BOUNDS FOR THE HH f-DIVERGENCE MEASURES IN TERMS OF χ^2-DIVERGENCE

SILVESTRU SEVER DRAGOMIR[1,2]

Abstract. In this paper we establish some inequalities for the Hermite-Hadamard (HH) f-divergence measures in terms of χ^2-divergence. An application for Kullback-Leibler divergence is also provided.

1. Introduction

Let the set X and the σ-finite measure μ be given and consider the set of all probability densities on μ to be defined on $\Omega := \left\{p | p : X \to \mathbb{R},\ p(x) \geq 0,\ \int_X p(x)\, d\mu(x) = 1\right\}$

The f-divergence is defined as follows [2], [3]

$$D_f(p,q) := \int_X p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x), \quad p,q \in \Omega, \tag{1.1}$$

where the function f is convex on $(0,\infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u = 1$. By appropriately defining this convex function, various divergences are derived. For instance, the following celebrated divergences are particular cases of f-divergence

$$D_{KL}(p,q) := \int_X p(x) \log\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \Omega, \tag{1.2}$$

(Kullback-Leibler divergence [9])

$$D_v(p,q) := \int_X |p(x) - q(x)|\, d\mu(x), \quad p,q \in \Omega; \tag{1.3}$$

(variation distance)

$$D_H(p,q) := \int_X \left|\sqrt{p(x)} - \sqrt{q(x)}\right| d\mu(x), \quad p,q \in \Omega; \tag{1.4}$$

(Hellinger distance [7])

$$D_{\chi^2}(p,q) := \int_X p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2 - 1\right] d\mu(x), \quad p,q \in \Omega; \tag{1.5}$$

(χ^2-*divergence*)

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.

RGMIA Res. Rep. Coll. 22 (2019), Art. 33, 13 pp. Received 21/03/19

$$D_J(p,q) := \int_X [p(x) - q(x)] \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \Omega; \tag{1.6}$$

(Jeffreys distance [8])

$$D_\Delta(p,q) := \int_X \frac{[p(x) - q(x)]^2}{p(x) + q(x)} d\mu(x), \quad p,q \in \Omega. \tag{1.7}$$

(triangular discrimination [12])

In [10], Lin and Wong (see also [11]) introduced the following divergence

$$D_{LW}(p,q) := \int_X p(x) \log\left[\frac{p(x)}{\frac{1}{2}p(x) + \frac{1}{2}q(x)}\right] d\mu(x), \quad p,q \in \Omega. \tag{1.8}$$

This can be represented as follows, using the Kullback-Leibler divergence:

$$D_{LW}(p,q) = D_{KL}\left(p, \frac{p+q}{2}\right).$$

Lin and Wong have established the following inequalities

$$D_{LW}(p,q) \le \frac{1}{2} D_{KL}(p,q); \tag{1.9}$$

$$D_{LW}(p,q) + D_{LW}(q,p) \le D_v(p,q) \le 2; \tag{1.10}$$

$$D_{LW}(p,q) \le 1. \tag{1.11}$$

In [11], Shioya and Da-te improved (1.9)-(1.11) by showing that

$$D_{LW}(p,q) \le \frac{1}{2} D_v(p,q) \le 1.$$

In the same paper [11], the authors introduced the generalised Lin-Wong f-divergence $D_f\left(p, \frac{1}{2}p + \frac{1}{2}q\right)$ and the Hermite-Hadamard (HH) f-divergence

$$D_{HH}^f(p,q) := \int_X p(x) \frac{\int_1^{\frac{q(x)}{p(x)}} f(t)\,dt}{\frac{q(x)}{p(x)} - 1} d\mu(x), \quad p,q \in \Omega \tag{1.12}$$

and, by use of the Hermite-Hadamard inequality for convex functions, proved the following basic inequality

$$D_f\left(p, \frac{p+q}{2}\right) \le D_{HH}^f(p,q) \le \frac{1}{2} D_f(p,q), \tag{1.13}$$

provided that f is convex and normalised, i.e., $f(1) = 0$.

In 2002, Barnett, Cerone & Dragomir [1] improved the inequality (1.13) as follows:

Theorem 1. *Assume that the function $f:(0,\infty)\to\mathbb{R}$ is convex and normalised, i.e. $f(1)=0$. Let $p,q\in\Omega$ then we have the inequality,*

$$\begin{aligned}
0 &\le D_f\left(p,\frac{p+q}{2}\right) \\
&\le \lambda D_f\left(p,p+\frac{\lambda}{2}(q-p)\right)+(1-\lambda)D_f\left(p,\frac{p+q}{2}+\frac{\lambda}{2}(q-p)\right) \\
&\le D_{HH}^f(p,q)\le\frac{1}{2}\left[D_f(p,(1-\lambda)p+\lambda q)+(1-\lambda)D_f(p,q)\right] \\
&\le \frac{1}{2}D_f(p,q),
\end{aligned}\tag{1.14}$$

for all $\lambda\in[0,1]$.

In particular,

$$\begin{aligned}
0 &\le D_f\left(p,\frac{p+q}{2}\right)\le\frac{1}{2}\left[D_f\left(p,\frac{3p+q}{4}\right)+D_f\left(p,\frac{p+3q}{4}\right)\right] \\
&\le D_{HH}^f(p,q)\le\frac{1}{2}\left[D_f\left(p,\frac{p+q}{2}\right)+\frac{1}{2}D_f(p,q)\right] \\
&\le \frac{1}{2}D_f(p,q).
\end{aligned}\tag{1.15}$$

In 2005, [5], the author obtained the following estimate for a differentiable convex and normalised function $f:(0,\infty)\to\mathbb{R}$

$$0\le D_{HH}^f(p,q)-D_f\left(p,\frac{p+q}{2}\right)\le\frac{1}{8}D_{f^\dagger}(p,q)\tag{1.16}$$

for $p,\ q\in\Omega$, where

$$f^\dagger(t):=(t-1)f'(t),\ t\in(0,\infty).\tag{1.17}$$

In the paper [6] we also obtained the dual inequality

$$0\le\frac{1}{2}D_f(p,q)-D_{HH}^f(p,q)\le\frac{1}{8}D_{f^\dagger}(p,q)\tag{1.18}$$

for $p,\ q\in\Omega$.

Motivated by the above results, we establish in this paper other inequalities for the HH f-divergence.

2. General Results

We start with the following useful representation fir the HH f-divergence:

Lemma 1. *Assume that the function $f:(0,\infty)\to\mathbb{R}$ is convex and normalised, then we have the representation*

$$\begin{aligned}
D_{HH}^f(p,q) &= \int_X p(x)\left(\int_0^1 f\left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)ds\right)d\mu(x) \\
&= \int_0^1 D_f(p,sq+(1-s)p)\,ds
\end{aligned}\tag{2.1}$$

for $p,\ q\in\Omega$.

Proof. Using the change of variable

$$t=\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)},\ s\in\left[0,1\right]$$

we have

$$\frac{\int_{1}^{\frac{q(x)}{p(x)}}f\left(t\right)dt}{\frac{q(x)}{p(x)}-1}=\int_{0}^{1}f\left(\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)}\right)ds$$

for $x\in X$ for which $p\left(x\right),\ q\left(x\right),\ q\left(x\right)-p\left(x\right)\neq 0.$

Therefore

$$\begin{aligned}D_{HH}^{f}\left(p,q\right)&:=\int_{X}p\left(x\right)\frac{\int_{1}^{\frac{q(x)}{p(x)}}f\left(t\right)dt}{\frac{q(x)}{p(x)}-1}d\mu\left(x\right)\\&=\int_{X}p\left(x\right)\left(\int_{0}^{1}f\left(\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)}\right)ds\right)d\mu\left(x\right)\\&=\int_{0}^{1}\left(\int_{X}p\left(x\right)f\left(\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)}\right)d\mu\left(x\right)\right)ds,\end{aligned}$$

where for the last equality we used Fubini's theorem.

Since

$$\int_{X}p\left(x\right)f\left(\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)}\right)d\mu\left(x\right)=D_{f}\left(p,sq+\left(1-s\right)p\right)$$

hence

$$\begin{aligned}&\int_{0}^{1}\left(\int_{X}p\left(x\right)f\left(\frac{sq\left(x\right)+\left(1-s\right)p\left(x\right)}{p\left(x\right)}\right)d\mu\left(x\right)\right)ds\\&=\int_{0}^{1}D_{f}\left(p,sq+\left(1-s\right)p\right)ds\end{aligned}$$

and the equalities in are proved. □

For $s\in\left[0,1\right]$ and the convex function $f:\left(0,\infty\right)\to\mathbb{R}$ we define the *s-weighted perspective* $\mathcal{P}_{f,s}:\left(0,\infty\right)\times\left(0,\infty\right)\to\mathbb{R}$ by

$$\mathcal{P}_{f,s}\left(u,v\right):=uf\left(\frac{sv+\left(1-s\right)u}{u}\right). \tag{2.2}$$

We have the following lemma that is of interest in itself as well:

Lemma 2. *Assume that the function $f:\left(0,\infty\right)\to\mathbb{R}$ is convex, then for all $s\in\left[0,1\right]$ the s-weighted perspective $\mathcal{P}_{f,s}$ is also convex as a function of two variables.*

Proof. Let $(u,v)\,,\,(w,z)\in(0,\infty)\times(0,\infty)$ and $\alpha,\,\beta\geq 0$ with $\alpha+\beta=1.$ Then

$$\begin{aligned}
&\mathcal{P}_{f,s}\left(\alpha\left(u,v\right)+\beta\left(w,z\right)\right)\\
&=\mathcal{P}_{f,s}\left(\alpha u+\beta w,\alpha v+\beta z\right)\\
&=\left(\alpha u+\beta w\right)f\left(\frac{s\left(\alpha v+\beta z\right)+\left(1-s\right)\left(\alpha u+\beta w\right)}{\alpha u+\beta w}\right)\\
&=\left(\alpha u+\beta w\right)f\left(\frac{\alpha\left(sv+\left(1-s\right)u\right)+\beta\left(sz+\left(1-s\right)w\right)}{\alpha u+\beta w}\right)\\
&=\left(\alpha u+\beta w\right)f\left(\frac{\alpha u\frac{sv+(1-s)u}{u}+\beta w\frac{sz+(1-s)w}{w}}{\alpha u+\beta w}\right)\\
&\leq\left(\alpha u+\beta w\right)\\
&\times\left[\frac{\alpha u}{\alpha u+\beta w}f\left(\frac{sv+\left(1-s\right)u}{u}\right)+\frac{\beta w}{\alpha u+\beta w}f\left(\frac{sz+\left(1-s\right)w}{w}\right)\right]\\
&=\alpha uf\left(\frac{sv+\left(1-s\right)u}{u}\right)+\beta wf\left(\frac{sz+\left(1-s\right)w}{w}\right)\\
&=\alpha\mathcal{P}_{f,s}\left(u,v\right)+\beta\mathcal{P}_{f,s}\left(w,z\right),
\end{aligned}$$

which proves the joint convexity of the perspective $\mathcal{P}_{f,s}$. □

Remark 1. *If we use the perspective concept, then by (2.1) we also have*

$$D_{HH}^{f}\left(p,q\right)=\int_{0}^{1}\left(\int_{X}\mathcal{P}_{f,s}\left(p\left(x\right),q\left(x\right)\right)d\mu\left(x\right)\right)ds. \tag{2.3}$$

The following joint convexity of the HH f-divergence holds:

Theorem 2. *Assume that the function $f:(0,\infty)\to\mathbb{R}$ is convex and normalised, then D_{HH}^{f} is convex as a mapping of two variables on $\Omega\times\Omega$.*

Proof. Let $(p_1,q_1)\,,\,(p_2,q_2)\in\Omega$ and and $\alpha,\,\beta\geq 0$ with $\alpha+\beta=1.$ Then by the representation (2.3) and Lemma 2 we have

$$\begin{aligned}
&D_{HH}^{f}\left(\alpha\left(p_1,q_1\right)+\beta\left(p_2,q_2\right)\right)\\
&=D_{HH}^{f}\left(\alpha p_1+\beta p_2,\alpha q_1+\beta q_2\right)\\
&=\int_{0}^{1}\left(\int_{X}\mathcal{P}_{f,s}\left(\alpha p_1\left(x\right)+\beta p_2\left(x\right),\alpha q_1\left(x\right)+\beta q_2\left(x\right)\right)d\mu\left(x\right)\right)ds\\
&=\int_{0}^{1}\left(\int_{X}\mathcal{P}_{f,s}\left(\alpha\left(p_1\left(x\right),q_1\left(x\right)\right)+\beta\left(p_2\left(x\right),q_2\left(x\right)\right)\right)d\mu\left(x\right)\right)ds\\
&\geq\int_{0}^{1}\left(\int_{X}\left[\alpha\mathcal{P}_{f,s}\left(p_1\left(x\right),q_1\left(x\right)\right)+\beta\mathcal{P}_{f,s}\left(p_2\left(x\right),q_2\left(x\right)\right)\right]d\mu\left(x\right)\right)ds\\
&=\alpha\int_{0}^{1}\left(\int_{X}\mathcal{P}_{f,s}\left(p_1\left(x\right),q_1\left(x\right)\right)d\mu\left(x\right)\right)ds\\
&+\beta\int_{0}^{1}\left(\int_{X}\mathcal{P}_{f,s}\left(p_2\left(x\right),q_2\left(x\right)\right)d\mu\left(x\right)\right)ds\\
&=\alpha D_{HH}^{f}\left(p_1,q_1\right)+\beta D_{HH}^{f}\left(p_2,q_2\right),
\end{aligned}$$

which proves the desired convexity. □

3. Bounds in Terms of χ^2-Divergence

The above definitions $D_f(p,q)$ and $D_{HH}^f(p,q)$ can be extended to continuous functions f defined on $(0,\infty)$, however, in this general case, the positivity properties of the divergences under consideration do not hold in general.

We have:

Theorem 3. *Assume that the function $f:(0,\infty)\to\mathbb{R}$ is twice differentiable and normalised. Let $0<r\leq 1\leq R<\infty$ and p, $q\in\Omega$ are such that*

$$r\leq\frac{q(x)}{p(x)}\leq R \text{ for } \mu\text{-almost every } x\in X. \tag{3.1}$$

(i) *If there exists a real number m such that*

$$m\leq f''(t) \text{ for all } t\in[r,R], \tag{3.2}$$

then we have the inequality

$$0\leq D_f\left(p,\frac{p+q}{2}\right)-\frac{1}{8}mD_{\chi^2}(p,q)\leq D_{HH}^f(p,q)-\frac{1}{6}mD_{\chi^2}(p,q). \tag{3.3}$$

(ii) *If there exists the real number M such that*

$$f''(t)\leq M \text{ for all } t\in[r,R], \tag{3.4}$$

then we have the inequality

$$0\leq\frac{1}{8}MD_{\chi^2}(p,q)-D_f\left(p,\frac{p+q}{2}\right)\leq\frac{1}{6}MD_{\chi^2}(p,q)-D_{HH}^f(p,q). \tag{3.5}$$

Proof. (i) Consider the auxiliary function $g_m:[r,R]\to\mathbb{R}$, $g_m(t):=f(t)-\frac{1}{2}m\left(\ell^2(t)-1\right)$, where $\ell(t)=t$ is the identity function. This function is convex and normalized on $[r,R]$, since g_m is twice differentiable and

$$g_m''(t):=f''(t)-m\geq 0 \text{ for all } t\in[r,R].$$

We have for p, $q\in\Omega$ that

$$\begin{aligned}
&D_{HH}^{g_m}(p,q)\\
&=D_{HH}^f(p,q)-\frac{1}{2}mD_{HH}^{\ell^2-1}(p,q)\\
&=D_{HH}^f(p,q)-\frac{1}{2}m\int_X p(x)\left(\int_0^1\left[\left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)^2-1\right]ds\right)d\mu(x)\\
&=D_{HH}^f(p,q)-\frac{1}{2}m\int_X p(x)\left(\int_0^1\left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)^2ds\right)d\mu(x)\\
&+\frac{1}{2}m\int_X p(x)\,d\mu(x)\\
&=D_{HH}^f(p,q)-\frac{1}{2}m\int_X p(x)\left(\int_0^1\left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)^2ds\right)d\mu(x)+\frac{1}{2}m.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\int_0^1 \left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)^2 ds \\
&= \int_0^1 \left[s^2\left(\frac{q(x)}{p(x)}\right)^2 + 2s(1-s)\frac{q(x)}{p(x)} + (1-s)^2\right] ds \\
&= \frac{1}{3}\left(\frac{q(x)}{p(x)}\right)^2 + \frac{1}{3}\frac{q(x)}{p(x)} + \frac{1}{3} = \frac{1}{3}\left[\left(\frac{q(x)}{p(x)}\right)^2 + \frac{q(x)}{p(x)} + 1\right]
\end{aligned}$$

and

$$\begin{aligned}
&\int_X p(x)\left(\int_0^1 \left(\frac{sq(x)+(1-s)p(x)}{p(x)}\right)^2 ds\right) d\mu(x) \\
&= \frac{1}{3}\int_X p(x)\left(\left(\frac{q(x)}{p(x)}\right)^2 + \frac{q(x)}{p(x)} + 1\right) d\mu(x) \\
&= \frac{1}{3}\left[\int_X p(x)\left(\frac{q(x)}{p(x)}\right)^2 d\mu(x) + \int_X p(x)\frac{q(x)}{p(x)} d\mu(x) + \int_X p(x)\,d\mu(x)\right] \\
&= \frac{1}{3}\left[\int_X \frac{q^2(x)}{p(x)} d\mu(x) + \int_X q(x)\,d\mu(x) + \int_X p(x)\,d\mu(x)\right] \\
&= \frac{1}{3}\left[\int_X \frac{q^2(x)}{p(x)} d\mu(x) + 1 + 1\right] = \frac{1}{3}\left[D_{\chi^2}(p,q) + 3\right] = \frac{1}{3}D_{\chi^2}(p,q) + 1.
\end{aligned}$$

Therefore

$$\begin{aligned}
D_{HH}^{g_m}(p,q) &= D_{HH}^{f}(p,q) - \frac{1}{2}m\left[\frac{1}{3}D_{\chi^2}(p,q) + 1\right] + \frac{1}{2}m \\
&= D_{HH}^{f}(p,q) - \frac{1}{6}mD_{\chi^2}(p,q).
\end{aligned}$$

We also have

$$\begin{aligned}
D_{g_m}\left(p, \frac{p+q}{2}\right) &= D_f\left(p, \frac{p+q}{2}\right) - \frac{1}{2}mD_{\ell^2-1}\left(p, \frac{p+q}{2}\right) \\
&= D_f\left(p, \frac{p+q}{2}\right) - \frac{1}{2}mD_{\chi^2}\left(p, \frac{p+q}{2}\right).
\end{aligned}$$

Now,

$$\begin{aligned}
D_{\chi^2}\left(p, \frac{p+q}{2}\right) &= \int_X p(x)\left[\left(\frac{\frac{p(x)+q(x)}{2}}{p(x)}\right)^2 - 1\right] d\mu(x) \\
&= \int_X p(x)\left[\left(\frac{p(x)+q(x)}{2p(x)}\right)^2 - 1\right] d\mu(x)
\end{aligned}$$

$$
\begin{aligned}
&= \int_X p(x)\left[\frac{1}{4}\left(\frac{q(x)}{p(x)}+1\right)^2-1\right]d\mu(x)\\
&= \int_X p(x)\left[\frac{1}{4}\left(\left(\frac{q(x)}{p(x)}\right)^2+2\frac{q(x)}{p(x)}+1\right)-1\right]d\mu(x)\\
&= \frac{1}{4}\int_X p(x)\left(\left(\frac{q(x)}{p(x)}\right)^2+2\frac{q(x)}{p(x)}+1\right)d\mu(x)-1\\
&= \frac{1}{4}\left[\int_X \frac{q^2(x)}{p(x)}d\mu(x)+2\int_X p(x)\frac{q(x)}{p(x)}d\mu(x)+\int_X p(x)\,d\mu(x)\right]-1\\
&= \frac{1}{4}D_{\chi^2}(p,q)+1-1=\frac{1}{4}D_{\chi^2}(p,q),
\end{aligned}
$$

therefore

$$
\begin{aligned}
D_{g_m}\left(p,\frac{p+q}{2}\right) &= D_f\left(p,\frac{p+q}{2}\right)-\frac{1}{2}mD_{\ell^2-1}\left(p,\frac{p+q}{2}\right)\\
&= D_f\left(p,\frac{p+q}{2}\right)-\frac{1}{8}mD_{\chi^2}(p,q).
\end{aligned}
$$

If we use the first inequality in (1.13) for g_m we have

$$0\le D_{g_m}\left(p,\frac{p+q}{2}\right)\le D_{HH}^{g_m}(p,q),$$

which by above calculations gives

$$0\le D_f\left(p,\frac{p+q}{2}\right)-\frac{1}{8}mD_{\chi^2}(p,q)\le D_{HH}^{f}(p,q)-\frac{1}{6}mD_{\chi^2}(p,q).$$

This proves (3.3).

(ii) Consider the auxiliary function $g_M:[r,R]\to\mathbb{R}$, $g_M(t):=\frac{1}{2}M\left(\ell^2(t)-1\right)-f(t)$, where $\ell(t)=t$ is the identity function. This function is convex and normalized on $[r,R]$, since g_M is twice differentiable and

$$g_M''(t)=M-f''(t)\ge 0 \text{ for all } t\in[r,R].$$

Now, by using a similar argument to the one for the auxiliary function g_m we deduce the desired result (3.5). □

Corollary 1. *With the assumptions of Theorem 3 and if*

$$0<m\le f''(t)\le M<\infty \textit{ for all } t\in[r,R], \tag{3.6}$$

then we have

$$\frac{1}{8}mD_{\chi^2}(p,q)\le D_f\left(p,\frac{p+q}{2}\right)\le\frac{1}{8}MD_{\chi^2}(p,q), \tag{3.7}$$

$$\frac{1}{6}mD_{\chi^2}(p,q)\le D_{HH}^{f}(p,q)\le\frac{1}{6}MD_{\chi^2}(p,q) \tag{3.8}$$

and

$$\frac{1}{24}mD_{\chi^2}(p,q)\le D_{HH}^{f}(p,q)-D_f\left(p,\frac{p+q}{2}\right)\le\frac{1}{24}MD_{\chi^2}(p,q). \tag{3.9}$$

We also have:

Theorem 4. *Assume that the function $f : (0, \infty) \to \mathbb{R}$ is twice differentiable and normalised. Let $0 < r \le 1 \le R < \infty$ and $p,\ q \in \Omega$ are such that the condition (3.1) is valid.*

(i) *If there exists a real number m such that the assumption (3.2) holds, then we have the inequality*

$$0 \le D_{HH}^{f}(p, q) - \frac{1}{6} m D_{\chi^2}(p, q) \le \frac{1}{2} D_f(p, q) - \frac{1}{4} m D_{\chi^2}(p, q). \tag{3.10}$$

(ii) *If there exists the real number M such that the assumption (3.4) holds, then we have the inequality*

$$0 \le \frac{1}{6} M D_{\chi^2}(p, q) - D_{HH}^{f}(p, q) \le \frac{1}{2} M D_{\chi^2}(p, q) - D_f(p, q). \tag{3.11}$$

Proof. (i) Consider the auxiliary function $g_m : [r, R] \to \mathbb{R},\ g_m(t) := f(t) - \frac{1}{2} m \left(\ell^2(t) - 1\right)$, where $\ell(t) = t$ is the identity function. This function is convex and normalized on $[r, R]$.

We have

$$D_{HH}^{g_m}(p, q) = D_{HH}^{f}(p, q) - \frac{1}{6} m D_{\chi^2}(p, q)$$

and

$$\begin{aligned} D_{g_m}(p, q) &:= \int_X p(x)\, g_m\left[\frac{q(x)}{p(x)}\right] d\mu(x) \\ &= \int_X p(x) \left[f\left(\frac{q(x)}{p(x)}\right) - \frac{1}{2} m \left(\ell^2\left(\frac{q(x)}{p(x)}\right) - 1\right)\right] d\mu(x) \\ &= D_f(p, q) - \frac{1}{2} m D_{\chi^2}(p, q). \end{aligned}$$

If we use the second inequality in (1.13) we have

$$0 \le D_{HH}^{g_m}(p, q) \le \frac{1}{2} D_{g_m}(p, q),$$

namely

$$\begin{aligned} 0 \le D_{HH}^{f}(p, q) - \frac{1}{6} m D_{\chi^2}(p, q) &\le \frac{1}{2}\left[D_f(p, q) - \frac{1}{2} m D_{\chi^2}(p, q)\right] \\ &= \frac{1}{2} D_f(p, q) - \frac{1}{4} m D_{\chi^2}(p, q), \end{aligned}$$

which proves (3.10).

(ii) Follows in a similar way for the auxiliary function $g_M : [r, R] \to \mathbb{R},\ g_M(t) := \frac{1}{2} M \left(\ell^2(t) - 1\right) - f(t)$. □

Corollary 2. *With the assumptions of Theorem 3 and if the condition (3.6) holds, then we have*

$$\frac{1}{2} m D_{\chi^2}(p, q) \le D_f(p, q) \le \frac{1}{2} M D_{\chi^2}(p, q) \quad \text{(see also [4])} \tag{3.12}$$

and

$$\frac{1}{12} m D_{\chi^2}(p, q) \le \frac{1}{2} D_f(p, q) - D_{HH}^{f}(p, q) \le \frac{1}{12} M D_{\chi^2}(p, q). \tag{3.13}$$

Further, we observe that by using the definitions of the auxiliary mappings $g_m(t)$ and $g_M(t)$ we have

$$g_m^{\dagger}(t) = (t-1)\left(f(t) - \frac{1}{2}m\left(t^2-1\right)\right)' = f^{\dagger}(t) - mt(t-1)$$

and

$$g_M^{\dagger}(t) = Mt(t-1) - f^{\dagger}(t).$$

This give

$$\begin{aligned} D_{g_m^{\dagger}}(p,q) &= D_{f^{\dagger}}(p,q) - m\int_X p(x)\frac{q(x)}{p(x)}\left(\frac{q(x)}{p(x)}-1\right)d\mu(x) \\ &= D_{f^{\dagger}}(p,q) - mD_{\chi^2}(p,q) \end{aligned} \tag{3.14}$$

and

$$D_{g_M^{\dagger}}(p,q) = MD_{\chi^2}(p,q) - D_{f^{\dagger}}(p,q). \tag{3.15}$$

Theorem 5. *Assume that the function $f:(0,\infty)\to\mathbb{R}$ is twice differentiable and normalised. Let $0<r\le 1\le R<\infty$ and $p,\, q\in\Omega$ are such that the condition (3.1) is valid.*

(i) *If there exists a real number m such that the assumption (3.2) holds, then we have the inequality*

$$\begin{aligned} 0 &\le D_{HH}^{f}(p,q) - D_f\left(p,\frac{p+q}{2}\right) - \frac{1}{24}mD_{\chi^2}(p,q) \\ &\le \frac{1}{8}\left[D_{f^{\dagger}}(p,q) - mD_{\chi^2}(p,q)\right] \end{aligned} \tag{3.16}$$

and

$$\begin{aligned} 0 &\le \frac{1}{2}D_f(p,q) - D_{HH}^{f}(p,q) - \frac{1}{12}mD_{\chi^2}(p,q) \\ &\le \frac{1}{8}\left[D_{f^{\dagger}}(p,q) - mD_{\chi^2}(p,q)\right]. \end{aligned} \tag{3.17}$$

(ii) *If there exists the real number M such that the assumption (3.4) holds, then we have the inequality*

$$\begin{aligned} 0 &\le \frac{1}{24}MD_{\chi^2}(p,q) - D_{HH}^{f}(p,q) + D_f\left(p,\frac{p+q}{2}\right) \\ &\le \frac{1}{8}\left[MD_{\chi^2}(p,q) - D_{f^{\dagger}}(p,q)\right] \end{aligned} \tag{3.18}$$

and

$$\begin{aligned} 0 &\le \frac{1}{12}MD_{\chi^2}(p,q) - \frac{1}{2}D_f(p,q) + D_{HH}^{f}(p,q) \\ &\le \frac{1}{8}\left[MD_{\chi^2}(p,q) - D_{f^{\dagger}}(p,q)\right]. \end{aligned} \tag{3.19}$$

Proof. (i) If we use the inequality (1.16) for g_m, then we have

$$0 \le D_{HH}^{g_m}(p,q) - D_{g_m}\left(p,\frac{p+q}{2}\right) \le \frac{1}{8}D_{g_m^{\dagger}}(p,q),$$

namely

$$\begin{aligned}0 &\le D_{HH}^{f}(p,q) - \frac{1}{6}mD_{\chi^2}(p,q) - D_f\left(p, \frac{p+q}{2}\right) + \frac{1}{8}mD_{\chi^2}(p,q) \\ &\le \frac{1}{8}\left[D_{f^\dagger}(p,q) - mD_{\chi^2}(p,q)\right],\end{aligned}$$

which is equivalent to (3.16).

If we use (1.18) for g_m, then we have

$$0 \le \frac{1}{2}D_{g_m}(p,q) - D_{HH}^{g_m}(p,q) \le \frac{1}{8}D_{g_m^\dagger}(p,q),$$

namely

$$\begin{aligned}0 &\le \frac{1}{2}\left[D_f(p,q) - \frac{1}{2}mD_{\chi^2}(p,q)\right] - D_{HH}^{f}(p,q) + \frac{1}{6}mD_{\chi^2}(p,q) \\ &\le \frac{1}{8}\left[D_{f^\dagger}(p,q) - mD_{\chi^2}(p,q)\right],\end{aligned}$$

(ii) Follows in a similar way for g_M. □

Finally, we have:

Corollary 3. *With the assumptions of Theorem 3 and if the condition (3.6) holds, then we have*

$$\begin{aligned}\frac{1}{12}mD_{\chi^2}(p,q) &\le \frac{1}{8}D_{f^\dagger}(p,q) - D_{HH}^{f}(p,q) + D_f\left(p, \frac{p+q}{2}\right) \\ &\le \frac{1}{12}MD_{\chi^2}(p,q)\end{aligned} \tag{3.20}$$

and

$$\frac{1}{24}mD_{\chi^2}(p,q) \le \frac{1}{8}D_{f^\dagger}(p,q) + D_{HH}^{f}(p,q) - \frac{1}{2}D_f(p,q) \le \frac{1}{24}MD_{\chi^2}(p,q). \tag{3.21}$$

4. An Example

We consider the convex and normalized function $f : (0, \infty) \to R$, $f(t) = -\ln t$. We have

$$D_f(p,q) := D_{KL}(p,q)$$

and

$$D_f\left(p, \frac{p+q}{2}\right) = D_{LW}(p,q)$$

for all p, $q \in \Omega$.

We define the *identric mean* of two positive numbers a, $b > 0$

$$I(a,b) := \begin{cases} \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{1/(b-a)} & \text{if } b \ne a, \\ a \text{ if } b = a. \end{cases}$$

We observe that

$$\frac{1}{b-a}\int_a^b \ln t dt = \frac{b\ln b - b - b\ln b + a}{b-a} = \ln I(a,b).$$

Therefore

$$D_{HH}^{f}\left(p,q\right)=-\int_{X}p\left(x\right)\frac{\int_{1}^{\frac{q(x)}{p(x)}}\ln tdt}{\frac{q(x)}{p(x)}-1}d\mu\left(x\right)=-\int_{X}p\left(x\right)\ln\left[I\left(\frac{q\left(x\right)}{p\left(x\right)},1\right)\right]d\mu\left(x\right)$$
$$=\int_{X}p\left(x\right)\ln\left[I\left(\frac{q\left(x\right)}{p\left(x\right)},1\right)\right]^{-1}d\mu\left(x\right)=:D_{HH}^{KL}\left(p,q\right),$$

where we call $D_{HH}^{KL}\left(p,q\right)$ the *Kullback-Leibler HH divergence.*

If $0<r<1<R<\infty$ then for $f\left(t\right)=-\ln t$,

$$\inf_{t\in[r,R]}f''\left(t\right)=\inf_{t\in[r,R]}\frac{1}{t^{2}}=\frac{1}{R^{2}},\quad \sup_{t\in[r,R]}f''\left(t\right)=\sup_{t\in[r,R]}\frac{1}{t^{2}}=\frac{1}{r^{2}}.$$

If $p,\ q\in\Omega$ satisfy the condition (3.1), then by using (3.7)-(3.9) for $m=\frac{1}{R^{2}}$ and $M=\frac{1}{r^{2}}$ we get

$$\frac{1}{8R^{2}}D_{\chi^{2}}\left(p,q\right)\leq D_{LW}\left(p,q\right)\leq\frac{1}{8r^{2}}D_{\chi^{2}}\left(p,q\right), \tag{4.1}$$

$$\frac{1}{6R^{2}}D_{\chi^{2}}\left(p,q\right)\leq D_{HH}^{KL}\left(p,q\right)\leq\frac{1}{6r^{2}}D_{\chi^{2}}\left(p,q\right) \tag{4.2}$$

and

$$\frac{1}{24R^{2}}D_{\chi^{2}}\left(p,q\right)\leq D_{HH}^{KL}\left(p,q\right)-D_{LW}\left(p,q\right)\leq\frac{1}{24r^{2}}D_{\chi^{2}}\left(p,q\right). \tag{4.3}$$

By (3.13), we also have

$$\frac{1}{12R^{2}}D_{\chi^{2}}\left(p,q\right)\leq\frac{1}{2}D_{KL}\left(p,q\right)-D_{HH}^{KL}\left(p,q\right)\leq\frac{1}{12r^{2}}D_{\chi^{2}}\left(p,q\right). \tag{4.4}$$

Now, if $f\left(t\right)=-\ln t$, then

$$f^{\dagger}\left(t\right):=-\left(\frac{t-1}{t}\right)=\frac{1}{t}-1$$

and

$$D_{f^{\dagger}}\left(p,q\right)=\int_{X}p\left(x\right)\left(\frac{p\left(x\right)}{q\left(x\right)}-1\right)d\mu\left(x\right)=\int_{X}\left(\frac{p^{2}\left(x\right)}{q\left(x\right)}-p\left(x\right)\right)d\mu\left(x\right)$$
$$=\int_{X}\frac{p^{2}\left(x\right)}{q\left(x\right)}d\mu\left(x\right)-1=D_{\chi^{2}}\left(q,p\right)$$

for all $p,\ q\in\Omega$.

Finally, by the (3.20) and (3.21) we also have

$$\begin{aligned}\frac{1}{12R^{2}}D_{\chi^{2}}\left(p,q\right)&\leq\frac{1}{8}D_{\chi^{2}}\left(q,p\right)-D_{HH}^{KL}\left(p,q\right)+D_{LW}\left(p,q\right)\\&\leq\frac{1}{12r^{2}}D_{\chi^{2}}\left(p,q\right)\end{aligned} \tag{4.5}$$

and

$$\begin{aligned}\frac{1}{24R^{2}}D_{\chi^{2}}\left(p,q\right)&\leq\frac{1}{8}D_{\chi^{2}}\left(q,p\right)+D_{HH}^{KL}\left(p,q\right)-\frac{1}{2}D_{KL}\left(p,q\right)\\&\leq\frac{1}{24r^{2}}D_{\chi^{2}}\left(p,q\right),\end{aligned} \tag{4.6}$$

provided $p,\ q\in\Omega$ satisfy the condition (3.1).

References

[1] N. S. BARNETT, P. CERONE, AND S. S. DRAGOMIR, Some new inequalities for Hermite-Hadamard divergence in information theory, in *Stochastic Analysis and Applications*, Volume 3, Edited by Yeol Je Cho, Jong Kyu Kim, Yong Kab Choi, Nova Science Publishers, New York, 2003, pp. 7-20. Preprint *RGMIA Res. Rep. Coll.* **5** (2002), No. 4, Art. 8. [Online https://rgmia.org/papers/v5n4/NIHHDIT.pdf].

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

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

[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, An Ostrowski type inequality for convex functions, *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat.* 16 (2005), 12–25. [Online http://pefmath2.etf.bg.ac.rs/files/125/963.pdf].

[6] S. S. DRAGOMIR, A generalised trapezoid type inequality for convex functions (June 2003). *Mathematics Preprint Archive* Vol. 2003, Issue 6, pp 56-65. Available at SSRN: https://ssrn.com/abstract=3177646, see also https://arxiv.org/pdf/math/0305374.pdf.

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

[8] H. JEFFREYS, An invariant form for the prior probability in estimating problems, *Proc. Roy. Soc. London*, **186** A (1946), 453-461.

[9] S. KULLBACK and R. A. LEIBLER, On information and sufficiency, *Ann. Math. Stat.*, **22** (1951), 79-86.

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

[11] 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.

[12] F. TOPSØE, Some inequalities for information divergence and related measures of discrimination, *IEEE Transactions on Information Theory,* Volume: **46** , Issue: 4, 1602 - 1609. Preprint *Res. Rep. Coll., RGMIA,* **2** (1) (1999), 85-98.

[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