ON VARIOUS CHARACTERIZATIONS OF GENERALIZED DIRECTED DIVERGENCE

by Pl. KANNAPPAN, *University of Waterloo, Waterloo, Ontario, Canada*

(Communicated by P. L. Bhatnagar, F.N.A.)

(*Received* 16 *July* 1973)

Many systems of postulates have been given to characterize the Shannon's entropy and directed divergence. In this paper, the measure generalized directed divergence which is connected to the above measures is characterized by four sets of postulates

INTRODUCTION

Let the true probabilities of a system of events, for instance the various outcomes of a certain experiment be given by

$P = (p_1, p_2, \ldots, P_n)$ with $p_i \geqslant 0$, $\sum_{i=1}^{n} p_i = 1$ (called the complete discrete probability distribution). Sometimes it may not be possible to forcast the physical phenomena exactly due to one reason or other; it is possible that some of the outcomes may not be observable; the persons who perform the experiment may not have taken all the relevant factors into consideration etc. Thus the probability distributions estimated by two independent observers A and B who assert that the probabilities associated with the same experiment, $Q = (q_1, \ldots, q_n)$ and $R = (r_1, \ldots, r_n)$ respectively may even be incomplete [called generalized probability distribution by Rényi (1961)] that is, $q_j, r_k \geqslant 0$ with $\sum_{j=1}^{n} q_j \leqslant 1$, $\sum_{k=1}^{n} r_k \leqslant 1$. The following question can be asked : What is the amount of directed divergence between the estimations Q and R of the experimentors A and B for the probability distribution of an experiment which actually is P ? An answer to this question can be given by the following measure called the generalized directed divergence

$$I_n(P \mid \mid Q \mid R) = \sum_{i=1}^{n} p_i \log \frac{q_i}{r_i}. \qquad \ldots(1)$$

For more information about the measure (1) refer to Nath (1968), Aczél and Nath (1972), Kannappan (1973), Kannappan and Rathie (1973). In this paper, the measure (1) is characterized by 4 sets of postulates. There are many algebraic properties satisfied by I_n. The postulates given below reflect some of the basic properties of I_n.

Remark : The following convention is followed throughout this paper. The logarithm is taken to the base 2; $O \log 0 = 0$, $O \log \frac{y}{0} = 0 = O. \log \frac{0}{z}$ and $q_i = 0$ or $r_i = 0$ imply the corresponding $p_i = 0$.

Systems of Postulates

Postulates A

A_1 — *Recursivity or Branching*

$$I_n(P \mid \mid Q \mid R) = I_{n-1}\begin{pmatrix} p_1 + p_2, p_3, \ldots, p_n \\ q_1 + q_2, q_3, \ldots, q_n \\ r_1 + r_2, r_3, \ldots, r_n \end{pmatrix} + (p_1+p_2)\, I_2\begin{pmatrix} \frac{p_1}{p_1+p_2}, & \frac{p_2}{p_1+p_2} \\ \frac{q_1}{q_1+q_2}, & \frac{q_2}{q_1+q_2} \\ \frac{r_1}{r_1+r_2}, & \frac{r_2}{r_1+r_2} \end{pmatrix},$$

for $p_1 + p_2, q_1 + q_2, r_1 + r_2 > 0$.

A_2 — *Symmetry*

$$I_4(P \mid \mid Q \mid R) \text{ is symmetric in } \begin{pmatrix} p_i \\ q_i \\ r_i \end{pmatrix}, \quad i = 1, 2, 3, 4.$$

$$A_3 — f(p, q, r) = I_2\begin{pmatrix} p, 1 - p \\ q, 1 - q \\ r, 1 - r \end{pmatrix} \text{ is continuous on } J,$$

where $J =]0, 1[\times]0, 1[\times]0, 1[\cup \{(0, y, z)\} \cup \{1, y', z')\}$ with $y, z \in [0, 1[$, $y', z' \in]0, 1]$.

A_4 — *Normalization*

$f(2/3, 2/3, 1/3) = 1/3$.

This determines the unit of divergence.

A_5 — *Nullity*

$f(1/2, 1/2, 1/2) = 0 = f(2/3, 1/3, 1/3)$.

Since we are measuring the amount of divergence between two estimations of the probabilities, it is natural to assume that the measure is zero, when there is no difference between the probabilities estimated by the two observers.

Postulates B

There is a real valued continuous function f defined on J such that :

B_1 — *Representability*

$$I_n(P \mid \mid Q \mid R) = \sum_{i=1}^{n} f(p_i, q_i, r_i), \text{ for all } n \geqslant 2.$$

B_2 — *Additivity*

$$I_{mn}(P^*U \mid \mid Q^*V \mid R^*W) = I_n(P \mid \mid Q \mid R) + I_m(U \mid \mid V \mid W),$$

where $U = (u_1, \ldots, u_m)$, $V = (v_1, \ldots, v_m)$, $W = (w_1, \ldots, w_m)$ are finite discrete probability distributions and $P^*U = (p_1u_1, \ldots p_1u_m, \ldots, p_nu_1, \ldots, p_nu_m)$ and similar expressions for Q^*V and Z^*W are the joint distributions.

B_3 — *Expansibility*

$$I_n\begin{pmatrix} p_1, \ldots, p_{n-1}, 0 \\ q_1, \ldots, q_{n-1}, q_n \\ r_1, \ldots, r_{n-1}, r_n \end{pmatrix} = I_{n-1}\begin{pmatrix} p_1, \ldots, p_{n-1} \\ q_1, \ldots, q_{n-1} \\ r_1, \ldots, r_{n-1} \end{pmatrix},$$

where $\sum_{i=1}^{n} p_i = 1$, $\sum_{i=1}^{n} q_i \leqslant 1$, $\sum_{i=1}^{n} r_i \leqslant 1$, for $n \geqslant 3$.

B_4 — *Normalization*

$$I_2\begin{pmatrix} 2/3, 1/3 \\ 2/3, 1/3 \\ 1/3, 2/3 \end{pmatrix} = 1/3.$$

B_5 — *Nullity*

$$I_2\begin{pmatrix} 1/2, 1/2 \\ 1/2, 1/2 \\ 1/2, 1/2 \end{pmatrix} = 0 = I_2\begin{pmatrix} 2/3, 1/3 \\ 1/3, 2/3 \\ 1/3, 2/3 \end{pmatrix}.$$

Postulates C

There is a real valued continuous function f on J such that

$$C_1 — I_n(P \mid \mid Q \mid R) = \sum_{i=1}^{n} f(p_i, q_i, r_i), \text{ for all } n \geqslant 2.$$

$$C_2 - I_{2n}\begin{pmatrix} p_1p, p_1(1-p), \ldots, p_np, p_n(1-p) \\ q_1q, q_1(1-q), \ldots, q_nq, q_n(1-q) \\ r_1r, r_1(1-r), \ldots, r_nr, r_n(1-r) \end{pmatrix} = I_n(P \mid \mid Q \mid R)$$

$$+ I_2\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}, \text{ for } n \geqslant 2.$$

C_3 — *Expansibility*

$$I_n\begin{pmatrix} p_1, \ldots, p_{n-1}, 0 \\ q_1, \ldots, q_{n-1}, q_n \\ r_1, \ldots, r_{n-1}, r_n \end{pmatrix} = I_{n-1}\begin{pmatrix} p_1, \ldots, p_{n-1} \\ q_1, \ldots, q_{n-1} \\ r_1, \ldots, r_{n-1} \end{pmatrix},$$

where $\sum_{i=1}^{n-1} p_i = 1, \ \sum_{i=1}^{n} q_i \leqslant 1, \ \sum_{i=1}^{n} r_i \leqslant 1$, for $n \geqslant 3$.

C_4 — *Normalization*

$$I_2\begin{pmatrix} 2/3, 1/3 \\ 2/3, 1/3 \\ 1/3, 2/3 \end{pmatrix} = 1/3.$$

C_5 — *Nullity*

$$I_2\begin{pmatrix} 1/2, 1/2 \\ 1/2, 1/2 \\ 1/2, 1/2 \end{pmatrix} = 0 = I_2\begin{pmatrix} 2/3, 1/3 \\ 1/3, 2/3 \\ 1/3, 2/3 \end{pmatrix}.$$

Postulates D

D_1 — $I_4(P \mid \mid Q \mid R)$ is symmetric in $\begin{pmatrix} p_i \\ q_i \\ r_i \end{pmatrix}$, $i = 1, 2, 3, 4$.

$$D_2 - I_n(P \mid \mid Q \mid R) - I_{n-1}\begin{pmatrix} p_1+p_2, p_3, \ldots, p_n \\ q_1+q_2, q_3, \ldots, q_n \\ r_1+r_2, r_3, \ldots, r_n \end{pmatrix} = \triangle_{n-1}\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix},$$

for all $n \geqslant 3$ and the function $\triangle_2$ defined on

$$D = \left\{ \begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} : 0 \leqslant p_1 + p_2 \leqslant 1, 0 \leqslant q_1 + q_2 \leqslant 1, 0 \leqslant r_1 + r_2 \leqslant 1 \right\}$$

is continuous.

D_3 — *Expansibility*

$$I_n \begin{pmatrix} p_1, \ldots, p_{n-1}, 0 \\ q_1, \ldots, q_{n-1}, q_n \\ r_1, \ldots, r_{n-1}, r_n \end{pmatrix} = I_{n-1} \begin{pmatrix} p_1, \ldots, p_{n-1} \\ q_1, \ldots, q_{n-1} \\ r_1, \ldots, r_{n-1} \end{pmatrix},$$

where $\sum_{i=1}^{n-1} p_i = 1, \sum_{i=1}^{n} q_i \leqslant 1, \sum_{i=1}^{n} r_i \leqslant 1$, for $n \geqslant 3$.

D_4 — *Normalization*

$$I_2 \begin{pmatrix} 2/3, 1/3 \\ 2/3, 1/3 \\ 1/3, 2/3 \end{pmatrix} = 1/3.$$

D_5 — *Nullity*

$$I_2 \begin{pmatrix} 1/2, 1/2 \\ 1/2, 1/2 \\ 1/2, 1/2 \end{pmatrix} = 0 = I_2 \begin{pmatrix} 2/3, 1/3 \\ 1/3, 2/3 \\ 1/3, 2/3 \end{pmatrix}.$$

$$D_6 - I_{2n} \begin{pmatrix} p_1 p, p_1(1-p), \ldots, p_n p, p_n(1-p) \\ q_1 q, q_1(1-q), \ldots, q_n q, q_n(1-q) \\ r_1 r, r_1(1-r), \ldots, r_n r, r_n(1-q) \end{pmatrix} = I_n(P \mid \mid Q \mid R)$$

$$+ I_2 \begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}, \text{ for } n \geqslant 2.$$

Theorem—The systems of postulates A, B, C and D are equivalent and characterize the quantity generalized directed divergence I_n given by (1).

PROOF : We will prove the theorem in stages as follows :

$B \Rightarrow C \Rightarrow D \Rightarrow A \Rightarrow B$.

(a) $B \Rightarrow C$ is trivial to verify.

(b) Now we will show that $C \Rightarrow D$.

Evidently $C_1 \Rightarrow D_1$. Again from C_1 follows D_2, by taking

$$\Delta_{n-1}\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} = f(p_1, q_1, r_1)+f(p_2, q_2, r_2) - f(p_1+p_2, q_1+q_2, r_1+r_2)$$

and noting the fact that f is continuous. Further, since C_3, C_4, C_5 and C_2 are the same as D_3, D_4, D_5 and D_6, $C \Rightarrow D$ is true.

(c) Now to prove that $D \Rightarrow A$. This is achieved through the following lemmas.

Lemma 1—If $\alpha\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} := \Delta_2\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix}$, for

$\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} \in D$, then $\Delta_n\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} = \alpha\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix}$, for $n \geqslant 2$.

PROOF : Choose p_3, q_3 and r_3 such that $p_1+p_2+p_3 = 1$, $q_1+q_2+q_3 \leqslant 1$, $r_1+r_2+r_3 \leqslant 1$. Then, from D_2 and D_3, by letting $p_4 = 0 \ldots = p_n = q_4 = \ldots = q_n = r_4 = \ldots = r_n$, we obtain

$$\Delta_{n-1}\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} = I_n\begin{pmatrix} p_1, p_2, p_3, 0, \ldots, 0 \\ q_1, q_2, q_3, 0, \ldots, 0 \\ r_1, r_2, r_3, 0, \ldots, 0 \end{pmatrix} - I_{n-1}\begin{pmatrix} p_1+p_2, p_3, 0, \ldots, 0 \\ q_1+q_2, q_3, 0, \ldots, 0 \\ r_1+r_2, r_3, 0, \ldots, 0 \end{pmatrix}$$

$$= I_3\begin{pmatrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \\ r_1, r_2, r_3 \end{pmatrix} - I_2\begin{pmatrix} p_1+p_2, p_3 \\ q_1+q_2, q_3 \\ r_1+r_2, r_3 \end{pmatrix}$$

$$= \Delta_2\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix},$$

which proves the lemma.

Since Δ_2 is continuous on D, α is also continuous on D.

Lemma 2—I_n is symmetric in $\begin{pmatrix} p_i \\ q_i \\ r_i \end{pmatrix}$, $i = 1, 2, \ldots, n$.

PROOF : From the symmetry of $I_4(D_1)$ and D_3, it follows that I_2, I_3 and α are symmetric. We will now prove that I_n is symmetric for $n \geqslant 2$, by the use of induction. Let us assume that I_{n-1} be symmetric for $n \geqslant 5$. Then by D_2, I_n will be symmetric in $\begin{pmatrix} p_i \\ q_i \\ r_i \end{pmatrix}$, $i = 1, 2, 3, \ldots, n$, provided I_n is symmetric in $\begin{pmatrix} p_2 \\ q_2 \\ r_3 \end{pmatrix}$ and $\begin{pmatrix} p_3 \\ q_3 \\ r_3 \end{pmatrix}$. This will happen, if

$$I_{n-1}\begin{pmatrix} p_1+p_2, p_3, \ldots, p_n \\ q_1+q_2, q_3, \ldots, q_n \\ r_1+r_2, r_3, \ldots, r_n \end{pmatrix} + \alpha\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} = I_{n-1}\begin{pmatrix} p_1+p_3, p_2, \ldots, p_n \\ q_1+q_3, q_2, \ldots, q_n \\ r_1+r_3, r_2, \ldots, r_n \end{pmatrix} + \alpha\begin{pmatrix} p_1, p_3 \\ q_1, q_3 \\ r_1, r_3 \end{pmatrix}$$

holds, that is, if

$$\alpha\begin{pmatrix} p_1+p_2, p_3 \\ q_1+q_2, q_3 \\ r_1+r_2, r_3 \end{pmatrix} + \alpha\begin{pmatrix} p_1, p_2 \\ q_1, q_2 \\ r_1, r_2 \end{pmatrix} = \alpha\begin{pmatrix} p_1+p_3, p_2 \\ q_1+q_3, q_2 \\ r_1+r_3, r_2 \end{pmatrix} + \alpha\begin{pmatrix} p_1, p_3 \\ q_1, q_3 \\ r_1, r_3 \end{pmatrix}$$

holds, which is indeed the case, as can be seen from the symmetry of I_4 by interchanging $\begin{pmatrix} p_2 \\ q_2 \\ r_2 \end{pmatrix}$ and $\begin{pmatrix} p_3 \\ q_3 \\ r_3 \end{pmatrix}$ in I_4 and using D_2 for $n = 4$ and lemma 1. This completes the proof of this lemma.

Lemma 3—For all $p, q, r \in [0, 1]$,

$$\alpha\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix} = I_2\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}.$$

PROOF : By Lemma 1, D_2 for $n = 3$ and D_3 follow

$$\alpha\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix} = I_3\begin{pmatrix} p, 1-p, 0 \\ q, 1-q, 0 \\ r, 1-r, 0 \end{pmatrix} - I_2\begin{pmatrix} 1, 0 \\ 1, 0 \\ 1, 0 \end{pmatrix}$$

$$= I_2\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix} - I_2\begin{pmatrix} 1, 0 \\ 1, 0 \\ 1, 0 \end{pmatrix}.$$

But D_6 for $n = 2$, D_1 and D_3 imply

$$I_4\begin{pmatrix} p.1, p.0(1-p).1, (1-p).0 \\ q, 0, 1-q, 0 \\ r, 0, 1-q, 0 \end{pmatrix} = I_2\begin{pmatrix} 1, 0 \\ 1, 0 \\ 1, 0 \end{pmatrix} + I_2\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}$$

that is, $I_2\begin{pmatrix} 1, 0 \\ 1, 0 \\ 1, 0 \end{pmatrix} = 0$. Thus this lemma is proved.

Lemma 4—For an arbitrary rational $s \in [0, 1]$.

$$\alpha\begin{pmatrix} sp, s(1-p) \\ tq, t(1-q) \\ ur, u(1-r) \end{pmatrix} = s\alpha\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix} \qquad \ldots(2)$$

for all $p, q, r, t, u \in [0, 1]$.

PROOF : From D_6 and the repeated use of D_2 and Lemma 1, we get

$$\sum_{i=1}^{n} \alpha\begin{pmatrix} p_i p, p_i(1-p) \\ q_i q, q_i(1-q) \\ r_i r, r_i(1-r) \end{pmatrix} = \alpha\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}. \qquad \ldots(3)$$

By taking $p_i = \frac{1}{n}$, $q_i = t \leqslant \frac{1}{n}$, $r_i = u \leqslant \frac{1}{n}$ $(i = 1, 2, \ldots, n)$ in (3), we have

$$\alpha\begin{pmatrix} \frac{1}{n}p, \frac{1}{n}(1-p) \\ tq, \quad t(1-q) \\ ur, \quad u(1-r) \end{pmatrix} = \frac{1}{n}\alpha\begin{pmatrix} p, 1-p \\ q, 1-q \\ r, 1-r \end{pmatrix}, \text{ for } t, u \leqslant \frac{1}{n}. \qquad \ldots(4)$$

Suppose $\frac{1}{n} \leqslant t, u < 1$. Then choose an integer $m > n$. Letting $p_j = \frac{1}{n}$, $q_1 = t$, $q_i = \frac{1-t}{m}$, $r_1 = u$, $r_i = \frac{1-u}{m}$, $j = 1, 2, \ldots, n$; $i = 2, 3, \ldots, n$ in (3) and using (4), it is easy to see that (4) holds for $t, u > \frac{1}{n}$ also, that is, (4) is true for all $t, u \in [0, 1]$.

For any rational $s = \frac{n_1}{n_2} \in [0, 1]$, (3) for $n = n_2 - n_1 + 1$, and for any integer $m > n_2$, with $p_1 = \frac{n_1}{n_2}$, $p_i = \frac{1}{n_2}$, $q_1 = t$, $q_i = \frac{1-t}{m}$, $r_1 = s$, $r_i = \frac{1-s}{m}$ $(i = 2, 3, \ldots, n_2 - n_1 + 1)$ where $t, u \in [0, 1]$ and (4) yield

$$\alpha \begin{pmatrix} \frac{n_1}{n_2} p, & \frac{n_1}{n_2}(1-p) \\ tq, & t(1-q) \\ ur, & u(1-r) \end{pmatrix} + (n_2 - n_1)\alpha \begin{pmatrix} \frac{p}{n_2}, & \frac{1-p}{n_2} \\ \frac{(1-t)\ q}{m}, & \frac{(1-t)(1-q)}{m} \\ \frac{(1-u)r}{m}, & \frac{(1-u)(1-r)}{m} \end{pmatrix}$$

$$= \alpha \begin{pmatrix} p, & 1-p \\ q, & 1-q \\ r, & 1-r \end{pmatrix}$$

so that (2) is true. Thus the proof of this lemma is complete.

Lemma 5—α defined on D has the following form, that is,

$$\alpha \begin{pmatrix} p_1, & p_2 \\ q_2, & q_2 \\ r_1, & r_2 \end{pmatrix} = (p_1 + p_2)\, I_2 \begin{pmatrix} \frac{p_1}{p_1+p_2}, & \frac{p_2}{p_1+p_2} \\ \frac{q_1}{q_1+q_2}, & \frac{q_2}{q_1+q_2} \\ \frac{r_1}{r_1+r_2}, & \frac{r_2}{r_1+r_2} \end{pmatrix} \qquad \ldots(5)$$

for $\begin{pmatrix} p_1, & p_2 \\ q_1, & q_2 \\ r_1, & r_2 \end{pmatrix} \in D$.

Proof: From Lemma 4 and the continuity of α, it follows that (2) holds for all $s, t, u, p, q, r \in [0, 1]$. Now allowing $sp = p_1$, $s(1-p) = p_2$, $tq = q_1$, $t(1-q) = q_2$, $ur = r_1$ and $u(1-r) = r_2$ in (2), we obtain (5).

Proof of $D \Rightarrow A$ — D_2 and Lemma 6 imply A_1 . D_1, D_4 and D_5 are the same as A_2, A_4 and A_5. Lemma 3 is same as A_3. Thus $D \Rightarrow A$.

Remark : The proof of $D \Rightarrow A$ given here is based on the proof of Daróczy (1967). Another proof of the same can be given based on Kannappan and Rathie (1971), under the weaker assumption of D_6 being true for $n = 2$ and 3 only.

(d) Let us now assume the postulates A to be true and prove that I_n indeed has the form given by (1). The proof is prepared through the following lemmas.

Lemma 6 — f defined in A_3, has the property

$$f(p, q, r) = f(1 - p, 1 - q, 1 - r), \text{ for } (p, q, r) \in J. \quad \ldots(6)$$

PROOF : By A_2, i.e. symmetry of I_4, we have

$$I_4\begin{pmatrix} p_1, p_2, p_3, p_4 \\ q_1, q_2, q_3, q_4 \\ r_1, r_2, r_3, r_4 \end{pmatrix} = I_4\begin{pmatrix} p_2, p_1, p_3, p_4 \\ q_2, q_1, q_3, q_4 \\ r_2, r_1, r_3, r_4 \end{pmatrix}$$

so that an application of A_1 for $n = 4$ implies

$$I_2\begin{pmatrix} \frac{p_1}{p_1 + p_2}, & \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, & \frac{q_2}{q_1 + q_2} \\ \frac{r_1}{r_1 + r_2}, & \frac{r_2}{r_1 + r_2} \end{pmatrix} = I_2\begin{pmatrix} \frac{p_2}{p_1 + p_2}, & \frac{p_1}{p_1 + p_2} \\ \frac{q_2}{q_1 + q_2}, & \frac{q_1}{q_1 + q_2} \\ \frac{r_2}{r_1 + r_2}, & \frac{r_1}{r_1 + r_2} \end{pmatrix}$$

which by A_3 proves the lemma.

Lemma 7—The function f defined in A_3 satisfies the functional equation

$$f(x, y, z) + (1 - x)\, f\left(\frac{u}{1 - x}, \frac{v}{1 - y}, \frac{w}{1 - z}\right) = f(u, v, w)$$
$$+ (1 - u)\, f\left(\frac{x}{1 - u}, \frac{y}{1 - v}, \frac{z}{1 - w}\right) \quad \ldots(7)$$

for $x, y, z, u, v, w \in [0, 1[$ with $x + u,\ y + v,\ z + w \in]0, 1]$.

PROOF : Applying A_1 for $n = 4$, we have

$$I_4\begin{pmatrix} p_1, p_2, p_3, p_4 \\ q_1, q_2, q_3, q_4 \\ r_1, r_2, r_3, r_4 \end{pmatrix} = I_3\begin{pmatrix} p_1 + p_2, p_3, p_4 \\ q_1 + q_2, q_3, q_4 \\ r_1 + r_2, r_3, r_4 \end{pmatrix} +$$

(*equation continued on p.* 665)

$$+ (p_1 + p_2) I_2 \begin{pmatrix} \frac{p_1}{p_1 + p_2}, & \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, & \frac{q_2}{q_1 + q_2} \\ \frac{r_1}{r_1 + r_2}, & \frac{r_2}{r_1 + r_2} \end{pmatrix} = I_2 \begin{pmatrix} p_1 + p_2 + p_3, p_4 \\ q_1 + q_2 + q_3, q_4 \\ r_1 + r_2 + r_3, r_4 \end{pmatrix}$$

$$+ (p_1 + p_2 + p_3) I_2 \begin{pmatrix} \frac{p_1 + p_2}{p_1 + p_2 + p_3}, & \frac{p_3}{p_1 + p_2 + p_3} \\ \frac{q_1 + q_2}{q_1 + q_2 + q_3}, & \frac{q_3}{q_1 + q_2 + q_3} \\ \frac{r_1 + r_2}{r_1 + r_2 + r_3}, & \frac{r_3}{r_1 + r_2 + r_3} \end{pmatrix}$$

$$+ (p_1 + p_2) I_2 \begin{pmatrix} \frac{p_1}{p_1 + p_2}, & \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, & \frac{q_2}{q_1 + q_2} \\ \frac{r_1}{r_1 + r_2}, & \frac{r_2}{r_1 + r_2} \end{pmatrix}. \qquad \ldots(8)$$

But by symmetry of I_4 from A_2 and A_1, we see that

$$I_4 \begin{pmatrix} p_1, p_2, p_3, p_4 \\ q_1, q_2, q_3, q_4 \\ r_1, r_2, p_3, p_4 \end{pmatrix} = I_4 \begin{pmatrix} p_1, p_2, p_3, p_4 \\ q_1, q_2, q_4, q_3 \\ r_1, r_2, r_4, r_3 \end{pmatrix} = I_2 \begin{pmatrix} p_1 + p_2 + p_4, p_3 \\ q_1 + q_2 + q_4, q_3 \\ r_1 + r_2 + r_4, r_3 \end{pmatrix}$$

$$+ (p_1 + p_2 + p_4) I_2 \begin{pmatrix} \frac{p_1 + p_2}{p_1 + p_2 + p_4}, & \frac{p_4}{p_1 + p_2 + p_4} \\ \frac{q_1 + q_2}{q_1 + q_2 + q_4}, & \frac{q_4}{q_1 + q_2 + q_4} \\ \frac{r_1 + r_2}{r_1 + r_2 + r_4}, & \frac{r_4}{r_1 + r_2 + r_4} \end{pmatrix}$$

$$+ (p_1 + p_2) I_2 \begin{pmatrix} \frac{p_1}{p_1 + p_2}, & \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, & \frac{q_2}{q_1 + q_2} \\ \frac{r_1}{r_1 + r_2}, & \frac{r_2}{r_1 + r_2} \end{pmatrix}. \qquad \ldots(9)$$

Now, use (8) and (9) with $x = p_4$, $1 - x = p_1 + p_2 + p_3$, $y = q_4$, $1 - y = q_1 + q_2 + q_3$, $z = r_4$, $1 - z = r_1 + r_2 + r_3$, $u = p_3$, $1 - u = p_1 + p_2 + p_4$, $v = q_3$, $1 - v = q_1 + q_2 + q_4$, $w = r_3$ and $1 - w = r_1 + r_2 + r_4$ to get, with the use of A_3 that

$$f(1-x,\ 1-y,\ 1-z)+(1-x)f\left(\frac{1-u-x}{1-x},\frac{1-v-y}{1-y},\frac{1-w-z}{1-z}\right)$$

$$=f(1-u,1-v,1-w)+(1-u)f\left(\frac{1-x-u}{1-u},\frac{1-y-v}{1-v},\frac{1-z-w}{1-w}\right)$$

which by Lemma 6 yields (7). This proves the lemma.

Lemma 8—f satisfying the functional equation (7) has the form

$$f(x,y,z)=x\log\frac{y}{z}+(1-x)\log\frac{1-y}{1-z},\ \text{for } (x,y,z)\in J. \qquad \ldots(10)$$

PROOF : Since f is continuous and satisfies (7), it follows from Kannappan and Ng (1974), f has the form

$$f(x,y,z)=a[x\log x+(1-x)\log(1-x)]+\left(b\log\frac{y}{1-y}\right.$$

$$\left.+c\log\frac{z}{1-z}+d\right)x+b\log(1-y)+c\log(1-z) \qquad \ldots(11)$$

where a, b, c and d are arbitrary constants.

The symmetry of f given by (6) shows that $d=0$. The conditions of nullity in A_5 give $a=0$ and $b+c=0$. That is,

$$f(x,y,z)=b\left(x\log\frac{y}{z}+(1-x)\log\frac{1-y}{1-z}\right).$$

But then the unit in A_4 implies $b=1$. Thus f has the form as given by (10). This proves the lemma.

By repeated application of A_1, using A_3, we get

$$I_n(P\mid\mid Q\mid R)=\sum_{i=2}^{n}p_i f\left(\frac{p_i}{P_i},\frac{q_i}{Q_i},\frac{r_i}{R_i}\right),$$

where $P_i=p_1+\ldots+p_i$, $Q=q_1+\ldots+q_i$, $R=r_1+\ldots+r_i$ for $i=1,2,\ldots,n$ with $P_n=Q_n=R_n=1$, which with the help of (10) gives (1). Thus $A\Rightarrow I_n$.

Evidently I_n given by (1) satisfies all the sets of postulates and in particular the set B, so that $A\Rightarrow B$.

Hence the postulates A, B, C and D are equivalent and characterize the generalized directed divergence. Thus the proof of the theorem is complete.

REFERENCES

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

Daróczy, Z. (1967). Über eine characterisierung der Shannonschen entropie. *Statistica, Bologna*, **27**, 199-206.

Kannappan, Pl. (1973). On a functional equation connected with generalized directed divergence. (to appear).

Kannappan, Pl., and Ng, C. T. (1974). A functional equation and its application to information theory. (to appear).

Kannappan, Pl., and Rathie, P. N. (1971). On various characterizations of directed divergence. *Proc. of the Sixth Prague Conference on Information Theory, Statistical Decision Functions and Random Process.*

Kannappan, Pl., and Rathie, P. N. (1973). An axiomatic characterization of generalized directed divergence. (to appear).

Nath, P. (1968). On measures of errors in information. *J. math. Sci.*, **3**, 1-16.

Rényi, A. (1961). On measures of entropy and information. *Proc. Fourth Berk. Symp. on Math. Stat. and Prob.*, **1**, 547-61.