Indian J. pure appl. Math., **11**(2) : 190–197, February 1980

ON GENERALIZED DIRECTED-DIVERGENCE IN INFORMATION THEORY

O. P. VINOCHA AND A. N. GOYAL

Department of Mathematics, University of Rajasthan, Jaipur 302004

(*Received* 18 *May* 1978; *after revision* 26 *November* 1979)

In this paper we have defined new generalized directed divergence for additive. A characterization theorem is proved by assuming a set of six postulates.

1. INTRODUCTION

We define a new generalized directed-divergence as

$$I_n\,[P_m] = \sum_{k=1}^{m-2} \sum_{i=1}^{n} p_{i,k} \log \frac{p_{i,k+1}}{p_{i,k+2}} \qquad \ldots(1.1)$$

where $P_k = [p_{1,k}, p_{2,k}, \ldots, p_{n,k}]$, $k = 1, 2, \ldots, m$; $p_{i,k} \geqslant 0$; $\sum_{i=1}^{n} p_{i,k} = 1$, be m finite discrete probability distribution.

Particular Case

If $m = 3$ and we write $p_{i,1} = p_i$; $p_{i,2} = q_i$; $p_{i,3} = r_i$ we obtain the definition of Theil (1967) who called it information improvement. Later on, for $m = 3$ and generalized probability distributions, it was interpreted as an error function by Nath (1968, 72), generalized directed divergence by Aczél and Nath (1972). It has also been studied by Kannappan and Rathie (1973), Sharma and Ram Autar (1974), Kannappan (1975), etc.

[*Note* : The convention $o \log o = 0$ is followed and logarithms will be to the base 2. Also, whenever $p_{i,k+1}$ or $p_{i,k+2}$ is zero, the corresponding $p_{i,k}$ is also zero.]

If $n = 2$, (1.1) takes the form

$$I_2 \begin{bmatrix} p_1, 1 - p_1 \\ \vdots \quad\quad \vdots \\ p_m, 1 - p_m \end{bmatrix} = \sum_{k=1}^{m-2} \left[p_k \log \frac{p_{k+1}}{p_{k+2}} + (1 - p_k) \log \frac{1 - p_{k+1}}{1 - p_{k+2}} \right]. \qquad \ldots(1.2)$$

For $(p_1, \ldots, p_m) \in D$; where $D = \underbrace{(0, 1) \times (0, 1) \times \ldots \times (0, 1)}_{m \text{ times}}$

$$\cup \{(0, x_2, \ldots, x_m)\} \cup \{(1, x'_2, \ldots, x'_m)\}$$

with $\quad x_2, \ldots, x_m \in [0, 1)$ and $x'_2, \ldots, x'_m \in (0, 1]$.

2. Postulates

(1) *Generalized Recursivity*

$$I_n\,[P_m] = I_{n-1}\begin{bmatrix} p_{1,1} + p_{2,1}, & p_{3,1}, & \ldots, & p_{n,1} \\ \vdots & \vdots & & \vdots \\ p_{1,m} + p_{2,m}, & p_{3,m}, & \ldots, & p_{n,m} \end{bmatrix}$$

$$+ \sum_{k=1}^{m-2} (p_{1,k} + p_{2,k})\, I_2 \begin{bmatrix} \dfrac{p_{1,k}}{p_{1,k} + p_{2,k}}, & \dfrac{p_{2,k}}{p_{1,k} + p_{2,k}} \\ \dfrac{p_{1,k+1}}{p_{1,k+1} + p_{2,k+1}}, & \dfrac{p_{2,k+1}}{p_{1,k+1} + p_{2,k+1}} \\ \dfrac{p_{1,k+2}}{p_{1,k+2} + p_{2,k+2}}, & \dfrac{p_{2,k+2}}{p_{1,k+2} + p_{2,k+2}} \end{bmatrix}$$

(2) *Symmetry*

$$I_3\begin{bmatrix} p_{1,1}, & p_{2,1}, & p_{3,1} \\ \vdots & \vdots & \vdots \\ p_{1,m}, & p_{2,m}, & p_{3,m} \end{bmatrix} = I_3\begin{bmatrix} p_{a,1}, & p_{b,1}, & p_{c,1} \\ \vdots & \vdots & \vdots \\ p_{a,m}, & p_{b,m}, & p_{c,m} \end{bmatrix}$$

where $\{a, b, c\}$ is an arbitrary permutation of $\{1, 2, 3\}$.

(3) *Derivative*

Let
$$f(p_1, \ldots, p_m) = I_2\begin{bmatrix} p_1, & 1 - p_1 \\ \vdots & \vdots \\ p_m, & 1 - p_m \end{bmatrix}$$

for all $(p_1, \ldots, p_m) \in D$ where D is as given in (1.2) and $k = 1, 2, \ldots, m$. Also let f have continuous first partial derivatives with respect to all the m variables

$$p_1, p_2, \ldots, p_m \in (0, 1).$$

(4) *Nullity*

$$f(p, p, \ldots, p) = 0 \quad \text{for} \quad p \in [0, 1].$$

(5) *Normalization*

$$I_2\begin{bmatrix} 2/3, & 1/3 \\ 2/3, & 1/3 \\ 1/3, & 2/3 \\ \vdots & \vdots \\ 1/3, & 2/3 \end{bmatrix} = 1/3 \quad \text{and} \quad I_2\begin{bmatrix} 2/3, & 1/3 \\ 1/3, & 2/3 \\ 1/3, & 2/3 \\ \vdots & \vdots \\ 1/3, & 2/3 \end{bmatrix} = 0.$$

(6) *Sum Representation*

$$\sum_{i=1}^{n} f(p_{i,1}, p_{i,2}, p_{i,3}, \ldots, p_{i,m}) = I_n [P_m].$$

3. Characterization Theorem

In this section we prove the following theorem with the help of the postulates given in section 2.

Theorem — The only function $I_n [P_m]$ satisfying the postulates (1) to (6) is the generalized directed divergence given by (1.1).

PROOF : The proof of the theorem depends on the following lemmas.

Lemma 1 — I_2 is symmetric.

PROOF : The postulate 1 for $n = 3$, $p_{1,k} + p_{2,k} > 0$, $k = 1, 2, \ldots, m$ gives

$$I_3 \begin{bmatrix} p_{1,1}, & p_{2,1}, & p_{3,1} \\ \vdots & \vdots & \vdots \\ p_{1,m}, & p_{2,m}, & p_{3,m} \end{bmatrix} = I_2 \begin{bmatrix} p_{1,1} + p_{2,1}, & p_{3,1} \\ \vdots & \vdots \\ p_{1,m} + p_{2,m}, & p_{3,m} \end{bmatrix}$$

$$+ \sum_{k=1}^{m-2} (p_{1,k} + p_{2,k})\, I_2 \begin{bmatrix} \dfrac{p_{1,k}}{p_{1,k} + p_{2,k}}, & \dfrac{p_{2,k}}{p_{1,k} + p_{2,k}} \\ \dfrac{p_{1,k+1}}{p_{1,k+1} + p_{2,k+1}}, & \dfrac{p_{2,k+1}}{p_{1,k+1} + p_{2,k+1}} \\ \dfrac{p_{1,k+2}}{p_{1,k+2} + p_{2,k+2}}, & \dfrac{p_{2,k+2}}{p_{1,k+2} + p_{2,k+2}} \end{bmatrix} \quad \ldots(3.1)$$

and

$$I_3 \begin{bmatrix} p_{2,1}, & p_{1,1}, & p_{3,1} \\ \vdots & \vdots & \vdots \\ p_{2,m}, & p_{1,m}, & p_{3,m} \end{bmatrix} = I_2 \begin{bmatrix} p_{2,1} + p_{1,1}, & p_{3,1} \\ \vdots & \vdots \\ p_{2,m} + p_{1,m}, & p_{3,m} \end{bmatrix}$$

$$+ \sum_{k=1}^{m-2} (p_{2,k} + p_{1,k})\, I_2 \begin{bmatrix} \dfrac{p_{2,k}}{p_{2,k} + p_{1,k}}, & \dfrac{p_{1,k}}{p_{2,k} + p_{1,k}} \\ \dfrac{p_{2,k+1}}{p_{2,k+1} + p_{1,k+1}}, & \dfrac{p_{1,k+1}}{p_{2,k+1} + p_{1,k+1}} \\ \dfrac{p_{2,k+2}}{p_{2,k+2} + p_{1,k+2}}, & \dfrac{p_{1,k+2}}{p_{2,k+2} + p_{1,k+2}} \end{bmatrix}. \quad \ldots(3.2)$$

Thus postulate (2), eqns. (3.1) and (3.2) prove Lemma 1 which is equivalent to

$$f(p_1, \ldots, p_m) = f[(1-p_1), \ldots, (1-p_m)] \text{ for } (p_1, \ldots, p_m) \in D. \quad \ldots(3.3)$$

In particular, (3.3) gives

$$f(0, 0, \ldots, 0) = f(1, 1, \ldots, 1). \quad \ldots(3.4)$$

Lemma 2 — f defined in postulate (3) satisfies the functional equation

$$f((x_m)) + \sum_{k=1}^{m-2} (1-x_k)\, g\left(\frac{u_k}{1-x_k}, \frac{u_{k+1}}{1-x_{k+1}}, \frac{u_{k+2}}{1-x_{k+2}}\right)$$

$$= f((u_m)) + \sum_{k=1}^{m-2} (1-u_k)\, g\left(\frac{x_k}{1-u_k}, \frac{x_{k+1}}{1-u_{k+1}}, \frac{x_{k+2}}{1-u_{k+2}}\right). \quad \ldots(3.5)$$

for $x_k, u_k \in [0, 1)$ with $x_k + u_k \in (0, 1]$ where $k = 1, 2, \ldots, m$; $f((x_m)) = f(x_1, x_2, \ldots, x_m)$ and that

$$f((x_m)) = \sum_{k=1}^{m-2} \left[x_k \log \frac{x_{k+1}}{x_{k+2}} + (1-x_k) \log \frac{1-x_{k+1}}{1-x_{k+2}}\right] \quad \ldots(3.6)$$

$$g(x_k, x_{k+1}, x_{k+2}) = x_k \log \frac{x_{k+1}}{x_{k+2}} + (1-x_k) \log \frac{1-x_{k+1}}{1-x_{k+2}} \quad \ldots(3.6a)$$

for $(x_1, \ldots, x_m) \in D$.

PROOF : The postulate (2) gives

$$I_3 \begin{bmatrix} x_{1,1}, & x_{2,1}, & x_{3,1} \\ \vdots & \vdots & \vdots \\ x_{1,m}, & x_{2,m}, & x_{3,m} \end{bmatrix} = I_3 \begin{bmatrix} x_{2,1}, & x_{3,1}, & x_{1,1} \\ \vdots & \vdots & \vdots \\ x_{2,m}, & x_{3,m}, & x_{1,m} \end{bmatrix}$$

$$= I_3 \begin{bmatrix} x_{3,1}, & x_{1,1}, & x_{2,1} \\ \vdots & \vdots & \vdots \\ x_{3,m}, & x_{1,m}, & x_{2,m} \end{bmatrix}. \quad \ldots(3.7)$$

Equations (3.7) and (3.3) alongwith postulates (3) and (1) yield

$$f((x_{1,m} + x_{2,m})) + \sum_{k=1}^{m-2} (x_{1,k} + x_{2,k})$$

$$\times\, g\left(\frac{x_{1,k}}{x_{1,k}+x_{2,k}}, \frac{x_{1,k+1}}{x_{1,k+1}+x_{2,k+1}}, \frac{x_{1,k+2}}{x_{1,k+2}+x_{2,k+2}}\right) =$$

(equation continued on p. 194)

$$= f((x_{1,m})) + \sum_{k=1}^{m-2} (1 - x_{1,k})$$

$$\times g\left(\frac{x_{2,k}}{1 - x_{1,k}}, \frac{x_{2,k+1}}{1 - x_{1,k+1}}, \frac{x_{2,k+2}}{1 - x_{1,k+2}}\right)$$

$$= f((x_{2,m})) + \sum_{k=1}^{m-2} (1 - x_{2,k})$$

$$\times g\left(\frac{x_{1,k}}{1 - x_{2,k}}, \frac{x_{1,k+1}}{1 - x_{2,k+1}}, \frac{x_{1,k+2}}{1 - x_{2,k+2}}\right). \quad ...(3.8)$$

for $x_{1,k}, x_{2,k} \in [0, 1)$, $x_{1,k} + x_{2,k} \in (0, 1]$, $k = 1, 2, ..., m$ and with the convention given in the note (section 1).

From the second and third equation pair in (3.8), we see that f, g satisfy the functional eqn. (3.5).

Let f_1, g_1 denote the partial derivatives of f, g with respect to the first variable viz. $x_{1,1}$. Differentiating partially the first and third equation pair in (3.8) with respect to $x_{1,1}$, we get

$$f_1((x_{1,m} + x_{2,m})) + g\left(\frac{x_{1,1}}{x_{1,1} + x_{2,1}}, \frac{x_{1,2}}{x_{1,2} + x_{2,2}}, \frac{x_{1,3}}{x_{1,3} + x_{2,3}}\right)$$

$$+ \frac{x_{2,1}}{x_{1,1} + x_{2,1}} g_1\left(\frac{x_{1,1}}{x_{1,1} + x_{2,1}}, \frac{x_{1,2}}{x_{1,2} + x_{2,2}}, \frac{x_{1,3}}{x_{1,3} + x_{2,3}}\right)$$

$$= g_1\left(\frac{x_{1,1}}{1 - x_{2,1}}, \frac{x_{1,2}}{1 - x_{2,2}}, \frac{x_{1,3}}{1 - x_{2,3}}\right) \quad ...(3.9)$$

for $x_{1,k} \in (0, 1)$, $x_{2,k} \in [0, 1)$ and $x_{1,k} + x_{2,k} \in (0, 1]$.

Now differentiating partially with respect to $x_{2,1}$ the first and second equation pairs in (3.8), we have

$$f_1((x_{1,m} + x_{2,m})) + \left[g\left(\frac{x_{1,1}}{x_{1,1} + x_{2,1}}, \frac{x_{1,2}}{x_{1,2} + x_{2,2}}, \frac{x_{1,3}}{x_{1,3} + x_{2,3}}\right)\right]$$

$$- \frac{x_{1,1}}{x_{1,1} + x_{2,1}}\left[g_1\left(\frac{x_{1,1}}{x_{1,1} + x_{2,1}}, \frac{x_{1,2}}{x_{1,2} + x_{2,2}}, \frac{x_{1,3}}{x_{1,3} + x_{2,3}}\right)\right]$$

$$= g_1\left(\frac{x_{2,1}}{1 - x_{1,1}}, \frac{x_{2,2}}{1 - x_{1,2}}, \frac{x_{2,3}}{1 - x_{1,3}}\right) \quad ...(3.10)$$

for $x_{2,k} \in (0, 1)$ and $x_{1,k} \in [0, 1)$, $x_{1,k} + x_{2,k} \in (0, 1]$. Subtracting (3.10) from (3.9) we have

$$g_1\left(\frac{x_{1,1}}{x_{1,1}+x_{2,1}}, \frac{x_{1,2}}{x_{1,2}+x_{2,2}}, \frac{x_{1,3}}{x_{1,3}+x_{2,3}}\right)$$
$$+ g_1\left(\frac{x_{2,1}}{1-x_{1,1}}, \frac{x_{2,2}}{1-x_{1,2}}, \frac{x_{2,3}}{1-x_{1,3}}\right)$$
$$= g_1\left(\frac{x_{1,1}}{1-x_{2,1}}, \frac{x_{1,2}}{1-x_{2,2}}, \frac{x_{1,3}}{1-x_{2,3}}\right). \quad \text{...(3.11)}$$

Integrating eqn. (3.11) with respect to the first variable (see Kannappan and Rathie 1973) we obtain

$$g(x_{1,1}, x_{1,2}, x_{1,3}) = x_{1,1}\log\frac{x_{1,2}}{x_{1,3}} + (1-x_{1,1})\log\frac{1-x_{1,2}}{1-x_{1,3}}. \quad \text{...(3.12)}$$

Equations (3.8) and (3.12) give

$$f((x_{1,m}+x_{2,m})) + \sum_{k=1}^{m-2}(x_{1,k}+x_{2,k})$$
$$\times\left[\frac{x_{1,k}}{x_{1,k}+x_{2,k}}\log\frac{x_{1,k+1}}{x_{1,k+2}}\frac{(x_{1,k+2}+x_{2,k+2})}{(x_{1,k+1}+x_{2,k+1})}\right.$$
$$\left.+\frac{x_{2,k}}{x_{1,k}+x_{2,k}}\log\frac{x_{2,k+1}}{x_{2,k+2}}\frac{(x_{1,k+2}+x_{2,k+2})}{(x_{1,k+1}+x_{2,k+1})}\right]$$
$$= f((x_{1,m})) + \sum_{k=1}^{m-2}(1-x_{1,k})$$
$$\times\left[\frac{x_{2,k}}{(1-x_{1,k})}\log\frac{x_{2,k+1}}{x_{2,k+2}}\frac{(1-x_{1,k+2})}{(1-x_{1,k+1})} + \frac{(1-x_{1,k}-x_{2,k})}{1-x_{1,k}}\right.$$
$$\left.\times\log\frac{(1-x_{1,k+1}-x_{2,k+1})(1-x_{1,k+2})}{(1-x_{1,k+2}-x_{2,k+2})(1-x_{1,k+1})}\right]$$
$$= f((x_{2,m})) + \sum_{k=1}^{m-2}(1-x_{2,k})$$
$$\times\left[\frac{x_{1,k}}{(1-x_{2,k})}\log\frac{x_{1,k+1}(1-x_{2,k+2})}{x_{1,k+2}(1-x_{2,k+1})} + \frac{(1-x_{1,k}-x_{2,k})}{(1-x_{2,k})}\right.$$
$$\left.\times\log\frac{(1-x_{1,k+1}-x_{2,k+1})(1-x_{2,k+2})}{(1-x_{1,k+2}-x_{2,k+2})(1-x_{2,k+1})}\right]. \quad \text{...(3.13)}$$

From second and third equation pair of (3.13) we have

$$f((x_{1,m})) - \sum_{k=1}^{m-2}\left[x_{1,k}\log\frac{x_{1,k+1}}{x_{1,k+2}} + (1-x_{1,k})\log\frac{(1-x_{1,k+1})}{(1-x_{1,k+2})}\right]$$
$$= f((x_{2,m})) - \sum_{k=1}^{m-2}\left[x_{2,k}\log\frac{x_{2,k+1}}{x_{2,k+2}} + (1-x_{2,k})\log\frac{(1-x_{2,k+1})}{(1-x_{2,k+2})}\right]. \quad \text{...(3.14)}$$

In eqn. (3.14), relation of all $x_{1,k}$ and relation of all $x_{2,k}$ are equal, where $k = 1, 2, ..., m$. Let

$$f((x_{1,m})) - \sum_{k=1}^{m-2} \left[x_{1,k} \log \frac{x_{1,k+1}}{x_{1,k+2}} + (1 - x_{1,k}) \log \frac{(1 - x_{1,k+1})}{(1 - x_{1,k+2})} \right]$$

$$= f((x_{2,m})) - \sum_{k=1}^{m-2} \left[x_{2,k} \log \frac{x_{2,k+1}}{x_{2,k+2}} + (1 - x_{2,k}) \frac{1 - x_{2,k+1}}{1 - x_{2,k+2}} \right]$$

$$= A \text{ (constant).} \qquad ...(3.15)$$

From the postulate (5) and (3.15) we obtain

$$A = 0.$$

Hence Lemma 2 is proved.

Proof of the Theorem

Applying the postulate (6), we have

$$I_n [P_m] = \sum_{i=1}^{n} f((x_{i,m})). \qquad ...(3.16)$$

Hence (3.16) and (3.6) give

$$= \sum_{i=1}^{n} \sum_{k=1}^{m-2} \left[p_{i,k} \log \frac{p_{i,k+1}}{p_{i,k+2}} + (1 - p_{i,k}) \log \frac{(1 - p_{i,k+1})}{(1 - p_{i,k+2})} \right]. \qquad ...(3.17)$$

The right-hand side of (3.17) on expansion will give two terms. The first term will correspond to the left-hand side whereas the second term can be expressed as log [1 + c (small)] which is eliminated by postulate (5).

$$I_n [P_m] = \sum_{k=1}^{m-2} \sum_{i=1}^{n} p_{i,k} \log \frac{p_{i,k+1}}{p_{i,k+2}}$$

which proves the theorem.

Acknowledgement

Our thanks are due to the referee for very valuable criticism which has considerably improved the paper. One of the authors (O.P.V.) is grateful to the U.G.C. for the award of a teacher fellowship.

REFERENCES

Aczél, J., and Nath, P. (1972). Axiomatic characterization of some measures of divergence in information. *Z. Wahns. Verw. Geb.*, **21**, 215–24.

Kannappan, Pl. (1975). On various characterization of generalized directed divergence. *Indian J. pure appl. Math.*, **6**, 655–67.

Kannappan, Pl., and Rathie, P. N. (1973). An axiomatic characterizations of generalized directed-divergence. *Kybernetika*, **9**, 330–37.

Nath, P. (1968). On the measures of errors in information. *J. Math. Sci., Delhi*, **3**, 1–16.

——————— (1972). Some axiomatic characterizations of a non-additive measure of divergence in information. R. S. Varma Memorial Volume. *J. Math. Sci., Delhi*, **7**, 57–68.

Sharma, B. D., and Ram Autar (1974). Information-improvement function. *Econometrica*, **42**, 103-12.

Theil, H. (1967). Economics and Information Theory. North-Holland Pub. Co , Amsterdam.

APPENDIX

This model is likely to be useful in physical situation where weather forcasting is based on data collection, revision of data and several re-revisions to arrive at the correct picture to save life and loss.

The data collector makes the prediction that cyclone observed in Australia is progressing towards India. Further data makes him revise his prediction by sending out information that northern part of India will be affected. When some more observations arrive in he re-revises his stand by sending out information that the capital of Rajasthan is likely to be hit by the cyclone. After revising and re-revising his analysis of more and more incoming data he predicts that only Hawa-Mahal building is likely to be erased by the cyclone. Thus in the above example this model is likely to be useful in making the forecast about the weather in and around the particular building so that the government can act in time to save life and loss.