Kybernetika

Abul Basar Khan

On an axiomatic characterization of entropy of order α (theoretic measure)

Kybernetika, Vol. 15 (1979), No. 4, (293)--298

Persistent URL: http://dml.cz/dmlcz/125196

KYBERNETIKA — VOLUME 15 (1979), NUMBER 4

On an Axiomatic Characterization of Entropy of Order α (Theoretic Measure)

A. BASAR KHAN*)

An axiomatic characterization of an information (entropy of order α) theoretic quantity associated with a pair of probability distributions having the same number of elements has been given, using some simple and clear postulates. This quantity under additional suitable condition leads to Kullback's relative information. Moreover the quantity also reduces to Pearson's χ^2-statistic.

1. INTRODUCTION

Vajda ([9], [10]) has investigated properties of α-entropy of a measures P with respect to Q defined as

$$H_\alpha(\mathsf{P}, \mathsf{Q}) = \int p^\alpha q^{1-\alpha} \, d\mu \,, \tag{1.1}$$

where p and q are Radon-Nikodym densities with respect to a dominating measure μ defined over the same measurable space. He has established the relation between $H_\alpha(\mathsf{P}, \mathsf{Q})$ and Bayes risk.

Rényi [6] introduced a measure of information of order α as

$$I_\alpha(\mathsf{P}, \mathsf{Q}) = \frac{1}{\alpha - 1} \log_2 H_\alpha(\mathsf{P}, \mathsf{Q}) \,, \quad \text{for} \quad \alpha \neq 0 \,. \tag{1.2}$$

*) The work is supported by Senior Research Fellowship award to the author by Council of Scientific and Industrial Research (India).

We take a measure of information (entropy) of order α in two discrete probability distributions $\mathsf{P} = (p_1, \ldots, p_n)$, $p_i \geqq 0$, $\sum_{i=1}^{n} p_i = 1$ and $\mathsf{Q} = (q_1, \ldots, q_n)$, $q_i \geqq 0$, $\sum_{i=1}^{n} q_i \leqq 1$ having the form

$$I_n^{(\alpha)}(\mathsf{P}, \mathsf{Q}) = \mu(\alpha) \left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1 \right], \quad \alpha \neq 0, \tag{1.3}$$

where $\mu(\alpha)$ $(\neq 0)$ is an arbitrary constant that depends on the parameter α.

An axiomatic characterization of information on theoretic measure associated with a pair of probability distributions having the same number of elements is given through some postulates in Section 2 and some of its special cases are also discussed.

2. CHARACTERIZATION OF $I_n^{(\alpha)}(\mathsf{P}, \mathsf{Q})$

Theorem. Let $K_n(p_1, \ldots, p_n; q_1, \ldots, q_n)$, $p_i \geqq 0$, $q_i \geqq 0$, $i = 1, \ldots, n$; $\sum_{i=1}^{n} p_i = 1$; $\sum_{i=1}^{n} q_i \leqq 1$ be a function of p_i's and q_i's satisfying the following postulates:

(i) *Continuity*: $K_n^{(\alpha)}(p_1, \ldots, p_n; q_1, \ldots, q_n)$ is continuous in the region.

(ii) *Symmetry*: $K_n^{(\alpha)}(p_1, \ldots, p_n; q_1, \ldots, q_n)$ is symmetric for any permutation of elements in P followed by the same permutation of elements in Q.

(iii) *Generalized Branching Property*:

$$K_{n+1}^{(\alpha)}(p_1, \ldots, p_{i-1}, u_{i1}, u_{i2}, p_{i+1}, \ldots, p_n; q_1, \ldots, q_{i-1}, v_{i1}, v_{i2}, q_{i+1}, \ldots, q_n) =$$

$$= K_n^{(\alpha)}(\mathsf{P}, \mathsf{Q}) + p_i^\alpha q_i^{1-\alpha} K_2^{(\alpha)} \left(\frac{u_{i1}}{p_i}, \frac{u_{i2}}{p_i}; \frac{v_{i1}}{q_i}, \frac{v_{i2}}{q_i} \right)$$

for every

$$\left. \begin{aligned} u_{i1} + u_{i2} &= p_i > 0 \\ v_{i1} + v_{i2} &= q_i > 0 \end{aligned} \right\}; \quad i = 1, \ldots, n, \quad \alpha \neq 0, \quad \alpha > 0;$$

then we have

$$K_n^{(\alpha)}(p_1, \ldots, p_n; q_1, \ldots, q_n) = \mu(\alpha) \left[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1 \right]. \tag{2.1}$$

First we prove the following lemmas:

Lemma 1. If $u_k \geqq 0$, $v_k \geqq 0$; $k = 1, \ldots, m$,

$$\sum_{k=1}^{m} u_k = p_i > 0; \quad \sum_{k=1}^{m} v_k = q_i > 0;$$

then

$$K^{(\alpha)}_{n+m-1}(p_1, \ldots, p_{i-1}, u_1, \ldots, u_m, p_{i+1}, \ldots, p_n; \; q_1, \ldots, q_{i-1}, v_1, \ldots, v_m, q_{i+1}, \ldots, q_n) = = K^{(\alpha)}_n(p_1, \ldots, p_n; q_1, \ldots, q_n) + p_i^\alpha q_i^{1-\alpha} K^{(\alpha)}_m(u_1/p_i, \ldots, u_m/p_i; v_1/q_i, \ldots, v_m/q_i) . \tag{2.2}$$

The proof of this lemma follows on the lines of the proof of Lemma 4 in the paper by Sharma and Taneja [8].

Lemma 2. If $u_{ij} \geqq 0$, $j = 1, \ldots, m_i$, $\sum_{j=1}^{m_i} u_{ij} = p_i > 0$, $\sum_{i=1}^{n} p_i = 1$ and $v_{ij} \geqq 0$, $j = 1, \ldots, m_i$, $\sum_{j=1}^{m_i} v_{ij} = q_i > 0$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} q_i = 1$ then

$$K^{(\alpha)}_{m_1+m_2+\ldots+m_n}(u_{11}, \ldots, u_{1m_1}, \ldots, u_{n1}, \ldots, u_{nm_n}; \; v_{11}, \ldots, v_{1m_1}, \ldots, v_{n1}, \ldots, v_{nm_n}) = = K^{(\alpha)}_n(p_1, \ldots, p_n; q_1, \ldots, q_n) + \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} K^{(\alpha)}_{m_i}(u_{i1}/p_i, \ldots, u_{im_i}/p_i; v_{i1}/q_i, \ldots, v_{im_i}/q_i). \tag{2.3}$$

When $m_i = m$ for all $i = 1, \ldots, n$, then the above lemma becomes

$$K^{(\alpha)}_{mn}(u_{11}, \ldots, u_{1m}, \ldots, u_{n1}, \ldots, u_{nm}; v_{11}, \ldots, v_{1m}, \ldots, v_{n1}, \ldots, v_{nm}) = = K^{(\alpha)}_n(p_1, \ldots, p_n; q_1, \ldots, q_n) + \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} \,. \, K^{(\alpha)}_m(u_{i1}/p_i, \ldots, u_{im}/p_i; v_{i1}/q_i, \ldots, v_{im}/q_i) . \tag{2.4}$$

Proof directly follows from postulate (iii).

Lemma 3. Let

$$\psi(m, r) = K^{(\alpha)}_m\left(\frac{1}{m}, \ldots, \frac{1}{m}; \frac{1}{r}, \ldots, \frac{1}{r}\right) \quad \text{for} \quad r \geqq m . \tag{2.5}$$

For $1 \leqq m \leqq r$; $1 \leqq n \leqq s$ where m, n, r and s are positive integers, we have

$$\psi(mn, rs) = \psi(n, s) + \left(\frac{1}{n}\right)^{\alpha-1} \left(\frac{1}{s}\right)^{1-\alpha} \psi(m, r) .$$

Proof. Taking

$$u_{ij} = \frac{1}{mn}, \quad v_{ij} = \frac{1}{rs}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m,$$

$$u_{ij} = 0 \quad \text{otherwise}, \quad q_i = \frac{1}{s}, \quad i = 1, \ldots, n,$$

in (2.4) and using (2.5) we have

$$(2.6) \qquad \psi(mn, rs) = \psi(n, s) + \left(\frac{1}{n}\right)^{\alpha-1}\left(\frac{1}{s}\right)^{1-\alpha}\psi(m, r)\,.$$

Symmetry in mn and rs implies

$$(2.7) \qquad \psi(mn, rs) = \psi(nm, sr) = \psi(m, r) + \left(\frac{1}{m}\right)^{\alpha-1}\left(\frac{1}{r}\right)^{1-\alpha}\psi(n, s)$$

which concludes the proof.

Lemma 4. If $\alpha \neq 0(>0)$ then

$$\psi(m, r) = \mu(\alpha)\left[\left(\frac{1}{m}\right)^{\alpha-1}\left(\frac{1}{r}\right)^{1-\alpha} - 1\right]$$

where $\mu(\alpha)\,(\neq 0)$ is an arbitrary constant depending on the parameter $\alpha(\neq 0)$.

Proof. On comparing (2.6) and (2.7) we get

$$\frac{\psi(m, r)}{\left[\left(\frac{1}{m}\right)^{\alpha-1}\left(\frac{1}{r}\right)^{1-\alpha} - 1\right]} = \frac{\psi(n, s)}{\left[\left(\frac{1}{n}\right)^{\alpha-1}\left(\frac{1}{s}\right)^{1-\alpha} - 1\right]} = \mu(\alpha) \quad \text{say}\,,$$

which gives

$$\psi(m, r) = \mu(\alpha)\left[\left(\frac{1}{m}\right)^{\alpha-1}\left(\frac{1}{r}\right)^{1-\alpha} - 1\right]$$

when $m < r$ and when $r = m$

$$\psi(m, r) = 0\,.$$

Proof of the Theorem. We prove the theorem for rationals and then continuity (axiom (i)) gives the result for reals. For this let m, r_i, and t_i be positive integers such that

$$r_i \leqq t_i\,, \quad i = 1, \ldots, n\,; \quad \sum_{i=1}^{n} r_i = m\,.$$

Now if we put

$$p_i = \frac{r_i}{m}\,, \quad q_i = \frac{t_i}{r}\,, \quad i = 1, \ldots, n\,,$$

where $\sum_{i=1}^{n} t_i \leqq r$; then by using Lemma 2

$$K_n^{(\alpha)}\left(\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}; \frac{1}{r}, \ldots, \frac{1}{r}\right) = K_n^{(\alpha)}(p_1, \ldots, p_n;\, q_1, \ldots, q_n) +$$

$$+ \sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha} K_n^{(\alpha)}\left(\frac{1}{r_i}, \ldots, \frac{1}{r_i}; \frac{1}{t_i}, \ldots, \frac{1}{t_i}\right)$$

by (2.5) and Lemma 4 becomes

$$\psi(m, r) = K_n^{(\alpha)}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) + \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\, \psi(r_i, q_i)$$

or

$$K_n^{(\alpha)}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = \psi(m, r) - \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\, \psi(r_i, t_i)$$

or

$$K_n^{(\alpha)}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = \mu(\alpha)\left[\left(\frac{1}{m}\right)^{\alpha-1}\left(\frac{1}{r}\right)^{1-\alpha} - 1 - \right.$$

$$\left. - \sum_{i=1}^{n} (p_i^\alpha / r_i^{\alpha-1})\,(q_i/t_i)^{1-\alpha} + \sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha}\right] =$$

$$= \mu(\alpha)\,[\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1]\,,$$

which conclude the proof.

Particular cases

Case I. Measure (1.3) under the condition

$$I_2(1, 0;\ \tfrac{1}{2}, \tfrac{1}{2}) = 1$$

reduces to

$$I_n^{(\alpha)}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = (2^{\alpha-1} - 1)^{-1}\,(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1)\,; \quad \alpha \neq 0\,.$$

Now if $\alpha \to 1$, we get

$$I_n(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = \sum_{i=1}^{n} p_i \log_2 \frac{p_i}{q_i}$$

which is Kullback's [4] relative information, which is characterized by Campbell [1], Hobson [2], Kannappan [3], Rathie and Kannappan [5], Sharma and Ram Autar [7], Ng [11].

Case II. Measure (1.3) under the condition

$$I_2(1, 0;\ \tfrac{1}{2}, \tfrac{1}{2}) = 2^{\alpha-1} - 1$$

reduces to

$$I_n^{\alpha}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = [\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1]\,.$$

If $\alpha = 2$, we get

$$I_n^{(2)}(p_1, \ldots, p_n;\ q_1, \ldots, q_n) = \sum_{i=1}^{n} (p_i^2/q_i) - 1$$

which is Pearson's χ^2-statistic and is a measure of discrepancy between the two discrete populations P and Q.

ACKNOWLEDGEMENT

The author is grateful to Dr. I. Vajda, Czechoslovak Academy of Sciences, Institute of Information Theory and Automation, Czechoslovakia, for making some suggestions for improvement of the paper.

(Received January 3, 1978.)

REFERENCES

[1] L. L. Campbell: Characterization of entropy of probability distributions on real lines. Information and control *21* (1972), 329—338.

[2] A. Hobson: A new theorem of information theory. J. Stat. Phys. *1* (1969), 383—391.

[3] P. L. Kannappan: On Shannon entropy, directed divergence and inaccuracy. Z. Wahrs. verw. Geb. *22* (1972), 95—100.

[4] S. Kullback: Information theory and statistics. Dover Publications, Inc., New York (1959).

[5] P. N. Rathie, P. L. Kannappan: A directed divergence function of type β. Information and Control *20* (1972), 38—45.

[6] A. Rényi: On measures of entropy and informations. Proc. 4th Berkley Symposium on Probability and Stat., Berkley 1961, Vol. 1, 547—561.

[7] B. D. Sharma, R. Autar: Relative information functions and their type (α, β). Generalizations. Metrika *21* (1973).

[8] B. D. Sharma, I. J. Taneja: On axiomatic characterization of information-theoretic measures. J. Statist. Phys. *10* (1974), 4, 337—346.

[9] I. Vajda: On the amount of information contained in a sequence of independent observations. Kybernetika *6* (1970), 4, 306—324.

[10] I. Vajda: Limit theorem for total variation of Cartesian product measures. Studia Scientiarum Mathematicarum Hungarica *6* (1971), 317—333.

[11] C. T. Ng: Representation for measures of information with the branching property. Information and Control *25* (1974), 45—56.

Prof. A. Basar Khan, Department of Statistics, A.M.U., Aligarh, India — 202001.