Computers, Materials & Continua
DOI:10.32604/cmc.2021.015047
Article

New Improved Ranked Set Sampling Designs with an Application to Real Data

Amer Ibrahim Al-Omari[1] and Ibrahim M. Almanjahie[2,3,*]

[1]Department of Mathematics, Faculty of Science, Al al-Bayt University, Mafraq, Jordan
[2]Department of Mathematics, College of Science, King Khalid University, Abha, 62529, Saudi Arabia
[3]Statistical Research and Studies Support Unit, King Khalid University, Abha, 62529, Saudi Arabia
[*]Corresponding Author: Ibrahim M. Almanjahie. Email: imalmanjahi@kku.edu.sa
Received: 04 November 2020; Accepted: 08 December 2020

Abstract: This article proposes two new Ranked Set Sampling (RSS) designs for estimating the population parameters: Simple Z Ranked Set Sampling (SZRSS) and Generalized Z Ranked Set Sampling (GZRSS). These designs provide unbiased estimators for the mean of symmetric distributions. It is shown that for non-uniform symmetric distributions, the estimators of the mean under the suggested designs are more efficient than those obtained by RSS, Simple Random Sampling (SRS), extreme RSS and truncation based RSS designs. Also, the proposed RSS schemes outperform other RSS schemes and provide more efficient estimates than their competitors under imperfect rankings. The suggested mean estimators under perfect and imperfect rankings are more efficient than the linear regression estimator under SRS. Our proposed RSS designs are also extended to cover the estimation of the population median. Real data is used to examine wthe usefulness and efficiency of our estimators.

Keywords: Ranked set sampling; unbiased estimator; simple random sampling; mean squared error; efficiency; imperfect ranking

1 Introduction

The Ranked Set Sampling (RSS) is originally derived by McIntyre [1] as a new design to increase the efficiency of pasture and forage yields estimates for fixed sample units. The RSS is considered when the study variable can simply be ranked than quantified. Takahasi et al. [2], independently, introduced the background of the RSS design, mathematically. It is shown that mean estimator by the RSS is unbiased, and provides more efficient estimates than the simple random sampling mean estimator. Even when the measured observations are ranked with errors, the RSS still provides an unbiased estimator, but the imperfect ranking is generally better than ordering based on random [3]. Stokes [4] considered the case of measuring the variable of interest and concluded that the study variable can be ranked by some concomitant variables. The competence of the estimator then depends on the relation between the study variables and the ancillary variables. With perfect ranking, the estimation based on RSS is more adequate as compared to the regression estimation based on SRS, especially, when the study variables and the ancillary variables are highly correlated (say $|\rho| > 0.85$) [5].

In the last few decades, many applications and modifications of the RSS design have been proposed. Halls et al. [6] for considering an application of forage yields using RSS. Samawi et al. [7] introduced the Extreme RSS (ERSS) design. Al-Omari et al. [8] introduced ratio estimators of the population mean with missing values using RSS. Al-Omari [9] considered the median estimation based on double robust extreme RSS. Al-Saleh et al. [10] extended the work further and provided Multistage RSS (MSRSS) design. They proved that as the number of stages increases, the efficiency of the mean estimator under MSRSS increases and vice versa. A Robust L RSS procedure based on the idea of L estimators is suggested by Al-Naseer [11]. Muttlak [12] introduced Median RSS (MRSS); he showed that it provides an unbiased estimator of the mean of symmetric distributions, and is more efficient than the SRS and RSS mean estimators. Jemain et al. [13] suggested multistage median RSS for estimating the population median and Jemain et al. [14] proposed some variations of RSS. Al-Omari [15] proposed ratio estimators of the population mean by considering ancillary information in SRS and median RSS and Al-Omari [16] considered the entropy estimation in RSS methods. Hossain et al. [17] suggested paired RSS for estimating the population mean. Shadid et al. [18] considered the BLUEs and BLIEs of the scale and location parameters together with the population mean using RSS. Al-Omari et al. [19] investigated the ratio estimation using a multi-stage median RSS approach. Al-Omari [20] proposed robust extreme RSS mean for mean estimation. Haq [21] proposed Shewhart control chart for monitoring process mean based on partially ordered judgment subset sampling. Haq et al. [22] suggested unbiased estimators for the basic linear regression model based on double RSS. Yu et al. [23] for investigating regression estimator in RSS. Al-Naseer et al. [24] proposed robust extreme RSS. Haq et al. [25] suggested some ratio estimators for the population mean in ERSS using two ancillary variables. Ozturk [26] studied sampling based on partially rank-ordered sets. Haq et al. [22] proposed the hybrid RSS method. Zamanzade et al. [27] introduced a new RSS estimation method for the population mean and variance. Haq [28] considered cluster sampling with hybrid RSS. Haq [29] studied the distribution function estimation under hybrid RSS. Al-Omari et al. [30,31] dealt with tests based on Laplace and logistic distributions. Al-Nasser et al. [32] studied information-theoretic weighted mean based on truncated RSS. Zamanzade et al. [33] used population proportion estimation in pair RSS. Haq [34] studied ordered partially subset sampling and consider the applications of this method to parametric inference. Al-Omari et al. [35] suggested a new RSS procedure called Truncation Based RSS (TBRSS), and showed that their estimator is unbaised of the population mean of symmetric distributions. Al-Nasser et al. [36] suggested minimax RSS method. Haq et al. [37] proposed the Hybrid ranked set sampling scheme. Wang et al. [38] investigated general ranked set sampling with cost consideration. Muttlak [39] introduced median ranked set sampling with concomitant variables and a comparison with ranked set sampling and regression estimators. For applications and new techniques based on RSS, we refer the readers to the references [40–44].

In this paper, we extended the work in this area and proposed two new improved RSS designs called the Simple Z Ranked Set Sampling (SZRSS) and the Generalized Z Ranked Set Sampling (GZRSS) methods. For some cases, SZRSS becomes a particular case of GZRSS design. The proposed sampling procedure estimator is unbiased of the population mean for symmetric distributions. It is shown, theoretically and numerically, that under perfect and imperfect rankings for symmetric non-uniform distributions, the proposed mean estimators under the GZRSS design are more efficient than those obtained by RSS and TBRSS. For asymmetric distributions, the proposed estimators based on GZRSS are more precise as compared to the estimators based on RSS and TBRSS. Also, we extended our sampling designs for estimating the population median. The efficiency of the suggested median estimators under GZRSS is better than that based on the RSS and TBRSS estimators, for symmetric non-uniform and asymmetric distributions. The GZRSS estimator of the population mean is investigated based on perfect and imperfect rankings, and is also compared to the SRS linear regression mean estimator. It is noteworthy that for small to moderate correlation between the auxiliary and study variables, the proposed estimators are more efficient than the SRS linear regression estimator of the population mean.

This paper is organized as follows. Some sampling methods are presented in Section 2. The proposed ZRSS designs are described in detail in Section 3. The problem of errors in ranking and a comparison with the SRS linear regression estimator is discussed in Section 4. The problem of estimating the population median is considered in Section 5. A detailed application to real data is given in Section 6, and finally, the paper is concluded in Section 7.

2 Sampling Methods

In this section, we explain some existing sampling schemes considered in this study.

2.1 Ranked Set Sampling

We describe the RSS design as follows:

Step 1: Given the value of sample size, say m, identify m^2 units from the corresponding population.

Step 2: These units are randomly allocated to m sets such that the size of each set is m.

Step 3: Now, rank the units within each set, this ranking can be done visually or by an inexpensive method with respect to the study variable. Then select the smallest ranked unit from the first set of m units. Similarly, select the second smallest ranked unit from the second set of m units. The procedure continues until the largest ranked unit is selected from the last set. This completes a cycle of a ranked set sample of size m.

Step 4: For a large sample size, say n, the above steps are repeated r times until size of the sample becomes $n = mr$, for $r \geq 1$.

Let Z be the variable of interest with a distribution function (cdf) $F(z)$ and a probability density function (pdf) $f(z)$. Suppose that Z has a mean μ and a variance σ^2. Let $Z_1, Z_2, \ldots, Z_m$ be a SRS of size m drawn from the pdf $f(y)$, i.e., $Z_i \sim f(z)$, for $i = 1, 2, \ldots, m$. Then, the mean of SRS is denoted by $\hat{\mu}_{\text{SRS}} = \frac{1}{m}\sum_{i=1}^{m} Z_i$. Here, $E\left(\hat{\mu}_{\text{SRS}}\right) = \mu$ (an unbiased estimator of μ), and variance is $\operatorname{Var}\left(\hat{\mu}_{\text{SRS}}\right) = \frac{\sigma^2}{m}$. Suppose that Z_{ij}, $i,j = 1, 2, 3, \ldots, m$ be m independent SRS each of size m. Let $Z_{i(1:m)}, Z_{i(2:m)}, \ldots, Z_{i(m:m)}$ denotes the order statistics of the ith sample $Z_{i1}, Z_{i2}, \ldots, Z_{im}$. Now, implement the RSS method to m selected samples. This gives a balanced RSS of size m, $Z_{1(1:m)}, Z_{2(2:m)}, \ldots, Z_{m(m:m)}$. The RSS mean estimator is denoted by $\hat{\mu}_{\text{RSS}} = \frac{1}{m}\sum_{i=1}^{m} Z_{i(i:m)}$. Assuming that $g_{(i:m)}(z)$ be the pdf of the ith order statistic $Z_{(i:m)}$, and noting that for each i, $Z_{i(i:m)} \overset{d}{=} Z_{(i:m)}$, where d stands for equality in distribution. The pdf of the $Z_{(i:m)}$ is given by $g_{(i:m)}(z) = m\binom{m-1}{i-1}F^{i-1}(z)\{1-F(z)\}^{m-i}f(z)$, $-\infty < z < \infty$, with mean $\mu_{(i:m)} = \int_{-\infty}^{\infty} z g_{(i:m)}(z)\,\mathrm{d}z$ and variance $\sigma^2_{(i:m)} = \int_{-\infty}^{\infty}\left(z-\mu_{(i:m)}\right)^2 g_{(i:m)}(z)\,\mathrm{d}z$. It is of interest to note that

$\hat{\mu}_{\text{RSS}}$ is unbiased estimator of μ and the corresponding variance is $\text{Var}\left(\hat{\mu}_{\text{RSS}}\right) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma_{(i:m)}^2 = \frac{\sigma^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2$. Takahasi et al. [2] introduced the foundation of the RSS design and proved that $f(z) = \frac{1}{m}\sum_{i=1}^{m} g_{(i:m)}(z)$ and $\mu = \frac{1}{m}\sum_{i=1}^{m}\mu_{(i:m)}$. The efficiency (Eff) of $\hat{\mu}_{\text{SRS}}$ and $\hat{\mu}_{\text{RSS}}$ is $1 \leq Eff\left(\hat{\mu}_{\text{RSS}}, \hat{\mu}_{\text{SRS}}\right) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{Var}(\hat{\mu}_{\text{RSS}})} \leq \frac{m+1}{2}$. For further details See Takahasi et al. [2].

2.2 Truncation Based RSS

As we mentioned in the introduction that Al-Omari et al. [35] derived the TBRSS design; its describtion is as follows:

Step 1: Choose m by SRS of size m each from the parent population.

Step 2: Within each choden sample, rank the units visually based on the variable of interest or by any inexpensive method.

Step 3: Define a coefficient $\delta = [\alpha m]$, for $0 \leq \alpha < 0.5$. Note that $[t]$ denotes the integer part of t.

Step 4: Choose the minumum ranked unit from the first δ samples and the maximum ranked unit from the last δ samples. From the remaining $m-2\delta$, choose the ith ranked unit from the ith sample for $i = \delta+1, \ldots, m-\delta$.

Step 5: This finalizes a cycle of a TBRSS. Steps 1–4 are repeated r times if needed to determine a sample of size $n = mr$.

The corresponding estimator of population mean based on TBRSS is $\hat{\mu}_{\text{TBRSS}} = \frac{1}{m}\left(\sum_{i=1}^{\delta} z_{i(1:m)} + \sum_{i=\delta+1}^{m-\delta} z_{i(i:m)} + \sum_{i=m-\delta+1}^{m} z_{i(m:m)}\right)$. The estimator $\hat{\mu}_{\text{TBRSS}}$ becomes unbiased if the population is symmetric. For symmetric populations, the variance of $\hat{\mu}_{\text{TBRSS}}$ is $\text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right) = \frac{\delta}{m^2}\left(\sigma_{(1:m)}^2 + \sigma_{(m:m)}^2\right) + \frac{1}{m^2}\sum_{i=\delta+1}^{m-\delta}\sigma_{(i:m)}^2$. Note that for $\delta = 0, 1$, $\text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right) = \text{Var}\left(\hat{\mu}_{\text{RSS}}\right)$. For samples of odd sizes, when $\delta = (m-1)/2$, The TBRSS and ERSS become equivalent. For further details and application of this method, see Al-Omari et al. [35].

3 New Sampling Designs

This section introduces two new RSS methods; namely, simple Z ranked set sampling (SZRSS) and generalized Z ranked set sampling (GZRSS) designs.

3.1 Simple ZRSS Design

The SZRSS procedure for both even and odd samples is described as follows. To get an SZRSS of m size, select m random samples each of size m. Without yet knowing the values in the samples, rank the units within each sample based on any inexpensive or cost free method.

i) For even m, choose the $(i+1)$th smallest ranked unit from the first $m/2$ samples, for $i = 1, \ldots, m/2$. Similarly, choose $(i-1)$th the smallest ranked unit from the last $m/2$ samples, for $i = (m/2)+1, \ldots, m$.

ii) For odd sample size m, choose the $(i+1)$th smallest ranked unit from the first $(m-1)/2$ samples, for $i = 1, \ldots, (m-1)/2$. Then select the median of the $((m+1)/2)$th sample. From the last $(m-1)/2$ samples, select the $(i-1)$th smallest ranked unit, for $i = (m+3)/2, (m+5)/2, \ldots, m$.

This process provides a cycle of an SZRSS of size m The cycle are repeated r times to determine the size $n = mr$. The SZRSS estimator of μ for an even m is

$$\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}} = \frac{1}{m}\left(\sum_{i=1}^{m/2} Z_{i(i+1:m)} + \sum_{i=m/2+1}^{m} Z_{i(i-1:m)}\right). \tag{1}$$

The variance of $\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}$ is $\mathrm{Var}\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) = \frac{1}{m^2}\left(\sum_{i=1}^{m/2}\sigma^2_{(i+1:m)} + \sum_{i=1}^{m/2}\sigma^2_{(i-1:m)}\right)$. For odd sample size m, the estimator based on SZRSS is $\hat{\mu}^{\mathrm{O}}_{\mathrm{SZRSS}} = \frac{1}{m}\left(\sum_{i=1}^{(m-1)/2} Z_{i(i+1:m)} + Z_{\{(m+1)/2\}((m+1)/2:m)} + \sum_{i=(m+3)/2}^{m} Z_{i(i-1:m)}\right)$. The variance of $\hat{\mu}^{\mathrm{O}}_{\mathrm{SZRSS}}$ is given by $\mathrm{Var}\left(\hat{\mu}^{\mathrm{O}}_{\mathrm{SZRSS}}\right) = \frac{1}{m^2}\left(\sum_{i=1}^{(m-1)/2}\sigma^2_{(i+1:m)} + \sigma^2_{((m+1)/2:m)} + \sum_{i=(m+3)/2}^{m}\sigma^2_{(i-1:m)}\right)$.

Lemma 1: (i) For symmetric distributions, the estimator $\hat{\mu}^{\mathrm{J}}_{\mathrm{SZRSS}}$ (J = E or O) of the population mean μ is unbiased. (ii) $\mathrm{Var}(\hat{\mu}^{\mathrm{J}}_{\mathrm{SZRSS}}) < \mathrm{Var}(\hat{\mu}_{\mathrm{RSS}})$ for symmetric (non-uniform) distributions.

Proof:

(i) For the estimator, given in Eq. (1), we have $E\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) = \frac{1}{m}\left(\sum_{i=1}^{m/2} E\left(Z_{i(i+1:m)}\right) + \sum_{i=m/2+1}^{m} E\left(Z_{i(i-1:m)}\right)\right) = \frac{1}{m}\left(\sum_{i=1}^{m/2}\mu_{(i+1:m)} + \sum_{i=m/2+1}^{m}\mu_{(i-1:m)}\right)$. For any symmetric distribution, $\mu_{(i:m)} - \mu = \mu - \mu_{(m-i+1:m)}$, for $i = 1, 2, \ldots, m$. After some simplications, we can write $E\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) = \frac{1}{m}\{2\mu\,(m/2)\} = \mu$. Follow the same process to prove that $E\left(\hat{\mu}^{\mathrm{O}}_{\mathrm{SZRSS}}\right) = \mu$.

(ii) The variance of Eq. (1) is defined as

$$\begin{aligned}\mathrm{Var}\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) &= \frac{1}{m^2}\left(\sum_{i=1}^{m/2}\mathrm{Var}\left(Z_{i(i+1:m)}\right) + \sum_{i=m/2+1}^{m}\mathrm{Var}\left(Z_{i(i-1:m)}\right)\right)\\ &= \frac{1}{m^2}\left(\sum_{i=1}^{m/2}\sigma^2_{(i+1:m)} + \sum_{i=m/2+1}^{m}\sigma^2_{(i-1:m)}\right).\end{aligned}$$

For any symmetric distribution, $\sigma^2_{(i:m)} = \sigma^2_{(m-i+1:m)}$, $i = 1, 2, \ldots, m$. With some algebraic operations, we can write $\mathrm{Var}\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2_{(i:m)} - \frac{2}{m^2}\left(\sigma^2_{(1:m)} - \sigma^2_{(m/2:m)}\right)$. Note that the variance decreases as i increases for symmetric (non-uniform) distributions, with minimum value occuring at $i = [(m+1)/2]$, i.e., $\sigma^2_{(i:m)} \le \sigma^2_{(j:m)}$ for $i, j = 1, 2, \ldots, [(m+1)/2]$ with $i \ge j$. Here, $[t]$ represents the greatest integer value of t. Therefore, $\sigma^2_{(1:m)} > \sigma^2_{(m/2:m)}$ and hence, $\mathrm{Var}\left(\hat{\mu}^{\mathrm{E}}_{\mathrm{SZRSS}}\right) < \mathrm{Var}\left(\hat{\mu}_{\mathrm{RSS}}\right)$, which completes the proof. Follow the same process to prove that $\mathrm{Var}\left(\hat{\mu}^{\mathrm{O}}_{\mathrm{SZRSS}}\right) < \mathrm{Var}\left(\hat{\mu}_{\mathrm{RSS}}\right)$.

3.2 Generalized ZRSS Design

Now, we propose a generalized ZRSS (GZRSS) design. The steps of selecting a GZRSS are given below in which the Steps 1–3 are similar to the TBRSS method.

Step 1: Choose m simple random samples, each with size m selected from the corresponding population.

Step 2: Within each sample, rank the units visually with respect to the variable of interest or by inexpensive or cost free method.

Step 3: Define a coefficient $\delta = [\alpha m]$, for $0 \le \alpha < 0.5$ and $[t]$ symbolizes the integer value of t.

Step 4: From the first δ samples, draw the $(i+1)$th smallest ranked unit. From the last δ samples, draw the $(i-1)$th smallest ranked unit. But from the remaining $m-2\delta$ samples, draw the ith ranked unit from the ith sample for $i = \delta+1, \ldots, m-\delta$.

Step 5: Previous steps finalize a cycle of a GZRSS of size m. Steps 1–4 are done r times if needed to determine a sample of size $n = mr$.

Let $Z_{11}, Z_{12}, \ldots, Z_{1m}, Z_{21}, Z_{22}, \ldots, Z_{2m}, \ldots, Z_{i1}, Z_{i2}, \ldots, Z_{im}, \ldots, Z_{m1}, Z_{m2}, \ldots, Z_{mm}$ be m independent simple random samples each of size m. The GZRSS estimator of population mean μ based on this sample is

$$\hat{\mu}_{\text{GZRSS}} = \frac{1}{m}\left(\sum_{i=1}^{\delta} Z_{i(i+1:m)} + \sum_{i=\delta+1}^{m-\delta} Z_{i(i:m)} + \sum_{i=m-\delta+1}^{m} Z_{i(i-1:m)}\right), \tag{2}$$

and the corresponding variance is $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = \frac{1}{m^2}\left(\sum_{i=1}^{\delta} \sigma^2_{(i+1:m)} + \sum_{i=\delta+1}^{m-\delta} \sigma^2_{(i:m)} + \sum_{i=m-\delta+1}^{m} \sigma^2_{(i-1:m)}\right)$. Note that for $\delta = 0$, we have $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = \text{Var}\left(\hat{\mu}_{\text{RSS}}\right)$.

Lemma 2: For symmetric distributions about its population mean μ we have

(i) $\hat{\mu}_{\text{GZRSS}}$ is an unbiased estimator of μ.
(ii) $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) \le \text{Var}\left(\hat{\mu}_{\text{RSS}}\right)$ and $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) \le \text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right)$.

Proof:

(i) Take the expectation of $\hat{\mu}_{\text{GZRSS}}$, given in Eq. (2), we have $E\left(\hat{\mu}_{\text{GZRSS}}\right) = \frac{1}{m}\left\{\sum_{i=1}^{\delta} E\left(Z_{i(i+1:m)}\right) + \sum_{i=\delta+1}^{m-\delta} E\left(Z_{i(i:m)}\right) + \sum_{i=m-\delta+1}^{m} E\left(Z_{i(i-1:m)}\right)\right\} = \frac{1}{m}\left\{\sum_{i=1}^{\delta} \mu_{(i+1:m)} + \sum_{i=\delta+1}^{m-\delta} \mu_{(i:m)} + \sum_{i=m-\delta+1}^{m} \mu_{(i-1:m)}\right\}$. As $\mu_{(i:m)} - \mu = \mu - \mu_{(m-i+1:m)}$, for $i = 1, 2, \ldots, m$. Therefore, we can write $E\left(\hat{\mu}_{\text{GZRSS}}\right) = \frac{1}{m}\{2\delta\mu + m\mu - 2\delta\mu\} = \mu$, which completes the proof.

(ii) Consider the estimator, given in Eq. (2), we have

$$\begin{aligned}\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) &= \frac{1}{m^2}\left\{\sum_{i=1}^{\delta} \text{Var}\left(Z_{i(i+1:m)}\right) + \sum_{i=\delta+1}^{m-\delta} \text{Var}\left(Z_{i(i:m)}\right) + \sum_{i=m-\delta+1}^{m} \text{Var}\left(Z_{i(i-1:m)}\right)\right\} \\ &= \frac{1}{m^2}\left\{\sum_{i=1}^{\delta} \sigma^2_{(i+1:m)} + \sum_{i=\delta+1}^{m-\delta} \sigma^2_{(i:m)} + \sum_{i=m-\delta+1}^{m} \sigma^2_{(i-1:m)}\right\}.\end{aligned}$$

For any symmetric distribution, $\sigma^2_{(i:m)} = \sigma^2_{(m-i+1:m)}$, $i = 1, 2, \ldots, m$. After some simplication, we can write $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2_{(i:m)} - \frac{2}{m^2}\left(\sigma^2_{(1:m)} - \sigma^2_{(k+1:m)}\right)$. As explain above, for symmetric (non-uniform) distributions, $\sigma^2_{(i:m)} \le \sigma^2_{(j:m)}$ for $i, j = 1, 2, \ldots, [(m+1)/2]$ with $i \ge j$. Therefore, $\sigma^2_{(1:m)} > \sigma^2_{(k+1:m)}$ and hence, $\text{Var}\left(\hat{\mu}^{\text{E}}_{\text{SZRSS}}\right) < \text{Var}\left(\hat{\mu}_{\text{RSS}}\right)$. The equality is attained when $k = 0$, which completes the proof.

(iii) $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) \le \text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right)$ if and only if $\text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right) - \text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) \ge 0$. This implies that $\frac{1}{m^2}\left\{\delta\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right) + \sum_{i=\delta+1}^{m-\delta}\sigma^2_{(i:m)}\right\} - \frac{1}{m^2}\left\{\sum_{i=1}^{\delta}\sigma^2_{(i+1:m)} + \sum_{i=\delta+1}^{m-\delta}\sigma^2_{(i:m)} + \sum_{i=m-\delta+1}^{m}\sigma^2_{(i-1:m)}\right\} \ge 0$ and also we have $\delta\left(\sigma^2_{(1:m)} + \sigma^2_{(m:m)}\right) - \sum_{i=1}^{\delta}\sigma^2_{(i+1:m)} - \sum_{i=m-\delta+1}^{m}\sigma^2_{(i-1:m)} \ge 0$. It can be written as $\left(\sigma^2_{(1:m)} - \sigma^2_{(2:m)}\right) + \left(\sigma^2_{(1:m)} - \sigma^2_{(3:m)}\right) + \ldots + \left(\sigma^2_{(1:m)} - \sigma^2_{(k+1:m)}\right) + \left(\sigma^2_{(m:m)} - \sigma^2_{(m-k:m)}\right) + \left(\sigma^2_{(m:m)} - \sigma^2_{(m-k+1:m)}\right) + \ldots + \left(\sigma^2_{(m:m)} - \sigma^2_{(m-1:m)}\right) \ge 0$.

As mentioned above, for symmetric (non-uniform) distributions, $\sigma^2_{(i:m)} \le \sigma^2_{(j:m)}$ for $i, j = 1, 2, \ldots, [(m+1)/2]$ with $i \ge j$. Therefore, all of the above differences are positive and hence $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) < \text{Var}\left(\hat{\mu}_{\text{TBRSS}}\right)$. The equality is attained when $\delta = 0$, which completes the proof.

In the case of symmetric of the parent distribution, the *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is defined by $Eff\left(\hat{\mu}_{GZRSS}, \hat{\mu}_{SRS}\right) = \frac{Var(\hat{\mu}_{SRS})}{Var(\hat{\mu}_{GZRSS})} = \frac{m\sigma^2}{\sum_{i=1}^{m}\sigma^2_{(i:m)} - 2\left(\sigma^2_{(1:m)} - \sigma^2_{(k+1:m)}\right)}$. For asymmetric populations, the *Eff* will be $Eff\left(\hat{\mu}_{GZRSS}, \hat{\mu}_{SRS}\right) = \frac{Var(\hat{\mu}_{SRS})}{MSE(\hat{\mu}_{GZRSS})} = \frac{m\sigma^2}{\sum_{i=1}^{\delta}\sigma^2_{(i+1:m)} + \sum_{i=\delta+1}^{m-\delta}\sigma^2_{(i:m)} + \sum_{i=m-\delta+1}^{m}\sigma^2_{(i-1:m)}}$. Now, to illustrate the method, some choices of the sample size m and the coefficient δ are considered for normal and Weibull distributions.

3.3 Examples

3.3.1 Normal Distribution

Let $Z \sim N(0, 1)$, where $f(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)$, $-\infty < z < \infty$. The cdf and pdf of the ith ranked unit from an RSS of size $m = 5$, respectively, are $G_{(i:5)}(z) = \text{BetaRegularized}\left[\frac{1}{2}\text{Erfc}\left(-z/\sqrt{2}\right), i, 5-i\right]$ and $g_{(i:5)}(z) = \frac{i}{8\sqrt{2\pi}}\binom{4}{i}\exp\left(-z^2/2\right)\left\{\text{Erfc}\left(-z/\sqrt{2}\right)\right\}^{i-1}\left\{\text{Erfc}\left(z/\sqrt{2}\right)\right\}^{4-i}$, where $\text{Erfc}(z) = 1 - \text{Erf}(z) = 1 - \frac{2}{\sqrt{\pi}}\int_0^z \exp\left(-w^2\right)\text{d}w$ is the complementary error function, $\text{BetaRegularized}[z, a, b] = \frac{1}{\text{Beta}(a,b)}\int_0^z t^{a-1}(1-t)^{b-1}\text{d}t$ is the regularized incomplete beta function and $\text{Beta}(a, b)$ is the beta function or Euler integral of the first kind. Consider the estimator given in Eq. (2), and let $m = 5$, we have $\hat{\mu}_{\text{GZRSS}} = \frac{1}{5}\left(\sum_{i=1}^{\delta} Z_{i(i+1:5)} + \sum_{i=\delta+1}^{5-\delta} Z_{i(i:5)} + \sum_{i=5-\delta+1}^{5} Z_{i(i-1:5)}\right)$. Based on the order statistics, the means and variances of the ith, for $i = 1, 2, \ldots, 5$, are $\mu_{(1:5)} = -1.1629$, $\mu_{(2:5)} = -0.4950$, $\mu_{(3:5)} = 0$,

$\mu_{(4:5)} = 0.4950$, $\mu_{(5:5)} = 1.1629$, $\sigma^2_{(1:5)} = 0.4475$, $\sigma^2_{(2:5)} = 0.3115$, $\sigma^2_{(3:5)} = 0.2868$, $\sigma^2_{(4:5)} = 0.3115$, and $\sigma^2_{(5:5)} = 0.4475$. Also, the cases below for δ can be treated as follows:

Case I: $\delta = 0$: The computed results for expectation and the variance of the estimator, given in Eq. (2), are, respectively, $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0$, $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0721$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is given by $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{Var}(\hat{\mu}_{\text{GZRSS}})} = 2.7701$.

Case II: $\delta = 1$: The expectation of $\hat{\mu}_{\text{GZRSS}}$ is $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0$ and variance $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0613$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{Var}(\hat{\mu}_{\text{GZRSS}})} = 3.2617$.

Case III: $\delta = 2$: The expectation of the estimator given in Eq. (2) is $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0$, with variance $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0593$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = 3.3703$.

3.3.2 Weibull Distribution

Let $Z \sim \text{Weibull}(2,1)$, having pdf $f(z) = 2z\exp\left(-z^2\right)$, $z > 0$. The cdf and pdf of the ith ranked unit from a ranked set sample for $m = 5$, respectively, are $G_{(i:5)}(z) = \text{BetaRegularized}\left[1 - \exp\left(-z^2\right), i, 6 - i\right]$ and $g_{(i:5)}(z) = 2i\binom{5}{i} z\exp\left[-z^2(6-i)\right]\left[1 - \exp\left(-z^2\right)\right]^{i-1}$.

For $m = 5$, we have $\hat{\mu}_{\text{GZRSS}} = \frac{1}{5}\left(\sum_{i=1}^{\delta} Z_{i(i+1:5)} + \sum_{i=\delta+1}^{5-\delta} Z_{i(i:5)} + \sum_{i=5-\delta+1}^{5} Z_{i(i-1:5)}\right)$. Based on the order statistics, the means and variances are $\mu_{(1:5)} = 0.3963$, $\mu_{(2:5)} = 0.6302$, $\mu_{(3:5)} = 0.8479$, $\mu_{(4:5)} = 1.0946$, $\mu_{(5:5)} = 1.4619$, and $\sigma^2_{(1:5)} = 0.0429$, $\sigma^2_{(2:5)} = 0.0528$, $\sigma^2_{(3:5)} = 0.0643$ $\sigma^2_{(4:5)} = 0.0850$, $\sigma^2_{(5:5)} = 0.1459$, respectively. Also, the cases below for δ can be treated as follows:

Case I: *Consider* $\delta = 0$:

The expectation of $\hat{\mu}_{\text{GZRSS}}$ is $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.8862$, which is an unbiased estimate with variance of $\hat{\mu}_{\text{GZRSS}}$ is $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0156$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = 2.7436$.

Case II: *Consider* $\delta = 1$: Similarly, the mean (expectation) and variance of $\hat{\mu}_{\text{GZRSS}}$ are $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.8595$ and $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0135$, respectively. As the estimator is not unbiased, therefore, the MSE of $\hat{\mu}_{\text{GZRSS}}$ is given by $\text{MSE}\left(\hat{\mu}_{\text{GZRSS}}\right) = \left(\text{Bias}\left(\hat{\mu}_{\text{GZRSS}}\right)\right)^2 + \text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right)$. The bias of $\hat{\mu}_{\text{GZRSS}}$ is $\text{Bias}\left(\hat{\mu}_{\text{GZRSS}}\right) = -0.0266$. Therefore, $\text{MSE}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0143$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{MSE}(\hat{\mu}_{\text{GZRSS}})} = 2.9989$.

Case III: *Consider*$\delta = 2$: The mean and variance of $\hat{\mu}_{\text{GZRSS}}$ are $E\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.8537$ and $\text{Var}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0132$. Again the estimator is biased with $\text{Bias}\left(\hat{\mu}_{\text{GZRSS}}\right) = -0.0324$ and $\text{MSE}\left(\hat{\mu}_{\text{GZRSS}}\right) = 0.0142$. The *Eff* of $\hat{\mu}_{\text{GZRSS}}$ with respect to $\hat{\mu}_{\text{SRS}}$ is $\mathit{Eff}\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{\text{SRS}}\right) = \frac{\text{Var}(\hat{\mu}_{\text{SRS}})}{\text{MSE}(\hat{\mu}_{\text{GZRSS}})} = 3.0034$.

Table 1: Exact *Eff* of mean estimators under symmetric distributions

	$m=4$		$m=7$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$k\delta=3$ GZRSS	$\delta=3$ TBRSS
Distributions	TBRSS	SZRSS	TBRSS				SZRSS	
N (0, 1)	2.3469	2.7743	3.5949	4.1746	4.3674	3.1566	4.4185	2.7323
Laplace (0, 1)	2.0383	3.8408	2.8676	4.8365	5.4700	2.0380	6.0144	1.5033
Logit (0, 1)	2.2164	3.1637	3.2667	4.5109	4.9153	2.5605	5.0195	2.0275
U (0, 1)	2.5000	2.0833	4.0000	3.5745	3.3600	4.5405	3.2941	5.7931
Beta (6, 6)	2.4026	2.6052	3.7412	4.0207	4.1242	3.4980	4.1526	3.2185
T (7)	2.1853	3.2286	3.1857	4.6091	5.0184	2.4339	5.1206	1.9029
	$m=6$		$m=5$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS
Distributions	TBRSS		SZRSS		TBRSS		SZRSS	
N (0, 1)	3.1857	3.7250	3.8860	2.7828	2.7702	3.2618	3.3703	2.4074
Laplace (0, 1)	2.6028	4.5481	5.3831	1.8230	2.3274	4.2355	4.8580	1.6046
Logit (0, 1)	2.9276	4.0931	4.4382	2.2787	2.5783	3.6514	3.8914	1.9927
U (0, 1)	3.5000	3.0625	2.8824	4.0833	3.0000	2.5610	2.4419	3.6207
Beta (6, 6)	3.2993	3.5594	3.6446	3.0747	2.8535	3.0895	3.1456	2.6509
T (7)	2.8646	4.1812	4.5287	2.1786	2.5321	3.7283	3.9694	1.9171

Now, we consider the mean estimation for some symmetric distributions, and also for some asymmetric distributions. The exact relative efficiencies of our proposed estimators are presented in Tabs. 1 and 2.

Tabs. 1 and 2 show that, for symmetric distributions, the efficiency of the GZRSS increases as the δ value increases except in the case of the uniform distribution. In the case of asymmetric distributions, generally, the efficiencies increase when δ increases for 0 to 1, and they decrease function when $\delta > 1$. For both asymmetric and symmetric distributions, the relative efficiency of mean estimators under GZRSS is an increasing function of the sample size. For all considered cases, GZRSS is more efficient than RSS and TBRSS except that TBRSS is more adequate than GZRSS when the considered distribution is standard uniform.

4 Errors in Ranking and Comparison with SRS Regression Estimator

We investigate the fulfillment of the suggested estimators for the mean under both GZRSS design and imperfect rankings. The suggested estimators under both rankings' schemes are also compared with the SRS for the population mean based on the linear regression estimator.

4.1 Errors in Ranking

Accurate ranking increases the efficiency of the RSS. However, Dell et al. [3] show that even if the ranking has some errors, the estimator under RSS still remains unbiased and performs at least as well as the SRS estimator. Here, we study the performance of the estimators under the proposed RSS designs, when ranking has some errors. The mostly used RSS model to study the effect of errors in ranking is based on the ranking with respect to a concomitant variable that is correlated with the study variable. The efficiency of the estimator now depends on the correlation value between the study variable Z and the concomitant or ancillary variable W. Stokes [4] suggested a model for imperfect ranking assuming that an ancillary variable W is available, can be simply measured and is correlated with the interest variable Z. For further details see Stokes [4], Patil et al. [5] and Muttlak [39]. Stokes [4] imposed the following assumptions considered in developing the following model:

(i) The relationship between Z and the regressor W is linear,
(ii) $\frac{Z-\mu_Z}{\sigma_Z}$ and $\frac{W-\mu_W}{\sigma_W}$ variables follow the same distribution.

Table 2: Exact *Eff* comparison of mean estimators under asymmetric distributions

	$m=4$		$m=7$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$k\delta=3$ GZRSS	$\delta=3$ TBRSS
Distributions	TBRSS	SZRSS	TBRSS				SZRSS	
Exp (1)	1.9200	2.4407	2.6997	3.0237	2.5973	1.6434	2.4628	0.8125
Gamma (2, 1)	2.0958	2.5639	3.0515	3.4334	3.1541	2.1030	3.0533	1.1880
Gamma (3, 2)	2.1695	2.6230	3.2055	3.6317	3.4503	2.3504	3.3763	1.4440
Beta (9, 2)	2.2667	2.4835	3.4398	3.5941	3.4134	2.7576	3.3441	1.7990
Weibull (2, 1)	2.3251	2.5676	3.5609	3.8263	3.7826	3.0318	3.7585	2.2718
Half-Nor. (1)	2.2393	2.3701	3.3857	3.4191	3.1403	2.6310	3.0456	1.5988
	$m=6$		$m=5$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS
Distributions	TBRSS		SZRSS		TBRSS		SZRSS	
Exp (1)	2.449	2.8169	2.4948	1.4815	2.1898	2.6201	2.4423	1.3216
Gamma (2, 1)	2.7423	3.1464	2.9529	1.8836	2.4244	2.8576	2.7680	1.6650
Gamma (3, 2)	2.8691	3.3037	3.1890	2.0976	2.5245	2.9696	2.9282	1.8449
Beta (9, 2)	3.0551	3.2287	3.1088	2.4411	2.6645	2.8598	2.8075	2.1275
Weibull (2, 1)	3.1551	3.4167	3.3992	2.6824	2.7436	2.9990	3.0034	2.3327
Half-Nor. (1)	3.0100	3.0702	2.8712	2.3379	2.6284	2.7206	2.6172	2.0482

If (Z, W) follows the bivariate normal distribution, then both conditions are easily satisfied. Following Stokes [4] and (i), we can write $Z_{[i:m]} = \mu_Z + \rho \frac{\sigma_Z}{\sigma_W} \left(W_{(i:m)} - \mu_W \right) + \xi_i$, $i = 1, 2, \ldots, m$, where where ρ is the coefficient of correlation, σ_Z and σ_W are the population standard deviations, μ_Z and μ_W are the corresponding means. Note that the ranking of the auxiliary variable W is perfect whereas the ranking of Z is imperfect, i.e., the ranking of Z has some errors. Here, $W_{(i:m)}$ and $Z_{[i:m]}$ denote the ith order statistic and the ith judgment order statistic of a random sample of size m. ξ_i denotes the error term with zero mean and a constant variance, i.e., $E(\xi_i) = 0$ and $\mathrm{Var}(\xi_i) = \sigma_\xi^2 = \sigma_Z^2 \left(1 - \rho^2\right)$. As SZRSS becomes a special case of GZRSS, therefore, we consider the estimator based on GZRSS. Now, the mean of the study variable Z with ranking based on the auxiliary variable W under GZRSS can be written as $\hat{\mu}_{Z\text{GZRSS}} = \frac{1}{m} \left(\sum_{i=1}^{\delta} Z_{i[i+1:m]} + \sum_{i=\delta+1}^{m-\delta} Z_{i[i:m]} + \sum_{i=m-\delta+1}^{m} Z_{i[i-1:m]} \right)$, where $\hat{\mu}_{Y\text{GZRSS}}$ is unbiased estimaor of μ_Y, and it variance is $\mathrm{Var}\left(\hat{\mu}_{\mathrm{ZGZRSS}}\right)=\frac{1}{m^2}\left\{m\left(1-\rho^2\right)\sigma_Z^2+\rho^2\frac{\sigma_Z^2}{\sigma_W^2}\left(\sum_{i=1}^{\delta}\sigma_{W(i+1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(i-1:m)}^2\right)\right\}$. Note that if we consider $\delta=0$ in $\hat{\mu}_{\mathrm{ZGZRSS}}$, then it becomes the simple RSS estimator of population mean. The efficiency of $\hat{\mu}_{Y\mathrm{GZRSS}}$ with respect to $\hat{\mu}_{\mathrm{SRS}}$ is

$$Eff\left(\hat{\mu}_{\mathrm{ZGZRSS}},\hat{\mu}_{\mathrm{SRS}}\right)=\frac{\mathrm{Var}\left(\hat{\mu}_{\mathrm{SRS}}\right)}{\mathrm{Var}\left(\hat{\mu}_{\mathrm{ZGZRSS}}\right)}$$

$$=\frac{1}{1-\rho^2+\frac{\rho^2}{m\sigma_W^2}\left(\sum_{i=1}^{\delta}\sigma_{W(i+1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(i-1:m)}^2\right)}.$$

The efficiency of the $\hat{\mu}_{\mathrm{ZGZRSS}}$ with respect to $\hat{\mu}_{\mathrm{RSS}}$ is

$$Eff\left(\hat{\mu}_{\mathrm{ZGZRSS}},\hat{\mu}_{\mathrm{RSS}}\right)=\frac{\mathrm{Var}\left(\hat{\mu}_{\mathrm{RSS}}\right)}{\mathrm{Var}\left(\hat{\mu}_{\mathrm{ZGZRSS}}\right)}$$

$$=\frac{\sum_{i=1}^{m}\sigma_{W(i:m)}^2}{\sum_{i=1}^{\delta}\sigma_{W(i+1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(i-1:m)}^2},$$

which shows that $\hat{\mu}_{Y\mathrm{GZRSS}}$ is always better over $\hat{\mu}_{Y\mathrm{SRS}}$ and $\hat{\mu}_{\mathrm{RSS}}$, even when there are errors in ranking. Similarly, we can define the estimator of population mean μ_Z based on TBRSS by $\hat{\mu}_{\mathrm{ZTBRSS}}=\frac{1}{m}\left(\sum_{i=1}^{\delta}Z_{i[1:m]}+\sum_{i=k+1}^{m-\delta}Z_{i[i:m]}+\sum_{i=m-\delta+1}^{m}Z_{i[m:m]}\right)$, where $\hat{\mu}_{Z\mathrm{TBRSS}}$ is unbiased estimaor of μ_Z, and it variance is given by $\mathrm{Var}\left(\hat{\mu}_{Z\mathrm{TBRSS}}\right)=\frac{1}{m^2}\left[m\left(1-\rho^2\right)\sigma_Z^2+\rho^2\frac{\sigma_Z^2}{\sigma_W^2}\left(\sum_{i=1}^{\delta}\sigma_{W(1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(m:m)}^2\right)\right]$.

The efficiency of $\hat{\mu}_{\mathrm{ZGZRSS}}$ with respect to $\hat{\mu}_{Z\mathrm{TBRSS}}$ (based on imperfect ranking) is given by

$$Eff\left(\hat{\mu}_{\mathrm{ZGZRSS}},\hat{\mu}_{Z\mathrm{TBRSS}}\right)=\frac{\mathrm{Var}\left(\hat{\mu}_{Z\mathrm{TBRSS}}\right)}{\mathrm{Var}\left(\hat{\mu}_{\mathrm{ZGZRSS}}\right)}$$

$$=\frac{1+\frac{\rho^2}{m(1-\rho^2)\sigma_W^2}\left(\sum_{i=1}^{\delta}\sigma_{W(1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(m:m)}^2\right)}{1+\frac{\rho^2}{m(1-\rho^2)\sigma_W^2}\left(\sum_{i=1}^{\delta}\sigma_{W(i+1:m)}^2+\sum_{i=\delta+1}^{m-\delta}\sigma_{W(i:m)}^2+\sum_{i=m-\delta+1}^{m}\sigma_{W(i-1:m)}^2\right)}.$$

The exact efficiencies of our proposed estimators under GZRSS with respect to RSS and TBRSS are given in Tab. 3.

It is clear from the results given in Tab. 3 that, as the efficiencies under each design are a function of the correlation coefficient ρ, i.e., as the value of ρ increases, the relative efficiencies increase and vice versa. As expected, the increase in the sample size also increases the efficiency of the estimator under each of the RSS design. The proposed estimators are better than the existing counterparts.

Table 3: Exact *Eff* comparison of mean estimators under standard bivariate normal distribution

			ρ								
Design	m	δ	0.1	0.2	0.3	0.4	0.5	0.7	0.8	0.9	1
RSS	4	0	1.0057	1.0235	1.0544	1.1011	1.1675	1.3912	1.5805	1.8687	2.3469
GZRSS	4	1	1.0064	1.0262	1.0610	1.1139	1.1903	1.4564	1.6929	2.0748	2.7742
RSS	5	0	1.0064	1.0262	1.0610	1.1138	1.1901	1.4558	1.6919	2.0729	2.7701
GZRSS	5	1	1.0069	1.0285	1.0665	1.1247	1.2097	1.5146	1.7978	2.2813	3.2617
GZRSS	5	2	1.0070	1.0289	1.0675	1.1267	1.2133	1.5258	1.8185	2.3237	3.3703
TBRSS	5	2	1.0058	1.0239	1.0555	1.1031	1.1711	1.4014	1.5978	1.8994	2.4073
RSS	6	0	1.0069	1.0282	1.0658	1.1233	1.2070	1.5064	1.7828	2.2509	3.1856
GZRSS	6	1	1.0073	1.0301	1.0704	1.1325	1.2238	1.5587	1.8803	2.4542	3.7250
GZRSS	6	2	1.0074	1.0306	1.0716	1.1348	1.2280	1.5721	1.9058	2.5097	3.8860
TBRSS	6	2	1.0064	1.0263	1.0611	1.1142	1.1907	1.4575	1.6949	2.0786	2.7827
RSS	7	0	1.0072	1.0297	1.0694	1.1305	1.2201	1.5472	1.8586	2.4077	3.5949
GZRSS	7	1	1.0076	1.0313	1.0734	1.1385	1.2347	1.5939	1.9481	2.6039	4.1745
GZRSS	7	2	1.0077	1.0318	1.0745	1.1407	1.2387	1.6072	1.9741	2.6633	4.3673
TBRSS	7	2	1.0068	1.0281	1.0655	1.1227	1.2059	1.5032	1.7769	2.2391	3.1566
GZRSS	7	3	1.0078	1.0319	1.0748	1.1412	1.2398	1.6105	1.9808	2.6786	4.4185
TBRSS	7	3	1.0063	1.0260	1.0605	1.1128	1.1883	1.4506	1.6828	2.0556	2.7322

4.2 Comparison with Regression Estimator Based on SRS

Patil et al. [5] compared the estimator of a population mean under RSS with the regression estimator based on SRS. It is shown that for a small correlation between the study variable and the ancillary variable, the RSS mean estimator is better than the regression estimator under SRS. In this section, we compare the performance of the proposed mean estimator under GZRSS with respect to the SRS regression estimator. It is assumed that the population mean of the ancillary variable is known. Following Muttlak [39], the linear regression of Z on W is $Z_i = \alpha + \beta W_i + \xi_i$, $i = 1, 2, \ldots, m$, where α and β are the intercept and slope of the regression line. Here, ξ_i is error term with zero mean. The linear regression estimator of the population mean μ_Z when μ_W is known is

$$\hat{\mu}_{Zlr} = \overline{Z} + \hat{\beta}\left(\mu_W - \overline{W}\right), \tag{3}$$

where $\overline{Z}$ and $\overline{W}$ are the corresponding sample mean of Z and W, based on an SRS of size m. Note that, $\hat{\beta}$ is the least square estimator of the slope β of the regression line. Sukhatme and Sukhatme (1970) showed that the regression estimator given in Eq. (3) of the population mean μ_Z is an unbiased estimator once the joint distribution of Z and W is a bivariate normal distribution. The variance of $\hat{\mu}_{Ylr}$ is given by $\text{Var}\left(\hat{\mu}_{Zlr}\right) = \frac{\sigma_Z^2(1-\rho^2)}{m}\left(1 + \frac{1}{m-3}\right)$. In case of perfect ranking, the *Eff* of $\hat{\mu}_{\text{GZRSS}}$ relative to $\hat{\mu}_{Zlr}$ is given by $Eff\left(\hat{\mu}_{\text{GZRSS}}, \hat{\mu}_{Zlr}\right) = \frac{\text{Var}(\hat{\mu}_{Zlr})}{\text{Var}(\hat{\mu}_{\text{GZRSS}})} = \frac{m\sigma_Z^2(1-\rho^2)\left(1+\frac{1}{m-3}\right)}{\sum_{i=1}^{\delta}\sigma_{(i+1:m)}^2 + \sum_{i=\delta+1}^{m-\delta}\sigma_{(i:m)}^2 + \sum_{i=m-\delta+1}^{m}\sigma_{(i-1:m)}^2}$. Similarly, in case of imperfect ranking, the *Eff* of $\hat{\mu}_{Z\text{GZRSS}}$

relative to $\hat{\mu}_{Ylr}$ is given by

$$Eff\left(\hat{\mu}_{ZGZRSS}, \hat{\mu}_{Zlr}\right) = \frac{\text{Var}\left(\hat{\mu}_{Zlr}\right)}{\text{Var}\left(\hat{\mu}_{ZGZRSS}\right)}$$

$$= \frac{1+\frac{1}{m-3}}{1+\frac{\rho^2}{m\sigma_W^2(1-\rho^2)}\left(\sum_{i=1}^{\delta}\sigma^2_{W(1:m)}+\sum_{i=\delta+1}^{m-\delta}\sigma^2_{W(i:m)}+\sum_{i=m-\delta+1}^{m}\sigma^2_{W(m:m)}\right)}.$$

In Tab. 4, we provide exact relative efficiencies of the proposed estimators with respect to the classical linear regression estimator of mean. Note that the proposed mean estimator with perfect ranking under GZRSS outperforms other competitor estimators when the value of ρ is less than 0.9.

Similarly, in Tab. 5, we compared the performance of the suggested estimators under imperfect ranking with respect to the linear regression estimator. It is worth mentioning that even when there are errors in ranking, the proposed estimator is still more efficient than the linear regression estimator when the value of ρ is less than 0.8. The efficiencies of the newly estimators are high based on perfect ranking as compared with the case of imperfect ranking. Note that here RE is a decreasing function of sample size because the performance of linear regression estimator is increasing with the increasing of the sample size. For all of the cases, GZRSS mean estimator always performs better than the TBRSS estimator.

Table 4: The *Eff* of the SRS linear regression estimator with respect to the GZRSS estimator based on perfect ranking

ρ											
Design	m	δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
RSS	4	0	4.6469	4.5061	4.2714	3.9428	3.5204	3.0040	2.3938	1.6898	0.8918
GZRSS	4	1	5.4930	5.3266	5.0491	4.6607	4.1614	3.5510	2.8297	1.9974	1.0542
RSS	5	0	4.1137	3.9890	3.7812	3.4904	3.1164	2.6593	2.1191	1.4958	0.7895
GZRSS	5	1	4.8437	4.6969	4.4523	4.1098	3.6694	3.1313	2.4952	1.7613	0.9296
GZRSS	5	2	5.0049	4.8532	4.6004	4.2466	3.7916	3.2355	2.5782	1.8199	0.9605
TBRSS	5	2	3.5749	3.4665	3.2860	3.0332	2.7082	2.3110	1.8416	1.2999	0.6861
RSS	6	0	4.2050	4.0776	3.8652	3.5679	3.1856	2.7184	2.1662	1.5291	0.8070
GZRSS	6	1	4.9170	4.7680	4.5196	4.1720	3.7250	3.1786	2.5330	1.7880	0.9436
GZRSS	6	2	5.1295	4.9740	4.7150	4.3523	3.8860	3.3160	2.6424	1.8652	0.9844
TBRSS	6	2	3.6732	3.5619	3.3764	3.1166	2.7827	2.3746	1.8922	1.3357	0.7049
RSS	7	0	4.4487	4.3139	4.0892	3.7746	3.3702	2.8759	2.2917	1.6177	0.8537
GZRSS	7	1	5.1660	5.0094	4.7485	4.3832	3.9136	3.3396	2.6612	1.8785	0.9914
GZRSS	7	2	5.4046	5.2408	4.9678	4.5857	4.0944	3.4938	2.7841	1.9653	1.0372
TBRSS	7	2	3.9063	3.7879	3.5906	3.3144	2.9593	2.5253	2.0123	1.4204	0.7497
GZRSS	7	3	5.4678	5.3022	5.0260	4.6394	4.1423	3.5348	2.8167	1.9883	1.0493
TBRSS	7	3	3.3812	3.2787	3.1079	2.8689	2.5615	2.1858	1.7418	1.2295	0.6489

Table 5: The *Eff* of the SRS linear regression estimator with respect to the GZRSS estimator based on imperfect ranking

			ρ								
Design	m	δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
RSS	4	0	1.9914	1.9651	1.9191	1.8498	1.7512	1.6133	1.4190	1.1379	0.7101
GZRSS	4	1	1.9927	1.9704	1.9311	1.8715	1.7854	1.6628	1.4855	1.2189	0.7884
RSS	5	0	1.4945	1.4777	1.4482	1.4035	1.3388	1.2468	1.1137	0.9136	0.5908
GZRSS	5	1	1.4953	1.4810	1.4558	1.4172	1.3609	1.2793	1.1587	0.9708	0.6501
GZRSS	5	2	1.4955	1.4816	1.4572	1.4197	1.3650	1.2854	1.1672	0.9820	0.6622
TBRSS	5	2	1.4937	1.4744	1.4408	1.3900	1.3175	1.2159	1.0721	0.8628	0.5413
RSS	6	0	1.3291	1.3161	1.2931	1.2581	1.2070	1.1332	1.0243	0.8557	0.5702
GZRSS	6	1	1.3297	1.3185	1.2988	1.2684	1.2238	1.1584	1.0599	0.9025	0.6217
GZRSS	6	2	1.3298	1.3191	1.3002	1.2710	1.2280	1.1647	1.0690	0.9148	0.6358
TBRSS	6	2	1.3285	1.3136	1.2875	1.2479	1.1907	1.1091	0.9911	0.8135	0.5266
RSS	7	0	1.2465	1.2356	1.2165	1.1871	1.1439	1.0808	0.9863	0.8363	0.5718
GZRSS	7	1	1.2469	1.2376	1.2210	1.1954	1.1575	1.1015	1.0161	0.8766	0.6184
GZRSS	7	2	1.2471	1.2381	1.2223	1.1977	1.1613	1.1073	1.0246	0.8883	0.6325
TBRSS	7	2	1.2460	1.2337	1.2120	1.1788	1.1306	1.0609	0.9583	0.7996	0.5317
GZRSS	7	3	1.2471	1.2383	1.2226	1.1983	1.1623	1.1088	1.0267	0.8913	0.6361
TBRSS	7	3	1.2454	1.2312	1.2063	1.1685	1.1140	1.0365	0.9248	0.7572	0.4882

5 Estimation of Population Median

Estimation of the population median based on the sampling methods, studied in this paper, is presented in this section. Let Q be the population median and $Z_1, Z_2, \ldots, Z_m$ be an SRS of size m. Then, the median estimator is given by $\hat{Q}_{\text{SRS}} = Z_{\left(\frac{m+1}{2}:m\right)}$, for odd m, and $\hat{Q}_{\text{SRS}} = \frac{1}{2}\left(Z_{\left(\frac{m}{2}:m\right)} + Z_{\left(\frac{m+2}{2}:m\right)}\right)$, for even m. From RSS units of size m, i.e., $Z_{1(1:m)}, Z_{2(2:m)}, \ldots,$ $Z_{m(m:m)}$, the population median estimator based on the RSS is $\hat{Q}_{\text{RSS}} = \mathit{median}\left\{Z_{1(1:m)}, Z_{2(2:m)}, \ldots, Z_{m(m:m)}\right\}$. The corresponding population median estimator based on the GZRSS is $\hat{Q}_{\text{GZRSS}} = \text{median}\left\{Z_{1(2:m)}, \ldots, Z_{\delta(\delta+1:m)}, Z_{\delta+1(\delta+1:m)}, \ldots, Z_{m-\delta(m-\delta:m)}, Z_{m-\delta+1(m-\delta:m)}, \ldots, Z_{m(m-1:m)}\right\}$. The efficiencies of $\hat{Q}_{\text{GZRSS}}$ and $\hat{Q}_{\text{RSS}}$ with respect to $\hat{Q}_{\text{SRS}}$, are given by $\mathit{Eff}\left(\hat{Q}_h, \hat{Q}_{SRS}\right) = \frac{MSE\left(\hat{Q}_{SRS}\right)}{MSE\left(\hat{Q}_h\right)}$, where $h = $ GZRSS, RSS. The estimated MSE of any median estimator is defined as $\text{MSE}\left(\hat{Q}_h\right) = \frac{1}{10^6}\sum_{i=1}^{10^6}\left(\hat{Q}_{ih} - \hat{Q}\right)^2$, $h =$ GZRSS, RSS, SRS. The median estimation of some symmetric and asymmetric distributions is considered here based on extensive Monte Carlo simulations. The obtained results are presented in Tabs. 6–8.

Table 6: *Eff* comparison of median estimators under symmetric distributions

	$m=4$		$m=7$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$k\delta=3$ GZRSS	$\delta=3$ TBRSS
Distributions	TBRSS	SZRSS	TBRSS				SZRSS	
Exp (1)	2.2047	2.7001	2.5002	2.8383	3.3952	2.1690	3.6972	1.3014
Gamma (2, 1)	2.5013	3.5738	3.0637	3.5165	4.3106	2.5335	4.7074	1.4043
Gamma (3, 2)	2.2825	2.8636	2.6048	2.9246	3.5516	2.1990	3.8096	1.3322
Beta (9, 2)	1.9823	2.2368	2.2379	2.5384	2.9668	1.9591	3.2303	1.2409
Weibull (2, 1)	2.1341	2.6238	2.4543	2.7646	3.2906	2.0837	3.6019	1.2979
Half-Nor. (1)	2.2876	2.8669	2.5555	2.9373	3.5187	2.2148	3.8230	1.3252
	$m=6$		$m=5$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS
Distributions	TBRSS		SZRSS		TBRSS		SZRSS	
Exp (1)	2.7410	3.1322	3.5320	2.2490	2.1106	2.7561	3.1107	1.4950
Gamma (2, 1)	3.2110	3.9148	4.5626	2.4605	2.5528	3.4580	4.1257	1.6981
Gamma (3, 2)	2.8315	3.3106	3.7088	2.3122	2.1780	2.8629	3.2754	1.5494
Beta (9, 2)	2.4445	2.7741	3.0185	2.0924	1.8650	2.3594	2.6441	1.3926
Weibull (2, 1)	2.6832	3.0811	3.4603	2.1935	2.0519	2.6942	3.0366	1.4784
Half-Nor. (1)	2.7820	3.3051	3.7001	2.2896	2.1881	2.8440	3.2737	1.5437

Table 7: *Eff* comparison of median estimators under asymmetric distributions

	$m=4$		$m=7$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$\delta=3$ GZRSS	$\delta=3$ TBRSS
Distributions	TBRSS	SZRSS	TBRSS				SZRSS	
Exp (1)	2.2343	3.1597	2.6781	3.1437	3.7520	2.3118	4.0702	1.3278
Gamma (2, 1)	2.2526	2.8805	2.5710	2.9520	3.5825	2.2039	3.8493	1.3000
Gamma (3, 2)	2.2182	2.8238	2.5536	2.9132	3.4760	2.1673	3.7753	1.3068
Beta (9, 2)	2.1758	2.7205	2.4799	2.8430	3.3866	2.1622	3.7585	1.3132
Weibull (2, 1)	2.1849	2.6422	2.4935	2.7935	3.3578	2.1300	3.6298	1.2904
Half-Nor. (1)	2.1544	2.6805	2.4791	2.8175	3.3463	2.1317	3.6332	1.3135
	$m=6$		$m=5$					
	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS	$\delta=0$ RSS	$\delta=1$ GZRSS	$\delta=2$ GZRSS	$\delta=2$ TBRSS
Distributions	TBRSS		SZRSS		TBRSS		SZRSS	
Exp(1)	2.9243	3.4698	3.9542	2.3130	2.3296	3.0752	3.5579	1.6061
Gamma (2, 1)	2.7792	3.3336	3.6999	2.2621	2.2133	2.8910	3.3331	1.5371
Gamma (3, 2)	2.7371	3.2919	3.6542	2.2696	2.1561	2.8117	3.2924	1.5337
Beta (9, 2)	2.7039	3.1430	3.5346	2.2543	2.0905	2.7394	3.1401	1.4998
Weibull (2, 1)	2.6740	3.1567	3.4869	2.2274	2.0889	2.7228	3.0921	1.4738
Half-Nor. (1)	2.6672	3.1059	3.4576	2.1861	2.0869	2.6910	3.0727	1.4856

The results, given in Tabs. 6 and 7, reveal that the attainment in efficiency determined by using the GZRSS method. For instance, when $m = 7$ and $\delta = 2$, the RE of the GZRSS is 3.5187 for estimating the median of the student's t distribution. Also, GZRSS is more efficient than RSS and TBRSS based on the same sample size for a fixed value of δ. To study the performance of the proposed median estimators under GZRSS for imperfect rankings, we have considered standard bivariate normal distribution. The relative efficiencies of the median estimators are obtained for different values of correlation coefficient using extensive Monte Carlo simulations and are displayed in Tab. 8.

According to the results given in Tab. 8, the median estimators under proposed designs are at least as efficient as compared with the SRS median estimator. Here, the relative efficiencies are also increasing function of m and ρ. The results under GZRSS are efficient as compared to RSS and TBRSS under perfect and imperfect rankings.

Table 8: *Eff* comparison of median estimators under bivariate normal distribution

ρ											
Design	m	δ	0.1	0.2	0.3	0.4	0.5	0.7	0.8	0.9	1
RSS	4	0	1.0010	1.0166	1.0430	1.0777	1.1343	1.3106	1.4740	1.7317	2.1975
GZRSS	4	1	1.0051	1.0274	1.0598	1.1131	1.1837	1.4396	1.6699	2.0362	2.7102
RSS	5	0	1.0034	1.0175	1.0413	1.0700	1.1162	1.2765	1.4205	1.6569	2.0987
GZRSS	5	1	1.0049	1.0244	1.0586	1.1113	1.1851	1.4411	1.6693	2.0320	2.7185
GZRSS	5	2	1.0105	1.0294	1.0642	1.1193	1.1944	1.4891	1.7597	2.2077	3.1121
TBRSS	5	2	1.0019	1.0102	1.0159	1.0331	1.0482	1.1180	1.1733	1.2791	1.5079
RSS	6	0	1.0039	1.0220	1.0468	1.0909	1.1493	1.3698	1.5728	1.9323	2.7094
GZRSS	6	1	1.0052	1.0313	1.0663	1.1184	1.1977	1.4819	1.7389	2.1988	3.1845
GZRSS	6	2	1.0099	1.0278	1.0703	1.1266	1.2160	1.5304	1.8227	2.3395	3.5345
TBRSS	6	2	1.0001	1.0137	1.0323	1.0656	1.1081	1.2604	1.4043	1.6591	2.2450
RSS	7	0	1.0039	1.0146	1.0469	1.0827	1.1364	1.3231	1.4958	1.8040	2.4819
GZRSS	7	1	1.0046	1.0244	1.0625	1.1103	1.1865	1.4494	1.6797	2.0671	2.8352
GZRSS	7	2	1.0065	1.0286	1.0651	1.1233	1.2071	1.5091	1.7870	2.2875	3.3949
TBRSS	7	2	1.0000	1.0115	1.0216	1.0488	1.0792	1.2038	1.3166	1.5467	2.1409
GZRSS	7	3	1.0066	1.0323	1.0673	1.1259	1.2118	1.5280	1.8295	2.3861	3.6701
TBRSS	7	3	1.0022	1.0018	1.0059	1.0045	1.0091	1.0190	1.0275	1.0800	1.3046

6 An Application to Real Data

To illustrate the use of the GZRSS method in the field, a real data set is considered for both mean and median estimation. This real data set is considered by Platt et al. [45] and it is related to the height and diameter of 399 conifers (Pinus Palustris) trees. The data consists of 7 variables of which we have considered only 2 variables. Let the variable of interest Z represents the height of the conifer tree measured in feet while the ancillary variable W is the diameter of the tree at breast height. In Tab. 9, we provide the summary statistics of the data, and the corresponding plots of the data are displayed in Fig. 1.

Table 9: Statistics summary of the trees data

Variable	μ	σ^2	Skewness	Kurtosis	Q	ρ
Diameter (W) in cm	20.84	310.11	0.884	–0.423	14.5	0.908
Height (Z) in feet	52.36	325.14	1.619	1.776	29	

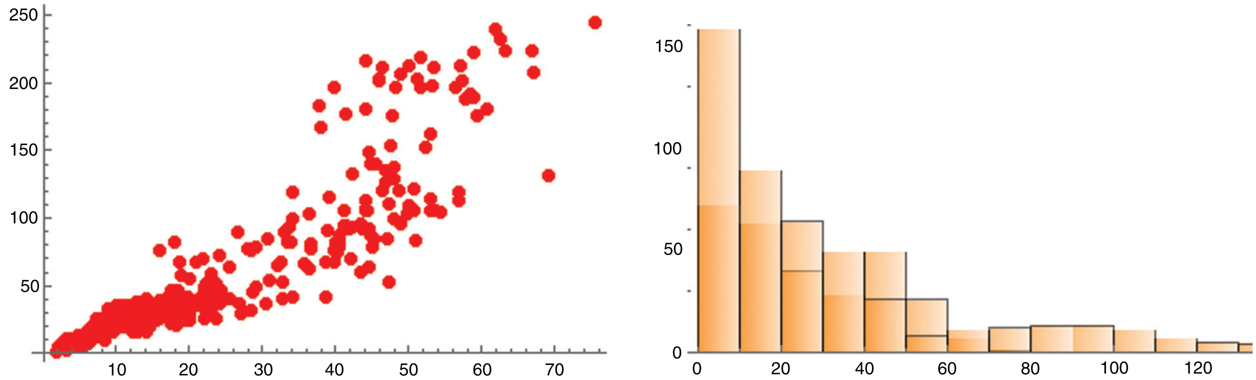

Figure 1: List plot (left) and histogram (right) of the 399-tree data

Table 10: The *Eff* of estimating of the population mean and median of the study variable based on perfect and imperfect rankings

		$\delta=0$	$\delta=1$	$\delta=2$		$\delta=3$	
		RSS	GZRSS	GZRSS	TBRSS	GZRSS	TBRSS
Population mean							
W (ranking on W)	$m=4$	1.9328	1.9562	–	–	–	–
W (ranking on Z)		1.9146	1.9488	–	–	–	–
W (ranking on W)	$m=5$	2.2201	2.0432	1.8576	1.3707	–	–
W (ranking on Z)		2.2083	2.0329	1.8514	1.3630	–	–
W (ranking on W)	$m=6$	2.5159	2.1750	1.8590	1.6181	–	–
W (ranking on Z)		2.4931	2.1654	1.8562	1.6102	–	–
W (ranking on W)	$m=7$	2.7956	2.3424	1.9183	1.8810	1.7934	0.7801
W (ranking on Z)		2.7658	2.3259	1.9182	1.8706	1.7930	0.7886
Population median							
W (ranking on X)	$m=4$	2.3463	4.0284	–	–	–	–
W (ranking on Z)		2.3222	3.9729	–	–	–	–
W (ranking on X)	$m=5$	3.0024	4.4195	5.5350	1.8235	–	–
W (ranking on Z)		2.9751	4.4076	5.4368	1.7997	–	–
W (ranking on X)	$m=6$	3.4954	4.7359	6.2749	2.3903	–	–
W (ranking on Z)		3.4450	4.6950	6.1556	2.3598	–	–
W (ranking on W)	$m=7$	3.9783	4.8876	6.3597	3.1500	7.2088	1.5035
W (ranking on Z)		3.9036	4.8217	6.2867	3.0682	7.0979	1.4719

For the diameter and the height, the coefficients of skewness are 0.884 and 1.619 respectively, indicating that these data are non-symmetric. The MSEs for various estimators (under SRS, RSS, TBRSS and GZRSS methods) were calculated by one million iterations. The obtained results are summarized in Tab. 10. The samples were drawn using SRS without replacement. The results given in Tab. 10 are the mean and median estimation values of the trees' heights under perfect and imperfect rankings. These results demonstrate that the GZRSS estimators are more efficient than their competitors. As we concluded in the above sections, the RE increases as sample size increases and vice versa. The perfect ranking provides efficient estimates than imperfect ranking. Also, the relative efficiencies under GZRSS in median estimation are greater than mean estimation because the data is asymmetrically distributed. The GZRSS is recommended for estimating the mean and median of the trees data.

7 Conclusions

We propose two new efficient RSS sampling methods for estimating the population mean and median. The proposed estimators based on the new designs are compared with their competitors using SRS, RSS and TBRSS techniques based on the same number of quantified units. It turns out that the GZRSS estimators of the population mean for symmetric populations are unbiased. It is worth mentioning that for non-uniform symmetric distributions, under perfect and imperfect rankings, the mean estimators under the proposed GZRSS are more efficient than those under SRS, RSS and TBRSS methods. We also compare the performance of the mean estimator under GZRSS with the SRS linear regression estimators. It is observed that for small and moderate correlation between the study and ancillary variables, the suggested estimators are more efficient than the SRS linear regression estimator for perfect and imperfect rankings. Therefore, we recommend the use of the proposed sampling methods over the existing RSS methods, considered here. The proposed methods, in this paper, can be considered in many real applications, such as mean estimation in case of missing data [46], quality control charts for monitoring the process mean [47], and in acceptance sampling plans [48,49].

Funding Statement: The authors extend their appreciation to Deanship of Scientific Research at King Khalid University for funding this work through Research Groups Program under Grant No. R. G. P. 2/68/41. I. M. A. and A. I. A. received the grant.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] G. A. McIntyre, "A method for unbiased selective sampling using ranked sets," *Australian Journal of Agricultural Research*, vol. 3, no. 4, pp. 385–390, 1952.
[2] K. Takahasi and K. Wakimoto, "On unbiased estimates of the population mean based on the sample stratified by means of ordering," *Annals Institute of Statistics and Mathematics*, vol. 20, pp. 1–31, 1986.
[3] T. R. Dell and J. L. Clutter, "Ranked set sampling theory with order statistics background," *Biometrics*, vol. 28, pp. 545–553, 1972.
[4] S. L. Stokes, "Ranked set sampling with concomitant variables," *Communication in Statistics Theory and Methods*, vol. 6, pp. 1207–1211, 1977.
[5] G. P. Patil, A. K. Sinha and C. Taillie, "Relative precision of ranked set sampling: A comparison with the regression estimator," *Environmetrics*, vol. 4, no. 4, pp. 399–412, 1993.
[6] L. K. Halls and T. R. Dell, "Trial of ranked-set sampling for forage yields," *Forest Sciences*, vol. 12, no. 1, pp. 22–26, 1996.
[7] H. M. Samawi, M. S. Ahmed and W. Abu-Dayyeh, "Estimating the population mean using extreme ranked set sampling," *Biometrical Journal*, vol. 38, no. 5, pp. 577–586, 1996.

[8] A. I. Al-Omari and C. N. Bouza, "Ratio estimators of the population mean with missing values using ranked set sampling," *Environmetrics*, vol. 26, no. 2, pp. 67–76, 2015.

[9] A. I. Al-Omari, "Estimation of the population median of symmetric and asymmetric distributions using double robust extreme ranked set sampling," *Revista Investigacion Operacional*, vol. 31, no. 3, pp. 200–208, 2010.

[10] M. F. Al-Saleh and A. I. Al-Omari, "Multistage ranked set sampling," *Journal of Statistical Planning and Inference*, vol. 102, no. 2, pp. 273–286, 2002.

[11] A. D. Al-Nasser, "L ranked set sampling: A generalized procedure for robust visual sampling," *Communications in Statistics–Simulation and Computation*, vol. 36, no. 1, pp. 33–43, 2007.

[12] H. A. Muttlak, "Median ranked set sampling," *Journal of Applied Statistical Science*, vol. 6, pp. 245–255, 1997.

[13] A. A. Jemain, A. I. Al-Omari and K. Ibrahim, "Multistage median ranked set sampling for estimating the population median," *Journal of Mathematics and Statistics*, vol. 3, no. 2, pp. 58–64, 2007.

[14] A. A. Jemain, A. I. Al-Omari and K. Ibrahim, "Some variations of ranked set sampling," *Electronic Journal of Applied Statistical Analysis*, vol. 1, no. 1, pp. 1–15, 2008.

[15] A. I. Al-Omari, "Ratio estimation of population mean using auxiliary information in simple random sampling and median ranked set sampling," *Statistics and Probability Letters*, vol. 82, no. 11, pp. 1883–1890, 2012.

[16] A. I. Al-Omari, "Estimation of entropy using random sampling," *Journal of Computational and Applied Mathematics*, vol. 261, pp. 95–102, 2014.

[17] S. S. Hossain and H. A. Muttlak, "Paired ranked set sampling: A more efficient procedure," *Environmetrics*, vol. 10, pp. 195–212, 1999.

[18] M. R. Shadid, M. Z. Raqaband and A. I. Al-Omari, "Modified BLUEs and BLIEs of the location and scale parameters and the population mean using ranked set sampling," *Journal of Statistical Computation and Simulation*, vol. 81, no. 3, pp. 261–274, 2011.

[19] A. I. Al-Omari and A. D. Al-Nasser, "Ratio estimation using multistage median ranked set sampling approach," *Journal of Statistical Theory and Practice*, vol. 12, no. 3, pp. 512–529, 2018.

[20] A. I. Al-Omari, "Estimation of mean based on modified robust extreme ranked set sampling," *Journal of Statistical Computation and Simulation*, vol. 81, no. 8, pp. 1055–1066, 2011.

[21] A. Haq, "A new Shewhart control chart for monitoring process mean based on partially ordered judgment subset sampling," *Quality & Quantity*, vol. 49, no. 3, pp. 1185–1202, 2015.

[22] A. Haq, J. Brown and E. Moltchanova, "Improved best linear unbiased estimators for the simple linear regression model using double ranked set sampling schemes," *Communications in Statistics Theory and Methods*, vol. 45, no. 12, pp. 3541–3561, 2015.

[23] P. L. H. Yu and K. Lam, "Regression estimator in ranked set sampling," *Biometrics*, vol. 53, no. 3, pp. 1070–1080, 1997.

[24] A. D. Al-Nasser and A. B. Mustafa, "Robust extreme ranked set sampling," *Journal of Statistical Computation and Simulation*, vol. 79, no. 7, pp. 859–867, 2009.

[25] A. Haq and J. Shabbir, "A family of ratio estimators for population mean in extreme ranked set sampling using two auxiliary variables," *SORT*, vol. 34, no. 1, pp. 45–64, 2010.

[26] O. Ozturk, "Sampling from partially rank-ordered sets," *Environmental and Ecological Statistics*, vol. 18, no. 4, pp. 757–779, 2011.

[27] E. Zamanzade and A. I. Al-Omari, "New ranked set sampling for estimating the population mean and variance," *Hacettepe Journal of Mathematics and Statistics*, vol. 45, no. 6, pp. 1891–1905, 2016.

[28] A. Haq, "Two-stage cluster sampling with hybrid ranked set sampling in the secondary sampling frame," *Communications in Statistics Theory and Methods*, vol. 46, no. 17, pp. 8450–8467, 2017.

[29] A. Haq, "Estimation of the distribution function under hybrid ranked set sampling," *Journal of Statistical Computation and Simulation*, vol. 87, no. 2, pp. 313–327, 2017.

[30] A. I. Al-Omari and E. Zamanzade, "Goodness of-fit-tests for Laplace distribution in ranked set sampling," *Revista Investigacin Operacional*, vol. 38, no. 4, pp. 366–276, 2017.

[31] A. I. Al-Omari and E. Zamanzade, “Goodness of fit tests for logistic distribution based on Phi-divergence,” *Electronic Journal of Applied Statistical Analysis*, vol. 11, no. 1, pp. 185–195, 2018.

[32] A. D. Al-Nasser and A. I. Al-Omari, “Information theoretic weighted mean based on truncated ranked set sampling,” *Journal of Statistical Theory and Practice*, vol. 9, no. 2, pp. 313–329, 2014.

[33] E. Zamanzade and M. Mahdizadeh, “Estimating the population proportion in pair ranked set sampling with application to air quality monitoring,” *Journal of Applied Statistics*, vol. 45, no. 3, pp. 426–437, 2018.

[34] A. Haq, “Ordered partially ordered judgment subset sampling with applications to parametric inference,” *Journal of Statistical Computation and Simulation*, vol. 89, no. 18, pp. 3354–3376, 2019.

[35] A. I. Al-Omari and M. Z. Raqab, “Estimation of the population mean and median using truncation-based ranked set samples,” *Journal of Statistical Computation and Simulation*, vol. 83, no. 8, pp. 453–471, 2012.

[36] A. D. Al-Nasser and A. I. Al-Omari, “MiniMax ranked set sampling,” *Revista Investigacin Operacional*, vol. 39, no. 4, pp. 560–570, 2018.

[37] A. Haq, J. Brown and E. Moltchanova, “Hybrid ranked set sampling scheme,” *Journal of Statistical Computation and Simulation*, vol. 86, no. 1, pp. 1–28, 2014.

[38] Y. G. Wang, Z. H. Chen and J. Liu, “General ranked set sampling with cost consideration,” *Biometrics*, vol. 60, no. 2, pp. 556–561, 2004.

[39] H. A. Muttlak, “Median ranked set sampling with concomitant variables and a comparison with ranked set sampling and regression estimators,” *Environmetrics*, vol. 9, pp. 225–267, 1998.

[40] E. Zamanzade and M. Mahdizadeh, “A more efficient proportion estimator in ranked set sampling,” *Statistics & Probability Letters*, vol. 129, pp. 28–33, 2017.

[41] M. Mahdizadeh and E. Zamanzade, “Smooth estimation of a reliability function in ranked set sampling,” *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 52, no. 4, pp. 750–768, 2018.

[42] M. Mahdizadeh and E. Zamanzade, “Estimation of a symmetric distribution function in multistage ranked set sampling,” *Statistical Papers*, vol. 61, no. 2, pp. 851–867, 2020.

[43] L. Dümbgen and E. Zamanzade, “Inference on a distribution function from ranked set samples,” *Annals of the Institute of Statistical Mathematics*, vol. 72, no. 1, pp. 157–185, 2020.

[44] E. Zamanzade and M. Mahdizadeh, “Using ranked set sampling with extreme ranks in estimating the population proportion,” *Statistical Methods in Medical Research*, vol. 29, no. 1, pp. 165–177, 2020.

[45] W. J. Platt, G. W. Evans and S. L. Rathbun, “The population dynamics of a long lived conifer (pinus palustris),” *American Nature*, vol. 131, no. 4, pp. 491–525, 1988.

[46] C. N. Bouza and A. I. Al-Omari, “Ranked set estimation with imputation of the missing observations: The median estimator,” *Revista Investigacion Operacional*, vol. 32, no. 1, pp. 30–37, 2011.

[47] A. I. Al-Omari and A. Haq, “Improved quality control charts for monitoring the process mean, using double-ranked set sampling methods,” *Journal of Applied Statistics*, vol. 39, no. 4, pp. 745–763, 2012.

[48] A. I. Al-Omari, “Improved acceptance sampling plans based on truncated life tests for Garima distribution,” *International Journal of System Assurance Engineering and Management*, vol. 9, no. 6, pp. 1287–1293, 2018.

[49] A. I. Al-Omari, “The transmuted generalized inverse Weibull distribution in acceptance sampling plans based on life tests,” *Transactions of the Institute of Measurement and Control*, vol. 40, no. 16, pp. 4432–4443, 2018.