

Modelagem estatística da captura da broca-do-café *(Hypothenemus hampei)* ao longo do tempo ¹

Alessandra dos Santos ²

Fernando de Pol Mayer ²

Rafael de Andrade Moral ²

Luiz Henrique Costa Mota ³

Clarice Garcia Borges Demétrio ²

John Hinde ⁴

1 Introdução

Em estudos envolvendo insetos, é comum a observação de variáveis respostas que consistem de contagens ao longo de um período de tempo. Um modelo simples que pode ser utilizado para analisar esse tipo de dados é o modelo de Poisson, um caso particular de modelo linear generalizado (McCullagh e Nelder, 1989) para o qual a média é igual à variância. Entretanto, comumente a variabilidade dos dados de contagem é maior do que a esperada pelo modelo de Poisson, fenômeno conhecido como superdispersão. Dentre as possíveis causas da superdispersão, podem ser citadas respostas agregadas, amostragem em multiestágios, excesso de zeros ou ainda correlação entre respostas individuais (Hinde e Demétrio, 1998).

Quando os dados apresentam superdispersão, modelos apropriados devem ser utilizados. Uma das consequências de não levar em conta a superdispersão é a subestimação dos erros padrões das estimativas dos parâmetros (Hinde e Demétrio, 1998). Diferentes estratégias podem ser adotadas para modelar a superdispersão, como a especificação de uma função de variância mais geral ou ainda utilização de modelos de misturas de distribuições. Outra estratégia envolve a inclusão de efeitos aleatórios no preditor linear. Ao utilizar um modelo de Poisson com um ou mais efeitos aleatórios normalmente distribuídos, tem-se um modelo Poisson-normal.

A interpretação de um efeito como sendo fixo ou aleatório depende do objetivo e da condição do experimento. Se o interesse está em testar diferenças entre os níveis particulares de um fator, considera-se o fator como de efeito fixo. Caso os níveis de um fator sejam uma amostra aleatória de uma população, assume-se que o fator é de efeito aleatório. Ainda assim, dependendo do contexto, o mesmo fator pode ser considerado como fixo ou aleatório. Em estudos longitudinais, nos quais variáveis respostas são observadas em uma mesma parcela repetidas vezes ao longo

¹Agradecimento ao CNPq e à FAPESP pelo apoio financeiro.

²LCE - ESALQ/USP. e-mail: *alessandra.s@usp.br*

³LEA - ESALQ/USP.

⁴National University of Ireland Galway.

do tempo, pode haver a necessidade de incluir efeitos aleatórios para modelar a correlação entre observações tomadas na mesma parcela (Pinheiro e Bates, 2002).

O objetivo deste trabalho foi analisar dados referentes a um estudo longitudinal, no qual foram observados o número de espécimes de um inseto praga capturados em diferentes armadilhas, durante 24 semanas.

2 Material e métodos

A broca-do-café *Hypothenemus hampei* (Coleoptera: Scolytidae) é uma das mais importantes pragas do café, pois ataca o fruto em desenvolvimento, causando grandes perdas na produção e qualidade dos grãos. Em um contexto de controle biológico, no qual estratégias alternativas à aplicação de pesticidas são desenvolvidas, podem ser levados em conta aspectos comportamentais da praga. Sabe-se que esse besouro é atraído por etanol e pela cor vermelha. Com base nisso, Mota (2012) conduziu um experimento para comparar três tipos de armadilhas, instaladas em quatro linhas (blocos) de uma plantação de café. Os insetos eram removidos das armadilhas e contados, semanalmente, por 24 semanas, caracterizando um estudo longitudinal.

Inicialmente, o modelo utilizado foi o de Poisson com função de ligação logarítmica e preditor linear dado por

$$\eta_{ijk} = \mu + \beta_i + \alpha_j + \delta_k + (\alpha\delta)_{jk} + \sigma_Z z_{ij}, \quad i = 1, \dots, 4, \quad j = 1, 2, 3, \quad k = 1, \dots, 24$$

sendo β_i o efeito fixo do i -ésimo bloco, α_j o efeito fixo da j -ésima armadilha, δ_k o efeito fixo da k -ésima semana, $(\alpha\delta)_{jk}$ o efeito da interação entre a j -ésima armadilha e a k -ésima semana, σ_Z^2 a variância do efeito aleatório associado às parcelas e Z_{ij} variáveis aleatórias com distribuição normal padrão. Porém, como os dados apresentam superdispersão, optou-se por incluir um efeito aleatório a nível observacional, reescrevendo o preditor linear como

$$\eta_{ijk} = \beta_i + \alpha_j + \delta_k + (\alpha\delta)_{jk} + \sigma_Z z_{ij} + \sigma_W w_{ijk},$$

sendo σ_W^2 a variância do efeito aleatório associado ao nível observacional. Têm-se $\text{Var}(\eta_{ijk}) = \sigma_Z^2 + \sigma_W^2$, $\text{Cov}(\eta_{ijk}, \eta_{ijk'}) = \sigma_W^2$ e as demais covariâncias iguais a zero. O efeito aleatório associado às parcelas modela a dependência entre observações tomadas na mesma parcela ao longo do tempo. Também foram ajustados um modelo sem efeitos aleatórios e um sem efeito da interação, para fins de comparação utilizando o critério de informação AIC e o teste da razão de verossimilhança.

Para verificar a qualidade do ajuste dos modelos considerados, foram utilizados gráficos meio-normais com envelope de simulação para os resíduos condicionais, obtidos por meio da função `hnp()` (Moral, 2014), no *software* R (R Core Team, 2013). Todas os modelos foram ajustados por meio da função `glmer()`, do pacote `lme4` (Bates *et al.*, 2013).

3 Resultados e discussão

Nota-se que o comportamento do número de brocas coletadas por semana não apresenta uma tendência linear (Figura 1), e além disso, há muitos picos ao longo do intervalo estudado. Assim, seria incorreto considerar um modelo de regressão linear. Segundo Mota (2012), isso se deve a fatores climáticos, como umidade e temperatura, que afetam o comportamento da broca-do-café, bem como pode estar relacionado diretamente ao ciclo de vida do inseto. Pode-se observar, também, que não há grande variabilidade entre as observações tomadas no mesmo bloco. Além disso, os blocos são considerados como de efeito fixo no modelo visto que são linhas igualmente espaçadas em uma mesma plantação de café.

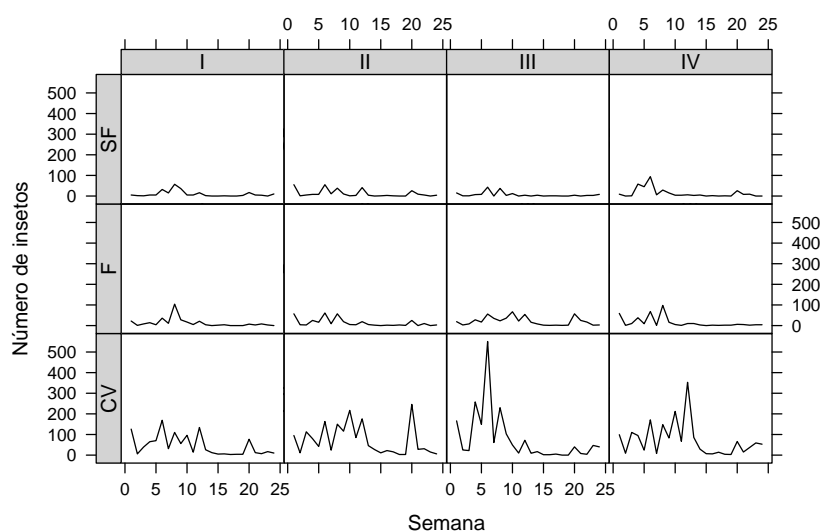


Figura 1: Número de insetos coletados por armadilha (SF = autoinoculação sem fungo, F = autoinoculação com fungo e CV = novo modelo proposto), em cada bloco, durante as 24 semanas do experimento.

Observando a relação entre as médias e as variâncias das combinações entre armadilhas e semanas (Figura 2), há indícios de que os dados apresentam superdispersão. Pelos resultados das estatísticas de ajuste dos modelos, observa-se que a inclusão de efeitos aleatórios é necessária (Tabela 1).

Tabela 1: Estatísticas dos ajustes dos modelos de Poisson sem e com efeitos aleatórios ajustados ao conjunto de dados de número de insetos coletados ao longo do tempo.

Preditor linear (η_{ijk})	$-2 \times \log \text{Lik}$	Deviance	X^2	GL	AIC
$\mu + \beta_i + \alpha_j + \delta_k + (\alpha\delta)_{jk}$	3847,8	2729,5	2695,7	213	3997,8
$\mu + \beta_i + \alpha_j + \delta_k + (\alpha\delta)_{jk} + \sigma_Z z_{ij}$	3763,0	2598,5	2503,6	212	3915,0
$\mu + \beta_i + \alpha_j + \delta_k + (\alpha\delta)_{jk} + \sigma_Z z_{ij} + \sigma_W w_{ijk}$	1900,7	115,1	86,4	211	2054,7
$\mu + \beta_i + \alpha_j + \delta_k + \sigma_Z z_{ij} + \sigma_W w_{ijk}$	1961,6	109,8	78,2	257	2023,5

Verifica-se, na Tabela 1, que ao se incluírem os dois efeitos aleatórios no modelo com interação houve uma grande redução no valor da deviance e no X^2 de Pearson. Pelo AIC também

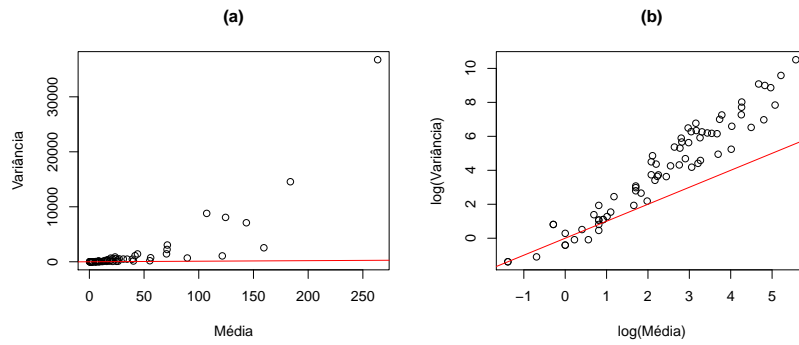


Figura 2: Gráficos das (a) médias *versus* variâncias das combinações entre armadilha e semana e do (b) logaritmo das médias *versus* o logaritmo das variâncias.

se constata essa redução, sendo que o modelo com a presença de dois efeitos aleatórios mas sem interação é o que apresenta o melhor resultado pelo critério, além disso, este apresenta menor X^2 de Pearson. Analisando a qualidade do ajuste pelos gráficos meio-normais (Figura 3), nota-se que apenas os modelos com dois efeitos aleatórios, com e sem interação, indicaram bom ajuste.

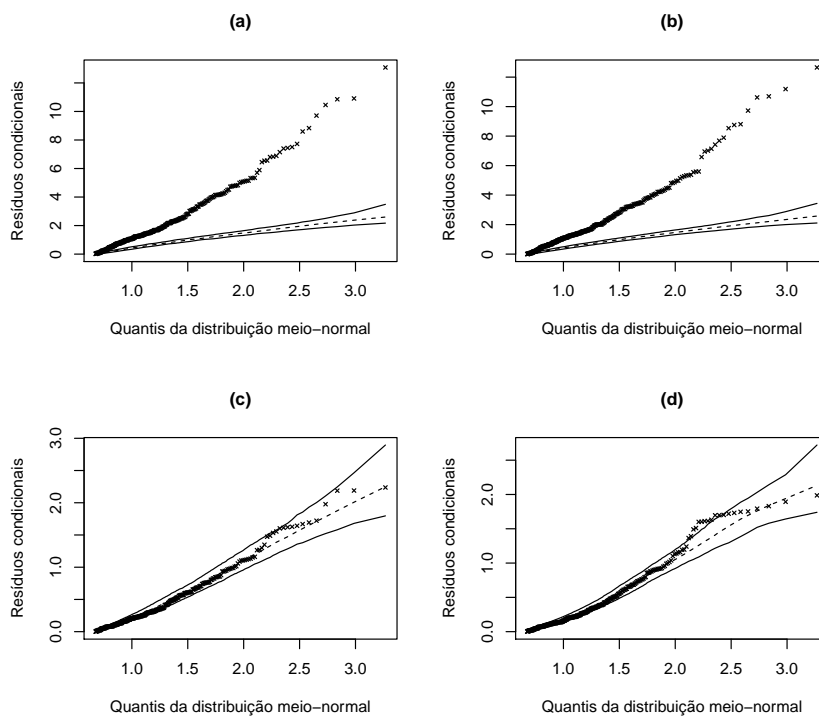


Figura 3: Gráfico meio normal para os modelos: (a) sem efeito aleatório (b) com efeito aleatório por indivíduo (c) com efeitos aleatórios por indivíduo e observação (d) com efeitos aleatórios, mas sem interação.

O teste da razão de verossimilhanças para os modelos com e sem interação, resultou em $-2(l_c - l_r) = 60,841$ o que corresponde ao nível descritivo de 0,07 com 46 graus de liberdade da diferença, o que indica não haver diferença significativa entre eles. No entanto, para se

testarem estes modelos encaixados, σ_Z^2 e σ_W^2 foram reestimados no ajuste do modelo reduzido absorvendo a variabilidade referente à interação. Tem-se que para o modelo completo $\sigma_Z^2 = 0,038$ e $\sigma_W^2 = 0,327$, enquanto que para o modelo reduzido $\sigma_Z^2 = 0,034$ e $\sigma_W^2 = 0,432$.

4 Conclusões

1. A modelagem dos dados por meio da distribuição poisson-normal é adequada para este estudo longitudinal, no qual foram observados o número de espécimes do inseto praga capturados em diferentes armadilhas.
2. A inclusão de efeito aleatório por observação captura o efeito da superdispersão dos dados melhorando a qualidade dos ajuste.
3. A interação dos fatores armadilha e semana não é significativa, sendo que o efeito aleatório por observação absorve uma maior variabilidade devido a ausência do efeito da interação.

Referências

- [1] BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5. <http://CRAN.R-project.org/package=lme4>, 2013.
- [2] HINDE, J.; DEMÉTRIO, C. G. B. Overdispersion: models and estimation. **Computational Statistics & Data Analysis**, Amsterdam, v. 27, p. 151-170, 1998.
- [3] McCULLAGH, P.; NELDER, J. A. **Generalized linear models**, Londres: Chapman and Hall, 1989. 511 p.
- [4] MORAL, R. de A. **Modelagem estatística e ecológica de relações tróficas em pragas e inimigos naturais**. 2013. 173 p. Dissertação (Mestrado em Ciências. Área de concentração: Estatística e Experimentação Agronômica), Universidade de São Paulo, Piracicaba, 2013.
- [5] MOTA, L. H. C. **Desenvolvimento de armadilhas de auto-inoculação para o controle de *Hypothenemus hampei* (Ferrari, 1867) (Coleoptera:Curculionidae) com *Beauveria bassiana* (Bals.) Vuil (Ascomycota: Hypocreales) em tecido sintético**. 2012. 80 p. Dissertação (Mestrado em Ciências. Área de concentração: Entomologia), Universidade de São Paulo, Piracicaba, 2012.
- [6] PINHEIRO, J. C.; BATES, D. M. **Mixed Effects Models in S and S-Plus**. New York: Springer, 2002. 537 p.