Journal of Statistical Physics, Vol. 11, No. 2, 1974

A Joint Characterization of Directed Divergence, Inaccuracy, and Their Generalizations

Inder Jeet Taneja[1]

Received December 21, 1973; revised April 12, 1974

Starting with an additive property for distributions of two statistically independent random variates in terms of different sum functions, we have characterized two general measures associated with two distributions of a discrete random variate. One of these measures is logarithmic, while the other contains powers of variables. An interesting aspect is that under suitable additional boundary conditions the logarithmic measure leads to measure of information (directed divergence) studied by Kullback and measure of inaccuracy studied by Kerridge, while the other solution leads to their parametric generalizations.

KEY WORDS: Directed divergence; inaccuracy; functional equation; additive measures; nonadditive measures; probability theory; information theory.

1. INTRODUCTION

Let $P = (p_1,..., p_n)$, $p_i \geqslant 0$, $\sum_{i=1}^{n} p_i = 1$, and $U = (u_1,..., u_n)$, $u_i > 0$, $\sum_{i=1}^{n} u_i \leqslant 1$, be two probability distributions associated with a discrete finite random variate X. Corresponding to the distributions P and U consider the

Supported by CSIR (India) research fellowship.

[1] Faculty of Mathematics, University of Delhi, Delhi, India.

measures

$$I(P/U) = A \sum_{i=1}^{n} p_i \log p_i + B \sum_{i=1}^{n} p_i \log u_i \tag{1}$$

and

$$I_{(\alpha,\beta)}(P/U) = C^{-1}\left(\sum_{i=1}^{n} p_i^{\alpha} u_i^{\beta} - 1\right) \tag{2}$$

where A, B, and C ($\neq 0$) are arbitrary constants and α and β are parameters.

Measures (1) and (2) under the conditions

$$I(\{1, 0\}/\{\tfrac{1}{2}, \tfrac{1}{2}\}) = 1, \qquad I(\{\tfrac{1}{2}, \tfrac{1}{2}\}/\{\tfrac{1}{2}, \tfrac{1}{2}\}) = 0 \tag{3}$$

reduce respectively to

$$_1I(P/U) = \sum_{i=1}^{n} p_i \log(p_i/u_i) \tag{4}$$

and

$$_1I_{\alpha}(P/U) = (2^{\alpha-1} - 1)^{-1}\left(\sum_{i=1}^{n} p_i^{\alpha} u_i^{1-\alpha} - 1\right), \qquad \alpha \neq 1 \tag{5}$$

When $\alpha \to 1$, (5) reduces to (4), which is Kullback's[7] directed divergence characterized by many authors.[2,4,5,7,10]

Again measures (1) and (2) under the conditions

$$I(\{1, 0\}/\{\tfrac{1}{2}, \tfrac{1}{2}\}) = 1, \qquad I(\{\tfrac{1}{2}, \tfrac{1}{2}\}/\{\tfrac{1}{2}, \tfrac{1}{2}\}) = 1 \tag{6}$$

reduce respectively to

$$_2I(P/U) = -\sum_{i=1}^{n} p_i \log u_i \tag{7}$$

and

$$_2I_{\beta}(P/U) = (2^{-\beta} - 1)^{-1}\left(\sum_{i=1}^{n} p_i u_i^{\beta} - 1\right), \qquad \beta \neq 0 \tag{8}$$

When $\beta \to 0$, (8) reduces to (7), which is Kerridge's[6] measure of inaccuracy characterized by many authors.[5,6,9,10]

Thus the measures studies by Kullback and Kerridge and their generalized forms are included in (1) and (2) respectively.

Moreover, (1) and (2) have the sum property

$$I(P/U) = \sum_{i=1}^{n} h(p_i, u_i) \tag{9}$$

where $h(p, u) = Ap \log p + Bp \log u$ and $h(p, u) = C^{-1}(p^{\alpha}u^{\beta} - p)$, respectively.

Further, measures (1) and (2) respectively have the properties

$$I(P^*Q/U^*V) = I(P/U) + I(Q/V) \tag{10}$$

and

$$I_{(\alpha,\beta)}(P^*Q/U^*V) = I_{(\alpha,\beta)}(P/U) + I_{(\alpha,\beta)}(Q/V) + CI_{(\alpha,\beta)}(P/U)I_{(\alpha,\beta)}(Q/V) \tag{11}$$

where for $P = (p_1, p_2,..., p_n)$, $\sum_{i=1}^{n} p_i = 1$; $Q = (q_1, q_2,..., q_m)$, $\sum_{j=1}^{m} q_j \leqslant 1$; $P^*Q = (p_1q_1,..., p_1q_m;...; p_nq_1,..., p_nq_m)$, etc.

The property (10) is contained in (11) for $C = 0$. Thus a joint study could be made from (9) and (11). But in (10) and (11) the measure associated with P^*Q and U^*V is expressed in terms of same function of P, U and Q, V. It is of some interest to express $I(P^*Q/U^*V)$ in terms of three different sum functions $R(P/U)$, $S(Q/V)$, and $T(P/U)$, satisfying the relation

$$I(P^*Q/U^*V) = R(P/U)S(Q/V) + T(P/U) \tag{12}$$

In this communication we shall characterize the measures through a functional equation arising from the relation (12); the functional equation so obtained is in a way a generalization of Chaundy and McLeod's[3] functional equation for two variables studied by Kannappan[5] for characterizing (4) and (7).

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

2. THE FUNCTIONAL EQUATION

Let f, g, h, and k: $[0, 1] \times (0, 1] \to R$ (reals) be continuous functions such that the functions R, S, I, and T, respectively, could be expressed in terms of them as in (9).

Thus (12) gives rise to following functional equation:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} h(p_iq_j, u_iv_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(p_i, u_i)g(q_j, v_j) + \sum_{i=1}^{n} k(p_i, u_i) \tag{13}$$

where $\sum_{i=1}^{n} p_i = \sum_{j=1}^{m} q_j = 1$, $\sum_{i=1}^{n} u_i \leqslant 1$, and $\sum_{j=1}^{m} v_j \leqslant 1$.

Now we will obtain all the continuous solutions of the functional equation (13) in the following theorem.

Theorem 2.1. The functional equation (13) admits of the following two sets of solutions.

First set of solutions:

$$h(p, u) = Lp + Ap \log p + Bp \log u \tag{14}$$

$$f(p, u) = Mp \tag{15}$$

$$g(p, u) = Np + (A/M)p \log p + (B/M)p \log u \tag{16}$$

$$k(p, u) = (L - MN)p + Ap \log p + Bp \log u \tag{17}$$

Second set of solutions

$$h(p, u) = Lp + \mu(p^\alpha u^\beta - p) \tag{18}$$

$$f(p, u) = Mp^\alpha u^\beta \tag{19}$$

$$g(p, u) = Np + (\mu/M)(p^\alpha u^\beta - p) \tag{20}$$

$$k(p, u) = (L - MN)p^\alpha u^\beta + (\mu - L)(p^\alpha u^\beta - p) \tag{21}$$

Here L, M, N, A, B, and μ $(\neq 0)$ are arbitrary constants and α and β are parameters $(\alpha \neq 1; \beta \neq 0, \alpha, \beta > 0)$.

Remarks. Solutions which would also arise are (i) $h(p, u) = Lp$, $f(p, u) = 0$, $g(p, u)$ arbitrary, and $k(p, u) = Lp$; and (ii) $h(p, u) = Lp$, $f(p, u)$ arbitrary, $g(p, u) = 0$, and $k(p, u) = Lp$. These have not been included in the statement of the theorem because of their triviality and futility.

Proof. Let m, n, r, s, x, y, a, b, z, and c be any positive integers such that $1 \leqslant m \leqslant n$, $1 \leqslant r \leqslant s$, $1 \leqslant x < y$, $1 \leqslant a < b$, $z \geqslant (n - m)y/(y - x)$, and $c \geqslant (s - r)b/(b - a)$.

Setting

$$p_1 = m/n, \quad p_2 = \cdots = p_{n-m+1} = 1/n; \quad q_1 = r/s, \quad q_2 = \cdots = q_{s-r+1} = 1/s$$

$$u_1 = x/y, \quad u_2 = \cdots = u_{n-m+1} = 1/z; \quad v_1 = a/b, \quad v_2 = \cdots = v_{s-r+1} = 1/c$$

in (13) (taking $n = n - m + 1$ and $m = s - r + 1$), we get

$$\begin{aligned} h(mr/ns, xa/yb) &+ (s - r)h(m/ns, x/yc) \\ &+ (n - m)h(r/ns, a/zb) + (n - m)(s - r)h(1/ns, 1/zc) \\ &= \{f(m/n, x/y) + (n - m)f(1/n, 1/z)\} \\ &\quad \times \{g(r/s, a/b) + (s - r)g(1/s, 1/c)\} \\ &\quad + k(m/n, x/y) + (n - m)k(1/n, 1/z) \end{aligned} \tag{22}$$

Now taking $r = a = 1$, $b = c$, $m = x = 1$, and $y = z$ in (22), we get

$$h(1/ns, 1/zc) = f(1/n, 1/z)g(1/s, 1/c) + (1/s)k(1/n, 1/z) \tag{23}$$

Again taking $r = a = 1$ and $b = c$ in (22) and using (23), we get

$$h(m/ns, x/yc) = f(m/n, x/y)g(1/s, 1/c) + (1/s)k(m/n, x/y) \tag{24}$$

Similarly putting $m = x = 1$ and $y = z$ in (22) and once again using (23), we get

$$h(r/ns, a/zb) = f(1/n, 1/z)g(r/s, a/b) + (r/s)k(1/n, 1/z) \tag{25}$$

Finally, (22) together with (23), (24), and (25) gives

$$h(mr/ns, xa/yb) = f(m/n, x/y)g(r/s, a/b) + (r/s)k(m/n, x/y)$$

i.e.,

$$h(pq, uv) = f(p, u)g(q, v) + qk(p, u) \tag{26}$$

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

From the continuity of f, g, h, and k it follows that (26) is valid for all real numbers $x, y \in [0, 1]$ and $u, v \in (0, 1]$.

Now setting $p = u = 1$ in (26), we get

$$h(q, v) = f(1, 1)g(q, v) + qk(1, 1) \tag{27}$$

Again setting $q = v = 1$ and then $p = u = 1$ in (26), we get

$$h(p, u) = f(p, u)g(1, 1) + k(p, u) \tag{28}$$

$$h(1, 1) = f(1, 1)g(1, 1) + k(1, 1) \tag{29}$$

Now when $f(1, 1) = 0$, (29) and (27) give

$$h(p, u) = h(1, 1)p \tag{30}$$

Thus in this case an arbitrary g and $k = h$ is a solution. Similarly by symmetry the other solution mentioned in the remarks follows.

Thus in this case $h(p, u)$ becomes a homogeneous linear function.

Next when $f(1, 1) \neq 0$, (26) together with (27)–(29) gives

$$h(pq, uv) = \frac{f(p, u)}{f(1, 1)}[h(q, v) - qh(1, 1)] + qh(p, u)$$

i.e.,

$$h_1(pq, uv) = \frac{f(p, u)}{f(1, 1)}h_1(q, v) + qh_1(p, u) \tag{31}$$

where $h_1(p, u) = h(p, u) - ph(1, 1)$.

Now because of symmetry we have

$$h_1(pq, uv) = h_1(qp, vu)$$

Putting $f_1(p, u) = f(p, u) - pf(1, 1)$, this gives

$$f_1(p, u)h_1(q, v) = f_1(q, v)h_1(p, u) \tag{32}$$

i.e.,

$$f_1(p, u) = \lambda h_1(p, u) \tag{33}$$

where λ is an arbitrary constant.

Now there arise three cases:

Case I. When $\lambda = 0$, (33) gives (15),

$$f_1(p, u) = 0, \quad \text{i.e.,} \quad f(p, u) = f(1, 1)p = Mp$$

In this case (31) becomes

$$h_1(pq, uv) = qh_1(p, u) + ph_1(q, v) \tag{34}$$

Dividing both sides of (34) by pq and setting $h_1(p, u)/p = \phi(p, u)$, we get

$$\phi(pq, uv) = \phi(p, u) + \phi(q, v) \tag{35}$$

The most general continuous solutions of (35) are[1]

$$\phi(p, u) = A \log p + B \log u \tag{36}$$

where A and B are arbitrary constants.

Equations (36) and (27)–(29) give the first set of solutions.

Case II. When $\lambda \neq 0$, $h_1(p, u) \not\equiv 0$, (31) together with (33) gives

$$h_1(pq, uv) = ph_1(q, v) + qh_1(p, u) + \mu^{-1}h_1(p, u)h_1(q, v) \tag{37}$$

where $\mu = f(1, 1)/\lambda$ $(\neq 0)$.

Now setting $p + \mu^{-1}h_1(p, u) = E(p, u)$ in (37), we get

$$E(pq, uv) = E(p, u)E(q, v) \tag{38}$$

The most general continuous solutions of (38) are given by

$$E(p, u) = p^\alpha u^\beta \qquad \text{and} \qquad E(p, u) = 0 \tag{39}$$

where α and β are arbitrary parameters.

Now the solution $E(p, u) = p^\alpha u^\beta$ in (39) gives (18),

$$h(p, u) = h(1, 1)p + \mu(p^\alpha u^\beta - p) = Lp + \mu(p^\alpha u^\beta - p)$$

The case $E(p, u) = 0$ is already covered.

Now (38) together with (27)–(29) and (33) gives (19)–(21).

Case III. When $h_1(p, u) \equiv 0$, i.e., $h(p, u) = h(1, 1)p$, in this case also h becomes a homogeneous linear function which is covered as a particular case in the solutions already obtained.

3. CHARACTERIZATION OF DIRECTED DIVERGENCE AND INACCURACY

Consider a discrete random variate X taking finite number of values $x_1,..., x_n$ and let there be two distributions $P = (p_1,..., p_n)$, $\sum_{i=1}^n p_i = 1$, $p_i \geqslant 0$, and $U = (u_1,..., u_n)$, $\sum_{i=1}^n u_i \leqslant 1$, $u_i > 0$, associated with it.

Corresponding to the solution $f(p, u)$ we shall associate an information-theoretic measure involving the distributions P and U given by

$$I(P/U) = \sum_{i=1}^{n} h(p_i, u_i) \tag{40}$$

under suitable boundary and normalizing conditions.

Theorem 3.1. (Characterization of Directed Divergence). The measures associated with the distributions P and U corresponding to the continuous solutions (14) and (18) of the functional equation (13) under the conditions

$$h(1, 1) = 0, \qquad h(1, \tfrac{1}{2}) = 1, \qquad h(\tfrac{1}{2}, \tfrac{1}{2}) = 0 \tag{41}$$

are given respectively by ${}_1I(P/U)$ and ${}_1I_\alpha(P/U)$.

Proof. Condition $h(1, 1) = 0$ in the solutions (14) and (18) gives $L = 0$. Thus solutions (14) together with $h(1, \frac{1}{2}) = 1$ and $h(\frac{1}{2}, \frac{1}{2}) = 0$ give $A = 1$, $B = -1$ and then ${}_1I(P/U)$ follows from (40).

Again solution (18) together with $h(1, \frac{1}{2}) = 1$ gives $\mu = (2^{-\beta} - 1)^{-1}$; then $h(\frac{1}{2}, \frac{1}{2}) = 0$ gives $\alpha + \beta = 1$ and thus ${}_1I_\alpha(P/U)$ follows from (40).

Quantity ${}_1I_\alpha(P/U)$ has also been studied by Rathie and Kannappan.[8]

Theorem 3.2. (Characterization of Inaccuracy). The measures associated with the distributions P and U corresponding to the continuous solutions (14) and (18) of the functional equation (13) under the conditions

$$h(1, 1) = 0, \qquad h(1, \tfrac{1}{2}) = 1, \qquad h(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{1}{2} \tag{42}$$

are given respectively by ${}_2I(P/U)$ and ${}_2I_\beta(P/U)$.

Proof. Condition $h(1, 1) = 0$ gives $L = 0$ for the solutions (14) and (18).

Thus solution (14) together with $h(1, \frac{1}{2}) = 1$ and $h(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$ gives $A = 0$ and $B = -1$ and then ${}_2I(P/U)$ follows from (40).

Again solution (18) together with $h(1, \frac{1}{2}) = 1$ gives $\mu = (2^{-\beta} - 1)^{-1}$; then $h(\frac{1}{2}, \frac{1}{2}) = \frac{1}{2}$ gives $\alpha = 1$ and thus ${}_2I_\beta(P/U)$ follows from (40).

ACKNOWLEDGMENTS

The author is thankful to Dr. B. D. 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 Prof. U. N. Singh, Dean, Faculty of Mathematics, University of Delhi for providing facilities of the department.

REFERENCES

1. J. Aczel, *Lectures on Functional Equations and their Applications*, Academic Press, New York (1966).
2. L. L. Campbell, *Information and Control* **21**:329–338 (1972).
3. T. W. Chaundy and J. B. McLeod, *Proc. Edin. Math. Soc. Edin. Math. Notes* **43**:7–8 (1960).
4. A. Hobson, *J. Stat. Phys.* **1**: 383–391 (1969).

5. Pl. Kannappan, *Z. Warhs. verw. Geb.* **22**:95–100 (1972).
6. D. F. Kerridge, *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, *Information and Control* **20**:38–45 (1972).
9. B. D. Sharma and Ram Autar, *J. Appl. Probability* **10**:464–468 (1973).
10. B. D. Sharma and I. J. Taneja, *J. Stat. Phys.* **10** (4) (1974).