



Cumulative residual extropy of minimum ranked set sampling with unequal samples



Mohammad Reza Kazemi^a, Saeid Tahmasebi^b, Camilla Calì^c,
Maria Longobardi^{c,*}

^a Department of Statistics, Faculty of Science, Fasa University, Fasa, Iran

^b Department of Statistics, Persian Gulf University, Bushehr, Iran

^c Department of Biology, University of Napoli Federico II, Napoli, Italy

ARTICLE INFO

Article history:

Received 6 November 2020

Received in revised form 24 January 2021

Accepted 1 February 2021

Available online xxxx

MSC:

62B10

60E15

62D05

94A17

Keywords:

Cumulative residual extropy

Discrimination measure

Minimum ranked set sampling

Stochastic ordering

ABSTRACT

Recently, an alternative measure of uncertainty called cumulative residual extropy (CREX) was proposed by Jahanshahi et al. (2019). In this paper, we consider uncertainty measures of minimum ranked set sampling procedure with unequal samples (MinRSSU) in terms of CREX and its dynamic version and we compare the uncertainty and information content of CREX based on MinRSSU and simple random sampling (SRS) designs. Also, using simulation, we study on new estimators of CREX for MinRSSU and SRS designs in terms of bias and mean square error. Finally, we provide a new discrimination measure of disparity between the distribution of MinRSSU and parental data SRS.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Ranked set sampling (RSS) design is a cost-effective sampling for situations where taking actual measurements on units is expensive but ranking units is easy. For the first time, based on the RSS sampling design, [1] provided a more efficient estimator of the population mean comparing to the simple random sampling (SRS) counterpart. To learn more about this concept, the readers can refer to [2]. There are many available studies that have developed and generalized the method of sampling used in RSS scheme and they efficiently estimate the population parameter comparing to the SRS scheme. Recently, [3] studied information content of minimum ranked set sampling procedure with unequal samples (MinRSSU) as useful modification of RSS procedure in terms of extropy. In the MinRSSU, we draw m simple random samples, where the size of the i th samples is i , $i = 1, \dots, m$. The one-cycle MinRSSU involves an initial ranking of m samples of size m as follows:

$$\begin{array}{llll}
 1 : & \underline{X}_{(1:1)1} & & \rightarrow \tilde{X}_1 = X_{(1:1)1} \\
 2 : & \underline{X}_{(1:2)2} & X_{(2:2)2} & \rightarrow \tilde{X}_2 = X_{(1:2)2} \\
 \vdots & \vdots & \vdots & \vdots \\
 m : & \underline{X}_{(1:m)m} & X_{(2:m)m} \cdots X_{(m:m)m} & \rightarrow \tilde{X}_m = X_{(1:m)m}
 \end{array}$$

* Corresponding author.

E-mail address: malongob@unina.it (M. Longobardi).

where $X_{(i:i)j}$ denotes the i th order statistic from the j th SRS of size i . The resulting sample is called one-cycle MinRSSU of size m and denoted by $\mathbf{X}_{\text{MinRSSU}}^{(m)} = \{\tilde{X}_i, i = 1, \dots, m\}$. The parameter m should be kept small because the ranking should not be difficult in this sense and the ranking may be done for example by using an easily measurable covariate, then it is not difficult to identify the minimum of ranked individuals in each subset. Note that \tilde{X}_i has the same distribution as $X_{(1)i}$ which is the smallest order statistic in a set of size i with probability density function (pdf) $f_{(1)i}(x) = if(x)[1 - F(x)]^{i-1}$ and survival function $\bar{F}_{(1)i}(x) = [1 - F(x)]^i = \bar{F}^i(x)$, where $f(\cdot)$, $F(\cdot)$ and $\bar{F}(\cdot)$ are the underlying pdf, cumulative distribution function (cdf) and survival function. In reliability theory, \tilde{X}_i measures the lifetime of a series system.

Several authors have worked on measures of information for RSS and its variants. [4] explored the notions of information content of RSS data and compared them with their counterparts in SRS data. [5] obtained some results of residual (past) entropy for ranked set samples. [6] studied information measures for record ranked set sampling. [7] considered information measures of maximum ranked set sampling procedure with unequal samples in terms of Shannon entropy, Rényi entropy and Kullback–Leibler information, instead [8] in terms of Tsallis entropy. More recently, [3] studied information content of MinRSSU in terms of extropy and [9] considered the problems of uncertainty and information content of RSS data based on extropy measure and the related monotonic properties and stochastic comparisons.

Let X denotes a continuous random variable with pdf f . [10] introduced a new measure termed by extropy associated with X as

$$J(X) = -\frac{1}{2} \int_{-\infty}^{+\infty} [f(x)]^2 dx = -\frac{1}{2} \int_0^1 f(F^{-1}(u)) du, \tag{1.1}$$

where $F^{-1}(\cdot)$ is the quantile function of X . [11] explored some characterization results, monotone properties, and lower bounds of extropy of order statistics and record values. Also, [3] and [9] considered the information measure of extropy $J(X)$ based on MinRSSU and RSS schemes, respectively and compared the results with their counterpart under SRS design.

This paper is organized as follows: Section 2 deals with the results of cumulative residual extropy (CREX) for MinRSSU data by comparing to its counterpart under SRS data. In Section 3, new estimators are proposed for CREX in SRS and MinRSSU designs using empirical approach. Also, by using simulation study, the behaviour of estimators of CREX in MinRSSU and SRS are compared in terms of bias and mean square error. Furthermore, we show that how MinRSSU scheme can efficiently reduce the uncertainty measure comparing to SRS design. Section 4 provides a new discrimination measure of disparity between the distribution of MinRSSU and parental data SRS. Section 5 concludes the paper.

2. Cumulative residual extropy of MinRSSU

Let X denotes the lifetime of a system with survival function \bar{F} . Recently, a new measure of information is proposed by [12] with substituting the function \bar{F} in extropy formula (1.1). This new measure is called CREX and defined as

$$\xi \mathcal{J}(X) = -\frac{1}{2} \int_0^{+\infty} \bar{F}^2(x) dx. \tag{2.1}$$

Note that $-\infty < \xi \mathcal{J}(X) \leq 0$. If the CREX of X is less than that of another random variable, say Y , i.e. $\xi \mathcal{J}(X) \leq \xi \mathcal{J}(Y)$, then X has less uncertainty than Y . Now let $\xi \mathcal{J}(X) < +\infty$. Then, for the MinRSSU and SRS designs, we have

$$\xi \mathcal{J}(\mathbf{X}_{\text{MinRSSU}}^{(m)}) = -\frac{1}{2} \prod_{i=1}^m [-2\xi \mathcal{J}(X_{(1:i)})] = -\frac{1}{2} \prod_{i=1}^m \int_0^{+\infty} \bar{F}^{2i}(x) dx \tag{2.2}$$

$$\begin{aligned} &= -\frac{1}{2} \prod_{i=1}^m \int_0^1 \frac{(1-u)^{2i}}{f(F^{-1}(u))} du \\ &= -\frac{1}{2} \prod_{i=1}^m \mathbb{E} \left[\frac{(1-U)^{2i}}{f(F^{-1}(U))} \right], \end{aligned} \tag{2.3}$$

and

$$\xi \mathcal{J}(\mathbf{X}_{\text{SRS}}^{(m)}) = -\frac{1}{2} \left[\int_0^{+\infty} \bar{F}^2(x) dx \right]^m = -\frac{1}{2} [-2\xi \mathcal{J}(X)]^m. \tag{2.4}$$

To compare the above measures, let us consider the following examples.

Example 2.1. If $U \sim \text{Uniform}(0, 1)$, then

$$\xi \mathcal{J}(\mathbf{U}_{\text{MRSSU}}^{(m)}) = -\frac{1}{2} \prod_{i=1}^m \frac{1}{2i+1} = -\frac{1}{2} \left(\frac{\sqrt{\pi}}{2^m \Gamma(m + \frac{1}{2})} \right) < \xi \mathcal{J}(\mathbf{U}_{\text{SRS}}^{(m)}) = -0.5 \left(\frac{1}{3} \right)^m. \tag{2.5}$$

Example 2.2. If Z is exponentially distributed with mean $\frac{1}{\lambda}$, then we have

$$\xi \mathcal{J}(\mathbf{Z}_{MRSSU}^{(m)}) = -\frac{1}{2} \prod_{i=1}^m \frac{1}{2i\lambda} < \xi \mathcal{J}(\mathbf{Z}_{SRS}^{(m)}) = -\frac{1}{2} \left(\frac{1}{2\lambda}\right)^m. \tag{2.6}$$

Example 2.3. Let X is finite range distribution with $\bar{F}(x) = (1 - ax)^b$, $0 < x < \frac{1}{a}$, $a > 0$, $b > 0$. Then, we have

$$\xi \mathcal{J}(\mathbf{X}_{MRSSU}^{(m)}) = -\frac{1}{2} \prod_{i=1}^m \frac{1}{a(1 + 2ib)} < \xi \mathcal{J}(\mathbf{X}_{SRS}^{(m)}) = -\frac{1}{2} \left(\frac{1}{a(1 + 2b)}\right)^m. \tag{2.7}$$

Theorem 2.1. Let $\mathbf{X}_{MinRSSU}^{(m)}$ be the MinRSSU from population X with pdf f and cdf F . Then, $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \leq \xi \mathcal{J}(\mathbf{X}_{SRS}^{(m)})$ for $m > 1$.

Proof. Since $\bar{F}^2(x) \geq \bar{F}^{2i}(x)$ for $i \geq 1$, we have

$$\left(\int_0^{+\infty} \bar{F}^2(x) dx\right)^m \leq \prod_{i=1}^m \int_0^{+\infty} \bar{F}^{2i}(x) dx.$$

The proof follows by recalling (2.2) and (2.4). \square

Remark 2.1. If $f(F^{-1}(u)) \geq 1$, $0 < u < 1$, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)})$ is increasing in $m \geq 1$.

Proof. From (2.2), we get

$$\frac{\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m+1)})}{\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)})} = \int_0^1 \frac{(1-u)^{2m+2}}{f(F^{-1}(u))} du \leq \frac{1}{2m+3} \leq 1.$$

The result follows readily, since the entropy is negative. \square

In the following, we provide some results on the cumulative residual entropy of $\mathbf{X}_{MinRSSU}^{(m)}$ in terms of stochastic ordering properties. Now, we state important properties of $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)})$ using the stochastic ordering. For that we present the following definitions:

Definition 2.2 (Shaked and Shanthikumar, 2007). Let X and Y be two non-negative random variables with pdfs f and g , cdfs F and G , and hazard functions $\lambda_X(x) = \frac{f(x)}{\bar{F}(x)}$ and $\lambda_Y(y) = \frac{g(y)}{\bar{G}(y)}$, respectively. Then

1. X is said to be smaller than Y in the usual stochastic order (denoted by $X \leq_{st} Y$) if $P(X \geq x) \leq P(Y \geq x)$ for all $x \in \mathbb{R}$.
2. X is smaller than Y in the hazard rate order (denoted by $X \leq_{hr} Y$) if $\lambda_X(x) \geq \lambda_Y(x)$ for all x .
3. X is smaller than Y in the dispersive order (denoted by $X \leq_{disp} Y$) if $f(F^{-1}(u)) \geq g(G^{-1}(u))$ for all $u \in (0, 1)$, where F^{-1} and G^{-1} are right continuous inverses of F and G , respectively.
4. X is said to have decreasing failure rate (DFR) if $\lambda_X(x)$ is decreasing in x .
5. X is smaller than Y in the convex transform order (denoted by $X \leq_c Y$) if $G^{-1}F(x)$ is a convex function on the support of X .
6. X is smaller than Y in the star order (denoted by $X \leq_* Y$) if $\frac{G^{-1}F(x)}{x}$ is increasing in $x \geq 0$.
7. X is smaller than Y in the superadditive order (denoted by $X \leq_{su} Y$) if $G^{-1}(F(t + u)) \geq G^{-1}(F(t)) + G^{-1}(F(u))$ for $t \geq 0, u \geq 0$.
8. X is said to have an increasing reversed hazard rate (IRHR) if $\tilde{\lambda}_X(x) = \frac{f(x)}{\bar{F}(x)}$ is increasing in x .

Theorem 2.3. If $X \leq_{st} Y$, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\mathbf{Y}_{MinRSSU}^{(m)})$, $m > 1$.

Proof. By the assumption of the stochastic order, $\bar{F}^{2i}(x) \leq \bar{G}^{2i}(x)$ for all $x \geq 0$. Now using (2.2), for $m > 1$, we get the desired result. \square

Theorem 2.4. Let X and Y be two non-negative random variable. If $X \leq_{disp} Y$, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\mathbf{Y}_{MinRSSU}^{(m)})$ for $m > 1$.

Proof. By the assumption of the dispersive order, $f(F^{-1}(u)) \geq g(G^{-1}(u))$ for all $u \in (0, 1)$. Using (2.2), for $m > 1$ the result follows. \square

Theorem 2.5. If $X \leq_{hr} Y$, and X or Y is DFR, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\mathbf{Y}_{MinRSSU}^{(m)})$ for $m > 1$.

Proof. If $X \leq_{hr} Y$, and X or Y is DFR, then $X \leq_{disp} Y$, due to [13]. Thus, from Theorem 2.4 the desired result follows. \square

Theorem 2.6. Let X and Y be two non-negative random variable with pdfs f and g , respectively, such that $f(0) \geq g(0) > 0$. If $X \leq_{su} Y (X \leq_* Y \text{ or } X \leq_c Y)$, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\mathbf{Y}_{MinRSSU}^{(m)})$ for $m > 1$.

Proof. If $X \leq_{su} Y (X \leq_* Y \text{ or } X \leq_c Y)$, then $X \leq_{disp} Y$, due to [14]. So, from Theorem 2.4 the desired result follows. \square

Proposition 2.7. Let $\mathbf{X}_{MinRSSU}^{(m)}$ and $\mathbf{X}_{RSS}^{(m)}$ be MinRSSU and RSS data from distribution X with DFR ageing property, respectively. Then for $m > 1$ we have

$$\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\mathbf{X}_{SRS}^{(m)}).$$

Remark 2.2. Let φ be a non-negative function such that the derivative $\varphi'(x) \geq 1$ for all x . Then $X \leq_{disp} \varphi(X)$. Thus,

$$\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) \geq \xi \mathcal{J}(\varphi(\mathbf{X})_{MinRSSU}^{(m)}).$$

If $X \leq_{disp} Y$, Theorem 3.B.26 in [15] claims that $X_{i:m} \leq_{disp} Y_{i:m}$, $i = 1, 2, \dots, m$. Thus, according to Theorems 4.5 and 4.6 in [11], we obtain the following Proposition .

Proposition 2.8. Let $\mathbf{X}_{MinRSSU}^{(m)}$ be a sample from MinRSSU design.

(i) If X is DFR ageing property, then $\xi \mathcal{J}(X_{(1)m})$ is increasing in $m \geq 1$.

(ii) If X is IRHR property, then $\xi \mathcal{J}(X_{(1)m})$ is decreasing in $m \geq 1$.

Proposition 2.9. Let $\mathbf{Y}_{MinRSSU}^{(m)} = a\mathbf{X}_{MinRSSU}^{(m)} + b$ with $a > 0$ and $b \geq 0$. Then, $\xi \mathcal{J}(\mathbf{Y}_{MinRSSU}^{(m)}) = a\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)})$.

Proposition 2.10. Let X be a symmetric random variable with respect to the finite mean $\mu = \mathbb{E}(X)$. Then

$$\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}) = \mathcal{C} \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}),$$

where $\mathcal{C} \mathcal{J}(X) = -\frac{1}{2} \int_0^{+\infty} [F_X(x)]^2 dx$ is the cumulative extropy (see [12]).

Let X be the random lifetime of a system, recall that $X_{[t]} = [X - t | X \leq t]$ describes the residual lifetime of a system. For all $t \geq 0$ the mean residual lifetime is given by

$$\mu(t) = \mathbb{E}[X - t | X \geq t] = \frac{1}{\bar{F}(t)} \int_t^{+\infty} \bar{F}(x) dx.$$

Now, we can define a generalized measure of cumulative residual extropy as

$$\xi \mathcal{J}(X; t) = -\frac{1}{2} \int_t^{+\infty} \left[\frac{\bar{F}(x)}{\bar{F}(t)} \right]^2 dx. \tag{2.8}$$

Note that $\xi \mathcal{J}(X; t) \geq -\mu(t)/(2\bar{F}(t))$. Moreover , we have

$$\xi \mathcal{J}(\mathbf{X}_{SRS}^{(m)}; t) = -\frac{1}{2} [-2\xi \mathcal{J}(X, t)]^m. \tag{2.9}$$

Under the MinRSSU design, it is clear that

$$\begin{aligned} \xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}; t) &= -\frac{1}{2} \prod_{i=1}^m [-2\xi \mathcal{J}(X_{(i);i}; t)] = -\frac{1}{2} \prod_{i=1}^m \int_t^{+\infty} \left[\frac{\bar{F}(x)}{\bar{F}(t)} \right]^{2i} dx \\ &= -\frac{1}{2} \prod_{i=1}^m \mathbb{E} \left[\frac{U^{2i} \bar{F}(t)}{f(F^{-1}(1 - U\bar{F}(t)))} \right]. \end{aligned} \tag{2.10}$$

Theorem 2.11. Let X be a random lifetime variable with cdf $F(\cdot)$. Then, for $m > 1$

$$\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}; t) \leq \xi \mathcal{J}(\mathbf{X}_{SRS}^{(m)}; t). \tag{2.11}$$

Proof. The proof is similar to Theorem 2.1. \square

Remark 2.3. If $f(F^{-1}(1 - u\bar{F}(t))) \geq 1$, $0 < u < 1$, then $\xi \mathcal{J}(\mathbf{X}_{MinRSSU}^{(m)}; t)$ is increasing in $m \geq 1$.

3. Results on empirical measure of CREX

This section focuses on the estimation of the $\xi \mathcal{J}(X)$ based on the SRS and MinRSSU schemes. Let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ be the order statistics of the random sample X_1, X_2, \dots, X_n from cdf F . Then the empirical measure of F is defined as

$$\hat{F}_n(x) = \begin{cases} 0, & x < X_{(1)}, \\ \frac{k}{n}, & X_{(k)} \leq x \leq X_{(k+1)}, \\ 1, & x > X_{(n)}. \end{cases} \quad k = 1, 2, \dots, n - 1$$

Thus, the empirical measure of $\xi \mathcal{J}(X)$ is obtained by replacing the distribution function F by the empirical distribution function \hat{F}_n as

$$\begin{aligned} V_n &= -\frac{1}{2} \int \hat{F}_n^2(x) dx = -\frac{1}{2} \sum_{k=1}^{n-1} \int_{X_{(k)}}^{X_{(k+1)}} \left(1 - \frac{k}{n}\right)^2 dx \\ &= -\frac{1}{2} \sum_{k=1}^{n-1} U_{k+1} \left(1 - \frac{k}{n}\right)^2, \end{aligned} \tag{3.1}$$

where $U_{k+1} = X_{(k+1)} - X_{(k)}$, $k = 1, \dots, n - 1$. [12] showed that V_n almost surely converges to the CREX of X , i.e.

$$V_n \xrightarrow{a.s.} \xi \mathcal{J}(X), \quad \text{as } n \rightarrow +\infty.$$

The problem of estimation of $\xi \mathcal{J}(X)$ based on MinSSU scheme can be deduced in the same line of the estimation based on SRS design V_n . In this part, we assume that instead of one-cycle MinRSSU, the process is repeated l cycles to have a sample of size $n = ml$. In this case, the resulting MinRSSU is denoted by $\{X_{(1:i)j}, i = 1, \dots, m; j = 1, \dots, l\}$, where $X_{(1:i)j}$ is the lowest order statistic from the i th sample in the j th cycle. Let $Y_{(1)}, \dots, Y_{(n)}$ be the ordered values of the MinRSSU design $\{X_{(1:i)j}, i = 1, \dots, m; j = 1, \dots, l\}$. Then, the natural estimation of $\xi \mathcal{J}(X)$ based on the MinRSSU, can be obtained as

$$R_n = -\frac{1}{2} \sum_{k=1}^{n-1} Z_{k+1} \left(1 - \frac{k}{n}\right)^2,$$

where $Z_{k+1} = Y_{(k+1)} - Y_{(k)}$, $k = 1, 2, \dots, n - 1$. Our preliminary computations and simulations showed that R_n has some deficiencies to be unbiased and have low mean square error (MSE) to estimate the $\xi \mathcal{J}(X)$. We observed that the term $\left(1 - \frac{k}{n}\right)$ should be slightly modified so that the estimator has optimal properties to estimate $\xi \mathcal{J}(X)$. We propose to modify this term with $\left(1 - \frac{k}{n+m+w}\right)$, where w is a number that resulting estimator has optimally low bias and MSE. The resulting estimator of the $\xi \mathcal{J}(X)$ has the following form

$$R_{m,n} = -\frac{1}{2} \sum_{k=1}^{n-1} Z_{k+1} \left(1 - \frac{k}{n+m+w}\right)^2. \tag{3.2}$$

3.1. A new estimation

In the previous section, the estimator V_n of $\xi \mathcal{J}(X)$ is a linear function of sample spacing $U_{k+1} = X_{(k+1)} - X_{(k)}$, $k = 1, \dots, n - 1$. The asymptotic distribution of this linear function of sample spacing can be found, for example in [16] and [17] only for exponential and standard uniform distributions. So we provide another estimator for $\xi \mathcal{J}(X)$ which is a linear function of order statistics.

Proposition 3.1. *Let X be an absolutely continuous non-negative random variable with survival function \bar{F} , then*

$$\xi \mathcal{J}(X) = -\int_0^{+\infty} x \bar{F}(x) dF(x). \tag{3.3}$$

Proof. By (2.1) and Fubini's theorem, we obtain

$$\begin{aligned} -\int_0^{+\infty} x \bar{F}(x) dF(x) &= -\int_0^{+\infty} \left(\int_0^x dt\right) \bar{F}(x) f(x) dx \\ &= -\int_0^{+\infty} \left(\int_t^{+\infty} \bar{F}(x) f(x) dx\right) dt = -\frac{1}{2} \int_0^{+\infty} \bar{F}^2(t) dt. \end{aligned}$$

Hence, the proof is completed. \square

The new estimator can be obtained replacing $\bar{F}(x)$ with $\hat{F}_n(x)$ in (3.3), so $\widehat{\xi\mathcal{J}}(X)$ has the following form

$$\widehat{\xi\mathcal{J}}(X) = - \int_0^{+\infty} x \hat{F}_n(x) d\hat{F}_n(x) = - \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{i}{n}\right) X_{(i)}. \tag{3.4}$$

Let $J(x) = (1 - x)$. Then (3.4) has the form

$$\widehat{\xi\mathcal{J}}(X) = - \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n}\right) X_{(i)},$$

which is a linear function of order statistics. The natural estimation of $\xi\mathcal{J}(X)$ based on the MinRSSU, can be obtained as

$$\widehat{\xi\mathcal{J}}(Y) := \widehat{\xi\mathcal{J}}(X_{\text{MinRSSU}}) = - \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n}\right) Y_{(i)}.$$

[18] showed that asymptotic distribution of such a linear combination is normal distribution. The results of [18] also hold if the independent observations are not identically distributed. These properties help us to obtain the asymptotic distribution of $\widehat{\xi\mathcal{J}}(X)$ for both SRS (observations are independent and identical) and MinRSSU (observations are only independent) designs.

Theorem 3.2. Assume that $\mathbb{E}(X^2) < +\infty$. Then

$$\begin{aligned} \sqrt{n} (\widehat{\xi\mathcal{J}}(X) - \xi\mathcal{J}(X)) &\xrightarrow{d} N(0, \sigma^2(J, F)), \\ \sqrt{n} (\widehat{\xi\mathcal{J}}(Y) - \xi\mathcal{J}(X)) &\xrightarrow{d} N(0, \sigma_{\text{MinRSSU}}^2(J, \tilde{F}, K)), \end{aligned}$$

where

$$\sigma^2(J, F) = \int_0^{+\infty} \int_0^{+\infty} J(F(x))J(F(y))[F(\min(x, y)) - F(x)F(y)]dx dy,$$

$$\sigma_{\text{MinRSSU}}^2(J, \tilde{F}, K) = \int_0^{+\infty} \int_0^{+\infty} J(\tilde{F}(x))J(\tilde{F}(y))K(x, y)dx dy,$$

and $\tilde{F}(\cdot)$ and $K(x, y)$ are given as

$$\begin{aligned} \tilde{F}(x) &= \frac{1}{m} \sum_{i=1}^m F_{(1)i}(x), \\ K(x, y) &= \frac{1}{m} \sum_{i=1}^m [F_{(1)i}(\min(x, y)) - F_{(1)i}(x)F_{(1)i}(y)]. \end{aligned}$$

Proof. For $\tilde{F}(x)$ and $K(x, y)$, we have

$$\begin{aligned} \tilde{F}(x) &= \lim_{n \rightarrow +\infty} \frac{l}{n} \sum_{i=1}^m F_{(1)i}(x) = \frac{1}{m} \sum_{i=1}^m F_{(1)i}(x), \\ K(x, y) &= \lim_{n \rightarrow +\infty} \frac{l}{n} \sum_{i=1}^m [F_{(1)i}(\min(x, y)) - F_{(1)i}(x)F_{(1)i}(y)] \\ &= \frac{1}{m} \sum_{i=1}^m [F_{(1)i}(\min(x, y)) - F_{(1)i}(x)F_{(1)i}(y)], \end{aligned}$$

where as before l is the size of the cycle of MinRSSU design with $n = ml$. The rest of the proof is done by using the results of [18] for both SRS and MinRSSU designs. \square

As we formerly stated for providing the estimator $R_{m,n}$, some adjusted forms of estimator $\widehat{\xi\mathcal{J}}(Y)$ can be used since the estimator $\widehat{\xi\mathcal{J}}(Y)$ has some deficiencies to estimate CREX which need to be fixed. We again observed that the term $(1 - \frac{i}{n})$ needs to be slightly adjusted so that the resulting estimator has optimal properties to estimate CREX. This new term has the following form

$$1 - \frac{i}{n + \psi(m, w)}, \tag{3.5}$$

Table 1
The biases and MSEs of the different estimators: Exponential distribution.

m	w	Results based on $R_{m,n}$				m	w	Results based on $\widehat{\xi\mathcal{J}}_{m,n}(Y)$			
		l = 2		l = 3				l = 2		l = 3	
		Bias	RMSE	Bias	RMSE			Bias	RMSE	Bias	RMSE
2	-2	0.321	0.407	0.370	0.477	2	-11	0.632	0.650	0.536	0.562
	-1	0.131	0.354	0.133	0.460		-10	0.402	0.483	0.354	0.427
	0	-0.051	0.423	-0.069	0.580		-9	0.263	0.435	0.223	0.370
	1	-0.216	0.556	-0.234	0.733		-8	0.171	0.434	0.125	0.361
	R_n	0.321	0.407	0.370	0.477		$\widehat{\xi\mathcal{J}}(Y)$	0.402	0.483	0.354	0.427
3	-1	0.124	0.367	0.164	0.304	3	-7	0.033	0.374	0.006	0.307
	0	-0.017	0.429	0.040	0.315		-6	0.012	0.382	-0.020	0.317
	1	0.124	0.367	0.164	0.304		-5	-0.006	0.391	-0.044	0.328
	2	-0.253	0.632	-0.191	0.472		-4	-0.022	0.399	-0.065	0.340
	R_n	0.437	0.478	0.411	0.444		$\widehat{\xi\mathcal{J}}(Y)$	0.452	0.489	0.430	0.458
4	0	0.024	0.383	0.118	0.291	4	-3	0.029	0.329	-0.007	0.269
	1	-0.077	0.452	0.029	0.313		-2	0.020	0.332	-0.019	0.274
	2	-0.170	0.532	-0.057	0.363		-1	0.012	0.334	-0.030	0.278
	3	-0.256	0.616	-0.139	0.427		0	0.004	0.337	-0.040	0.283
	R_n	0.486	0.509	0.474	0.490		$\widehat{\xi\mathcal{J}}(Y)$	0.509	0.527	0.489	0.503
5	1	-0.004	0.390	0.114	0.281	5	1	0.072	0.277	0.050	0.232
	2	-0.082	0.448	0.046	0.296		2	0.067	0.277	0.043	0.233
	3	-0.156	0.512	-0.021	0.329		3	0.063	0.278	0.036	0.233
	4	-0.224	0.578	-0.085	0.373		4	0.058	0.278	0.030	0.234
	R_n	0.533	0.545	0.524	0.532		$\widehat{\xi\mathcal{J}}(Y)$	0.554	0.563	0.540	0.547

where the function $\psi(\cdot, \cdot)$ is a challenging factor that can be specifically determined for each given distribution and consequently the resulting estimator has optimal low bias and MSE. In this case, the resulting estimator has the following form

$$\widehat{\xi\mathcal{J}}_{m,n}(Y) = -\frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n + \psi(m, w)}\right) Y_{(i)}. \tag{3.6}$$

In the next section, we determine the form of the function $\psi(\cdot, \cdot)$ for exponential, uniform and beta distributions and we show that the optimal choice of this function can reduce the bias and MSE in the estimate of CREX.

3.2. Simulation study

In advance, we explain the role of the parameter w and function $\psi(m, w)$ for which $R_{m,n}$ and $\widehat{\xi\mathcal{J}}_{m,n}(X_{MinRSSU})$ have optimally low bias and MSE. For this purpose, we examine some distributions to obtain the function $\psi(m, w)$ and optimal value of w in $R_{m,n}$ and $\widehat{\xi\mathcal{J}}_{m,n}(Y)$. In Tables 1–3, for four estimators R_n , $R_{m,n}$, $\widehat{\xi\mathcal{J}}(Y)$ and $\widehat{\xi\mathcal{J}}_{m,n}(Y)$, we compute the bias and MSE to estimate the parameter $\xi\mathcal{J}(X)$ for some different values of w . Here, the exponential ($Exp(\lambda)$), uniform ($Unif(0, b)$) and beta ($Beta(\alpha, 1)$, $\alpha > 1$) distributions are considered. For each configuration, the simulation study was carried out with 5000 repetitions. The number of cycle and size of the sample in each cycle are taken as $l = 2, 3$ and $m = 2, \dots, 5$, respectively. We compute the bias and root of MSE (RMSE) of each estimator of parameter $\xi\mathcal{J}(X)$. In Tables 1–3, it can be seen that results of biases and RMSEs of estimator $\widehat{\xi\mathcal{J}}(Y)$ are not comparable to those of $\widehat{\xi\mathcal{J}}_{m,n}(Y)$ for different values of w . We intuitively obtain the form of the function $\psi(m, w)$ for different distributions. We found that the function ψ has the form $5m - 4k_m + w$ with $k_2 = 3, \dots, k_5 = 0$, $3m - (2k_m + 1) + w$ with $k_2 = -1, \dots, k_5 = 2$ and $m - w$ for exponential, uniform and beta distributions, respectively. It is observed that changing the value of l has no effect on the whole results verified from biases and RMSEs. In all tables, we see that choosing the proper function $\psi(m, w)$ and parameter w as the challenging factors can improve the efficiency of estimators $R_{m,n}$ and $\widehat{\xi\mathcal{J}}_{m,n}(Y)$ against R_n and $\widehat{\xi\mathcal{J}}(Y)$, respectively, in estimating the parameter $\xi\mathcal{J}$.

4. Discrimination information

This section considers a new discrimination measure of disparity between the distribution of MinRSSU and parental data SRS. [9] defined the discrimination information between the density function of the i th order statistic $f_{(i)m}$ and the underlying density function f as

$$D_m(f_{(i)m} : f) = \frac{1}{2} \int_{-\infty}^{+\infty} f_{(i)m}(x) (f_{(i)m}(x) - f(x)) dx. \tag{4.1}$$

Table 2
The biases and MSEs of the different estimators: Uniform distribution.

m	w	Results based on $R_{m,n}$				m	w	Results based on $\widehat{\xi}_{\mathcal{T}_{m,n}}(Y)$			
		l = 2		l = 3				l = 2		l = 3	
		Bias	RMSE	Bias	RMSE			Bias	RMSE	Bias	RMSE
2	-2	0.360	0.392	0.298	0.333	2	-4	0.098	0.275	0.058	0.227
	-1	0.249	0.305	0.212	0.264		-3	0.045	0.281	-0.001	0.236
	0	0.158	0.253	0.134	0.214		-2	0.005	0.293	-0.049	0.254
	1	0.085	0.233	0.066	0.186		-1	-0.026	0.306	-0.088	0.275
	R_n	0.360	0.392	0.298	0.333		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.284	0.340	0.238	0.293
3	-1	0.191	0.243	0.201	0.237	3	-2	-0.006	0.247	-0.023	0.200
	0	0.125	0.204	0.147	0.196		-1	-0.026	0.255	-0.047	0.209
	1	0.067	0.183	0.096	0.165		0	-0.043	0.263	-0.068	0.220
	2	0.016	0.181	0.055	0.148		1	-0.058	0.270	-0.087	0.231
	R_n	0.349	0.372	0.318	0.339		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.318	0.350	0.290	0.317
4	0	0.14	0.195	0.175	0.204	4	0	-0.012	0.205	-0.008	0.164
	1	0.092	0.171	0.135	0.173		1	-0.023	0.209	-0.021	0.168
	2	0.049	0.159	0.098	0.148		2	-0.034	0.213	-0.034	0.172
	3	0.018	0.159	0.062	0.131		3	-0.043	0.217	-0.046	0.177
	R_n	0.372	0.387	0.349	0.360		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.352	0.369	0.337	0.350
5	1	0.116	0.17	0.175	0.198	5	2	0.020	0.181	0.028	0.146
	2	0.079	0.153	0.144	0.172		3	0.013	0.183	0.019	0.146
	3	0.044	0.144	0.114	0.150		4	0.006	0.184	0.010	0.147
	4	0.012	0.144	0.085	0.132		5	-0.001	0.186	0.002	0.148
	R_n	0.392	0.401	0.381	0.388		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.387	0.398	0.375	0.383

Table 3
The biases and MSEs of the different estimators: Beta distribution.

m	w	Results based on $R_{m,n}$				m	w	Results based on $\widehat{\xi}_{\mathcal{T}_{m,n}}(Y)$			
		l = 2		l = 3				l = 2		l = 3	
		Bias	RMSE	Bias	RMSE			Bias	RMSE	Bias	RMSE
2	-2	0.296	0.298	0.276	0.279	2	-3	0.202	0.204	0.146	0.149
	-1	0.280	0.283	0.264	0.267		-2	0.108	0.116	0.081	0.089
	0	0.267	0.271	0.253	0.257		-1	0.053	0.070	0.035	0.055
	1	0.257	0.261	0.244	0.248		0	0.015	0.053	0.000	0.045
	R_n	0.296	0.298	0.276	0.279		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.108	0.116	0.081	0.089
3	-1	0.251	0.255	0.238	0.242	3	-3	0.096	0.102	0.079	0.085
	0	0.241	0.245	0.230	0.234		-2	0.052	0.064	0.045	0.056
	1	0.233	0.238	0.222	0.227		-1	0.019	0.045	0.017	0.039
	2	0.225	0.231	0.216	0.221		0	-0.007	0.043	-0.006	0.037
	R_n	0.275	0.278	0.256	0.259		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.096	0.102	0.079	0.085
4	0	0.227	0.231	0.218	0.222	4	-3	0.059	0.068	0.056	0.063
	1	0.220	0.225	0.212	0.216		-2	0.031	0.047	0.033	0.044
	2	0.214	0.219	0.206	0.211		-1	0.007	0.038	0.013	0.033
	3	0.208	0.213	0.201	0.206		0	-0.012	0.041	-0.004	0.031
	R_n	0.264	0.267	0.246	0.249		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.095	0.100	0.082	0.087
5	1	0.217	0.221	0.210	0.214	5	-3	0.043	0.053	0.047	0.054
	2	0.211	0.216	0.205	0.209		-2	0.022	0.039	0.030	0.040
	3	0.206	0.210	0.200	0.204		-1	0.004	0.034	0.015	0.031
	4	0.201	0.205	0.196	0.201		0	-0.012	0.036	0.002	0.028
	R_n	0.257	0.259	0.239	0.242		$\widehat{\xi}_{\mathcal{T}}(Y)$	0.097	0.101	0.088	0.091

Analogously to (4.1), we define the discrimination information between the survival function of the smallest ordered statistic $\bar{F}_{(1)i}$ and the underlying survival function \bar{F} as

$$D(\bar{F}_{(1)i} : \bar{F}) = -\frac{1}{2} \int_{-\infty}^{+\infty} \bar{F}_{(1)i}(x) (\bar{F}_{(1)i}(x) - \bar{F}(x)) dx. \tag{4.2}$$

The discrimination information in (4.2) may be rewritten in a simpler way. It can be shown that

$$D(\bar{F}_{(1)i} : \bar{F}) = -\frac{1}{2} [\mathbb{E}(X_{(1)2i}) - \mathbb{E}(X_{(1)i+1})], \tag{4.3}$$

where $X_{(1)j}$ is the smallest order statistic in a random sample of size j .

Example 4.1. Let $U \sim \text{Uniform}(0, 1)$. We know that the order statistics from the standard uniform distribution follow the beta distribution. Then discrimination information based on (4.2) is

$$D(\bar{F}_{(1)i} : \bar{F}) = -\frac{1}{2} \left[\frac{1}{2i+1} - \frac{1}{i+2} \right] = \frac{i-1}{2(2i+1)(i+2)}.$$

In the following theorem, we obtain the discrimination information D between MinRSSU and SRS designs.

Theorem 4.1. For $\mathbf{X}_{\text{MinRSSU}}^{(m)}$ and $\mathbf{X}_{\text{SRS}}^{(m)}$, we have

$$D(\mathbf{X}_{\text{MinRSSU}}^{(m)} : \mathbf{X}_{\text{SRS}}^{(m)}) = -\frac{1}{2} \left(\prod_{i=1}^m \mathbb{E}(X_{(1)2i}) - \prod_{i=1}^m \mathbb{E}(X_{(1)i+1}) \right).$$

Proof. From (2.1) and (4.2) we have

$$\begin{aligned} D(\mathbf{X}_{\text{MinRSSU}}^{(m)} : \mathbf{X}_{\text{SRS}}^{(m)}) &= -\frac{1}{2} \left(\prod_{i=1}^m \int_{-\infty}^{+\infty} \bar{F}^{2i}(x_1) - \prod_{i=1}^m \int_{-\infty}^{+\infty} \bar{F}^{i+1}(x_1) \right) \\ &= -\frac{1}{2} \left(\prod_{i=1}^m \mathbb{E}(X_{(1)2i}) - \prod_{i=1}^m \mathbb{E}(X_{(1)i+1}) \right). \end{aligned}$$

The proof is completed. \square

Example 4.2. Let $U \sim \text{Uniform}(0, 1)$. Using the results of Example 4.1, the discrimination information D for the MinRSSU and SRS designs of the same size m is

$$D(\mathbf{U}_{\text{MinRSSU}}^{(m)} : \mathbf{U}_{\text{SRS}}^{(m)}) = -\frac{1}{2} \left(\prod_{i=1}^m \frac{1}{2i+1} - \prod_{i=1}^m \frac{1}{i+2} \right).$$

5. Conclusion

This paper has introduced the uncertainty measure of the cumulative residual entropy based on the MinRSSU and SRS data. Several results of the CREX measure including stochastic orders were obtained for MinRSSU and SRS data. Also, we provided two estimators of CREX measure for both SRS and MinRSSU data. Furthermore, it was shown that MinRSSU scheme can efficiently reduce the uncertainty measure of CREX. Also, by providing a discrimination measure, we derived the distance size between MinRSSU and SRS data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

C. Cali and M. Longobardi are partially supported by the GNAMPA research group of INDAM (Istituto Nazionale di Alta Matematica), Italy and MIUR-PRIN 2017, Project "Stochastic Models for Complex Systems" (No. 2017JFFHSH).

References

- [1] McIntyre GA. A method for unbiased selective sampling, using ranked sets. *Aust J Agric Res* 1952;3:385–90.
- [2] Patil G, Sinha A, Taillie C. Ranked set sampling: a bibliography. *Environ Ecol Stat* 1999;6:91–8.
- [3] Qiu G, Eftekharian A. Entropy information of maximum and minimum ranked set sampling with unequal samples. *Comm Statist Theory Methods* 2020. <http://dx.doi.org/10.1080/03610926.2019.1678640>.
- [4] Jozani MJ, Ahmadi J. On uncertainty and information properties of ranked set samples. *Inform Sci* 2014;264:291–301.
- [5] Tahmasebi S, Jafari AA, Eskandarzadeh M. Some results on residual entropy of ranked set samples. *Statist Probab Lett* 2016;112:137–45.
- [6] Eskandarzadeh M, Tahmasebi S, Afshari M. Information measures for record ranked set samples. *Ciência e Natura* 2016;38:554–63.
- [7] Eskandarzadeh M, Di Crescenzo A, Tahmasebi S. Measures of information for maximum ranked set sampling with unequal samples. *Comm Statist Theory Methods* 2018;47(19):4692–709.
- [8] Tahmasebi S, Longobardi M, Kazemi MR, Alizadeh M. Cumulative tsallis entropy for maximum ranked set sampling with unequal samples. *Physica A* 2020. <http://dx.doi.org/10.1016/j.physa.2020.124763>.
- [9] Raqab M, Qiu G. On entropy properties of ranked set sampling. *Statistics* 2019;53:210–26.
- [10] Frank L, Sanfilippo G, Agrò G. Entropy: Complementary dual of entropy. *Statist Sci* 2015;30:40–58.
- [11] Qiu G. The entropy of order statistics and record values. *Statist Probab Lett* 2017;120:52–60.
- [12] Jahanshahi SMA, Zarei H, Khammar AH. On cumulative residual entropy. *Probab Engrg Inform Sci* 2019;264:1–21. <http://dx.doi.org/10.1017/S0269964819000196>.

- [13] Bagai I, Kochar SC. On tail-ordering and comparison of failure rates. *Comm Statist Theory Methods* 1986;15:1377–88.
- [14] Ahmed AN, Alzaid A, Bartoszewicz J, Kochar SC. Dispersive and superadditive ordering. *Adv Appl Probab* 1986;18:1019–22.
- [15] Shaked M, Shanthikumar JG. *Stochastic orders*. New York: Springer; 2007.
- [16] Di Crescenzo A, Longobardi M. On cumulative entropies. *J Statist Plann Inference* 2009;139:4072–87.
- [17] Tahmasebi S. Weighted extensions of generalized cumulative residual entropy and their applications. *Comm Statist Theory Methods* 2019. <http://dx.doi.org/10.1080/03610926.2019.1615094>.
- [18] Stigler S. Linear functions of order statistics with smooth weight functions. *Ann Statist* 1974;2:676–93.