Kybernetika

Om Parkash; Radhey S. Singh

On characterization of useful information-theoretic measures

Kybernetika, Vol. 23 (1987), No. 3, 245--251

Persistent URL: http://dml.cz/dmlcz/125603

KYBERNETIKA — VOLUME 23 (1987), NUMBER 3

ON CHARACTERIZATION OF USEFUL INFORMATION-THEORETIC MEASURES

OM PARKASH, R. S. SINGH

A characterization of the unified measure associated with a pair of probability distributions and a utility distribution, under a set of axioms has been provided. An interesting aspect is that under suitable additional boundary conditions, this unified measure gives rise to two useful information-theoretic quantities which lead to Kullback's information and Kerridge's inaccuracy concepts.

1. INTRODUCTION

Let $P = (p_1, p_2, \ldots, p_n)$, $0 < p_i \leqq 1$, $\sum_{i=1}^{n} p_i = 1$, be a finite discrete probability distribution of a set of n events $E = (E_1, E_2, \ldots, E_n)$ on the basis of an experiment whose predicted probability distribution is $Q = (q_1, q_2, \ldots, q_n)$, $0 < q_i \leqq 1, \sum_{i=1}^{n} q_i = 1$.

There are two information-theoretic measures associated with a pair of probability distributions which are of great significance in Statistical estimation and Physics. One of these two measures is the measure of information known as Kullback's information or directed divergence [3] given by

$$I_n[P; Q] = \sum_{i=1}^{n} p_i \log (p_i/q_i) , \tag{1.1}$$

and the other is Kerridge's inaccuracy [2] given by

$$I_n[P; Q] = - \sum_{i=1}^{n} p_i \log q_i \tag{1.2}$$

Now we attach a utility distribution $U = (u_1, u_2, \ldots, u_n)$ to the random experiment $E = (E_1, E_2, \ldots, E_n)$, where $u_i > 0$ is the utility of the ith outcome E_i.

Thus we have two utility information schemes:

$$(1.3) \qquad S = \begin{bmatrix} E_1 & E_2 & \dots & E_n \\ p_1 & p_2 & \dots & p_n \\ u_1 & u_2 & \dots & u_n \end{bmatrix}, \quad p_i, u_i > 0, \quad \sum_{i=1}^{n} p_i = 1,$$

of a set of n events after an experiment, and

$$(1.4) \qquad S^* = \begin{bmatrix} E_1 & E_2 & \dots & E_n \\ q_1 & q_2 & \dots & q_n \\ u_1 & u_2 & \dots & u_n \end{bmatrix}, \quad q_i, u_i > 0, \quad \sum_{i=1}^{n} q_i = 1,$$

of the same set of n events before the experiment.

In both the schemes (1.3) and (1.4) the utility distribution is same, because we assume that the utility u_i of an outcome E_i is independent of its probability of occurrence p_i or predicted probability q_i; u_i is only a 'utility' or 'value' of the outcome E_i for an observer relative to some specified goal.

After attaching the utility distribution, Taneja and Tuteja [5], characterized a measure corresponding to (1.1), given by

$$(1.5) \qquad I_n[P; Q; U] = \sum_{i=1}^{n} u_i p_i \log (p_i/q_i).$$

A similar type of quantitative-qualitative measure corresponding to (1.2), has been characterized by Taneja and Tuteja [6] given by

$$(1.6) \qquad I_n[P; Q; U] = - \sum_{i=1}^{n} u_i p_i \log q_i.$$

The object of this paper is to characterize a measure which jointly contains (1.5) and (1.6). Also by imposing certain conditions on this measure, we obtain these two measures separately and further on ignoring the utility distribution, we obtain Kullback's measure [3] and Kerridge's inaccuracy [2].

In what follows we shall assume that $0 \log 0 = 0 \log (0/0) = 0$ and all logarithms are considered to the base 2.

2. AXIOMS FOR QUANTITATIVE-QUALITATIVE MEASURES OF INFORMATION

Let $I_n[p_1, p_2, \dots, p_n; q_1, q_2, \dots, q_n; u_1, u_2, \dots, u_n]$ be the quantitative-qualitative measure of information associated with the goal oriented experiment $E = (E_1, E_2, \dots, E_n)$. In order to characterize the $I_n[P; Q; U]$ function, we consider the following three axioms:

Axiom I. The function $I_n[P; Q; U]$ is continuous with respect to its arguments p_i's, q_i's and u_i's.

Axiom II. (Branching Property.) The function $I_n[P; Q; U]$ satisfies the following:

$$I_n[p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n; u_1, u_2, \ldots, u_n] =$$

$$= I_{n-1}\left[p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n; \frac{u_1 p_1 + u_2 p_2}{p_1 + p_2}, u_3, \ldots, u_n\right] +$$

$$+ (p_1 + p_2) I_2\left[\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}; \frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2}; u_1, u_2\right]$$

Axiom III. The quantitative-qualitative measure of information provided by an outcome E_i is proportional to its utility u_i, i.e. for each non-negative λ, the following holds:

$$I[p_i; q_i; \lambda u_i] = \lambda I[p_i; q_i; u_i] .$$

Now before proving the main result, we give some results as lemmas based on the above axioms:

Lemma 1. If

$$v_k \geqq 0 , \quad k = 1, 2, \ldots, m_i , \quad \sum_{k=1}^{m_i} v_k = p_i > 0 ; \quad h_k \geqq 0 , \quad k = 1, 2, \ldots, m_i ,$$

$$\sum_{k=1}^{m_i} h_k = q_i > 0 ;$$

and

$$r_k \geqq 0 , \quad k = 1, 2, \ldots, m_i , \quad \sum_{k=1}^{m_i} \frac{r_k v_k}{\sum_{k=1}^{m_i} v_k} = u_i > 0 , \quad \text{for every} \quad i = 1, 2, \ldots, n ,$$

then

(2.1)
$$I_{m_i+n-1}[p_1, p_2, \ldots, p_{i-1}, v_1, v_2, \ldots, v_{m_i}, p_{i+1}, \ldots, p_n ;$$
$$q_1, q_2, \ldots, q_{i-1}, h_1, h_2, \ldots, h_{m_i}, q_{i+1}, \ldots, q_n ;$$
$$u_1, u_2, \ldots, u_{i-1}, r_1, r_2, \ldots, r_{m_i}, u_{i+1}, \ldots, u_n] = I_n[P; Q; U] +$$
$$+ p_i I_{m_i}\left[\frac{v_1}{p_i}, \frac{v_2}{p_i}, \ldots, \frac{v_{m_i}}{p_i}; \frac{h_1}{q_i}, \frac{h_2}{q_i}, \ldots, \frac{h_{m_i}}{q_i}; r_1, r_2, \ldots, r_{m_i}\right]$$

Proof. We shall prove the lemma by induction. For $m_i = 2$, (2.1) reduces to Axiom II i.e. our lemma is true for $m_i = 2$.

Now applying (2.1) for m_i in I_{m_i+n}, we get

(2.2)
$$I_{m_i+n}[p_1, p_2, \ldots, p_{i-1}, v_1, v_2, \ldots, v_{m_i+1}, p_{i+1}, \ldots, p_n ;$$
$$q_1, q_2, \ldots, q_{i-1}, h_1, h_2, \ldots, h_{m_i+1}, q_{i+1}, \ldots, q_n ;$$
$$u_1, u_2, \ldots, u_{i-1}, r_1, r_2, \ldots, r_{m_i+1}, u_{i+1}, \ldots, u_n]$$

$$= I_{n+1}[p_1, p_2, \ldots, p_{i-1}, v_1, \bar{p}, p_{i+1}, \ldots, p_n;$$

$$q_1, q_2, \ldots; q_{i-1}, h_1, \bar{q}, q_{i+1}, \ldots, q_n; u_1, u_2, \ldots, u_{i-1}, r_1, \bar{u}, u_{i+1}, \ldots, u_n]$$

$$+ \bar{p}\, I_{m_i}\left[\frac{v_2}{\bar{p}}, \ldots, \frac{v_{m_i+1}}{\bar{p}}; \frac{h_2}{\bar{q}}, \ldots, \frac{h_{m_i+1}}{\bar{q}}; r_2, \ldots, r_{m_i+1}\right]$$

$$(2.3) \qquad = I_n[P; Q; U] + p_i I_2\left[\frac{v_1}{p_i}, \frac{\bar{p}}{p_i}; \frac{h_1}{q_i}, \frac{\bar{q}}{q_i}; r_1, \bar{u}\right] +$$

$$+ \bar{p}\, I_{m_i}\left[\frac{v_2}{\bar{p}}, \ldots, \frac{v_{m_i+1}}{\bar{p}}; \frac{h_2}{\bar{q}}, \ldots, \frac{h_{m_i+1}}{\bar{q}}; r_2, \ldots, r_{m_i+1}\right]$$

(Using Axiom II in (2.2)) where

$$\bar{p} = (v_2 + v_3 + \ldots + v_{m_i+1}), \quad \bar{q} = (h_2 + h_3 + \ldots + h_{m_i+1})$$

and

$$\bar{u} = \frac{(r_2 v_2 + r_3 v_3 + \ldots + r_{m_i+1} v_{m_i+1})}{(v_2 + v_3 + \ldots + v_{m_i+1})}.$$

Now for $n = 2$, Axiom II is

$$(2.4) \qquad I_{m_i+1}\left[\frac{v_1}{p_i}, \ldots, \frac{v_{m_i+1}}{p_i}; \frac{h_1}{q_i}, \ldots, \frac{h_{m_i+1}}{q_i}; r_1, \ldots, r_{m_i+1}\right] =$$

$$= I_2\left[\frac{v_1}{p_i}, \frac{\bar{p}}{p_i}; \frac{h_1}{q_i}, \frac{\bar{q}}{q_i}; r_1, \bar{u}\right]$$

$$+ \left(\frac{\bar{p}}{p_i}\right) I_{m_i}\left[\frac{v_2}{\bar{p}}, \ldots, \frac{v_{m_i+1}}{\bar{p}}; \frac{h_2}{\bar{q}}, \ldots, \frac{h_{m_i+1}}{\bar{q}}; r_2, \ldots, r_{m_i+1}\right]$$

Using (2.4) in (2.3), we see that the result of the lemma is true for $(m_i + 1)$. Hence by induction, lemma follows. □

The above lemma can be extended easily in the following form:

Lemma 2. If

$$v_{ij} \geqq 0, \quad j = 1, 2, \ldots, m_i, \quad \sum_{j=1}^{m_i} v_{ij} = p_i > 0, \quad \sum_{i=1}^{n} p_i = 1 \quad \text{and} \quad h_{ij} \geqq 0,$$
$$j = 1, 2, \ldots, m_i,$$

$$\sum_{j=1}^{m_i} h_{ij} = q_i > 0, \quad \sum_{i=1}^{n} q_i = 1 \quad \text{and} \quad r_{ij} \geqq 0, \quad j = 1, 2, \ldots, m_i,$$

$$\frac{\sum_{j=1}^{m_i} r_{ij} v_{ij}}{\sum_{j=1}^{m_i} v_{ij}} = u_i > 0, \quad \text{for every} \quad i = 1, 2, \ldots, n,$$

then

$$I_{nm_n}[V; H; R] = I_n[P; Q; U] + \tag{2.5}$$

$$+ \sum_{i=1}^{n} p_i I_{m_i}\left[\frac{v_{i1}}{p_i}, \ldots, \frac{v_{im_i}}{p_i}; \frac{h_{i1}}{q_i}, \ldots, \frac{h_{im_i}}{q_i}; r_{i1}, \ldots, r_{im_i}\right]$$

Now we come to the main result of this paper.

Theorem 1. The function $I_n[P; Q; U]$ satisfying Axiom I–III determine the function I_n as

$$I_n[P; Q; U] = A\sum_{i=1}^{n} u_i p_i \log p_i + B\sum_{i=1}^{n} u_i p_i \log q_i\,, \tag{2.6}$$

where A and B are arbitrary constants.

Proof. In Lemma 1, if we replace m_i by m, and substitute

$$v_{ij} = 1/mn\,, \quad h_{ij} = 1/rs\,, \quad r_{ij} = 1$$

and

$p_i = 1/m\,, \quad q_i = 1/r\,, \quad u_i = 1\,,$ for every $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$

where m, n, r, s are positive integers such that $1 \leqq m \leqq r$, $1 \leqq n \leqq s$, then we obtain

$$F[mn; rs; 1] = F[m; r; 1] + F[n; s; 1] \tag{2.7}$$

where

$$F[m; r; 1] = I[1/m, \ldots, 1/m; 1/r, \ldots, 1/r; 1, \ldots, 1] \tag{2.8}$$

Now (2.7) is Cauchy's functional equation in two variables and its most general bounded solution ([1], Chapter 5), is given by

$$F[m; r; 1] = A' \log m + B' \log r \tag{2.9}$$

where A' and B' are arbitrary constants.

Now we prove Theorem 1 for rationals and the continuity of I_n proves the result for reals.

If m, r, r_i and t_i are positive integers such that $\sum_{i=1}^{n} r_i = m$, $\sum_{i=1}^{n} t_i = r$ and if we put $v_{ij} = 1/m$, $h_{ij} = 1/r$, $r_{ij} = 1$, and $p_i = r_i/m$, $q_i = t_i/r$, $u_i = 1$, for every $i = 1, 2, \ldots$ $\ldots, n$, then an application of Lemma 2, gives

$$I[1/m, \ldots, 1/m; 1/r, \ldots, 1/r; 1, \ldots, 1] = I_n[P; Q; 1] + \tag{2.10}$$

$$+ \sum_{i=1}^{n} p_i I[1/r_i, \ldots, 1/r_i; 1/t_i, \ldots, 1/t_i; 1, \ldots, 1]$$

or

$$F[m; r; 1] = I_n[P; Q; 1] + \sum_{i=1}^{n} p_i F[r_i; t_i; 1] \tag{2.11}$$

Using (2.9), (2.11) gives

$$I_n[P; Q; 1] = (A' \log m + B' \log r) - \sum_{i=1}^{n} p_i(A' \log r_i + B' \log t_i) \tag{2.12}$$

Since $\sum_{i=1}^{n} p_i = 1$, we have

$$(2.13) \qquad I_n[P; Q; 1] = A \sum_{i=1}^{n} p_i \log p_i + B \sum_{i=1}^{n} p_i \log q_i$$

where $A = -A'$ and $B = -B'$, are arbitrary constants.

Now in Axiom III, setting $u_i = 1$ and $\lambda = u_i$, for each i, we get

$$(2.14) \qquad I[p_i; q_i; u_i] = u_i I[p_i, q_i; 1]$$

Using (2.14) in (2.13), we get (2.6). □

On ignoring the utility i.e. taking $u_i = 1$ for every i, we get

$$I_n[P; Q] = A \sum_{i=1}^{n} p_i \log p_i + B \sum_{i=1}^{n} p_i \log q_i ,$$

which is an information-theoretic quantity associated with a pair of probability distributions characterized by Sharma and Taneja [4].

3. APPLICATIONS TO INFORMATION THEORY

As remarked earlier, Kullback's information and Kerridge's innaccuracy are two information-theoretic measures which are particular cases of the results studied by Taneja and Tuteja [5], [6] and their characterizations are given below:

Theorem 2. The function $I_n[P; Q; U]$ under Axioms I–III and satisfying

$$(3.1) \qquad I_2[P; P; U] = 0 , \quad p \in (0, 1) \quad \text{and} \quad u > 0$$

and

$$(3.2) \qquad I_2[1, 0; \tfrac{1}{2}, \tfrac{1}{2}; 1, 1] = 1$$

is given by

$$(3.3) \qquad I_n[P; Q; U] = \sum_{i=1}^{n} u_i p_i \log (p_i/q_i)$$

Proof. Using (3.1) in (2.6), we get $A + B = 0$.

Also using (3.2), (2.6) gives $A = 1$ and $B = -1$. Substituting these values of A and B in (2.6), we get (3.3), which is a result studied by Taneja and Tuteja [5]. Further on ignoring the utility (3.3) gives Kullback's information [3]. □

Theorem 3. The function $I_n[P; Q; U]$ under Axioms I–III and satisfying

$$(3.4) \qquad I_3[p_1, p_2, p_3; q_1, q_2, q_2; u_1, u_2, u_3] = = I_2\left[p_1, p_2 + p_3; q_1, q_2; u_1, \frac{u_2 p_2 + u_3 p_3}{p_2 + p_3}\right]$$

and

$$(3.5) \qquad I_2[\tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2}; 1, 1] = 1 ,$$

is given by

$$I_n[P; Q; U] = -\sum_{i=1}^{n} u_i p_i \log q_i \tag{3.6}$$

Proof. Using (3.4) and (3.5) in (2.6), we get $A = 0$ and $B = -1$. Thus (2.6) reduces to (3.6), which is a result studied by Taneja and Tuteja [6].

Further on ignoring the utility, (3.6) gives Kerridge's inaccuracy [2]. □

ACKNOWLEDGEMENTS

The authors are thankful to Dr. H. C. Taneja and Prof. R. K. Tuteja, Department of Mathematics, M. D. University, Rohtak and Prof. T. D. Narang, Department of Mathematics G. N. D. University, Amritsar, for their constant encouragements.

REFERENCES

[1] J. Aczel: Lectures on Functional Equations and Their Applications. Academic Press, New York 1966.

[2] D. F. Kerridge: Inaccuracy and inference. J. Roy. Statist. Soc. Ser B *23* (1961), 184—194.

[3] S. Kullback: Information Theory and Statistics. Wiley, New York 1959.

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

[5] H. C. Taneja and R. K. Tuteja: Characterization of quantitative-qualitative measure of relative information. Inform. Sci. *33* (1984), 217—222.

[6] H. C. Taneja and R. K. Tuteja: Characterization of a quantitative-qualitative measure of inaccuracy. Kybernetika *22* (1985), 393—402.

Dr. Om Parkash, Department of Mathematics, Guru Nanak Dev University, Amritsar-143005. India.

Dr. R. S. Singh, Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1.