



APLICAÇÃO DE MODELO MISTO NA PREDIÇÃO DE VALORES GENÉTICOS DE BOVINOS DA RAÇA SIMENTAL¹

GILMARA BRUSCHI SANTOS², HENRIQUE NUNES DE OLIVEIRA³, GUILHERME JORDÃO DE MAGAÑES ROSA⁴, LUIS FERNANDO AARÃO MARQUES⁵

¹ Parte da dissertação de mestrado do primeiro autor, bolsita da FAPESP

² Doutoranda em produção animal, Faculdade de Medicina Veterinária e Zootecnia, UNESP, Botucatu

³ Professor da Faculdade de Medicina Veterinária e Zootecnia, UNESP, Botucatu

⁴ Professor da Universidade de Michigan, EUA

⁵ Professor da Universidade Estadual do Espírito Santos

RESUMO - Foram utilizados para este trabalho 3.559 registros de peso aos 730 dias de bovinos Simental. Estes registros foram analisados por três modelos para predição dos valores genéticos. Os modelos utilizados foram: modelo gaussiano, implementado pelo método frequentista REML, modelo gaussiano implementado sob análise Bayesiana e modelo robusto aplicado sob pressuposição de distribuição normal contaminada para os dados. Os resultados encontrados permitiram inferir que o modelo robusto pode levar a um aumento na acurácia da predição e, conseqüentemente, a uma maior resposta à seleção.

PALAVRAS-CHAVE: Heterogeneidade de Variâncias, Inferência Bayesiana, Modelo Robusto, Peso

MIXED MODEL APPLIED TO PREDICT BREEDING VALUES OF SIMMENTAL CATTLE

ABSTRACT - Were used for this study 3,559 registers of 730 days weight of Simmental cattle. This data were analysed with three modelos to predict their breeding values. Models were: gaussian model, implemented to REML, gaussian modelo with Bayesian Inference and robust model with contaminated normal distribution. The results let us to suppose that robust model can led to a rise in accuracy of prediction and, a higher selection response.

KEYWORDS: Bayesian Inference, Robust Model, Variance Heterogenity, Weight

INTRODUÇÃO

Quando há heterogeneidade de variâncias nos dados utilizados para avaliação genética, os animais dos grupos mais variáveis tendem a ter valores genéticos preditos mais variáveis, ocorrendo maior probabilidade de os animais destes grupos serem selecionados. Se a heterogeneidade for de origem ambiental, a maior seleção de animais de grupos mais variáveis poderia estar associada à redução da acurácia das predições dos valores genéticos e, conseqüentemente, à redução do ganho genético (Hill, 1984 e Vison, 1987). Se as variâncias aumentam com a média de produção dos rebanhos, animais pertencentes a rebanhos de maior variabilidade, podem ser geneticamente superestimados. Dependendo da pressão de seleção, da variância fenotípica e da herdabilidade nos diferentes rebanhos, pode haver redução no ganho genético esperado (Vison, 1987). Uma alternativa viável para corrigir o problema é o uso das distribuições normal-independentes que podem ser aplicadas por meio de modelos Bayesianos. Estas distribuições, são uma mistura de distribuições normais e são bem mais flexíveis à presença de heterogeneidade de variância.

MATERIAL E MÉTODOS

Foram analisados registros do serviço de Genealogia e de Controle de Desenvolvimento Ponderal, dos arquivos da Associação Brasileira de Criadores da Raça Simental (ABCRS, contendo registros de pedigree de 29.872 animais. O arquivo referente ao peso aos 730 dias continha dados de 3.559 animais, filhos de 526 touros e 1.885 vacas, distribuídos em 574 grupos de contemporâneos, os quais foram definidos como animais de mesmo sexo, nascidos no mesmo ano-estação, criados sob igual regime alimentar, na mesma fazenda. Foi utilizado, para verificação da presença de heterogeneidade de variâncias, o teste de Levene. As análises para predição dos valores genéticos foram realizadas por três procedimentos. Primeiramente a estimação dos componentes de variância foi feita pelo método frequentista REML (restricted maximum likelihood) sendo as análises implementadas por meio do software MTDFREML (*Multiple Trait Derivative-Free Restricted Maximum Likelihood*) desenvolvido por Boldman et al.(1993). Foi utilizado um modelo animal, supondo-se distribuição gaussiana dos resíduos. Três reinícios, utilizando-se os resultados da rodada anterior

como valor inicial na rodada subsequente, foram necessários para garantir a convergência a um máximo global. Foi utilizada a sigla GML para representação deste modelo. Para o uso da abordagem bayesiana na estimação dos componentes de (co) variância sob modelo animal também com distribuição gaussiana dos resíduos foi usado o *software* MTGSAM (*Multiple Trait Gibbs Sampling in Animal Models*), desenvolvido por Van Tassel e Van Vleck (1995). Assumiu-se que as distribuições a priori eram desconhecidas (priors flat ou não informativas). Os valores para iteração na rodada inicial foram obtidos na literatura. A inspeção gráfica e o programa Gibanal (VanKaam, 1998) foram usados para determinar a convergência. Foram realizadas no total 750.000 iterações do amostrador de Gibbs. As 1.000 primeiras iterações foram descartadas para permitir que a distribuição inicial, fornecida como priori, não interferisse nos resultados; e para evitar a redundância das informações, causada pela correlação serial entre amostras geradas subsequentemente e foi tomada apenas uma amostra a cada 350 geradas. Utilizou-se para representação deste modelo a sigla BG. Para o modelo robusto foi utilizado um programa de computador específico aplicando-se aqui a distribuição normal contaminada. Este programa é uma modificação efetuada por Pereira (2001) em Fortran 77, a partir de um programa desenvolvido pelo pesquisador Daniel Sorensen no Instituto Dinamarquês de Ciência Animal, para análise Bayesiana com modelos Gaussianos. As mesmas condições usadas para o modelo gaussiano foram também adotadas neste modelo. Para o modelo robusto utilizou-se a sigla BM para representação. Para comparação entre os valores genéticos preditos pelos três procedimentos, após procedidas as análises, foram calculados os coeficientes de correlação de Pearson e Spearman entre os valores genéticos preditos por cada um dos métodos de estimação utilizados, e também a inspeção gráfica foi utilizada para comparar os resultados esperados pela seleção baseada nos dois procedimentos.

RESULTADOS

Na Tabela 1 encontram-se as correlações entre os valores genéticos preditos a partir dos três procedimentos de estimação. As correlações para a população toda entre os três modelos foram altas. Para os animais do primeiro percentil as correlações são menores, principalmente aquelas envolvendo o modelo BM. Dos 242 animais do primeiro percentil de BG existiam 146 que também estavam no primeiro percentil para BM e para GML havia apenas 154 animais que também pertenciam ao primeiro percentil em BM. A Figura 1 mostra as diferenças entre as médias dos valores genéticos preditos pelo modelo BM quando os animais são selecionados por este modelo e pelo modelo GML. Nota-se que as diferenças chegam a cerca de 12 % quando a fração selecionada é de 0,5%.

Os animais com registro de produção são, em cada iteração do amostrador de Gibbs, classificados como sendo de uma das subpopulações. A observação do número proporcional de vezes que o animal foi classificado na população de variância mais alta permitiu identificar erros na formação de grupos de contemporâneos para o peso aos 730 dias. Os animais classificados erroneamente apareciam como observações discrepantes dentro dos grupos, e nos modelos gaussianos, parte do desvio em relação aos contemporâneos era atribuída ao valor genético destes animais. Este dado empírico indica que o modelo BM pode acomodar melhor as observações discrepantes do que os modelos gaussianos. Pereira (2001) em estudo de simulação já havia observado que este modelo é bem superior ao modelo gaussiano em situações em que há heterogeneidade de variância não sistemática.

Considerando os resultados de Pereira (2001), em estudo de simulação, que mostram que quando não existe contaminação os resultados dos modelos robustos e gaussianos são semelhantes e supondo-se que não há heterogeneidade de variância genética, mas apenas de variância residual, então, a utilização do modelo robusto pode levar a maior acurácia na identificação dos valores genéticos preditos para esta característica, e levar a melhoras importantes na resposta à seleção.

CONCLUSÕES

Os resultados obtidos indicam que a utilização do modelo robusto para estimação dos componentes de variância e predição dos valores genéticos no peso aos 730 dias da raça Simental analisados no presente estudo pode levar a um aumento na acurácia da predição e, conseqüentemente, a uma maior resposta à seleção.

REFERÊNCIAS BIBLIOGRÁFICAS

- BOLDMAN, K. G.; KRIESE, L. A.; Van VLECK, L. D.; Van TASSELL, C. P.; KACHMAN, S. D. **A Manual for of MTDFREML: A set of Programs to obtain estimates of variances and covariances** (DRAFT). U.S. Department of Agriculture, Agricultural Reserch Service. 1995. 115p.
- HILL, W. G. On selection among group with heterogeneous variance. **Anim. Prod.**, v.36, p.59-68,1983.
- PEREIRA, I. G. **Estudo se simulação e aplicação de modelos lineares mistos com distribuição normal contaminada no melhoramento genético animal**. Botucatu, FMVZ/UNESP, 2001. 91p. (Tese – Doutorado em Zootecnia).
- VANKAAM, J. B. C. H. M. (1998). Disponível em :
<<http://www.student.wau.nl/~janthijs/breedingsite/eadgibanal.html>>
- VISON, W. E. Potential bias in genetic evaluation from differences in variation within herds. **J. Dairy Sci.**, v. 70, p. 2450-2455. 1987.

TABELA 1. Coeficientes de correlação de Pearson e Spearman entre os valores genéticos preditos obtidos pelos modelos GML, e pelas médias a posteriori para os modelos BG e BM, para o população toda e para o primeiro percentil^{A,B} para a característica peso aos 730 dias. Correlações para a população toda e para o primeiro percentil estão acima e abaixo da diagonal respectivamente. O primeiro percentil refere-se aos animais com maiores valores genéticos preditos pelo modelo BM

	Correlações de Pearson			Correlações de Spearman		
	GML	BG	BM	GML	BG	BM
GML	–	0,98	0,95	–	0,98	0,95
BG	0,92	–	0,94	0,86	–	0,94
BM	0,49	0,26	–	0,34	0,12	–

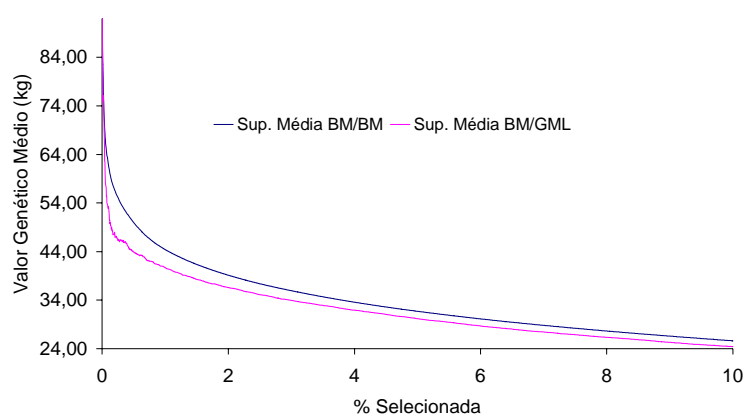


FIGURA 1. Superioridade média em valores genéticos para peso aos 730 dias preditos pelo modelo robusto, para indivíduos selecionados pelo valor genético predito pelo próprio modelo robusto (BM/BM) ou pelo modelo gaussiano (BM/GML), de acordo com a fração selecionada