

X Simpósio Brasileiro de Melhoramento Animal

Uberaba, MG – 18 a 23 de agosto de 2013

Comparação de modelos usando diferentes cenários de censura de dados para característica reprodutiva em gado de corte¹

Diogo Anastácio Garcia², Roberto Carvalheiro³, Lucia Galvão de Albuquerque^{3,4}

¹Trabalho financiado pela FAPESP 2009/16118-5 e CNPq 559631/2009-0

²Programa de Pós-Graduação em Genética e Melhoramento Animal – UNESP, Jaboticabal. Bolsista da FAPESP - Processo 2009/15010-6. e-mail: diogo.agarcia@gmail.com

³Departamento de Zootecnia – UNESP, Jaboticabal. e-mail: rcar@fcav.unesp.br

⁴Pesquisadora do CNPq e do INCT-CA e-mail: lgalb@fcav.unesp.br

Resumo: O objetivo deste trabalho foi comparar a habilidade de predição de modelos nos quais a expressão fenotípica da característica idade ao primeiro parto (IPP) pode ser considerada como não censurada e censurada. A IPP foi analisada da seguinte forma: 1) Animais sem informação fenotípica foram retirados da análise (SC); 2) Penalização de 21 dias (PEN); 3) Simulação dos valores censurados utilizando uma distribuição normal truncada (SIM); 4) Modelo de limiar bi-linear (THR). Correlações de Pearson (CP) e a porcentagem de coincidência na classificação dos 10% melhores touros (TOP10%) foram calculadas para avaliar o efeito de cada metodologia na predição dos valores genéticos. Além disso, o quadrado médio do erro de predição (QMEP) para cada modelo foi obtido através de um esquema de validação cruzada. As estimativas de herdabilidade foram de 0,14, 0,08, 0,08 e 0,09 utilizando SC, PEN, SIM e THR, respectivamente. As CP e TOP10% entre SC e os demais modelos foram, respectivamente, altas e moderadas (0,96-0,99; 73-85%). As médias do QMEP foram de 6,851, 8,070, 7,493 e 7,491 para os modelos SC, PEN, THR e SIM, respectivamente. O modelo em que as informações censuradas são desconsideradas apresentou uma maior habilidade de predição, logo para a avaliação genética da idade ao primeiro parto nesta população este modelo seria o recomendado.

Palavras-chave: idade ao primeiro parto, reprodução, Nelore

Comparison of models using different censored data scenarios for reproductive trait in beef cattle

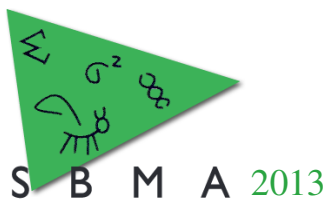
Abstract: The aim of this work was to compare prediction ability of models in which the phenotypic expression of age at the first calving (AFC) was considered to be censored and uncensored. The AFC was analyzed as following: 1) Animals without phenotypic information were excluded from the analysis (WC); 2) Penalization of 21 days (PEN); 3) The censored records were simulated using a truncated normal distribution (SIM); 4) Threshold bi-linear model (THR). Pearson correlations (PC) and the percentage of bulls in common, considering only 10% of sires with higher genetic values (TOP10%), were calculated to evaluate the effect of each methodology in breeding value prediction. Moreover, the mean square prediction error (MSPE) for each model was performed through a 10-fold cross-validation. The heritability estimates were 0.14, 0.08, 0.08, and 0.09 for WC, PEN, SIM, and THR, respectively. PC and TOP10% between WC and the other models were, respectively, high and moderate (0.96-0.99; 73-85%). The average MSPE mean were 6.851, 8.070, 7.493, and 7.491 for the models WC, PEN, THR, and SIM, respectively. The model ignoring the censored records showed the best prediction ability, therefore to perform the genetic evaluation of the age at first calving in this population the WC approach would be recommended.

Keywords: age at first calving, reproduction, Nelore

Introdução

A eficiência reprodutiva apresenta-se como um dos principais objetivos de seleção na bovinocultura de corte devido ao seu grande impacto econômico no sistema produtivo. Entre as principais vantagens do aumento da precocidade reprodutiva das novilhas estão a redução do tempo para o retorno do investimento, aumento da vida produtiva da matriz e o maior número de bezerros produzidos (Short et al., 1994). Um dos critérios de seleção mais utilizados visando o progresso genético da eficiência reprodutiva em gado de corte é a idade ao primeiro parto (IPP).

Na avaliação genética da IPP, uma parte das fêmeas, normalmente, não apresentam fenótipos devido a falhas reprodutivas, descarte voluntário e involuntário anterior à estação reprodutiva e, em



X Simpósio Brasileiro de Melhoramento Animal

Uberaba, MG – 18 a 23 de agosto de 2013

alguns casos, por registros inconsistentes. Estas observações parciais ou desconhecidas do fenótipo dos animais são denominadas censuras. De acordo com Guo et al. (2001), a inclusão adequada destas observações nas avaliações genéticas poderiam evitar distorções nas inferências e melhoria na predição dos valores genéticos.

Logo, o objetivo deste trabalho foi comparar a habilidade de predição de modelos nos quais a expressão fenotípica da característica idade ao primeiro parto foi considerada como não censurada ou censurada.

Material e Métodos

Os dados utilizados são oriundos do arquivo zootécnico de uma empresa que se dedica à bovinocultura de corte, com animais da raça Nelore criados no noroeste do estado de São Paulo e sul da Bahia. Fêmeas nascidas entre 1993 e 2007, que possuíam registros ao sobreano, foram consideradas na consistência dos dados (34.115 animais).

Os grupos de contemporâneos (GC) foram definidos a partir da combinação de ano, fazenda e grupo de manejo ao nascimento, fazenda e grupo de manejo a desmama e fazenda e grupo de manejo ao sobreano. Os GC com mais de cinco animais com informação fenotípica permaneceram no banco de dados para análise.

A IPP foi analisada utilizando a metodologia convencional de avaliação genética, onde os animais sem informação fenotípica são excluídos do banco de dados (SC). Além disso, outras três alternativas que levam em consideração a censura dos dados foram avaliadas: 1) Fêmeas sem informação foram penalizadas utilizando-se o maior valor de IPP do GC mais 21 dias (PEN); 2) Os valores censurados foram simulados a partir de uma distribuição normal truncada (SIM), sendo o maior valor de IPP observado no GC adotado como ponto de truncamento; 3) Aos animais censurados foram atribuídos o maior valor de IPP observado no GC e, além disso, uma variável auxiliar binária indicadora de censura foi definida (THR). Os componentes de variância e os valores genéticos foram estimados e preditos via modelo animal. Os programas utilizados foram os seguintes: 1) GIBBS1F90 para os modelos SC e PEN; 2) GIBBS2CEN para SIM; 3) THRGIBBS1F90 para o modelo THR.

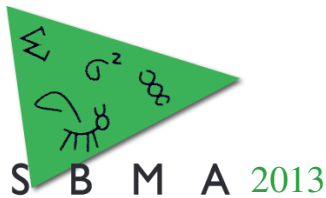
Correlações de Pearson e a porcentagem de coincidência na classificação dos 10% melhores touros foram calculadas para avaliar o efeito de cada metodologia na predição dos valores genéticos. Além disso, a habilidade de predição dos modelos foi observada através do quadrado médio do erro de predição obtido a partir de um esquema de validação cruzada com 10 repetições, em que, 80% e 20% das fêmeas com informações fenotípicas pertenciam aos arquivos de treinamento e validação, respectivamente. A alocação dos animais foi realizada aleatoriamente respeitando a restrição de pelo menos 3 animais por GC no banco de dados do treinamento.

Resultados e Discussão

As estimativas de herdabilidade para os modelos estudados foram de 0,14, 0,08, 0,08 e 0,09 para SC, PEN, SIM e THR, respectivamente. Os modelos que consideram a inclusão de informações censuradas apresentaram estimativas similares. Silva et al. (2010) reportaram resultados semelhantes, sendo estes, 0,07, 0,07 e 0,08 para PEN, SIM e THR, respectivamente. Entretanto, SC apresentou uma estimativa de herdabilidade superior aos demais. Guo et al. (2001), relataram um comportamento semelhante, aos das estimativas para IPP, trabalhando com vida produtiva (VP) e prolificidade ao longo da vida produtiva (PVP) em suínos. As estimativas de herdabilidade encontradas pelos autores para o modelo SC em relação a SIM declinaram de 0,25 a 0,18 e 0,34 a 0,16, respectivamente, para VP e PVP.

As estimativas de correlações de Pearson entre os valores genéticos preditos foram de alta magnitude nas duas situações delineadas (Tabela 1). Logo, estes resultados indicam uma forte associação entre as predições dos diferentes modelos. Guo et al. (2001), reportaram correlações entre as predições dos modelos SC-PEN e SC-SIM inferiores, variando de 0,77 a 0,86. Em gado de leite, González-Recio et al. (2006) obtiveram 0,77 de correlação entre as predições dos modelos SC-SIM.

A porcentagem de coincidência na classificação dos 10% melhores touros entre SC e os modelos que tratam a censura das observações foram inferiores às obtidas entre PEN-SIM, PEN-THR e SIM-THR (Tabela 1). Silva et al. (2010) também relataram evidências de mudança de classificação de touros, entre os modelos PEN, SIM e THR e a metodologia convencional nas avaliações genéticas de IPP.



X Simpósio Brasileiro de Melhoramento Animal

Uberaba, MG – 18 a 23 de agosto de 2013

Tabela 1. Correlações de Pearson e porcentagem de coincidência na classificação dos 10% melhores touros (TOP10%), usando médias a posteriori dos valores genéticos de touros com progênie, entre distintos modelos de avaliação genética para idade ao primeiro parto

Modelos ^a	Todos os Touros ^b	>10 filhos ^c	TOP10%
SC-PEN	0,97	0,96	73
SC-SIM	0,98	0,97	76
SC-THR	0,98	0,97	78
PEN-SIM	0,99	0,99	85
PEN-THR	0,99	0,99	85
SIM-THR	0,99	0,99	85

^aSC: somente fêmeas com fenótipo; PEN: penalização de 21 dias; SIM: utilizando uma normal truncada; THR: modelo threshold bi-linear; ^b412 animais; ^c249 animais.

As médias do quadrado médio do erro de predição dos modelos SIM e THR foram similares e menores ao do modelo PEN (Tabela 2). Entretanto, estes modelos apresentaram uma habilidade de predição inferior ao método convencional adotado nas avaliações genéticas (SC). Portanto, o modelo SC seria o mais indicado para a avaliação genética da idade ao primeiro parto nesta população.

Tabela 2. Médias dos quadrados médios dos erros de predição (QMEP), num esquema de validação cruzada de 10 repetições, entre os distintos modelos de avaliação genética para a idade ao primeiro parto

Modelos ^a	QMEP
SC	6,851
PEN	8,070
SIM	7,491
THR	7,493

^aSC: somente fêmeas com fenótipo; PEN: penalização de 21 dias; SIM: utilizando uma normal truncada; THR: modelo threshold bi-linear.

Conclusões

O modelo convencional, em que animais com informações censuradas são desconsiderados na análise, apresentou uma maior habilidade de predição em relação aos modelos que permitem a inclusão de censura na análise. Portanto, para a avaliação genética da idade ao primeiro parto nesta população o modelo convencional seria recomendado.

Agradecimentos

A Fundação de Amparo à Pesquisa do Estado de São Paulo pela concessão da bolsa de estudos, processo 2009/15010-6.

Literatura citada

- GONZÁLEZ-RECIO, O.; CHANG, Y. M.; GIANOLA, D.; WEIGEL, K. A.; Comparison of models using different censoring scenarios for days open in Spanish Holstein cows. **British Society of Animal Science**, v.82, p.223-239, 2006.
- GUO, S. F.; GIANOLA, D.; REKAYA, R.; SHORT, T. Bayesian analysis of lifetime performance and prolificacy in Landrace sows using linear models with censoring. **Livestock Production Science**, v.72, p.243-252, 2001.
- SHORT, R. E.; STAIGMILLER, R. B.; BELLOWS, R. A.; GREER, R. C. Breeding heifers at one year of age: biological and economic considerations. In: Factors Affecting Calf Crop, 1994, London. **Proceeding...** London, 1994. p.55-68.
- SILVA, L. O. C.; TSURUTA, S.; BERTRAND, J. K. et al. Including data on culled heifers on breeding values for age at first calving, Nelore cattle. In: World Congress on Genetics Applied to Livestock Production, 2010, Leipzig. **Proceeding...** Leipzig, 2010.