [here](#) – please note that it is a PDF file

[Proc GLM Results](#)

[Proc MIXED Results](#)

[Proc GLIMMIX Results](#)