Asian Journal of Statistical Sciences
Vol. 5, No-2, (2025), pp. 179-189

ARF INDIA
Academic Open Access Publishing
Journal homepage link:
http://arfjournals.com/ajss

Estimating the location and scale parameters of a symmetric distribution by *U*-statistics with kernels as the best linear unbiased estimators based on systematic statistics

N.V. Sreekumar[a] and Dhanya Nair R.[b]

[a]*Department of Statistics, Government College for Women, Thiruvananthapuram - 695 014, India*
[b]*Department of Statistics, Government Arts College, Thiruvananthapuram - 695 014, India.*

ARTICLE HISTORY
Compiled July 13, 2026

Received: 17 June 2025; Accepted: 07 July 2026

ABSTRACT
In this paper, we estimate the location parameter (μ) and scale parameter (σ) of a distribution symmetric about μ using U-statistics with kernels as the BLUE's of μ and σ based on systematic statistics, viz., quasi-midrange and quasi-range of the sample. We also provide explicit expressions for the variance of these U-statistics. The advantage of the estimators obtained in this paper is that the order of the covariance matrix involved in expressions for the variance of the U-statistics is less, thus making the computation of the variance of U-statistics much easier.

KEYWORDS
best linear unbiased estimators; location and scale parameters; quasi-midrange; quasi-range; U-statistics

1. Introduction

Lloyd (1952) established the Best Linear Unbiased Estimator (BLUE) for both the location parameter (μ) and scale parameter (σ) of a distribution belonging to the location-scale family of distributions. This estimator is derived from the order statistics of a random sample drawn from the distribution. The exact expressions for their variances were also obtained. Further, if the parent distribution is symmetric about the location parameter, Thomas (1990) has shown that the Lloyd's BLUE's of the μ and σ based on order statistics reduce to those based on the quasi-midrange and quasi-range, respectively, of the sample, thereby achieving a considerable reduction in the order of the covariance matrix involved.

A major drawback of Lloyd's BLUE's of μ and σ is that in order to obtain these estimators, one requires the values of means, variances, and covariances of the entire order statistics of a random sample of size n arising from the distribution. However, for most of the distributions belonging to the location-scale family of distributions, explicit expressions do not exist for the means, variances, and covariances of these or-

Dhanya Nair R[b]. Email: rdhanyanair2019@gmail.com

der statistics. Consequently, numerical computation becomes necessary, and therefore these values are computed numerically for $n \leq 20$, and the BLUE's and their variances are obtained for sample sizes up to 20. But for large samples ($n > 20$), numerical computation is typically infeasible. To address this issue, Thomas and Sreekumar (2008) proposed a method for estimating μ and σ for large samples using U-statistics. These U-statistics utilise the BLUE's of μ and σ based on order statistics of an initial sample of small size as kernels. The U-statistics estimators thus obtained are highly preferred as they leverage the optimality conditions of BLUE as well as U-statistics.

Based on the principles elucidated in Thomas (1990) and Thomas and Sreekumar (2008), this study focuses on deriving the BLUE's of location and scale parameters of distributions symmetric about the location parameter. We achieve this by constructing U-statistics with kernel as the BLUE's of μ and σ based on the quasi-midrange and quasi-range of the sample. A notable advantage of this approach lies in the considerable reduction of the covariance matrix's order during the evaluation of the U-statistic's variance, thereby making the computation of the variance much easier.

Section 2 of this paper focuses on the formulation of U-statistics with kernels based on the quasi-midrange and quasi-range of the sample. The expressions for the U-statistics and their variances have also been derived. To demonstrate the efficacy of this method, we apply it to two specific members of the family of symmetric distributions: the U-shaped distribution and the Cauchy distribution, discussed in Sections 3 and 4, respectively.

2. U-statistics with kernels based on Quasi-midrange and Quasi-range

Consider a random variable Z with probability density function (pdf) given by

$$f(z;\mu,\sigma) = \frac{1}{\sigma} f_0\left(\frac{z-\mu}{\sigma}\right), -\infty < \mu < \infty, \sigma > 0$$

where the form of $f_0(\cdot)$ is known. The family of distributions with pdf as defined above is called the location-scale family of distributions with location parameter μ and scale parameter σ. Let $Z_{1:n} \leq Z_{2:n} \leq ... \leq Z_{n:n}$ be the order statistics of a sample of size n, drawn from an absolutely continuous distribution symmetric about the location parameter μ and with scale parameter σ. Then, $X_{r:n} = \frac{Z_{r:n}-\mu}{\sigma}$, for $r = 1, 2, ..., n$, are order statistics of a random sample of size n arising from a distribution that is free of μ and σ. Let $Z = (Z_{1:n}, Z_{2:n}, ..., Z_{n:n})^T$ and $X = (X_{1:n}, X_{2:n}, ..., X_{n:n})^T$. Let the expectation of the vector X and its covariance matrix $D(X)$ be given by $E(X) = a$ and $D(X) = A$. Then the Lloyd's BLUE of μ and σ along with their variances are given by

$$\mu^* = (1^T A^{-1} 1)^{-1} 1^T A^{-1} Z \tag{1}$$

$$Var(\mu^*) = (1^T A^{-1} 1)^{-1} \sigma^2 \tag{2}$$

$$\sigma^* = (a^T A^{-1} a)^{-1} a^T A^{-1} Z \tag{3}$$

$$Var(\sigma^*) = (a^T A^{-1} a)^{-1}\sigma^2 \tag{4}$$

Now, let $M_{i:n} = \frac{Z_{n-i+1:n}+Z_{i:n}}{2}$, for $i = 1, 2, ..., \left[\frac{n+1}{2}\right]$ denote the i^{th} quasi-midrange, and $R_{i:n} = Z_{n-i+1:n} - Z_{i:n}$, for $i = 1, 2, ..., \left[\frac{n}{2}\right]$ denote the i^{th} quasi-range of the sample. Then using the distributional symmetric properties we have, $E(M_{i:n}) = \mu$, for $i = 1, 2, ..., \left[\frac{n+1}{2}\right]$ and $E(R_{i:n}) = 2\, a_i\, \sigma$, for $i = 1, 2, ..., \left[\frac{n}{2}\right]$, where $a_i = E(X_{n-i+1:n})$.

Let $M = \left(M_{1:n}, M_{2:n}, ..., M_{\left[\frac{n+1}{2}\right]:n}\right)^T$ and $R = \left(R_{1:n}, R_{2:n}, ..., R_{\left[\frac{n}{2}\right]:n}\right)^T$. Additionally, let

1= $(1, 1, ..., 1)^T$ denote a column vector of $\left[\frac{n}{2}\right]$ ones. Let us denote a_{ij}=$cov(X_{i:n}, X_{j:n})$. Then,

$$\begin{aligned} v_{ij} &= cov(M_{i:n}, M_{j:n}) \\ &= 2^{-1}(a_{ij} + a_{i(n-j+1)})\sigma^2 \\ \text{and } w_{ij} &= cov(R_i, R_j) \\ &= 2\left(a_{ij} - a_{i(n-j+1)}\right)\sigma^2. \\ \text{Also, let } V &= [v_{ij}] \text{ where } i, j = 1, 2, ..., \left[\frac{n+1}{2}\right] \text{ and} \\ W &= [w_{ij}] \text{ where } i, j = 1, 2, ..., \left[\frac{n}{2}\right]. \end{aligned}$$

Thomas (1990) has shown that the Lloyd's BLUE of μ and σ of a symmetric distribution given by (1) and (3), respectively, reduces to that based on quasi-midranges and quasi-ranges of the sample. The BLUE's of μ and σ based on quasi-midranges and quasi-ranges along with their variances (see, Thomas, 1990), are given by

$$\mu^* = (1^1 V^{-1} 1)^{-1} 1^1 V^{-1} M = \sum_{i=1}^{\left[\frac{n+1}{2}\right]} c_i M_{i:n} \tag{5}$$

$$Var(\mu^*) = (1^1 V^{-1} 1)^{-1}\sigma^2 \tag{6}$$

$$\sigma^* = (b^1 W^{-1} b)^{-1} b^1 W^{-1} R = \sum_{i=1}^{\left[\frac{n}{2}\right]} d_i R_{i:n} \tag{7}$$

$$Var(\sigma^*) = (b^T W^{-1} b)^{-1}\sigma^2 \tag{8}$$

The advantage of using quasi-midranges to estimate μ and σ is that the order of covariance matrix involved in Equations (6) and (8) are only $\left[\frac{n+1}{2}\right]$ and $\left[\frac{n}{2}\right]$ respectively while the order of the covariance matrix A in the corresponding Lloyd's estimate is n.

In order to obtain Lloyd's BLUE's of μ and σ, one requires the means, variances, and covariances of the entire order statistics of a random sample of size n emerging from the distribution. Since for most of the distributions belonging to location-scale family, explicit expressions fail to exist for the means, variances, and covariances of the order statistics, their values are computed numerically for $n \leq 20$, and the BLUE's and their variances are obtained for sample sizes up to 20. However, numerical computation is impractical when the random sample size is greater than 20. To deal with such situations, Thomas and Sreekumar (2008) suggested the estimation of μ and σ by constructing the appropriate U-statistics with kernel based on Lloyd's BLUE of μ and σ, which in turn are based on the order statistics of an initial sample of small size $m,\ m \geq 2$.

Hoeffding (1948) introduced U-statistics and its utility in estimating the parameters of a distribution. Consider independent observations $Z_1, Z_2, ..., Z_n$ drawn from a population with cumulative distribution function F. Then the U-statistic for estimating the parameter θ, with the symmetric kernel $h(\cdot)$ of degree m, is given by:

$$U(Z_1, Z_2, ..., Z_n) = \frac{1}{\binom{n}{m}} \sum_{\beta \epsilon B} h(Z_{\beta_1}, Z_{\beta_2}, ..., Z_{\beta_m})$$

where, $B = \{\beta/\beta = (\beta_1, \beta_2, ..., \beta_m),\ \beta_1 < \beta_2 < ... < \beta_m,$ is one of the $\binom{n}{m}$ combinations of m integers chosen without replacement from the set $(1, 2, ..., n)\}$.
Suppose $E[h^2(Z_1, Z_2, ..., Z_m)] < \infty$. Let $h(Z_1, Z_2, ..., Z_c, Z_{c+1}, ..., Z_m)$ and $h(Z_1, Z_2, ..., Z_c, Z_{m+1}, ..., Z_{2m-c})$ be two random variables having exactly c sample observations in common, where $c = 1, 2, ..., m$. Let $\xi_c^{(m)}$ be the covariance between these two random variables. Then Hoeffding (1948) derived the variance of the U-statistic as:

$$Var\left[U(Z_1, Z_2, ..., Z_n)\right] = \frac{1}{\binom{n}{m}} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \xi_c^{(m)}$$

Clearly, the U-statistic is an unbiased estimator of θ.

Using the BLUE's of μ and σ given by equations (1) and (3), based on a small sample, as kernels of degree m, Thomas and Sreekumar (2008) constructed the U-statistics that estimate the parameters explicitly. They also proved the unbiasedness of the proposed estimators. The approach is highly advantageous because inorder to obtain the value of the variances of the U-statistics for any sample size n, evaluation of means, variances, and covariances of order statistics of sample sizes up to $2m-1$ arising from the standard form of the parent distribution alone are necessary . Thus, the usual mathematical intractability in obtaining the variance of the U-statistics based on order statistics is avoided.

Combining the concepts used in Thomas (1990) and Thomas and Sreekumar (2008), in this work, we estimate the location and scale parameters of a distribution symmetric about the location parameter by constructing U-statistics with kernel as the BLUE's of μ and σ based on the quasi-midrange and quasi-range of the sample. Utilizing the BLUE's of μ and σ given by equations (5) and (7) as kernels of degree m, we can construct appropriate U-statistics as follows.
The BLUE of μ can be expressed as:

$$h_1(Z_1, Z_2, ..., Z_m) = \sum_{i=1}^{\left[\frac{m+1}{2}\right]} c_i M_{i:m} \tag{9}$$

and that of σ can be written as:

$$h_2(Z_1, Z_2, ..., Z_m) = \sum_{i=1}^{\left[\frac{m}{2}\right]} d_i R_{i:m} \tag{10}$$

where $c_1, c_2, \ldots, c_{\left[\frac{m+1}{2}\right]}$ and $d_1, d_2, \ldots, d_{\left[\frac{m}{2}\right]}$ are constants.
Then, the U-statistic for estimating μ based on kernel (9) is obtained as:

$$U_{1:n}^{(m)} = \frac{1}{\binom{n}{m}} \sum_{k=1}^{\left[\frac{m+1}{2}\right]} \left(\sum_{i=k}^{n-(m-k)} \binom{i-1}{k-1} \binom{n-i}{m-k} M_{i:n} \right) c_k \tag{11}$$

The combinatorial coefficients of $M_{i:n}$ in the second sum is obtained by the principle of mathematical induction. Now, using the fact that $E(M_{i:n}) = \mu$, for $i = 1, 2, ..., \left[\frac{n+1}{2}\right]$, and $\sum_{i=k}^{n-(m-k)} \binom{i-1}{k-1}\binom{n-i}{m-k} = \binom{n}{m}$, we obtain that $U_{1:n}^{(m)}$ given by (11) is an unbiased estimator of μ. Similarly, the U-statistic for estimating σ based on kernel (10) is obtained as:

$$U_{2:n}^{(m)} = \frac{1}{\binom{n}{m}} \sum_{k=1}^{\left[\frac{m}{2}\right]} \left(\sum_{i=k}^{n-(m-k)} \binom{i-1}{k-1} \binom{n-i}{m-k} R_{i:n} \right) d_k, \tag{12}$$

which is an unbiased estimator of σ.
The variances of $U_{1:n}^{(m)}$ and $U_{2:n}^{(m)}$ are given by

$$Var\left[U_{1:n}^{(m)}\right] = \frac{1}{\binom{n}{m}} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \varepsilon_c^{(m)} \tag{13}$$

and

$$Var\left[U_{2:n}^{(m)}\right] = \frac{1}{\binom{n}{m}} \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \zeta_c^{(m)}, \tag{14}$$

where,

$$\varepsilon_c^{(m)} = cov\left[h_1(Z_1, Z_2, ..., Z_c, Z_{c+1}, ..., Z_m), h_1(Z_1, Z_2, ..., Z_c, Z_{m+1}, ..., Z_{2m-c})\right],$$

$$\zeta_c^{(m)} = cov\left[h_2(Z_1, Z_2, ..., Z_c, Z_{c+1}, ..., Z_m), h_2(Z_1, Z_2, ..., Z_c, Z_{m+1}, ..., Z_{2m-c})\right],$$

for $c = 1, 2, ..., m$.
Clearly, $\varepsilon_m^{(m)} = Var[h_1(Z_1, Z_2, ..., Z_m)]$ and $\zeta_m^{(m)} = Var[h_2(Z_1, Z_2, ..., Z_m)]$ and are given by equations (6) and (8) respectively.
To determine $\varepsilon_c^{(m)}$, for $c = 1, 2, ..., m-1$, we proceed as follows.
For $p = 1, 2, ..., m-1$, using equation (13) we have

$$Var\left[U_{1:m+p}^{(m)}\right] = \frac{1}{\binom{m+p}{m}} \sum_{c=m-p}^{m} \binom{m}{c} \binom{p}{m-c} \varepsilon_c^{(m)} \tag{15}$$

Returning to equation (11), if we write $m^* = \left[\frac{m+1}{2}\right]$, we have

$$\begin{aligned} U^{(m)}_{1:m+p} &= \frac{1}{\binom{m+p}{m}} \sum_{k=1}^{m^*} \sum_{i=1}^{m^*+p} \binom{i-1}{k-1}\binom{m+p-i}{m-k} M_{i:m+p} c_k, \\ &= \frac{1}{\binom{m+p}{m}} \sum_{i=1}^{m^*+p} \left[\sum_{k=1}^{m^*} \binom{i-1}{k-1}\binom{m+p-i}{m-k} c_k \right] M_{i:m+p}. \\ &= b^T_{m^*+p} M_{m^*+p} \end{aligned} \tag{16}$$

where the vector

$$b^T_{m^*+p} = \left(\sum_{k=1}^{m^*} \frac{\binom{1-1}{k-1}\binom{m+p-1}{m-k}}{\binom{m+p}{m}} c_k, \sum_{k=1}^{m^*} \frac{\binom{2-1}{k-1}\binom{m+p-2}{m-k}}{\binom{m+p}{m}} c_k, ..., \sum_{k=1}^{m^*} \frac{\binom{m^*+p-1}{k-1}\binom{m-m^*}{m-k}}{\binom{m+p}{m}} c_k \right) \tag{17}$$

and

$$M_{m^*+p} = (M_{1:m+p}, M_{2:m+p}, ..., M_{m^*+p:m+p}). \tag{18}$$

Hence, from (16), we have

$$Var(U^{(m)}_{1:m+p}) = (b^T_{m^*+p} V_{m^*+p} b_{m^*+p})\sigma^2, \text{ for } p = 1, 2, ..., (m-1) \tag{19}$$

where V_{m^*+P} is the coefficient of σ^2 involved in the variance co-variance matrix of the vector M_{m^*+p}.
From equations (15) and (19), it follows that for $p = 1, 2, ..., m-1$,

$$\sum_{c=m-p}^{m-1} \binom{m}{c}\binom{p}{m-c} \varepsilon^{(m)}_c = \binom{m+p}{m} \left(b^T_{m^*+p} V_{m^*+p} b_{m^*+p}\right) \sigma^2 - \varepsilon^{(m)}_m.$$

If we write
$q_p = \binom{m+p}{m} \left(b^T_{m^*+p} V_{m^*+p} b_{m^*+p}\right) \sigma^2 - \varepsilon^{(m)}_m$, for $p = 1, 2, ...(m-1)$, then the above equation can be represented in matrix form as

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & \binom{m}{m-1}\binom{1}{1} \\ 0 & 0 & \cdots & \binom{m}{m-2}\binom{2}{2} & \binom{m}{m-1}\binom{2}{1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \binom{m}{1}\binom{m-1}{m-1} & \binom{m}{2}\binom{m-1}{m-2} & \cdots & \binom{m}{m-2}\binom{m-1}{2} & \binom{m}{m-1}\binom{m-1}{1} \end{pmatrix} \begin{pmatrix} \varepsilon^{(m)}_1 \\ \varepsilon^{(m)}_2 \\ \cdots \\ \cdots \\ \varepsilon^{(m)}_{m-1} \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ \cdots \\ \cdots \\ q_{m-1} \end{pmatrix}. \tag{20}$$

If we denote the matrix and the column vector on the left side of (20) by H and E, respectively, and the column vector on the right side of (20) by Q, then the equation (20) may be re-written as $H.E = Q$. Subsequently, the values of $\varepsilon^{(m)}_c$, for $c = 1, 2, ..., m-1$

can be determined from the equation

$$E = H^{-1}Q.$$

Similarly the values of $\zeta_c^{(m)}, c = 1, 2, ..., (m-1)$ can be computed from the equation

$$\begin{aligned} C &= H^{-1}J, \\ \text{where } C &= \left(\zeta_1^{(m)}, \quad \zeta_2^{(m)} \quad \dots \quad \zeta_{m-1}^{(m)}\right)^T \\ J &= (J_1, J_2, ..., J_{m-1})^T \text{ with} \\ J_p &= \binom{m+p}{m}\left(g_{\hat{m}+p}^T W_{\hat{m}+p} g_{\hat{m}+p}\right)\sigma^2 - \zeta_m^{(m)} \text{ for } p = 1, 2, 3, ..., (m-1). \end{aligned} \tag{21}$$

Here $\hat{m} = \left[\frac{m}{2}\right]$,

$$g_{\hat{m}+p}^T = \left(\sum_{k=1}^{\hat{m}} \frac{\binom{1-1}{k-1}\binom{m+p-1}{m-k}}{\binom{m+p}{m}} d_k, \sum_{k=1}^{\hat{m}} \frac{\binom{2-1}{k-1}\binom{m+p-2}{m-k}}{\binom{m+p}{m}} d_k, ..., \sum_{k=1}^{\hat{m}} \frac{\binom{\hat{m}+p-1}{k-1}\binom{m-\hat{m}}{m-k}}{\binom{m+p}{m}} d_k\right) \tag{22}$$

and W_{m^*+p} is the coefficient of σ^2 in the variance covariance matrix of the vector $R_{\hat{m}+p}$ where $R_{\hat{m}+p} = (R_{1:m+p}, R_{2:m+p}, ..., R_{\hat{m}+p:m+p})$.

Once the values of $\varepsilon_c^{(m)}$ and $\zeta_c^{(m)}$ for $c = 1, 2..., (m-1)$ are determined, the exact variances of the U-statistics for estimating μ and σ based on any sample size n can be obtained from (13) and (14). An advantage of employing this method is that the maximum order of the covariance matrix involved in equation (20) and (21) is $m^* + m - 1$ and $\hat{m} + m - 1$, respectively. Hence, in order to obtain the explicit variance of the U-statistics $U_{1:n}^{(m)}$ and $U_{2:n}^{(m)}$ for any $n \geq 2m$, one only needs to evaluate the covariances of the order statistics of sample sizes up to $m^* + m - 1$ and $\hat{m} + m - 1$, respectively. Moreover, the range of k in the summation considered in (17) and (22), is from 1 to m^* and 1 to $\hat{m}$, respectively, whereas in the corresponding equations used for determining the variance of the U-Statistic using Lloyd's BLUE based on order statistics as the kernel, the maximum order of the covariance matrix is $2m - 1$, and the range of k in the sum is from 1 to m. Thus, from a computational perspective, this method proves to be much easier for symmetric distributions.

3. Estimation of location and scale parameters of U-shaped distribution by U-statistics

The probability density of U-shaped distribution is given by

$$f(x) = \begin{cases} \frac{3(x-\mu)^2}{2\sigma^3}, \mu - \sigma \leq x \leq \mu + \sigma; \mu \in R \text{ and } \sigma > 0 \\ 0, \text{otherwise} \end{cases} \tag{23}$$

Samuel and Thomas (1996) have computed the means, variances, and covariances of the order statistics arising from the standard form of the distribution defined in (23).

Additionally, they estimated the BLUE of the location and scale parameters involved in (23) utilizing quasi midranges and quasi ranges. The coefficients of order statistics in the BLUE of μ and σ pertaining to the distribution (23) along with their variances have also been tabulated for sample sizes, $n = 2(1)20$. Now making use of the findings from Section 2, we can develop appropriate U-Statistics. Using the values provided in the tables given in Samuel and Thomas (1996), we have computed the values of $\varepsilon_c^{(m)}$ and $\zeta_c^{(m)}$ for $m = 2(1)5$, which are presented in Table 1. Using these values, we derived the variances of $U_{1:n}^{(m)}$ and $U_{2:n}^{(m)}$ for $n = 5(5)20(20)60, 100$ and $m = 2(1)5$, as displayed in Tables 2 and 3.
Furthermore, the values obtained for $\sigma^{-2}var\left(\bar{X}\right)$ are 0.20, 0.10, 0.67, 0.05, 0.03, 0.02 and 0.01 for $n = 5, 10, 15, 20, 40, 60$ and 100 respectively. Hence, we can conclude that the proposed estimator of μ is highly efficient than the sample mean. Now, for the scale parameter σ, explicit expression for the variance of moment estimator or MLE is not available in literature. However, the exact variances of the suggested U-statistics for estimating μ and σ based on any sample size n can be obtained from (13) and (14).

Table 1. Values of $\varepsilon_c^{(m)}$ and $\zeta_c^{(m)}$ for $c = 1, 2, 3, ..., m$ and $m = 2(1)5$.

m	c	$\sigma^{-2}\varepsilon_c^{(m)}$	$\sigma^{-2}\zeta_c^{(m)}$
2	1	0.150000	0.007576
	2	0.300000	0.633333
3	1	0.037642	0.003367
	2	0.075661	0.075421
	3	0.150078	0.216162
4	1	0.010050	0.001664
	2	0.020460	0.015997
	3	0.039838	0.043104
	4	0.076790	0.095504
5	1	0.002514	0.002021
	2	0.006200	0.004673
	3	0.011774	0.011414
	4	0.021852	0.023784
	5	0.040528	0.047021

Table 2. Variance of $U_{1:n}^{(m)}$

n	5	10	15	20	40	60	100
$\sigma^{-2}var\left(U_{1:n}^{(2)}\right)$	0.120000	0.060000	0.040000	0.030000	0.015000	0.010000	0.006000
$\sigma^{-2}var\left(U_{1:n}^{(3)}\right)$	0.071697	0.034253	0.022697	0.016988	0.008477	0.005649	0.003389
$\sigma^{-2}var\left(U_{1:n}^{(4)}\right)$	0.047228	0.017515	0.011146	0.008229	0.004050	0.002691	0.001611
$\sigma^{-2}var\left(U_{1:n}^{(5)}\right)$	0.040528	0.009710	0.005000	0.003800	0.001730	0.001120	0.000653

Table 3. Variance of $U_{2:n}^{(m)}$

n	5	10	15	20	40	60	100
$\sigma^{-2}var\left(U_{2:n}^{(2)}\right)$	0.067900	0.016800	0.007910	0.004770	0.001550	0.000854	0.000428
$\sigma^{-2}var\left(U_{2:n}^{(3)}\right)$	0.067879	0.016768	0.007908	0.004769	0.001550	0.000854	0.000428
$\sigma^{-2}var\left(U_{2:n}^{(4)}\right)$	0.053584	0.012900	0.006130	0.003740	0.001250	0.000701	0.000358
$\sigma^{-2}var\left(U_{2:n}^{(5)}\right)$	0.047021	0.009130	0.004700	0.003150	0.001380	0.000888	0.000520

4. Estimation of the Scale Parameter of the Cauchy Distribution

Cauchy distribution does not possesses the first two moments and as a result it is not possible to estimate its parameters by the method of moments. Moreover the exact variance of the maximum likelihood estimates are also not known for small samples drawn from this distribution. For these reasons the method of estimating the parameters of a Cauchy distribution by order statistics attracts special importance. Barnett (1966) has evaluated all existing means, variances, and covariances of order statistics drawn from a Cauchy distribution,

$$f(x,\mu,\sigma) = \frac{1}{\pi\sigma}\frac{1}{[1+(\frac{x-\mu}{\sigma})^2]}, -\infty < x < \infty,\ \mu \in R \text{ and } \sigma > 0$$

where μ and σ are the location and scale parameters, respectively.
Further, realising the fact that the estimate of the scale parameter of a Cauchy distribution is not seen published like the location parameter, Thomas (1990) concentrated on the estimator of scale parameter. Thomas (1990) estimated the BLUE of the scale parameter σ based on quasi-ranges and computed its variance for $n = 6(1)16(2)20$. Since the variances of the first two and last two order statistics of a Cauchy distribution do not exist, the BLUE of σ based on quasi-ranges is formulated as $\sum_{i=3}^{[\frac{n}{2}]} c_i R_{i:n}$, with the values of c_i determined for $n \leq 20$. Utilizing this BLUE as the kernel, we construct a U-statistic and is obtained as

$$U_{2:n}^{(m)} = \frac{1}{\binom{n}{m}} \sum_{k=3}^{\hat{m}} \sum_{i=k}^{n-(m-k)} \binom{i-1}{k-1}\binom{n-i}{m-k} R_{i:n} d_k$$

and its variance is given by

$$Var[U_{2:n}^{(m)}] = \frac{1}{\binom{n}{m}} \sum_{c=1}^{m} \binom{m}{c}\binom{n-m}{m-c} \zeta_c^{(m)}.$$

The values of $\zeta_c^{(m)}$ for $c = 1, 2, ..., m-1$ are obtained as

$$\begin{aligned} C &= H^{-1}J \\ \text{where } C &= \begin{pmatrix} \zeta_1^{(m)}, & \zeta_2^{(m)} & \dots & \zeta_{m-1}^{(m)} \end{pmatrix}^T \end{aligned} \tag{24}$$

$$\begin{aligned} J &= (J_1, J_2, ..., J_{m-1})^T \text{ with} \\ J_p &= \binom{m+p}{m}\left(g_{\hat{m}+p-2}^T W_{\hat{m}+p-2} g_{\hat{m}+p-2}\right)\sigma^2 - \zeta_m^{(m)},\ \text{ for } p = 1, 2, ..., m-1 \end{aligned}$$

Also $g_{\hat{m}+p-2}^T = \left\{ \sum_{k=3}^{\hat{m}} \frac{\binom{3-1}{k-1}\binom{m+p-3}{m-k}}{\binom{m+p}{m}} d_k, \sum_{k=3}^{\hat{m}} \frac{\binom{4-1}{k-1}\binom{m+p-4}{m-k}}{\binom{m+p}{m}} d_k, ..., \sum_{k=3}^{\hat{m}} \frac{\binom{\hat{m}+p-1}{k-1}\binom{m-\hat{m}}{m-k}}{\binom{m+p}{m}} d_k \right\}$

and $W_{\hat{m}+p-2}$ is the variance covariance matrix of the vector $R_{\hat{m}+p-2}$ where $R_{\hat{m}+p-2} = (R_{3:m+p}, R_{4:m+p}, ..., R_{\hat{m}+p:m+p})$.
The values of $\zeta_c^{(m)}$ for $c = 1, 2, ..., m-1;\ m = 6, 7, 8$ are provided in Table 4. Using these

values, we computed the variances of $U_{2:n}^{(m)}$ for $n = 10(5)20(20)60, 100$ and $m = 6, 7, 8$ and are given in Table 5. Kravchuk and Pollett (2012) have computed the asymptotic variance of MLE $\tilde{\sigma}$ of the scale parameter of the Cauchy distribution. The efficiency of $U_{2:n}^{(m)}$ relative to the MLE of scale parameter σ for $m = 6$ and $n = 10, 15$ and 20 is therefore evaluated and is given in Table 6. It can be observed that the relative efficiency increases as sample size increases.

Table 4. Values of $\zeta_c^{(m)}$ for c = 1,2,3,...,m and m=6(1)8.

m	c	$\sigma^{-2}\zeta_c^{(m)}$
6	1	0.561777
	2	0.152341
	3	0.284150
	4	0.496706
	5	0.877723
	6	1.755200
7	1	0.031820
	2	0.107058
	3	0.183643
	4	0.290977
	5	0.441143
	6	0.662512
	7	1.003200
8	1	0.077560
	2	0.081303
	3	0.126042
	4	0.190115
	5	0.269148
	6	0.370510
	7	0.502695
	8	0.684800

Table 5. Variance of $U_{2:n}^{(m)}$

n	10	15	20	40	60	100
$\sigma^{-2}var\left(U_{2:n}^{(6)}\right)$	0.440672	0.224803	0.149234	0.062040	0.038762	0.022034
$\sigma^{-2}var\left(U_{2:n}^{(7)}\right)$	0.440668	0.224310	0.149225	0.059211	0.035470	0.019163
$\sigma^{-2}var\left(U_{2:n}^{(8)}\right)$	0.424493	0.218198	0.145112	0.049755	0.023464	0.007706

Table 6. Relative efficiency of $U_{2:n}^{(m)}$ with respect to MLE $\tilde{\sigma}$

n	10	15	20
$\sigma^{-2}var\left(U_{2:n}^{(6)}\right)$	0.44	0.22	0.15
$\sigma^{-2}var\left(\tilde{\sigma}\right)$	0.20	0.13	0.10
$e\left(U_{2:n}^{(6)} \mid \tilde{\sigma}\right)$	0.4545	0.5909	0.6666

5. Conclusion

In this study, we have explored the estimation of location and scale parameters for symmetric distributions using U-statistics with kernels based on systematic statistics. We began by revisiting the foundational work of Lloyd (1952), Thomas (1990) and Thomas and Sreekumar (2008) regarding the Best Linear Unbiased Estimators

(BLUE) for location and scale parameters, specifically focusing on distributions symmetric about the location parameter.

By integrating the principles of U-statistics introduced by Hoeffding (1948) and the BLUE estimators, we devised novel approaches to estimate the parameters of symmetric distributions. Through meticulous derivations and computations, we derived explicit expressions for the variances of these estimators, simplifying the computation process significantly.

Our investigation encompassed applications to specific distributions, namely the U-shaped distribution and the Cauchy distribution. We utilized the insights from previous research, such as Barnett (1966), Thomas (1990), Samuel and Thomas (1996) to compute essential statistics and construct appropriate U-statistics for estimation.

Furthermore, we provided comprehensive tables detailing the computed values of parameters and their corresponding variances for various sample sizes and distributional settings. These tables serve as valuable references for practitioners and researchers alike, facilitating the practical application of our proposed methodologies.

In conclusion, our study underscores the efficacy of combining U-statistics with BLUE estimators for robust and efficient parameter estimation in symmetric distributions. The methodologies presented herein offer practical solutions for handling challenges associated with estimating parameters in distributions where explicit expressions for statistics are not readily available. Future research avenues may explore extensions of these techniques to other distributional families and investigate their performance under different sampling scenarios.

Acknowledgement(s)

The authors would like to express their gratitude to the editor-in-chief and reviewers for their valuable comments, which have considerably improved the quality of the article.

References

[1] Lloyd, E.H. (1952). Least-squares estimation of location and scale parameters using order statistics, Biometrika 39, 88-95.

[2] Thomas, P.Y. (1990). Estimating location and scale parameters of a symmetric distribution by systematic statistics, Journal of the Indian Society of Agricultural Statistics, 42(2), 250-256.

[3] Thomas, P.Y. and Sreekumar, N.V. (2008). Estimation of location and scale parameters of a distribution by U-statistics based on best linear functions of order statistics, Journal of Statistical Planning and Inference, 138(7), 2190-2200.

[4] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. The Annals of Mathematical Statistics. 19(3), 293-325.

[5] Samuel, P. and Thomas, P.Y. (1996). Estimation of location and scale parameters of U-shaped distribution, Journal of the Indian Society of Agricultural Statistics, 50(1), 75-94.

[6] Barnett, V.D. (1966). Order statistics estimators of the location of the Cauchy Distribution, Journal of the American Statistical Association, 61(316), 1205-1218.

[7] Kravchuk, O. Y. and Pollett, P. K. (2012). Hodges-Lehmann Scale Estimator for Cauchy Distribution. Communications in Statistics - Theory and Methods, 41(20), 3621–3632. https://doi.org/10.1080/03610926.2011.563016