Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 2, Issue 1, Article 5, 2001

FURTHER REVERSE RESULTS FOR JENSEN'S DISCRETE INEQUALITY AND APPLICATIONS IN INFORMATION THEORY

I. BUDIMIR, S.S. DRAGOMIR, AND J. PEČARIĆ

DEPARTMENT OF MATHEMATICS, FACULTY OF TEXTILE TECHNOLOGY, UNIVERSITY OF ZAGREB, CROATIA.

SCHOOL OF COMMUNICATIONS AND INFOMATICS, VICTORIA UNIVERSITY OF TECHNOLOGY, P.O. BOX 14428, MELBOURNE CITY MC, VICTORIA 8001, AUSTRALIA.
sever.dragomir@vu.edu.au
URL: http://rgmia.vu.edu.au/SSDragomirWeb.html

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF ADELAIDE, ADELAIDE, 5001.
URL: http://mahazu.hazu.hr/DepMPCS/indexJP.html

Received 12 April, 2000; accepted 6 October, 2000
Communicated by L.-E. Persson

ABSTRACT. Some new inequalities which counterpart Jensen's discrete inequality and improve the recent results from [4] and [5] are given. A related result for generalized means is established. Applications in Information Theory are also provided.

Key words and phrases: Convex functions, Jensen's Inequality, Entropy Mappings.

2000 *Mathematics Subject Classification.* 26D15, 94Xxx.

1. INTRODUCTION

Let $f : X \to \mathbb{R}$ be a convex mapping defined on the linear space X and $x_i \in X$, $p_i \geq 0$ $(i = 1, ..., m)$ with $P_m := \sum_{i=1}^{m} p_i > 0$.

The following inequality is well known in the literature as Jensen's inequality

$$f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \leq \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i). \tag{1.1}$$

There are many well known inequalities which are particular cases of Jensen's inequality, such as the weighted arithmetic mean-geometric mean-harmonic mean inequality, the Ky-Fan inequality, the Hölder inequality, etc. For a comprehensive list of recent results on Jensen's inequality, see the book [25] and the papers [9]-[15] where further references are given.

In 1994, Dragomir and Ionescu [18] proved the following inequality which counterparts (1.1) for real mappings of a real variable.

ISSN (electronic): 1443-5756

007-00

Theorem 1.1. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be a differentiable convex mapping on $\overset{\circ}{I}$ ($\overset{\circ}{I}$ is the interior of I), $x_i \in \overset{\circ}{I}$, $p_i \geq 0$ $(i = 1, ..., n)$ and $\sum_{i=1}^{n} p_i = 1$. Then we have the inequality*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &\leq \sum_{i=1}^{n} p_i x_i f'(x_i) - \sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i f'(x_i), \end{aligned} \tag{1.2}$$

where f' is the derivative of f on $\overset{\circ}{I}$.

Using this result and the discrete version of the Grüss inequality for weighted sums, S.S. Dragomir obtained the following simple counterpart of Jensen's inequality [5]:

Theorem 1.2. *With the above assumptions for f and if $m, M \in \overset{\circ}{I}$ and $m \leq x_i \leq M$ $(i = 1, ..., n)$, then we have*

$$0 \leq \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \frac{1}{4}(M - m)(f'(M) - f'(m)). \tag{1.3}$$

This was subsequently applied in Information Theory for Shannon's and Rényi's entropy.

In this paper we point out some other counterparts of Jensen's inequality that are similar to (1.3), some of which are better than the above inequalities.

2. Some New Counterparts for Jensen's Discrete Inequality

The following result holds.

Theorem 2.1. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be a differentiable convex mapping on $\overset{\circ}{I}$ and $x_i \in \overset{\circ}{I}$ with $x_1 \leq x_2 \leq \cdots \leq x_n$ and $p_i \geq 0$ $(i = 1, ..., n)$ with $\sum_{i=1}^{n} p_i = 1$. Then we have*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &\leq (x_n - x_1)(f'(x_n) - f'(x_1)) \max_{1 \leq k \leq n-1} \left\{P_k \bar{P}_{k+1}\right\} \\ &\leq \frac{1}{4}(x_n - x_1)(f'(x_n) - f'(x_1)), \end{aligned} \tag{2.1}$$

where $P_k := \sum_{i=1}^{k} p_i$ and $\bar{P}_{k+1} := 1 - P_k$.

Proof. We use the following Grüss type inequality due to J. E. Pečarić (see for example [25]):

$$\left|\frac{1}{Q_n}\sum_{i=1}^{n} q_i a_i b_i - \frac{1}{Q_n}\sum_{i=1}^{n} q_i a_i \cdot \frac{1}{Q_n}\sum_{i=1}^{n} q_i b_i\right| \leq |a_n - a_1| |b_n - b_1| \max_{1 \leq k \leq n-1}\left[\frac{Q_k \bar{Q}_{k+1}}{Q_n^2}\right], \tag{2.2}$$

provided that a, b are two monotonic $n-$tuples, q is a positive one, $Q_n := \sum_{i=1}^{n} q_i > 0$, $Q_k := \sum_{i=1}^{k} q_i$ and $\bar{Q}_{k+1} = Q_n - Q_{k+1}$.

If in (2.2) we choose $q_i = p_i$, $a_i = x_i$, $b_i = f'(x_i)$ (and a_i, b_i will be monotonic nondecreasing), then we may state that

$$\begin{aligned}(2.3)\quad \sum_{i=1}^{n} p_i x_i f'(x_i) - \sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i f'(x_i) \\ \le (x_n - x_1)\left(f'(x_n) - f'(x_1)\right) \max_{1\le k\le n-1}\left\{P_k \bar{P}_{k+1}\right\}.\end{aligned}$$

Now, using (1.2) and (2.3) we obtain the first inequality in (2.1).

For the second inequality, we observe that

$$P_k \bar{P}_{k+1} = P_k (1 - P_k) \le \frac{1}{4}(P_k + 1 - P_k)^2 = \frac{1}{4}$$

for all $k \in \{1, ..., n-1\}$ and then

$$\max_{1\le k\le n-1}\left\{P_k \bar{P}_{k+1}\right\} \le \frac{1}{4},$$

which proves the last part of (2.1). □

Remark 2.2. It is obvious that the inequality (2.1) is an improvement of (1.3) if we assume that the order for x_i is as in the statement of Theorem 2.1.

Another result is embodied in the following theorem.

Theorem 2.3. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be a differentiable convex mapping on $\mathring{I}$ and $m, M \in \mathring{I}$ with $m \le x_i \le M$ $(i = 1, ..., n)$ and $p_i \ge 0$ $(i = 1, ..., n)$ with $\sum_{i=1}^{n} p_i = 1$. If S is a subset of the set $\{1, ..., n\}$ minimizing the expression*

$$(2.4)\qquad \left|\sum_{i\in S} p_i - \frac{1}{2}\right|,$$

then we have the inequality

$$\begin{aligned}(2.5)\qquad 0 &\le \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &\le Q(M-m)\left(f'(M) - f'(m)\right) \le \frac{1}{4}(M-m)\left(f'(M) - f'(m)\right),\end{aligned}$$

where

$$Q = \sum_{i\in S} p_i \left(1 - \sum_{i\in S} p_i\right).$$

Proof. We use the following Grüss type inequality due the Andrica and Badea [2]:

$$(2.6)\quad \left|Q_n \sum_{i=1}^{n} q_i a_i b_i - \sum_{i=1}^{n} q_i a_i \cdot \sum_{i=1}^{n} q_i b_i\right| \le (M_1 - m_1)(M_2 - m_2) \sum_{i\in S} q_i \left(Q_n - \sum_{i\in S} q_i\right)$$

provided that $m_1 \le a_i \le M_1$, $m_2 \le b_i \le M_2$ for $i = 1, ..., n$, and S is the subset of $\{1, ..., n\}$ which minimises the expression

$$\left|\sum_{i\in S} q_i - \frac{1}{2} Q_n\right|.$$

Choosing $q_i = p_i$, $a_i = x_i$, $b_i = f'(x_i)$, then we may state that

$$\begin{aligned} 0 &\le \sum_{i=1}^{n} p_i x_i f'(x_i) - \sum_{i=1}^{n} p_i x_i \sum_{i=1}^{n} p_i f'(x_i) \\ &\le (M-m)\left(f'(M) - f'(m)\right) \sum_{i \in S} p_i \left(1 - \sum_{i \in S} p_i\right). \end{aligned} \tag{2.7}$$

Now, using (1.2) and (2.7), we obtain the first inequality in (2.5). For the last part, we observe that

$$Q \le \frac{1}{4}\left(\sum_{i \in S} p_i + 1 - \sum_{i \in S} p_i\right)^2 = \frac{1}{4}$$

and the theorem is thus proved. □

The following inequality is well known in the literature as the arithmetic mean-geometric mean-harmonic-mean inequality:

$$A_n(p,x) \ge G_n(p,x) \ge H_n(p,x), \tag{2.8}$$

where

$$\begin{aligned} A_n(p,x) &:= \sum_{i=1}^{n} p_i x_i \text{ - \textit{the arithmetic mean,}} \\ G_n(p,x) &:= \prod_{i=1}^{n} x_i^{p_i} \text{ - \textit{the geometric mean,}} \\ H_n(p,x) &:= \frac{1}{\sum_{i=1}^{n} \frac{p_i}{x_i}} \text{ - \textit{the harmonic mean,}} \end{aligned}$$

and $\sum_{i=1}^{n} p_i = 1$ $\left(p_i \ge 0,\, i = \overline{1,n}\right)$.

Using the above two theorems, we are able to point out the following reverse of the AGH - inequality.

Proposition 2.4. *Let* $x_i > 0$ $(i = 1, ..., n)$ *and* $p_i \ge 0$ *with* $\sum_{i=1}^{n} p_i = 1$.

(i) *If* $x_1 \le x_2 \le \cdots \le x_{n-1} \le x_n$, *then we have*

$$\begin{aligned} 1 &\le \frac{A_n(p,x)}{G_n(p,x)} \\ &\le \exp\left[\frac{(x_n - x_1)^2}{x_1 x_n} \max_{1 \le k \le n-1}\left\{P_k \bar{P}_{k+1}\right\}\right] \\ &\le \exp\left[\frac{1}{4} \cdot \frac{(x_n - x_1)^2}{x_1 x_n}\right]. \end{aligned} \tag{2.9}$$

(ii) *If the set* $S \subseteq \{1, ..., n\}$ *minimizes the expression (2.4), and* $0 < m \le x_i \le M < \infty$ $(i = 1, ..., n)$, *then*

$$1 \le \frac{A_n(p,x)}{G_n(p,x)} \le \exp\left[Q \cdot \frac{(M-m)^2}{mM}\right] \le \exp\left[\frac{1}{4} \cdot \frac{(M-m)^2}{mM}\right]. \tag{2.10}$$

The proof goes by the inequalities (2.1) and (2.5), choosing $f(x) = -\ln x$. A similar result can be stated for G_n and H_n.

Proposition 2.5. *Let $p \geq 1$ and $x_i > 0$, $p_i \geq 0$ $(i = 1, ..., n)$ with $\sum_{i=1}^{n} p_i = 1$.*

(i) *If $x_1 \leq x_2 \leq \cdots \leq x_{n-1} \leq x_n$, then we have*

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{n} p_i x_i^p - \left(\sum_{i=1}^{n} p_i x_i \right)^p \\
&\leq p\left(x_n - x_1\right)\left(x_n^{p-1} - x_1^{p-1}\right) \max_{1 \leq k \leq n-1} \left\{ P_k \bar{P}_{k+1} \right\} \\
&\leq \frac{p}{4}\left(x_n - x_1\right)\left(x_n^{p-1} - x_1^{p-1}\right).
\end{aligned} \tag{2.11}
$$

(ii) *If the set $S \subseteq \{1, ..., n\}$ minimizes the expression (2.4), and $0 < m \leq x_i \leq M < \infty$ $(i = 1, ..., n)$, then*

$$
\begin{aligned}
0 &\leq \sum_{i=1}^{n} p_i x_i^p - \left(\sum_{i=1}^{n} p_i x_i \right)^p \\
&\leq pQ\left(M - m\right)\left(M^{p-1} - m^{p-1}\right) \\
&\leq \frac{1}{4} p\left(M - m\right)\left(M^{p-1} - m^{p-1}\right).
\end{aligned} \tag{2.12}
$$

Remark 2.6. The above results are improvements of the corresponding inequalities obtained in [5].

Remark 2.7. Similar inequalities can be stated if we choose other convex functions such as: $f(x) = x \ln x$, $x > 0$ or $f(x) = \exp(x)$, $x \in \mathbb{R}$. We omit the details.

3. A Converse Inequality for Convex Mappings Defined on $\mathbb{R}^n$

In 1996, Dragomir and Goh [15] proved the following converse of Jensen's inequality for convex mappings on $\mathbb{R}^n$.

Theorem 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex mapping on $\mathbb{R}^n$ and*

$$
(\nabla f)(x) := \left(\frac{\partial f(x)}{\partial x^1}, ..., \frac{\partial f(x)}{\partial x^n} \right),
$$

the vector of the partial derivatives, $x = (x^1, ..., x^n) \in \mathbb{R}^n$.
If $x_i \in \mathbb{R}^m$ $(i = 1, ..., m)$, $p_i \geq 0$, $i = 1, ..., m$, with $P_m := \sum_{i=1}^{m} p_i > 0$, then

$$
\begin{aligned}
0 &\leq \frac{1}{P_m} \sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m} \sum_{i=1}^{m} p_i x_i \right) \\
&\leq \frac{1}{P_m} \sum_{i=1}^{m} p_i \left\langle \nabla f(x_i), x_i \right\rangle - \left\langle \frac{1}{P_m} \sum_{i=1}^{m} p_i \nabla f(x_i), \frac{1}{P_m} \sum_{i=1}^{m} p_i x_i \right\rangle.
\end{aligned} \tag{3.1}
$$

The result was applied to different problems in Information Theory by providing different counterpart inequalities for Shannon's entropy, conditional entropy, mutual information, conditional mutual information, etc.

For generalizations of (3.1) in Normed Spaces and other applications in Information Theory, see Matić's Ph.D dissertation [23].

Recently, Dragomir [4] provided an upper bound for Jensen's difference

$$
\Delta(f, p, x) := \frac{1}{P_m} \sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m} \sum_{i=1}^{m} p_i x_i \right), \tag{3.2}
$$

which, even though it is not as sharp as (3.1), provides a simpler way, and for applications, a better way, of estimating the Jensen's differences Δ. His result is embodied in the following theorem.

Theorem 3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex mapping and $x_i \in \mathbb{R}^n$, $i = 1, ..., m$. Suppose that there exists the vectors $\phi, \Phi \in \mathbb{R}^n$ such that*

$$\phi \le x_i \le \Phi \quad \text{(the order is considered on the co-ordinates)} \tag{3.3}$$

and $m, M \in \mathbb{R}^n$ are such that

$$m \le \nabla f(x_i) \le M \tag{3.4}$$

for all $i \in \{1, ..., m\}$. Then for all $p_i \ge 0$ $(i = 1, ..., m)$ with $P_m > 0$, we have the inequality

$$0 \le \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \le \frac{1}{4}\left\|\Phi - \phi\right\|\left\|M - m\right\|, \tag{3.5}$$

where $\|\cdot\|$ is the usual Euclidean norm on $\mathbb{R}^n$.

He applied this inequality to obtain different upper bounds for Shannon's and Rényi's entropies.

In this section, we point out another counterpart for Jensen's difference, assuming that the $\nabla-$operator is of Hölder's type, as follows.

Theorem 3.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex mapping and $x_i \in \mathbb{R}^n$, $p_i \ge 0$ $(i = 1, ..., m)$ with $P_m > 0$. Suppose that the $\nabla-$operator satisfies a condition of $r-H-$Hölder type, i.e.,*

$$\left\|\nabla f(x) - \nabla f(y)\right\| \le H\left\|x - y\right\|^r, \text{ for all } x, y \in \mathbb{R}^n, \tag{3.6}$$

where $H > 0$, $r \in (0, 1]$ and $\|\cdot\|$ is the Euclidean norm.
Then we have the inequality:

$$\begin{aligned} 0 &\le \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \\ &\le \frac{H}{P_m^2}\sum_{1\le i<j\le m} p_i p_j \left\|x_i - x_j\right\|^{r+1}. \end{aligned} \tag{3.7}$$

Proof. We recall Korkine's identity,

$$\frac{1}{P_m}\sum_{i=1}^{m} p_i \langle y_i, x_i\rangle - \left\langle \frac{1}{P_m}\sum_{i=1}^{m} p_i y_i, \frac{1}{P_m}\sum_{i=1}^{m} p_i x_i \right\rangle = \frac{1}{2P_m^2}\sum_{i,j=1}^{n} p_i p_j \langle y_i - y_j, x_i - x_j\rangle, \ x, y \in \mathbb{R}^n,$$

and simply write

$$\begin{aligned} &\frac{1}{P_m}\sum_{i=1}^{m} p_i \langle \nabla f(x_i), x_i\rangle - \left\langle \frac{1}{P_m}\sum_{i=1}^{m} p_i \nabla f(x_i), \frac{1}{P_m}\sum_{i=1}^{m} p_i x_i \right\rangle \\ &\qquad = \frac{1}{2P_m^2}\sum_{i,j=1}^{n} p_i p_j \langle \nabla f(x_i) - \nabla f(x_j), x_i - x_j\rangle. \end{aligned}$$

Using (3.1) and the properties of the modulus, we have

$$
\begin{aligned}
0 &\le \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \\
&\le \frac{1}{2P_m^2}\sum_{i,j=1}^{m} p_i p_j \left|\langle \nabla f(x_i) - \nabla f(x_j), x_i - x_j\rangle\right| \\
&\le \frac{1}{2P_m^2}\sum_{i,j=1}^{m} p_i p_j \left\|\nabla f(x_i) - \nabla f(x_j)\right\| \left\|x_i - x_j\right\| \\
&\le \frac{H}{P_m^2}\sum_{i,j=1}^{m} p_i p_j \left\|x_i - x_j\right\|^{r+1}
\end{aligned}
$$

and the inequality (3.7) is proved. □

Corollary 3.4. *With the assumptions of Theorem 3.3 and if* $\Delta = \max_{1\le i<j\le m}\|x_i - x_j\|$*, then we have the inequality*

$$
0 \le \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \le \frac{H\Delta^{r+1}}{2P_m^2}\left(1 - \sum_{i=1}^{m} p_i^2\right). \tag{3.8}
$$

Proof. Indeed, as

$$
\sum_{1\le i<j\le m} p_i p_j \left\|x_i - x_j\right\|^{r+1} \le \Delta^{r+1} \sum_{1\le i<j\le m} p_i p_j.
$$

However,

$$
\sum_{1\le i<j\le m} p_i p_j = \frac{1}{2}\left(\sum_{i,j=1}^{m} p_i p_j - \sum_{i=j} p_i p_j\right) = \frac{1}{2}\left(1 - \sum_{i=1}^{m} p_i^2\right),
$$

and the inequality (3.8) is proved. □

The case of Lipschitzian mappings is embodied in the following corollary.

Corollary 3.5. *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a differentiable convex mapping and* $x_i \in \mathbb{R}^n$, $p_i \ge 0$ $(i = 1, ..., n)$ *with* $P_m > 0$. *Suppose that the* $\nabla-$*operator is Lipschitzian with the constant* $L > 0$, *i.e.,*

$$
\left\|\nabla f(x) - \nabla f(y)\right\| \le L\left\|x - y\right\|, \quad \text{for all } x, y \in \mathbb{R}^n, \tag{3.9}
$$

where $\|\cdot\|$ *is the Euclidean norm. Then*

$$
\begin{aligned}
0 &\le \frac{1}{P_m}\sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right) \\
&\le L\left[\frac{1}{P_m}\sum_{i=1}^{m} p_i \left\|x_i\right\|^2 - \left\|\frac{1}{P_m}\sum_{i=1}^{m} p_i x_i\right\|^2\right].
\end{aligned} \tag{3.10}
$$

Proof. The argument is obvious by Theorem 3.3, taking into account that for $r = 1$,

$$
\sum_{1\le i<j\le m} p_i p_j \left\|x_i - x_j\right\|^2 = P_m\sum_{i=1}^{m} p_i \left\|x_i\right\|^2 - \left\|\sum_{i=1}^{m} p_i x_i\right\|^2,
$$

and $\|\cdot\|$ is the Euclidean norm. □

Moreover, if we assume more about the vectors $(x_i)_{i=\overline{1,n}}$, we can obtain a simpler result that is similar to the one in [4].

Corollary 3.6. *Assume that f is as in Corollary 3.5. If*

$$\phi \leq x_i \leq \Phi \quad \text{(on the co-ordinates)}, \ \phi, \Phi \in \mathbb{R}^n \ \ (i = 1, .., m), \tag{3.11}$$

then we have the inequality

$$\begin{aligned} 0 &\leq \frac{1}{P_m} \sum_{i=1}^{m} p_i f(x_i) - f\left(\frac{1}{P_m} \sum_{i=1}^{m} p_i x_i\right) \\ &\leq \frac{1}{4} \cdot L \cdot \|\Phi - \phi\|^2 . \end{aligned} \tag{3.12}$$

Proof. It follows by the fact that in $\mathbb{R}^n$, we have the following Grüss type inequality (as proved in [4])

$$\frac{1}{P_m} \sum_{i=1}^{m} p_i \|x_i\|^2 - \left\| \frac{1}{P_m} \sum_{i=1}^{m} p_i x_i \right\|^2 \leq \frac{1}{4} \|\Phi - \phi\|^2 , \tag{3.13}$$

provided that (3.11) holds. □

Remark 3.7. For some Grüss type inequalities in Inner Product Spaces, see [7].

4. SOME RELATED RESULTS

Start with the following definitions from [3].

Definition 4.1. Let $-\infty < a < b < \infty$. Then $CM[a, b]$ denotes the set of all functions with domain $[a, b]$ that are continuous and strictly monotonic there.

Definition 4.2. Let $-\infty < a < b < \infty$, and let $f \in CM[a, b]$. Then, for each positive integer n, each $n-$tuple $x = (x_1, ..., x_n)$, where $a \leq x_j \leq b$ $(j = 1, 2, ..., n)$, and each n-tuple $p = (p_1, p_2, ..., p_n)$, where $p_j > 0$ $(j = 1, 2, ..., n)$ and $\sum_{j=1}^{n} p_j = 1$, let $M_f(x, y)$ denote the (weighted) mean

$$f^{-1}\left\{\sum_{j=1}^{n} p_j f(x_j)\right\}.$$

We may state now the following result.

Theorem 4.1. *Let S be the subset of $\{1, ..., n\}$ which minimizes the expression $\left|\sum_{i \in S} p_i - 1/2\right|$. If $f, g \in CM[a, b]$, then*

$$\sup_x \{|M_f(x, p) - M_g(x, p)|\} \leq Q \cdot \left\|\left(f^{-1}\right)'\right\|_\infty \cdot \left\|\left(f \circ g^{-1}\right)''\right\|_\infty \cdot |g(b) - g(a)|^2 ,$$

provided that the right-hand side of the inequality is finite, where, as above,

$$Q = \left(\sum_{i \in S} p_i\right)\left(1 - \sum_{i \in S} p_i\right),$$

and $\|\cdot\|_\infty$ is the usual sup-norm.

Proof. Let, as in [3], $h = f \circ g^{-1}$, $n > 1$,

$$x = (x_1, x_2, ..., x_n) \text{ and } p = (p_1, p_2, ..., p_n)$$

be as in the Definition 4.2, and $y_j = g(x_j)$ $(j = 1, 2, ..., n)$. By the mean-value theorem, for some α in the open interval joining $f(a)$ to $f(b)$, we have

$$\begin{aligned} M_f(x,p) - M_g(x,p) &= f^{-1}\left\{\sum_{j=1}^{n} p_j f(x_j)\right\} - f^{-1}\left[h\left\{\sum_{j=1}^{n} p_j g(x_j)\right\}\right] \\ &= \left(f^{-1}\right)'(\alpha)\left[\sum_{j=1}^{n} p_j f(x_j) - h\left\{\sum_{j=1}^{n} p_j g(x_j)\right\}\right] \\ &= \left(f^{-1}\right)'(\alpha)\left[\sum_{j=1}^{n} p_j h(y_j) - h\left\{\sum_{j=1}^{n} p_j y_j\right\}\right] \\ &= \left(f^{-1}\right)'(\alpha)\left[\sum_{j=1}^{n} p_j \left\{h(y_j) - h\left(\sum_{k=1}^{n} p_k y_k\right)\right\}\right]. \end{aligned}$$

Using the mean-value theorem a second time, we conclude that there exists points $z_1, z_2, ..., z_n$ in the open interval joining $g(a)$ to $g(b)$, such that

$$\begin{aligned} M_f(x,p) - M_g(x,p) &= \left(f^{-1}\right)'(\alpha)\big[p_1\left\{(1-p_1)y_1 - p_2y_2 - \cdots - p_ny_n\right\}h'(z_1) \\ &\quad +p_2\left\{-p_1y_1 + (1-p_2)y_2 - \cdots - p_ny_n\right\}h'(z_2) \\ &\quad +\cdots \\ &\quad +p_n\left\{-p_1y_1 - p_2y_2 - \cdots + (1-p_n)y_n\right\}h'(z_n)\big] \\ &= \left(f^{-1}\right)'(\alpha)\big[p_1\left\{p_2(y_1-y_2) + \cdots + p_n(y_1-y_n)\right\}h'(z_1) \\ &\quad +p_2\left\{p_1(y_2-y_1) + \cdots + p_n(y_2-y_n)\right\}h'(z_2) \\ &\quad +\cdots \\ &\quad +p_n\left\{p_1(y_n-y_1) + \cdots + p_{n-1}(y_n-y_{n-1})\right\}h'(z_n)\big] \\ &= \left(f^{-1}\right)'(\alpha)\sum_{1\le i<j\le n} p_ip_j(y_i-y_j)\left\{h'(z_i)-h'(z_j)\right\}. \end{aligned}$$

Using the mean value theorem a third time, we conclude that there exists points ω_{ij} $(1 \le i < j \le n)$ in the open interval joining $g(a)$ to $g(b)$, such that

$$\begin{aligned} &\left(f^{-1}\right)'(\alpha)\sum_{1\le i<j\le n} p_ip_j(y_i-y_j)\left\{h'(z_i)-h'(z_j)\right\} \\ &\qquad\qquad = \left(f^{-1}\right)'(\alpha)\sum_{1\le i<j\le n} p_ip_j(y_i-y_j)(z_i-z_j)\,h''(\omega_{ij}). \end{aligned}$$

Consequently,

$$\begin{aligned} \left|M_f(x,p) - M_g(x,p)\right| &\le \left|\left(f^{-1}\right)'(\alpha)\right|\sum_{1\le i<j\le n} p_ip_j\left|y_i-y_j\right|\cdot\left|z_i-z_j\right|\cdot\left|h''(\omega_{ij})\right| \\ &\le \left\|\left(f^{-1}\right)'\right\|_\infty\cdot\left\|h''\right\|_\infty\cdot\sum_{1\le i<j\le n} p_ip_j\left|y_i-y_j\right|\cdot\left|z_i-z_j\right| \end{aligned}$$

$$\begin{aligned}
&\leq \text{(by the Cauchy-Buniakowski-Schwartz inequality)}\\
&\leq \left\|\left(f^{-1}\right)'\right\|_{\infty}\cdot\left\|\left(f\circ g^{-1}\right)''\right\|_{\infty}\cdot\sqrt{\sum_{1\leq i<j\leq n}p_ip_j\left|y_i-y_j\right|^2}\cdot\sqrt{\sum_{1\leq i<j\leq n}p_ip_j\left|z_i-z_j\right|^2}\\
&\leq \text{(by the Andrica and Badea result)}\\
&\leq \left\|\left(f^{-1}\right)'\right\|_{\infty}\cdot\left\|\left(f\circ g^{-1}\right)''\right\|_{\infty}\cdot\sqrt{\left(\sum_{i\in S}p_i\right)\left(1-\sum_{i\in S}p_i\right)\left|g(b)-g(a)\right|^2}\\
&\quad\cdot\sqrt{\left(\sum_{i\in S}p_i\right)\left(1-\sum_{i\in S}p_i\right)\left|g(b)-g(a)\right|^2}\\
&= Q\left\|\left(f^{-1}\right)'\right\|_{\infty}\cdot\left\|\left(f\circ g^{-1}\right)''\right\|_{\infty}\cdot\left|g(b)-g(a)\right|^2,
\end{aligned}$$

and the theorem is proved. □

Corollary 4.2. *If $f,g\in CM\,[a,b]$, then*

$$\sup_x\left\{\left|M_f\left(x,p\right)-M_g\left(x,p\right)\right|\right\}\leq Q\cdot\left\|\frac{1}{f'}\right\|_{\infty}\cdot\left\|\frac{1}{g'}\left(\frac{f'}{g'}\right)'\right\|_{\infty}\cdot\left|g(b)-g(a)\right|^2,$$

provided that the right hand side of the inequality exists.

Proof. This follows at once from the fact that

$$\left(f^{-1}\right)'=\frac{1}{f'\circ f^{-1}}$$

and

$$\left(f\circ g^{-1}\right)''=\frac{\left(g'\circ g^{-1}\right)\left(f''\circ g^{-1}\right)-\left(f'\circ g^{-1}\right)\left(g''\circ g^{-1}\right)}{\left(g'\circ g^{-1}\right)^3}=\left[\frac{1}{g'}\left(\frac{f'}{g'}\right)'\right]\circ g^{-1}.$$

□

Remark 4.3. This establishes Theorem 4.3 from [3] and replaces the multiplicative factor $\frac{1}{4}$ by Q. In Corollary 4.2, we also replaced the multiplicative factor $\frac{1}{4}$ by Q.

5. APPLICATIONS IN INFORMATION THEORY

We give some new applications for Shannon's entropy

$$H_b\left(X\right):=\sum_{i=1}^{r}p_i\log_b\frac{1}{p_i},$$

where X is a random variable with the probability distribution $\left(p_i\right)_{i=\overline{1,r}}$.

Theorem 5.1. *Let X be as above and assume that $p_1\geq p_2\geq\cdots\geq p_r$ or $p_1\leq p_2\leq\cdots\leq p_r$. Then we have the inequality*

$$0\leq\log_b r-H_b\left(X\right)\leq\frac{\left(p_1-p_r\right)^2}{p_1p_r}\max_{1\leq k\leq r}\left\{P_k\bar{P}_{k+1}\right\}. \tag{5.1}$$

Proof. We choose in Theorem 2.1, $f(x)=-\log_b x$, $x>0$, $x_i=\frac{1}{p_i}$ $(i=1,...,r)$. Then we have $x_1\leq x_2\leq\cdots\leq x_r$ and by (2.1) we obtain

$$0\leq\log_b r-H_b\left(X\right)\leq\left(\frac{1}{p_r}-\frac{1}{p_1}\right)\left(\frac{1}{-\frac{1}{p_r}}+\frac{1}{\frac{1}{p_1}}\right)\max_{1\leq k\leq r}\left\{P_k\bar{P}_{k+1}\right\},$$

which is equivalent to (5.1). The same inequality is obtained if $p_1 \leq p_2 \leq \cdots \leq p_r$. □

Theorem 5.2. *Let X be as above and suppose that*

$$\begin{aligned} p_M &: = \max\{p_i | i = 1, ..., r\}, \\ p_m &: = \min\{p_i | i = 1, ..., r\}. \end{aligned}$$

If S is a subset of the set $\{1, ..., r\}$ minimizing the expression $\left|\sum_{i\in S} p_i - 1/2\right|$, then we have the estimation

$$0 \leq \log_b r - H_b(X) \leq Q \cdot \frac{(p_M - p_m)^2}{\ln b \cdot p_M p_m}. \tag{5.2}$$

Proof. We shall choose in Theorem 2.3,

$$f(x) = -\log_b x,\ x > 0,\ x_i = \frac{1}{p_i} \quad \left(i = \overline{1, r}\right).$$

Then $m = \frac{1}{p_M}$, $M = \frac{1}{p_m}$, $f'(x) = -\frac{1}{x \ln b}$ and the inequality (2.3) becomes:

$$\begin{aligned} 0 &\leq \log_b r - \sum_{i=1}^{r} p_i \log_b \frac{1}{p_i} \\ &\leq Q \frac{1}{\ln b} \left(\frac{1}{p_m} - \frac{1}{p_M}\right) \left(-\frac{1}{\frac{1}{p_m}} + \frac{1}{\frac{1}{p_M}}\right) \\ &= Q \cdot \frac{1}{\ln b} \cdot \frac{(p_M - p_m)^2}{p_M p_m}, \end{aligned}$$

hence the estimation (5.2) is proved. □

Consider the Shannon entropy

$$H(X) := H_e(X) = \sum_{i=1}^{r} p_i \ln \frac{1}{p_i} \tag{5.3}$$

and Rényi's entropy of order α $(\alpha \in (0, \infty) \setminus \{1\})$

$$H_{[\alpha]}(X) := \frac{1}{1-\alpha} \ln \left(\sum_{i=1}^{r} p_i^{\alpha}\right). \tag{5.4}$$

Using the classical Jensen's discrete inequality for convex mappings, i.e.,

$$f\left(\sum_{i=1}^{r} p_i x_i\right) \leq \sum_{i=1}^{r} p_i f(x_i), \tag{5.5}$$

where $f : I \subseteq \mathbb{R} \to \mathbb{R}$ is a convex mapping on I, $x_i \in I$ $(i = 1, ..., r)$ and $(p_i)_{i=\overline{1,r}}$ is a probability distribution, for the convex mapping $f(x) = -\ln x$, we have

$$\ln \left(\sum_{i=1}^{r} p_i x_i\right) \geq \sum_{i=1}^{r} p_i \ln x_i. \tag{5.6}$$

Choose $x_i = p_i^{\alpha-1}$ $(i = 1, ..., r)$ in (5.6) to obtain

$$\ln \left(\sum_{i=1}^{r} p_i^{\alpha}\right) \geq (\alpha - 1) \sum_{i=1}^{r} p_i \ln p_i,$$

which is equivalent to

$$(1-\alpha)\left[H_{[\alpha]}(X) - H(X)\right] \geq 0.$$

Now, if $\alpha \in (0,1)$, then $H_{[\alpha]}(X) \leq H(X)$, and if $\alpha > 1$ then $H_{[\alpha]}(X) \geq H(X)$. Equality holds iff $(p_i)_{i=\overline{1,r}}$ is a uniform distribution and this fact follows by the strict convexity of $-\ln(\cdot)$. This inequality also follows as a special case of the following well known fact: $H_{[\alpha]}(X)$ *is a nondecreasing function of* α. See for example [26] or [22].

Theorem 5.3. *Under the above assumptions, given that* $p_m = \min_{i=\overline{1,r}} p_i$, $p_M = \max_{i=\overline{1,r}} p_i$, *then we have the inequality*

$$0 \leq (1-\alpha)\left[H_{[\alpha]}(X) - H(X)\right] \leq Q \cdot \frac{\left(p_M^{\alpha-1} - p_m^{\alpha-1}\right)^2}{p_M^{\alpha-1} p_m^{\alpha-1}}, \tag{5.7}$$

for all $\alpha \in (0,1) \cup (1,\infty)$.

Proof. If $\alpha \in (0,1)$, then

$$x_i := p_i^{\alpha-1} \in \left[p_M^{\alpha-1}, p_m^{\alpha-1}\right]$$

and if $\alpha \in (1,\infty)$, then

$$x_i = p_i^{\alpha-1} \in \left[p_m^{\alpha-1}, p_M^{\alpha-1}\right], \text{ for } i \in \{1,...,n\}.$$

Applying Theorem 2.3 for $x_i := p_i^{\alpha-1}$ and $f(x) = -\ln x$, and taking into account that $f'(x) = -\frac{1}{x}$, we obtain

$$\begin{aligned}
&(1-\alpha)\left[H_{[\alpha]}(X) - H(X)\right] \\
&\leq \begin{cases} Q\left(p_m^{\alpha-1} - p_M^{\alpha-1}\right)\left(-\frac{1}{p_m^{\alpha-1}} + \frac{1}{p_M^{\alpha-1}}\right) & \text{if} \quad \alpha \in (0,1), \\ Q\left(p_M^{\alpha-1} - p_m^{\alpha-1}\right)\left(-\frac{1}{p_M^{\alpha-1}} + \frac{1}{p_m^{\alpha-1}}\right) & \text{if} \quad \alpha \in (1,\infty) \end{cases} \\
&= \begin{cases} Q \cdot \frac{\left(p_m^{\alpha-1} - p_M^{\alpha-1}\right)^2}{p_m^{\alpha-1} p_M^{\alpha-1}} & \text{if} \quad \alpha \in (0,1), \\ Q \cdot \frac{\left(p_M^{\alpha-1} - p_m^{\alpha-1}\right)^2}{p_M^{\alpha-1} p_m^{\alpha-1}} & \text{if} \quad \alpha \in (1,\infty) \end{cases} \\
&= Q \cdot \frac{\left(p_M^{\alpha-1} - p_m^{\alpha-1}\right)^2}{p_M^{\alpha-1} p_m^{\alpha-1}}
\end{aligned}$$

for all $\alpha \in (0,1) \cup (1,\infty)$ and the theorem is proved. □

Using a similar argument to the one in Theorem 5.3, we can state the following direct application of Theorem 2.3.

Theorem 5.4. *Let* $(p_i)_{i=\overline{1,r}}$ *be as in Theorem 5.3. Then we have the inequality*

$$0 \leq (1-\alpha) H_{[\alpha]}(X) - \ln r - \alpha \ln G_r(p) \leq Q \cdot \frac{\left(p_M^{\alpha-1} - p_m^{\alpha-1}\right)^2}{P_M^{\alpha-1} p_m^{\alpha-1}}, \tag{5.8}$$

for all $\alpha \in (0,1) \cup (1,\infty)$.

Remark 5.5. The above results improve the corresponding results from [5] and [4] with the constant Q which is less than $\frac{1}{4}$.

Acknowledgement 1. The authors would like to thank the anonymous referee for valuable comments and for the references [26] and [22].

REFERENCES

[1] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkley Symp. Math. Statist. Prob.*, **1** (1961), 547–561, Univ. of California Press, Berkley.

[2] D. ANDRICA AND C. BADEA, Grüss' inequality for positive linear functionals, *Periodica Math. Hung.*, **19**(2) (1988), 155–167.

[3] G.T. CARGO AND O. SHISHA, A metric space connected with general means, *J. Approx. Th.*, **2** (1969), 207–222.

[4] S.S. DRAGOMIR, A converse of the Jensen inequality for convex mappings of several variables and applications. (Electronic Preprint: `http://matilda.vu.edu.au/~rgmia/InfTheory/ConverseJensen.dvi`)

[5] S.S. DRAGOMIR, A converse result for Jensen's discrete inequality via Grüss' inequality and applications in information theory, *Analele Univ. Oradea, Fasc. Math.*, **7** (1999-2000), 178–189.

[6] S.S. DRAGOMIR, A further improvement of Jensen's inequality, *Tamkang J. Math.*, **25**(1) (1994), 29–36.

[7] S.S. DRAGOMIR, A generalisation of Grüss's inequality in inner product spaces and applications, *J. Math. Anal. Appl.*, **237** (1999), 74–82.

[8] S.S. DRAGOMIR, A new improvement of Jensen's inequality, *Indian J. Pure Appl. Math.*, **26**(10) (1995), 959–968.

[9] S.S. DRAGOMIR, An improvement of Jensen's inequality, *Bull. Math. Soc. Sci. Math. Romania.*, **34** (1990), 291–296.

[10] S.S. DRAGOMIR, On some refinements of Jensen's inequality and applications, *Utilitas Math.*, **43** (1993), 235–243.

[11] S.S. DRAGOMIR, Some refinements of Jensen's inequality, *J. Math. Anal. Appl.*, **168** (1992), 518–522.

[12] S.S. DRAGOMIR, Some refinements of Ky Fan's inequality, *J. Math. Anal. Appl.*, **163** (1992), 317–321.

[13] S.S. DRAGOMIR, Two mappings in connection with Jensen's inequality, *Extracta Math.*, **8**(2-3) (1993), 102–105.

[14] S.S. DRAGOMIR AND S. FITZPARTICK, s - Orlicz convex functions in linear spaces and Jensen's discrete inequality, *J. Math. Anal. Appl.*, **210** (1997), 419–439.

[15] S.S. DRAGOMIR AND C.J. GOH, A counterpart of Jensen's discrete inequality for differentiable convex mappings and applications in Information Theory, *Math. Comput. Modelling*, **24**(2) (1996), 1–11.

[16] S.S. DRAGOMIR AND C.J. GOH, Some bounds on entropy measures in information theory, *Appl. Math. Lett.*, **10**(3) (1997), 23–28.

[17] S.S. DRAGOMIR AND C.J. GOH, Some counterpart inequalities for a functional associated with Jensen's inequality, *J. Inequal. Appl.*, **1** (1997), 311–325.

[18] S.S. DRAGOMIR AND N.M. IONESCU, Some Converse of Jensen's inequality and applications, *Anal. Num. Theor. Approx. (*Cluj-Napoca), **23** (1994), 71–78.

[19] S.S. DRAGOMIR, C.E.M. PEARCE AND J. E. PEČARIĆ, New inequalities for logarithmic map and their applications for entropy and mutual information, *Kyungpook Math. J.*, in press.

[20] S.S. DRAGOMIR, C.E.M. PEARCE AND J. E. PEČARIĆ, On Jensen's and related inequalities for isotonic sublinear functionals, *Acta Sci. Math.*, (Szeged), **61** (1995), 373–382.

[21] S.S. DRAGOMIR, J. E. PEČARIĆ AND L.E. PERSSON, Properties of some functionals related to Jensen's inequality, *Acta Math. Hungarica*, **20**(1-2) (1998), 129–143.

[22] T. KOSKI AND L.E. PERSSON, Some properties of generalized entropy with applications to compression of data, *J. of Int. Sciences*, **62** (1992), 103–132.

[23] M. MATIĆ, *Jensen's Inequality and Applications in Information Theory*, (in Croatian), Ph.D. Dissertation, Split, Croatia, 1998.

[24] R.J. McELIECE, *The Theory of Information and Coding*, Addison Wesley Publishing Company, Reading, 1977.

[25] J. E. PEČARIĆ, F. PROSCHAN AND Y.L. TONG, *Convex Functions, Partial Orderings and Statistical Applications*, Academic Press, 1992.

[26] J. PEETRE AND L.E. PERSSON, A general Beckenbach's inequality with applications, In: Function Spaces, Differential Operators and Nonlinear Analysis, *Pitman Research Notes in Math.*, **211** (1989), 125–139.