A unified formulation of entropy and its application

A unified formulation of entropy and its application

Narayanaswamy Balakrishnan[1], Francesco Buono[1,], Maria Longobardi[1]

[a]McMaster University, Canada
[b]Università di Napoli Federico II, Italy

Abstract

In this paper, a general formulation of entropy is proposed. It depends on two parameters and includes Shannon, Tsallis and fractional entropy, all as special cases. This measure of information is referred to as fractional Tsallis entropy and some of its properties are then studied. Furthermore, the corresponding entropy in the context of Dempster-Shafer theory of evidence is proposed and referred to as fractional version of Tsallis-Deng entropy. Finally, an application to two classification problems is presented.

Keywords: Measures of information, Shannon entropy, Tsallis entropy, Fractional entropy, Deng entropy, Dempster-Shafer theory of evidence
2010 MSC: 62H30, 94A17

1. Introduction

Let X be a discrete random variable, whose support S has cardinality N, with probability vector $(p_1, \ldots, p_N)$. In the context of information theory, Shannon [**?**] introduced a measure of uncertainty or discrimination about X, known as Shannon entropy, of the form

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

where log denotes the natural logarithm. Since then, several generalizations of Shannon entropy have been proposed. Among them, one of the most important one is the Tsallis entropy introduced in [**?**]. It is defined as

$$S_\alpha(X) = \frac{1}{\alpha - 1}\left(1 - \sum_{i=1}^{N} p_i^\alpha\right), \tag{2}$$

where $\alpha > 0$ and $\alpha \neq 1$, and it is evident that

$$\lim_{\alpha \to 1} S_\alpha(X) = H(X).$$

Several applications of Tsallis entropy have been discussed in the literature. For example, some relations to stochastic orders and order statistics have been studied in [**?**] while some connections to k-record statistics have been discussed in [**?**]. Moreover, several measures based on (**??**) have been introduced including the cumulative residual Tsallis entropy [**?**] and the cumulative Tsallis entropy [**?**].

Another important generalization of Shannon entropy has been introduced in [**?**] in the context of fractional calculus. It is known as fractional entropy and is defined as

$$H_q(X) = \sum_{i=1}^{N} p_i[-\log p_i]^q, \quad 0 < q \leq 1. \tag{3}$$

*Corresponding author

Evidently, it reduces to Shannon entropy, when $q = 1$, i.e., $H_1(X) = H(X)$.

Several measures of uncertainty have also been introduced and studied in the context of Dempster-Shafer theory (DST) of evidence [? ?]. This theory is a generalization of the classical probability theory and enables describing situations characterized by a major uncertainty. Here, a finite set X of mutually exclusive and collectively exhaustive events is considered. It is named frame of discernment and in DST, it is possible to give a positive mass to all the non-empty subsets of X. In particular, a mass function or a basic probability assignment m is a function defined on the power set of X, 2^X, with values in $[0, 1]$ such that

$$m(\emptyset) = 0, \quad \sum_{A \in 2^X} m(A) = 1. \tag{4}$$

The subsets of X with positive mass are called focal elements. Observe that if all the focal elements are singletons, then a BPA is a discrete probability distribution on the elements of X. In analogy with the definitions given for a discrete random variable, it is possible to define measures of uncertainty based on BPA. Nevertheless, while dealing with BPAs, a new type of uncertainty emerges relating to the cardinality of the focal elements. In this viewpoint, Deng [?] introduced a new entropy, known as Deng entropy, of the form

$$ED(m) = -\sum_{A \subseteq X : m(A) > 0} m(A) \log \left(\frac{m(A)}{2^{|A|} - 1} \right). \tag{5}$$

If the BPA reduces to a discrete probability distribution, i.e., $m(A) > 0$ implies $|A| = 1$, then Deng entropy coincides with Shannon entropy. Several generalizations of Deng entropy have been discussed in the literature. In specific situations, the prediction of the information volume contained in the future is still an open issue. Recently, several new measures have been introduced for evaluating the information volume (see [? ? ? ?]). Moreover, Liu et al. [?] proposed an entropy, called Tsallis-Deng entropy, as

$$SD_\alpha(m) = \frac{1}{\alpha - 1} \sum_{A \subseteq X : m(A) > 0} m(A_i) \left[1 - \left(\frac{m(A_i)}{2^{|A_i|} - 1} \right)^{\alpha - 1} \right], \tag{6}$$

where $\alpha > 0$ and $\alpha \neq 1$. Evidently, Tsallis-Deng entropy reduces to Deng entropy when α goes to 1. Further, Kazemi et al. [?] proposed the fractional version of Deng entropy as

$$ED_q(m) = \sum_{A \subseteq X : m(A) > 0} m(A_i) \left(-\log \frac{m(A_i)}{2^{|A_i|} - 1} \right)^q, \tag{7}$$

where $0 < q \leq 1$. This is indeed a generalization of Deng entropy as $ED_1(m) = ED(m)$. For a detailed overview of measures of uncertainty in the context of DST, one may refer to [?].

In the last few decades, several new measures of information have been studied and applied in many problems including language model, medical diagnosis, multi-sensor information fusion, fault diagnosis, pattern recognition, failure mode and effects analysis, risk assessment, decision-making, emergency management, quantum decision (see [?] and the references therein). For this reason, we thought about introducing a unified formulation of entropy.

We would like to emphasize that this paper is not a survey. The proposed results are new as the definitions; only in particular cases, the reader can find well-known entropies. The rest of this paper is organized as follows. In Section 2, the definition of fractional Tsallis entropy is given and some properties and examples are studied. In Section 3, the fractional version of Tsallis-Deng entropy is introduced and then it is shown that it includes several other entropies as limiting or particular cases. In Section 4, an application of the fractional version of Tsallis-Deng entropy to two classification problems is demonstrated. It is used to analyze two datasets, one about qualities of Italian wine and the other about types of iris flowers, and to classify each instance into one of three possible classes. Finally, in Section 5, some concluding remarks are made including the suggestion of some open problems.

2. Fractional Tsallis entropy

In this section, we introduce fractional Tsallis entropy of X as

$$S_\alpha^q(X) \quad = \quad \frac{1}{\alpha - 1} \sum_{i=1}^{N} p_i (1 - p_i^{\alpha - 1})(-\log p_i)^{q-1}, \tag{8}$$

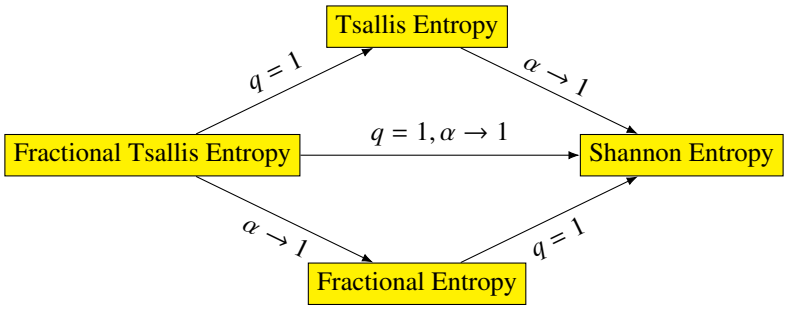

Figure 1: Relationships among different entropies in classical probability theory.

where $\alpha > 0$, $\alpha \neq 1$ and $0 < q \leq 1$. A distinct advantage of this definition is that it includes fractional, Tsallis and Shannon entropies, all as special cases.

Remark 1. The fractional Tsallis entropy in (**??**) is always non-negative. This is due to the fact that $1 - p_i^{\alpha-1}$ is positive for $\alpha > 1$ and negative for $0 < \alpha < 1$, and so the sum in (**??**) has a definite sign and it is the same as that of $\alpha - 1$.

Proposition 2.1. *The fractional Tsallis entropy in* (**??**) *coincides with Tsallis entropy when* $q = 1$.

Proof. When $q = 1$, from (**??**) we get

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

as required. □

Proposition 2.2. *As* α *tends to* 1*, fractional Tsallis entropy converges to fractional entropy.*

Proof. Taking the limit as α tends to 1 in (**??**), and by using L'Hôpital's rule, we get

$$\begin{aligned}
\lim_{\alpha\to 1} S_\alpha^q(X) &= \lim_{\alpha\to 1} \frac{1}{\alpha - 1} \sum_{i=1}^{N} p_i(1 - p_i^{\alpha-1})(-\log p_i)^{q-1} \\
&= \lim_{\alpha\to 1} \sum_{i=1}^{N} p_i\left(-p_i^{\alpha-1}\right) \log p_i(-\log p_i)^{q-1} \\
&= \lim_{\alpha\to 1} \sum_{i=1}^{N} p_i^\alpha(-\log p_i)^q = \sum_{i=1}^{N} p_i(-\log p_i)^q = H_q(X),
\end{aligned}$$

as required. □

Corollary 2.1. *If both parameters of the fractional Tsallis entropy tend to* 1*, fractional Tsallis entropy in* (**??**) *converges to Shannon entropy, i.e.,*

$$\lim_{\alpha,q\to 1} S_\alpha^q(X) = H(X).$$

To summarize the results given in Propositions **??** and **??** and Corollary **??**, the relationships among different kinds of entropies are depicted in the form of a schematic diagram, in Figure **??**.

Example 1. Let X be uniformly distributed over a support S of cardinality N. When N changes, the values of fractional Tsallis entropy were computed for different choices of α and q, and these are presented in Table **??**.

Table 1: Values of fractional Tsallis entropy as N changes, for different choices of q and α.

N	$q = 0.5, \alpha = 0.5$	$q = 0.5, \alpha = 2$	$q = 1, \alpha = 2$	$q = 0.3, \alpha = 0.75$	$q = 0.3, \alpha = 1.5$	$q = 0.5, \alpha = 5$
2	0.9950	0.6006	0.5	0.9782	0.7571	0.2815
3	1.3968	0.6360	0.6667	1.1837	0.7914	0.2356
4	1.6986	0.6370	0.75	1.3182	0.7956	0.2115
5	1.9487	0.6306	0.8	1.4200	0.7923	0.1967
6	2.1657	0.6226	0.8333	1.5207	0.7868	0.1866
7	2.3596	0.6145	0.8571	1.5727	0.7807	0.1791
8	2.5359	0.6068	0.875	1.6336	0.7745	0.1733
9	2.6985	0.5997	0.8889	1.6877	0.7685	0.1686
10	2.8499	0.5931	0.9	1.7364	0.7628	0.1647

From Table **??**, we observe that fractional Tsallis entropy does not always exhibit the same monotonic behavior as a function of N. For the uniform distribution, fractional Tsallis entropy is given by

$$S_\alpha^q\left(\frac{1}{p},\dots,\frac{1}{p}\right) = \frac{1}{\alpha-1}\left(1-\frac{1}{N^{\alpha-1}}\right)(\log N)^{q-1}.$$

Upon differentiating with respect to N (treating N as continuous), we obtain

$$\begin{aligned}\frac{\partial S_\alpha^q}{\partial N} &= \frac{(\log N)^{q-2}}{N^\alpha}\frac{1}{\alpha-1}\left[(\alpha-1)\log N + (q-1)(N^{\alpha-1}-1)\right] \\ &\overset{sgn}{=} \frac{1}{\alpha-1}\left[(\alpha-1)\log N + (q-1)(N^{\alpha-1}-1)\right]. \end{aligned} \tag{9}$$

By observing that

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

we see that we need to consider the function

$$\rho(v) = \log v + (q-1)(v-1), \quad v > 0,$$

which is increasing for $v < \frac{1}{1-q}$. Here, v represents $N^{\alpha-1}$ which is in $(0,1)$ for $\alpha \in (0,1)$ and greater than $2^{\alpha-1}$ for $\alpha > 1$. As $\frac{1}{1-q} > 1$, with $q \in (0,1)$, ρ is increasing in $(0,1)$ and reaches the maximum value for $v = 1$, that is, $\rho(1) = 0$. Hence, for $\alpha \in (0,1)$, the sign in (**??**) is given by the ratio of two negative quantities and so fractional Tsallis entropy is increasing in N. For $q = 1$, it is easy to observe from (**??**) that it is increasing in N regardless of α. Finally, when $\alpha > 1$ and $q \in (0,1)$, there are two possible scenarios. In fact, the fractional Tsallis entropy may be always decreasing in N, or simply definitely decreasing as seen in Table **??**.

Theorem 2.1. *The supremum of fractional Tsallis entropy, as a function of $q \in (0,1]$, is attained at one of the extremes of the interval, and the infimum is attained at one of the extremes of the interval or it is a minimum at a unique $q_0 \in (0,1)$.*

Proof. The fractional Tsallis entropy is a convex function of q. Hence, it can be strictly increasing, strictly decreasing or decreasing up to $q_0 \in (0,1)$, and then increasing. In the first case, the infimum is attained at 0 and the maximum at $q = 1$. In the second case, the minimum is reached at $q = 1$ and the supremum at 0. In the last case, the minimum is attained at q_0 and the supremum is reached at one of the extremes of the interval $(0,1)$. □

3. Fractional version of Tsallis-Deng entropy

In this section, we introduce fractional version of Tsallis-Deng entropy for a BPA m as

$$SD_\alpha^q(m) = \frac{1}{\alpha-1}\sum_{A\subseteq X: m(A)>0} m(A_i)\left[1-\left(\frac{m(A_i)}{2^{|A_i|}-1}\right)^{\alpha-1}\right]\left(-\log\frac{m(A_i)}{2^{|A_i|}-1}\right)^{q-1}, \tag{10}$$

where $\alpha > 0$, $\alpha \neq 1$, $0 < q \leq 1$. In analogy with fractional Tsallis entropy, this is a general expression of entropy as it includes several versions of entropy measure both in the context of DST and in the classical probability theory viewpoint.

Remark 2. In analogy with Remark **??**, fractional version of Tsallis-Deng entropy (**??**) is non-negative too.

Proposition 3.1. *When $q = 1$, fractional version of Tsallis-Deng entropy in* (**??**) *is equal to Tsallis-Deng entropy.*

Proof. Upon taking $q = 1$ in (**??**), we get

$$SD_\alpha^1(m) = \frac{1}{\alpha - 1} \sum_{A \subseteq X : m(A) > 0} m(A_i) \left[1 - \left(\frac{m(A_i)}{2^{|A_i|} - 1} \right)^{\alpha - 1} \right] = SD_\alpha(m),$$

as required. □

Proposition 3.2. *As α tends to 1, fractional version of Tsallis-Deng entropy in* (**??**) *reduces to fractional Deng entropy.*

Proof. Upon letting α tend to 1 in (**??**), and by using L'Hôpital's rule, we get

$$\begin{aligned}
\lim_{\alpha \to 1} SD_\alpha^q(m) &= \lim_{\alpha \to 1} \sum_{A \subseteq X : m(A) > 0} m(A_i) \left[-\left(\frac{m(A_i)}{2^{|A_i|} - 1} \right)^{\alpha - 1} \right] \left(-\log \frac{m(A_i)}{2^{|A_i|} - 1} \right)^{q-1} \log \frac{m(A_i)}{2^{|A_i|} - 1} \\
&= \lim_{\alpha \to 1} \sum_{A \subseteq X : m(A) > 0} m(A_i) \left(\frac{m(A_i)}{2^{|A_i|} - 1} \right)^{\alpha - 1} \left(-\log \frac{m(A_i)}{2^{|A_i|} - 1} \right)^{q} \\
&= \sum_{A \subseteq X : m(A) > 0} m(A_i) \left(-\log \frac{m(A_i)}{2^{|A_i|} - 1} \right)^{q} = ED_q(m),
\end{aligned}$$

as required. □

Corollary 3.1. *When both parameters α and q in* (**??**) *tend to 1, fractional version of Tsallis-Deng entropy in* (**??**) *converges to Deng entropy, i.e.,*

$$\lim_{\alpha, q \to 1} SD_\alpha^q(m) = ED(m).$$

Remark 3. If the BPA m is a discrete probability distribution, then for each focal element $|A| = 1$; in this case, the fractional version of Tsallis-Deng entropy reduces to fractional Tsallis entropy, i.e.,

$$SD_\alpha^q(m) = \frac{1}{\alpha - 1} \sum_{A \subseteq X : m(A) > 0} m(A_i) \left[1 - (m(A_i))^{\alpha - 1} \right] (-\log m(A_i))^{q-1} = S_\alpha^q(X),$$

where X is a discrete random variable with probability vector m.

To summarize the results given in Propositions **??** and **??**, Corollary **??** and Remark **??**, the relationships among different kinds of entropies are depicted in the form of a schematic diagram, in Figure **??**.

Theorem 3.1. *The supremum of the fractional verson of Tsallis-Deng entropy in* (**??**)*, as a function of $q \in (0, 1]$, is attained at one of the extremes of the interval, and the infimum is attained at one of the extremes of the interval or it is a minimum at a unique $q_0 \in (0, 1)$.*

Proof. The proof is similar to that of Theorem **??** and is therefore omitted for brevity. □

Example 2. Consider a frame of discernment $X = \{1, 2, \ldots, 15\}$ and a BPA m such that $m(3, 4, 5) = 0.05$, $m(6) = 0.05$, m(A)=0.8, $m(X) = 0.1$. When A changes, the values of fractional version of Tsallis-Deng entropy in (**??**) have been computed for different choices of α and q, are these are presented in Table **??**.

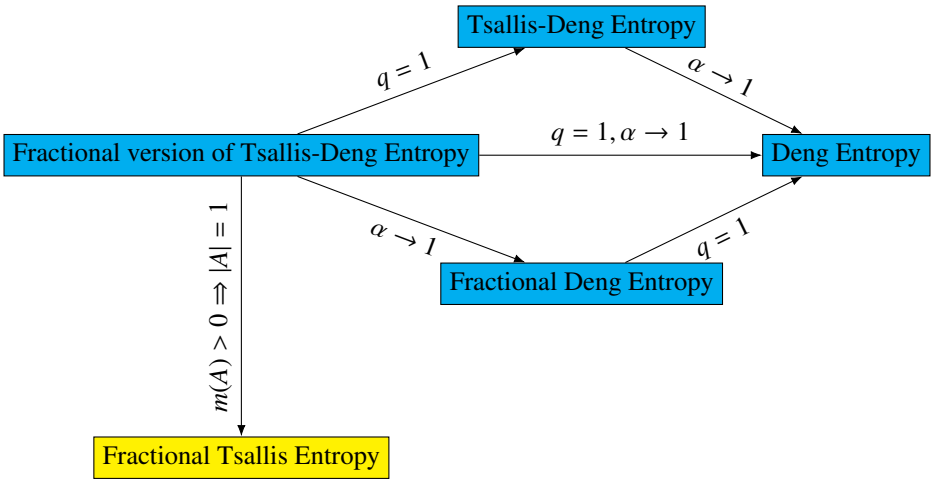

Figure 2: Relationships among different entropies in DST theory (blue) and in classical probability theory (yellow).

Table 2: Values of fractional version of Tsallis-Deng entropy as A changes, for different choices of q and α.

A	$q = 0.5, \alpha = 0.5$	$q = 0.5, \alpha = 2$	$q = 1, \alpha = 2$	$q = 0.3, \alpha = 0.75$	$q = 0.3, \alpha = 1.5$	$q = 0.5, \alpha = 5$
$\{1\}$	35.1571	0.4165	0.3571	2.3353	0.5823	0.2698
$\{1,2\}$	34.0606	0.5881	0.7838	2.8415	0.7361	0.1929
$\{1,2,3\}$	34.8845	0.5590	0.9057	3.1505	0.7156	0.1556
$\{1,\ldots,4\}$	35.8694	0.5202	0.9545	3.4400	0.6793	0.1367
$\{1,\ldots,5\}$	37.1288	0.4854	0.9765	3.7409	0.6415	0.1244
$\{1,\ldots,6\}$	38.7866	0.4558	0.9870	4.0676	0.6056	0.1156
$\{1,\ldots,7\}$	41.0020	0.4310	0.9921	4.4306	0.5726	0.1087
$\{1,\ldots,8\}$	43.9887	0.4100	0.9946	4.8390	0.5428	0.1032
$\{1,\ldots,9\}$	48.0383	0.3921	0.9959	5.3023	0.5160	0.0986
$\{1,\ldots,10\}$	53.5510	0.3767	0.9965	5.8305	0.4919	0.0946
$\{1,\ldots,11\}$	61.0779	0.3633	0.9968	6.4350	0.4704	0.0913
$\{1,\ldots,12\}$	71.3801	0.3515	0.9970	7.1288	0.4512	0.0883
$\{1,\ldots,13\}$	85.5090	0.3411	0.9971	7.9267	0.4339	0.0857
$\{1,\ldots,14\}$	104.9202	0.3317	0.9971	8.8461	0.4183	0.0833

Example 3. In this example, we consider a well-known BPA defined for all $A \subseteq X$ as

$$m^*(A) = \frac{2^{|A|} - 1}{\sum_{B \subseteq X} (2^{|B|} - 1)},$$

so,

$$\frac{m^*(A)}{2^{|A|} - 1} = \frac{1}{\sum_{B \subseteq X} (2^{|B|} - 1)} = K,$$

where $K \in (0, 1)$ is a constant. The interest of this BPA is due to the fact that it gives a degree of belief to each non-empty subset of the frame of discernment and the mass of a focal element depends only on its cardinality. Then, the fractional version of Tsallis-Deng entropy can be evaluated for the BPA m^* as

$$\begin{aligned} SD_\alpha^q(m^*) &= \frac{1}{\alpha - 1} \sum_{A \subseteq X} \left(2^{|A|} - 1\right) K \left[1 - K^{\alpha-1}\right] (-\log K)^{q-1} \\ &= \frac{1}{\alpha - 1} \left[1 - K^{\alpha-1}\right] (-\log K)^{q-1}. \end{aligned} \tag{11}$$

It is a decreasing function in α since the partial derivative with respect to α of the function in (**??**) has the same sign as that of the function

$$g(x) = x - 1 - x \log x$$

which is non-positive for each $x > 0$. Based on Proposition **??**, $\alpha = 1$ is a removable discontinuity and then the supremum of $SD^q_\alpha(m^*)$ is attained for $\alpha \to 0^+$ and it is given by

$$\lim_{\alpha\to 0^+} \frac{1}{\alpha - 1}\left[1 - K^{\alpha-1}\right](-\log K)^{q-1} = \left[K^{-1} - 1\right](-\log K)^{q-1}.$$

4. Application to classification problems

In this section, the fractional version of Tsallis-Deng entropy in (**??**) is made use of for two classification problems. First, it is used to analyze a dataset given in [**?**] about typical qualities of Italian wines composed of 178 instances and, for each one, thirteen attributes are given. The instances of the dataset are divided into three classes of wine: Class 1, Class 2 and Class 3. Six attributes (Alcohol, Malic acid, Ash, OD280/OD315 of diluted wines (OD), Color intensity (CI) and Proline) are used to classify each instance into the correct class. The method of max–min values is applied to generate a model of interval numbers. For a fixed attribute, the interval of variability in a single class is determined, and then the intervals of more classes are intersected. The model of interval numbers is presented in Table **??**.

Table 3: The model of interval numbers.

Class	Alcohol	Malic Acid	Ash	OD	CI	Proline
1	[12.850, 14.830]	[1.3500, 4.0400]	[2.0400, 3.2200]	[2.5100, 4.0000]	[3.5200, 8.9000]	[680, 1680]
2	[11.030, 13.860]	[0.7400, 5.8000]	[1.3600, 3.2300]	[1.5900, 3.6900]	[1.2800, 6.0000]	[278, 985]
3	[12.200, 14.340]	[1.2400, 5.6500]	[2.1000, 2.8600]	[1.2700, 2.4700]	[3.8500, 13.0000]	[415, 880]
1, 2	[12.850, 13.860]	[1.3500, 4.0400]	[2.0400, 3.2200]	[2.5100, 3.6900]	[3.5200, 6.0000]	[680, 985]
1, 3	[12.850, 14.340]	[1.3500, 4.0400]	[2.1000, 2.8600]	–	[3.8500, 8.9000]	[680, 880]
2, 3	[12.200, 13.860]	[1.2400, 5.6500]	[2.1000, 2.8600]	[1.5900, 2.4700]	[3.8500, 6.0000]	[415, 880]
1, 2, 3	[12.850, 13.860]	[1.3500, 4.0400]	[2.1000, 2.8600]	–	[3.8500, 6.0000]	[680, 880]

Suppose the selected instance is $(12.330, 1.1000, 2.2800, 1.6700, 3.2700, 680)$. It belongs to Class 2 and our aim is to classify it correctly. Six BPAs, one for each attribute, are generated by using a method based on the similarity of interval numbers proposed by Kang et al. [**?**]. Given two intervals $A = [a_1, a_2]$ and $B = [b_1, b_2]$, the similarity index $S(A, B)$ defined by these authors is

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

where $\alpha > 0$ is the coefficient of support (we used $\alpha = 5$ here), and $D(A, B)$ is the distance between intervals A and B defined in [**?**] as

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

For each attribute, seven values of similarity index are obtained by choosing as A the intervals given in Table **??** and as B the corresponding singleton of the selected instance. Then, by normalizing the obtained values, six BPAs are computed, and these are presented in Table **??**.

Without any additional information, the final BPA is determined by giving the same weight to each attribute, i.e., by summing the six values that are related to a focal element and then dividing by six. In this way, we get the final BPA as presented in Table **??**.

Given a BPA, there are different methods for choosing the most appropriate focal element. Obviously, a simple method is to consider the focal element with highest mass. In this example, as our objective is to discriminate between singletons, it will be reasonable to consider the pignistic probability transformation (PPT). The PPT is a point estimate of belief towards focal elements and is defined for $A \subseteq X$ as

$$PPT(A) = \sum_{B:A\subseteq B} \frac{m(B)}{|B|}; \tag{12}$$

Table 4: BPAs based on Kang's method.

Class	Alcohol	Malic Acid	Ash	OD	CI	Proline
$m(1)$	0.1008	0.1645	0.1180	0.1062	0.0977	0.0394
$m(2)$	0.1785	0.1153	0.1098	0.1455	0.2137	0.1084
$m(3)$	0.1381	0.1132	0.1635	0.3259	0.0575	0.1646
$m(1,2)$	0.1445	0.1645	0.1180	0.1171	0.1859	0.1292
$m(1,3)$	0.1191	0.1645	0.1635	0.0000	0.0949	0.1969
$m(2,3)$	0.1745	0.1132	0.1635	0.3053	0.1751	0.1646
$m(1,2,3)$	0.1445	0.1645	0.1635	0.0000	0.1751	0.1969

Table 5: Final BPA.

Class	Final BPA
$m(1)$	0.1045
$m(2)$	0.1452
$m(3)$	0.1605
$m(1,2)$	0.1432
$m(1,3)$	0.1232
$m(2,3)$	0.1827
$m(1,2,3)$	0.1408

see [**?**]. Upon considering the PPT only for singletons, we end up getting a discrete probability distribution. Now, based on the BPA in Table **??**, the PPT of the singleton classes is computed to be

$$PPT(1) = 0.2846, \quad PPT(2) = 0.3551, \quad PPT(3) = 0.3603.$$

Thus, the focal element with the highest PPT is Class 3, and would therefore be our final decision, which is not the correct one in this case.

We now try to improve the described method by using fractional version of Tsallis-Deng entropy in (**??**). Fix the values of $q = 0.5$ and $\alpha = 4$. The fractional version of Tsallis-Deng entropy of BPAs given in Table **??** is then evaluated and the corresponding results are shown in Table **??**.

Table 6: Fractional versions of Tsallis-Deng entropies of BPAs in Table **??**.

Attribute	Alcohol	Malic Acid	Ash	OD	CI	Proline
SD_α^q	0.2085	0.2058	0.2057	0.2464	0.2087	0.2032

Because a greater value of SD_α^q means a higher uncertainty, it is reasonable to give more weight to the attributes with lower SD_α^q. Specifically, we define the weights by normalizing to 1 the reciprocal values of the fourth power of fractional versions of Tsallis-Deng entropies. The weights so determined are reported in Table **??**.

Based on the weights in Table **??**, a weighted version of the final BPA is obtained, as given in Table **??**. Then, based on the weighted BPA in Table **??**, we compute the PPT of the singleton classes to be

$$PPT(1) = 0.2956, \quad PPT(2) = 0.3533, \quad PPT(3) = 0.3510.$$

Thus, the focal element with the highest PPT is Class 2, and would therefore be our final decision, which is indeed the correct decision in this case.

Table 7: The weights of attributes based on fractional version of Tsallis-Deng entropy.

Attribute	**Alcohol**	**Malic Acid**	**Ash**	**OD**	**CI**	**Proline**
Weight	0.1745	0.1838	0.1843	0.0895	0.1741	0.1938

Table 8: Final weighted BPA.

Class	**Final Weighted BPA**
$m(1)$	0.1037
$m(2)$	0.1438
$m(3)$	0.1431
$m(1,2)$	0.1451
$m(1,3)$	0.1358
$m(2,3)$	0.1711
$m(1,2,3)$	0.1542

In Table **??**, the recognition rates of the non-weighted method and methods based on fractional version of Tsallis-Deng entropy are presented for different choices of q and α.

Table 9: The recognition rate for different choice of q and α.

Non-Weighted Method	q	α	**Fractional Tsallis-Deng Method**
93.26%	0.5	4	93.82%
	0.6	3	93.82%
	0.1	0.8	91.57%
	1	5	93.26%

Next, let us consider a different dataset given in [**?**] about Iris flowers which is composed of 150 instances and, for each one, four attributes are given, namely, the sepal length in cm (SL), the sepal width in cm (SW), the petal length in cm (PL) and the petal width in cm (PW). The instances of the dataset are divided into three kind of flowers: Iris Setosa (Se), Iris Versicolour (Ve) and Iris Virginica (Vi). By using the method of max–min values, the model of interval numbers is obtained and is presented in Table **??**.

Suppose the selected instance is $(6.3, 2.7, 4.9, 1.8)$. It belongs to the kind Iris Virginica and our aim is then to classify it correctly. Four BPAs, one for each attribute, are generated by using the similarity of interval numbers as above. Without any additional information, the final BPA is determined by giving the same weight to each attribute, i.e., by summing the four values that are related to a focal element and then dividing by four. In this way, we get the final BPA as presented in Table **??**.

In order to discriminate among classes, we evaluate the PPT of singleton classes for the BPA given in Table **??** and the results are

$$PPT(Se) = 0.1826, \quad PPT(Ve) = 0.4131, \quad PPT(Vi) = 0.4043.$$

Thus, the focal element with the highest PPT is the type Iris Versicolour, and would therefore be our final decision, which is not the correct one in this case. We now try to improve the method by using fractional version of Tsallis-Deng entropy in (**??**). Fix the values of $q = 0.5$ and $\alpha = 0.5$. The fractional version of Tsallis-Deng entropy of BPAs obtained by using the similarity of interval numbers is then evaluated and the corresponding results are shown in Table **??**.

As stated earlier, because a greater value of SD_α^q means a higher uncertainty, it is reasonable to give more weight

Table 10: The model of interval numbers.

Class	SL	SW	PL	PW
Se	$[4.3, 5.8]$	$[2.3, 4.4]$	$[1.0, 1.9]$	$[0.1, 0.6]$
Ve	$[4.9, 7.0]$	$[2.0, 3.4]$	$[3.0, 5.1]$	$[1.0, 1.8]$
Vi	$[4.9, 7.9]$	$[2.2, 3.8]$	$[4.5, 6.9]$	$[1.4, 2.5]$
Se, Ve	$[4.9, 5.8]$	$[2.3, 3.4]$	–	–
Se, Vi	$[4.9, 5.8]$	$[2.3, 3.8]$	–	–
Ve, Vi	$[4.9, 7.0]$	$[2.2, 3.4]$	$[4.5, 5.1]$	$[1.4, 1.8]$
Se, Ve, Vi	$[4.9, 5.8]$	$[2.3, 3.4]$	–	–

Table 11: Final BPA.

Class	Final BPA
$m(Se)$	0.0872
$m(Ve)$	0.1891
$m(Vi)$	0.1861
$m(Se, Ve)$	0.0759
$m(Se, Vi)$	0.0643
$m(Ve, Vi)$	0.3215
$m(Se, Ve, Vi)$	0.1759

to the attributes with lower SD_{α}^{q}. In this case, we define the weights by normalizing to 1 the exponential function of fractional versions of Tsallis-Deng entropies multiplied by one. The weights so determined are reported in Table **??**.

Based on the weights in Table **??**, a weighted version of the final BPA is obtained, as given in Table **??**. Then, based on the weighted BPA in Table **??**, we compute the PPT of the singleton classes to be

$$PPT(1) = 0.1156, \quad PPT(2) = 0.4360, \quad PPT(3) = 0.4485.$$

Thus, the focal element with the highest PPT is the type Iris Virginica, and would therefore be our final decision, which is indeed the correct one in this case.

In Table **??**, the recognition rates of the non-weighted method and methods based on fractional version of Tsallis-Deng entropy are presented for different choices of q and α.

5. Conclusions

In this paper, we have put forward a general formulation of entropy measure which depends on two parameters and includes Shannon, Tsallis and fractional entropy all as special cases. Some properties of this new measure, referred to as fractional Tsallis entropy, have been studied. Moreover, the fractional version of Tsallis-Deng entropy has been defined and analyzed in the context of Dempster-Shafer theory of evidence. Finally, we have presented two applications to classification problems. Based on the results of this paper, it would similarly be of interest to unify the notion through information volume, and we plan to consider this as our future research.

Declaration of interests

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.

Table 12: Fractional versions of Tsallis-Deng entropies of BPAs based on similarity of interval numbers.

Attribute	**Alcohol**	**Malic Acid**	**Ash**	**OD**
SD_{α}^{q}	3.6184	3.7153	2.1226	2.0932

Table 13: The weights of attributes based on fractional version of Tsallis-Deng entropy.

Attribute	**SL**	**SW**	**PL**	**PW**
Weight	0.0912	0.0828	0.4069	0.4191

Acknowledgments

Francesco Buono and Maria Longobardi are members of the research group GNAMPA of INdAM (Istituto Nazionale di Alta Matematica), are partially supported by MIUR - PRIN 2017, project "Stochastic Models for Complex Systems", no. 2017 JFFHSH. The present work was developed within the activities of the project 000009_ALTRI _CDA_75_2021_FRA_LINEA_B funded by "Programma per il finanziamento della ricerca di Ateneo - Linea B" of the University of Naples Federico II. Narayanaswamy Balakrishnan acknowledges the support from the Natural Sciences and Engineering Research Council of Canada through an Individual Discovery Grant. Finally, the authors express their sincere thanks to the Editor and the anonymous reviewers for their incisive comments on an earlier version of this manuscript which led to this improved version.

References

[1] Baratpour, S., Khammar, A. (2016). Tsallis entropy properties of order statistics and some stochastic comparisons. *J. Statist. Res. Iran*, **13**, 25–41.

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

[3] Dempster, A.P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* **38**, 325–339.

[4] Deng, Y. (2016). Deng entropy. *Chaos Solitons Fractals* **91**, 549–553.

[5] Deng, Y. (2020). Information volume of mass function. *International Journal of Computers, Communications & Control*, **15**(6).

[6] Deng, Y. (2020). Uncertainty measure in evidence theory. *Sci China Inf Sci*, **63**(11): 210201, https://doi.org/10.1007/s11432-020-3006-9.

[7] Dua, D.; Graff, C. UCI Machine Learning Repository (2019). Available online: http://archive.ics.uci.edu/ml.

[8] Gao, X., Pan, L., Deng, Y. (2022). A generalized divergence of information volume and its applications. *Engineering Applications of Artificial Intelligence*, **108**, 104584.

[9] Gao, Q., Wen, T., Deng, Y. (2021). Information volume fractal dimension. *Fractals*, doi:10.1142/S0218348X21502637.

[10] 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. *Acta Electron. Sin.* **40**, 1092–1096.

[11] Kazemi, M.R., Tahmasebi, S., Buono, F., Longobardi, M. (2021). Fractional Deng entropy and extropy and some applications. *Entropy*, **23**, 623.

[12] Kumar, V. (2016). Some results on Tsallis entropy measure and k-record values. *Physica A*, **462**, 667–673.

[13] Liu, F., Gao, X., Deng, Y. (2019). Generalized belief entropy and its application in identifying conflict evidence. *IEEE Access*, **7**, 126625–126633, doi: 10.1109/ACCESS.2019.2939332.

[14] Sati, M.M., Gupta, N. (2015). Some characterization results on dynamic cumulative residual Tsallis entropy. *J Probab. Stat.*, **2015**, 694203.

[15] Shafer, G.*A Mathematical Theory of Evidence*; Princeton University Press: Princeton, NJ, USA, 1976.

[16] Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.

[17] Smets, P. (2000). Data fusion in the transferable belief model. In: *Proceedings of the Third International Conference on Information Fusion*, Paris, France, 10–13 July 2000, **1**, PS21–PS33.

[18] Tran, L., Duckstein, L. (2002). Comparison of fuzzy numbers using a fuzzy distance measure. *Fuzzy Sets Syst.* **130**, 331–341.

[19] Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistic. *Journal of Statistical Physics*, **52**, 479–487.

[20] Ubriaco M.R. (2009). Entropies based on fractional calculus. *Physics Letters A*, **373**, 2516–19.

[21] Zhou, Q., Deng, Y. (2021). Higher order information volume of mass function. *Information Sciences.,* arXiv:2012.07507v1.

Table 14: Final weighted BPA.

Class	**Final Weighted BPA**
Se	0.0825
Ve	0.2050
Vi	0.2194
Se, Ve	0.0262
Se, Vi	0.0224
Ve, Vi	0.4182
Se, Ve, Vi	0.0262

Table 15: The recognition rate for different choice of q and α.

Non-Weighted Method	q	α	**Fractional Tsallis-Deng Method**
94%	0.5	0.5	96.67%
	0.25	0.6	94.67%
	0.8	0.75	96.67%
	0.75	2	94%
	1	5	94%