sciendo

Ann. Univ. Paedagog. Crac. Stud. Math. 23 (2024), 5-12
DOI: 10.2478/aupcsm-2024-0001

FOLIA 386

Annales Universitatis Paedagogicae Cracoviensis
Studia Mathematica 23 (2024)

Slimane Benmahmoud

Some theoretical results on fractional-order continuous information measures

Abstract. By rewriting the differential entropy in a form of a differ-integral function's limit, and deforming the ordinary derivative to a fractional-order one, we derive in this paper a novel generalized fractional-order differential entropy along with its related information measures. When the order of fractional differentiation $\alpha \to 1$, the ordinary Shannon's differential entropy is recovered, which corresponds to the results from first-order ordinary differentiation.

1. Introduction

Inspired by the concept of entropy introduced by Clausius [2] thermodynamics and Boltzmann [1] in classical statistical mechanics, Shannon proposed his entropy in the context of communication theory [12]. It consists of a measure of surprise or uncertainty associated with the probability distribution of a random variable (RV). For a discrete RV X taking values in $\mathcal{X} = \{x_1, x_2, \dots, x_q\}$ and having a probability mass function $p_i = P(X = x_i)$ with $\sum_{i=1}^{q} p_i = 1$ and $p_i \geq 0$ for $i = 1, \dots, q$, it is given by

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

AMS (2010) Subject Classification: 44-XX, 54C70, 94A17, 42A38, 42B10.
Keywords and phrases: fractional differential entropy, fractional joint/conditional differential entropy, fractional relative entropy, fractional mutual information, Riemann-Liouville fractional integral/derivative.
ISSN: 2081-545X, e-ISSN: 2300-133X.

The principal properties of this suggested measure of uncertainty (i.e. Eq (1)) has shown a perfect agreement with the intuitive notions of randomness and justified its usefulness with respect to statistical problems in communication theory.

The continuous analogue of (1) is known as the differential entropy. For a continuous RV having a probability density function (pdf) $f_X(x)$ and a support $S_X \subseteq \mathbb{R}$, it is given by

$$h(X) := -\int_{S_X} f_X(x) \log f_X(x) dx. \tag{2}$$

However, this straightforward extension of the concept of entropy from discrete to continuous schemes raises some concerns including

1) it may be negative;
2) it may become infinity large;
3) it doesn't remain necessarily invariant under the transformation of the coordinate systems.

For further discussion on these and more concerns one may refer to [5].

In this paper, inspired by Ubriaco's work in [11], we propose a fractional-order continuous entropy (and its related fractional information measures) by rewriting the differential entropy in (2) in the form of a differ-integral function, then we deform the ordinary derivative to a fractional-order one by recourse to the fractional calculus (FC) theory. The use of FC allows to measure the information in a generalized metric space.

The remainder of this paper is organized as follows. In section 2, we give a brief survey on the theory of FC, and we re-call some definitions concerning the RL fractional integral and derivative. The main results on fractional-order information measures are then presented in section 3. Finally, section 4 concludes the paper.

2. Fractional calculus

2.1. A brief survey

Fractional calculus (FC) is a mathematical analysis branch which studies different possible approaches of defining fractional-order integrals and derivatives. Based on FC, the theory of classical integer-order differential equations has been then generalized to the broader theory of fractional-order differential equations. FC can be traced back to a letter written to l'Hopital by Leibniz in 1695 [7]. In 1832, Liouville carried out a heavy-handed investigation on FC [8]. After that, the Riemann-Liouville (RL) fractional integro-differential operator was introduced by Riemann in [10] along with a comprehensive theory of FC. FC has led to many breakthroughs in different fields of physics and engineering where various processes can be modeled in a more accurate and authentic way [4].

2.2. A review on Rieman-Liouville fractional integral and derivative

The left-sided Rieman-Liouville (RL) fractional integral ${}^{RL}I^{\alpha}_{a+}f$ of order $\alpha \in \mathbb{R}$, $\alpha > 0$ of an integrable function $f\colon [a,b] \to \mathbb{R}$, $0 \le a < b \le \infty$ is defined as [6],

$$({}^{RL}I^{\alpha}_{a+}f)(t) = {}^{RL}I^{\alpha}_{a+}[f(x)](t) = \frac{1}{\Gamma(\alpha)}\int_a^t (t-x)^{\alpha-1}f(x)dx$$

for $a \in \mathbb{R},\ t > a,\ \alpha > 0$, where $\Gamma(.)$ is Euler's gamma function defined as $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}dx$ for $\alpha \in \mathbb{R}$, $\alpha > 0$.

The left-sided Rieman-Liouville (RL) fractional derivative ${}^{RL}D^{\alpha}_{a+}f$ of order $\alpha \in \mathbb{R}$, $\alpha > 0$ of an integrable and differentiable function $f\colon [a,b] \to \mathbb{R}$, $0 \le a < b \le \infty$ is defined as [6],

$$({}^{RL}D^{\alpha}_{a+}f)(t) = {}^{RL}D^{\alpha}_{a+}[f(x)](t) = \left(\frac{d}{dt}\right)^n ({}^{RL}I^{n-\alpha}_{a^+}f)(t)$$

for $a \in \mathbb{R},\ t > a,\ \alpha > 0,\ n = [\alpha] + 1$. When $0 < \alpha < 1$, we get

$$({}^{RL}D^{\alpha}_{a+}f)(t) = \frac{d}{dt}({}^{RL}I^{1-\alpha}_{a^+}f)(t), \qquad a \in \mathbb{R},\ t > a. \tag{3}$$

3. Fractional calculus-based information measures

Definition 1 (Shannon's differential entropy, see [3])
The differential entropy of a continuous random variable (RV) X with a probability density function (pdf) $f_X(x)$ and a support $S_X \subseteq \mathbb{R}$ is defined as

$$h(X) := -\int_{S_X} f_X(x)\log f_X(x)dx = E[-\log f_X(X)], \tag{4}$$

when the integral exists.

Through the whole paper, $E[X]$ is the expected value of the variable X and the base of the logarithm will be set to Euler's number $e = \sum_{n=0}^{\infty}\frac{1}{n!}$.

Our basic idea consists of rewriting (4) as follows

$$h(X) := \lim_{t \to -1}\frac{d}{dt}\int_{S_X} f_X^{-t}(x)dx. \tag{5}$$

Then, we deform the ordinary differential operator $\frac{d}{dt}$ in (5) to the RL-fractional differential operator ${}^{RL}D^{\alpha}_{a+}$ defined in (3) (which reduces to $\frac{d}{dt}$ in the limit $\alpha \to 1$). Based on these ideas, we define the following fractional-order information measures.

3.1. The fractional differential entropy

Theorem 1 (The fractional differential entropy)
The fractional differential entropy of a continuous RV X with a pdf $f_X(x)$ and a support $S_X \subseteq \mathbb{R}$ is defined as

$$h^{\alpha}(X) := \int_{S_X} f_X(x)(-\log f_X(x))^{\alpha}dx = E[-(\log f_X(X))^{\alpha}], \qquad 0 < \alpha < 1. \tag{6}$$

Proof. Using the operator defined in (3) (for $a = -\infty$), (5) can be rewritten as follows

$$h^\alpha(X) := \lim_{t\to -1} \frac{d}{dt}\left({}^{RL}I^{1-\alpha}_{-\infty}\left[\int_{S_X} e^{-t\log f_X(x)}dx\right](t)\right), \qquad 0<\alpha<1.$$

Therefore, we need to solve the following integral

$$h^\alpha(X) := \frac{1}{\Gamma(1-\alpha)} \lim_{t\to -1} \frac{d}{dt} \int_{S_X}\left(\int_{-\infty}^{t} (t-y)^{-\alpha} e^{-y\log f_X(x)}dy\right)dx.$$

By letting $w = t - y$, using the definition of the $\Gamma(.)$ function, taking the ordinary derivative and setting $t = 1$, we get (6).

In the following, we give few examples of the proposed fractional entropy for some common continuous probability distributions.

Example 1
Let X be a uniformly distributed RV in $[0, a]$, $a > 0$, i.e. $X \sim \text{Uniform}(0, a)$. Then, its fractional entropy is given by

$$h^\alpha(X) = (\log a)^\alpha.$$

Example 2
Let X be an exponentially distributed RV with a rate parameter λ, i.e. $X \sim \text{Exp}(\lambda)$. Then, its fractional entropy is given by

$$h^\alpha(X) = \frac{\Gamma(\alpha+1, -\log\lambda)}{\lambda},$$

where $\Gamma(., .)$ is the upper incomplete gamma function (See (8.2.2) in [9]).

Example 3
Let X be a normally distributed RV with a mean μ and a standard deviation σ, i.e. $X \sim \mathcal{N}(\mu, \sigma)$. Then, its fractional entropy is given by

$$h^\alpha(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \left(\frac{1}{2}\log(2\pi\sigma^2) + \frac{(x-\mu)^2}{2\sigma^2}\right)^\alpha dx.$$

3.2. The fractional joint and conditional differential entropy

Theorem 2 (The fractional joint differential entropy)
Let X and Y be two jointly distributed continuous RVs with a joint pdf $f_{X,Y}(x, y)$ and a support $S_{X,Y} \subseteq \mathbb{R}^2$. Then, the fractional joint differential entropy of X and Y is defined as

$$\begin{aligned} h^\alpha(X,Y) &:= \int_{S_{X,Y}} f_{X,Y}(x,y)(-\log f_{X,Y}(x,y))^\alpha dxdy \\ &= E[-(\log f_{X,Y}(X,Y))^\alpha], \qquad 0<\alpha<1. \end{aligned} \tag{7}$$

Proof. For two jointly distributed continuous RVs X and Y with a joint pdf $f_{X,Y}(x,y)$ and a support $S_{X,Y} \subseteq \mathbb{R}^2$, the Shannon's joint differential entropy is defined as

$$h(X,Y) := -\int_{S_{X,Y}} f_{X,Y}(x,y) \log f_{X,Y}(x,y) dxdy.$$

Considering the idea in (5) and using the operator defined in (3) (for $a = -\infty$), the previous equation can be rewritten as follows

$$h^{\alpha}(X,Y) := \lim_{t\to -1} \frac{d}{dt}\left({}^{RL}I^{1-\alpha}_{-\infty}\left[\int_{S_{X,Y}} e^{-t\log f_{X,Y}(x,y)} dxdy \right](t) \right), \qquad 0<\alpha<1.$$

Following the same thoughts in the proof of (5), we get (7).

Theorem 3 (The fractional conditional differential entropy)
Let X and Y be two jointly distributed continuous RVs with a joint pdf $f_{X,Y}(x,y)$ and a support $S_{X,Y} \subseteq \mathbb{R}^2$ such that the conditional pdf of X given Y, given by

$$f_{X/Y}(x/y) = \frac{f_{X,Y}(x,y)}{f_Y(y)},$$

is well defined for all $(x,y) \in S_{X,Y}$, where $f_Y(y)$ is the marginal pdf of Y. Then, the fractional conditional differential entropy of X given Y is defined as

$$\begin{aligned} h^{\alpha}(X/Y) &:= \int_{S_{X,Y}} f_{X,Y}(x,y)(-\log f_{X/Y}(x/y))^{\alpha} dxdy \\ &= E[-(\log f_{X/Y}(X/Y))^{\alpha}], \qquad 0<\alpha<1. \end{aligned} \tag{8}$$

Proof. For the two RVs in Theorem 3, the Shannon's conditional differential entropy is defined as

$$h(X/Y) := -\int_{S_{X,Y}} f_{X,Y}(x,y) \log f_{X/Y}(x/y) dxdy.$$

Considering the idea in (5) and using the operator defined in (3) (for $a = -\infty$), the previous equation can be rewritten as follows

$$h^{\alpha}(X/Y) := \lim_{t\to -1} \frac{d}{dt}\left({}^{RL}I^{1-\alpha}_{-\infty}\left[\int_{S_{X,Y}} e^{-t\log f_{X/Y}(x/y)} dxdy \right](t) \right), \qquad 0<\alpha<1.$$

Following the same thoughts in the previous proofs, we get (8).

3.3. The fractional divergence and mutual information

Theorem 4 (The fractional divergence)
Let X and Y be two continuous RVs with marginal pdfs $f_X(x)$ and $f_Y(y)$, respectively, such that their supports satisfy $S_X \subseteq S_Y \subseteq \mathbb{R}$. Then, the fractional

divergence (or fractional relative entropy or fractional Kullback-Leibler distance) between X and Y is defined as

$$\begin{aligned} D^\alpha(X\|Y) &:= \int_{S_X} f_X(x)\Big(\log\frac{f_X(x)}{f_Y(x)}\Big)^\alpha dx \\ &= E\Big[\Big(\log\frac{f_X(X)}{f_Y(X)}\Big)^\alpha\Big], \qquad 0<\alpha<1. \end{aligned} \tag{9}$$

Proof. For the two RVs in Theorem 4, the Kullback-Leibler's distance is defined as

$$D(X\|Y) := \int_{S_X} f_X(x)\Big(\log\frac{f_X(x)}{f_Y(x)}\Big)dx.$$

Considering the idea in (5) and using the operator defined in (3) (for $a=-\infty$), the previous equation can be rewritten as follows

$$D^\alpha(X\|Y) := \lim_{t\to-1}\frac{d}{dt}\Big({}^{RL}I^{1-\alpha}_{-\infty}\Big[\int_{S_X} f_Y(x)e^{-t\log\frac{f_X(x)}{f_Y(y)}}dx\Big](t)\Big), \qquad 0<\alpha<1.$$

Following the same thoughts in the previous proofs, we get (9).

Theorem 5 (The fractional mutual information)
Let X and Y be two jointly distributed continuous RVs with a joint pdf $f_{X,Y}(x,y)$ and a support $S_{X,Y}\subseteq\mathbb{R}^2$. Then, the fractional mutual information between X and Y is defined as

$$\begin{aligned} I^\alpha(X;Y) &:= \int_{S_X} f_{X,Y}(x,y)\Big(\log\frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}\Big)^\alpha dx \\ &= E\Big[\Big(\log\frac{f_{X,Y}(X,Y)}{f_X(X)f_Y(Y)}\Big)^\alpha\Big], \qquad 0<\alpha<1. \end{aligned} \tag{10}$$

assuming the integral exists, where $f_X(x)$ and $f_Y(y)$ are the marginal pdfs of X and Y, respectively.

Proof. For the two RVs in Theorem 5, the Shannon's mutual information is defined as follows

$$I(X;Y) := \int_{S_X} f_{X,Y}(x,y)\Big(\log\frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}\Big)dx.$$

Considering the idea in (5) and using the operator defined in (3) (for $a=-\infty$), the previous equation can be rewritten as follows

$$I^\alpha(X\|Y) := \lim_{t\to-1}\frac{d}{dt}\Big({}^{RL}I^{1-\alpha}_{-\infty}\Big[\int_{S_X} f_X(x)f_Y(y)e^{-t\log\frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}}dxdy\Big](t)\Big)$$

for $0<\alpha<1$. Following the same thoughts in the previous proofs, we get (10).

Remark 1
$I^\alpha(X;Y)$ is symmetrical in X and Y.

Remark 2
The expressions in Theorems 1-5 allow us to calculate the fractional-order information measures for different values of α and compare them to their conventional counterparts (i.e. when $\alpha=1$) to get further insights on their behaviours.

4. Conclusion

In this work, we have introduced a new fractional-order differential entropy $h^{\alpha}(X)$ along with its related fractional-order information measures, which generalizes the Shannon's conventional differential entropy $h(X)$, in the context of FC. Some illustrative examples on the fractional entropy of some common continuous probability distributions are given. The information measures – we have developed in this work – may find their applications in many technical fields such as computer vision and information theory where the information measure in general and the entropy in special, are aspects with a great importance.

References

[1] Boltzmann, Ludwig. "Lectures on the Principles of Mechanics." In *Theoretical Physics and Philosophical Problems* editeg by B. McGuinness. Dordrecht: Springer, 1974. Cited on 5.

[2] Clausius, Rudolf and Thomas Archer Hirst. *The Mechanical Theory of Heat: With its Applications to the Steam-Engine and to the Physical Properties of Bodies*. London: J. Van Voorst, 1867. Cited on 5.

[3] Cover, Thomas M., and Joy A. Thomas. *Elements of Information Theory*. New-York: John Wiley & Sons, INC., 1991. Cited on 7.

[4] Hilfer, Rudolf. *Applications Of Fractional Calculus In Physics*. Universität Mainz & Universität Stuttgart, 2000. Cited on 6.

[5] Jumarie, Guy. *Relative Information: Theories and Applications*. Vol. 47 of *Springer Series in Synergetics*. Heidelberg: Springer Berlin, 2011. Cited on 6.

[6] Kilbas, Anatoly A., and Hari Mohan Srivastava, and Juan J. Trujillo. *Theory and Applications of Fractional Differential Equations*. Oxford: Elsevier LTD, 2006. Cited on 7.

[7] Leibniz G.W. "Letter from Hanover; Germany; to G.F.A. L'Hopital, September 30; 1695." In: *Mathematische Schriften*, 301-302. Hidesheim: Olms Verlag, 1962. Cited on 6.

[8] Liouville, Joseph. *Mémoire sur le calcul des différentielles à indices quelconques*. 1832. Cited on 6.

[9] Olver, Frank W.J. et al. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010. Cited on 8.

[10] Riemann, Bernhard. "Versuch Einer Allgemeinen Auffassung Der Integration Und Differentiation. (1847.)." In *Bernard Riemann's Gesammelte Mathematische Werke Und Wissenschaftlicher Nachlass*, edited by Richard Dedekind and Heinrich Martin Weber, 331-344. Cambridge: Cambridge University Press, 2013. Cited on 6.

[11] Ubriaco, Marcelo R. "Entropies based on fractional calculus." *Physics Letters A* 373 (2009): 2516-2519. Cited on 6.

[12] Shannon, Claude E. "A mathematical theory of communication." *Bell Syst. Tech. J.* 27 (1948): 379-423. Cited on 5.

Slimane Benmahmoud
Faculty of Technology, Department of Electronic Engineering
University of M'sila
M'sila
Algeria
E-mail: slimane.benmahmoud@univ-msila.dz

Received: December 27, 2022; final version: December 12, 2023; available online: January 24, 2024.