# Antithetic Sampling

Mohammad Jafari Jozani

Université de Sherbrooke, Sherbrooke, Qc, CANADA

François Perron

Université de Montréal, Montréal, Qc, CANADA

**Summary.** Consider the problem of estimating the mean of a population in a nonparametric framework. Assume that we can rank all of the observations but we are aloud to measure only $n$ of them. In this paper we introduce new sampling algorithms using antithetic variables. We propose unbiased estimators which reduce the variance in comparison with the ranked set sampling estimator. Theoretical results and numerical comparisons are included.

*Keywords:* Antithetic sampling; Ranked set sampling.

## 1 Introduction

In most sampling surveys, a reasonable number of the sampling units can be fairly accurately ordered with respect to a variable of interest without actual measurement and at little cost. On the other hand, exact measurements of these units may be very tedious and/or expensive. For example, for environmental risks such as radiation (soil contamination, disease clusters) or pollution (water contamination, root disease of crops) we commonly find that exact measurements involves substantial scientific processing of materials and correspondingly high attendant cost, while the variable of interest from the experimental (sampling) units can easily be ranked. Ranked set sampling, as proposed by McIntyre (1952) in estimating the mean of the pasture yields, provides an interesting alternative to simple random sampling in these situations.

Compared to simple random sampling, ranked set sampling has been proven theoretically (e.g., Takahasi and Wakimoto, 1968) and shown empirically (see, Kaur et al., 1995 and Chen et al., 2004

Address for correspondence: Département de Mathématiques et de Statistique, Université de Montréal, P.O. Box 6128, Succursale Centre-Ville, Montréal, Que., Canada H3C 3J7.
E-mail address: perronf@dms.umontreal.ca (F. Perron).

for more details) to yield more precise estimator of the population mean. This is especially useful for small surveys or for situations where it is expensive or destructive to obtain data.

Since the pioneering articles of McIntyre (1952) and Takahasi and Wakimoto (1968) several variations of ranked set sampling method has been proposed and developed by researchers to come up with more efficient estimators of a population mean. For example, Samawi et al. (1996) introduced extreme ranked set sampling and obtained an unbiased estimator of the mean which outperforms the usual mean of a simple random sample of the same size for symmetric distributions. Muttlak (1997) suggested median ranked set sampling to increase the efficiency and to reduce ranking errors over ranked set sampling method and proved its better performance in estimating the mean of a variable of interest for some symmetric distributions. Some relevant references here, in addition to those mentioned above, are among others: Bhoj (1997) for a new parametric ranked set sampling, Li et al. (1999) for random selection in ranked set sampling, Hossain and Muttlak (1999) for paired ranked set sampling, Al-Saleh and Al-Kadiri (2000) for double ranked set sampling, Hossain and Muttlak (2001) for selected ranked set sampling and Al-Saleh and Al-Omari (2002) for multistage ranked set sampling. Al-Nasser (2007) introduced L-Ranked set sampling design as a generalization of some of the above mentioned ranked set type sampling methods and proved the optimal property of his proposed estimators for symmetric family of distributions.

A review of these articles reveals that most of them concentrate mainly on situations where specific assumptions are made concerning the parent distribution of the underlying population and the proposed estimators are beneficiated from those assumptions to provide their optimal properties such as unbiasedness or having smaller variance.

The aim of this paper is to propose better estimators than $\hat{\mu}_R$, the ranked set sampling estimator, under the assumption that we can rank the observations before measurements, in a completely non parametric basis and without making any assumption about the distribution of the underlying population. The organization of the remaining sections is as follows. In Section 2, we introduce a new estimator, $\hat{\mu}_A$, of the population mean called the antithetic estimator, based on a new proposed ranked set type sampling method. We show that our estimator is unbiased and we derive an explicit expression for its variance. It is then proved that $\hat{\mu}_A$ dominates $\hat{\mu}_S$, the simple random sampling estimator, by a direct comparison of the variances. However, it is not clear how to handle the comparisons of the variances between the $\hat{\mu}_R$ and $\hat{\mu}_A$ based simply on a direct calculation. In

Section 3, we introduce a class of random estimators. The main idea is conditioning on the order statistics. It is interesting to see that $\hat{\mu}_S$ and $\hat{\mu}_R$ are now viewed as random estimators. A general dominance result leading to a sufficient condition (see Theorem 3.1), which we believe is novel, for a random estimator $\hat{\mu}_1$ to dominate another random estimator $\hat{\mu}_2$ is established. This method works for showing that $\hat{\mu}_A$ dominates $\hat{\mu}_S$. Similarly, we show that $\hat{\mu}_A$ dominates $\hat{\mu}_R$ when the sample size is equal to two. Finally, in Section 4, we propose a better estimator than $\hat{\mu}_R$ for all sample sizes. A simulation study is carried out in Section 5.

## 2 The antithetic estimator

Consider $X_1, \ldots, X_{n^2}$ as a simple random sample of size $n^2$ from a distribution with mean $\mu$ and finite variance $\sigma^2$ with $X_{(1)}, \ldots, X_{(n^2)}$ as its corresponding order statistics. Let $X_{ij} = X_{n(i-1)+j}$ and define $X_{(ij)} = X_{(n(i-1)+j)}$, $i, j = 1, \ldots, n$, be the $n(i-1)+j$ th order statistics in the sample of size $n^2$. Also, we define $X_{i(j)}$ to be the $j$ th order statistics based on $X_{i1}, \ldots, X_{in}$, $i, j = 1, \ldots, n$. We adopt the following notations in the rest of the paper:

$$X_{(i\cdot)} = \frac{1}{n}\sum_j X_{(ij)}, \quad X_{(\cdot j)} = \frac{1}{n}\sum_i X_{(ij)}, \quad X_{(\cdot\cdot)} = \frac{1}{n^2}\sum_{ij} X_{(ij)}, \quad i, j = 1, \ldots, n.$$

$$\mu = \mathrm{E}[X_1], \quad \sigma^2 = \mathrm{Var}[X_1], \quad \mu_{1(i)} = \mathrm{E}[X_{1(i)}], \quad \sigma^2_{1(i)} = \mathrm{Var}[X_{1(i)}], \quad i = 1, \ldots, n.$$

$$\mu_{(ij)} = \mathrm{E}[X_{(ij)}], \quad \mu_{(i\cdot)} = \mathrm{E}[X_{(i\cdot)}], \quad \mu_{(\cdot j)} = \mathrm{E}[X_{(\cdot j)}], \quad \mu_{(\cdot\cdot)} = \mathrm{E}[X_{(\cdot\cdot)}], \quad i, j = 1, \ldots, n.$$

$$\sigma^2_{(ij)} = \mathrm{Var}[X_{(ij)}], \quad \sigma^2_{(i\cdot)} = \mathrm{Var}[X_{(i\cdot)}], \quad \sigma^2_{(\cdot j)} = \mathrm{Var}[X_{(\cdot j)}], \quad \sigma^2_{(\cdot\cdot)} = \mathrm{Var}[X_{(\cdot\cdot)}], \quad i, j = 1, \ldots, n.$$

For estimating the population mean $\mu$, let $\hat{\mu}_S$ be the mean of a simple random sample of size $n$, i.e., $\{X_{1(1)}, \ldots, X_{1(n)}\}$, which is given by

$$\hat{\mu}_S = \frac{1}{n}\sum_{i=1}^{n} X_{1(i)}. \tag{1}$$

with $\mathrm{E}[\hat{\mu}_S] = \mu$ and $\mathrm{Var}[\hat{\mu}_S] = \frac{\sigma^2}{n}$. Also, let $\hat{\mu}_R$ be the mean of a ranked set sample of size $n$, i.e., $\{X_{1(1)}, \ldots, X_{n(n)}\}$, from the underlying population as

$$\hat{\mu}_R = \frac{1}{n}\sum_{i=1}^{n} X_{i(i)}. \tag{2}$$

It is shown that (e.g., Takahasi and Wakimoto, 1968; Dell, 1969; Dell and Clutter, 1972), $\mathrm{E}[\hat{\mu}_R] = \mu$ with

$$\mathrm{Var}[\hat{\mu}_R] = \frac{\sigma^2}{n} - \frac{1}{n^2}\sum_{i=1}^{n}(\mu_{1(i)} - \mu)^2. \tag{3}$$

Let $\pi = \{\pi_1, \ldots, \pi_n\}$ be a random permutation of $\{1, \ldots, n\}$ independent of the sample and select $\{X_{(1\pi_1)}, \ldots, X_{(n\pi_n)}\}$ as a new ranked set type sample called the antithetic sample. For estimating the population mean, based on the antithetic sample, we propose the following estimator

$$\hat{\mu}_A = \frac{1}{n}\sum_{i=1}^{n} X_{(i\pi_i)}. \tag{4}$$

Some basic properties of $\hat{\mu}_A$ are given in Theorem 2.1.

**Theorem 2.1** *Let $\hat{\mu}_A$ be the antithetic estimator defined in* (4) *and $\hat{\mu}_S$ be the sample mean of a simple random sample of the same size as in* (1)*. Then, $\hat{\mu}_A$ is an unbiased estimator of the population mean $\mu$ and*

$$\frac{\sigma^2}{n^2} \le Var[\hat{\mu}_A] = \frac{\sigma^2}{n} + \frac{1}{n(n-1)}\left[2\sigma^2 - \sum_{k=1}^{n}\{\sigma^2_{(k\cdot)} + (\mu_{(k\cdot)} - \mu)^2 + \sigma^2_{(\cdot k)} + (\mu_{(\cdot k)} - \mu)^2\}\right] \le \frac{\sigma^2}{n}.$$

**Proof.** Let $Y_{ij} = X_{(ij)} - \mu$. Consider$\mathcal{P}$, the class of all permutations on $\{1, \ldots, n\}$. We obtain

$$\begin{aligned} n\mathrm{E}[\hat{\mu}_A - \mu] &= \frac{1}{n!}\sum_{\pi\in\mathcal{P}} \mathrm{E}[\sum_i Y_{i\pi_i}] \\ &= \mathrm{E}[\frac{1}{n}\sum_{i\,k} Y_{ik}] \\ &= n\mathrm{E}[\bar{X} - \mu] \\ &= 0. \end{aligned}$$

Furthermore, to obtain $\mathrm{Var}(\hat{\mu}_A)$, we have

$$
\begin{aligned}
n^2\mathrm{Var}(\hat{\mu}_A) &= \frac{1}{n!}\sum_{\pi\in\mathcal{P}} \mathrm{E}[\{\sum_i Y_{i\pi_i}\}^2] \\
&= \frac{1}{n!}\sum_{\pi\in\mathcal{P}} \mathrm{E}[\sum_i Y_{i\pi_i}^2 + \sum_{i\neq j} Y_{i\pi_i}Y_{j\pi_j}] \\
&= \mathrm{E}[\frac{1}{n}\sum_{i\,k} Y_{ik}^2 + \frac{1}{n(n-1)}\sum_{i\neq j}\sum_{k\neq\ell} Y_{ik}Y_{j\ell}] \\
&= \mathrm{E}[\frac{1}{n}\sum_{i\,k} Y_{ik}^2 + \frac{1}{n(n-1)}\{\sum_{i\,j\,k\,\ell} Y_{ik}Y_{j\ell} - \sum_{i\,j\,k} Y_{ik}Y_{jk} - \sum_{i\,k\,\ell} Y_{ik}Y_{i\ell} + \sum_{i\,k} Y_{ik}^2\}] \\
&= \mathrm{E}[\frac{1}{n}\sum_{i\,k} Y_{ik}^2 + \frac{1}{n(n-1)}\{(\sum_{i\,k} Y_{ik})^2 - \sum_k(\sum_i Y_{ik})^2 - \sum_i(\sum_k Y_{ik})^2 + \sum_{i\,k} Y_{ik}^2\}] \\
&= \mathrm{E}[\frac{1}{n}\sum_{i\,k} Y_{ik}^2 + \frac{1}{n(n-1)}\{n^4 Y_{\cdot\cdot}^2 - n^2\sum_k Y_{\cdot k}^2 - n^2\sum_k Y_{k\cdot}^2 + \sum_{i\,k} Y_{ik}^2\}] \\
&= \mathrm{E}[n^2 Y_{\cdot\cdot}^2 + \frac{1}{(n-1)}\sum_{i\,k}\{Y_{ik} - Y_{\cdot k} - Y_{i\cdot} + Y_{\cdot\cdot}\}^2] \\
&= \sigma^2 + \frac{1}{(n-1)}\mathrm{E}[\sum_{i\,j}\{Y_{ij} - Y_{i\cdot} - Y_{\cdot j} + Y_{\cdot\cdot}\}^2] \\
&= \sigma^2 + \frac{1}{(n-1)}\sum_{i\,j}[\{\sigma_{(ij)}^2 - \sigma_{(i\cdot)}^2 - \sigma_{(\cdot j)}^2 + \sigma_{(\cdot\cdot)}^2\} + \{\mu_{(ij)} - \mu_{(i\cdot)} - \mu_{(\cdot j)} + \mu_{(\cdot\cdot)}\}^2] \\
&= \frac{n}{(n-1)}[(n+1)\sigma^2 - \sum_k\{\sigma_{(k\cdot)}^2 + (\mu_{(k\cdot)} - \mu)^2 + \sigma_{(\cdot k)}^2 + (\mu_{(\cdot k)} - \mu)^2\}],
\end{aligned}
$$

which leads to the result. In particular, note that

$$\mathrm{Var}(\hat{\mu}_A) = \frac{\sigma^2}{n^2} + \frac{1}{n^2(n-1)}\mathrm{E}[\sum_{i\,j}\{Y_{ij} - Y_{i\cdot} - Y_{\cdot j} + Y_{\cdot\cdot}\}^2] \geq \frac{\sigma^2}{n^2}.$$

In order to show that $\mathrm{Var}(\hat{\mu}_A) \leq \frac{\sigma^2}{n}$ we will verify that

$$\frac{1}{n-1}\sum_{i\,k}\{Y_{ik} - Y_{\cdot k} - Y_{i\cdot} + Y_{\cdot\cdot}\}^2 \leq \frac{1}{n+1}\sum_{i\,k}\{Y_{ik} - Y_{\cdot\cdot}\}^2,$$

and use the fact that

$$\mathrm{Var}(\sum_{i\,k}\{Y_{ik} - Y_{\cdot\cdot}\}^2) = \mathrm{Var}(\sum_i (X_i - \bar{X})^2) = (n^2-1)\sigma^2.$$

First of all, if $\alpha_1 \leq \cdots \leq \alpha_n$ and $\beta_1 \leq \cdots \leq \beta_n$ then $\sum_k (\alpha_k - \bar{\alpha})(\beta_k - \bar{\beta}) \geq 0$. This implies that

$$\sum_j (Y_{ij} - Y_{i\cdot})(Y_{kj} - Y_{k\cdot}) \geq 0, \quad i, k = 1, \ldots, n,$$

and

$$\sum_k (Y_{ki} - Y_{k\cdot})(Y_{kj} - Y_{k\cdot}) \geq 0, \quad i, j = 1, \ldots, n.$$

Notice also that $a_1 \leq a_2$ and $b_1 \leq b_2$ implies that $(a_1 + b_1)^2 + (a_2 + b_2)^2 \geq (a_1 + b_2)^2 + (a_2 + b_1)^2$. Similarly, if $Z_1 \leq \cdots \leq Z_{n^2}$ then $(Z_1 + \cdots + Z_n)^2 + \cdots + (Z_{n(n-1)+1} + \cdots + Z_{n(n-1)+n})^2 \geq (Z_{\pi_1} + \cdots + Z_{\pi_n})^2 + \cdots + (Z_{\pi_{n(n-1)+1}} + \cdots + Z_{\pi_{n(n-1)+n}})^2$ for all permutation $\pi = \{\pi_1, \ldots, \pi_n\} \in \mathcal{P}$. This is useful to show that

$$\begin{aligned} \sum_{ij} (Y_{ij} - Y_{i\cdot})^2 &= \sum_{ij} Y_{ij}^2 - n \sum_i Y_{i\cdot}^2 \\ &\leq \sum_{ij} Y_{ij}^2 - n \sum_j Y_{\cdot j}^2 \\ &= \sum_{ij} (Y_{ij} - Y_{\cdot j})^2. \end{aligned}$$

Finally,

$$
\begin{aligned}
\frac{1}{n-1}\sum_{ij}(Y_{ij}-Y_{i\cdot}-Y_{\cdot j}+Y_{\cdot\cdot})^2 &= \frac{1}{n-1}[\sum_{ij}(Y_{ij}-Y_{i\cdot})^2+\sum_{ij}(Y_{\cdot j}-Y_{\cdot\cdot})^2] \\
&= \frac{1}{n}\sum_{ij}(Y_{ij}-Y_{i\cdot})^2 \\
&\quad -\frac{1}{n(n-1)}[n\sum_{ij}(Y_{\cdot j}-Y_{\cdot\cdot})^2-\sum_{ij}(Y_{ij}-Y_{i\cdot})^2] \\
&= \frac{1}{n}\sum_{ij}(Y_{ij}-Y_{i\cdot})^2 \\
&\quad -\frac{1}{n(n-1)}[\sum_{j}\{\sum_{i}(Y_{ij}-Y_{i\cdot})\}^2-\sum_{ij}(Y_{ij}-Y_{i\cdot})^2] \\
&= \frac{1}{n}\sum_{ij}(Y_{ij}-Y_{i\cdot})^2 \\
&\quad -\frac{1}{n(n-1)}[\sum_{j}\{\sum_{i\neq k}(Y_{ij}-Y_{i\cdot})(Y_{kj}-Y_{k\cdot})\}] \\
&= \frac{1}{n}\sum_{ij}(Y_{ij}-Y_{i\cdot})^2 \\
&\quad -\frac{1}{n(n-1)}[\sum_{i\neq k}\{\sum_{j}(Y_{ij}-Y_{i\cdot})(Y_{kj}-Y_{k\cdot})\}] \\
&\leq \frac{1}{n}\sum_{ij}(Y_{ij}-Y_{i\cdot})^2 \\
&\leq \frac{1}{n+1}[\sum_{ij}(Y_{ij}-Y_{i\cdot})^2+\frac{1}{n}\sum_{ij}(Y_{ij}-Y_{\cdot j})^2] \\
&\leq \frac{1}{n+1}[\sum_{ij}(Y_{ij}-Y_{i\cdot})^2+\frac{1}{n}\sum_{ij}(Y_{ij}-Y_{\cdot j})^2] \\
&\quad +\frac{1}{n(n+1)}\sum_{i\neq j}(\sum_{k}(Y_{ki}-Y_{\cdot i})(Y_{kj}-Y_{\cdot j})) \\
&= \frac{1}{n+1}[\sum_{ij}(Y_{ij}-Y_{i\cdot})^2+\frac{1}{n}\sum_{i}\{\sum_{j}(Y_{ij}-Y_{\cdot j})\}^2)] \\
&= \frac{1}{n+1}[\sum_{ij}(Y_{ij}-Y_{i\cdot})^2+\sum_{ij}(Y_{i\cdot}-Y_{\cdot\cdot})^2] \\
&= \frac{1}{n+1}\sum_{ij}(Y_{ij}-Y_{\cdot\cdot})^2.
\end{aligned}
$$

We conclude this section with an application of the obtained results to the special case of the uniform distribution because here the variances have very simple forms.

**Example 2.1** *Consider a $\mathcal{U}(0,1)$ population of size $n$. Here, $\sigma^2 = \frac{1}{12}$,*

$$E[X_{(i)}] = \frac{i}{n^2+1} \quad \textit{and} \quad Cov(X_{(i)}, X_{(j)}) = \frac{i(n^2-j+1)}{(n^2+1)^2(n^2+2)}, \quad 1 \le i \le j \le n^2,$$

*so*

$$Var(\hat{\mu}_S) = \frac{1}{n}\sigma^2, \quad Var(\hat{\mu}_R) = \frac{2}{n(n+1)}\sigma^2, \quad Var(\hat{\mu}_A) = \frac{n^2+5}{(n^2+1)(n^2+2)}\sigma^2.$$

*We have already shown, in Theorem 2.1, that $\frac{\sigma^2}{n^2}$ is a lower bound for $Var(\hat{\mu}_A)$. Also, it is shown (see Theorem 3.1) that, $\frac{\sigma^2}{n^2}$ is a lower bound for $Var(\hat{\mu}_R)$ as well. Therefore we can analyse how much do we lose in comparison to the lower bound $\frac{\sigma^2}{n^2}$ by using $\hat{\mu}_A$ or $\hat{\mu}_R$ in estimating the population mean $\mu = \frac{1}{2}$. For the $\mathcal{U}(0,1)$ distribution*

$$(n^2\, Var(\hat{\mu}_R) - \sigma^2) = \frac{n-1}{n+1} = O(1),$$

*while*

$$(n^2\, Var(\hat{\mu}_A) - \sigma^2) = 2\frac{(n^2-1)}{(n^2+1)(n^2+2)} = O(\frac{1}{n^2}).$$

# 3 A special class of random estimators

In this section we propose a class of random estimators of the mean and develop a general dominance result leading to a sufficient condition for a random estimator $\hat{\mu}_1$ to dominate another random estimator $\hat{\mu}_2$. Our treatment permits consideration of a more general type of estimators than that discussed in the preceding section and includes the estimators $\hat{\mu}_S$, $\hat{\mu}_R$ and $\hat{\mu}_A$ as special cases. We are concern with random estimators $\hat{\mu}$ that can be written in the form

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n^2} I_j X_{(j)}, \tag{5}$$

where $I_1, \ldots, I_{n^2}$ are Bernoulli($\frac{1}{n}$) random variables, independent of $X_{(1)}, \ldots, X_{(n^2)}$, satisfying the regularity condition $\sum_{j=1}^{n^2} I_j = n$. In the above representation of the random estimator $\hat{\mu}$ the inclusion probabilities of the first and second orders are given by $p_i = \mathrm{E}[I_i] = \frac{1}{n}$ and $p_{ij} = \mathrm{E}[I_i I_j]$, $i, j = 1, \ldots, n^2$, respectively. In particular, we have $\mathrm{Cov}(I_i, I_j) = p_{ij} - \frac{1}{n^2}$ and $\sum_{i=1}^{n^2} p_{ij} = \sum_{j=1}^{n^2} p_{ij} = 1$, $i, j = 1, \ldots, n^2$.

**Remark 3.1** *Let $R_1, \ldots, R_{n^2}$ be the rank statistics corresponding to the sample $X_1, \ldots, X_{n^2}$ and assume that the ties are broken randomly. The rank statistics are independent of $X_{(1)}, \ldots, X_{(n^2)}$. Therefore, the construction of the random variables $I_1, \ldots, I_{n^2}$ could involve the rank statistics.*

**Example 3.1** *(Simple Random Sampling). The estimator $\hat{\mu}_S$ as in* (1)*, corresponding to a simple random sample of size $n$ from the underlying population, is in the form of* (5) *with Bernoulli random variables given by*

$$I_j = \begin{cases} 1 & \text{if } j \in \{R_i \colon i = 1 \ldots, n\} \\ 0 & \text{otherwise} \end{cases}, \quad j = 1, \ldots, n^2,$$

*and $p_{ij} = \frac{1}{n(n+1)}$, $1 \leq i < j \leq n^2$.*

**Example 3.2** *(Ranked Set Sampling). Let $\{R_{i(k)} \colon k = 1, \ldots, n\}$ be the order statistics based on $\{R_{ik} \colon k = 1, \ldots, n\}$, $i = 1, \ldots, n$. The estimator $\hat{\mu}_R$ as in* (2)*, corresponding to a ranked set sample of size $n$ from the underlying population, is in the form of* (5) *with Bernoulli random variables given by*

$$I_j = \begin{cases} 1 & \text{if } j \in \{R_{i(i)} \colon i = 1 \ldots, n\} \\ 0 & \text{otherwise} \end{cases}, \quad j = 1, \ldots, n^2.$$

*To obtain the corresponding probability of inclusion of the second order, consider a ranked set sampling algorithm with $n = 2$. There are six ways to put $\{X_{(1)}, \ldots, X_{(4)}\}$ in two rows which are,*

$$\begin{pmatrix} X_{(1)} & X_{(2)} \\ X_{(3)} & X_{(4)} \end{pmatrix} \begin{pmatrix} X_{(1)} & X_{(3)} \\ X_{(2)} & X_{(4)} \end{pmatrix} \begin{pmatrix} X_{(1)} & X_{(4)} \\ X_{(2)} & X_{(3)} \end{pmatrix}$$

$$\begin{pmatrix} X_{(2)} & X_{(3)} \\ X_{(1)} & X_{(4)} \end{pmatrix} \begin{pmatrix} X_{(2)} & X_{(4)} \\ X_{(1)} & X_{(3)} \end{pmatrix} \begin{pmatrix} X_{(3)} & X_{(4)} \\ X_{(1)} & X_{(2)} \end{pmatrix}$$

*leading to the following selections, $(X_{(1)}, X_{(4)})$, $(X_{(1)}, X_{(4)})$, $(X_{(1)}, X_{(3)})$, $(X_{(2)}, X_{(4)})$, $(X_{(2)}, X_{(3)})$, $(X_{(2)}, X_{(3)})$. Therefore, $p_{12} = 0$, $p_{13} = \frac{1}{6}$, $p_{14} = \frac{1}{3}$, $p_{23} = \frac{1}{3}$, $p_{24} = \frac{1}{6}$ and $p_{34} = 0$. For $n > 2$ finding the $p_{ij}$ is a combinatorics problem which leads to the general solution*

$$p_{ij} = \frac{1}{n^2 - 1}[1 - \frac{A}{\binom{n^2-2}{n-1 \; n-1 \; n(n-2)}}], \quad 1 \leq i < j \leq n^2,$$

*with*

$$A = \sum_{1\le k\le \ell\le n} \binom{i-1}{k-1}\binom{j-i-1}{\ell-k}\binom{n^2-j}{n-\ell}\binom{j-\ell-1}{k-1}\binom{n(n-1)+\ell-j}{n-k}.$$

**Example 3.3** *(Antithetic Sampling). Let $\pi = \{\pi_1, \ldots, \pi_n\}$ be a random permutation on $\{1, \ldots, n\}$. The antithetic estimator $\hat{\mu}_A$ as in* (4) *is in the form of* (5) *and the corresponding Bernoulli random variables are given by*

$$I_j = \begin{cases} 1 & \textit{if } j \in \{n(i-1)+\pi_i : i = 1\ldots, n\} \\ 0 & \textit{otherwise} \end{cases}, \quad j = 1, \ldots, n^2,$$

*with*

$$p_{ij} = \begin{cases} 0 & \textit{if } \lceil \frac{i}{n} \rceil = \lceil \frac{j}{n} \rceil \textit{ and } i \not\equiv j \ (\textit{mod } n) \\ 0 & \textit{if } \lceil \frac{i}{n} \rceil \neq \lceil \frac{j}{n} \rceil \textit{ and } i \equiv j \ (\textit{mod } n) \\ \frac{1}{n(n-1)} & \textit{if } \lceil \frac{i}{n} \rceil \neq \lceil \frac{j}{n} \rceil \textit{ and } i \not\equiv j \ (\textit{mod } n) \end{cases}, \quad 1 \le i < j \le n^2.$$

We pursue with a Theorem and a Corollary regarding to the properties of the random estimators $\hat{\mu}$ as in (5); the latter one giving simple and general sufficient condition for dominance results.

**Theorem 3.1** *The random estimator $\hat{\mu}$ as in* (5) *is unbiased, in estimating the population mean $\mu$, with*

$$Var[\hat{\mu}] = \frac{1}{n^2}\{\sigma^2 + \sum_{i=1}^{n^2-1}\sum_{j=1}^{n^2-1} \tau_{ij} E[S_i S_j]\} \ge \frac{\sigma^2}{n^2},$$

*where*

$$\tau_{ij} = Cov[\sum_{k=1}^{i} I_k, \sum_{\ell=1}^{j} I_\ell] = \sum_{k=1}^{i}\sum_{\ell=1}^{j}[p_{k\ell} - \frac{1}{n^2}] \textit{ and } S_i = X_{(i+1)} - X_{(i)}, \quad i, j = 1, \ldots, n^2-1.$$

**Proof.** The expectation of $\hat{\mu}$ is given by

$$\begin{aligned} \mathrm{E}[\hat{\mu}] &= \mathrm{E}[\mathrm{E}[\hat{\mu}|X_{(1)}, \ldots, X_{(n^2)}]] \\ &= \mathrm{E}[\bar{X}] \\ &= \mu. \end{aligned}$$

To obtain the variance of $\hat{\mu}$, we have

$$\begin{aligned}
\mathrm{Var}[\hat{\mu}] &= \mathrm{Var}[\mathrm{E}[\hat{\mu}|X_{(1)},\ldots,X_{(n^2)}]] + \mathrm{E}[\mathrm{Var}[\hat{\mu}|X_{(1)},\ldots,X_{(n^2)}]] \\
&= \mathrm{Var}[\bar{X}] + \frac{1}{n^2}\mathrm{E}[\sum_{i=1}^{n^2}\sum_{j=1}^{n^2}(p_{ij}-\frac{1}{n^2})X_{(i)}X_{(j)}] \\
&= \frac{1}{n^2}\{\sigma^2 + \mathrm{E}[\sum_{i=1}^{n^2-1}\sum_{j=1}^{n^2-1}\tau_{ij}S_iS_j]\} \\
&\geq \frac{\sigma^2}{n^2},
\end{aligned}$$

and the inequality holds because $\mathrm{Var}[\hat{\mu}] - \frac{\sigma^2}{n^2} = \mathrm{E}[\mathrm{Var}[\hat{\mu}|X_{(1)},\ldots,X_{(n^2)}]] \geq 0$.

**Corollary 3.2** *If $\hat{\mu}$ and $\hat{\mu}^*$ are two random estimators, in the form of* (5)*, such that $\tau_{ij} \geq \tau^*_{ij}$; for all $i, j = 1,\ldots,n^2-1$; then $Var[\hat{\mu}] \geq Var[\hat{\mu}^*]$.*

**Proof.** The proof is based on the fact that $S_i \geq 0$ for $i = 1,\ldots,n^2-1$.

We pursue with various applications of the above results.

**Example 3.4** *In this example we provide a new proof for superiority of $\hat{\mu}_A$ over $\hat{\mu}_S$ in estimating the population mean in the spirit of Corollary 3.2. By virtue of Corollary 3.2, we only need to obtain $\tau_{Aij}$ and $\tau_{Sij}$ corresponding to $\hat{\mu}_A$ and $\hat{\mu}_S$ and show that $\tau_{Aij} \leq \tau_{Sij}$ for all $i, j = 1,\ldots,n^2-1$. Let*

$$i_1 = \lceil\frac{i}{n}\rceil, i_2 = i - n(i_1 - 1),$$

$$j_1 = \lceil\frac{j}{n}\rceil, j_2 = j - n(j_1 - 1).$$

*We obtain*

$$\tau_{Sij} = \frac{1}{n+1}(i\wedge j)(1-\frac{i\vee j}{n^2}),$$

$$\tau_{Aij} = \begin{cases} \frac{1}{n^2}(i_2\wedge j_2)(n-(i_2\vee j_2)) & \text{if} \quad i_1 = j_1 \\ -\frac{1}{n^2(n-1)}(i_2\wedge j_2)(n-(i_2\vee j_2)) & \text{if} \quad i_1 < j_1 \end{cases}.$$

*When* $1 \le i_1 < j_1 \le n$ *we have* $\tau_{Aij} \le 0 \le \tau_{Sij}$. *Assume that* $i_1 = j_1$. *In this case*

$$\tau_{Sij} = \frac{1}{n^2(n+1)}[n(i_1-1)+(i_2 \wedge j_2)][n(n-i_1)+(n-(i_2 \vee j_2))].$$

*If* $i_1 > (i_2 \wedge j_2)$ *then* $[n(i_1-1)+(i_2 \wedge j_2)] \ge (n+1)(i_2 \wedge j_2)$, *and*

$$[n(n-i_1)+(n-(i_2 \vee j_2))] \ge (n-(i_2 \vee j_2)).$$

*So* $\tau_{Sij} \ge \tau_{Aij}$. *If* $i_1 \le (i_2 \wedge j_2)$ *then* $[n(n-i_1)+(n-(i_2 \vee j_2))] \ge (n+1)(n-(i_2 \vee j_2))$, *and*

$$[n(i_1-1)+(i_2 \wedge j_2)] \ge (i_2 \wedge j_2).$$

*So* $\tau_{Sij} \ge \tau_{Aij}$. *Therefore,* $\tau_{Sij} \ge \tau_{Aij}$ *for all* $1 \le i \le j \le n^2-1$.

In the reminding of this section, we discuss the problem of superiority of the antithetic sampling over the ranked set sampling in estimating the population mean. First, the following Example highlights the attractive feature of $\hat{\mu}_A$ in dominating $\hat{\mu}_R$ for $n=2$. As we see in the next section, this latter property of $\hat{\mu}_A$ leads us to introduce a new variation of the ranked set sampling and correspondingly a new estimator, $\hat{\mu}_V$, which outperforms $\hat{\mu}_R$.

**Example 3.5** . *If* $n=2$ *then*

$$Var[\hat{\mu}_R|X_{(1)},\ldots,X_{(4)}] - Var[\hat{\mu}_A|X_{(1)},\ldots,X_{(4)}] = \frac{1}{24}[(X_{(2)}-X_{(1)})(X_{(4)}-X_{(3)})] \ge 0.$$

In the following two Examples we consider the performance of $\hat{\mu}_A$ in comparison with $\hat{\mu}_R$ for $n=3$ and $n=4$. Let $T=(\tau_{ij})_{i,j=1,\ldots,n^2-1}$. Note that for all $n \ge 2$ we have

$$(T_R-T_A)_{ij} = (T_R-T_A)_{ji} = (T_R-T_A)_{n^2-i,n^2-j}, \quad i,j=1,\ldots,n^2-1.$$

**Example 3.6** . *For $n = 3$, the comparison between $\hat{\mu}_A$ and $\hat{\mu}_R$ leads to*

$$100(T_R - T_A) = \begin{pmatrix} 0 & 0 & 3.57 & 10.12 & 2.38 & -3.57 & 8.33 & 4.17 \\ 0 & 0 & 10.71 & 8.33 & 9.05 & -4.29 & 1.19 & 8.33 \\ 3.57 & 10.71 & 21.43 & 10.71 & 2.86 & -2.14 & -4.29 & -3.57 \\ 10.12 & 8.33 & 10.71 & 0 & 0 & 2.86 & 9.05 & 2.38 \\ 2.38 & 9.05 & 2.86 & 0 & 0 & 10.71 & 8.33 & 10.12 \\ -3.57 & -4.29 & -2.14 & 2.86 & 10.71 & 21.43 & 10.71 & 3.57 \\ 8.33 & 1.19 & -4.29 & 9.05 & 8.33 & 10.71 & 0 & 0 \\ 4.17 & 8.33 & -3.57 & 2.38 & 10.12 & 3.57 & 0 & 0 \end{pmatrix}.$$

*In particular,*

$$\sum_{i,j \in \{1,6\}} (\tau_{Rij} - \tau_{Aij}) S_i S_j \to -\infty \text{ when } S_6 = 1 \text{ and } S_1 \to \infty.$$

*Therefore, at this point, we cannot claim that $Var[\hat{\mu}_A] \leq Var[\hat{\mu}_R]$ for $n = 3$. On the other hand, for Bernoulli(p) trials, $S_i S_j = 0$ if $i \neq j$ and*

$$Var[\hat{\mu}] = \frac{pq}{n^2} + \sum_{k=1}^{n^2-1} \tau_{kk} P[Binomial(n^2, p) = n^2 - k + 1].$$

*We obtain that*

$$\begin{aligned} \frac{Var[\hat{\mu}_R] - Var[\hat{\mu}_A]}{Var[\hat{\mu}_S]} &= \frac{3 \times 0.2143}{pq} \{P[Binomial(9,p) = 3] + P[Binomial(9,p) = 6]\} \\ &= 54(pq)^2(1 - 3pq). \end{aligned}$$

*In this application we see that $(pq)^2(1 - 3pq)$ reaches its maximum when $p = \frac{1}{3}$ and $p = \frac{2}{3}$.*

**Example 3.7** . *For $n = 4$, the comparison between $\hat{\mu}_A$ and $\hat{\mu}_R$ leads to*

$$100(T_R - T_A)_{i=1,\ldots,15, j=1,\ldots,8} = \begin{pmatrix} 0 & 0 & 1.43 & 4.07 & 7.71 & 3.85 & 0.64 & -2.05 \\ 0 & 0 & 4.29 & 10.33 & 9.56 & 10.09 & 3,38 & -2.42 \\ 1.43 & 4.29 & 8.57 & 18.57 & 13.61 & 10.17 & 8.10 & -1.12 \\ 4.07 & 10.33 & 18.57 & 28.57 & 19.60 & 12.22 & 6.34 & 1.86 \\ 7.71 & 9.56 & 13.61 & 19.60 & 10.62 & 7.71 & 6.34 & 6.48 \\ 3.85 & 10.09 & 10.17 & 12.22 & 7.71 & 4.80 & 7.99 & 12.73 \\ 0.64 & 3.38 & 8.10 & 6.34 & 6.34 & 7.99 & 11.19 & 20.57 \\ -2.05 & -2.42 & -1.12 & 1.86 & 6.48 & 12.73 & 20.57 & 29.96 \\ 3.97 & 0.84 & -0.97 & -1.36 & 8.06 & 10.68 & 14.84 & 20.57 \\ 1.94 & 4.68 & 0.06 & -3.43 & 2.66 & 10.11 & 10.68 & 12.73 \\ 0.09 & 0.63 & 1.82 & -4.49 & -1.48 & 2.66 & 8.06 & 6.48 \\ -1.65 & -3.10 & -4.17 & -4.68 & -4.49 & -3.43 & -1.36 & 1.86 \\ 5.00 & 1.72 & -1.38 & -4.17 & 1.82 & 6.32 & -0.97 & -1.12 \\ 3.33 & 6.67 & 1.72 & -3.10 & 6.33 & 4.68 & 8.39 & -2.42 \\ 1.67 & 3.33 & 5.00 & -1.65 & 0.09 & 1.94 & 3.97 & -2.05 \end{pmatrix}$$

*Here again, $\tau_{Rii} \geq \tau_{Aii}$ for $i = 1, \ldots, 15$. Therefore, $Var(\hat{\mu}_R) \geq Var(\hat{\mu}_A)$ for Bernoulli$(p)$ trials.*

# 4 A variation of ranked set sampling

In this section, following the optimal property of $\hat{\mu}_A$ in dominating $\hat{\mu}_R$ for $n = 2$, explored in Example 3.5, we introduce another variation of the ranked set sampling and correspondingly another estimator of the population mean called $\hat{\mu}_V$ which outperforms $\hat{\mu}_R$. In fact, when $n = 2$, $\hat{\mu}_V$ is the antithetic estimator. Let $m_1 = \lfloor \frac{n}{2} \rfloor$ and $m_2 = \lceil \frac{n}{2} \rceil$. To construct $\hat{\mu}_V$ we start with a useful representation of the ranked set sampling estimator $\hat{\mu}_R$. Note that $\hat{\mu}_R$ in (2) can be written as follows

$$
\begin{aligned}
n\hat{\mu}_R &= \sum_{i=1}^{n} X_{i(i)} \\
&= \begin{cases} \sum_{k=1}^{m_1}(X_{k(k)} + X_{k+m_2\ (k+m_2)}) & \text{if } n \text{ is even} \\ \sum_{k=1}^{m_1}(X_{k(k)} + X_{k+m_2\ (k+m_2)}) + X_{m_2(m_2)} & \text{if } n \text{ is odd} \end{cases},
\end{aligned}
$$

Or equivalently

$$\hat{\mu}_R = \frac{1}{n}\sum_{k=1}^{m_2} Z_k, \tag{6}$$

with $Z_k = (X_{k(k)} + X_{k+m_2\ (k+m_2)})$; $k = 1, \ldots, m_1$. Now, we define our new proposed estimator as

$$\hat{\mu}_V = \frac{1}{n}\sum_{k=1}^{m_2} V_k, \tag{7}$$

with

$$V_k = \begin{cases} (X_{k(k)} \wedge X_{k+m_2\ (k)}) + (X_{k\ (k+m_2)} \vee X_{k+m_2\ (k+m_2)}) \text{ with probability } \frac{1}{2} \\ (X_{k(k)} \vee X_{k+m_2\ (k)}) + (X_{k\ (k+m_2)} \wedge X_{k+m_2\ (k+m_2)}) \text{ with probability } \frac{1}{2} \end{cases},$$

and $V_{m_2} = X_{m_2,(m_2)}$ if $n$ is even.

We now are ready to pursue with the following key Theorem of this section regarding to the optimal properties of $\hat{\mu}_V$ such as its superiority over $\hat{\mu}_R$ in estimating the population mean.

**Theorem 4.1** *The estimator $\hat{\mu}_V$ as in* (7) *is an unbiased estimator of the population mean and $Var(\hat{\mu}_V) \leq Var(\hat{\mu}_R)$.*

**Proof.** First of all, note that

$$\mathrm{E}[\hat{\mu}_R] = \frac{1}{n}\sum_{k=1}^{m_2}\mathrm{E}[Z_k], \mathrm{Var}[\hat{\mu}_R] = \frac{1}{n^2}\sum_{k=1}^{m_2}\mathrm{Var}[Z_k].$$

Similarly

$$\mathrm{E}[\hat{\mu}_V] = \frac{1}{n}\sum_{k=1}^{m_2}\mathrm{E}[V_k], \mathrm{Var}[\hat{\mu}_V] = \frac{1}{n^2}\sum_{k=1}^{m_2}\mathrm{Var}[V_k].$$

Let $1 \leq \ell_1 < \ell_2 \leq n$, and define

$$Z = X_{\ell_1(\ell_1)} + X_{\ell_2(\ell_2)} \quad V = \begin{cases} X_{\ell_1(\ell_1)} \wedge X_{\ell_2(\ell_1)} + X_{\ell_1(\ell_2)} \vee X_{\ell_2(\ell_2)} \text{ with probatility } \frac{1}{2} \\ X_{\ell_1(\ell_1)} \vee X_{\ell_2(\ell_1)} + X_{\ell_1(\ell_2)} \wedge X_{\ell_2(\ell_2)} \text{ with probatility } \frac{1}{2} \end{cases}$$

Now, consider the event $A_{ijkl}$, $1 \leq i < j < k < l \leq n^2$, given by

$$\{R_{\ell_1(\ell_1)}, R_{\ell_1(\ell_2)}, R_{\ell_2(\ell_1)}, R_{\ell_2(\ell_2)}\} = \{i, j, k, l\},$$

and assume that $(i, j, k, l)$ is such that $\mathrm{P}[A_{ijkl}] > 0$. When the event $A_{ijkl}$ occurs it implies that

$$(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) \in \{(i, l), (j, k), (i, k), (j, l)\}.$$

Since $\big((R_{\ell_1(\ell_1)}, R_{\ell_1(\ell_2)}), (R_{\ell_2(\ell_1)}, R_{\ell_2(\ell_2)})\big)$ and $\Big((R_{\ell_2(\ell_1)}, R_{\ell_2(\ell_2)}), (R_{\ell_1(\ell_1)}, R_{\ell_1(\ell_2)})\Big)$ have the same distribution we obtain that

$$\mathrm{P}[(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) = (i, l) | A_{ijkl}] = \mathrm{P}[(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) = (j, k) | A_{ijkl}].$$

Similarly

$$\mathrm{P}[(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) = (i, k) | A_{ijkl}] = \mathrm{P}[(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) = (j, l) | A_{ijkl}].$$

Let

$$\begin{aligned} \alpha_{ijkl} &= \mathrm{P}[(R_{\ell_1(\ell_1)} \wedge R_{\ell_2(\ell_2)}, R_{\ell_1(\ell_1)} \vee R_{\ell_2(\ell_2)}) \in \{(i, k), (j, l)\} | A_{ijkl}], \\ \mu_{ijkl} &= \frac{X_{(i)} + X_{(j)} + X_{(k)} + X_{(l)}}{2}. \end{aligned}$$

Conditionally on the event $A_{ijkl}$ we obtain that

$$Z = \begin{cases} X_{(i)} + X_{(l)} & \text{with probability} \quad \frac{(1-\alpha_{ijkl})}{2} \\ X_{(j)} + X_{(k)} & \text{with probability} \quad \frac{(1-\alpha_{ijkl})}{2} \\ X_{(i)} + X_{(k)} & \text{with probability} \quad \frac{\alpha_{ijkl}}{2} \\ X_{(j)} + X_{(l)} & \text{with probability} \quad \frac{\alpha_{ijkl}}{2} \end{cases}, V = \begin{cases} X_{(i)} + X_{(l)} & \text{with probability} \quad \frac{1}{2} \\ X_{(j)} + X_{(k)} & \text{with probability} \quad \frac{1}{2} \end{cases}.$$

Now,

$$\mathrm{E}[Z|A_{ijkl}] = \mu_{ijkl} = \mathrm{E}[V|A_{ijkl}],$$

establishing the unbiasedness of $\hat{\mu}_V$. Finally,

$$\begin{aligned} \mathrm{Var}[Z|A_{ijkl}] &= (X_{(i)} + X_{(l)} - \mu_{ijkl})^2(1-\alpha_{ijkl}) + (X_{(i)} + X_{(k)} - \mu_{ijkl})^2\alpha_{ijkl} \\ &= (X_{(i)} + X_{(l)} - \mu_{ijkl})^2 + \frac{\alpha_{ijkl}}{2}(X_{(j)} - X_{(i)})(X_{(l)} - X_{(k)}) \\ &= \mathrm{Var}[V|A_{ijkl}] + \frac{\alpha_{ijkl}}{2}(X_{(j)} - X_{(i)})(X_{(l)} - X_{(k)}) \\ &\geq \mathrm{Var}[V|A_{ijkl}]. \end{aligned}$$

# 5 Simulation study

In this section, we carry out a simulation study to compare numerically the performance of the proposed estimators $\hat{\mu}_A$ and $\hat{\mu}_V$ with that of $\hat{\mu}_R$. It has already been established, in Theorem (3.1), that $\frac{\sigma^2}{n^2}$ is a lower bound for the variances of all these estimators. The aim is to reduce, as much as possible, the gap between the variance of the proposed estimators and the lower bound $\frac{\sigma^2}{n^2}$. For this reason, given a distribution $P$ and a sample size $n$, the comparison between an estimator $\hat{\mu}$ and $\hat{\mu}_R$ will be based on the measure $\mathcal{R}_n$ given by

$$\mathcal{R}_n(P, \hat{\mu}) = \frac{\mathrm{Var}_P(\hat{\mu}) - \frac{\sigma^2}{n^2}}{\mathrm{Var}_P(\hat{\mu}_R) - \frac{\sigma^2}{n^2}}.$$

In our simulation study we worked on four underlying distributions: the standard normal distribution $\mathcal{N}(0,1)$, the exponential distribution $Exp(1)$, the uniform distribution $\mathcal{U}(0,1)$, and the Poisson distribution $Poisson(1)$.

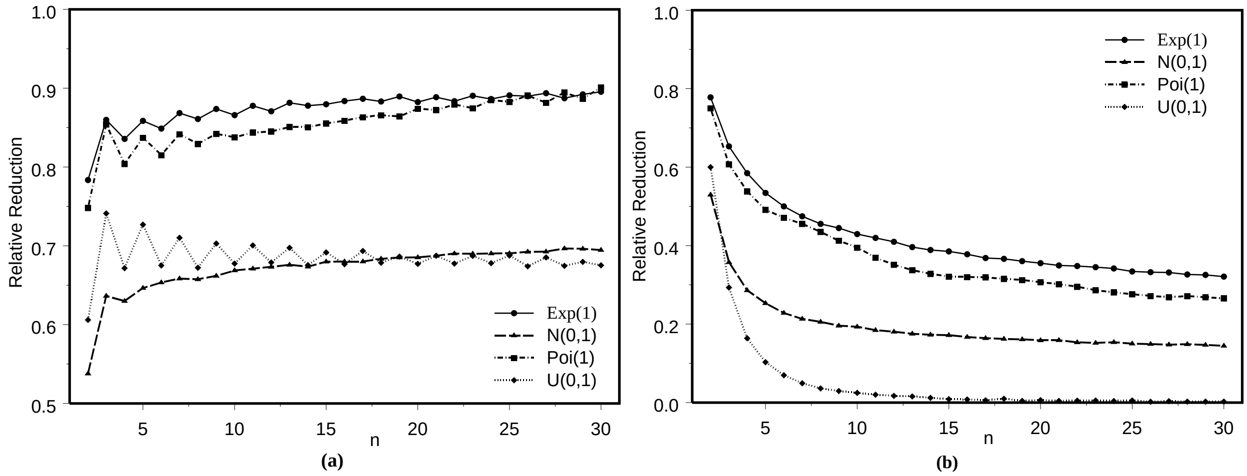


Figure 1: Simulated values of $\mathcal{R}_n(P, \hat{\mu})$ for considered distributions: (a) $\mathcal{R}_n(P, \hat{\mu}_V)$ (b) $\mathcal{R}_n(P, \hat{\mu}_A)$.

For a given sample size $n$, $n \in \{2, \ldots 30\}$, and a given distribution $P$, Figure 1 displays the simulated values of $\mathcal{R}_n(P, \hat{\mu}_A)$ and $\mathcal{R}_n(P, \hat{\mu}_V)$ based on one million iterations for the considered distributions. The simulations show that $\mathcal{R}_n(P, \hat{\mu}_V)$ varies a little bit with $n$ for small values of $n$ and becomes more stable when $n$ gets larger. Overall, the gab between $\mathrm{Var}_P(\hat{\mu}_R)$ and $\frac{\sigma^2}{n^2}$ has been reduced by a factor of at least 10% using $\hat{\mu}_V$.

On the other hand, the simulations indicate that $\hat{\mu}_A$ is always the best estimator. The theory tells us that $\mathcal{R}_2(P, \hat{\mu}_A) \leq 1$ for all $P$. Moreover, the simulations indicate that $\mathcal{R}_n(P, \hat{\mu}_A)$ decreases in $n$ when $P$ is fixed. This behavior has been observed in all of the other simulations that we have conducted so far. Perhaps $\mathcal{R}_n(P, \hat{\mu}_A) \leq 1$ for all $P$, $n \geq 1$.

# References


Al–Nasser, A.D. (2007) L–Ranked Set Sampling: A Generalization Procedure For Robust Visual Sampling. *Communications in Statistics: Theory and Methods*, **36**, 33–43.

Al–Saleh, M.F. and Al–Omari, A. (2002) Multistage Ranked Set Sampling. *Journal of Statistical Planning and Inference* , **102**, 273–286.

Al–Saleh, M.F. and Al–Kadiri, M.A. (2000) Double Ranked Set Sampling. *Statistics & Probability Letters*, **48**, 205–212.

Bhoj, D.S. (1997) New Parametric Ranked Set Sampling. *Journal of Applied Statistical Science*, **6** , 275–289.

Chen, Z., Bai, Z.D., and Sinha, B.K. (2004) Ranked Set Sampling: Theory and Applications. *Lecture Notes in Statistics*, **176**, Springer–Verlag, New York.

Dell, T.R. (1969) The Theory of Some Applications of Ranked Set Sampling. *Ph.D. Thesis*, University of Georgia, Athens, GA.

Dell, T.R. and Clutter, J.L. (1972) Ranked Set Sampling With Order Statistics Background. *Biometrics*, **28**, 545–553.

Hossain, S.S. and Muttlak, H.A. (1999) Paired Ranked Set Sampling: A More Efficient Procedure. *Environmetrics*, **10**, 195–212.

Hossain, S.S. and Muttlak, H.A. (2001) Selected Ranked Set Sampling. *Australian & New Zealand Journal of Statistics*, **43**, 311–325.

Kaur, A., Patil, G.P., Sinha, A.K., and Taillie, C. (1995) Ranked Set Sampling: An Annotated Bibliography. *Environmental and Ecological Statistics*, **2**, 25–54.

Li., D., Sinha, B.K., and Perron, F. (1999) Random Selection in Ranked Set Sampling and Its Applications. *Journal of Statistical Planning and Inference*, **76**, 185–201.

McIntyre, G.A. (1952) A Method for Unbiased Selective Sampling, Using Ranked Sets. *Australian Journal of Agricultural Research*, **3**, 385–390.

Mutllak, H.A. (1997) Median Ranked Set Sampling. *Journal of Applied Statistical Science*, **6**, 245–255.

Samawi, H., Abu–Daayeh, H.A., and Ahmed, S. (1996) Estimating the Population Mean Using Extreme Ranked Set Sampling. *Biometrical Journal*, **38**, 577–586.

Takahasi, K. and Wakimoto, K. (1968) On Unbiased Estimates of the Population Mean Based on the Sample Stratified by Means of Ordering. *Annals of the Institute of Statistical Mathematics*, **20**, 1–31.