AN AXIOMATIC DEFINITION OF SHANNON'S ENTROPY

K. K. Nambiar   Pramod Varma K.   Vandana Saroch

School of Computer and Systems Sciences
Jawaharlal Nehru University, New Delhi 110067

(Submitted September 1991)

Communicated by E. Y. Rodin

Abstract–An axiomatic definition of Shannon's entropy H is given under the assumption that it is an analytic function. From the two simple axioms given, $H = -\sum p_k \; log \; p_k$ follows in a straightforward manner.

1. NOTATIONS

Many definitions of entropy have been suggested in the literature during the last four decades [**1-3**], the one given here we consider to be particularly simple. First we explain the definitions used in the sequel.

1. $\mathbf{Z} = [z_{ij}]$: a matrix of order $m \times n$ with complex elements. As is conventional, we have used i for row index and j for column index and this use is invariably adhered to in what follows.

2. $\mathbf{P} = [p_{ij}]$: a matrix of order $m \times n$ with nonnegative real elements adding upto 1, called *distribution matrix*

3. $H[z_{ij}]$: the same as $H(\mathbf{Z})$

4. $\mathbf{p} = [p_i]$: a column matrix

5. $H[p_i]$: the same as $H(\mathbf{p})$

6. $\mathbf{q}_k = [q_{kj}]$: a row matrix

7. $H[q_{kj}]$: the same as $H(\mathbf{q}_k)$

2. AXIOMATIC DEFINITION

For the definition of entropy we want an analytic function $H[z_{ij}]$ such that its restriction $H[p_{ij}]$ satisfies the following conditions.

Axiom 1 : $H[p_{ij}] = f(c)$, if all nonzero p_{ij}'s are equal to c. Here $f(z)$ is an analytic function with $f(1/2) = 1$.

Axiom 2 : $H[p_{ij}] = H[p_i] + \sum_{k=1}^{m} p_k \; H[q_{kj}]$ where $p_i = \sum_{j=1}^{n} p_{ij}$ and $q_{kj} = p_{kj}/p_k$.

Theorem : $H[p_i] = -\sum_{k=1}^{m} p_k \log \; p_k$ where logarithm is to the base 2.

Proof : Let $p_{ij} = 1/mn$. Then $p_i = 1/m$, $q_{kj} = p_{kj}/p_k = 1/n$, and by axiom 1, $H[p_{ij}] = f\,(1/mn)$.

By axiom 2,

$$f\left(\frac{1}{mn}\right) = f\left(\frac{1}{m}\right) \; + \; \sum_{k=1}^{m} \frac{1}{m} f\left(\frac{1}{n}\right)$$

$$f\left(\frac{1}{mn}\right) = f\left(\frac{1}{m}\right) + f\left(\frac{1}{n}\right)$$

$$f\left(\frac{1}{n^2}\right) = 2\; f\left(\frac{1}{n}\right)$$

$$\vdots$$

$$f\left(\frac{1}{n^k}\right) = k\ f\left(\frac{1}{n}\right)$$

$$f\left(\frac{1}{2^k}\right) = k\ f\left(\frac{1}{2}\right) = k = -\log\left(\frac{1}{2^k}\right)$$

The values of $f(z)$ and $-\log z$ coincide at an infinite number of points converging to zero. Since both $\log z$ and $f(z)$ are analytic functions, they must be the same. Thus $f(z) = -\log z$.

Now, consider a distribution matrix $[p_i]$, $i = 1, 2, 3, \cdots, m$ where the p_i's are rational numbers. Let all p_i's have the same denominator M. Now construct a distribution matrix $[p_{ij}]$ of order $m \times s$ where s is the value of the highest numerator among all p_i's. Give the value $1/M$ for the first $p_i M$ elements of the i^{th} row of the matrix and the value zero for the rest. For this matrix, nonzero p_{ij}'s and q_{kj}'s are $p_{ij} = 1/M$ and $q_{kj} = p_{kj}/p_k = 1/Mp_k$
By axiom 2,

$$H[p_{ij}] = H[p_i] + \sum_{k=1}^{m} p_k\ H[q_{kj}]$$

$$f\left(\frac{1}{M}\right) = H[p_i] + \sum_{k=1}^{m} p_k\ f\left(\frac{1}{Mp_k}\right)$$

$$\log M = H[p_i] + \sum_{k=1}^{m} p_k\ \log\ p_k + \sum_{k=1}^{m} p_k\ \log M$$

$$H[p_i] = -\sum_{k=1}^{m} p_k\ \log\ p_k$$

Since $H[z_{ij}]$ is analytic, it follows that

$$H[p_i] = -\sum_{k=1}^{m} p_k\ \log\ p_k$$

even for irrational values of p_i.

REFERENCES

[1] Shannon, C. E. and Weaver, W. , *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

[2] Khinchin, A. I. , *Mathematical Foundations of Information Theory*, Dover Publications, New York, 1957.

[3] Mathai, A. M. and Rathie, P. N. , *Basic Concepts in Information Theory and Statistics : Axiomatic Foundations and Applications*, Wiley, New York, 1975.