SEQUENTIAL ESTIMATION OF THE SIZE OF A FINITE POPULATION

by

P.K. Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1804

September 1986

SEQUENTIAL ESTIMATION OF THE SIZE OF A FINITE POPULATION

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill, N.C., USA

SUMMARY

For the estimation of the total number of individuals in a finite population, the capture, mark, release and recapture method and its variants for inverse as well as sequential sampling schemes have been extensively worked out in the literature. These are systematically reviewed in the context of sequential (point as well as interval) estimation of the size of a population. With special emphasis placed on appropriate martingale constructions for suitable sequences of statistics arising in this context, invariance principles for sequential tagging are considered, and their role in the proposed sequential analysis is critically discussed.

1. INTRODUCTION

In estimation of animal abundance (or extinction of an endangered species) and in various other biological, ecological or environmental investigations, a complete census may not be operationally feasible (or practicable), and based on suitable sampling schemes, the estimation of the total size of a population (say, N) is of considerable importance. In actual practice, because of migration and natural birth and death effects, N may not remain stationary (over any period of time), so that an estimation procedure incorporating these provisions may become more involved. However, if the investigation is carried out in a relatively

---

AMS Subject classifications: 62L12, 62L15

Key Words and Phrases: Asymptotic distributional risk; CMRR; coefficient of variation; fixed percentage width confidence interval; minimum risk estimation; maximum likelihood estimation; sequential tagging; stopping rule.

short interval of time, this nonstationarity can be ignored, and under a plausible assumption of a closed population, objective sampling methods can be effectively used to estimate N in a covenient way.

The capture, mark, release and recapture (CMRR) technique has been quite extensively used for the estimation of the size N of a closed population. The Petersen (1896) two sample CMRR estimator and its various modification (and extensions) are indeed very useful in a broad spectrum of practical applications. Multi-sample, inverse as well as sequential sampling schemes have also been considered by a host of researchers; we shall provide a brief account of these traditional developments in Section 2.

Typically, an estimator of N based on a sample of size n, denoted by $\hat{N}_n$, has sampling fluctuation decreasing with n, but large, for large values of N. In many problems, N (through unknown) is a reasonably large number, so that, to reduce the margin of error of $\hat{N}_n$, at least intuitively, one would suggest a large value of the sample size n. However, the CMRR technique involves an operational cost (for drawing sample observations as well as marking and releasing them), so that in choosing a suitable sampling scheme, one needs to take into account this cost function in conjunction with some conventional risk function (depicting the expected level of error in the estimator). While this cost function may not depend (sensibly) on the size of the population (but on the sample size and the sampling design), the risk function is generally an involved function of N and n. Since N is not known in advance, an optimal solution (for n) may not be known a priori, and one may therefore take recourse to sequential methods to achieve this optimality, in (at least) some asymptotic setup. Our main interest lies in the formulation of such sequential procedures for the estimation of the population size N.

Two basic sequential estimation problems will be considered here. First, the sequential interval estimation problem. Typically, for a parameter θ (in a parametric or nonparametric setup), one needs to find out an interval I_n (based on a sample of size n), such that the probability that I_n covers θ is equal to some prespecified $1 - \alpha$ $(0<\alpha<1)$ and the width of the interval I_n is also bounded from above by some prefixed number $(2d\ ,\ d>0)$. Generally, no fixed-sample size solution exists for this problem, and the sample size is determined sequentially by using some well-defined stopping rule. In the context of estimation of N, this formulation may not be totally meaningful. Since N is a positive integer, we need to restrict ourselves to integer values of d. For any fixed $d(\geq 1)$, as N becomes large, a sequential solution may demand an indefinitely large value of the (average) sample number, which may run contrary to the practical considerations. In most of the cases, it may be quite reasonable to choose $d(=d_N)$, such that the width of the interval is bounded from above by $2N\varepsilon$, for some prefixed $\varepsilon(0<\varepsilon<1)$, where ε is usually taken small. This may be termed a fixed percentage width (sequential) confidence interval. We shall mainly study such intervals. Secondly, we shall consider the problem of mimimum risk point estimation of N. Here also, the risk function depends on the cost function as well as another function (depicting the expected level of error) involving the unknown size N, and hence, a fixed-sample size solution generally does not exist. Under an appropriate asymptotic setup, we shall consider a sequential procedure and study its properties.

Following the treatment of the classical estimators of N in Section 2, we shall discuss some general martingale characterizations relating to estimates and estimating functions in Section 3. Section 4 deals with the confidence interval problem, while the point estimation problem is treated in the concluding section.

2. ESTIMATION OF N BASED ON THE CMRR TECHNIQUE

Consider first the simple two-sample model. A sample of n_1 units are drawn from the population of N (unknown) units, these are marked conveniently and released subsequently, so that they mingle freely with the rest. A second random sample of n_2 observations is then drawn from the same population, and let there be r (random) units which are observed to be marked before (so that $0 \le r \le n_2$). Then, given N, n_1 and n_2, the probability function of r is given by

$$p(r|N,n_1,n_2) = \binom{n_1}{r}\binom{N-n_1}{n_2-r}\Big/\binom{N}{n_2}, \qquad 0 \le r \le n_2 \wedge n_1. \tag{2.1}$$

Therefore,

$$p(r|N,n_1,n_2)/p(r|N-1,n_1,n_2) = (N-n_1)(N-n_2)/N(N-n_1-n_2+r) \tag{2.2}$$

Thus, $\hat{N}_P$, the <u>maximum likelihood estimator</u> (MLE) of N, satisfies the inequalities $n_1n_2/r - 1 \le \hat{N}_P \le n_1n_2/r$, and noting that N is a positive integer, we may take

$$\hat{N}_P = [n_1n_2/r] = \text{largest integer} \le n_1n_2/r. \tag{2.3}$$

This is the classical Petersen (1896) estimator of N. We may remark that for every finite N, n_1 and n_2, $p(0|N,n_1,n_2) = \binom{N-n_1}{n_2}\Big/\binom{N}{n_2} > 0$, so that $\hat{N}_P = +\infty$ with a positive probability, and further, $\hat{N}_P$ does not have finite moment of any positive order. A modification of $\hat{N}_P$, to eliminate this problem, due to Chapman (1951), is the following:

$$\tilde{N}_{\epsilon} = [(n_1+1)(n_2+1)/(r+1) - 1]. \tag{2.4}$$

Also, another suggestion is to use inverse sampling on the second occasion, so that units may be drawn one by one until a prespecified member (say, $m(\geq 1)$) of marked units are observed. In this setup, n_2, the number of units required to be drawn on the second occasion to yield exactly m marked units is itself a positive integer valued random variable and we have

$$p(n_2|N,n_1,m) = \binom{N}{n_2-1}^{-1}\binom{n_1}{m-1}\binom{N-n_1}{n_2-m}(n_1-m+1)/(N-n_2+1), \tag{2.5}$$

for $n_2 \geq m$. Again,

$$\frac{p(n_2|N,n_1,m)}{p(n_2|N-1,n_1,m)} = \frac{(N-n_2)(N-n_1)}{N(N-n_1-n_2+m)}, \tag{2.6}$$

so that the MLE of N, as in (2.3) is given by

$$\hat{\hat{N}}_C = [n_1 n_2/m] \tag{2.7}$$

Note, however, that m (≥ 1) is prefixed, while n_2 is a random variable, and using (2.5), it can be shown that $\hat{\hat{N}}_C < \infty$ with probability one, and $\hat{\hat{N}}_C$ has a finite moment of any (finite) positive order.

The Petersen (1896) estimator has also been extended to the multi-sample case by a host of workers [led by Schnabel (1938)]. In this case, each sample captured (commencing from the second) is examined for marked members and then every member of the sample is given another mark before being released to the population. Let $s(\geq 2)$ be the number of samples (of sizes $n_1, \ldots, n_s$, respectively), and let m_i be the number of marked units found in the ith sample, $u_i=n_i-m_i$, for $i=2, \ldots, s$. Also, let $u_1=n_1$, $m_1=0$ and let $M_i=\Sigma_{j\le i-1}\, u_j$, for $i=1, \ldots, s+1$ $(M_1=0)$. Then, we obtain that

$$p(m_2,\ldots,m_s|n_1,\ldots,n_s,N) = \prod_{i=2}^{s} \binom{M_i}{m_i}\binom{N-M_i}{n_i-m_i}\binom{N}{n_i}^{-1}, \qquad (2.8)$$

so that

$$\frac{p(m_2,\ldots,m_s|n_1,\ldots,n_s,N)}{p(m_2,\ldots,m_s|n_1,\ldots,n_s,N-1)} = \prod_{i=2}^{s} \frac{(N-M_i)(N-n_i)}{N(N-n_i-M_i+m_i)}$$

$$= \prod_{i=2}^{s} \frac{(N-M_i)(N-n_i)}{N(N-M_{i+1})} = N^{-s+1}(N-M_{s+1})^{-1} \prod_{i=2}^{s} (N-n_i)(N-M_2). \qquad (2.9)$$

Thus, for the MLE $\hat{N}_s$, we have

$$(1-\hat{N}_s^{-1}M_{s+1}) \le (1-\hat{N}_s^{-1}n_1)\ldots(1-\hat{N}_s^{-1}n_s), \qquad (2.10)$$

while the opposite inequality holds when $\hat{N}_s$ is replaced by $\hat{N}_s+1$. For $s>2$, this non-linear equation generally calls for an iterative solution. Alternatively, at the ith stage $(i=2, \ldots, s)$, one may consider the usual two sample Petersen estimator n_iM_i/m_i, and combine these $s-1$ estimators by an weighted average. Generally, such Schumacher and Eschmeyer type alternative estimators entails some loss of efficiency, and detailed study of some related results is due to Sen and Sen (1981).

Sequential sampling tagging for the estimation of N has been considered by Chapman, (1952), Goodman (1953), Darroch (1958) and Sen (1982a,b), among others. Individuals are drawn randomly one by one marked and released before the next drawing is made. Let M_k be the number of marked individuals in the population just before the kth drawal, for $k\ge 1$. We may write

$$M_{k+1} = M_k + (1-X_k),\ k\geq 1 \qquad (2.11)$$

where

$$X_k = \begin{cases} 1, & \text{if the kth drawal yields a marked individuals} \\ 0, & \text{otherwise.} \end{cases} \qquad (2.12)$$

Thus, the likelihood function at the nth stage ($n\geq 2$), is given by

$$L_n(N) = \prod_{k=2}^{n} \{M_k^{X_k} (N-M_k)^{1-X_k} N^{-1}\}, \qquad (2.13)$$

so that, we have

$$L_n(N)/L_n(N-1) = (1-N^{-1})^{n-1} / \prod_{k=2}^{n} (1-(N-M_k)^{-1})^{1-X_k}. \qquad (2.14)$$

and the MLE $\hat{N}_n$ satisfies the condition that

$$\prod_{k=2}^{n} \{1-(\hat{N}_n-M_k)^{-1}\}^{1-X_k} \leq (1-\hat{N}_n^{-1})^{n-1}. \qquad (2.15)$$

While the above treatment holds for non-stochastic n, in the context of the two sequential estimation problems (referred to in Section 1), we would have generally a stochastic n (assuming positive integer values), for which study of the properties of the MLE would require knowledge on the nature of the stopping rules as well as some deeper asymptotic results on the sequence $\{L_n(N), n\geq 1\}$. Some of these were studied in Sen (1982 a,b) and we shall find it convenient to extend these results further in the current investigation. Towards this study, we consider first (in Section 3) some asymptotic results on the CMRR estimators, which will be needed in the sequel.

3. CMRR ESTIMATES: ASYMPTOTIC THEORY

We consider first the Petersen estimator. Assume that N is large and for some positive α, β,

$$n_1/N \to \alpha \text{ and } n_2/N \to \beta \; : \; 0<\alpha,\ \beta \le 1. \tag{3.1}$$

Then, using the normal approximation to the hypergeometric law along with the Slutzky theorem, we obtain from (2.3) that as N increases

$$N^{-\frac{1}{2}} (\hat{N}_p - N) \sim \mathcal{N}(0,\ (1-\alpha)(1-\beta)/\alpha\beta). \tag{3.2}$$

The same asymptotic distribution pertains to the estimator $\hat{\hat{N}}$ in (2.7) when we let $M \sim N\alpha\beta$ (so that the expected value of n_2 will be close to $N\beta$ and comparable to (3.1)). For the multi-sample CMRR estimator in (2.9)-(2.10), we allow N to be large, subject to

$$n_i/N \to \alpha_i \ (0<\alpha_i<1), \text{ for } i = 1, \ldots, s. \tag{3.3}$$

Also, we let

$$\beta_i = (1-\alpha_1) \ldots (1-\alpha_i), \text{ for } i = 1, \ldots, s. \tag{3.4}$$

Then, proceeding as in Sen and Sen (1981), we obtain that for large values of N, under (3.3)

$$N^{-\frac{1}{2}} (\hat{N}_s - N) \sim \mathcal{N}(0, \sigma_s^2), \tag{3.5}$$

where

$$\sigma_s^2 = \{\Sigma_{i=2}^{s} \alpha_i \beta_i^{-1} (1-\beta_{i-1})\}^{-1}. \tag{3.6}$$

Note that $\sigma_2^2 = (\alpha_2 \beta_2^{-1} (1-\beta_1))^{-1} = (1-\alpha_1)(1-\alpha_2)/\alpha_1\alpha_2$ and this agrees with (3.2) if we let $\beta=\alpha_2$. To compare σ_s^2 with σ_2^2 (for $s>2$), first, we consider the case of $s=3$. Let us write then

$$\alpha_1=\alpha,\ \alpha_2+\alpha_3=\beta,\ n_2+n_3=n_2'. \tag{3.7}$$

Thus, we need to compare (3.2) with (3.6), under (3.7)(which relates to the equality of the total sample sizes n_1+n_2' and $n_1+n_2+n_3$ for the two and three sample models). It is then easy to verify that

$$\alpha_2\beta_2^{-1}(1-\beta_1)+\alpha_3\beta_3^{-1}(1-\beta_2) \geq \alpha_1(\alpha_2+\alpha_3)/(1-\alpha_1)(1-\alpha_2-\alpha_3), \quad (3.8)$$

for all $0<\alpha_1, \alpha_2, \alpha_3, < 1$ $(\alpha_2+\alpha_3 \leq 1)$, so that $\sigma_2^2 \geq \sigma_3^2$. A similar case hold when $\alpha=\alpha_1+\alpha_2$ and $\beta=\alpha_3$. By induction, the proof can easily be extended to any $s>2$. This shows that the multi-sample estimators $\hat{N}_s$ have smaller (asymptotic) variances than the classical Petersen estimator (based on comparable total sample sizes). Further, looking at (3.2), we observed that $(1-\alpha)(1-\beta)/\alpha\beta$ is a mimimum (for a given $\alpha+\beta=\gamma>0$) when $\alpha=\beta=\gamma/2$. Thus, for the two-sample model, ideally, the two samples sizes n_1 and n_2 should be close to each other (if not, equal). For some related (asymptotic) variance inequalities in the multi-sample case, we may refer to Sen and Sen (1981).

Let us consider next the case of sequential tagging. Note that [Viz., Sen (1982b)] for every $k\geq 1$,

$$M_{k+1}^{o} = E(M_{k+1}) = N\{1-(1-N^{-1})^{k}\}. \quad (3.9)$$

Further, if we let $n=N\alpha$, $\alpha>0$ (here α need not be less than 1, although, in practice, α is small compared to 1), then

$$N^{-1}M_n \rightarrow 1-e^{-\alpha}, \text{ in probability, as } n\rightarrow\infty. \quad (3.10)$$

Proceeding as in Sen (1982b), we observe that as N increases,

$$N^{-\frac{1}{2}}(\hat{N}_n-N) \sim \mathcal{N}(0,(e^{\alpha}-\alpha-1)^{-1}). \quad (3.11)$$

Thus, if we let $n=n_1+n_2 \sim N(\alpha_1+\alpha_2)$ i.e., $\alpha=\alpha_1+\alpha_2$, and compare σ_2^2 with $(e^{\alpha}-\alpha-1)^{-1}$, we have the <u>asymptotic relative efficiency (ARE)</u>

of the estimator $\hat{N}_n$ with respect to the Petersen estimator:

$$\sigma_2^2\ (e^{\alpha}-\alpha-1) = \frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1\alpha_2}\ (e^{\alpha_1+\alpha_2}-\alpha_1-\alpha_2-1) \tag{3.12}$$

and, for a given $\alpha(=\alpha_1+\alpha_2)$, (3.12) is a minimum for $\alpha_1=\alpha_2=\alpha/2$, and this minimum value is given by

$$(2-\alpha)^2\alpha^{-2}(e^{\alpha}-\alpha-1). \tag{3.13}$$

As $\alpha \downarrow 0$, (3.13) converges to 2. Further, for all $\alpha \leq 0.793$, (3.13) is greater than 1, but as α goes beyond 0.793 (up to the upper bound 2), the ARE in (3.13) remains less than 1 and converges to 0 as $\alpha \to 2$. This clearly exhibits the superiority of sequential CMRR to the conventional two-sample CMRR, when α is less than 0.793 (i.e., each sample size $n_1=n_2 \leq 0.395N$). In actual practice, generally n_1/N and n_2/N are both small, and hence, whenever practicable, the sequential tagging seems to be more efficient than the classical CMRR scheme. On the other hand, operationally, the sequential tagging may be less adaptable (because of trap-shyness or other factors), and the consequences of such effects should be given due consideration in choosing a sequential tagging model over a conventional CMRR model.

In (3.11), we have considered the case where the sample size n is a non-stochastic and $n/N \to \alpha$ for some $\alpha > 0$. For both the sequential estimation problems (to be studied in Sections 4 and 5), we have a well-defined stopping rule which yields a stochastic sample size n (assuming non-negative integer values). To cover this more general situation, we need some deeper asymptotic results (mostly, studied in Sen (1982a,b)), and to incorporate them in our main (sequential) analysis.

The probability law in (2.13) may also be conveniently described in terms of a simple <u>urn model</u>. Suppose that in an urn there are N balls, all white, were N is not known. We repeatedly draw a ball at random, observe its color and replace it by a black ball, so that before each draw, there are N balls in the urn. Let W_n be the number of white balls observed in the first n draws, $n \geq 1$. Note that $W_n \leq n$ for every $n \geq 1$ and $W_1 = 1$. Also, W_n is nondecreasing in n (≥ 1). We may then define the M_k as in (2.11)-(2.12), and note that

$$W_k = M_{k+1}, \text{ for every } k \geq 1. \tag{3.14}$$

Also, note that for every k (≥ 1), $k - W_k$ refers to the number of trials (in the first k draws) in which a black ball appeared in the draw. The sequential tagging scheme can be equivalently described in terms of the sequence $\{W_k, k \geq 1\}$. Indeed, in the context of sequential estimation of N, Samuel (1968) considered the stopping variable

$$\begin{aligned} t_c &= \inf\{n: n \geq (c+1) W_n\} \\ &= \inf\{n: (n - W_n)/W_n \geq c\}, \quad c > 0, \end{aligned} \tag{3.15}$$

where t_c can take on only the values $[k(c+1)]$, $k = 1, 2, \ldots$, and $W_{t_c} = m$ whenever $t_c = [m(c+1)]$, $m \geq 1$. Here $[s]$ denotes the largest integer $\leq s$. Then, for every N (≥ 1) and k ($0 < k < \infty$), we may consider a stochastic process

$$Z_n = \{Z_n(t), 0 \leq t \leq k\} \tag{3.16}$$

where

$$Z_n(t) = N^{-1/2}\{W_{[nt]} - N(1 - (1 - N^{-1})^{[Nt]}\}, \quad t \in [0, k]. \tag{3.17}$$

[Note that (3.15) and (3.9) provide the appropriate centering sequence in (3.17)]. Then Z_n belongs to the $D[0,k]$ space, endowed with the

Skorokhod J_1-topology. Also, let $Z=\{Z(t),\ t\varepsilon[0,k]\}$ be a Gaussian function (belonging to the C[0,k]space) with zero drift and covariance function

$$EZ(s)Z(t) = e^{-t}\{1-(1+s)e^{-s}\},\ \text{for } 0\le s\le t\le k. \tag{3.18}$$

Then, through a suitable martingale construction, the following result was established in Sen (1982a):

For every $k:0<k<\infty$, as $N\to\infty$,

$$Z_n \xrightarrow{\mathcal{D}} Z,\ \text{in the } J_1\text{-topology on } D[0,k]. \tag{3.19}$$

Next, we may recall that $(1-(1-n^{-1})^n) \sim 1-e^{-n/N}$, so that for $W_{[Nt]}$ in (3.17), the centering constant may also be taken as $N(1-e^{-t})$. For the asymptotic behavior of t_c in (3.16), we consider a sub-interval $I^* = [a,b]$ of $[0,1]$, where $0<a<b<1$. For every $m\varepsilon I^*$, define t_m by the solution of the equation

$$m = (1-e^{-t_m})/t_m,\ m\varepsilon I^* \tag{3.20}$$

Note that $mt_m \le 1$, $\forall\ m\varepsilon[0,1]$; $t_1=0$ and as m moves from 1 to 0, t_m monotonically goes to $+\infty$. Then for every $N(\ge 1)$, we consider a stochastic process $U_N = \{U_N(m),\ m\varepsilon I^*\}$ by letting

$$\tau_{Nm} = \inf\{n\ge 1:mn\ge W_n\},\ U_N(m)=N^{-1/2}\{\tau_{Nm}-Nt_m\},\ m\varepsilon I^*. \tag{3.21}$$

Also, let $U = \{U(m),\ m\varepsilon I^*\}$ be a Gaussian function with zero drift and covariance function

$$EU(m)U(m') = e^{-tm}\{1+(1+t_{m'})e^{-tm'}\}/\{(m-e^{-tm})(m'-e^{-tm'})\}, \tag{3.22}$$

for $m\le m'$. Then, the following invariance principle has been established in Sen (1982a).

For every $0<a<b<1$, as N increases

$$U_N \underset{\mathcal{D}}{\to} U, \text{ in the } J_1\text{-topology on } D[a,b]. \tag{3.23}$$

Let us now go back to (2.13) and note that for every $n \geq 1$,

$$Z_n^{(1)}(N) = N^{1/2}(\partial/\partial N) \log L_n(N)$$
$$= N^{1/2}\{\Sigma_{k=2}^{n}(1-X_k)/(N-M_k)-(n-1)/N\}; \tag{3.24}$$
$$Z_n^{(2)}(N) = -N(\partial^2/\partial N^2) \log L_n(N)$$
$$= N \Sigma_{k=2}^{n} (1-X_k)/(N-M_k)^2 - (n-1)/N. \tag{3.25}$$

We also define

$$Z_n^*(n) = \Sigma_{k=2}^{n} (N-M_k)^{-1} - (n-1)/N, \; n \geq 2. \tag{3.26}$$

Then, it follows from Sen (1982b) that for every $\alpha: 0<\alpha\leq k<\infty$,

$$n/N \to \alpha \Rightarrow | Z_n^*(N) - (e^{\alpha}-\alpha-1)| \overset{p}{\to} 0, \text{ as } N\to\infty. \tag{3.27}$$

Further, if we consider a stochastic process $V_{Nn} = \{V_{Nn}(t), t\varepsilon [0,1]\}$ by letting

$$V_{Nn}(t) = Z_{n(t)}^{(1)} (N)/\{Z_n^*(N)\}^{1/2}; \tag{3.28}$$

$$n(t) = \text{man } \{k: Z_k^*(N) \leq t \, Z_n^*(N)\}, \; t\varepsilon [0,1], \tag{3.29}$$

then for every $\alpha\varepsilon(0,k)$, $n/N\to\alpha$, ensures that

$$V_{Nn} \underset{\mathcal{D}}{\to} V, \text{ in the } t_1\text{-topology on } D[0,1], \tag{3.30}$$

where V is a standard Wiener process on [0,1]. The proof of (3.30) is again based on suitable martingale constructions. It was actually shown in Sen (1982b) that for every $N(\geq 1)$, $\{Z_k^{(1)}(N), k\geq 1\}$ and

$\{Z_k^{(2)}(N), Z_k^*(N), k\geq 1\}$ are both (zero mean) martingales, and this basic result paved the way for a simple proof of (3.30). Actually, it was shown there that for every $\varepsilon>0$, whenever $n/N\to\alpha:0<\alpha<k(<\infty)$,

$$\max_{n\varepsilon\leq k\leq n} | N^{-\frac{1}{2}} (\hat{N}_k-N) - Z_k^{(1)}(N) / Z_k^{(2)}(N)| \overset{p}{\to} 0, \tag{3.31}$$

so that defining n(t) as in (3.29) and letting

$$V_{Nn}^*(t) = \{ N^{-\frac{1}{2}} (\hat{N}_{n(t)}-N)(e^{\alpha}-\alpha-1)^{1/2}\}, \varepsilon\leq t\leq 1, \tag{3.32}$$

$V_{Nn}^* = \{V_{Nn}^*(t), \varepsilon\leq t\leq 1\}$ and $V_\varepsilon^* = \{V^*(t) = V(t)/t, \varepsilon\leq t\leq 1\}$, we arrive at the following:

For every $\varepsilon>0$, whenever $n/N\to\alpha:0<\alpha\leq k(<\infty)$,

$$V_{Nn}^{*\varepsilon} \underset{\mathcal{D}}{\to} V_\varepsilon^*, \text{ in the } J_1\text{-topology on } D[\varepsilon,1]. \tag{3.33}$$

A direct corollary to (3.33) is the following:

If $\{\nu_n\}$ is any sequence of positive integer valued random variables, such that as n increases, $n^{-1}\nu_n \to 1$, in probability, then whenever $n/N \to \alpha(0<\alpha\leq k<\infty)$

$$N^{-\frac{1}{2}}(\hat{N}_{\nu_n} - N) \sim \mathcal{N}(0,(e^{\alpha}-\alpha -1)^{-1}). \tag{3.34}$$

In the development sketched above, we have tacitly assumed that $n/N\to\alpha$ for some $\alpha>0$. As we shall see in the next two sections, this condition may not hold for the sequential problems under considerations. What we would have more generally is that as N increases $n(=n_N) \to +\infty$, but $N^{-1}n_N$ may converge to 0. The martingale characterization based proof, considered in Sen (1982b), remains intact under this condition too, and parallel to (3.34), we have the following result (whose proof is omitted):

If $\{\nu_n\}$ is any sequence of positive integer valued random variables, such that $n^{-1}\nu_n \overset{p}{\to} 1$, as $n\to\infty$, and $n\{=n_N\}$ is such that $N^{-1}n_N^2 \to\infty$ as $N\to\infty$ (but $N^{-1}n_N$ may not converge to a positive number), then in the limiting degenerate case: $N^{-1}n_N \to 0$,

$$N^{-\frac{1}{2}} n(N^{-1}\hat{N}_{\nu_n}-1) \sim N(0,2) \qquad (3.35)$$

In passing, we may also remark that for the Petersen two-sample estimator, whenever $n_1=n_2=n/2$, and $n(=n_N)$ increases but $N^{-1}n_N \to 0$ as $N\to\infty$, we have, parallel to (3.2),

$$N^{-\frac{1}{2}} n(N^{-1}\hat{N}_P-1) \sim N(0,4). \qquad (3.36)$$

Again, the proof follows by the convergence of the hypergeometric law to a normal law (when n is large). A similar modification of the asymptotic normality result in (3.5) for $N^{-1}n_N \to 0$ follows. We shall find these results very convenient for our (proposed) sequential analysis.

4. FIXED PERCENTAGE WIDTH CONFIDENCE INTERVAL FOR N

As has been discussed in Section 1, our goal is to construct a confidence interval I_n, based on n units drawn from a population of size N, such that for some predetermined $1-\alpha$ (the <u>confidence coefficient</u>, $0<\alpha<1$) and $d(>0)$,

(i) $P\{N\varepsilon I_n\} \geq 1 - \alpha$, (4.1)

(ii) The width of $I_n \leq 2dN$. (4.2)

We are naturally tempted to use suitable estimates of N to provide such a solution to (4.1)-(4.2). For some solutions to this problem, we may refer to Darling and Robbins (1967) and Samuel (1968). We shall be mainly concerned here with the asymptotic case, where d is chosen small, and in this case, the results in the preceding sections can be incorporated with advantage to provide simpler solutions.

If we are able to use the asymptotic normality (of the standardized form) of an estimator $\hat{N}$ of N, then the solution for n can be derived in terms of d and the asymptotic variance function. However, such an asymptotic variance function, generally, depends on the unknown N, and hence, no fixed smaple size solution may exist for all N. For this reason, we take recourse to plausible sequential procedures which provide such a solution, at least, in the asymptotic case where $d \downarrow 0$, and possess some other desirable properties too. To motivate such a sequential procedure, first, we consider an asymptotically optimal fixed-sample size procedure, where the sample size $n_d (=n_{dN})$ depends on N as well.

For every d(>0), we may consider a d'(>0), such that

$$2d = (1-d')^{-1} - (1+d')^{-1}. \tag{4.3}$$

Then, we define n_d, by letting

$$n_d = \inf\{ n \geq 2 : n \geq (2N)^{\frac{1}{2}}(d')^{-1}\tau_{\alpha/2} \}, \; d > 0 \tag{4.4}$$

where $\tau_{\alpha/2}$ is the upper 50 α °/. point of the standard normal distribution. Note that n_d in (4.4) depends on N as well. Note that be (3.35), for large N, $P\{ N^{-\frac{1}{2}} n_d |N^{-1}\hat{N}_{n_d} - 1| \leq \sqrt{2}\, \tau_{\alpha/2}\} \to 1 - \alpha$, so that $P\{|N^{-1}\hat{N}_{nd}-1| \leq \sqrt{2}\, \tau_{\alpha/2} \sqrt{N}/n_d \; (\leq d')\} \to 1-\alpha$, and hence,

$$I_{n_d} = [\hat{N}_{n_d}(1+d')^{-1}, \hat{N}_{n_d}(1-d')^{-1}] \tag{4.5}$$

provides a confidence interval for which (4.1) and (4.2) hold, at least, for $d \downarrow 0$. Since n_d depends on d as well as the unknown N, we may estimate n_d in a sequential setup where we use the sequential tagging scheme along with the Chow and Robbins (1965) general approach. The necessary adjustments can easily be based on the results in Section 3.

Let $\{\hat{N}_n, n \geq 2\}$ be the sequence of estimators of N, based on the sequential tagging scheme in Section 2. Keeping (4.4) in mind, define a <u>stopping variable</u> $\{\nu_d; d>0\}$ by letting

$$\nu_d = \inf\{n \geq n_0 : \hat{N}_n^{-\frac{1}{2}} \cdot d'n \geq \sqrt{2}\, \tau_{\alpha/2}\}, \; d>0, \tag{4.6}$$

where d' is defined by (4.3) and n_0 is a suitable positive integer (≥ 2). Based on this stopping variable, the proposed (sequential) confidence interval for N is

$$I_{\nu_d} = [(1+d')^{-1} \hat{N}_{\nu_d} , \; (1-d')^{-1} \hat{N}_{\nu_d}] . \tag{4.7}$$

Thus, the width of the interval I_{ν_d} is equal to $2d\hat{N}_{\nu_d} = (2dN)(N^{-1}\hat{N}_{\nu_d})$, so that for (4.2) to hold for small $d(>0)$, we need to show that

$$N^{-1}\hat{N}_{\nu_d} \to 1, \text{ in probability, as } d \downarrow 0. \tag{4.8}$$

Further, if we are able to show that

$$\nu_d / n_d \to 1, \text{ in probability, as } d \downarrow 0, \tag{4.9}$$

then, by an appeal to (3.35), we are able to claim that (4.1) holds for $d \downarrow 0$. Verification of (4.8) is also facilitated by (4.9) and (3.35).

Thus, (4.9) is the key to our solutions to (4.1) and (4.2). To establish (4.9), we need to strengthen some of the asymptotic results in Sen (1982a,b), and these are considered first.

Note that $M_k \le k-1$, for every $k \ge 1$, so that by (3.26), for every $n \ge 2$

$$\begin{aligned} 0 \le Z_n^*(N) &\le \Sigma_{k=2}^{n} (N-k+1)^{-1} - (n-1)/N \\ &\sim \log (N-1) - \log (N-n+1) - (n-1)/N \qquad (4.10) \\ &\sim \tfrac{1}{2}(n/N)^2, \text{ whenever } n/N \text{ is small.} \end{aligned}$$

Further, if we let $n^0 \sim \varepsilon N$ for some $\varepsilon > 0$, then proceeding as in Sen (1982b), we have

$$|Z_{n^0}^*(N) - {}^1/_2(n^0/N)^2| = o_p((n^0/N)^2). \qquad (4.11)$$

Therefore, for every $k: 2 \le k \le n^0$ and η $(0 < \eta < {}^1/_2)$,

$$\{Z_k^*(N)/Z_{n^0}^*(N)\}^{\frac{1}{2}-\eta} \le (k/n^0)^{1-2\eta}(1+o(1)), \qquad (4.12)$$

with a probability converging to 1, as $N \to \infty$. Next, by virtue of the martingale property of $\{Z_k^{(1)}(N); k \ge 1\}$, we may extend Theorem 3.2 of Sen (1982b) wherein the Skorokhod J_1-topology may be strengthened to the d_q-metric defined by

$$d_q(x,y) = \sup\{|x(t)-y(t)|/q(t): 0<t<1\} \qquad (4.13)$$

where $q = \{q(t), t \in [0,1]\}$ may be taken as

$$q(t) = t^{\frac{1}{2}-\eta}, \text{ for some } \eta > 0;\ t > 0; \qquad (4.14)$$

See Theorem 2.4.8 of Sen (1981) in this respect. This yields that for every (fixed) $\varepsilon > 0$, $n^0 \sim \varepsilon N$,

$$\max_{k\le n^{\circ}} |Z_k^{(1)}(N)(n^{\circ}/k)^{1-2\eta}| = O_p((Z_n^{*}(N))^{1/2})$$
$$= O_p(n^{\circ}/N), \qquad (4.15)$$

so that with probability approaching to unity(as $N\to\infty$),

$$|Z_k^{(1)}(N)| \le C(N^{-1}(n^{\circ})^{2\eta} k^{1-2\eta}),\ \forall k \le n^{\circ}, \qquad (4.16)$$

where $C(<\infty)$ is a suitable constant. In a similar manner, Theorem 3.3 of Sen (1982b) can also be extended under the d_q-metric and this yields that

$$\max_{k\le n^{\circ}} |(n^{\circ}/k)^{1-2\eta} \{Z_k^{(2)}(N) - Z_k^{*}(N)\}| = O_p(n^{\circ}N^{-3/2}). \qquad (4.17)$$

We use these results to derive some (crude) a.s. lower bounds for the $\hat{N}_n$ which will be needed, in turn, for the verification of (4.9). For this, we consider the following <u>estimating function</u>:

$$w_{Nn}(L) = N^{1/2}\Sigma_{k\le n}\{(1-X_k)L(L-M_k)^{-1}-1\},\ n\ge 2,\ L \text{ positive integer}. \qquad (4.18)$$

Note that $w_{nN}(L)$ is $\searrow$ in L (where $L\ge M_n$) and $w_{Nn}(L)=0$ provides the MLE $\hat{N}_n$. Side by side, we let

$$w_{Nn}^{*}(L) = (1-L/N)\ N^{\frac{1}{2}}\ \Sigma_{k\le n} M_k(L-M_k)^{-1},\ n\ge 2,\ L\ge M_n. \qquad (4.19)$$

Then, it is easy to verify that for every N and L,

$$\{w_{Nn}(L)-w_{Nn}^{*}(L);n\ge 2\} \text{ is a zero-mean martingale.} \qquad (4.20)$$

Using the Hájek-Rényi-Chow inequality (for submartingales), we obtain that for every $\eta>0$ (and $n^{\circ}\sim N\varepsilon$), there exists a finite positive $c(\eta)$, such that

$$P\{m^{-1-\eta}|w_{Nm}(L)-w_{Nm}^{*}(L)|>1,\ \text{for some } m:n\le m\le n^{\circ}\}$$
$$\le c(\eta)n^{-2\eta},\ \text{where } c(\eta) \text{ does not depend on } n. \qquad (4.21)$$

Also, note that for every $k\geq 2$, $M_k(L-M_k)^{-1} = L(L-M_k)^{-1} - 1 = (1-L^{-1}M_k)^{-1}-1$, where $M_k\leq k-1$, with probability 1. Hence, proceeding as in Sen (1982a,b), it can be shown that for L<N

$$w^*_{Nn}(L) \sim {}^1/_2(1-L/N)N^{1/2}L^{-1}n^2,\ \forall\ n \leq n^o, \tag{4.22}$$

with a probability converging to 1 as $N\to\infty$. Further, note that $w_{Nn}(L)$ is $\gtreqless 0$ according as $N_n \gtreqless L$. Hence, on letting $L = N^{1/2}n^{1-\eta}$ or n^o according as $n^{1-\eta}$ is less than $n^o/N^{1/2}$ or not, we obtain from (4.21) and (4.22) that

$$2\hat{N}_m \geq (N^{1/2}n^{1-\eta}) \wedge n^o, \text{ for every } m\text{: } n\leq m\leq n^o, \tag{4.23}$$

with probability converging to 1, as $n\to\infty$.

Now, looking at (4.6), we note that $\sqrt{2}\tau_{\alpha/2}/d'\to\infty$ as $d\downarrow 0$. So using (4.23), we conclude that in (4.6), we may replace n_o by $N^{1/4}$ (a.s., for $d\downarrow 0$), and this immediately leads us to

$$\nu_d \geq (\tau_{\alpha/2}(d')^{-1})N^{(3-\eta)/8} \text{ a.s., as } d\downarrow 0 \tag{4.24}$$

At this stage, we note that for (4.1)-(4.2), we consider an asymptotic setup where N is large and d(or d') is small. To combine these two features into a single one, we let for an arbitrarily small γ (>0),

$$d = d_N = N^{-\gamma}, \tag{4.25}$$

so that we may take the limit $N\to\infty$, treating γ as an (unknown) constant. We shall restrict ourselves to $\gamma \in (0, \frac{1}{2})$, as for $\gamma > \frac{1}{2}$, $\nu_d \geq n^o$ a.s., as $d\downarrow 0$ so that the results would follow more easily from Sen (1982a,b). Then, in (4.23), on choosing $n\sim N^{\gamma+(3-\eta)/8}$ along with (4.26) and (4.25), we obtain that

$$\nu_d \geq (2\tau_{\alpha/2})N^{7/16+3\gamma/2-\eta/4+\eta^2/16} \text{a.s., as } d\downarrow 0. \tag{4.26}$$

In the next step, in (4.23), we choose $n \sim N^{7/16+3\gamma/2-\eta/4+\eta^2/16}$ and and reiterate the process, leading to a sharper bound for ν_d, where in (4.26), we may replace the exponent of N by $15/32+7\gamma/4-\eta/8+\eta^2/16 -\eta(7/32+3\gamma/2-\eta/4+\eta^2/16)$. Depending on the choice of $\eta(>0)$ and $\gamma(>0)$, this chain can be repeated a (finite) number of times, until in (4.23), we have $\hat{N}_m \geq n^o$ (a.s.), and this leads us to

$$\nu_d \geq \sqrt{2\varepsilon}\tau_{\alpha/2} N^{\frac{1}{2}+\gamma} \text{ a.s., as } d\downarrow 0. \tag{4.27}$$

Note that by (4. 4) and (4.25),

$$n_d \sim \sqrt{2}\tau_{\alpha/2} N^{\frac{1}{2}+\gamma}, \text{ as } d\downarrow 0 \ (N\to\infty). \tag{4.28}$$

Next, we again appeal to (4.21)-(4.22), where we choose $L=N(1-\varepsilon)$, $\varepsilon>0$ and $n \sim \sqrt{\varepsilon}\, N^{\frac{1}{2}+\gamma}$. This leads us to that with probability $\to 1$

$$\hat{N}_m \geq (1-\varepsilon)N, \text{ for every } m: \sqrt{\varepsilon}n^{\frac{1}{2}+\gamma} \leq m \leq n^o. \tag{4.29}$$

[Note that this choice of L is different from that in (4.23), but is consistent with the system in (4.21)-(4.22).] Similarly, letting $L = N(1+\varepsilon)$ and $n \sim \sqrt{\varepsilon}n^{\frac{1}{2}+\gamma}$, $\varepsilon>0$, we have by (4.21)-(4.22),

$$\hat{N}_m \leq (1+\varepsilon)N, \text{ for every } m: \sqrt{\varepsilon}N^{\frac{1}{2}+\gamma} \leq m \leq n^o. \tag{4.30}$$

For $m>n^o$, we use (3.31)(with $\alpha\geq 1$). and obtain that

$$\max_{n^o\leq n\leq n_\alpha} |(\hat{N}_n/N-1)| = O_p(N^{-1/2}) \to 0, \text{ as } N\to\infty. \tag{4.31}$$

Thus, using (4.6), (4.29), (4.30) and (4.31), we claim that (for $\gamma \leq \frac{1}{2}$)

$$\nu_d/n_d \leq 1+\varepsilon \text{ a.s. as } d\downarrow 0, \tag{4.32}$$

and further,

$$\max_{\varepsilon n_d \le n \le (1+\varepsilon) n_d} |\hat{N}_n/N-1| \overset{p}{\to} 0, \text{ as } d\downarrow 0. \qquad (4.33)$$

Then (4.27), (4.32) and (4.33) ensure (4.8). Further, note that by definition in (4.6),

$$\nu_d/n_d \geq (N^{-1}\hat{N}_{\nu_d})^{1/2}; \ (\nu_d-1)/n_d < (N^{-1}\hat{N}_{\nu_d-1}). \qquad (4.34)$$

As such, (4.27), (4.32), (4.33) and (4.36) ensure that (4.9) holds. This completes the proof of (4.1)-(4.2). Note that in this context, the choice of η in (4.21) and γ in (4.25) are quite arbitrary. In the literature, (4.1) is termed the asymptotic consistency and (4.9) as the asymptotic efficiency, although, for the later, usually one uses the stronger form that $E\nu_d/n_d \to 1$ as $d\downarrow 0$. The later result can also be deduced by more elaborate analysis, but we shall not do it. For the statistical interpretations, (4.1), (4.2) and (4.9) suffices.

5. ASYMPTOTICALLY OPTIMAL POINT ESTIMATION OF N.

As explained in Section 1, we want to estimate N incorporating a cost-function as well as a risk-function in this formulation. Towards this, we consider the simple cost function

$$c(n) = c_0 + cn, \ n\geq 2, \ c_0>0, \ c>0, \qquad (5.1)$$

where c_0 denotes the overall management cost and c is the cost per unit sample (for capture, mark and release operations). Also, for the risk functions we consider the coefficient of variation measure of the loss functions (this is comparable to the percent width problem treated in Section 4.) Thus, the compound risk function of an estimator N_n^* (of N) based on n units is

$$N^{-2}E(N_n^*-N)^2 + c_0 + cn, \ n\geq 2. \qquad (5.2)$$

Our goal is to choose n in such a way that (5.2) is a minimum. This may be termed the minimum risk point estimation of N. Since $E(N_c^*-N)^2/N^2$ depends on the the unknown N, an optimal n in this respect also depends on N. As such, we may again take recourse to a sequential scheme to achieve such an optimal solution in (at least) an asymptotic setup where N is large and c is small. In this context, we may not have a finite second moment [See for example, (2.3)], and hence, employing the second moment may induce some necessary changes in the estimators, such that they would have finite second moments to qualify for study. These complications may easily be avoided by working with the asymptotic distributional risk (ADR) where the second moment is computed from the asymptotic distribution of a normalized version of N_n^*. This concept works out well for small $c(>0)$ and has certain advantages too [viz. Sen (1986)].

Using the results of Sections 2,3, and 4, for the MLE, the ADR version of (5.2) for n (large but) $\leq n^o \sim \varepsilon N$ is given by

$$\frac{2N}{n^2} + c_0 + cn, \qquad c>0,\ c_0>0, \qquad (5.3)$$

while, for $n\geq n^o$, we have the parallel expression

$$N^{-1}(e^{n/N}-n/N-1)^{-1} + c_0 + cn. \qquad (5.4)$$

Thus, if N were known (but large), an optimal n (minimizing the ADR) is given by

$$n_c \sim N^{1/3}\, c^{-1/3}, \quad c>0 \tag{5.5}$$

and the corresponding minimum ADR is given by

$$c_0 + 3c^{2/3}\, N^{1/3}. \tag{5.6}$$

Our goal is to consider a sequential procedure, for which (5.5)-(5.6) are attained for small values of c(>0). Keeping (5.5) in mind, we define a stopping number by

$$V_c = \inf \{n \geq n_0 : n \geq (\hat{N}_n c^{-1})^{1/3}\}, \; c>0. \tag{5.7}$$

Note that as in the case of ν_d in Section 4, ν_c is $\searrow$ in c and $\lim_{c\downarrow 0} \nu_c = +\infty$ a.s.. Then, our proposed sequential point estimator of N is $\hat{N}_{\nu_c}$, where for every n, $\hat{N}_n$ is defined as in Section 2. The ADR of this sequential estimator is

$$N^{-2}\, \tilde{E}(\hat{N}_{\nu_c} - N)^2 + c_0 + c\tilde{E}\, \nu_c, \tag{5.8}$$

where $\tilde{E}$ stands for the expectation computed from the appropriate asymptotic distribution. Then, our basic goal is to show that as $c\downarrow 0$, (5.8) approaches (5.6), and, moreover, $\nu_c/n_c \xrightarrow{p} 1$, as $c\downarrow 0$. As in Section 4, here also, to tackle simultaneously the operations of two limits $N\to\infty$ and $c\downarrow 0$, we let

$$c = c_N = N^{-\gamma}, \text{ for some } \gamma(>0). \tag{5.9}$$

By virtue of (5.5) and (5.9), we have

$$n_c \sim c^{-(1+\gamma)/3\gamma} \text{ as } c\downarrow 0. \tag{5.10}$$

In the classical minimum risk (sequential) estimation problem [See for example, Sen and Ghosh (1981)] , usually, we have $n_c \sim c^{-\frac{1}{2}}$, which, in (5.10), leads to $\gamma = 2$. However, in our case, we do not have to restrict ourselves, to $\gamma = 2$. We assume that

$$c = c_N = N^{-\gamma}, \text{ for some } \gamma > \tfrac{1}{2} . \tag{5.11}$$

In that case, we have $n_c \sim N^{1/3} N^{\gamma/3} = N^{(1+\gamma)/3} = N^{\frac{1}{2}+\xi}$, for some $\xi > 0$. As such, we may virtually repeat the steps in Section 4 and conclude that

$$\nu_c/n_c \to 1, \text{ in probability, as } c \downarrow 0. \tag{5.12}$$

Thus, to show that (5.8) and (5.6) are convergent equivalent (for $c \downarrow 0$), if suffices to show that

$$N^{-2}\tilde{E}(\hat{N}_{\nu_c} - N)^2 \sim 2c^{(2\gamma-1)/3\gamma}, \text{ as } c \downarrow 0. \tag{5.13}$$

Towards this, we note that by (3.35) and (5.12), as $c \downarrow 0$,

$$n_c N^{-1/2}(N^{-1}\hat{N}_{\nu_c} - 1) \sim \mathcal{N}(0,2), \tag{5.14}$$

so that the mean square computed form this asymptotic distribution is equal to

$$2N\, n_c^{-2} \sim 2N^{1/3} c^{2/3} \sim 2c^{(2\gamma-1)/3\gamma}, \text{ as } c \downarrow 0. \tag{5.15}$$

This completes the proof of (5.13), and hence the asymptotic (as $c \downarrow 0$) equivalence of (5.8) and (5.6) is established.

Looking at (5.15), we may observe that for $\gamma < \frac{1}{2}$, (5.13) blows up as $c \downarrow 0$, so that the risk of the estimator $\hat{N}_{\nu_c}$ (or even $\hat{N}_{\nu_c}$) can not be made to converge to 0 when $c \downarrow 0$. This runs contrary to the normal phenomenon that as $c \downarrow 0$, $n_c \to +\infty$ and the risk of the MLE $\hat{N}_{n_c}$ (or $\hat{N}_{\nu_c}$) should converge to 0 (presumably at the rate of some positive power of c). This apparent anomaly is due to the restraint of γ in (5.9), where a very small $\gamma(<\frac{1}{2})$ induces a rather slow rate of decay of c (i.e., $n_c = o(N^{\frac{1}{2}})$), and for $n < \varepsilon N^{\frac{1}{2}}$, $\varepsilon>0$, the asymptotic distributional results on $\hat{N}_n$ are not that strong to ensure a small ADR. Critically, we need that $N^{-1}n_c^2$ becomes large as $c\downarrow 0$ (so that $\hat{N}_{\nu_n}/N \to 1$, in probability), as otherwise the stochastic convergence of $N^{-1}\hat{N}_{\nu_n}$ (to 1) may not hold and this, in turn, may distort the basic requirement that $\nu_c/n_c \to 1$ in probability, as $c\downarrow 0$. If (5.9) holds for some $\gamma \leq \frac{1}{2}$, one possibility is to consider a somewhat more general form of the risk function:

$$a\tilde{E}(\hat{N}_n/N-1)^2 + c_0 + cn,\ n\geq 1,\ a>0,\ c_0>0,\ c>0, \tag{5.16}$$

where $a(=a_N)$ is large compared to $c\ (=c_N)$, in the sense that $c_N/a_N \sim N^{-\gamma}$ for some $\gamma > \frac{1}{2}$. This would then ensure that $\nu_c/n_c \to 1$, in probability, as $c\downarrow 0$, although the asymptotic risk $(c_0 + 3a(c/a)^{(2\gamma-1)/3\gamma})$ still may not converge when $c/a\downarrow 0$ (but $a \uparrow\infty$), as $c\downarrow 0$. This means that in such a case, the asymptotic minimum risk property can be attained, but the asymptotic risk may not be finite. This can further be modified by using a preliminary estimator of N based on a pilot sample of size n_0 (which may even be drawn sequentially, so that $n_0 = \inf\{k\geq 2:\ k-W_k\geq m\}$ for some specified $m(\geq 1)$), and having such a rough estimate of N, to consider an (adaptive) risk function for which the prescribed solutions workout well. In many practical problems, a (crude) estimate of N can be obtained from previous surveys, so that such a formulation of an adaptive risk function may be done effectively.

This work was supported by the Office of Naval Research, Contract No. N00014-83-K0387.

REFERENCES

Chapman, D. (1951). Some properties of the hypergeometric distribution with applications to zoological census. *Univ. Calif. Publ. Statist.* 1, 131-160.

Chapman, D. (1952). Inverse, multiple and sequential sample censuses. *Biometrics* 8, 286-306.

Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence interval for the mean. *Ann. Math. Statist.* 36, 457-462.

Darling, D.A. and Robbins, H. (1967). Finding the size of a finite population. *Ann. Math. Statisti.* 38, 1392-1398.

Darroch, J.N. (1958). The multiple recapture census, I: Estimation of closed population. *Biometrika* 45, 343-359.

Goodman, L. A. (1953). Sequential sampling tagging for population size proclaims. *Ann. Math. Statist* 24, 56-69.

Petersen, C. G. J. (1896). The yearly immigration of young plaice into the Limford from the German Sea. *Rep. Danish Biol. Stat.* 6, 1-68.

Samuel, E. (1968). Sequential maximum likelihood estimation of the size of a population. *Ann. Math. Statist.* 39, 1057-1068.

Schnabel, Z.E. (1938). The estimation of the total fish population of a lake. *Amer. Math. Mem.* 45, 348-352.

Sen, A. R. and Sen, P. K. (1981). Schnabel-type estimators for closed populations, with multiple marking. *Sankhyā, Ser B.* 43, 68-80.

Sen, P. K. (1981). *Sequential Nonparametrics*. New York: Wiley.

Sen, P. K. (1982a). A renewal theorem for an urn model. *Ann. Probability* 10, 838-843.

Sen, P. K. (1982b). On asymptotic normality in sequential sampling tagging. *Sankhyā, Ser A.* 44, 352-363.

Sen, P. K. (1986). On the asymptotic distributional risks of shrinkage and preliminary test versions of maximum likelihood estimators. *Sankhya, Ser. A 48*, in press.

Sen, P. K. and Ghosh, M. (1981). Sequential point estimation of estimable parameters based on U-statistics. *Sankhyā, Ser A* 43, 331-344.