

Estimations of divergence measures for majorization inequalities via Peano's representation of Hermite's polynomial

Awais Rasheed

Department of Mathematics, Govt. Associate College Sillanwali Sargodha, Sargodha, Pakistan

Khuram Ali Khan

Department of Mathematics, University of Sargodha, Sargodha, Pakistan

Josip Pečarić

Croatian Academy of Sciences and Arts, Zagreb, Croatia, and

Đilda Pečarić

Department of Communication, Media and Journalism, University North, Koprivnica, Croatia

Abstract

Purpose – In this paper applications to information theory are presented. Generalized majorization theorem is presented in term of different entropies and divergences. So that obtained results are generalized and comprehensive.

Design/methodology/approach – The analytical approach is used to prove the results. i.e. Different n -convex functions are used to prove results. Generalized majorization theorem and Peanos's representation of Hermite polynomial are also imply.

Findings – The results related to Csizsár f divergence using generalized majorization theorem via Peanos's representation of Hermite polynomial are proved. In seek of applications, generalized results are established using Shannon entropy, Kullback-Leibler distance. Moreover, the unique cases of obtained results are derived in terms of Bhattacharyya coefficient, Jeffrey's distance and Triangular discrimination.

Originality/value – The results are new and never submitted before.

Keywords Csizsár f -divergence, Shannon entropy, Kullback–Leibler distance, Hermite polynomial

Paper type Research article

1. Introduction and preliminaries

The divergence measure is a measure that is used to compare the difference of two probability distributions. Many problems in probability theory are resolved using the concept of divergence measure. In literature, different types of divergence measures are used to calculate two probability distributions. Sensor networks [1], examining the order in a Markov chain [2], finance [3], economics [4], and approximation of probability distributions [5] are all areas where information and divergence measures are highly valuable and play a significant role.

© Awais Rasheed, Khuram Ali Khan, Josip Pečarić and Đilda Pečarić. Published in *Arab Journal of Mathematical Sciences*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at [Link to the terms of the CC BY 4.0 licence](#).

Funding: There is no funding for this work.

Competing interests: The authors declare that there is no conflict of interests regarding the publication of this paper.



Data transmission, quantification, and storage are all topics covered in the newly developed scientific field of information theory. Shannon [6] introduced the idea of information theory and identified entropy to be the fundamental variable for information and also defined the premise of information theory. A large class of divergences known as Csiszár's f -divergence is used to compare two probability densities. One remarkable aspect of Csiszár's f -divergence is that by substituting the appropriate values for the convex function f , various known divergence measures may be derived from it. Several scholars have put a lot of effort into defining bounds on divergences and distances in recent years; for examples, see [7–9]. Siddique *et al.* [10] generalized the majorization-type inequalities by employing Green's functions and Fink's identity. The obtained results are used in relation to Shannon entropy and Kullback–Leibler divergence.

In Niaz *et al.* [11], used Taylor's formula to estimate various entropies by applying Jensen-type functionals. Latif *et al.* [12] used the polynomial of Taylor in association with recently defined Green's functions to conclude majorization inequality. As majorization theory is often used in a wide range of areas, Khan *et al.* [13] explored a lot of results on majorization as well as gave its applications in information theory. For higher convex functions, Adeel *et al.* [23] derived various types of inequalities by applying Hermite interpolating polynomial and Green functions and also stated inequalities for the Shannon entropy and f -divergence. In Mehmood *et al.* [24], explained discrete and continuous cyclic refinements of Jensen's inequality and modified these for higher order convex function by applying Hermite polynomial whose error term is approximated using Peano's kernel. In Rasheed *et al.* [14], used Hermite interpolation to find the bounds for Shannon entropy and f -divergence. Rasheed *et al.* [15] defined novel Green functions to present the bounds for Shannon entropy and f -divergence. They also proved various inequalities involving Green functions for the class of higher-order convex functions. In Bilal *et al.* [16], presented a generalized form of Shannon-type inequalities via diamond integrals. In Bilal *et al.* [17], introduced Csiszár's f -divergence for diamond integrals. They also presented inequalities associated with Csiszár's f -divergence involving diamond integrals.

In [18], majorization is defined as

Definition 1.1. For fixed $m \geq 2$. Consider two non-increasing real-number sequences $\mathbf{x} = (\xi_1, \dots, \xi_m)$, $\mathbf{y} = (\eta_1, \dots, \eta_m)$. Then \mathbf{x} majorizes \mathbf{y} , symbolically we have $\mathbf{x} > \mathbf{y}$, if we have

$$\sum_{l=1}^{\rho} \xi_l \geq \sum_{l=1}^{\rho} \eta_l, \quad 1 \leq \rho \leq m - 1, \tag{1}$$

and

$$\sum_{l=1}^m \eta_l = \sum_{l=1}^m \xi_l. \tag{2}$$

The following theorem is given by Marshall–Olkin–Arnold [18, 19], which is also called Classical Majorization Theorem.

Theorem 1.1. (Classical Majorization Theorem) Consider two non-increasing real m -tuples $\mathbf{x} = (\xi_1, \dots, \xi_m)$, $\mathbf{y} = (\eta_1, \dots, \eta_m)$ such that $\xi_j, \eta_j \in [\zeta_1, \zeta_2]$ for $j = 1, \dots, m$. Then \mathbf{y} is majorized by \mathbf{x} if and only if the subsequent inequality is true, for any convex, continuous function, $f : [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$,

$$\sum_{l=1}^m f(\eta_l) \leq \sum_{l=1}^m f(\xi_l). \tag{3}$$

The Weighted Majorization Theorem [18, 20] has the following form:

4

Theorem 1.2. (Weighted Majorization Theorem) Consider two non-increasing real p -tuples $\mathbf{x} = (\xi_1, \dots, \xi_p)$, $\mathbf{y} = (\eta_1, \dots, \eta_p)$ such that $\xi_l, \eta_l \in [\zeta_1, \zeta_2]$ for $l = 1, \dots, p$. Let $\mathbf{q} = (q_1, \dots, q_p)$ be real p -tuple such that

$$\sum_{l=1}^j q_l \xi_l \geq \sum_{l=1}^j q_l \eta_l, \quad j \in \{1, \dots, p-1\} \tag{4}$$

and

$$\sum_{l=1}^p q_l \xi_l = \sum_{l=1}^p q_l \eta_l. \tag{5}$$

Then

$$\sum_{l=1}^p q_l f(\eta_l) \leq \sum_{l=1}^p q_l f(\xi_l) \tag{6}$$

is true for each convex, continuous function $f : [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$.

Remark 1.1. If all the assumptions of [Theorem 1.2](#) are satisfied, then

$$\mathfrak{M}(\mathbf{x}, \mathbf{y}, \mathbf{q}, f(\cdot)) := \sum_{i=1}^p q_i f(\xi_i) - \sum_{i=1}^p q_i f(\eta_i) \geq 0. \tag{7}$$

Let $[\zeta_1, \zeta_2] \subset (-\infty, \infty)$ and $\lambda \in \{1, 2, 3, 4\}$. In, Pečarić *et al.* [21] gave a new type of Green functions, $G_\lambda : [\zeta_1, \zeta_2] \times [\zeta_1, \zeta_2] \rightarrow \mathbb{R}$, which are given as:

$$G_1(w, v) = \begin{cases} \zeta_1 - v, & \zeta_1 \leq v \leq w, \\ \zeta_1 - w, & w \leq v \leq \zeta_2. \end{cases} \tag{8}$$

$$G_2(w, v) = \begin{cases} w - \zeta_2, & \zeta_1 \leq v \leq w, \\ v - \zeta_2, & w \leq v \leq \zeta_2. \end{cases} \tag{9}$$

$$G_3(w, v) = \begin{cases} w - \zeta_1, & \zeta_1 \leq v \leq w, \\ v - \zeta_1, & w \leq v \leq \zeta_2. \end{cases} \tag{10}$$

$$G_4(w, v) = \begin{cases} \zeta_2 - v, & \zeta_1 \leq v \leq w, \\ \zeta_2 - w, & w \leq v \leq \zeta_2. \end{cases} \quad (11)$$

For $\theta < \phi$. Let $\theta, \phi \in (-\infty, \infty)$ and $\phi \geq a_r > \dots > a_2 > a_1 \geq \theta$, ($r \geq 2$). A polynomial $p_H(s)$ of degree $(n - 1)$ exists for $f \in C^n[\theta, \phi]$, if any of the criteria below are met.

Hermite conditions:

$$f^{(i)}(a_j) = p_H^{(i)}(a_j); \quad k_j \geq i \geq 0, \quad r \geq j \geq 1, \quad \sum_{j=1}^r k_j + r = n.$$

Hermite conditions contain the following specific cases, which are of major interest:

Type $(m, n - m)$ conditions:

($k_1 = m - 1, k_2 = n - m - 1, n - 1 \geq m \geq 1, r = 2$)

$$\begin{aligned} p_{(m,n)}^{(i)}(\theta) &= f^{(i)}(\theta), \quad i \in \{0, 1, \dots, m - 1\}, \\ p_{(m,n)}^{(i)}(\phi) &= f^{(i)}(\phi), \quad i \in \{0, 1, \dots, n - m - 1\}, \end{aligned}$$

Two-point Taylor conditions:($r = 2, n = 2m, k_1 = k_2 = m - 1$)

$$f^{(i)}(\theta) = p_{2T}^{(i)}(\theta), \quad f^{(i)}(\phi) = p_{2T}^{(i)}(\phi), \quad i \in \{0, 1, \dots, m - 1\}.$$

In [22], the following result is given.

Theorem 1.3. Suppose $\theta, \phi \in (-\infty, \infty)$ and $\theta \leq a_1 < a_2, \dots \leq \phi$, ($r \geq 2$). If $f \in C^n([\theta, \phi])$ then

$$f(\gamma) = \rho_H(\gamma) + R_{H,n}(f, \gamma)$$

where

$$\rho_H(\gamma) = \sum_{j=1}^r \sum_{i=0}^{k_j} f^{(i)}(a_j) H_{ij}(\gamma)$$

is the Hermite interpolating polynomial.

The H_{ij} are fundamental Hermite basis given by

$$H_{ij}(\gamma) = \frac{w(\gamma)(\gamma - a_j)^{i-k_j-1}}{i!} \sum_{k=0}^{k_j-i} \frac{1}{k!} \frac{d^k}{d\gamma^k} \left(\frac{(\gamma - a_j)^{k_j+1}}{w(\gamma)} \right) \Big|_{\gamma=a_j} (\gamma - a_j)^k, \quad (12)$$

$$w(\gamma) = \prod_{j=1}^r (\gamma - a_j)^{k_j+1}, \quad (13)$$

and

$$R_{H,n}(f, \gamma) = \int_{\theta}^{\phi} G_{H,n}(\gamma, s) f^n(s) ds$$

6

is remainder. For all $a_{l+1} \geq s \geq a_l; l \in \{0, \dots, r\}$ with $\theta = a_0$ and $\phi = a_{r+1}$, $G_{H,n}(\gamma, s)$ (Peano's kernel) is written as

$$G_{H,n}(\gamma, s) = \begin{cases} \sum_{j=1}^l \sum_{i=0}^{k_j} \frac{(a_j - s)^{n-i-1}}{(n-i-1)!} H_{ij}(\gamma); & s \leq \gamma, \\ - \sum_{j=l+1}^r \sum_{i=0}^{k_j} \frac{(a_j - s)^{n-i-1}}{(n-i-1)!} H_{ij}(\gamma); & s \geq \gamma, \end{cases} \quad (14)$$

$(m, n - m)$ interpolating polynomial $\rho_{m,n}$ is given by

$$\rho_{m,n}(\gamma) = \sum_{i=0}^{m-1} \tau_i(\gamma) f^{(i)}(\theta) + \sum_{i=0}^{n-m-1} \nu_i(\gamma) f^{(i)}(\phi),$$

with

$$\tau_i(\gamma) = \frac{1}{i!} (\gamma - \theta)^i \left(\frac{\gamma - \phi}{\theta - \phi} \right)^{n-m-1-i} \sum_{k=0}^{m-1-i} \binom{n-m+k-1}{k} \left(\frac{\gamma - \theta}{\phi - \theta} \right)^k \quad (15)$$

and

$$\nu_i(\gamma) = \frac{1}{i!} (\gamma - \phi)^i \left(\frac{\gamma - \phi}{\phi - \theta} \right)^{m-n-1-i} \sum_{k=0}^{m-1-i} \binom{m+k-1}{k} \left(\frac{\gamma - \phi}{\theta - \phi} \right)^k. \quad (16)$$

The interpolating polynomial of Taylor for two points $\rho_{2T}(t)$ is defined by

$$\rho_{2T}(\gamma) = \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{m+k-1}{k} \left[\frac{(\gamma - \theta)^i}{i!} \left(\frac{\gamma - \phi}{\theta - \phi} \right)^m \left(\frac{\gamma - \theta}{\phi - \theta} \right)^k f^{(i)}(\theta) + \frac{(\gamma - \phi)^i}{i!} \left(\frac{\gamma - \theta}{\phi - \theta} \right)^m \left(\frac{\gamma - \phi}{\theta - \phi} \right)^k f^{(i)}(\phi) \right]. \quad (17)$$

Throughout this paper, the following assumptions are taken as:

A: Let $[\theta, \phi] \subset (-\infty, \infty)$ and $f : [\theta, \phi] \rightarrow \mathbb{R}$ be n -convex function. Consider $\mathbf{q} = (q_1, \dots, q_p)$ be p -tuple and $\mathbf{x} = (\xi_1, \dots, \xi_p), \mathbf{y} = (\eta_1, \dots, \eta_p)$ be non-increasing p -tuples such that $\xi_i, \eta_i \in [\theta, \phi]$ and $q_i \in \mathbb{R}(i = 1, \dots, p)$, which satisfy conditions (4) and (5).

B: Assume that

$$\sum_{i=1}^{\lambda} r_i \leq \sum_{i=1}^{\lambda} s_i, \quad 1 \leq \lambda \leq m-1 \tag{18}$$

and

$$\sum_{i=1}^m r_i = \sum_{i=1}^m s_i, \tag{19}$$

with $\frac{s_i}{w_i}, \frac{r_i}{w_i} \in [\theta, \phi] (i = 1, \dots, p)$. Let $\frac{s}{w}$ and $\frac{r}{w}$ are decreasing and $G_\lambda (\lambda = 1, 2, 3, 4)$ be as defined in (8)-(11).

Throughout this paper, b denotes the base of log function.

In Latif *et al.* [25], give the following weighted majorization inequality by using $(m, n - m)$ conditions.

Theorem 1.4. Let A hold. Consider τ_i, ν_i and $G_\lambda (\lambda = 1, 2, 3, 4)$ as defined in (15), (16) and (8)-(11), respectively. The following inequality exists if $n - m$ is even:

$$\begin{aligned} \mathfrak{R}(\mathbf{x}, \mathbf{y}, \mathbf{q}, f(\cdot)) &\geq \int_{\theta}^{\phi} \mathfrak{R}(\mathbf{x}, \mathbf{y}, \mathbf{q}, G_\lambda(\cdot, \gamma)) \\ &\times \left(\sum_{i=0}^{m-1} f^{(i+2)}(\theta) \tau_i(\gamma) + \sum_{i=0}^{n-m-1} f^{(i+2)}(\phi) \nu_i(\gamma) \right) d\gamma. \end{aligned} \tag{20}$$

In Latif *et al.* [25], give the following weighted majorization inequality by using two-point Taylor conditions.

Theorem 1.5. Let A hold and $G_\lambda (\lambda = 1, 2, 3, 4)$ be as defined in (8)-(11). If m is even, then

$$\begin{aligned} \mathfrak{R}(\mathbf{x}, \mathbf{y}, \mathbf{q}, f(\cdot)) &\geq \int_{\theta}^{\phi} \mathfrak{R}(\mathbf{x}, \mathbf{y}, \mathbf{q}, G_\lambda(\cdot, t)) \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{m+k-1}{k} \\ &\times \left[\frac{(\gamma - \phi)^i}{i!} \left(\frac{\gamma - \theta}{\phi - \theta} \right)^m \left(\frac{\gamma - \phi}{\theta - \phi} \right)^k f^{(i+2)}(\phi) \right. \\ &\left. + \frac{(\gamma - \theta)^i}{i!} \left(\frac{\gamma - \phi}{\theta - \phi} \right)^m \left(\frac{\gamma - \theta}{\phi - \theta} \right)^k f^{(i+2)}(\theta) \right] d\gamma. \end{aligned} \tag{21}$$

hold.

2. Main results

In this part, the generalization of [Theorem 1.5](#) are given in the sense of information theory. Firstly, Csizsar f -divergence is used to obtain new extended and modified results for generalized majorization theorem via Hermite interpolation. Then, by choosing a suitable higher-order convex function, bounds for Shannon entropy are estimated. Furthermore, the new results associated with Kullback–Leibler (K-L) divergence, the Bhattacharyya coefficient, Jeffrey’s distance and triangular discrimination are proved by utilizing Hermite interpolation.

2.1 Csiszár divergence

The inequalities provided in Theorems 1.4 and 1.5 are connected with Csiszár f -divergence in this subsection. The results of these sections are used to present various information theory results.

In Csiszár [26, 27], defined the f -divergence functional as: Consider a convex function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. If the probability distributions are $\mathbf{r}, \mathbf{w} \in \mathbb{R}_+^m$ then the f -divergence functional is given by

$$I_f(\mathbf{r}, \mathbf{w}) := \sum_{i=1}^m w_i f\left(\frac{r_i}{w_i}\right). \tag{22}$$

Horváth *et al.* [28] gave an extended form of (22) i.e.,

Let $J \subset \mathbb{R}$ and $f : J \rightarrow \mathbb{R}$ be a convex function. If $\mathbf{r}, \mathbf{w} \in \mathbb{R}_+^p$ is p -tuples such that $\frac{r_i}{w_i} \in J$, $i = 1, \dots, p$. Then

$$\widehat{I}_f(\mathbf{r}, \mathbf{w}) := \sum_{i=1}^p w_i f\left(\frac{r_i}{w_i}\right). \tag{23}$$

Theorem 2.1. Let A and B hold. If $n - m$ is even, then we have the following inequalities:

$$\begin{aligned} \widehat{I}_f(\mathbf{s}, \mathbf{w}) &\geq \widehat{I}_f(\mathbf{r}, \mathbf{w}) \\ &+ \int_{\theta}^{\phi} \left\{ \widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right\} \\ &\times \left[\sum_{i=0}^{m-1} \tau_i(t) f^{(i+2)}(\theta) + \sum_{i=0}^{n-m-1} \nu_i(\gamma) f^{(i+2)}(\phi) \right] d\gamma \end{aligned} \tag{24}$$

and

$$\begin{aligned} \widehat{I}_f(\mathbf{s}, \mathbf{w}) &\geq \widehat{I}_f(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \\ &\times \left[f^{(i+2)}(\phi) (\gamma - \phi)^i \left(\frac{\gamma - \theta}{\phi - \theta} \right)^m \left(\frac{\gamma - \phi}{\theta - \phi} \right)^k \right. \\ &\left. + (\gamma - \theta)^i \left(\frac{\gamma - \phi}{\theta - \phi} \right)^m \left(\frac{\gamma - \theta}{\phi - \theta} \right)^k f^{(i+2)}(\theta) \right] \frac{1}{i!} d\gamma. \end{aligned} \tag{25}$$

Where

$$\widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) = \sum_{i=1}^p w_i G_\lambda\left(\frac{s_i}{w_i}, \gamma\right)$$

and τ_i, ν_i are defined in (15) and (16).

Proof. Taking $\xi_i = \frac{s_i}{w_i}, \eta_i = \frac{r_i}{w_i}$ and $q_i = w_i > 0 (i = 1, \dots, p)$, then (7) becomes

$$\begin{aligned} \mathfrak{R}\left(\frac{\mathbf{s}}{\mathbf{w}}, \frac{\mathbf{r}}{\mathbf{w}}, \mathbf{w}, f(\cdot)\right) &:= \sum_{i=1}^p w_i f\left(\frac{s_i}{w_i}\right) - \sum_{i=1}^p w_i f\left(\frac{r_i}{w_i}\right) \\ &= \widehat{I}_f(\mathbf{s}, \mathbf{w}) - \widehat{I}_f(\mathbf{r}, \mathbf{w}). \end{aligned} \quad (26)$$

Similarly,

$$\mathfrak{R}\left(\frac{\mathbf{s}}{\mathbf{w}}, \frac{\mathbf{r}}{\mathbf{w}}, \mathbf{w} G_\lambda(\cdot, \gamma)\right) = \widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \quad (27)$$

Also, the constraints (18) and (19) imply (4) and (5), respectively. Therefore, using (26) and (27), we obtain (24) and (25), respectively. \square

Theorem 2.2. Let g be a mapping from $[\theta, \phi]$ to \mathbb{R} . If all the premises of Theorem 2.1 are true and $f(v) = vg(v), v \in [\theta, \phi]$, then

$$\begin{aligned} \check{I}_g(\mathbf{s}, \mathbf{w}) &:= \sum_{i=1}^p s_i g\left(\frac{s_i}{w_i}\right) \geq \check{I}_g(\mathbf{r}, \mathbf{w}) \\ &+ \int_{\theta}^{\phi} \left\{ \widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right\} \times \left(\sum_{i=0}^{m-1} \tau_i(\gamma) (xg)^{(i+2)}(\theta) + \sum_{i=0}^{n-m-1} \nu_i(\gamma) (xg)^{(i+2)}(\phi) \right) d\gamma. \end{aligned} \quad (28)$$

and

$$\begin{aligned} \check{I}_g(\mathbf{s}, \mathbf{w}) &:= \sum_{i=1}^p s_i g\left(\frac{s_i}{w_i}\right) \geq \check{I}_g(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left\{ \widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right\} \\ &\times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\phi}{\theta-\phi}\right)^m \left(\frac{\gamma-\theta}{\phi-\theta}\right)^k \right] \\ &\times \left[(xg)^{(i+2)}(\theta) + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta}\right)^m \left(\frac{\gamma-\phi}{\theta-\phi}\right)^k (xg)^{(i+2)}(\phi) \right] d\gamma. \end{aligned} \quad (29)$$

Proof. Putting $f(x) = xg(x)$ in Theorem 2.1, (28) and (29) are obtained, respectively. \square

2.2 Shannon entropy

In this subsection results related to Shannon entropy are presented. Majorization theory and entropic disorder measurement theory are closely related theories. Special cases of entropic inequalities associated with majorization are then presented.

In [29], for positive probability distribution $\mathbf{k} = (k_1, \dots, k_n)$, the Shannon entropy is given by

$$H(\mathbf{k}) := - \sum_{\rho=1}^n k_{\rho} \log(k_{\rho}).$$

Corollary 2.1. Consider A, B holds and log has a base greater than or equal to 1. If $n > 3$ is even, then

$$\begin{aligned} H(\mathbf{w}) \leq & \sum_{i=1}^p w_i \log\left(\frac{r_i}{w_i}\right) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_{\lambda}}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_{\lambda}}(\mathbf{s}, \mathbf{w}, \gamma) \right) \\ & \times \frac{(-1)^{i+2} (i+1)!}{\ln b} \left(\sum_{i=0}^{m-1} \theta^{-(i+2)} \tau_i(\gamma) + \sum_{i=0}^{n-m-1} \phi^{-(i+2)} \nu_i(\gamma) \right) d\gamma, \end{aligned} \tag{30}$$

and

$$\begin{aligned} H(\mathbf{w}) \leq & \sum_{i=1}^p w_i \log\left(\frac{r_i}{w_i}\right) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_{\lambda}}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_{\lambda}}(\mathbf{s}, \mathbf{w}, \gamma) \right) \\ & \times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \frac{(-1)^{i+2} (i+1)!}{\ln b} \\ & \times \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\phi}{\theta-\phi} \right)^m \left(\frac{\gamma-\theta}{\phi-\theta} \right)^k (\theta)^{-(i+2)} \right. \\ & \left. + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta} \right)^m \left(\frac{\gamma-\phi}{\theta-\phi} \right)^k (\phi)^{-(i+2)} \right] d\gamma. \end{aligned} \tag{31}$$

Where

$$\widehat{I}_{G_{\lambda}}(\mathbf{s}, \mathbf{w}, \gamma) = \sum_{i=1}^p w_i G_{\lambda} \left(\frac{s_i}{w_i}, \gamma \right)$$

and τ_i, ν_i are defined in (15) and (16).

Proof. Take $f(x) = -\log(x)$ and $s_i = 1 (i = 1, \dots, p)$, in Theorem 2.1 then (24) and (25) becomes (30) and (31), respectively. Here,

$$\begin{aligned}\widehat{I}_f(1, \mathbf{w}) &= -\sum_{i=1}^p w_i \log\left(\frac{1}{w_i}\right) = \sum_{i=1}^p w_i \log(w_i) \\ &= H(\mathbf{w}).\end{aligned}$$

□

Remark 2.1. Inequalities (30) and (31) are reversed if $b \in (0, 1)$.

Corollary 2.2. Let A, B hold and log has base $b > 1$, then the following is the association between the Shannon entropy of \mathbf{s} and \mathbf{r} :

$$\begin{aligned}H(\mathbf{s}) &\geq H(\mathbf{r}) + \int_0^\phi \left(\widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\quad \times \left(\sum_{i=0}^{n-m-1} \nu_i(\gamma) \frac{(-1)^{i+2} i!}{\phi^{i+1} \ln b} + \sum_{i=0}^{m-1} \tau_i(\gamma) \frac{(-1)^{i+2} i!}{\theta^{i+1} \ln b} \right) d\gamma\end{aligned}\quad (32)$$

and

$$\begin{aligned}H(\mathbf{s}) &\geq H(\mathbf{r}) + \int_\theta^\phi \left(\widehat{I}_{G_\lambda}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_\lambda}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\quad \times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \left[\left(\frac{\gamma-\theta}{\phi-\theta} \right)^k \left(\frac{\gamma-\phi}{\theta-\phi} \right)^m \frac{(\gamma-\theta)^i (-1)^{i+2}}{\theta^{i+1} \ln b} \right. \\ &\quad \left. + \left(\frac{\gamma-\theta}{\phi-\theta} \right)^m \left(\frac{\gamma-\phi}{\theta-\phi} \right)^k \frac{(\gamma-\phi)^i (-1)^{i+2}}{\phi^{i+1} \ln b} \right] d\gamma.\end{aligned}\quad (33)$$

Proof. Taking n -convex function $xg(x) := x \log x$ for even $n > 3$ in (28) and (29), respectively and $s_i = 1 (i = 1, \dots, p)$, we obtain the desired results. □

Remark 2.2. If $b \in (0, 1)$, then inequalities (32) and (33) are reversed.

2.3 Kullback–Leibler divergence

The results of Theorems 2.1 and 2.2 are described in terms of Kullback–Leibler (K-L) divergence in this subsection. In [30], Kullback–Leibler (K-L) divergence for $\mathbf{w} = (w_1, \dots, w_p)$ and $\mathbf{v} = (v_1, \dots, v_p) \in \mathbb{R}_+^p$ is defined by

$$L(\mathbf{v}, \mathbf{w}) := \sum_{i=1}^p v_i \log\left(\frac{v_i}{w_i}\right).$$

Corollary 2.3. Consider the hypothesis of Corollary 2.1 hold. Then the following estimates hold:

$$\begin{aligned} \sum_{i=1}^m w_i \log\left(\frac{s_i}{w_i}\right) &\leq \sum_{i=1}^m w_i \log\left(\frac{r_i}{w_i}\right) \\ &+ \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma)\right) \frac{(-1)^{i+2}(i+1)!}{\ln b} \\ &\times \left(\sum_{i=0}^{m-1} \tau_i(\gamma)\theta^{-(i+2)} + \sum_{i=0}^{n-m-1} \nu_i(\gamma)\phi^{-(i+2)}\right) d\gamma, \end{aligned} \tag{34}$$

and

$$\begin{aligned} \sum_{i=1}^m w_i \log\left(\frac{s_i}{w_i}\right) &\leq \sum_{i=1}^m w_i \log\left(\frac{r_i}{w_i}\right) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma)\right) \\ &\times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{m+k-1}{k} \frac{(i+1)!(-1)^{i+2}}{\ln b} \\ &\times \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\phi}{\theta-\phi}\right)^m \left(\frac{\gamma-\theta}{\phi-\theta}\right)^k (\theta)^{-(i+2)} \right. \\ &\left. + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta}\right)^m \left(\frac{\gamma-\phi}{\theta-\phi}\right)^k (\phi)^{-(i+2)} \right] d\gamma. \end{aligned} \tag{35}$$

Proof. Putting n -convex function $f(x) := -\log x$ (even $n > 3$) in (24) and (25), we get (34) and (35), respectively. \square

Remark 2.3. If $b \in (0, 1)$ then the inequalities in (34) and (35) are reversed.

Corollary 2.4. Consider all the hypotheses of Corollary 2.2 hold, then the K-L divergence of \mathbf{s}, \mathbf{w} and \mathbf{r}, \mathbf{w} is defined by

$$\begin{aligned} L(\mathbf{s}, \mathbf{w}) &\geq L(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma)\right) \\ &\times \left(\sum_{i=0}^{m-1} \tau_i(\gamma) \frac{(-1)^{i+2} i!}{\theta^{i+1} \ln b} + \sum_{i=0}^{n-m-1} \nu_i(\gamma) \frac{(-1)^{i+2} i!}{\phi^{i+1} \ln b}\right) d\gamma \end{aligned} \tag{36}$$

and

$$\begin{aligned} L(\mathbf{s}, \mathbf{w}) &\geq L(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma)\right) \\ &\times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\phi}{\theta-\phi}\right)^m \left(\frac{\gamma-\theta}{\phi-\theta}\right)^k \right. \\ &\left. \times \frac{(-1)^{i+2} i!}{\theta^{i+1} \ln b} + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta}\right)^m \left(\frac{\gamma-\phi}{\theta-\phi}\right)^k \frac{(-1)^{i+2} i!}{\phi^{i+1} \ln b} \right] d\gamma. \end{aligned} \tag{37}$$

If $b \in (0, 1)$ then the inequalities (36) and (37) are conversed.

Proof. Using the n -convex function $xg(x) := x \log x$ for even $n > 3$, in (28) and (29), we obtain (36) and (37). \square

2.4 Bhattacharyya coefficient

The Bhattacharyya coefficient is used to express generalised majorization-type inequality in this section. In [30], Bhattacharyya coefficient for $\mathbf{w}, \mathbf{v} \in \mathbb{R}_+^\rho$ is defined by

$$\mathbb{B}(\mathbf{v}, \mathbf{w}) = \sum_{k=1}^{\rho} \sqrt{v_k w_k}. \tag{38}$$

Corollary 2.5. If all assumptions of Corollary (2.1) are true, then the following inequalities hold:

$$\begin{aligned} \mathbb{B}(\mathbf{s}, \mathbf{w}) \leq & \mathbb{B}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) \right) \\ & \times \left[\sum_{i=0}^{m-1} \frac{(\theta)^{-\left(\frac{2i+3}{2}\right)} \left(\prod_{u=1}^{i+1} (2u-1) \right) (-1)^{i+2}}{2^{i+2}} \tau_i(\gamma) \right. \\ & \left. + \sum_{i=0}^{n-m-1} \frac{(\phi)^{-\left(\frac{2i+3}{2}\right)} \left(\prod_{u=1}^{i+1} (2u-1) \right) (-1)^{i+2}}{2^{i+2}} \nu_i(\gamma) \right] d\gamma, \end{aligned} \tag{39}$$

and

$$\begin{aligned} \mathbb{B}(\mathbf{s}, \mathbf{w}) \leq & \mathbb{B}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) \right) \\ & \times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} \frac{(-1)^{i+2} \left(\prod_{u=1}^{i+1} (2u-1) \right)}{2^{i+2}} \\ & \times \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\phi}{\theta-\phi} \right)^m \left(\frac{\gamma-\theta}{\phi-\theta} \right)^k (\theta)^{-\left(\frac{2i+3}{2}\right)} \right. \\ & \left. + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta} \right)^m \left(\frac{\gamma-\phi}{\theta-\phi} \right)^k (\phi)^{-\left(\frac{2i+3}{2}\right)} \right] d\gamma. \end{aligned} \tag{40}$$

Proof. Taking the n -convex function $f(x) := -\sqrt{x}$ (even $n > 3$) in (24) and (25), respectively, then we get the desired result. \square

2.5 Jeffrey's distance

The generalised majorization-type inequality in terms of Jeffrey's distance is provided in this subsection. In [30], Jeffrey's distance for $\mathbf{w} = (w_1, \dots, w_\rho)$ and $\mathbf{v} = (v_1, \dots, v_\rho)$ is defined by

$$\mathbb{J}(\mathbf{v}, \mathbf{w}) = \sum_{k=1}^{\rho} \log\left(\frac{v_k}{w_k}\right)(v_k - w_k). \tag{41}$$

Corollary 2.6. If A and B hold and $b > 1$, then we get the following estimates, respectively:

$$\begin{aligned} \mathbb{J}(\mathbf{s}, \mathbf{w}) &\geq \mathbb{J}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_{\lambda}}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_{\lambda}}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\quad \times \frac{1}{\ln b} \left(\sum_{i=0}^{m-1} \tau_i(\gamma) \frac{(1 + \theta + i)(-1)^{i+2} i!}{(\theta)^{i+2}} \right. \\ &\quad \left. + \sum_{i=0}^{n-m-1} \frac{(i + \phi + 1) i! (-1)^{i+2}}{(\phi)^{i+2}} \nu_i(\gamma) \right) d\gamma \end{aligned} \tag{42}$$

and

$$\begin{aligned} \mathbb{J}(\mathbf{s}, \mathbf{w}) &\geq \mathbb{J}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_{\lambda}}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_{\lambda}}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\quad \times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{m+k-1}{k} \frac{(-1)^{i+2} (i)!}{\ln b} \\ &\quad \times \left[\frac{(\gamma - \theta)^i}{i!} \left(\frac{\gamma - \phi}{\theta - \phi} \right)^m \left(\frac{\gamma - \theta}{\phi - \theta} \right)^k (i + 1 + \theta)(\theta)^{-(i+2)} \right. \\ &\quad \left. + \frac{(\gamma - \phi)^i}{i!} \left(\frac{\gamma - \theta}{\phi - \theta} \right)^m \left(\frac{\gamma - \phi}{\theta - \phi} \right)^k (i + 1 + \phi)(\phi)^{-(i+2)} \right] d\gamma. \end{aligned} \tag{43}$$

If $0 < b < 1$, then the inequalities (42) and (43) are reversed.

Proof. Putting n -convex function $f(x) := (x - 1) \log x$ (even $n > 3$) in (24) and (25), we obtain (42) and (43), respectively. \square

2.6 Triangular discrimination

In [30], triangular discrimination for two positively distributed probabilities $\mathbf{w} = (w_1, \dots, w_\rho)$ and $\mathbf{v} = (v_1, \dots, v_\rho)$ is defined by

$$\mathbb{T}(\mathbf{v}, \mathbf{w}) = \sum_{k=1}^{\rho} \frac{(v_k - w_k)^2}{v_k + w_k}. \tag{44}$$

Corollary 2.7. If the assumptions of Corollary 2.1 are true, then

$$\begin{aligned} \mathbb{T}(\mathbf{s}, \mathbf{w}) &\geq \mathbb{T}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\times \left(\sum_{i=0}^{m-1} \tau_i(\gamma) \frac{4(i+2)!(-1)^{i+2}}{(1+\theta)^{(i+3)}} + \sum_{i=0}^{n-m-1} \nu_i(\gamma) \frac{4(i+2)!(-1)^{i+2}}{(1+\phi)^{(i+3)}} \right) d\gamma, \end{aligned} \quad (45)$$

and

$$\begin{aligned} \mathbb{T}(\mathbf{s}, \mathbf{w}) &\geq \mathbb{T}(\mathbf{r}, \mathbf{w}) + \int_{\theta}^{\phi} \left(\widehat{I}_{G_2}(\mathbf{s}, \mathbf{w}, \gamma) - \widehat{I}_{G_2}(\mathbf{r}, \mathbf{w}, \gamma) \right) \\ &\times \sum_{i=0}^{m-1} \sum_{k=0}^{m-1-i} \binom{k+m-1}{k} 4(-1)^{i+2} (i+2)! \\ &\times \left[\frac{(\gamma-\theta)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta} \right)^k \left(\frac{\gamma-\phi}{\theta-\phi} \right)^m (1+\theta)^{-(i+3)} \right. \\ &\left. + \frac{(\gamma-\phi)^i}{i!} \left(\frac{\gamma-\theta}{\phi-\theta} \right)^m \left(\frac{\gamma-\phi}{\theta-\phi} \right)^k (1+\phi)^{-(i+3)} \right] d\gamma. \end{aligned} \quad (46)$$

hold.

Proof. Taking n -convex function $f(x) := \frac{(x-1)^2}{x+1}$ (even $n > 3$) in (24) and (25), we get (45) and (46). \square

3. Conclusion

The main goal of the work is to provide various applications of information theory for higher-order convex functions in terms of different divergence measures and entropies. For this, Csiszár f -divergence is applied to majorization theorem via Hermite interpolation to establish novel bounds. Moreover, various new inequalities are stated by using various divergence measures including; Shannon entropy, Kullback-Leibler distance, Bhattacharyya coefficient, Jeffrey's distance and Triangular discrimination. All the results are modified and extended as they are derived for the class of higher-order convex functions. The key findings can be used in future studies to explore other distances and divergences, including: the Zipf-Mandelbrot law, Rényi divergence and Rényi entropy etc.

Authors contribution

AR initiated the work and made calculations. KAK supervised and validated the draft. JP dealt with the formal analysis and investigation. GP included the applications to information theory. All the authors read and approved the final manuscript.

Availability of data and materials

Data sharing is not applicable to this paper, as no datasets were generated or analyzed during the current study.

Acknowledgments

The authors wish to thank the editor for possible publication.

References

1. Pardo L. Statistical inference based on divergence measures. New York: Chapman and Hall/CRC; 2018.
2. Hibbard LS. Region segmentation using information divergence measures. *Med Image Ana.* 2004; 8(3): 233-44. doi: [10.1016/j.media.2004.06.003](https://doi.org/10.1016/j.media.2004.06.003).
3. Sen A, Sen MA, Foster JE, Amartya S, Foster JE. On economic inequality. Oxford: Oxford University Press; 1997.
4. Theil H. Economics and information theory. Amsterdam: North-Holland; 1967.
5. Chow CKCN, Liu C. Approximating discrete probability distributions with dependence trees. *IEEE Trans Inf Theory.* 1968; 14(3): 462-7. doi: [10.1109/tit.1968.1054142](https://doi.org/10.1109/tit.1968.1054142).
6. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J.* 1948; 27(3): 379-423. doi: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x).
7. Dragomir SS. Other inequalities for Csiszár divergence and applications. Pre Print, RGMIA Res Rep Coll. 2000.
8. Dragomir SS. An upper bound for the Csiszár f -divergence in terms of the variational distance and applications. *Panam Math J.* 2002; 12(4): 43-54.
9. Lovričević N, Pečarić Đ, Pečarić J. Zipf–Mandelbrot law, f -divergences and the Jensen-type interpolating inequalities. *J Inequal Appl.* 2018; 2018(1): 1-20. doi: [10.1186/s13660-018-1625-y](https://doi.org/10.1186/s13660-018-1625-y).
10. Siddique N, Imran M, Khan KA, Pečarić J. Majorization inequalities via green functions and Fink's identity with applications to Shannon entropy. *J Inequal Appl.* 2020; 2020(1): 1-14. doi: [10.1186/s13660-020-02455-0](https://doi.org/10.1186/s13660-020-02455-0).
11. Niaz T, Khan KA, Pečarić Đ, Pečarić J. Estimation of different entropies via Taylor one point and Taylor two points interpolations using Jensen type functionals. *Int J Anal Appl.* 2019; 17(5): 686-710.
12. Latif N, Siddique N, Pečarić J. Generalization of majorization theorem-II. *J Math Inequal.* 2018; 12(3): 731-52. doi: [10.7153/jmi-2018-12-56](https://doi.org/10.7153/jmi-2018-12-56).
13. Khan MA, Bradanovic SI, Latif N, Pečarić Đ, Pečarić J. Majorization inequality and information theory. Zagreb: Element; 2019.
14. Rasheed A, Khan KA, Pečarić J, Pečarić Đ. Generalizations of Levinson-type inequalities via new green functions and Hermite interpolating polynomial. *J Inequal Appl.* 2024; 2024(1): 70. doi: [10.1186/s13660-024-03146-w](https://doi.org/10.1186/s13660-024-03146-w).
15. Rasheed A, Khan KA, Pečarić J, Pečarić Đ. Generalization of the Levinson inequality via new green functions with applications to information theory. *J Inequal Appl.* 2023; 2023(1): 124. doi: [10.1186/s13660-023-03040-x](https://doi.org/10.1186/s13660-023-03040-x).
16. Bilal M, Khan KA, Nosheen A, Pečarić J. Generalizations of Shannon type inequalities via diamond integrals on time scales. *J Inequal Appl.* 2023; 2023(1): 24. doi: [10.1186/s13660-023-02930-4](https://doi.org/10.1186/s13660-023-02930-4).
17. Bilal M, Khan KA, Nosheen A, Pečarić J. Some inequalities related to Csiszár divergence via diamond integral on time scales. *J Inequal Appl.* 2023; 2023(1): 55. doi: [10.1186/s13660-023-02966-6](https://doi.org/10.1186/s13660-023-02966-6).
18. Pečarić JE, Proschan F, Tong YL. Convex functions, partial orderings, and statistical applications. Academic Press; 1992. p. 187.
19. Liese F, Vajda I. Convex statistical distances, teubner-texte zur mathematik, 95. Leipzig: Teubner; 1987.
20. Fuchs L. A new proof of an inequality of Hardy-Littlewood-Polya; 2003. *Mat. Tidsskr, B.*
21. Mehmood N, Agarwal RP, Butt SI, Pečarić J. New generalizations of Popoviciu type inequalities via new green's functions and Montgomery identity. *J Inequal Appl.* 2017; 2017: 1-21. doi: [10.1186/s13660-017-1379-y](https://doi.org/10.1186/s13660-017-1379-y).

22. Agarwal RP, Wong PJY. Error inequalities in polynomial interpolation and their applications. Dordrecht/Boston/London: Kluwer Academic; 1993.
23. Adeel M, Khan KA, Pečarić Đ, Pečarić J. Entropy results for Levinson-type inequalities via green functions and Hermite interpolating polynomial. *Aequ Math.* 2022; 96(1): 1-16. doi: [10.1007/s00010-021-00845-3](https://doi.org/10.1007/s00010-021-00845-3).
24. Mehmood N, Butt SI, Pečarić Đ, Pečarić J. New bounds for Shannon, relative and Mandelbrot entropies via Hermite interpolating polynomial. *Demonstr Math.* 2018; 51(1): 112-30. doi: [10.1515/dema-2018-0011](https://doi.org/10.1515/dema-2018-0011).
25. Latif N, Siddique N, Pečarić J. Majorization inequalities via Peano's representation of Hermite's polynomial. *Int J Anal Appl.* 2018; 16: 374-99.
26. Csiszár I. Information measures: a critical survey. In: Tans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Random Process and 8th European Meeting of Statist., B. Academia Prague; 1978. p. 73-86.
27. Csiszár I. Information-type measures of difference of probability distributions and indirect observations. *Stud Sci Math Hungar.* 1967; 2: 299-318.
28. Horváth L, Pečarić Đ, Pečarić J. Estimations of f-and Rényi divergences by using a cyclic refinement of the Jensen's inequality. *Bull Malays Math Sci Soc.* 2019; 42(3): 933-46. doi: [10.1007/s40840-017-0526-4](https://doi.org/10.1007/s40840-017-0526-4).
29. Pearce C, Pecaric J. Shannon's and related inequalities in information theory. *Survey on classical inequalities*; 2000. p. 127.
30. Khan, MA, Pečarić, Đ, Pečarić, J, (2021). A new refinement of the Jensen inequality with applications in information theory, *Bull Malays Math Sci Soc*, 44(1), 267-78. doi: [10.1007/s40840-020-00944-5](https://doi.org/10.1007/s40840-020-00944-5).

Corresponding author

Awais Rasheed can be contacted at: awais10mathematicians@gmail.com