



Asian Journal of Mathematics & Statistics

ISSN 1994-5418

Some new Results on Fuzzy Directed Divergence Measures and their Inequalities

M.A.K. Baig and Mohd Javid Dar

Department of Statistics, University of Kashmir, Srinagar, 190006, India

Corresponding Author: M.A.K. Baig, Department of Statistics, University of Kashmir, Srinagar, 190006, India

ABSTRACT

In this study, we propose some new fuzzy directed divergence measure and study its particular cases. We also establish some fuzzy coding theorems. Some of the known results are the particular cases of our proposed divergence measure.

Key words: Fuzzy set, fuzzy directed divergence measure, fuzzy entropy, fuzzy average codeword length, fuzzy generalized average codeword length

INTRODUCTION

Classical information theoretic divergence measures have witnessed the need to study them. Kullback and Leibler (1951) first studied the measure of divergence. Jaynes (1957) introduced the Principle of Maximum Entropy (PME) and emphasized that "choose a distribution which is consistent to with the information available and is uniform as possible". For implementation of this consideration another advance was needed in the form of a measure of nearness of two probability distribution and it was already provided by Kullback Leibler (1951) in the form of:

$$I(P, Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} \quad (1)$$

If the distribution Q is uniform. This becomes:

$$I(P, Q) = \sum_{i=1}^n p_i \log p_i + \log n \quad (2)$$

where, $P, Q \in T_n$ and:

$$T_n = \{P = (p_1, p_2, \dots, p_n), p_i > 0, \sum_{i=1}^n p_i = 1; \forall i=1, 2, \dots, n, n \geq 2\}$$

Since (Shannon, 1948) Entropy:

$$H(P) = - \sum_{i=1}^n p_i \log p_i \quad (3)$$

was already available in the literature, so maximizing H is equivalent to minimizing $I(P, Q)$. This is one of the interpretations of PME.

Analyzing Eq. 1 in the following way:

$$I(P, Q) = \sum_{i=1}^n (p_i \log p_i - p_i \log q_i) \quad (4)$$

The second term present in Eq. 4 is called the Kerridge Inaccuracy which is:

$$= - \sum_{i=1}^n p_i \log q_i \quad (5)$$

Considering (Kerridge, 1961) inaccuracy, we can interpret (Kullback and Leibler, 1951) measure of divergence as:

$$I(P, Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}$$

i.e.,

= Difference of Kerridge inaccuracy and Shannon's entropy:

$$= \sum_{i=1}^n \{-p_i \log q_i - (-p_i \log p_i)\} \quad (6)$$

Since $I(P, Q)$ provides a measure of nearness of P from Q . Take the case of Reliability Theory; here we can consider how much the information is reliable. Because the distribution is the revised distribution/strategies to achieve the goal/objective/target with certain constraints, so optimization theory takes the birth, which is the need of every one.

Hence whenever we come across divergence measures, we are interested to minimize the divergence to make the information available, reliable. Every walk of life is governed with the reliability of information under certain constraints.

Analogous to information theoretic approach, when we arrive at fuzzy sets or fuzziness, we need to study fuzzy divergence measures. As presently, the vast applications of fuzzy information in life and social sciences, interpretational communication, Engineering, Fuzzy Aircraft Control, Medicine, Management and decision making, Computer Sciences, Pattern Recognition and Clustering. Hence the wide applications motivate us to consider Divergence Measures for fuzzy set theory to minimize or maximize or optimize the fuzziness.

Let $A = \{x_i: \mu_A(x_i), \forall i = 1, 2, \dots, n\}$ and $B = \{x_i: \mu_B(x_i), \forall i = 1, 2, \dots, n\}$ where, $0 < \mu_A(x_i) < 1$ and $0 < \mu_B(x_i) < 1$, be two fuzzy sets. The fuzzy divergence corresponding to Kullback and Leibler (1951) has been defined by Bhandari and Pal (1993) as:

$$D(A||B) = \sum_{i=1}^n \left[\mu_A(x_i) \log \frac{\mu_A(x_i)}{\mu_B(x_i)} + \{1 - \mu_A(x_i)\} \log \frac{1 - \mu_A(x_i)}{1 - \mu_B(x_i)} \right] \quad (7)$$

The fundamental properties of fuzzy divergence are as follows:

- Non-negativity, i.e., $D(A||B) \geq 0$
- $D(A||B) = 0$, if $A = B$

- $D(A||B)$ is a convex function in $(0, 1)$
- $D(A||B)$ should not change, when $\mu_A(x_i)$ is changed to $1-\mu_A(x_i)$ and $\mu_B(x_i)$ to $1-\mu_B(x_i)$

Bhandari and Pal (1993) have established some properties such as:

- $D(A||B) = I(A||B) + I(B||A)$
Where:

$$I(A||B) = [\mu_A(x_i) - \mu_B(x_i)] \log \frac{\mu_A(x_i)}{\mu_B(x_i)}$$

- $D(A \cup B || A \cap B) = D(A||B)$
- $D(A \cup B || C) \leq D(A||C) + D(B||C)$
- $D(A||B) \geq D(A \cup B || A)$
- $D(A||B)$ is maximum if B is the farthest non-fuzzy set of A

Havrda and Charvat (1967) has given the measure of directed divergence as:

$$D_\alpha(P, Q) = \frac{1}{\alpha(\alpha-1)} (\sum_{i=1}^n p_i^\alpha q_i^{1-\alpha} - 1) \tag{8}$$

Corresponding to Eq. 8, the average code word length can be taken as:

$$L_\alpha(P, Q) = \frac{1}{\alpha(\alpha-1)} (\sum_{i=1}^n p_i q_i D^{(\alpha-1)n_i} - 1) \tag{9}$$

Corresponding to Eq. 8, the fuzzy measure of directed divergence between two fuzzy sets $\mu_A(x_i)$ and $\mu_B(x_i)$ can taken as:

$$D_\alpha = \frac{1}{\alpha(\alpha-1)} \left(\sum_{i=1}^n \left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{(\alpha-1)n_i} - 1 \right)$$

And its corresponding fuzzy average code word length as:

$$L_\alpha = \frac{1}{\alpha(\alpha-1)} \left(\sum_{i=1}^n \left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{(\alpha-1)n_i} - 1 \right)$$

Remark:

- As $\alpha \rightarrow 1$, Eq. 8 tends to Eq. 1
- As $\alpha \rightarrow 1$ and $q_i = 1$, Eq. 8 tends to Eq. 3
- As $\alpha \rightarrow 1$ Eq. 9 tends to average codeword length given as:

$$L = \sum_{i=1}^n p_i q_i n_i \tag{10}$$

- As $\alpha \rightarrow 1$ and $q_i = 1$, Eq. 9 tends to average codeword length corresponding to Shannon's entropy given as:

$$L = \sum_{i=1}^n p_i n_i \tag{11}$$

NOISELESS DIRECTED DIVERGENCE CODING THEOREMS

Theorem 1: For all uniquely decipherable codes:

$$D_\alpha \leq L_\alpha \tag{12}$$

Where:

$$L_\alpha = \frac{1}{\alpha(\alpha-1)} \left(\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} D^{(\alpha-1)n_i} - 1 \right)$$

Proof: By Holders inequality, we have:

$$\sum_{i=1}^n x_i y_i \geq \left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n y_i^q \right)^{\frac{1}{q}}, 0 < p < 1, q < 0 \text{ or } 0 < q < 1, p < 0 \tag{13}$$

Set:

$$x_i = \left[f(\mu_A(x_i), \mu_B(x_i)) \right]^{\frac{-1}{t}} D^{-n_i}$$

$$y_i = \left[f(\mu_A(x_i), \mu_B(x_i)) \right]^{\frac{1}{t}}$$

and:

$$p = -t \Rightarrow 0 < p < 1, q = \frac{t}{t+1} \Rightarrow q < 0$$

Thus Eq. 13 becomes:

$$\sum_{i=1}^n \left[\left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\}^{\frac{-1}{t}} D^{-n_i} \left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\}^{\frac{1}{t}} \right] \geq \left[\sum_{i=1}^n \left\{ \left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\}^{\frac{-1}{t}} D^{-n_i} \right\}^{-t} \right]^{\frac{-1}{t}} \left[\sum_{i=1}^n \left\{ \left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\}^{\frac{1}{t}} \right\}^{\frac{t+1}{t}} \right]^{\frac{1}{t+1}}$$

Using Kraft's inequality, we have:

$$\left[\sum_{i=1}^n \left\{ \left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} \right\}^{\frac{t+1}{t}} \right]^{\frac{-1}{t}} \leq \left[\sum_{i=1}^n \left\{ \left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} D^{-n_i} \right\}^{-t} \right]^{\frac{-1}{t}}$$

or:

$$\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} \right]^{\frac{1}{t}} \leq \sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} D^{n_i} \right]^{\frac{-1}{t}}$$

or:

$$\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} \right]^{\frac{1}{t}} \leq \sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} D^{n_i} \right]^{\frac{-1}{t}} \tag{14}$$

dividing both sides by t, we get:

$$\frac{\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} \right]^{\frac{1}{t}}}{t} \leq \frac{\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} D^{n_i} \right]^{\frac{-1}{t}}}{t}$$

Subtracting n from both sides, we have:

$$\frac{\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} \right]^{\frac{1}{t}} - n}{t} \leq \frac{\sum_{i=1}^n \left[\left(f(\mu_A(x_i), \mu_B(x_i)) \right)^{\frac{-1}{t}} D^{n_i} \right]^{\frac{-1}{t}} - n}{t} \tag{15}$$

Taking $\alpha = t+1$, $t = \alpha-1$ and:

$$f(\mu_A(x_i), \mu_B(x_i)) = \left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\}$$

Equation 15 becomes:

$$\frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} - 1 \right]}{\alpha-1} \leq \frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{n_i(\alpha-1)} - 1 \right]}{\alpha-1} \tag{16}$$

Dividing both sides by α , we get:

$$\frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} - 1 \right]}{\alpha(\alpha-1)} \leq \frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} - 1 \right]}{\alpha(\alpha-1)}$$

that is $D_{\alpha} \leq L_{\alpha}$ which proves the theorem.

Theorem 2: For all uniquely decipherable codes:

$$D_{\alpha, \beta} \leq L_{\alpha, \beta} \tag{17}$$

Where:

$$L_{\alpha, \beta} = \frac{1}{\beta - \alpha} \sum_{i=1}^n \left[\frac{\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} D^{n(\alpha-1)}}{\left\{ \left(\mu_A(x_i)^{\beta} + (1 - \mu_A(x_i))^{\beta} \right) \left(\mu_B(x_i)^{\beta} + (1 - \mu_B(x_i))^{\beta} \right) \right\} D^{n(\beta-1)}} \right] \tag{18}$$

where, either $\alpha \geq 1, \beta \leq 1$ or $\alpha \leq 1, \beta \geq 1$

Proof: Since from Eq. 16, we have:

$$\frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} - 1 \right]}{\alpha - 1} \leq \frac{\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} D^{n(\alpha-1)} - 1 \right]}{\alpha - 1} \tag{19}$$

Multiplying both sides by $(\alpha-1)$, we get:

$$\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} - 1 \right] \leq \sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} D^{n(\alpha-1)} - 1 \right] \tag{20}$$

Changing α to β , Eq. 20 becomes:

$$\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\beta} + (1 - \mu_A(x_i))^{\beta} \right) \left(\mu_B(x_i)^{\beta} + (1 - \mu_B(x_i))^{\beta} \right) \right\} - 1 \right] \leq \sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\beta} + (1 - \mu_A(x_i))^{\beta} \right) \left(\mu_B(x_i)^{\beta} + (1 - \mu_B(x_i))^{\beta} \right) \right\} D^{n(\beta-1)} - 1 \right] \tag{21}$$

Subtracting Eq. 21 from 20, we have:

$$\sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} - \left\{ \left(\mu_A(x_i)^{\beta} + (1 - \mu_A(x_i))^{\beta} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\beta} \right) \right\} \right] \leq \sum_{i=1}^n \left[\left\{ \left(\mu_A(x_i)^{\alpha} + (1 - \mu_A(x_i))^{\alpha} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\alpha} \right) \right\} D^{n(\alpha-1)} - \left\{ \left(\mu_A(x_i)^{\beta} + (1 - \mu_A(x_i))^{\beta} \right) \left(\mu_B(x_i)^{\alpha} + (1 - \mu_B(x_i))^{\beta} \right) \right\} D^{n(\beta-1)} \right]$$

Dividing both sides by $(\beta-\alpha)$, we have:

$$\frac{1}{\beta-\alpha} \sum_{i=1}^n \left[\frac{\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} - 1}{\left\{ \left(\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\beta \right) \right\}} \right] \leq$$

$$\frac{1}{\beta-\alpha} \sum_{i=1}^n \left[\frac{\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{n(\alpha-1)}}{\left\{ \left(\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\beta \right) \right\} D^{n(\beta-1)}} \right]$$

That is $D_{\alpha, \beta} \leq L_{\alpha, \beta}$ which proves the theorem.

Theorem 3: For all uniquely decipherable codes:

$$D'_{\alpha, \beta} \leq L'_{\alpha, \beta}$$

where:

$$D'_{\alpha, \beta} = \frac{1}{\beta-\alpha} \left[\frac{\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \right\} - 1}{\left\{ \left(\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \left(\mu_B(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \right\} - 1} \right] \tag{22}$$

and:

$$L'_{\alpha, \beta} = \frac{1}{\beta-\alpha} \left[\frac{\left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{n(\alpha-1)} - 1}{\left\{ \left(\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \left(\mu_B(x_i)^\beta + (1-\mu_A(x_i))^\beta \right) \right\} D^{n(\beta-1)} - 1} \right] \tag{23}$$

To prove this theorem, we first prove the following lemma:

Lemma 1: For all uniquely decipherable codes:

$$\left[\sum_{i=1}^n \left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} - 1 \right]$$

$$\leq \left[\sum_{i=1}^n \left\{ \left(\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha \right) \left(\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha \right) \right\} D^{n(\alpha-1)} - 1 \right]$$

Proof of the lemma: From (3) we have:

$$\sum_{i=1}^n \left[\left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\} \right] \leq \sum_{i=1}^n \left[\left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\} D^{n\alpha} \right]$$

Subtracting 'n' from both sides, we have:

$$\sum_{i=1}^n \left[\left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\} - 1 \right] \leq \sum_{i=1}^n \left[\left\{ f(\mu_A(x_i), \mu_B(x_i)) \right\} D^{n\alpha} - 1 \right]$$

Taking $\alpha = t+1$, $t = \alpha-1$ and:

$$f(\mu_A(x_i), \mu_B(x_i)) = \left\{ (\mu_A(x_i)^\alpha + \mu_B(x_i)^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\}$$

we have:

$$\begin{aligned} & \left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1 \right] \\ & \leq \left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} D^{n(\alpha-1)} - 1 \right] \end{aligned} \tag{24}$$

Which proves the Lemma.

Proof of the Theorem 3: Changing α to β in (Eq. 24), we get:

$$\begin{aligned} & \left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1 \right] \\ & \leq \left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\beta + (1-\mu_B(x_i))^\beta) \right\} D^{n(\beta-1)} - 1 \right] \end{aligned} \tag{25}$$

dividing Eq. 25 from 24, we have:

$$\begin{aligned} & \frac{\left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1 \right]}{\left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1 \right]} \leq \\ & \frac{\left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} D^{n(\alpha-1)} - 1 \right]}{\left[\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\beta + (1-\mu_B(x_i))^\beta) \right\} D^{n(\beta-1)} - 1 \right]} \end{aligned}$$

Dividing both sides by $\beta-\alpha$ we have:

$$\begin{aligned} & \frac{1}{\beta-\alpha} \left[\frac{\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1}{\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} - 1} \right] \leq \\ & \frac{1}{\beta-\alpha} \left[\frac{\sum_{i=1}^n \left\{ (\mu_A(x_i)^\alpha + (1-\mu_A(x_i))^\alpha) (\mu_B(x_i)^\alpha + (1-\mu_B(x_i))^\alpha) \right\} D^{n(\alpha-1)} - 1}{\sum_{i=1}^n \left\{ (\mu_A(x_i)^\beta + (1-\mu_A(x_i))^\beta) (\mu_B(x_i)^\beta + (1-\mu_B(x_i))^\beta) \right\} D^{n(\beta-1)} - 1} \right] \end{aligned}$$

that is $D'_{\alpha, \beta} \leq L'_{\alpha, \beta}$, which proves the theorem.

REFERENCES

Bhandari, D. and N.R. Pal, 1993. Some new information measures for fuzzy sets. Inform. Sci., 67: 209-228.

- Havrda, J. and F. Charvat, 1967. Quantification method of classification processes: Concept of structural α -entropy. *Kybernetika*, 3: 30-35.
- Jaynes, E.T., 1957. Information theory and statistical mechanics. *Phys. Rev.*, 106: 620-630.
- Kerridge, D.F., 1961. Inaccuracy and inference. *J. R. Stat. Soc. Ser. B*, 23: 184-194.
- Kullback, S. and R.A. Leibler, 1951. On information and sufficiency. *Ann. Math. Statist.*, 22: 1-164.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27: 379-423.