

THE INCLUSION PROBABILITIES OF MEDIAN RANKED SET SAMPLING UNDER DIFFERENT SELECTION PROCEDURES.

YAPRAK ARZU OZDEMIR AND FIKRI GOKPINAR

ABSTRACT. In this paper, we developed generalized formulas to compute the inclusion probabilities of a median ranked set sample in a finite population setting under Level 0 and Level 2 sampling procedures given by Deshpande et al.(2006). We also compared the inclusion probabilities of these sampling procedures with the inclusion probabilities of Level 1 given by Ozdemir and Gokpinar(2008) under different population and sample sizes.

1. INTRODUCTION

McIntyre [4] introduced a sampling design, called Ranked Set Sampling (RSS) which has a better design than the Simple Random Sampling (SRS) design for the estimation of the population mean. RSS is preferred for use in some fields such as the environment, ecology, agriculture and medicine in which measurement of the sampling units in terms of the variable of interest is quite difficult or expensive in terms of cost, time and other factors.

In order to increase the efficiency of RSS, modified RSS designs have in the past been suggested for different distribution types. Some of the modified RSS designs are Median RSS (MRSS), Extreme RSS (ERSS) and Multistage RSS (MSRSS) ([5], [8], [1]). MRSS is used to reduce the errors in ranking and to increase the efficiency of the estimator for symmetric unimodal distributions, such as the normal distribution. These studies were based on the assumption of an infinite population setting. In recent years, RSS is also investigated under a finite population setting. Takahasi and Futatsuya ([9], [10]) were the first authors to give finite population theory in RSS. Al-Saleh and Samawi [2] gave an adjusted selection procedure for RSS. Also, Deshpande et al. [3] described three different sample selection procedure in RSS-called Level 0, Level 1 and Level 2- to construct nonparametric confidence

Received by the editors Dec. 12, 2009, Accepted: Feb. 26, 2010.

2000 *Mathematics Subject Classification.* Primary 62D05; Secondary 65C60.

Key words and phrases. Finite Population, Inclusion Probability, Median Ranked Set Sampling.

intervals for the quantile of a finite population. The Level 1 sampling procedure is equivalent to Al-Saleh and Samawi's [2] adjusted selection procedure.

In SRS, all units in a population have the same inclusion probabilities. However, in RSS and in its modifications some units in the population may have different inclusion probabilities. In sampling theory, the inclusion probabilities give an insight into how the RSS designs control the inclusion of the units in the sample. On the other hand, the unequal inclusion probabilities must be taken into account in order to come up with reasonable estimates of population parameters. Estimators like Horvitz-Thompson (HT) which is unbiased for the population mean and total for any sampling design, often depend on the inclusion probability of each unit in the sample. The HT estimator of the population total T is defined as below,

$$\hat{T}_{HT} = \sum \frac{u_i}{\pi_i},$$

where π_i is the inclusion probability of the unit u_i in the sample. This estimator is often feasible even for very complex sampling designs. In order to calculate HT estimator in RSS, the inclusion probabilities of each unit in the population and also their rank with respect to a variable of interest must be known. The rank of each unit in the population can be determined by using a concomitant variable which is highly correlated with the variable of interest. However, it is hard and complex to calculate the inclusion probabilities in RSS. Al-Saleh and Samawi [2] obtained the inclusion probabilities with respect to their adjusted selection procedure in a finite population setting only when the sample size was 2 or 3. Using the same selection procedure, Ozdemir and Gokpinar [6] derived a generalized formula for computing the inclusion probabilities in RSS for any sample and population size. Ozdemir and Gokpinar [7] developed a new formula to calculate the inclusion probabilities of population units based on the MRSS design for any sample and population size under the Level 1 sampling procedure. However, in practical use it may be necessary to know the inclusion probabilities for Level 0, Level 1 and Level 2 sampling procedures. In this study, we developed generalized formulas for the inclusion probabilities for Level 0 and Level 2 sampling procedures and shortened the formulas for the inclusion probabilities for the Level 1 sampling procedure which is given by Ozdemir and Gokpinar [7]. We also compared Level 0, Level 1 and Level 2 sampling procedures to the SRS sampling procedure.

The paper is organized as follows: In Section 2, we give generalized formulas for computing the inclusion probabilities in MRSS under Level 0, Level 1 and Level 2 sampling procedures. In Section 3, the inclusion probabilities under these sampling procedures based on MRSS and SRS are compared for different sample and population sizes. Section 4 draws the main conclusions.

2. INCLUSION PROBABILITIES OF THE UNITS

Let $u_1 < u_2 < \dots < u_N$ be the distinct ordered population units. Suppose that $n = mr$ units are to be chosen from this population with MRSS, where m and r are the set and cycle sizes, respectively. We use a modified version of the "Level 0 sampling", "Level 1 sampling" and "Level 2 sampling" procedures given in Deshpande et al. [3].

In order to calculate the inclusion probabilities which are based on the MRSS design for any set size m , cycle size r or population size N , the required definitions are given as follows:

$A_{c,j}$; the event of choosing u_k in the c^{th} cycle and j^{th} selection,
 $y_{c,j}$; the unit selected in the c^{th} cycle and j^{th} selection,

Using these definitions, the inclusion probability of the k^{th} unit u_k , $\pi_N(k)$, for $k = 1, 2, \dots, N$ can be defined as follows:

$$\pi_N(k) = \sum_{c=1}^r \pi_N^{(c)}(k) = \sum_{c=1}^r \sum_{j=1}^m \pi_N^{(c,j)}(k), \quad (1)$$

where $\pi_N^{(c)}(k)$ is the inclusion probability of u_k in the c^{th} cycle ($c = 1, 2, \dots, r$) and $\pi_N^{(c,j)}(k)$ is the inclusion probability of u_k in the j^{th} selection ($j = 1, 2, \dots, m$) and c^{th} cycle.

2.1. Inclusion Probabilities under the Level 0 Sampling Procedure. In the Level 0 sampling procedure, all sample units are selected without replacement but all of these units are returned to the population. So in this selection procedure, a unit can be observed more than one time in the final sample. The algorithm of the Level 0 sampling procedure is given below:

Case 1 (Odd set size m) : In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.
2. The sampled units are ranked with respect to the variable of interest and the $\left(\frac{m+1}{2}\right)^{th}$ order statistic is selected for measurement.
3. All other m units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

Case 2 (Even set size m): In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.

2. The sampled units are ranked with respect to the variable of interest and the $\left(\frac{m}{2}\right)^{th}$ order statistic and $\left(\frac{m+2}{2}\right)^{th}$ order statistic are selected for measurement when $j \leq \frac{m}{2}$ and $j > \frac{m}{2}$, respectively.
3. All other m units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

The entire cycle in both cases may be repeated, if necessary, r times to produce a median ranked set sample of size $n = mr$.

In this sampling procedure, since the selected units for ranking and measurement are returned to the population, the probabilities of $A_{c,j}$ are independent of previous

selections and can be written as follows:

Case 1:

$$P(A_{c,j}) = P(A) = \frac{\binom{k-1}{\frac{m-1}{2}} \binom{N-k}{\frac{m-1}{2}}}{\binom{N}{m}} \quad c = 1, \dots, r, \quad j = 1, \dots, m \quad (2)$$

Case 2:

$$P(A_{c,j}) = \begin{cases} P(A^1) = \frac{\binom{k-1}{\frac{m}{2}-1} \binom{N-k}{\frac{m}{2}}}{\binom{N}{m}} & c = 1, \dots, r, \quad j \leq \frac{m}{2} \\ P(A^2) = \frac{\binom{k-1}{\frac{m}{2}} \binom{N-k}{\frac{m}{2}-1}}{\binom{N}{m}} & c = 1, \dots, r, \quad j > \frac{m}{2}. \end{cases}$$

In case 1, u_k cannot be selected with the probability $1 - P(A)$. In case 2, u_k cannot be selected with the probability $1 - P(A^1)$ when $j \leq \frac{m}{2}$ or $1 - P(A^2)$ when $j > \frac{m}{2}$. In this way the probability that u_k is not selected in any r cycle and m selection is $(1 - P(A))^{rm}$ for case 1 and $(1 - P(A^1))^{rm/2}(1 - P(A^2))^{rm/2}$ for case 2. Therefore $1 - (1 - P(A))^{rm}$ and $1 - (1 - P(A^1))^{rm/2}(1 - P(A^2))^{rm/2}$ are the probabilities of choosing u_k at least one time in all r cycle and m selection. So $\pi_N(k)$ can be written as,

$$\pi_N(k) = \begin{cases} 1 - (1 - P(A))^{rm} & \text{for case 1} \\ 1 - (1 - P(A^1))^{rm/2}(1 - P(A^2))^{rm/2} & \text{for case 2} \end{cases} \quad (3)$$

2.2. Inclusion Probabilities under the Level 1 Sampling Procedure. In the Level 1 sampling procedure, all sample units are selected without replacement and all of these units except the measured one are returned to the population. So in this selection procedure, a unit cannot be observed more than one time in the final sample but it can be used for ranking purpose. The algorithm of the Level 1 sampling procedure is given below:

Case 1 (Odd set size m) : In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.
2. The sampled units are ranked with respect to the variable of interest and the $\left(\frac{m+1}{2}\right)^{th}$ order statistic is selected for measurement.
3. All other $m - 1$ units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

Case 2 (Even set size m): In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.
2. The sampled units are ranked with respect to the variable of interest and the $\left(\frac{m}{2}\right)^{th}$ order statistic is selected for measurement when $j \leq \frac{m}{2}$ and the $\left(\frac{m+2}{2}\right)^{th}$ order statistic is selected for measurement when $j > \frac{m}{2}$.
3. All other $m - 1$ units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

The entire cycle in both cases may be repeated, if necessary, r times to produce a median ranked set sample of size $n = mr$.

For obtaining the inclusion probabilities, we required some other definitions for the Level 0 sampling procedure. In this procedure, since only units selected for ranking are returned to the population and the measured one is not returned to the population, the probabilities of $A_{c,j}$ are dependent on previous selections. For this reason, we must know how many measured units are greater or smaller than u_k in the previous selections. Let $l_{c,j}$ indicate that the selected unit is greater or smaller than u_k in the c^{th} cycle and j^{th} selection as follows:

$$l_{c,j} = \begin{cases} 0 & y_{c,j} > u_k \\ 1 & y_{c,j} < u_k \end{cases},$$

$B_{c,j}^{l_{c,j}} = B_j^0$; the event of $\{y_{c,j} > u_k\}$ in the c^{th} cycle and j^{th} selection,

$B_{c,j}^{l_{c,j}} = B_{c,j}^1$; the event of $\{y_{c,j} < u_k\}$ in the c^{th} cycle and j^{th} selection.

The number of units smaller than u_k that can be chosen in the previous selections from the j^{th} selection in the c^{th} cycle is defined by

$$a_1 = \begin{cases} 0 & c = 1, j = 1 \\ \sum_{z=1}^{j-1} l_{1,z} & c = 1, j > 1 \\ \sum_{v=1}^{c-1} \sum_{z=1}^m l_{v,z} & c > 1, j = 1 \\ \sum_{v=1}^{c-1} \sum_{z=1}^m l_{v,z} + \sum_{z=1}^{j-1} l_{c,z} & c > 1, j > 1. \end{cases}$$

and the number of units greater than u_k that can be chosen in the previous selections from the j^{th} selection in the c^{th} cycle is defined by

$$b_1 = \begin{cases} 0 & c = 1, j = 1 \\ j - 1 - \sum_{z=1}^{j-1} l_{1,z} & c = 1, j > 1 \\ (c-1)m - \sum_{v=1}^{c-1} \sum_{z=1}^m l_{v,z} & c > 1, j = 1 \\ (c-1)m + j - 1 - \left(\sum_{v=1}^{c-1} \sum_{z=1}^m l_{v,z} + \sum_{z=1}^{j-1} l_{c,z} \right) & c > 1, j > 1. \end{cases}$$

Thus, $\pi_N^{(c,j)}(k)$ is defined as follows

$$\begin{aligned} \pi_N^{(c,j)}(k) &= \sum P(A_{c,j} / B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \cap \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) \\ &\quad P(B_{c,j-1}^{l_{c,j-1}} / B_{c,j-2}^{l_{c,j-2}} \cap \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) \dots P(B_{1,1}^{l_{1,1}}) \end{aligned} \quad (4)$$

where the summation includes all $2^{(c-1)m+j-1}$ possible permutations of

$$(l_{c,j}, \dots, l_{c,1}, \dots, l_{1,m}, \dots, l_{1,1}).$$

Case 1: When the set size m is odd, in cycle c before the j^{th} selection, $(c-1)m +$

$j - 1 = a_1 + b_1$ units must be selected from the population so that the number of remaining units in the population is $N - (a_1 + b_1)$. These remaining units contain $k - a_1 - 1$ units smaller than u_k and $N - b_1 - k$ units greater than u_k . In the j^{th} selection, we now consider the probability of selecting a unit $y_{c,j}$ greater than u_k ($l_{c,j} = 0$) given that the unit u_k is not selected prior to the cycle c and selection j . This probability can be computed from

$$P(B_{c,j}^0/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \frac{\sum_{i=\frac{m+1}{2}}^m \binom{k-a_1}{m-i} \binom{N-b_1-k}{i}}{\binom{N-(a_1+b_1)}{m}}.$$

When $l_{c,j} = 1$, the probability of choosing a unit smaller than u_k in the c^{th} cycle and j^{th} selection when u_k is not included in the previous selections, $P(B_{c,j}^1/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}})$, is written as

$$P(B_{c,j}^1/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \frac{\sum_{i=0}^{\frac{m-1}{2}} \binom{k-a_1-1}{m-i} \binom{N-b_1-k+1}{i}}{\binom{N-(a_1+b_1)}{m}}.$$

Finally, for choosing u_k in the c^{th} cycle and j^{th} selection, the sample must have $\frac{m-1}{2}$ units smaller than u_k and $\frac{m-1}{2}$ units greater than u_k . Thus, the probability of choosing u_k in the c^{th} cycle and j^{th} selection is given by

$$P(A_{c,j}/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \frac{\binom{k-a_1-1}{\frac{m-1}{2}} \binom{N-b_1-k}{\frac{m-1}{2}}}{\binom{N-(a_1+b_1)}{m}}.$$

Case 2: When the set size m is even the probabilities $P(B_{c,j}^1/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}})$ and $P(B_{c,j}^0/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}})$ can be obtained in a fashion similar to Case 1. However, these probabilities depend on whether $j \leq \frac{m}{2}$ or not. Therefore, the probability $P(B_{c,j}^0/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}})$ for $j \leq \frac{m}{2}$ and $j > \frac{m}{2}$ can be computed from

$$P(B_{c,j}^0/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \begin{cases} \frac{\sum_{i=\frac{m}{2}+1}^m \binom{k-a_1}{m-i} \binom{N-b_1-k}{i}}{\binom{N-(a_1+b_1)}{m}} & j \leq \frac{m}{2} \\ \frac{\sum_{i=\frac{m}{2}}^m \binom{k-a_1}{m-i} \binom{N-b_1-k}{i}}{\binom{N-(a_1+b_1)}{m}} & j > \frac{m}{2}. \end{cases}$$

When $l_{c,j} = 1$, the probability $P(B_{c,j}^1/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}})$ is given for $j \leq \frac{m}{2}$ and $j > \frac{m}{2}$ as follows:

$$P(B_{c,j}^1/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \begin{cases} \frac{\sum_{i=0}^{\frac{m}{2}} \binom{k-a_1-1}{m-i} \binom{N-b_1-k+1}{i}}{\binom{N-(a_1+b_1)}{m}} & j \leq \frac{m}{2} \\ \frac{\sum_{i=0}^{\frac{m}{2}-1} \binom{k-a_1-1}{m-i} \binom{N-b_1-k+1}{i}}{\binom{N-(a_1+b_1)}{m}} & j > \frac{m}{2}. \end{cases}$$

Finally, for choosing u_k in the c^{th} cycle and j^{th} selection, when $j \leq \frac{m}{2}$, the sample must have $\frac{m}{2} - 1$ units smaller than u_k and $\frac{m}{2}$ units greater than u_k . Moreover, when $j > \frac{m}{2}$, the sample must have $\frac{m}{2}$ units smaller than u_k and $\frac{m}{2} - 1$ units greater than u_k . As a result, the probability of choosing u_k in the c^{th} cycle and j^{th} selection is given by

$$P(A_{c,j}/B_{c,j-1}^{l_{c,j-1}} \cap B_{c,j-2}^{l_{c,j-2}} \dots \cap B_{c,1}^{l_{c,1}} \cap \dots \cap B_{1,1}^{l_{1,1}}) = \begin{cases} \frac{\binom{k-a_1-1}{\frac{m}{2}-1} \binom{N-k-b_1}{\frac{m}{2}}}{\binom{N-(a_1+b_1)}{m}} & j \leq \frac{m}{2} \\ \frac{\binom{k-a_1-1}{\frac{m}{2}} \binom{N-k-b_1}{\frac{m}{2}-1}}{\binom{N-(a_1+b_1)}{m}} & j > \frac{m}{2}. \end{cases}$$

Using these formulas, the inclusion probabilities for all the units in the population can be derived easily.

2.3. Inclusion Probabilities under the Level 2 Sampling Procedure. In the Level 2 sampling procedure, all sample units are selected without replacement and all of these units are not returned to the population. So in this selection procedure, a unit can not be observed more than one time in the final sample and cannot be used for ranking. The algorithm of Level 2 sampling procedure is given below:

Case 1 (Odd set size m): In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.
2. The sampled units are ranked with respect to the variable of interest and the $\left(\frac{m+1}{2}\right)^{th}$ order statistic is selected for measurement.
3. No units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

Case 2 (Even set size m): In the j^{th} selection,

1. A simple random sample of size m is selected without replacement from the population.

2. The sampled units are ranked with respect to the variable of interest and the $(\frac{m}{2})^{th}$ order statistic is selected for measurement when $j \leq \frac{m}{2}$ and the $(\frac{m+2}{2})^{th}$ order statistic is selected for measurement when $j > \frac{m}{2}$.
3. No units are returned to the population.
4. Steps 1-3 are repeated for $j = 1, 2, \dots, m$ to obtain a sample of size m .

The entire cycle in both cases may be repeated, if necessary, r times to produce a median ranked set sample of size $n = mr$.

For obtaining the inclusion probabilities we required some definitions. $t_{c,j}$ indicates that the number of units smaller than u_k in the c^{th} cycle and j^{th} selection and $t_{c,j} = 0, 1, 2, \dots, m$. So $B_{c,j}^{t_{c,j}}$ is the event that $t_{c,j}$ units are chosen which are smaller than u_k in the c^{th} cycle and j^{th} selection.

The number of units smaller than u_k that can be chosen in the selections previous to the j^{th} selection in the c^{th} cycle is defined by

$$a_2 = \begin{cases} 0 & c = 1, j = 1 \\ \sum_{z=1}^{j-1} t_{1,z} & c = 1, j > 1 \\ \sum_{v=1}^{c-1} \sum_{z=1}^m t_{v,z} & c > 1, j = 1 \\ \sum_{v=1}^{c-1} \sum_{z=1}^m t_{v,z} + \sum_{z=1}^{j-1} t_{c,z} & c > 1, j > 1. \end{cases}$$

and the number of units greater than u_k that can be chosen in the selections previous to the j^{th} selection in the c^{th} cycle is defined by

$$b_2 = \begin{cases} 0 & c = 1, j = 1 \\ (j-1)m - \sum_{z=1}^{j-1} t_{1,z} & c = 1, j > 1 \\ (c-1)m^2 - \sum_{v=1}^{c-1} \sum_{z=1}^m t_{v,z} & c > 1, j = 1 \\ (c-1)m^2 + (j-1)m - \left(\sum_{v=1}^{c-1} \sum_{z=1}^m t_{v,z} + \sum_{z=1}^{j-1} t_{c,z} \right) & c > 1, j > 1. \end{cases}$$

Thus, $\pi_N^{(c,j)}(k)$ is defined as follows:

$$\begin{aligned} \pi_N^{(c,j)}(k) &= \sum P(A_{c,j} / B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}}). \\ &P(B_{c,j-1}^{t_{c,j-1}} / B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}}) \dots P(B_{1,1}^{t_{1,1}}) \end{aligned} \quad (5)$$

where the summation includes all $(m+1)^{(c-1)m+j-1}$ possible permutations of $(t_{c,j}, \dots, t_{c,1}, t_{c-1,m}, \dots, t_{c-1,1}, \dots, t_{1,m}, \dots, t_{1,1})$. The conditional probabilities for the Level 2 sampling procedure, $P(B_{c,j}^{t_{c,j}}/B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}})$ is different from calculated conditional probabilities for the Level 1 sampling procedure and is independent of cases 1 and 2 as given below:

$$P(B_{c,j}^{t_{c,j}}/B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}}) = \frac{\binom{k-1-a_2}{t_{c,j}} \binom{N-k-b_2}{m-t_{c,j}}}{\binom{N-(a_2+b_2)}{m}}.$$

But the probability of $P(A_{c,j}/B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}})$ is dependent on these cases. So these probabilities are given for case 1 and 2 as follows:

Case1:

$$P(A_{c,j}/B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}}) = \frac{\binom{k-1-a_2}{\frac{m-1}{2}} \binom{N-k-b_2}{\frac{m-1}{2}}}{\binom{N-(a_2+b_2)}{m}}$$

Case2:

$$P(A_{c,j}/B_{c,j-1}^{t_{c,j-1}} \cap B_{c,j-2}^{t_{c,j-2}} \cap \dots \cap B_{c,1}^{t_{c,1}} \cap \dots \cap B_{1,1}^{t_{1,1}}) = \begin{cases} \frac{\binom{k-a_2-1}{\frac{m}{2}-1} \binom{N-k-b_2}{\frac{m}{2}}}{\binom{N-(a_2+b_2)}{m}} & j \leq \frac{m}{2} \\ \frac{\binom{k-a_2-1}{\frac{m}{2}} \binom{N-k-b_2}{\frac{m}{2}-1}}{\binom{N-(a_2+b_2)}{m}} & j > \frac{m}{2}. \end{cases}$$

3. COMPARISON OF THE SAMPLING PROCEDURES

In this section, we investigate the effects of the sample size m and population size N on the inclusion probability of elements in the population for Level 0, Level 1 and Level 2 sampling procedures. The inclusion probabilities are calculated using MATLAB 7.0. The calculated inclusion probabilities for Level 0, Level 1 and Level 2 sampling procedures are compared to the inclusion probabilities derived from SRS with the same sample and population sizes. It is well known that the inclusion probability for all elements in the population is m/N for SRS. For this comparison, Figures 1- 6 are constructed for $N = 20, 50, 500$ and sample size $m = 3, 4$.

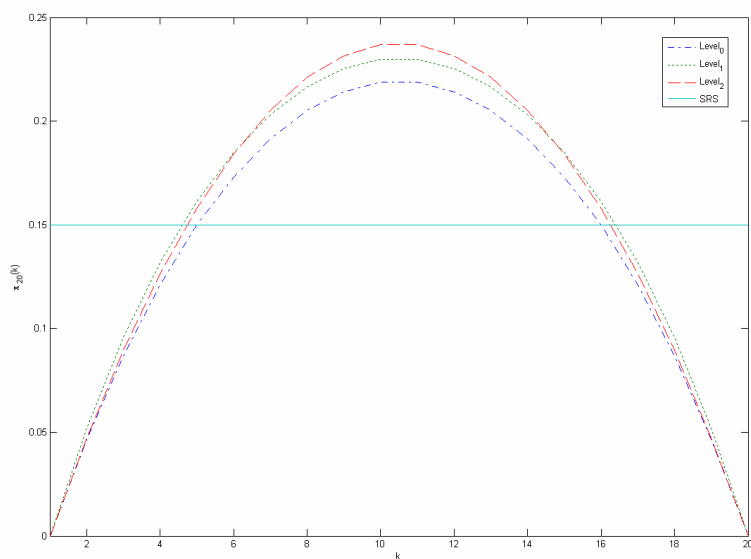


Figure 1: Inclusion Probabilities for $m=3$ and $N=20$

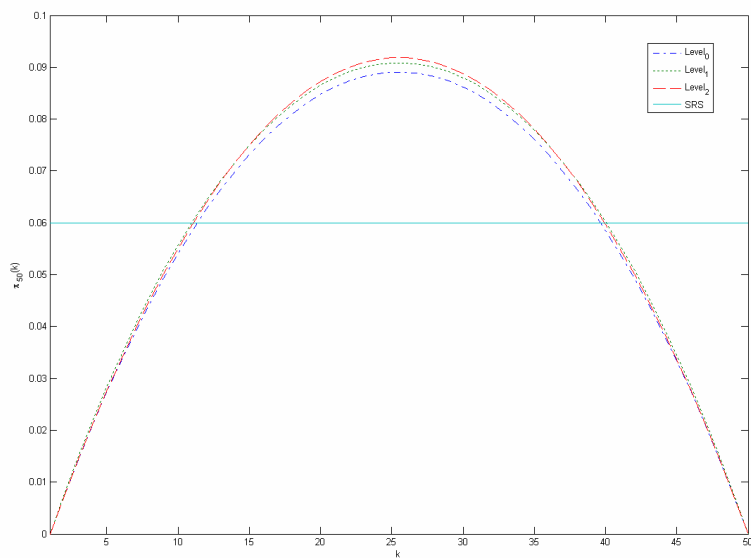


Figure 2: Inclusion Probabilities for $m=3$ and $N=50$

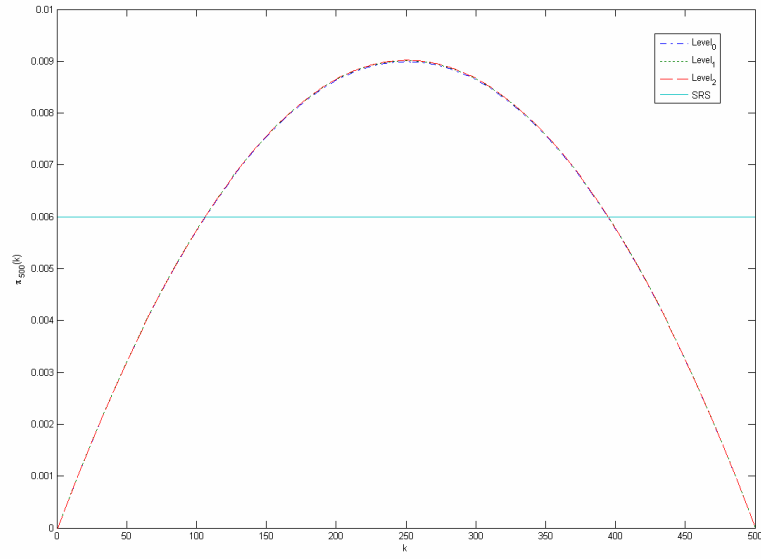


Figure 3: Inclusion Probabilities for $m=3$ and $N=500$

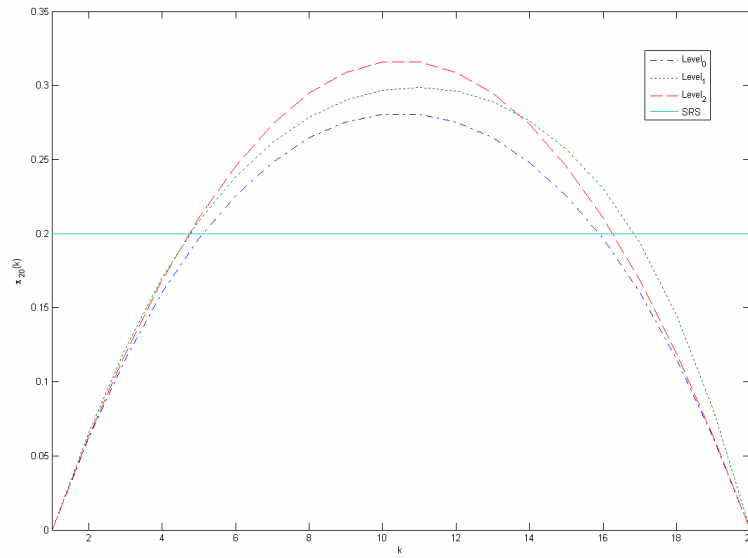


Figure 4: Inclusion Probabilities for $m=4$ and $N=20$

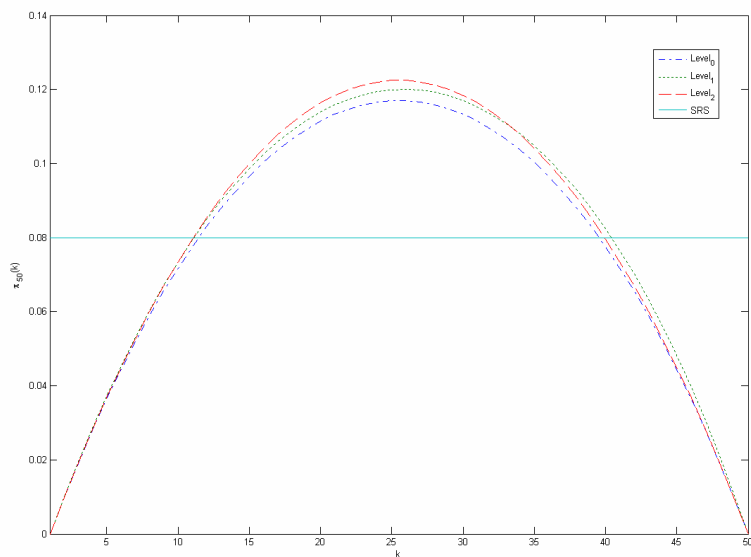


Figure 5: Inclusion Probabilities for $m=4$ and $N=50$

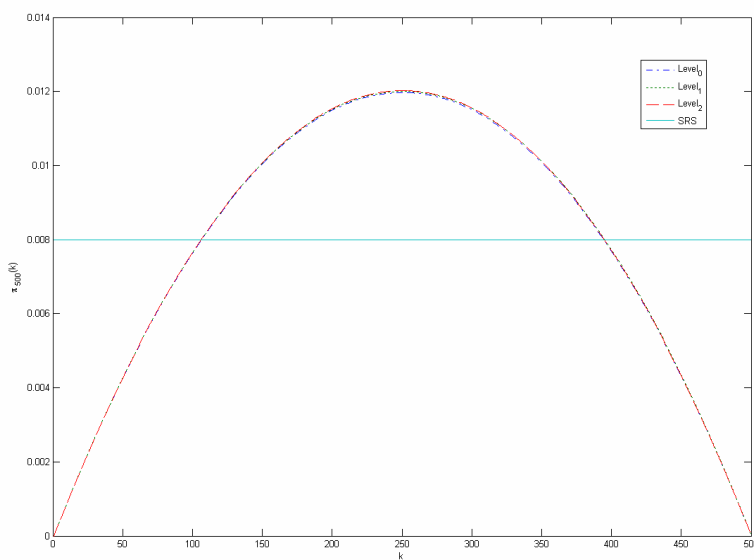


Figure 6: Inclusion Probabilities for $m=4$ and $N=500$

As shown in Figures 1 to 6, the inclusion probabilities of all the population units are equal in SRS when the sample and population sizes are fixed. In all MRSS procedures, the $\lfloor \frac{m-1}{2} \rfloor$ units at the extremes will have zero inclusion probabilities for all sample and population sizes. In addition, the middle values of the population units will have greater inclusion probabilities than the other units in the population. When the set size is 3 the inclusion probabilities are symmetric around the median of the population. However, when the set size is 4, this property is not valid since the second greatest unit of the first two selections and the third greatest unit of the last two selections are chosen. For odd set sizes, on the other hand, just the median values of the sets are chosen for all selections. Furthermore, when the sample size increases, the inclusion probability of any unit in the population increases for all the sampling designs considered. For all set sizes, the inclusion probabilities of middle values under Level 2 sampling procedure are greater than the other sampling procedures. But the inclusion probabilities of the extreme values under Level 1 sampling procedure are greater than the others. When the set size is 3 the difference between the inclusion probabilities under Level 0, Level 1 and Level 2 sampling procedures are symmetrical, but when the set size is 4 these differences are not symmetrical. In all sample sizes, the Level 0 sampling procedure gives the units smallest inclusion probabilities. Also, when the population size increases, the inclusion probabilities under Level 0, Level 1 and Level 2 sampling approach equivalence.

4. CONCLUSIONS

MRSS, a modification of RSS, is more efficient than RSS for estimating the population mean, as long as the underlying distribution is symmetric. In MRSS, there are three sampling procedures-Level 0, Level 1 and Level 2- which can be used for sample selection. In this study, we provided formulas for calculating the inclusion probabilities in MRSS under these three sampling procedures. These inclusion probabilities indicate that the Level 2 sampling procedure generates greater inclusion probabilities for the middle values in the population than the other procedures. By using the Level 2 sampling procedure, we get more information about the population than in the other procedures.

ÖZET: Bu çalışmada, Deshpande ve diğ.(2006) tarafından sonlu yığın varsayımı altında önerilen 0. ve 2. düzey örnek seçim yöntemleri kullanılarak, medyan sıralı küme örneklemesinde içerilme olasılıklarının hesaplanmasına ilişkin genel bir formül geliştirilmiştir. Ayrıca, adı geçen örnek seçim yöntemleri ile elde edilen içerilme olasılıkları Özdemir ve Gökpinar (2008) tarafından verilen 1. düzey seçim yöntemine göre elde edilen içerilme olasılıkları ile farklı yığın ve örnek çapları altında karşılaştırılmıştır.

REFERENCES

- [1] Al-Saleh M.F., and Al-Omari, A. (2002). Multistage ranked set sampling. *Statistical Planning and Inference* 102:273-286.
- [2] Al Saleh, M.F., Samawi H.M. (2007). A note on inclusion probability in ranked set sampling and some of its variations. *Sociedad de Estadística e Investigación Operativa* 16:198-209.
- [3] Deshpande, J.V., Frey, J., Oztürk, O. (2006). Nonparametric ranked-set sampling confidence intervals for quantiles of a finite population. *Environmental and Ecological Statistics* 13:25-40.
- [4] McIntyre, G.A. (1952). A method of unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research* 3:385-390.
- [5] Muttalak, H.A. (1997). Median ranked set sampling. *Applied Statistical Science* 6:245-255.
- [6] Ozdemir, Y.A., Gokpinar, F. (2007). A Generalized formula for inclusion probabilities in ranked set sampling. *Hacettepe University Bulletin of Natural Sciences & Engineering Series B: Mathematics & Statistics* 36:89-99.
- [7] Ozdemir, Y.A., Gokpinar, F. (2008). A new Formula for inclusion probabilities in median ranked set Sampling. *Communications In Statistics:Theory and Methods*. 37, 2022-2033.
- [8] Samawi ,H.M., Ahmed, M.S. and Abu-Dayyeh, W. (1996). Estimation the population mean using extreme ranked set sampling. *Biometrical Journal* 38:577-586.
- [9] Takahasi, K., Futatsuya, M. (1988). Ranked set sampling from a finite population. *Proceeding of the Institute Statistical Mathematics* 36:55-68.
- [10] Takahasi, K., Futatsuya, M. (1998). Dependence between order statistics in samples from finite population and its application to ranked set sampling. *Proceeding of The Institute Statistical Mathematics* 50:49-70.

Current address: Gazi University Faculty of Arts and Science Department of Statistics Teknikokullar 06500 Ankara Turkey.

E-mail address: yaprak@gazi.edu.tr, fikri@gazi.edu.tr