

Ajuste via modelos lineares generalizados para avaliação do controle de insetos-praga

Allana L. B. Paulino^{1†}, Ricardo A. Olinda², Sidcleide B. Sousa³, Isolina A. Vilas Bôas⁴, José Djair Vendramim⁵, Leandro P. Ribeiro⁶.

¹Universidade Federal Rural de Pernambuco, UFRPE.

²Univesidade Estadual da Paraíba, UEPB. E-mail: ricardo.estat@yahoo.com.br.

³Escola Superior de Agricultura “Luiz de Queiroz”/Universidade de São Paulo, ESALQ/USP. E-mail: sbarbosas1987@yahoo.com.br.

⁴Universidade Federal de Lavras, UFLA. E-mail: linavilasboas@gmail.com.

⁵Centro de pesquisa para Agricultura Familiar/Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina, CEPAF/EPAGRI. E-mail: jdvendra@usp.br.

⁶Centro de pesquisa para Agricultura Familiar/Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina, CEPAF/EPAGRI. E-mail: leandroribeiro@epagri.sc.gov.br.

Resumo: Neste trabalho, aplicou-se o método estatístico dos modelos lineares generalizados em um experimento com dados de proporção de mortalidade, em um bioensaio de contato residual utilizando o gorgulho-do-milho, *Sitophilus zeamais* (Coleoptera: Curculionidae) como bioindicador. Estes dados são frequentemente superdispersos, ou seja, são caracterizados por terem uma variação maior do que a assumida pelo modelo binomial. Uma das vantagens de modelar a superdispersão adequadamente é evitar inferências incorretas. Diante disso, objetivou-se ajustar uma distribuição de probabilidade aos dados, utilizando a função de ligação logística, cujo modelo ajustado foi o Quase-binomial logístico, sendo, por sua vez, avaliada a qualidade do ajuste por intermédio do gráfico half-Normal, usando o software R. Desta forma, o modelo empregado permitiu a separação dos tratamentos mais efetivos, contornando a superdispersão presente nos dados. Assim, poderá ser utilizado em estudos que visam à seleção de compostos alternativos com propriedade inseticida frente ao gorgulho-do-milho ou na avaliação da eficácia de produtos comerciais.

Palavras-chave: Modelo linear generalizado, Superdispersão, Bioensaios.

Abstract: In this study, the statistical method of generalized linear models was applied in an experiment with mortality proportion data, in a residual contact bioassay using maize weevil, *Sitophilus zeamais* (Coleoptera: Curculionidae), as a bioindicator. These data are frequently overdispersed, that is, they are characterized by having a greater variation than that assumed by the binomial model. One of the advantages of properly modeling overdispersion is to avoid incorrect inferences. Therefore, the objective was to adjust a probability distribution to the data, using the logistic link function, whose adjusted model was the quasi-binomial logistic and, in turn, the quality of the adjustment was evaluated through the half-Normal plot, using the R software. Thus, the model used allowed the separation of the most effective treatments, bypassing the overdispersion present in the data. Consequently, it can be used in studies aimed at the selection of alternative compounds with insecticidal properties against maize weevil or in the evaluation of the effectiveness of commercial products.

Keywords: Generalized linear models, Overdispersion, Bioassays.

[†] Autora correspondente: allivia_2@hotmail.com.

Introdução

Durante o armazenamento de grãos, a ocorrência de infestação por insetos-praga é comum, vindo a prejudicar o grão qualitativamente e quantitativamente. Uma das principais espécie-praga do milho armazenado é o gorgulho-de-milho, *Sitophilus zeamais* Motschulsky, 1855 (Coleoptera: Curculionidae) (CERUTI; LAZZARI, 2005), pois apresenta um grande número de hospedeiros que atacam os grãos tanto no campo como nos armazéns (FARONI, 1992).

Para o controle dos insetos (*S. zeamais*), o método mais utilizado é a aplicação de inseticidas sintéticos (GUEDES et al., 1995). Porém, essa técnica pode apresentar riscos à saúde humana, animais domésticos e meio ambiente, além dos resíduos presentes nos produtos (SIMIONI et al., 2010).

Como método alternativo para o manejo de insetos-praga em grãos armazenados, o uso de formulações à base de terra de diatomácea (pós inertes) tem sido constituída na principal e única alternativa para aplicação em maiores escalas (LORINI, 1998).

As vantagens de se utilizar tal método (sedimento de carapaças de algas diatomáceas), em relação aos inseticidas convencionais é que não possibilita o desenvolvimento de resistência nos insetos, tem eficiência duradoura, não deixa resíduos tóxicos nos grãos e não contamina o meio ambiente (solo, água e ar) (PINTO JR, 2008).

Uma proposta para análise estatística com dados de proporções são os modelos lineares generalizados (MLG) que foram introduzidos por Nelder e Wedderburn (1972). A importância dos MLGs não é apenas de índole prática. Do ponto de vista teórico, a sua importância advém, essencialmente, do fato de a metodologia destes modelos constituir uma abordagem unificada de muitos procedimentos estatísticos atualmente utilizados nas aplicações e promover o papel central da verossimilhança na teoria da inferência (TURKMAN; SILVA, 2000).

Para esse tipo de dados, utiliza-se, primeiramente, o modelo binomial, porém, a maioria desses dados são superdispersos, os quais são caracterizados por uma variabilidade maior do que a descrita por tal modelo. Dessa forma, se faz necessário adaptar modelos que levem em conta a superdispersão, a fim de evitar inferências incorretas (HINDE; DEMÉTRIO, 1998). Existem várias causas da superdispersão, entre elas, têm-se: variabilidade do material experimental, omissão de variáveis não observadas, correlações entre indivíduos etc. Em alguns casos, é difícil inferir as causas da ocorrência de superdispersão.

Diante do exposto, o propósito deste artigo, primeiramente, é ajustar um modelo com intuito de detectar as falhas com dados de proporção e, conseqüentemente, considerar extensões de modelos que expliquem a superdispersão via MLG, como o modelo quase-binomial, considerando função de ligação logística, verificar a qualidade do ajuste por meio do gráfico *half-normal* (MORAL et al., 2017), e, por fim, avaliar sua aplicação em um conjunto de dados de proporção de mortalidade obtido em um bioensaio de contato residual, visando a avaliação da eficácia de tratamentos alternativos frente ao gorgulho-do-milho.

Material e métodos

Os dados utilizados para essa análise foram disponibilizados pelo Departamento de Plantas e Inseticidas de Entomologia e Acarologia, da Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), membro do Instituto Nacional de Ciência e Tecnologia de Controle Biorracional de Insetos Pragas, sediado na UFSCar.

Entre os extratos vegetais testados foram selecionados os dois extratos, (EHamuc) e (EDAmom), os quais apresentaram os resultados mais promissores, refletidos nas menores concentrações letais média (CL_{30}) e (CL_{50}) (RIBEIRO et al., 2013;2014). Da mesma forma, selecionou-se a formulação de terra de diatomácea nas mesmas concentrações.

No início do experimento, planejamento experimental, apesar dos tratamentos representarem combinações de fatores, os pesquisadores não pensaram numa estrutura fatorial. Neste sentido,

foi considerado um único fator, sendo este a combinação de fatores que tinha interesse do ponto de vista prático (RIBEIRO et al., 2018).

Assim, foi elaborado um delineamento inteiramente casualizado (DIC), com duas concentrações letais (CL) e cinco inseticidas, visando avaliar os efeitos dos tratamentos. Cada amostra de 50 g de grãos de milho foi infestada por 30 a 40 insetos (*S. zeamais*), não sexados e com idade entre 10 e 20 dias, com 10 repetições por tratamento. Cada tratamento foi representado da seguinte maneira:

T_0 : EHAMuc - Extrato em hexano de *Annona mucosa* (CL₃₀)

T_1 : EHAMuc - Extrato em hexano de *Annona mucosa* (CL₅₀);

T_2 : EDAMon - Extrato em diclorometano de *Annona Montana* (CL₃₀);

T_3 : EDAMon - Extrato em diclorometano de *Annona Montana* (CL₅₀);

T_4 : TD - Terra de diatomácea (CL₃₀);

T_5 : TD - Terra de diatomácea (CL₅₀);

T_6 : TD(CL₃₀) + EHAMuc (CL₅₀);

T_7 : TD(CL₅₀) + EHAMuc (CL₃₀);

T_8 : TD(CL₃₀) + EDAMon (CL₅₀);

T_9 : TD(CL₅₀) + EDAMon (CL₃₀).

Modelo linear generalizado (MLG)

A transformação de dados era uma ferramenta bastante utilizada na análise estatística, a qual tinha como finalidade utilizar a teoria de modelos lineares clássicos, porém, nem sempre as pressuposições de normalidade e homogeneidade de variância eram atendidas (BOX; COX, 1964).

Com isso, os modelos lineares generalizados (MLG) foram introduzidos por Nelder e Wedderburn (1972), os quais têm uma teoria unificadora de modelagem estatística, como uma classe de modelos de regressão formulados que reuniram os modelos normal, binomial, Poisson, gama e normal inverso, que eram estudados separadamente.

Em tal modelagem, a distribuição da variável resposta do modelo era restrita somente à distribuição normal. Supondo as variáveis aleatórias Y_1, \dots, Y_n com função de probabilidade ou função densidade de probabilidade, associadas a um conjunto de variáveis explanatórias x_1, \dots, x_p , o MLG envolve três componentes

i) Componente aleatório: é representado pelo conjunto de variáveis aleatórias Y_1, \dots, Y_n , proveniente da mesma distribuição, pertencente à família exponencial na forma canônica com função de probabilidade ou função densidade de probabilidade expressa por

$$f(y_i, \theta_i, \phi) = \exp\{\phi^{-1}[y_i\theta_i - b(\theta_i)] + c(y_i, \phi)\}, \quad (1)$$

em que, $\phi > 0$ é um parâmetro de dispersão do modelo e a sua inversa ϕ^{-1} é um parâmetro de precisão, $b(\cdot)$ e $c(\cdot)$ são funções conhecidas e θ_i é chamado de parâmetro canônico, tendo como média e variância, respectivamente,

$$E(Y_i) = \mu_i = b'(\theta_i)$$

e

$$\text{Var}(Y_i) = \phi b''(\theta_i) = \phi V_i$$

sendo, $V_i = V(\mu_i) = d\mu_i/d\theta_i$ é chamada de função de variância, que depende apenas da média μ_i .

ii) Componente sistemático: conhecido como preditor linear $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)'$, sendo $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$ é a matriz do modelo e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)'$ é o vetor de parâmetros desconhecidos é dada pela soma linear dos efeitos das variáveis explanatórias, isto é, $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}$.

iii) Função de ligação: é uma função que relaciona o componente aleatório ao componente sistemático, vinculando a média μ_i ao preditor linear η_i , ou seja, $\eta_i = g(\mu_i)$.

De acordo com Cordeiro e Demétrio (2008), um caso particular dos MLGs é quando a variável de interesse é uma proporção Y , em que se tem Y insetos em um grupo de tamanho m , dessa forma, assume que $Y \sim \text{binomial}(m, \pi)$, com $0 < \pi < 1$ e m o número conhecido de ensaios independentes, a qual é usada para análise de dados na forma de proporções e tem função de probabilidade dada por:

$$f(y; \pi) = \binom{m}{y} \pi^y (1 - \pi)^{m-y} = \binom{m}{y} \exp \left[y \log \left(\frac{\pi}{1 - \pi} \right) + m \log(1 - \pi) \right] \quad (2)$$

em que, a variável aleatória Y_i tem média

$$E[Y_i] = \mu_i = m_i \pi_i$$

e variância

$$\text{Var}[Y_i] = m_i \pi_i (1 - \pi_i) = \mu_i \left(1 - \frac{\mu_i}{m_i} \right)$$

com $\theta = \log[\pi/(1 - \pi)]$, $\phi = 1$, $b(\theta) = -m \log(1 - \pi)$ e $V(\mu) = \pi(1 - \pi)$, sendo, portanto, membro da família exponencial descrita na equação (1) e função de ligação $g(\mu) = \log[\pi/(1 - \pi)]$, conhecida como função logit.

Definindo um preditor linear $\boldsymbol{\eta}$ em particular, tem-se, então, o conhecido modelo de regressão logística. Outra alternativa é considerar a função de ligação probit, definida por $g(\mu) = \Phi^{-1}(\mu)$, em que, Φ é a função de distribuição acumulada da distribuição normal padronizada. Como também, pode-se considerar a função de ligação complemento log - log, definida por $g(\mu) = \log[-\log(1 - \pi)]$.

Segundo Cordeiro e Demétrio (2008), é difícil propor uma estratégia geral para o processo de escolha de um MLG a ser ajustado aos dados que se dispõe. Isso está intimamente relacionado ao problema fundamental da estatística que, segundo Fisher (1922), “o que se deve fazer com os dados?”. Em geral, o algoritmo de ajuste deve ser aplicado não a um MLG isolado, mas a vários modelos de um conjunto bem amplo que deve ser, realmente, relevante para o tipo de dados que se pretende analisar.

Se o processo é aplicado a um único modelo, não levando em conta possíveis modelos alternativos, existe o risco de não se obter um dos modelos mais adequados aos dados. Esse conjunto de modelos pode ser formulado de várias maneiras:

- i) Definindo-se uma família de funções de ligação;
- ii) Considerando-se diferentes opções para a escala de medição;
- iii) Adicionando-se (ou retirando) vetores colunas independentes a partir de uma matriz básica original.

A suposição do modelo binomial não é satisfatória quando a variabilidade dos dados é maior do que a média, pois este modelo não é capaz de acomodar a superdispersão. Então, se faz necessário o uso de extensões do modelo binomial para analisar dados superdispersos. Pode-se lidar com a superdispersão por meio de uma abordagem de quase-verossimilhança, que requer a especificação dos dois primeiros momentos da distribuição, ou seja, média e variância. No MLG a variância do modelo binomial é substituída por

$$\text{Var}(Y_i) = \phi m_i \pi_i (1 - \pi_i)$$

em que ϕ é o parâmetro de dispersão desconhecido. Uma maneira de estimar este parâmetro é por meio de

$$\tilde{\phi} = \frac{X^2}{n - p},$$

ou seja, $X^2 = \sum_{i=1}^n (y_i - m_i \hat{\pi}_i)^2 / (m_i \hat{\pi}_i (1 - \hat{\pi}_i))$ é a estatística generalizada de Pearson do modelo binomial. É verificado o bom ajuste do modelo por meio da deviance residual. Um método de estimação para o vetor β é o método da máxima verossimilhança, o qual consiste em obter uma solução para o sistema de equações não-lineares para tais modelos, cujo j -ésimo componente do vetor escore é dado por

$$U_j = \frac{\partial l(\theta)}{\partial \beta_j} = \sum_{i=1}^n \frac{dl_i}{d\theta_i} \frac{d\theta_i}{d\mu_i} \frac{d\mu_i}{d\eta_i} \frac{\partial \eta_i}{\partial \beta_j} = \sum_{i=1}^n \phi^{-1}(y_i - \mu_i) \frac{1}{V(\mu_i)} \frac{d\mu_i}{d\eta_i} x_{ij}. \quad (3)$$

Observa-se que apenas μ_i e $V(\mu_i)$ são utilizados no processo de estimação, não sendo necessária a forma funcional completa da família exponencial. Para obter as estimativas de máxima verossimilhança de β , o algoritmo *Iterative Reweighted Least Square* (IRLS) é utilizado (MORALENBERGHS et al., 2010).

Uma outra forma de se verificar a qualidade do ajuste em um MLG é por meio do gráfico *Half-Normal* com o envelope simulado (*hnp*) (MORAL et al., 2017). Para a obtenção de *hnp*, é necessário plotar de forma ordenada os valores absolutos de uma medida de diagnóstico, como por exemplo, a deviance, versus valores esperados das estatísticas de ordem da distribuição *Half-Normal* expressa por

$$\Phi^{-1} \left(\frac{1 + n - \frac{1}{8}}{2n + \frac{1}{2}} \right),$$

em que, $\Phi^{-1}(\cdot)$ é a função acumulada da distribuição normal padrão, e n refere-se ao tamanho da amostra obtida, com $i = 1, \dots, n$. Para obtenção dos resultados utilizou-se o software R (R Core Team, 2020) de computação estatística.

Resultados e discussão

A análise exploratória dos dados vista na Figura 1 mostra a proporção observada de insetos mortos em diferentes tratamentos. Observa-se que o tratamento 0 foi o mais eficiente devido a uma maior proporção de insetos mortos, já o tratamento 6 foi o menos eficiente devido a uma menor proporção de insetos mortos. Pode-se observar que praticamente todos os tratamentos tiveram uma proporção de mortalidade acima de 0,6, apenas o tratamento 6 obteve uma proporção inferior a 0,3. Com isso, verifica-se a autenticidade das diferenças das proporções médias do fator em estudo, destacando-se os tratamentos 0, 1, 2, e 5.

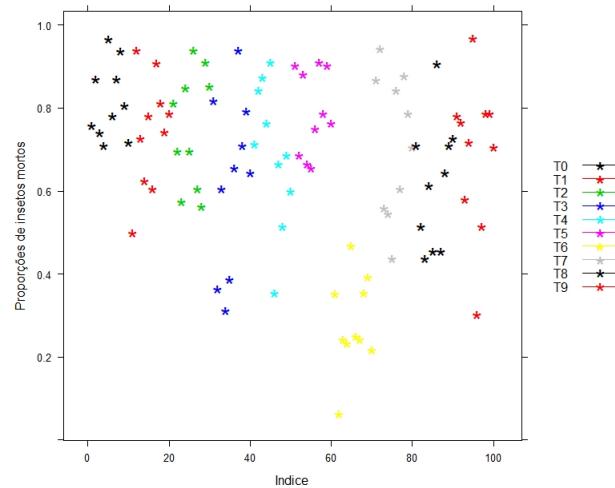


Figura 1: Gráfico de dispersão observando a proporção de inseto mortos.

Ajustando-se o modelo binomial padrão logístico com diferentes tratamentos e intercepto no preditor linear, a análise da deviance a bondade de ajuste é dada na tabela 1.

Tabela 1: Análise da deviance para os dados de gorgulhos de milho, usando o modelo binomial logístico

Fator de variação	gl	Deviance	p -valor	X^2	p -valor
Tratamentos	9	290,54	<0,0001		
Resíduo	90	346,62	<0,0001	326,93	<0,0001

Observa-se que, de acordo com os valores da deviance residual e a X^2 , o modelo não se ajusta aos dados. Há uma variação maior do que a captada pelo modelo binomial, pode-se ver também por meio do gráfico *half-Normal* mostrado na Figura 2.

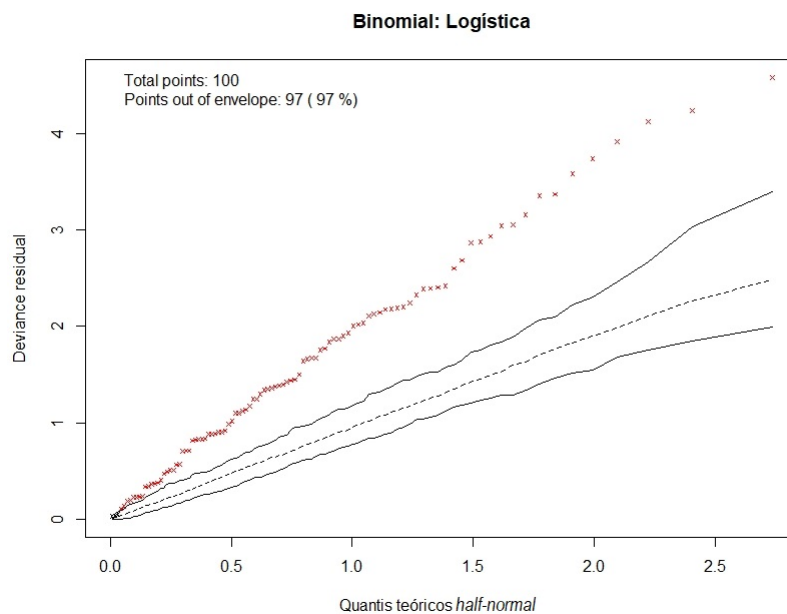


Figura 2: Gráfico *half-Normal* com envelope de simulação e deviance residual por meio do modelo binomial logístico.

Ao observar que o conjunto de dados pode exibir superdispersão, podemos estender o modelo básico. Existem várias causas possíveis de superdispersão, e vários modelos e métodos de estimação foram propostos por Hinde e Demétrio (1998). Para os dados de gorgulhos de milho foi ajustado o modelo Quase-binomial logístico para acomodar a variabilidade, o valor estimado de ϕ é $\tilde{\phi} = 3,63$ ($326,93/90 = X^2/gl$) e a análise da deviance associado ao teste F pode ser visto na tabela 2.

Tabela 2: Análise da deviance para os dados de gorgulhos de milho, usando o modelo Quase-binomial logístico

Fator de variação	gl	Deviance	F	P-valor
Tratamentos	9	290,54	8,89	<0,0001
Resíduo	90	346,62		

O gráfico na Figura 3 mostra fortes evidências que o modelo Quase-binomial logístico teve um bom ajuste com todos os pontos dentro do envelope de simulação.

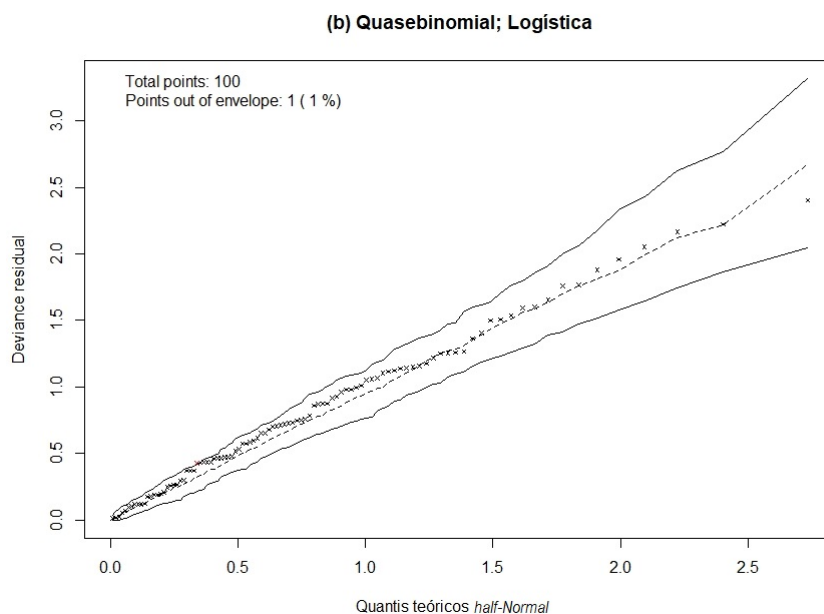


Figura 3: Gráfico *half-Normal* com envelope de simulação e deviance residual por meio do modelo Quase-binomial logístico.

O teste F para os tratamentos tem um valor observado de 8,89, como o *P-valor* foi menor que 0,05 rejeita-se a hipótese nula, ou seja, existe pelo menos um contraste (combinação linear) entre os níveis de tratamentos que é diferente de zero. Para identificar qual o melhor tratamento, foi apresentado um gráfico, Figura 4, com as proporções ajustadas com aproximadamente intervalos de confiança de 95% do modelo quase-binomial logístico e observa-se que o tratamento preparado com “extrato em hexano de *Annona mucosa*” mostrou-se mais eficaz no controle dos insetos expostos.

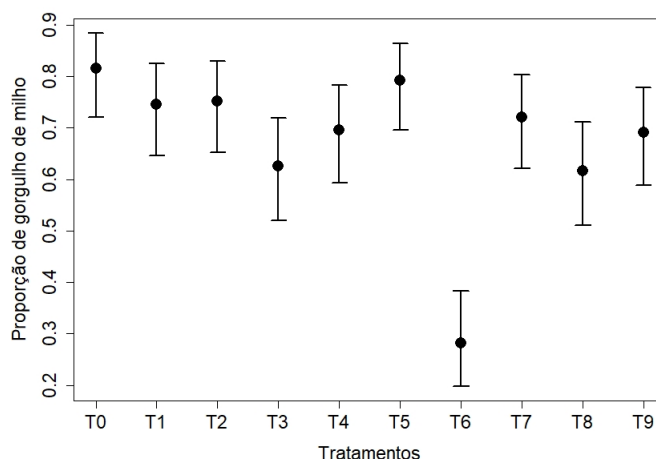


Figura 4: Gráfico de ajuste dos dados de proporção com os intervalos de confiança da distribuição de probabilidade Quase-binomial.

Conclusão

A distribuição de probabilidade que melhor se ajustou aos dados foi a Quase-binomial logístico, e a sua função de ligação foi a logística. O modelo empregado permitiu a separação dos tratamentos mais efetivos, contornando a superdispersão presente nos dados. Assim, poderá ser utilizado em estudos que visam a seleção de compostos alternativos com propriedade inseticida frente ao gorgulho-do-milho ou na avaliação da eficácia de produtos comerciais.

Referências bibliográficas

- BOX, G. E.; COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B*, JSTOR, v. 26, p. 211-252, 1964.
- CERUTI, F.C.; LAZZARI, S.M.N. Combination of diatomaceous earth and powder deltamethrin for insect control in stored corn. *Revista Brasileira de Entomologia*, São Paulo, v. 49, n. 4, p. 580-583, 2005.
- CORDEIRO, G. M.; DEMÉTRIO, C. G. *Modelos lineares generalizados e extensões*. v. 33, 159 2008. Disponível em: <https://docs.ufpr.br/taconeli/CE22517/LivClarice.pdf>. Acesso em: 10 de nov de 2017.
- FARONI, L.R.D'A. Manejo das pragas dos grãos armazenados e sua influência na qualidade do produto final. *Revista Brasileira de Armazenamento*, Viçosa, v. 17, n. 1/2, p. 36-43, 1992.
- FISHER, M. On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. Lond. A*, The Royal Society, v. 222, n. 594-604, p. 309-368, 1922.
- GUEDES, R.N.C.; LIMA, J.O.G.; SANTOS, J.P.; CRUZ, C.D. Resistance to DDT and pyrethroids in Brazilian populations of *Sitophilus zeamais* Motsch. (Coleoptera: Curculionidae). *Journal of Stored Product Research*, Oxford, v. 31, n. 8, p. 145-150, 1995.

HINDE, J.; DEMÉTRIO, C. G. Overdispersion: models and estimation. *Computational & Statistics Data Analysis*, Elsevier, v. 27, n. 2, p. 151-170, 1998.

LORINI, I. Controle integrado de pragas de grãos armazenados. *Embrapa Trigo-Documents (INFOTECA-E)*, Passo Fundo: EMBRAPA-CNPT, 1998.

MOLENBERGHS, G.; VERBEKE, G.; DEMÉTRIO, C. G.; VIEIRA, A. M. A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical science*, JSTOR, p. 325-347, 2010.

MORAL RA, HINDE J, DEMÉTRIO, CGB. hnp: *Half-Normal Plots with Simulation Envelopes*. R package version 1.2-4, URL <https://CRAN.R-project.org/package=hnp>, 2017.

NELDER, J.; WEDDERBURN, R. Generalized linear models. *Royal Statistical Society A*, v. 135, n. 3, p. 370-384, 1972.

PINTO JR., A.R. Eficiência de terra de diatomáceas no controle de algumas pragas de milho armazenado a granel. *Revista da FZVA*, v. 15, n. 1, 2008.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 2020. Disponível em: <https://www.R-project.org/>.

RIBEIRO, L.P.; VENDRAMIM, J.D.; BICALHO, K.U.; ANDRADE, M.S.; FERNANDES, J.B.; MORAL, R.A.; DEMÉTRIO, C.G.B. *Annona mucosa* Jacq. (Annonaceae): A promising source of bioactive compounds against *Sitophilus zeamais* Mots. (Coleoptera: Curculionidae). *Journal of Stored Products Research*, v. 55, p. 6-14, 2013.

RIBEIRO, L.P.; VENDRAMIM, J.D.; ANDRADE, M.S.; BICALHO, K.U.; SILVA, M.F.G.F.; VIEIRA, P.C.; FERNANDES, J. B. Tropical plant extracts as sources of grain-protectant compounds against *Sitophilus zeamais* Motschulsky (Coleoptera: Curculionidae). *Neotropical Entomology* (Impresso), v. 43, p. 470-482, 2014.

RIBEIRO, L.P.; LOVATTO, M.; VENDRAMIM, J.D. Avaliação da eficácia de duas formulações comerciais de terra de diatomácea no controle do gorgulho-do-milho com base em parâmetros toxicológicos. *Agropecuária Catarinense*, v.31, p. 56-60, 2018.

SIMIONI, N.R. *et al.* Propriedades físicas de grãos de milho e trigo tratadas com terra de diatomácea. *Varia Scientia Agrárias*, v. 1, n. 2, p. 131-140, 2010.

TURKMAN, M. A. A.; SILVA, G. L. Modelos lineares generalizados da teoria à prática. *In: VIII Congresso Anual da Sociedade Portuguesa de Estatística*, Lisboa, 2000.