POLITECNICO DI TORINO
Repository ISTITUZIONALE

Mutual Information and Nonadditive Entropies: The Case of Tsallis Entropy

Original
Mutual Information and Nonadditive Entropies: The Case of Tsallis Entropy / Sparavigna, A.C.. - In: INTERNATIONAL JOURNAL OF SCIENCES. - ISSN 2305-3925. - 4:10(2015), pp. 1-4. [10.18483/ijSci.845]

Availability:
This version is available at: 11583/2620811 since: 2015-11-30T16:23:32Z

Publisher:
Alkhaer

Published
DOI:10.18483/ijSci.845

Terms of use:

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

Publisher copyright

(Article begins on next page)

07 November 2026

International Journal of Sciences

Research Article Volume 4 – October 2015 (10)

Mutual Information and Nonadditive Entropies: The Case of Tsallis Entropy

Amelia Carolina Sparavigna[1]

[1]Department of Applied Science and Technology, Politecnico di Torino, Torino, Italy

Abstract: Mutual information of two random variables can be easily obtained from their Shannon entropies. However, when nonadditive entropies are involved, the calculus of the mutual information is more complex. Here we discuss the basic matter about information from Shannon entropy. Then we analyse the case of the generalized nonadditive Tsallis entropy.

Keywords: Mutual Information, Entropy, Tsallis Entropy, Generalized Additivity, Image Registration

1. Introduction

In many applications of engineering and telecommunication, it is often desired to increase or decrease the dependency of two random variables. This dependency is linked to the mutual information, which is its measure. The mutual information can be easily decomposed into a sum (with signs) of entropies [1], when the Shannon entropy is used. This entropy is additive, that is, when we have independent subsystems, X and Y, the entropy of their union is S(X,Y)=S(X)+S(Y). When nonadditive entropies are involved, the approach to find the mutual information is not so simple. Moreover, it is often preferred to use the term "mutual entropy", instead of "mutual information" [2,3].

The mutual entropy contains conditional entropies, which must be carefully defined when entropies are nonadditive [3]. Among the nonadditive entropies, we have the Tsallis entropy. It is a generalization of the standard Boltzmann–Gibbs entropy, introduced in 1988 as a basis for generalizing the standard statistical mechanics [4,5]. Due to its entropic index, which can be used as tuning parameter, this entropy is involved in several applications, in particular for image processing and image registration [6]. Here we discuss the basic matter concerning the mutual information when Tsallis entropy is involved.

2. Mutual information

The mutual information of two random variables in subsystems X,Y is providing a measure of the mutual dependence of the variables. This can be viewed intuitively as a measure of information that X and Y are sharing. In particular, it measures how much knowing one of these variables reduces uncertainty about the other [7-9]. In [9], examples are explaining this quantity. If X and Y are independent, knowing X does not give any information about Y and vice versa: the mutual information is zero. If Y=f(X) or X=f(Y), where f indicates a deterministic function, all information conveyed by X (or Y) is shared with Y (or X): the mutual information is the same as the uncertainty contained in Y (or X) alone, which is measured by the entropy of Y (or X). The mutual information is then the entropy of Y (or X) [9].

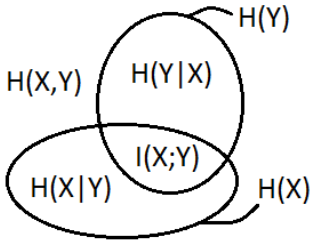

Figure 1: Venn diagram of mutual information I(X;Y) associated with correlated variables X and Y. The area contained by both circles is the joint entropy H(X,Y). The lower circle is the individual entropy H(X), such as the upper circle is that of H(Y).

The physical meaning of the mutual information I(X;Y) as "the reduction of the uncertainty of X due to knowledge of Y" (or vice versa) [9], can be depicted in a Venn diagram (Figure 1). In this

This article is published under the terms of the Creative Commons Attribution License 4.0
Author(s) retain the copyright of this article. Publication rights with Alkhaer Publications.
Published at: http://www.ijsciences.com/pub/issue/2015-10/
DOI: 10.18483/ijSci.845; Online ISSN: 2305-3925; Print ISSN: 2410-4477

Amelia Carolina Sparavigna (Correspondence)
d002040@polito.it
+39-011-090-7360

diagram, the single-variable entropies H(X), H(Y) are represented by two overlapping sets, whereas the two-variable entropy is represented by the union of these sets and the mutual information common to X and Y is represented by their intersection. Note that H(X)=I(X;X), so entropy is the "self-information". Also note that conditioning of entropies in Venn diagrams is indicated by set subtraction, so that, for example, the set representing H(X|Y) results from subtracting the set representing H(Y) from the set representing H(X).
The mutual information is given by [10]: I(X;Y)=H(X)+H(Y)−H(X,Y), with the following properties, I(X;Y)=I(Y;X) and I(X;X)=H(X). Note that H(Y|H) is the conditional entropy [11]. Let us assume the joint entropy H(X,Y) for the combined system determined by two random variables X and Y. We need H(X,Y) "bits of information" to describe its exact state [12]. If we first learn the value of X, we have gained H(X) bits of information. "Once X is known, we only need H(X,Y)−H(X) bits to describe the state of the whole system" [12]. This quantity is exactly H(Y|X), which gives the chain rule of conditional entropy: H(Y|X)=H(X,Y)−H(X).
If we are using the Shannon entropy S:

$$\begin{aligned} I(X;Y) &= S(X) - S(X \mid Y) \\ &= S(Y) - S(Y \mid X) \\ &= S(X) + S(Y) - S(X,Y) \quad (1) \end{aligned}$$

If X,Y are independent, we have that S(X,Y)=S(X)+S(Y), and therefore I(X;Y)=0.

3. Using Tsallis entropy

In fact, besides the Shannon entropy we have generalized entropies too, and, among them, the nonadditive entropies. How can we generalize the mutual information in this case? In [3], it is preferred the use of the so-called Tsallis mutual entropy, defined as:

$$\begin{aligned} MT(X;Y) &= T(X) - T(X \mid Y) \\ &= T(Y) - T(Y \mid X) \quad (2) \end{aligned}$$

In (2), T is referring to the Tsallis entropy. According to [4], T(X,Y)=T(X)+T(Y|X) and T(Y,X)=T(Y)+T(X|Y). Let us remember that Tsallis entropy T and Rényi entropy R, [13], are linked by the following equation:

$$R = \frac{\ln[1+(1-q)T]}{1-q} \quad (3)$$

Here q is the entropic index. As q approaches 1, the Tsallis entropy becomes the Shannon entropy.
Let us try defining the Tsallis mutual entropy as:

$$MT^{*}(X;Y) = T(X) + T(Y) - T(X,Y) \quad (4)$$

If X,Y are independent, we must have a mutual information equal to zero. However, from (4), we find:

$$MT^{*}(X;Y) = (q-1)T(X)T(Y) \quad (5)$$

This happens because the generalized additivity for independent subsystems is:

$$T(X,Y) = T(X) + T(Y) + (1-q)T(X)T(Y). \quad (6)$$

As a result, for an entropic index different from 1, (4) should give a result different from zero. Therefore, (4) is not good for representing the mutual information I(X;Y).
In his paper, Tsallis is discussing the problem of correlated systems too [4]. He used the Rényi entropy for correlated systems:

$$\Gamma = R(X) + R(Y) - R(X,Y) \quad (7)$$

Since the Rényi entropy is additive, if X,Y are independent, Γ is equal to zero. Let us note that it is function Γ which seems working as the mutual information.
However, for the non-additive Tsallis conditional entropy, a quite simple formula was given in [14]:

$$T(X \mid Y) = \frac{T(X,Y) - T(Y)}{1+(1-q)T(Y)} \quad (8a)$$

$$T(Y \mid X) = \frac{T(X,Y) - T(X)}{1+(1-q)T(X)} \quad (8b)$$

We could define the mutual entropy, as for the Shannon entropy, in the following manner:

$$\begin{aligned} MT^{**}(X;Y) &= T(X) - T(X \mid Y) \\ &= T(X) - \frac{T(X,Y) - T(Y)}{1+(1-q)T(Y)} \quad (9a) \end{aligned}$$

$$\begin{aligned} MT^{**}(Y;X) &= T(Y) - T(Y \mid X) \\ &= T(Y) - \frac{T(X,Y) - T(X)}{1+(1-q)T(X)} \quad (9b) \end{aligned}$$

For X,Y independent variables, using (9*a*) for instance:

$$MT^{**}(X;Y) = \frac{T(X) + F_{X,Y} - T(X) - T(Y) - F_{X,Y} + T(Y)}{1+(1-q)T(Y)} = 0$$

Where $F_{X,Y} = (1-q)T(X)T(Y)$. The same for (9*b*).

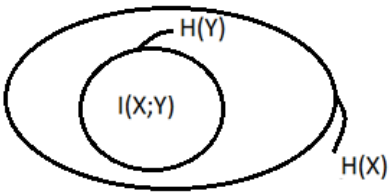

Figure 2: Venn diagram when Y is completely dependent on X.

4. The problem of symmetry

Let us note that I(X;Y) must be symmetric, that is I(X;Y)=I(Y;X), "otherwise it would not be mutual information" [15]. Are (9*a*) and (9*b*) giving us the same mutual information? Let us consider the case when Y is completely dependent on X, as depicted in the Figure 2.

In this case H(X,Y)=H(X), so we have T(X,Y)=T(X), and the mutual information is T(Y). In the following discussion, we have X larger that Y. Then T(X) lager than T(Y).

Let us consider (9*a*) and (9*b*):

$$MT^{*}(X;Y) = T(Y)\frac{1+(1-q)T(X)}{1+(1-q)T(Y)} \qquad (10a)$$

$$MT^{*}(Y;X) = T(Y) \qquad (10b)$$

Therefore, (9*a*)-(9*b*) do not fulfil the required symmetry. In fact, if we consider the Figure 2, it is clear that a non-symmetric situation exists,

Aiming to solve the case of Figure 2, we could modify these mutual Tsallis entropies so that:

$$MT(X;Y) = (1+(1-q)T(Y))\left[T(X) - \frac{T(X,Y) - T(Y)}{1+(1-q)T(Y)}\right]\frac{1}{1+(1-q)T(X)} \qquad (11a)$$

$$MT(Y;X) = (1+(1-q)T(X))\left[T(Y) - \frac{T(X,Y) - T(X)}{1+(1-q)T(X)}\right]\frac{1}{1+(1-q)T(X)} \qquad (11b)$$

We can see that:

$$MT(Y;X) = \frac{T(X) + T(Y) - T(X,Y) + (1-q)T(X)T(Y)}{1+(1-q)T(X)} \qquad (12)$$

So we have a mutual information, which is symmetric. If X and Y are independent: MT(X;Y)=0.

For completely dependent variable and T(X,Y)=T(X):

$$MT(X;Y) = MT(Y;X) = \frac{T(Y) + (1-q)T(X)T(Y)}{1+(1-q)T(X)} = T(Y) \qquad (13)$$

In the case X is coincident with Y, we have T(X)=T(Y):

$$MT(X;Y) = \frac{T(X) + (1-q)T(X)T(Y)}{1+(1-q)T(X)} = T(Y) = T(X) \qquad (14)$$

However, we can have also the case when X is totally inside Y (Figure 3).

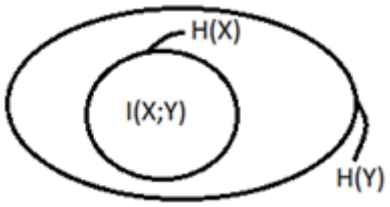

Figure 3: Venn diagram when Y is completely dependent on X.

Then, let us write the join entropy as T(Y,X), to remark this situation. We have that the mutual information must be equal to T(X). We obtain this in a symmetric entropy, when we define the mutual entropy as:

$$MT(X;Y) = MT(Y;X) = \frac{T(X) + T(Y) - T(Y,X) + (1-q)T(X)T(Y)}{1 + (1-q)T(Y)} \quad (15)$$

Then, when T(Y,X)=T(Y), from (15) we obtain MT(X;Y)=T(X). We can define, in the same manner as in Ref.4 is proposed the correlation term in the mutual Tsallis entropy:

$$MT(X;Y) = MT(Y;X) = \frac{T(X) + T(Y) - T(Y,X) + (1-q)T(X)T(Y)}{1 + (1-q)\left[\max\{T(X), T(Y)\}\right]} \quad (16)$$

This can be an expression of the mutual information for Tsallis entropy, which is properly answering to the fundamental requirements of computation. We have discussed such an approach, because it can be easily applied to another nonadditive entropy, the Kaniadakis entropy, to determine its mutual information [16]. This problem will be addressed in a following paper.

References

1) J.W. Kay, D. M. Titterington (1999). Statistics and Neural Networks: Advances at the Interface, Oxford University Press. ISBN-13: 978-0198524229, ISBN-10: 0198524226
2) R. W. Hamming (1980). Coding and Information Theory, Prentice-Hall. ISBN-13: 978-0521196819, ISBN-10: 0521196817
3) S. Furuichi (2006). Information Theoretical Properties of Tsallis Entropies, J. Math. Phys. 47:023302. DOI:10.1063/1.2165744
4) C. Tsallis (1988). Possible Generalization of Boltzmann-Gibbs Statistics, Journal of Statistical Physics 52:479. DOI:10.1007/BF01016429
5) A.C. Sparavigna (2015). Shannon, Tsallis and Kaniadakis Entropies in Bi-Level Image Thresholding, Int. J. Sci. 4(2):35. DOI: 10.18483/ijSci.626
6) S. Martin, G. Morison, W. Nailon, T. Durrani (2004). Fast and Accurate Image Registration Using Tsallis Entropy and Simultaneous Perturbation Stochastic Approximation, Electronics Letters, 40(10):595. DOI: 10.1049/el:20040375
7) Yue Deng (2014). High-Dimensional and Low-Quality Visual Information Processing, Springer. ISBN: 978-3-662-44526-6
8) V. Christophides, V. Efthymiou, K, Stefanidis (2015). Entity Resolution in the Web of Data, Morgan & Claypool Publishers. DOI: 10.2200/S00655ED1V01Y201507WBE013
9) Vv. Aa., Mutual Information, Wikipedia, Retrieved 3 October 2015.
10) G.J. Pottie, W.J. Kaiser (2005). Principles of Embedded Networked Systems Design, Cambridge University Press. Online ISBN: 9780511541049, DOI: 10.1017/CBO9780511541049
11) T.M. Cover, J.A. Thomas (1991). Elements of Information Theory, John Wiley & Sons, Inc. Print ISBN 0-471-06259-6, Online ISBN 0-471-20061-1
12) Vv. Aa., Conditional Entropy, Wikipedia, Retrieved 3 October 2015.
13) A. Rényi (1960). On Measures of Information and Entropy, Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability, pp. 547–561.
14) S. Abe, A.K. Rajagopal (2000). Nonadditive Conditional Entropy and its Significance for Local Realism, arXiv:quant-ph/0001085, 24 Jan 2000.
15) J.P.W. Pluim, J.B. Antoine Maintz, M.A. Viergever (2003). Mutual-Information-Based Registration of Medical Images: A Survey, IEEE Transactions on Medical Imaging, 22(8):986. DOI: 10.1109/tmi.2003.815867
16) A.C. Sparavigna (2015). Conditional Kaniadakis Entropy: a Preliminary Discussion. PHILICA.COM Article number 524.