

Aplicação dos modelos lineares generalizados na análise de dados da brotação e enraizamento de estacas de amoreira coletadas no Planalto Sul Catarinense

Daniela Tomazelli^{1†}, Simone S. Werner², Tássio D. Rech³, Murilo D. Costa⁴

¹Doutoranda do Programa de Pós-Graduação em Ciências do Solo, Centro de Ciências Agrovetinárias, Universidade do Estado de Santa Catarina - CAV/UDESC.

²Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina/Estação Experimental de Lages. E-mail: simonewerner@epagri.sc.gov.br.

³Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina/Estação Experimental de Lages. E-mail: tassior@epagri.sc.gov.br.

⁴Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina/Estação Experimental de Lages. E-mail: murilodc@epagri.sc.gov.br.

Resumo: A amoreira-negra é uma espécie arbórea de crescimento rápido, com facilidade de adaptação às diferentes condições climáticas do Brasil, sendo uma alternativa de utilização como mourão vivo. O presente trabalho objetivou estudar a aplicação de modelos lineares generalizados na análise dos dados de enraizamento e brotação de estacas de amoreira-negra provenientes de diferentes locais de origem do Planalto Sul Catarinense. Para isso, um experimento foi conduzido em delineamento experimental inteiramente casualizado, sendo os tratamentos constituídos por seis locais de origens do material vegetativo: Araçá- São José do Cerrito; Produtor 1- São José do Cerrito; Produtor 2- São José do Cerrito; Epagri-Lages; Frei Rogério-Lages; e Bocaína do Sul. Cada parcela foi constituída de 10 estacas, sendo utilizadas 4 repetições por tratamento. Após 85 dias de estaquia, avaliou-se o número de estacas enraizadas e o número de estacas com brotos em cada parcela. Para análises dos dados foram considerados os modelos binomial e binomial com parâmetro de dispersão constante. Todas as análises foram realizadas com o auxílio do ambiente R. Em ambos os casos observou-se subdispersão e o modelo com o parâmetro de dispersão constante (quasibinomial) foi o que proporcionou o melhor ajuste.

Palavras-chave: Modelo binomial; subdispersão; dados de proporção.

Abstract: Black's mulberry is a fast growing tree species with ease adaptation to the different climatic conditions of Brazil, being an alternative of use as living mourão. The present work aimed to study the application of generalized linear models in the analysis of rooting and sprouting data of black's mulberry cuttings' from different locations of the Southern Plateau of Santa Catarina. An experiment was carried out in a completely randomized experimental design, considering as treatments six places of origin of the vegetative material. Each experimental plot consisted of 10 cuttings, with 4 replicates per treatment. The treatments were evaluated after 85 days by the number of rooted cuttings and the number of cuttings with shoots in each plot. To analyze the data, were considered the binomial and quasibinomial models. All analyzes were performed with the aid of the R environment. In both cases subdispersion was observed and the model with the constant dispersion parameter (quasibinomial) was the one that provided the best fit.

Keywords: Binomial model; underdispersion; proportion data.

[†]Autora correspondente: danitomazelli@hotmail.com.

Introdução

Muitas são as avaliações realizadas na área agrônômica que geram dados de proporções de inteiros. O uso de modelos lineares clássicos, em geral, não é apropriado para analisar tais dados, pois as pressuposições do modelo não são atendidas. Uma alternativa para a análise desse tipo de dados é a utilização da teoria de modelos lineares generalizados, sendo a distribuição binomial, um caso particular, indicada para essas situações (McCULLAGH, NELDER, 1989; DEMÉTRIO, 2001).

Ao supormos que uma variável aleatória segue uma distribuição binomial de índice m_i e parâmetro π_i , ou seja $Y_i \sim Bin(m_i, \pi_i)$ assumimos que a esperança e a variância são dadas por: $E(Y_i) = \mu_i = m_i\pi_i$ e $Var(Y_i) = m_i\pi_i(1 - \pi_i) = V(\mu_i)$, nesse caso o parâmetro de escala é conhecido $\phi = 1$. No entanto tal suposição nem sempre é satisfeita, existindo muitos casos nos quais a variância observada não é aquela prevista pela distribuição binomial (HINDE, DEMÉTRIO, 1998). Nos casos conhecidos como subdispersão, a variância observada é menor do que a prevista pelo modelo binomial enquanto que nos casos de superdispersão a variância observada é maior do que a prevista pelo modelo. Essa dispersão pode ser acomodada assumindo uma forma mais geral para variância de Y_i , por exemplo: $Var(Y_i) = \phi V_i(\mu_i)$. Nesse caso, a estimação pode ser feita a partir do método da máxima quase-verossimilhança (LOGAN, 2011).

A amoreira-negra, de acordo com Lorenzi (2002), é uma espécie arbórea de crescimento rápido, com facilidade de adaptação às diferentes condições climáticas do Brasil. No sul do Brasil pode ser uma alternativa para utilização como mourão vivo, permitindo redução do tempo de isolamento das áreas ao acesso dos animais e contribuir com a melhoria no conforto e incremento de produtividade animal, proteção do solo, redução de custos e dos riscos associados ao uso de madeiras tratadas. Com essa finalidade, existe a necessidade de conhecer o potencial da amoreira-negra quanto ao enraizamento e brotação de suas estacas.

O presente trabalho objetivou estudar a aplicação de modelos lineares generalizados na análise dos dados de enraizamento e brotação de estacas de amoreira-negra provenientes de diferentes locais de origem do Planalto Sul Catarinense.

Material e Métodos

As estacas de amoreira foram obtidas de plantas em estágio de dormência, nos meses de junho e julho de 2017, nos municípios de São José do Cerrito, Bocaína do Sul e Lages. Foram selecionadas ramos de plantas adultas e com aspecto saudável, diâmetro de base entre 8 e 30 mm. O material selecionado foi plantado em mistura (1:1, volume:volume) de substrato comercial (PlantmaxT) e solo agrícola (Cambissolo húmico), em caixas de madeira com 1m por 4m, em casa de vegetação.

O experimento foi realizado na Epagri-Lages, conduzido no delineamento experimental inteiramente casualizado, sendo os tratamentos constituídos por seis locais de origens do material vegetativo: Araçá- São José do Cerrito; Produtor 1- São José do Cerrito; Produtor 2- São José do Cerrito; Epagri-Lages; Frei Rogério-Lages; e Bocaína do Sul. Cada parcela foi constituída de 10 estacas, sendo utilizadas 4 repetições por tratamento.

Após 85 dias de estaquia, avaliou-se o número de estacas enraizadas e o número de estacas com brotos em cada parcela. As duas variáveis resposta, Y_i , foram a proporção de estacas enraizadas e a proporção de estacas com brotos.

A distribuição considerada inicialmente foi a binomial; como função de ligação empregou-se a função logística, como parte sistemática adotou-se um delineamento inteiramente casualizado e como tratamentos as diferentes origens. Utilizou-se o diâmetro inicial das estacas como covariável no modelo. O preditor linear, para o modelo maximal, dado por:

$$\eta_i = \text{logit}(\pi_i) = \mu + \beta d_i + \alpha_k$$

em que, $i = 1, \dots, n$, μ representa o efeito associado a média geral, β é o coeficiente associado ao diâmetro inicial (d) e α_k o efeito associado ao k -ésimo local de origem, $k = 1; \dots; 6$.

Verificada a subdispersão, incorporou-se um parâmetro de dispersão constante (ϕ), utilizando para estimação o método da máxima quase-verossimilhança (modelo *quasibinomial*) (LOGAN, 2011).

O parâmetro de dispersão, ϕ , foi estimado a partir da estatística de Pearson generalizada, por:

$$\hat{\phi} = \frac{1}{n-p} \sum_{i=1}^n \frac{(y_i - m_i \hat{\pi}_i)^2}{m_i \hat{\pi}_i (1 - \hat{\pi}_i)}$$

Para seleção do modelo, usou-se a estatística F:

$$F = \frac{(D_2 - D_1)/(f_2 - f_1)}{\hat{\phi}} \sim F_{f_2 - f_1, f_3}$$

em que $\hat{\phi}$ é estimado a partir do modelo maximal, com f_3 graus de liberdade, e D_1 e D_2 são valores da estatística desvio para modelos encaixados com $f_1 < f_2$ parâmetros (JØRGENSEN, 1987). Para verificar o ajuste do modelo utilizou-se o gráfico normal de probabilidades com envelope simulado disponível no pacote *hnp* (MORAL, HINDE, DEMÉTRIO, 2017) e o gráfico de valores ajustados versus resíduos.

Quando observado efeito significativo de tratamento, a comparação entre os locais de acesso foi realizada utilizando contrastes com auxílio do pacote *multcomp* (HOTHORN, BRETZ, WESTFALL, 2008). Todas as análises foram realizadas considerando o nível de 5% de significância e utilizando o ambiente R: A Language and Environment for Statistical Computing (2018).

Resultados e Discussão

Ajustando o modelo logístico padrão aos dados da proporção de estacas com brotos, obteve-se para o desvio residual o valor de 2,45, com 17 graus de liberdade do resíduo. De acordo com Hinde e Demétrio (1998), no caso da distribuição binomial, espera-se que o valor do desvio residual seja aproximadamente o número de graus de liberdade associado a esse resíduo. Nesse caso o número de graus de liberdade associado ao desvio residual foi aproximadamente 7 vezes o desvio residual estimado, indicando subdispersão, o que foi confirmado pelo gráfico normal de probabilidades conforme mostra a Figura 1(a).

O modelo logístico com dispersão constante ($\hat{\phi} = 0,14$) ajustou-se bem aos dados de proporções de estacas com brotos, conforme mostra o gráfico normal de probabilidades na Figura 1(b).

Comportamento semelhante foi observado para os dados da proporção de estacas com raízes, que apresentaram para o desvio residual o valor de 6,96, com 17 graus de liberdade. A Figura 2 apresenta o gráfico normal de probabilidades com envelope simulado

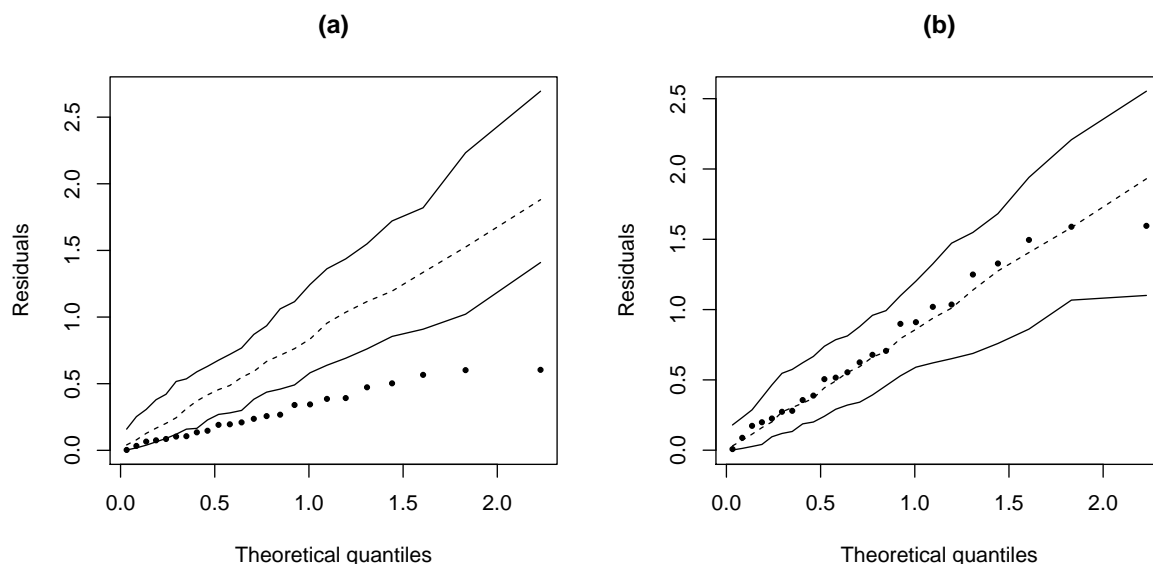


Figura 1: Gráfico normal de probabilidades com envelope simulado para o modelo logístico padrão (a) e para o modelo logístico com dispersão constante ($\hat{\phi} = 0,14$) (b) para proporção de estacas com brotos

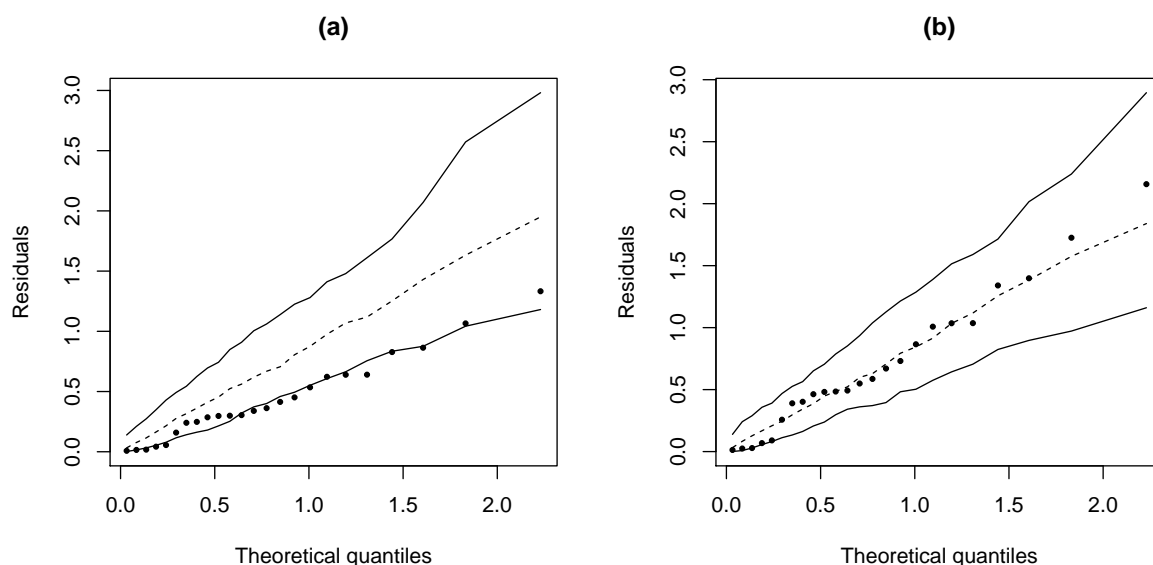


Figura 2: Gráfico normal de probabilidades com envelope simulado para o modelo logístico padrão (a) e para o modelo logístico com dispersão constante ($\hat{\phi} = 0,38$) (b) para proporção de estacas enraizadas

considerando os ajustes do modelo binomial (a) e *quasibinomial* (b). Nesse caso, quando ajustou-se o modelo com parâmetro de dispersão obteve-se $\hat{\phi} = 0,38$.

Observou-se efeito significativo do local de acesso das estacas de amoreiras-pretas do Planalto Sul Catarinense tanto para a proporção de estacas com brotos (p -valor = 0,0430), como para proporção de estacas enraizadas (p -valor = 0,0012).

O maior índice de emissão de brotos foi do Produtor 2-São Joaquim com 0,95, diferindo

do percentual de brotos observados nas estacas provenientes de Frei Rogério e Araçá ambas com 0,65 de brotação e da propriedade 1 de SJC com 0,67. O material coletado na Epagri, município de Lages, obteve a proporção de 0,82 de estacas enraizadas, no entanto não diferiu do Produtor 2-São Joaquim que obteve a segunda maior proporção de enraizamento (0,70).

Novas coletas e estudos são necessários para determinar quais os fatores locais que podem influenciar na taxa de enraizamento e brotação da amoreira negra e sua utilização como mourão vivo.

Considerações Finais

Os dados de enraizamento e brotação das estacas de amoreira negra apresentaram dispersão menor do que a prevista pela distribuição binomial.

O modelo *quasibinomial* ajustou-se bem aos dados de acordo com os gráficos de diagnóstico, sendo adequado nesse caso.

Referências bibliográficas

- DEMÉTRIO, C.G.B. Modelos lineares generalizados em experimentação agrônômica. *In: Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria*, 46., 2001, Piracicaba. Anais... Piracicaba: ESALQ/USP, 2001. p.113.
- HINDE, J.P.; DEMÉTRIO, C.G.B. Overdispersion: models and estimation. *In: SINAPE*, 13., 1998, Caxambu. Anais... Caxambu: SINAPE, 1998. 73p.
- HOTHORN, Torsten; BRETZ, Frank e WESTFALL, Peter. Simultaneous Inference in General Parametric Models. *Biometrical Journal*. 2008. 50(3), 346–363.
- JØRGENSEN, B. Small dispersion asymptotics. *Revista Brasileira de Probabilidade e Estatística*, 1987.
- LORAN, M. *Biostatistical Design and Analysis Using R: A Practical Guide*. John Wiley & Sons, 2011.
- LORENZI, H. *Árvores brasileiras*. 4ed. Nova Odessa : SP - Instituto Plantarum de Estudos da Flora LTDA, 2002.
- MORAL R.A.; HINDE J.; DEMÉTRIO C.G.B. Half-Normal Plots and Overdispersed Models in R: The hnp Package. *Journal of Statistical Software*, v.81, n.10, p.1-23, 2017.
- McCULLAGH, P.; NELDER, J. A. *Generalized linear models*. 2nd. ed. London : Chapman & Hall, 1989. 511p.
- R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. Disponível em: <https://www.R-project.org/>.