**Journal of Statistical and Mathematical Sciences**
**ISSN 3065-100X**
Volume 13 Issue 1
January – March 2025
Impact Factor: 8.99
Enquiry: contact@continentalpub.online
Published by Continental Publication | http://continentalpub.online/index.php/jsms

# METHODS FOR ESTIMATING HIDDEN POPULATIONS: TOOLS AND TECHNIQUES

***Dr. Ahmad Ziad Al-Khatib and Dr. Rana Hisham Al-Momani***
Department of Statistics, Yarmouk University, Jordan
Currently at: Department of Mathematics and Statistics, Jordan University of Science and Technology, Jordan

***Abstract:*** *Statistics, as the science of drawing meaningful conclusions about populations from representative samples, employs various sampling techniques, including simple, stratified, systematic, and cluster random sampling. The choice of sampling method hinges on research objectives, available population information, and budget constraints. From these samples, valuable insights into population parameters are derived, with a focus on mean (μ), proportion (p), variance (σ²), and total (τ). Additionally, the population size (N) serves as a crucial yet less common parameter.*

*In the context of finite populations, estimating these parameters and their variances relies on a known value for N. However, when N is unknown, it becomes necessary to estimate it beforehand to accurately assess population totals and estimator variances.*

*This paper delves into the intricacies of statistics, emphasizing the significance of appropriate sampling techniques and parameter estimation, particularly in scenarios where the population size remains uncertain.*

***Keywords:*** *Statistics, sampling techniques, parameter estimation, population parameters, population size.*

## 1. Introduction and Some Related Topics

Statistics is the science of making inference about a population using the information contained in a sample selected from it. The sample can be chosen by one of the techniques such as simple, stratified, systematic, cluster random sampling, etc. The choice of the technique depends on the objectives of the study, information available about the population of interest and the budget. The information obtained from the chosen sample is used to estimate the population parameters. For finite population, the main population parameters are the mean (□), proportion (p), variance (□$^2$ ), total (□). Less common parameter is the population size ($N$ ). Estimation of the parameters and their variances depend on $N$ , which is usually known. If $N$ is unknown, then it has to be estimated first so that the population total and the variances of the estimators can be estimated.

Assume that we have a population of size $N$ (known). Let $O_1,\ldots,O_N$ be the population measurements. The population mean, total and variance are:

$N$

□□□ $Oi$ $N$ 2 □$i$□1 ($Oi$ □□)2 □□ $i$ 1 , □□□$O_i$ =$N$μ, □ □ .

DOI:https://doi.org/10.5281/zenodo.15910173

□ $i$□1 $N$

A simple random sample, SRS, of size $n$ is a sample obtained from the population in such a way so that all possible samples of size $n$ have equal chance of being the chosen sample, i.e. $P$( a subset of size $n$ from a population of size $N$ is the chosen sample )= 1/□ $N_n$ □. Let $X_1,\ldots,X_n$ be a SRS of size $n$ from this population.

The usual estimators of the population mean& to talare, respectively:

$n$

□ˆ □□$i$ □1$X$ $i$ , □ ˆ□$N$□ˆ .

$n$

It is well known that □ˆ &□ˆ are unbiased estimators of □&□, respectively. Their variances are:

2 $N$ □$n$□

$Var$(□ˆ)□□$n$ □□□ $N$ □1□□ & $Var$(□ˆ)□$N$2$Var$□□ˆ□.

The main concern in this work is the estimation of the population total □ when $N$ is unknown. If $N$ is known, then some estimators of $N$ can be used as a *guard* against unsuitable estimates of □. Capture-Recapture technique is the main method used to estimate $N$ . There are two main procedures of this technique; Capture- Recapture with Direct Sampling and Capture- Recapture with Indirect (or Inverse) Sampling:

□ **Direct Sampling**

The direct Capture- Recapture sampling, known as *Petersen's method*, goes back to 1894. Assume that there is a closed population with equal chance of each member to be selected. A random sample of $m$ elements is drawn, tagged (marked) and then released back into the population. Then, after some enough period of time necessary for the marked units to mix with the remaining elements of the population, a second random sample of size $n$ is drawn. Let □ be the number of recaptured elements in the second sample. Then, the Petersen estimator of the population size is:

ˆ □ $nm$.

$N_P$

$T$

An approximate estimator of the variance of $\hat{N}$ is (Sekar and Deming, 1949):

$Var$ˆ ($N$ˆ$P$) □ $mn$($m$ □$T$3)($n$□$T$) ; $T$ □ 0 .

$T$

□

Actually, $N\ P$ is the maximum likelihood estimator (MLE) and also the method of moments estimator (MME) of $N$ . A modified estimator of $N$ was proposed by Chapman (1951):

$$\hat{N}_C □ \frac{(m+1)(n+1)}{T+1} □1,$$

with variance estimated by $Var$ˆ ($N$ˆ$C$) □ ($m$ □1)(($Tn$□1) 1)(2$m$($T$□□$T$2))($n$□$T$).

□

(Scheaffer et al., 1995). This estimator has the advantage of being valid even when $T$□ 0.

□ **Inverse Sampling** (Indirect Sampling)

**Journal of Statistical and Mathematical Sciences**
Vol. 13 No. 1 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15910173

Inverse sampling is another method for estimating $N$. In this method, a random sample of size $m$ is chosen, marked and released. Later, elements are randomly selected from the population until $k$ (fixed in advance) elements are being recaptured, then

$$\hat{N} = \frac{T m}{k}, \quad k \le m,$$

Where, $T$ is the total number of elements selected in the second random sample to obtain $k$ previously captured elements. The variance of $\hat{N}$ is estimated by

$$\widehat{Var}(\hat{N}) = \frac{m^2 T (T - k)}{k^2 (k + 1)}, \text{ (Scheaffer et al., 1995).}$$

Capture-Recapture technique is an old method used to estimate the size of fish and wildlife population. Later on, the method was used for estimating other population sizes. Azevedo-Silva et al. (2009) analyzed the number of cases and incidence of childhood acute lymphoblastic leukemia by using two source capture-recapture procedures in three different cities in Brazil. Estimating of birth and death rates in India was considered by SeKar and Deming (1949). Estimating the population size of Injecting Drug Users (IDU) was discussed by Luan et al. (2005). Estimating the number of people eligible for health service was studied by Smith et al. (2002). In their graduation project, Mohammad and Abdullah (2007) compared some Capture-Recapture techniques and cluster sampling for estimating the total number of times the word "Allah" "ﷲ" appears in the Holy Quran. For more details about the estimation of population total and size, see also Gutierrez and Breidt (2009), Otienoet al. (2005), andArnab(2004).

In many situations, the ratio estimator is used to estimate $\tau$ of the variable of interest for a population of size $N$ (unknown). One way to overcome the difficulty of not knowing $N$ is to use a suitable auxiliary variable. Let $O_1,\ldots,O_N$ be the population measurements of the main variable of interest (O) and the corresponding values of an auxiliary variable (V) be $V_1,\ldots,V_N$; the population measurements are $(O_1,V_1),\ldots,(O_n,V_n)$. Assume that there is a fair degree of association between $O$ & $V$. Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be the elements of a SRS, from this population. Now, using the relation

$$\tau_O = \frac{\mu_O}{\mu_V}\tau_V,$$

we can estimate $\tau_O$ by $\hat{\tau}_O = \frac{\bar{Y}}{\bar{X}}\tau_V$.

Let $r = \frac{\bar{Y}}{\bar{X}}$, an estimate of the variance of $\hat{\tau}_O$ is given by:

$$\widehat{Var}(\hat{\tau}_O) = \tau_V^2 \frac{N - n}{N n}\frac{1}{\mu_V^2} S_r^2, \text{ where } S_r^2 = \frac{\sum_{i=1}^{n}(Y_i - rX_i)^2}{n - 1}, \text{ (Scheaffer et al., 1995).}$$

Ahmad et al. (2000) introduced another method to estimate the population total and the population size. They used sequential sampling with replacement until a fixed ($k$) members are repeated.

DOI:https://doi.org/10.5281/zenodo.15910173

In this paper, the estimation of population total (□) utilizing estimators of the population size is considered. In Section 2, we consider the estimation of the population total when $N$ is unknown using Direct Sampling. Two estimators are suggested for □; one is based on Chapman estimator of $N$ and the other is based on a suggested modified estimator of $N$ . In Section 3, □ is estimated using Indirect Sampling. The suggested estimators are compared. Concluding remarks and suggested future works are outlined in Section 4.

**2. Estimation of □When *N* is Unknown Using Capture-Recapture- Direct Sampling**

Assume that we have a closed population with an equal chance of each member to be selected in a random sample. In Direct Sampling, a SRS of $n_1$ elements is drawn, marked(tagged) and the value of the random variable(r.v.) of interest($Y$ ) is noted for each element, the $n_1$ items are released back into the population. After waiting enough period of time so that the marked elements mixed with the remaining population elements(this necessary when the population is a mobile one), a second SRS of $n_2$ elements is drawn. Let $T$ be the number of recaptured elements in the second sample. The values of the variable $Y$ are noted for each of the $n_2$ □$T$ elements. From the first and second sample we obtain a net random sample of size $n$ , where

$n$ □ $n1$ □ $n2$ □$T$. (2.1)

Note that $n$ is a r.v. (not fixed). $T$ has a hypergeometric distribution with probability function:

□$n_1$ □□$N$ □$n_1$ □
□□ □
$f(t)$ □ $P(T$ □ $t)$ □ □ $t$ □□□□□ $n2$ □$t$ □ □□ , $t$□0,1,2,...,min($n1$,$n2$) .
□ $N$ □
□□□ $n_2$ □ □□

Actually, the smallest value of t is max(0,$n_1$ □ $n_2$ □ $N$ ) , but in practice $N$ is large and max(0,$n_1$ □ $n_2$ □ $N$ ) =0. Now, using the properties of hypergeometric distribution we have:

$nn$
$E(n)$ □$E(n_1$ □$n_2$ □$T)$ □$n_1$ □$n_2$ □ 1$N$ 2 (2.2)
$Var(n)$ □$Var(n1$ □$n2$ □$T)$ □$n2$□□$Nn1$ □□□□□□□1□$Nn1$ □□□□□□□$NN$□□$n12$ □□ □
(2.3)

□ **<u>Petersen estimator</u>** of the population size ( $N$ ) is:

$\hat{N}_P$ □ $n1n2$ ,$T$ □ 0,1,...,min($n1$,$n2$). (2.4) $T$

Note that $T$ may equal zero with positive probability; in this case, $\hat{N}_P$ is □. To overcome this difficulty, an alternative estimator of $N$ was proposed by Chapman (1951) as:

$$\hat{N}_C □ \frac{(n_1+1)(n_2+1)}{T+1} □ 1 \quad (2.5)$$

Now,

$E(\hat{N}_C$ )□ $E$□□($n1$ □1)($n2$ □1) □1□□ □ ($n1$ □□□ □1$n$)($N_2n$□ □□□ 2 □1) min(□$t$□$n10$,$n2$)
□□□□□□□□$t$ □1 1□□□□□□□□$nt1$ □□□□□□□□□$Nn2$□□$nt1$ □□□□□□□□□□1

□ $T$ □1 □ □
min($n1$,$n2$) □□□□□□□$nt1$□□11□□$N$ □$n1$ □□ 1□ 2□ □□ □$n1$ □1□□ $N$ □$n1$ □□
□□□□ $n2$ □$t$ □ □□□ □ min($n1$,$n$ 1) □□
□□ □ □ ($n2$ □1) □$t$□0 □□□ □□□□ $nN2$ □□□□ □ □ □ □□ 1□ □$t$□1
□□□□□□□ $t$ □□□□□□□□□□□□$nNn22$□□□□□1□$t$□□□□ □□□□□□□1

DOI:https://doi.org/10.5281/zenodo.15910173

□□$n_1$ □1□□ $N$ □$n_1$ □□
□
□($N$ □1)min($n$□1$t$□□11 ,$n$2□1) □□□□□□□□ $t$ □ □□□□$nN$□□₂ $n$□□211
□□□□1□$t$□□□□□ □□1□ $N$ □ ($n$2 □□1)$nN$2□□ $nN$2□□$n$11 □ (2.6)

□□
Clearly, $\hat{N}_C$ is negatively biased. Also,
□ □$n_1$ □□$N$ □$n_1$ □□
□ □□ □
2 □
$Var(\hat{N}_C)$ □$E$□$\hat{N}_C$□$E(\hat{N}_C)$□ □ min(□$t$□$n$10,$n$2 ) □ □□□□$\hat{N}_C$ □ $E(\hat{N}_C)$□2 □ $t$ □□□□□
□□□□$nNn_2$2□□□□□$t$ □□□ □ □□ □ □ □ (2.7)
Using the two estimators (2.4) and (2.5), we suggest the following new estimator of $N$ :
$\hat{N}_S$□□□ □ $\hat{N}_P$ if $T$□0 □ $\hat{N}_P I(T$ □0)□$\hat{N}_C I(T$ □0) (2.8)
□□ $\hat{N}_C$ if $T$ □0
Now,
$\hat{}$ ) □ $n_1 n_2$ min(□$n$□₁,$n_2$) □□□□1 $t$ □□$P(T$ □ $t$)□ □ □□□($n_1$ □1)($n_2$ □1)□1□$P(T$ □0)
$E(N_S$

$t$ 1 □□
□ □$n_1$ □□$N$ □$n_1$ □□ □$N$ □$n_1$ □ min($n$1,$n$2 ) □ 1 □□ $t$ □□□□ $n$ □$t$ □□□ □□ $n$
□ □
□ $n_1 n_2$ □ □□ □□ 2 □□ □ 2 □ . (2.9)
$t$□1 □□ □ $N$ □ □□□($n_1$ □1)($n_2$ □1)□1□ $N$ □
□□$t$ □□$n_2$ □ □□□$n_2$ □ □□
□□ □□□ □
□□ □ □ □
□ $Var(\hat{N}_S)$ □$E(\hat{N}_S$ □$E(\hat{N}_S))^2$ .
□
But,
2 2

□$\hat{N}_S$□$E(\hat{N}_S)$□ □□□$\hat{N}_P I(T$ □ 0) □$E(\hat{N}_S)I(T$ □ 0)□□□$\hat{N}_C I(T$ □ 0) □$E(\hat{N}_S)I(T$ □ 0) □□
□ □$\hat{N}_p$ $I(T$ □ 0) □$E(\hat{N}_S)I(T$ □ 0)□² □ □$\hat{N}_C$ $I(T$ □ 0) □$E(\hat{N}_S)I(T$ □ 0)□² □□$\hat{N}_p$ □$E(\hat{N}_S$ )□²$I(T$ □ 0) □□$\hat{N}_C$□$E(\hat{N}_S)$□²$I(T$ □ 0).
Thus,
min($n$ ,$n$ ) $n$1$n$2 $E(\hat{N}$ )□2
$Var(\hat{N}_S)$ □ □□1 2 □□ $t$ □ $S$ □□ $P(T$ □ $t$)□(($n$1 □1)($n$2 □1)□1□ $E(\hat{N}_S))$2 $P(T$ □0) $t$ 1 □
□$n_1$ □□$N$ □$n_1$ □ □$N$ □$n_1$ □
□
□ min(□$tn$□11,$n$2) □ □$n$1$tn$2 □ $E(\hat{N}_S)$□□□2 □□□ $t$ □□□□□□□□$Nn$2□□$t$ □□□□(($n$1 □1)($n$2 □1)□1□ $E(\hat{N}_S))$2 □□□ □$nN$2 □ □ □□ (2.10)
□□$n_2$ □□□ □□□$n_2$ □ □□
□ The above results are summarized in the following lemma:

**Lemma (2.1)**

**Journal of Statistical and Mathematical Sciences**
Vol. 13 No. 1 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15910173

The expected value and variance of the suggested estimator, $\hat{N}_S$, of the population size $N$ are given by

□ □$n_1$ □□$N$ □$n_1$ □□ □$N$ □$n_1$ □

min($n$ ,$n$ ) $E$□$\hat{N}$ $S$ □□ $n_1n_2$ □ □1 2 □□□ □□□□□ 1$t$ □□□□□□□ $t$ □□□□□ □□□$Nn_2$□□$t$ □□□□□□□□□□(($n_1$ □1)($n_2$ □1) □1□□□□□ □□□$nnN$22 □ □ □□ □□ □,

$t$ 1 □ □ □□ □

□$n_2$ □ □ □

□$n_1$ □□$N$ □$n_1$ □ □$N$ □$n_1$ □

min($n$ ,$n$ ) $Var$□$\hat{N}$ $S$ □□ □□1 2 □□$n_1tn_2$ □ $E(\hat{N}_S)$□□□2 □□□ $t$ □□□□□□□□$Nn_2$□□$t$ □□□ □(($n_1$ □1)($n_2$ □1)□1□ $E(\hat{N}_S))^2$ □□□ □$nN_2$ □ □ □□ .

$t$ 1 □

□□$n_2$ □□□ □□□$n_2$ □ □□

□

Given $n$ , let $Y_1,Y_2,...,Y_n$ be the values of the variable $Y$ for the sample elements. The suggested estimators of the population total (□) are:

$\hat{□}_C$ □$\hat{N}_C\bar{Y}$, (2.11)

and

$\hat{□}_S$ □$\hat{N}_S\bar{Y}$, (2.12)

**We conjecture here that given $n$ , $Y_1,Y_2,...,Y_n$ is a SRS from the population. We have not been able to prove this conjecture yet.**

Now,

$E(\hat{□}_S)$ □ $E(E(\hat{□}_S|n))$ □ $E(E(\hat{N}_S\bar{Y} \mid n))$ □ $E(\hat{N}_S E(\bar{Y} \mid n))$ □ $E(\hat{N}_S)$□

□□ □ □$n_1$ □□$N$ □ $n_1$ □□ □$N$ □ $n_1$ □□ □

□ □ □□□□ □□

□ $t$

$n_2$ □□ □ □□ $n_2$ □$t$ □□□ □□ □□ min($n$□$_1$,$n^2$) □□□□1$t$ □□□ □□ $N$ □ □□□($n_1$ □1)($n_2$ □1)□1□□ $n_2$ □□□□ (2.13)

□□□□□$n_1$ □$t$ □□ □$n_2$ □□□ □□□ □□□□$nN_2$□□□□ □□ □ □

□

Similarly, for $\hat{□}_C$ we have

□□ ($n_2$ □1)□$nN_2$□□$n_{11}$ □□ □

$E(\hat{□}_C)$ □ $E(\hat{N}_C)$□□□□$N$□ □$N_2$ □ □

□ $n$ □

DOI:https://doi.org/10.5281/zenodo.15910173

□*N*□*n*
□($n_2$□1) □ $^{n2\,□1}$ 1 □ □ □ (2.14)
□□ □ □
□*N*□
□$n_2$□ □ □

Also,

*Var*(□^*s*) □ *E*(*Var*(□^*s* |*n*)) □*Var*(*E*(□^*s* |*n*)) □ *E*(*Var*(*N*^ *sY* | *n*)) □*Var*(*E*(*N*^ *sY* | *n*))
□ *E*□*N*^ $s^2$*Var*(*Y* | *n*)□□*Var*(*N*^ *s*□),

but

*Var*(*Y* | *n*) □□2 □□*N* □*n*□ □ ,
*n* □ *N* □1 □

therefore, □ □ ^ 2□2 *N* □□*n*□□□□ □□ 2*Var*(*N*^ *S* ) □ *N*□□2 1 *E*□□□□*N*^
*S*2□□□*Nn* □1□□□□□□□□□□2*Var*(*N*^ *S* )
*Var*(□^*s*) □ *E*□□*Ns* *n* *N* 1
□2 *E*□□*N*^*S* 2 *N* □*N*^*S*2 □□□□2*Var*(*N*^*S*).
□
*N* □1 □ *n* □

Now,

2 *E*□□ ^*S*2 *N*□□□ □ *NE*□□□□*N*^*nS* □ □ □□
□*N* *n*
□ □$n_1$□□*N* □$n_1$□□ □*N* □$n_1$□
min(*n* ,*n* □ *Nn*2*n*2 1 2 ) □ 1 □□□ *t* □□□□□□ *n*2 □*t* □□□□□□ *N* □(*n*1 □1)(*n*2
□1)□1□2 □□□ *n*2 □□□ .
□

1 2 □*t*□1 □□□□ *t*2(*n*1 □*n*2 □*t*) □□□□*nN*2 □□□□ □□□□ *n*1 □*n*2 □
□□□*nN*2 □ □ □□

Thus,

□ □ □$n_1$□□*N* □$n_1$□□ □*N* □$n_1$□
1
□ □ □ *t* □□□□ *n* □*t* □□□□□ *N* □(*n* □1)(*n* □1) □1□ □□ *n* □□ □ *Var*(*N*^ )
□□*E* (*N*^ )□ □
□ □

Journal of Statistical and Mathematical Sciences
Vol. 13 No. 1 | Imp. Factor: 8.99
DOI:https://doi.org/10.5281/zenodo.15910173

Var (□^S ) □ N□□2 1 □□Nn12n22 min(tn□11,n2) □□ t 2(n1 □ n2 □t ) □N 2□ □ 1 n1 □2 n2 2 □N2 □ □ S S 2 □□□□□2Var (N^S ).

□□ □ □$n_2$□□ □□ □□$n_2$□□
□ □ □

(2.15)

Similarly for □^C, we have

□ □
□ □
□2 □ min($n_1$,$n_2$) □ N^ 2
Var(□^C) □ N □1□□N □t□o □□
n

□□□$nt_1$ □□□□□□□□□Nn2□□$nt_1$ □□□□□□□□□□
□□Var(N^ C ) □□E(N^ C )□2 □□□□ □ □□□2Var(N^ C ).
(2.16)
□
□ □□□ $nN_2$□□□□ □□□ □ □□ □ □

□ □
□ □

The above results are given in the following lemma:

**Lemma (2.2)**

The expected value and the variance of the estimators of the population total are given by

E(□^S) □□E(N^ S), E(□^C)□□E(N^C)

Var (□^ ) □ N□ 1□□□Nn² n² □□1 2 □□□ □□$n_1$ □□□□□□□Nn2□□$nt_1$ □□□□□□□□ N □(n1 □1)(n2 □1)□1□2 □□□□□Nn□2 n 1 □ □□□ □ 2 min(n ,n ) 1 □□ t □□
S □□ 1 2 □□t2(n1 □ n2 □t)□□nN □ □□ n1 □ n2 □□□□ nN2 □□
□ □
□ □ t 1 □ □ □□ □
□
□ □ 2 □ □
□Var (N^ S)□□E(N^ S)□² □□□□²Var (N^ S)
N□2 1 □□□ min(□n1,n2) □□□ N^ 2 □□ □□$nt_1$ □□□□□□□□□Nn2□□$nt_1$ □□□□□□□
□□Var(N^ C )□□E(N^ C )□2 □□□□□ □ □□2Var(N^ C ).
Var(□^C ) □ □N □ □ n □N □ □□□□ □□

□ □□ t o □□ □ □□ □ □
□ □ □□ n2 □ □ □

Now, if N is known, then □ can be estimated based on SRS of size n □ $n_1$ □$n_2$ □T by

□^ □NY, (2.17)

with

E(□^) □E(E(□^ | n)) □N□□□.

Thus, □^ is an unbiased estimator of □.

Var(□^)□Var(NY) □ E(Var(NY | n))□Var(E(NY | n))
□ E□□□□N2□n2 □□□ NN□□1n□□□□□□□□ □Var□N□□□ NN2□□12 □□□□NE□□□ 1n□□□□1□□□□

□ □$n_1$ □□$N$ □ $n_1$ □ □
$N$ 2□2 □ min($n$ ,$n$ ) □ 1 □□
□ $N$ □1 □□□ $N$ □$t$□10 2 □□ □□□ $t$ □□□□□□ $n2$ □$t$ □□□ □ 1□ □ □ (2.18)

□ □ $n1$ □ $n2$ □$t$ □□ □□□□ $nN2$ □□□□ □ □ □
□ $K$□²,
where,
□ □$n_1$ □□$N$ □ $n_1$ □ □
□ □
$K$ □ $NN$□21□□□ □□$N$ min(□$t$□$n$10,$n$2) □□□$n1$ □1$n2$ □$t$ □ □□□□□ $t$ □□□□□□□
□□□□$nNn$2₂□□□□□$t$ □□□ □1□ □ □ □□□ .

□
The efficiency of □ˆ$S$ with respect to□ˆ is
$Eff$ (□ˆ$s$;□ˆ) □ $MSE$( □ˆ) ,

$MSE$(□ˆ$s$)
where,
$MSE$(□ˆ) □$Var$(□ˆ), $MSE$(□ˆ$s$) □□$bias$($N$ˆ$s$)□²□² □$Var$(□ˆ$s$).
Now, let $Var$(□ˆ$S$ ) □ $L$□2 □$Var$($N$ˆ $S$ )□2,where
□ 2 min($n$ ,$n$ ) $L$ □ □□$Nn$ $t$□11 2 □□$t$ 2 ($n1$ □ $n2$ □$t$) □ $t$)□□□□ $N$
1 □1$n$)(1$n$□2 $n$□21)□1□2 $P$($T$ □ o)□□□
1 1 $n22$ □ □□1 $P$($T$ □ □($n$

$N$ □1□□□□$Var$($N$ˆ $s$ )□□$E$($N$ˆ $s$ )□2 □ □ □ □
□
then,
$Var$(□ˆ) $K$□2

$Eff$ (□ˆ$S$;□ˆ) □ $MSE$($N$ ˆ )□2 □ $L$□2 □$MSE$($N$ˆ$S$)□2 □$L$□2 .
$S$
The efficiency can be rewritten in terms of the coefficient of variation (CV) given by $CV$ □□ , (□□ o ) as

□
$K$ 2
□$CV$($y$)□
$MSE$($N$ˆ$s$) (2.19)
$Eff$ (□ˆ$s$;□ˆ) □
$L$ 2

DOI:https://doi.org/10.5281/zenodo.15910173

1□ □$CV(y)$□
$MSE(\hat{N}_S)$
The efficiency of □ˆ$_S$ and □ˆ$_C$ w.r.t. □ˆ are given in Tables (2.1) and (2.2) respectively. Also, the efficiency
of □ˆ$_S$ w.r.t. □ˆ$_C$ is given in Table (2.3). Based on these tables, we can see that □ˆ$_S$ is more efficient than □ˆ$_C$ for small expected sample size. □ˆ$_S$ &□ˆ$_C$ are more efficient than □ˆ when $E(n)$ is small and for large $CV$.

**3. Estimation of □When $N$ is Unknown Using Capture-Recapture- Indirect Sampling**

Indirect sampling is another Capture -Recapture method for estimating □ . In this method, sampling continues until a fixed number ($T$) of recaptured elements are obtained. So, a random sample of size $n_1$ is chosen, marked and released. Later, we select elements randomly from the population until $T$ elements are being recaptured. Let $n_2$ be the total number of elements selected in the second random sample to obtain $T$ previously captured elements, then from the first and the second sample we obtain a random sample of size $n$ elements, where

$n$ □ $n_1$ □$n_2$ □$T$. (3.1)

Here, $n_2$ is a random variable (not fixed); it has the negative hypergeometric distribution with probability function given by:

□$n_2$ □1□□$N$ □ $n_2$ □
□□ $T$ □1□□□□□□ $n1$ □$T$ □ □□ , $n$
□
$f(n2)$ □ 2 □$T$,$T$ □1,...,$N$ □ $n1$ □$T$. (3.2)
□$N$ □
□□□ $n_1$□ □□

Also, $E(n_2)$ □$T(N$ □1) (Balakrishnan 2003) $n_1$ □1

and $Var(n_2)$ □$T(N$ □(1$n$)($N$1□) 2$n$(1$n$)(1 $n$ □12□)1□$T$) (Khan, 1994).

1 □
Now,
$E(n)$ □$E(n1$ □$n2$ □$T$) □$n1$ □$T$□$T(nN$ 1 □□1 1) □$n1$ □$T$□□□$Nn1$□□$n11$ □□□□ (3.3)

□
Also,
$Var(n)$ □$Var(n_2)$ □ $T(N$ □(1$n$)(1 □$N$1□)2$n$( 1$n$)(1 $n$□12□) 1□$T$) (3.4)

Table (3.1) contains the expected value of the sample size for different values of $n_1,T,N$. It can be seen that $E(n)$ is increasing in $T$ for fixed $n_1$ and decreasing in $n_1$ for fixed $T$.

The estimator of the population size $N$ is $N_I$ˆ □ $n_1 T n_2$, 1□$T$□$n_1$,$T$ is integer. (3.5)

Now,
$E(\hat{N}_I)$ □ $E$□□$n_1 n_2$□□ □ $n_1(N$ □1) (3.6)

□ $T$ □ $n_1$ □1

Clearly, $E(\hat{N}_I)$ does not depend on $T$ and increasing in $n_1$, $\hat{N}_I$ is negatively biased. Now,
$Var(\hat{N}_I)$ □$Var$□□$n_1 n_2$ □□ □ $T$ $n$ 12 2 $Var(n2)$

□ T □
$n_1^2 (N□1)(N□n_1)(n_1 □1□T)$

□ T $(n_1 □1$ $2(n_1 □ 2)$ . (3.7)
)
Hence,
1 $MSE(\hat{N}) □ (n_1 □$ 2 □□□$n_T{}^{12} (N □1)(N(_n□_1n□^1)(_2)n^1 □1□T)$ □□$n_1 □N□^2$ □ □□ □ .
(3.8)
$I$

1) □
It can be seen from (3.6) that this estimator of $N$ can be corrected to be unbiased estimator of $N$ as follows:
$E(\hat{N}_I)$ □ $n1(N$ □1) , $n_1 □1$

which gives
$E(\hat{N}_I n^{1□1})$ □ $N$ □1.

$n_1$
So, the estimator
$\hat{N}_I$ □□□$n_1n□_1$ 1□□□□ □1 □ $n_2(nT_1 □1)$ □1

□
is an unbiased estimator of $N$ . Let
$\hat{N}_I^* □ n^2(n^1 □1)$ □1, (3.9) $T$

then
$E(\hat{N}_I^*) □ N$ .
Therefore, the mean square error of $\hat{N}_I^*$ is
(
$MSE(\hat{N} I^*) □Var(\hat{N} I^*) □ n\ 1T □21)2$ $Var(n2) □ (N □1)(NT(□n1n1□)(2n)1 □1□T)$ . (3.10)

2
Clearly,$Var(\hat{N} I)$ □□□□$n1n$ □1 1□□□□ $Var(\hat{N} I^*)$,thus, for any values of $N$,$n1$ and $T$ ,$Var(\hat{N} I)$□$Var(\hat{N} I^*)$ , however,

□
the $MSE(\hat{N}I)$ is not necessary less than $MSE(\hat{N}I^*)$
The efficiency of $\hat{N}_I$ w.r.t. $\hat{N}_I^*$ is:
$Eff (NI$
$\hat{}\ ^*, \hat{N} I) □ MSEMSE((NN\hat{}\hat{}_I I^*))$ .

DOI:https://doi.org/10.5281/zenodo.15910173

Table (3.2) contains some numerical values of the efficiency of $\hat{N}_I$ w.r.t. $\hat{N}_I^*$ . Clearly, the two estimators almost have the same performance.
Given $T$, $Y_1, Y_2, \ldots, Y_n$ are the values of variable for the sample element. Two estimators of the population total, □, are: $□\hat{}_I □ \hat{N}_I Y$ and $□\hat{}_I^* □ \hat{N}_I^* Y$ ,
Now,
$E□□\hat{}_I$ □□ $E□E□□\hat{}m_2□□□$ | $_E□_E□\hat{N}_I Y n_2$□|□□$^E$□$\hat{N}_I$$^E$□|$Y n_2$□□
$n$ $(N$
□□$E□\hat{N}_I$ □ □ 1 □1) □.
$n_1$□1

Also, $E(□\hat{}_I^*)$ □ $N$□□□. The variance of □$\hat{}_I$ and □$\hat{}_I^*$ can be derived as follow
$Var□□\hat{}_I$ □□ $E$|□$Var$□□$\hat{}_I$ | $n_2$□□□$Var$□|$E$□□$\hat{}m_2$□□□ $^E$□$^{Var}$□$\hat{N}$ $_I Y$
$n_2$□□□$^{Var}$□□$\hat{N}_I$□_|
□$E$ □$\hat{N}_I^2 Var$□$Y n_2$□|□□□$^2 Var$□$\hat{N}_I$□□$E$ $(\hat{N}_I^2 Var(Y \mid n_2))$ □□$^2 Var$□$\hat{N}_I$□
□$n_1 2 n_2 2$ □2 $N$ □ $(n1$ □$n2$ □$T)$ □□□□$2Var$□$\hat{N}$ $I$ □
□ $E$□□ $T2$ ~~$n1$ □$n2$ □$T$ $N$ □1 □□~~
□
$n12$□2 $E$□□$n22(N$ □$(n1$ □$n2$ □$T))$ □□
□ $T$ $2(N$ □1) □□ $n1$ □$n2$ □$T$ □ □□□$2Var$□$\hat{N}$ $I$ □
□ □$n_2$ □1□□$N$ □$n_2$ □□
2 2 $N$□$n1$□$T$ □$n22(N$ □$(n1$ □$n2$ □$T))$ □□□$T$ □1□□□□□□□$n1$ □$T$ □□□□ □
$2Var$□$\hat{N}$ □ $n1$□ □

□ $T(N$ □1) $_n$□2□$_T$□□ $n1$ □$n2$ □$T$ □□□ □$nN$ $_1$□□□□ □ □□ □□□ $_I$
□
□
□$B$□$^2$ □□$^2 Var(\hat{N}_I)$, (3.11)

where,
□ □$n_2$ □1□□$N$ □$n_2$ □□
□ □ □□ □□

$B$ □ $n12$ $N$□□$2n$□1□$T$□□$n22(Nn$□1 □$(nn12$□□$nT2$ □$T))$ □□$T$ □1□□□$N$□□$n$□1 □$T$
□ □□ □□ . (3.12)
$T(N$ □1) $n$ $T$ □□ □□□□ $n1$□□□ □ □
Now,
$Var(□\hat{}_I^*)$ □$E(Var(□\hat{}_I^* n_2))$□$Var(E(□\hat{}_I^* n_2))$ □$E(Var(\hat{N}_I^* Y n_2))$ □$Var(E(\hat{N}_I^* Y$
$n_2))$
□
□$E(N$ $\hat{}^* 2Var(Y$ $n2))$□□$2Var(\hat{N}$ $I^*)$ □ $E$□□□□□□$n2(nT1$ □1) □1□□□2 $n1$ □□$n22$ □$T$ $N$
□□$nN1$ □□$n12$ □$T$□□□ □ □
$I$
□
□ $(N$□1$)(N$□$n1)(n$ 1 □1□$T)$□2

DOI:https://doi.org/10.5281/zenodo.15910173

T(n₁ □ 2)
□² □ ₂□ N □□
□T ₂(N □1) E□□□□n₂(n₁ □1)□T□ □□□n₁ □n₂ □T □1□□□□□□

(N□1)(N□n )(n (3.18)
□ 1 1 □1□T)□ 2
T(n₁ □2)
□ Z□² □ J□ ²,
where

Z □T 2(N1 □1) E((n2(n1 □1) □T)2(n1 □Nn2 □ T □1)) & J □ (N□1)(NT(□n1n1□)(2n)1 □1□T) .

Note that, the previous derivations depend on the fact that $n_2$ has a negative hypergeometric distribution.

Now, if $N$ is known, then □ can be estimated based on a simple random sample of size $n$ □ $n_1$ □$n_2$ □$t$ by □^□$NY$ .

2 N □n□□ Var□N□□
Var(□^) □Var(NY) □ E(Var(NY | n))□Var(E(NY | n)) □ E□□□N ² □n □□□ N □1□□□□□ □

□
N2□2 □N □ □ E□ □1□ N □1 □ n □

□ N 2□2 E□□ N □ □□ N 2□2 □ N3□2 E□□1□□□ N2□2

N □1 □ n □ N □1 N □1 □n□ N □1
□ □n₂ □1□□ N □ n₂ □□
1 □
□ NN3□□12 Nn□□2n□1T□T □□□ □ n1 □ n2 □ t □□ T □1□□□□N□□□ n□1 □T
□□□□□□ □□ NN2□□1 2 □H□2,

□□ □
□□ □n1 □□ □ □
where,
□ □ □n2 □1□□N □n
N □ □ 1 □ T □1□□□□□□ n 2 □□ □ H □ N □21□□NN□□₂n□1T □T □□n₁ □ n₂ □t □□ □N □¹ □T
□□□□□□□ □1□ □□ □ .

□ n □□ □ □□ □ □ □□ □ □□ n₁□ □ □

DOI:https://doi.org/10.5281/zenodo.15910173

The efficiency of □ˆ$_I$ with respect to □ˆ (obtained for a sample size equal the expected sample size) is:
*Eff* (□ˆ$_I$;□ˆ) □ *MSE*( □ˆ) ,

*MSE*(□ˆ$_I$)
*MSE*(□ˆ$_I$) □□*bias*(□ˆ$_I$)□² □*Var*(□ˆ$_I$) □□*bias*($\hat{N}_I$)□□² □ *B*□² □*Var*($\hat{N}_I$)□²
□*B*□² □*MSE*($\hat{N}_I$)□² ,
*Eff* (□ˆ$_I$;□ˆ) □ *B*□2 □ *MSEH*□ 2 ($\hat{N}_I$)□2 .

The efficiency can be rewritten in terms of the coefficient of variation (*CV*), given by *CV* □□ , (assume
□
□□ 0 ), as:
*H* 2
□*CV*(*y*)□ .
*MSE*($\hat{N}_I$)
*Eff* (□ˆ$_I$;□ˆ) □
*B* 2
1□ □*CV*(*y*)□
*MSE*($\hat{N}_I$)
Some values of the efficiency of □ˆ$_I$ w.r.t. □ˆ are given in Table (3.3). Similarly,
*Eff* (□ˆ$_I^*$;□ˆ$_I$) □ *B*□*CV*(*y*)□² □ *MSE*2 ($\hat{N}_I$).

*Z*□*CV*(*y*)□ □ *J*
The efficiency of □ˆ$_I^*$ w.r.t. □ˆ$_I$ is given in Table (3.4). Based on the tables, we have the following conclusions:

1. From Table (3.3), the efficiency of □ˆ$_I$ w.r.t. □ˆ increases when the CV increases.
2. For small sample size, the efficiency of □ˆ$_I$ w.r.t. □ˆ$_S$ is greater than or close to one. Also, the efficiency of □ˆ$_I$ w.r.t. □ˆ$_C$ is greater than one for small sample size and small value of the CV.
3. For large sample size, the efficiency of □ˆ$_I$ w.r.t. □ˆ$_S$ is less than one.
4. For large value of the coefficient of variation, the efficiency of □ˆ$_I$ w.r.t. □ˆ$_C$ is less than one, and it decreases when the sample size increases.
5. □ˆ$_I^*$ is more efficient than □ˆ$_I$ for large value of the expected sample size and CV.

## 4. Conclusions and Suggestions for Further Research

In this paper, four different estimators for the population total are discussed and from the results obtained we can conclude the following:

1. For large expected sample size, when we use Direct Sampling we found that the estimator of the population total □ˆ$_C$ based on Chapman estimator $\hat{N}_C$ is better than the estimator □ˆ$_S$ based on the suggested modified estimator $\hat{N}_S$. On the other hand, if the expected sample size is small then □ˆ$_S$ is more efficient than □ˆ$_C$.
2. For small expected sample size and CV, we found that it is better to use Indirect Sampling to estimate □ than using Direct Sampling.
3. The bias in $\hat{N}_I$ can be corrected to obtain an unbiased estimator of $N$, $\hat{N}_I^*$, and also an unbiased estimator of □.

## Suggestions for Future Work

- If $N$ is known, then an estimator of $N$ (pretending it is unknown) can be used as a guard against unsuitable or insufficient sample. So, we may suggest estimator of □ conditioning on $\hat{N}$ to be between $N$□□ and $N$□□ for some □.
- Estimation of □ based on other sampling techniques when $N$ is unknown can also be considered next.

## References

Ahmad, M., Alalouf, S.and Chaubey, Y. (2000). Estimation of the population total when the population size is unknown. Statistics and Probability Letters 49, 211-216.

Arnab,R. (2004). Optimum estimation of a finite population total in PPS sampling with replacement for multicharacter surveys. Journal of the Indian Society of Agricultural Statistics58,231-243.

Azevedo-Silva, F., Reis, R., Santos, M., Luiz, R.and Pombo-de-oliveira, M. (2009). Evaluation of childhood acute leukemia incidence and underreporting in Brazil by capture-recapture methodology. Cancer Epidemiology 33, 403-405.

Balakrishnan, N., Charalambides, C.and Papadatos, N. (2003). Bounds on expectation of order statistics from a finite population. Journal of Statistical Planning and Inference 113, 569 – 588.

Chapman, D. (1951). Some properties of the Hyper-geometric distribution with application to zoological censuses. University of California Publication in Statistics.

Gutierrez, H.and Breidt, J. (2009). Estimation of the population total using the generalized difference estimator and Wilcoxon ranks. RevistaColombianadeEstadistica32,123-143.

Khan, R. A.(1994). A Note on the generating function of a negative hypergeometric distribution. Sankhya,56,309-313.

Luan, R., Liang, B., Yuan, P.,Fan, L., Huang, Y., Zeng, G., Wang, L., and Wang, S. (2005). A study on the Capture-Recapture method for estimating the population size of injecting drug users in Southwest China. Journal of Health Science 51, 405-409.

Mohammad, M.and Abdullah, M. (2007). Undergraduate graduation project. Qatar University.

Otieno, R., Mwita, P., Obudho, G. and Wafula, C. (2005).Model-based estimation of finite population total in stratified sampling. East African Journal of Stat.1, 23-40.

Scheaffer, Mendenhall and Ott (1995). Elementary survey sampling. Fifth edition, Duxbury Press.

Sekar, C. and Deming, W. (1949). On a method of estimation birth and death rates and the extent of registration. JASA 44, 101-115.

Smit, F., Reinking, D. and Reijerse, M. (2002). Estimating the number of people eligible for health use. Evaluation and Program Planning 25, 101-105.

DOI:https://doi.org/10.5281/zenodo.15910173

**Table(2.1):** $Eff(\hat{N}_S; \hat{N}_C)$

| $N$ | $E(n)$ | $MSE(\hat{N}_s)$ | $MSE(\hat{N}_c)$ | $Eff(\hat{N}_s; \hat{N}_c)$ |
|---|---|---|---|---|
| 1000 | 51.9 | 728205.9 | 732439.8804 | 1.005814 |
| 1000 | 57.6 | 332593.5929 | 402913.9604 | 1.21143 |
| 1000 | 61.4 | 177120.7456 | 289004.9081 | 1.631683 |
| 1000 | 67.1 | 84362.2724 | 213053.84 | 2.525464 |
| 1000 | 72.8 | 117772.4201 | 201699.6164 | 1.712622 |
| 1000 | 73.8 | 132095.56 | 203386.25 | 1.53969 |
| 1000 | 80.4 | 275199.96 | 229010.9225 | 0.832162 |
| 1000 | 92.8 | 598569.24 | 286622.6784 | 0.478846 |
| 1000 | 116.5 | 918583.24 | 311101.1376 | 0.338675 |
| 5000 | 227.6 | 16746267.56 | 7491881 | 0.447376 |
| 5000 | 237.3 | 19768303.04 | 7903480.49 | 0.399806 |
| 5000 | 247.0 | 22055514.24 | 8134961 | 0.36884 |
| 5000 | 252.82 | 23062512.25 | 8193942.25 | 0.355293 |
| 5000 | 253.79 | 23204389.44 | 8198340.81 | 0.35331 |
| 5000 | 266.4 | 24406592.01 | 8135765.61 | 0.333343 |
| 5000 | 290.65 | 23995928.16 | 7566230.24 | 0.315313 |
| 5000 | 314.9 | 21416904.64 | 6725680.81 | 0.314036 |
| 5000 | 344.0 | 17364010.41 | 5689186.49 | 0.327642 |

**Table (2.2):** $Eff(\square\hat{}C; \square\hat{})$

| $N$ | $E(n)$ | $CV\square$ 0.5 | $CV\square 1$ | $CV\square$ 5 | $CV\square 10$ |
|---|---|---|---|---|---|
| 1000 | 51.9 | 0.006241 | 0.024954 | 0.615966 | 2.370284 |
| 1000 | 57.6 | 0.010147 | 0.040398 | 0.878713 | 2.499718 |
| 1000 | 61.4 | 0.01319 | 0.052189 | 0.968865 | 2.147734 |
| 1000 | 67.1 | 0.016208 | 0.063373 | 0.920717 | 1.595043 |
| 1000 | 72.8 | 0.015642 | 0.060687 | 0.773051 | 1.220907 |
| 1000 | 73.75 | 0.015294 | 0.059299 | 0.748042 | 1.17425 |
| 1000 | 80.4 | 0.012367 | 0.04792 | 0.598649 | 0.93414 |
| 1000 | 92.75 | 0.008463 | 0.032934 | 0.440701 | 0.718821 |
| 1000 | 116.5 | 0.006056 | 0.023686 | 0.346428 | 0.603322 |
| 5000 | 227.6 | 0.003486 | 0.01378 | 0.250864 | 0.542584 |
| 5000 | 237.3 | 0.003163 | 0.012512 | 0.230536 | 0.506165 |
| 5000 | 247.0 | 0.002947 | 0.011661 | 0.216907 | 0.482064 |
| 5000 | 252.82 | 0.002855 | 0.011299 | 0.211125 | 0.471948 |
| 5000 | 253.79 | 0.002842 | 0.011248 | 0.21031 | 0.470536 |
| 5000 | 266.4 | 0.002721 | 0.010773 | 0.202774 | 0.457653 |
| 5000 | 290.65 | 0.002668 | 0.010566 | 0.199714 | 0.453306 |
| 5000 | 314.9 | 0.002756 | 0.010912 | 0.205493 | 0.464102 |
| 5000 | 344.0 | 0.002964 | 0.011728 | 0.218439 | 0.486251 |

**Table (2.3):** $Eff(\square\hat{}S; \square\hat{})$

DOI:https://doi.org/10.5281/zenodo.15910173

| *N* | *E*(*n*) | *CV*□ 0.5 | *CV*□1 | *CV*□ 5 | *CV*□10 |
|---|---|---|---|---|---|
| 1000 | 51.9 | 0.006277 | 0.025099 | 0.619341 | 2.380999 |
| 1000 | 57.6 | 0.012284 | 0.0488 | 1.002176 | 2.573034 |
| 1000 | 61.4 | 0.021431 | 0.083741 | 1.20335 | 2.066933 |
| 1000 | 67.1 | 0.040074 | 0.147658 | 1.047878 | 1.294508 |
| 1000 | 72.8 | 0.026292 | 0.096826 | 0.684161 | 0.844183 |
| 1000 | 73.75 | 0.023162 | 0.08576 | 0.634498 | 0.793077 |
| 1000 | 80.4 | 0.010235 | 0.039044 | 0.393256 | 0.548862 |
| 1000 | 92.75 | 0.004054 | 0.015796 | 0.215683 | 0.356763 |
| 1000 | 116.5 | 0.002056 | 0.008094 | 0.134857 | 0.264126 |
| 5000 | 227.6 | 0.00156 | 0.006172 | 0.114769 | 0.254933 |
| 5000 | 237.3 | 0.001265 | 0.005014 | 0.096394 | 0.223911 |
| 5000 | 247.0 | 0.001088 | 0.004314 | 0.084914 | 0.204024 |
| 5000 | 252.82 | 0.001015 | 0.004028 | 0.080156 | 0.195766 |
| 5000 | 253.79 | 0.001005 | 0.003988 | 0.079481 | 0.194601 |
| 5000 | 266.4 | 0.000908 | 0.003606 | 0.073122 | 0.183947 |
| 5000 | 290.65 | 0.000842 | 0.003348 | 0.069215 | 0.179724 |
| 5000 | 314.9 | 0.000867 | 0.003445 | 0.071808 | 0.188995 |
| 5000 | 344.0 | 0.000973 | 0.003866 | 0.080548 | 0.211892 |

**Table (2.4):***Eff* (□ˆ*s*;□ˆ*c*)

| *N* | *E*(*n*) | *CV* □ 0.5 | *CV* □1 | *CV* □ 5 | *CV* □10 |
|---|---|---|---|---|---|
| 1000 | 51.9 | 1.005811 | 1.005801 | 1.005478 | 1.004521 |
| 1000 | 57.6 | 1.210561 | 1.207977 | 1.140504 | 1.02933 |
| 1000 | 61.4 | 1.624743 | 1.604566 | 1.24202 | 0.962378 |
| 1000 | 67.1 | 2.472407 | 2.329969 | 1.138111 | 0.811582 |
| 1000 | 72.8 | 1.680817 | 1.595495 | 0.885014 | 0.69144 |
| 1000 | 73.75 | 1.514448 | 1.446229 | 0.848211 | 0.675391 |
| 1000 | 80.4 | 0.827601 | 0.814762 | 0.656905 | 0.587559 |
| 1000 | 92.75 | 0.479045 | 0.47962 | 0.489408 | 0.496317 |
| 1000 | 116.5 | 0.339446 | 0.341712 | 0.389279 | 0.437785 |
| 5000 | 227.6 | 0.447514 | 0.44792 | 0.457494 | 0.46985 |
| 5000 | 237.3 | 0.400046 | 0.400759 | 0.418129 | 0.442368 |
| 5000 | 247.0 | 0.36913 | 0.36999 | 0.391477 | 0.42323 |
| 5000 | 252.82 | 0.355601 | 0.356517 | 0.379659 | 0.414804 |
| 5000 | 253.79 | 0.353621 | 0.354545 | 0.377924 | 0.413574 |
| 5000 | 266.4 | 0.333682 | 0.334688 | 0.36061 | 0.401936 |
| 5000 | 290.65 | 0.315694 | 0.316825 | 0.34657 | 0.396475 |
| 5000 | 314.9 | 0.314463 | 0.315735 | 0.349444 | 0.407227 |
| 5000 | 344.0 | 0.328139 | 0.329615 | 0.368744 | 0.435766 |

**Table (3.1):** Expectation of $n$

DOI:https://doi.org/10.5281/zenodo.15910173

| $N$ | $T$ | $n_1$ | $E(n)$ |
|---|---|---|---|
| 1000 | 1 | 50 | 68.627 |
| 1000 | 3 | 50 | 105.88 |
| 1000 | 4 | 50 | 124.51 |
| 1000 | 5 | 100 | 144.55 |
| 1000 | 10 | 100 | 189.11 |
| 1000 | 15 | 100 | 233.66 |
| 1000 | 25 | 250 | 324.7 |
| 1000 | 63 | 250 | 438.25 |
| 1000 | 75 | 250 | 474.1 |
| 5000 | 1 | 100 | 148.51 |
| 5000 | 2 | 100 | 197.03 |
| 5000 | 8 | 400 | 491.77 |
| 5000 | 32 | 400 | 767.08 |
| 5000 | 63 | 400 | 812.97 |
| 5000 | 50 | 500 | 949.1 |
| 5000 | 65 | 500 | 1083.8 |

**Table(3.2):** $Eff\,(\hat{N}_I\,;\,\hat{N}^*_I)$

| $N$ | $E(n)$ | $Eff\,(\hat{N}_I, \hat{N}_I^*)$ |
|---|---|---|

| | | |
|---|---|---|
| 1000 | 68.627 | 0.96155 |
| 1000 | 105.88 | 0.962362 |
| 1000 | 124.51 | 0.962765 |
| 1000 | 144.55 | 0.980766 |
| 1000 | 189.11 | 0.981287 |
| 1000 | 233.66 | 0.981863 |
| 1000 | 324.7 | 0.992361 |
| 1000 | 438.25 | 0.993052 |
| 1000 | 474.1 | 0.993324 |
| 5000 | 148.51 | 0.980392 |
| 5000 | 197.03 | 0.980514 |
| 5000 | 491.77 | 0.995068 |
| 5000 | 767.08 | 0.995215 |
| 5000 | 812.97 | 0.995449 |
| 5000 | 949.1 | 0.99621 |
| 5000 | 1083.8 | 0.996266 |

**Table(3.3):** $Eff(\hat{\square}_I;\hat{\square})$

| $N$ | $E(n)$ | $CV$ □ 0.5 | $CV$ □1 | $CV$ □ 5 | $CV$□10 |
|---|---|---|---|---|---|
| 1000 | 68.627 | 0.004084 | 0.016098 | 0.275107 | 0.553307 |
| 1000 | 105.88 | 0.008043 | 0.030201 | 0.254925 | 0.332162 |
| 1000 | 124.51 | 0.009028 | 0.032967 | 0.217609 | 0.263777 |
| 1000 | 144.55 | 0.0087 | 0.030742 | 0.162453 | 0.187566 |
| 1000 | 189.11 | 0.012318 | 0.036371 | 0.096949 | 0.102273 |
| 1000 | 233.66 | 0.013651 | 0.034478 | 0.067367 | 0.069437 |
| 1000 | 324.7 | 0.013146 | 0.026527 | 0.039341 | 0.039944 |
| 1000 | 438.25 | 0.011182 | 0.014516 | 0.016046 | 0.016099 |
| 1000 | 474.1 | 0.01013 | 0.012492 | 0.013499 | 0.013533 |
| 5000 | 148.51 | 0.001876 | 0.007452 | 0.1538 | 0.398104 |
| 5000 | 197.03 | 0.002887 | 0.011358 | 0.186294 | 0.359178 |
| 5000 | 491.77 | 0.003979 | 0.014462 | 0.092053 | 0.110597 |
| 5000 | 767.08 | 0.007933 | 0.018096 | 0.030669 | 0.031349 |
| 5000 | 812.97 | 0.008032 | 0.0173 | 0.027427 | 0.027938 |
| 5000 | 949.1 | 0.007608 | 0.014631 | 0.019868 | 0.020093 |
| 5000 | 1083.8 | 0.008022 | 0.012369 | 0.015381 | 0.015499 |