Investigation of Biases for Variance Components on Multiple Traits with Varying Number of Categories in Threshold Models Using Bayesian Inferences

D. H. Lee*

Department of Animal Life and Resources**,** Hankyong National University, Korea

ABSTRACT : Gibbs sampling algorithms were implemented to the multi-trait threshold animal models with any combinations of multiple binary, ordered categorical, and linear traits and investigate the amount of bias on these models with two kinds of parameterization and algorithms for generating underlying liabilities. Statistical models which included additive genetic and residual effects as random and contemporary group effects as fixed were considered on the models using simulated data. The fully conditional posterior means of heritabilities and genetic (residual) correlations were calculated from 1,000 samples retained every 10th samples after 15,000 samples discarded as "burn-in" period. Under the models considered, several combinations of three traits with binary, multiple ordered categories, and continuous were analyzed. Five replicates were carried out. Estimates for heritabilities and genetic (residual) correlations as the posterior means were unbiased when underlying liabilities for a categorical trait were generated given by underlying liabilities of the other traits and threshold estimates were rescaled. Otherwise, when parameterizing threshold of zero and residual variance of one for binary traits, heritability estimates were inflated 7-10% upward. Genetic correlation estimates were biased upward if positively correlated and downward if negatively correlated when underling liabilities were generated without accounting for correlated traits on prior information. Residual correlation estimates were, consequently, much biased downward if positively correlated and upward if negatively correlated in that case. The more categorical trait had categories, the better mixing rate was shown. *(Asian-Aust. J. Anim. Sci. 2002. Vol 15, No. 7 : 925-931)*

Key Words : Categorical Traits, Bayesian, Variance Components, Biases

INTRODUCTION

Several traits of importance in animal breeding are recorded with discrete categories. The records of some of these traits interested are measured by ordered categories like litter size, calving ease, conformation and type score or binary like return rate, stillbirth, etc. We postulate that these traits were also correlated with each other and with the other continuous traits. Analysis of such variables can be performed based on the threshold model conception (Wright, 1934). Some studies (Gianola, 1982; Harville and Mee, 1984; Gianola and Foulley, 1983 and Foulley et al., 1983) showed a theoretical prospectives of threshold methodology. Gianola and Fernando (1986) proposed the Bayesian approach as a conceptual strategy to solve problems arising in animal breeding with respect to nonlinear merit functions and models. The usefulness of Bayesian statistics via Gibbs sampling (GS) (Gelman and Gelman, 1984) is not required the numerical integration of complicated density functions by taking the repeated samples from the posterior distributions. Sorensen et al. (1995) described the methodology of Bayesian inference in univariate threshold model to implement a GS algorithm. Wang et al. (1997) extended the work of Sorensen et al. (1995) to one multiple ordered categorical trait and one continuous trait with all possible missing patterns of data from a Bayesian perspective. They estimated one threshold (t_3) because the

simulated data used were generated with four categories in a categorical trait and fixed to $t_1 = 0$, $t_2 = $ *fixed* or possibly, residual variance=1. VanTassell et al. (1998) again extended Sorensen et al. (1995) and Wang et al. (1997) showed algorithms to generalized multi-trait animal model for categorical variables that have binomial or multinomial outcomes and for any combination of categorical and continuous traits. They used standard or alternate parameterization referred by Sorensen et al. (1995) of the thresholds and residual variance on their programs (MTGSAM) to guarantee identifiable. That is, "Standard" parameterization ($t_{1,i} = 0$ and residual variance of 1) was assumed for binary data and the alternative ($t_{1,i} = 0$ and $t_{2,i} = 1$) was assumed for at least three categories with unknown residual variance (Sorensen et al., 1995). Lee et al. (2002) proposed the other parameterization of thresholds. They suggested that rescaling method on thresholds was needed to sustain the properties of distribution of underlying variables and guarantee identifiable. However, there is still some obstacle to overcome. Main problem is biased estimates for (co) variance components, especially, in binary trait model under the Bayesian inference. Moreno et al. (1997) examined on biased inference about variance components in the binary threshold model. They suggested an iterative bootstrap method for correcting the bias problems. Luo et al. (2001) tried to implement GS for a multiple-trait Bayesian model with linear, binary, and categorical traits and described Monte-Carlo errors with respect to the frequentist properties. They concluded that

the animal model was improper for analysis of categorical traits using a threshold model and the Gibbs sampler when using one categorical and one continuous trait.

The objective of this study is to investigate the bias for heritabilities and genetic correlations in a Bayesian prospective with any combinations of continuous, ordered categorical, and binary multiple traits under an animal model.

METERIALS AND METHODS

Models and prior distributions

Assuming multiple traits with binary, multi-categories and/or continuous observations, m dichotomous or polychotomous random variables without missing traits are observed for each record j(j=1,2,…,n). The vector of observed variables can be the expression of the corresponding vector which consisted of underlying continuous variables (U_{ij}) with unknown threshold vector (t_i) corresponding to the traits. Let the observed random vectors denote y_i^0, i=1,…,k with ($y_1…y_m$) for categorical traits and ($U_{m+1},…,U_k$) for continuous traits, interchangeably for notational convenience. Now we consider the following model

$$U_i = X\beta_i + Za_i + e_i, \quad \text{for } (i=1,…,k) \qquad (1)$$

where U_i is a vector of the ith observed continuous (or underlying) variables, β_i is a vector of fixed effects, a_i is a vector of random effects, e_i is a vector of residuals, and X and Z are matrices associated with effects. Let t_{ij} be the jth threshold on i^{th} categorical trait. Then, a mixed model for the observed continuous and underlying liability variables can be $U = W\theta$, where $\theta = (\beta', a')'$ and $W = [X \quad Z]$. Furthermore, the conditional probability of U given θ and R regardless of proper or improper were

$$
\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_t \end{bmatrix} \sim MVN \left(\begin{bmatrix} W\theta_1 \\ W\theta_2 \\ \vdots \\ W\theta_3 \end{bmatrix}, R\otimes I \right) \qquad R = \begin{bmatrix} r_{11} & \cdots & r_{1t} \\ \vdots & \ddots & \vdots \\ r_{t1} & \cdots & r_{tt} \end{bmatrix}
$$

,Where $\qquad (2)$

The traditional model assumptions can be achieved in Bayesian model by putting the prior distributions appropriately. We refer to the works by VanTasell et al. (1998) for assuming the prior distributions of location parameters, variance components and thresholds to non-informatives.

Conditional posterior distributions and Gibbs sampling

To implement Gibbs sampling with the conditional posterior distributions, we will use a minus subscript to delete an appropriate element from a matrix or a vector. For example, U_{-i} will be denoted the vector U with U_i deleted. Similarly and U_{-ij} be denoted U with U_{ij} deleted.

Under Model (1), according to the work by VanTassell et al. (1998), the underlying liability variable of jth observation for ith categorical trait given θ, Y_{ij}, t_i, U_{-ij} and R can be drawn from a truncated normal distribution (TN) as follow:

$$u_{ij} \mid U_{-ij}, Y_{ij} = k, \theta, R, t_i \sim TN_{t_{ik-1}, t_{ik}} \left(\xi_{ij}, \sigma_{i\cdot}^2 \right) \qquad (3)$$

where $\xi_{ij} = w'_j\theta_i + R'_{i(-i)}R^{-1}_{(-i)(-i)} \begin{pmatrix} U_{1j} - w'_j\theta_1 \\ U_{2j} - w'_j\theta_2 \\ M \\ U_{kj} - w'_j\theta_k \end{pmatrix}_{-i}$;

$\sigma_{i\cdot}^2 = r_{ii} - R_{i(-i)}R^{-1}_{(-i)(-i)}R_{(-i)i}$, with w'_j being the jth row of W on (2).

Conditionally, thresholds for ith categorical trait can be drawn from a uniform distribution given by Sorensen et al. (1995).

Location parameters ($\theta = (\beta', a')'$) given θ_{-i}, G, R, and U were sampled from a multivariate normal distribution for the traits within effects with means and variances as:

$$\beta_i \mid \beta_{-i}, a, R, U \sim MVN_k \left(\hat{\beta}_i, (x'_i R^{-1} x_i)^{-1} \right) \qquad (4)$$

$$a_i \mid a_{-i}, \beta, G, R, U_i \sim MVN_k(\hat{a}_i, (z'_i R^{-1} z_i + G^{-1} A_{ii}^{-1})^{-1}) \qquad (5)$$

where $\hat{\beta}_i(\hat{a}_i)$ is a $k \times 1$ vector for the estimates of ith level of fixed (random) effect. The more details for estimating location parameters can be referred to the note by Sorensen (1996). Dispersion parameters were sampled from inverted Wishart distributions (IW) assuming the flat prior distributions as follow:

$$G \mid \theta, R, t, U, y \sim IW_k(SS_a) \qquad (6)$$

$$R \mid \theta, G, t, U, y \sim IW_k(SS_e) \qquad (7)$$

where $SS_a(SS_e)$ is the sum of squares for genetic (residual) effects. More details were described on Sorensen's (1996) note.

Liability

According to (3), underlying liabilities can be generated with two possibilities as follow:

Step by step method (ESTEP) : The distributions on (3) can be stated in terms of the residual $e_{ij} = U_{ij} - w'_{ij}\theta_i$ because of a one-to-one relationship with the liability given the fixed and random effects. Then residual effects for i^{th} trait on j^{th} observation were generated with taking into account of thresholds for each trait by (3).

$$e_{1j} \sim TN_{t_{1\lambda-1}, t_{1\lambda}}(0, r_{11})$$
$$e_{2j} \mid e_{1j} \sim TN_{t_{2\lambda-1}, t_{2\lambda}}(r_{12}r_{11}^{-1}e_1, r_{22} - r_{21}r_{11}^{-1}r_{12}) \tag{8}$$
$$\vdots$$
$$e_{ij} \mid e_{1j}, \ldots, e_{i-1,j} \sim TN_{t_{i\lambda-1}, t_{i\lambda}}\left(r_{i,1:i-1}r_{1:i-1,1:i-1}^{-1}e_{1:i-1}, r_{ii} - r_{i,1:i-1}r_{1:i-1,1:i-1}^{-1}r_{1:i-1,i}\right)$$

Full method (EFULL) : Residual terms of underlying liabilities were generated as follow:

$$e_i \mid e_{-i} \sim TN_{t_{i,\lambda-1}, t_{i,\lambda}}(r_{i,-i}r_{-i,-i}^{-1}e_{-i}, r_{ii} - r_{i,-i}r_{-i,-i}^{-1}r_{-i,i}) \tag{9}$$
M

where $e_{-i} = U_{-i} - W_{-i}\theta_{-i}$ and U_{-i} is the information gotten from previous round or previous step within same round. That is,

$$E(e_1^n) = r_{1,-1}r_{-1,-1}^{-1}\begin{pmatrix} e_2^{n-1} \\ e_3^{n-1} \\ \vdots \\ e_t^{n-1} \end{pmatrix}, E(e_2^n) = r_{2,-2}r_{-2,-2}^{-1}\begin{pmatrix} e_1^n \\ e_3^{n-1} \\ \vdots \\ e_t^{n-1} \end{pmatrix}, \ldots, E(e_i^n) = r_{i,-i}r_{-i,-i}^{-1}\begin{pmatrix} e_1^n \\ \vdots \\ e_{i-1}^n \\ e_{i+1}^{n-1} \\ \vdots \\ e_t^{n-1} \end{pmatrix}$$

Now, underlying liability for i^{th} categorical trait can be $w'_{ij}\theta_i + e_i$ on equation of (8) or (9).

Parameterizations

The Gibbs sampler for thresholds and residual variances were constrained or reparameterized as follow:

Standard parameterization (ST) : Standard parameterization was following the works by Wang et al. (1997) and Korsgaard et al. (1999). That is, on binary traits, thresholds were fixed to zero and residual variances of 1's. On three-category trait, the first and second thresholds were fixed to zero and one, respectively. Likewise, on five-categories trait, two thresholds were fixed to zero and one, and the other two thresholds were sampled from uniform distributions. The Gibbs samplers of residual covariances between binary and multiple category(ies) or linear trait were rescaled a method extended the work by Korsgaard et al. (1999) with taking into account for correlated traits.

Constrains and rescaling thresholds (RT) : The thresholds sampled regardless of binary or multiple categorical traits was(were) rescaled without restriction for residual variances as follow:

$$t_{ij}^* = (t_{ij} - t_{i1}) / \sigma_{ie}, \text{ for } j = 1, \ldots, C_i\text{-}1, \text{ where } \sigma_{ie} \text{ is}$$

standard deviation for residual effect of ith categorical

trait and C_i is the number of categories for i^{th} categorical trait.

Simulation study

Data were simulated using the following 3-trait linear model:

$$y_{ijk} = CG_{ij} + a_{ik} + e_{ijk} \text{ for i} = 1, \ldots, 3$$

where y_{ijk} is the observation in j^{th} contemporary group in animal k^{th} on i^{th} trait; CG_{ij} is the jth contemporary group effect on i^{th} trait; a_{ik} is the kth animal genetic effect; and e_{ijk} is the random residual. The random effects were distributed as:

$$a \sim N(0, G_0 \otimes A), \ e \sim N(0, R_0 \otimes I_n)$$

where A is the numerator relationship matrix and $\otimes$ is kronecker product. We assumed as:

$$G_0 = \begin{bmatrix} 30 & -5 & 10 \\ & 15 & 30 \\ Symm & & 100 \end{bmatrix}, R_0 = \begin{bmatrix} 50 & -10 & 20 \\ & 85 & 40 \\ Symm & & 400 \end{bmatrix}$$

The base population consisted of 10 sires and 1,000 dams. Mating within each generation was random, and 10% males and 90% females were randomly selected for each next generation. A total of 5 generations of animals with 10 contemporary groups (CGs) per generation were generated. Each animal randomly pertained to a CG of 10 CGs on each generation. Animals were randomly mated on every generation and selected 10% males and 90% females. The value of every record was discretized to 2 categories using 1 on standard scale. Likewise, for 3-, 4- and 5-categorical traits, (1,2), (0,1,2) and (0,1,2,3) on standard scale were used, respectively. Variance-covariance components and their parameters by GS with *ESTEP* were compared with *EFULL* in two threshold models. One is a three categorical trait model with one binary, one 3-categorical and one 5-categorical trait (235) (Model 1) and the other is multiple categorical trait model with two binary and one linear trait (22L) (Model 2). Furthermore, the other three different models with *EFULL* were setup and implemented ST and RT algorithms as: two binary and one 4-category model (224) (Model 3), one binary, one three-category and one linear model (23L) (Model 4), one three-category, one five-category, and one linear model (35L) (Model 5). Five replicates for Gibbs sampling on each model were carried out. Finally, 11,127 animals from 220 sires with 10,272 records were generated (table 1).

Table 1. Average numbers of animals, sires, contemporary groups (CG), and number of progeny per sire and CG in simulated data (5 replicates)

	Average numbers		Average numbers
Animals	11,127	Records	10,272
Sire	220	Progeny per sire	47
CG	50	Progeny per CG	205

RESULTS

Analysis of gibbs samples

1,000 cycles were apparently enough to assume "burn-in" in each model and replicate. Nevertheless, 1,000 Gibbs samples which were retained every 10th cycles after 15,000 cycles were discarded as "burn-in" period were used to calculate statistics of the posterior distributions of dispersion parameters. Mixing rate as percentage of Monte-Carlo standard deviations in the posterior means and the number of effective samples, which were calculated using a time series method provided by Geyer (1992), were presented. As considering the mixing rate by traits, the worst mixing rate in linear model with GS was shown at genetic correlations between highly correlated traits. For example in present study, correlation between binary trait and three-categorical trait ($r_{g(23)}$) was shown the worst mixing rate in all model considered. This characteristic was also shown in the general pattern although the amounts of mixing rates were depended on models.

Comparison of parameters

Parameter estimates by Gibbs samplers were compared

Table 2. Statistics for heritability and genetic (residual) correlation estimates by REML and Gibbs sampling in three traits linear model (5 replicates)

	REML	Gibbs sampler
h_1^2	0.396 (0.030) [1]	0.397 (0.029) [2]
h_2^2	0.143 (0.019)	0.145 (0.021)
h_3^2	0.203 (0.014)	0.206 (0.012)
$r_{g(12)}$	-0.233 (0.058)	-0.233 (0.062)
$r_{g(13)}$	0.200 (0.050)	0.202 (0.048)
$r_{g(23)}$	0.754 (0.030)	0.750 (0.037)
$r_{e(12)}$	-0.150 (0.013)	-0.151 (0.012)
$r_{e(13)}$	0.133 (0.012)	0.132 (0.011)
$r_{e(23)}$	0.220 (0.008)	0.219 (0.007)

[1] Average of estimates by REML.
[2] Average of the posterior means by Gibbs samples, values in parenthesis are empirical standard deviations.

to estimates by REML in linear models (table 2). Means of the posterior heritabilities and genetic correlations by GS in the model with all continuous traits were a good agreement with estimates by REML and indicated good mixing rates for all parameters. This implied these distributions to be proper.

In Model 1 and Model 2, there were no differences for heritability estimates even though the marginal posterior distributions for covariance components should be different due to different algorithms. For example, on *ESTEP*, residual correlations between categorical traits were much biased downward and genetic correlations would be slightly inflated consequently (table 3 and 4). These phenomenon might look like uncorrelated for residual effects among categorical traits. Otherwise, residual correlation between categorical trait and continuous trait were unbiased and close to "true" value (table 4). These results indicated that underlying values had to be generated with taken into account for the fully conditional posterior distribution in a multiple threshold model.

For two binary and one 4-categorical traits model (Model 3) with RT parameterization, the distributions of thresholds would be inflated slightly and showed in a good mixing rate if comparing ST method in same model. However, the distributions of thresholds should be highly peaked. We could not find out any difference for mixing rates of posterior heritabilities and genetic (residual) correlations between two different parameterizations (*ST* vs *RT*). Otherwise, the posterior heritabilities for binary traits should be different between ST and RT parameterterion. Heritability estimates with ST parameterization, restriction on residual variance of 1, showed 3-7% biased upward on binary traits (table 4). The posterior genetic correlations between two binary traits or binary and categorical trait was slightly inflated on standard parameterization (ST). The reason might be caused from inflation for the distributions of genetic covariances due to deflation for the distributions of residual covariances between both binary traits or binary and categorical trait. Otherwise, with RT parameterization, parameter estimates (heritabilities, genetic and residual correlations) from the posterior means were close to "true" values. In this case, variance-covariance matrix sampled from inverted Wishart distribution without any restriction was the approximated polygenic matrix and might be well defined as inverted Wishart distribution. In the Bayesian properties, RT parameterization ($t_{ij}^* = (t_{ij} - t_{i1})/\sigma_{ie}$) is identifiable with respect to scaling and location and this parameterization was in the agreement with the work by Wang et al. (1997).

For the one binary, one 3-categories, and one continuous trait threshold model (23L), the posterior mean for threshold on 3-categorical trait with RT was 0.978. This estimate was close to 1 as assuming at ST parameterization. This might imply that the scale of the posterior distribution for variance component using this method would be same to that with ST

Table 3. Average of the posterior means (PM) and differences for heritabilities, genetic, and residual correlations between by REML in linear model and by Gibbs sampling in threshold animal model on the different combinations of traits using simulated data (5 replicates)

	235-ESTEP-ST[1]		23L-ESTEP-RT[2]	
	PM	Diff	PM	Diff
h_1^2	0.385 (0.052)	-0.011 (0.037) [3]	0.394 (0.051)	-0.002 (0.035) [3]
h_2^2	0.148 (0.012)	0.005 (0.018)	0.163 (0.021)	0.020 (0.022)
h_3^2	0.219 (0.009)	0.016 (0.008)	0.209 (0.013)	0.006 (0.003)
$r_{g(12)}$	-0.231 (0.096)	0.002 (0.039)	-0.300 (0.122)	-0.067 (0.066)
$r_{g(13)}$	0.233 (0.100)	0.033 (0.054)	0.218 (0.086)	0.018 (0.041)
$r_{g(23)}$	0.773 (0.070)	0.019 (0.049)	0.714 (0.063)	-0.040 (0.047)
$r_{e(12)}$	-0.033 (0.012)	0.118 (0.013)	0.131 (0.030)	0.281 (0.028)
$r_{e(13)}$	0.033 (0.008)	-0.100 (0.005)	0.121 (0.019)	-0.012 (0.010)
$r_{e(23)}$	0.087 (0.009)	-0.133 (0.005)	0.215 (0.017)	-0.005 (0.012)

[1] ESTEP for Gibbs samples with standard parameterization in three-trait model on first trait with binary, second trait with 3-categories and third trait with 5-categoires.

[2] ESTEP for Gibbs samples with standard parameterization in three-trait model on first trait with binary, second trait with 3-categories and third trait with linear (more details were described on text).

[3] Difference between estimates by Gibbs samples in threshold models and estimates by REML in original linear scale, values in parenthesis are empirical standard deviations.

Table 4. Comparison of biases for genetic parameters between using standard parameterization (ST) and parameterization of rescaling thresholds (RT) in different combinations of traits models (average from 5 replicates)

	224[1]		23L[2]		35L[3]	
	ST	RT	ST	RT	ST	RT
h_1^2	0.034 (0.026)	0.003 (0.039)	0.036 (0.027)	-0.002 (0.042)	0.017 (0.037)	0.007 (0.031)
h_2^2	0.065 (0.056)	0.005 (0.027)	0.002 (0.021)	-0.001 (0.025)	-0.001 (0.008)	0.005 (0.008)
h_3^2	0.009 (0.010)	0.004 (0.008)	0.007 (0.004)	0.005 (0.004)	0.005 (0.004)	0.004 (0.003)
$r_{g(12)}$	0.063 (0.073)	0.019 (0.046)	0.019 (0.069)	0.029 (0.041)	0.015 (0.064)	0.003 (0.074)
$r_{g(13)}$	0.019 (0.061)	0.004 (0.061)	0.019 (0.030)	0.001 (0.024)	0.016 (0.041)	0.019 (0.045)
$r_{g(23)}$	0.000 (0.073)	0.043 (0.060)	0.010 (0.042)	0.036 (0.058)	0.010 (0.023)	-0.016 (0.033)
$r_{e(12)}$	-0.006 (0.062)	0.014 (0.041)	0.013 (0.034)	0.012 (0.034)	0.005 (0.016)	0.010 (0.021)
$r_{e(13)}$	-0.021 (0.017)	-0.015 (0.015)	-0.015 (0.010)	-0.009 (0.009)	-0.009 (0.014)	-0.011 (0.014)
$r_{e(23)}$	-0.010 (0.006)	-0.006 (0.010)	-0.001 (0.011)	-0.003 (0.008)	-0.005 (0.006)	-0.003 (0.004)

[1] First and second traits with binary and third trait with 4-categories.

[2] First trait with binary, second trait with 3-categories and third trait with linear.

[3] First trait with 3-categories, second trait with 5-categories and third trait with linear, values in parenthesis are empirical standard deviations.

parameterization. There were no differences for mixing rates of heritabilities, genetic and residual correlations between two different parameterizations (*ST* vs *RT*). Otherwise, the posterior heritability for binary trait with ST was also shown 3.6% upward biased. With RT parameterization, the posterior genetic and residual correlations as well as heritabilities regardless of the types of traits were close to estimates by REML.

As considering the mixing rate by traits, genetic correlation between binary trait and 3-categorical trait having true value of 0.775 was shown the worst mixing rate with both parameterizations. Residual variance for continuous trait was shown the best mixing rate and residual covariance between continuous trait and 3-categorical trait was well defined with respect to mixing rate.

For the two multi-categorical traits and one continuous trait threshold model (35L) (Model 5), threshold estimates were shown the better mixing rates with RT parameterization than with ST parameterization. We could not find out any differences for the posterior heritabilities, genetic and residual correlations between two different parameterizations (ST vs RT) although the posterior distributions for residual variance-covariance matrix should be different between two parameterizations.

DISCUSSION

Heritability estimates by means of the posterior distributions with RT parameterization were close to estimates by REML within 3% in each model. So we can interpret that these estimates were unbiased. However, Moreno et al. (1997) described that the sign of the bias in binary threshold model depends on the amount of information associated with either fixed effects or with random effects and GS can produce positively biased inferences when the amount of data per fixed effect is small. They also presented that, when fixed effects are poorly estimated, the bias persists even if posterior distributions are guaranteed to be proper and expected that, in the case of the animal model, the bias problems should be accentuated. In this study, we tried to implement binary traits analysis to multiple binary traits and joint analysis with binary, ordered categorical and linear traits and figure out bias problems on binary traits. In all case, the estimates of heritability were in a good agreement with the true values assuming that REML estimates were true values. One possibility for the reason of bias problems, the posterior distributions for residual covariance matrix would not be the property of inverted Wishart distributions anymore if an element on matrix was restricted, for example, residual variance of one for binary trait as standard parameterization.

Genetic covariances were biased upward although genetic variances were unbiased with *ESTEP*. Moreover, residual covariances would be much biased. These biases would be originated from not taking into account of the fully conditional posterior distribution. The amount of bias is larger on binary traits than on multiple categorical traits and larger on categorical (binary) traits in joint analysis model than in threshold models. Hoeschele and Tier (1995) got results that significantly overestimated heritability

under the sire model using Monte-Carlo posterior mode and mean in simulated data with 50 sires and 40 progeny on average. Our results about bias were much better than their results. However, further studies with several models are needed.

CONCLUSION

Genetic parameters (heritability and genetic correlations) and location parameters (breeding values) could be estimated in multi-trait threshold model with any combination of binary, multiple categories and linear traits in Bayesian prospective. Underlying arbitrary values as liability scale with correlated traits should be generated with taken into account for correlated effects of residual. One possibility is that underlying values of correlated traits, which were generated at previous round or step, can be conditioned if these values were generated sequentially according to traits. When residual variance of one was restricted, heritability estimates always cause to bias. The alternative would be to rescale the thresholds by residual variance or to find the other way if possible.

REFERENCES

Foulley, J. L., D. Gianola and R. Thompson. 1983. Prediction of genetic merit from data on binary and quantitative variates with an application to calving difficulty, birth weight and pelvic opening. Genet. Sel. Evol. 15:401-424.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Robin. 1995. Bayesian data analysis. Chapman and Hall/CRC. pp. 52-57.

Geman, S. and D. Geman. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. 6:721-741.

Geyer, C. J. 1992. Practical Markov Chain Monte Carlo (with discussion). Stat. Sci. 7:467-511.

Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079-1096.

Gianola, D. and R. L. Fernando. 1986. Bayesian methods in animal breeding theory. J. Anim. Sci. 63:217-244.

Gianola, D. and J. L. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:201-244.

Harville, D. A. and R. W. Mee. 1984. A mixed-model procedure for analyzing ordered categorical data. Biometrics. 40:393-408.

Hoeschele, I. and B. Tier. 1995. Estimation of variance components of threshold characters by marginal posterior modes and means via Gibbs sampling. Genet. Sel. Evol. 27:519-540.

Korsgaard, I. R., A. H. Andersen and D. Sorensen. 1999. A useful reparameterisation to obtain samples from conditional inverse Wishart distributions. Genet. Sel. Evol. 31:177-181.

Luo, M. F., P. J. Boettcher, L. R. Schaeffer and J. C. M. Dekkers. 2001. Bayesian inference for categorical traits with an application to variance component estimation. J. Dairy Sci. 84:694-704.

Moreno, C., D. Sorensen, L. A. Garcia-Cortes, L. Varona and J. Altarriba. 1997. On biased inferences about variance components in the binary threshold model. Genet. Sel. Evol. 29:145-160.

Sorensen, D. 1996. Gibbs sampling in quantitative genetics. Internal report no. 82. Danish institute of animal science. Denmark.

Sorensen, D. A., S. Andersen, D. Gianola and I. Korsgaard. 1995. Bayesian inference in threshold models using Gibbs sampling. Genet. Sel. Evol. 27:229-249.

VanTassell, C. P., L. D. Van Vleck and K. E. Gregory. 1998. Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling. J. Anim. Sci. 76:2048-2061.

Wang, C. S., R. L. Quaas and E. J. Pollak. 1997. Bayesian analysis of calving ease scores and birth weights. Genet. Sel. Evol. 29:117-143.

Wright, S. 1934. An analysis of variability in number of digits in an inbred strain of guinea pigs. Genetics 19:506-536.