Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2026, 14(3), 1102-1105
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Reexamination of Continuous Shannon Entropy

Pirooz Mohazzabi*, **Nathan Hugh Jones**

Department of Mathematics and Physics, University of Wisconsin-Parkside, Kenosha, WI, USA
Email: *mohazzab@uwp.edu

How to cite this paper: Mohazzabi, P. and Jones, N.H. (2026) A Reexamination of Continuous Shannon Entropy. *Journal of Applied Mathematics and Physics*, **14**, 1102-1105.
https://doi.org/10.4236/jamp.2026.143051

Received: January 24, 2026
Accepted: March 6, 2026
Published: March 9, 2026

Abstract

Extension of discrete Shannon entropy to continuous or differential entropy is reexamined. It is shown that the differential entropy, which is applied in many fields, is not the direct limit of discrete entropy, which is also responsible for its shortcomings.

Keywords

Shannon, Entropy, Discrete, Continuous, Differential

1. Introduction

Consider a random experiment with n different outcomes, with the corresponding probabilities $\{p_i\}_{i=1}^{n}$ with $\sum_{i=1}^{n} p_i = 1$. For such an experiment, the discrete Shannon entropy, also known as the information theoretic entropy S is defined by [1]-[3]

$$S = -\sum_{i=1}^{n} p_i \ln p_i . \tag{1}$$

Here we restrict ourselves to the natural logarithm, but in general, Shannon entropy can have the logarithm of any base. If all possible outcomes are equally likely, Shannon entropy is also related to thermodynamic configurational entropy, or Boltzmann entropy, by [4] [5]

$$S = k_B \ln \Omega \tag{2}$$

where k_B is the Boltzmann constant and Ω is the number of possible outcomes or microstates.

For a continuous random variable x, the common practice is to replace the discrete probabilities p_i in Equation (1) by the probability density function of the variable $f(x)$, and change the summation to integral. Therefore, the continuous Shannon entropy, also known as differential entropy, for the continuous random variable, is written as [6]-[8]

$$S = -\int_a^b f(x) \ln f(x) \mathrm{d}x \tag{3}$$

where the limits a and b define the support set or the interval of the random variable x. This extension of Shannon entropy, however, suffers from certain shortcomings. For example, it is dimensionally inconsistent. Thus, if the random variable has a physical dimension such as length L, then the dimension of $f(x)$ will be L^{-1}. However, the argument of a logarithmic function must be dimensionless. More specifically, if the variable x is length, depending on whether we use centimeters or meters for x, the value of the entropy changes. Another shortcoming is that for certain functions $f(x)$, the entropy becomes negative whereas the discrete Shannon entropy is always non-negative. For example, for the probability density function

$$f(x) = \begin{cases} 2 & 0 \le x \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

the continuous Shannon entropy of Equation (3) becomes negative. Nevertheless, continuous Shannon entropy has been applied in many areas, including statistical mechanics, thermodynamics, as well as signal processing and information theory.

Some of the inconsistencies described above have been pointed out in the literature [9]. In this article, we re-examine the transition from discrete to continuous Shannon entropy in a rigorous way. We clearly show that the continuous Shannon entropy described by Equation (3) is not the direct limit of the discrete entropy.

2. Transition from Discrete to Continuous Shannon Entropy

Let $f(x)$ be the probability density function of a random variable x defined on the interval $[a,b]$. Then the probability of the outcome being between x and $x+\Delta x$ is approximately $f(x)\Delta x$, which becomes exact when $\Delta x \to 0$.

Let us divide the interval $[a,b]$ into n equal slices, each of width

$$\Delta x_i = \Delta x = \frac{b-a}{n} \tag{5}$$

Then the discrete Shannon entropy of the outcomes would be

$$S_n = -\sum_{i=1}^{n} f(x_i)\Delta x_i \ln\left[f(x_i)\Delta x_i \right] \tag{6}$$

which can be written as

$$S_n = -\sum_{i=1}^{n} f(x_i) \ln f(x_i) \Delta x_i - \sum_{i=1}^{n} f(x_i)\Delta x_i \ln \Delta x_i \tag{7}$$

where S_n indicates approximation of S when $f(x)$ is discretized by n terms. Now, if we let $n \to \infty$ such that $\Delta x_i \to 0$, then each term in the second sum vanishes since $\lim_{u\to 0} u \ln u = 0$. The first sum, on the other hand, by definition becomes a Riemann integral,

$$\lim_{n\to\infty}\sum_{i=1}^{n} f(x_i)\ln f(x_i)\Delta x_i = \int_a^b f(x)\ln f(x)\,\mathrm{d}x \tag{8}$$

Therefore, it seems plausible that the discrete Shannon entropy reduces to the continuous entropy given by Equation (3).

The problem, however, is that if we compare the corresponding terms in the two sums of Equation (7), we see that as $\Delta x_i \to 0$, the terms of the second sum are much larger than those in the first sum. This is because the ratio of a term in the second sum to the corresponding term in the first sum is

$$\frac{f(x_i)\Delta x_i \ln \Delta x_i}{f(x_i)\ln f(x_i)\Delta x_i} = \frac{\ln \Delta x_i}{\ln f(x_i)} \tag{9}$$

Then as $\Delta x_i \to 0$, the denominator of this fraction remains finite whereas the numerator approaches $-\infty$. Consequently, in the above transition from discrete to continuous Shannon entropy, the main term is discarded. This accounts for the shortcomings of the continuous Shannoon entropy given by Equation (3).

We now take a closer look at a more fundamental question: Can the definition of Shannon entropy be extended to continuous random variables as described above? To do so, we start with Equation (6), which is the correct Shannon entropy for a discretized continuous probability distribution. Using Equation (5), this equation becomes

$$S_n = -\sum_{i=1}^{n} f(x_i)\left(\frac{b-a}{n}\right)\ln\left[f(x_i)\left(\frac{b-a}{n}\right)\right] \tag{10}$$

which reduces to

$$S_n = -(b-a)\left[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\ln f(x_i)\right] - (b-a)\left[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right]\ln\left(\frac{b-a}{n}\right) \tag{11}$$

or

$$S_n = -(b-a)\,\overline{f(x_i)\ln f(x_i)} - (b-a)\,\overline{f(x_i)}\ln\left(\frac{b-a}{n}\right) \tag{12}$$

where the over-bar represents average value. Now, as $n \to \infty$, every term on the right hand side of this expression remains finite except $\ln\left(\frac{b-a}{n}\right)$ which becomes $-\infty$. Therefore,

$$S = \lim_{n\to\infty} S_n = \infty \tag{13}$$

Consequently, as stated earlier, differential entropy is not the direct limit of discrete entropy [9]. We also verified this result computationally.

3. Conclusions

Although continuous Shannon entropy is applied in many areas, it lacks some of the properties of entropy. Extending the expression of discrete Shannon entropy to a continuous random variable results in an infinite entropy; therefore, such a simple extension is not possible. Furthermore, Equation (3), which is commonly used to express differential or continuous entropy is dimensionally inconsistent, scale dependent, and for certain probability density functions, becomes negative. Despite the fact that its shortcomings have been pointed out in some references, the expression continues to be applied in many fields.

It is also worth pointing out that some authors have suggested an alternative definition for continuous entropy by introducing an *invariant factor* $m(x)$ into Equation (3) [7],

$$H = -\int_a^b f(x) \ln \frac{f(x)}{m(x)} \mathrm{d}x. \tag{14}$$

However, because of this added function, this equation should not be called continuous "Shannon" entropy.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Shannon, C.E. (1948) A Mathematical Theory of Communication. *Bell System Technical Journal*, **27**, 623-656. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

[2] Rioul, O. (2021) This Is IT: A Primer on Shannon's Entropy and Information. In: Duplantier, B. and Rivasseau, V., Eds., *Information Theory*, Birkhäuser, 49-86.

[3] Viznyuk, S. (2015) Shannon's Entropy Revisited. arXiv: 1504.01407. https://arxiv.org/pdf/1504.01407v1

[4] Mohazzabi, P., Jones, N.H., Tucker, R.H. and Winter, N.J. (2025) The Entropy of a DNA Strand. *Journal of Applied Mathematics and Physics*, **13**, 4490-4497. https://doi.org/10.4236/jamp.2025.1312246

[5] Mandl, F. (1991) Statistical Physics. 2nd Edition, Wiley, 61 p.

[6] Ihara, S. (1993) Information Theory for Continuous Systems. World Scientific Publishing Co. Pte. Ltd. https://doi.org/10.1142/9789814355827

[7] Marsh, C. (2013) Introduction to Continuous Entropy. https://www.crmarsh.com/pdf/Charles_Marsh_Continuous_Entropy.pdf

[8] Cover, T.M. and Thomas, J.A. (1991) Elements of Information Theory. Wiley, 224 p.

[9] Dubois, D.M. (2019) Generalization of the Discrete and Continuous Shannon Entropy by Positive Definite Functions Related to the Constant of Motion of the Non-Linear Lotka-Volterra System. *Journal of Physics*: *Conference Series*, **1251**, Article ID: 012012. https://doi.org/10.1088/1742-6596/1251/1/012012