ESTIMATIVAS EXATAS DO
TAMANHO DE UMA POPULAÇÃO
FINITA E FECHADA

José Galvão Leite

TESE APRESENTADA

AO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DA

UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE DOUTOR

EM

ESTATÍSTICA

ORIENTADOR. PROF. DR. CARLOS ALBERTO DE BRAGANÇA PEREIRA

- São Paulo, Fevereiro de 1986 -

AGRADECIMENTOS

Quero apresentar os meus agradecimentos a todos aqueles que, de algum modo, contribuíram para a realização deste trabalho.

Ao Professor *Antonio Galves*, pelo apoio que me prestou durante o período em que participei, no IME USP, das atividades do grupo de pesquisa Sistemas de Partículas.

Ao Professor *Josemar Rodrigues*, com quem mantive as primeiras discussões sobre este trabalho, pelas sugestões apresentadas.

Aos Professores *Seiji Hariki* e *Luiz Augusto Fernandes de Oliveira* pela atenção e paciência que sempre me dedicaram.

Ao Professor *Flávio Wagner Rodrigues* cujo apoio, atenção e amizade me foram de grande significado durante a realização deste trabalho.

Ao Professor *Jorge Oishi*, que desenvolveu toda a programação computacional relativa a este trabalho, pelas importantíssimas sugestões apresentadas.

A Srta. *Antônia Soares*, que datilografou este trabalho, pela atenção que me dispensou.

Sou particularmente grato ao Professor *Carlos Alberto de Bragança Pereira*, que me acolheu sob sua orientação e de cujo convívio recebi amizade, dedicação e incentivo e que me passou a sabedoria de não complicar o simples.

ooo

RESUMO

Este trabalho trata da estimação do tamanho de uma população finita utilizando amostras obtidas pelo método da captura-recaptura. O tamanho desconhecido da população permanece fixo durante o processo. Um modelo de distribuição aleatória de bolas em celas é usado para obter as distribuições de probabilidades dos dados amostrais. Formas explícitas para as estimativas de máxima verossimilhança e as estimativas de Bayes, com relação à perda quadrática, são deduzidas. Propriedades e o comportamento assintótico das estimativas e dos estimadores de Bayes são também analisados.

ooo

ABSTRACT

This work deals with the estimation of the size of a finite population using samples obtained through the capture-recapture method. The unknown population size remains fixed throughout the process. A random distribution model of balls in cells is used to obtain the probabilistic distributions of the sampling data. Explicit expressions for the maximum likelihood estimates and the Bayes estimates, under quadratic loss, are derived. Properties and the asymptotic behavior of the Bayes estimates and estimators are also analysed.

oOo

Í N D I C E

	pág.
INTRODUÇÃO	1
CAPÍTULO I - UM MODELO DE URNA PARA O PROBLEMA DA CAPTURA-RECAPTURA	4
I.1 - Introdução	4
I.2 - Processo de seleção um a um	5
I.3 - Processo de seleção generalizado..	13
I.4 - Comentários e conclusões	27
REFERÊNCIAS PARA O CAPÍTULO I	29
CAPÍTULO II - ESTIMATIVAS DE MÁXIMA VEROSSIMILHANÇA ...	30
II.1 - Introdução	30
II.2 - Seleções um a um	31
II.3 - Seleções com igual número de capturas	39
II.4 - Seleções com diferentes números de capturas	48
II.5 - Comentários e conclusões	56
REFERÊNCIAS PARA O CAPÍTULO II	58

CAPÍTULO III - ESTIMATIVAS DE BAYES	59
III.1 - Introdução	59
III.2 - Obtenção das estimativas	61
III.3 - Propriedades das estimativas de Bayes	67
REFERÊNCIAS PARA O CAPÍTULO III	72
 CAPÍTULO IV - RESULTADOS ASSINTÓTICOS	73
IV.1 - Introdução	73
IV.2 - Propriedades assintóticas das estimativas de Bayes	74
IV.3 - Modelo estatístico	79
IV.4 - Propriedades assintóticas do estimador de Bayes	85
REFERÊNCIAS PARA O CAPÍTULO IV	91
 REFERÊNCIAS COMPLEMENTARES	92

Introdução

O objetivo deste trabalho é desenvolver estimativas do tamanho de uma população finita e fechada. A suposição básica é de que durante o processo de estimação a população não se altera em tamanho e forma, isto é, a população é estável. As estimativas serão baseadas em amostras obtidas pelo processo de captura-recaptura. Este processo consiste em selecionarmos, em um certo número de etapas, amostras aleatórias sem reposição (uma amostra em cada etapa) da população. Todos os elementos selecionados na primeira etapa serão marcados e devolvidos à população; após cada etapa, a partir da segunda, os elementos selecionados não marcados anteriormente serão marcados e devolvidos à população, juntamente com os elementos selecionados que já haviam sido marcados.

As estimativas do tamanho da população serão baseadas no número de elementos não marcados selecionados nas diversas etapas.

O problema da estimação do tamanho de uma população pelo processo da captura-recaptura foi considerado por inúmeros autores, notadamente Craig (1953), Goodman (1953), Chapman (1954), Lewontin & Prout (1956), Darroch (1958), Jolly (1965), Seber (1965), Darling & Robbins (1967), Samuel (1968), Freeman (1972), Zacks (1984) e Pollock, Hines & Nichols (1985).

Pelo nosso conhecimento, os únicos autores que tratam do problema sob o ponto de vista Bayesiano são Freeman

(1972) e Zacks (1984). Ambos os autores consideram o caso em que as amostras, em cada etapa, são de tamanho um (seleções um a um).

O presente trabalho foi motivado pelo relatório técnico "Bayes sequential estimation of the size of a finite population" desenvolvido pelo Prof. Zacks quando da sua visita, em 1984, ao IME-USP (Zacks (1984)).

O 1º capítulo trata das leis probabilísticas dos dados obtidos pelo processo da captura-recaptura. Nossa contribuição aqui foi a construção de um modelo de distribuição de bolas em urnas para determinar as leis amostrais, que originalmente (veja Zacks (1984)) são obtidas através de um processo indireto de equações a diferenças. Com o objetivo de facilitar a compreensão do texto, apresentamos em primeiro lugar o processo de seleção um a um, para posteriormente tratarmos do modelo generalizado.

No 2º capítulo, referente as estimativas de máxima verossimilhança, também por motivos didáticos decidimos apresentar, por ordem de dificuldade, o caso de seleções um a um, o caso de seleções com o mesmo número de elementos e o caso geral. Nossa contribuição neste capítulo são os teoremas 14, 18 e 22 que além de apresentar formas explícitas para as estimativas da máxima verossimilhança, mostram também a existência de, no máximo, duas estimativas de máxima verossimilhança.

No caso de seleções um a um provamos que a estimativa de máxima verossimilhança é única.

No 3º capítulo apresentamos as estimativas de Bayes apenas para o caso geral, pois o caso de seleções um a um já havia sido discutido em Freeman (1972) e Zacks (1984). Nossa contribuição neste capítulo foi a de provar, a menos de duas, as conjecturas de Zacks (1984). Para as duas restantes, através de contra-exemplos, mostramos que são falsas.

Finalmente, no 4º capítulo, este de natureza teórica, como não temos conhecimento de trabalho semelhante na literatura, apresentamos em detalhes a teoria assintótica das estimativas e dos estimadores de Bayes. Para os resultados sobre o comportamento assintótico dos estimadores de Bayes utilizamos os resultados de Doob (1953) relativos aos martin-gais e supermartingais.

oOo

CAPÍTULO I

UM MODELO DE URNA PARA O PROBLEMA DA CAPTURA-RECAPTURA

I. 1 - Introdução

O problema da captura-recaptura surge quando se deseja estimar o tamanho, N , de uma população finita e estável. Por estável entende-se uma população que, durante o estudo, não apresenta mortes, nascimentos e nem migrações. Considera-se k seleções com reposição de elementos dessa população. Em cada uma destas seleções pode-se considerar a captura de um ou mais elementos de uma só vez. A variável de interesse, denotada por T_k , é o número de elementos distintos nas k seleções. Para que um elemento capturado seja identificável, este recebe uma marca na primeira vez em que for capturado. Assim, poderá ser reconhecido no caso de ser recapturado.

O modelo de seleção mais simples é descrito a seguir. Um primeiro elemento é capturado, marcado e devolvido à população. Uma nova captura é efetuada; se o elemento não for marcado, ele receberá a marca antes de ser devolvido à população; se for marcado, ele será simplesmente devolvido à população. Este processo é repetido até a k -ésima seleção. Assim, T_k representa o número de elementos não marcados (e que receberam a marca) capturados nas k seleções.

O problema da captura-recaptura consiste na determinação da distribuição de probabilidades de T_k .

O processo mais geral de seleção pode ser descrito como o anterior, onde na j -ésima ($j=1,2,\dots,k$) seleção, em vez de apenas um elemento, são capturados m_j ($m_j \geq 1$) elementos de uma só vez. Como antes, os elementos marcados retornam imediatamente à população e os restantes são marcados antes de serem devolvidos à população. Da mesma forma, T_k é o número total de elementos distintos capturados nas k seleções. Resumindo, $m_1+m_2+\dots+m_k$ é o número total de capturas e T_k é este número subtraído do número total de recapturas.

Para determinar a distribuição de probabilidades da variável T_k , utilizamos um modelo de distribuição aleatória de bolas em urnas, que corresponde ao processo de seleção descrito.

Na seção I.2 apresentamos a solução para o caso de seleções um a um. Na seção I.3 apresentamos a solução para o caso geral.

Alguns comentários complementares são feitos na seção I.4.

I.2 - Processo de seleção um a um

Nesta seção consideramos o caso onde as k ($k \in \mathbb{N}^* = \{1,2,\dots\}$) seleções são unitárias. Isto é, apenas um elemen

to é capturado em cada uma das seleções.

Considere um conjunto de N ($N \in \mathbb{N}^*$) urnas e k bolas numeradas de 1 a k . À cada elemento da população associe biunivocamente uma urna e à j -ésima ($j=1,2,\dots,k$) seleção faça corresponder a j -ésima bola. Considere o experimento que consiste em se distribuir aleatoriamente as k bolas nas N urnas. Capturar o elemento i da população na j -ésima seleção é equivalente a se observar a j -ésima bola na i -ésima urna, quando o experimento da distribuição aleatória das bolas nas urnas é realizado.

Como na sequência deste trabalho o enfoque estatístico utilizado é o Bayesiano, poderemos tecnicamente considerar N como uma variável aleatória. Lembremos que N é um elemento fixado e desconhecido de $\mathbb{N}^*$. Contudo, para descrever o grau de desconhecimento sobre o valor de N , o estatístico Bayesiano se utiliza de uma distribuição de probabilidades sobre $\mathbb{N}^*$. Assim, é conveniente o uso da notação de probabilidade condicional, $P\{. | N=n\}$ ($n \in \mathbb{N}^*$), toda vez que se deseje descrever distribuições de probabilidades de variáveis aleatórias, onde N é considerado como fixado e desconhecido. Ou seja, quando N for o parâmetro desconhecido do modelo.

Consideremos a sequência de variáveis aleatórias $\{U_j\}_{j \geq 1}$ definidas como:

$$U_j = \begin{cases} 1 & \text{se o } j\text{-ésimo elemento selecionado é não} \\ & \text{marcado (1ª captura),} \\ 0 & \text{se o } j\text{-ésimo elemento selecionado é marca} \\ & \text{do (recaptura),} \end{cases}$$

$j = 1, 2, \dots$

Fixado $k \in \mathbb{N}^*$, consideremos o vetor aleatório $D_k = (U_1, U_2, \dots, U_k)$ e a variável aleatória $T_k = \mathbb{1}_{D_k} = \sum_{j=1}^k U_j$, onde $\mathbb{1}$ é o vetor de unidades de ordem k . Assim, para todo $n \in \mathbb{N}^*$,

$$P\{U_1 = 1 | N=n\} = P\{T_1 = 1 | N=n\} = 1.$$

Para a determinação das distribuições de probabilidades (condicionais a $\{N=n\}$) de D_k e T_k , no caso $k \geq 2$, utilizaremos a equivalência entre o processo de seleção e a distribuição aleatória de k bolas em n urnas. Denotemos por $t_1, t_2, \dots, t_k$ as somas $t_j = \sum_{i=1}^j u_i$, onde $t_1 = u_1 = 1$ e $u_j \in \{0, 1\}$, $j=2, 3, \dots, k$.

Teorema 1 - Para todo $k \geq 2$ e para todo $n \in \mathbb{N}^*$,

$$\begin{aligned} P\{D_k = (1, u_2, \dots, u_k) | N=n\} &= \\ &= \frac{n!}{(n-t_k)! n^k} \prod_{j=2}^k (t_{j-1})^{1-u_j} I_k(t_k) \end{aligned} \quad (I.1)$$

onde $I_k(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^*: 1 \leq x \leq \min\{k, n\}\}$.

Prova - Para todo $u_j \in \{0, 1\}$, $j=2, 3, \dots, k$, satisfazendo a condição

$$1 \leq t_k \leq \min\{k, n\},$$

a ocorrência do evento $\{D_k = (1, u_2, \dots, u_k)\}$ é equivalente, no caso da distribuição das k bolas em n urnas, à ocorrência do evento definido pelos itens i) e ii) abaixo.

- i) A bola de número 1 é colocada aleatoriamente em uma das n urnas;
- ii) para todo j , $2 \leq j \leq k$, a bola de número j é colocada aleatoriamente em uma das urnas vazias se $u_j = 1$ e, colocada aleatoriamente em uma das urnas já ocupadas se $u_j = 0$.

Note que, em relação aos itens i) e ii) acima, existem $\binom{n}{1}$ maneiras de se colocar a bola de número 1 em uma das n urnas e que, para todo j , $2 \leq j \leq k$, o número de possíveis alocações da bola de número j em uma das n urnas é dado por

$$\binom{n-t_{j-1}}{u_j} \binom{t_{j-1}}{1-u_j} = \begin{cases} n-t_{j-1} & \text{se } u_j = 1 \\ t_{j-1} & \text{se } u_j = 0 \end{cases} .$$

Como n^k é o número total de possíveis alocações das k bolas nas n urnas, temos que

$$P\{D_k = (1, u_2, \dots, u_k) | N=n\} = \frac{\binom{n}{1}}{n^k} \prod_{j=2}^k \binom{n-t_{j-1}}{u_j} \binom{t_{j-1}}{1-u_j} .$$

Fazendo $t_0=0$ podemos escrever

$$\begin{aligned}
 P\{D_k = (1, u_2, \dots, u_k) | N=n\} &= \\
 &= \frac{1}{n^k} \left[\prod_{j=1}^k \binom{n-t_{j-1}}{u_j} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{1-u_j} \right].
 \end{aligned}$$

Para concluir o teorema basta notar que

$$\prod_{j=1}^k \binom{n-t_{j-1}}{u_j} = \frac{n!}{(n-t_k)!},$$

$$\prod_{j=2}^k \binom{t_{j-1}}{1-u_j} = \prod_{j=2}^k (t_{j-1})^{1-u_j}$$

e que

$$P\{D_k = (1, u_2, \dots, u_k) | N=n\} = 0 \quad \text{se } t_k > \min\{k, n\}.$$

□

O resultado que apresentamos a seguir pode ser encontrado em Feller (1968) e será de grande importância na determinação da distribuição de probabilidades de T_k .

Denotemos por $P_v(k, n)$ a probabilidade de que exatamente v urnas ($v \in \mathbb{N} = \{0, 1, 2, \dots\}$) resultem vazias, quando k bolas ($k \geq 1$) são distribuídas aleatoriamente em n urnas.

Lema 2 - Para todo $k \geq 1$ e para todo $n \in \mathbb{N}^*$,

$$P_v(k, n) = \frac{\binom{n}{v}}{n^k} \sum_{i=0}^{n-v} (-1)^i \binom{n-v}{i} (n-v-i)^k I(v) \quad (I.2)$$

onde $I(\cdot)$ é a função indicadora do conjunto

$$\{x \in \mathbb{N} : n - \min\{k, n\} \leq x \leq n-1\}.$$

Prova - Considere o evento $A_i =$ "a i -ésima urna resulta vazia"

$i=1, 2, \dots, n$ e represente a soma sobre o conjunto

$J_i = \{(j_1, j_2, \dots, j_i) : 1 \leq j_1 < j_2 < \dots < j_i \leq n\}$ por Σ_{J_i} . A probabilidade de que pelo menos uma urna resulte vazia é dada por

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n (-1)^{i-1} \Sigma_{J_i} P(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_i}) = \\ &= \sum_{i=1}^n (-1)^{i-1} \Sigma_{J_i} \frac{(n-i)^k}{n^k} = \\ &= \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} \frac{(n-i)^k}{n^k} = \\ &= n^{-k} \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} (n-i)^k. \end{aligned}$$

Logo, a probabilidade de que todas as urnas resultem ocupadas é dada por

$$P_0(k, n) = 1 - P\left(\bigcup_{i=1}^n A_i\right) = n^{-k} \sum_{i=0}^n (-1)^i \binom{n}{i} (n-i)^k.$$

Consideremos agora o evento "exatamente v urnas fixadas resultam vazias", onde

$$n - \min\{k, n\} \leq v \leq n-1.$$

Para este evento as k bolas ocupam as restantes $n-v$ urnas, on-

de cada uma destas está ocupada. O número de casos possíveis satisfazendo este evento é

$$(n-v)^k P_0(k, n-v),$$

uma vez que $(n-v)^k$ é o número total de possíveis distribuições das k bolas nas $n-v$ urnas. Como v urnas podem ser escolhidas de $\binom{n}{v}$ maneiras, segue-se que

$$P_v(k, n) = n^{-k} \binom{n}{v} (n-v)^k P_0(k, n-v).$$

Para concluir o lema basta notar que,

$$P_0(k, n-v) = (n-v)^{-k} \sum_{i=0}^{n-v} (-1)^i \binom{n-v}{i} (n-v-i)^k$$

e

$$P_v(k, n) = 0 \quad \text{se } v \notin \{x \in \mathbb{N}: n - \min\{k, n\} \leq x \leq n-1\}.$$

A distribuição de probabilidades de $T_k = 1D'_k = \sum_{j=1}^k U_j$, principal resultado desta seção, será obtida através de uma aplicação direta deste lema. Apresentamo-la como um teorema pela importância que terá na sequência.

Teorema 3 - Para todo $k \geq 1$ e para todo $n \in \mathbb{N}^*$,

$$P\{T_k = t | N=n\} = \frac{n!}{(n-t)! n^k} \sum_{j=0}^t (-1)^{t-j} \frac{j^k}{j! (t-j)!} I_k(t) \quad (I.3)$$

onde $I_k(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^*: 1 \leq x \leq \min\{k, n\}\}$.

Prova - Ao evento $\{T_k=t\}$ corresponde o evento "exatamente t ($n-t$) urnas resultam ocupadas (vazias)" no modelo de urnas que estamos considerando. Logo, $P\{T_k=t|N=n\} = P_{n-t}(k, n)$ e, como $I(n-t) = I_k(t)$, pelo lema 2 segue-se que

$$\begin{aligned} P\{T_k = t|N=n\} &= \frac{\binom{n}{n-t}}{n^k} \sum_{i=0}^t (-1)^i \binom{t}{i} (t-i)^k I_k(t) = \\ &= \frac{\binom{n}{t}}{n^k} \sum_{j=0}^t (-1)^{t-j} \binom{t}{t-j} j^k I_k(t) = \\ &= \frac{n!}{(n-t)! n^k} \sum_{j=0}^t (-1)^{t-j} \frac{j^k}{j!(t-j)!} I_k(t). \end{aligned}$$

□

Considerando-se N como um parâmetro desconhecido e de interesse, D_k como os dados amostrais e T_k como uma estatística, os teoremas 1 e 3 nos permitem concluir o seguinte resultado importante.

Corolário 4 - Para inferências sobre N , para todo $k \geq 1$, T_k é uma estatística suficiente com relação a D_k . Para todo $k \geq 2$, a distribuição de probabilidades condicional de D_k dado T_k é dada por

$$\begin{aligned} P\{D_k = (1, u_2, \dots, u_k) | T_k=t\} &= \\ &= \left[\prod_{j=2}^k (t_{j-1})^{1-u_j} \right] \left[\sum_{j=0}^t (-1)^{t-j} \frac{j^k}{j!(t-j)!} \right]^{-1} I_{\{t\}}(t_k), \end{aligned}$$

onde $I_{\{t\}}(\cdot)$ é a função indicadora do conjunto $\{t\}$, com $1 \leq t \leq k$.

O fator $\frac{n!}{(n-t)!n^k}$, comum às distribuições de probabilidades de D_k e T_k , denomina-se "kernel da verossimilhança", pois é o único fator relevante para inferências Bayesianas ou estimativas de máxima verossimilhança para N .

I.3 - Processo de seleção generalizado

Nesta seção consideramos um processo de seleção mais geral que o da seção anterior. Consideramos o caso em que o processo de seleção é realizado em k ($k \in \mathbb{N}^*$) etapas, onde na j -ésima ($j=1,2,\dots,k$) etapa, m_j ($m_j \geq 1$ e $N \geq \max\{m_1, m_2, \dots, m_k\}$) elementos são capturados aleatoriamente de uma só vez. Como:

- 1º) o resultado obtido na j -ésima etapa pode ser considerado como uma amostra aleatória, não ordenada, $\{i_1, i_2, \dots, i_{m_j}\}$ de elementos da população, quando m_j elementos são capturados aleatoriamente um de cada vez e sem reposição e,
- 2º) em cada etapa, no caso de capturas uma de cada vez e sem reposição, considerar amostras aleatórias não ordenadas ou ordenadas de elementos da população é irrelevante na determinação da distribuição de probabilidades de T_k (número total de elementos distintos nas k etapas),

vamos considerar, no lugar do processo de seleção original, que na j -ésima etapa são capturados aleatoriamente m_j elementos, um de cada vez sem reposição e, para cada captura, o elemento capturado é anotado.

Considere um conjunto de N urnas e um conjunto de k distintas cores. Para cada j ($j=1,2,\dots,k$) considere m_j bolas coloridas com a j -ésima cor e numeradas de 1 a m_j . A cada elemento da população associe biunivocamente uma urna e, para cada etapa j , à i_j -ésima ($1 \leq i_j \leq m_j$) seleção faça corresponder a i_j -ésima bola de cor j . Considere o experimento que consiste em se distribuir aleatoriamente, em k etapas, as $m_1+m_2+\dots+m_k$ bolas nas N urnas do seguinte modo. Na j -ésima etapa, as m_j bolas de cor j são distribuídas aleatoriamente nas N urnas, de tal modo que, exatamente m_j urnas resultem ocupadas pelas bolas utilizadas nesta etapa. Isto é, após a j -ésima etapa cada urna conterá, no máximo, uma bola de cada uma das cores $1,2,\dots,j$. Na j -ésima etapa, capturar aleatoriamente os elementos $i_1, i_2, \dots, i_{m_j}$ da população nas $1^a, 2^a, \dots, m_j$ -ésimas seleções, respectivamente, é equivalente a se observar as $1^a, 2^a, \dots, m_j$ -ésimas bolas de cor j nas urnas $i_1, i_2, \dots, i_{m_j}$, respectivamente, quando o experimento da distribuição aleatória das bolas nas urnas é realizado. Lembremos que N é considerado como uma variável aleatória (seção I.2).

Consideremos a sequência de variáveis aleatórias $\{U_j\}_{j \geq 1}$ definidas como

U_j = número de elementos não marcados (capturados)
selecionados na j -ésima etapa,
 $j = 1, 2, \dots$

A variável U_1 assume o valor m_1 e, para $j=2, 3, \dots$, a variável U_j assume um valor no conjunto $\{0, 1, \dots, m_j\}$. Fixado $k \in \mathbb{N}^*$, considere o vetor aleatório $D_k = (U_1, U_2, \dots, U_k)$ e a variável aleatória $T_k = \mathbb{1} D'_k = \sum_{j=1}^k U_j$, onde $\mathbb{1}$ é o vetor de unidades de ordem k .

Assim, para todo $n \in \mathbb{N}^*$ com $n \geq m_1$,

$$P\{U_1=m_1 | N=n\} = P\{T_1=m_1 | N=n\} = 1.$$

Para determinar, no caso $k \geq 2$, as distribuições de probabilidades (condicionais a $\{N=n\}$ onde $n \geq \max\{m_1, m_2, \dots, m_k\}$) de D_k e T_k utilizaremos a equivalência entre o processo de seleção e a distribuição das $m_1+m_2+\dots+m_k$ bolas coloridas em n urnas acima descritos.

Denotemos por $t_1, t_2, \dots, t_k$ as somas $t_j = \sum_{i=1}^j u_i$, onde $t_1=u_1=m_1$ e $u_j \in \{0, 1, \dots, m_j\}$, $j=2, 3, \dots, k$.

Teorema 5 - Para todo $k \geq 2$ e para todo $n \geq \max\{m_1, m_2, \dots, m_k\}$,

$$\begin{aligned} P\{D_k = (m_1, u_2, \dots, u_k) | N=n\} = \\ = \frac{n!}{(n-t_k)! \prod_{j=1}^k \binom{n}{m_j}} \left[\prod_{j=1}^k \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j - u_j} \right] I_k(t_k) \end{aligned} \quad (I.4)$$

onde $I_k(\cdot)$ é a função indicadora do conjunto

$$\{x \in \mathbb{N}^*: \max\{m_1, m_2, \dots, m_k\} \leq x \leq \min\{\sum_{i=1}^k m_i, n\}\}.$$

Prova - Para todo $u_j \in \{0, 1, \dots, m_j\}$ satisfazendo a condição

$$\max\{m_1, m_2, \dots, m_j\} - m_1 \leq u_2 + u_3 + \dots + u_j \leq \min\{\sum_{i=1}^j m_i, n\} - m_1,$$

ou a condição

$$\max\{m_1, m_2, \dots, m_j\} \leq t_j \leq \min\{\sum_{i=1}^j m_i, n\},$$

para $j=2, 3, \dots, k$, a ocorrência do evento $\{D_k = (m_1, u_2, \dots, u_k)\}$ é equivalente, no caso da distribuição aleatória, em k etapas, das $m_1 + m_2 + \dots + m_k$ bolas coloridas em n urnas, à ocorrência do evento definido pelos itens i) e ii) abaixo.

- i) Na 1ª etapa, as bolas $1, 2, \dots, m_1$, coloridas com a cor 1, são distribuídas aleatoriamente nas n urnas, de tal modo que, venham a ocupar exatamente m_1 urnas;
- ii) na j -ésima ($2 \leq j \leq k$) etapa, as bolas $1, 2, \dots, m_j$, coloridas com a cor j , são distribuídas aleatoriamente nas n urnas, de tal modo que, u_j dessas bolas venham a ocupar exatamente u_j urnas dentre as $n - t_{j-1}$ urnas desocupadas e, as restantes $m_j - u_j$ bolas venham a ocupar exatamente $m_j - u_j$ urnas dentre as t_{j-1} urnas já ocupadas anteriormente.

No que segue, se $m \in \mathbb{N}^*$, usaremos a notação

$$(m)_i = \begin{cases} m(m-1) \dots (m-i+1) & \text{se } i=1, 2, \dots, m \\ 1 & \text{se } i=0 \\ 0 & \text{se } i=m+1, m+2, \dots \end{cases}$$

Note que, em relação aos itens i) e ii) acima, na 1ª etapa existem $(n)_{m_1} = \binom{n}{m_1} (m_1)!$ maneiras de se distribuir as m_1 bolas nas n urnas e que, na j -ésima ($2 \leq j \leq k$) etapa, como existem $\binom{m_j}{u_j}$ maneiras de se escolher u_j bolas dentre as m_j , o número de possíveis alocações das m_j bolas nas n urnas é dado por

$$\binom{m_j}{u_j} (n-t_{j-1})_{u_j} (t_{j-1})_{m_j-u_j} = (m_j)! \binom{n-t_{j-1}}{u_j} \binom{t_{j-1}}{m_j-u_j}.$$

Como $\prod_{j=1}^k (n)_{m_j} = \prod_{j=1}^k \binom{n}{m_j} (m_j)!$ é o número total de possíveis alocações das $m_1+m_2+\dots+m_k$ bolas nas n urnas, temos que

$$P\{D_k = (m_1, u_2, \dots, u_k) | N=n\} = \frac{\binom{n}{m_1}}{\prod_{j=1}^k \binom{n}{m_j}} \prod_{j=2}^k \binom{n-t_{j-1}}{u_j} \binom{t_{j-1}}{m_j-u_j}.$$

Fazendo $t_0=0$ podemos escrever

$$\begin{aligned} P\{D_k = (m_1, u_2, \dots, u_k) | N=n\} &= \\ &= \frac{1}{\prod_{j=1}^k \binom{n}{m_j}} \left[\prod_{j=1}^k \binom{n-t_{j-1}}{u_j} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j-u_j} \right]. \end{aligned}$$

Para concluir o teorema basta notar que a igualdade acima também se verifica se $t_j < \max\{m_1, m_2, \dots, m_j\}$ ou $t_j > \min\{\sum_{i=1}^j m_i, n\}$, para algum j , $j=2, 3, \dots, k$ e que

$$\begin{aligned} & \left[\prod_{j=1}^k \binom{n-t_{j-1}}{u_j} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j - u_j} \right] = \\ & = \begin{cases} \frac{n!}{(n-t_k)!} \left[\prod_{j=1}^k \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j - u_j} \right] & \text{se } I_k(t_k) = 1 \\ 0 & \text{se } I_k(t_k) = 0 \end{cases} \end{aligned}$$

□

Para todo $k \geq 2$, suponha que $m_1 = m_2 = \dots = m_k = m$ ($m \geq 1$). Neste caso o teorema 5 nos permite concluir o seguinte resultado.

Corolário 6 - Para todo $k \geq 2$ se $m_1 = m_2 = \dots = m_k = m$, então, para todo $n \geq m$,

$$\begin{aligned} P\{D_k = (m, u_2, \dots, u_k) \mid N=n\} &= \\ &= \frac{n!}{(n-t_k)! \left[\binom{n}{m} \right]^k} \left[\prod_{j=1}^k \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m - u_j} \right] I_k(t_k) \quad (I.5) \end{aligned}$$

onde $I_k(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : m \leq x \leq \min\{km, n\}\}$.

Note que, se $m=1$, a fórmula (I.5) do corolário 6 se reduz a fórmula (I.1) do teorema 1.

Os dois exemplos que apresentamos a seguir ilustram as aplicações das fórmulas (I.5) do corolário 6 e (I.4) do teorema 5.

Exemplo 1 - Consideremos o caso onde $k=3$ e $m_1=m_2=m_3=2$. Para todo $u_j \in \{0,1,2\}$, $j=2,3$, temos $t_1=u_1=2$, $t_2=2+u_2$ e $t_3=2+u_2+u_3$. Para todo $n \geq 2$ segue-se, de (I.5), que

$$\begin{aligned} P\{D_3 = (2, u_2, u_3) | N=n\} &= \\ &= \frac{n!}{(n-t_3)! \left[\binom{n}{2} \right]^3} \left[\prod_{j=1}^3 \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^3 \binom{t_{j-1}}{2-u_j} \right] I_3(t_3) = \\ &= \frac{4n!}{(n-t_3)! [n(n-1)]^3} \frac{\binom{2}{u_2} \binom{t_2}{2-u_3}}{(u_2)! (u_3)!} I_3(t_3), \end{aligned}$$

onde $I_3(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : 2 \leq x \leq \min\{6, n\}\}$.

Exemplo 2 - Consideremos o caso onde $k=3$, $m_1=2, m_2=3$ e $m_3=1$. Para quaisquer $u_2 \in \{0,1,2,3\}$ e $u_3 \in \{0,1\}$ temos $t_1=u_1=2$, $t_2=2+u_2$ e $t_3=2+u_2+u_3$.

Para todo $n \geq 3$, segue-se, de (I.4), que

$$\begin{aligned}
 P\{D_3 = (2, u_2, u_3) | N=n\} &= \\
 &= \frac{n!}{2(n-t_3)! n \binom{n}{2} \binom{n}{3}} \frac{\binom{2}{3-u_2} (t_2)^{1-u_3}}{(u_2)!} I_3(t_3)
 \end{aligned}$$

onde $I_3(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : 3 \leq x \leq \min\{6, n\}\}$.

O lema que apresentamos a seguir é uma generalização do lema 2 e será de grande utilidade na determinação da distribuição de probabilidades de T_k ($k \geq 2$). Denotemos por $P_v(m_1, m_2, \dots, m_k; n)$ a probabilidade de que exatamente v urnas ($v \in \mathbb{N} = \{0, 1, 2, \dots\}$) resultem vazias, quando $M = m_1 + m_2 + \dots + m_k$ bolas ($k \geq 1$) coloridas são distribuídas aleatoriamente em n urnas ($n \geq \max\{m_1, m_2, \dots, m_k\}$), segundo o modelo descrito nesta seção. Lembremos que, as $M = m_1 + m_2 + \dots + m_k$ bolas são distribuídas aleatoriamente nas n urnas, em k etapas, do seguinte modo. Na j -ésima ($1 \leq j \leq k$) etapa, as m_j bolas coloridas com a j -ésima cor são distribuídas aleatoriamente nas n urnas, de maneira que, exatamente m_j urnas resultem ocupadas pelas bolas desta cor.

Lema 7 - Para todo $k \geq 1$ e para todo $n \geq \max\{m_1, m_2, \dots, m_k\}$,

$$\begin{aligned}
 P_v(m_1, m_2, \dots, m_k; n) &= \\
 &= \frac{\binom{n}{v}}{\prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^{n-v} (-1)^i \binom{n-v}{i} \prod_{j=1}^k \binom{n-v-i}{m_j} I(v) \quad (I.6)
 \end{aligned}$$

onde $I(\cdot)$ é a função indicadora do conjunto

$$\{x \in \mathbb{N} : n - \min\{M, n\} \leq x \leq n - \max\{m_1, m_2, \dots, m_k\}\}.$$

Prova - Considere o evento A_i : "a i -ésima urna resulta vazia", $i = 1, 2, \dots, n$ e represente a soma sobre o conjunto

$$J_i = \{(j_1, j_2, \dots, j_i) : 1 \leq j_1 < j_2 < \dots < j_i \leq n\}$$

por $\sum_{J_i}$.

A probabilidade de que pelo menos uma urna resulte vazia é dada por

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n (-1)^{i-1} \sum_{J_i} P(A_{j_1} \cap A_{j_2} \cap \dots \cap A_{j_i}) =$$

$$= \sum_{i=1}^n (-1)^{i-1} \sum_{J_i} \left[\frac{\prod_{j=1}^k (n-i)_{m_j}}{\prod_{j=1}^k (n)_{m_j}} \right] = m_j = m \text{ de bolas na urna } j$$

$(n-i)$ = nº de urnas que ficaram ocupadas.

$$= \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} \left[\frac{\prod_{j=1}^k (n-i)_{m_j}}{\prod_{j=1}^k (n)_{m_j}} \right] =$$

$$= \frac{1}{\prod_{j=1}^k \binom{n}{m_j}} \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} \prod_{j=1}^k \binom{n-i}{m_j}.$$

Logo, a probabilidade de que todas as urnas resultem ocupadas é dada por

$$P_0(m_1, m_2, \dots, m_k; n) = 1 - P\left(\bigcup_{i=1}^n A_i\right) =$$

$$= \frac{1}{\prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^n (-1)^i \binom{n}{i} \prod_{j=1}^k \binom{n-i}{m_j}.$$

Consideremos agora o evento "exatamente v urnas fixadas resultam vazias", onde

$$n - \min\{M, n\} \leq v \leq n - \max\{m_1, m_2, \dots, m_k\}.$$

Para este evento as $M = m_1 + m_2 + \dots + m_k$ bolas ocupam as restantes $n - v$ urnas, onde cada uma destas está ocupada. O número de casos possíveis satisfazendo este evento é

$$P_0(m_1, m_2, \dots, m_k; n - v) \prod_{j=1}^k \binom{n-v}{m_j},$$

uma vez que $\prod_{j=1}^k \binom{n-v}{m_j}$ é o número total de possíveis distribuições das M bolas nas $n - v$ urnas. Como v urnas podem ser escolhidas de $\binom{n}{v}$ maneiras, segue-se que

$$\begin{aligned}
P_v(m_1, \dots, m_k; n) &= \\
&= \frac{\binom{n}{v}}{\prod_{j=1}^k \binom{n}{m_j}} P_0(m_1, m_2, \dots, m_k; n-v) \prod_{j=1}^k \binom{n-v}{m_j} = \\
&= \frac{\binom{n}{v}}{\prod_{j=1}^k \binom{n}{m_j}} P_0(m_1, m_2, \dots, m_k; n-v) \prod_{j=1}^k \binom{n-v}{m_j}.
\end{aligned}$$

Para concluir o lema basta notar que

$$\begin{aligned}
P_0(m_1, \dots, m_k; n-v) &= \\
&= \frac{1}{\prod_{j=1}^k \binom{n-v}{m_j}} \sum_{i=0}^{n-v} (-1)^i \binom{n-v}{i} \prod_{j=1}^k \binom{n-v-i}{m_j}
\end{aligned}$$

e

$$P_v(m_1, m_2, \dots, m_k; n) = 0$$

se $v \notin \{x \in \mathbb{N} : n - \min\{M, n\} \leq x \leq n - \max\{m_1, m_2, \dots, m_k\}\}$.

□

Note que, se $m_1 = m_2 = \dots = m_k = 1$, a fórmula (I.6) do lema 7 se reduz a fórmula (I.2) do lema 2.

O principal resultado desta seção, a distribuição de probabilidades de $T_k = 1D'_k = \sum_{j=1}^k U_j$, será dado a seguir.

Teorema 8 - Para todo $k \geq 1$ e para todo $n \geq \max\{m_1, m_2, \dots, m_k\}$,

$$P\{T_k=t | N=n\} = \frac{n!}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^t (-1)^{t-i} \frac{\prod_{j=1}^k \binom{i}{m_j}}{i!(t-i)!} I_k(t) \quad (I.7)$$

onde $I_k(\cdot)$ é a função indicadora do conjunto

$$\{x \in \mathbb{N}^*: \max\{m_1, m_2, \dots, m_k\} \leq x \leq \min\{M, n\}\}.$$

Prova - Ao evento $\{T_k=t\}$ corresponde o evento "exatamente $t(n-t)$ urnas resultam ocupadas (vazias)" no modelo de urnas que estamos considerando. Logo,

$$P\{T_k=t | N=n\} = P_{n-t}(m_1, m_2, \dots, m_k; n)$$

e, como $I(n-t) = I_k(t)$, pelo lema 7 segue-se que

$$\begin{aligned} P\{T_k=t | N=n\} &= \frac{\binom{n}{n-t}}{\prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^t (-1)^i \binom{t}{i} \prod_{j=1}^k \binom{t-i}{m_j} I_k(t) = \\ &= \frac{\binom{n}{t}}{\prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^t (-1)^{t-i} \binom{t}{t-i} \prod_{j=1}^k \binom{i}{m_j} I_k(t) = \\ &= \frac{n!}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^t (-1)^{t-i} \frac{\prod_{j=1}^k \binom{i}{m_j}}{i!(t-i)!} I_k(t) . \end{aligned}$$

□

Note que, se $m_1=m_2=\dots=m_k = 1$, a fórmula (I.7) do teorema 8 se reduz a fórmula (I.3) do teorema 3.

Os dois exemplos que apresentamos a seguir ilustram a aplicação da fórmula (I.7) do teorema 8.

Exemplo 3 - (Continuação do exemplo 1)

Neste caso $k=3$ e $m_1=m_2=m_3= 2$. Para todo $n \geq 2$

$$\begin{aligned} P\{T_3=t|N=n\} &= \frac{n!}{(n-t)! \left[\binom{n}{2} \right]^3} \sum_{i=0}^t (-1)^{t-i} \frac{\left[\binom{i}{2} \right]^3}{i!(t-i)!} I_3(t) = \\ &= \frac{n!}{(n-t)! [n(n-1)]^3} \sum_{i=0}^t (-1)^{t-i} \frac{[i(i-1)]^3}{i!(t-i)!} I_3(t), \end{aligned}$$

onde $I_3(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^*: 2 \leq x \leq \min\{6, n\}\}$.

Exemplo 4 - (Continuação do exemplo 2)

Neste caso $k=3$, $m_1=2$, $m_2=3$ e $m_3=1$. Para todo $n \geq 3$,

$$\begin{aligned} P\{T_3=t|N=n\} &= \\ &= \frac{n!}{(n-t)! n \binom{n}{2} \binom{n}{3}} \sum_{i=0}^t (-1)^{t-i} \frac{i \binom{i}{2} \binom{i}{3}}{i!(t-i)!} I_3(t) = \\ &= \frac{n!}{(n-t)! n^3 (n-1)^2 (n-2)} \sum_{i=0}^t (-1)^{t-i} \frac{i^3 (i-1)^2 (i-2)}{i!(t-i)!} I_3(t), \end{aligned}$$

onde $I_3(\cdot)$ é a função indicadora do conjunto

$$\{x \in \mathbb{N}^* : 3 \leq x \leq \min\{6, n\}\}.$$

Considerando-se N como um parâmetro desconhecido e de interesse, D_k como os dados amostrais e T_k como uma estatística, os teoremas 5 e 8 nos permitem concluir, como na seção I.2, o seguinte resultado importante.

Corolário 9 - Para inferências sobre N , para todo $k \geq 1$, T_k é uma estatística suficiente com relação a D_k . Para todo $k \geq 2$ a distribuição de probabilidades condicional de D_k dado T_k é dada por

$$\begin{aligned} P\{D_k = (m_1, u_2, \dots, u_k) \mid T_k = t\} &= \\ &= \left[\prod_{j=1}^k \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j - u_j} \right] \left[\sum_{i=0}^t (-1)^{t-i} \frac{\prod_{j=1}^k \binom{i}{m_j}}{i! (t-i)!} \right]^{-1} I_{\{t\}}(t_k), \end{aligned}$$

onde $I_{\{t\}}(\cdot)$ é a função indicadora do conjunto $\{t\}$, com $\max\{m_1, m_2, \dots, m_k\} \leq t \leq M$.

O fator $\frac{n!}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}}$, comum às distribuições de

probabilidades de D_k e T_k , denomina-se "kernel da verossimilhança", pois é o único fator relevante para inferências Bayesianas ou estimativas de máxima verossimilhança para N .

I. 4 - Comentários e conclusões

A prova do teorema 5 pode ser simplificada utilizando-se diretamente, em cada uma das k etapas, a distribuição Hipergeométrica. Contudo, a obtenção da distribuição de probabilidades condicional de T_k dado N , à partir da distribuição de probabilidades condicional de $D_k = (U_1, U_2, \dots, U_k)$ dado N , é bastante complicada. Note que T_k é igual a soma das variáveis aleatórias dependentes e não identicamente distribuídas $U_1, U_2, \dots, U_k$. Em Zacks (1984) e Craig (1953) os autores consideram a distribuição de probabilidades condicional de T_k dado N como a solução da equação às diferenças satisfeita por essa distribuição. Nestes trabalhos apenas o caso particular de seleções um a um é considerado. Em Johnson e Kotz (1977) os autores obtêm a distribuição de probabilidades condicional de T_k dado N como uma aplicação dos chamados "problemas dos comitês".

O modelo de distribuição de bolas em urnas que consideramos neste capítulo nos permite obter a distribuição de probabilidades condicional de T_k dado N , utilizando apenas argumentos probabilísticos. Embora, para inferências sobre N possamos considerar a estatística T_k no lugar dos dados amostrais D_k , pois T_k é suficiente com relação a D_k , trabalharemos algumas vezes com D_k no lugar de T_k e em outras com T_k no lugar de D_k .

A vantagem de se trabalhar com D_k no lugar de T_k é decorrente do fato da sequência $\{D_k\}_{k \geq 1}$ gerar uma sequência

encaixante de σ -álgebras, fato que utilizaremos na sequência deste trabalho. Por outro lado, a vantagem de se trabalhar com T_k no lugar de D_k vem do fato de T_k ter a mesma dimensão de N , o parâmetro de interesse a ser estimado. A sequência deste trabalho é voltada para a estimação de N . Muitos problemas surgirão pelo fato do espaço paramétrico variar com o valor observado de T_k . Assim, qualquer distribuição atribuída a N terá suporte variando com T_k .

Um problema que naturalmente surge na estimação do valor de N é a escolha do esquema de amostragem. Dado o tamanho da amostra M ($M \geq 2$), como $M = \sum_{j=1}^k m_j$, podemos supor $k=2, 3, \dots, M$. Isto é, há $M-1$ esquemas de amostragem. Os casos extremos são os casos em que $k=2$ e $k=M$. No caso $k=2$ temos $m_1+m_2=M$, isto é, m_1 elementos são selecionados na 1ª etapa e $m_2 = M-m_1$ elementos são selecionados na 2ª etapa. No caso $k=M$ temos $m_1=m_2=\dots=m_k=1$, ou seja, M elementos são selecionados, com reposição, um de cada vez. Para k fixado ($2 \leq k \leq M$), o número de possibilidades em que o esquema de amostragem correspondente pode ser adotado é igual a $\binom{M-1}{k-1}$. Note que $\binom{M-1}{k-1}$ é igual ao número de soluções inteiras estritamente positivas da equação $m_1 + m_2 + \dots + m_k = M$.

REFERÊNCIAS PARA O CAPÍTULO I

- CRAIG, C.C. 1953. On utilization of marked specimens in estimating the population of flying insects. *Biometrika*, 40: 170-6.
- FELLER, W. 1968. *An introduction to probability theory and its applications*. 3.ed. New York, John Wiley. 2v.
- JOHNSON, N.L. & KOTZ, S. 1977. *Urn models and their application: an approach to modern discrete probability theory*. New York, John Wiley. 402p.
- ZACKS, S. 1984. *Bayes sequential estimation of the size of a finite population*. São Paulo, IME-USP. 21p.
(RT-MAE-8404)

CAPÍTULO II

ESTIMATIVAS DE MÁXIMA VEROSSIMILHANÇA

II. 1 - Introdução

Como foi mencionado anteriormente, o esquema de seleções por captura-recaptura tem como objetivo a estimação do tamanho, N , da população. Na realidade, este esquema produz uma função de verossimilhança que é o elemento fundamental para as inferências sobre o valor de N . Por outro lado, como não são os valores mas sim a forma da verossimilhança que caracteriza a estatística suficiente mínima (Zacks (1981)), podemos nos restringir à análise do kernel da verossimilhança.

Para $t \in \mathbb{N}^*$ satisfazendo a condição

$\max\{m_1, m_2, \dots, m_k\} \leq t \leq \sum_{j=1}^k m_j$, o kernel da verossimilhança foi definido como

$$K(N) = \frac{N!}{(N-t)! \prod_{j=1}^k \binom{N}{m_j}} I_t(N) ,$$

onde $I_t(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : x \geq t\}$.

No que segue vamos supor $k \geq 2$. Uma estimativa de máxima verossimilhança de N é um valor $\hat{N}$ de N que maximiza a função $K(\cdot)$. O objetivo deste capítulo é justamente o de

descrever como as estimativas de máxima verossimilhança, EMV, de N podem ser obtidas. A referência básica para este capítulo é Samuel (1968). Nas próximas seções apresentaremos as EMV de N para os casos onde $m_1 = m_2 = \dots = m_k = 1$ (seção II.2), $m_1 = m_2 = \dots = m_k = m$, com $m \geq 2$ (seção II.3) e para o caso de seleções com diferentes números de capturas (seção II.4). Comentários complementares são feitos na seção II.5.

II. 2 - Seleções um a um

Para $t \in \mathbb{N}^*$ satisfazendo a condição $1 \leq t \leq k$, o kernel da verossimilhança no caso onde $m_1 = m_2 = \dots = m_k = 1$ se reduz a

$$K(N) = \frac{N!}{(N-t)! N^k} I_t(N) ,$$

onde $I_t(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^*: x \geq t\}$.

Assim, determinar a EMV de N corresponde a determinar o ponto $\hat{N} \in \{x \in \mathbb{N}^*: x \geq t\}$ que maximiza a função $K(\cdot)$.

Um resultado imediato ocorre quando $t=1$ e quando $t=k$. O seguinte lema apresenta este resultado.

Lema 10 (*) - A EMV de N quando $t=1$ é $\hat{N}=1$ e quando $t=k$ é $\hat{N}=\infty$.

Prova - Se $t=1$, então $K(N) = \frac{1}{N^{k-1}}$. Logo $\hat{N}=1$. Se $t=k$, então

(*) - Nos casos em que a função de verossimilhança for estritamente crescente vamos supor que N assume valores no conjunto $\mathbb{N}^* \cup \{\infty\}$ e definimos $K(\infty) = \lim_{N \rightarrow \infty} K(N)$.

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

Logo $K(N) \uparrow 1$, quando $N \rightarrow \infty$, e $\hat{N} = \infty$.

Note que $\hat{N}=1$ quando $t=1$, independentemente do número de etapas k . O mesmo ocorre quando t assume seu valor máximo k . Assim, $\hat{N}=\infty$ no caso em que $t=k=2$ e no caso em que $t=k=2^{80}$. Por outro lado, $\hat{N}=1$ no caso em que $t=1$ e $k=2$ e no caso em que $t=1$ e $k=2^{80}$. Estes fatos prejudicam um pouco a total aceitação do método de máxima verossimilhança.

Analisaremos, na sequência, as EMV de N quando $1 < t < k$. Chamamos a atenção para o fato de que N assume valores no conjunto dos números inteiros maiores ou igual a t e desse modo, para determinar o ponto de máximo de $K(N)$ a razão

$$\frac{K(N+1)}{K(N)} = \left(1 - \frac{t}{N+1}\right)^{-1} \left(1 - \frac{1}{N+1}\right)^k,$$

para todo $N \in \mathbb{N}^*$ com $N \geq t$, é de interesse fundamental. Considere a função $g(\cdot)$ definida por

$$g(x) = (1-tx)^{-1} (1-x)^k,$$

para todo $x \in [0, \frac{1}{t})$. A função $g(\cdot)$ é contínua e na origem assume o valor $g(0) = 1$. Por outro lado, $\lim_{x \rightarrow \frac{1}{t}-} g(x) = \infty$ e,

para todo $N \in \mathbb{N}^*$ com $N \geq t$,

$$g\left(\frac{1}{N+1}\right) = \frac{K(N+1)}{K(N)}.$$

O seguinte lema descreve o comportamento da função $g(\cdot)$.

Lema 11 - Se $1 < t < k$ e $k_t = \frac{k-t}{kt-t}$, então $g(\cdot)$ é uma função monótona decrescente no intervalo $[0, k_t]$ e monótona crescente no intervalo $[k_t, \frac{1}{t})$.

Prova - A função derivada $g'(\cdot)$ de $g(\cdot)$ é dada por

$$g'(x) = [t-k+ktx-tx](1-tx)^{-2}(1-x)^{k-1}, \text{ para todo } x \in (0, \frac{1}{t}).$$

Como $(1-tx)^{-2}(1-x)^{k-1} > 0$, para todo $x \in (0, \frac{1}{t})$, segue-se

$$\text{que } g'(x) > 0 \Leftrightarrow t-k+(kt-t)x > 0 \Leftrightarrow x > \frac{k-t}{kt-t} = k_t \quad \text{e}$$

$$g'(x) < 0 \Leftrightarrow x < \frac{k-t}{kt-t} = k_t.$$

Como $g(\cdot)$ é contínua o lema está provado.

Note que k_t é o ponto de mínimo de $g(\cdot)$ no intervalo $[0, \frac{1}{t})$ e como $g(0) = 1$, então $g(k_t) < 1$.

A figura abaixo é o gráfico da função $g(\cdot)$ no caso em que $k=4$ e $t=2$.

FIGURA 1 - Gráfico da função $g(\cdot)$ com $k=4$ e $t=2$

Uma característica importante da função $g(\cdot)$ é que existe um único ponto x_0 , $k_t < x_0 < \frac{1}{t}$, tal que $g(x_0) = 1$. Ou seja, a equação $g(x) = 1$ tem duas soluções: uma nula e a outra, que denotamos por x_0 , satisfaz a condição $k_t < x_0 < \frac{1}{t}$. Assim, podemos garantir que existe um número $n \in \{x \in \mathbb{N}^* : x \geq t\}$ tal que

$$\frac{1}{n+1} \leq x_0 < \frac{1}{n} . \quad (\text{II.1})$$

O resultado que apresentamos a seguir mostra que a igualdade na relação (II.1) não pode ocorrer.

Lema 12 - Se $1 < t < k$, então as soluções da equação $g(x) = 1$ não são da forma $\frac{1}{n}$, onde $n \in \mathbb{N}^*$.

Prova - Note que $g(x) = 1 \Leftrightarrow (1-x)^k = 1-tx \Leftrightarrow \sum_{j=0}^k \binom{k}{j} (-1)^j x^j = 1-tx \Leftrightarrow 1-kx + \binom{k}{2} x^2 + \dots + (-1)^k x^k - 1+tx = 0 \Leftrightarrow$
 $\Leftrightarrow x[t-k+\binom{k}{2} x + \dots + (-1)^k x^{k-1}] = 0.$

Como já havíamos visto $x=0$ é uma solução de $g(x) = 1$. A outra solução, x_0 , não pode ser da forma $\frac{1}{n}$ com $n \in \mathbb{N}^*$, pois $x_0 < \frac{1}{t} < 1$ e para todo $n > 1$, n não é um divisor do coeficiente dominante $(-1)^k$ do polinômio

$$(-1)^k x^{k-1} + \dots + \binom{k}{2} x + (t-k).$$

O resultado acima mostra que $K(N+1) \neq K(N)$, para todo $N \in \{x \in \mathbb{N}^* : x \geq t\}$. Este fato nos permite enunciar o seguinte resultado que contradiz uma afirmação devida a Samuel (1968).

Teorema 13 - A EMV de N , $\hat{N}$, é única.

Prova - Se $t=1$ ou $t=k$ temos, pelo lema 10, que $\hat{N}=1$ ou $\hat{N}=\infty$, respectivamente.

Se $1 < t < k$, então, pelo lema 12, temos que, para todo $n \in \mathbb{N}^*$, $\frac{1}{n}$ não é solução da equação $g(x)=1$. Se x_0 é a solução não nula desta equação, então existe um número $n_0 \in \{x \in \mathbb{N}^*: x \geq t\}$ tal que $\frac{1}{n_0+1} < x_0 < \frac{1}{n_0}$. Se $n_0 = t$, isto é, se não existe $n \in \mathbb{N}^*$ tal que $x_0 < \frac{1}{n} < \frac{1}{t}$, então, para todo $n \in \mathbb{N}^*$, $n > t$, temos, pelo lema 11, que $g(\frac{1}{n}) < 1$. Logo,

$$K(t) > K(t+1) > K(t+2) > \dots$$

e a única EMV de N é $\hat{N} = t$.

Se $n_0 > t$, para todo $n \in \mathbb{N}^*$ com $t < n \leq n_0$, temos, pelo lema 11, que $g(\frac{1}{n}) > 1$. Logo,

$$K(n_0) > K(n_0-1) > K(n_0-2) > \dots > K(t).$$

Por outro lado, para todo $n \in \mathbb{N}^*$ com $n \geq n_0+1$, temos, pelo lema 11, que $g(\frac{1}{n}) < 1$. Logo, $K(n_0) > K(n_0+1) > K(n_0+2) > \dots$. Portanto, n_0 é o único ponto de máximo de $K(N)$. Assim, $\hat{N}=n_0$ é a única EMV de N .

A sequência de resultados apresentados nesta seção nos permite enunciar o seguinte resultado fundamental, que introduz a forma explícita da EMV de N .

Teorema 14 - Para todo $k \geq 2$, a EMV de N , $\hat{N}$, existe, é única e se expressa como:

$$\hat{N} = \begin{cases} 1 & \text{se } t=1 \\ t+N_0-1 & \text{se } 1 < t < k \\ \infty & \text{se } t=k \end{cases}$$

onde

$$N_0 = \min\{n \in \mathbb{N}^* : (t+n-1)^k < n(t+n)^{k-1}\}.$$

Prova - A existência e unicidade de $\hat{N}$ está garantida pelo teorema 13. Pelo lema 10, $\hat{N}=1$ se $t=1$ e $\hat{N}=\infty$ se $t=k$. Resta provar que $\hat{N} = t+N_0-1$ se $1 < t < k$. Se $1 < t < k$, pelos lemas 11 e 12, podemos concluir o seguinte fato com relação a função $g(\cdot)$. Existe um número $N_0 \in \mathbb{N}^*$ tal que, para todo $n \in \mathbb{N}^*$ com $n \geq N_0$, temos $g(\frac{1}{t+n}) < 1$ e, para todo $n \in \mathbb{N}^*$ com $n < N_0$, temos $g(\frac{1}{t+n}) > 1$.

Logo,

$$\begin{aligned} N_0 &= \min\{n \in \mathbb{N}^* : g(\frac{1}{t+n}) < 1\} = \\ &= \min\{n \in \mathbb{N}^* : (t+n-1)^k < n(t+n)^{k-1}\}, \end{aligned}$$

e temos que

$$K(t+N_0-1) > K(t+N_0-2) > K(t+N_0-3) > \dots > K(t)$$

e

$$K(t+N_0-1) > K(t+N_0) > K(t+N_0+1) > \dots$$

Portanto, $\hat{N} = t+N_0-1$ e o teorema está provado.

Note que se $t^k < (t+1)^{k-1}$, então $N_0 = 1$ e neste caso $\hat{N}=t$.

A tabela abaixo apresenta as EMV de N para
 $2 \leq k \leq 12$ e $1 \leq t \leq k$.

$\begin{matrix} k \\ t \end{matrix}$	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1
2	∞	2	2	2	2	2	2	2	2	2	2
3		∞	5	3	3	3	3	3	3	3	3
4			∞	8	6	5	4	4	4	4	4
5				∞	13	8	7	6	5	5	5
6					∞	19	11	9	8	7	7
7						∞	25	15	12	10	9
8							∞	33	19	15	12
9								∞	42	24	18
10									∞	51	29
11										∞	62
12											∞

TABELA 1 - Estimativas de Máxima Verossimilhança de N para $2 \leq k \leq 12$ e $1 \leq t \leq k$.

Determinamos as EMV de N para $2 \leq k \leq 2.000$ e constatamos o mesmo comportamento evidenciado na tabela 1.

A forma explícita do estimador de máxima verossimilhança é obtida, à partir do teorema 14, considerando-se T_k no lugar de t .

II.3 - Seleções com igual número de capturas

Nesta seção apresentamos as EMV de N para o caso onde $m_1 = m_2 = \dots = m_k = m$, com $m \geq 2$. Neste caso veremos que a EMV de N pode, dependendo dos valores de k e m , não ser única.

No que segue vamos supor $m \in \mathbb{N}^*$ com $m \geq 2$. Para todo $t \in \mathbb{N}^*$ com $m \leq t \leq km$, o kernel da verossimilhança se reduz a

$$K(N) = \frac{N!}{(N-t)! \left[\binom{N}{m} \right]^k} I_t(N),$$

onde $I_t(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : x \geq t\}$.

Nos casos em que $t=m$ e $t=km$ o seguinte lema apresenta as EMV de N .

Lema 15 - A EMV de N quando $t=m$ é $\hat{N}=m$ e quando $t=km$ é $\hat{N}=\infty$.

Prova - Se $t=m$, então, para todo $N \geq m$,

$$K(N) = \frac{m!}{\left[\binom{N}{m} \right]^{k-1}}.$$

Logo, $\hat{N}=m$. Se $t=km$, então, para todo $N \geq km$,

$$K(N) = (m!)^k \frac{\prod_{j=1}^{km-1} (1 - \frac{j}{N})}{\prod_{j=1}^{m-1} (1 - \frac{j}{N})^k} .$$

Por outro lado, como para todo número real não nulo x , $x < 1$, e para todo número inteiro k , $k \geq 2$,

$$(1-x)^k > 1-kx$$

temos que, para todo $N \geq km$, a razão

$$\frac{K(N+1)}{K(N)} = (1 - \frac{km}{N+1})^{-1} (1 - \frac{m}{N+1})^k > 1 .$$

Logo, $K(N) \uparrow (m!)^k$, quando $N \rightarrow \infty$, e $\hat{N} = \infty$.

Pelas mesmas razões dadas após o lema 10 (seção II.2), o fato de que $\hat{N}=m$ quando $t=m$, independentemente do número de etapas k , e $\hat{N}=\infty$ quando $t=km$, independentemente dos valores de k e m , prejudica um pouco a total aceitação do método de máxima verossimilhança.

Analisaremos, na sequência, as EMV de N quando $m < t < km$. Como na seção II.2, para determinar o ponto de máximo de $K(N)$ nos interessa a razão

$$\frac{K(N+1)}{K(N)} = (1 - \frac{t}{N+1})^{-1} (1 - \frac{m}{N+1})^k ,$$

para todo $N \in \mathbb{N}^*$ com $N \geq t$. Considere a função $g(\cdot)$ definida por

$$g(x) = (1-tx)^{-1} (1-mx)^k,$$

para todo $x \in [0, \frac{1}{t})$. A função $g(\cdot)$ é contínua, $g(0) = 1$,
 $\lim_{x \rightarrow \frac{1}{t}-} g(x) = \infty$ e, para todo $N \in \mathbb{N}^*$ com $N \geq t$,

$$g\left(\frac{1}{N+1}\right) = \frac{K(N+1)}{K(N)}.$$

O seguinte lema descreve o comportamento da função $g(\cdot)$.

Lema 16 - Se $m < t < km$ e $k_t = \frac{km-t}{m(kt-t)}$, então $g(\cdot)$ é uma função monótona decrescente no intervalo $[0, k_t]$ e monótona crescente no intervalo $[k_t, \frac{1}{t})$.

Prova - A função $g(\cdot)$ é derivável. A função derivada $g'(\cdot)$ de $g(\cdot)$ é dada por,

$$g'(x) = [t - km + m(kt - t)x] (1-tx)^{-2} (1-mx)^{k-1},$$

para todo $x \in (0, \frac{1}{t})$. Como $(1-tx)^{-2} (1-mx)^{k-1} > 0$, para todo $x \in (0, \frac{1}{t})$, segue-se que

$$g'(x) > 0 \Leftrightarrow t - km + m(kt - t)x > 0 \Leftrightarrow x > \frac{km-t}{m(kt-t)} = k_t$$

e

$$g'(x) < 0 \Leftrightarrow x < \frac{km-t}{m(kt-t)} = k_t.$$

Como $g(\cdot)$ é contínua o lema está provado.

Note que k_t é o ponto de mínimo de $g(\cdot)$ no intervalo $[0, \frac{1}{t})$ e como $g(0)=1$, então $g(k_t) < 1$.

A figura abaixo é o gráfico da função $g(\cdot)$ no caso em que $k=m=2$ e $t=3$.

FIGURA 2 - Gráfico da função $g(\cdot)$ com $k=m=2$ e $t=3$

Pelo que já sabemos sobre a função $g(\cdot)$ podemos concluir que existe um único ponto x_0 , $k_t < x_0 < \frac{1}{t}$, tal que $g(x_0) = 1$. Ou seja, a equação $g(x) = 1$ tem duas soluções: uma nula e a outra, que denotamos por x_0 , satisfaz a condição $k_t < x_0 < \frac{1}{t}$. Desse modo, podemos garantir que existe um número $n \in \{x \in \mathbb{N}^* : x \geq t\}$ tal que

$$\frac{1}{n+1} \leq x_0 < \frac{1}{n}. \quad (\text{II.2})$$

Note, por exemplo, que no caso em que $k=m=2$ e $t=3$ (veja figura 2) a igualdade na relação (II.2) é satisfeita para $n=3$. O resultado que apresentamos a seguir ilustra a condição em que a igualdade na relação (II.2) ocorre.

Lema 17 - Para $m < t < km$, se existe um número $n_0 \in \mathbb{N}^*$ que satisfaz a relação

$$(t+n_0-m)^k = n_0(t+n_0)^{k-1} \quad (\text{II.3})$$

então, $\frac{1}{t+n_0}$ é a única raiz não nula da equação $g(x)=1$.

$$\begin{aligned} \text{Prova - } g\left(\frac{1}{t+n_0}\right) &= \left(1 - \frac{t}{t+n_0}\right)^{-1} \left(1 - \frac{m}{t+n_0}\right)^k = \\ &= \frac{(t+n_0-m)^k}{n_0(t+n_0)^{k-1}} = 1. \end{aligned}$$

Apresentamos, para alguns valores de k , m e t , os respectivos números $n_0 \in \mathbb{N}^*$ que satisfazem a relação (II.3).

Para $k=2$, $m=4$ e $t=6$, $n_0=2$ e $g(\frac{1}{8}) = 1$.

Para $k=3$, $m=4$ e $t=7$, $n_0=1$ e $g(\frac{1}{8}) = 1$.

Para $k=4$, $m=64$ e $t=175$, $n_0=81$ e $g(\frac{1}{256}) = 1$.

Portanto, dados k , m e t , se existe um número $n_0 \in \mathbb{N}^*$ que satisfaz a relação (II.3), então,

$$\frac{K(t+n_0)}{K(t+n_0-1)} = g\left(\frac{1}{t+n_0}\right) = 1,$$

ou seja,

$$K(t+n_0) = K(t+n_0-1).$$

Este fato implica, como veremos, que a EMV de N não é única.

Os resultados apresentados nesta seção nos permitem enunciar o seguinte resultado fundamental, que introduz a forma explícita da EMV de N .

Teorema 18 - Para todo $k \geq 2$, uma EMV de N , $\hat{N}$, existe e se expressa como

$$\hat{N} = \begin{cases} m & \text{se } t=m \\ t+N_0-1 & \text{se } m < t < km \\ \infty & \text{se } t=km \end{cases}$$

onde

$$N_0 = \min\{n \in \mathbb{N}^*: (t+n-m)^k < n(t+n)^{k-1}\}.$$

Além disso, a unicidade desta estimativa é garantida exceto quando

$$(t+N_0-1-m)^k = (N_0-1)(t+N_0-1)^{k-1}.$$

Neste caso, as duas estimativas são $t+N_0-1$ e $t+N_0-2$.

Prova - Pelo lema 15, $\hat{N} = m$ se $t=m$ e $\hat{N} = \infty$ se $t=km$. Se $m < t < km$ os lemas 16 e 17 nos permitem concluir o seguinte fato, com respeito à função $g(\cdot)$. Existe um número $N_0 \in \mathbb{N}^*$ tal que,

$$g\left(\frac{1}{t+n}\right) < 1, \text{ para todo } n \in \mathbb{N}^* \text{ com } n \geq N_0;$$

$$g\left(\frac{1}{t+N_0-1}\right) \geq 1, \text{ se } N_0 \geq 2 \text{ e}$$

$$g\left(\frac{1}{t+n}\right) > 1, \text{ para todo } n \in \mathbb{N}^* \text{ com } n < N_0-1.$$

$$\text{Logo, } N_0 = \min\{n \in \mathbb{N}^*: g\left(\frac{1}{t+n}\right) < 1\} =$$

$$= \min\{n \in \mathbb{N}^*: (t+n-m)^k < n(t+n)^{k-1}\}$$

e temos que, se $(t+N_0-1-m)^k > (N_0-1)(t+N_0-1)^{k-1}$, então

$$g\left(\frac{1}{t+n}\right) < 1, \text{ para todo } n \in \mathbb{N}^* \text{ com } n \geq N_0 \text{ e}$$

$g\left(\frac{1}{t+n}\right) > 1$, para todo $n \in \mathbb{N}^*$ com $n < N_0$;

se $(t+N_0-1-m)^k = (N_0-1)(t-N_0-1)^{k-1}$, então $g\left(\frac{1}{t+n}\right) < 1$, para todo $n \in \mathbb{N}^*$ com $n \geq N_0$, $g\left(\frac{1}{t-N_0-1}\right) = 1$ e

$g\left(\frac{1}{t+n}\right) > 1$, para todo $n \in \mathbb{N}^*$ com $n < N_0-1$.

Portanto,

se $(t+N_0-1-m)^k > (N_0-1)(t+N_0-1)^{k-1}$, então

$$K(t+N_0-1) > K(t+N_0-2) > K(t+N_0-3) > \dots > K(t)$$

e

$$K(t+N_0-1) > K(t+N_0) > K(t+N_0+1) > \dots$$

Logo, $\hat{N} = t+N_0-1$;

se $(t+N_0-1-m)^k = (N_0-1)(t+N_0-1)^{k-1}$, então

$$K(t+N_0-1) = K(t+N_0-2) > K(t+N_0-3) > \dots > K(t)$$

e

$$K(t+N_0-1) > K(t+N_0) > K(t+N_0+1) > \dots$$

Logo, $\hat{N} = t+N_0-1$ e $\hat{N}-1$ são as duas EMV de N .

Apenas para efeito de ilustração a tabela abaixo apresenta as EMV de N para alguns valores de K , m e t .

k	m	t	EMV
2	2	3	3 e 4
2	3	5	6
2	4	6	7 e 8
3	4	7	7 e 8
3	4	9	14
4	6	11	11
4	16	30	31 e 32
4	32	60	63 e 64
4	64	175	255 e 256
5	32	62	63 e 64
6	2	6	6
10	15	76	91

TABELA 2 - Estimativas de Máxima Verossimilhança de N para alguns valores de k, m e t.

Determinamos as EMV de N para diversos valores de k, m e t. O comportamento dessas estimativas é análogo ao das apresentadas na tabela 2.

A forma explícita do estimador de máxima verossimilhança é obtida, neste caso, a partir do teorema 18, considerando-se T_k no lugar de t.

Na tabela 2 em alguns casos nos quais existiam duas EMV de N, verificamos que estas eram do tipo

$$2^{k+A+B} \quad \text{e} \quad 2^{k+A+B} - 1,$$

onde A e B são inteiros satisfazendo $A \geq 1-k$ e $B \geq 1$. Este fato nos levou a estabelecer o seguinte resultado.

Corolário 19 - Se $m = 2^{k+A}$ e $t = 2^{k+A+B} [1 - (1-2^{-B})^k]$ então, as duas EMV são $m2^B$ e $m2^B - 1$.

A demonstração deste resultado segue imediatamente do Teorema 18.

Na próxima seção discutiremos o caso onde os números das capturas nas diversas etapas podem ser diferentes.

II.4 - Seleções com diferentes números de capturas

Nesta seção discutiremos o caso geral onde os números de capturas nas diversas etapas são representados por $m_1, m_2, \dots, m_k$. Neste caso, para todo $t \in \mathbb{N}^*$ com $\max\{m_1, m_2, \dots, m_k\} \leq t \leq \sum_{j=1}^k m_j$, o kernel da verossimilhança é dado por

$$K(N) = \frac{N!}{(N-t)! \prod_{j=1}^k \binom{N}{m_j}} I_t(N),$$

onde $I_t(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^*: x \geq t\}$.

O seguinte lema apresenta as EMV de N nos dois casos extremos.

Lema 20 - A EMV de N quando $t = \max\{m_1, \dots, m_k\}$ é $\hat{N} = t$ e quando $t = \sum_{j=1}^k m_j$ é $\hat{N} = \infty$.

Prova - Podemos supor, sem perda de generalidade, que $m_k = \max\{m_1, m_2, \dots, m_k\}$. Se $t = m_k$, então, para todo $N \geq m_k$,

$$K(N) = \frac{(m_k)!}{\prod_{j=1}^{k-1} \binom{N}{m_j}}.$$

Logo, $\hat{N} = m_k$. Denotemos por s a soma $s = \sum_{j=1}^k m_j$.

Se $t = s$, então, para todo $N \geq s$,

$$K(N) = \left[\prod_{j=1}^k (m_j)! \right] \left[\frac{\prod_{j=1}^{s-1} \left(1 - \frac{j}{N}\right)}{\prod_{j=1}^k \prod_{r=0}^{m_j-1} \left(1 - \frac{r}{N}\right)} \right].$$

Por outro lado, como para todo $k \geq 2$ e para quaisquer números reais $x_1, x_2, \dots, x_k$ tais que $0 < x_j < 1$, $j=1, 2, \dots, k$,

$$\prod_{j=1}^k (1-x_j) > 1 - \sum_{j=1}^k x_j$$

temos que, para todo $N \geq s$, a razão

$$\frac{K(N+1)}{K(N)} = \left(1 - \frac{s}{N+1}\right)^{-1} \prod_{j=1}^k \left(1 - \frac{m_j}{N+1}\right) > 1.$$

Logo, $K(N) \uparrow \prod_{j=1}^k (m_j)!$, quando $N \rightarrow \infty$, e $\hat{N} = \infty$.

Quando $\max\{m_1, \dots, m_k\} < t < s = \sum_{j=1}^k m_j$, para todo $N \in \mathbb{N}^*$ com $N \geq t$,

$$\frac{K(N+1)}{K(N)} = \left(1 - \frac{t}{N+1}\right)^{-1} \prod_{j=1}^k \left(1 - \frac{m_j}{N+1}\right).$$

Seguindo a mesma linha das seções anteriores considere a função $g(\cdot)$ definida por

$$g(x) = (1-tx)^{-1} \prod_{j=1}^k (1-m_j x),$$

para todo $x \in [0, \frac{1}{t})$. A função $g(\cdot)$ é contínua, $g(0) = 1$,

$\lim_{x \rightarrow \frac{1}{t}^-} g(x) = \infty$ e, para todo $N \in \mathbb{N}^*$ com $N \geq t$,

$$\frac{K(N+1)}{K(N)} = g\left(\frac{1}{N+1}\right).$$

O resultado abaixo descreve o comportamento da função $g(\cdot)$.

Lema 21 - Se $\max\{m_1, \dots, m_k\} < t < s$, então a equação $g(x) = 1$ tem apenas uma solução, x_0 , não nula. Para todo $x \in (0, x_0)$ temos $g(x) < 1$ e para todo $x \in (x_0, \frac{1}{t})$ temos $g(x) > 1$.

Prova - Considere a função $g_1(\cdot)$ definida por

$g_1(x) = \prod_{j=1}^k (1-m_j x)$, para todo x real e a reta $g_2(x) = 1-tx$.

A função $g_1(\cdot)$ é contínua, decrescente e côncava para cima no intervalo $[0, \frac{1}{t}]$. Com efeito, para todo $x \in (0, \frac{1}{t})$, as funções derivada primeira e segunda de $g_1(\cdot)$ são

$$g_1'(x) = \sum_{i=1}^k (-m_i) \prod_{\substack{j=1 \\ j \neq i}}^k (1-m_j x) < 0$$

e

$$g_2'(x) = \sum_{i=1}^k \sum_{\substack{j=1 \\ j \neq i}}^k (m_i m_j) \prod_{\substack{r=1 \\ r \neq i, j}}^k (1-m_r x) > 0 .$$

Na origem e no ponto $\frac{1}{t}$ temos $g_1(0) = g_2(0) = 1$, $g_1(\frac{1}{t}) > 0$ e $g_2(\frac{1}{t}) = 0$. Por outro lado, a função derivada de $g_2(\cdot) - g_1(\cdot)$

na origem vale $g_2'(0) - g_1'(0) = -t + \sum_{j=1}^k m_j > 0$. Logo, como a função $g_2'(\cdot) - g_1'(\cdot)$ é contínua, segue-se que, existe um número real δ , $\delta > 0$, tal que $g_2'(x) - g_1'(x) > 0$, para todo $x \in [0, \delta)$. Então, pelo teorema do valor médio, para todo $x \in (0, \delta)$, temos que

$$g_2(x) > g_1(x) .$$

Portanto, existe um único ponto $x_0 \in (0, \frac{1}{t})$ tal que

$$g_1(x_0) = g_2(x_0)$$

$$g_1(x) < g_2(x), \text{ para todo } x \in (0, x_0)$$

$$g_1(x) > g_2(x), \text{ para todo } x \in (x_0, \frac{1}{t}).$$

Como $g(\cdot)$ é a restrição da função $\frac{g_1(\cdot)}{g_2(\cdot)}$ no intervalo $[0, \frac{1}{t}]$, o teorema está provado.

A. figura abaixo ilustra o comportamento das funções $g_1(\cdot)$ e $g_2(\cdot)$ restritas ao intervalo $[0, \frac{1}{t}]$.

FIGURA 3 - Gráfico ilustrativo do comportamento de $g_1(\cdot)$ e $g_2(\cdot)$ restritas ao intervalo $[0, \frac{1}{t}]$

O seguinte teorema resume os resultados desta seção e introduz a forma geral das EMV de N .

Teorema 22 - Para todo $k \geq 2$, uma EMV de N , $\hat{N}$, existe e se expressa como

$$\hat{N} = \begin{cases} M = \max\{m_1, \dots, m_k\} & \text{se } t=M \\ t+N_0-1 & \text{se } M < t < \sum_{j=1}^k m_j \\ \infty & \text{se } t = \sum_{j=1}^k m_j \end{cases}$$

onde

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

Além disso, a unicidade desta estimativa é garantida exceto quando

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

Neste caso, as duas estimativas são $t+N_0-1$ e $t+N_0-2$.

A demonstração deste teorema é análoga à demonstração do teorema 18.

Apenas para efeito de ilustração a tabela abaixo apresenta as EMV de N para alguns valores de k , $m_1, m_2, \dots, m_k$ e t .

k	m_j	t	EMV
2	$m_1 = 40$	62	63
	$m_2 = 60$	80	119 e 120
3	$m_1 = 1$	10	12
	$m_2 = 5$	11	16
	$m_3 = 8$	12	25
3	$m_1 = 40$	90	92
	$m_2 = 60$	120	152
	$m_3 = 80$	150	327
		179	10381
5	$m_1 = 3$	6	6
	$m_2 = 3$	7	7
	$m_3 = 4$	10	11
	$m_4 = 4$	18	139
	$m_5 = 5$		
5	$m_1 = 15$	60	61
	$m_2 = 20$	61	63
	$m_3 = 25$	80	95
	$m_4 = 30$	90	120
	$m_5 = 50$	100	159
		104	180
		120	347

TABELA 3 - Estimativas de Máxima Verossimilhança de N para alguns valores de $k, m_1, m_2, \dots, m_k$ e t .

A forma explícita do estimador de máxima verossimilhança é obtida, neste caso, a partir do teorema 22, considerando-se T_k no lugar de t .

Nos casos em que $k=2$ e $k=3$ o seguinte resultado mostra que as EMV de N se expressam de uma forma bastante simples.

Corolário 23 - a) Para $k=2$ e $\max\{m_1, m_2\} < t < m_1 + m_2$ denotemos por

$$A = \frac{m_1 m_2}{m_1 + m_2 - t}.$$

Se $A \in \mathbb{N}^*$, então as duas EMV de N são A e $A-1$.

Se $A \notin \mathbb{N}^*$, então a única EMV de N é $[A]$.

($[A]$ é o maior inteiro não superior à A)

b) Para $k=3$ e $\max\{m_1, m_2, m_3\} < t < \sum_{j=1}^3 m_j$, denotemos por

$$B = \frac{2a}{b - \sqrt{b^2 - 4ac}}$$

onde

$$a = m_1 m_2 m_3$$

$$b = m_1 m_2 + m_1 m_3 + m_2 m_3$$

$$c = m_1 + m_2 + m_3 - t.$$

Se $B \in \mathbb{N}^*$, então as duas EMV de N são B e $B-1$.

Se $B \notin \mathbb{N}^*$, então a única EMV de N é $[B]$.

A demonstração deste resultado segue imediatamente do teorema 22.

No caso em que $k \geq 4$ a obtenção de uma expressão para a estimativa, como às obtidas nos casos $k=2$ e $k=3$, é um problema complexo, pois depende da solução de uma equação algébrica de grau $k-1$.

Na próxima seção fazemos alguns comentários sobre os estimadores considerados neste capítulo.

II. 5 - Comentários e conclusões

O objetivo deste capítulo foi a simples obtenção das estimativas de máxima verossimilhança de N sem contudo analisar as propriedades distribucionais dos estimadores envolvidos. Note que os métodos utilizados não são padrões, uma vez que as restrições de contorno usuais não estão satisfeitas. Lembremos que além do espaço paramétrico ser enumerável, o suporte das distribuições dos estimadores dependem do valor do parâmetro. Isto nos impede de utilizar as propriedades assintóticas de normalidade das EMV. Gostaríamos de nos referir a Good (1950), página 73, onde encontramos uma sugestão para um problema distinto, do procedimento que utilizamos para a obtenção da EMV no caso de seleções um a um.

Evidentemente, para que a análise estatística seja mais completa, intervalos de confiança deveriam ser construí-

dos. Porém, devido aos problemas distribucionais acima referidos, esta construção torna-se bastante complexa. Além disso não é difícil provar que o estimador tem todos os momentos infinitos. Este fato nos levou a considerar a análise Bayesiana descrita nos capítulos subsequentes. Evidentemente, ao se considerar distribuições no espaço paramétrico de N , intervalos de credibilidade (o correspondente Bayesiano do intervalo de confiança) podem ser obtidos, pelo menos de forma numérica.

Por outro lado, as EMV aqui obtidas podem ser consideradas como estimativas Bayesianas. Basta que utilizemos como priori para N uma distribuição imprópria uniforme nos inteiros positivos e consideremos como estimativa a moda da distribuição à posteriori. É notável o fato de que a obtenção numérica dessas estimativas exige métodos iterativos bem simples e de fácil utilização. Isto é evidenciado pela própria definição das estimativas.

Uma outra razão para a utilização dos procedimentos Bayesianos é o fato de se poder eliminar estimativas patológicas como $\hat{N}=1$ e $\hat{N}=\infty$. Esta última estimativa é absurda, por exemplo, quando o espaço onde se encontra a população em estudo for limitado, uma vez que, somente um número finito de elementos cabem em um espaço limitado.

REFERÊNCIAS PARA O CAPÍTULO II

- DARROCH, J.N. 1958. The multiple-recapture census I. Estimation of a closed population. *Biometrika*, 45:343-59
- GOOD, I.J. 1950. *Probability and the weighing of evidence*. London, Charles Griffin. 119p.
- GOODMAN, L.A. 1953. Sequential sampling tagging for population size problems. *Annals of Mathematical Statistics*, 24:56-69.
- LEWONTIN, R.C. & PROUT, T. 1956. Estimation of the number of different classes in a population. *Biometrics*, 12: 211-23.
- SAMUEL, E. 1968. Sequential maximum likelihood estimation of the size of a population. *Annals of Mathematical Statistics*, 39:1057-68.
- SEBER, G.A.F. 1962. The multi-sample single recapture census. *Biometrika*, 49: 339-50.
- ZACKS, S. 1981. *Parametric statistical inference: basic theory and modern approaches*. Oxford, Pergamon Press, 387p.

CAPÍTULO III

ESTIMATIVAS DE BAYES

III.1 - Introdução

Os problemas que surgem com o uso das EMV, como descritos no capítulo anterior, nos levaram ao estudo, objeto deste capítulo, das estimativas de Bayes, EB, para o tamanho da população N .

Lembremos que, embora N seja uma constante fixada e desconhecida, segundo a perspectiva Bayesiana, para descrever o grau de desconhecimento sobre N devemos considerar uma distribuição de probabilidades sobre os possíveis valores de N , que no máximo é N^* . Isto equivale a considerarmos tecnicamente N como sendo uma variável aleatória assumindo valores em N^* . Antes de observarmos os dados, denotaremos por π a distribuição (a priori) de probabilidades de N .

Com a mesma notação dos capítulos anteriores, $\pi(n|D_k=\underline{d})$ é a probabilidade a posteriori do evento $\{N=n\}$, $n \in N^*$, quando são observados os dados $D_k=(m_1, u_2, \dots, u_k)=\underline{d}$. Assim, $\pi(\cdot|D_k)$ é uma distribuição de probabilidades de N condicionada no vetor aleatório D_k .

Como foi visto anteriormente, $T_k = 1D_k'$ é uma estatística suficiente para N . Logo, se $1\underline{d}'=t$, então, para todo $n \in N^*$,

$$\pi(n|D_k=d) = \pi(n|T_k=t). \quad (\text{III.1})$$

Na realidade esta relação entre D_k e T_k é a definição Bayesiana de estatística suficiente. Os conceitos clássico e Bayesiano de suficiência são equivalentes se o modelo $\{P(D_k|N=n) : n \in \mathbb{N}^*\}$ for dominado (o que é o nosso caso) e se (III.1) for válido para qualquer π considerada. Como T_k é suficiente, trabalharemos indiscriminadamente tanto com $\pi(\cdot|D_k)$ como com $\pi(\cdot|T_k)$, dependendo apenas da conveniência do momento.

O estimador de Bayes utilizado será a média da distribuição a posteriori. Isto é, $\bar{N} = E(N|D_k) = E(N|T_k)$. Na verdade esta é a função dos dados que minimiza o erro médio quadrático, ou seja,

$$E(N - \bar{N})^2 \leq E(N - F(D_k))^2$$

onde F é qualquer função real, mensurável, dos dados D_k . Sob o ponto de vista da teoria da decisão isto equivale a considerarmos uma função de perda quadrática.

A primeira referência de nosso conhecimento que trata do problema da estimação de N sob o ponto de vista Bayesiano é Freeman (1972). Nesse trabalho o autor considerou apenas o modelo simples de seleções um a um. Mais recentemente, Zacks (1984), para o mesmo modelo simples, tenta estudar as propriedades das EB. Neste mesmo trabalho Zacks apresenta algumas conjecturas sobre o comportamento das EB, pa

ra as quais apresentamos provas ou contra-exemplos.

Os resultados aqui obtidos, dos quais os resultados de Zacks (1984) são casos particulares, se referem ao modelo generalizado.

Na próxima seção introduzimos as EB com alguns exemplos. Na seção III.3 discutimos as propriedades gerais das EB.

III.2 - Obtenção das estimativas

Nesta seção determinamos as EB para o caso em que os números de capturas nas k etapas são representados por $m_1, m_2, \dots, m_k$.

No que segue vamos supor $k \geq 2$. Façamos

$s_k = \sum_{j=1}^k m_j$ e $M_k = \max\{m_1, m_2, \dots, m_k\}$. Para todo $t \in \mathbb{N}^*$ com $M_k \leq t \leq s_k$, lembremos que o kernel da verossimilhança é dado por

$$K(N) = \frac{N!}{(N-t)! \prod_{j=1}^k \binom{N}{m_j}} I_t(N),$$

onde $I_t(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : x \geq t\}$.

Seja π uma distribuição de probabilidades a priori de N . Para $t \in \mathbb{N}^*$ seja $A_{\pi, t} = \{x \in \mathbb{N}^* : x \geq t \text{ e } \pi(x) > 0\}$.

Para todo $t \in \mathbb{N}^*$ tal que $M_k \leq t \leq s_k$ e $A_{\pi, t} \neq \emptyset$, a correspondente distribuição de probabilidades a posteriori de N é dada por: para todo $n \in A_{\pi, t}$,

$$\begin{aligned} \pi(n | T_k = t) &= \frac{K(n) \pi(n)}{\sum_{n=t}^{\infty} K(n) \pi(n)} = \\ &= \mathbb{C}(k, t) \frac{N! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}}, \end{aligned}$$

onde

$$\mathbb{C}(k, t) = \left[\sum_{n=t}^{\infty} \frac{n! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} \right]^{-1}.$$

Note que, para todo $n \geq t$,

$$\begin{aligned} \frac{n!}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} &= \left[\prod_{j=1}^k (m_j)! \right] \left[\frac{\prod_{j=0}^{t-1} (n-j)}{\prod_{j=1}^k \prod_{r=0}^{m_j-1} (n-r)} \right] \leq \\ &\leq \left[\prod_{j=1}^k (m_j)! \right] \left[\frac{\prod_{j=0}^{t-1} (n-j)}{\prod_{j=1}^k (n-M_k+1)^{m_j}} \right] = \\ &= \left[\prod_{j=1}^k (m_j)! \right] \left[\frac{\prod_{j=0}^{t-1} (n-j)}{(n-M_k+1)^{s_k}} \right] = \\ &= \left[\prod_{j=1}^k (m_j)! \right] \left[\frac{\prod_{j=0}^{t-1} (1 - \frac{j}{n})}{(1 - \frac{M_k-1}{n})^t (n-M_k+1)^{s_k-t}} \right] \leq \end{aligned}$$

$$\leq \left[\prod_{j=1}^k (m_j)! \right] \left[1 - \frac{M_k-1}{n} \right]^{-t} \leq$$

$$\leq \left[\prod_{j=1}^k (m_j)! \right] \left[1 - \frac{M_k-1}{t} \right]^{-t} \quad (\text{III.2})$$

o que implica

$$\sum_{n=t}^{\infty} \frac{n! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} \leq \left[\prod_{j=1}^k (m_j)! \right] \left[1 - \frac{M_k-1}{t} \right]^{-t} \sum_{n=t}^{\infty} \pi(n) \leq$$

$$\leq \left[\prod_{j=1}^k (m_j)! \right] \left[1 - \frac{M_k-1}{t} \right]^{-t} < \infty.$$

Portanto, $0 < \mathbb{E}(k, t) < \infty$ e $\pi(\cdot | T_k = t)$ está bem definida. Observe que a condição $M_k \leq t \leq s_k$ é natural, pois $t < m_j$, para algum j , implicaria em retiradas com reposição. Por outro lado, se $t > s_k$ teríamos que admitir a existência de elementos já marcados na população antes do início do processo.

Para o problema em questão não existe uma família conjugada de distribuições a priori. Se admitirmos que $\pi(n) = 1$, para todo $n \in \mathbb{N}^*$, isto é, π é uma medida (não de probabilidade) uniforme em $\mathbb{N}^*$, então, para o caso simples de seleções um a um ($m_1 = m_2 = \dots = m_k = 1$), temos que

$$[\mathbb{E}(k, t)]^{-1} = \sum_{n=t}^{\infty} \frac{n!}{(n-t)! n^k} \text{ converge somente se } 1 \leq t \leq k-2; \text{ ou}$$

seja, $\pi(\cdot | T_k = t)$ não está definida para $t = k-1$ e $t = k$.

Suponhamos que a variância da distribuição de probabilidades a priori π seja finita. Para todo $t \in \mathbb{N}^*$ tal que $M_k \leq t \leq s_k$ e $A_{\pi,t} \neq \emptyset$ a EB é dada por

$$\begin{aligned} E(k,t) &= E(N|T_k=t) = \sum_{n=t}^{\infty} n \pi(n|T_k=t) = \\ &= \mathbb{E}(k,t) \sum_{n=t}^{\infty} \frac{n n! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}}. \end{aligned}$$

Note que, pelo fato da média de π ser finita e usando-se a desigualdade (III.2),

$$E(k,t) < \infty.$$

Para o caso de seleções um a um, para $1 \leq t \leq k$ tal que $A_{\pi,t} \neq \emptyset$, temos que

$$E(k,t) = \frac{\mathbb{E}(k,t)}{\mathbb{E}(k-1,t)}$$

onde

$$\mathbb{E}(k,t) = \left[\sum_{n=t}^{\infty} \frac{n! \pi(n)}{(n-t)! n^k} \right]^{-1}.$$

A seguir apresentamos alguns exemplos de EB para o caso em que a distribuição de probabilidades a priori é uma distribuição de Poisson^(*). Nos exemplos 5, 6 e 7 supomos que a distribuição de probabilidades a priori é

(*) - Note, neste caso, que o suporte da distribuição de probabilidades a priori é $\mathbb{N} = \mathbb{N}^* \cup \{0\}$. O fato de incluirmos o zero no suporte da distribuição de probabilidades a priori não apresenta nenhum inconveniente teórico.

dada por $\pi(n) = \frac{e^{-\lambda} \lambda^n}{n!}$, $n=0,1,2,\dots$, onde $\lambda > 0$.

Exemplo 5 - Para o caso de seleções um a um as EB são dadas por

$$E(k,t) = \frac{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^{k-1}}}{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^k}} ; t=1,2,\dots,k .$$

Note que

$$E(k,t) = \frac{E[(X+t)^{-k+1}]}{E[(X+t)^{-k}]}$$

onde X é uma variável aleatória com distribuição de Poisson com média λ .

Para efeito de ilustração a tabela abaixo apresenta as EB para o caso de seleções um a um quando a distribuição de probabilidades a priori é Poisson de parâmetro $\lambda = 20$, para $k=10,12$ e 15 .

t	k=10	k=12	k=15
1	1,0450	1,0061	1,0006
2	4,7857	2,4358	2,0628
3	10,2049	6,0156	3,5704
4	13,1565	9,8999	6,0322
5	15,2625	12,6860	8,9280
6	17,0382	14,8565	11,5406
7	18,6396	16,7130	13,7788
8	20,1317	18,3865	15,7434
9	21,5481	19,9408	17,5215
10	22,9085	21,4111	19,1690
11		22,8189	20,7212
12		24,1781	22,2011
13			23,6244
14			25,0023
15			26,3430

TABELA 4 - Estimativas de Bayes para o caso de seleções um a um quando a distribuição de probabilidades a priori é Poisson de parâmetro $\lambda=20$, para $k=10,12$ e 15 .

Exemplo 6 - Para o caso de seleções com $m_1=m_2=\dots m_k=2$ as EB são dadas por

$$E(k,t) = \frac{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! (n-1)^k n^{k-1}}}{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! [(n-1)n]^k}} ; t=2,3,\dots,2k .$$

Exemplo 7 - Para o caso em que $m_1=1$, $m_2=2$ e $m_3=3$ as EB são dadas por

$$E(3,t) = \frac{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)!(n-1)^2(n-2)n^2}}{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)!(n-1)^2(n-2)n^3}}; \quad t=3,4,5,6.$$

III.3 - Propriedades das Estimativas de Bayes

Dada uma distribuição de probabilidades a priori π de N , para todo $k \geq 2$, a estimativa de Bayes $E(k,t)$ é uma função do par ordenado (k,t) , onde t é um número inteiro tal que $M_k \leq t \leq s_k$ e $A_{\pi,t} \neq \emptyset$.

O seguinte teorema estabelece que, para t fixado, $E(k,t)$ é uma função não crescente de k .

Teorema 24 - Para toda distribuição de probabilidades a priori π de N e para todo $k \geq 2$, se $m_{k+1} \leq s_k$ então, para todo $t \in N^*$ tal que $M_{k+1} \leq t \leq s_k$ e $A_{\pi,t} \neq \emptyset$, temos que

$$E(k,t) \geq E(k+1,t).$$

Vale a igualdade na relação acima se $\pi(\cdot | T_k=t)$ for degenerada em um ponto.

Prova - Suponha $m_{k+1} \leq s_k$. Para todo $t \in \mathbb{N}^*$ tal que $m_{k+1} \leq t \leq s_k$ e $A_{\pi, t} \neq \emptyset$, considere a função $h(\cdot)$ definida em $\mathbb{N}^*$ por

$$h(n) = \begin{cases} 0 & \text{se } n < t \\ \frac{1}{\binom{n}{m_{k+1}}} & \text{se } n \geq t. \end{cases}$$

Como a função $h(\cdot)$ restrita ao conjunto $A_{\pi, t}$ é decrescente, segue-se que (Lehmann (1966))

$$E(h(N) | T_k = t) E(N | T_k = t) \geq E(Nh(N) | T_k = t).$$

Por outro lado,

$$E(h(N) | T_k = t) = \mathbb{Q}(k, t) \sum_{n=t}^{\infty} \frac{n! \pi(n)}{(n-t)! \prod_{j=1}^{k+1} \binom{n}{m_j}} = \frac{\mathbb{Q}(k, t)}{\mathbb{Q}(k+1, t)}$$

e

$$\begin{aligned} E(Nh(N) | T_k = t) &= \mathbb{Q}(k, t) \sum_{n=t}^{\infty} \frac{n n! \pi(n)}{(n-t)! \prod_{j=1}^{k+1} \binom{n}{m_j}} = \\ &= \frac{\mathbb{Q}(k, t)}{\mathbb{Q}(k+1, t)} E(N | T_{k+1} = t). \end{aligned}$$

Logo, das relações acima segue-se que

$$\frac{\mathbb{Q}(k, t)}{\mathbb{Q}(k+1, t)} E(k, t) \geq \frac{\mathbb{Q}(k, t)}{\mathbb{Q}(k+1, t)} E(k+1, t),$$

ou seja,

$$E(k, t) \geq E(k+1, t)$$

e o teorema está provado. □

Seja π uma distribuição de probabilidades a priori de N . Para todo $t \in \mathbb{N}^*$ tal que $M_k \leq t \leq s_k$ e $A_{\pi, t} \neq \emptyset$ a variância da distribuição de probabilidades a posteriori é dada por

$$\begin{aligned} \text{Var}(N|T_k=t) &= E(N^2|T_k=t) - [E(N|T_k=t)]^2 = \\ &= E(N(N-t)|T_k=t) - E(N|T_k=t)E(N-t|T_k=t). \end{aligned}$$

Note que, pelo fato da variância de π ser finita e usando-se a desigualdade (III.2), $\text{Var}(N|T_k=t) < \infty$. Por outro lado, se $M_k \leq t \leq s_k - 1$, $A_{\pi, t} \neq \emptyset$ e $\pi(\cdot|T_k=t)$ não for degenerada no ponto t , então

$$\begin{aligned} E(N(N-t)|T_k=t) &= \mathbb{C}(k, t) \sum_{n=t+1}^{\infty} \frac{n(n-t)n! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} = \\ &= \mathbb{C}(k, t) \sum_{n=t+1}^{\infty} \frac{n n! \pi(n)}{(n-t-1)! \prod_{j=1}^k \binom{n}{m_j}} = \\ &= \frac{\mathbb{C}(k, t)}{\mathbb{C}(k, t+1)} E(N|T_k=t+1) \end{aligned}$$

e

$$\begin{aligned}
E(N-t | T_k=t) &= E(N | T_k=t) - t = \\
&= \mathbb{C}(k, t) \sum_{n=t+1}^{\infty} \frac{(n-t)n! \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} = \\
&= \mathbb{C}(k, t) \sum_{n=t+1}^{\infty} \frac{n! \pi(n)}{(n-t-1)! \prod_{j=1}^k \binom{n}{m_j}} = \frac{\mathbb{C}(k, t)}{\mathbb{C}(k, t+1)}.
\end{aligned}$$

Logo, das igualdades acima segue-se que

$$\text{Var}(N | T_k=t) = \frac{\mathbb{C}(k, t)}{\mathbb{C}(k, t+1)} [E(k, t+1) - E(k, t)]. \quad (\text{III.3})$$

O seguinte teorema estabelece que, para k fixado, $E(k, t)$ é uma função não decrescente de t .

Teorema 25 - Seja π uma distribuição de probabilidades a priori de N . Para todo $k \geq 2$ e para todo $t \in \mathbb{N}^*$ tal que $M_k \leq t \leq s_k - 1$ e $A_{\pi, t} \neq \emptyset$, se $\pi(\cdot | T_k=t)$ não for degenerada no ponto t , então

$$E(k, t) \leq E(k, t+1).$$

Vale a igualdade na relação acima se $\pi(\cdot | T_k=t)$ for degenerada em um ponto diferente de t .

Prova - Para todo $t \in \mathbb{N}^*$ tal que $M_k \leq t \leq s_k - 1$ e $A_{\pi, t} \neq \emptyset$, se

$\pi(\cdot|T_k=t)$ não for degenerada no ponto t , então $E(k,t)$ e $E(k,t+1)$ estão bem definidos. Como $\text{Var}(N|T_k=t) \geq 0$ o resultado segue da igualdade (III.3).

Com relação a $R(k,t) = \text{Var}(N|T_k=t)$, Zacks (1984) estabelece os seguintes fatos, para o caso de seleções um a um.

19) Para toda distribuição de probabilidades a priori e para t fixado, $R(k,t)$ é uma função não crescente de k .

20) Para toda distribuição de probabilidades a priori e para k fixado, $R(k,t)$ é uma função não decrescente de t .

Finalizando esta seção daremos um exemplo para mostrar que os fatos acima são falsos.

Exemplo 8 - Suponha que a distribuição de probabilidades a priori seja Poisson de parâmetro $\lambda=100$. Neste caso,

$$R(9,5) = 99,98439$$

$$R(10,5) = 100,0416$$

$$R(10,6) = 99,91608$$

$$R(10,7) = 99,79708.$$

Ou seja, $R(9,5) < R(10,5)$ e $R(10,6) > R(10,7)$, contrariando o estabelecido por Zacks (1984).

REFERÊNCIAS PARA O CAPÍTULO III

- FREEMAN, P.R. 1972. Sequential estimation of the size of a population. *Biometrika*, 59:9-17.
- LEHMANN, E.L. 1966. Some concepts of dependence. *Annals of Mathematical Statistics*, 37:1137-53.
- ZACKS, S. 1984. Bayes sequential estimation of the size of a finite population. São Paulo, IME-USP. 21p. (RT-MAE-8404).

oOo

CAPÍTULO IV

RESULTADOS ASSINTÓTICOS

IV. 1 - Introdução

Neste capítulo faremos um estudo das propriedades assintóticas das EB e do estimador de Bayes, $\bar{N}=E(N|D_k)$, para o tamanho da população. Os resultados aqui apresentados se referem ao modelo generalizado.

Para estudar o comportamento assintótico de $\bar{N}$ construiremos um espaço de probabilidades compatível com o modelo proposto, definiremos as variáveis aleatórias N , D_k e T_k nesse espaço e exploraremos o fato de que as sequências $\{E(N|D_k)\}_{k \geq 2}$ e $\{\text{Var}(N|D_k)\}_{k \geq 2}$ são, respectivamente, um martingal e um supermartingal.

Na próxima seção estudaremos as propriedades assintóticas das EB. Na seção IV.3 construiremos um espaço de probabilidades onde as variáveis aleatórias de interesse estão definidas e, finalmente, na seção IV.4 estudaremos o comportamento assintótico de $\bar{N}$.

IV. 2 - Propriedades assintóticas das estimativas de Bayes

Dada uma sequência limitada $\{m_n\}_{n \geq 1}$ de elementos de $\mathbb{N}^*$, seja $M = \max\{m_n : n \geq 1\}$. Para cada $n \in \mathbb{N}^*$, m_n representa o número de elementos capturados na n -ésima etapa. A razão de considerarmos $m_n \leq M$, para todo $n \in \mathbb{N}^*$, é devida ao fato de que a população é finita.

Com a mesma notação do capítulo anterior, dada uma distribuição de probabilidades a priori π , para cada $t \in \mathbb{N}^*$, fixado, satisfazendo a condição $t \geq M$ e $A_{\pi, t} \neq \emptyset$ seja $k_0 = \min\{j \in \mathbb{N}^* : j \geq 2 \text{ e } m_1 + m_2 + \dots + m_j \geq t\}$. Logo, para todo $k \geq k_0$, temos $M_k \leq t \leq s_k$ e $E(k, t)$ e $\text{Var}(N|T_k=t)$ estão bem definidos. Como, para todo $k \geq k_0$, $E(k, t) \geq 1$ e, pelo teorema 24, a sequência $\{E(k, t)\}_{k \geq k_0}$ é não crescente, então existe o limite de $\{E(k, t)\}_{k \geq k_0}$, quando $k \rightarrow \infty$.

O seguinte teorema estabelece o valor deste limite.

Teorema 26 - Para toda distribuição de probabilidades a priori π de N e para todo $t \in \mathbb{N}^*$, fixado, tal que $t \geq M$ e $A_{\pi, t} \neq \emptyset$, temos

$$\lim_{k \rightarrow \infty} E(k, t) = \ell,$$

onde $\ell = \min A_{\pi, t}$.

Prova - Para todo $k \geq k_0$,

$$\begin{aligned}
 E(k, t) &= \frac{\sum_{n=\ell}^{\infty} \frac{n \cdot n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}}}{\sum_{n=\ell}^{\infty} \frac{n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}}} = \\
 &= \frac{\ell + \frac{(\ell-t)!}{\ell! \cdot \pi(\ell)} \sum_{n=\ell+1}^{\infty} \frac{n \cdot n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right]}}{1 + \frac{(\ell-t)!}{\ell! \cdot \pi(\ell)} \sum_{n=\ell+1}^{\infty} \frac{n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right]}}.
 \end{aligned}$$

Para provar o teorema basta mostrar que

$$\sum_{n=\ell+1}^{\infty} \frac{n \cdot n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right]} \xrightarrow{k \rightarrow \infty} 0$$

e

$$\sum_{n=\ell+1}^{\infty} \frac{n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right]} \xrightarrow{k \rightarrow \infty} 0.$$

Contudo, para todo $k > t+1$

$$\begin{aligned}
 \sum_{n=\ell+1}^{\infty} \frac{n \cdot n! \cdot \pi(n)}{(n-t)! \prod_{j=1}^k \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right]} &\leq \sum_{n=\ell+1}^{\infty} \frac{n \cdot n! \cdot \pi(n)}{(n-t)!} \left(\frac{\ell}{n} \right)^k = \\
 &= \ell^{t+1} \sum_{n=\ell+1}^{\infty} \frac{n! \cdot \pi(n)}{(n-t)! n^t} \left(\frac{\ell}{n} \right)^{k-t-1} \leq
 \end{aligned}$$

$$\begin{aligned}
&\leq \ell^{t+1} \sum_{n=\ell+1}^{\infty} \pi(n) \left(\frac{\ell}{n}\right)^{k-t-1} \leq \\
&\leq \ell^{t+1} \left(\frac{\ell}{\ell+1}\right)^{k-t-1} \sum_{n=\ell+1}^{\infty} \pi(n) < \ell^{t+1} \left(\frac{\ell}{\ell+1}\right)^{k-t-1} = \\
&= (\ell+1)^{t+1} \left(\frac{\ell}{\ell+1}\right)^k \xrightarrow[k \rightarrow \infty]{} 0
\end{aligned}$$

e, para todo $k > t$.

$$\begin{aligned}
&\sum_{n=\ell+1}^{\infty} \frac{n! \pi(n)}{(n-t)!} \frac{k}{\prod_{j=1}^t n_j} \left[\frac{\binom{\ell}{m_j}}{\binom{n}{m_j}} \right] \leq \sum_{n=\ell+1}^{\infty} \frac{n! \pi(n)}{(n-t)!} \left(\frac{\ell}{n}\right)^k = \\
&= \ell^t \sum_{n=\ell+1}^{\infty} \frac{n! \pi(n)}{(n-t)! n^t} \left(\frac{\ell}{n}\right)^{k-t} \leq \ell^t \sum_{n=\ell+1}^{\infty} \pi(n) \left(\frac{\ell}{n}\right)^{k-t} \leq \\
&\leq \ell^t \left(\frac{\ell}{\ell+1}\right)^{k-t} \sum_{n=\ell+1}^{\infty} \pi(n) < (\ell+1)^t \left(\frac{\ell}{\ell+1}\right)^k \xrightarrow[k \rightarrow \infty]{} 0
\end{aligned}$$

e o teorema está provado. □

O seguinte corolário do Teorema 26 estabelece que a sequência $\{\text{Var}(N|T_k=t)\}_{k \geq k_0}$ converge para zero, quando $k \rightarrow \infty$.

Corolário 27- Para toda distribuição de probabilidades a priori π de N e para cada $t \in \mathbb{N}^*$, fixado, tal que $t \geq M$ e $A_{\pi,t} \neq \emptyset$, temos que

$$\lim_{k \rightarrow \infty} \text{Var}(N|T_k=t) = 0.$$

Prova - Se $\pi(\cdot|T_{k_0}=t)$ for degenerada em um ponto, então, para todo $k \geq k_0$, $\pi(\cdot|T_k=t)$ é degenerada em um ponto. Logo, para todo $k \geq k_0$, $\text{Var}(N|T_k=t) = 0$ e vale o resultado.

Se $\pi(\cdot|T_{k_0}=t)$ não for degenerada em um ponto, então, para todo $k > k_0$, $\pi(\cdot|T_k=t)$ não é degenerada no ponto t e, pela fórmula (III.3),

$$\begin{aligned} \text{Var}(N|T_k=t) &= \frac{\mathbb{C}(k,t)}{\mathbb{C}(k,t+1)} [E(k,t+1) - E(k,t)] = \\ &= [E(k,t) - t] [E(k,t+1) - E(k,t)]. \end{aligned}$$

Note que, como $M_k \leq t < s_k$, $E(k,t+1)$ está bem definido. Pelo teorema 26

$$[E(k,t) - t] [E(k,t+1) - E(k,t)] \xrightarrow[k \rightarrow \infty]{} 0$$

e o corolário está provado.

Com o objetivo de ilustrar o teorema 26 e o corolário 27, apresentamos o seguinte exemplo.

Exemplo 9 - Para o caso de seleções um a um e π uma distribuição de Poisson truncada no ponto zero, de parâmetro λ , $\lambda > 0$, temos

$$M = m_n = 1, \text{ para todo } n \in \mathbb{N}^* ;$$

$$\pi(n) = \frac{e^{-\lambda} \lambda^n}{(1-e^{-\lambda}) n!}, \quad n=1, 2, \dots;$$

$$\text{para todo } t \in \mathbb{N}^*, \text{ fixado, } A_{\pi, t} = \{t, t+1, \dots\}$$

e, para todo $k \geq \max\{2, t\}$,

$$E(k, t) = \frac{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^{k-1}}}{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^k}},$$

$$\text{Var}(N|T_k=t) = E(N^2|T_k=t) - [E(N|T_k=t)]^2 =$$

$$= \frac{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^{k-2}}}{\sum_{n=t}^{\infty} \frac{\lambda^n}{(n-t)! n^k}} - [E(k, t)]^2.$$

Pelo teorema 26 e corolário 27,

$$E(k, t) \xrightarrow[k \rightarrow \infty]{} t \quad \text{e} \quad \text{Var}(N|T_k=t) \xrightarrow[k \rightarrow \infty]{} 0.$$

IV.3 - Modelo estatístico

Como na seção anterior, supomos que $\{m_n\}_{n \geq 1}$ seja uma sequência limitada de elementos de $\mathbb{N}^*$, onde $M = \max\{m_n : n \geq 1\}$. Façamos $M_1 = s_1 = m_1$, $M_j = \max\{m_1, \dots, m_j\}$ e $s_j = \sum_{i=1}^j m_i$, $j=2,3,\dots$.

Sejam

$$\mathbb{C}_M = \{x \in \mathbb{N}^* : x \geq M\};$$

$\{A_n\}_{n \geq 1}$ a sequência de subconjuntos de $\mathbb{N}^*$ definida por $A_n = \{0, 1, \dots, m_n\}$, $n=1,2,\dots$;

$$\Omega = \{\omega = (\omega_1, \omega_2, \dots) : \omega_j \in A_j, j=1,2,\dots\}$$

e

$\mathcal{F}$ a σ -álgebra gerada pelos conjuntos da forma $\{\omega \in \Omega : \omega_1 \in B_1, \dots, \omega_n \in B_n\}$, $B_j \subset A_j$, $1 \leq j \leq n$, $n=1,2,\dots$.

O espaço mensurável $(\Omega, \mathcal{F})$ é o espaço das observações do experimento.

Como no teorema 5, denotemos por $t_1, t_2, \dots, t_k$ as somas $t_j = \sum_{i=1}^j u_i$, onde $t_1 = u_1 = m_1$ e $u_j \in A_j$, $j=2,3,\dots,k$.

Para todo $n \in \mathbb{C}_M$ seja P_n uma medida de probabilidade definida em $(\Omega, \mathcal{F})$ tal que

$$P_n(\{\omega \in \Omega : \omega_1 = m_1\}) = 1 \quad (\text{IV.1}) ;$$

para todo $k \geq 2$

$$\begin{aligned}
P_n(\{\omega \in \Omega : \omega_1 = m_1, \omega_2 = u_2, \dots, \omega_k = u_k, u_j \in A_j, j=2, \dots, k\}) = \\
= \frac{n!}{(n-t_k)! \prod_{j=1}^k \binom{n}{m_j}} \left[\prod_{j=1}^k \frac{1}{(u_j)!} \right] \left[\prod_{j=2}^k \binom{t_{j-1}}{m_j - u_j} \right] I_k(t_k) \quad (\text{IV.2})
\end{aligned}$$

onde $I_k(\cdot)$ é a função indicadora do conjunto $\{x \in \mathbb{N}^* : M_k \leq x \leq \min\{s_k, n\}\}$.

Note que, para todo $n \in \mathbb{C}_M$ e para todo $k \geq 2$, por (IV.2) e pelas fórmulas (I.4) do teorema 5 e (I.7) do teorema 8, para cada $t \in \mathbb{N}^*$

$$\begin{aligned}
P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = t\}) = \\
= P_n(\{\omega \in \Omega : M_j \leq \sum_{i=1}^j \omega_i \leq \min\{s_j, n\}, j=1, 2, \dots, k-1 \text{ e } \sum_{i=1}^k \omega_i = t\}) = \\
= \begin{cases} \frac{n!}{(n-t)! \prod_{j=1}^k \binom{n}{m_j}} \sum_{i=0}^t (-1)^{t-i} \frac{\prod_{j=1}^k \binom{i}{m_j}}{i!(t-i)!} & \text{se } M_k \leq t \leq \min\{s_k, n\} \\ 0 & \text{caso contrário.} \end{cases} \quad (\text{IV.3})
\end{aligned}$$

O espaço $(\Omega, F, \{P_n : n \in \mathbb{C}_M\})$ se denomina modelo estatístico. Note que a fórmula (IV.2) representa a probabilidade de observarmos o vetor $(m_1, u_2, \dots, u_k)$ quando o tamanho da população é n . Por outro lado, a fórmula (IV.3) representa a probabilidade de observarmos exatamente t elementos da população, em k etapas, quando o tamanho da população é n .

Sejam

G a σ -álgebra dos subconjuntos de $\mathbb{C}_M$;

π uma medida de probabilidade a priori de N ,

cujo suporte $A_\pi = \{x \in \mathbb{N}^* : \pi(x) > 0\} \subseteq \mathbb{C}_M$;

$\Omega^* = \Omega \times \mathbb{C}_M$;

$F^* = F \otimes G$: σ -álgebra produto de F e G

e

$\mathbb{P}$ uma medida de probabilidade definida em

(Ω^*, F^*) tal que, para todo $A \in F$ e para todo $B \in G$,

$$\mathbb{P}(A \times B) = \sum_{n \in B} P_n(A) \pi(n).$$

O espaço de probabilidades $(\mathbb{C}_M, G, \pi)$ é conhecido como modelo a priori.

Pelo que apresentamos até agora, podemos supor que N, T_k e D_k são variáveis aleatórias definidas em $(\Omega^*, F^*, \mathbb{P})$, a valores em $\mathbb{N}$ e $\mathbb{N}^k$, do seguinte modo:

$$N(\omega, n) = n;$$

$$T_1(\omega, n) = \omega_1 ;$$

para todo $k \geq 2$

$$D_k(\omega, n) = (\omega_1, \omega_2, \dots, \omega_k) \text{ e } T_k(\omega, n) = \sum_{i=1}^k \omega_i$$

para todo $(\omega, n) \in \Omega^*$.

Não é difícil verificar os seguintes fatos:

- 1) $\mathbb{P}\{N=n\} = \pi(n)$, para todo $n \in \mathbb{C}_M$;
- 2) se $n \in A_\pi$, então $\mathbb{P}\{T_1=m_1 | N=n\} = 1$;
- 3) se $n \in A_\pi$, para todo $k \geq 2$

$$\mathbb{P}\{D_k = (m_1, u_2, \dots, u_k) | N=n\} =$$

$$= P_n(\{\omega \in \Omega : \omega_1=m_1, \omega_2=u_2, \dots, \omega_k = u_k\})$$

e

$$\mathbb{P}\{T_k=t | N=n\} = P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = t\}) =$$

$$= P_n(\{\omega \in \Omega : M_j \leq \sum_{i=1}^j \omega_i \leq \min\{s_j, n\}, j=1, 2, \dots, k-1 \text{ e } \sum_{i=1}^k \omega_i = t\}).$$

Logo, dos fatos acima e de (IV.2) e (IV.3) podemos concluir que, o modelo probabilístico que construímos é coerente com o problema que estamos considerando.

O seguinte lema estabelece que a sequência de variáveis aleatórias $\{T_k\}_{k \geq 1}$ converge em probabilidade para N , quando $k \rightarrow \infty$.

Lema 28 - i) Para todo $k \geq 1$, $T_k \leq N$ q.c. (quase certamente) ($\mathbb{P}$).

ii) $T_k \xrightarrow{\mathbb{P}} N$; isto é, T_k converge em probabilidade para N .

Prova - i) Para todo $n \in A_\pi$ e para todo $k \geq 1$, de (IV.1) e (IV.3) segue-se que

$$P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i > n\}) = 0.$$

Logo, para todo $k \geq 1$,

$$\mathbb{P}\{T_k > N\} = \sum_{n \in A_\pi} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i > n\}) \pi(n) = 0$$

o que implica $T_k \leq N$ q.c. ($\mathbb{P}$).

(ii) Para cada $n \in A_\pi$ provemos inicialmente que

$$\lim_{k \rightarrow \infty} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = n\}) = 1.$$

De fato, para cada $n \in A_\pi$ existe um número inteiro positivo k_1 (que depende de n e da sequência $\{m_n\}_{n \geq 1}$), tal que, para todo número inteiro $k \geq k_1$, $n \leq s_k$. Logo, por (IV.3) segue-se que, para todo $k \geq k_1$

$$\begin{aligned} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = n\}) &= \\ &= n! \sum_{i=0}^n (-1)^{n-i} \frac{1}{i! (n-i)!} \prod_{j=1}^k \left[\frac{\binom{i}{m_j}}{\binom{n}{m_j}} \right] = \\ &= 1 + n! \sum_{i=0}^{n-1} (-1)^{n-i} \frac{1}{i! (n-i)!} \prod_{j=1}^k \left[\frac{\binom{i}{m_j}}{\binom{n}{m_j}} \right]. \end{aligned}$$

Logo,

$$\begin{aligned} \lim_{k \rightarrow \infty} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = n\}) &= \\ &= 1 + n! \sum_{i=0}^{n-1} (-1)^{n-i} \frac{1}{i! (n-i)!} \lim_{k \rightarrow \infty} \prod_{j=1}^k \left[\frac{\binom{i}{m_j}}{\binom{n}{m_j}} \right] = 1, \end{aligned}$$

uma vez que, para $i=0,1,\dots,n-1$,

$$0 \leq \prod_{j=1}^k \left[\frac{\binom{i}{m_j}}{\binom{n}{m_j}} \right] \leq \left(\frac{i}{n} \right)^k \xrightarrow[k \rightarrow \infty]{} 0.$$

Para completar a prova temos que, para todo $k \geq 1$

$$\mathbb{P}\{T_k=N\} = \sum_{n \in A_\pi} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = n\}) \pi(n)$$

o que implica, pelo teorema da convergência dominada,

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathbb{P}\{T_k=N\} &= \sum_{n \in A_\pi} \lim_{k \rightarrow \infty} P_n(\{\omega \in \Omega : \sum_{i=1}^k \omega_i = n\}) \pi(n) = \\ &= \sum_{n \in A_\pi} \pi(n) = 1 \end{aligned}$$

e o lema está provado. □

O seguinte teorema estabelece um resultado mais forte que o estabelecido acima.

Teorema 29 - $T_k \xrightarrow[k \rightarrow \infty]{q.c.} N$, isto é, T_k converge quase certamente para N .

Prova- A sequência de variáveis aleatórias $\{T_k\}_{k \geq 1}$ é não decrescente e, pelo lema 28, para todo $k \geq 1$, $T_k \leq N$ q.c.($\mathbb{P}$). Logo, $T_k \xrightarrow[k \rightarrow \infty]{q.c.} L$ onde L é uma variável aleatória definida em Ω^* com valores em $\mathbb{N}$. Como, pelo lema 28, $T_k \xrightarrow[k \rightarrow \infty]{\mathbb{P}} N$ segue que $L=N$ q.c. ($\mathbb{P}$) e o teorema está provado. □

IV. 4 - Propriedades assintóticas do estimador de Bayes

Para todo $k \geq 2$, seja F_k a σ -álgebra de subconjuntos de Ω^* induzida por D_k , isto é,

$$F_k = \{D_k^{-1}(A) : A \subset \mathbb{N}^k\}.$$

A sequência $\{F_k\}_{k \geq 2}$ é uma sequência crescente de sub- σ -álgebras de F^* .

O estimador de Bayes $\bar{N} = E(N|D_k)$ é aqui representado na linguagem de σ -álgebras por $E(N|F_k)$. De acordo como o capítulo anterior, seção III.2,

$$\bar{N} = E(k, T_k) I_{A_{k, \pi}}$$

onde $I_{A_{k, \pi}}(\cdot)$ é a função indicadora do conjunto

$$A_{k, \pi} = \{(\omega, n) \in \Omega^* : M_j \leq \sum_{i=1}^j \omega_i \leq \min\{s_j, n\}, j=1, 2, \dots, k \text{ e } A_{\pi, T_k}(\omega, n) \neq \emptyset\}.$$

Lembremos que estamos supondo $\text{Var}(N) < \infty$.

O seguinte lema estabelece que a sequência $\{E(N|F_k)\}_{k \geq 2}$ é um martingal.

Lema 30 - $\{E(N|F_k)\}_{k \geq 2}$ é um martingal relativamente a $\{F_k\}_{k \geq 2}$.

Prova - Para todo $k \geq 2$, $F_k \subset F_{k+1}$ e $\{E(N|F_k)\}_{k \geq 2}$ é adaptada a $\{F_k\}_{k \geq 2}$, isto é, $E(N|F_k)$ é F_k -mensurável;

$$E(E(N|F_k)) = E(N) < \infty \quad e$$

$$E(N|F_k) = E(E(N|F_{k+1})|F_k) \quad \text{q.c.} (\mathbb{P})$$

e o lema está provado. □

Como $\lim_{k \rightarrow \infty} E(E(N|F_k)) = E(N) < \infty$, segue-se

que $\{E(N|F_k)\}_{k \geq 2}$ converge quase certamente para uma variável aleatória. (Doob (1953), pág. 319, Teorema 4.1).

O seguinte teorema estabelece que esta v.a. é igual a N quase certamente.

Teorema 31 - $E(N|F_k) \xrightarrow[k \rightarrow \infty]{\text{q.c.}} N$; isto é, $E(N|F_k)$ converge quase certamente para N .

Prova - Como $E(N) < \infty$, pelo Teorema 4.3, pág. 331 de Doob (1953), temos

$$\lim_{k \rightarrow \infty} E(N|F_k) = E(N|F_\infty) \quad \text{q.c.} (\mathbb{P})$$

onde F_∞ é a menor σ -álgebra de subconjuntos de Ω^* que contém $\bigcup_{k \geq 2} F_k$. Desde que, para todo $k \geq 2$, T_k é F_∞ -mensurável (T_k é F_k -mensurável e $F_k \subset F_\infty$) e, pelo Teorema 29, $T_k \xrightarrow[k \rightarrow \infty]{\text{q.c.}} N$ segue-se que $\limsup_k T_k$ é F_∞ -mensurável e $N = \limsup_k T_k$ q.c. ($\mathbb{P}$).

Logo, $E(N|F_\infty) = N$ q.c. ($\mathbb{P}$) e o teorema está provado. $\square$

A $\text{Var}(N|D_k)$ é aqui representada na linguagem de σ -álgebras por $\text{Var}(N|F_k)$.

O seguinte lema estabelece que a sequência $\{\text{Var}(N|F_k)\}_{k \geq 2}$ é um supermartingal.

Lema - 32 - $\{\text{Var}(N|F_k)\}_{k \geq 2}$ é um supermartingal relativamente a $\{F_k\}_{k \geq 2}$.

Prova - Como, para todo $k \geq 2$, $\{\text{Var}(N|F_k)\}_{k \geq 2}$ é adaptada a $\{F_k\}_{k \geq 2}$ e $E(\text{Var}(N|F_k)) \leq \text{Var}(N) < \infty$, resta provar que

$$\text{Var}(N|F_k) \geq E(\text{Var}(N|F_{k+1})|F_k) \text{ q.c. } (\mathbb{P}) \text{ (IV.4).}$$

Como

$$\begin{aligned} E(\text{Var}(N|F_{k+1})|F_k) &= \\ &= E[E(N^2|F_{k+1})|F_k] - E[(E(N|F_{k+1}))^2|F_k] = \\ &= E(N^2|F_k) - E[(E(N|F_{k+1}))^2|F_k] \end{aligned}$$

e, pela desigualdade de Jensen

$$E[(E(N|F_{k+1}))^2|F_k] \geq \{E[E(N|F_{k+1})|F_k]\}^2 = (E(N|F_k))^2 \text{ q.c. } (\mathbb{P})$$

segue-se que

$$E(\text{Var}(N|F_{k+1})|F_k) \leq E(N^2|F_k) - (E(N|F_k))^2 = \text{Var}(N|F_k) \text{ q.c. } (\mathbb{P})$$

o que prova (IV.4) e o lema está provado. $\square$

O seguinte teorema estabelece que a sequência $\{E(\text{Var}(N|F_k))\}_{k \geq 2}$ é não crescente.

Teorema 33 - Para todo $k \geq 2$,

$$E(\text{Var}(N|F_k)) \geq E(\text{Var}(N|F_{k+1})).$$

Prova - Pelo lema anterior, para todo $k \geq 2$

$$\text{Var}(N|F_k) \geq E(\text{Var}(N|F_{k+1})|F_k) \text{ q.c. } (\mathbb{P}).$$

Logo,

$$\begin{aligned} E(\text{Var}(N|F_k)) &\geq E[E(\text{Var}(N|F_{k+1})|F_k)] = \\ &= E(\text{Var}(N|F_{k+1})) \end{aligned}$$

e o teorema está provado. □

$$\text{Como } \sup_{k \geq 2} E(\text{Var}(N|F_k)) \leq \text{Var}(N) < \infty,$$

pelo Corolário do Teorema 9.4.4, pg. 334 de Chung (1974), o supermartingal $\{\text{Var}(N|F_k)\}_{k \geq 2}$ converge quase certamente para uma variável aleatória.

O seguinte teorema estabelece que esta variável aleatória é igual a zero quase certamente.

Teorema 34 - $\text{Var}(N|F_k) \xrightarrow[k \rightarrow \infty]{\text{q.c.}} 0$, isto é, $\text{Var}(N|F_k)$ converge quase certamente para 0.

Prova - Como

$$\text{Var}(N|F_k) = E(N^2|F_k) - (E(N|F_k))^2$$

e, pelo Teorema 31, $(E(N|F_k))^2 \xrightarrow[k \rightarrow \infty]{\text{q.c.}} N^2$ para provar o teorema basta provar que

$$E(N^2|F_k) \xrightarrow[k \rightarrow \infty]{\text{q.c.}} N^2 \quad (\text{IV.5}).$$

Como $E(N^2) < \infty$, pelo Teorema 4.3, pg. 331 de Doob (1953), temos

$$\lim_{k \rightarrow \infty} E(N^2|F_k) = E(N^2|F_\infty)$$

onde F_∞ é a menor σ -álgebra que contém $\bigcup_{k \geq 2} F_k$. Por outro

lado, desde que T_k^2 é F_∞ -mensurável e, pelo teorema 29,

$T_k^2 \xrightarrow[k \rightarrow \infty]{\text{q.c.}} N^2$ segue-se que $\limsup_k T_k^2$ é F_∞ -mensurável e

$N^2 = \limsup_k T_k^2$ q.c. ($\mathbb{P}$). Logo, $E(N^2|F_\infty) = N^2$ q.c. ($\mathbb{P}$) o que prova (IV.5) e o teorema está provado. □

Pelo teorema 33, a sequência

$\{\text{Var}(E(N|F_k))\}_{k \geq 2}$ é não decrescente e, para todo $k \geq 2$,

$\text{Var}(E(N|F_k)) \leq \text{Var}(N) < \infty$. Logo, existe o limite de

$\text{Var}(E(N|F_k))$. O seguinte teorema estabelece o valor deste limite.

Teorema 35 - $\lim_{k \rightarrow \infty} \text{Var}(E(N|F_k)) = \text{Var}(N)$

Prova - Como $\{\text{Var}(E(N|F_k))\}_{k \geq 2}$ é não decrescente e, para todo $k \geq 2$, $\text{Var}(E(N|F_k)) \leq \text{Var}(N)$, segue-se que

$$\lim_{k \rightarrow \infty} \text{Var}(E(N|F_k)) \leq \text{Var}(N) \quad (\text{IV.6}).$$

Por outro lado,

$$\text{Var}(E(N|F_k)) = E((E(N|F_k))^2) - (E(N))^2$$

e, pelo lema de Fatou,

$$\begin{aligned} \liminf_k E((E(N|F_k))^2) &\geq E(\liminf_k (E(N|F_k))^2) = \\ &= E(N^2). \end{aligned}$$

Logo,

$$\begin{aligned} \lim_{k \rightarrow \infty} \text{Var}(E(N|F_k)) &= \liminf_k \text{Var}(E(N|F_k)) = \\ &= \liminf_k E((E(N|F_k))^2) - (E(N))^2 \geq E(N^2) - (E(N))^2 = \\ &= \text{Var}(N), \end{aligned}$$

ou seja,

$$\lim_{k \rightarrow \infty} \text{Var}(E(N|F_k)) \geq \text{Var}(N) \quad (\text{IV.7}).$$

De (IV.6) e (IV.7) segue-se que $\lim_{k \rightarrow \infty} \text{Var}(E(N|F_k)) = \text{Var}(N)$

e o teorema está provado.

REFERÊNCIAS PARA O CAPÍTULO IV

BASU, D. & PEREIRA, C.A. de B. 1983. Conditional independence in statistics. *Sankhyā, Serie A*, 45(3): 324-37.

BREIMAN, L. 1968. *Probability*. Reading, Addison-Wesley. 421p.

CHUNG, K.L. 1974. *A course in probability theory*. 2.ed. New York, Academic Press. 365p.

DOOB, J.L. 1953. *Stochastic processes*. New York, John Wiley. 659p.

FLORENS, J.P.; MOUCHART, M.S.; ROLIN, J.M. 1983. *Elements of bayesian statistics*. Louvain-La-Neuve, Center for Operations Research and Econometrics. 1v. (varias paginações)

MOUCHART, M. 1976. A note on Bayes theorem. *Statistica* 36(2): 349-57.

NEVEU, J. 1965. *Mathematical foundations of the calculus of probability*. San Francisco, Holden-Day. 223p.

ZACKS, S. 1984. *Bayes sequential estimation of the size of a finite population*. São Paulo, IME-USP. (RT-MAE-8404)

REFERÊNCIAS COMPLEMENTARES

- CHAPMAN, D.G. 1954. The estimation of biological populations. *Annals of Mathematical Statistical*, 25:1-15.
- DARLING, D.A & ROBBINS, H. 1967. Finding the size of a finite population. *Annals of Mathematical Statistical*, 38:1392-8.
- DARROCH, J.N. 1959. The multiple recapture census II. Estimation when there is immigration or death. *Biometrika*, 46:336-51.
- JOLLY, G.M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, 52:225-47.
- JOLLY, G.M. 1982. Mark-recapture models with parameters constant in time. *Biometrics*, 38:301-321.
- POLLOCK, K.H.; HINES, J.E.; NICHOLS, J.D. 1985. Goodness-of-fit tests for open capture-recapture models. *Biometrics*, 41: 399-410.
- ROBSON, D.W. 1969. Mark-recapture methods of population estimation. IN: JOHNSON, N.L. & SMITH Jr., H., eds. *New developments in survey sampling*. New York, Wiley-Interscience. 732p.

SEBER, G.A.F. 1965. A note on the multiple-recapture census.
Biometrika, 52:249-59.

SEBER, G.A.F. 1982. *The estimation of animal abundance and related parameters*. 2.ed. London, Griffin.

SNELL, J.L. 1952. Applications of martingale system theorems. *Transactions of the American Mathematical Society*, 73:293-312.