



Avaliação do modelo de regressão aleatória para as produções de leite do dia do controle usando o amostrador de Gibbs e REML

Claúdio Napolis Costa^{1,2}, Jaime Araujo Cobuci³, Ary Ferreira de Freitas^{1,2}, Marta F. M. Guimarães¹.

1 Pesquisadores da Embrapa Gado de Leite, Juiz de Fora, MG. E-mail: cnc8@cnpgl.embrapa.br

2 Bolsistas do CNPq.

3 Professor da Universidade Federal do Rio Grande do Sul - UFRGS, Porto Alegre, RS. E-mail: jaime.cobuci@ufrgs.br

Resumo - Os registros de produção de leite no dia do controle da primeira lactação de 25,5 mil vacas da raça Holandesa foram utilizados para estimar variâncias genéticas e herdabilidades para produção de leite por meio do modelo de regressão aleatória com polinômio de Legendre de ordem cinco para modelar as curvas fixas (produção média da população) e aleatórias (genética aditiva e permanente de ambiente) em análises baseadas no método da Máxima Verossimilhança Restrita (REML) e Amostrador de Gibbs (GS). As estimativas de variância genética foram similares ao longo da lactação e as herdabilidades variaram de 0,23 a 0,32 pelo GS e de 0,23 a 0,34 pelo método REML.

Palavras-chave: herdabilidade, produção de leite no dia do controle, polinômio de Legendre

Evaluation of random regression test day milk yield model using Gibbs Sampling and REML

Abstract - Data consisting of 225,357 test-day records of 25,528 first lactation Holstein cows were used to estimate variance components and heritability for test-day milk yield using random regression models by Gibbs Sampling (GS) and Restricted Maximum Likelihood (REML). Legendre polynomial of order five was used to model fixed curves for average of production of the population and random additive genetic and permanent environmental effects. Genetic variances were similar across lactation and heritability estimates ranged from 0.23 to 0.32 by GS and from 0.23 to 0.34 by REML.

Key word: heritability, test-day milk yield, Legendre polynomial

Introdução

Os procedimentos de avaliação genética para a produção de leite de bovinos da raça Holandesa no Brasil têm sido baseados na estimativa da produção acumulada no período de 305 dias, a partir da das produções nos dias dos controles leiteiros (PDC), geralmente realizados em intervalos de 30 dias. Uma alternativa de avaliação genética é a utilização de modelos que consideram a PDC, os quais permitem uma definição mais precisa dos grupos contemporâneos e dos efeitos ambientais à elas associadas (Swalve, 2000; Jensen, 2001) e, dependendo do modelo, possibilita ainda, a inclusão de informações acerca do formato da curva de lactação dos animais (Bormann et al. 2003).

Vários modelos têm sido investigados, mas não existe um consenso sobre qual deles utilizar nas avaliações genéticas (Jensen, 2001). Entre as funções paramétricas utilizadas como covariáveis para ajuste das PDC pelos modelos de regressão aleatória (MRA), o polinômio de Legendre têm sido indicado por sua ortogonalidade, propriedade matemática bastante útil na análise dos padrões de variação genética. Adicionalmente, a ordem do polinômio de Legendre utilizado na regressão fixa e aleatória desses modelos é um aspecto grande importância, visto que as estimativas dos parâmetros genéticos podem diferir com a ordem desse polinômio (Cobuci et al., 2006 e Costa et al. 2008). Neste sentido, estudos realizados por Costa et al. (2005), na raça Gir, e por Costa et al. (2008), na raça Holandesa, indicaram que o polinômio de Legendre de ordem cinco como mais apropriado para ajuste da PDC.

Embora os MRA tenham se caracterizado com uma opção preferencial para análise de dados longitudinais, as demandas computacionais das análises pelo método da Máxima Verossimilhança Restrita (REML) têm limitado a sua aplicabilidade (Miszta et al. 2000). Neste contexto, as técnicas de inferência bayesiana e, particularmente o amostrador de Gibbs (GS), têm merecido atenção do melhoramento genético devido à sua utilidade na estimação de parâmetros, ao reduzir a demanda computacional pois utilizam algoritmos que requerem apenas o armazenamento das equações dos modelos mistos, demandando assim memória proporcional ao número de animais.

O uso do GS proporciona uma alternativa significativamente simples de implementar e requer consideravelmente menos recursos computacionais do que o REML, o que facilita as análises envolvendo modelos mais complexos e grandes bases de dados (Meyer, 2002) e, ainda, resulta em estimativas de distribuições completas ao invés de apenas estimativa pontual dos parâmetros de interesse.

Objetivou-se estimar parâmetros genéticos e ambientais para PDC por meio de modelos de regressão aleatória com polinômios de Legendre de ordem cinco, usando inferência bayesiana e REML.

Material e Métodos

Foram utilizados 225 mil registros da produção de leite no dia do controle leiteiro (PLC) de 25.528 primeiras lactações de vacas Holandesas de 384 rebanhos supervisionados no período de 1993/2004 pelo Serviço de Controle Leiteiro da ABCBRH. Para tornar a base de dados mais balanceada e informativa, foram estabelecidas restrições como: a) mínimo oito registros, com intervalos de no máximo 45 dias entre controles, b) mínimo de cinco registros por rebanho-mês de controle, entendido como grupo contemporâneo, c) animais com no mínimo quatro meio-irmãs paternas e d) animais com controle leiteiro realizados entre 20 e 48 meses de idade no parto.

O modelo de regressão aleatória ajustados aos registros da PLC entre o sexto e o 305º dia de lactação incluíram o efeito de rebanho-ano-mês do controle, coeficientes do polinômio de Legendre de ordem cinco para modelar as curvas fixa (formadas por classes de idade-estação de parto da vaca) e alatórias da regressão (efeitos genético aditivo e ambiente permanente) e efeito residual.

Os componentes de variância foram estimados utilizando o amostrador de Gibbs por meio do programa GIBBS2F90 (Misztal, 2001). Inicialmente realizou-se uma amostragem previa com cadeia de 35.000 ciclos. De posse desses resultados e mediante a inspeção visual dos gráficos e do diagnóstico de convergência feitos por meio dos programas GIBBSIT (Raftery e Lewis, 1995) e GIBANAL (Van Kaam, 1998), foram definidos o tamanho da cadeia, o período de descarte e o intervalo de amostragem para obtenção das densidades marginais a posteriores dos parâmetros, definindo-se precisões de 0,001 e 0,002, com níveis de confiança de 0,95 e 0,99. De posse dos novos resultados obtidos a partir da análise de convergência, uma nova amostragem de Gibbs foi realizada com cadeia única de 105.000 ciclos.

Segundo Sorensen (1996), os estimadores das características das distribuições posteriores estão sujeitas a erro de amostragem de Monte Carlo que pode ser reduzido pela extensão da cadeia de Gibbs. Assim, utilizando-se o programa POSTGIBBSF90 (Tsuruta, 2005), definiu-se descartar as 15.000 amostras iniciais e posteriormente foram obtidas uma amostra a cada 100 ciclos, totalizando 900 amostras finais para obtenção das densidades marginais a posteriores dos parâmetros de interesse.

As estimativas dos componentes de variância dos efeitos genético, de ambiente permanente e residual obtidas pelo modelo sob análise via GS foram comparadas com aquelas obtidas por REML utilizando-se do programa REMLF90 (Misztal, 2001).

Resultados e Discussão

Estimativas de variância genética e herdabilidade para a PDC obtidas pelo modelo de regressão aleatória em análises com base no GS e método REML estão apresentadas na Figura 1. Não se observou diferença significativa entre estimativas de variâncias genéticas obtidas pelo REML ou GS, embora o modelo analisado com base no GS tenha apresentado estimativas mais estáveis ao longo da lactação, principalmente no final do período de lactação. Observa-se ainda que as estimativas de variância genética pelo REML, ao longo da maior parte da lactação, foram ligeiramente superiores às obtidas pelo GS. Ao contrário, o valor da estimativa de variância residual obtido pelo método REML (6,22 kg²) foi superior ao valor estimado pelo GS (5,73 kg²).

De posse dos valores médios das distribuições *a posteriori* dos componentes de variância dos efeitos genéticos, de ambiente permanente e da variância residual, obtidos na análise pelo GS, foram estimadas as herdabilidades para produção de leite ao longo da lactação. As estimativas seguiram a mesma tendência das variâncias genéticas para ambas metodologias, com maiores valores na fase intermediária e, menores no início e final da lactação. Estas estimativas se assemelham aos valores reportados por Melo et al. (2005), Cobuci et al. (2006) e por Costa et al. (2008), na raça Holandesa, utilizando-se do método REML.

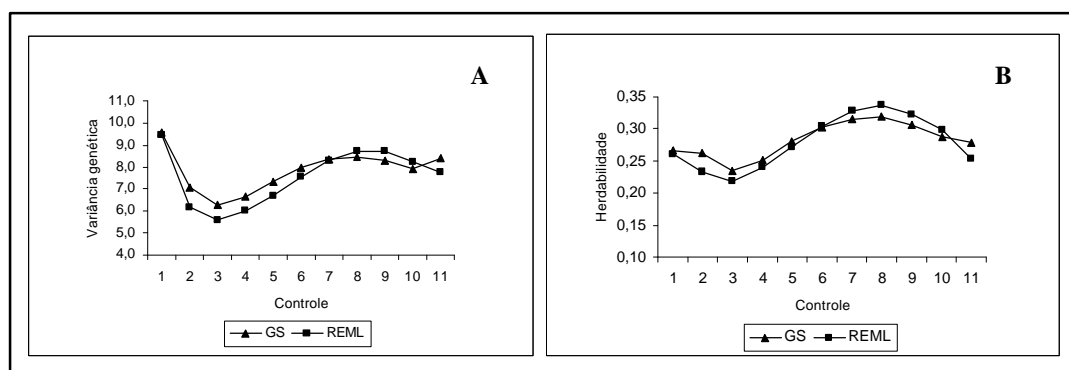


Figura 1 - Estimativa de variância genética (A) e herdabilidade (B) para produção de leite no dia do controle obtidas com o modelo de regressão aleatória através da Máxima Verossimilhança Restrita (REML) e do Amostrador de Gibbs (GS).

Conclusão

As estimativas de variância genética e de herdabilidade para as produções de leite no dia do controle obtidas pelo GS foram similares às obtidas pelo método REML.

Estudos adicionais são necessários para comparação destas metodologias com os modelos baseados na produção de leite da lactação, no que concerne a eventuais reclassificações dos valores genéticos de animais, para efetivar a sua implantação nas avaliações genéticas para as características produtivas (leite e os componentes gordura e proteína) da raça Holandesa no Brasil.

Literatura Citada

- BORMANN, J.; WIGGANS, G.R.; DRUET, T. et al. Within-herd effects of age at test day and lactation stage on test-day yields. *Journal of Dairy Science*, v.86, p.3765-3774, 2003.
- COBUCI, J.A.; COSTA, C.N.; TEIXEIRA, A.F. et al. Utilização dos polinômios de Legendre e da função de Wilkink em avaliações genéticas para persistência na lactação de animais da raça Holandesa. *Arq. Bras. Med. Vet. Zootec.*, v.58, n.4, p.614-623, 2006.
- COSTA, C.N.; MELO, C.M.R.; MACHADO, C.H.C. et al. Parâmetros genéticos para a produção de leite de controles individuais de vacas da raça Gir estimados com modelos de repetibilidade e regressão aleatória. *Revista Brasileira de Zootecnia*, v.34, n.05, p.1520-1531, 2005.
- COSTA, C.N.; MELO, C.M.R.; PACKER, I.U. et al. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. *Revista Brasileira de Zootecnia*, v.37, n.04, p.602-608, 2008.
- JENSEN, J. Genetic evaluation of dairy cattle using test-day models. *Journal of Dairy Science*, v.84, p.2803-2812, 2001.
- MELO, C.M.R.; COSTA, C.N., PACKER, U.I. et al. Parâmetros genéticos para as produções de leite no dia do controle e da primeira lactação de vacas da raça Holandesa. *Revista Brasileira de Zootecnia*, v.34, p.796-806, 2005.
- MEYER, K. RRGIBBS - A program for simple random regression analyses via Gibbs sampling. In: SEVENTH WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, Montpellier, France, 2002, Paper No. 28-27.
- MISZTAL, I. [2001], GIBBS2F90. Disponível em: <ftp://nce.ads.uga.edu/pub/ignacy/gibbs2/Gibbs2f90> Acesso em: janeiro de 2005.
- MISZTAL, I. [2001], REML90. Disponível em: <ftp://nce.ads.uga.edu/pub/ignacy/reml/remlf90> Acesso em: março de 2003.
- MISZTAL, I., T. STRABEL, E. A. MA`NTYSAARI, T. H. E. MEUWISSEN,, J. JAMROZIK. Strategies for estimating the parameters needed for different test-day models. *Journal of Dairy Science*, v.83, p.1125-1134, 2000.
- RAFTERY, A.E.; LEWIS, S.M. GIBBSIT (version 2.0). 1995. Disponível em: StatLib <ftp://lib.stat.cmu.edu/general/gibbsit> Acesso em: abril de 1993.
- SORENSEN, D. Gibbs sampling in quantitative genetics. Internal Report no. 82. Danish Institute of Animal Science. 186 p. 1996.
- SWALVE, H.H. Theoretical basis and computational methods for different test-day genetic evaluation methods. *Journal of Dairy Science*, v.83, p.1115-1124, 2000.
- TSURUTA, S. [2005], POSTGIBBSF90. Disponível em: <ftp://nce.ads.uga.edu/pub/ignacy/postgibbs/Postgibbsf90> Acesso em: maio 2005.
- VAN KAAM, J.B.C.H.M. GIBANAL: Analyzing program for Markov Chain Monte Carlo sequences (Version 2.1, July 1998). Manual. Department of Animal Sciences. Wageningen Agricultural University. 1998.