ON RANDOM SAMPLING WITHOUT REPLACEMENT FROM A FINITE POPULATION

SUBHASH C. KOCHAR[1] AND RAMESH KORWAR[2]*

[1]*Indian Statistical Institute, 7 SJS Sansanwal Marg, New Delhi-110016, India*
[2]*Department of Mathematics and Statistics, Lederle Graduate Research Center, University of Massachusetts, Amherst, MA 01003, U.S.A.*

(Received June 28, 1999; revised April 4, 2000)

Abstract. We consider the three progressively more general sampling schemes without replacement from a finite population: simple random sampling without replacement, Midzuno sampling and successive sampling. We (i) obtain a lower bound on the expected sample coverage of a successive sample, (ii) show that the vector of first order inclusion probabilities divided by the sample size is majorized by the vector of selection probabilities of a successive sample, and (iii) partially order the vectors of first order inclusion probabilities for the three sampling schemes by majorization. We also show that the probability of an ordered successive sample enjoys the arrangement increasing property and for sample size two the expected sample coverage of a successive sample is Schur convex in its selection probabilities. We also study the spacings of a simple random sample from a linearly ordered finite population and characterize in several ways a simple random sample.

Key words and phrases: Simple random sampling without replacement, Midzuno sampling, sample coverage, inclusion probabilities, arrangement increasing functions, exchangeable random variables, majorization, successive sampling, Schur convexity.

1. Introduction

Consider a finite population $\mathcal{U} = \{1, \dots, N\}$ of N distinct units. In this paper we present some new results on sampling without replacement from this population. There are a variety of such sampling schemes and the most general of them is successive sampling. In successive sampling, draws are made with replacement one by one, and at each draw a unit k has probability $p_k (> 0)$, $k = 1, \dots, N$, $\sum_{k=1}^{N} p_k = 1$, of being chosen. Draws are made until n distinct units are chosen, any repetitions being discarded. An *ordered* sample $s = (i_1, \dots, i_n)$ has probability

$$P(s) = \prod_{j=1}^{n} \frac{p_{i_j}}{(1 - \sum_{\ell=0}^{j-1} p_{i_\ell})} \tag{1.1}$$

of being chosen. Here $p_{i_0} \equiv 0$. The probability of an *unordered* sample $S = \{k_1, \dots, k_n\}$ is obtained by summing the probabilities of the $n!$ ordered samples given by the $n!$

*Partial work done while visiting ISI, New Delhi. The second author wishes to sincerely thank ISI for extending an invitation and making available all its facilities.

permutations of the elements of $\boldsymbol{S}$. That is,

$$P(\boldsymbol{S}) = \sum_r \prod_{j=1}^{n} \frac{p_{i_j}}{(1 - \sum_{\ell=0}^{j-1} p_{i_\ell})} \tag{1.2}$$

where $\sum_r$ is taken over all permutations $r = (i_1, \ldots, i_n)$ of the elements k_j, $k_j \in \boldsymbol{S}$. For more details on successive sampling, we refer the interested reader to Hájek (1981). We shall use upper case bold letters to denote *unordered* samples and lower case bold letters for *ordered* samples.

An interesting case of successive sampling is that of probability proportional to size without replacement (PPSWOR). In this case $p_k \propto x_k$, $x_k > 0$, a size measure of the k-th unit, for $k = 1, \ldots, N$. See Rao *et al.* (1991) for some interesting properties of PPSWOR sampling. In case of a simple random sample without replacement (SRSWOR) $p_k = 1/N$, for all k. Midzuno sampling, introduced by Midzuno (1950) is a cross between successive sampling and SRSWOR. As in successive sampling, draws are made one by one and with replacement. At the first draw a unit k has probability p_k, $k = 1, \ldots, N$ with $\sum_{k=1}^{N} p_k = 1$ of being chosen, but in subsequent draws has probability $1/N$ of being chosen, any repetitions being discarded. For Midzuno sampling it can be shown that $P(\boldsymbol{S}) = \sum_{k \in \boldsymbol{S}} p_k / \binom{N-1}{n-1}$.

In Section 2 we discuss some basic properties of successive sampling. We prove that the probability $P(\boldsymbol{s})$ of an ordered sample $\boldsymbol{s}$ is *arrangement increasing*, when $p_1 \geq \cdots \geq p_N$. We give a lower bound for $E[\sum_{k \in \boldsymbol{S}} p_k]$, the so called *expected sample coverage*. For sample size two, we show that the expected sample coverage is Schur convex in its selection probabilities.

Throughout the paper we use the word "majorization" to mean a special partial order relation among vectors (and not a component-wise partial order relation). For an arbitrary set $\boldsymbol{S}$ of $\mathcal{U}$ let $\pi(\boldsymbol{S})$ denote the probability of including $\boldsymbol{S}$ in the sample. For $\boldsymbol{S} = \{i\}$, $\pi(\boldsymbol{S})$ will be denoted by π_i and $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$ is called the vector of first order inclusion probabilities. A major result of Section 3 is that $\boldsymbol{\pi}/n$ is majorized by $\boldsymbol{p} = (p_1, \ldots, p_N)$, the vector of selection probabilities in case of successive sampling. Many important results previously discussed in the literature immediately follow from this result. In this section we also compare the three sampling schemes by partially ordering their vectors of first order inclusion probabilities by majorization.

In Section 4 we study spacings of a simple random sample without replacement (SRSWOR) from a linearly ordered finite population. We show that the spacings are exchangeable and the vector of spacings has multivariate increasing failure rate distribution. Section 5 contains some characterizations of a SRSWOR from a linearly ordered population by the exchangeability of spacings.

2. Properties of successive sampling

In this section we discuss some basic properties of ordered and unordered successive samples which will be used later on in the paper. We study the effect of changes in $\boldsymbol{p}$, the vector of selection probabilities on the vector $\boldsymbol{\pi}$ of inclusion probabilities.

The *sample coverage* for our problem is defined as $\Lambda(\boldsymbol{S}) = \sum_{k \in \boldsymbol{S}} p_k$ where $\boldsymbol{S}$ is a successive sample. In case of PPSWOR, $\Lambda(\boldsymbol{S}) = \sum_{k \in \boldsymbol{S}} x_k / \sum_{k=1}^{N} x_k$ is the proportion of the total measure captured by the sample. This quantity is of interest in many applications. Andreatta and Kaufman (1986) describe a situation of prospecting for oil.

The successive sample consists of magnitudes of discovered deposits and the expected proportion of total discovered deposits to the total (discovered and undiscovered) deposits is of interest. Below we obtain a lower bound on $E[\Lambda(S)]$. First we prove some preliminary results.

Let S_1 and S_2 be two unordered successive samples of the same size. One intuitively feels that if S_1 has higher sample coverage probability than S_2, then $P(S_1) > P(S_2)$. The following example shows that this is not true in general.

Example 2.1. Let $N = 5$ with $p_1 = 1/26$, $p_2 = 3/26$, $p_3 = 10/26$, $p_4 = p_5 = 6/26$. Let $S_1 = \{1,2,3\}$ and $S_2 = \{1,4,5\}$. Then $\Lambda(S_1) = 14/26$ while $\Lambda(S_2) = 13/26$. Thus $\Lambda(S_1) > \Lambda(S_2)$ and yet $P(S_1) = .0217 < .023 = P(S_2)$.

However, the following special case is true.

THEOREM 2.1. *Let S_1 and S_2 be two successive samples of the same size. Suppose S_1 and S_2 have all but one unit in common, with the nonoverlapping units being i and j, respectively. Then*

$$P(S_1) \geq P(S_2) \quad \text{if and only if} \quad p_i \geq p_j,$$

with equality holding if and only if $p_i = p_j$.

Its proof can be found in Rao *et al.* (1991) and it also follows from Lemma 3.1 of Kochar and Korwar (1996).

It follows from the above discussion that, in general, it may not be possible to completely order all possible successive samples of same size according to their probabilities of inclusion. The proof of the next corollary easily follows from the above theorem.

COROLLARY 2.1. *In successive sampling, for $i \neq j$,*

$$\pi_i \geq \pi_j \text{ if and only if } p_i \geq p_j,$$

with equality holding if and only if $p_i = p_j$ for $i, j = 1, \ldots, N$.

A similar result holds for the second order inclusion probabilities $\pi(ij)$ that both i and j are in the successive sample.

COROLLARY 2.2. *In successive sampling, for $i \neq j$, $i \neq k, i, j, k = 1, \ldots, N$,*

$$\pi(ik) > \pi(jk) \quad \text{if and only if } p_i \geq p_j,$$

with equality holding if and only if $p_i = p_j$ for $i, j = 1, \ldots, N$.

The next problem we consider is to see whether it is possible to rank all $n!$ possible ordered samples of the same set of n units according to their probabilities. To answer this question we need the following concept of *arrangement increasing* functions.

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ be two vectors. We say that $\boldsymbol{x}$ *is better arranged than* $\boldsymbol{y}$ (written as $\boldsymbol{x} \succeq^a \boldsymbol{y}$) if $\boldsymbol{x}$ can be obtained from $\boldsymbol{y}$ through successive pairwise interchanges of its components, with each interchange resulting in an increasing order of the two interchanged components; e.g. $(4,1,5,3) \succeq^a (4,3,5,1) \succeq^a (4,5,3,1)$. Note that $\succeq^a$ ordering is only a partial ordering over n-tuples. A function $g : \mathcal{R}^n \to \mathcal{R}$ that preserves the ordering $\succeq^a$ is called an *arrangement increasing* function and is denoted by $g \in \mathcal{AI}$ if $x \succeq^a y \Rightarrow g(x) \geq g(y)$. See Marshall and Olkin ((1979), p. 158) for further properties of such functions.

We prove in the next theorem that the probability of an ordered sample $s = (i_1, \ldots, i_n)$ as given by (1.1) is *arrangement increasing* when the selection probabilities p_i's are ordered from the largest to the smallest.

THEOREM 2.2. *Let $p_1 \geq \cdots \geq p_N$. Then the probability $P(s)$ of an ordered sample s is $\mathcal{AI}$.*

PROOF. Let $s = (i_1, \ldots, i_n)$ and $s' = (i'_1, \ldots, i'_n)$ be two ordered samples such that s' contains exactly one inversion of a pair of coordinates which occur in the natural order in s. Let $k < \ell$ and $i'_j = i_j$, $j \neq k$, $j \neq \ell$, $j = 1, \ldots, n$; $i'_k = i_\ell$, $i'_\ell = i_k$, $i_k < i_\ell$. Then from (1.1),

$$
\begin{aligned}
P(s') &= \frac{\prod_{j=1}^n p_{i'_j}}{\prod_{j=1}^n \{1 - \sum_{r=1}^{j-1} p_{i'_r}\}} \\
&= \left(\prod_{j=1}^n p_{i_j}\right) \Big/ \left[\left(\prod_{j=1}^k \left\{1 - \sum_{r=1}^{j-1} p_{i_r}\right\}\right)\right. \\
&\qquad \left. \times \left(\prod_{j=k+1}^{\ell} \left\{1 - \sum_{\substack{r=1 \\ r\neq k}}^{j-1} p_{i_r} - p_{i_\ell}\right\}\right) \left(\prod_{j=\ell+1}^{n} \left\{1 - \sum_{r=1}^{j-1} p_{i_r}\right\}\right)\right] \\
&\leq \left(\prod_{j=1}^n p_{i_j}\right) \Big/ \left[\left(\prod_{j=1}^k \left\{1 - \sum_{r=1}^{j-1} p_{i_r}\right\}\right)\right. \\
&\qquad \left. \times \left(\prod_{j=k+1}^{\ell} \left\{1 - \sum_{\substack{r=1 \\ r\neq k}}^{j-1} p_{i_r} - p_{i_k}\right\}\right) \left(\prod_{j=\ell+1}^{n} \left\{1 - \sum_{r=1}^{j-1} p_{i_r}\right\}\right)\right] \\
&= P(s),
\end{aligned}
$$

where in going over from equality to inequality above we used the fact that $p_{i_\ell} \leq p_{i_k}$. This proves the required result.

For a successive sample of size n from a finite population of size N, let $I(S, k) = 1$, if $k \in S$ and 0, otherwise. Then $\Lambda(S)$ can be expressed as $\Lambda(S) = \sum_{k=1}^N p_k I(S, k)$. On taking expectations, we get

$$
\begin{aligned}
E[\Lambda(S)] &= \sum_{i=1}^N p_i \pi_i \\
&\equiv \phi(p).
\end{aligned}
\tag{2.1}
$$

Now we establish an inequality for $E[\Lambda(S)]$, the expected sample coverage. We shall be using the following lemma to prove it.

LEMMA 2.1. (Čebyšev's inequality. Theorem 1, p. 36 of Mitrinović (1970)) Let $a_1 \leq \cdots \leq a_n$ and $b_1 \leq \cdots \leq b_n$ be two increasing sequences of real numbers. Then

$$n \sum_{i=1}^{n} a_i b_i \geq \left(\sum_{i=1}^{n} a_i \right) \left(\sum_{i=1}^{n} b_i \right),$$

with equality holding if, and only if $a_1 = \cdots = a_n$ or $b_1 = \cdots = b_n$.

THEOREM 2.3. *Let S be a successive sample of size n from a finite population of size N. Then*

$$E[\Lambda(S)] \geq \frac{n}{N} \tag{2.2}$$

with equality holding if and only if S is a SRSWOR.

PROOF. Since for any sampling design with a fixed sample size n, $\sum_{i=1}^{N} I(S,i) = n$, on taking expectations we get $\sum_{i=1}^{N} \pi_i = n$.

Now observe from (2.1) that $E[\Lambda(S)] = \phi(p)$ is symmetric in p_k's and thus without loss of generality, we assume that they are ordered from the largest to the smallest. Then it follows from Corollary 2.1 that π_k's are also ordered from the largest to the smallest, the ordering being strict if the ordering of p_i's is strict. Applying Lemma 2.1, we get

$$E[\Lambda(S)] \geq \left\{ \sum_{i=1}^{N} p_i \right\} \left\{ \sum_{i=1}^{N} \pi_i \right\} \Big/ N = \frac{n}{N}. \tag{2.3}$$

Equality in (2.3) holds if, and only if π_k's are equal or p_k's are equal. By Corollary 2.1, π_k's are equal if and only if p_k's are equal, that is, if and only if S is a SRSWOR. □

The above result suggests that perhaps the expected sample coverage is greater when the p_k's are more dispersed. We need to introduce the concepts of majorization among vectors and Schur convexity to make this statement more precise.

DEFINITION 2.1. Let $\boldsymbol{x} = (x_1, \ldots, x_N)$ and $\boldsymbol{y} = (y_1, \ldots, y_N)$ be two vectors in the N-dimensional real space R^N. Let $\{x_{[1]} \geq \cdots \geq x_{[N]}\}$ denote the decreasing arrangement of the components of the vector $\boldsymbol{x}$. Then the vector $\boldsymbol{y}$ is said to majorize vector $\boldsymbol{x}$ (written as $\boldsymbol{x} \prec^m \boldsymbol{y}$), if

$$\sum_{k=1}^{j} x_{[k]} \leq \sum_{k=1}^{j} y_{[k]}, \qquad j = 1, \ldots, N-1 \quad \text{and} \quad \sum_{k=1}^{N} x_{[k]} = \sum_{k=1}^{N} y_{[k]}.$$

Observe that majorization as defined here is a partial order relation among vectors as opposed to component-wise partial order relation.

DEFINITION 2.2. A real valued function ϕ defined on a set $\mathcal{A} \subset \mathcal{R}^n$ is said to be Schur convex (Schur concave) on $\mathcal{A}$ if $\boldsymbol{x} \prec^m \boldsymbol{y} \Rightarrow \phi(\boldsymbol{x}) \leq (\geq) \phi(\boldsymbol{y})$.

We conjecture that the expected sample coverage $\phi(\boldsymbol{p})$ of a successive sample as given by (2.1) is Schur convex in $\boldsymbol{p}$ for any sample size, but have a proof only for $n = 2$.

THEOREM 2.4. *Consider a successive sample of size $n = 2$ with selection probabilities vector $\boldsymbol{p}$. Then the expected sample coverage $\phi(\boldsymbol{p}) = E[\Lambda(S)]$ is Schur convex in $\boldsymbol{p}$.*

PROOF. Note that we can write π_k as

$$\pi_k = t(1,k) + t(2,k),$$

where

$$t(j,k) = P(\text{unit } k \text{ is the } j\text{-th distinct unit drawn}), \qquad j = 1,2; \quad k = 1,\ldots,N.$$

Note that $t(1,k) = p_k$ and $t(2,k) = p_k I(k)$, where

$$I(k) = \sum_{j=1}^{N}\{p_j/(1-p_j)\} - p_k/(1-p_k), \qquad k = 1,\ldots,N.$$

Hence

$$\phi(\boldsymbol{p}) = \sum_{k=1}^{N} p_k^2 + \sum_{k=1}^{N} p_k^2 I(k)$$

which can be written, after a bit of rearranging, as

$$\phi(\boldsymbol{p}) = (N+1) + \left[\sum_{k=1}^{N} p_k^2 - 1\right] J - (N-2)\sum_{k=1}^{N} p_k^2,$$

where $J = \sum_{k=1}^{N} 1/(1-p_k)$. From this we get

$$\begin{aligned}
\frac{\partial\phi}{\partial p_1} - \frac{\partial\phi}{\partial p_2} &= 2\{p_1 - p_2\}\left\{\sum_{k=3}^{N}\frac{1}{1-p_k} - (N-2)\right\} + \left\{\sum_{k=1}^{N} p_k^2 - 1\right\} \\
&\quad \cdot \left\{\frac{1}{(1-p_1)^2} - \frac{1}{(1-p_2)^2}\right\} + 2\{p_1 - p_2\}\left\{\frac{1}{1-p_1} + \frac{1}{1-p_2}\right\} \\
&= 2\{p_1 - p_2\}\left\{\sum_{k=3}^{N}\frac{1}{1-p_k} - (N-2)\right\} + \frac{(p_1-p_2)(2-p_1-p_2)}{(1-p_1)^2(1-p_2)^2} \\
&\quad \times \left[\sum_{k=1}^{N} p_k^2 - 1 + 2(1-p_1)(1-p_2)\right] \\
&= \{p_1 - p_2\}\left[2\left\{\sum_{k=3}^{N}\frac{1}{1-p_k} - (N-2)\right\}\right. \\
&\quad \left. + \frac{(2-p_1-p_2)}{(1-p_1)^2(1-p_2)^2}\left\{\sum_{k=3}^{N} p_k^2 + (p_1+p_2-1)^2\right\}\right]
\end{aligned}$$

which has the same sign as $(p_1 - p_2)$ since $1/(1-p_i) \geq 1$ for each $i = 3,\ldots,N$. Thus $\phi(\boldsymbol{p})$ is Schur convex in $p_1,\ldots,p_N$ by Theorem A.4, page 57 of Marshall and Olkin (1979). □

3. Some majorization results

Rao *et al.* (1991) proved that in case of successive sampling $np_{(1)} \leq \pi_{(1)}$ and $np_{(N)} \geq \pi_{(N)}$, where $p_{(1)}$ $(\pi_{(1)})$ and $p_{(N)}(\pi_{(N)})$ are the minimum and the maximum of p_i's (π_i's). Cochran ((1977), p. 259) considers the case $n = 2$ and asserts that $\pi_i/2$'s are "always closer to equality than the original p_i's ". However, he does not make it clear in what sense they are closer. We generalize these results in Theorem 3.1 to prove that the vector $\boldsymbol{\pi}$ is majorized by the vector $n\boldsymbol{p}$. We also compare the three without replacement sampling schemes by partially ordering the vectors of first order inclusion probabilities by majorization. First we prove some preliminary results in the next two lemmas.

LEMMA 3.1. *Let $p_1 \geq \cdots \geq p_N$ be a set of probabilities and let*

$$q_k = \frac{\sum_{j=1}^{k} \frac{p_j}{1-p_j}}{\sum_{j=1}^{k} p_j}, \qquad k = 1, \ldots, N. \tag{3.1}$$

Then q_k is a decreasing function of k.

PROOF. We have, after a bit of simplification,

$$\left(\sum_{j=1}^{k} p_j\right)\left(\sum_{j=1}^{k+1} p_j\right)[q_{k+1} - q_k] = \left(\frac{p_{k+1}}{1-p_{k+1}}\right)\left(\sum_{j=1}^{k} p_j\right) - \left(\sum_{j=1}^{k} \frac{p_j}{1-p_j}\right) p_{k+1}. \tag{3.2}$$

Now, $\sum_{j=1}^{k}\{p_j/(1-p_j)\} \geq \{1/(1-p_{k+1})\}\sum_{j=1}^{k} p_j$, since $(1-p_j) \leq (1-p_{k+1})$, $j = 1, \ldots, k$, as p_i's are ordered. Thus, the right hand side of (3.2) is less than or equal to 0. This completes the proof. □

For successive sampling from the finite population $\mathcal{U}$ with selection probabilities p_i's and first order inclusion probabilities π_i's, let

$$t(j,\ell) = P(\text{unit } \ell \text{ is the } j\text{-th distinct unit drawn}), \qquad j = 1, \ldots, N; \quad \ell = 1, \ldots, N.$$

Note that we have defined $t(j,\ell)$ for $j = 1, \ldots, N$. Although in practice we draw only n distinct units, this extended definition will be useful for theoretical purposes later on. Observe that $t(1,\ell) = p_\ell$, $\ell = 1, \ldots, N$. Conditioning on the first (distinct) unit drawn we have the following representation for $t(j,\ell)$ for other values of j,

$$t(j,\ell) = \sum_{\substack{i=1 \\ i \neq \ell}}^{N} p_i t(j-1, \ell \mid i), \qquad j = 2, \ldots, N, \tag{3.3}$$

where for each i $(i = 1, \ldots, N)$, $t(j-1, \ell \mid i)$ is the conditional probability of drawing unit ℓ as the j-th distinct unit drawn given that unit i was the first unit drawn. It turns out that $t(j-1,\ell \mid i)$ is the probability of drawing unit ℓ as the $(j-1)$-th distinct unit in the sample from the truncated population $\mathcal{U}(i) = \{1, \ldots, N\} - \{i\}$ and with selection probabilities $p(\ell \mid i) = p_\ell/(1-p_i)$, $\ell = 1, \ldots, N$, $\ell \neq i$. The corresponding first order inclusion probabilities $\pi(\ell \mid i)$'s satisfy the relation,

$$\pi_\ell = p_\ell + \sum_{\substack{i=1 \\ i \neq \ell}}^{N} p_i \pi(\ell \mid i), \qquad \ell = 1, \ldots, N.$$

This follows from the definitions of $\pi_\ell, \pi(\ell \mid i)$ and the identity $\pi_\ell = \sum_{j=1}^{n} t(j,\ell)$.

Lemma 3.2 below, which is also of independent interest, leads to the main result of this section.

LEMMA 3.2. *Suppose that in a successive sample from a finite population of size N, the selection probabilities p_i's are ordered from the largest to the smallest. Then for $j = 1,\dots,N-1$,*

$$\sum_{\ell=1}^{k} t(j+1,\ell) \le \sum_{\ell=1}^{k} t(j,\ell), \qquad k = 1,\dots,N. \tag{3.4}$$

In particular,

$$\sum_{\ell=1}^{k} t(j,\ell) \le \sum_{\ell=1}^{k} t(1,\ell) = \sum_{\ell=1}^{k} p_\ell, \qquad k = 1,\dots,N.$$

PROOF. We will use induction on N to prove the result. From the definition of $t(j,\ell)$, we have for $k = 1,\dots,N$

$$\begin{aligned} \sum_{\ell=1}^{k} t(2,\ell) &= \left(\sum_{\ell=1}^{k} p_\ell\right)\left(\sum_{i=1}^{N} \frac{p_i}{1-p_i}\right) - \sum_{\ell=1}^{k} \frac{p_\ell^2}{1-p_\ell} \qquad (3.5)\\ &= \sum_{\ell=1}^{k} p_\ell + \left(\sum_{\ell=1}^{k} p_\ell\right)\left[\sum_{i=1}^{N} \frac{p_i}{1-p_i} - \left(\sum_{\ell=1}^{k} \frac{p_\ell}{1-p_\ell}\right)\Big/\left(\sum_{\ell=1}^{k} p_\ell\right)\right] \\ &\quad \text{(writing } p_\ell/(1-p_\ell) \text{ as } -1+1/(1-p_\ell)\text{)} \\ &\le \sum_{\ell=1}^{k} t(1,\ell), \qquad k = 1,\dots,N. \qquad (3.6) \end{aligned}$$

The last inequality follows since by Lemma 3.1,

$$\left(\sum_{\ell=1}^{k} \frac{p_\ell}{1-p_\ell}\right)\Big/\left(\sum_{\ell=1}^{k} p_\ell\right) \ge \left(\sum_{i=1}^{N} \frac{p_i}{1-p_i}\right)\Big/\sum_{i=1}^{N} p_i = \sum_{i=1}^{N} \frac{p_i}{1-p_i}.$$

In particular, for $N = 2$,

$$\sum_{\ell=1}^{k} t(2,\ell) \le \sum_{\ell=1}^{k} t(1,\ell), \qquad k = 1,2. \tag{3.7}$$

Inequality (3.7) represents the initial step in the proof by induction on N. Suppose now that the result is true for $N-1$. From (3.3) it follows at once that

$$\begin{aligned} \sum_{\ell=1}^{k} t(j,\ell) &= \sum_{\ell=1}^{k} \sum_{\substack{i=1 \\ i\ne\ell}}^{N} p_i t(j-1,\ell \mid i) \\ &= \sum_{i=1}^{N} p_i \sum_{\substack{\ell=1 \\ \ell\ne i}}^{k} t(j-1,\ell \mid i) \end{aligned}$$

$$\le \sum_{i=1}^{N} p_i \sum_{\substack{\ell=1 \\ \ell \ne i}}^{k} t(j-2, \ell \mid i)$$

$$= \sum_{\ell=1}^{k} t(j-1, \ell), \qquad j = 3, \ldots, N.$$

This and (3.7) complete the proof. □

REMARK 3.1. Parts (i) and (ii) of Theorem 2.2 of Rao *et al.* (1991) follow as immediate consequences of Lemma 3.2.

Now we prove the main result of this section.

THEOREM 3.1. *Consider a successive sample of size n with selection probability vector $\boldsymbol{p}$ and first order inclusion probability vector $\boldsymbol{\pi}$. Then $\boldsymbol{\pi} \prec^m n\boldsymbol{p}$.*

PROOF. Let the p_i's be ordered from the largest to the smallest. Then by Corollary 2.1 the π_i's are also ordered the same way. We can write π_ℓ as

$$\pi_\ell = \sum_{j=1}^{n} t(j, \ell), \tag{3.8}$$

where as in Lemmas 3.1 and 3.2, $t(j, \ell)$ denotes the probability of including unit ℓ in the sample as the j-th distinct unit drawn. To complete the proof, we sum the above identity for π_ℓ from 1 to k, and use Lemma 3.2. Thus for $k = 1, \ldots, N$,

$$\begin{aligned} \sum_{\ell=1}^{k} \pi_\ell &= \sum_{\ell=1}^{k} \sum_{j=1}^{n} t(j, \ell) \\ &= \sum_{j=1}^{n} \sum_{\ell=1}^{k} t(j, \ell) \\ &\le \sum_{j=1}^{n} \sum_{\ell=1}^{k} p_\ell \qquad (3.9) \\ &= n \sum_{\ell=1}^{k} p_\ell. \end{aligned}$$

(3.9) follows from (3.4) since $t(1, \ell) = p_\ell$ for all ℓ. □

Our next result characterizes simple random sampling without replacement by the result in Theorem 3.1.

PROPOSITION 3.1. *Let $\boldsymbol{S}$ be a successive sample of size n from a finite population of size N with selection probabilities $p_1, \ldots, p_N$ and first order inclusion probabilities $\pi_1, \ldots, \pi_N$. Then $np_i = \pi_i$, $i = 1, \ldots, N$ if and only if, $p_i = 1/N$, $i = 1, \ldots, N$, that is, if and only if, $\boldsymbol{S}$ is a SRSWOR.*

PROOF. "Only if" part. Let the p_i's be ordered from the largest to the smallest. Let k be a nonnegative integer between 1 and N, both inclusive. Assume $np_i = \pi_i$ for $i = 1, \ldots, N$. We then have by Lemma 3.2

$$n\sum_{\ell=1}^{k} p_\ell = \sum_{\ell=1}^{k} \pi_\ell \leq \sum_{\ell=1}^{k} p_\ell + (n-1)\sum_{\ell=1}^{k} t(2,\ell) \leq n\sum_{\ell=1}^{k} p_\ell.$$

Thus

$$\sum_{\ell=1}^{k} t(2,\ell) = \sum_{\ell=1}^{k} p_\ell, \qquad k = 1, \ldots, N. \tag{3.10}$$

This implies

$$t(2, k+1) = p_{k+1}, \qquad \text{for} \quad k = 0, \ldots, N-1.$$

That is,

$$p_{k+1}\{I^* - p_{k+1}/(1-p_{k+1})\} = p_{k+1}, \qquad k = 0, \ldots, N-1 \tag{3.11}$$

where $I^* = \sum_{\ell=1}^{N}\{p_\ell/(1-p_\ell)\}$. From (3.11) we get

$$p_{k+1}/\{1-p_{k+1}\} = I^* - 1, \qquad k = 0, \ldots, N-1$$

which yields the desired result, concluding the proof of the "only if" part.

Proof of "if part". For a SRSWOR of size n, we have $p_\ell = 1/N$, $\pi_\ell = n/N$, $\ell = 1, \ldots, N$. Thus, $np_\ell = \pi_\ell$, $\ell = 1, \ldots, N$, completing the proof of "if" part. □

Our last result of this section compares the vectors of first order inclusion probabilities and the expected sample coverages for the three sampling designs SRSWOR, Midzuno sampling and successive sampling.

Note that since SRSWOR is a special case of successive sampling, the interpretation of $\phi(\boldsymbol{p})$ (defined by the last equality in (2.1)) as the expected sample coverage for SRSWOR makes perfect sense. In the case of Midzuno sampling, we define $\phi(\boldsymbol{p})$ by the same equality, the p_i's being the selection probabilities of the units for the first draw.

THEOREM 3.2. (a) *Let π^{SRSWOR}, π^{M} and π^{SS} be the vectors of first order inclusion probabilities for SRSWOR, Midzuno sampling and successive sampling respectively. Then*

$$\pi^{SRSWOR} \overset{m}{\prec} \pi^{M} \overset{m}{\prec} \pi^{SS}.$$

(b) *Let $\phi^{SRSWOR}(\boldsymbol{p})$, $\phi^{M}(\boldsymbol{p})$ and $\phi^{SS}(\boldsymbol{p})$ be the expected sample coverages for a SRSWOR, a Midzuno sample and a successive sample, respectively. Then,*

$$\phi^{SRSWOR}(\boldsymbol{p}) \leq \phi^{M}(\boldsymbol{p}) < \phi^{SS}(\boldsymbol{p}).$$

PROOF. (a) The result that π^{M} majorizes π^{SRSWOR} follows from the fact that for any real numbers a_i's, $(a_1, \ldots, a_N)$ majorizes $(\bar{a}, \ldots, \bar{a})$, where $\bar{a} = \sum_{i=1}^{N} a_i/N$.

Now we prove the other part of the assertion of the theorem. From (3.8) and the definition of $t(j,\ell)$, we have for successive sampling

$$\sum_{\ell=1}^{k} \pi_\ell^{SS} = \sum_{\ell=1}^{k} p_\ell + \sum_{\ell=1}^{k}\sum_{j=2}^{n} t(j,\ell) \tag{3.12}$$

and the inclusion probabilities for the Midzuno sampling are given by

$$\pi_\ell^M = p_\ell + \{1 - p_\ell\}\{(n-1)/(N-1)\}, \qquad \ell = 1, \ldots, N. \tag{3.13}$$

From (3.12) and (3.13) it follows that

$$\sum_{\ell=1}^{k}\{\pi_\ell^{SS} - \pi_\ell^M\} = \sum_{\ell=1}^{k}\sum_{j=2}^{n} t(j,\ell) - \left(k - \sum_{\ell=1}^{k} p_\ell\right)(n-1)/(N-1). \tag{3.14}$$

Now, the first term on the right hand side of (3.14) can be written as

$$\sum_{\ell=1}^{k}\sum_{j=2}^{n} t(j,\ell) = \sum_{i=1}^{N} p_i \sum_{j=1}^{n-1}\sum_{\substack{\ell=1\\ \ell\neq i}}^{k} t(j,\ell \mid i). \tag{3.15}$$

We now note that $\sum_{j=1}^{n-1}\sum_{\ell=1,\ell\neq i}^{k} t(j,\ell \mid i)$ is the sum of the first k largest π_ℓ's of a successive sample of size $n-1$ from the finite population consisting of all the original N units except i, and selection probabilities $p_k/\{1-p_i\}$, $k = 1,\ldots,N$, $k \neq i$. Since $((n-1)/(N-1),\ldots,(n-1)/(N-1))$ is majorized by any $(N-1)$-vector of nonnegative numbers adding to $n-1$, it follows that

$$\begin{aligned}\sum_{i=1}^{N} p_i \sum_{j=1}^{n-1}\sum_{\substack{\ell=1\\ \ell\neq i}}^{k} t(j,\ell \mid i) &\geq \sum_{i=1}^{N} p_i \sum_{\substack{\ell=1\\ \ell\neq i}}^{k} \frac{n-1}{N-1}\\ &= \sum_{\ell=1}^{k} \frac{n-1}{N-1}\sum_{\substack{i=1\\ i\neq \ell}}^{N} p_i\\ &= \left(k - \sum_{\ell=1}^{k} p_\ell\right)\frac{n-1}{N-1}.\end{aligned} \tag{3.16}$$

Thus, (3.14), (3.15) and (3.16) complete the proof of the second assertion of (a).

(b) To prove the first inequality, we have from (3.13),

$$\begin{aligned}\phi^M(\mathbf{p}) &= \frac{(N-n)}{(N-1)}\sum_{k=1}^{N} p_k^2 + \frac{n-1}{N-1}\\ &= \frac{N-n}{N-1}\sum_{k=1}^{N}\left(p_k - \frac{1}{N}\right)^2 + \frac{n}{N}\\ &\geq n/N\\ &= \phi^{SRSWOR}(\mathbf{p}).\end{aligned} \tag{3.17}$$

To prove the second inequality, we have from (3.8) and the definition of $t(j,k)$ that

$$\pi_k^{SS} = p_k + \sum_{\substack{i=1\\ i\neq k}}^{N} p_i \left\{\sum_{j=1}^{n-1} t(j,k \mid i)\right\} \tag{3.18}$$

$$= p_k + \sum_{\substack{i=1 \\ i \neq k}}^{N} p_i \pi(k \mid i).$$

From this we get

$$\phi^{SS}(\boldsymbol{p}) = \sum_{k=1}^{N} p_k^2 + \sum_{i=1}^{N} p_i(1 - p_i)\phi^{SS}(\boldsymbol{p} \mid i), \tag{3.19}$$

where $\phi^{SS}(\boldsymbol{p} \mid i)$ is the function $\phi(\boldsymbol{p})$ for the successive sample from $\mathcal{U}(i)$. Now the result follows from applying Theorem 2.3 to $\phi^{SS}(\boldsymbol{p} \mid i)$ and (3.17). □

Remark 3.2. Note that the lower bound $\phi^M(\boldsymbol{p})$ on $\phi^{SS}(\boldsymbol{p})$ is sharper than $\phi^{SRSWOR}(\boldsymbol{p})$ as provided by Theorem 3.2. However, unlike $\phi^M(\boldsymbol{p})$ which depends on p_i's, $\phi^{SRSWOR}(\boldsymbol{p})$ is a universal bound (the same for all successive samples of the same size).

4. Spacings of a simple random sample without replacement

In this section we study some properties of spacings corresponding to a SRSWOR from a linearly ordered finite population without multiplicities.

THEOREM 4.1. *Let $X_1, \ldots, X_n$ be a SRSWOR from a linearly ordered finite population without multiplicities and let $X_{(1)}, \ldots, X_{(n)}$ be the corresponding order statistics. Let $D_i = X_{(i)} - X_{(i-1)}$, $i = 1, \ldots, n$ be the spacings with $X_{(0)} \equiv 0$. Then D_i's are exchangeable.*

PROOF. We have

$$\begin{aligned} &P(D_i = s_i, i = 1, \ldots, n) \\ &\quad = P(X_{(i)} = \sum_{j=1}^{i} s_j, i = 1, \ldots, n) \\ &\quad = 1 \Big/ \binom{N}{n}, \qquad n \le s_1 + \cdots + s_n \le N, \quad \text{each } s_i \ge 1. \end{aligned}$$ □

COROLLARY 4.1.

$$P(D_i = s_i, i = 1, \ldots, k) = \binom{N - \sum_{i=1}^{k} s_i}{n-k} \Big/ \binom{N}{n} \tag{4.1}$$

and

$$P(D_i > t_i, i = 1, \ldots, k) = \binom{N - \sum_{i=1}^{k} t_i}{n} \Big/ \binom{N}{n}. \tag{4.2}$$

PROOF. We have

$$\begin{aligned} P(D_i = s_i, i = 1, \ldots, k) &= P(X_{(i)} = \sum_{j=1}^{i} s_j, i = 1, \ldots, k) \\ &= \binom{N - \sum_{i=1}^{k} s_i}{n-k} \Big/ \binom{N}{n}. \end{aligned}$$

The proof of (4.2) follows by repeated application on (4.1) of the well-known identity,

$$\sum_{k=t+1}^{m-1} \binom{m-k}{\ell} = \binom{m-t}{\ell+1}. \tag{4.3}$$ □

THEOREM 4.2. *The p.m.f. of D_i is logconcave for $i = 1, \ldots, n$.*

PROOF. From Corollary 4.1 it follows that

$$f(s) = P(D_i = s) = \binom{N-s}{n-1} \Big/ \binom{N}{n}.$$

Hence

$$\begin{aligned} f(s)/f(s-1) &= \binom{N-s}{n-1} \Big/ \binom{N-s+1}{n-1} \\ &= 1 - \frac{n-1}{N-s+1}. \end{aligned}$$

Now $1-(n-1)/(N-s+1)$ is a decreasing function of s. This completes the proof. □

Corollary 4.1 implies that each D_i has an IFR (*increasing failure rate*) distribution. We next prove a multivariate version of this result. First we give the definition of a *multivariate increasing failure rate* (MIFR) distribution.

DEFINITION 4.1. A random vector $(X_1, \ldots, X_n)$ with survival function $S(x_1, \ldots, x_n)$ is said to have a *multivariate increasing failure rate* (MIFR) distribution if the marginal survival function $S_{i_1,\ldots,i_k}(x_{i_1}, \ldots, x_{i_k})$ of $\{X_{i_1}, \ldots, X_{i_k}\}$ satisfies the condition that

$$\frac{S_{i_1,\ldots,i_k}(x_{i_1}+t, \ldots, x_{i_k}+t)}{S_{i_1,\ldots,i_k}(x_{i_1}, \ldots, x_{i_k})}$$

is decreasing in x_{i_1} for $t > 0$ for each subset $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, n\}$.

THEOREM 4.3. *The vector of spacings $\boldsymbol{D} = (D_1, \ldots, D_n)$ has a Multivariate Increasing Failure Rate (MIFR) distribution.*

PROOF. From Corollary 4.1 we have, for $1 \le k \le n$ and $x > 0$, that

$$\begin{aligned} &P(D_i > t_i + x, i = 1, \ldots, k \mid D_i > t_i, i = 1, \ldots, k) \\ &= \binom{N - \sum_{i=1}^k t_i - kx}{n} \Big/ \binom{N - \sum_{i=1}^k t_i}{n} \\ &= \prod_{j=0}^{n-1} \left\{ 1 - kx \Big/ \left(N - \sum_{i=1}^k t_i - j \right) \right\} \end{aligned}$$

which is decreasing in t_i's, since for each j, $1 - kx/(N - \sum_{i=1}^k t_i - j)$ is decreasing in t_i's. This completes the proof. □

Our last result in this section shows that the spacings are negatively dependent in the sense that the joint distribution of any pair of them is *reverse regular of order 2* (RR_2) (see Karlin and Rinott (1980)).

THEOREM 4.4. *For $1 \leq i < j \leq n$, the joint p.m.f. $f(s,t)$ of D_i and D_j is reverse regular of order two (RR_2). That is,*

$$d = \begin{vmatrix} f(s_1,t_1) & f(s_1,t_2) \\ f(s_2,t_1) & f(s_2,t_2) \end{vmatrix} \leq 0$$

for all $s_1 < s_2$ and $t_1 < t_2$.

PROOF. From Corollary 4.1 it follows that

$$\binom{N}{n}^2 d = \binom{N-s_1-t_1}{n-2}\binom{N-s_2-t_2}{n-2} - \binom{N-s_1-t_2}{n-2}\binom{N-s_2-t_1}{n-2}$$

which is < 0 for $s_1 < s_2$ and $t_1 < t_2$, since $\binom{m_1}{k}\binom{m_2}{k}$ is Schur concave in m_1 and m_2. This completes the proof. □

5. Characterizations of SRSWOR by exchangeability of spacings

In this section we use exchangeability of the spacings to characterize SRSWOR.

Let $\boldsymbol{S}$ be a Midzuno sample from a linearly ordered finite population without multiplicities, and with initial selection probability vector $\boldsymbol{p}$. Then the spacings D_i's have the joint p.m.f. given by

$$(5.1) \qquad \begin{aligned} P(D_i = s_i, i = 1,\ldots,n) &= P\left(X_{(i)} = \sum_{j=1}^{i} s_j, i = 1,\ldots,n\right) \\ &= P(\boldsymbol{S}) \\ &= \prod_{i=1}^{n} p_{(\sum_{j=1}^{i} s_j)} \Big/ \binom{N-1}{n-1}, \end{aligned}$$

where

$$(5.2) \qquad \boldsymbol{S} = \left\{\sum_{j=1}^{i} s_j, i = 1,\ldots,n\right\}.$$

THEOREM 5.1. *Let $\boldsymbol{S}$ be a Midzuno sample of size n from a linearly ordered finite population without multiplicities, and with initial selection probabilities $p_1,\ldots,p_N$. Let $p_N = 1/N$. Then the spacings are exchangeable if, and only if S is a SRSWOR.*

PROOF. "If" part. Suppose $\boldsymbol{S}$ is a SRSWOR. Then it follows from Theorem 4.1 that the D_i's are exchangeable.

"Only if" part. Let D_i's be exchangeable. Then interchanging s_i and $s_{i-1}(i = 2, \ldots, n)$ in (5.1), we have $P(S) = P(S^*)$, where

$$S^* = (s_1^*, \ldots, s_n^*), \tag{5.3}$$

with $s_j^* = \sum_{k=1}^{j} s_k$, $j = 1, \ldots, n$, $j \neq i-1$, $s_{i-1}^* = \sum_{k=1}^{i-2} s_k + s_i$ and S is given by (5.2). Now, S and S^* differ in only one element. Hence we have, by Theorem 2.1, that

$$p_{\sum_{j=1}^{i-1} s_j} = p_{\sum_{j=1}^{i-2} s_j + s_i}. \tag{5.4}$$

Putting, for $1 \leq k \leq N-n+1$

$$s_j = 1, \quad j = 1, \ldots, n, \quad j \neq i, \;\; s_i = k,$$

we have

$$p_{i-1} = p_{i-2+k}, \quad 1 \leq k \leq N-n+1, \quad 2 \leq i \leq n.$$

Thus, finally, we have

$$p_1 = \cdots = p_{N-1} = \{1 - p_N\}/(N-1) = 1/N,$$

the last equality following from the assumption that $p_N = 1/N$. This completes the proof of the "if only" part, and in turn that of the theorem. □

Now let S be a successive sample of size n from a linearly ordered population without multiplicities, and selection probabilities $p_1, \ldots, p_N$. Then, as in the case of the Midzuno sample

$$P(D_i = s_i, i = 1, \ldots, n) = P(S) \tag{5.5}$$

where S is given by (5.2) and $P(S)$ by (1.2) (with appropriate notational changes).

THEOREM 5.2. *Let S be a successive sample of size n from a linearly ordered finite population without multiplicities, and selection probabilities $p_1, \ldots, p_N$. Let $p_N = 1/N$. Then the spacings D_i's are exchangeable if and only if S is a SRSWOR.*

PROOF. "If" part. This is Theorem 4.1.

"Only if" part. Let D_i's be exchangeable. Then, interchanging s_i and s_{i-1} in (5.5), we get $P(S) = P(S^*)$, where S and S^* are given by (5.2) and (5.3) respectively. Now, S and S^* differ only in one element and hence from Theorem 2.1, (5.4) must hold. The proof from this point on is exactly the same as that of Theorem 5.2. This completes the proof of the theorem. □

Acknowledgements

We are thankful to the two referees for their careful reading of the original draft of the paper and for making some useful comments and suggestions.

References

Andreatta, G. and Kaufman, G. M. (1986). Estimation of finite population properties when sampling is without replacement and proportional to magnitude, *J. Amer. Statist. Assoc.*, **81**, 657–666.

Cochran, W. G. (1977). *Sampling Techniques*, 3rd ed., Wiley, New York.

Hájek, J. (1981). *Sampling from a Finite Population*, Marcel Dekker, New York.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities II. Multivariate reverse rule distributions, *J. Multivariate Anal.*, **10**, 499–516.

Kochar, S. C. and Korwar, R. M. (1996). Stochastic orders for spacings of heterogeneous exponential random variables, *J. Multivariate Anal.*, **57**, 69–83.

Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York.

Midzuno, H. (1950). An outline of the theory of sampling systems, *Ann. Inst. Statist. Math.*, **1**, 149–156.

Mitrinović, D. S. (1970). *Analytic Inequalities*, Springer, New York.

Rao, T. J. , Sengupta, S. and Sinha, B. K. (1991). Some order relations between selection and inclusion probabilities for PPSWOR sampling scheme, *Metrika*, **38**, 335–343.