SOME NEW f-DIVERGENCE MEASURES AND THEIR BASIC PROPERTIES

SILVESTRU SEVER DRAGOMIR[1,2]

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 intyroduced by Csiszar in 1963. Applications for the *dichotomy class* of convex functions are provided as well.

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 [3] introduced the concept of f-divergence as follows.

Definition 1. *Let P, $Q\in\mathcal{P}$. Then*

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

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

Remark 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]).

1991 *Mathematics Subject Classification.* 94A17, 26D15.

Key words and phrases. f-divergence measures, HH f-divergence measures, Kullback-Leibler divergence, Hellinger discrimination, χ^2-divergence, Jeffrey's distance.

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 $f_{\alpha} : [0,\infty) \to \mathbb{R}$

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

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 $\left[I_{\varphi_{\alpha}}(Q,P)\right]^{\alpha}$.

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 [19] that this class provides the distances $\left[I_{\Phi_{\alpha}}(Q,P)\right]^{\frac{1}{\alpha}}$.

1.5. **Divergences of Arimoto-type.** This class is generated by the functions

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

It has been shown in [21] 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 [20] (see also [2]).

Theorem 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 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 2 (see [2, Theorem 3]).

Theorem 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*

$$0 \le I_f(Q,P) \le \frac{1}{2}[f(0) + f^*(0)]V(Q,P) \tag{1.5}$$

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

For other inequalities for f-divergence see [1], [4]-[17].

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

$$0\le A_{f,t}(Q,P,W)\le(1-t)\,I_f(Q,W)+tI_f(P,W) \tag{2.3}$$

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

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

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 α, $\beta \geq 0$ and such that $\alpha + \beta = 1$. If (Q_1, P_1), $(Q_2, P_2) \in \mathcal{P} \times \mathcal{P}$, then

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

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

Using the convexity of f we have

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

namely

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

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

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

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 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*

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

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 1 (Dragomir, 2002 [6] and [7])**.** *Let $h:[a,b]\to\mathbb{R}$ be a convex function on $[a,b]$. Then*

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

and

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

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

We have the reverse inequalities:

Theorem 5. *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*

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

and

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

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

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

If we multiply this inequality by $w(x)\geq0$ 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 $w(x)\geq0$ and integrate on X we get (2.12). $\square$

Corollary 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*

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

then

$$0\leq A_f(Q,P,W)-M_f(Q,P,W)\leq\frac{1}{8}\left[f'(R)-f'(r)\right]d_1(Q,P) \tag{2.15}$$

and

$$0\leq\frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]-A_f(Q,P,W)\leq\frac{1}{8}\left[f'(R)-f'(r)\right]d_1(Q,P) \tag{2.16}$$

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

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

which proves the desired inequalities (2.15) and (2.16). □

Corollary 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 (r,R) holds and*

$$\|f''\|_{[r,R],\infty} := \sup_{t\in[r,R]} |f''(t)| < \infty, \tag{2.17}$$

then

$$0 \leq A_f(Q,P,W) - M_f(Q,P,W) \leq \frac{1}{8}\|f''\|_{[r,R],\infty}\, d_{\chi^2}(Q,P,W) \tag{2.18}$$

and

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

where

$$d_{\chi^2}(Q,P,W) := \int_X \frac{(q(x)-p(x))^2}{w(x)}\,d\mu(x). \tag{2.20}$$

Proof. We have

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

which proves the desired results (2.18) and (2.19). □

3. Further Results

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

Lemma 2. *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

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

which is equivalent to

$$\begin{aligned} &(1-\nu) f(a) + \nu f(b) - f((1-\nu)a+\nu b) = (1-\nu)\nu(b-a) \\ &\times \int_0^1 \left[f'((1-t)b + t((1-\nu)a+\nu b)) - f'((1-t)a + t((1-\nu)a+\nu b))\right] dt. \end{aligned} \tag{3.6}$$

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

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

while

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

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

By the monotonicity of the derivative we have

$$f'_{+}((1-\nu)a+\nu b)\le f'((1-t)b+t((1-\nu)a+\nu b))\le f'_{-}(b) \tag{3.7}$$

and

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

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

By integrating the inequalities (3.7) and (3.8) we get

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

and

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

which implies that

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

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 4. *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*

$$\begin{aligned}0&\le(1-\nu)f(a)+\nu f(b)-f((1-\nu)a+\nu b)\\&\le\nu(1-\nu)(b-a)[f'(b)-f'(a)].\end{aligned} \tag{3.9}$$

Proof. If $a<b$, then the inequality (3.9) follows by (3.1). If $b<a$, then by (3.1) we get

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

for any $\nu\in[0,1]$. If we replace ν by $1-\nu$ in (3.10), then we get (3.9). □

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

Theorem 6. *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*

$$0\le(1-t)\,I_f(Q,W)+tI_f(P,W)-A_{f,t}(Q,P,W)\le t(1-t)\,\Delta_{f'}(Q,P,W)\tag{3.11}$$

and

$$0\le S_{f,t}(Q,P,W)-M_f(Q,P,W)\le\frac{1}{2}\left(t-\frac{1}{2}\right)\Delta_{f',t}(Q,P,W),\tag{3.12}$$

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 (3.11) we get

$$\begin{aligned}0&\le(1-t)\,f\left(\frac{q(x)}{w(x)}\right)+tf\left(\frac{p(x)}{w(x)}\right)-f\left((1-t)\frac{q(x)}{w(x)}+t\frac{p(x)}{w(x)}\right)\\&\le t(1-t)\left[f'\left(\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)}{w(x)}\right)\right]\left(\frac{q(x)}{w(x)}-\frac{p(x)}{w(x)}\right).\end{aligned}\tag{3.13}$$

If we multiply this inequality by $w(x)\ge0$ and integrate on X we get (3.11).

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

$$0\le\frac{f(x)+f(y)}{2}-f\left(\frac{x+y}{2}\right)\le\frac{1}{4}(x-y)\left[f'(x)-f'(y)\right].\tag{3.14}$$

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

$$\begin{aligned}0&\le\frac{f((1-t)\,a+tb)+f((1-t)\,b+ta)}{2}-f\left(\frac{a+b}{2}\right)\\&\le\frac{1}{4}((1-t)\,a+tb-(1-t)\,b-ta)\\&\quad\times\left[f'((1-t)\,a+tb)-f'((1-t)\,b+ta)\right]\\&=\frac{1}{2}\left(t-\frac{1}{2}\right)(b-a)\left[f'((1-t)\,a+tb)-f'((1-t)\,b+ta)\right].\end{aligned}\tag{3.15}$$

From this inequality we have

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

If we multiply this inequality by $w(x)\ge0$ and integrate on X we get (3.11). $\square$

Corollary 5. *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 ((r,R)) holds, then*

$$0\leq(1-t)\,I_f(Q,W)+tI_f(P,W)-A_{f,t}(Q,P,W)\leq t(1-t)\left[f'(R)-f'(r)\right]d_1(Q,P)\tag{3.16}$$

and

$$0\leq S_{f,t}(Q,P,W)-M_f(Q,P,W)\leq\frac{1}{2}\left|t-\frac{1}{2}\right|\left[f'(R)-f'(r)\right]d_1(Q,P)\tag{3.17}$$

Proof. The inequality (3.16) is obvious. For (3.17), we have

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

□

Corollary 6. *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 ((r,R)) and (2.17) hold, then*

$$0\leq(1-t)\,I_f(Q,W)+tI_f(P,W)-A_{f,t}(Q,P,W)\leq t(1-t)\left\|f''\right\|_{[r,R],\infty}d_{\chi^2}(Q,P,W)\tag{3.18}$$

and

$$0\leq S_{f,t}(Q,P,W)-M_f(Q,P,W)\leq\left|t-\frac{1}{2}\right|^2\left\|f''\right\|_{[r,R],\infty}d_{\chi^2}(Q,P,W).\tag{3.19}$$

Proof. We have

$$\begin{aligned}\frac{1}{2}\left(t-\frac{1}{2}\right)\Delta_{f',t}(Q,P,W)&\leq\frac{1}{2}\left|t-\frac{1}{2}\right|\int_X|q(x)-p(x)|\\&\times\left|f'\left((1-t)\frac{p(x)}{w(x)}+t\frac{q(x)}{w(x)}\right)-f'\left((1-t)\frac{q(x)}{w(x)}+t\frac{p(x)}{w(x)}\right)\right|d\mu(x)\end{aligned}$$

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

which proves (3.19). □

4. Examples

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

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

We have

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

and

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

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*

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

e) The *logarithmic mean*

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

f) The *p-logarithmic mean*

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

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 $p \in \mathbb{R}\setminus\{-1,0\}$ 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

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

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

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

According to Corollary 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 5 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

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

and

$$\begin{aligned} 0 &\le \frac{1}{2}\left[I_f(Q,W) + I_f(P,W)\right] - A_f(Q,P,W) \\ &\le \frac{1}{8}d_1(Q,P)\begin{cases} \frac{R-r}{rR} & \text{for } \ \alpha = 0; \\ \frac{R^{\alpha-1}-r^{\alpha-1}}{\alpha-1} & \text{for } \ \alpha \in \mathbb{R}\backslash\{0,1\}; \\ \ln\left(\frac{R}{r}\right) & \text{for } \ \alpha = 1. \end{cases} \end{aligned} \tag{4.6}$$

Further, since

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

hence by Corollary 3 we have

$$\begin{aligned} 0 &\le A_f(Q,P,W) - M_f(Q,P,W) \\ &\le \frac{1}{8}d_{\chi^2}(Q,P,W)\begin{cases} \frac{1}{r^2} & \text{for } \ \alpha = 0; \\ R^{\alpha-2} & \text{for } \alpha \ge 2; \\ r^{\alpha-2} & \text{for } \alpha < 2, \ \alpha \in \mathbb{R}\backslash\{0,1\}; \\ \frac{1}{r} & \text{for } \ \alpha = 1, \end{cases} \end{aligned} \tag{4.7}$$

and

$$0 \leq \frac{1}{2}\left[I_f(Q,W)+I_f(P,W)\right]-A_f(Q,P,W) \tag{4.8}$$
$$\leq \frac{1}{8} d_{\chi^2}(Q,P,W)\begin{cases} \frac{1}{r^2} & \text{for } \alpha=0; \\ R^{\alpha-2} & \text{for } \alpha \geq 2; \\ r^{\alpha-2} & \text{for } \alpha<2,\ \alpha \in \mathbb{R}\setminus\{0,1\}; \\ \frac{1}{r} & \text{for } \alpha=1. \end{cases}$$

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.

References

[1] P. Cerone and S. S. Dragomir, Approximation of the integral mean divergence and f-divergence via mean results. *Math. Comput. Modelling* **42** (2005), no. 1-2, 207–219.

[2] P. Cerone, S. S. Dragomir and F. Österreicher, Bounds on extended f-divergences for a variety of classes, *Kybernetika* (Prague) **40** (2004), no. 6, 745–756. Preprint, *RGMIA Res. Rep. Coll.* **6**(2003), No.1, Article 5. [ONLINE: http://rgmia.vu.edu.au/v6n1.html].

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

[4] S. S. Dragomir, Some inequalities for (m,M)-convex mappings and applications for the Csiszár Φ-divergence in information theory. *Math. J. Ibaraki Univ.* **33** (2001), 35–50.

[5] S. S. Dragomir, Some inequalities for two Csiszár divergences and applications. *Mat. Bilten* No. **25** (2001), 73–90.

[6] S. S. Dragomir, An inequality improving the first Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products, *J. Inequal. Pure and Appl. Math.*, **3** (2) (2002), Art. 31.

[7] S. S. Dragomir, An inequality improving the second Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products, *J. Inequal. Pure and Appl. Math.*, **3** (3) (2002), Art. 35.

[8] S. S. Dragomir, An upper bound for the Csiszár f-divergence in terms of the variational distance and applications. *Panamer. Math. J.* **12** (2002), no. 4, 43–54.

[9] S. S. Dragomir, Upper and lower bounds for Csiszár f-divergence in terms of Hellinger discrimination and applications. *Nonlinear Anal. Forum* **7** (2002), no. 1, 1–13

[10] S. S. Dragomir, Bounds for f-divergences under likelihood ratio constraints. *Appl. Math.* **48** (2003), no. 3, 205–223.

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

[12] S. S. Dragomir, A generalized f-divergence for probability vectors and applications. *Panamer. Math. J.* **13** (2003), no. 4, 61–69.

[13] S. S. Dragomir, Some inequalities for the Csiszár φ-divergence when φ is an L-Lipschitzian function and applications. *Ital. J. Pure Appl. Math.* No. **15** (2004), 57–76.

[14] S. S. Dragomir, A converse inequality for the Csiszár Φ-divergence. *Tamsui Oxf. J. Math. Sci.* **20** (2004), no. 1, 35–53.

[15] S. S. Dragomir, Some general divergence measures for probability distributions. *Acta Math. Hungar.* **109** (2005), no. 4, 331–345.

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

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

[18] S. S. Dragomir, A generalization of f-divergence measure to convex functions defined on linear spaces. *Commun. Math. Anal.* **15** (2013), no. 2, 1–14.
[19] P. Kafka, F. Österreicher and I. Vincze, On powers of f-divergence defining a distance, *Studia Sci. Math. Hungar.,* **26** (1991), 415-422.
[20] F. Liese and I. Vajda, *Convex Statistical Distances,* Teubuer – Texte zur Mathematik, Band **95**, Leipzig, 1987.
[21] F. Österreicher and I. Vajda, A new class of metric divergences on probability spaces and its applicability in statistics. *Ann. Inst. Statist. Math.* **55** (2003), no. 3, 639–653.

[1]Mathematics, College 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]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