TESTING VARIANCE COMPONENTS IN BALANCED LINEAR GROWTH CURVE MODELS

REZA DRIKVANDI[1], AHMAD KHODADADI[1] AND GEERT VERBEKE[2,3]

[1] *Department of Statistics, Shahid Beheshti University, 1983963113 G.C, Tehran, Iran*
[2] *I-BioStat, Katholieke Universiteit Leuven, Kapucijnenvoer 35, B-3000 Leuven, Belgium*
[3] *I-BioStat, Universiteit Hasselt, Agoralaan 1, B-3590 Diepenbeek, Belgium*

Abstract. It is well known that the testing of zero variance components is a non-standard problem since the null hypothesis is on the boundary of the parameter space. The usual asymptotic chi-square distribution of the likelihood ratio and score statistics under the null does not necessarily hold because of this null hypothesis. To circumvent this difficulty in balanced linear growth curve models, we introduce an appropriate test statistic and suggest a permutation procedure to approximate its finite-sample distribution. The proposed test alleviates the necessity of any distributional assumptions for the random effects and errors and can easily be applied for testing multiple variance components. Our simulation studies show that the proposed test has Type I error rate close to the nominal level. The power of the proposed test is also compared with the likelihood ratio test in the simulations. An application on data from an orthodontic study is presented and discussed.

Keywords: *Balanced linear growth curve model; Efficiency; Likelihood ratio; Linear mixed effects model; Permutation test; Variance components.*

1. Introduction

The Linear Growth Curve (LGC) model (Rao, 1965; Laird and Ware, 1982; von Rosen, 1991) is an important type of Linear Mixed Effects (LME) model is used in human and animal studies for a variety of applications such as: tumor biology, pharmacokinetics, and clinical studies of height, weight and pulmonary function development (Demidenko and Stukel, 2002).

Given N distinct individuals, the LGC model is typically expressed in two stages. The first stage consists of the linear regression model with random coefficients

$$Y_i = Z_i a_i + \varepsilon_i, \qquad i = 1, ..., N, \tag{1.1}$$

where $Y_i = (Y_{i1}, ..., Y_{in_i})'$ is an $n_i \times 1$ vector of repeated measurements on the ith individual, Z_i is an $n_i \times k$ design matrix of full rank k, a_i is a $k \times 1$ random vector of individual coefficients and ε_i is an $n_i \times 1$ vector of random errors with $E(\varepsilon_i) = 0$ and $cov\,(\varepsilon_i) = \sigma^2 I_{n_i}$. Random vectors a_i and ε_i are mutually independent and independent across i. The second stage consists of the linear model

$$a_i = A_i \beta + b_i, \qquad i = 1, ..., N, \tag{1.2}$$

where A_i is a $k \times m$ design matrix, β is an $m \times 1$ vector of parameters of interest and b_i is a $k \times 1$ vector of random effects with $E(b_i) = 0$ and $cov\,(b_i) = \sigma^2 D$. Scalar σ^2 and matrix D are called variance components and are unknown.

The LGC model, (1.1) and (1.2), is a balanced linear growth curve (BLGC) model if matrices A_i can be represented as (see Demidenko, 2004)

$$A_i = (I_k \otimes q_i') \,, \qquad i = 1, ..., N, \tag{1.3}$$

where q_i is a $p \times 1$ design vector and $Z_i = Z$ for $i = 1, ..., N$. Thus, the BLGC model can be written as

$$Y_i = ZA_i\beta + Zb_i + \varepsilon_i, \qquad i = 1, ..., N. \tag{1.4}$$

When $p = 1$ and $q_i = 1$, the BLGC model (1.4) reduces to the balanced random-coefficient (BRC) model

$$Y_i = Z\beta + Zb_i + \varepsilon_i, \qquad i = 1, ..., N. \tag{1.5}$$

Two special cases of the BRC model (1.5) are the balanced mixed one-way ANOVA model (3.2) and the linear trend model with random intercepts and random slopes (3.1) which are important in many applications.

A number of reasons makes it desirable to test the presence of random effects in the model. Inconsistent estimators of the covariance matrix of fixed effect estimators are generally obtained if random effects are neglected. On the other hand, estimators of models with random effects are inefficient in the absence of random effects. Also, since random effects represent the variability between individuals, in many applications such as the orthodontic growth data presented in Section 4, it may be of interest to test for the need of random slopes and random intercepts in the model.

In Section 2, we introduce an appropriate test statistic and suggest a permutation procedure to approximate its finite-sample distribution. In Section 3, we conduct simulation studies to evaluate and compare the efficiency of the proposed test with respect to the likelihood ratio test. In Section 4, we apply the proposed test to an orthodontic growth data set. We conclude the paper in Section 5.

2. **The Test Statistic**

Initially we should test whether all random effects can be left out of the BLGC model (1.4). In statistical language, this translates into testing the hypothesis

$$H_0 : D = 0 \tag{2.1}$$

versus the alternative hypothesis that D is a non-negative definite matrix.

In the literature, the likelihood ratio (LR), score and F tests have been suggested for testing variance components in LME models; see, for example, Stram and Lee (1994), Lin (1997), Verbeke and Molenberghs (2003), Demidenko (2004) and Giampaoli and Singer (2009). The LR, score and F tests are all based on the normality assumption for both the random effects and errors. Unfortunately, the normality assumption of the random effects and errors often does not hold in practice and may not always give robust results. On the other hand, it is well known that the usual asymptotic chi-square distribution of the LR and score statistics under the null does not necessarily hold because $D = 0$ is the boundary point of the parameter space. Instead, the large sample distribution is a mixture of chi-square distributions which may not always be determined for testing multiple random effects. For more details, see, Miller (1977), Self and Liang (1987), Dean (1992), Stram and Lee (1994), Lin (1997), Gueorguieva (2001), Verbeke and Molenberghs (2003) and Fitzmaurice et al. (2007), for example. To avoid the issues with testing on the boundary of the parameter space, Fitzmaurice et al. (2007) proposed a permutation test which can only be used for testing a single variance component. In this paper, we propose a permutation test which can be applied for testing multiple variance components in BLGC models.

To explain the rationale for the new test statistic proposed in this paper, we use the notation $D_* = \sigma^2 D$ and consider the statistic

$$T = \frac{1}{N} tr\left(Z\hat{D}_* Z'\right), \tag{2.2}$$

where $\hat{D}_*$ is any distribution-free unbiased estimator of D_* (e.g., minimum norm quadratic unbiased estimator (MINQUE) or method of moments (MM) estimator). It can easily be shown that under H_0, $E(T) = 0$. Hence, one can reject H_0, if T deviates much from zero. In general in the LME models, there is no closed-form expression for $\hat{D}_*$. However, particularly for the BLGC model (1.4), a closed-form expression for $\hat{D}_*$ exists.

Demidenko and Stukel (2002) have shown that an unbiased MM estimator of D_* for the BLGC model (1.4) is

$$\hat{D}_* = \frac{1}{N-p}\hat{S} - \frac{\hat{\sigma}^2}{N-p}\left(Z'Z\right)^{-1}\sum_{i=1}^{N}\left(1 - q_i'Q^{-1}q_i\right), \tag{2.3}$$

in which $\hat{\sigma}^2 = \frac{1}{N(n-k)}\sum_{i=1}^{N}\left\|Y_i - Za_i^0\right\|^2$, $Q = \sum_{i=1}^{N} q_i q_i'$ and

$$\hat{S} = \sum_{i=1}^{N}\left(a_i^0 - A_i\hat{\beta}_{OLS}\right)\left(a_i^0 - A_i\hat{\beta}_{OLS}\right)',$$

where $a_i^0 = \left(Z'Z\right)^{-1} Z'Y_i$ is the OLS estimator of a_i in model (1.1) and

$$\hat{\beta}_{OLS} = \left(\sum_{i=1}^{N} A_i'Z'ZA_i\right)^{-1}\left(\sum_{i=1}^{N} A_i'Z'Y_i\right).$$

The estimator (2.3) coincides with both the Bayes empirical estimator developed by Reinsel (1985) and the MINQUE proposed by Demidenko (2004).

Substituting (2.3) into (2.2), we have

$$T = \frac{1}{N(N-p)}\sum_{i=1}^{N}\left\{\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)' P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right) - \frac{k}{N}\hat{\sigma}^2 Q^*\right\}, \tag{2.4}$$

where $Q^* = \sum_{i=1}^{N}\left(1 - q_i'Q^{-1}q_i\right)$ and $P_Z = Z\left(Z'Z\right)^{-1}Z'$ is a projection matrix. For proof of equation (2.4), see Appendix.

Without any distributional assumption for the random effects and errors, finding the exact distribution of T is difficult. Thus we approximate the finite-sample distribution of T using a permutation procedure.

Let $\{Y_{ij} : i = 1, ..., N; j = 1, ..., n\}$ be the original sample from the BLGC model (1.4) where Y_{ij} is the jth repeated measurement for the ith individual. To approximate the distribution of T, the permutation procedure randomly permutes the individual indices for each fixed j. Under H_0, the individual indices are simply random labels and any permutation of the individual indices is equally likely. Using this invariance property under H_0, the proposed test can be set up by the following steps:

(1) Compute the test statistic T for the original sample, denoted T.
(2) Randomly permute the individual indices while holding fixed the covariates A_i, and compute the test statistic T for this permutation sample.
(3) Repeat the process a large number, say B, of times, giving B test statistics, say $T^b, b = 1, ..., B$.
(4) Compute the proportion of permutation samples with T^b greater than or equal to T_{obs}.
(5) Given the significance level α, if this proportion is greater than α, accept H_0, otherwise reject H_0.

In the next section, we conduct simulation studies to evaluate and compare the efficiency of the proposed test (T-test) with respect to the LR test. We note that the MM estimator (2.3) is not necessarily non-negative definite. In the simulations, if this estimator is not non-negative definite, we replace it with $\hat{D}_*^+ = P\Lambda_+P'$, where P and Λ are, respectively, the matrix of eigenvectors and the diagonal matrix of eigenvalues of $\hat{D}_*$, and also $\Lambda_+ = \max(0, \Lambda)$. $\hat{D}_*^+$ is the closest matrix to $\hat{D}_*$ among all non-negative definite matrices (see Demidenko, 2004, p. 176).

In certain situations, it is of interest to test whether some random effects can be omitted, while keeping others in the model. For example, it may be of interest to test for the need of random slopes for linear time effects in a model with random intercepts and slopes. Suppose D_{*l} is the covariance matrix of the random effects we want to leave out of the model and Z_l is the appropriate columns from the original design matrix Z. Now, by choosing $T = \frac{1}{N}tr\left(Z_l\hat{D}_{*l}Z_l'\right)$ where $\hat{D}_{*l}$ is the appropriate block of the matrix $\hat{D}_*$, and applying the mentioned procedure, the desired test can be carried out.

3. **Simulation Study**

In this section, we summarize simulation studies conducted with the objective of evaluating the behavior of the proposed test. First, the efficiency of the T-test under different distributions for the random effects and errors is examined. Next, we compare the efficiency of the proposed T-test with respect to the LR test.

3.1. **Efficiency of the Proposed Test.** To evaluate the behavior of the T-test, we considered the linear trend model with random parameters

$$\begin{aligned} Y_{ij} &= \xi_i + \eta_i t_{ij} + \varepsilon_{ij} \\ \xi_i &= \alpha + a_i, \quad \eta_i = \beta + b_i \qquad i = 1, ..., N, \;\; j = 1, ..., n, \end{aligned} \tag{3.1}$$

where t_{ij} is the jth observation time for the ith individual, α and β are fixed effects, and a_i and b_i are random intercept and random slope, respectively.

In the simulations, we set $\beta_1 = 0.25$, $\beta_2 = 0.5$ and $t_{ij} = j$, and assumed that $\varepsilon_{ij} \sim N(0, 1)$. For simplicity in simulations, we assumed that the random intercept a_i and the random slope b_i are independent. In practice, this assumption may not be appropriate. We considered three types of distributions, normal, t and log-normal for the two random effects. For normal random effects we assumed $a_i \sim N\left(0, \sigma_1^2\right)$ and $b_i \sim N\left(0, \sigma_2^2\right)$, and for non-normal random effects, we assumed that $a_i \sim \left\{(X_1 - E(X_1)) / \sqrt{var(X_1)}\right\} \times \sigma_1$ and $b_i \sim \left\{(X_2 - E(X_2)) / \sqrt{var(X_2)}\right\} \times$

σ_2, where the random variables X_1 and X_2 are distributed as both t(3) and log-normal(0,1), so that $E\left(a_i\right)=0$, $E\left(b_i\right)=0$, $var(a_i)=\sigma_1^2$, and $var(b_i)=\sigma_2^2$. We generated $1,000$ Monte-Carlo samples under model (3.1), for different numbers of individuals $N=10,15$ and the numbers of repeated measurements $n=3,5$ for each individual and selected $B=1,000$ permutation samples for each setting. In the simulation study, the variance components (σ_1^2,σ_2^2) set to $(0,0)$ (to estimate the size of the test), $(0.02,0.02)$, $(0.05,0.05)$, $(0.1,0.1)$ and $(0.2,0.2)$. The empirical power of the test under each setting evaluated for a significance level of $\alpha=0.05$.

The results, displayed in Table 1, indicate that the Type I error of the T-test is stable across all distributions and is close to the nominal 0.05 level. Furthermore, the power of the test is high even for these small values of N, n, and (σ_1^2,σ_2^2), and increases with (σ_1^2,σ_2^2), as expected. In addition, the results indicate that the test is not liberal with respect to heavy-tailed distributions such as t and asymmetric distributions such as log-normal.

TABLE 1. Rejection rates (expressed as percentages) for the 5% level T-test in the linear trend model with random parameters (3.1), with $\varepsilon_{ij}\sim N(0,1)$.

		$N=10$		$N=15$	
Distribution of (b_{1i},b_{2i})	(σ_1^2,σ_2^2)	$n=3$	$n=5$	$n=3$	$n=5$
(Normal,Normal)	$(0,0)$	5.3	4.6	5.8	4.9
	$(0.02,0.02)$	8.0	21.9	8.6	25.7
	$(0.05,0.05)$	14.9	45.4	17.0	59.0
	$(0.1,0.1)$	25.5	66.4	30.1	81.5
	$(0.2,0.2)$	40.6	80.8	56.0	94.6
(t,t)	$(0,0)$	5.4	4.8	5.2	5.5
	$(0.02,0.02)$	7.6	18.8	10.0	26.6
	$(0.05,0.05)$	12.9	36.1	17.7	49.5
	$(0.1,0.1)$	20.1	56.0	28.5	72.5
	$(0.2,0.2)$	29.2	75.0	43.6	86.2
(Log-Normal,Log-Normal)	$(0,0)$	4.3	5.2	5.2	5.6
	$(0.02,0.02)$	7.8	18.5	9.0	23.3
	$(0.05,0.05)$	13.5	31.7	15.5	42.2
	$(0.1,0.1)$	20.1	47.9	23.6	60.8
	$(0.2,0.2)$	30.4	64.8	42.4	81.1

3.2. **Comparison between the T-test and the LR test.** To compare the two tests, we considered the balanced mixed one-way ANOVA model

$$Y_{ij}=\mu+b_i+\varepsilon_{ij},\qquad i=1,...,N,\quad j=1,...,n, \tag{3.2}$$

where μ is a general mean and b_i is a random effect with $E(b_i)=0$ and $var(b_i)=\sigma_b^2$. For model (3.2), the hypothesis (2.1) becomes $H_0:\sigma_b^2=0$ versus $H_1:\sigma_b^2>0$, and the test statistic (2.4) simplifies to

$$T=\frac{n}{N(N-1)}\sum_{i=1}^{N}\left(Y_{i.}-\bar{Y}_{..}\right)^2-\frac{\hat{\sigma}^2}{N},$$

where $\bar{Y}_{i.} = \sum\limits_{j=1}^{n} Y_{ij}/n$, $\bar{Y}_{..} = \sum\limits_{i=1}^{N} \bar{Y}_{i.}/N$ and

$$\hat{\sigma}^2 = \frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} \left(Y_{ij} - \bar{Y}_{i.}\right)^2.$$

Also, by assuming normal distribution for the random effect and errors, the asymptotic distribution of the LR statistic for the mixed one-way ANOVA model (3.2), when $N \to \infty$, is the mixture (see Stram and Lee, 1994)

$$0.5\chi_0^2 + 0.5\chi_1^2$$

where χ_0^2 is a point mass at 0 and χ_1^2 denotes a chi-square distribution with one degree of freedom.

In the simulations, we fixed $\mu = 2$ and considered three types of random effect distributions, normal, t and log-normal, and assumed that $\varepsilon_{ij} \sim N(0,1)$. For normal random effects we assumed $b_i \sim N\left(0, \sigma_b^2\right)$, and for non-normal random effects, we assumed that $b_i \sim \left\{(X - E(X))/\sqrt{var(X)}\right\} \times \sigma_b$, where the distribution of random variable X is t(3) and log-normal(0,1), so that $var(b_i) = \sigma_b^2$. We generated $1,000$ Monte-Carlo samples under model (3.2), for different numbers of individuals $N = 7, 15, 25, 50, 100$ and the number of repeated measurements $n = 5$ for each individual and selected $B = 1,000$ permutation samples for each setting. First, we set $\sigma_b^2 = 0$ to examine the sizes of the two tests at the significance level of 0.05. The sizes of the two tests are presented in Table 2 alongside the bias and the variance of the estimates of σ_b^2. The results suggest that the size of the T-test is much closer to the nominal 0.05 level than is that of the LR test. Although, the size of the LR test gets closer to the nominal level of 0.05 as the number of individuals, N, gets larger, but it is never greater than 0.035 in any of the simulations configurations.

TABLE 2. Type I error rates (expressed as percentages) of the T-test and LR test for the balanced mixed one-way ANOVA model (3.2), with $n = 5$ and $\varepsilon_{ij} \sim N(0,1)$. The two numbers in the parentheses are, respectively, bias and variance of the estimates of σ_b^2 which have been multiplied by 1000.

Distribution of b_i		$N = 7$	$N = 15$	$N = 25$	$N = 50$	$N = 100$
Normal	LR	$2.2(19-3)$	$1.7(15-1)$	$2.4(11-0)$	$1.8(6-0)$	$3.4(6-0)$
	T	$5.7(45-7)$	$5.1(32-2)$	$4.8(22-1)$	$5.0(18-0)$	$5.1(13-0)$
t	LR	$1.7(24-5)$	$1.7(12-1)$	$2.2(13-1)$	$2.5(6-0)$	$3.5(7-0)$
	T	$5.6(46-6)$	$4.8(32-2)$	$4.9(27-1)$	$5.6(15-0)$	$5.5(12-0)$
Log-Normal	LR	$1.9(35-16)$	$1.1(23-4)$	$1.7(14-2)$	$1.6(9-0)$	$3.2(6-0)$
	T	$5.5(53-7)$	$4.5(34-3)$	$4.3(25-1)$	$5.4(21-0)$	$5.1(13-0)$

Then, according to Fitzmaurice et al. (2007) and our experience in the simulation study, we varied σ_b^2 from 0.02 to 0.1, using smaller value of σ_b^2 in simulation configurations with larger individual numbers, N, to compare the powers of the two tests at the significance level of 0.05. Specifically, we set σ_b^2 equal to 0.1, 0.07, 0.05, 0.04, and 0.02 for N equal to 7, 15, 25, 50, and 100, respectively. The powers of the two tests are presented in Table 3 alongside the bias and the variance of the estimates of σ_b^2. The results suggest that the LR test appears to be more powerful than the T-test for normal random effect but for non-normal random effects the T-test appears to be more powerful than the LR test. Also, it should be noted that increasing N leads to better performance for the LR test in comparison to the T-test.

Overall, the results of the simulations indicate that the proposed T-test has the correct Type I error rate and is more powerful than the LR test for non-normal random effects. Moreover, a power advantage of the proposed T-test with respect to the LR test is that the proposed test can easily be applied for testing for multiple random effects. We note that the LR test can not be used to test for the need of both the random intercept and the random slope in the model (3.1).

TABLE 3. Powers (expressed as percentages) of the T-test and LR test for the balanced mixed one-way ANOVA model (3.2), with $n = 5$ and $\varepsilon_{ij} \sim N(0, 1)$. The two numbers in the parentheses are, respectively, bias and variance of the estimates of σ_b^2 which have been multiplied by 1000.

Distribution of b_i		$N = 7$	$N = 15$	$N = 25$	$N = 50$	$N = 100$
Normal	LR	19.2(18 − 21)	22.0(3 − 7)	18.2(13 − 3)	25.4(5 − 2)	17.7(4 − 0)
	T	17.8(29 − 25)	19.2(9 − 7)	18.6(16 − 4)	24.2(5 − 2)	16.1(6 − 0)
t	LR	11.8(68 − 81)	15.8(34 − 20)	16.4(27 − 133)	20.2(31 − 39)	19.1(6 − 1)
	T	14.2(32 − 113)	16.5(104 − 221)	17.1(16 − 36)	20.0(20 − 27)	16.0(5 − 0)
Log-Normal	LR	13.8(11 − 119)	19.1(43 − 35)	17.9(24 − 13)	22.6(20 − 6)	20.3(2 − 0)
	T	14.1(45 − 156)	19.3(17 − 30)	18.4(15 − 11)	23.0(18 − 5)	17.4(3 − 0)

4. **Real Data Example**

In this section, we consider the well-known growth data introduced by Potthoff and Roy (1964). For this data set, an orthodontic study was conducted on 27 children, 11 girls and 16 boys, all who were 8 years of age at the beginning of the study. On each child, the distance from the center of the pituitary to the pterygomaxillary fissure was measured (in mm) every two years through age 14. The study objectives were to determine if the distances were longer for boys than girls and if the rate of change of distance differed between boys and girls. This data set was analyzed by Fearn (1975), Rao (1987), Lee (1991) and Verbeke and Molenberghs (2000).

Figure 1 presents plot of the distance versus age for boys and girls. Clearly, the profiles differ on the intercept but the slopes are not evidently different. At the outset, we consider a model where the distance varies linearly with age, and with intercept and slope random effects to account for individual-to-individual variation. The balanced linear growth curve model we consider here is

$$Y_{ij} = (\beta_1 + \beta_2 t_j)\, x_i + b_{1i} + b_{2i} t_j + \varepsilon_{ij}, \qquad i = 1, ..., 27,\ j = 1, ..., 4, \tag{4.1}$$

where Y_{ij} is the distance for the ith child at time j, t_j is the jth point in time (in ages) in which the distance was recorded, and x_i is a dummy variable that equals 1 if the ith subject is boy and 0 otherwise. Also, b_{1i} and b_{2i} are random effects with zero mean, and with $var(b_{1i}) = \sigma_1^2$, $var(b_{2i}) = \sigma_2^2$ and $cov(b_{1i}, b_{2i}) = \sigma_{12}$.

The MM estimates of the variance components are $\hat{\sigma}_1^2 = 132.8$, $\hat{\sigma}_2^2 = 0.1$ and $\hat{\sigma}_{12} = 3.2$. To test for the need of both random intercept b_{1i} and random slope b_{2i} in the model (4.1), the proposed test produces a test statistic of 32.71. Based on 1,000 permutations, the p-value of the test is a value of 0.01. Thus, the proposed test rejects the null hypothesis at the 5% nominal level, i.e, it suggests that there are some random effects in the model. The likelihood ratio test can not handle

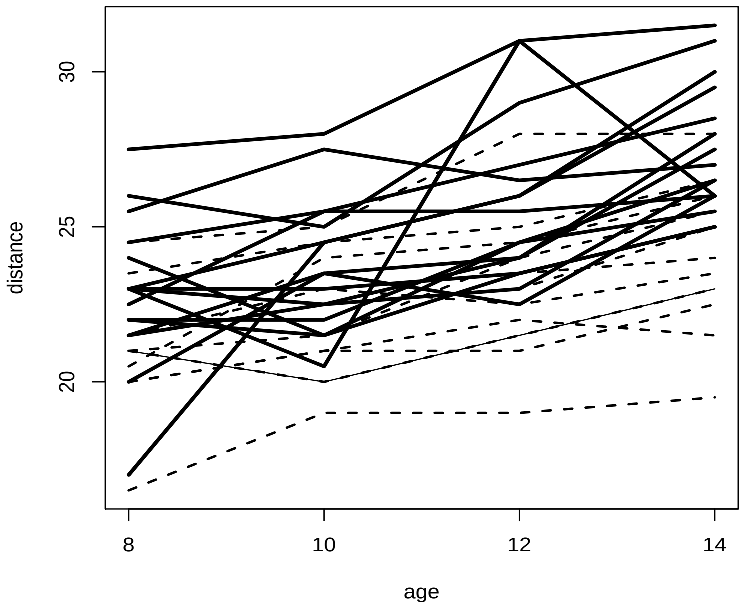

FIGURE 1. The distance versus age for boys and girls. (Boys are indicated with solid lines. Girls are indicated with dashed lines.)

this test. Now, we wish to test for the need of random slope in the model, i.e., we want to test $H_0 : D = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix}$ against $H_1 : D = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$. The proposed test produces a test statistic of 2.33. Based on $1,000$ permutations, the p-value of the test is a value of 0.19. Thus, the proposed test accepts the null hypothesis at the 5% nominal level, i.e, it suggests that there is no random slope in the model. Assuming normal distribution for the random effects and errors, the LR test yields a p-value of 0.2, and hence, this test also accepts the null hypothesis at the 5% nominal level.

We now consider the following reduced model

$$Y_{ij} = (\beta_1 + \beta_2 t_j)\, x_i + b_{1i} + \varepsilon_{ij}, \quad i = 1, ..., 27,\ j = 1, ..., 4.$$

We wish to test $H_0 : \sigma_1^2 = 0$ against $H_1 : \sigma_1^2 > 0$. Note that under the null hypothesis, there is no random intercept in the model. The proposed test assumes a test statistic of 20.08. Based on $1,000$ permutations, the p-value of the test equals 0.02. Also, by assuming normal distribution for the random intercept and errors, the LR test produces a p-value of 0.03. Thus, the two tests reject the null hypothesis at the 5% nominal level. Therefore, we conclude that a random intercept model is appropriate to analyze these data. This conclusion agrees with the findings of Verbeke and Molenberghs (2000).

5. **Concluding Remarks**

In this paper, we have proposed a distribution-free test for testing variance components in balanced linear growth curve models. A permutation procedure has been described to approximate the finite-sample distribution of the test statistic. The proposed test can easily be applied for testing for multiple variance components due to the simple test statistic and the permutation procedure we used. The performance of the proposed test has been illustrated via simulation studies and a real data example. The simulation results suggest that the proposed test has Type I error rate close to the nominal level and is more powerful than the LR test for non-normal random effects. The R codes of this work are available upon request from the correspondent author.

A disadvantage of available tests for variance components is the difficulty in verifying the required regularity conditions as shown in Giampaoli and Singer (2009). The derivation of the proposed T-test is not affected by such difficulties.

Bootstrap tests can also be applied for testing variance components but we note that bootstrap tests give approximated significance levels converging to exact when the number of iterations B tends to infinity. When the labels of observations are exchangeable under the null hypothesis, it is safe and preferable to use a permutation test because permutation tests yield exact significance levels.

For unbalanced linear growth curve models and generally for LME models, we can consider the T-statistic (2.2), and use a permutation procedure for testing variance components. While the benefit of this approach is its distribution-free nature, the only problem is to derive $\hat{D}_*$, because for these models, $\hat{D}_*$ is usually computed by numerical methods which may result a biased estimator of D_*. Since the unbiasedness of $\hat{D}_*$ is very important in this approach, it seems that further research is needed to extend the procedure presented in this paper to LME models. We finally note that in the empirical analysis, growth curves are not always linear, so it would be of interest to generalize this test toward non-linear cases.

Acknowledgments

R. Drikvandi would like to thank Professor M. Asgharian and the Dept. of Math. & Stat., McGill University for their hospitality. His research was partially funded by Prof. Asgharian through his research grant from NSERC, Canada. The authors are grateful to the editor and the referees for their helpful comments which significantly improved the presentation of the paper.

Appendix: **Proof of equation (2.4)**

First, for $\hat{D}_*$ in (2.3), we have

$$
\begin{aligned}
Z\hat{D}_*Z' &= \frac{1}{N-p}\left(Z\hat{S}Z' - \hat{\sigma}^2 Z(Z'Z)^{-1}Z'\sum_{i=1}^{N}\left(1 - q_i{}'Q^{-1}q_i\right)\right)\\
&= \frac{1}{N-p}\left(\sum_{i=1}^{N} Z\left(a_i{}^0 - A_i\hat{\beta}_{OLS}\right)\left(a_i{}^0 - A_i\hat{\beta}_{OLS}\right)'Z' - \hat{\sigma}^2 P_Z Q^*\right)\\
&= \frac{1}{N-p}\left(\sum_{i=1}^{N} Z\left((Z'Z)^{-1}Z'Y_i - A_i\hat{\beta}_{OLS}\right)\left((Z'Z)^{-1}Z'Y_i - A_i\hat{\beta}_{OLS}\right)'Z' - \hat{\sigma}^2 P_Z Q^*\right)\\
&= \frac{1}{N-p}\left(\sum_{i=1}^{N} P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z - \hat{\sigma}^2 P_Z Q^*\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
T &= \frac{1}{N}tr(Z\hat{D}_*Z')\\
&= \frac{1}{N(N-p)}tr\left(\sum_{i=1}^{N} P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z - \hat{\sigma}^2 P_Z Q^*\right)\\
&= \frac{1}{N(N-p)}\left(\sum_{i=1}^{N} tr\left(P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z\right) - k\hat{\sigma}^2 Q^*\right)\\
&= \frac{1}{N(N-p)}\left(\sum_{i=1}^{N} tr\left(\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)\right) - k\hat{\sigma}^2 Q^*\right)\\
&= \frac{1}{N(N-p)}\left(\sum_{i=1}^{N}\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right) - k\hat{\sigma}^2 Q^*\right)\\
&= \frac{1}{N(N-p)}\sum_{i=1}^{N}\left\{\left(Y_i - ZA_i\hat{\beta}_{OLS}\right)'P_Z\left(Y_i - ZA_i\hat{\beta}_{OLS}\right) - \frac{k}{N}\hat{\sigma}^2 Q^*\right\}.
\end{aligned}
$$

Then the equation follows.

References

Dean, C. (1992). Testing for overdispersion in Poisson and binomial regression models. *Journal of the American Statistical Association,* **87**, 451-457.

Demidenko, E. (2004). *Mixed Models: Theory and Applications.* New York: Wiley.

Demidenko, E. and Stukel, T. (2002). Efficient estimation of general linear mixed effects models. *Journal of Statistical Planning and Inference,* **104**, 197-219.

Fearn, T. (1975). A Bayesian approach to growth curves. *Biometrika,* **62**, 89-100.

Fitzmaurice, G. M., Lipsitz, S. R. and Ibrahim, J. G. (2007). A note on permutation tests for variance components in multilevel generalized linear mixed models.

Biometrics, **63**, 942-946.

Giampaoli, V. and Singer, J. M. (2009). Likelihood ratio tests for variance components in linear mixed models. *Journal of Statistical Planning and Inference,* **139**, 1435-1448.

Gueorguieva, R. (2001). A multivariate generalized linear mixed model for joint modelling of clustered outcomes in the exponential family. *Statistical Modelling,* **1**, 177-193.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics,* **38**, 963-974.

Lee, J. C. (1991). Tests and model selection for the general growth curve model. *Biometrics,* **47**, 147-159.

Lin, X. (1997). Variance component testing in generalised linear models with random effects. *Biometrika,* **84**, 309-326.

Miller, J. J. (1977). Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. *Annals of Statistics,* **5**, 746-762.

Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika,* **51**, 313-326.

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika,* **52**, 447-458.

Rao, C. R. (1987). Prediction of future observations in growth curve models. *Statistical Science,* **2**, 434-471.

Reinsel, G. C. (1985). Mean squared error properties of empirical Bayes estimators in a multivariate random effects general linear model. *Journal of the American Statistical Association,* **80**, 642-650.

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.

Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics,* **50**, 1171-1177.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data.* Berlin: Springer.

Verbeke, G. and Molenberghs, G. (2003). The use of score tests for inference on variance components. *Biometrics,* **59**, 254-262.

von Rosen, D. (1991). The growth curve model: a review. *Communications in Statistics- theory and Methods,* **20**, 2791-2822.