Revista Colombiana de Matemáticas
Volumen 50(2016)1, páginas 17-39

A Refinement and a divided difference reverse of Jensen's inequality with applications

Un refinamiento y diferencias divididas, la desigualdad inversa de Jensen's con aplicaciones

S. Sever Dragomir[1]

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

Abstract. A refinement and a new sharp reverse of Jensen's inequality for convex functions in terms of divided diferences is obtained. Applications for means, the Hölder inequality and for f-divergence measures in information theory are also provided.

Key words and phrases. Jensen's inequality, Measurable functions, Lebesgue integral, Divergence measures, f-Divergence, Hölder inequality.

2010 Mathematics Subject Classification. Primary 26D15, 26D20; Secondary 94A05..

Resumen. Se optimiza la desigualdad inversa de Jensen para funciones convexas en términos de diferencias divididas vía un refinamiento. Se proveen aplicaciones de la desigualdad de Hölder para medias y para medidas f-divergentes en teoría de la información.

Palabras y frases clave. desigualdad de Jensen, funciones medibles, integral de Lebesgue, medidas de divergencia, f-divergente, desigualdad de Hölder.

1. Introduction

Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space consisting of a set Ω, a $\sigma-$ algebra $\mathcal{A}$ of parts of Ω and a countably additive and positive measure μ on $\mathcal{A}$ with values in $\mathbb{R} \cup \{\infty\}$.

For a $\mu-$measurable function $w : \Omega \to \mathbb{R}$, with $w(x) \geq 0$ for μ − a.e. (almost every) $x \in \Omega$, consider the Lebesgue space $L_w(\Omega, \mu) := \{f : \Omega \to \mathbb{R},\ f$

is $\mu-$measurable and $\int_\Omega w(x)|f(x)|\,d\mu(x) < \infty\}$. For simplicity of notation we write everywhere in the sequel $\int_\Omega wd\mu$ instead of $\int_\Omega w(x)\,d\mu(x)$.

An useful result that is used to provide simpler upper bounds for the difference in Jensen's inequality is the Gruss' inequality.

We recall now some facts related to this famous result.

If $f, g : \Omega \to \mathbb{R}$ are $\mu-$measurable functions and $f, g, fg \in L_w(\Omega, \mu)$, then we may consider the *Čebyšev functional*

$$T_w(f,g) := \int_\Omega wfgd\mu - \int_\Omega wfd\mu \int_\Omega wgd\mu. \tag{1}$$

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

$$|T_w(f,g)| \le \frac{1}{4}(\Gamma - \gamma)(\Delta - \delta), \tag{2}$$

provided

$$-\infty < \gamma \le f(x) \le \Gamma < \infty, \quad -\infty < \delta \le g(x) \le \Delta < \infty \tag{3}$$

for μ − a.e. a. $x \in \Omega$.

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

Note that if $\Omega = \{1, \ldots, n\}$ and μ is the discrete measure on Ω, then we obtain the discrete Grüss inequality

$$\left|\sum_{i=1}^n w_i x_i y_i - \sum_{i=1}^n w_i x_i \cdot \sum_{i=1}^n w_i y_i\right| \le \frac{1}{4}(\Gamma - \gamma)(\Delta - \delta), \tag{4}$$

provided $\gamma \le x_i \le \Gamma$, $\delta \le y_i \le \Delta$ for each $i \in \{1, \ldots, n\}$ and $w_i \ge 0$ with $W_n := \sum_{i=1}^n w_i = 1$.

With the above assumptions, if $f \in L_w(\Omega, \mu)$ then we may define

$$D_w(f) := D_{w,1}(f) := \int_\Omega w\left|f - \int_\Omega wfd\mu\right| d\mu. \tag{5}$$

In 2002, Cerone & Dragomir [5] have obtained the following refinement of the Grüss inequality (2):

Theorem 1.1. *Let $w, f, g : \Omega \to \mathbb{R}$ be $\mu-$measurable functions with $w \ge 0$ $\mu-$ a.e. (almost everywhere) on Ω and $\int_\Omega wd\mu = 1$. If $f, g, fg \in L_w(\Omega, \mu)$ and there exists the constants δ, Δ such that*

$$-\infty < \delta \le g(x) \le \Delta < \infty \quad \textit{for} \quad \mu - \textit{a.e. } x \in \Omega, \tag{6}$$

then we have the inequality

$$|T_w(f,g)| \le \frac{1}{2}(\Delta-\delta) D_w(f). \tag{7}$$

The constant $\frac{1}{2}$ is sharp in the sense that it cannot be replaced by a smaller quantity.

***Remark* 1.2.** The inequality (7) was obtained for the particular case $\Omega=[a,b]$ and the uniform weight $w(t)=1, t\in[a,b]$ by X. L. Cheng and J. Sun in [7]. However, in that paper the authors did not prove the sharpness of the constant $\frac{1}{2}$.

For $f \in L_{p,w}(\Omega,\mathcal{A},\mu) := \left\{f:\Omega\to\mathbb{R}, \int_\Omega w|f|^p d\mu<\infty\right\}$, $p\ge 1$ we may also define

$$D_{w,p}(f) := \left[\int_\Omega w\left|f-\int_\Omega wfd\mu\right|^p d\mu\right]^{\frac{1}{p}} = \left\|f-\int_\Omega wfd\mu\right\|_{\Omega,p} \tag{8}$$

where $\|\cdot\|_{\Omega,p}$ is the usual p-norm on $L_{p,w}(\Omega,\mathcal{A},\mu)$, namely,

$$\|h\|_{\Omega,p} := \left(\int_\Omega w|h|^p d\mu\right)^{\frac{1}{p}}, \quad p\ge 1.$$

Using Hölder's inequality we get

$$D_{w,1}(f) \le D_{w,p}(f) \text{ for } p\ge 1,\ f\in L_{p,w}(\Omega,\mathcal{A},\mu); \tag{9}$$

and, in particular for $p=2$

$$D_{w,1}(f) \le D_{w,2}(f) := \left[\int_\Omega wf^2d\mu - \left(\int_\Omega wfd\mu\right)^2\right]^{\frac{1}{2}}, \tag{10}$$

if $f\in L_{2,w}(\Omega,\mathcal{A},\mu)$.

For $f\in L_\infty(\Omega,\mathcal{A},\mu) := \left\{f:\Omega\to\mathbb{R},\ \|f\|_{\Omega,\infty} := ess\sup_{x\in\Omega}|f(x)|<\infty\right\}$ we also have

$$D_{w,p}(f) \le D_{w,\infty}(f) := \left\|f-\int_\Omega wfd\mu\right\|_{\Omega,\infty}. \tag{11}$$

The following corollary may be useful in practice.

Corollary 1.3. *With the assumptions of Theorem 1.1, we have*

$$\begin{aligned} |T_w(f,g)| &\le \frac{1}{2}(\Delta-\delta)D_w(f) \qquad (12)\\ &\le \frac{1}{2}(\Delta-\delta)D_{w,p}(f) \quad \text{if } f\in L_p(\Omega,\mathcal{A},\mu),\ 1<p<\infty;\\ &\le \frac{1}{2}(\Delta-\delta)D_{w,\infty}(f) \quad \text{if } f\in L_\infty(\Omega,\mathcal{A},\mu). \end{aligned}$$

***Remark* 1.4.** The inequalities in (12) are in order of increasing coarseness. If we assume that $-\infty < \gamma \leq f(x) \leq \Gamma < \infty$ for μ − a.e. $x \in \Omega$, then by the Grüss inequality for $g = f$ we have for $p = 2$

$$\left[\int_\Omega wf^2 d\mu - \left(\int_\Omega wf d\mu\right)^2\right]^{\frac{1}{2}} \leq \frac{1}{2}(\Gamma - \gamma). \tag{13}$$

By (12), we deduce the following sequence of inequalities

$$\begin{aligned} |T_w(f,g)| &\leq \frac{1}{2}(\Delta - \delta)\int_\Omega w\left|f - \int_\Omega wf d\mu\right| d\mu \\ &\leq \frac{1}{2}(\Delta - \delta)\left[\int_\Omega wf^2 d\mu - \left(\int_\Omega wf d\mu\right)^2\right]^{\frac{1}{2}} \\ &\leq \frac{1}{4}(\Delta - \delta)(\Gamma - \gamma) \end{aligned} \tag{14}$$

for $f, g : \Omega \to \mathbb{R}$, μ − measurable functions and so that $-\infty < \gamma \leq f(x) < \Gamma < \infty$, $-\infty < \delta \leq g(x) \leq \Delta < \infty$ for μ − a.e. $x \in \Omega$. Thus, the inequality (14) is a refinement of Grüss' inequality (2).

In order to provide a reverse of the celebrated Jensen's integral inequality for convex functions, S.S. Dragomir obtained in 2002 [14] the following result:

Theorem 1.5. *Let $\Phi : [m, M] \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on (m, M) and $f : \Omega \to [m, M]$ so that $\Phi \circ f$, f, $\Phi' \circ f$, $(\Phi' \circ f) f \in L_w(\Omega, \mu)$, where $w \geq 0$ μ-a.e. on Ω with $\int_\Omega w d\mu = 1$. Then we have the inequality:*

$$\begin{aligned} 0 &\leq \int_\Omega w(\Phi \circ f) d\mu - \Phi\left(\int_\Omega wf d\mu\right) \\ &\leq \int_\Omega w(\Phi' \circ f) f d\mu - \int_\Omega w(\Phi' \circ f) d\mu \int_\Omega wf d\mu \\ &\leq \frac{1}{2}[\Phi'(M) - \Phi'(m)]\int_\Omega w\left|f - \int_\Omega wf d\mu\right| d\mu. \end{aligned} \tag{15}$$

For a generalization of the first inequality when differentiability is not assumed and the derivative Φ' is replaced with a selection φ from the subdifferential $\partial\Phi$, see the paper [28] by C.P. Niculescu.

***Remark* 1.6.** If $\mu(\Omega) < \infty$ and $\Phi \circ f,\ f,\ \Phi' \circ f,\ (\Phi' \circ f) \cdot f \in L(\Omega, \mu)$, then we have the inequality:

$$\begin{aligned} 0 &\leq \frac{1}{\mu(\Omega)} \int_\Omega (\Phi \circ f)\, d\mu - \Phi\left(\frac{1}{\mu(\Omega)} \int_\Omega f d\mu\right) \\ &\leq \frac{1}{\mu(\Omega)} \int_\Omega (\Phi' \circ f) f d\mu - \frac{1}{\mu(\Omega)} \int_\Omega (\Phi' \circ f)\, d\mu \cdot \frac{1}{\mu(\Omega)} \int_\Omega f d\mu \\ &\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \frac{1}{\mu(\Omega)} \int_\Omega \left| f - \frac{1}{\mu(\Omega)} \int_\Omega f d\mu \right| d\mu. \end{aligned} \tag{16}$$

***Remark* 1.7.** On making use of (15) and (14), one can state the following string of reverse inequalities for the Jensen's difference

$$\begin{aligned} 0 &\leq \int_\Omega w(\Phi \circ f)\, d\mu - \Phi\left(\int_\Omega w f d\mu\right) \\ &\leq \int_\Omega w(\Phi' \circ f) f d\mu - \int_\Omega w(\Phi' \circ f)\, d\mu \int_\Omega w f d\mu \\ &\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \int_\Omega w \left| f - \int_\Omega w f d\mu \right| d\mu \\ &\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \left[\int_\Omega w f^2 d\mu - \left(\int_\Omega w f d\mu\right)^2\right]^{\frac{1}{2}} \\ &\leq \frac{1}{4}\left[\Phi'(M) - \Phi'(m)\right](M - m). \end{aligned} \tag{17}$$

We notice that the inequality between the first, second and last term from (17) was proved in the general case of positive linear functionals in 2001 by S.S. Dragomir in [13].

The discrete case is as follows.

Let $\bar{\mathbf{a}} = (a_1, \ldots, a_n)$, $\bar{\mathbf{b}} = (b_1, \ldots, b_n)$, $\bar{\mathbf{p}} = (p_1, \ldots, p_n)$ be $n-$tuples of real numbers with $p_i \geq 0$ $(i \in \{1, \ldots, n\})$ and $\sum_{i=1}^n p_i = 1$. If $b \leq b_i \leq B$, $\quad i \in \{1, \ldots, n\}$, then one has the inequality

$$\begin{aligned} \left|\sum_{i=1}^n p_i a_i b_i - \sum_{i=1}^n p_i a_i \cdot \sum_{i=1}^n p_i b_i\right| &\leq \frac{1}{2}(B - b) \sum_{i=1}^n p_i \left| a_i - \sum_{j=1}^n p_j a_j \right| \\ &\leq \frac{1}{2}(B - b)\left[\sum_{i=1}^n p_i \left| a_i - \sum_{j=1}^n p_j a_j \right|^p\right]^{\frac{1}{p}} \\ &\leq \frac{1}{2}(B - b) \max_{i=\overline{1,n}} \left| a_i - \sum_{j=1}^n p_j a_j \right|, \end{aligned} \tag{18}$$

where $1 < p < \infty$. The constant $\frac{1}{2}$ is sharp in the first inequality.

If more information about the vector $\bar{\mathbf{a}} = (a_1, \dots, a_n)$ is available, namely, if there exists the constants a and A such that $a \le a_i \le A$, $\quad i \in \{1, \dots, n\}$, then

$$\left| \sum_{i=1}^{n} p_i a_i b_i - \sum_{i=1}^{n} p_i a_i \cdot \sum_{i=1}^{n} p_i b_i \right| \le \frac{1}{2} (B - b) \sum_{i=1}^{n} p_i \left| a_i - \sum_{j=1}^{n} p_j a_j \right| \tag{19}$$
$$\le \frac{1}{2} (B - b) \left[\sum_{i=1}^{n} p_i \left| a_i - \sum_{j=1}^{n} p_j a_j \right|^2 \right]^{\frac{1}{2}}$$
$$\le \frac{1}{4} (B - b) (A - a),$$

with the constants $\frac{1}{2}$ and $\frac{1}{4}$ best possible.

Corollary 1.8. *Let $\Phi : [m, M] \to \mathbb{R}$ 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 reverse of Jensen's weighted discrete inequality:*

$$0 \le \sum_{i=1}^{n} w_i \Phi (x_i) - \Phi \left(\sum_{i=1}^{n} w_i x_i \right) \tag{20}$$
$$\le \sum_{i=1}^{n} w_i \Phi' (x_i) x_i - \sum_{i=1}^{n} w_i \Phi' (x_i) \sum_{i=1}^{n} w_i x_i$$
$$\le \frac{1}{2} [\Phi' (M) - \Phi' (m)] \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right|.$$

***Remark* 1.9.** We notice that the inequality between the first and second term in (20) was proved in 1994 by Dragomir & Ionescu, see [16].

***Remark* 1.10.** On utilizing (20) and (19) we can state the string of inequalities

$$\begin{aligned} 0 &\leq \sum_{i=1}^{n} w_i \Phi(x_i) - \Phi\left(\sum_{i=1}^{n} w_i x_i\right) \qquad (21) \\ &\leq \sum_{i=1}^{n} w_i \Phi'(x_i)\, x_i - \sum_{i=1}^{n} w_i \Phi'(x_i) \sum_{i=1}^{n} w_i x_i \\ &\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j x_j \right| \\ &\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \left[\sum_{i=1}^{n} w_i x_i^2 - \left(\sum_{i=1}^{n} w_i x_i\right)^2\right]^{1/2} \\ &\leq \frac{1}{4}\left[\Phi'(M) - \Phi'(m)\right](M - m). \end{aligned}$$

We notice that the inequality between the first, second and last term in (21) was proved in 1999 by S.S. Dragomir in [12].

Motivated by the above results, a refinement and a new sharp reverse of Jensen's integral inequality for convex functions in terms of divided differences is obtained. Applications for means, the Hölder inequality and for f-divergence measures in information theory are also provided.

2. A refinement and a new 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

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

In what follows, we assume that $w : \Omega \to \mathbb{R}$, with $w(x) \geq 0$ for $\mu -$ a.e. $x \in \Omega$, is a μ-measurable function with $\int_\Omega w d\mu = 1$.

Theorem 2.1. *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 $[m, M] \subset \mathring{I}$, $\mathring{I}$ the interior of I. If $f : \Omega \to \mathbb{R}$, is $\mu-$measurable, satisfying the bounds*

$$-\infty < m \leq f(x) \leq M < \infty \quad \text{for} \quad \mu - \text{a.e. } x \in \Omega \qquad (22)$$

and such that $f, \Phi \circ f \in L_w(\Omega, \mu)$, then by denoting

$$\overline{f}_{\Omega,w} := \int_\Omega w f d\mu \in [m, M]$$

and assuming that $\overline{f}_{\Omega,w} \neq m, M$*, we have*

$$\begin{aligned} &\left| \int_{\Omega} \left| \Phi\left(f\right) - \Phi\left(\overline{f}_{\Omega,w}\right) \right| sgn\left[f - \overline{f}_{\Omega,w}\right] wd\mu \right| \\ &\leq \int_{\Omega} \left(\Phi \circ f\right) wd\mu - \Phi\left(\overline{f}_{\Omega,w}\right) \\ &\leq \frac{1}{2}\left(\left[\overline{f}_{\Omega,w}, M; \Phi\right] - \left[m, \overline{f}_{\Omega,w}; \Phi\right]\right) D_{w}\left(f\right) \\ &\leq \frac{1}{2}\left(\left[\overline{f}_{\Omega,w}, M; \Phi\right] - \left[m, \overline{f}_{\Omega,w}; \Phi\right]\right) D_{w,2}\left(f\right) \\ &\leq \frac{1}{4}\left(\left[\overline{f}_{\Omega,w}, M; \Phi\right] - \left[m, \overline{f}_{\Omega,w}; \Phi\right]\right) \left(M - m\right). \end{aligned} \tag{23}$$

The constant $\frac{1}{2}$ *in the second inequality from (23) is best possible.*

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

$$\Phi_{\alpha}\left(t\right) := \left[\alpha, t; \Phi\right] := \frac{\Phi\left(t\right) - \Phi\left(\alpha\right)}{t - \alpha}$$

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

For f as considered in the statement of the theorem we can assume that that it is not constant μ –almost every where, since for that case the inequality (23) is trivially satisfied.

For $\overline{f}_{\Omega,w} \in (m, M)$, we consider now the function defined μ -almost everywhere on Ω by

$$\Phi_{\overline{f}_{\Omega,w}}\left(x\right) := \frac{\Phi\left(f\left(x\right)\right) - \Phi\left(\overline{f}_{\Omega,w}\right)}{f\left(x\right) - \overline{f}_{\Omega,w}}.$$

We will show that $\Phi_{\overline{f}_{\Omega,w}}$ and $h := f - \overline{f}_{\Omega,w}$ are synchronous μ-a.e. on Ω.

Let $x, y \in \Omega$ with $f\left(x\right), f\left(y\right) \neq \overline{f}_{\Omega,w}$. Assume that $f\left(x\right) \geq f\left(y\right)$, then

$$\begin{aligned} \Phi_{\overline{f}_{\Omega,w}}\left(x\right) &= \frac{\Phi\left(f\left(x\right)\right) - \Phi\left(\overline{f}_{\Omega,w}\right)}{f\left(x\right) - \overline{f}_{\Omega,w}} \\ &\geq \frac{\Phi\left(f\left(y\right)\right) - \Phi\left(\overline{f}_{\Omega,w}\right)}{f\left(y\right) - \overline{f}_{\Omega,w}} = \Phi_{\overline{f}_{\Omega,w}}\left(y\right) \end{aligned} \tag{24}$$

and

$$h\left(x\right) \geq h\left(y\right) \tag{25}$$

which shows that

$$\left[\Phi_{\overline{f}_{\Omega,w}}\left(x\right) - \Phi_{\overline{f}_{\Omega,w}}\left(y\right)\right]\left[h\left(x\right) - h\left(y\right)\right] \geq 0. \tag{26}$$

If $f(x) < f(y)$, then the inequalities (24) and (25) reverse but the inequality (26) still holds true.

This show that for $\mu-$a.e. $x, y \in \Omega$ we have (26) and the claim is proven as stated.

Utilising the continuity property of the modulus we have

$$\begin{aligned}
&\left|\left[\left|\Phi_{\overline{f}_{\Omega,w}}(x)\right| - \left|\Phi_{\overline{f}_{\Omega,w}}(y)\right|\right][h(x) - h(y)]\right| \\
&\leq \left|\left[\Phi_{\overline{f}_{\Omega,w}}(x) - \Phi_{\overline{f}_{\Omega,w}}(y)\right][h(x) - h(y)]\right| \\
&= \left[\Phi_{\overline{f}_{\Omega,w}}(x) - \Phi_{\overline{f}_{\Omega,w}}(y)\right][h(x) - h(y)]
\end{aligned}$$

for μ-a.e. $x, y \in \Omega$.

Multiplying with $w(x), w(y) \geq 0$ and integrating over $\mu(x)$ and $\mu(y)$ we have

$$\begin{aligned}
&\left|\int_\Omega \int_\Omega \left[\left|\Phi_{\overline{f}_{\Omega,w}}(x)\right| - \left|\Phi_{\overline{f}_{\Omega,w}}(y)\right|\right]\right. \\
&\times [h(x) - h(y)]\, w(x)\, w(y)\, d\mu(x)\, d\mu(y)| \\
&\leq \int_\Omega \int_\Omega \left[\Phi_{\overline{f}_{\Omega,w}}(x) - \Phi_{\overline{f}_{\Omega,w}}(y)\right] \\
&\times [h(x) - h(y)]\, w(x)\, w(y)\, d\mu(x)\, d\mu(y).
\end{aligned} \tag{27}$$

A simple calculation shows that

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \int_\Omega \left[\left|\Phi_{\overline{f}_{\Omega,w}}(x)\right| - \left|\Phi_{\overline{f}_{\Omega,w}}(y)\right|\right] \\
&\times [h(x) - h(y)]\, w(x)\, w(y)\, d\mu(x)\, d\mu(y) \\
&= \int_\Omega \left|\Phi_{\overline{f}_{\Omega,w}}(x)\right| h(x)\, w(x)\, d\mu(x) \\
&- \int_\Omega \left|\Phi_{\overline{f}_{\Omega,w}}(x)\right| w(x)\, d\mu(x) \int_\Omega w(x)\, h(x)\, d\mu(x) \\
&= \int_\Omega \left|\frac{\Phi(f(x)) - \Phi\left(\overline{f}_{\Omega,w}\right)}{f(x) - \overline{f}_{\Omega,w}}\right| \left[f(x) - \overline{f}_{\Omega,w}\right] w(x)\, d\mu(x) \\
&= \int_\Omega \left|\Phi(f(x)) - \Phi\left(\overline{f}_{\Omega,w}\right)\right| sgn\left[f(x) - \overline{f}_{\Omega,w}\right] w(x)\, d\mu(x)
\end{aligned} \tag{28}$$

and

$$\begin{aligned}&\frac{1}{2}\int_{\Omega}\int_{\Omega}\left[\Phi_{\overline{f}_{\Omega,w}}(x)-\Phi_{\overline{f}_{\Omega,w}}(y)\right] \\ &\times[h(x)-h(y)]\,w(x)\,w(y)\,d\mu(x)\,d\mu(y) \\ &=\int_{\Omega}\Phi_{\overline{f}_{\Omega,w}}(x)\,h(x)\,w(x)\,d\mu(x) \\ &-\int_{\Omega}\Phi_{\overline{f}_{\Omega,w}}(x)\,w(x)\,d\mu(x)\int_{\Omega}h(x)\,w(x)\,d\mu(x) \\ &=\int_{\Omega}\frac{\Phi(f(x))-\Phi\left(\overline{f}_{\Omega,w}\right)}{f(x)-\overline{f}_{\Omega,w}}\left[f(x)-\overline{f}_{\Omega,w}\right]w(x)\,d\mu(x) \\ &=\int_{\Omega}\left[\Phi(f(x))-\Phi\left(\overline{f}_{\Omega,w}\right)\right]w(x)\,d\mu(x) \\ &=\int_{\Omega}w(\Phi\circ f)\,d\mu-\Phi\left(\overline{f}_{\Omega,w}\right).\end{aligned} \tag{29}$$

On making use of the identities (28) and (29) we obtain from (27) the first inequality in (23).

Now, since f satisfies the condition (22) then we have that

$$\begin{aligned}\left[m,\overline{f}_{\Omega,w};\Phi\right]&=\frac{\Phi\left(\overline{f}_{\Omega,w}\right)-\Phi(m)}{\overline{f}_{\Omega,w}-m}\leq\Phi_{\overline{f}_{\Omega,w}}(x) \\ &\leq\frac{\Phi(M)-\Phi\left(\overline{f}_{\Omega,w}\right)}{M-\overline{f}_{\Omega,w}}=\left[\overline{f}_{\Omega,w},M;\Phi\right]\end{aligned} \tag{30}$$

for $\mu-$a.e. $x\in\Omega$.

Applying now the Grüss' type inequality (7) and taking into account the second part of the equality in (28) we have that

$$\begin{aligned}&\int_{\Omega}w(\Phi\circ f)\,d\mu-\Phi\left(\overline{f}_{\Omega,w}\right) \\ &\leq\frac{1}{2}\left(\left[\overline{f}_{\Omega,w},M;\Phi\right]-\left[m,\overline{f}_{\Omega,w};\Phi\right]\right)\int_{\Omega}w\left|f-\overline{f}_{\Omega,w}\right|d\mu\end{aligned}$$

which proves the second inequality in (23).

The other two bounds are obvious from the comments in the introduction.

It is obvious that from (23) we get the following reverse of the first Hermite-Hadamard inequality for the convex function $\Phi:[a,b]\to\mathbb{R}$

$$\begin{aligned}&\frac{1}{b-a}\int_a^b\Phi(t)\,dt-\Phi\left(\frac{a+b}{2}\right) \\ &\leq\frac{1}{2}\left(\left[\frac{a+b}{2},b;\Phi\right]-\left[a,\frac{a+b}{2};\Phi\right]\right)D_w(e)\end{aligned} \tag{31}$$

where $e(t) = t, t \in [a, b]$.

Since a simple calculation shows that

$$\frac{1}{2}\left(\left[\frac{a+b}{2}, b; \Phi\right] - \left[a, \frac{a+b}{2}; \Phi\right]\right) = \frac{2}{b-a}\left[\frac{\Phi(a) + \Phi(b)}{2} - \Phi\left(\frac{a+b}{2}\right)\right]$$

and

$$D_w(e) = \frac{1}{b-a}\int_a^b \left|t - \frac{a+b}{2}\right| dt = \frac{1}{4}(b-a),$$

and we get from (31) that

$$0 \le \frac{1}{b-a}\int_a^b \Phi(t)\,dt - \Phi\left(\frac{a+b}{2}\right) \le \frac{1}{2}\left[\frac{\Phi(a) + \Phi(b)}{2} - \Phi\left(\frac{a+b}{2}\right)\right]. \tag{32}$$

To prove the sharpness of the constant $\frac{1}{2}$ in the second inequality from (23) we need now only to show that the equality case in (32) is realized.

If we take, for instance $\Phi(t) = \left|t - \frac{a+b}{2}\right|$, $t \in [a, b]$, then we observe that Φ is convex and we get in both sides of (32) the same quantity $\frac{1}{4}(b-a)$. ☑

Corollary 2.2. *With the assumptions in Theorem 2.1 and if the lateral derivatives $\Phi'_+(m)$ and $\Phi'_-(M)$ are finite, then we have the inequalities*

$$\begin{aligned} 0 &\le \int_\Omega (\Phi \circ f)\, w d\mu - \Phi\left(\overline{f}_{\Omega,w}\right) \\ &\le \frac{1}{2}\left(\left[\overline{f}_{\Omega,w}, M; \Phi\right] - \left[m, \overline{f}_{\Omega,w}; \Phi\right]\right) D_w(f) \\ &\le \frac{1}{2}\left(\Phi'_-(M) - \Phi'_+(m)\right) D_w(f) \\ &\le \frac{1}{2}\left(\Phi'_-(M) - \Phi'_+(m)\right) D_{w,2}(f) \\ &\le \frac{1}{4}\left(\Phi'_-(M) - \Phi'_+(m)\right)(M - m). \end{aligned} \tag{33}$$

The constant $\frac{1}{2}$ in the second and third inequality from (33) is best possible.

Proof. We need to prove only the third inequality.

By the convexity of Φ we have the gradient inequalities

$$\frac{\Phi(M) - \Phi\left(\overline{f}_{\Omega,w}\right)}{M - \overline{f}_{\Omega,w}} \le \Phi'_-(M)$$

and

$$\frac{\Phi\left(\overline{f}_{\Omega,w}\right)-\Phi\left(m\right)}{\overline{f}_{\Omega,w}-m}\geq\Phi'_{+}\left(m\right).$$

These imply that

$$\left[\overline{f}_{\Omega,w},M;\Phi\right]-\left[m,\overline{f}_{\Omega,w};\Phi\right]\leq\Phi'_{-}\left(M\right)-\Phi'_{+}\left(m\right)$$

and the proof is concluded.

We observe that from (33) we get the following reverse of the Hermite-Hadamard inequality for the convex function $\Phi:[a,b]\to\mathbb{R}$ having finite lateral derivative $\Phi'_{+}\left(a\right)$ and $\Phi'_{-}\left(b\right)$

$$\begin{aligned}&\frac{1}{b-a}\int_{a}^{b}\Phi\left(t\right)dt-\Phi\left(\frac{a+b}{2}\right) \\ &\leq\frac{1}{2}\left[\frac{\Phi\left(a\right)+\Phi\left(b\right)}{2}-\Phi\left(\frac{a+b}{2}\right)\right]\leq\frac{1}{8}\left[\Phi'_{-}\left(b\right)-\Phi'_{+}\left(a\right)\right]\left(b-a\right).\end{aligned}\tag{34}$$

We observe that the convex function $\Phi\left(t\right)=\left|t-\frac{a+b}{2}\right|$ has finite lateral derivatives

$$\Phi'_{-}\left(b\right)=1\text{ and }\Phi'_{+}\left(a\right)=-1$$

and replacing this function in (34) we get in all terms the same quantity $\frac{1}{4}\left(b-a\right)$.

This proves that the constant $\frac{1}{2}$ in the second and third inequality from (33) is best possible. ☑

***Remark* 2.3.** 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 $[m,M]\subset\mathring{I}$, $\mathring{I}$ the interior of I. Let $\bar{\mathbf{a}}=(a_1,\ldots,a_n)$, $\bar{\mathbf{p}}=(p_1,\ldots,p_n)$ be $n-$tuples of real numbers with $p_i\geq0$ $(i\in\{1,\ldots,n\})$ and $\sum_{i=1}^{n}p_i=1$. If $m\leq a_i\leq M$, $i\in\{1,\ldots,n\}$, with $\sum_{i=1}^{n}p_ia_i\neq m,M$, then

$$\begin{aligned}&\left|\sum_{i=1}^{n}p_i\left|\Phi\left(a_i\right)-\Phi\left(\sum_{i=1}^{n}p_ia_i\right)\right|sgn\left(a_i-\sum_{j=1}^{n}p_ja_j\right)\right| \\ &\leq\sum_{i=1}^{n}p_i\Phi\left(a_i\right)-\Phi\left(\sum_{i=1}^{n}p_ia_i\right) \\ &\leq\frac{1}{2}\left(\left[\sum_{i=1}^{n}p_ia_i,M;\Phi\right]-\left[m,\sum_{i=1}^{n}p_ia_i;\Phi\right]\right)\sum_{i=1}^{n}p_i\left|a_i-\sum_{j=1}^{n}p_ja_j\right|.\end{aligned}\tag{35}$$

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

$$0 \le \sum_{i=1}^{n} p_i \Phi(a_i) - \Phi\left(\sum_{i=1}^{n} p_i a_i\right) \tag{36}$$
$$\le \frac{1}{2}\left(\left[\sum_{i=1}^{n} p_i a_i, M; \Phi\right] - \left[m, \sum_{i=1}^{n} p_i a_i; \Phi\right]\right) \sum_{i=1}^{n} p_i \left| a_i - \sum_{j=1}^{n} p_j a_j \right|$$
$$\le \frac{1}{2}\left(\Phi'_-(M) - \Phi'_+(m)\right) \sum_{i=1}^{n} p_i \left| a_i - \sum_{j=1}^{n} p_j a_j \right|.$$

***Remark* 2.4.** Define the weighted arithmetic mean of the positive n-tuple $x = (x_1, \ldots, x_n)$ with the nonnegative weights $w = (w_1, \ldots, w_n)$ 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

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

Applying the inequality (36) for the convex function $\Phi(t) = -\ln t$, $t > 0$ we have the following reverse of the arithmetic mean-geometric mean inequality

$$1 \le \frac{A_n(w, x)}{G_n(w, x)} \tag{37}$$
$$\le \left[\frac{\left(\frac{A_n(w,x)}{m}\right)^{A_n(w,x)-m}}{\left(\frac{M}{A_n(w,x)}\right)^{M-A_n(w,x)}}\right]^{\frac{1}{2}A_n(w,|x-A_n(w,x)|)}$$
$$\le \exp\left[\frac{1}{2}\frac{M-m}{mM} A_n(w, |x - A_n(w, x)|)\right],$$

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

3. Applications for the Hölder Inequality

It is well known that if $f \in L_p(\Omega, \mu)$, $p > 1$, where the Lebesgue space $L_p(\Omega, \mu)$ is defined by

$$L_p(\Omega, \mu) := \{f : \Omega \to \mathbb{R},\ f \text{ is } \mu\text{-measurable and } \int_\Omega |f(x)|^p \, d\mu(x) < \infty\}$$

and $g\in L_q(\Omega,\mu)$ with $\frac{1}{p}+\frac{1}{q}=1$ then $fg\in L(\Omega,\mu):=L_1(\Omega,\mu)$ and the *Hölder inequality* holds true

$$\int_\Omega |fg|\,d\mu \le \left(\int_\Omega |f|^p\,d\mu\right)^{1/p}\left(\int_\Omega |g|^p\,d\mu\right)^{1/q}.$$

Assume that $p>1$. If $h:\Omega\to\mathbb{R}$ is μ-measurable, satisfies the bounds

$$-\infty < m \le |h(x)| \le M < \infty \text{ for } \mu\text{-a.e. } x\in\Omega$$

and is such that $h,|h|^p\in L_w(\Omega,\mu)$, for a μ-measurable function $w:\Omega\to\mathbb{R}$, with $w(x)\ge 0$ for μ -a.e. $x\in\Omega$ and $\int_\Omega wd\mu>0$, then from (23) we have

$$\begin{aligned}
&\left|\int_\Omega \left||h|^p-\overline{|h|}_{\Omega,w}^p\right| sgn\left[|h|-\overline{|h|}_{\Omega,w}\right] wd\mu\right| \qquad (38)\\
&\le \frac{\int_\Omega |h|^p\,wd\mu}{\int_\Omega wd\mu}-\left(\frac{\int_\Omega |h|\,wd\mu}{\int_\Omega wd\mu}\right)^p\\
&\le \frac{1}{2}\left(\left[\overline{|h|}_{\Omega,w},M;(\cdot)^p\right]-\left[m,\overline{|h|}_{\Omega,w};(\cdot)^p\right]\right)\tilde{D}_w(|h|)\\
&\le \frac{1}{2}\left(\left[\overline{|h|}_{\Omega,w},M;(\cdot)^p\right]-\left[m,\overline{|h|}_{\Omega,w};(\cdot)^p\right]\right)\tilde{D}_{w,2}(|h|)\\
&\le \frac{1}{4}\left(\left[\overline{|h|}_{\Omega,w},M;(\cdot)^p\right]-\left[m,\overline{|h|}_{\Omega,w};(\cdot)^p\right]\right)(M-m),
\end{aligned}$$

where $\overline{|h|}_{\Omega,w}:=\frac{\int_\Omega |h|wd\mu}{\int_\Omega wd\mu}\in[m,M]$ and

$$\tilde{D}_w(|h|):=\frac{1}{\int_\Omega wd\mu}\int_\Omega w\left||h|-\frac{\int_\Omega |h|\,wd\mu}{\int_\Omega wd\mu}\right|d\mu$$

while

$$\tilde{D}_{w,2}(|h|)=\left[\frac{\int_\Omega w\,|h|^2\,d\mu}{\int_\Omega wd\mu}-\left(\frac{\int_\Omega |h|\,wd\mu}{\int_\Omega wd\mu}\right)^2\right]^{\frac{1}{2}}.$$

The following result related to the Hölder inequality holds:

Proposition 3.1. *If $f\in L_p(\Omega,\mu)$, $g\in L_q(\Omega,\mu)$ with $p>1$, $\frac{1}{p}+\frac{1}{q}=1$ and there exists the constants $\gamma,\Gamma>0$ and such that*

$$\gamma\le\frac{|f|}{|g|^{q-1}}\le\Gamma \ \mu\text{-}a.e \text{ on } \Omega,$$

then we have

$$\left|\int_\Omega \left|\frac{|f|^p}{|g|^q} - \left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}\right)^p\right| sgn\left[\frac{|f|}{|g|^{q-1}} - \frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}\right] |g|^q\,d\mu\right| \tag{39}$$
$$\leq \frac{\int_\Omega |f|^p\,d\mu}{\int_\Omega |g|^q\,d\mu} - \left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}\right)^p$$
$$\leq \frac{1}{2}\left(\left[\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}; (\cdot)^p\right]\right) \tilde{D}_{|g|^q}\left(\frac{|f|}{|g|^{q-1}}\right)$$
$$\leq \frac{1}{2}\left(\left[\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}; (\cdot)^p\right]\right) \tilde{D}_{|g|^q,2}\left(\frac{|f|}{|g|^{q-1}}\right)$$
$$\leq \frac{1}{4}\left(\left[\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}, \Gamma; (\cdot)^p\right] - \left[\gamma, \frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}; (\cdot)^p\right]\right) (\Gamma - \gamma),$$

where

$$\tilde{D}_{|g|^q}\left(\frac{|f|}{|g|^{q-1}}\right) = \frac{1}{\int_\Omega |g|^q\,d\mu}\int_\Omega |g|^q\left|\frac{|f|}{|g|^{q-1}} - \frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}\right| d\mu$$

and

$$\tilde{D}_{|g|^q,2}\left(\frac{|f|}{|g|^{q-1}}\right) = \left[\frac{1}{\int_\Omega |g|^q\,d\mu}\int_\Omega \frac{|f|^2}{|g|^{q-2}} d\mu - \left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q\,d\mu}\right)^2\right]^{\frac{1}{2}}.$$

Proof. The inequalities (40) follow from (38) by choosing

$$h = \frac{|f|}{|g|^{q-1}} \text{ and } w = |g|^q.$$

The details are omitted. ☑

***Remark* 3.2.** We observe that for $p=q=2$ we have from the first inequality in (39) the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality

$$\left|\int_\Omega \left|\frac{|f|^2}{|g|^2}-\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2\,d\mu}\right)^2\right| sgn\left[\frac{|f|}{|g|}-\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2\,d\mu}\right]|g|^2\,d\mu\right| \tag{40}$$
$$\le \frac{\int_\Omega |f|^2\,d\mu}{\int_\Omega |g|^2\,d\mu}-\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2\,d\mu}\right)^2$$
$$\le \frac{1}{2}(\Gamma-\gamma)\frac{1}{\int_\Omega |g|^2\,d\mu}\int_\Omega |g|^2\left|\frac{|f|}{|g|}-\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2\,d\mu}\right|d\mu$$
$$\le \frac{1}{2}(\Gamma-\gamma)\left[\frac{1}{\int_\Omega |g|^2\,d\mu}\int_\Omega |f|^2\,d\mu-\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2\,d\mu}\right)^2\right]^{\frac{1}{2}}$$
$$\le \frac{1}{4}(\Gamma-\gamma)^2,$$

provided that $f,g\in L_2(\Omega,\mu)$, and there exists the constants $\gamma,\Gamma>0$ such that

$$\gamma\le\frac{|f|}{|g|}\le\Gamma \ \mu\text{-a.e on } \Omega.$$

4. Applications for f-divergence

One of the important issues in many applications of probability theory is finding an appropriate measure of *distance* (or *difference* or *discrimination*) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [20], Kullback and Leibler [25], Rényi [31], Havrda and Charvat [18], Kapur [23], Sharma and Mittal [33], Burbea and Rao [4], Rao [30], Lin [26], Csiszár [9], Ali and Silvey [1], Vajda [39], Shioya and Da-te [34] and others (see for example [27] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [30], genetics [27], finance, economics, and political science [32], [36], [37], biology [29], the analysis of contingency tables [17], approximation of probability distributions [8], [24], signal processing [21], [22] and pattern recognition [2], [6]. A number of these measures of distance are specific cases of Csiszár f-divergence and so further exploration of this concept will have a flow on effect to other measures of distance and to areas in which they are applied.

Assume that a set Ω and the $\sigma-$finite measure μ are given. Consider the set of all probability densities on μ to be $\mathcal{P}:=\{p|p:\Omega\to\mathbb{R},\, p(x)\ge 0,\ \int_\Omega p(x)\,d\mu(x)=1\}$.

Csiszár $f-$divergence is defined as follows [10]

$$I_f(p,q) := \int_\Omega p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}, \tag{41}$$

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 Kullback-Leibler divergence [25] is well known among the information divergences. It is defined as:

$$D_{KL}(p,q) := \int_\Omega p(x) \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}, \tag{42}$$

where ln is to base e.

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are the: *variation distance* D_v, *Hellinger distance* D_H [19], χ^2-*divergence* D_{χ^2}, $\alpha-$*divergence* D_α, *Bhattacharyya distance* D_B [3], *Harmonic distance* D_{Ha}, *Jeffrey's distance* D_J [20], *triangular discrimination* D_Δ [38], etc.... They are defined as follows:

$$D_v(p,q) := \int_\Omega |p(x) - q(x)| \, d\mu(x), \quad p,q \in \mathcal{P}; \tag{43}$$

$$D_H(p,q) := \int_\Omega \left|\sqrt{p(x)} - \sqrt{q(x)}\right| d\mu(x), \quad p,q \in \mathcal{P}; \tag{44}$$

$$D_{\chi^2}(p,q) := \int_\Omega p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2 - 1\right] d\mu(x), \quad p,q \in \mathcal{P}; \tag{45}$$

$$D_\alpha(p,q) := \frac{4}{1-\alpha^2}\left[1 - \int_\Omega [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x)\right], \quad p,q \in \mathcal{P}; \tag{46}$$

$$D_B(p,q) := \int_\Omega \sqrt{p(x)\, q(x)} d\mu(x), \quad p,q \in \mathcal{P}; \tag{47}$$

$$D_{Ha}(p,q) := \int_\Omega \frac{2p(x)\, q(x)}{p(x)+q(x)} d\mu(x), \quad p,q \in \mathcal{P}; \tag{48}$$

$$D_J(p,q) := \int_\Omega [p(x) - q(x)] \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}; \tag{49}$$

$$D_\Delta(p,q) := \int_\Omega \frac{[p(x) - q(x)]^2}{p(x)+q(x)} d\mu(x), \quad p,q \in \mathcal{P}. \tag{50}$$

For other divergence measures, see the paper [23] by Kapur or the book on line [35] by Taneja.

Most of the above distances $(42) - (50)$, are particular instances of Csiszár $f-$divergence. There are also many others which are not in this class (see for example [35]). For the basic properties of Csiszár $f-$divergence see [10], [11] and [39].

Before we apply the results obtained in the previous section we observe that, by employing the inequalities from (17) we can state the following theorem:

Theorem 4.1. *Let $f:(0,\infty)\to\mathbb{R}$ be a convex function with the property that $f(1)=0$. Assume that $p,q\in\mathcal{P}$ and there exists the constants $0<r<1<R<\infty$ such that*

$$r\leq\frac{q(x)}{p(x)}\leq R \text{ for } \mu\text{-a.e. } x\in\Omega. \tag{51}$$

Then we have

$$\begin{aligned} 0 &\leq I_f(p,q)\leq\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]D_v(p,q) \\ &\leq\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]\left[D_{\chi^2}(p,q)\right]^{1/2} \\ &\leq\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right]. \end{aligned} \tag{52}$$

Proof. From (17) we have

$$\begin{aligned} &\int_\Omega p(x)f\left(\frac{q(x)}{p(x)}\right)d\mu(x)-f\left(\int_\Omega q(x)\,d\mu(x)\right) \\ &\leq\frac{1}{2}\left[f'_-(R)-f'_+(r)\right] \\ &\times\int_\Omega p(x)\left|\frac{q(x)}{p(x)}-\int_\Omega q(y)\,d\mu(y)\right|d\mu(x) \\ &\leq\frac{1}{2}\left[f'_-(R)-f'_+(r)\right] \\ &\times\left[\int_\Omega p(x)\left(\frac{q(x)}{p(x)}\right)^2d\mu-\left(\int_\Omega q(x)\,d\mu\right)^2\right]^{\frac{1}{2}} \\ &\leq\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right], \end{aligned} \tag{53}$$

and since

$$\int_\Omega p(x)\left|\frac{q(x)}{p(x)}-\int_\Omega q(y)\,d\mu(y)\right|d\mu(x)=D_v(p,q)$$

and

$$\int_\Omega p(x)\left(\frac{q(x)}{p(x)}\right)^2d\mu-\left(\int_\Omega q(x)\,d\mu\right)^2=D_{\chi^2}(p,q),$$

then we get from (53) the desired result (52). ☑

***Remark* 4.2.** The inequality

$$I_f(p,q) \le \frac{1}{4}(R-r)\left[f'_-(R) - f'_+(r)\right] \tag{54}$$

was obtained for the discrete divergence measures in 2000 by S.S. Dragomir, see [15].

Theorem 4.3. *With the assumptions in Theorem 4.1 we have*

$$\begin{aligned} \left|I_{|f|(sgn(\cdot)-1)}(p,q)\right| &\le I_f(p,q) \\ &\le \frac{1}{2}([1,R;f]-[r,1;f])\, D_v(p,q) \\ &\le \frac{1}{2}([1,R;f]-[r,1;f])\left[D_{\chi^2}(p,q)\right]^{1/2} \\ &\le \frac{1}{4}([1,R;f]-[r,1;f])(R-r), \end{aligned} \tag{55}$$

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*

$$I_{|f|(sgn(\cdot)-1)}(p,q) := \int_\Omega \left|f\left(\frac{q(x)}{p(x)}\right)\right| sgn\left[\frac{q(x)}{p(x)}-1\right] p(x)\, d\mu. \tag{56}$$

***Proof*.** From the inequality (23) we have

$$\begin{aligned} &\left|\int_\Omega \left|f\left(\frac{q(x)}{p(x)}\right)\right| sgn\left[\frac{q(x)}{p(x)}-1\right] p(x)\, d\mu.\right| \\ &\le \int_\Omega p(x) f\left(\frac{q(x)}{p(x)}\right) d\mu(x) - f\left(\int_\Omega q(x)\, d\mu(x)\right) \\ &\le \frac{1}{2}([1,R;f]-[r,1;f]) \\ &\times \int_\Omega p(x)\left|\frac{q(x)}{p(x)} - \int_\Omega q(y)\, d\mu(y)\right| d\mu(x) \\ &\le \frac{1}{2}([1,R;f]-[r,1;f]) \\ &\times \left[\int_\Omega p(x)\left(\frac{q(x)}{p(x)}\right)^2 d\mu - \left(\int_\Omega q(x)\, d\mu\right)^2\right]^{\frac{1}{2}} \\ &\le \frac{1}{4}([1,R;f]-[r,1;f])(R-r), \end{aligned} \tag{57}$$

from where we get the desired result (55). ☑

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*

$$D_{KL}(p,q) := \int_{\Omega} p(x) \ln \left[\frac{p(x)}{q(x)} \right] d\mu(x), \quad p,q \in \mathcal{P},$$

which is an f-divergence for the convex function $f:(0,\infty) \to \mathbb{R}$, $f(t) = -\ln t$.

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

$$r \le \frac{q(x)}{p(x)} \le R \text{ for } \mu\text{-a.e. } x \in \Omega, \tag{58}$$

then we get from (52) that

$$\begin{aligned} D_{KL}(p,q) &\le \frac{R-r}{2rR} D_v(p,q) \\ &\le \frac{R-r}{2rR} \left[D_{\chi^2}(p,q) \right]^{1/2} \le \frac{(R-r)^2}{4rR} \end{aligned} \tag{59}$$

and from (55) that

$$\begin{aligned} D_{KL}(p,q) &\le \frac{1}{2} D_v(p,q) \ln \left(\frac{1}{R^{R-1} r^{1-r}} \right) \\ &\le \frac{1}{2} \left[D_{\chi^2}(p,q) \right]^{1/2} \ln \left(\frac{1}{R^{R-1} r^{1-r}} \right) \\ &\le \frac{1}{4} (R-r) \ln \left(\frac{1}{R^{R-1} r^{1-r}} \right). \end{aligned} \tag{60}$$

Acknowledgements

The author would like to thank the anonymous referee for valuables comments that have been implemented in the final version of this paper.

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] M. BETH BASSAT, *f-entropies, probability of error and feature selection*, Inform. Control **39** (1978), 227–242.

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

[4] I. BURBEA and C. R. RAO, *On the convexity of some divergence measures based on entropy function*, IEEE Trans. Inf. Th. **28** (1982), no. 3, 489–495.

[5] 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].

[6] C. H. CHEN, *Statistical pattern recognition*, Rocelle Park, New York, Hoyderc Book Co., 1973.

[7] X. L. CHENG and J. SUN, *A note on the perturbed trapezoid inequality*, J. Ineq. Pure. & Appl. Math. **3** (2002), no. 2, Article 29.

[8] C. K. CHOW and C. N. LIN, *Approximating discrete probability distributions with dependence trees*, IEEE Trans. Inf. Th. **14** (1968), no. 3, 462–467.

[9] I. CSISZÁR, *Information-type measures of difference of probability distributions and indirect observations*, Studia Math. Hungarica **2** (1967), 299–318.

[10] ———, *On topological properties of $f-$divergences*, Studia Math. Hungarica **2** (1967), 329–339.

[11] I. CSISZÁR and J. KÖRNER, *Information theory: Coding theorem for discrete memoryless systems*, Academic Press, New York, 1981.

[12] S. S. DRAGOMIR, *A converse result for jensen's discrete inequality via Gruss' inequality and applications in information theory*, An. Univ. Oradea Fasc. Mat. **7** (1999/2000), 178–189.

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

[14] ———, *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].

[15] ———, *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].

[16] 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.

[17] D. V. GOKHALE and S. KULLBACK, *Information in contingency tables*, New York, Marcel Decker, 1978.

[18] J. H. HAVRDA and F. CHARVAT, *Quantification method classification process: concept of structural α-entropy*, Kybernetika **3** (1967), 30–35.

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

[20] H. JEFFREYS, *An invariant form for the prior probability in estimating problems*, Proc. Roy. Soc. London **186 A** (1946), 4563–461.

[21] 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.

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

[23] J. N. KAPUR, *A comparative assessment of various measures of directed divergence*, Advances in Management Studies **3** (1984), 1–16.

[24] 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.

[25] S. KULLBACK and R. A. LEIBLER, *On information and sufficiency*, Annals Math. Statist. **22** (1951), 79–86.

[26] J. LIN, *Divergence measures based on the shannon entropy*, IEEE Trans. Inf. Th. **37** (1991), no. 1, 145–151.

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

[28] C. P. NICULESCU, *An extension of Chebyshev's inequality and its connection with Jensen's inequality. (English summary)*, J. Inequal. Appl. **6** (2001), no. 4, 451–462.

[29] E. C. PIELOU, *Ecological diversity*, Wiley, New York, 1975.

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

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

[32] A. SEN, *On economic inequality*, Oxford University Press, London, 1973.

[33] B. D. SHARMA and D. P. MITTAL, *New non-additive measures of relative information*, Journ. Comb. Inf. Sys. Sci. **2** (1977), no. 4, 122–132.

[34] 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** (1995), no. 7, 37–40.

[35] I. J. TANEJA, *Generalised information measures and their applications*, (`http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html`).

[36] H. THEIL, *Economics and information theory*, North-Holland, Amsterdam, 1967.

[37] ______, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.

[38] F. TOPSOE, *Some inequalities for information divergence and related measures of discrimination*, Res. Rep. Coll., RGMIA **2** (1999), no. 1, 85–98.

[39] I. VAJDA, *Theory of statistical inference and information*, Dordrecht-Boston, Kluwer Academic Publishers, 1989.

(Recibido en octubre de 2014. Aceptado en enero de 2015)

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

e-mail: `sever.dragomir@vu.edu.au`