

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

FÁBIO JANONI CARVALHO

ANÁLISE DE EXPERIMENTOS DE GERMINAÇÃO USANDO OS MODELOS
LINEARES GENERALIZADOS

UBERLÂNDIA
MINAS GERAIS – BRASIL
2016

FÁBIO JANONI CARVALHO

ANÁLISE DE EXPERIMENTOS DE GERMINAÇÃO USANDO OS MODELOS
LINEARES GENERALIZADOS

Dissertação apresentada à Universidade Federal de
Uberlândia, como parte das exigências do Programa de Pós-
graduação em Agronomia – Mestrado, área de concentração
em Fitotecnia, para obtenção do título de “Mestre”.

Orientadora

Prof. Dra. Denise Garcia de Santana

UBERLÂNDIA
MINAS GERAIS – BRASIL
2016

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

- C331a
2016
- Carvalho, Fábio Janoni, 1991
Análise de experimentos de germinação usando os modelos lineares generalizados / Fábio Janoni Carvalho. - 2016.
102 f. : il.
- Orientadora: Denise Garcia de Santana.
Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Agronomia.
Inclui bibliografia.
1. Agronomia - Teses. 2. Sementes - Germinação - Teses. 3. Modelos lineares (Estatística) - Teses. I. Santana, Denise Garcia de. II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Agronomia. III. Título.

CDU: 631

FÁBIO JANONI CARVALHO

ANÁLISE DE EXPERIMENTOS DE GERMINAÇÃO USANDO OS MODELOS
LINEARES GENERALIZADOS

Dissertação apresentada à Universidade Federal de
Uberlândia, como parte das exigências do Programa de Pós-
graduação em Agronomia – Mestrado, área de concentração
em Fitotecnia, para obtenção do título de “Mestre”.

APROVADO em 27 de janeiro de 2016.



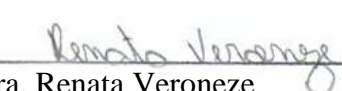
Prof. Dr. Gabriel Mascarenhas Maciel

docente/UFU



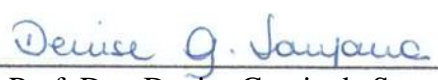
Prof. Dr. Marcelo Tavares

docente/UFU



Dra. Renata Veroneze

Pesquisadora/ EPAMIG



Prof. Dra. Denise Garcia de Santana
ICIAG-UFU
(Orientadora)

UBERLÂNDIA
MINAS GERAIS – BRASIL
2016

AGRADECIMENTOS

Esta dissertação foi construída com o suporte de diversas pessoas, direta e indiretamente. Primeiramente agradeço a Deus, que sempre me proporcionou saúde, força, fé e paz de espírito para terminar esta trajetória.

A minha Orientadora Denise Garcia de Santana pelos anos de orientação, companheirismo, ensinamentos, amizade, e pela paciência em me guiar pelo caminho correto da ciência. Ao Lúcio Borges de Araújo pelos atendimentos e auxílios estatísticos. A todos os membros do Instituto de Ciências Agrárias, tanto do Campus Umuarama quanto de Monte Carmelo. Pela banca avaliadora, por aceitarem participar e sugerirem melhorias ao trabalho.

Agradeço também aos meus pais Maria Edna Janoni Carvalho e Antonio Manoel Carvalho pelo amor, ajuda e proteção nesta trajetória da vida. Ao meu irmão Rodrigo Janoni Carvalho pelo suporte, ajuda e estabilidade emocional que sempre me passou. A todo restante da minha família pelo suporte e carinho em especial minhas madrinhas Maria Neuza, Gisele, Marina, minha vó Erminia e meus falecidos avós Antero, Ana Carolina e João.

Aos meus amigos que acompanharam toda esta trajetória de árduo trabalho, pela amizade, paciência e palavras de apoio. Agradeço em especial Paulo Vinícius Ribeiro pela ajuda na estruturação do trabalho.

Cabe aqui o meu muito obrigado. Este trabalho também é de vocês!

SUMÁRIO

	Página
CAPÍTULO 1: As contribuições de Gauss e Fisher para os Modelos Lineares Generalizados.....	01
Referências Bibliográficas.....	04
CAPÍTULO 2: Como analisar experimentos de germinação usando os Modelos Lineares Generalizados?.....	06
Resumo.....	06
Abstract.....	07
Introdução.....	08
Material e Métodos.....	10
Análise de variância (ANOVA).....	12
Modelos Lineares Generalizados (MLG).....	14
Análise de <i>deviance</i> (ANODEV).....	17
Qualidade do ajuste do modelo pelos critérios de AIC e BIC.....	20
Qualidade do ajuste pela Distância de Cook.....	21
Análise gráfica de “q-q plot”.....	21
Resultados.....	22
Discussão.....	25
Conclusão.....	29
Referências Bibliográficas.....	30
Anexos.....	36
CAPÍTULO 3: Por que analisar experimentos de germinação usando os Modelos Lineares Generalizados?.....	38
Resumo.....	38
Abstract.....	39
Introdução.....	40
Material e Métodos.....	42
Planejamento e execução.....	42
Retirada de <i>outliers</i> pelo método de Box-Plot.....	46
Pressuposições de Normalidade dos resíduos e homogeneidade das variâncias.....	46
ANOVA e MLG com distribuição Normal e função de ligação identidade..	46
MLG com distribuição Binomial e função de ligação logit.....	47
Análise de <i>deviance</i> (ANODEV).....	47
Ajuste do modelo.....	48
Resultados.....	48
Discussão.....	59
Referências Bibliográficas.....	64
Anexos.....	70

CAPÍTULO I

OS ESTATÍSTICOS QUE CONTRIBUÍRAM PARA OS MODELOS LINEARES GENERALIZADOS

Fábio Janoni Carvalho

Os tratamentos estatísticos são a base da ratificação científica em qualquer área em que são aplicados. Nas Ciências Agrárias, essa situação não seria diferente: a primeira técnica elaborada para dados, com vistas a seguir uma distribuição Normal foi a regressão linear múltipla, criada por Legendre em 1805. Em 1809, o modelo foi ajustado por Gauss à distribuição Normal, com média zero e variância homogênea para os resíduos.

As contribuições de Gauss advieram da necessidade astronômica de reduzir do número de observações. O trabalho desse estudioso assumiu os coeficientes de regressão e mostrou o cálculo dos valores estimados pelo método dos mínimos quadrados. Além disso, o método de inferência, considerado Bayesiano, atrelou-se a amostras com distribuição Normal, sendo a principal limitação de seu modelo. A regressão linear múltipla foi o marco para o início dos modelos mais refinados, mas ainda se afastava dos problemas práticos da área agrícola.

O cenário da experimentação agrícola começou a se movimentar com os trabalhos de Fisher (1920) iniciados na Estação de Pesquisa Agrícola de Rothamsted no Reino Unido. Em 1922, criou o Método da Máxima Verossimilhança (FISHER, 1922) e, com dados adquiridos em anos de pesquisa, avançou com a Análise de Variância (ANOVA) e os experimentos fatoriais (FISHER, 1925), pois até então os fatores eram analisados apenas de maneira individual. De fato, o modelo proposto por Fisher divergiu da regressão linear múltipla, pois as variáveis eram contínuas e não restritas a valores entre zero e um. Fisher alertava, porém, sobre os pressupostos exigidos pela ANOVA, principalmente no que se refere às variâncias homogêneas e à distribuição Normal dos resíduos.

Vale ressaltar que a ANOVA foi absorvida aos Modelos Lineares Generalizados (MLG) propostos por Nelder e Wedderburn em 1972. Apesar da eficiência encontrada, ainda se tratava de um modelo limitado, por ser voltado a rígidas pressuposições que

fizeram Fisher continuar suas análises e desenvolver modelos para outras famílias de distribuição.

Ensaio com diluições fizeram com que Fisher criasse um modelo de estimação da máxima verossimilhança para a distribuição de Poisson (FISHER, 1920). Além disso, ele expandiu os tratamentos estatísticos ao definir a família exponencial, uma classe de distribuições com propriedades específicas, e a estimação de seus parâmetros, teoria imprescindível dentro dos MLG (FISHER, 1934).

A partir das teorias elaboradas por Fisher, novas metodologias estatísticas foram criadas. A análise *probit* foi uma técnica estabelecida por Bliss (1935) que ascendeu entre os trabalhos biológicos, tendo envolvido pela primeira vez a distribuição Binomial, em que a variável assumia valores entre zero e um, expandindo sua aplicabilidade para ensaios de toxicologia, com dose-resposta, e para a germinação de sementes – essa importante análise também foi incorporada aos MLG. Nesse entremeio, Dyke e Patterson (1952), ao trabalharem com dados de proporção, ajustaram os Modelos Lineares Logísticos, o que garantiu outro tratamento aos dados de natureza Binomial.

Ademais, ajustaram-se os modelos log-lineares para contagens, envolvendo distribuição Poisson e Multinomial (BIRCH, 1963), e os modelos de regressão para dados de sobrevivência, abrangendo a distribuição exponencial (FEIGL; ZELEN, 1965; ZIPPIN; ARMITAGE, 1966; GASSER, 1967) – eles foram integrados ao MLG por Aitkin e Clayton (1980).

Diante de modelos que apresentavam a mesma estrutura básica analítica, Nelder e Wedderburn (1972) elaboraram um modelo generalizado que absorvesse as mais diversas distribuições dentro da família exponencial. O modelo foi complementado por McCullagh e Nelder (1989) e encarava a generalização em dois aspectos: na distribuição dos dados e no componente que se envolve com a média, a função de ligação.

Nesses termos, a popularização de tratamentos estatísticos mais refinados e uma diversidade de possibilidades analíticas foram possíveis somente com o advento dos avanços computacionais. Assim, os MLG começaram a ganhar destaque, pois os computadores economizavam grandiosas fórmulas matemáticas do modelo e as resolviam em segundos, possibilitando ao avaliador a análise do banco de dados por diversas vezes.

Mesmo com o advento computacional, os tratamentos estatísticos nas Ciências Agrárias estacionaram-se no modelo da ANOVA proposto por Fisher. Apesar de um modelo eficiente, os pressupostos que o regem impossibilitam a análise de alguns dados agrícolas. Há, porquanto, o “forçamento” ao ajuste com a transformação de dados ou o fato de o analista ignorar se uma das pressuposições não é atendida no banco de dados.

Para a área de sementes, o ajuste dos dados de germinação para a distribuição Binomial foi estudado, o que popularizou a análise *probit* dos dados (FINNEY, 1971). Trabalhos com a análise de sobrevivência permitiram a expansão e utilização de outros métodos estatísticos (LAWLESS, 1982), o que descarta a justificativa de que os dados possam ser tratados apenas pela ANOVA.

Na maioria dos experimentos predomina o tratamento convencional dos dados (ANOVA), sendo escassos os trabalhos que se utilizem de outros modelos, a exemplo daqueles absorvidos pelos MLG. Gauss e Fisher foram os percussores dos primeiros métodos estatísticos para inferência dos dados, por desencadearem outras técnicas, sendo a contribuição desses autores fundamental para a futura constituição dos MLG.

Pelo enquadramento da variável germinação em outras distribuições pertencentes à família exponencial, novas técnicas podem e devem ser utilizadas para descobrir, sobretudo, quais delas se ajustam melhor aos dados. Para tanto, esta dissertação conta com outros dois capítulos: no Capítulo II os Modelos Lineares são introduzidos ao leitor, assim como sua aplicação na germinação de sementes utilizando-se de um exemplo ilustrativo; e no Capítulo III se objetiva, dentro de um banco de dados de germinação de 50 espécies florestais, verificar qual dos modelos apresentados, um com a distribuição Normal e outro com distribuição Binomial, obteve o melhor ajuste dentro do mesmo banco de dados original.

REFERÊNCIAS BIBLIOGRÁFICAS

AITKIN, M.; CLAYTON, D. The fitting of exponential, Weibull, and extreme value distributions to complex censored survival data using GLIM. **Journal of the Royal Statistical Society C29**, p.156-163, 1980.

BIRCH, M. W. Maximum likelihood in three-way contingency tables. **Journal of the Royal Statistical Society**, v.25, n.1, p.220-233, 1963.

BLISS, C. I. The calculation of the dosage-mortality curve. **Annals of Applied Biology**, v.22, p.134-167, 1935.

DYKE, G. V.; PATTERSON, H. D. Analysis of factorial arrangements when the data are proportions. **Biometrics**, v.8, p.1-12, 1952.

FEIGL, P.; ZELEN, M. Estimation of exponential survival probabilities with concomitant information. **Biometrics**, v.21, p.826-838, 1965.

FINNEY, D. J. **Probit analysis**, 3rd ed. Cambridge: Cambridge University Press, 1971. 272 p.

FISHER, R. A. A Mathematical examination of the methods of determining the accuracy of an observation by the mean error, and by the mean square error. **Monthly Notices of the Royal Astronomical Society**, v.80, n.8, p.758-770, 1920.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society**, v.222, p.309-368, 1922.

FISHER, R. A. **Statistical methods for research workers**. Edinburgh: Oliver and Boyd, 1925. 319 p.

FISHER, R. A. Two new properties of mathematical likelihood. **Proceedings of the Royal Society A144**, v.852, p.285-307, 1934.

GASSER, M. Exponential survival with covariance. **Journal of the American Statistical Association**, v.62, p.561-568, 1967.

GAUSS, C. F. **Theoria motus corporum coelestium in sectionibus conicis solem ambientium**. Hamburg: Perthes et Besser, 1809. 280 p.

LAWLESS, J. F. **Statistical models and methods for lifetime**. Nova York: John Wiley & Sons, 1982. 580 p.

LEGENDRE, A. M. **Nouvelles methodes pour la determination des orbites de comètes.** Paris: F. Didot, 1805. 290 p.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**, 2nd ed. London: Chapman & Hall, 1989. 493p.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**, v.135, p.370-384, 1972.

ZIPPIN, C.; ARMITAGE, P. Use of concomitant variables and incomplete survival information in the estimation of an exponential survival parameter. **Biometrics**, v.22, p.665-672, 1966.

CAPÍTULO II

COMO ANALISAR EXPERIMENTOS DE GERMINAÇÃO USANDO OS MODELOS LINEARES GENERALIZADOS?

Fábio Janoni Carvalho

Resumo: A análise de variância (ANOVA) é um dos principais modelos estatísticos aplicados em experimentos agronômicos, especialmente na área de sementes. Fundamentada em rígidas pressuposições, ela ainda perdura com o auxílio de técnicas como a transformação de dados. Por ser a ANOVA uma particularização dos Modelos Lineares Generalizados (MLG), um experimento clássico de germinação envolvendo sementes da espécie florestal *Copaifera langsdorffii* Desf. mostrará o espelhamento entre ambos os métodos de análise, e esse é um dos objetivos dessa pesquisa. Visou-se também comparar a qualidade do ajuste e a eficiência dos modelos para a germinação das sementes expressa em porcentagem com distribuição Normal e em número de sementes germinadas com distribuição Normal. Para atender a esses objetivos, sementes de *C. langsdorffii* foram arranjadas em delineamento inteiramente casualizado (DIC) com quatro repetições de 25 sementes em um esquema fatorial 4 x 3, sendo o primeiro fator relativo aos métodos para superação de dormência (M_1 , M_2 , M_3 e M_4) e o segundo, referente às amostras (A_1 , A_2 e A_3). Nos resultados expressos em porcentagem de germinação foram testadas as pressuposições da normalidade e independência dos resíduos e homocedasticidade por Shapiro-Wilk, Durbin-Watson e Levene, respectivamente. Aplicou-se o modelo de ANOVA para o experimento em esquema fatorial e DIC, assim como o MLG para distribuição Normal e função de ligação identidade. No que tange aos dados expressos em número de sementes germinadas, foi aplicado o MLG para distribuição Binomial e função de ligação logística. Para ambas as distribuições, a qualidade do ajuste foi determinada pelos critérios de informação de Akaike (AIC) e Bayesiano (BIC), distância de Cook e análise de “q-q plot”. Como esperado, para a porcentagem de germinação de sementes de copaíba, o modelo da ANOVA se igualou ao MLG para distribuição Normal, e eles indicaram efeito significativo para amostra e interação, sendo que uma análise prévia comprovou o atendimento a todas as pressuposições do modelo. Resultados similares apontaram para o MLG com distribuição Binomial, com significância para os mesmos efeitos do modelo Normal. Contudo, AIC e BIC indicaram que o modelo Binomial obteve melhor ajuste aos dados, assim como ocorreu maior acomodação dos valores no envelope simulado, com 95% de confiança. A distância de Cook não distinguiu os modelos, uma vez que eles se aproximaram em relação à quantidade de pontos influentes.

Palavras-chave: copaíba, critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC), espécie florestal, pressuposições da ANOVA, MLG

HOW TO ANALYZE GERMINATION EXPERIMENTS USING GENERALIZED LINEAR MODELS?

Fábio Janoni Carvalho

Abstract: Analysis of variance (ANOVA) is one of the most important statistical models applied in agronomic experiments, especially in the seeds area. Based on strong assumptions, it lasted for many years with the support of techniques such as data transformation. As ANOVA being a special case of Generalized Linear Models (GLM), a classic experiment of seeds germination of tree species *Copaifera langsdorffii* Desf. can show the mirroring between both methods of analysis, and this is one of the goals of this research. It also aimed to compare the quality of the adjustment and the efficiency of the models for the germination, expressed in percentage with Normal distribution and number of germinated seeds with Binomial distribution. To meet these objectives, seeds of *C. langsdorffii* were arranged in a completely randomized design with four replications of 25 seeds in a 4 x 3 factorial scheme, in which the first factor refers to the methods to overcome dormancy (M_1 , M_2 , M_3 and M_4) and the second effect is related to samples (A_1 , A_2 and A_3). For the results expressed in percentage of germination, the assumptions of normality and independence of residuals and homoscedasticity were tested by Shapiro-Wilk, Durbin-Watson and Levene, respectively. Then, it was applied an ANOVA model, as well as GLM with Normal distribution and identity link function. About the data expressed as number of germinated seeds, GLM was performed with Binomial distribution and logistics link function. For both distributions, the quality of the adjustment was determined by Akaike information criterion (AIC) and Bayesian information criterion (BIC), Cook's distance and "q-q plot" analysis. As expected, ANOVA model was equal to GLM with Normal distribution for the percentage of copaiba seed germination, and they indicated a significant effect of sample and interaction, as a previous analysis confirmed that all assumptions of the model were held. The GLM with Binomial distribution had the same significance of the effects as the Normal GLM. However, AIC and BIC indicated that Binomial model was better adjusted to data, and the accommodation of values to the simulated envelope with 95% confidence was greater. Cook's distance did not discriminate the models, since they approached to the same amount of influential points.

Keywords: copaiba, Akaike's information criterion (AIC), Bayesian's information criterion (BIC), forest species, ANOVA assumptions, GLM

INTRODUÇÃO

A análise de variância (ANOVA) é um dos principais modelos estatísticos aplicados em experimentos agronômicos, sendo que na área de sementes não é diferente. Atribuída a Fisher em 1925, popularizou-se nas pesquisas científicas a partir da publicação do livro *A análise de variância* (SCHEFFÉ, 1959).

Tal método enfatiza o papel crucial da repetição, da casualização e do controle local na eficiência dos experimentos, e está alicerçado no modelo linear normal (FISHER, 1925; STEEL; TORRIE, 1980; SOKAL; ROHLF, 2012). Sua consolidação em germinação se deve, entre outros fatores, à baixa variabilidade residual das espécies cultivadas, alcançada por décadas de melhoramento genético, e a uma certa facilidade de ajuste dos dados aos modelos lineares propostos por Fisher.

O auxílio de técnicas como a transformação de dados fez com que modelos de análise de variância perdurassem por muitos anos, mesmo sendo fundamentados segundo rígidas pressuposições. De uso e interpretação problemáticos e com o agravante de poder corrigir uma das pressuposições e violar outra atendida na escala original, a transformação é alvo de críticas severas pelos pesquisadores (AHRENS et al., 1990; WARTON; HUI, 2011; SILESHI, 2012).

Os Modelos Lineares Generalizados (MLG), por englobarem distribuições com menores exigências que a Normal, não dependem necessariamente de atendimento às pressuposições da ANOVA (WILSON; HARDY, 2002; CRAWLEY, 2007; WARTON; HUI, 2011). Ao agruparem diversas técnicas estatísticas, são definidos por uma distribuição de probabilidade pertencente à família exponencial paramétrica (NELDER; WEDDERBURN, 1972). Convém salientar que a palavra “generalizado” indica a possibilidade de utilização de um maior número distribuições para a variável resposta (CORDEIRO, 1986).

Nesse sentido, os MLG são a generalização ou a flexibilização dos modelos lineares clássicos de variáveis contínuas como a ANOVA, de forma que toda a estrutura para a estimação e predição pode ser estendida para modelos com outras distribuições aplicáveis também a variáveis discretas (LINDSEY, 1997; CRAWLEY, 2007; DOBSON; BARNETT, 2008). Portanto, são eficientes em Ciências Agrárias e Ciências Florestais, especialmente nas pesquisas com sementes, pela natureza discreta da variável germinação, que pode seguir a distribuição Binomial.

Os modelos são formados por um componente aleatório, correspondente à variável resposta; um componente sistemático, relacionado às variáveis explanatórias inseridas na forma de uma estrutura linear; e uma função de ligação. Esta última visa associar o componente aleatório ao sistemático (LINDSEY, 1997; FARAWAY, 2006; CRAWLEY, 2007; DOBSON; BARNETT, 2008).

Embora seja comum a representação do número de sementes germinadas em porcentagem, uma variável de natureza contínua (a variável original) é discreta e segue todos os critérios da distribuição Binomial. Para o ajuste dessa distribuição aos dados de germinação, exigem-se número fixo de sementes, independência de uma germinação em relação à outra, existência de apenas dois resultados possíveis (a semente germina ou não), a chance de a germinação ser constante (que todas as sementes recebam as mesmas condições para germinação, como temperatura e umidade) e contabilização do número de sementes germinadas pela variável aleatória (NETER et al., 1996; LEE et al., 2006; LARSON; FARBER, 2010). Desse modo, mesmo que o modelo Binomial exija atendimento a tais pressupostos, a montagem correta do experimento consegue contemplar as necessidades do modelo. Com a distribuição, os dados podem receber outros tratamentos estatísticos, sem alteração de comportamento e permitindo melhor ajuste (DOBSON; BARNETT, 2008).

Na análise de variância, é incomum a checagem dos valores ajustados pelo modelo em relação aos dados originais. A única medida comumente utilizada é o coeficiente de variação, que permite avaliar apenas a magnitude do erro experimental, sem qualquer implicação sobre a qualidade do modelo. Nos MLG, a checagem do ajuste do modelo, especialmente quanto à função de ligação, pode ser verificada pelo *deviance* (NELDER; WEDDERBURN, 1972; MCCULLAGH; NELDER, 1989), mas essa não é a única forma. Critérios adicionais como os de informação de Akaike (AKAIKE, 1973) e Bayesiano (SCHWARZ, 1978), assim como a distância de Cook (COOK, 1977; COOK; WEISBERG, 1982; HINES; HINES, 1995), ao apresentarem princípios distintos, auxiliam na análise relativa ao ajuste do modelo.

Por ser a ANOVA uma particularização dos MLG (RESENDE; BIELE, 2002), um experimento clássico de germinação envolvendo sementes da espécie florestal *Copaifera langsdorffii* Desf. mostrará o espelhamento entre ambos os métodos de análise, e esse foi um dos objetivos desta pesquisa. Visou-se também comparar a qualidade do ajuste e a eficiência do modelo para a germinação das sementes expressa

em porcentagem, com distribuição Normal; e em número de sementes germinadas, com distribuição Binomial.

MATERIAL E MÉTODOS

Uma pesquisa com sementes de 13 indivíduos de copaíba (*C. langsdorffii*) foi escolhida para representar experimentos clássicos de germinação com sementes florestais e nortear a aplicação dos Modelos Lineares Generalizados (MLG). Cabe destacar que a aplicação dos modelos apresentados pode ser estendida a qualquer experimento de germinação com estrutura fatorial e variáveis expressas em porcentagem (germinabilidade, plântulas normais, anormais, sementes mortas e outras). Pelo foco nos modelos, a metodologia do experimento de germinação foi resumidamente descrita.

Sementes de copaíba foram arranjadas em delineamento inteiramente casualizado com quatro repetições ($r = 4$) de 25 sementes ($n = 25$) em um esquema fatorial 4×3 , sendo o primeiro fator relativo aos métodos para superação de dormência (M_1 , M_2 , M_3 e M_4) e o segundo, referente às amostras (A_1 , A_2 e A_3), perfazendo 48 parcelas experimentais ($n = 48$). Os métodos consistiram em testemunha (M_1), desinfestação das sementes com hipoclorito de sódio a 0,05% por cinco minutos (M_2), embebição das sementes em água destilada por 24 horas (M_3) e por 48 horas (M_4).

Tais sementes foram distribuídas de forma alternada em duas folhas de papel de filtro, cobertas com outras duas folhas, confeccionando-se rolos. Os rolos foram distribuídos em incubadora BOD regulada a 25 °C e submetida à luz branca fluorescente contínua. Trinta e cinco dias após a semeadura foi contabilizado, entre outras características, o número de sementes germinadas (critério de plântulas). Os percentuais de germinação e o número de sementes germinadas, parte integrante dos resultados, porém comumente expressos na forma de médias, foram declarados na íntegra neste capítulo para que possam ser reproduzidos (Tabela 1).

TABELA 1. Percentual de germinação e número de sementes germinadas para diferentes métodos de superação de dormência (M₁, M₂, M₃ e M₄) em amostras (A₁, A₂ e A₃) de copaíba (*Copaifera langsdorffii* Desf.).

Método (i)	Amostra (j)	Repetição* (k)	Percentual de germinação	Número de sementes germinadas	Método (i)	Amostra (j)	Repetição* (k)	Percentual de germinação	Número de sementes germinadas
M ₁	A ₁	1	56	14	M ₃	A ₁	1	64	16
		2	84	21			2	56	14
		3	60	15			3	68	17
		4	76	19			4	64	16
M ₁	A ₂	1	44	11	M ₃	A ₂	1	28	7
		2	28	7			2	28	7
		3	44	11			3	16	4
		4	36	9			4	32	8
M ₁	A ₃	1	84	21	M ₃	A ₃	1	100	25
		2	92	23			2	100	25
		3	88	22			3	92	23
		4	88	22			4	100	25
M ₂	A ₁	1	56	14	M ₄	A ₁	1	72	18
		2	56	14			2	80	20
		3	56	14			3	80	20
		4	64	16			4	76	19
M ₂	A ₂	1	60	15	M ₄	A ₂	1	20	5
		2	44	11			2	40	10
		3	68	17			3	40	10
		4	48	12			4	52	13
M ₂	A ₃	1	96	24	M ₄	A ₃	1	100	25
		2	88	22			2	84	21
		3	96	24			3	84	21
		4	96	24			4	88	22

*Cada repetição foi constituída de n=25 sementes, totalizando n = 48 parcelas.

De forma a resumir a notação, os modelos serão apresentados com número fixo de sementes por repetição ($n = 25$), assim como a quantidade fixa de repetições ($r = 4$). Entretanto, tal balanceamento não é uma condição obrigatória para aplicação da ANOVA e tampouco dos Modelos Lineares Generalizados (MLG).

Análise de variância (ANOVA)

A análise de variância foi realizada a fim de expor os resultados obtidos da verificação tradicionalmente utilizada nos experimentos de germinação e compará-los com os resultados conseguidos a partir dos Modelos Lineares Generalizados. Para tanto, os valores de percentual de germinação seguiram o modelo definido por:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad (1)$$

$$i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, r$$

em que:

y_{ijk} : porcentagem de germinação obtida do i -ésimo método, da j -ésima amostra na k -ésima repetição;

μ : percentual médio de germinação;

α_i : efeito do i -ésimo método;

β_j : efeito do j -ésima amostra;

γ_{ij} : efeito da interação da j -ésima amostra no i -ésimo método;

ε_{ijk} : resíduo independente e normalmente distribuído;

r : número de repetições;

a : número de métodos;

b : número de amostras.

Anteriormente à aplicação da ANOVA, as pressuposições ligadas ao modelo foram verificadas. Constatou-se a pressuposição de normalidade dos resíduos pelo teste de Shapiro-Wilk (SHAPIRO; WILK, 1965); para a homogeneidade de variâncias, aplicou-se o teste de Levene (LEVENE, 1960); e, para a independência dos resíduos, a estatística d de Durbin-Watson (DURBIN; WATSON, 1950). Após a verificação do atendimento às pressuposições, os procedimentos para a análise de variância convencional foram realizados (Tabela 2).

TABELA 2. Quadro de análise de variância para o percentual de germinação de sementes de copaíba (*Copaifera langsdorffii* Desf.), planejado em delineamento inteiramente casualizado em esquema fatorial para diferentes métodos de superação de dormência e amostras.

Fontes de Variação	gl	Soma de Quadrados (SQ)	Quadrado médio (QM)	F	$F_{Crítico}$
Método	$(a - 1)$	$SQ_{método} = br \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2$	$QM_{método} = \frac{SQ_{método}}{a - 1}$	$\frac{QM_{método}}{QM_{res}}$	$F_{\alpha, a-1, ab(r-1)}$
Amostra	$(b - 1)$	$SQ_{amostra} = ar \sum_{j=1}^b (\bar{y}_{.j.} - \bar{y}_{...})^2$	$QM_{amostra} = \frac{SQ_{amostra}}{b - 1}$	$\frac{QM_{amostra}}{QM_{res}}$	$F_{\alpha, b-1, ab(r-1)}$
Interação	$(a - 1)(b - 1)$	$SQ_{interação} = r \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{...})^2 - SQ_{método} - SQ_{amostra}$	$QM_{interação} = \frac{SQ_{interação}}{(a - 1)(b - 1)}$	$\frac{QM_{interação}}{QM_{res}}$	$F_{\alpha, (a-1)(b-1), ab(r-1)}$
Resíduo	$ab(r - 1)$	$SQ_{res} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{ij.})^2$	$QM_{res} = \frac{SQ_{res}}{ab(r - 1)}$		
Total	$abr - 1$	$SQ_{total} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (y_{ijk} - \bar{y}_{...})^2$			

y: percentual de sementes germinadas; a: número de métodos; b: número de amostras; r: número de repetições; F: Estatística do teste de Snedecor.

Modelos Lineares Generalizados (MLG)

Embora não restritos às distribuições Normal e Binomial, os MLG foram aplicados a elas em função da natureza da variável, ora expressa em porcentagem, ora em número de sementes germinadas de copaíba. Por pertencerem à família exponencial paramétrica, ambas as distribuições possuem a seguinte função de densidade (FISHER, 1922):

$$f(y; \theta) = h(y) [\eta(\theta) t(y) - b(\theta)] \quad (3)$$

em que:

$\eta(\theta)$, $b(\theta)$, $t(y)$ e $h(y)$: funções que assumem valores em subconjuntos dos reais;

θ : parâmetros da distribuição.

Para estender o Modelo Linear Generalizado às distribuições Normal e Binomial com parâmetro de perturbação $\phi > 0$, a forma canônica possui a seguinte função de densidade de probabilidade (NELDER; WEDDERBURN, 1972):

$$f(y; \theta; \phi) = \exp \left\{ \phi^{-1} [y(\theta) - b(\theta)] + c(y; \phi) \right\} \quad (4)$$

em que:

$b(\cdot)$ e $c(\cdot)$: funções conhecidas das duas distribuições e apresentadas na Tabela 3.

TABELA 3. Termos da família exponencial para distribuição Normal e Binomial.

Distribuição	ϕ	θ	$b(\theta)$	$c(y; \phi)$	$\mu(\theta)$	$V(\mu)$
Normal: $N(\mu, \sigma^2)$	σ^2	μ	$\frac{\theta^2}{2}$	$-\frac{1}{2} \left[\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2) \right]$	θ	1
Binomial: $B(n, \pi); 0 < \pi < 1$	1	$\log\left(\frac{\pi}{1-\pi}\right)$	$n \log(1 + e^\theta)$	$\log\left(\frac{n}{y}\right)$	$\frac{ne^\theta}{1 + e^\theta}$	$n\pi(1 - \pi)$

μ : média; σ^2 : variância; ϕ : parâmetro de dispersão; θ : parâmetro da distribuição; y : percentual ou número de sementes germinadas de copaíba; π : para a distribuição Binomial, é a proporção de sementes germinadas, e, para a Normal, é a constante de aproximadamente 3,14; n : número de sementes por repetição – para o experimento, o valor é fixo (25 sementes).

No tocante à germinação das sementes de copaíba expressa em porcentagem (Tabela 1), a função de densidade de probabilidade da distribuição Normal foi definida por:

$$f(y; \mu; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{y - \mu}{\sigma} \right)^2 \right] \quad Y_i \sim N(\mu_i, \sigma^2) \quad (5)$$

em que:

y : percentual de germinação;

μ : percentual médio de germinação;

σ^2 : variância;

π : constante 3,14.

No que diz respeito à germinação da espécie expressa em número de sementes germinadas (Tabela 1), a função de densidade de probabilidade da distribuição Binomial foi definida por:

$$f(y; \pi) = \binom{n}{y} \pi^y (1 - \pi)^{n-y} \quad (6)$$

em que:

y : número de sementes germinadas;

n : número de sementes por repetição (fixo em 25 sementes);

π : proporção de sementes germinadas.

O componente aleatório do modelo foi o percentual de germinação para a distribuição Normal ou o número de sementes germinadas para a distribuição Binomial; enquanto isso, o componente sistemático, para ambas as distribuições, se refere aos fatores método e amostra e à interação (DEMÉTRIO; CORDEIRO, 2007). A combinação linear de seus efeitos foi dada por:

$$\eta_i = \sum_{j=1}^p x_{ij} \beta_j = x_i^T \beta \quad \text{ou} \quad \eta = X\beta \quad (7)$$

em que:

$X = (x_1, x_2, \dots, x_n)^T$: matriz do modelo composta pelos componentes explanatórios (métodos, amostras e interação);

x_i^T : i -ésima linha da matriz experimental X ;

$\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$: vetor dos parâmetros estimados pelo modelo;

$\eta = (\eta_1, \eta_2, \dots, \eta_n)^T$: preditor linear.

A associação entre os componentes sistemático e aleatório foi realizada pela função de ligação que relacionou a média da variável y (percentual ou número) ao preditor linear por meio da seguinte expressão:

$$\eta = g(\mu) \quad (8)$$

em que:

$g(.)$: função monótona e diferenciável.

O preditor linear η surge a partir do modelo como a soma dos termos de cada um dos parâmetros p , sendo obtido ao transformar a variável percentual de germinação ou número de sementes germinadas pela função de ligação identidade e logística, respectivamente. O valor esperado de y foi obtido ao aplicar o inverso da função de ligação em η :

$$\eta = \mu_p + \alpha_i + \beta_j + \gamma_{ij} \quad (9)$$

em que:

μ_p : média do preditor;

α_i : efeito do i -ésimo método;

β_j : efeito da j -ésima amostra;

γ_{ij} : efeito da interação do i -ésimo método com a j -ésima amostra.

Para a variável do percentual de germinação, a função de ligação utilizada foi a identidade, em que $g(\mu_i) = \mu_i$. Nesse modelo, foi imprescindível a checagem dos resíduos (e_i 's) quanto à independência e à normalidade, assim como da estabilidade das variâncias. O modelo foi definido como:

$$\begin{aligned} y &= X\beta + e \\ E(Y_i) &= \mu_i = x_i^T \beta \quad Y_i \sim N(\mu_i, \sigma^2) \end{aligned} \quad (10)$$

em que:

$y = (Y_1, Y_2, \dots, Y_N)$: vetor composto pelos percentuais de germinação;

$X = (x_1, x_2, \dots, x_n)^T$: matriz do modelo composta pelos componentes explanatórios (métodos, amostras e interação);

$\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$: vetor dos parâmetros estimados pelo modelo;

$e = (e_1, e_2, \dots, e_N)$: vetor composto pelos resíduos.

Em relação ao número de sementes germinadas, a função de ligação foi a logística com os dados ajustados ao modelo:

$$g(\pi_i) = x_i^T \beta \quad E(Y_i) = n\pi_i \quad (11)$$

em que:

π_i : vetor das proporções de sementes germinadas;

x_i^T : i -ésima linha da matriz experimental X ;

n : número de sementes por repetição (fixo em 25 sementes).

Análise de *deviance* (ANODEV)

No que concerne à medição do ajuste dos fatores ao modelo, foi utilizada a estatística do *log* da taxa de máxima de verossimilhança, chamada de *deviance*, nos MLG (LEE et al., 2006) – tal técnica foi aplicada para estimar os efeitos dos fatores método, amostra e interação. Nelder e Wedderburn (1972) descreveram *deviance* como:

$$S_p = 2(\hat{\ell}_n - \hat{\ell}_p) = \frac{D_p}{\phi} = \frac{2}{\phi} \sum_{i=1}^n [y_i(\tilde{\theta}_i - \hat{\theta}_i) + b(\hat{\theta}_i) - b(\tilde{\theta}_i)] \quad (12)$$

em que:

$\hat{\ell}_n$: valor máximo do logaritmo da função de verossimilhança no modelo saturado.

$\hat{\ell}_p$: valor máximo do logaritmo da função de verossimilhança no modelo corrente;

S_p : desvio escalonado;

D_p : desvio;

$\hat{\phi}$: parâmetro de dispersão estimado;

$\tilde{\theta} = q(y_i)$: estimativa de máxima verossimilhança do parâmetro canônico sob o modelo saturado.

$\hat{\theta} = q(\hat{\mu}_i)$: estimativa de máxima verossimilhança do parâmetro canônico sob o modelo corrente.

O *deviance* é estimado de acordo com a família da distribuição dos dados, simplificando a equação original (LEE et al., 2006; DOBSON; BARNETT, 2008). Para o modelo com distribuição Normal a *deviance* foi definida por:

$$\begin{aligned}
S_p &= \frac{1}{\sigma^2} \sum_{i=1}^n 2 \left[y_i \left(y_i - \hat{\mu}_i - \frac{y_i^2}{2} + \frac{\mu_i^2}{2} \right) \right] = \frac{1}{\sigma^2} \sum_{i=1}^n (2y_i^2 - 2y_i\hat{\mu}_i - y_i^2 + \mu_i^2) = \\
&= \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \hat{\mu}_i)^2 = \frac{SQ_{res}}{\sigma^2}
\end{aligned} \tag{13}$$

No que se refere à distribuição Binomial, a *deviance* foi definida por:

$$\begin{aligned}
S_p &= \sum_{i=1}^n 2 \left\{ y_i \left[\log\left(\frac{y_i}{m_i - y_i}\right) - \log\left(\frac{\hat{\mu}_i}{m_i - \hat{\mu}_i}\right) \right] \right\} + \sum_{i=1}^n 2 \left[m_i \log\left(\frac{m_i - y_i}{m_i}\right) - m_i \log\left(\frac{m_i - \hat{\mu}_i}{m_i}\right) \right] = \\
&= 2 \sum_{i=1}^n \left[y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) + (m_i - y_i) \log\left(\frac{m_i - y_i}{m_i - \hat{\mu}_i}\right) \right]
\end{aligned} \tag{14}$$

em que:

S_p : desvio escalonado;

y_i é o valor observado no fator i ;

$\mu_i = x_i^T \beta$;

$m_i = n$: tamanho da amostra (25);

i : fator sob análise (método, amostra ou interação).

Os *deviances* foram calculados conforme cada fator isoladamente, para a interação, assim como para os modelos nulo e saturado (Tabela 4). Por terem parâmetro de dispersão conhecido ($\phi = 1$), as inferências da análise de *deviance* (ANODEV) para a distribuição Binomial se basearam na estatística Qui-Quadrado (χ^2). Quando desconhecida e com estimativa consistente $\hat{\phi}$, a exemplo da distribuição Normal, a inferência foi baseada na estatística F (DEMÉTRIO; CORDEIRO, 2007; PAULA, 2010).

No modelo Binomial, compararam-se as diferenças de *deviance* aos valores críticos respectivos de cada fator e interação ($\chi_{Crítico}^2$). Para o modelo Normal, as diferenças de *deviance* foram utilizadas para o cálculo das estatísticas F , e os valores obtidos foram comparados aos valores críticos correspondentes ($F_{Crítico}$). Quando os valores do *deviance* ou do $F_{Calculado}$ foram menores que $\chi_{Crítico}^2$ ou $F_{Crítico}$, respectivamente, o efeito foi considerado não significativo.

Com vistas a comparar a qualidade de ajuste dos dois modelos executados pela ANODEV para os dados de percentuais de germinação ou o número de sementes germinadas, utilizaram-se quatro indicadores: os critérios de informação de Akaike e Bayesiano, a distância de Cook e a análise gráfica do q-q plot.

Qualidade do ajuste do modelo pelos critérios de AIC e BIC

Dentre os métodos mais precisos para avaliar a qualidade do ajuste estão o critério de informação de Akaike (AIC) (AKAIKE, 1973) e o critério de informação Bayesiano (BIC), também chamado de critério Bayesiano de Schwarz (SCHWARZ, 1978). O AIC e o BIC representaram a ausência de generalidade do modelo e penalizaram tanto a falta de ajuste aos dados quanto a alta complexidade do modelo. Por isso foram preferidos os menores valores de ambos, definidos por:

$$AIC = -2\log L + 2p \quad (15)$$

e,

$$BIC = -2\log L + p \log n \quad (16)$$

em que:

p : número de parâmetros do modelo;

L : valor da razão de máxima verossimilhança para o modelo Normal ou Binomial

n : número de observações.

TABELA 4. Modelo de análise de *deviance* (ANODEV) de um delineamento inteiramente casualizado, com dois fatores e interação, para o percentual e número de plântulas normais, além dos valores dos testes Qui-quadrado e *F*-Snedecor.

Fontes de Variação	gl	Distribuição Binomial Ligação Logística		Distribuição Normal Ligação Identidade	
		<i>Deviance</i>	$\chi^2_{Crítico}$	<i>F</i>	$F_{Crítico}$
Nulo	0	D_0	-	-	-
Método	$(a-1)$	$D_{\text{método}}$	$\chi^2_{\alpha, (a-1)}$	$\frac{D_o - D_{\text{método}}}{\frac{a-1}{\hat{\phi}}}$	$F_{\alpha, (a-1), ab(r-1)}$
Amostra	$(b-1)$	D_{amostra}	$\chi^2_{\alpha, (b-1)}$	$\frac{D_{\text{método}} - D_{\text{amostra}}}{\frac{b-1}{\hat{\phi}}}$	$F_{\alpha, (b-1), ab(r-1)}$
Método x Amostra	$(a-1)(b-1)$	$D_{\text{interação}}$	$\chi^2_{\alpha, (a-1)(b-1)}$	$\frac{D_{\text{interação}} - D_{\text{resíduo}}}{\frac{(a-1)(b-1)}{\hat{\phi}}}$	$F_{\alpha, (a-1)(b-1), ab(r-1)}$
Resíduo	$ab(r-1)$	D_{saturado}			

gl : graus de liberdade; a : número de métodos; b : número de lotes; r : número de repetições; D : *deviance*; $\hat{\phi}$: parâmetro de dispersão estimado; F : estatística da distribuição de Snedecor; $F_{Crítico}$: quantil da distribuição F de Snedecor à significância α e aos respectivos graus de liberdade; $\chi^2_{Crítico}$: quantil da distribuição Qui-quadrado à significância α e aos respectivos graus de liberdade.

Qualidade do ajuste pela distância de Cook

A distância de Cook (COOK, 1977; COOK; WEISBERG, 1982; HINES; HINES, 1995), utilizada para a detecção de observações influentes, foi realizada para verificar o modelo com maior quantidade de pontos atípicos – quanto maior o número dessas observações, menor é o ajuste do modelo aos dados de germinação. O cálculo da distância de Cook (D_i) foi a medida geral da mudança na estimativa do parâmetro, quando uma ou mais observações foram retiradas. A estatística empregou a medida de leverage (h_{ii}), o parâmetro de dispersão e a função de variância, variando de acordo com a família exponencial determinada (Binomial ou Normal).

Os valores de D_i (eixo y) foram plotados versus o valor da observação (eixo x), sendo utilizados dois critérios para classificação de pontos atípicos, de acordo com a literatura. O primeiro indicou valores de D_i próximos ou maiores que um como observações atípicas (COOK; WEISBERG, 1982; CORDEIRO; LIMA, 2006); pelo segundo critério, observações atípicas apresentaram valores maiores que quatro, dividido pelo número de observações (BOLLEN; JACKMAN, 1990) – no caso de copaíba, obteve-se 0,083.

Análise gráfica de “q-q plot”

O gráfico quantil-quantil (“q-q plot”) ou Normal de probabilidade (“Normal plots”) utilizou o componente do desvio estudatizado e os quantis observados da amostra (WEISBERG, 2005) para identificar o ajuste dos dados à distribuição e à ocorrência de valores que se destacaram no conjunto. Para garantir melhor inferência na análise visual, geraram-se os gráficos com um envelope simulado, com intervalos de 95% de confiança. Os resíduos fora dos limites ou muito próximos destes refletiram observações influentes ou um modelo com baixo ajuste para os dados.

Nesse sentido, os dados foram analisados pelo *software* estatístico R, por ser um programa livre e de fácil utilização. O script para os modelos propostos de copaíba com as respectivas orientações encontra-se no Anexo A, ao final do capítulo.

RESULTADOS

A primeira etapa da análise do conjunto de dados das sementes de copaíba seguiu o modo usual, com testes das pressuposições do modelo da ANOVA. Nesse contexto, as probabilidades de 0,424 e 0,020 dos testes de Shapiro-Wilk e Levene, respectivamente, indicaram resíduos do percentual de germinação das sementes com distribuição Normal e variâncias homogêneas, descartando a indicação de transformação (Tabela 5).

Além disso, a independência dos resíduos pressumida pela casualização do experimento foi comprovada pelo teste de Durbin-Watson, com probabilidade de 0,476. O modelo da ANOVA indicou, para os efeitos principais, apenas a significância para o fator amostra. Com a interação significativa, descartaram-se as inferências sobre os efeitos principais.

TABELA 5. Análise de variância (ANOVA) para o percentual de germinação de sementes de copaíba (*Copaifera langsdorffii* Desf.) de um modelo em delineamento inteiramente casualizado em esquema fatorial (método e amostra).

Fonte de variação	gl	ANOVA			
		<i>SQ</i>	<i>QM</i>	¹ <i>F</i>	<i>valor-p</i>
Método	3	329,00	109,67	1,79	0,166
Amostra	2	22482,67	11241,33	183,62	< 0,001
Método x Amostra	6	2440,00	406,67	6,64	< 0,001
Resíduo	36	2204,00			
Total	47	27455,67			

²*W* = 0,976 (*P*=0,424); ²*F* = 2,517 (*P*=0,018); ²*DW* = 2,63 (*P*=0,476);

gl: grau de liberdade; *SQ*: Soma de quadrados; *QM*: Quadrado médio; ¹*F*: Estatística de Snedecor; ²*W*, *F*, *DW*: Estatísticas dos testes de Shapiro-Wilk, Levene e Durbin-Watson, respectivamente; *P*: Probabilidade.

Na análise desse conjunto de dados (percentual de germinação) pelo Modelo Linear Generalizado (MLG), a decisão pelo uso da distribuição Normal e da função de ligação identidade foi tomada mediante o atendimento às pressuposições (Tabela 5). Como esperado, os valores das diferenças de *deviance* da ANODEV (Tabela 6) foram os mesmos das somas de quadrados da ANOVA (Tabela 5). Em virtude dessa igualdade, as probabilidades associadas aos efeitos principais e a interação foram idênticas (Tabela 6).

TABELA 6. Análise de desvio (ANODEV) para germinação de copaíba (*Copaifera langsdorffii* Desf.), com dois fatores e interação modelados pela distribuição Normal, com função de ligação identidade; e pela Binomial, com função de ligação logística.

Fonte de variação	<i>Gl</i>	Diferença <i>gl</i>	Modelos Lineares Generalizados (ANODEV)						
			Distribuição Normal Ligação Identidade				Distribuição Binomial Ligação Logística		
			<i>Deviance</i>	Diferença de <i>deviance</i>	<i>F</i>	<i>valor-p</i>	<i>Deviance</i>	Diferença de <i>deviance</i>	<i>valor-p</i>
Nulo	0	47	27456,0				349,10		
Método	3	44	329,0	27127,0	1,79	0,166	3,67	345,44	0,300
Amostra	2	42	22483,0	4644,0	183,62	< 0,001	275,40	70,04	< 0,001
Interação	6	36	2440,0	2204,0	6,64	< 0,001	33,01	37,03	< 0,001
Saturado	11	36	2204,0 ¹				37,03 ¹		
		AIC	345,91	$\hat{\phi}$: 61,22			205,87	ϕ : 1	
		BIC	370,23				266,95		

gl: grau de liberdade; *F*: Estatística do teste de Snedecor; ϕ : parâmetro de perturbação; $\hat{\phi}$: parâmetro de perturbação estimado.¹Também chamado de *deviance* residual.

Quando a variável foi expressa em sua forma original, em que o número de sementes germinadas segue a distribuição Binomial, com função de ligação logística, as inferências foram as mesmas. Houve efeito não significativo para método e significativo para amostra e interação, ainda que as probabilidades associadas a esses efeitos tenham sido diferentes do modelo Normal (Tabela 6).

Num primeiro momento, a igualdade das inferências pode levar ao entendimento de que os modelos têm a mesma qualidade de ajuste. A redução do AIC de 345,91, do modelo Normal para 205,87, do Binomial (Tabela 6) indicou que o segundo modelo ajustou melhor aos dados – a mesma analogia foi feita para o BIC.

Os valores associados à distância de Cook em ambos modelos abaixo de 0,25 (Figura 1) não revelaram nenhum ponto extremo, *sensu* Cook e Weisberg (1982), autores que consideram valores iguais ou superiores a um como discrepantes. No entanto, de acordo com Bollen e Jackman (1990), cinco observações no modelo Normal e quatro, no Binomial situaram-se acima de 0,083 (4/48) e foram consideradas influentes – mesmo com um limite mais restritivo, os modelos se aproximaram na quantidade de pontos influentes. Para a germinação de copaíba, a distância de Cook não foi capaz de distinguir os modelos, equiparando-os.

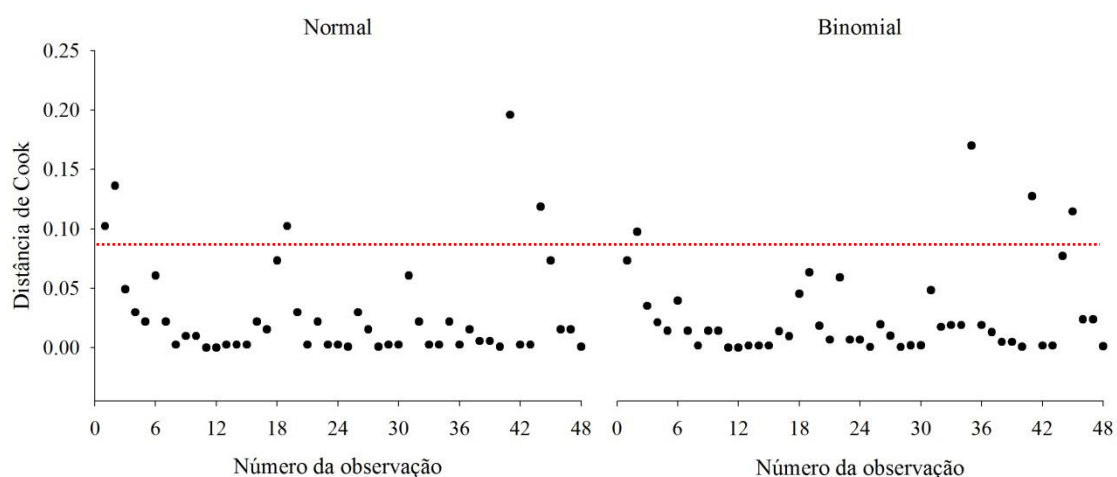


FIGURA 1. Representação de Cook para diagnóstico de dados discrepantes da germinação de sementes de copaíba (*Copaifera langsdorffii* Desf), para as distribuições Normal e Binomial. A linha tracejada representa o limite de 0,083, *sensu* Bollen e Jackman (1990).

Os gráficos de “Normal plots” permitiram complementar o que foi previamente observado por AIC e BIC e mostraram melhor ajuste dos resíduos estudentizados do modelo Binomial, que apresentaram maior linearidade e menor quantidade de

observações se distanciando da reta, tanto para as observações próximas da média (Quantil 0) quanto às mais distantes (Figura 2). Houve também maior acomodação dos valores no envelope simulado com 95% de confiança, quando comparados ao modelo Normal.

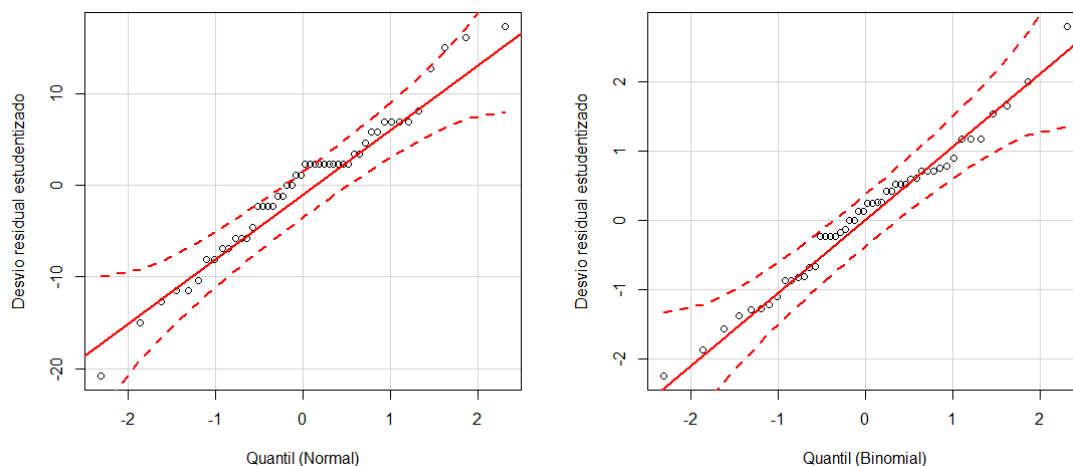


FIGURA 2. “Normal plots” com envelope simulado a 95% de confiança para diagnóstico da germinação de sementes de copaíba, em relação às distribuições Normal e Binomial.

DISCUSSÃO

A análise de variância tem sido bastante discutida por pesquisadores para dados expressos em porcentagem (BLISS, 1938; AHRENS et al., 1990), especialmente quando a escala é convertida em angular (SAKIA, 1992; SILESHI, 2007; WARTON; HUI, 2011; SILESHI, 2012), sendo comuns os relatos de problemas de interpretação de resultados na escala transformada (FERNANDEZ, 1992; SAKIA, 1992; DRUMMOND; TOM, 2011). Outro agravante é a execução da análise pelos programas estatísticos sem qualquer questionamento no que tange às pressuposições ou exigências da ANOVA, levando alguns pesquisadores a ignorarem a verificação destas (SAKIA, 1992). Entretanto, quando não atendidas, as inferências sobre a ANOVA tornam-se errôneas, o que leva a falsas conclusões e à aceitação de hipóteses falsas.

Copaíba é uma espécie emergente do dossel das matas ciliares no Distrito Federal (LEITE; SALOMÃO, 1992) e no Triângulo Mineiro (SCHIAVINI, 1997; RESENDE et al., 1997) e, por essa característica, é foco de pesquisa dos biólogos e

ecólogos. Por ser uma espécie florestal, não há como inferir que os percentuais de germinação não se ajustam ao modelo Normal. Alguns pesquisadores têm evidenciado que a não normalidade é prevalente para essas espécies (HAMPEL et al., 1986; AUSTIN, 1987; BIONDINI et al., 1988), mas essa afirmativa foi contestada.

Mesmo sendo expressos em porcentagem, os resíduos do modelo de copaíba convergiram para a distribuição Normal. Muitos autores alertam que, para dados com natureza Binomial, não há indicação do uso da ANOVA (ZHAO et al., 2001; AGRESTI, 2002; WARTON; HUI, 2011). Tal afirmação precisa ser melhor compreendida, pois, se atendidas as pressuposições, não há impedimentos para execução da ANOVA.

Ainda recai sobre o uso da ANOVA a ideia equivocada de que, se o coeficiente de variação está dentro dos limites aceitáveis, o modelo é inquestionável. É quase uma unanimidade na pesquisa científica o uso do coeficiente de variação para avaliar a qualidade do experimento (JUDICE et al., 1999; PIMENTEL-GOMES, 2000; JUDICE et al., 2002), e pesquisas acerca de valores de referência têm sido relatadas para características agrônômicas de cana (COUTO et al., 2013), soja (CARVALHO et al., 2003), milho (CARGNELUTTI FILHO; STORCK, 2007) e feijão (OLIVEIRA et al., 2009). Vale ressaltar que, quando estão baixos, os valores não têm qualquer relação com a qualidade do modelo.

A combinação de medidas descritivas e análises gráficas permite conclusões mais assertivas sobre o uso da Binomial com função de ligação logit para copaíba. Todos os métodos convergiram para essa distribuição, porém eles divergiram na forma como avaliaram o modelo. Para o BIC, os erros Tipo I e II declinam lentamente com o aumento de repetições, ao passo que, para o AIC, o erro Tipo II declina mais rapidamente, sem haver declínio do erro Tipo I (RAFTERY, 1995).

Sendo assim, para o AIC, um modelo subdimensionado (erro Tipo II) é considerado mais grave que um superdimensionado (erro Tipo I), e as chances de rejeição de qualquer hipótese falsa são essencialmente garantidas com grandes amostras, mas a taxa do erro Tipo I é constante e nunca chegará a zero. Como a redução do erro Tipo II é menor quando comparada ao AIC, há uma chance maior de se escolher um modelo subdimensionado com o BIC, mas também em selecionar um modelo parcimonioso e com menor número de parâmetros (SCHWARZ, 1978).

O BIC é valorizado devido à sua consistência, pois, à medida que a quantidade de repetições aumenta, a seleção do melhor modelo, o “modelo verdadeiro” (*true*

model) consegue chegar à precisão de 100%. Entretanto, esse resultado requer pressupostos, como o fato de que o “modelo verdadeiro” esteja sendo avaliado e que o número de parâmetros seja finito (VRIEZE, 2012). O AIC pode falhar na seleção do “modelo verdadeiro”, mesmo se ele estiver sob análise, dado que o AIC irá procurar modelos mais generalistas do que perfeitos (SHAO, 1997). Nesse entremeio, autores questionam que o modelo perfeito jamais conseguiria ser alcançado para comparação (BURNHAM; ANDERSON, 2003; MCDONALD, 2010).

A discussão sobre a criação de um “modelo verdadeiro” é tratada entre aspas em virtude das contraposições existentes entre estatísticos sobre tal fato. Alguns argumentam que, em princípio, esse modelo jamais conseguirá ser criado por completo (KOLMOGOROV, 1968; RISSANEN, 1987; MEEHL, 1990). Mas Vrieze (2012) relata que, quando um modelo é teoricamente simples, sem muitas covariáveis, e segue perfeitamente o tipo de distribuição que lhe é atribuído, o “modelo verdadeiro” poderia ser alcançado.

Nesses termos, AIC e BIC permitiram inferir que no modelo Binomial os dados foram ajustados harmonicamente, trazendo à proximidade do “modelo verdadeiro”. Apesar da convergência entre AIC e BIC no experimento de copaíba, poderá haver experimentos em que as estatísticas sejam divergentes; por conseguinte, o analista deve optar por um deles, levando em consideração as limitações existentes. Quando os modelos comparados são mais complexos e explicam diversas variáveis, com diferentes graus de interação, estudos mostraram que o AIC seria mais confiável que o BIC, e, quando os modelos são simples, o BIC se torna a preferência na escolha (NYLUND et al., 2007; YANG; YANG, 2007; VRIEZE, 2012).

De fato, os diagnósticos da distância de Cook e os “Normal plots” devem ser realizados para checagem do ajuste do modelo à função de variância (distribuição dos dados), da função de ligação ou ainda porque um ou mais valores observados são discrepantes em relação aos demais (DEMÉTRIO; CORDEIRO, 2007). Esta última situação pode ocorrer devido à existência de pontos nos extremos da amplitude da covariável, por erro de leitura ou transcrição equivocada, em razão de algum fator não controlado ter influenciado a sua obtenção.

Esses diagnósticos assumem importância quando se opta pela inserção de diferentes funções de ligação e distribuições de probabilidade para o banco de dados analisado. Para a variável germinação, a verificação da normalidade dos resíduos (distribuição Normal) e a plotagem destes como número de sementes germinadas

(distribuição Binomial) garantem confiabilidade ao analista de que as distribuições utilizadas são coerentes. Outras funções de ligação podem ser ajustadas à germinação, o que seria conferido em novos modelos; estes, por sua vez, deveriam proceder com a ANODEV da mesma forma.

Na presente pesquisa, utilizaram-se as funções de ligação canônicas de suas referidas distribuições, pois, por si só, simplificam o modelo (MYERS et al., 2002; CRAWLEY, 2007). Se os modelos colocados em teste não demonstrarem bons ajustes, e as análises gráficas evidenciarem irregularidade, modelos com novas funções de ligação precisam ser testados para a germinação, observadas as particularidades de cada um. Por exemplo, a função de ligação *probit* pode ser empregada dentro da distribuição Binomial para o ajuste da germinação (JAEGER, 2008).

Embora os resultados converjam ao modelo Binomial, a dificuldade deste se inicia com os testes pós-análise do ANODEV. Ao passo que, para a ANOVA, dados qualitativos possuem vastas pesquisas com testes cientificamente reconhecidos (Tukey, SNK, Scott-Knott, Duncan, Dunnett, entre outros), no Binomial apenas os contrastes ortogonais são aplicáveis (AGRESTI, 2002). Quando quantitativos, a regressão logística é utilizada para ajuste dos dados, o que confere uma aversão maior aos GLM, ao serem comparados ao clássico ANOVA.

Em um primeiro momento, dados na forma binária podem conferir maiores dificuldades para a análise, mas a maioria dos programas estatísticos executa todo o processo. Nesse sentido, a ANOVA continua sendo escolhida para a análise final da germinação – tal popularidade indica aderência à tradição e conforto em aplicá-la.

Sileshi (2012), ao relatar os trabalhos relacionados à germinação de 2000 a 2011, descobriu que, dos experimentos que utilizaram a ANOVA, apenas 19,5% testaram a pressuposição de normalidade dos resíduos. Isso ratifica a problemática contestada neste capítulo e demonstra que muitos dados da área de sementes estão sendo publicados sem confiabilidade. Assim, cabe ao analista se distanciar da ANOVA e buscar modelos mais parcimoniosos e concisos, mesmo que ele tenha de sair da zona de conforto.

CONCLUSÃO

O estudo permitiu demonstrar que os Modelos Lineares Generalizados podem ser aplicados em experimentos agronômicos de forma eficiente. Essa metodologia consegue enquadrar diferentes casos estatísticos e concede ao pesquisador novas inferências, por considerar outras famílias exponenciais.

As pressuposições necessárias ao modelo Normal podem ser resolvidas ao serem substituídas por uma distribuição na qual os dados consigam se ajustar, desde que também sejam atendidos os pressupostos dessa distribuição. A exemplo de *C. langsdorffi*, o MLG com distribuição Binomial e função de ligação logística se mostrou mais apropriado aos dados do que a Normal, com a função de ligação identidade.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRESTI, A. **Categorical data analysis**, 2nd ed. New York: John Wiley and Sons Press, 2002. 700 p.

AHRENS, W. H.; COX, D. J.; BUDHWAR, G. Use of the arcsine and square root transformations for subjectively determined percentage data. **Weed Science**, v.38, p.452-458, 1990.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: PETROV, B. N.; CSAKI, F. (Eds.). **Second International Symposium on Information Theory**. Budapest: Akademiai Kiado, 1973, p. 267-281.

AUSTIN, M. P. Models for the analysis of species response to environmental gradient. **Vegetation**, v.69, p.35-45, 1987.

BIONDINI, M. E.; MIELKE, P. W.; BERRY, K. J. Data-dependent permutation techniques for the analysis of ecological data. **Vegetation**, v.75, p.161-168, 1988.

BLISS, C. I. The Transformation of percentages for use in the analysis of variance. **The Ohio Journal of Science**, v.38, p.9-12, 1938.

BOLLEN, K. A.; JACKMAN, R. W. Regression diagnostics: An expository treatment of outliers and influential cases. In FOX, J.; LONG, J. S. (Eds.). **Modern Methods of Data Analysis**. Newbury Park: Sage, 1990, p. 257-291.

BURNHAM, K.; ANDERSON, D. **Model selection and multimodel inference: A practical-theoretic approach**. New York: SpringerVerlag, 2003. 488 p.

CARVALHO, C. G. P.; ARIAS, C. A. A.; TOLEDO, J. F. F.; ALMEIDA, L. A. A.; KIIHL, R. A. S.; OLIVEIRA, M. F.; HIROMOTO, D. M.; TAKEDA, C. Proposta de classificação dos coeficientes de variação em relação a produtividade e altura da planta de soja. **Pesquisa Agropecuária Brasileira**, v.38, p.187-193, 2003.

COOK, R. D. Detection of influential observations in linear regression, **Technometrics**, v.19, p.15-18, 1977.

COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. New York: Chapman & Hall, 1982. 229 p.

CORDEIRO, G. M. **Modelos lineares generalizados**. Campinas: 8º Simpósio Nacional de Probabilidade e Estatística, 1986. 286 p.

CORDEIRO, G. M.; LIMA, E. A. **Modelos Paramétricos**. Recife: Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, 2006. 248 p.

COUTO, M. F.; PETERNELLI, L. A.; BARBOSA, M. H. Classification of the coefficients of variation for sugarcane crops. **Ciência Rural**, Santa Maria, v.43, p.957-961, 2013.

CRAWLEY, M. J. **The R book**. England: Wiley, 2007. 942 p.

DEMÉTRIO, C. G. B.; CORDEIRO, G. M. Modelos lineares generalizados. In: SEAGRO 12 E A REUNIÃO ANUAL DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, v. 52, 2007, Santa Maria. **Anais...** Santa Maria: UFSM, 2007. 161 p.

DOBSON, A. J.; BARNETT, A. G. **An introduction to Generalized Linear Models**, 3rd ed. New York: Chapman & Hall, 2008. 307 p.

DRUMMOND, G.B.; TOM, B.D.M. Presenting data: can you follow a recipe? **Advances in Physiology Education**, v.35, p.365-368, 2011.

DURBIN, J.; WATSON, G. S. Testing for serial correlation in least squares regression, I. **Biometrika**, v.37, p.409-428, 1950.

FARAWAY, J. J. **Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models**. Florida: Chapman and Hall, 2006. 301 p.

FERNANDEZ, G. C. J. Residual analysis and data transformations: Important tools in statistical analysis. **HortScience**, v.27, p.297-300, 1992.

CARGNELUTTI FILHO, A.; STORCK, L. Estatísticas de avaliação da precisão experimental em ensaios de cultivares de milho. **Pesquisa Agropecuária Brasileira**, v.42, p.17-24, 2007.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society**, v.222, p.309-368, 1922.

FISHER, R. A. **Statistical methods for research workers**. Edinburgh: Oliver and Boyd, 1925. 319 p.

HAMPEL, F. R.; RONCHETTI, E. M.; BOUSSEEUW, P. J.; STAHEL, W. A. **Robust statistics: the approach based on influence functions**. New York: Wiley, 1986. 502 p.

HINES, R. J. O.; HINES, W. G. S. Exploring Cook's statistic graphically. **The American Statistician**, v.49, n. 4, p.389-394, 1995.

JAEGER, T.F. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. **Journal of Memory and Language**, v.59, p.434-446, 2008.

JUDICE, M. G.; MUNIZ, J. A.; AQUINO, L. H.; BEARZOTI, E. Avaliação da precisão experimental em ensaios com bovinos de corte. **Ciência e Agrotecnologia**, Lavras, v.26, n.5, p.1035-1040, 2002.

JUDICE, M. G.; MUNIZ, J. A.; CARVALHEIRO, R. Avaliação do coeficiente de variação na experimentação com suínos. **Ciência e Agrotecnologia**, Lavras, v.23, n.1, p.170-173, 1999.

KOLMOGOROV, A. Three approaches to the quantitative definition of information. International **Journal of Computer Mathematics**, v.2, p.157-168, 1968.

LARSON, R.; FARBER, B. **Estatística Aplicada**, 4^a ed. São Paulo: Pearson Education, 2010. 476 p.

LEE, Y.; NELDER, J. A.; PAWITAN, Y. **Generalized Linear Models with random effects**. New York: Chapman & Hall, 2006. 380 p.

LEITE, A. M. C.; SALOMÃO, A. N. Estrutura populacional de copaíba (*Copaifera langsdorffii* Desf.) em mata ciliar do Distrito Federal. **Acta Botanica Brasilica**, São Paulo, v.6, p.123-134, 1992.

LEVENE, H. Contributions to probability and statistics. Essays in honor of Harold Hotelling. In: OLKIN, I.; CHURYES, S. G.; HEOFFDING, W.; MADOW, W. G.; MANN, H. B. **Stanford studies in mathematics and statistics**. California: Stanford University Press, p.278-292, 1960.

LINDSEY, J. K. **Applying Generalized Linear Models**. New York: Springer, 1997. 256 p.

MCCULLAGH, P.; NELDER, J. A. **Generalized Linear Models**, 2nd ed. London: Chapman & Hall, 1989. 493 p.

MCDONALD, R. P. Structural models and the art of approximation. **Perspectives on Psychological Science**, v.5, p.675-686, 2010.

MEEHL, P. E. Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. **Psychological Inquiry**, v.1, p.108-141, 1990.

MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G. **Generalized Linear Models, with applications in engineering and the sciences**. New York: John Wiley and Sons Press, 2002. 342 p.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**, v.135, p.370-384, 1972.

NETER, J.; KUTNER, M.; NACHTSHEIM, C.; WASSERMAN, W. **Applied linear statistical models**, 4th ed. New York: McGraw-Hill Companies, 1996. 1408 p.

NYLUND, K. L.; ASPAROUHOV, T.; MUTHÉN, B. O. Deciding on the number of classes in latente class analysis and growth mixture modeling: A Monte Carlo simulation study. **Structural Equation Modeling**, v.14, p.535-569, 2007.

OLIVEIRA, R. L.; MUNIZ, J. A.; ANDRADE, M. J. B.; REIS, R. L. Precisão experimental em ensaios com a cultura do feijão. **Ciência e Agrotecnologia**, Lavras, v.33, p.113-119, 2009.

PAULA, G. A. **Modelos de regressão com apoio computacional**. São Paulo: IME-USP, 2010. 245 p.

PIMENTEL-GOMES, F. **Curso de estatística experimental**, 14^a ed. Piracicaba: Nobel, 2000. 477 p.

RAFTERY, A. E. Bayesian model selection in social research. **Sociological Methoology**, v.25, p.111-163, 1995.

RESENDE, J. C., SCHIAVINI, I.; KLINK, C. A. Ecological aspects of a *Copaifera langsdorffii* population in a gallery forest of central Brazil. In: ENCINAS, J.I.; KLEINN, C. (Eds). **Proceedings of the International Symposium on Assessment and Monitoring of Forest in Tropical Dry Regions with Special Reference to Gallery Forest**. Brasília: Universidade de Brasília, 1997, p.147-156.

RESENDE, M. D. V.; BIELE, J. Estimção e predição em modelos lineares generalizados mistos com variáveis binomiais. **Revista de Matemática e Estatística**, São Paulo, v.20, p.39-65, 2002.

RISSANEN, J. Stochastic complexity. **Journal of the Royal Statistical Society**, v.9, p.223-239, 1987.

SAKIA, R. M. The Box-Cox transformation technique: a review. **Journal of the Royal Statistical Society**, v.41, p169-178, 1992.

SCHEFFÉ H. **The analysis of variance**. New York: Wiley, 1959. 555 p.

SCHIAVINI, I. Environmental characterization and groups of species in gallery forest. In: ENCINAS, J.I.; KLEINN, C. (Eds). **Proceedings of the International Symposium on Assessment and Monitoring of Forest in Tropical Dry Regions with Special Reference to Gallery Forest**. Brasília: Universidade de Brasília, 1997, p.107-113.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v.6, n.2, p.461-464, 1978.

SHAO, J. An asymptotic theory for linear model selection (with discussion). **Statistica Sinica**, v.7, p.221-242, 1997.

SHAPIRO, S. S.; WILK, M. B. An Analysis of variance test for normality. **Biometrika**, v.52, p.591-611, 1965.

SOKAL, R. R.; ROHLF, F. J. **Biometry**: the principles and practice of statistics in biological research. 3rd ed. New York: W. H. Freeman, 1995. 937 p.

SILESHI, G. W. A critique of current trends in the statistical analysis of seed germination and viability data. **Seed Science Research**, v.22, p.145-159, 2012.

SILESHI, G. W. Evaluation of statistical procedures for efficient analysis of insect, disease and weed abundance and incidence data. **East African Journal of Science**, v.1, p.1-9, 2007.

STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics**. New York: McGraw-Hill, 1980. 633 p.

VRIEZE, S. I. Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). **Psychological Methods**, v.17, p.228-243, 2012.

WARTON, D.; HUI, F. The arcsine is asinine: the analysis of proportions in ecology. **Ecology**, v.92, p.3-10, 2011.

WEISBERG, S. **Applied linear regression**, 3rd ed. New York: John Wiley, 2005. 336 p.

WILSON, K.; HARDY, I. C .W. Statistical analysis of sex ratios: an introduction. In: HARDY, I. C. W. **Sex ratios**: Concepts and Research Methods. United Kingdom: Cambridge University Press, p.49-92, 2002.

YANG, C.; YANG, C. Separating latent classes by information criteria. **Journal of Classification**, v.24, p.183-203, 2007.

ZHAO, L.; CHEN, Y.; SCHAFFNER, D. W. Comparison of logistic regression and linear regression in modeling percentage data. **Applied and Environmental Microbiology**, v.67, p.2129-2135, 2001.

ANEXOS

Anexo A. Script comentado das funções do programa R para análise da germinação de copaíba.

```
###Estes são os pacotes exigidos para a execução das análises e que
devem ser instalados antes###
require(car)
require(designGG)
require(base)
require(fBasics)
require(nortest)
require(stats)
require(boot)

#####IMPORTAÇÃO DOS DADOS#####
###Os dados serão lidos de uma planilha escrita no bloco de dados, e
será atribuído quais colunas são fatores e quais são variáveis###
dados<-read.table("copaiba.txt", header=T)
MET<-as.factor(dados[,1])
AM<-as.factor(dados[,2])
REP<-as.factor(dados[,3])
PN<-dados[,4]
PNC<-dados[,5]
PNC2<-dados[,6]
dados.frame<-data.frame(MET, AM, REP, PN, PNC, PNC2)
attach(dados.frame)
resp<-cbind(PNC, PNC2)
dados

###MET corresponde aos métodos, AM as amostras, REP as repetições, PN a
porcentagem de plântulas normais, PNC número de sementes germinadas e
PNC2 número de sementes não germinadas###
###A linha dados irá mostrar os dados que serão analisados. Veja
parte do exemplo de copaíba:
```

```
###MET AM REP PN PNC PNC2###
###M1 A1 1 56 14 11###
###M1 A1 2 84 21 4###
###M1 A1 3 60 15 10###
### ... ###
###M4 A3 2 84 21 4###
###M4 A3 3 84 21 4###
###M4 A3 4 88 22 3###
```

```
#####PRESSUPOSIÇÕES#####
shapiro.test(residuals(lm(PN~MET+AM+MET*AM),data=dados.frame))
durbinWatsonTest(lm(PN~MET+AM+MET*AM),data=dados.frame)
leveneTest(PN,MET:AM, center=mean)
#####ANODEV NORMAL#####
mod1<-glm(PN~MET +AM + MET*AM, family=gaussian)
summary(mod1)
coef(mod1)
anova(mod1, test="F")
#####ANODEV BINOMIAL#####
mod2<- glm(resp~MET +AM + MET*AM, family=binomial)
summary(mod2)
coef(mod2)
anova(mod2, test="Chisq")
#####GRÁFICOS PARA O MODELO NORMAL#####
glm1<-glm(formula=PN~MET+AM+MET*AM,family=gaussian,maxit=50)
c<-cooks.distance(glm1)
plot(c,xlab="Índice", ylab="Distância de Cook")

h<-influence(glm1)$hat
rd<-resid(glm1,type="deviance")
td<-rd*sqrt(1/(1-h))
qqPlot(td,envelope=0.95, xlab="Quantil (Normal)", ylab="Desvio
residual estudentizado")
#####GRÁFICOS PARA O MODELO BINOMIAL#####
glm2<-glm(formula=resp~MET+AM+MET*AM,family=binomial,maxit=50)
c2<-cooks.distance(glm2)
plot(c2,xlab="Índice", ylab="Distância de Cook")

h2<-influence(glm2)$hat
rd2<-resid(glm2,type="deviance")
td2<-rd2*sqrt(1/(1-h2))
qqPlot(td2,envelope=0.95, xlab="Quantil (Binomial)", ylab="Desvio
residual estudentizado")
```

CAPÍTULO III

POR QUE ANALISAR EXPERIMENTOS DE GERMINAÇÃO UTILIZANDO OS MODELOS LINEARES GENERALIZADOS?

Fábio Janoni Carvalho

Resumo: Experimentos de germinação de sementes são constantemente analisados por meio da ANOVA, mas também se deparam com o não atendimento às pressuposições que, ao serem violadas, tem comprometida a confiabilidade de todos os testes paramétricos. Para solucionar essa problemática, alguns autores sugerem a transformação angular dos dados, ao passo que outros retratam o uso dessa técnica sem nenhum cuidado. Outra alternativa sugerida, de menor impacto aos dados, é a aplicação de metodologias estatísticas que não necessitam do atendimento a essas pressuposições. Dentre as metodologias existentes se destacam os Modelos Lineares Generalizados (MLG). Apesar da comum representação do número de sementes germinadas em porcentagem, a natureza original do dado é discreta e segue todos os critérios da distribuição Binomial. Assim, os MLG surgem como uma alternativa para solucionar as restrições da ANOVA e trazem diferentes técnicas estatísticas, permitindo melhor tratamento aos dados. Na área agrônômica, os MLG ainda são pouco conhecidos, sendo que, para a análise de sementes, não há trabalhos que investiguem a aplicabilidade e o ajuste dessa técnica, em comparação à ANOVA e à transformação de dados. Nesse sentido, o objetivo deste trabalho foi comparar a metodologia dos MLG com a da ANOVA, verificando os impactos ocasionados por elas dentro da variável de germinação de sementes. Visou-se, ainda, aplicar a transformação de dados e compará-la aos MLG, constatando a melhor alternativa para os dados estudados. As análises estatísticas se concentraram na característica das plântulas normais obtidas a partir do processo de validação de métodos para teste de germinação de sementes de 50 espécies florestais. A ANOVA faz parte dos MLG, e sua incorporação foi realizada assumindo a distribuição Normal do componente aleatório e a função de ligação identidade. O número de plântulas normais formadas seguiu a distribuição Binomial, correspondendo ao evento de sucesso com a função de ligação logística nesse MLG. Apenas 41% das espécies com pressuposições atendidas e 22% daquelas não atendidas obtiveram a mesma interpretação sobre o efeito dos fatores, o que demonstra que a mudança da análise dentro dos MLG foi radical até para espécies com pressuposições atendidas. Os registros dos valores de AIC permitem concluir que o modelo Binomial com função *logit* foi mais harmônico para o ajuste de dados e tem um menor número de parâmetros na explicação da variação, o que o tornou um modelo mais parcimonioso. Os gráficos “Normal plots” aludiram a uma linearidade melhor dos resíduos provenientes dos dados da distribuição Binomial. A transformação angular foi capaz de corrigir os problemas no atendimento às pressuposições por completo em apenas dez espécies, em relação às 23 estudadas. Isso comprova que a aplicação do MLG com distribuição Binomial de imediato foi indispensável para 13 delas.

Palavras-chave: Binomial, critério de informação de Akaike (AIC), q-q plots, validação, pressuposições da ANOVA, MLG

WHY ANALYZE GERMINATION EXPERIMENTS USING GENERALIZED LINEAR MODELS?

Fábio Janoni Carvalho

Abstract: Seed germination experiments are constantly analyzed using ANOVA, but it is also faced the problem of not holding the assumptions; when these ones are violated, the reliability of all parametric tests is compromised. To solve this problem, some authors suggest angular transformation of the data, as in many other cases the use of this technique with no care. Another suggested alternative, with less impact to the data, is the application of statistics methodologies that do not need to answer these assumptions. Among the existing methodologies, Generalized Linear Models (GLM) stands out. Despite the common representation of the number of germinated seeds in percentage, the original nature of data is discrete and follows all the criteria of Binomial distribution. Thus, GLM emerge as an alternative to solve ANOVA restrictions and to bring different statistical techniques, allowing a better data processing. GLM are poorly known in agronomy, and there are not works to the seed analysis that investigate the applicability and the adjustment of this technique, comparing to ANOVA and data transformation. In this way, the objective of this study was to compare the GLM methodology with ANOVA by checking the impact caused by them within seed germination variable. It was also aimed to apply the data transformation and compares it to GLM, checking which the best one for the studied data is. Statistical analysis focused on the characteristic of normal seedlings obtained from the process of validation of methods for germination test of 50 forest species seeds. ANOVA is a part of GLM, and its incorporation was made assuming the Normal distribution of random component and the identity link function. The number of normal seedlings followed a Binomial distribution, corresponding to the event of success with a logistic link function for this GLM. Only 41% of species that hold the assumptions and 22% of those which did not had the same interpretation about the effects of the factors, which proves that the analysis change within GLM was radical even for species that attended the assumptions. Registrations of AIC can conclude that the Binomial model with logit function was more harmonious for the data set and have fewer parameters to explain the variation, which made it a more parsimonious model. “Normal plots” graphics allude to a better linearity of the residuals from Binomial distribution data. The angular transformation was able to correct the problems in a completely meeting the assumptions in only ten species, in relation to the 23 that were studied. It proves that the application of GLM with an immediately Binomial distribution was essential for 13 of them.

Keywords: binomial, Akaike’s information criterion (AIC), q-q plots, validation, ANOVA assumptions, MLG

INTRODUÇÃO

Popularizada entre os experimentos agronômicos, a análise de variância (ANOVA) se tornou instrumento fundamental para a validação estatística de uma pesquisa frente à comunidade científica. Diante de diferentes conjuntos de dados, principalmente daqueles que não obedeciam à distribuição Normal, foi imprescindível a criação de novas metodologias estatísticas. Como exemplos disso, há o modelo de complemento log-log para ensaios de diluição (FISHER, 1922), os modelos probit para dados binomiais (BLISS, 1935) e os modelos log-lineares para contagens com distribuição de Poisson (BIRCH, 1963). Alguns deles conseguiram se popularizar na área agrícola, como os modelos probit, para análise de sobrevivência de insetos – eles são também necessários para o cálculo de doses letais (BLISS, 1935).

Apesar do constante alerta sobre o atendimento às pressuposições para a execução da ANOVA desde a sua elaboração (FISHER, 1925), a verificação desse aspecto só se tornou rotineira na última década, principalmente em razão das exigências da comunidade científica. As pressuposições para modelos de delineamento experimentais incluem normalidade e independência dos resíduos, variâncias homogêneas e, especificamente para delineamentos de blocos casualizados, efeitos aditivos entre tratamentos e blocos (STEEL; TORRIE, 1980; SOKAL; ROHLF, 1995). É bastante comum o não atendimento de uma ou mais pressuposições, podendo tornar errôneas as inferências a partir da análise de variância (KHAN; RAYNER, 2003; SILESHI, 2012).

Quando as pressuposições são violadas, a confiabilidade de todos os testes paramétricos fica comprometida, dadas as alterações na probabilidade de ocorrência dos erros Tipo I (rejeição da hipótese de nulidade quando ela é verdadeira) e Tipo II (não rejeição da hipótese de nulidade quando ela é falsa). Isso leva a falsas conclusões a respeito dos efeitos do modelo (STEEL; TORRIE, 1980; STORCK et al., 2011; LÚCIO et al., 2012).

Para solucionar esse problema, alguns autores sugerem a transformação angular dos dados (LITTLE; HILLS, 1978; ZAR, 1998; AHRENS et al., 1990), ao passo que outros alertam para o perigo do seu uso sem cuidados (WARTON; HUI, 2011; LÚCIO et al., 2012; SILESHI, 2012). De fato, a transformação dos dados pode causar dificuldade na interpretação dos resultados quando duas médias iguais nos dados originais se tornam diferentes. Além disso, para uma determinada média que antes era

maior do que a outra, pode ocorrer a inversão dessa relação (AHRENS et al., 1990). O maior agravante da transformação é que ela pode corrigir uma das pressuposições e violar outra atendida na escala original (SILESHI, 2012).

Outra alternativa sugerida para os casos com quebra de pressuposições é a aplicação de metodologias estatísticas flexíveis que se adequem melhor aos dados do experimento; com isso, pode-se evitar os problemas resultantes da transformação (WILSON; HARDY, 2002; CRAWLEY, 2007; WARTON; HUI, 2011). Dentre as metodologias existentes, destacam-se os Modelos Lineares Generalizados (MLG).

Os MLG surgiram em 1972, quando Nelder e Wedderburn criaram um modelo capaz de absorver diversos tipos de distribuição da família exponencial. Eles englobam diversas técnicas estatísticas, tais como a ANOVA, a regressão logística, a regressão de Poisson e a análise *probit*. Além da possibilidade de uso de diferentes distribuições do componente aleatório do modelo (a variável em análise), introduz-se a função de ligação ao modelo, vinculando a média da variável y ao preditor linear (CRAWLEY, 2007).

Experimentos de germinação de sementes são constantemente analisados por meio da ANOVA, mas eles também se deparam com a problemática do não atendimento às pressuposições. Apesar da comum representação do número de sementes germinadas em porcentagem, a natureza original do dado é discreta e segue todos os critérios da distribuição Binomial (NETER et al., 1996; LEE et al., 2006). Com essa distribuição, os dados podem receber outros tratamentos estatísticos, como a regressão logística, sem alteração de comportamento e permitindo melhor ajuste (DOBSON; BARNETT, 2008).

Assim, os MLG surgem como uma alternativa para solucionar as restrições da ANOVA e trazem diferentes técnicas estatísticas, o que permite um melhor tratamento dos dados. Na área agronômica eles ainda são pouco conhecidos, sendo que para a análise de sementes não há trabalhos que investiguem a aplicabilidade e o ajuste dos MLG, em comparação à ANOVA e à transformação de dados.

Vale destacar que o objetivo deste trabalho foi comparar a metodologia dos MLG com a dos modelos lineares, verificando os impactos ocasionados por elas dentro da variável de germinação de sementes. Tencionou-se, ainda, aplicar a transformação de dados e compará-la aos MLG, verificando a qualidade do ajuste.

MATERIAL E MÉTODOS

Planejamento e execução

As análises estatísticas se concentraram na característica das plântulas normais, obtida a partir do processo de validação de métodos para teste de germinação de sementes de 50 espécies florestais brasileiras (Tabela 1) oficializadas pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA) (BRASIL, 2010; 2011; 2012).

Nesse processo, lotes com, no mínimo, três qualidades distintas de cada uma das espécies foram enviados para, no mínimo, seis laboratórios credenciados pelo MAPA. Cada laboratório executou o teste de germinação com quatro repetições de 50 sementes, totalizando 200 sementes por lote. Apenas dois laboratórios receberam tamanhos de amostras duplicadas e realizaram quatro repetições de 100 sementes, totalizando 400 sementes para cada lote (valor verdadeiro). Após a execução do teste de germinação, as fichas com os resultados de plântulas normais por lote, laboratório e espécie foram encaminhadas.

Cada espécie apresentou metodologia específica a ser executada na montagem do teste (Tabela 1), com o intuito de viabilizar o potencial germinativo da semente. Essa metodologia, por sua vez, foi enviada a cada laboratório na forma de protocolo, juntamente com uma série de imagens para a distinção das plântulas normais de anormais. Dentro desse protocolo havia também o arranjo do experimento nas câmaras de germinação (BOD), que seguia um delineamento inteiramente casualizado (DIC).

A característica de plântulas normais foi descrita como uma variável representada pelo número de plântulas (natureza discreta) e percentual (natureza contínua) para a entrada no MLG, com vistas a validar as metodologias de cada espécie. Para tanto, as análises necessitaram demonstrar diferença significativa apenas entre os lotes, dada a diferença intrínseca em seus potenciais germinativos. Para confirmar a validação do método, os laboratórios deveriam ser estatisticamente iguais; os lotes, distintos (formados *a priori* com qualidades fisiológicas distintas), e a interação, não significativa – isso mostra que a metodologia é fixa e não varia de acordo com o laboratório que a executa.

TABELA 1. Relação de métodos para germinação das 50 espécies florestais incluindo substrato, modo/recipiente, temperatura, regime de luz, métodos pré-germinativos e contagens.

Espécie	Substrato	T (°C)/ Regime de luz	Métodos pré-germinativos	Contagens (dias)		
				1 ^a	2 ^a	Final
<i>Acacia polyphylla</i> DC.	RP	25 °C/contínua	Solução de detergente ¹	7	-	14
<i>Albizia hassleri</i> (Chodat) Burkart	RP	25 °C/ contínua	0,025% de NaClO + desponte na lateral/ terço superior + 0,025% de NaClO	7	-	14
<i>Anadenanthera colubrina</i> (Vell.) Brenan	RP (mais seco)	25 °C/ contínua	Solução de detergente ¹	4	-	10
<i>Anadenanthera macrocarpa</i> (Benth.) Brenan	RP (mais seco)	25 °C/ contínua	Solução de detergente ¹	4	-	10
<i>Apuleia leiocarpa</i> (Vogel) J.F. Macbr.	RP	25 °C/contínua	0,25% de NaClO + desponte na lateral/ terço superior + 0,025% de NaClO	7	-	10
<i>Astronium fraxinifolium</i> Schott ex Spreng.	RP	25 °C/ contínua	Solução de detergente ¹	7	-	10
<i>Cariniana estrellensis</i> (Raddi) Kuntze	RP	25 °C/contínua	Lavagem em água corrente + solução de detergente + 0,025% de NaClO	14	-	28
<i>Cariniana legalis</i> (Mart.) Kuntze	RP	25 °C/contínua	Lavagem em água corrente + solução de detergente + 0,025% de NaClO	14	-	28
<i>Cassia leptophylla</i> Vogel	RP	25 °C/ contínua	Solução de detergente ¹ + 0,05% de NaClO + desponte na lateral/ terço mediano + 0,05% de NaClO + embebição por 24 h	10	-	14
<i>Cedrela fissilis</i> Vell.	RP	25 °C/ contínua	0,025% de NaClO	14	-	21
<i>Cedrela odorata</i> L.	RP	25 °C/ contínua	0,025% de NaClO	14	21	28
<i>Ceiba speciosa</i> (A. St.-Hil.) Ravenna	RP	25 °C/ contínua	0,025% de NaClO	7	-	10
<i>Citharexylum myrianthum</i> Cham.	RP	25 °C/ contínua	Solução de detergente ¹ + 0,25% de NaClO	21	-	35
<i>Copaifera langsdorffii</i> Desf.	RP	25 °C/ contínua	0,05% de NaClO + escarificação na lateral/ terço superior + 0,05% de NaClO	14	21	28
<i>Cordia americana</i> (L.) Gottschling & J.S. Mill	SA	25 °C/ contínua	Solução de detergente ¹	10	-	17
<i>Cybistax antisiphilitica</i> (Mart.) Mart.	RP	25 °C/ contínua	Solução de detergente ¹	14	-	35
<i>Dalbergia miscolobium</i> Benth.	RP	25 °C/ contínua	Solução de detergente ¹ + 0,025% de NaClO	7	-	10
<i>Dalbergia nigra</i> (Vell.) Allemão ex Benth.	RP	25 °C/ contínua	Solução de detergente ¹ + 0,05% de NaClO	10	-	14
<i>Enterolobium contortisiliquum</i> (Vell.) Morong	RP	25 °C/ contínua	Escarificação na extremidade oposta à micrópila + solução de detergente ¹	7	-	14

“Continua...”

“Continuação...”

<i>Enterolobium maximum</i> Ducke	RP	25 °C/ contínua	0,25% de NaClO + escarificação na lateral/ terço superior + 0,125% de NaClO + tratamento térmico a 70 °C	7	-	14
<i>Erythrina speciosa</i> Andrews	RP	25 °C/ contínua	0,5% de NaClO + desponte na transversal do hilo + 0,05% de NaClO	7	-	14
<i>Gallesia integrifolia</i> (Spreng.) Harms	RP (mais seco)	25 °C/ contínua	Retirar o cálice + solução de detergente ¹	10	-	17
<i>Guazuma ulmifolia</i> Lam.	SP	25 °C/ contínua	Água a 90 °C + solução de detergente ¹	7	14	21
<i>Hymenaea courbaril</i> L.	RP	25 °C/ contínua	0,025% de NaClO + escarificação na região oposta à micrópila + 0,025% de NaClO + embebição por 48 h + solução de detergente ¹	21	-	28
<i>Hymenaea stigonocarpa</i> Mart. ex Hayne	RP	25 °C/ contínua	0,025% de NaClO + escarificação na região oposta à micrópila + 0,025% de NaClO + embebição por 48 h + solução de detergente ¹	21	-	28
<i>Jacaranda cuspidifolia</i> Mart.	RP	25 °C/ contínua	Solução de detergente ¹	21	-	28
<i>Jacaranda micranta</i> Cham.	RP (mais seco)	25 °C/ contínua	Solução de detergente ¹	21	28	42
<i>Lafoensia pacari</i> A. St.-Hil.	RP	25 °C/ contínua	Solução de detergente ¹	14	-	21
<i>Mimosa caesalpiniaefolia</i> Benth.	RP	25 °C/ contínua	Solução de detergente + desponte no terço superior/ região oposta à micrópila + solução de detergente	5	-	10
<i>Mimosa scabrella</i> Benth.	RP	25 °C/ contínua	0,05% de NaClO + tratamento térmico à 80 °C + embebição por 24 h + 0,05% de NaClO	5	-	10
<i>Ormosia arborea</i> (Vell.) Harms	RP	25 °C/ contínua	0,05% de NaClO + escarificação na lateral/ terço superior porção vermelha + 0,05% de NaClO + embebição por 24 h	21	-	28
<i>Parapiptadenia rigida</i> (Benth.) Brenan	RP	25 °C/ contínua	Solução de detergente ¹	7	-	14
<i>Parkia pendula</i> (Willd.) Benth. ex Walp.	RP	30 °C/ contínua	0,05% de NaClO + desponte na lateral/ terço superior + 0,05% de NaClO	7	-	14
<i>Peltogyne confertiflora</i> (Mart. ex Hayne) Benth.	RP	25 °C/ contínua	0,05% de NaClO + embebição por 24 h	21	-	28

“Continua...”

“Continuação...”

<i>Peltophorum dubium</i> (Spreng.) Taub.	RP	25 °C/ contínua	Desponte na região oposta à micrópila + solução de detergente ¹	7	-	14
<i>Plathymenia reticulata</i> Benth.	RP	25 °C/ contínua	0,5% de NaClO + desponte na lateral/ terço superior + 0,025% de NaClO	10	-	16
<i>Pseudobombax tomentosum</i> (Mart. & Zucc.) Robyns	RP	25 °C/ contínua	0,0125% de NaClO	10	-	17
<i>Pterogyne nitens</i> Tul.	RP	25 °C/ contínua	0,025% de NaClO + desponte na lateral/ terço superior + solução de detergente ¹	7	-	14
<i>Qualea grandiflora</i> Mart.	RP	25 °C/ contínua	Solução de detergente ¹	28	-	35
<i>Schinus terebinthifolius</i> Raddi	SP	25 °C/ contínua	Solução de detergente ¹	10	-	18
<i>Schizolobium parahyba</i> var. <i>amazonicum</i> (Huber ex Ducke) Barneby	RP	25 °C/ contínua	0,025% de NaClO + escarificação na região oposta a micrópila + 0,025% de NaClO + embebição por 24 h + solução de detergente	7	-	10
<i>Senna macranthera</i> (DC. ex Collad.) H.S. Irwin & Barneby	RP	25 °C/ contínua	0,05% de NaClO + desponte na lateral/ terço superior + 0,05% de NaClO	7	-	14
<i>Senna multijuga</i> (Rich.) H.S. Irwin & Barneby	RP	25 °C/ contínua	0,05% de NaClO + desponte na região oposta à micrópila + 0,05% de NaClO	4	-	7
<i>Stryphnodendron barbadetiman</i> (Vell.) Mart.	RP	25 °C/ contínua	0,125% de NaClO + desponte na lateral/ terço mediano + 0,125% de NaClO	10	-	14
<i>Stryphnodendron polyphyllum</i> Mart.	RP	25 °C/ contínua	0,05% de NaClO + desponte na lateral/ terço mediano + 0,05% de NaClO	10	-	14
<i>Tabebuia aurea</i> (Silva Manso) Benth. & Hook. f. ex S. Moore	RP	25 °C/ contínua	Solução de detergente ¹	10	15	21
<i>Tabebuia chrysotricha</i> (Mart. ex A. DC.) Standl.	RP	25 °C/ contínua	Solução de detergente ¹	7	-	14
<i>Tabebuia impetiginosa</i> (Mart. ex DC.) Standl.	RP	25 °C/ contínua	0,025% de NaClO	14	-	21
<i>Tabebuia roseoalba</i> (Ridl.) Sandwith	RP	25 °C/ contínua	Solução de detergente ¹	10	-	17
<i>Zeyheria tuberculosa</i> (Vell.) Bureau	RP	25 °C/ contínua	0,05% de NaClO	14	-	21

¹ Solução de detergente: proporção de 1 mL detergente neutro para 100 mL de água destilada, com permanência das sementes ou dos diásporos por 5 a 10 minutos na solução, seguida de lavagem em água corrente e permanência em água destilada por 3 minutos.

Retirada de *outliers* pelo método de Box-Plot

Antes da análise foi constatada a existência de *outliers* por lote, isto é, de valores de percentuais de plântulas normais discrepantes, sempre na escala percentual, pelo Box-Plot. Embora ela não seja uma pressuposição da análise de variância, é prescrita no Manual de Validação da Associação Internacional para Análise de Sementes (ISTA, 2007). Após a avaliação e retirada, foi feita uma nova representação gráfica para verificar se algum *outlier* apareceu. Caso a retirada desses novos *outliers* implicasse em novos valores na terceira representação gráfica do Box-Plot, haveria a remoção apenas dos *outliers* extremos.

Pressuposições de normalidade dos resíduos e homogeneidade das variâncias

Para a pressuposição de normalidade dos resíduos, executou-se o teste de Kolmogorov-Smirnov e, para a homogeneidade de variâncias, o teste de Levene, ambos a 0,01 e 0,05 de significância. Com os resultados obtidos, as espécies foram distribuídas em três grupos distintos:

- No primeiro foram incluídas aquelas que atendem a todas as pressuposições, sendo executados o MLG Normal, com função identidade, e Binomial, com função logística;
- No segundo se incluíram as espécies com uma ou mais pressuposições violadas e corrigidas pela transformação $\sqrt{x/100}$, em que x se refere ao percentual de plântulas normais;
- E no terceiro grupo foram incluídas aquelas em que, mesmo com a transformação arcoseno, os dados não atenderam às pressuposições.

ANOVA e MLG com distribuição Normal e função de ligação identidade

O percentual de plântulas normais seguiu o modelo de análise de variância de um delineamento inteiramente casualizado com dois fatores e interação (Equação 1), sendo definido por:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad (1)$$

$$i = 1, 2, \dots, a; \quad j = 1, 2, \dots, b; \quad k = 1, 2, \dots, r$$

em que:

y_{ijk} : porcentagem de plântulas normais obtida do i -ésimo lote, j -ésimo laboratório na k -ésima repetição;

μ : percentual médio de plântulas normais;

α_i : efeito do i -ésimo lote;

β_j : efeito do j -ésimo laboratório;

γ_{ij} : efeito da interação do j -ésimo laboratório no i -ésimo lote;

ε_{ijk} : resíduo independente e normalmente distribuído;

a : número de lotes,

b : número de laboratórios;

r : número de repetições.

Como o modelo exige resíduos com distribuição Normal e variâncias homogêneas, foi necessário testar o atendimento dessas pressuposições. A ANOVA é um caso particular dos Modelos Lineares Generalizados, e sua incorporação ao MLG foi realizada assumindo a distribuição Normal do componente aleatório e a utilização da função de ligação identidade. Como o MLG com distribuição Normal e função de ligação identidade se espelha na ANOVA, o atendimento às pressuposições também é necessário.

MLG com distribuição Binomial e função de ligação logit

O número de plântulas normais formadas seguiu a distribuição Binomial, correspondendo ao evento de sucesso. A linearização do modelo se deu pela forma:

$$\text{Logit}(p_i) = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (2)$$

em que:

Logit: função de ligação canônica para a distribuição Binomial;

a : intercepto aleatório;

b_1, b_2, \dots, b_n : parâmetros estimados para o modelo;

X_1, X_2, \dots, X_n : variáveis explanatórias.

Análise de *deviance* (ANODEV)

Observou-se a significância dos efeitos de laboratório, lote e interação na variável para ambos os MLG testados por meio da análise de *deviance* (ANODEV), que se utilizou do log da verossimilhança do modelo. Para o modelo Binomial, as hipóteses foram testadas pelo teste de Qui-quadrado (χ^2), e o Normal, pelo teste de F-Snedecor (F).

Para as espécies que necessitaram de transformação dos dados, a ANODEV foi executada nas escalas, até para aquelas espécies em que o atendimento às pressuposições não foi atingido com a transformação, no intuito de realizar a comparação final dos modelos. Por critérios de redação, foi designado como Modelo 1 aquele com distribuição Normal para os percentuais de germinação; Modelo 2, o que apresenta distribuição Normal para a transformação arcoseno dos percentuais de germinação; e Modelo 3, o que se refere à distribuição Binomial para o número de sementes germinadas.

Ajuste do modelo

Espécies que atenderam às pressuposições na escala original do percentual de plântulas normais possuíram dois modelos para comparação (Modelos 1 e 3). Para espécies em que a transformação dos dados foi executada, um modelo adicional também foi comparado (Modelo 2). Com vistas a verificar a qualidade do ajuste dos modelos, utilizaram-se o critério de informação de Akaike (AIC), a distância de Cook e os gráficos q-q plot.

RESULTADOS

Das 50 espécies, 27 apresentaram resíduos com distribuição Normal e variâncias estáveis para a variável do percentual de plântulas normais (Tabela 2). Nas situações de falha na escala original, a heterocedasticidade foi mais frequente (18 casos), quando comparada à normalidade de resíduos (três casos). Duas espécies tiveram ambos os pressupostos violados; logo, não se observou no banco de dados uma relação entre a não normalidade e a heterogeneidade das variâncias.

TABELA 2. Estatísticas e probabilidades associadas dos testes de Kolmogorov-Smirnov ($K-S$) e Levene (F) para os percentuais de plântulas normais na escala original e na escala transformada ($\arcsen\sqrt{x/100}$), para as 50 espécies florestais do processo de validação de métodos para teste de germinação.

Espécie	Dados originais				Dados transformados				Aplicar ANOVA?	
	$K-S$ (prob.)	Inferências	F (prob.)	Inferências	$K-S$ (prob.)	Inferências	F (prob.)	Inferências	Orig.	Transf.
<i>A. polyphylla</i>	0,073 (0,452)	Normal		Heterogêneas	0,052 (0,897)	Normal	2,167 (0,016)	Heterogêneas	Não aplicável	
<i>A. hassleri</i>	0,054 (0,723)	Normal	0,944 (0,544)	Homogêneas					X	
<i>A. colubrina</i>	0,067 (0,585)	Normal	2,369 (0,008)	Heterogêneas	0,058 (0,793)	Normal	2,170 (0,016)	Heterogêneas	Não aplicável	
<i>A. macrocarpa</i>	0,047 (0,632)	Normal	1,298 (0,201)	Homogêneas					X	
<i>A. leiocarpa</i>	0,087 (0,211)	Normal	1,965 (0,032)	Heterogêneas	0,101 (0,067)	Normal	1,345 (0,201)	Homogêneas		X
<i>A. fraxinifolium</i>	0,057 (0,812)	Normal	1,393 (0,177)	Homogêneas					X	
<i>C. estrellensis</i>	0,075 (0,311)	Normal	1,477 (0,123)	Homogêneas					X	
<i>C. legalis</i>	0,068 (0,559)	Normal	1,512 (0,126)	Homogêneas					X	
<i>C. leptophylla</i>	0,094 (0,126)	Normal	2,369 (0,009)	Heterogêneas	0,107 (0,062)	Normal	3,400 (0,003)	Heterogêneas	Não aplicável	
<i>C. fissilis</i>	0,108 (0,017)	não Normal	1,340 (0,189)	Homogêneas	0,055 (0,760)	Normal	1,138 (0,337)	Homogêneas		X
<i>C. odorata</i>	0,068 (0,261)	Normal	1,946 (0,013)	Heterogêneas	0,054 (0,652)	Normal	2,188 (0,004)	Heterogêneas	Não aplicável	
<i>C. speciosa</i>	0,078 (0,237)	Normal	1,255 (0,243)	Homogêneas						
<i>C. myrianthum</i>	0,111 (0,031)	não Normal	1,371 (0,188)	Homogêneas	0,080 (0,317)	Normal	1,057 (0,417)	Homogêneas		X
<i>C. langsdorffii</i>	0,070 (0,432)	Normal	1,683 (0,062)	Homogêneas					X	
<i>C. americana</i>	0,062 (0,710)	Normal	1,244 (0,264)	Homogêneas					X	
<i>C. antisiphilitica</i>	0,052 (0,699)	Normal	2,326 (0,002)	Heterogêneas	0,038 (0,969)	Normal	2,01 (0,010)	Heterogêneas	Não aplicável	
<i>D. miscolobium</i>	0,088 (0,287)	Normal	1,010 (0,462)	Homogêneas					X	
<i>D. nigra</i>	0,062 (0,716)	Normal	2,042 (0,024)	Heterogêneas	0,067 (0,580)	Normal	3,563 (0,000)	Heterogêneas	Não aplicável	
<i>E. contortisiliquum</i>	0,070 (0,129)	Normal	2,107 (0,003)	Heterogêneas	0,054 (0,469)	Normal	1,665 (0,030)	Heterogêneas	Não aplicável	
<i>E. maximum</i>	0,092 (0,141)	Normal	0,549 (0,914)	Homogêneas					X	
<i>E. speciosa</i>	0,062 (0,763)	Normal	1,469 (0,145)	Homogêneas					X	
<i>G. integrifolia</i>	0,098 (0,820)	Normal	0,722 (0,767)	Homogêneas					X	
<i>G. ulmifolia</i>	0,067 (0,633)	Normal	1,731 (0,070)	Homogêneas					X	
<i>H. chrysotrichus</i>	0,057 (0,663)	Normal	2,930 (0,000)	Heterogêneas	0,072 (0,249)	Normal	3,103 (0,000)	Heterogêneas	Não aplicável	
<i>H. impetiginosus</i>	0,065 (0,630)	Normal	1,534 (0,118)	Homogêneas					X	
<i>H. roseo-albus</i>	0,094 (0,127)	Normal	0,852 (0,629)	Homogêneas					X	

“Continua...”

“Continuação...”

Espécie	Dados originais				Dados transformados				Aplicar ANOVA?	
	K-S (prob.)	Inferências	F (prob.)	Inferências	K-S (prob.)	Inferências	F(prob.)	Inferências	Orig.	Transf.
<i>H. courbaril</i>	0,092 (0,170)	Normal	0,614 (0,864)	Homogêneas					X	
<i>H. stigonocarpa</i>	0,100 (0,154)	Normal	2,776 (0,003)	Heterogêneas	0,066 (0,613)	Normal	2,590 (0,004)	Heterogêneas	Não aplicável	
<i>J. cuspidifolia</i>	0,089 (0,120)	Normal	1,789 (0,044)	Heterogêneas	0,078 (0,276)	Normal	1,577 (0,091)	Homogêneas		X
<i>J. micranta</i>	0,091 (0,167)	Normal	1,113 (0,368)	Homogêneas					X	
<i>L. pacari</i>	0,110 (0,031)	não Normal	2,521 (0,005)	Heterogêneo	0,102 (0,064)	Normal	2,388 (0,008)	Heterogêneas	Não aplicável	
<i>M. caesalpiniaefolia</i>	0,059 (0,723)	Normal	2,328 (0,004)	Heterogêneas	0,046 (0,895)	Normal	1,539 (0,085)	Homogêneas		X
<i>M. scabrella</i>	0,079 (0,065)	Normal	1,278 (0,193)	Homogêneas					X	
<i>O. arborea</i>	0,083 (0,250)	Normal	1,306 (0,225)	Homogêneas					X	
<i>P. rigida</i>	0,099 (0,079)	Normal	2,161 (0,017)	Heterogêneas	0,107 (0,037)	Não Normal	1,335 (0,207)	Homogêneas	Não aplicável	
<i>P. pendula</i>	0,067 (0,603)	Normal	2,088 (0,021)	Heterogêneas	0,064 (0,652)	Normal	1,827 (0,048)	Heterogêneas	Não aplicável	
<i>P. confertiflora</i>	0,135 (0,000)	não Normal	1,678 (0,063)	Homogêneas	0,072 (0,342)	Normal	1,035 (0,437)	Homogêneas		X
<i>P. dubium</i>	0,091 (0,051)	Normal	1,039 (0,433)	Homogêneas					X	
<i>P. reticulata</i>	0,083 (0,249)	Normal	1,871 (0,042)	Heterogêneas	0,082 (0,267)	Normal	1,721 (0,067)	Homogêneas		X
<i>P. tomentosum</i>	0,080(0,201)	Normal	1,256 (0,243)	Homogêneas					X	
<i>P. nitens</i>	0,080 (0,181)	Normal	0,566 (0,902)	Homogêneas					X	
<i>Q. grandiflora</i>	0,094 (0,187)	Normal	2,736 (0,003)	Heterogêneas	0,083 (0,248)	Normal	2,223 (0,013)	Heterogêneas	Não aplicável	
<i>S. terebinthifolius</i>	0,086 (0,149)	Normal	2,094 (0,014)	Heterogêneas	0,060 (0,622)	Normal	1,717 (0,054)	Homogêneas		X
<i>S. parahyba</i>	0,104 (0,029)	não Normal	1,788 (0,043)	Heterogêneas	0,070 (0,415)	Normal	0,765 (0,742)	Homogêneas		X
<i>S. macranthera</i>	0,052 (0,985)	Normal	1,287 (0,211)	Homogêneas					X	
<i>S. multijuga</i>	0,071 (0,537)	Normal	1,661 (0,084)	Homogêneas					X	
<i>S. barbadetiman</i>	0,058 (0,786)	Normal	0,736 (0,752)	Homogêneas					X	
<i>S. polyphyllum</i>	0,083 (0,304)	Normal	2,026 (0,026)	Heterogêneas	0,075 (0,389)	Normal	1,597 (0,097)	Homogêneas		X
<i>T. aurea</i>	0,064 (0,698)	Normal	1,704 (0,073)	Homogêneas					X	
<i>Z. tuberculosa</i>	0,068 (0,587)	Normal	0,841 (0,641)	Homogêneas					X	

A transformação arcoseno foi empregada na tentativa de corrigir a falta de normalidade e/ou homogeneidade das 23 espécies com pressuposições violadas, sendo que em apenas dez delas esse recurso foi satisfatório (Tabela 2). Para as 20 espécies com heterocedasticidade, a transformação do tipo angular estabilizou a variância de somente oito, demonstrando sua incapacidade em corrigir todas as situações de violação desse pressuposto. Em contrapartida, para as cinco espécies com problemas na normalidade dos resíduos, a transformação garantiu a aderência dos dados à distribuição.

Uma síntese dos resultados das pressuposições levou ao agrupamento de 27 espécies (Grupo I), tendo em comum o atendimento a todas as pressuposições, e a essas espécies foram aplicados o MLG Normal, com função identidade, e Binomial, com função logística (Quadro 1). No segundo grupo, encontram-se as dez espécies com pressuposições atendidas após a transformação dos dados, e no terceiro, as 13 com pressuposições não atendidas, mesmo após a transformação. As espécies dos grupos II e III tiveram seus percentuais originais e transformados analisados por meio do MLG Normal com função identidade (ANOVA).

QUADRO 1. Grupos de espécies entre as 50 separadas em função dos resultados das pressuposições do modelo.

Grupo I- Espécies com pressuposições atendidas com os dados originais				
<i>A. hassleri</i>	<i>C. langsdorffii</i>	<i>G. ulmifolia</i>	<i>O. arborea</i>	<i>S. barbadetiman</i>
<i>A. macrocarpa</i>	<i>C. americana</i>	<i>H. impetiginosus</i>	<i>P. dubium</i>	<i>T. aurea</i>
<i>A. fraxinifolium</i>	<i>D. miscolobium</i>	<i>H. roseo-albus</i>	<i>P. tomentosum</i>	<i>Z. tuberculosa</i>
<i>C. estrellensis</i>	<i>E. maximum</i>	<i>H. courbaril</i>	<i>P. nitens</i>	
<i>C. legalis</i>	<i>E. speciosa</i>	<i>J. micranta</i>	<i>S. macranthera</i>	
<i>C. speciosa</i>	<i>G. integrifolia</i>	<i>M. scabrella</i>	<i>S. multijuga</i>	
Grupo II- Espécies com pressuposições atendidas com os dados transformados				
Ila- Não normalidade dos resíduos		I Ib- Variâncias heterogêneas		
<i>C. fissilis</i>	<i>P. confertiflora</i>	<i>A. leiocarpa</i>	<i>M. caesalpiniaefolia</i>	<i>S. terebinthifolius</i>
<i>C. myrianthum</i>	<i>S. parahyba</i>	<i>J. cuspidifolia</i>	<i>P. reticulata</i>	<i>S. polyphyllum</i>
Grupo III- Espécies com variâncias heterogêneas mesmo com transformação				
<i>A. polyphylla</i>	<i>C. odorata</i>	<i>E. contortisiliquum</i>	<i>L. pacari</i>	<i>Q. grandiflora</i>
<i>A. colubrina</i>	<i>C. antisiphilitica</i>	<i>H. chrysotrichus</i>	<i>P. rigida</i>	
<i>C. leptophylla</i>	<i>D. nigra</i>	<i>H. stigonocarpa</i>	<i>P. pendula</i>	

A partir do quadro de ANODEV gerado (Anexo A), observou-se que todas as espécies apresentaram diferença significativa para o fator lote ($p\text{-valor} < 0,001$), independentemente da distribuição, Normal e Binomial. Esse comportamento era

esperado, dada a grande diferença de potencial germinativo entre os lotes, com o intuito de verificar se a metodologia aplicada era eficaz na diferenciação dos lotes.

Ao comparar apenas o MLG Normal dos percentuais originais com o Binomial, 16 espécies apresentaram a mesma significância dos fatores no ANODEV gerado, a exemplo de *H. impetiginosus* (Tabela 3c). Destas, 11 pertenceram ao caso I e cinco, aos casos II e III. Diante disso, 41% das espécies com pressuposições atendidas e 22% daquelas não atendidas obtiveram a mesma interpretação sobre o efeito dos fatores; logo, a mudança da análise dentro dos MLG foi radical até para espécies com pressuposições atendidas. Nas outras 34, a análise Binomial aumentou a significância dos fatores relacionados a laboratório ou interação laboratório x lote, como *C. americana* (Tabela 3a), em que 18 se enquadram nos casos II e III.

O aumento de significância para o fator laboratório no modelo Binomial foi o mais comum entre as espécies. No caso I, oito espécies tiveram a significância do fator laboratório, e seis, tanto do fator laboratório quanto da interação. Para os casos II e III, três espécies aumentaram a significância do laboratório; sete, a interação; e cinco, ambas.

TABELA 3. ANODEV para o percentual de germinação (Distribuição Normal – função identidade) e número de sementes (Distribuição Binomial – função logística) para o modelo com dois fatores (laboratório e lote) e interação em *Cordia americana* (a), *Guazuma ulmifolia* (b) e *Handroanthus impetiginosus* (c).

<i>Cordia americana</i> (a)									
Normal							Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	24758,9				685,72		
Lab.	5	66	1889,0	22869,9	3,253	0,012	46,46	639,26	0,000
Lote	2	64	15609,1	7260,8	67,19	0,000	436,11	203,15	0,000
Lab*Lote	10	54	988,6	6272,2	0,851	0,583	30,69	172,45	0,000
Saturado	17	54	6272,2				30,69		

<i>Guazuma ulmifolia</i> (b)									
Normal							Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	65	12921,8				369,29		
Lab.	5	60	299,3	12622,5	2,459	0,046	8,37	360,92	0,137
Lote	2	58	11280,2	1342,3	231,7	0,000	322,09	38,84	0,000
Lab*Lote	10	48	174,2	1168,1	0,716	0,705	3,48	35,36	0,968
Saturado	17	48	1168,1				35,36		

<i>Handroanthus impetiginosus</i> (c)									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	17949,5				423,4		
Lab.	5	66	496,5	17453,0	1,975	0,097	10,0	413,41	0,077
Lote	2	64	14312,2	3140,8	142,3	0,000	340,1	73,33	0,000
Lab*Lote	10	54	425,2	2715,5	0,846	0,588	9,6	63,70	0,474
Saturado	17	54	2715,5				63,70		

F.V.: Fonte de variação; gl: grau de liberdade; F: Estatística de Snedecor; Lab.: Laboratório

Ao se atentar apenas às espécies pertencentes aos casos II e III, o ANODEV para as três situações (dados normais, dados normais transformados e dados binomiais – Anexo B) apresentou os mesmos resultados apenas para cinco espécies, a exemplo de *A. colubrina* (Tabela 4a), o que demonstra a provável ocorrência dos erros Tipo I e II.

Em sete espécies, o ANODEV para os dados Normais transformados e Binomiais foram iguais, diferenciando-se dos dados Normais originais. Para *A. polyphyla* e *C. myrianthum* (Tabela 4c), o fator laboratório tornou-se significativo a 0,05; em *P. reticulata*, o mesmo fator aumentou sua significância para 0,01; *C. odorata*, *H. chrysotrichus* (Tabela 4d) e *S. terebinthifolius* perderam a significância deste, atendido a 5% no modelo Normal com os dados originais; e para *P. pendula*, o fator interação não significativo para os percentuais originais se tornou a 0,01 para os outros dois modelos.

Enquanto isso, em *C. odorata*, o fator da interação tornou-se significativo a 0,05 com os percentuais transformados e a 0,01 no modelo Binomial. Já para três pertencentes ao caso II (*A. leiocarpa*, *P. reticulata*, *S. polyphyllum*) e seis ao caso III (*C. leptophylla*, *C. antisiphilitica*, *E. contortisiliquum*, *H. stigonocarpa*, *L. pacari*, *P. rigida*), os dois modelos Normais propostos se equipararam, mas o Binomial detectou diferença significativa para o fator interação. Ainda em *C. leptophylla* (Tabela 4b), além da interação, o laboratório se tornou significativo dentro do modelo Binomial.

Das cinco espécies com desvio de normalidade, apenas *P. confertiflora* obteve o mesmo resultado nos três casos analisados. Em *C. myrianthum*, a transformação de dados foi eficaz e obteve as mesmas significâncias constatadas no modelo Binomial. Para *L. pacari*, *C. fissilis* e *S. parahyba*, o modelo Binomial foi mais sensível na detecção de diferenças significativas que os demais. Apesar do pequeno número de espécies com problemas de normalidade, esse resultado demonstra como tal pressuposição pode afetar a análise de variância.

TABELA 4. ANODEV para o percentual de germinação original e transformado (Distribuição Normal – função identidade) e número de sementes (Distribuição Binomial – função logística) para o modelo com dois fatores (laboratório e lote) e interação para *A. colubrina* (a), *C. leptophylla* (b), *C. myrianthum* (c) e *H. chrysotrichus* (d).

<i>A. colubrina</i> (a)									
Normal				Normal (Transf. Arcoseno)			Binomial		
F.V.	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	33477,0			13063,5			856,24		
Lab.	427,6	33049	0,166	199,0	12864,5	0,134	8,76	847,48	0,119
Lote	29645,6	3404	0,000	11382,6	1481,9	0,000	756,18	91,30	0,000
Lab*Lote	580,9	2823	0,371	271,4	1210,5	0,305	16,27	75,03	0,092
Saturado	2823			1210,5			75,03		
<i>C. leptophylla</i> (b)									
Normal				Normal (Transf. Arcoseno)			Binomial		
F.V.	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	25716,8			9928,4			730,31		
Lab.	598,4	25118,4	0,138	328,8	9599,6	0,059	13,99	716,32	0,016
Lote	20567,7	4550,7	0,000	7647,3	1952,4	0,000	596,07	120,24	0,000
Lab*Lote	943,5	3607,2	0,212	400,9	1551,5	0,208	22,52	97,73	0,012
Saturado	3607,2			1551,5			97,73		
<i>C. myrianthum</i> (c)									
Normal				Normal (Transf. Arcoseno)			Binomial		
F.V.	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	18184,6			7847			516,34		
Lab.	457,1	17727,5	0,062	228,6	7618,4	0,030	13,89	502,46	0,016
Lote	15130,9	2596,6	0,000	6502	1116,4	0,000	427,03	75,43	0,000
Lab*Lote	454,1	2142,5	0,363	214,5	901,9	0,276	14,73	60,70	0,142
Saturado	2142,5			901,9			60,70		
<i>H. chrysotrichus</i> (d)									
Normal				Normal (Transf. Arcoseno)			Binomial		
F.V.	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	72518			35405			516,34		
Lab.	510	72009	0,028	254	35151	0,138	13,89	502,46	0,016
Lote	69559	2450	0,000	33124	2028	0,000	427,03	75,43	0,000
Lab*Lote	436	2014	0,431	435	1592	0,173	14,73	60,70	0,142
Saturado	2014			1592			60,70		

F.V.: Fonte de variação; gl: grau de liberdade; F: Estatística de Snedecor; Lab.: Laboratório. ¹Para esse modelo, os percentuais de germinação foram transformados pela expressão $\sqrt{x/100}$, em que x é o percentual de plântulas normais.

Os valores de AIC variaram de 414,95 (*G. ulmifolia*) a 815,97 (*M. scabrella*), para as espécies com distribuição Normal, e de 313,96 (*C. legalis*) a 680,47 (*M. scabrella*), para a Binomial (Tabela 5) – naquelas com os dados transformados, os valores foram de 419,96 (*C. myrianthum*) a 816,41 (*E. contortisiliquum*). Em todas as espécies, o valor do AIC foi menor no modelo com distribuição Binomial e função de ligação logística, e a espécie com menor redução no valor foi *C. speciosa* (10,96).

Ademais, a maioria apresentou redução maior que 50, e 11 ainda tiveram redução maior que 100 no valor de AIC – dessas onze, oito não atenderam a uma das pressuposições no modelo Normal. As espécies com os dados transformados também tiveram redução no valor de AIC, em comparação aos dados normais originais, com exceção de *E. contortisiliquum* e *H. chrysotrichus*. Entretanto, todas as reduções não foram da mesma grandeza que aquelas observadas no modelo Binomial, com nenhuma redução maior que 100.

TABELA 5. Valores para o critério de informação de Akaike (AIC) do MLG Normal, com função de ligação identidade na escala original (N); na escala transformada (NT); e do MLG Binomial com função de ligação logit (B), formadas para as 50 espécies florestais do processo de validação de métodos para teste de germinação.

Espécies com pressuposições atendidas (Grupo I)			Espécies com pressuposições não atendidas (Grupos II e III)			
Espécie	N	B	Espécie	N	B	NT
<i>A. hassleri</i>	602,14	513,36	<i>A. polyphylla</i>	488,58	390,26	427,01
<i>A. macrocarpa</i>	664,54	565,78	<i>A. colubrina</i>	506,48	415,55	455,52
<i>A. fraxinifolium</i>	487,45	413,69	<i>A. leiocarpa</i>	521,59	443,24	474,51
<i>C. estrellensis</i>	578,16	473,79	<i>C. leptophylla</i>	518,3	473,79	463,39
<i>C. legalis</i>	479,29	313,96	<i>C. fissilis</i>	602,97	487,46	534,78
<i>C. speciosa</i>	578,16	567,2	<i>C. odorata</i>	692,43	560,31	628,18
<i>C. langsdorffii</i>	508,94	410,71	<i>C. myrianthum</i>	481,39	390,5	419,96
<i>C. americana</i>	563,97	512,26	<i>C. antisiphilitica</i>	715,31	586,83	669,93
<i>D. miscolobium</i>	546,19	467,32	<i>D. nigra</i>	474,03	394,08	441,68
<i>E. maximum</i>	556,29	520,46	<i>E. contortisiliquum</i>	800,99	632,33	816,41
<i>E. speciosa</i>	484,16	434,21	<i>H. chrysotrichus</i>	590,08	451,73	592,05
<i>G. integrifolia</i>	536,69	512,23	<i>H. stigonocarpa</i>	519,9	453,33	503,96
<i>G. ulmifolia</i>	414,95	361,77	<i>J. cuspidifolia</i>	517,54	424,12	442,48
<i>H. impetiginosus</i>	503,69	414,95	<i>L. pacari</i>	495,69	417,15	437,7
<i>H. roseo-albus</i>	510,45	445,80	<i>M. caesalpiniaefolia</i>	605,66	516,27	575,24
<i>H. courbaril</i>	457,82	368,34	<i>P. rigida</i>	533,01	425,47	480,62
<i>J. micrantha</i>	492,77	444,62	<i>P. pendula</i>	489,15	409,71	459,39
<i>M. scabrella</i>	815,97	680,47	<i>P. confertiflora</i>	540,76	412,34	497,58
<i>O. arborea</i>	523,91	442,67	<i>P. reticulata</i>	519,03	435,04	467,98
<i>P. dubium</i>	614,94	525,18	<i>Q. grandiflora</i>	512,54	420,64	453,61
<i>P. tomentosum</i>	564,96	472,72	<i>S. terebinthifolius</i>	550,24	451,00	514,21
<i>P. nitens</i>	464,73	390,73	<i>S. parahyba</i>	540,87	401,28	497,3
<i>S. macranthera</i>	533,98	519,90	<i>S. polyphyllum</i>	493,67	424,62	441,65
<i>S. multijuga</i>	450,70	366,36				
<i>S. barbadetiman</i>	477,61	410,71				
<i>T. aurea</i>	452,64	375,39				
<i>Z. tuberculosa</i>	496,42	426,18				

Com o intuito de complementar os resultados encontrados pelo AIC, as análises gráficas dos “Normal plots” e da distância de Cook foram realizadas. Para os “Normal plots” (Anexos C e D), os resíduos estudantizados plotados se ajustaram em seus respectivos quantis, tanto na escala Normal quanto Binomial em 23 espécies, a exemplo de *C. antisiphilitica* (Figura 1a), sem haver diferenças em comparação às duas distribuições.

Para *C. speciosa* e *C. langsdorffii*, espécies com pressuposições atendidas, o gráfico para o modelo Binomial corrigiu os pontos que se encontravam fora do envelope simulado. Em *A. colubrina*, *D. nigra*, *L. pacari*, *S. terebinthifolius* e *S. parahyba* var. *amazonicum* (Figura 1e), espécies com pressuposições violadas, o gráfico para os modelos Binomial e Normal com transformação de dados corrigiu os pontos. Entretanto, para *A. leiocarpa*, *E. contortisiliquum* (Figura 1b), *H. chrysotrichus*, *H. stigonocarpa*, *P. rigida*, *P. reticulata* e *S. polyphyllum*, também com pressuposições violadas, apenas o modelo Binomial foi capaz de corrigir os pontos fora do envelope simulado.

Em seis espécies (por exemplo, *P. confertiflora* (Figura 1c)), apesar de os pontos fora do envelope não serem corrigidos por completo, o modelo Binomial também se apresentou com melhor ajuste. Porém, para outras seis espécies ocorreu o inverso (*Q. grandiflora* (Figura 1d)), e os resíduos se ajustaram melhor no modelo Normal.

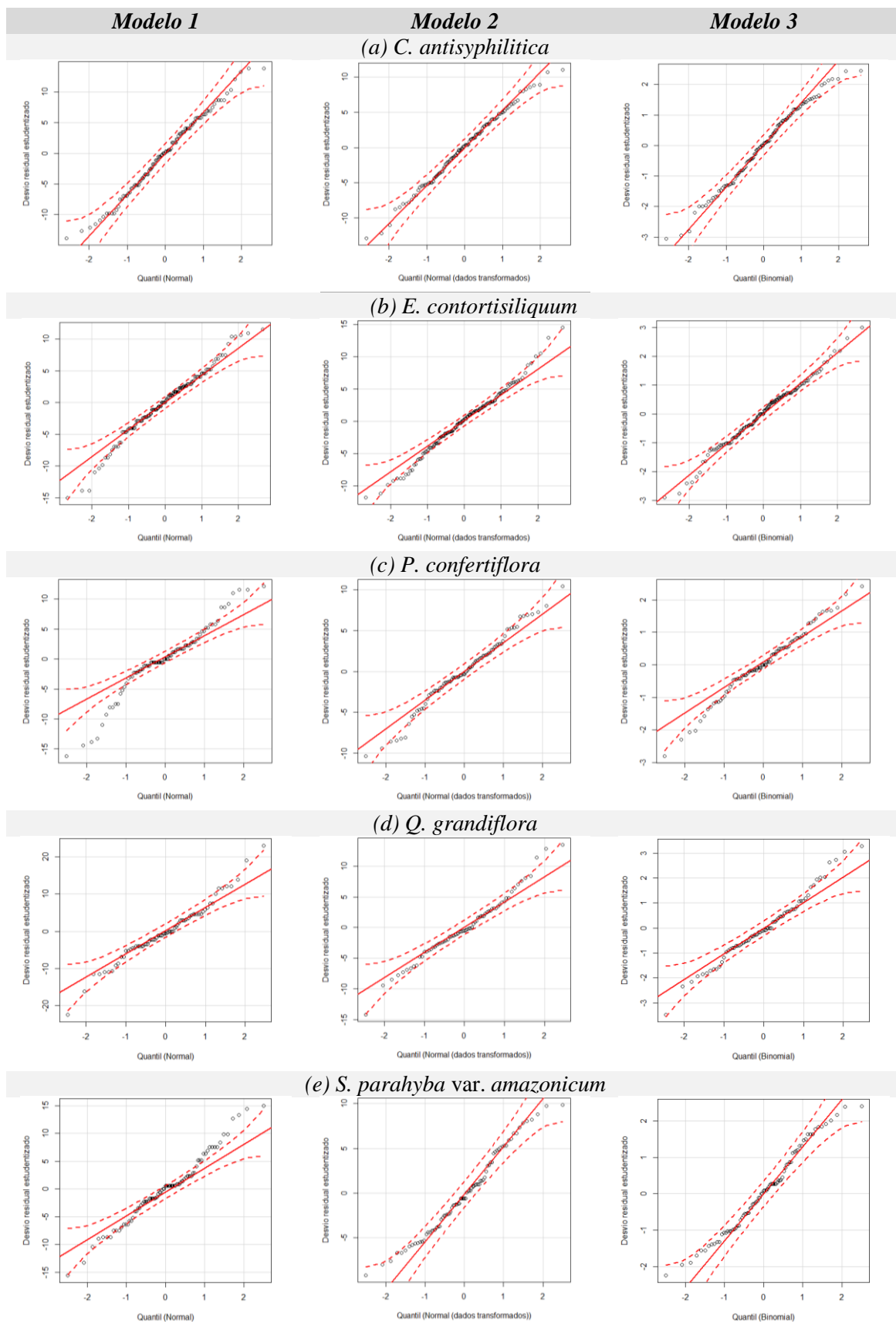


FIGURA 1. Gráfico q-q plot para os resíduos estudantizados de *C. antisiphilitica* (a), *E. contortisiliquum* (b), *P. confertiflora* (c), *Q. grandiflora* (d) e *S. parayiba* (e), em relação aos modelos Normal identidade (1ª e 2ª coluna) e Binomial *logit* (3ª coluna).

De forma geral, não ocorreram discrepâncias nas estatísticas de Cook nos modelos propostos, seja para os percentuais ou para o número de sementes germinadas (Anexos E e F), a exemplo da espécie *C. antisiphilitica* (Figura 2). Nem mesmo para os modelos Normais sem atendimento dos pressupostos houve a presença de pontos influentes com valores de D_i acima de um.

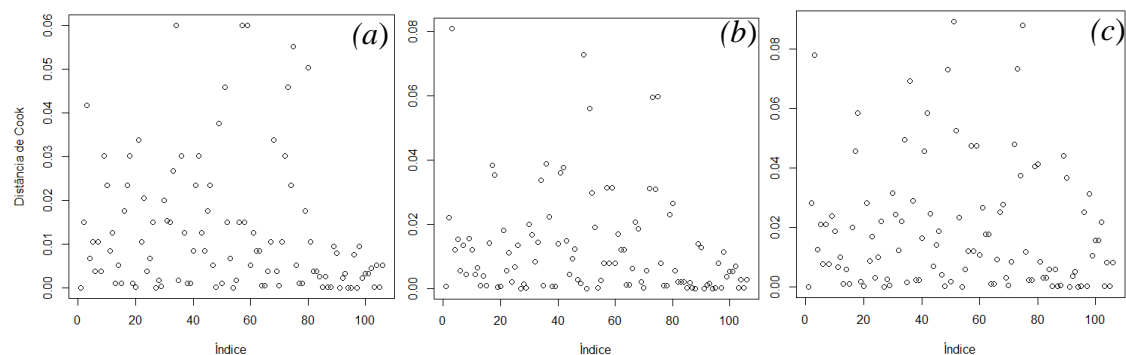


FIGURA 2. Gráfico da distância de Cook para MLG Normal dos percentuais de germinação (a), MLG Normal dos percentuais de germinação transformados, e MLG Binomial do número de sementes germinadas em *C. antisiphilitica*.

DISCUSSÃO

As problemáticas do não atendimento às pressuposições da ANOVA estão ligadas à perda de confiabilidade do teste, que pode resultar em erros Tipo I ou II (BRADLEY, 1978; GLASS et al., 1972; LEVINE; DUNLAP, 1983; RASMUSSEN, 1985), o que afeta diretamente o nível de significância do teste (GOMEZ; GOMEZ, 1984; KEMPTHORNE, 1952; LITTLE; HILLS, 1978).

Apesar de algumas referências indicarem a robustez da ANOVA para pequenos desvios de normalidade (SHARPE, 1970; HARWELL et al. 1992; FARAWAY, 2006), estudos mais recentes revelam o impacto causado na violação dessa pressuposição (KHAN; RAYNER, 2003; SCHMIDER et al., 2010; LANTZ, 2013a, 2013b). Isso, aliado à difícil determinação dos limites para pequenos desvios de normalidade, permite inferir que a ANOVA não deveria ser executada quando os resíduos não seguem a distribuição Normal.

Os resultados encontrados contrastam com o relatado por outros autores sobre a falta de homogeneidade comumente ligada à ausência de normalidade (BOX; COX, 1964; PEÑA, 1994; FONT et al., 2007). Segundo esses autores, os resíduos com

distribuição não Normal tendem a possuir variâncias heterogêneas. Mas o estudo mostrou que a maior parte dos casos de não normalidade não estava correlacionada à falta de homogeneidade e vice-versa.

Com a criação de lotes de diferentes padrões germinativos, os de maior e menor qualidade germinativa tenderam a possuir variâncias menores do que lotes intermediários, o que pode ter levado à heterogeneidade de variâncias em 40% das espécies. Isso é justificado em porcentagens baseadas em contagens, nas quais as variâncias de médias próximas a 0 e 100% tendem a ser menores que aquelas entre 30% e 70% (FINNEY, 1989; SILESHI, 2012). Há relatos de que dados de porcentagens não seguem a distribuição Normal, necessitando de uma transformação (LABOURIAU, 1983). Nesse caso, a elevada quantidade de espécies com resíduos normalmente distribuídos (90%) contrapõe essa afirmação.

A transformação angular é comumente utilizada na tentativa de estabilizar os dados e cumprir as pressuposições, mesmo que não seja recomendada por alguns estudiosos (WARTON; HUI, 2011; SILESHI, 2012). Outros autores indicam a transformação como possível e eficiente, quando os dados compõem tratamentos que não apresentam homocedasticidade (AHRENS et al., 1990; VIEIRA, 2006), mas, em geral, é um instrumento útil tanto para a correção da homogeneidade como para a da normalidade (ATKINSON, 1985; SOKAL; ROHLF, 1995; ZAR, 1998; GOTELLI; ELLISON, 2004). Os resultados obtidos complementam as duas vertentes a respeito da transformação, pois em alguns casos ela foi eficaz e em outros, não.

Nesse contexto, há a situação observada em *P. rigida*, em que a inversão da problemática dos pressupostos com a transformação já foi detectada por Sileshi (2007; 2012), que alertava que, ao corrigir uma pressuposição, a transformação pode violar outra. Para tanto, é injustificável qualquer transformação sem necessidade, e o analista deve estar ciente de que a transformação pode piorar ou não solucionar o problema existente no banco de dados.

Quando utilizada de forma consciente, a transformação angular se tornou um instrumento eficaz para o analista, numa possível solução do não atendimento às pressuposições. Apesar de ser apontada como solução para o uso seguro da ANOVA (COCHRAN, 1940; RAO, 1960; WINER et al., 1971), ela não pode ser tratada como única e nem ser empregada sem restrição ou análise crítica, pois pode piorar a análise, conforme mencionado anteriormente – isso demonstra que nem sempre a transformação de dados é suficiente para corrigir o não atendimento de uma pressuposição,

necessitando o emprego de outras metodologias estatísticas, a exemplo dos MLG (LIX et al., 1996; LANTZ, 2013a), e especificamente para dados com natureza binária como a germinação, da regressão logística (WILSON; HARDY, 2002; WARTON; HUI, 2011).

A transformação de dados se mostra eficiente na solução de problemas envolvendo a quebra de pressuposições nos modelos lineares, mas nem sempre é a melhor opção. Como mostra Montgomery (2000), a abordagem generalizada foi superior à de transformação e não paramétrica, e a técnica dos Modelos Lineares Generalizados explicou melhor a variabilidade dos dados estudados.

Houve maior sensibilidade do modelo Binomial em prever diferenças significativas entre os fatores analisados, o que mostra que, mesmo com dados oriundos do mesmo experimento, a mudança na análise confere diferenças nas conclusões finais – essa maior sensibilidade do modelo Binomial, em comparação ao Normal, também foi detectada por Sileshi (2007) e Mora et al. (2008). A maior sensibilidade do modelo Binomial poderia ser condicionada a uma menor taxa do erro Tipo II (aceitar hipóteses de nulidade quando falsas) presente no modelo, ou seja, ocorreu maiores tendências de o modelo Normal considerar os fatores analisados iguais quando estes foram estatisticamente diferentes. Nesse contexto, Lantz (2013a) demonstra que a probabilidade de se evitar o erro Tipo II é sempre maior quando os dados são de natureza discreta, em contraposição aos de natureza contínua.

A maior redução do AIC ocorre principalmente quando os modelos são subdimensionados (erro Tipo II), em detrimento daqueles superdimensionados (erro Tipo I) (SCHWARZ, 1978). Mais uma vez, a provável redução do erro Tipo II nos dados de natureza Binomial pode ter provocado reduções significativas no AIC. Estudos posteriores que comparam o poder do teste com o modelo utilizado são necessários, dada a escassez de literatura que aborda a temática (SILESHI, 2007).

Nota-se que as espécies com os dados transformados também apresentaram redução no valor de AIC, em comparação aos dados normais originais, com exceção de *E. contortisiliquum* e *H. chrysotrichus*. A redução da amplitude dos dados com a transformação reduz consequentemente a variação dos dados (BOX; COX, 1964), o que refletiu em redução no AIC – nesse caso, todas as reduções não foram da mesma grandeza que as do modelo Binomial, e nenhuma diminuição foi maior que 100. Os valores reduzidos do AIC para modelo Binomial indicam um modelo representativo e de menor complexidade, sendo esperado um melhor desempenho para a estimação de

novos dados a partir de modelos mais simples, ou seja, o AIC impõe uma penalidade para a complexidade (SAKAMOTO et al., 1989).

O AIC foi utilizado como forte ferramenta na seleção do modelo mais parcimonioso dentre aqueles analisados em outros trabalhos (FLORIANO et al., 2006; SILESHI, 2007; FARACO et al., 2008), mas é um valor de comparação único dentro de cada banco de dados. Por isso, não se encontram na literatura os valores recomendados tanto para o AIC quanto para outros coeficientes e estimadores, a exemplo do coeficiente de variação (CV) (PIMENTEL-GOMES, 2000).

Os gráficos quantil-quantil da distribuição Normal foram utilizados com o intuito de detectar pontos que não se ajustaram à curva (CIRILLO; FERREIRA, 2003). Modelos com muitos pontos fora do envelope simulado revelariam a sua dificuldade em ajustá-los (D'AGOSTINO; STEPHENS, 1986), posto que observações atípicas (*outliers*) já haviam sido retiradas pelo teste de Box-Plot.

Apesar da crítica de autores ao q-q plot, por ser uma análise subjetiva, dependendo de interpretação visual (ZAR, 1998; THODE, 2002; TORTAN et al., 2012), os gráficos permitiram complementar a discussão acerca do ajuste dos modelos, e, como foram gerados com envelopes simulados, eles padronizaram a inferência. Tais gráficos também foram utilizados de forma eficiente por Augustin et al. (2012), para comparação de ajuste entre três modelos propostos para dados de incidência de câncer – quando o modelo é correto, os desvios residuais possuem variância constante e distribuição próxima à Normal.

A distância de Cook é comumente utilizada para a detecção de observações influentes na análise de regressão (COOK, 1977), por ser uma medida geral da mudança dos parâmetros estimados quando uma ou mais observações são excluídas do conjunto de dados. Inicialmente desenvolvida para os modelos normais, essa estatística foi estendida para os MLG (PREGIBON, 1981). Dessa maneira, ela foi necessária no intuito de observar se os modelos propostos para as espécies foram capazes de descrever todo o conjunto de dados. Vale ressaltar que, mesmo sendo um excelente método para avaliação do modelo em análise (WILLIAMS, 1987), a distância de Cook não auxiliou na tomada de decisão entre os modelos propostos, fato explicado pela retirada dos *outliers*.

Resende e Biele (2002) também obtiveram modelos mais adequados com a função logística e a distribuição Binomial do que com a identidade e a distribuição Normal em experimentos de sobrevivência de plantas de duas espécies florestais

(*Eucalyptus grandis* e *Ilex paraguariensis*). Dados de germinação de sementes de pinhão-mansão (*Jatropha curcas*) também se ajustaram melhor ao modelo Binomial logístico (ARAÚJO, 2012), assim como na germinação de orquídeas (MORA et al, 2008).

Os resultados encontrados retratam o que Paolino (2001) e Ferrari e Cribari-Neto (2004) sugerem para a modelagem de proporções: outras metodologias devem ser preferidas no lugar dos modelos lineares normais, pois a esperança condicional só é definida sobre o intervalo de 0 a 1. Logo, sua função, no que tange às variáveis explicativas, é não linear.

CONCLUSÃO

Com base no critério de informação de Akaike (AIC) e no gráfico q-q plot, conclui-se que, para os dados de plântulas normais das 50 espécies, o modelo com distribuição Binomial e função de ligação *logit* foi mais eficiente e parcimonioso para o ajuste dos dados e, portanto, é de maior confiabilidade.

A transformação angular foi capaz de corrigir os problemas no atendimento às pressuposições por completo em apenas dez espécies, em relação às 23 estudadas. Isso comprova que a aplicação do MLG com distribuição Binomial de imediato foi indispensável para 13 delas.

Desconsiderando as metodologias utilizadas para verificação do ajuste dos modelos, apenas 16 das 50 espécies obtiveram o mesmo modelo validado independente da distribuição e função de ligação utilizada, o que revelou o impacto causado pelo modelo aplicado.

REFERÊNCIAS BIBLIOGRÁFICAS

AHRENS, W. H.; COX, D. J.; BUDHWAR, G. Use of the arcsine and square root transformations for subjectively determined percentage data. **Weed Science**, v.38, p.452-458, 1990.

ARAÚJO, G. L. D. **Métodos de estimação em regressão logística com efeito aleatório: aplicação em germinação de sementes**. 94 f. 2012. Dissertação de Mestrado. Viçosa, MG.

AUGUSTIN, N. H.; SAULEAU, E.; WOOD, S. N. On quantile quantile plots for generalized linear models. **Computational Statistics and Data Analysis**, v.56, p.2400-2409, 2012.

ATKINSON, A. C. **Plots, transformations and regression; an introduction to graphical methods of diagnostic regression analysis**. New York: Oxford University Press, 1985. 282 p.

BIRCH, M. W. Maximum likelihood in three-way contingency tables. **Journal of the Royal Statistical Society**, v.25, n.1, p.220-233, 1963.

BLISS, C. I. The calculation of the dosage-mortality curve. **Annals of Applied Biology**, v.22, p.134-167, 1935.

BOX, G. E. P.; COX, D. R. An analysis of transformations. **Journal of the Royal Statistical Society-Series B**, v.26, p.211-252, 1964.

BRADLEY, J. Y. Robustness? **British Journal of Mathematical and Statistical Psychology**, v.31, p.144-152, 1978.

BRASIL. Instrução Normativa nº 26 de 10 de Setembro de 2012. **Diário oficial da República Federativa do Brasil**, Poder executivo, Brasília, 11, Set. 2012.

BRASIL. Instrução Normativa nº 35 de 14 de Julho de 2011. **Diário oficial da República Federativa do Brasil**, Poder executivo, Brasília, 15, Jul. 2011.

BRASIL. Instrução Normativa nº 44 de 23 de Dezembro de 2010. **Diário oficial da República Federativa do Brasil**, Poder executivo, Brasília, 24, dez. 2010.

CIRILLO, M. A.; FERREIRA, D. F. Extensão do teste para normalidade univariado baseado no coeficiente de correlação quantil-quantil para o caso multivariado. **Revista de Matemática e Estatística**, São Paulo, v.21, n.3, p.67-84, 2003.

COCHRAN, W. G. The analysis of variances when experimental errors follow the Poisson or Binomial laws. **The Annals of Mathematical Statistics**, v.11, p.335–347, 1940.

CRAWLEY, M.J. **The R book**. England: Wiley, 2007. 942 p.

COOK, R. D. Detection of influential observations in Linear Regression, **Technometrics**, v.19, p.15-18, 1977.

D'AGOSTINO, R. B.; STEPHENS, M.A. **Goodness-of-fit techniques**. New York: M. Dekker, 1986. 576 p.

DOBSON, A. J.; BARNETT, A. G. **An introduction to Generalized Linear Models**, 3rd ed., New York: Chapman & Hall, 2008. 307 p.

FARACO, M. A.; URIBE-OPAZO, M. A.; SILVA, E. A. A.; JOHANN, J. A.; BORSSOI, J. A. Seleção de modelos de variabilidade espacial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. **Revista Brasileira de Ciência do Solo**, v.32, p.463-476, 2008.

FARAWAY, J. J. **Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models**. Florida: Chapman and Hall, 2006. 301 p.

FERRARI, S.; CRIBARI-NETO, F. Beta regression for modeling rates and proportions. **Journal of Applied Statistics**, Abingdon, v. 31, p. 799–815, 2004.

FINNEY, D. J. Was this in your statistical textbook? V. Transformation of data. **Experimental Agricultural**, v.25, p.165-175, 1989.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society**, v.222, p.309-368, 1922.

FISHER, R. A. **Statistical methods for research workers**. Edinburgh: Oliver and Boyd, 1925. 319 p.

FLORIANO, E. P.; MÜLLER, I.; FINGER, C. A. G; SCHNEIDER, P. R. Ajuste e seleção de modelos tradicionais para série temporal de dados de altura de árvores. **Ciência Florestal**, Santa Maria, v.16, n.2, p.177-199, 2006.

FONT, H.; TORRES, V.; HERRERA, M.; RODRÍGUEZ, R. Fulfillment of the normality and the homogeneity of the variance in frequencies of accumulated measurement of the egg production variable in White Leghorn hens. **Cuban Journal of Agricultural Science**, v.41, p.207-211, 2007.

GLASS, G. Y.; PECKHAM, P. D.; SANDERS, J. R. Consequences of failure to meet the assumptions underlying the fixed effects analysis of variance and covariance. **Review of Educational Research**, v.42, p.237-288, 1972.

GOMEZ, K. A; GOMEZ, A. A. **Statistical procedures for agricultural research**, 2nd ed. New York: Wiley, 1984. 680 p.

GOTELLI, N. J.; ELLISON, A. M. **A primer of ecological statistics**. Massachusetts: Sinauer Associates, 2004. 492 p.

HARWELL, M. R. Summarizing Monte Carlo results in methodological research. **Journal of Educational Statistics**, v.17, p.297-313, 1992.

INTERNATIONAL SEED TESTING ASSOCIATION. **ISTA method validation for seed testing**. Bassersdorf: ISTA, 2007. 70 p.

KHAN, A.; RAYNER G.D. Robustness to non-normality of common tests for the many-sample location problem. **Journal of Applied Mathematics and Decision Sciences**, v.7, p.187-206, 2003.

KEMPTHORNE, O. **Design and analysis of experiments**. New York: Wiley, 1952. 631 p.

LABOURIAU, L. G. **A germinação de sementes**. Washington: Secretaria Geral da Organização dos Estados Americanos, 1983. 175 p.

LANTZ, B. The impact of continuity violation on ANOVA and alternative methods. **Journal of Modern Applied Statistical Methods**, v.12, n.2, p.105-120, 2013a.

LANTZ, B. The impact of sample non-normality on ANOVA and alternative methods. **British Journal of Mathematical and Statistical Psychology**, v.66, p.224-244, 2013b.

LEE, Y.; NELDER, J. A.; PAWITAN, Y. **Generalized Linear Models with Random Effects**. New York: Chapman & Hall, 2006. 380 p.

LEVINE, D. W.; DUNLAP, W. P. Data transformation, power, and skew: A rejoinder to Games. **Psychological Bulletin**, v.93, p.596-599, 1983.

LITTLE, T. M.; HILLS, F. J. **Agricultural experimentations: Design and analysis**. New York: Wiley, 1978. 350 p.

LIX, L. M; KESELMAN, J. C.; KESELMAN, H. J. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance *F* test. **Review of Educational Research**, v.66, p.579-619, 1996.

LÚCIO, A. D.; SCHWERTNER, D. V.; HAESBAERT, F. M.; SANTOS, D.; BRUNES, R. R.; RIBEIRO, A. L. P.; LOPES, S. J. Violação dos pressupostos do modelo matemático e transformação de dados. **Horticultura Brasileira**, v.30, p.415-423, 2012.

MONTGOMERY, D. C. **Design and analysis of experiments**, 5th ed. New York: John Wiley & Sons, 2000.

MORA, F.; GONÇALVES, L. M.; SCAPIM, C. A.; MARTINS, E. L.; MACHADO, M. F. P. S. Generalized Lineal Models for the Analysis of Binary Data from Propagation Experiments of Brazilian Orchids. **Brazilian archives of Biology and Technology**, v. 51, n. 5, p.963-970, 2008.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. **Journal of the Royal Statistical Society**, v.135, p.370-384, 1972.

NETER, J.; KUTNER, M.; NACHTSHEIM, C.; WASSERMAN, W. **Applied Linear Statistical Models**, 4th ed. New York: McGraw-Hill Companies, 1996. 1408 p.

PAOLINO, P. Maximum likelihood estimation of models with beta-distributed dependent variables. **Political Analysis**, v.9, n.4, p.325-346, 2001.

PEÑA, S. de R. D. **Modelos lineales y series temporales**. Madrid: Alianza Editorial, 1994. 745 p.

PIMENTEL-GOMES, F. **Curso de estatística experimental**, 14^a ed. Piracicaba: Nobel, 2000. 477 p.

PREGIBON, D. Logistic Regression diagnostics. **Annals of Statistics**, v.9, p.705-724, 1981.

RAO, M. M. Some asymptotic results on transformations in the analysis of variance. **ARL Technical Note**, p.60-126, 1960.

RASMUSSEN, J. L. Data transformation maximizing homocedasticity and within-group normality. **Behavior Research Methods, Instruments, & Computers**, v.17, n.3, p.411-412, 1985.

RESENDE, M. D. V.; BIELE, J. Estimação e predição em modelos lineares generalizados mistos com variáveis binomiais. **Revista de Matemática e Estatística**, São Paulo, v.20, p.39-65, 2002.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. Reviews: Akaike Information Criterion statistics. **Psychometrika**, v.54, n.3, p.539-541, 1989.

SCHEFFÉ, H. **The analysis of variance**. Wiley: New York, 1959. 470p.

SCHMIDER, E.; ZIEGLER, M.; DANAY, E.; BEYER, L.; BÜHNER, M. Is it really robust? Reinvestigating the robustness of ANOVA against violations of the Normal Distribution assumption. **Methodology**, v.6, n.4, p.147-151, 2010.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v.6, n.2, p.461-464, 1978.

SHARPE, K. Robustness of normal tolerance intervals. **Biometrika**, v.57, n.1, p.71-78, 1970.

SILESHI, G. W. A critique of current trends in the statistical analysis of seed germination and viability data. **Seed Science Research**, v.22, p.145-159, 2012.

SILESHI, G. W. Evaluation of statistical procedures for efficient analysis of insect, disease and weed abundance and incidence data. **East African Journal of Science**, v.1, p.1-9, 2007.

SOKAL, R. R.; ROHLF, F. J. **Biometry: the principles and practice of statistics in biological research**, 3rd edition. New York: W. H. Freeman, 1995. 937 p.

STEEL, R. G. D.; TORRIE, J. H. **Principles and procedures of statistics**. New York: McGraw-Hill, 1980. 633 p.

STORCK, L.; GARCIA, D. C.; LOPES, S. J.; ESTEFANEL, V. **Experimentação vegetal**. Santa Maria: UFSM, 2011. 200 p.

TORTAN, V. B. L.; COSTER R.; RIBOLDI, J. Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação. **Revista HCPA**, v.32, n.2, p.227-234, 2012.

THODE, H. C. **Testing for normality**. New York: Marcel Dekker, 2002. 368 p.

VIEIRA, S. **Análise de variância (ANOVA)**. São Paulo: Editora Atlas, 2006. 204 p.

WARTON, D.; HUI, F. The arcsine is asinine: the analysis of proportions in ecology. **Ecology**, v.92, p.3-10, 2011.

WILSON, K.; HARDY, I. C. W. Statistical analysis of sex ratios: an introduction. In: HARDY, I. C. W. **Sex ratios: Concepts and Research Methods**. United Kingdom: Cambridge University Press, p.49-92, 2002.

WINER, B. J.; BROWN, D. R.; MICHELS, K. M. **Statistical principles in experimental design**. New York: McGraw-Hill, 1971. 928 p.

WILLIAMS, D. A. Generalized Linear Model diagnostics using the deviance and single case deletions. **Applied Statistics**, v.36, n.2, p.181-191, 1987.

ZAR, J. H. **Biostatistical analysis**, 4th ed. New Jersey: Prentice Hall, 1998. 662 p.

ANEXOS

Anexo A. ANODEV para percentual de germinação (Distribuição Normal- função identidade) e número de sementes (Distribuição Binomial- função logit) para modelo com dois fatores (laboratório e lote) e interação para as espécies que atenderam as pressuposições de normalidade e homogeneidade.

Albizia hassleri									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	91	21708.3				556.47		
Lab.	7	84	338.6	21369.7	1.511	0.178	8.00	548.46	0.333
Lote	2	82	18893.3	2476.4	295.14	0.000	481.91	66.56	< 0.001***
Lab*Lote	14	68	299.9	2176.5	0.669	0.796	11.26	55.29	0.665
Saturado	23	68	2176.5				55.29		
			$\hat{\phi}$: 32.00735	AIC: 602.14		AIC: 513.36			
Anadenanthera macrocarpa									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	95	38199.0				1004.66		
Lab.	7	88	950.0	37249.0	2.885	0.011	25.27	979.39	0.001***
Lote	2	86	32576.0	4493.0	348.08	< 0.001***	831.31	148.09	< 0.001***
Lab*Lote	14	72	1105.0	3388.0	1.678	0.079	48.14	99.95	< 0.001***
Saturado	23	72	3388.0				99.95		
			$\hat{\phi}$: 47.05208	AIC:664.54		AIC: 565.78			
Astronium fraxinifolium									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	30155.0				851.06		
Lab.	5	66	388.7	29766.2	1.937	0.103	11.51	839.55	0.042*
Lote	2	64	27092.2	2674.1	337.52	< 0.001***	761.10	78.45	< 0.001***
Lab*Lote	10	54	506.8	2167.2	1.263	0.275	15.86	62.59	0.104
Saturado	17	54	2167.2				62.59		
			$\hat{\phi}$: 40.13626	AIC:487.45		AIC: 413.69			

Cariniana estrellensis									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	81	20662.5				465.53		
Lab.	6	75	767.0	19895.5	2.408	0.037*	17.29	448.23	0.008**
Lote	2	73	15388.2	4507.4	144.91	< 0.001***	343.78	104.46	< 0.001***
Lab*Lote	12	61	1268.6	3238.8	1.991	0.041*	30.53	73.93	0.002**
Saturado	20	61	3238.8				73.93		
			$\hat{\phi}$: 53.09426 AIC:578.16				AIC: 473.79		
Cariniana legalis									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	25894.0				283.67		
Lab.	5	66	1069.5	24824.5	5.969	0.001***	10.18	273.49	< 0.001***
Lote	2	64	22176.6	2647.9	309.44	< 0.001***	237.12	36.37	< 0.001***
Lab*Lote	10	54	712.9	1935.0	1.990	0.053	8.31	28.07	< 0.001***
Saturado	17	54	1935.0				28.07		
			$\hat{\phi}$: 35.8333 AIC:479.29				AIC: 313.96		
Ceiba speciosa									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	83	32852.0				895.29		
Lab.	6	77	1399.1	31453.0	2.942	0.014*	35.84	859.45	< 0.001***
Lote	2	75	25811.6	5641.0	162.83	< 0.001***	687.49	171.97	< 0.001***
Lab*Lote	12	63	647.9	4994.0	0.681	0.763	14.44	157.52	0.273
Saturado	20	63	4994.0				157.52		
			$\hat{\phi}$: 53.09426 AIC:578.16				AIC: 567.2		
Copaifera langsdorffii									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	80	16998.5				589.33		
Lab.	6	74	199.8	16798.6	1.355	0.248	6.17	583.16	0.405
Lote	2	72	15154.9	1643.8	308.20	< 0.001***	517.73	65.43	< 0.001***
Lab*Lote	11	60	168.6	1475.2	0.572	0.856	10.75	54.68	0.551
Saturado	19	60	1475.2				54.68		
			$\hat{\phi}$: 24.58611 AIC:508.94				AIC: 410.71		

<i>Cordia americana</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	24758.9				685.72		
Lab.	5	66	1889.0	22869.9	3.253	0.012*	46.46	639.26	< 0.001***
Lote	2	64	15609.1	7260.8	67.192	< 0.001***	436.11	203.15	< 0.001***
Lab*Lote	10	54	988.6	6272.2	0.851	0.583	30.69	172.45	< 0.001***
Saturado	17	54	6272.2				172.45		
			$\hat{\phi}$: 116.1528 AIC:563.97				AIC: 512.26		
<i>Dalbergia miscolobium</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	19780.9				472.03		
Lab.	5	66	1497.3	18283.6	3.300	0.011*	34.324	437.71	< 0.001***
Lote	2	64	12690.8	5592.8	69.932	< 0.001***	304.15	133.56	< 0.001***
Lab*Lote	10	54	693.1	4899.8	0.764	0.662	15.40	118.16	0.118
Saturado	17	54	4899.8				118.16		
			$\hat{\phi}$: 90.73611 AIC:546.19				AIC: 467.32		
<i>Enterolobium maximum</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	198000.3				535.48		
Lab.	5	66	1482.4	18317.9	2.840	0.023*	34.212	501.27	< 0.001***
Lote	2	64	10817.9	7500.1	51.804	< 0.001***	282.74	218.53	< 0.001***
Lab*Lote	10	54	1861.8	5638.2	1.783	0.086	54.57	163.96	< 0.001***
Saturado	17	54	5638.2				163.96		
			$\hat{\phi}$: 104.412 AIC:556.29				AIC: 520.46		
<i>Erytrina speciosa</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	68	24982.6				718.85		
Lab.	5	63	994.4	23988.2	3.905	0.004*	26.88	691.97	< 0.001***
Lote	2	61	19961.1	4027.1	195.97	< 0.001***	546.35	145.62	< 0.001***
Lab*Lote	10	51	1429.6	2597.4	2.807	0.007*	46.67	98.96	< 0.001***
Saturado	17	51	2597.4				98.96		
			$\hat{\phi}$: 50.92974 AIC:484.16				AIC: 434.21		

<i>Gallesia integrifolia</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	23461.7				698.85		
Lab.	5	66	1113.4	22348.2	2.800	0.025*	28.18	670.67	< 0.001***
Lote	2	64	17741.4	4606.8	111.55	< 0.001***	503.71	166.96	< 0.001***
Lab*Lote	10	54	312.6	4294.2	0.393	0.944	8.16	158.80	0.613
Saturado	17	54	4294.2				158.80		
$\hat{\phi}$: 79.52315 AIC:536.69							AIC: 512.23		
<i>Guazuma ulmifolia</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	65	12921.8				369.29		
Lab.	5	60	299.3	12622.5	2.459	0.046*	8.37	360.92	0.137
Lote	2	58	11280.2	1342.3	231.77	< 0.001***	322.09	38.84	< 0.001***
Lab*Lote	10	48	174.2	1168.1	0.716	0.705	3.48	35.36	0.968
Saturado	17	48	1168.1				35.36		
$\hat{\phi}$: 24.33507 AIC:414.95							AIC: 361.77		
<i>Handroanthus impetiginosus</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	17949.5				423.4		
Lab.	5	66	496.5	17453.0	1.975	0.097	10.0	413.41	0.077
Lote	2	64	14312.2	3140.8	142.31	< 0.001***	340.1	73.33	< 0.001***
Lab*Lote	10	54	425.2	2715.5	0.846	0.588	9.6	63.70	0.474
Saturado	17	54	2715.5				63.7		
$\hat{\phi}$: 50.28704 AIC:503.69							AIC: 414.95		
<i>Handroanthus roseo-albus</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	43133.0				1101.23		
Lab.	5	65	735.0	42397.0	2.416	0.048*	14.74	1086.49	0.011*
Lote	2	63	38985.0	3413.0	320.22	< 0.001***	967.31	119.18	< 0.001***
Lab*Lote	10	53	187.0	3226.0	0.307	0.976	5.49	113.68	0.855
Saturado	17	53	3226.0				113.68		
$\hat{\phi}$: 60.87107 AIC:510.45							AIC: 445.8		

<i>Hymenaea courbaril</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	66	22969.9				115.09		
Lab.	5	61	713.5	22256.4	3.386	0.011*	3.361	11.65	0.645
Lote	2	59	19564.0	2692.3	232.15	< 0.001***	96.76	14.89	< 0.001***
Lab*Lote	10	49	627.7	2064.7	1.490	0.172	3.57	11.32	0.965
Saturado	17	49	2064.7				11.32		
$\hat{\phi}$: 42.13605 AIC:457.82							AIC: 368.34		
<i>Jacaranda micrantha</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	69	33478.0				987.37		
Lab.	5	64	461.2	33017.0	1.765	0.136	11.12	976.25	0.049*
Lote	2	62	29528.2	3489.0	282.55 8	< 0.001***	823.98	152.27	< 0.001***
Lab*Lote	10	52	771.6	2717.0	1.477	0.175	31.51	120.76	< 0.001***
Saturado	17	52	2717.0				120.76		
$\hat{\phi}$: 52.2516 AIC:492.77							AIC: 444.62		
<i>Mimosa scabrella</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	114	17283.1				419.64		
Lab.	9	105	896.5	16386.6	1.787	0.082	21.537	398.10	0.010*
Lote	2	103	10511.6	5875.0	94.290	< 0.001***	263.87	134.23	< 0.001***
Lab*Lote	18	85	1137.1	4737.9	1.133	0.336	25.77	108.46	0.105
Saturado	29	85	4737.9				108.46		
$\hat{\phi}$: 55.7402 AIC:815.97							AIC: 680.47		
<i>Ormosia arborea</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	21128.0				480.93		
Lab.	5	66	884.7	20243.0	2.657	0.032*	18.343	462.59	0.002**
Lote	2	64	14660.0	5583.0	110.08	< 0.001***	314.23	148.35	< 0.001***
Lab*Lote	10	54	1987.0	3596.0	2.984	0.004**	54.16	94.19	< 0.001***
Saturado	17	54	3596.0				94.19		
$\hat{\phi}$:66.59259 AIC:523.91							AIC: 442.67		

<i>Peltophorum dubium</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	93	33048.0				822.99		
Lab.	7	86	727.3	32321.0	3.243	0.004**	19.29	803.70	0.007**
Lote	2	84	29049.3	3272.0	453.34	< 0.001***	709.07	94.62	< 0.001***
Lab*Lote	14	70	1029.2	2243.0	2.294	0.011*	29.06	65.57	0.010*
Saturado	23	70	2243.0				65.57		
			$\hat{\phi}$: 32.03929 AIC:614.94				AIC: 525.18		
<i>Pseudobombax tomentosum</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	83	69091.0				2096.60		
Lab.	6	77	678.0	68413.0	2.933	0.013*	18.49	2078.11	0.005**
Lote	2	75	65083.0	3330.0	844.36	< 0.001***	1947.00	130.55	< 0.001***
Lab*Lote	12	63	902.0	2428.0	1.950	0.044*	41.49	89.07	< 0.001***
Saturado	20	63	2428.0				89.07		
			$\hat{\phi}$: 38.53968 AIC:564.96				AIC: 472.72		
<i>Pterogyne nitens</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	17837.5				445.25		
Lab.	5	65	440.4	17397.0	2.755	0.027*	11.60	433.65	0.040*
Lote	2	63	15069.1	235.7	235.67	< 0.001***	372.25	60.90	< 0.001***
Lab*Lote	10	53	633.4	1694.4	1.982	0.054	16.60	44.30	0.083
Saturado	17	53	1694.4				44.30		
			$\hat{\phi}$: 31.97013 AIC:464.73				AIC: 390.73		
<i>Senna macranthera</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	90	26443.8				622.47		
Lab.	7	83	863.6	25580.2	2.172	0.047*	20.34	602.13	0.004**
Lote	2	81	21296.4	4283.7	187.48	< 0.001***	498.45	103.68	< 0.001***
Lab*Lote	14	67	478.5	3805.2	0.602	0.854	11.65	92.03	0.634
Saturado	23	67	3805.2				92.03		
			$\hat{\phi}$: 85.49673 AIC:533.98				AIC: 519.9		

<i>Senna multijuga</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	67	27091.0				770.67		
Lab.	5	62	944.2	26146.8	5.486	< 0.001***	23.89	746.79	< 0.001***
Lote	2	60	24272.0	1874.7	352.55	< 0.001***	695.16	51.63	< 0.001***
Lab*Lote	10	50	153.6	1721.2	0.446	0.916	4.22	47.41	0.937
Saturado	17	50	1721.2				47.41		
$\hat{\phi}$: 34.42333 AIC:450.7							AIC: 366.36		
<i>Stryphnodendron barbadetiman</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	20855.9				621.95		
Lab.	5	66	263.3	20592.6	1.504	0.204	5.9	616.05	0.316
Lote	2	64	18341.3	2251.2	261.98	< 0.001***	546.18	69.88	< 0.001***
Lab*Lote	10	54	361.0	1890.2	1.031	0.431	12.04	57.84	0.283
Saturado	17	54	1890.2				57.84		
$\hat{\phi}$: 35.00463 AIC:477.61							AIC: 410.71		
<i>Stryphnodendron polyphyllum</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	31733.0				923.00		
Lab.	5	65	693.5	31041.0	2.882	0.025*	17.35	905.65	0.003**
Lote	2	63	28334.7	2706.0	294.79	< 0.001***	822.13	83.52	< 0.001***
Lab*Lote	10	43	159.0	2547.0	0.331	0.969	3.61	79.91	< 0.001***
Saturado	17	43	2547.0				79.91		
$\hat{\phi}$: 48.05975 AIC:493.67							AIC: 424.62		
<i>Tabebuia aurea</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	68	14511.0				443.65		
Lab.	5	63	141.8	14369.2	0.879	0.501	3.6	440.04	0.608
Lote	2	61	12290.8	2078.4	190.51	< 0.001***	364.03	76.01	< 0.001***
Lab*Lote	10	51	433.3	1645.1	1.343	0.234	22.89	53.12	0.011*
Saturado	17	51	1645.1				53.12		
$\hat{\phi}$: 32.25654 AIC:452.64							AIC: 375.39		

<i>Zeyeria tuberculosa</i>									
			Normal				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	27542.0				838.67		
Lab.	5	65	459.2	25082.7	1.839	0.121	10.00	828.67	0.075
Lote	2	63	23985.4	3097.3	240.07	< 0.001***	727.96	100.71	< 0.001***
Lab*Lote	10	53	449.7	2647.7	0.090	0.540	18.18	82.53	0.052
Saturado	17	53	2647.7				82.53		
			AIC: 496.42				AIC: 426.18		

Anexo B. ANODEV para percentual de germinação (Distribuição Normal- função identidade), transformação arcoseno do percentual de germinação (Distribuição Normal- função identidade) e número de sementes (Distribuição Binomial- função logit) para modelo com dois fatores (laboratório e lote) e interação para as espécies que não atenderam as pressuposições de normalidade e homogeneidade.

<i>Acacia polyphylla</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	29359.8				13056.0				823.5		
Lab.	5	66	452.4	28907.3	2.220	0.0655	265.9	12790.1	3.068	0.017*	12.11	811.39	0.033*
Lote	2	64	26547.2	2360.1	325.585	< 0.001***	11703.2	1086.9	337.577	< 0.001***	739.61	71.78	< 0.001***
Lab*Lote	10	54	158.6	2201.5	0.389	0.946	150.9	936.0	0.870	0.566	11.35	60.44	0.331
Saturado	17	54	2201.5				936.0				60.44		
			$\hat{\phi}$: 40.76852	AIC: 488.58					$\hat{\phi}$: 17.33408	AIC: 427.01	AIC: 390.26		

<i>Anadenanthera colubrina</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	33477.0				13063.5				856.24		
Lab.	5	66	427.6	33049.0	1.636	0.166	199.000	12864.5	1.776	0.134	8.76	847.48	0.119
Lote	2	64	29645.6	3404.0	283.564	< 0.001***	11382.6	1481.9	253.887	< 0.001***	756.18	91.30	< 0.001***
Lab*Lote	10	54	580.9	2823.0	1.111	0.371	271.400	1210.5	1.211	0.305	16.27	75.03	0.092
Saturado	17	54	2823.0				1210.5				75.03		
			$\hat{\phi}$: 52.27315	AIC: 506.48					$\hat{\phi}$: 22.41667	AIC: 445.52	AIC: 415.55		

Apuleia leiocarpa													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	14134.8				6738.2				471.49		
Lab.	5	65	1312.7	12822.1	3.687	0.006**	689.0	6049.2	4.110	0.003*	48.41	423.08	< 0.001***
Lote	2	63	8570.3	4251.7	60.176	< 0.001***	3947.2	2102.0	58.8637	< 0.001***	293.71	129.37	< 0.001***
Lab*Lote	10	53	477.6	3774.2	0.671	0.746	291.4	1810.5	0.869	0.567	19.17	110.20	0.038*
Saturado	17	53	3774.2				1810.5				110.20		
			$\hat{\phi}$: 71.21069	AIC:521.59			$\hat{\phi}$: 33.52863	AIC:474.51			AIC: 443.24		

Cedrela fissilis													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	83	42951.0				18739.6				1077.74		
Lab.	6	77	1095.0	41857.0	3.011	0.012*	476.100	18263.400	2.949	0.013*	23.55	1054.19	< 0.001***
Lote	2	75	36861.0	4996.0	304.157	< 0.001***	15834.7	2428.7	294.231	< 0.001***	907.81	146.38	< 0.001***
Lab*Lote	12	63	1178.0	3818.0	1.620	0.109	733.400	1695.200	2.271	0.018*	42.44	103.93	< 0.001***
Saturado	20	63	3818.0				1695.200				103.93		
			φ:60.59524 AIC:602.97				φ:26.90873 AIC: 534.78				AIC: 487.46		

Cassia leptophylla													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	25716.8				9928.4				730.31		
Lab.	5	65	598.4	25118.4	1.759	0.138	328.800	9599.600	2.289	0.059	13.99	716.32	0.016*
Lote	2	63	20567.7	4550.7	151.101	< 0.001***	7647.3	1952.4	133.084	< 0.001***	596.07	120.24	< 0.001***
Lab*Lote	10	53	943.5	3607.2	1.386	0.212	400.900	1551.500	1.395	0.208	22.52	97.73	0.012*
Saturado	17	53	3607.2				1551.5				97.73		
			φ : 68.05975 AIC:518.30				φ : 28.73099 AIC:463.39				AIC: 473.79		

Cedrela odorata													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	105	32889.0				15708.9				966.88		
Lab.	8	97	558.5	32330.0	2.194	0.037*	251.300	251.300	1.811	0.087	13.1	953.77	0.108
Lote	12	95	29212.3	3118.0	458.969	< 0.001***	13787.1	13787.1	397.271	< 0.001***	852.41	101.36	< 0.001***
Lab*Lote	16	79	604.0	2514.0	1.186	0.297	299.600	299.600	1.079	0.388	19.30	82.06	0.254
Saturado	36	79	2514.0				1370.8						
			φ: 31.82384 AIC: 692.43				φ: 17.35232 AIC: 628.18				AIC: 560.31		

<i>Citharexylum myrianthum</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	18184.6				7847				516.34		
Lab.	5	65	457.1	17727.5	2.262	0.062	228.600	7618.400	2.686	0.030*	13.89	502.46	0.016*
Lote	2	63	15130.9	2596.6	187.150	< 0.001***	6502	1116.4	191.044	< 0.001***	427.03	75.43	< 0.001***
Lab*Lote	10	53	454.1	2142.5	1.123	0.363	214.500	901.900	1.260	0.276	14.73	60.70	0.142
Saturado	17	53	2142.5				901.9				60.70		
			$\hat{\phi}$: 40.42453 AIC:481.39				$\hat{\phi}$: 17.01715 AIC:419.96				AIC: 390.5		

Cybistax antisyphilitica													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	105	104154.0				47141				2877.62		
Lab.	8	97	479.0	103675.0	1.516	0.165	238.000	46902.000	1.180	0.322	10.4	2867.18	0.238
Lote	2	95	99577.0	4099.0	1260.83	< 0.001***	44292	2610	876.923	< 0.001***	2706.97	160.24	< 0.001***
Lab*Lote	16	79	979.0	3120.0	1.550	0.103	615.000	1995.000	1.522	0.113	37.06	123.18	0.002*
Saturado	20	79	3120.0				1995				123.18		
			φ:39.4884 AIC:715.31				φ:25.25444 AIC:669.93				AIC: 586.83		

Dalbergia nigra													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	18203.8				11222.3				758.63		
Lab.	5	66	197.1	18006.7	1.184	0.329	143.2	11079.1	1.348	0.259	7.09	751.54	0.214
Lote	2	64	16007.7	1999.0	240.3	< 0.001***	9813.1	1265.9	230.881	< 0.001***	667.62	83.92	< 0.001***
Lab*Lote	10	54	200.5	1798.5	0.602	0.805	118.4	1147.6	0.557	0.841	8.00	75.92	0.628
Saturado	17	54	1798.5				1147.6				75.92		
			φ: 33.30556 AIC:474.03				φ: 21.25158 AIC: 441.68				AIC: 394.08		

Enterolobium contortisiliquum													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	126	136205.0				70376				3849.40		
Lab.	10	116	1968.0	134237.0	7.749	< 0.001***	447	69929	1.976	0.0438*	42.4	3807.00	< 0.001***
Lote	2	114	130923.0	3314.0	2577.60	< 0.001***	66951	2977	1478.86	< 0.001***	3661.50	145.40	< 0.001***
Lab*Lote	20	94	927.0	2387.0	1.827	0.029*	736.000	2241.000	1.627	0.060	38.60	106.80	0.007**
Saturado	32	94	2387.0				2241				106.80		
			φ: 25.39628 AIC:800.99				φ: 22.6361 AIC:816.41				AIC: 632.33		

<i>Handroanthus chrysotrichus</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	90	72518.0				35405				1985.16		
Lab.	7	83	510.0	72009.0	2.421	0.028*	254.000	35151.000	1.640	0.138	11.51	1973.65	0.118
Lote	2	81	69559.0	2450.0	1156.92	< 0.001***	33124	2028	748.978	< 0.001***	1893.22	80.43	< 0.001***
Lab*Lote	14	67	436.0	2014.0	1.035	0.431	435.000	1592.000	1.407	0.1725	20.75	59.69	0.108
Saturado	23	67	2014.0				1592				59.69		
			φ: 30.06219 AIC:590.08				φ: 22.11259 AIC:592.05				AIC: 451.73		

Jacaranda cuspidifolia													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	78	25480.7				9928.4				660.28		
Lab.	6	72	746.3	24734.5	3.889	0.002*	330.000	9598.400	4.447	< 0.001***	18.96	641.32	0.004*
Lote	2	79	22413.1	2321.4	350.362	< 0.001***	8690.1	908.3	351.304	< 0.001***	580.87	60.45	< 0.001***
Lab*Lote	12	58	466.2	1855.2	1.215	0.296	191.000	717.400	1.287	0.251	12.84	47.61	0.381
Saturado	20	58	1855.2				717.4				47.61		
			φ: 31.98563 AIC:517.54				φ: 12.36835 AIC:442.48				AIC: 424.12		

Lafoensia pacari													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	28607.8				11488				878.88		
Lab.	5	66	302.4	28305.3	1.344	0.260	103.200	11384.800	1.026	0.412	10.51	868.37	0.061
Lote	2	64	25205.8	3099.6	280.064	< 0.001***	9991.5	1393.3	248.412	< 0.001***	769.22	99.15	< 0.001***
Lab*Lote	10	54	669.6	2430.0	1.488	0.169	307.300	1086.000	1.528	0.155	24.79	74.36	0.005**
Saturado	17	54	2430.0				1086				74.36		
			ϕ: 45 AIC:495.69				ϕ: 18.5603 AIC:575.24				AIC: 516.27		

<i>Mimosa caesalpiniaefolia</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	91	51236.0				24214.9				1633.56		
Lab.	7	84	297.0	50939.0	1.276	0.276	240.700	23.974.2	1.853	0.090	13.16	1260.39	0.068
Lote	2	82	48335.0	2604.0	726.734	< 0.001***	22530.5	1443.7	606.961	< 0.001***	1504.58	115.81	< 0.001***
Lab*Lote	14	68	343.0	2261.0	0.737	0.730	107.300	1366.300	0.413	0.966	21.10	94.71	0.099
Saturado	23	68	2261.0				1336.3				94.71		
			$\hat{\phi}$: 33.2549 AIC:605.66				$\hat{\phi}$: 18.5603 AIC:575.24				AIC: 516.27		

Parapiptadenia rigida													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	32620.0				18082.2				998.20		
Lab.	5	66	203.1	32417.0	0.538	0.747	55.700	18026.500	0.305	0.908	4.84	993.26	0.436
Lote	2	64	27867.9	4549.0	184.397	< 0.001***	15745.5	2281	215.707	< 0.001***	857.82	135.53	< 0.001***
Lab*Lote	10	54	469.0	4080.0	0.620	0.789	310.100	1970.900	0.850	0.584	18.74	116.79	0.043*
Saturado	17	54	4080.0				1970.9				116.79		
			ϕ: 75.56481 AIC:533.01				ϕ: 36.4975 AIC:480.62				AIC: 425.47		

Parkia pendula													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	62473.0				29575.8				2021.19		
Lab.	5	65	1129.0	61345.0	5.006	< 0.001***	721.3	28854.5	5.3076	< 0.001***	36.13	1976.06	< 0.001***
Lote	2	63	58099.0	3245.0	644.198	< 0.001***	26356.5	2498	484.849	< 0.001***	1813.02	163.04	< 0.001***
Lab*Lote	10	53	855.0	2390.0	1.896	0.066	1030.200	1467.700	3.791	< 0.001***	66.54	96.50	< 0.001***
Saturado	17	53	2390.0				1467.7				96.50		
			phi: 45.09434 AIC:489.15				phi: 27.18019 AIC:459.39				AIC: 409.71		

<i>Peltogyne confertiflora</i>													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	82	50498.0				26900.3				1516.33		
Lab.	6	76	218.0	50280.0	1.167	0.335	162.600	26737.700	1.568	0.171	4.92	1511.41	0.554
Lote	2	74	48093.0	2187.0	771.919	< 0.001***	25385.8	1351.9	734.577	< 0.001***	1436.94	74.46	< 0.001***
Lab*Lote	12	62	255.0	1931.0	0.683	0.761	263.300	1088.600	1.270	0.259	12.35	62.12	0.418
Saturado	20	62	1931.0				1088.6				62.12		
			φ:31.15188 AIC:540.76				φ:17.27918 AIC: 497.58				AIC: 412.34		

Platymenia reticulata													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	39284.0				17521.1				1186.84		
Lab.	5	66	985.0	38299.0	3.166	0.014*	535.400	16985.700	3.497	0.008**	25.31	1161.53	< 0.001***
Lote	2	64	34506.0	3792.0	277.263	< 0.001***	15113.1	1872.6	246.745	< 0.001***	1039.80	121.74	< 0.001***
Lab*Lote	10	54	432.0	3360.0	0.694	0.725	218.000	1653.700	0.715	0.707	21.85	99.89	0.015*
Saturado	17	54	3360.0				1653.7				99.89		
			φ: 62.22685 AIC:519.03				φ: 30.62496 AIC:467.98				AIC: 435.04		

Qualea grandiflora													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	48588.0				19618.3				1261.71		
Lab.	5	66	895.0	47693.0	3.148	0.014*	373.300	19245.000	2.977	0.019*	18.04	1243.67	0.002**
Lote	2	64	43822.0	3871.0	385.340	< 0.001***	17583.4	1661.7	350.524	< 0.001***	1137.88	105.79	< 0.001***
Lab*Lote	10	54	800.0	3070.0	1.408	0.202	307.400	1354.400	1.225	0.296	17.07	88.71	0.073
Saturado	17	54	3070.0				1354.4				88.71		
			ϕ: 56.86111 AIC:512.54				ϕ:25.08154 AIC:453.61				AIC: 420.64		

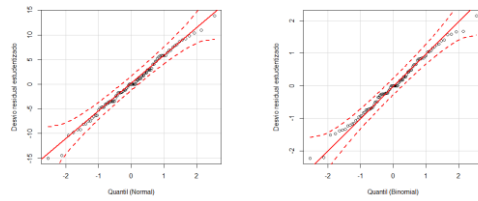
Schinus terebinthifolius													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	71	40676.0				15215.9				976.15		
Lab.	5	66	609.0	40067.0	2.479	0.032*	213.100	15002.800	1.686	0.139	12.29	963.86	0.055
Lote	2	64	37152.0	2915.0	453.730	< 0.001***	13479.3	1523.5	319.966	< 0.001***	885.31	78.55	< 0.001***
Lab*Lote	10	54	459.0	2456.0	0.934	0.520	196.400	1327.000	0.777	0.671	12.25	66.30	0.426
Saturado	17	54	2456.0				1327				66.30		
			ϕ : 40.94028 AIC:550.24				ϕ : 21.06369 AIC:514.21				AIC: 451.00		

Schizolobium parahyba var. amazonicum													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	80	17259.9				12969.2				721.22		
Lab.	6	74	618.9	16641.0	2.828	0.017*	395.400	12573.800	3.095	0.010*	19	702.22	0.004**
Lote	2	72	13524.7	3116.2	185.425	< 0.001***	10834.4	1739.5	254.371	< 0.001***	600.09	102.13	< 0.001***
Lab*Lote	12	60	928.1	2188.2	2.121	0.028*	461.700	1277.800	1.807	0.067	24.33	77.80	0.018*
Saturado	20	60	2188.2				1277.8				77.80		
			ϕ : 36.46944 AIC:540.87				ϕ : 21.29639 AIC:497.3				AIC: 401.28		

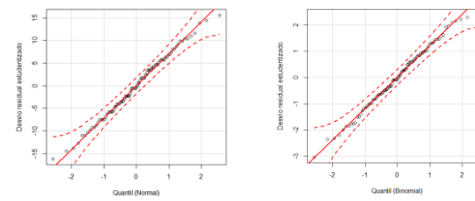
Stryphnodendron polyphyllum													
			Normal				Normal (Transformação Arcoseno)				Binomial		
F.V.	gl	Dif. de gl	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	F	valor-p	Deviance	Dif. de deviance	valor-p
Nulo	0	70	31733.0				11995.9				923.00		
Lab.	5	65	693.5	31041.0	2.882	0.025*	256.400	11739.400	2.414	0.047*	17.35	905.65	0.003**
Lote	2	63	28334.7	2706.0	294.786	< 0.001***	10534.6	1204.8	247.955	< 0.001***	822.13	83.52	< 0.001***
Lab*Lote	10	43	159.0	2547.0	0.331	0.969	57.600	1147.100	0.271	0.985	3.61	79.91	< 0.001***
Saturado	17	43	2547.0				1147.1				79.91		
			ϕ : 48.05975 AIC:493.67				ϕ : 21.24307 AIC: 441.65				AIC: 424.62		

Anexo C. Gráfico q-q plot do percentual e número de plântulas normais das espécies que atenderam as pressuposições de normalidade e homogeneidade., com as respectivas distribuições, Normal e Binomial.

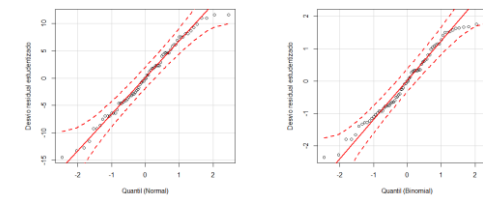
Albizia hassleri



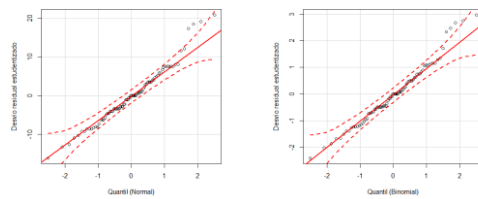
Anadenanthera macrocarpa



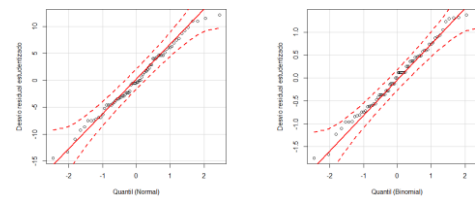
Astronium fraxinifolium



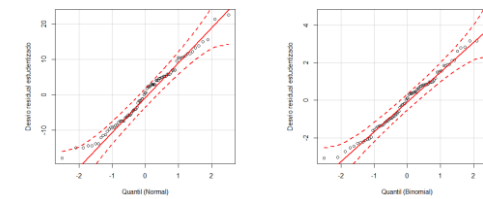
Cariniana estrellensis



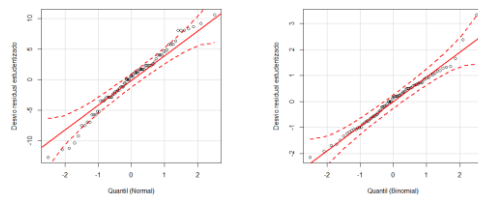
Cariniana legalis



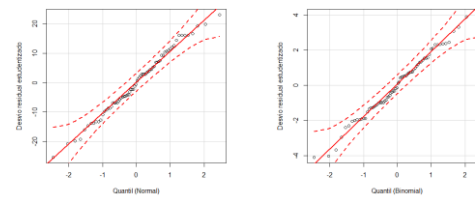
Ceiba speciosa



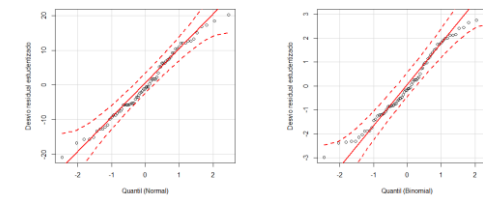
Copaifera langsdorffii



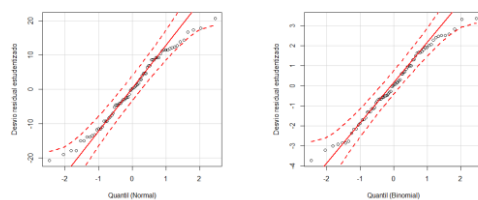
Cordia americana



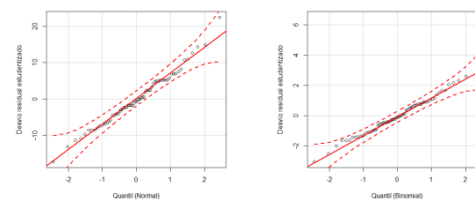
Dalbergia miscolobium



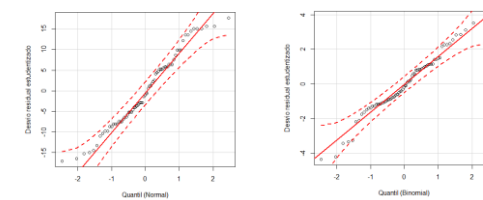
Enterolobium maximum



Erythrina speciosa



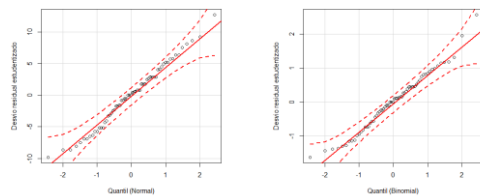
Gallesia integrifolia



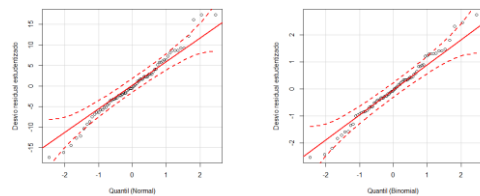
(...Continua...)

Continuação...

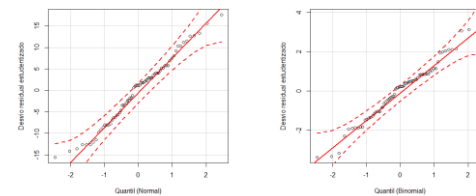
Guazuma ulmifolia



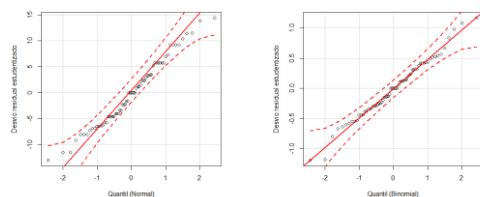
Handroanthus impetiginosus



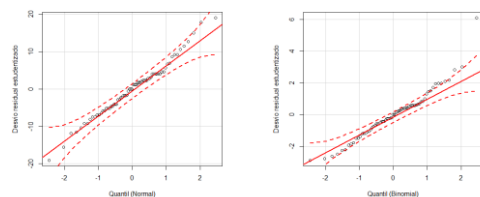
Handroanthus roseo-albus



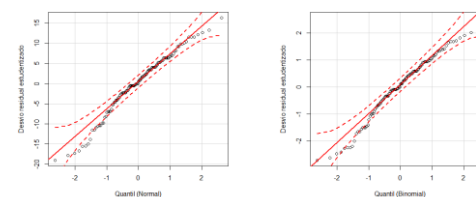
Hymenaea courbaril



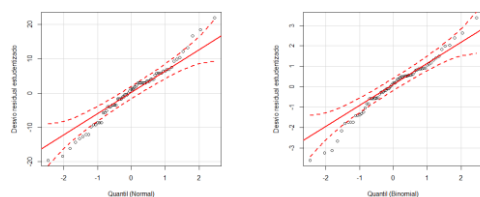
Jacaranda micrantha



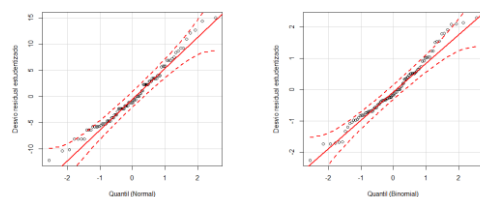
Mimosa scabrella



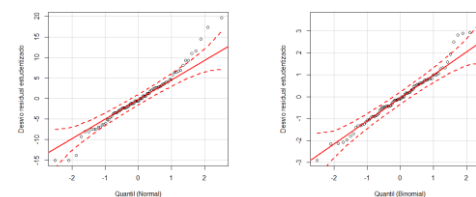
Ormosia arborea



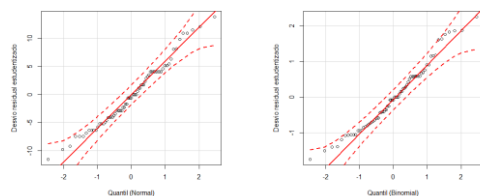
Peltophorum dubium



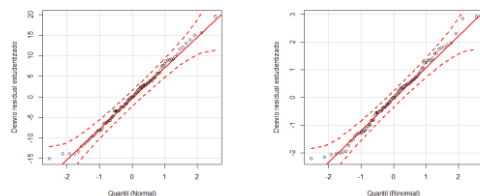
Pseudobombax tomentosum



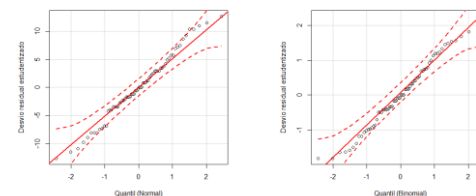
Pterogyne nitens



Senna macranthera



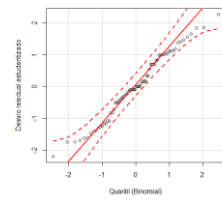
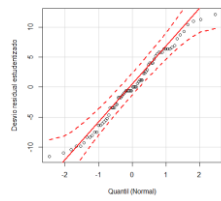
Senna multijuga



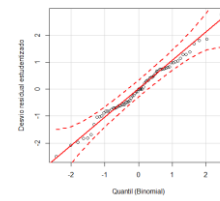
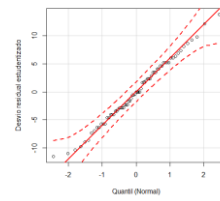
(...Continua...)

Continuação...

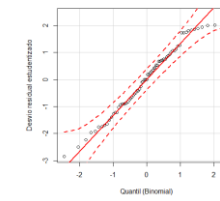
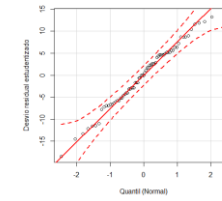
Stryphnodendron barbadetiman



Tabebuia aurea

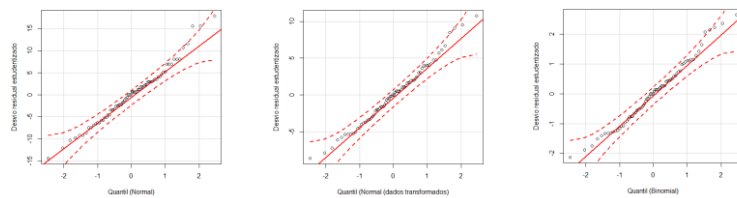


Zeyeria tuberculosa

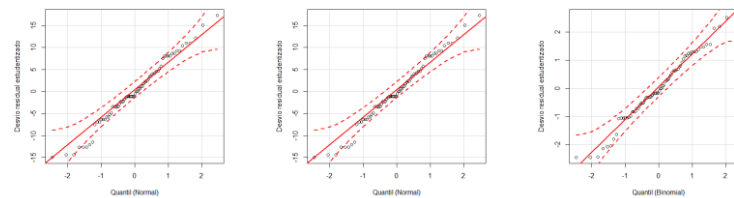


Anexo D. Gráfico q-q plot do percentual e número de plântulas normais das espécies que não atenderam as pressuposições de normalidade e homogeneidade., com as respectivas distribuições, Normal e Binomial.

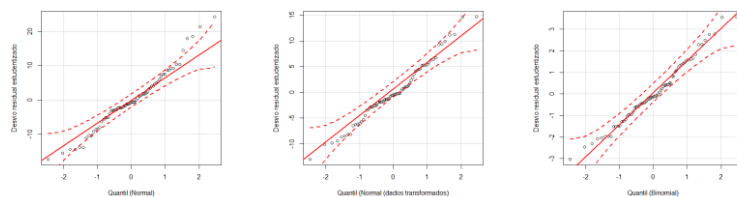
Acacia polyphylla



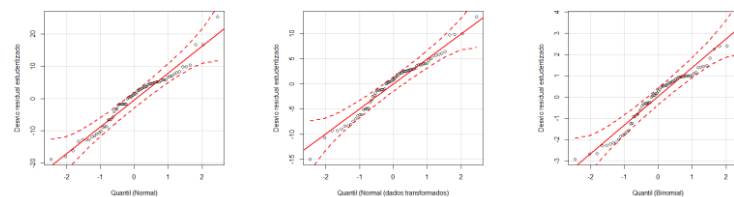
Anadenanthera colubrina



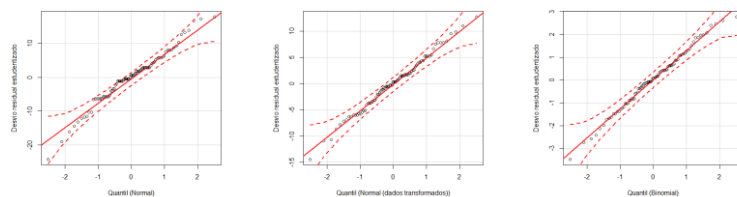
Apuleia leiocarpa



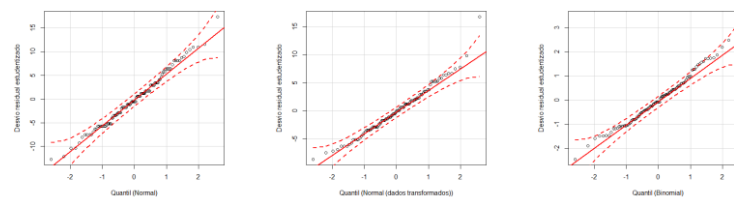
Cassia leptophylla



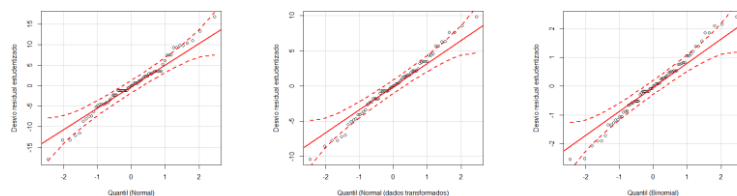
Cedrela fissilis



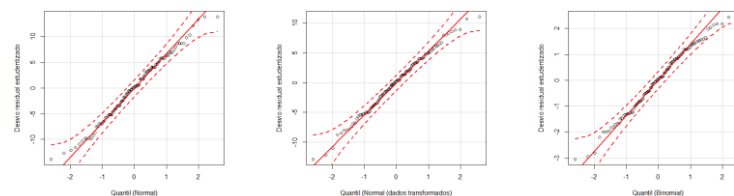
Cedrela odorata



Citharexylum myrianthum



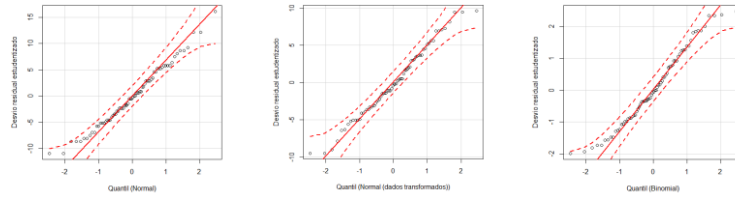
Cybistax antisiphilitica



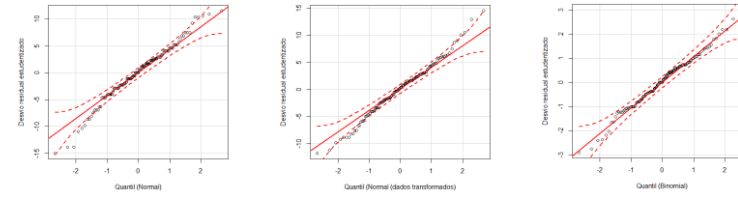
(...Continua...)

Continuação...

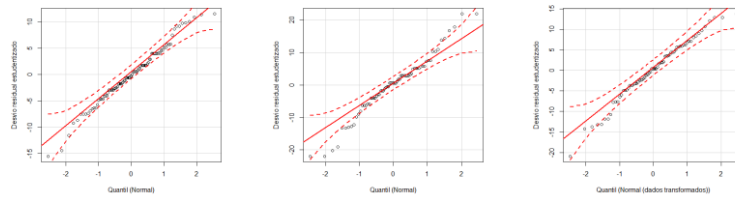
Dalbergia nigra



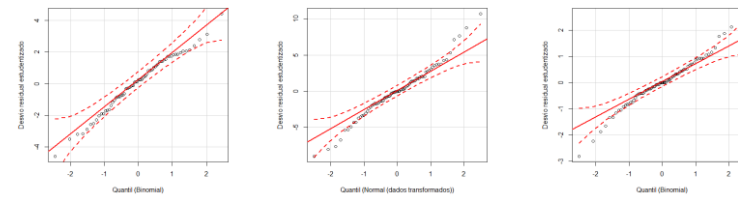
Enterolobium contortisiliquum



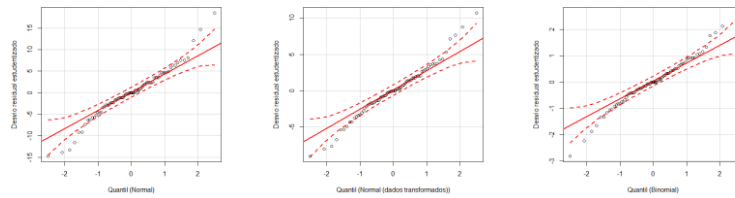
Handroanthus chrysotrichus



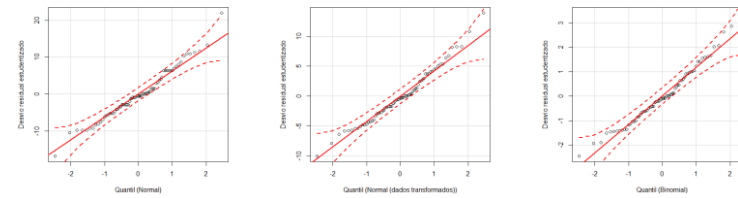
Hymenaea stigonocarpa



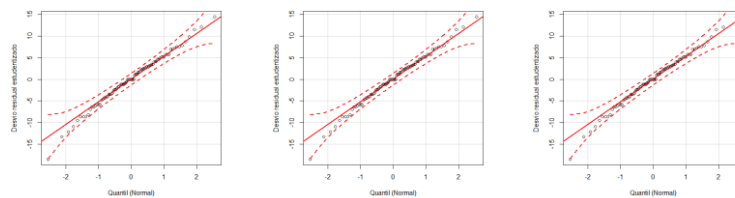
Jacaranda cuspidifolia



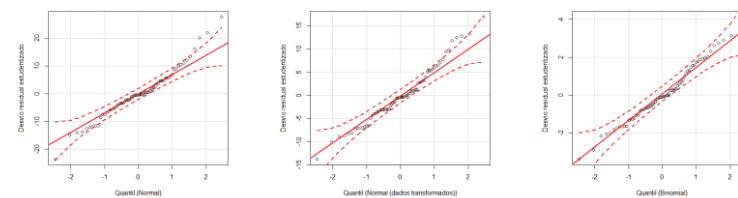
Lafoensia pacari



Mimosa caesalpiniaefolia



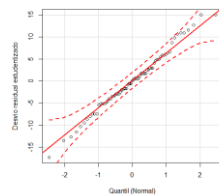
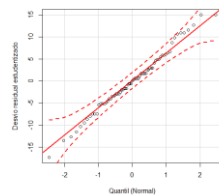
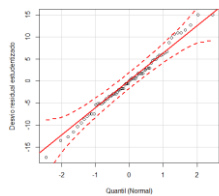
Parapiptadenia rigida



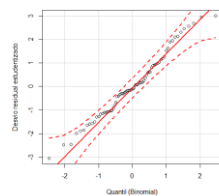
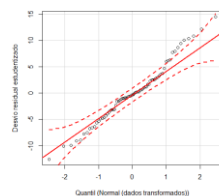
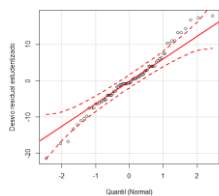
(...Continua...)

Continuação...

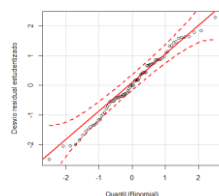
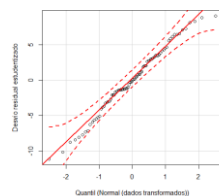
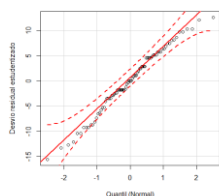
Parkia pendula



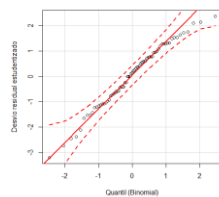
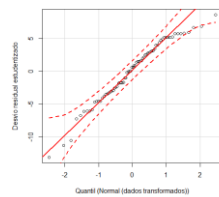
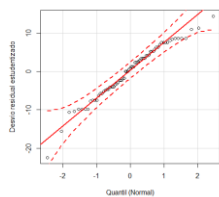
Platymenia reticulata



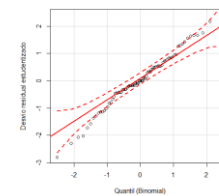
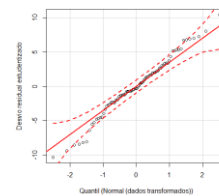
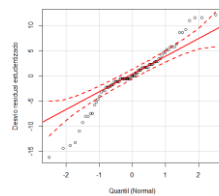
Schinus terebinthifolius



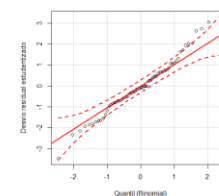
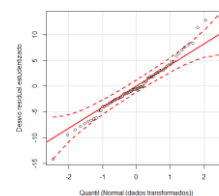
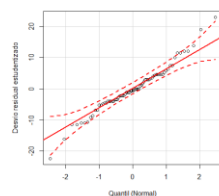
Stryphnodendron polyphyllum



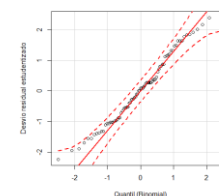
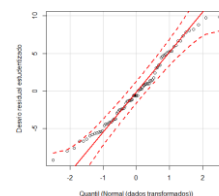
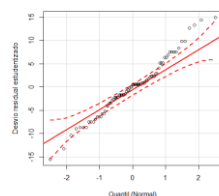
Peltogyne confertiflora



Qualea grandiflora

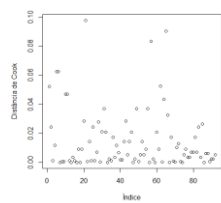


Schizolobium parahyba var. *amazonicum*

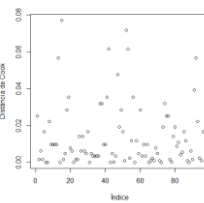


Anexo E. Gráfico da distância de Cook do percentual e número de plântulas normais das espécies que atenderam as pressuposições de normalidade e homogeneidade, com as respectivas distribuições, Normal e Binomial.

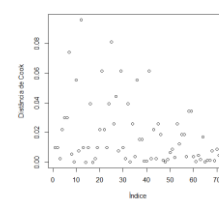
Albizia hassleri



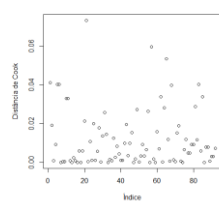
Anadenanthera macrocarpa



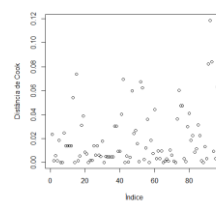
Astronium fraxinifolium



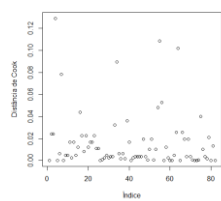
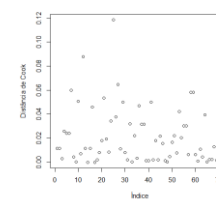
Cariniana estrellensis



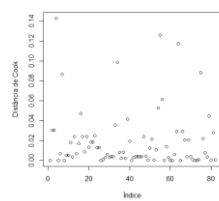
Cariniana legalis



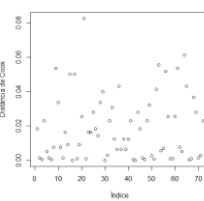
Ceiba speciosa



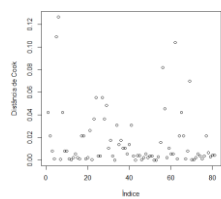
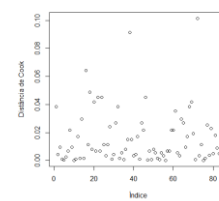
Copaifera langsdorffii



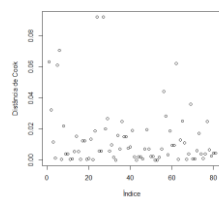
Cordia americana



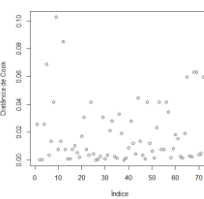
Dalbergia miscolobium



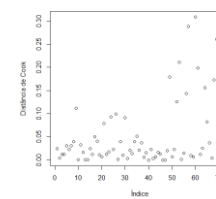
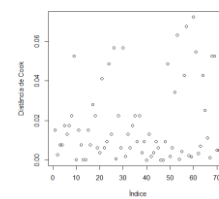
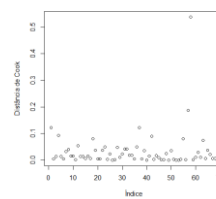
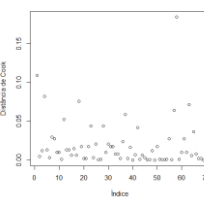
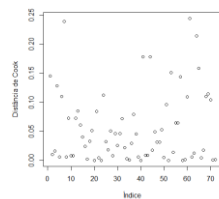
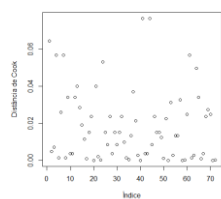
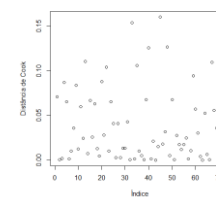
Enterolobium maximum



Erythrina speciosa



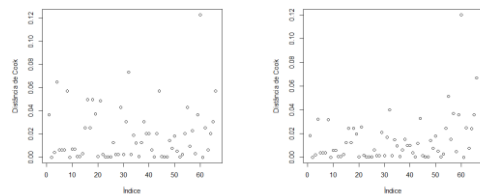
Galesia integrifolia



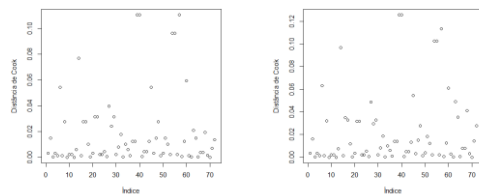
(...Continua...)

Continuação...

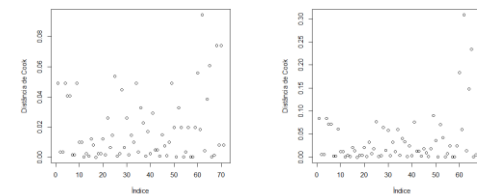
Guazuma ulmifolia



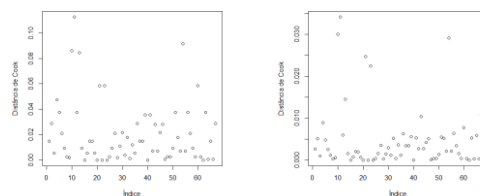
Handroanthus impetiginosus



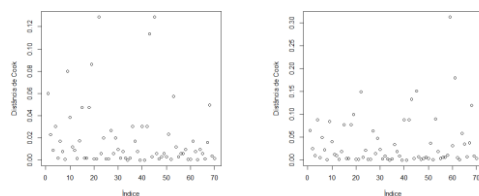
Handroanthus roseo-albus



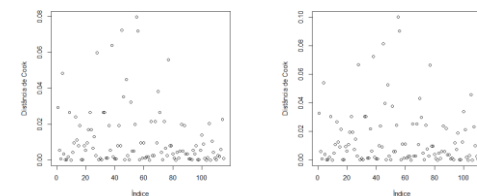
Hymenaea courbaril



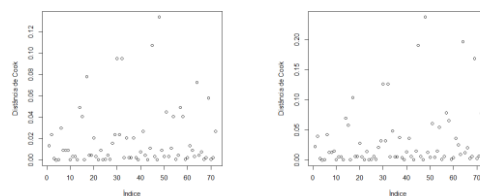
Jacaranda micrantha



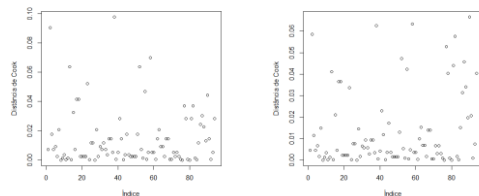
Mimosa scabrella



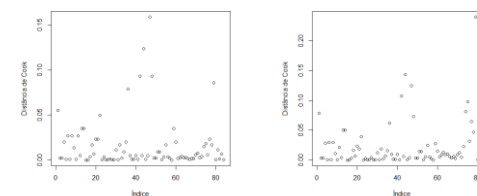
Ormosia arborea



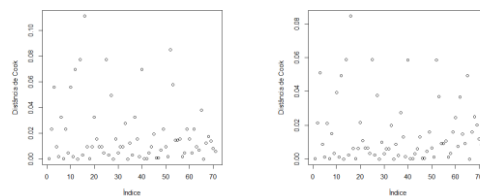
Peltophorum dubium



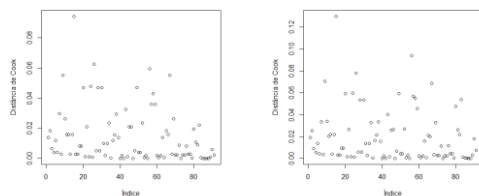
Pseudobombax tomentosum



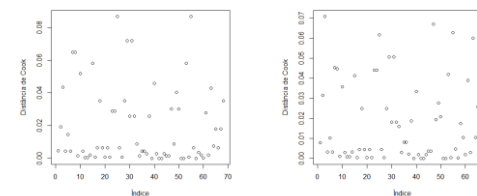
Pterogyne nitens



Senna macranthera



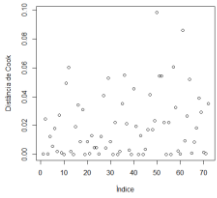
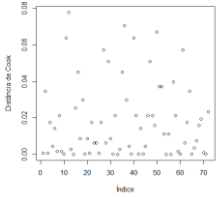
Senna multijuga



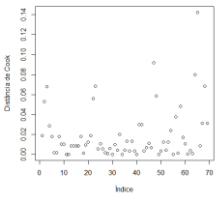
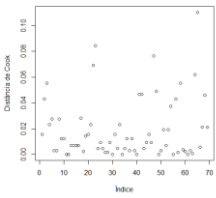
(...Continua...)

Continuação...

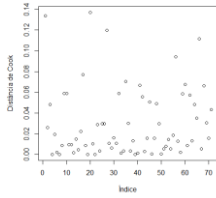
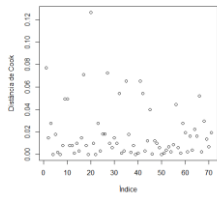
Stryphnodendron barbadetiman



Tabebuia aurea

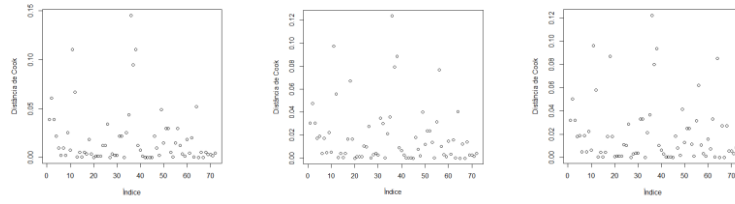


Zeyeria tuberculosa

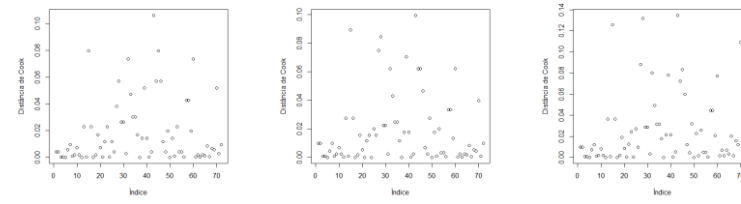


Anexo F. Gráfico da Distância de Cook para o percentual e número de plântulas normais de 50 espécies florestais, com as respectivas distribuições: Normal, Normal (dados com transformação arcoseno) e Binomial.

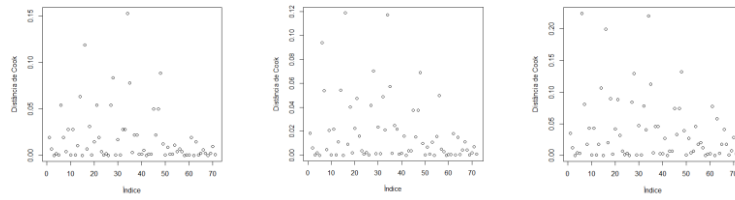
Acacia polyphylla



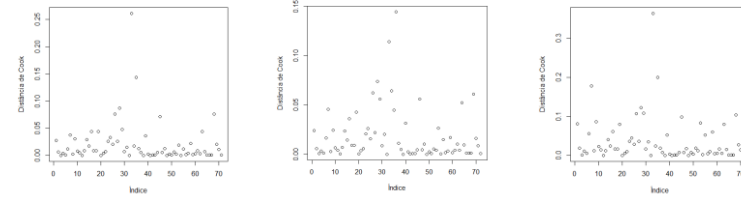
Anadenanthera colubrina



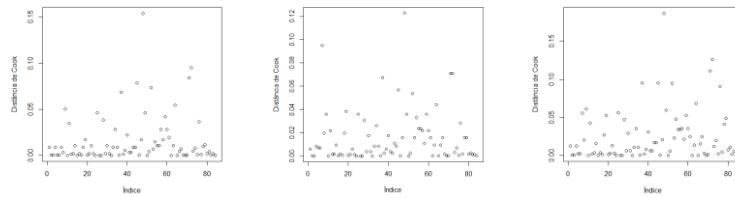
Apuleia leiocarpa



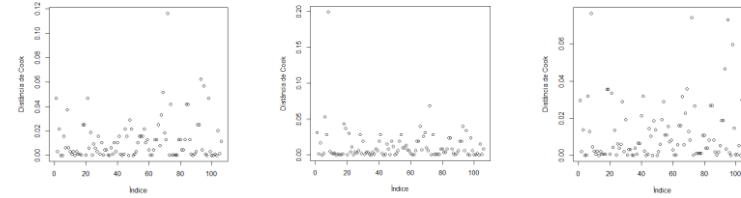
Cassia leptophylla



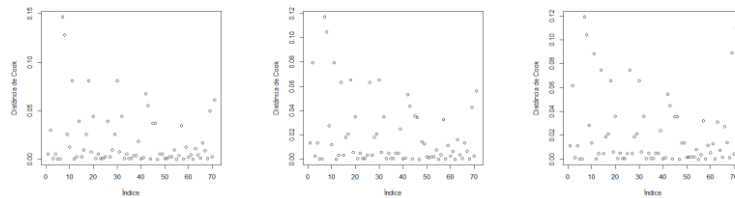
Cedrela fissilis



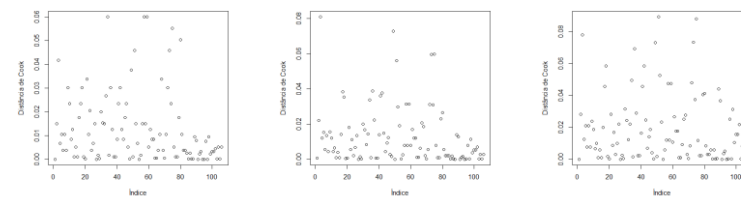
Cedrela odorata



Citharexylum myrianthum



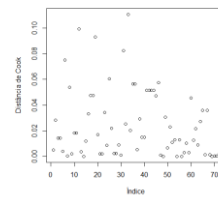
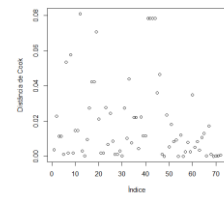
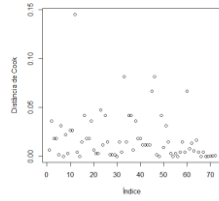
Cybistax antisyphilitica



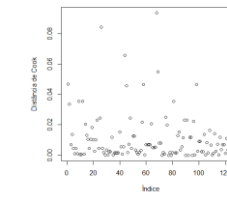
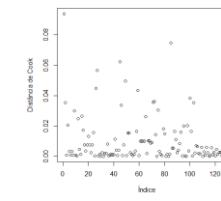
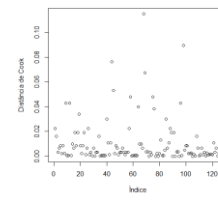
(...Continua...)

Continuação...

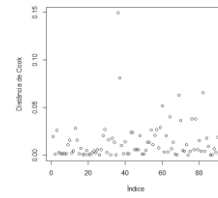
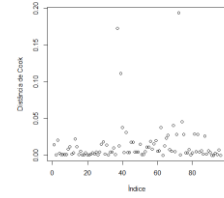
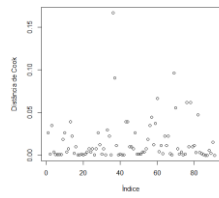
Dalbergia nigra



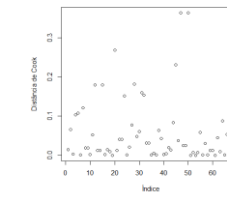
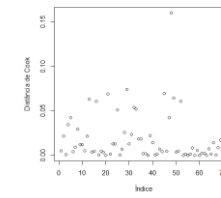
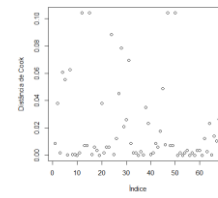
Enterolobium contortisiliquum



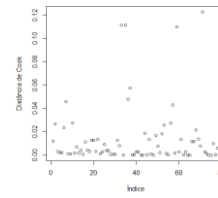
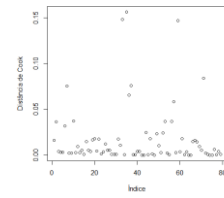
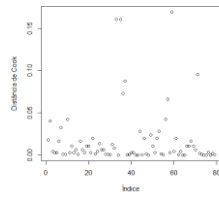
Handroanthus chrysotrichus



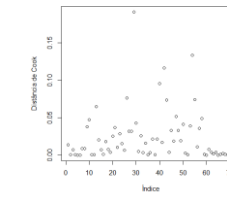
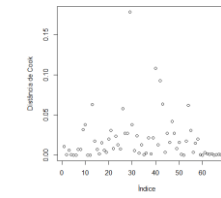
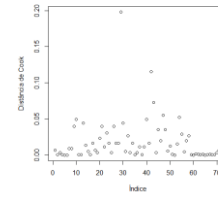
Hymenaea stigonocarpa



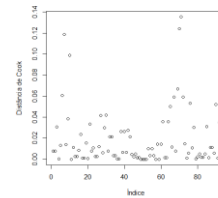
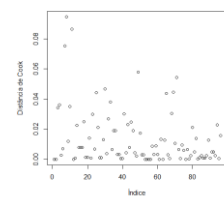
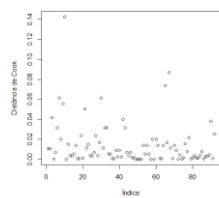
Jacaranda cuspidifolia



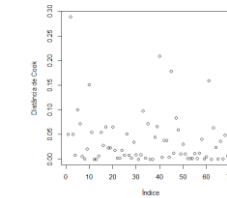
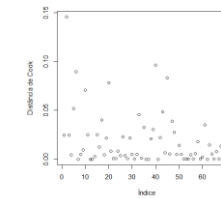
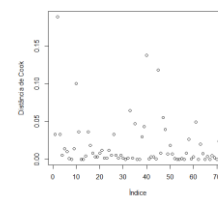
Lafoensia pacari



Mimosa caesalpiniaefolia



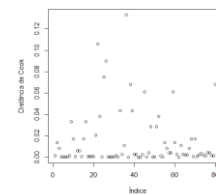
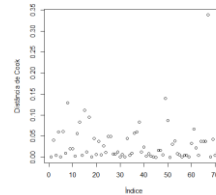
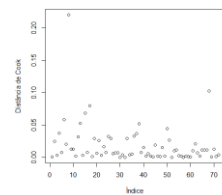
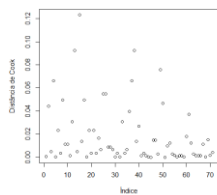
Parapiptadenia rigida



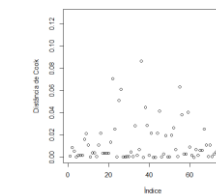
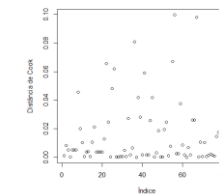
(...Continua...)

Continuação...

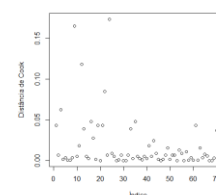
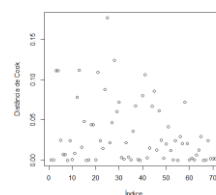
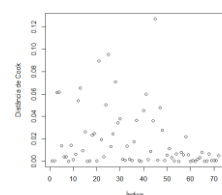
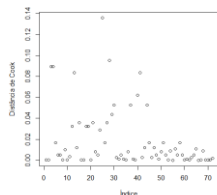
Parkia pendula



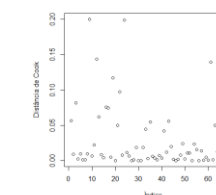
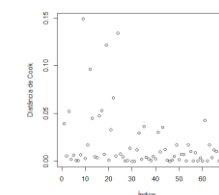
Peltogyne confertiflora



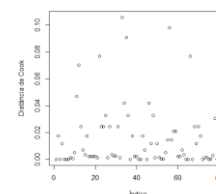
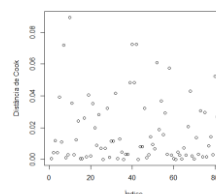
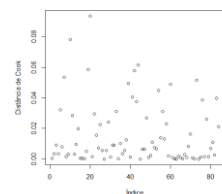
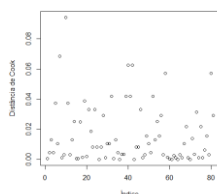
Platymenia reticulata



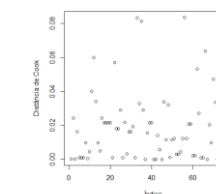
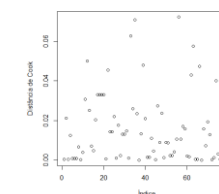
Qualea grandiflora



Schinus terebinthifolius



Schizolobium parahyba var. amazonicum



Stryphnodendron polyphyllum

