

On Tsallis extropy with an application to pattern recognition

Original

On Tsallis extropy with an application to pattern recognition / Balakrishnan, N., Buono, F., Longobardi, M.. - In: STATISTICS & PROBABILITY LETTERS. - ISSN 0167-7152. - 180:(2022), pp. 1-9. [10.1016/j.spl.2021.109241]

Availability:

This version is available at: 11583/2994636 since: 2024-11-25T11:34:05Z

Publisher:

Elsevier

Published

DOI:10.1016/j.spl.2021.109241

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

Elsevier postprint/Author's Accepted Manuscript

© 2022. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>. The final authenticated version is available online at:
<http://dx.doi.org/10.1016/j.spl.2021.109241>

(Article begins on next page)

On Tsallis extropy with an application to pattern recognition

Narayanaswamy Balakrishnan^a, Francesco Buono^{b,*}, Maria Longobardi^b

^aMcMaster University, Canada

^bUniversità di Napoli Federico II, Italy

Abstract

Recently, a new measure of information called extropy has been introduced by Lad, Sanfilippo and Agrò as the dual version of Shannon entropy. In the literature, Tsallis introduced a measure for a discrete random variable, named Tsallis entropy, as a generalization of Boltzmann-Gibbs statistics. In this work, a new measure of discrimination, called Tsallis extropy, is introduced and some of its properties are then discussed. The relation between Tsallis extropy and entropy is given and some bounds are also presented. Finally, an application of this extropy to pattern recognition is demonstrated.

Keywords: Measures of information, Shannon entropy, Tsallis entropy, Extropy, Pattern recognition
2020 MSC: 62H30, 94A17

1. Introduction

Let X be a discrete random variable with support $S = \{x_1, \dots, x_N\}$ and with corresponding probability vector $\mathbf{p} = (p_1, \dots, p_N)$, i.e., $\mathbb{P}(X = x_i) = p_i$, for $i = 1, \dots, N$. In 1948, Shannon [21] introduced a measure of information related to the information content and the uncertainty about an event associated with a discrete random variable. This measure, come to be known as Shannon entropy, is defined as

$$H(X) = - \sum_{i=1}^N p_i \log p_i, \quad (1)$$

where \log is the natural logarithm. The concept of entropy has since been generalized in different ways. Analogous to the discrete case, the Shannon entropy has been defined in the continuous case as well as

$$H(X) = - \int_0^{+\infty} f(x) \log f(x) dx,$$

where X is a non-negative random variable with probability density function f . Although the definitions are similar, the entropy is always non-negative in the discrete case while it could be negative in the continuous case. In the literature, different versions of entropy have been introduced, including residual and past entropy [7, 11], cumulative residual and past entropy [9, 10, 19, 23], weighted entropies [8], dynamic entropies [1], weighted cumulative residual entropy [17], generalized dynamic entropies [20], and also some relationships with other concepts in reliability theory [3, 16, 22].

Lad *et al.* [15] introduced the extropy, a measure of uncertainty, as a dual version of the entropy. It is useful for comparing the uncertainty between two random variables X and Y , i.e., if the extropy of X is less than that of Y , then X has less uncertainty than Y . For a discrete random variable X , the extropy $J(X)$ is defined as

$$J(X) = - \sum_{i=1}^N (1 - p_i) \log(1 - p_i), \quad (2)$$

*Corresponding author

Email address: francesco.buono3@unina.it (Francesco Buono)

and it is always non-negative. The entropy has subsequently been widely studied and several different versions have also been proposed in the literature; see [2, 12, 13, 18].

Among the different generalizations of Shannon entropy, the Tsallis entropy [24] has attracted considerable attention. For a discrete random variable, the Tsallis entropy $S_\alpha(X)$ is defined as

$$S_\alpha(X) = \frac{1}{\alpha - 1} \left(1 - \sum_{i=1}^N p_i^\alpha \right), \quad (3)$$

where $\alpha > 0$ and $\alpha \neq 1$. It is a generalization of Shannon entropy since it is evident that

$$\lim_{\alpha \rightarrow 1} S_\alpha(X) = H(X). \quad (4)$$

10 In the present work, we introduce a measure of uncertainty dual to the Tsallis entropy in (3), and it may be referred to as the Tsallis extropy. The rest of the paper proceeds as follows. In Section 2, the Tsallis extropy is defined, and some of its properties are given, and its relationships to other known measures are described. In Section 3, we study the maximum Tsallis extropy and establish an upper bound for it. In Section 4, we apply this new measure to a problem in pattern recognition, and compare some other known methods with the method based on the extropy. Finally, in
15 Section 5, we provide some concluding remarks and summarize the results of this work.

2. Tsallis extropy

In this section, we introduce the Tsallis extropy, dual to the Tsallis entropy, as a new measure of uncertainty. It is defined to preserve a relationship similar to the one between Shannon entropy and extropy. It is important to mention at this point that Lad *et al.* [15] proved the following property with regard to the sum of entropy and extropy:

$$H(\mathbf{p}) + J(\mathbf{p}) = \sum_{i=1}^N H(p_i, 1 - p_i) = \sum_{i=1}^N J(p_i, 1 - p_i), \quad (5)$$

where $H(p_i, 1 - p_i) = J(p_i, 1 - p_i) = -p_i \log p_i - (1 - p_i) \log(1 - p_i)$. Observe that the two terms on the RHS of this expression are the entropy and extropy of a discrete random variable taking on two values with the corresponding probabilities as $(p_i, 1 - p_i)$.

Definition 1. Let X be a discrete random variable with support $S = \{x_1, \dots, x_N\}$ and with corresponding probability vector $\mathbf{p} = (p_1, \dots, p_N)$, and let $\alpha > 0$, $\alpha \neq 1$. Then, the Tsallis extropy of X , $JS_\alpha(X)$, is defined as

$$JS_\alpha(X) = \frac{1}{\alpha - 1} \left(N - 1 - \sum_{i=1}^N (1 - p_i)^\alpha \right). \quad (6)$$

Remark 1. The definition in (6) is obtained in a different way from the Tsallis entropy in (3). In fact, using the normalization condition, we can rewrite the Tsallis entropy as

$$S_\alpha(X) = \frac{1}{\alpha - 1} \sum_{i=1}^N p_i (1 - p_i^{\alpha-1}).$$

We can then introduce the Tsallis extropy as

$$\frac{1}{\alpha - 1} \sum_{i=1}^N (1 - p_i) \left(1 - (1 - p_i)^{\alpha-1} \right) = \frac{1}{\alpha - 1} \left(\sum_{i=1}^N (1 - p_i) - \sum_{i=1}^N (1 - p_i)^\alpha \right), \quad (7)$$

20 from which the definition in (6) follows readily. Intuitively, the extropy corresponding to a fixed entropy could be simply introduced by replacing all the p_i by $(1 - p_i)$ as seen in (7), but we will see in Proposition 2.3 that our definition has a deeper meaning as it preserves the invariance property about the sum of entropy and extropy (5) studied by Lad *et al.* [15].

Proposition 2.1. *The Tsallis extropy in non-negative.*

Proof. Let us consider the expression of Tsallis extropy in the LHS of (7). For $\alpha > 1$, the function $h(x) = x^{\alpha-1}$ is increasing in $x > 0$, and so

$$1 - (1 - p_i)^{\alpha-1} \geq 0,$$

and hence the Tsallis extropy is non-negative. For $0 < \alpha < 1$, the function $h(x) = x^{\alpha-1}$ is decreasing in $x > 0$, i.e.,

$$1 - (1 - p_i)^{\alpha-1} \leq 0,$$

25 and hence the Tsallis extropy is non-negative due to the multiplicative factor $\frac{1}{\alpha-1}$ being negative. \square

Remark 2. Proposition 2.1 is logical since the situation characterized by the lowest uncertainty is the one in which we have $\mathbf{p} = (1, 0, \dots, 0)$, which corresponds to $JS_\alpha(X) = 0$.

We now present some examples to demonstrate the evaluation of Tsallis extropy.

30 **Example 1.** Let X be a discrete random variable uniformly distributed over $\{1, \dots, N\}$. Then, the Tsallis extropy is given by

$$\begin{aligned} JS_\alpha(X) &= \frac{1}{\alpha-1} \left\{ N - 1 - \sum_{i=1}^N \left(1 - \frac{1}{N}\right)^\alpha \right\} \\ &= \frac{1}{\alpha-1} \left\{ N - 1 - \frac{(N-1)^\alpha}{N^{\alpha-1}} \right\} \\ &= \frac{N-1}{\alpha-1} \frac{\{N^{\alpha-1} - (N-1)^{\alpha-1}\}}{N^{\alpha-1}}. \end{aligned} \quad (8)$$

In the following proposition, we show that the Tsallis extropy reduces to the extropy in (2) when α tends to 1. Bear in mind that this is a classical property of Tsallis and Shannon entropies.

Proposition 2.2. *Let X be a discrete random variable with finite support S and with corresponding probability vector \mathbf{p} . Then,*

$$\lim_{\alpha \rightarrow 1} JS_\alpha(X) = J(X). \quad (9)$$

Proof. From (6) and by using L'Hôpital's rule, we find

$$\begin{aligned} \lim_{\alpha \rightarrow 1} JS_\alpha(X) &= \lim_{\alpha \rightarrow 1} \frac{1}{\alpha-1} \left(N - 1 - \sum_{i=1}^N (1 - p_i)^\alpha \right) \\ &= - \lim_{\alpha \rightarrow 1} \sum_{i=1}^N (1 - p_i)^\alpha \log(1 - p_i) \\ &= - \sum_{i=1}^N (1 - p_i) \log(1 - p_i) = J(X). \end{aligned}$$

\square

35 Next, for discussing the sum of Tsallis entropy and extropy, similar to the one presented in (5), we need the following lemma about random variables taking on two values.

Lemma 2.1. *Let X be a discrete random variable taking on two values with corresponding probabilities $(p, 1 - p)$. Then,*

$$JS_\alpha(X) = S_\alpha(X). \quad (10)$$

Proof. From (6), with $N = 2$, we find

$$\begin{aligned} JS_\alpha(X) &= \frac{1}{\alpha-1} [1 - (1-p)^\alpha - \{1 - (1-p)\}^\alpha] \\ &= \frac{1}{\alpha-1} \{1 - p^\alpha - (1-p)^\alpha\} = S_\alpha(X). \end{aligned}$$

□

Proposition 2.3. *Let X be a discrete random variable with finite support S and with corresponding probability vector \mathbf{p} . Then,*

$$S_\alpha(X) + JS_\alpha(X) = \sum_{i=1}^N S_\alpha(p_i, 1-p_i) = \sum_{i=1}^N JS_\alpha(p_i, 1-p_i), \quad (11)$$

where $S_\alpha(p_i, 1-p_i)$ and $JS_\alpha(p_i, 1-p_i)$ are the Tsallis entropy and extropy of a discrete random variable taking on two values with corresponding probabilities $(p_i, 1-p_i)$.

Proof. We have to prove only the first equality, since the second one is given in Lemma 2.1. From Equations (3)–(6), we have

$$\begin{aligned} S_\alpha(X) + JS_\alpha(X) &= \frac{1}{\alpha-1} \left\{ N - \sum_{i=1}^N p_i^\alpha - \sum_{i=1}^N (1-p_i)^\alpha \right\} \\ &= \frac{1}{\alpha-1} \sum_{i=1}^N \{1 - p_i^\alpha - (1-p_i)^\alpha\} \\ &= \sum_{i=1}^N S_\alpha(p_i, 1-p_i), \end{aligned}$$

as required. □

In the following proposition, we will show that for the choice of the parameter $\alpha = 2$, the Tsallis entropy and extropy coincide.

Proposition 2.4. *Let X be a discrete random variable with finite support S of cardinality N . Then, $S_2(X) = JS_2(X)$.*

Proof. From (6), upon choosing $\alpha = 2$, we obtain

$$\begin{aligned} JS_2(X) &= \frac{1}{2-1} \left(N - 1 - \sum_{i=1}^N (1-p_i)^2 \right) \\ &= N - 1 - \sum_{i=1}^N (1 + p_i^2 - 2p_i) \\ &= N - 1 - N - \sum_{i=1}^N p_i^2 + 2 \sum_{i=1}^N p_i \\ &= 1 - \sum_{i=1}^N p_i^2 = S_2(X), \end{aligned}$$

as required. □

In the following theorem, we will prove that the Tsallis entropy is always greater than the Tsallis extropy for $\alpha < 2$ and that the reverse inequality holds for $\alpha > 2$.

Theorem 2.1. For any discrete random variable X with support of cardinality $N \geq 3$, we have

$$\begin{aligned} S_\alpha(X) &\geq JS_\alpha(X) \text{ if } 0 < \alpha < 2, \\ S_\alpha(X) &\leq JS_\alpha(X) \text{ if } \alpha > 2. \end{aligned}$$

Proof. First of all, we remark that for $\alpha = 1$ we mean the limit case in which we obtain the well known result about entropy and extropy. Let us consider the difference between Tsallis entropy and extropy given by

$$S_\alpha(X) - JS_\alpha(X) = \frac{1}{\alpha - 1} \left[2 - N - \sum_{i=1}^N p_i^\alpha + \sum_{i=1}^N (1 - p_i)^\alpha \right].$$

Then, we consider the Lagrange function L defined as

$$L = S_\alpha(X) - JS_\alpha(X) + \lambda \left(\sum_{i=1}^N p_i - 1 \right),$$

for which the partial derivatives with respect to p_i are

$$\frac{\partial L}{\partial p_i} = \frac{-\alpha}{\alpha - 1} \left(p_i^{\alpha-1} + (1 - p_i)^{\alpha-1} \right) + \lambda,$$

which vanish if and only if

$$p_i^{\alpha-1} + (1 - p_i)^{\alpha-1} = C, \quad (12)$$

where C is a constant. Then, we consider, for $0 \leq x \leq 1$, the function $h(x) = x^{\alpha-1} + (1-x)^{\alpha-1}$, such that $h(x) = h(1-x)$ and $h(0) = h(1) = 1$. The function h has a minimum at $x = \frac{1}{2}$, if $\alpha > 2$ or $0 < \alpha < 1$, and a maximum at the same point for $1 < \alpha < 2$. Then, in order to satisfy both (12) and the normalization condition, we have only two possibilities.

55 The first one is given by choosing one p_i equal to 1 and all the others equal to 0, whereas the second one is given by $p_i = \frac{1}{N}$, $i = 1, \dots, N$. These are the cases in which the difference between Tsallis entropy and extropy takes the maximum and the minimum values. In the first case, we have $S_\alpha(X) - JS_\alpha(X) = 0$. In the second case, we obtain

$$\begin{aligned} S_\alpha(X) - JS_\alpha(X) &= \frac{1}{\alpha - 1} \left[2 - N - \sum_{i=1}^N \frac{1}{N^\alpha} + \sum_{i=1}^N \left(1 - \frac{1}{N} \right)^\alpha \right] \\ &= \frac{2N^{\alpha-1} - N^\alpha - 1 + (N-1)^\alpha}{(\alpha - 1)N^{\alpha-1}}. \end{aligned} \quad (13)$$

Let us consider the numerator of (13) as a function of α , $g(\alpha) = 2N^{\alpha-1} - N^\alpha - 1 + (N-1)^\alpha$. We have

$$g'(\alpha) = N^{\alpha-1}(\log N)(2 - N) + (N-1)^\alpha \log(N-1),$$

which is non negative if, and only if,

$$\alpha \leq \frac{\log \left[\frac{N}{N-2} \frac{\log(N-1)}{\log N} \right]}{\log \left(\frac{N}{N-1} \right)} = G(N).$$

We have that, for $N \geq 3$,

$$1 < G(N) < 2 \iff \frac{N-2}{N-1} < \frac{\log(N-1)}{\log N} < \frac{N(N-2)}{(N-1)^2}, \quad (14)$$

which holds as one can see in Figure 1.

60 Then, the function g has a maximum between 1 and 2 and $g(1) = g(2) = 0$. Hence, we have $g(\alpha) > 0$ if $1 < \alpha < 2$ and $g(\alpha) < 0$ if $0 < \alpha < 1$ or $\alpha > 2$. By recalling the definition of g and (13), we obtain that the difference between Tsallis entropy and extropy for uniform distribution is greater than 0 if $0 < \alpha < 1$ or $1 < \alpha < 2$ and less than 0 if $\alpha > 2$. Hence, $S_\alpha(X) - JS_\alpha(X)$ has minimum of 0 and maximum for the uniform distribution if $0 < \alpha < 2$ and viceversa if $\alpha > 2$. \square

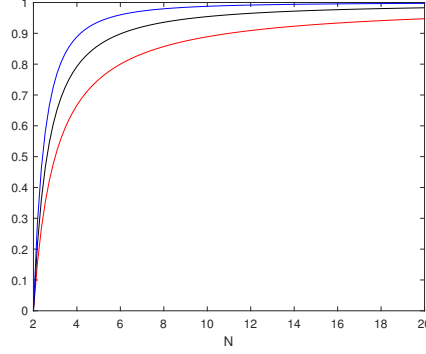


Figure 1: Plot of the functions on the RHS of (14) in red, black and blue, respectively.

3. Maximum Tsallis extropy

65 When we deal with a measure of information, it will be useful to know what its maximum value is. The Tsallis extropy reaches its maximum value when the random variable X is uniformly distributed, as established in the following theorem.

Theorem 3.1. *Let X be a discrete random variable with finite support S of cardinality N , and let $\alpha > 0$, $\alpha \neq 1$. Then, X has maximum Tsallis extropy for fixed N and α if, and only if, it is uniformly distributed.*

Proof. Let N and α be fixed. Then, we need to maximize the function of N variables given by

$$JS_{\alpha}(\mathbf{p}) = \frac{1}{\alpha - 1} \left(N - 1 - \sum_{i=1}^N (1 - p_i)^{\alpha} \right),$$

subject to the condition

$$\sum_{i=1}^N p_i = 1. \quad (15)$$

For this purpose, let us consider the Lagrange function defined by

$$JS_{\alpha}^*(\mathbf{p}) = \frac{1}{\alpha - 1} \left(N - 1 - \sum_{i=1}^N (1 - p_i)^{\alpha} \right) + \lambda \left(\sum_{i=1}^N p_i - 1 \right),$$

from which we readily find the partial derivatives respect to p_i , $i = 1, \dots, N$, as

$$\frac{\partial JS_{\alpha}^*(\mathbf{p})}{\partial p_i} = \frac{\alpha}{\alpha - 1} (1 - p_i)^{\alpha-1} + \lambda.$$

Then, we can determine the stationary points as

$$\frac{\alpha}{\alpha - 1} (1 - p_i)^{\alpha-1} + \lambda = 0 \iff p_i = 1 - \left(\frac{(1 - \alpha)\lambda}{\alpha} \right)^{\frac{1}{\alpha-1}} = K,$$

70 where K is a constant. For satisfying the condition in (15), we need to have $K = \frac{1}{N}$, in which case \mathbf{p} becomes the probability mass function vector of a discrete uniform distribution. \square

Remark 3. In Example 1, the Tsallis extropy has been evaluated for the uniform distribution over N elements, and so the maximum Tsallis extropy is as presented in (8).

Theorem 3.2. *The Tsallis entropy is less than 1.*

Proof. To establish this result, we show that the Tsallis entropy of a discrete uniform distribution increases to 1 as the size of the support N increases. Let X_N be a discrete random variable uniformly distributed over a finite support of size N . From (8), we know the corresponding Tsallis entropy is

$$JS_\alpha(X_N) = \frac{1}{\alpha - 1} \left(N - 1 - \frac{(N - 1)^\alpha}{N^{\alpha-1}} \right).$$

Let us now consider the function

$$g(N) = N - 1 - \frac{(N - 1)^\alpha}{N^{\alpha-1}}, \quad (16)$$

and show that it increases for $\alpha > 1$ and decreases for $0 < \alpha < 1$. This way, we will prove that $JS_\alpha(X_N)$ is increasing in N . Let us consider the derivative of g , by treating it as a function of a continuous variable N , given by

$$g'(N) = 1 - \frac{(N - 1)^{\alpha-1}}{N^\alpha} (\alpha + N - 1) = \frac{N^\alpha - (N - 1)^{\alpha-1} (\alpha + N - 1)}{N^\alpha},$$

whose sign is determined by

$$N^\alpha - (N - 1)^{\alpha-1} (\alpha + N - 1) = N^\alpha - (N - 1)^\alpha - \alpha(N - 1)^{\alpha-1},$$

which, by mean value theorem, is equal to

$$\alpha(N - 1 + \varepsilon)^{\alpha-1} - \alpha(N - 1)^{\alpha-1},$$

75 for some $\varepsilon \in (0, 1)$. Thence, by using the fact that the function $h(x) = x^{\alpha-1}$ is increasing in $x > 0$ for $\alpha > 1$ and decreasing for $0 < \alpha < 1$, we get the monotonicity of $g(N)$ in (16).

Now, we evaluate the limit of $JS_\alpha(X_N)$ as N tends to infinity. We have

$$\begin{aligned} \lim_{N \rightarrow +\infty} \frac{N - 1}{\alpha - 1} \frac{(N^{\alpha-1} - (N - 1)^{\alpha-1})}{N^{\alpha-1}} &= \lim_{N \rightarrow +\infty} \frac{N - 1}{\alpha - 1} \left[1 - \left(1 - \frac{1}{N} \right)^{\alpha-1} \right] \\ &= \lim_{N \rightarrow +\infty} \frac{N - 1}{N} = 1. \end{aligned}$$

Finally, upon using the result in Theorem 3.1 about the maximum Tsallis entropy, we conclude that the Tsallis entropy is less than 1 for any discrete random variable. \square

Corollary 3.1. *For any discrete random variable X , we have*

$$0 \leq JS_\alpha(X) < 1.$$

80 *Proof.* The result follows readily from Proposition 2.1 and Theorem 3.2. \square

4. Application to pattern recognition

In this section, we give an application of the Tsallis entropy in pattern recognition by using the well-known Iris dataset in [6]. We then compare the results obtained with those in the Dempster-Shafer theory of evidence [5] due to Kang et al. [14] and Bueno and Longobardi [4]. The objective is to classify among three classes of flowers: Iris Setosa (Se), Iris Versicolour (Ve) and Iris Virginica (Vi). The dataset consists of 150 samples, with 50 in each class. 85 The characteristics measured for each flower are: the sepal length in cm (SL), the sepal width in cm (SW), the petal length in cm (PL), the petal width in cm (PW) and the class (one of Se, Ve and Vi). We select 40 samples for each kind of Iris and then we find a sample of max-min value to generate a model of interval numbers, as shown in Table 1(a). Each element of the dataset can be regarded as an unknown test sample. Suppose the selected sample data is 90 (6.1, 3.0, 4.9, 1.8, Vi).

Table 1: (a) The interval numbers of the statistical model. (b) Probability distributions based on Kang's method.

(a)					(b)				
Item	SL	SW	PL	PW	Item	SL	SW	PL	PW
<i>Se</i>	[4.4,5.8]	[2.3,4.4]	[1.0,1.9]	[0.1,0.6]	$\mathbb{P}(Se)$	0.3058	0.2748	0.1391	0.1563
<i>Ve</i>	[4.9,7.0]	[2.0,3.4]	[3.0,5.1]	[1.0,1.7]	$\mathbb{P}(Ve)$	0.4148	0.3516	0.3801	0.3737
<i>Vi</i>	[4.9,7.9]	[2.2,3.8]	[4.5,6.9]	[1.4,2.5]	$\mathbb{P}(Vi)$	0.2794	0.3736	0.4808	0.4700

We then generate four discrete probability distributions using the method of Kang et al. [14] based on the similarity of interval numbers. Given two intervals $A = [a_1, a_2]$ and $B = [b_1, b_2]$, their similarity $S(A, B)$ is defined as

$$S(A, B) = \frac{1}{1 + \gamma D(A, B)},$$

where $\gamma > 0$ is the coefficient of support, and we used $\gamma = 5$, for example. Then, $D(A, B)$, the distance between intervals A and B , is defined to be

$$D^2(A, B) = \left[\left(\frac{a_1 + a_2}{2} \right) - \left(\frac{b_1 + b_2}{2} \right) \right]^2 + \frac{1}{3} \left[\left(\frac{a_2 - a_1}{2} \right)^2 + \left(\frac{b_2 - b_1}{2} \right)^2 \right].$$

To generate probability distributions, the intervals given in Table 1(a) are used for interval A and for interval B we use singletons given by the selected sample. For each one of the four characteristics measured, we get three values of similarity and then we obtain a probability distribution by normalizing them (see Table 1(b)). We then evaluate the Tsallis entropy of these probability distributions, as presented in Table 2(a), wherein we have used $\alpha = 0.5, 0.7, 1.5, 2$.

Table 2: Tsallis entropy (a) and the weights (b) for different choices of α .

(a)					(b)				
Item	SL	SW	PL	PW		$\omega(SL)$	$\omega(SW)$	$\omega(PL)$	$\omega(PW)$
$\alpha = 0.5$	0.8941	0.8965	0.8715	0.8759	$\alpha = 0.5$	0.2476	0.2470	0.2533	0.2522
$\alpha = 0.7$	0.8560	0.8592	0.8267	0.8324	$\alpha = 0.7$	0.2469	0.2461	0.2542	0.2528
$\alpha = 1.5$	0.7245	0.7291	0.6781	0.6871	$\alpha = 1.5$	0.2450	0.2439	0.2567	0.2544
$\alpha = 2$	0.6564	0.6613	0.6050	0.6150	$\alpha = 2$	0.2445	0.2433	0.2574	0.2548

We use the Tsallis entropies in Table 2(a) to generate other probability distributions. Observe that the higher the entropy, the higher the uncertainty, and so it would be reasonable to give more weight to observations related to characteristics with lower Tsallis entropy. We refer to the obtained Tsallis entropies as $JS_\alpha(SL)$, $JS_\alpha(SW)$, $JS_\alpha(PL)$, $JS_\alpha(PW)$. Due to the monotonicity of the exponential function, we choose as baseline weight the function $w(x) = e^{-x}$, and we can then obtain the weights ω by normalization. For example, for the sepal length, we have

$$\omega(SL) = \frac{e^{-JS_\alpha(SL)}}{e^{-JS_\alpha(SL)} + e^{-JS_\alpha(SW)} + e^{-JS_\alpha(PL)} + e^{-JS_\alpha(PW)}}.$$

The values of the weights are listed in Table 2(b) for different choices of the parameter α . We determine a final probability distribution in the following way: for each kind of flower, we have four probabilities, one for a specific characteristic; we multiply the probabilities given in Table 1(b) by the corresponding weights and then sum the values relating to the same class. For example, the probability of the class Iris Setosa is obtained as follows:

$$\mathbb{P}(Se) = 0.3058 \cdot \omega(SL) + 0.2748 \cdot \omega(SW) + 0.1391 \cdot \omega(PL) + 0.1563 \cdot \omega(PW).$$

Thus, by choosing $\alpha = 0.5$, we obtain the final probability distribution to be

$$\mathbb{P}(Se) = 0.2182, \quad \mathbb{P}(Ve) = 0.3800, \quad \mathbb{P}(Vi) = 0.4018,$$

95 and then the decision is that the selected flower belongs to the class with the higher probability, Iris Virginica, i.e., we thus made the correct decision, in this case.

In this manner, we tested all 150 samples for different values of α , and observed that the overall recognition rate of this method based on the Tsallis extropy to be 94.66%. The results obtained this way are compared with the recognition rates of the methods of Kang et al. [14] and Buono and Longobardi [4], and these are presented in Table 3. The proposed method is seen to present a slightly better performance in comparison to the other two methods.

Table 3: The recognition rates of different methods.

Item	Se	Ve	Vi	Overall
Kang's method	100%	96%	84%	93.33%
Buono and Longobardi's method	100%	96%	86%	94%
Method based on Tsallis extropy	100%	98%	86%	94.66%

5. Conclusions

In this paper we have studied the Tsallis extropy. It is a measure of uncertainty dual to the Tsallis entropy and it is introduced by preserving a classical invariance property about entropy and extropy. We have discussed some properties of the proposed measure and have given several examples. In particular, we have examined the problem of the maximum Tsallis extropy which is always of great interest in information theory. Finally, we have illustrated an application in the context of pattern recognition by using the proposed measure, and we have then compared the recognition rates with two other known methods. It will be of interest to introduce some other new measures with interesting properties and also discuss the optimal choice of parameter α in the proposed measure.

Acknowledgements

110 Narayanaswamy Balakrishnan thanks the Natural Sciences and Engineering Research Council of Canada for funding this research through an Individual Discovery Grant. Francesco Buono and Maria Longobardi are members of the research group GNAMPA of INdAM (Istituto Nazionale di Alta Matematica) and are partially supported by MIUR-PRIN 2017, project "Stochastic Models for Complex Systems", no. 2017 JFFHSH. Our sincere thanks also go to the anonymous reviewers and the Editor for their useful comments on an earlier version of the manuscript which led to this improved version.

References

- [1] Asadi, M., Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *Journal of Statistical Planning and Inference*, **137**, 1931–1941.
- [2] Balakrishnan, N., Buono, F., Longobardi, M. (2020). On weighted extropies. *Communications in Statistics – Theory and Methods*, DOI: 10.1080/03610926.2020.1860222.
- 120 [3] Balakrishnan, N., Buono, F., Longobardi, M. (2021). On cumulative entropies in terms of moments of order statistics. *Methodology and Computing in Applied Probability*, <https://doi.org/10.1007/s11009-021-09850-0>.
- [4] Buono, F., Longobardi, M. (2020). A dual measure of uncertainty: The Deng extropy. *Entropy*, **22**, 582.
- [5] Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, **38**, 325–339.
- [6] Dheeru, D., Karra Taniskidou, E. (2017). UCI Machine Learning Repository. Available online: <http://archive.ics.uci.edu/ml>.
- 125 [7] Di Crescenzo, A., Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied Probability*, **39**, 434–440.
- [8] Di Crescenzo, A., Longobardi, M. (2006). On weighted residual and past entropies. *Scientiae Mathematicae Japonicae*, **64**, 255–266.
- [9] Di Crescenzo, A., Longobardi, M. (2009a). On cumulative entropies. *Journal of Statistical Planning and Inference*, **139**, 4072–4087.

- 130 [10] Di Crescenzo, A., Longobardi, M. (2009b). On cumulative entropies and lifetime estimations, In: J. Mira, J.M. Ferrandez, J.R. Alvarez Sanchez, F. Paz, J. Toledo (Eds.), *Methods and Models in Artificial and Natural Computation*, IWINAC 2009, Part I, in: LNCS, vol. 5601, Springer-Verlag, Berlin, Heidelberg, pp. 132–141.
- [11] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhyā, Series A*, **58**, 48–56.
- [12] Jahanshani, S.M.A., Zarei, H., Khammar, A.H. (2020). On cumulative residual extropy. *Probability in the Engineering and Informational Sciences*, **34**, 605–625.
- 135 [13] Kamari, O., Buono, F. (2020). On extropy of past lifetime distribution. *Ricerche di Matematica*, DOI: 10.1007/s11587-020-00488-7.
- [14] Kang, B.Y., Li, Y., Deng, Y., Zhang, Y.J., Deng, X.Y. (2012) Determination of basic probability assignment based on interval numbers and its application. *Dianzi Xuebao (Acta Electronica Sinica)*, **40**, 1092–1096.
- [15] Lad, F., Sanfilippo, G., Agrò, G. (2015). Extropy: complementary dual of entropy. *Statistical Science*, **30**, 40–58.
- [16] Longobardi, M. (2014). Cumulative measures of information and stochastic orders. *Ricerche di Matematica*, **63**, 209–223.
- 140 [17] Mirali, M., Baratpour, S., Fakoor, V. (2017). On weighted cumulative residual entropy. *Communications in Statistics – Theory and Methods*, **46**, 2857–2869.
- [18] Qiu, G. (2017). The extropy of order statistics and record values. *Statistics & Probability Letters*, **120**, 52–60.
- [19] Rao, M., Chen, Y., Vemuri, B.C., Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory*, **50**, 1220–1228.
- 145 [20] Sekeh, S.Y., Mohtashami Borzadaran, G.R., Rezaei Roknabadi, A.H. (2014). Some results based on a version of the generalized dynamic entropies. *Communications in Statistics – Theory and Methods*, **43**, 2989–3006.
- [21] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.
- [22] Tahmasebi, S., Eskandarzadeh, M. (2017). Generalized cumulative entropy based on k th lower record values. *Statistics & Probability Letters*, **126**, 164–172.
- 150 [23] Tahmasebi, S., Jafari, A. K., Eskandarzadeh, M. (2016). Some results on residual entropy of ranked set samples. *Statistics & Probability Letters*, **112**, 137–145.
- [24] Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistic. *Journal of Statistical Physics*, **52**, 479–487.