On Cumulative Entropies in Terms of Moments of Order Statistics

(Article begins on next page)

22 November 2024

On Cumulative Entropies in Terms of Moments of Order Statistics

Narayanaswamy Balakrishnan[1] · Francesco Buono[2] · Maria Longobardi[2]

Abstract

In this paper, relations between some kinds of cumulative entropies and moments of order statistics are established. By using some characterizations and the symmetry of a non-negative and absolutely continuous random variable X, lower and upper bounds for entropies are obtained and illustrative examples are given. By the relations with the moments of order statistics, a method is shown to compute an estimate of cumulative entropies and an application to testing whether data are exponentially distributed is outlined.

1 Introduction

In reliability theory, to describe and study the information associated with a non-negative absolutely continuous random variable X, we use the Shannon entropy, or differential entropy of X, defined by Shannon (1948)

$$H(X) = -\mathbb{E}[\log(X)] = -\int_0^{+\infty} f(x) \log f(x) \mathrm{d}x,$$

where log is the natural logarithm and f is the probability density function (pdf) of X. In the following, we use F and $\overline{F}$ to indicate the cumulative distribution function (cdf) and the survival function (sf) of X, respectively.

✉ Maria Longobardi
maria.longobardi@unina.it

Narayanaswamy Balakrishnan
bala@mcmaster.ca

Francesco Buono
francesco.buono3@unina.it

[1] McMaster University, Hamilton, Ontario, Canada

[2] Università degli Studi di Napoli Federico II, Naples, Italy

In the literature, there are several different versions of entropy, each one suitable for a specific situation. In Rao et al. (2004) introduced the Cumulative Residual Entropy (CRE) of X as

$$\mathcal{E}(X) = -\int_0^{+\infty} \overline{F}(x) \log \overline{F}(x) \mathrm{d}x.$$

Di Crescenzo and Longobardi (2009) introduced the Cumulative Entropy (CE) of X as

$$\mathcal{CE}(X) = -\int_0^{+\infty} F(x) \log F(x) \mathrm{d}x. \tag{1}$$

If the random variable X is a lifetime of a system, this information measure is suitable when uncertainty is related to the past. It is a concept dual to the cumulative residual entropy which relates uncertainty on the future lifetime of the system.

Mirali et al. (2016) introduced the Weighted Cumulative Residual Entropy (WCRE) of X as

$$\mathcal{E}^w(X) = -\int_0^{+\infty} x(1 - F(x)) \log(1 - F(x)) \mathrm{d}x.$$

Mirali and Baratpour (2017) introduced the Weighted Cumulative Entropy (WCE) of X as

$$\mathcal{CE}^w(X) = -\int_0^{+\infty} x F(x) \log F(x) \mathrm{d}x.$$

It should be mentioned that the above measures can also be defined for random variables X with support over the entire real line provided the involved integrals exist (as in some examples discussed in later sections).

Recently, various authors have discussed different versions of entropy and their applications (see, for instance, Cali et al. 2017, 2019, 2020; Longobardi 2014).

The paper is organized as follows. In Section 2, we study relationships between some kinds of entropies and moments of order statistics and present various illustrative examples. In Section 3, bounds are given by using some characterizations and properties (as the symmetry) of the random variable X, and some examples and bounds for a few well-known distributions are also presented. In Section 4, we present a method based on moments of order statistics to estimate the cumulative entropies and an application to testing exponentiality of data is also outlined.

2 Relationships Between Entropies and Order Statistics

We recall that, if we have n i.i.d. random variables $X_1, \ldots, X_n$, we can introduce the order statistics $X_{k:n}, k = 1, \ldots, n$. The k-th order statistic is equal to the k-th smallest value from the sample. We know that the cdf of $X_{k:n}$ can be given in terms of the cdf of the parent distribution as

$$F_{k:n}(x) = \sum_{j=k}^{n} \binom{n}{j} [F(x)]^j [1 - F(x)]^{n-j},$$

while the pdf of $X_{k:n}$ is

$$f_{k:n}(x) = \binom{n}{k} k [F(x)]^{k-1} [1 - F(x)]^{n-k} f(x).$$

Choosing $k = 1$ and $k = n$, we get the smallest and largest order statistics, respectively. Their cdf and pdf are given by

$$F_{1:n}(x) = 1 - [1 - F(x)]^n, \qquad f_{1:n}(x) = n[1 - F(x)]^{n-1} f(x),$$
$$F_{n:n}(x) = [F(x)]^n, \qquad f_{n:n}(x) = n[F(x)]^{n-1} f(x).$$

In the following, we denote by μ and $\mu_{k:n}$ the expectation or mean of X and the mean of the k-th order statistic in a sample of size n with parent distribution as that of X, respectively, i.e., $\mu = \mathbb{E}(X)$ and $\mu_{k:n} = \mathbb{E}(X_{k:n})$. Moreover, we denote by $\mu^{(2)}$ and $\mu_{k:n}^{(2)}$ the second moment of X and the second moment of k-th order statistic in a sample of size n with parent distribution as that of X, respectively, i.e., $\mu^{(2)} = \mathbb{E}(X^2)$ and $\mu_{k:n}^{(2)} = \mathbb{E}(X_{k:n}^2)$. We recall that if X is a random variable with finite expectation μ, then the first moment of all order statistics is finite, and if the second moment of X is finite, then the second moment of all order statistics is also finite; see David and Nagaraja (2003) for further details.

2.1 Cumulative Residual Entropy

Let X be a random variable with finite expectation μ. The Cumulative Residual Entropy (CRE) of X can also be written in terms of order statistics as follows:

$$
\begin{aligned}
\mathcal{E}(X) &= -\int_0^{+\infty} (1 - F(x)) \log(1 - F(x)) \mathrm{d}x \\
&= -x(1 - F(x)) \log(1 - F(x))\big|_0^{+\infty} - \int_0^{+\infty} x \log(1 - F(x)) f(x) \mathrm{d}x \\
&\quad - \int_0^{+\infty} x f(x) \mathrm{d}x \\
&= \int_0^{+\infty} x[-\log(1 - F(x))] f(x) \mathrm{d}x - \mu \\
&= \int_0^{+\infty} x \left[\sum_{n=1}^{+\infty} \frac{F(x)^n}{n} \right] f(x) \mathrm{d}x - \mu \\
&= \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \mu_{n+1:n+1} - \mu,
\end{aligned}
\tag{2}
$$

provided that $\lim_{x \to +\infty} -x(1 - F(x)) \log(1 - F(x))$ exists and CRE is finite. In this case, the previous limit is equal to 0. We note that Eq. 2 can be rewritten as

$$\mathcal{E}(X) = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) \mu_{n+1:n+1} - \mu. \tag{3}$$

Remark 1 We want to emphasize that, under the assumptions made, the steps in Eq. 2 are correct. The improper integral can be written as

$$\lim_{t \to +\infty} \int_0^t x \lim_{N \to +\infty} \sum_{n=1}^N \frac{F(x)^n}{n} f(x) \mathrm{d}x. \tag{4}$$

Hence, we observe that the sequence $S_N(x) = \sum_{n=1}^N \frac{F(x)^n}{n}$ is increasing and converges pointwise to the continuous function $-\log(1 - F(x))$ for each $x \in [0, t]$ and, by applying Dini's theorem for uniform convergence (Bartle and Sherbert 2000), the convergence is

uniform. Then, Eq. 4 can be written as

$$\lim_{t\to+\infty}\lim_{N\to+\infty}\sum_{n=1}^{N}\int_{0}^{t}x\frac{F(x)^n}{n}f(x)\mathrm{d}x. \tag{5}$$

In order to apply Moore–Osgood theorem for the iterated limit (Taylor 2010), we have to show that

$$\lim_{t\to+\infty}\sum_{n=1}^{N}\int_{0}^{t}x\frac{F(x)^n}{n}f(x)\mathrm{d}x = \sum_{n=1}^{N}\frac{1}{n(n+1)}\mu_{n+1:n+1}$$

converges pointwise for each fixed N, and this is satisfied if X has finite mean. Hence, by applying Moore–Osgood theorem for the iterated limit, Eq. 5 can be written as

$$\lim_{N\to+\infty}\lim_{t\to+\infty}\sum_{n=1}^{N}\int_{0}^{t}x\frac{F(x)^n}{n}f(x)\mathrm{d}x = \sum_{n=1}^{+\infty}\frac{1}{n(n+1)}\mu_{n+1:n+1}.$$

In the following examples, we use Eq. 3 to evaluate the CRE for the standard exponential and uniform distributions.

Example 1 Consider the standard exponential distribution with pdf $f(x) = e^{-x}$, $x > 0$. Then, it is known that

$$\mu = 1 \text{ and } \mu_{n:n} = 1 + \frac{1}{2} + \cdots + \frac{1}{n};$$

see Arnold and Balakrishnan (1989) for further details. Then, from Eq. 3, we readily have

$$\begin{aligned}
\mathcal{E}(X) &= \left(1-\frac{1}{2}\right)\mu_{2:2} + \left(\frac{1}{2}-\frac{1}{3}\right)\mu_{3:3} + \left(\frac{1}{3}-\frac{1}{4}\right)\mu_{4:4} + \cdots - \mu \\
&= (\mu_{2:2}-\mu) + \frac{1}{2}(\mu_{3:3}-\mu_{2:2}) + \frac{1}{3}(\mu_{4:4}-\mu_{3:3}) + \ldots \\
&= \frac{1}{2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots \\
&= \sum_{n=1}^{+\infty}\frac{1}{n(n+1)} = 1.
\end{aligned}$$

Example 2 Consider the standard uniform distribution with pdf $f(x) = 1$, $0 < x < 1$. Then, it is known that

$$\mu = \frac{1}{2} \text{ and } \mu_{n:n} = \frac{n}{n+1}.$$

So, from Eq. 2, we readily find

$$\begin{aligned}
\mathcal{E}(X) &= \sum_{n=1}^{+\infty}\frac{1}{n(n+1)}\frac{n+1}{n+2} - \frac{1}{2} \\
&= \frac{1}{2}\sum_{n=1}^{+\infty}\left(\frac{1}{n}-\frac{1}{n+2}\right) - \frac{1}{2} \\
&= \frac{1}{2}\left(1+\frac{1}{2}\right) - \frac{1}{2} = \frac{1}{4}.
\end{aligned}$$

2.2 Cumulative Entropy

Let X be a random variable with finite expectation μ. The Cumulative Entropy (CE) of X can also be rewritten in terms of the mean of the minimum order statistic; in fact, from Eq. 1, we easily obtain

$$
\begin{aligned}
\mathcal{CE}(X) &= -xF(x)\log F(x)\big|_0^{+\infty} + \int_0^{+\infty} x\log F(x) f(x)\mathrm{d}x + \int_0^{+\infty} xf(x)\mathrm{d}x \\
&= \int_0^{+\infty} x\log[1-(1-F(x))]f(x)\mathrm{d}x + \mu \\
&= -\int_0^{+\infty} x\sum_{n=1}^{+\infty}\frac{(1-F(x))^n}{n}f(x)\mathrm{d}x + \mu \\
&= -\sum_{n=1}^{+\infty}\frac{1}{n(n+1)}\mu_{1:n+1} + \mu,
\end{aligned}
\tag{6}
$$

provided that $\lim_{x\to+\infty} -xF(x)\log F(x)$ exists and CE is finite. We note that Eq. 6 can be rewritten as

$$
\mathcal{CE}(X) = -\sum_{n=1}^{+\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)\mu_{1:n+1} + \mu.
\tag{7}
$$

In the following examples, we give an application of Eq. 7 to the standard exponential and uniform distributions.

Example 3 For the standard exponential distribution, it is known that

$$
\mu_{1:n} = \frac{1}{n},
$$

and so from Eq. 7, we readily have

$$
\begin{aligned}
\mathcal{CE}(X) &= -\sum_{n=1}^{+\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)\frac{1}{n+1} + 1 \\
&= -\sum_{n=1}^{+\infty}\frac{1}{n(n+1)} + \sum_{n=1}^{+\infty}\frac{1}{(n+1)^2} + 1 \\
&= \frac{\pi^2}{6} - 1,
\end{aligned}
$$

by the use of Euler's identity.

Example 4 For the standard uniform distribution, using the fact that

$$
\mu_{1:n} = \frac{1}{n+1},
$$

we obtain from Eq. 6 that

$$
\begin{aligned}
\mathcal{CE}(X) &= -\sum_{n=1}^{+\infty} \frac{1}{n(n+1)(n+2)} + \frac{1}{2} \\
&= \frac{1}{2} - \frac{1}{2}\sum_{n=1}^{+\infty}\left[\frac{1}{n} - \frac{2}{n+1} + \frac{1}{n+2}\right] \\
&= \frac{1}{2} - \frac{1}{2}\sum_{n=1}^{+\infty}\left[\frac{1}{n} - \frac{1}{n+1}\right] + \frac{1}{2}\sum_{n=1}^{+\infty}\left[\frac{1}{n+1} - \frac{1}{n+2}\right] \\
&= \frac{1}{4}
\end{aligned}
$$

by the use of Euler's identity.

Remark 2 If the random variable X has finite mean μ and is symmetrically distributed about μ, then it is known that

$$
\mu_{n:n} - \mu = \mu - \mu_{1:n},
$$

and so the equality $\mathcal{E}(X) = \mathcal{CE}(X)$ readily follows.

2.3 Weighted Cumulative Entropies

In a similar manner, the Weighted Cumulative Residual Entropy (WCRE) of X, with finite second moment, can be expressed as

$$
\begin{aligned}
\mathcal{E}^w(X) &= -\frac{x^2}{2}(1 - F(x))\log(1 - F(x))\Big|_0^{+\infty} - \frac{1}{2}\int_0^{+\infty} x^2\log(1 - F(x))f(x)\,\mathrm{d}x \\
&\quad - \frac{1}{2}\int_0^{+\infty} x^2 f(x)\,\mathrm{d}x \\
&= \frac{1}{2}\int_0^{+\infty} x^2[-\log(1 - F(x))]f(x)\,\mathrm{d}x - \frac{\mu^{(2)}}{2} \\
&= \frac{1}{2}\int_0^{+\infty} x^2\left[\sum_{n=1}^{+\infty}\frac{F(x)^n}{n}\right]f(x)\,\mathrm{d}x - \frac{\mu^{(2)}}{2} \\
&= \frac{1}{2}\sum_{n=1}^{+\infty}\frac{1}{n(n+1)}\mu_{n+1:n+1}^{(2)} - \frac{\mu^{(2)}}{2},
\end{aligned} \tag{8}
$$

provided that $\lim_{x\to+\infty} -\frac{x^2}{2}(1 - F(x))\log(1 - F(x))$ exists and WCRE is finite.

In the following example, we use Eq. 8 to evaluate the WCRE for the standard uniform distribution.

Example 5 For the standard uniform distribution, using the fact that

$$
\mu_{n+1:n+1}^{(2)} = \frac{n+1}{n+3} \quad \text{and} \quad \mu^{(2)} = \frac{1}{3},
$$

we obtain from Eq. 8 that

$$
\begin{aligned}
\mathcal{E}^w(X) &= \frac{1}{2} \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \frac{n+1}{n+3} - \frac{1}{6} \\
&= \frac{1}{6} \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+3} \right) - \frac{1}{6} \\
&= \frac{1}{6} \left(1 + \frac{1}{2} + \frac{1}{3} \right) - \frac{1}{6} = \frac{5}{36}.
\end{aligned}
$$

Moreover, we can derive the Weighted Cumulative Entropy (WCE) of X in terms of the second moment of the minimum order statistic as follows:

$$
\begin{aligned}
\mathcal{CE}^w(X) &= -\frac{x^2}{2} F(x) \log F(x) \Big|_0^{+\infty} + \frac{1}{2} \int_0^{+\infty} x^2 \log F(x) f(x) \mathrm{d}x \\
&\quad + \frac{1}{2} \int_0^{+\infty} x^2 f(x) \mathrm{d}x \\
&= \frac{1}{2} \int_0^{+\infty} x^2 \log[1 - (1 - F(x))] f(x) \mathrm{d}x + \frac{\mu^{(2)}}{2} \\
&= -\frac{1}{2} \int_0^{+\infty} x^2 \sum_{n=1}^{+\infty} \frac{(1 - F(x))^n}{n} f(x) \mathrm{d}x + \frac{\mu^{(2)}}{2} \\
&= -\frac{1}{2} \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \mu_{1:n+1}^{(2)} + \frac{\mu^{(2)}}{2},
\end{aligned}
\tag{9}
$$

provided that $\lim_{x \to +\infty} -\frac{x^2}{2} F(x) \log F(x)$ exists and WCE is finite.

3 Bounds

In the following, we use Z to denote the standard version of the random variable X, i.e.,

$$
Z = \frac{X - \mu}{\sigma},
$$

where σ is the standard deviation of X. By construction, the relation between a variable and its standard version holds for order statistics and so we have

$$
Z_{k:n} = \frac{X_{k:n} - \mu}{\sigma},
$$

for $k = 1, \ldots, n$. Hence, the mean of $X_{k:n}$ and the mean of $Z_{k:n}$ are directly related and, in particular, for the largest order statistic, we have

$$
\mathbb{E}(Z_{n:n}) = \frac{\mu_{n:n} - \mu}{\sigma}.
$$

We remark that this formula also holds by considering a generalization of the random variable Z with an arbitrary location parameter μ in place of the mean, and an arbitrary scale parameter σ in place of the standard deviation.

Let us consider a sample with parent distribution Z such that $\mathbb{E}(Z) = 0$ and $\mathbb{E}(Z^2) = 1$. Hartley and David (1954) and Gumbel (1954) have then shown that

$$\mathbb{E}(Z_{n:n}) \leq \frac{n-1}{\sqrt{2n-1}}.$$

Using the Hartley-David-Gumbel bound for a parent distribution with mean μ and variance σ^2, we get

$$\mu_{n:n} = \sigma \mathbb{E}(Z_{n:n}) + \mu \leq \sigma \frac{n-1}{\sqrt{2n-1}} + \mu. \qquad (10)$$

Theorem 1 *Let X be a random variable with mean μ and variance σ^2. Then, we obtain an upper bound for the CRE of X as*

$$\mathcal{E}(X) \leq \sum_{n=1}^{+\infty} \frac{\sigma}{(n+1)\sqrt{2n+1}} \simeq 1.21\,\sigma. \qquad (11)$$

Proof From Eqs. 2 and 10, we get

$$\begin{aligned}
\mathcal{E}(X) &= \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \mu_{n+1:n+1} - \mu \\
&\leq \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \left(\sigma \frac{n}{\sqrt{2n+1}} + \mu \right) - \mu \\
&= \sum_{n=1}^{+\infty} \frac{\sigma}{(n+1)\sqrt{2n+1}} \simeq 1.21\,\sigma,
\end{aligned}$$

which is the upper bound given in Eq. 11. □

Remark 3 If X is a non-negative random variable, we have $\mu_{n+1:n+1} \geq 0$, for all $n \in \mathbb{N}$. For this reason, using finite series approximations in Eq. 2, we get lower bounds for $\mathcal{E}(X)$ as

$$\mathcal{E}(X) \geq \sum_{n=1}^{m} \frac{1}{n(n+1)} \mu_{n+1:n+1} - \mu, \qquad (12)$$

for all $m \in \mathbb{N}$.

Remark 4 If X is a non-negative random variable, we have $\mu_{1:n+1} \geq 0$, for all $n \in \mathbb{N}$. For this reason, using finite series approximations in Eq. 6, we get upper bounds for $\mathcal{CE}(X)$ as

$$\mathcal{CE}(X) \leq -\sum_{n=1}^{m} \frac{1}{n(n+1)} \mu_{1:n+1} + \mu,$$

for all $m \in \mathbb{N}$.

In the following theorem, we consider a decreasing failure rate distribution (DFR). We recall that the failure rate, or hazard rate, function of X, $r(x)$, is defined as

$$r(x) = \lim_{\Delta x \to 0^+} \frac{\mathbb{P}(x < X \leq x + \Delta x | X > x)}{\Delta x}$$

$$= \frac{1}{1 - F(x)} \lim_{\Delta x \to 0^+} \frac{\mathbb{P}(x < X \leq x + \Delta x)}{\Delta x} = \frac{f(x)}{1 - F(x)},$$

where the last equality occurs under the assumption of absolute continuity. Then, the random variable X is said to be DFR when the function $r(x)$ is decreasing in x; see Barlow and Proschan (1996) for further details on hazard rate functions and DFR distributions.

Theorem 2 *Let X be DFR. Then, we have the following lower bound for $\mathcal{CE}(X)$:*

$$\mathcal{CE}(X) \geq \mu - \sqrt{\frac{\mu^{(2)}}{2} \left(2 - \frac{\pi^2}{6} \right)}. \tag{13}$$

Proof Let X be DFR. From Theorem 12 of Rychlik (2001), it is known that for a sample of size n, if

$$\delta_{j,n} = \sum_{k=1}^{j} \frac{1}{n+1-k} \leq 2 \quad j \in \{1, \ldots, n\},$$

then

$$\mu_{j:n} \leq \frac{\delta_{j,n}}{\sqrt{2}} \sqrt{\mu^{(2)}}.$$

For $j = 1$, we have $\delta_{1,n} = \frac{1}{n} \leq 2$ for all $n \in \mathbb{N}$, so that

$$\mu_{1:n} \leq \frac{\sqrt{\mu^{(2)}}}{\sqrt{2}\,n}.$$

Then, from Eq. 6, we get the following lower bound for $\mathcal{CE}(X)$:

$$\mathcal{CE}(X) \geq -\sum_{n=1}^{+\infty} \frac{1}{n(n+1)^2} \sqrt{\frac{\mu^{(2)}}{2}} + \mu$$

$$= \mu - \sqrt{\frac{\mu^{(2)}}{2} \left(2 - \frac{\pi^2}{6} \right)}.$$

□

Remark 5 We note that we can not provide an analogous bound for $\mathcal{E}(X)$ because $\delta_{n,n} \leq 2$ is not fulfilled for $n \geq 4$.

From Eqs. 2 and 6, we get the following expression for the sum of the cumulative residual entropy and the cumulative entropy:

$$\mathcal{E}(X) + \mathcal{CE}(X) = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} (\mu_{n+1:n+1} - \mu_{1:n+1}). \tag{14}$$

Calì et al. (2017) have shown a connection between (14) and the partition entropy studied by Bowden (2007).

Theorem 3 *We have the following bound for the sum of the CRE and the CE:*

$$\mathcal{E}(X) + \mathcal{CE}(X) \leq \sum_{n=1}^{+\infty} \frac{\sqrt{2}\,\sigma}{n\sqrt{n+1}} \simeq 3.09\,\sigma. \tag{15}$$

Proof From Theorem 3.24 of Arnold and Balakrishnan (1989), it is known that the bound for the difference between the expectations of the largest and smallest order statistics from a sample of size $n + 1$ is

$$\mu_{n+1:n+1} - \mu_{1:n+1} \leq \sigma\sqrt{2(n+1)}, \tag{16}$$

and so using Eq. 16 in Eq. 14, we get the following bound for the sum of the CRE and the CE:

$$\mathcal{E}(X) + \mathcal{CE}(X) \leq \sum_{n=1}^{+\infty} \frac{\sigma\sqrt{2(n+1)}}{n(n+1)} = \sum_{n=1}^{+\infty} \frac{\sqrt{2}\,\sigma}{n\sqrt{n+1}} \simeq 3.09\,\sigma,$$

as required.   □

In Table 1, we present some of the bounds obtained in this section for a few know distributions.

3.1 Symmetric Distributions

In this subsection, we obtain bounds for a symmetric parent distribution. David and Nagaraja (2003) have stated that if we have a sample $Z_1, \ldots, Z_n$ with parent distribution Z symmetric about 0 and with variance 1, then

$$\mathbb{E}(Z_{n:n}) \leq \frac{1}{2}\,nc(n), \tag{17}$$

where

$$c(n) = \left[\frac{2\left(1 - \frac{1}{\binom{2n-2}{n-1}}\right)}{2n-1} \right]^{\frac{1}{2}}.$$

Using the bound in Eq. 17 for a random variable X symmetric about mean μ and with variance σ^2, we have

$$\mu_{n:n} = \sigma\mathbb{E}(Z_{n:n}) + \mu \leq \frac{1}{2}\sigma nc(n) + \mu. \tag{18}$$

Table 1 Some bounds for known distributions

Support	CDF	$\mathcal{E}(X)$	Bound Thm.1	$\mathcal{E}(X) + \mathcal{CE}(X)$	Bound Thm.3
$x > 0$	$1 - \exp(-\lambda x)$	$\frac{1}{\lambda}$	$\frac{1.21}{\lambda}$	$\frac{\pi^2}{6\lambda}$	$\frac{3.09}{\lambda}$
$0 < x < a$	$\frac{x}{a}$	$\frac{a}{4}$	$\frac{1.21\,a}{2\sqrt{3}}$	$\frac{a}{2}$	$\frac{3.09\,a}{2\sqrt{3}}$
$x \in (0,1)$	$\frac{1}{x^2}\exp\left(2\left(1 - \frac{1}{x}\right)\right)$	0.1549	0.1999	0.2936	0.5105
$x > 0$	$1 - \frac{1}{(x+1)^3}$	0.75	1.0479	1.1115	2.6759
$x \in (0,1)$	x^2	0.1869	0.2852	0.4091	0.7283
$x \in (0,+\infty)$	$\exp\left(-\frac{1}{\exp(x)-1}\right)$	0.9283	1.1238	1.5246	2.87

Remark 6 It is clear that a non-negative random variable X that is symmetric about the mean has a bounded support.

Theorem 4 *Let X be a symmetric random variable with mean μ and variance σ^2. Then, an upper bound for the CRE of X is given by*

$$\mathcal{E}(X) \leq \frac{\sigma}{2} \sum_{n=1}^{+\infty} \frac{c(n+1)}{n}. \tag{19}$$

Proof From Eqs. 2 and 18, we get

$$
\begin{aligned}
\mathcal{E}(X) &= \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \mu_{n+1:n+1} - \mu \\
&\leq \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \left(\frac{1}{2}\sigma\ c(n+1)(n+1) + \mu \right) - \mu \\
&= \frac{\sigma}{2} \sum_{n=1}^{+\infty} \frac{c(n+1)}{n},
\end{aligned}
$$

which is the upper bound given in Eq. 19.   □

The bound in Eq. 17 is equivalent to (see Arnold and Balakrishnan 1989)

$$\mathbb{E}(Z_{n:n}) \leq \frac{n}{\sqrt{2}} \sqrt{\frac{1}{2n-1} - B(n,n)}, \tag{20}$$

where $B(n,n)$ is the complete beta function defined as

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

where $\alpha, \beta > 0$ (see Abramowitz and Stegun (1964) for further details). In fact

$$
\begin{aligned}
\frac{1}{2}nc(n) &= \frac{1}{2}n \left[\frac{2\left(1 - \frac{1}{\binom{2n-2}{n-1}}\right)}{2n-1} \right]^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2}}n \left[\frac{1 - \frac{(n-1)!(n-1)!}{(2n-2)!}}{2n-1} \right]^{\frac{1}{2}} \\
&= \frac{n}{\sqrt{2}} \left[\frac{1}{2n-1} - \frac{(n-1)!(n-1)!}{(2n-1)!} \right]^{\frac{1}{2}} \\
&= \frac{n}{\sqrt{2}} \left[\frac{1}{2n-1} - B(n,n) \right]^{\frac{1}{2}}.
\end{aligned}
$$

Then, the bound in Eq. 18 is equivalent to

$$\mu_{n:n} = \sigma \mathbb{E}(Z_{n:n}) + \mu \leq \sigma \frac{n}{\sqrt{2}} \sqrt{\frac{1}{2n-1} - B(n,n)} + \mu. \tag{21}$$

Hence, we have the following theorem which is equivalent to Theorem 4.

Theorem 5 *Let X be a symmetric random variable with mean μ and variance σ^2. Then, an upper bound for the CRE of X is*

$$\mathcal{E}(X) \leq \frac{\sigma}{\sqrt{2}} \sum_{n=1}^{+\infty} \frac{1}{n} \sqrt{\frac{1}{2n+1} - B(n+1, n+1)}. \tag{22}$$

Example 6 Let us consider a sample with parent distribution for $X \sim N(0, 1)$. From Harter (1961), we have the values of means of largest order statistics for samples of size up to 100. Hence, we compare the finite series approximations of Eqs. 2 and 22 and we expect the same result as the true value because truncated terms are negligible. We thus get the following result:

$$\mathcal{E}(X) \simeq \sum_{n=1}^{99} \frac{1}{n(n+1)} \mu_{n+1:n+1} \simeq 0.87486$$

$$< \frac{1}{\sqrt{2}} \sum_{n=1}^{99} \frac{1}{n} \sqrt{\frac{1}{2n+1} - B(n+1, n+1)} \simeq 0.94050,$$

computed by truncating the series in Theorem 5.

For a symmetric distribution, Arnold and Balakrishnan (1989) have stated that if we have a sample $Z_1, \ldots, Z_n$ from a parent distribution symmetric about mean μ and with variance 1, then

$$\mathbb{E}(Z_{n:n}) - \mathbb{E}(Z_{1:n}) \leq n\sqrt{2} \sqrt{\frac{1}{2n-1} - B(n, n)}, \tag{23}$$

where $B(n, n)$ is the complete beta function.

Using the bound in Eq. 23 for a parent distribution symmetric about mean μ and with variance σ^2, we have

$$\mu_{n:n} - \mu_{1:n} = \sigma \left(\mathbb{E}(Z_{n:n}) - \mathbb{E}(Z_{1:n}) \right) \leq \sigma n\sqrt{2} \sqrt{\frac{1}{2n-1} - B(n, n)}. \tag{24}$$

Theorem 6 *Let X be a symmetric random variable with mean μ and variance σ^2. Then, an upper bound for the sum of the CRE and the CE of X is given by*

$$\mathcal{E}(X) + \mathcal{CE}(X) \leq \sqrt{2}\, \sigma \sum_{n=1}^{+\infty} \frac{1}{n} \sqrt{\frac{1}{2n+1} - B(n+1, n+1)}. \tag{25}$$

Proof From Eqs. 14 and 24, we have

$$\mathcal{E}(X) + \mathcal{CE}(X) = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} (\mu_{n+1:n+1} - \mu_{1:n+1})$$

$$\leq \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \left(\sigma(n+1)\sqrt{2} \sqrt{\frac{1}{2n+1} - B(n+1, n+1)} \right)$$

$$= \sqrt{2}\, \sigma \sum_{n=1}^{+\infty} \frac{1}{n} \sqrt{\frac{1}{2n+1} - B(n+1, n+1)},$$

which is the upper bound given in Eq. 25.   □

4 Computation and Application

4.1 Numerical Evaluation

In this subsection, we present a method for the evaluation of an estimate of cumulative entropies based on the relationships with moments of order statistics derived in Section 2. The method is based on finite series approximations. About the CRE, we will use the result shown in Eq. 12:

$$\mathcal{E}(X) \geq \sum_{n=1}^{m} \frac{1}{n(n+1)} \mu_{n+1:n+1} - \mu,$$

where the difference between LHS and RHS is infinitesimal for $m \to +\infty$. Starting from a sample $\underline{X} = (X_1, \ldots, X_N)$, we estimate the CRE by

$$\hat{\mathcal{E}}(X) = \sum_{n=1}^{m} \frac{1}{n(n+1)} \hat{\mu}_{n+1:n+1} - \hat{\mu}, \tag{26}$$

where $\hat{\mu}$ is the arithmetic mean of the realization of the sample and $\hat{\mu}_{n+1:n+1}$ is an estimate for the mean of largest order statistic in a sample of size $n + 1$. In order to obtain an estimate of the mean of largest order statistic in a sample of size $n + 1$, we generate several simple random samples starting from $\underline{X}$ of size $n + 1$, for example, by using the function `randsample` of *MATLAB*. Then, we consider the maximum of each sample and by the mean we obtain $\hat{\mu}_{n+1:n+1}$. Specifically, for a fixed $n \in \{1, \ldots, m\}$, starting from $\underline{X}$, we generate K simple random samples of size $n + 1$, $(X_{1,k}, \ldots, X_{n+1,k})$, $k = 1, \ldots, K$. For each of them, we can consider the maximum, i.e., $X_{n+1:n+1,k}$, $k = 1, \ldots, K$. Then, we take as estimate of the mean of the maximum order statistic in a sample of size $n + 1$ from $\underline{X}$ by

$$\hat{\mu}_{n+1:n+1} = \frac{1}{K} \sum_{k=1}^{K} X_{n+1:n+1,k}.$$

Using similar techniques, we can get estimates for the second moment of the largest order statistic and the first and second moments of the smallest order statistic and, in analogy with Eq. 26, we obtain estimates for the other cumulative entropies, $\hat{\mathcal{CE}}(X), \hat{\mathcal{E}}^w(X), \hat{\mathcal{CE}}^w(X)$.

Example 7 By using the function `exprnd` of *MATLAB*, we generated a sample of size $N = 1000$ values from an exponential distribution with parameter 1. We obtain estimates for cumulative entropies stopping the sum at $m = 100$. For each sample size n from 2 to 101

Table 2 Approximations of cumulative entropies in the exponential case

Entropy	Theoretical value	Estimator
CRE	1	0.9220
CE	0.6449	0.6331
WCRE	2	2.1119
WCE	0.8470	0.8735

Table 3 Experimental results for testing exponentiality

Distribution	Success with threshold 0.1	Success with threshold 0.25
Exponential(1)	757	999
chi–square(3)	206	620
Weibull(3,4)	0	0

we generate $K = 100$ random samples of size n by the function `randsample`. In Table 2, we compare these results with the corresponding theoretical results.

4.2 Application to Test for Exponentiality

We recall a result proved in Rao et al. (2004) about the maximum cumulative residual entropy for fixed moments of the first and second order.

Theorem 7 *Let X be a non-negative random variable. Then,*

$$\mathcal{E}(X) \leq \mathcal{E}(X(\mu)),$$

where $X(\mu)$ is an exponentially distributed random variable with mean $\mu = \mathbb{E}(X^2)/2\mathbb{E}(X)$.

There are many situations in which we have a sample and we want to investigate about the underlying distribution, and in some of these situations, it may be of interest to test if the data are distributed according to an exponential distribution. In the literature, there are several papers about testing exponentiality of data; see, for example, Balakrishnan et al. (2007).

From Theorem 7, we know that the CRE of X is lower than $\mathbb{E}(X^2)/2\mathbb{E}(X)$ because, for an exponential random variable Y, we have $\mathcal{E}(Y) = \mathbb{E}(Y)$. If we have the realization of a sample $\underline{X} = (X_1, \ldots, X_N)$, we can consider the difference between the estimate of CRE, given in Eq. 26, and $\hat{\mu}^{(2)}/2\hat{\mu}$, where $\hat{\mu}$ and $\hat{\mu}^{(2)}$ are given by the arithmetic means of $X_1, \ldots, X_N$ and $X_1^2, \ldots, X_N^2$, respectively. If this difference is small enough in absolute value, we may suppose that the data are exponentially distributed and, as the difference increases, it becomes less plausible that the sample is exponentially distributed.

We fixed as threshold values 0.1 and 0.25 and we tested data generated in *MATLAB* from different distributions. For each distribution, we generated $K = 1000$ samples of size $N = 1000$, X_k, and the obtained results are presented in Table 3, where the success with threshold α, $S(\alpha)$, is given by

$$S(\alpha) = \#\{k \in \{1, \ldots, K\} : |\hat{\mathcal{E}}(X_k) - \hat{\mu}_k^{(2)}/2\hat{\mu}_k| < \alpha\}.$$

It will, of course, be of interest to develop a formal testing procedure along these lines.

Declarations

Competing interests The authors declare that they have no conflict of interest.

References

Abramowitz M, Stegun I (1964) Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York

Arnold BC, Balakrishnan N (1989) Relations, Bounds and Approximations for Order Statistics. Springer, New York

Balakrishnan N, Habibi Rad A, Arghami NR (2007) Testing exponentiality based on Kullback-Leibler information with progressively Type-II censored data. IEEE Trans Reliab 56:301–307

Barlow RE, Proschan FJ (1996) Mathematical Theory of Reliability. Society for Industrial and Applied Mathematics, Philadelphia

Bartle RG, Sherbert DR (2000) Introduction to Real Analysis, 3rd edn. Wiley, New York

Bowden R (2007) Information, measure shifts and distribution diagnostics. Statistics 46(2):249–262

Calì C, Longobardi M, Ahmadi J (2017) Some properties of cumulative Tsallis entropy. Physica A 486:1012–1021

Calì C, Longobardi M, Psarrakos G (2019) A family of weighted distributions based on the mean inactivity time and cumulative past entropies. Ricerche Mat. (in press). https://doi.org/10.1007/s11587-019-00475-7

Calì C, Longobardi M, Navarro J (2020) Properties for generalized cumulative past measures of information. Probab Eng Inform Sci 34:92–111

David HA, Nagaraja HN (2003) Order Statistics, 3rd edn. Wiley, Hoboken

Di Crescenzo A, Longobardi M (2009) On cumulative entropies. J Stat Plann Inference 139:4072–4087

Gumbel EJ (1954) The maxima of the mean largest value and of the range. Ann Math Stat 25:76–84

Harter HL (1961) Expected values of normal order statistics. Biometrika 48(1):151–165

Hartley HO, David HA (1954) Universal bounds for mean range and extreme observation. Ann Math Stat 25:85–99

Longobardi M (2014) Cumulative measures of information and stochastic orders. Ricerche Mat 63:209–223

Mirali M, Baratpour S, Fakoor V (2016) On weighted cumulative residual entropy. Commun Stat - Theory Methods 46(6):2857–2869

Mirali M, Baratpour S (2017) Some results on weighted cumulative entropy. J Iran Stat Soc 16(2):21–32

Rao M, Chen Y, Vemuri B, Wang F (2004) Cumulative residual entropy: A new measure of information. IEEE Trans Inf Theory 50:1220–1228

Rychlik T (2001) Projecting Statistical Functionals. Springer, New York

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423

Taylor AE (2010) General Theory of Functions and Integration. Dover Publications, Mineola

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.