Mathematics and Statistics 11(2): 263-276, 2023 http://www.hrpub.org
DOI: 10.13189/ms.2023.110204

Estimation of the Location Parameter of Cauchy Distribution Using Some Variations of the Ranked Set Sampling Technique

Arwa Salem Maabreh[1], Mohammad Fraiwan Al-Saleh[1,2,*]

[1]Department of Statistics, Yarmouk University, Jordan
[2]Department of Mathematics & Statistics, Jordan University of Science and Technology, Jordan

Received November 20, 2022; Revised January 4, 2023; Accepted January 12, 2023

Cite This Paper in the Following Citation Styles

(a): *[1] Arwa Salem Maabreh, Mohammad Fraiwan Al-Saleh, "Estimation of the Location Parameter of Cauchy Distribution Using Some Variations of the Ranked Set Sampling Technique," Mathematics and Statistics, Vol. 11, No. 2, pp. 263 - 276, 2023. DOI: 10.13189/ms.2023.110204.*

(b): *Arwa Salem Maabreh, Mohammad Fraiwan Al-Saleh (2023). Estimation of the Location Parameter of Cauchy Distribution Using Some Variations of the Ranked Set Sampling Technique. Mathematics and Statistics, 11(2), 263 - 276. DOI: 10.13189/ms.2023.110204.*

Abstract It is well known that ranked set sampling (RSS) technique and its variations, when applicable, are more efficient for estimating the population mean than the usual random sampling techniques. Despite the fascinating applications of Cauchy distribution, it has many unusual properties. For example: its moments either don't exist or exist but are infinite, and its minimal sufficient statistics are just the order statistics. Given that the shape of the Cauchy distribution is similar to the normal one, it would be advantageous to carry out some statistical studies to focus on estimating its parameters; in particular the location parameter which is the median. In this paper, the estimation of the location parameter of the Cauchy distribution using RSS and some of its variations; namely, Double RSS, Median RSS, Multistage RSS, and Steady-State RSS are considered. The estimators are compared with each other and with their counterparts using simple random sampling (SRS). The findings show that RSS or any of its variations, being evaluated in this study, is more efficient in estimating the location parameter compared to SRS. The comparison among the RSS variations reveals that the steady-state RSS is more efficient than other RSS variations. Moreover, to overcome some of the challenges of Cauchy distribution, such as the non-existence of moments, a truncated Cauchy distribution is used. For this distribution, all moments are finite as well as the moments of order statistics. Results show that RSS and Median RSS outperform the SRS in estimating the location parameter, even with the truncated version of Cauchy. Overall, the work of this paper identifies other advantages of RSS techniques.

Keywords Cauchy Distribution, Double Ranked Set Sampling, Median Ranked Set Sampling, Multistage Ranked Set Sampling, Ranked Set Sampling, Simple Random Sampling, Steady State Ranked Set Sampling, Truncated Cauchy Distribution

1. Introduction

The Cauchy distribution was named after Augustin Cauchy in 1853. It is symmetric about its location parameter and is very similar to the normal distribution but with a heavy tail. The distribution has many applications. "Simulating the energy distribution in an unstable state of an atom in quantum physics [1], modelling the relationship between two normal random variables [2], and modelling the financial returns in economics [3]" are just examples of its main applications.

Let X be a Cauchy random variable (r.v.), then the pdf of X is:

$$f(x;\theta,\lambda) = \frac{1}{\pi}\frac{\lambda}{\lambda^2 + (x-\theta)^2}, \quad -\infty < x,\theta < \infty, \ 0 < \lambda < \infty \quad (1)$$

θ is the location parameter and λ is the scale parameter. The cumulative distribution function (cdf) is:

$$F(x) = 0.5 + \frac{1}{\pi}tan^{-1}\left(\frac{x-\theta}{\lambda}\right). \quad (2)$$

For more details about the Cauchy distribution, see [4].

The Cauchy distribution is often used as an example to show the failure of the method of moments of estimating population parameters because of the undefined mean and variance [2]. Many researchers discussed the difficulties of maximum likelihood estimation of the parameters of Cauchy distribution, due to the presence of multiple roots of the likelihood equation [5, 6]. Barnett [7] discussed a numerical method for finding a global maximum to estimate θ when λ is known. Haas et al. [8] developed numerical methods for solving the likelihood equations, and Ferguson [9] obtained the closed forms of MLEs of the parameters for sample sizes 3 and 4. Copas [10] showed that the joint likelihood function has only one mode. Therefore, the situation when both parameters are unknown is simpler than the case of estimating θ when λ is known. Howlader and Weiss [11] considered the Bayesian estimation of the Cauchy Parameters; it turned out that the Bayes estimators are computationally difficult due to the instability of the estimators.

As an alternative approach to obtain more efficient estimators, several candidates were developed and investigated based on order statistics. The first was based on the median instead of the mean. However, the efficiency of the estimator w.r.t. Cramer-Rao lower bound is 0.81, when the sample size is large,. Rothenberg [12] considered another estimator which is the average of a "central subset of the sample order statistics". Bloch [13] suggested an unbiased estimator for θ based on 5-order Statistics, Chan [14] discussed the asymptotic best linear unbiased estimation of the parameters θ and λ based on k optimally selected order statistics, when at least one of the parameters is unknown. Chu ́v et al. [15] obtained the best linear unbiased estimator (BLUE) of θ using RSS. AlSubh [16] considered the maximum likelihood estimation based on RSS. It turned out that the MLE based on RSS is more efficient than the MLE based on SRS.

Given that the Cauchy distribution has so many real life applications, it would be advantageous to look for different ways of estimating the parameters. Simple random sampling is the most popular method of choosing the sample to estimate the parameters. However, other sampling techniques, such as ranked set sampling, might be more effective. The results in too many published work revealed that RSS is more efficient than SRS when estimating the parameters of some well-known distributions such as Normal distribution, Exponential distribution, and Pareto distribution. Therefore, we believe that comparing Simple Random Sampling and Ranked Set Sampling and its variations in estimating the Cauchy parameters would be of great interest.

McIntyre [17] introduced the RSS technique to choose a random sample. It turned out that the mean of the chosen sample using this technique as an estimator of the population's mean is more efficient than the corresponding one using SRS, provided that the units selected from a population are difficult or expensive to measure w.r.t. the characteristic of interest, but can be easily ranked by judgment. Takahasi and Wakimoto [18] established the theory of RSS. They have shown that the efficiency of the mean using RSS w.r.t. the mean of SRS is larger than or equal to 1 and less than or equal to *(m+1)/2*. Kaur et al. [19] provided an annotated bibliography of the literature on RSS. Al-Omari and Bouza [20] reviewed the modification of RSS and its applications.

The purpose of this paper is to consider different ranked set sampling techniques to estimate θ when λ is known, namely, RSS, Median RSS, Double RSS, Multistage RSS, and Steady-State RSS. The performance of the suggested estimators is compared with their counterparts using SRS. Throughout this paper, without loss of generality, we assume that $\lambda = 1$. The rest of work of this paper is organized as follows: SRS, RSS, MRSS, DRSS, MSRSS, and SSRSS procedures and the suggested estimators are discussed in Sections 27, respectively. The truncated Cauchy distribution used to estimate the location parameter based on SRS, RSS, and MRSS is discussed in Section 8. Results and discussion are the content of Section 9. Finally, the conclusions are given in Section 10.

2. Estimation of the Location Parameter Using SRS

Let $X_1, X_2, \ldots, X_n$ be SRS of size n from $C(\theta,1)$. Since θ is the median of the distribution, consider the sample median, $\hat{\theta}_{SRS}$, as an estimator of θ

$$\hat{\theta}_{SRS} = median\{X_i; i = 1,\cdots,n = rm\}.$$

Let $X_{(i)}$ be the i^{th} order statistic of $X_1, X_2, \ldots, X_n$, then

$$\hat{\theta}_{SRS} = \begin{cases} X_{\left(\frac{n+1}{2}\right)}, & if \ n \ is \ odd \\ \dfrac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}, & if \ n \ is \ even \end{cases} \tag{3}$$

Since the Cauchy distribution is symmetric around θ, then $\hat{\theta}_{SRS}$ is an unbiased estimator of θ, with variance:

$$\text{var}(\hat{\theta}_{SRS}) = \begin{cases} \sigma^2_{(\frac{n+1}{2})}, & if \ n \ is \ odd \\ \dfrac{1}{4}\left[\sigma^2_{(\frac{n}{2})} + \sigma^2_{(\frac{n}{2}+1)} + 2\,\text{cov}(X_{(\frac{n}{2})}, X_{(\frac{n}{2}+1)})\right], & if \ n \ is \ even \end{cases} \tag{4}$$

where $\sigma^2_{(i)}$ is the variance of the i^{th} order statistic of a random sample of size *n* from $C(\theta,1)$. According to [6], $\sigma^2_{(i)}$ is undefined for $i = 1, 2, n-1$ and n, and this imposes the constraint that the sample size must be greater than 4.

3. Estimation of the Location Parameter Using RSS

McIntyre [17] introduced RSS as a more efficient alternative to SRS. The sample is obtained by selecting m random samples of size m each; m should be small enough so that the units can be ranked by judgment without making a ranking error. The i^{th} ranked observation from the i^{th} sample is quantified after the m units are sorted by the characteristic of interest. The whole process can be performed as many times as needed (r cycles).

Therefore, the elements of RSS of size $n = rm$ from $C(\theta,1)$, where m is the set size and r is the number of cycles is: $\left\{X_{(i:m)}^{(j)}; i = 1,\ldots,m \;\; ; \; j = 1,\ldots,r\right\}$, $X_{(i:m)}^{(j)}$ is the i^{th} order statistic of the i^{th} random sample of size m at the j^{th} cycle.

It should be noted that a total of $m^2 r$ units have to be drawn from the population and only mr of them are to be quantified. Unlike the order statistics from SRS, the elements of RSS are independent order statistics. The suggested estimator of θ using RSS is

$$\hat{\theta}_{RSS} = median\left\{X_{(i:m)}^{(j)};\; i = 1,\ldots,m;\; j = 1,\ldots,r\right\}. \tag{5}$$

Numerical comparisons of $\hat{\theta}_{RSS}$ and $\hat{\theta}_{SRS}$ are given in Table 1.

4. Estimation of the Location Parameter Using MRSS

Muttlak [21] introduced a modification of RSS, called Median RSS. The MRSS can be obtained using the following steps:

<u>Case 1</u>: Odd m

Step (1): Select randomly m samples each of size m from $C(\theta,1)$.

Step (2): Rank (by Judgment) the units w.r.t. the variable of interest. The $(\frac{m+1}{2})^{th}$ smallest rank element, the median of the sample, is chosen for quantification.

Step (3): The above 2 steps may be repeated r times to get a MRSS of size $n = rm$ units.

The obtained MRSS of size $n = rm$ is $\left\{ X^{(j)}_{i(\frac{m+1}{2}:m)} ; i = 1,\dots,m ; j = 1,\dots,r \right\}$, where $X^{(j)}_{i(\frac{m+1}{2}:m)}$ is the median of the i^{th} sample $(i = 1,\dots,m)$ at the j^{th} cycle.

<u>**Case 2**</u>**:** Even m

The MRSS (even) can be obtained using the following steps:

Step (1): Select randomly m samples each of size m from $C(\theta,1)$.

Step (2): From the first $(\frac{m}{2})$ samples, rank (by Judgment) the units w.r.t. the variable of interest for each sample, and select the $(\frac{m}{2})^{th}$ smallest ranked unit for quantification.

Step (3): From the last $(\frac{m}{2})$ samples, rank (by Judgment) the units w.r.t. the variable of interest for each sample, and select the $(\frac{m}{2}+1)^{th}$ smallest ranked unit for quantification.

Step (4): The above 3 steps may be repeated r times to get a MRSS of size $n = rm$ units.

The obtained sample

$\left\{ X^{(j)}_{i(\frac{m}{2}:m)} ; i = 1,\dots,\frac{m}{2} \right\} \cup \left\{ X^{(j)}_{i(\frac{m}{2}+1:m)} ; i = \frac{m}{2}+1,\dots,m \right\}$ is called MRSS of size rm, where $X^{(j)}_{i(\frac{m}{2}:m)}$ is the $(\frac{m}{2})^{th}$ smallest ranked unit of the i^{th} sample $(i = 1,\dots,\frac{m}{2})$ at j^{th} cycle, and $X^{(j)}_{i(\frac{m}{2}+1:m)}$ is the $(\frac{m}{2}+1)^{th}$ smallest ranked unit of the i^{th} sample $(i = \frac{m}{2}+1,\dots,m)$ at j^{th} cycle.

Note that the first $\frac{m}{2}$ units are iid, and the second $\frac{m}{2}$ units are iid. However, $X^{(j)}_{1(\frac{m}{2}:m)},\dots,X^{(j)}_{\frac{m}{2}(\frac{m}{2}:m)}$, $X^{(j)}_{\frac{m}{2}+1(\frac{m}{2}+1:m)},\dots,X^{(j)}_{m(\frac{m}{2}+1:m)}$ are independent but not identically distributed. The suggested estimator for θ using MRSS is

$$\hat{\theta}_{MRSS} = median \begin{cases} X^{(j)}_{i(\frac{m+1}{2}:m)} & 1 \le i \le m,\ 1 \le j \le r \text{ , if m is odd} \\ X^{(j)}_{i(\frac{m}{2}:m)},\dots,X^{(j)}_{i+\frac{m}{2}(\frac{m}{2}+1:m)} & 1 \le i \le \frac{m}{2} \text{ ,} 1 \le j \le r \text{ , if m is even} \end{cases} \tag{6}$$

To study the properties of $\hat{\theta}_{MRSS}$, for *r=1*, let $\{D_1, D_2, \dots, D_m\}$ be the obtained MRSS of size *m* from $C(\theta,1)$, and defined as $\left\{ D_i = X_{i(\frac{m+1}{2}:m)} ;\ 1 \le i \le m \right\}$ when *m* is odd, and $\left\{ D_i = X_{i(\frac{m}{2}:m)} ; 1 \le i \le \frac{m}{2} \right\}$ $\cup \left\{ D_i = X_{i(\frac{m}{2}+1:m)} ;\ \frac{m}{2} \le i \le m \right\}$ when *m* is even.

In this case,

$$\hat{\theta}_{MRSS} = \begin{cases} D_{(\frac{m+1}{2})}, & m \text{ is odd} \\ \frac{1}{2}\left[D_{(\frac{m}{2})} + D_{(\frac{m}{2}+1)} \right], & m \text{ is even.} \end{cases}$$

Suppose that $W_{(i)} = D_{(i)} - \theta$, $W_{(i)}$ has the pdf of i^{th} order statistic of MRSS from $C(0,1)$, then, $\mathrm{E}(D_{(i)}) = \mathrm{E}(W_{(i)}) + \theta$.

To find the expected value of $\hat{\theta}_{MRSS}$, the derivative of the pdf of $W_{(\frac{m+1}{2})}$, $W_{(\frac{m}{2})}$, $W_{(\frac{m}{2}+1)}$ is required. Therefore, for odd *m*, assuming that the pdf of $W_{(\frac{m+1}{2})}$ is $g_{(\frac{m+1}{2})}(w)$. So, for $\{ W_i = X_{i(\frac{m+1}{2}:m)} \;;\; 1 \le i \le m \}$ the pdf is given by $f_{(\frac{m+1}{2})}(x) = \frac{m!}{\left[\left(\frac{m-1}{2}\right)!\right]^2} [F(x)]^{\frac{m-1}{2}} [1-F(x)]^{\frac{m-1}{2}} f(x)$, and the cdf is $F_{(\frac{m+1}{2})}(x) = \int_{-\infty}^{x} f_{(\frac{m+1}{2})}(t)\, dt$.

Thus, the pdf of $W_{(\frac{m+1}{2})}$ is

$$g_{(\frac{m+1}{2})}(w) = \frac{m!}{\left[\left(\frac{m-1}{2}\right)!\right]^2} [F_{(\frac{m+1}{2})}(w)]^{\frac{m-1}{2}} [1-F_{(\frac{m+1}{2})}(w)]^{\frac{m-1}{2}} f_{(\frac{m+1}{2})}(w). \; -\infty < w < \infty \tag{7}$$

For $X \sim C(\theta,1)$, we have $X_{(\frac{m+1}{2})} \overset{d}{=} -X_{(\frac{m+1}{2})}$ (i.e. the Cauchy is symmetric distribution, [22]). So,

$$f_{(\frac{m+1}{2})}(x) = f_{(\frac{m+1}{2})}(-x), \quad F_{(\frac{m+1}{2})}(-x) = 1 - F_{(\frac{m+1}{2})}(x), \quad \text{and } 1 - F_{(\frac{m+1}{2})}(-x) = F_{(\frac{m+1}{2})}(x). \tag{8}$$

Applying (8) in (7), we get

$$g_{(\frac{m+1}{2})}(-w) = \frac{m!}{\left[\left(\frac{m-1}{2}\right)!\right]^2} [1-F_{(\frac{m+1}{2})}(w)]^{\frac{m-1}{2}} [F_{(\frac{m+1}{2})}(w)]^{\frac{m-1}{2}} f_{(\frac{m+1}{2})}(w) = g_{(\frac{m+1}{2})}(w)$$

.

Therefore, $g_{(\frac{m+1}{2})}(w)$ is a symmetric pdf $\rightarrow W_{(\frac{m+1}{2})} \overset{d}{=} -W_{(\frac{m+1}{2})}$, and

$$\mathrm{E}(W_{(\frac{m+1}{2})}) = 0 \quad \rightarrow \quad \mathrm{E}(D_{(\frac{m+1}{2})}) = \theta$$

So, $\hat{\theta}_{MRSS,odd}$ is an unbiased estimator of θ, with

$$\mathrm{var}(\hat{\theta}_{MRSS,odd}) = \int_{-\infty}^{\infty} w^2\, g_{(\frac{m+1}{2})}(w)\, dw . \tag{9}$$

For the even case, which requires a more extensive theoretical justification, we employ simulation to examine the estimator's properties. The results are given in Table 2.

5. Estimation of the Location Parameter Using DRSS

Double RSS is variation of RSS introduced by Al-Saleh and Al-Kadiri [23]. As stated in their paper, the DRSS can be obtained using the following steps:

Step (1): Select at random m^3 elements from the population of interest and partition the set of elements into these elements into m subsets; each of the size m^2 elements.

$$\underset{\text{set}(1)}{\begin{bmatrix} X_{11}^{1} & \cdots & X_{1m}^{1} \\ \vdots & & \\ X_{m1}^{1} & \cdots & X_{mm}^{1} \end{bmatrix}} \quad \cdots \quad \underset{\text{set}(m)}{\begin{bmatrix} X_{11}^{m} & \cdots & X_{m1}^{m} \\ \vdots & & \\ X_{m1}^{m} & \cdots & X_{mm}^{m} \end{bmatrix}}$$

Step (2): For each subset obtained in Step 1, apply the original RSS technique to obtain m RSSs of size m each:

$$A_1 = \left\{X_{(i:m)}^{1}\ ;\ 1 \le i \le m\right\},\quad A_2 = \left\{X_{(i:m)}^{2}\ ;\ 1 \le i \le m\right\},\quad \cdots,\quad A_m = \left\{X_{(i:m)}^{m}\ ;\ 1 \le i \le m\right\},$$

A_i is the RSS of the i^{th} subset

Step (3): Apply the RSS technique again to the m RSSs obtained in Step (2) to obtain a DRSS of size m :

$$Y_{(1)} = \min(A_1),\qquad Y_{(2)} = 2^{nd}\min(A_2),\quad \cdots,\quad Y_{(m)} = \max(A_m).$$

Step (4): Repeat Steps (1-3) r times if needed, to obtain a DRSS of size $n = rm$.

The obtained sample $\left\{Y_{(i)}^{(j)}\ ; 1 \le i \le m\ ;\ \ 1 \le j \le r\right\}$, where $Y_{(i)}^{(j)}$ is the i^{th} smallest observation in A_i $(i = 1,\dots,m)$ at the j^{th} cycle $(\ j = 1,\dots,r\ )$ is called a DRSS of size $n = rm$. Clearly, $Y_{(i)}^{(j)}$ are independent but not identically distributed.

The estimator of θ using DRSS is:

$$\hat{\theta}_{DRSS} = median\left\{\ Y_{(i)}^{(j)}\ ;\ 1 \le i \le m;\ 1 \le j \le r\ \right\}\ . \tag{10}$$

The efficiency values of the estimator w.r.t $\hat{\theta}_{SRS}$ are given in Table 3.

6. Estimation of the Location Parameter Using MSRSS

As a generalization of DRSS, Al-Saleh and Al- Omari [24] introduced another variation of RSS coined by them "Multistage RSS". The following shows the steps to obtain the sample when $k = 3$ (number of stages) and $m = 5$ (sample size):

Step (1): Select 5^4 random units from the target population.

Step (2): Partition the selected set of units obtained in step 1 randomly into 5^2 subsets, each of size 25. The subsets are:

$$\underset{Subset(1)}{\begin{bmatrix} X_{11}^{1} & \cdots & X_{15}^{1} \\ \vdots & & \\ X_{51}^{1} & \cdots & X_{55}^{1} \end{bmatrix}} \quad \underset{\text{Subset}(2)}{\begin{bmatrix} X_{11}^{2} & \cdots & X_{15}^{2} \\ \vdots & & \\ X_{51}^{2} & \cdots & X_{55}^{2} \end{bmatrix}} \quad \cdots \quad \underset{\text{Subset}(25)}{\begin{bmatrix} X_{11}^{25} & \cdots & X_{51}^{25} \\ \vdots & & \\ X_{51}^{25} & \cdots & X_{55}^{25} \end{bmatrix}}$$

Step (3): For $k = 1$ (the first stage), apply the original RSS procedure for each generated set in Step (2) to obtain 5^2 samples, each of size 5.

$$A_1 = \left\{X_{(i:5)}^{1}\ ; i = 1,\cdots,5\right\},\quad A_2 = \left\{X_{(i:5)}^{2}\ ; i = 1,\cdots,5\right\},\ \cdots,\quad A_{25} = \left\{X_{(i:5)}^{25}\ ; i = 1,\cdots,m\ \right\}.$$

Step (4): For $k=2$ (second stage), apply the RSS procedure for each generated set in Step (3) to obtain 5 samples, each of size 5.

$$\begin{array}{lll} Y_{(1)} = \min(A_1), & Y_{(2)} = 2^{nd} \min(A_2), \quad \cdots & Y_{(5)} = \max(A_5), \\ Y_{(6)} = \min(A_6), & Y_{(7)} = 2^{nd} \min(A_7), \quad \cdots & Y_{(10)} = \max(A_{10}), \\ \vdots & \vdots & \vdots \\ Y_{(21)} = \min(A_{21}), & Y_{(22)} = 2^{nd} \min(A_{22}), \quad \cdots & Y_{(25)} = \max(A_{25}). \end{array}$$

Step (5): In the third stage, the final 5 units are measured exactly for the variable of interest.

$$\begin{array}{l} Z_{(1)} = \min\{Y_1, Y_2, \cdots, Y_5\}, \\ Z_{(2)} = 2^{nd} \min\{Y_6, Y_7, \cdots, Y_{10}\}, \\ \vdots \\ Z_{(5)} = \max\{Y_{21}, Y_{22}, \cdots, Y_{25}\}. \end{array}$$

The obtained sample $\{Z_{(1)}, Z_{(2)}, \cdots, Z_{(5)}\}$ is the 3rd Stage RSS of size 5. In general, the kth stage RSS of size m is denoted by $\{ Z_{(i)}^{k};\ i = 1, 2, \cdots, m\}$.

The estimator of θ using MSRSS at stage k is

$$\hat{\theta}_{MSRSS} = median\{Z_{(i)}^{k};\ i = 1, \cdots, m\}. \tag{11}$$

The numerical results are given in Table 4.

7. Estimation of the Location Parameter Using SSRSS

Steady-State RSS is another variation of RSS which is MSRSS as the number of stages goes to infinity. In this variation, the distribution is partitioned into *m* strata. According to [24], the MSRSS mean is unbiased for the population mean, and its efficiency w.r.t the SRS mean increases as *k* (the number of stages) increases. They showed that, as *k* goes to infinity, the maximum efficiency approaches m^2, which extremely larger than $(m+1)/2$ for RSS.

SSRSS was linked to stratified sampling [25], where the support of the density is divided into *m* non-overlapped portions, all with equal probability of $1/m$.

Assume that $m = 5$, then the support of the density $C(0,1)$ is divided into five partitions, the first partition is $(-\infty, -1.3764]$, the second one is $[-1.3764, -0.3249]$, the third one is $[-0.3249, 0.3249]$, the fourth one is $[0.3249, 1.3764]$, and the last one is $[1.3764, \infty)$. Figure 1 below shows the five portions of the density; each is of area 0.20.

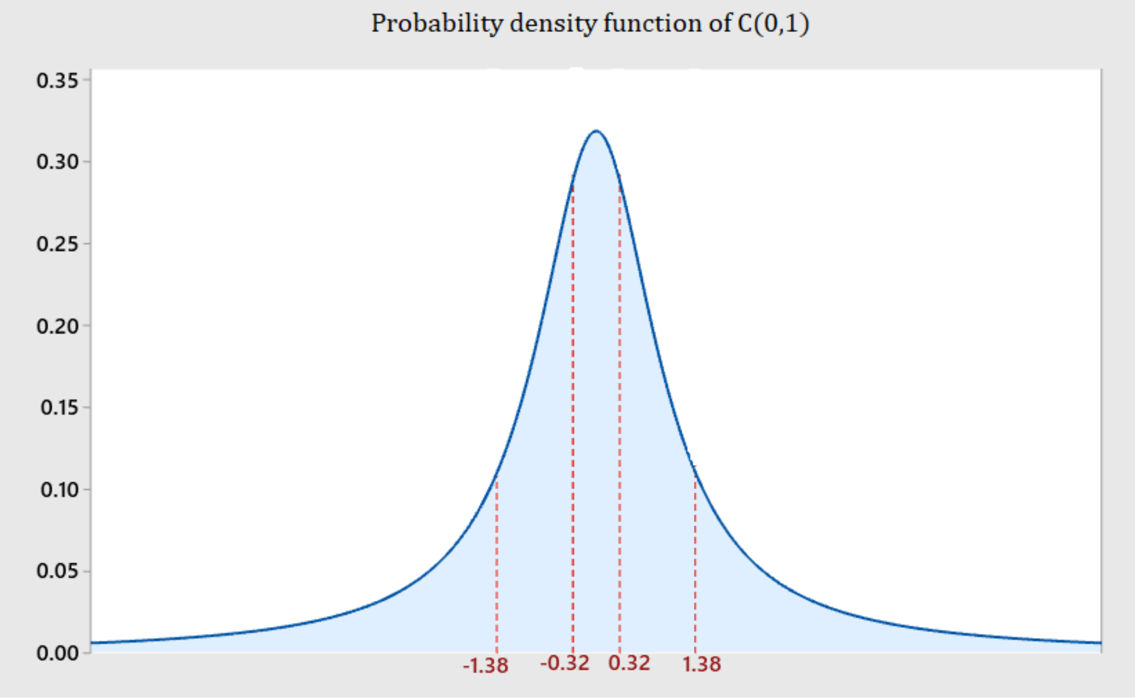

Figure 1. Standard Cauchy distribution partition into five non-overlapping strata

Now for SSRSS, we can take a random sample of size r from each truncated Cauchy distribution (each partition), denoted as $X_s^{(j)} \sim TC(\theta,1,a,b)$ for $s=1,\cdots,5$, and $j=1,\cdots,r$, where S is the number of partitions, and j is the sample size from each partition.

The estimator of θ using SSRSS is

$$\hat{\theta}_{SSRSS} = median\left\{ X_s^{(j)};\ s=1,2,\cdots,5;\ j=1,\cdots,r \right\}. \tag{12}$$

Numerical comparisons are given in Table 5.

8. Estimation of the Location Parameter Using Truncated Cauchy Distribution

To overcome the limitation of the non-existing moments, a truncated Cauchy distribution was introduced in 1970, and it was discussed by Nadarajah [26]. A continuous random variable X is distributed as truncated Cauchy, $TC(\theta,1,a,b)$, defined on the interval $[a,b]$ if its pdf is:

$$g(x) = \frac{1}{D}\frac{1}{\left(1+(x-\theta)^2\right)} \quad a \le x \le b, -\infty < \theta < \infty, \tag{13}$$

where $D = \tan^{-1}(b-\theta) - \tan^{-1}(a-\theta)$.

For simplicity, it will be assumed that the sample size is $n=m$, and the truncation is symmetric, such that the truncation is limited in the finite interval $[\theta-L, \theta+L]$, $L>0$. For large values of L, this truncated distribution is not significantly different from the Cauchy distribution. In this case, $D = 2\tan^{-1}(L)$, and the pdf is $g(x) = \frac{1}{D}\frac{1}{\left(1+(x-\theta)^2\right)} \quad \theta - L \le x \le \theta + L$.

1- Estimation of the Parameter Using SRS

Let $X_1, X_2, \ldots, X_m$ be SRS of size m from $TC\ (\theta, 1, \theta - L, \theta + L)$, then the estimator of θ using SRS is

$$\hat{\theta}_{SRS,TC} = median\ \{X_i ; i = 1, \cdots, m\} \tag{14}$$

It is an unbiased estimator of θ, with variance:

$$\text{var}(\hat{\theta}_{SRS,TC}) = \begin{cases} \sigma^{2^*}_{(\frac{m+1}{2})}, & m \text{ is odd} \\ \frac{1}{4}\left[\sigma^{2^*}_{(\frac{m}{2})} + \sigma^{2^*}_{(\frac{m}{2}+1)} + 2\,\text{cov}(X_{(\frac{m}{2})}, X_{(\frac{m}{2}+1)})\right], & m \text{ is even} \end{cases} \tag{15}$$

where $\sigma^{2^*}_{(i)}$ is the variance of the i^{th} order statistic of a random sample of size *m* from $TC\ (\theta, 1, -L, L)$. Note that unlike in the original Cauchy, all moments and moments of order statistics exist.

2- Estimation of the Parameter Using RSS

Assume that $\left\{X_{(i:m)}; i = 1, \ldots, m\right\}$ is RSS of size m from $TC(\theta, 1, \theta - L, \theta + L)$. The second proposed estimator of θ using RSS is $\hat{\theta}_{RSS,TC} = median\left\{X_{(i:m)}; i = 1, \ldots, m\right\}$.

3- Estimation of the Parameter Using MRSS

Assume that $\left\{X_{i(\frac{m+1}{2}:m)}; 1 \le i \le m\right\}$ is a MRSS of size *m* from $TC(\theta, 1, \theta - L, \theta + L)$ for odd sample size, and $\left\{X_{i(\frac{m}{2}:m)}, \ldots, X_{i+\frac{m}{2}(\frac{m}{2}+1:m)}; 1 \le i \le \frac{m}{2}\right\}$ is a MRSS of size *m* from $TC(\theta, 1, \theta - L, \theta + L)$ for even sample size. The obtained estimator of θ using MRSS is

$$\hat{\theta}_{MRSS,TC} = median\begin{cases} X_{i(\frac{m+1}{2}:m)}, & 1 \le i \le m,\ m \text{ is odd} \\ X_{i(\frac{m}{2}:m)}, \ldots, X_{i+\frac{m}{2}(\frac{m}{2}+1:m)}, & 1 \le i \le \frac{m}{2},\ m \text{ is even} \end{cases} \tag{16}$$

Numerical comparisons of the above estimators are given in Table 6 and Table 7.

9. Numerical Results and Discussions

A simulation study has been conducted to investigate the properties of the estimators. The simulation has been implemented using Minitab11 and R tools for different sample sizes. The $\text{var}(\hat{\theta}_{SRS})$ is calculated numerically. We use the efficiency criterion to measure the estimators' performance as follows:

$$eff\ (\hat{\theta}_1; \hat{\theta}_{SRS}) = \frac{\text{var}(\hat{\theta}_{SRS})}{MSE(\hat{\theta}_1)}, \quad \hat{\theta}_1 \in \{\hat{\theta}_{RSS}, \hat{\theta}_{MRSS}, \hat{\theta}_{DRSS}, \hat{\theta}_{MSRSS}, \hat{\theta}_{SSRSS}\}$$

$$eff\ (\hat{\theta}_{1^*}; \hat{\theta}_{SRS,TC}) = \frac{\text{var}(\hat{\theta}_{SRS,TC})}{MSE(\hat{\theta}_{1^*})}, \quad \hat{\theta}_{1^*} \in \{\hat{\theta}_{RSS,TC}, \hat{\theta}_{MRSS,TC}\}$$

$$eff\ (\hat{\theta}_2; \hat{\theta}_1) = \frac{eff\ (\hat{\theta}_2; \hat{\theta}_{SRS})}{eff\ (\hat{\theta}_1; \hat{\theta}_{SRS})}.$$

The results, based on 100,000 replications, are summarized in Tables 1-7

Table 1. The variance of $\hat{\theta}_{SRS}$, bias & MSE of $\hat{\theta}_{RSS}$ and $eff\left(\hat{\theta}_{RSS}; \hat{\theta}_{SRS}\right)$

r	m	$\text{var}\left(\hat{\theta}_{SRS}\right)$	$bias\ (\hat{\theta}_{RSS})$	$MSE\left(\hat{\theta}_{RSS}\right)$	$eff\left(\hat{\theta}_{RSS}; \hat{\theta}_{SRS}\right)$
Odd sample size					
1	5	1.2213	-0.0035	0.3174	3.8478
	7	0.6121	0.0014	0.1589	3.8521
	9	0.4087	0.0001	0.1060	3.8557
3	5	0.2048	0.0022	0.0879	2.3299
	7	0.1367	-0.0010	0.0509	2.6857
	9	0.1026	0.0003	0.0344	2.9826
Even sample size					
1	6	0.8607	-0.0076	0.1891	4.5516
	8	0.4776	-0.0005	0.1047	4.5616
	10	0.3362	0.0002	0.0733	4.5866
2	5	0.3362	0.00111	0.1185	2.8371
	6	0.2611	-0.0210	0.0870	3.0012
	7	0.2139	-0.0026	0.0681	3.1410
	8	0.1815	0.0012	0.0552	3.2880

Table 2. The variance of $\hat{\theta}_{MRSS}$, and $eff\left(\hat{\theta}_{MRSS}; \hat{\theta}_{SRS}\right)$

r	m	$\text{var}(\hat{\theta}_{MRSS})$	$eff\left(\hat{\theta}_{MRSS}; \hat{\theta}_{SRS}\right)$
Odd sample size			
1	5	0.1587	7.6957
	7	0.0762	8.0328
	9	0.0458	8.9236
3	5	0.0482	4.2490
	7	0.0248	5.5121
	9	0.0153	6.7059
Even sample size			
1	6	0.1020	8.4382
	8	0.0550	8.6836
	10	0.0353	9.5241
2	5	0.0695	4.8374
	6	0.0495	5.2747
	7	0.0357	5.9916
	8	0.0286	6.3462

Table 3. The bias, MSE of $\hat{\theta}_{DRSS}$, and $eff\left(\hat{\theta}_{DRSS};\hat{\theta}_{SRS}\right)$

r	m	$bias(\hat{\theta}_{DRSS})$	$MSE\left(\hat{\theta}_{DRSS}\right)$	$eff\left(\hat{\theta}_{DRSS};\hat{\theta}_{SRS}\right)$
Odd sample size				
1	5	-0.0031	0.1663	7.3440
	7	0.0032	0.0850	7.2011
	9	-0.0018	0.0546	7.4853
3	5	-0.0026	0.0564	3.6312
	7	0.0021	0.0302	4.5265
	9	0.001	0.0195	5.2615
Even sample size				
1	6	0.0006	0.0851	10.1140
	8	-0.0024	0.0470	10.1617
	10	0.0001	0.0317	10.6057
2	5	-0.0009	0.0685	4.9080
	6	0.0012	0.0471	5.5435
	7	0.0001	0.0361	5.9252
	8	-0.001	0.0282	6.4362

Table 4. The bias, MSE of $\hat{\theta}_{MSRSS}$, and $eff\left(\hat{\theta}_{MSRSS};\hat{\theta}_{SRS}\right)$

m	$bias(\hat{\theta}_{MSRSS})$	$MSE\left(\hat{\theta}_{MSRSS}\right)$	$eff\left(\hat{\theta}_{MSRSS};\hat{\theta}_{SRS}\right)$
Odd sample size			
5	0.0003	0.114	10.7132
7	-0.0027	0.0562	10.8915
9	0.0018	0.0356	11.4803
Even sample size			
6	-0.001	0.0531	16.2090
8	-0.0009	0.029	16.4690
10	-0.0021	0.0187	17.9786

Table 5. The bias, MSE of $\hat{\theta}_{SSRSS}$, and $eff\left(\hat{\theta}_{SSRSS};\hat{\theta}_{SRS}\right)$

r	m	$bias(\hat{\theta}_{SSRSS})$	$MSE\left(\hat{\theta}_{SSRSS}\right)$	$eff\left(\hat{\theta}_{SSRSS};\hat{\theta}_{SRS}\right)$
1	5	0.0001	0.0343	35.6064
2	5	-0.0004	0.0172	19.5465
3	5	0.0003	0.0203	10.0887

Table 6. The variance of $\hat{\theta}_{SRS,TC}$, bias, MSE of $\hat{\theta}_{RSS,TC}$, and $eff\left(\hat{\theta}_{RSS,TC};\hat{\theta}_{SRS,TC}\right)$

Truncation	m	$\hat{\text{var}}\left(\hat{\theta}_{SRS,TC}\right)$	$bias(\hat{\theta}_{RSS,TC})$	$MSE\left(\hat{\theta}_{RSS,TC}\right)$	$eff\left(\hat{\theta}_{RSS,TC};\hat{\theta}_{SRS,TC}\right)$
$\left[-10^5,10^5\right]$	3	12.1552	-0.0077	1.1697	10.3917
	4	6.0902	0.0051	0.5810	10.4823
	5	1.2213	0.0029	0.3026	4.0360
	6	0.8604	0.0003	0.1838	4.6812
$\left[-10^{10},10^{10}\right]$	3	26.1526	0.0043	1.1625	22.4969
	4	5.2271	0.0044	0.5842	8.9474
	5	1.3682	-0.0065	0.3065	4.4639
	6	0.8455	-0.0027	0.1810	4.6713
$\left[-10^{15},10^{15}\right]$	3	6.6208	0.0023	1.1753	5.6333
	4	3.6279	0.0030	0.6344	5.7187
	5	1.1988	-0.0031	0.3066	3.8200
	6	0.8526	0.0018	0.1827	4.6667
$\left[-10^{50},10^{50}\right]$	3	6.4333	0.0015	1.0837	5.9364
	4	3.6218	-0.0007	0.6151	6.1210
	5	1.2157	0.0023	0.3047	3.9898
	6	0.8313	0.0024	0.1816	4.5776

Table 7. The bias, MSE of $\hat{\theta}_{MRSS,TC}$, and $eff\left(\hat{\theta}_{MRSS,TC};\hat{\theta}_{SRS,TC}\right)$

Truncation	m	$bias(\hat{\theta}_{MRSS,TC})$	$MSE\left(\hat{\theta}_{MRSS,TC}\right)$	$eff\left(\hat{\theta}_{MRSS,TC};\hat{\theta}_{SRS,TC}\right)$
$\left[-10^5,10^5\right]$	3	0.0032	0.5741	21.1726
	4	-0.0023	0.2878	21.1612
	5	0.0007	0.1567	7.7939
	6	0.0027	0.1020	8.4353
$\left[-10^{10},10^{10}\right]$	3	0.0017	0.5556	47.0709
	4	0.0008	0.2891	18.0806
	5	-0.0006	0.1556	8.7931
	6	-0.0009	0.1025	8.2488
$\left[-10^{15},10^{15}\right]$	3	-0.0007	0.5701	11.6134
	4	0.0005	0.2931	12.3777
	5	-0.0045	0.1574	7.6163
	6	-0.0021	0.1021	8.3506
$\left[-10^{50},10^{50}\right]$	3	0.0046	0.5687	11.3123
	4	-0.0024	0.2965	12.2152
	5	0.0011	0.1564	7.7731
	6	0.0013	0.1028	8.0866

Based on Table 1 to Table 5, we have the following conclusions:

(a) The bias of each of $\hat{\theta}_{RSS}$, $\hat{\theta}_{DRSS}$, $\hat{\theta}_{MSRSS}$ and $\hat{\theta}_{SSRSS}$ is very close to zero; we may say that the estimators are nearly unbiased.

(b) The estimated MSE of $\hat{\theta}_{RSS}$ decreases as m increases for a given *r*.

(c) The variance of $\hat{\theta}_{MRSS}$ is always less than the variance of $\hat{\theta}_{SRS}$ and $\hat{\theta}_{RSS}$.

(d) The MSE of $\hat{\theta}_{DRSS}$ and $\hat{\theta}_{MSRSS}$ are decreasing in m.

(e) $Eff(\hat{\theta}_{RSS};\hat{\theta}_{SRS})>1$. Thus, $\hat{\theta}_{RSS}$ is more efficient than $\hat{\theta}_{SRS}$.

(f) Based on Table 1 and 2, $\hat{\theta}_{MRSS}$ is more efficient estimator than $\hat{\theta}_{SRS}$ and $\hat{\theta}_{RSS}$.

(g) The efficiency values of $\hat{\theta}_{DRSS}$ w.r.t $\hat{\theta}_{SRS}$ are greater than one and increase in m, but it is affected by r-values, causing a decrease in the efficiency.

(h) Based on Table 1 and Table 3, we can infer that $\hat{\theta}_{DRSS}$ is more efficient than $\hat{\theta}_{RSS}$, and the efficiency increases in m for a given r. The slight difference in efficiency values, for different values of r (number of cycles), indicates that efficiency is unaffected by the number of cycles.

(i) Based on Table 1 and Table 3, we can conclude that $Eff\,(\hat{\theta}_{DRSS};\hat{\theta}_{MRSS})$ is decreasing in m for a given r; it is less than one when the sample size is odd.

(j) $\hat{\theta}_{MSRSS}$ is a more efficient estimator $\hat{\theta}_{SRS}$, and we can infer from Table 1-Table 4 that $\hat{\theta}_{MSRSS}$ more efficient estimator than $\hat{\theta}_{RSS}$, $\hat{\theta}_{MRSS}$ and $\hat{\theta}_{DRSS}$.

(k) The efficiency values of $\hat{\theta}_{MSRSS}$ w.r.t. $\hat{\theta}_{SRS}$ are increasing in m.

(l) $\hat{\theta}_{SSRSS}$ is more efficient than $\hat{\theta}_{SRS}$, and we can infer from Table1-Table 5 that $\hat{\theta}_{SSRSS}$ is more efficient than $\hat{\theta}_{RSS}$, $\hat{\theta}_{MRSS}$, $\hat{\theta}_{DRSS}$ and $\hat{\theta}_{MSRSS}$.

From Table 6 and 7, we draw the following conclusions:

(a) The bias of $\hat{\theta}_{RSS,TC}$ and $\hat{\theta}_{MRSS,TC}$ are negligible.

(b) The estimated MSE of $\hat{\theta}_{RSS,TC}$ and $\hat{\theta}_{MRSS,TC}$ decrease as m increases, and the values of MSE for different truncations are practically the same.

(c) $\hat{\theta}_{RSS,TC}$ is a more efficient estimator than $\hat{\theta}_{SRS,TC}$ since the efficiency values are greater than one.

(d) $\hat{\theta}_{MRSS,TC}$ is a more efficient estimator than $\hat{\theta}_{SRS,TC}$.

(e) We can infer from Table 6 and Table 7 that $\hat{\theta}_{MRSS,TC}$ is a more efficient estimator than $\hat{\theta}_{RSS,TC}$ since the efficiency value is greater than one, and the efficiency values for various truncations are nearly the same.

10. Conclusions

In this paper, we investigated some popular sampling approaches, SRS, RSS, MRSS, DRSS, MSRSS, and SSRSS, for estimating the location parameter of the Cauchy distribution. The properties of the suggested estimators (bias, variance, and efficiency) were obtained and compared. The findings revealed that the estimator based on MRSS is more efficient than the estimators using SRS and RSS. The DRSS estimator is more efficient than the estimators using SRS and RSS, although it is still less efficient than the estimator based on MRSS when m is odd. The estimator using MSRSS outperforms the estimators using SRS, RSS, MRSS, and DRSS; the efficiency values are always greater than one. The estimators based on SSRSS have the highest efficiency values compared to the estimators using other variations.

Finally, to overcome some of the challenges of Cauchy distribution, several estimators have been evaluated using truncated Cauchy distribution, where all of the moments exist, as do all of the moments of order statistics. Results showed that the median of RSS and MRSS are more efficient than the median of SRS using truncated Cauchy.

Acknowledgement

The authors would like to thank the referees for their comments and suggestions.

REFERENCES

[1] Roe B. P., "Probability and statistics in experimental physics", 2nd Edition, Springer Science & Business Media, 2012.

[2] Johnson N. L., Kotz S., Balakrishnan N., "Cauchy Distribution in Continuous Univariate Distributions", Vol. 1, 2nd Edition, John Wiley & Sons Ltd., New York, pp.298-336, 1994.

[3] Nolan J. P., "Financial modeling with heavy - tailed stable distributions", Wiley Interdisciplinary Reviews: Computational Statistics, vol.6, no.1, pp.45-55, 2014. DOI: 10.1002/wics.1286.

[4] Maabreh A., Al-Saleh M. F., "A Rich Learning Statistical Lesson in the Cauchy Distribution", 2022. (submitted)

[5] Bai Z. D., Fu J. C., "On the maximum likelihood estimator for the location parameter of a Cauchy distribution", Canadian Journal of Statistics, vol.15, no. 2, pp.137-146, 1987. DOI: 10.2307/3315202.

[6] Barnett V. D., "Order statistics estimators of the location of the Cauchy distribution", Journal of the American Statistical Association, vol.61, no.316, pp.1205-1218, 1966. DOI: 10.1080/01621459.1966.10482205.

[7] Barnett V. D., "Evaluation of the maximum-likelihood estimator where the likelihood equation has multiple roots", Biometrika, vol.53, no. (1/2), pp.151-165, 1966. DOI: 10.2307/2334061.

[8] Haas G., Bain L., Antle C., "Inferences for the Cauchy distribution based on maximum likelihood estimators", Biometrika, vol.57, no.2, pp.403-408, 1970. DOI: 10.1093/biomet/57.2.403.

[9] Ferguson T. S., "Maximum likelihood estimates of the parameters of the Cauchy distribution for samples of size 3 and 4", Journal of the American Statistical Association, vol.73, no.361, pp.211-213, 1978. DOI: 10.1080/01621459.1978.10480031.

[10] Copas J. B., "On the unimodality of the likelihood for the Cauchy distribution", Biometrika, vol.62, no.3, pp.701-704, 1975. DOI: 10.1093/biomet/62.3.701.

[11] Howlader H. A., Weiss G., "On Bayesian estimation of the Cauchy parameters", Sankhya: The Indian Journal of Statistics, Series B, vol.50, no.3, pp. 350–61, 1988.

[12] Rothenberg T. J., Fisher F. M., Tilanus C. B., "A note on estimation from a Cauchy sample", Journal of the American Statistical Association, vol.59, no.306, pp.460-463, 1964. DOI: 10.1080/01621459.1964.10482170.

[13] Bloch D., "A note on the estimation of the location parameter of the Cauchy distribution", Journal of the American Statistical Association, vol.61, no.315, pp. 852-855, 1966. DOI: 10.1080/01621459.1966.10480912.

[14] Chan L. K., "Linear estimation of the location and scale parameters of the Cauchy distribution based on sample quantiles", Journal of the American Statistical Association, vol.65, no.330, pp. 851-859, 1970. DOI: 10.1080/01621459.1970.10481129.

[15] Chuiv N., Sinha B., Wu Z., "Estimation of the location parameter of a Cauchy distribution using a ranked set sample", Metrika: International Journal for Theoretical and Applied Statistics, vol.42, no.1, pp.234–235, 1995. DOI: 10.1007/BF01894304.

[16] Al-Subh, S., "Estimation of Cauchy parameters under ranked set sampling", International Journal of Sciences: Basic and Applied Research, vol.28, no.2, pp. 284-295. 2016.

[17] McIntyre G. A., "A method for unbiased selective sampling, using ranked sets", Australian Journal of Agricultural Research, vol.3, no.4, pp. 385-390, 1952. DOI: 10.1071/AR9520385.

[18] Takahasi K., Wakimoto K., "On unbiased estimates of the population mean based on the sample stratified by means of ordering", Annals of the Institute of Statistical Mathematics, vol.20, no.1, pp.1-31, 1968. DOI: 10.1007/BF02911622.

[19] Kaur A., Patil G. P., Sinha A. K., Taillie C., "Ranked set sampling: an annotated bibliography", Environmental and Ecological Statistics, vol.2, no.1, pp.25-54, 1995. DOI: 10.1007/BF00452930.

[20] Al-Omari A. I., Bouza C. N., "Review of ranked set sampling: modifications and applications," Investigation Operacional, vol.35, no.3, pp.215-235, 2014.

[21] Muttlak H. A., "Median ranked set sampling," Journal of Applied Statistical Sciences, vol. 6, pp.245-255, 1997.

[22] Arnold B. C., Balakrishnan N., Nagaraja H. N., "Basic Distribution Theory", A first course in order statistics, Society for Industrial and Applied Mathematics, 2008, pp.9-40.

[23] Al-Saleh M. F., Al-Kadiri M. A., "Double-ranked set sampling", Statistics & Probability Letters, vol.48, no.2, pp.205-212, 2000. DOI: 10.1016/S0167-7152(99)00206-0.

[24] Al-Saleh M. F., Al-Omari A. I., "Multistage ranked set sampling", Journal of Statistical Planning and Inference, vol.102, no.2, pp.273-286, 2002.DOI: 10.1016/S0378-3758(01)00086-6.

[25] Al-Saleh M. F., "Steady-state ranked set sampling and parametric estimation", Journal of statistical planning and inference, vol.123, no.1, pp.83-95, 2004. DOI: 10.1016/S0378-3758(03)00139-3.

[26] Nadarajah S., "Making the Cauchy work", Brazilian Journal of Probability and Statistics, vol.25, no.1, pp.99-120, 2011. DOI: 10.1214/09-BJPS112.