Stochastic Models in Probability and Statistics
Vol. 1, No. 2, 2024, pp. 171–189. Research Paper

Varinaccuracy of ranked set sample

Manoj Chacko[†] **and Varghese George**[‡*]

[†]*Department of Statistics, University of Kerala, Thiruvanthapuram 695 581, India*
[‡]*Department of Statistics, University of Kerala, Thiruvanthapuram 695 581, India*
[‡]*Department of Statistics, St. Stephen's College, Pathanapuram, Kerala, 689 695, India*
Email(s): manojchacko02@gmail.com, varghesessc@gmail.com

Abstract. In this paper, Kerridge's inaccuracy measure and varinaccuracy of the ranked set sample (RSS) are considered. By deriving the expression for Kerridge's inaccuracy measure and varinaccuracy of the rth order statistic, the expression for Kerridge's inaccuracy measure and varinaccuracy of RSS are obtained. Kerridge's inaccuracy measure and varinaccuracy of moving ranked set sample are also obtained.

Keywords: Inaccuracy measure; Order statistics; Ranked set sampling; Varinaccuracy.

1 Introduction

McIntyre (1952) introduced a sampling scheme named ranked set sampling (RSS) as a process of improving the precision of the sample mean as an estimate of the population mean. In some situations, the measurement of the variable of interest is costly and/or time-consuming, but the ranking of variables related to the study variable can be easily done by a judgment method (see Chen et al. (2004)). The procedure of ranked set sampling involves randomly choosing n set of units, each of size n from a population, and then the units in each set are ranked using some inexpensive methods. Then from the first set of n units, select the unit that has the lowest rank. From the second set of n units, select the unit with the second lowest rank. The process continues until the unit with nth rank is selected in the nth set. Then make measurements on the variable of interest of the selected units. Let $X_{(i:n)}$ be the measurement made on the ith selected units, then $X_{(1:n)}, X_{(2:n)}, \ldots, X_{(n:n)}$ constitute the ranked set sample.

Many authors modified the McIntyre (1952) method of ranked set sampling. Al-Odat and Al-Saleh (2001) introduced a modified ranked set sampling scheme named moving extreme ranked set sampling for improving the efficiency of the estimate of the population mean. The procedure

*Corresponding author

Received: 29 August 2024 / Accepted: 26 October 2024
DOI: 10.22067/smps.2024.45946

 https://smps.um.ac.ir

of moving the extreme ranked set sampling method is as follows. Choose n sets of random samples of sizes $1,2,...,n$ respectively, from the population. Rank the units in each set using the judgment method or some inexpensive methods, without making the actual measurement of the variable of interest. Select the unit with rank one from each set and then take actual measurement of selected units, we get a moving lower extreme ranked set sample (MLERSS) and the method is known as moving lower extreme ranked set sampling. If we select the units with maximum rank in each set and then take actual measurements of the selected units, we get a moving upper extreme ranked set sample (MUERSS) and the method is known as moving upper extreme ranked set sampling.

Let X be a non negative random variable with probability density function $f(x)$. Shannon (1948) introduced a measure of uncertainty as an average level of information associated with the random variable X, known as Shannon Entropy and it is defined as

$$H(X) = E_f\left[-\ln[f(X)]\right] = -\int_0^\infty f(x)\ln[f(x)]dx.$$

One of the generalization of Shannon entropy was done by Kerridge (1961). Let X and Y be two absolutely continuous non negative random variables with distribution functions F and G and probability density functions f and g, respectively. If F is the distribution function corresponding to the observations and G is the distribution assigned by the experimenter, then the inaccuracy measure of X and Y, proposed by Kerridge (1961) is given by

$$KI(f,g) = E_f\left[-\ln[g(X)]\right] = -\int_0^\infty f(x)\ln[g(x)]dx. \tag{1}$$

The inaccuracy measure defined in (1) can also be called as cross-entropy of Y on X or the relative distance between X and Y. In reliability analysis, Kerridge (1961) notion of inaccuracy delves into the concept by examining the potential errors in expressing probabilities related to different events in experiments. These errors can stem from two primary sources: one from missing data and another from misspecification of the model. Kerridge's inaccuracy measure is designed to address both types of errors. Kayal and Sunoj (2017) studied a generalized Kerridge's inaccuracy measure for conditionally specified models.Taneja and Tuteja (1986) discussed about the weighted inaccuracy measure. Ghosh and Kundu (2018) discussed the conditional cumulative past version of Kerridge's inaccuracy measure. Taneja et al. (2009) presented the dynamic version of inaccuracy between two residual lifetime distributions. Here the information contents are averaged over a known distribution. Kerridge's inaccuracy measure has a significant role in regression analysis for the Akaike information criteria. Applications of Kerridge's measure in coding theory can be seen in Nath (1968).

Recently, researchers have been interested in studying the scatter of information contents of a random variable. The variability of the information component cannot be explained by inaccuracy, which is the expected information content of suitability. Buono et al. (2021) introduced a dispersion index between two random variables X and Y based on cross-entropy named varianaccuracy defined as

$$\begin{aligned} VarKI(f,g) &= Var_f\big[-\ln[g(X)]\big] \\ &= E_f[\ln^2 g(X)] - \Big[E_f\big[\ln g(X)\big]\Big]^2 \\ &= \int_0^\infty f(x)\ln^2 g(x)dx - \Big[KI(f,g)\Big]^2. \end{aligned}$$

The variance of information content have many applications in reliability study, estimation etc. Varinaccuracy measures the discrepancy incurred in choosing a reference distribution. Sharma and Kundu (2024) have discussed the residual and past varinaccuracy measures and its application. Balakrishnan et al. (2024) discussed the dispersion indices based on Kerridge inaccuracy measure.

Thapliyal and Taneja (2013, 2015) discussed the inaccuracy and residual inaccuracy of order statistics. Ahmadi (2021) explained some results based on Kerridge's inaccuracy measures of records. Goel et al. (2018) discussed Kerridge's inaccuracy measure for record statistics. Mohammed (2019) discussed the inaccuracy measure in concomitants of ordered random variables under Farlie-Gumbel-Morgenstern family. Mohammadi and Hashempour (2023) studied extropy based inaccuracy measure in order statistics. Chacko and George (2023, 2024) discussed the extropy properties of ranked set samples when sampling is not perfect. George and Chacko (2023) discussed the cumulative residual extropy properties of ranked set samples for Cambanis type bivariate distributions. None of the previous work studied the varinaccuracy based on ranked set samples. In this paper, we discuss Kerridge's inaccuracy measure and varrinacuracy for RSS and its modifications.

The paper is organized as follows. In section 2, Kerridge's inaccuracy measure and varinaccuracy between the distribution of the rth order statistic and parent distribution are explained. Kerridge's measure of inaccuracy of ranked set sample is explained in section 3. In section 4, we explain the varriancccuracy of ranked set sample with examples. Section 5 gives Kerridge's inaccuracy measure of moving extreme ranked set sample. Varinaccuracy of moving extreme ranked set sample is given in section 6. Finally, section 7 gives the concluding remarks.

2 Varinaccuracy between distribution of the rth order statistic and the parent distribution

Let $X_1, X_2, ... X_n$ be a random sample from a population with pdf $f(x)$ and cdf $F(x)$. If we arrange X_i's in ascending order of magnitude as $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$, then $X_{(r)}$ is called the rth order statistic of X_i's. Then, the pdf of rth order statistic is given by

$$f_{r:n}(x) = \frac{1}{B(r, n-r+1)}(F(x))^{r-1}(1-F(x))^{n-r}f(x), \tag{2}$$

where $B(a,b)$ is the beta function. Therefore Kerridge's inaccuracy measure between $X_{(r)}$ and X can be written as

$$\begin{aligned} KI(f_{r:n}, f) &= -\int_0^\infty f_{r:n}(x)\ln f(x)dx \\ &= E\big[-\ln f(X_{(r)})\big]. \end{aligned}$$

Also, the varinaccuracy between distribution of $X_{(r)}$ and X can be written as

$$VarKI(f_{r:n}, f) = Var\big[-\ln f(X_{(r)})\big].$$

Lemma 1. *Let $M_{r:n}(t)$ be the moment generating function of $\ln f(X_{(r)})$. Then, $M_{r:n}(t)$ is given by*

$$M_{r:n}(t) = E[f(F^{-1}(U))]^t,$$

where U follows a beta distribution with parameters r and $n-r+1$.

Proof. From (2), we have

$$\begin{aligned} M_{r:n}(t) &= E[e^{t\ln f(X_{(r)})}] \\ &= E[f^t(X_{(r)})] \\ &= \int_0^\infty \frac{1}{B(r,n-r+1)}(F(x))^{r-1}(1-F(x))^{n-r}f^{t+1}(x)dx \\ &= \int_0^1 \frac{1}{B(r,n-r+1)}u^{r-1}(1-u)^{n-r}f^t(F^{-1}(u))du \\ &= E[f^t(F^{-1}(U))], \end{aligned}$$

where U follows beta distribution with parameters r and $n-r+1$. □

Theorem 1. *Let $K_{r:n}(t)$ be the cumulant generating function of $\ln f(X_{(r)})$. Then, $KI(f_{r:n}, f) = -K'_{r:n}(0)$ and $VarKI(f_{r:n}, f) = K''_{r:n}(0)$, where $K'_{r:n}(0)$ and $K''_{r:n}(0)$ are the first and second derivatives of $K_{r:n}(t)$ at $t=0$.*

Proof. We have $K_{r:n}(t) = \ln M_{r:n}(t)$. Therefore

$$K'_{r:n}(0) = E[\ln f(X_{(r)})], \quad \text{and} \quad K''_{r:n}(0) = Var[\ln f(X_{(r)})].$$

Also,

$$KI(f_{r:n}, f) = -K'_{r:n}(0) \quad \text{and} \quad VarKI(f_{r:n}, f) = K''_{r:n}(0).$$

Hence the theorem. □

Example 1. If X follows uniform distribution over (a,b), then $M_{r:n}(t) = \dfrac{1}{(b-a)^t}$. Therefore

$$KI(f_{r:n}, f) = -\ln(b-a) \quad \text{and} \quad VarKI(f_{r:n}, f) = 0.$$

Example 2. If X follows exponential distribution with pdf $f(x) = \theta e^{-\theta x}, x \geq 0, \theta > 0$, then

$$M_{r:n}(t) = \theta^{t-1}\frac{B(r,n-r+t)}{B(r,n-r+1)}.$$

Therefore

$$\begin{aligned} KI(f_{r:n}, f) &= -\Big[\ln\theta + \psi(n-r+1) - \psi(n+1)\Big] \\ &= -\ln\theta + \sum_{i=1}^{r}\frac{1}{n-i+1} \end{aligned}$$

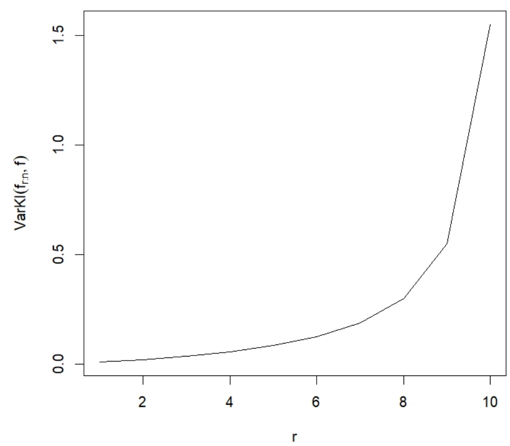

Figure 1: Graph of $VarKI(f_{r:n}, f)$ of exponential distribution.

and

$$VarKI(f_{r:n}, f) = \psi'(n-r+1) - \psi'(n+1)$$
$$= \sum_{i=1}^{r} \frac{1}{(n-i+1)^2},$$

where $\psi(.)$ is the digamma function and $\psi'(.)$ is the trigamma function.

Clearly $VarKI(f_{r:n}, f)$ of exponential distribution is free from parameter. We have drawn the graph of $VarKI(f_{r:n}, f)$ for exponential distribution when $n = 10$ and is given in Figure 1.

Example 3. If X follows Pareto distribution with pdf $f(x) = \frac{\lambda\beta^\lambda}{x^{\lambda+1}}, x \geq \beta, \lambda > 0, \beta > 0$, then

$$M_{r:n}(t) = \frac{\lambda^{t-1}}{\beta^{t-1}} \frac{B(r, n-r+(t-1)(1+\frac{1}{\lambda})+1)}{B(r, n-r+1)}.$$

Therefore

$$KI(f_{r:n}, f) = -\Big[\ln\lambda - \ln\beta + (1+\frac{1}{\lambda})\psi(n-r+1) - (1+\frac{1}{\lambda})\psi(n+1)\Big]$$
$$= \ln\beta - \ln\lambda + (1+\frac{1}{\lambda})\sum_{i=1}^{r} \frac{1}{n-i+1}$$

and

$$VarKI(f_{r:n}, f) = (1+\frac{1}{\lambda})^2\psi'(n-r+1) - (1+\frac{1}{\lambda})^2\psi'(n+1)$$
$$= (1+\frac{1}{\lambda})^2\sum_{i=1}^{r} \frac{1}{(n-i+1)^2}.$$

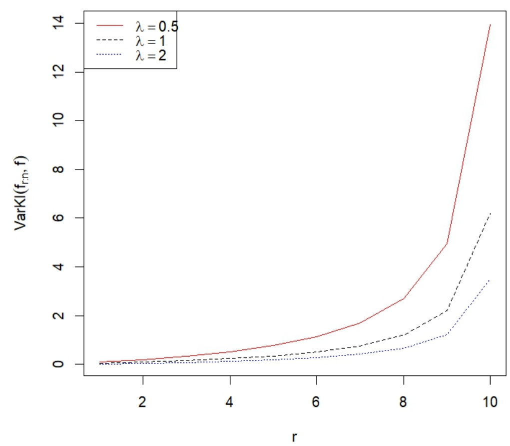

Figure 2: Graph of $VarKI(f_{r:n}, f)$ of Pareto distribution.

We have drawn the graph of $VarKI(f_{r:n}, f)$ for Pareto distribution for different values of λ when $n = 10$ and is given in Figure 2.

Example 4. If X follows standard power distribution with pdf $f(x) = \eta x^{\eta-1}, 0 < x < 1, \eta > 0$, then

$$M_{r:n}(t) = \eta^{t-1}\frac{B(r, n-r+(t-1)(1-\frac{1}{\eta})+1)}{B(r, n-r+1)}.$$

Therefore

$$\begin{aligned} KI(f_{r:n}, f) &= -\Big[\ln\eta + (1-\frac{1}{\eta})\psi(n-r+1) - (1-\frac{1}{\eta})\psi(n+1)\Big] \\ &= -\ln\eta + (1-\frac{1}{\eta})\sum_{i=1}^{r}\frac{1}{n-i+1} \end{aligned}$$

and

$$\begin{aligned} VarKI(f_{r:n}, f) &= (1-\frac{1}{\eta})^2\psi'(n-r+1) - (1-\frac{1}{\eta})^2\psi'(n+1) \\ &= (1-\frac{1}{\eta})^2\sum_{i=1}^{r}\frac{1}{(n-i+1)^2}. \end{aligned}$$

We have drawn the graph of $VarKI(f_{r:n}, f)$ for power distribution for different values of η when $n = 10$ and is given in Figure 3.

Remark 1. *From Table 1, it is observed that if X, Y and Z follow standard power distribution with pdf $f(x) = \eta x^{\eta-1}, 0 < x < 1, \eta > 0$, exponential distribution with pdf $g(y) = \theta e^{-\theta y}, y \geq 0, \theta > 0$ and Pareto distribution with pdf $h(z) = \frac{\lambda\beta^{\lambda}}{z^{\lambda+1}}, \lambda > 0, \beta > 0, z \geq \beta$, respectively and if $1 + \frac{1}{\lambda} \geq |1 - \frac{1}{\eta}|$, then $VarKI(f_{r:n}, f) \leq VarKI(g_{r:n}, g) \leq VarKI(h_{r:n}, h)$.*

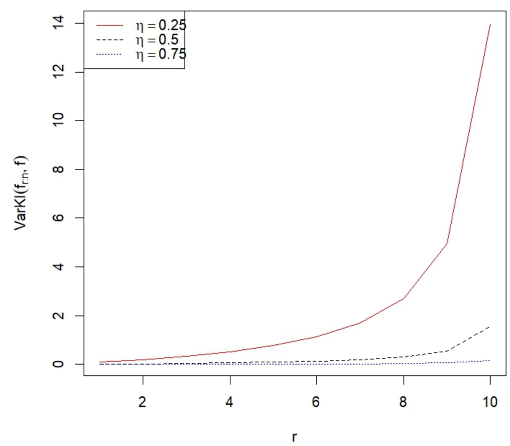

Figure 3: Graph of $VarKI(f_{r:n}, f)$ of standard power distribution.

Table 1: Expressions for $KI(f_{r:n}, f)$ and $VarKI(f_{r:n}, f)$.

Distribution	pdf	$KI(f_{r:n}, f)$	$VarKI(f_{r:n}, f)$
Uniform	$\frac{1}{b-a},\ a < x < b$	$-\ln(b-a)$	0
Exponential	$\theta e^{-\theta x}, x \geq 0, \theta > 0$	$-\ln\theta + \sum_{i=1}^{r} \frac{1}{n-i+1}$	$\sum_{i=1}^{r} \frac{1}{(n-i+1)^2}$
Pareto	$\frac{\lambda\beta^{\lambda}}{x^{\lambda+1}}, x \geq \beta, \lambda > 0, \beta > 0$	$\ln\beta - \ln\lambda + (1+\frac{1}{\lambda})\sum_{i=1}^{r} \frac{1}{n-i+1}$	$(1+\frac{1}{\lambda})^2 \sum_{i=1}^{r} \frac{1}{(n-i+1)^2}$
Standard power	$\eta x^{\eta-1}, 0 < x < 1, \eta > 0$	$-\ln\eta + (1-\frac{1}{\eta})\sum_{i=1}^{r} \frac{1}{n-i+1}$	$(1-\frac{1}{\eta})^2 \sum_{i=1}^{r} \frac{1}{(n-i+1)^2}$

So if $1 + \frac{1}{\lambda} \geq \mid 1 - \frac{1}{\eta} \mid$, discrepancy incurred in choosing the rth order statistic of power distribution is less compared to those of exponential distribution and Pareto distribution.

Theorem 2. *Let X and Y be two absolutely continuous non negative random variables with pdf $f_X(x)$ and $g_X(x)$, then*

$$VarKI(f,g) \leq \left(\int (KI(f,g) + \ln f_X(x))^4 dx \right)^{\frac{1}{2}} \left(E(f_X(X)) \right)^{\frac{1}{2}},$$

where $KI(f,g)$ is the Kerridge's inaccuracy measure of X and Y.
The equality is attained by

$$\left(\frac{KI(f,g) + \ln f_X(x)}{f_X(x)} \right)^2 E(f_X(X)) = \int \left[KI(f,g) + \ln f_X(x) \right]^2 dx.$$

Proof. We have

$$\begin{aligned} VarKI(f,g) &= E_g\left[-\ln f_X(X) - E_g(-\ln f_X(X))\right]^2 \\ &= E\big[-\ln f_X(X) - KI(f,g)\big]^2 \\ &= \int (KI(f,g) + \ln f_X(x))^2 f_X(x)dx. \end{aligned}$$

Then, by Cauchy- Schwartz inequality,

$$\begin{aligned} VarKI(f,g) &\leq \left(\int (KI(f,g) + \ln f_X(x))^4 dx\right)^{\frac{1}{2}} \left(\int (f(x))^2 dx\right)^{\frac{1}{2}} \\ &= \left(\int (KI(f,g) + \ln f_X(x))^4 dx\right)^{\frac{1}{2}} \Big(E(f_X(X))\Big)^{\frac{1}{2}}. \end{aligned}$$

The equality is attained if and only if there is a constant $\alpha \geq 0$ such that

$$(KI(f,g) + \ln f_X(x))^2 = \alpha (f_X(x))^2.$$

The constant α can be evaluated as

$$\alpha = \frac{\int \left[KI(f,g) + \ln f_X(x)\right]^2 dx}{E(f_X(X))}.$$

Hence the theorem. □

Remark 2. *Balakrishnan et al. (2024) showed that $VarKI(f,g)$ does not affect linear transformation. So clearly $VarKI(f_{r:n}, f)$ is independent of scale parameter. This can be evident from Examples 2 and 3 for exponential and Pareto distributions, respectively.*

3 Kerridge's inaccuracy measure of ranked set sample

Let $X_{(r:n)}, r = 1, 2, ..., n$ be a ranked set sample of size n from a population with pdf $f(x)$ and cdf $F(x)$. If the ranking is perfect, then $X_{(r:n)}$ is nothing but the rth order statistic of a random sample of size n.

Theorem 3. *If $X_{SRS} = \{X_1, X_2, ..., X_n\}$ and $X_{RSS} = \{X_{(1:n)}, X_{(2:n)}, ..., X_{(n:n)}\}$, where $X_1, X_2, ..., X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then the Kerridge's inaccuracy measure associated with ranked set sample and simple random sample can be written as*

$$KI(f_{RSS}, f_{SRS}) = \sum_{r=1}^{n} KI(f_{r:n}, f).$$

Table 2: Expression for $KI(f_{RSS}, f_{SRS})$.

Distribution	pdf	$KI(f_{RSS}, f_{SRS})$
Uniform	$\frac{1}{b-a}, a<x<b$	$-n\ln(b-a)$
Exponential	$\theta e^{-\theta x}, x\geq 0, \theta>0$	$n(1-\ln\theta)$
Pareto	$\frac{\lambda\beta^{\lambda}}{x^{\lambda+1}}, \lambda>0, \beta>0, x\geq\beta$	$n\left(1-\ln\lambda+\ln\beta+\frac{1}{\lambda}\right)$
Standard power	$\eta x^{\eta-1}, 0<x<1, \eta>0$	$n\left(1-\ln\eta-\frac{1}{\eta}\right)$

Proof. Let $f_{SRS}(\underline{x})$ and $f_{RSS}(\underline{x})$ be the pdfs of X_{SRS} and X_{RSS}, respectively, where $\underline{x} = (x_1, x_2, ..., x_n)$. Therefore, $f_{SRS}(\underline{x}) = \prod_{r=1}^{n} f(x_r)$ and $f_{RSS}(\underline{x}) = \prod_{r=1}^{n} f_{r:n}(x_r)$. Then

$$\begin{aligned} KI(f_{RSS}, f_{SRS}) &= -\int_0^{\infty}\int_0^{\infty}...\int_0^{\infty} f_{RSS}(\underline{x})\ln f_{SRS}(\underline{x})dx_1dx_2...dx_n \\ &= E\left[-\ln f_{SRS}(X_{RSS})\right] \\ &= \sum_{r=1}^{n} E\left[-\ln f(X_{(r:n)})\right] \\ &= \sum_{r=1}^{n} KI(f_{r:n}, f). \end{aligned}$$

Hence the theorem. ☐

Example 5. We obtained $KI(f_{RSS}, f_{SRS})$ for uniform, exponential, Pareto and standard power distributions with the pdfs given as in Table 1, the results are summarized in Table 2.

4 Measure of Varinaccuracy of ranked set sample

Theorem 4. *If $X_{SRS} = \{X_1, X_2, ..., X_n\}$ and $X_{RSS} = \{X_{(1:n)}, X_{(2:n)}, ..., X_{(n:n)}\}$, where $X_1, X_2, ..., X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then the varinaccuracy between ranked set sample and simple random sample can be written as*

$$VarKI(f_{RSS}, f_{SRS}) = \sum_{r=1}^{n} VarKI(f_{r:n}, f).$$

Proof. Let $f_{SRS}(\underline{x})$ and $f_{RSS}(\underline{x})$ be the pdfs of X_{SRS} and X_{RSS}, respectively, where $\underline{x} = (x_1, x_2, ..., x_n)$. Therefore, $f_{SRS}(\underline{x}) = \prod_{r=1}^{n} f(x_r)$ and $f_{RSS}(\underline{x}) = \prod_{r=1}^{n} f_{r:n}(x_r)$. Then

$$\begin{aligned} VarKI(f_{RSS}, f_{SRS}) &= Var\left[-\ln f_{SRS}(X_{RSS})\right] \\ &= Var\left[-\ln \prod_{r=1}^{n} f(X_{(r:n)})\right] \\ &= \sum_{r=1}^{n} Var\left[-\ln f(X_{(r:n)})\right] \\ &= \sum_{r=1}^{n} VarKI(f_{r:n}, f). \end{aligned}$$

Table 3: Expression for $VarKI(f_{RSS}, f_{SRS})$.

Distribution	pdf	$VarKI(f_{RSS}, f_{SRS})$
Uniform	$\frac{1}{b-a}, a < x < b$	0
Exponential	$\theta e^{-\theta x}, x \geq 0, \theta > 0$	$\sum_{i=1}^{n} \frac{1}{i}$
Pareto	$\frac{\lambda \beta^{\lambda}}{x^{\lambda+1}}, \lambda > 0, \beta > 0, x \geq \beta$	$(1+\frac{1}{\lambda})^2 \sum_{i=1}^{n} \frac{1}{i}$
Standard power	$\eta x^{\eta-1}, 0 < x < 1, \eta > 0$	$(1-\frac{1}{\eta})^2 \sum_{i=1}^{n} \frac{1}{i}$

□

Example 6. We obtained $VarKI(f_{RSS}, f_{SRS})$ for uniform, exponential, Pareto and standard power distributions with the pdfs given as in Table 1, the results are summarized in Table 3.

We have drawn the graph of $VarKI(f_{RSS}, f_{SRS})$ for exponential, Pareto and standard power distributions and displayed in Figures 4, 5 and 6, respectively.

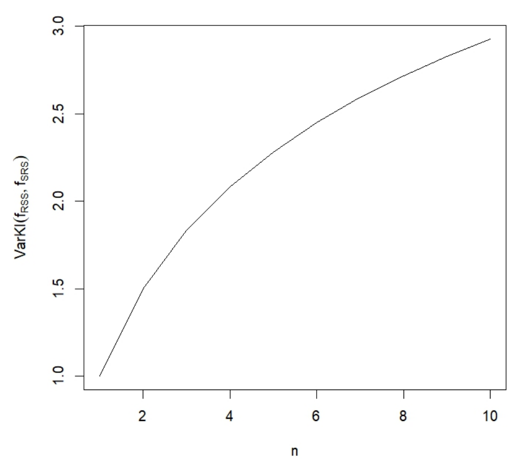

Figure 4: Graph of $VarKI(f_{RSS}, f_{SRS})$ for exponential distribution.

Remark 3. *From Table 3 and the same assumptions of Remark 1, we have*

$$VarKI(f_{RSS}, f_{SRS}) \leq VarKI(g_{RSS}, g_{SRS}) \leq VarKI(h_{RSS}, h_{SRS}).$$

So if $1+\frac{1}{\lambda} \geq \mid 1-\frac{1}{\eta} \mid$, discrepancy incurred in choosing RSS for power distribution is less compared to those of exponential distribution and Pareto distribution.

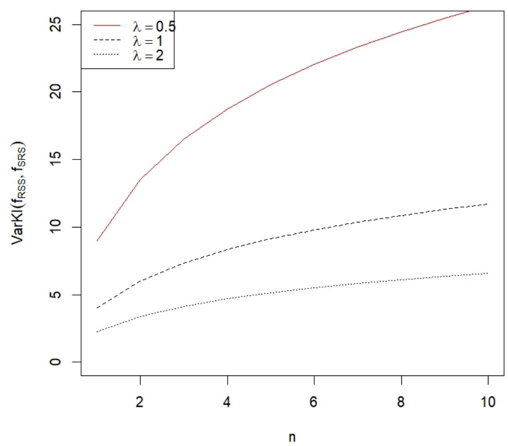

Figure 5: Graph of $VarKI(f_{RSS}, f_{SRS})$ for Pareto distribution.

5 Kerridge's inaccuracy measure of moving extreme ranked set sample

In this section, first we consider the Kerridge's inaccuracy measure of MLERSS. Let $X_{(1:j)}$, $j = 1,2,...n.$ be the measurement of the jth unit of MLERSS.

Theorem 5. *If $X_{SRS} = \{X_1, X_2, ..., X_n\}$ and $X_{MLERSS} = \{X_{(1:j)}, j = 1,2,...,n\}$, where $X_1, X_2, ..., X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then the Kerridge's inaccuracy measure between MLERSS and simple random sample can be written as*

$$KI(f_{MLERSS}, f_{SRS}) = \sum_{i=1}^{n} KI(f_{1:i}, f).$$

Proof. Let $f_{SRS}(\underline{x})$ and $f_{MLERSS}(\underline{x})$ be the pdfs of X_{SRS} and X_{MLERSS}, respectively, where $\underline{x} = (x_1, x_2, ..., x_n)$. Therefore, $f_{SRS}(\underline{x}) = \prod_{i=1}^{n} f(x_i)$ and $f_{MLERSS}(\underline{x}) = \prod_{i=1}^{n} f_{1:i}(x_i)$. Then

$$\begin{aligned} KI(f_{MLERSS}, f_{SRS}) &= -\int_0^{\infty}\int_0^{\infty}...\int_0^{\infty} f_{MLERSS}(\underline{x}) ln f_{SRS}(\underline{x}) dx_1 dx_2 ... dx_n \\ &= E\left[-\ln f_{SRS}(X_{MLERSS})\right] \\ &= \sum_{i=1}^{n} E\left[-\ln f(X_{(1:i)})\right] \\ &= \sum_{i=1}^{n} KI(f_{1:i}, f). \end{aligned}$$

Hence the theorem. □

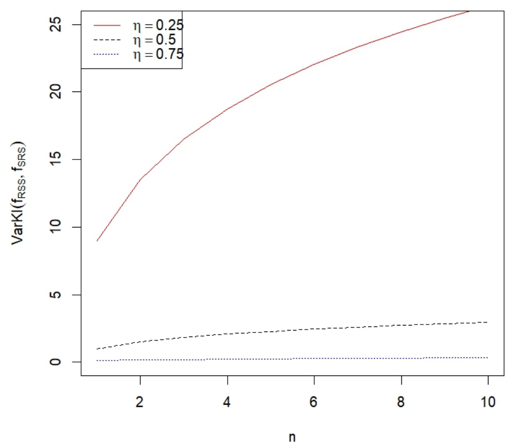

Figure 6: Graph of $VarKI(f_{RSS}, f_{SRS})$ for standard power distribution.

Next, we consider the Kerridge's inaccuracy measure of MUERSS. Let $X_{(j:j)}$, $j = 1, 2, ...n.$ be the measurement the jth unit of MUERSS.

Theorem 6. *Let $X_{SRS} = \{X_1, X_2, ..., X_n\}$ and $X_{MUERSS} = \{X_{(j:j)}, j = 1, 2, ..., n\}$ where $X_1, X_2, ..., X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then, the Kerridge's inaccuracy measure between MUERSS and simple random sample can be written as of*

$$KI(f_{MUERSS}, f_{SRS}) = \sum_{i=1}^{n} KI(f_{i:i}, f).$$

Proof. Let $f_{SRS}(\underline{x})$ and $f_{MUERSS}(\underline{x})$ be the pdfs of X_{SRS} and X_{MUERSS} respectively, where $\underline{x} = (x_1, x_2, ..., x_n)$. Therefore, $f_{SRS}(\underline{x}) = \prod_{i=1}^{n} f(x_i)$ and $f_{MUERSS}(\underline{x}) = \prod_{i=1}^{n} f_{i:i}(x_i)$. Then

$$\begin{aligned} KI(f_{MUERSS}, f_{SRS}) &= -\int_0^\infty \int_0^\infty ... \int_0^\infty f_{MUERSS}(\underline{x}) \ln f_{SRS}(\underline{x}) dx_1 dx_2 ... dx_n \\ &= E\left[-\ln f_{SRS}(X_{MUERSS})\right] \\ &= \sum_{i=1}^{n} E\left[-\ln f(X_{(i:i)})\right] \\ &= \sum_{i=1}^{n} KI(f_{i:i}, f). \end{aligned}$$

□

Example 7. We obtained $KI(f_{MLERSS}, f_{SRS})$ and $KI(f_{MUERSS}, f_{SRS})$ for uniform, exponential, Pareto and standard power distributions with the pdfs given as in Table 1, the results are summarized and presented in Table 4.

Table 4: Expression for $KI(f_{MLERSS}, f_{SRS})$ and $KI(f_{MUERSS}, f_{SRS})$.

Distribution	pdf	$KI(f_{MLERSS}, f_{SRS})$	$KI(f_{MUERSS}, f_{SRS})$
Uniform	$\frac{1}{b-a}, a < x < b$	$-n\ln(b-a)$	$-n\ln(b-a)$
Exponential	$\theta e^{-\theta x}, x \geq 0$	$-n\ln\theta + \sum_{i=1}^{n}\frac{1}{i}$	$(n+1)\sum_{i=1}^{n}\frac{1}{i} - n(1+\ln\theta)$
Pareto	$\frac{\lambda\beta^{\lambda}}{x^{\lambda+1}}, x \geq \beta$	$n\ln\beta - n\ln\lambda + (1+\frac{1}{\lambda})\sum_{i=1}^{n}\frac{1}{i}$	$n(\ln\beta - \ln\lambda) + (1+\frac{1}{\lambda})\sum_{i=1}^{n}\sum_{j=1}^{i}\frac{1}{j}$
Standard power	$\eta x^{\eta-1}, 0 < x < 1$	$-n\ln\eta + (1-\frac{1}{\eta})\sum_{i=1}^{n}\frac{1}{i}$	$-n\ln\eta + (1-\frac{1}{\eta})\sum_{i=1}^{n}\sum_{j=1}^{i}\frac{1}{j}$

6 Measure of Varinaccuracy of moving extreme ranked set sample

In this section, first we consider the varentropy of MLERSS. Let $X_{(1:j)}$, $j = 1, 2, \ldots n.$ be the measurement of the jth unit of MLERSS.

Theorem 7. *If $X_{SRS} = \{X_1, X_2, \ldots, X_n\}$ and $X_{MLERSS} = \{X_{(1:j)}, j = 1, 2, \ldots, n\}$, where $X_1, X_2, \ldots, X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then the varinaccuracy between MLERSS and simple random sample can be written as of*

$$VarKI(f_{MLERSS}, f_{SRS}) = \sum_{i=1}^{n} VarKI(f_{1:i}, f).$$

Proof. Let $f_{SRS}(\underline{x})$ and $f_{MLERSS}(\underline{x})$ be the pdfs of X_{SRS} and X_{MLERSS} respectively, where $\underline{x} = (x_1, x_2, \ldots, x_n)$. Therefore, $f_{SRS}(\underline{x}) = \prod_{i=1}^{n} f(x_i)$ and $f_{MLERSS}(\underline{x}) = \prod_{i=1}^{n} f_{i:i}(x_i)$. Then

$$\begin{aligned}
VarKI(f_{MLERSS}, f_{SRS}) &= Var\big[-\ln f_{SRS}(X_{MLERSS})\big] \\
&= Var\big[-\ln \prod_{r=1}^{n} f(X_{(1:i)})\big] \\
&= \sum_{i=1}^{n} Var\big[-\ln f(X_{(1:i)})\big] \\
&= \sum_{i=1}^{n} VarKI(f_{1:i}, f).
\end{aligned}$$

□

We have drawn the graph of $VarKI(f_{MLERSS}, f_{SRS})$ for exponential, Pareto and standard power distributions, is given in Figures 7, 8 and 9, respectively.

Next, we consider the varentropy of MUERSS. Let $X_{(j:j)}$, $j = 1, 2, \ldots n.$ be the measurement the jth unit of MUERSS.

Theorem 8. *Let $X_{SRS} = \{X_1, X_2, \ldots, X_n\}$ and $X_{MUERSS} = \{X_{(j:j)}, j = 1, 2, \ldots, n\}$ where $X_1, X_2, \ldots, X_n$ is a random sample from a population with pdf $f(x)$ and cdf $F(x)$, then, the varinaccuracy between MUERSS and simple random sample can be written as of*

$$VarKI(f_{MUERSS}, f_{SRS}) = \sum_{i=1}^{n} VarKI(f_{i:i}, f).$$

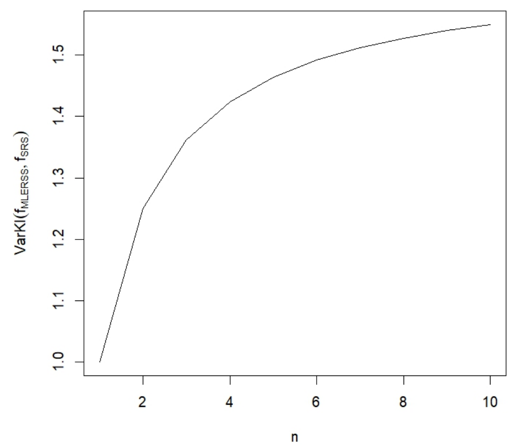

Figure 7: Graph of $VarKI(f_{MLERSS}, f_{SRS})$ for exponential distribution.

Proof. Let $f_{SRS}(\underset{\sim}{x})$ and $f_{MUERSS}(\underset{\sim}{x})$ be the pdfs of X_{SRS} and X_{MLERSS} respectively, where $\underset{\sim}{x} = (x_1, x_2, ..., x_n)$. Therefore, $f_{SRS}(\underset{\sim}{x}) = \prod_{i=1}^{n} f(x_i)$ and $f_{MUERSS}(\underset{\sim}{x}) = \prod_{i=1}^{n} f_{i:i}(x_i)$. Then,

$$\begin{aligned} VarKI(f_{MUERSS}, f_{SRS}) &= Var\big[-\ln f_{SRS}(X_{MUERSS})\big] \\ &= Var\Big[-\ln \prod_{i=1}^{n} f(X_{(i:i)})\Big] \\ &= \sum_{i=1}^{n} Var\big[-\ln f(X_{(i:i)})\big] \\ &= \sum_{i=1}^{n} VarKI(f_{i:i}, f). \end{aligned}$$

□

Example 8. We obtained $VarKI(f_{MLERSS}, f_{SRS})$ and $VarKI(f_{MUERSS}, f_{SRS})$ for uniform, exponential, Pareto and standard power distributions with the pdfs given as in Table 1, the results are summarized and presented in Table 5.

Remark 4. *From Table 5 and the same assumptions of Remark 1, we have*

$$VarKI(f_{MLERSS}, f_{SRS}) \leq VarKI(g_{MLERSS}, g_{SRS}) \leq VarKI(h_{MLERSS}, h_{SRS}),$$

and

$$VarKI(f_{MUERSS}, f_{SRS}) \leq VarKI(g_{MUERSS}, g_{SRS}) \leq VarKI(h_{MUERSS}, h_{SRS}).$$

Figures 10, 11 and 12 show the graph of $VarKI(f_{MUERSS}, f_{SRS})$ for exponential, Pareto and standard power distributions with the pdfs given as in Table 1, respectively.

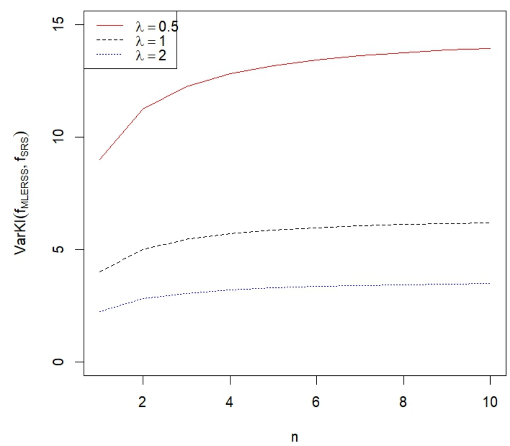

Figure 8: Graph of $VarKI(f_{MLERSS}, f_{SRS})$ for Pareto distribution.

Table 5: Expression for $VarKI(f_{MLERSS}, f_{SRS})$ and $VarKI(f_{MUERSS}, f_{SRS})$

Distribution	pdf	$VarKI(f_{MLERSS}, f_{SRS})$	$VarKI(f_{MUERSS}, f_{SRS})$
Uniform	$\frac{1}{b-a}, a < x < b$	0	0
Exponential	$\theta e^{-\theta x}, x \geq 0, \theta > 0$	$\sum_{i=1}^{n} \frac{1}{i^2}$	$\sum_{i=1}^{n} \frac{i}{(n-i+1)^2}$
Pareto	$\frac{\lambda \beta^{\lambda}}{x^{\lambda+1}}, \lambda > 0, \beta > 0, x \geq \beta$	$(1+\frac{1}{\lambda})^2 \sum_{i=1}^{n} \frac{1}{i^2}$	$(1+\frac{1}{\lambda})^2 \sum_{i=1}^{n} \frac{i}{(n-i+1)^2}$
Standard power	$\eta x^{\eta-1}, 0 < x < 1, \eta > 0$	$(1-\frac{1}{\eta})^2 \sum_{i=1}^{n} \frac{1}{i^2}$	$(1-\frac{1}{\eta})^2 \sum_{i=1}^{n} \frac{i}{(n-i+1)^2}$

7 Conclusion

In this work, we considered Kerridge's inaccuracy measure and varinaccuracy of ranked set sample and its different modifications. If we consider a ranked set sampling when ranking is not perfect, the measurement of ith unit of the ranked set sample is nothing but the ith order statistic, $X_{(i)}$ of the random sample. In this paper, we derived Kerridge's inaccuracy measure between $X_{(i)}$ and X and obtained its varinaccuracy. The expression for varinaccuracy between $X_{(i)}$ and X for the exponential, standard power, and Pareto distributions are derived. We also obtained an upper bound to varinaccuracy.

It is seen that Kerridge's inaccuracy measure between the ranked set sample and simple random sample can be written as the sum of Kerridge's inaccuracy measure between $X_{(i)}$ and X. Similarly, varinaccuracy between a ranked set sample and a simple random sample can be written as the sum of varinaccuracy between $X_{(i)}$ and X. We also obtained the expression for Kerridge's inaccuracy measure and varinacccuracy based on moving ranked set samples. The expression for Kerridge's inaccuracy measure and varinaccuracy of exponential, standard power, and Pareto distributions based on ranked set sample and moving extreme ranked set sample are obtained. It is observed that if $1+\frac{1}{\lambda} \geq \mid 1-\frac{1}{\eta} \mid$, the discrepancy incurred in choosing RSS and

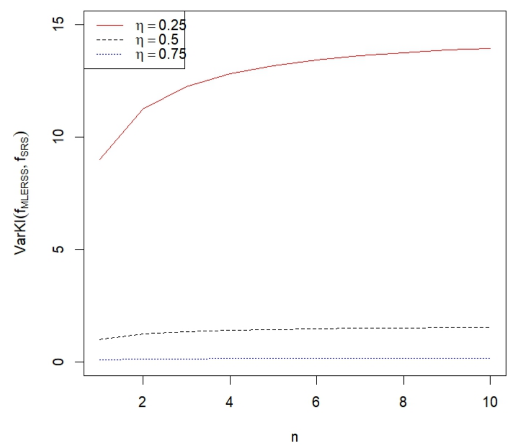

Figure 9: Graph of $VarKI(f_{MLERSS}, f_{SRS})$ for standard power distribution.

moving extreme ranked set sample for power distribution is less than those of exponential and Pareto distribution.

Acknowledgement

The authors are thankful to the anonymous referees for their valuable comments and suggestions which lead to an improved version of the manuscript.

References

Ahmadi, J. (2021). Characterization of continuous symmetric distributions using information measures of records. *Statistical Papers*, **62**, 2603–2626.

Al-Odat, M. T., and Al-Saleh, M. F. (2001). A variation of ranked set sampling. *Journal of Applied Statistical Science*, **10**, 137–146.

Balakrishnan, N., Buono, F., Calì, C., and Longobardi, M. (2024). Dispersion indices based on Kerridge inaccuracy measure and Kullback-Leibler divergence. *Communications in Statistics-Theory and Methods*, **53**, 5574–5592.

Buono, F., Calì, C., and Longobardi, M. (2021). Dispersion indices based on Kerridge inaccuracy and Kullback-Leibler divergence. *arXiv preprint arXiv*:2106. 12292.

Chacko, M., and George, V. (2023). Extropy properties of ranked set sample for Cambanis type bivariate distributions. *Journal of the Indian Society for Probability and Statistics,* **24**, 111–133.

]

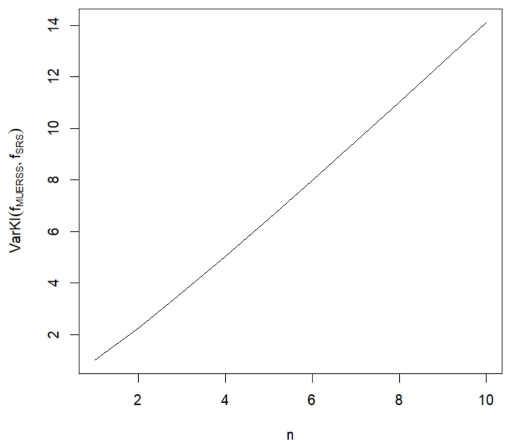

Figure 10: Graph of $VarKI(f_{MUERSS}, f_{SRS})$ for exponential distribution.

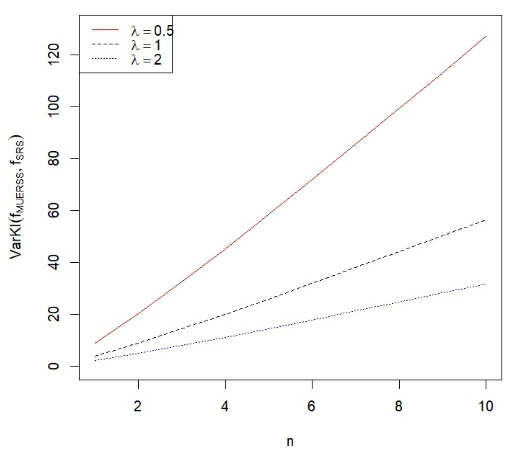

Figure 11: Graph of $VarKI(f_{MUERSS}, f_{SRS})$ for Pareto distribution.

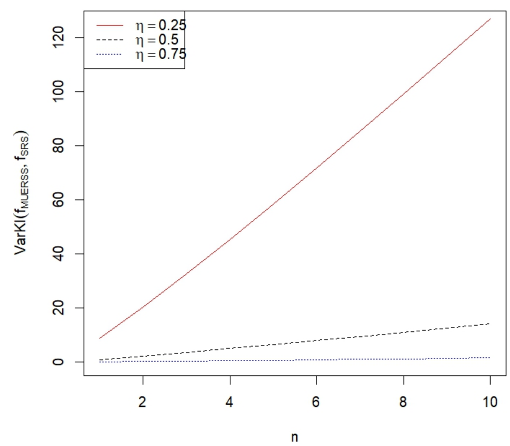

Figure 12: Graph of $VarKI(f_{MUERSS}, f_{SRS})$ for standard power distribution.

Chacko, M., and George, V. (2024). Extropy properties of ranked set sample when ranking is not perfect. *Communications in Statistics-Theory and Methods*, **53**, 3187–3210.

Chen, Z., Bai, Z., and Sinha, B. K. (2004). *Ranked Set Sampling: Theory and Applications* (Vol. 176). Springer, New York.

George, V., and Chacko, M. (2023). Cumulative residual extropy properties of ranked set sample for Cambanis type bivariate distributions. *Journal of the Kerala Statistical Association*, **33**, 50–70.

Ghosh, A., and Kundu, C. (2018). On generalized conditional cumulative past inaccuracy measure. *Applications of Mathematics*, **63**, 167–193.

Goel, R., Taneja, H. C., and Kumar, V. (2018). Kerridge measure of inaccuracy for record statistics. *Journal of Information and Optimization Sciences*, **39**, 1149–1161.

Kayal, S., and Sunoj, S. M. (2017). Generalized Kerridge's inaccuracy measure for conditionally specified models. *Communications in Statistics-Theory and Methods*, **46**, 8257–8268.

Kerridge, D. F. (1961). Inaccuracy and inference. *Journal of the Royal Statistical Society. Series B* (Methodological), 184–194.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, **3**, 385–390.

Mohamed, M. S. (2019). A measure of inaccuracy in concomitants of ordered random variables under Farlie-Gumbel-Morgenstern family. *Filomat*, **33**, 4931–4942.

Mohammadi, M., Hashempour, M. (2023). Extropy based inaccuracy measure in order statistics. *Statistics*, **57**, 1490–1510.

Nath, P. (1968). Inaccuracy and coding theory. *Metrika*, **13**, 123–135.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, **27**, 379–423.

Sharma, A., and Kundu, C. (2024). Residual and past varinaccuracy measures. *IMA Journal of Mathematical Control and Information*, **41**, 539–563.

Taneja, H. C., Kumar, V., and Srivastava, R. (2009). A dynamic measure of inaccuracy between two residual lifetime distributions. *International Mathematical Forum*, **4**, 1213–1220.

Taneja, H. C., and Tuteja, R. K. (1986). Characterization of a quantitative-qualitative measure of inaccuracy. *Kybernetika*, **22**, 393–402.

Thapliyal, R., and Taneja, H. C. (2013). A measure of inaccuracy in order statistics. *Journal of Statistical Theory and Applications,* **12**, 200–207.

Thapliyal, R., and Taneja, H. C. (2015). On residual inaccuracy of order statistics. *Statistics and Probability Letters*, **97**, 125–131.