



Nonnegative estimation of variance components in heteroscedastic one-way random effects ANOVA models

T. Mathew, T. Nahtman, D. von Rosen, and B. K. Sinha

**Research Report
Centre of Biostochastics**

Swedish University of
Agricultural Sciences

Report 2007:1
ISSN 1651-8543

Nonnegative estimation of variance components in heteroscedastic one-way random effects ANOVA models

THOMAS MATHEW

*Department of Mathematics and Statistics, University of Maryland,
Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA*

TATJANA NAHTMAN

*Institute of Mathematical Statistics, University of Tartu,
Liivi 2, 50409 Tartu, Estonia*

DIETRICH VON ROSEN¹

*Centre of Biostochastics, Swedish University of Agricultural Sciences, Box
7032, 750 07 Uppsala, Sweden*

BIMAL K. SINHA

*Department of Mathematics and Statistics, University of Maryland,
Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA*

Abstract

There is considerable amount of literature dealing with inference about the parameters in a heteroscedastic one-way random effects ANOVA model. In this paper we primarily address the problem of improved quadratic estimation of the random effect variance component. It turns out that such estimators with a smaller MSE compared to some standard unbiased quadratic estimators exist under quite general conditions. Improved estimators of the error variance components are also established.

Keywords: Variance components, heteroscedastic one-way ANOVA, perturbed estimators

MSC2000: 62F99, 62J99

¹E-mail address to the correspondence author: Dietrich.von.Rosen@bt.slu.se

1 Introduction

In the context of a heteroscedastic random effects one-way ANOVA model, there is a considerable amount of literature dealing with inference about the common mean, random effect variance component as well as the error variance components. The earliest work in this area is due to Cochran (1937, 1954). Other notable works are due to Rao (1977), Hartley and Rao (1978), Rao, Kaplan and Cochran (1981), Harville (1977), Vangel and Rukhin (1999), and Rukhin and Vangel (1998). Some related works are reported in Fairweather (1972), Jordan and Krishnamoorthy (1996) and Yu, Sun and Sinha (1999).

The basic premise underlying the model is that repeated measurements are made on the same quantity by several laboratories using different instruments of varying precisions. Often non-negligible between-laboratory variability may be present and the number of measurements made at each laboratory may also differ. The inference problems of interest are on the fixed common mean, inter-laboratory variance component and also the intra-laboratory variances. It should be mentioned that there is a huge literature on estimation of the common mean when inter-laboratory variance is assumed to be absent (see Yu, Sun and Sinha (1999) and the references therein).

Assume that there are k laboratories, and that there are n_i measurements from the i th laboratory, $i = 1, \dots, k$. Denoting by X_{ij} the j th replicate measurement obtained from the i th laboratory, the model is

$$X_{ij} = \mu + \tau_i + e_{ij}, \quad (1)$$

where μ is the common mean, τ_1, \dots, τ_k are the random laboratory effects, assumed to be independent normal with mean 0 and variance σ_τ^2 , and the laboratory measurement errors e_{ij} 's are assumed to be independent and normally distributed with $Var(e_{ij}) = \sigma_i^2$, $j = 1, \dots, n_i$, $i = 1, \dots, k$. Moreover, τ_i 's and e_{ij} 's are also assumed to be independent. Here σ_τ^2 is known as the inter-laboratory (between) variance, and $\sigma_1^2, \dots, \sigma_k^2$ are known as the intra-laboratory (within) variances. There are several papers dealing with the estimation of the mean μ (see Rukhin and Vangel (1998) and Vangel and Rukhin (1999)). Our interest in this paper is in the improved estimation of the between and within laboratory variances: $\sigma_\tau^2, \sigma_1^2, \dots, \sigma_k^2$.

By sufficiency, without any loss of generality, inference about the common mean and the variance components can be based on the overall sample mean $\bar{X} = \sum_{ij} X_{ij} / \sum n_i$, the individual lab means $Y_i = \sum_j X_{ij} / n_i$, and the within

lab corrected sum of squares $S_i^2 = \sum_j (X_{ij} - Y_i)^2$. Obviously,

- (a) $\bar{X} \sim N[\mu, \sigma_\tau^2 \sum n_i^2/n^2 + \sum n_i \sigma_i^2/n^2]$, where $n = \sum n_i$,
- (b) $Y_i \sim N[\mu, \sigma_\tau^2 + \sigma_i^2/n_i]$,
- (c) $S_i^2 \sim \sigma_i^2 \chi_{n_i-1}^2$.

We note that $\{S_i^2\}$ is independent of $\{Y_i\}$. Estimators of within lab variances σ_i^2 are usually based on S_i^2 , which are typically unbiased estimators or their best multiples. There are several unbiased quadratic estimators of σ_τ^2 . Notably among them are the two derived in Rao, Kaplan and Cochran (1981), given below. It should be mentioned that, following a general result of LaMotte (1973), any *unbiased* quadratic estimator of σ_τ^2 is bound to assume negative values for some data points. Moreover, the non-uniqueness of the unbiased estimators of σ_τ^2 follows from the fact that the set of minimal sufficient statistics $\{\bar{X}, Y_i, i = 1, \dots, k, S_i^2, i = 1, \dots, k\}$ is not complete. Define

$$\begin{aligned}\bar{y} &= \sum n_i Y_i / n, \\ \bar{y}^* &= \sum Y_i / k.\end{aligned}$$

Then the two unbiased estimators mentioned above can be written as

$$\hat{\sigma}_{\tau 1}^2 = \sum (Y_i - \bar{y}^*)^2 / (k - 1) - \sum \frac{S_i^2}{kn_i(n_i - 1)}, \quad (2)$$

$$\hat{\sigma}_{\tau 2}^2 = \frac{1}{n - \frac{\sum n_i^2}{n}} \left\{ \sum n_i (Y_i - \bar{y})^2 - \sum \frac{(n - n_i) S_i^2}{n(n_i - 1)} \right\}. \quad (3)$$

Observe that in the balanced case, i.e. $n_1 = n_2 = \dots = n_k$, the estimators are identical.

Our primary objective in this paper is to derive improved quadratic estimators of σ_τ^2 with a smaller mean squared error compared to any quadratic unbiased estimator. It turns out, that such improved estimators exist quite generally. In particular, we derive conditions under which our proposed quadratic estimator dominates the two unbiased estimators given in (2)–(3). This is in the same spirit as in Kelly and Mathew (1993, 1994), though in somewhat different contexts. However, our proposed improved estimators can also assume negative values although with a smaller probability. We recommend that suitable modifications along the lines of Kelly and Mathew (1993, 1994) be done to obtain nonnegative (non-quadratic) improved estimators.

Following arguments in Mathew et al. (1992), we also derive simple estimators of the error variances $\sigma_1^2, \dots, \sigma_k^2$, which have smaller mean squared errors compared to their unbiased estimators or best multiples of the unbiased estimators. We should point out that Vangel and Rukhin (1999) discuss the maximum likelihood estimators of the parameters in our model and also provide a brief Bayesian discussion on this problem. Naturally due to the complicated nature of the likelihood, exact inference based on the likelihood is impossible, and one has to depend on large sample theory.

The organization of the paper is as follows. In section 2, we address the problem of improved estimation of between lab variance σ_τ^2 . A brief Bayesian analysis of the problem is also taken up in this section. Rather than concentrating on the posterior distribution of the variance parameters via highest posterior density (HPD) regions as in Vangel and Rukhin (1999), our main focus here is to study numerically the frequentist properties (mean and variance) of the Bayes estimator of σ_τ^2 . In section 3 we deal with the problem of improved nonnegative estimation of within lab variances $\sigma_1^2, \dots, \sigma_k^2$. Finally, in section 4, we analyze two data sets arising in the context of our model.

2 Improved estimation of σ_τ^2

In this section we discuss the problem of improved estimation of the between lab variance σ_τ^2 . Like the unbiased estimators of σ_τ^2 which can assume negative values, our proposed improved estimators which are essentially some variations of the unbiased estimators can also assume negative values. We first deal with improved quadratic estimators in a non-Bayesian framework in section 2.1. In section 2.2 Edgeworth expansion is carried out and in section 3, we derive and study some properties of the Bayes estimator of σ_τ^2 .

2.1 Improved quadratic estimators of σ_τ^2

Writing $Y = (Y_1, \dots, Y_k)'$, $N = \text{diag}(n_1, \dots, n_k)$, $\mathbf{n} = (n_1, \dots, n_k)'$, and $\mathbf{1}$ to be a vector of ones, the two unbiased estimators $\hat{\sigma}_{\tau 1}^2$ and $\hat{\sigma}_{\tau 2}^2$ can be expressed as

$$\hat{\sigma}_{\tau 1}^2 = Y'(I - \mathbf{1}\mathbf{1}'/k)Y/(k-1) - \sum \frac{S_i^2}{kn_i(n_i-1)}, \quad (4)$$

$$\hat{\sigma}_{\tau 2}^2 = \frac{n}{n^2 - \sum n_i^2} \left\{ Y'(N - \frac{\mathbf{n}\mathbf{n}'}{n})Y - \sum \frac{(n-n_i)S_i^2}{n(n_i-1)} \right\}. \quad (5)$$

Consider now the general estimator

$$\hat{\sigma}_\tau^2 = Y'AY + \sum c_i S_i^2,$$

where we assume A to be symmetric and satisfying $A\mathbf{1} = \mathbf{0}$ (since we would like to have σ_τ^2 translation invariant). Then the unbiasedness of $\hat{\sigma}_\tau^2$ requires that $\text{tr}A = 1$ and $c_i = -a_{ii}/\{n_i(n_i - 1)\}$, $i = 1, \dots, n$. Hence a general form of a translation invariant quadratic unbiased estimator of σ_τ^2 is given by

$$\hat{\sigma}_\tau^2 = Y'AY - \sum \frac{a_{ii}S_i^2}{n_i(n_i - 1)} \quad (6)$$

with $\text{tr}A = 1$ and $A\mathbf{1} = \mathbf{0}$. The variance of such an unbiased estimator is easily obtained as

$$\begin{aligned} \text{Var}(Y'AY - \sum \frac{a_{ii}S_i^2}{n_i(n_i - 1)}) &= 2\{\sum_{ij} (\frac{\sigma_i^2}{n_i} + \sigma_\tau^2)(\frac{\sigma_j^2}{n_j} + \sigma_\tau^2)a_{ij}^2 \\ &\quad + \sum \frac{a_{ii}^2\sigma_i^4}{n_i^2(n_i - 1)}\}, \end{aligned} \quad (7)$$

where we have also used the fact that $Y'AY$ and $\{S_i^2\}$ are independently distributed. Later on we also need the third moment,

$$\begin{aligned} E[(\hat{\sigma}_\tau^2)^3] &= E[(Y'AY)^3] - E[(Y'AY)^2] \sum a_{ii}/n_i \\ &\quad + E[Y'AY]\{\text{Var}[\frac{a_{ii}S_i^2}{n_i(n_i - 1)}] + (\sum a_{ii}/n_i)^2\} \\ &\quad + E[(\sum \frac{a_{ii}}{n_i(n_i - 1)}S_i^2)^3]. \end{aligned} \quad (8)$$

Now $\text{Var}[\sum \frac{a_{ii}}{n_i(n_i - 1)}S_i^2]$ is obtained from (7), and

$$\begin{aligned} E[(\sum \frac{a_{ii}}{n_i(n_i - 1)}S_i^2)^3] &= \sum \frac{a_{ii}^3}{n_i^3(n_i - 1)^3}E[(S_i^2)^3] \\ &\quad + 3 \sum_{i,k:i \neq k} \frac{a_{ii}^2 a_{kk}}{n_i^2(n_i - 1)^2 n_k(n_k - 1)}E[S_i^2]E[(S_k^2)^2] \\ &\quad + 6 \sum_{i \neq j \neq k} \frac{a_{ii} a_{jj} a_{kk}}{n_i n_j n_k}. \end{aligned}$$

If $n = 2$ the last term disappears. Moreover, since S_i^2/σ_i^2 is $\chi_{n_i-1}^2$,

$$\begin{aligned} E[S_i^2/\sigma_i^2] &= n_i - 1, \\ E[(S_i^2/\sigma_i^2)^2] &= (n_i - 1)^2 + 2(n_i - 1) \end{aligned}$$

and

$$E[(S_i^2/\sigma_i^2)^3] = (n_i - 1)^3 + 6(n_i - 1)^2 + 8(n_i - 1).$$

It remains to express $E[Y'AY]$, $E([Y'AY]^2)$ and $E([Y'AY]^3)$ in (8). Because $A\mathbf{1} = \mathbf{0}$ we may assume that $Y \sim N_k(\mathbf{0}, \Sigma)$, where

$$\Sigma = \sigma_\tau^2 I_k + \text{diag}(\sigma_1^2/n_1, \sigma_2^2/n_2, \dots, \sigma_k^2/n_k).$$

Then

$$E[Y'AY] = \text{tr}(\Sigma A), \quad (9)$$

$$E([Y'AY]^2) = (\text{tr}(\Sigma A))^2 + 2\text{tr}(\Sigma A \Sigma A), \quad (10)$$

$$E([Y'AY]^3) = (\text{tr}(\Sigma A))^3 + 6\text{tr}(\Sigma A)\text{tr}(\Sigma A \Sigma A) + 8\text{tr}(\Sigma A \Sigma A \Sigma A). \quad (11)$$

Thus, $E[(\hat{\sigma}_\tau^2)^3]$ has been derived.

We now observe the following facts from the variance expression in (7).

- (a) The coefficient of σ_τ^4 in (7) is $2\text{tr}(AA) = 2\sum_{ij} a_{ij}^2$;
- (b) The coefficient of σ_i^4 in (7) is $2a_{ii}^2/n_i(n_i - 1)$;
- (c) The coefficient of $\sigma_i^2\sigma_\tau^2$ in (7) is $4\sum_j a_{ij}^2/n_i$.
- (d) The coefficient of $\sigma_i^2\sigma_j^2$ in (7) is $4a_{ij}^2/n_i n_j$.

If we choose $a_{ii} = 1/k$, $a_{ij} = -1/k(k - 1)$, $i \neq j$, i.e.

$$A = \frac{1}{k-1}(I - \mathbf{1}_k(\mathbf{1}'_k \mathbf{1}_k)^{-1} \mathbf{1}'_k),$$

we have $\hat{\sigma}_{\tau 1}^2$. The choice $a_{ii} = \frac{n_i(n-n_i)}{n^2 - \sum n_i^2}$, $a_{ij} = -\frac{n_i n_j}{n^2 - \sum n_i^2}$, $i \neq j$, leads to $\hat{\sigma}_{\tau 2}^2$. The Cauchy-Schwarz inequality yields, if $\text{tr}A = 1$ and $A\mathbf{1} = \mathbf{0}$,

$$1^2 = (\text{tr}A)^2 = (\text{tr}\{A(I - \mathbf{1}_k(\mathbf{1}'_k \mathbf{1}_k)^{-1} \mathbf{1}'_k)\})^2 \leq \text{tr}(AA)\text{tr}(I - \mathbf{1}_k(\mathbf{1}'_k \mathbf{1}_k)^{-1} \mathbf{1}'_k)$$

and equality holds if $A = (I - \mathbf{1}_k(\mathbf{1}'_k \mathbf{1}_k)^{-1} \mathbf{1}'_k)/(k-1)$. Thus, the minimization of $\sum_{ij} a_{ij}^2$ subject to $\text{tr}A = 1$ and $A\mathbf{1} = \mathbf{0}$ results in the unique solution $a_{ii} = 1/k$, $a_{ij} = -1/k(k-1)$ for $i \neq j$ and the following proposition has been established.

Proposition 2.1. *Among translation invariant quadratic unbiased estimators of σ_τ^2 , $\hat{\sigma}_{\tau 1}^2$ given in (4) minimizes the coefficient of σ_τ^4 in the variance.*

In order to obtain improved quadratic estimators of $\hat{\sigma}_\tau^2$, consider the perturbed estimator

$$\hat{\sigma}_{\tau p}^2 = c\{Y'AY - \sum \frac{d_i a_{ii} S_i^2}{n_i(n_i - 1)}\}, \quad (12)$$

where A again satisfies $\text{tr}A = 1$ and $A\mathbf{1} = \mathbf{0}$, and c, d_1, \dots, d_k are to be suitably chosen. Obviously, the bias of such an estimator is given by

$$\begin{aligned} \text{bias}[\hat{\sigma}_{\tau p}^2] &= E[c(Y'AY - \sum \frac{d_i a_{ii} S_i^2}{n_i(n_i - 1)})] - \sigma_\tau^2 \\ &= -(1 - c)\sigma_\tau^2 + c \sum \frac{(1 - d_i) a_{ii} \sigma_i^2}{n_i}. \end{aligned} \quad (13)$$

and, similar to the derivation of (7), the variance of this estimator can be obtained as

$$\text{Var}[c(Y'AY - \sum \frac{d_i a_{ii} S_i^2}{n_i(n_i - 1)})] = 2c^2 \{ \sum_{ij} (\frac{\sigma_i^2}{n_i} + \sigma_\tau^2) (\frac{\sigma_j^2}{n_j} + \sigma_\tau^2) a_{ij}^2 + \sum \frac{a_{ii}^2 d_i^2 \sigma_i^4}{n_i^2 (n_i - 1)} \}.$$

Hence, the mean squared error of (12) is readily obtained as

$$\begin{aligned} \text{MSE}[\hat{\sigma}_{\tau p}^2] &= 2c^2 \{ \sum_{ij} (\frac{\sigma_i^2}{n_i} + \sigma_\tau^2) (\frac{\sigma_j^2}{n_j} + \sigma_\tau^2) a_{ij}^2 + \sum \frac{a_{ii}^2 d_i^2 \sigma_i^4}{n_i^2 (n_i - 1)} \} \\ &\quad + \{ (1 - c)\sigma_\tau^2 - c \sum \frac{(1 - d_i) a_{ii} \sigma_i^2}{n_i} \}^2. \end{aligned} \quad (14)$$

Moreover, from the calculations concerning $E[(\hat{\sigma}_\tau^2)^3]$ we obtain

$$\begin{aligned} E[(\hat{\sigma}_{\tau p}^2)^3] &= c^3 E[(Y'AY)^3] - 3c^3 E[(Y'AY)^2] \sum_i d_i a_{ii} / n_i \\ &\quad + 3c^3 E[Y'AY] \{ \text{Var}[\sum_i \frac{d_i a_i S_i^2}{n_i(n_i - 1)}] + (\sum_i d_i a_{ii} / n_i)^2 \} \\ &\quad - E[(\sum_i \frac{d_i a_i S_i^2}{n_i(n_i - 1)})^3]. \end{aligned}$$

All moments can easily be obtained from the previously presented moment formulas, for example,

$$\begin{aligned} E[(\sum_i \frac{d_i a_i S_i^2}{n_i(n_i - 1)} / \sigma_i^2)^3] &= \sum_i \frac{d_i^3 a_{ii}^3}{n_i^3 (n_i - 1)^3} + 3 \sum_{i, i \neq k} \frac{d_i^2 d_j a_{ii}^2 a_{jj} E[(S_i^2)^2] E[S_j^2]}{n_i^2 (n_i - 1)^2 n_j (n_j - 1)} \\ &\quad + 6 \sum_{i \neq j \neq k} \frac{d_i d_j d_k a_{ii} a_{jj} a_{kk}}{n_i n_j n_k}. \end{aligned}$$

As before, we observe the following facts:

- (a) The coefficient of σ_τ^4 in (14) is $2c^2 \sum_{ij} a_{ij}^2 + (1-c)^2$;
(b) The coefficient of σ_i^4 in (14) is $\frac{c^2 a_{ii}^2}{n_i^2} \{2 + \frac{2d_i^2}{n_i-1} + (1-d_i)^2\}$;
(c) The coefficient of $\sigma_i^2 \sigma_\tau^2$ in (14) is $4c^2 \sum_j a_{ij}^2/n_i - 2c(1-c)a_{ii}(1-d_i)/n_i$;
(d) The coefficient of $\sigma_i^2 \sigma_j^2$, $i \neq j$, in (14) is $\frac{4c^2 a_{ij}^2}{n_i n_j} + \frac{2c^2 a_{ii} a_{jj}}{n_i n_j} (1-d_i)(1-d_j)$.

Obviously, $d_i = d_{i0} = (n_i - 1)/(n_i + 1)$ minimizes the coefficient of σ_i^4 , $i = 1, \dots, k$, which is independent of A , and $c = c_0 = (1 + 2 \sum a_{ij}^2)^{-1}$ minimizes the coefficient of σ_τ^4 .

If we compare the MSE of $\hat{\sigma}_\tau^2$ and $\hat{\sigma}_{\tau p}^2$ we obtain

- (i) The difference, say a , of the coefficients of σ_τ^4 in (7) and (14) equals

$$a = 2 \sum_{ij} a_{ij}^2 - 2c^2 \sum_{ij} a_{ij}^2 - (1-c)^2 \geq 0. \quad (15)$$

- (ii) The difference, say $2b_i$, of the coefficients of σ_i^4 in (7) and (14) equals

$$2b_i = \frac{2a_{ii}^2}{n_i(n_i-1)} - \frac{c^2 a_{ii}^2}{n_i^2} \{2 + \frac{2d_i^2}{n_i-1} + (1-d_i)^2\} \geq 0. \quad (16)$$

- (iii) The difference, say e_i , of the coefficients of $\sigma_i^2 \sigma_\tau^2$ in (7) and (14) equals

$$e_i = 4 \sum_j a_{ij}^2/n_i - 4c^2 \sum_j a_{ij}^2/n_i + 2c(1-c)a_{ii}(1-d_i)/n_i \geq 0. \quad (17)$$

- (iv) The difference, say $2f_{ij}$, of the coefficients of $\sigma_i^2 \sigma_j^2$, $i \neq j$, in (7) and (14) equals

$$2f_{ij} = \frac{4a_{ij}^2}{n_i n_j} - \frac{4c^2 a_{ij}^2}{n_i n_j} - 2c^2 a_{ii} a_{jj} (1-d_i)(1-d_j)/n_i n_j, \quad i \neq j. \quad (18)$$

Thus, besides σ_τ^4 and σ_i^4 the coefficient of $\sigma_i^2 \sigma_\tau^2$ is also smaller in the MSE of the estimator $\hat{\sigma}_{\tau p}^2$ compared to the unbiased estimator $\hat{\sigma}_\tau^2$ with $\text{tr}A = 1$ and $A\mathbf{1} = \mathbf{0}$. This may not be true for the coefficient of $\sigma_i^2 \sigma_j^2$ for $i \neq j$. However, an improvement in MSE over the unbiased estimator $\sigma_{\tau 1}^2$ is still possible. The risk difference, i.e. the difference in MSE, is non-negative if and only if $(u, v')B \begin{pmatrix} u \\ v \end{pmatrix} \geq 0$, where for all $u = \sigma_\tau^2$ and $v' = (\sigma_1^2, \dots, \sigma_k^2)$ with non-negative components and

$$B = \begin{pmatrix} a & b_1 & b_2 & \dots & b_k \\ b_1 & e_1 & f_{12} & \dots & f_{1k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_k & f_{k1} & f_{k2} & \dots & e_k \end{pmatrix}.$$

Thus, B has to be copositive. However, a and b_i are non-negative and therefore it is enough to consider when

$$v' B_1 v \geq 0,$$

where B_1 is the submatrix of B with the first column and row removed.

To give some useful necessary and sufficient conditions for B_1 to be copositive seems difficult. However, if $f_{ij} \geq 0$ then obviously B_1 is copositive, and with $c = (1 + 2 \sum_{ij} a_{ij}^2)^{-1}$ and $d_i = (n_i - 1)/(n_i + 1)$ this holds if and only if

$$\sum_{ij} a_{ij}^2 \geq -\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{2} \frac{a_{ii} a_{jj}}{a_{ij}^2} \frac{1}{(n_i + 1)(n_j + 1)}}, \quad (19)$$

for all i, j .

From now on we will consider some special cases. Let us start by assuming that all f_{ij} are equal, i.e. $f = f_{ij}$, as well as for $i = 1, 2, \dots, k$, $b = b_i$ and $e = e_i$. Thus

$$B = \begin{pmatrix} a & b \mathbf{1}'_{\mathbf{k}} \\ b \mathbf{1}_{\mathbf{k}} & (e - f) I_k + f \mathbf{1}_{\mathbf{k}} \mathbf{1}'_{\mathbf{k}} \end{pmatrix}. \quad (20)$$

Since a and b are non-negative it is enough to focus on

$$v'((e - f) I_k + f \mathbf{1}_{\mathbf{k}} \mathbf{1}'_{\mathbf{k}}) v. \quad (21)$$

Since $(e - f) I_k + f \mathbf{1}_{\mathbf{k}} \mathbf{1}'_{\mathbf{k}} = e I_k + f(\mathbf{1}_{\mathbf{k}} \mathbf{1}'_{\mathbf{k}} - I_k)$ and if $f \geq 0$ we always have that (21) is non-negative. Now turning to the case when $f < 0$ it is observed that by the Cauchy-Schwarz inequality

$$v'((e - f) I_k + f \mathbf{1}_{\mathbf{k}} \mathbf{1}'_{\mathbf{k}}) v \geq (e + (k - 1)f) v' v.$$

Thus, if $f < 0$ and (21) should be non-negative $(e + (k - 1)f) \geq 0$ must hold.

Theorem 2.2. *If $f = f_{ij}$, $b = b_i$ and $e = e_i$, $i, j = 1, 2, \dots, k$, in (16) – (18) a necessary and sufficient condition for the risk difference of (7) and (14) to be non-negative is that either $f \geq 0$ or if $f < 0$ then $(e + (k - 1)f) \geq 0$ must hold.*

When $n_i = n_0$, $i = 1, 2, \dots, k$, i.e. the balanced case, and $d_0 = (n_0 - 1)/(n_0 + 1)$ the two special unbiased estimators $\hat{\sigma}_{\tau_1}^2$ and $\hat{\sigma}_{\tau_2}^2$, given by (4) and (5), are identical and will be compared to $\hat{\sigma}_{\tau_p}^2$, given by (12).

Thus, with $a_{ii}=1/k$, $a_{ij} = -1/k(k-1)$, $i \neq j$, resulting in $c_0 = (k-1)/(k+1)$, and the above choice of d_0 , it follows from (15)–(18) that

$$a = \frac{4}{(k-1)(k+1)}, \quad (22)$$

$$b = \frac{2}{n_0 k(k-1)} - \frac{2(k-1)}{n_0 k(k+1)^2} \frac{n_0 - 1}{n_0 + 1}, \quad (23)$$

$$e = \frac{2}{k^2 n_0(n_0 - 1)} - \frac{2(k-1)^2}{n_0^2 k^2 (k+1)^2} \frac{n_0 + 2}{n_0 + 1}, \quad (24)$$

$$f = \frac{2}{k^2(k-1)^2 n_0^2} - \frac{2}{k^2(k+1)^2 n_0^2} \left\{ 1 + \frac{2(k-1)^2}{(n_0 + 1)^2} \right\}. \quad (25)$$

Now $f \geq 0$ means that $n \geq \frac{1}{\sqrt{k}}(k-1)^2 - 1$ holds, and if $f \leq 0$ the necessary and sufficient condition for risk improvement is that $(e + (k+1))f > 0$. The quantity $(e + (k+1))f$ can be shown to be to a third degree polynomial in k and n .

Theorem 2.3. *If $n_i = n_0$, $i = 1, 2, \dots, k$, and $a_{ii} = 1/k$, $a_{ij} = -1/k(k-1)$, $i \neq j$, a necessary and sufficient condition for the risk difference of $\hat{\sigma}_{\tau_1}^2 = \hat{\sigma}_{\tau_2}^2$, given by (4) and (5), and $\hat{\sigma}_{\tau_p}^2$, given by (12), to be positive is that either $f \geq 0$ or if $f < 0$ then $(e + (k-1))f > 0$ must hold, where the constants a , b , e and f are given by (22) – (25).*

If we briefly look at the unbalanced case, with $c = (k-1)/k+1$ and $d_i = (n_i - 1)/(n_i + 1)$, the estimator $\hat{\sigma}_{\tau_p}^2$ has a smaller MSE than $\hat{\sigma}_{\tau_1}^2$, if the following condition holds:

$$(n_i + 1)(n_j + 1) \geq (k-1)^4/2k, \quad i \neq j. \quad (26)$$

A sufficient condition for (26) to hold is $\min_i(n_i) + 1 \geq (k-1)^2/\sqrt{2k}$.

The theorem is illustrated in Section 4.

2.2 Edgeworth expansion

In order to compare the densities of $\hat{\sigma}_\tau^2$ with $\hat{\sigma}_{\tau p}^2$, ordinary Edgeworth expansions will be performed. From Kollo and von Rosen (1998) it follows that

$$\begin{aligned} f_y(x) &\approx f_x(x) + (E[y] - E[x])f_x^1(x) \\ &\quad + \frac{1}{2}\{Var[y] - Var[x] + (E[y] - E[x])^2\}f_x^2(x) \\ &\quad - \frac{1}{6}\{c_3[y] - c_3[x] + 3(Var[y] - Var[x])(E[y] - E[x]) \\ &\quad + ((E[y] - E[x])^3)\}f_x^3(x), \end{aligned}$$

where $f_y(x)$, $f_x(x)$ are the densities of y and x , respectively, $f_x^i(x)$, $i = 1, 2, 3$ is the i th derivative of $f_x(x)$ and $c_3[y]$ is the third order cumulant of y . Now, let $f_x(x)$ represent the normal density with mean σ_τ^2 and variance equal to $Var[\hat{\sigma}_\tau^2]$. Of course we could have chosen densities other than the normal, for example the chi-square, but there is no appropriate criteria for choosing between different distributions and therefore the normal was used. By using the normal distribution with mean and variance suggested above we obtain

$$f_{\hat{\sigma}_\tau^2}(x) \approx (1 + \frac{1}{6}c_3[\hat{\sigma}_\tau^2]((x - \sigma_\tau^2)^3 - 3Var[\hat{\sigma}_\tau^2])/Var[\hat{\sigma}_\tau^2]^3)f_x(x).$$

Here $c_3[\hat{\sigma}_\tau^2]$ is calculated via

$$c_3[\hat{\sigma}_\tau^2] = E[(\hat{\sigma}_\tau^2)^3] - 3E[(\hat{\sigma}_\tau^2)^2]\sigma_\tau^2 + 2(\sigma_\tau^2)^3.$$

The approximate density for $\hat{\sigma}_{\tau p}^2$ equals

$$\begin{aligned} f_{\hat{\sigma}_{\tau p}^2}(x) &\approx (1 - bias[\hat{\sigma}_{\tau p}^2](x - \sigma_\tau^2)/Var[\hat{\sigma}_\tau^2]) \\ &\quad + \frac{1}{2}(Var[\hat{\sigma}_{\tau p}^2] - Var[\hat{\sigma}_\tau^2] + bias[\hat{\sigma}_{\tau p}^2]^2)((x - \sigma_\tau^2)^2 - Var[\hat{\sigma}_\tau^2])/Var[\hat{\sigma}_\tau^2]^2 \\ &\quad + \frac{1}{6}\{c_3[\hat{\sigma}_{\tau p}^2] + 3(Var[\hat{\sigma}_{\tau p}^2] - Var[\hat{\sigma}_\tau^2])bias[\hat{\sigma}_{\tau p}^2] + bias[\hat{\sigma}_{\tau p}^2]^3\} \\ &\quad \times ((x - \sigma_\tau^2)^3 - 3Var[\hat{\sigma}_\tau^2])/Var[\hat{\sigma}_\tau^2]^3 f_x(x). \end{aligned}$$

Moreover, $c_3[\hat{\sigma}_{\tau p}^2]$ is calculated via

$$c_3[\hat{\sigma}_{\tau p}^2] = E[(\hat{\sigma}_{\tau p}^2)^3] - 3E[(\hat{\sigma}_{\tau p}^2)^2]E[\hat{\sigma}_{\tau p}^2] + 2(E[\hat{\sigma}_{\tau p}^2])^3.$$

When comparing the approximate densities we get

$$\begin{aligned}
f_{\hat{\sigma}_{\tau p}^2}(x) - f_{\hat{\sigma}_{\tau}^2}(x) &\approx \text{bias}[\hat{\sigma}_{\tau p}^2](x - \sigma_{\tau}^2)/\text{Var}[\hat{\sigma}_{\tau}^2]f_x(x) \\
&- \frac{1}{2}(\text{MSE}[\hat{\sigma}_{\tau p}^2] - \text{Var}[\hat{\sigma}_{\tau}^2])((x - \sigma_{\tau}^2)^2 \\
&- \text{Var}[\hat{\sigma}_{\tau}^2])/\text{Var}[\hat{\sigma}_{\tau}^2]^2 f_x(x). \tag{27}
\end{aligned}$$

Observe that the difference $\text{MSE}[\hat{\sigma}_{\tau p}^2] - \text{Var}[\hat{\sigma}_{\tau}^2]$ is negative. With the help of (27) we may evaluate the impact of $\text{bias}[\hat{\sigma}_{\tau p}^2]$ and $\text{MSE}[\hat{\sigma}_{\tau p}^2]$. If bias is not severe and if $\text{MSE}[\hat{\sigma}_{\tau p}^2]$ is significant smaller than $\text{Var}[\hat{\sigma}_{\tau}^2]$ we observe that the improved variance estimator is less skewed. Indeed this is a very good estimator although the new estimator may be somewhat biased. In Figure 1 in the next paragraph the distributions are presented in a particular simulation experiment.

2.3 Bayes estimators of σ_{τ}^2 ; a comparison study

Here we will briefly compare, in a small simulation study and in the balanced case, the unbiased estimator, the improved estimator, Bayesian estimators based on different priors for σ_{τ}^2 and the maximum likelihood estimator. As Bayes estimator we are going to take the posterior mean. The model together with a vague prior can easily be implemented in Bayesian estimation programs such as for example WinBUGS (Spiegelhalter et al., 2003; Sturtz et al., 2005) and proc mixed in SAS (SAS Institute, 2005). In WinBUGS, via suitable choices of parameters in the inverse gamma distribution we used a non-informative Jeffrey prior, where it was assumed that the variance parameters were independently distributed. For the mean a flat prior was supposed to hold, i.e. a constant. The Metropolis-Hastings algorithm was used to generate a chain consisting of 10000 observations where the last 1000 were used for calculating the posterior mean. When performing a Bayesian analysis in proc mixed in SAS either a complete flat prior was used, i.e. the posterior density equals the likelihood, or based on the information matrix a Jeffrey prior was used as a simultaneous prior for all variances which is somewhat different from WinBUGS, where independence was assumed to hold. For the mean a constant prior was used. In proc mixed the importance sampling algorithm was used and once again the posterior mean was calculated via the last 1000 observations in a generated chain of 10000 observation. In order to evaluate the distributions of the estimators 1000 data sets were generated.

In the artificially created data sets 9 labs with 2 repeated measurements on each were considered. Moreover, the variation differed between the labs,

i.e. we assumed a heteroscedastic model. Thus, the situation is rather extreme with 11 parameters and 18 observations with pairwise dependency. In Table 1 the set-up of the simulation study is given. The choice of parameter values follows the data presented in Vangel & Rukhin (1999, Table 1).

Table 1. Parameter values used in the simulation study.

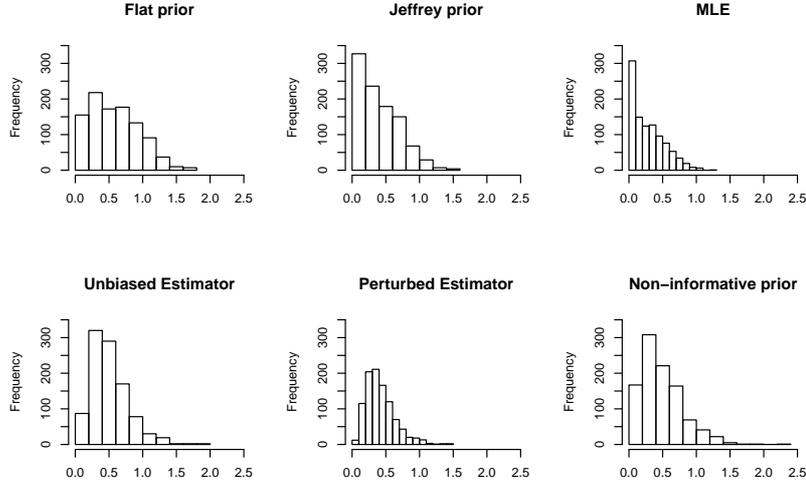
n_i	μ	σ_τ^2	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8	σ_9
2	13	.50	.03	.23	.32	.07	.34	.32	.06	.21	.10

It follows from Table 2 (see below) that the unbiased estimator is close to $\hat{\sigma}_{bn}^2$ and the perturbed estimator $\hat{\sigma}_{\tau p}^2$ is close to $\hat{\sigma}_{bJ}^2$. Moreover, $\hat{\sigma}_{bf}^2$ slightly overestimates the parameter whereas it is clear that the maximum likelihood is too biased to be of any use. We may also observe that $\hat{\sigma}_\tau^2$, $\hat{\sigma}_{\tau p}^2$ and $\hat{\sigma}_{bn}^2$ have similar quadratic variations. However, if we look at Figure 1, which is showing the posterior distributions for all estimators one can see that the distributions look rather different. The unbiased estimator, the perturbed one and $\hat{\sigma}_{bn}^2$ are more symmetric than the others with a smallest tail for $\hat{\sigma}_{\tau p}^2$. Thus, despite of some bias, $\hat{\sigma}_{\tau p}^2$ is competitive to the others.

Table 2. The mean, median, standard deviation, min and max values of 1000 simulations are presented, $\hat{\sigma}_\tau^2$ is given in (4), $\hat{\sigma}_{\tau p}^2$ is given in (12), $\hat{\sigma}_{bn}^2$ is the mean posterior Bayes estimator with non-informative prior calculated in WinBUGS, $\hat{\sigma}_{bf}^2$ and $\hat{\sigma}_{bJ}^2$ are the mean posterior Bayes estimators calculated in proc mixed in SAS with flat and Jeffrey prior distributions and $\hat{\sigma}_{MLE}^2$ is the maximum likelihood estimator calculated in proc mixed in SAS.

estimator	mean	median	std. dev.	minimum	maximum
$\hat{\sigma}_\tau^2$	0.507	0.460	0.272	0.045	1.840
$\hat{\sigma}_{\tau p}^2$	0.419	0.382	0.217	0.064	1.490
$\hat{\sigma}_{bn}^2$	0.381	0.433	0.212	0.109	1.537
$\hat{\sigma}_{bJ}^2$	0.406	0.344	0.299	0.013	1.568
$\hat{\sigma}_{bf}^2$	0.589	0.543	0.356	0.009	1.760
$\hat{\sigma}_{MLE}^2$	0.278	0.236	0.247	0.0001	1.241

Figure 1. The distribution (posterior for the Bayes estimators) based on 1000 simulations are presented for $\hat{\sigma}_\tau^2$, given in (4) and $\hat{\sigma}_{\tau p}^2$, given in (12), the



mean posterior Bayes estimator with non-informative prior, $\hat{\sigma}_{bn}^2$, calculated in WinBUGS, $\hat{\sigma}_{bf}^2$ and $\hat{\sigma}_{bJ}^2$, the mean posterior Bayes estimators calculated in proc mixed in SAS with flat and Jeffrey prior distributions, respectively, and $\hat{\sigma}_{MLE}^2$, the maximum likelihood estimator, calculated in proc mixed in SAS.

3 Improved estimators of within lab variances $\sigma_1^2, \dots, \sigma_k^2$

In this section, following ideas of Stein (1964) and Mathew et al. (1992), we derive improved estimates of the error variances $\sigma_1^2, \dots, \sigma_k^2$, in the sense of providing estimates with a smaller mean squared error. Towards this end, we first prove a general result.

As in section 2 let $Y_i \sim N[\mu, \sigma_\tau^2 + \theta_2]$ be independent of $S_i^2 \sim \theta^2 \chi_{m_i}^2$ where, $m_i = n_i - 1$, $\sigma_\tau^2 \geq 0$ and $\theta_2 = \sigma_i^2/n_i > 0$. Assume $\mu, \sigma_\tau^2, \theta_2$ all unknown. For estimating θ_2 , define $\hat{\theta}_2 = S_i^2/(m_i + 2)$ and $\tilde{\theta}_2 = (S_i^2 + Y_i^2)/(m_i + 3)$ if $(S_i^2 + Y_i^2)/(m_i + 3) \leq S_i^2/(m_i + 2)$, and $S_i^2/(m_i + 2)$, otherwise. Then the following result holds.

Proposition 3.1. $\tilde{\theta}_2$ has a smaller MSE than $\hat{\theta}_2$ uniformly in the unknown parameters.

Proof. Let $W = Y_i^2/S_i^2$, and consider the estimate $T = \phi(W)S_i^2$ where $\phi(\cdot)$ is a nonnegative function of W . Note that both $\hat{\theta}_2$ and $\tilde{\theta}_2$ are special cases of T .

Then the MSE of T can be expressed as

$$MSE(T) = E[T - \theta_2]^2 = E[E(\phi(W)S_i^2 - \theta_2)^2|W].$$

For a given $W = w$, an optimum choice of $\phi(w)$, minimizing the above MSE, is given by

$$\phi_{opt}(w) = E[S_i^2|w]\theta_2/E(S_i^2|w).$$

Suppose that

$$\phi_{opt}(w) \leq (1 + w)/(m_i + 3), \quad (28)$$

holds uniformly in the parameters. Because of the convexity of $MSE(T)$, it follows then that, given any $\phi(w)$, defining $\phi_0(w) = \min[\phi(w), (1+w)/(m_i+3)]$ results in $T_0 = \phi_0(W)S_i^2$ having a smaller MSE than T . The theorem follows upon taking $\phi(w) = 1/(m_i + 2)$.

Finally we prove (28). Noting that $Y^2/(\theta_1 + \theta_2)$, where the index i is omitted, has a noncentral chisquare distribution with 1 df and non-centrality parameter $\lambda = \mu^2/(\theta_1 + \theta_2)$, the joint pdf of Y^2 and V can be written as

$$\begin{aligned} f(y^2, v) &= K_m[\exp(-y^2/2(\theta_1 + \theta_2)) \sum_{j=0}^{\infty} d_j(y^2)^{j-\frac{1}{2}}(\theta_1 + \theta_2)^{-(j+\frac{1}{2})}] \\ &\times [\exp(-v/2\theta_2)v^{\frac{m}{2}-1}(\theta_2)^{-m/2}], \end{aligned} \quad (29)$$

where K_m and d_j are constants whose actual values are not necessary in our context. Making the transformation from (Y^2, V) to $(W = \frac{Y^2}{V}, V)$ yields the joint pdf of W, V as

$$\begin{aligned} f(w, v) &= K_m[v\exp(-wv/2(\theta_1 + \theta_2)) \sum_{j=0}^{\infty} d_j(wv)^{j-\frac{1}{2}}(\theta_1 + \theta_2)^{-(j+\frac{1}{2})}] \\ &\times [\exp(-v/2\theta_2)v^{\frac{m}{2}-1}(\theta_2)^{-m/2}]. \end{aligned} \quad (30)$$

It is now directly verified that

$$\phi_{opt}(w) = \frac{E[V|w]\theta_2}{E[V^2|w]}.$$

□

Remark. Stein's (1964) result follows as a special case when θ_1 is taken as 0. The result in Mathew et al. (1992) also follows as a special case when it is assumed that $\mu = 0$.

4 Applications

There are many practical applications of the assumed basic model (1) in the literature with a main focus on the estimation of the common mean μ . For example see Jordan and Krishnamoorthy (1996), Yu et al. (1999), Rukhin and Vangel (1998), and Vangel and Rukhin (1999). We consider two of them, i.e. Rukhin and Vangel (1998), and Vangel and Rukhin (1999), with the purpose to illustrate the estimation of the variance components.

Example 4.1. In this example we examine the data reported in Willie and Berman (1995) and analyzed in Rukhin and Vangel (1998) about concentration of several trace metals in oyster tissues. The data appear in Rukhin and Vangel (1998; Table 1). Here $k = 28$ and $n_3 = 2, n_1 = n_2 = n_4 \dots = n_{28} = 5$. Our object is to provide efficient estimates of σ_τ^2 and $\sigma_1^2, \dots, \sigma_{28}^2$. However, in order to have balanced data we exclude from the analysis the laboratory with $n_3 = 2$. Following the conditions and analysis in section 2, our proposed estimates equal (remember that in the balanced case $\hat{\sigma}_{\tau_1}^2 = \hat{\sigma}_{\tau_2}^2$) $\hat{\sigma}_{\tau_p}^2 = 1.55$ whereas $\hat{\sigma}_{\tau_1}^2 = 1.63$. Here $f < 0$ and $(e + (k - 1)f) > 0$ and thus $\hat{\sigma}_{\tau_p}^2$ improves $\hat{\sigma}_{\tau_1}^2$. We may note that for the complete data set, i.e. $k = 28$, $f < 0$ and we get $\hat{\sigma}_{\tau_1}^2 = 1.89$ and its improved estimator equals 1.80. Moreover, $\hat{\sigma}_{\tau_1}^2 = 1.74$ versus 1.66 for the improved version. Turning to the within (laboratory) variances we have that the observations are relatively large in relation to the variation. Therefore Y_i^2 in the definition of the estimators is larger than V_i and it follows that the estimators $\hat{\theta}_2$ and $\tilde{\theta}_2$ are identical.

Example 4.2. In this example we examine the data reported in Li and Cardozo (1994) and analyzed in Vangel and Rukhin (1999) about dietary fibre in apples. Here $k = 9$ and $n_1 = n_2 = \dots = n_9 = 2$, i.e. we have again a balanced case. Following the analysis in section 2, our proposed estimates equal $\hat{\sigma}_{\tau_p}^2 = 0.39$ whereas $\hat{\sigma}_{\tau_1}^2 = 0.47$. Here $f < 0$ and $(e + (k - 1)f) > 0$ and thus $\hat{\sigma}_{\tau_p}^2$ improves $\hat{\sigma}_{\tau_1}^2$. For the within variances we observe that $\hat{\theta}_2$ and $\tilde{\theta}_2$ are identical.

Acknowledgements

The work of T. Nahtman was supported by the grant GMTMS6702 of Estonian Research Foundation.

References

- [1] Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society*, **4**, 102–118.
- [2] Cochran, W.G. (1954). The combination of estimates from different experiments. *Biometrics*, **10**, 101–129.
- [3] Fairweather, W.R. (1972). A method of obtaining an exact confidence interval for the common mean of several normal populations. *Applied Statistics*, **21**, 229–233.
- [4] Hartley, H.O. and Rao, J.N.K. (1978). Estimation of nonsampling variance components in sample surveys. *In: Survey Sampling and Measurement* (ed. N.K. Namboodiri), 35–43, New York.
- [5] Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, **72**, 320–340.
- [6] Jordan, S.M. and Krishnamoorthy, K. (1996). Exact confidence intervals for the common mean of several normal populations. *Biometrics*, **52**, 77–86.
- [7] Kelly, R.J. and Mathew, T. (1993). Improved estimators of variance components with smaller probability of negativity. *Journal of the Royal Statistical Society, Series B*, **55**, 897–911.
- [8] Kelly, R.J. and Mathew, T. (1994). Improved nonnegative estimation of variance components in some mixed models with unbalanced data. *Technometrics*, **36**, 171–181.
- [9] Kollo, T. and von Rosen, D. (1998). A unified approach to the approximation of multivariate densities. *Scandinavian Journal of Statistics*, **25**, 93–109.
- [10] LaMotte, L.R. (1973). On non-negative quadratic unbiased estimation of variance components. *Journal of the American Statistical Association*, **680**, 728–730.
- [11] Li, B.W. and Cardozo, M.S. (1994). Determination of total dietary fiber in foods and products with little or no starch, nonenzymatic-gravimetric method: collaborative study. *Journal of Association of Analytical Chemists International*, **77**, 687–689.
- [12] Mathew, T., Sinha, B.K. and Sutradhar, B. (1992). Improved estimation of error variance in general balanced mixed models. *Statistics and Decisions*, **10**, 227–238.

- [13] Rao, J.N.K. (1977). Maximum likelihood approaches to variance component estimation and to related problems: Comment. *Journal of the American Statistical Association*, **72**, 338–339.
- [14] Rao, P.S.R.S., Kaplan, J. and Cochran, W.G. (1981). Estimators for the one-way random effects model with unequal error variances. *Journal of the American Statistical Association*, **76**, 89–97.
- [15] Rukhin, A.L. and Vangel, M.G. (1998). Estimation of a common mean and weighted means statistics. *Journal of the American Statistical Association*, **93**, 303–308.
- [16] SAS Institute. (2005). SAS OnlineDoc 9.1.3: SAS/STAT 9.1 user’s guide. Cary, NC: SAS Institute, Inc. Retrieved September 20, 2006, from SAS OnlineDoc 9.1.3 web site: <http://v9doc.sas.com/sasdoc/>.
- [17] Spiegelhalter, D., Thomas, A. and Best, N. (2003). WinBUGS Version 1.4 User Manual. Medical Research Council Biostatistics Unit, Cambridge, UK.
- [18] Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics*, **16**, 155–160.
- [19] Vangel, M.G. and Rukhin, A.L. (1999). Maximum likelihood analysis for heteroscedastic one-way random effects ANOVA in interlaboratory studies. *Biometrics*, **55**, 129–136.
- [20] Willie, S. and Berman, S. (1996). NOAA National Status and Trends Program tenth round intercomparison exercise results for trace metals in marine sediments and biological tissues. *NOAA Technical Memorandum NOS ORCA 93*. NOAA/NOS/ORCA, Silver Spring, MD.
- [21] Yu, P.L.H., Sun, Y. and Sinha, B.K. (1999). On exact confidence intervals for the common mean of several normal populations. *Journal of Statistical Planning and Inference*, **81**, 263–277.