Some Inequalities on 'Useful' Mean g −deviation with Applications in Information Theory

Pankaj Prasad Dwivedi[a], Dilip Kumar Sharma[b]
[a, b] Department of Mathematics Jaypee University of Engineering and Technology
A.B. Road, Raghogarh, Dist. Guna, 473226 India.

Abstract: The objective of this correspondence is to offer an elaboration of some latest inequalities' findings, in which we have given a new improvement of 'useful' Jensen's inequality, as well as utilization in the theory of information. In linear spaces, for convex functions constructed on a convex subset, an improvement inequality of Jensen's is provided. For 'useful mean g −deviation and 'useful' g −divergences, we provide robust lower bounds as well as the 'useful' mean h-absolute deviation, and lastly, we have given applications of divergence measure. Uniqueness for the 'useful' KL-Divergence and 'useful' Jeffreys divergence is obtained.

Keywords: Shannon information, Jensen's inequality, Local bounds, 'Useful' total variation distance, Utility distribution, 'Useful' information inequalities, KL-Divergence, 'Useful' Jeffreys divergence.

1. Introduction

Let $\Delta_k^+ = \{\boldsymbol{l} = (l_1, l_2, \dots, l_k);\ l_i \geq 0, \sum_i^k l_i = 1\}$, be a lot of all possible discrete likelihood distributions of a random variable $\boldsymbol{x} = \{x_1, x_2 \dots\dots\dots x_k\}$ and $\mho = \{(u_1, u_2, \dots\dots u_k);\ u_i > 0\ \forall i,\ \mho \in \mho_k\}$ are the utility distribution attached to each $\boldsymbol{l} \in \Delta_k^+$ such that $u_i > 0$ is the utility of an event having the probability of occurrence $l_i > 0$.

Let u_i be the utility or importance of the result x_i, as well as $\mho = (u_1, u_2, \dots\dots u_k)$ be the arrangement of non-negative actual numerals. In general, the utility is unaffected by the likelihood of encoding the source symbol x_i, i.e., l_i.

The pattern of information is provided by

$$\begin{bmatrix} x_1, x_2 \dots\dots\dots x_k \\ l_1, l_2 \dots\dots\dots l_k \\ u_1, u_2 \dots\dots\dots u_k \end{bmatrix}, \quad \text{Where } \sum_{i=1}^k l_i = 1, 0 < l_i \leq 1, u_i > 0, \tag{1}$$

The source is not completely identified by the distribution of likelihood $\boldsymbol{l}$ across the source symbols $\boldsymbol{x}$ without their qualitative character, according to Belis and Guiasu [5]. As a result of the experimenter's findings, it is indeed deduced that the source symbols or letters are given weights based on their significance or use. The following qualitative-quantitative measure of information was thus introduced by Belis and Guisau [5]:

$$H(\boldsymbol{l}; \mho) = -\sum_{i=1}^k u_i l_i \log l_i \tag{2}$$

The 'useful' information measure is called the quantitative-qualitative measure defined in (2). This calculation may be used as a statistic for the average amount of 'useful' information produced by the information system (1). It is clear that when utilities are ignored, (2) reduces to Shannon's information measure [25] which is given below:

$$H(\boldsymbol{l}) = -\sum_{i=1}^k l_i \log l_i \tag{3}$$

By using different postulates, several authors have specified the entropy of Shannon. By utilizing significant assumptions which Fadeev [14] deduced, Khinchin [17] made Shannon's statement more precise. Tverberg [28], Chandy and Mcliod [6], Kendall [16], etc., were further defined by the entropy of Shannon by considering various sets of postulates. For strongly convex and highly mid-convex functions, counterparts of the converse Jensen inequality were presented by Klaricic & Nikodem [18]. Dwivedi and Sharma [9] obtained the lower and upper bound for Renyi information rate in the terms of utility. In several fields of mathematics, convex functions play an essential role in information theory, Rashid et al., [23], further improvement of Jensen's inequality given by Dragomir [10], Ge-Jile et al. [15], and Sayyari, [24], the function is very important for the study of optimization problems, where it has many beneficial properties. Jeffreys-Renyi type divergences and Jensen-Renyi induced by convex functions and represented by Kluza [19], also, Kumari & Sharma [20] given 'useful' non-symmetric divergence for some cases.

Suppose g be a convex function on R, and also R be the convex subset of linear space $\mathcal{T}$. If $\boldsymbol{l} = (l_1, \ldots, l_k)$ is the sequence of chance distribution with utility distribution $\mho = \{(u_1, u_2, \ldots \ldots u_k);\ u_i > 0\ \forall i\}$ and $\boldsymbol{x} = (x_1, \ldots, x_k) \in R^k$, in this case, following 'useful' Jensen's inequality holds:

$$g\left(\frac{(\sum_{i=1}^{k} u_i l_i x_i)}{\sum_{i=1}^{k} u_i l_i}\right) \leq \frac{\sum_{i=1}^{k} u_i l_i g(x_i)}{\sum_{i=1}^{k} u_i l_i}\ . \tag{4}$$

It may be noted that when utilities are $u_i = 1$, then (4) reduces to Jensen's inequality. The inequality of Jensen is very significant of all inequalities because it has many mathematical and statistical applications and its special cases are several other well-known inequalities such as Holder's inequality, Cauchy's inequality, AGH inequalities, etc.

We have provided the following 'useful' refinement of (4):

$$g\left(\frac{\sum_{i=1}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i}\right) \leq \frac{\sum_{i_1,\ldots,i_{t+1}=1}^{k} u_{i_1} l_{i_1} \ldots u_{i_{t+1}} l_{i_{t+1}}}{\sum_{i=1}^{k} u_i l_i} g\left(\frac{x_{i_1} + \cdots + x_{i_{t+1}}}{t+1}\right)$$

$$\leq \frac{\sum_{i_1,\ldots,i_t=1}^{k} u_{i_1} l_{i_1} \ldots u_{i_t} l_{i_t}}{\sum_{i=1}^{k} u_i l_i} g\left(\frac{x_{i_1} + \cdots + x_{i_t}}{t}\right)$$

$$\leq \ldots \leq \frac{\sum_{i=1}^{k} u_i l_i g(x_i)}{\sum_{i=1}^{k} u_i l_i} \tag{5}$$

for $\boldsymbol{l} = (l_1, \ldots, l_k) \in \Delta_k^+, \boldsymbol{x} = (x_1, \ldots, x_k) \in R^k, \mho = (u_1, \ldots, u_k) \in \mho_k$ and $t \geq 1$.

The above measure (5) reduces to results which were obtained in 1989, by Pecaric and Dragomir [22], when utilities are ignored i.e., $u_i = 1$.

If r_1, ..., $r_t \geq 0$ with $\sum_{q=1}^{t} r_q = 1$, then we have the following refinement:

$$g\left(\frac{\sum_{i=1}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i}\right) \leq \frac{\sum_{i_1,\ldots,i_t=1}^{k} u_{i_1} l_{i_1} \ldots u_{i_t} l_{i_t}}{\sum_{i=1}^{k} u_i l_i} g\left(\frac{x_{i_1} + \cdots + x_{i_t}}{t}\right)$$

$$\leq \frac{\sum_{i_1,\ldots,i_t=1}^{k} u_{i_1} l_{i_1} \ldots u_{i_t} l_{i_t}}{\sum_{i=1}^{k} u_i l_i} g\left(r_1 x_{i_1} + \cdots + r_t x_{i_t}\right)$$

$$\leq \frac{\sum_{i=1}^{k} u_i l_i g(x_i)}{\sum_{i=1}^{k} u_i l_i} \tag{6}$$

Where $1 \leq t \leq k$, when utilities are ignored i.e., $u_i = 1$, then (6) is reduced to Jensen's inequality.

An improvement of the following Jensen's inequality:

$$g\left(\frac{\sum_{q=1}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right)$$

$$\leq \min_{t\in\{1,2,\ldots k\}}\left[(1-u_tl_t)g\left(\frac{\frac{\sum_{q=1}^{k}u_ql_qx_q}{\sum_{q=1}^{k}u_ql_q}-u_tl_tx_t}{1-u_tl_t}\right)+u_tl_tg(x_t)\right]$$

$$\leq \frac{1}{k}\left[\sum_{t=1}^{k}(1-u_tl_t)\,g\left(\frac{\frac{\sum_{q=1}^{k}u_ql_qx_q}{\sum_{q=1}^{k}u_ql_q}-u_tl_tx_t}{1-u_tl_t}\right)+\frac{\sum_{t=1}^{k}u_tl_tg(x_t)}{\sum_{t=1}^{k}u_tl_t}\right]$$

$$\leq \max_{t\in\{1,2,\ldots k\}}\left[(1-u_tl_t)g\left(\frac{\frac{\sum_{q=1}^{k}u_ql_qx_q}{\sum_{q=1}^{k}u_ql_q}-u_tl_tx_t}{1-u_tl_t}\right)+\ u_tl_tg(x_t)\right]$$

$$\leq \frac{\sum_{q=1}^{k}u_ql_qg(x_q)}{\sum_{q=1}^{k}u_ql_q} \tag{7}$$

We will call the above inequality as 'useful' Jensen's inequality, where g, x_t, and u_tl_t as above given, and after utilities are ignored i.e.,$u_i=1$, then (7) reduces to Jensen's inequalities.

A large body of work on Jensen's inequality and its various extensions, modifications, equivalents, and converse findings See, for example, [1,2,3,4,9,10,12,21,22] and [16,17] and also given by Simic [26], Tapus, and Popescu [27].

We provide purification of Inequality of Jensen's related through the functionals that are general in Section 2. We also currently acquired for 'useful' mean g −deviation of lower bound as well as the 'useful' mean h −absolute deviation in Section 3. last Section 4 has provided applications for 'useful' g −divergence measures in information theory, and Applications for norms especially for KL-divergence, χ^2- divergence, Absolute divergence, the 'useful' Jeffreys divergence, total variation divergence, etc. For the 'useful' g −divergences and 'useful' mean g −deviation, the bounds obtained are superior as compared to the bounds which are presented by Dragomir [11].

2. New improvements

In the real linear space $\mathcal{T}$, suppose that R is a convex subset, also let $g: R\to\mathbb{R}$ is a convex function on R. If $l_i>0$, with $\sum_{i=1}^{k}l_i=1$ where $1\leq i\leq k$ and $x_i\in R$, then we write $\bar{E}=\{1,2,\ldots,k\}\setminus E(\neq\emptyset)$ for any nonempty subset E of $\{1,2,\ldots,k\}$ and define $\Psi_E=\sum_{i\in E}^{k}l_i$ and $\bar{\Psi}_E=\Psi_{\bar{E}}=\sum_{q\in\bar{E}}l_q=1-\sum_{i\in E}^{k}l_i$. For the k −tuples $\boldsymbol{x}=(x_1,x_2,\ldots,x_k)$, $\boldsymbol{l}=(l_1,l_2,\ldots,l_k)$, the convex function g and $\mho=\{(u_1,u_2,\ldots\ldots u_k);\ u_i>0\ \forall i\}$ utilities are attached to each $\boldsymbol{l}\in\Delta_k^+$ shown before. We can define the following function

$$A(g,\boldsymbol{l},\boldsymbol{x};\ E,\mho)=\Psi_Eg\left(\frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k}u_il_ix_i}{\sum_{i=1}^{k}u_il_i}\right)+\bar{\Psi}_Eg\left(\frac{1}{\bar{\Psi}_E}\frac{\sum_{q\in\bar{E}}^{k}u_ql_qx_q}{\sum_{q=1}^{k}u_ql_q}\right) \tag{8}$$

where we will use, for $E\subset\{1,2,\ldots k\}$ with $E\neq\emptyset$ and $E\neq\{1,2,\ldots k\}$ here and everywhere also below:

It's worth noting that we have the function for $E=\{t\}, t\in\{1,2,\ldots,k\}$.

$$A_t(g,\boldsymbol{l},\boldsymbol{x};\mho)=A(g,\boldsymbol{l},\boldsymbol{x};\{t\},\mho)$$

$$= u_t l_t g(x_t) + (1 - u_t l_t) g\left(\frac{\sum_{i=1}^{k} u_i l_i x_i - u_t l_t x_t}{1 - u_t l_t}\right) \quad (9)$$

The above measure reduces to [11] when utilities aspects are ignored i.e., $u_i = 1$.

Theorem 1. Suppose that $g: R \rightarrow \mathbb{R}$ is a convex function on R and in the real linear space $\mathcal{T}$, let R be a convex subset, and $u_i > 0$ are the utilities attached to probabilities. For any nonempty subset $E = \{1,2, \dots, k\}$, if $l_i > 0$, with $\sum_{i=1}^{k} l_i = 1$, and $x_i \in R$, then we have

$$\frac{\sum_{t=1}^{k} u_t l_t g(x_t)}{\sum_{t=1}^{k} u_t l_t} \geq A(g, \boldsymbol{l}, \boldsymbol{x}; E, \mho) \geq g\left(\frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right) \quad (10)$$

Proof. We have the convexity of the function g

$$A(g, \boldsymbol{l}, \boldsymbol{x};\ E, \mho) = \Psi_E g\left(\frac{1}{\Psi_E} \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i}\right) + \bar{\Psi}_E g\left(\frac{1}{\bar{\Psi}_E} \frac{\sum_{q\in \bar{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right)$$

$$\geq g\left[\Psi_E\left(\frac{1}{\Psi_E} \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i}\right) + \bar{\Psi}_E\left(\frac{1}{\bar{\Psi}_E} \frac{\sum_{q\in \bar{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right)\right]$$

$$\geq g\left(\frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right)$$

This establishes the second 'useful' inequality in (10), for any E.

We also have by the Jensen inequality

$$\frac{\sum_{t=1}^{k} u_t l_t g(x_t)}{\sum_{t=1}^{k} u_t l_t} = \frac{\sum_{i\in E}^{k} u_i l_i g(x_i)}{\sum_{i=1}^{k} u_i l_i} + \frac{\sum_{q\in \bar{E}}^{k} u_q l_q g(x_q)}{\sum_{q=1}^{k} u_q l_q}$$

$$\geq \Psi_E g\left(\frac{1}{\Psi_E} \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i}\right) + \bar{\Psi}_E g\left(\frac{1}{\bar{\Psi}_E} \frac{\sum_{q\in \bar{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right)$$

$$= A(g, \boldsymbol{l}, \boldsymbol{x}; E, \mho)$$

This establishes the first 'useful' inequality in (10), for any E.

Remark 1. Here we notice that the inequality (10) may be expressed in the following ways:

$$\frac{\sum_{t=1}^{k} u_t l_t g(x_t)}{\sum_{t=1}^{k} u_t l_t} \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} A(g, \boldsymbol{l}, \boldsymbol{x}; E, \mho) \quad (11)$$

and

$$g\left(\frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right) \leq \min_{\emptyset \neq E \subset \{1,2,\dots k\}} A(g, \boldsymbol{l}, \boldsymbol{x}; E, \mho) \quad (12)$$

These inequalities imply the following findings, using a somewhat more difficult technique of proof:

$$\frac{\sum_{t=1}^{k} u_t l_t g(x_t)}{\sum_{t=1}^{k} u_t l_t} \geq \max_{t\in \{1,2,\dots,k\}} A_t(g, \boldsymbol{l}, \boldsymbol{x}; \mho) \quad (13)$$

and

$$g\left(\frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right) \leq \min_{t\in \{1,2,\dots,k\}} A_t(g, \boldsymbol{l}, \boldsymbol{x}; \mho) \quad (14)$$

Furthermore, since

$$\max_{\emptyset \neq E \subset \{1,2,\dots,k\}} A(g, \boldsymbol{l}, \boldsymbol{x}; E, \mho) \geq \max_{t\in \{1,2,\dots,k\}} A_t(g, \boldsymbol{l}, \boldsymbol{x}; \mho)$$

and

$$\min_{t\in\{1,2,\dots,k\}} A_t(g,\boldsymbol{l},\boldsymbol{x};\mho) \geq \min_{\emptyset\neq E\subset\{1,2,\dots k\}} A(g,\boldsymbol{l},\boldsymbol{x};E,\mho)$$

The resulting inequalities (11) and (12) are thus superior to the prior results from [11].

The case of uniform distribution, in which $l_i = \frac{1}{k}$ is the same for all $\{1,2,\dots,k\}$, is also interesting. When we take a natural integer a with it by $1 \leq a \leq k-1$, if we define

$$A_a(g,\boldsymbol{x}) := \frac{a}{k} g\left(\frac{1}{a}\sum_{i=1}^{a} x_i\right) + \frac{k-a}{k} g\left(\frac{1}{k-a}\sum_{q=a+1}^{k} x_q\right) \tag{15}$$

then we can come up with the following conclusion:

Corollary 1. Let $g: R \to \mathbb{R}$ is a convex function on R and in the real linear space $\mathcal{T}$, suppose R be a convex subset. If $x_i \in R$, then we have for any $a \in \{1,2,\dots,k-1\}$,

$$\frac{1}{k}\sum_{t=1}^{k} g(x_t) \geq A_a(g,\boldsymbol{x}) \geq g\left(\sum_{t=1}^{k} x_t\right) \tag{16}$$

We have the bounds, in particular

$$\frac{1}{k}\sum_{t=1}^{k} g(x_t) \geq \max_{a\in\{1,2,\dots,k-1\}} \left[\frac{a}{k} g\left(\frac{1}{a}\sum_{i=1}^{a} x_i\right) + \frac{k-a}{k} g\left(\frac{1}{k-a}\sum_{q=a+1}^{k} x_q\right)\right] \tag{17}$$

and

$$\min_{a\in\{1,2,\dots,k-1\}} \left[\frac{a}{k} g\left(\frac{1}{a}\sum_{i=1}^{a} x_i\right) + \frac{k-a}{k} g\left(\frac{1}{k-a}\sum_{q=a+1}^{k} x_q\right)\right] \geq g\left(\frac{1}{k}\sum_{t=1}^{k} x_t\right). \tag{18}$$

For symmetric convex functions, the subsequent variant of the inequality (10) may be useful:

Corollary 2. Suppose that R be a convex module with the characteristic that it has the value $0 \in R$. If $z_q \in \mathcal{T}$ such that for every $l_i > 0, 1 \leq i \leq k$, with $\sum_{i=1}^{k} l_i = 1$, we have $z_q - \frac{(\sum_{i=1}^{k} u_i l_i z_i)}{\sum_{i=1}^{k} u_i l_i} \in R$, for any $q \in \{1,2,\dots,k\}$ and $u_i > 0$ are the utilities attached to probabilities, then we have for any subset E of $\{1,2,\dots,k\}$

$$\begin{aligned}\frac{\sum_{t=1}^{k} u_t l_t g\left(z_t - \frac{\sum_{i=1}^{k} u_i l_i z_i}{\sum_{i=1}^{k} u_i l_i}\right)}{\sum_{t=1}^{k} u_t l_t} &\geq \Psi_E g\left[\overline{\Psi}_E\left(\frac{1}{\Psi_E}\frac{\sum_{i\in E} u_i l_i z_i}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\overline{\Psi}_E}\frac{\sum_{q\in \bar{E}} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q}\right)\right] \\ &\quad + \overline{\Psi}_E g\left[\Psi_E\left(\frac{1}{\overline{\Psi}_E}\frac{\sum_{q\in \bar{E}} u_q l_q z_q}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\Psi_E}\frac{\sum_{i\in E} u_i l_i z_i}{\sum_{q=1}^{k} u_q l_q}\right)\right] \\ &\geq g(0)\end{aligned} \tag{19}$$

Remark 2. If $E = \{t\}$, then the specific aspect we may deduce from the corollary is helpful as well (19) can be written as

$$\begin{aligned}\frac{\sum_{m=1}^{k} u_m l_m g\left(z_m - \frac{\sum_{i=1}^{k} u_i l_i z_i}{\sum_{i=1}^{k} u_i l_i}\right)}{\sum_{m=1}^{k} u_m l_m} &\geq u_t l_t g\left[(1-u_t l_t)\left(z_t - \frac{1}{1-u_t l_t}\left(\frac{\sum_{q=1}^{k} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q} - u_t l_t z_t\right)\right)\right] \\ &+(1-u_t l_t) g\left[u_t l_t\left(\frac{1}{1-u_t l_t}\left(\frac{\sum_{q=1}^{k} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q} - u_t l_t z_t\right) - z_t\right)\right] \geq g(0)\end{aligned} \tag{20}$$

which is equivalent with

$$\frac{\sum_{m=1}^{k} u_m l_m g\left(z_m - \frac{\sum_{i=1}^{k} u_i l_i z_i}{\sum_{i=1}^{k} u_i l_i}\right)}{\sum_{m=1}^{k} u_m l_m}$$

$$\geq u_t l_t g\left(z_t - \frac{\sum_{q=1}^{k} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q}\right) + (1 - u_t l_t) g\left[\frac{u_t l_t}{(1-u_t l_t)}\left(\frac{\sum_{q=1}^{k} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q} - z_t\right)\right]$$

$$\geq g(0) \tag{21}$$

where $t \in \{1,2,\dots,k\}$.

Remark 3. The continuous versions are contemplated in [13], for the Lebesgue integral.

3. Lower bound for 'useful' mean g −deviation

Suppose $\mathcal{T}$ be a real linear space and $u_i > 0$ are the utilities attached to probabilities. Define the 'useful' mean $g -$ deviation of an $k -$tuple of vectors $\boldsymbol{x} = (x_1, \dots, x_k) \in \mathcal{T}^k$ with the likelihood dispensation $\boldsymbol{l} = (l_1, \dots, l_k)$ by the nonnegative amount for a convex function $g: \mathcal{T} \to \mathbb{R}$ with the characteristics that $g(0) = 0$. then

$$M_g(\boldsymbol{l}, \boldsymbol{x}; \mho) = \frac{\sum_{i=1}^{k} u_i l_i g\left(x_i - \frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right)}{\sum_{i=1}^{k} u_i l_i} \tag{22}$$

Since $M_g(\boldsymbol{l}, \boldsymbol{x}; \mho)$ is positive, this follows that 'useful' inequality of Jensen's and can be expressed as

$$M_g(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq g\left(\frac{\sum_{i=1}^{k} u_i l_i\left(x_i - \frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right)}{\sum_{i=1}^{k} u_i l_i}\right) = g(0) = 0.$$

For convex function $g(x) = \|x\|^h$, $h \geq 1$, and $u_i > 0$ are the utilities attached to probabilities described on a normed linear space $(\mathcal{T}, \|\cdot\|)$ provides a natural example of such variations. This is denoted by

$$M_h(\boldsymbol{l}, \boldsymbol{x}; \mho) = \frac{\sum_{i=1}^{k} u_i l_i \left\|x_i - \frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right\|^h}{\sum_{i=1}^{k} u_i l_i} \tag{23}$$

The above measure called 'useful' mean $h -$ absolute variance with distribution $\boldsymbol{l} = (l_1, \dots, l_k)$ of the $k -$ tuple of vectors $\boldsymbol{x} \in \mathcal{T}^k$ and $u_i > 0$ are the utilities attached to probabilities.

For the 'useful' mean g −deviation the following findings give a better lower bound.

Theorem 2. Assume that a convex function $g: \mathcal{T} \to [0, \infty)$ with $g(0) = 0$. If $\boldsymbol{x} \in \mathcal{T}^k$ and $\boldsymbol{l} = (l_1, \dots, l_k)$ is a likelihood dispensation with all l_i are not zero, then

$$M_g(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{\Psi_E g\left[\bar{\Psi}_E\left(\frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\bar{\Psi}_E}\frac{\sum_{q\in \bar{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right)\right]\right.$$

$$\left. + \Psi_E g\left(\frac{1}{\bar{\Psi}_E}\frac{\sum_{q\in \bar{E}}^{k} u_q l_q z_q}{\sum_{q=1}^{k} u_q l_q} - \frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k} u_i l_i z_i}{\sum_{q=1}^{k} u_q l_q}\right)\right\} \geq 0 \tag{24}$$

In special, we have

$$M_g(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq \max_{t\in\{1,2,\dots,k\}} \left\{(1 - u_t l_t) g\left[\frac{u_t l_t}{(1-u_t l_t)}\left(\frac{\sum_{m=1}^{k} u_m l_m x_m}{\sum_{m=1}^{k} u_m l_m} - x_t\right)\right]\right.$$

$$+u_t l_t g\left(x_t - \frac{\sum_{m=1}^{k} u_m l_m x_m}{\sum_{m=1}^{k} u_m l_m}\right)\Bigg\} \geq 0 \quad (25)$$

Proof. Corollary 2 and Remark 2 provide the proof.

We have the following, in the particular case:

Corollary 3. Let a normed linear space to be $(\mathcal{T}, \|\cdot\|)$, and $u_i > 0$ are the utilities attached to probabilities. If $\boldsymbol{x} = (x_1, \dots, x_k) \in \mathcal{T}^k$ and $\boldsymbol{l} = (l_1, \dots, l_k)$ is a likelihood dispensation with all l_i are not zero, then we have for $h \geq 1$

$$M_h(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{\Psi_E \overline{\Psi}_E\left(\overline{\Psi}_E^{h-1} + \Psi_E^{h-1}\right)\left\|\frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\overline{\Psi}_E}\frac{\sum_{q\in \overline{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right\|^h\right\} \geq 0 \quad (26)$$

Remark 4. We can see that from the power function $g(s)^h, h \geq 1$ and then by the convexity of the function, we have

$$\Psi_E \overline{\Psi}_E\left(\overline{\Psi}_E^{h-1} + \Psi_E^{h-1}\right) = \Psi_E \overline{\Psi}_E^h + \overline{\Psi}_E^h \Psi_E \geq \left(\Psi_E \overline{\Psi}_E + \overline{\Psi}_E \Psi_E\right)^h = 2^h \Psi_E^h \overline{\Psi}_E^h$$

therefore

$$\Psi_E \overline{\Psi}_E\left(\overline{\Psi}_E^{h-1} + \Psi_E^{h-1}\right)\left\|\frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\overline{\Psi}_E}\frac{\sum_{q\in \overline{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right\|^h$$

$$\geq 2^h \Psi_E^h \overline{\Psi}_E^h \left\|\frac{1}{\Psi_E}\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \frac{1}{\overline{\Psi}_E}\frac{\sum_{q\in \overline{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right\|^h$$

$$= 2^h \left\|\overline{\Psi}_E \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \Psi_E \frac{\sum_{q\in \overline{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q}\right\|^h \quad (27)$$

Since

$$\overline{\Psi}_E \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \Psi_E \frac{\sum_{q\in \overline{E}}^{k} u_q l_q x_q}{\sum_{q=1}^{k} u_q l_q} = (1-\Psi_E)\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \Psi_E\left(\frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t} - \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{q=1}^{k} u_q l_q}\right)$$

$$= \frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{i=1}^{k} u_i l_i} - \Psi_E \frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}, \quad (28)$$

then we conclude (26)-(28) to get the related, however, more relevant lower bound.

$$M_h(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq 2^h \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{\left\|\frac{\sum_{i\in E}^{k} u_i l_i x_i}{\sum_{q=1}^{k} u_q l_q} - \Psi_E \frac{\sum_{t=1}^{k} u_t l_t x_t}{\sum_{t=1}^{k} u_t l_t}\right\|^h\right\} (\geq 0) \quad (29)$$

The following is an instance for 'useful' mean h −absolute deviation:

Corollary 4. Let a normed linear space $(\mathcal{T}, \|\cdot\|)$, and $u_i > 0$ are the utilities attached to probabilities. If $\boldsymbol{x} \in \mathcal{T}^k$ and $\boldsymbol{l} = (l_1, \dots, l_k)$ is a likelihood dispensation with all l_i are not zero, then we have for $h \geq 1$

$$M_h(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq \max_{t \in \{1,2,\dots,k\}} \left\{ [(1 - u_t l_t)^{1-h} . u_t^h l_t^h + u_t l_t] \left\| x_t - \frac{\sum_{m=1}^k u_m l_m z_m}{\sum_{m=1}^k u_m l_m} \right\|^h \right\} \tag{30}$$

Remark 5. Since the function is strictly increasing on [0,1) as well as $v_h(s) = (1 - s)^{1-h} s^h + s, h \geq 1, s \in [0,1)$, therefore

$$\min_{t \in \{1,2,\dots,k\}} \{(1 - u_t l_t)^{1-h} . u_t^h l_t^h + u_t l_t\} = u_a l_a + (1 - u_a l_a)^{1-h} . u_a^h l_a^h$$

Where $u_a l_a = \min_{t \in \{1,2,\dots,k\}} u_t l_t$, we get the next inequality by (30):

$$M_h(\boldsymbol{l}, \boldsymbol{x}; \mho) \geq [u_a l_a + (1 - u_a l_a)^{1-h} . u_a^h l_a^h] \max_{t \in \{1,2,\dots,k\}} \left\| x_t - \frac{\sum_{m=1}^k u_m l_m x_m}{\sum_{m=1}^k u_m l_m} \right\|^l \tag{31}$$

which may be better appropriate for usage (see also [12]).

4. Application for 'useful' mean g −divergence

If $u_i > 0$ are the utilities attached to probabilities and let $g: [0, \infty) \to \mathbb{R}$ be convex, then 'useful' mean g −divergence functional is

$$S_g(\boldsymbol{l}, \boldsymbol{r}; \mho) = \frac{\sum_{i=1}^k u_i r_i g\left(\frac{l_i}{r_i}\right)}{\sum_{i=1}^k u_i r_i} \tag{32}$$

A generalized measure of information was developed by Csiszar [7], a "distance function" on the set of likelihood dispensation $\mathbb{S}^k$. We defined a set of utility distributions $\mho^k$, let $\boldsymbol{l} = (l_1, l_2, \dots, l_k)$ and $\boldsymbol{r} = (r_1, r_2, \dots, r_k)$ are positive sequences and $\mho = (u_1, u_2, \dots \dots u_k)$. Undefined expressions are interpreted in the same way as in [7].

$$g(0) = \lim_{s \to 0^+} g(s), 0g\left(\frac{0}{0}\right) = 0,\ 0g\left(\frac{b}{0}\right) = \lim_{r \to 0^+} rg\left(\frac{b}{r}\right) = b \lim_{s \to \infty} \frac{g(s)}{s}, b > 0$$

Csiszar and Korner [8] were essentially stated the following results:

(i) $S_g(\boldsymbol{l}, \boldsymbol{r}; \mho)$ is jointly convex in $\boldsymbol{l}, \boldsymbol{r}$, if g is convex

(ii) We have, $\forall\ \boldsymbol{l}, \boldsymbol{r} \in T_+^k$

$$S_g(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq \frac{\sum_{q=1}^k u_q r_q g\left(\frac{\sum_{q=1}^k l_q}{\sum_{q=1}^k r_q}\right)}{\sum_{q=1}^k u_q r_q} \tag{33}$$

Equality holds in (33) if and only if $\frac{l_1}{r_1} = \frac{l_2}{r_2} = \dots = \frac{l_k}{r_k}$, if g is strictly convex.

We have the inequality, for every $\boldsymbol{l}, \boldsymbol{r} \in T_+^k$ and $u_i > 0$ are the utilities attached to probabilities with $\sum_{i=1}^{k} l_i = \sum_{i=1}^{k} r_i$, if $g(1) = 0$ i.e., g is normalized, then

$$S_g(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq 0. \tag{34}$$

If $\mathbf{l}, \boldsymbol{r} \in \mathbb{S}^k$ in a particular case, then (34) is valid. This is the g −divergence's well-known positivity characteristic.

As follows, we try to generalize this notion in linear space to the functions defined on a cone. To begin, we note that if the following two criteria are fulfilled, the subset N is a cone, in a linear space $\mathcal{T}$:

(i) we have $x + z \in N$, for any $x, z \in N$;

(ii) we have $\alpha x \in N$, for any $x \in N$ and any $\alpha \geq 0$

For the convex function $g: N \to \mathbb{R}$, we may define the following 'useful' g −divergence of $\boldsymbol{y}$ with $\boldsymbol{r}$ for a likelihood dispensation $\boldsymbol{r} \in \mathbb{S}^k$, and given $k-$ tuple of vectors $\boldsymbol{y} = (y_1, y_2, \dots, y_k) \in N^k$, $u_i > 0$ are the utilities attached to probabilities denoted by $\mho$ with all entries are not zero.

$$S_g(\boldsymbol{y}, \boldsymbol{r}; \mho) \geq \frac{\sum_{i=1}^{k} u_i r_i g\left(\frac{\sum_{q=1}^{k} y_i}{\sum_{q=1}^{k} r_i}\right)}{\sum_{i=1}^{k} u_i r_i} \tag{35}$$

If $\mathcal{T} = \mathbb{R}, \boldsymbol{x} = \boldsymbol{l} \in \mathbb{S}^k$, and $N = [0, \infty)$ then it is self-evident that we have the basic notion of the g −divergence connected with a function $g: [0, \infty) \to \mathbb{R}$.

Now, a likelihood dispensation $\boldsymbol{r} \in \mathbb{S}^k$ with all entries are not zero, for each nonvoid set E of $\{1,2, \dots, k\}$, for a given k −tuple of vectors $\boldsymbol{x} \in N^k$ and utility distributions $\mho \in \mho^k$, then

$\boldsymbol{r}_E = (K_E, \bar{K}_E) \in \mathbb{S}^2$, $\boldsymbol{x}_E = (\mathcal{T}_E, \bar{\mathcal{T}}_E) \in N^2$ and $\mho_E \in \mho^k$

It is obvious that

$$S_g(\boldsymbol{x}_E, \boldsymbol{r}_E; \mho_E) = \mho_E K_E g\left(\frac{\mathcal{T}_E}{K_E}\right) + \bar{\mho}_E \bar{K}_E g\left(\frac{\bar{\mathcal{T}}_E}{\bar{K}_E}\right).$$

Where $\mathcal{T}_E = \sum_{i \in E} x_i$, and $\bar{\mathcal{T}}_E = \mathcal{T}_{\bar{E}}$, as above. In a linear space, for the g −divergence of k −tuple of vectors the following inequality stands:

Theorem 3. On the cone N, let $g: N \to \mathbb{R}$ be a convex function. Then, for any nonempty subset E of $\{1,2, \dots, k\}$ and every k −tuple of vectors $\boldsymbol{x} \in N^k$, utility distributions $\mho \in \mho^k$ and a likelihood dispensation $\boldsymbol{r} \in \mathbb{S}^k$ with all values are not zero, we get

$$\begin{aligned} S_g(\boldsymbol{x}, \boldsymbol{r}; \mho) &\geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} S_g(\boldsymbol{x}_E, \boldsymbol{r}_E; \mho_E) \geq S_g(\boldsymbol{x}_E, \boldsymbol{r}_E; \mho_E) \\ &\geq \min_{\emptyset \neq E \subset \{1,2,\dots,k\}} S_g(\boldsymbol{x}_E, \boldsymbol{r}_E; \mho_E) \geq g(\mathcal{T}_k) \end{aligned} \tag{36}$$

Where $\mathcal{T}_k = \sum_{i=1}^{k} x_i$

Proof. The proof is the same as Theorem 1.

We conclude that $g(\mathcal{T}_k) \geq 0$ is a necessary stipulation for the validity of $S_g(\boldsymbol{x}, \boldsymbol{r}; \mho)$ about any likelihood dispensation $\boldsymbol{r} \in \mathbb{S}^k$ with all values are not zero for every k −tuple of vectors $\boldsymbol{x} \in N^k$. If $\boldsymbol{x} = \boldsymbol{l} \in \mathbb{S}^k$ in the scalar case, then $g(1) \geq 0$ is a necessary stipulation for the positiveness of the 'useful' g −divergence $S_g(\boldsymbol{l}, \boldsymbol{r}; \mho)$.

Corollary 5. Consider a normalized convex function to be $g: [0, \infty) \to \mathbb{R}$, and $u_i > 0$ are the utilities attached to probabilities. Therefore, we have for any $\boldsymbol{l}, \boldsymbol{r} \in \mathbb{S}^k$ and $\mho \in \mho^k$

$$S_g(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left[K_E g\left(\frac{\mho_E \Psi_E}{K_E}\right) + (1 - K_E) g\left(\frac{1-\mho_E \Psi_E}{1-K_E}\right)\right] \geq 0 \tag{37}$$

We have given lower bounds for several g −divergences utilized in Statistics, Probability Theory, including Information Theory in the following sections.

The convex function $g(s) = |s - 1|$, $s \in \mathbb{R}$ defines the 'useful' total variation distance, which is given in:

$$W(\boldsymbol{l}, \boldsymbol{r}; \mho) = \frac{\sum_{q=1}^k u_q r_q \left|\frac{u_q l_q}{u_q r_q} - 1\right|}{\sum_{q=1}^k u_q r_q} = \frac{\sum_{q=1}^k |u_q l_q - u_q r_q|}{\sum_{q=1}^k u_q r_q}. \tag{38}$$

For the total variation distance, the following improvement of the optimism disparity can be stated such as.

Proposition 1. For any $\boldsymbol{l}, \boldsymbol{r} \in \mathbb{S}^k$ and utility distributions $\mho \in \mho^k$, we have the inequality:

$$W(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq 2 \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} |\mho_E \Psi_E - K_E| \tag{39}$$

Proof. The inequality (37) for $g(s) = |s - 1|$, $s \in \mathbb{R}$ completes the proof.

For the convex function $g(s) = (s - 1)^2$, $s \in \mathbb{R}$, then 'useful' K. Pearson χ^2 −divergence is acquired and given by

$$\chi^2(\boldsymbol{l}, \boldsymbol{r}; \mho) = \frac{\sum_{q=1}^k u_q r_q \left(\frac{u_q l_q}{u_q r_q} - 1\right)^2}{\sum_{q=1}^k u_q r_q} = \frac{\sum_{q=1}^k \left(u_q l_q - u_q r_q\right)^2}{(u_q r_q) \sum_{q=1}^k u_q r_q} \tag{40}$$

Proposition 2. For any $\boldsymbol{l}, \boldsymbol{r} \in \mathbb{S}^k$ and utility distributions $\mho \in \mho^k$, we have the inequality:

$$\chi^2(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{\frac{(\mho_E \Psi_E - K_E)^2}{K_E(1 - K_E)}\right\}$$

$$\geq 4 \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} (\mho_E \Psi_E - K_E)^2 \ (\geq 0) \tag{41}$$

Proof. For the function $g(s) = (s - 1)^2$, $s \in \mathbb{R}$ on applying the inequality (37), we get

$$\chi^2(\boldsymbol{l}, \boldsymbol{r}; \mho) \geq \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{(1 - K_E) g\left(\frac{1-\mho_E \Psi_E}{1-K_E} - 1\right)^2 + K_E \left(\frac{\mho_E \Psi_E}{K_E} - 1\right)^2\right\}$$

$$= \max_{\emptyset \neq E \subset \{1,2,\dots,k\}} \left\{\frac{(\mho_E \Psi_E - K_E)^2}{K_E(1 - K_E)}\right\}.$$

Since

$$K_E(1 - K_E) \leq \frac{1}{4}[K_E + (1 - K_E]^2 = \frac{1}{4}$$

then

$$\frac{(\mho_E \Psi_E - K_E)^2}{K_E(1 - K_E)} \geq 4(\mho_E \Psi_E - K_E)^2$$

for each $E \subset \{1,2, \dots, k\}$, that establishes the last portion of (41).

For $g: (0, \infty) \to \mathbb{R}$, $g(s) = s \ln s$, the 'useful' Kullback-Leibler (KL-divergence) and denote it by $D(\boldsymbol{l}, \boldsymbol{r}; \mho)$ may be calculated as follows:

$$D(\boldsymbol{l},\boldsymbol{r};\mho)=\frac{\sum_{q=1}^{k} r_q.\frac{u_q l_q}{r_q}\ln\left(\frac{u_q l_q}{r_q}\right)}{\sum_{q=1}^{k} u_q r_q}=\frac{\sum_{q=1}^{k} u_q l_q\ln\left(\frac{u_q l_q}{r_q}\right)}{\sum_{q=1}^{k} u_q r_q} \tag{42}$$

Proposition 3. We have the inequality, for any $\boldsymbol{l},\boldsymbol{r}\in\mathbb{S}^k$ and utility distributions $\mho\in\mho^k$:

$$D(\boldsymbol{l},\boldsymbol{r};\mho)\geq\left[\max_{\emptyset\neq E\subset\{1,2,\dots,k\}}\left\{\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)^{1-\Psi_E}.\left(\frac{\mho_E\Psi_E}{K_E}\right)^{\mho_E\Psi_E}\right\}\right]\geq 0 \tag{43}$$

Proof. Corollary 5 demonstrates the first inequality by utilizing the harmonic mean and geometric mean inequality,

$$w^{\gamma}x^{1-\gamma}\geq\frac{1}{\frac{\gamma}{x}+\frac{1-\gamma}{z}},\ w,x>0,\gamma\in[0,1]$$

for $w=\frac{\mho_E\Psi_E}{K_E}$, $x=\frac{1-\mho_E\Psi_E}{1-K_E}$ and $\gamma=\mho_E\Psi_E$, we have

$$\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)^{1-\mho_E\Psi_E}.K_E\left(\frac{\mho_E\Psi_E}{K_E}\right)^{\mho_E\Psi_E}\geq 1$$

for any $E\subset\{1,2,\dots,k\}$, implying that the second portion of (43).

The 'useful' Jeffreys divergence (UJD) is another important divergence measure in Information Theory, for the function $g(s)=s\ln s, s>0$.

$$UJD(\boldsymbol{l},\boldsymbol{r};\mho)=\frac{\sum_{q=1}^{k} r_q.\left(\frac{u_q l_q}{r_q}-1\right)\ln\left(\frac{u_q l_q}{r_q}\right)}{\sum_{q=1}^{k} u_q r_q}=\frac{\sum_{q=1}^{k}(u_q l_q-r_q)\ln\left(\frac{u_q l_q}{r_q}\right)}{\sum_{q=1}^{k} u_q r_q}, \tag{44}$$

Proposition 4. For any $\boldsymbol{l},\boldsymbol{r}\in\mathbb{S}^k$ and utility distributions $\mho\in\mho^k$, then we have the inequality:

$$UJD(\boldsymbol{l},\boldsymbol{r};\mho)\geq\ln\left(\max_{\emptyset\neq E\subset\{1,2,\dots,k\}}\left\{\left[\frac{1-\mho_E\Psi_E}{1-K_E}\right]^{(K_E-\mho_E\Psi_E)}\right\}\right)$$

$$\geq\max_{\emptyset\neq E\subset\{1,2,\dots,k\}}\left[\frac{(K_E-\mho_E\Psi_E)^2}{K_E+\mho_E\Psi_E-2K_E.\mho_E\Psi_E}\right]\geq 0 \tag{45}$$

Proof. Using the inequality (37) for $g(s)=(s-1)\ln s$, we arrive at

$$UJD(\boldsymbol{l},\boldsymbol{r};\mho)\geq\max_{t\in\{1,2,\dots,k\}}\left\{(1-K_E)\left[\left(\frac{1-\mho_E\Psi_E}{1-K_E}-1\right)\ln\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)\right]\right.$$
$$\left.+K_E\left(\frac{\mho_E\Psi_E}{K_E}-1\right)\ln\left(\frac{\mho_E\Psi_E}{K_E}\right)\right\}$$
$$=\max_{t\in\{1,2,\dots,k\}}\left\{(K_E-\mho_E\Psi_E)\ln\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)-(K_E-\mho_E\Psi_E)\ln\left(\frac{\mho_E\Psi_E}{K_E}\right)\right\}$$
$$=\max_{t\in\{1,2,\dots,k\}}\left\{(K_E-\mho_E\Psi_E)\ln\left[\frac{(1-\mho_E\Psi_E)K_E}{(1-K_E)\mho_E\Psi_E}\right]\right\},$$

We use the basic disparity with positive numbers as a starting point, for proving the first inequality in (45).

$$\frac{\ln c-\ln d}{c-d}\geq\frac{2}{c+d}, c,d>0$$

We have

$$(K_E-\mho_E\Psi_E)\left[\ln\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)-\ln\left(\frac{\mho_E\Psi_E}{K_E}\right)\right]$$

$$= (K_E - \mho_E \Psi_E).\frac{\ln\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)-\ln\left(\frac{\mho_E\Psi_E}{K_E}\right)}{\frac{1-\mho_E\Psi_E}{1-K_E}-\frac{\mho_E\Psi_E}{K_E}}.\left[\frac{1-\mho_E\Psi_E}{1-K_E}-\frac{\mho_E\Psi_E}{K_E}\right]$$

$$= \frac{(K_E-\mho_E\Psi_E)^2}{K_E(1-K_E)}.\frac{\ln\left(\frac{1-\mho_E\Psi_E}{1-K_E}\right)-\ln\left(\frac{\mho_E\Psi_E}{K_E}\right)}{\frac{1-\mho_E\Psi_E}{1-K_E}-\frac{\mho_E\Psi_E}{K_E}}$$

$$\geq \frac{(K_E-\mho_E\Psi_E)^2}{K_E(1-K_E)}.\frac{2}{\frac{1-\mho_E\Psi_E}{1-K_E}-\frac{\mho_E\Psi_E}{K_E}} = \frac{2(K_E-\mho_E\Psi_E)^2}{K_E+\mho_E\Psi_E-2K_E.\mho_E\Psi_E} \geq 0$$

Given the second inequality in (45), for each $E \subset \{1,2,\dots,k\}$.

5.Conclusion

In both theory and practice, the classical Jensen's inequality is extremely essential. Using the generalized functional, we were able to refine Jensen's inequality (10) – (21) in real linear space. In addition, we discovered new and sharp bounds of Shannon's entropy as well as various $g-$divergence metrics in information theory. We will continue to investigate potential applications of the newly discovered inequalities in future research.

References

1. Adil Khan, M., G. Ali Khan, T. Ali, A. Kilicman. On the refinement of Jensen"s inequality,Applied Mathematics and Computation 2015; 262: 128–135.
2. Bakula, M.K., J. Pecaric. On the Jensen's inequality for convex functions on the coordinates in a rectangle from the plane, Taiwanese J. Math. 2006; 10: 1271-1292.
3. Bakula, M.K., J. Pecaric, J. Peric. On the converse Jensen's inequality, Appl. Math. Comput. 2012; 218: 6566-6575.
4. Beesack, P.R., Pecaric. On Jessen's inequality for convex functions, J. Math. Anal. Appl. 1985; 110,536-552.
5. Belis, M. and S. Guiasu. A quantitative-qualitative measure of information in Cybernetics System, IEEE Trans. Inform. Theory, IT 1968; 14: 593-594.
6. Chandy, T.W. And J.B. Mcliod. On a functional equation, Proc. Edinburgh Maths 1960; 43, 7-8.
7. Csiszar, I., Information-type measures of differences of probability distributions and indirect observations, Studia Sci. Math. Hung. 1967; 2: 229-318.
8. Csiszar, I., J. Korner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Academic Press, New York, 1981.
9. Dwivedi P.P., Sharma D.K. Lower and Upper Bounds for 'Useful' Renyi Information Rate. Advances in Computing and Data Sciences. ICACDS 2021. Communications in Computer and Information Science, vol 1441. Springer, Cham. 2021, doi:10.1007/978-3-030-88244-0_26
10. Dragomir, S.S. A further improvement of Jensen's inequality.' Tamkang J. Math. 1994; 25 (1): 29-36.
11. Dragomir, S.S. A new refinement of Jensen's inequality in linear spaces with applications, Math. Comput. Model. 2010; 52: 1497-1505.

12. Dragomir, S.S. A refinement of Jensen's inequality with applications for f-divergence measures, Taiwanese J. Math. 2010; 14 (1): 153-164.
13. Dragomir, S. & Khan, Muhammad & Abathun, Addisalem. Refinement of the Jensen integral inequality. Open Math. 2016; 14: 221–228. doi:14. 10.1515/math-2016-0020.
14. Fadeev, D.K. On the concept of entropies of finite probabilistic scheme, (Russian) Uspchi Math. Nauk, 1956; 11: 227-231.
15. JiLe G., H. Rashid., Farooq F., and Sultana, S. Some Inequalities for a New Class of Convex Functions with Applications via Local Fractional Integral, Journal of Function Spaces, vol. 2021; 2021: https://doi.org/10.1155/2021/6663971.
16. Kendall, D.G. Functional equations in information theory, Z.Wahrs Verw.Geb, 1964; 2: 225-229.
17. Khinchin, A.I. Mathematical Foundations of Information Theory, Dover Publications, New York 1957.
18. Klaricic Bakula, M., & Nikodem, K. On the converse Jensen inequality for strongly convex functions. Journal of Mathematical Analysis and Applications, 2016; 434(1): 516-522. doi: 10.1016/j.jmaa.2015.09.032.
19. Kluza, P. A. On Jensen-Renyi and Jeffreys-Renyi type f-divergences induced by convex functions. Physica A: Statistical Mechanics and Its Applications, 2019; 122527. doi:10.1016/j.physa.2019.122527.
20. Kumari, R. and Shrma D. K. Generalized 'useful' non-symmetric divergence measures and inequalities, Journal of mathematical inequalities, 2019; 13 (2): 451–466, doi.org/10.7153/jmi-2019-13-30.
21. Micic, J., J. Pecaric, P. Jurica. Refined Jensen's operator inequality with condition on spectra, Oper. Matrices 2013; 7 (2): 293-308.
22. Pecaric, J., S.S. Dragomir. A refinements of Jensen inequality and applications, Stud. Univ. Babe-Bolyai Math. 1989; 24 (1): 15-19.
23. Rashid S., T. Abdeljawad, F. Jarad, and M. A. Noor. Some estimates for generalized Riemann-Liouville fractional integrals of exponentially convex functions and their applications, Mathematics, 2019; 7(9): 807. doi:10.3390/math7090807.
24. Sayyari, Y. New entropy bounds via uniformly convex functions. Chaos, Solitons & Fractals, 2020; 141: 110360. doi:10.1016/j.chaos.2020.110360.
25. Shannon, C.E. A mathematical theory of communication, Bell System Technical Journal 1948; 27: 379–423(Part I) 623–656(Part II).
26. Simic, S. Best possible global bounds for Jensen's inequality, Appl. Math. Comput. 2009; 215 (6): 2224–2228. doi:10.1016/j.amc.2009.08.062.
27. Tapus, N., P.G. Popescu. A new entropy upper bound, Appl. Math. Lett. 2012; 25: 1887-1890.
28. Tverberg, H. A new derivation of the information function, Math. Scand, 1958; 6: 297-298.