

Programa R para inferência sobre parâmetros populacionais de artrópodes

Ricardo Antonio de Almeida Pazianotto^{1,2}

Aline de Holanda Nunes Maia¹

Alfredo José Barreto Luiz¹

Jeanne Scardini Marinho Prado¹

1 Introdução

Parâmetros populacionais são importantes indicadores em estudos envolvendo a interação entre artrópodes e o meio ambiente, como, por exemplo: quantificação de efeitos de concentrações de produtos (MARINHO-PRADO et al., 2011), avaliação de risco ambiental do uso de biopesticidas (NASCIMENTO et al., 1998), culturas transgênicas sobre insetos não-alvo (LIU et al, 2005). Nesses estudos, são realizados experimentos onde dados de oviposição e sobrevivência são coletados, ao longo do tempo, e condensados em tabelas de vida e fertilidade (TBVF) para posterior estimação dos parâmetros: taxa líquida de reprodução (R_0), taxa intrínseca de crescimento (R_m), tempo de duplicação (TD), intervalo entre gerações (TMG) e razão finita de crescimento (λ), para cada um dos tratamentos avaliados. Os parâmetros da TBVF são estimados para cada grupo (tratamento) e não para cada fêmea o que impossibilita a estimação da variância via métodos tradicionais. Uma abordagem largamente utilizada para quantificar a incerteza associada a esses parâmetros é o método *jackknife*, onde cada parâmetro em questão é recalculado retirando-se do conjunto de dados, as informações de uma fêmea a cada iteração (MEYER, 1986). Os pseudovalores, estimativas obtidas a cada iteração, são então usados na estimação da variância, informação necessária para construção de intervalos de confiança, realização de testes para comparação de tratamentos qualitativos ou ainda ajuste de modelos de regressão para avaliar o efeito de níveis de tratamentos quantitativos (ex. doses subletais de produtos) sobre os parâmetros da TBVF. HULTING et al (1990) desenvolveram um programa em linguagem Fortran para obter estimativas *jackknife* de R_0 , R_m , DT , TMD e λ , intervalo de confiança para o R_m , e testes para comparar grupos quanto ao R_m . Esse programa, no entanto, apresenta algumas limitações tais como: a)

¹ Embrapa Meio Ambiente, Jaguariúna, SP

² Contato: pazianotto@cnpma.embrapa.br

considera a sobrevivência na fase imatura igual para todos os grupos; b) o algoritmo para estimar a porcentagem de fêmeas na descendência não é adequado para casos onde alguma das fêmeas de quaisquer dos grupos avaliados tem postura zero durante todo o período de avaliação e c) não calcula os níveis de significância nominal (valores p) associados aos testes de comparação de grupos. Face a essas limitações, Maia et al (2000) desenvolveram um programa alternativo em ambiente SAS (Lifetable.sas, disponível em <http://www.cnpma.embrapa.br/forms/sastat.php3>) para estimação de parâmetros da TBVF e comparação entre tratamentos qualitativos.

No presente trabalho, é apresentada uma versão do referido programa SAS, em linguagem R (R DEVELOPMENT CORE TEAM, 2010), acrescida das seguintes inovações: a) uso do método iterativo de Brent para estimação do R_m , ao invés da busca entre valores pontuais num intervalo arbitrário em torno R_m , estimado pelo método aproximado, b) um módulo para testes de comparação múltipla e c) um módulo para ajuste de modelos de regressão linear quando os tratamentos são quantitativos, incluindo diagnósticos de resíduos e análise de influência. O programa em R é mais acessível aos usuários potenciais por tratar-se de um *software* livre de código aberto. Para demonstração da utilidade do programa e interpretação das análises descritivas e inferenciais implementadas, é apresentado um estudo de caso onde se avalia o efeito de concentrações de berenil, um inibidor de proteases, sobre parâmetros populacionais de *Thyriniteina arnobia* (Stoll) (Lepidoptera: Geometridae), lagarta desfolhadora da cultura do eucalipto (MARINHO-PRADO et al., 2011).

2 Material e Métodos

No presente trabalho, consideramos casos onde os experimentos para obtenção dos dados necessários à construção de tabelas de vida e fertilidade e conseqüente estimação de parâmetros associados, têm delineamento inteiramente casualizado. Assim, n fêmeas adultas são alocadas aleatoriamente a G grupos. As fêmeas são acasaladas e acompanhadas até a morte, registrando-se, a intervalos de tempo regulares preestabelecidos (dias, semanas, etc), todos os dados relativos à sobrevivência e oviposição. As informações sobre fertilidade e sobrevivência, necessárias para construção de TBVF são organizadas numa planilha Excel. O programa R lê a planilha Excel para cálculo de estatísticas descritivas e construção de gráfico dos padrões temporais de oviposição e *box-plots* para o número total de ovos por fêmea, em cada grupo. São utilizados algoritmos para cálculo da tabela de vida em cada tratamento, com os dados completos e para cada iteração (i) do método *jackknife*. No programa Lifetable.sas, o valor de R_m é obtido iterativamente a partir da Equação 1,

$$\sum_{t=t0_g}^{t_g} e^{-r.t} \times Ro_{gt} = 1 \quad (1)$$

na qual, $t_{gi}^{(0)}$ e $t_{gi}^{(f)}$ representam as idades da fêmea i ($i=1,2,\dots,n_g$), do grupo g , no início da fase adulta e na data da morte da mesma, respectivamente; r é a raiz da equação a ser estimada no processo iterativo e Ro_{gt} é o número líquido de fêmeas produzidas por fêmea no grupo g , no tempo t . Para o intervalo discreto entre $0,80 Rm$ e $1,20 Rm$, com Rm obtido pelo método aproximado, busca-se o valor de r que minimiza a diferença entre os dois lados da expressão (Maia et al (2000)).

No programa R, as estimativas de Rm são obtidas utilizando a função *uniroot*, baseada no método iterativo de Brent (BRENT, 1973), sem necessidade de discretização do intervalo de busca ($0,5Rm$ a $1,5 Rm$). Em cada passo i do método *jackknife*, calculam-se os pseudovalores correspondentes a cada parâmetro, subtraindo-se o valor da estimativa no passo zero do valor da estimativa no passo i , ponderando-se pelos respectivos números de fêmeas considerados na sua estimação. Para ilustrar a magnitude e a dispersão dos pseudovalores em cada grupo, são construídos *box-plots* para cada parâmetro. É gerada uma tabela com as estimativas *jackknife* dos parâmetros e respectivos erros-padrão em cada grupo.

Para fatores qualitativos como, por exemplo, tipos de dieta (NASCIMENTO et al., 1998, foram implementado o testes t para comparação de parâmetros entre tratamentos, dois a dois, usando a desigualdade de Sidák para correção dos níveis de significância no caso de haver mais de dois tratamentos ($G>2$). Para fatores quantitativos como, por exemplo, temperatura (Laumman, 2008), ou inibidores de protease (MARINHO-PRADO et al., 2011) foi implementado um algoritmo para ajuste de modelos de regressão linear. Para cada parâmetro, são gerados gráficos que mostram os pseudovalores calculados e os respectivos modelos de regressão ajustados com intervalos de confiança de 95%. São geradas tabelas com as estimativas dos coeficientes linear (β_0) e angular (β_1) dos modelos com respectivos erros-padrão e níveis de significância nominal (valores p) derivados dos testes t para hipóteses nulas $\beta_0=0$ e $\beta_1=1$.

3 Resultados e Discussão

O programa R, além da maior facilidade de acesso pelos usuários, por ser desenvolvido num ambiente de software livre e código aberto; inclui algoritmos para uso de métodos não disponíveis na versão em SAS (Lifetable.SAS). O uso do método iterativo de Brent para cálculo do Rm a partir da igualdade descrita na equação 1 torna o processo de busca da raiz

mais rápido e eficiente do que o utilizado no programa Lifetable.sas. Como ilustração dos métodos utilizados e resultados gerados pelo Lifetable.R apresentamos, a seguir, um estudo onde foi avaliado o efeito do berenil (inibidor de proteases), sobre parâmetros da TBVF da lagarta desfolhadora do eucalipto.

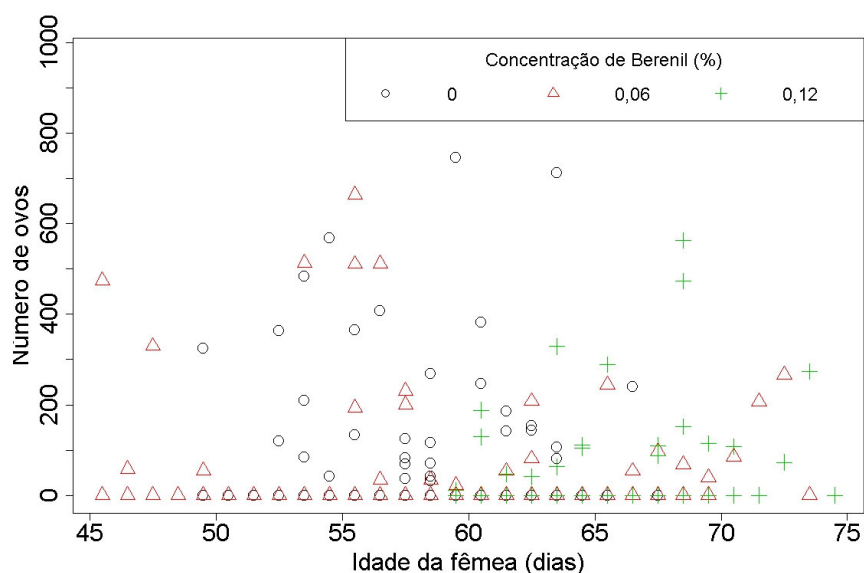


Figura 1. Oviposição das fêmeas para as três concentrações de berenil ao longo do tempo.

Tabela 3. Estimativas dos parâmetros β_0 (intercepto) e β_1 (coeficiente angular) dos modelos de regressão para R_0 , TG , TD , λ e R_m .

Parâmetro da TBVF	Coefficiente	Estimativa	Erro-padrão	Valor p
Taxa líquida de reprodução (R_0)	β_0	223,95	18,59	<0,0001
	β_1	-1010,48	268,28	0,0009
Intervalo entre gerações (TG)	β_0	55,32	1,91	<0,0001
	β_1	69,17	27,59	0,1900
Tempo de duplicação (TD)	β_0	7,03	0,33	<0,0001
	β_1	-19,52	4,80	0,0004
Razão finita de crescimento (λ)	β_0	0,974	0,004	<0,0001
	β_1	-0,192	0,051	0,0010
Taxa intrínseca de crescimento (R_m)	β_0	1,102	0,004	<0,0001
	β_1	-0,208	0,056	0,0010

As estimativas dos parâmetros β_0 e β_1 dos modelos de regressão para Ro, TG, TD, λ e Rm são apresentados na Tabela 3. Os coeficientes angulares são positivos para o intervalo entre gerações e tempo de duplicação e negativos para os demais parâmetros; os correspondentes níveis de significância nominal inferiores a 0,001, indicam uma redução do potencial de crescimento populacional da praga com o aumento da dose de berenil, comportamento consistente com efeito esperado para um inibidor de proteases.

4 Conclusões

Neste trabalho, apresentamos um programa R para estimação de parâmetros populacionais, importante ferramenta para os estudos de efeitos subletais de tratamentos em artrópodes. Estão disponíveis métodos inferenciais para tratamentos quantitativos e qualitativos, com geração de relatórios, informações relevantes para análise e gráficos em alta resolução.

5 Referências

- [1] BRENT, R. **Algorithms for Minimization without Derivatives**. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- [2] LAUMANN R A ; MORAES M C B ; PAREJA M ; G C ALARCÃO ; BOTELHO, A C ; MAIA, A H N ; LEONARDECZ, E ; BORGES, M . Comparative biology and functional response of *Trissolcus spp.*(Hymenoptera: Scelionidae) and implications for stink bugs (Hemiptera: Pentatomidae) biological control. **Biological Control**, v. 44, p. 32-41, 2008.
- [3] LIU, X.D.; BAO; P. Z., XIAO, X. Z.; J. M. ZONG. Impact of transgenic cotton plants on a non-target pest, *Aphis gossypii* Glover. **Ecological Entomology**, 30: 307–315, 2005.
- [4] MAIA, A. DE H. N.; LUIZ, A. J. B.; CAMPANHOLA, C. Statistical Inferences on associated lifetable parameters using jackknife technique: computational aspects. **J. Econ. Entomol.** 93(2):511-518, 2000.
- [5] MEYER, J. S., C. G. ITERSOLL, L. L. MACDONALD AND M. S. BOYCE. Estimating uncertainty in population growth rates: Jackknife vs. Bootstrap techniques. **Ecology** 67: 1156-1166, 1986.
- [6] NASCIMENTO, M. L., D. F. Capalbo, G. J. Moraes, E. A. B. De Nardo, A. de H. N. Maia and R. C. A. L. Oliveira. Effect of a formulation of *Bacillus thuringiensis* Berliner var. *kurstaki* on *Podisus nigrispinus* Dallas (Heteroptera: Pentatomidae asopinae). **J. Invert. Path.** 72 : 178-180, 1998.
- [7] R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [8] MARINHO-PRADO, J. S.; A. L. LOURENÇÃO, J. A. OLIVEIRA; R. N. C. GUEDES; M. G. A. OLIVEIRA. Survival and feeding avoidance of the eucalyptus defoliator *Thyrinteina arnobia* exposed to the proteinase inhibitor berenil. *Journal of Applied Entomology*, 135: 763–770, 2011.