Ramji P. Singh; Ravindar Kumar Khanna
On characterization of directed divergence of type $\beta$ through information equation


Persistent URL: [http://dml.cz/dmlcz/124490](http://dml.cz/dmlcz/124490)

Terms of use:

© Institute of Information Theory and Automation AS CR, 1984

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these Terms of use.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project DML-CZ: The Czech Digital Mathematics Library [http://project.dml.cz](http://project.dml.cz)
ON CHARACTERIZATION OF DIRECTED DIVERGENCE
OF TYPE $\beta$ THROUGH INFORMATION EQUATION

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

The directed-divergence of type $\beta$ ($\beta > 0$, $\beta \neq 1$) 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 $\beta$ to Information Theory have been discussed.

I. 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 $\beta$ ($\beta > 0$, $\beta \neq 1$) 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:

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

Let the $Q = (q_1, q_2, \ldots, q_n), q_i \geq 0, \sum_{i=1}^{n} q_i = 1$ 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

$$I_{\alpha}(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i^{-\alpha}}$$

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

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

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

Here we consider the 'information equation' given by

\[ I(x, y, z; l, m, n) = I(x + y, 0, n) + I(l, m, 0) \] (1.3)

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.

\[ H(x, y, z) = H(x + y, 0, z) + H(x, y, 0) \] (1.4)

\( (x, y, z \geq 0, xy + yz + zx > 0) \).

The homogeneity condition considered here is defined as follows:

\[ I^\lambda \left( x_1, y_1, \ldots, x_n, y_n \right) = \lambda I(x_1, y_1, \ldots, x_n, y_n) \] (1.5)

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

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 \( \beta \) under the homogeneity condition (1.5). Let the measure (1.2) satisfy the following postulates:

**Postulate 1.** Branching property i.e.

\[ I^{\lambda \mu} (p_1, p_2, \ldots, p_n) = I^{\lambda \mu} (p_1, p_2, \ldots, p_n) + (p_1 + p_2) \left( I^{\lambda \mu} (p_1, q_1 + q_2) \right) \] (2.1)

**Postulate 2.** Symmetry i.e.

\[ I^\beta (p_1, p_2, \ldots, p_n) = I^\beta (p_{k(1)}, p_{k(2)}, \ldots, p_{k(n)}) \] (2.2)

where \( k(1), k(2), \ldots, k(n) \) is a permutation of \( 1, 2, \ldots, n \).
Postulate 3. Nullity i.e.

\[ I^e \left( 1, 0, 0 \right) = 0. \]  

Postulate 4. Unit i.e.

\[ I^e \left( 1, 0, 0 \right) = 1. \]

Lemma 1. The function

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

\[(x, y, z \geq 0, x + y + z > 0, l, m, n \geq 0, l + m + n > 0; \beta + 1, \beta > 0)\]

satisfies the information equation (1.3).

Proof. Set \( n = 3 \) in Postulate 1,

\[ I^e \left( p_1, p_2, p_3 \right) = I^e \left( p_1 + p_2, p_3 \right) + \]

\[ + (p_1 + p_2)^{\beta} \left( q_1 + q_2 \right)^{1-\beta} I^e \left( \frac{p_1}{q_1 + q_2}, \frac{p_2}{q_1 + q_2} \right) \]

Letting \( p_1 = p_2 = \frac{1}{2}, p_3 = 0 \)
and then \( q_1 = q_2 = \frac{1}{2}, q_3 = 0 \)
and using Postulate 2 we get

\[ I^e \left( 1, 0, 0 \right) = 0. \]  

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

\[ I^e \left( p_1, p_2, 0 \right) = I^e \left( p_1 + p_2, 0 \right) + (p_1 + p_2)^{\beta} \left( q_1 + q_2 \right)^{1-\beta} \]

\[ \left( \frac{p_1}{q_1 + q_2}, \frac{p_2}{q_1 + q_2} \right) \]
which in accordance with $p_1 + p_2 = 1 = q_1 + q_2$ and (2.7) yields

\[(2.9) \quad P_5 \left( \frac{p_1, p_2, 0}{q_1, q_2, 0} \right) = P_5 \left( \frac{p_1, p_2}{q_1, q_2} \right).\]

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

\[(2.10) \quad P_5 \left( \frac{p_1, p_2, p_3}{q_1, q_2, q_3} \right) = P_5 \left( \frac{p_1 + p_2, 0, p_3}{q_1 + q_2, 0, q_3} \right) + (p_1 + p_2)^\theta (q_1 + q_2)^{1-\theta}.

Next setting

\[\begin{align*}
p_1 &= \frac{x}{x + y + z}, & p_2 &= \frac{y}{x + y + z}, & p_3 &= \frac{z}{x + y + z}, \\
q_1 &= \frac{1}{l + m + n}, & q_2 &= \frac{m}{l + m + n}, & q_3 &= \frac{n}{l + m + n}
\end{align*}\]

in (2.10), we get on simplification

\[(2.11) \quad P_5 \left( \frac{x}{l + m + n}, \frac{y}{l + m + n}, \frac{z}{l + m + n} \right) = \frac{1}{(x + y + z)^\theta (l + m + n)^{1-\theta}} \left[ (x + y + z)^\theta (l + m + n)^{1-\theta} P_5 \left( \frac{x + y}{l + m + n}, \frac{z}{x + y + z}, 0 \right) \right. \]

\[\left. + (x + y)^\theta (l + m)^{1-\theta} P_5 \left( \frac{x}{l + m}, \frac{y}{l + m}, 0 \right) \right].\]

Finally using the functional relation (2.5), we mark that $P_5 \left( \frac{x, y, z}{l, m, n} \right)$ satisfies (1.3) and this proves Lemma 1.

Next we prove the main theorem which relaxes the regularity condition.
Theorem 1. The symmetric and homogeneous solution of type \( \beta \), of (1.3) satisfying Postulate 3 and Postulate 4 is given by

\[
(2.12)
\]

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

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

Proof. By homogeneity, we have

\[
(2.13)
\]

\[
\phi^{(1)}(x, 0, 0) = x^{\beta m} \phi^{(1)}(1, 0, 0) = 0
\]

\[
x, \ l > 0, \ \beta + 1, \ \beta > 0.
\]

Define a function \( f : [0, 1] \times [0, 1] \to \mathbb{R} \) such that

\[
(2.14)
\]

\[
f(x; l) = \phi^{(1)} \left( \frac{1-x}{1-l}, l, 0 \right) = \frac{x^{\lambda}}{\lambda} \mu^{1-\beta} \phi^{(1)} \left( \frac{1-x}{1-l}, l, 0 \right)
\]

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

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

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

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

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

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

or

\[
(2.15)
\]

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

\[
= f \left( \frac{y}{x + y + z} ; \frac{m}{l + m + n} \right) + \left( \frac{x + z}{x + y + z} \right)^{\beta} \left( \frac{l + n}{l + m + n} \right)^{1-\beta} f \left( \frac{z}{x + z} ; \frac{n}{l + n} \right)
\]
Putting
\[ a = \frac{z}{x + y + z}, \quad b = \frac{y}{x + y + z}, \quad x = \frac{n}{l + m + n}, \quad \theta = \frac{m}{l + m + n} \]

(2.15) becomes
\[
(2.16) \quad f(a; x) + (1 - a)\beta (1 - a)^{1 - \beta} f\left(\frac{b}{1 - a}; \frac{\theta}{1 - \theta}\right) = f(b; \theta) + (1 - b)\beta (1 - \theta)^{1 - \beta} f\left(\frac{a}{1 - b}; \frac{\alpha}{1 - \alpha}\right)
\]

which is a functional equation which has the following solution
\[
(2.17) \quad f(a; x) = A_\beta [a^\beta a^{1 - \beta} + (1 - a)^\beta (1 - a)^{1 - \beta} - 1], \quad \beta > 0, \quad \beta \neq 1
\]
(c.f. [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
\[
(2.21) \quad f^\beta \left(\frac{1 - a, a, 0}{1 - x, x, 0}\right) = A_\beta [a^\beta a^{1 - \beta} + (1 - a)^\beta (1 - x)^{1 - \beta} - 1].
\]

Returning to the substitution,
\[ a = \frac{z}{x + y + z}, \quad b = \frac{y}{x + y + z}, \quad x = \frac{n}{l + m + n}, \quad \theta = \frac{m}{l + m + n} \]

(2.21) takes the form
\[
(2.22) \quad f^\beta \left(\frac{x + y, z, 0}{l + m, n, 0}\right) = A_\beta \left[ z^\beta (x + y + z)^{1 - \beta} + (x + y)^\beta \left(\frac{l + m + n}{l + m + n}\right)^{1 - \beta} - 1 \right]
\]
or
\[
(2.22) \quad f^\beta \left(\frac{x + y, z, 0}{l + m, n, 0}\right) = A_\beta [z^\beta (x + y + z)^{1 - \beta} - (x + y)^\beta (l + m)^{1 - \beta} - (x + y + z)^\beta (l + m + n)^{1 - \beta}].
\]
Also (2.22) and (2.2) gives

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

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

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

where \( A_{\beta} = (2^{\beta - 1} - 1)^{-1} \).
This completes the proof of the Theorem 1.

3. APPLICATIONS TO INFORMATION THEORY

**Theorem 2.** Let \( P = (p_1, p_2, \ldots, p_n) \in A_n \) and \( Q = (q_1, q_2, \ldots, q_n) \in A_n \) be two complete probability distributions; their directed divergence of type \( \beta \) satisfying Postulates 1–5 is given by

\[ I_\beta (p_1, p_2, \ldots, p_n) = (2^{\beta - 1} - 1)^{-1} \left[ \sum_{i=1}^{n} p_i q_i^{1-\beta} - 1 \right] \]

where \( \beta > 0 \).

**Proof.** For probability distributions \( (p_1, p_2, p_3) \in A_3 \) and \( (q_1, q_2, q_3) \in A_3, \)
\[ \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1 \), we have from (2.12)

\[ I_\beta (p_1, p_2, p_3) = A_{\beta} [p_1 q_1^{1-\beta} + p_2 q_2^{1-\beta} + p_3 q_3^{1-\beta} - 1]. \]

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

\[ I_\beta (p_1, p_2) = A_{\beta} [p_1 q_1^{1-\beta} + p_2 q_2^{1-\beta} - 1]. \]

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

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

\[ I_\beta \left( \sum_{i=1}^{m} p_i q_{i1}, \sum_{i=1}^{m} p_i q_{i2}, \ldots, \sum_{i=1}^{m} p_i q_{im} \right) \leq \sum_{i=1}^{m} p_i I_\beta (q_{i1}, q_{i2}, \ldots, q_{im}). \]
Proof.

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

Refer, [3] (p. 532)

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

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

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

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

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

Thus

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

This completes the proof of the theorem.

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


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.