Metric, Topological, and Graph Entropies

Angel Garrido

Facultad de Ciencias de la UNED

Abstract

Our paper analyzes here some new lines to introduce the evolving concept of an important Uncertainty Measure, the so-called Entropy. We need to obtain these new ways to model adequate conditions, departing from vague pieces of information. For this, it will be very necessary to analyze certain type of such measures, with very interesting applications, as Graph Entropy, Metric Entropy and Topological entropy.

Keywords: Measure Theory, Fuzzy Measures, Entropy.

Mathematics Subject Classification: 68R10, 68R05, 05C78, 78M35.

1. Introduction

The study of concepts such as Entropy and Symmetry is very important in the current science, not only on Physics, but also on Information Theory and Mathematical Sciences, in general. This paper attempt to show this fact.

For instance, it will be very suprising the way open by the Romanian mathematician *Nicolae Georgescu-Roegen* [15], a very inspirate and unortodox disciple of Karl Pearson and Joseph Schumpeter, suggesting the application of the 2nd Law of Thermodynamics to Economics. Later developed into Evolutionary Economics. This and subsequent developments give rise to currently very essential fields as Bioeconomics, or Ecological Entropy. Also to many important studies, as on Equilibrium Theory, and so on.

Also may be a very useful tool on Biocomputing, for instance, or in many others, as studying Environmental Sciences.

2. Metric Entropy

We may consider the *Metric Entropy,* also called *Kolmogorov Entropy,* or *Kolmogorov-Sinai Entropy,* in acronym *K-S Entropy.* Its name is given in honour of the great Russian mathematician Andrei N. Kolmogorov and its disciple, Yakov Sinai [22-24].

In a dynamical system, the metric entropy is equal to zero, for nonchaotic motion, and is strictly greater than zero, for chaotic motion. So, it is a simple indicator of the complexity of a dynamical system.

[1] AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

To define the *K-S Entropy,* we need to divide the phase space into n-dimensional hypercubes of content ε^n.

Let $P_{i_0,i_1,\dots,i_n}$ the probability that is in the hypercube expressed by

$$i_j \ at\ T = jT,\ for\ all\ j \in \{0,1,\dots,n\}$$

Then, we introduce

$$K_n = -\sum_{i_0,i_1,\dots,i_n} P_{i_0,i_1,\dots,i_n} \ln P_{i_0,i_1,\dots,i_n}$$

as the information needed to predict which hypercube the trajectory will be in at $(n+1)T$, given trajectories up to nT.

The *K-S entropy* is then defined by

$$K = \lim_{T\to 0}\left[\lim_{\varepsilon\to 0}\left[\lim_{n\to +\infty}\sum_{j=0}^{n-1}(k_{j+1}-k_j)\right]\right]$$

And now, from a measure theoretical view of point.

Let $(X,\Omega,\ \mu)$ be a probability space, or in a more general way, a fuzzy measurable space.

Recall that a *measurable partition* of X is such that each of their elements is a measurable set; therefore, an element of the fuzzy $\sigma-$algebra, Ω. And let I^X be the set of mappings from X to the closed unit interval, $[0,1]$.

A fuzzy $\sigma-$algebra, Σ, on a nonempty set, X, is a subfamily of I^X satisfying that

(1) $\mathbf{1} \in \Sigma$.

(2) If $\alpha \in \Sigma,$ then $1-\alpha \in \Sigma$.

(3) If $\{\alpha_i\}$ is a sequence in Σ, then $\vee_{i=1}^{\infty}\alpha_i = \sup \alpha_i \in \Sigma$.

A *fuzzy probability measure,* on a fuzzy $\sigma-$algebra, Σ, is a function

$$m : \Sigma \to [0,1]$$

which holds

[1] $m(\mathbf{1}) = 1$

[2] for all $\alpha \in \Sigma,\ m(1-\alpha) = 1 - m(\alpha)$

[3] for all $\alpha,\ \beta \in \Sigma,\ m(\alpha \vee \beta) + m(\alpha \wedge \beta) = m(\alpha) + m(\beta)$

[4] If $\{\alpha_i\}$ is a sequence in Σ, such that $\alpha_i \uparrow \alpha$, with $\alpha \in \Sigma$, then

$$m(\alpha) = \sup\ m(\alpha_i)$$

We call $(X, \Omega,\ \mu)$ a *fuzzy-probability measure space,* and the elements of Ω are called *measurable fuzzy sets.*

The notion of "fuzzy partition" was introduced by E. Ruspini.

Given a finite measurable partition, $\wp$, we can define its *Entropy* by

$$H_\mu(\wp) = \sum_{p\in\wp} -\mu(p)\log\mu(p)$$

As usually in these cases, we take as convention that 0 log 0 = 0.

Let $T : X \to X$ be a measure-preserving transformation. Then, the *Entropy of T* w.r.t. a finite measurable partition, $\wp$, is expressed as

$$h_\mu(T,\ \wp) = \lim_{n\to\infty} H_\mu\left(\vee_{k=0}^{n-1} T^{-k}\wp\right)$$

with H_μ the entropy of a partition, and $\vee$ denotes here the join of partitions.
Such limit always exists. So,we may define the *Entropy of T* by

$$h_\mu(T) = \sup_\wp h_\mu(T,\ \wp)$$

Taking the supremum over all finite measurable partitions.

Many times $h_\mu(T)$ is named the *Metric Entropy of T.* So,we may to differentiate this mathematical object from the Topological Entropy.

We may to see the mutual relationship between Metric Entropy and the Covering Numbers.

Let (X, d) a metric space, and let $Y \subseteq X$ a subset of X.

We says that $Y^* \subseteq X$ is an $\epsilon - cover\ of\ Y$, if for each $y \in Y$, there exists a $y^* \in Y^*$ such that

$$d(y,\ y^*) \leq \varepsilon$$

There are many different covers of Y.

But we are specially interested here in one which contains the lesser number of elements. We call [2] the cardinal, or size, of such a cover its *Covering Number.*

Mathematically expressed, the $\epsilon - covering\ number\ of\ Y$ is

$$N(\varepsilon, Y, d) = \min\{card(Y^*) :\ Y^*\ is\ an\ \epsilon - cover\}$$

A *proper cover* is one where $Y^* \subseteq Y$. And a *proper covering number* is defined in terms of the cardinality of the minimum proper cover.

Both, covering numbers and proper covering numbers are related by

$$N\left(\epsilon, Y\right) \leq N_{proper}\left(\epsilon, Y\right) \leq N\left(\frac{\varepsilon}{2}, Y\right)$$

Furthermore, we recall that the Metric Entropy, $H\left(\epsilon, Y\right)$, is a natural representation of the cardinal of the set of bits needed to send, in order to identify an element of the set up to precision ε.

It will be expressed by

$$H\left(\epsilon, Y\right) = \log N\left(\epsilon, Y\right)$$

3. Topological Entropy

Let $(X,\ d)$ be a compact metric space, and let $f : X \rightarrow\ X$ be a continuous map.

For each $n > 0$, we define a new metric, d_n, by

$$d_n(x, y) = max\{d(f^i(x), f^i(y)) : 0 \leq i < n\}$$

Two points, x and y, are *close* with respect to (w. r. t.) this metric, if their first n iterates (given by f^i, $i=1,2,\ldots$) are close.

For $\epsilon > 0$, and $n \in N^*$, we say that $S \subset X$ is an $(n,\epsilon)-$ *separated set*, if for each pair, $x,\ y$, of points of S, we have

$$d_n(x, y) > \epsilon$$

Denote by $N(n, \epsilon)$ the *maximum cardinality of a (n, ϵ)-separated set.*

It must be finite, because X is compact. In general, this limit may exists, but it could be infinite.

A possible interpretation of this number [3] is as a measure the average exponential growth of the number of distinguishable orbit segments. So, we could say that *the higher the topological entropy is, the more essentially different orbits we have.*

The topological entropy is a continuous, monotonically increasing function.

This concept was introduced, in 1965, by Adler, Konheim and McAndrew [1].

$N(n, \epsilon)$ shows the number of "distinguishable" orbit segments of length n, assuming we cannot distinguish points that are less than ϵ apart.

The *topological entropy of f* is then defined by

$$H_{top} = \lim_{\epsilon \rightarrow 0} \lim\ \sup_{n \rightarrow \infty}\left[\frac{1}{n}\log N\left(n, \epsilon\right)\right]$$

We can see now tsome basic *properties of topological entropy*

$[I]$ $h\,(T) \geq h\,(S)\,$, if (Y,S) is a topological factor of (X,T).

I. e., if $\phi : X \rightarrow Y$ is any continuous and surjective application.

$[II]$ If (X,T) and (Y,S) are topologically conjugate, then $h(T) = h(S)$.

Recall that T and S are topologically conjugate when $T = \phi S \phi^{-1}$, being ϕ a homeomorphism.

$[III]$ $h(T^n) = n\ h\,(T)\,,$ if $n \geq 0$.

$[IV]$ $h(T) = h(T^{-1})$, if T is a homeomorphism.

$[V]$ $h\,(T \times S) = h(T) + h(S)$.

$[VI]$ $h\,(T \circ S) = h\,(S \circ T)$

4. Graph Entropy

The entropy of a system represent [21] the amount of uncertainty one observer has about the state of the system. The simplest example of a system will be a random variable, which can be shown by a node into the graph, being their edges the representation of the mutual relationship between them. Information measures the amount of correlation between two systems, and it reduces to a mere difference in entropies.

We consider a functional on a graph, $G = (V, E)$, with P a probability distribution on its node (or vertex) set, V. These mathematical construct will be denoted by GE.

It is a concept introduced by Körner, as solution of a coding problem formulated in Information Theory. Because its sub-additivity, has become a useful tool in proving some lower bounds results in Computational Complexity Theory.

The search for exact additivity has produced certain interesting combinatorial structures. One of such results is the characterization of perfect graphs by the additivity of GE.

It will be defined as

$$H(G,P) = min \sum p_i \ \log \ p_i$$

Observe that such function is convex. It tends to $+\infty$ on the boundary of the non-negative orthant of R^n , and monotonically to $-\infty$ along rays from the origin.

So, such minimum is always achieved and it will be finite.

The properties of Graph Entropy [21] may be very essential in aplications.

Let V be their set of nodes, or vertices, and E their set of edges, or links.

These properties may be

- *Monotonicity:* If F and G are two graphs, with

$$V(F) = V(G)$$
$$and$$
$$E(F) \subseteq E(G)$$

Then, it holds

$$H(F, P) \leq H(G, P)$$

- *Subadditivity:* Let F and G be as above. And let $F \cup G$ the graph with node set

$$E(F) \cup E(G)$$

Then,for any fixed probability distribution, P, we have

$$H(F \cup G) \leq H(F, P) + H(G, P)$$

- *Additivity of Substitution:* Let F and G be two node disjoint graphs, and let v be a node of G. By *substituting F for v*, we said deleting v and joining everynode of F to those nodes of G whichhave been adjacent with v. The resulting graph is denoted by

$$G_{v \leftarrow F}$$

It will be possible to extend such concept to probability distributions.
So,

$$P_{v \leftarrow Q}(u) = \; P(u), \; if \; u \in V(G) - \{v\}; \; and \; equal \; to \; P(v)Q(u), \; if \; u \in V(F)$$

The *Substitution Lemma* says that being F and G two node disjoint graphs, $v \in V(G)$, with P and Q two probability distributions on the respective set of nodes, then

$$H(G_{v \leftarrow F}, P_{v \leftarrow Q}) = H(G, P) + P(v) \; H(F, P)$$

As example, we refer to the entropy of some special graphs,

i) the entropy of the empty graph is always null

$$H(G_{\varnothing}) = 0$$

ii) the entropy of K_n, the complete graph on n nodes is given by

$$H(K_n, P) = H(P)$$

5. Conclusions

Statistical entropy is a probabilistic measure of uncertainty [19], or ignorance about data. Whereas *Information* is a measure of a reduction in that uncertainty [20].

The Entropy of a probability distribution is just the expected value of the information of such distribution [19].

All these improved tools must permits to advance not only in fields as Optimization Theory, but also on Generalized Fuzzy Measures [7], Economics [15], modeling in Biology, and so on.

References

[1] Adler, Konheim and Mac Andrew: "Topological Entropy". *Transactions of the AMS.* Vol. 114, pp. 309-319, 1965.

[2] Bartlett: "Covering Numbers and Metric Entropy". *Statistical Learning Theory.* Lecture, Spring 2006.

[3] Burbea, and Rao: Entropy Differential Metric, Distance and Divergence Measures in Probabilistic Spaces: A Unified Approach. *J. Multi. Analysis,* Vol. 12, pp. 575-596, 1982.

[4] Burillo, and Bustince: "Estructuras algebraicas en conjuntos IFS". Proc. II Congreso Nacional de Lógica y Tecnología Fuzzy, pp. 134-147. Boadilla, 1992.

[5] Ibid.: "Entropy on Intuitionistic Fuzzy Sets and Interval-Valued Fuzzy Sets". *FSS,* Vol. 78, pp. 305-316, 1996.

[6] Csiszár: Information Measures: A Critical Survey. *Trans. Of the 7th Prague Conference,* pp. 83-86, 1974.

[7] Dubois and Prade: *Fundamentals of Fuzzy Sets.* Series: The Handbooks of Fuzzy Sets. Vol.7. Springer-Kluwer, 2000.

[8] Dumitrescu, D.: Entropy of a fuzzy dynamical system, FSS, Vol. 70, pp. 45-57, 1995.

[9] Ibid.: Entropy of a fuzzy process. FSS. Vol. 55, pp. 169-177, 1993.

[10] Ibid.: Fuzzy measures and the entropy of fuzzy partitions. J. Math. Anal. Appl. Vol. 175, pp. 359-373, 1993.

[11] Garmendia: "The Evolution of the Concept of Fuzzy Measure". *Studies in Computational Intelligence,* Vol. 5, pp. 186-200, 2005.

[12] Garrido: "Additivity and Monotonicity in Fuzzy Measures". Plenary Talk in ICMI45, at Bacau University, then published at *SCSSM journal,* Vol. 16, pp. 445-458, 2006.

[13] Garrido: "Information and Entropy Measures". Plenary Talk in the 6th ICTAMI, at Alba Iulia University, accepted; must to be published at *AUA journal,* Vol. 19 (special issue), pp. 18, 2009.

[14] Garrido: "Analyzing the concept of Entropy". *EIJ-AMO* journal, Vol. 11, Issue No. 4, pp.525-530, 2009.

[15] Georgescu-Roegen: *The Entropy Law and the Economic Process.* Harvard University Press. Cambridge, MA,1971.

[16] Hung and Yang: "Fuzzy entropy on Intuitionistic Fuzzy Sets". *Intern. J. of Intelligent Systems.* Vol. 21, pp. 443-451, 2006.

[17] Liu:"Entropy, distance and similarity measures of fuzzy sets and their relations. *FSS,* Vol. 52, pp. 305-318, 1992.

[18] Luca and Termini: "A definition of non-probabilistic entropy, in the setting of fuzzy theory. Information and Control. Vol. 20, pp. 301-318, 1972.

[19] Rényi: "On measures of information and entropy". *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability,* pp. 547-561, 1961.

[20] Shannon: "A Mathematical Theory of Communication", Bell System Technical Journal, pp. 379-423 and 623-656, 1948. Later, it appears as book, at Illinois Press, 1963. Also by E-print.

[21] Simonyi: "Graph Entropy: A Survey". DIMACS Series in Discrete Mathematics and Theoretical Computer Science.

[22] Sinai: "On the concept of Entropy of a Dynamical System". *Doklady Akademii Nauk SSSR,* Vol. 124, pp. 768-771, 1959.

[23] Ibid.: "On the foundations of the Ergodic Hypothesis for a Dynamical System of Statistical Mechanics". *Doklady Akademii Nauk SSSR,* Vol. 153, pp. 1261-1264, 1963.

[24] Ibid.:"Gibbs measures in Ergodic Theory". *Russian Math. Surveys,* Vol. 27, pp. 21-69, 1972.

[22] Titchener, Nicolescu, Staiger, Gulliver, and Speidel: "Deterministic Complexity and Entropy". *Fundamenta Informaticae,* Vol. 64, Issue 1-4, Contagious Creativity-In Honor of the 80th Birthday of Prof. Solomon Marcus, 2004.

[23] Titchener: "A measure of information". *Proc. Data Compression Conference,* pp. 353-362, Snowbird, UT, 2000.

[24] Ibid.: "A deterministic theory of Complexity, Information and Entropy". *IEEE Information Theory Workshop.* San Diego, CA, 1998.

[25] Volkenstein: *Entropy and Information.* Series: Progress in Mathematical Physics, Vol. 57, Birkhäuser Verlag, 2009.

[26] Wang, and Klir: *Generalized Measure Theory.* Springer Verlag, Berlin-New York, 2008.

[27] Yager: "On the measure of fuzziness and negation". *Int. J. Gral.Systems,* Vol.5, pp. 189-200, 1979.

[28] You, and Gao: "Maximum entropy membership functions for discrete fuzzy variables". *Information Sciences,* 179, Issue No. 14, pp. 2353-2361, 2009.

[29] Yu, and Hu: "Entropies of fuzzy indiscernibility relation and its operations". *Intern. J. of Uncertainty, Fuzziness and Knowledge-Based Systems.* Vol. 12, Issue No. 5, pp. 575-589, 2005.

[30] Zeng and Li: "Relations between Similarity and Entropy of Interval Valued Fuzzy Sets". *FSS,* Vol.157, pp.1477-1484, 2006.