Fractals -10

**Overview of entropy measures**

**Shannon entropy**

A statistical concept of entropy was introduced by Shannon 1948 [1] while working in Bell Laboratories on the problem of how to transmit information most efficiently through a given channel.

The entropy of a random variable is a measure of the uncertainty associated with that random variable; it is a measure of the amount of information required on the average to describe the random variable. The entropy of a discrete random variable  taking value from a finite set  and probability mass function (pmf)is defined by



The log in the above equation is taken to be to the base 2, and entropy is expressed in *bits*. If the base of logarithm is , the entropy is expressed in *nats*. By convention  since  as . The entropy is a functional of the distribution of . It does not depend on the values of , but only on the probabilities. It is seen from the definition that has following properties [2]:

1. 
2. 

**Joint Entropy**

The joint entropy of a pair of discrete random variables with a joint distribution  is defined as



The joint entropy measures how much uncertainty there is in the two random variables  and  taken together [2].

**Conditional Entropy**

The conditional entropy of given is defined as



We can write this as



Conditional entropy is a measure of how much uncertainty remains about the random variable  when we know the value of  [2]. Conditional entropy is asymmetric , .

Joint entropy and conditional entropy are related trough the chain rule that states that the total uncertainty about the value of  and  is equal to the uncertainty about  plus the (average) uncertainty about  once you know .



**Relative Entropy**

The relative entropy is a measure of the statistical distance between two different probability distributions. The relative entropy or *Kullback Leibler distance* between two probability mass functions  and is defined as



It is a measure of the inefficiency of assuming that the distribution of random variable is  when the true distribution is  . The relative entropy is always non-negative and it is zero if and only if the two distributions are the same [2].

**Mutual Information**

Mutual information is a measure of information that one random variable contains about other random variable. It represents the reduction of uncertainty of one random variable due to the knowledge of the other [2]. The mutual information between two random variables and (with marginal probability mass functions  and  and joint probability mass function ) is the relative entropy between the joint distribution and the product distribution :



It can be shown that has following properties:

1. 
2. 
3. 
4. 

Figure 1 shows the relationship between  and expressed in a Venn diagram [2].

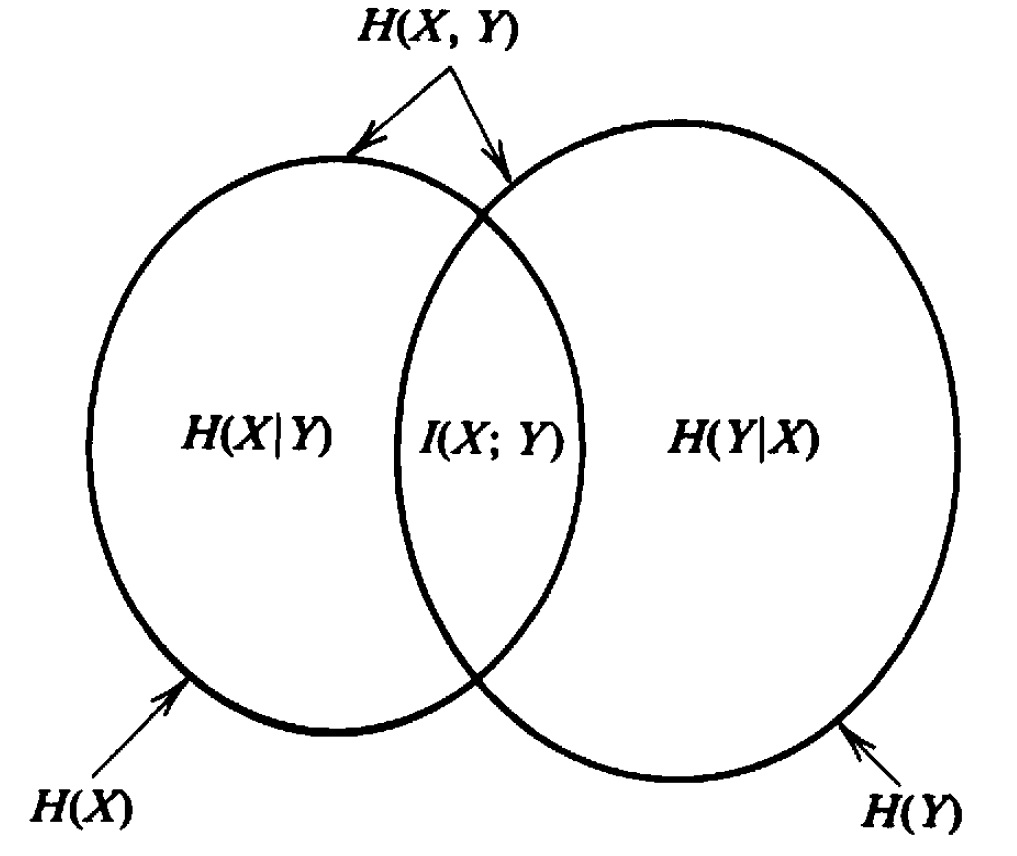


Figure1. Relationship between entropy and mutual information (adapted from [2])

Entropy measures listed above were extensively used in studies of wide range of phenomena such as in physiology [3,4], geophysics [5,6], hydrology [7,8], image analysis [9], ecology [10] and finances [11].

**Bibliography**

[1] C. Shanon, A mathematical theory of communication, The Bell systems technical journal 27, 379-423,623-656, 1948.

[2]T.M. Cover, J.A. Thomas, Entropy, relative entropy and mutual information, in: Elements of Information Theory, Wiley, New York, 1991, pp. 12–49.

[3] D. Bernardi , B. Lindner, A frequency-resolved mutual information rate and its application to neural systems, Journal of Neurophysiology  113, 1342-1357, 2015.

[4] K. Koch, J. McLean, M. Berry, P. Sterling, V. Balasubramanian, M. A. Freed, Efficiency of Information Transmission by Retinal Ganglion Cells, Current Biology 14, 1523-1530,2004.

[5] R. Kleeman, A. Majda, 2005: Predictability in a Model of Geophysical Turbulence, J. Atmos. Sci. 62, 2864–2879, 2005.

[6] L. Telesca, M. Lovallo, A. E. A. Mohamed, M. ElGabry, S. El-hady, K. M. A. Elenean, R. E. F. ElBary, Informational analysis of seismic sequences by applying the Fisher Information Measure and the Shannon entropy: An application to the 2004–2010 seismicity of Aswan area (Egypt), Physica A 391, 2889-2897, 2012.

[7] Q. Liu, Z. Yang, B. Cui, Spatial and temporal variability of annual precipitation during 1961–2006 in Yellow River Basin, China, Journal of Hydrology 361, 330–338, 2008.

[8] [V. P. Singh](http://ascelibrary.org/action/doSearch?ContribStored=Singh%2C+V+P), F. Asce,  Hydrologic Synthesis Using Entropy Theory: Review,  J. Hydrol. Eng. 16, 421–433, 2011.

[9] J. P. W. Pluim, J. B. A. Maintz, M.A. Viergever, Mutual information based registration of medical images: a survey, IEEE Transactions On Medical Imaging 22, 986-1004, 2003.

[10] H.R. Maier, A.C. Zecchin, L. Radbone, P. Goonan, Optimising the mutual information of ecological data clusters using evolutionary algorithms, Mathematical and Computer Modelling 44, 439-450,2006.

[11] R. Zhou, R. Cai, G. Tong, Applications of Entropy in Finance: A Review, Entropy 15, 4909-4931, 2013.