MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models

Mihaela Ene, University of South Carolina, Columbia, SC

Whitney Smiley, University of South Carolina, Columbia, SC

Bethany A. Bell, University of South Carolina, Columbia, SC

ABSTRACT

When estimating multilevel models it is important for researchers to make sure their models fit their data. However, examining model fit can be quite cumbersome. We have developed the macro MIXED_FIT to help researchers assess model fit in a simple yet comprehensive way. Specifically, this paper provides a SAS macro that incorporates changes in model fit statistics [i.e., -2 log likelihood (-2LL), AIC, and BIC] as well as changes in pseudo-R^2. By utilizing data from PROC MIXED ODS tables, the macro produces a comprehensive table of changes in model fit measures and allows SAS users to examine model fit in both nested and non-nested models both in terms of statistical and practical significance without having to calculate these values by hand.

Keywords: MULTILEVEL MODELING, MODEL FIT, PROC MIXED, SAS MACRO

INTRODUCTION

MLMs account for dependent data by modeling data at different levels, such as children nested within schools or patients nested within hospitals. Because MLMs can account for dependencies in nested data, they are frequently used in social and health sciences where data are typically hierarchical in nature.

An important part of the multilevel analysis is assessing model fit. This evaluation consists of comparisons between different models considered in the study and allows researchers to select the best model in terms of its ability to explain the data. Considered a "crucial part of the multilevel modeling process" (McCoach & Black, 2008, p.246), model selection can be determined through several fit indices such as deviance (aka -2LL), AIC, and BIC. Whereas the differences in deviance can be used only when comparing nested models, the changes in AIC and BIC can be used to compare both nested and non-nested models (McCoach & Black, 2008). Considering that the efficacy of these statistics in the selection of the most parsimonious model depends on the sample size and the number of parameters included in the study, it is recommended that all three statistics be examined when analyzing nested models (McCoach & Black, 2008). In addition, the difference in pseudo-R^2 can be used to determine the proportional reduction of variance that can be attributed to more complex models, thus, serving as a measure of practical significance. The examination of these changes adds useful information to be considered in researchers' decisions of model selection.

Within SAS, PROC MIXED produces the -2LL, AIC, and BIC statistics, but does not provide any information about the differences in these fit statistics between two competing models. Furthermore, PROC MIXED does not provide the pseudo-R^2 statistic or changes in pseudo-R^2; these values have to be calculated by hand for each model examined. Bardenheier (2009) developed a macro that can be used with PROC MIXED in order to obtain the test statistic for the difference in -2LLvalues along with the *p*-value of the LRT. However, her macro only provided information useful in assessing model fit for nested models and does not address the changes in AIC or BIC which are the fit indices used when comparing non-nested models. In the absence of information related to these other model fit indices, novice researchers might incorrectly interpret the difference in -2LL values when assessing model fit for non-nested models.

In an effort to provide SAS users a simpler way of obtaining the changes in all fit indices and of making an informed decision of model selection for both nested and non-nested models, this paper provides a comprehensive SAS macro to assess model fit in two-level linear models estimated in PROC MIXED. This macro is an extension of the previous macro developed by Bardenheier (2009) that includes changes in AIC, BIC, and pseudo-R^2 in addition to the changes in deviance. We hope that this macro is a useful tool for researchers and SAS users to acquire model fit information without having to calculate the values by hand.

EXAMINING MODEL FIT

NESTED MODELS

When models are nested, that is if model 2 is an extension of model 1, then the likelihood ratio test (LRT) can be

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

used to assess model fit (Snijders & Bosker, 2012). The likelihood ratio test calculates the change in deviance (-2LL) between the smaller model and the larger model and statistically tests the change using a χ^2 distribution, with degrees of freedom equal to the change in number of parameters from the smaller model to the larger model.

Changes in deviance and the LRT can be used to nested test models that differ in fixed effects as well as models that differ in random effects. When two nested models differ in their random effects but not their fixed effects (i.e., a random intercept and slope model compared to a random intercept model), in order for the LRT to be valid, the models must be estimated using restricted maximum likelihood. Conversely, when two nested models differ in the number of fixed effects, but do not differ in their random effect, full maximum likelihood estimation must be used to ensure a valid LRT. In PROC MIXED, the default estimation method is restricted maximum likelihood. Therefore, if comparing the -2LL between nested models that differ in fixed effects, researchers need to be sure to specify "method = ML" in their PROC MIXED code.

Although the LRT cannot be directly tested in SAS, Bardenheier (2009) developed a macro that can be used with PROC MIXED to test for statistically significant change in model fit between two nested linear models. However, the macro does not provide any information related to the practical significance of the MLMs being compared.

In addition to examining if improvement in model fit is statistically significant, it is also important to examine changes in model fit in terms of practical significance. In MLM research, one effect size commonly used to examine practical significance is the Pseudo-R^2 statistic (Singer & Willett, 2003). This statistic calculates the amount of ICC that can be explained by the model; it is a function of the reduction of the intercept variance from the unconditional model compared to a larger model. The following equation denotes the general calculations used to calculate Pseudo-R^2 for any model beyond the unconditional model:

$$\frac{\tau_{00unconditional} - \tau_{00new\,model}}{\tau_{00unconditional}}$$

After calculating Pseudo-R^2 for each of the nested models, calculating the magnitude of the change in Pseudo-R^2 from a larger model to a smaller model can inform you how much explained variance is gained by the larger model. A detailed example of how changes in Pseudo-R^2 can be used to assess model fit is provided below in the illustrated example of the MIXED_FIT macro.

Because model selection is a fine balance between parsimony and model fit, examining changes in Pseudo-R^2 in conjunction with the LRT gives you not only a measure of statistical significance but also a measure of practical significance.

NON-NESTED MODELS

Whereas most model comparisons occur between nested models, there are situations in which researchers want to compare model fit between two non-nested models (i.e., situations where the models being compared are not extensions of one another). In these instances, the LRT cannot be used. Instead researchers must rely on other fit indices such as Akaike Information Criterion (AIC) and Schwarz's Bayesian Information Criterion (BIC). AIC can be expressed using the following equation:

$$AIC = 2k - 2\ln(L)$$

where k is equal to the number of parameters in the model and L is the maximized value of the likelihood function.

BIC can be expressed using this equation:

$$BIC = 2\ln(L) + k\ln(n)$$

where k is equal to the number of parameters in the model, L is the maximized value of the likelihood function, and n is equal to the sample size.

From these equations model fit for the AIC is quantified using the -2 Log Likelihood as well as the number of parameters used in the model, and BIC uses the same information as well as sample size. As a rule of thumb, the more attractive model should have smaller AIC and BIC (Kwok, Underhill, Berry, Luo, Elliott, & Yoon, 2008). Additionally, if we provide the change in AIC and BIC between the non-nested models, the researcher will have a better understanding of the difference in magnitude of model fit. Although the usual rule of thumb for examining changes in AIC and BIC is simply "smaller is better", some general recommendations on BIC change between models were provided by Raftery (as cited in McCoach & Black, 2008). For example, Raftery states that there is no substantial difference between two models if the BIC difference is less than 2 and there is a substantial difference

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

between two models if the difference between the two BIC values is greater than 10.

MACRO MIXED_FIT

The MIXED_FIT macro provides a comprehensive approach for comparing model fit for nested and non-nested two-level multilevel linear models. The inputs associated with the macro consist of ODS table names generated from PROC MIXED including (a) FitStatistics, (b) SolutionF, and (c) CovParms. The MIXED_FIT macro produces numeric output consisting of a table that includes the difference in -2 log likelihood values along with the *p*-value of the LRT, the difference in the AIC and BIC, as well as the change in pseudo-R^2.

Important information about the MIXED_FIT macro is listed below:

- MIXED_FIT has been created and tested for use in SAS v9.2.

- MIXED_FIT is designed for two-level linear models estimated in PROC MIXED.

- In order to provide the changes in all fit indices, MIXED_FIT uses three models: unconditional, reduced, and full model.

- Users should set the parameters for each model in the ODS OUTPUT statement. These parameters will be used when invoking the macro to compare two models.

- Users need to invoke the macro for each of the two models that are being compared.

The most recent version of the MIXED_FIT macro is available for download from http://www.ed.sc.edu/bell/.

EXAMPLE OF MACRO MIXED_FIT

The use of the MIXED_FIT macro to test changes in nested models that differ in their fixed effects is demonstrated below. In this example, the hierarchical data structure consists of students nested within schools. Using student data (level-1) and school data (level-2), two-level nested models are used to investigate the relationship between math achievement and various characteristics at the student and school level. As is commonly done in the social and behavior sciences, three models are estimated: an unconditional model, a model that contains only level-1 predictors, and a model that contains both level-1 and level-2 predictors. Moreover, in this example, a researcher would be interested in examining if the model that contains level-1 and level-2 predictors is a better fitting model than the model that contains only level-1 predictors.

As shown below, the three models are estimated using PROC MIXED, followed by the call to the MIXED_FIT macro.

```
Title 'Model 1: Unconditional Model';
PROC MIXED data = HSB covtest noclprint method = ml empirical;
class SCHOOLID;
model MATHACH = /solution ddfm = SATTERTHWAITE s;
random intercept / sub=SCHOOLID s;
ods output Fitstatistics=FS_Model_1 SolutionF=SF_Model_1;
ods output CovParms=CovModel_1; run;

Title 'Model 2: Level-1 Model Random Intercept Only';
PROC MIXED data = HSB covtest noclprint method = ml empirical;
class SCHOOLID;
model MATHACH = FEMALE MINORITY SES/solution ddfm = SATTERTHWAITE s;
random intercept / sub=SCHOOLID s;
ods output Fitstatistics=FS_Model_2 SolutionF=SF_Model_2;
ods output CovParms=CovModel_2; run;

Title 'Model 3: Level-1 and Level-2 Model Random Intercept Only';
PROC MIXED data = HSB covtest noclprint method = ml empirical;
class SCHOOLID;
model MATHACH = FEMALE MINORITY SES HIMINTY MEANSES/solution ddfm = SATTERTHWAITE s;
random intercept / sub=SCHOOLID s;
ods output Fitstatistics=FS_Model_3 SolutionF=SF_Model_3;
ods output CovParms=CovModel_3; run;
```

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

```
%Mixed_Fit(fullmodel=FS_Model_3,redmodel=FS_Model_2,DFfull=SF_Model_3,DFred=SF_Model_2
, CovarianceUnconditionalModel=CovModel_1, CovarianceReducedModel=CovModel_2,
CovarianceFullModel=CovModel_3); run;
```

Output 1 presents the SAS output that will be obtained when using this example of the PROC MIXED code and MIXED_FIT macro to compare the level-1 model (Model 2) with the level-2 model (Model 3).

```
             Model Fit Statistics for changes in Fixed Effects



        Change in AIC        Change in BIC        Change in -2LL
         Reduced-Full         Reduced-Full         Reduced-Full      p-value LRT
        Useful with non-     Useful with non-     For nested models    For nested
Obs       nested models        nested models           only          models only
 1          57.193               51.043               61.1933            0.0000



                          Change in Pseudo-R2

                              Full-Reduced
                               For nested
                              models only
                       Obs

                        1        0.15951
```
Output 1. Output from MIXED_FIT Summary Table

Output 1 contains changes in all three model fit statistics (-2LL, AIC and BIC), as well as the p-value for the LRT and the change in pseudo-R^2. This statistical output can be used to compare the two models used in this example in order to determine whether or not the more complex model with predictors at both levels is a better fit for the data. As shown in Output 1, the model with predictors at level-1 (student) and level-2 (school) appears to fit the data better than the model with just level-1 predictors. Based on the changes in the -2LL, the improved fit of the full model is significantly better than the fit of the smaller model (p<.01). Moreover, in this example, the full model also provided a meaningful improvement in model fit as evidenced by the large change in pseudo-R^2 values.

A similar combination of the PROC MIXED code and MIXED_FIT macro should be used when comparing two-level non-nested linear models. The obtained output will be similar with the one produced for the nested models example; however, the user should only use the changes in AIC and BIC when comparing the model fit for the two non-nested models in order to select the best model in terms of its ability to explain the data.

The complete macro code that was invoked to for this example is provided in the Appendix.

CONCLUSION

The MIXED_FIT macro is a tool created to facilitate the process of examining model fit for two-level linear models by providing researchers and SAS users with a simpler way of obtaining the changes in all fit indices necessary for comparison of both nested and non-nested two-level models. Indices provided in the SAS output include changes in the -2LL and its corresponding LRT along with its p-value, as well as changes in AIC, BIC, and pseudo-R^2. The accessibility of this information helps the researcher to examine model fit in terms of both statistical and practical significance and select the model that represents the best fit for the data considering all appropriate indices for the models that are being considered. In addition, the fairly straightforward way of invoking the macro makes MIXED_FIT a simple and valuable tool to have when conducting multilevel analysis. Plus, utilizing the macro when estimating models via PROC MIXED prevents users from having to calculate changes in model fit by hand.

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

APPENDIX

```
%macro Mixed_Fit(fullmodel=, redmodel=, DFfull=, DFred=,
CovarianceUnconditionalModel=, CovarianceReducedModel=, CovarianceFullModel=);

/*Obtaining the AIC and BIC for the full model*/
proc transpose data=&fullmodel out=fullmodel; run;
proc print;run;

proc sql;
create table fullfit as select COL2 as AICfull, COL4 as BICfull from fullmodel;
quit;

/*Original syntax used to get Chi squared for the full model. Added the (obs=1) to
make merging easier later*/
data fullmodel4; set &fullmodel (obs=1);
if descr='-2 Res Log Likelihood' then chisqfullRML=value; *RML;
else if descr='-2 Log Likelihood' then chisqfullFML=value; *FML;
run;

/*Obtaining the AIC and BIC outside of the restricted model*/
proc transpose data=&redmodel out=redmodel; run;
proc sql;
create table redfit as
select COL2 as AICred, COL4 as BICred from redmodel;
quit;

/*Original syntax used to get Chi squared for the reduced model. Added the (obs=1) to
make merging easier later*/
data redmodel4; set &redmodel (obs=1);
if descr='-2 Res Log Likelihood' then chisqredRML=value; *RML;
else if descr='-2 Log Likelihood' then chisqredFML=value; *FML;
run;

/*Original syntax used. Added the (obs=1) to help with merging later*/
proc sql;
CREATE TABLE flDF AS SELECT effect, count(*) AS fullDF from
work.&DFfull; quit; *Count number of effects to get degrees of freedom;
proc sql;
create table rdDF as select effect, count(*) as redDF from
work.&DFred; quit;
data degfree (obs=1);
merge work.flDF work.rdDF;
if _n_=1 then LRTDF= fullDF - redDF; run;
proc print data=degfree; run;

/*Added this to get the change in AIC and BIC*/
data fit;
merge fullfit redfit;
if _n_=1 then AICchg= AICred-AICfull;
if _n_=1 then BICchg= BICred-BICfull;
run;

/*Updated this to help merge the fit data I just computed to the rest of the table*/
data likelihood;
merge fit fullmodel4 redmodel4 degfree;
testintRML=abs((chisqredRML)-(chisqfullRML)); **Models can yield
negative LLs, those that are smaller in absolute value -ie, closer to 0
fit better (pg 116-117, Singer);
*7;
testintFML=abs((chisqredFML)-(chisqfullFML)); **Models can yield
negative LLs, those that are smaller in absolute value -ie, closer to 0
fit better (pg 116-117, Singer);
```

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

```sas
pvaluemixture=(( .5*(1-probchi(testintRML,2)) + .5*(1-
probchi(testintRML,1)) )); *for random terms;
pvalueLRT=(1-probchi(testintFML,LRTDF)); *For fixed terms;
run;

/*Creating the table that will be displayed. Added the AIC and BIC information, a
description of it and formatted it.*/
proc print data=likelihood (obs=1) split='*';
var AICchg BICchg testintFML pvalueLRT;
format pvalueLRT 6.4;
format AICchg 7.3;
format BICchg 7.3;
label AICchg= 'Change in AIC*Reduced-Full*Useful with non-nested*models'
BICchg= 'Change in BIC*Reduced-Full*Useful with non-nested*models'
testintFML='Change in -2LL*
Reduced-Full*For nested models*only'
pvalueLRT='p-value LRT*For nested models*only';
title "Model Fit Statistics for changes in Fixed Effects"; run;
proc print data=likelihood split='*' noobs;
var AICchg BICchg testintRML pvaluemixture;
format pvaluemixture 6.4;
format AICchg 7.3;
format BICchg 7.3;
where testintRML ne . ;
label AICchg= 'Change in AIC*Reduced-Full*Useful with non-nested*models'
BICchg= 'Change in BIC*Reduced-Full*Useful with non-nested*models'
testintRML='Change in -2LL*
Reduced-Full*For nested models*only'
pvaluemixture='mixture method p-value*For nested models*only';
title "Mixture method Test for random effects";
run;

/*Extract tau00 from the Covariance Parameter Estimates table of the unconditional
model*/
proc sql;
create table tauum as select CovParm, Sum(Estimate)as Estimateum from
&CovarianceUnconditionalModel where CovParm='Intercept' group by CovParm;
quit;

/*Extract tau00 from the Covariance Parameter Estimates table of the reduced model*/
proc sql;
create table taurm as select CovParm, Sum(Estimate)as Estimaterm from
&CovarianceReducedModel where CovParm='Intercept' group by CovParm;
quit;

/*Extract tau00 from the Covariance Parameter Estimates table of the full model*/
%if &CovarianceFullModel ne %then
%do;
proc sql;
create table taufm as select CovParm, Sum(Estimate)as Estimatefm from
&CovarianceFullModel where CovParm='Intercept' group by CovParm;
quit;
%end;

/*Summary table with tau00 from the unconditional and reduced model*/
data tauSummary;
Merge tauum taurm;
by CovParm;
run;

/*Calculate Pseudo-R2 for the unconditional model*/
proc sql;
```

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

```
create table pseudoR2um as select (Estimateum-Estimateum)/Estimateum as
PseudoR2UnconditionalModel from tauSummary;
quit;

/*Calculate Pseudo-R2 for the reduced model*/
proc sql;
create table pseudoR2rm as select (Estimateum-Estimaterm)/Estimateum as
PseudoR2ReducedModel from tauSummary;
quit;

/*Summary table with Pseudo-R2 for the unconditional and reduced model*/
data pseudoR2Summary;
Merge pseudoR2um pseudoR2rm;
run;

/*Calculate changes in Pseudo-R2*/
%if &CovarianceFullModel eq %then
%do;
proc sql;
create table ChangePseudoR2 as select PseudoR2ReducedModel-PseudoR2UnconditionalModel
as ChangePseudoR2 from pseudoR2Summary;
quit;
proc print data=ChangePseudoR2;
run;
%end;

/*Summary table with tau00 from the unconditional and full model*/
%if &CovarianceFullModel ne %then
%do;
data tauSummary;
Merge tauum taufm;
by CovParm;
run;

/*Calculate Pseudo-R2 for the unconditional model*/
proc sql;
create table pseudoR2um as select (Estimateum-Estimateum)/Estimateum as
PseudoR2UnconditionalModel from tauSummary;
quit;

/*Calculate Pseudo-R2 for the full model*/
proc sql;
create table pseudoR2fm as select (Estimateum-Estimatefm)/Estimateum as
PseudoR2FullModel from tauSummary;
quit;

/*Summary table with Pseudo-R2 for the reduced and full model*/
data pseudoR2Summary;
Merge pseudoR2rm pseudoR2fm;
run;

/*Calculate changes in Pseudo-R2*/
proc sql;
create table ChangePseudoR2 as select PseudoR2FullModel-PseudoR2ReducedModel as
ChangePseudoR2 from pseudoR2Summary;
quit;
proc print data=ChangePseudoR2 split='*';
title 'Change in Pseudo-R2';
label ChangePseudoR2 = 'Full-Reduced*For nested models*only';
run;
%end;
%mend Mixed_Fit;
```

7

MIXED_FIT: A SAS® Macro to Assess Model Fit and Adequacy for Two-Level Linear Models, continued

REFERENCES

Bardenheier, B. H. (2009). Proc Mixed: Macro to assess fixed and random effects for significance using the Likelihood Ratio test and the approximate Mixture Method. *SESUG 2009: The Proceedings of the SouthEast SAS Users Group.*

Kwok, O., Underhill, A. T., Berry, J. W., Luo, W., Elliott, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models: An example with individuals living with lower extremity intra-articular fractures. *Rehabilitative Psychology, 53 (3)*, 370-386.

McCoach, D. B. & Black, A. C. (2008). Evaluation of model fit and adequacy. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (245-272). Charlotte, NC: Information Age Publishing, Inc.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence* (1st ed.). New York, NY: Oxford University Press, Inc.

Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Bethany A. Bell
University of South Carolina
College of Education, Wardlaw 133
820 Main Street
Columbia, SC 29208
Work Phone: 803-777-2387
E-mail: babell@sc.edu
Web: http://www.ed.sc.edu/bell/