Received 05/07/14

A SURVEY OF REVERSE INEQUALITIES FOR f-DIVERGENCE MEASURE IN INFORMATION THEORY

S.S. DRAGOMIR[1,2]

Abstract. In this paper we survey some discrete inequalities for the f-divergence measure in Infromation Theory by the use of recent reverses of the celebrated Jensen's inequality. Applications in connection with Hölder's inequality and for particular measures such as Kullback-Leibler divergence measure, Hellinger discrimination, χ^2-distance and variation distance are provided as well.

1. Introduction

Given a convex function $f:\mathbb{R}_+\to\mathbb{R}_+$, the *$f$-divergence functional,* or *f-divergence measure*

$$I_f(p,q) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{1.1}$$

was introduced by Csiszár in [13], [14] as a generalized measure of information, a "distance function" on the set of probability distributions $\mathbb{P}^n$.

The restriction to discrete distributions is only for convenience, similar results hold for more general distributions.

The definition (1.1) can be extended for nonconvex function, however in this case the positivity property of $I_f(p,q)$ is not always assured.

As in Csiszár [14], we interpret the following, otherwise undefined expressions, as indicated:

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

The immediately following results were essentially given by Csiszár and Körner [15].

Theorem 1 (Csiszár & Körner, 1981 [15]). *If $f:\mathbb{R}_+\to\mathbb{R}$ is convex, then $I_f(p,q)$ is jointly convex in p and q.*

The following lower bound for the f-divergence functional also holds.

Theorem 2 (Csiszár & Körner, 1981 [15]). *Let $f:\mathbb{R}_+\to\mathbb{R}_+$ be convex, then for every $p,q\in\mathbb{R}_+^n$, we have the inequality:*

$$I_f(p,q) \geq \sum_{i=1}^{n} q_i f\left(\frac{\sum_{i=1}^{n} p_i}{\sum_{i=1}^{n} q_i}\right). \tag{1.2}$$

Key words and phrases. Convex functions, Jensen's inequality, Reverse of Jensen's inequality, Reverse of Hölder's inequality, f-divergence measure, Kullback-Leibler divergence measure, Hellinger discrimination, χ^2-distance, Variation distance, Grüss' type inequality.

If f is strictly convex, equality holds in (1.2) iff

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \cdots = \frac{p_n}{q_n}. \tag{1.3}$$

Corollary 1. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be convex and normalized, i.e.,*

$$f(1) = 0, \tag{1.4}$$

then, for any $p, q \in \mathbb{R}_+^n$ with

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i, \tag{1.5}$$

we have the inequality,

$$I_f(p, q) \geq 0. \tag{1.6}$$

If f is strictly convex, equality holds in (1.6) iff $p_i = q_i$ for all $i \in \{1, \ldots, n\}$.

In particular, if p, q are probability vectors, then (1.5) is assured. Corollary 1 then shows that, for strictly convex and normalized $f : \mathbb{R}_+ \to \mathbb{R}$,

$$I_f(p, q) \geq 0 \quad \text{for all } p, q \in \mathbb{P}^n. \tag{1.7}$$

The equality holds in (1.7) iff $p = q$.

These are "distance properties", however, I_f is not a metric since it violates the triangle inequality, and is asymmetric, i.e, for general $p, q \in \mathbb{R}_+^n$, $I_f(p, q) \neq I_f(q, p)$.

2. Some Examples

In the examples below we obtain, for suitable choices of the kernel f, some of the best known distance functions I_f used in mathematical statistics, information theory and signal processing, see [1]-[12], [16], [52]-[60] and [65]-[92].

Example 1. *(Kullback-Leibler) For*

$$f(t) := t \log t, \; t > 0 \tag{2.1}$$

the f-divergence is

$$I_f(p, q) = KL(p, q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right), \tag{2.2}$$

called the Kullback-Leibler distance [63]-[64].

Example 2. *(Hellinger) Let*

$$f(t) = \frac{1}{2}\left(1 - \sqrt{t}\right)^2, \; t > 0. \tag{2.3}$$

Then I_f gives the Hellinger distance [70]

$$I_f(p, q) = h^2(p, q) = \frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{p_i} - \sqrt{q_i}\right)^2, \tag{2.4}$$

which is symmetric.

Example 3. *(Renyi) For $\alpha > 1$, let*

$$f(t) = t^{\alpha},\ t > 0. \tag{2.5}$$

Then

$$I_f(p,q) = D_{\alpha}(p,q) = \sum_{i=1}^{n} p_i^{\alpha} q_i^{1-\alpha}, \tag{2.6}$$

which is the α-order entropy [80].

Example 4. *(χ^2-distance) Let*

$$f(t) = (t-1)^2,\ t > 0. \tag{2.7}$$

Then

$$\begin{aligned} I_f(p,q) = D_{\chi^2}(p,q) &= \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i} \\ &= \sum_{i=1}^{n} \frac{p_i^2}{q_i} - 2P_n + Q_n \\ &\left(= \sum_{i=1}^{n} \frac{p_i^2 - q_i^2}{q_i} \quad \text{if } P_n = Q_n \right) \end{aligned} \tag{2.8}$$

is the χ^2-distance between p and q, where $P_n = \sum_{i=1}^{n} p_i$ and $Q_n = \sum_{i=1}^{n} q_i$.

Finally, we have

Example 5. *(Variation distance). Let $f(t) = |t-1|,\ t > 0$. The corresponding f-divergence, called the variation distance, is symmetric,*

$$V(p,q) = \sum_{i=1}^{n} |p_i - q_i|.$$

For other examples of divergence measures, see the paper [61] by J. N. Kapur, where further references are given.

For other examples of divergence measures and further references, see [61] and [85].

In this paper we survey some discrete inequalities for the f-divergence measure in Infromation Theory by the use of recent reverses of the celebrated Jensen's inequality. Applications in connection with Hölder's inequality and for particular measures such as Kullback-Leibler divergence measure, Hellinger discrimination, χ^2-distance and variation distance are provided as well.

3. A Reverse Inequality Due to Dragomir & Ionescu

If $x_i, y_i \in \mathbb{R}$ and $w_i \geq 0$ $(i = 1, \ldots, n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$ then we may consider the *Čebyšev functional*

$$T_w(x,y) := \sum_{i=1}^{n} w_i x_i y_i - \sum_{i=1}^{n} w_i x_i \sum_{i=1}^{n} w_i y_i. \tag{3.1}$$

The following result is known in the literature as the *Grüss inequality*

$$|T_w(x,y)| \leq \frac{1}{4}(\Gamma - \gamma)(\Delta - \delta), \tag{3.2}$$

provided

$$-\infty < \gamma \le x_i \le \Gamma < \infty, \ \ -\infty < \delta \le y_i \le \Delta < \infty \tag{3.3}$$

for $i = 1, \dots, n$.

The constant $\frac{1}{4}$ is sharp in the sense that it cannot be replaced by a smaller constant.

If we assume that $-\infty < \gamma \le x_i \le \Gamma < \infty$ for $i = 1, \dots, n$, then by the Grüss inequality for $y_i = x_i$ and by the Schwarz's discrete inequality, we have

$$\sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right| \le \left[\sum_{i=1}^{n} w_i x_i^2 - \left(\sum_{j=1}^{n} w_j x_j \right)^2 \right]^{\frac{1}{2}} \le \frac{1}{2} (\Gamma - \gamma) . \tag{3.4}$$

In order to provide a reverse of the celebrated Jensen's inequality for convex functions, S.S. Dragomir obtained in 2002 [28] the following result:

Theorem 3. *Let $f : [m, M] \to \mathbb{R}$ be a differentiable convex function on (m, M). If $x_i \in [m, M]$ and $w_i \ge 0 \ \ (i = 1, \dots, n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$, then one has the counterpart of Jensen's weighted discrete inequality:*

$$\begin{aligned} 0 &\le \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i \right) \\ &\le \sum_{i=1}^{n} w_i f'(x_i) x_i - \sum_{i=1}^{n} w_i f'(x_i) \sum_{i=1}^{n} w_i x_i \\ &\le \frac{1}{2} \left[f'(M) - f'(m) \right] \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right| . \end{aligned} \tag{3.5}$$

Remark 1. *We notice that the inequality between the first and the second term in (3.5) was proved in 1994 by Dragomir & Ionescu, see* [45].

On making use of (3.4), we can state the following string of reverse inequalities

$$\begin{aligned} 0 &\le \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i \right) \\ &\le \sum_{i=1}^{n} w_i f'(x_i) x_i - \sum_{i=1}^{n} w_i f'(x_i) \sum_{i=1}^{n} w_i x_i \\ &\le \frac{1}{2} \left[f'(M) - f'(m) \right] \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right| \\ &\le \frac{1}{2} \left[f'(M) - f'(m) \right] \left[\sum_{i=1}^{n} w_i x_i^2 - \left(\sum_{j=1}^{n} w_j x_j \right)^2 \right]^{\frac{1}{2}} \\ &\le \frac{1}{4} \left[f'(M) - f'(m) \right] (M - m) , \end{aligned} \tag{3.6}$$

provided that $f : [m, M] \subset \mathbb{R} \to \mathbb{R}$ is a differentiable convex function on (m, M), $x_i \in [m, M]$ and $w_i \ge 0 \ \ (i = 1, \dots, n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$.

Remark 2. *We notice that the inequality between the first, second and last term from (3.6) was proved in the general case of positive linear functionals in 2001 by S. S. Dragomir in* [24].

For various Jensen's type inequalities see [17]-[51].

4. Further Reverse Inequalities

The following reverse of the Jensen's inequality holds:

Theorem 4 (Dragomir, 2013 [43])**.** *Let $f : I \to \mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \mathring{I}$, $\mathring{I}$ is the interior of I. If $x_i \in [m, M]$ and $w_i \geq 0 \quad (i = 1, \ldots, n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$, then*

$$\begin{aligned}
(4.1)\qquad 0 &\leq \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i\right) \\
&\leq \frac{\left(M - \sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M - m} \Psi_f\left(\sum_{i=1}^{n} w_i x_i; m, M\right) \\
&\leq \frac{\left(M - \sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M - m} \sup_{t \in (m,M)} \Psi_f(t; m, M) \\
&\leq \left(M - \sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right) \frac{f'_-(M) - f'_+(m)}{M - m} \\
&\leq \frac{1}{4}(M - m)\left[f'_-(M) - f'_+(m)\right],
\end{aligned}$$

where $\Psi_f(\cdot; m, M) : (m, M) \to \mathbb{R}$ is defined by

$$\Psi_f(t; m, M) = \frac{f(M) - f(t)}{M - t} - \frac{f(t) - f(m)}{t - m}.$$

We also have the inequality

$$\begin{aligned}
(4.2)\qquad 0 &\leq \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i\right) \\
&\leq \frac{\left(M - \sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M - m} \Psi_f\left(\sum_{i=1}^{n} w_i x_i; m, M\right) \\
&\leq \frac{1}{4}(M - m)\Psi_f\left(\sum_{i=1}^{n} w_i x_i; m, M\right) \\
&\leq \frac{1}{4}(M - m) \sup_{t \in (m,M)} \Psi_f(t; m, M) \\
&\leq \frac{1}{4}(M - m)\left[f'_-(M) - f'_+(m)\right],
\end{aligned}$$

provided that $\sum_{i=1}^{n} w_i x_i \in (m, M)$.

Proof. By the convexity of f we have that

$$\begin{aligned}
(4.3)\qquad & \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i\right) \\
&= \sum_{i=1}^{n} w_i f\left[\frac{m(M-x_i)+M(x_i-m)}{M-m}\right] \\
&- f\left(\sum_{i=1}^{n} w_i \left[\frac{m(M-x_i)+M(x_i-m)}{M-m}\right]\right) \\
&\leq \sum_{i=1}^{n} w_i \frac{(M-x_i) f(m) + (x_i-m) f(M)}{M-m} \\
&- f\left(\frac{m\left(M-\sum_{i=1}^{n} w_i x_i\right) + M\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M-m}\right) \\
&= \frac{\left(M-\sum_{i=1}^{n} w_i x_i\right) f(m) + \left(\sum_{i=1}^{n} w_i x_i - m\right) f(M)}{M-m} \\
&- f\left(\frac{m\left(M-\sum_{i=1}^{n} w_i x_i\right) + M\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M-m}\right) := B.
\end{aligned}$$

By denoting

$$\Delta_f(t; m, M) := \frac{(t-m) f(M) + (M-t) f(m)}{M-m} - f(t), \quad t \in [m, M]$$

we have

$$\begin{aligned}
(4.4)\qquad \Delta_f(t; m, M) &= \frac{(t-m) f(M) + (M-t) f(m) - (M-m) f(t)}{M-m} \\
&= \frac{(t-m) f(M) + (M-t) f(m) - (M-t+t-m) f(t)}{M-m} \\
&= \frac{(t-m)[f(M)-f(t)] - (M-t)[f(t)-f(m)]}{M-m} \\
&= \frac{(M-t)(t-m)}{M-m} \Psi_f(t; m, M)
\end{aligned}$$

for any $t \in (m, M)$.

Therefore we have the equality

$$(4.5)\qquad B = \frac{\left(M-\sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M-m} \Psi_f\left(\sum_{i=1}^{n} w_i x_i; m, M\right)$$

provided that $\sum_{i=1}^{n} w_i x_i \in (m, M)$.

For $\sum_{i=1}^{n} w_i x_i = m$ or $\sum_{i=1}^{n} w_i x_i = M$ the inequality (4.1) is obvious. If $\sum_{i=1}^{n} w_i x_i \in (m, M)$, then

$$\begin{aligned}
\Psi_f\left(\sum_{i=1}^{n} w_i x_i; m, M\right) &\leq \sup_{t\in(m,M)} \Psi_f(t; m, M) \\
&= \sup_{t\in(m,M)} \left[\frac{f(M)-f(t)}{M-t} - \frac{f(t)-f(m)}{t-m}\right] \\
&\leq \sup_{t\in(m,M)} \left[\frac{f(M)-f(t)}{M-t}\right] + \sup_{t\in(m,M)} \left[-\frac{f(t)-f(m)}{t-m}\right] \\
&= \sup_{t\in(m,M)} \left[\frac{f(M)-f(t)}{M-t}\right] - \inf_{t\in(m,M)} \left[\frac{f(t)-f(m)}{t-m}\right] \\
&= f'_-(M) - f'_+(m)
\end{aligned}$$

which by (4.3) and (4.5) produces the desired result (4.1).

Since, obviously

$$\frac{\left(M - \sum_{i=1}^{n} w_i x_i\right)\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M-m} \leq \frac{1}{4}(M-m),$$

then by (4.3) and (4.5) we deduce the second inequality (4.2). The last part is clear. □

Corollary 2. *Let $f : I \to \mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \mathring{I}$. If $x_i \in [m, M]$, then we have the inequalities*

$$\begin{aligned}
(4.6)\qquad 0 &\leq \frac{1}{n}\sum_{i=1}^{n} f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \\
&\leq \frac{\left(M - \frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} x_i - m\right)}{M-m} \Psi_f\left(\frac{1}{n}\sum_{i=1}^{n} x_i; m, M\right) \\
&\leq \frac{\left(M - \frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} x_i - m\right)}{M-m} \sup_{t\in(m,M)} \Psi_f(t; m, M) \\
&\leq \left(M - \frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} x_i - m\right)\frac{f'_-(M) - f'_+(m)}{M-m} \\
&\leq \frac{1}{4}(M-m)\left[f'_-(M) - f'_+(m)\right],
\end{aligned}$$

and

$$0 \leq \frac{1}{n}\sum_{i=1}^{n} f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \tag{4.7}$$
$$\frac{\left(M - \frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} x_i - m\right)}{M-m}\Psi_f\left(\frac{1}{n}\sum_{i=1}^{n} x_i; m, M\right)$$
$$\leq \frac{1}{4}(M-m)\Psi_f\left(\frac{1}{n}\sum_{i=1}^{n} x_i; m, M\right)$$
$$\leq \frac{1}{4}(M-m)\sup_{t\in(m,M)}\Psi_f(t; m, M)$$
$$\leq \frac{1}{4}(M-m)\left[f'_-(M) - f'_+(m)\right],$$

where $\frac{1}{n}\sum_{i=1}^{n} x_i \in (m, M)$.

Remark 3. *Define the weighted arithmetic mean of the positive n-tuple* $x = (x_1, ..., x_n)$ *with the nonnegative weights* $w = (w_1, ..., w_n)$ *by*

$$A_n(w, x) := \frac{1}{W_n}\sum_{i=1}^{n} w_i x_i$$

where $W_n := \sum_{i=1}^{n} w_i > 0$ *and the weighted geometric mean of the same n-tuple, by*

$$G_n(w, x) := \left(\prod_{i=1}^{n} x_i^{w_i}\right)^{1/W_n}.$$

It is well know that the following arithmetic mean-geometric mean inequality holds true

$$A_n(w, x) \geq G_n(w, x).$$

Applying the inequality between the first and third term in (4.6) for the convex function $f(t) = -\ln t, t > 0$ *we have*

$$1 \leq \frac{A_n(w,x)}{G_n(w,x)} \leq \exp\left[\frac{1}{Mm}(M - A_n(w,x))(A_n(w,x) - m)\right] \tag{4.8}$$
$$\leq \exp\left[\frac{1}{4}\frac{(M-m)^2}{mM}\right],$$

provided that $0 < m \leq x_i \leq M < \infty$ *for* $i \in \{1, ..., n\}$.

Also, if we apply the inequality (4.7) for the same function f we get that

$$1 \leq \frac{A_n(w,x)}{G_n(w,x)} \tag{4.9}$$
$$\leq \left[\left(\frac{M}{A_n(w,x)}\right)^{M-A_n(w,x)}\left(\frac{m}{A_n(w,x)}\right)^{A_n(w,x)-m}\right]^{-\frac{1}{4}(M-m)}$$
$$\leq \exp\left[\frac{1}{4}\frac{(M-m)^2}{mM}\right].$$

The following result also holds:

Theorem 5 (Dragomir, 2013 [43]). *With the assumptions of Theorem 4, we have the inequalities*

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i\right) \\ &\leq 2\max\left\{\frac{M-\sum_{i=1}^{n} w_i x_i}{M-m}, \frac{\sum_{i=1}^{n} w_i x_i - m}{M-m}\right\} \\ &\times \left[\frac{f(m)+f(M)}{2} - f\left(\frac{m+M}{2}\right)\right] \\ &\leq \frac{1}{2}\max\left\{M-\sum_{i=1}^{n} w_i x_i, \sum_{i=1}^{n} w_i x_i - m\right\}\left[f'_-(M) - f'_+(m)\right]. \end{aligned} \tag{4.10}$$

Proof. First of all, we recall the following result obtained by the author in [36] that provides a refinement and a reverse for the weighted Jensen's discrete inequality:

$$\begin{aligned} & n \min_{i\in\{1,\dots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^{n} f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right] \\ &\leq \frac{1}{P_n}\sum_{i=1}^{n} p_i f(x_i) - f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \\ & n \max_{i\in\{1,\dots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^{n} f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\right], \end{aligned} \tag{4.11}$$

where $f : C \to \mathbb{R}$ is a convex function defined on the convex subset C of the linear space X, $\{x_i\}_{i\in\{1,\dots,n\}} \subset C$ are vectors and $\{p_i\}_{i\in\{1,\dots,n\}}$ are nonnegative numbers with $P_n := \sum_{i=1}^{n} p_i > 0$.

For $n = 2$ we deduce from (4.11) that

$$\begin{aligned} & 2\min\{t, 1-t\}\left[\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)\right] \\ &\leq t f(x) + (1-t) f(y) - f(tx + (1-t) y) \\ &\leq 2\max\{t, 1-t\}\left[\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)\right] \end{aligned} \tag{4.12}$$

for any $x, y \in C$ and $t \in [0,1]$.

If we use the second inequality in (4.12) for the convex function $f : I \to \mathbb{R}$ and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \mathring{I}$, we have for $t = \frac{M-\sum_{i=1}^{n} w_i x_i}{M-m}$ that

$$\begin{aligned} & \frac{\left(M-\sum_{i=1}^{n} w_i x_i\right) f(m) + \left(\sum_{i=1}^{n} w_i x_i - m\right) f(M)}{M-m} \\ & - f\left(\frac{m\left(M-\sum_{i=1}^{n} w_i x_i\right) + M\left(\sum_{i=1}^{n} w_i x_i - m\right)}{M-m}\right) \\ &\leq 2\max\left\{\frac{M-\sum_{i=1}^{n} w_i x_i}{M-m}, \frac{\sum_{i=1}^{n} w_i x_i - m}{M-m}\right\} \\ &\times \left[\frac{f(m)+f(M)}{2} - f\left(\frac{m+M}{2}\right)\right]. \end{aligned} \tag{4.13}$$

Utilizing the inequality (4.3) and (4.13) we deduce the first inequality in (4.10).

Since

$$\frac{\frac{f(m)+f(M)}{2}-f\left(\frac{m+M}{2}\right)}{M-m} = \frac{1}{4}\left[\frac{f(M)-f\left(\frac{m+M}{2}\right)}{M-\frac{m+M}{2}}-\frac{f\left(\frac{m+M}{2}\right)-f(m)}{\frac{m+M}{2}-m}\right]$$

and, by the gradient inequality, we have that

$$\frac{f(M)-f\left(\frac{m+M}{2}\right)}{M-\frac{m+M}{2}} \le f'_{-}(M)$$

and

$$\frac{f\left(\frac{m+M}{2}\right)-f(m)}{\frac{m+M}{2}-m} \ge f'_{+}(m),$$

then we get

$$\frac{\frac{f(m)+f(M)}{2}-f\left(\frac{m+M}{2}\right)}{M-m} \le \frac{1}{4}\left[f'_{-}(M)-f'_{+}(m)\right]. \tag{4.14}$$

On making use of (4.13) and (4.14) we deduce the last part of (4.10). □

Corollary 3. *With the assumptions in Corollary 2, we have the inequalities*

$$0 \le \frac{1}{n}\sum_{i=1}^{n} f(x_i) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \tag{4.15}$$

$$\le 2\max\left\{\frac{M-\frac{1}{n}\sum_{i=1}^{n}x_i}{M-m}, \frac{\frac{1}{n}\sum_{i=1}^{n}x_i - m}{M-m}\right\} \times\left[\frac{f(m)+f(M)}{2}-f\left(\frac{m+M}{2}\right)\right] \tag{4.16}$$

$$\le \frac{1}{2}\max\left\{M-\frac{1}{n}\sum_{i=1}^{n}x_i, \frac{1}{n}\sum_{i=1}^{n}x_i - m\right\}\left[f'_{-}(M)-f'_{+}(m)\right].$$

Remark 4. *Since, obviously,*

$$\frac{M-\sum_{i=1}^{n}w_i x_i}{M-m}, \frac{\sum_{i=1}^{n}w_i x_i - m}{M-m} \le 1,$$

then we obtain from the first inequality in (4.10) the simpler, however coarser inequality, namely

$$0 \le \sum_{i=1}^{n} w_i f(x_i) - f\left(\sum_{i=1}^{n} w_i x_i\right) \le 2\left[\frac{f(m)+f(M)}{2}-f\left(\frac{m+M}{2}\right)\right]. \tag{4.17}$$

This inequality was obtained in 2008 by S. Simic in [83].

Remark 5. *With the assumptions in Remark 3 we have the following reverse of the arithmetic mean-geometric mean inequality*

$$1 \le \frac{A_n(w,x)}{G_n(w,x)} \le \left(\frac{A(m,M)}{G(m,M)}\right)^{2\max\left\{\frac{M-A_n(w,x)}{M-m}, \frac{A_n(w,x)-m}{M-m}\right\}}, \tag{4.18}$$

where $A(m,M)$ *is the arithmetic mean while* $G(m,M)$ *is the geometric mean of the positive numbers* m *and* M*.*

5. Applications for the Hölder Inequality

If $x_i, y_i \geq 0$ for $i \in \{1,...,n\}$, then the *Hölder inequality* holds true

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \left(\sum_{i=1}^{n} y_i^q\right)^{1/q},$$

where $p > 1, \frac{1}{p} + \frac{1}{q} = 1$.

Assume that $p > 1$. If $z_i \in \mathbb{R}$ for $i \in \{1,...,n\}$, satisfies the bounds

$$0 < m \leq z_i \leq M < \infty \text{ for } i \in \{1,...,n\}$$

and $w_i \geq 0 \ \ (i = 1, \ldots, n)$ with $W_n := \sum_{i=1}^{n} w_i > 0$, then from (4.1) we have

$$\begin{aligned} 0 &\leq \frac{\sum_{i=1}^{n} w_i z_i^p}{W_n} - \left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n}\right)^p \\ &\leq \frac{\left(M - \frac{\sum_{i=1}^{n} w_i z_i}{W_n}\right)\left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n} - m\right)}{M-m} B_p(m,M) \\ &\leq p\frac{M^{p-1} - m^{p-1}}{M-m}\left(M - \frac{\sum_{i=1}^{n} w_i z_i}{W_n}\right)\left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n} - m\right) \\ &\leq \frac{1}{4} p (M-m)\left(M^{p-1} - m^{p-1}\right), \end{aligned} \tag{5.1}$$

where $\Psi_p(\cdot; m, M) : (m, M) \to \mathbb{R}$ is defined by

$$\Psi_p(t; m, M) = \frac{M^p - t^p}{M - t} - \frac{t^p - m^p}{t - m}$$

while

$$B_p(m,M) := \sup_{t \in (m,M)} \Psi_p(t; m, M). \tag{5.2}$$

From (4.2) we also have the inequality

$$\begin{aligned} 0 &\leq \frac{\sum_{i=1}^{n} w_i z_i^p}{W_n} - \left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n}\right)^p \\ &\leq \frac{1}{4}(M-m)\Psi_p\left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n}; m, M\right) \leq \frac{1}{4} p (M-m)\left(M^{p-1} - m^{p-1}\right). \end{aligned} \tag{5.3}$$

Proposition 1 (Dragomir, 2013 [43])**.** *If* $x_i \geq 0$, $y_i > 0$ *for* $i \in \{1,...,n\}$, $p > 1, \frac{1}{p} + \frac{1}{q} = 1$ *and there exists the constants* $\gamma, \Gamma > 0$ *and such that*

$$\gamma \leq \frac{x_i}{y_i^{q-1}} \leq \Gamma \text{ for } i \in \{1,...,n\},$$

then we have

$$0 \le \frac{\sum_{i=1}^n x_i^p}{\sum_{i=1}^n y_i^q} - \left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q}\right)^p \le \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q}\right)\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q} - \gamma\right) \le p\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\left(\Gamma - \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q}\right)\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q} - \gamma\right) \le \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right), \tag{5.4}$$

and

$$0 \le \frac{\sum_{i=1}^n x_i^p}{\sum_{i=1}^n y_i^q} - \left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q}\right)^p \le \frac{1}{4}(\Gamma-\gamma)\Psi_p\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^q};\gamma,\Gamma\right) \le \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right), \tag{5.5}$$

where $B_p(\cdot,\cdot)$ and $\Psi_p(\cdot;\cdot,\cdot)$ are defined above.

Proof. The inequalities (5.4) and (5.5) follow from (5.1) and (5.3) by choosing

$$z_i = \frac{x_i}{y_i^{q-1}} \text{ and } w_i = y_i^q.$$

The details are omitted. □

Remark 6. *We observe that for $p = q = 2$ we have $\Psi_2(t;\gamma,\Gamma) = \Gamma - \gamma = B_2(\gamma,\Gamma)$ and then from the first inequality in (5.4) we get the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality:*

$$\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n x_i y_i\right)^2 \le \left(\Gamma - \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2}\right)\left(\frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n y_i^2} - \gamma\right)\left(\sum_{i=1}^n y_i^2\right)^2 \tag{5.6}$$

provided that $x_i \ge 0$, $y_i > 0$ for $i \in \{1,...,n\}$ and there exists the constants $\gamma,\Gamma > 0$ such that

$$\gamma \le \frac{x_i}{y_i} \le \Gamma \text{ for } i \in \{1,...,n\}.$$

Corollary 4 (Dragomir, 2013 [43])**.** *With the assumptions of Proposition 1 we have the following additive reverses of the Hölder inequality*

$$
\begin{aligned}
(5.7)\qquad 0 &\leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p}\left(\sum_{i=1}^{n} y_i^q\right)^{1/q} - \sum_{i=1}^{n} x_i y_i \\
&\leq \left[\frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\right]^{1/p}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^{1/p}\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)^{1/p} \\
&\times \sum_{i=1}^{n} y_i^q \\
&\leq p^{1/p}\left(\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\right)^{1/p}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^{1/p}\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)^{1/p} \\
&\times \sum_{i=1}^{n} y_i^q \\
&\leq \frac{1}{4^{1/p}} p^{1/p}(\Gamma-\gamma)^{1/p}\left(\Gamma^{p-1}-\gamma^{p-1}\right)^{1/p}\sum_{i=1}^{n} y_i^q
\end{aligned}
$$

and

$$
\begin{aligned}
(5.8)\qquad 0 &\leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p}\left(\sum_{i=1}^{n} y_i^q\right)^{1/q} - \sum_{i=1}^{n} x_i y_i \\
&\leq \frac{1}{4^{1/p}}(\Gamma-\gamma)^{1/p}\,\Psi_p^{1/p}\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}; m, M\right)\sum_{i=1}^{n} y_i^q \\
&\leq \frac{1}{4^{1/p}} p^{1/p}(\Gamma-\gamma)^{1/p}\left(\Gamma^{p-1}-\gamma^{p-1}\right)^{1/p}\sum_{i=1}^{n} y_i^q
\end{aligned}
$$

where $p > 1$ and $\frac{1}{p}+\frac{1}{q}=1$.

Proof. By multiplying in (5.4) with $(\sum_{i=1}^{n} y_i^q)^p$ we have

$$
\begin{aligned}
&\sum_{i=1}^{n} x_i^p\left(\sum_{i=1}^{n} y_i^q\right)^{p-1} - \left(\sum_{i=1}^{n} x_i y_i\right)^p \\
&\leq \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)\left(\sum_{i=1}^{n} y_i^q\right)^p \\
&\leq p\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)\left(\sum_{i=1}^{n} y_i^q\right)^p \\
&\leq \frac{1}{4} p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right)\left(\sum_{i=1}^{n} y_i^q\right)^p,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
(5.9)\qquad & \sum_{i=1}^{n} x_i^p \left(\sum_{i=1}^{n} y_i^q\right)^{p-1} \\
& \leq \left(\sum_{i=1}^{n} x_i y_i\right)^p + \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right) \\
& \times \left(\sum_{i=1}^{n} y_i^q\right)^p \\
& \leq \left(\sum_{i=1}^{n} x_i y_i\right)^p + p\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right) \\
& \times \left(\sum_{i=1}^{n} y_i^q\right)^p \frac{\Gamma^{p-1} - \gamma^{p-1}}{\Gamma - \gamma} \\
& \leq \left(\sum_{i=1}^{n} x_i y_i\right)^p + \frac{1}{4} p (\Gamma - \gamma)\left(\Gamma^{p-1} - \gamma^{p-1}\right)\left(\sum_{i=1}^{n} y_i^q\right)^p .
\end{aligned}
$$

Taking the power $1/p$ with $p > 1$ and employing the following elementary inequality that state that for $p > 1$ and $\alpha, \beta > 0$,

$$(\alpha+\beta)^{1/p} \leq \alpha^{1/p} + \beta^{1/p}$$

we have from the first part of (5.9) that

$$
\begin{aligned}
(5.10)\qquad & \left(\sum_{i=1}^{n} x_i y_i\right)^{1/p} \left(\sum_{i=1}^{n} y_i^q\right)^{1-\frac{1}{p}} \\
& \leq \sum_{i=1}^{n} x_i y_i + \left[\frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\right]^{1/p}\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^{1/p}\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)^{1/p} \\
& \times \sum_{i=1}^{n} y_i^q
\end{aligned}
$$

and since $1 - \frac{1}{p} = \frac{1}{q}$ we get from (5.10) the first inequality in (5.7). The rest is obvious.

The inequality (5.8) can be proved in a similar manner, however the details are omitted. □

If $z_i \in \mathbb{R}$ for $i \in \{1, ..., n\}$, satisfies the bounds

$$0 < m \leq z_i \leq M < \infty \text{ for } i \in \{1, ..., n\}$$

and $w_i \geq 0 \ \ (i = 1, \ldots, n)$ with $W_n := \sum_{i=1}^{n} w_i > 0$, then from (4.10) we also have the inequality

$$
\begin{aligned}
0 &\leq \frac{\sum_{i=1}^{n} w_i z_i^p}{W_n} - \left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n}\right)^p \\
&\leq 2\left[\frac{m^p + M^p}{2} - \left(\frac{m+M}{2}\right)^p\right] \\
&\times \max\left\{\frac{M - \frac{\sum_{i=1}^{n} w_i z_i}{W_n}}{M-m}, \frac{\frac{\sum_{i=1}^{n} w_i z_i}{W_n} - m}{M-m}\right\} \\
&\leq \frac{1}{2} p\left(M^{p-1} - m^{p-1}\right) \max\left\{M - \frac{\sum_{i=1}^{n} w_i z_i}{W_n}, \frac{\sum_{i=1}^{n} w_i z_i}{W_n} - m\right\}.
\end{aligned} \tag{5.11}
$$

From the inequality (5.11) we can state:

Proposition 2 (Dragomir, 2013 [43])**.** *With the assumptions of Proposition 1 we have*

$$
\begin{aligned}
0 &\leq \frac{\sum_{i=1}^{n} x_i^p}{\sum_{i=1}^{n} y_i^q} - \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^p \\
&\leq 2 \cdot \frac{\frac{\gamma^p + \Gamma^p}{2} - \left(\frac{\gamma+\Gamma}{2}\right)^p}{\Gamma - \gamma} \max\left\{\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}, \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right\} \\
&\leq \frac{1}{2} p\left(\Gamma^{p-1} - \gamma^{p-1}\right) \max\left\{\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}, \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right\}.
\end{aligned} \tag{5.12}
$$

Finally, the following additive reverse of the Hölder inequality can be stated as well:

Corollary 5 (Dragomir, 2013 [43])**.** *With the assumptions of Proposition 1 we have*

$$
\begin{aligned}
0 &\leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \left(\sum_{i=1}^{n} y_i^q\right)^{1/q} - \sum_{i=1}^{n} x_i y_i \\
&\leq 2^{1/p} \cdot \left(\frac{\frac{\gamma^p + \Gamma^p}{2} - \left(\frac{\gamma+\Gamma}{2}\right)^p}{\Gamma - \gamma}\right)^{1/p} \\
&\times \max\left\{\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^{1/p}, \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)^{1/p}\right\} \sum_{i=1}^{n} y_i^q \\
&\leq \frac{1}{2^{1/p}} p^{1/p} \max\left\{\left(\Gamma - \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^{1/p}, \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \gamma\right)^{1/p}\right\} \\
&\times \left(\Gamma^{p-1} - \gamma^{p-1}\right)^{1/p} \sum_{i=1}^{n} y_i^q.
\end{aligned} \tag{5.13}
$$

Remark 7. *As a simpler, however coarser inequality we have the following result:*

$$0 \leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \left(\sum_{i=1}^{n} y_i^q\right)^{1/q} - \sum_{i=1}^{n} x_i y_i \leq 2^{1/p} \cdot \left[\frac{\gamma^p + \Gamma^p}{2} - \left(\frac{\gamma + \Gamma}{2}\right)^p\right]^{1/p} \sum_{i=1}^{n} y_i^q, \tag{5.14}$$

where x_i and y_i are as above.

6. Applications for f-Divergence

Consider the f-divergence

$$I_f(p,q) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{6.1}$$

defined on the set of probability distributions $p, q \in \mathbb{P}^n$, where f is convex on $(0, \infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u = 1$. By appropriately defining this convex function, various divergences are derived.

The following result holds:

Proposition 3 (Dragomir, 2013 [43])**.** *Let $f : (0, \infty) \to \mathbb{R}$ be a convex function with the property that $f(1) = 0$. Assume that $p, q \in \mathbb{P}^n$ and there exists the constants $0 < r < 1 < R < \infty$ such that*

$$r \leq \frac{p_i}{q_i} \leq R \text{ for } i \in \{1, ..., n\}. \tag{6.2}$$

Then we have the inequalities

$$0 \leq I_f(p,q) \leq \frac{(R-1)(1-r)}{R-r} \sup_{t \in (r,R)} \Psi_f(t; r, R) \leq (R-1)(1-r) \frac{f'_-(R) - f'_+(r)}{R-r} \leq \frac{1}{4}(R-r)\left[f'_-(R) - f'_+(r)\right], \tag{6.3}$$

and $\Psi_f(\cdot; r, R) : (r, R) \to \mathbb{R}$ is defined by

$$\Psi_f(t; r, R) = \frac{f(R) - f(t)}{R-t} - \frac{f(t) - f(r)}{t-r}.$$

We also have the inequality

$$I_f(p,q) \leq \frac{1}{4}(R-r) \frac{f(R)(1-r) + f(r)(R-1)}{(R-1)(1-r)} \leq \frac{1}{4}(R-r)\left[f'_-(R) - f'_+(r)\right]. \tag{6.4}$$

The proof follows by Theorem 4 by choosing $w_i = q_i$, $x_i = \frac{p_i}{q_i}$, $m = r$ and $M = R$ and performing the required calculations. The details are omitted.

Utilising the same approach and Theorem 5 we can also state that:

Proposition 4 (Dragomir, 2013 [43])**.** *With the assumptions of Proposition 3 we have*

$$
\begin{aligned}
0 &\leq I_f(p,q) \leq 2\max\left\{\frac{R-1}{R-r},\frac{1-r}{R-r}\right\} \\
&\times\left[\frac{f(r)+f(R)}{2}-f\left(\frac{r+R}{2}\right)\right] \\
&\leq \frac{1}{2}\max\{R-1,1-r\}\left[f'_-(R)-f'_+(r)\right].
\end{aligned} \tag{6.5}
$$

The above results can be utilized to obtain various inequalities for the divergence measures in Information Theory that are particular instances of f-divergence.

Consider the Kullback-Leibler divergence

$$KL(p,q)=\sum_{i=1}^{n}p_i\log\left(\frac{p_i}{q_i}\right),\ p,q\in\mathbb{P}^n.$$

For the convex function $f:(0,\infty)\to\mathbb{R}$, $f(t)=-\ln t$ we have

$$I_f(p,q):=\sum_{i=1}^{n}q_i f\left(\frac{p_i}{q_i}\right)=-\sum_{i=1}^{n}q_i\ln\left(\frac{p_i}{q_i}\right)=\sum_{i=1}^{n}q_i\ln\left(\frac{q_i}{p_i}\right)=KL(q,p)$$

If $p,q\in\mathbb{P}^n$ such that there exists the constants $0<r<1<R<\infty$ with

$$r\leq\frac{p_i}{q_i}\leq R \text{ for } i\in\{1,...,n\}, \tag{6.6}$$

then we get from the second inequality in (6.3) that

$$0\leq KL(q,p)\leq\frac{(R-1)(1-r)}{rR}, \tag{6.7}$$

from the first inequality in (6.4) that

$$0\leq KL(q,p)\leq\frac{1}{4}(R-r)\ln\left[R^{-\frac{1}{R-1}}r^{-\frac{1}{1-r}}\right]$$

and from the first inequality in (6.5) that

$$0\leq KL(q,p)\leq 2\max\left\{\frac{R-1}{R-r},\frac{1-r}{R-r}\right\}\ln\left(\frac{A(r,R)}{G(r,R)}\right) \tag{6.8}$$

where $A(r,R)$ is the arithmetic mean and $G(r,R)$ is the geometric mean of the positive numbers r and R.

For the convex function $f:(0,\infty)\to\mathbb{R}$, $f(t)=t\ln t$ we have

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

If $p,q\in\mathbb{P}^n$ such that there exists the constants $0<r<1<R<\infty$ with the property (6.6), then we get from the second inequality in (6.3) that

$$0\leq KL(p,q)\leq\frac{(R-1)(1-r)}{L(r,R)}, \tag{6.9}$$

where $L(r,R)$ is the Logarithmic mean of r, R, namely

$$L(r,R)=\frac{R-r}{\ln R-\ln r}.$$

From the first inequality in (6.4) we also have:

$$0 \leq KL(p,q) \leq \frac{1}{4}(R-r)\frac{R-r+\ln\left(R^{1-r}r^{R-1}\right)}{(R-1)(1-r)}. \tag{6.10}$$

Finally, by the first inequality in (6.5) we have

$$0 \leq KL(p,q) \leq 2\max\left\{\frac{R-1}{R-r}, \frac{1-r}{R-r}\right\}\ln\left[\frac{G\left(r^r, R^R\right)}{\left[A(r,R)\right]^{A(r,R)}}\right]. \tag{6.11}$$

7. More Reverse Inequalities

For the *Lebesgue measurable* function $g:[\alpha,\beta]\to\mathbb{R}$ we introduce the *Lebesgue p-norms* defined as

$$\|g\|_{[\alpha,\beta],p} := \left(\int_\alpha^\beta |g(t)|^p\,dt\right)^{1/p} \quad \text{if } g\in L_p[\alpha,\beta],$$

for $p\geq 1$ and

$$\|g\|_{[\alpha,\beta],\infty} := ess\sup_{t\in[\alpha,\beta]}|g(t)| \text{ if } g\in L_\infty[\alpha,\beta],$$

for $p=\infty$.

The following result also holds:

Theorem 6 (Dragomir, 2013 [44])**.** *Let $\Phi: I\to\mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m,M\in\mathbb{R}$, $m<M$ with $[m,M]\subset\mathring{I}$, $\mathring{I}$ is the interior of I. If $x_i\in I$ and $w_i\geq 0$ for $i\in\{1,...,n\}$ with $\sum_{i=1}^n w_i=1$, denote $\bar{x}_w:=\sum_{i=1}^n w_i x_i\in I$, then we have the inequality*

$$\begin{aligned} 0 &\leq \sum_{i=1}^n w_i\Phi(x_i)-\Phi(\bar{x}_w) \\ &\leq \frac{(M-\bar{x}_w)\int_m^{\bar{x}_w}|\Phi'(t)|\,dt+(\bar{x}_w-m)\int_{\bar{x}_w}^M|\Phi'(t)|\,dt}{M-m} := \Theta_\Phi(\bar{x}_w;m,M), \end{aligned} \tag{7.1}$$

where $\Theta_\Phi(\bar{x}_w;m,M)$ satisfies the bounds

$$\Theta_\Phi(\bar{x}_w;m,M) \leq \begin{cases} \left[\frac{1}{2}+\frac{\left|\bar{x}_w-\frac{m+M}{2}\right|}{M-m}\right]\int_m^M|\Phi'(t)|\,dt \\[2ex] \left[\frac{1}{2}\int_m^M|\Phi'(t)|\,dt+\frac{1}{2}\left|\int_{\bar{x}_w}^M|\Phi'(t)|\,dt-\int_m^{\bar{x}_w}|\Phi'(t)|\,dt\right|\right], \end{cases} \tag{7.2}$$

$$\begin{aligned} \Theta_\Phi(\bar{x}_w;m,M) &\leq \frac{(\bar{x}_w-m)(M-\bar{x}_w)}{M-m}\left[\|\Phi'\|_{[\bar{x}_w,M],\infty}+\|\Phi'\|_{[m,\bar{x}_w],\infty}\right] \\ &\leq \frac{1}{2}(M-m)\frac{\|\Phi'\|_{[\bar{w}_p,M],\infty}+\|\Phi'\|_{[m,\bar{w}_p],\infty}}{2} \leq \frac{1}{2}(M-m)\|\Phi'\|_{[m,M],\infty} \end{aligned} \tag{7.3}$$

and

$$\begin{aligned}(7.4)\qquad &\Theta_{\Phi}(\bar{x}_w;m,M)\\ &\leq \frac{1}{M-m}\Big[(\bar{x}_w-m)(M-\bar{x}_w)^{1/q}\|\Phi'\|_{[\bar{x}_w,M],p}\\ &+(M-\bar{x}_w)(\bar{x}_w-m)^{1/q}\|\Phi'\|_{[m,\bar{x}_w],p}\Big]\\ &\leq \frac{1}{M-m}\left[(\bar{x}_w-m)^q(M-\bar{x}_w)+(M-\bar{x}_w)^q(\bar{x}_w-m)\right]^{1/q}\|\Phi'\|_{[m,M],p}\end{aligned}$$

where $p>1$, $\frac{1}{p}+\frac{1}{q}=1$.

Proof. By the convexity of Φ we have that

$$\begin{aligned}(7.5)\qquad &\sum_{i=1}^{n}w_i\Phi(x_i)-\Phi(\bar{x}_w)\\ &=\sum_{i=1}^{n}w_i\Phi\left[\frac{m(M-x_i)+M(x_i-m)}{M-m}\right]-\Phi(\bar{x}_w)\\ &\leq\sum_{i=1}^{n}w_i\frac{(M-x_i)\Phi(m)+(x_i-m)\Phi(M)}{M-m}-\Phi(\bar{x}_w)\\ &=\frac{(M-\bar{x}_w)\Phi(m)+(\bar{x}_w-m)\Phi(M)}{M-m}-\Phi(\bar{x}_w)=B.\end{aligned}$$

By denoting

$$\Lambda_{\Phi}(t;m,M):=\frac{(t-m)\Phi(M)+(M-t)\Phi(m)}{M-m}-\Phi(t),\quad t\in[m,M]$$

we have

$$\begin{aligned}(7.6)\qquad \Lambda_{\Phi}(t;m,M)&=\frac{(t-m)\Phi(M)+(M-t)\Phi(m)}{M-m}-\Phi(t)\\ &=\frac{(t-m)\Phi(M)+(M-t)\Phi(m)-(M-m)\Phi(t)}{M-m}\\ &=\frac{(t-m)\Phi(M)+(M-t)\Phi(m)-(M-t+t-m)\Phi(t)}{M-m}\\ &=\frac{(t-m)[\Phi(M)-\Phi(t)]-(M-t)[\Phi(t)-\Phi(m)]}{M-m}\end{aligned}$$

for any $t\in[m,M]$. Also

$$B=\Lambda_{\Phi}(\bar{x}_w;m,M).$$

Taking the modulus on (7.6) and, noticing that, by the convexity of Φ we have

$$\begin{aligned}&\Lambda_{\Phi}(t;m,M)\\ &=\frac{(t-m)\Phi(M)+(M-t)\Phi(m)}{M-m}-\Phi\left(\frac{(t-m)M+(M-t)m}{M-m}\right)\geq 0\end{aligned}$$

for any $t \in [m, M]$, then we have

$$\Lambda_{\Phi}(t; m, M) \leq \frac{(t-m)|\Phi(M) - \Phi(t)| + (M-t)|\Phi(t) - \Phi(m)|}{M-m} \tag{7.7}$$
$$= \frac{(t-m)\left|\int_t^M \Phi'(s)\,ds\right| + (M-t)\left|\int_m^t \Phi'(s)\,ds\right|}{M-m}$$
$$\leq \frac{(t-m)\int_t^M |\Phi'(s)|\,ds + (M-t)\int_m^t |\Phi'(s)|\,ds}{M-m}$$

for any $t \in [m, M]$.

Finally, if we write the inequality (7.7) for $t = \bar{x}_w \in [m, M]$ and utilize the inequality (7.5), we deduce the desired result (7.1).

Now, we observe that

$$\frac{(t-m)\int_t^M |\Phi'(s)|\,ds + (M-t)\int_m^t |\Phi'(s)|\,ds}{M-m} \tag{7.8}$$
$$\leq \begin{cases} \max\{t-m, M-t\} \int_m^M |\Phi'(t)|\,dt \\ \\ \max\left\{\int_t^M |\Phi'(s)|\,ds, \int_m^t |\Phi'(s)|\,ds\right\}(M-m) \end{cases}$$
$$= \begin{cases} \left[\frac{1}{2}(M-m) + \left|t - \frac{m+M}{2}\right|\right] \int_m^M |\Phi'(t)|\,dt \\ \\ \left[\frac{1}{2}\int_m^M |\Phi'(s)|\,ds + \frac{1}{2}\left|\int_t^M |\Phi'(s)|\,ds - \int_m^t |\Phi'(s)|\,ds\right|\right](M-m) \end{cases}$$

for any $t \in [m, M]$. This proves the inequality (7.2).

By the Hölder's inequality we have

$$\int_t^M |\Phi'(s)|\,ds \leq \begin{cases} (M-t)\|\Phi'\|_{[t,M],\infty} \\ \\ (M-t)^{1/q}\|\Phi'\|_{[t,M],p} & \text{if } p > 1, \frac{1}{p} + \frac{1}{q} = 1 \end{cases}$$

and

$$\int_m^t |\Phi'(s)|\,ds \leq \begin{cases} (t-m)\|\Phi'\|_{[m,t],\infty} \\ \\ (t-m)^{1/q}\|\Phi'\|_{[m,t],p} & \text{if } p > 1, \frac{1}{p} + \frac{1}{q} = 1, \end{cases}$$

which give that

$$\frac{(t-m)\int_t^M |\Phi'(s)|\,ds + (M-t)\int_m^t |\Phi'(s)|\,ds}{M-m} \tag{7.9}$$
$$\leq \frac{(t-m)(M-t)\|\Phi'\|_{[t,M],\infty} + (M-t)(t-m)\|\Phi'\|_{[m,t],\infty}}{M-m}$$
$$= \frac{(t-m)(M-t)}{M-m}\left[\|\Phi'\|_{[t,M],\infty} + \|\Phi'\|_{[m,t],\infty}\right]$$
$$\leq \frac{1}{2}(M-m)\frac{\|\Phi'\|_{[t,M],\infty} + \|\Phi'\|_{[m,t],\infty}}{2}$$
$$\leq \frac{1}{2}(M-m)\max\left\{\|\Phi'\|_{[t,M],\infty}, \|\Phi'\|_{[m,t],\infty}\right\} = \frac{1}{2}(M-m)\|\Phi'\|_{[m,M],\infty}$$

and

$$\begin{aligned}(7.10)\qquad & \frac{(t-m)\int_t^M |\Phi'(s)|\,ds+(M-t)\int_m^t |\Phi'(s)|\,ds}{M-m} \\ & \leq \frac{(t-m)(M-t)^{1/q}\|\Phi'\|_{[t,M],p}+(M-t)(t-m)^{1/q}\|\Phi'\|_{[m,t],p}}{M-m} \\ & \leq \frac{1}{M-m}\left[\left((t-m)(M-t)^{1/q}\right)^q+\left((M-t)(t-m)^{1/q}\right)^q\right]^{1/q} \\ & \times\left[\|\Phi'\|_{[t,M],p}^p+\|\Phi'\|_{[m,t],p}^p\right]^{1/p} \\ & = \frac{1}{M-m}\left[(t-m)^q(M-t)+(M-t)^q(t-m)\right]^{1/q}\|\Phi'\|_{[m,M],p}\end{aligned}$$

for any $t\in[m,M]$.

These prove the desired inequalities (7.3) and (7.4). □

Remark 8. *Define the weighted arithmetic mean of the positive n-tuple* $x=(x_1,...,x_n)$ *with the nonnegative weights* $w=(w_1,...,w_n)$ *by*

$$A_n(w,x):=\frac{1}{W_n}\sum_{i=1}^n w_i x_i$$

where $W_n:=\sum_{i=1}^n w_i>0$ *and the weighted geometric mean of the same n-tuple, by*

$$G_n(w,x):=\left(\prod_{i=1}^n x_i^{w_i}\right)^{1/W_n}.$$

It is well know that the following arithmetic mean-geometric mean inequality holds true

$$A_n(w,x)\geq G_n(w,x).$$

On applying the inequality (7.1) for the convex function $\Phi(t)=-\ln t$, *we have the following reverse of the arithmetic mean-geometric mean inequality*

$$(7.11)\qquad 1\leq\frac{A_n(w,x)}{G_n(w,x)}\leq\left(\frac{A_n(w,x)}{m}\right)^{M-A_n(w,x)}\left(\frac{M}{A_n(w,x)}\right)^{A_n(w,x)-m}.$$

8. Applications for the Hölder Inequality

If $x_i,y_i\geq 0$ for $i\in\{1,...,n\}$, then the *Hölder inequality* holds true

$$\sum_{i=1}^n x_i y_i\leq\left(\sum_{i=1}^n x_i^p\right)^{1/p}\left(\sum_{i=1}^n y_i^q\right)^{1/q},$$

where $p>1$, $\frac{1}{p}+\frac{1}{q}=1$.

Assume that $p>1$. If $z_i\in\mathbb{R}$ for $i\in\{1,...,n\}$, satisfies the bounds

$$0<m\leq z_i\leq M<\infty \text{ for } i\in\{1,...,n\}$$

and $w_i \geq 0 \quad (i = 1, \ldots, n)$ with $W_n := \sum_{i=1}^{n} w_i > 0$, then from Theorem 6 we have amongst other the following inequality

$$0 \leq \frac{\sum_{i=1}^{n} w_i z_i^p}{W_n} - \left(\frac{\sum_{i=1}^{n} w_i z_i}{W_n} \right)^p \leq (M^p - m^p) \left[\frac{1}{2} + \frac{1}{M-m} \left| \frac{\sum_{i=1}^{n} w_i z_i}{W_n} - \frac{m+M}{2} \right| \right]. \tag{8.1}$$

From this inequality we can state that:

Proposition 5 (Dragomir, 2013 [44])**.** *If $x_i \geq 0$, $y_i > 0$ for $i \in \{1, ..., n\}$, $p > 1, \frac{1}{p} + \frac{1}{q} = 1$ and there exists the constants $\gamma, \Gamma > 0$ and such that*

$$\gamma \leq \frac{x_i}{y_i^{q-1}} \leq \Gamma \textit{ for } i \in \{1, ..., n\},$$

then we have

$$0 \leq \frac{\sum_{i=1}^{n} x_i^p}{\sum_{i=1}^{n} y_i^q} - \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} \right)^p \leq (\Gamma^p - \gamma^p) \left[\frac{1}{2} + \frac{1}{\Gamma - \gamma} \left| \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \frac{\gamma + \Gamma}{2} \right| \right]. \tag{8.2}$$

Finally, the following additive reverse of the Hölder inequality can be stated as well:

Corollary 6 (Dragomir, 2013 [44])**.** *With the assumptions of Proposition 5 we have*

$$0 \leq \left(\sum_{i=1}^{n} x_i^p \right)^{1/p} \left(\sum_{i=1}^{n} y_i^q \right)^{1/q} - \sum_{i=1}^{n} x_i y_i \leq (\Gamma^p - \gamma^p)^{1/p} \left[\frac{1}{2} + \frac{1}{\Gamma - \gamma} \left| \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q} - \frac{\gamma + \Gamma}{2} \right| \right]^{1/p} \sum_{i=1}^{n} y_i^q. \tag{8.3}$$

Remark 9. *We observe that for $p = q = 2$ we have from the first inequality in (8.2) the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality*

$$\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} x_i y_i \right)^2 \leq \left(\Gamma^2 - \gamma^2 \right) \left[\frac{1}{2} + \frac{1}{\Gamma - \gamma} \left| \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^2} - \frac{\gamma + \Gamma}{2} \right| \right] \left(\sum_{i=1}^{n} y_i^2 \right)^2 \tag{8.4}$$

provided that $x_i \geq 0$, $y_i > 0$ for $i \in \{1, ..., n\}$ and there exists the constants $\gamma, \Gamma > 0$ such that

$$\gamma \leq \frac{x_i}{y_i} \leq \Gamma \textit{ for } i \in \{1, ..., n\}.$$

9. Applications for f-Divergence

Consider the f-divergence

$$I_f(p, q) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i} \right) \tag{9.1}$$

defined on the set of probability distributions $p, q \in \mathbb{P}^n$, where f is convex on $(0, \infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u = 1$. By appropriately defining this convex function, various divergences are derived.

Proposition 6 (Dragomir, 2013 [44])**.** *Let $f : (0, \infty) \to \mathbb{R}$ be a convex function with the property that $f(1) = 0$. Assume that $p, q \in \mathbb{P}^n$ and there exists the constants $0 < r < 1 < R < \infty$ such that*

$$r \leq \frac{p_i}{q_i} \leq R \text{ for } i \in \{1, ..., n\}. \tag{9.2}$$

Then we have the inequalities

$$0 \leq I_f(p, q) \leq B_f(r, R) \tag{9.3}$$

where

$$B_f(r, R) := \frac{(R-1)\int_r^1 |f'(t)|\,dt + (1-r)\int_1^R |f'(t)|\,dt}{R-r}. \tag{9.4}$$

Moreover, we have the following bounds for $B_f(r, R)$

$$\begin{aligned} &B_f(r, R) \\ &\leq \begin{cases} \left[\frac{1}{2} + \frac{\left|1-\frac{r+R}{2}\right|}{R-r}\right]\int_r^R |f'(t)|\,dt \\ \\ \left[\frac{1}{2}\int_r^R |f'(t)|\,dt + \frac{1}{2}\left|\int_1^R |f'(t)|\,dt - \int_r^1 |f'(t)|\,dt\right|\right], \end{cases} \end{aligned} \tag{9.5}$$

and

$$\begin{aligned} &B_f(r, R) \\ &\leq \frac{(1-r)(R-1)}{R-r}\left[\|f'\|_{[1,R],\infty} + \|f'\|_{[r,1],\infty}\right] \\ &\leq \frac{1}{2}(R-r)\frac{\|f'\|_{[1,R],\infty} + \|f'\|_{[r,1],\infty}}{2} \leq \frac{1}{2}(R-r)\|f'\|_{[r,R],\infty} \end{aligned} \tag{9.6}$$

and

$$\begin{aligned} &B_f(r, R) \\ &\leq \frac{1}{R-r}\left[(1-r)(R-1)^{1/q}\|f'\|_{[1,R],p} + (R-1)(1-r)^{1/q}\|f'\|_{[r,1],p}\right] \\ &\leq \frac{1}{R-r}\left[(1-r)^q(R-1) + (R-1)^q(1-r)\right]^{1/q}\|f'\|_{[r,R],p} \end{aligned} \tag{9.7}$$

where $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$.

The proof follows by Theorem 6 by choosing $w_i = q_i$, $x_i = \frac{p_i}{q_i}$, $m = r$ and $M = R$ and performing the required calculations. The details are omitted.

The above results can be utilized to obtain various inequalities for the divergence measures in information theory that are particular instances of f-divergence.

Consider the Kullback-Leibler divergence

$$KL(p, q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right), \ p, q \in \mathbb{P}^n.$$

For the convex function $f:(0,\infty)\to\mathbb{R}$, $f(t)=-\ln t$ we have

$$I_f(p,q):=\sum_{i=1}^{n}q_i f\left(\frac{p_i}{q_i}\right)=-\sum_{i=1}^{n}q_i\ln\left(\frac{p_i}{q_i}\right)=\sum_{i=1}^{n}q_i\ln\left(\frac{q_i}{p_i}\right)=KL(q,p)$$

If $p,q\in\mathbb{P}^n$ such that there exists the constants $0<r<1<R<\infty$ with

$$r\le\frac{p_i}{q_i}\le R\text{ for }i\in\{1,...,n\},\tag{9.8}$$

then we get from the inequality (9.4)

$$0\le KL(q,p)\le\ln\left(\frac{R^{1-r}}{r^{R-1}}\right)^{\frac{1}{R-r}}.\tag{9.9}$$

For $\alpha>1$, let

$$f(t)=t^{\alpha},\ t>0.$$

Then

$$I_f(p,q)=D_\alpha(p,q)=\sum_{i=1}^{n}p_i^{\alpha}q_i^{1-\alpha},$$

which is the α-order entropy.

If $p,q\in\mathbb{P}^n$ such that (9.8) holds true, then by (9.4) we have

$$0\le D_\alpha(p,q)\le\frac{(R-1)(1-r^{\alpha})+(1-r)(R^{\alpha}-1)}{R-r}.$$

10. A Refinement and Another Reverse

For a real function $g:[m,M]\to\mathbb{R}$ and two distinct points $\alpha,\beta\in[m,M]$ we recall that the *divided difference* of g in these points is defined by

$$[\alpha,\beta;g]:=\frac{g(\beta)-g(\alpha)}{\beta-\alpha}.$$

Theorem 7 (Dragomir, 2011 [41])**.** *Let $f:I\to\mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m,M\in\mathbb{R}$, $m<M$ with $[m,M]\subset\mathring{I}$, $\mathring{I}$ the interior of I. Let $\bar{\mathbf{a}}=(a_1,\dots,a_n)$, $\bar{\mathbf{p}}=(p_1,\dots,p_n)$ be $n-$tuples of real numbers with $p_i\ge0$ $(i\in\{1,\dots,n\})$ and $\sum_{i=1}^{n}p_i=1$. If $m\le a_i\le M$, $i\in\{1,\dots,n\}$, with $\sum_{i=1}^{n}p_ia_i\ne m,M$, then*

(10.1)

$$\begin{aligned}&\left|\sum_{i=1}^{n}p_i\left|f(a_i)-f\left(\sum_{j=1}^{n}p_ja_j\right)\right|sgn\left(a_i-\sum_{j=1}^{n}p_ja_j\right)\right|\\&\le\sum_{i=1}^{n}p_if(a_i)-f\left(\sum_{i=1}^{n}p_ia_i\right)\\&\le\frac{1}{2}\left(\left[\sum_{i=1}^{n}p_ia_i,M;f\right]-\left[m,\sum_{i=1}^{n}p_ia_i;f\right]\right)\sum_{i=1}^{n}p_i\left|a_i-\sum_{j=1}^{n}p_ja_j\right|\\&\le\frac{1}{2}\left(\left[\sum_{i=1}^{n}p_ia_i,M;f\right]-\left[m,\sum_{i=1}^{n}p_ia_i;f\right]\right)\left[\sum_{i=1}^{n}p_ia_i^2-\left(\sum_{j=1}^{n}p_ja_j\right)^2\right]^{1/2}.\end{aligned}$$

If the lateral derivatives $f'_+(m)$ and $f'_-(M)$ are finite, then we also have the inequalities

$$0 \le \sum_{i=1}^{n} p_i f(a_i) - f\left(\sum_{i=1}^{n} p_i a_i\right) \tag{10.2}$$
$$\le \frac{1}{2}\left(\left[\sum_{i=1}^{n} p_i a_i, M; f\right] - \left[m, \sum_{i=1}^{n} p_i a_i; f\right]\right) \sum_{i=1}^{n} p_i \left|a_i - \sum_{j=1}^{n} p_j a_j\right|$$
$$\le \frac{1}{2}\left[f'_-(M) - f'_+(m)\right] \sum_{i=1}^{n} p_i \left|a_i - \sum_{j=1}^{n} p_j a_j\right|$$
$$\le \frac{1}{2}\left[f'_-(M) - f'_+(m)\right]\left[\sum_{i=1}^{n} p_i a_i^2 - \left(\sum_{j=1}^{n} p_j a_j\right)^2\right]^{1/2}.$$

Proof. We recall that if $f : I \to \mathbb{R}$ is a continuous convex function on the interval of real numbers I and $\alpha \in I$ then the *divided difference function* $f_\alpha : I \setminus \{\alpha\} \to \mathbb{R}$,

$$f_\alpha(t) := [\alpha, t; f] := \frac{f(t) - f(\alpha)}{t - \alpha}$$

is monotonic nondecreasing on $I \setminus \{\alpha\}$.

For $\bar{a}_p := \sum_{j=1}^{n} p_j a_j \in (m, M)$, we consider now the sequence

$$f_{\bar{a}_p}(i) := \frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}.$$

We will show that $f_{\bar{a}_p}(i)$ and $h_i := a_i - \bar{a}_p, \in \{1, \dots, n\}$ are synchronous.

Let $i, j \in \{1, \dots, n\}$ with $a_i, a_j \ne \bar{a}_p$. Assume that $a_i \ge a_j$, then by the monotonicity of f_α we have

$$f_{\bar{a}_p}(i) = \frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p} \ge \frac{f(a_j) - f(\bar{a}_p)}{a_j - \bar{a}_p} = f_{\bar{a}_p}(j) \tag{10.3}$$

and

$$h_i \ge h_j \tag{10.4}$$

which shows that

$$\left[f_{\bar{a}_p}(i) - f_{\bar{a}_p}(j)\right](h_i - h_j) \ge 0. \tag{10.5}$$

If $a_i < a_j$, then the inequalities (10.3) and (10.4) reverse but the inequality (10.5) still holds true.

Utilising the continuity property of the modulus we have

$$\left|\left[\left|f_{\bar{a}_p}(i)\right| - \left|f_{\bar{a}_p}(j)\right|\right](h_i - h_j)\right| \le \left|\left[f_{\bar{a}_p}(i) - f_{\bar{a}_p}(j)\right](h_i - h_j)\right|$$
$$= \left[f_{\bar{a}_p}(i) - f_{\bar{a}_p}(j)\right](h_i - h_j)$$

for any $i, j \in \{1, \dots, n\}$.

Multiplying with $p_i, p_j \geq 0$ and summing over i and j from 1 to n we have

$$\left|\sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \left[\left|f_{\bar{a}_p}(i)\right| - \left|f_{\bar{a}_p}(j)\right|\right](h_i - h_j)\right| \leq \sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \left[f_{\bar{a}_p}(i) - f_{\bar{a}_p}(j)\right](h_i - h_j). \tag{10.6}$$

A simple calculation shows that

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \left[\left|f_{\bar{a}_p}(i)\right| - \left|f_{\bar{a}_p}(j)\right|\right](h_i - h_j) \\
&= \sum_{i=1}^{n} p_i \left|f_{\bar{a}_p}(i)\right| h_i - \sum_{i=1}^{n} p_i \left|f_{\bar{a}_p}(i)\right| \sum_{i=1}^{n} p_i h_i \\
&= \sum_{i=1}^{n} p_i \left|\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right|(a_i - \bar{a}_p) \\
&- \sum_{i=1}^{n} p_i \left|\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right| \sum_{i=1}^{n} p_i (a_i - \bar{a}_p) \\
&= \sum_{i=1}^{n} p_i \left|\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right|(a_i - \bar{a}_p) \\
&= \sum_{i=1}^{n} p_i \left|f(a_i) - f(\bar{a}_p)\right| sgn(a_i - \bar{a}_p)
\end{aligned} \tag{10.7}$$

and

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \left[f_{\bar{a}_p}(i) - f_{\bar{a}_p}(j)\right](h_i - h_j) \\
&= \sum_{i=1}^{n} p_i f_{\bar{a}_p}(i) h_i - \sum_{i=1}^{n} p_i f_{\bar{a}_p}(i) \sum_{i=1}^{n} p_i h_i \\
&= \sum_{i=1}^{n} p_i \left(\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right)(a_i - \bar{a}_p) \\
&- \sum_{i=1}^{n} p_i \left(\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right) \sum_{i=1}^{n} p_i (a_i - \bar{a}_p) \\
&= \sum_{i=1}^{n} p_i \left(\frac{f(a_i) - f(\bar{a}_p)}{a_i - \bar{a}_p}\right)(a_i - \bar{a}_p) \\
&= \sum_{i=1}^{n} p_i f(a_i) - f\left(\sum_{i=1}^{n} p_i a_i\right).
\end{aligned} \tag{10.8}$$

On making use of the identities (10.7) and (10.8) we obtain from (10.6) the first inequality in (10.1).

Now, since $\bar{a}_p := \sum_{j=1}^{n} p_j a_j \in (m, M)$ then we have by the monotonicity of $f_{\bar{a}_p}(i)$ that

$$[m, \bar{a}_p; f] = \frac{f(\bar{a}_p) - f(m)}{\bar{a}_p - m} \leq f_{\bar{a}_p}(i) \leq \frac{f(M) - f(\bar{a}_p)}{M - \bar{a}_p} = [\bar{a}_p, M; f] \tag{10.9}$$

for any $i \in \{1, \ldots, n\}$.

Applying now the *Grüss' type inequality* obtained by Cerone & Dragomir in [9]

$$\left| \sum_{i=1}^{n} w_i x_i y_i - \sum_{i=1}^{n} w_i x_i \sum_{i=1}^{n} w_i y_i \right| \leq \frac{1}{2} (\Gamma - \gamma) \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right|$$

provided

$$-\infty < \delta \leq y_i \leq \Delta < \infty \tag{10.10}$$

for $i = 1, \ldots, n$, we have that

$$\sum_{i=1}^{n} p_i f(a_i) - f\left(\sum_{i=1}^{n} p_i a_i \right) \leq \frac{1}{2} \left([\bar{a}_p, M; f] - [m, \bar{a}_p; f] \right) \sum_{i=1}^{n} p_i \left| a_i - \sum_{j=1}^{n} p_j a_j \right|,$$

which proves the second inequality in (10.1).

The last bound in (10.1) is obvious by Cauchy-Bunyakovsky-Schwarz discrete inequality.

If the lateral derivatives $f'_+(m)$ and $f'_-(M)$ are finite, then by the convexity of f we have the *gradient inequalities*

$$\frac{f(M) - f(\bar{a}_p)}{M - \bar{a}_p} \leq f'_-(M)$$

and

$$\frac{f(\bar{a}_p) - f(m)}{\bar{a}_p - m} \geq f'_+(m),$$

where $\bar{a}_p \in (m, M)$. These imply that

$$[\bar{a}_p, M; f] - [m, \bar{a}_p; f] \leq f'_-(M) - f'_+(m)$$

and the proof of the third inequality in (10.2) is concluded.

The rest is obvious. □

Remark 10. *Define the weighted arithmetic mean of the positive n-tuple $x = (x_1, ..., x_n)$ with the nonnegative weights $w = (w_1, ..., w_n)$ by*

$$A_n(w, x) := \frac{1}{W_n} \sum_{i=1}^{n} w_i x_i$$

where $W_n := \sum_{i=1}^{n} w_i > 0$ and the weighted geometric mean of the same n-tuple, by

$$G_n(w, x) := \left(\prod_{i=1}^{n} x_i^{w_i} \right)^{1/W_n}.$$

It is well know that the following arithmetic mean-geometric mean inequality holds

$$A_n(w,x) \geq G_n(w,x).$$

Applying the inequality (10.2) for the convex function $f(t) = -\ln t, t > 0$ *we have the following reverse of the arithmetic mean-geometric mean inequality*

$$\begin{aligned}(10.11)\qquad 1 &\leq \frac{A_n(w,x)}{G_n(w,x)} \\ &\leq \left[\frac{\left(\frac{A_n(w,x)}{m}\right)^{A_n(w,x)-m}}{\left(\frac{M}{A_n(w,x)}\right)^{M-A_n(w,x)}}\right]^{\frac{1}{2}A_n(w,|x-A_n(w,x)|)} \\ &\leq \exp\left[\frac{1}{2}\frac{M-m}{mM}A_n(w,|x-A_n(w,x)|)\right],\end{aligned}$$

provided that $0 < m \leq x_i \leq M < \infty$ *for* $i \in \{1,...,n\}$.

11. Applications for the Hölder Inequality

If $x_i, y_i \geq 0$ for $i \in \{1,...,n\}$, then the *Hölder inequality* holds true

$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n x_i^p\right)^{1/p}\left(\sum_{i=1}^n y_i^q\right)^{1/q},$$

where $p > 1, \frac{1}{p} + \frac{1}{q} = 1$.

Assume that $p > 1$. If $z_i \in \mathbb{R}$ for $i \in \{1,...,n\}$, satisfies the bounds

$$0 < m \leq z_i \leq M < \infty \text{ for } i \in \{1,...,n\}$$

and $w_i \geq 0 \;\; (i = 1,\dots,n)$ with $W_n := \sum_{i=1}^n w_i > 0$, then from Theorem 7 we have amongst other the following inequality

$$\begin{aligned}(11.1)\qquad &\left|\frac{1}{W_n}\sum_{i=1}^n \left|z_i^p - \left(\frac{\sum_{i=1}^n w_i z_i}{W_n}\right)^p\right| w_i sgn\left[z_i - \frac{\sum_{i=1}^n w_i z_i}{W_n}\right] d\mu\right| \\ &\leq \frac{\sum_{i=1}^n w_i z_i^p}{W_n} - \left(\frac{\sum_{i=1}^n w_i z_i}{W_n}\right)^p \\ &\leq \frac{1}{2}\left(\left[\frac{\sum_{i=1}^n w_i z_i}{W_n}, M; (\cdot)^p\right] - \left[m, \frac{\sum_{i=1}^n w_i z_i}{W_n}; (\cdot)^p\right]\right)\tilde{D}_w(z) \\ &\leq \frac{1}{2}\left(\left[\frac{\sum_{i=1}^n w_i z_i}{W_n}, M; (\cdot)^p\right] - \left[m, \frac{\sum_{i=1}^n w_i z_i}{W_n}; (\cdot)^p\right]\right)\tilde{D}_{w,2}(z) \\ &\leq \frac{1}{4}\left(\left[\frac{\sum_{i=1}^n w_i z_i}{W_n}, M; (\cdot)^p\right] - \left[m, \frac{\sum_{i=1}^n w_i z_i}{W_n}; (\cdot)^p\right]\right)(M-m),\end{aligned}$$

where $\frac{\sum_{i=1}^n w_i z_i}{W_n} \in (m, M)$ and

$$\tilde{D}_w(z) := \frac{1}{W_n}\sum_{i=1}^n w_i \left|z_i - \frac{\sum_{j=1}^n w_j z_j}{W_n}\right|$$

while

$$\tilde{D}_{w,2}(z) = \left[\frac{\sum_{i=1}^n w_i z_i^2}{W_n} - \left(\frac{\sum_{i=1}^n w_i z_i}{W_n}\right)^2\right]^{\frac{1}{2}}.$$

The following result related to the Hölder inequality holds:

Proposition 7 (Dragomir, 2011 [41]). *If $x_i \geq 0,\ y_i > 0$ for $i \in \{1, ..., n\}$, $p > 1, \frac{1}{p} + \frac{1}{q} = 1$ and there exists the constants $\gamma, \Gamma > 0$ and such that*

$$\gamma \leq \frac{x_i}{y_i^{q-1}} \leq \Gamma \text{ for } i \in \{1, ..., n\},$$

then we have

$$\begin{aligned}
(11.2)\qquad & \left|\sum_{i=1}^{n}\left|\frac{x_i^p}{y_i^q} - \left(\frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} y_j^q}\right)^p\right| y_i^q \, sgn\left[\frac{x_i}{y_i^{q-1}} - \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} y_j^q}\right]\right| \\
& \leq \frac{\sum_{i=1}^{n} x_i^p}{\sum_{i=1}^{n} y_i^q} - \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}\right)^p \\
& \leq \frac{1}{2}\left(\left[\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}; (\cdot)^p\right]\right) \tilde{D}_{y^q}\left(\frac{x}{y^{q-1}}\right) \\
& \leq \frac{1}{2}\left(\left[\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}; (\cdot)^p\right]\right) \tilde{D}_{y^q,2}\left(\frac{x}{y^{q-1}}\right) \\
& \leq \frac{1}{4}\left(\left[\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^q}; (\cdot)^p\right]\right) (\Gamma - \gamma),
\end{aligned}$$

where

$$\tilde{D}_{y^q}\left(\frac{x}{y^{q-1}}\right) = \frac{1}{\sum_{i=1}^{n} y_i^q} \sum_{i=1}^{n} y_i^q \left|\frac{x_i}{y_i^{q-1}} - \frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} y_j^q}\right|$$

and

$$\tilde{D}_{y^q,2}\left(\frac{x}{y^{q-1}}\right) = \left[\frac{1}{\sum_{i=1}^{n} y_i^q} \sum_{i=1}^{n} \frac{x_i^2}{y_i^{q-2}} - \left(\frac{\sum_{j=1}^{n} x_j y_j}{\sum_{j=1}^{n} y_j^q}\right)^2\right]^{\frac{1}{2}}.$$

Proof. The inequalities (11.3) follow from (11.1) by choosing

$$z_i = \frac{x_i}{y_i^{q-1}} \text{ and } w_i = y_j^q.$$

The details are omitted. □

Remark 11. *We observe that for $p = q = 2$ we have from the first inequality in (11.2) the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality*

$$
\begin{aligned}
(11.3)\qquad & \left|\sum_{i=1}^{n}\left|\frac{x_i^2}{y_i^2}-\left(\frac{\sum_{j=1}^{n}x_jy_j}{\sum_{j=1}^{n}y_j^2}\right)^2\right| y_i^2 sgn\left(\frac{x_i}{y_i}-\frac{\sum_{j=1}^{n}x_jy_j}{\sum_{j=1}^{n}y_j^2}\right)\right| \\
& \le \frac{\sum_{i=1}^{n}x_i^2}{\sum_{i=1}^{n}y_i^2}-\left(\frac{\sum_{i=1}^{n}x_iy_i}{\sum_{i=1}^{n}y_i^2}\right)^2 \\
& \le \frac{1}{2}\left(\Gamma-\gamma\right)\frac{1}{\sum_{i=1}^{n}y_i^2}\sum_{i=1}^{n}y_i^2\left|\frac{x_i}{y_i}-\frac{\sum_{j=1}^{n}x_jy_j}{\sum_{j=1}^{n}y_j^2}\right| \\
& \le \frac{1}{2}\left(\Gamma-\gamma\right)\left[\frac{1}{\sum_{i=1}^{n}y_i^2}\sum_{i=1}^{n}x_i^2-\left(\frac{\sum_{j=1}^{n}x_jy_j}{\sum_{j=1}^{n}y_j^2}\right)^2\right]^{\frac{1}{2}} \\
& \le \frac{1}{4}\left(\Gamma-\gamma\right)^2,
\end{aligned}
$$

provided that there exists the constants $\gamma, \Gamma > 0$ such that

$$\gamma \le \frac{x_i}{y_i} \le \Gamma \text{ for } i \in \{1, ..., n\}.$$

12. Applications for f-Divergence

Consider the *f-divergence*

$$I_f(p,q) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) \tag{12.1}$$

defined on the set of probability distributions $p, q \in \mathbb{P}^n$, where f is convex on $(0,\infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u = 1$.

Proposition 8 (Dragomir, 2011 [41])**.** *Let $f : (0,\infty) \to \mathbb{R}$ be a convex function with the property that $f(1) = 0$. Assume that $p, q \in \mathbb{P}^n$ and there exists the constants $0 < r < 1 < R < \infty$ such that*

$$r \le \frac{p_i}{q_i} \le R \text{ for } i \in \{1, ..., n\}. \tag{12.2}$$

Then we have

$$
\begin{aligned}
(12.3)\qquad 0 \le I_f(p,q) & \le \frac{1}{2}\left([1,R;f]-[r,1;f]\right)D_v(p,q) \\
& \le \frac{1}{2}\left[f'_-(R)-f'_+(r)\right]D_v(p,q) \\
& \le \frac{1}{2}\left[f'_-(R)-f'_+(r)\right]\left[D_{\chi^2}(p,q)\right]^{1/2} \\
& \le \frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right],
\end{aligned}
$$

where $D_v(p,q) = \sum_{i=1}^{n}|p_i - q_i|$ and $D_{\chi^2}(p,q) = \sum_{i=1}^{n}\frac{p_i^2}{q_i} - 1$.

Proof. From (10.2) we have

$$\begin{aligned}
0 &\leq \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) - f(1) \\
&\leq \frac{1}{2}\left([1, R; f] - [r, 1; f]\right) \sum_{i=1}^{n} q_i \left|\frac{p_i}{q_i} - 1\right| \\
&\leq \frac{1}{2}\left[f'_-(R) - f'_+(r)\right] \sum_{i=1}^{n} q_i \left|\frac{p_i}{q_i} - 1\right| \\
&\leq \frac{1}{2}\left[f'_-(R) - f'_+(r)\right] \left(\sum_{i=1}^{n} \frac{p_i^2}{q_i} - 1\right)^{1/2} \leq \frac{1}{4}(R - r)\left[f'_-(R) - f'_+(r)\right]
\end{aligned}$$

i.e., the desired result (12.3). □

Remark 12. *The inequality*

$$I_f(p, q) \leq \frac{1}{4}(R - r)\left[f'_-(R) - f'_+(r)\right] \tag{12.4}$$

was obtained for the discrete divergence measures in 2000 by S.S. Dragomir, see [32]*.*

Proposition 9 (Dragomir, 2011 [41])**.** *With the assumptions in Proposition 8 we have*

$$\begin{aligned}
\left|I_{|f|(sgn(\cdot)-1)}(p, q)\right| &\leq I_f(p, q) \\
&\leq \frac{1}{2}\left([1, R; f] - [r, 1; f]\right) D_v(p, q) \\
&\leq \frac{1}{2}\left([1, R; f] - [r, 1; f]\right) \left[D_{\chi^2}(p, q)\right]^{1/2} \\
&\leq \frac{1}{4}\left([1, R; f] - [r, 1; f]\right)(R - r),
\end{aligned} \tag{12.5}$$

where $I_{|f|(sgn(\cdot)-1)}(p, q)$ is the generalized f-divergence for the non-necessarily convex function $|f|(sgn(\cdot) - 1)$ and is defined by

$$I_{|f|(sgn(\cdot)-1)}(p, q) := \sum_{i=1}^{n} q_i \left|f\left(\frac{p_i}{q_i}\right)\right| sgn\left(\frac{p_i}{q_i} - 1\right). \tag{12.6}$$

Proof. From the inequality (10.1) we have

$$\begin{aligned}
&\left|\sum_{i=1}^{n} q_i \left|f\left(\frac{p_i}{q_i}\right) - f(1)\right| sgn\left(\frac{p_i}{q_i} - 1\right)\right| \\
&\leq \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) - f(1) \\
&\leq \frac{1}{2}\left([1, R; f] - [r, 1; f]\right) \sum_{i=1}^{n} q_i \left|\frac{p_i}{q_i} - 1\right| \\
&\leq \frac{1}{2}\left([1, R; f] - [r, 1; f]\right) \left(\sum_{i=1}^{n} \frac{p_i^2}{q_i} - 1\right)^{1/2} \\
&\leq \frac{1}{4}\left([1, R; f] - [r, 1; f]\right)(R - r)
\end{aligned}$$

from where we get the desired result (12.5). □

The above results can be utilized to obtain various inequalities for the divergence measures in information theory that are particular instances of f-divergence.

Consider the *Kullback-Leibler divergence*

$$KL(p, q) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right), \ p, q \in \mathbb{P}^n.$$

For the convex function $f : (0, \infty) \to \mathbb{R}$, $f(t) = -\ln t$ we have

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

If $p, q \in \mathbb{P}^n$ such that there exists the constants $0 < r < 1 < R < \infty$ with

$$r \leq \frac{p_i}{q_i} \leq R \text{ for } i \in \{1, ..., n\}, \tag{12.7}$$

then we get from the first inequality in (12.3) that

$$0 \leq KL(q, p) \leq \frac{1}{2} D_v(p, q) \ln\left(\frac{1}{R^{R-1} r^{1-r}}\right).$$

For the convex function $f : (0, \infty) \to \mathbb{R}$, $f(t) = t \ln t$ we have

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

If $p, q \in \mathbb{P}^n$ are such that there exists the constants $0 < r < 1 < R < \infty$ with the property (12.7), then we get from the first inequality in (12.3) that

$$0 \leq KL(p, q) \leq \frac{1}{2} D_v(p, q) \ln\left(R^{\frac{R}{R-1}} r^{\frac{r}{1-r}}\right).$$

References

[1] S.M. Ali and S.D. Silvey, A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. Sec B,* **28** (1966), 131-142.

[2] N.S. Barnett, P. Cerone and S.S. Dragomir, Some new inequalities for Hermite-Hadamard difference in information theory, *Stochastic Analysis and Applications*, Y.J. Cho, J.K. Kim and Y.K. Choi (Eds.), Nova Science Publishers, 2003, pp. 7-20.

[3] N.S. Barnett, P. Cerone, S.S. Dragomir and A. Sofo, Approximating Csiszár $f-$divergence by the use of Taylor's formula with integral remainder, *Math. Ineq. & Appl.,* **5**(3) (2002), 417-434.

[4] N.S. Barnett, P. Cerone, S.S. Dragomri and A. Sofo, Approximating two mappings associated to Csiszár $f-$divergence via Taylor's expansion, *PanAmerican Math. J.,* **12**(4) (2002), 105-117.

[5] M. Beth Bassat, f-entropies, probability of error and feature selection, *Inform. Control*, **39** (1978), 227-242.

[6] R. Beran, Minimum Hellinger distance estimates for parametric models, *Ann. Statist.,* 5 (1977), 445-463.

[7] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions, *Bull. Calcutta Math. Soc.,* **35** (1943), 99-109.

[8] I. Burbea and C.R. Rao, On the convexity of some divergence measures based on entropy function, *IEEE Trans. Inf. Th.,* **28**(3) (1982), 489-495.

[9] P. Cerone and S. S. Dragomir, A refinement of the Grüss inequality and applications. *Tamkang J. Math.* **38** (2007), no. 1, 37–49. Preprint RGMIA *Res. Rep. Coll.* **5**(2002), No. 2, Art. 14. `[Online http://rgmia.org/v5n2.php]`.

[10] C.H. Chen, *Statistical Pattern Recognition,* Rocelle Park, New York, Hoyderc Book Co., 1973.

[11] C.K. Chow and C.N. Lin, Approximating discrete probability distributions with dependence trees, *IEEE Trans. Inf. Th.,* **14**(3) (1968), 462-467.

[12] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons Inc., 1991.

[13] I. Csiszár, Information measures: A critical survey, *Trans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Random Processes and 8th European Meeting of Statist.,* Volume B, Academia Prague, 1978, 73-86.

[14] I. Csiszár, Information-type measures of difference of probability functions and indirect observations, *Studia Sci. Math. Hungar,* **2** (1967), 299-318.

[15] I. Csiszár and J. Körner, *Information Theory: Coding Theorem for Discrete Memoryless Systems,* Academic Press, New York, 1981.

[16] D. Dacunha-Castelle, Ecole d'ete de Probabilités de Saint-Fleour, III-1997, Berlin, Heidelberg: Springer 1978.

[17] S.S. Dragomir, An improvement of Jensen's inequality, *Bull. Math. Soc. Sci. Math. Roumanie*, **34**(**82**) (1990), No. 4, 291-296.

[18] S.S. Dragomir, Some refinements of Ky Fan's inequality, *J. Math. Anal. Appl.*, **163**(2) (1992), 317-321.

[19] S.S. Dragomir, Some refinements of Jensen's inequality, *J. Math. Anal. Appl.*, **168**(2) (1992), 518-522.

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

[21] S.S. Dragomir, A new improvement of Jensen's inequality, *Indian J. Pure and Appl. Math.*, **26**(10) (1995), 959-968.

[22] S.S. Dragomir, Some inequalities for (m, M)-convex mappings and applications for the Csiszár f-divergence in Information Theory, *Math. J. Ibaraki Univ.,* 33 (2001), 35-50.

[23] S.S. Dragomir, Some inequalities for two Csiszár divergences and applications, *Mat. Bilten* (Macedonia), **25** (2001), 73-90.

[24] S.S. Dragomir, On a reverse of Jessen's inequality for isotonic linear functionals, *J. Ineq. Pure & Appl. Math.*, **2**(2001), No. 3, Article 36.

[25] S.S. Dragomir, An upper bound for the Csiszár f-divergence in terms of variational distance and applicaitons, *PanAmerican Math. J.,* **12**(4) (2002), 165-173.

[26] S.S. Dragomir, Upper and lower bounds for Csiszár f-divergence in terms of Hellinger discrimination and applications, *Nonlinear Anal. Forum* (Korea), **7**(1) (2002), 1-13.

[27] S.S. Dragomir, Bounds for f-divergences under likelihood ratio constraints, *Appl. of Math.* (Czech Republic), **48**(3) (2003), 205-223.

[28] S.S. Dragomir, A Grüss type inequality for isotonic linear functionals and applications. *Demonstratio Math.* **36** (2003), no. 3, 551–562. Preprint RGMIA *Res. Rep. Coll.* **5**(2002), Suplement, Art. 12. [Online http://rgmia.org/v5(E).php].

[29] S.S. Dragomir, Some inequalities for the Csiszár f-divergence, *J. KSIAM* (Korea), **7**(1) (2003), 63-77.

[30] S. S. Dragomir, New inequalities for Csiszár divergence and applications. *Acta Math. Vietnam.* **28** (2003), no. 2, 123–134.

[31] S.S. Dragomir, A converse inequality for the Csiszár f-divergence, *Tamsui Oxford J. Math. Sci.* (Taiwan), **20**(1) (2004), 35-53.

[32] S. S. Dragomir, A converse inequality for the Csiszár Φ-divergence. *Tamsui Oxf. J. Math. Sci.* **20** (2004), no. 1, 35–53. Preprint in S.S. Dragomir (Ed.): *Inequalities for Csiszár f-Divergence in Information Theory,* RGMIA Monographs, Victoria University, 2000, [http://rgmia.org/monographs/csiszar_list.html#chap1].

[33] S.S. Dragomir, Some inequalities for the Csiszár f-divergence when f is an $L-$Lipschitzian function and applications, *Italian J. Pure & Appl. Math.,* **15** (2004), 57-76.

[34] S. S. Dragomir, *Semi-inner Products and Applications*, Nova Science Publishers Inc., NY, 2004.

[35] S. S. Dragomir, *Discrete Inequalities of the Cauchy-Bunyakovsky-Schwarz Type*, Nova Science Publishers, NY, 2004.

[36] S.S. Dragomir, Bounds for the normalized Jensen functional, *Bull. Austral. Math. Soc.* **74**(3)(2006), 471-476.

[37] S.S. Dragomir, Bounds for the deviation of a function from the chord generated by its extremities. *Bull. Aust. Math. Soc.* **78** (2008), no. 2, 225–248.

[38] S.S. Dragomir, A refinement of Jensen's inequality with applications for f-divergence measures, *Taiwanese J. Math.* **14** (2010), no. 1, 153–164.

[39] S S. Dragomir, A new refinement of Jensen's inequality in linear spaces with applications. *Math. Comput. Modelling* **52** (2010), no. 9-10, 1497–1505.

[40] S. S. Dragomir, Inequalities in terms of the Gâteaux derivatives for convex functions on linear spaces with applications. *Bull. Aust. Math. Soc.* **83** (2011), no. 3, 500–517.

[41] S. S. Dragomir, A refinement and a divided difference reverse of Jensen's inequality with applications, Preprint *RGMIA Res. Rep. Coll.* **14** (2011), Art. 74 [Online http://rgmia.org/papers/v14/v14a74.pdf].

[42] S. S. Dragomir, Some Slater's type inequalities for convex functions defined on linear spaces and applications. *Abstr. Appl. Anal.* **2012**, Art. ID 168405, 16 pp.

[43] S. S. Dragomir, Some reverses of the Jensen inequality with applications. *Bull. Aust. Math. Soc.* **87** (2013), no. 2, 177–194.

[44] S. S. Dragomir, Reverses of the Jensen inequality in terms of first derivative and applications. *Acta Math. Vietnam.* **38** (2013), no. 3, 429–446

[45] S.S.Dragomir and N.M. Ionescu, Some converse of Jensen's inequality and applications. *Rev. Anal. Numér. Théor. Approx.* **23** (1994), no. 1, 71–78.

[46] 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.

[47] S.S. Dragomir and C.J. Goh, Some bounds on entropy measures in information theory, *Appl. Math. Lett.,* **10** (1997), 23-28.

[48] S.S. Dragomir and C.J. Goh, Some counterpart inequalities in for a functional associated with Jensen's inequality, *J. of Ineq. & Appl.,* **1** (1997), 311-325.

[49] S.S. Dragomir and C.J. Goh, A counterpart of Jensen's continuous inequality and applications in information theory, *Anal. St. Univ. "Al. I. Cuza", Iaşi,* **XLVII** (2001), 239-262.

[50] S.S. Dragomir, J. Pečarić and L.E. Persson, Properties of some functionals related to Jensen's inequality, *Acta Math. Hung.*, **70**(1-2) (1996), 129-143.

[51] S.S. Dragomir, M. Scholz and J. Šunde, Some upper bounds for relative entropy and applications, *Comp. and Math. with Appl.,* **39** (2000), 91-100..

[52] B.R. Frieden, Image enhancement and restoration, *Picture Processing and Digital Filtering* (T.S. Huang, Editor), Springer-Verlag, Berlin, 1975.

[53] J.H. Havrda and F. Charvat, Quantification method classification process: concept of structural α-entropy, *Kybernetika*, **3** (1967), 30-35.

[54] E. Hellinger, Neue Bergrüirdung du Theorie quadratisher Formerus von uneudlichvieleu Veränderlicher, *J. für reine and Augeur. Math.,* **36** (1909), 210-271.

[55] R.G. Gallager, *Information Theory and Reliable Communications,* J. Wiley, New York, 1968.

[56] D.V. Gokhale and S. Kullback, *Information in Contingency Tables*, New York, Merul Dekker, 1978.

[57] H. Jeffreys, An invariant form for the prior probability in estimation problems, *Proc. Roy. Soc. London, Ser. A*, **186** (1946), 453-461.

[58] J.H. Jistice (editor), *Maximum Entropy and Bayssian Methods in Applied Statistics,* Cambridge University Press, Cambridge, 1986.

[59] T.T. Kadota and L.A. Shepp, On the best finite set of linear observables for discriminating two Gaussian signals, *IEEE Trans. Inf. Th.,* **13** (1967), 288-294.

[60] T. Kailath, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Technology.,* Vol COM-15 (1967), 52-60.

[61] J.N. Kapur, A comparative assessment of various measures of directed divergence, *Advances in Management Studies,* **3** (1984), No. 1, 1-16.

[62] J.N. Kapur, On the roles of maximum entropy and minimum discrimination information principles in Statistics, *Technical Address of the 38th Annual Conference of the Indian Society of Agricultural Statistics,* 1984, 1-44.

[63] S. Kullback, *Information Theory and Statistics,* J. Wiley, New York, 1959.

[64] S. Kullback and R.A. Leibler, On information and sufficiency, *Annals Math. Statist.,* 22 (1951), 79-86.

[65] D. Kazakos and T. Cotsidas, A decision theory approach to the approximation of discrete probability densities, *IEEE Trans. Perform. Anal. Machine Intell.,* **1** (1980), 61-67.

[66] J.H.B. Kemperman, On the optimum note of transmitting information, *Ann. Math. Statist.,* **40** (1969), 2158-2177.

[67] C. Kraft, Some conditions for consistency and uniform consistency of statistical procedures, *Univ. of California Pub. in Statistics*, **1** (1955), 125-142.

[68] L. Lecam, *Asymptotic Methods in Statistical Decision Theory,* New York: Springer, 1986.

[69] R.M. Leahy and C.E. Goutis, An optimal technique for constraint-based image restoration and mensuration, *IEEE Trans. on Acoustics, Speech and Signal Processing,* **34** (1986), 1692-1642.

[70] F. Liese and I. Vajda, *Convex Statistical Distances,* Teubner Verlag, Leipzig, 1987.

[71] J. Lin, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.,* **37** (1) (1991), 145-151.

[72] J. Lin and S.K.M. Wong, A new directed divergence measure and its characterization, *Int. J. General Systems,* **17** (1990), 73-81.

[73] M. Matić, C.E.M. Pearce and J. Pečarić, Improvements of some bounds on entropy measures in information theory, *Math. Inequal. and Appl.*, 1 (1998), 295–304.

[74] M. Matić, J. Pečarić and N. Ujević, On new estimation of the remainder in generalised Taylor's formula, *Math. Ineq. and Appl.*, **2** (3) (1999), 343-361.

[75] H.P. Mckean, JR., Speed of approach to equilibrium for Koc's caricature of a Maximilian gas, *Arch. Ration. Mech. Anal.,* **21** (1966), 343-367.

[76] M. Mei, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.

[77] E.C. Pielou, *Ecological Diversity,* Wiley, New York, 1975.

[78] M.S. Pinsker, Information and Information Stability of Random variables and processes, (in Russian), Moscow: Izv. Akad. Nouk, 1960.

[79] C.R. Rao, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.

[80] A. Rényi, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.

[81] A. Sen, *On Economic Inequality,* Oxford University Press, London 1973.

[82] B.D. Sharma and D.P. MittalL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.,* **2** (4)(1977), 122-132.

[83] S. Simić, On a global upper bound for Jensen's inequality, *J. Math. Anal. Appl.* **343**(2008), 414-419.
[84] H. Shioya and T. Da-Te, A generalisation of Lin divergence and the derivative of a new information divergence, *Elec. and Comm. in Japan*, **78** (7) (1995), 37-40.
[85] I.J. Taneja, *Generalised Information Measures and their Applications* (http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html).
[86] H. Theil, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.
[87] H. Theil, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.
[88] G.T. Toussaint, Sharper lower bounds for discrimination in terms of variation, *IEEE Trans. Inf. Th.,* **21** (1975), 99-100.
[89] F. Topsoe, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll. RGMIA*, **2** (1) (1999), 85-98.
[90] I. Vajda, *Theory of Statistical Inference and Information,* Kluwer, Boston, 1989.
[91] I. Vajda, Note on discrimination information and variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.
[92] V.A. Volkonski and J.A. Rozanov, Some limit theorems for random function -I, (English Trans.), *Theory Prob. Appl., (*USSR), **4** (1959), 178-197.

[1]Mathematics, School of Engineering & Science, Victoria University, PO Box 14428, Melbourne City, MC 8001, Australia.

E-mail address: sever.dragomir@vu.edu.au

URL: http://rgmia.org/dragomir

[2]School of Computational & Applied Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa