COMPARISON OF METHODS OF ESTIMATING VARIANCE COMPONENTS IN BALANCED TWO-WAY RANDOM NESTED DESIGNS

Ebru KAYA BAŞAR 1, Mehmet Ziya FIRAT 2

Abstract

The objective of this paper is to compare different methods of estimating variance components, such as analysis of variance (ANOVA), maximum likelihood (ML), restricted maximum likelihood (REML) and Bayesian. The ANOVA method for estimating variance components is based on equating the mean squares to their expectations. However, the problem with this method is that it gives negative estimates of variance components. This can be overcome by the use of likelihood based methods and Bayesian. In this study, four different methods of estimating variance components were compared and also demonstrated how these methods overcome the problem of negative estimates of variance components in balanced two-way random nested designs.

Keywords: Variance components, Bayesian analysis, REML, ML.

1. INTRODUCTION

An understanding of variability and the nature of measurement error is a fundamental issue for the researchers. Estimating variance components in experimental design is a way to assess the amount of variation in a dependent variable which is associated with one or more random effects variables. Variance components are used extensively in developing many of the basic concepts of many fields such as animal breeding. Sources of variation in the analysis of variance context were partitioned into their expected components, which were particularly useful to the researchers [1].

A wide array of methods has been developed for estimating variance components, for example, Analysis of Variance (ANOVA) and likelihood based methods, such as, Maximum Likelihood (ML)

---

1 Akdeniz University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, Antalya, Turkey.
E-mail: ebrukaya10@gmail.com.tr

2 Corresponding Author: Akdeniz University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, Antalya, Turkey.
E-mail: mzfirat@akdeniz.edu.tr
and Restricted Maximum Likelihood (REML), and Bayesian methods. The ANOVA method for estimating variance components is based on equating the mean squares to their expectations. However, the problem with this method is that ANOVA estimates may lie outside the parameter space so it gives negative estimates of variance components. In light of this weakness, alternative methods were required and methods based on likelihood have become common [1,2]. The ML estimation of variance components does not take into account the loss of degrees of freedom caused by estimation of the fixed effects. This can be worrying when the model includes fixed effects with many levels. These problems can be overcome by the use of REML method which was first proposed by Thompson [3] and generalized by Patterson and Thompson [4]. REML method is more useful than ML method in some aspects. This approach allows for several random factors in the model and is based on maximizing with respect to the variances only the part of the likelihood function that does not depend on fixed effects. In contrast to ANOVA estimations, estimators which are based on likelihood functions are functions of every sufficient statistic and are consistent, asymptotically normal and efficient.

An alternative to the methods mentioned above is the Bayesian method for estimating the variance components. The Bayesian framework was introduced by Thomas Bayes and Bayesian estimation was first used by Laplace in 1786. Estimation of variance components using Bayesian framework is found in Hill [5] and have been applied to animal breeding by many researchers [2,6,7,8]. The likelihood corresponding to the data and the prior distribution of the parameters are the components of posterior distribution which is required in order to carry out a Bayesian analysis. Inferences about each parameter are based on the corresponding marginal posterior distributions. For many models, the joint distribution of the dispersion parameters can be derived, but numerical integration techniques are required to obtain the marginal posterior distribution of functions of interest. Gibbs sampling is a numerical integration method introduced by Geman and Geman [9] which operates by generating samples from a sequence of full conditional distributions. Bayesian method allows for the formal incorporation of prior knowledge into the process of inference. Within the Bayesian framework, it is possible to obtain the marginal posterior distributions of parameters of interest using Gibbs sampling. This, in general, is not possible in the classical approaches.

The objective of this study is to compare four different methods of estimating variance components ANOVA, ML, REML and Bayesian and also to demonstrate how these methods overcome the problem of negative estimates of variance components in balanced two-way random nested designs.

2. MODEL

In some experiments samples must be chosen in two or more steps. If we consider an experiment with two factors, let factor A has a levels and factor B b levels within each level of factor A. The levels of B are nested within levels of A and within each level of B some number n random samples are chosen. The model for this two-way random nested design is given as follows

\[ y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + e_{ijk} \quad i = 1,2,\ldots,a; \, j = 1,2,\ldots,b; \, k = 1,2,\ldots,n \]  

where \( y_{ijk} \) is the observation \( k \) in level \( i \) of factor A and level \( j \) of factor B; \( \mu \) is the overall mean, \( \alpha_i \) is the effect of level \( i \) of factor A; \( \beta_{j(i)} \) is the effect of level \( j \) of factor B within level \( i \) of factor A and \( e_{ijk} \) is the random error with \( e_{ijk} \sim N(0, \sigma^2_e) \). Also, \( \alpha_i \) and \( \beta_{j(i)} \) are both random effects in the model and these parameters assumed to be normally distributed \( \alpha_i \sim N(0, \sigma^2_\alpha) \), \( \beta_{j(i)} \sim N(0, \sigma^2_\beta) \), respectively.
3. METHODS OF ESTIMATING VARIANCE COMPONENTS

3.1. Analysis of Variance (ANOVA)

The analysis of variance (ANOVA) estimation of the variance components consists of equating mean squares to their respective expected value. The resulting equations are solved for the variance components and the solutions are the estimators of $\sigma^2_a$, $\sigma^2_{\beta(a)}$ and $\sigma^2_e$. These ANOVA estimators are given as

$$\hat{\sigma}^2_{e,\text{ANOVA}} = MS_E, \quad \sigma^2_{\beta(a),\text{ANOVA}} = \frac{MS_{\beta(A)} - MS_E}{n}, \quad \hat{\sigma}^2_{a,\text{ANOVA}} = \frac{MS_A - MS_{\beta(A)}}{bn}$$

(2)

where

$$MS_E = \frac{1}{ab(n-1)} \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ij})^2$$

$$MS_{\beta(A)} = \frac{1}{a(b-1)} \sum_{i=1}^{a} \sum_{j=1}^{b} n(\bar{y}_{ij} - \bar{y}_{..})^2$$

$$MS_A = \frac{1}{a-1} \sum_{i=1}^{b} bn(\bar{y}_{..i} - \bar{y}_{..})^2$$

However, it is known that variance estimates may lie outside their parameter space. Negative variance estimation problem is the most important issue for ANOVA method [10].

3.2. Maximum Likelihood (ML)

The conditional distribution of $\{y_{ijk}\}$ given $\mu$, $\{\alpha_i\}$, $\{\beta_{j(i)}\}$ and $\sigma^2_e$ is

$$\left(\{y_{ijk}\} \mid \mu, \{\alpha_i\}, \{\beta_{j(i)}\}, \sigma^2_e\right) \sim N \left( \mu + \alpha_i + \beta_{j(i)}, \sigma^2_e \right)$$

Similarly, the conditional distributions of $\{\alpha_i\}$ and $\{\beta_{j(i)}\}$ are respectively,

$$\left(\{\alpha_i\} \mid \sigma^2_a\right) \sim N \left( 0, \sigma^2_a \right) \quad \text{and} \quad \left(\{\beta_{j(i)}\} \mid \sigma^2_{\beta(a)}\right) \sim N \left( 0, \sigma^2_{\beta(a)} \right)$$

Under these assumptions of normality, the likelihood function is as follows:

$$L(\mu, \{\alpha_i\}, \{\beta_{j(i)}\}, \sigma^2_a, \sigma^2_{\beta(a)}, \sigma^2_e \mid \{y_{ijk}\}) \propto \left(\sigma^2_e\right)^{-n/2} \exp \left\{ \frac{1}{2} \left[ \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \frac{(y_{ijk} - \mu - \alpha_i - \beta_{j(i)})^2}{\sigma^2_e} + \sum_{i=1}^{a} \frac{\alpha_i^2}{\sigma^2_a} + \sum_{j=1}^{b} \frac{\beta_{j(i)}^2}{\sigma^2_{\beta(a)}} \right] \right\}$$

(3)

The log-likelihood function can be obtained from the equation (3) and then is solved for its ML estimators by equating to zero and taking the partial derivative of $\ln(L)$ with respect to $\mu$, $\left(bn\sigma^2_a + n\sigma^2_{\beta(a)} + \sigma^2_e\right)$, $\left(n\sigma^2_{\beta(a)} + \sigma^2_e\right)$ and $\sigma^2_e$. We can demonstrate these ML solutions as follows:

\[ \hat{\sigma}_{\alpha,ML}^2 = \frac{1 - \frac{1}{a}}{bn} MS_A - MS_{B(A)}, \quad \hat{\sigma}_{\beta(a),ML}^2 = \frac{MS_{B(A)} - MS_E}{n} \quad \text{and} \quad \hat{\sigma}_{e,ML}^2 = MS_E \]

Table 1. ML estimators of the variance components

<table>
<thead>
<tr>
<th>Conditions</th>
<th>( \hat{\sigma}_e^2 )</th>
<th>( \hat{\sigma}_{\beta(a)}^2 )</th>
<th>( \hat{\sigma}_\alpha^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\sigma}<em>a^2 \geq 0, \hat{\sigma}</em>{\beta(a)}^2 \geq 0 )</td>
<td>( MS_E )</td>
<td>( MS_{B(A)} - MS_E )</td>
<td>( \frac{1 - \frac{1}{a}}{bn} MS_A - MS_{B(A)} )</td>
</tr>
<tr>
<td>( \hat{\sigma}<em>a^2 \geq 0, \hat{\sigma}</em>{\beta(a)}^2 &lt; 0 )</td>
<td>( \frac{SS_E + SS_{B(A)}}{a(bn - 1)} )</td>
<td>0</td>
<td>( \frac{1}{bn} MS_E - \hat{\sigma}_e^2 )</td>
</tr>
<tr>
<td>( \hat{\sigma}<em>a^2 &lt; 0, \hat{\sigma}</em>{\beta(a)}^2 \geq 0 )</td>
<td>( \frac{1}{n} \left( \frac{SS_A + SS_{B(A)}}{ab} - MS_E \right) )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \hat{\sigma}<em>a^2 &lt; 0, \hat{\sigma}</em>{\beta(a)}^2 &lt; 0 )</td>
<td>( \frac{SS_{B(A)} + SS_A + SS_E}{abn} )</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

However, these are only the equations of ML solutions, not the ML estimators, which can be shown as \( \bar{\mu}, \bar{\sigma}_e^2, \bar{\sigma}_{\beta(a)}^2 \) and \( \bar{\sigma}_\alpha^2 \) because, especially, \( \hat{\sigma}_a^2 \) and \( \hat{\sigma}_{\beta(a)}^2 \) can be negative which is contrast to the definition of variance. So, generally if \( \hat{\sigma}_a^2 \) and \( \hat{\sigma}_{\beta(a)}^2 \) are non-negative, \( \bar{\sigma}_a^2 \) and \( \bar{\sigma}_{\beta(a)}^2 \) will be ML estimators. The ML estimators of \( \sigma_a^2, \sigma_{\beta(a)}^2 \) and \( \sigma_e^2 \) under various conditions on \( \hat{\sigma}_a^2 \) and \( \hat{\sigma}_{\beta(a)}^2 \) are given in Table 1 (See [2,10,11] for details of derivations).

3.3. Restricted Maximum Likelihood (REML)

REML estimators of \( \sigma_a^2, \sigma_{\beta(a)}^2 \) and \( \sigma_e^2 \) can be obtained by maximizing that part of the likelihood function which is invariant to the location parameters of the model [2,11]. That is, in contrast to ML method, REML method is not interested in the fixed effects of the model. However, it uses fixed effects’ degrees of freedoms. This is one of the features of REML method. In order to derive REML estimations, the likelihood function in (3) is separated into fixed and random effects of the model which is given as follows:

\[
L(\mu, \sigma_a^2, \sigma_{\beta(a)}^2, \sigma_e^2 | y_{(k)}) = L(\mu | \overline{\bar{y}}) \times L(\sigma_a^2, \sigma_{\beta(a)}^2, \sigma_e^2 | MS_A, MS_{B(A)}, MS_E)
\]

As REML method takes only random part of the model, likelihood function, corresponding to REML, is given as follows:

\[
L_{REML} = L(\sigma_a^2, \sigma_{\beta(a)}^2, \sigma_e^2 | MS_A, MS_{B(A)}, MS_E)
\]

\[
= (2\pi)^{-\frac{1}{2} \left(abn - 1\right)} \left(bn \sigma_a^2 + n \sigma_{\beta(a)}^2 + \sigma_e^2\right)^{-\frac{1}{2} \left(abn - 1\right)} \left(n \sigma_{\beta(a)}^2 + \sigma_e^2\right)^{-\frac{1}{2} \left(bn - 1\right)} \left(\sigma_e^2\right)^{-\frac{1}{2} \left(abn - 1\right)} \exp \left(-\frac{1}{2} \left[ \frac{SS_A}{bn \sigma_a^2 + n \sigma_{\beta(a)}^2 + \sigma_e^2} + \frac{SS_{B(A)}}{n \sigma_{\beta(a)}^2 + \sigma_e^2} + \frac{SS_E}{\sigma_e^2} \right] \right)
\]

\[
\times \left(abn\right)^{\frac{1}{2}} \exp \left(-\frac{1}{2} \left[ SS_A + SS_{B(A)} + SS_E \right] \right)
\]
Then we can obtain REML estimators by taking the log of the likelihood function in (5) and equating to zero and taking the partial derivatives of $\ln(L_{\text{REML}})$ with respect to $\left[bn\sigma_{a}^{2}+n\sigma_{(a)}^{2}+\sigma_{e}^{2}\right]$, $\left[n\sigma_{(a)}^{2}+\sigma_{e}^{2}\right]$ and $\sigma_{e}^{2}$, respectively. After editing of these equations, finally we obtain the REML estimators as follows:

$$\hat{\sigma}_{a,\text{REML}}^{2} = \frac{MS_{A} - MS_{(A)}}{bn}, \quad \hat{\sigma}_{(a)}^{2} = \frac{MS_{(A)} - MS_{A}}{bn} \quad \text{and} \quad \hat{\sigma}_{e,\text{REML}}^{2} = MS_{E}$$

However, as in the ML estimators, these solutions are not the REML estimators due to requirements of non-negativity. The REML estimators of $\hat{\sigma}_{a}^{2}$, $\hat{\sigma}_{(a)}^{2}$ and $\hat{\sigma}_{e}^{2}$ under various conditions on $\sigma_{a}^{2}$ and $\sigma_{(a)}^{2}$ are given in Table 2 (See [2,10,11] for details of derivations).

### Table 2. REML estimators of the variance components

<table>
<thead>
<tr>
<th>Conditions</th>
<th>$\hat{\sigma}_{e}^{2}$</th>
<th>$\hat{\sigma}_{(a)}^{2}$</th>
<th>$\hat{\sigma}_{a}^{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{a}^{2} \geq 0$, $\sigma_{(a)}^{2} \geq 0$</td>
<td>$MS_{E}$</td>
<td>$\frac{MS_{(A)} - MS_{A}}{n}$</td>
<td>$\frac{MS_{A} - MS_{(A)}}{bn}$</td>
</tr>
<tr>
<td>$\sigma_{a}^{2} \geq 0$, $\sigma_{(a)}^{2} &lt; 0$</td>
<td>$SS_{E} + SS_{(A)}$</td>
<td>0</td>
<td>$\frac{MS_{A} - \hat{\sigma}_{e,\text{REML}}^{2}}{bn}$</td>
</tr>
<tr>
<td>$\sigma_{a}^{2} &lt; 0$, $\sigma_{(a)}^{2} \geq 0$</td>
<td>$MS_{E}$</td>
<td>$\frac{1}{n}\left(\frac{SS_{A} + SS_{(A)}}{ab-1} - MS_{E}\right)$</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma_{a}^{2} &lt; 0$, $\sigma_{(a)}^{2} &lt; 0$</td>
<td>$\frac{SS_{(A)} + SS_{A} + SS_{E}}{abn-1}$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### 3.4. Bayesian Method

In order to carry out Bayesian analysis, in addition to the likelihood function we need prior distributions for each parameter of the model. For the fixed effect, a flat prior was used, so that $f(\mu) \propto \text{constant}$, indicating no prior knowledge about this parameter. The normal distribution (N) was used as the prior distribution of the random effects, $f(\alpha_{i} | \sigma_{a}^{2}) \sim N(0, \sigma_{a}^{2})$ and $f(\beta_{(a)i} | \sigma_{(a)}^{2}) \sim N(0, \sigma_{(a)}^{2})$. The inverse gamma (IG) distribution is assigned for the variance components, $f(\sigma_{a}^{2}) \sim IG(v_{a}, s_{a}^{2})$ The forms of the prior distributions are

$$f(\mu) \propto \text{constant}$$

$$f(\alpha_{i} | \sigma_{a}^{2}) \propto \left(\sigma_{a}^{2}\right)^{-\frac{1}{2}} \exp\left(-\frac{\sum_{i=1}^{a} \alpha_{i}^{2}}{2\sigma_{a}^{2}}\right)$$

$$f(\beta_{(a)i} | \sigma_{(a)}^{2}) \propto \left(\sigma_{(a)}^{2}\right)^{-\frac{1}{2}} \exp\left(-\frac{\sum_{i=1}^{b} \beta_{(a)i}^{2}}{2\sigma_{(a)}^{2}}\right)$$

(6)
\[
f(\sigma^2_i | \nu_i, s_i^2) \propto (\sigma^2_i)^{-\frac{1}{2}(\nu_i+2)} \exp \left\{ -\frac{\nu_i s_i^2}{2\sigma^2_i} \right\} \quad i = \alpha, \beta, e
\]

By multiplying likelihood function in (3) with the prior distributions of all the parameters in (6), the joint posterior density of parameters is obtained as:

\[
f(\theta | y) \propto L(\theta | y)f(\mu) f(\alpha_i) \frac{1}{\sigma^2_{\alpha}} \, f(\beta_{j(i)}) \frac{1}{\sigma^2_{\beta(a)}} \, f(\sigma^2_{\alpha}) \frac{1}{\sigma^2_{\beta(a)}} \, f(\sigma^2_{\beta(a)}) \frac{1}{\sigma^2_{\beta(a)}} \, f(\sigma^2_e) \frac{1}{\sigma^2_{\beta(a)}} \times \exp \left\{-\frac{1}{2} \left[ \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{d} (y_{ik} - \mu - \alpha_i - \beta_{j(i)})^2 + \sum_{i=1}^{a} \alpha_i^2 + \sum_{j=1}^{b} \beta_{j(i)}^2 + \sum_{i=1}^{a} \sum_{j=1}^{b} \beta_{j(i)}^2 \right] \right\}
\]

To implement the Gibbs sampling algorithm, we require the full conditional posterior distributions of \( \mu, \alpha, \beta, \sigma^2_{\alpha}, \sigma^2_{\beta(a)} \) and \( \sigma^2_e \). The full conditional posterior distribution of any parameter of interest can be obtained by integrating over the remaining parameters from joint posterior distribution. These are summarized as follows:

\[
[\mu | \alpha_i, \beta_{j(i)}, \sigma^2_{\alpha}, \sigma^2_{\beta(a)}, \sigma^2_e, y] \sim N\left( \bar{y}_i - \alpha_i - \bar{\beta}_{j(i)}, \sigma^2_e \right)
\]

\[
[\alpha_i | \mu, \beta_{j(i)}, \sigma^2_{\alpha}, \sigma^2_{\beta(a)}, \sigma^2_e, y] \sim N\left( \frac{\sigma^2_{\alpha}}{bn\sigma^2_{\alpha} + \sigma^2_e} n(b\bar{y}_i - \mu) - b\bar{\beta}_{j(i)} \frac{\sigma^2_e}{bn\sigma^2_{\alpha} + \sigma^2_e} \right)
\]

\[
[\beta_{j(i)} | \mu, \alpha_i, \sigma^2_{\alpha}, \sigma^2_{\beta(a)}, \sigma^2_e, y] \sim N\left( \frac{n^2 \sigma^2_{\beta(a)} \sum_{j=1}^{d} (y_{ij} - \mu - \alpha_i)}{n^2 \sigma^2_{\beta(a)} + \sigma^2_e} \frac{\sigma^2_e}{n^2 \sigma^2_{\beta(a)} + \sigma^2_e} \right)
\]

\[
[\sigma^2_{\alpha} | \mu, \alpha_i, \beta_{j(i)}, \sigma^2_{\beta(a)}, \sigma^2_e, y] \sim IG\left( a + \sum_{i=1}^{a} \alpha_i^2 + \nu_a \frac{s_a^2}{2} \right)
\]

\[
[\sigma^2_{\beta(a)} | \mu, \alpha_i, \beta_{j(i)}, \sigma^2_{\alpha}, \sigma^2_e, y] \sim IG\left( ab + \sum_{i=1}^{a} \sum_{j=1}^{b} \beta_{j(i)}^2 + \nu_{\beta(a)} \frac{s_{\beta(a)}^2}{2} \right)
\]
We consider 4 different data sets to determine variance components from such an experiment. For this reason, 4 sires were randomly chosen with 5 dams per sire and 2 offspring per dam. Then, the milk yield of offspring was recorded (305 days). This data set is shown in the first column of Table 3. Other data sets (data set 2, data set 3 and data set 4) were obtained by modifying the data set 1.

Table 3. Records for milk yields (kg)

<table>
<thead>
<tr>
<th>Sire</th>
<th>Dam</th>
<th>Offspring's yield</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Data set 1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4379</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5560</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4637</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5726</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4968</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5355</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4605</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4393</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5195</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6137</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6255</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5555</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6268</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>7112</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5840</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>6246</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5400</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>7301</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5455</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>7374</td>
</tr>
</tbody>
</table>
In the analysis of data sets 1 to 4, a normal distribution with means obtained from REML method (5932, 5707, 6582 and 6584, respectively) were used with a variance of 1000. An inverse gamma distribution with shape and scale parameters set to 0.001, IG(0.001,0.001), was chosen as the prior distribution for each of the variance components in models \( \left( \sigma^2_\alpha, \sigma^2_{\beta(a)}, \sigma^2_e \right) \) to reflect prior ignorance about the parameters.

Statistical analyses for ANOVA, ML and REML estimations were obtained using PROC MIXED procedure, and Bayesian analysis was conducted using the PROC MCMC procedure of SAS software [12]. A single chain of size 500000 iterations was run. The initial 10000 iterations were discarded as a burn-in, and every 50th sample was recorded to reduce the auto-correlation. In total, 9800 samples were stored for each parameter, and means of the sample values were used as an estimate of the parameters.

5. RESULTS

The results from four different methods are shown in Table 4. As can be seen from this table, the ANOVA estimates of the parameters outside the parameter space are treated as they are and data sets yield ANOVA estimates of \( \sigma^2_\alpha \) and \( \sigma^2_{\beta(a)} \) ranging from -16480 to 1101634. Although the data set 1 has positive ANOVA estimates for \( \sigma^2_\alpha \) and \( \sigma^2_{\beta(a)} \), the data sets 2, 3 and 4 are the most difficult ones and badly behaved, in that the estimates of \( \sigma^2_\alpha \) and \( \sigma^2_{\beta(a)} \) are negative, rendering inferences about these parameters very difficult. The ANOVA estimators of \( \sigma^2_e \) and \( \sigma^2_{\beta(a)} \) are unbiased yet they can be negative. The ML and REML estimators are obtained by deleting those negative values and substituting zero. Therefore, these estimators are biased upwards.

The posterior means for \( \sigma^2_e \), \( \sigma^2_\alpha \) and \( \sigma^2_{\beta(a)} \) from the analysis of data sets 1-4 in Table 4 are based on 9800 Gibbs sampler. These point estimates of variance components from Bayesian analysis are within the permissible parameter space in contrast to the estimates obtained from ANOVA. The Bayesian method is feasible computationally and appears to give much more sensible answer to the inferential problems than ANOVA and likelihood-based estimation methods.

Table 4. ANOVA, ML, REML and Bayesian estimations of variance components

<table>
<thead>
<tr>
<th>Methods</th>
<th>Data Set 1</th>
<th>Data Set 2</th>
<th>Data Set 3</th>
<th>Data Set 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \sigma^2_\alpha )</td>
<td>( \sigma^2_{\beta(a)} )</td>
<td>( \sigma^2_e )</td>
<td>( \sigma^2_\alpha )</td>
</tr>
<tr>
<td>ANOVA</td>
<td>259296</td>
<td>97566</td>
<td>689557</td>
<td>-10245</td>
</tr>
<tr>
<td>ML</td>
<td>172355</td>
<td>97567</td>
<td>689557</td>
<td>823985</td>
</tr>
<tr>
<td>REML</td>
<td>259297</td>
<td>97567</td>
<td>689557</td>
<td>1100051</td>
</tr>
</tbody>
</table>
It can be noted that the Bayesian method overestimates the variance components compared with the estimates of ANOVA, ML and REML. Variance components obtained by REML are only marginal with respect to fixed effects but conditionals to other nuisance parameters of the model. The Bayesian analysis allows further marginalization via Markov chain Monte Carlo methods. This approach is particularly interesting for models, as the present, with high number of variance components. In consequence, point estimates of variance components obtained in the Bayesian analysis under that priors presented some differences with the ML and REML estimates.

6. DISCUSSIONS

In this paper we have presented four different methods of estimating the variance components. Bayesian estimators depend not only on the information about the parameters contained in the data, but also on prior knowledge. This is one of the potential advantages of the Bayesian methods. Therefore, it is expected that the Bayesian method will do better than the classical procedures when the data contain little information about the parameters of interest. Moreover, the Bayesian method implicitly account for the uncertainty about the values at the parameters of interest.

The computations required to implement the Bayesian method are of the same order of magnitude as those required for the classical methods, and therefore the Bayesian method are likely to be computationally feasible whenever the classical procedures are computationally feasible.

One of the main differences between the Bayesian and maximum likelihood approaches is the way in which they deal with nuisance parameters. This is apparent from our results. The likelihood function is obtained by maximizing with respect to the nuisance parameters, whereas the conditional posterior density is obtained by a Monte Carlo numerical integration method, which is known as a Gibbs Sampler. In certain cases, the two operations may produce sharply contrasting results.

While variance components estimations of all parameters for each method are non-negative in the first data set, there are negative ANOVA estimations for the other data sets. So that, when ANOVA estimations are negative, ML and REML estimations are zero. REML estimators are quite different from ML estimators because REML method eliminates deviations which are originated from ML method. Moreover, REML estimations of first data set are almost similar with ANOVA estimations [13].

Based on the results from our four data sets, we can conclude that the estimates of ANOVA, ML and REML are accurate but the posterior point estimates from the Gibbs sampling can be overestimated depending on the nature of the data set. The differences in the results of different estimation methods occurred the most in the estimation of sire and dam variances. Data sets 2, 3 and 4 give highly negative estimates of sire and dam variance components from ANOVA, ML and REML methods truncate these estimates to zero, and the posterior point estimates from the Gibbs sampler tend to be biased. This is due to the fact that each data set provides relatively small amount of information for sire and dam variances than for the error variance.

Finally, we can conclude that the Bayesian method of estimation using the Gibbs sampling approach is suitable for estimating the variance components under balanced two-way nested design as compared to traditional methods, particularly for small sample data sets.

REFERENCES


