STUDY OF INFORMATION THEORETIC BASED MEASURES AND THEIR PROPERTIES

Submitted in

fulfillment of the requirements for the degree of

Doctor of Philosophy

by

Devendra Singh
(2013RMA9061)

Under the Supervision of
Prof. K. C. Jain and Prof. Rashmi Jain

DEPARTMENT OF MATHEMATICS
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY, JAIPUR

SEPTEMBER, 2019

Dedicated To

My loving daughters

Saumya & Kavya

Declaration

I, Devendra Singh, declare that this thesis titled**, “Study of Information Theoretic Based Measures and Their Properties”** and the work presented in it, are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this university.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself, jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

September 2019

Devendra Singh
(2013RMA9061)

Certificate

This is to certify that the thesis entitled **"Study of Information Theoretic Based Measures and Their Properties"** being submitted by Mr. Devendra Singh (2013RMA9061) is a bonafide research work carried out under our supervision and guidance in fulfillment of the requirement for the award of the degree of **Doctor of Philosophy** in the **Department of Mathematics,** Malaviya National Institute of Technology, Jaipur, India. The matter embodied in this thesis is original and has not been submitted to any other University or Institute for the award of any other degree.

Kamal Chand Jain
Professor (Ret.)
Department of Mathematics
Malaviya National Institute of Technology Jaipur

Place: Jaipur
Date: September, 2019

Rashmi Jain
Professor
Department of Mathematics
Malaviya National Institute of Technology Jaipur

Place: Jaipur
Date: September, 2019

Acknowledgements

Firstly, I pay homage to my parents for showering blessings and paving my way to success. Further, I feel immense pleasure to thank my esteemed supervisors Dr. K.C. Jain and Dr. Rashmi Jain, Professor in Department of Mathematics, Malaviya National Institute of Technology, Jaipur for their generosity, continuous attention, intellectual discussion and constructive criticism which constantly encouraged and enabled me to accomplish this huge task with full zeal and enthusiasm. I sincerely thank them for the judicious guidance, graciously lavished upon this work, in spite of their busy schedule.

It is a matter of great privilege for me to extend my sincere thanks to Dr. Vatsala Mathur, Dr. Sanjay Bhatter, Dr. Santosh Chaudhary, Dr. Nilima Gupta, Dr. Ritu Agarwal and Dr. Varun Jindal Department of Mathematics, Malaviya National Institute of Technology Jaipur for their valuable suggestions and constant encouragement during the course of this study.

With great pleasure, I express my most affable and cordial thanks to Prof. Udaykumar R Yaragatti, Director of Malaviya National Institute of Technology, Jaipur for rendering necessary help and research facilities provided in the Institute.

Lastly, I would like to extend my sincere thanks to my family, friends, especially Dheeraj Kumar Joshi, Kaushal Kishore Sharma and relatives, colleagues for their round the clock assured help and support, due to which I am able to complete my thesis.

Devendra Singh

Abstract

Many authors did a lot of study regarding information divergence measures and applied these divergences in several fields in information theory. Various measures of information have attracted the interest of scientific community recently primarily due to their use in several disciplines such as in Communication theory, Cybernetics, Biology, Psychology, Economics, Statistics, Thermodynamics, Questionnaire theory, Probability theory and many more.

In the present work, Various classes of non – parametric symmetric information measures of divergence, which belongs to the family of Csiszar's *f* – divergence have been studied. Most of these classes are made up of symmetric Chi-Square, Kullback –Leibler, Relative Jensen-Shanon and arithmetic mean divergence measures. It is further shown that most of the information divergence measures in these classes are closely associated with the already known divergence measures in the literature. Also the properties of new divergence measures has been discussed which, by suitable choice of the convex, normalized and other functions involved, leads to some recently proposed divergence measures of Csiszar's *f* – divergence class. Further bounds on these new divergence measures have been obtained.

The summary of the thesis is as follows:
Chapter 1 introduces the whole thesis.

Chapter 2 introduces a new class of, non-parametric symmetric, divergence measures. Further we found that these new information measures are meticulously associated with some familiar information divergence measures.

Chapter 3 introduces different series of some new classes of information divergence measures, which belong to the family of Csiszár's *f*- divergences.

Chapter 4 introduces two new series of information divergence measures using Jain and Saraswat generalized f- divergence measure to obtain various new information

inequalities on these new series of divergence measures with some well-known information measures.

Chapter 5 introduces a non-parametric theoretic based exponential information divergence measure.

Chapter 6 introduces a non-parametric symmetric information divergence measure and its properties are studied and discussed.

Chapter 7 includes the conclusion of the work reported in this thesis and also discuss the scope for further study.

Content

List of Figures

List of Tables

1

INTRODUCTION

1.1 Historical background

Information theory deals with the quantification, information storage, and communication of information. The most prominent feature of 20th century technology has been the development and exploitation of new communication media. Information theory is a mathematical demonstration of the state of affairs and constraints affecting the conduction and processing of information. The field of information theory is at the intersection of Mathematics, Statistics, Computer Science and Electrical Engineering. The landmark event that established the discipline of Information theory and brought it to immediate worldwide attention was the publication of Claude E. Shannon's characteristic paper "A mathematical Theory of Communication" in the Bell system Technical Journal in 1948 [128].

The first person who studied all information theory was Harry Nyquist in 1924 [109] & 1928 [110] and by Hartley in 1928[46] who discovered the logarithmic nature of the measure of information. Harry Nyquist published the paper 'Certain Factors Affecting Telegraph Speed' in which he gave the relation $W=K \log m$ where W is the speed of transmission of intelligence, m is the number of difference voltage levels to choose from at each time step and K is a constant. He quantified "intelligence" and the "line speed" at which it can be communicated by communication system.

In 1928 Ralph Hartley published a paper titled 'Transmission of Information' and used the word information as a measurable quantity and quantifying information as $H = \log S^n = n \log S$ where S was the number of possible symbols and n the number of symbols in a transmission. At that time Wiener [159, 160] also came up with results analogous to those of Shannon. After its beginning it has broadened to find applications in many other areas, including statistical inference, natural language processing, cryptography, networks other than communication networks - as in neurobiology, the evolution and function of molecular codes, model selection in ecology, thermal physics, quantum computing, plagiarism detection and other forms of data analysis. Applications of fundamental topics of information theory include lossless data compression (e.g. ZIP files), lossy data compression (e.g. MP3's and JPEG's), and channel coding (e.g. for DSL lines).

Information theory overlaps deeply with communication theory, but it is more oriented toward the necessary limitations on the processing and communication of information and less oriented toward the detailed operation of particular strategies and devices.

Its impact has been crucial to the success of the Voyager missions to deep space, the invention of the compact disc, the feasibility of mobile phones, the development of the Internet, the study of linguistics and of human perception, the understanding of black holes, and numerous other fields. Important sub-fields of information theory are source coding, channel coding, algorithmic complexity theory, algorithmic information theory, and measures of information.

The most important and direct application of Information Theory is in Coding Theory. In 1974 Dutta in his paper (39) showed that information theory can also be applied in Number Theory, Quantum Mechanics, Qualitative Dynamics and Approximation Theory. The concepts introduced by Shannon have also been applied with enormous degree of success in a number of fields such as Biology, Psychology, Economics, Statistics thermodynamics, Language Questionnaire theory, probability theory, communication theory, Cybernetics and many more. Since its inception it has broadened to find applications in many other areas, including statistical inference, natural language processing, cryptography, networks other than communication networks as in neurobiology, the evolution and function of molecular codes, model selection in ecology, thermal physics, quantum computing, plagiarism detection and other forms of data analysis. A key measure of information is known as entropy, which is usually expressed by the average number of bits needed to store or communicate one symbol in a message, entropy quantifies the uncertainty involved in predicting the value of random variable.

1.1.1 Shannon's entropy and other generalized entropy

Without essential loss of insight, we have restricted ourselves to discrete probability distributions, so let

$$\Gamma_n = \left\{ P = (p_{1,} p_{2,} .... p_n) \middle| p_i \geq 0, \sum_{i=1}^{n} p_i = 1 \right\}, n \geq 2 \tag{1.1.1}$$

be the set of all complete finite discrete probability distributions.

Shannon [128] introduced the following measure of information

$$H_n(P) = -\sum_{i=1}^{n} p_i \log p_i \tag{1.1.2}$$

For all $P = (p_{1,} p_{2,} .... p_n) \in \Gamma_n$.The expression (1.1.2) is famous as Shannon's entropy or measure of uncertainty. The function $H_n(P)$ represents the expected value of uncertainty associated with the given probability scheme and it is uniquely determined by some rather natural postulates. The Shannon entropy is the key concept in information theory. It satisfies many interesting properties [140]. There are different approaches to the derivation of Shannon entropy based on different postulates and axioms [2, 14, 15, 86, 98, 101 and 151]. The Shannon entropy has found wide applications in different fields of science and technology [12, 13, 28, 43, 45, 53, 100 and 151]. Applications of Shannon's entropy to music can be seen in Siromoney and Rajagopalan [135]. Further Shannon's entropy has also been used extensively in the analysis of the structure of languages. Various applications can be seen from Siromoney [135], Balasubrahmanyam and Siromoney [7].

The coding theorems of information theory provide such overwhelming evidence for the adequateness of the Shannon information measure that to look for essentially different measure of information might appear to make no sense at all. Still, all the evidence show that the Shannon information measure is the only possible one is valid, only within restricted scope of coding problems, considered by Shannon. As pointed out by Renyi [125] in his fundamental paper on generalized information measures, that in some specific cases, other quantities may serve just, as well, or even better, as measures of uncertainty. This should be supported either by their operational significance or by a set of natural postulates characterizing them, or, preferably by both. Thus the idea of generalized entropy arises in the literature. It started with Renyi [125] who generalized the Shannon's entropy by expressions

$$H_n^*(P) = \frac{-\sum_{i=1}^{n} p_i \log p_i}{\sum_{i=1}^{n} p_i} \tag{1.1.3}$$

$$H_n^\alpha(P) = \frac{1}{1-\alpha} \log\left[\sum_{i=1}^{n} p_i^\alpha\right],\ \alpha \neq 1,\ \alpha > 0. \tag{1.1.4}$$

For all $P = (p_{1,} p_{2,} .... p_n) \in \Gamma_n$ and called them as entropy of order 1 and order α respectively. Here α is a real parameter. We can easily verify that

$$\lim_{\alpha \to 1} H_n^\alpha(P) = H(P) \tag{1.1.5}$$

Where $H_n(P)$ is the Shannon's entropy defined by (1.1.2). This was the first systematic attempt to develop a generalization of Shannon's entropy. Renyi's entropy was generalized by Aczel and Dacroczy [1], Dacroczy [26], Verma [156, 157], Kapur [79, 80], Verma and Nath [158], Rathie [122] etc. For operational purposes, it seems more natural to consider, the expression $\sum_{i=1}^{n} p_i^\alpha$ as an information measure instead of Renyi's entropy of order α. So Havrda and Charvat [47] introduced a new type of entropy called structural α - entropy by the expression

$$\tilde{H}_n^{\alpha}(P) = \frac{\sum_{i=1}^{n} p_i^{\alpha} - 1}{2^{1-\alpha} - 1} \;, \alpha \neq 1, \alpha > 0 \tag{1.1.6}$$

Again we can verify that

$$\lim_{\alpha \to 1} \tilde{H}_n^{\alpha}(P) = H_n(P) \tag{1.1.7}$$

Where $H_n(P)$ is the Shannon's entropy defined by (1.1.2). This quantity (1.1.6) permits a simple characterization [47]. Dacrozy [27] gave an alternative way to characterize it. Another type of generalization identical to (1.1.6) can be seen in Nath [107] and Vajada [155]. Some other important generalizations of Shannon's entropy can be seen in Arimoto [4], Kapur [82, 83, 84].

From 1961, more than thirty measures of entropies have been introduced in the literature on information theory generalizing Shannon's entropy. These are famous as parametric, trigonometric and weighted entropies. Renyi [125] for the first time gave the idea of parametric entropy. The idea of the trigonometric entropies was initiated by Aczel and Dacrozy [1]. Later Sharma and Taneja [132] and Sant'anna & Taneja [127] studied it from different aspects. The idea of weighted entropies was given by Belis and Guaisu [9]. Longo [99] gave several interpretations for these measures. Later Picard [116] extended it for generalized measures. The list of these generalized measures including their unified forms can be seen in Kapur [84], Taneja [140]. For simplicity, from now onwards we will denote the Shannon entropy by *H (P)* instead of $H_n(P)$.

1.1.2 Directed divergence & inaccuracy

In 1951 Soloman Kullback and Richard-Leibler [91], two national security agency mathematicians, studied a measure of information from statistical overview, given by

$$K(P,Q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \tag{1.1.8}$$

for all $P = (p_{1,} p_{2,} .... p_n) \in \Gamma_n$ $Q = (q_{1,} q_{2,} .... q_n) \in \Gamma_n$. The measure (1.1.8) has many names given by different authors such as Relative Information, Directed Divergence, Cross Entropy, Measure of Discrimination etc. whenever any q_i is zero, then the corresponding p_i is also zero and we take $0 \log \frac{0}{0} = 0 \log 0 - 0 \log 0 = 0$. At the same time, Kullback and Leibler also studied a measure, called J-divergence, given by.

$$J(P,Q) = K(P,Q) + K(Q,P) = \sum_{i=1}^{n} (p_i - q_i) \log \frac{p_i}{q_i} \tag{1.1.9}$$

We can easily see that $K(P,Q)$ is non-symmetric information measure whereas $J(P,Q)$ is symmetric information measure with respect to probability distributions. The above measure $J(P,Q)$ was already studied by Jeffrey in 1996 [67]. When the two distributions are equal i.e $p_i = q_i \; \forall \; i = 1,2,3,...,n$, then the Directed Divergence becomes zero. This property is useful in differentiating Directed Divergence from other information measures. Some of the characterizations of Directed Divergence are given by Kannapan [76], Kannapan and Rathie [77, 78] etc.

Another important measure of information for a pair of probability distributions is the inaccuracy measure, introduced by Kerridge [87] and is given by

$$H(P,Q) = -\sum_{i=1}^{n} p_i \log q_i \tag{1.1.10}$$

For all $P = (p_{1,} p_{2,} .... p_n) \in \Gamma_n$, $Q = (q_{1,} q_{2,} .... q_n) \in \Gamma_n$.

When $p_i = q_i \; \forall \; i = 1,2,3,...,n$, the measure (1.1.10) becomes the Shannon's entropy. Therefore Kerridge's Inaccuracy, given by (1.1.10), is a generalization of Shannon's entropy. The Kerridge's inequaracy was characterized by Kerridge [87], Nath [107] etc. We can see that *H(P), K(P,Q) and H(P,Q)* also satisfy a very interesting relation given by

$$H(P,Q) = H(P) + K(P,Q) \tag{1.1.11}$$

Renyi (1961) [125] first presented a scalar parametric generalization of Kullback-Leibler's Directed Divergence given by (1.1.8). Several authors presented alternative ways of generalizing it. These generalizations are as follows

- Directed Divergence of order 'r' (Renyi [125])

$$K^r(P,Q) = (r-1)^{-1} \log\left(\sum_{i=1}^{n} p_i^r \, q_i^{1-r}\right),\ r \neq 1, r > 0 \tag{1.1.12}$$

- Directed Divergence of type 's'(Sharma & Autar [129])

$$^1K_s(P,Q) = [s-1]^{-1}\left(\sum_{i=1}^{n} p_i^s \, q_i^{1-s} - 1\right),\ s \neq 1, s > 0 \tag{1.1.13}$$

The modified version of the measure (1.1.13) is given by

$$^2K_s(P,Q) = [s(s-1)]^{-1}\left(\sum_{i=1}^{n} p_i^s \, q_i^{1-s} - 1\right),\ s \neq 0,1 \tag{1.1.14}$$

In particular, we have

$$\lim_{r\to 1} K^r(P,Q) = \lim_{s\to 1} {}^1K_s(P,Q) = \lim_{s\to 1} {}^2K_s = K(P,Q) \tag{1.1.15}$$

And

$$\lim_{s\to 0} {}^2K_s(P,Q) = K(P,Q) \tag{1.1.16}$$

Rathi and Kanappan [123] gave another generalization of (1.1.8) and called it directed divergence of order beeta which is given by

$$I^\beta(P,Q) = \left(\sum_{i=}^{n} p_i^\beta \, q_i^{\beta-1} - 1\right)\left(2^{1-\beta} - 1\right)^{-1} \quad \textit{for } \beta \neq 1 \tag{1.1.17}$$

Later Patni & Jain [112, 113, and 114] gave various different methods of characterizing measures (1.1.13) and (1.1.17) respectively. Some other important generalizations of directed divergence can be seen in [130, 149]. The concept of weighted directed divergence and weighted inaccuracy was introduced by Taneja and Tuteja [136, 137]. Further results in this direction can be seen in Bhaker and Hooda [11], Hooda and Tuteja [51], Hooda and Ram [50]. Similar generalizations of Kerridge's inaccuracy exist in literature and can be seen in Kapur [84], Sharma and Mittal [131], and Taneja [140].

1.2 A Review of Information and Divergence Measures

As a generalization of the uncertainty theory based on the notion of possibility (Hartley, 1928[46]), information theory considers the uncertainty of randomness perfectly. The concept of Shannon's entropy (1948[128]) is the central conception of information theory. Sometimes this measure is referred as the measure of uncertainty. The entropy of a random variable is defined in terms of its probability distribution and can be shown to be a good measure of randomness or uncertainty. Shannon's model used the formalized language of the classical set theory, so it is only suitable to be used in limitation of classical set theory. Kolmogorov [88] proposed the notion of ε-entropy to measure the uncertainty when the set has unlimited elements. As pointed out by Renyi [125] in his fundamental paper on generalized information measures, in other short of problems other quantities may serve just as well, or even better, as measures of information. This should be supported either by their operational significance or by a set of natural postulates characterizing them or preferably by both. Thus the idea of generalized entropies arises in the literature. It started with Renyi [125] who characterized scalar parametric entropy as entropy of order, which includes Shannon entropy as a limiting case.

As to the divergence and inaccuracy of information, Kullback and Leibler [91] studied a measure of information from statistical aspects of view involving two probability distributions associated with the same experiment, calling discrimination function, later different authors named as cross entropy, relative information etc. It is a non-symmetric measure of the difference between two probability distributions P and Q. At the same time they also developed the idea of the Harold (1946) invariant, famous as J-divergence. Kerridge [87] studied a different kind of measure calling inaccuracy measure involving again two probability distributions.

Sibson [134] studied another divergence measure involving two probability distributions, using mainly the concavity property of Shannon's entropy, calling information radius. Later, Burbea and Rao [16, 17] studied extensively the information radius and its parametric generalization, calling this measure as Jensen difference measure. Taneja [142, 143] studied a new measure of divergence and its two parametric generalizations involving two probability distributions based on arithmetic and geometric mean inequality.

Sharma and Taneja [132] and Sant’anna and Taneja [127] studied trigonometric entropies from different aspects. The idea of weighted entropies started by Belis and Guaisu [9], Later Picard [116] extended it for generalized measures. After Renyi [125], other researchers such as Havrda and Charvat [47], Arimoto [4], Sharma and Mittal [131] etc, were interested towards other kinds of expressions generalizing Shannon's entropy. Taneja [138] unified some of these. Taneja [139] introduced a new divergence measure called arithmetic-geometric mean divergence measure. Taneja [142, 143] studied symmetric and non-symmetric divergence measures and their generalizations based on different divergence measures.

The present chapter is an introductory one. In this chapter, we have listed different measures of information proposed by various mathematicians and researchers in associated fields. These measures are made up of one, two, and in some cases more than two probability distributions. Since our work deals with measures involving two probability distributions, our focus is more on these measures and generalizations. Some of the specific applications and interpretations of these measures discussed in the subsequent chapters of the thesis have been given in the following sections. Thereafter, an overview of the chapters that would constitute the major part of this thesis is provided. We begin with brief descriptions of measure of information made up of one probability distribution.

One of the important issues in many applications of Statistics & Probability Theory is finding an appropriate measure of distance (or difference or discrimination) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [67], Kullback and and leibler [91], Renyi [125], Havrda and Charvat [47], Kapur [81], Sharma and Mittal [131], Berbea and Rao [16], Rao [121], Csiszar [25], Ali and Sllvey [3], vajda [154], Shioya and Da-te [133] among others. These measures applied a variety of fields such as anthropology [121], genetic [104], economics and poltical science [149], biology [117], the analysis of contingency tables [44], approximation of probability distributions [21, 85], signal processing [70, 75] and pattern recognition [8, 19, 68]. In this section, we will focus on these divergence measures. Further we will discuss some specific applications and interpretations of these measures.

We start with the measure (1.3.1) introduced by Kullback and Leibler in 1951. The Kullback-Leibler's Divergence given by (1.3.1) is very important concept in quantum information theory as well as in statistical mechanics [119]. In Bayesian statistics the KL divergence (Kullback-Leibler Divergence) given by (1.3.1) can be used as a measure of the "distance" between the prior distribution and the posterior distribution. If the logarithms are taken to the base 2 the KL divergence is also the gain in Shannon information involved in going from the prior to the posterior. In Bayesian experiment design, a design which is optimized to maximize the KL divergence between the prior and the posterior is said to be Bayes d-optimal. In coding theory, the KL divergence can be interpreted as the needed extra message-length per datum for sending messages distributed as q, if the messages are encoded using a code that is optimal for distribution p.

In probability theory and information theory, the KL divergence is a measure of the difference between two probability distributions: from a "true" distribution P to an arbitrary probability distribution Q. Although it is often intuited as a distance metric, the KL divergence is not a true metric since it is not symmetric ('divergence' rather than 'distance'). Typically P represents data, observations, or a precise calculated probability distribution. The measure Q typically represents a theory, a model, a description or an approximation of P. However one of the most important applications of KL divergence has been in Image processing and Pattern recognition [19, 38, 75, 161]. Some other applications of KL divergence can be seen in Jumarie [71], Michel et al. [105] etc.

1.2.1 Csiszár's *f* - divergence and properties

The Csiszár's f - divergence is a general class of divergence measures that includes several divergences used in measuring the distance or affinity between two probability distributions. This class is introduced by using a convex function f, defined on $(0,\infty)$. An important property of this divergence is that many known divergences can be obtained from this measure by appropriately defining the convex function f. [Shannon (1958), Renyi (1961), Csiszár's (1967, 1974), Ali & Silvey (1966), Vajda (1972), Ferentimos & Papaiopannou (1981), Berbea & Rao (1982a, b), Taneja (1995)]

Further let F be the set of convex functions $f:[0,\infty)\to\mathbf{R}$ continuous at 0, i.e. $f(0)=\lim\limits_{u\to 0} f(u)$, $\mathsf{F}_0=\{f\in\mathsf{F}: f(1)=0\}$ and let $D_- f$ & $D_+ f$ denote the *left hand side* and *right hand side* derivatives of f, respectively.

Define

$f^*\in\mathsf{F}$, by $f^*(u)=u\,f(1/u)$, the * - conjugate (convex) function of f, let a function $f\in\mathsf{F}$, Satisfying $f^*\equiv f$ be called * - self conjugate and let $\tilde{f}=f+f^*$.

In order to avoid meaningless expressions in the sequel, let us agree in the following notational conventions

$$\begin{aligned}
&0\ f^{*}\left(\frac{u}{0}\right)=u\ f\left(\frac{0}{u}\right)=u\ f(0) \quad ,u\in(0,\infty).\\
&0\ f\left(\frac{u}{0}\right)=u\ f^{*}\left(\frac{0}{u}\right)=u\ f(0) \quad ,u\in(0,\infty).\\
&0\ f\left(\frac{0}{0}\right)=0\ f^{*}\left(\frac{0}{0}\right)=0.\\
&0\ f\left(\frac{a}{0}\right)=\lim_{\in\to 0^{+}}\in f\left(\frac{a}{\in}\right)=a\lim_{u\to\infty}\frac{f(u)}{u},\ a>0.
\end{aligned} \tag{1.2.1}$$

For a convex function $f:[0,\infty)\to\mathbf{R}$, the f-Divergence measure of the probability distributions P and Q, defined by Csiszár's [24, 25] and Ali & Silvey [3], is given by

$$C_f(P,Q)=\sum_{i=1}^{n}q_i\ f\left(\frac{p_i}{q_i}\right) \tag{1.2.2}$$

It is well known that $C_f(P,Q)$ is a versatile functional form, which results in a number of popular divergence measures [29, 111, 139]. Most common choices are satisfy $f(1)=0$, so that $C_f(P,Q)$=0. Convexity ensures that the divergence measure $C_f(P,Q)$ is always non-negative. Some examples are

- $f(u)=u\log u\left(f^{*}(u)=-\log u\right)$ provides the Kullback-Leibler's measure [90, 91].
- $f(u)=|u-1|=f^{*}(u)$ results in the variational distance [88].
- $f(u)=(u-1)^2\left(f^{*}(u)=\dfrac{(u-1)^2}{u}\right)$ yields the χ^2- divergence [115] and many more.

The properties (Uniqueness theorem, Symmetry theorem, Range of values theorem and Characterization theorem) of Csiszár's generalized divergence can be seen in literature by Osterreicher [111]. Osterreicher has discussed axiomatic properties and some important classes of generalized divergence measures. Now we are discussing the following fundamental properties of $C_f(P,Q)$, which are being used in this thesis

For $f, f^*, g \in \mathrm{F}, \forall P,Q \in \Gamma_n, u \in [0,\infty)$,

(i) $C_f(P,Q) = C_{f^*}(Q,P)$

(ii) UNIQUENESS THEOREM [98]

$C_g(P,Q) = C_{f^*}(P,Q)$ *iff* $\exists\, c \in R : g(u) - f(u) = c(u-1)$

(iii) Let $c \in [D_- f(1), D_+ f(1)]$, then $g(u) = f(u) - c(u-1)$ satisfies

$g(u) \geq f(1) \forall u \in [0,\infty)$

$g(u) \geq f(1) \forall u \in [0,\infty)$

While not change the $f-$ divergence. Hence without loss of generality

(iv) SYMMETRY THEOREM [98]

$C_{f^*}(P,Q) = C_f(Q,P)$ *iff* $\exists c \in R : f^*(u) - f(u) = c(u-1)$

i.e. an $f-$ divergence is symmetric *iff* apart from an additional linear term, $c(u-1) - f^*$ self-conjugate.

(v) RANGE OF VALUES THEOREM [152]

$f(1) \leq C_f(P,Q) \leq f(0) + f^*(0)$

In the first inequality, equality holds *iff* $P = Q$. The letter provides that f is strictly convex at *u = 1*.

Many authors introduced several divergence measures. These divergences are very useful in information theory for comparing discrete probability distributions. These are as follows :-

Relative Information or Kullback-Leibler distance (Kullback and Leibler [91])

$$K(P,Q)=\sum_{i=1}^{n} p_i \log\frac{p_i}{q_i} \tag{1.2.3}$$

Variational Distance or *l₁* distance (Kolmogorov [115])

$$V(P,Q)=\sum_{i=1}^{n} \left| p_i - q_i \right| \tag{1.2.4}$$

Chi-square divergence or Pearson divergence (Pearson [115])

$$\chi^2(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^2}{q_i}=\sum_{i=1}^{n}\frac{p_i^2}{q_i}-1 \tag{1.2.5}$$

Relative Jensen-Shannon divergence (Sibson [134])

$$F(P,Q)=\sum_{i=1}^{n} p_i \log\left(\frac{2p_i}{p_i+q_i}\right) \tag{1.2.6}$$

Relative Arithmetic- Geometric divergence (Taneja [139])

$$G(P,Q)=\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2p_i}\right) \tag{1.2.7}$$

Hellinger discrimination (Hellinger [49])

$$h(P,Q)=1-B(P,Q)=\frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{p_i}-\sqrt{q_i}\right)^2 \tag{1.2.8}$$

where $B(P,Q)=\sum_{i=1}^{n}\sqrt{p_i q_i}$ (1.2.9)

is known as Bhattacharyya divergence measure [12]

Triangular discrimination (Dacunha- Castelle etc. all [150])

$$\Delta(P,Q) = 2[1 - H(P,Q)] = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i + q_i} \tag{1.2.10}$$

$$\text{Where } H(P,Q) = \sum_{i=1}^{n} \frac{2\, p_i\, q_i}{p_i + q_i} \tag{1.2.11}$$

is known as harmonic mean divergence measure

Symmetric Chi-square divergence (Dragomir etc. all [37])

$$\psi(P,Q) = \chi^2(P,Q) + \chi^2(Q,P) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2 (p_i + q_i)}{p_i q_i} \tag{1.2.12}$$

J-divergence measure (Kullback and Leibler [67, 91]**)**

$$J(P,Q) = K(P,Q) + K(Q,P) = J_R(P,Q) + J_R(Q,R) = \sum_{i=1}^{n} (p_i - q_i) \log \frac{p_i}{q_i} \tag{1.2.13}$$

Relative J-divergence measure (Dragomir etc. all [36])

$$J_R(P,Q) = \sum_{i=1}^{n} (p_i - q_i) \log \frac{p_i + q_i}{2q_i} \tag{1.2.14}$$

Arithmetic-Geometric mean divergence (Taneja [139])

$$T(P,Q) = \frac{1}{2}\left[G(P,Q) + G(Q,P)\right] = \sum_{i=1}^{n} \left(\frac{p_i + q_i}{2}\right) \log\left(\frac{p_i + q_i}{2\sqrt{p_i q_i}}\right) \tag{1.2.15}$$

Where *G(P,Q)* is the Relative AG divergence (1.2.7)

Jensen- Shannon divergence measure (Burbea and Rao[17], Sibson [134])

$$I(P,Q) = \frac{1}{2}\left[F(P,Q) + F(Q,P)\right] = \frac{1}{2}\left[\sum_{i=1}^{n} p_i \log\left(\frac{2p_i}{p_i + q_i}\right) + \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right)\right] \tag{1.2.16}$$

Where *F(P,Q)* is the Relative JS divergence (1.2.6)

d- Divergence measure (Basseville [147])

$$d(P,Q)=1-\sum_{i=1}^{n}\left(\frac{\sqrt{p_i}+\sqrt{q_i}}{2}\right)\sqrt{\frac{p_i+q_i}{2}} \tag{1.2.17}$$

Jain and Srivastava divergence (Jain and Srivastava [54]**)**

$$E^*(P,Q)=\sum_{i=1}^{n}\frac{\left(p_i-q_i\right)^2}{\sqrt{p_i\ q_i}} \tag{1.2.18}$$

Kumar and Chhina divergence (Kumar and Chhina [92])

$$S(P,Q)=\sum_{i=1}^{n}\frac{\left(p_i+q_i\right)\left(p_i-q_i\right)^2}{p_i\ q_i}\log\frac{p_i+q_i}{2\sqrt{p_i\ q_i}} \tag{1.2.19}$$

Kumar and Hunter divergence (Kumar and Hunter [93])

$$L\left(P,Q\right)=\sum_{i=1}^{n}\frac{\left(p_i-q_i\right)^2}{p_i+q_i}\log\frac{p_i+q_i}{2\sqrt{p_i\ q_i}} \tag{1.2.20}$$

Kumar and Johnson divergence (Kumar and Johnson [94])

$$\psi_M\left(P,Q\right)=\sum_{i=1}^{n}\frac{\left(p_i^2-q_i^2\right)^2}{2\left(p_i\ q_i\right)^{\frac{3}{2}}} \tag{1.2.21}$$

Jain and Mathur divergence (Jain and Mathur [60])

$$P^*\left(P,Q\right)=\sum_{i=1}^{n}\frac{\left(p_i+q_i\right)\left(p_i-q_i\right)^4\left(p_i^2+q_i^2\right)}{p_i^3\ q_i^3} \tag{1.2.22}$$

Relative information of type s [147]

$$\Psi_s(P,Q)=\begin{cases} {}^1K_s(P,Q)=\left[s(s-1)\right]^{-1}\left[\sum_{i=1}^{n} p_i^s q_i^{1-s}-1\right], s\neq 1, s>0 \\ D(P,Q)=\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right), \qquad s=1 \end{cases} \tag{1.2.23}$$

and

$$\eta_s(P,Q)=\begin{cases} (s-1)^{-1}\sum_{i=1}^{n}(p_i-q_i)\left(\frac{p_i}{q_i}\right)^{s-1}, \qquad s\neq 1 \\ J(P,Q)=\sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i}{q_i}\right), \quad s=1 \end{cases} \tag{1.2.24}$$

The following measures and generalized particular cases of (1.2.23) are introduced by

$$\Phi_s(P,Q)=\begin{cases} {}^2K_s(P,Q)=\left[s(s-1)\right]^{-1}\left[\sum_{i=1}^{n} p_i^s q_i^{1-s}-1\right], s\neq 0,1 \\ D(Q,P)=\sum_{i=1}^{n} q_i \log\left(\frac{q_i}{p_i}\right), \qquad s=0 \\ D(P,Q)=\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right), \qquad s=1 \end{cases} \tag{1.2.25}$$

and

$$\eta_s(P,Q)=\begin{cases} (s-1)^{-1}\sum_{i=1}^{n}(p_i-q_i)\left(\frac{p_i}{q_i}\right)^{s-1}, \qquad s\neq 1 \\ J(P,Q)=\sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i}{q_i}\right), \quad s=1 \end{cases} \tag{1.2.26}$$

Unified relative Jensen-Shannon and Arithmetic-Geometric divergence of type s [139]

$$\Omega_s(P,Q)=\begin{cases} FG_s(P,Q)=\left[s(s-1)\right]^{-1}\left[\sum_{i=1}^{n} p_i\left(\dfrac{p_i+q_i}{2p_i}\right)^{s}-1\right] & s\neq 0,1\\ F(P,Q)=\sum_{i=1}^{n} p_i \log\left(\dfrac{2p_i}{p_i+q_i}\right), & s=0\\ G(P,Q)=\sum_{i=1}^{n}\left(\dfrac{p_i+q_i}{2}\right)\log\left(\dfrac{p_i+q_i}{2p_i}\right), & s=1 \end{cases} \tag{1.2.27}$$

The modified version of Unified relative Jensen-Shannon and Arithmetic-Geometric divergence of type s (1.2.27) is given by

$$W_s(P,Q)=\Omega_s(P,Q)+\Omega_s(Q,P)$$
$$=\begin{cases} IT_s(P,Q)=\left[s(s-1)\right]^{-1}\sum_{i=1}^{n}\left[\left(p_i^{1-s}+q_i^{1-s}\right)\left(\dfrac{p_i+q_i}{2}\right)^{s}-2\right], & s\neq 0,1\\ I(P,Q)=\dfrac{1}{2}\left[\sum_{i=1}^{n} p_i\log\left(\dfrac{2p_i}{p_i+q_i}\right)+\sum_{i=1}^{n} q_i\log\left(\dfrac{2q_i}{p_i+q_i}\right)\right], & s=0\\ T(P,Q)=\sum_{i=1}^{n}\left(\dfrac{p_i+q_i}{2}\right)\log\left(\dfrac{p_i+q_i}{2\sqrt{p_iq_i}}\right), & s=1 \end{cases} \tag{1.2.28}$$

J-divergence of type s [139]

$$\nu_s(P,Q)=\begin{cases} J_S(P,Q)=\left[s(s-1)\right]^{-1}\sum_{i=1}^{n}\left[p_i^s q_i^{1-s}+p_i^{1-s}q_i^{s}-2\right], & s=0,1\\ J(P,Q)=\sum_{i=1}^{n}(p_i-q_i)\log\dfrac{p_i}{q_i}, & s=0,1 \end{cases} \tag{1.2.29}$$

Relative J-divergence of type s [139]

$$\varsigma_s(P,Q)=\begin{cases} D_s(P,Q)=\left[s-1\right]^{-1}\sum_{i=1}^{n}(p_i-q_i)\left(\dfrac{p_i+q_i}{2q_i}\right)^{s-1}, & s\neq 1\\ J(P,Q)=\sum_{i=1}^{n} p_i\log\left(\dfrac{p_i}{q_i}\right), & s=1 \end{cases} \tag{1.2.30}$$

Sibson Information Radius (Sibson [134], Berbea & Rao [17])

$$I_r(P,Q)=\begin{cases}(r-1)^{-1}\left[\sum_{i=1}^{n}\left(\frac{p_i^{\,r}+q_i^{\,r}}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{1-r}-1\right], & r\neq 1, r>0\\ I(P,Q)=\sum_{i=1}^{n}\left(\frac{p_i\log p_i+q_i\log q_i}{2}\right)-\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2}\right), & if\ r=1\end{cases} \tag{1.2.31}$$

Taneja Divergence Measure (Taneja [152, 153])

$$T_r(P,Q)=\begin{cases}(r-1)^{-1}\left[\sum_{i=1}^{n}\left(\frac{p_i^{1-r}+q_i^{1-r}}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{r}-1\right], & r\neq 1, r>0\\ T(P,Q)=\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2\sqrt{p_iq_i}}\right), & if\ \ r=1\end{cases} \tag{1.2.32}$$

1.2.2 New generalized Divergence

We did a detail study about Csiszár's *f*- divergence in previous section. Similarly, Jain and Saraswat[57] introduced and characterized a new generalized divergence, given by

$$S_f(P,Q)=\sum_{i=1}^{n}q_i\, f\left(\frac{p_i+q_i}{2q_i}\right) \tag{1.2.33}$$

Where $f:(0,\infty)\to R$ (set of real numbers) is real, continuous, and convex function and $P,Q\in\Gamma_n$.

We can obtain several well-known divergence measures by suitably defining the convex function in (1.2.33).

The following results are presented by Jain & Saraswat [57]

Proposition 1.2.2.1: Let $f:[0,\infty)\to R$ be the convex function $P,Q\in\Gamma_n$. Then we have the following inequality

$$S_f(P,Q) \geq f(1) \tag{1.2.34}$$

If ***f*** is normalized i.e. $f(1) = 0$, then $S_f(P,Q) \geq 0$ and if f is strictly convex and equality holds *iff* $p_i = q_i \ \forall \ i = 1, 2, 3, ..., n$

$$\text{i.e. } S_f(P,Q) \geq 0 \quad \textit{and} \ S_f(P,Q) = 0 \ \textit{if} \quad P = Q \tag{1.2.35}$$

Proposition 1.2.2.2: If f_1 and f_2 are two convex functions and $F = c_1 \ f_1 + c_2 \ f_2$ then

$$S_F(P,Q) = c_1 S_{f_1}(P,Q) + c_2 S_{f_2}(P,Q) \tag{1.2.36}$$

Where c_1 and c_2 are constants and $P, Q \in \Gamma_n$

1.2.3 Classes of *f* - Divergences

In this unit, we present some of the more intensively studied classes of $f-$Divergences in terms of their convex function f. The historic references are intended to give some insinuations as to their making.

(I) The Class of χ^α-divergences

Total Variation Distance [115]

$$f(u) = |u - 1|$$

K. Pearson (1900) [115]

$$\chi^2(u) = (u-1)^2$$

Kagan (1963) [74], Vajda (1973) [154], Boekee (1977) [13]

$$\chi^\alpha(u) = |u-1|^\alpha, \alpha \geq 1$$

(II) Dichotomy Class

Kullback and Leibler (1951) [91]

$$f(u) = u \ln u$$

Likelihood Disparity

$$f^*(u) = -\ln u$$

K. Pearson (1900) [115]

$$\chi^2(u) = (u-1)^2$$

Neyman (1949) [108]

$$(\chi^2)^*(u) = \frac{(u-1)^2}{u}$$

Liese & Vajda (1987) [98]

$$f^{\alpha}(u) = \begin{cases} u - 1 - \ln u & \text{for } \alpha = 0 \\ \dfrac{\alpha u + 1 - \alpha - u^{\alpha}}{\alpha(1-\alpha)} & \text{for } \alpha \in \mathbf{R} / \{0,1\} \\ 1 - u + u \ln u & \text{for } \alpha = 1 \end{cases}$$

Read & Cressie (1988) [124]

$$f_{\lambda}(u) = \frac{u^{\lambda+1} - 1}{\lambda(1+\lambda)} \text{ for } \lambda = \alpha - 1 \in \mathrm{R} \setminus \{-1, 0\}$$

Jeffreys (1946) [68]

$$\tilde{f}(u) = (u-1)\ln u$$

Csiszár & Fischer (1962) [25]

$$\tilde{f}^{(s)}(u) = \begin{cases} f^{s}(u) = 1 + u - (u^{s} + u^{1-s}), & \text{for } 0 < s < 1 \\ (u-1)\ln u & \text{for } s = 1 \\ \dfrac{u + 1 - (u^{s} + u^{1-s})}{s(1-s)} & \text{for } s \in (0,1) U (1,\infty) \end{cases}$$

(III) Matusita's Divergences

Matusita's (1955) [102]

$$f^{1/2}(u) = (\sqrt{u} - 1)^2$$

Matusita's (1964) [103] Boekee (1977) [13]

$$f^{\alpha}(u) = |u^{\alpha} - 1|^{1/\alpha}, \; 0 < \alpha \leq 1$$

Renyi's Divergences (This class doesn't belong to the family of *f*–Divergences and the functions $g^{\alpha}(u) = u^{\alpha}, \alpha \in (0,1)$ are concave).

(Hellinger (1909): $g^{1/2}(u) = \sqrt{u}$)

Bhattacharyya (1946) [12]

$$-\ln\left(\sum_{x \in \Omega} \sqrt{p(x)q(x)}\right)$$

Chernoff (1952) [20]

$$-\min_{0 \leq \alpha \leq 1} \ln\left(\sum_{x \in \Omega} p(x)\left(\frac{p(x)}{q(x)}\right)^{\alpha}\right)$$

Renyi (1961) [125]

$$R_{\alpha}(Q,P) = \begin{cases} \sum\limits_{x \in \Omega} q(x) \ln\left(\dfrac{q(x)}{p(x)}\right) & \textit{for } \alpha = 1 \\ \dfrac{1}{\alpha - 1}\left[\sum\limits_{x \in \Omega} p(x)\left(\dfrac{q(x)}{p(x)}\right)^{\alpha}\right] & \text{for } \alpha \in \{0, \infty\} \setminus \{1\} \end{cases}$$

(IV) Elementary Divergences

Feldman & Osterreicher (1981) [40]

$$f_t(u) = \max(u - t, 0), \qquad \text{t} \geq 0$$

(V) Puri-Vincze Divergences

Le Cam (1986) [96], Topsoe (1999) [150]

$$\phi_2(u) = \frac{1}{2}\frac{(1-u)^2}{u+1}$$

Puri & Vincze (1990) [118], Kafka, Osterreicher & Vincze (1989) [73]

$$\phi_k(u) = \frac{1}{2}\frac{(1-u)^k}{(u+1)^{k-1}}, \quad \text{k} \geq 1$$

(V) Divergences of Arimoto-type Perimeter Divergence:

Osterreicher (1982) [111], Reschenhofer & Bomze (1991) [126]

$$f(u) = \sqrt{1+u^2} - \frac{1+u}{\sqrt{2}}$$

Perimeter-type Divergence: *Osterreicher* (1996) [111]

$$f_p(u) = \left\{ \begin{array}{ll} (1+u^p)^{1/p} - 2^{\frac{1}{p}-1}(1+u) & \text{for } p \in (1,\infty) \\ |1-u|/2 & \text{for } p = \infty \end{array} \right\}$$

Osterreicher & Vajda (1997) [111]

$$f_\beta(u) = \left\{ \begin{array}{ll} \dfrac{1}{1-\frac{1}{\beta}}\{(1+u^\beta)^{1/\beta} - 2^{1/\beta-1}(1+u)\} & \text{If } \beta \in \{0,\infty\}\backslash\{1\} \\ (1+u)\ln(2) + u\ln u - (1+u)\ln(1+u) & \text{if } \beta=1 \\ |1-u|/2 & \text{If } \beta=\infty \end{array} \right\}$$

1.3 Information inequalities

Various mathematicians and researchers have used these inequalities for a variety of determinations. Taneja has used it for obtaining bounds on symmetric and non-symmetric divergence measures in terms of relative information of type s [141], for obtaining relationships among mean divergence measures [144], for obtaining bounds on symmetric divergence measures in terms of non-symmetric divergence measures [142]. Further Taneja has shown that various one-parameter generalizations of symmetric and non-symmetric divergence measures can be written as particular cases of Csiszár's f - Divergences [95]. Osterreicher [111] has discussed basic general properties of f - divergences including their axiomatic properties and some important classes. Different kinds of bounds on the information divergence measures have been studied during the recent past [150, 29- 37]. In [147], Kumar and Taneja unified and generalized information bounds for $C_f\,(P,Q)$ studied by Dragomir [29- 37]**.** The main results [147] are given in the following theorem.

Theorem1.3.1: Let $f : R_+ \to R$ be a mapping which is normalized i.e. $f(1) = 0$ and suppose that

(i) f is twice differentiable function on (r, R), $0 \le r \le 1 \le R < \infty$

(ii) There exists real constants *m*, *M*, such that *m* < *M* and $m \le t^{2-S} f''(t) \le M,\ \forall\, t \in (r,R),\ s \in R$.

(f' *and* f'' denote the first and second order derivatives of function *f*)

If $P,Q \in \Gamma_n$ are discrete probability distributions with $0 < r \le \dfrac{p_i}{q_i} \le R < \infty$,

$$m\,\Phi_S(P,Q) \le \quad C_f(P,Q) \quad \le M\ \Phi_S(P,Q), \tag{1.3.1}$$

And

$$m\,\{\,\eta_s(P,Q) - \Phi_S(P,Q)\,\} \le C_\rho(P,Q) - C_f\,(P,Q) \le M\,\{\,\eta_s(P,Q) - \Phi_s(P,Q)\,\} \tag{1.3.2}$$

Where

$$\Phi_s(P,Q)=\begin{cases} {}^2K_s(P,Q), & s\neq 0,1, \\ K(Q,P), & s=0, \\ K(P,Q), & s=1, \end{cases} \tag{1.3.3}$$

$${}^2K_s(P,Q)=\left[s(s-1)\right]^{-1}\left[\sum p^s q^{1-s}-1\right], \quad s\neq 0,1, \tag{1.3.4}$$

$$K(P,Q)=\sum_{i=1}^{n} p_i \log\frac{p_i}{q_i} \tag{1.3.5}$$

$$C_\rho(P,Q)=\sum_{i-1}^{n}(p_i-q_i)\, f'\left(\frac{p_i}{q_i}\right), \tag{1.3.6}$$

$$\eta_s(P,Q)=C_{\Phi'_S}\left(\frac{P^2}{Q},P\right)-C_{\Phi'_S}(P,Q)$$

$$=\begin{cases} (s-1)^{-1}\sum_{i=1}^{n}(p_i-q_i)\left(\frac{p_i}{q_i}\right)^{s-1}, & s\neq 1 \\ \sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i}{q_i}\right), & s=1 \end{cases} \tag{1.3.7}$$

As a consequence of this theorem, following information inequalities which are interesting from the *information-theoretic* point of view are also obtained in [134].

(i) The case s = 2 provides the information bounds in terms of the Chi-square divergence $\chi^2(P,Q)$,

$$\frac{m}{2}\chi^2(P,Q)\leq C_f(P,Q)\leq\frac{M}{2}\chi^2(P,Q), \tag{1.3.8}$$

And

$$\frac{m}{2}\chi^2(P,Q)\leq C_\rho(P,Q)-C_f(P,Q)\leq\frac{M}{2}\chi^2(P,Q) \tag{1.3.9}$$

Where $\chi^2(P,Q)$ is given by (1.2.5).

(ii) For s = 1, the information bounds in terms of the Kullback-Leibler divergence, *K (P, Q),*

$$m\,K(P,Q) \le C_f\,(P,Q) \le M\ K(P,Q) \tag{1.3.10}$$

And

$$m\,K(P,Q) \le C\rho\,(P,Q) - C_f\,(P,Q) \le M\ K(P,Q) \tag{1.3.11}$$

(iii) For s = ½, yields the information bounds in terms of the Hellinger's discrimination

$$h\,(P,Q),\ 4\,m\,h(P,Q) \le C_f\,(P,Q) \le 4\,M\ h(P,Q) \tag{1.3.12}$$

And

$$4\,m\ \left(\frac{1}{4}\eta_{1/2}(P,Q) - h(P,Q)\right) \le C_\rho(P,Q) - C_f(P,Q) \le 4\,M\ \left(\frac{1}{4}\eta_{1/2}(P,Q) - h(P,Q)\right) \tag{1.3.13}$$

(iv) For s = 0, the information bounds in terms of the Kullback-Leibler and Chi-square divergence

$$m\,K(P,Q) \le C_f(P,Q) \le M\ K(P,Q) \tag{1.3.14}$$

And

$$m\left\{\chi^2(Q,P) - K(Q,P)\right\} \le C\rho(P,Q) - C_f\,(P,Q) \le M\ \left\{\chi^2(Q,P) - K\,(Q,P)\right\} \tag{1.3.15}$$

These inequalities have been considerably improved by Taneja [143]. Also very recently Baig and Dar [5, 6] have obtained bounds on divergence measures of Csiszár's $f-$divergence D class in terms of Relative J-Divergence of type s and Renyi's entropy of order α. Some other important inequalities related to Csiszár's $f-$Divergence can be seen in Csiszár's and Korner [25] and Dragomir [34].

1.3.1 Mean and their Applications in Information Theory

In the previous section, we have seen that lot of recently proposed symmetric divergence measures are based on arithmetic, geometric and other means and as such we could use the inequalities among these means for obtaining bounds on these divergence measures.

Now, let us define some means for $a, b > 0$

$$A(a,b) = \frac{a+b}{2} = \text{Arithmetic mean.} \tag{1.3.16}$$

$$B(a,b) = \sqrt{a\,b} = \text{Geometric mean.} \tag{1.3.17}$$

$$H(a,b) = \frac{2\,a\,b}{a+b} = \text{Harmonic mean.} \tag{1.3.18}$$

$$L^*(a,b) = \frac{a-b}{\log a - \log b}, a \neq b = \text{Logarithmic mean.} \tag{1.3.19}$$

Now for $P, Q \in \Gamma_n$, put $a = p_i$ and $b = q_i$ in above means and then sum over i = 1, 2, 3, …, n, we obtain

$$A(P,Q) = \sum \frac{p_i + q_i}{2} = 1 = \text{Arithmetic mean divergence} \tag{1.3.20}$$

$$B(P,Q) = \sum_{i=1}^{n} \sqrt{p_i\ q_i} = \text{Bhattacharya distance} \tag{1.3.21}$$

$$H(P,Q) = \sum_{i=1}^{n} \frac{2\,p_i\,q_i}{p_i + q_i} = \text{Harmonic mean divergence} \tag{1.3.22}$$

$$L^*(P,Q) = \sum \frac{p_i - q_i}{\log p_i - \log q_i}, p_i \neq q_i = \text{Logarithmic mean divergence} \tag{1.3.23}$$

Here *B(P,Q)* is well-known Bhattacharya distance(Bhattacharyya[12]). We can see some small inequality and equality relations

$$H(P,Q) \leq B(P,Q) \leq A(P,Q) \tag{1.3.24}$$

$$\Delta(P,Q) = 2[A(P,Q) - H(P,Q)] \tag{1.3.25}$$

$$h(P,Q) = [A(P,Q) - B(P,Q)] \tag{1.3.26}$$

(1.3.25) and (1.3.26) define small equality relations for Triangular discrimination and Hellinger discrimination respectively with above defined quantities.

Various information divergence measures are used in several disciplines such as in Communication theory, Cybernetics, Biology, Psychology, Economics, Statistics, Thermodynamics, Questionnaire theory, Probability theory etc. This makes Information Theory a useful tool for researchers of other disciplines

In the present work, various measures in information theory especially measures made up of two probability distributions have been studied. A number of new information measures have been proposed and their relationship with some well- known measures in information theory has been established. Some of the specific applications and interpretation of these measures are also discussed. Further a brief description of all the problems discussed in the subsequent chapters of the thesis has been given in the following sections.

Chapter 2 introduces a new class of, non-parametric symmetric, divergence measures. The properties of these divergence measures are studied and their bounds in terms of some familiar information divergence measures are derived. **Chapter 3** introduces different series of some new classes of information divergence measures, which belong to the family of Csiszár's f - divergences. Some inequalities & equalities among new divergence measures and chi-square divergence, triangular discrimination, Jain and Saraswat divergence, Hellinger discrimination, Variational distance, harmonic mean divergence, are present in results respectively. **Chapter 4** will introduce two new series of information divergence measures using Jain and Saraswat generalized f - divergence measure to obtain various new information inequalities on these new series of divergence measures with some well-known information measures. **Chapter 5** introduces a non-parametric theoretic based

exponential information divergence measure. Further for first time, we have derived some inequalities for this exponential information divergence measure in terms of some valuable information divergence measures. Some numerical illustrations are carried out, based on two distinct discrete probability distributions. **Chapter 6** introduces a non-parametric symmetric information divergence measure.. Its properties are studied and discussed. Further, we have derived some new bounds for this divergence measure in terms of some recognized divergence measures based on two distinct probability distributions. The information divergence measures are used to find out distance or difference or affinity between two probability distributions.
Chapters 7 include the conclusions of the work reported in the thesis and also underline the future scope of work.

Lastly, references, candidate's research profile and candidate's academic & personal profile.

2

New Information Divergence Measures of Csiszár’s *f*-Divergence Family and Their Bounds

2.1 Introduction

During past years, a large number of new divergence measures have been proposed and studied broadly by various researchers and mathematicians in their related fields of research. Lots of these information divergence measures belong to the family of Csiszár’s f - divergence and most of them are symmetric information divergence measures with respect to the used probability distributions.

In this chapter, a new class of, non-parametric symmetric, divergence measures are introduced. The properties of these divergence measures are studied and their bounds in terms of some familiar information divergence measures are derived. Further we found that these information measures are meticulously associated with some familiar information divergence measures.

The whole chapter is structured as follows. In section 2.2, a new class of symmetric divergence measures, have been studied, which belong to the Csiszár’s

f – Divergence family. In section 2.3, some new convex functions (2.3.2, 2.3.4, 2.3.6) have been extracted, from the convex function (2.2.1) studied in section 2.2, using the property that, sum of convex functions is also a convex function. In this section 2.3, some non-parametric information divergence measures (2.3.3, 2.3.5, 2.3.7), corresponding to convex functions (2.3.2, 2.3.4, 2.3.6), which belongs to the family of Csiszár's f – divergence is also derived. In section 2.4, some general new information inequalities among new f – divergence measure (2.3.3) and relative J - divergence, J - divergence, relative JS - divergence, relative AG divergence and Kullback-Leibler divergence are derived. In section 2.5, some inequalities & equalities among new divergence measures and chi-square divergence, triangular discrimination, logarithmic mean divergence, relative information, relative J-divergence, relative arithmetic-geometric divergence, relative Jenson-Shannon divergence, arithmetic mean divergence, arithmetic-geometric mean divergence, harmonic mean divergence, J-divergence and relations are present in results 2.5.1, 2.5.2, 2.5.3, 2.5.4 & 2.5.5 respectively. Section 2.6, concludes the results of the chapter.

2.2 New Information Divergence Measures

In this section we shall find out the new information divergence measure with the help of following convex function. Let us consider the function $f:(0,\infty)\rightarrow R$ such than

$$f_k(t)=\frac{(t-1)^{k+1}}{(t+1)^k};\quad k=1,3,5,\ldots \tag{2.2.1}$$

Then

$$f_k^{'}(t)=\frac{(t-1)^k(2k+t+1)}{(t+1)^{k+1}} \tag{2.2.2}$$

And

$$f_k^{''}(t)=\frac{(t-1)^{k-1}4k(k+1)}{(t+1)^{k+2}} \tag{2.2.3}$$

The function $f_k(t)$ is convex since $f_k^{'}(t) \geq 0$, t > 0, k =1, 3, 5… and normalized also since $f(1) = 0$. Figure: 2.1 shows the behavior of the function $f_k(t)$, which is always convex if k = 1, 3, 5… (Odd numbers) with t > 0

Ali- Silvey [3] and Csiszár's [24, 25] introduced the generalized measure of information using *f*-divergence measure given by

$$C_f(P,Q) = \sum_{i=1}^{n} q_i \; f\left(\frac{p_i}{q_i}\right) \tag{2.2.4}$$

Applying Csiszár's *f*-divergence properties on equation (2.2.1), we get

$$Y_k^c(P,Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^{k+1}}{(p_i + q_i)^k}; \qquad k = 1,3,5,... \tag{2.2.5}$$

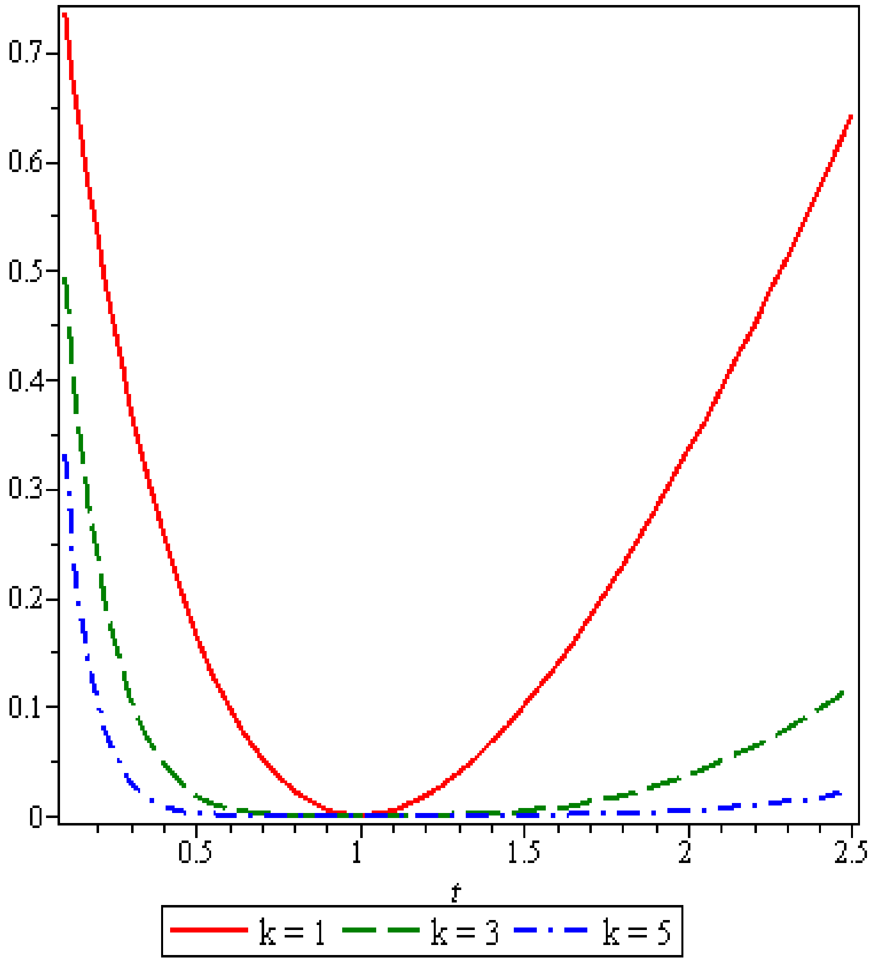

FIGURE: 2.1 Graph of the convex function $f_k(t)$

New information divergence measure $Y_k^c(P,Q)$ is symmetric divergence measure, since

$$Y_k^c(P,Q) = Y_k^c(Q,P) \tag{2.2.6}$$

2.3 Extraction of Some Other Divergence Measures

From equation (2.1.1) for k = 1, 3, 5… we get the following convex functions

$$f_1 = \frac{(t-1)^2}{(t+1)},\ f_3 = \frac{(t-1)^4}{(t+1)^3},\ f_5 = \frac{(t-1)^6}{(t+1)^5},\ f_7 = \frac{(t-1)^8}{(t+1)^7}\ldots$$

We know that, sum of convex functions is also a convex function

$c_1 f_1(t) + c_3 f_3(t) + c_5 f_5(t) + c_7 f_7 + \ldots$, is also a convex function. Where $c_1, c_3, c_{5,} c_7 \ldots$

are arbitrary positive constants and at least one c_i (i = 1, 3, 5,…) is not equal to zero.

Now

$$F_k(t) = c_1 f_1(t) + c_3 f_3(t) + c_5 f_5(t) + c_7 f_7(t) + \ldots$$

$$= c_1 \frac{(t-1)^2}{(t+1)^1} + c_2 \frac{(t-1)^4}{(t+1)^3} + c_3 \frac{(t-1)^6}{(t+1)^5} + c_4 \frac{(t-1)^8}{(t+1)^7} + \ldots \tag{2.3.1}$$

Now taking, $c_1 = 1, c_3 = 1/3, c_5 = 1/5, c_7 = 1/7 \quad \ldots$

Therefore

$$F_1(t) = \frac{(t-1)^2}{(t+1)^1} + \frac{1}{3}\frac{(t-1)^4}{(t+1)^3} + \frac{1}{5}\frac{(t-1)^6}{(t+1)^5} + \frac{1}{7}\frac{(t-1)^8}{(t+1)^7} + \ldots$$

$$= (t-1)\left[\left(\frac{t-1}{t+1}\right)^1 + \frac{1}{3}\left(\frac{t-1}{t+1}\right)^3 + \frac{1}{5}\left(\frac{t-1}{t+1}\right)^5 + \frac{1}{7}\left(\frac{t-1}{t+1}\right)^7 + \ldots\right]$$

$$= (t-1)\frac{1}{2}\log\left\{\frac{1+\left(\frac{t-1}{t+1}\right)}{1-\left(\frac{t-1}{t+1}\right)}\right\}$$

$$F_1(t) = \frac{1}{2}(t-1)\log(t) \tag{2.3.2}$$

Now corresponding divergence measure of Csiszár's *f*- divergence class

$$Y_1^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right) \tag{2.3.3}$$

Next taking $c_1 = 0,\ c_3 = 1,\ c_5 = 1/3,\ c_7 = 1/5\ ...$ in equation (2.3.1), we get

$$F_3(t) = \frac{1}{2}\frac{(t-1)^3}{(t+1)^2}\log(t) \tag{2.3.4}$$

Hence, corresponding divergence measure of Csiszar's f-divergence class is as under

$$Y_3^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}\frac{(p_i - q_i)^3}{(p_i + q_i)^2}\log\left(\frac{p_i}{q_i}\right) \tag{2.3.5}$$

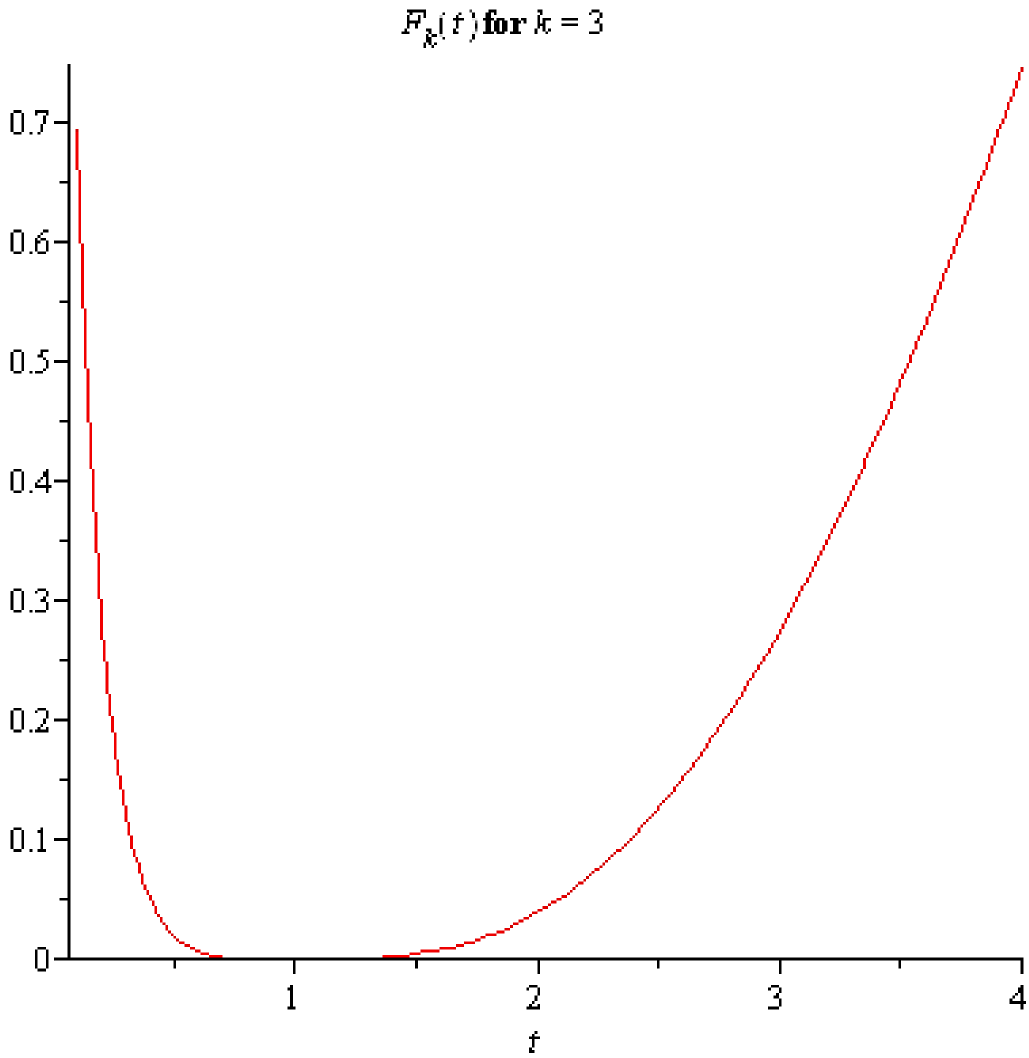

FIGURE 2.2: Graph of convex function $F_k(t)$ for k = 3 corresponding to $Y_3^*(P,Q)$

And the corresponding series of divergence measures of Csiszár’s *f*-divergence class

$$Y_k^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}\frac{(p_i - q_i)^k}{(p_i + q_i)^{k-1}}\log\left(\frac{p_i}{q_i}\right); k = 1,3,5... \tag{2.3.7}$$

2.4 Some Bounds for $Y_1^*(P,Q)$

In this section, we will derive some divergence inequalities and equalities for $Y_1^*(P,Q)$ in terms of other well-known information divergence measures.

Now, new divergence measure (2.3.3),

$$Y_1^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right)$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left[p_i \log\left(\frac{p_i}{q_i}\right) - q_i \log\left(\frac{p_i}{q_i}\right)\right]$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left[p_i \log\left(\frac{p_i}{q_i}\right) + q_i \log\left(\frac{q_i}{p_i}\right)\right]$$

$$= \frac{1}{2}\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) + \frac{1}{2}\sum_{i=1}^{n} q_i \log\left(\frac{q_i}{p_i}\right)$$

Using equation (1.2.3), we get, the result

$$Y_1^*(P,Q) = \frac{1}{2}\left[K(P,Q) + K(Q,P)\right] \tag{2.4.1}$$

From equation (1.2.13), we know that

$$K(P,Q) + K(Q,P) = J(P,Q) = J_R(P,Q) + J_R(Q,P)$$

So that

$$Y_1^*(P,Q) = \frac{1}{2}J(P,Q) \tag{2.4.2}$$

And

$$Y_1^*(P,Q) = \frac{1}{2}\left[J_R(P,Q) + J_R(Q,P)\right] \tag{2.4.3}$$

Therefore, (2.4.1), (2.4.2) and (2.4.3), shows the equality relations among new divergence measure (2.3.3), relative information (1.2.3), J- divergence measure (1.2.13) and relative J- divergence measure (1.2.14) respectively.

Next, we know that,

$$0 < p_i \leq 1;\ 0 < q_i \leq 1\ ;\ 0 < \frac{p_i}{q_i}$$

$$\Rightarrow \quad \frac{p_i}{q_i} < \left(1 + \frac{p_i}{q_i}\right)$$

$$\log\left(\frac{p_i}{q_i}\right) < \log\left\{\frac{2}{2}\left(1 + \frac{p_i}{q_i}\right)\right\}$$

$$\Rightarrow \qquad (p_i - q_i)\log\left(\frac{p_i}{q_i}\right) < (p_i - q_i)\log\left(2\,\frac{p_i + q_i}{2q_i}\right)$$

$$\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right) < \sum_{i=1}^{n}(p_i - q_i)\left[\log 2 + \log\left(\frac{p_i + q_i}{2\,q_i}\right)\right]$$

$$\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right) < \sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i + q_i}{2\,q_i}\right) + \sum_{i=1}^{n}(p_i - q_i)\log(2)$$

Hence, using (1.2.6) (1.2.7) and (1.2.14), we get the inequalities (2.4.4) & (2.4.5) among new divergence measure, Relative JS- divergence, Relative AG- divergence and Relative J- divergence

$$2\,Y_1^*(P,Q) < F(Q,P) + G(Q,P) + \log 2 \sum_{i=1}^{n}(p_i - q_i) \tag{2.4.4}$$

And

$$2\,Y_1^*(P,Q) < J_R(P,Q) + \log 2 \sum_{i=1}^{n}(p_i - q_i) \tag{2.4.5}$$

Equality holds for $P = Q$ only.

2.5 Inequalities and Equalities

In this section we will derive some other inequalities and equalities for $Y_k^*(P,Q)$(for the case k = 1, k = 3) in terms of some well-known divergence measures.

Proposition 2.5.1: Let $P,Q \in \Gamma_n$, then we have the following new inequality

$$Y_3^*(P,Q) \leq \frac{1}{2}\Delta(P,Q)\frac{1}{A(P,Q)}\left[K(P,Q)+K(Q,P)\right] \tag{2.5.1}$$

And

$$Y_3^*(P,Q) \leq \frac{1}{2}\Delta(P,Q)\quad\frac{1}{A(P,Q)}\quad\left[J_R(P,Q)+J_R(Q,P)\right] \tag{2.5.2}$$

Where $Y_3^*(P,Q)$ new divergence measure, *A(P,Q)* Arithmetic mean divergence, $\Delta(P,Q)$ Triangular discrimination, *K(P,Q)* Kullback-Leibler divergence (Relative information) and $J_R(P,Q)$ Relative J- divergence measure are given by (2.3.5), (1.3.19), (1.2.10), (1.2.3) and (1.2.14) respectively.

Proof: From (2.3.5), we have

$$Y_3^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}\frac{(p_i - q_i)^3}{(p_i + q_i)^2}\log\left(\frac{p_i}{q_i}\right)$$

$$\leq \sum_{i=1}^{n}\frac{(p_i - q_i)^3}{(p_i + q_i)^2}\log\left(\frac{p_i}{q_i}\right)$$

$$= \frac{1}{2}\sum_{i=1}^{n}\frac{(p_i - q_i)^2}{p_i + q_i}\quad\sum_{i=1}^{n}\frac{2}{p_i + q_i}\quad\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right)$$

$$Y_3^*(P,Q)\ \leq\ \frac{1}{2}\Delta(P,Q)\quad\frac{1}{A(P,Q)}\quad\left[K(P,Q)+K(Q,P)\right]$$

Hence the required inequality

We know that, $K(P,Q)+K(Q,P)=J_R(P,Q)+J_R(Q,P)$

So, $Y_3^*(P,Q)\le \frac{1}{2}\,\Delta(P,Q)\quad \frac{1}{A(P,Q)}\quad [J_R(P,Q)+J_R(Q,P)]$

Hence the required inequality

Proposition 2.5.2: Let $P,Q\in\Gamma_n$ and then we have the following new relation

$$Y_3^*(P,Q)=\frac{1}{8}\left[\chi^2(P,Q)+\chi^2(Q,P)\right]\ \frac{H(P,Q)}{[A(P,Q)]^2}\left[G(P,Q)-G(Q,P)+K(Q,P)\right] \qquad (2.5.3)$$

Where $Y_3^*(P,Q)$ new divergence measure, $A(P,Q)$ Arithmetic mean divergence, $\chi^2(P,Q)$ Chi-Square divergence, $K(P,Q)$ Kullback-Leibler divergence (Relative information), $H(P,Q)$ Harmonic mean divergence and $G(P,Q)$ Relative AG- divergence are given by (2.3.5), (1.3.19), (1.2.5), (1.2.3), (1.3.21) and (1.2.7) respectively.

Proof: From (2.3.5), we have

$$Y_3^*(P,Q)=\frac{1}{2}\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i}{q_i}\right)$$

$$Y_3^*(P,Q)=\frac{1}{2}\sum_{i=1}^{n}\frac{(p_i-q_i)^2(p_i+q_i)}{p_iq_i}\quad \frac{p_iq_i}{(p_i+q_i)^3}\quad (p_i-q_i)\log\left(\frac{p_i}{q_i}\right)$$

$$Y_3^*(P,Q)=\frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^2(p_i+q_i)}{p_iq_i}\quad \frac{2p_iq_i}{(p_i+q_i)}\quad \left(\frac{1}{p_i+q_i}\right)^2\quad (p_i+q_i-2q_i)\log\left(\frac{2p_i}{p_i+q_i}\ \frac{p_i+q_i}{2q_i}\right)$$

$$Y_3^*(P,Q)=\frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^2(p_i+q_i)}{p_iq_i}\quad\sum_{i=1}^{n}\frac{2p_iq_i}{p_i+q_i}\;\left(\sum_{i=1}^{n}\frac{1}{p_i+q_i}\right)^2$$
$$[\sum_{i=1}^{n}(p_i+q_i)\log\left(\frac{2p_i}{p_i+q_i}\right)+\sum_{i=1}^{n}(p_i+q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)-2q_i\log\left(\frac{p_i}{q_i}\right)]$$

$$Y_3^*(P,Q)=\frac{1}{8}\left[\chi^2(P,Q)+\chi^2(Q,P)\right]\quad H(P,Q)\quad\left[\frac{1}{A(P,Q)}\right]^2\quad\left[G(Q,P)-G(P,Q)+K(Q,P)\right]$$

Hence the result

Proposition 2.5.3: Let $P,Q\in\Gamma_n$, and then we have the following new relation

$$Y_3^*(P,Q)\le L^*(P,Q)\left[\frac{1}{A(P,Q)}\right]^2\;\left[J(P,Q)\right]^2 \tag{2.5.4}$$

Where $Y_3^*(P,Q)$ new divergence measure, $A(P,Q)$ Arithmetic mean divergence, $L^*(P,Q)$ Logarithmic mean divergence, and $J(P,Q)$ J- divergence measure are given by (2.3.5), (1.3.20), (1.3.23) and (1.2.13) respectively.

Proof: From (2.3.5), we have

$$Y_3^*(P,Q)=\frac{1}{2}\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i}{q_i}\right)$$

$$\le\quad\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i}{q_i}\right)$$

$$=\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i}{q_i}\right)\quad\frac{\log\left(\frac{p_i}{q_i}\right)}{\log\left(\frac{p_i}{q_i}\right)}$$

$$=\frac{1}{4}\left[\sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i}{q_i}\right)\right]^2\left(\sum_{i=1}^{n}\frac{p_i-q_i}{\log p_i-\log q_i}\right)\left(\sum_{i=1}^{n}\frac{2}{p_i+q_i}\right)^2$$

$$= \frac{1}{4}\left[\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right)\right]^2 \left(\sum_{i=1}^{n}\frac{p_i - q_i}{\log p_i - \log q_i}\right)\left(\sum_{i=1}^{n}\frac{2}{p_i + q_i}\right)^2$$

$$Y_3^*(P,Q) \le [J(P,Q)]^2 \quad L^*(P,Q) \quad \left[\frac{1}{A(P,Q)}\right]^2$$

Hence the inequality

Proposition 2.5.4: let $P,Q \in \Gamma_n$ and then we have the following new inequality

$$2Y_1^*(P,Q) - K(P,Q) \le J_R(P,Q) - F(Q,P) + \log 2 \tag{2.5.5}$$

Where $Y_1^*(P,Q)$ new divergence measure, *F(P,Q* Relative JS- divergence, *K(P,Q)* Kullback-Leibler divergence and $J_R(P,Q)$ Relative J- divergence measure, are given by (2.3.3), (1.2.6), (1.2.3) and (1.2.14) respectively.

Proof: We know that,

$$0 \le p_i \le 1, 0 \le q_i \le 1$$

$$\frac{p_i}{q_i} < 1 + \frac{p_i}{q_i}$$

$$\log\left(\frac{p_i}{qi}\right) < \log\left(1 + \frac{p_i}{q_i}\right)$$

$$\log\left(\frac{p_i}{qi}\right) < \log\left(\frac{p_i + q_i}{2q_i}\right) + \log 2$$

$$\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{qi}\right) < \sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2q_i}\right) + \sum_{i=1}^{n} p_i \quad (\log 2)$$

$$\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{qi}\right) + \sum_{i=1}^{n} q_i \log\left(\frac{p_i}{q_i}\right) < \sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i + q_i}{2q_i}\right) - \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right) + (\log 2)$$

$$2Y_1^*(P,Q) - K(Q,P) < J_R(P,Q) - F(Q,P) + \log 2$$

Hence the inequality

Proposition 2.5.5: let $P,Q \in \Gamma_n$ and then we have the following new equality

$$Y_1^*(P,Q) = 2\,[I(P,Q) + T(P,Q)] \tag{2.5.6}$$

Where $Y_1^*(P,Q)$ new divergence measure, *T(P,Q)* AG- mean divergence and *I(P,Q)* JS- divergence measure, are given by (2.3.3), (.1.2.15), and (.1.2.16) respectively

Proof: From (2.3.3), we have

$$Y_1^*(P,Q) = \frac{1}{2}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i}{q_i}\right)$$

$$= \frac{1}{2}\left[\sum_{i=1}^{n}(p_i - q_i)\left\{\log\left(\frac{2p_i}{p_i + q_i}\right) + \log\left(\frac{p_i + q_i}{2q_i}\right)\right\}\right]$$

$$= \frac{1}{2}\left[\sum_{i=1}^{n} p_i \log\left(\frac{2p_i}{p_i + q_i}\right) + \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right)\right]$$

$$+ \frac{1}{2}\left[\sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2q_i}\right) + \sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2p_i}\right) - \sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2p_i}\right)\right]$$

$$+ \frac{1}{2}\left[\sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2p_i}\right) + \sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2q_i}\right) - \sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2q_i}\right)\right]$$

$$= \left[\sum_{i=1}^{n} p_i \log\left(\frac{2p_i}{p_i + q_i}\right) + \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i + q_i}\right)\right]$$

$$+\frac{1}{2}\left[\sum_{i=1}^{n} p_i \log\left(\frac{p_i+q_i}{2\sqrt{p_i q_i}}\right)^2\right]+\frac{1}{2}\left[\sum_{i=1}^{n} q_i \log\left(\frac{p_i+q_i}{2\sqrt{p_i q_i}}\right)^2\right]$$

$$=\left[\sum_{i=1}^{n} p_i \log\left(\frac{2p_i}{p_i+q_i}\right)+\sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i+q_i}\right)\right]+2\left[\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2\sqrt{p_i q_i}}\right)\right]$$

$$Y_1^*(P,Q)=2\left[I(P,Q)+T(P,Q)\right]$$

Hence the equality

3

Series of Convex Functions and Analogous Information Divergence Measures

3.1 Introduction

In chapter 2, we have calculated some new classes of divergence measures, which belong to the family of Csiszár’s *f* – divergence. In this chapter, we have studied different series of some new classes of information divergence measures, which again; belong to the family of Csiszár’s *f*-divergence.

This chapter is organized as follows: After this introduction section 3.1, we introduce the new series of convex functions (3.2.1) in section 3.2 and get corresponding series of the new information divergence measures (3.2.4) of Csiszár’s class. In section 3.3, the new series of exponential convex functions is extracted from previous series (3.2.1) of the new convex functions and using, the Csiszár’s generalized *f*- divergence concept; we get the new series of exponential divergence measures. In section 3.4, some new convex functions have been extracted, from the convex function (3.2.1), using the property that, sum of convex functions is also a convex function. In this section 3.4, some non-parametric information divergence measures (3.4.5, 3.4.6, 3.4.7, 3.4.8), corresponding to convex functions (3.4.1, 3.4.2, 3.4.3, 3.4.4), which belongs to the family of Csiszár's f – divergence also derived. In section 3.5, some inequalities & equalities among new divergence measures and chi-square divergence, triangular discrimination, Jain and Saraswat divergence, Hellinger discrimination, Variational distance , harmonic mean divergence, are present in results (3.5.4, 3.5.5, 3.5.6, 3.5.7, 3.5.8, 3.5.9, 3.5.10, 3.5.11) respectively.

3.2 Series of New Convex Functions and Information Measures

In this section we shall find out the series of new information divergence measures with the help of the following new series of convex functions.

Let us consider the function $f:(0,\infty)\to R$

Such that
$$f_k(t)=\frac{(t-1)^{2k+2}}{t^{2k}} \quad , \quad k = 1, 2, 3, \ldots \tag{3.2.1}$$

Then
$$f_k'(t)=\frac{2(t-1)^{2k+1}(t+k)}{t^{2k+1}} \tag{3.2.2}$$

And
$$f_k''(t)=\frac{2(t-1)^{2k}\left[(t+k)^2+k^2+k\right]}{t^{2k+2}} \tag{3.2.3}$$

Since $f_k''(t)\geq 0\ \forall t>0$ and k = 1, 2, 3… Therefore $f_k(t)$ are convex functions for each k, and normalized also, since $f_k(1)=0$. Now for convex functions (3.2.1), we get the following new series of divergences

$$C_f(P,Q)=J_k^c(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^{2k+2}}{p_i^{2k}q_i} \quad ; \ k = 1, 2, 3, \ldots, \tag{3.2.4}$$

$$J_1^c(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^4}{p_i^2 q_i} \tag{3.2.5}$$

$$J_2^c(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^6}{p_i^4 q_i} \tag{3.2.6}$$

$$J_3^c(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^8}{p_i^6 q_i} \tag{3.2.7}$$

Where $C_f(P,Q)$ is well known Csiszár's [24, 25]generalized divergence measure (1.2.2). Divergence measure J_k^c(P, Q) is non-symetric divergence measure since

$$J_k^c(P,Q)\neq J_k^c(Q,P) \tag{3.2.8}$$

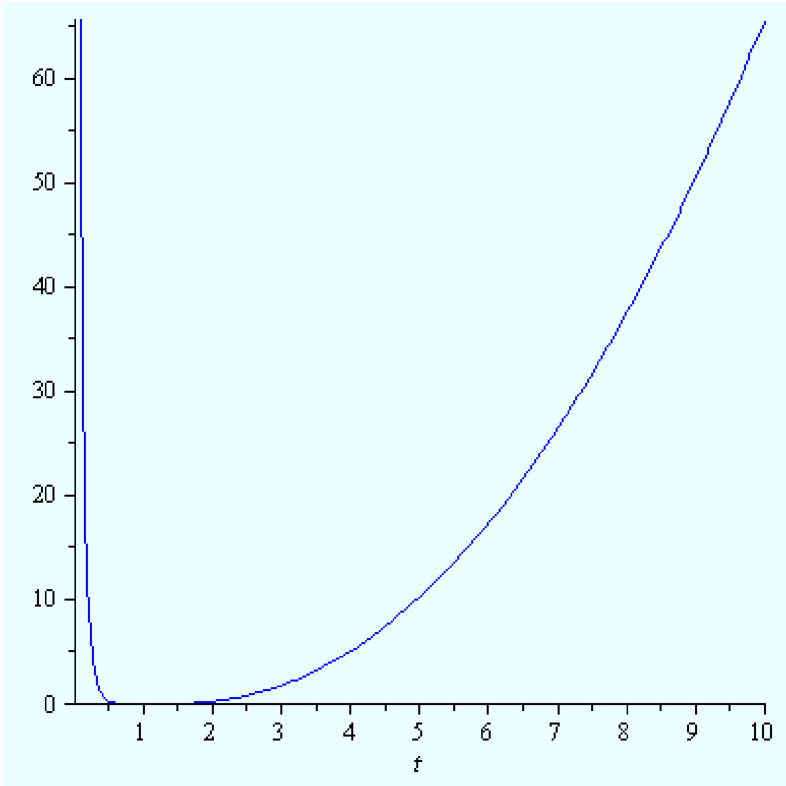

FIGURE 3.1: Graph of the convex function f_k (t) for k =1

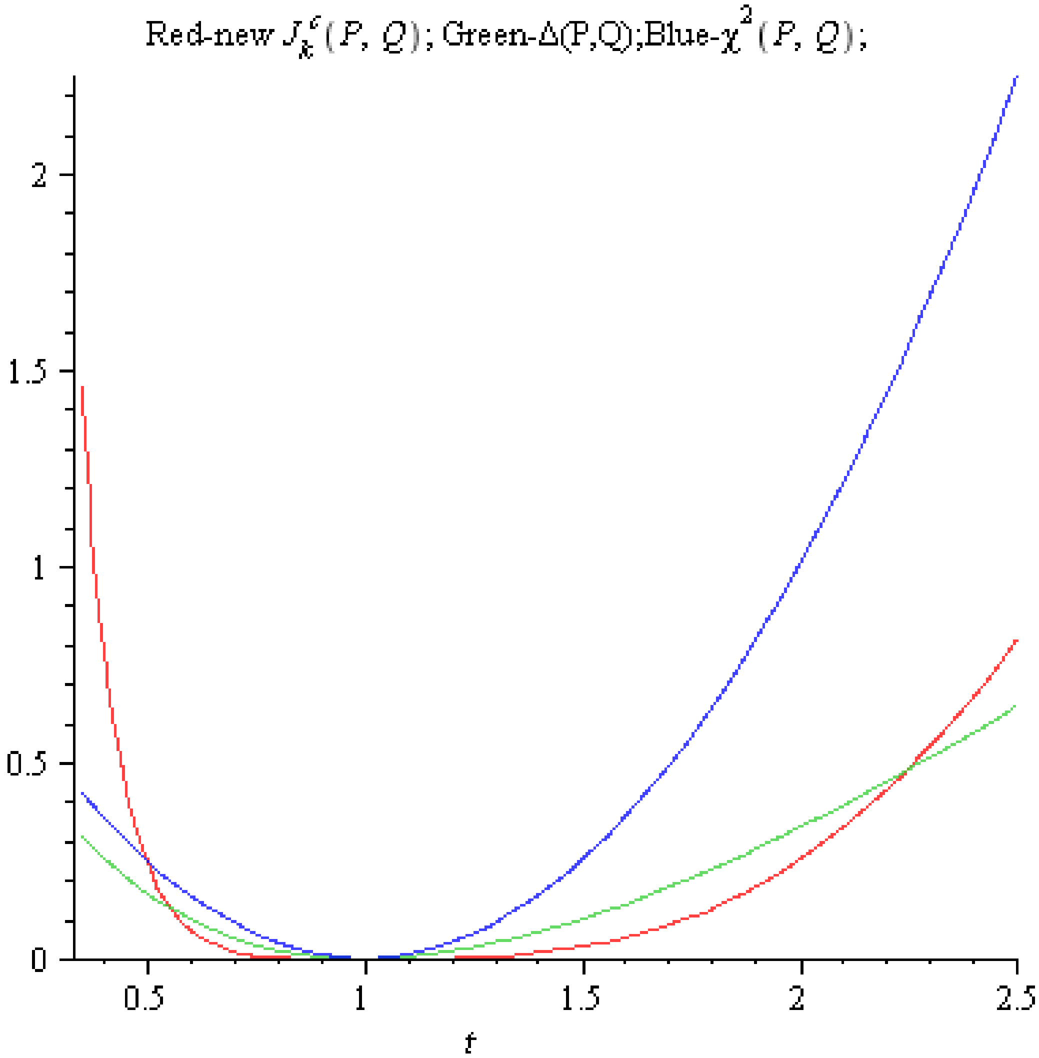

FIGURE 3.2: Comparison of new divergence measure $J_k^c(P,Q)$ with some well-known measures

3.3 Extraction of New Series of Convex Functions and Information Measures

Now, from equation (3.2.1) for k = 1, 2, 3, 4 ... we get the following convex functions

$$f_1(t)=\frac{(t-1)^4}{t^2},\ f_2(t)=\frac{(t-1)^6}{t^4},\ f_3(t)=\frac{(t-1)^8}{t^6},\ f_4(t)=\frac{(t-1)^{10}}{t^8},...$$

We know that, sum of convex functions is also a convex function

i.e. $$c_1 f_1(t)+c_2 f_2(t)+c_3 f_3(t)+c_4 f_4(t)+...$$

is also a convex function, where $c_1, c_2, c_3, c_4,...$ are arbitrary positive constants and at least one

c_i (*i=1, 2, 3…*) is not equal to zero.

Now,

$$F_k(t)=\sum_{i=1}^{n} c_i f_i(t)=c_1 f_1(t)+c_2 f_2(t)+c_3 f_3(t)+c_4 f_4(t)+...$$

$$\Rightarrow F_k(t)=c_1\frac{(t-1)^4}{t^2}+c_2\frac{(t-1)^6}{t^4}+c_3\frac{(t-1)^8}{t^6}+c_4\frac{(t-1)^{10}}{t^8}+... \tag{3.3.1}$$

Taking $c_1=1,\ c_2=\frac{1}{1!},\ c_3=\frac{1}{2!},\ c_4=\frac{1}{3!},...$

Then

$$F_1(t)=1\,\frac{(t-1)^4}{t^2}+\frac{1}{1!}\frac{(t-1)^6}{t^4}+\frac{1}{2!}\frac{(t-1)^8}{t^6}+\frac{1}{3!}\frac{(t-1)^{10}}{t^8}+...$$

$$=\frac{(t-1)^4}{t^2}\left[1+\frac{1}{1!}\left\{\left(\frac{t-1}{t}\right)^2\right\}+\frac{1}{2!}\left\{\left(\frac{t-1}{t}\right)^2\right\}^2+\frac{1}{3!}\left\{\left(\frac{t-1}{t}\right)^2\right\}^3+...\right]$$

$$\therefore\ F_1(t) = \frac{(t-1)^4}{t^2}\exp\left\{\left(\frac{t-1}{t}\right)^2\right\} \tag{3.3.2}$$

Now, for (3.3.2), Divergence measure of Csiszár's f-divergence class

$$J_1^*(P,Q) = \sum_{i=1}^{n}\frac{(p_i-q_i)^4}{p_i^2 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.3}$$

Next, taking $c_1 = 0,\ c_2 = 1,\ c_3 = \frac{1}{1!},\ c_4 = \frac{1}{2!},\ldots$

We get

$$F_2(t) = \frac{(t-1)^6}{t^4}\exp\left\{\left(\frac{t-1}{t}\right)^2\right\} \tag{3.3.4}$$

Therefore, corresponding information divergence measure of Csiszár's f – divergence class

$$J_2^*(P,Q) = \sum_{i=1}^{n}\frac{(p_i-q_i)^6}{p_i^4 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.5}$$

Similarly, by appropriate selection of constants, we get the following convex functions

$$F_k(t) = \frac{(t-1)^{2k+2}}{t^{2k}}\exp\left\{\left(\frac{t-1}{t}\right)^2\right\};\ \text{k = 1, 2, 3 …} \tag{3.3.6}$$

Hence corresponding series of information measures of Csiszár's f – divergence class

$$J_k^*(P,Q) = \sum_{i=1}^{n}\frac{(p_i-q_i)^{2k+2}}{p_i^{2k} q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.7}$$

Further $F_k(1) = 0$, so that $J_k^*(P,P) = 0$ and the convexity of the function $F_k(t)$ ensure that the measure $J_k^*(P,Q)$ is non-negative.

Thus we can say that the measure (3.3.7) is non-negative and non-symmetric in the pair of probability distributions $(P,Q)\in\Gamma_n$, since

$$J_1^*(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^4}{p_i^2 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.8}$$

$$J_2^*(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^6}{p_i^4 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.9}$$

$$J_3^*(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^8}{p_i^6 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.10}$$

$$J_4^*(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^{10}}{p_i^8 q_i}\exp\left\{\left(\frac{p_i-q_i}{p_i}\right)^2\right\} \tag{3.3.11}$$

And so on.

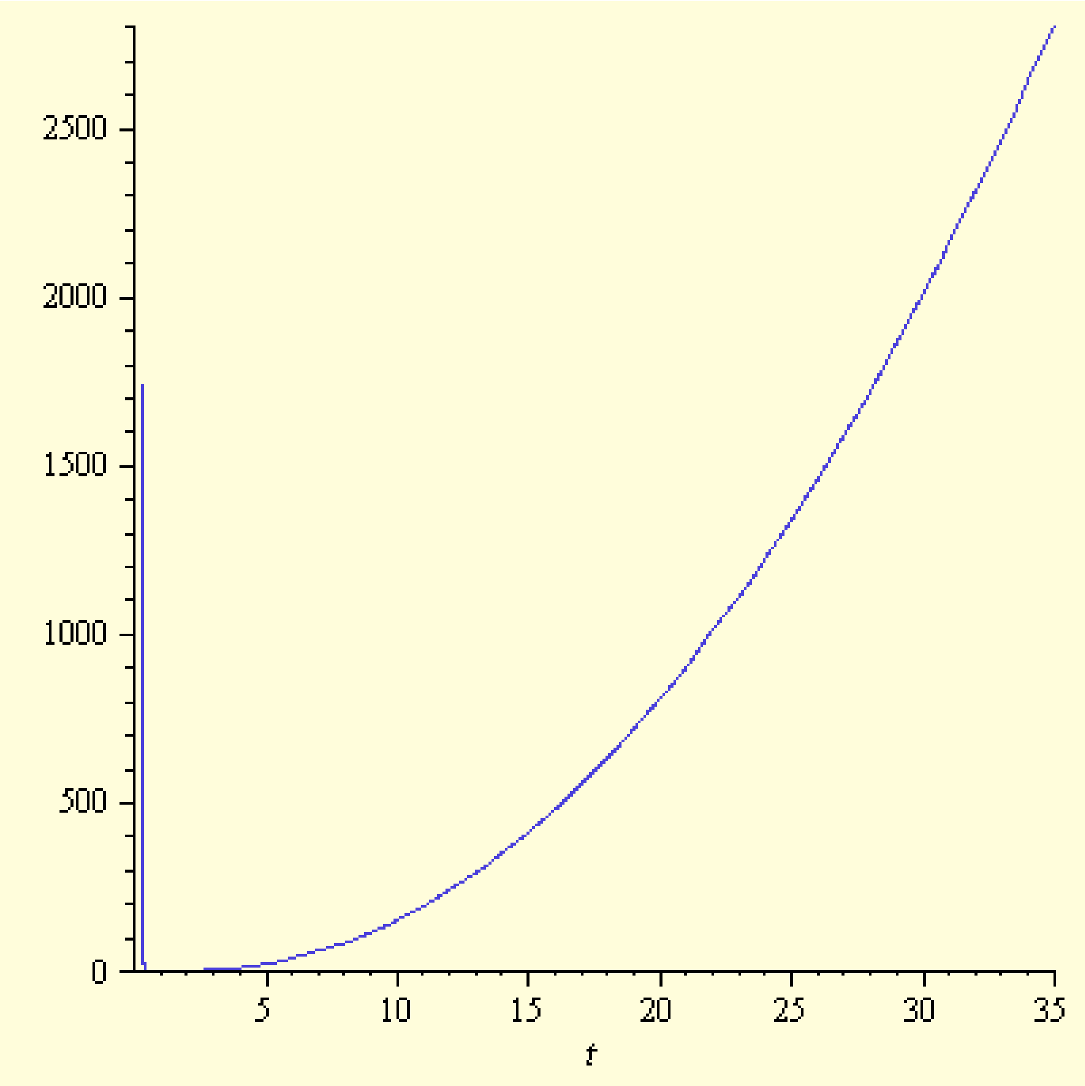

FIGURE 3.3: Graph of the convex function $F_k(t)$ corresponding to $J_k^*(P,Q)$ for k = 1

3.4 Other New Series of Information Measures

In this section, we will derive series of new information divergence measures, using the series of convex functions. We know that, the sum of convex functions is again a convex function; therefore we have the following series of convex functions

$$\frac{(t-1)^4}{t^2}+\frac{(t-1)^6}{t^4}=\frac{(t-1)^4\left(2t^2-2t+1\right)}{t^4} \tag{3.4.1}$$

$$\frac{(t-1)^6}{t^4}+\frac{(t-1)^8}{t^6}=\frac{(t-1)^6\left(2t^2-2t+1\right)}{t^6} \tag{3.4.2}$$

$$\frac{(t-1)^8}{t^6}+\frac{(t-1)^{10}}{t^8}=\frac{(t-1)^8\left(2t^2-2t+1\right)}{t^8} \tag{3.4.3}$$

$$\frac{(t-1)^{10}}{t^8}+\frac{(t-1)^{12}}{t^{10}}=\frac{(t-1)^{10}\left(2t^2-2t+1\right)}{t^{10}} \tag{3.4.4}$$

And so on

Hence, the new series of information divergence measures of Csiszár's f – divergence class corresponding to the above new series of convex functions, given as

$$J_1(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^4\left(2p_i^2-2p_iq_i+q_i^2\right)}{p_i^4q_i} \tag{3.4.5}$$

$$J_2(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^6\left(2p_i^2-2p_iq_i+q_i^2\right)}{p_i^6q_i} \tag{3.4.6}$$

$$J_3(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^8\left(2p_i^2-2p_iq_i+q_i^2\right)}{p_i^8q_i} \tag{3.4.7}$$

$$J_4(P,Q)=\sum_{i=1}^{n}\frac{(p_i-q_i)^{10}\left(2p_i^2-2p_iq_i+q_i^2\right)}{p_i^{10}q_i} \tag{3.4.8}$$

And so on.

Similarly, we can generate various other series of information divergence measures, using the properties of convex functions.

3.5 Some Bounds for the New Series of Information Measures

In this section, we will derive some bounds for new information divergence measures, which are derived earlier in (3.2.4), (3.3.3), (3.3.5), (3.3.7), (3.4.5), (3.4.6), (3.4.7), and (3.4.8) , with some familiar divergence measures

$\Delta(P,Q)$: Triangular discrimination (1.2.10);

$V(P,Q)$: Variational Distance or l_1 distance (1.2.4);

$h(P,Q)$: Hellinger Discrimination (1.2.8);

$H(P,Q)$: Harmonic mean divergence (1.3.22);

Jain and Saraswat divergence [61]

$$R_k^*(P,Q) = \sum \frac{(p-q)^{k+1}}{p^k} \exp\left\{\frac{(p-q)^2}{p^2}\right\}, \text{ k = 1, 3, …} \tag{3.5.1}$$

For k = 1,

$$R_1^*(P,Q) = \sum \frac{(p-q)^2}{p} \exp\left\{\frac{(p-q)^2}{p^2}\right\} \tag{3.5.2}$$

$\chi^2(P,Q)$: Chi-square divergence or Pearson divergence (1.2.5)

Other form of Chi-square divergence

$$\chi_1^2(Q,P) = \sum \frac{(p-q)^2}{p}, \quad \chi_1^4(Q,P) = \sum \frac{(p-q)^4}{p^3}, \tag{3.5.3}$$

Now, from (3.2.4) for k = 1, we have

$$J_1^c(P,Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^4}{p_i^2 \; q_i}$$

$$J_1^c(P,Q) = 2\left(\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{(p_i + q_i)}\right)\left(\sum_{i=1}^{n} \frac{p_i + q_i}{2 \; p_i \; q_i}\right)\left(\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i}\right)$$

Using (1.2.10), (1.2.5), and (1.3.22), we get

$$J_1^c(P,Q) = 2\left[\frac{\Delta(P,Q) \;\; \chi^2(Q,P)}{H(P,Q)}\right] \tag{3.5.4}$$

From (1.2.8) and (1.2.10), we know that [59],

$$2 \, h(P,Q) \le \Delta(P,Q)$$

So, from (3.5.4), we get

$$J_1^c(P,Q) \ge 4\left[\frac{h(P,Q) \;\; \chi^2(Q,P)}{H(P,Q)}\right] \tag{3.5.5}$$

And, from (1.2.4) and (1.2.10) [54], we know that,

$$V^2(P,Q) \le 2\,\Delta\,(P,Q)$$

Then, from (3.5.4), we get

$$J_1^c(P,Q) \ge \left[\frac{V^2(P,Q) \;\; \chi^2(Q,P)}{H(P,Q)}\right] \tag{3.5.6}$$

Next, from (3.3.3), the outcome is

$$J_1^*(P,Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^4}{p_i^2 q_i} \exp\left\{\left(\frac{p_i - q_i}{p_i}\right)^2\right\}$$

$$= \sum_{i=1}^{n} \left[\frac{(p_i - q_i)^2}{(p_i + q_i)} \frac{p_i + q_i}{p_i \; q_i} \frac{(p_i - q_i)^2}{p_i} \exp\left\{ \left(\frac{p_i - q_i}{p_i} \right)^2 \right\} \right]$$

$$= \left(\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{(p_i + q_i)} \right) \; 2 \left(\sum_{i=1}^{n} \frac{(p_i + q_i)}{2 \; p_i \; q_i} \right) \left[\sum_{i=1}^{n} \frac{(p_i - q_i)^2}{p_i} \exp\left\{ \left(\frac{p_i - q_i}{p_i} \right)^2 \right\} \right]$$

Now, using (1.2.10), (3.5.2) and (1.3.22), we get

$$J_1^*(P,Q) = 2 \left[\frac{\Delta(P,Q) \;\; R_1^*(P,Q)}{H(P,Q)} \right] \tag{3.5.7}$$

And we know that [54],

$$V^2(P,Q) \le 2 \Delta \, (P,Q)$$

Then, from (3.5.7), we get

$$J_1^*(P,Q) \ge \left[\frac{V^2(P,Q) \;\; R_1^*(Q,P)}{H(P,Q)} \right] \tag{3.5.8}$$

Next,

$$J_1^*(P,Q) \le 2 \left[\frac{\chi^2(P,Q) \;\; R_1^*(P,Q)}{H(P,Q)} \right] \tag{3.5.9}$$

by [58] we have

$$\Delta(P,Q) \le \chi^2(Q,P)$$

Further, we know that [59], $2 \, h(P,Q) \le \Delta(P,Q)$

So, from (3.5.7), we get

$$J_1^*(P,Q) \ge 4 \left[\frac{h(P,Q) \;\; R_1^*(Q,P)}{H(P,Q)} \right] \tag{3.5.10}$$

And, from (1.2.10) & (3.5.3),

$$\Delta(P,Q) \le \left[\chi_1^2(Q,P) + \chi_1^4(Q,P)\right]$$

Therefore, from (3.5.7), we get

$$J_1^*(P,Q) \le \left[\frac{\left\{\chi_1^2(Q,P) + \chi_1^4(Q,P)\right\} R_1^*(Q,P)}{H(P,Q)}\right] \tag{3.5.11}$$

4

New Series of *f*-Divergence Measures and Their Bounds

4.1 Introduction

In this chapter, we will introduce two new series of information divergence measures using Jain and Saraswat generalized *f*- divergence measure to obtain various new information inequalities on these new series of divergence measures with some well known information measures.

The complete chapter is divided in sections: In section 4.2, we give some well-known inequalities, which are established in literature of pure and applied mathematics. Using these inequalities, we have derived important bounds of divergence measures. In section 4.3, we have introduced, Jain and Saraswat[57] divergence measure and it's properties. In section 4.4, we shall find out two new series of information divergence measures (4.4.4, 4.4.11) of Jain and Saraswat generalized *f*- divergence class. Finally, new information inequalities and equalities, in terms of new information divergence measures and other well known divergence measures (e.g. Chi-Square divergence, Relative J- divergence measure, Triangular discrimination, Arithmetic mean divergence, Relative JS- divergence, Relative AG-divergence, AG- mean divergence, JS- divergence measure, Relative Information) are obtained in section 4.5.

4.2 Well known inequalities

In this section we give some well-known inequalities which are established in literature of pure and applied mathematics. These are very useful to derive some bounds of well-known information divergence measure in literature of information theory and statistics. Using following inequalities we have derived important bounds of well-known divergence measures

$$\frac{x}{1+x} \le \log(1+x) \le x \; ; x > 0 \tag{4.2.1}$$

$$x - \frac{x^2}{2} \le \log(1+x) \le x - \frac{x^2}{2(1+x)} \; ; x > 0 \tag{4.2.2}$$

4.3 New f-Divergence Measure and Properties

A new f-divergence measure introduced by Jain & Saraswat [57], which is given by

$$S_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i + q_i}{2q_i}\right) \tag{4.3.1}$$

where $f:(0,\infty) \to R$ (set of real numbers) is a convex function and $P = (p_1, p_2, p_3, \ldots p_n), Q = (q_1, q_2, q_3, \ldots q_n) \in \Gamma_n$, where p_i *and* q_i are discreate probability mass function.

An important property of this divergence is that many known divergences can be obtained from this measure by appropriately defining the convex function *f*.

The following results are presented by Jain & Saraswat [57]

Proposition 4.3.1: Let *f*: [0,∞)→R be the convex function $P,Q \in \Gamma_n$. Then we have the following inequality

$$S_f(P,Q) \ge f(1) \tag{4.3.2}$$

If ***f*** is normalized i.e. $f(1) = 0$, then $S_f(P,Q) \geq 0$ and if ***f*** is strictly convex and equality holds iff

$$p_i = q_i \;\forall\; i=1, 2, 3\ldots n$$

$$\text{i.e. } S_f(P,Q) \geq 0 \quad and \;\; S_f(P,Q) = 0 \; if \quad P = Q \tag{4.3.3}$$

Proposition 4.3.2: If f_1 and f_2 are two convex functions and $F = c_1 f_1 + c_2 f_2$ then

$$S_F(P,Q) = c_1 S_{f_1}(P,Q) + c_2 S_{f_2}(P,Q) \tag{4.3.4}$$

Where c_1 and c_2 are constants and $P,Q \in \Gamma_n$

4.4 Divergence Measure of New f-Divergence Measure's Class

In this section, we will find out the new series of information divergence measures with the help of following series of the convex functions.

Let us consider the function $f:(0,\infty) \to R$, defined as

$$f_k(t) = \frac{(t-1)^{k+1}}{(t+1)^k} \qquad : k = 1,3,5,\cdots \tag{4.4.1}$$

$$\text{The } \; f_k'(t) = \frac{(t-1)^k(2k+t+1)}{(t+1)^{k+1}} \tag{4.4.2}$$

$$\text{An} \quad f_k''(t) = \frac{(t-1)^{k-1}4k(k+1)}{(t+1)^{k+2}} \tag{4.4.3}$$

Since $f_k''(t) \geq 0 \;\forall\; t > 0$;k=1,3,5,… therefore the function $f_k(t)$ is convex and normalized also since f(1) = 0.

Figure 4.1, shows the behavior of the function $f_k(t)$ is always convex if k =1, 3, 5,… $\forall\; t > 0$.

Now putting function [4.4.1] in [4.3.1], we obtain,

$$S_f(P,Q)=M_k^c(P,Q)=\frac{1}{2}\sum_{i=1}^{n}\frac{(p_i-q_i)^{k+1}}{(p_i+3q_i)^k} \qquad ;k=1,3,5,\cdots \tag{4.4.4}$$

Which is the new series of information divergence measure for k = 1, 3, 5, … Divergence measure $M_k^c(P,Q)$ is non-symmetric divergence measure, since

$$M_k^c(P,Q)\neq M_k^c(Q,P)$$

Moreover

$$M_k^c(P,Q)\geq 0 \quad \forall\, P,Q\in\Gamma_n \ \ and \ \ M_k^c(P,Q)=0 \ \ iff \ \ P=Q$$

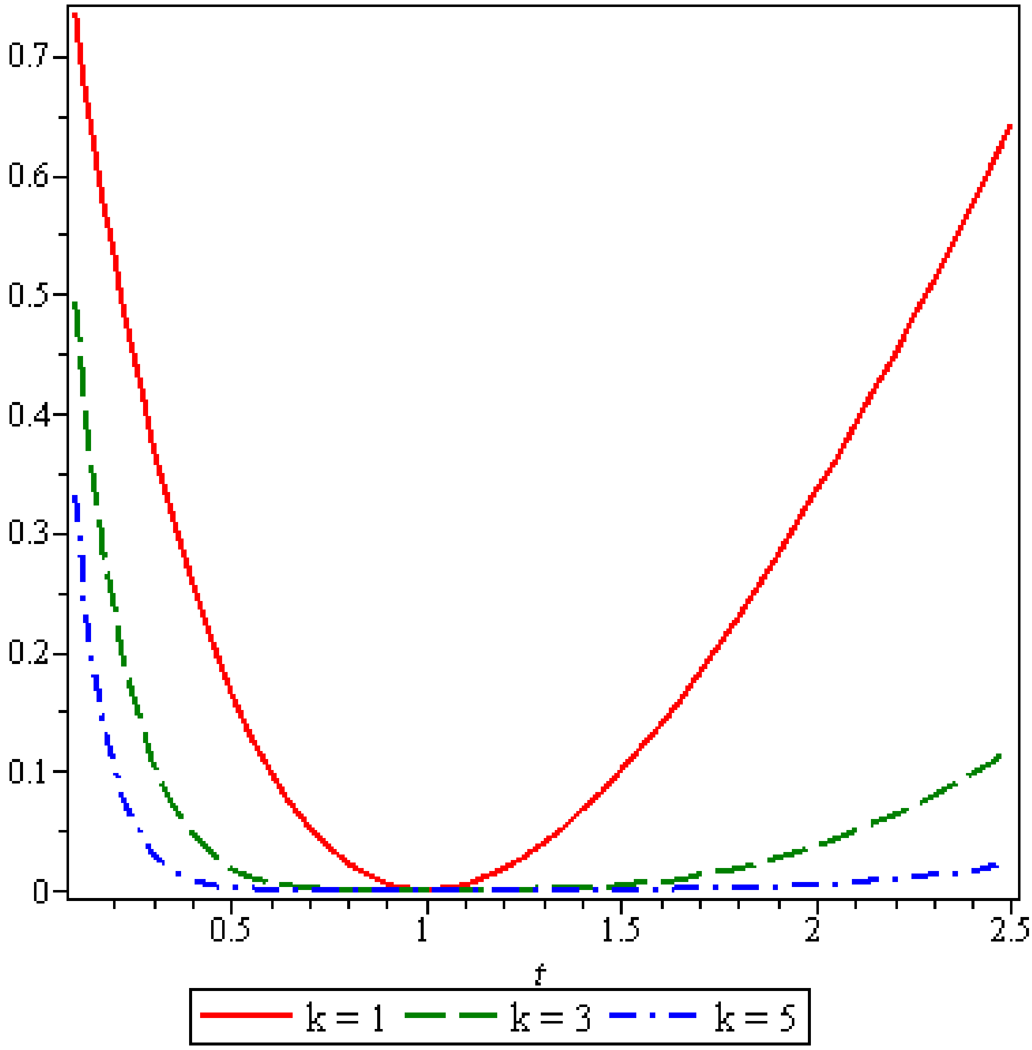

FIGURE 4.1: Behavior of the convex function $f_k(t)$

Now, from equation (4.4.1) for k= 1, 3, 5…we get the following convex functions

$$f_1 = \frac{(t-1)^2}{(t+1)},\; f_3 = \frac{(t-1)^4}{(t+1)^3},\; f_5 = \frac{(t-1)^6}{(t+1)^5},\; f_7 = \frac{(t-1)^8}{(t+1)^7}\ldots$$

We know that, sum of convex functions is also a convex function

$c_1 f_1(t) + c_3 f_3(t) + c_5 f_5(t) + c_7 f_7 + \ldots$ is also a convex function. Where $c_1, c_3, c_{5,} c_7 \ldots$ are arbitrary positive constants and at least one c_k (k = 1, 3, 5, 7 …) is not equal to zero.

Now

$$F_k(t) = c_1 f_1(t) + c_3 f_3(t) + c_5 f_5(t) + c_7 f_7(t) + \ldots$$

$$= c_1 \frac{(t-1)^2}{(t+1)^1} + c_3 \frac{(t-1)^4}{(t+1)^3} + c_5 \frac{(t-1)^6}{(t+1)^5} + c_7 \frac{(t-1)^8}{(t+1)^7} + \ldots \qquad (4.4.5)$$

Now taking, $c_1 = 1, c_3 = 1/3,\; c_5 = 1/5, c_7 = 1/7 \quad \ldots$

Therefore

$$F_1(t) = \frac{(t-1)^2}{(t+1)^1} + \frac{1}{3}\frac{(t-1)^4}{(t+1)^3} + \frac{1}{5}\frac{(t-1)^6}{(t+1)^5} + \frac{1}{7}\frac{(t-1)^8}{(t+1)^7} + \ldots$$

$$= (t-1)\left[\left(\frac{t-1}{t+1}\right)^1 + \frac{1}{3}\left(\frac{t-1}{t+1}\right)^3 + \frac{1}{5}\left(\frac{t-1}{t+1}\right)^5 + \frac{1}{7}\left(\frac{t-1}{t+1}\right)^7 + \ldots\right]$$

$$= (t-1)\frac{1}{2}\log\left\{\frac{1+\left(\frac{t-1}{t+1}\right)}{1-\left(\frac{t-1}{t+1}\right)}\right\}$$

$$F_1(t) = \frac{1}{2}(t-1)\log(t) \qquad (4.4.6)$$

From (4.3.1), Divergence measure of *f*- divergence class for (4.4.6),

$$M_1^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i + q_i}{2q_i}\right) \tag{4.4.7}$$

Next, taking $c_1 = 0, c_3 = 1, c_5 = 1/3, c_7 = 1/5$ …, then from equation (4.4.5),

We get

$$F_3(t) = \frac{1}{2}\frac{(t-1)^3}{(t+1)^2}\log(t) \tag{4.4.8}$$

Figure 4.2, shows the convex behavior of the new convex function $F_k(t)$ for k = 3. Hence, Divergence measure, for (4.4.8) from (4.3.1),

$$M_3^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i - q_i)^3}{(p_i + 3q_i)^2}\log\left(\frac{p_i + q_i}{2q_i}\right) \tag{4.4.9}$$

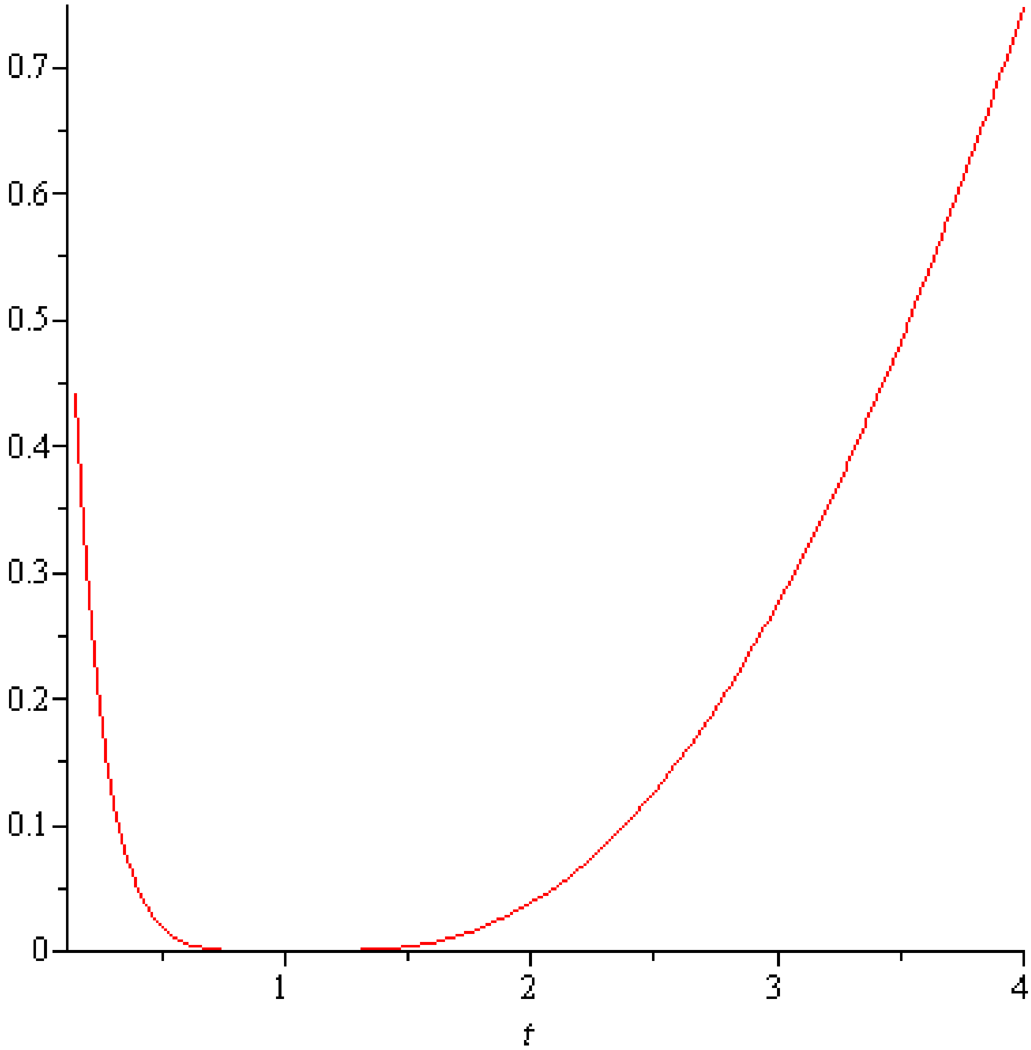

FIGURE 4.2: Graph of the convex function $F_k(t)$ corresponding to $M_k^*(P,Q)$ for k = 1

Similarly, by appropriate selection of constants, we get the following convex function

$$F_k(t) = \frac{1}{2}\frac{(t-1)^k}{(t+1)^{k-1}}\log(t) \;\; ; k = 1,3,5... \tag{4.4.10}$$

And the corresponding series of divergence measures of f-divergence class

$$M_k^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i - q_i)^k}{(p_i + 3\,q_i)^{k-1}}\log\left(\frac{p_i + q_i}{2\,q_i}\right); \; k = 1,3,5... \tag{4.4.11}$$

It may be noted that, $F_k(t)$ in (4.4.10) satisfies $F_k(1) = 0$, so that $M_k^*(P,P) = 0$, Convexity of $F_k(t)$ make sure that divergence measure $M_k^*(P,Q)$ is non-negative. Thus, we have

$(a) M_k^*(P,Q) \geq 0 \;\; \forall\, P,Q \in \Gamma_n \; and \; M_k^*(P,Q) = 0 \; iff \;\; P = Q$

$(b)\, M_k^*(P,Q)$ is non-symmetric with respect to probability distributions.

4.5 New Information Inequalities and Equalities

We now derive information divergence inequalities and equalities providing bounds for the new series of information divergences $M_k^*(P, Q)$ in terms of the well-known divergence measures in the following propositions

Proposition 4.5.1: Let (P, Q)$\in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$M_1^*(P, Q) \leq [F\,(Q, P) + G\,(Q, P)] \tag{4.5.1}$$

Where M_1^*(P, Q) new divergence measure, *F(Q, P)* Relative J- divergence measure and *G (P, Q)* Relative Arithmetic- Geometric divergence are given by (4.4.7), (1.2.6), and (1.2.7) respectively.

Proof. From (4.4.7), we have

$$M_1^*(P, Q) = \frac{1}{4}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$= \frac{1}{4}\sum_{i=1}^{n}(p_i + q_i - 2q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left\{\left(\frac{p_i+q_i}{2}\right) - q_i\right\}\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$= \frac{1}{2}\left[\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2q_i}\right) - \sum_{i=1}^{n} q_i\log\left(\frac{p_i+q_i}{2q_i}\right)\right]$$

$$= \frac{1}{2}\left[\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2q_i}\right) + \sum_{i=1}^{n} q_i\log\left(\frac{2q_i}{p_i+q_i}\right)\right]$$

$$M_1^*(P, Q) = \frac{1}{2}[G(Q,P) + F(Q,P)]$$

$$2M_1^*(P, Q) = [F(Q,P) + G(Q,P)]$$

$$M_1^*(P, Q) \le [F(Q,P) + G(Q,P)]$$

Hence the inequality is proved.

Proposition 4.5.2: Let $(P, Q) \in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$M_3^*(P, Q) \le \frac{\Delta(P,Q) J_R(P,Q)}{2\, A(P,Q)} \tag{4.5.2}$$

Where $M_3^*(P,Q)$ new divergence, $\Delta(P,Q)$ Triangular discrimination, $J_R(P,Q)$ Relative J- divergence measure & A(P,Q) Arithmetic mean divergence are given by (4.4.9), (1.2.10), (1.2.14) and (1.3.20) respectively.

Proof. We have

$$M_3^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+3q_i)^2}\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$\Rightarrow M_3^*(P,Q) \quad \leq \sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i+q_i}{2q_i}\right)$$

i.e.$M_3^*(P,Q)$

$$\leq \sum_{i=1}^{n}\frac{(p_i-q_i)^2}{(p_i+q_i)}\sum_{i=1}^{n}\frac{1}{(p_i+q_i)}\sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$= \Delta(P,Q)\frac{1}{2A(P,Q)}J_R(P,Q)$$

$$M_3^*(P,Q) \leq \frac{\Delta(P,Q)J_R(P,Q)}{2\,A(P,Q)}$$

Hence the inequality is proved.

Proposition 4.5.3: Let (P, Q)$\in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$4\,M_1^*(P,Q) = J_R(P,Q) = 2[F(Q,P) + G(Q,P)] \tag{4.5.3}$$

$$M_3^*(P,Q) \leq \frac{1}{16}\chi^2(P,Q)\left[\frac{1}{A(P,Q)}\right]^2 J_R(P,Q) \tag{4.5.4}$$

And $$M_3^*(P,Q) \leq \frac{1}{8}\chi^2(P,Q)\left[\frac{1}{A(P,Q)}\right]^2 [F(Q,P) + G(Q,P)] \tag{4.5.5}$$

Where M_1^*(P,Q) ,M_3^*(P,Q) new divergence measures, *F(Q,P)* Relative JS-divergence, *G(P,Q)* Relative AG- divergence, *A(P,Q)* Arithmetic mean divergence,

$J_R(P,Q)$ Relative J- divergence measure and $\chi^2(P,Q)$ Chi- Square divergence are given(4.4.7), (4.4.9), (1.2.6), (1.2.7), (1.3.20), (1.2.14) and (1.2.5) respectively.

Proof. From (25), we have

$$M_1^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}(p_i - q_i)\, log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$\Rightarrow \qquad 4\, M_1^*(P,Q) = \sum_{i=1}^{n}(p_i - q_i)\, log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$\Rightarrow \qquad 4M_1^*(P,Q) = J_R(P,Q)$$

And we have

$$J_R(P,Q) = 2[F(Q,P) + G(Q,P)]$$

$$\therefore 4M_1^*(P,Q) = J_R(P,Q) = 2[F(Q,P) + G(Q,P)]$$

Hence the equality is proved.

Next, from (4.4.9), we have

$$M_3^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+3q_i)^2}\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$\leq \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^3}{(p_i+q_i)^2}\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$= \frac{1}{4}\sum_{i=1}^{n}\frac{(p_i-q_i)^2}{q_i}\sum_{i=1}^{n}\left(\frac{1}{p_i+q_i}\right)^2\left(\sum_{i=1}^{n}q_i\right)\sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$\therefore M_3^*(P,Q) \leq \frac{1}{16}\chi^2(P,Q)\left[\frac{1}{A(P,Q)}\right]^2 J_R(P,Q)$$

Hence the inequality is proved.

We have

$$J_R(P, Q) = 2[F(Q, P) + G(Q, P)]$$

So,

$$M_3^*(\mathrm{P, Q}) \leq \frac{1}{8}\chi^2(\mathrm{P, Q}) \left[\frac{1}{A(P,Q)}\right]^2 \ [F(Q,P) + G(Q,P)]$$

Hence the inequality is proved.

Proposition 4.5.4: Let (P, Q)$\in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$H(P, Q) + 4\, M_1^*(P,Q) - 2\, G(Q, P) \leq 2\log 2 \leq 2 + 4\, M_1^*(P,Q) - 2\, G(Q, P) \quad (4.5.6)$$

Where M_1^*(P, Q) new divergence measure, G (Q, P) Relative Arithmetic-Geometric divergence and H (P, Q) Harmonic mean divergence are given by (4.4.7), (1.2.7) and (1.3.22) respectively.

Proof: From (4.2.1), we have the inequality

$$\frac{x}{1+x} \leq \log(1+x) \leq x\ ; \quad x > 0$$

Taking, $x = \frac{p_i}{q_i}$, then we get

$$\frac{p_i}{p_i + q_i} \leq \log\left(\frac{p_i + q_i}{q_i}\right) \leq \frac{p_i}{q_i}$$

$$\frac{p_i}{p_i + q_i} \leq \log\left(\frac{p_i + q_i}{2q_i}.2\right) \leq \frac{p_i}{q_i}$$

$$\frac{p_i}{p_i+q_i} \leq \log\left(\frac{p_i+q_i}{2q_i}\right) + \log 2 \leq \frac{p_i}{q_i} \quad (4.5.7)$$

Multiplying by $2q_i$ and taking summation, we get

$$\sum_{i=1}^{n} \frac{2p_i q_i}{p_i + q_i} \le 2 \sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2q_i}\right) + 2\left(\sum_{i=1}^{n} q_i\right) \log 2 \le 2 \sum_{i=1}^{n} p_i$$

$$\sum_{i=1}^{n} \frac{2p_i q_i}{p_i + q_i} \le 2 \sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2q_i}\right) - \sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2q_i}\right) + \sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2q_i}\right) + 2\log 2 \le 2$$

$$\sum_{i=1}^{n} \frac{2p_i q_i}{p_i + q_i} \le 2 \sum_{i=1}^{n} \left(\frac{p_i + q_i}{2}\right) \log\left(\frac{p_i + q_i}{2q_i}\right) - \sum_{i=1}^{n} (p_i - q_i) \log\left(\frac{p_i + q_i}{2q_i}\right) + 2\log 2 \le 2$$

$H(P, Q) \le 2\,G(Q, P) - 4\,M_1^*(P,Q) + 2\log 2 \le 2$

$H(P, Q) + 4M_1^*(P, Q) - 2G(Q, P) \le 2\log 2 \le 2 + 4\,M_1^*(P, Q) - 2G(Q, P)$

Hence the result

Proposition 4.5.5: Let $(P, Q) \in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$M_1^*(P, Q) = \frac{1}{2} I(P, Q) + \frac{T(P,Q)}{2\,A(P,Q)} \tag{4.5.8}$$

And

$$M_1^*(P, Q) + M_1^*(Q, P) = I(P,Q) + T(P,Q) \tag{4.5.9}$$

Where $M_1^*(P, Q)$ new divergence measure, *I(P,Q)* Jensen- Shannon divergence measure and *T(P,Q)* Arithmetic- Geometric mean divergence are given by (4.4.7), (4.1.12), and (4.1.13) respectively.

Proof: We have

$$M_1^*(P,Q) = \frac{1}{4}\sum_{i=1}^{n}(p_i - q_i)\log\left(\frac{p_i+q_i}{2q_i}\right)$$

$$M_1^*(P,Q) = \frac{1}{4}\left[\sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2q_i}\right) - \sum_{i=1}^{n} q_i\log\left(\frac{p_i+q_i}{2q_i}\right)\right]$$

$$M_1^*(P,Q) = \frac{1}{4}\left[\sum_{i=1}^{n} q_i\log\left(\frac{2q_i}{p_i+q_i}\right) - \sum_{i=1}^{n} p_i\log\left(\frac{2q_i}{p_i+q_i}\right) + \sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2p_i}\right) - \sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2p_i}\right)\right]$$

$$M_1^*(P,Q) = \frac{1}{4}\left[\sum_{i=1}^{n} q_i\log\left(\frac{2q_i}{p_i+q_i}\right) + \sum_{i=1}^{n} p_i\log\left(\frac{2p_i}{p_i+q_i}\right) + \sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2p_i}\right) + \sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2q_i}\right)\right]$$

$$M_1^*(P,Q) = \frac{1}{4}\left[\sum_{i=1}^{n} q_i\log\left(\frac{2q_i}{p_i+q_i}\right) + \sum_{i=1}^{n} p_i\log\left(\frac{2p_i}{p_i+q_i}\right)\right] + \frac{1}{4}\sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2p_i}\frac{p_i+q_i}{2q_i}\right)$$

$$M_1^*(P,Q) = \frac{1}{4}[F(Q,P)+F(P,Q)] + \frac{1}{2}\textstyle\sum_{i=1}^{n} p_i\log\left(\frac{p_i+q_i}{2\sqrt{p_iq_i}}\right) \tag{4.5.10}$$

But $\mathrm{I}(P,Q) = \frac{1}{2}[F(P,Q)+F(Q,P)]$

$$\therefore M_1^*(P,Q) = \frac{1}{2}\mathrm{I}(P,Q) + \frac{1}{2}\sum_{i=1}^{n}(p_i)\sum_{i=1}^{n}\left(\frac{2}{p_i+q_i}\right)\sum_{i=1}^{n}\left(\frac{p_i+q_i}{2}\right)\log\left(\frac{p_i+q_i}{2\sqrt{p_iq_i}}\right)$$

$$\Rightarrow \qquad M_1^*(P,Q) = \frac{1}{2}\mathrm{I}(P,Q) + \frac{1}{2\,A(P,Q)}\,T(P,Q)$$

Hence prove the required equality.

Again from (4.5.10),

$$M_1^*(P,Q) = \frac{1}{4}[F(Q,P) + F(P,Q)] + \frac{1}{2}\sum_{i=1}^{n} p_i \log\left(\frac{p_i + q_i}{2\sqrt{p_i q_i}}\right)$$

$$M_1^*(Q,P) = \frac{1}{4}[F(Q,P) + F(P,Q)] + \frac{1}{2}\sum_{i=1}^{n} q_i \log\left(\frac{p_i + q_i}{2\sqrt{p_i q_i}}\right)$$

Hence

$$M_1^*(P,Q) + M_1^*(Q,P) = \frac{1}{2}[F(Q,P) + F(P,Q)] + \sum_{i=1}^{n}\left(\frac{p_i + q_i}{2}\right)\log\left(\frac{p_i + q_i}{2\sqrt{p_i q_i}}\right)$$

$$\mathrm{M}_1^*(\mathrm{P},\mathrm{Q}) + \mathrm{M}_1^*(\mathrm{Q},\mathrm{P}) = \mathrm{I}(\mathrm{P},\mathrm{Q}) + \mathrm{T}(\mathrm{P},\mathrm{Q})$$

Hence the equality is proved.

Proposition 4.5.6: Let (P, Q)$\in \Gamma_n \times \Gamma_n$, then we have the following new inter-relation

$$K(P,Q) + F(Q,P) - 4M_1^*(P,Q) < \log 2 \tag{4.5.11}$$

Where $M_1^*(P,Q)$, $K(P,Q)$ *and* $F(Q,P)$ are new divergence measure, Relative information and Relative Jensen- Shannon divergence respectively and given by (4.4.7), (1.2.3) and (1.2.6) respectively.

Proof: We know that,

$$0 \le p_i \le 1 \ \ and \ \ 0 \le \ q_i \ \le 1$$

$$\frac{p_i}{q_i} < 1 + \frac{p_i}{q_i}$$

$$\log\left(\frac{p_i}{q_i}\right) < \log\left(1 + \frac{p_i}{q_i}\right)$$

$$\log\left(\frac{p_i}{q_i}\right) < \log\left(\frac{p_i+q_i}{2q_i}.\,2\right)$$

$$\log\left(\frac{p_i}{q_i}\right) < \log\left(\frac{p_i+q_i}{2q_i}\right) + \log 2$$

$$\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) < \sum_{i=1}^{n} p_i \log\left(\frac{p_i+q_i}{2q_i}\right) + \left(\sum_{i=1}^{n} p_i\right)\log 2$$

$$\sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) < \sum_{i=1}^{n}(p_i-q_i)\log\left(\frac{p_i+q_i}{2q_i}\right) - \sum_{i=1}^{n} q_i \log\left(\frac{2q_i}{p_i+q_i}\right) + \log 2$$

$$K(P,Q) < 4M_1^*(P,Q) - F(Q,P) + \log 2$$

$$K(P,Q) - 4M_1^*(P,Q) + F(Q,P) < \log 2$$

Hence the desired result is verified.

5

Theoretic Based Exponential Information Divergence Measure and Inequalities

5.1 Introduction

In this chapter, a non-parametric theoretic based exponential information divergence measure is proposed. This measure belongs to the category of Csiszár's $f-$divergences. Further for first time, we have derived some inequalities for this exponential information divergence measure in terms of some valuable information divergence measures. Some numerical illustrations are carried out, based on two distinct discrete probability distributions.

Let $\Gamma_n = \{P = (p_1, p_2, p_3, \ldots p_n); p_i > 0, \sum_{i=1}^{n} p_i = 1\}$, n≥2, be the set of all complete finite discrete probability distributions. Ali & Silvey [3] and Csiszár [24, 25] introduced a generalized measure of information using f-divergence measure

$$C_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{5.1.1}$$

Where $f:(0,\infty)\rightarrow$R (set of real numbers) is a convex function. Most common choices of function f satisfy $f(1) = 0$, so that $C_f(P,Q)=0.$ Convexity of function f ensures that divergence measure $C_f(P,Q)$ is nonnegative. An important characteristic of this divergence measure is that many known divergences can be obtained from this measure by appropriately defining the convex function f, like [Shannon (1958), Renyi (1961), Ali & Silvey (1966), Vajda (1972), Jain and Patni (1976), Berbea & Rao (1982a, b), Taneja (1995), Kumar & Chhina (2005), Kumar & Johnson (2005), Jain and Shrivastva(2007), Jain and Mathur (2011), Jain and Chhabra (2014) and many more].

Further, these information divergence measures are used to find out distance or affinity between two probability distributions. Non parametric divergence measures give the amount of information supplied by the data for discriminating in favor of a probability distribution P $=\{p_1, p_2, p_3, \ldots p_n\}$ against another Q $=\{q_1, q_2, q_3, \ldots q_n\}$,where P, Q $\in \Gamma_n$. The construction of information divergence measure for two distinct probability distributions is not an easy task.

In this research work, we are introducing a new theoretic based non-parametric exponential information divergence measure which fits to the category of Csiszár's $f-$divergences [25, 28].

In 5.2; we will discuss some advantageous inequalities. New exponential information divergence measure is achieved in section 5.3. In section 5.4, we get some information on inequalities for the new exponential information divergence measure in terms of some recognized and valued divergence measures. Some numerical illustrations of new exponential information measure are shown in section 5.5. Section 5.6 concludes the chapter.

For rapidity, we will denote p_i, q_i and $\sum_{i=1}^{n}$ by p, q and $\sum$ respectively.

During past years P. Kumar and others [92, 93, 94] have contributed a lot of work providing different kinds of information, bounds on the distance and divergence measures. His existing information divergence measures are as under

P. Kumar and S. Chhina [92]

$$f(t) = \frac{(t+1)(t-1)^2}{t} \ln\left(\frac{t+1}{2\sqrt{t}}\right) \tag{5.1.2}$$

Then

$$S(P,Q) = \sum \frac{(p+q)(p-q)^2}{(pq)} \ln\left(\frac{p+q}{2\sqrt{pq}}\right) \tag{5.1.3}$$

And

$$S_\rho(P,Q) = \sum \left(\frac{(p-q)^2}{2p^2 q}\right)\left[2\left(2p^2 + pq + q^2\right)\ln\left(\frac{p+q}{2\sqrt{pq}}\right) + (p^2 - 2pq + q)\right] \tag{5.1.4}$$

P. Kumar and A. Johnson [94]

$$f(t) = \frac{\left(t^2-1\right)^2}{2t^{3/2}} \tag{5.1.5}$$

Then,

$$\Psi M(P,Q) = \sum \frac{\left(p^2-q^2\right)^2}{2(pq)^{3/2}} \tag{5.1.6}$$

And

$$\Psi M_\rho(P,Q) = \sum \frac{(p-q)\left(p^2-q^2\right)\left(5p^2+3q^2\right)}{4p^{5/2}q^{3/2}} \tag{5.1.7}$$

5.2 Well Know Inequalities

In this section, we will discuss some well-known inequalities which are established in literature of pure and applied mathematics. Using following inequalities we have derived important bounds of well-known divergence measures

$$\frac{x}{1+x} \le \ln(1+x) \le x, \quad x > 0 \tag{5.2.1}$$

$$x - \frac{x^2}{2} \le \ln(1+x) \le x - \frac{x^2}{2(1+x)}, \quad x > 0 \tag{5.2.2}$$

$$1 + x \le \exp\{x\}, \qquad x > 0 \tag{5.2.3}$$

$$x \quad < \exp\{x\}, \qquad x > 0 \tag{5.2.4}$$

5.3 New Exponencial Information Divergence Measure

Now, we consider the function $f:(0,\infty) \to R$ given by

$$f(t) = \frac{(t^2-1)^2}{2t^{3/2}} \exp\{t\} \tag{5.3.1}$$

And thus the new theoretic exponential information divergence measure

$$\Phi D(P,Q) = \sum \frac{(p^2-q^2)^2}{2(pq)^{3/2}} \exp\left\{{p}/{q}\right\} \tag{5.3.2}$$

Next,

$$f'(t) = \frac{1}{4}(t^2-1)\left(\frac{2t^3+5t^2-2t+3}{t^{5/2}}\right)\exp\{t\} \tag{5.3.3}$$

And

$$f''(t) = \frac{1}{8}\ \frac{\left[20t^5 + 7t^4 + 4\left(t^3 - 1\right)^2 + 6\left(t - 1\right)^2 + 5\right]}{t^{7/2}}\ \exp\{t\} \tag{5.3.4}$$

The function f (t) is convex since $f''(t) > 0$ for all t > 0 and normalized also since f (1) = 0.

Figure 5.1, Shows the behavior of the function $f(t)$ which is continuously convex. Thus the measure is nonnegative and convex in the pair of discrete probability distributions $(P, Q) \in \Gamma_n$.

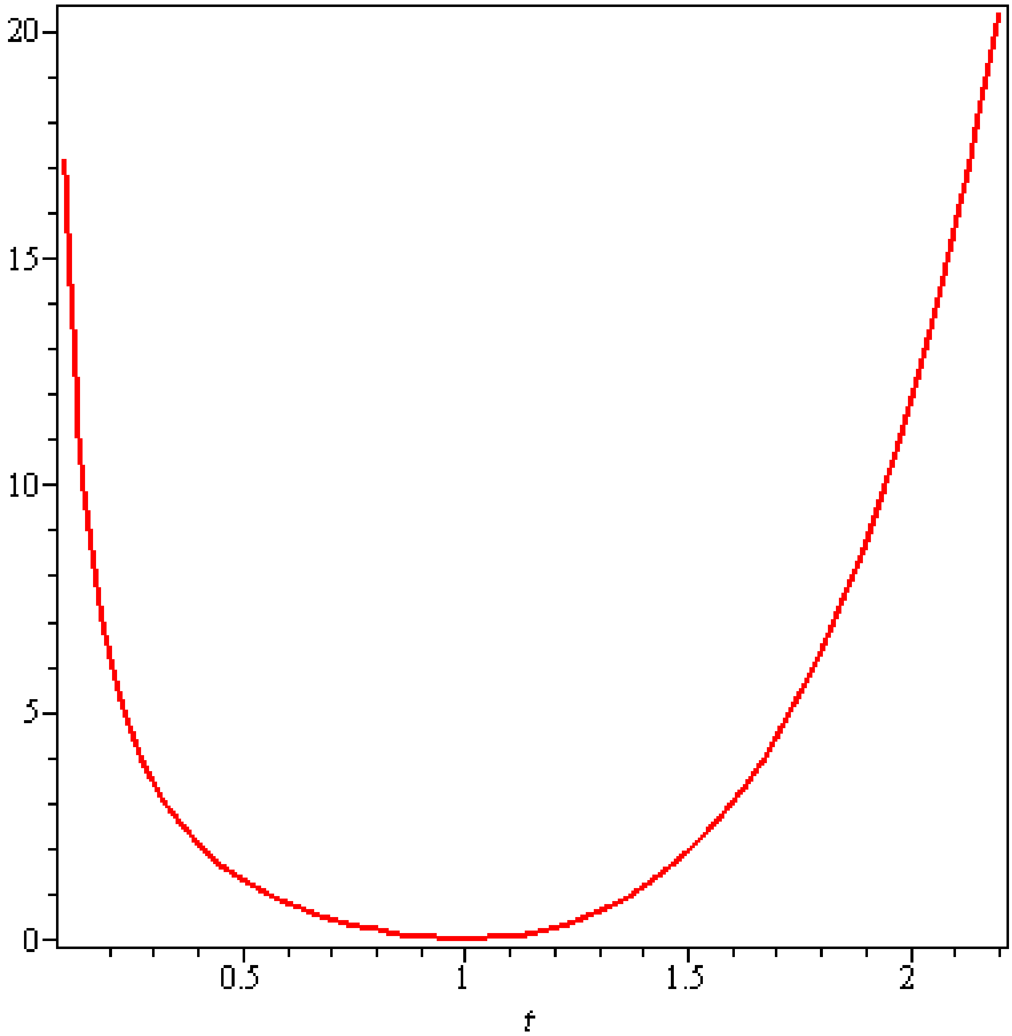

FIGURE 5.1: Graph of the convex function *f* (t)

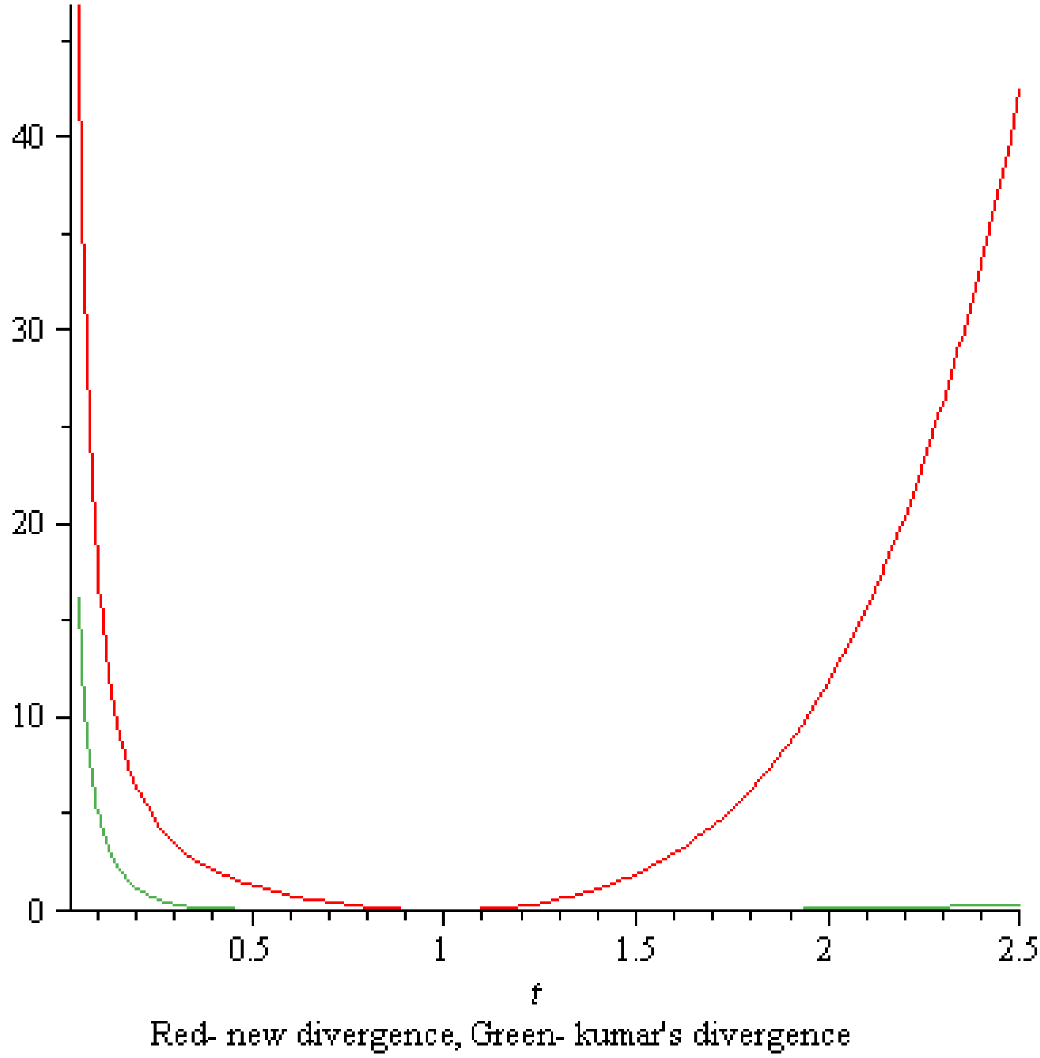

FIGURE 5.2: comparison graph of conxex functions of new divergence $\Phi D(P,Q)$ and Kumar's divergence $S(P,Q)$

5.4 BOUNDS FOR $\Phi D(P,Q)$

We now develop information inequalities providing bounds for $\Phi D\,(P,Q)$ in terms of the recognized information divergence measures in the following propositions.

Proposition 5.4.1: Let $\Phi D\,(P,Q)$ and $\Psi M\,(P,Q)$ be defined as (5.3.2) and (5.1.6) respectively and the symmetric χ^2 -divergence

$$\Psi\,(P,Q) = \chi^2(P,Q) + \chi^2(Q,P) = \sum \frac{(p+q)(p-q)^2}{pq} \tag{5.4.1}$$

Then inequality

$$\Psi\,(P,Q) \le \Psi M\,(P,Q) \le \Phi D\,(P,Q), \tag{5.4.2}$$

Holds and equality, *iff* $P = Q$.

Proof: Considering the Harmonic mean (HM) and Geometric mean (GM) inequality,

$$\text{HM} \le GM$$

Or,

$$\frac{2pq}{p+q} \le \sqrt{pq}\,,$$

Or,

$$\frac{p+q}{2\sqrt{pq}} \le \left(\frac{p+q}{2\sqrt{pq}}\right)^2,$$

$$\frac{(p+q)(p-q)^2}{pq} \le \frac{(p^2-q^2)^2}{2(pq)^{3/2}}, \tag{5.4.3}$$

Further, for

x > 0,

exp {*x*} > 1,

Thus for, $0 < p \leq 1$ *and* $0 < q \leq 1, \dfrac{p}{q} > 0,$

$1 < \exp\left\{ p \big/ q \right\},$

$$\frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \leq \frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \exp\left\{ p \big/ q \right\} \tag{5.4.4}$$

Now, from (4.3) and (4.4), we get

$$\frac{(p+q)(p-q)^2}{pq} \leq \frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \leq \frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \exp\left\{ p \big/ q \right\}$$

Summing over all terms we get,

$$\sum \frac{(p+q)(p-q)^2}{pq} \leq \sum \frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \leq \sum \frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}} \exp\left\{ p \big/ q \right\}$$

Therefore, we get

$$\Psi\left(P,Q\right) \leq \Psi M\left(P,Q\right) \leq \Phi D\left(P,Q\right),$$

Hence, the inequality

Proposition 5.4.2: Let $S\left(P,Q\right)$ *and* $\Phi D\left(P,Q\right)$ be defined as (5.1.3) and (5.3.2), respectively. Then inequality

$$S\left(P,Q\right) \leq \Phi D\left(P,Q\right) \tag{5.4.5}$$

And equality holds for $P = Q$

Proof: From inequality (5.2.1) and (5.2.4), we get,

$$\ln(1+x) \le x < \exp\{x\}, \qquad x>0$$

$$\Rightarrow \ln\left(\frac{1+x}{2\sqrt{x}}\right) < \exp\{x\}, \qquad x>0$$

$$\Rightarrow \frac{(x^2-1)^2}{2\,x^{3/2}} \ln\left(\frac{1+x}{2\sqrt{x}}\right) < \frac{(x^2-1)^2}{2\,x^{3/2}} \exp\{x\} \tag{5.4.6}$$

Replacing, x by p/q , we get,

$$\frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \ln\left(\frac{p+q}{2\sqrt{pq}}\right) \le \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\} \tag{5.4.7}$$

Or, $$\left(\frac{p+q}{2\sqrt{pq}}\right) \quad \frac{(p+q)(p-q)^2}{pq} \ln\left(\frac{p+q}{2\sqrt{pq}}\right) \le \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\}$$

We have, $\dfrac{2\sqrt{pq}}{p+q} \le 1 \Rightarrow \dfrac{p+q}{2\sqrt{pq}} \ge 1$

And thus,

$$\frac{(p+q)(p-q)^2}{pq} \ln\left(\frac{p+q}{2\sqrt{pq}}\right) \le \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\}$$

Summing over both sides, we get,

$$\sum \frac{(p+q)(p-q)^2}{pq} \ln\left(\frac{p+q}{2\sqrt{pq}}\right) \le \sum \frac{(p^2-q^2)^2}{2\,(pq)^{3/2}} \exp\left\{p/q\right\}$$

Hence, the required inequality,

$S(P,Q) \le \Phi D\,(P,Q)$.

Proposition 5.4.3: Let $\Phi D\left(P,Q\right), \Psi M\left(P,Q\right)$ *and* $\Psi\left(P,Q\right)$ be defined as (5.3.2), (5.1.6) and (5.4.1) respectively. Then inequality

$$2\,\Psi\left(P,Q\right)-\frac{3}{2}\,\Psi M\left(P,Q\right)-\sum\frac{\left(p-q\right)^2}{\sqrt{pq}}\le\Phi D\left(P,Q\right) \qquad (5.4.8)$$

And equality holds for $P=Q$.

Proof: From inequality (5.4.6), we have

$$\frac{\left(x^2-1\right)^2}{2\,x^{3/2}}\ln\left(\frac{1+x}{2\sqrt{x}}\right)<\frac{\left(x^2-1\right)^2}{2\,x^{3/2}}\exp\{x\}$$

Or, $\frac{\left(x^2-1\right)^2}{2\,x^{3/2}}\ln\left(\frac{2\sqrt{x}}{1+x}\right)<\frac{\left(x^2-1\right)^2}{2\,x^{3/2}}\exp\{x\}$

Replacing, x by p/q, we get,

$$\frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}}\ln\left(\frac{2\sqrt{pq}}{p+q}\right)\le\frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}}\exp\left\{p/q\right\} \qquad (5.4.9)$$

Further, we know that

$$\ln\left(\frac{2\sqrt{pq}}{p+q}\right)\approx\frac{4\sqrt{pq}}{p+q}-\frac{2pq}{\left(p+q\right)^2}-\frac{3}{2} \qquad (5.4.10)$$

Now, from (5.4.9) and (5.4.10), we get

$$\frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}}\left(\frac{4\sqrt{pq}}{p+q}-\frac{2pq}{\left(p+q\right)^2}-\frac{3}{2}\right)\le\frac{\left(p^2-q^2\right)^2}{2\left(pq\right)^{3/2}}\exp\left\{p/q\right\}$$

Arranging in appropriate forms and summing over all terms, we get

$$2\sum\frac{(p+q)(p-q)^2}{pq}-\frac{3}{2}\sum\frac{\left(p^2-q^2\right)^2}{2(pq)^{3/2}}-\sum\frac{(p-q)^2}{\sqrt{pq}}\le\sum\frac{\left(p^2-q^2\right)^2}{2(pq)^{3/2}}\exp\left\{p/q\right\}$$

Using (5.4.1), (5.1.6), and (5.3.2), we get

$$2\,\Psi\,(P,Q)-\frac{3}{2}\,\Psi M\,(P,Q)-\sum\,\frac{(p-q)^2}{\sqrt{pq}}\le\Phi D\,(P,Q)$$

Proposition 5.4.4: Let $\Phi D\,(P,Q)$, $\Psi M\,(P,Q)$ *and* $\Psi M_\rho(P,Q)$ be defined as (5.3.2), (5.1.6) and (5.1.7), respectively. Then inequality

$$\Psi M_\rho(P,Q)-\Psi M\,(P,Q)\le\Phi D_\rho(P,Q)-\Phi D\,(P,Q)\qquad(5.4.11)$$

Holds and equality i*ff P =Q*.

Where

$$\Phi D_\rho(P,Q)=\frac{1}{4}\frac{(p-q)\left(p^2-q^2\right)\left(2p^3+5p^2q-2pq^2+3q^3\right)}{(pq)^{5/2}}\exp\left\{p/q\right\}\qquad(5.4.12)$$

Proof: From (5.3.3), we have

$$f'(t)=\frac{1}{4}\left(t^2-1\right)\left(\frac{2t^3+5t^2-2t+3}{t^{5/2}}\right)\exp\{t\}$$

And, thus

$$\Phi D_\rho(P,Q)=\sum(p-q)f'\left(\frac{p}{q}\right)$$

$$\Phi D_\rho(P,Q)=\frac{1}{4}\sum\frac{(p-q)\left(p^2-q^2\right)\left(2p^3+5p^2q-2pq^2+3q^3\right)}{(pq)^{5/2}}\exp\left\{p/q\right\}\qquad(5.4.13)$$

Further, from (5.3.2) and (5.4.6), we get

$$\Phi D_\rho(P,Q)-\Phi D\,(P,Q)=\frac{1}{4}\sum\frac{(p-q)\left(p^2-q^2\right)\left(2p^3+5p^2q-2pq^2+3q^3\right)}{(pq)^{5/2}}\exp\left\{p/q\right\}\;(5.4.14)$$

From (5.1.6) and (5.1.7), we get

$$\Psi M_{\rho}(P,Q) - \Psi M\ (P,Q) = \frac{1}{4}\sum \frac{(p-q)\left(p^2-q^2\right)q\left(3p^2-2pq+3q^2\right)}{(pq)^{5/2}} \qquad (5.4.15)$$

From (5.4.14) and (5.4.15), we get inequality (5.4.11).

5.5 Numerical Illustration

In this section, we consider an example of symmetrical probability distributions. We will numerically verify the bounds achieved in the earlier section. For this, we calculate measures

$$\Phi D\ (P,Q), \Psi M\ (P,Q), S(P,Q), \textit{ and } \Psi\ (P,Q).$$

Let P be the binomial probability distribution for the random variable X with parameter (n = 8, p = 0.5) and Q its approximated normal probability distribution.

Table 5.1 Binomial Probability Distribution (n = 8, p = 0.5)

X	0	1	2	3	4	5	6	7	8
$p(x)$	0.0040	0.0310	0.1090	0.2190	0.2740	0.2190	0.1090	0.0310	0.0040
$q(x)$	0.0050	0.0300	0.1040	0.2200	0.2820	0.2200	0.1040	0.0300	0.0050
$\frac{p(x)}{q(x)}$	0.8000	1.0333	1.0481	0.9955	0.9716	0.9955	1.0481	1.0333	0.8000

The divergence measures $\Phi D\ (P,Q), \Psi M\ (P,Q), S(P,Q), \textit{ and } \Psi\ (P,Q)\ \textit{are}:$

$\Phi D\ (P,Q) = 0.08103214,$ $\qquad \Psi M\ (P,Q) = 0.00306097$

$S(P,Q) = 0.00001030,$ $\qquad \Psi\ (P,Q) = 0.00305063$

It is noted that

$$0.8000\ \leq\ \left(\frac{p}{q}\right)\ \leq\ 1.0481$$

These numerical values of measures verified the propositions of previous section.

6

An Information Measure of *f*- Divergence Family and Its Properties

6.1 Introduction

In this chapter, a non-parametric symmetric information divergence measure is proposed. This information measure belongs to the family of Csiszár's $f-$divergences. Its properties are studied and discussed. Further, we have derived some new bounds for this information divergence measure in terms of some recognized divergence measures based on two distinct probability distributions. The information divergence measures are used to find out distance or difference or affinity between two probability distributions. The construction of information divergence measure for two distinct probability distributions is always a herculean task.

Let $\Gamma_n = \left\{ P = (p_1, p_2, p_3, \ldots, p_n);\ p_i > 0, \sum_{i=1}^{n} p_i = 1 \right\}, n \geq 2$ be the set of all discrete probability distributions. For convex function $f : (0,\infty) \to R$ (set of real numbers) and for probability distributions P, Q ϵ Γ_n, the $f-$divergence measure , by Csiszár's, [24, 25] and Ali & Silvey, [3], is defined as

$$C_f(P,Q)=\sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{6.1.1}$$

An important property of this divergence measure is that many known divergence measures can be obtained from this information measure by appropriately defining the convex function f.

The complete chapter is organized as follows:. In section 6.2; we discuss information inequalities of Csiszár's $f-$divergences with other known divergence measures. New symmetric information divergence measure is obtained in section 6.3. In section 6.4, we have derived some information inequalities for the new information divergence measure in terms of some well-known divergence measures. Section 5, concludes the paper.

For brevity, we will denote $C_f(P,Q)$, p_i, q_i and $\sum_{i=1}^{n}$ by $C(P,Q)$, p, q and $\sum$, respectively.

6.2 Information Inequalities

Different kinds of bounds on the information divergence measures have been studied during the recent past [29- 37]. In [148], Kumar and Taneja unified and generalized information bounds for *C (P, Q)* studied by Dragomir [29- 37]**.** The main results in [148] given in the following theorem.

Theorem 6.2.1: Let $f: R_+ \to R$ be a mapping which is normalized and suppose that

(iii) f is twice differentiable on (r, R), $0 \le r \le 1 \le R < \infty$

(iv) There exists real constants *m, M,* such that *m* < *M* and $m \le t^{2-s} f''(t) \le M, \forall\, t \in (r,R), s \in \mathbb{R}$.

If $P,Q \in \Gamma_n$ are discrete probability distributions with $0 < r \le p/q \le R < \infty$,

$$m\,\Phi_s(P,Q) \le \quad C(P,Q) \quad \le M\ \Phi_s(P,Q), \tag{6.2.1}$$

And

$$m\{\eta_s(P,Q)-\Phi_S(P,Q)\} \le C_\rho(P,Q)-C(P,Q) \le M\ \{\eta_s(P,Q)-\Phi_s(P,Q)\} \tag{6.2.2}$$

Where

$$\Phi_s(P,Q)=\begin{cases} {}^2K_s(P,Q), & s \ne 0,1, \\ K(Q,P), & s=0, \\ K(P,Q), & s=1, \end{cases} \tag{6.2.3}$$

$${}^2K_s(P,Q)=[s(s-1)]^{-1}\left[\sum p^s q^{1-s}-1\right], \quad s \ne 0,1, \tag{6.2.4}$$

$$K(P,Q)=\sum p \ln\left(\frac{p}{q}\right),$$

$$C_\rho(P,Q)=\sum (p-q)\ f'\left(\frac{p}{q}\right), \tag{6.2.5}$$

$$\eta_s(P,Q)=C_{\phi_s'}\left(\frac{P^2}{Q},P\right)-C_{\phi_s'}(P,Q)$$

$$=\begin{cases} (s-1)^{-1}\sum (p-q)\left(\frac{p}{q}\right)^{s-1}, & s \ne 1, \\ \sum (p-q)\ln\left(\frac{p}{q}\right), & s=1, \end{cases} \tag{6.2.6}$$

As a consequence of this theorem, following information inequalities which are interesting from the *information-theoretic* point of view are also obtained in [25].

(i) The case s = 2 provides the information bounds in terms of the Chi-square divergence $\chi^2(P,Q)$,

$$\frac{m}{2}\chi^2(P,Q) \le C(P,Q) \le \frac{M}{2}\chi^2(P,Q), \tag{6.2.7}$$

And

$$\frac{m}{2}\chi^2(P,Q) \le C_\rho(P,Q) - C(P,Q) \le \frac{M}{2}\chi^2(P,Q) \tag{6.2.8}$$

(ii) For s = 1, the information bounds in terms of the Kullback-Leibler divergence, *K (P, Q),*

$$m\,K(P,Q) \le C(P,Q) \le M\ K(P,Q) \tag{6.2.9}$$

And

$$m\,K(P,Q) \le C\rho(P,Q) - C(P,Q) \le M\ K(P,Q) \tag{6.2.10}$$

(iii) For s = ½, yields the information bounds in terms of the Hellinger's discrimination

$$h\,(P,Q),\ 4m\,h(P,Q) \le C(P,Q) \le 4M\ h(P,Q) \tag{6.2.11}$$

And

$$4\,m\left(\frac{1}{4}\eta_{1/2}(P,Q) - h(P,Q)\right) \le C_\rho(P,Q) - C(P,Q) \le 4\,M\left(\frac{1}{4}\eta_{1/2}(P,Q) - h(P,Q)\right) \tag{6.2.12}$$

(iv) For s = 0, the information bounds in terms of the Kullback-Leibler and Chi-square divergence

$$mK(P,Q) \le C(P,Q) \le MK(P,Q) \tag{6.2.13}$$

And

$$m\left\{\chi^2(Q,P) - K(Q,P)\right\} \le C\rho(P,Q) - C(P,Q) \le M\left\{\chi^2(Q,P) - K(Q,P)\right\} \tag{6.2.14}$$

6.3 New Symmetric Information Divergence Measure

Now, we consider the function $f:(0,\infty)\to R$ given by

$$f(t) = \frac{(t-1)^4\left(t^2+1\right)(t+1)\left(t^2+3t+1\right)}{t^4} \tag{6.3.1}$$

Then

$$f'(t) = \frac{(t-1)^3\left(5t^6+15t^5+15t^4+15t^3+14t^2+12t+4\right)}{t^5} \tag{6.3.2}$$

And

$$f''(t) = \frac{10(t+1)^2(t-1)^2\left(2t^5+t^3+t^2+2\right)}{t^6} \tag{6.3.3}$$

The function f (t) is convex since f'' (t) > 0 for all t >0 and normalized also since f (1) = 0.

Figure 5.1, Shows the behavior of the function $f(t)$ and which is always convex.

Now, we have the following new information divergence measure belonging to the Csiszár's f-divergence family,

$$Z(P,Q)=\sum\frac{(p-q)^4\left(p^2+q^2\right)(p+q)\left(p^2+3pq+q^2\right)}{p^4q^4} \tag{6.3.4}$$

We have,

(a) Divergence measure *Z (P, Q)* is symmetric with respect to probability distributions.

(b) $Z(P,Q)\geq 0$ *and* $Z(P,Q)=0,$ *iff* $P=Q$

(c) The function f (t) is *-self conjugate since $f^*(t)\equiv t\ f\left(\frac{1}{t}\right)=f(t)$.

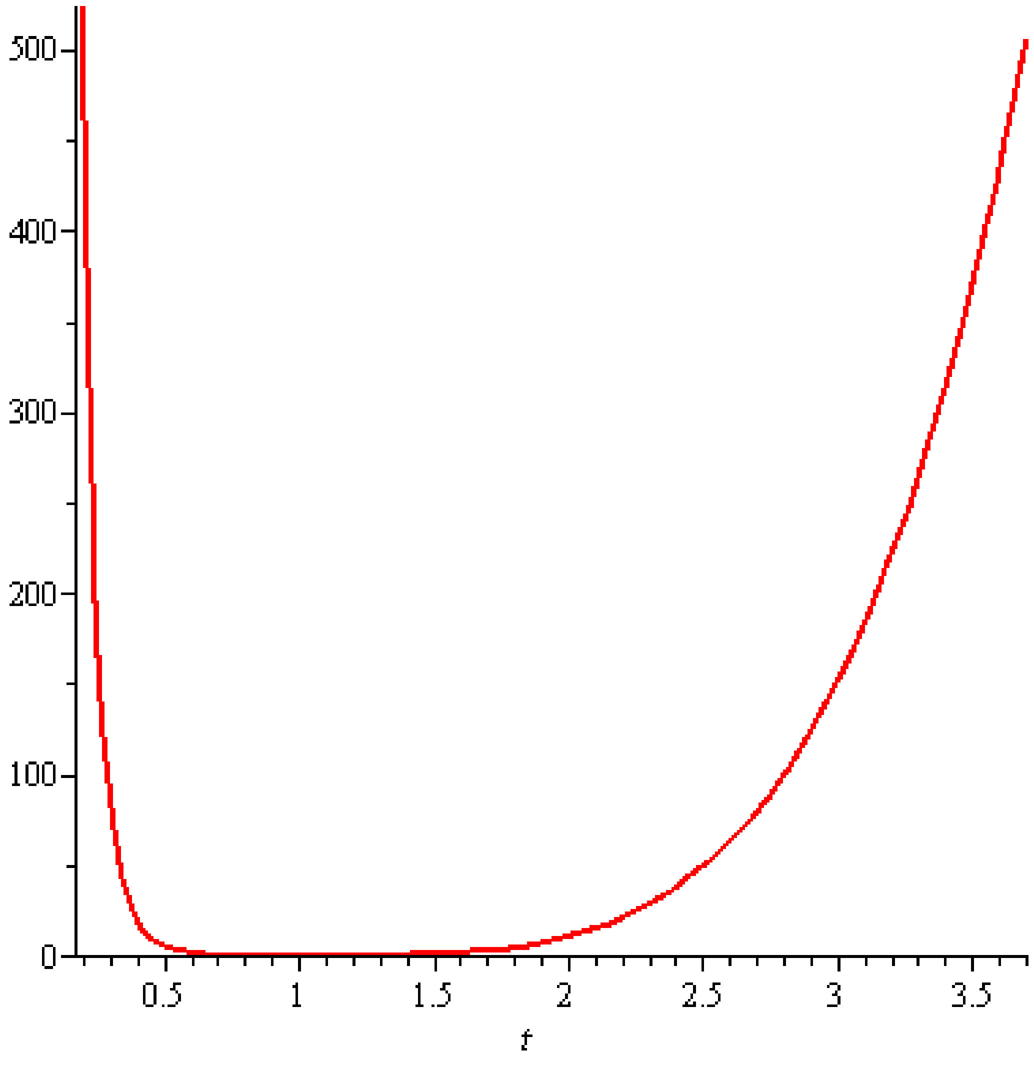

FIGURE 6.1: Graph of convex function *f* (t)

6.4 Bounds For $Z(P,Q)$

We now derive information inequalities providing bounds for Z (P, Q) in terms of the well-known divergence measures in the following propositions.

Proposition 6.4.1: Let (P, Q) $\in \Gamma_n \times \Gamma_n$, then we have the following new inequality

$$Z(P,Q) \le \left[\chi^2(P,Q)+\chi^2(Q,P)\right]^2\left[\left(\frac{2}{W(P,Q)}\right)^2+\left(\frac{1}{B(P,Q)}\right)^2\right] \tag{6.4.1}$$

Where $\chi^2(P,Q)+\chi^2(Q,P), W(P,Q)$ *and* $B(P,Q)$ are given by (6.1.3), (6.1.8) and (6.1.7) respectively.

Proof: From (6.3.4), we have,

$$Z(P,Q)=\sum\frac{(p-q)^4(p+q)(p^2+q^2)(p^2+3pq+q^2)}{p^4q^4}$$

$$\le\sum\frac{(p-q)^4(p+q)(p+q)(p^2+3pq+q^2)}{p^4q^4} \text{ ;Since } p^2+q^2\le p+q$$

$$=\sum\frac{(p-q)^4(p+q)^2}{p^2q^2}\sum\frac{(p^2+3pq+q^2)}{p^2q^2}$$

$$=\left[\sum\frac{(p-q)^2(p+q)}{pq}\right]^2\left[4\left(\sum\frac{p+q}{2pq}\right)^2+\left(\sum\frac{1}{\sqrt{pq}}\right)^2\right]$$

$$Z(P,Q)\le\left[\chi^2(P,Q)+\chi^2(Q,P)\right]^2\left[\left(\frac{2}{W(P,Q)}\right)^2+\left(\frac{1}{B(P,Q)}\right)^2\right]$$

Hence the result

Proposition 6.4.2: Let (P, Q) $\in \Gamma_n \times \Gamma_n$, then we have the following new inequality

$$Z(P,Q) \le 8\ \Psi M(P,Q)\,\Delta(P,Q)\left[\frac{1}{W(P,Q)}\right]^2 \frac{1}{B(P,Q)} \tag{6.4.2}$$

Where $Z(P,Q), \Psi M(P,Q), \Delta(P,Q), W(P,Q)$ *and* $B(P,Q)$ are given by (6.3.4), (6.1.9), (6.1.6), (6.1.8) and (6.1.7) respectively.

Proof: From (6.3.4), we have,

$$Z(P,Q) = \sum \frac{(p-q)^4 (p+q)\left(p^2+q^2\right)\left(p^2+3pq+q^2\right)}{p^4 q^4}$$

$$\le \sum \frac{(p-q)^4 (p+q)(p+q)\left(p^2+3pq+q^2\right)}{p^4 q^4}$$

$$= 2\sum \frac{\left(p^2-q^2\right)^2}{2(pq)^{3/2}} \sum \frac{1}{\sqrt{pq}} \quad \sum \frac{(p-q)^2}{p+q} \quad \sum \frac{(p+q)\left(p^2+3pq+q^2\right)}{p^2q^2}$$

$$\le 2\sum \frac{\left(p^2-q^2\right)^2}{2(pq)^{3/2}} \sum \frac{1}{\sqrt{pq}} \quad \sum \frac{(p-q)^2}{p+q} \quad 4\left(\sum \frac{p+q}{2pq}\right)^2$$

Since $p^2+3pq+q^2 \le p+q$

$$Z(P,Q) \le 8\ \Psi M(P,Q)\ \ \Delta(P,Q)\ \left[\frac{1}{W(P,Q)}\right]^2 \ \frac{1}{B(P,Q)}$$

Hence the result

Proposition 6.4.3: Let $\chi^2(P,Q)$ and *Z (P, Q)* be defined as in (6.1.2) and (6.3.4) respectively.

For $P, Q \in \Gamma_n \times \Gamma_n$ and $0 < r \leq p/q \leq R < \infty,$ we have

$$\frac{5(R+1)^2(R-1)^2(2R^5+R^3+R^2+2)}{R^6}\chi^2(P,Q) \leq Z(P,Q) \leq \frac{5(r+1)^2(r-1)^2(2r^5+r^3+r^2+2)}{r^6}\chi^2(P,Q) \tag{6.4.3}$$

And

$$\frac{5(R+1)^2(R-1)^2(2R^5+R^3+R^2+2)}{R^6}\chi^2(P,Q) \leq Z_\rho(P,Q) - Z(P,Q) \leq \frac{5(r+1)^2(r-1)^2(2r^5+r^3+r^2+2)}{r^6}\chi^2(P,Q) \tag{6.4.4}$$

Where

$$Z_\rho(P,Q) = \sum \frac{(p-q)^4(5p^6+15p^5q+15p^4q^2+15p^3q^3+14p^2q^4+12p\,q^5+4q^6)}{p^5q^4} \tag{6.4.5}$$

Proof. From (6.3.2), we have

$$f'(t) = \frac{(t-1)^3(5t^6+15t^5+15t^4+15t^3+14t^2+12t+4)}{t^5} \tag{6.4.6}$$

So that

$$Z_\rho(P,Q) = \sum (p-q) f'\left(p/q\right) = \sum \frac{(p-q)^4(5p^6+15p^5q+15p^4q^2+15p^3q^3+14p^2q^4+12p\,q^5+4q^6)}{p^5q^4} \tag{6.4.7}$$

Further, from (6.3.3), we have

$$f''(t) = \frac{10(t+1)^2(t-1)^2(2t^5+t^3+t^2+2)}{t^6} \tag{6.4.8}$$

If $t \in [a, b] \subset (0, \infty)$, *then*

$$\frac{10(b+1)^2(b-1)^2(2b^5+b^3+b^2+2)}{b^6} \le f''(t) \le \frac{10(a+1)^2(b-1)^2(2a^5+a^3+a^2+2)}{a^6}$$

Or, consequently

$$\frac{10(R+1)^2(R-1)^2(2R^5+R^3+R^2+2)}{R^6} \le f''(t) \le \frac{10(r+1)^2(r-1)^2(2r^5+r^3+r^2+2)}{r^6} \quad (6.4.9)$$

Hence, from (6.2.7) and (6.2.8), we get inequalities (6.4.3) and (6.4.4), respectively.

Proposition 6.4.4: Let $K(P,Q), Z(P,Q)$ *and* $Z_\rho(P,Q)$ be defined as in (6.1.4), (6.3.4) and (6.4.5) respectively. For $P, Q \in \Gamma_n \times \Gamma_n$ and $0 < r \le p/q \le R < \infty$, we have

$$\begin{aligned} \frac{10(R+1)^2(R-1)^2(2R^5+R^3+R^2+2)}{R^5} K(P,Q) &\le Z(P,Q) \\ &\le \frac{10(r+1)^2(r-1)^2(2r^5+r^3+r^2+2)}{r^5} K(P,Q) \end{aligned} \quad (6.4.10)$$

And

$$\begin{aligned} \frac{10(R+1)^2(R-1)^2(2R^5+R^3+R^2+2)}{R^5} K(P,Q) &\le Z_\rho(P,Q) - Z(P,Q) \\ &\le \frac{10(r+1)^2(r-1)^2(2r^5+r^3+r^2+2)}{r^5} K(P,Q) \end{aligned} \quad (6.4.11)$$

Proof: From (6.3.3), we have

$$f''(t) = \frac{10(t+1)^2(t-1)^2(2t^5+t^3+t^2+2)}{t^6} \quad (6.4.12)$$

Let the function $h : [r, R] \to \mathbb{R}$, such that

$$h(t) = t\, f''(t) = \frac{10(t+1)^2(t-1)^2(2t^5+t^3+t^2+2)}{t^5} \quad (6.4.13)$$

Then

$$\inf_{t \in [r,R]} h(t) = \frac{10(R+1)^2(R-1)^2\left(2R^5 + R^3 + R^2 + 2\right)}{R^5} \tag{6.4.14}$$

And

$$\sup_{t \in [r,R]} h(t) = \frac{10(r+1)^2(r-1)^2\left(2r^5 + r^3 + r^2 + 2\right)}{r^5} \tag{6.4.15}$$

Hence, from (6.2.9) and (6.2.10), we get inequalities (6.4.10) and (6.4.11), using (6.4.14) and (6.4.15), respectively.

7

Conclusion and Future Scope

In this chapter we conclude the work reported in this thesis and also discuss the scope for further study which can be carried out on the basis of the work reported.

7.1 Conclusion of the work reported

Several generalized divergences had been introduced in information theory for comparing two probability distributions at time, like: Csiszar'sdivergence([24], [25]), Burbea–Rao's ([17]), Renyi's divergence ([125]), Taneja and Tuteja's measure of inaccuracy ([137]), Taneja and Kumar's divergence ([148]), Jain and Saraswat's divergence ([57]) etc. Motivated by the findings of various authors have studied information measures.

We have obtained thedifferent equalities and inequalities of derived non-parametric symmetric*f*-divergence measures in terms of other well-known information divergence measures, which are very interesting in the field of the information theory.

In the next chapter, we have introduced different series of information divergence measures using properties of Csiszar's*f* - divergence and convex function. Divergence measures give most important key results for information theory because we can derive different information divergence measures for different values of k.

Further, we have obtained the different information measures using Jain and Saraswat's generalized f - divergence which are very interesting in the field of the information theory. We have also described different equalities and inequalities of derived *f*-divergence measures in terms of other well-known information divergence measures.

During past years Dragomir [29-37], Teneja [136-146], Kumar and others [147-148] gave the idea of divergence measures, their properties, bounds and relations with other measures. Kumar and other did a lot of work especially in the field of information theory. In [147-148], he derived new bounds in terms of different symmetric and non-symmetric divergence measures. We have introduced a new exponential non-parametric divergence measure, in the Csiszár's*f*-divergence category [24-25], by considering a convex function*f*, defined on $(0, \infty)$. This chapter also defines the bounds and properties of new exponential information measure with the work of Kumar's divergence measures and some other well-known measures.

In the last chapter, in this new work, we have obtained the new information divergence measure using properties of Csiszár's *f*- divergence category. This work is very interesting in the field of the information theory. We have also described different bounds for new derived *f*-divergence measure in terms of other recognized information divergence measures. Work on a

7.2 Future Scope

While compiling this thesis some thoughts have originated in our mind, with several new directions that open with the study reported here, which could be of great potential to study further.

a. The findings of Thesis (Information measures) can be used as practical application in different research fields of engineering and sciences as in Signal Processing in communication system, Cryptography etc.

b. One can generalize our new information divergence measures and their properties for his further research in the field of information theory.

c. Study of divergence measures in fuzzy mathematics as fuzzy directed divergences and fuzzy entropies, which are very useful to find the amount of difficulty in making a decision whether an element belongs to a set or not (Hooda [52]).

d. Study of new information inequalities in mutual information sense (Dragomiretc [36])

References

[1] **Aczel, J. and Daroezy, Z. (1975),** On measures of information and their characterizations, New York- San Francisco- Landon. Academic Press, XII, 234 S. (Mathematics in Science and Engineering 115) (page 581)

[2] **Aczel, J., Forte, B. and Ng, C.T.(1974),** Why the Shanon and Hartley entropies are natural, *Adv. Appl. Prob.*, 6, 131-146.

[3] **Ali, S.M. and Silvey, S.D. (1966),** A general class of coefficients of divergences of one distribution from another, Jour. Roy. Statist. Soc. B, 28, 131-142.

[4] **Arimoto, S. (1971)**, Information theoretic considerations on estimation Problems, Information and Control, 19, 181-190.

[5] **Baig, M.A.K. and Dar, R.A. (2007),** A new generalized bound for Csiszar's f- divergnece in terms of Renyi's entropy of order α and its applications" , Antarctica Journal of Mathematics, 4(1), 65-77

[6] **Baig, M.A.K. and Dar, R.A. (2005),** On JS and AG divergence measure and Information inequalities, Twelth Annual International Confrence on Statistics, Combinatorics, Mathematics and Applications, Auburn University (USA).

[7] **Balasubrahmanyam, C. and Siromoney, G. (1968),** A note on entropy of Telungu prose", Information and Control, 13, 281-285.

[8] **Bassat Beth, M. (1978),** *f*- entropies, probability of error and feature selection, Information Control, 39, 227-242

[9] **Belis, M. and Guiasu, S. (1968),** A quantitative qualitative measure of Information in cybernetics systems, IEEE Trans. on Information Theory, 14, 593-594.

[10] **Bercher, J.F. (2008)**, On some entropy functionals derived from Renyi information divergence, Information Sciences, 178, 2489-2506

[11] **Bhaker, U.S. and Hooda, D.S. (1993),** On mean value characterization of useful information measure, Tamkang Journal of Mathematics, vol. 24(4), 383-384.

[12] **Bhattacharyya, A. (1946),** On some analogues to the amount of information and their uses in statistical estimation, Sankhya: The Indian Journal of Statistics (1933- 1960) , vol. 8 (3), 201-218.

[13] **Boekee, D.E. (1977),** A generalization of the Fisher Information measure, Delft University Press, Delft.

[14] **Borges, R. (1967),** Zur Herleitung der Shannonschen information, Math. Z., 96, 282-287.

[15] **Brillouin, L. (1956),** Science and Information theory, Academic Press, New York.

[16] **Burbea, J. and Rao, C.R. (1982),** Entropy diffrential metric, distance and divergence measures in Probability spaces: A unified approach, J. Multivariate Analysis, vol. 12, 575-596.

[17] **Burbea, J. and Rao, C.R. (1982),** On the convexity of some divergence measures based on Entropy functions, IEEE Trans. on Information Theory, IT-28, 489- 495.

[18] **Chaundy, T.W. and MclLeod, J.B. (1960),** On a functional equation, Proc. Edinburgh Math. Soc. Notes, vol. 43, 7- 8.

[19] **Chen, H.C. (1973),** Statistical Pattern Recognition, Hoyderc Book Co., Rocelle Park, New York.

[20] **Chernoff, H. (1952),** A measure of asymptotic efficiency for test of a hypothesis based on a sum of observations, Ann. Math. Stat., vol. 30, 493-507.

[21] **Chow, C.K. and Lin, C.N. (1968),** Approximating discrete probability distributions with dependence tress", IEEE Trans. on Information Theory. vol. 14(3), 462- 467.

[22] **Cover, T.M. and Thomas, J.A. (1991),** Elements of Information Theory, Willey Series in Telecommunication, John Willey and Sons, New York.

[23] **Cozzolono, J.M. and Zahner, M.J. (1973),** The Maximum entropy distibution of the future market price of a stock, Operations Res., vol. 21(6) , 1200-1211.

[24] **Csiszár's, I. (1963)**, Eine informations theoretische ungleichung und ithre anwendung auf den beweis der ergodizitat von markoffschen ketten. Publ. Math. Inst. Hunger. Acad. Sci., 8, 85-107.

[25] **Csiszár's, I**. **(1974),** Information measures: A critical survey, In Trans. seventh prague conf. on information Theory, volume A, Academia, Prague, 73- 86 .

[26] **Darocozy, Z. (1970)**, Generalized Information Measures, Information and Control, vol. 16, 36- 51.

[27] **Darocozy, Z. (1964),** Einige Ungleichungen uber die mit Gewichtsfunktionen gebideten Mittelwerte, Monatsh. Math., 68, 102- 112.

[28] **Debibh, K.G. and Denbibh, S.J. (1983),** Entropy in relation to incomplete knowledge, Cambridge University Press, Cambridge

[29] **Dragomir, S. S. (2000),** Some inequalities for the Csiszar's ϕ- divergence, Inequalities for the Csiszar's *f*- divergence in Information Theory, **Chapter 1, Article 1**, (Edited by S. S. Dragomir), Available online at: http://rgmia.vu.edu.au/monographs/csiszar.html.

[30] **Dragomir, S. S. (2000),** A converse inequality for the Csiszar's ϕ-divergence, Inequalities for the Csiszar's *f*- divergence in Information Theory, **Chapter 1, Article 2**, (Edited by S. S. Dragomir), Available online at: http://rgmia.vu.edu.au/monographs/csiszar.html.

[31] **Dragomir, S. S. (2000),** Some inequalities for (m, M)-convex mappings and applications for the Csiszar's ϕ- divergence, in information theory, Inequalities for the Csiszar's *f*- divergence in Information Theory, **Chapter**

1, Article 3, (Edited by S. S. Dragomir), Available online at: http://rgmia.vu.edu.au/monographs/csiszar.html.

[32] **Dragomir, S.S. (2000)**, Upper and lower bounds for Csiszar's *f*- divergence in terms of the Kullback- Leibler distance and applications, Inequalities for the Csiszar's *f*- divergence in Information Theory, **Chapter 2, Article 1,** (Edited by S.S. Dragomir), Available online at: http://rgmia.vu.edu.au/monographs/csiszar.html.

[33] **Dragomir, S. S. (2000)**, Upper and lower bounds for Csiszar's *f*- divergence in terms of the Hellinger discrimination and applications, Inequalities for the Csiszar's *f*- divergence in Information Theory, **Chapter 2, Article 2**, (Edited by S.S. Dragomir), Available online at: http://rgmia.vu.edu.au/monographs/csiszar.html.

[34] **Dragomir, S.S. and Ionescu, N.M. (1994),** Some converse of Jensen inequality and Applications, Anal. Num. Theor. Approx, 23, 71-78.

[35] **Dragomir, S. S. (2013),** Inequalities of Pompeiu's type for absolutely continuous functions with applications to Ostrowski's inequality, RGMIA Res. Rep. Coll., vol. 16, no. 66, online: http://rgmia.org/papers/v16/v16a66.pdf

[36] **Dragomir, S.S., Gluscevic, V. and Pearce**, **C.E.M. (2001),** Approximation for the Csiszar *f*– divergence via mid- point inequalities, in inequality theory and applications- Y. J. Cho, J. K. Kim and S.S. dragomir (Eds.), Nova Science Publishers, Inc., Huntington, new York, vol. 1, 139-154.

[37] **Dragomir, S.S., Sunde, J., and Buse, C. (2000),** New Inequalities for Jeffrey Divergence Measures, Tamusi Oxford Journal of Mathematical Sciences, vol. 16(2), 295- 309.

[38] **Duda, R., and Hart, P. (1973),** Pattern classification and scene analysis, Wiley, New York

[39] **Dutta, M. (1974),** On some applications of information theoretic measures in Mathematical Sciences, Proceedings of Conference on Measures of Information and their Applications, HT Bombay, 107- 115.

[40] **Feldman, D. and Osterreicher, F. (1981),** Divergenzen von Wahrscheinlichkeitsverteilungen- integralgeometrisch betrachtet, Acta Math, Acad. Sci. Hungar., vol. 37 (4), 329- 337.

[41] **Ferentimos, K. and Papaiopannou, T. (1981),** New Parametric measure of Information, Information and Control,vol. 51, 193- 208.

[42] **Fisher, R.A. (1925),** Theory of statistical estimation, Proc. Cambridge Philos. Soc., vol. 22, 700– 725.

[43] **Gallager, R.G. (1968),** Information theory and reliable Communication, J. Wiley & Sons, New York.

[44] **Gokhale, D.V. and Kullback, S. (1978),** Information in Contingency Tables", New York, Marcel Dekker.

[45] **Goldman, S. (1953)**, Information Theory, Prentice- Hall, Englewood, Cliffs, NJ, New York.

[46] **Hartley, R. (1928)**, Transmission of information, in Bell System Technical Journal, vol. 7, 535- 563.

[47] **Havrda, J. and Charvat, F. (1967),** Quantification method of classification Process, Concept of Structural α - Entropy, Kybernetika, vol. 3, 30- 35.

[48] **He Yun, Hamza, A.B., and Krim, H. (2003),** A generalized divergence measure for robust image registration, IEEE Transaction on signal processing, vol. 51 (5), 1211- 1220.

[49] **Helinger, E. (1967),** Neue Begrundung der theorie der quadratischen formen von unendlichen vielen veraderlichen, J. Rein. Aug. Math., vol. 136, 210- 271.

[50] **Hooda, D.S. and Ram, A. (2002),** on useful relative information and J-divergence measures, Tamkang Journal of Mathematics, vol. 33, 146- 160.

[51] **Hooda, D.S. and Tuteja, R.K. (1985),** on characterization of non-additive measures of relative information and inaccuracy, Bull. Cal. Math. Soc., vol. 77, 363- 369.

[52] **Hooda, D.S. (2004),** On generalized measures of fuzzy entropy, Mathematica Slovaca, vol. 54(3), 315- 325.

[53] **Horowitz, I. (1970),** Employment concentration in the common market- an entropy approach, J. Roy. Statist. Soc. Ser. A., 133, 463- 479.

[54] **Jain, K.C. and Srivastava, A. (2007),** On symmetric information divergence measure of Csiszar's *f*- divergence class, Journal of Applied Mathematics, Statistics and Informatics, vol. 3 (1), 85- 102.

[55] **Jain, K. C. and Saraswat, R. N. (2011),** Some bounds of Csiszar's *f*-divergence measure in terms of the well- known divergence measures of Information Theory, International Journal of Mathematical Sciences and Engineering Applications, vol. 5. no. 5 (September), 1- 11.

[56] **Jain, K. C. and Saraswat, R. N. (2011),** Relations among some *f*-divergence measures using a new information inequality, Ganita Sandesh, vol. 5, no. 1, 1- 13.

[57] **Jain, K. C. and Saraswat, R. N. (2012),** Some new information inequalities and its applications in information theory, International Journal of Mathematics Research, vol. 4, no. 3, 295- 307.

[58] **Jain, K. C. and Saraswat, R. N. (2012),** A new information inequality and its application in establishing relation among various *f*- divergence measures, Journal of Applied Mathematics, Statistics and Informatics , vol. 8(1), 17-32.

[59] **Jain, K. C. and Saraswat, R. N. (2012),** inequalities among well- known divergence measures in terms of Triangular discrimination, International Journal of Computational and Applied Mathematics, vol. 4, no. 3, 215- 224.

[60] **Jain, K. C. and Mathur, R. (2011),** A symmetric divergence measure and its bounds, Tamkang Journal of Mathematics, vol. 42 (4), 493- 503.

[61] **Jain, K. C. and Saraswat, R. N. (2012),** Series of Information divergence measures using new *f*- divergence, convex functions and inequalities, International Journal of Modern Engineering Research, vol. 2, 3226- 3231.

[62] **Jain, K .C. and Chhabra, P. (2014)**, Series of new information divergences, properties and corresponding series of metric spaces, International Journal of Innovative Research in Sciences, Engineering and Technology, vol. 3(5), 12124- 12132

[63] **Jain, K .C. and Chhabra, P. (2014)**, New information inequalities and its special cases, Journal of Rajasthan Academy of Physical Sciences, vol. 13 (1), 39- 50.

[64] **Jain, K .C. and Chhabra, P. (2014)**, Establishing relations among various measures by using well known inequalities, International Journal of Modern Engineering Research, vol. 4(1), 238- 246.

[65] **Jain, K. C, and Singh, D. (2018),** New f-Divergence Measure and its Inequalities, Ganita Sandesh**,** vol. 30 (1), 39- 52.

[66] **Jain, K. C. and Singh, D. (2017),** Theoretic based exponential information divergence measure and inequalities, International Journal of Creative Research Thoughts, vol. 5(12), 295- 305.

[67] **Jeffrey's, H. (1946),** An invariant form for the prior probability in estimation Problems, Proc. Roy. Soc. Lon. Series A, 186, 453- 461.

[68] **Jensen, J.L.W.V. (1905),** Om konvexe Funktioner og uligheder mellem middelvaerdier, Nyt. Tidsskr. Math. 16 B, 49- 69.

[69] **Jha, P. and Mishra, V. K. (2012),** Some new trigonometric, hyperbolic and exponential measures of fuzzy entropy and fuzzy directed divergence, International Journal of Scientific and Engineering Research, vol. 3 (4), 1- 5.

[70] **Jones, L. and Byrne, C. (1990),** General entropy criteria for inverse problems, with application to data compression, pattern classification and cluster analysis, IEEE Trans. Inform. Theory, vol. 36, 23- 30.

[71] **Jumarie, G. (1990),** Relative information theories and Applications, Springer Series in Synergetics, vol. 47, Springer, Berlin.

[72] **Kadota, T.T. and Shepp, L.A. (1967),** On the best finite set of linear observables for discriminating two Gaussian signals, IEEE Trans. Infor. Theory, vol. 13, 288- 294.

[73] **Kafka, P., Osterriecher, F., and Vincze, I. (1991),** On Powers of *f*-divergence defining a distance, Studia Sci. Math. Hungar., vol. 26, 415- 422.

[74] **Kagan, A.M. (1963),** On the Theory of Fisher"s amount of Information (in Russian), Dokl. Akad. Nauk SSSR, vol. 151, 277- 278.

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

[76] **Kanappan, P.L. (1972)**, On Directed Divergence and Inaccuracy, Z. Wahrs. Verw. Geb., vol. 25, 49- 55.

[77] **Kanappan, P.L. and Rathie, P.N. (1972),** On a functional equation connected with Sahnnons entropy, Funkcialaj Ekvaciaj, vol. 14, 153- 159.

[78] **Kanappan, P.L. and Rathie, P.N. (1973),** On a Characterization of Directed Divergence, Information and Control, vol. 22, 163-171.

[79] **Kapur, J.N. (1967),** Generalized entropy of order α and Type β, The Math. Seminar, vol. 4, 78-94.

[80] **Kapur, J. N. (1968),** On information of order a and type β, Proc. Indian Academy of Sciences, vol. 23A, 65- 75.

[81] **Kapur, J. N. (1984),** A comparative assessment of various measures of directed divergence, Advances in Management Studies, vol. 3, 1- 16.

[82] **Kapur, J. N. (1988b),** Some new non additive measures of entropy, Boll. U.M.I., vol. 7(2-B), 253- 266.

[83] **Kapur, J. N. (1990),** On trignometric Measures of Information, J. Math. Phys. Sci., vol. 24, 1- 10.

[84] **Kapur, J. N. (1994),** Measures of information and their applications, Wiley & Sons., New York, ISBN- 13: 978-0470220641.

[85] **Kazakos, D. and Cotside, T. (1980),** A decision theory approach to the approximation of discrete probability densities, IEEE Trans. Perform. Anal. Machine Intell., vol.1, 61- 67.

[86] **Kendall, D.G. (1964),** Functional equations in information theory. Z. Wahrscheinlichkeits theorie und Verw. Gebiete., vol. 2:225-229.

[87] **Kerridge, D.F. (1961),** Inaccuracy and inference, Journal of Royal Statistical Society, Series B, vol. 23(1), 184- 194.

[88] **Kolmogorav, A.N. (1963),** On the approximation of distributions of sums of independent summands by infinitely divisible distributions, Sankhya, vol. 25, 159-174.

[89] **Kolmogorav, A.N. (1958),** Entropy per unit time as a metric invariant of automorphism, Dokl. Akad. Nauk USSR, vol. 120, 754- 758.

[90] **Kullback, S. (1968),** Information theory and Statistics, Dover Publication, New York.

[91] **Kullback, S. and Leibler, R. A., (1951),** On information and sufficiency, Ann. Math. Statist., vol. 22, 79- 86.

[92] **Kumar, P. and Chhina, S. (2005),** A symmetric information divergence measure of the Csiszar's *f*- divergence class and its bounds, Computers and Mathematics with Applications, vol. 49, 575- 588.

[93] **Kumar, P. and Hunter, L. (2004),** On an information divergence measure and information inequalities, Carpathian Journal of Math., vol. 20 (1), 51- 66.

[94] **Kumar, P. and Johnson, A. (2005),** On a symmetric divergence measure and information inequalities, Journal of Inequalities in pure and applied Mathematics, vol. 6 (3), article 65, 1- 13.

[95] **Kumar, P. and Taneja, I.J. (2005),** On unified generalization of Relative Jensen- Shannon and Arithmetic- Geometric Divergence measure and their properties", Indian Journal of Mathematics and Mathematical Sciences, vol. 1 (1), 77- 97.

[96] **Le Cam, L. (1986),** Asymptotic methods in statistical Decision Theory, Springer- Verlag, New York-Berlin.

[97] **Lee, P.M. (1964),** On the axioms of information theory, Ann. Math. Stat., vol. 35, 414- 418.

[98] **Liese, F. and Vajda, I. (1987),** Convex Statistical Distances, Teubner- Texte Zur Mathematik, Band 95, Leipzig.

[99] **Longo, G. (1972),** Quantitative- qualitative measure of information, Springer New York.

[100] **Majernik, V. (2001),** Elementry Theory of Organization, Palacky University Press, Olomoue, ISBN: 80-244-0262-9.

[101] **Mathai, A. and Rathie, P.N. (1974),** Basic Concept in Information Theory and Statistics- Axiomatic Foundations and Applications, Wiley Eastern Limited, New Delhi.

[102] **Matusita, K. (1955),** Decision rules based on problem of fit, two samples and estimation, Ann. Inst. Statist. Math., vol.26, 631- 640.

[103] **Matusita, K. (1964),** Distance and Decision Rules, Ann. Inst. Statist. Math., vol.16, 305- 320.

[104] **Mei, M. (1978),** The theory of genetic distance and evaluation of human races, Japan Journal of Human Genetics, vol. 23, 341-369.

[105] **Michel, O., Baraniuk, R.G. and Flandrin, P. (1994),** Time-frequency based distance and divergence measures, Proceedings of IEEE- SP International Symposium on Time- Frequency and Time- Scale Analysis, 64-67.

[106] **Nath, P. (1969),** Some recent results in information theory, presented at A seminar on scientific information, organized by Defence Scientific Evaluation and Documentation Centre, Metcalfe House, Ministry of Defence, India.

[107] **Nath, P. (1970),** An axiomatic characterization of inaccuracy for discrete generalized Probability distribution, Opsearch, vol. 7, 115- 133.

[108] **Neyman, J. (1949),** Contribution to the theory of χ^2 – test, Proc. I. Berkeley Symp. On Math. Statist. , 239-273, Univ. Calif. Press Berkeley.

[109] **Nyquist, H., (1924),** Certain factors affecting telegraphy speed, Bell system Tech.Journal, vol. 3, 324-346.

[110] **Nyquist, H. (1928),** Certain topics in telegraphy transmission theory, A.I.E.E. Trans., Vol. 47, 617-644.

[111] **Osterreicher, F. (2002),** Csiszar's *f*-divergences basic properties, Res. Report Collection http://www.sbg.ac.at/mat/home.html.

[112] **Patni, G. C. and Jain, K. C. (1976),** On some information measures, Information and Control, vol. 31 (2), 185-192.

[113] **Patni, G. C. and Jain, K. C. (1976),** Generalized functional equation and its application in Information theory, Proc. Indian Acad. Sci. , vol. 86 A (4), 353- 364.

[114] **Patni, G. C. and Jain, K. C. (1977),** On axiomatic characterization of some non-additive measures of information, Mertrika, 24, S-23-34.

[115] **Pearson, K. (1900),** On the criterion that a given system of deviations from the probable in the case of correlated system of variables in such that it can be reasonable supposed to have arisen from random sampling, Phil. Mag., vol. 50, 157- 172.

[116] **Picard, C.F. (1979),** Weighted Probabilistic Information Measures, Journal of Inform. And System Sciences, vol. 4, 343- 356.

[117] **Pielou, E.C. (1975),** Ecological Diversity, John Wiley and Sons, New York.

[118] **Puri, M.L. and Vincze, I. (1990),** Measures of Information and contiguity, Statist. Probab. Letters, vol. 9, 223- 228.

[119] **Qian, H. (2000)**, Relative entropy: free energy associated with equilibrium fluctuations and on equilibrium deviations, 8 Jul 2008. https://arxiv.org/abs/math-ph/0007010/.

[120] **Rajagopalan, K.R. (1965),** A note on entropy of Kannada, Information and Control, vol. 8, 640-644.

[121] **Rao, C. R. (1982),** Diversity and dissimilarity coefficient: a unified approach, Theoretical Population Biology, Vol. 21, 24-43

[122] **Rathie, P.N. (1970),** On a generalized entropy and a coding theorem, Journal of Appl. Probl. vol. 7, 124-133.

[123] **Rathie, P. N. and Kannappan, P. I. (1972),** A directed- divergence function of type ß, Information and Control, vol. 20, no. 1, 38- 45.

[124] **Read, T.C.R. and Cressie, N.A. (1988),** Goodness-of-Fit statistics for discrete multivariate data, Springer, New York.

[125] **Renyi, A. (1961),** On measures of entropy and information, In Proceedings 4th Berkeley Symposium on Stat. and Prob., University of California Press,Berkeley, vol. 1, 547-561.

[126] **Reschenhofer, E. and Bomze, I.M. (1991),** Length tests for goodness of fit, Bio-metrika, vol. 78, 207- 216.

[127] **Sant'anna, A.P. and Taneja, I.J. (1985)**, Trignometric entropies Jensen difference divergence measures and error bounds, Information Sciences, vol. 35, 145- 156.

[128] **Shannon, C. E. (1948),** A mathematical theory of communications, Bell System Technical Journal, vol. 27, 379- 423.

[129] **Sharma, B.D. and Autar, R. (1974),** Relative Information Function and their Type (α, β) generalizations, Metrika, vol. 21, 41-50

[130] **Sharma, B. D. (1972),** On amount of information of type ß and other measure, Metrika, vol. 19, no. 1, 1-10.

[131] **Sharma, B. D. and Mittal, D. P. (1975),** New non- additive measures of entropy of discrete probability distribution, Journal of Mathematical Sciences, vol. 10, 28-40.

[132] **Sharma, B. D. and Taneju, I. J. (1977),** Three generalized additive measures of entropy, Elektronische Information Verarbeitung and Kybernetika, vol. 13, 419-433.

[133] **Shioya, H. and Da-te, T. (1995),** A generalization of Lin divergence and the derivative of a new I formation divergence, Elec. And Comm. In Japan, vol. 78 (70), 37-40.

[134] **Sibson, R. (1969),** Information radius, Z. wahrs und verw geb. , vol. 14(2), 149- 160.

[135] **Siromoney, G. and Rajgopalan, K.R. (1964),** Style as information in Karnatic music, The Journal of Music Theory, vol. 8(2), 267- 272.

[136] **Taneja, H. C.and Tuteja, R. K. (1984),** Characterization of quantitative and qualitative measure of relative information, Information Sciences, vol. 33, 217- 222.

[137] **Taneja, H. C. and Tuteja, R. K. (1986)**, Characterization of quantitative and qualitative measure of inaccuracy, Kybernetika, vol. 22, 393-402.

[138] **Taneja, I. J. (1989),** On generalized information measures and their application, Chapter in Ad. Elect. And Elct. Physics (Ed. P.W. Hawkes), vol. 76, 327-413

[139] **Taneja, I.J. (1995),** New developments in generalized information measures, Chapter in: Advances in imaging and Electron Physics, Ed. P. W. Hawkes ,vol. 91, 37-135

[140] **Taneja, I.J. (2001),** Generalized Information Measures and Their Apploications, https://www.mtm.ufsc.br/-taneja/book/book.html.

[141] **Taneja, I.J. (2004),** Relative Divergence Measures and Information inequalities, Inequality theory and applications, Vol. 4, Y.J. Cho., J.K. Kim and S.S. Dragomir (Eds.), Nova Science Publishers, Inc Hungtington, New York.

[142] **Taneja, I.J. (2005),** Bounds on Non-Symmetric Divergence measures in terms of Symmetric Divergence Measures, Journal of Combinatorics, Information and System Sciences, vol.29 (1-4), 115-134.

[143] **Taneja, I.J. (2005),** On symmetric and non-symmetric divergence measures and their generalization, Chapter in: Advances in Imaging and electron Physics, vol. 138, 177- 250.

[144] **Taneja, I.J. (2013),** Seven means, generalized triangular discrimination and generating divergence measures, Information, vol. 4 (2), 198-239.

[145] **Taneja, I.J. (2011),** Inequalities among logarithmic mean measures, Available online: http://arxiv.org/abs/1103.25880v1

[146] **Taneja, I.J. (2012),** Inequality having seven means and proportionality relations, Available online: http://arxiv.org/abs/120.2288

[147] **Taneja, I.J. and Kumar, P. (2004),** Relative informaton of Types-s, Csiszar's f-divergence and information inequalities, Information Sciences, vol. 166, 105-115.

[148] **Taneja, I.J. and Kumar, P. (2005),** Generalized non-symmetric divergence measures and inequalities, Available online: arxiv.math/0501300v1 [math ST].

[149] **Theil, H. (1967),** Economics and Information Theory, North-Holland Amsterdam, 1967.488 P. [Univ.Chicago, IL].

[150] **Topose, F. (1999),** Res. Rep. Collection, RGMIA, vol. 2(1), 85- 98.

[151] **Tverberg, G.H.(1958),** A new derivation of the information functions, Maths Scand, vol.6, 297-298.

[152] **Vajda, I. (1968),** Axioms for a-entropy of a generalized probability scheme, Kybernetika, vol. 4, 105-112.

[153] **Vajda, I. (1972),** On *f*-divergence and singularity of probability measures, Period. Math. Hungar., vol. 2, 223-234.

[154] **Vajda, I. (1973),** χ^{α} – divergences and generalized Fisher's Information, In:Trans. 6th Prague Conf. on Information Theory, Vol. B, Academia Prague, 873-876.

[155] **Vajda, I. (1989),** Theory of Statistical Infernece and Information, Kluwer Academic Publishers, Dordrecht-Boston, Series: Theory and Decision Library B, vol. 11.

[156] **Varma, R.S. (1965),** Entropy in Information Theory, Ganita, vol. 16, 1- 6.

[157] **Varma, R.S. (1968),** Generalization of Entropy of order α, Journal of Math. Sc. , vol. 1, 34- 38.

[158] **Varma, R.S. and Nath, P. (1967),** Information Theory: a survey, Journal of Math. Sci., vol. 2, 75-109.

[159] **Wiener, N. (1948),** Cybernatics, The MIT Press and Wiley, New York.

[160] **Wiener, N. and Aaron, D. (2001),** The significance of Shanon's work, ACM SIGMOBILE Mobile Computing and Communication Review vol. 5 (1), 56-57.

[161] **Zhu, Y.M. (2002),** Volume image registration by cross-entropy optimization, IEEE Transactions on Medical Imaging, vol. 21(2), 174-180.

Research Profile

Research Area: Information Theory

Research Topic: Study of information theoretic based measures and their properties.

Department: Mathematics, Malaviya National Institute of Technology, Jaipur.

Supervisor's: 1. K.C. Jain, Professor in Department of Mathematics, MNIT, Jaipur.

2. Rashmi Jain, Professor in Department of Mathematics, MNIT, Jaipur.

Research work published:

1. K.C. Jain and Devendra Singh, New f-Divergence Measure and its Inequalities, **GANITA SANDESH,** Volume 30, 39-52, 2018.
2. K.C. Jain and Devendra Singh, New logarithmic information divergence measure, **International Bulletin of Mathematical Research,** vol.4 (3), 45-49, **2017**
3. K.C. Jain and Devendra Singh, Theoretic based exponential information divergence measure and inequalities, **International Journal of Creative Research Thoughts**, Volume 5, Issue 12, 295- 305, 2017.
4. K.C. Jain and Devendra Singh, Inequalities and Equalities for New *f*-divergence Measure, **International Journal of Computer and Mathematical Sciences,** volume 6, issue5, 114-120, 2017.
5. K.C. Jain and Devendra Singh, New Information Divergence Measure of Csiszár's F-Divergence Class and it's Relations to other Well Known Divergence Measures, **International Journal of Engineering Management & Sciences**, Volume 2, Issue 12, 10-12, 2015.
6. K.C. Jain and Devendra Singh, An information measure of $f-$divergence family and its properties, Jnanabha, Special Issue, 72-80, 2018.
7. K.C. Jain and Devendra Singh, Series of new divergence measures and relation to other well-known divergence measures, **communicated.**

Papers Presented in Conferences

1. Presented a research paper **"Theoretic Based Exponential Information Divergence Measure & Inequalities**" at an International Conference on "**Communication and Computational Technologies**" from Dec 8-9, 2017, Organized by **RIET**, Jaipur.
2. Presented a research paper **"An information measure of *f*-divergence family and its properties**" at National Conference on "**Mathematical Sciences and Scientific Computing for Industrial Development**" from Nov 24-26, 2017, Organized by **Manipal University**, Jaipur.
3. Presented a research paper**"Inequalities and equalities for New *f*-Divergence Measure"**at an International Conference on "**New Frontiers of Engineering, Science, Management and Humanities**" on May 21, 2017, Organized by **National Institute of Technical Teachers Training & Research,** Chandigarh, India.

WORSHOPS/ SEMINAR/ CONFERENCES

1. 2017 – Attended one week **short term** course on "**Modeling & Simulation Tools"** from July 03-07, 2017 Organized by **MNIT**, Jaipur.
2. 2016 – Attended National level training programme on "**Advances in OFC Technology**" from 19-23 December 2016 Organized by **RIET**, Jaipur and **NITTTR** Chandigarh, India.
3. 2015 – Attended National Conference on "**Mathematical Analysis and Computation**" Organized by **MNIT**, Jaipur and **NITTTR** Chandigarh, India.
4. 2014 – Attended five days **short term** course on "**Analysis and Applications"** Organized by **MNIT**, Jaipur and **NITTTR**, Chandigarh, India.
5. 2013 – Attended one week **short term** course on "**Computational Techniques in Engineering and Sciences"** Organized by **MNIT**, Jaipur.
6. 2012 – Attended one week **short term** course on "**Mathematical Methods and its Application in Engineering and Sciences"** Organized by **MNIT**, Jaipur.
7. 2012 – Attended National Conference on "**Emerging Trends in Mathematics**" Organized by Dept. of Mathematics, **M.S.J. Govt. College, Bharatpur** and Rajasthan Ganita Parishad, Rajasthan.
8. 2011 – Attended National Conference on "**Emerging Trends in Wireless Communication**" Organized by Dept. of ECE, **RCERT**, Jaipur.
9. 2010 - Attended 21st National Conference of **Rajasthan GanitaParishad** on "**Recent Trends in Mathematics& Actuarial Sciences**" Organized by **Central University of Rajasthan** and **MNIT**, Jaipur.
10. 2010-Attended National Conference on "**Advances in Microwave Communication, Devices and Application**" Organized by Dept. of ECE, **RIET**, Jaipur.
11. 2008 - Attended 19th annual Conference of **Rajasthan GanitaParishad** on "**Numerical Techniques and their Application in Mathematics**" conducted by Department of Mathematics, **Jaipur Engineering College**, Jaipur.

Academic and Personal Profile

Devendra Singh

Date of birth: 1July 1981

Address: 3/338, Chitrakoot, Vaishali Nagar, Jaipur-302021

Email: devendra.mnit2013@gmail.com, devendrasingh_maths@yahoo.com

Contact Number: 7240000849/ 9588017548

Academic Profile:

Examination	University/ Board	Year of Passing	Percentage
M.Tech	Suresh GyanVihar University	2013	63.16
M.Phil	Madurai Kamaraj University	2009	60.60
M.Sc	University of Rajasthan	2003	67.60
B.Sc	University of Rajasthan	2001	61.70
12^{th}	Rajasthan Board	1998	55.85
10th	Rajasthan Board	1996	72 .00

Core Knowledge and Skill Area: Pure and Applied Mathematics, Information Theory and Coding, Statistics and Probability Theory, Operation Research, Engineering Mathematics etc.

Books Authored

- **Engineering Mathematics - I** published by College Book House, Jaipur.
- **Engineering Mathematics - II** published by College Book House, Jaipur.
- **Engineering Mathematics - III** published by V&T Publication, Jaipur.
- **Advanced Engineering Mathematics - IV** published by College Book House, Jaipur.
- **Discrete Mathematical Structure** published by College Book House, Jaipur.
- **Differential Equations** published by College Book House, Jaipur.

Teaching Experience

1. Malaviya National Institute of Technology, Jaipur August, 2018 to May 2019.
2. Rajasthan Institute of Engineering and Technology, Jaipur June 2016 to July 2017
3. Apex Institute of Engineering and Technology, Jaipur March 2013 to October 2015.
4. Rajasthan College of Engineering for Women, Jaipur October 2007 to February, 2013.
5. Regional College of Education Research and Technology, Jaipur July 2006 to Sept 2007.
6. Government College, Bandikui, Dausa September 2003 to April 2006.

Strengths: Positivism, Punctuality and Dedication for Work and Peers.

I hereby declare that the information furnished above is true and correct to the best of my knowledge.

Devendra Singh