

ON CHARACTERIZATION OF DIRECTED DIVERGENCE OF TYPE β THROUGH INFORMATION EQUATION

R. P. SINGH*, R. K. KHANNA

The directed-divergence of type β ($\beta > 0, \beta \neq 1$) has been characterized through an 'Information Equation' and its solution, under the homogeneity (of type $\beta, \beta > 0, \beta \neq 1$) has been obtained. Some applications of the directed-divergence of type β to Information Theory have been discussed.

1. INTRODUCTION

The information theoretic concepts as envisaged in various measures, namely Kullback's information or directed-divergence [5], Kerridge's inaccuracy [7] and Theil's information improvement [12], have found many applications in behavioural sciences. Characterizations of these measures in arbitrary probability spaces and continuous analogs have been discussed earlier by Campbell [2], Rathie and Kannappan [8], [6], Sharma and Autar [10], Sharma and Soni [11] and Renyi [9] etc.

The object of this contribution is to characterize the directed-divergence of type β ($\beta > 0, \beta \neq 1$) through 'information equation' and to discuss some applications of it to information theory. Let the true probabilities of a system of events be given by the complete probability distribution:

$$P = (p_1, p_2, \dots, p_n), \quad p_i \geq 0, \quad \sum_{i=1}^n p_i = 1.$$

Let the $Q = (q_1, q_2, \dots, q_n), q_i \geq 0, \sum_{i=1}^n q_i = 1$ be the revised probability distribution.

The measures of error made by the observer or the measures of information gain, estimating the discrete probability distribution Q from the probability distribution P are given by

$$(1.1) \quad I_n \left(\begin{matrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{matrix} \right) = \sum_{i=1}^n p_i \log (p_i q_i^{-1})$$

* First author is thankful to U.G.C. (India) for financial assistance.

and

$$(1.2) \quad I_n^\beta(p_1, p_2, \dots, p_n) = (2^{\beta-1} - 1)^{-1} \left[\sum_{i=1}^n p_i^\beta q_i^{1-\beta} - 1 \right]$$

(cf. [5], and [8]) where $\beta > 0, \beta \neq 1$.

Here we consider the 'information equation' given by

$$(1.3) \quad I \begin{pmatrix} x, y, z \\ l, m, n \end{pmatrix} = I \begin{pmatrix} x+y, 0, z \\ l+m, 0, n \end{pmatrix} + I \begin{pmatrix} x, y, 0 \\ l, m, 0 \end{pmatrix}$$

in the domain $D^2 = \{(x, y, z; l, m, n); x, y, z \geq 0, l, m, n \geq 0, xy + yz + zx > 0, lm + mn + nl > 0\}$, a generalization of entropy equation [4] viz.

$$(1.4) \quad H(x, y, z) = H(x+y, 0, z) + H(x, y, 0) \\ (x, y, z \geq 0, xy + yz + zx > 0).$$

The homogeneity condition considered here is defined as follows:

$$(1.5) \quad I^\beta \begin{pmatrix} \lambda x, \lambda y, \lambda z \\ \mu l, \mu m, \mu n \end{pmatrix} = \lambda^\beta \mu^{1-\beta} I_3^\beta \begin{pmatrix} x, y, z \\ l, m, n \end{pmatrix}, \quad \lambda, \mu > 0, \quad \beta > 0, \quad \beta \neq 1.$$

The symmetric and homogeneous (of type $\beta, \beta > 0, \beta \neq 1$) solution of (1.3) has been given in Section 2 and its applications to information theory have been discussed in Section 3.

2. SOLUTION OF INFORMATION EQUATION AND CHARACTERIZATION OF DIRECTED-DIVERGENCE

In this section we solve the information equation (1.3) and characterize the directed-divergence of type β under the homogeneity condition (1.5). Let the measure (1.2) satisfy the following postulates:

Postulate 1. Branching property i.e.

$$(2.1) \quad I_n^\beta \begin{pmatrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{pmatrix} = I_{n-1}^\beta \begin{pmatrix} p_1 + p_2, p_3, \dots, p_n \\ q_1 + q_2, q_3, \dots, q_n \end{pmatrix} + \\ + (p_1 + p_2)^\beta (q_1 + q_2)^{1-\beta} I_2^\beta \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$$

Postulate 2. Symmetry i.e.

$$(2.2) \quad I_n^\beta \begin{pmatrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{pmatrix} = I_n^\beta \begin{pmatrix} p_{k(1)}, p_{k(2)}, \dots, p_{k(n)} \\ q_{k(1)}, q_{k(2)}, \dots, q_{k(n)} \end{pmatrix}$$

where $k(1), k(2), \dots, k(n)$ is a permutation of $1, 2, \dots, n$.

Postulate 3. Nullity i.e.

$$(2.3) \quad I_n^\beta \begin{pmatrix} 1, 0, 0 \\ 1, 0, 0 \end{pmatrix} = 0.$$

Postulate 4. Unit i.e.

$$(2.4) \quad I_n^\beta \begin{pmatrix} 1, 0, 0 \\ \frac{1}{2}, \frac{1}{2}, 0 \end{pmatrix} = 1.$$

Lemma 1. The function

$$(2.5) \quad I^\beta \begin{pmatrix} x, y, z \\ l, m, n \end{pmatrix} = (x + y + z)^\beta (l + m + n)^{1-\beta} I_3^\beta \begin{pmatrix} \frac{x}{x+y+z}, \frac{y}{x+y+z}, \frac{z}{x+y+z} \\ \frac{l}{l+m+n}, \frac{m}{l+m+n}, \frac{n}{l+m+n} \end{pmatrix}$$

($x, y, z \geq 0, x + y + z > 0, l, m, n \geq 0, l + m + n > 0$) $\beta \neq 1, \beta > 0$ satisfies the information equation (1.3).

Proof. Set $n = 3$ in Postulate 1,

$$(2.6) \quad I_3^\beta \begin{pmatrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \end{pmatrix} = I_2^\beta \begin{pmatrix} p_1 + p_2, p_3 \\ q_1 + q_2, q_3 \end{pmatrix} + (p_1 + p_2)^\beta (q_1 + q_2)^{1-\beta} I_2^\beta \begin{pmatrix} \frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2} \end{pmatrix}$$

Letting $p_1 = p_2 = \frac{1}{2}, p_3 = 0$

$$q_1 = q_2 = \frac{1}{2}, q_3 = 0$$

and then $p_1 = p_3 = \frac{1}{2}, p_2 = 0$

$$q_1 = q_3 = \frac{1}{2}, q_2 = 0$$

and using Postulate 2 we get

$$(2.7) \quad I_2^\beta \begin{pmatrix} 1, 0 \\ 1, 0 \end{pmatrix} = 0.$$

Next setting $p_3 = 0, q_3 = 0$ in (2.6), we get

$$(2.8) \quad I_3^\beta \begin{pmatrix} p_1, p_2, 0 \\ q_1, q_2, 0 \end{pmatrix} = I_2^\beta \begin{pmatrix} p_1 + p_2, 0 \\ q_1 + q_2, 0 \end{pmatrix} + (p_1 + p_2)^\beta (q_1 + q_2)^{1-\beta} \cdot I_2^\beta \begin{pmatrix} \frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2} \\ \frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2} \end{pmatrix}$$

which in accordance with $p_1 + p_2 = 1 = q_1 + q_2$ and (2.7) yields

$$(2.9) \quad I_3^\beta \left(\begin{matrix} p_1, p_2, 0 \\ q_1, q_2, 0 \end{matrix} \right) = I_2^\beta \left(\begin{matrix} p_1, p_2 \\ q_1, q_2 \end{matrix} \right).$$

Therefore (2.6), on using (2.9) takes the form

$$(2.10) \quad I_3^\beta \left(\begin{matrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \end{matrix} \right) = I_3^\beta \left(\begin{matrix} p_1 + p_2, 0, p_3 \\ q_1 + q_2, 0, q_3 \end{matrix} \right) + (p_1 + p_2)^\beta (q_1 + q_2)^{1-\beta} \cdot \\ \cdot I_3^\beta \left(\begin{matrix} \frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}, 0 \\ \frac{q_1}{q_1 + q_2}, \frac{q_2}{q_1 + q_2}, 0 \end{matrix} \right).$$

Next setting

$$p_1 = \frac{x}{x + y + z}, \quad p_2 = \frac{y}{x + y + z}, \quad p_3 = \frac{z}{x + y + z}$$

and

$$q_1 = \frac{l}{l + m + n}, \quad q_2 = \frac{m}{l + m + n}, \quad q_3 = \frac{n}{l + m + n}$$

in (2.10), we get on simplification

(2.11)

$$I_3^\beta \left(\begin{matrix} \frac{x}{x + y + z}, \frac{y}{x + y + z}, \frac{z}{x + y + z} \\ \frac{l}{l + m + n}, \frac{m}{l + m + n}, \frac{n}{l + m + n} \end{matrix} \right) = \frac{1}{(x + y + z)^\beta (l + m + n)^{1-\beta}} \cdot \\ \cdot \left[(x + y + z)^\beta (l + m + n)^{1-\beta} I_3^\beta \left(\begin{matrix} \frac{x + y}{x + y + z}, 0, \frac{z}{x + y + z} \\ \frac{l + m}{l + m + n}, 0, \frac{n}{l + m + n} \end{matrix} \right) + \right. \\ \left. + (x + y)^\beta (l + m)^{1-\beta} I_3^\beta \left(\begin{matrix} \frac{x}{x + y}, \frac{y}{x + y}, 0 \\ \frac{l}{l + m}, \frac{m}{l + m}, 0 \end{matrix} \right) \right].$$

Finally using the functional relation (2.5), we mark that $I^\beta \left(\begin{matrix} x, y, z \\ l, m, n \end{matrix} \right)$ satisfies (1.3) and this proves Lemma 1. \square

Next we prove the main theorem which relaxes the regularity condition.

Theorem 1. The symmetric and homogeneous solution of type β , of (1.3) satisfying Postulate 3 and Postulate 4 is given by

(2.12)

$$I^\beta \begin{pmatrix} x, y, z \\ l, m, n \end{pmatrix} = A_\beta [x^\beta l^{1-\beta} + y^\beta m^{1-\beta} + z^\beta n^{1-\beta} - (x + y + z)^\beta (l + m + n)^{1-\beta}]$$

where $A_\beta = (2^{\beta-1} - 1)^{-1}$.

Proof. By homogeneity, we have

$$(2.13) \quad I^\beta \begin{pmatrix} x, 0, 0 \\ l, 0, 0 \end{pmatrix} = x^\beta l^{1-\beta} I_3^\beta \begin{pmatrix} 1, 0, 0 \\ 1, 0, 0 \end{pmatrix} = 0$$

$x, l > 0, \quad \beta \neq 1, \quad \beta > 0.$

Define a function $f: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ such that

$$(2.14) \quad f(x; l) = I^\beta \begin{pmatrix} 1-x, x, 0 \\ 1-l, l, 0 \end{pmatrix} =$$

$$= \lambda^\beta \mu^{1-\beta} I^\beta \begin{pmatrix} \frac{1-x}{\lambda}, \frac{x}{\lambda}, 0 \\ \frac{1-l}{\mu}, \frac{l}{\mu}, 0 \end{pmatrix} = \lambda^\beta \mu^{1-\beta} I^\beta \begin{pmatrix} \frac{1-x}{\lambda}, 0, \frac{x}{\lambda} \\ \frac{1-l}{\mu}, 0, \frac{l}{\mu} \end{pmatrix}$$

$x \in [0, 1], \quad l \in [0, 1], \quad \lambda, \mu > 0.$

With this substitution on the right hand side of (1.3) and then using symmetry, we get

$$(x + y + z)^\beta (l + m + n)^{1-\beta} f\left(\frac{z}{x + y + z}; \frac{n}{l + m + n}\right) +$$

$$+ (x + y)^\beta (l + m)^{1-\beta} f\left(\frac{y}{x + y}; \frac{m}{l + m}\right) =$$

$$= (x + y + z)^\beta (l + m + n)^{1-\beta} f\left(\frac{y}{x + y + z}; \frac{m}{l + m + n}\right) +$$

$$+ (x + z)^\beta (l + n)^{1-\beta} f\left(\frac{z}{x + z}; \frac{n}{l + n}\right)$$

or

(2.15)

$$f\left(\frac{z}{x + y + z}; \frac{n}{l + m + n}\right) + \left(\frac{x + y}{x + y + z}\right)^\beta \left(\frac{l + m}{l + m + n}\right)^{1-\beta} f\left(\frac{y}{x + y}; \frac{m}{l + m}\right) =$$

$$= f\left(\frac{y}{x + y + z}; \frac{m}{l + m + n}\right) + \left(\frac{x + z}{x + y + z}\right)^\beta \left(\frac{l + n}{l + m + n}\right)^{1-\beta} f\left(\frac{z}{x + z}; \frac{n}{l + n}\right)$$

Putting

$$a = \frac{z}{x+y+z}, \quad b = \frac{y}{x+y+z}, \quad \alpha = \frac{n}{l+m+n}, \quad \theta = \frac{m}{l+m+n}$$

(2.15) becomes

$$(2.16) \quad f(a; \alpha) + (1-a)^\beta (1-\alpha)^{1-\beta} f\left(\frac{b}{1-a}; \frac{\theta}{1-\alpha}\right) = \\ = f(b; \theta) + (1-b)^\beta (1-\theta)^{1-\beta} f\left(\frac{a}{1-b}; \frac{\alpha}{1-\theta}\right)$$

which is a functional equation which has the following solution

$$(2.17) \quad f(a; \alpha) = A_\beta [a^\beta \alpha^{1-\beta} + (1-a)^\beta (1-\alpha)^{1-\beta} - 1], \quad \beta > 0, \quad \beta \neq 1$$

(cf. [8]) under the boundary conditions

$$(2.18) \quad f(1; 1) = f(0; 0)$$

and

$$(2.19) \quad f(1, \frac{1}{2}) = f(0; \frac{1}{2}) = 1.$$

From (2.18) we have

$$(2.20) \quad A_\beta = (2^{1-\beta} - 1)^{-1}.$$

Next (2.14) and (2.17) gives

$$(2.21) \quad I^\beta \left(\frac{1-a, a, 0}{1-\alpha, \alpha, 0} \right) = A_\beta [a^\beta \alpha^{1-\beta} + (1-a)^\beta (1-\alpha)^{1-\beta} - 1].$$

Returning to the substitution,

$$a = \frac{z}{x+y+z}, \quad b = \frac{y}{x+y+z}, \quad \alpha = \frac{n}{l+m+n}, \quad \theta = \frac{m}{l+m+n}$$

(2.21) takes the form

$$I^\beta \left(\frac{\frac{x+y}{x+y+z}, \frac{z}{x+y+z}, 0}{\frac{l+m}{l+m+n}, \frac{n}{l+m+n}, 0} \right) = \\ = A_\beta \left[\left(\frac{z}{x+y+z} \right)^\beta \left(\frac{n}{l+m+n} \right)^{1-\beta} + \left(\frac{x+y}{x+y+z} \right)^\beta \left(\frac{l+m}{l+m+n} \right)^{1-\beta} - 1 \right]$$

or

$$(2.22) \quad I^\beta \left(\frac{x+y, z, 0}{l+m, n, 0} \right) = A_\beta [z^\beta n^{1-\beta} + (x+y)^\beta (l+m)^{1-\beta} - \\ - (x+y+z)^\beta (l+m+n)^{1-\beta}].$$

Also (2.22) and (2.2) gives

$$(2.23) \quad I^\beta \begin{pmatrix} x, y, 0 \\ l, m, 0 \end{pmatrix} = A_\beta [x^\beta l^{1-\beta} + y^\beta m^{1-\beta} - (x+y)^\beta (l+m)^{1-\beta}].$$

Finally (2.22), (2.23), (2.2) and (2.5) with (1.3) give the required result i.e.

$$I^\beta \begin{pmatrix} x, y, z \\ l, m, n \end{pmatrix} = A_\beta [x^\beta l^{1-\beta} + y^\beta m^{1-\beta} + z^\beta n^{1-\beta} - (x+y+z)^\beta (l+m+n)^{1-\beta}]$$

where $A_\beta = (2^{\beta-1} - 1)^{-1}$.

This completes the proof of the Theorem 1. \square

3. APPLICATIONS TO INFORMATION THEORY

Theorem 2. Let $P = (p_1, p_2, \dots, p_n) \in \mathcal{A}_n$ and $Q = (q_1, q_2, \dots, q_n) \in \mathcal{A}_n$ be two complete probability distributions; their directed divergence of type β satisfying Postulates 1–5 is given by

$$(3.1) \quad I_n^\beta \begin{pmatrix} p_1, p_2, \dots, p_n \\ q_1, q_2, \dots, q_n \end{pmatrix} = (2^{\beta-1} - 1)^{-1} \left[\sum_{i=1}^n p_i^\beta q_i^{1-\beta} - 1 \right]$$

where $\beta \neq 1, \beta > 0$.

Proof. For probability distributions $(p_1, p_2, p_3) \in \mathcal{A}_3$ and $(q_1, q_2, q_3) \in \mathcal{A}_3$, $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$, we have from (2.12)

$$(3.2) \quad I_3^\beta \begin{pmatrix} p_1, p_2, p_3 \\ q_1, q_2, q_3 \end{pmatrix} = A_\beta [p_1^\beta q_1^{1-\beta} + p_2^\beta q_2^{1-\beta} + p_3^\beta q_3^{1-\beta} - 1].$$

Also from (2.9), when $p_1 + p_2 = 1, q_1 + q_2 = 1$, we have

$$(3.3) \quad I_2^\beta \begin{pmatrix} p_1, p_2 \\ q_1, q_2 \end{pmatrix} = A_\beta [p_1^\beta q_1^{1-\beta} + p_2^\beta q_2^{1-\beta} - 1].$$

Applying the mathematical induction, we get the required result i.e. (3.1). Hence the theorem. \square

Theorem 3. Let $P = (p_1, p_2, \dots, p_m), \sum_{i=1}^m p_i = 1, Q = (q_1, q_2, \dots, q_n), \sum_{j=1}^n q_j = 1$ and $Q_i = (q_{i1}, q_{i2}, \dots, q_{in}), \sum_{j=1}^n q_{ij} = 1, i = 1, 2, \dots, m$ be the probability distributions, then we have

$$(3.4) \quad I_n^\beta \begin{pmatrix} \sum_{i=1}^m p_i q_{i1}, & \sum_{i=1}^m p_i q_{i2}, & \dots, & \sum_{i=1}^m p_i q_{in} \\ q_1, & q_2, & \dots, & q_n \end{pmatrix} \leq \sum_{i=1}^m p_i I_n^\beta \begin{pmatrix} q_{i1}, q_{i2}, \dots, q_{in} \\ q_1, q_2, \dots, q_n \end{pmatrix}$$

Proof.

$$I_n^\beta \left(\sum_{i=1}^m p_i q_{i1}, \dots, \sum_{i=1}^m p_i q_{in} \right) = (2^{\beta-1} - 1)^{-1} \left[\sum_{j=1}^n \left(\sum_{i=1}^m p_i q_{ij} \right)^\beta q_j^{1-\beta} - 1 \right].$$

Refer, [3] (p. 532)

$$(3.5) \quad \left(\sum_{i=1}^m p_i q_{ij} \right)^\beta \geq \sum_{i=1}^m p_i q_{ij}^\beta, \quad \text{for } \beta < 1 \\ \leq \sum_{i=1}^m p_i q_{ij}^\beta, \quad \text{for } \beta > 1.$$

Multiplying by $q_j^{1-\beta}$ and summing over all j 's, we have

$$\sum_{j=1}^n \left(\sum_{i=1}^m p_i q_{ij} \right)^\beta q_j^{1-\beta} \geq \sum_{j=1}^n \left(\sum_{i=1}^m p_i q_{ij}^\beta \right) q_j^{1-\beta}$$

according as $\beta \geq 1$. However since $(2^{\beta-1} - 1) \geq 0$ according as $\beta \geq 1$, we have when $\beta \neq 1$

$$(3.6) \quad (2^{\beta-1} - 1)^{-1} \left[\sum_{j=1}^n \left(\sum_{i=1}^m p_i q_{ij} \right)^\beta q_j^{1-\beta} - 1 \right] \leq \\ \leq (2^{\beta-1} - 1)^{-1} \left[\sum_{i=1}^m \sum_{j=1}^n p_i q_{ij}^\beta q_j^{1-\beta} - 1 \right].$$

Thus

$$I_n^\beta \left(\sum_{i=1}^m p_i q_{i1}, \dots, \sum_{i=1}^m p_i q_{in} \right) \leq (2^{\beta-1} - 1)^{-1} \sum_{i=1}^m p_i \left[\sum_{j=1}^n q_{ij}^\beta q_j^{1-\beta} - 1 \right] = \\ = \sum_{i=1}^m p_i I_n^\beta \left(q_{i1}, q_{i2}, \dots, q_{in} \right).$$

This completes the proof of the theorem. \square

We conclude that the characterization of directed divergence under homogeneity condition have applications in Mathematical Economics, Production Theory and Utility Theory.

ACKNOWLEDGEMENT

Authors are thankful to Prof. B. D. Sharma, Department of Mathematics, The University of the West Indies, St. Augustine, Trinidad (W.I.) for guidance in the preparation of this paper.

(Received September 9, 1982.)

REFERENCES

- [1] J. Aczel: Results on entropy equation. *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 25 (1977), 12–17.
- [2] L. L. Campbell: Characterization of Entropy in Arbitrary Probability Spaces. Preprint No. 1970–32, Queen's University, Kingston, Ontario 1970.
- [3] D. G. Gallager: *Information Theory and Reliable Communication*. John Wiley and Sons, New York 1968, 523–524.
- [4] A. Kamiński and J. Mikusiński: On the entropy equation. *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 22 (1974), 319–323.
- [5] S. Kullback: *Information Theory and Statistics*. John Wiley and Sons, New York 1959.
- [6] P. L. Kannappan and P. N. Rathie: An application of a functional equation to information theory. *Ann. Polon. Math.* XXVI (1972).
- [7] D. F. Kerridge (1961): Inaccuracy and inference. *J. Roy. Statist. Soc. Ser. A* 23 (1961), 184–194.
- [8] P. N. Rathie and P. L. Kannappan: A directed divergence function of type β . *Inform. and Control* 20 (1972), 38–45.
- [9] A. Rényi: On measures of entropy and information. *Proc. Fourth Berkeley Symposium Math. Statist. and Probability I* (1961), 547–561.
- [10] B. D. Sharma and Ram Autar: Relative information functions and their type (α, β) generalizations. *Metrika* 21 (1974), 41–50.
- [11] B. D. Sharma and R. S. Soni: A new generalized functional equation for inaccuracy and entropy of kind β . *Funkcial. Ekvac.* 17 (1974), 1–11.
- [12] H. Theil: *Economics and Information Theory*. North-Holland, Amsterdam 1967.

Dr. R. P. Singh, Department of Mathematics, L. R. College, Sahibabad, Ghaziabad - 201005. India.

Dr. R. K. Khanna, SAG, R and D Organization, Ministry of Defence, Delhi - 110054. India.