

Módulo 7(continuação)

Estimação de parâmetros

7.9 Intervalos de confiança

Quando a população é sondada e não recenseada, as estatísticas da LT, designadamente S_x e l_x , são também estimativas e, por isso, requerem que as suas propriedades estatísticas sejam tidas em atenção. Quando é efectuada uma estimativa da sobrevivência, esta deve ser apresentada com a informação estatística necessária para o leitor julgar por si próprio a confiança que pode depositar na referida estimativa. Por outras palavras, deve ser apresentada com uma variância e um intervalo de confiança.

Recorde-se S_x e q_x , definidas por [4.1] e [4.5], respectivamente. Estas taxas definem a probabilidade de que cada indivíduo que chega vivo à idade x pertença a uma de duas categorias, sobreviver ou não-sobreviver, enquanto tem x a $x+\Delta x$ anos de idade. Se N_x fôr uma *amostra aleatória* de todos os indivíduos chegados à idade x da população, então a probabilidade de k destes indivíduos sobreviverem até $x+\Delta x$ e N_x-k morrerem, é dada pela distribuição binomial, com parâmetros S_x e $1-S_x=q_x$. A variância de uma estimativa de S_x ou de q_x é dada pela variância da binomial:

$$\text{Var}(S_x) = \text{Var}(q_x) = \frac{S_x q_x}{N_x} \quad [7.16]$$

Como S_x e q_x são eles próprios estimativas, o estimador centrado da variância é (Zar 1984),

$$\tilde{\text{Var}}(S_x) = \tilde{\text{Var}}(q_x) = \frac{\tilde{S}_x \tilde{q}_x}{N_x - 1} \quad [7.17]$$

onde o til significa que se trata duma estimativa do correspondente valor na população.

O mesmo raciocínio pode ser usado para l_x . Isto é, a probabilidade de que cada recém-nascido pertença a uma de duas categorias, sobreviver ou não sobreviver, enquanto tem 0 a x anos de idade. Se N_0 fôr uma amostra aleatória de todos os recém-nascidos, a probabilidade de k destes sobreviverem até x , é dada pela distribuição binomial, com parâmetros l_x e $1-l_x$. A variância de l_x é estimada por,

$$\tilde{\text{Var}} l_x = \frac{\tilde{l}_x (1 - \tilde{l}_x)}{N_0 - 1} \quad [7.18]$$

Podem-se obter limites de confiança para os parâmetros S_x , q_x e l_x , directamente de [7.15] e [7.18], usando a aproximação da distribuição normal à binomial. Para um valor de N_x suficientemente grande (diga-se >30), os limites inferior e superior do intervalo de confiança para S_x são aproximadamente dados por,

$$S_x \pm z_{\alpha/2} \sqrt{\text{Var}(S_x)} \quad [7.19]$$

onde S_x é a estimativa obtida pelos métodos da secção 4.1 e $z_{\alpha/2}$ é o quantil da distribuição normal reduzida, $N(0,1)$ para lá do qual a área sob a curva é $\alpha/2$. Para um intervalo de confiança de 95%, tem-se $\alpha=0.05$ e $\alpha/2=0.025$, a consulta de uma tabela da distribuição $N(0,1)$ indica que $z_{\alpha/2}=1.96$. Finalmente, o termo com a raiz quadrada é o desvio-padrão de S_x , calculado a partir de [7.17]. O mesmo raciocínio pode ser aplicado para construir intervalos de confiança para q_x e l_x , usando ainda o desvio-padrão derivado de, respectivamente, [7.17] e [7.19], e o valor de $z_{\alpha/2}$.

A tabela 7.4 tem um exemplo numérico que ilustra a construção de intervalos de confiança a 95% para S_x e l_x , baseados na distribuição binomial. Partindo de um número inicial de indivíduos, $N_0=100$, que decresce sucessivamente (99, 50, ...), foram calculadas as taxas de sobrevivência, os desvios-padrões (usando [7.16] e [7.18]) e os intervalos de confiança usando [7.19].

Tabela 7.3. Cálculo de intervalos de confiança para as taxas de sobrevivência, S_x e l_x , numa coorte que se inicia com $N_x=100$ indivíduos. As variâncias binomiais (Var Binm) são calculados usando [7.16] e [7.18] e os desvios-padrões (dp Binm) são a raiz quadrada da variância. Os limites inferior (L inf) e superior (L sup) dos intervalos de confiança (IC Binm) para S_x e l_x são calculados usando [7.19].

N_x	S_x		l_x		IC Binm p/ S_x		IC Binm p/ l_x			
	S_x	l_x	Var Binm	dp Binm	Var Binm	dp Binm	L Inf	L Sup	L Inf	L Sup
100	0.99	1	0.0001	0.0099	0	0	0.9705	1.0095	1	1
99	0.51	0.99	0.0025	0.0502	0.0001	0.0099	0.4115	0.6085	0.9705	1.0095
50	0.4	0.5	0.0048	0.0693	0.0025	0.0500	0.2642	0.5358	0.4020	0.5980
20	0.25	0.2	0.0094	0.0968	0.0016	0.0400	0.0602	0.4398	0.1216	0.2784
5	0.2	0.05	0.0320	0.1789	0.0005	0.0218	-0.1506	0.5506	0.0073	0.0927
1	0	0.01	0	0	0.0001	0.0099	0.0000	0.0000	-0.0095	0.0295
0		0			0	0	0	0	0	0

Infelizmente, há dois problemas com os intervalos de confiança calculados por [7.19] que são visíveis na Tabela 7.4. Primeiro, os limites dados por [7.19] podem gerar intervalos que saiam para fora dos limites $[0, 1]$ admissíveis para uma taxa de sobrevivência. O segundo problema surge quando uma das estimativas, S_x ou l_x , é igual a 0 ou 1. Neste caso, a variância dada por [7.16] ou [7.18] também é zero e os limites do intervalo de confiança ficam iguais ao próprio parâmetro estimado. Uma forma simplista de resolver o primeiro problema, consiste em substituir o limite do intervalo de confiança que ultrapassa 0 ou 1 por, respectivamente, 0 ou 1. Este procedimento, porém, tem o inconveniente de tornar a amplitude do intervalo (i.e. a diferença entre o limite superior e inferior) artificialmente pequena. Há formas melhores de lidar com o problema, algumas das quais bastante sofisticadas (Ebert 1999). A técnica apresentada a seguir, foi seleccionada pela facilidade com que pode ser programada numa folha de cálculo (excel ou outra) e porque foi recentemente recomendada por Agresti and Coull (1998) e

por Brown *et al* (2001). Além disso, os seus méritos não dependem muito do tamanho da amostra ou do valor verdadeiro do parâmetro que se pretende estimar (n e p da distribuição binomial). A sua dedução é directa, embora requiera longa manipulação algébrica. No fim, tudo se resume a uma fórmula que será apresentada sem explicação teórica.

Designa-se por θ_x a estimativa da taxa da idade x (θ_x pode ser uma estimativa de S_x , q_x ou l_x). Esta estimativa foi obtida a partir de um número inicial de indivíduos, N_x (o tamanho da amostra). Seja $z_{\alpha/2}$, como há pouco, o quantil da distribuição normal reduzida, $N(0,1)$. Por exemplo, para um IC de 95%, já vimos que $z_{\alpha/2}=1.96$. Os limites inferior e superior que delimitam o IC dentro do qual o verdadeiro valor da taxa deve estar com $(1-\alpha)\%$ de confiança são dados por,

$$\frac{\theta_x + \frac{z_{\alpha/2}^2}{2N_x} \pm z_{\alpha/2} \sqrt{\frac{\theta_x(1-\theta_x)}{N_x} + \frac{z_{\alpha/2}^2}{4N_x^2}}}{1 + \frac{z_{\alpha/2}^2}{N_x}} \quad [7.20]$$

Para exemplificar a aplicação deste método, retomo dois exemplos numéricos da Tabela 7.4. No primeiro, tomo $N_0=100$ e $N_3=20$, donde $l_3=0.20$, e vou calcular os limites de um intervalo de confiança a 95% ($\alpha/2=0.025$) para l_3 . Neste caso, $\theta_3= 0.20$, N_0 é o tamanho da amostra e $z_{\alpha/2}=1.96$. A aplicação de [7.20] dá directamente o seguinte IC: [0.1334, 0.2888], o qual está um pouco mais “encostado à direita” do que o intervalo [0.1216, 0.2784] da Tabela 7.4. Além disso, ao contrário do IC binomial, não é simétrico em torno de 0.2.

Num segundo exemplo, considero $N_0=100$ e $N_1=99$, donde $S_0=0.99$, e calculo os limites para um IC a 95% em torno de S_0 . Recorde-se que, na Tabela 7.4, o limite superior do IC neste caso deu superior a 1, o que não é admissível. A aplicação de [7.20] dá o seguinte IC: [0.9455, 0.9982], o qual já é admissível.

A Tabela 7.5 retoma a coorte da Tabela 7.4 e recalcula os intervalos de confiança para S_x e os quais são comparados com os anteriores intervalos binomiais.

Tabela 7.4. Intervalos de confiança para a taxa de sobrevivência, S_x , duma coorte que se inicia com $N_0=100$ indivíduos. Apresentam-se os limites inferior (L inf) e superior (L sup) dos intervalos baseados na distribuição binomial (da Tab. 7.4) e baseados na equação [7.20] de Agresti-Coull.

N_x	S_x	Intervalo binomial		Intervalo Agresti-Coull	
		L inf	L sup	L inf	L sup
100	0,99	0,970	1,010	0,946	0,998
99	0,51	0,412	0,608	0,408	0,601
50	0,4	0,264	0,536	0,276	0,538
20	0,25	0,060	0,440	0,112	0,469
5	0,2	-0,151	0,551	0,036	0,624
1	0	0	0	0	0,793

Para valores intermédios da sobrevivência (~ 0.4-0.6), não existe grande diferença na amplitude e nos limites dos intervalos obtidos pelos dois métodos. Perto dos limites 0 e 1, porém, os intervalos Agresti-Coull são assimétricos e não saem para fora daqueles limites, ao contrário dos intervalos binomiais que mantêm a simetria a que a equação [7.19] obriga. Nos casos extremos em que a sobrevivência é igual a 0 ou a 1, só os intervalos Agresti-Coull são aceitáveis.

7.10 Variância nos métodos Kaplan-Meier e Nelson-Aalan

Para estimar a variância em torno de uma estimativa da sobrevivência no método Kaplan-Meier, há a considerar dois casos. Primeiro, a variância para as estimativas de sobrevivência por idade, S_x , pode ser estimada pela fórmula habitual, [7.17], para a variância de proporções,

$$Var(S_x) = \frac{S_x(1-S_x)}{N_x - 1}$$

Por exemplo, nos ursos (sec 7.2), a variância da sobrevivência da idade 10,5 é,

$$Var(S_{10.5}) = \frac{S_{10.5}(1-S_{10.5})}{N_{10.5} - 1} = \frac{(0.889)(0.111)}{9 - 1} = 0.012$$

A construção de intervalos de confiança para S_x pode depois prosseguir normalmente, seguindo as orientações da secção anterior.

O segundo caso é o da variância de estimativas de sobrevivência até à idade x (l_x) que resultam, elas próprias, do produto de outras sobrevivências, como é o caso de l_x no método KM, em que não se acompanha sempre a mesma coorte. Suponhamos que l_x foi estimado a partir das sobrevivências das idades $i, i+1, i+2, \dots, x-1$. A estimativa da variância de l_x pode ser obtida pela chamada “fórmula de Greenwood” (Skalski *et al* 2005)¹,

$$Var(l_x) = l_x^2 \sum_i^{x-1} \frac{D_i}{N_i N_{i+1}} \quad [7.21]$$

Sendo D_i o número de mortes com idade i . No exemplo dos ursos, a variância da sobrevivência entre as idades 2.5 e 5.5, por exemplo, é dada por,

$$Var(l_{[2.5, 5.5]}) = (0.612)^2 \left[\frac{0}{(7)(7)} + \frac{2}{(7)(5)} + \frac{1}{(7)(6)} \right] = 0.030$$

No método Nelson-Aalan, o estimador da sobrevivência era, recorde-se,

¹ A fórmula de Greenwood é a mais popular, mas Cox and Oakes (1984) *Analysis of Survival Data*, Chapman And Hall, argumentam que a seguinte fórmula tem um comportamento preferível: $Var(l_x) = l_x^2(1-l_x)/N_x$

$$l_x = e^{-\sum_{i=0}^x \frac{D_i}{N_i}}$$

A variância deste estimador, pode ser aproximativamente calculada por (Skalski *et al* 2005),

$$Var(l_x) = l_x^2 \sum_{i=0}^x \frac{D_i}{N_i^2} \quad [7.22]$$

7.11 Tamanho da amostra

Considere-se agora o tamanho da amostra necessário ao cálculo duma taxa de sobrevivência. Suponhamos que estamos interessados em estimar a taxa de sobrevivência entre a idade x e $x+\Delta x$, mas não queremos cometer um erro maior do que a quantidade E (por exemplo, $E=0.1$). Suponhamos ainda que o risco que estamos dispostos a correr de que o referido erro seja superior a E é $\alpha=0.05$, isto é, em cada 20 vezes que fizéssemos a experiência, só uma vez o erro seria superior a E . Qual é o número mínimo de animais, N_x , que temos de amostrar para assegurar este objectivo? A teoria estatística da amostragem (Cochran 1977) indica-nos o valor mínimo que N_x deve ter:

$$N_x = \frac{z_\alpha^2 S_x (1 - S_x)}{E^2} \quad [7.23]$$

Sendo z_α o valor tabelado da distribuição normal reduzida ($z_\alpha=1.96$, para $\alpha=0.05$). Na prática, como S_x não é conhecido, é necessário substituí-lo em [7.21] por um valor estimado. Caso não haja a mínima indicação do valor provável de S_x , uma alternativa de último recurso a [7.23] é,

$$N_x = \frac{0.25 z_\alpha^2}{E^2} \quad [7.24]$$

Se houver razões para crer que o valor de N_x , estimado por [7.23] ou [7.24], não é muito inferior ao número total de indivíduos que na população têm x anos de idade, então a amostra não necessita de ser tão grande. Designando por $N_{x \text{ pop}}$ o número total na população, o tamanho final da amostra, M_x , pode ser corrigido por:

$$M_x = \frac{N_x}{1 + \frac{N_x - 1}{N_{x \text{ pop}}}} \quad [7.25]$$

7.12 Populações exploradas

Frequentemente, o biólogo tem de recorrer a dados de capturas efectuadas sobre a população, por razões comerciais ou desportivas, para estimar parâmetros demográficos e, a partir daí, modelar a dinâmica da população e, eventualmente, geri-la. É importante portanto compreender as vantagens e desvantagens deste tipo de dados. A vantagem mais óbvia é que os caçadores (ou pescadores) ajudam à colecta dos dados. A captura de animais pode decorrer ao longo áreas geográficas muito extensas e/ou ocupar uma grande parte do ano. Uma sondagem que fosse efectuada para fins científicos com uma ocupação espaço-temporal semelhante, seria proibitivamente cara. Comparativamente, a amostragem das capturas é um procedimento muito mais barato e, em geral, logisticamente mais simples. Por esta razão, os dados de capturas pela caça e pela pesca são com frequência uma das principais fontes de informação em que se baseia a gestão de populações exploradas.

Apesar destas vantagens, há importantes desvantagens no uso dos dados de capturas. Os caçadores e os pescadores não tomam amostras aleatórias da população explorada. Há vários factores que podem influenciar as características dos animais que são capturados, nomeadamente, a legislação relativa à exploração, os meios de captura de que se dispõe (armadilhas ou armas, anzóis ou redes de arrasto ...), o valor económico dos diferentes tipos de animais, os factores climáticos e a própria abundância dos animais. A disponibilidade e a vulnerabilidade à captura também não é a mesma entre animais de diferentes tamanhos, sexos, ou ao longo dos anos. Evidentemente, a eventual relutância de caçadores e pescadores em relatar o que capturaram com exactidão, como e onde, também não ajuda, mesmo que a legislação obrigue a que o façam. Por estas razões, a gestão de grandes populações exploradas é, em geral, feita com base na combinação das "espreitadelas ao saco do caçador" com sondagens científicas independentes, realizadas de forma a obter amostras tão representativas da população que está no campo quanto possível.

Populações marinhas exploradas

A estimação da abundância das grandes populações marinhas coloca dificuldades muito particulares. É em geral impossível observar os indivíduos directamente, a sua distribuição geográfica é em geral muito extensa e, quando se efectuam marcações de animais, as taxas de recaptura são habitualmente muito baixas (<2%). Estas dificuldades, aliadas ao facto de muitas destas populações terem grande interesse económico, são talvez a razão por que os biólogos que trabalham com populações marinhas foram, ao longo dos últimos 50 anos, particularmente activos na elaboração de teoria e métodos alternativos para estimar taxas de sobrevivência (Gulland 1955, 1969; Ricker 1975, Pauly 1984, King 1995)

A estimação da sobrevivência de populações exploradas pela pesca ou pela caça, depende fortemente da informação recolhida a partir dos animais capturados. É intuitivo que a estrutura etária das capturas deve ser muito informativa acerca da estrutura etária da população, *desde que as capturas sejam uma amostra representativa* da população. Em geral é razoável pressupôr que as capturas são uma amostra representativa a partir de determinada idade, a chamada idade de recrutamento à exploração. A partir desta idade, C_x/C_T , isto é, a proporção de animais capturados com idade x , C_x , relativamente à captura total, C_T , deve ser idêntica à correspondente proporção na população, $c_x=N_x/N_T$, desde que as capturas ocorram aleatoriamente. Uma das técnicas mais utilizadas nas décadas de 1960-70 foi a chamada "curva de capturas" (the catch curve). Baseia-se na relação que, numa população estacionária, deve existir entre o decréscimo das capturas por idade (C_x versus x)

e a sobrevivência por idade, assumindo que a distribuição etária das capturas é representativa da distribuição etária na população. Ricker (1975) expõe os detalhes do método.

A utilização de relações empíricas que envolvem taxas de sobrevivência ou de mortalidade também tem sido comum em espécies aquáticas. Uma relação empírica é uma relação estatística que se constata que existe entre os parâmetros biológicos ou ambientais das espécies. A explicação para estas relações é em geral mal compreendida, mas o seu carácter predictivo confere-lhes grande utilidade prática (Peters 1983). Por exemplo, Hoenig (1984) reuniu dados de grande número de moluscos, peixes e cetáceos, e concluiu que nas espécies marinhas existe a seguinte relação entre a taxa média de sobrevivência total, S_t , e a idade máxima observada na espécie ($x_{\text{máx}}$):

$$\ln S_t = -e^{1.44 - 0.984 \ln x_{\text{máx}}}$$

A aplicação directa da fórmula indica que, se o indivíduo mais velho que foi observado na espécie tivesse 1, 5 e 10 anos de idade, a sobrevivência total média da espécie (eqs 2.4 e 2.5) seria, respectivamente, 0.015, 0.42 e 0.645.

Num outro exemplo, Hoenig *et al.* (1983) apresentaram uma relação empírica entre S_t , a mediana dos comprimentos dos indivíduos da população, Y_m (estimada a partir de amostras de capturas), dois dos parâmetros da equação de crescimento de Von Bertalanffy (K , L_∞) e, finalmente, o comprimento médio com que os indivíduos são capturados pela primeira vez, l_c :

$$\ln S_t = -\frac{K \ln 2}{Y_m - \ln\left(1 - \frac{l_c}{L_\infty}\right)}$$

7.13 O método Michod-Anderson para estimar r e sobrevivência a partir de frequências etárias.

(NÃO leccionado nas aulas, NÃO vem para avaliação de conhecimentos)

No módulo 4 introduzi o conceito de população estacionária e discuti a possibilidade de aplicar a teoria da coorte à população, assumindo que esta se encontra estacionária. Esta aplicação tem bastante interesse, dado ser frequente o ecologista não dispôr de várias observações da estrutura etária da população espaçadas no tempo, de forma a poder seguir a sobrevivência das coortes. De facto, quando existem estimativas do número de indivíduos por idade num único ano, t , do calendário (simbolicamente, $N_{0,t}$, $N_{1,t}$, ..., $N_{x,t}$, ...), o método habitualmente usado para estimar as taxas de sobrevivência consiste em assumir que a população atingiu DEE e que $r=0$, i.e., está estacionária. Como a estrutura etária da população nestas circunstâncias é igual à da coorte, as taxas de sobrevivências são então calculadas por,

$$S_x = \frac{N_{x+1,t}}{N_{x,t}} \quad \text{e} \quad l_x = \frac{N_{x,t}}{N_{0,t}} \quad [A1]$$

O pressuposto de estacionaridade é bastante forte e, se possível, convém dispôr de alternativas menos restritivas que também permitam extrair informação demográfica relevante a partir da estrutura etária vertical da população num único ano. Nesta secção apresento um método que permite estimar o parâmetro r e as taxas de sobrevivência a partir da estrutura etária ($N_{0,t}$, $N_{1,t}$, ..., $N_{x,t}$, ...) observada no ano t , sem que seja necessário assumir $r=0$. Basta assumir que a população tem crescido com LT constante, tendo atingido a DEE (mas não necessariamente a estacionaridade) e, simultâneamente, dispôr de informação sobre a fertilidade da população por idade, isto é m_x . O método deve-se a Michod and Anderson (1980)² e é concebido para reprodutores sazonais.

Se a população se encontra em DEE, o número esperado de indivíduos no *inicio* da idade x , no instante t , é igual aos nascimentos ocorridos há x anos atrás, B_{t-x} , multiplicados pela probabilidade de sobreviver até à idade x ,

$$B_{t-x} l_x$$

Em geral, porém, a sondagem da população não é efectuada imediatamente após a época de reprodução, i.e., no inicio da idade. Seja t o instante em que se dá a sondagem e designe-se por p a *fracção* da unidade de tempo considerada que decorre desde a época de reprodução até ao momento t . Por exemplo, se o intervalo de tempo considerado, $(t, t+1)$, são 12 meses e a sondagem ocorreu 3 meses após a época de reprodução, então $p=3/12=0.25$. Designe-se por l_p a probabilidade de um recém-nascido permanecer vivo até à sondagem. O número esperado de indivíduos na idade $x=0$, na altura da sondagem em t , deve ser igual ao número de nascimentos ocorridos no intervalo anterior $(t-1, t)$, multiplicado por l_p :

$$N_{0,t} = B_{t-1} l_p \quad [A2]$$

O número esperado de indivíduos na idade $x=1$, na altura da sondagem, deve ser,

$$N_{1,t} = B_{t-2} l_{1+p}$$

onde l_{1+p} é a probabilidade de um recém-nascido fazer 1 ano de idade e depois permanecer vivo até à sondagem. O número esperado na idade $x=2$ será,

$$N_{2,t} = B_{t-3} l_{2+p}$$

Em geral, na idade x , espera-se o seguinte número de indivíduos,

$$N_{x,t} = B_{t-(x+1)} l_{x+p}$$

² Michod, RE and Anderson WW. 1980. On calculating demographic parameters from age frequency data. *Ecology* 61:265-269

Recordando a relação [7.5], válida para uma população em DEE, isto é, $B_t = B_{t-x}e^{rx}$, o número esperado de indivíduos na idade x no instante da sondagem é,

$$N_{x,t} = B_t e^{-r(x+1)} l_{x+p} \quad [A3]$$

No intervalo de tempo $(t, t+1)$, cada indivíduo da idade x contribui com m_x descendentes para a população. O número total de nascimentos ocorridos no intervalo $(t, t+1)$, isto é B_t , deve ser o somatório das contribuições de todas as idades,

$$B_t = \sum_{x=0}^L B_t e^{-r(x+1)} l_{x+p} m_x$$

dividindo tudo por B_t ,

$$1 = \sum_{x=0}^L e^{-r(x+1)} l_{x+p} m_x \quad [A4]$$

A equação [A4] corresponde à equação de Lotka [7.8], quando o instante em que se dá a sondagem do número de indivíduos por idade está “deslocado” a fracção p “para a frente”, relativamente à época de reprodução.

Retome-se a equação [A3] e resolva-se em ordem a l_{x+p} ,

$$l_{x+p} = \frac{N_{x,t} e^{r(x+1)}}{B_t}$$

De novo, uma vez que $B_t = B_{t-x}e^{rx}$,

$$l_{x+p} = \frac{N_{x,t} e^{r(x+1)}}{B_{t-1} e^r} = \frac{N_{x,t} e^{rx}}{B_{t-1}}$$

Recordando [A2], tem-se finalmente,

$$l_{x+p} = \frac{N_{x,t} e^{rx}}{N_{0,t}} l_p \quad [A5]$$

Quer dizer, se fôr conhecido r e l_p , a equação [A5] permite obter estimativas de l_{x+p} , a taxa de sobrevivência dos recém-nascidos até à idade $x+p$, a partir da estrutura etária da população no instante t . Com o objectivo de obter uma estimativa de r , substitua-se [A5] em [A4],

$$1 = \sum_{x=0}^L \frac{N_{x,t}}{N_{0,t}} e^{-r} l_p m_x$$

isto é,

$$e^r = \sum_{x=0}^L \frac{N_{x,t}}{N_{0,t}} l_p m_x \quad [A6]$$

A equação [A6] permite estimar r a partir da estrutura etária da população, da tabela de fertilidade por idade e de l_p . Uma vez conhecido r , este pode ser usado na equação [A5] para estimar sobrevivência por idade.

Para estes procedimentos é necessário ter uma estimativa independente de l_p , a probabilidade de um recém nascido sobreviver a fração de tempo p . Quando este valor não é conhecido, muitas vezes será mais razoável assumir um valor numérico para l_p do que assumir que a população se encontra estacionária ($r=0$) como muitas vezes é feito. Se, no entanto, o ecologista preferir assumir estacionaridade, a sobrevivência deve ser estimada por [A5] e não por [A1]. Note-se que a utilização de [A1] pressupõe que $l_p=1$, isto é, recenseamento imediatamente após reprodução, ou então o que se está a estimar é l_{x+p}/l_p .

Finalizo recapitulando os pressupostos e os cálculos. Assumindo que a população está em DEE e que são conhecidos: $(N_{0,t}, N_{1,t}, \dots, N_{x,t}, \dots)$, $(m_0, m_1, \dots, m_x, \dots)$ e l_p . Primeiro, obter o valor de r a partir de [A6], a seguir obter a sobrevivência das idades $x+p$, a partir de [A5].

Literatura Citada

- Aalen, O. 1978. Nonparametric inference for a family of counting processes. *Annals of Statistics* **6**:701-726.
- Agresti, A, and BA Coull. 1998. Approximate is better than “exact” for interval estimation of binomial proportions. *Amer Statistician* **52**(2): 119-126
- Brown, LD, TT Cai, and A DasGupta. 2001. Interval estimation for a binomial proportion. *Statistical Science* **16**(2): 101-133
- Carey, J.R. 1993. *Applied Demography for Biologists*. Oxford Univ. Press, Oxford.
- Caswell, H. 1989. *Matrix Population Models*. Sinauer, Sunderland, Mass.
- Cochran, W. 1977 (3rd ed). *Sampling Techniques*. John Wiley.
- Dempster, JR. 1975. *Animal Population Ecology*, Academic Press, New York.
- Ebert, T.A. 1999. *Plant and Animal Populations. Methods in Demography*. Academic Press, San Diego, Calif.
- Gulland, J.A. 1955. Estimation of growth and mortality in commercial fish populations. *U.K. Min. Agric. Fish., Fishery Invest. (Ser. 2)* **18**(9): 46 pp.
- Gulland, J.A. 1969. *Manual of Methods for Fish Stock Assessment. Part 1. Fish Population Analysis*. Food and Agriculture Organization of the UN, Man. Fish. Sci. **4**:1-154.
- Hoenig, J.M. 1984. Empirical use of longevity data to estimate mortality rates. *Fishery Bulletin* **81**(4).
- Hoenig, J.M., W.D. Lawing, and N.A. Hoenig. 1983. Using mean age, mean length, and median length data to estimate the total mortality rate. *Int. Counc. Explor. Sea. C.M.* **23**:1-11.
- INE. 1997. *Estimativas de População Residente*, Instituto Nacional de Estatística, Lisboa.
- INE. 1998. *Estatísticas Demográficas*, Instituto Nacional de Estatística, Lisboa.
- King M. 1995. *Fisheries Biology, Assessment and Management*. Fishing News Books, Blackwell Sci., London.
- Klein, JP and ML Moeschberger. 1997. *Survival Analysis. Techniques for Censored and Truncated Data*. Springer-Verlag, NY.
- May RM. 1976. Estimating r: A pedagogical note. *American Naturalist* **110**:496-499.
- Michod RE and Anderson WW. 1980. On calculating demographic parameters from age frequency data. *Ecology* **61**:265-269.
- Nelson, W. 1972. Theory and applications of hazard plotting for censored failure data. *Technometrics* **14**:946-965.

- Pauly, D. 1984. *Fish Population Dynamics in Tropical Waters: a Manual for Use with Programmable Calculators*. ICLARM, Manila.
- Peters, RH. 1983. *The Ecological Implications of Body Size*. Cambridge Univ. Press, Cambridge.
- Pianka ER. 1978. 2nd Ed. *Evolutionary Ecology*. Harper & Row, NY.
- Pollock, KH, SR Winterstein, CM Bunck, and PD Curtis. 1989. Survival analysis in telemetry studies: The staggered entry design. *Journal of Wildlife Management* **53**:7-15.
- Pollard JH. 1973. *Mathematical Models for the Growth of Human Populations*. Cambridge Univ. Press, Cambridge.
- Ricker, W.E. 1975. *Computation and Interpretation of Biological Statistics of Fish Populations*. Fisheries Research Board of Canada, Bull. 191.
- Roughgarden J. 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. MacMillan Pub., NY.
- Skalski JR, KE Ryding, and JJ Millspaugh. 2005. *Wildlife Demography. Analysis of Sex, Age, and Count Data*. Elsevier, Academic Press, Amsterdam.
- Southwood TRE, 1978. *Ecological Methods*. Chapman and Hall, London.
- Yodzis P. 1989. *Introduction to Theoretical Ecology*. Harper & Row Publ., NY.
- Zar, JH. *Biostatistical Analysis*, 2nd Ed. Prentice Hall, New Jersey.