Test (2007) 16: 198–209
DOI 10.1007/s11749-006-0009-7

ORIGINAL PAPER

A note on inclusion probability in ranked set sampling and some of its variations

M.F. Al-Saleh · H.M. Samawi

Received: 27 July 2004 / Accepted: 25 February 2005 /
Published online: 6 March 2007

Abstract Inclusion probabilities are design dependent and should be furnished with the design elements. Inclusion probability of an element in the population is the probability that the element will be chosen in a sample. In this paper the inclusion probabilities in the case of ranked set sampling design and some of its variations are furnished. This paper provides good and interesting examples of sampling designs for which the inclusion probabilities are not equal.

Keywords Inclusion probability · Ranked set sampling · Moving extreme ranked set sampling

Mathematics Subject Classification (2000) 46N30

1 Introduction

Simple random sampling (SRS) is the most basic sampling technique. It is the sampling technique for which all possible samples of the same size have the same probability of being chosen. If the population consists of N distinct elements, then a subset of size n is a SRS, if it is chosen so that each subset of size n has a probability of $1/\binom{N}{n}$ of being the chosen sample. As a consequence of this definition, each element of the population has a probability of $\frac{n}{N}$ of being included in the chosen sample. This probability is called the *inclusion probability*. Thus, in SRS, all elements of the population have equal inclusion probability. Other sampling plans also have this property; for example, systematic sampling and some special cases of cluster and stratified sampling. In general, probability sampling is a sampling technique for which each

M.F. Al-Saleh is currently at Qatar University (on leave), e-mail: malsaleh@qu.edu.qa.

M.F. Al-Saleh (✉) · H.M. Samawi
Department of Statistics, Yarmouk University, Irbid, Jordan
e-mail: m-saleh@yu.edu.jo

element in the population has a known inclusion probability. Let $\pi_N(k)$ be the probability that an element u_k of a population of size N will be chosen in a sample of size n. It is well known that $\sum_{k=1}^{N} \pi_N(k) = n$; thus, in the above mentioned sampling techniques the inclusion probabilities are all equal to the average inclusion probability of $\frac{n}{N}$. In SRS, all samples of size n are equally likely; hence, a very unrepresentative sample is as likely to appear as does a very representative one. For example, if a sample of size $n = 3$ is to be taken from a population of trees for the purpose of estimating the population average height then the likelihood of obtaining very short 3 trees or very tall 3 trees in the sample is the same as that of obtaining one tall, one short and one medium. This is because in SRS there is no control on which element enters the sample.

A more controlled sample (more representative) can be obtained using ranked set sampling technique. Ranked set sampling (RSS) was first suggested by McIntyre (1952) as a method for estimating pasture yields. The supporting mathematical theory was later provided by Takahasi and Wakimoto (1968). The RSS procedure consists of drawing m random samples of size m each from the population, and ranking each of them by judgment with respect to the characteristic of interest. Then the ith smallest observation from the ith set is chosen for actual quantification. The RSS consists of these m selected units. Although only m units out of m^2 are chosen for quantification, all units contribute information to the m quantified ones. In practice, to be able to rank by judgment, m should be small, 2 or 3, say. To get a sample of larger size the entire cycle may be repeated, if necessary, r times to produce a RSS of size $n = rm$. The mean of the RSS is an unbiased estimator of the population mean (μ) and is found to have smaller variance than the mean of a SRS of the same size. Assume that sampling is from an infinite population. Let $\hat{\mu}_{\text{RSS}}$ be the average of a RSS sample of size $n = rm$, and $\hat{\mu}_{\text{SRS}}$ be the average of a SRS of the same size. It was shown by Takahasi and Wakimoto (1968) that the relative efficiency of $\hat{\mu}_{\text{RSS}}$ w.r.t. $\hat{\mu}_{\text{SRS}}$ satisfies the following relation:

$$1 \le \text{Var}(\hat{\mu}_{\text{RSS}}; \hat{\mu}_{\text{SRS}}) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{Var}(\hat{\mu}_{\text{RSS}})} \le \frac{m+1}{2},$$

where the upper bound is achieved if and only if the distribution is uniform. However, relative efficiencies of other unimodal distributions are also close to the upper bound (Dell and Clutter 1972). For more about RSS see Kaur et al. (1995) and Patil et al. (1999). For recent work see Al-Saleh (2004), Perron and Sinha (2004), Carlos (2000), Al-Saleh and Zheng (2003), Al-Saleh and Al-Omary (2002), Zheng and Al-Saleh (2002) and Al-Saleh and Al-Kadiri (2000).

All of the authors above assume that the RSS is taken from an infinite population. In this paper we consider sampling from a finite population. Only few authors have discussed RSS in finite populations. It was first considered by Takahasi and Futatsuya (1988). They showed that $\hat{\mu}_{\text{RSS}}$ is unbiased for the population mean and found explicit formula for its efficiency for a set size $m = 2$. Their results were extended to more general finite populations and to larger set-sizes by Patil et al. (1999). The efficiency was shown to be ≥ 1, under some mild conditions, by Takahasi and Futatsuya (1998).

To obtain the RSS sample from a finite population these authors used the same procedure that is used for infinite populations: to obtain an RSS of size $n = rm$ a SRS of a sample of size rm^2 is partitioned into mr sets of size m each. For each of the first

r sets the minimum is identified; for each of the second r sets the second minimum is identified; etc. The RSS obtained consists of r minima, r second minima, $\dots, r$ maxima. Thus, from the rm^2 elements only rm are chosen for actual quantifications and the others are ignored. While this seems to be okay if the population is infinite, it might be a waste and may cause some problems if the population is finite and N is not very large. Therefore, the following adjusted procedure will be used for obtaining the RSS:

1. A SRS of size m is selected (without replacement) from the population and the minimum of the sample with respect to the characteristic of interest is identified by judgment. All other elements are returned to the population.
2. A second SRS of size m is selected (without replacement) from the population and the second minimum of the sample is identified by judgment. All other elements are returned to the population.
3. In the ith step, the ith minimum of the ith chosen SRS is identified by judgment, $i = 1, 2, \dots, m$.

The m identified elements make up a RSS of size m. The entire cycle may be repeated, if necessary, r times to produce a RSS sample of size $n = rm$.

Note 1 The only difference between this adjusted procedure and the previous one is that after each step, all elements *except the one chosen for quatification*, are returned to the population before the start of the process of choosing the next element.

Each sampling method implies a set of inclusion probabilities. In this paper, our main interest is in the inclusion probability of this sampling technique and some of its variations. To the best of our knowledge, the inclusion probability of this technique has not been obtained yet.

2 Inclusion probability

Knowing the value of the inclusion probability for each element of the population is very important. The following are some of the reasons for the need of knowing the inclusion probability in RSS:

1. Inclusion probabilities give an insight on how the RSS design has more control on which element enters the sample as compared with the SRS.
2. Inclusion probabilities are needed when one wants to consider some unbiased estimators of the population mean or total such as the Horvitz–Thompson estimator. If π_i is the inclusion probability of the element u_i in the population and s is a sample taken from the population then the Horvitz–Thompson estimator of the population total T is

$$\hat{T}_{\mathrm{HT}} = \sum_{i \in s} \frac{u_i}{\pi_i}.$$

$\hat{T}_{\mathrm{HT}}$ is the only unbiased estimator of T in the class of estimators

$$\left\{ \sum_{i \in s} b_i u_i : b_i \in \mathbb{R} \right\}.$$

To use $\hat{T}_{\text{HT}}$ in practice π_i should be known in advance for each u_i in the population; which is rarely the case. It will be seen that, in order to find π_i, we need to find the rank of u_i in the population. However, there are some situations in which the rank of u_i can be approximately obtained using some other concomitant variable w_i. For example, if u_i is the size of tree i in a population of N trees then the rank of u_i can be approximated using the height w_i, which is highly correlated with u_i. (Largest trees tend to produce the most timber and contribute the most to the variability of volumes.) See Mukhopadhyay (2000).

3. Knowing the inclusion probabilities of a design facilitates the sampling from an available population. For example, if the values of a large population are given and stored on a computer, then one of the main purposes is to summarize the data using some suitable methods of data reduction. If it is decided to summarize the data set using repeated RSS, as in Al-Saleh (2004), it is easier, in terms of time and effort, to write a program to repeatedly sample RSSs from the population using the inclusion probabilities rather than using the above procedure. Also the same applies if we are studying the properties of some complex estimator using simulation.
4. Most of the examples taught in a classroom for unequal probability samples are artificial. The content of the paper provides good and interesting non-artificial examples of sampling designs for which the inclusion probabilities are not equal.
5. Knowing the inclusion probabilities of two or more sampling designs can help one decide which design suits one's need. In the trees example, see point 2 above, since large trees contribute more to the total size of the population, one may choose a design that assigns larger inclusion probabilities to trees of larger heights; hence MERSS, described below with moving maximum instead of minimum, is a more suitable design. If prior information suggests that the sampled population is symmetric then a design that gives large inclusion probabilities to extreme values is preferred, for example extreme RSS. If the variability among the elements of the population is low then a design that assigns equal inclusion probabilities is preferable; hence SRS is a more suitable design.

In the next section the inclusion probability for each element in the population is obtained for the most common cases of $m = 2$ and $m = 3$, when the design is RSS. The case of a general m is also addressed. In Sect. 3 this probability is obtained for two variations of RSS. Conclusions are given in Sect. 4.

3 Inclusion probability in RSS

Assume that the population of interest consists of the elements $\{u_1, u_2, \ldots, u_N\}$. An RSS sample of size $n = rm$ is to be taken from the population using the procedure outlined above, where m is the set size and r is the number of cycles. The interest is in the inclusion probability of each u_i, $i = 1, \ldots, N$. For simplicity, assume that the population consists of distinct elements ($u_1 < u_2 < \cdots < u_N$, say). Assume first that $r = 1$ and let S be the chosen RSS sample. Throughout the paper it is found instructive to use binomial coefficients without simplification.

Case 1: $m=2$.

For $k=1,2,\dots,N$, let $\pi_N(k)=p(u_k\in S)$ be the inclusion probability of the kth element of the population. For u_k to be in S, u_k must be the minimum of the first SRS of size 2 taken from the N elements (A_1) or the maximum of the second sample of size 2 taken from the $N-1$ remaining elements (A_2). Let the probabilities of these two events be $\pi_N^{(1)}(k)$ and $\pi_N^{(2)}(k)$, respectively. Clearly,

$$\pi_N(k)=\pi_N^{(1)}(k)+\pi_N^{(2)}(k) \tag{3.1}$$

and

$$\pi_N^{(2)}(k)=p\big(A_2|A_1^c\big)p\big(A_1^c\big)=p\big(A_2|A_1^c\big)\big(1-\pi_N^{(1)}(k)\big). \tag{3.2}$$

Now, for u_k to be the minimum of the first sample, the other element should be chosen from the $N-k$ elements; thus, using binomial coefficients notation, we have

$$\pi_N^{(1)}(k)=\frac{\binom{N-k}{1}}{\binom{N}{2}}. \tag{3.3}$$

To obtain the value of $\pi_N^{(2)}(k)$, A_1^c can be further partitioned to $A_{11}(A_{12})$: the minimum of the first SRS of size 2 taken from the N elements is $<u_k$ ($>u_k$). Then,

$$\pi_N^{(2)}(k)=p(A_2|A_{11})p(A_{11})+p(A_2|A_{12})p(A_{12}). \tag{3.4}$$

It can be easily seen using some elementary combinatorics that

$$p(A_2|A_{11})=\frac{\binom{k-2}{1}}{\binom{N-1}{2}} \quad \text{and} \quad p(A_2|A_{12})=\frac{\binom{k-1}{1}}{\binom{N-1}{2}}. \tag{3.5}$$

The event A_{11} consists of those samples for which u_k is an element of (there are $\binom{k-1}{1}$ of them) and those samples for which u_k is not an element of (there are $\binom{k-1}{2}+\binom{k-1}{1}\binom{N-k}{1}$ of them). Thus,

$$p(A_{11})=\frac{\binom{k-1}{2}+\binom{k-1}{1}\binom{N-k}{1}+\binom{k-1}{1}}{\binom{N}{2}}. \tag{3.6}$$

Similarly, $p(A_{12})$ can be shown to be

$$p(A_{12})=\frac{\binom{N-k}{2}}{\binom{N}{2}}. \tag{3.7}$$

Therefore, the value of $\pi_N^{(2)}(k)$ is given by

$$\pi_N^{(2)}(k)=\frac{\binom{k-2}{1}}{\binom{N-1}{2}}\times\frac{\binom{k-1}{2}+\binom{k-1}{1}\binom{N-k}{1}+\binom{k-1}{1}}{\binom{N}{2}}+\frac{\binom{k-1}{1}}{\binom{N-1}{2}}\times\frac{\binom{N-k}{2}}{\binom{N}{2}}. \tag{3.8}$$

Table 1 RSS inclusion probability for $m = 2$

N	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
3	0. 667	0.333	1.00	–	–	–	–	–	–	–
4	0.500	0.389	0.444	0.667	–	–	–	–	–	–
5	0.400	0.350	0.350	0.400	0.500	–	–	–	–	–
6	0.333	0.307	0.300	0.313	0.347	0.400	–	–	–	–
7	0.286	0.270	0.263	0.267	0.280	0.302	0.333	–	–	–
8	0.250	0.240	0.235	0.235	0.240	0.250	0.265	0.286	–	–
9	0.222	0.215	0.211	0.210	0.212	0.217	0.225	0.236	0.250	–
10	0.200	0.195	0.192	0.191	0.191	0.194	0.198	0.204	0.212	0.222

Actually, if Y_1 and Y_2 are, respectively, the first and second element of the RSS sample then the pdf of Y_1 is $f_1(k) = \pi_N^{(1)}(k)$, (i.e. $\pi_N^{(1)}(k)$ is the pdf of the first order statistic) and the pdf of the second order statistic Y_2 is $f_2(k) = \pi_N^{(2)}(k)$.

Using (3.1) and (3.6), the inclusion probability for any element u_k, $k = 1, \ldots, N$, of the population is given by

$$\begin{aligned}
\pi_N(k) &= \pi_N^{(1)}(k) + \pi_N^{(2)}(k) \\
&= \frac{\binom{N-k}{1}}{\binom{N}{2}} + \frac{\binom{k-2}{1}}{\binom{N-1}{2}} \times \frac{\binom{k-1}{2} + \binom{k-1}{1}\binom{N-k}{1} + \binom{k-1}{1}}{\binom{N}{2}} \\
&\quad + \frac{\binom{k-1}{1}}{\binom{N-1}{2}} \times \frac{\binom{N-k}{2}}{\binom{N}{2}}.
\end{aligned} \tag{3.9}$$

It can be simplified to

$$\pi_N(k) = 2\left(\frac{3k^2 - (2N+5)k + 7N - 4N^2 + N^3}{N(N-1)^2(N-2)}\right). \tag{3.10}$$

It is not difficult to show that $\pi_N(k)$ is decreasing for $k < \frac{1}{3}N + \frac{5}{6}$ and increasing for $k > \frac{1}{3}N + \frac{5}{6}$ with the minimum at the nearest integer to $\frac{1}{3}N + \frac{5}{6}$. Also it can be checked easily that $\sum_{k=1}^{N} \pi_N(k) = n = 2$ (as expected); $\pi_N(1) = \frac{2}{N}$ and $\pi_N(N) = \frac{2}{N-1}$. For illustration, Table 1, gives the inclusion probability for some values of N. Clearly, the inclusion probabilities are not the same for the population elements. In this sense RSS is different from the other sampling plans discussed in the introduction. Elements in the two extremes get higher probabilities of selection than elements in the middle of the population. For example, when $N = 3$ the inclusion probabilities are, respectively, 0.667, 0.333 and 1.00, which means that u_3 is an element of any RSS of size 2; u_1 has more chances (twice as much) to be the second element of the RSS sample than u_2.

Case 2: $m = 3$.

For u_k to be in the chosen sample, u_k must be, for some $i \in \{1, 2, 3\}$, the ith order statistics of the ith SRS of size 3 taken from the remaining $N - i + 1$ elements (A_i). Let the probabilities of these three events be $\pi_N^{(1)}(k)$, $\pi_N^{(2)}(k)$ and $\pi_N^{(3)}(k)$, respectively.

Now,

$$\pi_N^{(1)}(k) = \frac{\binom{N-k}{2}}{\binom{N}{3}}, \tag{3.11}$$

$\pi_N^{(2)}(k) = p(A_2|A_1^c)p(A_1^c) = p(A_2\ A_{11})p(A_{11}) + p(A_2\ A_{12})p(A_{12})$, where A_{11} and A_{12} are as in Case 1 (except that the sample size is 3 instead of 2). Thus, using the same argument as in Case 1, $\pi_N^{(2)}(k)$ can be shown to be

$$\begin{aligned}\pi_N^{(2)}(k) = {} & \frac{\binom{k-2}{1}\binom{N-k}{1}}{\binom{N-1}{3}} \frac{\binom{k}{3} + \binom{k-1}{1}\binom{N-k}{2} + \binom{k}{2}\binom{N-k}{1}}{\binom{N}{3}} \\ & + \frac{\binom{k-1}{1}\binom{N-k-1}{1}}{\binom{N-1}{3}} \frac{\binom{N-k}{3}}{\binom{N}{3}}. \end{aligned} \tag{3.12}$$

Let B: [u_k is the maximum of the third chosen sample], B_{11}: [min $< u_k$], B_{12}: [median $< u_k$], B_{21}: [min $> u_k$], B_{22}: [median $> u_k$]. Then

$$\begin{aligned}\pi_N^{(3)}(k) &= p(B|B_{11} \cap B_{12})p(B_{11} \cap B_{12}) \\ &\quad + p(B|B_{21} \cap B_{22})p(B_{21} \cap B_{22}) \\ &\quad + p(B|B_{21} \cap B_{12})p(B_{21} \cap B_{12}) \\ &\quad + p(B|B_{11} \cap B_{22})p(B_{11} \cap B_{22}) \\ &= p_1 + p_2 + p_3 + p_4. \end{aligned} \tag{3.13}$$

It can be shown that

$$p_1 = \frac{\binom{k-3}{2}}{\binom{N-2}{3}} \frac{\binom{k-2}{2} + \binom{k-2}{2}\binom{N-k}{1} + \binom{k-2}{3}}{\binom{N-1}{3}} \frac{\binom{k}{3} + \binom{k-1}{1}\binom{N-k}{2} + \binom{k}{2}\binom{N-k}{1}}{\binom{N}{3}},$$

$$p_2 = \frac{\binom{k-1}{2}}{\binom{N-2}{3}} \frac{\binom{N-k-1}{2} + \binom{N-k-1}{3} + \binom{k-1}{1}\binom{N-k-1}{2}}{\binom{N-1}{3}} \frac{\binom{N-k}{3}}{\binom{N}{3}},$$

$$p_3 = \frac{\binom{k-2}{2}}{\binom{N-2}{3}} \frac{\binom{k-1}{2} + \binom{k-1}{2}\binom{N-k-1}{1} + \binom{k-1}{3}}{\binom{N-1}{3}} \frac{\binom{N-k}{3}}{\binom{N}{3}}$$

and

$$p_4 = \frac{\binom{k-2}{2}}{\binom{N-2}{3}} \frac{\binom{N-k}{2} + \binom{N-k}{2}\binom{k-2}{1} + \binom{N-k}{3}}{\binom{N-1}{3}} \frac{\binom{k}{3} + \binom{k-1}{1}\binom{N-k}{2} + \binom{k}{2}\binom{N-k}{1}}{\binom{N}{3}}.$$

Table 2 RSS inclusion probability for $m = 3$

N	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
5	0.600	0.350	0.550	0.500	1.00	–	–	–	–	–
6	0.500	0.360	0.414	0.501	0.475	0.750	–	–	–	–
7	0.428	0.343	0.354	0.407	0.427	0.440	0.600	–	–	–
8	0.375	0.319	0.317	0.344	0.368	0.377	0.400	0.500	–	–
9	0.333	0.295	0.288	0.302	0.320	0.331	0.340	0.362	0.428	–
10	0.300	0.272	0.265	0.272	0.283	0.294	0.301	0.309	0.329	0.375

Hence,

$$\pi_N(k) = \pi_N^{(1)}(k) + \pi_N^{(2)}(k) + \pi_N^{(3)}(k). \tag{3.14}$$

Table 2, gives the values of the inclusion probability for some values of N.

Case 3: general m.

Let $\{Y_1, Y_2, \ldots, Y_m\}$ be the elements of the RSS. Let E_{ik} denotes the event $[Y_i = u_k]$, then

$$\pi_N(k) = \sum_{i=1}^{m} p(E_{ik}) = \sum_{i=1}^{m} p\left(E_{ik} | \bigcap_{j=1}^{i-1} E_{jk}^c\right) p\left(\bigcap_{j=1}^{i-1} E_{jk}^c\right), \tag{3.15}$$

with $E_{0k} = \phi$. Thus, the inclusion probability for each element in the population can be, in principle, obtained for any given m. However, finding a closed form for $\pi_N(k)$ will be tedious for large m. For all practical purposes of using RSS, $m = 2$ or 3 are sufficient.

If $r > 1$, then $P(u_k \in S) = \pi_N(k) = \sum_{i=1}^{r} P(u_k \text{ is chosen in the } i\text{th cycle})$, which can be evaluated in a similar fashion. Note that in the ith cycle, an RSS of size m is obtained from a population of size $N - m(i-1)$.

4 Inclusion probability for some variations of RSS

4.1 Moving extreme RSS

Moving extreme RSS (MERSS) was introduced and discussed by Al-Odat and Al-Saleh (2000) and also by Al-Saleh and Al-Hadrami (2003). MERSS is a variation of RSS that allows the use of a larger set size m. In MERSS, for $i = 1, \ldots, m$, a SRS of size i is taken from a population and the minimum of each sample, with respect to the characteristic of interest, is identified by judgment obtaining a moving minimum RSS. Alternatively, we can identify the maximum at each step, getting a moving maximum RSS of size m instead. In this section we consider the inclusion probability of MERSS when identifying the minimum at each step. After each step all remaining elements are returned to the population before the next step.

Case 1: $m = 2$.

In this case the first element of the MERSS is obtained by taking one element randomly from the population; the second element is the minimum of a SRS of size 2 taken from the $N - 1$ remaining elements of the population. The inclusion probability of u_k can be easily shown to be (in binomial notation):

$$\pi_N(k) = \frac{\binom{1}{1}}{\binom{N}{1}} + \frac{\binom{k-1}{1}}{\binom{N}{1}} \frac{\binom{N-k}{1}}{\binom{N-1}{2}} + \frac{\binom{N-k}{1}}{\binom{N}{1}} \frac{\binom{N-k-1}{1}}{\binom{N-1}{2}} \tag{4.1}$$

which can be simplified to $\frac{3N-2k-1}{N(N-1)}$.

Clearly, $\pi_N(k)$ is decreasing in k with the maximum value of $\frac{3}{N}$ at $k = 1$ and the minimum value of $\frac{1}{N}$ at $k = N$.

Case 2: $m = 3$.

In this case the probability that u_k is in the chosen sample is the probability that u_k is the minimum of a SRS of size 1 taken from the N elements, the minimum of a SRS of size 2 taken from the $N - 1$ remaining elements or the minimum of a SRS of size 3 taken from the $N - 2$ remaining elements. Denote these three probabilities by $\pi_N^{(1)}(k)$, $\pi_N^{(2)}(k)$, $\pi_N^{(3)}(k)$, respectively, then $\pi_N(k) = \pi_N^{(1)}(k) + \pi_N^{(2)}(k) + \pi_N^{(3)}(k)$. Clearly, $\pi_N^{(1)}(k)$ is the first term of Eq. (4.1) above and $\pi_N^{(2)}(k)$ is the last two terms of Eq. (4.1). Also $\pi_N^{(3)}(k) = p(u_k = Y_3)$, where Y_3 is the third element in the sample. This probability can be factored as

$$\begin{aligned}
\pi_N^{(3)}(k) &= p(Y_1 < u_k, Y_2 < u_k, Y_3 = u_k) \\
&\quad + p(Y_1 < u_k, Y_2 > u_k, Y_3 = u_k) \\
&\quad + p(Y_1 > u_k, Y_2 < u_k, Y_3 = u_k) \\
&\quad + p(Y_1 > u_k, Y_2 > u_k, Y_3 = u_k) \\
&= p_1 + p_2 + p_3 + p_4,
\end{aligned} \tag{4.2}$$

where

$$p_1 = \frac{\binom{k-1}{1}}{\binom{N}{1}} \frac{\binom{k-2}{2} + \binom{k-2}{1}\binom{N-k+1}{1}}{\binom{N-1}{2}} \frac{\binom{N-k}{2}}{\binom{N-2}{3}}, \tag{4.3}$$

$$p_2 = \frac{\binom{k-1}{1}}{\binom{N}{1}} \frac{\binom{N-k}{2}}{\binom{N-1}{2}} \frac{\binom{N-k-1}{2}}{\binom{N-2}{3}}, \tag{4.4}$$

$$p_3 = \frac{\binom{N-k}{1}}{\binom{N}{1}} \frac{\binom{k-1}{2} + \binom{k-1}{1}\binom{N-k}{1}}{\binom{N-1}{2}} \frac{\binom{N-k-1}{2}}{\binom{N-2}{3}} \tag{4.5}$$

Table 3 MERSS inclusion probability for $m = 2, 3$

N	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9
$m = 2$									
3	1.00	0.667	0.333	–	–	–	–	–	–
4	0.750	0.583	0.417	0.250	–	–	–	–	–
5	0.600	0.500	0.400	0.300	0.200	–	–	–	–
6	0.500	0.433	0.367	0.300	0.233	0.167	–	–	–
7	0.428	0.381	0.333	0.286	0.238	0.190	0.143	–	–
$m = 3$									
5	1.00	0.900	0.600	0.300	0.200	–	–	–	–
6	0.875	0.758	0.579	0.388	0.233	0.167	–	–	–
7	0.771	0.667	0.539	0.406	0.284	0.190	0.143	–	–
8	0.688	0.598	0.500	0.400	0.305	0.223	0.161	0.125	–
9	0.619	0.544	0.465	0.386	0.311	0.242	0.184	0.139	0.111

and

$$p_4 = \frac{\binom{N-k}{1}}{\binom{N}{1}} \frac{\binom{N-k-1}{2}}{\binom{N-1}{2}} \frac{\binom{N-k-2}{2}}{\binom{N-2}{3}}. \tag{4.6}$$

Table 3 contains the inclusion probability for some values of N.

Case 3: general m.

As in RSS, in this case $\pi_N(k)$ can be written as $\pi_N(k) = \sum_{i=1}^{m} \pi_N^{(i)}(k)$, where $\pi_N^{(i)}(k) = p(Y_i = u_{k)}$. Each of $\pi_N^{(i)}(k)$ can be obtained in the same manner as in Case 2 above.

Note that if instead we use the moving maximum RSS then the inclusion probability of u_k is $\pi_N^*(k) = \pi_N(N - k + 1)$.

4.2 Extreme RSS

Extreme RSS, ERSS, is similar to MERSS except that the set size m is kept fixed. This method was suggested first by Stokes (1980). To obtain a sample of size m using ERSS, a SRS of size m is taken and the maximum (or the minimum) is identified by judgment. This step is repeated m times yielding a ERSS of size m. The procedure was modified and investigated further by Samawi et al. (1996). The inclusion probability can be easily obtained using the same method as in Sect. 3. For example, if the minimum is the one identified at each step then, for $m = 2$, the inclusion probability can be easily obtained as:

$$\pi_N(k) = \frac{\binom{N-k}{1}}{\binom{N}{2}} + \frac{\binom{k-1}{2} + \binom{k-1}{1}\binom{N-k+1}{1}}{\binom{N}{2}} \frac{\binom{N-k}{1}}{\binom{N-1}{2}} + \frac{\binom{N-k}{1}}{\binom{N}{2}} \frac{\binom{N-k-1}{1}}{\binom{N-1}{2}}. \tag{4.7}$$

Note that $\pi_N(N) = 0$ and, in general, $\pi_N(k) = 0$ for $k = N, N-1, \ldots, N-m+2$. The modification of Samawi et al. (1996) and Al-Saleh and Al-Hadrami (2003), overcomes this difficulty.

5 Concluding remarks

The inclusion probability of each element in the population is important to know in advance for each used sampling design. It is determined by the design and should be furnished with the design points. "Figuratively speaking, the inclusion probabilities can be described as the sampling method's "footprint" on the population elements; each sampling method implies a set of inclusion probabilities" (Tryfos 1996). For example when $N = 4$, $m = 2$, $(0.500, 0.500, 0.500, 0.500)$; $(0.500, 0.389, 0.444, 0.667)$; $(0.750, 0.583, 0.417, 0.250)$; $(0.833, 0.722, 0.444, 0.000)$ are respectively, the *foot prints* of SRS, RSS, MERSS and ERSS on the population elements. The content of the paper provides good and interesting non-artificial examples of sampling designs for which the inclusion probabilities are not equal.

Acknowledgements Most of the material in Sect. 2 of this revised version is motivated by comments of the editor, the associate editor and a referee. The authors wish to express appreciation to them for their valuable comments and constructive suggestions on a preliminary draft of this paper.

References

Al-Odat M, Al-Saleh MF (2000) A variation of ranked set sampling. J Appl Stat Sci 10:137–146
Al-Saleh MF (2004) Steady state ranked set sampling and parametric estimation. J Stat Plan Inference 123:83–95
Al-Saleh MF, Al-Hadrami S (2003) Estimation of the mean of the exponential distribution using moving extremes ranked set sampling. Stat Pap 44:367–382
Al-Saleh MF, Al-Kadiri M (2000) Double ranked set sampling. Stat Probab Lett 48:205–212
Al-Saleh MF, Al-Omary A (2002) Multistage ranked set sampling. J Stat Plan Inference 102:273–286
Al-Saleh MF, Zheng G (2003) Controlled sampling using ranked set sampling. J Nonparametr Stat 15:505–516
Carlos NB (2000) Estimation of the mean in ranked set sampling with non responses. Metrika 56:171–179
Dell TR, Clutter JL (1972) Ranked set sampling theory with order statistics background. Biometrics 28:545–555
Kaur A, Patil GP, Sinha AK, Taillie C (1995) Ranked set sampling: an annotated bibliography. Environ Ecol Stat 2:25–54
McIntyre GA (1952) A method for unbiased selective sampling using ranked sets. Aust J Agric Res 3:385–390
Mukhopadhyay P (2000) Topics in survey sampling. Springer, New York
Patil GP, Sinha AK, Taillie C (1999) Ranked set sampling: a bibliography. Environ Ecol Stat 6:91–98
Perron F, Sinha BK (2004) Estimation of variance based on ranked set sampling. J Stat Plan Inference 120:21–28
Samawi HM, Ahmad MS, Abu-Dayyeh W (1996) Estimating the population mean using extreme ranked set sampling. Biom J 38:577–586
Stokes SL (1980) Inferences on the correlation coefficient in bivariate normal population from ranked set sampling. J Am Stat Assoc 75:989–995
Takahasi K, Futatsuya M (1988) Ranked set sampling from a finite population. Proc Inst Stat Math 36:55–68

Takahasi K, Futatsuya M (1998) Dependence between order statistics in samples from finite population and its application to ranked set sampling. Ann Inst Stat Math 50:49–70

Takahasi K, Wakimoto K (1968) On unbiased estimates of the population mean based on the sample stratified by means of ordering. Ann Inst Stat Math 20:1–31

Tryfos P (1996) Sampling methods for applied research. Wiley, New York

Zheng G, Al-Saleh MF (2002) Modified maximum likelihood estimators based on ranked set sampling. Ann Inst Stat Math 54:641–658