AVALIAÇÃO DAS CURVAS DE CRESCIMENTO DE CULTIVARES DE FEIJOEIRO: UMA ABORDAGEM BAYESIANA

Sebastião Martins Filho, Fabyano Fonseca e Silva

Universidade Federal de Viçosa/Informática: Área estatística, 36570-000, Viçosa-MG
smartins@dpi.ufv.br, fabyano@dpi.ufv.br

Resumo- Foi realizada uma avaliação das curvas de crescimento das cultivares de feijoeiro Neguinho e Carioca utilizando o modelo não linear de crescimento Logístico por meio da metodologia Bayesiana. A metodologia utilizada permitiu comparar curvas de crescimento de forma eficiente, sem utilizar aproximações assintóticas, além disso, ficou demonstrado que a cultivar Carioca apresentou um incremento em altura maior que o da cultivar Neguinho.

Palavras-chave: *Phaseolus vulgaris, Inferência bayesiana, modelos não lineares, Modelo logístico.*
Área do Conhecimento: Ciências Agrárias

Introdução

Geralmente o estudo de curvas de crescimento de espécies vegetais tem sido conduzido por meio de uma abordagem freqüentista, mediante ajuste de modelos não-lineares que buscam sintetizar as informações em poucos parâmetros interpretáveis biologicamente. Esta abordagem é fundamentada em processos iterativos, como o Gauss-Newton, DUD e Marquardt, devido à não-linearidade dos parâmetros, através da minimização da soma de quadrados dos resíduos. Porém, quando se trata de ajustes individuais, ou seja, para várias unidades experimentais, de modelos matematicamente complexos, ou se dispõe de poucas observações longitudinais, os métodos iterativos muitas vezes estimam valores irreais para os parâmetros, levando à confecção de curvas de crescimento atípicas. Além disso, quando se trata de comparações de curvas provenientes de tratamentos diferentes, por exemplo, cultivares, na maioria das vezes, a distribuição dos parâmetros de modelos não-lineares não segue uma distribuição normal. Deste modo torna-se complexo o processo de formulação de testes estatísticos por meio do método freqüentista, pois não serão atendidas pressuposições relacionadas à utilização da teoria assintótica (Silva et al., 2005).

Em recentes estudos envolvendo ajuste de modelos de regressão não-linear (Blasco et al., 2003 e Silva et al., 2005) a inferência Bayesiana foi utilizada com sucesso, pois reduziu o número de estimativas viesadas, mesmo utilizando poucas informações. A estimação por intervalo apresentou-se mais precisa em relação à obtida pela metodologia freqüentista e possibilitou a comparação estatística direta e simplificada entre curvas por meio da obtenção de uma distribuição para as diferenças entre parâmetros de duas curvas diferentes.

Toda teoria da inferência Bayesiana está fundamentada no teorema de Bayes, o qual é um resultado simples de probabilidade condicional. Para a sua utilização é necessário especificar $p(\theta)$, que é conhecida como distribuição *a priori* de θ. Tal distribuição é utilizada para representar probabilisticamente o conhecimento que se tem sobre θ antes de os dados serem obtidos. Outra exigência é a especificação de uma distribuição conjunta para os dados amostrais, denominada de função de verossimilhança, $L(\theta \mid y_1,...,y_n)$, a qual representa a informação sobre θ que foi obtida dos dados.

A expressão matemática do Teorema de Bayes é:

$$P(\theta \mid Y) = \frac{L(\theta \mid Y)P(\theta)}{\int L(\theta \mid Y)P(\theta)d\theta}, \qquad (1)$$

sendo $Y = \{y_1, y_2, ..., y_n\}$.

Uma forma equivalente da expressão (1), visto que o denominador não depende de θ, é dada por:

$$P(\theta \mid Y) \propto L(\theta \mid Y)P(\theta) \qquad (2)$$

A expressão (2) é entendida como:

$$Dist.\ Posteriori \propto Verossimilhança \times Dist.\ Priori$$

A distribuição *a posteriori* de um parâmetro contém toda a informação probabilística a respeito do mesmo. Dessa forma, toda a inferência com respeito ao parâmetro é realizada por meio desta distribuição, pois é a partir de seus valores de média, mediana e moda que se têm as estimativas de interesse.

Rosa (1998) afirma que para se inferir com relação a qualquer elemento de θ, a distribuição *a posteriori* conjunta dos parâmetros, $p(\theta \mid Y)$, deve ser integrada em relação a todos os outros elementos que a constituem. Assim, se o interesse do pesquisador se concentra sobre determinado conjunto de θ, por exemplo, θ_1, tem-se a necessidade da obtenção da distribuição $p(\theta_1 | Y)$, dada por:

$$P(\boldsymbol{\theta}_1 | Y) = \int_{\theta \neq \theta_1} P(\boldsymbol{\theta} | Y) d\boldsymbol{\theta}_{\theta \neq \theta_1}$$

A integração da distribuição conjunta *a posteriori* para a obtenção das marginais geralmente não é analítica, necessitando de algoritmos iterativos especializados como o Gibbs Sampler e o Metropolis Hastings, os quais são denominados de algoritmos MCMC (Markov Chain-Monte Carlo). Uma forma prática de implementar estes algoritmos é utilizar o software Winbugs (Spiegelhalter et al., 2000)

Materiais e Métodos

O experimento foi conduzido em casa de vegetação utilizando vasos plásticos de 20kg, com solo Aluvial Eutrófico textura média da camada superficial (0,0 a 0,2m), peneirado em malhas de 4mm. Na preparação dos vasos, efetuou-se a adubação de plantio correspondente a dosagem de 250kg.ha-1 da fórmula 04-14-08 (N, P2O5 e K2O) e 30 dias após a emergência das plantas foi feita uma adubação de cobertura com nitrogênio na forma de uréia, equivalente a 30kg.ha-1. O delineamento experimental utilizado foi o inteiramente casualizado com vinte repetições no esquema de parcelas subdivididas. Nas parcelas foram colocadas as cultivares de feijão Neguinho e Carioca e nas subparcelas foram coletados os dados de altura da plantas de cinco em cinco dias até aos 85 dias após o plantio, correspondendo assim a 17 observações longitudinais.

Para descrever o crescimento utilizou-se o modelo não-linear de crescimento Logístico.

$$y_{ij} = \boldsymbol{a}_i \left[1 + \boldsymbol{b}_i \exp(\boldsymbol{k}_i t_{ij})\right]^{-1} + e_{ij}$$

em que: y_{ij} é a altura da planta avaliada na repetição i na idade j, a_i é a altura máxima da planta, b_i é o parâmetro de integração (não tem interpretação biológica) e k_i é a taxa de maturidade (quanto maior, maior a precocidade), e e_{ij} é o termo de erro aleatório, $e_{ij} \sim N(0,\sigma_e^2)$.

A metodologia Bayesiana foi empregada mediante as especificações seguintes.

A distribuição dos dados amostrais foi dada por:

$$y_{ij} \big| \theta, \sigma_e^2 \sim N(\boldsymbol{A}_i \left[1 + \boldsymbol{b}_i \exp(-\boldsymbol{K}_i t_{ij})\right]^{-1}, \sigma_e^2), \ \theta = \left[\boldsymbol{A}_i, \boldsymbol{b}_i, \boldsymbol{K}_i\right]'$$

Portanto, a função de verossimilhança é:

$$p(\boldsymbol{y} | \theta, \sigma_e^2) \propto \prod_{i=1}^{N} \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp\left\{ -\frac{\left[y_{ij} - \boldsymbol{A}_i \left[1 + \boldsymbol{b}_i \exp(-\boldsymbol{K}_i t_{ij})\right]^{-1} \right]^2}{2\sigma_e^2} \right\}$$

As distribuições a priori utilizadas foram:

$$\theta \sim N(\boldsymbol{\mu}, \sigma_\theta^2), \ \boldsymbol{\mu} = \left[\mu_{\mathrm{A}}, \mu_{\mathrm{b}}, \mu_{\mathrm{K}}\right]$$

$$\boldsymbol{\mu} \sim Uniforme\left(Linf, Lsup\right)$$

$$\frac{1}{\sigma_e^2} \sim GI\left(\alpha, \beta\right), \ \frac{1}{\sigma_{\theta_i}^2} \sim GI\left(\alpha_1, \beta_1\right),$$ em que:

Linf, *Lsup*, α, β, α_1, β_1, são os parâmetros das distribuições a priori, também chamados de hiperparâmetros.

Toda a análise foi conduzida no software Winbugs, utilizando-se 25.000 iterações na implementação dos algoritmos MCMC. A rotina desenvolvida é apresentada no Quadro 1.

Quadro 1. Código do Winbugs para análise de curva de crescimento das cultivares de feijoeiro.

```
model {
for (i in 1:K) {
for (j in 1:n) {
Y[i, j] ~ dnorm(eta[i, j], tauC)
eta[i, j] <- phi[i, 1] / (1 + phi[i, 2] * exp(phi[i, 3] * x[j]))
Y1[i, j] ~ dnorm(eta1[i, j], tauC1)
eta1[i, j] <- phi1[i, 1] / (1 + phi1[i, 2] * exp(phi1[i, 3] * x[j]))
for (k in 1:3) {
phi[i, k] ~ dnorm(mu[k], tau[k])
phi1[i, k] ~ dnorm(mu1[k], tau1[k]) }}
tauC ~ dgamma(1.0, 1.0E-3)  VarC <- 1 / tauC
tauC1 ~ dgamma(1.0, 1.0E-3)VarC1 <- 1 / tauC1
mu[1] ~ dunif(130, 190) mu[2] ~ dunif(10,400)
mu[3] ~ dunif(0.05, 0.52) mu1[1] ~ dunif(130, 190)
mu1[2] ~ dunif(10,400) mu1[3] ~ dunif(0.05, 0.52)
tau[1] ~ dgamma(1, 1E-3) tau[2] ~ dgamma(1, 1E-3)
tau[3] ~ dgamma(1, 1E-3) tau1[1] ~ dgamma(1, 1E-3)
tau1[2] ~ dgamma(1.0, 1.0E-3)
tau1[3] ~ dgamma(1.0, 1.0E-3)
dif[1]<-mu[1]-mu1[1]
dif[2]<-mu[3]-mu1[3] }
```

As amostras das distribuições marginais a posteriori para as diferenças dos parâmetros foram obtidas por meio da utilização das amostras geradas para as distribuições marginais dos parâmteros da curva representativa de cada população, ou seja, considerou-se a diferença como uma função, dada por:

$$f(x) = (\boldsymbol{\mu}_{\mathbf{cultivar1}} - \boldsymbol{\mu}_{\mathbf{cultivar2}})$$

Assim, é possível testar a hipótese de igualdade dos parâmetros mediante avaliação do intervalo de credibilidade para as diferenças, pois, se este vier a conter o valor zero conclui-se que, em média, os parâmetros das duas populações envolvidoas no contraste são iguais. Esta metodologia foi apresentada por Silva et al. (2005) para comparar parâmetros de curvas de lactação de cabras referentes a duas diferentes populações.

Resultados

Na Tabela 1 estão apresentadas as estimativas dos parâmetros do modelo de crescimento logístico ajustado aos dados das cultivares de feijoeiro Neguinho e Carioca.

Tabela 1 – Estimativas dos parâmetros do modelo de crescimento logístico ajustado aos dados das cultivares de feijoeiro Neguinho e Carioca.

Parâmetros	Média	Desvio padrão	Erro[2/] MC	Intervalo de credibilidade	
				2.5%	97.5%
A_1[1/]	163,2000	5,7620	0,0437	151,6000	174,7000
B_1	42,9000	0,6322	0,0560	41,9400	43,8000
K_1	0,0606	0,0113	0,0001	0,0502	0,0900
σ^2_1	57,7300	4,5270	0,0582	46,6100	64,3600
A_2	160,2000	17,4800	0,1866	131,4000	188,6000
B_2	205,4000	113,2000	1,1460	20,2000	390,6000
K_2	0,2871	0,1355	0,0012	0,0614	0,5083
σ^2_2	0,0082	0,0694	0,0005	0,0003	0,0429
$A_1 - A_2$	3,2870	18,41	0,0137	-28,6400	35,2900
$K_1 - K_2$	-0,2255	0,1353	0,0009	-0,4494	-0,0013

1/ Índice 1 e 2 correspondem às cultivares de feijoeiro Neguinho e Carioca, respectivamente; 2/ Erro MC: erro de Monte Carlo.

As curvas de crescimento representativas de cada cultivar estão apresentadas na Figura 1.

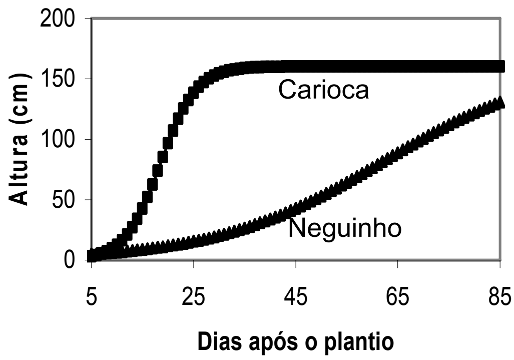

Figura 1. Curvas de crescimento estimadas para cada cultivar estudada.

As Figuras 2, 3, 4 e 5 mostram as distribuições marginais dos parâmetros do modelo de crescimento Logístico, da variância residual e das diferenças entre os parâmetros A e K. De forma geral estas figuras retratam de forma prática a utilização da metodologia Bayesiana, a qual está fundamentada em distribuições de probabilidade para estimar os parâmetros de interesse.

Discussão

De acordo com os resultados apresentados na Tabela 1 é possível concluir que, em média, o parâmetro A, ou seja, a altura máxima atingida pela planta, não apresentou diferença significativa entre as duas cultivares comparadas. O mesmo não foi verificado para o parâmetro K, taxa de maturidade, pois se observou que o intervalo de credibilidade contém apenas valores negativos, o que permite afirmar que a cultivar Carioca apresentou maior precocidade que a cultivar Neguinho.

Conforme Vieira, et al. (2005) informações a respeito de cultivares de feijoeiros mais precoces em relação ao crescimento é importante para avaliar sistemas de plantio e para recomendar adubações referentes a cada sistema. Estes autores compararam a evolução do crescimento das cultivares Talismã e Ouro Negro sob os sistemas de plantio direto e convencional e concluíram que esta última apresentou maior eficiência de crescimento ao se utilizar plantio direto.

Além da comparação entre os parâmetros das curvas pode-se também inferir em relação à precisão do ajuste do modelo não-linear aos dados de cada cultivar estudada. Dessa forma tem-se que os dados provenientes da cultivar Carioca apresentaram menor variabilidade em relação a forma da curva de crescimento, uma vez que a variância (σ^2_2) é menor que aquela apresentada pela outra cultivar.

Conclusões

A metodologia utilizada permitiu comparar curvas de crescimento de forma eficiente, sem utilizar aproximações assintóticas. Além disso, ficou demonstrado que a cultivar Carioca apresenta um incremento em altura maior que o da cultivar Neguinho, podendo esta informação ser explorada em experimentos futuros.

Referências

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.

ROSA, G.J.M. Análise Bayesiana de Modelos Mistos Robustos via Amostrador de Gibbs, ESALQ-USP:1998.

SILVA, F.F.; MUNIZ, J. A.; AQUINO, L.H.; SÁFADI, T. Abordagem Bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordem de parto. **Pesquisa Agrop. Brasileira**, v.40, n.1. p.27-33, jan. 2005.

SPIEGELHALTER, D. J., THOMAS, A. AND BEST, N. G. (2000) WinBUGS Version 1.3 User Manual.Cambridge:Medical Research Council Biostatistics Unit. www.mrc-bsu.cam.ac.uk/bugs

VIEIRA, N.M.B., ALVES JUNIOR, J., ANDRADE, M.J.B., CARVALHO, J.G. MORAES, A.R. **Altura de planta do feijoeiro cvs. talismã e ouro negro em plantio direto e convencional**. In: Congresso Nacional de Pesquisa de Feijão, 8., 2005. Goiânia-GO.

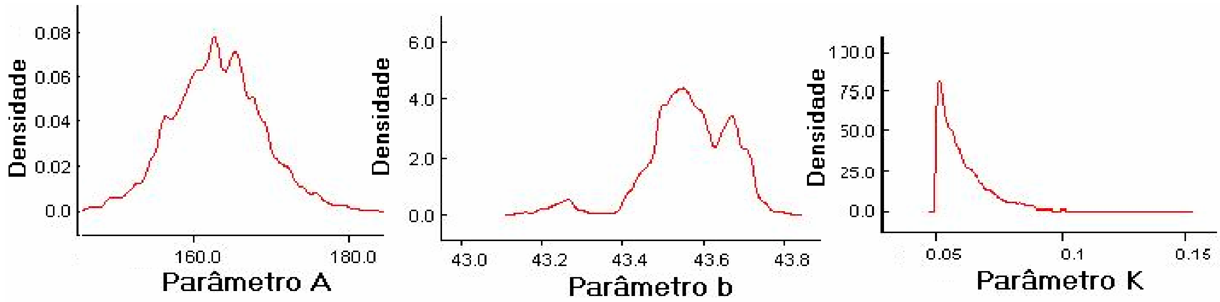

Figura 2. Amostras das distribuições marginais dos parâmetros do modelo de crescimento logístico ajustado aos dados da cultivar de feijoeiro Neguinho.

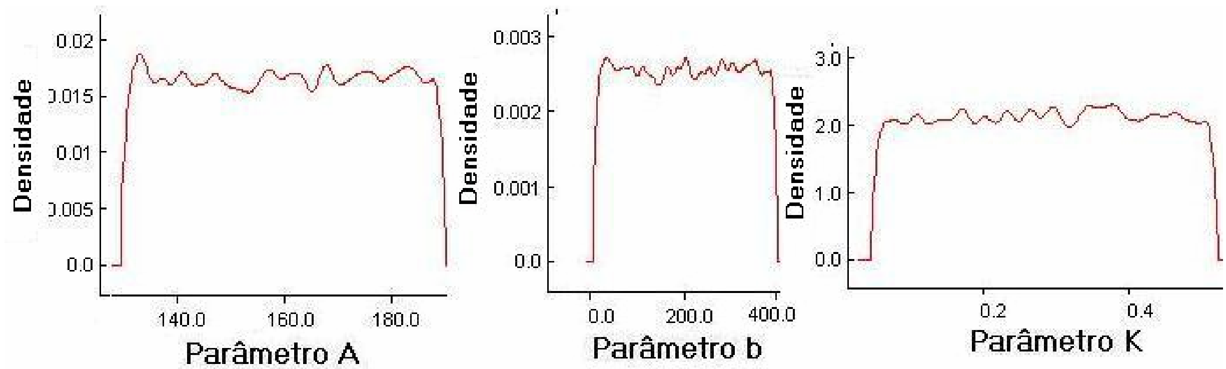

Figura 3. Amostras das distribuições marginais dos parâmetros do modelo de crescimento logístico ajustado aos dados da cultivar de feijoeiro Carioca.

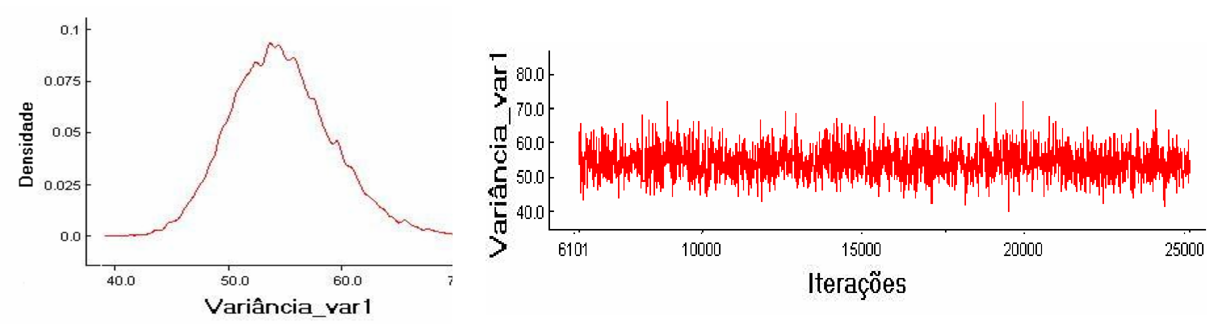

Figura 4. Amostras da distribuição marginal e cadeia de valores da variância do erro para a cultivar de feijoeiro Neguinho.

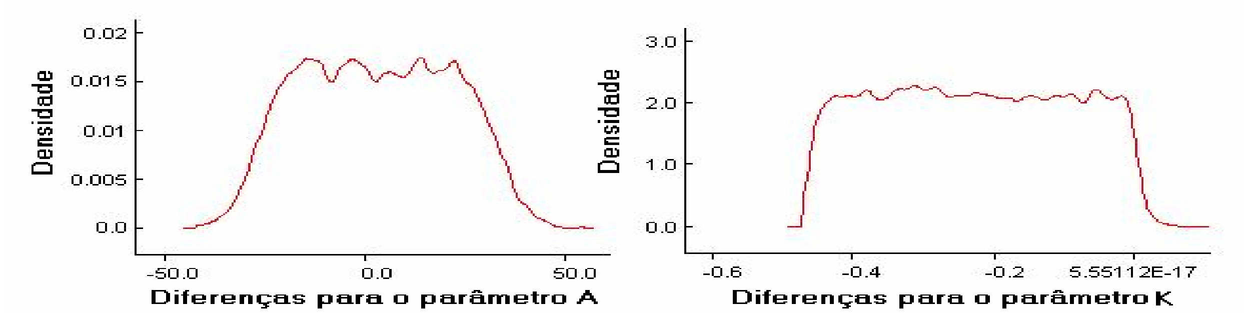

Figura 5. Amostras das distribuições marginais das diferenças para os parâmetros A e K.