Aequat. Math.

https://doi.org/10.1007/s00010-024-01068-y

Aequationes Mathematicae

Shannon's entropy and its bounds for some a priori known equiprobable states

Eleutherius Symeonidis and Flavia-Corina Mitroi-Symeonidis

Abstract. It is known that Shannon's entropy is nonnegative and its maximum value is reached for equiprobable events. Adding or removing impossible events does not affect Shannon's entropy. However, if we increase the number of events and consider not necessarily all of them equiprobable, but at least as many of them as the initial number of equiprobable events, how does Shannon's entropy change? We study the lower bound of the interval where the probability value of the a priori assumed equiprobable states must belong when the entropy increases.

Mathematics Subject Classification. Primary 94A17.

Keywords. Shannon's entropy, Permutation entropy, Time series analysis.

1. Introduction

Shannon's entropy [9] is defined as

$$H(P) \equiv -\sum_{i=1}^{n} p_i \log p_i,$$

where $P = (p_1, \ldots, p_n)$ is a finite probability distribution. (Here and elsewhere in this paper, log denotes the natural logarithm.) It is nonnegative and its maximum value is $H(U) = \log n$, where $U = (1/n, \ldots, 1/n)$. Throughout the paper we use the convention $0 \log 0 = 0$.

The known recursivity (grouping) property of Shannon's entropy (see for instance [1,2]) states that

$$H\left(p_1, p_2, \ldots, p_n\right) = H\left(p_1 + p_2, \ldots, p_n\right) + \left(p_1 + p_2\right) H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right). \tag{1.1}$$

Published online: 15 May 2024

Apparently the simple question "how do slight modifications of the probabilities affect the entropy?" does not have many answers in the literature, and we stated in [6] the following open problem.

Open Problem. Find the lower bound (threshold) $a(k) \geq 0$ such that, if the probability distribution $P = (p_1, \ldots, p_n)$ has at least k nonzero and equal components $\geq a(k)$, then the Shannon entropy $H(P)$ attains its minimum when $n-k$ components of P are zero. In other words, find the best (smallest) $a(k)$ such that

$$H(p_1, \ldots, p_{n-k}, p, \ldots, p) \geq H(0, \ldots, 0, 1/k, \ldots, 1/k)$$

for all probability distributions $P = (p_1, \ldots, p_{n-k}, p, \ldots, p) \in \mathbb{R}_+^n$ such that $p > 0$ and $a(k) \leq p \leq 1/k$ $(k \leq n-1)$. Obviously $a(k) \leq 1/k$.

2. Main results

Our starting point now is the following answer given in [3], useful for computer assisted analysis of the experimental data.

Proposition 1. *Let the probability distribution $P = (p_1, \ldots, p_n)$ be such that it has at least k nonzero and equal components $p_{n-k+1} = \cdots = p_n = p$. The best (smallest) $a(k) \geq 0$ such that*

$$H(p_1, \ldots, p_{n-k}, p, \ldots, p) \geq H(0, \ldots, 0, 1/k, \ldots, 1/k), \tag{2.1}$$

for all P for which additionally $a(k) \leq p \leq 1/k$ holds, is the value of the abscisse of the first intersection of the horizontal line $y = \log(k)$ and the graph of the function

$$f_k(p) = -kp\log(p) - (1-kp)\log(1-kp), \ 0 \leq p \leq 1/k.$$

Figure 1 shows these intersections for $k = 1, \ldots, 5$.

In [3], the proof of this result was reduced to the fact that $a(k)$ is given as the smallest solution p of the equation

$$-kp\log(p) - (1-kp)\log(1-kp) = \log(k). \tag{2.2}$$

The maximum of the function $f_k(p)$ is $\log(k+1)$. Therefore, we are interested in the part of the graph which is in between the horizontal lines $y = \log(k)$ and $y = \log(k+1)$. The line $y = \log(k)$ meets the graph of $f_k(p)$ twice: one point has as abscisse the required bound $a(k)$, the other is situated at the right endpoint of the domain of $f_k(p)$, $p = 1/k$.

In [3] we also provided some particular estimates of interest for $a(k)$, found with the computer package MATLAB, needed for practical purposes, as in Fig. 1.

In what follows, we look for a nicer formula (however still implicit) of the first solution of the equation (2.2), $a(k)$. As a result, the equation (2.2) is

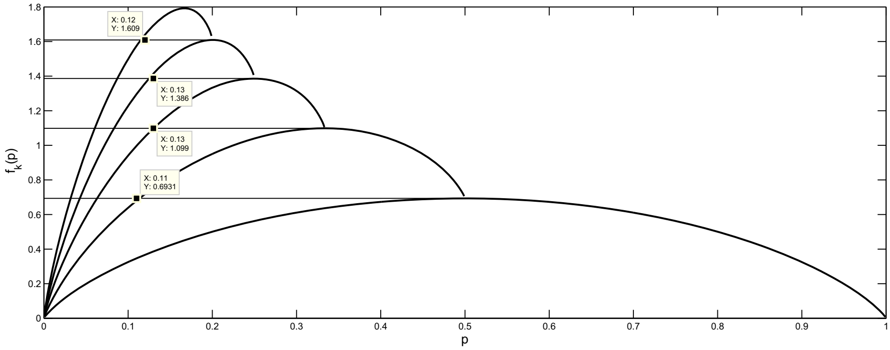

FIGURE 1. The plot of the function $f_k\left(p\right)$ for $k = 1, \dots, 5$

solved also in the case when k is not an integer, and we consider this fact of some theoretical importance.
If $x := kp$, equation (2.2) takes the form

$$F(k,x) := \log k + x \log \frac{x}{k} + (1-x)\log(1-x) = 0. \tag{2.3}$$

Since $F(k,x) = (1-x)\log k + x\log x + (1-x)\log(1-x)$, (2.3) is solvable in k, and the solution is

$$k = \frac{x^{\frac{x}{x-1}}}{1-x}, \qquad 0 < x < 1.$$

As a result we obtain $p = a(k)$ as a function of $x = kp$:

$$p = p(x) = \frac{x}{k} = (1-x)x^{\frac{1}{1-x}}. \tag{2.4}$$

In Fig. 2 (generated with MATLAB as well) we plot the function $(1-x)x^{\frac{1}{1-x}}$ and the straight lines $\frac{x}{k}$ for $k = 1, \dots, 5$. The intersections correspond to $a\left(k\right)$ for $k = 1, \dots, 5$.

Proposition 2. *With the above notation, it holds that*

$$0 < a\left(k\right) \leq 1/(k+1),$$

for $k \geq 2$.

Proof. It is straightforward to observe, as an immediate consequence of the recursivity of Shannon's entropy (1.1), that

$$H\left(p_1, p_2, \dots, p_{k+2}\right) \geq H\left(p_1 + p_2, \dots, p_{k+2}\right).$$

In the case $p_1 + p_2 = 1/(k+1)$, $p_3 = \cdots = p_{k+2} = 1/(k+1)$ this yields

$$H(p_1, p_2, \underbrace{1/(k+1), \dots, 1/(k+1))}_{k \text{ terms}}) \geq H\left(1/(k+1), \dots, 1/(k+1)\right) = \log(k+1),$$

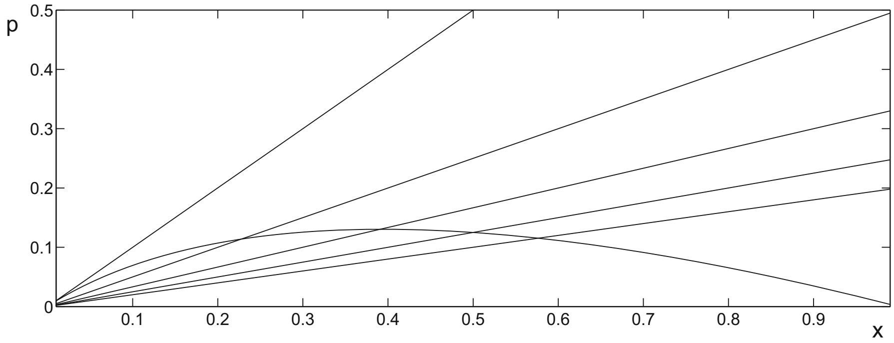

Figure 2. The plot of the function $(1-x)x^{\frac{1}{1-x}}$ and the straight lines $\frac{x}{k}$ for $k = 1, \dots, 5$

and we can infer that for all $n > k$ it holds that

$$H(p_1, \dots, p_{n-k}, \underbrace{1/(k+1), \dots, 1/(k+1)}_{k \text{ terms}})$$
$$\geq \log(k+1) \geq H\left(0, \dots, 0, 1/k, \dots, 1/k\right)$$

for all positive $p_1, \dots, p_{n-k}$ such that $p_1 + \dots + p_{n-k} = 1/(k+1)$.

Then for $p = 1/(k+1)$ inequality (2.1) holds true, therefore $a\left(k\right) \leq 1/(k+1)$. □

Geometrically speaking, this means that the intersection of the graph of the function $(1-x)x^{\frac{1}{1-x}}$ with the straight line $\frac{x}{k}$ has a lower ordinate than the intersection of the straight line $\frac{x}{k+1}$ with the vertical line $x = 1$.

Remark 1. Note that, according to Corollary 2 in [3], one also has

$$-kp\log p - (1-kp)\log(1-kp) \leq H\left(p_1, \dots, p_{n-k}, p, \dots, p\right)$$
$$\leq -kp\log p - (1-kp)\log\frac{1-kp}{n-k} \leq \log n.$$

The first equality holds true for $p_1 = \dots = p_{n-k-1} = 0$, $p_{n-k} = 1-kp$, the second equality is valid for $p_1 = \dots = p_{n-k} = \frac{1-kp}{n-k}$. The last equality holds true for $p = 1/n$. In this paper we studied an alternative way to determine the domain of p such that

$$\log k \leq -kp\log p - (1-kp)\log(1-kp).$$

Such studies become of practical interest when one uses redistributing algorithms to analyze the time series, as in the papers [4–8].

Author contributions All authors have seen and approved the manuscript and have contributed significantly in the paper. There are no conflicts of interest to disclose.

Declarations

Conflict of interest The authors declare no Conflict of interest.

References

[1] Fadeev, D.K.: On the concept of entropy of a finite probabilistic scheme. Uspekhi Math. Nauk. **11**(1), 227–31 (1956)
[2] Khinchin Ya A.: On the concept of entropy in the theory of probabilities. Uspekhi Mat. Nauk VIII **3**(55) (1953)
[3] Mitroi-Symeonidis, F.-C., Symeonidis, E.: Redistributing algorithms and Shannon's Entropy. Aequat. Math. **96**, 267–277 (2022)
[4] Mitroi-Symeonidis, F.-C., Anghel, I.: The permutation entropy and the assessment of compartment fire development: growth and decay, (ROMFIN2019). Math. Rep. **23**(73), 203–210 (2021)
[5] Mitroi-Symeonidis, F.-C., Anghel, I., Minculete, N.: Parametric Jensen-Shannon statistical complexity and its applications on full-scale compartment fire data. Symmetry (Special Issue: Symmetry in Applied Mathematics) **12**(1), 22 (2020)
[6] Mitroi-Symeonidis, F.-C., Anghel, I., Lalu, O., Popa, C.: The permutation entropy and its applications on fire tests data. J. Appl. Comput. Mech. **6**(SI), 1380–1393 (2020)
[7] Mitroi-Symeonidis, F.-C., Anghel, I., Furuichi, S.: Encodings for the calculation of the permutation hypoentropy and their applications on full-scale compartment fire data. Acta Technica Napocensis **2**(6), 607–616 (2019)
[8] Mitroi-Symeonidis, F.-C., Anghel, I.: The PYR-algorithm for time series modeling of temperature values and its applications on full-scale compartment fire data. In: Acta Technica Napocensis, Series: Applied Mathematics, Mechanics, and Engineering, Vol. 63(IV), pp. 403–410 (2020)
[9] Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)

Eleutherius Symeonidis
Mathematisch-Geographische Fakultät
Katholische Universität Eichstätt-Ingolstadt
Eichstätt 85071
Germany
e-mail: e.symeonidis@ku.de

Flavia-Corina Mitroi-Symeonidis
Department of Applied Mathematics
Academy of Economic Studies
Calea Dorobanti 15-17, sector 1
Bucharest 010552
Romania
e-mail: fcmitroi@yahoo.com

Received: November 3, 2023
Revised: April 4, 2024
Accepted: April 5, 2024