



---

Variance Components Testing in the Longitudinal Mixed Effects Model

Author(s): Daniel O. Stram and Jae Won Lee

Reviewed work(s):

Source: *Biometrics*, Vol. 50, No. 4 (Dec., 1994), pp. 1171-1177

Published by: [International Biometric Society](#)

Stable URL: <http://www.jstor.org/stable/2533455>

Accessed: 20/02/2013 05:32

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*International Biometric Society* is collaborating with JSTOR to digitize, preserve and extend access to *Biometrics*.

<http://www.jstor.org>

## Variance Components Testing in the Longitudinal Mixed Effects Model

Daniel O. Stram and Jae Won Lee

Department of Preventive Medicine, University of Southern California,  
Childrens Cancer Group, P. O. Box 60012,  
Arcadia, California 91066-6012, U.S.A.

### SUMMARY

This article discusses the asymptotic behavior of likelihood ratio tests for nonzero variance components in the longitudinal mixed effects linear model described by Laird and Ware (1982, *Biometrics* **38**, 963–974). Our discussion of the large-sample behavior of likelihood ratio tests for nonzero variance components is based on the results for nonstandard testing situations by Self and Liang (1987, *Journal of the American Statistical Association* **82**, 605–610).

### 1. Introduction

The use of likelihood ratio methods for constructing tests for nonzero variance components is a nonstandard problem in the use of maximum likelihood because the null hypothesis, that such a component is zero, places the true value of the variance parameters on the boundary of the parameter space defined by the alternative hypothesis. This has an effect on the large-sample behavior of likelihood ratio tests so that the limiting distribution of  $-2$  times the logarithm of the likelihood ratio (denoted  $-2 \ln \lambda_N$ ) cannot be treated as that of a  $\chi^2$  random variable. This paper discusses the problem of testing for nonzero variance components in the linear mixed effects model of Laird and Ware (1982). We describe the parameter-space geometry of such tests in sufficient detail as to be able to apply methods for nonstandard likelihood ratio testing described by Self and Liang (1987).

To motivate this paper consider the well-known growth-curve data set of Potthoff and Roy (1964). The data consist of measurements of the distance from the pituitary to the pterygomaxillary fissure for 27 children. All subjects have four observations, at years 8, 10, 12, and 14. The object is to model the growth in this distance as a function of age and sex of the children. We wish to allow for individual variability in growth as a function of age, with a mean intercept and growth rate depending on the sex of the individual. The model that we fit is of the form

$$y_i = [1 \text{ Age Sex Age} \times \text{Sex}] \alpha + [1 \text{ Age}] \beta_i + e_i. \quad (1)$$

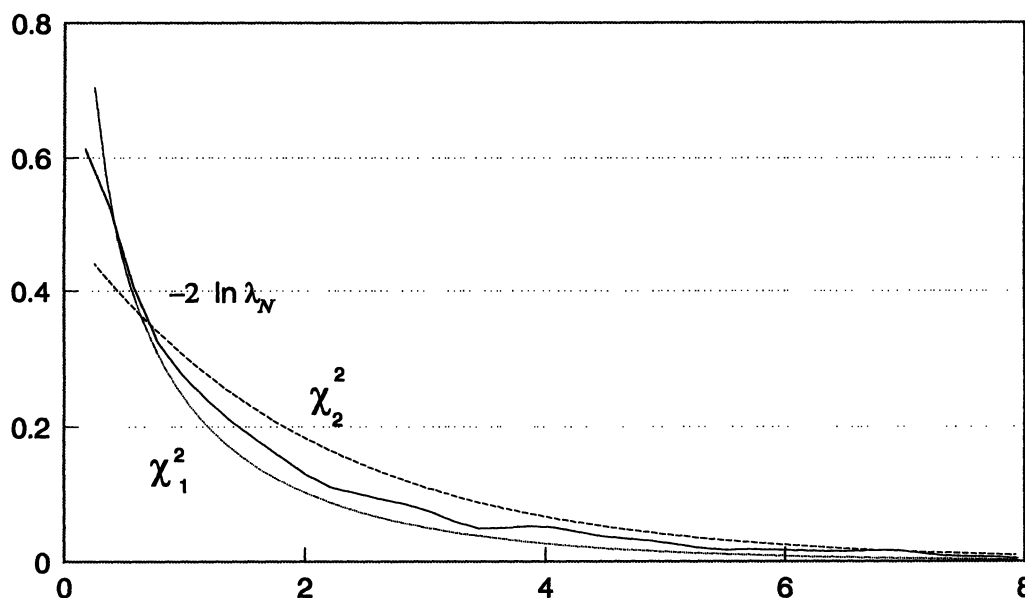
Here  $1$  denotes a  $4 \times 1$  vector of ones,  $\text{Sex}$  is a  $4 \times 1$  vector with components equal to 0 for boys and 1 for girls,  $\text{Age}$  is a  $4 \times 1$  vector of ages,  $\text{Age} = (8 \ 10 \ 12 \ 14)^T$ . The parameter vector  $\alpha$  corresponds to fixed effects, i.e., mean intercept and slope parameters for boys and girls,  $\beta_i$  is a random intercept and slope, allowing for individual differences in these parameters, and  $e_i$  are the errors around the individual regression lines. We assume that  $e_i$  are independent, normally distributed random variables with zero mean and common variance,  $\sigma^2$ , and that the  $\beta_i$  are distributed as multivariate normal with mean zero and covariance matrix  $D$ .

The values of the log-likelihoods when sequentially fitting three models to the Potthoff and Roy data are considered. Fitting under the assumption that  $D = 0$ , i.e., assuming there is no between-individual variability in  $\beta_i$ , the log-likelihood value is  $-239.12$ . Freeing  $D_{1,1}$ , to allow for individual variability in growth-curve intercept, gives a log-likelihood value of  $-214.32$ , so that  $-2 \ln \lambda_N = 49.6$ . Finally, freeing both  $D_{2,2}$  and  $D_{1,2}$  to allow for individual variation in growth rate as well as intercept, gives a log-likelihood of  $-213.90$ , so that  $-2 \ln \lambda_N = .833$ .

For testing the significance of individual variability in intercepts, the change in log-likelihood

---

**Key words:** Asymptotic distribution; Likelihood ratio tests; Variance components.



**Figure 1.** Smoothed empirical density of  $-2\ln\lambda_N$  for multivariate normal data simulated under the null hypothesis that  $d_{2,2} = d_{1,2} = 0$  using the parameter estimates obtained under the null from the Potthoff and Roy (1964) data.

criterion,  $-2\ln\lambda_N = 49.6$ , is so large as to leave little doubt that we may reject the hypothesis that  $D_{1,1} = 0$  with a high degree of confidence. In contrast, the evidence for individual variation in growth rate, as well as intercepts, appears very slight. Nevertheless it is important to compare the observed  $-2\ln\lambda_N = .833$  with the appropriate distribution of this statistic under the null hypothesis. To get a sense of what this appropriate distribution looks like, Figure 1 displays a smoothed estimate of the density of simulated values of  $-2\ln\lambda_N$  for testing that  $D_{2,2} = 0$ , from 1,000 sets of data simulated from the model in equation (1) using the parameter values for  $\alpha$ ,  $\sigma^2$ , and  $D_{1,1}$  estimated from the real data under the null hypothesis. Also shown in the figure are graphs of the density functions of  $\chi^2$  random variables, with 1 and 2 degrees of freedom, respectively. As made explicit in the following, it is no accident that the simulated distribution appears to fall approximately midway between these two distributions, and this paper treats the question of the appropriate asymptotic distribution of likelihood ratio tests for the Laird-Ware model in some detail. Note that the common practice (cf. Morrell and Brant, 1991) of relating the change  $-2\ln\lambda_N$  to a  $\chi^2$  random variable with the number of degrees of freedom equal to the increase in the number of variance and covariance parameters in  $D$  appears, in this case, to be somewhat overconservative; 485 of the 1,000 simulated values were above .833 compared to an expected 659 under the  $\chi^2_2$  distribution.

In general, we write the linear mixed effects model as

$$y_i = X_i\alpha + Z_i\beta_i + e_i, \quad (2)$$

where  $y_i$  is an  $n_i \times 1$  vector of observations for the  $i$ th independent subject,  $X_i$  and  $Z_i$  are known covariate matrices of order  $n_i \times p$  and  $n_i \times q$ , respectively, and  $\alpha$  is a  $p \times 1$  vector of mean parameters to be estimated. The  $\beta_i$  are independent  $q \times 1$  random vectors that are assumed to follow a multivariate Gaussian distribution with mean zero and variance matrix  $D$ , and the  $n_i \times 1$  vectors of residuals  $e_i$  are assumed to be independent normal with mean zero and variance matrix  $\sigma^2 I$ .

Many authors have described the methods for maximum likelihood and restricted maximum likelihood estimation of the parameters  $\alpha$ ,  $\sigma^2$ , and the unique elements of  $D$  in this model (cf. Laird and Ware, 1982; Jennrich and Schluchter, 1986; Lindstrom and Bates, 1988).

In the linear mixed effects model, the primary constraint upon the parameters is that  $D$ , the covariance matrix of the random effects  $\beta_i$ , be positive-semidefinite. In addition,  $\sigma^2$  is required to be positive, and we assume from now on that this is always the case. Null hypotheses concerning diagonal elements of  $D$ , i.e., hypotheses of the form  $d_{i,i} = 0$ , where  $d_{i,j}$  is the  $(i, j)$ th element of  $D$ , place  $D$  on the boundary of the parameter space, i.e., where  $D$  is positive-semidefinite, and of course impose the requirement that the corresponding off-diagonal elements of  $D$ ,  $d_{i,j}$  ( $j = 1, 2, \dots, i-1$ ), also be zero.

In this article, we discuss the asymptotic distribution of likelihood ratio tests of the null hypothesis that components or submatrices of  $D$  are equal to zero. In Section 2, the asymptotic behavior of likelihood ratio tests under nonstandard conditions, i.e., when the null hypothesis places parameters on the boundary of the parameter space defined by the alternative hypothesis, is briefly reviewed. The application of this theory to the tests concerning elements of  $D$  in the longitudinal mixed effects model is described in Section 3.

## 2. Large-Sample Distribution of Likelihood Ratio Statistics Under Nonstandard Conditions

In many cases, we want to test the hypothesis that  $\theta_0$  lies in a subset of  $p$ -dimensional set  $\Omega$ , denoted by  $\Omega_0$ , versus the alternative that  $\theta_0$  lies in the complement of  $\Omega_0$  in  $\Omega$ , denoted by  $\Omega_1$ . When  $\Omega_0$  is an  $r$ -dimensional subset of  $\Omega$  ( $r < p$ ),  $\theta_0$  is a boundary point of both  $\Omega_0$  and  $\Omega_1$  but  $\theta_0$  is an interior point of  $\Omega$ , and under the regularity conditions, the asymptotic distribution of the likelihood ratio test statistic,  $-2\ln\lambda_N$ , is  $\chi^2$  with  $p - r$  degrees of freedom. Chernoff (1954) provided a representation of the asymptotic distribution of  $-2\ln\lambda_N$  when both  $\Omega_0$  and  $\Omega_1$  have the same dimension as  $\Omega$ ,  $\theta_0$  is a boundary point of both  $\Omega_0$  and  $\Omega_1$ , but  $\theta_0$  is an interior point of  $\Omega$ . The asymptotic distribution of a class of test statistics (which includes likelihood ratio statistics) when  $\theta_0$  is on the boundary of  $\Omega_0$  but is an interior point of  $\Omega$  was also investigated by Shapiro (1985). He characterized this asymptotic distribution as a mixture of  $\chi^2$  distributions. Self and Liang (1987) generalized these results to the case in which  $\theta_0$  is a boundary point of  $\Omega$ .

We first introduce the following definition (cf. Chernoff, 1954) to discuss the theorem by Self and Liang (1987).

**Definition.** The set  $\Omega \subset R^p$  is approximated at  $\theta_0$  by a cone with vertex at  $\theta_0$ ,  $C_\Omega$ , if

$$(1) \inf_{y \in C_\Omega} \|x - y\| = o(\|y - \theta_0\|) \quad \text{for all } y \in \Omega$$

and

$$(2) \inf_{y \in \Omega} \|x - y\| = o(\|x - \theta_0\|) \quad \text{for all } x \in C_\Omega.$$

Note that a cone with vertex at  $\theta_0$ ,  $C$ , is a set of points such that if  $x \in C$  then  $a(x - \theta_0) + \theta_0 \in C$ , where  $a$  is any nonnegative real number.

**Theorem** (Self and Liang, 1987): Let  $Z$  be a random variable with a multivariate Gaussian distribution with mean  $\theta$  and covariance matrix  $I^{-1}(\theta_0)$ , and let  $C_{\Omega_0}$  and  $C_{\Omega_1}$  be non-empty cones approximating  $\Omega_0$  and  $\Omega_1$  at  $\theta_0$ , respectively. Then under the regularity conditions, the asymptotic distribution of the likelihood ratio statistic,  $-2\ln\lambda_N$ , is the same as the distribution of the likelihood ratio test of  $\theta \in C_{\Omega_0}$  versus the alternative  $\theta \in C_{\Omega_1}$  based on a single realization of  $Z$  when  $\theta = \theta_0$ .

The asymptotic representation of the likelihood ratio statistic given by the above theorem can be written as

$$\inf_{\theta \in C_0} \|\bar{Z} - \theta\|^2 - \inf_{\theta \in \bar{C}} \|\bar{Z} - \theta\|^2,$$

with  $\bar{C} = \{\bar{\theta}: \bar{\theta} = \Lambda^{1/2}P^T\theta \text{ for all } \theta \in C_{\Omega-\theta_0}\}$  and  $\bar{C}_0 = \{\bar{\theta}: \bar{\theta} = \Lambda^{1/2}P^T\theta \text{ for all } \theta \in C_{\Omega_0-\theta_0}\}$ , where  $\bar{Z}$  ( $= \Lambda^{1/2}P^TZ$ ) has a multivariate Gaussian distribution with mean 0 and identity covariance matrix, and  $P\Lambda P^T$  represents the spectral decomposition of  $I(\theta_0)$ . Self and Liang (1987) presented some special cases and used the above representation to calculate the distribution of likelihood ratio statistics.

## 3. Tests for Nonzero Variance Components in the Longitudinal Mixed Effects Model

In this section we use the theorem by Self and Liang (1987) to investigate the asymptotic behavior of the log-likelihood statistic in the longitudinal mixed effects model:

$$-2\ln\lambda_N = -2(l_N(\hat{\theta}_{\Omega_1}) - l_N(\hat{\theta}_{\Omega_0})),$$

where  $\theta$  represents the parameters, i.e.,  $\alpha$ ,  $\sigma^2$ , and the lower triangle of  $D$ , to be estimated, with  $\hat{\theta}_{\Omega_0}$  and  $\hat{\theta}_{\Omega_1}$  the maximum likelihood estimates under the null and alternative hypotheses, respectively, and where  $l_N(\theta)$  is the log-likelihood from model (1), i.e.,

$$\sum_{i=1}^N -\frac{n_i}{2} \ln(2\pi) - \frac{1}{2} \ln(\det(\Sigma_i^{-1})) - \frac{1}{2} \text{tr}(\Sigma_i^{-1}S_i), \quad (3)$$

where  $\Sigma_i = \sigma^2 I + Z_i D Z_i^T$  and  $S_i = (y_i - X_i \alpha)(y_i - X_i \alpha)^T$ .

The theorem by Self and Liang (1987) holds under regularity conditions that can be established for the mixed effects model under reasonable restrictions on the covariate matrices,  $X_i$  and  $Z_i$ , in model (1), which are needed to ensure that the information matrix,  $I(\theta)$ , for the parameters  $\alpha$ ,  $\sigma^2$ , and unique components of  $D$  increases with  $N$ . [See Szatrowski (1983) for detailed discussion of the form of this matrix in the linear structured covariance model, which includes the model of Laird and Ware as a special case.] The standardized score statistic is then asymptotically a mean-zero normal random variable with variance  $I^{-1}(\theta)$ , as required in the Self and Liang theorem. We are assuming here that the true value of  $\sigma^2$  is nonzero and that there are no additional constraints being imposed upon the estimates of  $\alpha$ . Because  $\sigma^2$  and  $\alpha$  thus lie in the interior of the admissible region for these parameters, we can restrict our descriptions of the geometry of  $\Omega_0$  and  $\Omega_1$  to deal only with  $D$  (a consequence of Self and Liang's Theorem 3; see also Case 4 and Case 5 of that paper for similar examples). The application of Self and Liang's results to testing for nonzero variance components in the mixed effects model thus requires an understanding of the geometry of the constraints on  $D$  so that the proper approximating cones can be constructed.

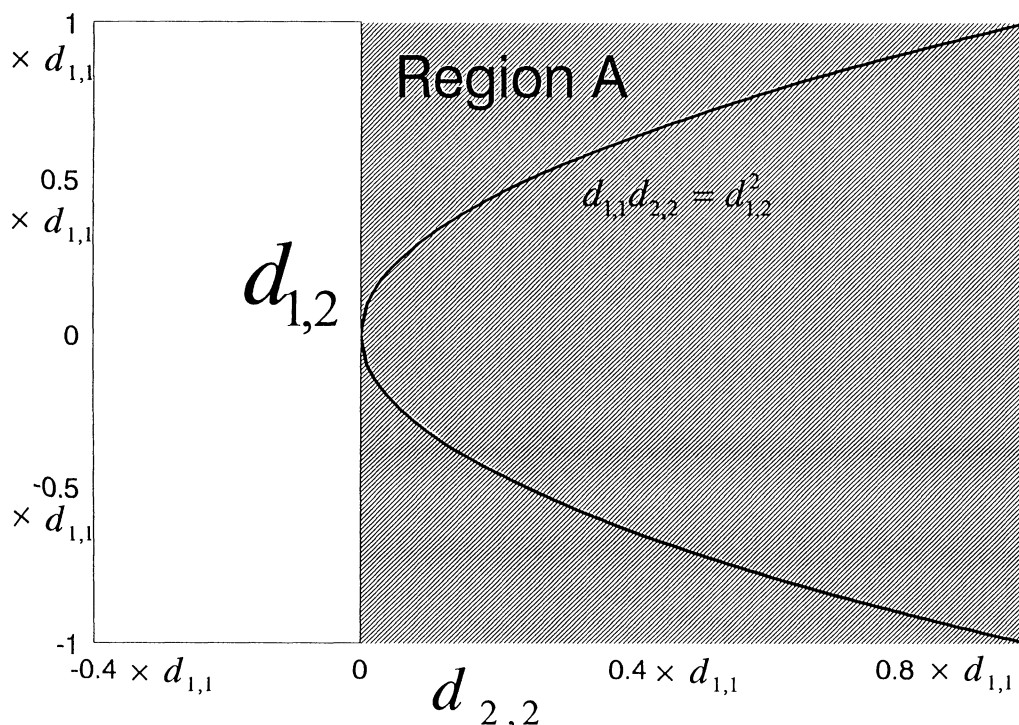
We now discuss the particulars of a number of specific hypotheses concerning  $D$ .

**Case 1:** Testing  $D = 0$  versus  $D = d_{1,1}$ , where  $d_{1,1}$  is an unspecified positive scalar.

Here,  $\theta = d_{1,1}$ ,  $C_{\Omega_0} = \{0\}$ , and  $C_{\Omega_1} = [0, \infty)$ . This corresponds to Case 5 of Self and Liang and the asymptotic distribution of  $-2\ln\lambda_N$  is a 50:50 mixture of  $\chi_0^2$  and  $\chi_1^2$ .

**Case 2:** Testing  $D = \begin{pmatrix} d_{1,1} & 0 \\ 0 & 0 \end{pmatrix}$  against  $D$  positive semidefinite.

Understanding the parameter-space geometry of this case is key to understanding the distribution of the likelihood ratio statistic, both for this problem and for more complicated hypothesis testing. Note that, with  $\theta = (d_{1,1}, d_{1,2}, d_{2,2})$ , the approximating cone at  $\theta_0$  is  $C_{\Omega_0} = (0, \infty) \times \{0\} \times \{0\}$ . Since  $d_{1,1}$  is assumed positive under both  $H_0$  and  $H_1$ , there is only one relevant constraint,  $d_{1,1}d_{2,2} \geq d_{1,2}^2$ . Figure 2 plots the parameter space for  $d_{1,1}$  and  $d_{1,2}$ . Because the parabola defined by the second constraint is smooth, i.e., its derivative,  $U(\theta)$ , exists at  $\Omega_0$ , the approximating cone at  $\theta_0$  is  $C_{\Omega_1} = (0, \infty) \times R^1 \times [0, \infty)$  (Region A, shaded). Thus by the above theorem, the large-sample distribution of  $-2\ln\lambda_N$  is a 50:50 mixture of  $\chi_2^2$  and  $\chi_1^2$ .



**Figure 2.** Parameter-space geometry of null and alternative hypotheses for testing that  $d_{2,2} = d_{1,2} = 0$ .

Case 3: Testing  $D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}$  with  $q \times q$  positive-definite matrix  $D_1$ , against the hypothesis that  $D$  is a  $(q + 1) \times (q + 1)$  positive-semidefinite matrix.

Here a total of  $q + 1$  parameters,  $d_{q+1,1}, d_{q+1,2}, \dots, d_{q+1,q+1}$ , are being added in the alternative hypothesis. Note that if  $D_1$  is positive-definite, the constraint that  $D$  be positive-semidefinite is equivalent to the system of  $(q + 1)$  constraints

$$\begin{aligned} d_{q+1,1}^2 &\leq d_{1,1}d_{q+1,q+1}, \\ d_{q+1,2}^2 &\leq d_{2,2}d_{q+1,q+1}, \\ &\vdots \\ d_{q+1,q}^2 &\leq d_{q,q}d_{q+1,q+1}. \end{aligned} \tag{4}$$

Here,  $\theta = (d_{1,1}, d_{1,2}, \dots, d_{q,q}, d_{q+1,1}, d_{q+1,2}, \dots, d_{q+1,q+1})$  and  $C_{\Omega_0} = C^* \times \{0\}^{q+1}$ , where  $C^* = \{(d_{1,1}, d_{1,2}, \dots, d_{q,q}); d_{i,i} > 0, i = 1, \dots, q\}$ . Now as in Case 2, since each of  $d_{1,1}$  to  $d_{q,q}$  is positive, the derivative of the boundary of the region defined by (4) exists and is defined as the plane  $d_{q+1,q+1} \geq 0$  so that  $C_{\Omega_1} = C^* \times R^q \times [0, \infty)$ . Thus by the above theorem the large-sample distribution of  $-2\ln\lambda_N$  is a 50:50 mixture of  $\chi_{q+1}^2$  and  $\chi_q^2$ .

Case 4: Testing  $D = \begin{pmatrix} D_{11} & 0 \\ 0 & 0 \end{pmatrix}$  against  $D = \begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix}$ .

Here we assume that  $D_{11}$  is a positive-definite matrix of order  $q \times q$ , and that  $D_{22}$  is of order  $k \times k$ . Under the alternative hypothesis we require  $D$  to be at least positive-semidefinite. In order to understand the large-sample distribution of  $-2\ln\lambda_N$  under these circumstances, it is best to deal first with the case when  $D_{22}$  is required to be a diagonal matrix; note, however, that we are not restricting  $D_{12}$  to be zero in this case. In this case the restriction that  $D$  be positive-semidefinite is equivalent to the set of restrictions

$$\begin{aligned} d_{q+i,1}^2 &\leq d_{1,1}d_{q+i,q+i}, \\ &\vdots \\ d_{q+i,q}^2 &\leq d_{q,q}d_{q+i,q+i}, \end{aligned} \tag{5}$$

for all  $i = 1, 2, \dots, k$ . Now by the same arguments as before, since each of  $d_{1,1}$  to  $d_{q,q}$  is required, by the null hypothesis, to be positive, the approximating cones are defined as  $C_{\Omega_0} = \{d_{q+1,q+1} = \dots = d_{q+k,q+k} = 0\}$  and  $C_{\Omega_1} = \{d_{q+i,q+i} \geq 0, i = 1, \dots, k\}$ . Asymptotically then we have, by the above theorem, that all constraints except those that  $d_{q+i,q+i}$  be nonnegative, vanish, and the large-sample distribution of  $-2\ln\lambda_N$  is a mixture of  $\chi^2$  random variables with degrees of freedom  $kq, kq - 1, kq - 2, \dots, (k - 1)q$ . The mixing probabilities for the  $\chi_{q(k-1)+j}^2$  components are sums of region volumes in  $(k + 1)q$ -dimensional space, and may in principle be calculated from the information for  $D_{22}$ , evaluated under the null hypothesis. In the very special case when the information matrix,  $I(\theta_0)$ , is the identity matrix, then the mixing probabilities are equal to  $\binom{q}{j-(k-1)q}2^{-q}$ , since this is equivalent to Case 9 of Self and Liang. For other  $I(\theta_0)$ , simulation methods may be applied to either estimating the mixing probabilities, or directly to the distribution of  $-2\ln\lambda_N$  as we have done in Figure 1.

For general  $D$  there are additional restrictions on  $D$  other than those in (4) in order that  $D$  be positive-semidefinite. For each off-diagonal element of  $D_{22}$  we have the new restriction that

$$d_{q+j,q+i}^2 \leq d_{q+j,q+j}d_{q+i,q+i}.$$

Note that as the true value,  $\theta_0$ , of the parameter vector approaches a point where  $d_{q+j,q+j}$  and  $d_{q+i,q+i}$  are zero, the admissible region collapses in on the off-diagonal elements  $d_{q+i,q+j}$  so that the approximating cone  $C_{\Omega_1}$  is equal to

$$C_{\Omega_1} = \{d_{q+i,q+i} \geq 0, i = 1, \dots, k; \quad d_{q+i,q+j} = 0, \quad i \neq j \quad (i = 1, \dots, k)\}.$$

Therefore we see that estimation of the off-diagonal elements of  $D_{22}$  has no asymptotic effect on the large-sample distribution of  $-2\ln\lambda_N$ , and this distribution remains the same mixture of  $\chi^2$  random variables with degrees of freedom  $kq, kq - 1, kq - 2, \dots, (k - 1)q$ , as when  $D_{22}$  is assumed to be diagonal.



#### 4. Discussion

We have focused here on the asymptotic distribution of likelihood ratio tests under the null hypothesis; other tests in common use, such as score tests, admit similar or identical asymptotic approximations, so that the description of their large-sample properties involves the same considerations as does  $-2\ln\lambda_N$ . In problems with boundary constraints, two distinct types of asymptotic approximations are required. First is the approximation by the Gaussian distribution of the distribution of the score, which is the same as in the unconstrained case. Second is the local approximation to the parameter space by cones. In Figure 2, as the information in the sample increases, the dispersion of the vector  $I^{-1}(\theta_0)S(\theta_0)$  (where  $S$  is the score) becomes small relative to the curvature of the alternative parameter space, and the projection of  $I^{-1}(\theta_0)S(\theta_0)$  on this parabola becomes well approximated by the projection onto the cone (the half-plane tangent to the parabola). In small samples, however, this approximation may be poor. Simulation studies can be useful for deciding whether the information in a specific sample is large enough for the asymptotic distribution to be relied upon. As illustrated in Figure 1, for the Potthoff and Roy data set, simulated values of  $-2\ln\lambda_N$  appear to indicate a close correspondence between the actual distribution of this statistic and its asymptotic distribution as a mixture of two  $\chi^2$  random variables.

The common practice in testing for the significance of new variance components in the longitudinal mixed model of relating the change,  $-2\ln\lambda_N$ , to a  $\chi^2$  random variable, with the number of degrees of freedom for the test equal to the increase in the number of variance and covariance terms in  $D$ , is asymptotically conservative. In the simplest kind of tests, where one new random effect (one more row of  $\beta_i$ ) is being added, the degree of bias in the significance level of a naive test, i.e., comparing  $-2\ln\lambda_N$  to a  $\chi^2_{q+1}$  rather than to a 50:50 mixture of  $\chi^2_q$  and  $\chi^2_{q+1}$ , is typically small. In certain complicated testing situations where many new random effects are being simultaneously added to  $\beta_i$ , however, the degree of asymptotic bias of the naive test could be quite substantial.

On the other hand, in certain hierarchical models for marginal association between elements of the outcome vector, the requirement that  $D$  be positive-definite may not be necessary, and  $D$  may even be allowed to have negative diagonal elements, so long as each of the marginal covariance matrices,  $\sigma^2 I + Z_i D Z_i^T$ , remains positive-definite. In such models the issues here are not relevant for testing hypotheses involving zero elements of  $D$ , since these are then standard rather than non-standard hypotheses. Even if one is actually interested in positive-definite  $D$ , the estimation of unconstrained  $D$  may be of utility for the purpose of calculating confidence intervals. First estimating the model with  $D$  unconstrained, and then constructing confidence intervals for  $D$  as the intersection of confidence intervals for unconstrained  $D$  with the admissible region, corresponds to the suggestion of Feng and McCulloch (1992).

In most cases the random effects interpretation of the linear mixed effects model, with its constraint upon  $D$ , is implicit in the use of this form of covariance matrix. This is seen in the computation procedures of most available software for fitting the mixed model: both the SAS procedure MIXED (SAS, 1992) and LME, which was developed by Lindstrom and Bates to run under S (Becker, Chambers, and Wilks, 1988), enforce this constraint upon  $D$ . The SAS procedure stops running when an indefinite value of  $D$  is encountered in the estimation, whereas the LME program appears to be very effective at finding the maximum likelihood value of a positive-semidefinite  $D$  lying on the constraint surface. Users of computer programs that behave like MIXED will have a somewhat different asymptotic distribution of the likelihood values that they can actually attain, than we have described here. Consider again the problem of adding one additional variance component to the model. Since MIXED forces the user to simplify the model, dropping the new random effect variance and all associated covariances, when an indefinite estimate of  $D$  is found, the realizable asymptotic distribution of  $-2\ln\lambda_N$  under the null hypothesis will be a 50:50 mixture of a  $\chi^2_{q+1}$  and a  $\chi^2_0$ , rather than a 50:50 mixture of  $\chi^2_{q+1}$  and  $\chi^2_q$ .

Recently there has been interest in the development of quasi-likelihood methods for several types of models (cf. Breslow, 1990; Prentice and Zhao, 1991; Breslow and Clayton, 1993). In these models likelihood ratio tests are often replaced with what may be termed "empirical" score tests. The Laird-Ware model (since it is a special case of the linear structured covariance matrix model) can be generalized within the GEE-2 framework of Prentice and Zhao (1991). Using these techniques, it is possible to weaken the assumption that the random vectors,  $\beta_i$  and  $e_i$ , are multivariate Gaussian while still providing consistent estimates of the variance-covariance matrix,  $D$ , of  $\beta_i$ . The asymptotic distribution of empirical score tests for "nonstandard" hypotheses concerning  $D$  will be subject to the same basic considerations as outlined here for the likelihood ratio test.

## ACKNOWLEDGEMENTS

This research was supported by National Cancer Institute Grant No. CA13539. The authors thank the referees and an associate editor for helpful suggestions which led to better organization and presentation of the results of this paper.

## RÉSUMÉ

Cet article traite du comportement asymptotique des tests du rapport de vraisemblance de la non-nullité des composantes de variances dans le modèle longitudinal à effets mixtes décrit par Laird et Ware (1982, *Biometrics* **38**, 963–974). Notre propos qui porte sur beaucoup de comportements de ce type de test, est fondé sur les résultats des tests en situation inhabituelle décrits par Self et Liang (1987, *Journal of the American Statistical Association* **82**, 605–610).

## REFERENCES

- Becker, R. A., Chambers, J. M., and Wilks, A. R. (1988). *The New S Language*. Pacific Grove, California: Wadsworth and Brooks/Cole.
- Breslow, N. (1990). Tests of hypotheses in overdispersed Poisson regression and other quasi-likelihood models. *Journal of the American Statistical Association* **85**, 565–571.
- Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear models. *Journal of the American Statistical Association* **88**, 9–25.
- Chernoff, H. (1954). On the distribution of the likelihood ratio. *Annals of Mathematical Statistics* **25**, 573–578.
- Feng, Z. and McCulloch, C. E. (1992). Statistical inference using maximum likelihood estimation and the generalized likelihood ratio when the true parameter is on the boundary of the parameter space. *Statistics and Probability Letters* **13**, 325–332.
- Jennrich, R. I. and Schluchter, M. (1986). Unbalanced repeated measures models with structural covariance matrices. *Biometrics* **42**, 802–820.
- Laird, N. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963–974.
- Lindstrom, M. J. and Bates, D. M. (1988). Newton–Raphson and EM algorithms for linear mixed effects models for repeated measures data. *Journal of the American Statistical Association* **83**, 1014–1022.
- Morrell, C. H. and Brant, L. J. (1991). Modelling hearing thresholds in the elderly. *Statistics in Medicine* **10**, 1453–1464.
- Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance useful especially in growth curve problems. *Biometrika* **51**, 313–326.
- Prentice, R. L. and Zhao, L. P. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics* **47**, 825–829.
- SAS (1992). Technical Report P-229. *SAS/STAT Software Changes and Enhancements, Release 6.07*. Cary, North Carolina: SAS Institute Inc.
- Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**, 605–610.
- Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika* **72**, 133–144.
- Szatrowski, T. H. (1983). Missing data in the one-population multivariate normal patterned mean and covariance matrix testing and estimation problem. *Annals of Statistics* **11**, 947–958.

*Received May 1993; revised July 1994; accepted July 1994.*