



**VII Simpósio Brasileiro de Melhoramento Animal**  
**São Carlos, SP, 10 e 11 de julho de 2008**

**Estudo preliminar de análises de parâmetros genéticos em dados simulados de escores visuais com diferentes distribuições, por meio de inferência bayesiana**

Fabiana Barichello<sup>1</sup>, Maurício Mello de Alencar<sup>2</sup>, Roberto Augusto de Almeida Torres Júnior<sup>3</sup>

<sup>1</sup>Doutoranda do Programa de Pós-graduação em Genética e Melhoramento Animal – UNESP/Jaboticabal. Bolsista da FAPESP. e-mail: [fabiana\\_barichello@yahoo.com.br](mailto:fabiana_barichello@yahoo.com.br)

<sup>2</sup>Pesquisador da Embrapa Pecuária Sudeste. Bolsista do CNPq. e-mail: [mauricio@cnpqse.embrapa.br](mailto:mauricio@cnpqse.embrapa.br)

<sup>3</sup>Pesquisador da Embrapa Gado de Corte. e-mail: [rtorres@cnpqgc.embrapa.br](mailto:rtorres@cnpqgc.embrapa.br)

**Resumo** – O objetivo neste estudo foi avaliar o efeito de diferentes distribuições de frequência de dados e a utilização de modelo linear sobre os valores genéticos aditivos na avaliação de características categóricas. Foram simulados rebanhos com 40 touros e 1.200 fêmeas (1.000 vacas e 200 novilhas), acasalados aleatoriamente, acompanhados por 20 anos, e gerados efeitos aditivos direto e materno e de ambiente permanente materno. Para os produtos, foram geradas informações de grupos de manejo e o efeito da idade da vaca ao parto, os quais juntamente com os demais efeitos e um erro aleatório independente foram combinados para formar o valor fenotípico do animal na escala subjacente. Os dados na escala observada (escores visuais) foram gerados de forma a se obterem: distribuição normal relativa, distribuição normal fixa e distribuição normal relativa assimétrica. Foram geradas cinco repetições de cada distribuição, os componentes de (co)variância e os valores genéticos foram estimados por inferência bayesiana e foi obtida a correlação de Spearman entre os valores genéticos estimados e os valores genéticos verdadeiros dentro de cada categoria animal (touro, vaca e produto), para cada repetição e distribuição. Os valores de herdabilidade direta e materna estimados foram semelhantes para todas as distribuições, variando, quanto à repetição, de 0,19 a 0,26 e 0,07 a 0,12, respectivamente. As médias, das cinco repetições, das correlações de Spearman dos valores genéticos aditivos direto e materno verdadeiros com os estimados foram de 0,85 e 0,56 (touros), 0,52 e 0,45 (vacas) e 0,55 e 0,33 (produtos), respectivamente, em todas as distribuições. A distribuição relativa assimétrica se mostrou levemente inferior às outras duas formas de distribuições analisadas.

**Palavras-chave:** correlação de Spearman, distribuição, escores visuais, modelo linear

**Preliminary study of analyses of genetic parameters of visual score simulated data with different distributions, using bayesian inference**

**Abstract** – The objective of this study was to evaluate the effect of different distributions of data and the utilization of linear model on breeding value estimates for categorical traits. Herds with 40 bulls and 1,200 females (1,000 cows and 200 heifers), mated at random,

were simulated for a period of 20 years and additive direct and maternal effects and maternal permanent environmental effects were generated. For the offspring, informations on management groups and age of cow effect were generated, which were combined with the other effects and an independent error to form a phenotypic value in the underlying scale. The data in the observed scale (visual score) were generated to obtain: a “normal” relative distribution, a “normal” fixed distribution, and “normal” asymmetric relative distribution. Five replicates of each distribution were generated, the variance and the covariance components and the breeding values were estimated by bayesian inference, and the Spearman rank correlation between the estimated and the true breeding values were obtained for each animal class (sires, dams and offspring), for each repetition and distribution. The estimated direct and maternal heritabilities were similar for all distributions, ranging, across replicates, from 0.19 to 0.26 and from 0.07 to 0.12, respectively. The average, across the five replicates, for the Spearman rank correlations of the estimated and true direct and maternal breeding values were, respectively, 0.85 and 0.56 (sires), 0.52 and 0.45 (dams) and 0.55 and 0.33 (offspring), for all three distributions. The “normal” asymmetric relative distribution showed to be inferior to the other two distributions analyzed.

**Keywords:** distributions, linear models, Spearman rank correlation, visual scores

### Introdução

Nos programas de avaliação genética de bovinos de corte no Brasil, várias características são avaliadas por meio de escores (estrutura, precocidade, musculosidade, etc.), para os quais são atribuídas notas (escores) que têm natureza discreta. Os sistemas utilizados na avaliação dos animais variam quanto ao número de classes de notas, quanto ao referencial a que são dadas as notas e quanto à dispersão das notas nas várias classes, dependendo da característica e do programa de avaliação.

Além destes fatores, outra questão é a metodologia empregada para se fazer as análises. O BLUP é considerado a melhor metodologia para a estimação dos efeitos aleatórios quando a variável resposta é contínua, entretanto, vários trabalhos utilizam a metodologia BLUP na predição de valores genéticos para características categóricas.

Os objetivos neste estudo foram explorar o efeito das diferentes distribuições de frequência dos dados e a utilização de modelo linear sobre os valores genéticos aditivos diretos e maternos na avaliação de características categóricas.

### Material e Métodos

Os dados utilizados neste trabalho foram obtidos por meio de simulação, gerados a partir de um rebanho com 40 touros e 1.200 fêmeas (1.000 vacas e 200 novilhas), acasalados aleatoriamente. Das 1.200 matrizes acasaladas foram escolhidas aleatoriamente 1.000 fêmeas prenhes que permaneceram no rebanho. Os touros foram escolhidos aleatoriamente entre os machos de um ano do ano anterior e mantidos na reprodução por dois anos. A simulação contemplou 20 anos de acompanhamento do rebanho.

Foram gerados efeitos aditivos direto e materno para todos os animais e de ambiente permanente materno para as vacas. Para os produtos, foram geradas informações de grupos de manejo (cinco classes) utilizadas para a formação dos grupos de contemporâneos. Foi também gerado, para cada bezerro, o efeito da idade de sua mãe ao parto, que juntamente com os demais efeitos e um erro aleatório independente foi combinado para formar o valor fenotípico do animal na escala subjacente. Os valores para a simulação foram:

herdabilidade aditiva direta (0,25), herdabilidade materna (0,09), efeito de ambiente permanente materno (0,09), correlação dos efeitos aditivos direto e materno (-0,20) e variância de grupo de contemporâneos (0,04).

Os dados na escala observada foram gerados de forma a se obterem: distribuição normal relativa (Y1), em que as notas são atribuídas em relação ao grupo de contemporâneos, com maiores frequências nas classes centrais e menores nas extremas; distribuição normal fixa (Y2), em que as notas são atribuídas em relação a um padrão fixo; e distribuição normal relativa assimétrica (Y3), em que há tendência para maiores frequências de notas favoráveis. Para cada distribuição foram geradas cinco repetições.

Os componentes de (co)variância e os valores genéticos foram estimados por inferência bayesiana, utilizando-se o programa computacional MTGSAM (Van Tassel & Van Vleck, 1995). O modelo linear estatístico adotado foi  $y = Xb + Za + Wm + Qc + e$  em que,  $y$  = vetor das variáveis dependentes;  $b$  = vetor de efeitos fixos de grupos de contemporâneos e idade da vaca;  $a$ ,  $m$  e  $c$  = vetores de efeitos genéticos aditivos diretos e maternos e de ambiente permanente materno, respectivamente;  $e$  = vetor de erros aleatórios residuais associados às observações; e  $X$ ,  $Z$ ,  $W$  e  $Q$  são as respectivas matrizes de incidência para cada efeito. Para as análises foram computadas cadeias de 300.000 ciclos com descarte (*burn-in*) amostral dos 5.000 ciclos e a retirada amostral a cada 150 ciclos. Distribuições “flat” foram assumidas para todos os componentes de (co)variâncias. O critério de convergência foi de  $10^{-12}$ , verificada pelos testes de Geweke e Halfwidht disponibilizado no pacote estatístico *boa* (SMITH, 2007). Uma vez obtidos os valores genéticos estimados, foi calculada para cada repetição a correlação de Spearman entre os valores genéticos estimados e os verdadeiros, dentro de cada categoria animal (touro, vaca e produto).

### Resultados e Discussão

Na Tabela 1 são apresentadas as médias dos valores estimados de herdabilidade aditiva direta e materna em cada distribuição, assim como a moda, a mediana e a região de credibilidade.

Na distribuição assimétrica (Y3), a repetição três não convergiu e na repetição quatro a herdabilidade materna também não apresentou convergência, diferente das demais distribuições em que todas as repetições convergiram. Essa maior dificuldade de convergência das cadeias pode ser efeito da violação da normalidade dos dados, podendo ser resolvida com a utilização de cadeias mais longas. Abdel-Azim e Berger (1999) encontraram efeito da distribuição extrema dos dados em relação às distribuições mais próximas à normal, em que as correlações entre os valores verdadeiros e os estimados foram inferiores para a distribuição extrema.

Os valores de herdabilidade direta e materna estimados (Tabela 1) foram semelhantes para todas as distribuições, variando de 0,19 a 0,26 e 0,07 a 0,12, respectivamente, observando-se ligeira tendência de diminuição nos valores de herdabilidade direta estimados em Y3. Embora o intervalo contenha o valor verdadeiro dos parâmetros, os valores foram subestimados pelo modelo utilizado, resultado este que concorda com Hossein-Zadeh et al. (2006) que verificaram que modelos lineares em características categóricas subestimam os valores de herdabilidade verdadeiros.

As médias, das cinco repetições, das correlações de Spearman dos valores genéticos aditivos direto e materno verdadeiros com os valores estimados foram de 0,85 e 0,56 (tours), 0,52 e 0,45 (vacas) e 0,55 e 0,33 (produtos), respectivamente, em todas as

distribuições. Estes valores indicam que, apesar das herdabilidades subestimadas, a seleção para touros não seria muito prejudicada.

A pouca divergência encontrada entre as distribuições pode ser efeito do baixo valor simulado para a variância de grupos de contemporâneos (0,04).

Tabela 1- Estimativas de herdabilidade direta ( $h^2_a$ ) e materna ( $h^2_m$ ) para as distribuições (D) normal relativa (Y1), normal fixa (Y2) e relativa assimétrica (Y3).

D	R	$h^2_a$				$h^2_m$			
		Média	Moda	Med	RC (95%)	Média	Moda	Med	RC (95%)
Y1	1	0,20	0,20	0,20	0,16 – 0,24	0,12	0,12	0,12	0,09 – 0,16
	2	0,20	0,20	0,20	0,16 – 0,24	0,07	0,07	0,07	0,04 – 0,10
	3	0,21	0,21	0,21	0,17 – 0,25	0,08	0,08	0,08	0,05 – 0,11
	4	0,26	0,26	0,26	0,22 – 0,31	0,07	0,07	0,07	0,04 – 0,10
	5	0,22	0,22	0,21	0,17 – 0,25	0,08	0,09	0,08	0,05 – 0,11
Y2	1	0,20	0,20	0,20	0,17 – 0,24	0,11	0,12	0,11	0,07 – 0,14
	2	0,20	0,21	0,20	0,17 – 0,24	0,07	0,07	0,07	0,04 – 0,11
	3	0,22	0,21	0,22	0,18 – 0,25	0,09	0,09	0,09	0,06 – 0,12
	4	0,26	0,25	0,26	0,22 – 0,31	0,07	0,62	0,07	0,04 – 0,10
	5	0,21	0,20	0,21	0,17 – 0,25	0,08	0,07	0,08	0,05 – 0,11
Y3	1	0,19	0,19	0,19	0,15 – 0,22	0,12	0,12	0,12	0,08 – 0,15
	2	0,19	0,19	0,19	0,15 – 0,23	0,07	0,07	0,07	0,04 – 0,10
	4	0,26	0,25	0,26	0,21 – 0,30	-	-	-	-
	5	0,20	0,19	0,20	0,16 – 0,24	0,09	0,09	0,09	0,08 – 0,16

R = repetição; RC = região de credibilidade; Med = mediana;

### Conclusões

A distribuição relativa assimétrica se mostrou levemente inferior às outras duas formas de distribuição analisadas. O modelo linear subestimou os valores estimados para as herdabilidades, refletindo a sua baixa adequabilidade para analisar dados categóricos. A variância de grupo de contemporâneos deve ser revista para possibilitar melhor compreensão do seu efeito sobre as distribuições e os parâmetros genéticos.

### Literatura Citada

- ABDEL-AZIM, G.A.; BERGER, P.J. Properties of threshold model predictions. **Journal of Animal Science**, v.77, p.582-590, 1999.
- HOSSEIN-ZADEH, N.G.; NEJATI-JAVAREMI, A.; MIRAEI-ASHTIANI, S.R.; MEHRABANI-YEGANEH, H. Effect of the threshold nature of traits on heritability estimates obtained by linear model. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 8., 2006, Belo Horizonte. **Anais...** Belo Horizonte: Sociedade Brasileira de Melhoramento Animal, 2006. Cd room.
- SMITH, B.J. Boa: an R package for MCMC output convergence assessment and posterior inference. **Journal of Statistical Software**, v.21, n.11, p.1-37, 2007.
- VAN TASSEL, 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. USDA, ARS, 1995.