

Estimation of the Population Total-Utilizing Estimators of the Population Size

Mohammad Fraiwan Al-Saleh¹& Ma'aly Salah Ababneh²

Abstract

The estimation of the population total when the population size is unknown is considered. Two estimators using Direct Capture- Recapture sampling technique and another two using Indirect Capture- Recapture sampling technique are proposed. These estimators are compared numerically. In addition, they are compared with the estimator of the population total, when the population size is assumed to be known. It turned out that the suggested estimators can be useful in practice.

Mathematics Subject Classification: S97K80

Keywords: Population Total, Population Size, Capture-Recapture Sampling, Petersen Estimator, Chapman Estimator, Hypergeometric Distribution.

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 (μ), proportion (p), variance (σ^2), total (τ). 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, \dots, O_N be the population measurements. The population mean, total and variance are:

$$\mu = \frac{\sum_{i=1}^N O_i}{N}, \quad \tau = \sum_{i=1}^N O_i = N\mu, \quad \sigma^2 = \frac{\sum_{i=1}^N (O_i - \mu)^2}{N}.$$

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. $P(\text{a subset of size } n \text{ from a population of size } N \text{ is the chosen sample}) = 1/\binom{N}{n}$. Let X_1, \dots, X_n be a SRS of size n from this population. The usual estimators of the population mean & total are, respectively:

¹ Department of Statistics, Yarmouk University, Jordan, m-saleh@yu.edu.jo

Currently at Dept. of Math. and Stat., Jordan Univ. of Science and Technology, Jordan, mfalsaleh@just.edu.jo

² Department of Statistics, Yarmouk University, Jordan, maaliababneh@yahoo.com

$$\hat{\mu} = \frac{\sum_{i=1}^n X_i}{n}, \hat{\tau} = N \hat{\mu}.$$

It is well known that $\hat{\mu}$ & $\hat{\tau}$ are unbiased estimators of μ & τ , respectively. Their variances are:

$$Var(\hat{\mu}) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \& Var(\hat{\tau}) = N^2 Var(\hat{\mu}).$$

The main concern in this work is the estimation of the population total τ when N is unknown. If N is known, then some estimators of N can be used as a *guard* against unsuitable estimates of τ . 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 T be the number of recaptured elements in the second sample. Then, the Petersen estimator of the population size is:

$$\hat{N}_P = \frac{nm}{T}.$$

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

$$\hat{Var}(\hat{N}_P) = \frac{mn(m-T)(n-T)}{T^3}; T \neq 0.$$

Actually, \hat{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

$$\hat{Var}(\hat{N}_C) = \frac{(m+1)(n+1)(m-T)(n-T)}{(T+1)^2(T+2)}.$$

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

- **Inverse Sampling** (Indirect Sampling)

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

$$\hat{N} = \frac{T^* m}{k}, k \leq m,$$

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

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

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 Holly Quran. For more details about the estimation of population total and size, see also Gutierrez and Breidt (2009), Otieno et al. (2005), and Arnab (2004).

In many situations, the ratio estimator is used to estimate τ 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, \dots, 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, \dots, V_N ; the population measurements are $(O_1, V_1), \dots, (O_n, V_n)$. Assume that there is a fair degree of association between O & V . Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be the elements of a SRS, from this population. Now, using the relation

$$\frac{\tau_o}{\tau_v} = \frac{\mu_o}{\mu_v},$$

we can estimate τ_o by $\hat{\tau}_o = \frac{\bar{X}}{\bar{Y}} \tau_v$.

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

$$Var(\hat{\tau}_o) = \tau_v^2 \frac{N-n}{Nn} \frac{1}{\mu_v^2} S_r^2, \text{ where } S_r^2 = \frac{\sum_{i=1}^n (Y_i - rX_i)^2}{n-1}, \text{ (Scheaffer et al., 1995).}$$

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.

In this paper, the estimation of population total (τ) 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 τ ; one is based on Chapman estimator of N and the other is based on a suggested modified estimator of N . In Section 3, τ is estimated using Indirect Sampling. The suggested estimators are compared. Concluding remarks and suggested future works are outlined in Section 4.

2. Estimation of τ 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

$$n = n_1 + n_2 - T. \tag{2.1}$$

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

$$f(t) = P(T = t) = \frac{\binom{n_1}{t} \binom{N - n_1}{n_2 - t}}{\binom{N}{n_2}}, \quad t=0,1,2,\dots,\min(n_1, n_2).$$

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:

$$E(n) = E(n_1 + n_2 - T) = n_1 + n_2 - \frac{n_1 n_2}{N} \tag{2.2}$$

$$Var(n) = Var(n_1 + n_2 - T) = n_2 \left(\frac{n_1}{N} \right) \left(1 - \frac{n_1}{N} \right) \left(\frac{N - n_2}{N - 1} \right) \tag{2.3}$$

Petersen estimator of the population size (N) is:

$$\hat{N}_P = \frac{n_1 n_2}{T}, T = 0, 1, \dots, \min(n_1, n_2). \tag{2.4}$$

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:

$$\hat{N}_C = \frac{(n_1 + 1)(n_2 + 1)}{T + 1} - 1 \tag{2.5}$$

Now,

$$\begin{aligned} E(\hat{N}_C) &= E\left(\frac{(n_1 + 1)(n_2 + 1)}{T + 1} - 1\right) = \frac{(n_1 + 1)(n_2 + 1)}{\binom{N}{n_2}} \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{1}{t + 1} \binom{n_1}{t} \binom{N - n_1}{n_2 - t} \right) - 1 \\ &= (n_2 + 1) \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{\binom{n_1 + 1}{t + 1} \binom{N - n_1}{n_2 - t}}{\binom{N}{n_2}} \right) - 1 = \sum_{t=1}^{\min(n_1 + 1, n_2 + 1)} \left(\frac{\binom{n_1 + 1}{t} \binom{N - n_1}{n_2 + 1 - t}}{\binom{N}{n_2}} \right) - 1 \\ &= (N + 1) \sum_{t=1}^{\min(n_1 + 1, n_2 + 1)} \left(\frac{\binom{n_1 + 1}{t} \binom{N - n_1}{n_2 + 1 - t}}{\binom{N + 1}{n_2 + 1}} \right) - 1 = N - \frac{(n_2 + 1) \binom{N - n_1}{n_2 + 1}}{\binom{N}{n_2}} \end{aligned} \tag{2.6}$$

Clearly, \hat{N}_C is negatively biased. Also,

$$Var(\hat{N}_C) = E(\hat{N}_C - E(\hat{N}_C))^2 = \sum_{t=0}^{\min(n_1, n_2)} \left((\hat{N}_C - E(\hat{N}_C))^2 \frac{\binom{n_1}{t} \binom{N - n_1}{n_2 - t}}{\binom{N}{n_2}} \right) \tag{2.7}$$

Using the two estimators (2.4) and (2.5), we suggest the following new estimator of N :

$$\hat{N}_S = \begin{cases} \hat{N}_P & \text{if } T \neq 0 \\ \hat{N}_C & \text{if } T = 0 \end{cases} = \hat{N}_P I(T \neq 0) + \hat{N}_C I(T = 0) \tag{2.8}$$

Now,

$$\begin{aligned}
 E(\hat{N}_S) &= n_1 n_2 \sum_{t=1}^{\min(n_1, n_2)} \left(\left(\frac{1}{t} \right) P(T=t) \right) + ((n_1 + 1)(n_2 + 1) - 1) P(T=0) \\
 &= n_1 n_2 \sum_{t=1}^{\min(n_1, n_2)} \left(\left(\frac{1}{t} \right) \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + ((n_1 + 1)(n_2 + 1) - 1) \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}}. \quad (2.9) \\
 \text{Var}(\hat{N}_S) &= E(\hat{N}_S - E(\hat{N}_S))^2.
 \end{aligned}$$

But,

$$\begin{aligned}
 (\hat{N}_S - E(\hat{N}_S))^2 &= \left((\hat{N}_P I(T \neq 0) - E(\hat{N}_S) I(T \neq 0)) + (\hat{N}_C I(T = 0) - E(\hat{N}_S) I(T = 0)) \right)^2 \\
 &= (\hat{N}_P I(T \neq 0) - E(\hat{N}_S) I(T \neq 0))^2 + (\hat{N}_C I(T = 0) - E(\hat{N}_S) I(T = 0))^2 \\
 &= (\hat{N}_P - E(\hat{N}_S))^2 I(T \neq 0) + (\hat{N}_C - E(\hat{N}_S))^2 I(T = 0).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 \text{Var}(\hat{N}_S) &= \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{n_1 n_2}{t} - E(\hat{N}_S) \right)^2 P(T=t) + ((n_1 + 1)(n_2 + 1) - 1 - E(\hat{N}_S))^2 P(T=0) \\
 &= \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{n_1 n_2}{t} - E(\hat{N}_S) \right)^2 \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} + ((n_1 + 1)(n_2 + 1) - 1 - E(\hat{N}_S))^2 \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}} \quad (2.10)
 \end{aligned}$$

The above results are summarized in the following lemma:

Lemma (2.1)

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

$$\begin{aligned}
 E(\hat{N}_S) &= n_1 n_2 \sum_{t=1}^{\min(n_1, n_2)} \left(\left(\frac{1}{t} \right) \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + ((n_1 + 1)(n_2 + 1) - 1) \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}}, \\
 \text{Var}(\hat{N}_S) &= \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{n_1 n_2}{t} - E(\hat{N}_S) \right)^2 \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} + ((n_1 + 1)(n_2 + 1) - 1 - E(\hat{N}_S))^2 \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}}.
 \end{aligned}$$

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

$$\hat{\tau}_C = \hat{N}_C \bar{Y}, \quad (2.11)$$

and

$$\hat{\tau}_S = \hat{N}_S \bar{Y}, \quad (2.12)$$

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

Now,

$$E(\hat{\tau}_S) = E(E(\hat{\tau}_S | n)) = E(E(\hat{N}_S \bar{Y} | n)) = E(\hat{N}_S E(\bar{Y} | n)) = E(\hat{N}_S) \mu$$

$$= \mu \left(n_1 n_2 \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{1}{t} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + ((n_1+1)(n_2+1)-1) \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}} \right) \tag{2.13}$$

Similarly, for $\hat{\tau}_C$ we have

$$\begin{aligned} E(\hat{\tau}_C) &= E(\hat{N}_C)\mu = \mu \left(N - \frac{(n_2+1)\binom{N-n_1}{n_2+1}}{\binom{N}{n_2}} \right) \\ &= \tau - \frac{\mu(n_2+1)\binom{N-n_1}{n_2+1}}{\binom{N}{n_2}} \end{aligned} \tag{2.14}$$

Also,

$$\begin{aligned} Var(\hat{\tau}_S) &= E(Var(\hat{\tau}_S | n)) + Var(E(\hat{\tau}_S | n)) = E(Var(\hat{N}_S \bar{Y} | n)) + Var(E(\hat{N}_S \bar{Y} | n)) \\ &= E(\hat{N}_S^2 Var(\bar{Y} | n)) + Var(\hat{N}_S \mu), \end{aligned}$$

but

$$Var(\bar{Y} | n) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right),$$

therefore,

$$\begin{aligned} Var(\hat{\tau}_S) &= E\left(\hat{N}_S^2 \frac{\sigma^2}{n} \frac{N-n}{N-1} \right) + \mu^2 Var(\hat{N}_S) = \frac{\sigma^2}{N-1} E\left(\hat{N}_S^2 \left(\frac{N}{n} - 1 \right) \right) + \mu^2 Var(\hat{N}_S) \\ &= \frac{\sigma^2}{N-1} E\left(\hat{N}_S^2 \frac{N}{n} - \hat{N}_S^2 \right) + \mu^2 Var(\hat{N}_S). \end{aligned}$$

Now,

$$\begin{aligned} E\left(\hat{N}_S^2 \frac{N}{n} \right) &= NE\left(\frac{\hat{N}_S^2}{n} \right) \\ &= N n_1^2 n_2^2 \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{1}{t^2(n_1+n_2-t)} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + N \frac{((n_1+1)(n_2+1)-1)^2}{n_1+n_2} \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}}. \end{aligned}$$

Thus,

$$Var(\hat{\tau}_S) = \frac{\sigma^2}{N-1} \left(N n_1^2 n_2^2 \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{1}{t^2(n_1+n_2-t)} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + N \frac{((n_1+1)(n_2+1)-1)^2}{n_1+n_2} \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}} - (Var(\hat{N}_S) + (E(\hat{N}_S))^2) \right) + \mu^2 Var(\hat{N}_S).$$

(2.15)

Similarly for $\hat{\tau}_C$, we have

$$Var(\hat{\tau}_C) = \frac{\sigma^2}{N-1} \left(N \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{\hat{N}^2}{n} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) - \left(Var(\hat{N}_C) + (E(\hat{N}_C))^2 \right) \right) + \mu^2 Var(\hat{N}_C). \quad (2.16)$$

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

$$E(\hat{\tau}_S) = \mu E(\hat{N}_S), \quad E(\hat{\tau}_C) = \mu E(\hat{N}_C)$$

$$Var(\hat{\tau}_S) = \frac{\sigma^2}{N-1} \left(N n_1^2 n_2^2 \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{1}{t^2(n_1+n_2-t)} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) + N \frac{((n_1+1)(n_2+1)-1)^2}{n_1+n_2} \frac{\binom{N-n_1}{n_2}}{\binom{N}{n_2}} - \right.$$

$$\left. \left(Var(\hat{N}_S) + (E(\hat{N}_S))^2 \right) \right) + \mu^2 Var(\hat{N}_S)$$

$$Var(\hat{\tau}_C) = \frac{\sigma^2}{N-1} \left(N \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{\hat{N}^2}{n} \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} \right) - \left(Var(\hat{N}_C) + (E(\hat{N}_C))^2 \right) \right) + \mu^2 Var(\hat{N}_C).$$

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

$$\hat{\tau} = N\bar{Y}, \quad (2.17)$$

with

$$E(\hat{\tau}) = E(E(\hat{\tau} | n)) = N\mu = \tau.$$

Thus, $\hat{\tau}$ is an unbiased estimator of τ .

$$\begin{aligned} Var(\hat{\tau}) &= Var(N\bar{Y}) = E(Var(N\bar{Y} | n)) + Var(E(N\bar{Y} | n)) \\ &= E\left(N^2 \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \right) + Var(N\mu) = \frac{N^2 \sigma^2}{N-1} \left(NE\left(\frac{1}{n} \right) - 1 \right) \\ &= \frac{N^2 \sigma^2}{N-1} \left(N \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{1}{n_1+n_2-t} \right) \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} - 1 \right) \quad (2.18) \\ &= K\sigma^2, \end{aligned}$$

where,

$$K = \frac{N^2}{N-1} \left(N \sum_{t=0}^{\min(n_1, n_2)} \left(\frac{1}{n_1+n_2-t} \right) \frac{\binom{n_1}{t} \binom{N-n_1}{n_2-t}}{\binom{N}{n_2}} - 1 \right).$$

The efficiency of $\hat{\tau}_S$ with respect to $\hat{\tau}$ is

$$Eff(\hat{\tau}_S; \hat{\tau}) = \frac{MSE(\hat{\tau})}{MSE(\hat{\tau}_S)},$$

where,

$$MSE(\hat{\tau}) = Var(\hat{\tau}), MSE(\hat{\tau}_S) = (bias(\hat{N}_S))^2 \mu^2 + Var(\hat{\tau}_S).$$

Now, let $Var(\hat{\tau}_S) = L\sigma^2 + Var(\hat{N}_S)\mu^2$, where

$$L = \frac{1}{N-1} \left(\begin{aligned} & Nn_1^2 n_2^2 \sum_{t=1}^{\min(n_1, n_2)} \left(\frac{1}{t^2(n_1+n_2-t)} P(T=t) \right) + N \frac{((n_1+1)(n_2+1)-1)^2}{n_1+n_2} P(T=0) \\ & - \left(Var(\hat{N}_S) + (E(\hat{N}_S))^2 \right) \end{aligned} \right)$$

then,

$$Eff(\hat{\tau}_S; \hat{\tau}) = \frac{Var(\hat{\tau})}{MSE(\hat{N}_S)\mu^2 + L\sigma^2} = \frac{K\sigma^2}{MSE(\hat{N}_S)\mu^2 + L\sigma^2}.$$

The efficiency can be rewritten in terms of the coefficient of variation (CV) given by $CV = \frac{\sigma}{\mu}$, ($\mu \neq 0$) as

$$Eff(\hat{\tau}_S; \hat{\tau}) = \frac{\frac{K}{MSE(\hat{N}_S)}(CV(y))^2}{1 + \frac{L}{MSE(\hat{N}_S)}(CV(y))^2} \tag{2.19}$$

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

3. Estimation of τ When N is Unknown Using Capture-Recapture- Indirect Sampling

Indirect sampling is another Capture -Recapture method for estimating N . 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

$$n = n_1 + n_2 - T. \tag{3.1}$$

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

$$f(n_2) = \frac{\binom{n_2-1}{T-1} \binom{N-n_2}{n_1-T}}{\binom{N}{n_1}}, \quad n_2 = T, T+1, \dots, N-n_1+T. \tag{3.2}$$

Also, $E(n_2) = \frac{T(N+1)}{n_1+1}$ (Balakrishnan 2003)

and

$$Var(n_2) = \frac{T(N+1)(N-n_1)(n_1+1-T)}{(n_1+1)^2(n_1+2)} \tag{Khan, 1994}.$$

Now,

$$E(n) = E(n_1 + n_2 - T) = n_1 - T + \frac{T(N+1)}{n_1+1} = n_1 + T \left(\frac{N-n_1}{n_1+1} \right) \tag{3.3}$$

Also,

$$Var(n) = Var(n_2) = \frac{T(N+1)(N-n_1)(n_1+1-T)}{(n_1+1)^2(n_1+2)} \quad (3.4)$$

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

$$\hat{N}_I = \frac{n_1 n_2}{T}, \quad 1 \leq T \leq n_1, T \text{ is integer.} \quad (3.5)$$

Now,

$$E(\hat{N}_I) = E\left(\frac{n_1 n_2}{T}\right) = \frac{n_1(N+1)}{n_1+1} \quad (3.6)$$

Clearly, $E(\hat{N}_I)$ does not depend on T and increasing in n_1 , \hat{N}_I is negatively biased. Now,

$$\begin{aligned} Var(\hat{N}_I) &= Var\left(\frac{n_1 n_2}{T}\right) = \frac{n_1^2}{T^2} Var(n_2) \\ &= \frac{n_1^2}{T} \frac{(N+1)(N-n_1)(n_1+1-T)}{(n_1+1)^2(n_1+2)}. \end{aligned} \quad (3.7)$$

Hence,

$$MSE(\hat{N}_I) = \frac{1}{(n_1+1)^2} \left(\frac{n_1^2}{T} \frac{(N+1)(N-n_1)(n_1+1-T)}{(n_1+2)} + (n_1 - N)^2 \right). \quad (3.8)$$

It can be seen from (3.6) that this estimator of N can be corrected to be unbiased estimator of N as follows:

$$E(\hat{N}_I) = \frac{n_1(N+1)}{n_1+1},$$

which gives

$$E\left(\hat{N}_I \frac{n_1+1}{n_1}\right) = N+1.$$

So, the estimator

$$\hat{N}_I \left(\frac{n_1+1}{n_1} \right) - 1 = \frac{n_2(n_1+1)}{T} - 1$$

is an unbiased estimator of N . Let

$$\hat{N}_I^* = \frac{n_2(n_1+1)}{T} - 1, \quad (3.9)$$

then

$$E(\hat{N}_I^*) = N.$$

Therefore, the mean square error of \hat{N}_I^* is

$$MSE(\hat{N}_I^*) = Var(\hat{N}_I^*) = \frac{(n_1+1)^2}{T^2} Var(n_2) = \frac{(N+1)(N-n_1)(n_1+1-T)}{T(n_1+2)}. \quad (3.10)$$

Clearly, $Var(\hat{N}_I) = \left(\frac{n_1}{n_1+1}\right)^2 Var(\hat{N}_I^*)$, thus, for any values of N, n_1 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:

$$Eff(\hat{N}_I^*, \hat{N}_I) = \frac{MSE(\hat{N}_I)}{MSE(\hat{N}_I^*)}$$

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, \dots, Y_n are the values of variable for the sample element. Two estimators of the population total, τ , are: $\hat{\tau}_I = \hat{N}_I \bar{Y}$ and $\hat{\tau}_I^* = \hat{N}_I^* \bar{Y}$,

Now,

$$\begin{aligned} E(\hat{\tau}_I) &= E(E(\hat{\tau}_I | n_2)) = E(E(\hat{N}_I \bar{Y} | n_2)) = E(\hat{N}_I E(\bar{Y} | n_2)) \\ &= \mu E(\hat{N}_I) = \frac{n_1(N+1)}{n_1+1} \mu. \end{aligned}$$

Also, $E(\hat{\tau}_I^*) = N\mu = \tau$. The variance of $\hat{\tau}_I$ and $\hat{\tau}_I^*$ can be derived as follow

$$\begin{aligned} Var(\hat{\tau}_I) &= E(Var(\hat{\tau}_I | n_2)) + Var(E(\hat{\tau}_I | n_2)) = E(Var(\hat{N}_I \bar{Y} | n_2)) + Var(\mu \hat{N}_I) \\ &= E(\hat{N}_I^2 Var(\bar{Y} | n_2)) + \mu^2 Var(\hat{N}_I) = E(\hat{N}_I^2 Var(\bar{Y} | n_2)) + \mu^2 Var(\hat{N}_I) \\ &= E\left(\frac{n_1^2 n_2^2}{T^2} \frac{\sigma^2}{n_1 + n_2 - T} \frac{N - (n_1 + n_2 - T)}{N - 1}\right) + \mu^2 Var(\hat{N}_I) \\ &= \frac{n_1^2 \sigma^2}{T^2 (N - 1)} E\left(\frac{n_2^2 (N - (n_1 + n_2 - T))}{n_1 + n_2 - T}\right) + \mu^2 Var(\hat{N}_I) \\ &= \frac{n_1^2 \sigma^2}{T(N - 1)} \sum_{n_2=T}^{N-n_1+T} \left[\frac{n_2^2 (N - (n_1 + n_2 - T))}{n_1 + n_2 - T} \frac{\binom{n_2 - 1}{T - 1} \binom{N - n_2}{n_1 - T}}{\binom{N}{n_1}} \right] + \mu^2 Var(\hat{N}_I) \\ &= B\sigma^2 + \mu^2 Var(\hat{N}_I), \end{aligned} \tag{3.11}$$

where,

$$B = \frac{n_1^2}{T(N - 1)} \sum_{n_2=T}^{N-n_1+T} \left[\frac{n_2^2 (N - (n_1 + n_2 - T))}{n_1 + n_2 - T} \frac{\binom{n_2 - 1}{T - 1} \binom{N - n_2}{n_1 - T}}{\binom{N}{n_1}} \right]. \tag{3.12}$$

Now,

$$\begin{aligned} Var(\hat{\tau}_I^*) &= E(Var(\hat{\tau}_I^* | n_2)) + Var(E(\hat{\tau}_I^* | n_2)) = E(Var(\hat{N}_I^* \bar{Y} | n_2)) + Var(E(\hat{N}_I^* \bar{Y} | n_2)) \\ &= E(\hat{N}_I^{*2} Var(\bar{Y} | n_2)) + \mu^2 Var(\hat{N}_I^*) = E\left(\left(\frac{n_2(n_1 + 1)}{T} - 1\right)^2 \frac{\sigma^2}{n_1 + n_2 - T} \frac{N - (n_1 + n_2 - T)}{N - 1}\right) \\ &\quad + \frac{(N + 1)(N - n_1)(n_1 + 1 - T)}{T(n_1 + 2)} \mu^2 \\ &= \frac{\sigma^2}{T^2 (N - 1)} E\left(\left(n_2(n_1 + 1) - T\right)^2 \left(\frac{N}{n_1 + n_2 - T} - 1\right)\right) \end{aligned}$$

$$+ \frac{(N + 1)(N - n_1)(n_1 + 1 - T)}{T(n_1 + 2)} \mu^2 \tag{3.18}$$

$$= Z\sigma^2 + J\mu^2,$$

where

$$Z = \frac{1}{T^2(N - 1)} E((n_2(n_1 + 1) - T)^2 \left(\frac{N}{n_1 + n_2 - T} - 1\right)) \text{ \& } J = \frac{(N + 1)(N - n_1)(n_1 + 1 - T)}{T(n_1 + 2)}.$$

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

Now, if N is known, then τ can be estimated based on a simple random sample of size $n = n_1 + n_2 - t$ by $\hat{\tau} = N\bar{Y}$.

$$\begin{aligned} \text{Var}(\hat{\tau}) &= \text{Var}(N\bar{Y}) = E(\text{Var}(N\bar{Y} | n)) + \text{Var}(E(N\bar{Y} | n)) = E\left(N^2 \frac{\sigma^2}{n} \left(\frac{N - n}{N - 1}\right)\right) + \text{Var}(N\mu) \\ &= \frac{N^2 \sigma^2}{N - 1} E\left(\frac{N}{n} - 1\right) \\ &= \frac{N^2 \sigma^2}{N - 1} E\left(\frac{N}{n}\right) - \frac{N^2 \sigma^2}{N - 1} = \frac{N^3 \sigma^2}{N - 1} E\left(\frac{1}{n}\right) - \frac{N^2 \sigma^2}{N - 1} \\ &= \frac{N^3 \sigma^2}{N - 1} \sum_{n_2=T}^{N-n_1+T} \left(\frac{1}{n_1 + n_2 - t} \frac{\binom{n_2 - 1}{T - 1} \binom{N - n_2}{n_1 - T}}{\binom{N}{n_1}} \right) - \frac{N^2 \sigma^2}{N - 1} = H\sigma^2, \end{aligned}$$

where,

$$H = \frac{N^2}{N - 1} \left(N \sum_{n_2=T}^{N-n_1+T} \left(\frac{1}{n_1 + n_2 - t} \frac{\binom{n_2 - 1}{T - 1} \binom{N - n_2}{n_1 - T}}{\binom{N}{n_1}} \right) - 1 \right).$$

The efficiency of $\hat{\tau}_I$ with respect to $\hat{\tau}$ (obtained for a sample size equal the expected sample size) is:

$$\text{Eff}(\hat{\tau}_I; \hat{\tau}) = \frac{MSE(\hat{\tau})}{MSE(\hat{\tau}_I)},$$

$$\begin{aligned} MSE(\hat{\tau}_I) &= (\text{bias}(\hat{\tau}_I))^2 + \text{Var}(\hat{\tau}_I) = (\text{bias}(\hat{N}_I))^2 \mu^2 + B\sigma^2 + \text{Var}(\hat{N}_I) \mu^2 \\ &= B\sigma^2 + MSE(\hat{N}_I) \mu^2, \end{aligned}$$

$$\text{Eff}(\hat{\tau}_I; \hat{\tau}) = \frac{H\sigma^2}{B\sigma^2 + MSE(\hat{N}_I) \mu^2}.$$

The efficiency can be rewritten in terms of the coefficient of variation (CV), given by $CV = \frac{\sigma}{\mu}$, (assume $\mu \neq 0$), as:

$$\text{Eff}(\hat{\tau}_I; \hat{\tau}) = \frac{H}{1 + \frac{B}{MSE(\hat{N}_I)} (CV(y))^2}.$$

Some values of the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}$ are given in Table (3.3). Similarly,

$$\text{Eff}(\hat{\tau}_I^*; \hat{\tau}_I) = \frac{B(CV(y))^2 + MSE(\hat{N}_I)}{Z(CV(y))^2 + J}.$$

The efficiency of $\hat{\tau}_I^*$ w.r.t. $\hat{\tau}_I$ is given in Table (3.4). Based on the tables, we have the following conclusions:

1. From Table (3.3), the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}$ increases when the CV increases.

2. For small sample size, the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}_S$ is greater than or close to one. Also, the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}_C$ is greater than one for small sample size and small value of the CV.
3. For large sample size, the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}_S$ is less than one.
4. For large value of the coefficient of variation, the efficiency of $\hat{\tau}_I$ w.r.t. $\hat{\tau}_C$ is less than one, and it decreases when the sample size increases.
5. $\hat{\tau}_I^*$ is more efficient than $\hat{\tau}_I$ 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 $\hat{\tau}_C$ based on Chapman estimator \hat{N}_C is better than the estimator $\hat{\tau}_S$ based on the suggested modified estimator \hat{N}_S . On the other hand, if the expected sample size is small then $\hat{\tau}_S$ is more efficient than $\hat{\tau}_C$.
2. For small expected sample size and CV, we found that it is better to use Indirect Sampling to estimate τ 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 τ .

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 τ conditioning on \hat{N} to be between $N - \varepsilon$ and $N + \varepsilon$ for some ε .
- Estimation of τ 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 multi-character surveys. *Journal of the Indian Society of Agricultural Statistics* **58**, 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. *Revista Colombiana de Estadística* **32**, 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.

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): $Eff(\hat{\tau}_C; \hat{\tau})$

N	$E(n)$	$CV = 0.5$	$CV = 1$	$CV = 5$	$CV = 10$
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): $Eff(\hat{\tau}_S; \hat{\tau})$

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): $Eff(\hat{\tau}_S; \hat{\tau}_C)$

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

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): $Eff(\hat{\tau}_I; \hat{\tau})$

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

Table(3.4): $Eff(\hat{\tau}_i^*; \hat{\tau}_I)$

N	$E(n)$	$CV = 0.5$	$CV = 1$	$CV = 5$	$CV = 10$
1000	68.627	0.961554074	0.961567	0.961839	0.962132
1000	105.88	0.976538358	1.017843	1.782482	2.402905
1000	124.51	0.986293669	1.054636	2.265405	3.177549
1000	144.55	1.016554076	1.120112	2.862745	4.070053
1000	189.11	1.098159758	1.430944	6.043467	8.431787
1000	233.66	1.199760755	1.813903	9.464754	12.88522
1000	324.7	1.447352181	2.713834	16.74673	22.13178
1000	438.25	3.130990361	8.738736	49.20767	58.61586
1000	474.1	3.790414517	11.03604	59.69973	70.17584
5000	148.51	0.98039283	0.980394	0.980422	0.98047
5000	197.03	0.983243013	0.99134	1.194488	1.497824
5000	491.77	1.025178903	1.11397	3.103942	5.416284
5000	767.08	1.315245288	2.23671	16.78553	25.90941
5000	812.97	1.711857968	3.694578	24.88031	32.22224
5000	949.1	1.632535803	3.445735	28.77917	42.04627
5000	1083.8	1.939294326	4.608143	39.26858	55.66183