J. S. Chawla

A note on characterizations of entropies

Kybernetika, Vol. 13 (1977), No. 3, (194)--199

Persistent URL: http://dml.cz/dmlcz/125048

KYBERNETIKA — VOLUME 13 (1977), NUMBER 3

A Note on Characterizations of Entropies

J. S. CHAWLA

In this note we discuss and characterize four entropies each involving a parameter, by using the concept of generalized probability distributions. We shall call them entropies of order α, β, γ and δ.

1. GENERALIZED PROBABILITY DISTRIBUTIONS

Let $(\Omega, \mathscr{R}, \mathbf{P})$ be a probability space, that is, Ω an arbitrary non-empty set, called the set of elementary events, $\mathscr{R}$ a σ-field of subsets of Ω and $\mathbf{P}$ a probability measure defined on $\mathscr{R}$. Let us call a function $\xi = \xi(w)$ which is defined for $w \in \Omega_1$ where $\Omega_1 \in \mathscr{R}$ and $\mathbf{P}(\Omega_1) > 0$, and which is measurable w.r.t. $\mathscr{R}$, a *generalized random variable*. If $\mathbf{P}(\Omega_1) = 1$ we call ξ an *ordinary* (or *complete*) random variable, while if $0 < \mathbf{P}(\Omega_1) < 1$ we call ξ an *incomplete* random variable. Clearly, an incomplete random variable can be interpreted as a quantity describing the result of an experiment depending on chance which is not always observable. The distribution of a generalized random variable will be called a *generalized probability distribution*. In particular, in the case when ξ takes on only a finite number of different values $x_1, x_2, \ldots, x_n$ the distribution of ξ consists of the set of numbers $p_k = \mathbf{P}\{\xi = x_k\}$ for $k = 1, 2, \ldots, n$. Thus a finite discrete generalized probability distribution is simply a sequence $p_1, p_2, \ldots, p_n$ of non-negative numbers such that putting $\mathsf{P} = (p_1, p_2, \ldots, p_n)$ and $W(\mathsf{P}) = \sum_{k=1}^{n} p_k$, we have, $0 < W(\mathsf{P}) \leqq 1$.

We shall call $W(\mathsf{P})$ the *weight* of the distribution. Thus the weight of an ordinary distribution is equal to 1. A distribution which has a weight less than 1 will be called an *incomplete* distribution. If $\mathsf{P} = (p_1, p_2, \ldots, p_n)$ and $\mathsf{Q} = (q_1, q_2, \ldots, q_m)$ are two finite discrete generalized probability distributions, then $\mathsf{P} * \mathsf{Q}$ is the *direct product* of the distributions P and Q and is a finite discrete generalized probability distribution consisting of the numbers $p_j q_k$ with $j = 1, 2, \ldots, n$; $k = 1, 2, \ldots, m$.

If $W(\mathsf{P}) + W(\mathsf{Q}) \leqq 1$, we put

$$\mathsf{P} \cup \mathsf{Q} = (p_1, p_2, \ldots, p_n, q_1, q_2, \ldots, q_m) .$$

2. CHARACTERIZATION OF ENTROPIES

Let Δ denote the set of all finite discrete generalized probability distributions, that is, Δ is the set of all sequences $\mathsf{P} = (p_1, p_2, \ldots, p_n)$ of non-negative numbers such that $0 < \sum_{k=1}^{n} p_k \leqq 1$. It is assumed that the entropy $H[\mathsf{P}]$ of a finite discrete generalized probability distribution is defined for all $\mathsf{P} \in \Delta$. We define the following postulates:

Postulate 1. $H[\mathsf{P}]$ is a symmetric function of the elements of P.

Postulate 2. If $\{p\}$ denotes the generalized probability distribution consisting of the probability p, then $H[\{p\}]$ is a continuous function of p in the interval $0 < p \leqq 1$.

Postulate 3. $H[\{\frac{1}{2}\}] = 1$.

Postulate 4. $H[\{1\}] = 0$.

Postulate 5. If $\mathsf{P} = (p_1, p_2, \ldots, p_n)$ and $\mathsf{Q} = (q_1, q_2, \ldots, q_m)$, then $H[\mathsf{P} * \mathsf{Q}] = H[\mathsf{P}] + H[\mathsf{Q}]$.

Postulate 6. If $\mathsf{P} = (p_1, p_2, \ldots, p_n)$ and $\mathsf{Q} = (q_1, q_2, \ldots, q_m)$, such that $W(\mathsf{P}) + W(\mathsf{Q}) \leqq 1$, then

$$H[\mathsf{P} \cup \mathsf{Q}] = \frac{W(\mathsf{P})\, H(\mathsf{P}] + W(\mathsf{Q})\, H[\mathsf{Q}]}{W(\mathsf{P}) + W(\mathsf{Q})} .$$

It has been shown by Renyi [8] that if $H[\mathsf{P}]$ satisfies the Postulates 1, 2, 3, 5 and 6, then

$$H[\mathsf{P}] = \frac{\sum_{k=1}^{n} p_k \log_2 (1/p_k)}{\sum_{k=1}^{n} p_k} ,$$

which is a well known Shannon's entropy. We define another postulate as follows:

Postulate 7. There exists a strictly monotonic and continuous function $y = g(x)$ such that if $\mathsf{P} = (p_1, p_2, \ldots, p_n)$, $\mathsf{Q} = (q_1, q_2, \ldots, q_m)$ and $W(\mathsf{P}) + W(\mathsf{Q}) \leqq 1$, then

$$H[\mathsf{P} \cup \mathsf{Q}] = g^{-1} \left| \frac{W(\mathsf{P})\, g(H[\mathsf{P}]) + W(\mathsf{Q})\, g(H[\mathsf{Q}])}{W(\mathsf{P}) + W(\mathsf{Q})} \right| .$$

It was also shown by Renyi [8] that if $g(x) = g_\alpha(x) = 2^{(1-\alpha)x}$ where $\alpha > 0$, $\alpha \neq 1$, then Postulates 1, 2, 3, 5 and 7 characterize Renyi's entropy of order α. In other words if $H[\mathsf{P}]$ satisfies postulates 1, 2, 3, 5 and 7 with $g(x) = g_\alpha(x) = 2^{(1-\alpha)x}$, $\alpha > 0$ and $\alpha \neq 1$, then, for $\mathsf{P} = (p_1, p_2, \ldots, p_n)$, we have

$$H_\alpha[\mathsf{P}] = \frac{1}{1-\alpha} \log_2 \left[\left(\sum_{k=1}^{n} p_k^\alpha \right) \Big/ \left(\sum_{k=1}^{n} p_k \right) \right],$$

which is well known Renyi's entropy.

M. Behara and P. Nath [3] introduced the following postulate for characterizing a new class of entropies.

Postulate 5′. For every $\mathsf{P} = (p_1, p_2, \ldots, p_n) \in \Delta$, $n = 1, 2, 3, \ldots$ and $\mathsf{Q} = \{q\}$

$$H[(p_1, p_2, \ldots, p_n) * q] = H[p_1 q, p_2 q, \ldots, p_n q] = a\, H[p_1, p_2, \ldots, p_n]\, H[q] + b\, H[p_1, p_2, \ldots, p_n] + b\, H[q] + c\,, \quad \text{where} \quad ac = b^2 - b \quad \text{and} \quad a \neq 0\,.$$

They proved that if $H[\mathsf{P}]$ is defined for all $\mathsf{P} \in \Delta$ and satisfies the Postulates 2, 3, 4, 5′ and 6, then

$$H[\mathsf{P}] = H_\beta[\mathsf{P}] = \frac{1 - \left(\sum_{k=1}^{n} p_k^\beta \big/ \sum_{k=1}^{n} p_k\right)}{1 - 2^{1-\beta}}\,, \quad \beta \neq 1 \quad \text{and} \quad \beta > 0\,.$$

This entropy was first of all proposed by Havrda and Charvát [7]. Later on, Daroczy [6] and M. Behara and P. Nath [2, 3] studied these entropies. Vajda [9] also characterized this entropy for finite discrete generalized probability distributions.

In the following theorems we characterize two more entropies and to avoid confusion we call them entropies of order γ and δ.

Theorem 2.1. If $H[\mathsf{P}]$ is defined for all $\mathsf{P} \in \Delta$ and satisfies Postulates 2, 3, 4 and 7 with $g(x) = g_\gamma(x) = [1 - x(1 - 2^{\gamma-1})]^{1/\gamma}$, $\gamma \neq 1$, $\gamma > 0$ and the following postulate

Postulate 5″. $H[pq] = a H[p]\, H[q] + b H[p] + b H[q] + c$, where $ac = b^2 - b$ and $a \neq 0$,

then $H[\mathsf{P}] = H_\gamma[\mathsf{P}]$, where, putting $\mathsf{P} = (p_1, p_2, \ldots, p_n)$, we have

$$H_\gamma[\mathsf{P}] = \left\{ 1 - \left[\left(\sum_{k=1}^{n} p_k^{1/\gamma} \right) \Big/ \left(\sum_{k=1}^{n} p_k \right) \right]^\gamma \right\} \Big/ (1 - 2^{\gamma-1})\,.$$

Proof. From Postulate 5″, we have

$$a\,H[pq] = a^2\,H[p]\,H[q] + ab\,H[p] + ab\,H[q] + ac \Rightarrow$$
$$\Rightarrow a\,H[pq] = a^2\,H[p]\,H[q] + ab\,H[p] + ab\,H[q] + b^2 - b \Rightarrow$$
$$\Rightarrow (a\,H[pq] + b) = (a\,H[p] + b)(a\,H[q] + b)\,.$$

Replacing $a\,H[p] + b$ by $h[p]$, we obtain

(1) $$h[pq] = h[p]\,h[q]\,.$$

The continuity of $H[p] \Rightarrow$ continuity of $h[p]$. Hence non-identically vanishing continuous solutions of (1) are of the form

$$h[p] = p^{1-\gamma}\,, \quad \gamma \neq 1$$

so that

$$H[p] = \frac{p^{1-\gamma} - b}{a}\,, \quad \gamma \neq 1\,.$$

Using Postulates 3 and 4, it can be easily seen that $b = 1$ and $a = 2^{\gamma-1} - 1$, $\gamma \neq 1$, so that

$$H[p] = H_\gamma[p] = \frac{1 - p^{1-\gamma}}{1 - 2^{\gamma-1}}\,.$$

From Postulate 7, we have

$$H[p_1, p_2, \ldots, p_n] = g^{-1}\left|\frac{\sum_{k=1}^{n} p_k\, g(H[p_k])}{\sum_{k=1}^{n} p_k}\right| =$$

$$= g^{-1}\left|\frac{\sum_{k=1}^{n} p_k\, g\left(\frac{1 - p_k^{1-\gamma}}{1 - 2^{\gamma-1}}\right)}{\sum_{k=1}^{n} p_k}\right| = g^{-1}\left|\frac{\sum_{k=1}^{n} p_k \left\{1 - (1 - 2^{\gamma-1})\frac{1 - p_k^{1-\gamma}}{1 - 2^{\gamma-1}}\right\}^{1/\gamma}}{\sum_{k=1}^{n} p_k}\right| =$$

$$= g^{-1}\left|\frac{\sum_{k=1}^{n} p_k p_k^{(1-\gamma)/\gamma}}{\sum_{k=1}^{n} p_k}\right| = g^{-1}\left|\frac{\sum_{k=1}^{n} p_k^{1/\gamma}}{\sum_{k=1}^{n} p_k}\right| =$$

$$= \left\{1 - \left(\frac{\sum_{k=1}^{n} p_k^{1/\gamma}}{\sum_{k=1}^{n} p_k}\right)^{\gamma}\right\} \Big/ (1 - 2^{\gamma-1});\ \gamma \neq 1,\ \gamma > 0\,.$$

 Because, if $y = g(x) = [1 - x(1 - 2^{\gamma-1})]^{1/\gamma}$, then

$$y^\gamma = 1 - x(1 - 2^{\gamma-1}) \Rightarrow x(1 - 2^{\gamma-1}) = 1 - y^\gamma \Rightarrow x = \frac{1 - y^\gamma}{1 - 2^{\gamma-1}} \Rightarrow$$

$$\Rightarrow g^{-1}(y) = \frac{1 - y^\gamma}{1 - 2^{\gamma-1}}.$$

The above entropy of order γ has been studied by M. Behara and J. S. Chawla [4].

Suguru Arimoto [1] while investigating the finite parameter estimation problem deduced the entropy of order γ using the concept of generalized information measure and discussed its relationship with Renyi's entropy of order α.

Theorem 2.2. If $H[\mathsf{P}]$ is defined for all $\mathsf{P} \in \Delta$ and satisfies the Postulates 2, 3, 6 and the following:

Postulate 5‴. $H^{1/\delta}[pq] = H^{1/\delta}[p] + H^{1/\delta}[q]$; $\delta \neq 1$, $\delta > 0$,

then, for $\mathsf{P} = (p_1, p_2, \ldots, p_n)$

$$H[\mathsf{P}] = H_\delta[\mathsf{P}] = \sum_{k=1}^{n} p_k |\log_2 p_k|^\delta / \sum_{k=1}^{n} p_k .$$

Proof. Replacing $H^{1/\delta}[p]$ by $h[p]$ in Postulate 5‴, we obtain

$$h[pq] = h[p] + h[q] . \tag{2}$$

Now, by Postulate 2, $H[p]$ is a continuous function of p, $p \in (0, 1]$ and hence $h[p]$ is a continuous function of p, $p \in (0, 1]$. Thus the only continuous solutions of (2) are of the form $h[p] = c \log_2 p$ and hence $H[p] = c^\delta (\log_2 p)^\delta$.

Now, $H[\frac{1}{2}] = 1 \Rightarrow 1 = c^\delta (\log_2 \frac{1}{2})^\delta \Rightarrow 1 = c^\delta (-\log_2 2)^\delta \Rightarrow 1 = c^\delta (-1)^\delta \Rightarrow (-c)^\delta = = 1 \Rightarrow c = -1$. Thus $H^{1/\delta}[p] = -\log_2 p \Rightarrow H[p] = |\log_2 p|^\delta$. For $(p_1, p_2) \in \Delta$, we have

$$H[p_1, p_2] = \frac{p_1 H[p_1] + p_2 H[p_2]}{p_1 + p_2} = \frac{p_1 |\log_2 p_1|^\delta + p_2 |\log_2 p_2|^\delta}{p_1 + p_2},$$

$$H[p_1, p_2, p_3] = \frac{(p_1 + p_2) H[p_1, p_2] + p_3 H[p_3]}{p_1 + p_2 + p_3} =$$

$$= \frac{(p_1 + p_2) \dfrac{p_1 H[p_1] + p_2 H[p_2]}{(p_1 + p_2)} + p_3 H[p_3]}{(p_1 + p_2 + p_3)} = \frac{\sum_{i=1}^{3} p_i |\log_2 p_i|^\delta}{\sum_{i=1}^{3} p_i}.$$

Using the method of mathematical induction, we obtain

$$H[p_1, p_2, \dots, p_n] = \frac{\sum_{i=1}^{n} p_i |\log_2 p_i|^\delta}{\sum_{i=1}^{n} p_i}.$$

This entropy has been studied by the author [5] for complete probability distributions. It has also been shown in [5] that the entropy of order δ like Shannon entropy is also an invariant for isomorphic Bernoulli shifts.

It is important to note that if $\mathsf{P} = (p_1, p_2, \dots, p_n)$, then

$$\lim_{\alpha \to 1} H_\alpha[\mathsf{P}] = \lim_{\beta \to 1} H_\beta[\mathsf{P}] = \lim_{\gamma \to 1} H_\gamma[\mathsf{P}] = \lim_{\delta \to 1} H_\delta[\mathsf{P}] = \frac{-\sum_{k=1}^{n} p_k \log_2 p_k}{\sum_{k=1}^{n} p_k}$$

which is Shannon's entropy for a finite discrete generalized probability distribution.

The author wishes to express his sincere appreciation to Dr. M. Behara for his encouragement and helpful suggestions during the course of this research.

(Received June 11, 1976.)

REFERENCES

[1] Suguru Arimoto: Information-Theoretical Considerations on Estimation Problems. Informamation and Control *19* (1971), 181—194.

[2] M. Behara, P. Nath: Additive and Non-Additive Entropies of Finite Measurable Partitions. Probability and Information Theory II, 1970, Springer-Verlag, 102—138.

[3] M. Behara, P. Nath: An Axiomatic Characterization of Entropy of a Finite Discrete Generalized Probability Distribution. McMaster Univ. Math. report no. 40.

[4] M. Behara, J. S. Chawla: Generalized Gamma-Entropy: "Entropy and Ergodic Theory". Selecta Statistica Canadiana, University Press of Canada/Hindustan Publishing Corporation (India), 1974, 15—38.

[5] J. S. Chawla: Ph. D. Dissertation. McMaster University, Hamilton, Ontario, Canada 1974.

[6] Z. Daroczy: Generalized Information Functions. Information and Control *16* (1970), 1, 36—51.

[7] J. Havrda, F. Charvát: Quantification Method of Classification Processes. The Concept of Structural a-Entropy. Kybernetika *3* (1967), 1, 30—35.

[8] A. Rényi: On Measures of Entropy and Information. Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, 1961, 547—561.

[9] I. Vajda: Axioms of α-Entropy of a Generalized Probability Scheme. Kybernetika *4* (1968), 2, 105—110.

Prof. J. S. Chawla, Department of Mathematics, McMaster University, Hamilton, Ontario. Canada.