

Sobre o procedimento de Zippin para estimar populações de animais.

Jair Licio Ferreira Santos*, Waldyr Muniz Oliva**

SANTOS, J.L. F. & OLIVA, W.M. Sobre o procedimento de Zippin para estimar populações de animais. *Rev. Saúde públ.*, S. Paulo, 25: 53-5, 1991. O procedimento de Zippin é uma das técnicas mais precisas para estimar populações de animais, além de fornecer estimativas do erro padrão. Entretanto, a solução das equações de máxima verossimilhança não é algebricamente explícita. Sugere-se um procedimento iterativo utilizando a técnica de Newton Raphson para encontrar as soluções desejadas. Comparações com o procedimento clássico mostram sensíveis diferenças nas estimativas do erro padrão e diferenças menores nos demais parâmetros.

Descritores: Estimativas de população. Técnicas de estimação.

Introdução

Na seção referente a estimativas de populações através de remoção por captura, Southwood² (1978) menciona três abordagens diferentes. Entre elas, a técnica de Zippin, além de proporcionar resultados mais precisos é a única que possibilita calcular o erro padrão da estimativa populacional.

Na notação de Southwood², sendo Y(i) o total capturado na i-ésima ocasião, k o número total de capturas, e T o total capturado nas k capturas, deve-se calcular

$$R = \frac{\sum_{i=1}^k (i - 1) \cdot Y(i)}{T} \quad (1)$$

O valor de R obtido em (1) se relaciona com a probabilidade de captura em cada ocasião — denotada por 'p' através da expressão

$$F(q) = R \cdot \frac{q}{p} + \frac{k \cdot q^k}{(1 - q^k)} = 0 \quad (2)$$

onde q é a probabilidade de não captura, q = 1 - p.

Os estimadores de máxima verossimilhança para o total populacional e seu erro padrão são respectivamente:

$$N = T / (1 - q^k) \quad (3)$$

$$S = \{W \cdot T / [T \cdot T - W \cdot Z]\}^{1/2} \quad (4)$$

onde W e Z valem respectivamente:

$$W = N(N - T), \quad Z = \frac{(kp)^2}{(1-p)}$$

Já que da expressão (2) para F(q) não se pode obter explicitamente o valor de q, (ou p) necessário para o cálculo de N e seu erro padrão, a técnica prevê a utilização de gráficos. Para k = 3, 4, 5 e 7, Southwood² (1978) publica gráficos para estimar p e q a partir de R.

É evidente, portanto, a dupla limitação até agora existente para a aplicação da técnica: a limitação a um número pré-fixado de capturas e a natural imprecisão na utilização dos gráficos, cuja menor divisão em R e em p é de 0.05.

Solução Alternativa

O presente trabalho propõe uma solução numérica para a expressão (2), baseada no tradicional e bem conhecido método de Newton-Raphson. (Claudio e Marins¹, 1989). Trata-se de processo iterativo brevemente esboçado a seguir, com condições de convergência bem conhecidas, de ordem 2 para raízes de multiplicidade um.

Nesta técnica calcula-se inicialmente a derivada de primeira ordem da expressão (2) para F(q), obtendo-se:

$$(5) \quad F'(q) = \frac{-1}{(1 - q)^2} + \frac{k^2 \cdot q^{k-1}}{(1 - q^k)^2}$$

Dado um valor inicial, arbitrário, para q, um valor mais aproximado é obtido pela expressão (6) abaixo, e o processo reiniciado, até que a diferença entre dois valores de q consecutivos não ultrapasse

* Departamento de Epidemiologia da Faculdade de Saúde Pública da Universidade de São Paulo - Av. Dr. Arnaldo, 715 - 01255 - São Paulo, SP - Brasil.

** Departamento de Matemática e Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo - Rua do Matão, 1010 - 05508 - São Paulo, SP - Brasil.

uma dada tolerância. O valor de q no passo $(j+1)$ -ésimo, denotado por $q(j+1)$, é relacionado com o valor no passo anterior, $q(j)$, pela expressão:

$$q(j+1) = q(j) - \frac{F(q(j))}{F'(q(j))} \quad (6)$$

No presente trabalho variou-se a tolerância de 0.01 até 0.00000001. Isto é, os ciclos prosseguiram até que a diferença em valor absoluto de dois valores consecutivos de q se situasse na abaixo da tolerância fixada. Southwood² (1978) trabalha até a segunda casa decimal, ou seja, até a menor tolerância aqui estudada.

Note-se também que a escolha do valor inicial para q e a própria convergência do processo não apresentam maiores problemas, já que para $k > 2$ e $R < (k-1)/2$ a equação $F(q) = 0$ tem uma única raiz no intervalo $0 < q < 1$ (Ver Apêndice).

Comparação de Resultados

O exemplo ilustrado por Southwood² (1978) servirá para as presentes comparações: tratam-se de cinco capturas que resultaram na coleção respectivamente de 65, 43, 34, 18 e 12 espécimes, resultando num valor de R , segundo a equação (1), de 1,2384. Na Tabela, são comparados os resultados exemplificados pelo autor com os obtidos pela técnica proposta.

Observa-se que com tolerância de 10^{-4} os resultados praticamente se estabilizam, obtendo-se um valor de $F(q)$ bem próximo de zero para todos os efeitos práticos. Os 27 ou 37 ciclos necessários para tolerâncias maiores são alcançados em frações de segundo nos computadores comuns. Cabe observar, no entanto, que com apenas um ciclo, isto é, com um trabalho que se pode realizar em calculadoras comuns, o valor de $F(q)$ é mais próximo de zero do que na técnica tradicional.

A dependência do número de ciclos com o valor inicial não é crítica face ao exposto acima. Nestes cálculos o valor inicial foi fixado em 0,65, o que parece ser uma boa sugestão de ordem geral, já que é difícil de esperar-se probabilidades de captura superiores a 50%.

Cabe ainda observar que o valor do total populacional N na nova técnica é cerca de 2% inferior ao tradicional. Já a diferença de valores nos erros-padrão chega próximo dos 40%.

O exemplo referido parece ter servido para demonstrar maior simplicidade e melhor qualidade da solução numérica aqui proposta em relação a técnica tradicional.

Tabela . Valores obtidos pela técnica de Zippin tradicional e pela nova técnica, e n° ciclos necessários.

Valores	Tradicional	Tolerância da Nova Técnica			
		10^{-2}	10^{-4}	10^{-6}	10^{-8}
P	0,33	0,3454	0,3349	0,3347	0,3347
F(q)	0,09	0,0375	0,0007	0,0000	0,0000
N	202	195	198	198	198
S	14,46	9,5551	10,3747	10,3917	10,3917
Ciclos		1	13	27	37

SANTOS, J. L. F. & OLIVA, W. M. [On Zippin's procedure for the estimation of animal populations]. *Rev. Saúde públ.*, S. Paulo, 25: 53-5, 1991. Zippin's procedure is one of the most accurate methods used to estimate total populations from removal trapping data. However, explicit algebraic solutions of the maximum likelihood equations do not exist. Hitherto they have been obtained from graphs known as Zippin charts. In this paper a Newton-Raphson procedure is used to obtain numerical solutions to the equations. Comparisons with the classical method show some differences in the parameters, particularly in the estimation of the standard error.

Keywords: Populations estimates. Estimation technics.

Referências Bibliográficas

- 1 - CLAUDIO, D.M. & MARINS, J.M. *Cálculo numérico computacional*. São Paulo, Atlas, 1989.
- 2 - SOUTHWOOD, T.R.E. *Ecological methods*. London, Chapman & Hall, 1978.

Recebido para publicação em 18/5/1990.
Reapresentado em 9/10/1990.
Aprovado para publicação em 10/10/1990.

Apêndice

Neste Apêndice são apresentados um Lema e um Teorema que mostram a completa viabilidade da utilização da técnica de Newton Raphson como alternativa numérica para o procedimento de Zippin.

LEMA:

Se $k \geq 2$, a derivada da função $F(q)$ é negativa no intervalo $0 < q < 1$. Além disso, $F'(0) = -1$ e

$$\lim_{q \rightarrow 1} F'(q) = 0$$

PROVA

$$F'(q) = \frac{-1}{(1-q)^2} + \frac{k^2 \cdot q^{k-1}}{(1-q^k)^2} = \frac{k^2 q^{k-1} \cdot (1-q^k)^2}{(1-q)^2 (1-q^k)^2} \quad (7)$$

como $(1-q^k) = (1-q) \{1 + q + q^2 + \dots + q^{k-1}\}$ temos:

$$F'(q) = \{k^2 q^{k-1} - (1 + q + q^2 + \dots + q^{k-1})^2\} / (1-q^k)^2$$

$$= N(q, k) \{k \cdot q^{(k-1)/2} + (1 + q + \dots + q^{k-1})\} / (1-q^k)^2 \quad (8)$$

onde $N(q, k) = \{k \cdot q^{(k-1)/2} - (1 + q + \dots + q^{k-1})\}$

É fácil mostrar diretamente que, para $k = 2, 3, 4, \dots$, o fator $N(q, k)$ é negativo no intervalo $0 < q < 1$. Mostremos, por indução finita sobre k , que ele é negativo para todo $k \geq 2$ e $0 < q < 1$. Para isso, admitindo ser $N(q, k) < 0$ deveremos provar que $N(q, k+1)$ também é negativo. De fato,

$$N(q, k + 1) = (k + 1)q^{k/2} - (1 + q + \dots + q^k) =$$

$$= N(q, k) + \{ (k+1)q^{k/2} - q^k - kq^{(k-1)/2} \} =$$

$$= N(q, k) - \{ q^{(k+1)/2} - (k+1)q^{1/2} + k \} \cdot q^{(k-1)/2}$$

Denotando $x = q^{1/2}$ o lema estará provado desde que se mostre que a função

$$f(x) = x^{k+1} - (k+1) \cdot x + k$$

permanece positiva para todo x entre zero e um. Para a verificação deste último fato basta observar que $f(0) = k, f(1) = 0$ e que

$$f'(x) = (k+1)x^k - (k+1) = (k+1)(x^{k-1} - 1) < 0$$

para $0 < x < 1$. Finalmente, pela regra de L'Hospital é facilmente verificado que

$$\lim_{q \rightarrow 1} F'(q) = 0$$

TEOREMA

Se $k \geq 2$ e $R < (k-1)/2$ a equação $F(q) = 0$ tem uma única raiz no intervalo $0 < q < 1$.

PROVA

Pela fórmula (2) verifica-se que $F(0) = R$ que é positivo pela definição de R . Para a conclusão da prova é suficiente utilizar o lema, mostrando antes que o limite de $F(q)$ para q tendendo a 1 vale $R - (k-1)/2$, e este último é, por hipótese, negativo. Mas,

$$F(q) - R = \frac{-q}{1-q} + k \cdot q^k / (1-q^k) =$$

$$= \frac{-q \cdot (1 + q + \dots + q^{k-1}) + k \cdot q^k}{(1-q^k)} =$$

$$= \frac{(k-1) \cdot q^k - (q + q^2 + \dots + q^{k-1})}{(1-q^k)} \quad (9)$$

Usando a regra de L'Hospital vê-se que

$$\lim_{q \rightarrow 1} F(q) - R =$$

$$= \frac{\lim_{q \rightarrow 1} \{k(k-1)q^{k-1} - (1 + 2q + 3q^2 + \dots + (k-1)q^{k-2})\}}{\lim_{q \rightarrow 1} [-kq^{k-1}]}$$

$$= \frac{k(k-1) - [1 + 2 + 3 + \dots + (k-1)]}{-k} =$$

$$= \frac{k(k-1) - k(k-1)/2}{-k} = -\frac{(k-1)}{2}$$