

Using Higher Order Asymptotics to Improve P-values and Confidence Intervals Given by PROC MIXED

Ben Lyons, Oregon State University
Dawn Peters, Oregon State University

Abstract

P-values and confidence intervals given by the MIXED procedure are often based on asymptotic properties and may not be accurate for small sample sizes. This paper describes a simple adjustment to the signed square root of the log-likelihood ratio statistic yielding a test that can be very accurate for small samples. This adjustment, based on recent results in higher order asymptotics, is readily obtained in SAS[®] with PROC MIXED and SAS/IML[®].

Introduction

SAS/STAT[®] offers several methods for obtaining p-values and confidence intervals in mixed linear models. Standard asymptotic tests provided by the MIXED procedure include the Wald test and the likelihood ratio test (LRT). SAS also offers specialized approximate and exact tests for some mixed linear models. For testing certain hypotheses in variance component models, for example, the GLM procedure gives exact p-values when the data is balanced and p-values based on the Satterthwaite approximation when the data is unbalanced.

The likelihood based tests are appealing because they offer a general method for obtaining p-values and confidence intervals. However in some settings, both the Wald test and the LRT can have poor small sample properties when compared with specialized approximate tests. For a comparison of the Wald test and the specialized Hotelling-Lawley-Pillai-Samson trace statistic in repeated measures models see Wright (1995).

Skovgaard (1996) introduced a test statistic for one dimensional hypotheses, called \tilde{r} , that can give very accurate p-values for moderate sample sizes. The \tilde{r} statistic is often more accurate than specialized approximate tests and, in addition, it can be applied

in settings where specialized tests have not been developed. The \tilde{r} statistic has a compact closed form, does not require numerically intensive calculations, and is straight forward to implement in SAS with PROC MIXED and PROC IML.

In the following sections we describe procedures for the general mixed linear model with Y a multivariate normal random vector having mean $X\beta$ and covariance matrix $V(\rho)$, where β and ρ are vectors of unknown parameters of dimension p and k respectively. We will assume the matrix V is positive definite for all values of ρ and that the true value of ρ is in the interior of the parameter space.

Asymptotic Procedures

The asymptotic procedures discussed here are modifications of the directed log-likelihood statistic, that is the signed square root of the LRT. To define the directed log-likelihood statistic, r , rewrite $\theta = (\beta, \rho)$ in terms of ψ , a scalar parameter of interest, and the nuisance parameter vector λ . For the hypothesis $\psi = \psi_o$, we will write the full and restricted MLEs of θ as $\hat{\theta} = (\hat{\psi}, \hat{\lambda})$ and $\tilde{\theta} = (\psi_o, \tilde{\lambda})$ respectively. The restricted nuisance parameter estimate $\tilde{\lambda}$ maximizes the log-likelihood $l(\theta)$ when $\psi = \psi_o$. The directed log-likelihood statistic, or directed deviance, is then

$$r(\psi_o) = \text{sign}(\hat{\psi} - \psi_o) (2(l(\hat{\theta}) - l(\tilde{\theta})))^{\frac{1}{2}}.$$

Under the null hypothesis r is approximately standard normal with the error of the approximation being $O_p(n^{-\frac{1}{2}})$ or first order (See Barndorff-Nielsen and Cox, 1994). One can calculate r in SAS by making two calls to PROC MIXED to find the LRT. For more details on calculating the LRT with the MIXED procedure see SAS Institute Inc. (1995).

The statistic of interest here, \tilde{r} , is an approximation to Barndorff-Nielsen's (1986) r^* , a highly accurate

statistic but one that is often difficult to obtain. This statistic is a modification of r having the form

$$r^* = r - \frac{1}{r} \log\left(\frac{r}{u}\right),$$

where u is a complicated expression depending on often intractable sample space derivatives (see Barndorff-Nielsen and Cox, 1994). Since r^* is accurate to third order it will frequently provide substantially more accurate results than does r ; for large samples the adjustment becomes negligible with r and r^* being nearly equal.

The difficulty involved in the calculation of u led Skovgaard to an approximation \tilde{u} , described below, and the associated statistic

$$\tilde{r} = r - \frac{1}{r} \log\left(\frac{r}{\tilde{u}}\right),$$

which is generally standard normal to second order when $\psi = \psi_0$. For mixed models with block diagonal covariance matrices, such as repeated measures models, asymptotic results suggest that \tilde{r} will perform well as the number of subjects increases. In mixed linear models \tilde{u} is straight forward to obtain and in some models, e.g. balanced variance component models, $\tilde{r} = r^*$.

The improvement in the order of the error does not fully capture the benefits of using r^* or \tilde{r} instead of r . The accuracy of r decreases as models become increasingly complex with many nuisance parameters to estimate. When testing a covariance parameter in mixed linear models, r^* and \tilde{r} , like REML, correct for the "nuisance" fixed effects. Unlike REML, however, they also correct for the "nuisance" covariance parameters. Similarly, when testing fixed effects, where there is no equivalent to REML, r^* and \tilde{r} adjust for the nuisance fixed and covariance parameters. See Pierce and Peters (1992) for information regarding the performance of r and r^* in the presence of nuisance parameters.

Calculation of \tilde{r} for Mixed Linear Models

Although \tilde{r} is defined for testing any scalar function of (β, ρ) the calculations presented only apply to tests of individual coordinates of β or ρ . The calculation of \tilde{u} requires:

1) Full and Restricted MLEs:

$$(\hat{\beta}, \hat{\rho}) \text{ and } (\tilde{\beta}, \tilde{\rho}),$$

2) Observed Information:

$$j(\theta) = - \left[\frac{\partial^2 l(\theta)}{\partial \theta_j \partial \theta_i} \right]_{ij},$$

3) Expected Information:

$$i(\theta) = E(j(\theta)),$$

and

4) Derivatives of the covariance matrix V :

$$V_{\rho}^{-1} = \frac{\partial}{\partial \rho} V^{-1}(\rho).$$

The expression for \tilde{u} is then

$$\tilde{u} = [S^{-1}q]_{\psi} |j_{\lambda\lambda}(\tilde{\theta})|^{-\frac{1}{2}} |S| |j(\tilde{\theta})|^{\frac{1}{2}} |i(\tilde{\theta})|^{-1},$$

where S and q are described below and $j_{\lambda\lambda}(\tilde{\theta})$ is the portion of the observed information matrix associated with the nuisance parameters. The $(p+k) \times (p+k)$ matrix S , and the $(p+k)$ dimensional vector q depend on the full and restricted MLEs and the derivative matrix V_{ρ}^{-1} . The quantity $[S^{-1}q]_{\psi}$ is the component of the product of S^{-1} and q associated with the parameter of interest. The components of

$$S = \begin{bmatrix} S_{\beta\beta} & S_{\beta\rho} \\ S_{\rho\beta} & S_{\rho\rho} \end{bmatrix} \text{ and } q = \begin{pmatrix} q_{\beta} \\ q_{\rho} \end{pmatrix}$$

needed to obtain \tilde{r} have the following simple form in mixed linear models (see Lyons and Peters, 1996):

$$S_{\beta\beta} = X'V^{-1}(\tilde{\rho})X,$$

$$S_{\rho\beta} = 0,$$

$$S_{\beta\rho_j} = X'V_{\rho_j}^{-1}(\tilde{\rho})X(\tilde{\beta} - \hat{\beta})$$

for $j = 1$ to k ,

$$[S_{\rho\rho}]_{ij} = \frac{1}{2} \text{tr}(V(\hat{\beta})V_{\rho_i}^{-1}(\hat{\beta})V(\hat{\beta})V_{\rho_j}^{-1}(\tilde{\rho}))$$

for $i, j = 1$ to k ,

$$q_{\beta} = X'\tilde{V}^{-1}X(\hat{\beta} - \tilde{\beta}) \text{ and}$$

$$q_{\rho_j} = \frac{1}{2} \text{tr}(V(\hat{\beta})V_{\rho_j}^{-1}(\hat{\beta})V(\hat{\beta})(V^{-1}(\hat{\beta}) - V^{-1}(\tilde{\rho})))$$

$j = 1$ to k .

One can calculate \tilde{r} in SAS by using PROC MIXED to obtain the full and restricted MLEs and then passing these estimates to PROC IML to carry out the necessary matrix calculations. The following example demonstrates the accuracy of \tilde{r} in a simple mixed linear model. A program for calculating \tilde{r} in this model is given in the appendix.

Example: One-Way Balanced Variance Component Model

Consider the one-way balanced variance component model:

$$Y_{ij} = \mu + b_i + \varepsilon_{ij}$$

for $i = 1, \dots, s$, and $j = 1, \dots, t$,

where b and ε are independent random vectors. In this example Y is a multivariate normal random vector with mean $\mu 1_n$, $n = s \times t$. The covariance matrix of Y , which depends on the "within" variance component, σ_ε^2 , and the "between" variance component, σ_b^2 , can be written as

$$V = \sigma_\varepsilon^2 I_n + \sigma_b^2 ZZ'$$

where Z is the classification matrix for the random effects b_i and I_n is the identity matrix. The variance components are constrained so that V is positive definite which is equivalent to the conditions

$$\sigma_\varepsilon^2 > 0 \text{ and } \sigma_\varepsilon^2 + t\sigma_b^2 > 0.$$

Even in this simple model there is no exact procedure for testing $\sigma_b^2 = c \neq 0$ or for constructing confidence intervals for σ_b^2 . In PROC MIXED an approximate interval can be obtained by inverting the Wald test, based on either the full or residual likelihood. Using r or \tilde{r} one can construct typically more accurate intervals by performing a series of hypothesis tests and then inverting these to numerically find confidence limits. This can be done in SAS by making repeated calls to PROC MIXED. The program in the appendix calculates \tilde{r} for testing $\sigma_b^2 = c$.

In order to assess the performance of r and \tilde{r} we compare them to a specialized approximate test introduced by Ting et. al. (1990) and the REML directed deviance statistic. Simulation results were obtained with $s = 10$, $t = 2$, $\mu = 1$, and $\sigma_\varepsilon^2 = 1$ and

100,000 repetitions. Simulations with 10,000 repetitions conducted for t equal to 3 and 4 give similar results. The tables below contain the observed rejection proportions for testing $\sigma_b^2 = .5$ against both upper and lower alternatives.

Lower Alternatives

Nominal Level	10%	5%	1%	.1%
r	17.48%	10.04%	2.63%	.35%
REML	12.29%	6.44%	1.40%	.16%
Ting	10.15%	5.16%	1.05%	.13%
\tilde{r}	10.12%	5.08%	.99%	.11%

Upper Alternatives

Nominal Level	10%	5%	1%	.1%
r	6.04%	2.87%	.55%	.05%
REML	8.90%	4.22%	.82%	.08%
Ting	9.60%	4.70%	.85%	.07%
\tilde{r}	9.81%	4.97%	1.01%	.11%

In this model there are only two nuisance parameters so for moderate sample sizes one might expect that first order methods, such as r , REML, and the Wald statistic, would perform well. However, the simulation results suggest that these tests can be misleading while \tilde{r} performs quite well for both alternatives. The Wald test was uniformly inferior to r and hence is not shown. Ting's test is superior to the first order tests, r and REML, but slightly inferior to \tilde{r} for the lower nominal levels. Since Ting's test is based on the Cornish-Fisher approximation, which is known to perform poorly in the tails of the distribution, the inferiority of Ting's test in this region is not surprising.

Conclusion

The p-values and confidence intervals given by the MIXED procedure can often be misleading for small samples. However, for many hypotheses exact tests or accurate specialized approximations are not available. The \tilde{r} statistic is easy to obtain and often provides a significant improvement over available methods for one dimensional hypotheses. In SAS \tilde{r} can be calculated with PROC MIXED and PROC IML.

Appendix

For the balanced one-way random model S and q have simple forms:

$$S = i(\tilde{\beta}, \tilde{\sigma})$$

$$q_{\sigma} = -\frac{1}{2} \begin{pmatrix} \text{tr}(ZZ' (V^{-1}(\hat{\sigma}) - V^{-1}(\tilde{\sigma}))) \\ \text{tr}(V^{-1}(\hat{\sigma}) - V^{-1}(\tilde{\sigma})) \end{pmatrix}.$$

The term \tilde{u} then simplifies to

$$\tilde{u} = [S^{-1}q]_{\sigma_b^2} |X'V^{-1}(\tilde{\sigma})X|^{\frac{1}{2}} |j_{\sigma_{\varepsilon}\sigma_{\varepsilon}}(\tilde{\sigma})|^{-\frac{1}{2}} \times |i_{\sigma\sigma}(\tilde{\sigma})| |X'V^{-1}(\hat{\sigma})X|^{-\frac{1}{2}} |i_{\sigma\sigma}(\hat{\sigma})|^{-\frac{1}{2}},$$

where $i_{\sigma\sigma}$ is the variance component block of the expected information and $j_{\sigma_{\varepsilon}\sigma_{\varepsilon}}$ is the entry of the observed information corresponding to σ_{ε}^2 . The terms $|X'V^{-1}(\tilde{\sigma})X|$, $|X'V^{-1}(\hat{\sigma})X|$, and $|i_{\sigma\sigma}(\hat{\sigma})|$ can be obtained directly from PROC MIXED. The other terms are calculated in PROC IML using output from PROC MIXED. Below is a program that calculates \tilde{r} to test $\sigma_b^2 = c$.

```
% let sigO = c;
```

```
* Full MLEs
```

```
proc mixed method = ML nobound asycov;
class subj;
model y = / s xpvixi;
random subj;
make 'CovParms' out=randful;
make 'SolutionF' out=fixful;
make 'xpvixi' out=xpvixful;
make 'AsyCov' out=covful;
make 'fitting' out=fullfit(rename=(value=fullrt));
run;
```

```
* Restricted MLEs
```

```
proc mixed method = ML nobound asycov;
class subj;
model y = / s xpvixi pm;
random subj;
make 'CovParms' out=randred;
make 'SolutionF' out=fixred;
make 'xpvixi' out=xpvixred;
parms (&sigO) (1) / eqcons=1;
make 'PredMeans' out=resid(keep=resid);
make 'fitting' out=resfit(rename=(value=reslrt));
run;
```

```
*Calculate the LRT
```

```
data maker;
merge fullfit resfit;
lrt = 2*(fullrt - reslrt);
if descr = 'Log Likelihood' then output;
run;
```

```
proc iml;
```

```
/* Read in ZZ' and I (Identity Matrix)
Z output from PROC GLMMOD
Calculation of ZZ' and I in IML not shown. */
```

```
use ZZp;read all into ZZp;
use I;read all into I;
```

```
* Read in full MLEs of ( $\sigma_b^2$ ,  $\sigma_{\varepsilon}^2$ )
use randful; read all into randful;
* Read in restricted MLEs
use randred;read all into randred;
```

```
* Import  $X'\hat{V}^{-1}X$ 
use xpvixful; read all into xpvixful;
* Import  $X'\hat{V}^{-1}X$ 
use xpvixred; read all into xpvixred;
```

```
* Import Covariance
use covful;read all into covful;
```

```
* Import (y - X $\tilde{\beta}$ ) to calculate  $\tilde{j}$ 
use resid;read all into resid;
```

```
* Import LRT
use maker;
read all into maker;
```

```
*calculate r
sgnr = -1*((randful[1,2] - randred[1,2]) < 0) +
((randful[1,2] - randred[1,2]) > 0);
r = sgnr*sqrt(maker[1,3]);
```

```
* Calculate Variance with full and res. MLES
Vfuli = inv(I*randful[2,2] + randful[1,2]*ZZp);
Vredi = inv(I*randred[2,2] +randred[1,2]*ZZp);
```

```
* Calculate  $i_{\sigma\sigma}(\tilde{\sigma}) = S_{\sigma\sigma}$ 
rexpinf = repeat(0,2,2);
rexpinf[2,2] = .5#trace(Vredi**2);
rexpinf[1,2] = .5#trace(Vredi*ZZp*Vredi);
rexpinf[1,1] = .5#trace(Vredi*ZZp*Vredi*ZZp);
rexpinf[2,1] = rexpinf[1,2];
```

```
* Calculate q
q = repeat(0,2,1);
q[1,1] = - trace(ZZp*(Vfuli - Vredi));
q[2,1] = - trace(Vfuli - Vredi);
```

```

* Calculate  $\hat{j}_{\sigma_\varepsilon\sigma_\varepsilon}$ 
jss = -.5*(trace(Vredi**2) -
t(resid)*Vredi**3*resid);

* Calculate  $\tilde{u}$ 
Sq = inv(rexpinf)*q;
utild = Sq[1,1]*(xpvixred[1,2]** -.5)
*(1/sqrt(jss))*det(rexpinf)*
(xpvixful[1,2]**.5)*(det(covful)**.5);

* Calculate  $\tilde{r}$ 
rtild = r - (1/r)*sqrt(r/utild);
quit;

```

References

Barndorff-Nielsen, O.E. (1986), "Inference on full or partial parameters based on the standardized signed log-likelihood ratio," *Biometrika*, 73, 307-322.

Barndorff-Nielsen, O.E. and Cox, D.R., (1994), *Inference and Asymptotics*. New York: Chapman and Hall.

Lyons, B. and Peters, D., (1996), "A modified directed deviance statistic for mixed linear models," Technical Report 164, Department of Statistics, Oregon State University.

Pierce, D.A. and Peters, D., (1992), "Practical use of higher order asymptotics for multiparameter exponential families," *Journal of the Royal Statistical Society, Series B* 54, 701-737.

SAS Institute Inc. *SAS/STAT Software: Changes and Enhancements through Release 6.11*. Chapter 18. Cary, NC: SAS Institute Inc.

Skovgaard, I.M., (1996), "An explicit large-deviation approximation to one-parameter tests," *Journal of the Bernoulli Society*, 2, 145 - 165.

Ting, N., Burdick, R.K., Graybill, F.A., Jeyaratnam, S., and Lu T.-F.C., (1990), "Confidence intervals on linear combinations of variance components that are unrestricted in sign," *Journal of Statistical Computation and Simulation* 35, 135-141.

Wright, S. P., (1994), "Adjusted F Tests for Repeated Measures with the MIXED Procedure,"

328 SMC-Statistics Department, University of Tennessee.

SAS, SAS/STAT, and SAS/IML are registered trademarks or trademarks of the SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

For more information contact: Ben Lyons,
Department of Statistics, Oregon State University,
Corvallis, OR 97331. Phone: (541) 737-2578.
email: lyons@stat.orst.edu.