MATHEMATICAL SCIENCES AND APPLICATIONS E-NOTES
https://doi.org/10.36753/mathenot.1362706
12 (2) 43-59 (2024) - Research Article
ISSN: 2147-6268

Some New f-Divergence Measures and Their Basic Properties

Silvestru Sever Dragomir*

Abstract
In this paper, we introduce some new f-divergence measures that we call *t-asymmetric/symmetric divergence measure* and *integral divergence measure,* establish their joint convexity and provide some inequalities that connect these f-divergences to the classical one introduced by Csiszar in 1963. Applications for the *dichotomy class* of convex functions are provided as well.

Keywords: *f-divergence measures, Hellinger discrimination, HH f-divergence measures, Jeffrey's distance, Kullback-Leibler divergence, χ^2-divergence*

AMS Subject Classification (2020): *94A17; 26D1*

Corresponding author

1. Introduction

Let $(X, \mathcal{A})$ be a measurable space satisfying $|\mathcal{A}| > 2$ and μ be a σ-finite measure on $(X, \mathcal{A})$. Let $\mathcal{P}$ be the set of all probability measures on $(X, \mathcal{A})$ which are absolutely continuous with respect to μ. For $P, Q \in \mathcal{P}$, let $p = \frac{dP}{d\mu}$ and $q = \frac{dQ}{d\mu}$ denote the *Radon-Nikodym* derivatives of P and Q with respect to μ.
Two probability measures $P, Q \in \mathcal{P}$ are said to be *orthogonal* and we denote this by $Q \perp P$ if

$$P(\{q = 0\}) = Q(\{p = 0\}) = 1.$$

Let $f : [0, \infty) \to (-\infty, \infty]$ be a convex function that is continuous at 0, i.e., $f(0) = \lim_{u \downarrow 0} f(u)$.
In 1963, I. Csiszár [1] introduced the concept of f-divergence as follows.

Definition 1.1. Let $P, Q \in \mathcal{P}$. Then

$$I_f(Q, P) = \int_X p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x), \tag{1.1}$$

Received : *19-09-2023,* ***Accepted :*** *04-01-2024,* ***Available online :*** *21-01-2024*

*(**Cite as** "S. S. Dragomir, Some New f-Divergence Measures and Their Basic Properties, Math. Sci. Appl. E-Notes, 12(2) (2024), 43-59")*

is called the f-divergence of the probability distributions Q and P.

Remark 1.1. Observe that, the integrand in the formula (1.1) is undefined when $p(x) = 0$. The way to overcome this problem is to postulate for f as above that

$$0f\left[\frac{q(x)}{0}\right] = q(x)\lim_{u\downarrow 0}\left[uf\left(\frac{1}{u}\right)\right], \; x \in X. \tag{1.2}$$

We now give some examples of f-divergences that are well-known and often used in the literature (see also [2]).

1.1 The class of χ^α-divergences

The f-divergences of this class, which is generated by the function χ^α, $\alpha \in [1,\infty)$, defined by

$$\chi^\alpha(u) = |u-1|^\alpha, \quad u \in [0,\infty)$$

have the form

$$I_f(Q,P) = \int_X p\left|\frac{q}{p}-1\right|^\alpha d\mu = \int_X p^{1-\alpha}|q-p|^\alpha d\mu. \tag{1.3}$$

From this class only the parameter $\alpha = 1$ provides a distance in the topological sense, namely the *total variation distance* $V(Q,P) = \int_X |q-p|\,d\mu$. The most prominent special case of this class is, however, *Karl Pearson's* χ^2-*divergence*

$$\chi^2(Q,P) = \int_X \frac{q^2}{p}d\mu - 1$$

that is obtained for $\alpha = 2$.

1.2 Dichotomy class

From this class, generated by the function $f_\alpha : [0,\infty) \to \mathbb{R}$

$$f_\alpha(u) = \begin{cases} u - 1 - \ln u & \text{for } \alpha = 0; \\ \dfrac{1}{\alpha(1-\alpha)}\left[\alpha u + 1 - \alpha - u^\alpha\right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ 1 - u + u\ln u & \text{for } \alpha = 1; \end{cases}$$

only the parameter $\alpha = \frac{1}{2}$ $\left(f_{\frac{1}{2}}(u) = 2(\sqrt{u}-1)^2\right)$ provides a distance, namely, the *Hellinger distance*

$$H(Q,P) = \left[\int_X (\sqrt{q}-\sqrt{p})^2 d\mu\right]^{\frac{1}{2}}.$$

Another important divergence is the *Kullback-Leibler divergence* obtained for $\alpha = 1$,

$$KL(Q,P) = \int_X q\ln\left(\frac{q}{p}\right)d\mu.$$

1.3 Matsushita's divergences

The elements of this class, which is generated by the function φ_α, $\alpha \in (0,1]$ given by

$$\varphi_\alpha(u) := |1-u^\alpha|^{\frac{1}{\alpha}}, \quad u \in [0,\infty),$$

are prototypes of metric divergences, providing the distances $[I_{\varphi_\alpha}(Q,P)]^\alpha$.

1.4 Puri-Vincze divergences

This class is generated by the functions Φ_α, $\alpha \in [1, \infty)$ given by

$$\Phi_\alpha(u) := \frac{|1-u|^\alpha}{(u+1)^{\alpha-1}}, \quad u \in [0,\infty).$$

It has been shown in [3] that this class provides the distances $[I_{\Phi_\alpha}(Q,P)]^{\frac{1}{\alpha}}$.

1.5 Divergences of Arimoto-type

This class is generated by the functions

$$\Psi_\alpha(u) := \begin{cases} \dfrac{\alpha}{\alpha-1}\left[(1+u^\alpha)^{\frac{1}{\alpha}} - 2^{\frac{1}{\alpha}-1}(1+u)\right] & \text{for } \alpha \in (0,\infty)\setminus\{1\}; \\ (1+u)\ln 2 + u\ln u - (1+u)\ln(1+u) & \text{for } \alpha = 1; \\ \dfrac{1}{2}|1-u| & \text{for } \alpha = \infty. \end{cases}$$

It has been shown in [4] that this class provides the distances $[I_{\Psi_\alpha}(Q,P)]^{\min\left(\alpha,\frac{1}{\alpha}\right)}$ for $\alpha \in (0,\infty)$ and $\frac{1}{2}V(Q,P)$ for $\alpha = \infty$.

For f continuous convex on $[0,\infty)$ we obtain the $*$-*conjugate* function of f by

$$f^*(u) = uf\left(\frac{1}{u}\right), \quad u \in (0,\infty)$$

and

$$f^*(0) = \lim_{u\downarrow 0} f^*(u).$$

It is also known that if f is continuous convex on $[0,\infty)$ then so is f^*.

The following two theorems contain the most basic properties of f-divergences. For their proofs we refer the reader to Chapter 1 of [5] (see also [2]).

Theorem 1.1 (Uniqueness and Symmetry Theorem)**.** *Let f, f_1 be continuous convex on $[0,\infty)$. We have*

$$I_{f_1}(Q,P) = I_f(Q,P),$$

for all $P, Q \in \mathcal{P}$ if and only if there exists a constant $c \in \mathbb{R}$ such that

$$f_1(u) = f(u) + c(u-1),$$

for any $u \in [0,\infty)$.

Theorem 1.2 (Range of Values Theorem)**.** *Let $f : [0,\infty) \to \mathbb{R}$ be a continuous convex function on $[0,\infty)$.*

For any $P, Q \in \mathcal{P}$, we have the double inequality

$$f(1) \le I_f(Q,P) \le f(0) + f^*(0). \tag{1.4}$$

(i) If $P = Q$, then the equality holds in the first part of (1.4).

If f is strictly convex at 1, then the equality holds in the first part of (1.4) if and only if $P = Q$;

(ii) If $Q \perp P$, then the equality holds in the second part of (1.4).

If $f(0) + f^*(0) < \infty$, then equality holds in the second part of (1.4) if and only if $Q \perp P$.

The following result is a refinement of the second inequality in Theorem 1.2 (see [2, Theorem 3]).

Theorem 1.3. *Let f be a continuous convex function on $[0,\infty)$ with $f(1)=0$ (f is normalised) and $f(0)+f^{*}(0)<\infty$. Then*

$$0\leq I_{f}(Q,P)\leq\frac{1}{2}\left[f(0)+f^{*}(0)\right]V(Q,P) \tag{1.5}$$

for any $Q,P\in\mathcal{P}$.

For other inequalities for f-divergence see [6], [7]-[21].

2. Some basic properties

Let f be a continuous convex function on $[0,\infty)$ with $f(1)=0$ and $t\in[0,1]$. We define the *t-asymmetric divergence measure* $A_{f,t}$ by

$$A_{f,t}(Q,P,W):=\int_{X}f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]w(x)\,d\mu(x) \tag{2.1}$$

and the *t-symmetric divergence measure* $S_{f,t}$ by

$$S_{f,t}(Q,P,W):=\frac{1}{2}\left[A_{f,t}(Q,P,W)+A_{f,1-t}(Q,P,W)\right] \tag{2.2}$$

for any $Q,P,W\in\mathcal{P}$.

For $t=\frac{1}{2}$ we consider the *mid-point divergence measure* M_f by

$$\begin{aligned}M_{f}(Q,P,W)&:=\int_{X}f\left[\frac{q(x)+p(x)}{2w(x)}\right]w(x)\,d\mu(x)\\&=A_{f,1/2}(Q,P,W)=S_{f,1/2}(Q,P,W),\end{aligned}$$

for any $Q,P,W\in\mathcal{P}$.

We can also consider the *integral divergence measure*

$$\begin{aligned}A_{f}(Q,P,W)&:=\int_{0}^{1}A_{f,t}(Q,P,W)\,dt=\int_{0}^{1}S_{f,t}(Q,P,W)\\&=\int_{X}\left(\int_{0}^{1}f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]dt\right)w(x)\,d\mu(x).\end{aligned}$$

The following result contains some basic facts concerning the divergence measures above:

Theorem 2.1. *Let f be a continuous convex function on $[0,\infty)$ with $f(1)=0$. Then for all $Q,P,W\in\mathcal{P}$ and $t\in[0,1]$*

$$0\leq A_{f,t}(Q,P,W)\leq(1-t)\,I_{f}(Q,W)+tI_{f}(P,W) \tag{2.3}$$

and the mapping

$$\mathcal{P}\times\mathcal{P}\ni(Q,P)\mapsto A_{f,t}(Q,P,W)\in[0,\infty) \tag{2.4}$$

is convex as a function of two variables.

We have the inequalities

$$0\leq M_{f}(Q,P,W)\leq S_{f,t}(Q,P,W)\leq\frac{1}{2}\left[I_{f}(Q,W)+I_{f}(P,W)\right] \tag{2.5}$$

for all $Q,P,W\in\mathcal{P}$ and the mapping

$$\mathcal{P}\times\mathcal{P}\ni(Q,P)\mapsto S_{f,t}(Q,P,W)\in[0,\infty) \tag{2.6}$$

is convex as a function of two variables.

Proof. Let $t\in[0,1]$ and $Q,P,W\in\mathcal{P}$. We use Jensen's integral inequality to get

$$\begin{aligned}A_{f,t}(Q,P,W)&=\int_X f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]w(x)\,d\mu(x)\\&\geq f\left(\int_X\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]w(x)\,d\mu(x)\right)\\&=f\left(\int_X[(1-t)\,q(x)+tp(x)]\,d\mu(x)\right)\\&=f\left((1-t)\int_X q(x)\,d\mu(x)+t\int_X p(x)\,d\mu(x)\right)=f(1)=0.\end{aligned}$$

By the convexity of f we also have

$$\begin{aligned}A_{f,t}(Q,P,W)&=\int_X f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]w(x)\,d\mu(x)\\&\leq(1-t)\int_X f\left[\frac{q(x)}{w(x)}\right]w(x)\,d\mu(x)+t\int_X f\left[\frac{p(x)}{w(x)}\right]w(x)\,d\mu(x)\\&=(1-t)\,I_f(Q,W)+tI_f(P,W)\end{aligned}$$

for $t\in[0,1]$ and $Q,P,W\in\mathcal{P}$, and the inequality (2.3) is proved.

Let $\alpha,\beta\geq0$ and such that $\alpha+\beta=1$. If $(Q_1,P_1),(Q_2,P_2)\in\mathcal{P}\times\mathcal{P}$, then

$$\begin{aligned}&A_{f,t}(\alpha(Q_1,P_1,W)+\beta(Q_2,P_2,W))\\=\;&A_{f,t}((\alpha Q_1+\beta Q_2,\alpha P_1+\beta P_2,W))\\=\;&\int_X f\left[\frac{(1-t)(\alpha Q_1+\beta Q_2)+t(\alpha P_1+\beta P_2)}{w(x)}\right]w(x)\,d\mu(x)\\=\;&\int_X f\left[\frac{\alpha[(1-t)\,Q_1+tP_1]+\beta[(1-t)\,Q_2+tP_2]}{w(x)}\right]w(x)\,d\mu(x)\\\leq\;&\alpha\int_X f\left[\frac{(1-t)\,Q_1+tP_1}{w(x)}\right]w(x)\,d\mu(x)+\beta\int_X f\left[\frac{(1-t)\,Q_2+tP_2}{w(x)}\right]w(x)\,d\mu(x)\\=\;&\alpha A_{f,t}(Q_1,P_1,W)+\beta A_{f,t}(Q_2,P_2,W),\end{aligned}$$

which proves the joint convexity of the mapping defined in (2.4).

Using the convexity of f we have

$$f\left(\frac{1}{2}\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}+\frac{(1-t)\,p(x)+tq(x)}{w(x)}\right]\right)\leq\frac{1}{2}\left\{f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]+f\left[\frac{(1-t)\,p(x)+tq(x)}{w(x)}\right]\right\},$$

namely

$$f\left(\frac{q(x)+p(x)}{2w(x)}\right)\leq\frac{1}{2}\left\{f\left[\frac{(1-t)\,q(x)+tp(x)}{w(x)}\right]+f\left[\frac{(1-t)\,p(x)+tq(x)}{w(x)}\right]\right\},\tag{2.7}$$

for $x\in X$.

By multiplying (2.7) with $w(x)$ and integrating over $\mu(x)$ we get the second inequality inequality in (2.5).

We have, by (2.3) that

$$\begin{aligned}S_{f,t}(Q,P,W)&=\frac{1}{2}[A_{f,t}(Q,P,W)+A_{f,1-t}(Q,P,W)]\\&\leq\frac{1}{2}\left[(1-t)\,I_f(Q,W)+tI_f(P,W)+tI_f(Q,W)+(1-tI)_f(P,W)\right]\\&=\frac{1}{2}[I_f(Q,W)+I_f(P,W)],\end{aligned}$$

which proves the third inequality in (2.5).

The convexity of the mapping defined by (2.6) follows by the same property of the mapping defined by (2.4). □

Corollary 2.1. *Let f be a continuous convex function on $[0,\infty)$ with $f(1)=0$. Then for all Q, P, $W\in\mathcal{P}$ we have the inequalities*

$$0\le M_f(Q,P,W)\le A_f(Q,P,W)\le\frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]. \tag{2.8}$$

The mapping

$$\mathcal{P}\times\mathcal{P}\ni(Q,P)\mapsto A_f(Q,P,W)\in[0,\infty) \tag{2.9}$$

is convex as a function of two variables.

Proof. The inequality (2.8) follows by integrating over t in the inequality (2.5). Since the mapping

$$\mathcal{P}\times\mathcal{P}\ni(Q,P)\mapsto S_{f,t}(Q,P,W)\in[0,\infty)$$

is convex as a function of two variables for all $t\in[0,1]$, then it remains convex if one takes the integral over $t\in[0,1]$. □

The following reverses of the Hermite-Hadamard inequality hold:

Lemma 2.1 (Dragomir, 2002 [9] and [10])**.** *Let $h:[a,b]\to\mathbb{R}$ be a convex function on $[a,b]$. Then*

$$\begin{aligned}0&\le\frac{1}{8}\left[h_+\left(\frac{a+b}{2}\right)-h_-\left(\frac{a+b}{2}\right)\right](b-a)\\&\le\frac{h(a)+h(b)}{2}-\frac{1}{b-a}\int_a^b h(x)\,dx\\&\le\frac{1}{8}\left[h_-(b)-h_+(a)\right](b-a)\end{aligned} \tag{2.10}$$

and

$$\begin{aligned}0&\le\frac{1}{8}\left[h_+\left(\frac{a+b}{2}\right)-h_-\left(\frac{a+b}{2}\right)\right](b-a)\\&\le\frac{1}{b-a}\int_a^b h(x)\,dx-h\left(\frac{a+b}{2}\right)\\&\le\frac{1}{8}\left[h_-(b)-h_+(a)\right](b-a).\end{aligned} \tag{2.11}$$

The constant $\frac{1}{8}$ is best possible in all inequalities.

We have the reverse inequalities:

Theorem 2.2. *Let f be a differentiable convex function on $[0,\infty)$ with $f(1)=0$. Then for all Q, P, $W\in\mathcal{P}$ we have*

$$0\le A_f(Q,P,W)-M_f(Q,P,W)\le\frac{1}{8}\Delta_{f'}(Q,P,W) \tag{2.12}$$

and

$$0\le\frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]-A_f(Q,P,W)\le\frac{1}{8}\Delta_{f'}(Q,P,W) \tag{2.13}$$

where

$$\Delta_{f'}(Q,P,W):=\int_X\left[f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right](q(x)-p(x))\,d\mu(x). \tag{2.14}$$

Proof. Let Q, P, $W\in\mathcal{P}$. By the inequality (2.11) we have

$$\begin{aligned}0&\le\int_0^1 f\left[\frac{(1-t)q(x)+tp(x)}{w(x)}\right]dt-f\left(\frac{q(x)+p(x)}{2w(x)}\right)\\&\le\frac{1}{8}\left[f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right]\left(\frac{q(x)}{w(x)}-\frac{p(x)}{w(x)}\right).\end{aligned}$$

If we multiply this inequality by $w(x) \geq 0$ and integrate on X we get (2.12).

From (2.10) we also have

$$
\begin{aligned}
0 &\leq \frac{1}{2}\left[f\left(\frac{q(x)}{w(x)}\right)+f\left(\frac{p(x)}{w(x)}\right)\right]-\int_0^1 f\left[\frac{(1-t)q(x)+tp(x)}{w(x)}\right]dt \\
&\leq \frac{1}{8}\left[f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right]\left(\frac{q(x)}{w(x)}-\frac{p(x)}{w(x)}\right).
\end{aligned}
$$

If we multiply this inequality by $w(x) \geq 0$ and integrate on X we get (2.12). □

Corollary 2.2. *Let f be a differentiable convex function on $[0,\infty)$ with $f(1)=0$ and $Q, P, W \in \mathcal{P}$. If there exists $0<r<1<R<\infty$ such that the following condition holds*

$$
r \leq \frac{q(x)}{w(x)}, \frac{p(x)}{w(x)} \leq R \text{ for } \mu\text{-a.e. } x \in X, \tag{2.15}
$$

then

$$
0 \leq A_f(Q,P,W)-M_f(Q,P,W) \leq \frac{1}{8}\left[f'(R)-f'(r)\right]d_1(Q,P) \tag{2.16}
$$

and

$$
0 \leq \frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]-A_f(Q,P,W) \leq \frac{1}{8}\left[f'(R)-f'(r)\right]d_1(Q,P) \tag{2.17}
$$

where

$$
d_1(Q,P) := \int_X |q(x)-p(x)|\, d\mu(x).
$$

Proof. Since f' is increasing on $[r,R]$, then

$$
|f'(t)-f'(s)| \leq f'(R)-f'(r)
$$

for all $t, s \in [r,R]$.

Therefore

$$
\begin{aligned}
\Delta_{f'}(Q,P,W) &:= \int_X \left[f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right](q(x)-p(x))\, d\mu(x) \\
&\leq \int_X \left|f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right| |q(x)-p(x)|\, d\mu(x) \\
&\leq \left[f'(R)-f'(r)\right]\int_X |q(x)-p(x)|\, d\mu(x) \\
&= \left[f'(R)-f'(r)\right]d_1(Q,P),
\end{aligned}
$$

which proves the desired inequalities (2.16) and (2.17). □

Corollary 2.3. *Let f be a twice differentiable convex function on $[0,\infty)$ with $f(1)=0$ and $Q, P, W \in \mathcal{P}$. If there exists $0<r<1<R<\infty$ such that the condition (2.15) holds and*

$$
\|f''\|_{[r,R],\infty} := \sup_{t\in[r,R]} |f''(t)| < \infty, \tag{2.18}
$$

then

$$
0 \leq A_f(Q,P,W)-M_f(Q,P,W) \leq \frac{1}{8}\|f''\|_{[r,R],\infty}\, d_{\chi^2}(Q,P,W) \tag{2.19}
$$

and

$$
0 \leq \frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]-A_f(Q,P,W) \leq \frac{1}{8}\|f''\|_{[r,R],\infty}\, d_{\chi^2}(Q,P,W), \tag{2.20}
$$

where

$$
d_{\chi^2}(Q,P,W) := \int_X \frac{(q(x)-p(x))^2}{w(x)}\, d\mu(x). \tag{2.21}
$$

Proof. We have

$$\begin{aligned}\Delta_{f'}(Q,P,W) &:= \int_X \left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{p(x)}{w(x)}\right)\right](q(x)-p(x))\,d\mu(x)\\ &\le \int_X \left|f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{p(x)}{w(x)}\right)\right| |q(x)-p(x)|\,d\mu(x)\\ &\le \|f''\|_{[r,R],\infty}\int_X \left|\frac{q(x)}{w(x)} - \frac{p(x)}{w(x)}\right| |q(x)-p(x)|\,d\mu(x)\\ &= \|f''\|_{[r,R],\infty}\int_X \frac{(q(x)-p(x))^2}{w(x)}d\mu(x),\end{aligned}$$

which proves the desired results (2.19) and (2.20). □

3. Further results

We have the following result for convex functions that is of interest in itself as well:

Lemma 3.1. *Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be a convex function on the interval I, $a, b \in \mathring{I}$, the interior of I, with $a < b$ and $\nu \in [0,1]$. Then*

$$\begin{aligned}&\nu(1-\nu)(b-a)\left[f'_+((1-\nu)a+\nu b) - f'_-((1-\nu)a+\nu b)\right]\\ &\le (1-\nu)f(a) + \nu f(b) - f((1-\nu)a+\nu b)\\ &\le \nu(1-\nu)(b-a)\left[f'_-(b) - f'_+(a)\right].\end{aligned} \tag{3.1}$$

In particular, we have

$$\begin{aligned}\frac{1}{4}(b-a)\left[f'_+\left(\frac{a+b}{2}\right) - f'_-\left(\frac{a+b}{2}\right)\right] &\le \frac{f(a)+f(b)}{2} - f\left(\frac{a+b}{2}\right)\\ &\le \frac{1}{4}(b-a)\left[f'_-(b) - f'_+(a)\right].\end{aligned} \tag{3.2}$$

The constant $\frac{1}{4}$ is best possible in both inequalities from (3.2).

Proof. The case $\nu = 0$ or $\nu = 1$ reduces to equality in (3.1).

Since f is convex on I it follows that the function is differentiable on $\mathring{I}$ except a countably number of points, the lateral derivatives $f'_\pm$ exists in each point of $\mathring{I}$, they are increasing on $\mathring{I}$ and $f'_- \le f'_+$ on $\mathring{I}$.

For any $x, y \in \mathring{I}$ we have for the Lebesgue integral

$$f(x) = f(y) + \int_y^x f'(s)\,ds = f(y) + (x-y)\int_0^1 f'((1-t)y+tx)\,dt. \tag{3.3}$$

Assume that $a < b$ and $\nu \in (0,1)$. By (3.3) we have

$$\begin{aligned}&f((1-\nu)a+\nu b)\\ &= f(a) + \nu(b-a)\int_0^1 f'((1-t)a + t((1-\nu)a+\nu b))\,dt\end{aligned} \tag{3.4}$$

and

$$\begin{aligned}&f((1-\nu)a+\nu b)\\ &= f(b) - (1-\nu)(b-a)\int_0^1 f'((1-t)b + t((1-\nu)a+\nu b))\,dt.\end{aligned} \tag{3.5}$$

If we multiply (3.4) by $1-\nu$, (3.4) by ν and add the obtained equalities, then we get

$$\begin{aligned} f\left((1-\nu)a+\nu b\right) &= (1-\nu) f(a) + \nu f(b) \\ &\quad + (1-\nu)\nu(b-a)\int_0^1 f'\left((1-t)a + t\left((1-\nu)a+\nu b\right)\right)dt \\ &\quad - (1-\nu)\nu(b-a)\int_0^1 f'\left((1-t)b + t\left((1-\nu)a+\nu b\right)\right)dt, \end{aligned}$$

which is equivalent to

$$(1-\nu) f(a) + \nu f(b) - f\left((1-\nu)a+\nu b\right) = (1-\nu)\nu(b-a) \tag{3.6}$$

$$\times \int_0^1 \left[f'\left((1-t)b + t\left((1-\nu)a+\nu b\right)\right) - f'\left((1-t)a + t\left((1-\nu)a+\nu b\right)\right)\right]dt. \tag{3.7}$$

That is an equality of interest in itself.

Since $a<b$ and $\nu \in (0,1)$, then $(1-\nu)a+\nu b \in (a,b)$ and

$$(1-t)a + t\left((1-\nu)a+\nu b\right) \in \left[a, (1-\nu)a+\nu b\right]$$

while

$$(1-t)b + t\left((1-\nu)a+\nu b\right) \in \left[(1-\nu)a+\nu b, b\right]$$

for any $t\in[0,1]$.

By the monotonicity of the derivative we have

$$f'_+\left((1-\nu)a+\nu b\right) \le f'\left((1-t)b + t\left((1-\nu)a+\nu b\right)\right) \le f'_-(b) \tag{3.8}$$

and

$$f'_+(a) \le f'\left((1-t)a + t\left((1-\nu)a+\nu b\right)\right) \le f'_-\left((1-\nu)a+\nu b\right) \tag{3.9}$$

for any $t\in[0,1]$.

By integrating the inequalities (3.8) and (3.9) we get

$$f'_+\left((1-\nu)a+\nu b\right) \le \int_0^1 f'\left((1-t)b + t\left((1-\nu)a+\nu b\right)\right)dt \le f'_-(b)$$

and

$$f'_+(a) \le \int_0^1 f'\left((1-t)a + t\left((1-\nu)a+\nu b\right)\right)dt \le f'_-\left((1-\nu)a+\nu b\right),$$

which implies that

$$\begin{aligned} &f'_+\left((1-\nu)a+\nu b\right) - f'_-\left((1-\nu)a+\nu b\right) \le \int_0^1 f'\left((1-t)b + t\left((1-\nu)a+\nu b\right)\right)dt \\ &- \int_0^1 f'\left((1-t)a + t\left((1-\nu)a+\nu b\right)\right)dt \le f'_-(b) - f'_+(a). \end{aligned}$$

Making use of the equality (3.6) we the obtain the desired result (3.1).

If we consider the convex function $f:[a,b]\to\mathbb{R}$, $f(x) = \left|x - \frac{a+b}{2}\right|$, then we have $f'_+\left(\frac{a+b}{2}\right) = 1$, $f'_-\left(\frac{a+b}{2}\right) = -1$ and by replacing in (3.2) we get in all terms the same quantity $\frac{1}{2}(b-a)$ which show that the constant $\frac{1}{4}$ is best possible in both inequalities from (3.2). □

Corollary 3.1. *If the function $f: I\subset\mathbb{R}\to\mathbb{R}$ is a differentiable convex function on $\mathring{I}$, then for any $a,b\in\mathring{I}$ and $\nu\in[0,1]$ we have*

$$\begin{aligned} 0 &\le (1-\nu) f(a) + \nu f(b) - f\left((1-\nu)a+\nu b\right) \\ &\le \nu(1-\nu)(b-a)\left[f'(b) - f'(a)\right]. \end{aligned} \tag{3.10}$$

Proof. If $a < b$, then the inequality (3.10) follows by (3.1). If $b < a$, then by (3.1) we get

$$
\begin{aligned}
0 &\leq (1-\nu) f(b) + \nu f(a) - f((1-\nu) b + \nu a) \\
&\leq \nu (1-\nu)(b-a)[f'(b) - f'(a)]
\end{aligned} \tag{3.11}
$$

for any $\nu \in [0,1]$. If we replace ν by $1-\nu$ in (3.11), then we get (3.10). □

We can prove now the following reverse of the second inequality in (2.3) and the first inequality in (2.5).

Theorem 3.1. *Let f be a differentiable convex function on $[0,\infty)$ with $f(1) = 0$. Then for all $Q, P, W \in \mathcal{P}$ and $t \in [0,1]$ we have*

$$
\begin{aligned}
0 &\leq (1-t) I_f(Q,W) + t I_f(P,W) - A_{f,t}(Q,P,W) \\
&\leq t(1-t) \Delta_{f'}(Q,P,W)
\end{aligned} \tag{3.12}
$$

and

$$
0 \leq S_{f,t}(Q,P,W) - M_f(Q,P,W) \leq \frac{1}{2}\left(t - \frac{1}{2}\right) \Delta_{f',t}(Q,P,W), \tag{3.13}
$$

where

$$
\Delta_{f',t}(Q,P,W) = \int_X (q(x) - p(x)) \times \left[f'\left((1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)}\right) - f'\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right)\right] d\mu(x).
$$

Proof. From the inequality (3.12) we get

$$
\begin{aligned}
0 &\leq (1-t) f\left(\frac{q(x)}{w(x)}\right) + t f\left(\frac{p(x)}{w(x)}\right) - f\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right) \\
&\leq t(1-t)\left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{p(x)}{w(x)}\right)\right]\left(\frac{q(x)}{w(x)} - \frac{p(x)}{w(x)}\right).
\end{aligned} \tag{3.14}
$$

If we multiply this inequality by $w(x) \geq 0$ and integrate on X we get (3.12).

For any $x, y \in \mathring{I}$ we have

$$
0 \leq \frac{f(x) + f(y)}{2} - f\left(\frac{x+y}{2}\right) \leq \frac{1}{4}(x-y)[f'(x) - f'(y)]. \tag{3.15}
$$

If in this inequality we take $x = (1-t)a + tb$, $y = (1-t)b + ta$ with $a, b \in \mathring{I}$ and $t \in [0,1]$, then we get

$$
\begin{aligned}
0 &\leq \frac{f((1-t)a + tb) + f((1-t)b + ta)}{2} - f\left(\frac{a+b}{2}\right) \\
&\leq \frac{1}{4}((1-t)a + tb - (1-t)b - ta) \\
&\quad \times [f'((1-t)a + tb) - f'((1-t)b + ta)] \\
&= \frac{1}{2}\left(t - \frac{1}{2}\right)(b-a)[f'((1-t)a + tb) - f'((1-t)b + ta)].
\end{aligned} \tag{3.16}
$$

From this inequality we have

$$
\begin{aligned}
0 &\leq \frac{1}{2}\left[f\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right) + f\left((1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)}\right)\right] \\
&\quad - f\left(\frac{q(x) + p(x)}{2w(x)}\right) \\
&\leq \frac{1}{2}\left(t - \frac{1}{2}\right)\left(\frac{q(x)}{w(x)} - \frac{p(x)}{w(x)}\right) \\
&\quad \times \left[f'\left((1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)}\right) - f'\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right)\right].
\end{aligned}
$$

If we multiply this inequality by $w(x) \geq 0$ and integrate on X we get (3.12). □

Corollary 3.2. *Let f be a differentiable convex function on $[0,\infty)$ with $f(1)=0$ and $Q,\ P,\ W\in\mathcal{P}$. If there exists $0<r<1<R<\infty$ such that the condition (2.15) holds, then*

$$\begin{aligned} 0 &\le (1-t)\, I_f(Q,W) + tI_f(P,W) - A_{f,t}(Q,P,W) \\ &\le t(1-t)\left[f'(R)-f'(r)\right] d_1(Q,P) \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} 0 &\le S_{f,t}(Q,P,W) - M_f(Q,P,W) \\ &\le \frac{1}{2}\left|t-\frac{1}{2}\right| \left[f'(R)-f'(r)\right] d_1(Q,P) \end{aligned} \tag{3.18}$$

Proof. The inequality (3.17) is obvious. For (3.18), we have

$$\begin{aligned} \frac{1}{2}\left(t-\frac{1}{2}\right)\Delta_{f',t}(Q,P,W) &= \frac{1}{2}\left|t-\frac{1}{2}\right| \left|\Delta_{f',t}(Q,P,W)\right| \\ &\le \frac{1}{2}\left|t-\frac{1}{2}\right| \int_X \left|q(x)-p(x)\right| \\ \times \left| f'\left((1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)}\right) - f'\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right)\right| d\mu(x) \\ &\le \frac{1}{2}\left[f'(R)-f'(r)\right]\left|t-\frac{1}{2}\right| \int_X \left|q(x)-p(x)\right| d\mu(x) \\ &= \frac{1}{2}\left|t-\frac{1}{2}\right| \left[f'(R)-f'(r)\right] d_1(Q,P). \end{aligned}$$

□

Corollary 3.3. *Let f be a twice differentiable convex function on $[0,\infty)$ with $f(1)=0$ and $Q,\ P,\ W\in\mathcal{P}$. If there exists $0<r<1<R<\infty$ such that the conditions (2.15) and (2.18) hold, then*

$$\begin{aligned} 0 &\le (1-t)\, I_f(Q,W) + tI_f(P,W) - A_{f,t}(Q,P,W) \\ &\le t(1-t)\left\|f''\right\|_{[r,R],\infty} d_{\chi^2}(Q,P,W) \end{aligned} \tag{3.19}$$

and

$$0 \le S_{f,t}(Q,P,W) - M_f(Q,P,W) \le \left|t-\frac{1}{2}\right|^2 \left\|f''\right\|_{[r,R],\infty} d_{\chi^2}(Q,P,W). \tag{3.20}$$

Proof. We have

$$\begin{aligned} \frac{1}{2}\left(t-\frac{1}{2}\right)\Delta_{f',t}(Q,P,W) \quad &\le \quad \frac{1}{2}\left|t-\frac{1}{2}\right| \int_X \left|q(x)-p(x)\right| \\ &\times \quad \left| f'\left((1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)}\right) - f'\left((1-t)\frac{q(x)}{w(x)} + t\frac{p(x)}{w(x)}\right)\right| d\mu(x) \\ &\le \quad \frac{1}{2}\left|t-\frac{1}{2}\right| \left\|f''\right\|_{[r,R],\infty} \int_X \left|q(x)-p(x)\right| \\ &\times \quad \left|(1-t)\frac{p(x)}{w(x)} + t\frac{q(x)}{w(x)} - (1-t)\frac{q(x)}{w(x)} - t\frac{p(x)}{w(x)}\right| d\mu(x) \\ &= \quad \left|t-\frac{1}{2}\right|^2 \left\|f''\right\|_{[r,R],\infty} \int_X \left|q(x)-p(x)\right| \frac{\left|q(x)-p(x)\right|}{w(x)} d\mu(x) \\ &= \quad \left|t-\frac{1}{2}\right|^2 \left\|f''\right\|_{[r,R],\infty} d_{\chi^2}(Q,P,W), \end{aligned}$$

which proves (3.20). □

4. Examples

Consider the *dichotomy class* generated by the function $f_\alpha : [0,\infty) \to \mathbb{R}$ that is given by

$$f_\alpha(u) = \begin{cases} u - 1 - \ln u & \text{for } \alpha = 0; \\ \dfrac{1}{\alpha(1-\alpha)}\left[\alpha u + 1 - \alpha - u^\alpha\right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ 1 - u + u\ln u & \text{for } \alpha = 1. \end{cases}$$

We have

$$\begin{aligned} A_{f_\alpha,t}(Q,P,W) &= \int_X f\left[\frac{(1-t)q(x)+tp(x)}{w(x)}\right]w(x)\,d\mu(x) \\ &= \begin{cases} -\int_X w(x)\ln\left[\dfrac{(1-t)q(x)+tp(x)}{w(x)}\right]d\mu(x) & \text{for } \alpha = 0; \\ \dfrac{1}{\alpha(1-\alpha)}\left[1 - \int_X [(1-t)q(x)+tp(x)]^\alpha w^{1-\alpha}(x)\,d\mu(x)\right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \int_X [(1-t)q(x)+tp(x)]\ln\left[\dfrac{(1-t)q(x)+tp(x)}{w(x)}\right]d\mu(x) & \text{for } \alpha = 1 \end{cases} \end{aligned}$$

and

$$\begin{aligned} M_{f_\alpha}(Q,P,W) &= \int_X f\left[\frac{q(x)+p(x)}{2w(x)}\right]w(x)\,d\mu(x) \\ &= \begin{cases} -\int_X w(x)\ln\left[\dfrac{q(x)+p(x)}{2w(x)}\right]d\mu(x) & \text{for } \alpha = 0; \\ \dfrac{1}{\alpha(1-\alpha)}\left[1 - \int_X \left[\dfrac{q(x)+p(x)}{2}\right]^\alpha w^{1-\alpha}(x)\,d\mu(x)\right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \int_X \left[\dfrac{q(x)+p(x)}{2}\right]\ln\left[\dfrac{q(x)+p(x)}{2w(x)}\right]d\mu(x) & \text{for } \alpha = 1. \end{cases} \end{aligned}$$

Let us recall the following special means:

a) The *arithmetic mean*

$$A(a,b) := \frac{a+b}{2}, \quad a,b > 0,$$

b) The *geometric mean*

$$G(a,b) := \sqrt{ab}; \quad a,b \geq 0,$$

c) The *harmonic mean*

$$H(a,b) := \frac{2}{\frac{1}{a}+\frac{1}{b}}; \quad a,b > 0,$$

d) The *identric mean*

$$I(a,b) := \begin{cases} \dfrac{1}{e}\left(\dfrac{b^b}{a^a}\right)^{\frac{1}{b-a}} & \text{if } b \neq a \\ a & \text{if } b = a \end{cases} ; \; a,b > 0$$

e) The *logarithmic mean*

$$L(a,b) := \begin{cases} \dfrac{b-a}{\ln b - \ln a} & \text{if} \quad b \neq a \\ a & \text{if} \quad b = a \end{cases} \quad ; \quad a,b>0$$

f) The *p-logarithmic mean*

$$L_p(a,b) := \begin{cases} \left(\dfrac{b^{p+1}-a^{p+1}}{(p+1)(b-a)}\right)^{\frac{1}{p}} & \text{if} \quad b \neq a, \ \ p \in \mathbb{R}\backslash\{-1,0\} \\ a & \text{if} \quad b = a \end{cases} \quad ; \quad a,b>0.$$

If we put $L_0(a,b) := I(a,b)$ and $L_{-1}(a,b) := L(a,b)$, then it is well known that the function $\mathbb{R} \ni p \mapsto L_p(a,b)$ is monotonic increasing on $\mathbb{R}$.

We observe that for $p \in \mathbb{R}\backslash\{-1,0\}$ we have

$$\int_0^1 [(1-t)a+tb]^p \, dt = L_p^p(a,b), \quad \int_0^1 [(1-t)a+tb]^{-1} \, dt = L^{-1}(a,b)$$

and

$$\int_0^1 \ln[(1-t)a+tb] \, dt = \ln I(a,b).$$

We also have

$$\begin{aligned} &\int_0^1 [(1-t)a+tb] \ln[(1-t)a+tb] \, dt \\ &= \frac{1}{b-a}\int_a^b t\ln t dt = \frac{1}{2}\frac{1}{b-a}\int_a^b \ln t d\left(t^2\right) \\ &= \frac{1}{2}\frac{1}{b-a}\left[b^2\ln b - a^2 \ln a - \frac{b^2-a^2}{2}\right] \\ &= \frac{1}{2}\frac{1}{b-a}\left[\frac{b^2\ln b^2 - a^2\ln a^2}{2} - \frac{b^2-a^2}{2}\right] \\ &= \frac{1}{2}\frac{1}{b-a}\frac{b^2-a^2}{2}\left[\frac{b^2\ln b^2 - a^2\ln a^2}{b^2-a^2} - 1\right] \\ &= \frac{1}{4}(b+a)\ln I\left(a^2,b^2\right) = \frac{1}{2}A(a,b)\ln I\left(a^2,b^2\right). \end{aligned}$$

Therefore

$$\begin{aligned} A_{f_\alpha}(Q,P,W) &:= \int_0^1 A_{f_\alpha,t}(Q,P,W)\,dt \\ &= \int_X \left(\int_0^1 f\left[\frac{(1-t)q(x)+tp(x)}{w(x)}\right]dt\right) w(x)\,d\mu(x) \end{aligned}$$

$$= \begin{cases} -\int_X \left(\int_0^1 \ln\left[\frac{(1-t)\,q(x) + tp(x)}{w(x)} \right] dt \right) w(x)\,d\mu(x) & \text{for } \alpha = 0; \\ \frac{1}{\alpha(1-\alpha)} \left[1 - \int_X \left(\int_0^1 \left[\frac{(1-t)\,q(x) + tp(x)}{w(x)} \right]^\alpha dt \right) w(x)\,d\mu(x) \right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \int_X \int_0^1 \left(\left[\frac{(1-t)\,q(x) + tp(x)}{w(x)} \right] \ln\left[\frac{(1-t)\,q(x) + tp(x)}{w(x)} \right] dt \right) w(x)\,d\mu(x) & \text{for } \alpha = 1 \end{cases}$$

$$= \begin{cases} -\int_X \ln I\left(\frac{q(x)}{w(x)}, \frac{p(x)}{w(x)} \right) w(x)\,d\mu(x) & \text{for } \alpha = 0; \\ \frac{1}{\alpha(1-\alpha)} \left[1 - \int_X L_\alpha^\alpha \left(\frac{q(x)}{w(x)}, \frac{p(x)}{w(x)} \right) w(x)\,d\mu(x) \right] & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \frac{1}{2} \int_X A\left(\frac{q(x)}{w(x)}, \frac{p(x)}{w(x)} \right) \ln I\left(\left(\frac{q(x)}{w(x)} \right)^2, \left(\frac{p(x)}{w(x)} \right)^2 \right) w(x)\,d\mu(x) & \text{for } \alpha = 1. \end{cases}$$

According to Corollary 2.1 we have

$$0 \le M_{f_\alpha}(Q,P,W) \le A_{f_\alpha}(Q,P,W) \le \frac{1}{2}\left[I_{f_\alpha}(Q,W) + I_{f_\alpha}(P,W) \right] \tag{4.1}$$

and the mapping

$$\mathcal{P} \times \mathcal{P} \ni (Q,P) \mapsto A_{f_\alpha}(Q,P,W) \in [0,\infty) \tag{4.2}$$

is convex.

Observe also that

$$f_\alpha'(u) = \begin{cases} 1 - \frac{1}{u} & \text{for } \alpha = 0; \\ \frac{1}{1-\alpha}\left(1 - u^{\alpha-1}\right) & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \ln u & \text{for } \alpha = 1, \end{cases}$$

which implies that

$$\Delta_{f_\alpha'}(Q,P,W) := \int_X \left[f_\alpha'\left(\frac{q(x)}{w(x)} \right) - f_\alpha'\left(\frac{p(x)}{w(x)} \right) \right] (q(x) - p(x))\,d\mu(x)$$

$$= \begin{cases} \int_X \frac{(q(x) - p(x))^2}{p(x)\,q(x)} w(x)\,d\mu(x) & \text{for } \alpha = 0; \\ \frac{1}{\alpha - 1} \int_X \frac{q^{\alpha-1}(x) - p^{\alpha-1}(x)}{w^\alpha(x)} (q(x) - p(x))\,d\mu(x) & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \int_X (q(x) - p(x)) \ln\left(\frac{q(x)}{p(x)} \right) d\mu(x) & \text{for } \alpha = 1. \end{cases}$$

For all Q, P, $W \in \mathcal{P}$ we have by Theorem 2.2 that

$$0 \le A_{f_\alpha}(Q,P,W) - Mf_\alpha(Q,P,W) \le \frac{1}{8}\Delta_{f_\alpha'}(Q,P,W) \tag{4.3}$$

and

$$0 \le \frac{1}{2}\left[I_{f_\alpha}(Q,W) + I_{f_\alpha}(P,W) \right] - A_{f_\alpha}(Q,P,W) \le \frac{1}{8}\Delta_{f_\alpha'}(Q,P,W). \tag{4.4}$$

If there exists $0 < r < 1 < R < \infty$ such that the following condition holds

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

then by Corollary 2.2

$$\begin{aligned} 0 &\leq A_{f_\alpha}(Q,P,W) - M_{f_\alpha}(Q,P,W) \qquad (4.5) \\ &\leq \frac{1}{8} d_1(Q,P) \begin{cases} \dfrac{R-r}{rR} & \text{for } \alpha = 0; \\ \dfrac{R^{\alpha-1} - r^{\alpha-1}}{\alpha - 1} & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \ln\left(\dfrac{R}{r}\right) & \text{for } \alpha = 1 \end{cases} \qquad (4.6) \end{aligned}$$

and

$$\begin{aligned} 0 &\leq \frac{1}{2}\left[I_f(Q,W) + I_f(P,W)\right] - A_f(Q,P,W) \qquad (4.7) \\ &\leq \frac{1}{8} d_1(Q,P) \begin{cases} \dfrac{R-r}{rR} & \text{for } \alpha = 0; \\ \dfrac{R^{\alpha-1} - r^{\alpha-1}}{\alpha - 1} & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \ln\left(\dfrac{R}{r}\right) & \text{for } \alpha = 1. \end{cases} \qquad (4.8) \end{aligned}$$

Further, since

$$f''_\alpha(u) = \begin{cases} \dfrac{1}{u^2} & \text{for } \alpha = 0; \\ u^{\alpha-2} & \text{for } \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \dfrac{1}{u} & \text{for } \alpha = 1, \end{cases}$$

hence by Corollary 2.3 we have

$$\begin{aligned} 0 &\leq A_f(Q,P,W) - M_f(Q,P,W) \qquad (4.9) \\ &\leq \frac{1}{8} d_{\chi^2}(Q,P,W) \begin{cases} \dfrac{1}{r^2} & \text{for } \alpha = 0; \\ R^{\alpha-2} & \text{for } \alpha \geq 2; \\ r^{\alpha-2} & \text{for } \alpha < 2,\ \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \dfrac{1}{r} & \text{for } \alpha = 1, \end{cases} \qquad (4.10) \end{aligned}$$

and

$$\begin{aligned} 0 &\leq \frac{1}{2}\left[I_f(Q,W) + I_f(P,W)\right] - A_f(Q,P,W) \qquad (4.11) \\ &\leq \frac{1}{8} d_{\chi^2}(Q,P,W) \begin{cases} \dfrac{1}{r^2} & \text{for } \alpha = 0; \\ R^{\alpha-2} & \text{for } \alpha \geq 2; \\ r^{\alpha-2} & \text{for } \alpha < 2,\ \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \dfrac{1}{r} & \text{for } \alpha = 1. \end{cases} \qquad (4.12) \end{aligned}$$

The interested reader may apply the above general results for other particular divergences of interest generated by the convex functions provided in the introduction. We omit the details.

Article Information

Acknowledgements: The authors are grateful to the referees for their careful reading of this manuscript and several valuable suggestions which improved the quality of the article.

Author's contributions: All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Conflict of Interest Disclosure: No potential conflict of interest was declared by the author.

Supporting/Supporting Organizations: No grants were received from any public, private or non-profit organizations for this research.

Ethical Approval and Participant Consent: It is declared that during the preparation process of this study, scientific and ethical principles were followed and all the studies benefited from are stated in the bibliography.

Plagiarism Statement: This article was scanned by the plagiarism program. No plagiarism detected.

Availability of data and materials: Not applicable.

References

[1] Csiszár, I.: *Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten*. Magyar Tud. Akad. Mat. Kutató Int. Közl. **8**, 85-108 (1963).

[2] Cerone, P., Dragomir, S. S., Österreicher, F.: *Bounds on extended f-divergences for a variety of classes*. Kybernetika (Prague). **40**(6), 745-756 (2004) Preprint, RGMIA Res. Rep. Coll. **6**(1), 5 (2003). http://rgmia.vu.edu.au/v6n1.html].

[3] Kafka, P., Österreicher, F., Vincze, I.: *On powers of f-divergence defining a distance*. Studia Scientiarum Mathematicarum Hungarica. **26**, 415-422 (1991).

[4] Österreicher, F. Vajda, I.: *A new class of metric divergences on probability spaces and its applicability in statistics*. Annals of the Institute of Statistical Mathematics. **55** (3), 639-653 (2003).

[5] Liese, F., Vajda, I.: *Convex Statistical Distances*. Teubuer-Texte zur Mathematik, Band, Leipzig. **95** 1987.

[6] Cerone, P., Dragomir, S. S.: *Approximation of the integral mean divergence and f-divergence via mean results*. Mathematical and Computer Modelling. **42**(1-2), 207-219 (2005).

[7] Dragomir, S. S.: *Some inequalities for (m, M)-convex mappings and applications for the Csiszár Φ-divergence in information theory*. Mathematical Journal of Ibaraki University. **33**, 35-50 (2001).

[8] Dragomir, S. S.: *Some inequalities for two Csiszár divergences and applications*. Matematichki Bilten. **25**, 73-90 (2001).

[9] Dragomir, S. S.: *An inequality improving the first Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products*. Journal of Inequalities in Pure and Applied Mathematics. **3** (2), 31 (2002).

[10] Dragomir, S. S.: *An inequality improving the second Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products*. Journal of Inequalities in Pure and Applied Mathematics. **3**(3), 35 (2002).

[11] Dragomir, S. S.: *An upper bound for the Csiszár f-divergence in terms of the variational distance and applications*. Panamerican Mathematical Journal. **12**(4), 43-54 (2002).

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

[13] Dragomir, S. S.: *Bounds for f-divergences under likelihood ratio constraints*. Applications of Mathematics. **48**(3), 205-223 (2003).

[14] Dragomir, S. S.: *New inequalities for Csiszár divergence and applications*. Acta Mathematica Vietnamica. **28**(2), 123-134 (2003).

[15] Dragomir, S. S.: *A generalized f-divergence for probability vectors and applications*. Panamerican Mathematical Journal. **13**(4), 61-69 (2003).

[16] Dragomir, S. S.: *Some inequalities for the Csiszár φ-divergence when φ is an L-Lipschitzian function and applications*. Italian Journal of Pure and Applied Mathematics. **15**, 57-76 (2004).

[17] Dragomir, S. S.: *A converse inequality for the Csiszár Φ-divergence*. Tamsui Oxford Journal of Mathematical Sciences. **20**(1), 35-53 (2004).

[18] Dragomir, S. S.: *Some general divergence measures for probability distributions*. Acta Mathematica Hungarica. **109**(4), 331-345 (2005).

[19] Dragomir, S. S.: *Bounds for the normalized Jensen functional*. Bulletin of the Australian Mathematical Society. **74**(3), 471-478 (2006).

[20] Dragomir, S. S.: *A refinement of Jensen's inequality with applications for f-divergence measures*. Taiwanese Journal of Mathematics. **14**(1), 153-164 (2010).

[21] Dragomir, S. S.: *A generalization of f-divergence measure to convex functions defined on linear spaces*. Communications in Mathematical Analysis. **15**(2), 1-14 (2013).

Affiliations

SILVESTRU SEVER DRAGOMIR
ADDRESS: Mathematics, College of Engineering & Science, Victoria University, PO Box 14428, Melbourne City, MC 8001, Australia.
DST-NRF Centre of Excellence in the Mathematical and Statistical Sciences, School of Computer Science& Applied Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa.
E-MAIL: sever.dragomir@vu.edu.au
ORCID ID:0000-0003-2902-6805