Journal of Statistical Physics, Vol. 14, No. 3, 1976

On Measures of Information and Inaccuracy

Inder Jeet Taneja[1]

Received December 12, 1974; revised August 14, 1975

The Kullback relative-information measure and Kerridge's inaccuracy measure and their generalized forms are consequences of different forms of the branching property that these measures are required to satisfy. We consider a seemingly more generalized form and show that it does not lead to new measures. We also form a functional equation in two variables through this generalized branching property and show that this leads to the same result.

KEY WORDS: Information theory; branching property; relative-information measure; inaccuracy measure; statistics.

1. INTRODUCTION

Sharma and Taneja[11] axiomatically characterized the measures

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

and

$$I_n^{(\alpha,\beta)}(P; Q) = C\left[\sum_{i=1}^{n} p_i^{\alpha} q_i^{\beta} - 1\right], \qquad \alpha > 0 \tag{2}$$

(A, B, and C are arbitrary constants and α and β are parameters such that $\alpha \neq 1$ when $\beta = 0$) corresponding to the probability distributions $P = (p_1,\ldots, p_n)$, $p_i \geqslant 0$, $\sum_{i=1}^{n} p_i = 1$, and $Q = (q_1,\ldots, q_n)$, $q_i > 0$, $\sum_{i=1}^{n} q_i \leqslant 1$,

---

[1] Senior Research Fellow (CSIR), Department of Mathematics, University of Delhi, Delhi, India.

associated with a discrete random variable X assuming finite number of values $\xi_1,\ldots,\xi_n$.

Measures (1) and (2) jointly have also been characterized by Taneja[13] by a generalized functional equation having four different functions.

Measures (1) and (2) under certain boundary conditions reduce to Kullback's[7] relative-information measure and Kerridge's[6] inaccuracy measure [see expressions (27) and (31), respectively] and their generalized forms given in (28) and (32), respectively. Thus the measures studied by Kullback and Kerridge, which have many uses in information theory, statistics, physics, economics, etc., and their respective generalized forms are included in (1) and (2).

These measures arise mainly due to a branching property which for (2) may be written as

$$I_n^{(\alpha,\beta)}(P;\, Q) - I_{n-1}^{(\alpha,\beta)}(p_1 + p_2, p_3,\ldots, p_n;\, q_1 + q_2, q_3,\ldots, q_n)$$
$$= p_i^{\alpha} q_i^{\beta} I_2^{(\alpha,\beta)}(\cdots) \tag{3}$$

where $p_i = p_1 + p_2 > 0$; $q_i = q_1 + q_2 > 0$.

The relation (3) when $\beta = 0$ and $\alpha = 1$ gives rise to a different case which leads to the measure (1).

In this communication, we start with a seemingly more generalized form of the branching property, taking a general continuous function $f(p_i\,;\, q_i)$ in place of the $p_i^{\alpha} q_i^{\beta}$ that occurs in (3). It is established that such a change does not give new measures and that (1) and (2) cover all the measures that can be so obtained. In fact, $f(p_i\,;\, q_i) = p_i^{\alpha} q_i^{\beta}$ is the most general form compatible with the generalized form of (3), provided we impose constraints of symmetry and continuity.

Remarks. In what follows we shall take $0 \log 0 = 0 \log (0/q_i) = 0$ for all $i = 1, 2,\ldots, n$ and all the logarithms are considered to the base 2.

2. CHARACTERIZATION THEOREM

Let $I_n^f(P;\, Q)$ be an information-theoretic measure associated with a pair of probability distributions $P = (p_1,\ldots, p_n)$, $p_i \geqslant 0$, $\sum_{i=1}^{n} p_i = 1$, and $Q = (q_1,\ldots, q_n)$, $q_i > 0$, $\sum_{i=1}^{n} q_i \leqslant 1$ of a discrete random variable. We consider that the function $I_n^f(P;\, Q)$ satisfies the following axioms:

(I) (Continuity). $I_n^f(P;\, Q)$ is a continuous function of its arguments.

(II) (Symmetry). $I_n^f(P;\, Q)$ is symmetric for any permutation of elements in P followed by the same permutation of elements in Q.

(III) (Generalized branching property).

$$I^f_{n+m-1}(p_1,\ldots,p_{i-1},v_1,\ldots,v_m,p_{i+1},\ldots,p_n;\\ \times q_1,\ldots,q_{i-1},h_1,\ldots,h_m,q_{i+1},\ldots,q_n)\\ = I_n^f(P;Q) + f(p_i;q_i)I_m^f(v_1/p_i,\ldots,v_m/p_i;h_1/q_i,\ldots,h_m/q_i)$$

where $v_k \geqslant 0$, $\sum_{k=1}^m v_k = p_i > 0$; $h_k > 0$, $\sum_{k=1}^m k_k = q_i > 0$ for every $i = 1, 2,\ldots, n$; and f is any continuous function defined in $[0, 1] \times (0, 1]$ such that $f(0; q) = 0$.

Theorem. The function $I_n^f(P; Q)$ determined by the axioms (I)–(III) can be only of the form (1) or (2).

Before proving the theorem, we give some intermediate results based on above axioms in the following lemmas:

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

$$I^f_{m_1+m_2+\cdots m_n}(v_{11},\ldots,v_{1m_1},\ldots,v_{n1},\ldots,v_{nm_n};h_{11},\ldots,h_{1m_1},\ldots,h_{n1},\ldots,h_{nm_n})\\ = I_n^f(P;Q) + \sum_{i=1}^n f(p_i;q_i)I^f_{m_i}(v_{i1}/p_i,\ldots,v_{im_i}/p_i;h_{i1}/q_i,\ldots,h_{im_i}/q_i) \qquad (4)$$

This lemma directly follows from axiom (III).

Lemma 2. If $F(m; r) = I_m^f(1/m,\ldots, 1/m; 1/r,\ldots, 1/r)$, then

$$F(m;r) = A'\log m + B'\log r \qquad \text{when} \quad f(1/m;1/r) \neq 1/m \qquad (5)$$

or

$$F(m;r) = C[mf(1/m;1/r) - 1] \qquad \text{when} \quad f(1/m;1/r) \neq 1/m \qquad (6)$$

where

$$f(1/mn;1/rs) = f(1/m;1/r)f(1/n;1/s) \qquad (7)$$

m, n, r, and s are arbitrary positive integers such that $1 \leqslant m \leqslant r$, $1 \leqslant n \leqslant s$; and A', B', and C are arbitrary constants.

Proof. In Lemma 1 replace m_i by m, set $v_{ij} = 1/mn$, $h_{ij} = 1/rs$, $q_i = 1/s$, $i = 1, 2,\ldots, n$; $j = 1, 2,\ldots, m$; where m, n, r, and s are positive integers such that $1 \leqslant m \leqslant r$, $1 \leqslant n \leqslant s$; then we obtain

$$F(mn;rs) = F(m;r) + mf(1/m;1/r)F(n;s) \qquad (8)$$

Now there are two cases to consider.

Case I. When $f(1/m; 1/r) = 1/m$. In this case (8) reduces to

$$F(mn;rs) = F(m;r) + F(n;s) \qquad (9)$$

The continuous solution of this number-theoretic functional equation (refer to Aczél[1]) is given by (5).

Case II. When $f(1/m;\, 1/r) \neq 1/m$. In this case symmetry in m,n and r,s implies

$$F(mn;\, rs) = F(nm;\, sr)$$

i.e.,

$$F(m;\, r) + mf(1/m;\, 1/r)F(n;\, s) = F(n;\, s) + nf(1/n;\, 1/s)F(m;\, r)$$

or

$$\frac{F(m;\, r)}{mf(1/m;\, 1/r) - 1} = \frac{F(n;\, s)}{nf(1/n;\, 1/s) - 1} = C \text{ (say)} \tag{10}$$

provided $f(1/m;\, 1/r) \neq 1/m$.

Thus, expression (10) gives

$$F(m;\, r) = C[mf(1/m;\, 1/r) - 1] \qquad \text{if} \quad f(1/m;\, 1/r) \neq 1/m$$

where C is any arbitrary constant.

Now substituting (6) in (8), we get (7).

Lemma 3. The function f in axiom (III) is such that it satisfies a functional equation

$$f(pu;\, qv) = f(p;\, q)f(u;\, v) \tag{11}$$

for all reals $p, u \in [0, 1]$ and $q, v \in (0, 1]$.

Proof. From axiom (III), we may write

$$\begin{aligned}
&I^f_{n+m-1}(p_1,\ldots,p_{i-1}, v_1,\ldots,v_m, p_{i+1},\ldots,p_n; q_1,\ldots,q_{i-1}, h_1,\ldots,h_m, q_{i+1},\ldots,q_n)\\
&\quad = I^f_{n+1}(p_1,\ldots,p_{i-1}, v_1, \bar{p}, p_{i+1},\ldots,p_n; q_1,\ldots,q_{i-1}, h_1, \bar{q}, q_{i+1},\ldots,q_n)\\
&\qquad + f(\bar{p};\bar{q})I^f_{m-1}(v_2/\bar{p},\ldots, v_m/\bar{p}; h_2/\bar{q},\ldots, h_m/\bar{q})\\
&\qquad\qquad \text{where} \quad \bar{p} = v_2 + \ldots + v_m > 0; \quad \bar{q} = h_2 + \ldots + h_m > 0\\
&\quad = I_n{}^f(P;\, Q) + f(p_i;q_i)I_2{}^f(v_1/p_i; \bar{p}/p_i; h_1/q_i, \bar{q}/q_i)\\
&\qquad + f(\bar{p};\bar{q})I^f_{m-1}(v_2/\bar{p},\ldots, v_m/\bar{p}; h_2/\bar{q},\ldots, h_m/\bar{q})\\
&\qquad\qquad \text{where} \quad p_i = v_1 + \bar{p} = v_1 + \ldots + v_m; \quad q_i = h_1 + \bar{q} = h_1 + \ldots + h_m
\end{aligned} \tag{12}$$

Alternatively, we can write, again from axiom (III),

$$\begin{aligned}
&I^f_{n+m-1}(p_1,\ldots,p_{i-1}, v_1,\ldots,v_m, p_{i+1},\ldots,p_n; q_1,\ldots,q_{i-1}, h_1,\ldots,h_m, q_{i+1},\ldots,q_n)\\
&\quad = I_n{}^f(P;\, Q) + f(p_i;q_i)I_m{}^f(v_1/p_i,\ldots, v_m/p_i; h_1/q_i,\ldots, h_m/q_i)\\
&\quad = I_n{}^f(P;\, Q) + f(p_i\,;q_i)\{I_2{}^f(v_1/p_i, \bar{p}/p_i; h_1/q_i, \bar{q}/q_i)\\
&\qquad + f(\bar{p}/p_i; \bar{q}/q_i)I^f_{m-1}(v_2/\bar{p},\ldots, v_m/\bar{p}; h_2/\bar{q},\ldots, h_m/\bar{q})\}\\
&\quad = I_n{}^f(P;\, Q) + f(p_i;q_i)I_2{}^f(v_1/p_i, \bar{p}/p_i; h_1/q_i, \bar{q}/q_i)\\
&\qquad + f(p_i;q_i)f(\bar{p}/p_i; \bar{q}/q_i)I^f_{m-1}(v_2/\bar{p},\ldots, v_m/\bar{p}; h_2/\bar{q},\ldots, h_m/\bar{q})
\end{aligned} \tag{13}$$

Comparing (12) and (13), we get

$$f(\bar{p}/p_i; \bar{q}/q_i) = f(\bar{p}; \bar{q})/f(p_i; q_i) \qquad \text{if} \quad f(p_i; q_i) \neq 0 \tag{14}$$

Now (14) together with (6) and the continuity of the function f gives (11).

Proof of the Theorem. We prove the theorem for rationals and then the continuity axiom (I) gives the result for reals. For this let m, the x_i and the y_i be positive integers such that $x_i \leqslant y_i$ for every $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} x_i = m$ and if we put $p_i = x_i/m$, $q_i = y_i/r$, $i = 1, 2, \ldots, n$, where $\sum_{i=1}^{n} y_i \leqslant r$, then Lemmas 1 and 2 give

$$I_n^f(P; Q) = F(m; r) - \sum_{i=1}^{n} f(p_i; q_i) F(x_i; y_i) \tag{15}$$

Now (15) together with (5) gives (1).

Again (15) together with (11) and (6) gives

$$I_n^f(P; Q) = C\left[\sum_{i=1}^{n} f(p_i; q_i) - 1\right] \qquad \text{if} \quad f(p; q) \neq p \tag{16}$$

where C is any arbitrary constant and f satisfies the functional equation (11).

The most general nonzero continuous solution of the functional equation (11) in $[0, 1] \times (0, 1]$ (refer to Aczél[1]) is given by

$$f(p; q) = p^{\alpha} q^{\beta} \tag{17}$$

where α and β are arbitrary parameters.

The condition of continuity of the function $f(p; q)$ at $p = 0$ requires that $\alpha \geqslant 0$. But when $\alpha = 0$, we get from (17) that $f(p; q) = q^{\beta}$. This violates our condition $f(0; q) = 0$ [refer to axiom (III)]. Therefore $\alpha \neq 0$, i.e., $\alpha > 0$. Further, when $\alpha = 1$ and $\beta = 0$, we get from (17) that $f(p; q) = p$, which together with (15) and (5) gives the measure (1). This case has been discussed separately. Therefore, we have the solution of (11) in which $f(p; q) = p^{\alpha} q^{\beta}$, $\alpha > 0$, and $f(p; q) \neq p$.

3. A FUNCTIONAL EQUATION

Let us take

$$h(p; q) = I_2^f(p, 1 - p; q, 1 - q), 0 \leqslant p \leqslant 1, 0 < q < 1; \tag{18}$$

then from symmetry, we have

$$h(p; q) = h(1 - p; 1 - q) \tag{19}$$

Again, if we consider the branching property for $n = 3$, this leads to

$$h(p; q) + f(1 - p; 1 - q)h\left(\frac{u}{1 - p}; \frac{v}{1 - q}\right)$$

$$= h(u; v) + f(1 - u; 1 - v)h\left(\frac{p}{1 - u}; \frac{q}{1 - v}\right) \tag{20}$$

for all $p, u \in [0, 1)$; $q, v \in (0, 1)$; and $p + u \leqslant 1, q + v \leqslant 1$.

Next, using the branching property for any n as in Lemma 3, we have

$$I_n^f(P; Q) = \sum_{1=2}^{n} f(s_i; t_i)h(p_i/s_i; q_i/t_i) \tag{21}$$

where $s_i = p_1 + \cdots + p_i$; $t_i = q_1 + \cdots + q_i$; $i = 2, 3, \ldots, n$; and f satisfies a functional equation (refer to Lemma 3) given by

$$f(pu; qv) = f(p; q)f(u; v) \tag{22}$$

for all $p, u \in [0, 1]$ and $q, v \in (0, 1]$.

The functional equation (20) when $f(p; q) = p$ (refer to Kannappan and Ng[5]) has the general continuous solution given by

$$h(p; q) = A[p \log p + (1 - p) \log (1 - p)] + B[p \log q + (1 - p) \log (1 - q)] \quad \text{when} \quad f(p; q) = p \tag{23}$$

Again, when $f(p; q) \neq p$, the functional equation (20) (refer to Soni[12]) has the general continuous solution given by

$$h(p; q) = C[f(p; q) + f(1 - p; 1 - q) - 1] \quad \text{if} \quad f(p; q) \neq p \tag{24}$$

where f satisfies the functional equation (22) by Lemma 3.

Now (21) together with (23) gives (1); while (21) together with (24) gives (16), which under the general continuous solution (17) of the functional equation (22) reduces to (2). This gives another characterization of the measures (1) and (2).

4. PARTICULAR CASES

Case I. (Kullback's relative-information measure): Measures (1) and (2) under the conditions

$$I_2(p, 1 - p; p, 1 - p) = 0, \qquad p \in (0, 1) \tag{25}$$

and

$$I_2(1, 0; \tfrac{1}{2}, \tfrac{1}{2}) = 1 \tag{26}$$

reduce to

$$_1I_n(P; Q) = \sum_{i=1}^{n} p_i \log(p_i/q_i) \tag{27}$$

and

$$ {}_1I_n^\alpha(P;Q) = (2^{\alpha-1}-1)^{-1}\left(\sum_{i=1}^{n} p_i^\alpha q_i^{1-\alpha} - 1\right), \qquad \alpha \neq 1, \quad \alpha > 0 \tag{28} $$

respectively.

Expression (28) reduces to (27) when $\alpha \to 1$, which is Kullback's relative-information measure as characterized by many authors.[2-4,7,8,10]

Case II. (Kerridge's inaccuracy measure): Measures (1) and (2) under the conditions

$$ I_3(p_1, p_2, p_3; q_1, q_2, q_2) = I_2(p_1, p_2 + p_3; q_1, q_2) \tag{29} $$

and

$$ I_2(\tfrac{1}{2}, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2}) = 1 \tag{30} $$

reduce to

$$ {}_2I_n(P;Q) = -\sum_{i=1}^{n} p_i \log q_i \tag{31} $$

and

$$ {}_2I_n^\beta(P;Q) = (2^{-\beta}-1)^{-1}\left(\sum_{i=1}^{n} p_i q_i^\beta - 1\right), \qquad \beta \neq 0 \tag{32} $$

respectively.

Expression (32) reduces to (31) when $\beta \to 0$, which is Kerridge's inaccuracy measure as studied by many authors.[4,6,9]

ACKNOWLEDGMENTS

The author is grateful to CSIR (India) for providing a Senior Research Fellowship.

The author is thankful to Dr. Bhu Dev Sharma, Reader in Mathematics, University of Delhi, for guidance in the preparation of this paper and for discussions at various stages. Thanks are also due to referee for helpful remarks on an earlier version of the paper.

REFERENCES

1. J. Aczél, *Lectures on Functional Equations and ther Applications*, Academic Press, New York (1966).
2. L. L. Campbell, Characterization of Entropy of Probability Distributions on the Real Line, *Information and Control* **21**:329–338 (1972).
3. A. Hobson, A New Theorem of Information Theory, *J. Stat. Phys.* **1**:383–391 (1969).

4. Pl. Kannappan, On Shannon's Entropy, Directed-Divergence and Inaccuracy, *Z. Wahrs. Verw Geb.* **22**:95–100 (1972).
5. Pl. Kannappan and C. T. Ng, Measurable Solutions of Functional Equations Related to Information Theory, *Proc. Am. Math. Soc.* **38**:303–310 (1973).
6. D. F. Kerridge, Inaccuracy and Inference, *J. Royal Statist. Soc. Ser. B* **23**:184–194 (1961).
7. S. Kullback, *Information Theory and Statistics*, Wiley, New York (1959).
8. P. N. Rathie and Pl. Kannappan, A Directed-Divergence Function of Type β, *Information and Control* **20**:38–45 (1972).
9. B. D. Sharma and Ram Autar, On Characterization of a Generalized Inaccuracy Measure in Information Theory, *J. Appl. Prob.* **10**:464–468 (1973).
10. B. D. Sharma and Ram Autar, Relative-Information Functions and their Type (α, β) Generalizations, *Metrika* **21**:41–50 (1973).
11. B. D. Sharma and I. J. Taneja, On Axiomatic Characterization of Information-Theoretic Measures, *J. Stat. Phys.* **10**:337–346 (1974).
12. R. S. Soni, On a Functional Equation Connected with Information and Information-Improvement, *Indian J. Pure Appl. Math.* (to be published).
13. I. J. Taneja, A Joint Characterization of Directed-Divergence, Inaccuracy, and their Generalizations, *J. Stat. Phys.* **11**(2):169–176 (1974).