Marzena Kosno; Dominik Szynal
On a characterization of the Shannon entropy

*Kybernetika*, Vol. 27 (1991), No. 5, 421--425

Persistent URL: [http://dml.cz/dmlcz/125855](http://dml.cz/dmlcz/125855)

**Terms of use:**


Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these Terms of use.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project DML-CZ: The Czech Digital Mathematics Library [http://project.dml.cz](http://project.dml.cz)
ON A CHARACTERIZATION OF THE SHANNON ENTROPY

MARZENA KOSNO, DOMINIK SZYNAL

We give a characterization of the Shannon entropy using less restrictive assumptions on symmetry than extreme symmetry and block symmetry of Prem Nath and Mohan Kaur [2].

1. INTRODUCTION

Let
\[ D_n = \{(p_1, p_2, \ldots, p_n) : p_i \geq 0, \quad i = 1, 2, \ldots, n, \quad \sum_{i=1}^{n} p_i = 1\}, \quad n \geq 1, \]
be the set of all finite discrete n-component probability distribution with nonnegative elements. There are different axioms for the Shannon entropy \( H_n : D_n \to \mathbb{R}, n \geq 1, \) defined by
\[ H_n(p_1, p_2, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log_2 p_k \tag{1} \]
with \( 0 \log_2 0 = 0. \) For instance, D. K. Fadeev [1] proposed the following postulates:

I. \( p \to h(p) := H_2(p, 1 - p) \) is a continuous function of \( p, 0 \leq p \leq 1. \)

II. \( H_n \) is a real symmetric function of \((p_1, p_2, \ldots, p_n)\) on \( D_n \) for \( n \geq 2. \)

III. \( H_n \) is recursive, that is
\[ H_n(p_1, p_2, \ldots, p_n) = H_{n-1}(p_1, p_2, \ldots, p_n) + (p_1 + p_2) H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2)), \quad p_1 + p_2 > 0 \]

IV. \( H_2(\frac{1}{2}, \frac{1}{2}) = 1 \)

H. Tverberg [4] has shown that (1) holds true when in I \( h \) is an integrable function while P. M. Lee [3] ordered only \( h \) to be measurable. Moreover, Prem Nath and Man Mohan Kaur [2] have shown that one can use in II an extreme symmetric function or a block symmetric function instead of symmetric function.

In this note we weaken the symmetry postulate II and we generalize results of Prem Nath and Man Mohan Kaur. Moreover, we use so called grouping axiom instead of III.
2. THE MAIN RESULTS

The following theorem characterizes the Shannon entropy.

Theorem 1. Let $H_1(1), H_2(p_1, p_2), \ldots, H_n(p_1, p_2, \ldots, p_n)$ be a sequence of real functions defined on $\mathcal{D}_n, n \geq 1$.

We assume the following three conditions as axioms:

I' $h(p) := H_2(p, 1 - p)$ is a Lebesgue integrable function on $[0, 1]$.

II' (the axiom of reduced symmetry).

$$H_2(p_1, \ldots, p_{n-2}, p_{n-1}, p_n) = H_2(p_1, \ldots, p_{n-2}, p_n, p_{n-1})$$

for all $(p_1, p_2, \ldots, p_n) \in \mathcal{D}_n, \ n \geq 2$.

III' (the grouping axiom).

Letting $P_n = \sum_{k=1}^{n} p_k$, $P_1 = 1$, we have

$$H_2(p_1, p_2, \ldots, p_n) = H_2(p_{n-1}, p_n) + H_2(p_1, \ldots, p_{n-2}, p_n, p_{n-1}).$$

Then

$$H_n(p_1, p_2, \ldots, p_n) = -C \sum_{k=1}^{n} p_k \log p_k$$

(2)

where $C$ is a positive constant.

Proof. By III' we have

$$H_2(p_1 + p_2, p) = H_2(p_1 + p_2, p) + (p_1 + p_2) H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2))$$

and

$$H_2(p_1, p, p_2) = H_2(p_1 + p, p_2) + (p_1 + p) H_2(p_1/(p_1 + p), p/(p_1 + p)).$$

Now, by II' we get

$$H_2(p_1 + p_2, p) + (p_1 + p_2) H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2)) =$$

$$= H_2(p_1 + p, p_2) + (p_1 + p) H_2(p_1/(p_1 + p), p/(p_1 + p)).$$

Again using II' we have

$$H_2(p_1 + p_2, p) + (1 - p) H_2(p_1/(1 - p), p_2/(1 - p)) =$$

$$= H_2(p_2, p_1 + p) + (1 - p_2) H_2(p/(1 - p_2), p_1/(1 - p_2)).$$

Hence we conclude that the function $h$ satisfies the following functional equations

(a) $h(p) = h(1 - p)$

(b) $h(p) + (1 - p) h(p_2/(1 - p)) = h(p_2) + (1 - p_2) h(p/(1 - p_2))$

Now following Tveberg’s arguments [4] we can get

$$h(p) = C[-p \log p - (1 - p) \log (1 - p)]$$

where $C$ is a positive constant, which proves (2) for $n = 2$. 422
Using now $\Pi_n'$ and the induction principle we get

\[ H_n(p_1, p_2, \ldots, p_n) = H_2(p_{n-1}, p_n) + P_{n-1}H_{n-1}(p_1/P_{n-1}, \ldots, p_{n-1}/P_{n-1}) = \]

\[ = -C[p_{n-1}\log p_{n-1} + p_n\log p_n + P_{n-1}\sum_{k=1}^{n-1}(p_k/P_{n-1})\log(p_k/P_{n-1})] = \]

\[ = -C\sum_{k=1}^{n}p_k\log p_k, \]

which completes the proof of (2).

It is not difficult to verify that the function

\[ f_n: \mathcal{D}_n \to \mathbb{R} \text{ defined by} \]

\[ f_n(x_1, x_2, \ldots, x_{n-2}, x_{n-1}, x_n) = x_{n-1} + x_n \]

satisfies the axiom of reduced symmetry but it is not symmetric and even it is no extreme symmetric neither block symmetric. Indeed, we see that

\[ f_4(\frac{1}{4}, \frac{1}{6}, \frac{1}{4}, \frac{1}{12}) = \frac{1}{4} + \frac{1}{12} = \frac{1}{3} = f_4(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4}), \]

but

\[ f_4(\frac{1}{6}, \frac{1}{12}, \frac{1}{2}, \frac{1}{4}) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4} = f_4(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4}). \]

Moreover,

\[ f_4(\frac{1}{2}, \frac{1}{6}, \frac{1}{4}, \frac{1}{2}) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4} = f_4(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4}), \]

and

\[ f_4(\frac{1}{4}, \frac{1}{12}, \frac{1}{2}, \frac{1}{6}) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3} = f_4(\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4}). \]

One can also state that the axiom of reduced symmetry is independent of the postulates of extreme symmetry and block symmetry. It is enough to take into account the functions

\[ f_n(x_1, x_2, \ldots, x_n) = x_1 + x_n \]

and

\[ f_{2n}(x_1, x_2, \ldots, x_{2n}) = x_1 + x_2 + x_{n+1} + x_{n+2}, \]

respectively.

3. GENERALIZATIONS

Note that in the proof of Theorem 1 we have used $\Pi_n'$ only for $n = 2$ and 3. Thus in fact we have proved the following result.

**Theorem 2.** Suppose that real functions $H_1(1), H_2(p_1, p_2), \ldots, H_n(p_1, p_2, \ldots, p_n)$ defined on $\mathcal{D}_n, n \geq 1$, satisfy $I', \Pi_n'$ and $\Pi''$. $H_2(p_1, p_2) = H_2(p_2, p_1)$

\[ H_3(p_1, p_2, p_3) = H_3(p_1, p_3, p_2). \]

Then (2) holds.
Note that the conditions $\text{II}''$ and $\text{III}_n'$ for $n = 3$ imply the symmetry of $H_3(p_1, p_2, p_3)$. Indeed, using $\text{III}_n''$ and $\text{II}''$, we have
\[
H_3(p_1, p_2, p_3) = H_2(p_1 + p_2, p_3) +
\]
\[
+ (p_1 + p_2) H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2)) =
\]
\[
= H_2(p_2 + p_1, p_3) + (p_2 + p_1) H_2(p_2/(p_2 + p_1), p_1/(p_2 + p_1)) =
\]
\[
= H_3(p_2, p_1, p_3) = H_3(p_2, p_3, p_1).
\]
On the other hand from $\text{II}'$ and $\text{III}_n$ we deduce that
\[
H_3(p_1, p_2, p_3) = H_3(p_3, p_1, p_2) =
\]
\[
= H_2(p_1 + p_3, p_2) + (p_1 + p_3) H_2(p_1/(p_1 + p_3), p_2/(p_1 + p_3)) =
\]
\[
= H_2(p_3 + p_1, p_2) + (p_3 + p_1) H_2(p_3/(p_3 + p_1), p_2/(p_3 + p_1)) =
\]
\[
= H_3(p_3, p_1, p_2) = H_3(p_3, p_2, p_1)
\]
Hence we get the following equalities
\[
H_3(p_3, p_2, p_1) = H_3(p_3, p_1, p_2) = H_3(p_1, p_3, p_2) =
\]
\[
= H_3(p_1, p_2, p_3) = H_3(p_2, p_1, p_3) = H_3(p_2, p_3, p_1)
\]
which prove the symmetry of $H_3(p_1, p_2, p_3)$.

The above observation leads us to a stronger version of P. M. Lee [3] characterization of the Shannon entropy, in which the symmetry of $H_2(p_1, p_2)$ and $H_3(p_1, p_2, p_3)$ is replaced by the symmetry of $H_2(p_1, p_2)$ and the reduced symmetry of $H_3(p_1, p_2, p_3)$.

Namely, we have the following result.

**Theorem 3.** Suppose that real functions $H_1(1), H_2(p_1, p_2), \ldots, H_n(p_1, p_2, \ldots, p_n), n \geq 1,$ satisfy the axioms $\text{II}''$ and $\text{III}_n'$ of Theorem 2, and $h(p) := H_2(p, 1 - p)$ is a Lebesgue measurable function on $(0, 1)$.

Then (2) holds true.

Note that $H_n$ for $n \geq 3$ can be expressed in terms of the single function $h$. The property $\text{III}_n'$ gives the following formula
\[
H_n(p_1, p_2, \ldots, p_n) = \sum_{k=2}^{n} P_k h(p_k/P_k).
\]
In the case $p_1 = p_2 = \ldots = p_n = 1/n$ we have
\[
f(n) := H_n(1/n, 1/n, \ldots, 1/n) = (1/n) \sum_{k=2}^{n} kh(1/k).
\]

(Received February 28, 1990.)

424
REFERENCES


Mgr Marzena Kosno, Prof. Dominik Szynal, Mathematical Institute, Maria Curie-Sklodowska University, Pl. Marii Curie-Sklodowskiej 1, 20-031 Lublin. Poland.