

Avaliação de Métodos de Estimação de Componentes de Variância Utilizando Dados Simulados

José Marques Carneiro Júnior¹, Ricardo Frederico Euclides², Paulo Sávio Lopes³,
Robledo de Almeida Torres⁴

RESUMO - Estudos de simulação foram conduzidos com o objetivo de realizar análise comparativa entre os métodos de estimação de componentes de variância: método da máxima verossimilhança restrita (REML), método da máxima verossimilhança (ML) e método III de Henderson. Todos os três métodos foram utilizados com modelo reprodutor, e o método REML foi utilizado também com modelo animal. Objetivou-se verificar, principalmente, o efeito do desbalanceamento na estimação dos componentes de variância. Foi simulado um genoma, constituído de uma única característica quantitativa governada por 200 locos. O genoma foi utilizado na construção de três populações-base com herdabilidades de 0,60, 0,30 e 0,10, constituídas de 1000 animais (500 machos e 500 fêmeas). A partir de cada população-base, foram produzidas as populações iniciais. Cada população inicial foi submetida a acasalamento ao acaso, que deu origem a quatro populações para cada herdabilidade estudada. Para o estudo comparativo entre os métodos, foram introduzidos nas populações diferentes tipos e níveis de desbalanceamento. Para população com alta herdabilidade e dados totalmente balanceados, os três métodos, empregados como modelos reprodutores, apresentaram resultados semelhantes entre si, e para populações com baixa herdabilidade, o método REML, utilizado como modelo animal, mostrou resultados mais próximos do valor verdadeiro para a variância genética aditiva. Os desbalanceamentos provocados não afetaram a estimação dos componentes de variância pelos métodos; as diferenças observadas foram conseqüências dos níveis de herdabilidade. O método REML, como modelo animal, pode ser considerado o mais apropriado para estimar componentes de variância para características de baixa herdabilidade.

Palavras-chave: componentes de variância, estimação – métodos, melhoramento animal, simulação

Evaluation of the Variance Component Estimation Methods Using Simulate Data

ABSTRACT - Studies on simulation were carried out aiming to achieve a comparative analysis between the following variance component estimation methods: Restrict Maximum Likelihood (REML), Maximum Likelihood (ML) and Method III of Henderson. All of them were used as a reproducer model, and the REML method was also used as a animal model. The main objective was to verify the effect of the unbalanced data on the estimation of the variance components. A genomic, made of a single quantitative characteristic, ruled by 200 loci was simulated. The genomic was used in the construction of three base-populations with heritability of 0.60, 0.30 and 0.10, constituted by 1000 animals (500 males and 500 females). From each base-population, it was constructed the initial populations. Each initial population was submitted to the mating at random, that producing four populations for each heritabilities studied. For a comparative study of the methods, different kinds and levels of unbalanced were introduced in the populations. For the population with high heritability and fully balanced data, the three methods, used as a reproducer models, presented similar results among themselves. For populations with low heritability, the REML method, used as an animal model, presented results closer to the real value for the additive genetic variance. The induced unbalanced data did not affect the estimation of the variance components by the methods. The differences observed were a consequence of the heritability levels. The REML method, as an animal model, could be considered the most appropriate to estimate the variance components for traits of low heritability.

Key Words: variance components, methods – estimation, animal breeding, simulation

Introdução

Segundo Henderson (1975), um dos problemas mais complicados no melhoramento animal é a comparação de métodos alternativos como, por exemplo, métodos de avaliação genética. O problema

fundamental é que os valores genéticos verdadeiros são desconhecidos, conseqüentemente, predição ou métodos alternativos não podem ser comparados com aqueles valores verdadeiros para verificar qual método é mais adequado. Conforme o autor, existem três maneiras de se avaliar a eficiência de

¹ Mestre em Genética e Melhoramento - UFV, 36571-000, Viçosa, MG (zemarques28@yahoo.com.br).

² Professor do DZO/UFV, Departamento de Zootecnia, 36571-000, Viçosa, MG (rbaja@ufv.br).

³ Professor do DZO/UFV, Departamento de Zootecnia, 36571-000, Viçosa, MG (plopess@ufv.br).

⁴ Professor do DZO/UFV, Departamento de Zootecnia, 36571-000, Viçosa, MG (rtorres@ufv.br).

um método: por meio de análise algébrica, de simulação e de aplicação a um conjunto de dados.

Análises algébricas determinam as propriedades dos métodos em competição, precisamente em termos dos parâmetros assumidos do modelo e do delineamento do experimento ou esquema de amostragem. As propriedades investigadas poderiam ser aquelas mais desejáveis e possíveis de calcular, como, por exemplo, minimização do erro de predição. Entretanto, análises algébricas freqüentemente envolvem manipulações matemáticas extremamente difíceis.

A simulação computacional tem se tornado uma ferramenta muito usada nos estudos recentes de melhoramento animal, especialmente na avaliação de metodologias utilizadas na estimação de componentes de variância e na predição de valores genéticos. Apesar de ser computacionalmente oneroso, apresenta a vantagem de ser possível para situações complexas em que a aplicação de outro método de avaliação seria inviável.

Basicamente, existem dois tipos de simulação, sendo um aquele que utiliza desvios normais, ou seja, simula vetores de números com médias e variâncias. Outro tipo de simulação é aquele que utiliza estrutura semelhante à do material genético, em que a determinação de uma característica é feita pela contribuição de uma série de genes, os quais podem apresentar várias formas alélicas e alterar a expressão genotípica, de acordo com um código de expressão genética determinado pelo pesquisador, em que podem ocorrer, à semelhança do processo biológico, dominância, epistasia, pleiotropia e ligação gênica. A grande vantagem da simulação em nível de genes é que se obtêm os verdadeiros valores dos parâmetros, permitindo análise comparativa mais adequada entre os métodos escolhidos.

Uma terceira forma de comparação de metodologias mais largamente utilizada é o uso de diferentes métodos em um mesmo conjunto de dados. No entanto, raramente produzem o mesmo resultado. Quando produzem resultados diferentes, não existem informações suficientes para escolha do melhor método, uma vez que esta metodologia apenas informa o grau de similaridade entre os métodos em comparação.

O melhoramento genético animal geralmente tem como base a análise de dados coletados em vários rebanhos, em diferentes grupos de animais, em vários anos e em diferentes locais, como fazendas, granjas e centros de pesquisa. Esse fato contribui para que os

dados utilizados em programas de melhoramento animal quase sempre se apresentem desbalanceados, isto é, com número desigual de informações nas subclasses.

No caso da predição dos valores genéticos, é necessário o conhecimento das magnitudes das variâncias genéticas aditivas e fenotípicas da característica avaliada ou de sua herdabilidade. Segundo Kennedy (1981), essas variâncias freqüentemente são desconhecidas, particularmente onde novas características ou populações provenientes de programas de melhoramento são avaliadas. Um recurso é estimar componentes de variância primeiro e, então, predizer os valores genéticos de interesse num mesmo conjunto de dados.

Swallow & Monahan (1984), para modelos aleatórios, compararam os métodos ANOVA, ML, REML e método de estimação quadrática não-viesada de norma mínima (MIVQUE), usando como critérios de comparação o quadrado médio do erro e o viés. Verificaram que, para dados não muito desbalanceados, o método ANOVA deveria ser preferido, por sua facilidade de aplicação, e não encontraram grandes diferenças entre ML e REML.

Smith et al. (1991) compararam quatro métodos de estimação de componentes de variância: o método III de Henderson, o método ML, o método REML e o método MIVQUE. Concluíram que os procedimentos de verossimilhança (ML e REML) apresentam maior eficiência, em comparação com os métodos III de Henderson e MIVQUE, para estimar componentes de variância e parâmetros genéticos para dados desbalanceados.

Beaumont (1991), usando dados simulados para características reprodutivas em aves, afirmou que o método REML é mais eficiente que o método I de Henderson nas seguintes situações: quando ocorre seleção seqüencial, quando muitas características são consideradas simultaneamente e quando a herdabilidade é baixa.

Gianola e Fernando (1986), trabalhando com métodos bayesianos, apresentaram razões consistentes para que o método REML seja considerado superior ao método ML.

O objetivo desse trabalho foi realizar uma análise comparativa entre os métodos REML, o método ML e método III de Henderson, na estimativa das variâncias genéticas aditivas e residuais em dados com diferentes estruturas de desbalanceamento e herdabilidade.

Material e Métodos

Os dados utilizados neste trabalho foram simulados com o sistema computacional GENESYS para ambiente Windows (Euclides, 1996). O sistema é constituído de 11 programas, escrito para o compilador FORTRAN, e permite a simulação de genomas de certa complexidade para estudos de métodos de seleção, testes de pressuposições, avaliações de novas metodologias de seleção, entre outros, dispensando, portanto, animais e laboratórios.

Simulação do genoma

Por meio do programa Genesys, foi simulado um genoma considerando uma única característica quantitativa governada por 200 locos. Este genoma foi utilizado para formação de três populações-base, nas quais a característica quantitativa possuía herdabilidade inicial de 0,10, 0,30 e 0,60. Estas herdabilidades foram simuladas pela modificação das variâncias residuais, mantendo-se os valores das variâncias genéticas aditivas próximos entre si. Além do efeito fixo de sexo, foi introduzido um efeito de local, com 10 subclasses. O genoma simulado para o presente trabalho apresentou as seguintes propriedades:

- Quinze pares de cromossomos de tamanhos aleatórios.
- Dois mil centimorgans de comprimento.
- Os efeitos aditivos dos locos quantitativos foram simulados segundo a distribuição normal com média zero e variância .
- Os locos quantitativos foram dialélicos e não possuíam desvios de dominância nem interações entre si.
- O genoma não possuía cromossomo sexual e as frequências alélicas iniciais eram iguais em ambos os sexos.
- A distribuição estatística usada para a simulação das frequências iniciais dos locos quantitativos nas populações foi a uniforme, com média de 0,50.
- Os efeitos fixos foram sexo e local, sendo o último com 10 níveis ou classes.
- Os efeitos de ambiente não-controláveis foram simulados por meio da distribuição normal.

Simulação das populações

Simulação das populações-base

Foram simuladas três populações-base, constituídas de 1.000 animais (500 machos e 500 fêmeas), segundo a estrutura genômica proposta. Na obtenção dos

animais das populações-base, foram simulados os gametas (espermatozoides e óvulos), para formação dos zigotos. Como os gametas foram produzidos ao acaso, e não como resultado de segregações mendelianas, não houve parentesco entre os indivíduos da população-base. O sexo dos indivíduos foi definido pela simulação de um número aleatório, uniformemente distribuído e dentro do intervalo de 0 a 1, sendo declarado como macho se o número aleatório estivesse no intervalo de 0 a 0,5 e fêmea, para valores simulados entre 0,5 e 1.

Simulação das populações iniciais

A partir de cada população-base foram escolhidos aleatoriamente 20 machos e 400 fêmeas para serem os genitores das três populações iniciais. Cada população inicial foi submetida ao acasalamento ao acaso, introduzindo-se diferentes tipos de desbalanceamento, produzindo assim quatro populações para cada herdabilidade estudada. Escolhidos os pais, iniciava-se o processo de gametogênese e, posteriormente, a zigotogênese, para formação das novas gerações.

Cálculo dos componentes de variância

O componente de variância genética aditiva para a característica estudada foi calculado com base nos valores genéticos, segundo a fórmula:

$$\hat{S}_a^2 = \frac{\sum_{j=1}^n a_j^2 - \left(\sum_{j=1}^n a_j \right)^2 / N}{N}$$

em que \hat{S}_a^2 representa o estimador da variância genética aditiva para a característica; a_j , o valor genético do indivíduo j ; e N , o número total de indivíduos na população.

Desse modo, é possível obter os componentes de variância verdadeiros.

Desbalanceamento nos dados

Para o estudo comparativo entre os métodos III de Henderson, ML e REML, foram introduzidos nas populações diferentes tipos e níveis de desbalanceamento: a) desbalanceamento no número de descendentes por fêmea; b) desbalanceamento no número de fêmeas por macho; e c) desbalanceamento no efeito fixo de local.

Desbalanceamento

Para obtenção dos diferentes tipos e níveis de desbalanceamento, utilizou-se o programa Genesys,

capaz de simular desbalanceamentos no efeito fixo de local, no número de fêmeas por macho e no número de descendentes por fêmea. De cada três populações iniciais, simuladas para as herdabilidades de 0,10, 0,30 e 0,60, respectivamente, produziram-se quatro cópias, sendo uma utilizada para controle; nas demais foram introduzidos os desbalanceamentos. As populações possuíam as seguintes estruturas: na população usada como referência não foi introduzido nenhum desbalanceamento. As outras três populações possuíam diferentes tipos de desbalanceamento: primeira população - desbalanceamento apenas no número de filhos por fêmea, este número apresentou variação de 1 até 10; segunda população - desbalanceamento no número de fêmeas por macho, este número variou de 1 até 30, mais o desbalanceamento anterior; terceira população - desbalanceamento no efeito fixo de local, mais os outros dois desbalanceamentos anteriores. A cada geração, foram escolhidos ao acaso 20 machos e determinado número de fêmeas, que foram acasalados ao acaso. Cada macho foi acasalado com um número de fêmeas, que variou de 1 até 30, e de cada acasalamento foram produzidos de 1 a 10 descendentes. Esse processo foi conduzido durante 10 gerações consecutivas, com uma repetição apenas.

Identificação dos desbalanceamentos

EEE - totalmente balanceada; EED - desbalanceamento no número de descendentes por fêmeas; EDD - desbalanceamento no número de descendentes por fêmeas e número de fêmeas por macho; DDD - desbalanceamento no efeito fixo de local, no número de descendentes por fêmea e no número de fêmeas por macho. Para indicar herdabilidade alta (0,60), foi utilizada a letra "A"; para média herdabilidade (0,30), a letra "M"; e para herdabilidade baixa (0,10), a letra "B".

Tabela 1 - Identificação das populações, herdabilidades e níveis de desbalanceamento nas populações a serem analisadas

Table 1 - Identification of the populations, heritabilities and unbalanced levels in the analyzed populations

Estrutura Structure	EEE	EED	EDD	DDD
h ² = 0,10	BEEE	BEED	BEDD	BDDD
h ² = 0,30	MEEE	MEED	MEDD	MDDD
h ² = 0,60	AEEE	AEED	AEDD	ADDD

Análise dos dados

Os dados com os diferentes tipos e níveis de desbalanceamento foram submetidos à análise para estimação de componentes de variância, por meio de três métodos de estimação: método III de Henderson (Henderson, 1953); método da máxima verossimilhança - ML (Hartley & Rao, 1967); e método da máxima verossimilhança restrita - REML (Patterson & Thompson, 1971). Os três métodos forneceram estimativas de componentes de variância por meio do modelo reprodutor. O método REML também foi comparado utilizando-se o modelo animal. Empregou-se a sigla REML (A) para designar o método REML com modelo animal e REML (R) para o método REML com modelo reprodutor.

Modelo animal

A estimação dos componentes de variância pelo método REML (A) foi realizada com base no seguinte modelo animal:

$$y = Xb + Za + e$$

em que: y = vetor das observações de características medidas nos indivíduos; X = matriz de incidência dos efeitos fixos de local e sexo; b = vetor de efeitos fixos; Z = matriz de incidência de efeitos aleatórios dos animais; a = vetor de efeitos aleatórios dos animais; e e = vetor de resíduos.

As pressuposições acerca do modelo são:

$$\begin{bmatrix} y \\ a \\ e \end{bmatrix} \sim N \left\{ \begin{bmatrix} Xb \\ 0 \\ 0 \end{bmatrix}; \begin{bmatrix} V & ZG & R \\ GZ' & G & f \\ R & f & R \end{bmatrix} \right\}$$

em que: $V = ZGZ' + R$; $G = A\mathbf{S}_a^2$, sendo A a matriz de parentesco; $R = I\mathbf{S}_e^2$, sendo I a matriz identidade de ordem igual ao número de observações.

Modelo reprodutor

A estimação dos componentes de variância para os métodos REML (R), ML e método III de Henderson foi realizada com base no seguinte modelo reprodutor:

$$y = Xb + Zr + e$$

em que: y = vetor das observações de características medidas nos indivíduos; X = matriz de incidência dos efeitos fixos de local e sexo; b = vetor de efeitos fixos; Z = matriz de incidência de efeitos aleatórios de reprodutor; r = vetor de efeitos aleatórios de

reprodutor; e \tilde{e} = vetor de resíduos.

Quando usado no método REML as pressuposições acerca do modelo de reprodutor foram:

$$\begin{bmatrix} \tilde{y} \\ \tilde{r} \\ \tilde{e} \end{bmatrix} \sim N \left\{ \begin{bmatrix} X \tilde{b} \\ \underline{0} \\ \underline{0} \end{bmatrix}; \begin{bmatrix} V & ZG & R \\ GZ' & G & f \\ R & f & R \end{bmatrix} \right\}$$

em que: $V = ZGZ' + R$; $G = A \frac{1}{4} S_a^2$, sendo A a matriz de parentesco entre reprodutores.

$$R = I S_e^2$$

Quando usado nos métodos ML e III de Henderson, as pressuposições acerca do modelo de reprodutor diferiram daquelas no método REML quanto à estrutura da matriz G, que passa a ser definida como $G = I \frac{1}{4} S_a^2$, sendo I a matriz identidade de ordem igual ao número de reprodutores.

Para o cálculo dos componentes de variância pelo método REML, método ML e método III de Henderson com o modelo reprodutor foi utilizado o procedimento VARCOMP do Statistical Analysis System - SAS (LITTEL et al., 1991). O programa do SAS utilizado não permitiu a inclusão da matriz de parentescos. Para o método REML com modelo animal, foi utilizado o sistema MTDFREML (Estimação pelo Método da Máxima Verossimilhança Restrita adotando-se o procedimento Livre de Derivadas), sendo possível a incorporação da matriz de parentescos. A pressuposição para a utilização do método III de Henderson é que os dados não sejam oriundos de um processo seletivo. O método ML não considera a perda dos graus de liberdade resultante da estimação dos efeitos fixos. O método REML evita erros provenientes de pequenas amostras associadas aos efeitos fixos e erros provenientes de seleção nos dados. Todos os métodos possuem a pressuposição básica de que os dados tenham distribuição normal.

A acurácia dos três métodos foi avaliada pela comparação entre os componentes de variância preditos e os verdadeiros. Por meio da porcentagem de aproximação entre os componentes de variâncias verdadeiros e os estimados pelos métodos propostos, foi possível verificar qual método estimou melhor os componentes de variância aditiva e residual. Quanto mais próximos os valores estiverem de 100%, mais acurada foi a estimativa.

Resultados e Discussão

Na Tabela 2, são apresentados os valores dos componentes de variância genética aditiva e residual, estimados para as herdabilidades de 0,10, 0,30 e 0,60 nas populações com diferentes níveis de desbalanceamento, bem como os valores verdadeiros dos componentes de variância calculados na população-base. Os resultados observados mostram, de forma geral, que os métodos REML com modelo reprodutor - REML (R), o método ML e o método III de Henderson apresentaram resultados semelhantes entre si, mesmo para os dados totalmente desbalanceados. Entretanto, o método REML com modelo animal - REML (A) apresentou estimativas diferentes de todos os outros métodos.

No caso de herdabilidades maiores, é possível observar, nos dados verdadeiros, que o componente de variância ambiental é inferior ao componente de variância genética aditiva, indicando que, para herdabilidades maiores, parte das diferenças entre os indivíduos pode ser observada diretamente pelo fenótipo. No entanto, para herdabilidades menores, a maior parte das diferenças entre os indivíduos não pode ser verificada por meio do fenótipo. Dessa forma, aumenta-se a importância da acurácia da predição para características de baixa herdabilidade.

Com relação ao componente de variância ambiental, pode-se observar que os métodos apresentaram valores semelhantes entre si.

Nas Figuras de 1 a 6, é apresentada a eficiência dos métodos em calcular os componentes de variância genética aditiva e residual. Os resultados são expressos em porcentagem de aproximação em relação aos valores verdadeiros dos componentes de variância.

Pode-se observar que, para os níveis de desbalanceamento estudados nas herdabilidades de 0,10, 0,30 e 0,60, os métodos III de Henderson, ML e o REML (R) para estimação dos componentes de variância foram similares entre si. O método REML (A) apresentou melhores resultados para variância genética aditiva em populações com baixa herdabilidade (0,10), enquanto os demais métodos superestimaram este valor. Esse comportamento se deve ao fato de que no modelo reprodutor não é considerada toda a matriz de parentesco. Segundo Sorensen & Kennedy (1984), quando a análise é feita pelo modelo animal de Henderson & Quaas (1976) e toda a matriz de parentesco é utilizada, estimadores suficientes, como o REML, condensam as informa-

ções de todos os graus de parentesco para obter uma única estimativa dos componentes de variância da população-base.

Na Figura 1, observa-se que, para herdabilidade de 0,60 e população totalmente balanceada, todos os métodos apresentaram resultados semelhantes entre si, concordando com as observações de Corbeil & Searle (1976), citados por Seraphin (1984), os quais afirmam que o método da máxima verossimilhança restrita - REML é semelhante ao método III de Henderson para dados balanceados, e com as observações de Anderson (1979), citado por Kennedy (1981). Segundo esse autor, para dados balanceados ou levemente desbalanceados, o REML e o método III de Henderson produziram estimativas semelhantes entre si. Ainda na Figura 1, para população com característica de alta herdabilidade e desbalanceamento apenas no número de descendentes por fêmea (Aeed), o método REML (A) foi o que melhor estimou a variância genética aditiva. Entretanto, para população com estrutura Addd, ou seja, totalmente desbalanceada e com alta herdabilidade, este método foi o que mais subestimou a variância genética aditiva. Este comportamento pode ser verificado também na Figura 2, na qual se observa que o

método REML (A) apresentou os piores resultados, superestimando a variância ambiental em todos os níveis de desbalanceamento.

Para a herdabilidade de 0,30, como mostrado nas Figuras 3 e 4, os métodos ML, REML (R) e o método III de Henderson apresentaram resultados semelhantes entre si, conduzindo a estimativas mais precisas em todos os níveis de desbalanceamento. Na estimação do componente de variância genética aditiva (Figura 3), maior diferença foi verificada para a população totalmente desbalanceada, em que os três métodos apresentaram aproximação em torno de 95%, enquanto o REML com modelo animal obteve apenas 66% de aproximação em relação ao valor verdadeiro. Esses resultados foram diferentes dos encontrados por Smith et al. (1991), que compararam quatro métodos de estimação de componentes de variância - métodos III de Henderson, método da máxima verossimilhança (ML), método da máxima verossimilhança restrita (REML) e método de estimação quadrática não-viesada de norma mínima (MIVQUE), concluindo que os procedimentos de verossimilhança (ML e REML) apresentam maior eficiência, em comparação ao método III de Henderson, para estimar componentes de variância e parâmetros genéticos para dados desbalanceados.

Tabela 2 - Estimativas dos componentes de variância genética aditiva, $\hat{\sigma}_a^2$, e ambiental, $\hat{\sigma}_e^2$, pelos diferentes métodos, e valores verdadeiros para populações com diferentes níveis de desbalanceamento, com herdabilidades de 0,60, 0,30 e 0,10

Table 2 - Estimates of the additive genetic variance, $\hat{\sigma}_a^2$, and environmental, $\hat{\sigma}_e^2$, components, for the different methods, and real value of the populations with different unbalanced levels, with 0.60, 0.30 and 0.10 heritabilities

	REML(R)		REML(A)		ML		Hend.III		Real	
	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_a^2$	$\hat{\sigma}_e^2$	σ_a^2	σ_e^2
Herdabilidade = 0,60										
<i>Heritability = 0.60</i>										
Aeee	66,46	84,30	66,85	86,36	66,10	84,50	66,46	84,30	80,94	38,56
Aeed	94,91	65,79	73,63	85,29	94,41	67,64	93,18	68,63	76,56	38,51
Aedd	79,62	78,80	72,10	84,85	79,19	77,52	80,42	76,66	87,60	38,52
Addd	68,92	81,83	59,62	90,22	68,55	82,04	72,52	79,13	81,28	38,56
Herdabilidade = 0,30										
<i>Heritability = 0.30</i>										
Meee	64,81	249,31	57,99	256,82	64,43	249,59	64,81	249,31	69,50	163,36
Meed	81,36	238,16	65,50	253,11	80,90	238,51	81,55	238,02	83,93	163,16
Medd	61,04	251,27	56,85	256,32	60,69	251,53	61,52	250,90	70,24	163,19
Mddd	80,50	239,17	56,71	259,56	80,04	239,52	79,83	239,67	85,63	163,56
Herdabilidade = 0,10										
<i>Heritability = 0.10</i>										
Beee	79,18	908,61	58,39	926,46	78,61	908,55	79,18	908,61	66,45	651,14
Beed	81,54	890,42	65,35	906,10	80,88	890,43	81,04	890,76	62,74	651,76
Bedd	59,47	915,31	62,37	915,06	59,00	915,18	60,43	914,70	68,04	650,86
Bddd	85,31	894,90	68,72	911,10	84,67	894,90	83,86	895,87	65,37	651,85

A Figura 5 mostra que o método REML (A), para todos os níveis de desbalanceamento, foi o que apresentou estimativas do componente de variância genética aditiva mais próximas do valor verdadeiro, enquanto os demais métodos algumas vezes superestimaram este valor, concordando com os resultados obtidos por vários autores (Groeneveld & Kovac, 1990; Boldman et al., 1991; Keele et al., 1991, entre outros). Isso se deve principalmente ao fato de que o REML, quando utiliza o modelo animal, aproveita toda a informação que os

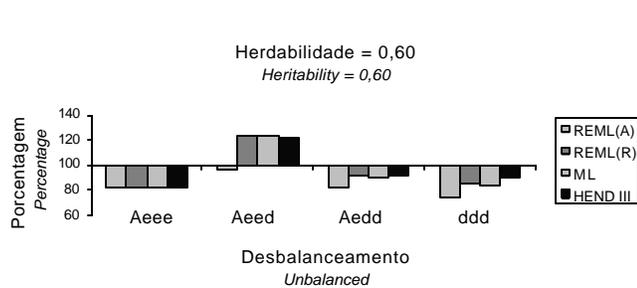


Figura 1 - Aproximação entre a variância genética aditiva estimada pelos métodos e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade de 0,60.

Figure 1 - Approach between the additive genetic variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .60.

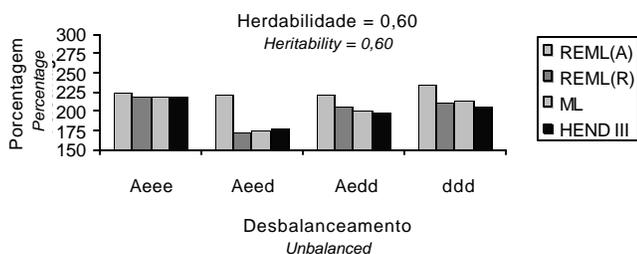


Figura 2 - Aproximação entre a variância ambiental estimada e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade de 0,60.

Figure 2 - Approach between the environmental variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .60.

dados contêm, devido ao uso da matriz de parentesco completa.

Pode-se observar que o método REML (A) apresentou maior acurácia com relação ao valor verdadeiro do componente de variância genética aditiva, para populações com herdabilidade de 0,10. Isso pode ser explicado porque, para herdabilidades menores, aumenta-se a importância da utilização da matriz de parentesco para estimação deste componente.

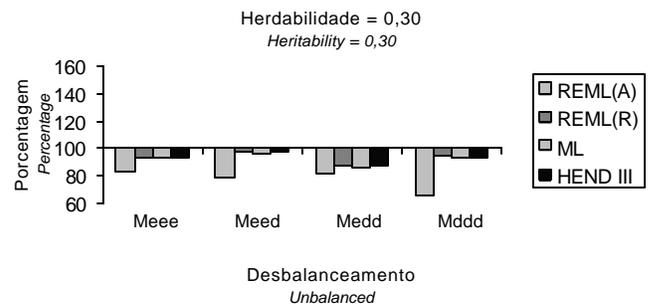


Figura 3 - Aproximação entre a variância genética aditiva estimada e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade 0,30.

Figure 3 - Approach between the additive genetic variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .30.

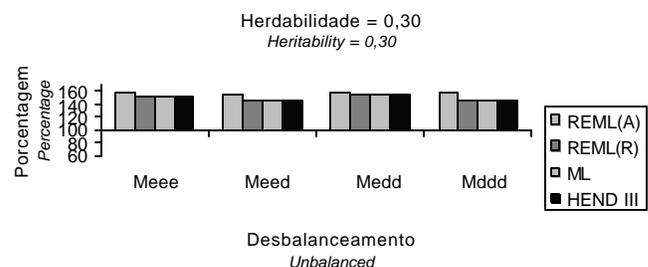


Figura 4 - Aproximação entre a variância ambiental estimada e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade de 0,30.

Figure 4 - Approach between the environmental variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .30.

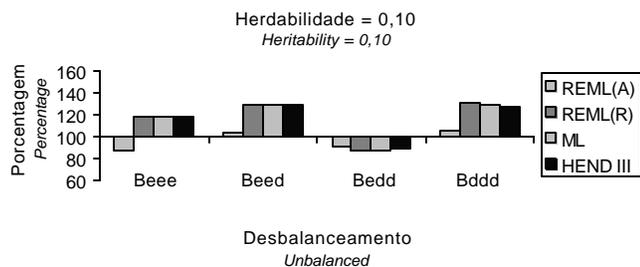


Figura 5 - Aproximação entre a variância genética aditiva estimada e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade de 0,10.

Figure 5 - Approach between the additive genetic variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .10.

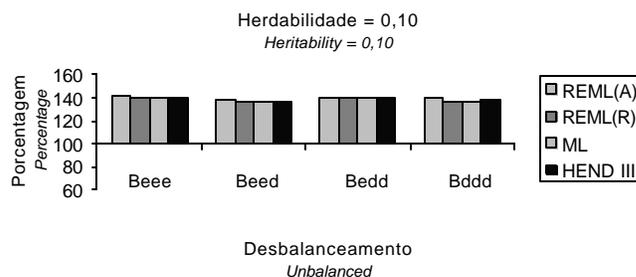


Figura 6 - Aproximação entre a variância ambiental estimada e a verdadeira, expressa em porcentagem, para as populações com diferentes níveis de desbalanceamento e herdabilidade de 0,10.

Figure 6 - Approach between the environmental variance estimated for the methods and the true values, expresses in percentage, for the populations with different unbalanced levels and heritability of .10.

No melhoramento animal, a maioria das características de interesse econômico apresenta herdabilidades baixas. Desse modo, o método REML (A) pode ser considerado o método preferido para se estimar componentes de variância para as características, no melhoramento genético, que apresentem baixa herdabilidade.

Conclusões

Para características com alta herdabilidade e população totalmente balanceada, os métodos de estimação apresentaram comportamento semelhante: aumentando-se o nível de desbalanceamento, para herdabilidade de 0,60, o método III de Henderson, o método ML e o método REML com modelo reprodutor foram mais precisos na estimação da variância genética aditiva em relação ao REML com modelo animal.

Para herdabilidade de 0,10, indiferentemente do nível de desbalanceamento, o método REML (A) apresentou estimativas do componente de variância genética aditiva mais próximas do valor verdadeiro que os demais métodos. No entanto, o desbalanceamento provocado não afetou as estimativas dos componentes de variância; as diferenças verificadas se deveram aos níveis de herdabilidades diferentes.

Recomenda-se o método REML, utilizando-se modelo animal, para estimar componentes de variância em populações com herdabilidades baixas.

Em se tratando da herdabilidade de 0,30, os

métodos ML, Henderson III e REML (R) conduziram a ganhos genéticos semelhantes entre si e superiores aos do método REML (A).

Agradecimento

À Universidade Federal de Viçosa.

Ao CNPq, pela concessão da bolsa de estudos ao primeiro autor durante seu curso de Mestrado.

Literatura Citada

- BEAUMONT, C. Comparison of Henderson's method I and restricted maximum likelihood estimation of genetic parameters of reproductive traits. **Poultry Science**, v.70, p.1462-1468, 1991.
- BOLDMAN, K.G.; Van VLECK, L.D., GREGORY, K.E. et al. Estimates of direct and maternal parameters for 200 d weight in purebred and composite lines of beef cattle. **Journal of Animal Science**, v.69 (Suppl.), p.203 (Abstract), 1991.
- EUCLYDES, R.F. **Uso do sistema para simulação Genesys na avaliação de métodos de seleção clássicos e associados a marcadores moleculares**. Viçosa, MG: Universidade Federal de Viçosa, 1996. 150p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, 1996.
- GIANOLA, D.; FERNANDO, L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v.63, p.217- 244, 1986.
- GROENEVELD, E.; KOVAC, M. A note on multiple solutions in multivariate restricted maximum likelihood covariance component estimation. **Journal of Dairy Animal**, v.73, p.2221-2229, 1990.
- HARTLEY, H.O.; RAO, J.N.K. Maximum likelihood estimation for the mixed analysis of variance model. **Biometrika**, v.54, n.1/2, p.93-108, 1967.

- HENDERSON, C.R.; QUAAS, R.L. Multiple trait evaluation using relative records. **Journal of Dairy Science**, v.43, p.1188, 1976.
- HENDERSON, C.R. Comparison of alternative sire evaluation methods. **Journal of Animal Science**, v.41, p.760-770, 1975.
- HENDERSON, C.R. Estimation of variance and covariance components. **Biometrics**, v.17, p.226-52, 1953.
- KEELE, J.W.; LONG, T.E.; JOHNSON, R.K. Comparison of methods of estimating variance components in pigs. **Journal of Animal Science**, v.69, p.1428-1434, 1991.
- KENNEDY, B.W. Variance component estimation and prediction of breeding values. **Canadian Journal of Genetics and Cytology**, v.23, n.4, p.565-578, 1981.
- KENNEDY, B.W. Variance component estimation and prediction of breeding values. **Canadian Journal of Genetics and Cytology**, v.23, n.4, p.565-578, 1981.
- LITTELL, R.C.; FREUND, R.J.; SPECTOR, P.C. **SAS®System for linear models**. 3.ed. Cary: 1991. 329p.
- PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block size are unequal. **Biometrics**, v.58, p.545-554, 1971.
- SERAPHIN, J.C. **Comparação numérica de três estimadores de componentes de variância em alguns modelos genéticos – estatísticos de cruzamentos**. Piracicaba: Escola Superior de Agricultura “Luiz Queiroz”, 1984. 74p. Tese (Mestrado) - Escola Superior de Agricultura “Luiz Queiroz”, 1984.
- SERAPHIN, J.C. **Comparação numérica de três estimadores de componentes de variância em alguns modelos genéticos – estatísticos de cruzamentos**. Piracicaba: Escola Superior de Agricultura “Luiz Queiroz”, 1984. 74p. Tese (Mestrado) - Escola Superior de Agricultura “Luiz Queiroz”, 1984.
- SMITH, E.J.; SAVAGE, T.F.; HARPER, J.A. Genetic variation in the incidence of pipped eggs in turkeys selected for low and high semen ejaculate volume. **Poultry Science**, v.70, p.2219-2222, 1991.
- SMITH, E.J.; SAVAGE, T.F.; HARPER, J.A. Genetic variation in the incidence of pipped eggs in turkeys selected for low and high semen ejaculate volume. **Poultry Science**, v.70, p.2219-2222, 1991.
- SORENSEN, D.A.; KENNEDY, B.W. Estimation of genetic variances from unselected and selected populations. **Journal of Animal Science**, v.59, p.1213-1223, 1984.
- SWALLOW, W.H.; MONAHAN, J.F. Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. **Technometrica**, v.26, p.47-57, 1984.

Recebido em: 26/04/02

Aceito em: 19/08/03