Estimadores de Máxima Verossimilhança para os Parâmetros de Alguns Modelos de Captura-Recaptura em População Animal

Mónica Inés Casajús

Dissertação Apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para Obtenção do Grau de Mestre
em
Estatística

Área de Concentração: **Estatística**

Orientador: **Prof. Dr. José Galvão Leite**

– SÃO PAULO, junho de 1990 –

A Carlos.

A minha mãe.

À professora Clara Badler.

Agradecimento

Desejo agradecer, nesta seção, a todos que de uma forma ou outra me ajudaram a realizar este trabalho.

Em primeiro lugar quero agradecer ao prof. Dr. José Galvão Leite por sua paciência e dedicação nestes dois anos e meio de orientação e por tudo o que aprendi com ele neste periodo.

À professora e amiga Elisabeth Kira por sua disposição para me corrigir o português.

Aos professores Antonio Luiz Pereira e Luiz Augusto F. de Oliveira pelo entusiasmo e empenho em descobrir as soluções dos problemas por mim apresentados e por sua ajuda na demonstração presente no Apêndice 5.

À professora Clara Badler por encorajar-me a todo momento, tanto no trabalho como no estudo.

A meu esposo Carlos por suas correções, por sua ajuda com os gráficos, por ter me ensinado a usar o editor de texto, e por seu apoio em todo momento.

Ao CNPq e à CAPES pelas bolsas concedidas, sem as quais este estudo não se realizaria.

Ao COSEAS que permitiu minha moradia no campus na Universidade de São Paulo facilitando, dessa forma, a elaboração deste trabalho.

SUMÁRIO

INTRODUÇÃO.

A estimação do tamanho de uma população animal é um problema de interêsse em ciências biológicas. Um dos métodos utilizados para estimar o tamanho de uma população animal é o método de captura-recaptura. Em geral, o método de captura-recaptura pode ser descrito da seguinte maneira. Duas ou mais amostras são selecionadas da população em estudo. Todos os animais capturados na primeira amostra são marcados (em estudos com aves, por exemplo, um anel é colocado em uma das patas). Em seguida, todos os animais marcados (ou uma parte deles) são devolvidos à população. Para cada uma das amostras subsequentes são registradas as capturas dos animais marcados e não marcados; todos os animais não marcados na amostra recebem uma marca e, em alguns casos, os animais anteriormente marcados (na amostra) também recebem uma nova marca, indicando a época de sua captura e, em seguida, todos os animais (ou uma parte deles) são devolvidos à população. Uma estimativa do tamanho da população pode ser obtida através do número de animais marcados recapturados em cada amostra.

Dependendo da situação pode-se considerar que o tempo utilizado para completar o procedimento de captura-recaptura é suficientemente pequeno para não ocorrerem nascimentos, mortes nem migrações na população. Neste caso, diz-se que a população é *fechada*. Por outro lado, se durante o procedimento a população se altera pela ocorrência de nascimentos, mortes e/ou migrações, a população se diz *aberta*. Neste trabalho consideraremos ambos os casos e os elementos que compõem a população serão referidos indistintamente por animais ou indivíduos.

O primeiro pesquisador a aplicar o método de captura-recaptura foi Petersen (1896), que estudou o fluxo migratório de peixes no mar Báltico. Contudo, Laplace (1783) já havia utilizado praticamente esse procedimento para estimar o tamanho da população francesa. Lincoln (1930) foi um dos primeiros pesquisadores, neste século, a usar o método de captura-recaptura para estimar o tamanho da população de patos da América do Norte. Entre otros autores que se dedicaram a este tema, além de Petersen e Lincoln, podemos citar Jackson (1953), Bailey (1951), Chapman (1951), Haldane (1953), Darroch (1958), Jolly (1965), Seber (1965), Cormack (1969), Manly (1968), Pollock (1976), Otis (1978), Burnham (1978), White (1978) e Anderson (1978).

O objetivo deste trabalho é fazer um estudo sistemático de alguns modelos baseados no método de captura-recaptura, assim como determinar os Estimadores de Máxima Verossimilhança para os parâmetros populacionais. O Capitulo 1 é dedicado à *amostragem simples* que consiste na seleção de duas amostras e o Capitulo 2 à *amostragem múltipla*: qualquer número (maior ou igual a dois) de amostras. Em ambos os casos estudamos os métodos de amostragem direta e inversa; o método de *amostragem direta* é aquele no qual os tamanhos das amostras são fixados, e o de *amostragem inversa* é aquele no qual o número de indivíduos marcados (ou não marcados) a ser capturados é fixado previamente. No Capitulo 3 consideramos um modelo alternativo com k amostras ($k \geq 2$) no qual cada indivíduo da população, independentemente dos demais, tem uma certa probabilidade p_i de ser capturado na i-ésima amostra ($1 \leq i \leq k$). Em todos esses capítulos supomos populações fechadas. No último capítulo consideramos populações abertas e determinamos os Estimadores de Máxima Verossimilhança (E.M.V.) para os parâmetros populacionais no caso de k amostras ($k \geq 3$), levando em conta, para cada amostragem, a última marca de cada indivíduo.

Neste trabalho N denotará sempre o tamanho da população em estudo.

CAPÍTULO 1

1. AMOSTRAGEM SIMPLES.

1.1. AMOSTRAGEM DIRETA SIMPLES.

O Método de Amostragem Direta Simples permite obter estimadores do tamanho da população e do inverso do tamanho da população. A técnica consiste em capturar sem reposição, marcar e soltar n_0 indivíduos da população. Depois de um certo tempo no qual espera-se que os indivíduos marcados misturem-se livremente com os não marcados, seleciona-se uma segunda amostra e observa-se a quantidade de marcados. Esta segunda amostra pode ser colhida com ou sem reposição.

Os indivíduos selecionados na segunda amostra não são marcados antes de serem devolvidos à população. Os tamanhos das amostras são fixados e o número de indivíduos marcados na segunda amostra é aleatório.

Baseados na informação destas duas amostras obtemos estimadores do tamanho da população, N, do inverso do tamanho da população, $\frac{1}{N}$, bem como as esperanças e variâncias destes estimadores.

a) COM REPOSIÇÃO

Sejam

n_0 : número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra;

n_1 : tamanho da segunda amostra, que é colhida com reposição;

m : número de indivíduos marcados observados na segunda amostra.

A função densidade de probabilidade de m é uma Binomial com parâmetros n_1 e $p = \frac{n_0}{N}$, isto é,

$$P(m \mid n_0, n_1, N) = \binom{n_1}{m} \left(\frac{n_0}{N}\right)^m \left(1 - \frac{n_0}{N}\right)^{n_1 - m} \quad ; \quad m = 0, 1, \cdots, n_1.$$

A média e variância de m são, respectivamente

$$E(m) = \frac{n_1 n_0}{N} \quad \text{e} \quad V(m) = n_1 \frac{n_0}{N}\left(1 - \frac{n_0}{N}\right)$$

Teorema 1. A estimativa de Máxima Verossimilhança (E.M.V.) para N, $\hat{N}_1$, é tal que $\hat{N}_1 = \frac{n_0 n_1}{m}$ se $\frac{n_0 n_1}{m}$ for um número inteiro e $\hat{N}_1 \cong \frac{n_0 n_1}{m}$ se $\frac{n_0 n_1}{m}$ não for um número inteiro.

Prova. A função de verossimilhança é

$$L(N \mid n_0, n_1, m) = \binom{n_1}{m} \left(\frac{n_0}{N}\right)^m \left(1 - \frac{n_0}{N}\right)^{n_1 - m}, \qquad N \geq n_0.$$

Logo,

$$L(N|n_0, n_1, m) \propto K(N) = \left(\frac{n_0}{N}\right)^m \left(1 - \frac{n_0}{N}\right)^{n_1 - m},$$

onde $K(N)$ é o kernel da verossimilhança e

1) se $n_1 = m \Rightarrow K(N) = (\frac{n_0}{N})^{n_1} \Rightarrow$ o ponto de máximo de $K(N)$ é

$$\hat{N}_1 = n_0 = \frac{n_0 n_1}{m};$$

2) se $n_1 > m \quad \Rightarrow \quad K(N) = (\frac{n_0}{N})^m (1 - \frac{n_0}{N})^{n_1 - m}$.

Seja $g(x) = x^m (1-x)^{n_1 - m} \quad , \quad x \in [0,1]$. Então, $g(\frac{n_0}{N}) = K(N)$ e derivando $g(x)$ temos:

$$\begin{aligned} g'(x) &= m x^{m-1}(1-x)^{n_1-m} - (n_1 - m) x^m (1-x)^{n_1-m-1} = \\ &= x^{m-1}(1-x)^{n_1-m-1}[m(1-x) - (n_1-m)x] = \\ &= x^{m-1}(1-x)^{n_1-m-1}(m - n_1 x), \quad x \in (0,1). \end{aligned}$$

Logo,

$$g'(x) > 0 \iff m - n_1 x > 0 \iff m > n_1 x \iff x < \frac{m}{n_1}.$$

$$g'(x) < 0 \iff m - n_1 x < 0 \iff m < n_1 x \iff x > \frac{m}{n_1}.$$

$$g'(x) = 0 \iff m - n_1 x = 0 \iff m = n_1 x \iff x = \frac{m}{n_1}.$$

Donde concluímos que $g(x)$ é crescente em $[0, \frac{m}{n_1}]$, decrescente em $[\frac{m}{n_1}, 1]$ e o ponto de máximo de $g(x)$ é $\hat{x} = \frac{m}{n_1}$.

Portanto, o ponto de máximo de $K(N)$, $\hat{N}_1$, é tal que $\hat{N}_1 = \frac{n_0 n_1}{m}$ se $\frac{n_0 n_1}{m}$ for um número inteiro e $\hat{N}_1 \cong \frac{n_0 n_1}{m}$ se $\frac{n_0 n_1}{m}$ não for um número inteiro, o que prova o teorema. ■

Considerando m como uma variável aleatória, notemos que o E.M.V. para N, $\hat{N}_1$ é aproximadamente o estimador de Petersen (1896). Este último estimador pode ser obtido igualando-se a proporção de animais marcados na população, antes da segunda amostra ser selecionada, $(\frac{n_0}{N})$, à proporção de animais marcados na segunda amostra, $(\frac{m}{n_1})$. Então,

$$\frac{n_0}{N} = \frac{m}{n_1} \Rightarrow \hat{N}_p = \frac{n_0 n_1}{m}.$$

O estimador de Petersen também pode ser obtido igualando-se a proporção de não marcados na população, antes da segunda amostra ser selecionada, $(\frac{N-n_0}{N})$, à proporção de não marcados na segunda amostra, $(\frac{n_1-m}{n_1})$. Com efeito,

$$\frac{N-n_0}{N}=\frac{n_1-m}{n_1} \quad \Rightarrow \quad n_1N-n_0n_1=n_1N-mN \quad \Rightarrow$$

$$\Rightarrow \quad \hat{N}_p=\frac{n_0n_1}{m}.$$

Este estimador tem esperança infinita pois m tem uma probabilidade positiva de assumir o valor zero e sua variância não existe.

Para resolver este problema, Bailey (1951) propôs o seguinte estimador para N:

$$\hat{N}_2=\frac{n_0(n_1+1)}{m+1}, \qquad \text{onde } m\sim B\left(n_1,\frac{n_0}{N}\right).$$

O seguinte lema nos dá a esperança de $\hat{N}_2$.

Lema 2 . A esperança de $\hat{N}_2$ é dada por

$$E(\hat{N}_2)=N-N\left(1-\frac{n_0}{N}\right)^{n_1+1}.$$

Prova.

$$E(\hat{N}_2)=E\left[\frac{n_0(n_1+1)}{m+1}\right]=\sum_{j=0}^{n_1}\frac{n_0(n_1+1)}{(j+1)}\binom{n_1}{j}\left(\frac{n_0}{N}\right)^j\left(1-\frac{n_0}{N}\right)^{n_1-j}$$

$$= N\sum_{j=0}^{n_1}\binom{n_1+1}{j+1}\left(\frac{n_0}{N}\right)^{j+1}\left(1-\frac{n_0}{N}\right)^{n_1+1-(j+1)}$$

$$= N\sum_{k=1}^{n_1+1}\binom{n_1+1}{k}\left(\frac{n_0}{N}\right)^{k}\left(1-\frac{n_0}{N}\right)^{n_1+1-k} = N\left[1-\left(1-\frac{n_0}{N}\right)^{n_1+1}\right]$$

$$= N - N\left(1-\frac{n_0}{N}\right)^{n_1+1}.$$

■

Teorema 3. Para todo K, $0 < K < N$, se $\frac{n_0 n_1}{N} \geq \ln(\frac{N}{K})$, então,

$$\left|E(\hat{N}_2) - N\right| < K.$$

Prova. Pelo Lema 2,

$$|E(\hat{N}_2) - N| = N\left(1-\frac{n_0}{N}\right)^{n_1+1} = N\left(1-\frac{n_0}{N}\right)^{n_1}\left(1-\frac{n_0}{N}\right)$$

$$= (N-n_0)\left(1-\frac{n_0}{N}\right)^{n_1} < N\left(1-\frac{n_0}{N}\right)^{n_1} < Ne^{\frac{-n_0 n_1}{N}} \leq Ne^{\ln\frac{K}{N}} = K.$$

■

A tabela 1 fornece valores de $ln(\frac{N}{K})$: o mínimo valor de $\frac{n_0 n_1}{N}$ para o qual $|E(\hat{N}_2) - N| < K$.

TABELA 1

$K \backslash N$	10^4	10^5	10^6	10^7	10^8	10^9
0.1	11.51	13.82	16.12	18.42	20.72	23.03
0.2	10.82	13.12	15.42	17.73	20.03	22.33
0.3	10.41	12.72	15.02	17.32	19.62	21.93
0.5	9.90	12.21	14.51	16.81	19.11	21.42
0.9	9.32	11.62	13.92	16.22	18.53	20.83
1	9.21	11.51	13.82	16.12	18.42	20.72
2	8.52	10.82	13.12	15.42	17.73	20.03

Valores de $\ln(\frac{N}{K})$: mínimo valor de $\frac{n_0 n_1}{N}$ para o qual $|E(\hat{N}_2) - N| < K$.

O seguinte lema nos dá um limite superior para a variância de $\hat{N}_2$.

Lema 4. A variância de $\hat{N}_2$ é tal que

$$V(\hat{N}_2) \leq \left(\frac{n_0 n_1}{2}\right)^2.$$

Prova. Como

$$0 \leq m \leq n_1 \quad \Rightarrow \quad 1 \leq m+1 \leq n_1 + 1 \quad \Rightarrow \quad \frac{1}{n_1+1} \leq \frac{1}{m+1} \leq 1,$$

então (ver Apêndice 1)

$$V\left(\frac{1}{m+1}\right) \quad \leq \quad \frac{\left(1 - \frac{1}{n_1+1}\right)^2}{4} = \left(\frac{n_1}{2(n_1+1)}\right)^2.$$

Logo,

$$V(\hat{N}_2) \leq \frac{{n_0}^2(n_1+1)^2{n_1}^2}{4(n_1+1)^2} = \left(\frac{n_0 n_1}{2}\right)^2.$$

■

O seguinte lema nos dá um estimador não viciado do inverso do tamanho da população.

Lema 5 . O inverso do estimador de Petersen,

$$\hat{I}_1 = \frac{1}{\hat{N}_p} = \frac{m}{n_0 n_1}, \qquad \text{onde } m \sim B\left(n_1, \frac{n_0}{N}\right)$$

é um estimador não viciado para $\frac{1}{N}$ e sua variância é

$$V(\hat{I}_1) = \frac{N - n_0}{n_0 n_1 N^2}.$$

Prova

$$E(\hat{I}_1) = E\left(\frac{m}{n_0 n_1}\right) = \frac{n_1}{n_0 n_1}\frac{n_0}{N} = \frac{1}{N} \quad \text{e}$$

$$V(\hat{I}_1) = V\left(\frac{m}{n_0 n_1}\right) = \frac{1}{n_0^2 n_1^2} V(m) = \frac{n_0 n_1}{{n_0}^2 {n_1}^2 N}\left(1 - \frac{n_0}{N}\right) = \frac{N - n_0}{n_0 n_1 N^2}.$$

■

b) SEM REPOSIÇÃO.

Sejam

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra;

n_1 : tamanho da segunda amostra, que é colhida sem reposição;

m : número de indivíduos marcados observados na segunda amostra.

A função densidade de probabilidade da variável aleatória m segue a lei hipergeométrica.

$$P(m \mid n_0, n_1, N) = \frac{\binom{n_0}{m}\binom{N-n_0}{n_1-m}}{\binom{N}{n_1}}, \qquad \max\{0, n_0 + n_1 - N\} \leq m \leq \min\{n_0, n_1\}.$$

A esperança e variância de m são, respectivamente:

$$E(m) = \frac{n_0 n_1}{N} \quad \text{e} \quad V(m) = \frac{n_0 n_1}{N}\left(1 - \frac{n_0}{N}\right)\left(\frac{N - n_1}{N - 1}\right).$$

Teorema 6. O E.M.V. de N é aproximadamente $\hat{N}_p = \frac{n_0 n_1}{m}$ se $\frac{n_0 n_1}{m}$ não for um número inteiro e é igual a $\hat{N}_p$ se $\frac{n_0 n_1}{m}$ for um número inteiro.

Prova . A função de verossimilhança é

$$L(N|n_0, n_1, m) = \frac{\binom{n_0}{m}\binom{N-n_0}{n_1-m}}{\binom{N}{n_1}} = \frac{n_0!(N-n_0)!n_1!(N-n_1)!}{m!(n_0-m)!(n_1-m)!(N-n_0-n_1+m)!N!},$$

onde $N \geq n_0 + n_1 - m$.

$$L(N \mid n_0, n_1, m) \propto K(N) = \frac{(N-n_0)!(N-n_1)!}{(N-n_0-(n_1-m))!N!}.$$

Para determinarmos o ponto de máximo de $K(N)$ consideremos o quociente

$$\frac{K(N+1)}{K(N)} = \frac{\big((N\text{-}n_0+1)!(N\text{-}n_1+1)!\big)/\big((N\text{-}n_0\text{-}n_1+m+1)!(N+1)!\big)}{\big((N\text{-}n_0)!(N\text{-}n_1)!\big)/\big((N\text{-}n_0\text{-}n_1+m)!N!\big)}$$

$$= \frac{(N-n_0+1)(N-n_1+1)}{(N-n_0-n_1+m+1)(N+1)}$$

Logo,

$$\frac{K(N+1)}{K(N)} > 1 \Longleftrightarrow$$

$K(N+1) > K(N)$

$$\Longleftrightarrow [(N+1)-n_0][(N+1)-n_1] > [(N+1)-n_0-n_1+m](N+1)$$

$$\Longleftrightarrow (N+1)^2 - n_1(N+1) - n_0(N+1) + n_0 n_1 >$$

$$> (N+1)^2 - n_0(N+1) - n_1(N+1) + m(N+1)$$

$$\Longleftrightarrow n_0 n_1 > mN + m \Longleftrightarrow \quad N < \frac{n_0 n_1}{m} - 1.$$

Ou seja,

$$\begin{array}{lllllll} K(N+1) & > & K(N) & \Longleftrightarrow & N & < & \frac{n_0 n_1}{m} - 1 \\ K(N+1) & < & K(N) & \Longleftrightarrow & N & > & \frac{n_0 n_1}{m} - 1 \\ K(N+1) & = & K(N) & \Longleftrightarrow & N & = & \frac{n_0 n_1}{m} - 1. \end{array}$$

Então,

1º) se $\frac{n_0 n_1}{m}$ não for inteiro, segue que $K(N)$ é crescente no intervalo $\left[n_0 + n_1 - m \ ; \ [\frac{n_0 n_1}{m}]\right] \cap N^*$ e $K(N)$ é decrescente no intervalo $\left[[\frac{n_0 n_1}{m}] \ ; \ \infty\right) \bigcap N^*$, onde $N^* = \{1, 2, \cdots\}$.

Logo a estimativa de Máxima Verossimilhança (E.M.V.) para N é

$$\left[\frac{n_0 n_1}{m}\right] \cong \hat{N}_p,$$

onde $[\ x\ ]$ representa o maior inteiro menor ou igual que x;

2º) se $\frac{n_0 n_1}{m}$ for inteiro segue que $K(N)$ é crescente no intervalo $\left[n_0 + n_1 - m, \ \frac{n_0 n_1}{m} - 1\right] \bigcap N^*$, $K\left(\frac{n_0 n_1}{m} - 1\right) = K\left(\frac{n_0 n_1}{m}\right)$, e $K(N)$ é decrescente no intervalo $\left[\frac{n_0 n_1}{m} \ , \ \infty\right) \bigcap N^*$. Logo há duas estimativas de Máxima Verossimilhança para N,

$$\frac{n_0 n_1}{m} = \hat{N}_p \qquad \text{e} \qquad \frac{n_0 n_1}{m} - 1 = \hat{N}_p - 1.$$

?

■

Como $\hat{N}_p$ tem esperança infinita e sua variância não existe, Chapman (1951) propôs o seguinte estimador

$$\hat{N}_3 = \frac{(n_0+1)(n_1+1)}{(m+1)} - 1, \qquad \text{onde } m \sim \text{Hip}(n_1,\ n_0,\ N).$$

Teorema 7. A esperança matemática de $\hat{N}_3$ é dada por

$$E(\hat{N}_3) = N - \frac{\binom{N-n_0}{n_1+1}(n_1+1)}{\binom{N}{n_1}}.$$

Prova.

$$E(\hat{N}_3) = E\left[\frac{(n_0+1)(n_1+1)}{m+1} - 1\right] = \sum_{j=0}^{n_1} \frac{(n_0+1)(n_1+1)\binom{n_0}{j}\binom{N-n_0}{n_1-j}}{(j+1)\binom{N}{n_1}} - 1$$

$$= (N+1)\sum_{j=0}^{n_1} \frac{\binom{n_0+1}{j+1}\binom{N+1-(n_0+1)}{n_1+1-(j+1)}}{\binom{N+1}{n_1+1}} - 1$$

$$= (N+1)\left[\sum_{k=0}^{n_1+1} \frac{\binom{n_0+1}{k}\binom{N+1-(n_0+1)}{n_1+1-k}}{\binom{N+1}{n_1+1}} - \frac{\binom{N-n_0}{n_1+1}}{\binom{N+1}{n_1+1}}\right] - 1$$

$$= (N+1)\left[1 - \frac{\binom{N-n_0}{n_1+1}}{\binom{N+1}{n_1+1}}\right] - 1 = N - \frac{(N+1)\binom{N-n_0}{n_1+1}}{\binom{N+1}{n_1+1}} = N - \frac{\binom{N-n_0}{n_1+1}(n_1+1)}{\binom{N}{n_1}}.$$

■

$n+1 > N-n$

n_0+n_1

Corolário 8.

i) Se $n_0 + n_1 \geq N$, então

$$E(\hat{N}_3) = N.$$

ii) Para todo K, $0 < K < N$, se $\frac{n_0 n_1}{N} \geq \ln(\frac{N}{K})$, então

$$|E(\hat{N}_3) - N| < K.$$

Prova de (i):

$$n_0 + n_1 \geq N \quad \Rightarrow \quad n_1 \geq N - n_0 \quad \Rightarrow \quad n_1 + 1 > n_1 \geq N - n_0 \Rightarrow \binom{N - n_0}{n_1 + 1} = 0$$

e pelo Teorema 7, (i) está provado.

Prova de (ii): suponhamos que $\frac{n_0 n_1}{N} \geq \ln(\frac{N}{K})$; então

a) Se $n_0 + n_1 \geq N$, por i), $\hat{N}_3$ é não viciado o que verifica (ii).

b) Se $n_0 + n_1 < N$, como

$$|E(\hat{N}_3) - N| = \frac{\binom{N-n_0}{n_1+1}(n_1 + 1)}{\binom{N}{n_1}} = \frac{(N - n_0)!n_1!(N - n_1)!(n_1 + 1)}{(n_1 + 1)!(N - n_0 - n_1 - 1)!N!}$$

$$= \frac{(N - n_0)(N - n_0 - 1)(N - n_0 - 2)\cdots(N - n_0 - n_1 + 1)[N - (n_0 + n_1)]}{N(N-1)(N-2)\cdots(N - n_1 + 2)(N - n_1 + 1)}$$

$$= \left(1 - \frac{n_0}{N}\right)\left(1 - \frac{n_0}{N-1}\right)\left(1 - \frac{n_0}{N-2}\right)\cdots\left(1 - \frac{n_0}{N - (n_1 - 1)}\right)[N - (n_0 + n_1)]$$

$$< [N - (n_0 + n_1)]\left(1 - \frac{n_0}{N}\right)^{n_1} \leq [N - (n_0 + n_1)](e^{\frac{-n_0}{N}})^{n_1}$$

$$= [N - (n_0 + n_1)]e^{\frac{-n_0 n_1}{N}} < N e^{\frac{-n_0 n_1}{N}},$$

e, para todo K, $0 < K < N$,

$$\frac{n_0 n_1}{N} \geq \ln\left(\frac{N}{K}\right) \quad \Rightarrow \quad \frac{-n_0 n_1}{N} \quad \leq \quad -\ln\left(\frac{N}{K}\right) = \ln\left(\frac{K}{N}\right),$$

segue que

$$|E(\hat{N}_3) - N| < Ne^{\frac{-n_0 n_1}{N}} \leq Ne^{\ln(\frac{K}{N})} = N\frac{K}{N} = K, \quad \text{o que prova (ii).}$$

■

Teorema 9. O inverso de $\hat{N}_p$,

$$\hat{I}_2 = \frac{m}{n_0 n_1}, \quad \text{onde} \quad m \sim H(n_1, n_0, N)$$

é um estimador não viciado para $\frac{1}{N}$, com variância

$$Var(\hat{I}_2) = \frac{(N - n_0)(N - n_1)}{n_0 n_1 N^2 (N-1)}.$$

Prova.

$$E(\hat{I}_2) = E\left(\frac{m}{n_0 n_1}\right) = \frac{1}{n_0 n_1} E(m) = \frac{n_1}{n_0 n_1}\frac{n_0}{N} = \frac{1}{N}$$

e

$$V(\hat{I}_2) = V\left(\frac{m}{n_0 n_1}\right) = \frac{1}{n_0^2 n_1^2} V(m) = \frac{n_1 n_0}{n_0^2 n_1^2 N}\left(1 - \frac{n_0}{N}\right)\left(\frac{N - n_1}{N - 1}\right)$$

$$= \frac{(N - n_0)(N - n_1)}{n_0 n_1 N^2 (N-1)}.$$

■

INTERVALO DE CONFIANÇA PARA N.

Nesta seção determinaremos um intervalo de confiança para N no caso de amostragem direta com reposição.

Como vimos no início da seção 1.1. $m \sim B(n_1, n_0/N)$. Pelo Teorema Central do Limite (T.C.L.)

$$\sqrt{n_1}(m/n_1 - n_0/N) \xrightarrow[n_1 \to \infty]{\mathcal{D}} N(0, (n_0/N)(1 - n_0/N))$$

Seja $g(x) = \frac{1}{x}$, $x > 0$. Logo

$$\begin{aligned} g\left(\frac{m}{n_1}\right) &= \frac{n_1}{m}, \\ g\left(\frac{n_0}{N}\right) &= \frac{N}{n_o}, \\ g'\left(\frac{n_0}{N}\right) &= -\frac{N^2}{n_0^2} \quad \text{e} \\ \left[g'\left(\frac{n_0}{N}\right)\right]^2 &= \frac{N^4}{n_0^4}. \end{aligned}$$

Pelo método Delta (Bishop (1975))

$$\sqrt{n_1}(n_1/m - N/n_0) \xrightarrow[n_1 \to \infty]{\mathcal{D}} N\left(0, [g'(\frac{n_0}{N})]^2 (n_0/N)(1 - n_0/N)\right),$$

ou seja,

$$\sqrt{n_1}(n_1/m - N/n_0) \xrightarrow[n_1 \to \infty]{\mathcal{D}} N\left(0, (1 - n_0/N)/(n_0/N)^3\right).$$

Este resultado nos permite construir para n_1 suficientemente grande um intervalo de confiança para N:

$$\left(\frac{n_0 n_1}{m} - z_\gamma \frac{n_0}{\sqrt{n_1}} \sqrt{\frac{1 - \frac{n_0}{N}}{\left(\frac{n_0}{N}\right)^3}} \; ; \; \frac{n_0 n_1}{m} + z_\gamma \frac{n_0}{\sqrt{n_1}} \sqrt{\frac{1 - \frac{n_0}{N}}{\left(\frac{n_0}{N}\right)^3}} \right),$$

onde z_γ é tal que $P(-z_\gamma \leq Z \leq z_\gamma) = \gamma$, com $Z \sim N(0,1)$.

Substituindo n_0/N por m/n_1 no intervalo de confiança acima, para n_1 suficientemente grande, o intervalo de confiança "otimista" para N é

$$\left(\frac{n_0 n_1}{m} - z_\gamma \frac{n_0}{\sqrt{n_1}} \sqrt{\frac{1 - \frac{m}{n_1}}{\left(\frac{m}{n_1}\right)^3}} \; ; \; \frac{n_0 n_1}{m} + z_\gamma \frac{n_0}{\sqrt{n_1}} \sqrt{\frac{1 - \frac{m}{n_1}}{\left(\frac{m}{n_1}\right)^3}} \right)$$

Exemplo 1. Para $n_0 = 100$, $n_1 = 250$ e $m = 25$, então $\frac{n_0 n_1}{m} = \frac{25.000}{25} = 1.000$ e os intervalos de 95 % e 99 % de confiança para N são, respectivamente,

$$\left(1.000 - 1,96\frac{100}{\sqrt{250}}\sqrt{\frac{1-0,1}{\left(\frac{25}{250}\right)^3}}\ ;\ 1.000 + 1,96\frac{100}{\sqrt{250}}\sqrt{\frac{1-0,1}{\left(\frac{25}{250}\right)^3}}\right)$$

$$= (629\ ;\ 1.372)$$

e

$$\left(1.000 - 2,58\frac{100 \times 30}{15,81139}\ ;\ 1.000 + 2,58\frac{100 \times 30}{15,81139}\right)$$

$$= (511\ ;\ 1.490)$$

AMOSTRAGEM INVERSA SIMPLES.

Até agora estudamos estimadores do tamanho da população pelo processo de captura-recaptura considerando fixos os tamanhos das amostras. Um outro método para obter estimadores do tamanho da população, pelo processo de captura-recaptura, é o método da amostragem inversa. No caso simples, tal método consiste em capturarmos na primeira amostragem um número n_0 (n_0 fixado) de animais da população de tamanho N, os quais são marcados e devolvidos à população. Fixado um número inteiro positivo, a segunda amostragem consiste na captura de animais da população, um de cada vez, até ocorrer esse número fixado de marcados (ou não marcados) Os animais não marcados capturados na segunda amostragem não recebem marca. No método da amostragem inversa, o tamanho da segunda amostra é uma variável aleatória.

AMOSTRAGEM INVERSA SIMPLES FIXANDO O NÚMERO DE MARCADOS.

Suponhamos que selecionemos e marquemos n_0 indivíduos de uma população de tamanho N. Os indivíduos marcados são devolvidos à população . Fixado um número inteiro $m \geq 1$, selecionamos numa segunda amostra, um indivíduo de cada vez, até que apareçam esses m indivíduos marcados. Tal procedimento pode ser realizado de duas maneiras: com ou sem reposição. Nas duas secções seguintes faremos um estudo desse método.

a) COM REPOSIÇÃO.

Sejam

n_0: número de indivíduos capturados sem reposição, marcados e soltos na primeira amostra, e

n_1: número de indivíduos capturados, com reposição, até que ocorra o m-ésimo indivíduo marcado, onde m é um número inteiro positivo fixado.

A variável aleatória n_1 tem distribuição de probabilidade Binomial Negativa com parâmetros m e $p = \frac{n_0}{N}$.

A distribuição de probabilidade de n_1 é dada por

$$P(n_1|n_0, m, N) = \binom{n_1 - 1}{m - 1}\left(\frac{n_0}{N}\right)^m \left(1 - \frac{n_0}{N}\right)^{n_1 - m}, \qquad n_1 \geq m$$

A média e a variância de n_1 são, respectivamente,

$$E(n_1) = \frac{m}{p} = \frac{mN}{n_0}$$

e

$$V(n_1) = \frac{m(1-p)}{p^2} = \frac{m}{n_0^2}\left(\frac{N - n_0}{N}\right)N^2 = \frac{mN(N - n_0)}{n_0^2}.$$

Se $n_0 <<< N$ (n_0 muito menor que N), então podemos aproximar a variância de n_1 por

$$Var(n_1) = \frac{mN^2}{{n_0}^2}\left(1 - \frac{n_0}{N}\right) \cong \frac{mN^2}{n_0^2}.$$

Teorema 10 . A estimativa de Máxima Verossimilhança para N é igual a $\hat{N}_p$ se $\frac{n_o n_1}{m}$ for um número inteiro e aproximadamente igual a $\hat{N}_p$ se $\frac{n_o n_1}{m}$ não for um número inteiro

?

Prova. A função de verossimilhança de N é

$$L(N \mid n_0, n_1, m) \propto K(N) = \left(\frac{n_0}{N}\right)^m \left(1 - \frac{n_0}{N}\right)^{n_1 - m} \quad , \quad N > n_0.$$

Então, de modo análogo ao caso de amostragem direta com reposição (ver Teorema 1), segue que a estimativa de Máxima Verossimilhança para N é igual ou aproximadamente igual à $\hat{N}_p$. ∎

Notamos que $\hat{N}_p$ é não viciado. Com efeito,

$$E(\hat{N}_p) = E\left(\frac{n_0 n_1}{m}\right) = \frac{n_0}{m} E(n_1) = \frac{n_0}{m}\frac{mN}{n_0} = N.$$

A variancia de $\hat{N}_p$ é

$$Var(\hat{N}_p) = \frac{N(N - n_0)}{m}.$$

Com efeito,

$$Var(\hat{N}_p) = Var\left(\frac{n_0 n_1}{m}\right) = \frac{n_0^2}{m^2} Var(n_1) = \frac{n_0^2}{m^2}\frac{mN(N - n_0)}{n_0^2} = \frac{N(N - n_0)}{m}.$$

Teorema 11. Um estimador não viciado da variância de $\hat{N}_p$ é

$$\hat{\sigma}^2_{\hat{N}_p} = \frac{n_0^2 n_1 (n_1 - m)}{m^2 (m + 1)}.$$

Prova.

$$E(\hat{\sigma}^2_{\hat{N}_p}) = E\left[\frac{n_0^2 n_1(n_1 - m)}{m^2(m+1)}\right] = \frac{n_0^2}{m^2(m+1)}[E(n_1^2) - mE(n_1)]$$

$$= \frac{n_0^2}{m^2(m+1)}\left[\frac{m(N-n_0)N}{n_0^2} + \frac{m^2N^2}{n_0^2} - \frac{m^2N}{n_0}\right]$$

$$= \frac{1}{m(m+1)}\left[N(N-n_0) + mN^2 - mn_0N\right]$$

$$= \frac{1}{m(m+1)}[N(N-n_0) + mN(N-n_0)] = \frac{N(N-n_0)(m+1)}{m(m+1)}$$

$$= \frac{N(N-n_0)}{m} = Var(\hat{N}_p).$$

∎

O seguinte teorema nos da a distribuição assintótica do estimador $\hat{N}_p$.

Teorema 12.

$$\frac{\hat{N}_p - N}{\sqrt{\frac{N(N-n_0)}{m}}} \xrightarrow[m \to \infty]{\mathcal{D}} Z, \quad \text{onde } Z \sim N(0,1).$$

Prova.

$$\frac{\hat{N}_p - N}{\sqrt{\frac{N(N-n_0)}{m}}} = \frac{\frac{n_0 n_1}{m} - N}{\sqrt{\frac{N(N-n_0)}{m}}} = \frac{n_1 - \frac{mN}{n_0}}{\sqrt{\frac{mN(N-n_0)}{n_0^2}}} = \frac{n_1 - E(n_1)}{\sqrt{Var(n_1)}}.$$

Notando que $n_1 = \sum_{j=1}^{m} X_j$, onde $X_1, X_2, \cdots$ é uma seqüencia de variáveis aleatórias independentes e identicamente distribuidas com distribuição geométrica de parâmetro $p = \frac{n_0}{N}$, segue, pelo Teorema Central do Limite, que

$$\frac{n_1 - E(n_1)}{\sqrt{Var(n_1)}} = \frac{\sum_{j=1}^m X_j - E(\sum_{j=1}^m X_j)}{\sqrt{Var(\sum_{j=1}^m X_j)}} \xrightarrow[m \to \infty]{\mathcal{D}} Z.$$

■

Corolário 13 .

$$\frac{\hat{N}_p - N}{\sqrt{\frac{\hat{N}_p(\hat{N}_p - n_0)}{m}}} \xrightarrow[m \to \infty]{\mathcal{D}} Z, \quad \text{onde } Z \sim N(0,1).$$

Prova. Pela Lei Fraca Dos Grandes Números,

$$\frac{n_1}{m} = \frac{\sum_{j=1}^m X_j}{m} \xrightarrow[m \to \infty]{\mathcal{P}} E(X_1) = \frac{N}{n_0}.$$

Então,

$$\hat{N}_p = \frac{n_0 n_1}{m} \xrightarrow[m \to \infty]{\mathcal{P}} \frac{n_0 N}{n_0} = N,$$

o que implica

$$\sqrt{\frac{N(N-n_0)}{\hat{N}_p(\hat{N}_p - n_0)}} \xrightarrow[m \to \infty]{\mathcal{P}} \sqrt{\frac{N(N-n_0)}{N(N-n_0)}} = 1.$$

Logo, pelos teoremas anterior e de Slutsky,

$$\frac{\hat{N}_p - N}{\sqrt{\frac{\hat{N}_p(\hat{N}_p - n_0)}{m}}} = \frac{\hat{N}_p - N}{\sqrt{\frac{N(N-n_0)}{m}}} \sqrt{\frac{N(N-n_0)}{\hat{N}_p(\hat{N}_p - n_0)}} \xrightarrow[m \to \infty]{\mathcal{D}} Z.$$

■

Corolário 14. Seja

$$\hat{\sigma}^2_{\hat{N}_p} = \frac{n_0^2 n_1(n_1 - m)}{m^2(m+1)} = \frac{\hat{N}_p(\hat{N}_p - n_0)}{m+1}; \text{ então}$$

$$\frac{\hat{N}_p - N}{\sqrt{\hat{\sigma}^2_{\hat{N}_p}}} \xrightarrow[m \to \infty]{\mathcal{D}} Z, \text{ onde } Z \sim N(0,1).$$

Observação: Notemos que $\hat{\sigma}^2_{\hat{N}_p}$, pelo Teorema 11, é um estimador não viciado da variância de $\hat{N}_p$.

Prova. Como vimos anteriormente,

$$\frac{n_1}{m} \xrightarrow[m \to \infty]{\mathcal{P}} \frac{N}{n_0}.$$

Logo,

$$\sqrt{\frac{N(N-n_0)m(m+1)}{n_0^2 n_1(n_1 - m)}} = \sqrt{\frac{N(N-n_0)(1+\frac{1}{m})}{n_0^2 \frac{n_1}{m}(\frac{n_1}{m} - 1)}} \xrightarrow[m \to \infty]{\mathcal{P}} 1$$

e, portanto, pelo Teorema 12 e de Slutsky, temos:

$$\frac{\hat{N}_p - N}{\sqrt{\hat{\sigma}^2_{\hat{N}_p}}} = \frac{\hat{N}_p - N}{\sqrt{\frac{n_0^2 n_1(n_1-m)}{m^2(m+1)}}}$$

$$= \frac{(\hat{N}_p - N)\sqrt{\frac{N(N-n_0)}{m}}}{\sqrt{(n_0^2 n_1(n_1 - m))/(m^2(m+1))}\sqrt{\frac{N(N-n_0)}{m}}}$$

$$= \frac{\hat{N}_p - N}{\sqrt{\frac{N(N-n_0)}{m}}} \sqrt{\frac{N(N-n_0)m(m+1)}{n_0^2 n_1(n_1 - m)}} \xrightarrow[m \to \infty]{\mathcal{D}} Z.$$

■

Os corolários acima nos permitem construir intervalos de confiança para N.

1.5. INTERVALOS DE CONFIANÇA PARA N.

Do Corolário 13 temos, para m suficientemente grande, que

$$\left(\hat{N}_p - z_\gamma \sqrt{\frac{\hat{N}_p(\hat{N}p - n_0)}{m}} \; ; \; \hat{N}_p + z_\gamma \sqrt{\frac{\hat{N}_p(\hat{N}_p - n_0)}{m}} \right)$$

é um intervalo de confiança para N com coeficiente de confiança γ, onde z_γ é tal que $P(-z_\gamma \leq Z \leq z_\gamma) = \gamma$.

Do Corolário 14 temos, para m suficientemente grande, que

$$\left(\hat{N}_p - z_\gamma \sqrt{\frac{\hat{N}_p(\hat{N}_p - n_0)}{m+1}} \; ; \; \hat{N}_p + z_\gamma \sqrt{\frac{\hat{N}_p(\hat{N}_p - n_0)}{m+1}} \right)$$

é um intervalo de confiança para N com coeficiente de confiança γ.

Exemplo 2. Para $n_0 = 1.000$, $n_1 = 1.000$ e $m = 500$ o intervalo de 95 % de confiança para N obtido com o resultado do Corolário 13 é

$$\left(2.000 - 1,96\sqrt{\frac{2.000(2.000 - 1.000)}{500}} \; ; \; 2.000 + 1,96\sqrt{\frac{2.000(2.000 - 1.000)}{500}} \right)$$

$$= (2.000 - 123,96 \; ; \; 2.000 + 123,96) = (1.877 \; ; \; 2.124).$$

O intervalo de 95 % de confiança para N obtido com o resultado do Corolário 14 é

$$\left(2.000 - 1,96\sqrt{\frac{2.000(2.000 - 1.000)}{501}} \; ; \; 2.000 + 1,96\sqrt{\frac{2.000(2.000 - 1.000)}{501}} \right)$$

$$= (1.877 \; ; \; 2.124).$$

Com o coeficiente de confiança de 99 %, os intervalos de confiança para N obtidos com os resultados dos Corolários 13 e 14 são, respectivamente,

$$\left(2.000 - 2,58\sqrt{\frac{2.000(2.000-1.000)}{500}} \; ; \; 2.000 + 2,58\sqrt{\frac{2.000(2.000-1.000)}{500}} \right)$$

$$= (2.000 - 163,174 \; ; \; 2.000 + 163,174) = (1.837 \; ; \; 2.164)$$

e

$$\left(2.000 - 2,58\sqrt{\frac{2.000(2.000-1.000)}{501}} \; ; \; 2.000 + 2,58\sqrt{\frac{2.000(2.000-1.000)}{501}} \right)$$

$$= (2.000 - 163,010 \; ; \; 2.000 + 163,010) = (1.837 \; ; \; 2.164).$$

Por outro lado, para $\frac{n_0}{N}$ pequeno a variável aleatória $\frac{2n_0n_1}{N}$ tem, aproximadamente, distribuição de probabilidade χ^2 com $2m$ graus de liberdade (ver Apêndice 3). Este fato nos permite construir um intervalo de confiança para N .

Se $\chi^2_{2m}(1-\gamma)$ é tal que

$$P(\chi^2_{2m} \leq \chi^2_{2m}(1-\gamma)) = 1-\gamma,$$

então, um intervalo de confiança para N com coeficiente de confiança γ é dado por

$$\left(\frac{2n_0n_1}{\chi^2_{2m}(\frac{1+\gamma}{2})} \; ; \; \frac{2n_0n_1}{\chi^2_{2m}(\frac{1-\gamma}{2})} \right).$$

Exemplo 3. Se $n_0 = 100$, $n_1 = 250$, $m = 25$ e se $\frac{100}{N}$ é pequeno, então um intervalo de confiança para N com um 95 % de confiança é

$$\left(\frac{2\ 100\ 250}{\chi^2_{50}(0,975)} \; ; \; \frac{2\ 100\ 250}{\chi^2_{50}(0,025)} \right) = (701; 1.546),$$

e com um 99 % de confiança é

$$\left(\frac{2\ 100\ 250}{\chi^2_{50}(0,995)} \ ; \ \frac{2\ 100\ 250}{\chi^2_{50}(0,005)} \right) = (630 \ ; \ 1.787).$$

b) SEM REPOSICAO.

Sejam

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra e

n_1: número de indivíduos capturados, sem reposição, até que ocorra o m-ésimo animal marcado, onde m é um número inteiro, $1 \leq m \leq n_0$, fixado.

A variável aleatória n_1 tem distribuição de probabilidade Hipergeométrica Negativa com parâmetros m, n_0 e N.

Notando que o evento "n_1 elementos são selecionados da população, sem reposição, até que ocorra o m-ésimo elemento marcado " é o evento "$(m-1)$ elementos marcados são selecionados da população, sem reposição, nas $(n_1 - 1)$ primeiras seleções, e na n_1-ésima seleção é selecionado um indivíduo marcado ", segue que a distribuição de probabilidades de n_1 é dada por

$$P(n_1 | \ m, n_0, N) = \frac{\binom{n_0}{m-1}\binom{N-n_0}{n_1-m}}{\binom{N}{n_1-1}} \frac{(n_0 - m + 1)}{(N - n_1 + 1)}$$

$$= \frac{n_0(n_0-1)!\binom{N-n_0}{n_1-m}(n_1-1)!(N-n_1+1)!(n_0-m+1)}{(m-1)!(n_0-m+1)!N(N-1)!(N-n_1+1)}.$$

$$= \frac{n_0}{N} \frac{\binom{n_0-1}{m-1}\binom{N-n_0}{n_1-m}}{\binom{N-1}{n_1-1}}, \quad \text{onde} \quad m \leq n_1 \leq (N - n_0) + m$$

Observando a última expressão da distribuição acima, notamos que ela é igual à probabilidade do evento "um elemento marcado é selecionado sem reposição, na primeira seleção, e $(m-1)$ elementos marcados são selecionados quando uma amostra de tamanho (n_1-1) é selecionada, sem reposição, da população, agora contendo (n_0-1) elementos marcados e $(N-n_0)$ elementos não marcados".

Lema 15: Se n_1 tem distribuição Hipergeométrica Negativa com parâmetros m, n_0 e N, então

$$i) \qquad E(n_1) = \frac{m(N+1)}{(n_0+1)}$$

e

$$ii) \qquad V(n_1) = \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)}.$$

Prova. Notemos que

$$\sum_{k=m}^{(N-n_0)+m} \frac{\binom{N-n_0}{k-m}}{\binom{N-1}{k-1}} = \frac{N}{n_0\binom{n_0-1}{m-1}}. \tag{1.1.}$$

Com efeito,

$$\sum_{k=m}^{(N-n_0)+m} \frac{\binom{N-n_0}{k-m}}{\binom{N-1}{k-1}} = \frac{N}{n_0\binom{n_0-1}{m-1}} \sum_{k=m}^{(N-n_0)+m} \left(\frac{n_0}{N}\right) \frac{\binom{n_0-1}{m-1}\binom{N-n_0}{k-m}}{\binom{N-1}{k-1}}$$

$$= \frac{N}{n_0\binom{n_0-1}{m-1}} \sum_{k=m}^{(N-n_0)+m} P(n_1 = k|m, n_0, N) = \frac{N}{n_0\binom{n_0-1}{m-1}}.$$

Logo,

$$E(n_1) = \sum_{k=m}^{(N-n_0)+m} k\frac{n_0}{N} \frac{\binom{n_0-1}{m-1}\binom{N-n_0}{k-m}}{\binom{N-1}{k-1}}$$

$$= n_0\binom{n_0-1}{m-1} \sum_{k=m}^{(N-n_0)+m} \frac{\binom{N-n_0}{k-m}}{\binom{N}{k}}. \tag{1.2.}$$

A soma do lado direito de (1.2) é obtida a partir de (1.1) escrevendo-se $(N+1)$, (n_0+1), $(k+1)$ e $(m+1)$ nos lugares de N, n_0, k e m, respectivamente. Então,

$$E(n_1) = \frac{n_0\binom{n_0-1}{m-1}(N+1)}{(n_0+1)\binom{n_0}{m}} = \frac{n_0(n_0-1)!(N+1)m!(n_0-m)!}{(m-1)!(n_0-m)!(n_0+1)n_0!} = \frac{m(N+1)}{n_0+1},$$

o que prova (i).

Por outro lado,

$$\begin{aligned} E[n_1(n_1+1)] &= \sum_{k=m}^{(N-n_0)+m} k(k+1)\frac{n_0}{N}\frac{\binom{n_0-1}{m-1}\binom{N-n_0}{k-m}}{\binom{N-1}{k-1}} \\ &= n_0(N+1)\binom{n_0-1}{m-1}\sum_{k=m}^{(N-n_0)+m}\frac{\binom{N-n_0}{k-m}}{\binom{N+1}{k+1}}. \qquad (1.3.) \end{aligned}$$

A soma do lado direito de (1.3.) é obtida a partir de (1.1.) escrevendo-se $(N+2)$, (n_0+2), $(k+2)$ e $(m+2)$ nos lugares de N, n_0, k e m, respectivamente. Então,

$$\begin{aligned} E[n_1(n_1+1)] &= \frac{n_0(N+1)\binom{n_0-1}{m-1}(N+2)}{(n_0+2)\binom{n_0+1}{m+1}} \\ &= \frac{n_0(N+1)(n_0\text{-}1)!(N+2)(m+1)!(n_0\text{-}m)!}{(n_0+2)(m\text{-}1)!(n_0\text{-}m)!(n_0+1)!} \\ &= \frac{m(m+1)(N+1)(N+2)}{(n_0+1)(n_0+2)}. \qquad (1.4.) \end{aligned}$$

Portanto, de (i) e (1.4.), segue que

$$V(n_1) = E[n_1(n_1+1)] - E(n_1) - E^2(n_1)$$

$$= \frac{m(m+1)(N+1)(N+2)}{(n_0+1)(n_0+2)} - \frac{m(N+1)}{(n_0+1)} - \frac{m^2(N+1)^2}{(n_0+1)^2}$$

$$= m(N+1)\left[\frac{(m+1)(N+2)(n_0+1)-(n_0+1)(n_0+2)-m(N+1)(n_0+2)}{(n_0+1)^2(n_0+2)}\right]$$

$$= \left(m(N+1)/(n_0+1)^2(n_0+2)\right)[(mN+2m+N+2)(n_0+1)-({n_0}^2+3n_0+2)-$$

$$-m(n_0N+2N+n_0+2)]$$

$$= \frac{m(N+1)}{(n_0+1)^2(n_0+2)}[mn_0+n_0N-n_0-mN+N-{n_0}^2]$$

$$= \frac{m(N+1)}{(n_0+1)^2(n_0+2)}[n_0(N-n_0)-m(N-n_0)+(N-n_0)]$$

$$= \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)}$$

$$= \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)},$$

o que prova (ii). ■

Notemos que $Var(n_1) \xrightarrow[N\to\infty]{} \infty$.

Como $1 \leq m \leq n_0$, de i) segue que

$$\frac{N+1}{n_0+1} \leq E(n_1) \leq \frac{n_0(N+1)}{n_0+1}.$$

O seguinte lema estabelece limites inferior e superior para a variância de n_1.

Lema 16. Se n_1 tem distribuição Hipergeométrica Negativa com parâmetros m, n_0 e N, então,

$$\left(\frac{m}{n_0+2}\right)\left(\frac{N-n_0}{n_0+1}\right)^2 < Var(n_1) \leq \min\left\{m\left(\frac{N}{n_0}\right)^2, \left(\frac{N-n_0}{2}\right)^2\right\}.$$

Prova .

$$Var(n_1) = \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)} \geq \frac{m(N+1)(N-n_0)}{(n_0+1)^2(n_0+2)}$$

$$\geq \frac{m(N-n_0)(N-n_0)}{(n_0+1)^2(n_0+2)} = \frac{m}{(n_0+2)}\left(\frac{N-n_0}{n_0+1}\right)^2.$$

Também temos que, como $(n_0-m+1) < (n_0+2)$, então,

$$Var(n_1) = \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)} < \frac{m(N+1)(N-n_0)}{(n_0+1)^2}$$

$$< \frac{m(N+1)(N-1)}{{n_0}^2} = \frac{m(N^2-1)}{{n_0}^2} < m\left(\frac{N}{n_0}\right)^2.$$

Por outro lado, como $m \leq n_1 \leq N-n_0+m$, então usando o resultado do Apêndice 1,

$$V(n_1) \leq \frac{(N-n_0)^2}{4}.$$

Logo,

$$Var(n_1) \leq \min\left\{m\left(\frac{N}{n_0}\right)^2, \left(\frac{N-n_0}{2}\right)^2\right\},$$

o que prova o teorema. ∎

O E.M.V. de N é dado pelo seguinte teorema.

Teorema 17. O E.M.V. de N é aproximadamente $\hat{N}_p$

Prova. A função de verossimilhança de N é

$$L(N|n_0, n_1, m) = \frac{\binom{n_0}{m-1}\binom{N-n_0}{n_1-m}(n_0 - m + 1)}{\binom{N}{n_1-1}(N - n_1 + 1)} \propto K(N) = \frac{\binom{N-n_0}{n_1-m}}{\binom{N}{n_1-1}(N - n_1 + 1)},$$

onde $N \geq n_0 + n_1 - m$.

Para determinarmos o ponto de máximo de $K(N)$ consideremos o quociente

$$\frac{K(N+1)}{K(N)} = \frac{\binom{N+1-n_0}{n_1-m}/\binom{N+1}{n_1-1}(N - n_1 + 2)}{\binom{N-n_0}{n_1-m}/\binom{N}{n_1-1}(N - n_1 + 1)}$$

$$= \frac{(N - n_0 + 1)!N!(N - n_1 + 1)(N - n_1 + 2)!(N - n_0 - n_1 + m)!}{(N + 1)!(N - n_1 + 2)(N - n_0)!(N - n_0 - n_1 + m + 1)!(N - n_1 + 1)!}$$

$$= \frac{(N - n_0 + 1)(N - n_1 + 1)}{(N + 1)(N - n_0 - n_1 + m + 1)}.$$

Portanto, seguindo os mesmos passos dados na prova do Teorema 6 temos:

1º) se $\frac{n_0 n_1}{m}$ não for um número inteiro, então a E.M.V. de N é

$$\left[\frac{n_0 n_1}{m}\right];$$

2º) se $\frac{n_0 n_1}{m}$ for um número inteiro, então há duas E.M.V.:

$$\frac{n_0 n_1}{m} = \hat{N}_p \quad e \quad \frac{n_0 n_1}{m} - 1 = \hat{N}_p - 1.$$

Portanto, o E.M.V. de N é aproximadamente $\hat{N}_p = \frac{n_0 n_1}{m}$. ∎

Pelo Lema 15 a esperança e a variância de $\hat{N}_p$ são, respectivamente,

$$E(\hat{N}_p) = E(\frac{n_0 n_1}{m}) = \frac{n_0}{m}E(n_1) = \frac{n_0}{m}m\frac{(N+1)}{(n_0+1)} = \frac{n_0}{(n_0+1)}(N+1)$$

e

$$Var(\hat{N}_p) = \frac{n_0^2}{m^2}Var(n_1) = \frac{n_0^2}{m^2}\frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)}$$

$$= \frac{n_0^2(N+1)(N-n_0)(n_0-m+1)}{m(n_0+1)^2(n_0+2)}.$$

Chapman (1952) propôs o seguinte estimador para N:

$$\hat{N}_4 = \frac{n_1(n_0+1)}{m} - 1.$$

Este estimador é não viciado como mostra o seguinte teorema:

Teorema 18. O estimador

$$\hat{N}_4 = \frac{n_1(n_0+1)}{m} - 1$$

é um estimador não viciado para N, com variância

$$Var(\hat{N}_4) = \frac{(N+1)(N-n_0)(n_0-m+1)}{m(n_0+2)}.$$

Prova.

$$E(\hat{N}_4) = E\left[\frac{n_1(n_0+1)}{m} - 1\right] = \frac{(n_0+1)}{m}E(n_1) - 1.$$

Pelo Lema 15,

$$E(\hat{N}_4) = \frac{(n_0+1)}{m}\frac{m(N+1)}{(n_0+1)} - 1 = N$$

e

$$V(\hat{N}_4) = V\left[\frac{n_1(n_0+1)}{m} - 1\right] = \frac{(n_0+1)^2}{m^2} V(n_1)$$

$$= \frac{(n_0+1)^2}{m^2} \frac{m(N+1)(N-n_0)(n_0-m+1)}{(n_0+1)^2(n_0+2)}$$

$$= \frac{(N+1)(N-n_0)(n_0-m+1)}{m(n_0+2)}.$$

■

Observemos que, como $n_0 - m + 1 < n_0 + 2$, então,

$$Var(\hat{N}_4) = \frac{(N+1)(N-n_0)(n_0-m+1)}{m(n_0+2)} < \frac{(N+1)(N-1)}{m} = \frac{N^2-1}{m} < \frac{N^2}{m}. \quad (1.5.)$$

Usando este último resultado e a desigualdade de Tchebychev, temos:

$$P\left(\left|\frac{\hat{N}_4}{N} - 1\right| \leq \lambda\right) \geq 1 - \frac{V(\hat{N}_4)}{N^2\lambda^2} > 1 - \frac{1}{m\lambda^2} = \frac{m\lambda^2 - 1}{m\lambda^2}.$$

A Tabela 2 fornece valores de $1 - \frac{1}{m\lambda^2}$: minimo valor da $P\left(\left|\frac{\hat{N}_4}{N} - 1\right| \leq \lambda\right)$ para alguns valores de m e λ.

TABELA 2

m \ λ	0,9	1	2	3
2	0,38	0,50	0,87	0,94
4	0,69	0,75	0,94	0,97
6	0,79	0,83	0,96	0,98
8	0,85	0,87	0,97	0,99
10	0,88	0,90	0,97	0,99
20	0,94	0,95	0,99	0,99

A relação $V(\hat{N}_4) < \frac{N^2}{m}$ pode ser usada na escolha do m. De fato, podemos escolher m de modo que $\frac{\sqrt{Var(\hat{N}_4)}}{N}\left(< \frac{1}{\sqrt{m}}\right)$ tenha um nível desejado.

Com relação aos estimadores $\hat{N}_4$ e $\hat{N}_p$, observamos que o primeiro é não viciado, enquanto que $\hat{N}_p$ não é. No entanto, a $Var(\hat{N}_4) > Var(\hat{N}_p)$. Com efeito,

$$Var(\hat{N}_4) = \frac{(n_0+1)^2}{m^2} Var(n_1) > \frac{{n_0}^2}{m^2} Var(n_1) = Var(\hat{N}_p).$$

Se o pesquisador não tem informação alguma com respeito a N, então, por uma escolha inadequada de n_0 e m, $E(n_1)$ e $Var(n_1)$ podem ser muito grandes. Um procedimento alternativo que pode melhorar esta situação é o processo de amostragem inversa, fixando o número de não marcados. Na próxima seção descreveremos este procedimento.

AMOSTRAGEM INVERSA SIMPLES FIXANDO O NÚMERO DE NÃO MARCADOS

Nesta seção estudaremos o método de amostragem inversa, no qual é fixado o número de indivíduos não marcados ($\delta \geq 1$). Ou seja, depois de capturar, marcar e devolver à população n_0 indivíduos, uma segunda amostra é retirada até encontrar δ indivíduos não marcados. Neste caso, n_1: número de tentativas até a captura do δ-ésimo indivíduo não marcado é aleatório. O procedimento pode ser realizado com ou sem reposição . Na segunda amostra, os animais capturados sem marca não são marcados.

a) COM REPOSIÇÃO.

Sejam

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra, e

n_1: número de indivíduos capturados com reposição na segunda amostra até que ocorra o δ-ésimo indivíduo não marcado, onde δ é um número inteiro positivo fixado.

A variável aleatória n_1 tem distribuição de probabilidade Binomial Negativa com parâmetros δ e $p = 1 - \frac{n_0}{N}$.

A distribuição de probabilidade de n_1 é dada por

$$P(n_1|n_0,\delta,N) = \binom{n_1-1}{\delta-1}\left(1-\frac{n_0}{N}\right)^{\delta}\left(\frac{n_0}{N}\right)^{n_1-\delta}, \quad n_1 \geq \delta$$

A média e a variância de n_1 são, respectivamente,

$$E(n_1) = \frac{\delta}{p} = \frac{\delta}{1-\frac{n_0}{N}} = \delta\left(1-\frac{n_0}{N}\right)^{-1}$$

e

$$Var(n_1) = \frac{\delta(1-p)}{p^2} = \frac{\frac{\delta n_0}{N}}{(1-\frac{n_0}{N})^2} = \frac{n_0\delta N}{(N-n_0)^2} = \frac{n_0\delta}{N}\left(1-\frac{n_0}{N}\right)^{-2}.$$

Observemos que se $n_0 < \frac{N}{2}$ e $\delta = m$, então a média e a variância de n_1, neste caso, são menores que a média e a variância de n_1 obtidas no caso inverso com reposição em que m é fixado (seção 1.4.a)

O seguinte teorema nos da a E.M.V. para N.

Teorema 19. A E.M.V. para N é igual a ∞ se $n_1 = \delta$; igual a $\frac{n_0 n_1}{n_1-\delta}$ se $\frac{n_0 n_1}{n_1-\delta}$ for um número inteiro e aproximadamente igual a $\frac{n_0 n_1}{n_1-\delta}$ se $\frac{n_0 n_1}{n_1-\delta}$ não for um número inteiro.

Prova.

A função de verossimilhança é dada por:

$$L(N|\delta,n_0,n_1) = \binom{n_1-1}{\delta-1}\left(1-\frac{n_0}{N}\right)^{\delta}\left(\frac{n_0}{N}\right)^{n_1-\delta} \propto K(N) = \left(1-\frac{n_0}{N}\right)^{\delta}\left(\frac{n_0}{N}\right)^{n_1-\delta},$$

onde $N \geq n_0 + 1$. Então,

1º) Se $n_1 = \delta$,

$$K(N) = \left(1 - \frac{n_0}{N}\right)^{\delta}.$$

Como $K(N) \uparrow 1$ quando $N \to \infty$, segue-se que o ponto de máximo de $K(N)$ é ∞ e, portanto, o E.M.V. de N é $\hat{N} = \infty$.*

2º) Se $n_1 > \delta$,

$$K(N) = \left(\frac{n_0}{N}\right)^{n_1-\delta}\left(1 - \frac{n_0}{N}\right)^{\delta}.$$

Logo, como este caso é análogo ao caso de amostragem direta, com reposição, estudado no Teorema 1, o teorema está provado. ∎

Como o estimador $\hat{N}_p$ tem esperança infinita e não tem variância, propomos o seguinte estimador para N

$$\hat{N}_5 = \frac{n_0(n_1+1)}{n_1-\delta+1}.$$

Teorema 20.

i) Se $\delta = 1$, então $E(\hat{N}_5) = n_0 - (N - n_0)\ln\left(1 - \frac{n_0}{N}\right)$.

ii) Se $\delta \geq 2$, então

$$E(\hat{N}_5) = n_0\left\{1 + \frac{\delta}{(\delta-1)}\frac{N}{n_0}\left(1 - \frac{n_0}{N}\right)\left[1 - \left(1 - \frac{n_0}{N}\right)^{\delta-1}\right]\right\}.$$

Prova .

i) Se $\delta = 1$,

* Nos casos em que a função de verossimilhança for estritamente crescente, vamos supor que N assume valores no conjunto $N^* = \{1, 2, \cdots\} \bigcup \{\infty\}$ e definimos: $K(\infty) = \lim_{N\to\infty} K(N)$.

$$E(\hat{N}_5) = n_0 E\left(\frac{n_1+1}{n_1}\right) = n_0 \sum_{k=1}^{\infty} \frac{(k+1)}{k}\left(1-\frac{n_0}{N}\right)\left(\frac{n_0}{N}\right)^{k-1}$$

$$= n_0\left(1-\frac{n_0}{N}\right)\sum_{j=0}^{\infty}\frac{(j+2)}{(j+1)}\left(\frac{n_0}{N}\right)^{j}$$

$$= n_0\left(1-\frac{n_0}{N}\right)\sum_{j=0}^{\infty}\left(1+\frac{1}{j+1}\right)\left(\frac{n_0}{N}\right)^{j}$$

$$= n_0\left[1+\left(1-\frac{n_0}{N}\right)\sum_{j=0}^{\infty}\frac{1}{j+1}\left(\frac{n_0}{N}\right)^{j}\right].$$

Como $ln(1-\frac{n_0}{N}) = -\sum\limits_{j=1}^{\infty}\frac{1}{j}(\frac{n_0}{N})^j$ (ver Feller (1980)), segue que,

$$\sum_{j=0}^{\infty}\frac{1}{j+1}\left(\frac{n_0}{N}\right)^{j} = \sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{n_0}{N}\right)^{j-1}$$

$$= \frac{N}{n_0}\sum_{j=1}^{\infty}\frac{1}{j}\left(\frac{n_0}{N}\right)^{j} = \frac{-N}{n_0} ln\left(1-\frac{n_0}{N}\right),$$

e portanto,

$$E(\hat{N}_5) = n_0\left[1-\left(1-\frac{n_0}{N}\right)\frac{N}{n_0}\ln\left(1-\frac{n_0}{N}\right)\right] = n_0-(N-n_0)ln\left(1-\frac{n_0}{N}\right).$$

o que prova (i).

ii) Se $\delta \geq 2$,

$$E(\hat{N}_5) = n_0 E\left[\frac{n_1+1}{n_1-\delta+1}\right] = n_0 \sum_{k=\delta}^{\infty} \frac{(k+1)}{(k-\delta+1)} \binom{k-1}{\delta-1} \left(1-\frac{n_0}{N}\right)^{\delta} \left(\frac{n_0}{N}\right)^{k-\delta}$$

$$= n_0 \left(1-\frac{n_0}{N}\right)^{\delta} \sum_{j=0}^{\infty} \left(\frac{\delta+j+1}{j+1}\right) \binom{\delta+j-1}{\delta-1} \left(\frac{n_0}{N}\right)^{j}$$

$$= n_0 \left(1-\frac{n_0}{N}\right)^{\delta} \sum_{j=0}^{\infty} \left(1+\frac{\delta}{j+1}\right) \binom{\delta+j-1}{j} \left(\frac{n_0}{N}\right)^{j}$$

$$= n_0 \left(1-\frac{n_0}{N}\right)^{\delta} \left[\sum_{j=0}^{\infty} \binom{-\delta}{j} \left(\frac{-n_0}{N}\right)^{j} + \delta \sum_{j=0}^{\infty} \frac{1}{j+1} \binom{-\delta}{j} \left(\frac{-n_0}{N}\right)^{j}\right]$$

$$= n_0 \left[1 + \delta \left(1-\frac{n_0}{N}\right)^{\delta} \sum_{j=0}^{\infty} \frac{1}{j+1} \binom{-\delta}{j} \left(\frac{-n_0}{N}\right)^{j}\right]. \qquad (1.6.)$$

Como

$$\sum_{j=0}^{\infty} \binom{-\delta}{j} (-1)^j x^j = (1-x)^{-\delta} \qquad \forall -1 < x < 1$$

segue que,

$$\sum_{j=0}^{\infty} \binom{-\delta}{j} (-1)^j \int_0^{\frac{n_0}{N}} x^j dx = \sum_{j=0}^{\infty} \binom{-\delta}{j} (-1)^j \frac{x^{j+1}}{j+1}\Bigg|_0^{\frac{n_0}{N}}$$

$$= \sum_{j=0}^{\infty} \binom{-\delta}{j} (-1)^j \frac{1}{j+1} \left(\frac{n_0}{N}\right)^{j+1}$$

$$= \frac{n_0}{N} \sum_{j=0}^{\infty} \binom{-\delta}{j} \frac{1}{j+1} \left(-\frac{n_0}{N}\right)^j$$

$$= \int_0^{\frac{n_0}{N}} (1-x)^{-\delta} dx == \left. \frac{-(1-x)^{-\delta+1}}{-\delta+1} \right|_0^{\frac{n_0}{N}}$$

$$= \frac{-1}{(-\delta+1)} \left(1 - \frac{n_0}{N}\right)^{-\delta+1} + \frac{1}{-\delta+1}$$

$$= \frac{1}{\delta-1} \left[\left(1 - \frac{n_0}{N}\right)^{-\delta+1} - 1 \right].$$

Logo ,

$$\sum_{j=0}^{\infty} \frac{1}{j+1} \binom{-\delta}{j} \left(\frac{-n_0}{N}\right)^j = \frac{N}{n_0} \frac{1}{(\delta-1)} \left[\left(1 - \frac{n_0}{N}\right)^{-\delta+1} - 1 \right].$$

Usando o resultado que acabamos de obter, em (1.6.), segue que

$$E(\hat{N}_5) = n_0 \left[1 + \frac{\delta N}{(\delta-1) n_0} \left(1 - \frac{n_0}{N}\right) \left(1 - \left(1 - \frac{n_0}{N}\right)^{\delta-1}\right) \right],$$

o que prova o teorema. ∎

O seguinte teorema fornece o limite de $E(\hat{N}_5)$ quando $N \to \infty$.

Teorema 21 . Para todo $\delta \geq 1$,

$$\lim_{N \to \infty} \quad E(\hat{N}_5) = n_0(1+\delta).$$

Prova .

a) Se $\delta = 1$,

$$\lim_{N\to\infty} E(\hat{N}_5) = \lim_{N\to\infty}\left[n_0 - (N-n_0)\ln\left(1-\frac{n_0}{N}\right)\right]$$

$$= n_0 - \ln \lim_{N\to\infty}\left(1-\frac{n_0}{N}\right)^{N-n_0}$$

$$= n_0 - \ln \lim_{N\to\infty}\left[\left(1-\frac{n_0}{N}\right)^{N}\left(1-\frac{n_0}{N}\right)^{-n_0}\right]$$

$$= n_0 - \ln e^{-n_0} = 2n_0.$$

b) Se $\delta \geq 2$, como

$$N\left[\left(1-\frac{n_0}{N}\right)^{1-\delta} - 1\right] = \frac{\left(1-\frac{n_0}{N}\right)^{1-\delta} - 1}{\frac{1}{N}},$$

aplicando a regra de L'Hospital temos:

$$\lim_{N\to\infty} N\left[\left(1-\frac{n_0}{N}\right)^{1-\delta} - 1\right] = \lim_{N\to\infty}\frac{(1-\delta)\left(1-\frac{n_0}{N}\right)^{-\delta}\frac{n_0}{N^2}}{\frac{-1}{N^2}}$$

$$= -n_0(1-\delta) = n_0(\delta-1)$$

e como

$$\lim_{N\to\infty}\left(1-\frac{n_0}{N}\right)^{\delta} = 1,$$

segue que

$$\lim_{N\to\infty} E(\hat{N}_5) = \lim_{N\to\infty}\left\{n_0 + \frac{\delta}{\delta-1}\left(1-\frac{n_0}{N}\right)^{\delta} N\left[\left(1-\frac{n_0}{N}\right)^{1-\delta} - 1\right]\right\}$$

$$= n_0 + \frac{\delta}{(\delta-1)} n_0(\delta-1) = n_0(1+\delta).$$

■

Um estimador para $\frac{1}{N}$ é o inverso de $\hat{N}_p$:

$$\frac{1}{\hat{N}_p} = \frac{n_1 - \delta}{n_0 n_1}.$$

Se $\delta > 1$, uma pequena modificação deste estimador nos fornece um estimador não viciado para $\frac{1}{N}$,

$$\hat{I}_3 = \frac{n_1 - \delta}{n_0(n_1 - 1)} = \frac{1}{n_0}\left(1 - \frac{\delta - 1}{n_1 - 1}\right)$$

Teorema 22. $\hat{I}_3$ é um estimador não viciado para $\frac{1}{N}$.

Prova . Como

$$\begin{aligned} E\left(\frac{\delta-1}{n_1-1}\right) &=(\delta-1)\sum_{k=\delta}^{\infty}\frac{1}{k-1}\binom{k-1}{\delta-1}\left(1-\frac{n_0}{N}\right)^{\delta}\left(\frac{n_0}{N}\right)^{k-\delta} \\ &= \left(1-\frac{n_0}{N}\right)\sum_{k=\delta}^{\infty}\binom{k-2}{\delta-2}\left(\frac{n_0}{N}\right)^{k-1-(\delta-1)}\left(1-\frac{n_0}{N}\right)^{\delta-1} \\ &= \left(1-\frac{n_0}{N}\right)\sum_{j=\delta-1}^{\infty}\binom{j-1}{(\delta-1)-1}\left(1-\frac{n_0}{N}\right)^{\delta-1}\left(\frac{n_0}{N}\right)^{j-(\delta-1)} = 1-\frac{n_0}{N}, \end{aligned}$$

segue que

$$E(\hat{I}_3) = \frac{1}{n_0} - \frac{1}{n_0} + \frac{1}{N} = \frac{1}{N}.$$

■

b) **SEM REPOSIÇÃO.**

Sejam

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população;

n_1: número de indivíduos capturados, sem reposição, até que ocorra o δ-ésimo indivíduo não marcado, onde δ é um número inteiro fixado, $1 \leq \delta \leq N - n_0$.

A variável aleatória n_1 tem distribuição de probabilidade Hipergeométrica Negativa com parâmetros δ, n_0 e N.

Notando que o evento "n_1 elementos são selecionados da população sem reposição, até que ocorra o δ-ésimo elemento não marcado " é o evento "$(\delta-1)$ elementos não marcados são selecionados da população sem reposição, nas $(n_1 - 1)$ primeiras seleções, e na n_1-ésima seleção é selecionado um indivíduo não marcado", segue que a distribuição de probabilidade de n_1 é

$$P(n_1|n_0,\delta,N) = \frac{\binom{N-n_0}{\delta-1}\binom{n_0}{n_1-\delta}}{\binom{N}{n_1-1}} \frac{[N - n_0 - (\delta - 1)]}{[N - (n_1 - 1)]}$$

$$= \frac{(N - n_0)!n_0!(n_1 - 1)!(N - n_1)!}{N!(n_1 - \delta)!(n_0 - n_1 + \delta)!(\delta - 1)!(N - n_0 - \delta)!}$$

$$= \frac{(N - n_0)}{N} \frac{\binom{N-n_0-1}{\delta-1}\binom{n_0}{n_1-\delta}}{\binom{N-1}{n_1-1}}, \quad \delta \leq n_1 \leq n_0 + \delta$$

Observando a última expresão da distribuição acima, notamos que ela é igual à probabilidade do evento "um elemento não marcado é selecionado, sem reposição, na primeira seleção e $(\delta - 1)$ elementos não marcados são capturados quando uma amostra de tamanho $(n_1 - 1)$ é selecionada, sem reposição, da população, agora contendo $(N - n_0 - 1)$ elementos não marcados e n_0 elementos marcados".

Usando os resultados do Lema 15,

$$E(n_1) = \frac{\delta(N+1)}{(N-n_0+1)} = \frac{\delta}{(1-\frac{n_0}{N+1})}$$

e

$$V(n_1) = \frac{\delta(N+1)n_0(N-n_0-\delta+1)}{(N-n_0+1)^2(N-n_0+2)}$$

$$= \frac{\delta n_0(1-\frac{n_0+\delta}{N+1})}{(1-\frac{n_0}{N+1})^2(N-n_0+2)}.$$

Notemos que $V(n_1) \xrightarrow[N\to\infty]{} 0.$

Notemos que, como no caso com reposição, se $n_0 < \frac{N}{2}$ e $\delta = m$, então a média de n_1 é menor que a média obtida no caso sem reposição, em que m é fixado (seção 1.5.b). Quanto à variância, notemos que $Var(n_1) \to \infty$ quando $N \to \infty$ com m fixado (seção 1.5.b) e $Var(n_1) \to 0$ quando $N \to \infty$ no caso em que δ é fixado.

Teorema 23. O E.M.V.de N é aproximadamente $\hat{N}_p = \frac{n_0 n_1}{n_1-\delta}$ se $\frac{n_0 n_1}{n_1-\delta}$ não for um número inteiro e é igual a $\hat{N}_p$ se $\frac{n_0 n_1}{n_1-\delta}$ for um número inteiro.

Prova. A função de verossimilhança é dada por:

$$L(N|\delta, n_0, n_1) = \frac{(N-n_0)!n_0!(n_1-1)!(N-n_1)!}{N!(n_1-\delta)!(n_0-n_1+\delta)!(\delta-1)!(N-n_0-\delta)!}$$

$$\propto K(N) = \frac{(N-n_0)!(N-n_1)!}{N!(N-n_0-\delta)!}, \qquad N \geq n_0+\delta.$$

Este caso é análogo ao caso de amostragem direta sem reposição estudado no Teorema 6. Logo,

1) se $\frac{n_0 n_1}{n_1-\delta}$ não for inteiro, então a Estimativa de Máxima Verossimilhança de N é

$$\left[\frac{n_0 n_1}{n_1 - \delta}\right] \cong \hat{N}_p.$$

2)se $\frac{n_0 n_1}{n_1 - \delta}$ for inteiro, segue que as estimativas de Máxima Verossimilhança para N são

$$\frac{n_0 n_1}{n_1 - \delta} = \hat{N}_p \qquad e \qquad \frac{n_0 n_1}{n_1 - \delta} - 1 = \hat{N}_p - 1.$$

■

Como $\hat{N}_p$ tem média infinita e sua variância não existe, Chapman (1952) propôs o estimador

$$\hat{N}_6 = \frac{(n_0 + 1)n_1}{n_1 - \delta + 1} - 1.$$

O seguinte teorema nos dá a esperança de $\hat{N}_6$.

Teorema 24. Desde que n_1 é uma variável aleatória com distribuição de probabilidade Hipergeométrica Negativa de parâmetros δ, $N - n_0$ e N, então

$$E(\hat{N}_6) = N - (N+1)\left(1 - \frac{\delta}{N+1}\right)\prod_{j=0}^{n_0-1}\left(1 - \frac{\delta}{N-j}\right).$$

Prova . Como

$$E\left(\frac{n_1}{n_1 - \delta + 1}\right) = \sum_{k=\delta}^{n_0+\delta} \frac{k}{(k - \delta + 1)} \frac{(N - n_0)!n_0!(k-1)!(N-k)!}{N!(k-\delta)!(n_0 - k + \delta)!(\delta - 1)!(N - n_0 - \delta)!}$$

$$= \frac{(N+1)}{(n_0+1)} \sum_{k=\delta}^{n_0+\delta} \frac{\binom{k}{\delta-1}\binom{N-k}{n_0-k+\delta}}{\binom{N+1}{n_0+1}}$$

$$= \frac{(N+1)}{(n_0+1)} \sum_{j=1}^{n_0+1} \frac{\binom{\delta+j-1}{j}\binom{N+1-\delta-j}{n_0-j+1}}{\binom{N+1}{n_0+1}}$$

$$= \frac{(N+1)}{(n_0+1)} \left(\sum_{j=0}^{n_0+1} \frac{\binom{\delta+j-1}{j}\binom{N+1-\delta-j}{n_0-j+1}}{\binom{N+1}{n_0+1}} - \frac{\binom{N+1-\delta}{n_0+1}}{\binom{N+1}{n_0+1}} \right),$$

usando o resultado do Apêndice 2 para $a = \delta$, $b = N+1-\delta-n_0$ e $k = n_0+1$, segue que

$$E\left(\frac{n_1}{n_1-\delta+1}\right) = \frac{(N+1)}{(n_0+1)}\left[1 - \frac{(N-n_0)!(N+1-\delta)!}{(N+1)!(N-n_0-\delta)!}\right]$$

$$= \frac{(N+1)}{(n_0+1)}\left[1 - \frac{(N-\delta+1)(N-\delta)(N-\delta-1)\cdots(N-\delta-n_0+1)}{(N+1)N(N-1)\cdots(N-(n_0-1))}\right]$$

$$= \frac{(N+1)}{(n_0+1)}\left[1 - \left(1-\frac{\delta}{N+1}\right)\prod_{j=0}^{n_0-1}\left(1-\frac{\delta}{N-j}\right)\right]. \qquad (1.7.)$$

Então, usando o resultado anterior, temos:

$$\begin{aligned} E(\hat{N}_6) &= E\left[\frac{(n_0+1)n_1}{n_1-\delta+1} - 1\right] \\ &= (n_0+1)E\left(\frac{n_1}{n_1-\delta+1}\right) - 1 \\ &= \frac{(n_0+1)(N+1)}{(n_0+1)}\left[1-\left(1-\frac{\delta}{N+1}\right)\prod_{j=0}^{n_0-1}\left(1-\frac{\delta}{N-j}\right)\right]-1 \\ &= N-(N+1)\left(1-\frac{\delta}{N+1}\right)\prod_{j=0}^{n_0-1}\left(1-\frac{\delta}{N-j}\right). \end{aligned}$$

■

Para controlar $|E(\hat{N}_6) - N|$ estabelecemos o seguinte teorema.

Teorema 25 . Para todo K, $0 < K < N$, se $\frac{n_0\delta}{N} \geq \ln\left(\frac{N}{K}\right)$, então

$$|E(\hat{N}_6) - N| < K.$$

Em particular, se $\frac{n_0\delta}{N} \geq \ln N$ então $|E(\hat{N}_6) - N| < 1$.

Prova. Pelo Teorema 24

$$
\begin{aligned}
|E(\hat{N}_6) - N| &= (N+1)\left(1 - \frac{\delta}{N+1}\right)\prod_{j=0}^{n_0-1}\left(1 - \frac{\delta}{N-j}\right) \\
&= [N - (\delta - 1)]\prod_{j=0}^{n_0-1}\left(1 - \frac{\delta}{N-j}\right) \\
&< N\left(1 - \frac{\delta}{N}\right)^{n_0} < Ne^{\frac{-\delta n_0}{N}}.
\end{aligned}
$$

Como $\frac{n_0\delta}{N} \geq \ln\left(\frac{N}{K}\right)$, então

$$|E(\hat{N}_6) - N| < Ne^{\frac{-\delta n_0}{N}} \leq Ne^{\ln\left(\frac{K}{N}\right)} = K.$$

■

Por exemplo, de acordo com a Tabela 1, se $N = 10^4$ e $K = 2$, o mínimo valor de $n_0\delta$ para o qual $|E(\hat{N}_6) - 10^4| < 2$ é $8,52 \times 10^4$.

Se $N = 10^5$ e $n_0\delta \geq 10^5 \times 11,51$, então, de acordo com a Tabela 1, $|E(\hat{N}_6) - 10^5| < 1$.

O seguinte Lema é útil para o cálculo da variância de $\hat{N}_6$.

Lema 26. Sejam

$$n_{(i)} = \begin{cases} n(n-1)(n-2)\cdots(n-i+1) & \text{se } i \geq 1 \\ 1 & \text{se } i = 0; \end{cases}$$

$$q_{ij} = \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}$$

e

$$\eta_{ij} = \sum_{r=0}^{j-1}\frac{\binom{\delta+r-j+i-1}{r}\binom{N-\delta-r+j}{n_0-r+j}}{\binom{N+i}{n_0+j}},$$

para $i = 1, 2$ e $j = 2, 3, 4$. Se $\delta \geq 2$, então,

$$E\left[\frac{(n_1+i-1)_{(i)}}{(n_1-\delta+j)_{(j)}}\right] = q_{ij}(1-\eta_{ij}).$$

Prova.

$$E\left[\frac{(n_1+i-1)_{(i)}}{(n_1-\delta+j)_{(j)}}\right] = \sum_{k=\delta}^{n_0+\delta} \frac{(k+i-1)_{(i)}(N-n_0)!n_0!(k-1)!(N-k)!}{(k-\delta+j)_{(j)}N!(k-\delta)!(n_0-k+\delta)!(\delta-1)!(N-n_0-\delta)!}$$

$$= \frac{(N\text{-}n_0)!n_0!}{N!(\delta\text{-}1)!(N\text{-}n_0\text{-}\delta)!}\sum_{k=\delta}^{n_0+\delta}\frac{(k+i\text{-}1)(k+i\text{-}2)\cdots k(k\text{-}1)!(N\text{-}k)!}{(k\text{-}\delta+j)(k\text{-}\delta+j\text{-}1)\cdots(k\text{-}\delta+1)(k\text{-}\delta)!(n_0\text{-}k+\delta)!}$$

$$= \frac{(N+i)(N+i\text{-}1)\cdots(N+1)(N\text{-}n_0)(N\text{-}n_0\text{-}1)\cdots(N\text{-}n_0\text{-}j+i+1)}{(N+i)!(n_0+j)(n_0+j\text{-}1)\cdots(n_0+1)(\delta\text{-}1)(\delta\text{-}2)\cdots(\delta\text{-}j+i)}\times$$

$$\times\frac{(N+i\text{-}n_0\text{-}j)!}{(\delta\text{-}j+i\text{-}1)!(N\text{-}n_0\text{-}\delta)!}\sum_{k=\delta}^{n_0+\delta}\frac{(k+i-1)!(N-k)!(n_0+j)!}{(k-\delta+j)!(n_0-k+\delta)!}$$

$$= \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}\sum_{k=\delta}^{n_0+\delta}\frac{\binom{k+i-1}{k-\delta+j}\binom{N-k}{n_0-k+\delta}}{\binom{N+i}{n_0+j}}$$

$$= \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}\sum_{l=0}^{n_0}\frac{\binom{\delta+l+i-1}{l+j}\binom{N-\delta-l}{n_0-l}}{\binom{N+i}{n_0+j}}$$

$$= \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}\left[\sum_{l=-j}^{n_0}\frac{\binom{\delta+l+i-1}{l+j}\binom{N-\delta-l}{n_0-l}}{\binom{N+i}{n_0+j}} - \sum_{l=-j}^{-1}\frac{\binom{\delta+l+i-1}{l+j}\binom{N-\delta-l}{n_0-l}}{\binom{N+i}{n_0+j}}\right]$$

$$= \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}\left[\sum_{r=0}^{n_0+j}\frac{\binom{\delta+r-j+i-1}{r}\binom{N-\delta-r+j}{n_0-r+j}}{\binom{N+i}{n_0+j}} - \sum_{r=0}^{j-1}\frac{\binom{\delta+r-j+i-1}{r}\binom{N-\delta-r+j}{n_0-r+j}}{\binom{N+i}{n_0+j}}\right].$$

Usando o resultado do Apêndice 2 para $a = \delta + i - j$, $b = N - \delta - n_0 + 1$ e $k = n_0 + j$, segue que

$$E\left[\frac{(n_1+i-1)_{(i)}}{(n_1-\delta+j)_{(j)}}\right] = \frac{(N+i)_{(i)}(N-n_0)_{(j-i)}}{(n_0+j)_{(j)}(\delta-1)_{(j-i)}}\left[1-\sum_{r=0}^{j-1}\frac{\binom{\delta+r-j+i-1}{r}\binom{N-\delta-r+j}{n_0-r+j}}{\binom{N+i}{n_0+j}}\right]$$

$$= q_{ij}(1-\eta_{ij}),$$

o que prova o Lema. ■

O seguinte teorema nos dá um valor aproximado da variância de $\hat{N}_6$, para $\delta \geq 2$.

Teorema 27. Sejam $n_{(i)}, q_{ij}$ e η_{ij} definidos como no Lema anterior. Se $\delta \geq 2$ e $n_0\delta > N \ln N$, então,

$$V(\hat{N}_6) \cong (n_0+1)^2[q_{22}(1-\eta_{22}) + q_{23}(1-\eta_{23}) + 2q_{24}(1-\eta_{24}) -$$

$$-q_{12}(1-\eta_{12}) - q_{13}(1-\eta_{13}) - 2q_{14}(1-\eta_{14})] - 2N - 1 - N^2.$$

Prova. Como

$$(\hat{N}_6+1)^2 = \frac{(n_0+1)^2 n_1^2}{(n_1-\delta+1)^2},$$

tomando os três primeiros termos da série fatorial inversa $(n_1-\delta+1)^{-2}$ (ver Apêndice 5), resulta que,

$$(\hat{N}_6+1)^2 \cong (n_0+1)^2[n_1(n_1+1) - n_1]\left[\frac{1}{(n_1-\delta+1)(n_1-\delta+2)} + \right.$$

$$\left. + \frac{1}{(n_1-\delta+1)(n_1-\delta+2)(n_1-\delta+3)} + \frac{2}{(n_1-\delta+1)(n_1-\delta+2)(n_1-\delta+3)(n_1-\delta+4)}\right].$$

Observamos que na expressão acima desprezamos o resto $R_2(n_1 - \delta) < 0,2839$ (ver observação feita após Apêndice 4).

Então

$$(\hat{N}_6 + 1)^2 = \hat{N}_6^2 + 2\hat{N}_6 + 1$$

$$\cong (n_0+1)^2\left[\frac{n_1(n_1+1)}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)} + \frac{n_1(n_1+1)}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)} + \right.$$

$$+\frac{2n_1(n_1+1)}{(n_1-\delta+1)(n_1-\delta+2)(n_1-\delta+3)(n_1-\delta+4)} - \frac{n_1}{(n_1-\delta+1)(n_1-\delta+2)} -$$

$$\left. - \frac{n_1}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)} - \frac{2n_1}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)(n_1\text{-}\delta+4)}\right].$$

Logo,

$$E(\hat{N}_6+1)^2 \cong (n_0+1)^2\left\{E\left[\frac{n_1(n_1+1)}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)}\right] + E\left[\frac{n_1(n_1+1)}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)}\right] + \right.$$

$$+2E\left[\frac{n_1(n_1+1)}{(n_1-\delta+1)(n_1-\delta+2)(n_1-\delta+3)(n_1-\delta+4)}\right] - E\left[\frac{n_1}{(n_1-\delta+1)(n_1-\delta+2)}\right] -$$

$$\left. -E\left[\frac{n_1}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)}\right] - 2E\left[\frac{n_1}{(n_1\text{-}\delta+1)(n_1\text{-}\delta+2)(n_1\text{-}\delta+3)(n_1\text{-}\delta+4)}\right]\right\}.$$

Usando o Lema anterior para $i = 1,2$ e $j = 2,3,4$ temos,

$$E(\hat{N}_6)^2 \cong (n_0+1)^2[q_{22}(1\text{-}\eta_{22}) + q_{23}(1\text{-}\eta_{23}) + 2q_{24}(1\text{-}\eta_{24})\text{-}q_{12}(1\text{-}\eta_{12})\text{-}q_{13}(1\text{-}\eta_{13})\text{-}2q_{14}(1\text{-}\eta_{14})].$$

Logo,

$$V(\hat{N}_6) = E(\hat{N}_6^2) - E^2(\hat{N}_6) = E(\hat{N}_6 + 1)^2 - 2E(\hat{N}_6) - 1 - E^2(\hat{N}_6)$$

$$\cong (n_0 + 1)^2[q_{22}(1 - \eta_{22}) + q_{23}(1 - \eta_{23}) + 2q_{24}(1 - \eta_{24}) - q_{12}(1 - \eta_{12}) - q_{13}(1 - \eta_{13}) -$$

$$-2q_{14}(1 - \eta_{14})] - 2E(\hat{N}_6) - 1 - E^2(\hat{N}_6).$$

Se $n_0\delta > N \ln N$, pelo Teorema 26, $|E(\hat{N}_6) - N| < 1$, e portanto,

$$V(\hat{N}_6) \cong (n_0 + 1)^2[q_{22}(1 - \eta_{22}) + q_{23}(1 - \eta_{23}) + 2q_{24}(1 - \eta_{24}) - q_{12}(1 - \eta_{12}) - q_{13}(1 - \eta_{13}) -$$

$$-2q_{14}(1 - \eta_{14})] - 2N - 1 - N^2,$$

o que prova o teorema . ∎

1.7. COMPARAÇÃO ENTRE O PROCESSO DIRETO SIMPLES E O PROCESSO INVERSO SIMPLES

Suponhamos que um pesquisador realize um experimento segundo o processo direto simples e que um segundo pesquisador realize um experimento segundo o processo inverso simples, fixando o número de indivíduos marcados aproximadamente igual ao número médio de indivíduos marcados obtidos pelo primeiro pesquisador. Uma questão natural que se coloca é: existe alguma vantagem de algum processo em relação ao outro?
Na próxima seção veremos sob que condições um pesquisador apresenta alguma vantagem sobre o outro.

1.8. Comparação entre o processo direto simples com reposição e o processo inverso simples (número de marcados fixado) com reposição.

Consideremos um processo direto simples com reposição tal que,

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra;

n_1: tamanho da segunda amostra selecionada com reposição e

m: número de indivíduos marcados observados na segunda amostra.

Pelo resultado da seção 1.1,

$$E(m) = \frac{n_0 n_1}{N}.$$

Consideremos agora um processo inverso simples tal que,

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra;

m': número de indivíduos marcados, observados na segunda amostra definido (fixado)por

$$m' = \begin{cases} 1 & \text{se } \frac{n_0 n_1}{N} = E(m) < 1 \\ \left[\frac{n_0 n_1}{N}\right] & \text{se } \frac{n_0 n_1}{N} \geq 1 (1.\ .) \end{cases}$$

e

n_1': tamanho (aleatório) da segunda amostra.

Com relação a estes dois processos vale o seguinte resultado.

Teorema 28.

$$i) Se \quad N > n_0 n_1, \text{ então } E(n_1') > n_1$$

$$ii) Se \quad N \leq n_0 n_1, \text{ então } E(n_1') \leq n_1.$$

Prova.

i) Se $N > n_0 n_1$, então

$$E(n_1') = \frac{m'N}{n_0} = \frac{N}{n_0} > n_1,$$

o que prova i).

ii) Se $N \leq n_0 n_1$, então

$$E(n_1') = \frac{m'N}{n_0} = \frac{\left[\frac{n_0 n_1}{N}\right] N}{n_0} \leq \frac{n_0 n_1}{N}\frac{N}{n_0} = n_1,$$

o que prova o teorema. ■

Este teorema nos permite concluir que, para $N \leq n_0 n_1$, o número médio de capturas do processo inverso simples (onde o número de marcados é fixado aproximadamente igual ao número médio de marcados do processo direto) é menor ou igual que o tamanho da segunda amostra do processo direto .

Por outro lado, se $N > n_0 n_1$ (nesse caso o número de marcados é fixado igual a 1), o número médio de capturas do processo inverso simples é maior que o tamanho da segunda amostra do processo direto.

Comparemos agora as esperanças e as variâncias dos estimadores de N apresentados nas seções 1.1.a e 1.4.a para ambos os processos.

Então, para o processo direto consideremos o estimador de N,

$$\hat{N}_2 = \frac{n_0(n_1+1)}{m+1}$$

Pelo Lema 2,

$$E(\hat{N}_2) = N - N\left(1 - \frac{n_0}{N}\right)^{n_1+1};$$

pelo Teorema 3, $\forall\ K,\ 0 < K < N$, se

$$\frac{n_0 n_1}{N} \geq \ln\frac{N}{K}, \quad \text{então } |E(\hat{N}_2) - N| < K,$$

e pelo Lema 4,

$$V(\hat{N}_2) \leq \frac{n_0^2 n_1^2}{4}.$$

Para o processo inverso consideremos o estimador $\hat{N}_p$, que é não viciado e cuja variância é

$$V(\hat{N}_p) = \frac{N(N-n_0)}{m'} \quad \text{(ver seção 1.4.a).}$$

Se $n_0n_1 < N$, como $m' = 1$ (ver a expressão 1.) então, para $n_1 \geq 2$,

$$V(\hat{N}_p) = \frac{N(N-n_0)}{m'} = N(N-n_0) > n_0n_1(n_0n_1 - n_0)$$

$$= {n_0}^2 n_1(n_1 - 1) \geq \frac{{n_0}^2{n_1}^2}{4} \geq V(\hat{N}_2)$$

e além disso, pelo Teorema 28, $E(n_1') > n_1$.

Logo, podemos concluir que, para $n_1 \geq 2$ e $n_0n_1 < N$, o processo direto nos fornece um estimador de N, $\hat{N}_2$, cuja esperança é igual a

$$E(\hat{N}_2) = N - N\left(1 - \frac{n_0}{N}\right)^{n_1+1}$$

e cuja variância é menor que a variância de $\hat{N}_p$: estimador não viciado de N obtido pelo processo inverso e, além disso, pelo Teorema 28, o tamanho da segunda amostra no processo direto é menor que o número médio de capturas do processo inverso.

Por outro lado, como $N \geq 3$ implica que $N \ln N > N$, então para $N \geq 3$ segue que, se $n_0n_1 \geq N \ln N$ temos:

i) $n_0n_1 > N$, o que implica, pelo Teorema 28, que $E(n_1') \leq n_1$;

ii) pelo Teorema 3, $|E(\hat{N}_2) - N| < 1$.

Comparemos agora as variâncias de $\hat{N}_p$ e $\hat{N}_2$. Lembremos que

$$V(\hat{N}_2) \leq (\frac{n_0n_1}{2})^2$$

e

$$V(\hat{N}_p) = \frac{N(N-n_0)}{m'}.$$

Se $n_0n_1 \geq 2N$ ($\Rightarrow n_0n_1 > N$), e como pelo Teorema 28, $m' = [\frac{n_0n_1}{N}]$, então $m' \geq 2$ ($1/m' \leq 1/2$). Logo,

$$V(\hat{N}_p) = \frac{N(N-n_0)}{m,} \leq \frac{(n_0n_1/2)(n_0n_1/2 - n_0)}{2}$$

$$= \frac{{n_0}^2 n_1}{4}(n_1/2 - 1) = \frac{{n_0}^2{n_1}^2}{4}(1/2 - 1/n_1) \quad < {n_0}^2{n_1}^2/4.$$

Portanto, podemos concluir que, se $N \geq 3$ e $n_0 n_1 \geq \max\{N \ln N, 2N\}$, o processo direto fornece um estimador para N, $\hat{N}_2$, quase não viciado; o número médio de capturas no processo inverso é menor ou igual ao tamanho da segunda amostra: $E(n'1) \leq n_1$ e, além disso, as variâncias dos estimadores obtidos por ambos os processos são limitadas superiormente por ${n_0}^2 {n_1}^2/4$.

1.9. Comparação entre o processo direto simples sem reposição e o processo inverso simples (número de marcados fixado) sem reposição.

Consideremos um processo direto simples sem reposição, tal que

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra,

n_1: tamanho da segunda amostra, selecionada sem reposição e

m: número de indivíduos marcados observados na segunda amostra.

Pelo resultado da seção 1.1.b

$$E(m) = \frac{n_0 n_1}{N}.$$

Como na seção anterior, consideremos agora um processo inverso simples sem reposição com

n_0: número de indivíduos capturados sem reposição, marcados e devolvidos à população na primeira amostra.

m': número de indivíduos marcados observados na segunda amostra definido (fixado) por

$$m' = \begin{cases} 1 & \text{se } \frac{n_0 n_1}{N} = E(m) < 1 \\ \left[\frac{n_0 n_1}{N}\right] & \text{se } \frac{n_0 n_1}{N} \geq 1 \end{cases}$$

e $n_1' =$ tamanho (aleatório) da segunda amostra.

Pelo Lema 15 ,

$$E(n_1') = \frac{m'(N+1)}{n_0 + 1}.$$

Com relação a estes dois processos, temos o seguinte resultado.

Teorema 29.

i)*Se* $N < (n_0 + 1)n_1$, então $E(n_1') \leq n_1$.

ii)*Se* $N \geq (n_0 + 1)n_1$, então $E(n_1') > n_1$.

Prova.

$$i)\text{Se} \qquad N \leq n_0 n_1, \text{ então} E(n_1') = \frac{m'(N+1)}{n_0+1} = \frac{[\frac{n_0 n_1}{N}](N+1)}{n_0+1}$$

$$\leq \frac{n_0 n_1}{N} \frac{(N+1)}{(n_0+1)} = \frac{(1+\frac{1}{N})}{1+\frac{1}{n_0}} n_1 \leq n_1;$$

se $n_0 n_1 < N < (n_0+1)n_1$, então $n_0 n_1 < N$ e $(N+1) \leq (n_0+1)n_1$

e, portanto

$$E(n_1') = \frac{m'(N+1)}{n_0+1} = \frac{(N+1)}{n_0+1} \leq n_1,$$

o que prova i).

$$ii)Se\ N \geq (n_0+1)n_1, \text{ então } \ n_0 n_1 < N \ e\ N+1 > (n_0+1)n_1$$

e, portanto,

$$E(n_1') = \frac{m'(N+1)}{n_0+1} = \frac{N+1}{n_0+1} > n_1,$$

o que prova o teorema. ∎

Este teorema nos permite concluir que se $N < (n_0+1)n_1$ o número médio de capturas do processo inverso simples (onde o número de marcados é fixado aproximadamente igual ao número médio de marcados do processo direto) é menor ou igual que o tamanho da segunda amostra do processo direto. Por outro lado, se $N \geq (n_0+1)n_1$, o número médio de capturas do processo inverso simples é maior que o tamanho da segunda amostra do processo direto.

Comparemos agora as esperanças e variâncias dos estimadores de N apresentados nas seções 1.1.b e 1.5.b para ambos os processos. Então, para o processo direto consideremos o estimador

$$\hat{N}_3 = \frac{(n_0+1)(n_1+1)}{m} - 1$$

cuja esperança, pelo Teorema 7, é

$$E(\hat{N}_3) = N - \frac{\binom{N-n_0}{n_1+1}(n_1+1)}{\binom{N}{n_1}}$$

e cuja variância aproximada, para $\frac{n_0 n_1}{N} > 10$ (Seber (1981)) é

$$V(\hat{N}_3) \cong N^2 \left[\left(\frac{n_0 n_1}{N} \right)^{-1} + 2\left(\frac{n_0 n_1}{N} \right)^{-2} + 6\left(\frac{n_0 n_1}{N} \right)^{-3} \right].$$

Para o processo inverso consideremos o estimador não viciado para N

$$\hat{N}_4 = \frac{n_1'(n_0+1)}{m'} - 1$$

cuja variância, pelo Teorema 18, é

$$V(\hat{N}_4) = \frac{(N+1)(N-n_0)(n_0-m'+1)}{m'(n_0+2)}.$$

Além disso (ver a expressão 1.5.)

$$V(\hat{N}_4) < \frac{N^2}{m'}.$$

Se $n_0 n_1 > 10N$, usando o resultado do Apêndice 4, temos,

$$V(\hat{N}_3) \cong N^2 \left[\left(\frac{n_0 n_1}{N} \right)^{-1} + 2\left(\frac{n_0 n_1}{N} \right)^{-2} + 6\left(\frac{n_0 n_1}{N} \right)^{-3} \right]$$

$$> N^2 \frac{1}{\left[\frac{n_0 n_1}{N}\right]} = \frac{N^2}{m'} > V(\hat{N}_4).$$

Por outro lado, pelo Corolário 8, se tomarmos $K = 1$, então,

$$n_0 n_1 \geq N \ln N \Rightarrow |E(\hat{N}_3) - N| < 1.$$

Como $N \geq 3$ implica que $N \ln N > N$, então para $N \geq 3$ segue que se $n_0 n_1 \geq N \ln N$ temos:

i) $n_0 n_1 > N$ o que implica $(n_0+1)n_1 > N$. Logo, pelo Teorema 29, $E(n_1') \leq n_1$.

ii)Pelo Corolário 8 $|E(\hat{N}_3) - N| < 1$.

Portanto, podemos concluir que se $N \geq 3$ e $n_0 n_1 \geq \max\{10N, N \ln N\}$, então o processo inverso fornece um etimador não viciado para N, $\hat{N}_4$, com variância menor que a variância do estimador dado pelo processo direto, e o número médio de capturas no processo inverso é menor ou igual ao tamanho da segunda amostra no processo direto.

CAPÍTULO 2

AMOSTRAGEM MÚLTIPLA

2.1. AMOSTRAGEM DIRETA MÚLTIPLA SEM REPOSIÇÃO

Nesta seção estudaremos o método de amostragem direta múltipla sem reposição, no qual em cada uma de k etapas ($k \geq 2$) são selecionados ao acaso e sem reposição um número fixado de indivíduos da população. Em cada etapa observamos o número de indivíduos marcados; aqueles que não estão marcados são marcados e, em seguida, todos os indivíduos selecionados são devolvidos à população. Para esta situação o modelo apropriado é caracterizado por

n_j : tamanho da j-ésima amostra ($1 \leq j \leq k$);

m_j : número de indivíduos marcados observados na j-ésima amostra ($m_1 = 0$ e $1 \leq j \leq k$);

M_j: número de indivíduos marcados na população no momento em que a j-ésima amostra é selecionada;

$M_1 = 0$, $M_2 = n_1$ e $M_j = \sum\limits_{r=1}^{j-1}(n_r - m_r)$, se $j = 3, 4, \cdots, k+1$.

Observemos que $M_{k+1} = \sum\limits_{r=1}^{k}(n_r - m_r)$ é o número de indivíduos distintos observados durante o experimento.

A função densidade de probabilidade conjunta de $(m_1,\ m_2, \cdots, m_k)$ condicionada a n_j, $1 \leq j \leq k, N$ é

$$P(0, m_2, m_3, \cdots, m_k | n_1, n_2, \cdots, n_k, N) = P(0, m_2, m_3, \cdots, m_k | \mathbf{n}, N)$$

$$= P(m_k|0, m_2, \cdots, m_{k-1}, \mathbf{n}, N) P(m_{k-1}|0, m_2, \cdots, m_{k-2}, \mathbf{n}, N) \cdots P(m_2|0, \mathbf{n}, N) P(0|\mathbf{n}, N)$$

$$= \frac{\binom{M_k}{m_k}\binom{N-M_k}{n_k-m_k}}{\binom{N}{n_k}} \frac{\binom{M_{k-1}}{m_{k-1}}\binom{N-M_{k-1}}{n_{k-1}-m_{k-1}}}{\binom{N}{n_{k-1}}} \cdots \frac{\binom{M_2}{m_2}\binom{N-M_2}{n_2-m_2}}{\binom{N}{n_2}} \frac{\binom{M_1}{m_1}\binom{N-M_1}{n_1-m_1}}{\binom{N}{n_1}}$$

$$= \prod_{i=1}^{k} \frac{M_i!(N-M_i)!}{\binom{N}{n_i} m_i!(n_i-m_i)!(M_i-m_i)!(N-M_i-n_i+m_i)!}$$

$$= \frac{(N-M_1)!}{(N-M_{k+1})!} \prod_{i=1}^{k} \frac{M_i!}{\binom{N}{n_i} m_i!(n_i-m_i)!(M_i-m_i)!}$$

$$= \frac{N!}{(N-M_{k+1})!} \prod_{i=1}^{k} \frac{M_i!}{\binom{N}{n_i} m_i!(n_i-m_i)!(M_i-m_i)!},$$

onde $0 \leq m_i \leq \min\{n_i, M_i\}$.

Teorema 30. O E.M.V. para N, no caso de amostragem direta múltipla sem reposição, é

$$\hat{N} = \begin{cases} M_{k+1} + N_0 - 1 & \text{se } M \leq M_{k+1} < \sum_{j=1}^{k} n_j \\ \infty & \text{se } M_{k+1} = \sum_{j=1}^{k} n_j, \end{cases}$$

onde $M = \max\{n_1, n_2, \cdots, n_k\}$,

e

$$N_0 = \min\{n \in N^* : \prod_{j=1}^{k} (M_{k+1} + n - n_j) < n(M_{k+1} + n)^{k-1}\}.$$

Além disso, a unicidade deste estimador é garantida exceto quando

$$\prod_{j=1}^{k}(M_{k+1} + N_0 - 1 - n_j) = (N_0 - 1)(M_{k+1} + N_0 - 1)^{k-1}. \tag{2.1.}$$

Neste caso os dois estimadores são $M_{k+1} + N_0 - 1$ *e* $M_{k+1} + N_0 - 2$.
Prova. A função de verossimilhança é

$$L(N|0, m_2, \cdots, m_k, \mathbf{n}) = \frac{N!}{(N - M_{k+1})!}\prod_{i=1}^{k}\frac{M_i!}{\binom{N}{n_i}m_i!(n_i - m_i)!(M_i - m_i)!}$$

$$\propto K(N) = \frac{N!}{(N - M_{k+1})!\prod_{j=1}^{k}\binom{N}{n_j}},$$

onde $N \geq M_{k+1}$ e $M \leq M_{k+1} \leq \sum_{j=1}^{k} n_j$.

A E.M.V. para N, $\hat{N}$, é solução da equação $\Delta \ln K(n) = \ln K(N) - \ln K(N-1) = 0$, ou equivalentemente, é solução da equação $K(N)/K(N-1) = 1$.
Como

$$\Delta \ln K(N) = \ln K(N) - \ln K(N-1) = 0$$

$$\Longleftrightarrow \ln N! - \ln(N - M_{k+1})! - \ln\prod_{i=1}^{k}\binom{N}{n_i} - \ln(N-1)! + \ln(N - M_{k+1} - 1)! + \ln\prod_{i=1}^{k}\binom{N-1}{n_i} = 0$$

$$\Longleftrightarrow \ln\left(\frac{N}{N - M_{k+1}}\right) + \ln\prod_{i=1}^{k}\left(\frac{N - n_i}{N}\right) = 0$$

$$\Longleftrightarrow \ln\prod_{i=1}^{k}\left(1 - \frac{n_i}{N}\right) = \ln\left(\frac{N - M_{k+1}}{N}\right)$$

$$\Longleftrightarrow \prod_{i=1}^{k}\left(1 - \frac{n_i}{N}\right) = 1 - \frac{M_{k+1}}{N}, \tag{2.2.}$$

segue que $\hat{N}$ é solução da equação acima. Contudo, Leite, J. G. et al. (1988) obtiveram $\hat{N}$ explicitamente. Mais precisamente:

$$\hat{N} = \begin{cases} M_{k+1} + N_0 - 1 & \text{se } M \leq M_{k+1} < \sum_{j=1}^{k} n_j \\ \infty & \text{se } M_{k+1} = \sum_{j=1}^{k} n_j, \end{cases}$$

onde

$$N_0 = \min\{n \in N^* : \prod_{j=1}^{k}(M_{k+1} + n - n_j) < n(M_{k+1} + n)^{k-1}\}.$$

Além disso, a unicidade desta estimativa é garantida exceto quando

$$\prod_{j=1}^{k}(M_{k+1} + N_0 - 1 - n_j) = (N_0 - 1)(M_{k+1} + N_0 - 1)^{k-1}.$$

Neste caso as duas estimativas são $M_{k+1} + N_0 - 1$ *e* $M_{k+1} + N_0 - 2$. ∎

Observemos que, se $M_{k+1} = \max\{n_1, n_2, \cdots, n_k\}$, supondo sem perda de generalidade que $n_1 = \max\{n_1, n_2, \cdots, n_k\}$, então a equação (2.2.) se expressa como

$$\prod_{i=1}^{k}\left(1 - \frac{n_i}{N}\right) = 1 - \frac{n_1}{N}.$$

A solução da equação acima é $\hat{N} = M_{k+1} = n_1 = \max\{n_1, n_2, \cdots, n_k\}$.

Observemos também que, se $M_{k+1} = \max\{n_1, n_2, \cdots, n_k\}$ então $N_0 = 1$ e o E.M.V. para N dado por Leite, J. G. et (1988), para este caso, é $\hat{N} = M_{k+1} = \max\{n_1, n_2, \cdots, n_k\}$.

Por outro lado, se $M_{k+1} = \sum_{j=1}^{k} n_j$, como $\prod_{j=1}^{k}(1 - x_j) > 1 - \sum_{j=1}^{k} x_j$, $\forall\ k \geq 2$, $\forall x_1, x_2, \cdots, x_k$ tal que $0 < x_j < 1$, $j = 1, 2, \cdots, k$, então a única solução da equação (2.2.) é $\hat{N} = \infty$.

Notemos que se $k = 2$ e $\max\{n_1, n_2\} < M_3 < n_1 + n_2$, então

$$\hat{N} = M_3 + N_0 - 1,$$

onde

$$N_0 = min\{n \in N^* : (M_3 + n - n_1)(M_3 + n - n_2) < n(M_3 + n)\}$$

$$= min\{n \in N^* : n(n_1 + n_2 - M_3) > M_3(M_3 - n_1\text{-}n_2) + n_1 n_2\}$$

$$= min\left\{n \in N^* : n > \frac{n_1 n_2}{n_1 + n_2 - M_3} - M_3\right\}$$

$$= \begin{cases} \frac{n_1 n_2}{n_1+n_2-M_3} - M_3 + 1 & \text{se } \frac{n_1 n_2}{n_1+n_2-M_3} \text{ for um número inteiro} \\ \left[\frac{n_1 n_2}{n_1+n_2-M_3}\right] - M_3 + 1 & \text{se } \frac{n_1 n_2}{n_1+n_2-M_3} \text{ não for um número inteiro.} \end{cases}$$

Por outro lado, se $\frac{n_1 n_2}{n_1+n_2-M_3}$ for um número inteiro, então

$$N_0 - 1 = \frac{n_1 n_2}{n_1 + n_2 - M_3} - M_3$$

satisfaz a relação (2.1.). Logo,

i) se $\frac{n_1 n_2}{n_1+n_2-M_3}$ for um número inteiro, então os E.M.V. para N são

$$M_3 + N_0 - 1 = \frac{n_1 n_2}{n_1 + n_2 - M_3} = \frac{n_1 n_2}{m_2}$$

e

$$M_3 + N_0 - 2 = \frac{n_1 n_2}{m_2} - 1;$$

ii) se $\frac{n_1 n_2}{n_1+n_2-M_3}$ não for um número inteiro, então o E.M.V. para N é

$$M_3 + N_0 - 1 = \left[\frac{n_1 n_2}{n_1 + n_2 - M_3}\right] = \left[\frac{n_1 n_2}{m_2}\right],$$

resultado este obtido na seção 1.1.b.

2.2. AMOSTRAGEM INVERSA MÚLTIPLA SEM REPOSIÇÃO

Nesta seção estudaremos o método de amostragem inversa múltipla sem reposição, que é uma generalização do caso simples para mais de duas amostras. Este procedimento consiste em capturar, sem reposição e ao acaso, n_0 indivíduos, marcá-los e devolvê-los à população. Em seguida, em cada uma de k etapas($k \geq 1$), fixamos o número de indivíduos marcados (ou não marcados) que serão selecionados. Capturamos então, sem reposição, indivíduos da população até obter esse número de indivíduos marcados (ou não marcados). O tamanho de cada amostra, n_i, é aleatório ($i = 1, 2, \cdots, k$). Os indivíduos não marcados são marcados e todos os indivíduos capturados são devolvidos à população.

a) Fixando o número de marcados.

Sejam

n_0: número de indivíduos capturados (sem reposição) pela primeira vez, marcados e devolvidos à população;

m_i: número fixado de indivíduos marcados capturados na i-ésima amostra ($m_0 = 0$).

n_i: tamanho (aleatório) da i-ésima amostra selecionada sem reposição até obter m_i indivíduos marcados ($i = 1, 2, \cdots, k$);

M_i: número de indivíduos marcados na população no momento em que a i-ésima amostra é selecionada, isto é, $M_i = \sum_{r=0}^{i-1}(n_r - m_r)$, $M_0 = 0$ e $i = 1, 2, \cdots, k+1$.

Observemos que $M_{k+1} = \sum_{r=0}^{k}(n_r - m_r)$ é o número de indivíduos distintos observados durante o experimento.

Neste modelo

i) $M_{i+1} - M_i = n_i - m_i$.

ii) n_0, N e $\mathbf{m} = (m_1, m_2, \cdots, m_k)$ são parâmetros.

iii) $M_2, M_3, \cdots, M_k; n_1, n_2, \cdots, n_k$ são variáveis aleatórias.

Notando que, para cada i ($i = 1, 2, \cdots, k$) o evento "n_i elementos são selecionados da população, sem reposição, até que ocorra o m_i-ésimo elemento marcado " é o evento "$(m_i - 1)$ elementos marcados são selecionados da população, sem reposição, nas $(n_i - 1)$ primeiras seleções e na n_i-ésima seleção é selecionado um indivíduo marcado", segue que a função densidade de probabilidade conjunta de $(n_1, n_2, \cdots, n_k)$ dado $(n_0, m_1, m_2, \cdots, m_k, N)$ é dada por

$$P(n_1, n_2, \cdots, n_k \mid n_0, m_1, m_2, \cdots, m_k, N) = P(n_1, n_2, \cdots, n_k \mid n_0, \mathbf{m}, N)$$

$$= P(n_k|n_0, n_1, n_2, \cdots, n_{k-1}, \mathbf{m}, N)P(n_{k-1}|n_0, n_1, \cdots, n_{k-2}, \mathbf{m}, N) \cdots P(n_1|n_0, \mathbf{m}, N)$$

$$= \prod_{i=1}^{k} \frac{\binom{M_i}{m_i-1}\binom{N-M_i}{n_i-m_i}(M_i - m_i + 1)}{\binom{N}{n_i-1}(N - n_i + 1)}$$

$$= \prod_{i=1}^{k} \frac{(n_i - 1)!M_i!(N - M_i)!(N - n_i)!}{(m_i - 1)!(M_i - m_i)!(n_i - m_i)!N!(N - n_i - M_i + m_i)!}$$

$$= \prod_{i=1}^{k} \frac{M_i}{N} \frac{\binom{M_i-1}{m_i-1}\binom{N-M_i}{n_i-m_i}}{\binom{N-1}{n_i-1}},$$

onde $m_i \leq n_i \leq N - M_i + m_i$, para $i = 1, 2, \cdots, k$.

Observando a última expressão da distribuição acima, notamos que ela é igual ao produto das probabilidades dos eventos A_i = "um elemento marcado é selecionado, sem reposição, na primeira seleção e $(m_i - 1)$ elementos marcados são selecionados quando uma amostra de tamanho $(n_i - 1)$ é selecionada, sem reposição, da população, agora contendo $(M_i - 1)$ elementos marcados e $(N - M_i)$ elementos não marcados ", para $i = 1, 2, \cdots, k$. Notemos também que, para $k = 1$ estamos no caso simples e a função densidade de probabilidade de n_1 coincide com a função densidade de probabilidade de n_1 estudada na seção 1.5.b. O seguinte teorema nos fornece o E.M.V. para N.

Teorema 31. No processo de amostragem inversa múltipla sem reposição fixando o número de marcados, o E.M.V. para N é

$$\hat{N} = M_{k+1} + N_0 - 1, \tag{2.3.}$$

onde

$$N_0 = \min\{n \in N^* : \prod_{j=0}^{k}(M_{k+1} + n - n_j) < n(M_{k+1} + n)^k\}.$$

Além disso, a unicidade deste estimador é garantida exceto quando

$$\prod_{j=0}^{k}(M_{k+1} + N_0 - 1 - n_j) = (N_0 - 1)(M_{k+1} + N_0 - 1)^k. \tag{2.4.}$$

Neste caso os dois estimadores são

$$M_{k+1} + N_0 - 1 \quad e \quad M_{k+1} + N_0 - 2.$$

Prova. A função de verossimilhança é

$$L(N|n_0, n_1, \cdots n_k, \mathrm{m})$$

$$= \prod_{i=1}^{k} \frac{(n_i - 1)!M_i!(N - M_i)!(N - n_i)!}{(m_i - 1)!(M_i - m_i)!(n_i - m_i)!N!(N - n_i - M_i + m_i)!}$$

$$= \frac{(N - n_0)!}{(N - M_{k+1})!} \prod_{i=1}^{k} \frac{(n_i - 1)!M_i!(N - n_i)!}{(m_i - 1)!(M_i - m_i)!(n_i - m_i)!N!}$$

$$\propto \frac{(N - n_0)!}{(N - M_{k+1})!} \prod_{i=1}^{k} \frac{(N - n_i)!}{N!} \propto K(N),$$

onde

$$K(N) = \frac{N!}{(N - M_{k+1})! \prod_{j=0}^{k} \binom{N}{n_j}}, \quad N \geq M_{k+1}.$$

Como este kernel é exatamente aquele obtido por Leite, J. G. et al (1988), segue que a E.M.V. para N é

$$\hat{N} = \begin{cases} M_{k+1} + N_0 - 1 & \text{se } M_{k+1} < \sum_{j=0}^{k} n_j \\ \infty & \text{se } M_{k+1} = \sum_{j=0}^{k} n_j \end{cases},$$

onde

$$N_0 = \min\{n \in N^* : \prod_{j=0}^{k}(M_{k+1} + n - n_j) < n(M_{k+1} + n)^k\}.$$

Mas, como neste caso cada $m_i \geq 1$, $i = 1, 2, \cdots, k$, então

$$M_{k+1} = \sum_{r=0}^{k}(n_r - m_r) = \sum_{r=0}^{k} n_r - \sum_{r=1}^{k} m_r < \sum_{r=0}^{k} n_r,$$

o que implica que a estimativa de Máxima Verossimilhança para N não assume o valor ∞. Logo, a estimativa de Máxima Verossimilhança para N é

$$\hat{N} = M_{k+1} + N_0 - 1.$$

Além disso, a unicidade desta estimativa é garantida exceto quando

$$\prod_{j=0}^{k}(M_{k+1} + N_0 - 1 - n_j) = (N_0 - 1)(M_{k+1} + N_0 - 1)^k.$$

Neste caso as duas estimativas são

$$M_{k+1} + N_0 - 1 \quad e \quad M_{k+1} + N_0 - 2.$$

■

Notemos que se $k = 1$ estamos no caso simples (m_1 aqui toma o lugar de m do caso simples) e a E.M.V. para N é a mesma que aquela obtida na seção 1.5.b. De fato, fazendo $k = 1$ na fórmula (2.3.) temos $\hat{N} = M_2 + N_0 - 1$, onde

$$N_0 = \min\{n \in N^* : (M_2 + n - n_0)(M_2 + n - n_1) < n(M_2 + n)\}$$

$$= \min\{n \in N^* : n(n_0 + n_1 - M_2) > M_2(M_2 - n_0\text{-}n_1) + n_0 n_1\}$$

$$= \min\left\{n \in N^* : n > \frac{n_0 n_1}{n_0 + n_1 - M_2} - M_2\right\}$$

$$= \begin{cases} \frac{n_0 n_1}{n_0+n_1-M_2} - M_2 + 1 & \text{se } \frac{n_0 n_1}{n_0+n_1-M_2} \text{ for um número inteiro} \\ \left[\frac{n_0 n_1}{n_0+n_1-M_2}\right] - M_2 + 1 & \text{se } \frac{n_0 n_1}{n_0+n_1-M_2} \text{ não for um número inteiro.} \end{cases}$$

Por outro lado, se $\frac{n_0 n_1}{n_0+n_1-M_2}$ for um número inteiro, então,

$$N_0 - 1 = \frac{n_0 n_1}{n_0 + n_1 - M_2} - M_2$$

satisfaz a relação (2.4.). Logo,

i) se $\frac{n_0 n_1}{n_0+n_1-M_2}$ for um número inteiro, então as E.M.V. para N são,

$$M_2 + N_0 - 1 = \frac{n_0 n_1}{n_0 + n_1 - M_2} = \frac{n_0 n_1}{m_1}$$

e

$$M_2 + N_0 - 2 = \frac{n_0 n_1}{m_1} - 1.$$

ii) Se $\frac{n_0 n_1}{n_0 + n_1 - M_2}$ não for um número inteiro, então a E.M.V. para N é

$$M_2 + N_0 - 1 = \left[\frac{n_0 n_1}{n_0 + n_1 - M_2}\right] = \left[\frac{n_0 n_1}{m_1}\right].$$

A **Tabela 3** fornece, para alguns valores de k, n_i e m_i, as E.M.V. para N.

TABELA 3

k	n_j	m_j	M_{k+1}	E.M.V.
2	$n_0 = 10$			
	$n_1 = 5$	$m_1 = 2$		
	$n_2 = 10$	$m_2 = 5$	$M_3 = 18$	25
2	$n_0 = 10$			
	$n_1 = 5$	$m_1 = 2$		
	$n_2 = 10$	$m_2 = 1$	$M_3 = 22$	64
2	$n_0 = 20$			
	$n_1 = 15$	$m_1 = 15$		
	$n_2 = 16$	$m_2 = 16$	$M_3 = 20$	20
2	$n_0 = 30$			
	$n_1 = 30$	$m_1 = 3$		
	$n_2 = 50$	$m_2 = 5$	$M_3 = 102$	475
3	$n_0 = 40$			
	$n_1 = 60$	$m_1 = 3$		
	$n_2 = 80$	$m_2 = 5$		
	$n_3 = 90$	$m_3 = 4$	$M_4 = 258$	2.173
3	$n_0 = 40$			
	$n_1 = 60$	$m_1 = 10$		
	$n_2 = 80$	$m_2 = 6$		
	$n_3 = 90$	$m_3 = 8$	$M_4 = 246$	1.064

Chapman (1952) propôs o seguinte estimador para N:

$$\hat{N}_7 = \frac{1}{k} \sum_{i=1}^{k} \left[\frac{n_i(M_i + 1)}{m_i} - 1 \right].$$

Observemos que, se para cada i $(i = 1, 2, \cdots, k)$ igualarmos a proporção de marcados na população à proporção de marcados na i-ésima amostra, ou seja, $\frac{M_i}{N_i} = \frac{m_i}{n_i}$, obtemos, para cada i, o estimador de Petersen para N, na época da i-ésima amostragem:

$$\hat{N}_i = \frac{M_i n_i}{m_i}.$$

O estimador $\hat{N}_7$ é a média aritmética de estimadores de Petersen modificados.

TEOREMA 32. O estimador $\hat{N}_7$ é um estimador não viciado para N.

Prova.

Como n_i dado $(n_j, 0 \leq j < i) \sim$ Hipergeométrica Negativa com parâmetros m_i, n_i e N, então,

$$E(n_i | n_j, 0 \leq j < i) = \frac{m_i(N+1)}{M_i + 1}, \qquad (1 \leq i \leq k).$$

Logo,

$$E(\hat{N}_7) = E\left[\frac{1}{k}\sum_{i=1}^{k}\left(\frac{n_i(M_i+1)}{m_i} - 1\right)\right]$$

$$= \frac{1}{k}\sum_{i=1}^{k} E\left[E\left(n_i\frac{(M_i+1)}{m_i}|n_j, 0 \leq j < i)\right)\right] - 1$$

$$= \frac{1}{k}\sum_{i=1}^{k} E\left[\frac{(M_i+1)}{m_i}E(n_i|n_j, 0 \leq j < i)\right] - 1$$

$$= \frac{1}{k}\sum_{i=1}^{k} E\left[\frac{(M_i+1)}{m_i}\frac{m_i(N+1)}{(M_i+1)}\right] - 1 = N.$$

■

TEOREMA 33. A variância do estimador $\hat{N}_7$ é limitada superiormente por $\frac{N^2}{k^2}\sum_{i=1}^{k}\frac{1}{m_i}$.

Prova.

$$V\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\right) = E\left[V\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle| n_j, 0 \leq j \leq k-1\right)\right] +$$

$$+V\left[E\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle|n_j, 0\leq j\leq k-1\right)\right]. \tag{2.5.}$$

Como

$$V\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle|n_j, 0\leq j\leq k-1\right)$$

$$=V\left(\sum_{i=1}^{k\text{-}1}\frac{n_i(M_i+1)}{m_i}+\frac{n_k(M_k+1)}{m_k}\middle|n_j, 0\leq j\leq k\text{-}1\right)$$

$$=V\left(\frac{n_k(M_k+1)}{m_k}\middle|n_j, 0\leq j\leq k-1\right)=\frac{(M_k+1)^2}{m_k^2}V(n_k|n_j, 0\leq j\leq k-1)$$

$$=\frac{(M_k+1)^2}{m_k^2}\frac{(N+1)m_k(N-M_k)(M_k-m_k+1)}{(M_k+1)^2(M_k+2)}\leq\frac{N^2}{m_k},\quad \text{(ver a expressão 1.5.)}$$

segue que

$$E\left[V\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle|n_j, 0\leq j\leq k-1\right)\right]\leq\frac{N^2}{m_k}. \tag{2.6.}$$

Como

$$E\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle|n_j, 0\leq j\leq k-1\right)$$

$$=\sum_{i=1}^{k-1}\frac{n_i(M_i+1)}{m_i}+\frac{(M_k+1)}{m_k}E(n_k|n_j, 0\leq j\leq k-1)$$

$$=\sum_{i=1}^{k-1}\frac{n_i(M_i+1)}{m_i}+\frac{(M_k+1)m_k(N+1)}{m_k(M_k+1)}$$

$$=\sum_{i=1}^{k-1}\frac{n_i(M_i+1)}{m_i}+(N+1),$$

temos que

$$V\left[E\left(\sum_{i=1}^{k}\frac{n_i(M_i+1)}{m_i}\middle|n_j, 0\leq j\leq k-1\right)\right]$$

$$= V\left(\sum_{i=1}^{k-1} \frac{n_i(M_i+1)}{m_i}\right). \qquad (2.7.)$$

De (2.6.) e (2.7.) temos que

$$V\left(\sum_{i=1}^{k} \frac{n_i(M_i+1)}{m_i}\right) \leq \frac{N^2}{m_k} + V\left(\sum_{i=1}^{k-1} \frac{n_i(M_i+1)}{m_i}\right).$$

De forma análoga ao que fizemos em (2.5.) e condicionando, agora, em n_j, $0 \leq j \leq k-2$, na expressão à direita da última relação, temos:

$$V\left(\sum_{i=1}^{k-1} \frac{n_i(M_i+1)}{m_i}\right) \leq \frac{N^2}{m_{k-1}} + V\left(\sum_{i=1}^{k-2} \frac{n_i(M_i+1)}{m_i}\right).$$

Logo,

$$V\left(\sum_{i=1}^{k} \frac{n_i(M_i+1)}{m_i}\right) \leq \frac{N^2}{m_k} + \frac{N^2}{m_{k-1}} + V\left(\sum_{i=1}^{k-2} \frac{n_i(M_i+1)}{m_i}\right),$$

e assim sucessivamente, condicionando em $n_j, 0 \leq j \leq k-3, k-4, \cdots, 0$, resulta

$$V\left(\sum_{i=1}^{k} \frac{n_i(M_i+1)}{m_i}\right) \leq \frac{N^2}{m_k} + \frac{N^2}{m_{k-1}} + \cdots + \frac{N^2}{m_1} = N^2 \sum_{i=1}^{k} \frac{1}{m_i}.$$

Portanto,

$$V(\hat{N}_7) = V\left(\frac{1}{k}\sum_{i=1}^{k} \frac{n_i(M_i+1)}{m_i} - 1\right) = \frac{1}{k^2} V\left(\sum_{i=1}^{k} \frac{n_i(M_i+1)}{m_i}\right)$$

$$\leq \frac{N^2}{k^2} \sum_{i=1}^{k} \frac{1}{m_i}.$$

■

Observemos que

$$V(\hat{N}_7) \leq \frac{N^2}{k}.$$

É de interesse, neste procedimento de amostragem inversa múltipla, o cálculo da $E(n')$, onde $n' = \sum_{j=0}^{k} n_j$, pois tal esperança fornece o esforço médio gasto no experimento.

— Você sabe = Maria ~~Beltriz~~ Beltrãniz

— alem do horizonte

Daremos a seguir uma fórmula aproximada para $E(n_i)$, a partir da qual teremos uma idéia da $E(n')$, o número médio de indivíduos capturados em todo o experimento. Portanto, seja

$$g(x) = \frac{1}{x+1}.$$

Expandindo $g(x)$ pela fórmula de Taylor em torno do ponto a, $a \neq -1$, temos

$$g(x) = g(a) + g'(a)(x-a) + R_2(x),$$

onde o resto $R_2(x) = o(|x-a|^2) \xrightarrow[x \to a]{} 0$. Logo, podemos aproximar

$$g(x) \cong g(a) + g'(a)(x-a).$$

Para $x = M_i$, e $a = \mu_i = E(M_i)$, resulta,

$$g(M_i) = \frac{1}{M_i+1} \cong \frac{1}{\mu_i+1} - \frac{1}{(\mu_i+1)^2}(M_i - \mu_i).$$

Logo,

$$E\left(\frac{1}{M_i+1}\right) \cong \frac{1}{\mu_i+1}.$$

Usando o resultado anterior temos,

$$E(n_i) = E[E(n_i|n_j; 0 \leq j < i)] = E\left[\frac{m_i(N+1)}{M_i+1}\right]$$

$$= m_i(N+1)E\left(\frac{1}{M_i+1}\right) \cong \frac{m_i(N+1)}{E(M_i)+1}$$

$$= (m_i(N+1)) \Big/ \left(E\left[\sum_{j=0}^{i-1}(n_j - m_j)\right] + 1\right)$$

$$= (m_i(N+1)) \Big/ \left(\sum_{j=0}^{i-1}[E(n_j) - m_j] + 1\right).$$

Esta fórmula recursiva nos dá, uma vez fixados k e os m_i, uma idéia do número médio de seleções necessárias para obter $m_1 + m_2 + \cdots + m_k$ marcados.

A **Tabela 4** fornece para dois tamanhos distintos de população ($N = 5000$ e $N = 10000$), e diferentes valores de k e m_i, o máximo valor de $\sigma_{\hat{N}_7}/N$ e o valor de $E\left(\sum_{j=0}^{k} n_j\right)$.

TABELA 4

Limite superior do quociente entre o desvio padrão de $\hat{N}_7$ e N, e número médio de capturas (Amostragem Inversa Múltipla Sem Reposição Fixando o Número de Marcados).

k	m_1	m_2	m_3	m_4	m_5	n_0	$\sigma_{\hat{N}_7/N} \leq \frac{1}{k}\sqrt{\sum_{i=1}^{k} \frac{1}{m_i}}$	$N = 5.000$ $E(n')$	$N = 10.000$ $E(n')$
1	10					300	0.316	467	633
2	20	20				300	0.158	796	1177
3	10	15	15			400	0.161	787	1059
4	10	15	20	15		500	0.133	461	1280
5	15	15	15	15	15	388	0.115	1008	1394
5	25	20	15	10	5	500	0.135	1058	1452

b) Fixando o número de não marcados

Nesta seção estudaremos o caso de amostragem inversa múltipla sem reposição fixando o número de não marcados. Este procedimento consiste em capturarmos, sem reposição e ao acaso, n_0 indivíduos, marcá-los e devolvê-los à população. Em seguida selecionam-se k amostras ($k \geq 1$) em cada uma das quais é fixado o número δ_i de indivíduos não marcados que serão selecionados. Depois de capturar os δ_i indivíduos não marcados eles são marcados e todos os indivíduos da amostra são devolvidos à população.

Sejam

n_0 : número de indivíduos capturados (sem reposição) pela primeira vez, marcados e devolvidos à população;

δ_i: número fixado de indivíduos não marcados capturados na i-ésima amostra ($i = 1, 2, \cdots, k$);

n_i: tamanho (aleatório) da i-ésima amostra selecionada, sem reposição, até obter δ_i indivíduos não marcados ($i = 1, 2, \cdots, k$);

m_i: número de indivíduos marcados capturados na i-ésima amostra ($i = 1, 2, \cdots, k$) e

M_i: número de indivíduos marcados na população no momento em que a i-ésima amostra é selecionada, isto é,

$M_i = \sum_{j=0}^{i-1} \delta_j$, onde $\delta_0 = n_0$ e ($i = 1, 2, \cdots, k+1$).

Observemos que $M_{k+1} = \sum_{r=0}^{k} \delta_r$ é o número de indivíduos distintos observados durante o experimento.

Neste modelo

i) $M_{i+1} - M_i = \delta_i$ ($i = 1, 2, \cdots, k$).

ii) $\mathbf{m} = m_1, m_2, \cdots, m_k$ e $n_1, n_2, \cdots, n_k$ são variáveis aleatórias; n_0 e $\delta = \delta_1, \delta_2, \cdots, \delta_k$ são parâmetros.

Como δ_i é um parâmetro em cada etapa do processo de selecionar as amostras, e portanto M_i também o é, há independência entre as sucessivas variáveis aleatórias n_i. Notando que para cada i ($i = 1, 2, \cdots, k$) o evento "n_i elementos são selecionados sem reposição da população até que ocorra o δ_i-ésimo elemento não marcado " é o evento "$(\delta_i - 1)$ elementos não marcados são selecionados sem reposição da população nas $(n_i - 1)$ primeiras seleções e na n_i-ésima é selecionado um indivíduo não marcado ", segue que a função densidade de probabilidade conjunta de $(n_1, n_2, \cdots, n_k)$ condicionada a (δ, n_0, N) é

$$P(n_1, n_2, \cdots, n_k | \delta, n_0, N) = \prod_{i=1}^{k} P(n_i | \delta, n_0, N)$$

$$= \prod_{i=1}^{k} \frac{\binom{N-M_i}{\delta_i - 1}\binom{M_i}{n_i - \delta_i}}{\binom{N}{n_i - 1}} \frac{[N - M_i - (\delta_i - 1)]}{[N - (n_i - 1)]}$$

$$= \prod_{i=1}^{k} \frac{(N - M_i)!(n_i - 1)!(N - n_i + 1)![N - M_i - (\delta_i - 1)]\binom{M_i}{n_i - \delta_i}}{(\delta_i - 1)!(N - M_i - \delta_i + 1)!N![N - (n_i - 1)]}$$

$$= \prod_{i=1}^{k} \frac{(N - M_i)!(n_i - 1)!(N - n_i)!\binom{M_i}{n_i - \delta_i}}{(\delta_i - 1)!(N - M_i - \delta_i)!N!}$$

$$= \prod_{i=1}^{k} \frac{(N - M_i)}{N} \frac{\binom{N-M_i-1}{\delta_i-1}\binom{M_i}{n_i-\delta_i}}{\binom{N-1}{n_i-1}},$$

onde $\delta_i \leq n_i \leq M_i + \delta_i$.

Observando a última expressão da distribuição acima, notamos que ela é igual ao produto das probabilidades dos eventos A_i = "um elemento não marcado é selecionado, sem reposição, na primeira seleção e $(\delta_i - 1)$ elementos não marcados são selecionados quando uma amostra de tamanho $(n_i - 1)$ é selecionada, sem reposição, da população, agora contendo $(N - M_i - 1)$ elementos não marcados e M_i elementos marcados".

Notemos que para $k = 1$ a função densidade de probabilidade de n_1 dado (δ_1, n_0, N) coincide com a função densidade de probabilidade de n_1 dado (δ, n_0, N) estudada no caso simples na seção 1.6.b.

O seguinte teorema nos da a E.M.V. para N.

Teorema 34. A E.M.V. para N no caso de amostragem inversa múltipla sem reposição fixando o número de não marcados,é

$$\hat{N} = \begin{cases} M_{k+1} + N_0 - 1 & \text{se } M_{k+1} < \sum_{j=0}^{k} n_j \\ \infty & \text{se } M_{k+1} = \sum_{j=0}^{k} n_j, \end{cases} \qquad (2.8.)$$

onde

$$N_0 = \min\{n \in N^* : \prod_{j=0}^{k}(M_{k+1} + n - n_j) < n(M_{k+1} + n)^k\}.$$

Além disso, a unicidade desta estimativa é garantida exceto quando

$$\prod_{j=0}^{k}(M_{k+1} + N_0 - 1 - n_j) = (N_0 - 1)(M_{k+1} + N_0 - 1)^k. \qquad (2.9.)$$

Neste caso as duas estimativas são

$$M_{k+1} + N_0 - 1 \quad \text{e} \quad M_{k+1} + N_0 - 2.$$

Prova. A função de verossimilhança é

$$L(N|n_0, n_1, n_2, \cdots, n_k, \delta) =$$

$$= \prod_{i=1}^{k} \frac{(N-M_i)!M_i!(n_i-1)!(N-n_i)!}{(\delta_i-1)!((N-M_i-M_{i+1}+M_i)!(n_i-\delta_i)!(M_i-n_i+\delta_i)!N!}$$

$$= \frac{(N-M_1)!}{(N-M_{k+1})!}\left[\prod_{i=1}^{k} \frac{M_i!(n_i-1)!(N-n_i)!}{(\delta_i-1)!(n_i-\delta_i)!(M_i-n_i+\delta_i)!N!}\right]$$

$$= \frac{(N-n_0)!}{(N-M_{k+1})!}\left[\prod_{i=1}^{k} \frac{M_i!(n_i-1)!(N-n_i)!}{(\delta_i-1)!(n_i-\delta_i)!(M_i-n_i+\delta_i)!N!}\right]$$

$$\propto \frac{(N-n_0)!}{(N-M_{k+1})!}\prod_{i=1}^{k}\frac{(N-n_i)!}{N!} \propto K(N),$$

onde

$$K(N) = \frac{N!}{(N-M_{k+1})!\prod\limits_{j=0}^{k}\binom{N}{n_j}}, \qquad N \geq M_{k+1}.$$

Como $K(N)$ é igual ao kernel da verossimilhança obtido por Leite, J. G. et al (1988), segue que a E.M.V. para N é a mesma que aquela obtida no mencionado artigo, isto é,

$$\hat{N} = \begin{cases} M_{k+1}+N_0-1 & \text{se } M_{k+1} < \sum\limits_{j=0}^{k} n_j \\ \infty & \text{se } M_{k+1} = \sum\limits_{j=0}^{k} n_j, \end{cases}$$

onde

$$N_0 = \min\{n \in N^* : \prod_{j=0}^{k}(M_{k+1}+n-n_j) < n(M_{k+1}+n)^k\}.$$

Além disso, a unicidade desta estimativa é garantida exceto quando

$$\prod_{j=0}^{k}(M_{k+1}+N_0-1-n_j) = (N_0-1)(M_{k+1}+N_0-1)^k.$$

Neste caso as duas estimativas são

$$M_{k+1}+N_0-1 \quad e \quad M_{k+1}+N_0-2.$$

■

Notamos que, se $k = 1$ e $M_2 = n_0 + \delta_1 < n_0 + n_1$, então, fazendo $k = 1$ em (2.8.), obtemos a E.M.V. para N que coincide com aquela obtida para o caso simples (ver seção 1.6.b). Observemos que δ_1 é igual a δ do caso simples. De forma análoga à seção anterior, temos que a E.M.V. para N é

$$\hat{N} = M_2 + N_0 - 1,$$

é não viciado?

onde

$$N_0 = \min\{n \in N^* : (M_2 + n - n_0)(M_2 + n - n_1) < n(M_2 + n)\}$$

$$= \begin{cases} \frac{n_0 n_1}{n_0+n_1-M_2} - M_2 + 1 & \text{se } \frac{n_0 n_1}{n_0+n_1-M_2} \text{ for um número inteiro} \\ \left[\frac{n_0 n_1}{n_0+n_1-M_2}\right] - M_2 + 1 & \text{se } \frac{n_0 n_1}{n_0+n_1-M_2} \text{ não for um número inteiro.} \end{cases}$$

Por outro lado, se $\frac{n_0 n_1}{n_0+n_1-M_2}$ for um número inteiro, então

$$N_0 - 1 = \frac{n_0 n_1}{n_0 + n_1 - M_2} - M_2$$

satisfaz a relação (2.9.). Logo,

i) se $\frac{n_0 n_1}{n_0+n_1-M_2}$ for um número inteiro, então as E.M.V. para N são,

$$M_2 + N_0 - 1 = \frac{n_0 n_1}{n_0 + n_1 - M_2} = \frac{n_0 n_1}{n_1 - \delta_1},$$

e

$$M_2 + N_0 - 2 = \frac{n_0 n_1}{n_1 - \delta_1} - 1.$$

ii) se $\frac{n_0 n_1}{n_0+n_1-M_2}$ não for um número inteiro, então a E.M.V. para N é

$$M_2 + N_0 - 1 = \left[\frac{n_0 n_1}{n_0 + n_1 - M_2}\right] = \left[\frac{n_0 n_1}{n_1 - \delta_1}\right].$$

Exemplo 4. Sejam $n_0 = 50$, $n_1 = 40$, $n_2 = 50$, $\delta_1 = 10$ e $\delta_2 = 20$, então $M_1 = n_0 = 50$, $M_2 = n_0 + \delta_1 = 60$ e $M_3 = n_0 + \delta_1 + \delta_2 = 80$. Logo, para $k = 1$ a E.M.V. para N é $\hat{N} = 66$ e para $k = 2$, $N_0 = 10$ e a E.M.V. para N é $\hat{N} = 89$.

Chapman(1952) propôs o seguinte estimador para N:

$$\hat{N}_8 = \frac{1}{k} \sum_{i=1}^{k} \frac{n_i(M_i + 1)}{n_i - \delta_i + 1} - 1.$$

Teorema 35. A esperança de N_6 é dada por:

$$E(\hat{N}_8) = N - \frac{(N+1)}{k}\sum_{i=1}^{k}\left(1 - \frac{\delta_i}{(N+1)}\right)\prod_{j=0}^{M_i-1}\left(1 - \frac{\delta_i}{N-j}\right).$$

Prova. A relação 1.7. da seção 1.7.b implica que, $\forall\, i, i = 1, 2, \cdots, k$,

$$E\left(\frac{n_i}{n_i - \delta_i + 1}\right) = \frac{(N+1)}{(M_i+1)}\left[1 - \left(1 - \frac{\delta_i}{N+1}\right)\prod_{j=0}^{M_i-1}\left(1 - \frac{\delta_i}{N-j}\right)\right]$$

Então,

$$E(\hat{N}_8) = E\left[\frac{1}{k}\sum_{i=1}^{k}\frac{n_i(M_i+1)}{n_i - \delta_i + 1} - 1\right] =$$

$$= \frac{1}{k}\sum_{i=1}^{k}(M_i+1)E\left(\frac{n_i}{n_i - \delta_i + 1}\right) - 1$$

$$= \frac{1}{k}\sum_{i=1}^{k}(M_i+1)\frac{(N+1)}{(M_i+1)}\left[1 - \left(1 - \frac{\delta_i}{N+1}\right)\prod_{j=0}^{M_i-1}\left(1 - \frac{\delta_i}{N-j}\right)\right] - 1$$

$$= N - \frac{(N+1)}{k}\sum_{i=1}^{k}\left(1 - \frac{\delta_i}{N+1}\right)\prod_{j=0}^{M_i-1}\left(1 - \frac{\delta_i}{N-j}\right).$$

■

O vício de $\hat{N}_8$ é dado pelo seguinte corolário do teorema anterior:

Corolário 36. Seja $\delta' = \min\{\delta_i : 1 \le i \le k\}$. Para todo K, $0 < K < N$, se $\frac{n_0\delta'}{N} \ge \ln(\frac{N}{K})$, então,

$$|E(\hat{N}_8 - N| < K.$$

Prova. Pelo Teorema anterior,

$$|E(\hat{N}_8) - N| = \frac{(N+1)}{k}\sum_{i=1}^{k}\left(1 - \frac{\delta_i}{N+1}\right)\prod_{j=0}^{M_i-1}\left(1 - \frac{\delta_i}{N-j}\right)$$

$$\le \frac{(N+1)}{k}\sum_{i=1}^{k}\left(1 - \frac{\delta'}{N+1}\right)\left(1 - \frac{\delta'}{N}\right)^{M_i}$$

$$< \frac{(N+1)}{k}k\left(1-\frac{\delta'}{N+1}\right)\left(1-\frac{\delta'}{N}\right)^{n_0}$$

$$= (N+1-\delta')\left(1-\frac{\delta'}{N}\right)^{n_0} \leq N\left(1-\frac{\delta'}{N}\right)^{n_0}$$

$$\leq Ne^{-\frac{n_0\delta'}{N}} \leq Ne^{\ln\frac{K}{N}} = K.$$

■

O seguinte teorema é útil para o cálculo da variância de $\hat{N}_8$.

Teorema 37. Sejam

$$n_{(i)} = \begin{cases} n(n-1)(n-2)\cdots(n-i+1) & \text{se } i \geq 1, \\ 1 & \text{se i=0;} \end{cases}$$

$$q_{ij}^{(r)} = \frac{(N+i)_{(i)}(N-M_r)_{(j-i)}}{(M_r+j)_{(j)}(\delta_r-1)_{(j-i)}} \text{ e}$$

$$\eta_{ij}^{(r)} = \sum_{l=0}^{j-1} \frac{\binom{\delta_r+l-j+i-1}{l}\binom{N-\delta_r-l+j}{M_r-l+j}}{\binom{N+i}{M_r+j}}$$

para $i=1,2$; $j=2,3,4$; $r=1,2,\cdots,k$ e δ' definido como no Corolário 36.

Se $\delta' \geq 2$ e $n_0\delta' \geq N\ln N$, então

$$V\left[\frac{(M_r+1)n_r}{n_r-\delta_r+1}\right] \cong (M_r+1)^2[q_{22}^{(r)}(1-\eta_{22}^{(r)})+q_{23}^{(r)}(1-\eta_{23}^{(r)})+2q_{24}^{(r)}(1-\eta_{24}^{(r)})-q_{12}^{(r)}(1-\eta_{12}^{(r)})-$$

$$-q_{13}^{(r)}(1-\eta_{13}^{(r)})-2q_{14}^{(r)}(1-\eta_{14}^{(r)})]-2N-1-N^2, \qquad 1 \leq r \leq k$$

Prova. Analogamente ao caso simples (ver seção 1.7.b), para cada $r=1,2,\cdots,k$, temos:

$$P(n_r|\delta_r,N,n_0) = \frac{\binom{N-M_r}{\delta_r-1}\binom{M_r}{n_r-\delta_r}}{\binom{N}{n_r-1}}\frac{[N-M_r-(\delta_r-1)]}{[N-(n_r-1)]}$$

e

$$E\left[\frac{(n_r-i-1)_{(i)}}{(n_r-\delta_r+j)_{(j)}}\right] = q_{ij}^{(r)}(1-\eta_{ij}^{(r)}),$$

para $i=1,2$ e $j=2,3,4$.

Usando o resultado do corolário anterior para $K = 1$, e analogamente ao caso simples,temos,

$$V\left[\frac{n_r(M_r+1)}{n_r-\delta_r+1}\right] \cong (M_r+1)^2[q_{22}^{(r)}(1-\eta_{22}^{(r)})+q_{23}^{(r)}(1-\eta_{23}^{(r)})+2q_{24}^{(r)}(1-\eta_{24}^{(r)})-q_{12}^{(r)}(1-\eta_{12}^{(r)})-$$

$$-q_{13}^{(r)}(1-\eta_{13}^{(r)})-2q_{14}^{(r)}(1-\eta_{14}^{(r)})-2N-1-N^2, \qquad 1 \leq r \leq k.$$

■

Teorema 38. Sejam $q_{ij}^{(r)}$, $\eta_{ij}^{(r)}$ $1 \leq r \leq k$, $1 \leq i \leq 2$, $2 \leq j \leq 4$, e δ' definido como no teorema anterior. Se $\delta' \geq 2$ e $n_0\delta' \geq N \ln N$, então

$$V(\hat{N}_8) \cong \frac{1}{k^2}\sum_{i=1}^{k}\Big[(M_r+1)^2[q_{22}^{(r)}(1-\eta_{22}^{(r)})+q_{23}^{(r)}(1-\eta_{23}^{(r)})+2q_{24}^{(r)}(1-\eta_{24}^{(r)})-q_{12}^{(r)}(1-\eta_{12}^{(r)})-$$

$$-q_{13}^{(r)}(1-\eta_{13}^{(r)})-2q_{14}^{(r)}(1-\eta_{14}^{(r)})]-2N-1-N^2\Big].$$

Prova. Usando o fato de que $n_1, n_2, \cdots, n_k$ são independentes e o resultado do Teorema 37, temos:

$$V(\hat{N}_8) = V\left(\frac{1}{k}\sum_{r=1}^{k}\frac{n_r(M_r+1)}{n_r-\delta_r+1}-1\right) = \frac{1}{k^2}\sum_{r=1}^{k}V\left(\frac{n_r(M_r+1)}{n_r-\delta_r+1}\right)$$

$$\cong \frac{1}{k^2}\sum_{r=1}^{k}\Big[(M_r+1)^2[q_{22}^{(r)}(1-\eta_{22}^{(r)})+q_{23}^{(r)}(1-\eta_{23}^{(r)})+2q_{24}^{(r)}(1-\eta_{24}^{(r)})-q_{12}^{(r)}(1-\eta_{12}^{(r)})-$$

$$-q_{13}^{(r)}(1-\eta_{13}^{(r)})-2q_{14}^{(r)}(1-\eta_{14}^{(r)})]-2N-1-N^2\Big].$$

■

CAPÍTULO 3

MODELOS DE CAPTURA-RECAPTURA CONSIDERANDO-SE A PROBABILIDADE DE CAPTURA DE CADA INDIVÍDUO

3.1. Probabilidades de captura variando de amostra para amostra

Nesta seção estudaremos um modelo de captura-recaptura baseado nos dados de k amostras ($k \geq 2$), selecionadas de uma população fechada. Cada membro da população tem uma probabilidade p_i ($0 < p_i < 1$) de ser capturado na i-ésima amostra ($i = 1, 2, \cdots, k$). Os parâmetros p_i são desconhecidos e cada indivíduo é capturado ou não em cada amostra, independentemente dos demais. Como antes, todos os membros da população não marcados selecionados, em cada amostra, são marcados antes de serem devolvidos à população. Os marcados são simplesmente devolvidos à população.

Sejam

n_i : tamanho da i-ésima amostra ($i = 1, 2, \cdots, k$);

m_i : número de indivíduos marcados observados na i-ésima amostra ($i = 1, 2, \cdots, k$; $m_1 = 0$);

M_i : número de indivíduos marcados na população no momento da seleção da i-ésima amostra ($M_1 = 0$ e $M_2 = n_1$).

Então, $M_i = \sum\limits_{j=1}^{i-1}(n_j - m_j)$, $2 \leq i \leq k+1$, onde

$M_{k+1} = \sum_{j=1}^{k}(n_j - m_j)$: número total de indivíduos distintos capturados durante o experimento.

A função densidade de probabilidade conjunta de $(m_1, m_2, \cdots, m_k; n_1, n_2, \cdots, n_k)$ dados $p_j, 1 \leq j \leq k$ e N é dada por:

$$P(0, m_2, \cdots, m_k; n_1, n_2, \cdots, n_k | p_j, 1 \leq j \leq k, N)$$

$$= P(m_k, n_k | m_{k-1}, n_{k-1}, \cdots, m_2, n_2, 0, n_1; p_j, 1 \leq j \leq k, N) \times$$

$$\times P(m_{k-1}, n_{k-1} | m_{k-2}, n_{k-2}, \cdots, m_2, n_2, 0, n_1; p_j, 1 \leq j \leq k, N) \times$$

$\mu_1 = m_2 - m_1$

$\mu_2 = n_2 - m_2$

$$\times \cdots \times P(m_2, n_2 | 0, n_1; p_j, 1 \leq j \leq k, N) P(0, n_1 | p_j, 1 \leq j \leq k, N)$$ $m_2 = n_2 - \mu_2$

$$= \binom{M_k}{m_k} {p_k}^{m_k} (1 - p_k)^{M_k - m_k} \binom{N - M_k}{n_k - m_k} {p_k}^{n_k - m_k} (1 - p_k)^{N - M_k - n_k + m_k} \times$$

$P(A,B) = P(A|B) \cdot P(B)$

$P(n_2|n_1)$

$$\times \binom{M_{k-1}}{m_{k-1}} {p_{k-1}}^{m_{k-1}} (1 - p_{k-1})^{M_{k-1} - m_{k-1}} \binom{N - M_{k-1}}{n_{k-1} - m_{k-1}} p_{k-1}^{n_{k-1} - m_{k-1}} \times$$

$$\times (1 - p_{k-1})^{N - M_{k-1} - n_{k-1} + m_{k-1}} \cdots \binom{M_2}{m_2} {p_2}^{m_2} (1 - p_2)^{M_2 - m_2} \binom{N - M_2}{n_2 - m_2} {p_2}^{n_2 - m_2} \times$$

$$\times (1 - p_2)^{N - M_2 - n_2 + m_2} \binom{N}{n_1} {p_1}^{n_1} (1 - p_1)^{N - n_1} = \prod_{j=1}^{k} \binom{M_j}{m_j} \binom{N - M_j}{n_j - m_j} {p_j}^{n_j} (1 - p_j)^{N - n_j}$$

$$= \frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} \binom{M_j}{m_j} \frac{{p_j}^{n_j} (1 - p_j)^{N - n_j}}{(n_j - m_j)!},$$

onde $0 \leq n_i \leq N$ e $\max\{0; M_i + n_i - N\} \leq m_i \leq \min\{n_i, M_i\}$, $i = 1, 2, \cdots, k$.

$(N - M_i)!$

$(N - M_i + n_j + m_j)!$

$N - M_i + n_j + m_j$

A função de verossimilhança é dada por

$$L(p_1, p_2, \cdots, p_k, N | m_1, m_2, \cdots, m_k; n_1, n_2, \cdots, n_k)$$

$$= \frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} \binom{M_j}{m_j} \frac{{p_j}^{n_j}(1-p_j)^{N-n_j}}{(n_j - m_j)!} \propto$$

$$\propto K(p_1, p_2, \cdots, p_k, N)$$

onde

$$K(p_1, p_2, \cdots, p_k, N) = \frac{N!}{(N - M_{k+1})!} \prod_{j=1}^{k} {p_j}^{n_j}(1-p_j)^{N-n_j}, \tag{3.1.}$$

$0 < p_j < 1$ $(j = 1, 2, \cdots, k)$ e $N \geq M_{k+1}$.

O seguinte teorema nos dá os E.M.V. para p_j, $1 \leq j \leq k$, e N.

Teorema 38. Os E.M.V. para p_j são

$$\hat{p}_j = \frac{n_j}{\hat{N}}, \quad 1 \leq j \leq k,$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\prod_{j=1}^{k} \left(1 - \frac{n_j}{N}\right) = 1 - \frac{M_{k+1}}{N}.$$

Prova. De (3.1.) segue que

$$\ln K(p_1, p_2, \cdots, p_k, N) = \ln N! - \ln(N - M_{k+1})! + \sum_{j=1}^{k} n_j \ln p_j + \sum_{j=1}^{k} (N - n_j) \ln(1 - p_j).$$

Então

$$\frac{\partial \ln p(p_1, p_2, \cdots, p_k)}{\partial p_j} = \frac{n_i}{p_j} - \frac{N - n_j}{1 - p_j} = 0$$

$$\Longleftrightarrow n_j - n_j p_j - N p_j + n_j p_j = 0$$

$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B)}$

$$\Longleftrightarrow p_j = \frac{n_j}{N}, \ 1 \leq j \leq k. \tag{3.2.}$$

Com relação ao E.M.V. para N temos:

$$\Delta \ln K(p_1, p_2, \cdots, p_k, N) = 0 \Longleftrightarrow \ln \frac{K(p_1, p_2, \cdots, p_k, N)}{K(p_1, p_2, \cdots, p_k, N-1)} = 0$$

$$\Longleftrightarrow \ln \frac{N!/(N-M_{k+1})! \prod_{j=1}^{k} p_j^{n_j}(1-p_j)^{N-n_j}}{(N-1)!/(N-M_{k+1}-1)! \prod_{j=1}^{k} p_j^{n_j}(1-p_j)^{N-n_j-1}} = 0$$

$$\Longleftrightarrow ln\left[\frac{N}{N-M_{k+1}} \prod_{j=1}^{k}(1-p_j)\right] = 0$$

$$\Longleftrightarrow \prod_{j=1}^{k}(1-p_j) = 1 - \frac{M_{k+1}}{N}. \tag{3.3.}$$

Logo, de (3.2.) e (3.3.), segue que as E.M.V. de p_j, $j = 1, 2, \cdots, k$, são

$$\hat{p}_j = \frac{n_j}{\hat{N}}$$

e $\hat{N}$ é solução da equação

$$\prod_{j=1}^{k}\left(1 - \frac{n_j}{N}\right) = 1 - \frac{M_{k+1}}{N}.$$

■

3.2. Probabilidades de captura iguais para todas as amostras

Nesta seção vamos considerar o caso em que, em cada amostra, cada indivíduo da população, independentemente dos demais, tem uma mesma probabilidade p de ser capturado.

Fazendo $p_1 = p_2 = \cdots = p_k = p$ em (3.1.) o kernel da função de Máxima Verossimilhança é

$$K(p,N) = \frac{N!}{(N-M_{k+1})!}\, p^{\sum\limits_{j=1}^{k} n_j} (1-p)^{kN-\sum\limits_{j=1}^{k} n_j},$$

onde $0 < p < 1$ e $N \geq M_{k+1}$.

O seguinte teorema nos dá os E.M.V. para p e N.

Teorema 39. O E.M.V. para p é

$$\hat{p} = \frac{\sum\limits_{j=1}^{k} n_j}{k\hat{N}}$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\left(1 - \frac{\sum\limits_{j=1}^{k} n_j}{kN}\right)^k = 1 - \frac{M_{k+1}}{N}.$$

Prova De modo análogo ao caso anterior, temos

$$\ln K(p,N) = \ln N! - \ln(N-M_{k+1})! + \sum_{j=1}^{k} n_j \ln p + (kN - \sum_{j=1}^{k} n_j)\ln(1-p).$$

Então

$$\frac{\partial \ln K(p,N)}{\partial p} = \frac{\sum\limits_{j=1}^{k} n_j}{p} - \frac{kN - \sum\limits_{j=1}^{k} n_j}{1-p} = 0$$

$$\Longleftrightarrow \sum_{j=1}^{k} n_j - p\sum_{j=1}^{k} n_j - kNp + p\sum_{j=1}^{k} n_j = 0$$

$$\Longleftrightarrow p = \frac{\sum\limits_{j=1}^{k} n_j}{kN} \qquad (3.4.)$$

e

$$\Delta K(p,N) = 0 \iff \ln\frac{K(p,N)}{K(p,N-1)} = 0$$

$$\iff \ln \frac{N!\Big/(N-M_{k+1})!\; p^{\sum\limits_{j=1}^{k} n_j}\,(1-p)^{kN-\sum\limits_{j=1}^{k} n_j}}{(N-1)!\Big/(N-M_{k+1}-1)!\; p^{\sum\limits_{j=1}^{k} n_j}\,(1-p)^{kN-k-\sum\limits_{j=1}^{k} n_j}} = 0$$

$$\iff \ln\left[\frac{N}{(N-M_{k+1})}(1-p)^k\right] = 0$$

$$\iff (1-p)^k = 1 - \frac{M_{k+1}}{N}. \qquad (3.5.)$$

Logo, de (3.4.) e (3.5.) segue que a E.M.V. para p é

$$\hat{p} = \frac{\sum\limits_{j=1}^{k} n_j}{k\hat{N}}$$

e $\hat{N}$ é solução da equação

$$\left(1 - \frac{\sum\limits_{j=1}^{k} n_j}{kN}\right)^k = 1 - \frac{M_{k+1}}{N}.$$

■

3.3. Probabilidades de recaptura dependendo da captura.

Nesta seção estudaremos o caso em que a captura inicial altera a probabilidade de captura em ocasiões posteriores. Isto é, cada animal apresenta um comportamento em relação à captura, convertendo-se em "viciado em captura" ou "fugidio devido à captura". Vários autores estudaram este caso. Dentre eles podemos citar Overton e Davis (1969), Geis (1952), Tanaka (1956, 1963), Flyger (1959), Bailey (1968) e Pucek (1969). Supomos que, nos momentos em que são selecionadas as amostras, cada animal da população não marcado tem uma probabilidade p $(0 < p < 1)$ de ser capturado e cada animal marcado tem uma probabilidade c $(0 < c < 1)$ de ser recapturado, isto é, as probabilidades de captura e recaptura não variam de amostra para amostra. Em cada amostra os animais não marcados são marcados e, em seguida, todos os animais selecionados são devolvidos à população.

Sejam

n_i : tamanho da i-ésima amostra ($i = 1, 2, \cdots, k$);

m_i : número de indivíduos marcados observados na i-ésima amostra ($i = 1, 2, \cdots, k$; $m_1 = 0$);

M_i : número de indivíduos marcados na população no momento da seleção da i-ésima amostra ($M_1 = 0$ e $M_2 = n_1$). Observemos que, $M_i = \sum_{j=1}^{i-1}(n_j - m_j)$, $2 \leq i \leq k+1$, onde

$M_{k+1} = \sum_{j=1}^{k}(n_j - m_j)$: número total de indivíduos distintos capturados durante o experimento.

A função densidade de probabilidade conjunta de $(m_1, m_2, \cdots, m_k; n_1, n_2, \cdots, n_k)$ dados p, c e N é dada por

$u_k = n_j - m_j$

$$P(0, m_2, \cdots, m_k; n_1, n_2, \cdots, n_k|p, c, N) = P(m_k, n_k|m_{k\text{-}1}, n_{k\text{-}1}, \cdots, m_2, n_2, 0, n_1, p, c, N)\times$$

$$\times P(m_{k-1}, n_{k-1}|m_{k-2}, n_{k-2}, \cdots, m_2, n_2, o, n_1, p, c, N)\times$$

$$\times \cdots \times P(m_2, n_2|o, n_1, p, c, N)P(0, n_1|p, c, N)$$

$u_2 = n_2 - m_2$

m_1

$$= \binom{M_k}{m_k} c^{m_k}(1-c)^{M_k - m_k}\binom{N - M_k}{n_k - m_k} p^{n_k - m_k}(1-p)^{N - M_k - n_k + m_k}\times$$

$$\times \binom{M_{k\text{-}1}}{m_{k\text{-}1}} c^{m_{k\text{-}1}}(1\text{-}c)^{M_{k\text{-}1}\text{-}m_{k\text{-}1}}\binom{N\text{-}M_{k\text{-}1}}{n_{k\text{-}1}\text{-}m_{k\text{-}1}} p^{n_{k\text{-}1}\text{-}m_{k\text{-}1}}(1\text{-}p)^{N\text{-}M_{k\text{-}1}\text{-}n_{k\text{-}1}+m_{k\text{-}1}}\times$$

$$\times \cdots \times \binom{M_2}{m_2} c^{m_2}(1-c)^{M_2 - m_2}\binom{N - M_2}{n_2 - m_2} p^{n_2 - m_2}\times$$

$$\times(1-p)^{N - M_2 - n_2 + m_2}\binom{N}{n_1} p^{n_1}(1-p)^{N - n_1}$$

$$= \prod_{i=1}^{k} \frac{M_i!(N-M_i)!}{m_i!(M_i-m_i)!(n_i-m_i)!(N-M_{i+1})!} c^{\sum_{i=1}^{k} m_i} (1-c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} \times$$

$$\times p^{M_{k+1}} (1-p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1}}$$

$$= \frac{N!}{(N-M_{k+1})!} \prod_{i=1}^{k} \frac{M_i!}{m_i!(M_i-m_i)!(n_i-m_i)!} c^{\sum_{i=1}^{k} m_i} \times$$

$$\times (1-c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} p^{M_{k+1}} (1-p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1}},$$

onde $0 \leq n_i \leq N$ e $\max\{0, M_i + n_i - N\} \leq m_i \leq \min\{n_i, M_i\}$.

A função de verossimilhança é dada por

$$L(p,c,N) = \frac{N!}{(N-M_{k+1})!} \prod_{i=1}^{k} \frac{M_i!}{m_i!(M_i-m_i)!(n_i-m_i)!} c^{\sum_{i=1}^{k} m_i} \times$$

$$\times (1-c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} p^{M_{k+1}} (1-p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1}} \propto K(p,c,N)$$

onde

$$K(p,c,N) = \frac{N!}{(N-M_{k+1})!} c^{\sum_{i=1}^{k} m_i} (1-c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} p^{M_{k+1}} (1-p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1}}, \quad (3.6.)$$

$0 < p < 1$, $0 < c < 1$ e $N \geq M_{k+1}$.

Teorema 40. Os E.M.V. para p e c são, respectivamente,

$$\hat{p} = \frac{M_{k+1}}{k\hat{N} - \sum_{i=1}^{k} M_i}$$

e

$$\hat{c} = \frac{\sum_{i=1}^{k} m_i}{\sum_{i=1}^{k} M_i},$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\left(1 - \frac{M_{k+1}}{kN - \sum_{i=1}^{k} M_i}\right)^k = 1 - \frac{M_{k+1}}{N}.$$

Prova. De (3.6.) segue que

$$\ln K(p,c,N) = \sum_{i=1}^{k} m_i \ln c + \left(\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i\right) \ln(1-c) + M_{k+1} \ln p +$$

$$+ \left(kN - \sum_{i=1}^{k} M_i - M_{k+1}\right) \ln(1-p) + \ln N! - \ln(N - M_{k+1})!.$$

Então

$$\frac{\partial K(p,c,N)}{\partial p} = \frac{M_{k+1}}{p} - \frac{kN - \sum_{i=1}^{k} M_i - M_{k+1}}{1-p} = 0$$

$$\Longleftrightarrow \frac{M_{k+1} - pM_{k+1} - p(kN - \sum_{i=1}^{k} M_i - M_{k+1})}{p(1-p)} = 0$$

$$\Longleftrightarrow M_{k+1} = p\left(M_{k+1} + kN - \sum_{i=1}^{k} M_i - M_{k+1}\right)$$

$$\Longleftrightarrow p = \frac{M_{k+1}}{kN - \sum_{i=1}^{k} M_i}. \qquad (3.7.)$$

e

$$\frac{\partial K(p,c,N)}{\partial c} = \frac{\sum_{i=1}^{k} m_i}{c} - \frac{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i}{1-c} = 0$$

$$\iff \frac{\sum_{i=1}^{k} m_i - c\sum_{i=1}^{k} m_i - c\sum_{i=1}^{k} M_i + c\sum_{i=1}^{k} m_i}{c(1-c)} = 0$$

$$\iff c = \frac{\sum_{i=1}^{k} m_i}{\sum_{i=1}^{k} M_i} \tag{3.8.}$$

Com relação ao E.M.V. para N temos:

$$\Delta K(p,c,N) = 0 \iff \ln \frac{K(p,c,N)}{K(p,c,N-1)} = 0$$

$$\iff ln \frac{N! \Big/ (N\text{-}M_{k+1})!\, c^{\sum_{i=1}^{k} m_i} (1\text{-}c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} p^{M_{k+1}} (1-p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1}}}{(N\text{-}1)! \Big/ (N\text{-}M_{k+1}\text{-}1)!\, c^{\sum_{i=1}^{k} m_i} (1\text{-}c)^{\sum_{i=1}^{k} M_i - \sum_{i=1}^{k} m_i} p^{M_{k+1}} (1\text{-}p)^{kN - \sum_{i=1}^{k} M_i - M_{k+1} - k}} = 0$$

$$\iff \frac{N}{N - M_{k+1}}(1-p)^k = 1$$

$$\iff (1-p)^k = 1 - \frac{M_{k+1}}{N}. \tag{3.9.}$$

Logo, de (3.7.), (3.8.) e (3.9.) segue que as E.M.V. para p e c são, respectivamente,

$$\hat{p} = \frac{M_{k+1}}{k\hat{N} - \sum_{i=1}^{k} M_i}$$

e

$$\hat{c} = \frac{\sum_{i=1}^{k} m_i}{\sum_{i=1}^{k} M_i},$$

e $\hat{N}$ é solução da equação

$$\Longleftrightarrow \left(1 - \frac{M_{k+1}}{kN - \sum_{i=1}^{k} M_i}\right)^k = 1 - \frac{M_{k+1}}{N}.$$

■

Com o objetivo de ilustrar os modelos propostos nas seções 3.1., 3.2. e 3.3., consideremos o caso em que são selecionadas duas amostras (k=2). Denotemos por N_1 o tamanho da primeira amostra, N_2 o tamanho da segunda amostra, M_3 o número de indivíduos distintos capturados durante o experimento e M o número de indivíduos marcados observados na segunda amostra. Notemos que $M_3 = N_1 + N_2 - M$. Cada indivíduo da população, independentemente dos demais, tem uma probabilidade p_1 $(0 < p_1 < 1)$ de ser capturado na primeira amostra e uma probabilidade p_2 $(0 < p_2 < 1)$ de ser capturado na segunda.

Então, pelo Teorema 38, os E.M.V. para p_1 e p_2 são, respectivamente,

$$\hat{p}_1 = \frac{N_1}{\hat{N}} \quad \text{e} \quad \hat{p}_2 = \frac{N_2}{\hat{N}}$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\left(1 - \frac{N_1}{N}\right)\left(1 - \frac{N_2}{N}\right) = 1 - \frac{N_1 + N_2 - M}{N}.$$

Resolvendo esta última equação temos:

$$N\left(1 - \frac{N_1}{N} - \frac{N_2}{N} + \frac{N_1 N_2}{N^2}\right) = N - N_1 - N_2 + M$$

$$\Longleftrightarrow \; -N\left(\frac{N_1}{N} + \frac{N_2}{N} - \frac{N_1 N_2}{N^2}\right) = -(N_1 + N_2 - M)$$

$$\Longleftrightarrow N(N_1 N + N_2 N - N_1 N_2) = N^2(N_1 + N_2 - M)$$

$$\Longleftrightarrow \; N(N_1 + N_2) - N_1 N_2 = N(N_1 + N_2) - MN$$

$$\Longleftrightarrow N = \frac{N_1 N_2}{M};$$

logo, o E.M.V. para N é

$$\hat{N} = \frac{N_1 N_2}{M}.$$

Suponhamos agora que $p_1 = p_2 = p$. Então, pelo Teorema 39, o E.M.V. para p é

$$\hat{p} = \frac{N_1 + N_2}{2\hat{N}}$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\left(1 - \frac{N_1 + N_2}{2N}\right)^2 = 1 - \frac{N_1 + N_2 - M}{N}.$$

Resolvendo a equação anterior temos:

$$\frac{(2N - N_1 - N_2)^2}{(2N)^2} = \frac{N - N_1 - N_2 + M}{N}$$

$$\Longleftrightarrow 4N^2 + N_1^2 + N_2^2 - 4N_1N - 4N_2N + 2N_1N_2 = 4N^2 - 4N(N_1 + N_2) + 4NM$$

$$\Longleftrightarrow N = \frac{(N_1 + N_2)^2}{4M};$$

logo, o E.M.V. para N é

$$\hat{N} = \frac{(N_1 + N_2)^2}{4M}.$$

Finalmente, supondo que cada indivíduo tem uma probabilidade p de ser capturado pela primeira vez, tanto na primeira quanto na segunda amostra, e uma probabilidade c de ser recapturado, segue, pelo Teorema 40, que os E.M.V. para p e c são, respectivamente,

$$\hat{p} = \frac{N_1 + N_2 - M}{2\hat{N} - N_1} \quad \text{e} \quad \hat{c} = \frac{M}{N_1}$$

onde $\hat{N}$: E.M.V. para N é solução da equação

$$\left(1 - \frac{N_1 + N_2 - M}{2N - N_1}\right)^2 = 1 - \frac{N_1 + N_2 - M}{N}.$$

Resolvendo esta última equação temos:

$$N\left(1 - \frac{N_1 + N_2 - M}{2N - N_1}\right)^2 = N - N_1 - N_2 + M$$

$$\Longleftrightarrow N(2N - N_1 - N_1 - N_2 + M)^2 = (2N - N_1)^2[N - N_1 - N_2 + M]$$

$$\iff N[(2N-N_1)^2+2(2N-N_1)(-N_1-N_2+M)+(-N_1-N_2+M)^2]$$

$$=(2N-N_1)^2N+(2N-N_1)^2(-N_1-N_2+M)$$

$$\iff (4N^2-2N_1N)(-N_1-N_2+M)+N(N_1^2+N_2^2+M^2+2N_1N_2-2N_1M-2N_2M)$$

$$=(4N^2-4N_1N+N_1^2)(-N_1-N_2+M)$$

$$\iff 2N_1^2N+2N_1N_2N-2N_1MN+N(N_1^2+N_2^2+M^2+2N_1N_2-2N_1M-2N_2M)$$

$$=4N_1^2N+4N_1N_2N-4N_1MN-N_1^3-N_1^2N_2+N_1^2M$$

$$\iff N_1^3+N_1^2N_2-N_1^2M=N(N_1^2-N_2^2-M^2+2N_2M)$$

$$\iff N=\frac{N_1^2[N_1+(N_2-M)]}{N_1^2-(N_2-M)^2}$$

$$\iff N=\frac{N_1^2}{N_1-(N_2-M)},\quad \text{se } N_1>N_2-M.$$

Logo, o E.M.V. para N é dado por:

$$\hat{N}=\frac{N_1^2}{N_1-(N_2-M)},\quad \text{se } N_1>N_2-M.$$

3.4. Teste de hipóteses para a igualdade das probabilidades de captura no caso de duas amostras.

Consideremos o caso de duas amostras, em que cada indivíduo da população tem uma probabilidade p_1 de ser capturado na primeira amostra e uma probabilidade p_2 de ser capturado na segunda. Nesta seção construiremos um teste de hipóteses para testar a igualdade das probabilidades de captura. Testaremos $H_0 : p_1 = p_2$ versus $H_1 : p_1 \neq p_2$. Para isto, usando a mesma notação do último parágrafo da seção anterior, consideremos a variável aleatória $T = N_1 - N_2$. Rejeitaremos H_0 se $|T|$ for maior que uma constante T_α, onde T_α depende da probabilidade α de erro do tipo I. O teste será baseado na distribuição de probabilidade condicional de T, dado que $M_3 = N_1 + N_2 - M$ indivíduos distintos foram capturados. Notemos que M indivíduos foram capturados duas vezes.

Como

$$P(T = t|M_3 = r, M = m) = P(N_1 - N_2 = t|N_1 + N_2 - M = r, M = m)$$

$$= \frac{P(N_1 - N_2 = t, N_1 + N_2 = r + m, M = m)}{\sum_{j=m}^{r} P(N_1 = j, N_1 + N_2 = r + m, M = m)}$$

$$= \frac{P(N_1 = \frac{r+t+m}{2}, N_2 = \frac{r+m-t}{2}, M = m)}{\sum_{j=m}^{r} P(N_1 = j, N_2 = r + m - j, M = m)}$$

$$= \frac{\frac{N!p_1^{\frac{r+t+m}{2}}(1-p_1)^{N-\frac{r+t+m}{2}} p_2^{\frac{r+m-t}{2}}(1-p_2)^{N-\frac{r+m-t}{2}}}{((r+t+m)/2-m)!((r+m-t)/2-m)!m!(N-r)!}}{\sum_{j=m}^{r} \frac{N!\ p_1{}^j(1-p_1)^{N-j}p_2^{r+m-j}(1-p_2)^{N-r-m+j}}{(j-m)!(r+m-j-m)!m!(N-r)!}}$$

$$= \frac{(p_1(1-p_2))^{\frac{t+r-m}{2}}(p_2(1-p_1))^{\frac{r-t-m}{2}}(r-m)!}{\left(\frac{r+t-m}{2}\right)!\left(\frac{r-t-m}{2}\right)! \sum_{j=m}^{r} \binom{r-m}{r-j}(p_2(1-p_1))^{r-j}(p_1(1-p_2))^{j-m}}$$

$$= \binom{r-m}{\frac{t+r-m}{2}} \frac{(p_2(1-p_1))^{\frac{r-t-m}{2}}(p_1(1-p_2))^{\frac{t+r-m}{2}}}{(p_1+p_2-2p_1p_2)^{r-m}}$$

$$= \binom{r-m}{\frac{t+r-m}{2}} \left(\frac{p_1(1-p_2)}{p_1+p_2-2p_1p_2}\right)^{\frac{t+r-m}{2}} \left(\frac{p_2(1-p_1)}{p_1+p_2-2p_1p_2}\right)^{\frac{-t+r-m}{2}}$$

$$= \binom{r-m}{\frac{t+r-m}{2}} \left(\frac{p_1(1-p_2)}{p_1+p_2-2p_1p_2} \right)^{\frac{t+r-m}{2}} \left(1 - \frac{p_1(1-p_2)}{p_1+p_2-2p_1p_2} \right)^{r-m-\frac{t+r-m}{2}},$$

para $t = 2l - r + m$, $l = 0, 1, 2, \cdots, r - m$, segue que a distribuição condicional de $X = \frac{T+A}{2}$ dados M_3 e M é Binomial(A, p'),
onde

$$A = M_3 - M \qquad \text{e} \quad p' = \frac{p_1(1-p_2)}{p_1+p_2-2p_1p_2}.$$

A região de rejeição para este teste é definida por $(|T| > T_\alpha)$, onde

$$P(|T| > T_\alpha | M_3, M; p_1 = p_2) = \alpha.$$

Logo, como

$$P(|T| \leq T_\alpha | M_3, M; p_1 = p_2)$$

$$= P\left(\frac{-T_\alpha + A}{2} \leq X \leq \frac{T_\alpha + A}{2} | M_3, M; p_1 = p_2 \right) = 1 - \alpha$$

e a distribuição condicional de X dados M_3, M e $p_1 = p_2$ é Binomial $(A, p' = 1/2)$, segue que $\frac{T_\alpha + A}{2} = X_{1-\frac{\alpha}{2}}$, ou

$$T_\alpha = 2X_{1-\frac{\alpha}{2}} - A,$$

onde $X_{1-\frac{\alpha}{2}}$ é o $(1 - \frac{\alpha}{2})$-ésimo quantil da distribuição Binomial $(A, 1/2)$. Se $A \geq 12$, (ou $Ap'(1-p') \geq 3$) podemos usar a aproximação Normal para a Distribuição Binomial para determinar T_α. Com efeito, dados M_3, M e $p_1 = p_2$,

$$\frac{X - \frac{A}{2}}{\sqrt{\frac{A}{4}}} = \frac{2X - A}{\sqrt{A}} \cong Z, \text{ onde } Z \sim N(0,1)$$

Logo,

$$T_\alpha = 2X_{1-\frac{\alpha}{2}} - A \cong \sqrt{A} Z_{1-\frac{\alpha}{2}},$$

onde $Z_{1-\frac{\alpha}{2}}$ é o $(1 - \frac{\alpha}{2})$-ésimo quantil da distribuição Normal Padrão.

A função poder do teste é dada por:

$$\pi(p') = P(\text{rejeitar } H_0|M_3, M, p') = P(|T| > T_\alpha|M_3, M, p') = 1 - P(|T| \leq T_\alpha|M_3, M, p')$$

$$= 1 - P(-T_\alpha \leq T \leq T_\alpha|M_3, M, p')$$

$$= 1 - P\left(\frac{A - T_\alpha}{2} \leq X \leq \frac{A + T_\alpha}{2}\middle| M_3, M, p'\right).$$

Logo, se A for suficientemente grande

$$\pi(p') \cong 1 - P\left(\left[\left(\frac{A - T_\alpha}{2} - Ap'\right)\Big/\left(Ap'(1 - p')\right)^{\frac{1}{2}}\right] \leq Z \leq\right.$$

$$\left.\leq \left[\left(\frac{A + T_\alpha}{2} - Ap'\right)\Big/\left(Ap'(1 - p')\right)^{\frac{1}{2}}\right]\middle| M_3, M, p'\right)$$

$$= 1\text{-}\left[\phi\left(\left(\frac{A + T_\alpha}{2}\text{-}Ap'\right)\Big/\left(Ap'(1\text{-}p')\right)^{\frac{1}{2}}\right)\text{-}\phi\left(\left(\frac{A - T_\alpha}{2}\text{-}Ap'\right)\Big/\left(Ap'(1\text{-}p')\right)^{\frac{1}{2}}\right)\right],$$

onde ϕ é a função de distribuição Normal Padrão.

Para concluir este capítulo veremos um exemplo no qual não é possível obter os E.M.V. para alguns parâmetros, pois estes resultam não identificáveis para o modelo. Consideremos o caso de duas amostras onde cada indivíduo da população, independentemente dos demais, tem uma probabilidade p_1 $(0 < p_1 < 1)$ de ser capturado na primeira amostra, uma probabilidade p_2 $(0 < p_2 < 1)$ de ser capturado pela primeira vez na segunda amostra, (ou seja, p_2 é a probabilidade de captura de cada indivíduo não marcado, na segunda amostra) e uma probabilidade c $(0 < c < 1)$ de ser recapturado (ou seja, c é a probabilidade de captura de cada indivíduo marcado). Usaremos a mesma notação do último parágrafo da seção anterior.

A função densidade de probabilidade de (N_1, N_2, M) dados N, p_1, p_2 e c é dada por:

$$P(N_1 = n_1, N_2 = n_2, M = m \mid p_1, p_2, c, N)$$

$$= P(N_1 = n_1 \mid p_1, p_2, c, N)P(M = m, N_2 = n_2|N_1 = n_1, p_1, p_2, c, N)$$

$$= \binom{N}{n_1} p_1{}^{n_1}(1-p_1)^{N-n_1}\binom{n_1}{m} c^m (1-c)^{n_1-m}\binom{N-n_1}{n_2-m} p_2{}^{n_2-m}(1-p_2)^{N-n_1-n_2+m}$$

$$= N! p_1{}^{n_1}(1-p_1)^{N-n_1} c^m (1-c)^{n_1-m} p_2{}^{n_2-m} \times$$

$$\times (1-p_2)^{N-n_1-n_2+m} / m!(n_1-m)!(n_2-m)!(N-n_1-n_2+m)!$$

$$\propto K(N,p_1,p_2,c) = N! p_1{}^{n_1}(1\text{-}p_1)^{N\text{-}n_1} c^m (1\text{-}c)^{n_1\text{-}m} p_2{}^{n_2\text{-}m}(1\text{-}p_2)^{N\text{-}n_1\text{-}n_2+m}/(N\text{-}n_1\text{-}n_2+m)!,$$

onde $N \geq M_3$.

Como

$$\ln K(p_1,p_2,c,N) = \ln N! + n_1 \ln p_1 + (N-n_1)\ln(1-p_1) + m \ln c + (n_1-m)\ln(1-c) +$$

$$+(n_2-m)\ln p_2 + (N-n_1-n_2+m)\ln(1-p_2) - \ln(N-n_1-n_2+m)!,$$

segue que

$$\frac{\partial \ln K(p_1,p_2,c,N)}{\partial p_1} = \frac{n_1}{p_1} - \frac{N-n_1}{1-p_1} = 0$$

$$\Longleftrightarrow n_1 - n_1 p_1 - p_1 N + n_1 p_1 = 0$$

$$\Longleftrightarrow p_1 = \frac{n_1}{N}.$$

Analogamente,

$$\frac{\partial \ln K(p_1,p_2,c,N)}{\partial p_2} = \frac{n_2-m}{p_2} - \frac{N-n_1-n_2+m}{1-p_2} = 0.$$

$$\Longleftrightarrow n_2 - n_2 p_2 - m + m p_2 - p_2 N + n_1 p_2 + n_2 p_2 - m p_2 = 0$$

$$\Longleftrightarrow p_2(n_1 - N) = m - n_2$$

$$\Longleftrightarrow p_2 = \frac{n_2 - m}{N - n_1}$$

e

$$\frac{\partial \ln K(N, p_1, p_2, c)}{\partial c} = \frac{m}{c} - \frac{n_1 - m}{1 - c} = 0.$$

$$\Longleftrightarrow m - mc - n_1 c + mc = 0$$

$$\Longleftrightarrow c = \frac{m}{n_1}.$$

Com relação ao parâmetro N temos,

$$\ln \frac{K(N, p_1, p_2, c)}{K(N-1, p_1, p_2, c)} = 0 \Longleftrightarrow \frac{K(N, p_1, p_2, c)}{K(N-1, p_1, p_2, c)} = 1$$

$$\Longleftrightarrow \frac{N(1-p_1)(1-p_2)}{N - n_1 - n_2 + m} = 1.$$

Logo, as E.M.V. para p_1, p_2 , c e N satisfazem o sistema

$$\begin{cases} \hat{p}_1 = \frac{n_1}{\hat{N}} \\ \hat{p}_2 = (n_2 - m)/(\hat{N} - n_1) \\ \hat{c} = m/n_1 \\ \frac{\hat{N}}{\hat{N} - n_1 - n_2 + m}(1 - \hat{p}_1)(1 - \hat{p}_2) = 1 \end{cases}$$

Contudo, a última equação é redundante, e portanto, as E.M.V. para p_1, p_2 e N não podem ser obtidas neste modelo.

Figura 1. Função de poder do teste para $\alpha = 0.01$.

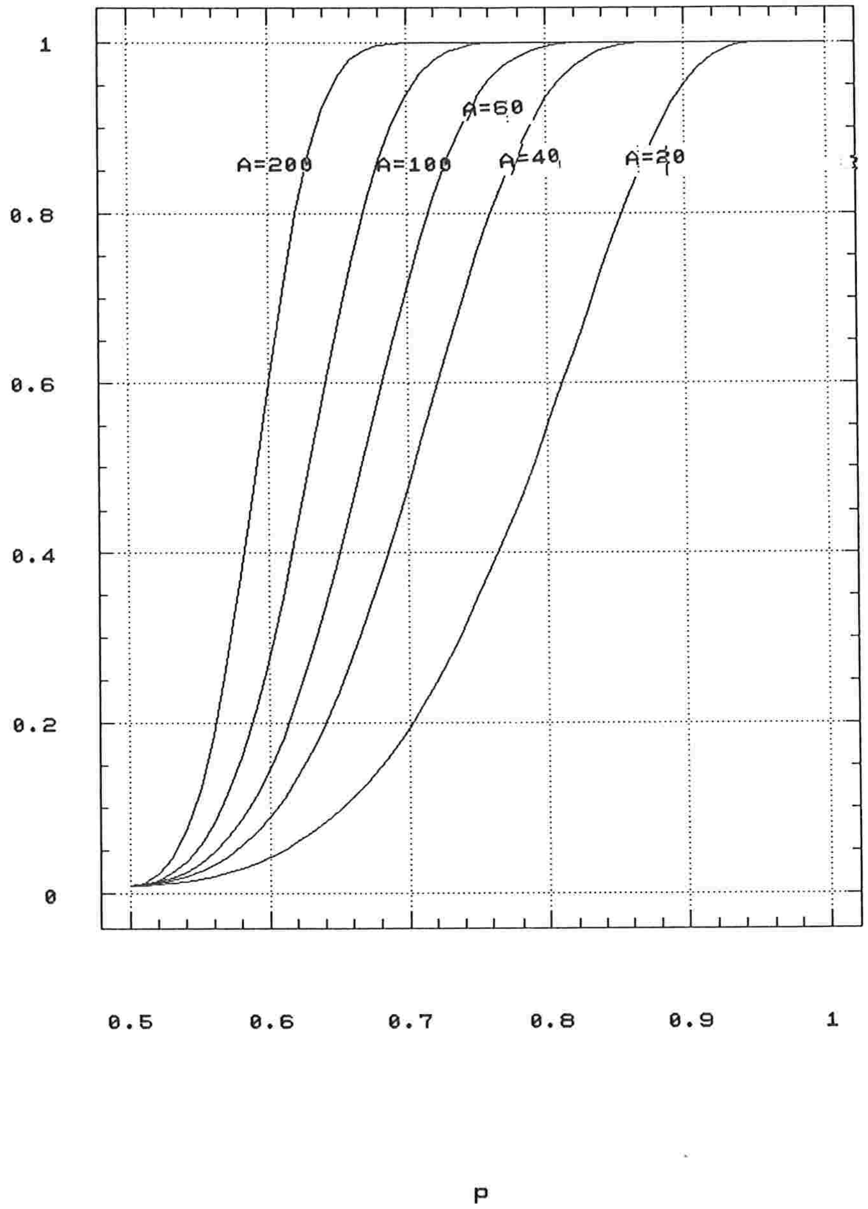


Figura 2. Função de poder do teste para $\alpha = 0.05$.

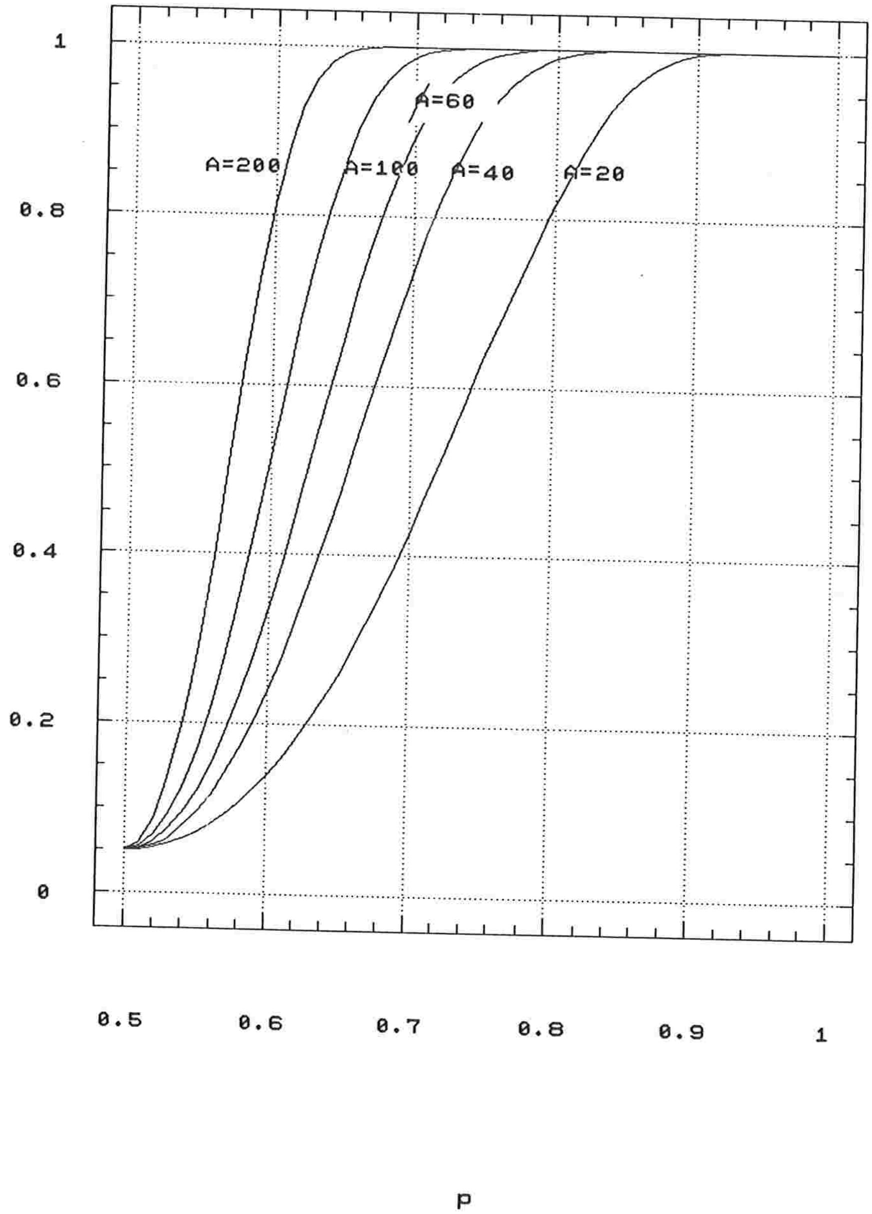


Figura 3. Função de poder do teste para $\alpha = 0.10$.

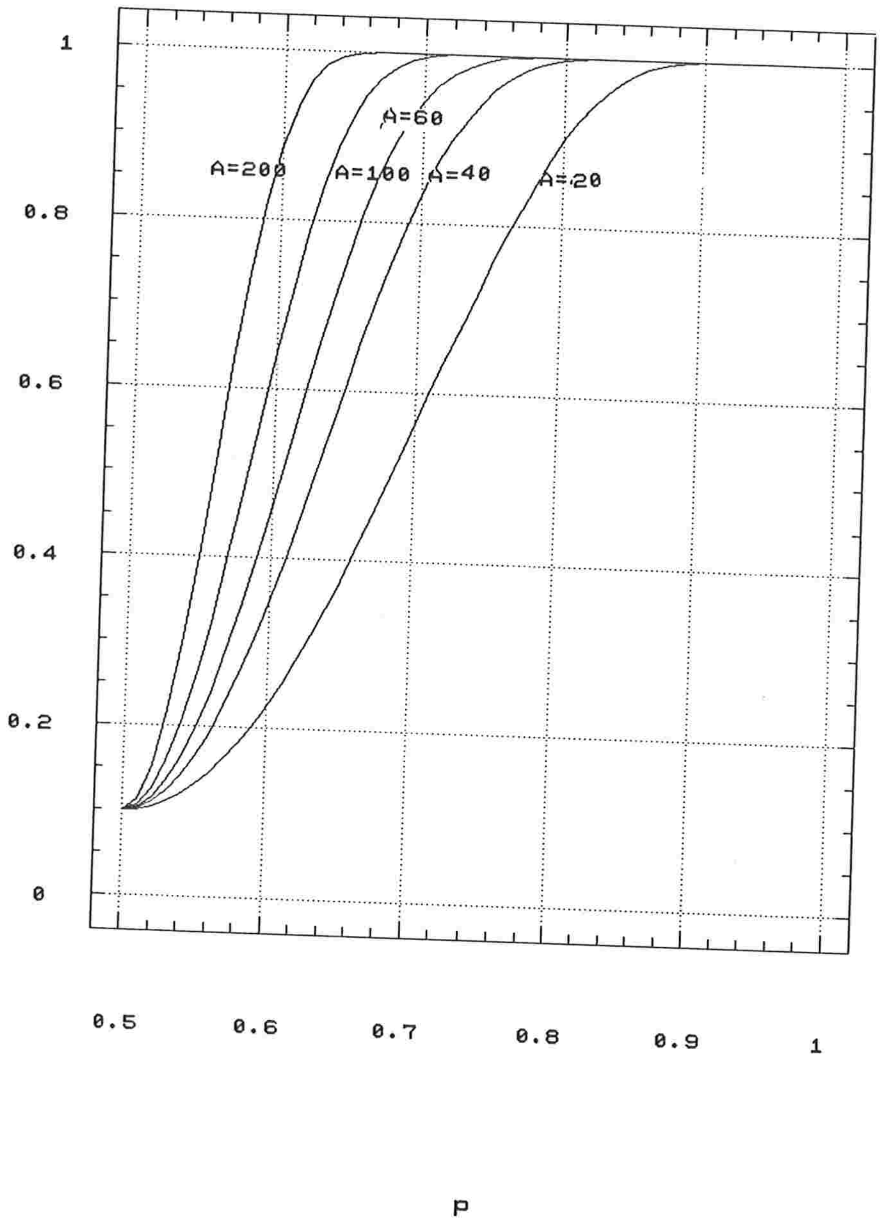
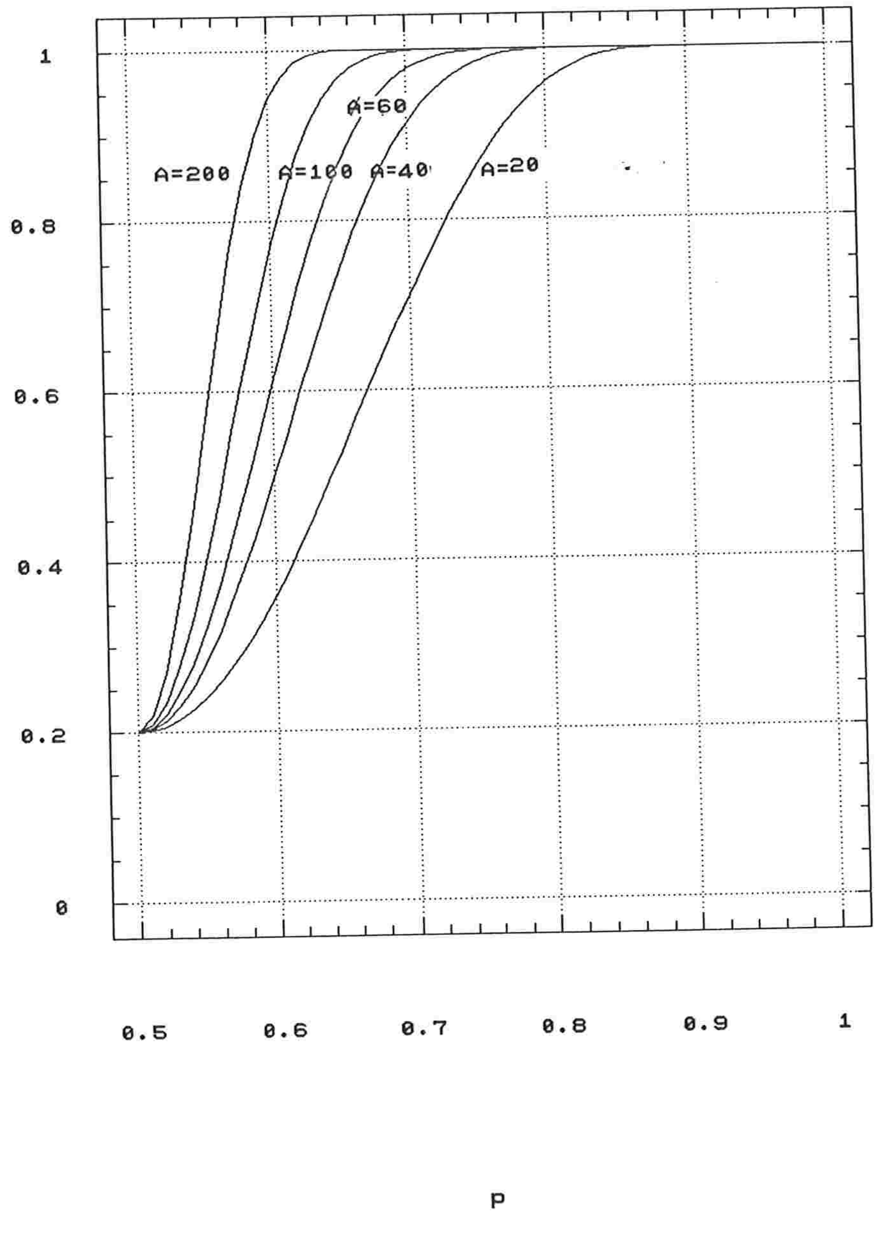


Figura 4. Função de poder do teste para $\alpha = 0.20$.



CAPÍTULO 4

MÉTODO DE CAPTURA-RECAPTURA EM POPULAÇÕES ABERTAS.

4.1. Suposições e Notação.

O objetivo deste capítulo é estudar o método de captura-recaptura para populações nas quais podem ocorrer nascimentos, mortes, imigrações e emigrações (populações abertas). Consideremos neste capítulo uma "única" população aberta. A palavra "única " significa uma população cobrindo uma área limitada onde os animais (ou, em geral, indivíduos) tem liberdade para se movimentar e se misturar com outros do mesmo tipo que estão na área, como também para entrar e sair da mesma. O tipo de situação excluída é aquela onde a população se divide em um certo número de áreas, tornando-se necessárias estimações dos parâmetros populacionais para cada área, como também para o número de animais que se movimentam de uma área para outra. Derivamos de maneira intuitiva estimadores para alguns parâmetros populacionais e, a seguir, mostramos que eles são os E.M.V. para tais parâmetros.

Consideramos k $(k \geq 3)$ tentativas de capturas de animais da população. Seja n_i o número de animais capturados na i-ésima tentativa $(i = 1, 2, \cdots, k)$. Em cada tentativa todos os animais capturados (marcados ou não) são marcados de modo a indicar que foram capturados nessa tentativa. Os intervalos de tempo entre as sucessivas tentativas não são necessariamente iguais. Dos n_i animais marcados na época da i-ésima tentativa $(i = 1, 2, \cdots, k)$, s_i são devolvidos à população $(s_i \leq n_i)$. O modelo leva em conta mortes acidentais por marcação ou manuseio dos animais.

Suposições:

Admitamos as seguintes hipóteses:

a) cada animal da população, marcado ou não, tem independentemente dos demais, uma probabilidade p_i de ser capturado na i-ésima tentativa ($i = 1, 2, \cdots, k$);

b) cada animal da população tem, independentemente dos demais, uma probabilidade ϕ_i de sobreviver entre a i-ésima e a $(i+1)$-ésima tentativa ($1 \leq i \leq k-1$);

c) cada animal capturado na i-ésima tentativa tem, independentemente dos demais, uma probabilidade η_i ($1 \leq i \leq k$) de ser devolvido à população;

d) os animais marcados não perdem suas marcas durante o processo e todas as marcas observadas são registradas;

e) os tempos de duração das tentativas são desprezíveis quando comparados com os intervalos entre as épocas das sucessivas tentativas;

f) cada animal que emigra da população o fará de maneira permanente.

Notação.

Sejam:

N_i: número total de indivíduos ou animais na população na época imediatamente anterior à da i-ésima tentativa, ($1 \leq i \leq k$);

M_i: número total de animais marcados na população na época imediatamente anterior à da i-ésima tentativa, ($1 \leq i \leq k$);

m_i: número de animais marcados observados na i-ésima tentativa, ($1 \leq i \leq k$);

B_i: número de novos animais que se juntam à população entre a i-ésima e a $(i+1)$-ésima tentativa, ($1 \leq i \leq k-1$);

N_{ij}: número total de animais na população na época imediatamente anterior à da i-ésima tentativa, capturados pela última vez na época j-ésima tentativa, ($2 \leq i \leq k$ e $1 \leq j \leq i-1$);

n_{ij}: número de animais observados na i-ésima tentativa, capturados pela última vez na j-ésima tentativa, ($2 \leq i \leq k$ e $1 \leq j \leq i-1$);

N_{i0}: número de animais não marcados na população na época imediatamente anterior à da i-ésima tentativa, ($1 \leq i \leq k$);

n_{i0}: número de animais não marcados observados na i-ésima tentativa, ($1 \leq i \leq k$);

Z_i : número de animais não capturados na i-ésima tentativa, mas capturados antes e após a i-ésima tentativa, $(2 \leq i \leq k-1)$;

R_i : número de animais capturados dentre os s_i após a época da i-ésima tentativa, $(1 \leq i \leq k-1)$.

Observemos que

$$M_1 = m_1 = 0;\ M_i = \sum_{j=1}^{i-1} N_{ij},\ (2 \leq i \leq k);\ N_i = \sum_{j=0}^{i-1} N_{ij},\ (1 \leq i \leq k);$$

$$n_i = \sum_{j=0}^{i-1} n_{ij},\ (1 \leq i \leq k);\ Z_i = \sum_{j=1}^{i-1} \sum_{r=i+1}^{k} n_{rj},\ (2 \leq i \leq k-1)$$

e

$$R_i = \sum_{j=i+1}^{k} n_{ji},\quad (1 \leq i \leq k-1).$$

Devido à natureza estocástica do modelo, o tamanho da população, N_i, na época da i-ésima tentativa, estará afetado pela variação aleatória do número de sobreviventes das épocas anteriores, sendo, portanto, uma variável aleatória. Somente na época da i-ésima tentativa é que N_i torna-se um parâmetro no sentido usual, sendo então um parâmetro de uma distribuição condicionada aos eventos ocorridos até essa época. Por esta razão, tais parâmetros se dizem parâmetros condicionais. Observemos que os parâmetros N_1, p_i, η_i, $(1 \leq i \leq k)$; ϕ_j e B_j, $(1 \leq j \leq k-1)$ não são parâmetros condicionais.

4.2. Estimadores intuitivos dos parâmetros N_i, ϕ_i e B_i.

Um estimador para N_i é o estimador de Petersen, obtido igualando-se a proporção dos animais marcados observados na i-ésima tentativa à proporção dos animais marcados na população imediatamente antes da i-ésima tentativa:

$$\frac{m_i}{n_i} = \frac{M_i}{N_i} \Rightarrow \hat{N}_i = \frac{n_i \hat{M}_i}{m_i}, \qquad (i = 2, 3, \cdots, k). \tag{4.1.}$$

Um estimador de ϕ_i pode ser definido como o quociente entre $\hat{M}_{i+1}$, um estimador de M_{i+1}, e $\hat{M}_i - m_i + s_i$, um estimador do número de animais marcados na população após a i-ésima tentativa:

$$\hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + s_i}, \quad (1 \le i \le k-1). \qquad (4.2.)$$

Observemos que, se

$$s_i = n_i \Rightarrow \hat{\phi}_i = \frac{\hat{M}_{i+1}}{\hat{M}_i + (n_i - m_i)} \quad (1 \le i \le k-1).$$

Como o tamanho da população imediatamente após a i-ésima tentativa é $N_i - n_i + s_i$, então X_i = número de sobreviventes entre a i-ésima e a $(i+1)$-ésima tentativa é uma variável aleatória Binomial $(N_i - n_i + s_i, \phi_i)$, $(1 \le i \le k-1)$.
Logo, o número esperado de sobreviventes entre a i-ésima e a $i+1$-ésima tentativa é

$$E(X_i) = (N_i - n_i + s_i)\phi_i, \quad (1 \le i \le k-1)$$

e, portanto, um estimador do número de nascimentos (nascimentos e imigrações) entre i-ésima e a $i+1$-ésima tentativa pode ser definido como:

$$\hat{B}_i = \hat{N}_{i+1} - \hat{E}(X_i) = \hat{N}_{i+1} - \hat{\phi}_i(\hat{N}_i - n_i + s_i), \ 1 \le i \le k-1. \qquad (4.3.)$$

Notemos que, imediatamente após a i-ésima tentativa, há na população dois grupos de animais marcados: o grupo dos $M_i - m_i$ animais marcados antes da i-ésima tentativa (grupo I) e o grupo dos s_i animais marcados exatamente na i-ésima tentativa (grupo II). Por outro lado, Z_i é o número de animais capturados da população constituida pelo grupo I, após a i-ésima tentativa e R_i é o número de animais capturados da população constituida pelo grupo II, após a i-ésima tentativa.

Logo, como as probabilidades de sobrevivência para os animais de cada grupo são iguais, segue que, um estimador para M_i pode ser obtido igualando-se as proporções dos animais capturados após a i-ésima tentativa para ambos os grupos I e II:

$$\frac{Z_i}{M_i - m_i} = \frac{R_i}{s_i}, \ (2 \le i \le k-1),$$

o que implica

$$\hat{M}_i = m_i + \frac{s_i Z_i}{R_i}, \quad (2 \le i \le k-1). \qquad (4.4.)$$

Das equações (4.1.) e (4.4.) segue que

$$\hat{N}_i = n_i + \frac{n_i s_i Z_i}{m_i R_i} = n_i[1 + s_i Z_i/(m_i R_i)], \quad (2 \le i \le k-1). \qquad (4.5.)$$

Usando as equações (4.2.) e (4.4.) temos

$$\hat{\phi}_i = \frac{m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1}}{m_i + s_iZ_i/Ri - m_i + s_i}$$

$$= \frac{m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1}}{s_iZ_i/R_i + s_i}, \quad (2 \leq i \leq k-2). \tag{4.6.}$$

Usando as equações (4.2.) para $i = 1$ e (4.4.) para $i = 2$ temos

$$\hat{\phi}_1 = \frac{\hat{M}_2}{s_1} = (m_2 + s_2Z_2/R_2)/s_1.$$

Finalmente, usando as equações (4.3.), (4.5.) e (4.6.) temos

$$\hat{B}_i = n_{i+1} + \frac{n_{i+1}s_{i+1}Z_{i+1}}{m_{i+1}R_{i+1}} - \frac{m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1}}{s_i + s_iZ_i/R_i}(n_i + n_is_iZ_i/(m_iR_i) - n_i + s_i)$$

$$= n_{i+1}[1 + s_{i+1}Z_{i+1}/(m_{i+1}R_{i+1})] - \frac{(m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1})}{1 + Z_i/R_i}(n_iZ_i/(m_iR_i) + 1), \tag{4.7.}$$

$2 \leq i \leq k-2$.

4.3. Função densidade de probabilidade conjunta das variáveis aleatórias associadas ao modelo .

Nesta seção deduziremos a função densidade de probabilidade conjunta de $\{\{n_{10}, s_1, N_{20}, N_{21}\}; \{n_{ij}, s_i, N_{i+1,r}\ ;\ 2 \leq i \leq k-1, 0 \leq j \leq i-1, 0 \leq r \leq i\}\ ;\ \{n_{kj}, s_k,$ $0 \leq j \leq k-1\}\}$ condicionado a $\{N_{10}, \{B_i, p_i, \phi_i, \eta_i, 1 \leq i \leq k-1\}, \{p_k, \eta_k\}\}$. Para uma melhor compreensão do assunto trataremos, inicialmente, do caso $k = 3$.

Notemos que de

$$P(n_{21}|N_{21}, p_2, \phi_2) = \binom{N_{21}}{n_{21}} p_2^{n_{21}}(1-p_2)^{N_{21}-n_{21}}$$

e

$$P(N_{31}|n_{21}, N_{21}, p_2, \phi_2) = \binom{N_{21} - n_{21}}{N_{31}} \phi_2^{N_{31}}(1-\phi_2)^{N_{21}-n_{21}-N_{31}}$$

segue que

$$P(n_{21}, N_{31}|N_{21}, p_2, \phi_2) = P(n_{21}|N_{21}, p_2, \phi_2)P(N_{31}|n_{21}, N_{21}, p_2, \phi_2)$$

$$= \binom{N_{21}}{n_{21}} p_2^{n_{21}} (1-p_2)^{N_{21}-n_{21}} \binom{N_{21}-n_{21}}{N_{31}} \phi_2^{N_{31}} (1-\phi_2)^{N_{21}-n_{21}-N_{31}}. \quad (4.8.)$$

Temos também que

$$P(n_{3j}|N_{3j}, p_3) = \binom{N_{3j}}{n_{3j}} p_3^{n_{3j}} (1-p_3)^{N_{3j}-n_{3j}} \quad (1 \leq j \leq 2), \quad (4.9.)$$

e

$$P(n_{i0}|N_{i0}, p_i, \phi_i, B_i) = \binom{N_{i0}}{n_{i0}} p_i^{n_{i0}} (1-p_i)^{N_{i0}-n_{i0}}, \quad (1 \leq i \leq 2).$$

Observemos que a distribuição de probabilidades de $N_{i+1,0}$ condicionada a $\{n_{i0}, N_{i0}, p_i, \phi_i, B_i,\ 1 \leq i \leq 2\}$ é análoga à de N_{31} condicionada a $\{n_{21}, N_{21}, p_2, \phi_2\}$ exceto que B_i, $(1 \leq i \leq 2)$ novos animais juntam-se à população, de modo que o número de animais que nunca foram capturados e sobreviveram entre as épocas i e $i+1$ é $N_{i+1,0} - B_i$, $(1 \leq i \leq 2)$. Logo,

$$P(N_{i+1,0}|n_{i0}, N_{i0}, p_i, \phi_i, B_i)$$

$$= \binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i} \phi_i^{N_{i+1,0}-B_i} (1-\phi_i)^{N_{i0}-n_{i0}-N_{i+1,0}+B_i}, \quad (1 \leq i \leq 2)$$

o que implica

$$P(n_{i0}, N_{i+1,0}|N_{i0}, p_i, \phi_i, B_i) = P(N_{i+1,0}|n_{i0}, N_{i0}, p_i, \phi_i, B_i) P(n_{i0}|N_{i0}, p_i, \phi_i, B_i)$$

$$= \binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i} \phi_i^{N_{i+1,0}-B_i} (1-\phi_i)^{N_{i0}-n_{i0}-N_{i+1,0}+B_i} \times$$

$$\times \binom{N_{i0}}{n_{i0}} p_i^{n_{i0}} (1-p_i)^{N_{i0}-n_{i0}}, \quad (1 \leq i \leq 2). \quad (4.10.)$$

Também temos que

$$P(n_{30}|N_{30}, p_3) = \binom{N_{30}}{n_{30}} p_3^{n_{30}} (1-p_3)^{N_{30}-n_{30}}. \quad (4.11.)$$

De

$$P(s_i|n_i,\phi_i,\eta_i)=\binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i},\quad (1\le i\le 2)$$

e

$$P(N_{i+1,i}|s_i,n_i,\phi_i,\eta_i)=\binom{s_i}{N_{i+1,i}}\phi_i^{N_{i+1,i}}(1-\phi_i)^{s_i-N_{i+1,i}},\quad (1\le i\le 2)$$

segue que

$$P(s_i,N_{i+1,i}|n_i,\phi_i,\eta_i)=P(s_i|n_i,\phi_i,\eta_i)P(N_{i+1,i}|s_i,n_i,\phi_i,\eta_i)$$

$$=\binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i}\binom{s_i}{N_{i+1,i}}\phi_i^{N_{i+1,i}}(1-\phi_i)^{s_i-N_{i+1,i}},\quad (1\le i\le 2). \qquad (4.12.)$$

Também temos

$$P(s_3|n_3,\eta_3)=\binom{n_3}{s_3}\eta_3^{s_3}(1-\eta_3)^{n_3-s_3} \qquad (4.13.)$$

Notemos que, associadas à primeira amostra existem as variáveis n_{10},s_1,N_{20} e N_{21}; associadas à segunda amostra existem as variáveis $n_{20},n_{21},s_2,N_{30},N_{31}$ e N_{32} e associadas à terceira amostra existem as variáveis n_{30},n_{31},n_{32} e s_3. Sejam

$$A_1=\{n_{10},s_1,N_{20},N_{21}\},\quad A_2=\{n_{20},n_{21},s_2,N_{30},N_{31},N_{32}\},$$

$$A_3=\{n_{30},n_{31},n_{32},s_3\}\ \text{ e }\ A_4=\{N_{10},B_1,p_1,\phi_1,\eta_1,B_2,p_2,\phi_2,\eta_2,p_3,\eta_3\}.$$

Então, a função densidade de probabilidade conjunta de $\{\{n_{10},s_1,N_{20},N_{21}\};\{n_{20},n_{21},s_2,N_{30},N_{31},N_{32}\};\{n_{30},n_{31},n_{32},s_3\}\}$ condicionada a $\{N_{10},B_1,p_1,\phi_1,\eta_1,B_2,p_2,\phi_2,\eta_2,p_3,\eta_3\}$ é dada por

$$P(\{n_{10},s_1,N_{20},N_{21}\}\ ;\ \{n_{20},n_{21},s_2,N_{30},N_{31},N_{32}\}\ ;\{n_{30},n_{31},n_{32},s_3\}|$$

$$N_{10},B_1,p_1,\phi_1,\eta_1,B_2,p_2,\phi_2,\eta_2,p_3,\eta_3)$$

$$= P(A_1, A_2, A_3|A_4) = P(A_3|A_1, A_2, A_4)P(A_2|A_1, A_4)P(A_1|A_4). \qquad (4.14.)$$

Usando (4.9.), (4.11.) e (4.13.) temos:

$$P(A_3|A_1, A_2, A_4) = P(s_3|\ \{n_{30}, n_{31}, n_{32}\}, A_1, A_2, A_4)\times$$

$$\times P(n_{30}|\ \{n_{31}, n_{32}\}, A_1, A_2, A_4)P(n_{31}|\ \{n_{32}\}, A_1, A_2, A_4)P(n_{32}|A_1, A_2, A_4)$$

$$= \binom{n_3}{s_3}\eta_3^{s_3}(1-\eta_3)^{n_3-s_3}\binom{N_{30}}{n_{30}}p_3^{n_{30}}(1-p_3)^{N_{30}-n_{30}}\binom{N_{31}}{n_{31}}p_3^{n_{31}}(1-p_3)^{N_{31}-n_{31}}\times$$

$$\times\binom{N_{32}}{n_{32}}p_3^{n_{32}}(1-p_3)^{N_{32}-n_{32}} \qquad (4.15.)$$

Usando (4.8.), (4.10.) e (4.12.) temos:

$$P(A_2|A_1, A_4) = P(s_2, N_{32}|\{n_{20}, N_{30}, n_{21}, N_{31}\}, A_1, A_4)\times$$

$$\times P(n_{21}, N_{31}|\ \{n_{20}, N_{30}\}, A_1, A_4)P(n_{20}, N_{30}\ |A_1, A_4)$$

$$= \binom{n_2}{s_2}\eta_2^{s_2}(1-\eta_2)^{n_2-s_2}\binom{s_2}{N_{32}}\phi_2^{N_{32}}(1-\phi_2)^{s_2-N_{32}}\times$$

$$\times\binom{N_{21}}{n_{21}}p_2^{n_{21}}(1-p_2)^{N_{21}-n_{21}}\binom{N_{21}-n_{21}}{N_{31}}\phi_2^{N_{31}}(1-\phi_2)^{N_{31}-n_{21}-N_{31}}\times$$

$$\times\binom{N_{20}}{n_{20}}p_2^{n_{20}}(1-p_2)^{N_{20}-n_{20}}\binom{N_{20}-n_{20}}{N_{30}-B_2}\phi_2^{N_{30}-B_2}(1-\phi_2)^{N_{20}-n_{20}-N_{30}+B_2}. \qquad (4.16.)$$

Usando (4.10.) e (4.12.) temos:

$$P(A_1|A_4) = P(s_1, N_{21}|n_{10}, N_{20}; A_4)P(n_{10}, N_{20}|A_4) = \binom{n_1}{s_1}\eta_1^{s_1}(1-\eta_1)^{n_1-s_1}\times$$

$$\times\binom{s_1}{N_{21}}\phi_1^{N_{21}}(1-\phi_1)^{s_1-N_{21}}\binom{N_{10}}{n_{10}}p_1^{n_{10}}(1-p_1)^{N_{10}-n_{10}}\binom{N_{10}-n_{10}}{N_{20}-B_1}\phi_1^{N_{20}-B_1}\times$$

$$\times(1-\phi_1)^{N_{10}-n_{10}-N_{20}+B_1}. \qquad (4.17.)$$

Logo, usando (4.14.) a (4.17.) temos:

$$P(A_1,A_2,A_3|A_4)=\binom{n_3}{s_3}\eta_3^{s_3}(1-\eta_3)^{n_3-s_3}\binom{N_{30}}{n_{30}}p_3^{n_{30}}(1-p_3)^{N_{30}-n_{30}}\binom{N_{31}}{n_{31}}p_3^{n_{31}}\times$$

$$\times(1-p_3)^{N_{31}-n_{31}}\binom{N_{32}}{n_{32}}p_3^{n_{32}}(1\text{-}p_3)^{N_{32}-n_{32}}\binom{n_2}{s_2}\eta_2^{s_2}(1-\eta_2)^{n_2-s_2}\binom{s_2}{N_{32}}\phi_2^{N_{32}}\times$$

$$\times(1-\phi_2)^{s_2-N_{32}}\binom{N_{21}}{n_{21}}p_2^{n_{21}}(1-p_2)^{N_{21}-n_{21}}\binom{N_{21}-n_{21}}{N_{31}}\phi_2^{N_{31}}(1-\phi_2)^{N_{21}-n_{21}-N_{31}}\binom{N_{20}}{n_{20}}\times$$

$$\times p_2^{n_{20}}(1-p_2)^{N_{20}-n_{20}}\binom{N_{20}-n_{20}}{N_{30}-B_2}\phi_2^{N_{30}-B_2}(1-\phi_2)^{N_{20}-n_{20}-N_{30}+B_2}\binom{n_1}{s_1}\eta_1^{s_1}(1-\eta_1)^{n_1-s_1}\times$$

$$\times\binom{s_1}{N_{21}}\phi_1^{N_{21}}(1\text{-}\phi_1)^{s_1\text{-}N_{21}}\binom{N_{10}}{n_{10}}p_1^{n_{10}}(1\text{-}p_1)^{N_{10}\text{-}n_{10}}\binom{N_{10}\text{-}n_{10}}{N_{20}\text{-}B_1}\phi_1^{N_{20}\text{-}B_1}(1\text{-}\phi_1)^{N_{10}\text{-}n_{10}-N_{20}+B_1}$$

$$=\prod_{i=1}^{3}\binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i}\binom{N_{i0}}{n_{i0}}p_i^{n_{i0}}(1-p_i)^{N_{i0}-n_{i0}}\prod_{j=1}^{2}\binom{s_j}{N_{j+1,j}}\phi_j^{N_{j+1,j}}(1-\phi_j)^{s_j-N_{j+1,j}}\times$$

$$\times\binom{N_{j0}-n_{j0}}{N_{j+1,0}-B_j}\phi_j^{N_{j+1,0}-B_j}(1-\phi_j)^{N_{j0}-n_{j0}-N_{j+1,0}+B_j}\binom{N_{21}}{n_{21}}p_2^{n_{21}}\times$$

$$\times(1-p_2)^{N_{21}-n_{21}}\binom{N_{21}-n_{21}}{N_{31}}\phi_2^{N_{31}}(1-\phi_2)^{N_{21}-n_{21}-N_{31}}\prod_{r=1}^{2}\binom{N_{3r}}{n_{3r}}p_3^{n_{3r}}(1-p_3)^{N_{3r}-n_{3r}}$$

$$= \binom{N_{21}}{n_{21}}\binom{N_{21}-n_{21}}{N_{31}}\prod_{i=1}^{3}\binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i}\binom{N_{i0}}{n_{i0}}p_i^{n_i}(1-p_i)^{N_i-n_i}\times$$

$$\times\prod_{r=1}^{2}\binom{s_r}{N_{r+1,r}}\phi_r^{N_{r+1}-B_r}(1-\phi_r)^{s_r-n_r+B_r+N_r-N_{r+1}}\binom{N_{3r}}{n_{3r}}\binom{N_{r0}-n_{r0}}{N_{r+1,0}-B_r}.$$

A seguir determinaremos a função densidade de probabilidade conjunta de $\{\{n_{10},s_1,N_{20},N_{21}\};\{n_{ij},s_i,N_{i+1,r};2\leq i\leq k-1,0\leq j\leq i-1,0\leq r\leq i\};\{n_{kj},s_k,$ $0\leq j\leq k-1\}\}$ condicionada a $\{\{N_{10},B_i,p_i,\phi_i,\eta_i,1\leq i\leq k-1\},\{p_k,\eta_k\}\}$ (caso geral).

Notemos que de

$$P(n_{ij}|N_{ij},p_i,\phi_i)=\binom{N_{ij}}{n_{ij}}p_i^{n_{ij}}(1-p_i)^{N_{ij}-n_{ij}},\quad (1\leq i\leq k-1,1\leq j\leq i-1)$$

e de

$$P(N_{i+1,j}|n_{ij},N_{ij},p_i,\phi_i)$$

$$=\binom{N_{ij}-n_{ij}}{N_{i+1,j}}\phi_i^{N_{i+1,j}}(1-\phi_i)^{N_{ij}-n_{ij}-N_{i+1,j}},\quad (1\leq i\leq k-1,1\leq j\leq i-1)$$

segue que

$$P(n_{ij},N_{i+1,j}|N_{ij},p_i,\phi_i)=P(n_{ij}|N_{ij},p_i,\phi_i)P(N_{i+1,j}|n_{ij},N_{ij},p_i,\phi_i)$$

$$=\binom{N_{ij}}{n_{ij}}p_i^{n_{ij}}(1-p_i)^{N_{ij}-n_{ij}}\binom{N_{ij}-n_{ij}}{N_{i+1,j}}\times$$

$$\times\phi_i^{N_{i+1,j}}(1-\phi_i)^{N_{ij}-n_{ij}-N_{i+1,j}},\quad (1\leq i\leq k-1,1\leq j\leq i-1)\qquad (4.18.)$$

Também temos que

$$P(n_{kj}|N_{kj},p_k)=\binom{N_{kj}}{n_{kj}}p_k^{n_{kj}}(1-p_k)^{N_{kj}-n_{kj}}\quad (1\leq j\leq k-1).\qquad (4.19.)$$

De

$$P(n_{i0}|N_{i0},p_i,\phi_i,B_i)=\binom{N_{i0}}{n_{i0}}p_i^{n_{i0}}(1-p_i)^{N_{i0}-n_{i0}},\quad (1\le i\le k-1)$$

e de

$$P(N_{i+1,0}|n_{i0},N_{i0},p_i,\phi_i,B_i)$$

$$=\binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i}\phi_i^{N_{i+1,0}-B_i}(1-\phi_i)^{N_{i0}-n_{i0}-N_{i+1,0}+B_i},\quad (1\le i\le k-1)$$

segue que

$$P(n_{i0},N_{i+1,0}|N_{i0},p_i,\phi_i,B_i)=P(n_{i0}|N_{i0},p_i,\phi_i,B_i)P(N_{i+1,0}|n_{i0},N_{i0},p_i,\phi_i,B_i)$$

$$=\binom{N_{i0}}{n_{i0}}p_i^{n_{i0}}(1-p_i)^{N_{i0}-n_{i0}}\times$$

$$\times\binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i}\phi_i^{N_{i+1,0}-B_i}(1-\phi_i)^{N_{i0}-n_{i0}-N_{i+1,0}+B_i},\quad (1\le i\le k-1) \qquad (4.20.)$$

Também temos

$$P(n_{k0}|N_{k0},p_k)=\binom{N_{k0}}{n_{k0}}p_k^{n_{k0}}(1-p_k)^{N_{k0}-n_{k0}}. \qquad (4.21.)$$

De

$$P(s_i|n_i,\phi_i,\eta_i)=\binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i},\quad (1\le i\le k-1)$$

e de

$$P(N_{i+1,i}|s_i,n_i,\phi_i,\eta_i)=\binom{s_i}{N_{i+1,i}}\phi_i^{N_{i+1,i}}(1-\phi_i)^{s_i-N_{i+1,i}},\quad (1\le i\le k-1)$$

segue que

$$P(s_i,N_{i+1,i}|n_i,\phi_i,\eta_i)=P(s_i|n_i,\phi_i,\eta_i)P(N_{i+1,i}|s_i,n_i,\phi_i,\eta_i)$$

$$= \binom{n_i}{s_i}\eta_i^{s_i}(1-\eta_i)^{n_i-s_i}\binom{s_i}{N_{i+1,i}}\phi_i^{N_{i+1,i}}(1-\phi_i)^{s_i-N_{i+1,i}}, \quad (1 \le i \le k-1) \qquad (4.22.)$$

e

$$P(s_k|n_k,\eta_k) = \binom{n_k}{s_k}\eta_k^{s_k}(1-\eta_k)^{n_k-s_k}. \qquad (4.23.)$$

Definindo

$$A_1 = \{n_{10}, s_1, N_{20}, N_{21}\},$$

$$A_i = \{n_{ij}, s_i, N_{i+1,r}, 0 \le j \le i-1, 0 \le r \le i\}, \ (2 \le i \le k-1),$$

$$A_k = \{n_{kj}, s_k, 0 \le j \le k-1\}$$

e

$$V_1 = \{N_{10}, B_i, p_i, \phi_i, \eta_i, 1 \le i \le k-1\}, \quad V_2 = \{p_k, \eta_k\},$$

temos:

$$P(\{n_{10}, s_1, N_{20}, N_{21}\}; \{n_{ij}, s_i, N_{i+1,r}; 2 \le i \le k-1, 0 \le j \le i-1, 0 \le r \le i\}; \{n_{kj}, s_k,$$

$$0 \le j \le k-1\}|\{N_{10}, B_i, p_i, \phi_i, \eta_i, 1 \le i \le k-1\}, \{p_k, \eta_k\}) = P(A_1, A_2, \cdots, A_k|V_1, V_2)$$

$$= P(A_1|V_1,V_2)P(A_2|A_1,V_1,V_2)\cdots P(A_k|A_1,A_2,\cdots,A_{k-1},V_1,V_2) \qquad (4.24.)$$

Usando (4.20.) e (4.22.) temos

$$P(A_1|V_1,V_2) = P(n_{10}, s_1, N_{20}, N_{21}|V_1,V_2) = P(s_1, N_{21}|n_{10}, N_{20}, V_1, V_2)\times$$

$$\times P(n_{10}, N_{20}|V_1, V_2)$$

$$= \binom{n_1}{s_1}\eta_1^{s_1}(1-\eta_1)^{n_1-s_1}\binom{s_1}{N_{21}}\phi_1^{N_{21}}(1-\phi_1)^{s_1-N_{21}}\binom{N_{10}}{n_{10}}p_1^{n_{10}}(1-p_1)^{N_{10}-n_{10}}\times$$

$$\times\binom{N_{10}-n_{10}}{N_{20}-B_1}\phi_1^{N_{20}-B_1}(1-\phi_1)^{N_{10}-n_{10}-N_{20}+B_1} \qquad (4.25.)$$

Usando (4.18.), (4.20.) e (4.22.) temos,

$$P(A_2|A_1,V_1,V_2) = P(\{n_{2j},s_2,N_{3r},0\le j\le 1,0\le r\le 2\}|A_1,V_1,V_2)$$

$$= P(n_{20},n_{21},s_2,N_{30},N_{31},N_{32}|A_1,V_1,V_2) = P(s_2,N_{32}|n_{20},n_{21},N_{30},N_{31},A_1,V_1,V_2)\times$$

$$\times P(n_{21},N_{31}|n_{20},N_{30},A_1,V_1,V_2)P(n_{20},N_{30}|A_1,V_1,V_2)$$

$$= \binom{n_2}{s_2}\eta_2^{s_2}(1-\eta_2)^{n_2-s_2}\binom{s_2}{N_{32}}\phi_2^{N_{32}}(1-\phi_2)^{s_2-N_{32}}\binom{N_{21}}{n_{21}}p_2^{n_{21}}(1-p_2)^{N_{21}-n_{21}}\times$$

$$\times\binom{N_{21}-n_{21}}{N_{31}}\phi_2^{N_{31}}(1-\phi_2)^{N_{21}-n_{21}-N_{31}}\binom{N_{20}}{n_{20}}p_2^{n_{20}}(1-p_2)^{N_{20}-n_{20}}\times$$

$$\times\binom{N_{20}-n_{20}}{N_{30}-B_2}\phi_2^{N_{30}-B_2}(1-\phi_2)^{N_{20}-n_{20}-N_{30}+B_2}. \qquad (4.26.)$$

Usando (4.18.), (4.20.) e (4.22.) temos:

$$P(A_3|A_2,A_1,V_1,V_2) = P(\{n_{3j},s_3,N_{4r},0\le j\le 2,0\le r\le 3\}|A_2,A_1,V_1,V_2)$$

$$= P(\{n_{30},n_{31},n_{32},s_3,N_{40},N_{41},N_{42},N_{43}\}|A_2,A_1,V_1,V_2)$$

$$= P(\{s_3,N_{43}\}|\{n_{30},n_{31},n_{32},N_{40},N_{41},N_{42}\},A_2,A_1,V_1,V_2)\times$$

$$\times P(\{n_{32}, N_{42}\}|\{n_{30}, n_{31}, N_{40}, N_{41}\}, A_2, A_1, V_1, V_2)\times$$

$$\times P(\{n_{31}, N_{41}\}|\{n_{30}, N_{40}\}, A_2, A_1, V_1, V_2)P(n_{30}, N_{40}|A_2, A_1, V_1, V_2)$$

$$= \binom{n_3}{s_3}\eta_3^{s_3}(1-\eta_3)^{n_3-s_3}\binom{s_3}{N_{43}}\phi_3^{N_{43}}(1-\phi_3)^{s_3-N_{43}}\binom{N_{32}}{n_{32}}p_3^{n_{32}}\times$$

$$\times(1-p_3)^{N_{32}-n_{32}}\binom{N_{32}-n_{32}}{N_{42}}\phi_3^{N_{42}}(1-\phi_3)^{N_{32}-n_{32}-N_{42}}\binom{N_{31}}{n_{31}}\times$$

$$\times p_3^{n_{31}}(1-p_3)^{N_{31}-n_{31}}\binom{N_{31}-n_{31}}{N_{41}}\phi_3^{N_{41}}(1-\phi_3)^{N_{31}-n_{31}-N_{41}}\times$$

$$\times\binom{N_{30}}{n_{30}}p_3^{n_{30}}(1-p_3)^{N_{30}-n_{30}}\binom{N_{30}-n_{30}}{N_{40}-B_3}\phi_3^{N_{40}-B_3}(1-\phi_3)^{N_{30}-n_{30}-N_{40}+B_3}. \qquad (4.27.)$$

. . . .

. . . .

. . . .

Usando (4.18.), (4.20.) e (4.22.) temos

$$P(A_{k-1}|A_{k-2}, A_{k-3}, \cdots, A_2, A_1, V_1, V_2)$$

$$= P(\{n_{k-1,j}, s_{k-1}, N_{kr}, 0 \le j \le k-2, 0 \le r \le k-1\}|A_{k-2}, A_{k-3}, \cdots, A_2, A_1, V_1, V_2)$$

$$= P(\{n_{k-1,0}, n_{k-1,1}, \cdots, n_{k-1,k-2}, s_{k-1}, N_{k0}, N_{k1}, \ldots, N_{k,k-1}|A_{k-2}, A_{k-3}, \cdots, A_2, A_1, V_1, V_2)$$

$$= P(s_{k-1}, N_{k,k-1}|n_{k-1,0}, \cdots, n_{k-1,k-2}, N_{k0}, N_{k1}, \ldots, N_{k,k-2}, A_{k-2}, A_{k-3}, \cdots, A_2, A_1, V_1, V_2)\times$$

$$\times P(n_{k-1,k-2}, N_{k,k-2}|n_{k-1,0}, \cdots, n_{k-1,k-3}, N_{k0}, N_{k1}, \ldots, N_{k,k-3}, A_{k-2}, A_{k-3}, \cdots, A_1, V_1, V_2)\times$$

$$\times \cdots \times P(n_{k-1,1}, N_{k1} | n_{k-1,0}, N_{k0}, A_{k-2}, A_{k-3}, \cdots, A_1, V_1, V_2) \times$$

$$\times P(n_{k-1,0}, N_{k0} | A_{k-2}, \cdots, A_1, V_1, V_2)$$

$$= \binom{n_{k-1}}{s_{k-1}} \eta_{k-1}^{s_{k-1}} (1-\eta_{k-1})^{n_{k-1}-s_{k-1}} \binom{s_{k-1}}{N_{k,k-1}} \phi_{k-1}^{N_{k,k-1}} (1-\phi_{k-1})^{s_{k-1}-N_{k,k-1}} \binom{N_{k-1,k-2}}{n_{k-1,k-2}} \times$$

$$\times p_{k-1}^{n_{k-1,k-2}} (1-p_{k-1})^{N_{k-1,k-2}-n_{k-1,k-2}} \binom{N_{k-1,k-2}-n_{k-1,k-2}}{N_{k,k-2}} \phi_{k-1}^{N_{k,k-2}} \times$$

$$\times (1-\phi_{k-1})^{N_{k-1,k-2}-n_{k-1,k-2}-N_{k,k-2}} \times \cdots \times \binom{N_{k-1,1}}{n_{k-1,1}} p_{k-1}^{n_{k-1,1}} (1-p_{k-1})^{N_{k-1,1}-n_{k-1,1}} \times$$

$$\times \binom{N_{k-1,1}-n_{k-1,1}}{N_{k1}} \phi_{k-1}^{N_{k1}} (1-\phi_{k-1})^{N_{k-1,1}-n_{k-1,1}-N_{k1}} \binom{N_{k-1,0}}{n_{k-1,0}} \times$$

$$\times p_{k-1}^{n_{k-1,0}} (1-p_{k-1})^{N_{k-1,0}-n_{k-1,0}} \binom{N_{k-1,0}-n_{k-1,0}}{N_{k0}-B_{k-1}} \times$$

$$\times \phi_{k-1}^{N_{k0}-B_{k-1}} (1-\phi_{k-1})^{N_{k-1,0}-n_{k-1,0}-N_{k0}+B_{k-1}}. \qquad (4.28.)$$

Usando (4.19.),(4.21.) e (4.23.) temos

$$P(A_k | A_{k-1}, \cdots, A_2, A_1, V_1, V_2) = P(\{n_{kj}, s_k, 0 \le j \le k-1\} | A_{k-1}, \cdots, A_2, A_1, V_1, V_2)$$

$$= P(n_{k0}, n_{k1}, \cdots, n_{k,k-1}, s_k | A_{k-1}, \cdots, A_2, A_1, V_1, V_2)$$

$$= P(s_k | n_{k0}, n_{k1}, \cdots, n_{k,k-1}, A_{k-1}, \cdots, A_2, A_1, V_1, V_2) \times$$

$$\times P(n_{k,k-1} | n_{k0}, n_{k1}, \cdots, n_{k,k-2}, A_{k-1}, \cdots, A_2, A_1, V_1, V_2) \times \cdots \times$$

$$\times P(n_{k0}|A_{k-1},\cdots,A_2,A_1,V_1,V_2)$$

$$= \binom{n_k}{s_k}\eta_k^{s_k}(1-\eta_k)^{n_k-s_k}\binom{N_{k,k-1}}{n_{k,k-1}}p_k^{n_{k,k-1}}(1-p_k)^{N_{k,k-1}-n_{k,k-1}}\times\cdots\times$$

$$\times\binom{N_{k1}}{n_{k1}}p_k^{n_{k1}}(1-p_k)^{N_{k1}-n_{k1}}\binom{N_{k0}}{n_{k0}}p_k^{n_{k0}}(1-p_k)^{N_{k0}-n_{k0}}. \tag{4.29.}$$

Finalmente, usando (4.24.) a (4.29.) temos,

$$P(\{n_{10},s_1,N_{20},N_{21}\};\{n_{ij},s_i,N_{i+1,r};2\leq i\leq k-1,0\leq j\leq i-1,0\leq r\leq i\};\{n_{kj},s_k,$$

$$0\leq j\leq k-1\}|\{N_{10},B_i,p_i,\phi_i,\eta_i,1\leq i\leq k-1\},\{p_k,\eta_k\})$$

$$=P(A_1,A_2,\cdots,A_k|V_1,V_2)=P(n_{10},N_{20}|V_1,V_2)P(s_1,N_{21}|n_{10},N_{20},V_1,V_2)\times$$

$$\times\prod_{i=2}^{k-1}P(s_i,N_{i+1,i}|n_{ir},N_{i+1,r},0\leq r\leq i-1,A_1,A_2,\cdots,A_{i-1},V_1,V_2)\times$$

$$\times P(n_{i0},N_{i+1,0}|A_1,A_2,\cdots,A_{i-1},V_1,V_2)\times$$

$$\times\prod_{j=1}^{i-1}P(n_{ij},N_{i+1,j}|n_{i0},n_{i1},\cdots,n_{i,j-1},N_{i+1,0},\cdots,N_{i+1,j-1},A_1,A_2,\cdots,A_{i-1},V_1,V_2)\times$$

$$\times P(s_k|n_{k0},\cdots,n_{k,k-1},A_1,A_2,\cdots,A_{k-1},V_1,V_2)\times$$

$$\times P(n_{k0}|A_1,\cdots,A_{k-1},V_1,V_2)\prod_{l=1}^{k-1}P(n_{kl}|n_{k0},n_{k1},\cdots,n_{k,l-1},A_1,A_2,\cdots,A_{k-1},V_1,V_2)$$

$$= \prod_{i=1}^{k} \binom{n_i}{s_i} \eta_i^{s_i}(1-\eta_i)^{n_i-s_i} \binom{N_{i0}}{n_{i0}} p_i^{n_{i0}}(1-p_i)^{N_{i0}-n_{i0}} \prod_{i=1}^{k-1} \binom{s_i}{N_{i+1,i}} \phi_i^{N_{i+1,i}}(1-\phi_i)^{s_i-N_{i+1,i}} \times$$

$$\times \binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i} \phi_i^{N_{i+1,0}-B_i}(1-\phi_i)^{N_{i0}-n_{i0}-N_{i+1,0}+B_i} \times$$

$$\times \prod_{i=2}^{k-1}\prod_{j=1}^{i-1} \binom{N_{ij}}{n_{ij}} p_i^{n_{ij}}(1-p_i)^{N_{ij}-n_{ij}} \binom{N_{ij}-n_{ij}}{N_{i+1,j}} \times$$

$$\times \phi_i^{N_{i+1,j}}(1-\phi_i)^{N_{ij}-n_{ij}-N_{i+1,j}} \prod_{r=1}^{k-1} \binom{N_{kr}}{n_{kr}} p_k^{n_{kr}}(1-p_k)^{N_{kr}-n_{kr}}$$

$$= \prod_{i=1}^{k} \binom{n_i}{s_i} \binom{N_{i0}}{n_{i0}} p_i^{n_i}(1-p_i)^{N_i-n_i} \eta_i^{s_i}(1-\eta_i)^{n_i-s_i} \times$$

$$\times \prod_{i=1}^{k-1} \binom{s_i}{N_{i+1,i}} \binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i} \phi_i^{N_{i+1}-B_i}(1-\phi_i)^{s_i-n_i+B_i+N_i-N_{i+1}} \times$$

$$\times \prod_{i=2}^{k-1}\prod_{j=1}^{i-1} \binom{N_{ij}}{n_{ij}} \binom{N_{ij}-n_{ij}}{N_{i+1,j}} \prod_{r=1}^{k-1} \binom{N_{kr}}{n_{kr}}$$

$$= \prod_{i=1}^{k-1}\prod_{j=1}^{i-1} C_{ij}C_{i0}C_i p_i^{n_i}(1-p_i)^{N_i-n_i} \eta_i^{s_i}(1-\eta_i)^{n_i-s_i} \phi_i^{N_{i+1}-B_i}(1-\phi_i)^{s_i-n_i+B_i+N_i-N_{i+1}} \times$$

$$\times \prod_{j=0}^{k} C_{kj} p_k^{n_k}(1-p_k)^{N_k-n_k} \eta_k^{s_k}(1-\eta_k)^{n_k-s_k}.$$

onde $C_{11} = 1$,

$$C_{ij} = \binom{N_{ij}}{n_{ij}} \binom{N_{ij}-n_{ij}}{N_{i+1,j}} = \frac{N_{ij}!}{n_{ij}!N_{i+1,j}!(N_{ij}-n_{ij}-N_{i+1,j})!},$$

$i = 2, 3, \cdots, k-1$; $j = 1, 2, \cdots, i-1$,

$$C_{i0} = \binom{N_{i0}}{n_{i0}} \binom{N_{i0}-n_{i0}}{N_{i+1,0}-B_i} = \frac{N_{i0}!}{n_{i0}!(N_{i+1,0}-B_i)!(N_{i0}-n_{i0}-N_{i+1,0}+B_i)!},$$

$i = 1, 2, \cdots, k-1,$

$$C_i = \binom{n_i}{s_i}\binom{s_i}{N_{i+1,i}} = \frac{n_i!}{(n_i - s_i)!N_{i+1,i}!(s_i - N_{i+1,i})!}, \; (1 \leq i \leq k-1)$$

$$C_{kj} = \binom{N_{kj}}{n_{kj}} = \frac{N_{kj}!}{n_{kj}(N_{kj} - n_{kj})!}, \quad (0 \leq j \leq k-1)$$

e

$$C_{kk} = \binom{n_k}{s_k} = \frac{n_k!}{s_k!(n_k - s_k)!}.$$

4.4. Estimadores de Máxima Verossimilhança.

Nesta seção determinaremos os E.M.V. para N_i, p_i, ϕ_i, B_i e η_i. Na seqüência demonstraremos uma série de teoremas, cujos resultados serão utilizados na determinação dos E.M.V.(Teorema 51). Nesta seção, $\hat{N}_{ij}$, $\hat{N}_i$, $\hat{p}_i$, $\hat{\phi}_i$, $\hat{B}_i$ e $\hat{\eta}_i$ denotam os E.M.V. para $\hat{N}_{ij}$, N_i, p_i, ϕ_i, B_i e η_i, respectivamente.

Teorema 41.

Os E.M.V. para N_{ij}, $\hat{\phi}_i$ e $\hat{p}_i$, satisfazem as relações

$$\frac{\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}}{\hat{N}_{ij} - n_{ij} - \hat{N}_{i+1,j}} \cong \frac{1 - \hat{\phi}_{i-1}}{(1 - \hat{p}_i)\hat{\phi}_{i-1}(1 - \hat{\phi}_i),} \tag{4.30.}$$

$(3 \leq i \leq k-1, \; 1 \leq j \leq i-2)$;

$$\frac{\hat{N}_{k-1,j} - n_{k-1,j} - \hat{N}_{kj}}{\hat{N}_{kj} - n_{kj}} \cong \frac{1 - \hat{\phi}_{k-1}}{\hat{\phi}_{k-1}(1 - \hat{p}_k)}, \quad (1 \leq j \leq k-2); \tag{4.31.}$$

$$\frac{s_{i-1} - \hat{N}_{i,i-1}}{\hat{N}_{i,i-1} - n_{i,i-1} - \hat{N}_{i+1,i-1}} \cong \frac{1 - \hat{\phi}_{i-1}}{\hat{\phi}_{i-1}(1 - \hat{\phi}_i)(1 - \hat{p}_i)}, \quad (2 \leq i \leq k-1); \tag{4.32.}$$

e

$$\frac{s_{k-1} - \hat{N}_{k,k-1}}{\hat{N}_{k,k-1} - n_{k,k-1}} \cong \frac{1 - \hat{\phi}_{k-1}}{\hat{\phi}_{k-1}(1 - \hat{p}_k)}. \tag{4.33.}$$

Prova. A função de verossimilhança é dada por

$$L[\{N_i, p_i, \eta_i, 1 \leq i \leq k; B_j, \phi_j, 1 \leq j \leq k-1\} | \{n_i, s_i, 1 \leq i \leq k\}]$$

$$= \prod_{i=1}^{k-1} \prod_{j=1}^{i-1} C_{ij} C_{i0} C_i p_i^{n_i} (1-p_i)^{N_i - n_i} \phi_i^{N_{i+1} - B_i} (1-\phi_i)^{N_i - N_{i+1} + B_i - n_i + s_i} \times$$

$$\times \eta_i^{s_i} (1-\eta_i)^{n_i - s_i} \prod_{j=0}^{k} C_{kj} p_k^{n_k} (1-p_k)^{N_k - n_k} \eta_k^{s_k} (1-\eta_k)^{n_k - s_k}. \qquad (4.34.)$$

onde C_{ij}, $(i = 1, 2, \cdots, k \ ; \ j = 0, 1, \cdots, i)$ e $C_i, (i = 1, 2, \cdots, k-1.)$ são definidos como no final da seção anterior.

Para maximizar a função de verossimilhança,L, igualamos $L(\theta)$ a $L(\theta - 1)$, onde $\theta = N_{ij}$, $(3 \leq i \leq k-1, \ 1 \leq j \leq i-2)$. Logo,

$$\frac{L(N_{ij})}{L(N_{ij} - 1)} = 1 \Longrightarrow$$

$$\frac{\phi_{i-1}^{N_{ij}} (1-\phi_{i-1})^{-N_{ij}} / (N_{ij} - n_{ij} - N_{i+1,j})! (N_{i-1,j} - n_{i-1,j} - N_{ij})!}{\phi_{i-1}^{N_{ij}-1} (1-\phi_{i-1})^{-N_{ij}+1} / (N_{ij} - n_{ij} - N_{i+1,j} - 1)! (N_{i-1,j} - n_{i-1,j} - N_{ij} + 1)!} \times$$

$$\times \frac{N_{ij}! (1-p_i)^{N_{ij}} (1-\phi_i)^{N_{ij}} / N_{ij}!}{(N_{ij}-1)! (1-p_i)^{N_{ij}-1} (1-\phi_i)^{N_{ij}-1} / (N_{ij}-1)!} = 1$$

$$\Longrightarrow \frac{N_{i-1,j} - n_{i-1,j} - N_{ij} + 1}{N_{ij} - n_{ij} - N_{i+1,j}} = \frac{1 - \phi_{i-1}}{(1-p_i)\phi_{i-1}(1-\phi_i)}.$$

Supondo que estamos tratando com grandes populações substituimos $N_{i-1,j} - n_{i-1,j} - N_{ij} + 1$ por $N_{i-1,j} - n_{i-1,j} - N_{ij}$ na última equação obtendo

$$\frac{\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}}{\hat{N}_{ij} - n_{ij} - \hat{N}_{i+1,j}} \cong \frac{1 - \hat{\phi}_{i-1}}{(1-\hat{p}_i)\hat{\phi}_{i-1}(1-\hat{\phi}_i)}, \quad (3 \leq i \leq k-1, \ 1 \leq j \leq i-2)$$

o que prova a relação (4.30.).

De

$$\frac{L(N_{kj})}{L(N_{kj} - 1)} = 1, \ (1 \leq j \leq k-2)$$

segue que

$$\frac{(1-\phi_{k-1})^{-N_{kj}} N_{kj}! (1-p_k)^{N_{kj}} / (N_{k-1,j} - n_{k-1,j} - N_{kj})!}{(1-\phi_{k-1})^{-N_{kj}+1} (N_{kj}-1)! (1-p_k)^{N_{kj}-1} / (N_{k-1,j} - n_{k-1,j} - N_{kj} + 1)!} \times$$

$$\times \frac{\phi_{k-1}^{N_{kj}}/(N_{kj}-n_{kj})!N_{kj}!}{\phi_{k-1}^{N_{kj}-1}/(N_{kj}-n_{kj}-1)!(N_{kj}-1)!} = 1$$

$$\Longrightarrow \frac{N_{k-1,j}-n_{k-1,j}-N_{kj}+1}{N_{kj}-n_{kj}} = \frac{1-\phi_{k-1}}{\phi_{k-1}(1-p_k)}.$$

Substituindo $N_{k-1,j}-n_{k-1,j}-N_{kj}+1$ por $N_{k-1,j}-n_{k-1,j}-N_{kj}$ na última equação segue que

$$\frac{\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj}}{\hat{N}_{kj}-n_{kj}} \cong \frac{1-\hat{\phi}_{k-1}}{\hat{\phi}_{k-1}(1-\hat{p}_k)}, \quad (1 \le j \le k-2)$$

o que prova a relação (4.31.).

De

$$\frac{L(N_{i,i-1})}{L(N_{i,i-1}-1)} = 1$$

segue que

$$\frac{(1-\phi_{i-1})^{-N_{i,i-1}}N_{i,i-1}!(1-p_i)^{N_{i,i-1}}/(N_{i,i-1}-n_{i,i-1}-N_{i+1,i-1})!}{(1-\phi_{i-1})^{-N_{i,i-1}+1}(N_{i,i-1}-1)!(1-p_i)^{N_{i,i-1}-1}/(N_{i,i-1}-n_{i,i-1}-N_{i+1,i-1}-1)!} \times$$

$$\times \frac{\phi_{i-1}^{N_{i,i-1}}(1-\phi_i)^{N_{i,i-1}}/(s_{i-1}-N_{i,i-1})!N_{i,i-1}!}{\phi_{i-1}^{N_{i,i-1}-1}(1-\phi_i)^{N_{i,i-1}-1}/(s_{i-1}-N_{i,i-1}+1)!(N_{i,i-1}-1)!} = 1$$

$$\Longrightarrow \frac{s_{i-1}-N_{i,i-1}+1}{N_{i,i-1}-n_{i,i-1}-N_{i+1,i-1}} = \frac{1-\phi_{i-1}}{\phi_{i-1}(1-\phi_i)(1-p_i)}.$$

Supondo que estamos tratando com grandes populações, substituimos na última equação $s_{i-1}-N_{i,i-1}+1$ por $s_{i-1}-N_{i,i-1}$. Logo,

$$\frac{s_{i-1}-\hat{N}_{i,i-1}}{\hat{N}_{i,i-1}-n_{i,i-1}-\hat{N}_{i+1,i-1}} \cong \frac{1-\hat{\phi}_{i-1}}{\hat{\phi}_{i-1}(1-\hat{\phi}_i)(1-\hat{p}_i)}, \quad (2 \le i \le k-1)$$

o que prova a relação (4.32.).

De modo análogo, de

$$\frac{L(N_{k,k-1})}{L(N_{k,k-1}-1)} = 1$$

segue que

$$\frac{(1-\phi_{k-1})^{-N_{k,k-1}}N_{k,k-1}!(1-p_k)^{N_{k,k-1}}/(N_{k,k-1}-n_{k,k-1})!}{(1-\phi_{k-1})^{-N_{k,k-1}+1}(N_{k,k-1}-1)!(1-p_k)^{N_{k,k-1}-1}/(N_{k,k-1}-n_{k,k-1}-1)!}\times$$

$$\times\frac{\phi_{k-1}^{N_{k,k-1}}/(s_{k-1}-N_{k,k-1})!N_{k,k-1}!}{\phi_{k-1}^{N_{k,k-1}-1}/(s_{k-1}-N_{k,k-1}+1)!(N_{k,k-1}-1)!}=1$$

$$\Longrightarrow \frac{s_{k-1}-N_{k,k-1}+1}{N_{k,k-1}-n_{k,k-1}}=\frac{1-\phi_{k-1}}{\phi_{k-1}(1-p_k)}.$$

Substituindo $s_{k-1}-N_{k,k-1}+1$ por $s_{k-1}-N_{k,k-1}$ na última equação obtemos

$$\frac{s_{k-1}-\hat{N}_{k,k-1}}{\hat{N}_{k,k-1}-n_{k,k-1}}\cong\frac{1-\hat{\phi}_{k-1}}{\hat{\phi}_{k-1}(1-\hat{p}_k)}$$

e o teorema está provado. ∎

O teorema seguinte relaciona $\hat{N}_{ij}$ com $\hat{\phi}_i$.

Teorema 42. Os E.M.V para N_{ij} e ϕ_i satisfazem as equações

$$\hat{N}_{ij}-n_{ij}-\hat{N}_{i+1,j}\cong\frac{(1-\hat{\phi}_i)\left(\hat{N}_{ij}-\sum_{r=i}^{k}n_{rj}\right)}{A_i},\ (2\le i\le k-1,\ 1\le j\le i-1) \quad (4.35.)$$

onde $A_k=1$ e

$$A_i=1-\hat{\phi}_i[1-(1-\hat{p}_{i+1})A_{i+1}]=1-\hat{\phi}_i[\hat{p}_{i+1}+(1-\hat{p}_{i+1})(1-A_{i+1})]. \quad (4.36.)$$

Prova. Provaremos este teorema por indução, isto é, verificaremos que ele vale para $i=k-1$ e, supondo que ele vale para todo i $(3\le i\le k-1)$, provamos que ele vale para $i-1$.

Da relação (4.31.) segue que

$$\hat{\phi}_{k-1}(\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj})-\hat{\phi}_{k-1}\hat{p}_k(\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj})+(\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj})$$

$$\cong\hat{N}_{kj}-n_{kj}-\hat{N}_{kj}\hat{\phi}_{k-1}+\hat{\phi}_{k-1}n_{kj}+\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj}$$

$$\Longrightarrow(1-\hat{\phi}_{k-1}\hat{p}_k)(\hat{N}_{k-1,j}-n_{k-1,j}-\hat{N}_{kj})\cong(1-\hat{\phi}_{k-1})(\hat{N}_{k-1,j}-n_{k-1,j}-n_{kj})$$

$$\Longrightarrow \hat{N}_{k-1,j} - n_{k-1,j} - \hat{N}_{kj} \cong \frac{(1-\hat{\phi}_{k-1})}{(1-\hat{\phi}_{k-1}\hat{p}_k)}(\hat{N}_{k-1,j} - \sum_{r=k-1}^{k} n_{rj}), \qquad (4.37.)$$

para $1 \leq j \leq k-2$, isto é, (4.35.) vale para $i = k-1$.

Supomos agora que (4.35.) vale para todo i $(3 \leq i \leq k-1)$. Usando este fato e a relação (4.30.) segue que $\forall\, i$, $\forall j$, $3 \leq i \leq k-1$, $1 \leq j \leq i-2$,

$$(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij})(1-\hat{p}_i)\hat{\phi}_{i-1}(1-\hat{\phi}_i)[1-\hat{\phi}_i + \hat{\phi}_i(1-\hat{p}_{i+1})A_{i+1}] \cong$$

$$\cong (1-\hat{\phi}_{i-1})(1-\hat{\phi}_i)\left(\hat{N}_{ij} - \sum_{r=i}^{k} n_{rj}\right)$$

$$\Longrightarrow (\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij})\{1-\hat{\phi}_{i-1}[1-(1-p_i)A_i]\} \cong \hat{N}_{ij} - \sum_{r=i}^{k} n_{rj} - \hat{\phi}_{i-1}\hat{N}_{ij} +$$

$$+\hat{\phi}_{i-1}\sum_{r=i}^{k} n_{rj} - \hat{\phi}_{i-1}(1-\hat{p}_i)(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) + \hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) -$$

$$-\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(1-\hat{p}_{i+1})A_{i+1}(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) + \hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij} - \hat{\phi}_{i-1}(\hat{N}_{i-1,j} -$$

$$-n_{i-1,j} - \hat{N}_{ij}) + \hat{\phi}_{i-1}(1-\hat{p}_i)A_i(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) = (1-\hat{\phi}_{i-1})(\hat{N}_{i-1,j} - \sum_{r=i-1}^{k} n_{rj}) -$$

$$-\hat{\phi}_{i-1}(1-\hat{p}_i)(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) + \hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) -$$

$$-\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(1-\hat{p}_{i+1})A_{i+1}(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) + \hat{\phi}_{i-1}(1-\hat{p}_i)[1-\hat{\phi}_i +$$

$$+\hat{\phi}_i(1-p_{i+1})A_{i+1}](\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij}) = (1-\hat{\phi}_{i-1})\Big(\hat{N}_{i-1,j}-\sum_{r=i-1}^{k} n_{rj}\Big)-$$

$$-\hat{\phi}_{i-1}(1-\hat{p}_i)(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})+\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})-$$

$$-\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(1-\hat{p}_{i+1})A_{i+1}(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})+\hat{\phi}_{i-1}(1-\hat{p}_i)(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})-$$

$$-\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})+\hat{\phi}_{i-1}\hat{\phi}_i(1-\hat{p}_i)(1-\hat{p}_{i+1})A_{i+1}(\hat{N}_{i-1,j}-n_{i-1,j}-\hat{N}_{ij})$$

$$= (1-\hat{\phi}_{i-1})\Big(\hat{N}_{i-1,j}-\sum_{r=i-1}^{k} n_{rj}\Big),$$

isto é, (4.35.) vale para $i-1$. ∎

Teorema 43. Os E.M.V. para N_i, N_{i0}, p_i, η_i, ϕ_i e B_i satisfazem as equações

$$\hat{\phi}_i = \frac{\hat{N}_{i+1}-\hat{B}_i}{\hat{N}_i-n_i+s_i}, \ (1\le i\le k-1), \tag{4.38.}$$

$$\hat{p}_i = \frac{n_i}{\hat{N}_i}, \ (1\le i\le k), \tag{4.39.}$$

$$\hat{\eta}_i = \frac{s_i}{n_i}, \ (1\le i\le k) \tag{4.40.}$$

e

$$\hat{B}_i \cong \hat{N}_{i+1,0}-\hat{\phi}_i(\hat{N}_{i0}-n_{i0}), \ \ (1\le i\le k-1). \tag{4.41.}$$

Prova. Da relação (4.34.) segue que

$$\ln L[\{N_i,p_i,\eta_i,1\le i\le k;B_j,\phi_j,1\le j\le k-1\}|\{n_i,s_i,1\le i\le k\}]$$

$$=\sum_{i=1}^{k}[\ln n_i!-\ln s_i!-\ln(n_i-s_i)!+\ln N_{i0}!-\ln n_{i0}!-\ln(N_{i0}-n_{i0})!+n_i\ln p_i+(N_i-n_i)\ln(1-p_i)+$$

$$+s_i \ln \eta_i + (n_i - s_i)\ln(1-\eta_i)] + \sum_{i=1}^{k-1}[\ln s_i! - \ln(s_i - N_{i+1,i})! - \ln N_{i+1,i}! + \ln(N_{i0} - n_{i0})! -$$

$$- \ln(N_{i+1,0} - B_i)! - \ln(N_{i0} - n_{i0} - N_{i+1,0} + B_i)! + (N_{i+1} - B_i)\ln\phi_i + (s_i - n_i + B_i + N_i -$$

$$-N_{i+1})\ln(1-\phi_i)] + \sum_{i=2}^{k-1}\sum_{j=1}^{i-1}[\ln N_{ij}! - \ln n_{ij}! - \ln N_{i+1,j}! - \ln(N_{ij} - n_{ij} - N_{i+1,j})! +$$

$$+\sum_{r=1}^{k-1}[\ln N_{kr}! - \ln n_{kr}! - \ln(N_{kr} - n_{kr})!].$$

Logo,

$$\begin{cases} \frac{\partial \ln L}{\partial \phi_i} = \frac{N_{i+1}-B_i}{\phi_i} - \frac{N_i - N_{i+1} + B_i - n_i + s_i}{1-\phi_i} = 0 \\ \frac{\partial \ln L}{\partial p_i} = \frac{n_i}{p_i} - \frac{N_i - n_i}{1-p_i} = 0 \\ \frac{\partial \ln L}{\partial \eta_i} = \frac{s_i}{\eta_i} - \frac{n_i - s_i}{1-\eta_i} = 0, \end{cases}$$

ou

$$\begin{cases} (1-\phi_i)(N_{i+1} - B_i) - \phi_i(N_i - N_{i+1} + B_i - n_i + s_i) = 0 \\ n_i - n_i p_i - p_i N_i + n_i p_i = 0 \\ s_i - s_i\eta_i - n_i\eta_i + s_i\eta_i = 0, \end{cases}$$

ou

$$\begin{cases} N_{i+1} - B_i - \phi_i N_{i+1} + \phi_i B_i - \phi_i N_i + \phi_i N_{i+1} - \phi_i B_i + \phi_i n_i - \phi_i s_i = 0 \Rightarrow \\ \Rightarrow \hat{\phi}_i = \frac{\hat{N}_{i+1} - \hat{B}_i}{\hat{N}_i - n_i + s_i} \quad (1 \le i \le k-1) \\ \hat{p}_i = \frac{n_i}{\hat{N}_i} \quad (1 \le i \le k) \\ \hat{\eta}_i = \frac{s_i}{n_i} \quad (1 \le i \le k), \end{cases}$$

o que prova as relações (4.38.), (4.39.) e (4.40.). Com relação a $\hat{B}_i$, $(1 \le i \le k-1)$ temos

$$\Delta \ln L(B_i) = \ln L(B_i) - \ln L(B_i - 1) = 0 \implies$$

$$\implies \ln \frac{L(B_i)}{L(B_i-1)} = 0 \implies \frac{L(B_i)}{L(B_i-1)} = 1$$

$$\implies \frac{\phi_i^{-B_i}(1-\phi_i)^{B_i}/(N_{i+1,0}-B_i)!(N_{i0}-n_{i0}-N_{i+1,0}+B_i)!}{\phi_i^{-B_i+1}(1-\phi_i)^{B_i-1}/(N_{i+1,0}-B_i+1)!(N_{i0}-n_{i0}-N_{i+1,0}+B_i-1)!} = 1$$

$$\implies \frac{(1-\phi_i)}{\phi_i}\frac{(N_{i+1,0}-B_i+1)}{(N_{i0}-n_{i0}-N_{i+1,0}+B_i)} = 1$$

$$\implies \frac{N_{i+1,0}-B_i+1}{N_{i0}-n_{i0}-N_{i+1,0}+B_i} = \frac{\phi_i}{1-\phi_i}.$$

Substituindo $N_{i+1,0} - B_i + 1$ por $N_{i+1,0} - B_i$ na última equação temos

$$\frac{N_{i+1,0}-B_i}{N_{i0}-n_{i0}-N_{i+1,0}+B_i} \cong \frac{\phi_i}{1-\phi_i}$$

$$\implies N_{i+1,0} - N_{i+1,0}\phi_i - B_i + B_i\phi_i \cong \phi_i N_{i0} - \phi_i n_{i0} - \phi_i N_{i+1,0} + B_i\phi_i$$

$$\implies \hat{B}_i \cong \hat{N}_{i+1,0} - \hat{\phi}_i(\hat{N}_{i0} - n_{i0}) \; (1 \le i \le k-1),$$

o que prova o teorema. ∎

Teorema 44. Os E.M.V. para N_{i0} e p_i satisfazem as equações

$$\hat{N}_{i0} \cong \frac{n_{i0}}{\hat{p}_i}, \quad (1 \le i \le k). \tag{4.42.}$$

Prova. Como $\hat{N}_i = \sum_{j=0}^{i-1} \hat{N}_{ij}$, $(1 \le i \le k)$, temos

$$\ln \frac{L(N_{i0})}{L(N_{i0}-1)} = 0 \;\; (1 \le i \le k-1) \implies \frac{L(N_{i0})}{L(N_{i0}-1)} = 1$$

$$\implies \frac{(1\text{-}\phi_{i\text{-}1})^{-N_{i0}}/(N_{i0}\text{-}n_{i0}\text{-}N_{i+1,0}+B_i)!(N_{i\text{-}1,0}\text{-}n_{i\text{-}1,0}\text{-}N_{i0}+B_{i\text{-}1})!}{(1\text{-}\phi_{i\text{-}1})^{-N_{i0}+1}/(N_{i0}\text{-}n_{i0}\text{-}N_{i+1,0}+B_i\text{-}1)!(N_{i-1,0}\text{-}n_{i-1,0}\text{-}N_{i0}+B_{i-1}+1)!} \times$$

$$\times \frac{{\phi_{i-1}}^{N_{i0}} N_{i0}!(1-p_i)^{N_{i0}}(1-\phi_i)^{N_{i0}}/(N_{i0}-B_{i-1})!}{{\phi_{i-1}}^{N_{i0}-1}(N_{i0}-1)!(1-p_i)^{N_{i0}-1}(1-\phi_i)^{N_{i0}-1}/(N_{i0}-B_{i-1}-1)!} = 1$$

$$\Longrightarrow \frac{N_{i0}(1-p_i)(1-\phi_i)\phi_{i-1}(N_{i-1,0}-n_{i-1,0}-N_{i0}+B_{i-1}+1)}{(N_{i0}-B_{i-1})(N_{i0}-n_{i0}-N_{i+1,0}+B_i)(1-\phi_{i-1})} = 1.$$

Substituindo $N_{i-1,0}-n_{i-1,0}-N_{i0}+B_{i-1}+1$ por $N_{i-1,0}-n_{i-1,0}-N_{i0}+B_{i-1}$ na última equação e usando a relação (4.41.) temos:

$$N_{i0}(1-p_i)(1-\phi_i)\phi_{i-1}[N_{i-1,0}-n_{i-1,0}-N_{i0}+N_{i0}-\phi_{i-1}(N_{i-1,0}-n_{i-1,0})] \cong$$

$$\cong [N_{i0}-N_{i0}+\phi_{i-1}(N_{i-1,0}-n_{i-1,0})](N_{i0}-n_{i0}-N_{i+1,0}+N_{i+1,0}-\phi_i(N_{i0}-n_{i0})](1-\phi_{i-1})$$

$$\Longrightarrow N_{i0}(1\text{-}p_i)(1\text{-}\phi_i)(N_{i\text{-}1,0}\text{-}n_{i\text{-}1,0})(1\text{-}\phi_{i\text{-}1}) \cong (N_{i\text{-}1,0}\text{-}n_{i\text{-}1,0})(N_{i0}\text{-}n_{i0})(1\text{-}\phi_i)(1\text{-}\phi_{i\text{-}1})$$

$$\Longrightarrow N_{i0}(1-p_i) \cong N_{i0}-n_{i0} \Longleftrightarrow N_{i0}(1-1+p_i) \cong n_{i0}$$

$$\Longrightarrow \hat{N}_{i0} \cong \frac{n_{i0}}{\hat{p}_i},\ (1 \le i \le k-1). \tag{4.43.}$$

De $\Delta \ln L(N_{k0}) = 0$ segue que

$$\ln \frac{L(N_{k0})}{L(N_{k0}-1)} = 0 \Longrightarrow \frac{L(N_{k0})}{L(N_{k0}-1)} = 1$$

$$\Longrightarrow \frac{(1-\phi_{k-1})^{-N_{k0}}/(N_{k0}-n_{k0})!(N_{k-1,0}-n_{k-1,0}-N_{k0}+B_{k-1})!}{(1-\phi_{k-1})^{-N_{k0}+1}/(N_{k0}-n_{k0}-1)!(N_{k-1,0}-n_{k-1,0}-N_{k0}+B_{k-1}+1)!}\times$$

$$\times \frac{{\phi_{k-1}}^{N_{k0}} N_{k0}!(1-p_k)^{N_{k0}}/(N_{k0}-B_{k-1})!}{{\phi_{k-1}}^{N_{k0}-1}(N_{k0}-1)!(1-p_k)^{N_{k0}-1}/(N_{k0}-B_{k-1}-1)!} = 1$$

$$\Longrightarrow N_{k0}(1-p_k)\phi_{k-1}(N_{k-1,0}-n_{k-1,0}-N_{k0}+B_{k-1}+1) = (N_{k0}-B_{k-1})(N_{k0}-n_{k0})(1-\phi_{k-1}).$$

Substituindo $N_{k-1,0} - n_{k-1,0} - N_{k0} + B_{k-1} + 1$ por $N_{k-1,0} - n_{k-1,0} - N_{k0} + B_{k-1}$ na última equação e usando a relação (4.41.) temos:

$$(N_{k-1,0} - n_{k-1,0} - N_{k0} + N_{k0} - \phi_{k-1}(N_{k-1,0} - n_{k-1,0}))N_{k0}(1 - p_k)\phi_{k-1} \cong$$

$$\cong (N_{k0} - n_{k0})(1 - \phi_{k-1})(N_{k0} - N_{k0} + \phi_{k-1}(N_{k-1,0} - n_{k-1,0}))$$

$$\Longrightarrow (N_{k-1,0} - n_{k-1,0})(1 - \phi_{k-1})N_{k0}(1 - p_k)\phi_{k-1} \cong$$

$$\cong (N_{k0} - n_{k0})(1 - \phi_{k-1})\phi_{k-1}(N_{k-1,0} - n_{k-1,0})$$

$$\Longrightarrow \quad N_{k0}p_k \cong n_{k0} \quad \Longrightarrow \quad \hat{N}_{k0} \cong \frac{n_{k0}}{\hat{p}_k}.$$

De (4.43.) e da última relação segue que

$$\hat{N}_{i0} \cong \frac{n_{i0}}{\hat{p}_i}, \ (1 \leq i \leq k)$$

o que prova o teorema. ∎

Teorema 45. Os E.M.V. para M_i, N_i, ϕ_i e $N_{i+1,i}$ satisfazem as relações

$$\hat{\phi}_i \cong \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + s_i}, \ (1 \leq i \leq k-1) \tag{4.44.}$$

$$\hat{N}_i = \frac{n_i \hat{M}_i}{m_i}, \ (2 \leq i \leq k) \tag{4.45.}$$

$$\hat{N}_{i+1,i} \cong \hat{\phi}_i s_i, \quad (1 \leq i \leq k-1). \tag{4.46.}$$

Prova. Das relações (4.30.) e (4.32.) segue que

$$\frac{\sum\limits_{j=1}^{i-2}(\hat{N}_{i-1,j} - n_{i-1,j} - \hat{N}_{ij}) + s_{i-1} - \hat{N}_{i,i-1}}{\sum\limits_{j=1}^{i-2}(\hat{N}_{ij} - n_{ij} - \hat{N}_{i+1,j}) + N_{i,i-1} - n_{i,i-1} - N_{i+1,i-1}} \cong \frac{1 - \hat{\phi}_{i-1}}{\hat{\phi}_{i-1}(1 - \hat{\phi}_i)(1 - \hat{p}_i)},$$

para $3 \leq i \leq k-1$.

Como $\sum_{j=1}^{i-2} \hat{N}_{i-1,j} = \hat{M}_{i-1}$, $\sum_{j=1}^{i-2} n_{i-1,j} = m_{i-1}$, $\sum_{j=1}^{i-1} \hat{N}_{ij} = \hat{M}_i$ e $\sum_{j=1}^{i} \hat{N}_{i+1,j} = \hat{M}_{i+1}$, da equação anterior resulta

$$\frac{\hat{M}_{i-1} - m_{i-1} + s_{i-1} - \hat{M}_i}{\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i}} \cong \frac{1-\hat{\phi}_{i-1}}{\hat{\phi}_{i-1}(1-\hat{\phi}_i)(1-\hat{p}_i)}, \quad (3 \leq i \leq k-1). \qquad (4.47.)$$

Como $M_1 = m_1 = 0$, $\hat{M}_2 = \hat{N}_{21}$, $n_{21} = m_2$ e $\hat{M}_3 - \hat{N}_{32} = \hat{N}_{31}$, da relação (4.32.) para $i = 2$, resulta

$$\frac{M_1 - m_1 + s_1 - \hat{M}_2}{\hat{M}_2 - m_2 - \hat{M}_3 + \hat{N}_{32}} \cong \frac{1-\hat{\phi}_1}{\hat{\phi}_1(1-\hat{\phi}_2)(1-\hat{p}_2)}.$$

Logo,

$$\frac{\hat{M}_{i-1} - m_{i-1} + s_{i-1} - \hat{M}_i}{\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i}} \cong \frac{1-\hat{\phi}_{i-1}}{\hat{\phi}_{i-1}(1-\hat{\phi}_i)(1-\hat{p}_i)}, \quad (2 \leq i \leq k-1). \qquad (4.48.)$$

Das relações (4.38.) e (4.41.) resulta

$$\hat{\phi}_i \cong \frac{\hat{N}_{i+1} - \hat{N}_{i+1,0} + \hat{\phi}_i(\hat{N}_{i0} - n_{i0})}{\hat{N}_i - n_i + s_i}$$

$$\Longrightarrow \hat{\phi}_i(\hat{N}_i - n_i + s_i - \hat{N}_{i0} + n_{i0}) \cong \hat{N}_{i+1} - \hat{N}_{i+1,0}$$

$$\Longrightarrow \hat{\phi}_i \cong \frac{\hat{M}_{i+1}}{\hat{M}_i - (n_i - n_{i0}) + s_i}$$

$$\Longrightarrow \hat{\phi}_i \cong \frac{\hat{M}_{i+1}}{\hat{M}_i - m_i + s_i}, \quad (1 \leq i \leq k-1),$$

o que prova a relação (4.44.).

Notemos que esta relação é exatamente a relação (4.2.) obtida na seção 4.2.

Das relações (4.39.), (4.44.) e (4.48.) segue que

$$\frac{\hat{M}_{i-1} - m_{i-1} + s_{i-1} - \hat{M}_i}{\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i}} \cong$$

$$\cong \frac{1 - \hat{M}_i/(\hat{M}_{i-1} - m_{i-1} + s_{i-1})}{\Big(\hat{M}_i/(\hat{M}_{i-1} - m_{i-1} + s_{i-1})\Big)\Big(1 - \hat{M}_{i+1}/(\hat{M}_i - m_i + s_i)\Big)\Big(1 - n_i/\hat{N}_i\Big)}$$

$$= \frac{(\hat{M}_{i-1} - m_{i-1} + s_{i-1} - \hat{M}_i)/(\hat{M}_{i-1} - m_{i-1} + s_{i-1})}{\hat{M}_i(\hat{M}_i - m_i + s_i - \hat{M}_{i+1})(\hat{N}_i - n_i)/(\hat{M}_{i-1} - m_{i-1} + s_{i-1})(\hat{M}_i - m_i + s_i)\hat{N}_i}$$

$$\Longrightarrow \frac{1}{\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i}} \cong \frac{\hat{N}_i(\hat{M}_i - m_i + s_i)}{\hat{M}_i(\hat{N}_i - n_i)(\hat{M}_i - m_i + s_i - \hat{M}_{i+1})}$$

$$\Longrightarrow \frac{(\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i})(\hat{M}_i - m_i + s_i)\hat{N}_i}{(\hat{M}_i - m_i + s_i - \hat{M}_{i+1})(\hat{N}_i - n_i)\hat{M}_i} \cong 1, \ (1 \leq i \leq k-1). \qquad (4.49.)$$

Como $\hat{N}_i = \hat{M}_i + \hat{N}_{i0}$, $(1 \leq i \leq k)$, segue desta última relação e da relação (4.42.) que $\hat{N}_i = \hat{M}_i + \frac{n_{i0}}{\hat{p}_i}$, $(1 \leq i \leq k)$. Desta última relação e da relação (4.39.) temos

$$\hat{N}_i = \hat{M}_i + \frac{n_{i0}\hat{N}_i}{n_i} \Longrightarrow \hat{N}_i(1 - n_{i0}/n_i) = \hat{M}_i$$

$$\Longrightarrow \hat{N}_i\left(\frac{n_i - n_{i0}}{n_i}\right) = \hat{M}_i$$

$$\Longrightarrow \hat{N}_i = \frac{n_i\hat{M}_i}{m_i}, \ (2 \leq i \leq k),$$

o que prova a relação (4.45.). Observemos que esta relação é exatamente a relação (4.1.) obtida na seção 4.2.

A última relação implica

$$\frac{\hat{N}_i - n_i}{\hat{N}_i} = \frac{\hat{M}_i - m_i}{\hat{M}_i}, \ (2 \leq i \leq k).$$

Usando a última relação e a relação (4.49.) temos

$$(\hat{M}_i - m_i - \hat{M}_{i+1} + \hat{N}_{i+1,i})(\hat{M}_i - m_i + s_i) \cong (\hat{M}_i - m_i + s_i - \hat{M}_{i+1})(\hat{M}_i - m_i)$$

$$\Longrightarrow \hat{N}_{i+1,i}(\hat{M}_i\text{-}m_i + s_i) \cong (\hat{M}_i\text{-}m_i + s_i\text{-}\hat{M}_{i+1})(\hat{M}_i\text{-}m_i) + (\hat{M}_i\text{-}m_i + s_i)(m_i + \hat{M}_{i+1}\text{-}\hat{M}_i)$$

$$\Longrightarrow \hat{N}_{i+1,i} \cong m_i + \hat{M}_{i+1} - \hat{M}_i + \hat{M}_i - m_i - \hat{M}_{i+1}\frac{(\hat{M}_i - m_i)}{\hat{M}_i - m_i + s_i}$$

$$\Longrightarrow \hat{N}_{i+1,i} \cong \hat{M}_{i+1}\left(1 - \frac{\hat{M}_i - m_i}{\hat{M}_i - m_i + s_i}\right)$$

$$\Longrightarrow \hat{N}_{i+1,i} \cong \frac{\hat{M}_{i+1}s_i}{\hat{M}_i - m_i + s_i}.$$

Usando a relação (4.44.) e a última relação temos

$$\hat{N}_{i+1,i} \cong \hat{\phi}_i s_i, \quad (2 \leq i \leq k-1).$$

Como $\hat{N}_{21} = \hat{M}_2$ e $\hat{M}_1 = m_1 = 0$, então

$$\hat{N}_{21}(\hat{M}_1 - m_1 + s_1) = \hat{M}_2 s_1$$

$$\Longrightarrow \hat{N}_{21} = \frac{\hat{M}_2 s_1}{\hat{M}_1 - m_1 + s_1}.$$

Usando a relação (4.44.) para $i = 1$ e a última relação, resulta $\hat{N}_{21} \cong \hat{\phi}_1 s_1$. Logo,

$$\hat{N}_{i+1,i} \cong \hat{\phi}_i s_i, \quad (1 \leq i \leq k-1),$$

o que prova o teorema. ∎

Teorema 46. Os E.M.V. para N_{ij}, M_i e ϕ_i satisfazem as relações

$$\sum_{j=1}^{i-1} \hat{N}_{i+1,j} \cong \hat{M}_{i+1} - \hat{\phi}_i s_i, \qquad (2 \leq i \leq k-1).$$

Prova. Como

$$\sum_{j=1}^{i-1} \hat{N}_{i+1,j} + \hat{N}_{i+1,i} = \hat{M}_{i+1} \quad (2 \leq i \leq k-1).$$

$$\Longrightarrow \sum_{j=1}^{i-1} \hat{N}_{i+1,j} = \hat{M}_{i+1} - \hat{N}_{i+1,i} \quad (2 \leq i \leq k-1)$$

Logo, da relação (4.46.) segue que

$$\sum_{j=1}^{i-1} \hat{N}_{i+1,j} = \hat{M}_{i+1} - \hat{N}_{i+1,i} \cong \hat{M}_{i+1} - \hat{\phi}_i s_i, \ (2 \leq i \leq k-1),$$

o que prova o teorema. ■

Teorema 47. Para todo i, $2 \leq i \leq k-1$,

$$\sum_{j=1}^{i-1} \sum_{r=i}^{k} n_{rj} = m_i + Z_i.$$

Prova. Como $Z_i = \sum\limits_{r=i+1}^{k} \sum\limits_{j=1}^{i-1} n_{rj}$ segue que

$$\sum_{j=1}^{i-1} \sum_{r=i}^{k} n_{rj} = \sum_{r=i}^{k} \sum_{j=1}^{i-1} n_{rj} = \sum_{j=1}^{i-1} n_{ij} + \sum_{r=i+1}^{k} \sum_{j=1}^{i-1} n_{rj} = m_i + Z_i, \quad (2 \leq i \leq k-1).$$

■

Teorema 48. Para todo i, $(2 \leq i \leq k-1)$,

$$R_i = m_{i+1} + Z_{i+1} - Z_i.$$

Prova. Pelo Teorema 47 temos

$$\begin{aligned} m_{i+1} + Z_{i+1} - Z_i &= m_{i+1} + \sum_{j=1}^{i} \sum_{r=i+1}^{k} n_{rj} - m_{i+1} - \sum_{j=1}^{i-1} \sum_{r=i}^{k} n_{rj} + m_i \\ &= \sum_{j=1}^{i-1} \sum_{r=i+1}^{k} n_{rj} + \sum_{r=i+1}^{k} n_{ri} - \sum_{j=1}^{i-1} n_{ij} - \sum_{j=1}^{i-1} \sum_{r=i+1}^{k} n_{rj} + \sum_{j=1}^{i-1} n_{ij} \\ &= \sum_{r=i+1}^{k} n_{ri} = R_i, \quad (2 \leq i \leq k-1). \end{aligned}$$

■

Teorema 49. Para todo i, $2 \leq i \leq k-1$, os E.M.V. para M_i verificam

$$\hat{M}_i \cong m_i + \frac{Z_i}{1 - A_i},$$

onde A_i, $(2 \leq i \leq k-1)$ é definido pela relação (4.36.).

Prova. Como $\sum\limits_{j=1}^{i-1} \hat{N}_{ij} = \hat{M}_i$ e $\sum\limits_{j=1}^{i-1} n_{ij} = m_i$, segue da relação (4.35.) e dos Teoremas 46 e 47 que

$$(\hat{M}_i - m_i) - (\hat{M}_{i+1} - \hat{\phi}_i s_i) \cong \frac{(1-\hat{\phi}_i)}{A_i}(\hat{M}_i - m_i - Z_i), \ (2 \leq i \leq k-1).$$

Logo, da relação (4.44.) segue que

$$\hat{M}_i - m_i \cong \frac{1}{A_i}(\hat{M}_i - m_i - Z_i)$$

$$\Longrightarrow (\hat{M}_i - m_i)(1 - 1/A_i) \cong -\frac{Z_i}{A_i}$$

$$\Longrightarrow \hat{M}_i - m_i \cong Z_i/(1 - A_i)$$

$$\Longrightarrow \hat{M}_i \cong m_i + \frac{Z_i}{1 - A_i}, \ (2 \leq i \leq k-1).$$

■

O seguinte teorema nos dá um estimador de M_i em função de m_i, s_i, Z_i e R_i.

Teorema 50. Para todo i, $2 \leq i \leq k-1$, o E.M.V. para M_i verifica a equação

$$\hat{M}_i = m_i + \frac{s_i Z_i}{R_i}.$$

Prova. Das relações (4.36.), (4.44.) e (4.45.) temos

$$1 - A_i = \frac{1}{\hat{M}_i - m_i + s_i}[m_{i+1} + (\hat{M}_{i+1} - m_{i+1})(1 - A_{i+1})].$$

Desta última relação e do Teorema 49 segue que

$$\frac{Z_i}{\hat{M}_i - m_i} = \frac{m_{i+1} + Z_{i+1}}{\hat{M}_i - m_i + s_i}$$

$$\Longrightarrow \hat{M}_i - m_i = \frac{Z_i(\hat{M}_i - m_i) + s_i Z_i}{m_{i+1} + Z_{i+1}}$$

$$\Longrightarrow (\hat{M}_i - m_i)\left(1 - \frac{Z_i}{m_{i+1} + Z_{i+1}}\right) = \frac{s_i Z_i}{m_{i+1} + Z_{i+1}}$$

$$\Longrightarrow (\hat{M}_i - m_i)\left(\frac{m_{i+1} + Z_{i+1} - Z_i}{m_{i+1} + Z_{i+1}}\right) = \frac{s_i Z_i}{m_{i+1} + Z_{i+1}}$$

$$\Longrightarrow \hat{M}_i - m_i = \frac{s_i Z_i}{m_{i+1} + Z_{i+1} - Z_i}.$$

Então, pelo Teorema 48

$$\hat{M}_i - m_i = \frac{s_i Z_i}{R_i}$$

$$\Longrightarrow \hat{M}_i = m_i + \frac{s_i Z_i}{R_i}, \ (2 \leq i \leq k-1).$$

∎

Logo $\hat{M}_i$, o estimador do número de animais marcados na população na época imediatamente anterior à da momento da i-ésima tentativa, envolve s_i, m_i e o quociente Z_i/R_i.

Finalmente apresentamos o principal resultado desta seção

Teorema 51. Os E.M.V. para η_i N_i, p_i, ϕ_i e B_i são, respectivamente,

$$\hat{\eta}_i = \frac{s_i}{n_i}, \quad (1 \leq i \leq k) \tag{4.50.}$$

$$\hat{N}_i = n_i[1 + s_i Z_i/(R_i m_i)], \ (2 \leq i \leq k-1) \tag{4.51.}$$

$$\hat{p}_i = 1/[1 + s_i Z_i/(R_i m_i)], \ (2 \leq i \leq k-1) \tag{4.52.}$$

$$\hat{\phi}_i = (m_{i+1} + s_{i+1} Z_{i+1}/R_{i+1})/(s_i + s_i Z_i/R_i), \ (2 \leq i \leq k-2) \tag{4.53.}$$

e

$$\hat{B}_i \cong n_{i+1}\left(1 + \frac{s_{i+1}Z_{i+1}}{R_{i+1}m_{i+1}}\right) - [(m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1})/(1 + Z_i/R_i)][n_i Z_i/(R_i m_i) + 1] \tag{4.54.}$$

$(2 \leq i \leq k-2)$.

Prova . A relação (4.50.) segue diretamente da relação (4.40.). Usando a relação (4.45.) e o resultado do Teorema 50 temos

$$\hat{N}_i = \frac{n_i}{m_i}(m_i + s_i Z_i/R_i) = n_i[1 + s_i Z_i/(m_i R_i)], \quad (2 \leq i \leq k-1),$$

o que prova a relação (4.51.). Notemos que esta relação é exatamente a relação (4.5.) obtida na seção 4.2.

Pelas relações (4.39.) e (4.51.) segue que

$$\hat{p}_i = n_i/[n_i(1 + s_i Z_i/(m_i R_i))]$$

$$= 1/[1 + s_i Z_i/(m_i R_i)] = m_i R_i/(m_i R_i + s_i Z_i), \ (2 \leq i \leq k-1)$$

o que prova a relação (4.52.).

Pelo Teorema 50 e a relação (4.44.) temos

$$\hat{\phi}_i = (m_{i+1} + s_{i+1} Z_{i+1}/R_{i+1})/(s_i Z_i/R_i + m_i - m_i + s_i)$$

$$= (m_{i+1} + s_{i+1} Z_{i+1}/R_{i+1})/(s_i + s_i Z_i/R_i), \quad (2 \leq i \leq k-2).$$

o que prova a relação (4.53.). Observemos que esta relação é exatamente a relação (4.6.) obtida na seção 4.2.

Como $\hat{N}_{i0} = \hat{N}_i - \hat{M}_i$ e $n_{i0} = n_i - m_i$ $(1 \leq i \leq k)$, pela relação (4.41.) temos

$$\hat{B}_i \cong \hat{N}_{i+1} - \hat{M}_{i+1} - \hat{\phi}_i(\hat{N}_i - \hat{M}_i - n_i + m_i), \quad (1 \leq i \leq k-1).$$

Logo, pela relação (4.44.) temos

$$\hat{B}_i \cong \hat{N}_{i+1} - \hat{\phi}_i(\hat{M}_i - m_i + s_i) - \hat{\phi}_i(\hat{N}_i - \hat{M}_i - n_i + m_i)$$

$$= \hat{N}_{i+1} - \hat{\phi}_i(\hat{N}_i - n_i + s_i), \quad (1 \leq i \leq k-1).$$

Finalmente, usando esta última relação e as relações (4.51.) e (4.53.) temos

$$\hat{B}_i \cong n_{i+1}[1+s_{i+1}Z_{i+1}/(R_{i+1}m_{i+1})] - \frac{(m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1})}{s_i + s_i Z_i/R_i}\{n_i[1+s_i Z_i/(m_i R_i)] - n_i + s_i\}$$

$$= n_{i+1}[1 + s_{i+1}Z_{i+1}/(R_{i+1}m_{i+1})] - \frac{(m_{i+1} + s_{i+1}Z_{i+1}/R_{i+1})}{1 + Z_i/R_i}[n_i Z_i/(R_i m_i) + 1]$$

para $2 \leq i \leq k-2$, o que prova o teorema. Observemos que a última relação é exatamente a relação (4.7.) obtida na seção 4.2.

■.

APÊNDICE 1

LIMITE SUPERIOR PARA A VARIÂNCIA

Se X é uma variável aleatória tal que $P(a \leq X \leq b) = 1$, com $-\infty < a < b < \infty$, então

$$V(X) \leq \frac{(b-a)^2}{4}$$

Prova. Se $a = 0$ e $b = 1$, então $0 \leq X^2 \leq X \leq 1$. Logo,

$$0 \leq E(X^2) \leq E(X) \leq 1$$

o que implica

$$V(X) = E(X^2) - E^2(X) \leq E(X) - E^2(X) = E(X)[1 - E(X)] \leq \frac{1}{4}.$$

Se $a \neq 0$ ou $b \neq 1$, como a variável aleatória $Y = \frac{X-a}{b-a}$ assume valores em $[0,1]$ com probabilidade 1, então, pelo resultado anterior,

$$V(Y) = \frac{V(X)}{(b-a)^2} \leq \frac{1}{4}$$

ou

$$V(X) \leq \frac{(b-a)^2}{4}$$

■

APÊNDICE 2

SOMA DE DUAS VARIÁVEIS BINOMIAIS NEGATIVAS INDEPENDENTES

Se todo a, b e k são números inteiros positivos, então

$$\sum_{j=0}^{k} \binom{a+j-1}{j} \binom{b+k-j-1}{k-j} = \binom{a+b+k-1}{k}$$

Prova.

Se X e Y são variáveis aleatórias independentes com distribuição de probabilidades Binomial Negativa com parâmetros $(a,\ p)$ e $(b,\ p)$, respectivamente, então $(X+Y)$ tem distribuição Binomial Negativa com parâmetros $(a+b,\ p)$.

Com efeito, a função característica de $X+Y$ é, para todo t real,

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = \left(\frac{p}{1-qe^{it}}\right)^a \left(\frac{p}{1-qe^{it}}\right)^b = \left(\frac{p}{1-qe^{it}}\right)^{a+b},$$

onde $q = 1-p$.

Então, para todo número inteiro positivo k,

$$P(X=j/X+Y=k) = \frac{P(X=j, Y=k-j)}{P(X+Y=k)}$$

$$= \frac{P(X=j)P(Y=k-j)}{P(X+Y=k)} = \frac{\binom{a+j-1}{j}\binom{b+k-j-1}{k-j}}{\binom{a+b+k-1}{k}}, \quad 0 \leq j \leq k$$

Logo,

$$\sum_{j=0}^{k} P(X = j | X + Y = k) = 1 \Rightarrow$$

$$\Rightarrow \quad \sum_{j=0}^{k} \binom{a+j-1}{j} \binom{b+k-j-1}{k-j} = \binom{a+b+k-1}{k}.$$

■

APÊNDICE 3

APROXIMAÇÃO À DISTRIBUIÇÃO χ^2.

Sejam n_0, N números inteiros positivos. Para $p = \frac{n_0}{N}$ pequeno e $n_1 \sim$ Bin.Neg.$(m,\, p)$, a variável aleatória $\frac{2n_0n_1}{N}$ tem distribuição de probabilidade aproximadamente χ^2 com $2m$ graus de liberdade.

Prova . A função característica da variável $(2n_1p)$ é

$$\begin{aligned}
\varphi_{2n_1p}(t) = E\Big(e^{2n_1ipt}\Big) &= \sum_{j=m}^{\infty}\binom{j-1}{m-1}e^{2jipt}p^m(1-p)^{j-m} \\
&= p^m e^{2pmit}\sum_{j=m}^{\infty}\binom{j-1}{m-1}[(1-p)e^{2ipt}]^{(j-m)} \\
&= p^m e^{2pmit}\sum_{k=0}^{\infty}\binom{m+k-1}{k}[(1-p)e^{2ipt}]^{k} \\
&= p^m e^{2pmit}\sum_{k=0}^{\infty}\binom{-m}{k}[-(1-p)e^{2ipt}]^{k} \\
&= p^m e^{2pmit}[1-(1-p)e^{2ipt}]^{-m} \\
&= e^{2pmit}\left[\frac{1-(1-p)e^{2ipt}}{p}\right]^{-m}, \qquad \forall\, t \in R
\end{aligned}$$

.

Fazendo $p \to 0$ temos, pela regra de L'Hospital,

$$\lim_{p\to 0} \varphi_{2n_1p}(t) = \left[\lim_{p\to 0} \frac{1-(1-p)e^{2ipt}}{p}\right]^{-m}$$

$$= \left[\lim_{p\to 0} \frac{e^{2pti} - (1-p)e^{2pti}2it}{1}\right]^{-m}$$

$$= (1-2it)^{-m}, \qquad \forall\, t \in R.$$

que é a função característica de uma variável aleatória χ^2 com $2m$ graus de liberdade . ■

APÊNDICE 4

Se $x > \frac{3}{2}$**, então** $\frac{1}{x} + \frac{2}{x^2} + \frac{6}{x^3} > \frac{1}{[x]}$.

Prova : Como $0 < x - 1 < [x]$, então

$$[x](x^2 + 2x + 6) > (x - 1)(x^2 + 2x + 6) = x^3 + 2x^2 + 6x - x^2 - 2x - 6 =$$

$$= x^3 + x^2 + 4x - 6 = x^3 + x^2 + 2(2x - 3) > x^3,$$

ou seja,

$$[x](x^2 + 2x + 6) > x^3.$$

Logo, dividindo-se por $x^3[x]$ ambos os membros da última desigualdade, temos:

$$\frac{1}{x} + \frac{2}{x^2} + \frac{6}{x^3} > \frac{1}{[x]}.$$

∎

APÊNDICE 5

DESENVOLVIMENTO DE UMA SÉRIE FATORIAL INVERSA.

Para todo inteiro $m > 0$,

$$(m+1)^{-2} = \sum_{k=0}^{\infty} \frac{k!}{(m+1)(m+2)\cdots(m+k+2)}.$$

Prova: Como

$$\frac{1}{(m+1)} = \frac{1}{m+n} + \frac{n-1}{(m+1)(m+n)}, \quad n \geq 1, \tag{1}$$

então, aplicando este resultado para $n = 2$ em $\frac{1}{(m+1)^2} = \frac{1}{(m+1)(m+1)}$ temos,

$$\frac{1}{(m+1)^2} = \frac{1}{(m+1)}\left[\frac{1}{m+2} + \frac{1}{(m+1)(m+2)}\right] =$$

$$= \frac{1}{(m+1)(m+2)} + \frac{1}{(m+1)(m+2)}\frac{1}{(m+1)}.$$

Usando (1) para $n = 3$ no último termo da direita da relação anterior, temos:

$$\frac{1}{(m+1)^2} = \frac{1}{(m+1)(m+2)} + \frac{1}{(m+1)(m+2)}\left[\frac{1}{(m+3)} + \frac{2}{(m+1)(m+3)}\right]$$

$$= \frac{1}{(m+1)(m+2)} + \frac{1}{(m+1)(m+2)(m+3)} + \frac{2}{(m+1)(m+2)(m+3)} \frac{1}{(m+1)}.$$

Usando (1) para $n = 4$ no último termo da direita da última relação, resulta

$$\frac{1}{(m+1)^2} = \frac{1}{(m+1)(m+2)} + \frac{1}{(m+1)(m+2)(m+3)} +$$

$$+ \frac{2}{(m+1)(m+2)(m+3)} \left[\frac{1}{(m+4)} + \frac{3}{(m+1)(m+4)} \right]$$

$$= \frac{0!}{(m+1)(m+2)} + \frac{1!}{(m+1)(m+2)(m+3)} + \frac{2!}{(m+1)(m+2)(m+3)(m+4)} +$$

$$+ \frac{3!}{(m+1)(m+2)(m+3)(m+4)} \frac{1}{(m+1)}.$$

Análogamente, aplicando sucessivamente (1) para $n = 5, 6, \cdots$ temos

$$\frac{1}{(m+1)^2} = \sum_{k=0}^{\infty} \frac{k!}{(m+1)(m+2)\cdots(m+k+2)}$$

■

Observemos que, como

$$\frac{1}{(m+1)^2} = \sum_{k=0}^{n} \frac{k!}{(m+1)(m+2)...(m+k+2)} + \sum_{k=n+1}^{\infty} \frac{k!}{(m+1)(m+2)...(m+k+2)},$$

chamando de $R_n(m)$ o segundo termo da direita da relação anterior, temos:

$$R_n(m) = \sum_{k=n+1}^{\infty} \frac{k!}{(m+1)(m+2)...(m+k+2)} =$$

$$= \sum_{j=1}^{\infty} \frac{(n+j)!}{(m+1)(m+2)...(m+n+j+2)} <$$

$$< \sum_{j=1}^{\infty} \frac{(n+j)!}{(n+j+2)!} = \sum_{j=1}^{\infty} \frac{1}{(n+j+1)(n+j+2)} < \sum_{j=1}^{\infty} \frac{1}{(n+j+1)^2}.$$

Para $n = 2$,

$$R_2(m) < \sum_{j=1}^{\infty} \frac{1}{(j+3)^2} = \frac{\pi^2}{6} - 1 - \frac{1}{2^2} - \frac{1}{3^2} = \frac{6\pi^2 - 49}{36} \cong 0,283822955.$$

BIBLIOGRAFIA

Adams, L. (1951) . Confidence limits for the Petersen or Lincoln Index used in animal population studies.*The Journal of Wildlife Management.* vol. 15, 1, 13–19.

Bailey,G. N. A.(1968) . Trap-shyness in a woodland population of bank voles (Clethrionomys Glareolus). *J. Zool.* 156 (10). 517–521.

Bailey, Norman T. J. (1951) . On estimating the size of mobile populations from recapture data. *Biometrika* 38, 293–306.

Begon, M. (1959) .*Investigating animal abundance: capture-recapture for biologists.* London, Willmer.

Blower, J. G. ; Cook, L. M. ; Bishop, J. A. (1981) *Estimating the size of animal populations.* London, G. Allen and Unwin.

Chapman, Douglas G. (1951) . Some properties of the hypergeometric distribution with applications to zoological sample censuses. *Univ. Calif. Public. Stat..* 1, 131–60.

Chapman, Douglas G. (1952). Inverse, multiple and sequential sample censuses. *Biometrics* 8, 286–306.

Clopper, C ; Pearson, E. S. (1934) . The use of confidence or fiducial limits illustrates in the case of the Binomial. *Biometrika* 26, 404–413.

Cormack, R. M. (1969) . The statistics of capture-recapture methods. *Oceanogr. Mar. Biol. Ann. Rev.* 6, 455–506.

Cormack, R. M. (1973) . Commonsense estimates from capture-recapture studies. In the mathematical theory of the dynamics of biological populations.

Cramér, H. (1974) . *Mathematical methods of statistics.* (13a. ed.). Princeton University Press.

Darroch, J. N. (1958) . The multiple-recapture census. I. Estimation of a closed Population. *Biometrika* 45, 343–59.

Darroch, J. N. (1961) . The two-sample capture-recapture census when tagging and sampling are stratified. *Biometrika* 47, 241–60.

Dowdeswell, W. H; Fisher, R. A.; Ford, E. B. (1940) . The quantitative study of populations in the Lepidoptera. I Polyommatus icarus(Rott). *Ann. Eugen. Lond.* 10, 123–36.

Dowdeswell, W. H.; Fisher, R. A.; Ford, E. B. (1949) . The quantitative study of populations in the Lepidoptera. 2. Maniola furtina. *L. Heredity* 3, 67–84.

Feller, W. (1980) .*Introducción a la Teoría de Probabilidades.* vol. I, México, Ed. Limusa.

Fisher, R. A. ; Ford, E. B. (1947) . The spread of a gene in natural conditions in a colony of the moth Panaxia dominula. *L. Heredity* 1, 143–74.

Flyger, V. R.(1959) . Movements and home range of the gray squirrel Sciurus carolinensis, in two Maryland woodlots. *Ecology* 41 (3), 365–369.

Geis, A. D. (1952) . Trap response of the cottontail rabbit and its effect on censusing. *J. Wildl. Mange.*.19(4), 466–472.

Haldane, J. M. (1953) . Capture-recapture analysis. *Biometrika* 40, 265–78.

Jackson, C. H. N. (1953) . On the true density of tsetse flies. *J. Anim. Ecol.* 2, 204–9.

Johnson, N. L. ; Kotz, S. (1969) . *Discrete Distributions.* Boston, Houghton Mifflin.

Jolly, G. M. (1965) . Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* 52, 225–47.

Le Cren, E. D. (1965) . A note on the history of mark-recapture population estimates. *J. Anim. Ecol.* 34, 453–4.

Lehmann, E. L. (1983) . *Theory of Point Estimation.* N. York, Wiley.

Leite, J. G.; Oishi, J.; Pereira, C. A. B. (1988) . A note on the exact maximum likelihood estimation of the size of a finite and closed population. *Biometrika.* 75, 178–80.

Leslie, P. H.; Chitty, D. (1951) . The estimation of population parameters from data obtained by means of the capture-recapture method. I. The maximum likelihood equations for estimating the death rate. *Biometrika.* 38, 269–92.

Lincoln, F. C. (1930) . Calculating waterfowl abundance on the basis of banding returns.*U. S. Dept. Agric. Circ.* 118, 1–4.

Manly, B. F. J.; Parr, M. J. (1968) . A new method of estimating population size, survivorship, and birth rate from capture-recapture data. *Trans. Soc. Brit. Ent.* 18, 81–9.

McDonald, J. F., Selby, M. A., Wong, C. S., Favro, L. D. and Kuo, P. K. (1983). Test, point estimations, and confidence sets for a capture-recapture model.

Journal of the American Statistical Association 78, 913–919.

Otis, D. L.; Burnham, K. P.; White, G. C.; Anderson, D. R. (1978) Statistical Inference from capture data on closed animal populations. *Wildlife Monographs* 62, 1-135.

Overton, W. S. ; Davies, D. E. (1969) . Estimating the number of animals in wildlife populations. Wildlife Management techniques. 3rd. ed. The Wildlife Society, Wash., D. C. Ed. R. H. Giles, Jr. p.403–455.

Parr, M. J. (1965) . A population study of a colony of imaginal Ischeneura elegans(van der Linden)(Odonata. Coenagrudae) at Dale. *Pembrokeshire Fld. Studies* 2, 237–82.

Petersen, G. G. J. (1896) . The yearly immigration of young plaice into the Limfjord from the German Sea. *Rept. Danish Biol. Sta.* 6, 1–48.

Pollock, K. H. (1976) . Building models of capture-recapture experiments. *The Statistician* 25, 4, 253-259.

Pucek, Z. (1969) . Trap response and estimation of numbers of shrews in removal catches. *Acta Theriol.* 14, 403–426.

Seber, G. A. F. (1965) . A note on the multiple-recapture census. *Biometrika* 52, 249–59.

Seber, G. A. F. (1973) . The estimation of animal abundance and related parameters. London, Griffin.

Tanaka, R. (1956) . On differential response to live traps of marked and unmarked small mammals. *Ann. Zool. Jap..* 29 (1), 44–51.

Tanaka, R. (1963) . On the problem of trap-response types of small mammal populations. *Res. POp. Ecol.* 2, 139–146.