

ESTUDO DE PARÂMETROS DE CRESCIMENTO DE BEZERROS NELORE POR MEIO DE UM MODELO DE REGRESSÃO LINEAR: UMA ABORDAGEM BAYESIANA

NATASCHA ALMEIDA MARQUES DA SILVA,¹ LUIZ HENRIQUE DE AQUINO,² FABYANO FONSECA E SILVA³
E JOEL AUGUSTO MUNIZ⁴

-
1. Aluna do 8.º período de Zootecnia, bolsista do PIBIC/CNPq, Departamento de Ciências Exatas, UFLA – natascha@zootecnia.ufla.br.
 2. Professor do Departamento de Ciências Exatas, Universidade Federal de Lavras, UFLA – lhaquino@ufla.br
 3. Professor do Departamento de Ciências Exatas, Universidade Federal de Lavras, UFLA – fabyanobr@ufla.br
 4. Professor do Departamento de Ciências Exatas, Universidade Federal de Lavras, UFLA – jamuniz@ufla.br.

RESUMO

O estudo da fase inicial de crescimento de bovinos é de grande importância na identificação de eficiência de crescimento. Utilizou-se o método Bayesiano no ajuste do modelo de regressão linear a dados de peso–idade de novilhos Nelore nascidos nas épocas chuvosa e seca. As amostras das distribuições marginais *a posteriori* dos parâmetros do modelo foram obtidas através do algoritmo Gibbs sampler. Realizou-se um estudo de simulação de dados para

validar o programa computacional utilizado, o qual indicou a eficiência deste. A amostra da distribuição marginal *a posteriori* para a diferença entre as taxas de crescimento indicou a maior eficiência de crescimento dos animais nascidos na época seca. Em relação à variação das taxas de crescimento, verificou-se que a dos animais nascidos na época chuvosa é mais baixa do que a dos nascidos no período da seca.

PALAVRAS-CHAVE: Bezerros Nelore, método Bayesiano, regressão linear.

ABSTRACT

GROWTH STUDY OF NELORE CATTLE MALE USING A LINEAR REGRESSION MODEL: A BAYESIAN APPROACH

The initial phase of bovine growth is very important in precocity study. The Bayesian method was used to fit the linear regression model in weight-age data of Nelore cattle male born in two period: rainy and dry season. The samples of marginal posterior distributions were obtained via Gibbs sampler algorithm. A simulation study was carried on for testing the validation of the software. This study

showed efficient results. The sample of marginal posterior distribution for difference between growth rates indicated the highest growth efficiency of the animals that were born in dry period. Related to the variability of growth rate was found that animals born in rainy season show less variation than does the animals born in dry season.

KEYS WORD: Nelore cattle, linear regression, Bayesian method.

INTRODUÇÃO

O estudo das fases iniciais de crescimento de bovinos de corte é de grande importância quando

se deseja obter um crescimento rápido e eficiente dos animais. Geralmente esse estudo é realizado por meio de modelos de regressão, que, ajustados a dados de peso–idade, geram parâmetros interpretáveis

biologicamente. Desse modo, as estimativas dos parâmetros podem ser interpretadas como variáveis biológicas e utilizadas principalmente na identificação de animais com maior eficiência de crescimento.

Para identificar os animais com maior eficiência de crescimento é necessário realizar um ajuste de modelos de regressão para cada animal, o que se torna um empecilho para esse tipo de estudo, pois a análise de regressão pelo método freqüentista (quadrados mínimos) é muito sensível à flutuação dos dados individuais, e na maioria das vezes produz estimativas irreais (VARONA et al., 1997). Uma solução adequada para esse problema é ajustar modelos de regressão por meio de um método robusto à variabilidade dos dados individuais de peso-idade e que seja capaz de produzir estimativas que representem bem a precocidade dos animais.

Nos últimos anos o método Bayesiano tem sido aplicado com sucesso no ajuste de modelos de regressão a dados individuais de crescimento animal, como em VARONA et al. (1997), BLASCO et al. (2003), entre outros. Segundo tais autores, o método se caracteriza por considerar todos os parâmetros como variáveis aleatórias, utilizando conceitos probabilísticos que reduzem a dependência do ajuste em relação aos dados, o que diminui substancialmente o número de estimativas atípicas. Além disso, a estimação por intervalo é realizada de forma direta, sendo portanto mais precisa do que aquela apresentada pelo método freqüentista, que em geral utiliza variâncias assintóticas e aproximadas.

A eficiência do método Bayesiano no estudo de modelos de regressão linear foi confirmado por SIMÕES (1981). Este autor usou técnica de simulação e concluiu que as estimativas obtidas tendem mais rapidamente ao valor verdadeiro do parâmetro ao se fazerem repetições dos experimentos quando comparadas com o método freqüentista. MENDOZA (1990) também utilizou simulação de dados para estimar o quociente entre taxas e verificar o comportamento desse método em estudos de regressão linear em ensaios biológicos. Segundo o autor, a estimativa de índices biológicos tem sido bastante controvertida quando a análise é conduzida sob uma abordagem freqüentista.

WAKEFIELD et al. (1994) apresentaram uma

estratégia computacional Bayesiana bastante eficiente no estudo de modelos biométricos lineares e não-lineares de crescimento por meio do algoritmo Gibbs sampler (Amostrador de Gibbs), e concluíram que a metodologia utilizada estimou com sucesso os parâmetros dos modelos de regressão estudados.

VARONA et al. (1997) generalizaram os estudos de WAKEFIELD et al. (1994) para analisar as variáveis biológicas de qualquer função de produção (índice de crescimento, curva de persistência de lactação etc.), e descreveram o algoritmo de Gibbs sampler para implementação de tais procedimentos. Estes autores mostraram a aplicação da metodologia em uma função de crescimento linear com dois parâmetros: o intercepto e a inclinação, ajustada a dados de peso-idade de bovinos de corte da raça Pirenaica.

GELFAND et al. (1990) usaram a simulação para ilustrar a facilidade com que inferências sobre funções de parâmetros de modelos de componentes de variâncias, como o quociente entre variâncias ($\sigma_\theta/\sigma_\epsilon$) e o coeficiente de correlação ($\sigma_{\theta\epsilon}/(\sigma_\theta + \sigma_\epsilon)$), são feitas mediante o Amostrador de Gibbs.

O objetivo deste trabalho foi utilizar o método Bayesiano no ajuste de um modelo de regressão linear a dados de crescimento de machos da raça Nelore, estabelecendo-se uma estrutura de simulação para validar o programa computacional.

MATERIAL E MÉTODOS

O modelo usado para descrever a variação no peso dos novilhos e a variação entre os parâmetros foi um modelo hierárquico de três estágios (GELFAND et al., 1990; WAKEFIELD et al., 1994; GROENEWALD et al., 1996; CHANG et al., 2001; SILVA, 2003).

O primeiro estágio descreve como o peso varia com a idade do animal em determinados valores dos parâmetros. Este estágio é representado por:

$$y_{ij} = f(\boldsymbol{\omega}_i, \mathbf{t}_{ij}) + \mathbf{e}_{ij}$$

em que y_{ij} é o peso do novilho i na idade j (dias), $i = 1, 2, \dots, k$ (número de animais estudados) e

$j = 1, 2, \dots, n$ (número de medidas tomadas por animal). Além disso, ω_i é um vetor de parâmetros para cada animal, $\omega_i = [a_i, b_i]'$ e t_{ij} é a idade j em que o peso i é observado. Assume-se que os erros e_{ij} são independentes e normalmente distribuídos com média zero e variância σ_e^2 ou seja,

$$e_{ij} \sim N(0, \sigma_e^2)$$

Supõe-se ainda que a variância do erro é independente de ω_i . Logo, condicionado aos parâmetros, o peso do novilho i na idade t_{ij} se distribui segundo uma normal:

$$y_{ij} | \omega_i, \sigma_e^2 \sim N[f(\omega_i, t_{ij}), \sigma_e^2]$$

em que $f(\omega_i, t_{ij}) = \hat{w}_i$ é o peso esperado, $E(y_{ij})$, do animal i no tempo t_{ij} dado w_i . Tem-se, então,

$$f(\omega_i, t_{ij}) = E(y_{ij}) = a_i + b_i t_{ij}$$

Portanto, o modelo linear para descrever o crescimento nesse estágio é:

$$y_{ij} = a_i + b_i t_{ij} + e_{ij}$$

O parâmetro a_i é o intercepto que indica o peso ao nascer e b_i é a inclinação da reta representando a taxa de crescimento, sendo, pois, uma medida de precocidade.

Segundo GROENEWALD et al. (1996), SILVA (2003) e outros, as pesagens de cada novilho (y_{ij}), condicionado aos parâmetros, que definem o primeiro estágio da modelagem hierárquica Bayesiana relacionado em (3), podem ser expressas em notação matricial, sendo a média $= E(y_{ij}) = f(\omega_i, t_{ij}) = X\omega_i$ e variância $= \sigma_e^2 I_n$, como se vê a seguir:

$y_i | \omega_i, \sigma_e^2 \sim N(X\omega_i, \sigma_e^2 I_n)$, $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n$ em que:

y_i é o vetor das pesagens de cada animal i ;

X é a matriz ($n \times 2$) de delineamento com colunas $(1 \ t_j)$, sendo t_j os períodos de coleta de dados;

ω_i é o já mencionado vetor (2×1) de parâmetros a_i e b_i , $\omega_i = [a_i, b_i]'$ para o animal i ;

é a variância do erro;

I_n é a matriz identidade de ordem n .

O segundo estágio da hierarquia descreve a variação dos parâmetros entre os animais, sendo necessário nessa fase definir uma distribuição para

ω_i , que é o vetor de parâmetros a ser estimado. Essa distribuição é modelada por uma normal bivariada dada por:

$$\omega_i \sim N_p(\mu, \Sigma), \quad i = 1, 2, \dots, k,$$

em que:

p é o número de parâmetros do modelo de crescimento,

μ é o vetor de médias populacionais de a e b , $[\alpha, \beta]'$, Σ é a matriz de covariância de dimensão 2×2 .

O terceiro estágio da hierarquia trata da incerteza *a priori* sobre todos os parâmetros, σ_e^2, μ, Σ envolvidos nos estágios anteriores. Nessa fase especificam-se as distribuições *a priori* para σ_e^2, μ e Σ^{-1} e considera-se, como padrão, a independência entre eles. Dessa forma, segundo GROENEWALD et al. (1996), tem-se:

$$p(\mu, \Sigma^{-1}, \sigma_e^2) = p(\mu) \cdot p(\sigma_e^2) \cdot p(\Sigma^{-1})$$

em que $p(\mu) \propto$ constante, $p(\sigma_e^2) \propto 1/\sigma_e^2$ *a priori* não informativa de Jeffreys (JEFFREYS, 1961) e $p(\Sigma^{-1})$ a inversa da matriz de covariância. Esta última se distribui segundo uma Wishart com média R^{-1} e ρ graus de liberdade, ou seja,

$$\Sigma^{-1} \sim W[(\rho R)^{-1}, \rho]$$

No contexto da inferência Bayesiana, todos os parâmetros, fixos ou aleatórios, são vistos como variáveis aleatórias, conforme o conceito subjetivo de probabilidade (O'HAGAN, 1994). Neste conceito as informações *a priori* sobre os parâmetros a serem estimados são utilizadas em associação com os dados amostrais por meio da função de verossimilhança, gerando assim uma distribuição conjunta *a posteriori*, de forma que: *posteriori* \propto verossimilhança \times *priori*. Logo, pode-se escrever:

$$p(\omega | y) \propto p(y | \omega)p(\omega)$$

em que $p(\omega | y)$ e $p(\omega)$ são as distribuições *a posteriori* e *a priori* respectivamente e $p(y | \omega)$ é a distribuição de amostragem (função de verossimilhança). Substituindo-se as distribuições *a priori* relacionadas para $\mu, \sigma_e^2, \Sigma^{-1}$ e a função de verossimilhança

na expressão básica da inferência Bayesiana mostrada em (10), obtém-se a distribuição conjunta *a posteriori*. As distribuições marginais dos parâmetros, obtidas da distribuição conjunta *a posteriori*, fornecem os estimadores do método Bayesiano (BOX & TIAO, 1973).

$$p(\varpi_i | \mu, \Sigma) = (2\pi)^{-k/2} \cdot \exp[-1/2 \sum_{i=1}^k (\varpi_i - \mu)' \Sigma^{-1} (\varpi_i - \mu)]$$

$$p(\Sigma | R, \rho) \propto |\Sigma|^{-(\rho+p+1)} \cdot \exp[-1/2 \text{tr}(\Sigma^{-1} \rho R)]$$

As distribuições *a priori* relacionadas em (7) e (9) têm por expressões:

$$p(y | \varpi) \propto \frac{1}{(\sigma_e^2)^{(k/2)+1}} \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{i=1}^k [(y_i - X\varpi_i)' (y_i - X\varpi_i)]\right\}$$

A função de verossimilhança mencionada em (10) pode ser escrita como:

De (8), (11), (12) e (13), obtém-se a distribuição conjunta *a posteriori*:

$$p(\varpi_i, \mu, \sigma_e^2, \Sigma^{-1} | y_i) \propto \frac{1}{(\sigma_e^2)^{(k/2)+1}} \exp\left\{\frac{1}{2\sigma_e^2} \sum_{i=1}^k [(y_i - X\varpi_i)' (y_i - X\varpi_i)]\right\} \cdot |\Sigma|^{-k/2} \cdot \exp[1/2 \sum (\varpi_i - \mu)' \Sigma^{-1} (\varpi_i - \mu)] \cdot |\Sigma|^{-(\rho+p+1)/2} \cdot \exp[1/2 \text{tr}(\Sigma^{-1} \rho R)]$$

A integração da distribuição conjunta *a posteriori* para a obtenção das marginais geralmente não é analítica, e necessita de algoritmos especializados como o Gibbs sampler, o qual depende das distribuições condicionais completas para cada parâmetro a ser estimado. Esse algoritmo utiliza a simulação de Monte Carlo para gerar valores a partir de distribuições de probabilidades conhecidas e a teoria das Cadeias de Markov para representar a dependência entre os parâmetros, portanto o Gibbs Sampler faz uso do método conhecido como MCMC (Monte Carlo – Markov Chain) (GAMERMAN, 1996; SORENSEN, 1996).

As distribuições condicionais completas para ω, Σ^{-1}, μ e σ_e^2 estão disponíveis em WAKEFIELD et al. (1994). Esses autores consideraram R e r como hiperparâmetros e relataram que os valores assumidos por eles não influenciam diretamente os resultados, sendo utilizados apenas para manter as estimativas dentro de uma determinada região paramétrica, o que otimiza o uso do algoritmo. Os hiperparâmetros usados neste estudo foram os mesmos recomendados por eles, ou seja,

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 0,1 \end{bmatrix}, \rho = 2$$

Definido, então, $y = [y_1, y_2, \dots, y_k]$,

$$\varpi = [\varpi_1, \varpi_2, \dots, \varpi_k], \quad D = [X' X_i / \sigma_e^2 + \Sigma^{-1}],$$

$$\bar{\varpi} = \sum_{i=1}^k \varpi_i / k \quad \text{e} \quad S = \sum_{i=1}^k (\varpi_i - \mu)(\varpi_i - \mu)' \quad \text{têm-se as}$$

seguintes distribuições condicionais completas dos parâmetros, obtidas a partir da distribuição conjunta *a posteriori* (14), as quais são expressas por:

$$\varpi_i | y_i, \mu, \Sigma^{-1}, \sigma_e^2 \sim N[D^{-1}(X' y_i / \sigma_e^2 + \Sigma^{-1} \mu), D^{-1}]$$

$$\mu | y_i, \varpi, \Sigma^{-1}, \sigma_e^2 \sim N(\bar{\varpi}, 1/k \Sigma)$$

$$\Sigma^{-1} | y_i, \varpi, \mu, \sigma_e^2 \sim W[S + \rho R]^{-1}, k = \rho]$$

$$\sigma_e^2 | y_i, \varpi, \mu \sim \text{GI}\{nk/2, 1/2 \{ \sum_{i=1}^k \varpi_i' (y_i - X\varpi_i) \}'\}$$

As funções dos parâmetros de interesse prático são $b_1 - b_2$, que é a diferença nas taxas de crescimento na população, e $\sigma_{\beta_1}^2 / \sigma_{\beta_2}^2$, que é a variação nas taxas de crescimento nas duas populações.

O algoritmo Gibbs sampler foi implementado pela linguagem IML (*iterative matrix language*) disponível no *software* SAS (SAS, 1996) com 10.000 iterações. O aquecimento (*burn-in*), ou seja, o número de iterações excluídas para minimizar o efeito dos valores iniciais, foi de 4.000 e utilizou-se um espaçamento entre os pontos amostrados (*thin*) de vinte observações.

Sendo o Gibbs sampler um algoritmo iterativo, a constatação de sua convergência torna-se imprescindível para a validade do estudo. Portanto, além da análise gráfica (monitoramento informal) utilizaram-se o critério de RAFTERY & LEWIS (1992), o teste de Geweke (GEWEKE, 1992) e o fator escala de Gelman & Rubin (GELMAN & RUBIN, 1992) disponíveis no pacote computacional BOA (Bayesian Output Analysis) do *software* R.

Para avaliar os recursos computacionais utilizados, realizou-se um estudo de simulação de dados atribuindo-se valores pré-fixados para os parâmetros a e b do modelo de regressão linear

simples e para a variância do erro σ_e^2 . Dessa forma, é possível comparar as estimativas obtidas com os valores paramétricos utilizados na simulação. Foram simulados dez experimentos, os quais continham dez dados de peso-idade de dez animais, realizando-se análises individuais de cada animal dentro de cada experimento. Os valores paramétricos usados no modelo de regressão linear foram $\alpha = 30$, $\beta = 0,8$ e $\sigma_e^2 = 4$.

O método Bayesiano também foi aplicado a dados reais, os quais constam de seis pesagens, do nascimento até aproximadamente um ano de idade, período em que o crescimento do animal mostra um comportamento linear. Foram estudadas duas populações de 387 novilhos Nelore nascidos nas épocas chuvosa (outubro-março) e seca (abril-setembro), pertencentes a um mesmo rebanho localizado na região de Araguari, MG, cujos dados foram fornecidos pela Associação Brasileira de Criadores de Zebu (ABCZ) localizada em Uberaba, MG.

RESULTADOS E DISCUSSÃO

A Tabela 1 mostra o comportamento dos parâmetros a e b do modelo de regressão linear simples e da variância do erro σ_e^2 no estudo de simulação de dados.

Através da Tabela 1, é possível observar que o método Bayesiano mostrou-se efetivo para estimar os parâmetros do modelo de regressão linear simples e a variância do erro, pois os valores paramétricos utilizados na simulação encontram-se, em média, dentro do intervalo de credibilidade estimado em todos os experimentos. Resultados semelhantes foram obtidos por MENDOZA (1990), mediante a mesma técnica de simulação, o que possibilitou validar a metodologia e o programa computacional.

A Figura 1 mostra a curva de crescimento média dos animais Nelore.

Na Figura 1, é possível observar que, em média, os animais apresentam um padrão de crescimento linear até aproximadamente um ano de idade, justificando assim o ajuste do modelo de regressão linear simples. Este fato também foi observado

por VARONA et al. (1997), MAZZINI (2001) e por BERGAMASCO et al. (2001).

TABELA 1. Estimativas pontuais e intervalo de credibilidade de 95% para o parâmetro α , β e σ_e^2 nos dez experimentos simulados.

Experi- mento	Média	Desvio padrão	Limite inferior	Limite superior
Parâmetro a ($\alpha = 30$) ¹				
1	30,0653	1,5315	26,9981	33,1326
2	29,9951	1,3731	27,2705	32,7608
3	30,1902	1,6137	26,9713	33,3918
4	30,0869	1,8998	26,3108	33,8706
5	29,5379	1,4872	26,5814	32,5244
6	30,0682	1,6377	26,8034	33,3568
7	29,0966	1,5899	25,8835	32,2524
8	29,9131	1,5950	26,7253	33,1026
9	30,2424	1,6741	26,9117	33,6128
10	30,2539	1,6575	26,9676	33,5591
Parâmetro β ($\beta = 0,8$)				
1	0,7398	0,2469	0,2462	1,2325
2	0,7735	0,2212	0,3282	1,2108
3	0,7521	0,2602	0,2339	1,2709
4	0,7848	0,3062	0,1785	1,3996
5	0,8541	0,2393	0,3757	1,3293
6	0,7427	0,2634	0,2157	1,2727
7	0,9740	0,2568	0,4607	1,4917
8	0,8545	0,2570	0,3391	1,3709
9	0,8062	0,2702	0,2664	1,3468
10	0,7201	0,2668	0,1886	1,2496
Parâmetro σ_e^2 ($\sigma_e^2 = 4$)				
1	5,2214	1,6864	1,7859	6,4050
2	4,1551	0,8812	1,4138	5,3873
3	5,9312	1,5603	2,0244	5,3873
4	6,8400	2,5071	2,6812	6,7070
5	4,8692	1,4383	1,6687	6,4024
6	5,9552	2,1789	2,0226	6,4077
7	5,7316	1,9801	1,9664	6,8773
8	5,8949	1,1373	2,0181	5,3623
9	6,5261	2,5525	2,2296	6,8665
10	6,0132	1,2488	2,0550	6,7414

¹ = valores paramétricos utilizados na simulação de dados.

Os resultados obtidos, pela análise das amostras da distribuição marginal *a posteriori* dos parâmetros, das variâncias e covariâncias juntamente com os critérios de convergência utilizados, estão descritos nas Tabelas 2 e 3.

A convergência foi alcançada, pois os valores do teste de Gelman e Rubin (\sqrt{R}) foram todos iguais a 1 e os valores de z do teste de Geweke não foram superiores a 1,96 ou inferiores a -1,96 como indicado por BLASCO et al. (2003). Nota-se também que os valores de EMC foram de pequena magnitude, mostrando que o processo proporcionou estimativas confiáveis para os parâmetros. O aquecimento (*burn in*), o número total de interações e o espaço entre as amostragens (*thin*) usados foram

bem maiores que o mínimo recomendado pelo teste de Raftery e Lewis. Portanto, nenhuma irregularidade foi observada no processo de amostragem, o que assegura a qualidade dos resultados. Observa-se, pela Tabela 1, que a estação chuvosa apresenta maior peso ao nascer (α_1), enquanto que a estação seca registrou maior taxa de crescimento (β_2).

TABELA 2. Médias, desvios padrões (DP), intervalos de maior probabilidade (HPD), erro de Monte Carlo (EMC), valor de z para o teste de Geweke (Z) e fator escala de Gelman e Rubin (\sqrt{R}) obtidos das amostras das distribuições *a posteriori* dos parâmetros do modelo de crescimento linear para a população de animais nascidos nas épocas chuvosa e seca.

Parâmetro ⁽¹⁾	Média	DP	HPD		EMC	Z	\sqrt{R}
			Limite inferior	Limite superior			
α_1	37,99	0,66	37,2900	38,6300	$3,9 \times 10^{-3}$	0,77	1,00
α_2	36,99	0,36	36,2500	37,6700	3×10^{-3}	0,72	1,00
$\alpha_1 - \alpha_2$	1,00	0,03	0,9300	1,0700	3×10^{-3}	-0,07	1,00
β_1	0,42	1×10^{-3}	0,4210	0,4280	1×10^{-5}	-0,44	1,00
β_2	0,49	1×10^{-3}	0,4900	0,5000	1×10^{-5}	-0,28	1,00
$\beta_1 - \beta_2$	-0,07	2×10^{-4}	-0,0730	-0,0720	2×10^{-6}	-0,97	1,00

I = Subscritos 1 e 2 indicam épocas chuvosa e seca, respectivamente.

TABELA 3. Médias, desvios padrões (DP), intervalos de maior probabilidade (HPD), erro de Monte Carlo (EMC), valor de z para o teste de Geweke (Z) e fator escala de Gelman e Rubin (\sqrt{R}) obtidos das amostras das distribuições *a posteriori* das variâncias e covariâncias das populações referentes às épocas chuvosa e seca.

Parâmetro ⁽¹⁾	Média	DP	HPD		EMC	Z	\sqrt{R}
			Limite inferior	Limite superior			
	68,21	0,24	67,7200	68,7000	$2,4 \times 10^{-3}$	0,34	1,000
	76,45	0,28	75,9000	77,0000	2×10^{-3}	0,33	1,00
	$3,8 \times 10^{-3}$	1×10^{-5}	$3,836 \times 10^{-3}$	$3,893 \times 10^{-3}$	1×10^{-7}	-0,30	1,00
	4×10^{-3}	1×10^{-5}	0,0046	0,00470	1×10^{-5}	-0,30	1,00
	273,58	0,40	272,7800	274,3700	$3,7 \times 10^{-3}$	1,67	1,00
	326,50	0,14	325,5500	327,4500	4×10^{-3}	1,67	1,00
$\text{Cov}(\alpha_1, \beta_1)$	-0,23	1×10^{-3}	-0,2356	-0,2298	1×10^{-5}	-0,49	1,00
$\text{Cov}(\alpha_2, \beta_2)$	-0,25	1×10^{-3}	-0,2594	-0,2527	1×10^{-7}	-0,07	1,00
/	0,81	$1,1 \times 10^{-16}$	0,816414	0,816415	1×10^{-10}	$1,2 \times 10^{-2}$	1,00

Subscritos 1 e 2 indicam épocas chuvosa e seca, respectivamente

A Figura 2 indica a amostra da distribuição *a posteriori* para a diferença da taxa de crescimento entre as populações de animais nascidos nas estações chuvosa (β_1) e seca (β_2).

Nota-se, pela Tabela 1, que a população referente aos animais nascidos na época seca apresenta maior taxa de crescimento do que aqueles nascidos na chuvosa, o que pode ser explicado pela melhor condição das pastagens, quando os animais nascidos na seca atingem a idade de desmama. Isso também foi verificado por SILVA et al. (2001) e por SILVA (2003), em estudos de crescimento de gado Nelore.

Na Figura 3, é mostrado o processo de obtenção das distribuições marginais para as variâncias de β_1 e β_2 pelo Gibbs sampler e também a amostra da distribuição *a posteriori* para a razão dessas variâncias.

Observa-se pela Figura 3 que o algoritmo Gibbs sampler não demonstrou problemas de con-

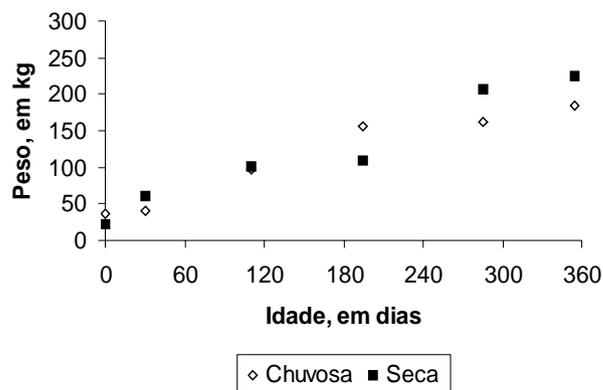


FIGURA 1. Crescimento linear dos animais até aproximadamente um ano de idade

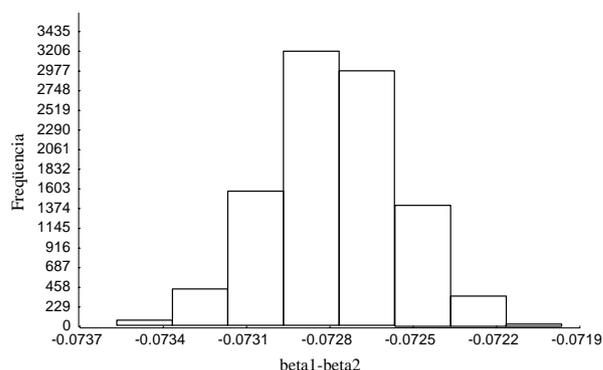


FIGURA 2. Amostra da distribuição *a posteriori* para $\beta_1 - \beta_2$

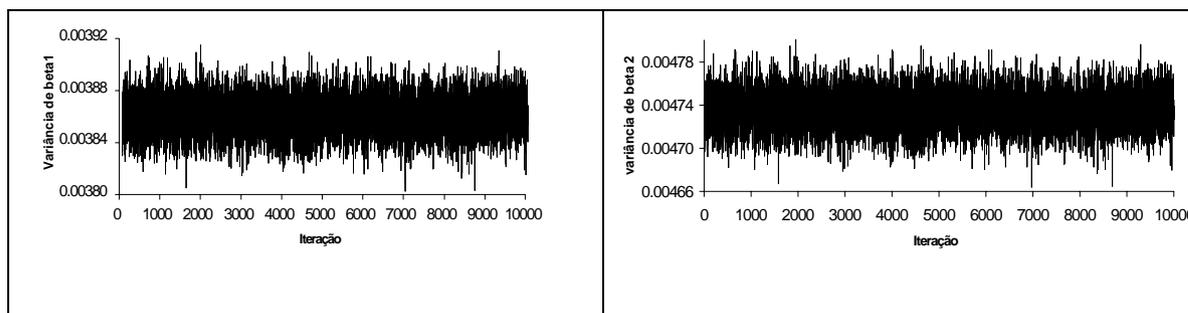


FIGURA 3. Gibbs sampler para $\sigma^2_{\beta_1}$ e $\sigma^2_{\beta_2}$

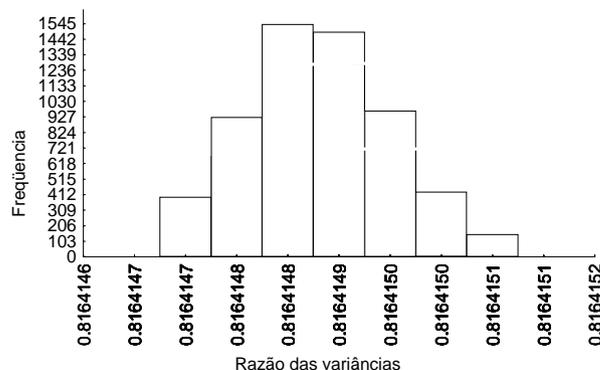


FIGURA 4. Razão das variâncias ($\sigma^2_{\beta_1} / \sigma^2_{\beta_2}$)

vergência na obtenção da amostra da distribuição marginal de $\sigma_{\beta 1}^2$ e $\sigma_{\beta 2}^2$. Verifica-se, por meio da amostra da distribuição *a posteriori* para a razão $\sigma_{\beta 1}^2 / \sigma_{\beta 2}^2$, que a população dos animais nascidos na época chuvosa apresentou-se mais variável (Tabela 3 e Figura 4) em relação à taxa média de crescimento populacional que a população de animais nascidos na época seca.

CONCLUSÃO

Mediante o estudo realizado, conclui-se que o método Bayesiano é eficiente no estudo do crescimento de novilhos Nelore do nascimento a aproximadamente 1 ano de idade. Além disso, verificou-se que o peso dos animais, ao nascer, foi maior na estação chuvosa e que a taxa de crescimento foi maior e mais variável na estação seca. Por fim, assinala-se que a simulação validou o programa computacional.

REFERÊNCIAS

- BERGAMASCO, A.F.; AQUINO, L. H. de; MUNIZ, J. A. Ajuste de modelos não-lineares a dados de crescimento de fêmeas da raça holandesa. **Ciência e Agrotecnologia**, Lavras, v. 25, n. 2, p. 235-241, 2001.
- BLASCO, A.; PILES, M.; VARONA, L. A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. **Genet. Sel. Evol.**, Valencia, v. 35, p. 21-41, 2003.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: John Wiley & Sons, 1973. 360p.
- CHANG, Y.M.; REKAYA, R.; GIANOLA, D.; THOMAS, D.L. Genetic variation of lactation curves in dairy sheep: a Bayesian analysis of Wood's function. **Livestock Production Science**, Amsterdam, v. 71, n. 2-3, p. 241-251, 2001.
- GAMERMAN, D. Simulação estocástica via cadeias de Markov. In: SIMPÓSIO NACIONAL DE **Ciência Animal Brasileira**, v. 7, n. 1, p. 57-65, jan./mar. 2006
- PROBABILIDADE ESTATÍSTICA (SINAPE), 12., 1996, Caxambu, MG. **Anais...** Caxambu, MG: ABE, 1996. 196p.
- GELFAND, A.E.; HILLS, S.; RACINE POON, A.; SMITH, A.E.M. Illustration of Bayesian inference in normal data models using Gibbs sampling. **Journal of the American Statistical Association**, Alexandria, v. 85, n. 410, p. 972-985, 1990.
- GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequences. **Statistical Science**, Hayward, v. 7, n. 4, p. 457-511, 1992.
- GEWEKE, J. Evaluating the accuracy of Sampling-based approaches to the calculation of posterior moments (with discussion). In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P.; SMITH, A. F.(Eds.). **Bayesian Statistics**. Oxford: University Press, p. 169-193, 1992. 4 p.
- GROENEWALD, P.C.N.; FERREIRA, A.V.; MERWE, H.J.; VAN DER, SLIPPERS, S.C. Application of Bayesian inference in the comparison of lactation curves of Merino ewes. **Animal Science**, London, v. 62, p. 63-69, 1996.
- JEFFREYS, H. **Theory of probability**. Oxford: Clarendon Press, 1961. 325p.
- MAZZINI, A.R.A. **Análise da curva de crescimento de machos Hereford considerando heterogeneidade de variâncias e autocorrelação dos erros**. 2001. 94 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.
- MENDOZA, M A Bayesian analysis of the slope ratio bioassay. **Biometrics**, Alexandria, v. 46, p. 1059-1069, 1990.
- O'HAGAN, A. **Kendall's advanced theory of statistics: Bayesian inference**. Edward Arnold,

Cambridge, v. 2b, 1994. 452 p.

RAFTERY, A. E.; LEWIS, S. M. How many iterations in the Gibbs Sampler? In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P.; SMITH, A.F. (Eds.). **Bayesian Statistics 4**. Oxford: University Press, 1992. p. 763-773.

SAS INSTITUTE SAS / STAT. **SAS / IML software**: usage and reference, version 6. Cary, 1996. 501p.

SILVA, F.F. **Análise Bayesiana da curva de lactação de cabras Saanen**. 2003. 64 f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.

SILVA, F.F.; AQUINO, L.H de; OLIVEIRA, J.A. Influência de fatores genéticos e ambientais sobre as estimativas dos parâmetros das funções de crescimento em gado Nelore. **Ciência e Agrotecnologia**, Lavras, v. 25, n. 5, p. 1195-1205, 2001.

SIMÕES, N.R. **Estimadores de Bayes**: aplicação ao modelo de regressão linear simples. 1981. 138 f. Dissertação (Mestrado em Estatística e Experimentação Agrônômica) – Universidade de São Paulo, Piracicaba, SP.

SORENSEN, D. **Gibbs sampling in quantitative genetic**. Copenhagen: Foulun, 1996. 186p.

VARONA, L.; MORENO, C.; GARCIA CORTÉS, L.A.; ALTARRIBA, J. Multiple trait genetic analysis of underlying biological variables of production functions. **Livestock Production Science**, Amsterdam, v. 47, p. 201-209, 1997.

WAKEFIELD, J.C.; SMITH, A.F.M.; RACINE-POON, A.; GELFAND, A.E. Bayesian analysis of linear and non-linear population models by using the Gibbs Sampler. **Journal of Applied Statistics**, Série C, Abingdon, v. 43, n. 1, p. 201-221, 1994.

Protocolado em: 10 maio 2004. Aprovado em: 30 mar. 2005.