Research Article

Testing the Correlation and Heterogeneity for Hierarchical Nonlinear Mixed-Effects Models

Qingming Zou[1] and Zhongyi Zhu[2]

[1] *College of Economics and Management, University of South China, Hengyang, Hunan 421001, China*
[2] *Department of Statistics, Fudan University, Shanghai 200433, China*

Correspondence should be addressed to Zhongyi Zhu, zhuzy@fudan.edu.cn

Nonlinear mixed-effects models are very useful in analyzing repeated-measures data and have received a lot of attention in the field. It is of common interest to test for the correlation within clusters and the heterogeneity across different clusters. In this paper, we address these problems by proposing a class of score tests for the null hypothesis that all components of within- and between-subject variance are zeros in a kind of nonlinear mixed-effects model, and the asymptotic properties of the proposed tests are studied. The finite sample performance of this test is examined through simulation studies, and an illustrative example is presented.

1. Introduction

Repeated-measures data are frequent observations in different areas of investigation, such as economics and pharmacokinetics. For instance, in longitudinal studies, observations on the same subject are usually made at different times. Analysis of such data requires accounting for the within-cluster correlation and the between-subject heterogeneity of the data. Random-effects models are commonly used for analyzing clustered and repeated-measures data. It is of no doubt that the linear mixed-effects models play an important role in evaluating and analyzing the repeated-measures and clustered data. For an example, see Laird and Ware [1]. However, many repeated-measures data, such as growth data, dose-response and pharmacokinetics data, are inherently nonlinear with respect to a given response regression function. Several different nonlinear mixed-effects model and various inference procedures have been proposed [2–5]. For models considered in these literatures, one of the interesting issues is whether there exists correlation within clusters or/and heterogeneity between subjects. It is well known that the misspecification of model can have serious impact

on statistical inference. In regression model, two of the common assumptions are that of homogeneity and independence; the violation of these two assumptions can have adverse consequence for the efficiency of the estimators. So it is important to check these two assumptions whenever possible. For repeated-measures data and cluster data, if one can identify that there are no heterogeneity and correlation, which are caused by random effects, then a simple model can be used to fit the data and the efficient statistical inference can be obtained. However, if there are heterogeneity and correlation among outcomes and one does not identify their existence, then the overestimate or underestimate will be obtained for the parameters in model, and the precision and confidence of statistical inference are affected and even the mistaken conclusion is obtained. Therefore, It is an important and meaningful work to test the heteroscedasticity and correlation among outcomes in nonlinear mixed-effects model with repeated-measures data.

For testing explainable heterogeneity of variance, Cook and Weisberg [6] and Simonoff and Tsai [7] investigated the score test in classical linear models. Cai et al. [8], Eubank and Thomas [9], and Oyet and Sutradhar [10] considered some problems in nonparametric regression models. A number of tests for the unexplainable heterogeneity of variance have been proposed. Liang [11] applied the score test for homogeneity in different groups. Jacqmin-Gadda and Commenges [12] extended Liang's work to canonical generalized linear models(GLMs) with random effects and GEEs with a single correlation parameter. Lin [13] developed a unified theory for testing for correlation and heterogeneity in the framework of GLMs with random effects, using the Laplace expansion of the integrated log-quasilikelihood. Zhu and Fung [14] extended Lin's work [13] to semiparametric mixed models. Zhang and Weiss [15] divided the heterogeneity of the variance into explainable and unexplainable and considered the test of explainable heterogeneity. However, for a nonlinear mixed-effects model with repeated-measures data, such tests have not been developed. Our aim here is to develop a class of score tests for correlation within clusters and heterogeneity across different subjects in a nonlinear mixed-effects model.

Following Lin [13], we use the Laplace expansion to derive a class of score tests of homogeneity and correlation for hierarchical nonlinear mixed-effects model. The advantage of the score test is that it does not require the user to specify the joint distribution of the random effects. Thus this test is rather easy to be carried out. We give the asymptotic distribution of the test statistic under the null hypothesis and examine the finite-sample performance of the test through a Monte Carlo simulation study. It is found that the performance of this test is satisfactory in terms of both size and power even when the samples are of medium size. The rest of this paper is organized as follows. In Section 2, nonlinear mixed-effects model is introduced. In Section 3, we derive the score test statistic for the testing of correlation within clusters and heterogeneity across different subjects and give the asymptotic distribution of the test statistic under the null hypothesis. In Section 4, we examine the finite-sample performance of these tests through a Monte Carlo simulation study and a real example is analyzed.

2. Nonlinear Mixed-Effects Model

In this paper, we consider the following nonlinear mixed-effects model as proposed by Pinheiro and Bates [5]. In the first stage the jth observation on the ith subject is modeled as

$$y_{ij} = f(\phi_{ij}, x_{ij}) + \varepsilon_{ij}, \quad i = 1, \ldots, m, \ j = 1, \ldots, n_i, \tag{2.1}$$

where $f(\cdot, \cdot)$ is a twice-differentiable smooth nonlinear function of a subject-specific parameter vector ϕ_{ij} and the predictor x_{ij}, ε_{ij} is a normally distributed noised term, m is the total number of subjects, and n_i is the number of observations on the ith subject, and $n = \sum_{i=1}^{m} n_i$ is the total number of observations. In the second stage, the subject-specific vector is modeled as

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i, \tag{2.2}$$

where β is a $p \times 1$ vector of unknown parameters, $b_i, i = 1, \ldots, m$, are independent $q_i \times 1$ of random effects associated with the ith subject, and A_{ij} and B_{ij} are design matrices for fixed and random effects, respectively. The random effects b_i account for the correlation with the same cluster. Furthermore, we make the following distributional assumptions. The ε_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, are independently distributed as $N(0, \sigma^2 m_{ij})$, where $m_{ij} = m(z_{ij}, \gamma)$, $m(\cdot, \cdot)$ is a twice differentiable function, γ is a $r \times 1$ vector of unknown parameters, z_{ij} are the known design vectors, and $m(z_{ij}, \gamma_0) = 1$ if and only if $\gamma = \gamma_0$, and they are independent of b_i. Let $b = (b_1^T, \ldots, b_m^T)^T$ denote the vector obtained from stacking up the m cluster-specific entries and assume that b is generated from some distribution F with mean zero and covariance matrix $D(\theta)$ with $D(\theta) = \text{diag}(D_1(\theta), \ldots, D_m(\theta))$, where θ is a $s \times 1$ vector of unknown variance component. The magnitude of θ can be used to measure the degree of correlation and heterogeneity of the cluster-specific response vector within each subject. We postulate that each component of $D_i(\theta) = \text{cov}(b_i)$ is a function of θ such that $D_i(\theta) = 0 (i = 1, \ldots, m)$ if $\theta = \theta_0$. We further assume that the third- and higher-order moments of the random effects b_i are of order $o(\|\theta\|)$. These conditions are consistent with Lin [13], Hall and Praestgaard [16], and Zhu and Fung [14]. This model can be regarded as a hierarchical model that in some aspects generalized the linear mixed-effects model of Laird and Ware [1], the usual nonlinear model for independent data of Bates and Watts [17].

It should be pointed out that the distribution of the random error ε_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, need not to be normal. In this paper, we still assume the ε_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, to be normal. The main reason is to have the computation and deduction in mathematics become relatively simple. From the results presented in the below of this paper, it can be found that the derivation of the score test statistics is on the basis of the complete-data log likelihood, which requires making some specifications for the random error in model and further the random error is directly related to the complexity and difficulty of computation and deduction in mathematics. In nonlinear mixed-effects model, the random effects entering nonlinearly in model make the likelihood analysis of nonlinear mixed-effects model more difficult and complicated than that of their linear counterpart. It can be seen that even if the error distribution is assumed to be normal, the computation and deduction of the test score are cumbersome, let alone the error distribution is the others. In addition, the score test does not depend on the normality of the random error, but some good properties of the normal distribution help us obtain the test statistics. Therefore, we choose the normal distribution as that of the random error.

For model (2.1) and (2.2), there is either the correlation within clusters or heterogeneity among observations. First, we investigate whether there exists correlation among observations in the same subject; it is equivalent to test $D(\theta) = 0$ or not. Thus we can use the hypothesis

$$H_0 : \theta = \theta_0, \qquad H_1 : \theta \neq \theta_0 \tag{2.3}$$

for the test of the correlation within clusters in nonlinear mixed model. Second, we study whether there exists heterogeneity of between-subject variance and correlation within clusters at the same time. To address this problem, we can use the composite hypothesis

$$H_0 : \theta = \theta_0, \quad \gamma = \gamma_0, \quad H_1 : \theta \neq \theta_0, \quad \text{or} \quad \gamma \neq \gamma_0. \tag{2.4}$$

The testing of random effects in the nonlinear mixed-effects models has been discussed in some literatures (e.g., Jacqmin-Gadda and Commenges [12], Hall and Praestgaard [16], and Zhu and Fung [14]). However, the models they investigated are confined to the additive nonlinear mixed-effects model; that is, the random effects added to a nonlinear function or the rand noise is i.i.d random. Moreover, they generally studied the hypothesis (2.3), which tested whether there exists the correlation within clusters. In our paper, we consider the testing of random effects under some considerably general conditions for hierarchical nonlinear mixed-effects models. We not only study the testing correlation within clusters, but also investigate whether there exists heterogeneity of between-subject variance at the same time.

Let $Y_i = (y_{i1}, \ldots, y_{in_i})^T$, $f_i = (f_{i1}, \ldots, f_{in_i})^T$, $f_{ij} = f(\phi_{ij}, x_{ij})$, $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})^T$, and $M_i = \text{diag}(m_{i1}, \ldots, m_{in_i})$. Denote the the vector obtained from stacking up the m cluster-specific entries of the same symbol by $Y = (Y_1^T, \ldots, Y_m^T)^T$, $\mathbf{f} = (f_1^T, \ldots, f_m^T)^T$, $M = \text{diag}(M_1, \ldots, M_m)$, and $\varepsilon = (\varepsilon_1^T, \ldots, \varepsilon_m^T)^T$. To derive score tests for null hypothesis in which $\theta = \theta_0$ and $\theta = \theta_0$, $\gamma = \gamma_0$. We firstly study the properties of the log likelihood in model (2.1) and (2.2). For a given b, the conditional log-likelihood of nonlinear mixed-effects model and its derivative are as follows:

$$l^*\left(y, \beta, \sigma^2, \gamma, b\right) = -\frac{(Y - \mathbf{f})^T M^{-1} (Y - \mathbf{f})}{2\sigma^2} - \frac{n \ln \sigma^2}{2} - \frac{1}{2} \sum_{i=1}^{m} \ln|M_i|,$$

$$\left.\frac{\partial l^*(y, \beta, \sigma^2, \gamma, b)}{\partial b}\right|_{b=0} = -\frac{1}{\sigma^2} \left.\frac{\partial (Y - \mathbf{f})^T}{\partial b} M^{-1} (Y - \mathbf{f})\right|_{b=0},$$

$$\frac{\partial (Y - \mathbf{f})^T}{\partial b} = \left(\frac{\partial (Y_1^T - f_1^T)}{\partial b}, \ldots, \frac{\partial (Y_m^T - f_m^T)}{\partial b}\right), \tag{2.5}$$

$$\frac{\partial (Y_i^T - f_i^T)}{\partial b} = -\left(\frac{\partial f(\phi_{i1})}{\partial b}, \ldots, \frac{\partial f(\phi_{in_i})}{\partial b}\right) = -\begin{pmatrix} 0 \\ \Omega_i \\ 0 \end{pmatrix}_{q \times n_i},$$

where $\Omega_i = (B_{i1}^T \dot{f}_{i1}, \ldots, B_{in_i}^T \dot{f}_{in_i})_{q_i \times n_i}$, $q = \sum_{i=1}^{m} q_i$, $\dot{f}_{ij}$ is the first partial derivative of $f(\phi_{ij}, x_{ij})$ with respect to ϕ_{ij}, and

$$\left.\frac{\partial l^*(y, \beta, \sigma^2, \gamma, b)}{\partial b}\right|_{b=0} = \frac{1}{\sigma^2} \text{diag}(\Omega_1, \ldots, \Omega_m) M^{-1} (Y - \mathbf{f})|_{b=0},$$

$$\frac{\partial^2 l^*(y, \beta, \sigma^2, \gamma, b)}{\partial b \partial b^T} = \frac{1}{\sigma^2} \cdot \frac{\partial P}{\partial (b_1^T, \ldots, b_m^T)},$$

$$P = \left((Y_1 - f_1)^T M_1^{-1} \Omega_1^T, \dots, (Y_m - f_m)^T M_m^{-1} \Omega_m^T \right)^T,$$

$$\frac{\partial \Omega_i M_i^{-1}(Y_i - f_i)}{\partial (b_1^T, \dots, b_m^T)} = \left(\frac{\partial \Omega_i M_i^{-1}(Y_i - f_i)}{\partial b_1^T}, \dots, \frac{\partial \Omega_i M_i^{-1}(Y_i - f_i)}{\partial b_m^T} \right)$$

$$= \left(0, \dots, \frac{\partial \Omega_i M_i^{-1}(Y_i - f_i)}{\partial b_i^T}, \dots, 0 \right),$$

$$\left. \frac{\partial \Omega_i M_i^{-1}(Y_i - f_i)}{\partial b_i^T} \right|_{b=0} = \frac{\partial \sum_{j=1}^{n_i} B_{ij}^T \dot{f}_{ij} m_{ij}^{-1}(y_{ij} - f(\phi_{ij}, x_{ij}))}{\partial b_i^T} = Q_i,$$

(2.6)

where $Q_i = \sum_{j=1}^{n_i} B_{ij}^T [\ddot{f}_{ij} m_{ij}^{-1}(y_{ij} - f(\phi_{ij})) - \dot{f}_{ij} m_{ij}^{-1} \dot{f}_{ij}^T] B_{ij}$ and $\ddot{f}_{ij}$ is the second partial derivative of $f(\phi_{ij}, x_{ij})$ with respect to ϕ_{ij} evaluated at $b = 0$, and then

$$\left. \frac{\partial^2 l^*(y, \beta, \sigma^2, \gamma, b)}{\partial b \partial b^T} \right|_{b=0} = \frac{1}{\sigma^2} \operatorname{diag}(Q_1, \dots, Q_m).$$

(2.7)

3. Score Test for Correlation and Heterogeneity within Clusters

In this section, we firstly use the Laplace expansions to develop a score test for the null hypothesis $H_0 : \theta = 0$ under model (2.1) and (2.2), which corresponds to having no correlation within clusters. Then using the same approach, we obtain a score test statistic for the null hypothesis $H_0 : \theta = 0$, $\gamma = 0$ under model (2.1) and (2.2), which corresponds to having no correlation within clusters and heterogeneity of between-subject variance at the same time.

Let $\xi = (\beta^T, \sigma^2, \gamma^T)^T$, for an $n \times 1$ data vector Y following model (2.1) and (2.2). The likelihood function is

$$L(\xi, \theta) = \int \exp\{l^*(y, \xi, \theta, b)\} dF(b).$$

(3.1)

Using the moment assumptions on the random effects b and the Laplace expansion similar to Lin [13] and Hall and Praestgaard [16], we expand the integrated likelihood (3.1) and obtain the marginal log-likelihood for (ξ, θ) as follows:

$$l(y, \xi, \theta) = l^*(y, \xi, 0) + \frac{1}{2} \operatorname{tr}\left[D(\theta) \left(\frac{\partial l^*(y, \xi, 0)}{\partial b} \frac{\partial l^*(y, \xi, 0)}{\partial b^T} + \frac{\partial^2 l^*(y, \xi, 0)}{\partial b \partial b^T} \right) \right] + o(\|\theta\|),$$

(3.2)

where $l^*(y, \xi, 0) = l^*(y, \xi, \theta, b)|_{\theta=\theta_0}$, $\partial l^*(y, \xi, 0)/\partial b = (\partial l^*(y, \xi, \theta, b)/\partial b)|_{b=0}$ and $\partial^2 l^*(y, \xi, 0)/\partial b \partial b^T = (\partial^2 l^*(y, \xi, \theta, b)/\partial b \partial b^T)|_{b=0}$. Note that (3.2) mimics a Laplace expansion of $l(y, \xi, \theta)$ [18]. Many authors extended this expansion to a variety of models; see Lin [13] for an

example. It is worth pointing out that (3.2) is a second-order expansion of the marginal distribution of Y about $b = 0$. For model (2.1) and (2.2), from the log-quasilikelihood expansion (3.2), some calculations give the efficient score $U_\theta(\hat{\xi}_0) = (U_{\theta_1}(\hat{\xi}_0), \ldots, U_{\theta_s}(\hat{\xi}_0))$ under the null hypothesis $H_0 : \theta = \theta_0$ as follows:

$$
\begin{aligned}
U_{\theta_k}\left(\hat{\xi}_0\right) &= \left. \frac{\partial l(\xi, \theta)}{\partial \theta_k} \right|_{\xi = \hat{\xi}_0, \theta = \theta_0} \\
&= \frac{1}{2\hat{\sigma}^4} \hat{e}^T M^{-1} \operatorname{diag}\left(\Omega_1^T, \ldots, \Omega_m^T\right) \dot{D}_k(0) \operatorname{diag}(\Omega_1, \ldots, \Omega_m) M^{-1} \hat{e} \\
&\quad + \frac{1}{2\sigma^2} \operatorname{tr}\left(\dot{D}_k(0) \operatorname{diag}(Q_1, \ldots, Q_m)\right), \quad k = 1, \ldots, s,
\end{aligned}
$$
(3.3)

where $\hat{e}$ is the magnitude of $e = Y - \mathbf{f}$ evaluated at $(\xi, \theta) = (\hat{\xi}_0, \theta_0)$, $\dot{D}_k(0) = (\partial D / \partial \theta_k)|_{\theta=\theta_0}$ ($k = 1, \ldots, s$) and $\hat{\xi}_0$ is the maximum likelihood estimator of ξ under $\theta = \theta_0$.

To test for $H_0 : \theta = \theta_0$, we construct a score statistics as follows:

$$
\mathrm{ST} = U_\theta\left(\hat{\xi}_0\right)^T I_\theta\left(\hat{\xi}_0\right)^{-1} U_\theta\left(\hat{\xi}_0\right),
$$
(3.4)

where $I_\theta(\hat{\xi}_0) = I_{\theta\theta} - I_{\theta\xi} I_{\xi\xi}^{-1} I_{\xi\theta}$ is the efficient information matrix of θ evaluated under H_0. Here,

$$
I_{\theta\theta} = E\left(\frac{\partial l}{\partial \theta} \frac{\partial l}{\partial \theta^T}\right), \qquad I_{\xi\theta} = E\left(\frac{\partial l}{\partial \xi} \frac{\partial l}{\partial \theta^T}\right) = I_{\theta\xi}^T, \qquad I_{\xi\xi} = E\left(\frac{\partial l}{\partial \xi} \frac{\partial l}{\partial \xi^T}\right),
$$
(3.5)

where $l = l(\xi, \theta)$, and the scores $\partial l / \partial \xi$, $\partial l / \partial \theta$ and the expectations are all calculated at $\theta = \theta_0$. Noting the properties of the normal distribution, after some computations, we obtain, for $h, k = 1, \ldots, s$, $u, v = 1, \ldots, r$,

$$
I_{\theta_h \theta_k}\left(\hat{\xi}_0\right) = \frac{1}{2\hat{\sigma}^4} \operatorname{tr}\left(M^{-1} G M^{-1} H\right) + \frac{1}{4\hat{\sigma}^2} \sum_{i=1}^m \sum_{j=1}^{n_i} \operatorname{tr}\left(\dot{D}_{k_i}(0) B_{ij}^T \ddot{f}_{ij} B_{ij}\right) \operatorname{tr}\left(\dot{D}_{h_i}(0) B_{ij}^T \ddot{f}_{ij} B_{ij}\right) m_{ij}^{-1},
$$
(3.6)

$$I_{\beta\theta_k}\left(\widehat{\xi}_0\right) = \frac{1}{2\widehat{\sigma}^2}(\Sigma_1,\ldots,\Sigma_m)M^{-1}\begin{pmatrix} \text{tr}(\dot{D}_{k_1}(0)B_{11}^T\ddot{f}_{11}B_{11}) \\ \vdots \\ \text{tr}(\dot{D}_{k_m}(0)B_{mn_m}^T\ddot{f}_{mn_m}B_{mn_m}) \end{pmatrix},$$

$$I_{\sigma^2\theta_k}\left(\widehat{\xi}_0\right) = \frac{1}{2\widehat{\sigma}^4}\,\text{tr}\left(HM^{-1}\right), \qquad I_{\beta\beta}\left(\widehat{\xi}_0\right) = \frac{1}{\widehat{\sigma}^2}\sum_{i=1}^{m}\Sigma_i M_i^{-1}\Sigma_i^T,$$

$$I_{\gamma_h\theta_k}\left(\widehat{\xi}_0\right) = -\frac{1}{2\widehat{\sigma}^2}\,\text{tr}\left(M_{(h)}^{-1}\,\text{diag}\left(\widehat{\Omega}_1^T,\ldots,\widehat{\Omega}_m^T\right)\dot{D}_k(0)\,\text{diag}\left(\widehat{\Omega}_1,\ldots,\widehat{\Omega}_m\right)\right),$$

(3.7)

$$I_{\beta\sigma^2}\left(\widehat{\xi}_0\right) = 0, \qquad I_{\beta\gamma}\left(\widehat{\xi}_0\right) = 0,$$

$$I_{\sigma^2\sigma^2}\left(\widehat{\xi}_0\right) = \frac{n}{2\widehat{\sigma}^4}, \qquad I_{\sigma^2\gamma_u}\left(\widehat{\xi}_0\right) = -\frac{1}{2\widehat{\sigma}^2}\,\text{tr}\left(M_{(u)}^{-1}M\right),$$

$$\begin{aligned} I_{\gamma_u\gamma_v}\left(\widehat{\xi}_0\right) &= \frac{1}{2}\,\text{tr}\left(M_{(u)}^{-1}MM_{(v)}^{-1}M\right) + \frac{1}{4}\Big(\text{tr}\left(M_{(u)}^{-1}M\right)\text{tr}\left(M_{(v)}^{-1}M\right) \\ &\quad + \text{tr}\left(M_{(u)}^{-1}M\right)\text{tr}(M_{(v)}M^{-1}) \\ &\quad + \text{tr}(M_{(u)}M^{-1})\,\text{tr}\left(M_{(v)}^{-1}M\right) \\ &\quad + \text{tr}(M_{(u)}M^{-1})\,\text{tr}(M_{(v)}M^{-1})\Big), \end{aligned}$$

where $\Sigma_i = (A_{i1}^T\dot{f}_{i1},\ldots,A_{in_i}^T\dot{f}_{in_i})$, $\dot{D}_{k_i}(0) = \partial D_i/\partial\theta_k{}_{\theta=\theta_0}$ ($i = 1,\ldots,m$, $k = 1,\ldots,s$), $G = \text{diag}(\widehat{\Omega}_1^T,\ldots,\widehat{\Omega}_m^T)\dot{D}_h(0)\,\text{diag}(\widehat{\Omega}_1,\ldots,\widehat{\Omega}_m)$, $\ddot{f}_{ij} = \partial^2 f(x_{ij},\phi_{ij})/(\partial\phi_{ij}\partial\phi_{ij}^T)|_{b=0}$, $H = \text{diag}(\widehat{\Omega}_1^T,\ldots,\widehat{\Omega}_m^T)\dot{D}_k(0)\,\text{diag}(\widehat{\Omega}_1,\ldots,\widehat{\Omega}_m)$, $M_{(k)}^{-1} = \partial M^{-1}/\partial\gamma_k$, and $m_{ij}^{(k)} = (\partial m(z_{ij},\gamma))/\partial\gamma_k$, $M_{(k)} = \text{diag}(m_{ij}^{(k)})$, $i = 1,\ldots,m$, $j = 1,\ldots,n_i$. The detailed derivation of (3.6) is supplied in Appendix 4.2, and the others are similar and omitted.

One important feature of the proposed score test statistic ST is that a detailed specification of the distribution $F(b,\theta)$ of the random effects is not necessary. Therefore, the test is robust against arbitrary mixed model alternation where only the first two moments of the random effects are specified. The following gives the asymptotic properties of the proposed score test statistic. The "asymptotic" in the theorem refers to the number of clusters $m \to \infty$ with cluster sizes n_i bounded in model (2.1) and (2.2).

Theorem 3.1. *For model* (2.1) *and* (2.2), *under regularity conditions in Appendix 4.2, when H_0 in* (2.3) *is true, the asymptotic distribution of the score test statistic ST is a χ^2-distribution with s degrees of freedom.*

Now, one considers testing the composite hypothesis

$$H_0 : \theta = \theta_0, \qquad \gamma = \gamma_0, \tag{3.8}$$

for detecting whether there exists heterogeneity of between-subject variance and correlation within clusters at the same time. Let $\alpha = (\beta^T,\sigma^2)^T$, $\eta = (\theta^T,\gamma^T)^T$, $\eta_0 = (\theta_0^T,\gamma_0^T)^T$, and $\phi = (\alpha^T,\eta^T)^T$, and

let $\hat{\alpha}_0$ be the maximizer of $l(y, \alpha, 0)$ under H_0. The Fisher information matrix I_ϕ can be obtained by differentiating (3.2) twice, taking the expectation under H_0, and evaluating at $\phi = (\hat{\alpha}_0^T, \eta_0^T)$. If one partitions the information matrix as

$$I_\phi = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\eta} \\ I_{\eta\alpha} & I_{\eta\eta} \end{pmatrix}, \tag{3.9}$$

here

$$I_{\alpha\alpha} = E\left(\frac{\partial l}{\partial \alpha} \frac{\partial l}{\partial \alpha^T} \right), \qquad I_{\alpha\eta} = E\left(\frac{\partial l}{\partial \alpha} \frac{\partial l}{\partial \eta^T} \right) = I_{\eta\alpha}^T, \qquad I_{\eta\eta} = E\left(\frac{\partial l}{\partial \eta} \frac{\partial l}{\partial \eta^T} \right), \tag{3.10}$$

where $l = l(\alpha, \eta)$ and the score $\partial l / \partial \alpha$, $\partial l / \partial \eta$ and the expectations are all evaluated at $\eta = \eta_0$, then, similar to (3.4), one constructs a score statistic as follows:

$$ST1 = U_\eta(\hat{\alpha}_0)^T \hat{I}_\eta(\hat{\alpha}_0)^{-1} U_\eta(\hat{\alpha}_0), \tag{3.11}$$

where $U_\eta(\hat{\alpha}_0) = (\partial l / \partial \theta^T, \partial l / \partial \gamma^T)^T|_{\alpha=\hat{\alpha}_0, \eta=\eta_0} = (U_\theta^T, U_\gamma^T)^T|_{\alpha=\hat{\alpha}_0, \eta=\eta_0}$, $\hat{I}_\eta = I_{\eta\eta} - I_{\eta\alpha} I_{\alpha\alpha}^{-1} I_{\alpha\eta}$ is the efficient information matrix of η. Similar to the derivation of ST, the expression form of the score statistics ST1 can be obtained. Therefore, for the sake of space, the detailed derivation of ST1 is omitted. The asymptotic properties of the score test statistic ST1 are given as follows.

Corollary 3.2. *For model (2.1) and (2.2), under some regularity conditions, when H_0 in (2.4) is true, the asymptotic distribution of the score test statistic ST1 is a χ^2-distribution with $s + r$ degrees of freedom.*

Remark 3.3. It is obvious that obtaining MLEs of parameter under null hypothesis is very crucial in the score test. To obtain the MLEs of the unknown parameters in the hierarchical nonlinear mixed-effects models (2.1) and (2.2), we can take the Newton-Raphson method or the method of score according to the score functions $\partial l / \partial \beta$, $\partial l / \partial \sigma^2$ and $\partial l / \partial \gamma_k$ (see Appendix 4.2).

Denote by $\xi = (\beta, \sigma^2, \gamma)^T$ the unknown parameter vector, and let $l(\cdot)$ be the likelihood function defined in (3.2) under the null hypothesis $\theta = \theta_0$ and $\dot{l}(\cdot), \ddot{l}(\cdot)$ are the score function and Hessian matrix, respectively. If $\tilde{\xi}$ is the value of parameter vector ξ at the previous iteration, the New-Raphson methods gets the new estimates $\hat{\xi}$ such that

$$\hat{\xi} = \tilde{\xi} - \ddot{l}\left(\tilde{\xi} \right)^{-1} \dot{l}\left(\tilde{\xi} \right), \tag{3.12}$$

and the methods of score get

$$\hat{\xi} = \tilde{\xi} - E\left[\ddot{l}\left(\tilde{\xi} \right)^{-1} \right] \dot{l}\left(\tilde{\xi} \right). \tag{3.13}$$

Applying (3.12) or (3.13) iteratively, we can obtain the MLEs $\hat{\xi}$. For the asymptotic distribution of the estimate $\hat{\xi}$ of ξ, we can use the Taylor expansion to show that

$$\hat{\xi} - \xi_0 = I_{\xi\xi}^{-1}(\xi_0)U_{\xi}(\xi_0) + O_p\left(n^{-1}\right), \tag{3.14}$$

where ξ_0 is the true value of ξ under the null hypothesis and $I^{-1}(\cdot)$ is the Fisher information matrix.

It should be pointed out that the estimating parameters in the nonlinear model are a challenge. In general, the iterating procedure of obtaining the estimates of unknown parameters in a nonlinear model is a popular way. For example, Wong et al. [19] used the Newton-Raphson iteration approach to get the maximum likelihood estimation of ARMA model with error process for replicated observation. In this paper, we adopt a similar method to obtain the MLEs of hierarchical nonlinear mixed-effects model. We do not show in detail the convergence of the algorithm and only give the idea of proof, but we explore the behavior of this algorithm both in the simulation study and the analysis of a real data set in Section 4 and find that the convergence of this algorithm is guaranteed and the precision of convergence is well. The practical results show that the algorithm is reasonable and feasible.

4. Empirical Investigations

4.1. Simulation Studies

We perform some simulation studies to evaluate the sizes and the powers of the score tests proposed in Section 3. We first draw data for m subjects with $n_i = 10$ measurements on each unit from the model

$$y_{ij} = \frac{\beta_1 x_{ij}}{\beta_2 + b_{i1} + x_{ij}} + \left(\beta_3 + b_{i2}\right)x_{ij} + \varepsilon_{ij} \quad (i = 1, \ldots, m, \ j = 1, \ldots, 10), \tag{4.1}$$

where x_{ij} are independent random variates from $N(0, 1/4)$, the ε_{ij} are random noise having the $N(0, \sigma^2 m_{ij})$, and $m_{ij} = \exp(z_{ij}\gamma)$ with z_{ij} as independent random drawing from $N(0, 1/4)$, and the random effects b_i $(i = 1, \ldots, m)$ are independent random from distribution $N_2(0, \theta I_2)$, where I_2 is an 2×2 identity matrix. We vary θ from 0 to 0.2, to study the sizes and powers of the score test ST and ST1.

We consider four different sample sizes, $m = 10, 20, 40, 60$. The experiment is replicated 2000 times for each parameter configuration. The nominal sizes of the tests are set to be 0.05. The empirical sizes and powers of test statistics ST and ST1 are presented in Tables 1 and 2 respectively.

The results in Tables 1 and 2 show that the empirical sizes of the tests are very close to 0.05. As the θ, γ, and m increase, the power of the test increases quickly and approaches 1. Furthermore, we notice that, for the score test ST1, when $\theta = 0$ but $\gamma \neq 0$, the test has lower power. We speculate that the higher power is obtained because the discrepancy between the model and the postulated model is larger when there exist the random effects in the model; otherwise, the lower power is obtained. These findings are consistent with the theoretical results and also show that random effects in the models may be main factors influencing the inference performance.

Table 1: Empirical sizes and powers of the score test statistics ST in 2000 replications.

	θ	sizes and powers			
m		10	20	40	60
	0.000	0.0571	0.0557	0.0539	0.0503
	0.050	0.5832	0.6650	0.7655	0.8187
	0.075	0.6894	0.7413	0.8358	0.8861
	0.100	0.7302	0.8036	0.9027	0.9435
	0.125	0.7621	0.8602	0.9587	0.9749
	0.150	0.7836	0.8774	0.9695	0.9891
	0.175	0.8449	0.9195	0.9785	0.9927
	0.200	0.8575	0.9875	0.9947	1

Table 2: Empirical sizes and powers of the score test statistics ST1 in 2000 replications.

	θ	γ	sizes and powers			
m			10	20	40	60
	0.00	0.00	0.0577	0.0556	0.0548	0.0533
	0.00	0.20	0.4574	0.6175	0.6602	0.6798
	0.00	0.30	0.5519	0.7993	0.8271	0.8636
	0.00	0.50	0.6602	0.8127	0.8630	0.8784
	0.05	0.00	0.9097	0.9679	0.9894	0.9910
	0.05	0.20	0.9473	0.9795	0.9920	0.9963
	0.05	0.30	0.9890	0.9951	0.9972	1
	0.05	0.50	0.9983	0.9996	1	1
	0.10	0.00	0.9320	0.9741	0.9875	0.9926
	0.10	0.20	0.9548	0.9803	0.9915	0.9987
	0.10	0.30	0.9855	0.9925	0.9997	1
	0.10	0.50	0.9925	0.9995	1	1

To further confirm the performance of the test ST and ST1, we present the *Q-Q* plots of ST and ST1 with $m = 20, 40$ in Figures 1 and 2, respectively. The others are omitted for the sake of space. The *Q-Q* plots also confirm that the asymptotic distribution of the test ST and ST1 is χ^2-distribution, which is consistent with the theoretical results.

4.2. An Illustrative Example

We illustrate the use of the test in analysis of longitudinal study. The data of the example are taken from the guinea pig data in Johansen [20]. In the experiment, 50 tissue samples were taken from the intestine of each of eight guinea pigs. For each guinea pig, five tissue samples were assigned randomly to each different concentration of B-methyl-glucoside and the uptake volume was measured in micromoles per milligram of fresh tissue per 2 min. Only the means of the five tissue samples at each concentration for each animal were used. The data is previously analyzed by Lee and Xu [21], to investigate the diagnostic measures through
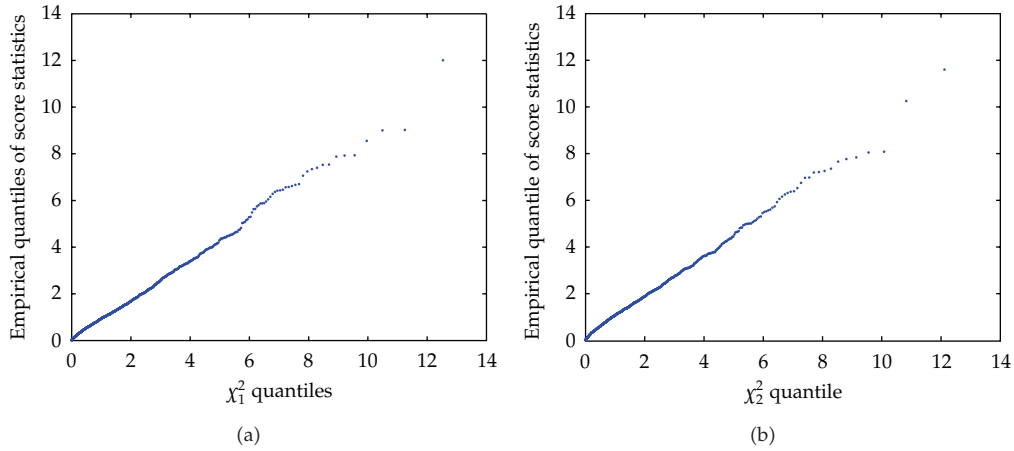
(a)

(b)

Figure 1: Q-Q plots of ordered score statistics against χ_1^2 quantiles based on 2000 replications generated from nonlinear mixed-effects model (2.1) under $H_0 : \theta = 0$.
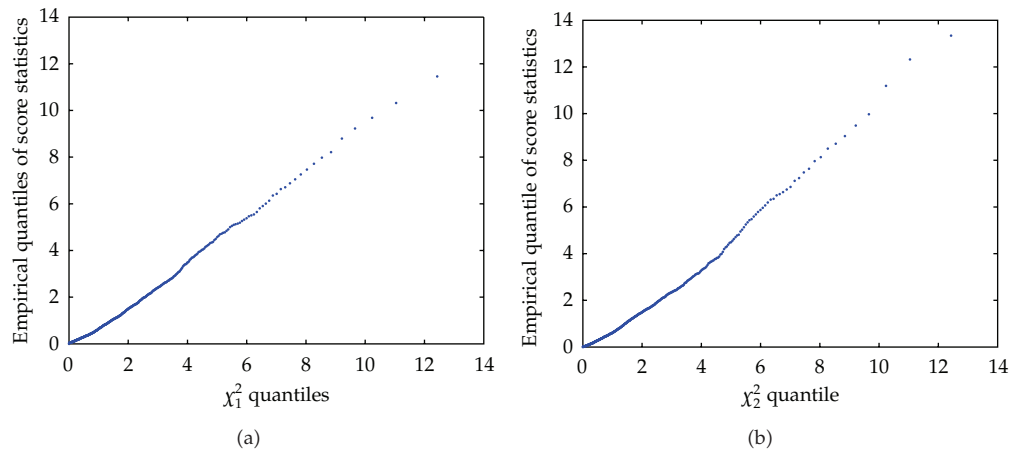

(a)

(b)

Figure 2: Q-Q plots of ordered score statistics against χ_2^2 quantiles based on 2000 replications generated from nonlinear mixed-effects model (2.1) under $H_0 : \theta = 0, \gamma = 0$.

local influence analysis. They proposed nonlinear mixed-effects model for uptake volume as function of concentration is

$$ y_{ij} = \log \left[\frac{\beta_1 x_{ij}}{\beta_2 + b_{i1} + x_{ij}} + (\beta_3 + b_{i2}) x_{ij} \right] + \varepsilon_{ij}, \tag{4.2} $$

where y_{ij} is the jth uptake volume for individual i, x_{ij} is the jth concentration level for individual i, $\varepsilon_{ij} \sim N(0, \sigma^2)$, and $b_i = (b_{i1}, b_{i2})^T$ is a vector of individual random effects with $b_i \sim N(0, \sigma^2 D)$, where $D = \begin{pmatrix} \theta_1 & \theta_2 \\ \theta_2 & \theta_3 \end{pmatrix}$. In this paper, we assume that $\varepsilon_{ij} \sim N(0, \sigma^2 \exp(z_{ij}\gamma))$ and $z_{ij} = x_{ij}/100$. According to the algorithm proposed in Section 3, under the null hypothesis $\theta = (\theta_1, \theta_2, \theta_3) = 0$; that is, there exist no random effects in the model, we obtain

β_1 = 0.2002 (0.0042), β_2 = 2.3844 (0.0671), β_3 = 0.0044 (0.0086), σ^2 = 0.0012, γ = 2.4004, and ST = 104.1653 (P = 0.0000). The degree of freedom of ST is 3. We also calculate the test statistic ST1 and obtain β_1 = 0.2023 (0.0041), β_2 = 2.4175 (0.0795), β_3 = 0.0043 (0.0002), σ^2 = 0.0017, and ST1 = 91.4559 (P = 0.0000) under $\theta = 0, \gamma = 0$. The degree of freedom of ST1 is 4. The small P-values of the test statistics suggest strongly rejecting the null hypothesis of independence and homogeneity $\theta = 0$ and $\gamma = 0$. These results demonstrate that there exists heterogeneity of the between-subject variance and correlation within clusters at the same time for the guinea pig data. It should be pointed out that our results are consistent with those of Lee and Xu [21]. They also illustrated the dependence and heterogeneity for the guinea pig data. These results may suggest that the original model having both correlation within clusters and the heterogeneity across different clusters for the guinea pig data should be taken. The illustrative example also demonstrates that the score test can efficiently detect the random effects among the outcomes in practice.

It is worth noticing that the estimate of the random effects is an interesting topic when the existence of random effects is proved. However, in this paper, our interests focus on testing whether the random effects exist. So the estimate of random effects is out of our study scope. Lee and Xu [21] gave an estimate of $(\theta_1, \theta_2, \theta_3)$ through SA-MCMC algorithm.

Appendices

Appendix A

In what follows, we assume that the expectations are taken under $\theta = \theta_0$. According to (3.3), the (h, k)th component $(h, k = 1, \ldots, s)$ of $I_{\theta\theta}$ is,

$$
\begin{aligned}
I_{\theta_h \theta_k} = E(U_{\theta_h} U_{\theta_k}) = \frac{1}{4} E \Bigg[& \frac{1}{\sigma^8} e^T M^{-1} G M^{-1} e e^T M^{-1} H M^{-1} e + \frac{1}{\sigma^6} e^T M^{-1} G M^{-1} e \\
& \cdot \operatorname{tr}(\dot{D}_k(0) \operatorname{diag}(Q_1, \ldots, Q_m)) \\
& + \frac{1}{\sigma^6} e^T M^{-1} H M^{-1} e \operatorname{tr}(\dot{D}_h(0) \operatorname{diag}(Q_1, \ldots, Q_m)) \\
& + \frac{1}{\sigma^4} \operatorname{tr}(\dot{D}_h(0) \operatorname{diag}(Q_1, \ldots, Q_m)) \operatorname{tr}(\dot{D}_k(0) \operatorname{diag}(Q_1, \ldots, Q_m)) \Bigg]
\end{aligned}
$$

$$
= I_1 + I_2 + I_3 + I_4,
$$

$$\tag{A.1}$$

where $G, H, \dot{D}_i(0), Q_j, \Omega_j, i = 1, \ldots, s, \ j = 1, \ldots, m$, all being the same as that of in Section 3, respectively.

Note the following properties of the normal distribution: if $\varepsilon \sim N(0, \Sigma)$, then

$$
E\left(\varepsilon^T G \varepsilon \varepsilon^T H \varepsilon\right) = 2 \operatorname{tr}(\Sigma G \Sigma H) + \operatorname{tr}(\Sigma G) \operatorname{tr}(\Sigma H), \qquad E\left(\varepsilon^T G \varepsilon\right) = \operatorname{tr}(G \Sigma), \qquad \text{(A.2)}
$$

for any nonnegative definite matrices G and H. After some calculations, we obtain

$$I_1 = \frac{1}{4\sigma^4}\left[2\operatorname{tr}(M^{-1}GM^{-1}H) + \operatorname{tr}(M^{-1}G)\operatorname{tr}(M^{-1}H)\right],$$

$$I_2 = I_3 = -\frac{1}{4\sigma^4}\operatorname{tr}(M^{-1}G)\operatorname{tr}(M^{-1}H),$$

$$I_4 = \frac{1}{4\sigma^4}\operatorname{tr}(M^{-1}G)\operatorname{tr}(M^{-1}H) + \frac{1}{4\hat{\sigma}^2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\operatorname{tr}(\dot{D}_{k_i}(0)B_{ij}^T\ddot{f}_{ij}B_{ij})\operatorname{tr}(\dot{D}_{h_i}(0)B_{ij}^T\ddot{f}_{ij}B_{ij})m_{ij}^{-1}.$$

$$(A.3)$$

Therefore, (3.6) holds immediately.

Appendix B

Proof of the Asymptotic Distribution of ST.

Here we study the asymptotic distribution of ST under $H_0 : \theta = \theta_0$. Let

$$U = \begin{pmatrix} U_\xi \\ U_\theta \end{pmatrix}, \qquad I = \begin{pmatrix} I_{\xi\xi} & I_{\xi\theta} \\ I_{\xi\theta}^T & I_{\theta\theta} \end{pmatrix}, \tag{B.1}$$

where

$$U_\xi = \left(\left(\frac{\partial l}{\partial \beta}\right)^T, \frac{\partial l}{\partial \sigma^2}, \left(\frac{\partial l}{\partial \gamma}\right)^T\right)^T,$$

$$\frac{\partial l}{\partial \beta} = -\frac{1}{\sigma^2}(\Sigma_1, \ldots, \Sigma_m)M^{-1}(Y - \mathbf{f}),$$

$$\frac{\partial l}{\partial \sigma^2} = \frac{1}{2\sigma^4}(Y - \mathbf{f})^T M^{-1}(Y - \mathbf{f}) - \frac{n}{2\sigma^2},$$

$$\frac{\partial l}{\partial \gamma_k} = \frac{1}{2\sigma^2}(Y - \mathbf{f})^T M_{(k)}^{-1}(Y - \mathbf{f}) - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{m_{ij}^{(k)}}{m_{ij}},$$

$$(B.2)$$

and U_θ is given in (3.6). Assume ξ_0 is the true value of ξ. For obtaining the asymptotic properties of the score test statistic ST. We assume the following regularity conditions under $\theta = \theta_0$. These assumptions are similar to those given in [13] by Lin.

Condition 1. The size of cluster is a finite sequence of positive integers, the first- and second-order partial derivatives of $f(\cdot, \cdot)$ with respect to parameter are bounded. The components of $\dot{D}_k(0),\ k = 1, \ldots, s$, are uniformly bounded.

Condition 2. There exists a neighborhood $O(\xi_0)$. The components of $\partial l/\partial \beta,\ \partial l/\partial \gamma$ and $\partial l/\partial \sigma^2$ are bounded in $O(\xi_0)$.

Condition 3. The log-quasilikelihood $l(\xi, \theta_0)$ of $\xi = (\beta^T, \sigma^2, \gamma^T)^T$ has the usual asymptotic properties, including consistency of $\hat{\xi}$ and the linear expansion

$$n^{1/2}\left(\hat{\xi}_0 - \xi_0\right) = \left\{nI_{\xi\xi}^{-1}(\xi_0)\right\}\left\{n^{-1/2}U_{\xi}(\xi_0)\right\} + o_p(1). \tag{B.3}$$

Condition 4. There exists a positive definite matrix $I^0 = \begin{pmatrix} I_{\xi\xi}^0 & I_{\xi\theta}^0 \\ I_{\xi\theta}^{0T} & I_{\theta\theta}^0 \end{pmatrix}$, such that

$$\lim_{n \to \infty} n^{-1}I = I^0. \tag{B.4}$$

Proof. For any given $((p + 1 + r) + s) \times 1$ constant vector $\lambda = (\lambda_1^T, \lambda_2, \lambda_3^T, \lambda_4^T)^T$, where λ_1 is an $p \times 1$ vector, λ_2 is a constant, λ_3 is an $r \times 1$ vector, and λ_4 is an $s \times 1$ vector, we have

$$\lambda^T U = \left(\lambda_1^T, \lambda_2, \lambda_3^T\right)U_{\xi} + \lambda_4^T U_{\theta} = -\frac{1}{\sigma^2}\sum_{i=1}^{m}\lambda_1^T \Sigma_i M_i (y_i - f_i)$$

$$+ \frac{1}{2\sigma^4}\sum_{i=1}^{m}(Y_i - f_i)^T \lambda_2 M_i (Y_i - f_i) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{k=1}^{r}\lambda_{3k}(Y_i - f_i)^T M_i^{-1(k)}(Y_i - f_i)$$

$$+ \frac{1}{2\sigma^4}\sum_{i=1}^{m}\sum_{k=1}^{s}\lambda_{4k}(Y_i - f_i)^T M_i^{-1}\Omega_i^T \dot{D}_{k_i}(0)\Omega_i M_i^{-1}(Y_i - f_i) \tag{B.5}$$

$$+ \sum_{k=1}^{s}\lambda_{4k}\,\mathrm{tr}\left(\dot{D}_k(0)\,\mathrm{diag}(Q_1, \ldots, Q_m)\right) - \frac{n\lambda_2}{2\sigma^2} - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k=1}^{r}\lambda_{3k}\frac{m_{ij}^{(k)}}{m_{ij}}.$$

Because $\dot{D}_k(0)$ is a block diagonal matrix, according to the definition of m-dependent sequence, then $\lambda^T U$ is the summation of a sequence of m-dependent random variables; that is, $\lambda^T U$ can be written as

$$\lambda^T U = \sum_{i=1}^{m}U_{\lambda,i}, \tag{B.6}$$

where $\{U_{\lambda,i}\}$ is an m-dependent sequence. From Conditions 1 and 2, we have that $\{U_{\lambda,i}\}$ are uniformly bounded in $O(\xi_0)$ for any given λ. It can be shown that $E(U_{\lambda,i}) = 0$. By applying Theorem 7.3.1 of Chung [22] and applying the Condition 4 to $\lambda^T U$, we have

$$n^{-1/2}\lambda^T U(\xi_0) \longrightarrow N\left(0, \lambda^T I^0(\xi_0)\lambda\right) \tag{B.7}$$

in distribution as $n \to \infty$. Using the Cramer-Wald device, we conclude that $n^{-1/2}U(\xi_0) \to N(0, I^0(\xi_0))$ in distribution.

   Note the linear expansion of the efficient score statistic $U_\theta(\hat{\xi}_0)$ about $\xi = \xi_0$ and use Condition 3. It follows that $n^{-1/2}U_\theta(\hat{\xi}_0) = n^{-1/2}U_\theta(\xi_0) - \{n^{-1}I_{\xi\theta}(\xi_0)\}^T\{nI_{\xi\xi}^{-1}(\xi_0)\}\{n^{-1/2}U_{\xi}(\xi_0)\} +$

$o_p(1)$. This implies that $n^{-1/2}U_\theta(\widehat{\xi}_0) \rightarrow N(0, I_\theta^0(\xi_0))$ in distribution, where $I_\theta^0 = I_{\theta\theta}^0 - I_{\xi\theta}^{0T}(I_{\xi\xi}^0)^{-1}I_{\xi\theta}^0$. According to the consistency of $\widehat{\xi}_0$ and Slutsky's theorem, we have that

$$\text{ST} = U_\theta\left(\widehat{\xi}_0\right)^T I_\theta\left(\widehat{\xi}_0\right)^{-1} U_\theta\left(\widehat{\xi}_0\right) = \left[n^{-1/2}U_\theta\left(\widehat{\xi}_0\right)\right]^T \left[n^{-1}I_\theta\left(\widehat{\xi}_0\right)\right]^{-1}\left[n^{-1/2}U_\theta\left(\widehat{\xi}_0\right)\right] \qquad \text{(B.8)}$$

converges in distribution to a chi-square distribution with s degrees of freedom as $n \rightarrow \infty$, that is, $m \rightarrow \infty$. $\qquad\square$

Acknowledgments

References

[1] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.

[2] L. B. Sheiner and S. L. Beal, "Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data," *Journal of Pharmacokinetics and Biopharmaceutics*, vol. 8, no. 6, pp. 553–571, 1980.

[3] M. J. Lindstrom and D. M. Bates, "Nonlinear mixed effects models for repeated measures data," *Biometrics*, vol. 46, no. 3, pp. 673–687, 1990.

[4] W. K. Wong and R. B. Miller, "Analysis of ARIMA-Noise models with repeated time series," *Journal of Business and Economic Statistics*, vol. 8, pp. 243–250, 1990.

[5] J. C. Pinheiro and D. M. Bates, "Approximations to the log-likelihood function in the nonlinear mixed-effects model," *Journal of Computational and Graphical Statistics*, vol. 4, pp. 12–35, 1995.

[6] R. D. Cook and S. Weisberg, "Diagnostics for heteroscedasticity in regression," *Biometrika*, vol. 70, no. 1, pp. 269–274, 1983.

[7] J. S. Simonoff and C. L. Tsai, "Improved tests for nonconstant variance in regression based on the modified profile likelihood," *Applied Statistics*, vol. 42, pp. 31–41, 1994.

[8] Z. W. Cai, C. M. Hurvich, and C. L. Tsai, "Score tests for heteroscedasticity in wavelet regression," *Biometrika*, vol. 85, no. 1, pp. 229–234, 1998.

[9] R. L. Eubank and W. Thomas, "Detecting heteroscedasticity in nonparametric regression," *Journal of the Royal Statistical Society. Series B*, vol. 55, no. 1, pp. 145–155, 1993.

[10] A. J. Oyet and B. Sutradhar, "Testing variances in wavelet regression models," *Statistics & Probability Letters*, vol. 61, no. 1, pp. 97–109, 2003.

[11] K. Y. Liang, "A locally most powerful test for homogeneity with many strata," *Biometrika*, vol. 74, no. 2, pp. 259–264, 1987.

[12] H. Jacqmin-Gadda and D. Commenges, "Tests of homogeneity for generalized linear models," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1237–1246, 1995.

[13] X. Lin, "Variance component testing in generalised linear models with random effects," *Biometrika*, vol. 84, no. 2, pp. 309–326, 1997.

[14] Z. Y. Zhu and W. K. Fung, "Variance component testing in semiparametric mixed models," *Journal of Multivariate Analysis*, vol. 91, no. 1, pp. 107–118, 2004.

[15] F. Zhang and R. E. Weiss, "Diagnosing explainable heterogeneity of variance in random-effects models," *The Canadian Journal of Statistics*, vol. 28, no. 1, pp. 3–18, 2000.

[16] D. B. Hall and J. T. Præstgaard, "Order-restricted score tests for homogeneity in generalised linear and nonlinear mixed models," *Biometrika*, vol. 88, no. 3, pp. 739–751, 2001.

[17] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, New York, NY, USA, 1988.

[18] P. J. Solomon and D. R. Cox, "Nonlinear component of variance models," *Biometrika*, vol. 79, no. 1, pp. 1–11, 1992.

[19] W. K. Wong, R. B. Miller, and K. Shrestha, "Maximum likelihood estimation of ARMA model with error processes for replicated observations," *Journal of Applied Statistical Science*, vol. 10, no. 4, pp. 287–297, 2001.

[20] S. Johansen, *Functional Relations, Random Coefficients, and Nonlinear Regression with Application to Kinetic Data*, vol. 22 of *Lecture Notes in Statistics*, Springer, New York, NY, USA, 1984.

[21] S. K. Lee and L. Xu, "Influence analyses of nonlinear mixed-effects models," *Computational Statistics & Data Analysis*, vol. 45, no. 2, pp. 321–341, 2004.

[22] K. L. Chung, *A Course in Probability Theory*, London, UK, 2nd edition, 1974.