Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Transformations for Estimation of Linear Models with Nested-Error Structure

Wayne A. Fuller & George E. Battese

To cite this article: Wayne A. Fuller & George E. Battese (1973) Transformations for Estimation of Linear Models with Nested-Error Structure, Journal of the American Statistical Association, 68:343, 626-632

To link to this article: https://doi.org/10.1080/01621459.1973.10481396

Published online: 05 Apr 2012.

Transformations for Estimation of Linear Models with Nested-Error Structure

WAYNE A. FULLER and GEORGE E. BATTESE*

Two linear models with error structure of the nested type are considered. Transformations are presented by which uncorrelated errors with constant variances are obtained. The transformed observations are differences between the original observations and multiples of averages of subsets of the observations. The transformations permit the calculation of the generalized least-squares estimators and their covariance matrices by ordinary least-squares regression. Regression-type estimators are presented for use when the variance components are unknown. Sufficient conditions are presented under which the estimated generalized least-squares estimator is unbiased and asymptotically equivalent to the generalized least-squares estimator.

1. INTRODUCTION

For the linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{1.1}$$

where

- $\mathbf{X}$ is an $(n \times p)$ matrix of fixed constants of rank p;
- β is a vector of p unknown parameters;
- $E(\mathbf{u}) = 0$; and
- $E(\mathbf{uu}') = \mathbf{V}$, where $\mathbf{V}$ is positive-definite,

it is well known that the best-linear-unbiased estimator for $\boldsymbol{\beta}$ is the generalized least-squares estimator

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} \tag{1.2}$$

which has covariance matrix

$$\operatorname{cov}(\tilde{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{1.3}$$

Except for special forms of the covariance matrix $\mathbf{V}$, the generalized least-squares estimates and their variances and covariances are not obtained by the ordinary least-squares regression of $\mathbf{Y}$ on $\mathbf{X}$. To obtain the generalized least-squares estimates by the matrix operations suggested by (1.2) may not be computationally convenient. With a large number of observations on the variables in a linear model, it is possible that the storage required by the available programs exceeds that of the computer facilities.

If, however, a transformation matrix $\mathbf{T}$ can be found such that the transformed errors

$$\mathbf{u}^* = \mathbf{Tu} \tag{1.4}$$

are uncorrelated with constant variances, the generalized least-squares estimator (1.2) is given by the regression of the transformed dependent variable $\mathbf{TY}$ on the transformed independent variables $\mathbf{TX}$. Further, the statistics obtained by an ordinary least-squares regression program are appropriate for hypothesis testing. Numerous ways are available for constructing the $\mathbf{T}$ matrix of (1.4). The utility of any particular transformation depends on its computational convenience. Lemma 1 gives a procedure for constructing the matrix $\mathbf{V}^{-\frac{1}{2}}$ which satisfies the relationship

$$\mathbf{V}^{-\frac{1}{2}}\mathbf{V}\mathbf{V}^{-\frac{1}{2}} = \mathbf{I}_n.$$

With use of the transformation matrix $\mathbf{T} = \mathbf{V}^{-\frac{1}{2}}$, the transformed errors $\mathbf{Tu}$ are uncorrelated with variances equal to one.

Lemma 1: If $\mathbf{V}$ is an $(n \times n)$ positive definite symmetric matrix with distinct characteristic roots $\lambda_1, \lambda_2, \cdots, \lambda_r$ having multiplicity $m_1, m_2, \cdots, m_r$, respectively, where $\sum_{i=1}^r m_i = n$, then $\mathbf{V}$ can be expressed as $\mathbf{V} = \sum_{i=1}^r \lambda_i \mathbf{A}_i$, where the $\mathbf{A}_i$ are mutually orthogonal, symmetric idempotent matrices which satisfy the conditions $\mathbf{VA}_i = \lambda_i \mathbf{A}_i$, $i = 1, 2, \cdots, r$. The matrix $\mathbf{V}^{-\frac{1}{2}}$ defined by

$$\mathbf{V}^{-\frac{1}{2}} = \textstyle\sum_{i=1}^r \lambda_i^{-\frac{1}{2}}\mathbf{A}_i$$

satisfies the relationship $\mathbf{V}^{-\frac{1}{2}}\mathbf{V}\mathbf{V}^{-\frac{1}{2}} = \mathbf{I}_n$.

Proof: Let $\mathbf{B}_i$ denote an $(n \times m_i)$ matrix of m_i mutually orthogonal, normalized characteristic vectors associated with λ_i and let the $(n \times n)$ matrix $\mathbf{A}_i$ be defined by $\mathbf{A}_i = \mathbf{B}_i\mathbf{B}_i'$, $i = 1, 2, \cdots, r$. It follows that $\mathbf{VA}_i = \lambda_i\mathbf{A}_i$ since $\mathbf{B}_i$ satisfies the condition $\mathbf{VB}_i = \lambda_i\mathbf{B}_i$. The $\mathbf{A}_i$ are mutually orthogonal, symmetric idempotent matrices and the matrix $\sum_{i=1}^r \mathbf{A}_i$ is thus idempotent and has rank, the sum of the ranks of the $\mathbf{A}_i$ (e.g., see Theorems 1.68 and 1.69 of [5]). Hence the matrix $\sum_{i=1}^r \mathbf{A}_i$ is the identity matrix of order n. Therefore, the matrix $\mathbf{Q} \equiv \sum_{i=1}^r \lambda_i\mathbf{A}_i$ has inverse $\mathbf{Q}^{-1} = \sum_{i=1}^r \lambda_i^{-1}\mathbf{A}_i$. Since

$$\mathbf{VQ}^{-1} = \textstyle\sum_{i=1}^r \lambda_i^{-1}\mathbf{VA}_i = \sum_{i=1}^r \lambda_i^{-1}(\lambda_i\mathbf{A}_i) = \mathbf{I}_n,$$

the matrix $\sum_{i=1}^r \lambda_i\mathbf{A}_i$ is equal to $\mathbf{V}$. Further,

$$\mathbf{V}^{-\frac{1}{2}} \equiv \textstyle\sum_{i=1}^r \lambda_i^{-\frac{1}{2}}\mathbf{A}_i$$

satisfies the condition $\mathbf{V}^{-\frac{1}{2}}\mathbf{V}\mathbf{V}^{-\frac{1}{2}} = \mathbf{I}_n$.

* Wayne A. Fuller is professor and George E. Battese is assistant professor, both with the Department of Statistics at Iowa State University of Science and Technology, Ames, Iowa 50010. This research was partly supported by a Joint Statistical Agreement with the U.S. Bureau of the Census, J.S.A. 72-4. The authors wish to thank the referee who suggested the inclusion of Lemma 1 and the associate editor who suggested the compact proof of Theorem 2.

Journal Paper No. J-7463 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 1806.

September 1973, Volume 68, Number 343
Theory and Methods Section

For linear models with nested-error structure the decomposition $\mathbf{V} = \sum_{i=1}^{r} \lambda_i \mathbf{A}_i$ is easily obtained. We suggest a transformation matrix that is a multiple of the matrix $\mathbf{V}^{-\frac{1}{2}}$. The transformation has the advantage that it can be performed by relatively simple arithmetic operations.

In most practical problems, the nonzero elements of the matrix $\mathbf{V}$ are unknown and must be estimated from sample data. If the estimated covariance matrix is denoted by $\hat{\mathbf{V}}$, we obtain the values of the estimated, generalized least-squares estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Y} \tag{1.5}$$

by the ordinary least-squares regression of $\hat{\mathbf{T}}\mathbf{Y}$ on $\hat{\mathbf{T}}\mathbf{X}$, where $\hat{\mathbf{T}}$ is the estimate for the transformation matrix $\mathbf{T}$ obtained from the elements of $\hat{\mathbf{V}}$.

2. THE ONE-FOLD, NESTED-ERROR MODEL

2.1 Model Presentation

Frequently data arise from the random selection of "individuals" for which several "measurements" are made. For example, a sample of women may be selected and height measurements taken for the individuals over several years in a study of the relationship between height and age of adult women. This "nesting" pattern, by which the data are generated, has, in general, a significant bearing on the statistical model that is appropriate for satisfactory analyses of the data.

In the presentation of the statistical model for the analysis of observations that arise in a one-fold, nested structure, we denote the variable under study by the letter y with two subscripts. The first subscript distinguishes the "individual" in the sample, and the second subscript distinguishes the "measurement" for the particular individual. We assume that t "individuals" are selected at random and that n_i "measurements" are made on the ith individual. We do not require that the number of measurements be the same for all individuals. The linear model is expressed as

$$y_{ij} = \textstyle\sum_{k=1}^{p} x_{ijk}\beta_k + u_{ij},$$
$$j = 1, \cdots, n_i; \quad i = 1, \cdots, t, \tag{2.1}$$

and

$$u_{ij} = v_i + e_{ij} \tag{2.2}$$

where

y_{ij} denotes the value of the jth measurement for the ith individual;

x_{ijk}, $k = 1, \cdots, p$, denote the levels of the p control variables at which the observation y_{ij} is obtained (the x_{ijk} are assumed fixed constants);

β_k, $k = 1, \cdots, p$, denote the unknown parameters to be estimated; and

u_{ij}, the random error associated with y_{ij}, is assumed the sum of the random effect associated with the ith sample individual (v_i) and the random effect associated with the jth measurement for the ith individual in the sample (e_{ij}).

The random errors v_i and e_{ij} are assumed independently distributed with zero means and variances σ_v^2 and σ_e^2, respectively, where $\sigma_v^2 \geq 0$ and $\sigma_e^2 > 0$. The covariance structure for the random errors u_{ij}, defined by (2.2), is thus expressed by

$$\begin{aligned} E(u_{ij}u_{i'j'}) &= \sigma_v^2 + \sigma_e^2, && \text{if } i = i', \quad j = j' \\ &= \sigma_v^2, && \text{if } i = i', \quad j \neq j' \\ &= 0, && \text{if } i \neq i'. \end{aligned} \tag{2.3}$$

Subsamples from primary sampling units and split-plot experiments are classical examples of situations in which the error model (2.2) may arise. Thus, for these examples, t is interpreted as the number of primary sampling units in the sample or the number of whole-plots in the experiment, and n_i is interpreted as the number of elements subsampled in the ith primary sampling unit or the number of subplots in the ith whole-plot. In the split-plot, experimental-design situation, the number of subplots in each whole-plot is generally the same for all whole-plots.

2.2 Transformation for the One-Fold, Nested-Error Model

We suggest the transformation of (2.1) into the regression equation

$$y_{ij} - \alpha_i \bar{y}_{i.} = \textstyle\sum_{k=1}^{p} (x_{ijk} - \alpha_i \bar{x}_{i.k})\beta_k + u_{ij}^* \tag{2.4}$$

where

$$\alpha_i = 1 - [\sigma_e^2/(\sigma_e^2 + n_i\sigma_v^2)]^{\frac{1}{2}} \tag{2.4a}$$

and $\bar{y}_{i.}$, $\bar{x}_{i.k}$, $k = 1, 2, \cdots, p$, denote the averages of the n_i y- and x-measurements on the ith individual. In Theorem 1 we prove that the errors, u_{ij}^*, are uncorrelated and have variances σ_e^2.

This transformation was suggested by Fuller in 1965 and presented in [11, p. 95]. It has been applied in several agronomic studies (e.g., [2, 12]) and in a nutritional study of individuals over time [4].

Theorem 1: For the linear model (2.1)–(2.3), the transformed errors, $u_{ij}^* \equiv u_{ij} - \alpha_i \bar{u}_{i.}$, where

$$\bar{u}_{i.} = \textstyle\sum_{j=1}^{n_i} u_{ij}/n_i,$$

are uncorrelated and have variances σ_e^2.

Proof: The result can be proved by direct evaluation of the expectations. An alternative approach is to express the model in matrix notation and show that the transformed model (2.4) is obtained with use of the results of Lemma 1. We write the linear model (2.1) as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{2.5}$$

where $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \cdots, \mathbf{Y}_t')$, $\mathbf{Y}_i' = (y_{i1}, y_{i2}, \cdots, y_{in_i})$ and the $\mathbf{X}$ and $\mathbf{u}$ matrices are constructed similarly. We assume that the $(N_1 \times p)$ matrix $\mathbf{X}$ in (2.5) is a matrix of fixed constants and has rank p, where N_1 denotes the number of observations on the y-variable for the one-fold, nested-error model (i.e., $N_1 = \sum_{i=1}^{t} n_i$).

Given the assumptions on the random errors v_i and e_{ij}, the covariance matrix for the vector $\mathbf{u}$ is the block-diagonal matrix

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{V}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{V}_t \end{bmatrix}, \tag{2.6}$$

where

$$\mathbf{V}_i = \sigma_e^2 \mathbf{I}_{n_i} + \sigma_v^2 \mathbf{J}_{n_i} \tag{2.6a}$$

and $\mathbf{I}_{n_i}$ denotes the identity matrix of order n_i, and $\mathbf{J}_{n_i}$ denotes the $(n_i \times n_i)$ matrix with all elements equal to one.

It is readily verified that the characteristic roots of $\mathbf{V}_i$ are σ_e^2 and $(\sigma_e^2 + n_i\sigma_v^2)$, having multiplicity $(n_i - 1)$ and 1, respectively. Further, $\mathbf{V}_i$ is equivalently expressed as

$$\mathbf{V}_i = \sigma_e^2\left(\mathbf{I}_{n_i} - \frac{1}{n_i}\mathbf{J}_{n_i}\right) + (\sigma_e^2 + n_i\sigma_v^2)\frac{1}{n_i}\mathbf{J}_{n_i} \tag{2.7}$$

where $[\mathbf{I}_{n_i} - (1/n_i)\mathbf{J}_{n_i}]$ and $(1/n_i)\mathbf{J}_{n_i}$ are symmetric idempotent matrices that are mutually orthogonal. From Lemma 1, or by direct multiplication, we find that the matrix

$$\mathbf{V}_i^{-\frac{1}{2}} = \left(\mathbf{I}_{n_i} - \frac{1}{n_i}\mathbf{J}_{n_i}\right)\Big/\sigma_e + \frac{1}{n_i}\mathbf{J}_{n_i}\Big/(\sigma_e^2 + n_i\sigma_v^2)^{\frac{1}{2}} \tag{2.8}$$

transforms the error vector $\mathbf{u}_i$, associated with the covariance matrix (2.7), into a vector of errors that are uncorrelated with variances equal to one. Thus the matrix $\mathbf{T}_i \equiv \sigma_e V_i^{-\frac{1}{2}}$ transforms the error vector $\mathbf{u}_i$ to a vector of uncorrelated random variables with variances σ_e^2. The matrix $\mathbf{T}_i$ is equivalently expressed as

$$\mathbf{T}_i = \mathbf{I}_{n_i} - \{1 - [\sigma_e^2/(\sigma_e^2 + n_i\sigma_v^2)]^{\frac{1}{2}}\}\frac{1}{n_i}\mathbf{J}_{n_i}. \tag{2.9}$$

Thus, by pre-multiplying the one-fold, nested-error model (2.5) by the block-diagonal matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{T}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{T}_t \end{bmatrix} \tag{2.9a}$$

we obtain the transformed model that is expressed algebraically by equation (2.4).

2.3 Estimation of Variance Components

When the variance components σ_v^2 and σ_e^2 are unknown, the values of the transformation factors α_i defined in (2.4a) must be estimated from estimates of σ_v^2 and σ_e^2. The different techniques available for estimation of variance components are reviewed by Searle [10], but no definitive results are available to suggest the "best" method. To estimate the variance components in our model, we use the well known "fitting-of-constants" method suggested by Henderson [6] and discussed by Searle [10, p. 54].

By regressing the y-deviations, $y_{ij} - \bar{y}_{i.}$, on the x-deviations, $x_{ijk} - \bar{x}_{i.k}$, $k = 1, 2, \cdots, p$, that are not identically zero, we obtain the unbiased estimator for σ_e^2

$$\hat{\sigma}_e^2 = \hat{\mathbf{e}}'\hat{\mathbf{e}}/(N_1 - t - p + \lambda_1) \tag{2.10}$$

where $\hat{\mathbf{e}}'\hat{\mathbf{e}}$ denotes the residual sum of squares obtained from the regression and λ_1 is the number of x-variables which are a linear combination of the indicator variables for individuals. By use of the deviations from individual means, the "individual effects" are removed. If the only x-variable which is a linear combination of the indicator variables for individuals is a column of ones, then the regression is performed on the $p - 1$ columns of x-deviations $x_{ijk} - \bar{x}_{i.k}$, $k = 2, \cdots, p$.

The variance component σ_v^2 is unbiasedly estimated by

$$\tilde{\sigma}_v^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}} - (N_1 - p)\hat{\sigma}_e^2}{N_1 - \operatorname{tr}[(\mathbf{X}'\mathbf{X})^{-1}\sum_{i=1}^{t} n_i^2 \bar{x}_{i.}'\bar{x}_{i.}]} \tag{2.11}$$

where $\hat{\mathbf{u}}'\hat{\mathbf{u}}$ denotes the residual sum of squares from the regression of $\mathbf{Y}$ on $\mathbf{X}$, and $\bar{x}_{i.}$ denotes the $(1 \times p)$ vector having kth element $\bar{x}_{i.k}$, $k = 1, 2, \cdots, p$.

We suggest the non-negative estimator of σ_v^2

$$\begin{aligned} \hat{\sigma}_v^2 &= \tilde{\sigma}_v^2 \quad \text{if} \quad \tilde{\sigma}_v^2 > 0 \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{2.11a}$$

The estimators for σ_e^2 and σ_v^2, given by (2.10) and (2.11a), are substituted into (2.4a) to obtain estimators for the transformation factors.

In the classical split-plot experiment (where $n_i = n$ for all i), the expectations of the split-plot error mean square and the whole-plot error mean square are σ_e^2 and $\sigma_e^2 + n\sigma_v^2$, respectively, in our notation (see [8, p. 375]). This implies that, for data from split-plot experiments, the transformation factor (2.4a) can be estimated by use of the split-plot and whole-plot error mean square as computed from the split-plot, analysis-of-variance table.

2.4 Properties of the Estimated, Generalized Least-Squares Estimator

The estimated, generalized least-squares estimator (1.5), obtained by use of the estimators $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$, remains unbiased for β in a wide range of situations.

Theorem 2: If the errors, u_{ij}, in the linear model (2.5)–(2.6) have fourth moments and are symmetrically distributed, and if the expectation of $(\hat{\sigma}_e^2)^{-1}$ exists, then

$$E(\hat{\boldsymbol{\beta}}) \equiv E\{(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Y}\} = \boldsymbol{\beta}.$$

Proof: Since $\mathbf{u}$ is distributed symmetrically about zero and the estimators $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$ are even functions of $\mathbf{u}$, it follows from the result of Kakwani [7] that $\hat{\boldsymbol{\beta}}$ is unbiased if its expectation exists. To demonstrate that the expectation exists, we consider an arbitrary linear combination of $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$. Let $\boldsymbol{\eta}$ denote any vector of real numbers from n-dimensional space. We consider the expectation of

$|\boldsymbol{\eta}'\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})|$. Now

$$|\boldsymbol{\eta}'\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta})| = |\boldsymbol{\eta}'\hat{\mathbf{V}}^{\frac{1}{2}}\hat{\mathbf{V}}^{-\frac{1}{2}}\mathbf{X}(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{u}|$$
$$\leq (\boldsymbol{\eta}'\hat{\mathbf{V}}\boldsymbol{\eta})^{\frac{1}{2}}[\mathbf{u}'\hat{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{u}]^{\frac{1}{2}}$$
$$\leq (\boldsymbol{\eta}'\hat{\mathbf{V}}\boldsymbol{\eta})^{\frac{1}{2}}(\mathbf{u}'\hat{\mathbf{V}}^{-1}\mathbf{u})^{\frac{1}{2}}$$

where the last inequality follows from the fact that $\hat{\mathbf{V}}^{-\frac{1}{2}}\mathbf{X}(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-\frac{1}{2}}$ is a symmetric idempotent matrix.

The minimum and maximum characteristic roots of $\hat{\mathbf{V}}$ are $\hat{\sigma}_e^2$ and $\hat{\sigma}_e^2 + n_L\hat{\sigma}_v^2$, respectively, where n_L is the largest of the n_i, $i = 1, 2, \cdots, t$. Therefore

$$(\boldsymbol{\eta}'\hat{\mathbf{V}}\boldsymbol{\eta})^{\frac{1}{2}}(\mathbf{u}'\hat{\mathbf{V}}^{-1}\mathbf{u})^{\frac{1}{2}} \leq \left(\frac{\hat{\sigma}_e^2 + n_L\hat{\sigma}_v^2}{\hat{\sigma}_e^2}\right)^{\frac{1}{2}} (\boldsymbol{\eta}'\boldsymbol{\eta})^{\frac{1}{2}}(\mathbf{u}'\mathbf{u})^{\frac{1}{2}}.$$

Since $(\hat{\sigma}_e^2 + n_L\hat{\sigma}_v^2)$ is bounded by a multiple of $\mathbf{u}'\mathbf{u}$, the multiple depending on the matrix $\mathbf{X}$, and since the expectation of $(\hat{\sigma}_e^2)^{-1}$ exists, and the u_{ij} have fourth moments, it follows that $E|\boldsymbol{\eta}'\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta})|$ is finite and thus $E(\boldsymbol{\beta}) = \boldsymbol{\beta}$.

We note that if the errors u_{ij} are normally distributed and if $(N_1 - t - p + \lambda_1) > 2$, then the expectation of $(\hat{\sigma}_e^2)^{-1}$ exists, and hence $\hat{\boldsymbol{\beta}}$ is unbiased for $\boldsymbol{\beta}$.

To consider the large-sample properties of the estimated, generalized least-squares estimator, we first set the problem in a somewhat more general context. We express the model as

$$\mathbf{Y}_n = \mathbf{X}_n\boldsymbol{\beta} + \mathbf{u}_n \tag{2.12}$$

where

$\mathbf{X}_n$ denotes an $(n \times p)$ matrix of fixed constants of rank p;
β is a vector of p unknown parameters;
$E(\mathbf{u}_n) = 0$ for all n; and
$E(\mathbf{u}_n\mathbf{u}_n') = \mathbf{V}_n$ and $\mathbf{V}_n^{-1}$ exists for all n.

We define the regularity conditions:

(1) The elements of $\mathbf{V}_n$ are functions of a q-dimensional vector of parameters γ, such that the elements of the matrices

$$\mathbf{G}_{nr}(\gamma) = \frac{\partial}{\partial\gamma_r}\mathbf{V}^{-1}(\gamma), \quad r = 1, 2, \cdots, q$$

are continuous functions of γ in an open sphere $\mathbf{S}$ of γ^0, the true value of the parameter vector γ;

(2) The sequences of matrices $\{\mathbf{X}_n\}$ and $\{\mathbf{V}_n\}$ are such that

$$\lim_{n\to\infty}\frac{1}{n}\mathbf{X}'\mathbf{V}_n^{-1}(\gamma)\mathbf{X}_n = \mathbf{M}(\gamma),$$

where $\mathbf{M}(\gamma)$ is a $(p \times p)$ matrix of fixed constants such that $\mathbf{M}^{-1}(\gamma)$ exists for all γ in $\mathbf{S}$, and

$$\lim_{n\to\infty}\frac{1}{n}\mathbf{X}_n'\mathbf{G}_{nr}(\gamma)\mathbf{X}_n = \mathbf{H}_r(\gamma),$$

where $\mathbf{H}_r(\gamma)$ is a matrix whose elements are continuous functions of γ, $r = 1, 2, \cdots, q$; and

(3) An estimator, $\hat{\mathbf{V}}_n \equiv \mathbf{V}_n(\hat{\gamma})$, for $\mathbf{V}_n \equiv \mathbf{V}_n(\gamma^0)$ is available such that $\mathbf{V}_n^{-1}(\hat{\gamma})$ exists for all n and $\hat{\gamma}$ satisfies the condition

$$\hat{\gamma} = \gamma^0 + O_p(n^{-\delta}), \quad \delta > 0.$$

Theorem 3: The regularity conditions (1)–(3) for the linear model (2.12) are sufficient conditions for the estimator

$$\hat{\boldsymbol{\beta}}_n = (\mathbf{X}_n'\hat{\mathbf{V}}_n^{-1}\mathbf{X}_n)^{-1}\mathbf{X}_n'\hat{\mathbf{V}}_n^{-1}\mathbf{Y}_n$$

to have the same asymptotic distribution as the estimator

$$\tilde{\boldsymbol{\beta}}_n = (\mathbf{X}_n'\mathbf{V}_n^{-1}\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{V}_n^{-1}\mathbf{Y}_n.$$

Proof: A sketch of the proof is the following: The estimator $\hat{\boldsymbol{\beta}}_n$ can be written

$$\hat{\boldsymbol{\beta}}_n = \boldsymbol{\beta} + [\mathbf{X}_n'\mathbf{V}_n^{-1}(\hat{\boldsymbol{\gamma}})\mathbf{X}_n]^{-1}\mathbf{X}_n'\mathbf{V}_n^{-1}(\hat{\boldsymbol{\gamma}})\mathbf{u}_n.$$

By a Taylor's expansion with remainder about the true parameter $\boldsymbol{\gamma}^0$, we obtain

$$(\mathbf{X}_n'\mathbf{V}_n^{-1}(\hat{\boldsymbol{\gamma}})\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{V}_n^{-1}(\hat{\boldsymbol{\gamma}})\mathbf{u}_n$$
$$= \left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^0)\mathbf{X}_n\right)^{-1}\left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^0)\mathbf{u}_n\right)$$
$$+ \sum_{r=1}^{q}\left\{\left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^*)\mathbf{X}_n\right)^{-1}\left(\frac{1}{n}\mathbf{X}_n'\mathbf{G}_{nr}(\boldsymbol{\gamma}^*)\mathbf{u}_n\right)\right.$$
$$- \left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^*)\mathbf{X}_n\right)^{-1}\left(\frac{1}{n}\mathbf{X}_n'\mathbf{G}_{nr}(\boldsymbol{\gamma}^*)\mathbf{X}_n\right)$$
$$\left.\cdot\left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^*)\mathbf{X}_n\right)^{-1}\left(\frac{1}{n}\mathbf{X}_n'\mathbf{V}_n^{-1}(\boldsymbol{\gamma}^*)\mathbf{u}_n\right)\right\}(\hat{\gamma}_r - \gamma_r^0)$$

where $\boldsymbol{\gamma}^*$ is between $\boldsymbol{\gamma}^0$ and $\boldsymbol{\gamma}$. By the regularity conditions (1)–(3), it follows that

$$(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = (\tilde{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) + O_p[n^{-(\frac{1}{2}+\delta)}]$$

and so

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = \sqrt{n}(\tilde{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) + O_p(n^{-\delta}).$$

Corollary 3: Given the model $\mathbf{Y}_n = \mathbf{X}_n\beta + \mathbf{u}_n$, where the elements of $\mathbf{u}_n$ are independent random variables with the covariance structure of (2.3), the u_{ij} have finite fourth moments, t^{-1} and $[\sum_{i=1}^{t}(n_i - 1)]^{-1}$ are both of order $n^{-\delta}$, $\delta > 0$, where $n \equiv \sum_{i=1}^{t} n_i$, $\mathbf{X}_n'\mathbf{X}_n$ is nonsingular for all n, and $\lim_{n\to\infty}(1/n)\mathbf{X}_n'\mathbf{X}_n$ and $\lim_{n\to\infty}(1/n)\mathbf{X}_n'\mathbf{V}_n^{-1}\mathbf{X}_n$ exist and are positive definite $(p \times p)$ matrices; then

$$\hat{\boldsymbol{\beta}}_n = (\mathbf{X}_n'\hat{\mathbf{V}}_n^{-1}\mathbf{X}_n)^{-1}\mathbf{X}_n'\hat{\mathbf{V}}_n^{-1}\mathbf{Y}_n$$

has the same asymptotic distribution as the estimator

$$\tilde{\boldsymbol{\beta}}_n = (\mathbf{X}_n'\mathbf{V}_n^{-1}\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{V}_n^{-1}\mathbf{Y}_n.$$

Proof: The matrix $\mathbf{V}_n$, associated with the nested model, is a function of the two parameters σ_e^2 and σ_v^2. The matrix $\mathbf{V}_n$ is block-diagonal with ith block

$$\mathbf{V}_{ni} = \sigma_e^2\left(\mathbf{I}_{ni} - \frac{1}{n_i}\mathbf{J}_{ni}\right) + (\sigma_e^2 + n_i\sigma_v^2)\frac{1}{n_i}\mathbf{J}_{ni}.$$

The matrices denoted by $\mathbf{G}_{nr}(\gamma)$ in condition (1) of Theorem 3 are block diagonal with the matrices

$$\mathbf{G}_{ni1} = -\left(\frac{1}{\sigma_e^2}\right)^2\left(\mathbf{I}_{ni} - \frac{1}{n_i}\mathbf{J}_{ni}\right) - \frac{1}{(\sigma_e^2 + n_i\sigma_v^2)^2}\frac{1}{n_i}\mathbf{J}_{ni},$$
$$i = 1, 2, \cdots, t$$

and

$$\mathbf{G}_{ni2} = -\frac{1}{(\sigma_e^2 + n_i\sigma_v^2)^2}\mathbf{J}_{ni}, \quad i = 1, 2, \cdots, t$$

as blocks for $\mathbf{G}_{ni}(\gamma)$ and $\mathbf{G}_{n2}(\gamma)$, respectively.

Clearly, the elements of these matrices are continuous functions of $\sigma_e^2 > 0$ and $\sigma_v^2 \geq 0$ in a sphere containing the true parameters. Since $\lim_{n\to\infty} (1/n) \sum_{i=1}^t n_i^2 \bar{\mathbf{x}}_{i.}' \bar{\mathbf{x}}_{i.}$ exists by assumption, then condition (2) of Theorem 3 is satisfied. Further, condition (3) of Theorem 3 is satisfied since the variance of $\hat{\sigma}_e^2$ is of order $[\sum_{i=1}^t (n_i - 1)]^{-1}$ and the variance of $\hat{\sigma}_v^2$ is of order

$$\max \{[\sum_{i=1}^t (n_i - 1)]^{-1}, t^{-1}\}.$$

The existence of the fourth moments of the u_{ij} implies that the variances of $\hat{\sigma}_e^2$ and $\hat{\sigma}_v^2$ exist. The result of Corollary 3 thus follows from Theorem 3.

3. THE TWO-FOLD, NESTED-ERROR MODEL

3.1 Model Presentation

We extend the model considered in Section 2 to the case where there is a two-fold nesting structure in the data. The variable under study, y, thus has three subscripts. The first subscript distinguishes the "individual" in the sample, the second subscript distinguishes the "measurement" for the particular individual, and the third subscript distinguishes the "determination" of the measurements on the individuals. We assume that t "individuals" are chosen at random and n_i "measurements" are taken on the ith individual, each of these measurements having K_i "determinations." We thus assume that, for a given individual, the number of determinations is the same for each measurement. We do not require that the number of measurements or the number of determinations be the same for all individuals. The linear model is expressed as

$$y_{ijk} = \sum_{m=1}^p x_{ijkm}\beta_m + u_{ijk}, \quad k = 1, \cdots, K_i; \quad j = 1, \cdots, n_i; \quad i = 1, \cdots, t; \tag{3.1}$$

and

$$u_{ijk} = v_i + e_{ij} + \epsilon_{ijk} \tag{3.2}$$

where

y_{ijk} denotes the value of the variable obtained at the kth determination of the jth measurement on the ith individual in the sample;

x_{ijkm}, $m = 1, \cdots, p$, denote the levels of the p control variables at which the observation y_{ijk} is obtained (the x_{ijkm} are assumed fixed constants);

β_m, $m = 1, \cdots, p$, denote the unknown parameters to be estimated; and

u_{ijk}, the random error associated with the observation y_{ijk}, is assumed the sum of the random effects associated with the ith individual (v_i), the jth measurement on the ith individual (e_{ij}), and the kth determination of the jth measurement on the ith individual (ϵ_{ijk}).

We assume that the random errors v_i, e_{ij}, and ϵ_{ijk} are independently distributed with zero means and variances σ_v^2, σ_e^2, and σ_ϵ^2, respectively, where $\sigma_v^2 \geq 0$, $\sigma_e^2 \geq 0$, and $\sigma_\epsilon^2 > 0$. The covariance structure for the random errors u_{ijk} in the two-fold, nested-error model (3.1) is thus expressed by

$$\begin{aligned} E(u_{ijk}u_{i'j'k'}) &= \sigma_v^2 + \sigma_e^2 + \sigma_\epsilon^2, && \text{if } i = i', \; j = j', \; k = k' \\ &= \sigma_v^2 + \sigma_e^2, && \text{if } i = i', \; j = j', \; k \neq k' \\ &= \sigma_v^2, && \text{if } i = i', \; j \neq j; \\ &= 0, && \text{if } i \neq i'. \end{aligned}$$

A classical example of the two-fold, nested-error model is the split-split-plot experiment. In our notation, t would represent the number of whole-plots in the experiment, n (where $n_i = n$ for all i) would represent the number of split-plots per whole-plot, and K (where $K_i = K$ for all i) would represent the number of split-split-plots per split-plot in the trials. A second example is subsampling within primary, secondary, and tertiary sampling units.

3.2 Transformation for the Two-Fold, Nested-Error Model

For the two-fold, nested-error model, we suggest the transformation

$$\begin{aligned} &(y_{ijk} - \alpha_{1i}\bar{y}_{ij.} - \alpha_{2i}\bar{y}_{i..}) \\ &= \sum_{m=1}^p (x_{ijkm} - \alpha_{1i}\bar{x}_{ij.m} - \alpha_{2i}\bar{x}_{i..m})\beta_m + u_{ijk}^*, \\ &\quad k = 1, \cdots, K_i; \quad j = 1, \cdots, n_i; \quad i = 1, \cdots, t, \end{aligned} \tag{3.4}$$

where

$$\alpha_{1i} = 1 - [\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2)]^{\frac{1}{2}} \tag{3.4a}$$

$$\alpha_{2i} = [\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2)]^{\frac{1}{2}} - [\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2 + n_iK_i\sigma_v^2)]^{\frac{1}{2}}. \tag{3.4b}$$

$\bar{y}_{ij.}$, $\bar{x}_{ij.m}$, $m = 1, 2, \cdots, p$, denote the averages of the y- and x-determinations for the jth measurement on the ith individual; and $\bar{y}_{i..}$, $\bar{x}_{i..m}$, $m = 1, 2, \cdots, p$, denote the averages of the y- and x-measurements for the ith individual. In Theorem 4 we prove that the errors, u_{ijk}^*, are uncorrelated and have variances σ_ϵ^2.

Theorem 4: For the linear model (3.1)–(3.3), the transformed errors, $u_{ijk}^* \equiv u_{ijk} - \alpha_{1i}\bar{u}_{ij.} - \alpha_{2i}\bar{u}_{i..}$, where $\bar{u}_{ij.} = \sum_{k=1}^{K_i} u_{ijk}/K_i$ and $\bar{u}_{i..} = \sum_{j=1}^{n_i}\sum_{k=1}^{K_i} u_{ijk}/n_iK_i$, are uncorrelated and have variances σ_ϵ^2.

Proof: It can be shown that the transformation (3.4) is obtained by twice applying the one-fold, nested-error transformation (2.4). An alternative approach is to demonstrate that the transformation arises from the use of Lemma 1. This approach is outlined with the use of the matrix notation. We express the linear model (3.1) as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{3.5}$$

where

$$\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2', \cdots, \mathbf{Y}_t')', \quad \mathbf{Y}_i' = (\mathbf{Y}_{i1}', \mathbf{Y}_{i2}', \cdots, \mathbf{Y}_{in_i}'),$$

$\mathbf{Y}_{ij}' = (y_{ij1}, y_{ij2}, \cdots, y_{ijK_i})$ and the $\mathbf{X}$ and $\mathbf{u}$ matrices are similarly constructed.

The number of observations in this linear model, $\sum_{i=1}^{t} n_i K_i$, is denoted N_2. As in the one-fold, nested-error model, we denote $\sum_{i=1}^{t} n_i$ by N_1. The covariance matrix for the vector $\mathbf{u}$ in the linear model (3.5) is the block-diagonal matrix

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{V}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{V}_t \end{bmatrix} \tag{3.6}$$

where

$$\mathbf{V}_i = \sigma_\epsilon^2 \mathbf{I}_{n_iK_i} + \sigma_e^2 \mathbf{M}_{n_iK_i} + \sigma_v^2 \mathbf{J}_{n_iK_i} \tag{3.6a}$$

and $\mathbf{I}_{n_iK_i}$ denotes the identity matrix of order n_iK_i, $\mathbf{M}_{n_iK_i}$ denotes the Kronecker-product matrix of $\mathbf{I}_{n_i}$ and $\mathbf{J}_{K_i}$, and $\mathbf{J}_{n_iK_i}$ denotes the square matrix of order n_iK_i with all elements equal to one.

The covariance matrix $\mathbf{V}_i$ is equivalently expressed as

$$\begin{aligned} \mathbf{V}_i = \sigma_\epsilon^2 \left(\mathbf{I}_{n_iK_i} - \frac{1}{K_i} \mathbf{M}_{m_iK_i} \right) \\ + (\sigma_\epsilon^2 + K_i\sigma_e^2)\left(\frac{1}{K_i} \mathbf{M}_{n_iK_i} - \frac{1}{n_iK_i} \mathbf{J}_{n_iK_i} \right) \\ + (\sigma_\epsilon^2 + K_i\sigma_e^2 + n_iK_i\sigma_v^2) \frac{1}{n_iK_i} \mathbf{J}_{n_iK_i} \end{aligned} \tag{3.7}$$

where it can be verified that the three matrices in this expression are symmetric idempotent matrices that are mutually orthogonal. The coefficients of the idempotent matrices in (3.7) are the characteristic roots of $\mathbf{V}_i$. From Lemma 1, the matrix

$$\begin{aligned} \mathbf{V}_i^{-\frac{1}{2}} = \left(\mathbf{I}_{n_iK_i} - \frac{1}{K_i} \mathbf{M}_{n_iK_i} \right) \Big/ \sigma_\epsilon \\ + \left(\frac{1}{K_i} \mathbf{M}_{n_iK_i} - \frac{1}{n_iK_i} \mathbf{J}_{n_iK_i} \right) \Big/ (\sigma_\epsilon^2 + K_i\sigma_e^2)^{\frac{1}{2}} \\ + \frac{1}{n_iK_i} \mathbf{J}_{n_iK_i} \Big/ (\sigma_\epsilon^2 + K_i\sigma_e^2 + n_iK_i\sigma_v^2)^{\frac{1}{2}} \end{aligned} \tag{3.8}$$

transforms the error vector $\mathbf{u}_i$ of model (3.5) to one with errors that are uncorrelated with variances equal to one. Thus the matrix $\mathbf{T}_i \equiv \sigma_\epsilon \mathbf{V}_i^{-\frac{1}{2}}$ transforms the error vector $\mathbf{u}_i$ to variables with variances equal to σ_ϵ^2. The matrix $\mathbf{T}_i$ is equivalently expressed as

$$\begin{aligned} \mathbf{T}_i = \mathbf{I}_{n_iK_i} - \{1 - [\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2)]^{\frac{1}{2}}\} \frac{1}{K_i} \mathbf{M}_{n_iK_i} \\ - \{[\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2)]^{\frac{1}{2}} \\ - [\sigma_\epsilon^2/(\sigma_\epsilon^2 + K_i\sigma_e^2 + n_iK_i\sigma_v^2)]^{\frac{1}{2}}\} \frac{1}{n_iK_i} \mathbf{J}_{n_iK_i}. \end{aligned} \tag{3.9}$$

Thus, by pre-multiplying the two-fold, nested-error model (3.5) by the block-diagonal matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{T}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{T}_t \end{bmatrix} \tag{3.9a}$$

we obtain the transformed model that is expressed algebraically by equation (3.4).

3.3 Estimation of Variance Components

We present the fitting-of-constants estimators for the variance components in the two-fold, nested-error model.

By regressing the y-deviations, $y_{ijk} - \bar{y}_{ij.}$, on the x-deviations, $x_{ijkm} - \bar{x}_{ij.m}$, $m = 1, 2, \cdots, p$, that are not identically zero, we obtain the unbiased estimator for σ_ϵ^2

$$\hat{\sigma}_\epsilon^2 = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}/(N_2 - N_1 - p + \lambda_{12}) \tag{3.10}$$

where $\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$ is the residual sum of squares from the regression and λ_{12} is the number of x-variables that have constant values for all determinations of measurements for individuals.

From the regression of the y-deviations, $y_{ijk} - \bar{y}_{i..}$, on the x-deviations, $x_{ijkm} - \bar{x}_{i..m}$, $m = 1, 2, \cdots, p$, that are not identically zero, we obtain the unbiased estimator for σ_e^2

$$\hat{\sigma}_e^2 = \frac{\hat{\mathbf{e}}'\hat{\mathbf{e}} - (N_2 - t - p + \lambda_1)\hat{\sigma}_\epsilon^2}{N_2 - \operatorname{tr}\{[(\mathbf{X} - \overline{\mathbf{X}}_{(1..)})'(\mathbf{X} - \overline{\mathbf{X}}_{(1..)})]^{-1} \sum_{i=1}^{t} \sum_{j=1}^{n_i} K_i^2 (\overline{\mathbf{x}}_{ij.} - \overline{\mathbf{x}}_{i..})'(\overline{\mathbf{x}}_{ij.} - \overline{\mathbf{x}}_{i..})\}} \tag{3.11}$$

where

$\hat{\mathbf{e}}'\hat{\mathbf{e}}$ denotes the residual sum of squares from the regression;

λ_1 denotes the number of x-variables that have constant values for measurements of individuals;

$(\mathbf{X} - \overline{\mathbf{X}}_{(1..)})$ denotes the $[N_2 \times (p - \lambda_1)]$ matrix of the values of the independent variables in the regression; and

$(\overline{\mathbf{x}}_{ij.} - \overline{\mathbf{x}}_{i..})$ denotes the $(p - \lambda_1)$ row vector of the non-zero deviations,

$\overline{\mathbf{x}}_{ij.m} - \overline{\mathbf{x}}_{i..m}$, for the jth measurement on the ith individual.

The estimator for σ_v^2 is

$$\hat{\sigma}_v^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}} - (N_2 - p)\hat{\sigma}_\epsilon^2 - \{N_2 - \operatorname{tr}[(\mathbf{X}'\mathbf{X})^{-1} \sum_{i=1}^{t} \sum_{j=1}^{n_i} K_i^2 \overline{\mathbf{x}}_{ij.}' \overline{\mathbf{x}}_{ij.}]\}\hat{\sigma}_e^2}{N_2 - \operatorname{tr}[(\mathbf{X}'\mathbf{X})^{-1} \sum_{i=1}^{t} n_i^2 K_i^2 \overline{\mathbf{x}}_{i..}' \overline{\mathbf{x}}_{i..}]} \tag{3.12}$$

where

$\hat{\mathbf{u}}'\hat{\mathbf{u}}$ denotes the residual sum of squares from the regression of $\mathbf{Y}$ on $\mathbf{X}$ in model (3.5);

$\bar{\mathbf{x}}_{ij.}$ denotes the $(1 \times p)$ vector of means of the x-values for the determinations of the jth measurement for the ith individual; and

$\bar{\mathbf{x}}_{i..}$ denotes the $(1 \times p)$ vector of means of the x-values for the determinations and measurements on the ith individual.

In practice, if a negative value is obtained for $\hat{\sigma}_e^2$ or $\hat{\sigma}_v^2$, the corresponding variance component is estimated by zero.

In balanced split-split-plot experiments, the transformation factors α_{1i} and α_{2i}, defined by (3.4a) and (3.4b), are functions of expectations of mean squares from the associated analysis-of-variance table. That is, σ_ϵ^2, $(\sigma_\epsilon^2 + K\sigma_e^2)$, and $(\sigma_\epsilon^2 + K\sigma_e^2 + nK\sigma_v^2)$ are the expectations (in our notation) of the split-split-plot, split-plot, and whole-plot error mean squares, respectively (see [8, p. 381]). This implies that for data from a traditional split-split-plot experiment, the transformation factors (3.4a) and (3.4b) can be estimated directly from the analysis-of-variance table for the data.

A proof completely analogous to that of Theorem 2 can be used to show that the estimated, generalized least-squares estimator for the two-fold nested model is unbiased under the condition of Theorem 2. Likewise, the conditions of Corollary 3 are sufficient for this estimator to have the same limiting distribution as the generalized least-squares estimator with known variance components.

4. CONCLUSION

The transformations suggested for use with the nested-error models considered in this article are such that the variances of the transformed errors are the same as that of the error term in the final state of "nesting." This fact is useful for goodness-of-fit testing in certain research situations. For example, if the data arising from a split-plot field experiment are used for the estimation of a crop-response function, the split-plot and whole-plot error mean squares from the analysis of variance are used to estimate the variance components unbiasedly. Under the hypothesis of a given response function (assumed to be a linear model), the residual mean square from the regression with the transformed data estimates the variance of the split-plot error. An approximate F-test for goodness-of-fit for the hypothesized response function is thus obtained, given that the errors are normally distributed (e.g., see [2, 12]).

In this article, attention has been confined to transformations useful in the estimation of linear models with nested-error structure. In recent years the linear model with the error decomposition $u_{ij} = v_i + e_j + \epsilon_{ij}$, $i = 1, 2, \cdots, N$; $j = 1, 2, \cdots, T$, has been considered for the combining of cross-section and time-series data in econometrics (e.g., [1, 9, 13]). In another paper [3], the authors discuss the estimation of linear models with this error structure.

[*Received March 1972. Revised January 1973.*]

REFERENCES

[1] Balestra, P. and Nerlove, M., "Pooling Cross-Sections and Time-Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas," *Econometrica*, 34 (July 1966), 585–612.

[2] Battese, G.E. and Fuller, W.A., "Determination of Economic Optima from Crop-Rotation Experiments," *Biometrics*, 28 (September 1972), 781–92.

[3] Fuller, W.A., and Battese, G.E., "Estimation of the Linear Model with Crossed Error Structure," (submitted for publication).

[4] Garcia, P., Brewer, W.D., Battese, G.E. and Hotchkiss, D.K., "Nutritional Status of Women During Aging: A Longitudinal Study," Unpublished paper, Department of Food and Nutrition, Iowa State University, Ames, 1972.

[5] Graybill, F.A., *An Introduction to Linear Statistical Models*, Volume 1, New York: McGraw Hill Book Co., 1961.

[6] Henderson, C.R., "Estimation of Variance and Covariance Components," *Biometrics*, 9 (June 1953), 226–52.

[7] Kakwani, N.C., "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equations Estimators," *Journal of the American Statistical Association*, 62 (March 1967), 141–2.

[8] Kempthorne, O., *Design and Analysis of Experiments*, New York: John Wiley and Sons, Inc., 1952.

[9] Nerlove, M., "A Note on Error Component Models," *Econometrica*, 39 (March 1971), 383–96.

[10] Searle, S.R., "Topics in Variance Component Estimation," *Biometrics*, 27 (March 1971), 1–76.

[11] Shih, Chang Sheng, "Interval Estimation for the Exponential Model and the Analysis of Rotation Experiments," Ph.D. thesis, 126 p., Iowa State University, Ames, 1966.

[12] Shrader, W.D., Fuller, W.A., and Cady, F.B., "Estimation of a Common Nitrogen Response Function for Corn in Different Crop Rotations," *Agronomy Journal*, 58 (July–August 1966), 397–401.

[13] Wallace, T.D., and Hussain, A., "The Use of Error Components Models in Combining Cross Section with Time Series Data," *Econometrica*, 37 (January 1969), 55–72.