Advances and Applications in Mathematical Sciences
Volume 19, Issue 8, June 2020, Pages 683-708

A GENERALIZATION OF *J*-DIVERGENCE MEASURE BASED ON RENYI'S-TSALLIS ENTROPY WITH APPLICATION IN FAULT DETECTION

RATIKA KADIAN and SATISH KUMAR

Department of Mathematics
Maharishi Markandeshwar
(Deemed To Be University)
Mullana-Ambala 133207, India
E-mail: kadian.ratika@mmumullana.org
profsatish74@mmumullana.org

Abstract

In this communication, we have characterized the sum of two general measures associated with two distributions with discrete random variables. One of these measures is logarithmic, while other contains the power of variables, named as J-divergence based on Renyi's-Tsallis entropy measure and establish their validity as well as discuss their basic properties. To show the efficiency of proposed measure, we apply it to pattern recognition and fault detection. Some illustrative examples are given to support the findings and further exhibit their practicality and adequacy of measure between probability sets.

1. Introduction

The Shannon entropy [25] and Kullback-Leibler divergence [16] are the most significant and most generally utilized quantities in information theory. Many information-theoretic divergence measures between two probability distributions have been introduced and extensively studied [1], [8], [13], [16], [17], [18]. Because of their successful use, many attempts have been made to generalize them. It is known that their significant generalizations are the Renyi entropy and Renyi divergence [23], Tsallis entropy and Tsallis divergence [27], R-norm information measure and R-norm divergence [3] respectively. These quantities have many significant applications, for

2010 Mathematics Subject Classification: 94A15, 94A24, 26D15
Keywords: Shannon, Renyi and Tsallis divergence measures, inaccuracy measure, fault detection.
Received November 2, 2019; Accepted November 22, 2019

example, in statistics, magnetic resonance image analysis, cost sensitive classification for medical diagnosis [24], pattern recognition [6], approximation of probability distributions [7], color image segmentation , signal processing [14], biology and economics, model validation, robust detection [22], quantum information theory [21] and industrial engineering etc. We can use divergence measures in fuzzy mathematics as fuzzy directed divergences and fuzzy entropies. Also, divergence measures in which we measure average ambiguity or difficulty in deciding whether a particular component has a place with a set or not.

In the period of last few years, the literature on advances and applications of information and divergence measures between probability distributions has elongated measurably, still there is a scope that the better divergence measures can be created which will find applications in the variety of fields. Inspired by the above mentioned work, we introduce a new directed divergence that overcomes the previous difficulties and discuss the basic properties of this measure. A symmetric form of the new directed divergence is defined. The new generalized measure has graceful characteristics which are proven in the paper to depict the efficiency of the proposed measure. The proposed symmetric measure has been demonstrated by its applications in the context of pattern recognition and fault detection.

The paper is sorted out as follows: In the section 'Introduction' we have presented the work done by earlier researchers in the field. In section 2, some required introductory out-comes of new directed divergence measure is introduced and its validity is set up. Section 3 recalls a new generalized J-divergence measure and their properties with respect to relative information measure. In section 4, few applications for proposed symmetric divergence measure are introduced and a numerical example is presented to illustrate applications of pattern recognition and fault detection has been examined to determine the possible fault suffered by any machine. The performance of proposed symmetric measure is contrasted with some other existing measures in section 5. At last, the paper is finished with Concluding Remarks in segment 6.

In the following section, we review the generalization of Shannon entropy [25] and Kullback-Leibler [16] information maesure.

2. Preliminaries

Let $\Delta_k = \{E = (e_1, e_2, \ldots e_k) : e_p \geq 0, p = 1, 2, \ldots, k; \sum_{p=1}^{k} e_p = 1\}, k \geq 2$ be set of k-complete probability distributions. For any probability distributions $E = (e_1, e_2, \ldots, e_k) \in \Delta_k$, Shannon [25] defined an entropy as:

$$H(E) = -\sum_{p=1}^{k} (e_p) \log(e_p). \tag{2.1}$$

All through this paper, it is assumed (0) log (0) = 0 and all logarithms are to the base 2.

Corresponding to (2.1), for any $E, F \in \Delta_k$, Kullback and Leibler [16] defined a divergence measure as:

$$D^{K-L}(E, F) = \sum_{p=1}^{k} (e_p) \log\left(\frac{e_p}{f_p}\right). \tag{2.2}$$

It is well known that $D^{K-L}(E, F)$ is nonnegative, additive but not symmetric [17]. To obtain symmetric measure, one can define

$$J(E, F) = D^{K-L}(E, F) + D^{K-L}(F, E) = \sum_{p=1}^{k} (e_p - f_p) \log\left(\frac{e_p}{f_p}\right). \tag{2.3}$$

which is called the J divergence [18]. Clearly, D^{K-L} and J divergences share most of their properties. It should be noted that $D^{K-L}(E, F)$ is undefined if $f_p = 0$ and $e_p \neq 0$. This means that distribution E has to be absolutely continuous [17] with respect to distribution F for $D^{K-L}(E, F)$ to be defined. Similarly, J requires that E and F be absolutely continuous with respect to each other. This is one of the problems with these divergence measures.

In the next section, we present the idea of J-divergence based on Renyi's-Tsallis entropy for the probability distributions we will give basic properties of this measure and for illustration, we provide some applications in fault detection and pattern recognition in section 4.

3. A new Generalized *J*-Divergence Measure

For some $E = (e_1, e_2, \ldots, e_k) \in \Delta_k$. Recently, Litegebe and Satish [28] define a new information measure as :

$$H_{R-T}^{\alpha}(E) = \begin{cases} \frac{1}{\alpha^{-1} - \alpha}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha}\right) + \sum_{p=1}^{k} e_p^{\alpha} - 1\right], \alpha > 0 (\neq 1); \\ -\sum_{p=1}^{k} (e_p) \log(e_p), \alpha = 1. \end{cases} \tag{3.1}$$

The quantity (3.1) introduced a joint representation of Renyi's-Tsallis entropy of order α and second case is well-known Shannon entropy [25].

Further, corresponding to (3.1), we define an inaccuracy measure as:

$$H_{R-T}^{\alpha}(E/F) = \begin{cases} \frac{1}{\alpha^{-1} - \alpha}\left[\log\left(\sum_{p=1}^{k} e_p f_p^{\alpha-1}\right) + \sum_{p=1}^{k} e_p f_p^{\alpha-1} - 1\right], \alpha > 0 (\neq 1); \\ -\sum_{p=1}^{k} (e_p) \log(f_p), \alpha = 1. \end{cases} \tag{3.2}$$

Remark 1. If $e_p = f_p$ for $p = 1, 2, \ldots, k,$ then (3.2) becomes (3.1). Further, second case in (3.2) is a Kerridge [15] inaccuracy measure.

For some $E, F \in \Delta_k$, corresponding to (3.1), we define a new symmetric divergence measure based on joint representation of Renyi's-Tsallis divergence measure as:

$$D_{sym}^{\alpha}(E, F) = \begin{cases} \frac{1}{\alpha - \alpha^{-1}}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right. \\ \left. + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right]; \alpha > 0 (\neq 1). \\ \sum_{p=1}^{k} (e_p - f_p) \log\left(\frac{e_p}{f_p}\right), \alpha = 1. \end{cases} \tag{3.3}$$

Remark 2. For $\alpha = 2$, (3.3) becomes a generalization of Pearson's chi-square statistic the measure of discrepancy between two populations.

$$D_{sym}^{2}(E, F) = \frac{2}{3}\left[\log\left(\sum_{p=1}^{k} e_p^2 f_p^{-1}\right) + \log\left(\sum_{p=1}^{k} f_p^2 e_p^{-1}\right)\right.$$
$$\left. + \left(\sum_{p=1}^{k} e_p^2 f_p^{-1}\right) + \left(\sum_{p=1}^{k} f_p^2 e_p^{-1}\right) - 2\right]. \tag{3.4}$$

It might be noticed that (3.3) does not fulfill triangular property. However, $D^{\alpha}_{sym}(E, F)$ fulfills non-negativity, symmetricity and convexity property. This convexity property guarantees that in every situation, a local minimum will be global. For this, we need to prove the following lemmas:

Lemma 2.1. *Let* $E = (e_1, e_2, \ldots, e_k)$ *and* $F = (f_1, f_2, \ldots f_k) \in \Delta_k$. *Then,* $D^{\alpha}_{sym}(E, F) \geq 0$ *for* $\alpha > 0 (\neq 1)$, *with the inequality if and only if* $e_p = f_p$, *for* $p = 1, 2, \ldots, k$.

Proof: The inequality pursues by applying the Jensen inequality to the function χ characterized by $\chi(x) = x^{1-\alpha}$, $x \in [0, \infty)$ and putting $c_p = e_p$, $a_p = \dfrac{f_p}{e_p}$, for $p = 1, 2, \ldots, k$. Consider the case of $\alpha \in (1, \infty)$. Then, $1 - \alpha < 0$, therefore the function χ is convex. By the Jensen inequality, we get

$$1 = \left(\sum_{p=1}^{k} f_p\right)^{1-\alpha} = \left(\sum_{p=1}^{k} e_p \frac{f_p}{e_p}\right)^{1-\alpha} \leq \sum_{p=1}^{k} e_p \left(\frac{f_p}{e_p}\right)^{1-\alpha} = \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right), \tag{3.5}$$

and consequently

$$\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) \geq \log 1 = 0, \tag{3.6}$$

and

$$\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) - 1 \geq 0. \tag{3.7}$$

Changing e_p to f_p in (3.6) and (3.7), we have

$$\log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) \geq \log 1 = 0, \tag{3.8}$$

and

$$\log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 1 \geq 0. \tag{3.9}$$

Adding (3.6), (3.7), (3.8) and (3.9) we have

$$\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2 \geq 0, \quad (3.10)$$

since $\alpha - \alpha^{-1} > 0$ for $1 < \alpha < \infty$.

It follows that

$$\frac{1}{\alpha - \alpha^{-1}}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right.$$
$$\left. + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right] \geq 0; \quad (3.11)$$

Let $\alpha \in (0, 1)$, then the function χ is concave, and therefore we get $\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha} \leq 1$,

$$\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2 \leq 0, \quad (3.12)$$

since $\alpha - \alpha^{-1} < 0$, for $0 < \alpha < 1$. It follows that

$$\frac{1}{\alpha - \alpha^{-1}}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right.$$
$$\left. + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right] \geq 0; \quad (3.13)$$

i.e., $D_{sym}^{\alpha}(E, F) \geq 0$.

The equality in (3.5) holds if and only if $\frac{f_p}{e_p}$ is constant, for $p = 1, 2, \ldots k$, i.e., if and only if $f_p = t \times e_p$, for $p = 1, 2, \ldots, k$. By summing over all

$p = 1, 2, \dots, k,$ we get $\sum_{p=1}^{k} f_p = t \times \sum_{p=1}^{k} e_p,$ which implies that $t = 1.$ Hence, $e_p = f_p,$ for $p = 1, 2, \dots, k.$ Therefore, we conclude that $D_{sym}^{\alpha}(E, F) = 0$ if and only if $e_p = f_p,$ for $p = 1, 2, \dots, k.$

As $\alpha = 1,$ $D_{sym}^{\alpha}(E, F) \to \sum_{p=1}^{k} (e_p - f_p) \log \frac{e_p}{f_p},$ which is due to the fact the measure (3.3) is a continuous function of α.

In the proof of the next lemma, we shall use the Jensen inequality which states that for a function $\chi(\cdot)$ over a set I is said to be convex if for all choices of $x_1, x_2, \dots, x_k \in I$ and for all scalers $\lambda_1, \lambda_2, \dots, \lambda_k$ such that $\lambda_p \geq 0, \sum_{p=1}^{k} \lambda_p = 1,$ the following holds:

$$\sum_{p=1}^{k} \lambda_p \chi(x_p) \geq \chi\left(\sum_{p=1}^{k} \lambda_p x_p\right). \tag{3.14}$$

and the inequality is reversed if χ is a real concave function. The equality holds if and only if $x_1 = x_2 = \dots x_k$ or χ is linear.

Lemma 2.2.

$$D_{sym}^{\alpha}(E, F) = \frac{1}{\alpha - \alpha^{-1}}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right.$$
$$\left. + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right] \tag{3.15}$$

is a convex function of E and F for $\alpha \in (0, 1)$.

Proof. Associated with a random variable $X' = (x_1, x_2, \dots, x_k),$ let us consider r distributions

$$E_t(X') = \{e_t(x_1), e_t(x_2), \dots e_t(x_k)\}, \tag{3.16}$$

where

$$\sum_{p=1}^{k} e_t(x_p) = 1, t = 1, 2, \dots, r. \tag{3.17}$$

Now, let there be r numbers $\lambda_1, \lambda_2, \ldots, \lambda_r$ such that $\lambda_t \geq 0, \sum_{t=1}^{r} \lambda_t = 1$ and define

$$E_0(X') = \{e_0(x_1), e_0(x_2), \ldots, e_0(x_k)\}, \tag{3.18}$$

where

$$e_0(x_p) = \sum_{t=1}^{r} (\lambda_t e_t(x_p)), \; p = 1, 2, \ldots, k. \tag{3.19}$$

Obviously, $\sum_{p=1}^{k} e_0(x_p) = \sum_{p=1}^{k} f_0(x_p) = 1$, and thus $E_0(X')$ is a Bonafide distribution of X'. Consider the function

$$\begin{aligned} T = \sum_{t=1}^{r} \lambda_t \Bigg[\frac{1}{\alpha - \alpha^{-1}} \Bigg[\log\left(\sum_{p=1}^{k} (e_t(x_p))^{\alpha} f_p^{1-\alpha} \right) + \log\left(\sum_{p=1}^{k} (f_t(x_p))^{\alpha} e_p^{1-\alpha} \right) \\ + \left(\sum_{p=1}^{k} (e_t(x_p))^{\alpha} f_p^{1-\alpha} \right) + \left(\sum_{p=1}^{k} (f_t(x_p))^{\alpha} e_p^{1-\alpha} \right) - 2 \Bigg] \Bigg] \\ - \frac{1}{\alpha - \alpha^{-1}} \Bigg[\log\left(\sum_{p=1}^{k} (e_0(x_p))^{\alpha} f_p^{1-\alpha} \right) + \log\left(\sum_{p=1}^{k} (f_0(x_p))^{\alpha} e_p^{1-\alpha} \right) \\ + \left(\sum_{p=1}^{k} (f_0(x_p))^{\alpha} e_p^{1-\alpha} \right) - 2 \Bigg]. \end{aligned} \tag{3.20}$$

The function

$$\begin{aligned} D_{sym}^{\alpha}(E, F) = \frac{1}{\alpha - \alpha^{-1}} \Bigg[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha} \right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha} \right) \\ + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha} \right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha} \right) - 2 \Bigg], \end{aligned}$$

will be convex if T is greater than zero for $\alpha \in (0, 1)$.

Using Jensen's inequality for $\alpha \in (0, 1)$, we have

$$\left[\sum_{t=1}^{r} (\lambda_t e_t(x_p))\right]^{\alpha} > \sum_{t=1}^{r} (\lambda_t (e_t(x_p))^{\alpha}), \tag{3.21}$$

$$\Rightarrow \sum_{p=1}^{k} \left[\sum_{t=1}^{r} (\lambda_t e_t(x_p)\right]^{\alpha} f_p^{1-\alpha} > \sum_{p=1}^{k} \left[\sum_{t=1}^{r} (\lambda_t (e_t(x_p))^{\alpha})\right] f_p^{1-\alpha}, \tag{3.22}$$

$$\Rightarrow \sum_{p=1}^{k} ((e_0(x_p))^{\alpha} f_p^{1-\alpha}) > \sum_{t=1}^{r} \lambda_t \left[\sum_{p=1}^{k} ((e_t(x_p))^{\alpha} f_p^{1-\alpha})\right], \tag{3.23}$$

$$\Rightarrow \sum_{p=1}^{k} ((e_0(x_p))^{\alpha} f_p^{1-\alpha}) - 1 > \sum_{t=1}^{r} \lambda_t \left[\sum_{p=1}^{k} ((e_t(x_p))^{\alpha} f_p^{1-\alpha}) - 1\right], \tag{3.24}$$

$$\Rightarrow \log \sum_{p=1}^{k} ((e_0(x_p))^{\alpha} f_p^{1-\alpha}) > \log \sum_{t=1}^{r} \lambda_t \left[\sum_{p=1}^{k} ((e_t(x_p))^{\alpha} f_p^{1-\alpha})\right], \tag{3.25}$$

since $\log(\cdot)$ is a concave function for $\alpha \in (0, \infty)$.

$$\Rightarrow \log \sum_{p=1}^{k} ((e_0(x_p))^{\alpha} f_p^{1-\alpha}) > \sum_{t=1}^{r} \lambda_t \left[\log \sum_{t=1}^{k} ((e_t(x_p))^{\alpha} f_p^{1-\alpha})\right]. \tag{3.26}$$

Changing e_p to f_p in (3.24) and (3.26) we have

$$\Rightarrow \sum_{p=1}^{k} ((f_0(x_p))^{\alpha} e_p^{1-\alpha}) - 1 > \sum_{t=1}^{r} \lambda_t \left[\sum_{p=1}^{k} ((f_t(x_p))^{\alpha} e_p^{1-\alpha}) - 1\right], \tag{3.27}$$

and

$$\Rightarrow \log \sum_{p=1}^{k} ((f_0(x_p))^{\alpha} e_p^{1-\alpha}) > \sum_{t=1}^{r} \lambda_t \left[\log \sum_{p=1}^{k} ((f_t(x_p))^{\alpha} e_p^{1-\alpha})\right], \tag{3.28}$$

adding (3.24), (3.26), (3.27) and (3.28) and multiplying both sides by $\frac{1}{\alpha - \alpha^{-1}} < 0$, we get

$$\sum_{t=1}^{r}\lambda_t\left[\frac{1}{\alpha-\alpha^{-1}}\left[\log\left(\sum_{p=1}^{k}(e_t(x_p))^{\alpha}f_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(e_t(x_p))^{\alpha}f_p^{1-\alpha}\right)\right.\right.$$

$$\left.\left.+\log\left(\sum_{p=1}^{k}(f_t(x_p))^{\alpha}e_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(f_t(x_p))^{\alpha}e_p^{1-\alpha}\right)-2\right]\right]$$

$$>\frac{1}{\alpha-\alpha^{-1}}\left[\log\left(\sum_{p=1}^{k}(e_0(x_p))^{\alpha}f_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(e_0(x_p))^{\alpha}f_p^{1-\alpha}\right)\right.$$

$$\left.+\log\left(\sum_{p=1}^{k}(f_0(x_p))^{\alpha}e_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(f_0(x_p))^{\alpha}e_p^{1-\alpha}\right)-2\right]. \tag{3.29}$$

Using (3.19), we get

$$\sum_{t=1}^{r}\lambda_t\left[\frac{1}{\alpha-\alpha^{-1}}\left[\log\left(\sum_{p=1}^{k}(e_t(x_p))^{\alpha}f_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(e_t(x_p))^{\alpha}f_p^{1-\alpha}\right)\right.\right.$$

$$\left.\left.+\log\left(\sum_{p=1}^{k}(f_t(x_p))^{\alpha}e_p^{1-\alpha}\right)+\left(\sum_{p=1}^{k}(f_t(x_p))^{\alpha}e_p^{1-\alpha}\right)-2\right]\right]$$

$$>\frac{1}{\alpha-\alpha^{1}}\left[\log\left(\sum_{p=1}^{k}\left(\left(\sum_{t=1}^{r}(\lambda_t e_t(x_p))\right)^{\alpha}f_p^{1-\alpha}\right)\right)+\sum_{p=1}^{k}\left(\left(\sum_{t=1}^{r}(\lambda_t e_t(x_p))\right)^{\alpha}f_p^{1-\alpha}\right)\right.$$

$$\left.+\log\left(\sum_{p=1}^{k}\left(\left(\sum_{t=1}^{r}(\lambda_t f_t(x_p))\right)^{\alpha}e_p^{1-\alpha}\right)\right)+\sum_{p=1}^{k}\left(\left(\sum_{t=1}^{r}(\lambda_t f_t(x_p))\right)^{\alpha}e_p^{1-\alpha}\right)-2\right]. \tag{3.30}$$

This gives

$$\sum_{t=1}^{r}\lambda_t D_{sym}^{\alpha}(E_t,\ F)>D_{sym}^{\alpha}(E_o,\ F); \tag{3.31}$$

which implies the function

$$T=\sum_{t=1}^{r}\lambda_t D_{sym}^{\alpha}(E_k,\ F)-D_{sym}^{\alpha}(E_o,\ F)>0. \tag{3.32}$$

Therefore, $D^{\alpha}_{sym}(E, F)$ is a convex function of E. This completes the proof of lemma (2.2).

4. Application to Fault Detection Problem

Now, we introduce the application in the context of fault detection problem of the proposed divergence measure.

4.1. Case study of fault detection

To remove any fault, its proper detection is necessary. Symptoms are the strong indicators of any fault, so major investigations starts from analyzing symptoms. But, since each time engineer/expert is not readily available then symptoms alone cannot be the guaranteed indicators of the detect the fault and thus uncertainty prevails. Therefore, coping up with the uncertainty is essential for proper detection of fault. One such way to deal with uncertainty in technical analysis is fuzzy set theory. This uncertain information can be expressed as fuzzy sets. Therefore, different researchers have considered different approaches to fuzzy sets and their applications in technical faults. Now, we present an illustrative example as an application of proposed symmetric divergence measure.

Example 1. Consider an example of a factory where a machine is not functioning appropriately. Such break downs are normal in large factories because of which production of the factory gets affected. In this manner, it is always desirable to detect and correct the fault in such a machine at the earliest so that factory have not to suffer especially due to such break downs. We know that to fix a fault in a machine, an engineer relies upon the symptoms indicators by the machine. But sometimes the different faults in the machine have normal indications. In such a situation, the engineer needs to use its intuition to distinguish the part of the machine in which fault is located. Therefore, with the help of proposed symmetric measure, we try to fix the fault in the ill functioning machine. Broadly, we partition the machine into four parts denoted by $\xi = \{\xi_1, \xi_2, \xi_3, \xi_4\}$ and a set of faults say $\eta = \{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5\}$. Let the number of symptoms indicated by machine parts be five denoted by $\rho = \{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\}$. Characteristics symptoms for detect the faults are presented in Table 1 with lines speaking to symptoms and columns speaking to faults. Table 2 exhibits the symptoms for every machine with lines speaking to machines and sections representing symptoms, every entry in the tables is given as fuzzy numbers. To identify each fault appropriately, we evaluate proposed symmetric divergence measure for detect the faults and all machines in context of symptoms. The process is repeated for all faults. Finally, we recommend the faults to the machine whose symptoms have lowest fuzzy divergence measure from machine's symptoms.

In this way, the proposed symmetric divergence measure $D_{sym}^{\alpha}(E, F)$ is

$$= \frac{1}{\alpha - \alpha^{-1}} \left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha} \right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha} \right) \right.$$
$$\left. + \left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha} \right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha} \right) - 2 \right. \quad (4.1)$$

Table 1. Symptoms characteristics for faults.

	η_1	η_2	η_3	η_4	η_5
ρ_1	0.2	0.4	0.1	0.2	0.1
ρ_2	0.60	0.02	0.25	0.02	0.11
ρ_3	0.1	0.1	0.2	0.5	0.1
ρ_4	0.30	0.12	0.26	0.21	0.11
ρ_5	0.2	0.1	0.1	0.2	0.4

Table 2. Symptoms characteristics for machine parts.

	ρ_1	ρ_1	ρ_1	ρ_1	ρ_1
ξ_1	0.3	0.2	0.1	0.1	0.3
ξ_2	0.1	0.4	0.2	0.2	0.1
ξ_3	0.6	0.1	0.1	0.1	0.1
ξ_4	0.70	0.05	0.12	0.11	0.02

Detection Results for Proposed Symmetric Divergence Measure

Table 3. ($\alpha = 0.3$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	0.161	1.557	0.232	0.018	0.214
ξ_2	0.282	1.361	0.396	0.053	0.116
ξ_3	0.591	2.216	0.681	0.544	0.591
ξ_4	1.362	3.096	1.502	1.357	1.415
	Results	$\xi_1(\eta_4)$,	$\xi_2(\eta_4)$,	$\xi_3(\eta_4)$,	$\xi_4(\eta_4)$,

Table 4. ($\alpha = 0.35$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	1.886	3.591	1.993	1.713	1.944
ξ_2	2.027	3.389	2.176	1.758	1.845
ξ_3	2.418	4.371	2.546	2.349	2.418
ξ_4	3.367	5.437	3.548	3.317	3.364

Table 5. ($\alpha = 0.4$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	3.817	5.872	3.968	3.611	3.878
ξ_2	3.976	5.671	4.171	3.668	3.783
ξ_3	4.467	6.781	4.640	4.370	4.467
ξ_4	5.618	8.056	5.847	5.567	5.590
	Results:	$\xi_1(\eta_4)$,	$\xi_2(\eta_4)$,	$\xi_3(\eta_4)$,	$\xi_4(\eta_4)$,

Table 6. ($\alpha = 0.45$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	4.292	5.754	4.428	4.147	4.324
ξ_2	4.389	5.665	4.548	4.191	4.289
ξ_3	4.765	6.363	4.914	4.678	4.765
ξ_4	5.595	7.257	5.773	5.544	5.570
	Results:	$\xi_1(\eta_4)$,	$\xi_2(\eta_4)$,	$\xi_3(\eta_4)$,	$\xi_4(\eta_4)$,

Table 7. ($\alpha = 0.5$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	8.584	11.509	8.857	8.295	8.648
ξ_2	8.778	11.331	9.097	8.383	8.578
ξ_3	9.531	12.727	9.282	9.356	9.537
ξ_4	11.191	14.515	11.546	11.088	11.139
	Results:	$\xi_1(\eta_4)$,	$\xi_2(\eta_4)$,	$\xi_3(\eta_4)$,	$\xi_4(\eta_3)$,

Table 8. Results of proposed symmetric divergence measure.

Machine Parts	$\alpha = 0.3$	$\alpha = 0.35$	$\alpha = 0.4$	$\alpha = 0.45$	$\alpha = 0.5$
ξ_1	η_4	η_4	η_4	η_4	η_4
ξ_2	η_4	η_4	η_4	η_4	η_4
ξ_3	η_4	η_4	η_4	η_4	η_4
ξ_4	η_4	η_4	η_4	η_4	η_4

5. Comparisons with other Existing Measures

In this section, we compare the performance of proposed symmetric measure with different measures. Firstly, we consider the symmetric divergence measure corresponding to Renyi's [23] which is given by

$$D_R^{\alpha}(E,\ F) = \frac{1}{\alpha - 1}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right], \tag{5.1}$$

where $\alpha > 0 (\neq 1)$.

Presently, we conclude the above example using (5.1) for various values of parameters.

Table 9. ($\alpha = 0.1$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	1.248	1.017	0.713	1.590	1.248
ξ_2	1.401	1.028	0.866	1.743	1.401
ξ_3	2.674	1.457	2.139	3.017	2.674
ξ_4	5.103	3.754	3.931	3.710	5.103
	Results:	$\xi_1(\eta_3)$,	$\xi_2(\eta_3)$,	$\xi_3(\eta_2)$,	$\xi_4(\eta_4)$,

Table 10. ($\alpha = 0.2$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	3.016	0.486	2.643	3.255	3.016
ξ_2	3.152	0.622	2.779	3.391	3.152
ξ_3	4.284	1.754	3.911	4.523	4.284
ξ_4	6.443	2.457	3.119	4.336	6.44
	Results:	$\xi_1(\eta_2)$,	$\xi_2(\eta_2)$,	$\xi_3(\eta_2)$,	$\xi_4(\eta_2)$,

Table 11. ($\alpha = 0.3$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	5.285	4.036	5.101	5.403	5.285
ξ_2	5.404	4.155	5.220	5.522	5.404
ξ_3	6.394	5.146	6.210	6.512	6.394
ξ_4	8.284	5.461	6.102	6.740	8.284
	Results:	$\xi_1(\eta_2)$,	$\xi_2(\eta_2)$,	$\xi_3(\eta_2)$,	$\xi_4(\eta_2)$,

Table 12. ($\alpha = 0.4$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	8.304	8.611	8.349	8.275	8.304
ξ_2	8.406	8.713	8.451	8.377	8.406
ξ_3	9.255	9.562	9.300	9.226	9.255
ξ_4	10.874	9.625	9.701	9.621	10.874
	Results:	$\xi_1(\eta_4)$,	$\xi_2(\eta_4)$,	$\xi_3(\eta_4)$,	$\xi_4(\eta_4)$,

On investigating the Tables 9-12, we see that machine part ξ_1 is recommended for fault η_3 and η_4 in Tables 9 and 12 respectively. At long last, Tables 10 and 11 recommend detection for fault η_2 for machine part ξ_1.

Now, we proceed above example using new measure of probabilistic symmetric divergence corresponding to Tsallis [27] and Havdra Charvat [10] is

$$D_T^{\alpha}(E, F) = \frac{1}{\alpha - 1}\left[\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right]. \tag{5.2}$$

Solving above example using equation (5.2), we get following tables.

Table 13. ($\alpha = 0.1$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	5.047	5.113	5.114	5.046	5.020
ξ_2	5.013	5.249	5.070	5.058	5.094
ξ_3	5.089	5.041	5.159	5.043	5.089
ξ_4	5.140	5.042	5.198	5.070	5.152
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_1)$,	$\xi_3(\eta_2)$,	$\xi_4(\eta_2)$,

Table 14. ($\alpha = 0.2$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	5.784	5.878	5.886	5.783	5.745
ξ_2	5.734	6.070	5.818	5.800	5.855
ξ_3	5.848	5.774	5.949	5.780	5.848
ξ_4	5.926	5.778	6.003	5.819	5.947
	Results:	$\xi_1(\eta_5)$,	$\xi_1(\eta_1)$,	$\xi_1(\eta_2)$,	$\xi_1(\eta_2)$,

Table 15. ($\alpha = 0.2$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	6.759	6.878	6.899	6.758	6.708
ξ_2	6.694	7.120	6.802	6.779	6.854
ξ_3	6.845	6.744	6.974	6.754	6.845
ξ_4	6.952	6.751	7.042	6.807	6.987
	Results:	$\xi_1(\eta_5)$,	$\xi_1(\eta_1)$,	$\xi_1(\eta_2)$,	$\xi_1(\eta_2)$,

Table 16. ($\alpha = 0.4$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	8.935	8.994	9.009	8.935	8.909
ξ_2	8.902	9.124	8.957	8.945	8.984
ξ_3	8.980	8.927	9.046	8.933	8.980
ξ_4	9.039	8.932	9.080	8.960	9.067
	Results:	$\xi_1(\eta_5)$,	$\xi_1(\eta_1)$,	$\xi_1(\eta_2)$,	$\xi_1(\eta_2)$,

On breaking down the Tables 13-16, we see that machine parts ξ_1, ξ_2, ξ_3 and ξ_4 are recommended for faults η_5, η_1, η_2 and η_2 respectively.

R-norm symmetric divergence measure corresponding to Boekee and Lubbe [3] which is as follows:

$$D_\alpha(E, F) = \frac{\alpha}{\alpha - 1}\left[\left(\sum_{p=1}^{k} e_p^\alpha f_p^{1-\alpha}\right)^{\frac{1}{\alpha}} + \left(\sum_{p=1}^{k} f_p^\alpha e_p^{1-\alpha}\right)^{\frac{1}{\alpha}} - 2\right]. \quad (5.3)$$

Solving above example using equation (5.3), we obtain the following tables.

Table 17. ($\alpha = 0.1$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	0.701	0.703	0.697	0.706	0.694
ξ_2	0.695	0.697	0.691	0.700	0.688
ξ_3	0.694	0.696	0.690	0.698	0.687
ξ_4	0.709	0.711	0.704	0.713	0.702
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_5)$,	$\xi_3(\eta_5)$,	$\xi_4(\eta_5)$,

Table 18. ($\alpha = 0.2$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	0.778	0.782	0.767	0.788	0.761
ξ_2	0.764	0.767	0.753	0.775	0.747
ξ_3	0.760	0.765	0.751	0.770	0.764
ξ_4	0.796	0.799	0.784	0.806	0.778
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_5)$,	$\xi_3(\eta_5)$,	$\xi_4(\eta_5)$,

Table 19. ($\alpha = 0.3$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	1.346	1.352	1.329	1.360	1.321
ξ_2	1.327	1.331	1.310	1.341	1.300
ξ_3	1.321	1.327	1.305	1.335	0.996
ξ_4	1.372	1.377	1.354	1.386	1.345
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_5)$,	$\xi_3(\eta_5)$,	$\xi_4(\eta_5)$,

Table 20. ($\alpha = 0.4$).

	η_1	η_2	η_3	η_4	η_5
ξ_1	1.801	1.810	1.769	1.826	1.752
ξ_2	1.767	1.775	1.735	1.792	1.718
ξ_3	1.756	1.765	1.724	1.781	1.708
ξ_4	1.847	1.855	1.814	1.872	1.798
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_5)$,	$\xi_3(\eta_5)$,	$\xi_4(\eta_5)$,

On breaking down the Tables 17-20, we see that each machine part ξ_1, ξ_2, ξ_3 and ξ_4 is recommended for fault η_5.

Now, we sum up the above example utilizing symmetric divergence measure corresponding to Lin [20] given as:

$$D^{K-L}(E, F) = \sum_{p=1}^{k}(e_p - f_p)\log\left(\frac{e_p}{f_p}\right). \tag{5.4}$$

Table 21. Detection Result.

	η_1	η_2	η_3	η_4	η_5
ξ_1	0.675	1.557	1.762	0.672	0.300
ξ_2	0.200	3.550	0.996	0.828	1.400
ξ_3	1.334	0.571	2.321	0.644	1.334
ξ_4	1.334	0.571	2.321	0.644	1.134
	Results:	$\xi_1(\eta_5)$,	$\xi_2(\eta_1)$,	$\xi_3(\eta_2)$,	$\xi_4(\eta_2)$,

From Table 21, we see that machine parts ξ_1, ξ_2, ξ_3 and ξ_4 are recommended for faults η_5, η_1, η_2 and η_4 respectively.

Conclusion

From the above discussion, it can be concluded that

1. Measure condition (4.1), proposes that η_4 is right for all machine parts ξ_1, ξ_2, ξ_3, and ξ_4 for different parametric quantities.

2. Measure condition (5.1), suggest the fault η_2 for machine part ξ_1 but in other case fault η_4 is recommend for machine part ξ_1 for different values of parameters.

3. Measure equation (5.2) suggests that η_5, η_1, η_2 and η_2 are correct for machine parts ξ_1, ξ_2, ξ_3 and ξ_4 respectively.

4. Measure equation (5.3) suggests that η_5 is correct for all machine parts ξ_1, ξ_2, ξ_3 and ξ_4 for different values of parameter α.

5. Measure equation (5.4) suggests that η_5, η_1, η_2 and η_2 are correct for machine parts ξ_1, ξ_2, ξ_3 and ξ_4 respectively.

Correspondingly, symmetric divergence measure proposed by different researchers might be used to evaluate the above example. The results are summarized in the following Table 22.

Table 22. Comparison of Results.

Machine parts	Proposed sym. Measure	Lin	Renyi	Tsallis	Boekee	Boekee and Lubbe
	ξ_1	η_4	η_5	either η_2 or η_4	η_5	η_5
	ξ_2	η_4	η_1	η_4	η_1	η_5
	ξ_3	η_4	η_2	either η_2 or η_4	η_2	η_5
	ξ_4	η_4	η_2	η_4	η_2	η_5

Discussion

Above analysis shows that the performance of proposed symmetric measure equation (4.1) is more consistent when contrasted with equation (5.1), (5.2), (5.3) and (5.4). Because proposed measure gives consistent result according to the sequence, while other comparative measures gave no consistency in any case. i.e., Renyi's symmetric measure does not able to discriminate which fault either η_2 or η_4 is correct for machine part ξ_1 and ξ_3. Thus, based on above examination, we can say that determining a detection based on symptoms is a difficult task. Only some particular symptoms lead to some specific detection. A thorough analysis ought to be done to arrive at a conclusion. Aside from symptoms, engineer's advice is also advisable to determine the detection.

5.1. Case study of pattern recognition

In the present subsection, we give the application of proposed symmetric divergence measure in pattern recognition. In pattern recognition problem, we need to distinguish the pattern out of given patterns which resembles the most with the given pattern by looking at their features. The procedure is as under:

Let there be k-known patterns $\xi_1, \xi_2, \ldots, \xi_k$ with m respective classification $L_1, L_2, \ldots, L_m$. The Patterns are represented by the k-probability distributions associated with discrete random variables $X = \{z_1, z_2, z_3, \ldots z_k\}$. Given an unknown pattern $Q = (g_1, g_2, \ldots g_k) \in \Delta_k$. Our point is to characterize Q to one of the classes $L_1, L_2, \ldots, L_m$. As per the principle of minimum divergence/discrimination information between probability distribution on a set of probability distributions, the way of assigning Q to L_t^* is described by

$$t^* = \arg\min\{\xi_p, Q\}, \ for\ p = 1, 2, \ldots k.$$

As indicated by this algorithm, the given pattern can be recognized so that the best class can be chosen. It is a practical application of minimum divergence measure rule to pattern recognition.

Illustrative Example 2. Suppose $X = \{z_1, z_2, z_3,\}$ ba a finite set of random variables. Three known patterns ξ_1, ξ_2, ξ_3 with classifications L_1, L_2, L_3 respectively are given. Let ξ_1, ξ_2, ξ_3 be represented by the following set of probability distributions

$$\xi_1 = \{(z_1, 0.6), (z_2, 0.3), (z_3, 0.1)\},$$

$$\xi_2 = \{(z_1, 0.4), (z_2, 0.2), (z_3, 0.4)\},$$

$$\xi_3 = \{(z_1, 0.3), (z_2, 0.1), (z_3, 0.6)\}.$$

Unknown pattern Q is given as pursues:

$$Q = \{(z_1, 0.3), (z_2, 0.5), (z_3, 0.2)\},$$

Our goal is to find one of the classes L_1, L_2, and L_3 that Q belongs to. From formula (3.13), we can calculate the values of divergence measure $T(P_t, Q)$, $t = 1, 2, 3$ for any $\alpha \in (0, 1)$ and are presented in Table as follows:

Table 23. Computed numerical values of proposed probabilistic symmetric divergence measure $T(\xi_t, Q)$, $t = 1, 2, 3$ for any $\alpha \in (0, 1)$.

	$\alpha = 0.4$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
$T(\xi_1, Q)$	1.078	3.142	6.142	10.819	20.061	45.383
$T(\xi_2, Q)$	0.668	2.574	5.293	9.667	18.206	43.445
$T(\xi_3, Q)$	0.919	2.976	5.926	10.678	19.952	47.321

Hence, the unknown pattern is classified into class L_2 for different parametric values of α.

Comparison with other existing symmetric measures

We currently demonstrate the effectiveness of proposed measure equation (4.1) by contrasting its performance and measure in literature through example 2. In the present comparison, we have considered Renyi [23], Tsallis [27], Boekee and Lubbe [3] and Lin [20] etc.

The symmetric divergence measure corresponding to Renyi's [23] which is given by

$$D_R^{\alpha}(E, F) = \frac{1}{\alpha - 1}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \log\left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)\right]. \tag{5.5}$$

The symmetric divergence measure corresponding to Tsallis [27] is as follows:

$$D_T^{\alpha}(E, F) = \frac{1}{\alpha - 1}\left[\log\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right) + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right) - 2\right]. \tag{5.6}$$

R-norm symmetric divergence measure corresponding to Boekee and Lubbe [3] which is as follows:

$$D_{B-L}^{\alpha}(E, F) = \frac{\alpha}{\alpha - 1}\left[\left(\sum_{p=1}^{k} e_p^{\alpha} f_p^{1-\alpha}\right)^{\frac{1}{\alpha}} + \left(\sum_{p=1}^{k} f_p^{\alpha} e_p^{1-\alpha}\right)^{\frac{1}{\alpha}} - 2\right]. \tag{5.7}$$

The probability symmetric divergence measure corresponding to Lin [20] is

$$D_L(E, F) = \sum_{p=1}^{k} (e_p - f_p)\log\left(\frac{e_p}{f_p}\right). \tag{5.8}$$

Computed values of symmetric divergence measure equation (4.1), equation (5.5), equation (5.6), equation (5.7) and equation (5.8) are presented in the Table 24.

Table 24. Computed values of D_{sym}^{α}, D_R^{α}, D_T^{α}, D_{B-L}^{α} and D_L.

divergence measure	ξ_1	ξ_2	ξ_3
$D_{sym}^{\alpha}(\xi_p, Q)$	1.078	0.668	0.919
$D_R^{\alpha}(\xi_p, Q)$	3.421	3.366	3.421
$D_T^{\alpha}(\xi_p, Q)$	6.817	6.841	7.083
$D_{B-L}^{\alpha}(\xi_p, Q)$	0.547	0.638	1.562
$D_L(\xi_p, Q)$	0.483	0.207	0.324

Thus, Table 24 shows a comparison of the classification result of the proposed symmetric measure with the ones of the existing symmetric measures. From Table 24, we can know the proposed measure D_{sym}^{α}, D_{R}^{α} and D_L that support the pattern ξ_2 resembles the most with Q. While, D_{T}^{α} and D_{B-L}^{α} support the pattern ξ_1. Although, final classifications results $\xi_2(D_{sym}^{\alpha}, D_{R}^{\alpha}, D_L)$ and $\xi_1(D_{T}^{\alpha}, D_{B-L}^{\alpha})$ are distinguish in unknown pattern.

6. Conclusions

In this study, we have proposed J-divergence measure based on Renyi's-Tsallis entropy measure. Some of its basic properties are examined and comparison between proposed symmetric measure and some other existing measures also discussed. In addition to this, its efficiency is established utilizing an example on fault detection and pattern recognition. For fault detection problem in example 1 and pattern recognition problem in example, a correlation of results of existing measures is exhibiting in Table 22 and 24 respectively. In future, we will utilize the symmetric divergence measure in another ways such as data mining and decision making. In addition, as this paper is just an applied result focus on the divergence measure for probability distributions. We shall attempt to design some softwares to preferably relies the initiated information measure in every day life. Meanwhile, we will also bring them into various fuzzy environment.

References

[1] S. M. Ali and S. D. Silvey, A general class of coefficients of divergence of one distribution from another, J. Roy. Statist. Soc. Ser. B 28 (1966), 131-142.

[2] D. Bhandari, N. R. Pal and D. D. Majumdar, Fuzzy divergence, probability measure of fuzzy events and image thresholding, Inf. Sci. 13 (1992), 857-867.

[3] D. E. Boekee and J. C. A. Van Der Lubbe, Some aspect of error bounds in feature selection, Pattern Recognition 1 (1979), 353-360.

[4] J. Burbea and C. R. Rao On the convexity of some divergence measures based on entropy functions, IEEE Trans. Inf. Th. 28 (1982), 489-495.

[5] M. Ben Bassat, *f*-entropies, probability of error and feature selection, Inform, Contr. 39 (1978), 227-242.

[6] C. H. Chen, Statistical Pattern Recognition, Rochelle Park, NJ: Hayden Book Co., Ch. 4, 1973.

[7] C. K. Chow and C. N. Lin, Approximating discrete probability distributions with dependence trees, IEEE Trans. Inform. Theory 14(3) (1968), 462-467.

[8] I. Csiszar, Information-type measures of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar 2 (1967), 299-318.

[9] S. S. Dragomir, A generalized f-divergence for probability vectors and applications, Research report collection, 5, 2000.

[10] J. Havdra and F. Charvat, Quantification method of classification processes, concept of structural α-entropy, Kybernetika 3 (1967), 30-35.

[11] D. S. Hooda, On generalized measures of fuzzy entropy, Mathematica Slovaca 54 (2004), 315-325.

[12] D. Garg, and S. Kumar, Parametric R-norm directed divergence convex function, Infinite dimensional analysis, quantum probability and related topics 19(2) (2016), 1-12.

[13] R. W. Johnson, Axiomatic characterization of the directed divergences and their linear combinations, IEEE Trans. Inform. Theory IT-25, 6 (1979), 709-716.

[14] T. Kailath, The divergence and Bhattacharyya distance measures in signal selection, IEEE Trans. Comm. Technology COM-15 (1967), 52- 60.

[15] D. F. Kerridge, Inaccuracy and Inference. J. Roy. Statist. Soc. Ser. B 23 (1961), 184-194.

[16] S. Kullback and R. A. Leibler, On information and suficiency, Ann. Math. Stat. 22 (1951), 79-86. 17

[17] S. Kullback, Information theory and Statistics, Wiley, New York, 1959.

[18] H. Jefireys, An invarient form for the prior probability in estimation problem, Proc Roy. Soc. Lon. Ser. A 186 (1946), 453-461.

[19] P. Kumar, and A. Johnson, On a symmetric divergence measure and information inequalities, J. Inequal. in pure and Appl. Math. 6 (2005), 1-13.

[20] J. Lin, Divergence measure based on Shannon entropy, IEEE Trans. Inf. Th. 37(1) (1991), 145-151.

[21] M. A. Nielsen and M. A. Chuang, I. L., Quantum computation and information, Cambridge University Press, Cambridge UK, 3(8) (2000), 9.

[22] H. V. Poor, Robust decision design using a distance criterion, IEEE Trans. Inf. Th. IT-26 (1980) 575-587.

[23] A. Renyi, On measures of entropy and information, Proc. 4th Barkley symp. on math. Stat. and probability, University of California Press 1 (1961) 547.

[24] R. Santos-Rodriguez, D. Garcia-Garcia, and J. Cid-Sueiro, Costsensitive classification based on Bregman divergences for medical diagnosis, In M. A. Wani, editor, Proceedings of the 8th International Conference on Machine Learning and Applications (ICMLA'09), Miami beach, Fl., USA, December 13-15 (2009), 551-556.

[25] C. E. Shannon, A mathematical theory of communication, Bell Syst. Technol. J. 27 (378-423), (1948), 623-656.

[26] I. J. Taneja, Generalized symmetric divergence measures and inequalities, RGMIA research report collection. http://rgmia.vu.edu.au, 7, Art. 9. Available online at: arXiv:math.ST/0501301, 2005.

[27] C. Tsallis, Possible generalization of Boltzman-Gibbs statistics, J. Stat. Phys. 52 (1988), 480-487.

[28] L. Wondie and S. Kumar, A joint representation of Renyi's-Tsallis entropy with application in coding theory, Int J. Math. Math Sci., Art. ID 2683293 (2017), 1-5.

[29] R. R. Yager, On the measure of fuzziness and negation, Part 1: Membership in the unitinterval, Int. J. Gen. Syst. 5 (1979), 221-229.

[30] L. A. Zadeh, Fuzzy sets, Inf. Control 8 (1965), 221-229.

[31] L. A. Zadeh, Probability measures of fuzzy events, J. Math. Anal. Appl. 23 (1968), 421-427.