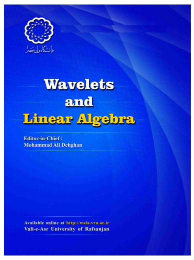
Wavelets and Linear Algebra 11(2) (2024) 76-87

Vali-e-Asr University
of Rafsanjan

Wavelets and Linear Algebra

http://wala.vru.ac.ir



Jensen's inequality and p-convex functions with Applications in information Theory

H. Barsam*[a], Y. Sayyari[b], L. Ciurdariu[c]

[a] *Department of Mathematics, Faculty of Science, University of Jiroft, Jiroft, Iran.*
[b] *Sirjan University of Technology, Sirjan, Iran.*
[c] *Politehnica University of Timisoara, Department of Mathematics, Timisoara,Romania.*

Article Info

Article history:
Received 27 April 2024
Accepted 16 november 2024
Available online 16 November 2024

Keywords:
Jensen's inequality
p-convex function
Convex function
Inequalit
Shannon entropy

2000 MSC:
26A51, 26B25, 26D20, 26D07

Abstract

In this paper we give extensions of Jensens discrete inequality considering the class of p-convex functions. We also introduce lower and upper bounds for Jensens inequality (for p-convex functions), and we apply this results in information theory and obtain new and strong bounds for Shannons entropy of a probability distribution.

*Corresponding author
Email addresses: barsam65@gmail.com (H. Barsam), ysayyari@gmail.com (Y. Sayyari), 16lciu@gmail.com (L. Ciurdariu)

https://doi.org/10.22072/wala.2024.2027512.1451

1. Introduction

The study of functions with a convex concept has been of interest to researchers for a long time. In fact, many functions have been defined derived from the concept of convexity, which can be seen as examples such as m-convex, uniformly convex; s-convex; p-convex and etc. (see [1, 5, 9, 12]). Inequalities are one of the important concepts that were created by these functions and were considered by many researchers as a branch of mathematics. Inequalities such as Hermite-Hadamard, Jensen, Hermite-Fejer, Hardy and Sherman [1, 2, 3, 11, 13] are one of the most investigated in recent years and play an important role in different areas of applied mathematics. Among these inequalities, Jensen's inequality has been used in information theory and entropy. The study of entropy problems is an active research area in dynamical systems, information theory, statistics, thermodynamics, code theory, physics, and also in some other fields of mathematics [8, 6]. The entropy have been studied by many researchers and have done so much to calculate this criterion [4, 6, 8]. But numerical calculations of entropy are still difficult. In Simic [16], Tapus and Popescu [17], the authors presented a strong upper bound for the classical Shannon entropy. In [14], the author obtain upper bound and lower bound for Shannons entropy of information sources. In Simic [14], the author obtained new and more precise bounds for Shannons entropy. An upper global bound for a differentiable convex function was given by Dragomir in [6]. In Simic [16] introduced the characteristic C(f) of a convex function f and improved previous results.

In this paper, we introduce Jensens inequality for p- convex functions and present some new and strong bounds for Shannons entropy improving the results from [16], and [17]. The paper is structured on five sections: first is an introduction where several important recent results on the inequalities which will be studied here are recalled, then in Section 2 a brief summary of the fundamental notions utilized here is given. Section 3 contain main results where several extensions of Jensens discrete inequality considering the class of p-convex functions and applications for Shannon's entropy are proved. In Section 4 new Sherman and Hardy inequalities for p-convex functions are established. Sections 5 and 6 contain aproximations and some applications. Last section is dedicated to discussions and conclusions.

2. basic notions

We start by recalling the wel-known concept of convex function and then some related notions.

Definition 2.1. Let $h : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ be a function. h is called convex if the inequality

$$h(\lambda x + (1-\lambda)y) \leq \lambda h(x) + (1-\lambda)h(y)$$

holds for every $x, y \in I$ and every $\lambda \in [0, 1]$.

Definition 2.2. ([9]) $f : [a, b] \longrightarrow \mathbb{R}$ is called a p-convex function if the inequality

$$f(tx + (1-t)y) \leq f(x) + f(y)$$

holds for every $x, y \in [a, b]$ and every $t \in [0, 1]$.

It is known that if $x_1, ..., x_n \in I$ are points, and $p_1, ..., p_n \in [0, 1]$ are coefficients such that $\sum_{i=1}^{n} p_i = 1$ then the sum

$$\sum_{i=1}^{n} p_i x_i$$

is called the convex combination of points x_i (with coefficients p_i) .

Theorem 2.3. *(**Jensen's Inequality**) Let $f : I \longrightarrow \mathbb{R}$ be a function. Then the inequality*

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1}^{n} p_i f(x_i)$$

holds for every convex combination $\sum_{i=1}^{n} p_i x_i$ of points $x_i \in I$.

Theorem 2.4. *Simic in [16] proved that if f is convex on I, then*

$$0 \leq \max_{0\leq\mu\leq\nu\leq n} \left\{ p_\mu f(x_\mu) + p_\nu f(x_\nu) - (p_\mu + p_\nu) f\left(\frac{p_\mu x_\mu + p_\nu x_\nu}{p_\mu + p_\nu}\right)\right\}$$
$$\leq \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right).$$

In order to prove main theorems, we need the following lemma that has been proved in [7].

Lemma 2.5. *Dragomir in [7] proved that for $f : I^o \to \mathbb{R}$ a differentiable function on I^o, with $a, b \in I^o$, $a < b$ for which $f' \in L[a, b]$ the following identity is satisfied:*

$$\frac{f(a) + f(b)}{2} - \frac{1}{b-a}\int_a^b f(t)dt = \frac{b-a}{2}\int_0^1 (1-2t) f'(ta + (1-t)b)dt.$$

According to [19], we recall that $y = (y_1, ..., y_n) \in \mathbb{R}^n$ is majorized by $x = (x_1, ..., x_n) \in \mathbb{R}^n$ $(y \prec x)$, when $\sum_{i=1}^{l} x_{[i]} \geq \sum_{i=1}^{l} y_{[i]}$, for $l = \overline{1, n}$ and $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$ where $x_{[1]} \geq ... \geq x_{[n]}$, $y_{[1]} \geq ... \geq y_{[n]}$ are the component of x and y rearranged in decreasing order.

In addition in [19] are defined the row stochastic matrix as the $n \times m$ real matrix $P = (p_{ij})$ for which $p_{ij} \geq 0$ for $i = \overline{1, n}$, $j = \overline{1, m}$ and all row sums of P are equal to 1. Similary, the column stochastic matrix are the $n \times m$ real matrix $P = (p_{ij})$ for which $p_{ij} \geq 0$ for $i = \overline{1, n}$, $j = \overline{1, m}$ but, this time, all column sums of P are equal to 1. P is said to be doubly stochastic matrix if P is column stochastic and also row stochastic.

By using the $n \times n$ doubly stochastic matrix, in [19], we have that for any $x, y \in \mathbb{R}^n$, $y \prec xl$ iff $y = Px$.

Theorem 2.6. *(HardyLittlewoodPlya) According to Hardy et all, [10], if $f : J \subset \mathbb{R} \to \mathbb{R}$ is a real convex function, and $x = (x_1, \cdots, x_n)^T \in J^n$, $y = (y_1, \cdots, y_n)^T \in J^n$ so that $y \prec x$ then we have*

$$\sum_{i=1}^{n} f(x_i) \geq \sum_{i=1}^{n} f(y_i).$$

The below Sherman's theorem is a generalization of Theorem 2.6.

Theorem 2.7. *Sherman [15] proved that for a real convex function $f : J \subset \mathbb{R} \to \mathbb{R}$ defined on an interval J and $a = (a_1, \cdots, a_m)^T \in \mathbb{R}^m$, $b = (b_1, \cdots, b_n)^T \in \mathbb{R}^n$, $x = (x_1, \cdots, x_m)^T \in J^m$, $y = (y_1, \cdots, y_n)^T \in J^n$ if $y = Sx$ and $a = S^T b$ for some $n \times m$ row stochastic matrix $S = (s_{ij})$, then*

$$\sum_{j=1}^{m} a_j f(x_j) \geq \sum_{i=1}^{n} b_i f(y_i).$$

3. main results

In this section, we present some new results on p-convex function.

Theorem 3.1. *Let f be a p-convex function and $0 \leq r < s \leq n$ are arbitrary, then*

$$0 \leq \max_{1\leq r<s\leq n} \left\{ f(x_r) + f(x_s) - f\left(\frac{p_r x_r + p_s x_s}{p_r + p_s}\right) \right\} \leq \sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right).$$

Proof.

$$\begin{aligned} f\left(\sum_{i=1}^{n} p_i x_i\right) &= f\left(\sum_{i=1, i\neq r,s}^{n} p_i x_i + (p_r + p_s)\left(\frac{p_r x_r + p_s x_s}{p_r + p_s}\right)\right) \\ &\leq \sum_{i=1, i\neq r,s}^{n} f(x_i) + f\left(\frac{p_r x_r + p_s x_s}{p_r + p_s}\right), \end{aligned}$$

so, we have

$$\sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \geq f(x_r) + f(x_s) - f\left(\frac{p_r x_r + p_s x_s}{p_r + p_s}\right)$$

which completes the proof. □

Theorem 3.2. *Let f be a p-convex function and $0 \leq r < s \leq n$ are arbitrary, then*

$$\sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \leq (n+1)(f(a) + f(b)) - f\left(\frac{a+b}{2}\right).$$

Proof.

$$\begin{aligned}\sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) &= \sum_{i=1}^{n} f(\lambda_i a + (1-\lambda_i)b) \\ &\quad - f\left(\sum_{i=1}^{n} p_i(\lambda_i a + (1-\lambda_i)b)\right) \\ &\leq n\left(f(a) + nf(b)\right) - f\left(a\left(\sum_{i=1}^{n} \lambda_i p_i\right) + b\left(1 - \sum_{i=1}^{n} \lambda_i p_i\right)\right) \\ &= n(f(a) + f(b)) - f(pa + qb) \\ &\leq (n+1)(f(a) + f(b)) - f(pa+qb) - f(qa+pb) \\ &\leq (n+1)(f(a) + f(b)) - f\left(\frac{pa+qb}{2} + \frac{qa+pb}{2}\right) \\ &\leq (n+1)(f(a) + f(b)) - f\left(\frac{a+b}{2}\right)\end{aligned}$$

which $p = \sum_{i=1}^{n} \lambda_i p_i, q = 1 - \sum_{i=1}^{n} \lambda_i p_i$. □

Theorem 3.3. *Let f be a p-convex function and $0 \leq r < s \leq n$ are arbitrary, then*

$$\begin{aligned}f(\mu) + f(\nu) - f\left(\frac{\nu+\mu}{2}\right) &\leq \sum_{i=1}^{n} f(x_i) - f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \\ &\leq (n+1)(f(\mu) + f(\nu)) - f\left(\frac{\mu+\nu}{2}\right).\end{aligned}$$

Proof. In Theorem 3.1, put $x_r = \mu, x_s = \nu, p_1 = \cdots = p_n = \frac{1}{n}$ the left inequality is obtained. For right inequality. Set $a = \mu, b = \nu, p_1 = \cdots = p_n = \frac{1}{n}$ in Theorem 3.2. □

Lemma 3.4. *Assume that $a \geq e^{\frac{3}{2}}$. The function $\frac{\log x}{x} : [a, b] \to \mathbb{R}$ is p-convex.*

Proof. Assume that $a \geq e^{\frac{3}{2}}$, $x, y \geq a$. Define $h : [0, 1] \to \mathbb{R}$,

$$h(t) := \frac{\log(tx + (1-t)y)}{tx + (1-t)y} - \frac{\log x}{x} - \frac{\log y}{y}.$$

Now, we have $h(0) \leq 0, h(1) \leq 0$. Also,

$$\begin{aligned}\frac{d^2h}{dt^2} &= (x-y)^2(tx + (1-t)y)\frac{2\log(tx + (1-t)y) - 3}{(tx + (1-t)y)^4} \\ &= (x-y)^2\frac{2\log(tx + (1-t)y) - 3}{(tx + (1-t)y)^3} \geq 0.\end{aligned}$$

Hence, $h(t) \leq 0$. We conclude that $\frac{\log x}{x}$ is a p-convex function. □

Lemma 3.5. *Assume that* $0 < p_i \leq e^{-\frac{3}{2}}, \mu = \min p_i, \nu = \max p_i, \quad P = \{p_1, \cdots, p_n\}$ *then*

$$H(P) - \frac{\log n}{n} \geq \mu \log \frac{\mu + \nu}{2\mu} + \nu \log \frac{\nu + \mu}{2\nu}.$$

Proof. In Theorem 3.1 put $f(x) = \frac{\log x}{x}$, $x_i = \frac{1}{p_i}$. Now using Lemma 3.4 we have

$$-\sum p_i \log p_i - \frac{\log n}{n} \geq -\mu \log \mu - \nu \log \nu - \frac{\nu + \mu}{2} \log \frac{2}{\mu + \nu}.$$

So, we have

$$H(P) - \frac{\log n}{n} \geq \mu \log \frac{\mu + \nu}{2\mu} + \nu \log \frac{\mu + \nu}{2\mu},$$

which completes the proof. □

Lemma 3.6. *Assume that* $0 < a < b \leq 1$. *Then the function* $f(x) = (x-1)\log x$ *is* $p-$*convex.*

Proof. For $a \leq x, y \leq b$ the function $h : [0,1] \to \mathbb{R}$ define by

$$h(t) := (tx + (1-t)y - 1)\log(tx + (1-t)y) - (x-1)\log x - (y-1)\log y$$

we have $h(0) = -(x-1)\log x < 0$, $h(1) = -(y-1)\log y < 0$ Also,

$$h''(t) = (x-y)^2 \left[\frac{1}{tx + (1-t)y} + \frac{1}{(tx + (1-t)y)^2} \right] \geq 0.$$

So, $h(t) \leq 0$ for each $0 \leq t \leq 1$ hence $f(x) = (x-1)\log x$ is a p-convex function. □

Theorem 3.7. *Assume that* $\mu = \min p_i, \ \nu = \max p_i$ *then*

$$\begin{aligned} &\left(1 - \frac{\mu + \nu}{2}\right) \log \frac{2}{\mu + \nu} + (n+1)\left[(1-\mu)\log \mu + (1-\nu)\log \nu\right] \\ &\leq \frac{n-1}{n} \log n + H(P) + \log \prod_{i=1}^{n} p_i \\ &\leq \mu \log \frac{\sqrt{\mu + \nu}}{\mu\sqrt{2}} + \nu \log \frac{\sqrt{\mu + \nu}}{\nu\sqrt{2}} + \log \frac{2\mu\nu}{\mu + \nu}. \end{aligned}$$

Proof. In Theorem 3.3, put $f(x) = (x-1)\log x$, $x_i = p_i$. Now using Lemma 3.6 we have

$$\begin{aligned} &(\mu - 1)\log \mu + (\nu - 1)\log \nu - \left(\frac{\nu + \mu}{2} - 1\right) \log \left(\frac{\mu + \nu}{2}\right) \\ &\leq \sum_{i=1}^{n} (p_i - 1)\log p_i - \left(\frac{1}{n} - 1\right) \log \left(\frac{1}{n} - 1\right) \\ &\leq (n+1)((\mu - 1)\log \mu + (\nu - 1)\log \nu) - \left(\frac{\nu + \mu}{2} - 1\right) \log \left(\frac{\mu + \nu}{2}\right), \end{aligned}$$

now with some calculation, the result is obtained. □

An analogue of Tapus and Popescu result, see [17] for p-convex functions, which is a refinement of Simic's inequality [16], is given below.

Theorem 3.8. *Let f be a p-convex function on an interval I, $x_i \in I$, $1 \le i \le n$ and $\sum_{i=1}^{n} p_i = 1$. Then we have*

$$\begin{aligned} 0 &\le \max_{1\le\mu<\nu\le n}\left[f(x_\mu)+f(x_\nu)-f\left(\frac{p_\nu x_\nu+p_\mu x_\mu}{p_\nu+p_\mu}\right)\right] \\ &\le \max_{1\le\mu<\nu<\eta\le n}\left[f(x_\mu)+f(x_\nu)+f\left(x_\eta\right)-f\left(\frac{p_\nu x_\nu+p_\mu x_\mu+p_\eta x_\eta}{p_\nu+p_\mu+p_\eta}\right)\right] \\ &\le \sum_{i=1}^{n} f(x_i)-f\left(\sum_{i=1}^{n} p_i x_i\right). \end{aligned}$$

Proof. As in the proof of Theorem in [17] we have to show that for all $t \in \{1,2,...,n\}-\{r,s\}$

$$f(x_r)+f(x_s)-f\left(\frac{p_s x_s+p_r x_r}{p_r+p_s}\right) \le f(x_r)+f(x_e)+f(x_t)-f\left(\frac{p_r x_r+p_s x_s+p_t x_t}{p_r+p_s+p_t}\right),$$

which is equivalent to

$$f(x_t)+f\left(\frac{p_s x_s+p_r x_r}{p_r+p_s}\right) \ge f\left(\frac{p_r x_r+p_s x_s+p_t x_t}{p_r+p_s+p_t}\right).$$

But we have

$$\frac{p_r x_r+p_s x_s+p_t x_t}{p_r+p_s+p_t} = \frac{p_s x_s+p_r x_r}{p_r+p_s}\,\frac{p_r+p_s}{p_r+p_s+p_t}+x_t\frac{p_t}{p_r+p_s+p_t}$$

and using now the definition pf p-convex functions, we get

$$f\left(\frac{p_r x_r+p_s x_s+p_t x_t}{p_r+p_s+p_t}\right) \le f(x_t)+f\left(\frac{p_s x_s+p_r x_r}{p_r+p_s}\right).$$

Taking into account that we consider the maximum $1 \le \mu < \nu < \eta \le n$, of the expression

$$f(x_\mu)+f(x_\nu)+f(x_\eta)-f\left(\frac{p_\nu x_\nu+p_\mu x_\mu+p_\eta x_\eta}{p_\nu+p_\mu+p_\eta}\right)$$

which is greater or equal than

$$f(x_r)+f(x_s)+f(x_t)-f\left(\frac{p_r x_r+p_s x_s+p_t x_t}{p_r+p_s+p_t}\right)$$

and then the maximum of first expression we check the first inequality. For the second part,

$$\max_{1\le\mu<\nu<\eta\le n}\left[f(x_\mu)+f(x_\nu)+f(x_\eta)-f\left(\frac{p_\nu x_\nu+p_\mu x_\mu+p_\eta x_\eta}{p_\nu+p_\mu+p_\eta}\right)\right] \le \sum_{i=1}^{n} f(x_i)-f\left(\sum_{i=1}^{n} p_i x_i\right)$$

we take arbitrary three different items $r, s, t \in \{1, 2, ..., n\}$ and we will prove that

$$f(x_r) + f(x_s) + f(x_t) - f\left(\frac{p_r x_r + p_s x_s + p_t x_t}{p_r + p_s + p_t}\right) \leq \sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right).$$

Previous inequality is equivalent to

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1, i\neq r,s,t}^{n} f(x_i) + f\left(\frac{p_r x_r + p_s x_s + p_t x_t}{p_r + p_s + p_t}\right)$$

which results directly from the definition of p-convex functions, using the expression

$$\sum_{i=1}^{n} p_i x_i = \sum_{i=1, i\neq r,s,t}^{n} p_i x_i + (p_r + p_s + p_t)\frac{p_r x_r + p_s x_s + p_t x_t}{p_r + p_s + p_t}.$$

This expression thus becomes,

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1, i\neq r,s,t}^{n} f(x_i) + f\left(\frac{p_r x_r + p_s x_s + p_t x_t}{p_r + p_s + p_t}\right).$$

It is known that if f is p-convex then $f(\sum_{i=1}^{n} p_i x_i) \leq \sum_{i=1}^{n} f(x_i)$, where $\sum_{i=1}^{n} p_i = 1$. The proof can be given by mathematical induction for example. Hypothesis is satisfied for $n = 2$ by the definition of p-convexity. We suppose the affirmation true for $n - 1$ and we check for n. We write

$$\begin{aligned} f(\sum_{i=1}^{n} p_i x_i) &= f\left(\sum_{i=1}^{n-1} p_i x_i + p_n x_n\right) \\ &= f\left(\sum_{i=1}^{n-1} \frac{p_i x_i}{p_1 + ... + p_{n-1}}(p_1 + ... + p_{n-1}) + p_n x_n\right) \\ &\leq f\left(\sum_{i=1}^{n-1} \frac{p_i x_i}{p_1 + ... + p_{n-1}}(p_1 + ... + p_{n-1})\right) + f(x_n) \leq \sum_{i=1}^{n} f(x_i). \end{aligned}$$

□

Theorem 3.9. *If f is a p-convex function on an interval I, $1 \leq i \leq n$ and $\sum_{i=1}^{n} p_i = 1$ then*

$$0 \leq P_2 \leq P_3 \leq ... \leq P_{n-1} \leq \sum_{i=1}^{n} f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right),$$

where

$$P_{n-1} = \max_{1\leq \mu_1 < \mu_2 < ... < \mu_{n-1} \leq n} \left[\sum_{k=1}^{n-1} f(x_{\mu_k}) - f\left(\frac{\sum_{k=1}^{n-1} p_{\mu_k} x_{\mu_k}}{\sum_{k=1}^{n-1} p_{\mu_k}}\right)\right]$$

and $\mu_1, \mu_2, ..., \mu_{n-1} \in \{1, 2, ..., n\}$ are different items.

Proof. We follows the same steps as in the proof of Theorem 3.8. □

We consider the probability distribution F given by $P(X = i) = p_i, \ p_i > 0, \ 1 \leq i \leq n$ with $\sum_{i=1}^{n} p_i = 1$. The Shannon's entropy, is given by $H(P) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$.

Theorem 3.10. *Let $P = \{p_1, ..., p_n\}$ with $\sum_{i=1}^{n} p_i = 1$. Then we have,*

$$H(P) - \log n \geq \max_{1 \leq \mu_1 < \mu_2 < ... < \mu_{n-1} \leq n} \log \left[\frac{\prod_{k=1}^{n-1} \frac{1}{p_{\mu_k}^{p_{\mu_k}}}}{\frac{n-1}{\sum_{k=1}^{n-1} p_{\mu_k}}} \right].$$

Proof. We take $f(x) = \frac{\log x}{x}$, $x_i = \frac{1}{p_i}$, $1 \leq i \leq n$ in the last inequality because this function is p-convex, see Lemma 3.4 for $a \geq e^{\frac{3}{2}}$ and after some small calculus we get the desired result. □

Lemma 3.11. *(a)The function $f : (0, 1] \to \mathbb{R}$ defined by $f(x) = -\log x$ is a p-convex function. (b) The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = (x-1)^2$ is a p-convex function.*

Proof. (a) If we define $h : [0, 1] \to \mathbb{R}$ by $h(t) = \log(tx+(1-t)y)+\log x+\log y$ then $h(0) = \log x < 0$ and $h(1) = \log y < 0$. In addition,

$$\frac{d^2h}{dt^2} = \frac{(x-y)^2}{(tx + (1-t)y)^2} > 0$$

hence $h(t) \leq 0$ i.e. f is a p-convex function. (b) If we define $h : [0, 1] \to \mathbb{R}$ by $h(t) = (tx + (1 - t)y - 1)^2 - (x-1)^2 - (y-1)^2$ then $h(0) = -(x-1)^2 < 0$ and $h(1) = -(y-1)^2 < 0$. In addition,

$$\frac{d^2h}{dt^2} = 2(x-y)^2 > 0$$

hence $h(t) \leq 0$ i.e. f is a p-convex function. □

4. Sherman and Hardy inequalities for *p*-convex function

Theorem 4.1. *Assume that $y \prec x$, $f : J \to \mathbb{R}$ is a p-convex function. Also, $x, y \in J^n, y = Px$ then*

$$\sum_{i=1}^{n} f(y_i) \leq n \sum_{i=1}^{n} f(x_i)$$

where P is a stochastic matrix.

Proof. Since $y \prec x$ hence $y = Px$ so, $y = \sum_{j=1}^{n} p_{ij} x_j$. Now, by the definition of p-convexity, we have

$$f(y_i) = f\left(\sum_{j=1}^{n} p_{ij} x_j\right) \leq \sum_{j=1}^{n} f(x_j)$$

this conclude

$$\sum_{i=1}^{n} f(y_i) \leq \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_j) = n \sum_{i=1}^{n} f(x_j).$$

That completes the proof. □

Theorem 4.2. *Assume that $y \prec x$, $f : J \to \mathbb{R}$ is a p-convex function. Also, $y \in J^n, x \in J^m, b \in \mathbb{R}^n, y = Px$ then*

$$\sum_{i=1}^{n} b_i f(y_i) \leq \left(\sum_{i=1}^{n} b_i\right)\left(\sum_{j=1}^{m} f(x_j)\right).$$

Proof. Since $y \prec x$ hence $y = Px$ (P is a row stochastic matrix) so, $y = \sum_{j=1}^{m} p_{ij} x_j$. Now, by the same reason, we have

$$f(y_i) = f\left(\sum_{j=1}^{m} p_{ij} x_j\right) \leq \sum_{j=1}^{m} f(x_j),$$

also,

$$b_i f(y_i) \leq \sum_{j=1}^{m} b_i f(x_j),$$

this conclude

$$\sum_{i=1}^{n} b_i f(y_i) \leq \sum_{i=1}^{n}\sum_{j=1}^{m} b_i f(x_j) = \left(\sum_{i=1}^{n} b_i\right)\left(\sum_{j=1}^{m} f(x_j)\right).$$

That completes the proof. □

5. Approximation with *p*-convex functions

Theorem 5.1. *Let $f : [a, b] \to \mathbb{R}$ be a function. If $|f'|$ is a p-convex function then*

$$\left|\frac{f(a) + f(b)}{2} - \frac{1}{b-a}\int_a^b f(x)dx\right| \leq \frac{b-a}{4}\left[|f'(a)| + |f'(b)|\right].$$

Proof. Using Lemma 2.5 we have

$$\begin{aligned}\left|\frac{f(a) + f(b)}{2} - \frac{1}{b-a}\int_a^b f(x)dx\right| &\leq \frac{b-a}{2}\int_0^1 |1-2t||f'(ta + (1-t)b)|dt \\ &\leq \frac{b-a}{2}\left(|f'(a)| + |f'(b)|\right)\int_0^1 |1-2t|dt \\ &= \frac{b-a}{4}\left(|f'(a)| + |f'(b)|\right).\end{aligned}$$

Note that

$$\int_0^1 |1-2t|dt = \frac{1}{2}.$$

□

Theorem 5.2. *Let $f : [a,b] \to \mathbb{R}$ be a function and $p,q \in \mathbb{R}$ be two positive numbers with $\frac{1}{p} + \frac{1}{q} = 1$ and $p > 1$. If $|f'|^q$ is p-convex then*

$$\left|\frac{f(a)+f(b)}{2} - \frac{1}{b-a}\int_a^b f(x)dx\right| \leq \frac{b-a}{2(p+1)^{\frac{1}{p}}}\left[|f'(a)|^q + |f'(b)|^q\right].$$

Proof. In view of Holder inequality, p-convexity of $|f'|^q$ and of Lemma 2.5, we conclude that

$$\begin{aligned}\left|\frac{f(a)+f(b)}{2} - \frac{1}{b-a}\int_a^b f(x)dx\right| &\leq \frac{b-a}{2}\left(\int_0^1 |1-2t|^p dt\right)^{\frac{1}{p}}\left(\int_0^1 |f'(ta+(1-t)b)|^q\right)^{\frac{1}{q}}\\ &\leq \frac{b-a}{2}\times\left(\frac{1}{p+1}\right)^{\frac{1}{p}}\times\left(|f'(a)|^q + |f'(b)|^q\right)^{\frac{1}{q}}.\end{aligned}$$

We used here that

$$\int_0^1 |1-2t|^p dt = \frac{1}{p+1}.$$

. □

That completes the proof.

6. Some applications of *p*-convex functions

Proposition 6.1. *Let $a,b \in \mathbb{R}$ such that $0 \leq a < b$ then*

$$b \leq a + \frac{6(a^q+b^q)^{\frac{1}{q}}}{(p+1)^{\frac{1}{p}}}.$$

Proof. In view of the Lemma 3.11, $f(x) = x^2$ is a p-convex function. Now using Theorem 5.2 we have

$$\begin{aligned}\left|\frac{a^2+b^2}{2} - \frac{b^3-a^3}{3(b-a)}\right| &= \frac{(b-a)^2}{6}\\ &\leq \frac{b-a}{2(p+1)^{\frac{1}{p}}}\times\left(2^q a^q + 2^q b^q\right)^{\frac{1}{q}}\\ &\leq \frac{b-a}{(p+1)^{\frac{1}{p}}}\left(a^q+b^q\right)^{\frac{1}{q}},\end{aligned}$$

which completes the proof. □

Proposition 6.2. *Let $a,b \in \mathbb{R}$ such that $0 \leq a < b \leq 1$. Then we get*

$$\left|ab\frac{\ln b - \ln a}{b-a} - \frac{a+b}{2}\right| \leq \frac{a-b}{2}\ln(ab).$$

Proof. Put $f(x) = x - x\ln x$ in Theorem 5.1 and using Lemma 3.11(b) we find the desired result. □

Proposition 6.3. *Let $a, b \in \mathbb{R}$ such that $0 \leq a < b$. Then we have*

$$\left|\frac{(a-1)^3+(b-1)^3}{6}-\frac{(b-1)^4-(a-1)^4}{12(b-a)}\right| \leq \frac{b-a}{4}\left((a-1)^2+(b-1)^2\right).$$

Proof. Put $f(x) = \frac{(x-1)^3}{3}$ in Theorem 5.1 and using Lemma 3.11(a) we obtain the desired result. □

7. Conclusions

New refinements of Jensens discrete inequality for the class of p-convex functions were presented in this paper and then we apply this results in information theory for obtaining new and strong bounds for Shannons entropy of a probability distribution. In addition, Sherman's and Hardy's type inequalities for p- convex functions were given and also some applications.

References

[1] H. Barsam and A. R. Sattarzadeh, Hermite-Hadamard inequalities for uniformly convex functions and Its Applications in Means,Miskolc Math. Notes, 21(2) (2020), 621-630.

[2] H. Barsam, Y. Sayyari, On some inequalities of differentable uniformly convex mapping with applications. Numer. Funct. Anal. Optim. 44(2), (2023),368381.

[3] H. Barsam, S. M. Ramezani, Y. Sayyari, On the new HermiteHadamard type inequalities for s-convex functions. Afr. Mat. 13(2021), 13551367.

[4] I. Budimir, S.S Dragomir, J. Pecaric, Further reverse results for Jensens discrete inequality and applications in information theory. J Inequal Pure Appl Math, 2(1).(2001) Art. 5.

[5] S.S. Dragomir and C.E.M. Pearce, Selected Topics on Hermite-Hadamard Inequalities and Applications, RGMIA Monographs, Victoria University, 2000.

[6] S.S. Dragomir, A converse result for Jensen's discrete inequality via Grss inequality and applications in information theory, An. Univ. Oradea. Fasc. Mat.7 (1999-2000), 178-189.

[7] S. Dragomir and R. Agarwal, Two inequalities for differentiable mappings and applications to special means of real numbers and to trapezoidal formula, Applied Mathematics Letters, (11) 5(1998) 91–95.

[8] S.S. Dragomir and C.J. Goh, Some bounds on entropy measures in Information Theory, Appl. Math. Lett. 10 (3) (1997) 23-28.

[9] S. S. Dragomir, J. Pecaric, and L. E. Persson, Some inequalities of Hadamard type, Soochow Journal of Mathematics, 21 (3)(1995)335341.

[10] G. M. Hardy, J. E. Littlewood, G. Plya: Inequalities, 2nd edn. Cambridge University Press, Cambridge (1952).

[11] Y. Sayyari and H. Barsam, Hermite-Hadamard type inequality for m-convex functions by using a new inequality for differentiable functions, J. Mahani Math. Res. Cent. 9(2), (2020), 55-67.

[12] Y. Sayyari, H. Barsam, Jensen-Mercer inequality for uniformly convex functions with some applications, Afr. Mat. (2023), DOI:https://doi.org/10.1007/s13370-023-01084-2.

[13] Y. Sayyari, H. Barsam, A.R. Sattarzadeh, On new refinement of the Jensen inequality using uniformly convex functions with applications. Appl. Anal. (2023),1-9. https://doi.org/10.1080/00036811.2023.2171873.

[14] Y. Sayyari, New bounds for entropy of information sources. Wavelet and Linear Algebra 7(2) (2020) 1-9.

[15] Sherman, S.: On a theorem of Hardy, Littlewood, Plya, and Blackwell. Proc. Natl. Acad. Sci. USA 37(1951) 826831.

[16] S. Simic, Jensens inequality and new entropy bounds, Appl. Math. Lett. 22 (8) (2009), 1262-1265.

[17] N. Tapus, PG. Popescu, A new entropy upper bound. Appl Math Lett, 25(11) (2012) 18871890.

[18] A.W. Marshall, I. Olkin, Inequalities:Theory of Majorization and its Applications, Academic Press, New York, 1979.

[19] N., Niezgoda, A Jensen-Sherman type inequality with a control map for G-invariant convex functions, Positivity, 25:431-446, 2021.