ANÁLISE DE MODELOS LINEARES MISTOS VIA INFERÊNCIA BAYESIANA

Marcos Deon Vilela de RESENDE[1]
Laércio Luíz DUDA[2]
Paulo Ricardo Bitencourt GUIMARÃES[3]
José Sebastião Cunha FERNANDES[4]

▪ RESUMO: O presente trabalho teve como objetivos apresentar os fundamentos teóricos e práticos da estimação Bayesiana de variáveis aleatórias (valores genéticos, efeitos fixos e componentes de variância) e aplicar esta abordagem na análise de um conjunto de dados experimentais, comparando-a com o procedimento REML/BLUP. Os resultados revelaram que: a análise Bayesiana propiciou resultados adicionais àqueles obtidos pela abordagem freqüentista, destacando-se os intervalos de confiança Bayesianos para as estimativas de componentes de variância, valores genéticos e ganhos genéticos; as estimativas dos componentes de variância, valores genéticos e ganhos genéticos pelo procedimento Bayesiano (implementado via amostragem de Gibbs) foram mais precisas do que pelo procedimento REML/BLUP; a análise Bayesiana é uma técnica elegante e flexível que permite a simultânea estimação dos componentes de variância, efeitos fixos e valores genéticos de maneira precisa, mesmo para amostras de tamanho finito.

▪ PALAVRAS-CHAVE: Componentes de variância; variáveis aleatórias; simulação estocástica; probabilidades condicionais; amostrador de Gibbs.

1 EMBRAPA, Caixa Postal 319, CEP 83411-000 – Colombo – PR.
2 Champion Papel e Celulose S.A. Mogi-Guaçu – SP.
3 Departamento de Estatística – Universidade Federal do Paraná, Curitiba – PR, Caixa Postal 19071, CEP 81531-990 – Curitiba – PR.
4 Departamento de Genética – Universidade Federal do Paraná, Curitiba – PR, Caixa Postal 19071, CEP 81531-990 – Curitiba – PR.

1 Introdução

A predição de variáveis aleatórias e a estimação de componentes de variância e efeitos fixos via metodologia de modelos mistos apresentam grande relevância em diversas áreas do conhecimento, especialmente na biometria, econometria, engenharia, física e genética quantitativa.

Embora a metodologia de modelos lineares mistos tratada pela abordagem da inferência estatística freqüentista apresente várias propriedades desejáveis (Henderson, 1984; Searle et al, 1992; Rao, 1997), publicações recentes (Gianola et al., 1990; Sorenson, 1996) tem mostrado vantagens da metodologia de modelos mistos quando abordada do ponto de vista da inferência estatística Bayesiana.

Na análise de modelos lineares mistos pela abordagem freqüentista, rotineiramente a predição/estimação baseia-se no procedimento BLUP/REML (melhor predição linear não viciada/máxima verossimilhança restrita). Entretanto, através deste procedimento, a distribuição e variância dos estimadores não são conhecidas, de forma que questões referentes a acurácia e precisão das predições e estimações não podem ser respondidas com rigor. O método REML propicia apenas intervalos de confiança aproximados para os componentes de variância, através do uso de aproximações e suposições de normalidade (argumentos assintóticos) (Gianola et al., 1990; Resende & Rosa Perez, 1999).

Por outro lado, a análise Bayesiana de modelos lineares mistos baseia-se no conhecimento da distribuição a posteriori dos parâmetros a serem estimados, fato que possibilita a construção de intervalos de confiança exatos para as estimativas das variáveis aleatórias, componentes de variância e efeitos fixos. Assim, tal abordagem propicia uma descrição mais completa sobre a confiabilidade das estimativas do que o procedimento BLUP/REML. Segundo Sun et al. (1996), uma análise Bayesiana exata pode ser obtida para os modelos de componentes de variância baseados na teoria da normalidade, permitindo, para qualquer parâmetro de interesse, uma detalhada inferência para amostras de tamanho finito.

Em inferência Bayesiana não existe qualquer distinção entre efeitos fixos ou aleatórios, sendo que os parâmetros a serem estimados são considerados variáveis aleatórias (Bibby & Toutenburg, 1977; Gianola & Fernando, 1986) que devem ser estimadas considerando as incertezas associadas a elas. Em termos de estimação, enquanto para a

inferência freqüentista vários estimadores para um parâmetro podem existir, para a inferência Bayesiana, existe, a princípio, um único estimador, o qual conduz a estimativas que maximizam a função densidade de probabilidade a posteriori. Dessa forma, os dados são fixados na distribuição a posteriori e a estimação Bayesiana permite a integrada estimação-decisão e a análise exata de amostras de tamanho finito (Gianola et al., 1990), a qual não pode ser obtida pela metodologia clássica de modelos mistos.

Devido às excelentes propriedades teóricas e práticas da análise Bayesiana, acredita-se que a mesma tornar-se-á rotineira na área de genética quantitativa, transformando-se no procedimento padrão para a estimação de componentes de variância e de valores genéticos. Com base no exposto, o presente trabalho teve como objetivos implementar a análise Bayesiana via simulação estocástica (amostrador de Gibbs, no caso) e compará-la ao procedimento REML/BLUP para a estimação/predição de componentes de variância e valores genéticos, empregando dados experimentais.

2 Metodologia

2.1 Fundamentos Bayesianos da Predição de Variáveis Aleatórias

O Teorema de Bayes definido em termos de densidades de probabilidade, tem a seguinte formulação para a distribuição de uma variável aleatória contínua:

$$f\left(\theta|y\right) = \frac{f\left(y|\theta\right) f\left(\theta\right)}{\int_R f\left(y|\theta\right) f\left(\theta\right) d\theta}. \qquad (1)$$

Onde:

θ – vetor de parâmetros

y – vetor de dados ou de informações obtidas por amostragem.

$f(\theta/y)$ – distribuição condicional de θ dado y, ou *distribuição a posteriori* (que é a base da estimação e predição Bayesiana).

$f(\theta/y)$ – função densidade de probabilidade da distribuição condicional de uma observação (y) dado θ (denominada função de verossimilhança ou modelo para os dados).

$f(\theta)$ – função densidade de probabilidade da *distribuição a priori*, que é também a densidade marginal de θ. Esta função denota o grau de conhecimento acumulado sobre θ, antes da observação de y.

$f(y/\theta)\, f(\theta)$ – função densidade conjunta de y e θ.

$$f(y) = \int_R f(y,\theta)\, d\theta = \int_R f(y|\theta)\, f(\theta)\, d\theta = E_\theta\left[f(y|\theta)\right] \quad -$$

distribuição marginal ou preditiva de y com respeito a θ, onde R é a amplitude da distribuição de θ. E_θ significa esperança com respeito à distribuição de θ. (A integração da distribuição conjunta, no espaço paramétrico, produz a marginal de y).

Como $f(y)$ não é função de θ (ou seja, $f(y)$ é constante para qualquer θ), a forma usual da formulação de Bayes é: $f(\theta/y) \propto f(y/\theta)\, f(\theta)$, onde $\propto$ indica proporcionalidade.

A expressão (1) advém das expressões $f(\theta,y) = f(y/\theta)\, f(\theta)$ e $f(\theta,y)=f(\theta/y)\, f(y)$, as quais são obtidas a partir do teorema da probabilidade condicional.

Verifica-se pela expressão (1) que um fator lógico que diferencia o enfoque Bayesiano da abordagem freqüentista refere-se ao tipo de informação utilizada. Na concepção Bayesiana, toda informação de que se dispõe é útil e deve ser utilizada. Por outro lado, a estatística clássica utiliza apenas observações de dados reais, desprezando-se as informações subjetivas (Gamerman & Migon, 1993).

A componente amostral ou experimental é comum aos modelos clássicos e Bayesianos, mas com interpretações diferentes. Embora os modelos Bayesianos passem por uma extensão dos modelos clássicos, existe uma divergência fundamental entre os dois enfoques: no modelo clássico o parâmetro é um escalar ou um vetor desconhecido, porém fixo, ao passo que no modelo Bayesiano o parâmetro é considerado como escalar ou vetor aleatório (não observável), pois para os Bayesianos tudo o que é desconhecido é incerto e, portanto, toda a incerteza deve ser quantificada em termos de probabilidade (Murteira, 1988; 1990). Em função disto, os modelos Bayesianos tratam formalmente a informação a priori, através da distribuição de probabilidade (subjetiva ou lógica) a priori. As informações a priori e amostrais permitem a atualização periódica da distribuição de probabilidade a posteriori e, portanto, permitem modificar e atualizar as estimativas dos parâmetros.

A expressão (1) fornece a regra de atualização de probabilidades sobre θ, partindo de $f(\theta)$ e chegando a $f(\theta/y)$. Assim, a distribuição a

posteriori é proporcional a verossimilhança X priori, ou seja, a função de verossimilhança conecta a priori a posteriori usando para isto os dados experimentais (amostrais). Dessa forma, a distribuição a posteriori contempla o grau de conhecimento prévio sobre o parâmetro *[f(θ)]* e também as informações adicionais propiciadas pelo experimento *[f(y/θ)]*.

Assumindo distribuição normal para a priori e para as observações, ou seja, $\theta \sim N(\mu, r^2)$ e $(Y|\theta) \sim N(\theta, \sigma^2)$ com σ^2 conhecido, a distribuição a posteriori de θ é também normal, ou seja $(\theta|Y = y) \sim N(\mu_1, r_1^2)$, onde:

$$\mu_1 = \frac{r^{-2}\mu + \sigma^{-2}y}{r^{-2} + \sigma^{-2}}; \quad r_1^{-2} = r^{-2} + \sigma^{-2}$$

Assim, verifica-se que a precisão ou inverso da variância conota informação, ou seja, a relação $w = r^{-2} / (r^{-2} + \sigma^{-2})$, $w \in (0,1)$, mede a informação contida na priori em relação à informação total (priori + verossimilhança). Dessa forma, pode-se rescrever $\mu_1 = w\mu + (1-w)\, y$, de forma que a média da posteriori eqüivale à média ponderada pela certeza na priori e na verossimilhança. Verifica-se também que a precisão (r_1^{-2}) da posteriori eqüivale à soma das precisões da priori e da verossimilhança.

Usando esta distribuição a posteriori (distribuição a priori, normal, conjugada a posteriori normal), pode-se demonstrar que os fundamentos básicos da predição de valores genéticos (que são variáveis aleatórias desconhecidas) são essencialmente de natureza Bayesiana, conforme apresentado inicialmente por Robertson (1955).

Os efeitos genéticos aditivos são definidos como desvios e uma população de efeitos genéticos apresenta zero como média e σ_A^2 como variância. A melhor predição do efeito genético de um indivíduo sem nenhuma informação, tomado aleatoriamente da população é a média populacional μ, a qual pode ser tomada como o estimador a priori, cuja variância é σ_A^2. Tomando uma informação y do indivíduo, um segundo (dado observado) estimador do efeito genético é o desvio *(y-u)* fenotípico em relação à média populacional, o qual possui variância $\sigma_E^2 = \sigma_F^2 - \sigma_A^2$, em que σ_F^2 é a variância de *y*. Estes dois estimadores independentes podem ser combinados linearmente da melhor maneira

possível tomando as recíprocas das respectivas variâncias como pesos. Sob o enfoque Bayesiano, a esperança da distribuição a posteriori corresponde à média ponderada pela precisão, das médias da priori e da verossimilhança. Assim, tem-se:

$$\hat{a} = \left(\frac{estimador`a\ \ priori}{\sigma_A^2} + \frac{estimador\ inf ormação observada}{\sigma_F^2 - \sigma_A^2} \right) / \left(\frac{1}{\sigma_A^2} + \frac{1}{\sigma_F^2 - \sigma_A^2} \right) =$$

$$= \left(\frac{0}{\sigma_A^2} + \frac{y - \mu}{\sigma_F^2 - \sigma_A^2} \right) / \left(\frac{1}{\sigma_A^2} + \frac{1}{\sigma_F^2 - \sigma_A^2} \right) = h^2\ (y - \mu),$$

como no enfoque freqüentista, onde h^2 é a herdabilidade.

2.2 Relação entre BLUP e Estimadores Bayesianos

Duas derivações da metodologia BLUP sob o enfoque Bayesiano são apresentadas a seguir, uma delas com base em Ronningen (1971) e Dempfle (1977) e outra com base em Robertson (1955).

Para ambas derivações tem-se o seguinte modelo e suposições, assumindo *G* e *R* conhecidos:

$y = X\beta + Za + e = Wt + e; \quad W = [X\ Z]; \quad t' = [\beta'\ a'];$
$E(a) = 0; \quad E(e) = 0; \quad E(y) = X\beta;$
$Var(a) = G; \quad var(e) = R; \quad cov\ (a, e) = 0; \quad var(y) = ZGZ' + R.$

A primeira derivação (Dempfle, 1977) usa o teorema de Bayes e assume-se que adicionalmente $t \sim N\ (r, M); \ e \sim N\ (0, R); \ y \sim N(X\beta, WMW' + R)$, onde:

$$E(t) = E \begin{pmatrix} \beta \\ a \end{pmatrix} = r = \begin{pmatrix} r_1 \\ 0 \end{pmatrix} \quad e \quad var(t) = M = \begin{pmatrix} S & 0 \\ 0 & G \end{pmatrix}$$

Dado $t,\ y \sim N(Wt, R)$, e dado $r,\ t \sim N(r, M)$. A verossimilhança de *y* dado *t* é

$L\ (t;\ y) = k_1\ exp[-(y - Wt)'R^{-1}\ (y - Wt)/2] = f\ (y/t)$, e a densidade da distribuição a priori de t é

$f(t) = k_2\ exp\ [-(t - r)'\ M^{-1}\ (t - r)/2].$

Assim, a distribuição a posteriori de t é proporcional a

$f(t/y) \propto L(t;\ y)\ f(t) \propto k_3\ exp\{-[(y - Wt)'R^{-1}\ (y - Wt) + (t - r)'M^{-1}\ (t-r)]/2\}.$

Esta equação pode ser expressa também em termos da densidade conjunta a posteriori dada por:

$$f(a, \beta) \propto exp\ \{ - \tfrac{1}{2} (y - X\beta - Za)' R^{-1} (y - X\beta - Za)\}\ x$$
$$x\ exp\ \{ - \tfrac{1}{2} (\beta - r_1)' S^{-1} (\beta - r_1)\}\ x$$
$$x\ exp\ \{ - \tfrac{1}{2} (a - E(a))' G^{-1} (a - E(a))\}.$$

Uma vez que esta distribuição é simétrica e unimodal (normal), a moda, a mediana e a média são idênticas e uma grande classe de funções de perda comum (função de perda quadrática, função de perda absoluta ou função de perda uniforme) conduz ao mesmo estimador. Determinando a moda obtém-se o vetor médio da distribuição conjunta a posteriori, por maximização e não integração. Diferenciando a expressão com respeito a t (β e a) e igualando a zero obtém-se:

$$\left[W' R^{-1} W + M^{-1}\right] \hat{t} = W' R^{-1} y + M^{-1} r.$$

Este sistema é equivalente a:

$$\begin{bmatrix} X' R^{-1} X + S^{-1} & X' R^{-1} Z \\ Z' R^{-1} X & Z' R^{-1} Z + G^{-1} \end{bmatrix} \begin{bmatrix} E(\beta | y) \\ E(a | y) \end{bmatrix} = \begin{bmatrix} X' R^{-1} Y + S^{-1}\ r_1 \\ Z' R^{-1} Y + G^{-1} 0 \end{bmatrix} \quad (2)$$

onde $r_1 = E(\beta)$ e $0 = E(a)$.

Tomando a distribuição a priori sobre os efeitos fixos como não informativa (expressa como $S \to \infty$ e então $S^{-1} \to 0$) , tem-se que esta equação resultante eqüivale as equações do modelo misto (EMM):

$$\begin{bmatrix} X' R^{-1} X & X' R^{-1} Z \\ Z' R^{-1} X & Z' R^{-1} Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X' R^{-1} y \\ Z' R^{-1} y \end{bmatrix}$$

A equação para $\hat{t}$, pode ser derivada também pela maximização de $f(y, t)$ para variações em t, sendo o estimador, neste caso, denominado máximo a posteriori (MAP) (Henderson, 1984).

A segunda derivação baseia-se em Robertson (1955) e fundamenta-se na combinação de dois estimadores (fontes de informação) independentes:

(i) o estimador BLUE ou de mínimos quadrados generalizados de t

$\hat{t}_1 = (W' R^{-1} W)^{-1}\ W' R^{-1} y$ e

(ii) o estimador a priori de t

$\hat{t}_2 = E(t) = r.$

Estes estimadores apresentam as matrizes de variância-covariância

$$var(\hat{t}_1) = (W' R^{-1} W)^{-1} \; M \quad \text{e} \quad var(\hat{t}_2) = 0,$$

com variâncias do erro de estimação

$$var(\hat{t}_1 - t) = (W' R^{-1} W)^{-1} \quad \text{e} \quad var(\hat{t}_2 - t) = M.$$

Combinando os dois estimadores obtém-se:

$$\hat{t}_c = [var(\hat{t}_1 - t)^{-1} + var(\hat{t}_2 - t)^{-1}]^{-1} \{var(\hat{t}_1 - t)^{-1} \; \hat{t}_1 + var(\hat{t}_2 - t)^{-1} \; \hat{t}_2\},$$

que eqüivale a $[W' R^{-1} W + M^{-1}] \; \hat{t}_c = W' R^{-1} y + M^{-1} r.$

Fazendo as mesmas suposições sobre os efeitos fixos, esta expressão eqüivale as equações de modelo misto. As equações de modelo misto apresentadas em (**2**) são denominadas equações de modelo misto de Robertson.

2.3 Relação entre Estimadores de Máxima Verossimilhança e Estimadores Bayesianos

Assumindo como uniforme a distribuição a priori dos parâmetros a serem estimados e maximizando (obtendo a moda) a distribuição a posteriori, o estimador resultante é de máxima verossimilhança (ML) (Henderson, 1984; Gianola & Fernando, 1986). De fato maximizando *f (a, β)* (tópico anterior) (mas considerando uma priori não informativa para β) com respeito a a e β obtém-se um estimador denominado de máxima verossimilhança, por Henderson et al. (1959), embora *f(a, β)* (tópico anterior) não seja uma função de verossimilhança e sim uma densidade a posteriori. Mesmo assim, pode ser obtido a partir das EMM que

$$E(\beta | y) = (X' V^{-1} X)^{-1} \; X' V^{-1} y = \hat{\beta} \text{ e}$$

$$E(a | y) = GZ' V^{-1} \; [y - X\beta] = \hat{a}$$

$\hat{\beta}$ é um estimador GLS e também ML de β e $\hat{a}$ é um estimador ML de *E(a/β, y)*, eqüivalendo à média da distribuição condicional na qual β é fixado.

2.4 Estimação Bayesiana de Componentes de Variância e Relação com ML e REML

No contexto dos modelos lineares mistos, os valores genéticos (θ_1) são preditos simultaneamente à estimação dos efeitos fixos (θ_2) e dos componentes de variância (θ_3). Na abordagem Bayesiana, a avaliação genética pode ser obtida, de maneira geral, pela construção da densidade a posteriori $f(\theta_1, \theta_2, \theta_3/y)$ e, se necessário, pela integração de $f(\theta_1, \theta_2, \theta_3/y)$ em relação a θ_2 e θ_3. θ_2 e θ_3 são denominados parâmetros de "nuissance" e por isso devem ser integrados, exceto θ_2 em alguns casos, onde o mesmo constitui-se em uma parte da função de mérito total (neste caso, a função de mérito depende da combinação linear de θ_1 e θ_2).

A obtenção de θ_1, requer a integração ou o conhecimento de θ_2 e θ_3. Henderson (1973) propôs o método BLUP para situações em que θ_3 é conhecido e θ_2 não o é. Para situações em que θ_3 não é conhecido, este autor sugeriu que o procedimento de máxima verossimilhança (ML) propiciaria estimativas razoáveis. Conforme Gianola & Fernando (1986), argumentos Bayesianos, que não requerem normalidade e linearidade permitem validar a intuição de Henderson.

A distribuição de θ_1, θ_2, e θ_3, dado y é proporcional a

$f(\theta_1, \theta_2, \theta_3/y) = f(y/\theta_1, \theta_2, \theta_3) \, . \, f(v\theta_1, \theta_2, \theta_3)$

Concentrando o interesse em θ_1 (o vetor de valores genéticos), deve-se integrar fora θ_2 e θ_3 através de

$$f(\theta_1 \mid y) = \int_{R_{\theta_2}} \int_{R_{\theta_3}} f(\theta_1 \mid \theta_2, \theta_3, y) \cdot f(\theta_2, \theta_3 \mid y) \cdot d\theta_2 \, d\theta_3 .$$

Tomando a distribuição conjunta a posteriori de forma que a maioria da densidade esteja na moda $(\hat{\theta}_2, \hat{\theta}_3)$, tem-se

$$f(\theta_1 \mid y) \doteq f(\theta_1 \mid \theta_2 = \hat{\theta}_2, \ \theta_3 = \hat{\theta}_3, \ y) .$$

Usando prioris não informativas para θ_2 e θ_3, tem-se que $\hat{\theta}_2$ *e* $\hat{\theta}_3$ são precisamente estimadores ML de θ_2 e θ_3, pois neste caso $f(\theta_2, \theta_3/y) \propto f(y/\theta_2, \theta_3)$, ou seja a densidade de θ_2 e θ_3 dado y é proporcional a função de verossimilhança, de forma que a moda da posteriori conjunta corresponde ao máximo da função de verossimilhança, produzindo estimadores ML.

Uma abordagem alternativa para inferência sobre θ_1 consiste em obter

$f(\theta_1, \theta_2 \mid y) \doteq f(\theta_1, \theta_2 \mid \theta_3 = \hat{\hat{\theta}}_3, y)$, *onde* $\hat{\hat{\theta}}_3$ refere-se à moda da densidade marginal de θ_3, dado y. Para obtenção de $\hat{\hat{\theta}}_3$ deve-se integrar θ_2 em $f(\theta_2, \theta_3 \mid y) \propto f(y \mid \theta_2, \theta_3)$ e então maximizar $f(\theta_3/y)$. Usando-se uma priori não informativa para θ_3, sob normalidade $\hat{\hat{\theta}}_3$ é um estimador de máxima verossimilhança restrita (REML) para θ_3 (Harville, 1977). Assim, se o interesse reside na inferência conjunta para θ_1 e θ_2 basta usar $f(\theta_1, \theta_2 \mid y) \doteq f(\theta_1, \theta_2 \mid \theta_3 = \hat{\hat{\theta}}_3, y)$, que sob normalidade é equivalente a solução das equações de modelo misto com θ_3 substituído pelas estimativas REML de θ_3 (desde que se tenha usado prioris não informativas para θ_2 e θ_3).

Inferências sobre componentes de variância devem ser baseadas em $f(\theta_3/y) \propto f(y/\theta_3) \cdot f(\theta_3)$, onde θ_3 contém variâncias e, portanto, $f(\theta_3/y)$ é definida na amplitude $(0, \infty)$ para cada um dos elementos de θ_3, de forma que nunca surgem problemas de estimativas negativas de componentes de variância (Box & Tiao, 1973). $f(\theta_3/y)$ é obtida integrando-se θ_1 em $f(\theta_1, \theta_2, \theta_3/y)$ produzindo $f(\theta_2, \theta_3/y)$ e integrando-se θ_2 nesta última. Neste caso, $f(\theta_2, \theta_3/y)$ conduz aos estimadores ML de θ_2 e θ_3 e $f(\theta_3/y)$ conduz a um estimador REML de θ_3. Segundo Gianola & Fernando (1986), isto (eliminação das influências de θ_2 ou dos efeitos fixos) mostra precisamente porque REML deve ser preferido em relação a ML, ou seja, estes argumentos são mais fortes do que os apresentados por Patterson & Thompson (1971), que enfatizaram a propriedade de vício do ML.

2.5 Problemas da Abordagem Freqüentista e Vantagens da Abordagem Bayesiana

As propriedades desejáveis dos estimadores pontuais e intervalares na estatística clássica são baseadas em hipotéticas repetições do experimento em um número infinito de vezes e não considera apenas os dados disponíveis. Isto significa que na teoria de amostragem, θ é uma quantidade conhecida fixa e o intervalo de confiança é aleatório. Assim, sob hipotética repetição de y, espera-se que o intervalo contenha θ em uma certa proporção das amostras (repetições). Entretanto, tendo-se os dados observados (uma repetição do vetor y), θ estará

dentro ou fora do intervalo. Em análise Bayesiana, inferências são realizadas diretamente da distribuição a posteriori, baseando-se apenas nos dados disponíveis e na distribuição a priori. A partir da posteriori, estimativas pontuais podem ser obtidas para resumir as características de tal distribuição e também estimativas intervalares e inferências probabilísticas podem ser realizadas.

Os procedimentos freqüentistas de estimação, ML e REML, apresentam propriedades desejáveis que são bem definidas somente para amostras muito grandes, ou seja, somente apresentam justificativas assintóticas. Entretanto, a maioria dos programas de melhoramento de espécies perenes baseia-se em amostras de tamanho finito (pequeno). O teorema de Bayes propicia soluções precisas para o problema de amostras de tamanho finito, pois para cada conjunto de dados, pequeno ou grande, existe uma distribuição a posteriori exata para realização de inferências. Existe também uma teoria assintótica Bayesiana, de forma que com grandes amostras obtém-se resultados similares aos obtidos com o método ML ou REML.

2.6 Implementação Prática da Análise Bayesiana

Os resultados de interesse gerados pela análise Bayesiana são, em geral, as distribuições marginais a posteriori dos parâmetros. Assim, inferências baseadas na média, mediana e moda e desvios padrões destas distribuições devem ser realizadas na prática.

O problema básico da implementação da análise Bayesiana refere-se à integração numérica. A integração (no espaço do parâmetro) da função densidade de probabilidade a posteriori, por exemplo:

$E\ [\ g\ (\theta)|y\ ] = \int_{R_\theta}\ g\ (\theta)\ \ p(\theta|y\ )\ d\theta$, onde:

$g\ (\theta) = \theta$, para obtenção da média a posteriori e

$g\ (\theta) = (\theta - \mu)^2,\ \ \mu = E(\theta|y\ )$, para obtenção da variância a posteriori ou risco de Bayes, pode ser realizada através dos métodos (Gamerman, 1997): (*i*) analítico para aproximação de integral; (*ii*) automáticos ou de quadratura; (*iii*) simulação estocástica para obtenção de distribuições a posteriori. Os métodos analíticos para aproximação de integrais, tais como a Discrepância de Kulback Liebler, a Aproximação Assintótica à Normal e o Método de Laplace são funcionais até determinada dimensão do problema. O método de quadratura é

puramente matemático e vem sendo substituído vantajosamente por métodos computacionalmente intensivos, os quais são essencialmente estatísticos, tais como os métodos de simulação estocástica. Dentre os métodos de simulação estocática destacam-se: (*i*) Monte Carlo; (*ii*) Monte-Carlo com Função de Importância; (*iii*) Reamostragem ou Bootstrap Bayesiano; (*iv*) Monte-Carlo – Cadeias de Markov.

Dentre estas três grandes classes de algorítmos para aproximar as integrais, os métodos de Monte-Carlo são largamente indicados e utilizados para integração multivariada. Os métodos de Monte-Carlo referem-se a processos de aproximação de valores esperados (integrais com respeito a uma distribuição de probabilidade) por meio de amostras, podendo ser referidos também como um caso especial de simulação de um processo estocástico.

Em genética quantitativa, para implementação prática da análise Bayesiana, uma das maiores dificuldades técnicas é a marginalização. A obtenção de distribuições marginais por processos analíticos é praticamente impossível (Sorensen, 1996). Assim, a obtenção da distribuição marginal a posteriori (marginalização da distribuição conjunta a posteriori) tem sido obtida pelo método da amostragem de Gibbs (GS) através da amostragem e atualização das distribuições condicionais. O método da amostragem de Gibbs pertence à classe de métodos, denominada Monte-Carlo – Cadeias de Markov, a qual é sustentada em propriedades das Cadeias de Markov. O nome Gibbs advém da distribuição de Gibbs, muito utilizada na área de mecânica estatística (Gamerman, 1996).

Para ilustrar a aplicação da técnica da amostragem de Gibbs na avaliação genética de espécies perenes será considerado o modelo individual univariado, conforme Sorensen (1996).

Modelo:

$y = X\beta + Za + e$, onde:

y = vetor de dados, de ordem n;
β = vetor de efeitos fixos, de ordem p;
a = vetor de valores genéticos aditivos, de ordem q;
e = vetor de erros, de ordem n;
X, Z = matrizes de incidência que associam β e a aos dados (y).

Assume-se inicialmente que a distribuição condicional dos dados, dados β, a e σ_e^2 é normal multivariada:

$y|\beta,\ a,\ \sigma_e^2 \sim N\ (X\beta + Za,\ I\ \sigma_e^2)$, onde I é a matriz identidade e σ_e^2 a variância residual.

Assumindo o modelo quantitativo infinitesimal, tem-se que a distribuição de a é também normal multivariada:

$a|A,\ \sigma_A^2 \sim N(O,\ A\ \sigma_A^2)$, onde A é a matriz de parentesco genético aditivo e σ_A^2 é a variância genética aditiva na população base.

Os parâmetros de interesse para inferências são: $\beta,\ a,\ \sigma_A^2$ e σ_e^2. Para conduzir a análise Bayesiana torna-se necessário especificar as distribuições a priori para $\beta,\ \sigma_A^2$ e σ_e^2 (a distribuição de a já foi especificada).

Como priori para β pode-se assumir $p(\beta) \propto$ constante, que especifica aproximadamente a noção de conhecimento a priori vago para β. Esta distribuição a priori é imprópria, mas pode-se tornar própria, desde que se especifique os limites superior e inferior para $p(\beta)$.

As distribuições a priori dos componentes de variância (σ_e^2 e σ_A^2) poderiam ser uniforme da forma $p\ (\sigma_i^2) \propto$ constante, $0 \leq \sigma_i^2 < \sigma_{i\ max}^2\ (i = e,\ A)$, onde, de acordo com o conhecimento acumulado sobre o caráter, $\sigma_{i\ max}^2$ seria o valor máximo que σ_i^2 poderia assumir, a priori. Alternativamente, poderia ser especificada uma priori mais informativa para os componentes de variância, assumindo uma distribuição qui-quadrado escalonada invertida, da forma:

$$p\ (\sigma_i^2|\upsilon_i,\ S_i^2)\ \propto\ (\sigma_i^2)^{-((\upsilon_i/2)+1)}\ exp\left[\frac{-\upsilon_i\ S_i^2}{2\ \sigma_i^2}\right] \qquad (i = e,\ A),$$

onde υ são os graus de liberdade da distribuição qui-quadrado e S_i^2, o valor inicial da variância. Esta distribuição reduz-se a uma distribuição uniforme imprópria se $\upsilon_i = -2\ e\ S_i^2 = 0$.

Definidas estas distribuições, pode-se agora escrever a distribuição conjunta a posteriori dos parâmetros do modelo.

$$p\ (\beta,\ a,\ \sigma_A^2,\ \sigma_e^2|y)\ \propto\ p\ (\beta,\ a,\ \sigma_A^2,\ \sigma_e^2)\ p\ (y|\beta,\ a,\ \sigma_A^2,\ \sigma_e^2)$$

$$= p\ (\beta)\ p\ (a|\sigma_A^2)\ p\ (\sigma_A^2)\ p\ (\sigma_e^2)\ p\ (y|\beta,\ a,\ \sigma_A^2,\ \sigma_e^2),$$

onde omitiu-se o condicionamento nos hiperparâmetros (parâmetros

que auxiliam na especificação da priori) e na conhecida matriz de parentesco *A*.

Considerando a distribuição a priori dos componentes de variância como uma qui-quadrado escalonada invertida, tem-se que a distribuição conjunta a posteriori pode ser reescrita:

$$p(\beta,\ a,\ \sigma_A^2,\ \sigma_e^2|y) \propto \sigma_e^{2-\left(\frac{n+\upsilon_e}{2}+1\right)}\ exp\left[-\frac{(y-X\beta-Za)'(y-X\beta-Za)+\upsilon_e\ S_e^2}{2\ \sigma_e^2}\right]$$
$$\sigma_A^{2-\left(\frac{q+\upsilon_A}{2}+1\right)}\ exp\left[-\frac{(a'A^{-1}a+\upsilon_A\ S_A^2}{2\ \sigma_A^2}\right]$$

Desejando assumir distribuição a priori uniforme para σ_A^2 e σ_e^2, basta fazer $\upsilon_i = -2$ e $S_i^2 = 0$ $(i = A,\ e)$ na expressão acima.

Para implementação do GS, deve-se derivar todas as distribuições condicionais a posteriori a partir da distribuição conjunta a posteriori apresentada acima.

Denominando-se $X\beta + Za = W\theta,\ onde\ W = [X\ \ Z]$ e $\theta' = [\beta'\ \ a']$, tem-se que a matriz dos coeficientes das equações de modelo misto é dada por $C = W'W + \Sigma$, onde $\Sigma = \begin{bmatrix} 0 & 0 \\ 0 & A^{-1}\sigma_e^2/\sigma_A^2 \end{bmatrix}$. A distribuição condicional a posteriori de θ' é:

$\theta \mid \sigma_A^2,\ \sigma_e^2,\ y \sim N(\hat{\theta},\ C^{-1}\sigma_e^2)$, *em que* $\hat{\theta}$ *é dado por* $C\hat{\theta} = W'y$, ou seja, pelas equações de modelo misto.

Para derivar a distribuição condicional a posteriori β_i (o *i*-ésimo elemento do vetor β) deve-se fazer as partições:

$$\theta = (\theta_i,\ \theta_{-i});\ W' = (W_i\ W_{-i})';\ C = \begin{bmatrix} C_{i,i} & C_{i,-i} \\ C_{-i,i} & C_{-i,-i} \end{bmatrix} = \begin{bmatrix} X'_iX_i & X'_iX_{-i} & X'_iZ \\ X'_{-i}X_i & X'_{-i}X_{-i} & X'_{-i}Z \\ Z'X_i & Z'X_{-i} & Z'Z+A^{-1}\alpha \end{bmatrix},$$

onde: $\alpha = \sigma_e^2/\sigma_A^2$, β_{-i} é o vetor de efeitos fixos com o *i*-ésimo elemento excluído, de ordem *p-1*, X_i é o i-ésimo vetor coluna da matriz *X* e X_{-i} é a matriz *X* com X_i excluído.

Assim, tem-se:

$\theta_i|\theta_{-i},\ \sigma_A^2,\ \sigma_e^2,\ y \sim N(\hat{\theta}_i, C_{i,i}^{-1}\sigma_e^2)$, *em que* $\hat{\theta}_i$ *é dado por* $C_{i,i}\ \hat{\theta}_i = (W'_i\ y - C_{i,-i}\ \theta_{-i})$.

Tem-se também:

$W'_i\ y = X'_i\ Y;\ (X'_i\ \ X_i)\ \hat{\beta}_i = X'_i\ y - X'_i\ \ X_{-i}\ \beta_{-i} - X'_i\ \ Za$ e finalmente pode-se escrever a distribuição condicional a posteriori de β_i:

$$\beta_i|\beta_{-i},\ a,\ \sigma_A^2,\ \sigma_e^2,\ y \sim N\ (\hat{\beta}_i, (X'_i X_i)^{-1}\sigma_e^2),$$
$$em\ \ que:\ \ \hat{\beta}_i = (X'_i X_i)^{-1} X'_i (y - X_{-i}\beta_{-i} - Za) \qquad (3)$$

Para derivar a distribuição condicional a posteriori de a_i, deve-se proceder de maneira similar. Assim, tem-se:

$$A^{-1} = \begin{bmatrix} A_{ii}^{-1} & A_{i,-i}^{-1} \\ A_{-i,i}^{-1} & A_{-i,-i}^{-1} \end{bmatrix};\quad C = \begin{bmatrix} X'X & X'z_i & X'Z_{-i} \\ z'_i X & z'_i z_i + A_{i,i}^{-1}\alpha & z'_i Z_{-i} + A_{i,-i}^{-1}\alpha \\ Z'_{-i} X & Z'_{-i} z_i + A_{-i,i}^{-1}\alpha & Z'_{-i} Z_{-i} + A_{-i,-i}^{-1}\alpha \end{bmatrix}$$

$$W'_i\ y = z'_i\ y;\ (z'_i\ z_i + A_{i,i}^{-1}\alpha)\ \hat{a}_i = z'_i\ y - z'_i X\beta - (z'_i\ Z_{-i} + A_{i,-i}^{-1}\alpha)\ a_{-i}$$
$$= z'_i\ y - z'_i X\beta - A_{i,-i}^{-1}\alpha\ a_{-i},\ \ pois\ \ z'_i\ Z_{-i} = 0$$

Finalmente, a distribuição condicional a posteriori de a_i é:

$$a_i|\beta,\ a_{-i},\ \sigma_A^2,\ \sigma_e^2,\ y \sim N\ (\hat{a}_i, (z'_i z_i + A_{i,i}^{-1}\alpha)^{-1}\sigma_e^2) \qquad (4)$$

A obtenção das distribuições condicionais a posteriori completas dos componentes de variância será descrita a seguir. Para σ_A^2, deve-se tomar da distribuição conjunta a posteriori somente os termos que envolvem σ_A^2. Assim:

$$p\ (\sigma_A^2|\beta,\ a,\ \sigma_e^2,\ y)\ \alpha\ (\sigma_A^2)^{-\frac{q+\upsilon_A}{2}+1}\ \ exp\ \ -\frac{a'A^{-1}a + \upsilon_A\ S_A^2}{2\ \sigma_A^2}$$

$$= (\sigma_A^2)^{-\frac{\tilde{\upsilon}_A}{2}+1}\ \ exp\ \ -\frac{\tilde{\upsilon}_A\ \tilde{S}_A^2}{2\ \sigma_A^2},\ \text{em que}$$

$$\tilde{S}_A^2 = (a'A^{-1}a + \upsilon_A\ S_A^2)/\tilde{\upsilon}_A\ \ e\ \ \tilde{\upsilon}_A = q + \upsilon_A.$$

A expressão de $p\ (\sigma_A^2|\beta,\ a,\ \sigma_e^2,\ y)$ é proporcional à distribuição qui-quadrado (χ^2) escalonada invertida, com $\tilde{\upsilon}_a$ graus de liberdade e parâmetro de escala $a'A^{-1}a + \upsilon_A\ S_A^2$.

Assim:

$$\sigma_A^2|\beta,\ a,\ \sigma_e^2,\ y \sim \tilde{\upsilon}_A\ \tilde{S}_A^2\ \chi^2_{\tilde{\upsilon}_A} \qquad (5)$$

e para amostrar desta distribuição deve-se inicialmente trabalhar com um qui-quadrado com $\tilde{\upsilon}_A = q + \upsilon_A$ graus de liberdade, inverter este número e posteriormente multiplicá-lo por $a' A^{-1} a + \upsilon_A S_A^2$.

De maneira similar tem-se para σ_e^2:

$$p(\sigma_e^2 | \beta, a, \sigma_A^2, y) \propto (\sigma_e^2)^{-\left(\frac{n+\upsilon_e}{2}+1\right)} exp\left[\frac{(y - X\beta - Za)'(y - X\beta - Za) + \upsilon_e S_e^2}{2\sigma_e^2}\right]$$

$$= (\sigma_e^2)^{-\frac{\tilde{\upsilon}_e}{2}+1} exp\left[\frac{-\tilde{\upsilon}_e \tilde{S}_e^2}{2\sigma_e^2}\right], \text{ onde}$$

$\tilde{\upsilon}_e = n + \upsilon_e$ e $\tilde{S}_e^2 = ((y - X\beta - Za)'(y - X\beta - Za) + \upsilon_e S_e^2)/\tilde{\upsilon}_e$.

A expressão de $p(\sigma_e^2 | \beta, a, \sigma_A^2, y)$ é proporcional à distribuição qui-quadrado escalonada invertida, com $\tilde{\upsilon}_e = n + \upsilon_e$ graus de liberdade e parâmetro de escala $(y - X\beta - Za)'(y - X\beta - Za) + \upsilon_e S_e^2$.

Assim, tem-se:

$$\sigma_e^2 | \beta, a, \sigma_A^2, y \sim \tilde{\upsilon}_e \tilde{S}_e^2 \chi^{-2}_{\tilde{\upsilon}_e} \qquad (6)$$

A implementação da amostragem de Gibbs consiste em amostrar sucessivamente de (3), (4), (5) e (6).

Considere o exemplo:

Indivíduo	Bloco	Pai	Mãe	Y_i
1	1	-	-	-
2	2	-	-	-
3	1	1	-	Y_3
4	2	1	2	Y_4
5	1	3	4	Y_5
6	2	1	4	Y_6
7	1	5	6	-

Considerando o modelo individual tem-se $\beta = (\beta_1, \beta_2)'$ para os blocos 1 e 2 e $a = (a_1, \ldots, a_7)'$ para os valores genéticos aditivos dos sete indivíduos. Assumindo distribuições a priori uniforme para os

componentes de variância, tais que $\upsilon_i = -2,\ S_i^2 = 0,\ (i = A,\ e)$ e valores iniciais $\sigma_A^2 = 5{,}0$ *e* $\sigma_e^2 = 5{,}0$, deve-se trabalhar com as equações de modelo misto atualizando o valor $\alpha = \sigma_e^2 / \sigma_A^2$, a cada iteração. A matriz dos coeficientes das equações de modelo misto em função de α eqüivale a:

$$C = \begin{array}{c} \\ \beta_1 \\ \beta_2 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \end{array} \begin{array}{c} \begin{array}{ccccccccc} \beta_1 & \beta_2 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{array} \\ \left[\begin{array}{ccccccccc} 2{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 1{,}00 & 0{,}00 & 1{,}00 & 0{,}00 & 0{,}00 \\ 0{,}00 & 2{,}00 & 0{,}00 & 0{,}00 & 0{,}00 & 1{,}00 & 0{,}00 & 1{,}00 & 0{,}00 \\ 0{,}00 & 0{,}00 & 2{,}33\alpha & 0{,}50\alpha & -0{,}66\alpha & -0{,}50\alpha & 0{,}00\alpha & -1{,}00\alpha & 0{,}00\alpha \\ 0{,}00 & 0{,}00 & 0{,}50\alpha & 1{,}50\alpha & 0{,}00\alpha & -1{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha \\ 1{,}00 & 0{,}00 & -0{,}66\alpha & 0{,}00\alpha & 1+2{,}83\alpha & 0{,}50\alpha & -1{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha \\ 0{,}00 & 1{,}00 & -0{,}50\alpha & -1{,}00\alpha & 0{,}50\alpha & 1+3{,}00\alpha & -1{,}00\alpha & -1{,}00\alpha & 0{,}00\alpha \\ 1{,}00 & 0{,}00 & 0{,}00\alpha & 0{,}00\alpha & -1{,}00\alpha & -1{,}00\alpha & 1+2{,}62\alpha & 0{,}62\alpha & -1{,}23\alpha \\ 0{,}00 & 1{,}00 & -1{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha & -1{,}00\alpha & 0{,}62\alpha & 1+2{,}62\alpha & -1{,}23\alpha \\ 0{,}00 & 0{,}00 & 0{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha & 0{,}00\alpha & -1{,}23\alpha & -1{,}23\alpha & 2{,}46\alpha \end{array}\right] \end{array}$$

Os parâmetros de locação são $\theta = (\beta_1,\ \beta_2,\ a_1,\ \ldots\ a_7)$ e assumindo que os valores iniciais são zero, tem-se $\theta^{(0)} = 0$ e pode-se iniciar a amostragem com uma cadeia de comprimento igual a *n (i = 1, ..., n)*.

A primeira amostragem para $\beta_1^{(1)}$ é obtida de **(3)** ou seja obtém-se $\beta_1^{(1)}$ de $\beta_1|\beta_2,\ a,\ \sigma_A^2,\ \sigma_e^2,\ y \sim N\ (\hat{\beta}_1, (2)^{-1}\sigma_e^2)$. Dados os valores iniciais, a média da distribuição normal é $\hat{\beta}_1 = (Y_3 + Y_5)/2$ e a variância é $(2)^{-1}\sigma_e^2$, ou seja, 5/2.

$\beta_2^{(1)}$ por sua vez é retirado de uma distribuição normal com média $\hat{\beta}_2 = (Y_4 + Y_6)/2$ e variância 5/2.

$a_1^{(1)}$ por sua vez é retirado de $a_1|\beta,\ a_{-1},\ \sigma_A^2,\ \sigma_e^2,\ y \sim N\ (\hat{a}_1, (2{,}33\alpha)^{-1}\sigma_e^2)$, que provém de **(4)**, onde $\hat{a}_1 = 0/2{,}33\quad \alpha = 0,\ \sigma_e^2 = 5\ e\ \alpha = 1$.

$a_2^{(1)}$ é retirado de uma distribuição normal $N\ (\hat{a}_2, (1{,}50\ \alpha)^{-1}\sigma_e^2)$, onde $\hat{a}_2 = (0 - 0{,}50\alpha\ a_1^{(1)})/(1{,}50\alpha),\ \sigma_e^2 = 5\ e\ \alpha = 1$, onde 0,50 provém de C.

Para o indivíduo número 6, o valor genético na primeira iteração, $a_6^{(1)}$, é retirado de:

$N\ (\hat{a}_6, (1+2{,}62\ \alpha)^{-1}\sigma_e^2)$, onde

$$\hat{a}_6 = \frac{Y_6 - \beta_2^{(1)} - (-1{,}00\ \alpha\ a_1^{(1)} - 1{,}00\ \alpha\ a_4^{(1)} + 0{,}62\ \alpha\ a_5^{(1)})}{1 + 2{,}62\ \alpha},\ \sigma_e^2 = 5\ e\ \alpha = 1.$$

Para o indivíduo número 7, o valor genético na primeira iteração, $a_7^{(1)}$, é retirado de:

$N\ (\hat{a}_7,\ (2{,}46\ \alpha)^{-1}\sigma_e^2)$, onde

$$\hat{a}_7 = \frac{0 - (-1{,}23\ \alpha\ a_5^{(1)} - 1{,}23\ \alpha\ a_6^{(1)})}{2{,}46\ \alpha},\ \sigma_e^2 = 5\ e\ \alpha = 1.$$

Tendo-se amostrados todos os parâmetros de locação do modelo, deve-se computar:

$$SS_e^{(1)} = (y - X\beta^{(1)} - Za^{(1)})'(y - X\beta^{(1)} - Za^{(1)})$$

$$SS_A^{(1)} = (a^{(1)})' A^{-1} a^{(1)}$$

A primeira iteração do amostrador é completada, retirando-se os componentes de variância, usando $SS_A^{(1)}$ e $SS_e^{(1)}$:

$$\sigma_A^2 | \beta,\ a,\ \sigma_e^2, y \sim SS_a^{(1)}\ \chi_{q-2}^{-2} \tag{7}$$

$$\sigma_e^2 | \beta,\ a,\ \sigma_A^2, y \sim SS_e^{(1)}\ \chi_{n-2}^{-2} \tag{8}$$

A segunda iteração inicia-se através de atualizações das equações de modelo misto com $\alpha = \sigma_e^2 / \sigma_A^2$, onde σ_e^2 e σ_A^2 são os valores amostrados de **(7)** e **(8)** e prossegue com a amostragem de $\theta^{(2)}$ da mesma maneira realizada para $\theta^{(1)}$ e com o cômputo de $SS_e^{(2)}$ e $SS_A^{(2)}$ da mesma maneira realizada para $SS_e^{(1)}$ e $SS_A^{(1)}$, e assim sucessivamente.

Em termos mais simples, o algorítmo GS pode ser apresentado de forma resumida:

1. Fornecer os valores iniciais dos parâmetros de locação e dispersão do modelo. Estes valores iniciais podem ser calculados através de procedimentos padrões tais como a estimação de componentes de variância por REML ou quadrados mínimos. Assumindo a média geral $\bar{y}$ como único efeito fixo, pode-se calcular $\bar{y}$ como a média aritmética das observações e $a_i = h^2(y_i - \bar{y})$. Devem ser fornecidos os valores iniciais para $\bar{y}_i,\ a_i,\ \sigma_e^2,\ \sigma_A^2\ e\ \alpha = \sigma_e^2 / \sigma_A^2$.

2. Gerar valores para os efeitos fixos. Sendo o único efeito fixo, a média geral, tem-se:

$$\hat{\bar{y}} = \bar{y} + rnd \;\; \sigma_e /(n)^{1/2}$$

3. Gerar valores para os efeitos aleatórios:

$\hat{a} = a_i + rnd \;\; [(1 - r_{IA}^2) \;\; \sigma_A^2 \,]^{1/2}$, onde r_{IA} é a acurácia dada por $r_{IA} = (1 - PEV_i / \sigma_A^2)^{1/2}$, onde PEV_i é o iésimo elemento da inversa da matriz dos coeficientes das EMM multiplicado por σ_e^2.

4. Calcular a soma de quadrados do resíduo (SSE) e a variância residual σ_e^2. Assumindo que a distribuição a priori para a variância residual é a inversa de uma qui-quadrado, tem-se:

$$SSE = \sum (y_i - \hat{\bar{y}} - \hat{a}_i)^2$$

$$\sigma_e^2 = \frac{SSE}{X_n^2}$$

5. Gerar um valor para a variância dos efeitos aleatórios de valores genéticos.

$$\sigma_A^2 = \frac{\hat{a}' A^{-1} \hat{a}}{X_q^2}$$

6. Calcular o novo valor do parâmetro

$$\hat{\alpha} = \frac{\sigma_e^2}{\sigma_A^2}$$

7. Repetir os passos de (2) a (6) até que se obtenha a convergência da cadeia

Devido ao fato de que valores aleatórios são utilizados inicialmente como realização do conjunto de parâmetros, é necessário um período de descarte de amostras até que as amostras de GS possam ser consideradas como provenientes da distribuição conjunta a posteriori, ou seja da distribuição em equilíbrio estacionário. Em geral, tem sido utilizado o esquema tradicional de cadeia longa (única) de Gibbs, onde o processo de reamostragem é contínuo. Assim, de maneira geral, um grande (da ordem de 10.000 ou 1.000.000) número de ciclos tem sido utilizado, sendo descartadas as primeiras amostras (da ordem de poucos milhares) e amostras de cada parâmetro são salvas a cada

pequeno (da ordem de 50 a 100) número de iterações. O intervalo entre amostras salvas é necessário como forma de obtenção de amostras independentes, visto que amostras sucessivas apresentam correlação serial. O número total de amostras salvas são utilizadas para cômputo das estimativas pontuais e intervalares de interesse. O software MTGSAM (Van Tassell & Van Vleck, 1995) tem sido utilizado com sucesso para a implementação da análise Bayesiana.

2.7 Aplicação a dados experimentais

Foram utilizados dados de quatro experimentos (testes de progênies de irmãos germanos) desbalanceados de *Pinus caribaea* var. *hondurensis* com um total de 89 progênies (tratamentos), sendo 65 diferentes e algumas comuns aos experimentos. As 65 progênies foram obtidas de 76 genitores e no total foram avaliados 3.886 indivíduos. Os experimentos foram implantados pela Champion Papel e Celulose S.A., no município de Santana no Estado do Amapá, no delineamento em blocos ao acaso com 9 repetições e 5 plantas por parcela. Considerou-se a variável diâmetro avaliada aos 9,5 anos e o modelo aditivo infinitesimal.

O modelo linear univariado eqüivale a:

$y = X\beta + Za + Wc + e$, em que:

y, β, a, c e e = vetores de dados, de efeitos fixos (médias de blocos), de valores genéticos, de efeitos de parcela e erros aleatórios, respectivamente.

X, Z e W = matrizes de incidência para β, a e c, respectivamente.

As seguintes distribuições são assumidas:

$$y \left| \beta, a, c, \sigma_e^2 \right. \sim N \left(X\beta + Za + Wc, I\sigma_e^2 \right)$$
$$a \left| A, \sigma_A^2 \right. \sim N \left(0, A\sigma_A^2 \right)$$
$$c \left| \sigma_c^2 \right. \sim N \left(0, I\sigma_c^2 \right)$$
$$e \left| \sigma_e^2 \right. \sim N \left(0, I\sigma_e^2 \right)$$

$COV(a, c') = 0;\ COV(a, e') = 0;\ COV(c, e') = 0$, em que:

A = matriz de parentesco genético aditivo.
I = matriz identidade de ordem apropriada.

σ_A^2, σ_c^2 e σ_e^2 = variâncias genética aditiva, entre parcelas e residual, respectivamente.

Sob este modelo, a herdabilidade (h^2) e correlação devida ao ambiente comum (c^2) são, respectivamente:

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_c^2 + \sigma_e^2}; \qquad c^2 = \frac{\sigma_c^2}{\sigma_A^2 + \sigma_c^2 + \sigma_e^2}$$

A análise Bayesiana deste modelo linear mais geral (incluindo o efeito de parcela) é apresentada a seguir:

Distribuição conjunta a posteriori dos parâmetros do modelo

$$p(\beta, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2|y) \propto p(\beta, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2)\, p(y|\beta, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2)$$
$$= p(\beta)\, p(a|\sigma_A^2)\, p(c|\sigma_c^2)\, p(\sigma_A^2)\, p(\sigma_c^2)\, p(\sigma_e^2)y|\beta, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2),$$

Assumindo a distribuição a priori dos componentes da variância como uma qui-quadrado escalonada invertida, tem-se que:

Distribuição conjunta a posteriori

$$p(\beta, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2|y) \propto \sigma_e^{-\left(\frac{n+\upsilon_e}{2}+1\right)} exp\left[(y-X\beta-Za-Wc)'(y-X\beta-Za-Wc)+\upsilon_e S_e^2\right]$$

$$\sigma_A^{-\left(\frac{q+\upsilon_A}{2}+1\right)} exp\left[-\frac{(a'A^{-1}a+\upsilon_A S_A^2}{2\sigma_A^2}\right]$$

$$\sigma_c^{-\left(\frac{p+\upsilon_c}{2}+1\right)} exp\left[-\frac{(c'c+\upsilon_c S_c^2}{2\sigma_c^2}\right]$$

Incorporando na matriz *C* (item 2.6) as submatrizes relacionando *W* com *Z* e *X*, obtém-se as distribuições condicionais a posteriori para cada parâmetro:

Distribuição condicional a posteriori de β_i

$$\beta_i|\beta_{-i}, a, c, \sigma_A^2, \sigma_c^2, \sigma_e^2, y \sim N(\hat{\beta}_i, (X'_i X_i)^{-1}\sigma_e^2),$$
$$em\ que: \hat{\beta}_i = (X'_i X_i)^{-1} X'_i (y - X_{-i}\beta_{-i} - Za - Wc)$$

Distribuição condicional a posteriori de a_i

$$a_i|\beta, a_{-i},\ c,\ \sigma_A^2,\ \sigma_c^2,\ \sigma_e^2,\ y \sim N\ (\hat{a}_i, (Z'_i Z_i + A^{-1}_{i,\ i}\alpha)^{-1}\sigma_e^2)$$

Distribuição condicional a posteriori de c_i

$$c_i|\beta, a,\ c_{-i},\ \sigma_A^2,\ \sigma_c^2,\ \sigma_e^2,\ y \sim N\ (\hat{c}_i, (W'_i W_i + \alpha^*)^{-1}\sigma_e^2)$$

Distribuição condicional a posteriori de σ_A^2

$$\sigma_A^2|\beta,\ a,\ c,\ \sigma_c^2,\ \sigma_e^2,\ y \sim \tilde{\upsilon}_A\ \tilde{S}_A^2\ \chi^2_{\tilde{\upsilon}_A}$$

Distribuição condicional a posteriori de σ_C^2

$$\sigma_c^2|\beta,\ a,\ c,\ \sigma_A^2,\ \sigma_e^2,\ y \sim \tilde{\upsilon}_C\ \tilde{S}_c^2\ \chi^2_{\tilde{\upsilon}_c}$$

Distribuição condicional a posteriori de σ_e^2

$$\sigma_e^2|\beta,\ a,\ c,\ \sigma_A^2,\ \sigma_c^2,\ y \sim \tilde{\upsilon}_e\ \tilde{S}_e^2\ \chi^2_{\tilde{\upsilon}_e} \text{ em que } \alpha^* = \frac{\sigma_e^2}{\sigma_c^2}$$

e os demais parâmetros são definidos analogamente ao item 2.6.

Além das distribuições assumidas para os efeitos aleatórios no modelo linear clássico e para a verossimilhança do vetor de observações, a abordagem Bayesiana requer atribuições para as distribuições a priori dos efeitos fixos e componentes de variância. Uma distribuição a priori não informativa ou uniforme foi atribuída a β, refletindo conhecimento a priori vago sobre os efeitos fixos, de forma que $p(\beta) \propto K$ (constante). Para os componentes de variância, distribuições χ^2 inversas foram assumidas como priori, conforme especificado no tópico anterior. Considerou-se adicionalmente $\upsilon_i = -2$ e $S_i^2 = 0$, de forma que a distribuição se tornasse uniforme, e, portanto, não informativa.

Para a inferência Bayesiana sobre os parâmetros de interesse empregou-se a técnica da amostragem de Gibbs. O principal aspecto deste procedimento refere-se ao fato das inferências basearem-se na distribuição marginal a posteriori dos parâmetros, sendo que a marginalização da distribuição conjunta a posteriori é obtida via o amostrador de Gibbs através de amostragens e atualizações das distribuições condicionais. Em resumo, a abordagem Bayesiana baseia-se na construção da distribuição marginal a posteriori de um parâmetro de interesse tratando-o como uma variável aleatória e aplicando cálculo de

probabilidades. Este procedimento implica problemas multidimensionais, uma vez que todos os outros parâmetros do modelo devem ser integrados (eliminados), fato que raramente é possível usando os métodos numéricos padrões. O procedimento iterativo da amostragem de Gibbs refere-se a uma técnica de integração estocástica que cria uma cadeia de Markov, que é uma distribuição (conjunta a posteriori) estacionária associada à distribuição a posteriori de interesse. Tomando-se amostras, iterativamente, das distribuições condicionais a posteriori, com contínua atualização, obtém-se a distribuição conjunta a posteriori em equilíbrio e, após um número de iterações suficientemente grande, a última amostra desta seqüência e qualquer amostra subseqüente é uma amostra da distribuição conjunta a posteriori. Este resultado implica que cada coordenada do vetor de amostras retiradas, $\theta^n = [\beta^n \; a^n \; c^n, \; \sigma_A^{2(n)}, \; \sigma_c^{2(n)}, \; \sigma_e^{2(n)}]$, é uma amostra da distribuição marginal a posteriori apropriada.

Para a análise Bayesiana empregou-se o software MTGSAM (Van Tassell & Van Vleck, 1995) e para a análise freqüentista empregou-se o software MTDFREML (Boldman et al., 1995). Foram avaliados vários tamanhos de cadeia de Markov e intervalos entre amostragens.

3 Resultados e discussão

Para análise de convergência do processo de estimação foram consideradas cadeias de tamanhos 10.000, 15.000 e 20.000 ciclos, intervalos de 50 e 100 entre amostras dos parâmetros de interesse e período de descarte de 2.000 ciclos. Alguns resultados referentes aos parâmetros genéticos são apresentados na Tabela 1.

Tabela 1 – Resultados comparativos entre algumas estratégias de amostragem de Gibbs empregadas, para estimação (médias) dos parâmetros genéticos herdabilidade (h^2) e correlação de ambiente comum (c^2).

Tamanho da cadeia	Intervalo entre amostras	h^2	c^2
10.000	100	0,34	0,032
20.000	100	0,34	0,032

15.000	50	0,38	0,029
20.000	50	0,36	0,030

Os resultados apresentados na Tabela 1 juntamente com outros não apresentados permitiram concluir que a convergência do processo de amostragem ocorreu com intervalos de 100 entre amostras e com tamanho de cadeia de 10.000 ciclos (visto que a cadeia de 10.000 ciclos encontra-se dentro da cadeia de 20.000 ciclos e que ambas conduziram à resultados idênticos).

Outra forma de análise de convergência refere-se à estimação do erro de Monte Carlo, que é uma estatística associada ao erro de estimação de determinado parâmetro devido ao número de amostras utilizadas na cadeia de Gibbs, sendo que este erro é inversamente proporcional ao tamanho da cadeia (Van Tassel & Van Vleck, 1996). Este erro pode ser calculado através da variância dos parâmetros amostrados sucessivamente a cada intervalo dividida pelo número de amostras salvas, sendo que a raiz quadrada deste erro fornece uma aproximação para o desvio padrão do erro associado ao comprimento da cadeia. Considerando a cadeia de tamanho 20.000 e intervalo de 100 entre amostras o desvio padrão do erro associado ao tamanho da cadeia foi de 0,0004 para a herdabilidade (Tabela 2), resultado que revela que os 18.000 ciclos após o período de descarte, foram suficientes para obtenção de precisas estimativas das médias a posteriori. Dessa forma, optou-se por apresentar os resultados referentes ao tamanho de cadeia de 20.000 ciclos e com amostras salvas a cada 100 ciclos. O tempo de processamento para esta análise foi de cerca de 10 horas em um computador Pentium 100 Mhz, com 16 MB de memória.

Na Tabela 2 são apresentadas as estimativas de parâmetros genéticos pelos procedimentos REML e GS.

Os resultados pontuais obtidos pelas abordagens freqüentista e Bayesiana foram similares (Tabela 2), conforme esperado. Uma vez que foram utilizadas distribuições a priori não informativas na análise Bayesiana, as modas das distribuições marginais a posteriori dos parâmetros genéticos foram similares às correspondentes estimativas REML. Do ponto de vista Bayesiano, as estimativas obtidas por REML correspondem às modas das distribuições conjuntas a posteriori dos componentes de variância, dada a utilização de prioris uniformes para os efeitos fixos e componentes de variância. A técnica GS possibilitou estimativas de componentes de variância pelo método VEIL (ou da verossimilhança integrada) apresentado por Gianola & Foulley (1990).

Tabela 2 – Estimativas descritivas das distribuições marginais a posteriori da herdabilidade (h^2), correlação devida ao ambiente comum (c^2) e variância fenotípica (σ_F^2) obtidas pelo procedimento da amostragem de Gibbs (GS), bem como estimativas destes parâmetros obtidas pelo método da máxima verossimilhança restrita (REML). Caráter diâmetro em *Pinus caribaea* var. *hondurensis*.

	GS					REML
Parâmetro	Média ± Desvio Padrão	Moda	Mediana	Desvio Padrão Monte Carlo	Intervalo de Confiança (95%)	Estimativa Pontual
h^2	0,34 ± 0,005	0,34	0,34	0,0004	0,33 – 0,35	0,32 ± 0,06
c^2	0,032 ± 0,003	0,032	0,032	0,0002	0,026 – 0,038	0,031 ± 0,01
σ_F^2	8,2441 ± 0,02	-	-	-	-	8,0624

A grande vantagem da análise Bayesiana, neste caso, refere-se à obtenção dos desvios padrões e intervalos de confiança exatos para os parâmetros genéticos e valores genéticos preditos. Verificou-se que a técnica GS propiciou estimativas muito mais precisas que aquelas obtidas por REML. Na Tabela 3 são apresentados os valores genéticos preditos dos melhores genitores obtidos pelos procedimentos GS e BLUP, bem como os ganhos genéticos associados.

As diferenças observadas nos valores genéticos preditos pelos procedimentos GS e BLUP/REML foram pequenas, conduzindo a uma ligeira alteração na ordem dos melhores pelos dois procedimentos (Tabela 3). Do ponto de vista Bayesiano, uma vez que nenhuma distinção existe entre efeitos fixos e aleatórios, as soluções das equações de modelo misto do BLUP correspondem às médias das distribuições marginais a posteriori dos parâmetros de locação (efeitos fixos e aleatórios), dados os componentes de variância, ou parâmetros de dispersão. As condições para que haja esta correspondência entre BLUP e GS para os parâmetros de locação são: atribuição de prioris não informativas para os efeitos fixos, prioris normais para os efeitos aleatórios e verossimilhança normal para o vetor de observações. Estas premissas foram atendidas no presente trabalho, fato que explica os resultados obtidos.

Tabela 3 – Estimativas descritivas das distribuições marginais a posteriori dos valores genéticos preditos dos 10 melhores genitores obtidos pelo procedimento da amostragem de Gibbs (GS), bem como valores genéticos preditos pelo procedimento BLUP. Caráter diâmetro em *Pinus caribaea* var. *hondurensis.*

Genitor	Média ± Desvio Padrão	Ordem (GS)	BLUP ± SEP*	Acurácia (BLUP)	Ordem (BLUP)
130	2,70 ± 0,03	1	2,61 ± 0,80	0,86	1
49	2,23 ± 0,07	2	2,15 ± 0,94	0,81	3
104	2,19 ± 0,06	3	2,26 ± 0,86	0,84	2
73	2,07 ± 0,07	4	1,96 ± 0,99	0,78	5
114	2,03 ± 0,09	5	1,96 ± 1,06	0,75	6
56	1,99 ± 0,07	6	1,89 ± 0,96	0,80	7
100	1,97 ± 0,03	7	1,97 ± 0,78	0,87	4
92	1,90 ± 0,03	8	1,82 ± 0,79	0,87	8
76	1,80 ± 0,05	9	1,81 ± 0,96	0,80	9
125	1,79 ± 0,03	10	1,69 ± 0,65	0,91	10
		GS			BLUP
Ganho genético (cm)	2,066 + 0,04		(1,99 – 2,14)**	2,011 + 0,28	(1,46 – 2,56)***
Ganho genético (%)	13,93 + 0,27		(13,40 – 14,46)**	13,56 + 1,88	(9,88 – 17,24)***

* desvio padrão do erro de predição.
** intervalo com 95% de confiança.
*** cálculo considerando uma acurácia média de 0,88.

Segundo Wang et al. (1994), uma deficiência do procedimento BLUP/REML é que os erros de estimação dos componentes de dispersão não são considerados por ocasião da predição de valores genéticos. A abordagem Bayesiana e a teoria de probabilidade à ela associada impõem que as inferências devem ser baseadas nas distribuições marginais a posteriori dos parâmetros de interesse, de forma que toda a incerteza sobre os parâmetros são consideradas completamente. Isto é conseguido a partir da distribuição conjunta a posteriori de todos os parâmetros. Desta distribuição, a distribuição marginal a posteriori do valor genético de um indivíduo é obtida por sucessivas

integrações (eliminações) de todos os parâmetros de "nuissance" do modelo, tais como efeitos fixos (blocos no caso), outros efeitos aleatórios (parcela no caso) e componentes de variância e parâmetros associados (h^2 e c^2 no caso).

O aspecto ressaltado por Wang et al. (1994) é muito relevante considerando, principalmente, que para que as propriedades desejáveis do BLUP sejam asseguradas é necessário o conhecimento exato dos componentes de variância. Dessa forma, a análise Bayesiana, que permite estimativas quase exatas dos componentes de variância, tende a conduzir também a uma maior aproximação entre ganhos genéticos preditos e realizados com seleção. Este último aspecto pode, em parte, ser visualizado através dos desvios padrões dos ganhos genéticos apresentados na Tabela 3, sendo que o procedimento GS propiciou uma inferência muito mais segura (menor desvio padrão ou risco) sobre o ganho genético a ser capitalizado com a seleção dos 10 melhores genitores.

Os parâmetros genéticos iniciais utilizados foram obtidos pelo procedimento REML e as distribuições a priori foram tomadas como não informativas. Se as distribuições a priori, para estes parâmetros fossem tomadas como informativas, menores desvios padrões para os parâmetros seriam obtidos. Quanto às estimativas pontuais (médias a posteriori), tem sido verificado que os valores obtidos são similares quando se consideram prioris muito informativas e não informativas (Van Tassell & Van Vleck, 1996).

É importante relatar que o uso da abordagem Bayesiana em genética quantitativa não se restringe aos modelos lineares Gaussianos mas, engloba também os modelos lineares generalizados (Tempelman, 1998), os modelos não lineares e os modelos robustos (Gianola, 2000).

4 Conclusões

- A análise Bayesiana propiciou resultados adicionais àqueles obtidos pela abordagem freqüentista, destacando-se os intervalos de confiança Bayesianos para as estimativas de componentes de variância, valores genéticos e ganhos genéticos.
- As estimativas dos componentes de variância (ou funções destes), valores genéticos e ganhos genéticos pelo procedimento da amostragem de Gibbs foram mais precisas do que pelo procedimento REML/BLUP.

- A análise Bayesiana é uma técnica elegante e flexível que permite a simultânea estimação dos componentes de variância, efeitos fixos e valores genéticos de maneira precisa, mesmo para amostras de tamanho finito.

RESENDE, M. D. V. de; DUDA, L. L.; GUIMARÃES, P. R. B.; FERNANDES, J. S. C. Mixed linear models analysis using Bayesian inference. *Rev. Mat. Estat*. (São Paulo), v.19, p.41-70, 2001.

▪ *ABSTRACT: The objectives of this paper were to present theoretical and applied aspects of Bayesian parameter estimation and to apply this procedure to experimental data aiming to compare it with the REML/BLUP procedure. The results showed that: Bayesian analysis provided additional results in relation to frequentist approach, like Bayesian confidence ranges for the parameter estimates, the parameter estimates were more precise by the Gibbs sampling procedure than by the REML/BLUP, Bayesian analysis is a suitable technique for precise estimation and finite-sample inference*.

▪ *KEYWORDS: Variance components; random variables; stochastic simulation; conditional probabilities; Gibbs sampler.*

Referências bibliográficas

BIBBY, J., TOUTENBURG, H. *Prediction and improved estimation in linear models.* New York: John Wiley , 1977. p.188.

BOLDMAN, K. G. et al. *A manual for use of MTDFREML:* a set of programs to obtain estimates of variances and covariances. Washington: ARS-USDA, 1995. 120p.

BOX, G. E. P., TIAO, G. C. *Bayesian Inference in Statistical Analysis*. Reading: Addison – Wesley, 1973. p.588.

DEMPFLE, L. Relation entre BLUP (Best linear unbiased prediction) et estimateurs bayesiens. *Ann. Genet. Sel. Anim*., v.9, p.27-32, 1977.

GAMERMAN, D. *Markov chain monte carlo: stochastic simulation for bayesian inference*. Boca Raton: CRC Press, 1997. p.230.

GAMERMAN, D. *Simulação estocástica via cadeias de Markov*. Caxambú: Associação Brasileira de Estatística, 1996. 196p.

GAMERMAN, D., MIGON, H. S. *Inferência estatística:* uma abordagem integrada*.* Rio de Janeiro: Instituto de Matemática, 1993. 207p.

GIANOLA, D. Statistics in animal breeding. *J. Am. Stat. Assoc.*, v.95, p.296-9, 2000.

GIANOLA, D., FERNANDO, R. L. Bayesian methods in animal breeding theory. *J. Anim. Sci.*, v.63, p.217-44, 1986.

GIANOLA, D., FOULLEY, J. L. Variance estimation from integrated likelihood (VEIL). *Genet. Sel. Evol.*, v.22, p.403-17, 1990.

GIANOLA, D., IM, S., MACEDO, F. W. A framework for prediction of breeding value. In: GIANOLA, D., HAMMOND, K. (Ed.). *Advances in statistical methods for genetic improvement of Livestock.* Berlin: Springer Verlag, 1990. p.210-38.

HARVILLE, D. A. Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Stat. Assoc.*, v.72, p.320-8, 1977.

HENDERSON, C. R. *Aplications of linear models in animal breeding*. Guelph: University of Guelph, 1984. 462p.

HENDERSON, C. R. Sire evaluation and genetic trends. In: ANIMAL BREEDING AND GENETICS SYMPOSIUM, 3., 1973. Champaign: American Society of Annimal Science, 1973. p.10-41.

HENDERSON, C. R. et al. The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, v.15, p.192, 1959.

MURTEIRA, B. J. F. *Estatística: inferência e decisão*. Lisboa: Imprensa Nacional – Casa da Moeda, 1988. p.380.

MURTEIRA, B. J. F. *Probabilidade e estatística*: inferência estatística. Lisboa: Mc Graw-Hill, 1990. v.2. p.480.

PATTERSON, H. D., THOMPSON, R. Recovery of inter-block information when block sizes are unequal. *Biometrika*, v.58, p.545-54, 1971.

RAO, P. S. *Variance Components: Mixed models methodologies and applications*. Boca Raton: CRC Press, 1997. 350p.

RESENDE, M. D. V.de, ROSA-PEREZ, J. R. H. *Genética quantitativa e estatística no melhoramento animal*. Curitiba: Imprensa Universitária, Universidade Federal do Paraná, 1999. 494p.

ROBERTSON, A. Prediction equations in quantitative genetics. *Biometrics,* v.11, p.95-8, 1955.

RONNINGEN, K. Some properties of the selection index derived by "Henderson's mixed model method". *Z. Tierz. Zuchtungsbiol.*, v.88, p.186, 1971.

SEARLE, S. R., CASELLA, G., Mc CULLOCH, C. E. *Variance components*. New York: J. Wiley, 1992. 528p.

SORENSEN, D. A. *Gibbs sampling in quantitative genetics*. Koppenhagen: Danish Institute of Animal Science, 1996. p.186 (Intern Report, n.82).

SUN, L. et al. Bayesian methods for variance component models. *J. Am. Stat. Assoc.*, v.91, n.434, p.743-52, 1996.

TEMPELMAN, R. J. Generalized linear mixed models in dairy cattle breeding. *J. Dairy Sci.*, v.81, p.1428-44, 1998.

VAN TASSELL, C. P., VAN VLECK, L. D. A manual for use of MTGSAM: A set of FORTRAN programs to apply Gibbs sampling to animal models for variance component estimation. Lincoln: USDA, ARS, 1995. p.82.

VAN TASSELL, C. P., VAN VLECK, L. D. Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based covariance component inference. *J. Anim. Sci.*, v.74, p.2586-97, 1996.

WANG, C. S., RUTLEDGE, J. J., GIANOLA, D. Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. *Genet. Sel. Evol.,* v.26, p.91-115, 1994.

Recebido em 14.4.2000