

Paper SC-013

PROC MIXED: Macro to assess fixed and random effects for significance using the Likelihood Ratio test and the approximate Mixture Method

B. H. Bardenheier, Centers for Disease Control & Prevention and Emory University, Atlanta, GA

ABSTRACT

A macro written for use with PROC MIXED to obtain the test statistic for the difference in deviances between two nested models along with the p-values for the Likelihood Ratio Test (LRT) and the approximate Mixture Method will be demonstrated. Some researchers believe these tests may have better statistical properties than the approximate Wald test¹ which is given in the SAS[®] output. Further, for random terms, SAS[®] gives the two-tailed *p*-value of the Wald test; it must be halved to get the approximate *p*-value for a one-tailed test (the random term is squared and therefore can never be negative, hence it is one-sided). If users are not aware of this they may fail to reject the hypothesis of no association when they should (ie, there may be significant variation between groups but if the *p*-value of the random term is not halved there may be type II error). It is also possible to get a type I error using the Wald test. Use of the macro for assessing random effects and fixed effects will be explained with data for a multilevel model. When and why maximum likelihood and restricted maximum likelihood methods should be used will be discussed. Mathematical models, null hypotheses, and SAS[®] code will be shown simultaneously for best understanding. This macro is a tool to give the user a simpler way to get the best significance testing for both fixed and random effects when fitting mixed linear models.

INTRODUCTION

A macro written for use with PROC MIXED to obtain the test statistic for the difference in deviances between two nested models along with the p-values for the Likelihood Ratio Test (LRT) and the approximate Mixture Method will be demonstrated. Some researchers believe these tests may have better statistical properties than the approximate Wald test¹ which is given in the SAS[®] output. For many statisticians, the LRT is preferable to the Wald test for several reasons: 1) the Wald test uses two approximations (the standard error is known, and the distribution is chi-squared), whereas the likelihood ratio test uses one approximation (that the distribution is chi-squared); 2) the Wald test has been found to have poor power;¹ and 3) the Wald test can give different answers to the same question, according to how the question is asked, i.e., it depends on the parameter used.² Further, when assessing random terms, the p-value obtained from the approximate mixture method is preferable to the p-value obtained from both the Wald test and also the LRT in which the degrees of freedom are equal to *n* (number of random terms tested), because the test statistic is distributed more like 2 chi-squared distributions with *n* and *n*+1 degrees of freedom than 1 chi-squared distribution, where *n* is the number of random terms tested.³ Thus the test statistic is the same for both the LRT and the approximate mixture method, but the degrees of freedom used in determining the p-values differ.

First, I will give background information on the data used, then describe use of the macro for assessing random effects and then for assessing fixed effects. When and why maximum likelihood and restricted maximum likelihood methods should be used will be discussed. Mathematical models, null hypotheses, and SAS[®] code will be shown simultaneously for best understanding.

Data

Data for 5,930 nursing home facilities with ≥ 10 African-American residents in 41 states and the District of Columbia are analyzed. The type of mixed model being demonstrated is a multilevel model. Nursing facility data are level-1 and state is level-2. The dependent variable is racial disparity in vaccination coverage. The variable, DIFFCOV, was calculated by taking the difference: vaccination coverage for all eligible white nursing home residents in each nursing home – vaccination coverage for all eligible African-American nursing home residents in each nursing home.

Examining estimated variance components

To assess the significance of a random term, the full model with the random term is run as well as the reduced model without the random term. ODS statements should be included with those models as well as the macro call (see code below). It is important to keep fixed effects constant in both the full and reduced models. Only one random effect should be tested at a time when using this macro. Restricted maximum likelihood should be used because that maximizes the likelihood of the sample residuals. The reason for this is that random terms are variables, like the error term, with a multivariate normal distribution with a mean of zero and variance of sigma-squared. The null hypothesis is the sigma-squared equals zero. Therefore the residuals should be maximized to test this by using the option, method=REML on the PROC MIXED statement.

1) Testing Random intercept

The model and null hypothesis are:

$$\text{DIFFCOV}_{ij} = \gamma_{00} + u_{0j} + \varepsilon_{ij}$$

where i =nursing home and j =state, $u_{0j} \sim \text{MVN}(0, \sigma^2_{00})$, $\varepsilon_{ij} \sim \text{MVN}(0, \sigma^2_e)$,

$$H_0: \sigma^2_{00} = 0$$

The entire code below should be run at the same time.

```
%include '\\cdc\private\mixture method pvalue macro1.sas';
proc mixed data=allUS2levge10blk method=reml covtest empirical
noclprint ;
class state_id;
model diffcov=/ ddfm=contain s;
random intercept /subject=state_id s ;
ods output FitStatistics=fm SolutionF=SFfm ;
run;

proc mixed data=allUS2levge10blk method=reml noclprint;
class state_id;
model diffcov= /ddfm=contain s;
```

```
ods output FitStatistics=rm SolutionF=SFrM ;
run;
%LRMixtureTest (fullmodel=fm, redmodel=rm, DFfull=SFfm, DFred=SFrM) ;
```

When running a mixed model in SAS[®], the covariance parameters for the random terms are estimated, but the COVTEST option must be used to have the results displayed. The results include the estimate, standard error, Z value, and *p*-value. Results for the values corresponding to the **G** matrix (random terms or between group variance) are given first, then the residual value which corresponds to the variance estimate in the **R** matrix (error term or within group variance).

Covariance Parameter Estimates					
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
Intercept	state_id	1.9168	0.7297	2.63	0.0043
Residual		129.61	2.3887	54.26	<.0001

Using the SAS[®] output we should assess the approximate Wald test for the random intercept with caution. The *p*-value is 0.0043 but should be divided by 2 (0.0043/2=0.0022) because SAS[®] gives the two-tailed *p*-value for the Wald test; it must be halved to get the approximate *p*-value for a one-tailed test (the parameter being tested is squared and therefore can never be negative, hence it is one-sided). If users are not aware of this they may fail to reject the hypothesis of no association when they should (ie, there may be significant variation between groups but if the *p*-value of the random term is not halved there may be type II error). It is also possible to get a type I error using the Wald test. The mixture method *p*-value is best for assessing random terms.⁴ This *p*-value is preferred by some statisticians because the test statistic is distributed more like 2 chi-squared distributions with *n* and *n*+1 degrees of freedom than 1 chi-squared distribution, where *n* is the number of random terms tested. The macro output is below; the actual macro being invoked is in appendix A.

likelihood ratio test statistic comparing reduced model to full model	p-value mixture test
49.8116	0.0000

For the data in this example, both methods found the random intercept to be significant. Note that although there are no random terms in the reduced model, the containment method (the default denominator degrees of freedom (ddfm) when the random statement is used) is used in the reduced model to be consistent with the full model.

2) Testing fixed effect

This macro does not limit the number of fixed effects being tested. However, the reduced model (ie, model without the effects being tested) should be nested within the full model. Random effects should be held constant when testing the fixed effects. Next we assess

the proportion African-Americans in the nursing home as a fixed effect. The method option on the PROC MIXED statement should be maximum likelihood (method=ML) because that maximizes the likelihood of observing all the sample data actually observed.⁵ This is appropriate because the null hypothesis is that the parameter of the fixed effect is equal to zero. This test is based on estimation of the parameter using the sample data actually observed.

The proportion African-Americans in the nursing home was added to both models as a random effect. Now we test this variable as a fixed effect. The model and null hypothesis are:

$$\text{DIFFCOV}_{ij} = \gamma_{00} + \gamma_{10} * \text{PCTBLK}_{ij} + u_{0j} + u_{1j} * \text{PCTBLK}_{ij} + \varepsilon_{ij}$$

where i =nursing home and j =state, $u_{0j} \sim \text{MVN}(0, \sigma^2_{00})$, $u_{1j} \sim \text{MVN}(0, \sigma^2_{11})$, $\varepsilon_{ij} \sim \text{MVN}(0, \sigma^2_e)$

$H_0: \gamma_{10}=0$

```
%include '\\cdc\private\mixture method pvalue macro1.sas';
proc mixed data=bhb7.allUS2levge10blk method=ml covtest empirical
noclprint ;
class state_id;
model diffcov=pctblk/ ddfm=contain s;
random intercept pctblk /subject=state_id type=ar(1) s ;
ods output FitStatistics=fm SolutionF=SFfm ;
run;

proc mixed data=bhb7.allUS2levge10blk method=ml covtest empirical
noclprint;
class state_id;
model diffcov=/ddfm=contain s;
random intercept pctblk /subject=state_id type=ar(1) s;
ods output FitStatistics=rm SolutionF=SFrm ;
run;
%LRMixtureTest (fullmodel=fm, redmodel=rm, DFfull=SFfm, DFred=SFrm) ;
```

When assessing the significance of fixed effects variables the SAS[®] output gives values for the t-test and for the F-test.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	3.6333	0.4672	41	7.78	<.0001
pctblk	-5.7544	1.2451	40	-4.62	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
pctblk	1	40	21.36	<.0001

Statisticians prefer the LRT to the Wald test for fixed effects as well.⁵ The LRT p-value is given in the macro output.

Likelihood Ratio Test for fixed effects

likelihood ratio test statistic comparing reduced model to full model	p-value LRT
38.1589	0.0000

In this case, the fixed effect, proportion African-Americans in the nursing home, was found significant using both methods.

CONCLUSION

We present use of a macro to obtain the LRT for examining fixed effects and the approximate Mixture Method for examining random effects using methods not easily obtainable from the output in SAS®. This macro is a tool to give the user a simpler way to get the best significance testing for both fixed and random effects when fitting mixed linear models.

Acknowledgements

I would like to thank John Stevenson, MA, for help with writing the macro, Steve Benoit, MD, MPH, for help with programming, and David G. Kleinbaum, PhD, for encouraging me to write the macro for his longitudinal class.

Contact Information:

Barbara Bardenheier, MPH, MA
Centers for Disease Control and Prevention
1600 Clifton Rd, MS E-52
Atlanta, GA 30333
Tel : (404) 639-8789
Fax : (404) 639-8614
Email : bfb7@cdc.gov

APPENDIX A

```
%macro LRMixtureTest(fullmodel=,redmodel=,DFfull=,DFred=);
/*
This macro calculates the Likelihood Ratio Test for fixed effects in a
mixed model
See Singer Applied Longitudinal Data Analysis pp 116-120.
and also calculates the mixture method pvalue for random effects in a
mixed model -see KKNM chapter 26.
The macro argument &fullmodel refers to the output dataset produced by
running the full model in proc mixed and using the statement:
    ods output FitStatistics=fm;
The macro argument &DFfull refers to the output dataset produced by
running the full model in proc mixed and adding to the statement:
    ods output FitStatistics=fm SolutionF=SFfm; The number of
parameters being tested (ie, fixed effects) determine the degrees of
freedom in the LRT and come from this dataset.
The argument &redmodel refers to the output dataset produced by running
the reduced model in proc mixed using the statement:
    ods output FitStatistics=rm SolutionF=SFrm ;
Maximum Likelihood should be used when comparing fixed effects because
it maximizes the likelihood of the sample data. Using the Restricted
Maximum likelihood should only be used when assessing a random term
(ie, the fixed effects should be the same in both models)
because in RML we maximize the likelihood of the residuals. The degrees
of freedom would not be correct if the number of fixed effects were
different in each model when using RML -which is the default in PROC
MIXED.
*/
data &fullmodel; set &fullmodel;
    if descr='-2 Res Log Likelihood' then chisqfullRML=value; *RML;
    else if descr='-2 Log Likelihood' then chisqfullFML=value; *FML;
run;
data &redmodel; set &redmodel;
    if descr='-2 Res Log Likelihood' then chisqredRML=value; *RML;
    else if descr='-2 Log Likelihood' then chisqredFML=value; *FML;
run;
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 (drop=effect);
    merge work.flDF work.rdDF;
if _n_=1 then LRTDF= fullDF - redDF; run;

data likelihood;
    merge &fullmodel &redmodel degfree;
    testinRML=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);
```

```

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

proc print data=likelihood split='*' noobs;
    var testintFML pvalueLRT;
    format pvalueLRT 6.4;
    where testintFML ne .;
    label testintFML='likelihood ratio*test statistic*
        comparing*reduced model to full model'
        pvalueLRT='p-value LRT';
    title "Likelihood Ratio Test for fixed effects"; run;

proc print data=likelihood split='*' noobs;
    var testintrML pvaluemixture;
    format pvaluemixture 6.4;
    where testintrML ne .;
    label testintrML='likelihood ratio*test statistic*
        comparing*reduced model to full model'
        pvaluemixture='mixture method p-value';
    title "Mixture method Test for random effects";
run;
%mend LRMixtureTest;

```

Reference List

1. Hauck WW, Donner A. Wald's Test as Applied to Hypotheses in Logit Analysis. *Journal of the American Statistical Association*. 1977; 72:851-853.
2. Fears T, Benichou J, Gail MH. A Reminder of the Fallibility of the Wald Statistic. *The American Statistician*. 1996; 50:226-227.
3. Verbeke G, Molenberghs G. A Review on Linear Mixed Models for Longitudinal Data, Possibly Subject to Dropout. *Statistical Modelling*. 2001; 1:235-269.
4. Kleinbaum DG, Kupper LL, Nizam A, Muller KE. Analysis of correlated data Part 2: Random Effects and other issues. *Applied Regression Analysis and Multivariable Methods*. New York, NY: Duxbury, 2007.
5. Singer JD, Willett JB. *Applied Longitudinal Data Analysis: Modeling change and event occurrence*. 1st ed. New York, NY: Oxford University Press, Inc., 2003.

SAS and all other SAS Institute Inc. product or services names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.