TROUBLES WITH SHANNON'S ACCOUNT OF THE ENTROPY IN THE CONTINUOUS CASE

Arturo Tozzi
Center for Nonlinear Science, Department of Physics, University of North Texas, Denton, Texas, USA
tozziarturo@libero.it

Abstract

According to Shannon, the entropy for continuous stochastic processes has many properties analogous to the entropy for discrete processes. Nevertheless, there are some mathematical and physical problems with Shannon's account. The obstacle arises in chapter 21 (Shannon 1948), when Shannon describes the continuous case of the entropy of ergodic ensemble of functions (i.e., the entropy of a set of functions together with a probability measure).

Two possibilities may occur:

1) The entropy in the discrete case is related to the logarithm of the PROBABILITY of long sequences of samples and to the NUMBER of the probable sequences of long length,
2) the entropy in the continuous case is related to the logarithm of the PROBABILITY DENSITY of long sequences of samples and to the VOLUME of the probable sequences of long length in the function space.

This divergence follows from Shannon's requirement of systems under assessment to be ergodic, thus allowing the continuous space to be divided into many small cells. Indeed, the key property and main tenet of the Shannon's account of entropy, both discrete and continuous, is the occurrence of ergodicity, such that every generated sequence displays the same statistical properties. If an ensemble is ergodic, each function is the set typical of the ensemble: each function can be expected to go through all the convolutions of any of the functions of the set. The presence of ergodicity leads to the statistical homogeneity required by Shannon to treat the entropy of discrete noiseless channels, of discrete channels with noise and of continuous channels.

In the case of white noise and large n, there is a well-defined volume (at least in the logarithmic sense) of high probability, and within this volume the probability density is relatively uniform (in the logarithmic sense). Shannon demonstrates that the region of high probability is a sphere of radius $\sqrt{nN}$ in the discrete case, while in the continuous case the volume of high probability is the squared radius of a sphere having the same volume.

However, it is well-known that the more the dimensions of the sphere, the less its volume (Kůrková V. 2019). In higher-dimensional spheres, the ball volume tends to be concentrated near the equator, approaching the zero as their dimensionality tends to infinity. If the volume of the higher-dimensional ball is progressively "shrunken" inside the sphere, this means that the system's ergodicity gets lost. This also means that the volume of the probable sequences of long length in the function space, which is a typical parameter of the entropy in the continuous case, does not hold anymore in multidimensions.

When the ergodicity is not guaranteed, the operation performed by Shannon of dividing the continuum of messages and signals into a large but finite number of homogeneous small regions is doomed to failure. Therefore, the concept of Shannon entropy is undermined by its very foundations since its ergodic tenet is not viable in the real multidimensional world.

Furthermore, Shannon portrays his stochastic ergodic systems in terms of random walks, white noise, Markov paths. It has been shown that this approach is not profitable when random walks occur in phase spaces of dimensions higher than two. The more the dimensions, the more the (seemingly) stochastic paths are constrained, because their trajectories cannot resume to the starting point. This means that high-dimensional tracks, ubiquitous in real world physical/biological phenomena, cannot be operationally treated in terms of closed paths. This also means that memoryless events disconnected from the past such as Markov chains cannot exist in high dimensions. For further details, see: Tozzi (2020).

REFERENCES

1) Kůrková V. 2019. Some insights from high-dimensional spheres: Comment on "The unreasonable effectiveness of small neural ensembles in high-dimensional brain" by Alexander N. Gorban et al. Phys Life Rev. https://doi.org/10.1016/j.plrev.2019.03.014.
2) Shannon CE. 1948. A Mathematical Theory of information. The Bell System Technical Journal 27:379-423.
3) Tozzi A. 2021. Random Walks Are Not So Random, After All. viXra:2010.0002.