Research Article

Lei Xiao and Guoxiang Lu*

A new refinement of Jensen's inequality with applications in information theory

https://doi.org/10.1515/math-2020-0123
received January 26, 2020; accepted November 25, 2020

Abstract: In this paper, we present a new refinement of Jensen's inequality with applications in information theory. The refinement of Jensen's inequality is obtained based on the general functional in the work of Popescu et al. As the applications in information theory, we provide new tighter bounds for Shannon's entropy and some f-divergences.

Keywords: refinements, Jensen's inequality, information theory, Shannon's entropy, f-divergences, bounds

MSC 2020: 26B25, 26D15, 94A17

1 Introduction

Let C be a convex subset of the linear space X and f a convex function on C. If $\mathbf{p} = (p_1, \ldots, p_n)$ is a probability sequence and $\mathbf{x} = (x_1, \ldots, x_n) \in C^n$, then the well-known Jensen's inequality

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

holds [1]. If f is concave, then the preceding inequality is reversed.

Jensen's inequality probably plays a crucial role in the theory of mathematical inequalities. It is applied widely in mathematics, statistics, and information theory and can deduce many important inequalities such as arithmetic-geometric mean inequality, Hölder inequality, Minkowski inequality, and Ky Fan's inequality.

In 2010, Dragomir obtained a refinement of Jensen's inequality as follows [2]:

Theorem 1.1. *If f, $\mathbf{x}$, $\mathbf{p}$ are defined as above, then*

$$\begin{aligned} f\left(\sum_{i=1}^{n} p_i x_i\right) &\le \min_{k\in\{1,2,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{i=1}^{n} p_i x_i - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \\ &\le \frac{1}{n}\left[\sum_{k=1}^{n} (1-p_k) f\left(\frac{\sum_{i=1}^{n} p_i x_i - p_k x_k}{1-p_k}\right) + \sum_{k=1}^{n} p_k f(x_k)\right] \\ &\le \max_{k\in\{1,2,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{i=1}^{n} p_i x_i - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \le \sum_{i=1}^{n} p_i f(x_i). \end{aligned} \tag{2}$$

* **Corresponding author: Guoxiang Lu,** School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, P. R. China, e-mail: lgxmath@zuel.edu.cn
Lei Xiao: School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, P. R. China, e-mail: xiaoleiyouhua@sina.com

The same year, Dragomir has also obtained a different refinement of Jensen's inequality as follows [3]:

Theorem 1.2. (S. S. Dragomir) *Let C be a convex subset in the real linear space X and assume that* $f : C \to \mathbb{R}$ *is a convex function on C. If* $x_k \in C$ *and* $p_k > 0$, $k \in \{1, 2, \ldots, n\}$ *with* $\sum_{k=1}^{n} p_k = 1$, *then for any nonempty subset J of* $\{1, 2, \ldots, n\}$, *we have*

$$\sum_{k=1}^{n} p_k f(x_k) \geq D(f, \mathbf{p}, \mathbf{x}; J) \geq f\left(\sum_{k=1}^{n} p_k x_k\right), \tag{3}$$

where $D(f, \mathbf{p}, \mathbf{x}; J)$ *is a functional defined as follows:*

$$D(f, \mathbf{p}, \mathbf{x}; J) := P_J f\left(\frac{1}{P_J}\sum_{i\in J} p_i x_i\right) + \bar{P}_J f\left(\frac{1}{\bar{P}_J}\sum_{j\in \bar{J}} p_j x_j\right)$$

with J a nonempty subset of $\{1, 2, \ldots, n\}$, $\bar{J} := \{1, 2, \ldots, n\} \setminus J$, $P_J := \sum_{i\in J} p_i$ *and* $\bar{P}_J := P_{\bar{J}} = \sum_{j\in\bar{J}} p_j = 1 - \sum_{i\in J} p_i$, *where* $J \neq \{1, 2, \ldots, n\}$.

It is easy to find that if $J = \{k\}$, then inequalities (3) imply inequalities (2).

In 2016, Popescu et al. defined a new refined functional as follows [4]:

$$D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_m) := \sum_{i=1}^{m} P_{J_i} f\left(\frac{1}{P_{J_i}}\sum_{j\in J_i} p_j x_j\right) + \bar{P}_J f\left(\frac{1}{\bar{P}_J}\sum_{j\in\bar{J}} p_j x_j\right),$$

where $J_1, J_2, \ldots, J_m$ are nonempty, pairwise disjoint subsets of J, with $J = \bigcup_i J_i$ and $P_{J_i} := \sum_{j\in J_i} p_j$. It is easy to observe that

$$\sum_{i=1}^{m} P_{J_i} + \bar{P}_J = 1,$$

and in order to make sense, m should be less or equal with the cardinal of J, that is, $1 \leq m \leq |J|$. *If* $m = 1$, then

$$D(f, \mathbf{p}, \mathbf{x}; J, J_1) = D(f, \mathbf{p}, \mathbf{x}; J).$$

Then Theorem 1.2 can be generalized as follows:

Theorem 1.3. (P. G. Popescu et al.) *Let C be a convex subset in the real linear space X and assume that* $f : C \to \mathbb{R}$ *is a convex function on C. If* $x_k \in C$ *and* $p_k > 0$, $k \in \{1, 2, \ldots, n\}$ *with* $\sum_{k=1}^{n} p_k = 1$, *then for any nonempty subset J of* $\{1, 2, \ldots, n\}$, *we have*

$$\sum_{k=1}^{n} p_k f(x_k) \geq D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_m) \geq D(f, \mathbf{p}, \mathbf{x}; J) \geq f\left(\sum_{k=1}^{n} p_k x_k\right), \tag{4}$$

where $J_1, J_2, \ldots, J_m$ *are nonempty, pairwise disjoint subsets of J, with* $J = \bigcup_i J_i$ *and* $m \leq |J|$.

In [5], Horváth developed a general method to refine the discrete Jensen's inequality in the convex and mid-convex cases. The main part of the inequalities in Theorems 1.2 and 1.3 are special cases of Theorem 1 in the paper. Recently, Horváth et al. [6] presented new upper bounds for the Shannon entropy (see Corollary 1) and defined an extended *f*-divergence functional (see Definition 2) by applying a cyclic refinement of Jensen's inequality. For more other refinements and applications related to Jensen's inequality, see [7–17].

The main aim of this paper is to extend the results of Dragomir [3] and Popescu et al. [4] by the aforementioned functional. In Section 2, we give refinement of Jensen's inequality associated with the general functionals. The refinement demonstrates some estimates of Jensen's gap and tightens the inequalities (4). In Section 3, we show the applications in information theory. We propose and prove new tighter upper bounds for Shannon's entropy compared to the bound given in [4]. At last, we obtain new bounds for some *f*-divergences better than the bounds given in [3].

2 General inequalities by generalization

We continue to use the aforementioned definition and show the main results.

Theorem 2.1. *Let C be a convex subset in the real linear space X and assume that $f : C \to \mathbb{R}$ is a convex function on C. If $x_k \in C$ and $p_k > 0, k \in \{1, 2, \ldots, n\}$ with $\sum_{i=1}^{n} p_k = 1$, then for any nonempty subset J of $\{1, 2, \ldots, n\}$, we have*

$$\begin{aligned}\sum_{k=1}^{n} p_k f(x_k) &\geq \max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_{m+1}) \\ &\geq \max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_m) \geq f\left(\sum_{k=1}^{n} p_k x_k\right).\end{aligned} \tag{5}$$

Proof. We assume the value of $\max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_m)$ is obtained for $J_i = J_i^{(m)}, 1 \leq i \leq m$.

If $m + 1 = n$ and each subset $J_i^{(m)} (1 \leq i \leq m)$ and $\bar{J}^{(m)}$ contain one element, then we can easily obtain that the inequalities (5) hold as follows:

$$\begin{aligned}\sum_{k=1}^{n} p_k f(x_k) &= \max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_{m+1}) \\ &= \max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_m) \geq f\left(\sum_{k=1}^{n} p_k x_k\right).\end{aligned}$$

Otherwise, there exists a subset $J_i^{(m)} (1 \leq i \leq m)$ or $\bar{J}^{(m)}$, which contains more than one element. Without loss of generality we assume $J_m^{(m)}$ contains more than one element. Then we find two nonempty subsets $J_m^{(m+1)}, J_{m+1}^{(m+1)}$ such that $J_m^{(m+1)} \cup J_{m+1}^{(m+1)} = J_m^{(m)}$ and $J_m^{(m+1)} \cap J_{m+1}^{(m+1)} = \emptyset$. By using Jensen's inequality, we have

$$\begin{aligned}&\frac{P_{J_m^{(m+1)}}}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} f\left(\frac{1}{P_{J_m^{(m+1)}}} \sum_{j \in J_m^{(m+1)}} p_j x_j\right) + \frac{P_{J_{m+1}^{(m+1)}}}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} f\left(\frac{1}{P_{J_{m+1}^{(m+1)}}} \sum_{j \in J_{m+1}^{(m+1)}} p_j x_j\right) \\ &\geq f\left(\frac{P_{J_m^{(m+1)}}}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} \cdot \frac{1}{P_{J_m^{(m+1)}}} \sum_{j \in J_m^{(m+1)}} p_j x_j + \frac{P_{J_{m+1}^{(m+1)}}}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} \cdot \frac{1}{P_{J_{m+1}^{(m+1)}}} \sum_{j \in J_{m+1}^{(m+1)}} p_j x_j\right) \\ &= f\left(\frac{1}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} \sum_{j \in J_m^{(m+1)} \cup J_{m+1}^{(m+1)}} p_j x_j\right).\end{aligned}$$

The aforementioned inequality can be rewritten as:

$$P_{J_m^{(m+1)}} f\left(\frac{1}{P_{J_m^{(m+1)}}} \sum_{j \in J_m^{(m+1)}} p_j x_j\right) + P_{J_{m+1}^{(m+1)}} f\left(\frac{1}{P_{J_{m+1}^{(m+1)}}} \sum_{j \in J_{m+1}^{(m+1)}} p_j x_j\right) \geq P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}} f\left(\frac{1}{P_{J_m^{(m+1)} \cup J_{m+1}^{(m+1)}}} \sum_{j \in J_m^{(m+1)} \cup J_{m+1}^{(m+1)}} p_j x_j\right).$$

So let $J_i^{(m+1)} = J_i^{(m)}, 1 \leq i \leq m - 1$, we can deduce that

$$\begin{aligned}&\max_{\emptyset \neq J \subset \{1,\ldots,n\}} D(f, \mathbf{p}, \mathbf{x}; J, J_1, J_2, \ldots, J_{m+1}) \\ &\geq D(f, \mathbf{p}, \mathbf{x}; J^{(m+1)}, J_1^{(m+1)}, J_2^{(m+1)}, \ldots, J_{m-1}^{(m+1)}, J_m^{(m+1)}, J_{m+1}^{(m+1)}) \\ &= \sum_{i=1}^{m-1} P_{J_i^{(m+1)}} f\left(\frac{1}{P_{J_i^{(m+1)}}} \sum_{j \in J_i^{(m+1)}} p_j x_j\right) + P_{J_m^{(m+1)}} f\left(\frac{1}{P_{J_m^{(m+1)}}} \sum_{j \in J_m^{(m+1)}} p_j x_j\right) \\ &\quad + P_{J_{m+1}^{(m+1)}} f\left(\frac{1}{P_{J_{m+1}^{(m+1)}}} \sum_{j \in J_{m+1}^{(m+1)}} p_j x_j\right) + \bar{P}_{J^{(m+1)}} f\left(\frac{1}{\bar{P}_{J^{(m+1)}}} \sum_{j \in \bar{J}^{(m+1)}} p_j x_j\right)\end{aligned}$$

$$
\begin{aligned}
&\geq \sum_{i=1}^{m-1} P_{J_i^{(m+1)}} f\left(\frac{1}{P_{J_i^{(m+1)}}} \sum_{j\in J_i^{(m+1)}} p_j x_j\right) + P_{J_m^{(m+1)}\cup J_{m+1}^{(m+1)}} f\left(\frac{1}{P_{J_m^{(m+1)}\cup J_{m+1}^{(m+1)}}} \sum_{j\in J_m^{(m+1)}\cup J_{m+1}^{(m+1)}} p_j x_j\right)\\
&\quad + \bar{P}_{J^{(m+1)}} f\left(\frac{1}{\bar{P}_{J^{(m+1)}}} \sum_{j\in \bar{J}^{(m+1)}} p_j x_j\right)\\
&= \sum_{i=1}^{m-1} P_{J_i^{(m)}} f\left(\frac{1}{P_{J_i^{(m)}}} \sum_{j\in J_i^{(m)}} p_j x_j\right) + P_{J_m^{(m)}} f\left(\frac{1}{P_{J_m^{(m)}}} \sum_{j\in J_m^{(m)}} p_j x_j\right) + \bar{P}_{J^{(m)}} f\left(\frac{1}{\bar{P}_{J^{(m)}}} \sum_{j\in \bar{J}^{(m)}} p_j x_j\right)\\
&= D(f,\mathbf{p},\mathbf{x}; J^{(m)}, J_1^{(m)}, J_2^{(m)},\ldots, J_{m-1}^{(m)}, J_m^{(m)})\\
&= \max_{\varnothing\neq J\subset\{1,\ldots,n\}} D(f,\mathbf{p},\mathbf{x}; J, J_1, J_2,\ldots, J_m).
\end{aligned}
$$

So the middle inequality in (5) holds.

The first inequality and the last inequality in (5) can be seen from Theorem 1.3. □

Theorem 2.2. *Let C be a convex subset in the real linear space X and assume that $f: C \to \mathbb{R}$ is a convex function on C. If $x_k \in C$ and $p_k > 0$, $k \in \{1, 2,\ldots, n\}$ with $\sum_{i=k}^{n} p_k = 1$, then for any nonempty subset J of $\{1, 2,\ldots, n\}$, we have*

$$
\begin{aligned}
\sum_{k=1}^{n} p_k f(x_k) &\geq \min_{\varnothing\neq J\subset\{1,\ldots,n\}} D(f,\mathbf{p},\mathbf{x}; J, J_1, J_2,\ldots, J_{m+1})\\
&\geq \min_{\varnothing\neq J\subset\{1,\ldots,n\}} D(f,\mathbf{p},\mathbf{x}; J, J_1, J_2,\ldots, J_m) \geq f\left(\sum_{k=1}^{n} p_k x_k\right).
\end{aligned} \tag{6}
$$

Proof. We assume the value of $\min_{\varnothing\neq J\subset\{1,\ldots,n\}} D(f,\mathbf{p},\mathbf{x}; J, J_1, J_2,\ldots, J_{m+1})$ is obtained for $J_i = J_i^{(m+1)}$, $1 \leq i \leq m+1$. Then we let two nonempty subsets $J_m^{(m+1)}$, $J_{m+1}^{(m+1)}$ such that $J_m^{(m+1)} \cup J_{m+1}^{(m+1)} = J_m^{(m)}$. Using the similar method in Theorem 2.1, inequalities (6) can be obtained. □

Now we say that $S_1, S_2,\ldots, S_m$ generate a partition of the set $S \neq \varnothing$ if they are pairwise disjoint and non-empty sets with $\cup_{i=1}^{m} S_i = S$. Then the main results above are given as follows:

Theorem 2.3. *Let C be a convex subset in the real linear space X and assume that $f: C \to \mathbb{R}$ is a convex function on C. Assume further that $x_k \in C$ and $p_k > 0$, $k \in \{1, 2,\ldots, n\}$ with $\sum_{i=1}^{n} p_k = 1$. If $\mathcal{A}_m$ denotes all partitions of the set $\{1, 2,\ldots, n\}$ with m elements ($m = 1, 2,\ldots, n$), then*

$$
\begin{aligned}
\sum_{k=1}^{n} p_k f(x_k) &\geq \max_{\{J_1,J_2,\ldots,J_{n-1}\}\in\mathcal{A}_{n-1}} D(f,\mathbf{p},\mathbf{x}; J_1, J_2,\ldots, J_{n-1})\\
&\geq\cdots\geq \max_{\{J_1,J_2,\ldots,J_m\}\in\mathcal{A}_m} D(f,\mathbf{p},\mathbf{x}; J_1, J_2,\ldots, J_m)\\
&\geq\cdots\geq \max_{\{J_1,J_2\}\in\mathcal{A}_2} D(f,\mathbf{p},\mathbf{x}; J_1, J_2) \geq f\left(\sum_{k=1}^{n} p_k x_k\right),
\end{aligned} \tag{7}
$$

where

$$
D(f,\mathbf{p},\mathbf{x}; J_1, J_2,\ldots, J_m) := \sum_{i=1}^{m} P_{J_i} f\left(\frac{1}{P_{J_i}} \sum_{j\in J_i} p_j x_j\right), \quad m = 1, 2,\ldots, n.
$$

Proof. Since the first inequality and the last inequality follow from Theorem 1.3, we can suppose that $n \geq 4$, and we need only to prove that

$$\max_{\{J_1,J_2,\ldots,J_{m+1}\}\in\mathcal{A}_{m+1}} D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_{m+1}) \geq \max_{\{J_1,J_2,\ldots,J_m\}\in\mathcal{A}_m} D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_m).$$

for every $m = 2,\ldots,n-2$. It is enough to show that for each fixed $\{J_1,J_2,\ldots,J_m\}\in\mathcal{A}_m$ there exists $\{K_1,K_2,\ldots,K_{m+1}\}\in\mathcal{A}_{m+1}$ such that

$$D(f,\mathbf{p},\mathbf{x};K_1,K_2,\ldots,K_{m+1}) \geq D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_m).$$

Since $n \geq 4$ and $m \in \{2,\ldots,n-2\}$, one of the sets $J_1,J_2,\ldots,J_m$ contains at least two elements. We can suppose that

$$J_m = K_m \cup K_{m+1},$$

where K_m and K_{m+1} are disjoint and nonempty sets. Then $\{J_1,J_2,\ldots,J_{m-1},K_m,K_{m+1}\}\in\mathcal{A}_{m+1}$ and

$$P_{J_m} f\left(\frac{1}{P_{J_m}}\sum_{j\in J_m} p_j x_j\right) = P_{J_m} f\left(\frac{P_{K_m}}{P_{J_m}}\left(\frac{1}{P_{K_m}}\sum_{j\in K_m} p_j x_j\right) + \frac{P_{K_{m+1}}}{P_{J_m}}\left(\frac{1}{P_{K_{m+1}}}\sum_{j\in K_{m+1}} p_j x_j\right)\right).$$

By $\frac{P_{K_m}}{P_{J_m}} + \frac{P_{K_{m+1}}}{P_{J_m}} = 1$, Jensen's inequality can be applied, and we obtained from the aforementioned equality that

$$P_{J_m} f\left(\frac{1}{P_{J_m}}\sum_{j\in J_m} p_j x_j\right) \leq P_{K_m} f\left(\frac{1}{P_{K_m}}\sum_{j\in K_m} p_j x_j\right) + P_{K_{m+1}} f\left(\frac{1}{P_{K_{m+1}}}\sum_{j\in K_{m+1}} p_j x_j\right),$$

and this gives the result.

The proof is complete. □

Theorem 2.4. *Let C be a convex subset in the real linear space X and assume that $f: C \to \mathbb{R}$ is a convex function on C. Assume further that $x_k \in C$ and $p_k > 0$, $k \in \{1,2,\ldots,n\}$ with $\sum_{i=1}^n p_k = 1$. If $\mathcal{A}_m$ denotes all partitions of the set $\{1,2,\ldots,n\}$ with m elements $(m = 1,2,\ldots,n)$, then*

$$\begin{aligned}\sum_{k=1}^n p_k f(x_k) &\geq \min_{\{J_1,J_2,\ldots,J_{n-1}\}\in\mathcal{A}_{n-1}} D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_{n-1})\\ &\geq\cdots\geq \min_{\{J_1,J_2,\ldots,J_m\}\in\mathcal{A}_m} D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_m)\\ &\geq\cdots\geq \min_{\{J_1,J_2\}\in\mathcal{A}_2} D(f,\mathbf{p},\mathbf{x};J_1,J_2) \geq f\left(\sum_{k=1}^n p_k x_k\right),\end{aligned} \tag{8}$$

where

$$D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_m) := \sum_{i=1}^m P_{J_i} f\left(\frac{1}{P_{J_i}}\sum_{j\in J_i} p_j x_j\right), \quad m = 1,2,\ldots,n.$$

Proof. Analyzing the proof of Theorem 2.3, we can see the next: if $\{K_1,K_2,\ldots,K_{m+1}\}\in\mathcal{A}_{m+1}$ is a refinement of $\{J_1,J_2,\ldots,J_m\}\in\mathcal{A}_m$ (every element of $\{K_1,K_2,\ldots,K_{m+1}\}$ is contained in an element of $\{J_1,J_2,\ldots,J_m\}$), then

$$D(f,\mathbf{p},\mathbf{x};K_1,K_2,\ldots,K_{m+1}) \geq D(f,\mathbf{p},\mathbf{x};J_1,J_2,\ldots,J_m)$$

holds. Since each partition from $\mathcal{A}_{m+1}$ is a refinement of a partition from $\mathcal{A}_m$, the result follows.

The proof is complete. □

3 Applications in information theory

3.1 New upper bounds for Shannon's entropy

As the consistent work, bounds for Shannon's entropy [18] can be found in [4,8,10,15]. For further discussion, we present the definition of Shannon's entropy first. If the discrete probability distribution $\mathbb{P}^n$ is given by $P(X = i) = p_i, p_i > 0,\ i = 1, 2, \dots, n,\ s.\ t.\ \sum_{i=1}^{n} p_i = 1$, then Shannon's entropy is defined as

$$H(X) := \sum_{i=1}^{n} p_i \log \frac{1}{p_i}.$$

In [4], Popescu et al. obtained a new upper bound for entropy as follows:

$$H(X) \le \min_{J, J_1, J_2, \dots, J_m} \log \left[\prod_{i=1}^{m} \left(\frac{|J_i|}{P_{J_i}} \right)^{P_{J_i}} \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right]. \tag{9}$$

Furthermore, considering the aforementioned results the following tighter bounds for Shannon's entropy are presented.

Theorem 3.1. *Let $H(X)$ be defined as above, under the assumptions of* Theorem 2.1, *the following inequalities hold:*

$$\begin{aligned} H(X) &\le \cdots \le \min_{J, J_1, J_2, \dots, J_{m+1}} \log \left[\prod_{i=1}^{m+1} \left(\frac{|J_i|}{P_{J_i}} \right)^{P_{J_i}} \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right] \\ &\le \min_{J, J_1, J_2, \dots, J_m} \log \left[\prod_{i=1}^{m} \left(\frac{|J_i|}{P_{J_i}} \right)^{P_{J_i}} \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right] \\ &\le \cdots \le \min_{J, J_1} \log \left[\left(\frac{|J_1|}{P_{J_1}} \right)^{P_{J_1}} \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right] \le \log n. \end{aligned} \tag{10}$$

Proof. Taking into consideration the inequalities of Theorem 2.1 applied for the convex function $f(x) = -\log x$ and $x_i = 1/p_i$, $1 \le i \le n$, then

$$\begin{aligned} -\sum_{k=1}^{n} p_k \log \frac{1}{p_k} &\ge \max_{J, J_1, J_2, \dots, J_{m+1}} \left[-\sum_{i=1}^{m+1} P_{J_i} \log \left(\frac{1}{P_{J_i}} \sum_{j \in J_i} p_j \cdot \frac{1}{p_j} \right) - \bar{P}_J \log \left(\frac{1}{\bar{J}} \sum_{j \in \bar{J}} p_j \cdot \frac{1}{p_j} \right) \right] \\ &\ge \max_{J, J_1, J_2, \dots, J_m} \left[-\sum_{i=1}^{m} P_{J_i} \log \left(\frac{1}{P_{J_i}} \sum_{j \in J_i} p_j \cdot \frac{1}{p_j} \right) - \bar{P}_J \log \left(\frac{1}{\bar{P}_J} \sum_{j \in \bar{J}} p_j \cdot \frac{1}{p_j} \right) \right] \ge -\log \left(\sum_{k=1}^{n} p_k \cdot \frac{1}{p_k} \right). \end{aligned}$$

Those inequalities are equivalent with

$$\begin{aligned} H(X) &\le \min_{J, J_1, J_2, \dots, J_{m+1}} \left[\sum_{i=1}^{m+1} \log \left(\frac{|J_i|}{P_{J_i}} \right)^{P_{J_i}} + \log \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right] \\ &\le \min_{J, J_1, J_2, \dots, J_m} \left[\sum_{i=1}^{m} \log \left(\frac{|J_i|}{P_{J_i}} \right)^{P_{J_i}} + \log \left(\frac{|\bar{J}|}{\bar{P}_J} \right)^{\bar{P}_J} \right] \\ &\le \log n. \end{aligned}$$

Let m have the value from 1 to $n - 1$ and the inequalities (10) are deduced. □

Theorem 3.2. *Let $H(X)$ be defined as above, under the assumptions of* Theorem 2.3*, the following inequalities hold:*

$$\begin{aligned} H(X) &\le \min_{\{J_1,J_2,\dots,J_{n-1}\}\in\mathcal{A}_{n-1}} \log\left[\prod_{i=1}^{n-1}\left(\frac{|J_i|}{P_{J_i}}\right)^{P_{J_i}}\right] \le\cdots\le \min_{\{J_1,J_2,\dots,J_m\}\in\mathcal{A}_m} \log\left[\prod_{i=1}^{m}\left(\frac{|J_i|}{P_{J_i}}\right)^{P_{J_i}}\right] \\ &\le\cdots\le \min_{\{J_1,J_2\}\in\mathcal{A}_2} \log\left[\left(\frac{|J_1|}{P_{J_1}}\right)^{P_{J_1}}\left(\frac{|J_2|}{P_{J_2}}\right)^{P_{J_2}}\right] \le \log n. \end{aligned} \tag{11}$$

Proof. Taking into consideration the inequalities of Theorem 2.3, we have the inequalities (11) by the similar method above. □

3.2 New lower bounds for *f*-divergence measures

Given a convex function $f : [0,\infty) \to \mathbb{R}$, the f-divergence functional

$$I_f(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right), \tag{12}$$

where $\mathbf{p} = (p_1,\dots,p_n)$, $\mathbf{q} = (q_1,\dots,q_n)$ are positive sequences, was introduced by Csiszár in [19], as a generalized measure of information, a "distance function" on the set of probability distributions $\mathbb{P}^n$. As in [19], we interpret undefined expressions by

$$f(0) = \lim_{t\to 0+} f(t); \quad 0f\left(\frac{0}{0}\right) = 0; \quad 0f\left(\frac{a}{0}\right) = \lim_{q\to 0+} qf\left(\frac{a}{q}\right) = a\lim_{t\to\infty}\frac{f(t)}{t}, \quad a > 0.$$

The following results were essentially given by Csiszár and Körner [20]:

(i) If f is convex, then $I_f(\mathbf{p},\mathbf{q})$ is jointly convex in $\mathbf{p}$ and $\mathbf{q}$;

(ii) For every $\mathbf{p},\mathbf{q} \in \mathbb{R}_+^n$, we have

$$I_f(\mathbf{p},\mathbf{q}) \ge \sum_{i=1}^{n} q_i f\left(\frac{\sum_{i=1}^n p_i}{\sum_{i=1}^n q_i}\right). \tag{13}$$

If f is strictly convex, equality holds in (13) iff

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \cdots = \frac{p_n}{q_n}.$$

If f is normalized, i.e., $f(1) = 0$, then for every $\mathbf{p},\mathbf{q} \in \mathbb{R}_+^n$ with $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i$, we have the inequality

$$I_f(\mathbf{p},\mathbf{q}) \ge 0. \tag{14}$$

In particular, if $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$, then (14) holds. This is the well-known nonnegative property of the f-divergence.

Dragomir gives the concept for functions defined on a cone in a linear space as follows [3]:

In the first place, we recall that the subset K in a linear space X is a cone if the following two conditions are satisfied:

(i) for any $x, y \in K$ we have $x + y \in K$;

(ii) for any $x \in K$ and any $\alpha \ge 0$ we have $\alpha x \in K$.

For a given n-tuple of vectors $\mathbf{z} = (z_1,\dots,z_n) \in K^n$ and a probability distribution $\mathbf{q} \in \mathbb{P}^n$ with all values nonzero, we can define, for the convex function $f : K \to \mathbb{R}$, the following f-divergence of $\mathbf{z}$ with the distribution $\mathbf{q}$

$$I_f(\mathbf{z},\mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{z_i}{q_i}\right). \tag{15}$$

It is obvious that if $X \in \mathbb{R}$, $K = [0, \infty)$ and $x = p \in \mathbb{P}^n$, then we obtain the usual concept of the f-divergence associated with a function $f : [0, \infty) \to \mathbb{R}$. Now, for a given n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in K^n$, a probability distribution $\mathbf{q} \in \mathbb{P}^n$ with all values nonzero and for any nonempty pairwise disjoint subsets $J_1, J_2, \ldots, J_m, \bar{J}$ of $\{1, \ldots, n\}$ we have

$$\mathbf{q}_{J^{(m)}} := (Q_{J_1}, Q_{J_2}, \ldots, Q_{J_m}, \bar{Q}_J) \in \mathbb{P}^{m+1}$$

and

$$\mathbf{x}_{J^{(m)}} := (X_{J_1}, X_{J_2}, \ldots, X_{J_m}, \bar{X}_J) \in \mathbb{P}^{m+1},$$

where $Q_I = \sum_{i \in I} q_i$, $\bar{Q}_J = Q_{\bar{J}}$, and $X_I = \sum_{i \in I} x_i$, $\bar{X}_J = X_{\bar{J}}$.

Let

$$I_f(\mathbf{x}_{J^{(m)}}, \mathbf{q}_{J^{(m)}}) := \sum_{i=1}^{m} Q_{J_i} f\left(\frac{X_{J_i}}{Q_{J_i}}\right) + \bar{Q}_J f\left(\frac{\bar{X}_J}{\bar{Q}_J}\right). \tag{16}$$

The following inequalities for the f-divergence of an n-tuple of vectors in a linear space holds, which are better than the inequalities given in [3].

Theorem 3.3. *Let $f : K \to \mathbb{R}$ be a convex function on the cone K. Then for any n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in K^n$, a probability distribution $\mathbf{q} \in \mathbb{P}^n$ with all values nonzero and for any nonempty pairwise disjoint subsets $J_1, J_2, \ldots, J_m, \bar{J}$ of $\{1, \ldots, n\}$ we have*

$$\begin{aligned} I_f(\mathbf{x}, \mathbf{q}) &\geq \cdots \geq \max_{J, J_1, J_2, \ldots, J_{m+1}} (I_f(\mathbf{x}_{J^{(m+1)}}, \mathbf{q}_{J^{(m+1)}})) \\ &\geq \max_{J, J_1, J_2, \ldots, J_m} (I_f(x_{J^{(m)}}, q_{J^{(m)}})) \\ &\geq \cdots \geq \max_{J, J_1} (I_f(\mathbf{x}_{J^{(1)}}, \mathbf{q}_{J^{(1)}})) \geq f(X_n), \end{aligned} \tag{17}$$

where $X_n := \sum_{i=1}^{n} x_i$.

Proof. The aforementioned inequalities are obtained directly from Theorem 2.1 by letting $p_i \to q_i$ and $x_i \to \frac{x_i}{q_i}$. □

Theorem 3.4. *Let $f : K \to \mathbb{R}$ be a convex function on the cone K. Then for any n-tuple of vectors $\mathbf{x} = (x_1, \ldots, x_n) \in K^n$, a probability distribution $q \in \mathbb{P}^n$ with all values nonzero and for any nonempty any nonempty pairwise disjoint subsets $J_1, J_2, \ldots, J_m, \bar{J}$ of $\{1, \ldots, n\}$ we have*

$$\begin{aligned} I_f(\mathbf{x}, \mathbf{q}) &\geq \cdots \geq \min_{J, J_1, J_2, \ldots, J_{m+1}} (I_f(x_{J^{(m+1)}}, q_{J^{(m+1)}})) \\ &\geq \min_{J, J_1, J_2, \ldots, J_m} (I_f(x_{J^{(m)}}, q_{J^{(m)}})) \\ &\geq \cdots \geq \min_{J, J_1} (I_f(x_{J^{(1)}}, q_{J^{(1)}})) \geq f(X_n), \end{aligned} \tag{18}$$

where $X_n := \sum_{i=1}^{n} x_i$.

Proof. The aforementioned inequalities are obtained directly from Theorem 2.2 by letting $p_i \to q_i$ and $x_i \to \frac{x_i}{q_i}$. □

In the scalar case and if $\mathbf{x} = \mathbf{p} \in \mathbb{P}^n$, a sufficient condition for the positivity of the f-divergence $I_f(\mathbf{p}, \mathbf{q})$ is that $f(1) \geq 0$. The case of functions of a real variable that is meaningful for applications is involved in the following:

Corollary 3.1. *Let* $I_f(\mathbf{x}, \mathbf{q})$ *be defined as above, under the assumptions of* Theorem 3.2, *the following inequalities hold:*

$$
\begin{aligned}
I_f(\mathbf{p}, \mathbf{q}) &\geq \cdots \geq \max_{J, J_1, J_2, \ldots, J_{m+1}} \left(\sum_{i=1}^{m+1} Q_{J_i} f\left(\frac{P_{J_i}}{Q_{J_i}}\right) + \bar{Q}_J f\left(\frac{\bar{P}_J}{\bar{Q}_J}\right) \right) \\
&\geq \max_{J, J_1, J_2, \ldots, J_m} \left(\sum_{i=1}^{m} Q_{J_i} f\left(\frac{P_{J_i}}{Q_{J_i}}\right) + \bar{Q}_J f\left(\frac{\bar{P}_J}{\bar{Q}_J}\right) \right) \\
&\geq \cdots \geq \max_{J, J_1} \left(Q_{J_1} f\left(\frac{P_{J_1}}{Q_{J_1}}\right) + \bar{Q}_J f\left(\frac{\bar{P}_J}{\bar{Q}_J}\right) \right) \geq f(1) = 0.
\end{aligned} \tag{19}
$$

In what follows, we provide some lower bounds for a number of f-divergences that are used in various fields of information theory, probability theory and statistics.

The total variation distance is defined by the convex function $f(t) = |t - 1|$, $t \in \mathbb{R}$ and given by

$$
V(\mathbf{p}, \mathbf{q}) := \sum_{i=1}^{n} q_i \left| \frac{p_i}{q_i} - 1 \right| = \sum_{i=1}^{n} |p_i - q_i|. \tag{20}
$$

Proposition 3.1. *For any* $\mathbf{p}, \mathbf{q} \in \mathbb{P}^n$*, we have the inequality*

$$
\begin{aligned}
V(\mathbf{p}, \mathbf{q}) &\geq \cdots \geq \max_{J, J_1, J_2, \ldots, J_{m+1}} \left(\sum_{i=1}^{m+1} |P_{J_i} - Q_{J_i}| + |\bar{P}_J - \bar{Q}_J| \right) \\
&\geq \max_{J, J_1, J_2, \ldots, J_m} \left(\sum_{i=1}^{m} |P_{J_i} - Q_{J_i}| + |\bar{P}_J - \bar{Q}_J| \right) \geq \cdots \geq 2 \max_{J, J_1} |P_{J_1} - Q_{J_1}| (\geq 0).
\end{aligned} \tag{21}
$$

Proof. The proof follows by the inequalities (19) for the convex function $f(t) = |t - 1|, \; t \in \mathbb{R}$. □

The K. Pearson χ^2-divergence [21] is obtained for the convex function $f(t) = (1 - t)^2, \; t \in \mathbb{R}$ and given by

$$
\chi^2(\mathbf{p}, \mathbf{q}) := \sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i} - 1 \right)^2 = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i}. \tag{22}
$$

Proposition 3.2. *For any* $\mathbf{p}, \mathbf{q} \in \mathbb{P}^n$*, we have the inequality*

$$
\begin{aligned}
\chi^2(\mathbf{p}, \mathbf{q}) &\geq \cdots \geq \max_{J, J_1, J_2, \ldots, J_{m+1}} \left(\sum_{i=1}^{m+1} \frac{(P_{J_i} - Q_{J_i})^2}{Q_{J_i}} + \frac{(\bar{P}_J - \bar{Q}_J)^2}{\bar{Q}_J} \right) \\
&\geq \max_{J, J_1, J_2, \ldots, J_m} \left(\sum_{i=1}^{m+1} \frac{(P_{J_i} - Q_{J_i})^2}{Q_{J_i}} + \frac{(\bar{P}_J - \bar{Q}_J)^2}{\bar{Q}_J} \right) \\
&\geq \cdots \geq \max_{J, J_1} \frac{(P_{J_1} - Q_{J_1})^2}{Q_{J_1}(1 - Q_{J_1})} \geq 4 \max_{J, J_1} \left(P_{J_1} - Q_{J_1} \right)^2 (\geq 0).
\end{aligned} \tag{23}
$$

Proof. Using the inequalities (19) for the convex function $f(t) = (1 - t)^2, \; t \in \mathbb{R}$, we get the inequalities

$$
\begin{aligned}
\chi^2(\mathbf{p}, \mathbf{q}) &\geq \cdots \geq \max_{J, J_1, J_2, \ldots, J_{m+1}} \left(\sum_{i=1}^{m+1} \frac{(P_{J_i} - Q_{J_i})^2}{Q_{J_i}} + \frac{(\bar{P}_J - \bar{Q}_J)^2}{\bar{Q}_J} \right) \\
&\geq \max_{J, J_1, J_2, \ldots, J_m} \left(\sum_{i=1}^{m+1} \frac{(P_{J_i} - Q_{J_i})^2}{Q_{J_i}} + \frac{(\bar{P}_J - \bar{Q}_J)^2}{\bar{Q}_J} \right) \\
&\geq \cdots \geq \max_{J, J_1} \left(\frac{(P_{J_1} - Q_{J_1})^2}{Q_{J_1}} + \frac{(\bar{P}_J - \bar{Q}_J)^2}{\bar{Q}_J} \right) = \max_{J, J_1} \frac{(P_{J_1} - Q_{J_1})^2}{Q_{J_1}(1 - Q_{J_1})}.
\end{aligned}
$$

Since

$$Q_{J_1}(1-Q_{J_1}) \le \frac{1}{4}[Q_{J_1} + (1-Q_{J_1})]^2 = \frac{1}{4},$$

then

$$\frac{(P_{J_1}-Q_{J_1})^2}{Q_{J_1}(1-Q_{J_1})} \ge 4\left(P_{J_1}-Q_{J_1}\right)^2,$$

which proves the last part of inequalities (23). □

The Kullback-Leibler divergence [22] can be obtained for the convex function $f(t) = t\ln t,\ t > 0$ and given by

$$KL(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i \frac{p_i}{q_i}\ln\left(\frac{p_i}{q_i}\right) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right). \tag{24}$$

Proposition 3.3. *For any* $\mathbf{p},\mathbf{q} \in \mathbb{P}^n$, *we have the inequality*

$$\begin{aligned} KL(\mathbf{p},\mathbf{q}) &\ge \cdots \ge \ln\left\{\max_{J,J_1,J_2,\ldots,J_{m+1}}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} \ge \ln\left\{\max_{J,J_1,J_2,\ldots,J_m}\left[\prod_{i=1}^{m}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} \\ &\ge \cdots \ge \ln\left\{\max_{J,J_1}\left[\left(\frac{P_{J_1}}{Q_{J_1}}\right)^{P_{J_1}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} \ge 0. \end{aligned} \tag{25}$$

Proof. Using the inequalities (19) for the convex function $f(t) = t\ln t,\ \ t > 0$, we get the inequalities

$$\begin{aligned} KL(\mathbf{p},\mathbf{q}) &\ge \cdots \ge \ln\left\{\max_{J,J_1,J_2,\ldots,J_{m+1}}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} \ge \ln\left\{\max_{J,J_1,J_1,\ldots,J_m}\left[\prod_{i=1}^{m}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} \\ &\ge \cdots \ge \ln\left\{\max_{J,J_1}\left[\left(\frac{P_{J_1}}{Q_{J_1}}\right)^{P_{J_1}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J}\right]\right\} = \ln\left\{\max_{J,J_1}\left[\left(\frac{P_{J_1}}{Q_{J_1}}\right)^{P_{J_1}}\cdot\left(\frac{1-P_{J_1}}{1-Q_{J_1}}\right)^{1-P_{J_1}}\right]\right\}. \end{aligned}$$

Utilizing the geometric-harmonic mean inequality

$$x^w y^{1-w} \ge \frac{1}{\frac{w}{x} + \frac{1-w}{y}}, \quad x, y > 0, \quad 0 \le w \le 1,$$

we have for $x = \frac{P_{J_1}}{Q_{J_1}}$, $y = \frac{1-P_{J_1}}{1-Q_{J_1}}$, and $w = P_{J_1}$ that

$$\left(\frac{P_{J_1}}{Q_{J_1}}\right)^{P_{J_1}}\cdot\left(\frac{1-P_{J_1}}{1-Q_{J_1}}\right)^{1-P_{J_1}} \ge 1,$$

which proves the last part of inequalities (25). □

The Jeffrey's divergence [23] that has great importance in information theory can be obtained for the convex function $f(t) = (t-1)\ln t,\ t > 0$ and given by

$$J(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i\left(\frac{p_i}{q_i} - 1\right)\ln\left(\frac{p_i}{q_i}\right) = \sum_{i=1}^{n}(p_i - q_i)\ln\left(\frac{p_i}{q_i}\right). \tag{26}$$

Proposition 3.4. *For any* $\mathbf{p}, \mathbf{q} \in \mathbb{P}^n$, *we have the inequality*

$$\begin{aligned} J(\mathbf{p},\mathbf{q}) &\geq \cdots \geq \ln\left\{\max_{J,J_1,J_2,\ldots,J_{m+1}}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}-Q_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J-\bar{Q}_J}\right]\right\} \\ &\geq \ln\left\{\max_{J,J_1,J_2,\ldots,J_m}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}-Q_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J-\bar{Q}_J}\right]\right\} \\ &\geq \cdots \geq \ln\left\{\max_{J,J_1}\left[\left(\frac{(1-P_{J_1})Q_{J_1}}{(1-Q_{J_1})P_{J_1}}\right)^{Q_{J_1}-P_{J_1}}\right]\right\} \geq \max_{J,J_1}\left[\frac{2(Q_{J_1}-P_{J_1})^2}{P_{J_1}+Q_{J_1}-2P_{J_1}Q_{J_1}}\right] \geq 0. \end{aligned} \tag{27}$$

Proof. Applying the inequalities (19) for the convex function $f(t) = (t-1)\ln t$, $t > 0$, we get the inequalities

$$\begin{aligned} J(\mathbf{p},\mathbf{q}) &\geq \cdots \geq \ln\left\{\max_{J,J_1,J_2,\ldots,J_{m+1}}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}-Q_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J-\bar{Q}_J}\right]\right\} \\ &\geq \ln\left\{\max_{J,J_1,J_2,\ldots,J_m}\left[\prod_{i=1}^{m+1}\left(\frac{P_{J_i}}{Q_{J_i}}\right)^{P_{J_i}-Q_{J_i}}\cdot\left(\frac{\bar{P}_J}{\bar{Q}_J}\right)^{\bar{P}_J-\bar{Q}_J}\right]\right\} \\ &\geq \cdots \geq \ln\left\{\max_{J,J_1}\left[\left(\frac{(1-P_{J_1})Q_{J_1}}{(1-Q_{J_1})P_{J_1}}\right)^{Q_{J_1}-P_{J_1}}\right]\right\}. \end{aligned}$$

Utilizing the elementary inequality for positive numbers

$$\frac{\ln b - \ln a}{b-a} \geq \frac{2}{a+b}, \quad a, b > 0,$$

we have

$$\frac{(Q_{J_1}-P_{J_1})^2}{Q_{J_1}(1-Q_{J_1})}\cdot\frac{\ln\left(\frac{1-P_{J_1}}{1-Q_{J_1}}\right) - \ln\left(\frac{P_{J_1}}{Q_{J_1}}\right)}{\frac{1-P_{J_1}}{1-Q_{J_1}} - \frac{P_{J_1}}{Q_{J_1}}} \geq \frac{(Q_{J_1}-P_{J_1})^2}{Q_{J_1}(1-Q_{J_1})}\cdot\frac{2}{\frac{1-P_{J_1}}{1-Q_{J_1}} + \frac{P_{J_1}}{Q_{J_1}}}.$$

This inequality derives

$$(Q_{J_1}-P_{J_1})\left[\ln\left(\frac{1-P_{J_1}}{1-Q_{J_1}}\right) - \ln\left(\frac{P_{J_1}}{Q_{J_1}}\right)\right] \geq \frac{2(Q_{J_1}-P_{J_1})^2}{P_{J_1}(1-Q_{J_1}) + Q_{J_1}(1-P_{J_1})} \geq 0.$$

Rewriting the aforementioned inequalities the last part of the inequalities (27) can be obtained. □

Moreover, all the aforementioned theorems, corollaries, and propositions can also be changed into comparable versions according to Theorems 2.3 and 2.4.

4 Conclusion

The classical Jensen's inequality plays a very important role in both theory and applications. In this paper, we have obtained some refinements of Jensen's inequality (5)–(8) in real linear space using the generalized Popescu et al. functional. Moreover, we have obtained the new and sharp bounds of Shannon's entropy and several *f*-divergence measures in information theory. In the future work, we will continue to explore other applications on the inequalities newly obtained in Section 2.

Acknowledgement: The authors would like to thank the editor and referees for their very helpful suggestions and comments on the manuscript sincerely. This manuscript was supported by National Social Science Fund of China (17BTJ007), "the Fundamental Research Funds for the Central Universities," Zhongnan University of Economics and Law (2722020JCT031), MOE (Ministry of Education in China) Youth Foundation Project of Humanities and Social Sciences (19YJCZH111), Natural Science Foundation of Hubei Province (2017CFB145), and Technology Innovation Special Soft Science Research Program of Hubei Province (2019ADC136).

References

[1] J. L. W. V. Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math. **30** (1906), no. 1, 175–193.
[2] S. S. Dragomir, *A refinement of Jensen's inequality with applications for f-divergence measures*, Taiwanese J. Math. **14** (2010), no. 1, 153–164.
[3] S. S. Dragomir, *A new refinement of Jensen's inequality in linear spaces with applications*, Math. Comput. Model. **52** (2010), 1497–1505.
[4] P. G. Popescu, E. I. Slusanschi, V. Iancu, and F. Pop, *A new upper bound for Shannon entropy. A novel approach in modeling of Big Data applications*, Concurr. Comp.-Pract. Ex. **28** (2016), no. 2, 351–359.
[5] L. Horvăth, *A method to refine the discrete Jensen's inequality for convex and mid-convex functions*, Math. Comput. Model. **54** (2011), 2451–2459.
[6] L. Horváth, Đ. Pečarić and J. Pečarić, *Estimations of f- and Rényi divergences by using a cyclic refinement of the Jensens inequality*, Bull. Malays. Math. Sci. Soc. **42** (2019), 933–946.
[7] S. Simic, *Best possible global bounds for Jensen's inequality*, Appl. Math. Comput. **215** (2009), no. 6, 2224–2228.
[8] S. Simic, *Jensen's inequality and new entropy bounds*, Appl. Math. Lett. **22** (2009), no. 8, 1262–1265.
[9] L. Horvăth and J. Pečarić, *A refinement of the discrete Jensen's inequality*, Math. Inequal. Appl. **14** (2011), no. 4, 777–791.
[10] N. Ţăpuş and P. G. Popescu, *A new entropy upper bound*, Appl. Math. Lett. **25** (2012), no. 11, 1887–1890.
[11] L. Horvăth, *Weighted form of a recent refinement of the discrete Jensens inequality*, Math. Inequal. Appl. **17** (2014), no. 3, 947–961.
[12] S. G. Walker, *On a lower bound for the Jensen inequality*, SIAM J. Math. Anal. **46** (2014), no. 5, 3151–3157.
[13] S. S. Dragomir, M. A. Khan and A. Abathun, *Refinement of the Jensen integral inequality*, Open Math. **14** (2016), no. 1, 221–228.
[14] M. Sababheh, *Improved Jensen's inequality*, Math. Inequal. Appl. **20** (2017), no. 2, 389–403.
[15] G. Lu, *New refinements of Jensens inequality and entropy upper bounds*, J. Math. Inequal. **12** (2018), no. 2, 403–421.
[16] M. Adil Khan, M. Hanif, Z. A. Khan, K. Ahmad, and Y. M. Chu, *Association of Jensen inequality for s-convex function*, J. Inequal. Appl. **2019** (2019), art. 162, DOI: https://doi.org/10.1186/s13660-019-2112-9.
[17] M. Adil Khan, Z. Husain and Y. M. Chu, *New estimates for Csiszár divergence and Zipf-Mandelbrot entropy via Jensen-Mercers inequality*, Complexity **2020** (2020), art. 8928691, DOI: https://doi.org/10.1155/2020/8928691.
[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edn., John Wiley and Sons, Inc., New York, 2006.
[19] I. Csiszár, *Information-type measures of differences of probability distributions and indirect observations*, Studia Sci. Math. Hung. **2** (1967), 299–318.
[20] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.
[21] K. Pearson, *On the criterion that a give system of deviations from the probable in the case of correlated system of variables in such that it can be reasonable supposed to have arisen from random sampling*, Phil. Mag. **50** (1900), no. 302, 157–172.
[22] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Statist. **22** (1951), no. 1, 79–86.
[23] H. Jeffreys, *An invariant form for the prior probability in estimation problems*, Proc. Roy. Soc. Lon. Ser. A **186** (1946), no. 1007, 453–461.