POLITECNICO DI TORINO
Repository ISTITUZIONALE

Cumulative entropies and sums of moments of order statistics

(Article begins on next page)

22 November 2024
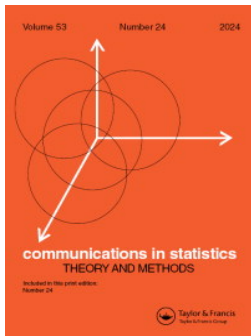
Cumulative entropies and sums of moments of order statistics

Francesco Buono, Udo Kamps & Maria Kateri

Published online: 28 Aug 2024.

Submit your article to this journal

Article views: 180

View related articles

View Crossmark data

Cumulative entropies and sums of moments of order statistics

Francesco Buono, Udo Kamps, and Maria Kateri

Institute of Statistics, RWTH Aachen University, Aachen, Germany

ABSTRACT
Generalized weighted cumulative residual entropies and generalized weighted cumulative entropies are represented by means of weighted sums of moments of upper and lower order statistics, respectively. A variety of examples is shown by specifying the weight function.

1. Introduction

Several notions of entropy have been discussed in the literature. In reliability theory, the Cumulative Residual Entropy (CRE) of a nonnegative random variable X with distribution function (cdf) F given by

$$\mathcal{E}(X) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dx \tag{1}$$

with $\bar{F} = 1 - F$ (see Rao et al. 2004) and the Cumulative Entropy (CE) of X given by

$$\mathcal{CE}(X) = -\int_0^\infty F(x) \log F(x) dx \tag{2}$$

see Di Crescenzo and Longobardi (2009) are of special interest. Throughout, a term $0 \log 0$ is defined to be 0. Weighted versions of these entropies of the form

$$-\int_0^\infty x\bar{F}(x) \log \bar{F}(x) dx \quad \text{and} \quad -\int_0^\infty xF(x) \log F(x) dx \tag{3}$$

have been introduced by Misagh et al. (2011).

In Balakrishnan, Buono, and Longobardi (2022), the above entropies are represented by sums of expectations of maxima and minima of random variables within a sequence of independent and identically distributed (iid), nonnegative random variables with absolutely continuous distribution function F.

We consider generalized weighted versions of (1) and (2) with an increasing nonnegative function φ, $\varphi(0) = 0$, of the form

CONTACT Maria Kateri ✉ maria.kateri@rwth-aachen.de ⬛ Institute of Statistics, RWTH Aachen University, 52056 Aachen, Germany.

$$\mathcal{E}_\varphi(X) = -\int_0^\infty \bar{F}^r(x) \log \bar{F}(x) d\varphi(x) \quad \text{and} \tag{4}$$

$$\mathcal{CE}_\varphi(X) = -\int_0^\infty F^r(x) \log F(x) d\varphi(x) \tag{5}$$

for some $r \in \mathbb{N}$, which are mentioned in Tahmasebi (2020) by applying a cumulative residual entropy with a general weight function to a minimum distribution and in Tahmasebi et al. (2020) by considering a generalized cumulative entropy of a maximum distribution. For the definition of other generalizations of cumulative entropy and cumulative residual entropy, we refer to, for example, Khorashadizadeh, Rezaei Roknabadi, and Mohtashami Borzadaran (2013), Tahmasebi and Eskandarzadeh (2017), and Kattumannil, Sreedevi, and Balakrishnan (2022). For $\varphi(x) = x$, $x \geq 0$, (4) and (5) are obtained by (1) and (2) by plugging in the distribution functions $1 - (1 - F)^r$ and F^r of the minimum and the maximum of r iid random variables with distribution function F, respectively, subject to the constant r (cf. Baratpour 2010; Mirali and Baratpour 2017). For $r = 1$ and $\varphi(x) = x$, $x \geq 0$, the quantities (1) and (2), and for $r = 1$ and $\varphi(x) = \frac{1}{2}x^2$, the quantities (3) are obtained.

It turns out that the entropies (4) and (5) can be expressed by means of sums of moments of upper r-th order statistics and lower r-th order statistics, based on the cdf F, respectively. Several particular cases are shown in the sequel.

2. Representation of a generalized weighted cumulative residual entropy

The generalized weighted cumulative residual entropy (4) can be represented via a weighted sum of moments of r-largest order statistics.

Theorem 2.1. *Let F be an absolutely continuous distribution function with density function f and $F(0) = 0$. Let further $(X_i)_{i \in \mathbb{N}}$ be a sequence of iid random variables with distribution function F and let $X_{j,n}$, $1 \leq j \leq n$, $n \in \mathbb{N}$, denote respective order statistics. Moreover, let φ be an increasing, nonnegative function with $\varphi(0) = 0$, and let, for some $r \in \mathbb{N}$, the series of the RHS of (6) be finite. Then*

$$\mathcal{E}_\varphi(X) = -\int_0^\infty \bar{F}^r(x) \log \bar{F}(x) d\varphi(x)$$

$$= \sum_{n=r}^\infty \left((n+1)\binom{n}{r} \right)^{-1} \mathbb{E}\varphi(X_{n-r+2,n+1}) - \frac{1}{r}\mathbb{E}\varphi(X_{1,r}). \tag{6}$$

Proof. Let $\alpha = \sup\{x \geq 0; F(x) = 0\}$ and $\omega = \inf\{x \geq 0; F(x) = 1\}$, $\inf \emptyset = \infty$, denote the lower and the upper endpoint of the support of F, respectively. By using the theorems of Levi and Fubini, we find

$$\sum_{n=r}^\infty \left((n+1)\binom{n}{r} \right)^{-1} \mathbb{E}\varphi(X_{n-r+2,n+1})$$

$$= \sum_{n=r}^\infty \frac{r!(n-r)!}{(n+1)!} \int_\alpha^\omega \varphi(x)(n-r+2)\binom{n+1}{n-r+2} F^{n-r+1}(x)\bar{F}^{r-1}(x)f(x)dx$$

$$= r \int_\alpha^\omega \varphi(x) \left(\sum_{n=r}^\infty \frac{F^{n-r+1}(x)}{n-r+1} \right) \bar{F}^{r-1}(x)dF(x)$$

$$= r \int_0^\alpha \int_\alpha^\omega \bar{F}^{r-1}(x)(-\log \bar{F}(x))dF(x)d\varphi(y) + r \int_\alpha^\omega \int_y^\omega \bar{F}^{r-1}(x)(-\log \bar{F}(x))dF(x)d\varphi(y)$$

$$= r \int_0^\alpha \int_0^1 x^{r-1}(-\log x)dxd\varphi(y) + r \int_\alpha^\omega \int_0^{\bar{F}(y)} x^{r-1}(-\log x)dxd\varphi(y)$$

$$= \frac{1}{r}\varphi(\alpha) - \int_\alpha^\omega \bar{F}^r(y)\left(\log \bar{F}(y) - \frac{1}{r}\right)d\varphi(y)$$

$$= \int_\alpha^\omega \bar{F}^r(y)\log \bar{F}(y)d\varphi(y) + \frac{1}{r}\int_0^\omega \bar{F}^r(y)d\varphi(y)$$

which proves the assertion, since

$$\int_0^\omega \bar{F}^r(y)d\varphi(y) = \mathbb{E}\varphi(X_{1,r}).$$

□

Example 1. Specific choices of the function φ in Theorem 2.1, lead to various relations, as shown in cases *(i)* to *(ix)* presented below.

(i)   For $\varphi(x) = x$ and $r = 1$, (6) reduces to

$$\mathcal{E}(X) = - \int_0^\infty \bar{F}(x)\log \bar{F}(x)dx = \sum_{n=1}^\infty \frac{1}{n(n+1)}\mathbb{E}X_{n+1,n+1} - \mathbb{E}X_1, \qquad (7)$$

which is relation (2) in Balakrishnan, Buono, and Longobardi (2022). In Sordo, Castaño-Martinez, and Pigueiras (2016), the authors introduce the CRE premium principle in insurance mathematics, which, for a risk X with distribution function F, is given by $\mathbb{E}X + \mathcal{E}(X)$.

(ii)   Case *(i)* can be further generalized as follows. Let $\varphi(x) = x^\alpha$, $x \geq 0$, for some $\alpha > 0$. Then Theorem 2.1 provides a relation involving moments of order statistics of order α. In particular, (6) leads to

$$-\alpha \int_0^\infty x^{\alpha-1}\bar{F}^r(x)\log \bar{F}(x)dx = \sum_{n=r}^\infty \left((n+1)\binom{n}{r}\right)^{-1}\mathbb{E}X^\alpha_{n-r+2,n+1} - \frac{1}{r}\mathbb{E}X^\alpha_{1,r}. \quad (8)$$

In Balakrishnan, Buono, and Longobardi (2022, p. 346), the weighted cumulative residual entropy (3) of $X \sim F$ is considered, which coincides with (8) by choosing $\alpha = 2$ and $r = 1$. Then, relation (8) in Balakrishnan, Buono, and Longobardi (2022), that is,

$$-2 \int_0^\infty x\bar{F}(x)\log \bar{F}(x)dx = \sum_{n=1}^\infty \frac{1}{n(n+1)}\mathbb{E}X^2_{n+1,n+1} - \mathbb{E}X^2_1$$

is obtained as a particular case of (6).

(iii) For $\varphi(x) = x$, $x \geq 0$, the cdf F in Balakrishnan, Buono, and Longobardi (2022), (see *(i)*), could be replaced by the distribution $1 - (1-F)^r$ of the minimum of r iid random variables according to F. Then, $\mathbb{E}X_1$ in the RHS of (7) becomes $\mathbb{E}X_{1,r}$ and the expression $\int_0^\infty \bar{F}^r(x)\log \bar{F}(x)dx$ appears in its LHS. Since the latter expression appears in (6) of Theorem 2.1 as well, we derive the following relation for expectations of order statistics

$$\frac{1}{r}\sum_{n=1}^\infty \frac{1}{n(n+1)}\mathbb{E}Y_{n+1,n+1} = \sum_{n=r}^\infty \left((n+1)\binom{n}{r}\right)^{-1}\mathbb{E}X_{n-r+2,n+1} \qquad (9)$$

where $X_{p,n+1}$, $p \leq n+1$, are order statistics from F and $Y_{n+1,n+1}$, $n \in \mathbb{N}$, are maxima from $1 - (1 - F)^r$.

(iv)   Let $F(x) = x$, $x \in (0, 1)$, and $\varphi(x) = x$, $x \geq 0$. Then, we find

$$- \int_0^\infty \bar{F}^r(x) \log \bar{F}(x) dx = - \int_0^1 x^r \log x \, dx = \frac{1}{(r+1)^2}.$$

For an order statistic $U_{p,q}$ from a standard uniform distribution it is well known that $EU_{p,q} = \frac{p}{q+1}$, $1 \leq p \leq q$. This yields the relation

$$\mathcal{E}_\varphi(X) = \frac{1}{(r+1)^2} = \sum_{n=r}^\infty \left((n+1) \binom{n}{r} \right)^{-1} \frac{n-r+2}{n+2} - (r(r+1))^{-1}$$

and thus

$$\sum_{n=r}^\infty \left((n-r+1) \binom{n+2}{n-r+2} \right)^{-1} = \sum_{n=1}^\infty \left(n \binom{n+r+1}{r} \right)^{-1} = \frac{1}{r(r+1)} + \frac{1}{(r+1)^2}.$$

This can be seen directly by using expressions for infinite sums with a reciprocal of a binomial coefficient (cf. Gould 1972).

(v)   Let $F(x) = 1 - e^{-x}$, $x \geq 0$, and $\varphi(x) = x$, $x \geq 0$. In this case, we obtain

$$- \int_0^\infty \bar{F}^r(x) \log \bar{F}(x) dx = \frac{1}{r^2}.$$

Then, making use of the relation $\mathbb{E}X_{p,q} = \sum_{i=q-p+1}^q \frac{1}{i}$, (6) leads to

$$\sum_{n=r}^\infty \left((n+1) \binom{n}{r} \right)^{-1} \sum_{i=r}^{n+1} \frac{1}{i} = \frac{2}{r^2} \quad 1 \leq r \leq n.$$

Thus, Theorem 2.1 yields a summation formula with a binomial coefficient in the denominator.

(vi)   The setting $\varphi \equiv F$, where $\varphi(X_{p,q}) \overset{d}{=} U_{p,q}$ with $U_{p,q}$ as in (iv), leads to the same relation as in (iv).

(vii)   By setting in (6) $\varphi(x) = \mathbb{1}_{[t,\infty)}(x)$, $x \in \mathbb{R}$, $t > 0$, we immediately obtain a relation for survival functions of order statistics based on F:

$$\sum_{n=r}^\infty \left((n+1) \binom{n}{r} \right)^{-1} P(X_{n-r+2,n+1} > t) = \frac{1}{r} P(X_{1,r} > t) - \bar{F}^r(t) \log \bar{F}(t),$$

$$= \bar{F}^r(t) \left(\frac{1}{r} - \log \bar{F}(t) \right), \quad t > 0.$$

(viii)   By choosing φ as $\varphi(x) = e^{tx} - 1$, for $t \in \mathbb{R}$, $x \geq 0$, moment generating functions of upper order statistics appear in the representation of Theorem 2.1.

(ix)   If φ is chosen to be the odds $\varphi(x) = F(x)/\bar{F}(x)$, $x \geq 0$, then $\varphi'(x) = f(x)/\bar{F}^2(x)$ (≥ 0) so that

$$\mathcal{E}_\varphi(X) = - \int_0^\infty \bar{F}^{r-2}(x) \log \bar{F}(x) f(x) dx = \frac{1}{(r-1)^2}, \tag{10}$$

for $r \geq 2$ and $\mathcal{E}_\varphi(X)$ is infinite for $r = 1$. For $r < n$, it can be easily verified that $\mathbb{E}\varphi(X_{r:n}) = \mathbb{E}\left(\frac{F(X_{r:n})}{1-F(X_{r:n})} \right) = \frac{r}{n-r}$, which for $r \geq 2$ gives $\mathbb{E}\varphi(X_{n-r+2:n+1}) = \frac{n-r+2}{r-1}$ and

$\mathbb{E}\varphi(X_{1:r}) = \frac{1}{r-1}$. Thus, for $r \geq 2$, the RHS of (6) reads

$$r \sum_{n=r}^{\infty} \left((n+1)(n+1-r) \binom{n}{r-2} \right)^{-1} - \frac{1}{r(r-1)}.$$

3. Representation of a generalized weighted cumulative entropy

By analogy with Theorem 2.1, we find a relation for $\int_0^{\infty} F^r(x) \log F(x) d\varphi(x)$ by means of sums of expectations of lower order statistics.

Theorem 3.1. *Let the conditions of Theorem 2.1 be given, and let, for some $r \in \mathbb{N}$, the series on the RHS of (11) be finite. Then*

$$\mathcal{CE}_{\varphi}(X) = -\int_0^{\infty} F^r(x) \log F(x) d\varphi(x)$$

$$= -\sum_{n=r}^{\infty} \left((n+1)\binom{n}{r} \right)^{-1} \mathbb{E}\varphi(X_{r,n+1}) + \frac{1}{r}\mathbb{E}\varphi(X_{r,r}). \tag{11}$$

Remark 1. If the random variable X has finite support $(0, 2\mu)$, mean $\mathbb{E}(X) = \mu$ and is symmetrically distributed about it, then, under the assumptions of Theorem 2.1 and Theorem 3.1 and for an absolutely continuous function φ on $(0, 2\mu)$ with φ' being symmetric about μ, the equality $\mathcal{E}_{\varphi}(X) = \mathcal{CE}_{\varphi}(X)$ readily follows. This is a generalization of the well-known equality between cumulative residual entropy and cumulative entropy for symmetric distributions (cf. Balakrishnan, Buono, and Longobardi 2022). In this case,

$$\mathcal{E}_{\varphi}(X) = \mathcal{CE}_{\varphi}(X) = -2 \int_0^{\mu} \bar{F}^r(x) \log \bar{F}(x) d\varphi(x).$$

Some examples of Theorem 3.1 are stated as in Example 1 by specifying the function φ.

Example 2. By analogy with Example 1, specific choices of the function φ in Theorem 3.1 lead to various relations, as shown in cases *(i)* to *(vii)* presented below.

(i) For $\varphi(x) = x$ and $r = 1$, (11) reduces to relation (6) in Balakrishnan, Buono, and Longobardi (2022)

$$-\int_0^{\infty} F(x) \log F(x) dx = -\sum_{n=1}^{\infty} \frac{1}{n(n+1)} \mathbb{E}X_{1,n+1} + \mathbb{E}X_1. \tag{12}$$

(ii) For $\varphi(x) = x^{\alpha}, x \geq 0$, with some $\alpha > 0$, Theorem 3.1 yields a relation involving moments of order statistics of order α. In particular, (11) becomes

$$-\alpha \int_0^{\infty} x^{\alpha-1} F^r(x) \log F(x) dx = -\sum_{n=r}^{\infty} \left((n+1)\binom{n}{r} \right)^{-1} \mathbb{E}X_{r,n+1}^{\alpha} + \frac{1}{r}\mathbb{E}X_{r,r}^{\alpha}. \tag{13}$$

In Balakrishnan, Buono, and Longobardi (2022, p. 346), the weighted cumulativel entropy (3) of $X \sim F$ is seen to coincide with (13) by choosing $\alpha = 2$ and $r = 1$. Then,

relation (9) in Balakrishnan, Buono, and Longobardi (2022), that is,

$$-2\int_0^\infty xF(x)\log F(x)dx = -\sum_{n=1}^\infty \frac{1}{n(n+1)}\mathbb{E}X_{1,n+1}^2 + \mathbb{E}X_1^2$$

follows as a particular case of (11).

(iii) For $\varphi(x) = x, x \geq 0$, the cdf F in Balakrishnan, Buono, and Longobardi (2022), (see *(i)*), could be replaced by the distribution F^r of the maximum of r iid random variables according to F. Then, $\mathbb{E}X_1$ in the RHS of (12) becomes $\mathbb{E}X_{r,r}$ and the expression $\int_0^\infty F^r(x)\log F(x)dx$ appears in its LHS. By analogy with Example 1*(iii)*, relation (11) in Theorem 3.1 yields

$$\frac{1}{r}\sum_{n=1}^\infty \frac{1}{n(n+1)}\mathbb{E}Y_{1,n+1} = \sum_{n=r}^\infty \left((n+1)\binom{n}{r}\right)^{-1}\mathbb{E}X_{r,n+1} \qquad (14)$$

where $X_{p,n+1}, p \leq n+1$, are order statistics from F and $Y_{1,n+1}, n \in \mathbb{N}$, are minima from F^r.

(iv) Let $F(x) = 1 - e^{-x}, x \geq 0$, and $\varphi(x) = x, x \geq 0$. Then, we have

$$-\int_0^\infty F^r(x)\log F(x)dx = -\Psi'(r+1) = \frac{\pi^2}{6} - \sum_{i=1}^r \frac{1}{i^2}$$

where Ψ is the Digamma function. Using the expression for the expectation of an order statistic from a standard exponential random sample (see Example 1*(iv)*), (11) takes the form

$$\sum_{n=r}^\infty \left((n+1)^{-1}\binom{n}{r}^{-1}\sum_{i=n-r+2}^{n+1}\frac{1}{i}\right) = \frac{\pi^2}{6} - \sum_{i=1}^r \frac{1}{i^2} + \frac{1}{r}\sum_{i=1}^r \frac{1}{i} \quad 1 \leq r \leq n.$$

Thus, Theorem 3.1 yields a summation formula involving a reciprocal binomial coefficient.

(v) Let $\varphi \equiv F$. In this case, as recalled in Example 1*(iv)*, $\varphi(X_{p,q}) \stackrel{d}{=} U_{p,q}$. Then

$$\int_0^\infty F^r(z)\log F(z)dF(z) = \int_0^1 x^r\log x\,dx = -\frac{1}{(r+1)^2} \text{ and } \frac{1}{r}\mathbb{E}U_{r,r} = \frac{1}{r+1}.$$

Thus, from (11) we obtain the relation

$$\sum_{n=r}^\infty \left((n+1)\binom{n}{r}\right)^{-1}\mathbb{E}U_{r,n+1} = \frac{r}{(r+1)^2}$$

which, by observing that $\mathbb{E}U_{r,n+1} = \frac{r}{n+2}$, can be expressed by

$$\sum_{n=r}^\infty \frac{(n-r)!}{(n+2)!} = \frac{1}{r!(r+1)^2}.$$

As in Example 1, the same relation results by choosing F to be the standard uniform distribution function and $\varphi(x) = x, x \geq 0$.

(vi)   By setting $\varphi(x) = \mathbb{1}_{[t,\infty)}(x)$, $x \in \mathbb{R}$, $t > 0$, in (11), we immediately obtain a relation for cumulative distributions functions of order statistics based on F, as follows

$$
\sum_{n=r}^{\infty} \left((n+1) \binom{n}{r} \right)^{-1} P(X_{r,n+1} > t) = \frac{1}{r} P(X_{r,r} > t) + F^r(t) \log F(t)
$$

$$
= \frac{1}{r}(1 - F^r(t)) + F^r(t) \log F(t)
$$

$$
= \frac{1}{r} - F^r(t) \left(\frac{1}{r} - \log F(t) \right), t > 0.
$$

(vii)   By choosing φ as $\varphi(x) = e^{tx} - 1$, for $t \in \mathbb{R}$, $x \geq 0$, moment generating functions of lower order statistics appear in the representation of Theorem 3.1.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors have not used generative AI and AI-assisted technologies and take full responsibility for the content of the publication.

References

Balakrishnan, N., F. Buono, and M. Longobardi. 2022. On cumulative entropies in terms of moments of order statistics. *Methodology and Computing in Applied Probability* 24 (1):345–59. doi:10.1007/s11009-021-09850-0.

Baratpour, S. 2010. Characterizations based on cumulative residual entropy of first-order statistics. *Communications in Statistics - Theory and Methods* 39 (20):3645–51. doi:10.1080/03610920903324841.

Di Crescenzo, A., and M. Longobardi. 2009. On cumulative entropies. *Journal of Statistical Planning and Inference* 139 (12):4072–87. doi:10.1016/j.jspi.2009.05.038.

Gould, H. W. 1972. *Combinatorial identities: A standardized set of tables listing 500 binomial coefficient summations*. Morgantown.

Kattumannil, S. K., E. P. Sreedevi, and N. Balakrishnan. 2022. A generalized measure of cumulative residual entropy. *Entropy* 24 (4):444.

Khorashadizadeh, M., A. H. Rezaei Roknabadi, and G. R. Mohtashami Borzadaran. 2013. Doubly truncated (interval) cumulative residual and past entropy. *Statistics & Probability Letters* 83 (5):1464–71. doi:10.1016/j.spl.2013.01.033.

Mirali, M., and S. Baratpour. 2017. Some results on weighted cumulative entropy. *Journal of the Iranian Statistical Society* 16:21–32.

Misagh, F., Y. Panahi, G. H. Yari, and R. Shahi. 2011. Weighted cumulative entropy and its estimation. *2011 IEEE International Conference on Quality and Reliability*, Bangkok, Thailand, 477–80. doi:10.1109/ICQR.2011.6031765.

Rao, M., Y. Chen, B. C. Vemuri, and F. Wang. 2004. Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory* 50 (6):1220–8. doi:10.1109/TIT.2004.828057.

Sordo, M. A., A. Castaño-Martínez, and G. Pigueiras. 2016. A family of premium principles based on mixtures of TVaRs. *Insurance: Mathematics and Economics* 70:397–405. doi:10.1016/j.insmatheco.2016.07.006.

Tahmasebi, S. 2020. Weighted extensions of generalized cumulative residual entropy and their applications. *Communications in Statistics - Theory and Methods* 49 (21):5196–219. doi:10.1080/03610926.2019.1615094.

Tahmasebi, S., and M. Eskandarzadeh. 2017. Generalized cumulative entropy based on k th lower record values. *Statistics & Probability Letters* 126:164–72. doi:10.1016/j.spl.2017.02.036.

Tahmasebi, S., M. Longobardi, F. Foroghi, and F. Lak. 2020. An extension of weighted generalized cumulative past measure of information. *Ricerche di Matematica* 69 (1):53–81. doi:10.1007/s11587-019-00448-w.