Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

Advanced Generalizations of Convex Function Inequalities: Implications for High-Order Divergence and Entropy Estimation in Information Theory

Memoona Mukhtar[1], Tasadduq Niaz[2], Waseem Abbasi[3], Wadood Abdul[4], Khursheed Aurangzeb[4], Muhammad Shahid Anwar[5*]

[1] Department of Mathematics and Statistics, The University of Lahore, Sargodha, 40100, Pakistan
[2] Department of Mathematics, Superior University, Lahore, 54000, Pakistan
[3] Department of Computer Science, The University of Lahore, Lahore, 54000, Pakistan
[4] Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, 11543, Saudi Arabia
[5] Department of AI and Software, Gachon University, Seongnam-Si, 13120, South Korea
E-mail: shahidanwar786@gachon.ac.kr

Received: 17 September 2024; **Revised:** 18 October 2024; **Accepted:** 18 November 2024

Abstract: The inequalities involving convex function have many applications in analysis and in recent years it has helped in estimating many entropies and divergences that are used in information theory. In this paper, an inequality constructed by the two inequalities Jensen inequality and Lah-Ribaric inequality is considered. The non-negative difference of the this inequality are used to construct the non-negative difference. Two identities Abel-Gontscharoff and Montgomery identity at a time are used in non-negative differences to construct new identities. These identities are used to generalized the inequality for higher order convex function. Furthermore for the sake of application in information theory these generalized results are used to estimate Csiszer divergence, Shannon entropy, Kullback Leibler divergence and Zipf-Mandelbrot laws.

Keywords: Jensen's inequality, Lah-Ribaric inequality, Lidstone identity, information theory, Zipf-Madelbrot law

MSC: 65L05, 34K06, 34K28

1. Introduction

In literature, there are many divergences that are used to find the distance between probability distribution by making an average value, one of them is Csiszár divergence. The Shannon entropy is actually a quantity in information theory that we associate with random variable and it gives an average level of uncertainty, information, or possible outcome of a variable. The researchers working on mathematical inequalities related to convex function have found it very interesting to find the limitations and bounds for the different functionals. Their work motivates us to find the bound for divergence and entropies from information theory. Pečarić and Tong [1] was the first to introduce the higher order convexity. It has been observed that some properties are not valid for higher order convex that are valid for convex function and this

DOI: https://doi.org/10.37256/cm.6420255733

motivates researchers to study higher-order convexity. A higher-order convex function can be identified by the result that "if $h^{(\lambda)}$ exists and h is λ-convex if and only if $h^{(\lambda)}$ is non-negative".

There has been significant growth in the study of classical inequalities like the Hölder, Jensen, and others. These inequalities were initially presented in discrete and integral forms, after which numerous generalizations and improvements have been explained (for example, see [2, 3]). Recent research has shown that they are extremely helpful in information theory (for example, see [4]). Researchers have generalized different inequalities for higher-order convex functions using different interpolations like Popoviciu inequality have been generalized by using Lidstone polynomial [5], Hermite interpolating polynomial [6], Fink's identity [7], Taylor's polynomial [8], Abel-Gontscharoff interpolation [5], and Montgomery identity [9]. In [10] Jensen's inequality is given as follows.

Let J be an interval in $\mathbb{R}$ and $h: J \to \mathbb{R}$ a convex function. If $z = (z_1, \ldots, z_n)$ is any n-tuple in J^n and $\mathbf{q} = (q_1, \ldots, q_n)$ a non-negative n-tuple such that $Q_n = \sum_{i=1}^{n} q_i > 0$, then the Jensen's inequality

$$h\left(\frac{1}{Q_n}\sum_{i=1}^{n} q_i z_i\right) \leq \frac{1}{Q_n}\sum_{i=1}^{n} q_i h(z_i), \tag{1}$$

holds.

Inequality (1) has remained of great interest among researchers for many years due to its application and its impact on other inequalities because many of the inequalities in literature are the consequence of (1). A lot of refinement of (1) has been constructed. Many researchers have generalized (1) and its refinement (see [11, 12]). Butt et al. [13] found the application in information theory by using the Jensen-Grüss type inequalities.

The Lah-Ribaric inequality is closely related to Jensen's inequality. In [14, 15] Lah-Ribarič inequality is given as follows,

$$\frac{1}{Q_n}\sum_{i=1}^{n} q_i h(z_i) \leq \frac{M-\bar{z}}{M-m} h(m) + \frac{\bar{z}-m}{M-m} h(M), \tag{2}$$

which holds when $h: J \to \mathbb{R}$ is a convex function on J, $[m, M]^n \subset J$, $-\infty < m < M < +\infty$, $\boldsymbol{q}$ is as in (1), $z = (z_1, \ldots, z_n)$ is any n-tuple in $[m, M]^n$ and $\bar{z} = \frac{1}{Q_n}\sum_{i=1}^{n} q_i z_i$. If h is strictly convex then (1) is strict unless $z_i \in m, M$ for all $i \in \{j : q_j > 0\}$.

The inequality (1) has been researched, and interested readers are able to find numerous related results in more recent research as well as in monographs like [11, 12]. It would be interesting to learn more about how the previously mentioned imbalance is possibly further developed. We also deal with the idea of f-divergences, which evaluates the separation between two probability distributions. One of the most significant is the Csiszár f-divergence, of which specific instances include the Shannon entropy and the Kullback-Leibler divergence. We determine the relationships for the above-mentioned f-divergences. We also examined more closely the Zipf-Mandelbrot law.

2. Preliminaries

Consider the following lemma (see [16]).

Lemma 1 Let h be a convex function on an interval J. If $l, m, n, o \in J$ such that $l \leq m < n \leq o$, then the inequality,

$$\frac{n-z}{n-m} h(m) + \frac{z-m}{n-m} h(n) \leq \frac{o-z}{o-l} h(l) + \frac{z-l}{o-l} h(o), \tag{3}$$

holds for any $z \in [m, n]$.

The result that gives the refinement of the Lah-Ribarič inequality (2) from [15].

Theorem 1 Suppose a convex function $h : J \to \mathbb{R}$, where $J \subset \mathbb{R}$ be an interval, $[m, M] \subset J$, $\mathbf{q} = (q_1, \ldots, q_n)$ be a non-negative n-tuple, $\mathbf{z} = (z_1, \ldots, z_n)$ is any n-tuple in $[m, M]^n$ and $\bar{z} = \frac{1}{Q_n} \sum_{i=1}^{n} q_i z_i$. Suppose $N_i \subseteq (1, \ldots, n)$, $i = 1, \ldots, m$ where $N_i \bigcap N_j = \phi$ for $i \neq j$, $\bigcup_{i=1}^{m} N_i = \{1, \ldots, n\}$, $\sum_j \in N_i q_j > 0$, for $i = 1, \ldots, m$ and $m_i = \min\{z_j : j \in N_i\}$, $M_i = \max\{x_j : j \in N_i\}$, for $i = 1, \ldots, m$. Then, the following intervals hold,

$$\frac{1}{Q_n} \sum_{i=1}^{n} q_i h(z_i) \leq \frac{1}{Q_n} \sum_{i=1}^{m} \left(\sum_{j \in N_i} q_j \right) \left[\frac{M_i - \bar{z}_i}{M_i - m_i} h(m_i) + \frac{\bar{z}_i - m_i}{M_i - m_i} h(M_i) \right]$$

$$\leq \frac{M - \bar{z}}{M - m} h(m) + \frac{\bar{z} - m}{M - m} h(M), \tag{4}$$

where if h is a concave function on J, then the inequalities in (4) are reversed.

Consider the following functionals from (4).

$$\theta_1(h) = \frac{1}{Q_n} \sum_{i=1}^{m} \left(\sum_{j \in N_i} q_j \right) \left[\frac{M_i - \bar{z}_i}{M_i - m_i} h(m_i) + \frac{\bar{z}_i - m_i}{M_i - m_i} h(M_i) \right] - \frac{1}{Q_n} \sum_{i=1}^{n} q_i h(z_i), \tag{5}$$

$$\theta_2(h) = \frac{M - \bar{z}}{M - m} h(m) + \frac{\bar{z} - m}{M - m} h(M) - \frac{1}{Q_n} \sum_{i=1}^{m} \left(\sum_{j \in N_i} p_j \right) \left[\frac{M_i - \bar{z}_i}{M_i - m_i} h(m_i) + \frac{\bar{z}_i - m_i}{M_i - m_i} h(M_i) \right], \tag{6}$$

$$\theta_3(h) = \frac{M - \bar{z}}{M - m} h(m) + \frac{\bar{z} - m}{M - m} h(M) - \frac{1}{Q_n} \sum_{i=1}^{n} q_i h(z_i), \tag{7}$$

If the assumption of Theorem 1 is valid then

$$\theta_i(h) \geq 0, \quad i = 1, 2, 3, \tag{8}$$

and inequalities symbols in (8) are reversed if h is a concave function.

Let $h_1, h_2, h_3 : I \to \mathbb{R}$ be three functions with the same assumption of Theorem 1 and μ_1, μ_2 and μ_3 be any scalars, then

$$\theta_i(\mu_1 h_1 + \mu_2 h_2 + \mu_3 h_3) = \mu_1 \theta_i(h_1) + \mu_2 \theta_i(h_2) + \mu_3 \theta_i(h_3), \quad i = 1, 2, 3. \tag{9}$$

For $n = 2$, the Abel-Gontscharoff [17] polynomial is given by

$$h(z) = h(\beta_1) + (z - \beta_1) h'(\beta_2) + \int_{\beta_1}^{\beta_2} G_1(z, v) h''(v) dv, \tag{10}$$

where

$$G_1(z,\ v) = \begin{cases} \beta_1 - v, & \beta_1 \le v \le z; \\ \beta_1 - z, & z \le v \le \beta_2. \end{cases} \tag{11}$$

In [7], the new Green functions are defined as follows.

$$G_2(z,\ v) = \begin{cases} z - \beta_2, & \beta_1 \le v \le z; \\ v - \beta_2, & z \le v \le \beta_2. \end{cases} \tag{12}$$

$$G_3(z,\ v) = \begin{cases} z - \beta_1, & \beta_1 \le v \le z; \\ v - \beta_1, & z \le v \le \beta_2. \end{cases} \tag{13}$$

$$G_4(z,\ v) = \begin{cases} \beta_2 - v, & \beta_1 \le v \le z; \\ \beta_2 - z, & z \le v \le \beta_2. \end{cases} \tag{14}$$

Remark 1 Observe that, since

$$\frac{d^2}{dv^2} G_k(.,\ v) \ge 0, \quad k = 1,\ \ldots,\ 4, \tag{15}$$

therefore by using this function in Theorem 1 we observe

$$\theta_i(G_k(.,\ v)) \ge 0, \quad k = 1,\ \ldots,\ 4, \quad i = 1,\ 2,\ 3. \tag{16}$$

The authors in [7] also use these new Green functions to construct the following three new identities for twice differentiable functions on $[\beta_1,\ \beta_2]$.

$$h(z) = h(\beta_2) + (\beta_2 - z)h'(\beta_1) + \int_{\beta_1}^{\beta_2} G_2(z,\ v)h''(v)dv, \tag{17}$$

$$h(z) = h(\beta_2) - (\beta_2 - \beta_1)h'(\beta_2) + (z - \beta_1)h'(\beta_1) + \int_{\beta_1}^{\beta_2} G_3(z,\ v)h''(v)dv, \tag{18}$$

$$h(z) = h(\beta_1) - (\beta_2 - \beta_1)h'(\beta_1) + (\beta_2 - z)h'(\beta_2) + \int_{\beta_1}^{\beta_2} G_4(z,\ v)h''(v)dv. \tag{19}$$

Let $\varphi \in C^\infty([0,\ 1])$, the Lidstone polynomial [17] is defined as

$$\varphi(\mu) = \sum_{s=0}^{\lambda-1} \left[\varphi^{(2s)}(0)\left(F_s(1-\mu)+F_s(\mu)\right)\right] + \int_0^1 G_p(\mu,\, l)\varphi^{(2q)}(l)dl, \tag{20}$$

where F_s is a $s+1$ degree polynomial defined by the following relations

$$F_0(\mu) = \mu, \quad F_\lambda''(\mu) = F_{\lambda-1}(\mu), \quad F_\lambda(0) = F_\lambda(1) = 0, \quad \lambda \geq 1, \tag{21}$$

and

$$\hat{G}_1(\mu,\, l) = \hat{G}(\mu,\, l) = \begin{cases} l(\mu-l), & \alpha_1 \leq l < \mu \leq \mu_2; \\ \mu(l-1), & \alpha_1 \leq \mu < l \leq \mu_2, \end{cases} \tag{22}$$

where (12) is a Green's function of the differential operator $\dfrac{d^2}{ds^2}$ on $[0,\, 1]$, and with the successive iterates of $\hat{G}(\mu,\, l)$

$$\hat{G}_\lambda(\mu,\, l) = \int_0^1 \hat{G}_1(\mu,\, t)\hat{G}_{\lambda-1}(\mu,\, t)(l)dt, \quad \lambda \geq 2. \tag{23}$$

The Lidstone polynomial can also be expressed as the following

$$P_\lambda(\mu) = \int_0^1 G_\lambda(\mu,\, l)l dl. \tag{24}$$

The Lidstone series representation [17] of $\varphi \in C^{2\lambda}[\beta_1,\, \beta_2]$ is given by

$$\varphi(\mu) = \sum_{s=0}^{\lambda-1} (\beta_2-\beta_1)^{(2s)} \left[\varphi^{(2s)}(\beta_1)F_s\left(\frac{\beta_2-\mu}{\beta_2-\beta_1}\right) + \varphi^{(2s)}(\beta_2)F_s\left(\frac{\mu-\beta_1}{\beta_2-\beta_1}\right)\right]$$

$$+ (\beta_2-\beta_1)^{(2s-1)} \int_{\beta_1}^{\beta_2} \hat{G}_\lambda\left(\frac{\beta_2-\mu}{\beta_2-\beta_1},\, \frac{l-\beta_1}{\beta_2-\beta_1}\right)\varphi^{(2\lambda)}(l)dl. \tag{25}$$

3. Generalization for higher-order convex function

In this section, the inequality (4) is generalized for higher order convex function using two interpolations at a time. In the first result of this section new identities are constructed using Abel-Gontscharoff and Lidstone identity at a time. The second results gives the generalization of (4). In the last result of this section the generalization of (4) is given for different domain.

Consider the following hypothesis:

($\mathbf{H_1}$) Suppose a function $\varphi \in C^{2\lambda}[\beta_1,\, \beta_2]$.

($\mathbf{H_2}$) Suppose $z_1,\, \ldots,\, z_n \in [\beta_1,\, \beta_2]$ and $q_1,\, \ldots,\, q_n$ are positive real numbers such that $q_1+\ldots+q_n = 1$.

In the following result, new identities are constructed using non-negative differences give in (5)-(7).

Theorem 2 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$, suppose that $\theta_i(i=1,\ 2,\ 3),\ G_k(k=1,\ 2,\ 3,\ 4)$ and $\hat{G}_m$ are the same as defined in (5)-(7), (11)-(14) and (23) respectively, then

$$\theta_i(\varphi)=\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i\left(G_k(.,\ v)\right)\left(\varphi^{(2s+2)}\left(\beta_1\right)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.$$

$$\left.+\varphi^{(2s+2)}\left(\beta_2\right)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv \tag{26}$$

$$+\left(\beta_2-\beta_1\right)^{2s-1}\int_{\beta_1}^{\beta_2}\varphi^{(2\lambda)}(l)\left(\int_{\beta_1}^{\beta_2}\theta_i\left(G_k(.,\ v)\right)\hat{G}_{\lambda-1}\left(\frac{v-\beta_1}{\beta_2-\beta_1},\ \frac{l-\beta_1}{\beta_2-\beta_1}\right)dv\right)dl.$$

Proof. Using (10), (17), (18), (19) in (5), (6), and (7), and using property (9) of $\theta_i(\varphi)$, we get

$$\theta_i(\varphi)=\int_{\beta_1}^{\beta_2}\varphi''(v)\theta_i(G_k(.,\ v)dv,\ i=1,\ 2,\ 3. \tag{27}$$

In (25), replace φ with φ'', λ by $\lambda-1$ and μ with v, we get

$$\varphi''(v)=\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{(2s)}\left[\varphi^{(2s+2)}(\beta_1)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)+\varphi^{(2s+2)}(\beta_2)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right]$$

$$+\left(\beta_2-\beta_1\right)^{(2s-1)}\int_{\beta_1}^{\beta_2}\hat{G}_{\lambda-1}\left(\frac{\beta_2-v}{\beta_2-\beta_1},\ \frac{l-\beta_1}{\beta_2-\beta_1}\right)\varphi^{(2\lambda)}(l)dl. \tag{28}$$

Using $\varphi''(v)$ in (27) and with the use of Fubini's theorem, the identity (26) will be obtained. □

The following result is the generalization of (4) for the 2λ-convex function.

Theorem 3 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$, suppose that $\theta_i(i=1,\ 2,\ 3),\ G_k(k=1,\ 2,\ 3,\ 4)$ and $\hat{G}_\lambda$ are the same as defined in (5)-(7), (11)-(14) and (23) respectively, further suppose φ is 2λ-convex function, if

$$\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\hat{G}_{\lambda-1}\left(\frac{v-\beta_1}{\beta_2-\beta_1},\ \frac{l-\beta_1}{\beta_2-\beta_1}\right)dv\geq 0, \tag{29}$$

for all $v\in[\beta_1,\ \beta_2]$, then

$$\theta_i(\varphi)\geq\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\left(\varphi^{(2s+2)}(\beta_1)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.$$

$$\left.+\varphi^{(2s+2)}(\beta_2)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv. \tag{30}$$

Proof. Since, φ is 2λ-convex function, so $\varphi^{(2\lambda)} \geq 0$ (since φ is λ-convex). By Theorem 1.41 gives in [1], the function $\varphi^{(k)}$ exists and $(2\lambda - k)$-convex for all $k \in \{1, \ldots, 2\lambda - 2\}$. So $\varphi^{(k)} \geq 0$ for all $k \in \{1, \ldots, 2\lambda - 2\}$. On applying Theorem 2, we get (30). □

Theorem 4 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$, let $\varphi : [\beta_1, \beta_2] \to \mathbb{R}$ be a function, if φ is 2λ-convex function then the following two results are valid.

(i) If $\lambda \geq 3$ is an odd integer, then (30) holds.

(ii) Suppose (30) be satisfied, if

$$T(v) := \sum_{s=0}^{\lambda-2} (\beta_2 - \beta_1)^{2s} \left(\varphi^{(2s+2)}(\beta_1) F_s \left(\frac{\beta_2 - v}{\beta_2 - \beta_1} \right) + \varphi^{(2s+2)}(\beta_2) F_s \left(\frac{v - \beta_1}{\beta_2 - \beta_1} \right) \right) \geq 0, \tag{31}$$

then

$$\theta_i(\varphi) \geq 0. \tag{32}$$

Proof. **(i)** For $\lambda \geq 3$ odd integers the function $G_{\lambda-1}$ is positive. Also from Remark 1, $\theta_i(G_k(., v)) \geq 0$, therefore (29) holds. Hence by using Theorem 3, we get (30).

(ii) Using the positivity of $\theta_i(G_k(., v))$ and (31), we get (32). □

4. Information theory

In this section, the new generalized results are used to estimate different divergences and entropies like; Csiszár divergence, Kullback-Leibler divergence, Shannon entropy and Zipf-Mandelbrot law.

Csiszár divergence [18, 19] is defined as follows: Suppose $\psi : (0, \infty) \to \mathbb{R}$ is a positive function. Let $r := (r_1, \ldots, r_n)$ and $s := (s_1, \ldots, s_n)$ be positive probability distributions. The Csiszár f-divergence $I_\psi(r, s)$ is defined as

$$I_\psi(r, s) := \sum_{i=1}^{n} s_i \psi\left(\frac{r_i}{s_i}\right). \tag{33}$$

Horv́ath et al. [20] generalized the Csiszár divergence as follows.

Suppose $\psi : (0, \infty) \to \mathbb{R}$ be a function. Let $r := (r_1, \ldots, r_n)$ be a real n-tuple and $s := (s_1, \ldots, s_n)$ be a positive n-tuple such that

$$\frac{r_j}{s_j} \in (0, \infty), \; j = 1, \ldots, n. \tag{34}$$

Then,

$$\hat{I}_\psi(r, s) := \sum_{j=1}^{n} s_j \psi\left(\frac{r_j}{s_j}\right). \tag{35}$$

Theorem 5 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$, let $\mathbf{r} = (r_1, \ldots, r_n)$ and $\mathbf{s} = (s_1, \ldots, s_n)$ are n-tuples lies in $(0, \infty)^n$, suppose I is an interval, with

$$\frac{r_k}{s_k} \in I, \; k = 1, \ldots, n.$$

If $\varphi : (0, \infty) \to \mathbb{R}$ is an 2λ-convex function, such that $\varphi \in C^{2\lambda}[\beta_1, \beta_2]$ (where $\lambda \geq 3$ is supposed to be an odd integer), then for $i = 1, 2, 3,$

$$\theta_{i,\, cis}(\varphi) \geq \sum_{s=0}^{\lambda-2} (\beta_2 - \beta_1)^{2s} \int_{\beta_1}^{\beta_2} \theta_i(G_k(., v)) \left(\varphi^{(2s+2)}(\beta_1) F_s\left(\frac{\beta_2 - v}{\beta_2 - \beta_1}\right) + \varphi^{(2s+2)}(\beta_2) F_s\left(\frac{v - \beta_1}{\beta_2 - \beta_1}\right)\right) dv, \tag{36}$$

where

$$\theta_{1,\, cis}(\varphi) = \sum_{i=1}^{m} \left(\sum_{j \in N_i} q_j\right) \left[\frac{M_i - \bar{z}_i}{M_i - m_i} \varphi(m_i) + \frac{\bar{z}_i - m_i}{M_i - m_i} \varphi(M_i)\right] - \sum_{k=1}^{n} s_k \varphi\left(\frac{r_k}{s_k}\right), \tag{37}$$

$$\theta_{2,\, cis}(\varphi) = (Q_n S_n) \left[\frac{M - \bar{z}}{M - m} \varphi(m) + \frac{\bar{z} - m}{M - m} \varphi(M)\right] - \sum_{i=1}^{m} \left(\sum_{j \in N_i} p_j\right) \left[\frac{M_i - \bar{z}_i}{M_i - m_i} \varphi(m_i) + \frac{\bar{z}_i - m_i}{M_i - m_i} \varphi(M_i)\right], \tag{38}$$

and

$$\theta_{3,\, cis}(\varphi) = (-Q_n S_n) \left[\frac{M - \bar{z}}{M - m} \varphi(m) + \frac{\bar{z} - m}{M - m} \varphi(M)\right] + \sum_{k=1}^{n} s_k \varphi\left(\frac{r_k}{s_k}\right). \tag{39}$$

Proof. Since the functions $G_k(., v)$ are convex for $k = 1, 2, 3, 4,$ therefore $\theta_i(G_k(., v)) \geq 0$. Also note that $\hat{G}_\lambda$ is positive for every odd number λ greater than 2, so (29) holds. Consider $q_k = \frac{s_k}{\sum_{k=1}^{n} s_k}$ and $z_k = \frac{r_k}{s_k}$ in (30) and multiplying by $-Q_n S_n$, we get (36). □

Remark 2 For $i = 3$ the inequality (36) gives

$$\hat{I}_\varphi(r, s) \geq \sum_{s=0}^{\lambda-2} (\beta_2 - \beta_1)^{2s} \int_{\beta_1}^{\beta_2} \theta_i(G_k(., v)) \left(\varphi^{(2s+2)}(\beta_1) F_s\left(\frac{\beta_2 - v}{\beta_2 - \beta_1}\right) + \varphi^{(2s+2)}(\beta_2) F_s\left(\frac{v - \beta_1}{\beta_2 - \beta_1}\right)\right) dv + (Q_n S_n) \left[\frac{M - \bar{z}}{M - m} \varphi(m) + \frac{\bar{z} - m}{M - m} \varphi(M)\right]. \tag{40}$$

Proof. Since

$$\hat{I}_{\varphi}(r,\ s)=\sum_{k=1}^{n} s_k \varphi\left(\frac{r_k}{s_k}\right), \tag{41}$$

therefore for $i=3$ in the inequality (36) implies (40). □

4.1 *Inequalities for shannon entropy*

For a positive probability distribution $r=(r_1,\ \ldots,\ r_n)$, the Shannon entropy S [20] is defined as

$$S:=-\sum_{s=1}^{n} r_s \ln(r_s). \tag{42}$$

Corollary 1 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$. If $r=(r_1,\ \ldots,\ r_n)$, then

$$S \leq \left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\left(\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_1^{2s+2}}F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$
$$\left.\left.+\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_2^{2s+2}}F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]+(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \tag{43}$$

Proof. Using $\varphi(x)=-\ln(x)$ in (40), we get

$$\sum_{k=1}^{n} s_k\left[-\ln\left(\frac{r_k}{s_k}\right)\right] \geq -\left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\left(\ln^{(2s+2)}(\beta_1)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$
$$\left.\left.+\ln^{(2s+2)}(\beta_2)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]-(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \tag{44}$$

Since for $\varphi(x)=\ln(x)$, the $(2s+2)$-th derivative is given as

$$\varphi^{(2s+2)}(x)=(-1)^{(2s+1)}\frac{(2s+1)!}{x^{2s+2}}. \tag{45}$$

Use (45) in (44) and multiply by -1 on both sides, we get

$$-\sum_{k=1}^{n} s_k \ln\left(\frac{s_k}{r_k}\right) \leq \left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\, v))\left(\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_1^{2s+2}}F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$

$$\left.\left.+\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_2^{2s+2}}F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]+(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \tag{46}$$

Taking $r = (1,\, \ldots,\, 1)$ and using the properties of ln in (46), and after simplification, it become (43). □

4.2 *Inequalities for Kullback-Leibler divergence*

Consider the two probability distributions $\mathbf{r} = (r_1,\, \ldots,\, r_n)$ and $\mathbf{s} = (s_1,\, \ldots,\, s_n)$, the function defined by

$$D(r,\, s) := \sum_{j=1}^{n} r_j \ln\left(\frac{r_j}{s_j}\right). \tag{47}$$

is called Kullback-Leibler divergence [20].

Corollary 2 Assume $(\mathbf{H_1})$ and $(\mathbf{H_2})$. If $r = (r_1,\, \ldots,\, r_n)$ and $s = (s_1,\, \ldots,\, s_n)$, then

$$D(s,\, r) \geq \left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\, v))\left(\frac{(-1)^{(2s+2)}(2s+1)!}{\beta_1^{2s+2}}F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$

$$\left.\left.+\frac{(-1)^{(2s+2)}(2s+1)!}{\beta_2^{2s+2}}F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]-(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \tag{48}$$

Proof. Using $\varphi(x) = -\ln(x)$ in (40), we get

$$\sum_{k=1}^{n} s_k\left[-\ln\left(\frac{r_k}{s_k}\right)\right] \geq -\left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\, v))\left(\ln^{(2s+2)}(\beta_1)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$

$$\left.\left.+\ln^{(2s+2)}(\beta_2)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]-(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right].$$

Using the property of ln which is $-\ln x = \ln(x)^{-1} = \ln\left(\frac{1}{x}\right)$, we get

$$\sum_{k=1}^{n} s_k \ln\left(\frac{s_k}{r_k}\right) \geq -\left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\, v))\left(\ln^{(2s+2)}(\beta_1)F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$

$$\left.\left.+\ln^{(2s+2)}(\beta_2)F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]-(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \tag{49}$$

Since for $\varphi(x) = \ln(x)$, the $(2s+2)$-th derivative is given as $\varphi^{(2s+2)}(x) = (-1)^{(2s+1)}\dfrac{(2s+1)!}{x^{2s+2}}$. Use the properties of ln on the larger side of the inequality (49),

$$\sum_{k=1}^{n} s_k \ln\left(\frac{s_k}{r_k}\right) \geq \left[\sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\left(\frac{(-1)^{(2s+2)}(2s+1)!}{\beta_1^{2s+2}}F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.\right.$$

$$\left.\left.+\frac{(-1)^{(2s+2)}(2s+1)!}{\beta_2^{2s+2}}F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv\right]-(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \qquad (50)$$

Hence, we get (48). □

4.3 *Inequalities for Zipf-Mandelbrot law*

The Zipf-Mandelbrot entropy for positive integer n, and for real numbers $r \in [0,\ \infty),\ s \in (0,\ \infty)$ is defined as

$$Z(H,\ r,\ s) = \frac{s}{H_{r,\ s}^{n}}\sum_{\kappa=1}^{n}\frac{\ln(\kappa+r)}{(\kappa+r)^s}+\ln(H_{r,\ s}^{n}), \qquad (51)$$

where

$$H_{r,\ s}^{n} = \sum_{\upsilon=1}^{n}\frac{1}{(\upsilon+r)^s}. \qquad (52)$$

Corollary 3 Assume ($\mathbf{H_1}$) and ($\mathbf{H_2}$). If $r = (r_1,\ \ldots,\ r_n)$ and $s = (s_1,\ \ldots,\ s_n)$, then

$$S = Z(H,\ r,\ s) \leq \sum_{s=0}^{\lambda-2}(\beta_2-\beta_1)^{2s}\int_{\beta_1}^{\beta_2}\theta_i(G_k(.,\ v))\left(\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_1^{2s+2}}F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right)\right.$$

$$\left.+\frac{(-1)^{(2s+1)}(2s+1)!}{\beta_2^{2s+2}}F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right)dv+(Q_nS_n)\left[\frac{M-\bar{z}}{M-m}\ln(m)+\frac{\bar{z}-m}{M-m}\ln(M)\right]. \qquad (53)$$

Proof. Consider the following function

$$r_\kappa = \phi(\kappa;\ n,\ r,\ s) = \frac{1}{(\kappa+r)^s H_{r,\ s}^{n}}. \qquad (54)$$

Note that $\sum_{\kappa=1}^{n} r_\kappa = \sum_{\kappa=1}^{n}\dfrac{1}{(\kappa+r)^s H_{r,\ s}^{n}} = \dfrac{H_{r,\ s}^{n}}{H_{r,\ s}^{n}} = 1$, therefore use r_κ in Shannon entropy gives Zipf-Mandelbrot entropy that is:

$$S_r = -\sum_{\kappa=1}^{n} r_\kappa \ln(r_\kappa) = Z(H,\ r,\ s). \tag{55}$$

Thus using r_κ in (43), gives (53). □

Corollary 4 Assume $(\mathbf{H_1})$. For $r_1,\ r_2 \geq 0$ and $s_1,\ s_2 \geq 0$, suppose $H^n_{r_1,\ s_1} = \sum_{\upsilon=1}^{n} \frac{1}{(\upsilon+r_1)^{s_1}}$ and $H^n_{r_2,\ s_2} = \sum_{\upsilon=1}^{n} \frac{1}{(\upsilon+r_2)^{s_2}}$. If $r_\kappa = \frac{1}{(\kappa+r_1)^{s_1} H^n_{r_1,\ s_1}}$ and $s_\kappa = \frac{1}{(\kappa+r_2)^{s_2} H^n_{r_2,\ s_2}}$, then

$$\begin{aligned}
D(r,\ s) &= \sum_{\kappa=1}^{n} \frac{1}{(\kappa+r_1)^{s_1} H^n_{r_1,\ s_1}} \ln \frac{(\kappa+r_2)^{s_2} H^n_{r_2,\ s_2}}{(\kappa+r_1)^{s_1} H^n_{r_1,\ s_1}} \\
&= -Z(H,\ r_1,\ s_1) + \frac{s_2}{H^n_{r_1,\ s_1}} \sum_{\kappa=1}^{n} \frac{\ln(\kappa+r_2)}{(\kappa+r_1)^{s_1}} + \ln(H_{r_2,\ s_2}) \\
&\geq \sum_{s=0}^{\lambda-2} (\beta_2-\beta_1)^{2s} \int_{\beta_1}^{\beta_2} \theta_i(G_k(.,\ v)) \left(\frac{(-1)^{(2s+2)}(2s+1)!}{\beta_1^{2s+2}} F_s\left(\frac{\beta_2-v}{\beta_2-\beta_1}\right) \right. \\
&\quad \left. + \frac{(-1)^{(2s+2)}(2s+2)!}{\beta_2^{2s+2}} F_s\left(\frac{v-\beta_1}{\beta_2-\beta_1}\right)\right) dv - (Q_n S_n) \left[\frac{M-\bar{z}}{M-m}\ln(m) + \frac{\bar{z}-m}{M-m}\ln(M)\right].
\end{aligned} \tag{56}$$

Proof. Consider the functions $r_\kappa = \frac{1}{(\kappa+r_1)^{s_1} H^n_{r_1,\ s_1}}$ and $s_\kappa = \frac{1}{(\kappa+r_2)^{s_2} H^n_{r_2,\ s_2}}$. Note that $\sum_{\kappa=1}^{n} r_\kappa = \frac{H_{r_1,\ s_1}}{H_{r_1,\ s_1}} = 1$ and $\sum_{\kappa=1}^{n} s_\kappa = \frac{H^n_{r_2,\ s_2}}{H^n_{r_2,\ s_2}} = 1$. Therefore using r_κ and s_κ in (48) gives (56). □

5. Conclusions

The inequality given by Pečarić et al. [15] is generalized for higher-order convex functions using two interpolations Abel-Gontscharoff and Lidstone Polynomials at a time. New identities are constructed using the non-negative difference of the inequality given in [15] and Abel-Gontscharoff and Lidstone Polynomials. The higher-order convexity is discussed for different cases of the domains. These generalizations are a new addition to the literature to study higher-order convexity. The generalized inequality is used to estimate the entropies of information theory like; Csiszár divergence, Kullback-Leibler divergence, Shannon entropy, and Zipf-Mandelbrot law for higher order convex function. Further, the similar result in Section 3 and Section 4 of [8] can be constructed.

Funding

This Research is funded by Ongoing Research Funding Program (ORF-2025-1051), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that there are not conflict of interest.

References

[1] Pečarić JE, Tong YL. *Convex Functions, Partial Orderings, and Statistical Applications*. New York: Academic Press; 1992.

[2] Dragomir SS, Khan MA, Abathun A. Refinement of the Jensen integral inequality. *Open Mathematics*. 2016; 14(1): 221-228.

[3] Jessen B. Remarks on convex functions and inequalities between mean values. *Matematisk Tidsskrift B*. 1931; 1: 17-28.

[4] Merhav N. Reversing Jensen's inequality for information-theoretic analyses. *Information*. 2022; 13(1): 39.

[5] Butt SI, Pečarić J. *Popoviciu's Inequality for N-Convex Functions*. Saarbrücken: LAP Lambert Academic Publishing; 2016.

[6] Butt SI, Khan KA, Pečarić J. Popoviciu type inequalities via hermite's polynomial. *Mathematical Inequalities & Applications*. 2016; 19(4): 1309-1318.

[7] Butt SI, Mehmood N, Pečarić J. New generalizations of popoviciu type inequalities via new green functions and fink's identity. *Transactions of A Razmadze Mathematical Institute*. 2017; 171(3): 293-303.

[8] Butt SI, Khan KA, Pečarić J. Generalization of popoviciu inequality for higher order convex function via Taylor's polynomial. *Acta Universitatis Apulensis Mathematica-Informatica*. 2015; 42: 181-200.

[9] Butt SI, Pečarić J. Weighted popoviciu type inequalities via generalized montgomery identities. *Rad Hrvatske Akademije Znanosti i Umjetnosti Matematičke Znanosti*. 2015; 19(523): 69-89.

[10] Franjić I, Khalid S, Pečarić J. On the refinements of the Jensen-steffensen inequality. *Journal of Inequalities and Applications*. 2011; 2011(1): 12.

[11] Niaz T, Khan KA, Pečarić J. On generalization of refinement of Jensen's inequality using Fink's identity and abel-gontscharoff green function. *Journal of Inequalities and Applications*. 2017; 2017(1): 254.

[12] Khan KA, Niaz T, Pečarić Đ, Pečarić J. Estimation of different entropies via Abel-Gontscharoff Green functions and Fink's identity using Jensen type functionals. *Arabian Journal of Mathematical Sciences*. 2020; 26(1/2): 15-39.

[13] Butt SI, Pečarić Đ, Pečarić J. Several Jensen-grüss inequalities with applications in information theory. *Ukrainian Mathematical Journal*. 2023; 74(12): 1888-1908.

[14] Lah P, Ribarič M. Converse of Jensen's inequality for convex functions. *Publikacije Elektrotehničkog Fakulteta Serija Matematika i Fizika*. 1973; (412/460): 201-205.

[15] Pečarić Đ, Pečarić J, Perić J. Refinement of discrete Lah-Ribarić inequality and applications on Csiszár divergence. *Mathematics*. 2022; 10(5): 755.

[16] Pečarić J, Perić J. Refinements of the integral form of Jensen's and the Lah-Ribarić inequalities and applications for Csiszár divergence. *Journal of Inequalities and Applications*. 2020; 1: 108.

[17] Rasheed T, Butt SI, Pečarić Đ, Pečarić J, Akdemir AO. Uniform treatment of Jensen's inequality by montgomery identity. *Journal of Mathematics*. 2021; 2021(1): 5564647.

[18] Csiszár I. Axiomatic characterizations of information measures. *Entropy*. 2008; 10(3): 261-273.

[19] Csiszár I. On information-type measure of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*. 1967; 2: 299-318.

[20] Horváth L, Pečarić Đ, Pečarić J. Estimations of f- and rényi divergences by using a cyclic refinement of the Jensen's inequality. *Bulletin of the Malaysian Mathematical Sciences Society*. 2019; 42(2): 933-946.