*Probability in the Engineering and Informational Sciences*, 1, 1987, 225–236. Printed in the U.S.A.

# EXACT MAXIMUM LIKELIHOOD ESTIMATE OF A FINITE POPULATION SIZE

## *CAPTURE/RECAPTURE SEQUENTIAL SAMPLE DATA*

JOSE GALVÃO LEITE, JORGE OISHI, AND
CARLOS ALBERTO DE BRAGANÇA PEREIRA
*IME – Universidade de São Paulo*
*C.P. 20570 – 01498*
*São Paulo, Brazil*

Using data obtained by the general capture/recapture sequential sampling process, an exact analytical expression for the maximum likelihood (ML) estimate of the population size, $N$, is introduced. As a consequence, it is shown that bounded likelihood functions have at most two maxima. For the simple one-by-one case the ML estimate is unique.

## 1. INTRODUCTION

The objective of this paper is to introduce a closed analytical expression for the maximum likelihood (ML) estimate of the size, $N$, of a finite (and closed) population when the data are obtained by the capture/recapture sequential sampling process. Inferences about $N$ based on data obtained by special cases of this sampling process were considered by many authors (see for instance Craig [2], Goodman [7], Chapman [1], Lewontin and Prout [15], Darroch [4,5], Jolly [12], Seber [19], Darling and Robbins [3], Samuel [18], Freeman [8], Robson

This research was partially supported by $CNP_q$, Brasilia, Brazil.

Present address: University of California, Operations Research Center, 3115 Etcheverry Hall, Berkeley, California 94720.

 0269-9648/87 $5.00 + .00 

[17], and Pollock, Hines, and Nichols [16]). For a more accurate reference list, see Seber [21a]. The techniques used in the present paper, however, were suggested by a related problem described in Good [6, p. 73]. The sampling design for the capture/recapture sequential process and its sampling probability distribution are described next. For complete details see Leite & Pereira [14].

Consider a population of finite size, $N$ ($\in \mathbb{N}^* = \{0,1,\ldots\}$), such that during the study time it changes neither in size nor in form; that is, the population is closed during the study time. From this population, $k$ ($>1$) random samples (without replacement) are sequentially selected from the population. Each of these samples is returned back to the population before the next is selected. To obtain the relevant data to estimate $N$, the following steps are performed:

i. The first random sample of size $m_1$ ($\geq 1$) is drawn, without replacement. After the sample units are marked they are returned to the population and the number $m_1 = U_1$ is recorded.
ii. The $j$th ($j > 1$) random sample of size $m_j$ ($\geq 1$) is drawn, without replacement. The sample units marked in earlier selected samples are immediately returned to the population. The remaining $U_j$ unmarked sample units are returned after being marked. The numbers $m_j$ and $U_j$ are recorded.

After the $k$ samples have been obtained, the data $D_k = (U_1, \ldots, U_k)$ are observed. Note that the statistic $T_k = (U_1 + \cdots + U_k)$ is the number of distinct population units selected in the whole sampling process. Leite and Pereira [14] show that this statistic is sufficient and that the smallest factor of the likelihood function that depends on the value of $N$, the *Likelihood Kernel*, which is a minimal sufficient statistic (Zacks [22]), is given by:

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

where $t$ is the observed value of $T_k$ and $I_t(\cdot)$ is the indicator function of $\mathbb{N}_t^* = \{n \in \mathbb{N}^*;\ n \geq t\}$. The probability distribution of $T_k$ (Johnson and Kotz [10] and Leite and Pereira [14]) is given by:

$$P\{T_k = t \mid N\} = K(N,t)I_{\mathbb{C}}(t)\sum_{i=0}^{t}\left\{(-1)^{t-i}[i!\,(t-i)!]^{-1}\prod_{j=1}^{k}\binom{i}{m_j}\right\},$$

where $I_{\mathbb{C}}(t)$ is the indicator function of the set

$$\mathbb{C} = \left\{x \in \mathbb{N}^*;\ \max\{m_1,\ldots,m_k\} \leq x \leq \min\left\{N, \sum_{j=1}^{k} m_j\right\}\right\}$$

evaluated at point $t$.

In the following sections, a detailed study of $K(N,t)$, as a function of $N$, is presented in order to obtain the ML estimates of $N$. To illustrate, numerical results are presented.

## 2. MAIN RESULTS

For convenience of notation, although $N$ is a finite nonnegative integer, we consider that $\infty$ is also a point of $\mathbb{N}^*$; that is $\mathbb{N}^* = \{0,1,\ldots,\infty\}$. For an observed point $t$ of $T_k$, an ML estimate of $N$ is a point $\bar{N} \in \mathbb{N}^*$ that maximizes the function $K(\cdot,t)$. The following simple result introduces the ML estimate for the two extreme cases.

*Proposition 1:* If $t = \max\{m_1,\ldots,m_k\}$, the minimum possible value of $t$, then $\bar{N} = t$. If $t = m_1 + \cdots + m_k$, the maximum possible value of $t$, then $\bar{N} = \infty$.

PROOF: (i) Without any loss of generality, let $t = m_1 = \max\{m_1,\ldots,m_k\}$. Then,

$$K(N,m_1) = I_{m_1}(N)\left\{\prod_{j=2}^{k}\binom{N}{m_j}\right\}^{-1} m_1!$$

has its maximum at point $\bar{N} = m_1$, the minimum possible value of $N$. (ii) Let $t = s = m_1 + \cdots + m_k$. Hence,

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

Note also that for any fixed $k \geq 2$ and for any real numbers $x_j$, $0 < x_j < 1$ and $j \in \{i,\ldots,k\}$,

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

Consequently, for all $N \geq s$,

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

concluding the proof. ■

Let $m = \max\{m_1,\ldots,m_k\}$ and, as above, $s = m_1 + \cdots + m_k$. We consider now the case where $m < t < s$. For all $N \in \mathbb{N}_t^* = \{n \in \mathbb{N}^*; n \geq t\}$, define the ratio

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

and consider the function

$$f_t(x) = (1-xt)^{-1}\prod_{j=1}^{k}(1-xm_j),$$

defined in $0 \leq x < 1/t$. This function is continuous in $[0,1/t)$, $f_t(0) = 1$, $f_t(x) \to \infty$ as $x \uparrow 1/t$, and if $N \in \mathbb{N}_t^*$,

$$f_t\left(\frac{1}{N+1}\right) = \frac{K(N+1,t)}{K(N,t)}. \tag{1}$$

The behavior of $f_t$ is described by the following result:

LEMMA 2: *For $m < t < s$, the equation $f_t(x) = 1$ has a unique positive solution $x_0$ in the open interval $(0,1/t)$. Also, $f_t(x) < 1$ if $0 < x < x_0$ and $f_t(x) > 1$ if $x_0 < x < 1/t$.*

PROOF: Let $f_t = g/h_t$ where $g$ and $h_t$ are two functions defined as:

$$g(x) = \prod_{j=1}^{k} (1 - xm_j) \text{ and } h_t(x) = 1 - xt.$$

The first and second derivatives of $g$ are, respectively,

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

and

$$g''(x) = \sum_{i=1}^{k} \left[ \sum_{\substack{j=1 \\ j \neq i}}^{k} (m_i m_j) \prod_{\substack{l=1 \\ l \neq i,j}}^{k} (1 - xm_l) \right] > 0.$$

Hence, $g(\cdot)$ is a decreasing convex continuous function. Also, $h_t(x)$ is a decreasing linear function, $g(0) = h_t(0) = 1$, $g(1/t) > 0$, and $h_t(1/t) = 0$. (Figure 1 illustrates these two functions for a particular case.) The derivative of $(h_t - g)$ evaluated at the origin is $h_t'(0) - g'(0) = \left[-t + \sum_{j=1}^{k} m_j\right] > 0$. Consequently, there is a positive real number, $\delta$, such that $h_t'(x) - g'(x) > 0$ for all $x$ in the interval $[0,\delta)$ and, from the mean value theorem, $h_t(x) > g(x)$ for all $x$ in the open interval $(0,\delta)$. The conclusion is that there is a unique point, $x_0$, belonging to the open interval $(0,1/t)$, such that

$$g(x_0) = h_t(x_0),$$

$$g(x) < h_t(x) \text{ if } x \in (0,x_0), \text{ and}$$

$$g(x) > h_t(x) \text{ if } x \in (x_0,1/t).$$

Since $f_t$ is the restriction of $g/h_t$ to the interval $[0,1/t)$, the proof is concluded. ■

The main result of the present paper is stated next. It introduces the general form of the ML estimate and highlights the region of unicity of such an estimate.

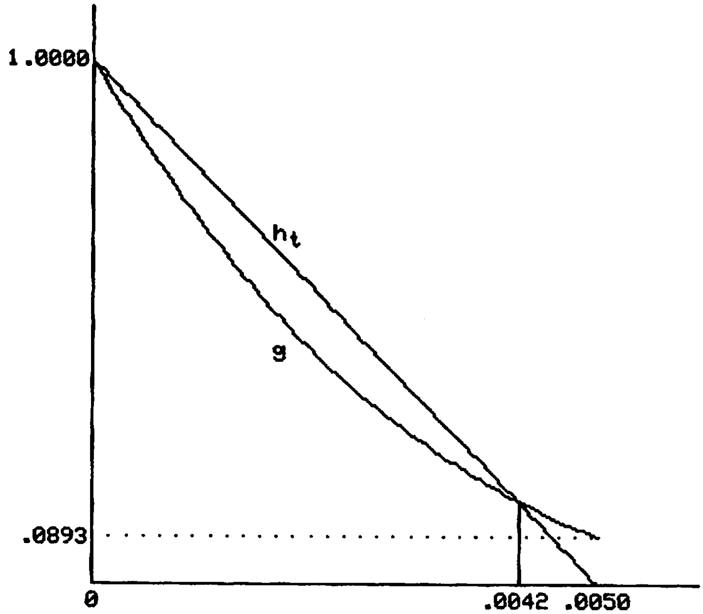


**FIGURE 1.** Functions $g$ and $h_t$ for the case $M = (150, 40, 50, 60, 30)$ and $t = 200$.

Let $M = (m_1, m_2, \ldots, m_k)$ be a vector composed of $k$ (>1) sample sizes defined by the capture/recapture sequential sampling design, $t$ be the observed value of $T_k$,

$$m = \max\{m_1, \ldots, m_k\}, \qquad s = m_1 + \cdots + m_k, \text{ and}$$

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

**THEOREM 3:** *An ML estimate of $N$, $\overline{N}$, exists and is defined as*

$$\overline{N} = \begin{cases} t & \text{if } t = m \\ t + n_t - 1 & \text{if } m < t < s \\ \infty & \text{if } t = s. \end{cases}$$

*Also, this estimate is unique except when*

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

*in which case the only two possible estimates are $(t + n_t - 1)$ and $(t + n_t - 2)$.*

PROOF: (i) From Proposition 1, $\{t = m\} \to \{\bar{N} = t\}$ and $\{t = s\} \to \{\bar{N} = \infty\}$.
(ii) Let $m < t < s$. From Lemma 2, we have that $\exists n_0 \in \mathbb{N}^*$, $n_0 > 1$, such that, for $n \in \mathbb{N}^*$,

$$f_t\left(\frac{1}{t+n}\right) \begin{cases} <1 & \text{if } n \geq n_0 \\ \geq 1 & \text{if } n = n_0 - 1 \\ >1 & \text{if } n < n_0 - 1. \end{cases}$$

Hence, $n_0 = \min\left\{n \in \mathbb{N}^*; f_t\left(\frac{1}{t+n}\right) < 1\right\}$. That is, $n_0 = n_t$ where $n_t$ is defined above. Note that $f_t\left(\frac{1}{t+n}\right) < 1$ is equivalent to $\prod_{j=1}^{k} (t + n - m_j) < n(t + n)^{k-1}$. The following facts conclude the proof:
(a) If

$$\prod_{j=1}^{k} (t + n_t - 1 - m_j) > (n_t - 1)(t + n_t - 1)^{k-1},$$

then

$$f_t\left(\frac{1}{t+n}\right) \begin{cases} < 1 & \text{if } n \geq n_t \\ > 1 & \text{if } n \leq n_t - 1. \end{cases}$$

Consequently, from Eq. (1), we conclude that

$$K(t + n_t - 1, t) > K(t + n_t - 2, t) > \cdots > K(t, t) \text{ and}$$
$$K(t + n_t - 1, t) > K(t + n_t, t) > K(t + n_t + 1, t) > \cdots.$$

That is, $\bar{N} = (t + n_t - 1)$ is the only ML estimate of $N$.
(b) If

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

then

$$f_t\left(\frac{1}{t+n}\right) \begin{cases} <1 & \text{if } n \geq n_t \\ =1 & \text{if } n = n_t - 1 \\ >1 & \text{if } n \leq n_t - 2. \end{cases}$$

From Eq. (1) we conclude that

$$K(t + n_t - 1, t) = K(t + n_t - 2, t) > \cdots > K(t, t) \text{ and}$$
$$K(t + n_t - 1, t) > K(t + n_t, t) > K(t + n_t + 1, t) > \cdots.$$

That is, $\bar{N} = (t + n_t - 1)$ and $\bar{N}' = t + n_t - 2$ are the only two ML estimates of $N$. ∎

Figure 2 presents the function $f_t$ for the case $M = (m_1, \ldots, m_5) = (150, 40,50,60,30)$ *and* $t = 200$. For this, the unique ML estimate is $\bar{N} = 236$. Table 1 presents some numerical examples in order to give an idea of the ML estimate behavior.

## 3. SPECIAL CASES

In this section we consider special cases of sample designs. We study in detail the one-by-one case; that is, $m_1 = \cdots m_k = 1$.

As direct consequences of Theorem 3, we present the following results which are ways of checking the unicity of the ML estimate.

COROLLARY 4: *In Theorem 3 let* $k = 2$ *and* $\bar{N}_2 = m_1 m_2 (m_1 + m_2 - t)^{-1}$. *If* $\bar{N}_2 \in \mathbb{N}^*$, *then the only two ML estimates of* $N$ *are* $\bar{N}_2$ *and* $\bar{N}_2 - 1$. *If* $\bar{N}_2 \notin \mathbb{N}^*$, *then* $[\bar{N}_2]$, *the larger integer not superior to* $\bar{N}_2$, *is the unique ML estimate of* $N$.

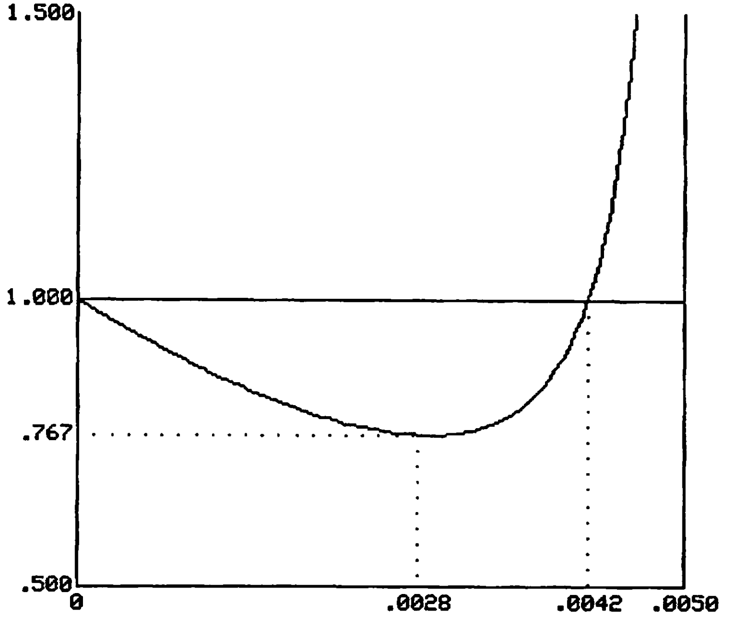


FIGURE 2. Function $f_t$ for the case $M = (150, 40, 50, 60, 30)$ and $t = 200$.

TABLE 1. Examples of ML Estimates

| $M$ | $t$ | ML Estimates |
|---|---|---|
| (40, 60) | 62 | 63 |
| | 80 | 119 & 120 |
| (1, 5, 8) | 10 | 12 |
| | 11 | 16 |
| | 12 | 25 |
| (40, 60, 80) | 90 | 92 |
| | 120 | 152 |
| | 140 | 239 & 240 |
| | 179 | 10381 |
| (3, 3, 4, 4, 5) | 6 | 6 |
| | 7 | 7 |
| | 10 | 11 |
| | 17 | 67 |
| | 18 | 139 |
| (15, 20, 25, 30, 50) | 60 | 61 |
| | 80 | 95 |
| | 98 | 149 & 150 |
| | 120 | 347 |
| | 139 | 7449 |

COROLLARY 5: *In Theorem 3 let* $k = 3$, $a = m_1 m_2 m_3$, $b = (m_1 m_2 + m_1 m_3 + m_2 m_3)$, $c = (m_1 + m_2 + m_3 - t)$, *and* $\bar{N}_3 = 2a[b - (b^2 - 4ac)^{1/2}]^{-1}$. *If* $\bar{N}_3 \in \mathbb{N}^*$, *then the only two ML estimates of N are* $\bar{N}_3$ *and* $\bar{N}_3 - 1$. *If* $\bar{N}_3 \notin \mathbb{N}^*$, *then* $[\bar{N}_3]$ *is the unique ML estimate of N.*

To obtain similar results for $k > 3$ one needs to handle complicated algebraic equations of degree $k - 1$. However, in the case of equal sample sizes, $m_1 = \cdots = m_k$, we have the following result.

COROLLARY 6: *Let* $\{n_1, n_2\} \subset \mathbb{N}^*$ *with* $n_2 \neq 0$. *If*

$$m_1 = \cdots = m_k = 2^{k+n_1} \text{ and } t = \{1 - [1 - 2^{-n_2}]^k\}2^{k+n_1+n_2},$$

*then the only two ML estimates are* $\bar{N} = 2^{k+n_1+n_2}$ *and* $\bar{N} - 1$.

From Theorem 3, the proofs of these results are straightforward. Table 2 presents some numerical examples for the case of equal sample sizes. These examples show that the conditions of Corollary 6 are not necessary.

The most important special case of sampling design is the one-by-one sequential sampling. That is, $m_1 = m_2 = \cdots = m_k = 1$. We end this section

**Table 2.** Examples of ML Estimates with Equal Sample Sizes

| $k$ | $m_1$ | $t$ | ML Estimates |
|---|---|---|---|
| 2 | 2 | 3 | 3 & 4 |
| 2 | 3 | 5 | 6 |
| 2 | 4 | 6 | 7 & 8 |
| 3 | 4 | 7 | 7 & 8 |
| 3 | 4 | 9 | 14 |
| 4 | 6 | 11 | 11 |
| 4 | 16 | 30 | 31 & 32 |
| 4 | 64 | 175 | 255 & 256 |
| 5 | 32 | 62 | 63 & 64 |
| 6 | 2 | 6 | 6 |
| 10 | 15 | 76 | 91 |

showing the unicity of the ML estimate for this sampling design. This is not in agreement with some other authors (Samuel [18]) who believe that there may exist more than one ML estimate in this case.

From Theorem 3 we know that, in the one-by-one case, to obtain two ML estimates when $1 < t < k$ we should have

$$(t + n_t - 2)^k = (n_t - 1)(t + n_t - 1)^{k-1}.$$

Let $x$ be the integer $t + n_t - 1$. The above equation is equivalent to

$$(x-1)^k = (x-t)x^{k-1} \text{ or } (t-k)x^{k-1} + \sum_{i=2}^{k-1} \binom{k}{i} (-1)^i x^{k-i} = (-1)^k.$$

Note also that this last expression is equivalent to

$$(t-k)x^{k-2} + \sum_{i=2}^{k-1} \binom{k}{i} (-1)^i x^{k-i-1} = \frac{(-1)^{k-1}}{x}.$$

Finally, since $x$ is an integer, the left-hand side of this equation must also be an integer. This is absurd because the right-hand side cannot be an integer number since $x > t > 1$. The conclusion is that the above equation does not have an integer solution and the following result holds.

Theorem 7: *For the one-by-one sequential sampling, there exists a unique ML estimate defined by*

$$\bar{N} = \begin{cases} t & \text{if } t = 1 \\ t + n_t - 1 & \text{if } 1 < t < k \\ \infty & \text{if } t = k, \end{cases}$$

*where*

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

To illustrate the behavior of the ML estimates Table 3 introduces some numerical examples.

## 4. COMMENTS AND CONCLUSION

Inferences about population size have been the object of some recent papers, see for instance Isaki [9]. However, they are usually restricted to the case of $k = 2$. We believe that, with the simple expression of the ML estimator obtained from $\bar{N}$ (with $T_k$ in the place of $t$), some of this work can be extended to the general case of $k \geq 2$.

Before using $\bar{N}$ one needs to observe the following list of limitations:

i. The parameter space $\mathbb{N}_t^*$ changes with the observed value $t$ of $T_k$.

ii. The random variables $U_i$, $i = 1, \ldots, k$, that form the data, $D_k$, are not independent and identically distributed. In fact, they are not even exchangeable.

iii. The ML estimator defined from $\bar{N}$ has no finite moments.

iv. When either event $\{T_k = m\}$ or $\{T_k = s\}$ occurs the value of $\bar{N}$ is not related to the value of $k$. For instance, let $m_1 = \cdots = m_k = 1$. If $\{T_k = 1\}$ (or $\{T_k = k\}$) obtains, then $\{\bar{N} = 1\}$ (or $\{\bar{N} = \infty\}$) whether $k = 2$ or $k = 2^{80}$.

**Table 3.** Examples of ML Estimates for the One-by-One Sequential Sampling

| $t$ | $K$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 |

Facts (i), (ii), and (iii) restrict the use of standard statistical procedures. The use of Bayesian procedures may be the way to contour these problems since they rely on the actual observed data rather than on the distributional properties of $D_k$ or $T_k$. In a subsequent note the authors intend to present a Bayesian discussion for the present problem. To the best of our knowledge, Freeman [8], Zacks [23], and Leite [13] are the available Bayesian references for the estimation problem of $N$.

From a practical point of view, $\bar{N} = \infty$ is to be understood as a very large number. Note that usually we are interested in the size of a population located in a limited region which may accommodate a large but finite number of units. Hence, with some little prior knowledge, one may easily figure out (with some desired exaggeration) the maximum possible value of $N$. This number is technically represented here by $\infty$. Thus, in fact, we can say that the moments of our estimator are very large but finite numbers. Restriction (iii) may then be eliminated. However, by posing the problem in this manner, many of the well-known asymptotical simplifier properties may not be easily adjusted to our problem. Restrictions (i) and (ii) indicate that, to have all kinds of inferences about $N$, one needs to perform a careful study of the distributional properties of the sequence $\{T_k\}_{k>1}$.

The strongest restriction on $\bar{N}$, in our opinion, is introduced by (iv). In order to make the length $k$ of the data vector be relevant when $T_k = m$ or $T_k = s$, one needs to use prior knowledge.

### References


1. Chapman, D. C. (1954). The estimation of biological populations. *The Annals of Mathematical Statistics* 25: 1–15.
2. Craig, C. C. (1953). On utilization of marked specimens in estimating the population of flying insects. *Biometrika* 4: 170–176.
3. Darling, D. A. and Robbins, H. (1967). Finding the size of a finite population. *The Annals of Mathematical Statistics* 38: 1392–1398.
4. Darroch, J. N. (1958). The multiple-recapture census I. Estimation of a closed population. *Biometrika* 45: 343–359.
5. Darroch, J. N. (1959). The multiple-recapture census II. Estimation when there is immigration or death. *Biometrika* 46: 336–351.
6. Good, I. J. (1950). *Probability and the Weighing of Evidence*, Charles Griffin, London.
7. Goodman, L. A. (1953). Sequential sampling tagging for population size problems. *The Annals of Mathematical Statistics* 24: 56–69.
8. Freeman, P. R. (1972). Sequential estimation of the size of a population. *Biometrika* 59: 9–17.
9. Isaki, C. T. (1986). Bias of the dual system estimator and some alternatives. *Communications in Statistics: Theory and Methods* 15(5): 1435–1450.
10. Johnson, N. L. and Kotz, S. (1977). *Urn Models and Their Applications: An Approach to Modern Discrete Probability Theory*, John Wiley, New York.
11. Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika* 52: 225–247.
12. Jolly, G. M. (1982). Mark-recapture models with parameters constant in time. *Biometrika* 38: 301–321.

13. Leite, J. G. (1986). Exact estimates of the size of a finite and closed population (in Portuguese). Doctoral Dissertation, São Paulo, Brazil, Universidade de São Paulo, 93p.
14. Leite, J. G. and Pereira, C. A. de B. (1986). An urn model for the capture/recapture sequential sampling process. *Sequential Analysis* 6, to appear.
15. Lewontin, R. C. and Prout, T. (1956). Estimation of the number of different classes in a population. *Biometrika* 12: 211-223.
16. Pollock, K. H., Hines, J. E., and Nichols, J. D. (1985). Goodness-of-fit tests for open capture-recapture models. *Biometrika* 41: 399-410.
17. Robson, D. W. (1969). Mark-recapture methods of population estimation. In: Johnson and Smith, eds., *New Developments in Survey Sampling*, John Wiley, New York.
18. Samuel, E. (1968). Sequential maximum likelihood estimation of the size of a population. *The Annals of Mathematical Statistics* 39: 1057-1068.
19. Seber, G. A. F. (1962). The multi-sample single recapture census. *Biometrika* 49: 339-350.
20. Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika* 52: 249-259.
21. Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters*, 2nd ed., Griffin, London.
21a. Seber, G. A. F. (1968). A review of estimating animal abundance. *Biometrics* 42:267-292.
22. Zacks, S. (1981). *Parametric Statistical Inference: Basic Theory and Modern Approaches.* Pergamon Press, Oxford.
23. Zacks, S. (1984). *Bayes Sequential Estimation of the Size of a Finite Population*, São Paulo, Brazil. IME-USP Technical Report 8404, 12p.