SOME f-DIVERGENCE MEASURES RELATED TO JENSEN'S ONE

SILVESTRU SEVER DRAGOMIR[1,2]

Abstract. In this paper we introduce some f-divergence measures that are related to the Jensen's divergence introduced by Burbea and Rao in 1982. We establish their joint convexity and provide some inequalities between these measures and a combination of Csiszár's f-divergence, f-midpoint divergence and f-integral divergence measures.

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 [4] 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*

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

We now give some examples of f-divergences that are well-known and often used in the literature (see also [3]).

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

Key words and phrases. Jensen divergence, f-divergence measures, HH f-divergence measures, χ^2-divergence.

RGMIA Res. Rep. Coll. 22 (2019), Art. 40, 21 pp. Received 17/04/19

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

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

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 [26] 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 [28] 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 [27] (see also [3]).

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 [3, 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 [2], [7]-[20].

2. Some Preliminary Facts

For a function f defined on an interval I of the real line $\mathbb{R}$, by following the paper by Burbea & Rao [1], we consider the $\mathcal{J}$*-divergence* between the elements t, $s \in I$ given by

$$\mathcal{J}_f(t,s) := \frac{1}{2}[f(t)+f(s)] - f\left(\frac{t+s}{2}\right).$$

As important examples of such divergences, we can consider [1],

$$\mathcal{J}_\alpha(t,s) := \begin{cases} (\alpha-1)^{-1}\left[\frac{1}{2}(t^\alpha+s^\alpha) - \left(\frac{t+s}{2}\right)^\alpha\right], & \alpha \neq 1, \\ \\ \left[t\ln(t)+s\ln(s)-(t+s)\ln\left(\frac{t+s}{2}\right)\right], & \alpha = 1. \end{cases}$$

If f is convex on I, then $\mathcal{J}_f(t,s) \geq 0$ for all $(t,s) \in I \times I$.

The following result concerning the joint convexity of $\mathcal{J}_f$ also holds:

Theorem 4 (Burbea-Rao, 1982 [1])**.** *Let f be a C^2 function on an interval I. Then $\mathcal{J}_f$ is convex (concave) on $I \times I$, if and only if f is convex (concave) and $\frac{1}{f''}$ is concave (convex) on I.*

We define the *Hermite-Hadamard trapezoid* and *mid-point divergences*

$$\mathcal{T}_f(t,s) := \frac{1}{2}[f(t)+f(s)] - \int_0^1 f((1-\tau)t+\tau s)\,d\tau \tag{2.1}$$

and

$$\mathcal{M}_f(t,s) := \int_0^1 f((1-\tau)t+\tau s)\,d\tau - f\left(\frac{t+s}{2}\right) \tag{2.2}$$

for all $(t,s) \in I \times I$.

We observe that

$$\mathcal{J}_f(t,s) = \mathcal{T}_f(t,s) + \mathcal{M}_f(t,s) \tag{2.3}$$

for all $(t,s) \in I \times I$.

If f is convex on I, then by *Hermite-Hadamard inequalities*

$$\frac{f(a)+f(b)}{2} \geq \int_0^1 f((1-\tau)a+\tau b)\,d\tau \geq f\left(\frac{a+b}{2}\right)$$

for all a, $b \in I$, we have the following fundamental facts

$$\mathcal{T}_f(t,s) \geq 0 \text{ and } \mathcal{M}_f(t,s) \geq 0 \tag{2.4}$$

for all $(t,s) \in I \times I$.

Using *Bullen's inequality*, see for instance [22, p. 2],

$$\begin{aligned} 0 &\leq \int_0^1 f((1-\tau)a+\tau b)\,d\tau - f\left(\frac{a+b}{2}\right) \\ &\leq \frac{f(a)+f(b)}{2} - \int_0^1 f((1-\tau)a+\tau b)\,d\tau \end{aligned}$$

we also have

$$0 \leq \mathcal{M}_f(t,s) \leq \mathcal{T}_f(t,s). \tag{2.5}$$

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*

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

f) The p-*logarithmic mean*

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

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

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

and

$$\int_0^1 \ln[(1-\tau)a+\tau b] \, d\tau = \ln I(a,b).$$

Using these notations we can define the following divergences for $(t,s) \in I^n \times I^n$ where I is an interval of positive numbers:

$$\mathcal{T}_p(t,s) := A(t^p,s^p) - L_p^p(t,s)$$

and

$$\mathcal{M}_p(t,s) := L_p^p(t,s) - A^p(t,s)$$

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

$$\mathcal{T}_{-1}(t,s) := H^{-1}(t,s) - L^{-1}(t,s)$$

and

$$\mathcal{M}_{-1}(t,s) := L^{-1}(t,s) - A^{-1}(t,s)$$

for $p=-1$ and

$$\mathcal{T}_0(t,s):=\ln\left(\frac{G(t,s)}{I(t,s)}\right)$$

and

$$\mathcal{M}_0(t,s):=\ln\left(\frac{I(t,s)}{A(t,s)}\right)$$

for $p=0$.

Since the function $f(\tau)=\tau^p$, $\tau>0$ is convex for $p\in(-\infty,0)\cup(1,\infty)$, then we have

$$\mathcal{T}_p(t,s)\,,\ \mathcal{M}_p(t,s)\geq 0 \tag{2.6}$$

for all $(t,s)\in I\times I$.

For $p\in(0,1)$ the function $f(\tau)=\tau^p$, $\tau>0$ and for $p=0$, the function $f(\tau)=\ln\tau$ are concave, then we have for $p\in[0,1)$ that

$$\mathcal{T}_p(t,s)\,,\ \mathcal{M}_p(t,s)\leq 0 \tag{2.7}$$

for all $(t,s)\in I\times I$.

Finally for $p=1$ we have both $\mathcal{T}_1(t,s)=\mathcal{M}_1(t,s)=0$ for all $(t,s)\in I\times I$.

We need the following convexity result that is a consequence of Burbea-Rao's theorem above:

Lemma 1. *Let f be a C^2 function on an interval I. Then $\mathcal{T}_f$ and $\mathcal{M}_f$ are convex (concave) on $I\times I$, if and only if f is convex (concave) and $\frac{1}{f''}$ is concave (convex) on I.*

Proof. If $\mathcal{T}_f$ and $\mathcal{M}_f$ are convex on $I\times I$ then the sum $\mathcal{T}_f+\mathcal{M}_f=\mathcal{J}_f$ is convex on $I\times I$, which, by Burbea-Rao theorem implies that f is convex and $\frac{1}{f''}$ is concave on I.

Now, if f is convex and $\frac{1}{f''}$ is concave on I, then by the same theorem we have that the function $\mathcal{J}_f:I\times I\to\mathbb{R}$

$$\mathcal{J}_f(t,s):=\frac{1}{2}[f(t)+f(s)]-f\left(\frac{t+s}{2}\right)$$

is convex.

Let $t,\ s,\ u,\ v\in I$. We define

$$\begin{aligned}\varphi(\tau)&:=\mathcal{J}_f((1-\tau)(t,s)+\tau(u,v))=\mathcal{J}_f(((1-\tau)t+\tau u,(1-\tau)s+\tau v))\\&=\frac{1}{2}[f((1-\tau)t+\tau u)+f((1-\tau)s+\tau v)]\\&-f\left(\frac{(1-\tau)t+\tau u+(1-\tau)s+\tau v}{2}\right)\\&=\frac{1}{2}[f((1-\tau)t+\tau u)+f((1-\tau)s+\tau v)]\\&-f\left((1-\tau)\frac{t+s}{2}+\tau\frac{u+v}{2}\right)\end{aligned}$$

for $\tau\in[0,1]$.

Let $\tau_1,\, \tau_2 \in [0,1]$ and $\alpha,\, \beta \geq 0$ with $\alpha + \beta = 1$. By the convexity of $\mathcal{J}_f$ we have

$$\begin{aligned}
&\varphi\left(\alpha\tau_1 + \beta\tau_2\right)\\
&= \mathcal{J}_f\left(\left(1 - \alpha\tau_1 - \beta\tau_2\right)(t,s) + \left(\alpha\tau_1 + \beta\tau_2\right)(u,v)\right)\\
&= \mathcal{J}_f\left(\left(\alpha + \beta - \alpha\tau_1 - \beta\tau_2\right)(t,s) + \left(\alpha\tau_1 + \beta\tau_2\right)(u,v)\right)\\
&= \mathcal{J}_f\left(\alpha\left(1 - \tau_1\right)(t,s) + \beta\left(1 - \tau_2\right)(t,s) + \alpha\tau_1(u,v) + \beta\tau_2(u,v)\right)\\
&= \mathcal{J}_f\left(\alpha\left[\left(1 - \tau_1\right)(t,s) + \tau_1(u,v)\right] + \beta\left[\left(1 - \tau_2\right)(t,s) + \tau_2(u,v)\right]\right)\\
&\leq \alpha\mathcal{J}_f\left(\left(1 - \tau_1\right)(t,s) + \tau_1(u,v)\right) + \beta\mathcal{J}_f\left(\left(1 - \tau_2\right)(t,s) + \tau_2(u,v)\right)\\
&= \alpha\varphi\left(\tau_1\right) + \beta\varphi\left(\tau_2\right),
\end{aligned}$$

which proves that φ is convex on $[0,1]$ for all $t,\, s,\, u,\, v \in I$.

Applying the Hermite-Hadamard inequality for φ we get

$$\frac{1}{2}\left[\varphi(0) + \varphi(1)\right] \geq \int_0^1 \varphi(\tau)\, d\tau \tag{2.8}$$

and since

$$\varphi(0) = \frac{1}{2}\left[f(t) + f(s)\right] - f\left(\frac{t+s}{2}\right),$$

$$\varphi(1) = \frac{1}{2}\left[f(u) + f(v)\right] - f\left(\frac{u+v}{2}\right)$$

and

$$\begin{aligned}
\int_0^1 \varphi(\tau)\, d\tau = &\frac{1}{2}\left[\int_0^1 f\left((1-\tau)t + \tau u\right) d\tau + \int_0^1 f\left((1-\tau)s + \tau v\right) d\tau\right]\\
&- \int_0^1 f\left((1-\tau)\frac{t+s}{2} + \tau\frac{u+v}{2}\right) d\tau,
\end{aligned}$$

hence by (2.8) we get

$$\begin{aligned}
&\frac{1}{2}\left\{\frac{1}{2}\left[f(t) + f(s)\right] - f\left(\frac{t+s}{2}\right) + \frac{1}{2}\left[f(u) + f(v)\right] - f\left(\frac{u+v}{2}\right)\right\}\\
&\geq \frac{1}{2}\left[\int_0^1 f\left((1-\tau)t + \tau u\right) d\tau + \int_0^1 f\left((1-\tau)s + \tau v\right) d\tau\right]\\
&- \int_0^1 f\left((1-\tau)\frac{t+s}{2} + \tau\frac{u+v}{2}\right) d\tau.
\end{aligned}$$

Re-arranging this inequality, we get

$$\begin{aligned}
&\frac{1}{2}\left[\frac{f(t) + f(u)}{2} - \int_0^1 f\left((1-\tau)t + \tau u\right) d\tau\right]\\
&+ \frac{1}{2}\left[\frac{f(s) + f(v)}{2} - \int_0^1 f\left((1-\tau)s + \tau v\right) d\tau\right]\\
&\geq \frac{1}{2}\left[f\left(\frac{t+s}{2}\right) + f\left(\frac{u+v}{2}\right) - \int_0^1 f\left((1-\tau)\frac{t+s}{2} + \tau\frac{u+v}{2}\right) d\tau\right],
\end{aligned}$$

which is equivalent to

$$\frac{1}{2}\left[\mathcal{T}_f(t,u)+\mathcal{T}_f(s,v)\right]\geq\mathcal{T}_f\left(\frac{t+s}{2},\frac{u+v}{2}\right)$$
$$=\mathcal{T}_f\left(\frac{1}{2}(t,u)+\frac{1}{2}(s,v)\right),$$

for all $(t,u),(s,v)\in I\times I$, which shows that $\mathcal{T}_f$ is Jensen's convex on $I\times I$. Since $\mathcal{T}_f$ is continuous on $I\times I$, hence $\mathcal{T}_f$ is convex in the usual sense on $I\times I$.

Now, if we use the second Hermite-Hadamard inequality for φ on $[0,1]$, we have

$$\int_0^1\varphi(\tau)\,d\tau\geq\varphi\left(\frac{1}{2}\right). \tag{2.9}$$

Since

$$\varphi\left(\frac{1}{2}\right)=\frac{1}{2}\left[f\left(\frac{t+u}{2}\right)+f\left(\frac{s+v}{2}\right)\right]-f\left(\frac{1}{2}\frac{t+s}{2}+\frac{1}{2}\frac{u+v}{2}\right)$$

hence by (2.9) we have

$$\frac{1}{2}\left[\int_0^1 f((1-\tau)t+\tau u)\,d\tau+\int_0^1 f((1-\tau)s+\tau v)\,d\tau\right]$$
$$-\int_0^1 f\left((1-\tau)\frac{t+s}{2}+\tau\frac{u+v}{2}\right)d\tau$$
$$\geq\frac{1}{2}\left[f\left(\frac{t+u}{2}\right)+f\left(\frac{s+v}{2}\right)\right]-f\left(\frac{1}{2}\left(\frac{t+s}{2}+\frac{u+v}{2}\right)\right),$$

which is equivalent to

$$\frac{1}{2}\left[\int_0^1 f((1-\tau)t+\tau u)\,d\tau-f\left(\frac{t+u}{2}\right)\right]$$
$$+\frac{1}{2}\left[\int_0^1 f((1-\tau)s+\tau v)\,d\tau-f\left(\frac{s+v}{2}\right)\right]$$
$$\geq\int_0^1 f\left((1-\tau)\frac{t+s}{2}+\tau\frac{u+v}{2}\right)d\tau-f\left(\frac{1}{2}\left(\frac{t+s}{2}+\frac{u+v}{2}\right)\right)$$

that can be written as

$$\frac{1}{2}\left[\mathcal{M}_f(t,u)+\mathcal{M}_f(s,v)\right]\geq\mathcal{M}_f\left(\frac{t+s}{2},\frac{u+v}{2}\right)$$
$$=\mathcal{M}_f\left(\frac{1}{2}(t,u)+\frac{1}{2}(s,v)\right)$$

for all $(t,u),(s,v)\in I\times I$, which shows that $\mathcal{M}_f$ is Jensen's convex on $I\times I$. Since $\mathcal{M}_f$ is continuous on $I\times I$, hence $\mathcal{M}_f$ is convex in the usual sense on $I\times I$. $\square$

The following reverses of the Hermite-Hadamard inequality hold:

Lemma 2 (Dragomir, 2002 [10] and [11])**.** *Let $h:[a,b]\to\mathbb{R}$ be a convex function on $[a,b]$. Then*

$$0\leq\frac{1}{8}\left[h_{+}\left(\frac{a+b}{2}\right)-h_{-}\left(\frac{a+b}{2}\right)\right](b-a)\tag{2.10}$$
$$\leq\frac{h(a)+h(b)}{2}-\frac{1}{b-a}\int_a^b h(\tau)\,d\tau$$
$$\leq\frac{1}{8}\left[h_{-}(b)-h_{+}(a)\right](b-a)$$

and

$$0\leq\frac{1}{8}\left[h_{+}\left(\frac{a+b}{2}\right)-h_{-}\left(\frac{a+b}{2}\right)\right](b-a)\tag{2.11}$$
$$\leq\frac{1}{b-a}\int_a^b h(\tau)\,d\tau-h\left(\frac{a+b}{2}\right)$$
$$\leq\frac{1}{8}\left[h_{-}(b)-h_{+}(a)\right](b-a).$$

The constant $\frac{1}{8}$ is best possible in all inequalities from (2.10) and (2.11).

We also have:

Lemma 3. *Let f be a C^1 convex function on an interval I. If $\mathring{I}$ is the interior of I, then for all $(t,s)\in\mathring{I}\times\mathring{I}$ we have*

$$0\leq\mathcal{M}_f(t,s)\leq\mathcal{T}_f(t,s)\leq\frac{1}{8}\mathcal{C}_{f'}(t,s)\tag{2.12}$$

where

$$\mathcal{C}_{f'}(t,s):=\left[f'(t)-f'(s)\right](t-s).\tag{2.13}$$

Proof. Since for $b\neq a$

$$\frac{1}{b-a}\int_a^b f(t)\,dt=\int_0^1 f((1-t)a+tb)\,dt,$$

then from (2.10) we get

$$\frac{f(t)+f(s)}{2}-\int_0^1 f((1-\tau)t+\tau s)\,dt\leq\frac{1}{8}\left[f'(t)-f'(s)\right](t-s)$$

for all $(t,s)\in\mathring{I}\times\mathring{I}$. □

Remark 2. *If*

$$\gamma=\inf_{t\in\mathring{I}}f'(t)\ \text{ and }\Gamma=\sup_{t\in\mathring{I}}f'(t)$$

are finite, then

$$\mathcal{C}_{f'}(t,s)\leq(\Gamma-\gamma)|t-s|$$

and by (2.12) we get the simpler upper bound

$$0\leq\mathcal{M}_f(t,s)\leq\mathcal{T}_f(t,s)\leq\frac{1}{8}(\Gamma-\gamma)|t-s|.$$

Moreover, if $t,\ s\in[a,b]\subset\mathring{I}$ and since f' is increasing on $\mathring{I}$, then we have the inequalities

$$0\leq\mathcal{M}_f(t,s)\leq\mathcal{T}_f(t,s)\leq\frac{1}{8}\left[f'(b)-f'(a)\right]|t-s|.\tag{2.14}$$

Since $\mathcal{J}_f(t,s) = \mathcal{T}_f(t,s) + \mathcal{M}_f(t,s)$, *hence*

$$0 \le \mathcal{J}_f(t,s) \le \frac{1}{4}\left[f'(b) - f'(a)\right]|t-s|.$$

Corollary 1. *With the assumptions of Lemma 3 and if the derivative* f' *is Lipschitzian with the constant* $K > 0$, *namely*

$$|f'(t) - f'(s)| \le K|t-s| \text{ for all } t, s \in \mathring{I},$$

then we have the inequality

$$0 \le \mathcal{M}_f(t,s) \le \mathcal{T}_f(t,s) \le \frac{1}{8}K(t-s)^2. \tag{2.15}$$

3. Main Results

Let $P, Q, W \in \mathcal{P}$ and $f : (0,\infty) \to \mathbb{R}$. We define the following f-divergence

$$\begin{aligned}\mathcal{J}_f(P,Q,W) &:= \int_X w(x)\,\mathcal{J}_f\left(\frac{p(x)}{w(x)}, \frac{q(x)}{w(x)}\right) d\mu(x) \\ &= \frac{1}{2}\left[\int_X w(x) f\left(\frac{p(x)}{w(x)}\right) d\mu(x) + \int_X w(x) f\left(\frac{q(x)}{w(x)}\right) d\mu(x)\right] \\ &\qquad - \int_X w(x) f\left(\frac{p(x)+q(x)}{2w(x)}\right).\end{aligned}\tag{3.1}$$

If we consider the *mid-point divergence measure* M_f defined by

$$M_f(Q,P,W) := \int_X f\left[\frac{q(x)+p(x)}{2w(x)}\right] w(x)\,d\mu(x)$$

for any $Q, P, W \in \mathcal{P}$, then from (3.1) we get

$$\mathcal{J}_f(P,Q,W) = \frac{1}{2}\left[I_f(P,W) + I_f(Q,W)\right] - M_f(Q,P,W). \tag{3.2}$$

We can also consider the *integral divergence measure*

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

We introduce the related f-divergences

$$\begin{aligned}\mathcal{T}_f(P,Q,W) &:= \int_X w(x)\,\mathcal{T}_f\left(\frac{p(x)}{w(x)}, \frac{q(x)}{w(x)}\right) d\mu(x) \\ &= \frac{1}{2}\left[I_f(P,W) + I_f(Q,W)\right] - A_f(Q,P,W)\end{aligned}\tag{3.3}$$

and

$$\begin{aligned}\mathcal{M}_f(P,Q,W) &:= \int_X w(x)\,\mathcal{M}_f\left(\frac{p(x)}{w(x)}, \frac{q(x)}{w(x)}\right) d\mu(x) \\ &= A_f(Q,P,W) - M_f(Q,P,W).\end{aligned}\tag{3.4}$$

We observe that

$$\mathcal{J}_f(P,Q,W) = \mathcal{T}_f(P,Q,W) + \mathcal{M}_f(P,Q,W).$$

If f is convex on $(0,\infty)$ then by the Hermite-Hadamard and Bullen's inequalities we have the positivity properties

$$0 \le \mathcal{M}_f(P,Q,W) \le \mathcal{T}_f(P,Q,W)$$

and

$$0 \leq \mathcal{J}_f(P, Q, W)$$

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

We have the following result:

Theorem 5. *Let f be a C^2 function on an interval $(0, \infty)$. If f is convex on $(0, \infty)$ and $\frac{1}{f''}$ is concave on $(0, \infty)$, then for all $W \in \mathcal{P}$, the mappings*

$$\mathcal{P} \times \mathcal{P} \ni (P, Q) \mapsto \mathcal{J}_f(P, Q, W), \ \mathcal{M}_f(P, Q, W), \ \mathcal{T}_f(P, Q, W)$$

are convex.

Proof. Let $(P_1, Q_1), (P_2, Q_2) \in \mathcal{P} \times \mathcal{P}$ and α, $\beta \geq 0$ with $\alpha + \beta = 1$. We have

$$\begin{aligned}
&\mathcal{J}_f(\alpha(P_1, Q_1, W) + \beta(P_2, Q_2, W)) = \mathcal{J}_f(\alpha P_1 + \beta P_2, \alpha Q_1 + \beta Q_2, W) \\
&\quad = \int_X w(x)\, \mathcal{J}_f\left(\frac{\alpha p_1(x) + \beta p_2(x)}{w(x)}, \frac{\alpha q_1(x) + \beta q_2(x)}{w(x)}\right) d\mu(x) \\
&\quad = \int_X w(x)\, \mathcal{J}_f\left(\alpha\frac{p_1(x)}{w(x)} + \beta\frac{p_2(x)}{w(x)}, \alpha\frac{q_1(x)}{w(x)} + \beta\frac{q_2(x)}{w(x)}\right) d\mu(x) \\
&\quad = \int_X w(x)\, \mathcal{J}_f\left(\alpha\left(\frac{p_1(x)}{w(x)}, \frac{q_1(x)}{w(x)}\right) + \beta\left(\frac{p_2(x)}{w(x)}, \frac{q_2(x)}{w(x)}\right)\right) d\mu(x) \\
&\quad \leq \int_X w(x)\left[\alpha\mathcal{J}_f\left(\frac{p_1(x)}{w(x)}, \frac{q_1(x)}{w(x)}\right) + \beta\mathcal{J}_f\left(\frac{p_2(x)}{w(x)}, \frac{q_2(x)}{w(x)}\right)\right] d\mu(x) \\
&= \alpha\int_X w(x)\, \mathcal{J}_f\left(\frac{p_1(x)}{w(x)}, \frac{q_1(x)}{w(x)}\right) d\mu(x) + \beta\int_X w(x)\, \mathcal{J}_f\left(\frac{p_2(x)}{w(x)}, \frac{q_2(x)}{w(x)}\right) d\mu(x) \\
&\qquad = \alpha\mathcal{J}_f(P_1, Q_1, W) + \beta\mathcal{J}_f(P_2, Q_2, W),
\end{aligned}$$

which proves the convexity of $\mathcal{P} \times \mathcal{P} \ni (P, Q) \mapsto \mathcal{J}_f(P, Q, W)$ for all $W \in \mathcal{P}$.

The convexity of the other two mappings follows in a similar way and we omit the details. □

Theorem 6. *Let f be a C^1 function on an interval $(0, \infty)$. If f is convex on $(0, \infty)$, then for all $W \in \mathcal{P}$*

$$0 \leq \mathcal{M}_f(P, Q, W) \leq \mathcal{T}_f(P, Q, W) \leq \frac{1}{8}\Delta_{f'}(Q, P, W) \tag{3.5}$$

where

$$\Delta_{f'}(Q, P, W) := \int_X \left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{p(x)}{w(x)}\right)\right](q(x) - p(x))\, d\mu(x). \tag{3.6}$$

Proof. From the inequality (2.12) we have

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

for all $x \in X$.

If we multiply by $w(x) > 0$ and integrate on X we get

$$\begin{aligned}
&\frac{1}{2}\left[I_f(P,W) + I_f(P,W)\right] - A_f(Q,P,W) \\
&\leq \frac{1}{8}\int_X w(x)\left(f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)}{w(x)}\right)\right)\left(\frac{p(x)}{w(x)} - \frac{q(x)}{w(x)}\right)d\mu(x) \\
&= \frac{1}{8}\int_X \left(f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)}{w(x)}\right)\right)(p(x) - q(x))\, d\mu(x),
\end{aligned}$$

which implies the desired inequality. □

Corollary 2. *With the assumptions of Theorem 6 and if f' is Lipschitzian with the constant $K > 0$, namely*

$$|f'(s) - f'(t)| \leq K|s - t| \text{ for all } t,\ s \in (0,\infty),$$

then

$$0 \leq \mathcal{M}_f(P,Q,W) \leq \mathcal{T}_f(P,Q,W) \leq \frac{1}{8}K d_{\chi^2}(Q,P,W), \tag{3.7}$$

where

$$d_{\chi^2}(Q,P,W) := \int_X \frac{(q(x) - p(x))^2}{w(x)}d\mu(x). \tag{3.8}$$

Remark 3. *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

$$0 \leq \mathcal{M}_f(P,Q,W) \leq \mathcal{T}_f(P,Q,W) \leq \frac{1}{8}\left[f'(R) - f'(r)\right]d_1(Q,P) \tag{3.9}$$

where

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

Moreover, if f is twice differentiable and

$$\|f''\|_{[r,R],\infty} := \sup_{t\in[r,R]}|f''(t)| < \infty \tag{3.10}$$

then

$$0 \leq \mathcal{M}_f(P,Q,W) \leq \mathcal{T}_f(P,Q,W) \leq \frac{1}{8}\|f''\|_{[r,R],\infty}\, d_{\chi^2}(Q,P,W). \tag{3.11}$$

We also have:

Theorem 7. *Let f be a C^2 function on an interval $(0,\infty)$. If f is convex on $(0,\infty)$ and $\frac{1}{f''}$ is concave on $(0,\infty)$, then for all $W \in \mathcal{P}$,*

$$0 \leq \mathcal{J}_f(P,Q,W) \leq \frac{1}{2}\left[\Psi_{f'}(P,Q,W) + \Psi_{f'}(Q,P,W)\right], \tag{3.12}$$

where

$$\begin{aligned}
&\Psi_{f'}(P,Q,W) \\
&:= \int_X \left[f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right](p(x) - w(x))\, d\mu(x).
\end{aligned}$$

Proof. It is well known that if the function of two independent variables $F : D \subset \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is convex on the convex domain D and has partial derivatives $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$ on D then for all $(t, s), (u, v) \in D$ we have the gradient inequalities

$$\begin{aligned} &\frac{\partial F(t,s)}{\partial x}(t-u) + \frac{\partial F(t,s)}{\partial y}(s-v) \\ &\geq F(t,s) - F(u,v) \\ &\geq \frac{\partial F(u,v)}{\partial x}(t-u) + \frac{\partial F(u,v)}{\partial y}(s-v). \end{aligned} \tag{3.13}$$

Now, if we take $F : (0, \infty) \times (0, \infty) \to \mathbb{R}$ given by

$$F(t,s) = \frac{1}{2}[f(t) + f(s)] - f\left(\frac{t+s}{2}\right)$$

and observe that

$$\frac{\partial F(t,s)}{\partial x} = \frac{1}{2}\left[f'(t) - f'\left(\frac{t+s}{2}\right)\right]$$

and

$$\frac{\partial F(t,s)}{\partial y} = \frac{1}{2}\left[f'(s) - f'\left(\frac{t+s}{2}\right)\right]$$

and since F is convex on $(0, \infty) \times (0, \infty)$, then by (3.13) we get

$$\begin{aligned} &\frac{1}{2}\left[f'(t) - f'\left(\frac{t+s}{2}\right)\right](t-u) + \frac{1}{2}\left[f'(s) - f'\left(\frac{t+s}{2}\right)\right](s-v) \\ &\geq \frac{1}{2}[f(t) + f(s)] - f\left(\frac{t+s}{2}\right) - \frac{1}{2}[f(u) + f(v)] + f\left(\frac{u+v}{2}\right) \\ &\geq \frac{1}{2}\left[f'(u) - f'\left(\frac{u+v}{2}\right)\right](t-u) + \frac{1}{2}\left[f'(v) - f'\left(\frac{u+v}{2}\right)\right](s-v). \end{aligned} \tag{3.14}$$

If we take $u = v = 1$ in (3.14), then we have

$$\begin{aligned} &\frac{1}{2}\left[f'(t) - f'\left(\frac{t+s}{2}\right)\right](t-1) + \frac{1}{2}\left[f'(s) - f'\left(\frac{t+s}{2}\right)\right](s-1) \\ &\geq \frac{1}{2}[f(t) + f(s)] - f\left(\frac{t+s}{2}\right) \geq 0 \end{aligned} \tag{3.15}$$

for all $(t, s) \in (0, \infty) \times (0, \infty)$.

If we take $t = \frac{p(x)}{w(x)}$ and $s = \frac{q(x)}{w(x)}$ in (3.15) then we obtain

$$\begin{aligned} &\frac{1}{2}\left[f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right]\left(\frac{p(x)}{w(x)} - 1\right) \\ &+ \frac{1}{2}\left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right]\left(\frac{q(x)}{w(x)} - 1\right) \\ &\geq \frac{1}{2}\left[f\left(\frac{p(x)}{w(x)}\right) + f\left(\frac{q(x)}{w(x)}\right)\right] - f\left(\frac{q(x)+p(x)}{2w(x)}\right) \geq 0. \end{aligned} \tag{3.16}$$

By multiplying this inequality with $w(x) > 0$ we get

$$\begin{aligned} 0 &\leq \frac{1}{2}\left[w(x) f\left(\frac{p(x)}{w(x)}\right) + w(x) f\left(\frac{q(x)}{w(x)}\right)\right] - w(x) f\left(\frac{q(x)+p(x)}{2w(x)}\right) \\ &\leq \frac{1}{2}\left[f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right](p(x) - w(x)) \\ &+ \frac{1}{2}\left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right](q(x) - w(x)) \end{aligned}$$

for all $x \in X$. □

Corollary 3. *With the assumptions of Theorem 6 and if f' is Lipschitzian with the constant $K > 0$, then*

$$\begin{aligned} (3.17)\quad 0 &\leq \mathcal{J}_f(P, Q, W) \\ &\leq \frac{1}{4} K \int_X |p(x) - q(x)| \left[\left|\frac{p(x)}{w(x)} - 1\right| + \left|\frac{q(x)}{w(x)} - 1\right|\right] d\mu(x). \end{aligned}$$

Proof. We have that

$$\begin{aligned} &\Psi_{f'}(P, Q, W) \\ &\leq \int_X \left|f'\left(\frac{p(x)}{w(x)}\right) - f'\left(\frac{q(x)+p(x)}{2w(x)}\right)\right| |p(x) - w(x)|\, d\mu(x) \\ &\leq K \int_X \left|\frac{p(x)}{w(x)} - \frac{q(x)+p(x)}{2w(x)}\right| |p(x) - w(x)|\, d\mu(x) \\ &= K \int_X \left|\frac{p(x) - q(x)}{2w(x)}\right| |p(x) - w(x)|\, d\mu(x) \\ &= \frac{1}{2} K \int_X \frac{|p(x) - q(x)|\, |p(x) - w(x)|\, d\mu(x)}{w(x)} \\ &= \frac{1}{2} K \int_X |p(x) - q(x)| \left|\frac{p(x)}{w(x)} - 1\right| d\mu(x) \end{aligned}$$

and similarly

$$\Psi_{f'}(P, Q, W) \leq \frac{1}{2} K \int_X |p(x) - q(x)| \left|\frac{q(x)}{w(x)} - 1\right| d\mu(x).$$

Finally, by the use of (3.12) we get the desired result. □

Remark 4. *If there exist $0 < r < 1 < R < \infty$ such that the following condition (r, R) holds and if f is twice differentiable and (3.10) is valid, then*

$$\begin{aligned} (3.18)\quad 0 \leq \mathcal{J}_f(P, Q, W) &\leq \frac{1}{4} \|f''\|_{[r,R],\infty} \\ &\times \int_X |p(x) - q(x)| \left[\left|\frac{p(x)}{w(x)} - 1\right| + \left|\frac{q(x)}{w(x)} - 1\right|\right] d\mu(x). \end{aligned}$$

Since

$$\left|\frac{p(x)}{w(x)} - 1\right|, \left|\frac{q(x)}{w(x)} - 1\right| \leq \max\{R - 1, 1 - r\}$$

and

$$\left|\frac{p(x)}{w(x)} - \frac{q(x)}{w(x)}\right| \leq R - r,$$

hence by (3.18) we get the simpler bound

$$0 \leq \mathcal{J}_f(P,Q,W) \leq \frac{1}{2}\|f''\|_{[r,R],\infty}(R-r)\max\{R-1,1-r\}. \tag{3.19}$$

We also have:

Theorem 8. *With the assumptions of Theorem 6 and if f' is Lipschitzian with the constant $K>0$, then*

$$\begin{aligned} 0 &\leq \mathcal{T}_f(P,Q,W) \\ &\leq \frac{1}{6}K\int_X |p(x)-q(x)|\left[\left|\frac{p(x)}{w(x)}-1\right|+\left|\frac{q(x)}{w(x)}-1\right|\right]d\mu(x). \end{aligned} \tag{3.20}$$

Proof. Let $(x,y), (u,v) \in (0,\infty)\times(0,\infty)$. If we take $F:(0,\infty)\times(0,\infty)\to\mathbb{R}$ given by

$$F(t,s)=\frac{f(t)+f(s)}{2}-\int_0^1 f((1-\tau)t+\tau s)\,d\tau$$

then

$$\begin{aligned} \frac{\partial F(t,s)}{\partial x} &= \frac{1}{2}f'(t)-\int_0^1(1-\tau)f'((1-\tau)t+\tau s)\,d\tau \\ &= \int_0^1(1-\tau)\left[f'(t)-f'((1-\tau)t+\tau s)\right]d\tau \end{aligned}$$

and

$$\begin{aligned} \frac{\partial F(t,s)}{\partial y} &= \frac{1}{2}f'(s)-\int_0^1 \tau f'((1-\tau)t+\tau s)\,d\tau \\ &= \int_0^1 \tau\left[f'(s)-f'((1-\tau)t+\tau s)\right]d\tau \end{aligned}$$

and since F is convex on $(0,\infty)\times(0,\infty)$, then by (3.1) we get

$$\begin{aligned} &(t-u)\int_0^1(1-\tau)\left[f'(t)-f'((1-\tau)t+\tau s)\right]d\tau \\ &+(s-v)\int_0^1\tau\left[f'(s)-f'((1-\tau)t+\tau s)\right]d\tau \\ &\geq \frac{f(t)+f(s)}{2}-\int_0^1 f((1-\tau)t+\tau s)\,d\tau \\ &-\frac{f(u)+f(v)}{2}+\int_0^1 f((1-\tau)u+\tau v)\,d\tau \\ &\geq (t-u)\int_0^1(1-\tau)\left[f'(u)-f'((1-\tau)u+\tau v)\right]d\tau \\ &+(s-v)\int_0^1\tau\left[f'(v)-f'((1-\tau)u+\tau v)\right]d\tau \end{aligned} \tag{3.21}$$

for all $(t,s), (u,v) \in (0,\infty)\times(0,\infty)$.

If we take $u = v = 1$ in (3.21), then we have

$$\begin{aligned}(t-1)\int_0^1 (1-\tau)\left[f'(t) - f'((1-\tau)t + \tau s)\right] d\tau \\ + (s-1)\int_0^1 \tau\left[f'(s) - f'((1-\tau)t + \tau s)\right] d\tau \\ \geq \frac{f(t)+f(s)}{2} - \int_0^1 f((1-\tau)t+\tau s)\, d\tau \geq 0\end{aligned} \tag{3.22}$$

for all $(u,v) \in (0,\infty)\times(0,\infty)$.

If we take $t = \frac{p(x)}{w(x)}$ and $s = \frac{q(x)}{w(x)}$ in (3.22) then we get

$$\begin{aligned}\left(\frac{p(x)}{w(x)} - 1\right)\int_0^1 (1-\tau)\left[f'\left(\frac{p(x)}{w(x)}\right) - f'\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right)\right] d\tau \\ + \left(\frac{q(x)}{w(x)} - 1\right)\int_0^1 \tau\left[f'\left(\frac{q(x)}{w(x)}\right) - f'\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right)\right] d\tau \\ \geq \frac{f\left(\frac{p(x)}{w(x)}\right) + f\left(\frac{q(x)}{w(x)}\right)}{2} - \int_0^1 f\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right) d\tau \geq 0.\end{aligned} \tag{3.23}$$

Since f' is Lipschitzian with the constant $K > 0$, hence

$$\begin{aligned}0 &\leq \frac{f\left(\frac{p(x)}{w(x)}\right) + f\left(\frac{q(x)}{w(x)}\right)}{2} - \int_0^1 f\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right) d\tau \\ &\leq \left|\frac{p(x)}{w(x)} - 1\right|\int_0^1 (1-\tau)\left|f'\left(\frac{p(x)}{w(x)}\right) - f'\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right)\right| d\tau \\ &+ \left|\frac{q(x)}{w(x)} - 1\right|\int_0^1 \tau\left|f'\left(\frac{q(x)}{w(x)}\right) - f'\left((1-\tau)\frac{p(x)}{w(x)} + \tau\frac{q(x)}{w(x)}\right)\right| d\tau \\ &\leq K\left|\frac{p(x)}{w(x)} - 1\right|\left|\frac{p(x)}{w(x)} - \frac{q(x)}{w(x)}\right|\int_0^1 (1-\tau)\tau d\tau \\ &+ K\left|\frac{q(x)}{w(x)} - 1\right|\left|\frac{p(x)}{w(x)} - \frac{q(x)}{w(x)}\right|\int_0^1 (1-\tau)\tau d\tau \\ &= \frac{1}{6}K\left|\frac{p(x)}{w(x)} - \frac{q(x)}{w(x)}\right|\left[\left|\frac{p(x)}{w(x)} - 1\right| + \left|\frac{q(x)}{w(x)} - 1\right|\right].\end{aligned}$$

If we multiply this inequality by $w(x) > 0$ and integrate, then we get the desired result (3.20). □

Corollary 4. *If there exist $0 < r < 1 < R < \infty$ such that the condition (r,R) holds and if f is twice differentiable and (3.10) is valid, then*

$$0 \leq \mathcal{T}_f(P,Q,W) \leq \frac{1}{3}\|f''\|_{[r,R],\infty}(R-r)\max\{R-1, 1-r\}. \tag{3.24}$$

Finally, we also have:

Theorem 9. *With the assumptions of Theorem 6 and if f' is Lipschitzian with the constant $K>0$, then*

$$\begin{aligned}(3.25)\qquad & 0\leq \mathcal{M}_{f}\left(P,Q,W\right) \\ & \leq \frac{1}{8}K\int_{X}\left\vert p\left(x\right)-q\left(x\right)\right\vert \left[\left\vert \frac{p\left(x\right)}{w\left(x\right)}-1\right\vert +\left\vert \frac{q\left(x\right)}{w\left(x\right)}-1\right\vert \right] d\mu\left(x\right).\end{aligned}$$

Proof. Let $(t,s),\ (u,v)\in(0,\infty)\times(0,\infty)$. If we take $F:(0,\infty)\times(0,\infty)\to\mathbb{R}$ given by

$$F\left(t,s\right)=\int_{0}^{1}f\left(\left(1-\tau\right)t+\tau s\right)d\tau-f\left(\frac{t+s}{2}\right)$$

then

$$\begin{aligned}\frac{\partial F\left(t,s\right)}{\partial x} &=\int_{0}^{1}\left(1-\tau\right)f'\left(\left(1-\tau\right)t+\tau s\right)d\tau-\frac{1}{2}f'\left(\frac{t+s}{2}\right) \\ &=\int_{0}^{1}\left(1-\tau\right)\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau,\end{aligned}$$

$$\begin{aligned}\frac{\partial F\left(t,s\right)}{\partial y} &=\int_{0}^{1}\tau f'\left(\left(1-\tau\right)t+\tau s\right)d\tau-\frac{1}{2}f'\left(\frac{t+s}{2}\right) \\ &=\int_{0}^{1}\tau\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau\end{aligned}$$

and since F is convex on $(0,\infty)\times(0,\infty)$, then by (3.1) we get

$$\begin{aligned}(3.26)\qquad & \left(t-u\right)\left[\int_{0}^{1}\left(1-\tau\right)\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau\right] \\ & +\left(s-v\right)\left[\int_{0}^{1}\tau\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau\right] \\ & \geq\int_{0}^{1}f\left(\left(1-\tau\right)t+\tau s\right)d\tau-f\left(\frac{t+s}{2}\right) \\ & -\int_{0}^{1}f\left(\left(1-\tau\right)u+\tau v\right)d\tau+f\left(\frac{u+v}{2}\right) \\ & \geq\left(t-u\right)\left[\int_{0}^{1}\left(1-\tau\right)\left[f'\left(\left(1-\tau\right)u+\tau v\right)-f'\left(\frac{u+v}{2}\right)\right]d\tau\right] \\ & +\left(s-v\right)\int_{0}^{1}\tau\left[f'\left(\left(1-\tau\right)u+\tau v\right)-f'\left(\frac{u+v}{2}\right)\right]d\tau.\end{aligned}$$

If we take $u=v=1$ in (3.26), then we have

$$\begin{aligned}(3.27)\qquad & \left(t-1\right)\left[\int_{0}^{1}\left(1-\tau\right)\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau\right] \\ & +\left(s-1\right)\left[\int_{0}^{1}\tau\left[f'\left(\left(1-\tau\right)t+\tau s\right)-f'\left(\frac{t+s}{2}\right)\right]d\tau\right] \\ & \geq\int_{0}^{1}f\left(\left(1-\tau\right)t+\tau s\right)d\tau-f\left(\frac{t+s}{2}\right)\geq 0\end{aligned}$$

for all $(t,s)\in(0,\infty)\times(0,\infty)$.

If we take $t=\frac{p(x)}{w(x)}$ and $s=\frac{q(x)}{w(x)}$ in (3.27) then we get

$$
\begin{aligned}
0 &\leq \int_0^1 f\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)d\tau - f\left(\frac{p(x)+q(x)}{2w(x)}\right) \\
&\leq \left(\frac{p(x)}{w(x)}-1\right) \\
&\quad\times\left[\int_0^1 (1-\tau)\left[f'\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)+q(x)}{2w(x)}\right)\right]d\tau\right] \\
&\quad+\left(\frac{q(x)}{w(x)}-1\right) \\
&\quad\times\left[\int_0^1 \tau\left[f'\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)+q(x)}{2w(x)}\right)\right]d\tau\right]
\end{aligned} \tag{3.28}
$$

$$
\begin{aligned}
&\leq \left|\frac{p(x)}{w(x)}-1\right| \\
&\quad\times\left[\int_0^1 (1-\tau)\left|f'\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)+q(x)}{2w(x)}\right)\right|d\tau\right] \\
&\quad+\left|\frac{q(x)}{w(x)}-1\right| \\
&\quad\times\left[\int_0^1 \tau\left|f'\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)-f'\left(\frac{p(x)+q(x)}{2w(x)}\right)\right|d\tau\right] \\
&\leq K\left|\frac{p(x)}{w(x)}-1\right|\left|\frac{p(x)}{w(x)}-\frac{q(x)}{w(x)}\right|\int_0^1(1-\tau)\left|\tau-\frac{1}{2}\right|d\tau \\
&\quad+K\left|\frac{q(x)}{w(x)}-1\right|\left|\frac{p(x)}{w(x)}-\frac{q(x)}{w(x)}\right|\int_0^1(1-\tau)\left|\tau-\frac{1}{2}\right|d\tau.
\end{aligned}
$$

Since

$$\int_0^1(1-\tau)\left|\tau-\frac{1}{2}\right|d\tau=\frac{1}{8},$$

hence

$$
\begin{aligned}
0 &\leq \int_0^1 f\left((1-\tau)\frac{p(x)}{w(x)}+\tau\frac{q(x)}{w(x)}\right)d\tau - f\left(\frac{p(x)+q(x)}{2w(x)}\right) \\
&\leq \frac{1}{8}K\left|\frac{p(x)}{w(x)}-\frac{q(x)}{w(x)}\right|\left[\left|\frac{p(x)}{w(x)}-1\right|+\left|\frac{q(x)}{w(x)}-1\right|\right]
\end{aligned}
$$

for all $x\in X$.

If we multiply this inequality by $w(x)>0$ and integrate, then we get the desired result (3.20). □

Corollary 5. *If there exist $0<r<1<R<\infty$ such that the condition (r,R) holds and if f is twice differentiable and (3.10) is valid, then*

$$0\leq\mathcal{M}_f(P,Q,W)\leq\frac{1}{4}\|f''\|_{[r,R],\infty}(R-r)\max\{R-1,1-r\}. \tag{3.29}$$

4. Some Examples

The Dichotomy Class of f-divergences are generated by the functions $f_\alpha : [0,\infty) \to \mathbb{R}$ defined as

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

Observe that

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

In this family of functions only the functions f_α with $\alpha \in [1,2)$ are both convex and with $\frac{1}{f''_\alpha}$ concave on $(0,\infty)$.

We have

$$\begin{aligned} I_{f_\alpha}(P,W) &= \int_X w(x) f_\alpha\left(\frac{p(x)}{w(x)}\right) d\mu(x) \\ &= \begin{cases} \frac{1}{\alpha(\alpha-1)}\left[\int_X w^{1-\alpha}(x)\, p^\alpha(x)\, d\mu(x) - 1\right], \ \alpha \in (1,2), \\ \int_X p(x) \ln\left(\frac{p(x)}{w(x)}\right) d\mu(x), \ \alpha = 1, \end{cases} \end{aligned}$$

and

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

We also have

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

Therefore

$$\begin{aligned} A_{f_\alpha}(Q,P,W) &:= \int_X \left(\int_0^1 f\left[\frac{(1-t)q(x) + tp(x)}{w(x)}\right] dt\right) w(x)\, d\mu(x) \\ &= \begin{cases} \frac{1}{\alpha(\alpha-1)}\left[\int_X L_\alpha^\alpha\left(\frac{q(x)}{w(x)}, \frac{p(x)}{w(x)}\right) w(x)\, d\mu(x) - 1\right], \ \alpha \in (1,2) \\ \frac{1}{2}\int_X A\left(\frac{q(x)}{w(x)}, \frac{p(x)}{w(x)}\right) \ln I\left(\left(\frac{q(x)}{w(x)}\right)^2, \left(\frac{p(x)}{w(x)}\right)^2\right) w(x)\, d\mu(x), \ \alpha = 1. \end{cases} \end{aligned}$$

We have

$$\mathcal{J}_{f_\alpha}(P,Q,W)=\frac{1}{2}\left[I_{f_\alpha}(P,W)+I_{f_\alpha}(Q,W)\right]-M_{f_\alpha}(Q,P,W),$$

$$\mathcal{T}_{f_\alpha}(P,Q,W)=\frac{1}{2}\left[I_{f_\alpha}(P,W)+I_{f_\alpha}(Q,W)\right]-A_{f_\alpha}(Q,P,W)$$

and

$$\mathcal{M}_{f_\alpha}(P,Q,W)=A_{f_\alpha}(Q,P,W)-M_{f_\alpha}(Q,P,W).$$

According to Theorem 5, for all $\alpha\in[1,2)$, the mappings

$$\mathcal{P}\times\mathcal{P}\ni(P,Q)\mapsto\mathcal{J}_{f_\alpha}(P,Q,W),\ \mathcal{M}_{f_\alpha}(P,Q,W),\ \mathcal{T}_{f_\alpha}(P,Q,W)$$

are *convex* for all $W\in\mathcal{P}$.

If $0<r<1<R$, then

$$\|f_\alpha''\|_{[r,R],\infty}=\sup_{t\in[r,R]}f_\alpha''(t)=\frac{1}{r^{2-\alpha}}\text{ for }\alpha\in[1,2).$$

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 by (3.19), (3.24) and (3.29) we get

$$0\leq\mathcal{J}_{f_\alpha}(P,Q,W)\leq\frac{1}{2}\|f''\|_{[r,R],\infty}(R-r)\max\{R-1,1-r\},\tag{4.1}$$

$$0\leq\mathcal{T}_{f_\alpha}(P,Q,W)\leq\frac{1}{3}\frac{(R-r)}{r^{2-\alpha}}\max\{R-1,1-r\}\tag{4.2}$$

and

$$0\leq\mathcal{M}_{f_\alpha}(P,Q,W)\leq\frac{1}{4}\frac{(R-r)}{r^{2-\alpha}}\max\{R-1,1-r\},\tag{4.3}$$

for all $\alpha\in[1,2)$ and $W\in\mathcal{P}$.

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] J. Burbea and C. R. Rao, On the convexity of some divergence measures based on entropy functions, *IEEE Tran. Inf. Theor.*, Vol. IT-**28**, No. 3, 1982, 489-495.

[2] 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.

[3] 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].

[4] 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.

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

[6] I. Csiszár, On topological properties of f-divergences, *Studia Math. Hungarica,* **2** (1967), 329-339.

[7] 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.

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

[9] 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.
[10] 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.
[11] 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.
[12] 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
[13] S. S. Dragomir, Bounds for f-divergences under likelihood ratio constraints. *Appl. Math.* **48** (2003), no. 3, 205–223.
[14] S. S. Dragomir, New inequalities for Csiszár divergence and applications. *Acta Math. Vietnam.* **28** (2003), no. 2, 123–134.
[15] S. S. Dragomir, A generalized f-divergence for probability vectors and applications. *Panamer. Math. J.* **13** (2003), no. 4, 61–69.
[16] 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.
[17] S. S. Dragomir, A converse inequality for the Csiszár Φ-divergence. *Tamsui Oxf. J. Math. Sci.* **20** (2004), no. 1, 35–53.
[18] S. S. Dragomir, Some general divergence measures for probability distributions. *Acta Math. Hungar.* **109** (2005), no. 4, 331–345.
[19] S. S. Dragomir, Bounds for the normalized Jensen functional, *Bull. Austral. Math. Soc.* **74**(3)(2006), 471-476.
[20] S. S. Dragomir, A refinement of Jensen's inequality with applications for f-divergence measures. *Taiwanese J. Math.* **14** (2010), no. 1, 153–164.
[21] 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.
[22] S. S. Dragomir and C. E. M. Pearce, *Selected Topics on Hermite-Hadamard Inequalities and Applications*, RGMIA Monographs, 2000. [Online https://rgmia.org/monographs/hermite_hadamard.html].
[23] S. Kullback and R. A. Leibler, On information and sufficiency, *Ann. Math. Stat.*, **22** (1951), 79-86.
[24] J. Lin and S. K. M. Wong, A new directed divergence measure and its characterization, *Int. J. General Systems,* **17** (1990), 73-81.
[25] H. Shioya and T. Da-Te, A generalisation of Lin divergence and the derivative of a new information divergence, *Elec. and Comm. in Japan*, **78** (7) (1995), 37-40.
[26] P. Kafka, F. Österreicher and I. Vincze, On powers of f-divergence defining a distance, *Studia Sci. Math. Hungar.,* **26** (1991), 415-422.
[27] F. Liese and I. Vajda, *Convex Statistical Distances,* Teubuer – Texte zur Mathematik, Band **95**, Leipzig, 1987.
[28] 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