Tatra Mt. Math. Publ. **39** (2008), 17–25

**Tatra Mountains**
*Mathematical Publications*

# COMPARISON OF APPROXIMATE TESTS OF FIXED EFFECTS IN LINEAR REPEATED MEASURES DESIGN MODELS WITH COVARIATES

Júlia Volaufová — Lynn R. LaMotte

ABSTRACT. The topic of testing linear hypotheses about parameters of fixed effects in models with variance-covariance components has been investigated extensively in the past decades. The main question is in determination of the degrees of freedom of an approximate $F$-test. The approximation is based on the choice of the estimate of the approximate variance-covariance matrix of the estimator of the fixed effect parameters. Various approximations have been suggested in the literature and some have already been implemented in statistical packages, such as the generalized Satterthwaite approximation of degrees of freedom, sandwich estimator, the Harville-Jeske-Kenward-Roger approximation, to name a few. For repeated measures designs there are some possibilities of modeling the covariance matrix as well as different choices of approximations. There are still open questions about how the different choices, either of the covariance structure or of the approximating distribution of the test statistic, affect the size (and power) of the tests. Here the sizes of different options of tests are determined and compared by a simulation study.

## 1. Introduction

A special case of mixed linear model, the so-called repeated measures design model, enjoys a special popularity and special treatment among investigators dealing with replicated observations on sampling units. In such cases the dependences between individual observations on the same unit cannot be ignored. Particular attention is paid to replications that are carried out over time. In such experiments, the interest is mostly in investigating the effect of intervention (treatment), of time, and possibly of the treatment-by-time interaction. Often the response is associated with a set of variables that are included in the

2000 Mathematics Subject Classification: Primary 62J10, 62F10.
Keywords: repeated measures design model, empirical BLUE, approximate $F$-test.

model and are referred to as covariates. In most cases, the covariates are measured only once at the beginning of the experiment, i.e., they are considered not to be time dependent.

In general, the model considered here can be expressed as follows. Denote by $Y_{i(j)}$ the $T_{i(j)}$ dimensional observation vector corresponding to the $i$th sampling unit in the $j$th (treatment) group. Then $Y_{i(j)}$ can be modeled by

$$Y_{i(j)} = (X_1, X_2)_{i(j)}\beta + \epsilon_{i(j)}, \tag{1}$$

where $E\big(\epsilon_{i(j)}\big) = 0$ and the covariance matrix of $Y_{i(j)}$ is denoted by $\operatorname{cov}\epsilon_{i(j)} = R_{i(j)}$. Then, when combining all observations on all sampling units, we create an $N$ dimensional vector of observations $Y$, for which we shall assume the following model.

$$E(Y) = (X_1 : X_2)\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \qquad \operatorname{cov}(Y) = V(\vartheta), \tag{2}$$

where $Y$ is the $N \times 1$ random vector of observations. Here we shall consider the case where $Y$ follows a multivariate normal distribution. The matrix $(X_1 : X_2)$ is a known $N \times q$ matrix partitioned into two parts $X_1$ and $X_2$. $X_1$ is the $N \times q_1$ matrix containing the observed covariates as columns, and $X_2$ is the $N \times q_2$ design matrix corresponding to a factorial treatment structure of the trial. Since observations on different sampling units are independent, the covariance matrix $V(\vartheta)$ has a block diagonal structure, with $R_{i(j)}$, the $T_{i(j)}$ dimensional blocks on the diagonal. It depends on an unknown $r$ dimensional vector parameter $\vartheta \in \Theta \subset R^r$, such that $V(\vartheta)$ is positive definite for all $\vartheta \in \Theta$. There are several possible ways of modeling the covariance matrix. Here are just three most widely used examples of types of structures of covariance matrices.

A. Compound symmetry (CS). The assumed structure is $R_{i(j)} = \vartheta_1 \mathbf{1}_{T_{i(j)}} \mathbf{1}'_{T_{i(j)}} + \vartheta_2 I_{T_{i(j)}}$, $\vartheta_2 > 0$, $\mathbf{1}_\ell$ being an $\ell$-vector of ones. The parameter $\vartheta_1$ denotes the common covariance between any two observations on the same sampling unit, and $\vartheta_2$ denotes the error variance.

B. Autoregressive type of dependencies (AR(1)). Here, $R_{i(j)}$ is a matrix with entries $R_{i(j)_{l,m}} = \sigma^2\rho^{|l-m|}$, with $0 \le \rho \le 1$ and $\sigma^2 > 0$, $l, m = 1, 2, \dots, T_{i(j)}$.

C. The most general is the case when all entries of the matrix $R_{i(j)}$ (and hence of $V(\theta)$) are unknown and we do not assume any particular structure of $R_{i(j)}$, other than that it is positive definite. In general, the $l, m$th entry of $R_{i(j)}$ is assumed to be $\sigma_{l,m}$. We refer to this case as unstructured.

## 2. Testing $P'\beta_2 = 0$

In the setting of model (2), we are interested in testing a linear estimable hypothesis $H_0 : P'\beta_2 = 0$ for a given $q_2 \times k$ full rank matrix $P$, i.e., $r(P) = k$. Let $F$ be a full column rank matrix for which $X_1'F = 0$. The unbiased estimability of $P'\beta_2$ is equivalent to the condition $P = X_2'FQ$ for some matrix $Q$.

If $V(\vartheta)$ is known, then the test statistic is based on the best linear unbiased estimator (BLUE), $\widehat{P'\beta_2}$, yielding $\widehat{\beta_2'P}\big[\mathrm{cov}(\widehat{P'\beta_2})\big]^{-1}\widehat{P'\beta_2}$ that under $H_0$ has a $\chi^2$ distribution with degrees of freedom equal to $k$. Let $W(\vartheta) = F'V(\theta)F$. Notice that here $\widehat{P'\beta_2} = P'\big(X_2'FW(\vartheta)^{-1}F'X_2\big)^{-}X_2'FW(\vartheta)^{-1}F'Y$, which is invariant with respect to the choice of $F$. It is an easy exercise to show that $FW(\vartheta)^{-1}F' = \big(M_{X_1}V(\vartheta)M_{X_1}\big)^{+}$ and hence that $\widehat{P'\beta_2}$ can be also expressed as $P'\big(X_2'(M_{X_1}V(\vartheta)M_{X_1})^{+}X_2\big)^{-}X_2'\big(M_{X_1}V(\vartheta)M_{X_1}\big)^{+}Y$, where $M_{X_1} = I - X_1X_1^{+}$ and where "+" denotes the Moore-Penrose generalized inverse. (See, e.g., [5].) Its covariance matrix $\mathrm{cov}\big(\widehat{P'\beta_2}\big)$, denoted by $C(\vartheta)$, can be expressed as

$$\begin{aligned} C(\vartheta) &= P'\left[X_2'FW(\vartheta)^{-1}F'X_2\right]^{-}P \\ &= P'\left[X_2'(M_{X_1}V(\vartheta)M_{X_1})^{+}X_2\right]^{-}P. \end{aligned} \tag{3}$$

In the case that $V(\vartheta)$ is proportional to a known matrix, say $V(\vartheta) = \theta V$, with $\theta > 0$ unknown, then the test is based on $\widehat{\beta_2'P}\big[\widehat{\mathrm{cov}}(\widehat{P'\beta_2})\big]^{-1}\widehat{P'\beta_2}$, where $\widehat{\mathrm{cov}}(\widehat{P'\beta_2}) = \hat{\theta}P'\big[X_2'(M_{X_1}VM_{X_1})^{+}X_2\big]^{-}P$. Here, $\hat{\theta}$ is chosen to be the minimum variance unbiased invariant estimator of $\theta$. Under our assumptions and under $H_0$,

$$F = \frac{1}{k}\widehat{\beta_2'P}\Big[\widehat{\mathrm{cov}}\big(\widehat{P'\beta_2}\big)\Big]^{-1}\widehat{P'\beta_2} \tag{4}$$

has an $F$-distribution with $N - r(X_1, X_2)$ degrees of freedom.

When $V(\vartheta)$ depends on unknown $\vartheta \in \Theta \subset R^r$, $r > 1$, the situation is more complicated. From now on let $\hat{\vartheta}$ denote the REML estimator of $\vartheta$ based on the full model (2). After plugging the estimates $\hat{\vartheta}$ into $V(\vartheta)$ and substituting $V(\hat{\vartheta})$ in the expression for $\widehat{P'\beta_2}$, we get an estimator denoted by $\widetilde{P'\beta_2}$ that is often referred to as *the empirical* BLUE (EBLUE) (see, e.g., [3], [2].) In spite of being nonlinear in $Y$, under normality, EBLUE remains unbiased for $P'\beta_2$, since the estimator of $\vartheta$ is an even translation invariant function of $Y$, as shown by several authors in a more general setting (see e.g., [2]). It is suggestive that even in this complex case, the test statistic might be approached through

$$\left[\widetilde{P'\beta_2}\right]'\left[\mathrm{cov}(\widetilde{P'\beta_2})\right]^{-}\widetilde{P'\beta_2}\,. \tag{5}$$

The problem here is that $\operatorname{cov}(\widetilde{P'\beta_2})$, even in some simple settings under model (2), in addition to the fact that it depends on unknown $\theta$, does not have a closed form. Instead, various approximations are suggested that still depend on $\vartheta$, and hence the estimated version of the approximation has to be used. The approximations considered by various authors are based on a decomposition

$$\operatorname{cov}(\widetilde{P'\beta_2}) = C(\vartheta) + B(\vartheta) \tag{6}$$

(see, e.g., [3], [6], [2], [4]).

In the particular case of our model (2), we get the approximation of $B(\vartheta)$ as follows. Denote by $X_2^* = F'X_2$ and by $S$ the (asymptotic) covariance matrix of the REML estimators of $\vartheta$ with entries $S_{ij}$, $i,j = 1,2,\dots,r$. Let $P_{X_2^*} = X_2^*(X_2^{*\prime}W(\vartheta)^{-1}X_2^*)^{-}X_2^{*\prime}W(\vartheta)^{-1}$. Then in our particular case, the approximation $B_A(\vartheta)$ of $B(\vartheta)$ is given by

$$B_A(\vartheta) = \sum_{i,j}^{r} S_{ij} Q' P_{X_2^*} \frac{\partial W(\vartheta)}{\partial \vartheta_i}(I - P'_{X_2^*})W(\vartheta)^{-1}\frac{\partial W(\vartheta)}{\partial \vartheta_j} P'_{X_2^*} Q\,. \tag{7}$$

(See and compare also with, e.g., Kenward and Roger [4].)

The empirical version of the asymptotic covariance matrix of $\widetilde{P'\beta_2}$, denoted by $C(\hat\vartheta)$, takes the form

$$C(\hat\vartheta) = P'\left[X_2'W(\hat\vartheta)^{+}X_2\right]^{-} P. \tag{8}$$

(See, e.g., [3], [6].) As pointed out by Harville and Jeske in [2] and subsequently by Kenward and Roger in [4], for the estimate of the approximated covariance matrix of $\widetilde{P'\beta_2}$, a correction term is needed that is exactly twice the empirical version of $B_A(\vartheta)$. The estimate of the approximated covariance matrix of $\widetilde{P'\beta_2}$ is then given by

$$\widehat{\operatorname{cov}}_A(\widetilde{P'\beta_2}) = C(\hat\vartheta) + 2B_A(\hat\vartheta), \tag{9}$$

which in our model takes the form

$$\begin{aligned}&\widehat{\operatorname{cov}}_A(\widetilde{P'\beta_2})\\ &= Q'\hat P_{X_2^*}\left[W(\hat\vartheta) + 2\sum_{ij}^{r} S_{ij}\frac{\partial W(\hat\vartheta)}{\partial\vartheta_i}(I - \hat P'_{X_2^*})W(\hat\vartheta)^{-1}\frac{\partial W(\hat\vartheta)}{\partial\vartheta_j}\right]\hat P'_{X_2^*}Q\,. \end{aligned}\tag{10}$$

Here

$$\frac{\partial W(\hat\vartheta)}{\partial\vartheta_j} = \left.\frac{\partial W(\vartheta)}{\partial\vartheta_j}\right|_{\vartheta=\hat\vartheta} \quad \text{and} \quad \hat P_{X_2^*} \text{ is } P_{X_2^*} \text{ at } \hat\vartheta.$$

At this point there are several choices for creating the basis for the test statistic. In each case we create the empirical version of the quadratic form by plugging in the REML estimates for $\vartheta$. The options we shall investigate here are listed below.

1. $F_1 = \frac{1}{k}\left[\widetilde{P'\beta_2}\right]' C(\hat{\vartheta})^- \widetilde{P'\beta_2}\,,$ (11)

   emphasizing the simplicity of computations of $C(\hat{\vartheta})$. Under $H_0$, $F_1$ is approximated by an $F$-distribution with $k$ and $\nu$ degrees of freedom. $\nu$ may be determined possibly in two ways.

   a) One possibility is to use the generalized Satterthwaite method (see, e.g., [1]).
   b) he other is the so-called between-within method of calculating the degrees of freedom, which is the default for a repeated measures design analysis in the statistical package SAS in the procedure PROC MIXED.

2. $F_3 = \lambda F_2\,,$ (12)

   where

   $$F_2 = \frac{1}{k}\left[\widetilde{P'\beta_2}\right]' \left[C(\hat{\vartheta}) + 2B_A(\hat{\vartheta})\right]^- \widetilde{P'\beta_2}\,, \tag{13}$$

   considering the suggested estimate of the approximation with the additional correction matrix $B_A(\hat{\vartheta})$. Kenward and Roger in [4] derived the proportionality coefficient $\lambda$ as a function of $\hat{\vartheta}$ as well as the approximate denominator degrees of freedom, say $\nu(\hat{\vartheta})$, so that the moments of $F_3$ match the moments of the $F$-distribution with $k$ and $\nu(\hat{\vartheta})$ degrees of freedom.

## 3. Simulation study

A simulation study was carried out in order to investigate two phenomena, the effect of the choice of the structure of the covariance matrix in modeling dependencies between observations on the same sampling unit, and the effect of the choice of type of approximate test of the hypothesis on the size of the test. We focus here on the behavior of the $p$-value (size of the test) and its accuracy (validity). For sampling units and covariates, the NHANES(2004) data set was used with 8192 distinct records as the basic source. Equal numbers of males and females were randomly selected with records of age, height and weight that are considered as covariates in the model and constitute the matrix $X_1$. Two levels of sample sizes were considered, $n = 6$ per treatment group and $n = 12$ per treatment group. There were three treatment groups considered, simulating a weight-loss study, starting from the weight from the record of each sampling unit. For each subject, repeated measures (up to 4, but with occasional missing values) were generated with compound symmetric (CS) variance-covariance structure.

The intraclass correlation, $\rho$, was set to four different levels, 0.1, 0.3, 0.6, and 0.9. The error variance $\sigma_e^2$ was set to 4. In the simulated models, there were no treatment effects and no time-by-treatment interaction effects. For each above mentioned configuration, there were 1000 simulations carried out and three different approximations of the test of the hypothesis of no treatment effect, corresponding to $F_1$ with both methods of calculating $\nu$ as indicated in a) and b) above and to $F_3$. Although the generated data all had CS structure, when performing the tests, we ran some simulations (Figures 1–3) in which the correct (CS) covariance structure was used, and some (Figures 4 and 5) in which the incorrect covariance structure was used (we used AR(1) for this purpose).

In the Figures 1–5, we plot the observed $p$-values on the vertical axis versus the empirical cumulative distribution function of the $p$-value on the horizontal axis from the 1000 simulations. Each panel corresponds to a particular approximation of the test. Since the range of interest for the size of the test is mainly bounded by 10%, the graphs show only the range from 0 to 0.1.

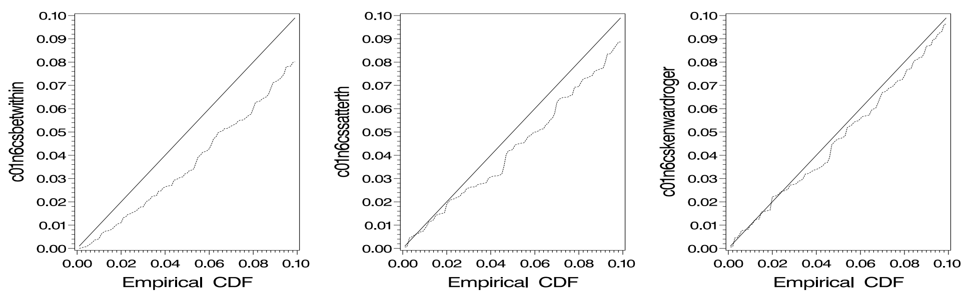


FIGURE 1. Observed $p$-values vs. empirical CDF of $p$-values for $\rho = 0.1$, $n = 6$ per group and the right choice of CS.

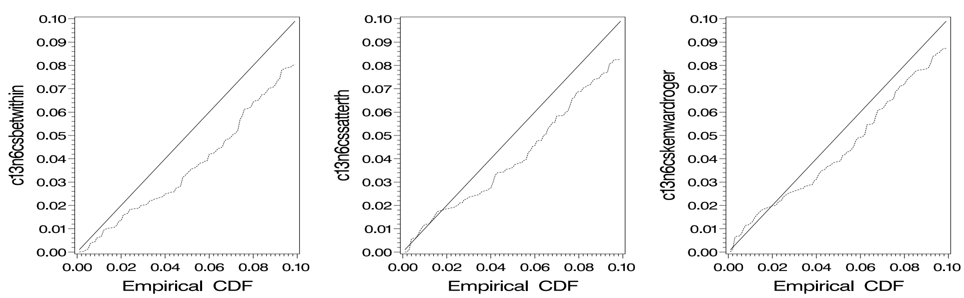


FIGURE 2. Observed $p$-values vs. empirical CDF of $p$-values for $\rho = 0.3$, $n = 6$ per group and the right choice of CS.

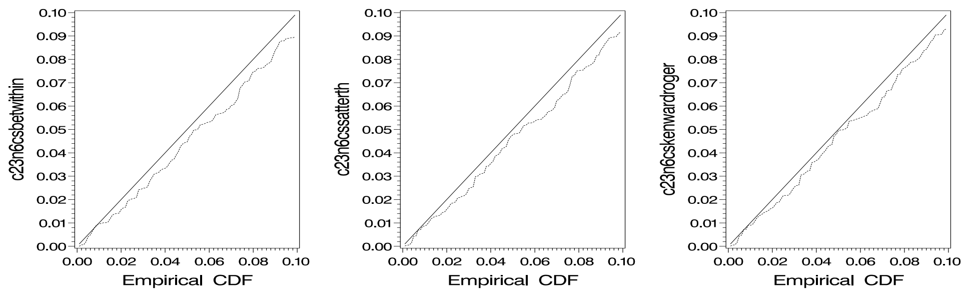


FIGURE 3. Observed $p$-values vs. empirical CDF of $p$-values for $\rho = 0.6$, $n = 6$ per group and the right choice of CS.

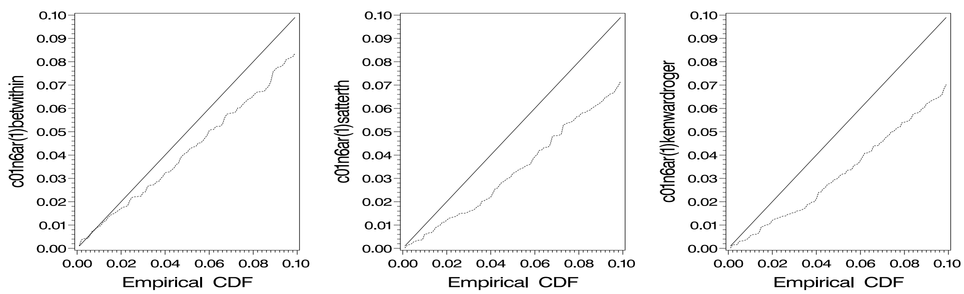


FIGURE 4. Observed $p$-values vs. empirical CDF, n= 6/group, $\rho = 0.1$, AR(1) chosen instead of the true CS.

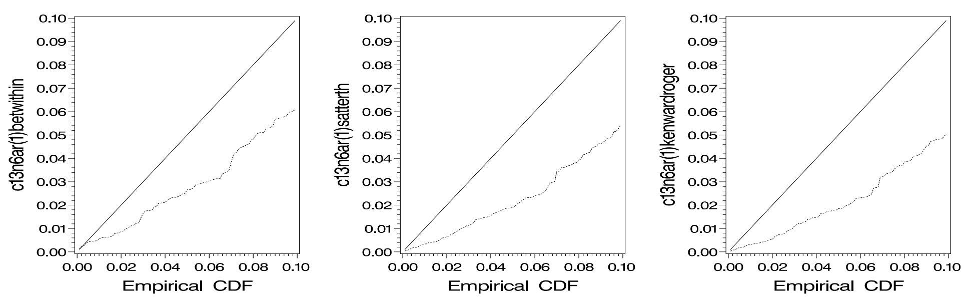


FIGURE 5. Observed $p$-values vs. empirical CDF, n=6/group, $\rho = 0.3$, AR(1) chosen instead of the true CS.

## 4. Conclusions

Shown here are results with CS as the correct, and AR(1) as the incorrect covariance structure. We also ran simulations examining unstructured as the

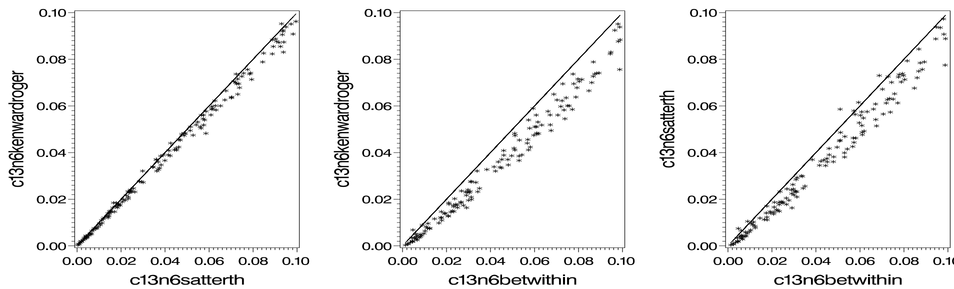


FIGURE 6. Scatter plots of $p$-values, AR(1) instead of true CS, n=6/group, $\rho = 0.3$ comparing methods.

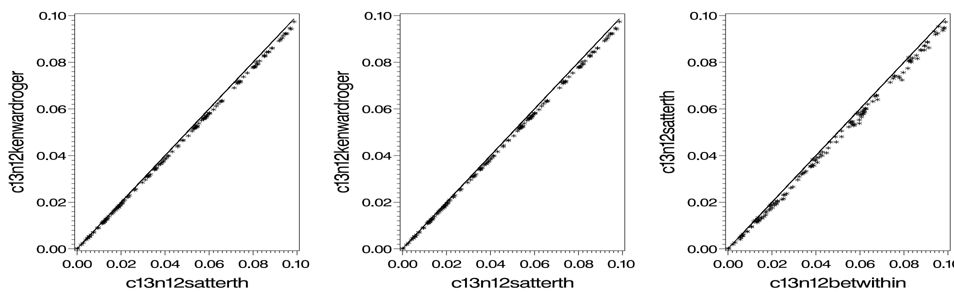


FIGURE 7. Scatter plots of $p$-values, AR(1) instead of true CS, n=12/group, $\rho = 0.3$ comparing methods.

incorrect covariance structure. We also generated data with AR(1) as the correct and CS and unstructured as incorrect covariance structure. In each case the results were similar to results reported here.

It seems that the major effect on the size of the test and consequently on the accuracy of the $p$-value is the choice of the (wrong) covariance matrix. This effect depends further on the tightness of dependencies between observations on the same sampling unit.

The different types of approximations do not seem to have a meaningful effect on the size of considered tests except when the choice of the covariance matrix happens to be the right one (that, in practice, we never have a way of verifying). See Figures 1–5.

The different types of approximate tests behave very similarly to each other. In most of the cases, the approximate tests are anti-conservative. This is a serious phenomenon since the observed $p$-values appear to be smaller than in fact they should be, which may lead to unwarranted rejections and inflating the probability of Type I error.

Figures 6 and 7 show $p$-values computed with incorrect (AR(1) instead of CS) covariance structure for different approximations ($F_3$, $F_1$ with a) and $F_1$ with b) ). These scatterplots illustrate that for larger sample size, it does not make a difference what type of approximation is used, since all behave in the same way. More research is needed in this area to make generally valid recommendations.

## REFERENCES

[1] FAI, A. H. T.—CORNELIUS, P. L.: *Approximate $F$-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments*, J. Stat. Comput. Simul. **54** (1996), 363–378.

[2] HARVILLE, D. A.—JESKE, D. R.: *Mean squared error of estimation or prediction under a general linear model*, J. Amer. Statist. Assoc. **87** (1992), 724-731.

[3] KACKAR, R. N.—HARVILLE, D. A.: *Approximations for standard errors of estimators of fixed and random effects in mixed linear models*, J. Amer. Statist. Assoc. **79** (1984), 853-862.

[4] KENWARD, M. G.—ROGER, J. H.: *Small sample inference for fixed effects from restricted maximum likelihood*, Biometrics **53** (1997), 983-997.

[5] SEARLE, S. R.—CASELLA, G.—MCCULLOCH, C. M.: *Variance Components*. Wiley & Sons, New York, 1992.

[6] VOLAUFOVÁ, J.: *On the variance of two-stage estimator in variance-covariance components model*, Applications of Mathematics **38** (1993), 1–9.

Received October 16, 2006

*Biostatistics Program*
*LSUHSC School of Public Health*
*1615 Poydras St., Suite 1400*
*New Orleans, LA 70112*
*U.S.A*

*E-mail*: jvolau@lsuhsc.edu
llamot@lsuhsc.edu