INTEGO GROUP

PROC MIXED: A Complicated Procedure in Simple Words

Kateryna Fedoryshyna

PHUSE 2021 EU Connect
Kyiv, Ukraine

15th–19th November 2021

INTEGO GROUP

Agenda

1. Introduction
2. Fundamentals of PROC MIXED (Syntax, Type of effects, Covariance structures)
3. Basic example
4. Efficacy analysis of continuous endpoints
5. Summary

INTEGO GROUP

Introduction

- ✓ Statistical analysis plays a crucial role in conducting clinical trials.
- ✓ SAS contains various statistical procedures for different types of statistical analysis.
- ✓ PROC MIXED is widely used for mixed linear models.

INTEGO GROUP

Fundamentals of PROC MIXED (Type of effects)

```
PROC MIXED options;
  BY variables;
  CLASS variables;
  MODEL dependent = fixed-effects / options;
  RANDOM random-effects / options;
  REPEATED repeated-effect / options;
  CONTRAST "label" fixed-effect values | random-effect values / options;
  ESTIMATE "label" fixed-effect values | random-effect values / options;
  LSMEANS fixed-effects / options;
  LSMESTIMATE model-effect lsmestimate-specification / options;
RUN;
```

INTEGO GROUP

Fundamentals of PROC MIXED (Type of effects)

Fixed Effects

are those factors whose levels are fixed before conducting the experiment, and the researcher is interested in the difference in the response variable among those levels included in the study.

Examples: arm, gender, age, stratification factors.

Random Effects

are those factors that have a lot of levels, and a random subset of them is selected from a large population of levels.

Example: study center.

Models in which some factors are ***fixed*** effects and other factors are ***random*** effects are called ***mixed*** models.

INTEGO GROUP

Fundamentals of PROC MIXED (Covariance structures)

The **standard** linear model :

$$Y = X\beta + \varepsilon$$

The **mixed** linear model :

$$Y = X\beta + Z\gamma + \varepsilon$$

Assuming that the random effect γ and the residuals ε are independently and normally distributed with

$$E[\gamma] = 0, E[\varepsilon] = 0, V[\gamma] = G, V[\varepsilon] = R,$$

we get:

$$E(Y) = X\beta,\ Var(Y) = ZGZ^T + R.$$

The various covariance structures for G and R matrices can be specified in the RANDOM/REPEATED statements in PROC MIXED within TYPE= option correspondingly.

Fundamentals of PROC MIXED Covariance Structures

- Variance components: type = VC

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

- Compound symmetry: type = CS

$$\begin{bmatrix} (\sigma^2 + \sigma_1) & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & (\sigma^2 + \sigma_1) & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & (\sigma^2 + \sigma_1) & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & (\sigma^2 + \sigma_1) \end{bmatrix}$$

- Unstructured: type = UN

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} & \sigma_{41} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} & \sigma_{42} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{43} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

- First-order autoregressive: type = AR(1)

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

INTEGO GROUP

Basic example

```
data score;
  input TRT01A $ SITEID $ SCORE;
  datalines;
ACT 101 56
ACT 102 34
ACT 101 21
ACT 103 67
ACT 102 54
PBO 102 90
PBO 101 53
PBO 103 78
PBO 102 89
PBO 102 79
;
run;

title "--- Mixed Model Using PROC MIXED ---";
proc mixed data = score;
  class TRT01A SITEID;
  model SCORE = TRT01A;
  random SITEID TRT01A*SITEID;
  lsmeans TRT01A / diff;
run;
```

Class Level Information

Class	Levels	Values
TRT01A	2	ACT PBO
SITEID	3	101 102 103

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
TRT01A	1	2	8.11	0.1043

The Type 3 test for treatment group effect is not significant at the 5% level.

INTEGO GROUP

Efficacy analysis of continuous endpoints

Assuming we have the following statement in Protocol or Statistical Analysis Plan (SAP):

"Continuous outcomes will be analyzed using a mixed model for repeated measures (MMRM). The estimates and confidence intervals (CIs) will be provided for the mean for each of the 3 treatment groups (X, Y, Z) and for the difference in means between pairwise comparisons of active comparator <Treatment Z> and each of the treatment groups <Treatment X>, < Treatment Y>. In addition, p-values for statistical tests will be provided. The efficacy analyses will be tested at the global-level at a significance level of 0.0496, and all CIs will be two-sided and at the 95.04% level."

Implementation Details:

Input data:

- ADSL (Subject Level Analysis Dataset),
- Analysis Dataset, e.g. ADOE (Ophthalmic Examinations).

Analysis Population: Intent-To-Treat

The following types of outputs need to be created:

- Composite Contrast over Weeks 48, 52, 56
- Over Time Through Week 56

INTEGO GROUP

Efficacy Analysis of Continuous endpoints, Composite Contrast over Weeks (48, 52, 56)

	Treatment X (N=xxx)	Treatment Y (N=xxx)	Treatment Z (N=xxx)
1 Year			
n	XXX	XXX	XXX
Adjusted Mean (SE)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)
95% CI for Adjusted Mean	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)
vs Treatment Z			
Difference in Adjusted Means (SE)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)	
95% CI for Difference in Adjusted Means	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)	
p-value	X.XXXX	X.XXXX	

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrast over Weeks 48, 52, 56)

Step 1: Data Selection

- Include all assessment results based on Endpoint specific information (e.g. “best corrected visual acuity” data, i.e. [ADOE.PARAMCD] = “SBCVA”) from Week 4 to Week 56 visits.

Step 2: Perform data modeling

- 3 different stratification factors: <strata 1>, <strata 2> and <strata-3>.
- PROC MIXED with significance level alpha=0.0496.
- type = UN (by default) → if any convergence problems → type = CS or AR (1).
- To find out the coefficients of L matrix the statement “ods output coef =” needs to be added for further constructing composite contrast.

Note: Set up of MMRM model is study specific and should be based on Protocol/SAP information.

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrasts over Weeks 18, 52, 56)

```
proc mixed data = <data selected at Step 1>;
  by paramcd param;
  class &sid. &trt. avisitn usubjid <strata factor variables>;
  model &dep. = &sid.
                &trt.
                avisitn
                &trt. * avisitn
                <strata factor variables>
                base /s ddfm=SAT;
  repeated avisitn / subject = usubjid(&trt.) type=&type.;
  lsmeans &trt. * avisitn / obsmargins e alpha=&alp.;
  ods output coef = _coef;
run;
```

Macro references:

&sid. – Study Identifier;

&trt. – Formatted planned treatment group;

&dep. – Dependent variable (e.g. change from baseline CHG variable);

&type. – Covariance structure type;

&alp. – Significance level.

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrasts over Weeks 48, 52, 56)

Sample of received **_coef** dataset (“treatment * analysis visit” interaction for <Treatment X>, <Treatment Z> only):

	&trt.	avisitn	X* (4...44)	X*48	X*52	X*56	Z* (4...44)	Z*48	Z*52	Z*56
Treatment X (4-56 visits)	⋮	⋮	0 ... 0	0	0	0	0 ... 0	0	0	0
	X	48	0 ... 0	1	0	0	0 ... 0	0	0	0
	X	52	0 ... 0	0	1	0	0 ... 0	0	0	0
	X	56	0 ... 0	0	0	1	0 ... 0	0	0	0
Treatment Z (4-56 visits)	⋮	⋮	0 ... 0	0	0	0	0 ... 0	0	0	0
	Z	48	0 ... 0	0	0	0	0 ... 0	1	0	0
	Z	52	0 ... 0	0	0	0	0 ... 0	0	1	0
	Z	56	0 ... 0	0	0	0	0 ... 0	0	0	1

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrast over Weeks 48, 52, 56)

Step 3: Processing with L matrix coefficients

Based on the "treatment * analysis visit" interaction table, the following calculation samples can be derived:

- combined coefficients for aggregated 1 year:

	&trt.	avisitn	X* (4...44)	X*48	X*52	X*56	Z* (4...44)	Z*48	Z*52	Z*56
	X	1 year (48,52,56)	0 ... 0	1	1	1	0 ... 0	0	0	0
	Z	1 year (48,52,56)	0 ... 0	0	0	0	0 ... 0	1	1	1

- difference for aggregated 1 year visit (Treatment X vs Treatment Z):

	avisitn	X* (4...44)	X*48	X*52	X*56	Z* (4...44)	Z*48	Z*52	Z*56
Difference X vs Z	1 year (48,52,56)	0 ... 0	1	1	1	0 ... 0	-1	-1	-1

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrasts over Weeks 18, 52, 56)

Step 4: Find out desired estimates for aggregated 1 year visit

a) Estimates for mean & CIs for each of the treatment groups.

b) Estimates for the difference in means & CIs between arms: X vs Z, Y vs Z + p-values.

To cover (a) point we can use several options.

Option 1:

```
lsmestimate &trt. * avisitn 'X' 0 … 0 1 1 1 0 … 0 0 0 0 / obsmargins divisor=3 cl alpha=&alp.;
```

Option 2:

```
estimate 'X' &trt. * avisitn 0 … 0 1 1 1 0 … 0 0 0 0 … <*> … / divisor=3 cl alpha=&alp.;
```

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrasts over Weeks 18, 52, 56)

<*> - we need to list each model effect with a specified vector of coefficients, which can be constructed as follows:

1) intercept 3
2) avisitn 0 … 0 1 1 1
3) &trt. 3 0
4) mean of baseline values among all data multiplied by 3
5) sfactor R1_num*3 R2_num*3 … Rn_num*3 where “Rn_num” corresponds to proportion of patients with “Rn” result among all data.

INTEGO GROUP

Efficacy analysis of continuous endpoints (Composite Contrast over Weeks 48, 52, 56)

To cover point (b) we can use the following options:

Option 1:

```
lsmestimate &trt. * avisitn 'X-Z' 0 … 0 1 1 1 0 … 0 -1 -1 -1 / obsmargins divisor=3 cl alpha=&alp.;
```

Option 2:

```
estimate 'X-Z' &trt. * avisitn 0 … 0 1 1 1 0 … 0 -1 -1 -1 &trt. 3 -3 / divisor=3 cl alpha=&alp.;
```

Option 3 (for p-values only):

```
contrast 'X-Z' &trt. * avisitn 0 … 0 1 1 1 0 … 0 -1 -1 -1 &trt. 3 -3;
```

INTEGO GROUP

Efficacy Analysis of Continuous Endpoints (Over Time Through Week xx)

	Treatment X (N=xxx)	Treatment Y (N=xxx)	Treatment Z (N=xxx)
Week xx			
n	XXX	XXX	XXX
Adjusted Mean (SE)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)
95% CI for Adjusted Mean	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)
vs Treatment Z			
Difference in Adjusted Means (SE)	XXX.XX (XX.XXX)	XXX.XX (XX.XXX)	
95% CI for Difference in Adjusted Means	(XXX.XX, XXX.XX)	(XXX.XX, XXX.XX)	
p-value	X.XXXX	X.XXXX	

INTEGO GROUP

Efficacy Analysis of Continuous Endpoints (Over Time Through Week XX)

```
proc mixed data = <data selected at Step 1>;
  by paramcd param;
  class &sid. &trt. (ref = "&ref.") avisitn usubjid <strata factor variables>;
  model &DEP = &sid.
               &trt.
               avisitn
               &trt. * avisitn
               <strata factor variables>
               base /s ddfm=SAT;
  repeated avisitn / subject = usubjid(&trt.) type=&type.;
  lsmeans &trt. * avisitn / obsmargins e alpha=&alp.;

  ods output lsmeans = lsm;
  ods output diffs = diff;
run;
```

&ref. – Reference treatment

To get individual estimates we are specifying: `lsmeans &trt. * avisitn / obsmargins e alpha=&alp.;`

For difference in means & CIs between treatment groups: X vs Z, Y vs Z + p-values we need to specify "ods output diffs =" statement.

INTEGO GROUP

Summary

- Identification of appropriate statistical procedure based on Protocol/SAP.
- PROC MIXED is a powerful procedure for construction of different mixed linear models.
- Study specific set up of MMRM model.
- Different ways to obtain the same statistics.

INTEGO GROUP

THANK YOU

Kateryna Fedoryshyna
Intego Group LLC
35-B Borychiv Tik Street
Kyiv 04070, Ukraine
Work Phone: +1 407.512.1006 (ext. 2447)
Email: kateryna.fedoryshyna@intego-group.com
Web: www.intego-group.com