

Article

Fractional Order Generalized Information

José Tenreiro Machado

Institute of Engineering, Polytechnic of Porto, Department of Electrical Engineering,
Rua Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal;
E-Mail: jtm@isep.ipp.pt; Tel.: +351-22-8340500; Fax: +351-22-8321159

Received: 3 April 2014; in revised form: 16 April 2014 / Accepted: 21 April 2014 /

Published: 24 April 2014

Abstract: This paper formulates a novel expression for entropy inspired in the properties of Fractional Calculus. The characteristics of the generalized fractional entropy are tested both in standard probability distributions and real world data series. The results reveal that tuning the fractional order allow an high sensitivity to the signal evolution, which is useful in describing the dynamics of complex systems. The concepts are also extended to relative distances and tested with several sets of data, confirming the goodness of the generalization.

Keywords: information; entropy; fractional calculus

1. Introduction

During the last decades the scientific community paid considerable attention to the generalization of concepts such as information, entropy [1–4] and differentiation [5–8]. Entropy was introduced in thermodynamics by Clausius and Boltzmann and was later adopted by Shannon and Jaynes in information theory [9–11]. Fractional Calculus (FC) was introduced by Leibniz in mathematics and found application in the areas of biology, physics and engineering [12–18]. The progress motivated the formulation of novel entropy indices and fractional operators, often relaxing some properties and allowing their application in complex dynamical systems [19–21].

The generalized concepts motivate further developments and new research avenues emerge. Bearing these ideas in mind, the present study combines both concepts and is organized as follows. Section 2 introduces entropy and fractional calculus in order to formulate the new generalized fractional entropy. Section 3 applies the new index in several types of data, namely two mathematical induced series, the digits of number π [22] and the Weierstrass function, two financial time series, the Dow Jones Industrial Average and the Europe Brent Spot Price [23,24], and one genomic series, the Human chromosome

Y [25]. The results are analysed and distinct entropy formulations, for several fractional orders, compared. Section 4 expands the proposed index towards the concepts of distance. The Kullback-Leibler and Jensen-Shannon divergence measures are revisited and rewritten in the light of the fractional perspective. The performance of the index is tested with two sets of data, namely 13 irrational numbers and the whole 24 Human chromosomes, adopting the fractional order that reveals higher sensitivity. Finally, Section 5 outlines the main conclusions.

2. Fractional Generalization of Entropy

Information theory was developed by Claude Shannon in 1948 [26,27] and has been applied in many scientific areas. The fundamental cornerstone is the information content of some event having probability of occurrence p_i :

$$I(p_i) = -\ln p_i \tag{1}$$

The expected value, called Shannon entropy [28,29], becomes:

$$S = E(-\ln p) = \sum_i (-\ln p_i)p_i \tag{2}$$

where $E(\cdot)$ denotes the expected value operator.

Expression (2) obeys the four Khinchin axioms [30,31] and several generalizations of entropy have been proposed, obeying only a sub-set of them.

Recently Ubriaco brought together information theory and FC and proposed [32] the expression:

$$S_q = E[(-\ln p)^q] = \sum_i (-\ln p_i)^q p_i \tag{3}$$

where $0 \leq q \leq 1$ denotes the “order” so that $q = 1$ yields Expression (2). This formulation obeys the same properties as the Shannon entropy except additivity and is the expected value of information content given by:

$$I_q(p_i) = (-\ln p_i)^q \tag{4}$$

It is well known in FC the adoption of a power function for obtaining intermediate values, that is, for “fractionating” classical integer operators. In brief, the Laplace transform of the fractional derivative of order $\alpha \in \mathbb{R}$ of a signal $x(t)$ with zero initial conditions is given by:

$$\mathcal{L}\{ {}_0D^\alpha x(t) \} = s^\alpha \mathcal{L}\{ x(t) \} \tag{5}$$

where t represents time, and $\mathcal{L}\{\cdot\}$ and s denote the Laplace operator and variable, respectively.

This property motivated the approximation of fractional derivatives by expanding the factor s^α both with the Fourier and the \mathcal{Z} transforms [33,34]. However, the adoption by means of a power function is related with transforms and we can design a distinct fractional approach for information and entropy. In fact, we can think of Shannon information $I(p_i) = -\ln p_i$ between the cases $D^{-1}I(p_i) = p_i(1 - \ln p_i)$ and $D^1I(p_i) = -\frac{1}{p_i}$, which, in the perspective of FC, leads to the proposal of information and entropy of order $\alpha \in \mathbb{R}$ given by [35]:

$$I_\alpha(p_i) = D^\alpha I(p_i) = -\frac{p_i^{-\alpha}}{\Gamma(\alpha + 1)} [\ln p_i + \psi(1) - \psi(1 - \alpha)] \tag{6}$$

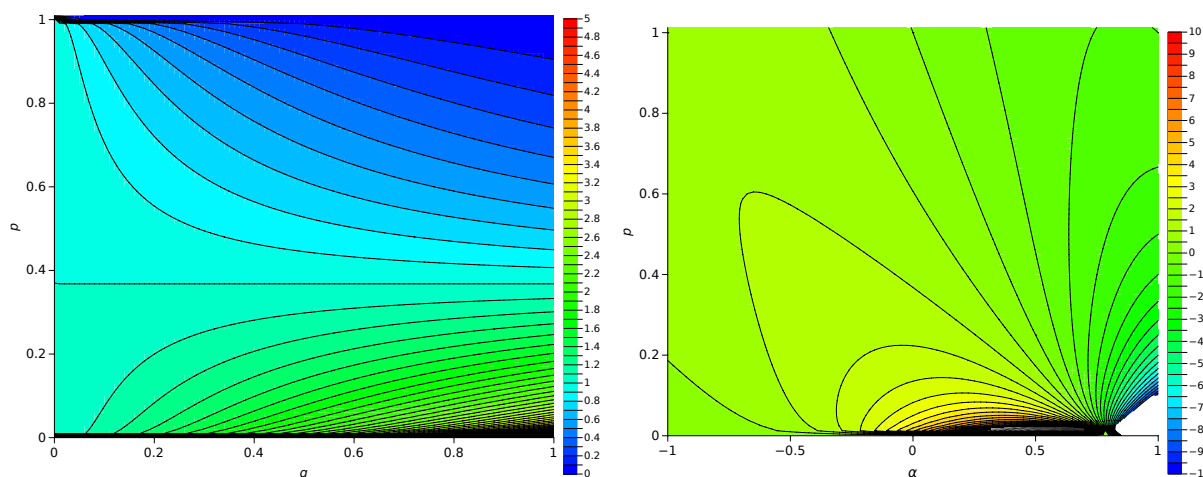
$$S_\alpha = \sum_i \left\{ -\frac{p_i^{-\alpha}}{\Gamma(\alpha + 1)} [\ln p_i + \psi(1) - \psi(1 - \alpha)] \right\} p_i \tag{7}$$

where $\Gamma(\cdot)$ and $\psi(\cdot)$ represent the gamma and digamma functions.

Expression (7) fails to obey some of the Khinchin axioms with exception of the case $\alpha = 0$ that leads to the classical Shannon entropy. This behaviour is in line with what occurs in FC, where fractional derivatives fail to obey some of the properties of integer-order operators. By other words, in both cases, by generalizing operators we loose some classical properties.

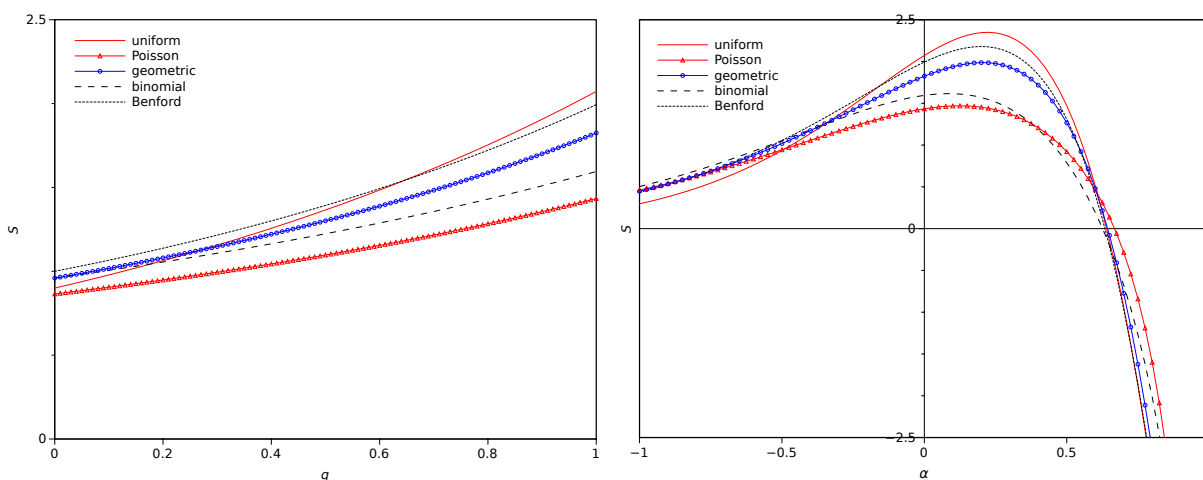
Figure 1 shows the locus of I_q versus (q, p) , $0 \leq q \leq 1$, and I_α versus (α, p) , $-1 \leq \alpha \leq 1$. We observe that I_q has a smaller amplitude excursion than I_α . Moreover, we verify that I_α takes not only positive, but also negative values for $\alpha > 0$. Therefore, Expression (6) assumes also the assumption that we can have negative information, that, for a given value of $\alpha > 0$, can be interpreted as derived from “misleading events”. While exploring the concept of “deception” is not the objective of the present paper, we should note that related ideas were addressed, in abstract terms, in the scope of negative probabilities [36–41] and, in practical terms, in the scope of robotics [42]. In short, we can say that the parameter α allows us to tune the level of confidence of the information varying from positive (trustworthy) up to negative (deceptive) information.

Figure 1. Variation of information: I_q versus (q, p) , $0 \leq q \leq 1$ (left) and I_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).



In order to illustrate the behaviour of the new index and to compare the two approaches Figure 2 depicts the entropies S_q and S_α for the uniform, Poisson ($\alpha = 2$), geometric ($p = 0.3$), binomial ($p = 0.3$), and Benford probability distributions. We verify that S_q has a much smaller variation with q than S_α with α . There is a large similarity between the shape of the curves for $0 < q \leq 1$ and $-0.5 < \alpha \leq 0$. This is natural since S_q tends to the traditional entropy when $q \rightarrow 1$, while S_α tends to the traditional entropy when $\alpha \rightarrow 0$. Furthermore, we verify that S_α has maxima for $0.07 < \alpha < 0.23$ and reaches null values for $0.62 < \alpha < 0.68$. Therefore, in a practical application we can adopt values for α in the first range if information is reliable, or we can consider values of α in the second range if data contains misleading information.

Figure 2. Entropy variation for the uniform, Poisson ($\alpha = 2$), geometric ($p = 0.3$), binomial ($p = 0.3$), Benford probability distributions: S_q^{bin} versus q (left) and S_α^{bin} versus α (right).



Usually it is of interest to investigate the evolution for a binary distribution, that in our case leads to the expressions:

$$S_q^{bin} = p(-\ln p)^q + (1-p)[-\ln(1-p)]^q \tag{8}$$

$$S_\alpha^{bin} = p \left\{ -\frac{p^\alpha}{\Gamma(\alpha+1)} [\ln(p) + \psi(1) - \psi(1-\alpha)] \right\} + (1-p) \left\{ -\frac{(1-p)^\alpha}{\Gamma(\alpha+1)} [\ln(1-p) + \psi(1) - \psi(1-\alpha)] \right\} \tag{9}$$

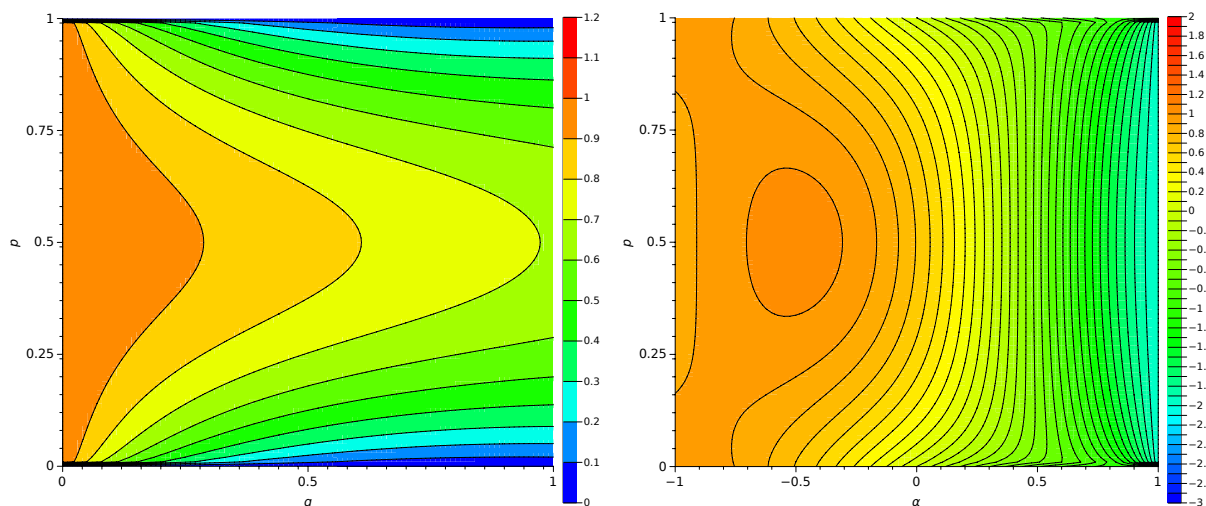
Figure 3 shows the locus of S_q and S_α versus (p, q) , $0 \leq q \leq 1$, and (α, p) , $-1 \leq \alpha \leq 1$, respectively. In both cases we have a symmetrical variation relatively to $p = 0.5$, but S_q is less sensitive than S_α to the variation of the order. In the case of S_α we observe that the chart passes from convex to concave in the region of $\alpha = 0.5$.

3. Application of the Generalized Entropy

This section applies S_q and S_α to the mathematical constant π , the Weierstrass function, the Dow Jones Industrial Average (closing values) and the Europe Brent Spot Price (USD per barrel) financial time series, and one genomic series, the Human chromosome Y. The mathematical constant π is expanded in base 10, and each digit is considered separately in the series. In the Weierstrass function, $f(\xi) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi \xi)$ are adopted the parameters $a = 0.5$, $b = 3$ and the range $-2 \leq \xi \leq 2$. The two financial series correspond to daily values during the period 18 May 1987 up to 14 March 2014. In the four cases we adopt a total of $L = 7000$ data values. For the calculation of the histograms of relative frequency a non-overlapping sliding time window of $W = 100$ points is adopted, producing a total of $k = 1, \dots, 70$ samples. In the case of the genomic series we have four bases denoted $\{A, C, T, G\}$ that are sampled in groups of 3 producing histograms with 4^3 bins. A small percentage of triplets involving the symbol N (considered as “not useful” in genomics) are not analysed. Therefore, a sequence of size $L = 872 \cdot 10^4$ is adopted and two distinct non-overlapping sliding windows, of

$W_1 = 10,000$ and $W_2 = 124,571$ points each, are considered producing a total of $k = 1, \dots, 872$ and $k = 1, \dots, 70$ samples, respectively.

Figure 3. Variation of entropy: S_q^{bin} versus (q, p) , $0 \leq q \leq 1$ (left) and S_α^{bin} versus (α, p) , $-1 \leq \alpha \leq 1$ (right).



Figures 4 and 5 represent S_q versus (q, p) , and S_α versus (α, p) , for the π series and the Weierstrass function, respectively. We observe that S_q has a low sensitivity to the dynamics of the series exhibiting significant variations only for q close to one, that is, when it reduces to the Shannon entropy. On the other hand, S_α detects clearly dynamical variations, being particularly sensitive in the region $0 < \alpha < 0.6$.

Figures 6 and 7 depict the plots of S_q and S_α for the Dow Jones Industrial Average and Europe Brent Spot Price, respectively. We verify a behaviour similar to the one pointed out previously.

Figure 4. Entropy variation for the π series: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).

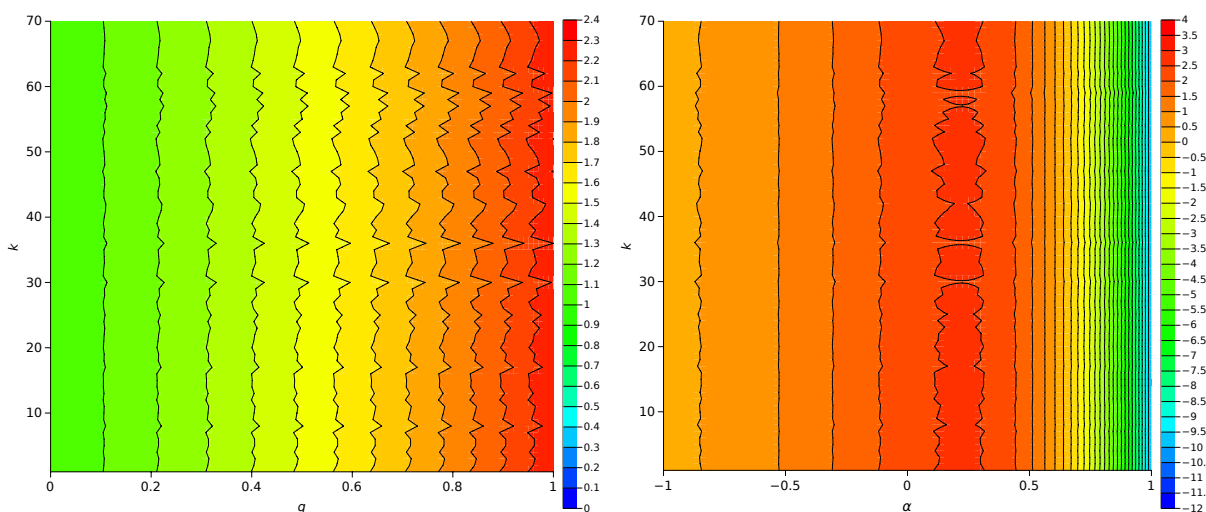


Figure 5. Entropy variation for the Weierstrass function: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).

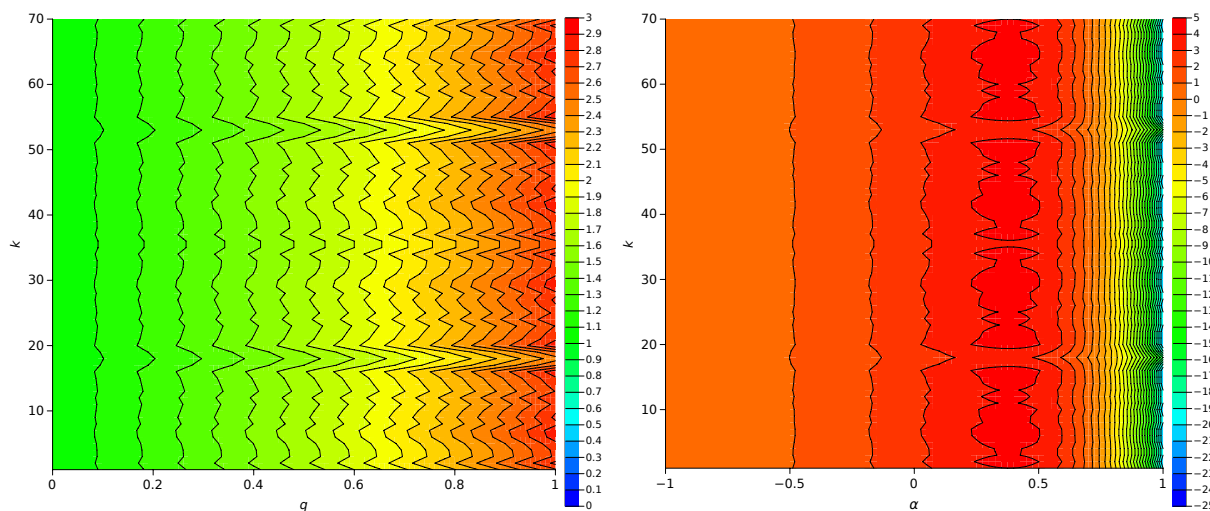
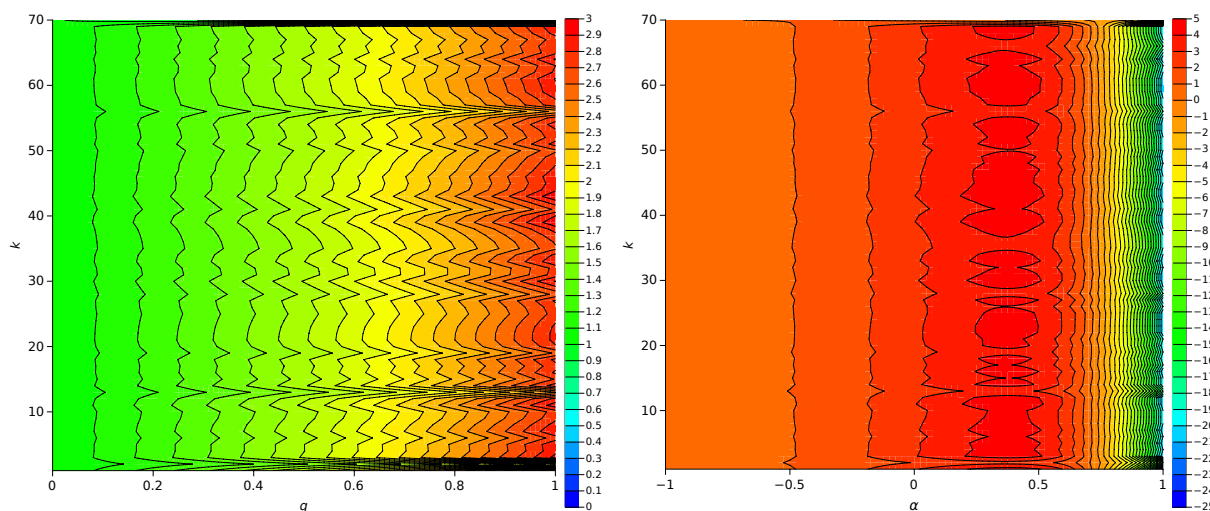


Figure 6. Entropy variation for the Dow Jones Industrial Average time series: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).



Finally, Figures 8 and 9 show S_q and S_α for the Human chromosome Y, with the only difference being the size and number of sliding windows. As previously the higher sensitivities occur for $q = 1$ with S_q and for $\alpha = 0.5$ with S_α . The sliding window W_1 is more appropriate for highlighting dynamical evolutions than window W_2 that is considerable large and leads to an “averaging” of the information content of the chromosome series.

It is interesting to note that the average entropy over the complete data series characterizes the type of embedded information. In fact, the maxima values are $(\alpha, S_\alpha^{av}) = (0.225, 2.52)$, $(\alpha, S_\alpha^{av}) = (0.375, 4.08)$ and $(\alpha, S_\alpha^{av}) = (0.40, 0.217)$, for the $\{\pi\}$, {Weierstrass, Dow Jones Industrial Average, Europe Brent Spot Price} and {Human chromosome Y} (both windows) data series, respectively. The results remain identical for other numerical constants and chromosomes to be discussed in the next section.

Figure 7. Entropy variation for the Europe Brent Spot Price time series: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).

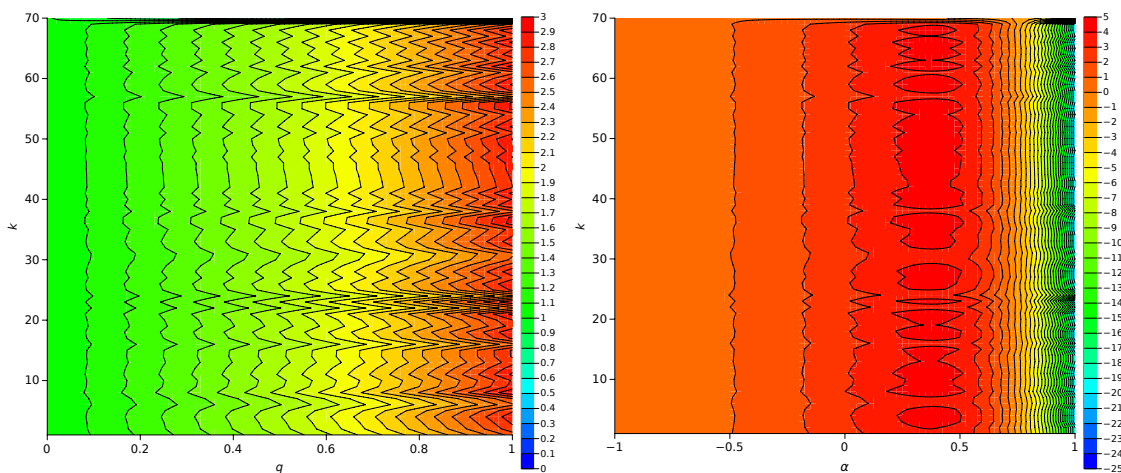


Figure 8. Entropy variation for the Human chromosome Y, $W_1 = 10,000$: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).

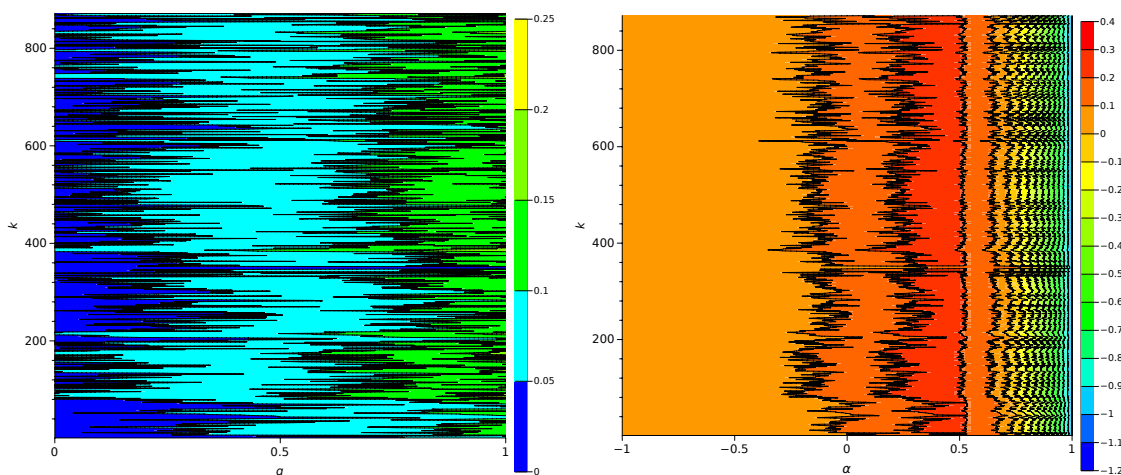
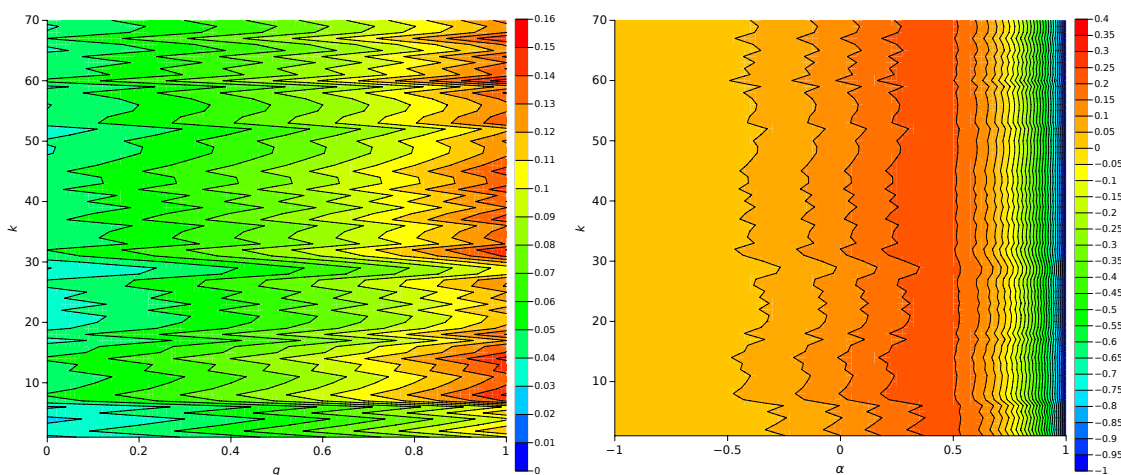


Figure 9. Entropy variation for the Human chromosome Y, $W_2 = 124,571$: S_q versus (q, p) , $0 \leq q \leq 1$ (left) and S_α versus (α, p) , $-1 \leq \alpha \leq 1$ (right).



4. Application of the Generalized Entropy

In this section we explore further the concept of generalized fractional information and entropy. We start by recalling the Kullback-Leibler divergence of Q from P defined as [43–47]:

$$D_{KL}(P \parallel Q) = \sum_i p_i \ln \frac{p_i}{q_i} \tag{10}$$

The Jensen-Shannon divergence $JSD(P \parallel Q)$ is defined as:

$$JSD(P \parallel Q) = \frac{1}{2} [D_{KL}(P \parallel M) + D_{KL}(Q \parallel M)] \tag{11}$$

where $M = \frac{P+Q}{2}$.

Alternatively, we can calculate $JSD(P \parallel Q)$ as:

$$JSD(P \parallel Q) = \frac{1}{2} \left[\sum_i p_i \ln p_i + \sum_i q_i \ln q_i \right] - \sum_i m_i \ln m_i \tag{12}$$

Having in mind Expressions (4), (6) and (12), the fractional JSD can be written as:

$$JSD_q(P \parallel Q) = \frac{1}{2} \sum_i p_i (-\ln p_i)^q + \frac{1}{2} \sum_i q_i (-\ln q_i)^q - \sum_i m_i (-\ln m_i)^q \tag{13}$$

$$JSD_\alpha(P \parallel Q) = \frac{1}{2} \sum_i p_i \left\{ -\frac{p_i^{-\alpha}}{\Gamma(\alpha + 1)} [\ln p_i + \psi(1) - \psi(1 - \alpha)] \right\} + \frac{1}{2} \sum_i q_i \left\{ -\frac{q_i^{-\alpha}}{\Gamma(\alpha + 1)} [\ln q_i + \psi(1) - \psi(1 - \alpha)] \right\} - \sum_i m_i \left\{ -\frac{m_i^{-\alpha}}{\Gamma(\alpha + 1)} [\ln m_i + \psi(1) - \psi(1 - \alpha)] \right\} \tag{14}$$

In order to illustrate the fractional-order distance we consider two examples, namely the set \mathcal{A} of $n = 13$ irrational numbers and the set \mathcal{B} of $n = 24$ Human chromosomes. Set \mathcal{A} consists of the numbers Pi ($\pi = 3.141\dots$), Nepper ($e = 2.718\dots$), Euler-Mascheroni ($\gamma = 0.577\dots$), Catalan ($G = 0.915\dots$), Hilbert or Gelfond-Schneider ($2^{\sqrt{2}} = 2.665\dots$), Khinchin ($K_0 = 2.685\dots$), Golden ratio ($\varphi = \frac{1+\sqrt{5}}{2} = 1.618\dots$), $\ln 2$, $\ln 3$, $\ln 5$, $\sqrt{2}$, $\sqrt{3}$ and $\sqrt{5}$ labelled in the sequel as {Pi, Nep, Eul, Cat, Hil, Khi, Gol, Ln2, Ln3, Ln5, St2, St3, St5}. Set \mathcal{B} consists of the whole set of Human chromosomes labelled in the sequel as {Hu1, ..., Hu22, HuX, HuY}. The irrational numbers are expanded up to 7000 digits and, for each one, groups of two digits feed 10^2 bins of histograms of relative frequency of occurrence. On the other hand, the chromosome bases are read in triplets feeding 4^3 bins of histograms of relative frequency of occurrence. In both cases, a comparison $n \times n$ symmetrical matrix \mathbf{D} of element to element relative distances is constructed, adopting the indices JSD_q and JSD_α . For simplifying comparisons all distances were converted to the interval between zero (minimum distance) and one (maximum distance). The results are visualized by means of Phylip [48,49] (plots using options “neighbor” and “drawtree”), a package of programs for inferring phylogenies. These algorithms produces a tree based on matrix \mathbf{D} , trying to accommodate the distances into the two dimensional space.

Figures 10 and 11 show the trees for sets \mathcal{A} and \mathcal{B} based of distances (13) and (14). We verify that not only the charts are qualitatively of the same type, but also that the generalization leads to results compatible with those produced by distinct methods [50–52] which confirms the goodness of the proposed concept.

Figure 10. Tree (Phylip with algorithm “neighbor” and visualization by “drawtree”) of the set \mathcal{A} of 13 irrational numbers, compared by means of the indices $I_q, q = 1$ (left) and $I_\alpha, \alpha = 0.5$ (right).

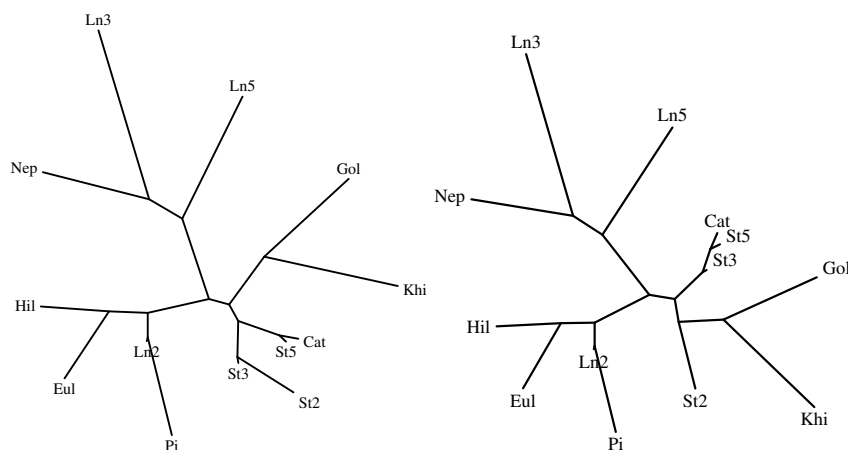
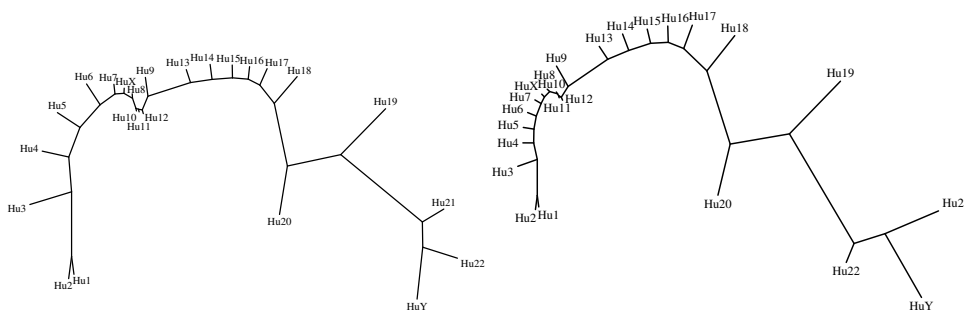


Figure 11. Tree (Phylip with algorithm “neighbor” and visualization by “drawtree”) of the set \mathcal{B} of 24 Human chromosomes compared by means of the indices $I_q, q = 1$ (left) and $I_\alpha, \alpha = 0.5$ (right).



5. Conclusions

This paper presented a generalization of the concept of entropy inspired in the properties of Fractional Calculus. The novel index follows the recent trend in expanding the scope of application of both mathematical tools, by relaxing some properties and allowing their application in new scientific areas. The generalized fractional entropy was first adopted with several typical probability distributions. In a second phase the index was also applied to several types of data, namely of mathematical, financial and biological nature. It was verified that the proposed entropy leads to an higher sensitivity to the signal evolution being useful in describing the dynamics of complex systems. Furthermore, the proposed generalization embeds the concept of positive and negative information, that is, with data either reliable or misleading, allowing the extension of entropy for deceptive cases. The new formulation is then

extended for measuring relative distances and tested with two distinct sets of data. The results reveal the goodness of the generalized fractional information concept.

Acknowledgments

The author thanks the anonymous reviewers for their constructive comments.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Plastino, A.; Plastino, A.R. Tsallis entropy and Jaynes' Information Theory formalism. *Braz. J. Phys.* **1999**, *29*, 50–60.
2. Li, X.; Essex, C.; Davison, M.; Hoffmann, K.H.; Schulzky, C. Fractional diffusion, irreversibility and entropy. *J. Non-Equilib. Thermodyn.* **2003**, *28*, 279–291.
3. Mathai, A.; Haubold, H. Pathway model, superstatistics, Tsallis statistics, and a generalized measure of entropy. *Physica A* **2007**, *375*, 110–122.
4. Anastasiadis, A. Special issue: Tsallis entropy. *Entropy* **2012**, *14*, 174–176.
5. Oldham, K.; Spanier, J. *The Fractional Calculus: Theory and Application of Differentiation and Integration to Arbitrary Order*; Academic Press: New York, NY, USA, 1974.
6. Samko, S.; Kilbas, A.; Marichev, O. *Fractional Integrals and Derivatives: Theory and Applications*; Gordon and Breach Science Publishers: Amsterdam, The Netherlands, 1993.
7. Miller, K.; Ross, B. *An Introduction to the Fractional Calculus and Fractional Differential Equations*; Wiley: New York, NY, USA, 1993.
8. Kilbas, A.; Srivastava, H.; Trujillo, J. *Theory and Applications of Fractional Differential Equations*; North-Holland Mathematics Studies; Elsevier: Amsterdam, The Netherlands, 2006; Volume 204.
9. Rényi, A. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; pp. 547–561.
10. Haubold, H.; Mathai, A.; Saxena, R. Boltzmann-Gibbs entropy versus Tsallis entropy: Recent contributions to resolving the argument of Einstein concerning “neither herr Boltzmann nor herr Planck has given a definition of W”? *Astrophys. Space Sci.* **2004**, *290*, 241–245.
11. Ben-Naim, A. *Statistical Thermodynamics Based on Information: A Farewell to Entropy*; World Scientific: Singapore, Singapore, 2008.
12. Podlubny, I. *Fractional Differential Equations, Volume 198: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution, Mathematics in Science and Engineering*; Academic Press: San Diego, CA, USA, 1998.
13. Hilfer, R. *Application of Fractional Calculus in Physics*; World Scientific: Singapore, Singapore, 2000.
14. Zaslavsky, G. *Hamiltonian Chaos and Fractional Dynamics*; Oxford University Press: Oxford, UK, 2005.

15. Tarasov, V. *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media*; Springer: New York, NY, USA, 2010.
16. Mainardi, F. *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*; Imperial College Press: London, UK, 2010.
17. Baleanu, D.; Diethelm, K.; Scalas, E.; Trujillo, J.J. *Fractional Calculus: Models and Numerical Methods*; Series on Complexity, Nonlinearity and Chaos; World Scientific Publishing Company: Singapore, Singapore, 2012.
18. Ionescu, C. *The Human Respiratory System: An Analysis of the Interplay between Anatomy, Structure, Breathing and Fractal Dynamics*; Series in BioEngineering; Springer: London, UK, 2013.
19. Machado, J.A.T. Entropy analysis of integer and fractional dynamical systems. *J. Appl. Nonlinear Dyn.* **2010**, *62*, 371–378.
20. Machado, J.A.T. Fractional dynamics of a system with particles subjected to impacts. *Commun. Nonlinear Sci. Numer. Simul.* **2011**, *16*, 4596–4601.
21. Machado, J.A.T. Entropy analysis of fractional derivatives and their approximation. *J. Appl. Nonlinear Dyn.* **2012**, *1*, 109–112.
22. Machado, J.T.; Galhano, A.M. Symbolic fractional dynamics. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2013**, *3*, 468–474.
23. Machado, J.T. Complex dynamics of financial indices. *Nonlinear Dyn.* **2013**, *74*, 287–296.
24. Machado, J.T. Relativistic time effects in financial dynamics. *Nonlinear Dyn.* **2014**, *75*, 735–744.
25. Machado, J.T.; Costa, A.; Quelhas, M. Entropy analysis of DNA code dynamics in human chromosomes. *Comput. Math. Appl.* **2011**, *62*, 1612–1617.
26. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
27. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656.
28. Gray, R.M. *Entropy and Information Theory*; Springer: New York, NY, USA, 1990.
29. Beck, C. Generalised information and entropy measures in physics. *Contemp. Phys.* **2009**, *50*, 495–510.
30. Khinchin, A.I. *Mathematical Foundations of Information Theory*; Dover: New York, NY, USA, 1957.
31. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
32. Ubriaco, M.R. Entropies based on fractional calculus. *Phys. Lett. A* **2009**, *373*, 2516–2519.
33. Machado, J.A.T.; Galhano, A.M.S. Approximating fractional derivatives in the perspective of system control. *Nonlinear Dyn.* **2009**, *56*, 401–407.
34. Machado, J.A.T.; Galhano, A.M.; Oliveira, A.M.; Tar, J.K. Approximating fractional derivatives through the generalized mean. *Commun. Nonlinear Sci. Numer. Simul.* **2009**, *14*, 3723–3730.
35. Valério, D.; Trujillo, J.J.; Rivero, M.; Machado, J.T.; Baleanu, D. Fractional calculus: A survey of useful formulas. *Eur. Phys. J. Spec. Top.* **2013**, *222*, 1827–1846.
36. Dirac, P.A.M. Bakerian Lecture. The physical interpretation of quantum mechanics. *Proc. R. Soc. Lond.* **1942**, *180*, 1–40.

37. Feynman, R.P. The concept of probability theory in quantum mechanics. In Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability Theory, Berkeley, CA, USA, 31 July–12 August 1950; University of California Press: Berkeley, CA, USA, 1950.
38. Feynman, R.P. Negative probability. In *Quantum Implications: Essays in Honour of David Bohm*; Basil, J., Hiley, F.D.P., Eds.; Routledge & Kegan Paul Ltd.: London, UK and New York, NY, USA, 1987; pp. 235–248.
39. Bartlett, M.S. Negative probability. *Math. Proc. Camb. Philos. Soc.* **1945**, *41*, 71–73.
40. Székely, G.J. Half of a Coin: Negative Probabilities. *Wilmott Magazine* **2005**, pp. 66–68.
41. Machado, J.T. Fractional coins and fractional derivatives. *Abstr. Appl. Anal.* **2013**, *2013*, doi:10.1155/2013/205097.
42. Shim, J.; Arkin, R.C. A taxonomy of robot deception and its benefits in HRI. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 13–16 October 2013; pp. 2328–2335.
43. Sibson, R. Information radius. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **1969**, *14*, 149–160.
44. Taneja, I.; Pardo, L.; Morales, D.; Ménendez, L. Generalized information measures and their applications: A brief survey. *Qüestió* **1989**, *13*, 47–73.
45. Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145–151.
46. Cha, S.H. Measures between probability density functions. *Int. J. Math. Models Methods Appl. Sci.* **2007**, *1*, 300–307.
47. Deza, M.M.; Deza, E. *Encyclopedia of Distances*; Springer: Berlin/Heidelberg, Germany, 2009.
48. PHYLIP. Available online: <http://evolution.genetics.washington.edu/phylip.html> (accessed on 3 April 2014).
49. Tuimala, J. *A Primer to Phylogenetic Analysis Using the PHYLIP Package*; CSC—Scientific Computing Ltd.: Espoo, Finland, 2006.
50. Costa, A.; Machado, J.T.; Quelhas, M. Histogram-based DNA analysis for the visualization of chromosome, genome and species information. *Bioinformatics* **2011**, *27*, 1207–1214.
51. Machado, J.T.; Costa, A.C.; Quelhas, M.D. Shannon, Rényi and Tsallis entropy analysis of DNA using phase plane. *Nonlinear Anal. Ser. B: Real World Appl.* **2011**, *12*, 3135–3144.
52. Machado, J.T.; Costa, A.C.; Quelhas, M.D. Wavelet analysis of human DNA. *Genomics* **2011**, *98*, 155–163.