

QUASILINEARITY OF THE JENSEN INTEGRAL INEQUALITY WITH APPLICATIONS

S.S. DRAGOMIR^{1,2}

ABSTRACT. The quasilinearity and monotonicity of two functionals associated to the celebrated Jensen's integral inequality for convex functions with applications for Hölder's inequality and f -divergence measures in information theory are given.

1. INTRODUCTION

Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space consisting of a set Ω , a σ -algebra \mathcal{A} of parts of Ω and a countably additive and positive measure μ on \mathcal{A} with values in $\mathbb{R} \cup \{\infty\}$. For a μ -measurable function $w : \Omega \rightarrow \mathbb{R}$, with $w(x) \geq 0$ for μ -a.e. (almost every) $x \in \Omega$, consider the Lebesgue space

$$L_w(\Omega, \mu) := \{f : \Omega \rightarrow \mathbb{R}, f \text{ is } \mu\text{-measurable and } \int_{\Omega} w(x) |f(x)| d\mu(x) < \infty\}.$$

For simplicity of notation we write everywhere in the sequel $\int_{\Omega} w d\mu$ instead of $\int_{\Omega} w(x) d\mu(x)$.

In order to provide a reverse of the celebrated Jensen's integral inequality for convex functions, S.S. Dragomir obtained in 2002 [12] the following result:

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

$$\begin{aligned} (1.1) \quad 0 &\leq \int_{\Omega} w(\Phi \circ f) d\mu - \Phi \left(\int_{\Omega} w f d\mu \right) \\ &\leq \int_{\Omega} w(\Phi' \circ f) f d\mu - \int_{\Omega} w(\Phi' \circ f) d\mu \int_{\Omega} w f d\mu \\ &\leq \frac{1}{2} [\Phi'(M) - \Phi'(m)] \int_{\Omega} w \left| f - \int_{\Omega} w f d\mu \right| d\mu. \end{aligned}$$

For a generalization of the first inequality in (1.1) without the differentiability assumption and the derivative Φ' replaced with a selection φ from the subdifferential $\partial\Phi$, see the paper [28] by C.P. Niculescu.

The following discrete inequality is valid as well.

Date: August, 2011.

1991 Mathematics Subject Classification. Primary 26D15, 26D20; Secondary 94A05..

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

Corollary 1. Let $\Phi : [m, M] \rightarrow \mathbb{R}$ be a differentiable convex function on (m, M) . If $x_i \in [m, M]$ and $w_i \geq 0$ ($i = 1, \dots, n$) with $W_n := \sum_{i=1}^n w_i = 1$, then one has the counterpart of Jensen's weighted discrete inequality:

$$(1.2) \quad \begin{aligned} 0 &\leq \sum_{i=1}^n w_i \Phi(x_i) - \Phi\left(\sum_{i=1}^n w_i x_i\right) \\ &\leq \sum_{i=1}^n w_i \Phi'(x_i) x_i - \sum_{i=1}^n w_i \Phi'(x_i) \sum_{i=1}^n w_i x_i \\ &\leq \frac{1}{2} [\Phi'(M) - \Phi'(m)] \sum_{i=1}^n w_i \left| x_i - \sum_{j=1}^n w_j x_j \right|. \end{aligned}$$

Remark 1. We notice that the inequality between the first and the second term in (1.2) was proved in 1994 by Dragomir & Ionescu, see [16].

Utilising a different approach than in [12], we obtained in [15] the following two results that provide other upper bounds for the Jensen's difference:

$$\int_{\Omega} w (\Phi \circ f) d\mu - \Phi\left(\int_{\Omega} w f d\mu\right).$$

Theorem 2. Let $\Phi : I \rightarrow \mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \overset{\circ}{I}$, $\overset{\circ}{I}$ is the interior of I . If $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable, satisfies the bounds

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

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

$$(1.3) \quad \begin{aligned} 0 &\leq \int_{\Omega} w (\Phi \circ f) d\mu - \Phi(\bar{f}_{\Omega, w}) \\ &\leq \frac{(M - \bar{f}_{\Omega, w})(\bar{f}_{\Omega, w} - m)}{M - m} \sup_{t \in (m, M)} \Psi_{\Phi}(t; m, M) \\ &\leq (M - \bar{f}_{\Omega, w})(\bar{f}_{\Omega, w} - m) \frac{\Phi'_-(M) - \Phi'_+(m)}{M - m} \\ &\leq \frac{1}{4} (M - m) [\Phi'_-(M) - \Phi'_+(m)], \end{aligned}$$

where $\bar{f}_{\Omega, w} := \int_{\Omega} w(x) f(x) d\mu(x) \in [m, M]$ and $\Psi_{\Phi}(\cdot; m, M) : (m, M) \rightarrow \mathbb{R}$ is defined by

$$\Psi_{\Phi}(t; m, M) = \frac{\Phi(M) - \Phi(t)}{M - t} - \frac{\Phi(t) - \Phi(m)}{t - m}.$$

We also have the inequality

$$(1.4) \quad \begin{aligned} 0 &\leq \int_{\Omega} w (\Phi \circ f) d\mu - \Phi(\bar{f}_{\Omega, w}) \leq \frac{1}{4} (M - m) \Psi_{\Phi}(\bar{f}_{\Omega, w}; m, M) \\ &\leq \frac{1}{4} (M - m) [\Phi'_-(M) - \Phi'_+(m)], \end{aligned}$$

provided that $\bar{f}_{\Omega, w} \in (m, M)$.

Theorem 3. *With the assumptions of Theorem 4, we have the inequalities*

$$\begin{aligned}
 (1.5) \quad 0 &\leq \int_{\Omega} w (\Phi \circ f) d\mu(x) - \Phi(\bar{f}_{\Omega,w}) \\
 &\leq 2 \max \left\{ \frac{M - \bar{f}_{\Omega,w}}{M - m}, \frac{\bar{f}_{\Omega,w} - m}{M - m} \right\} \left[\frac{\Phi(m) + \Phi(M)}{2} - \Phi\left(\frac{m+M}{2}\right) \right] \\
 &\leq \frac{1}{2} \max \{M - \bar{f}_{\Omega,w}, \bar{f}_{\Omega,w} - m\} [\Phi'_-(M) - \Phi'_+(m)].
 \end{aligned}$$

Motivated by these results we establish in the current paper some refinements and reverses of the Jensen integral inequality by capitalizing on the superadditivity and monotonicity properties of two associated functionals. Application in connection with the Hölder inequality and for f -divergence measures in Information Theory are provided as well.

2. SUPERADDITIVITY AND MONOTONICITY PROPERTIES

For a μ -measurable function $w : \Omega \rightarrow \mathbb{R}$, with $w(x) \geq 0$ for μ -a.e. $x \in \Omega$ and $\int_{\Omega} w d\mu > 0$ we consider the functional

$$(2.1) \quad J(w; \Phi, f) := \int_{\Omega} w (\Phi \circ f) d\mu - \Phi\left(\frac{\int_{\Omega} w f d\mu}{\int_{\Omega} w d\mu}\right) \int_{\Omega} w d\mu \geq 0,$$

where $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_w(\Omega, \mu)$.

Theorem 4. *Let $w_i : \Omega \rightarrow \mathbb{R}$, with $w_i(x) \geq 0$ for μ -a.e. (almost every) $x \in \Omega$ and $\int_{\Omega} w_i d\mu > 0, i \in \{1, 2\}$. If $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_{w_1}(\Omega, \mu) \cap L_{w_2}(\Omega, \mu)$, then*

$$(2.2) \quad J(w_1 + w_2; \Phi, f) \geq J(w_1; \Phi, f) + J(w_2; \Phi, f) \geq 0$$

i.e., J is a superadditive functional of weights.

Moreover, if $w_2 \geq w_1 \geq 0$ μ -a.e. on Ω , then

$$(2.3) \quad J(w_2; \Phi, f) \geq J(w_1; \Phi, f) \geq 0,$$

i.e., J is a monotonic nondecreasing functional of weights.

Proof. Utilising the convexity property of Φ we have successively

$$\begin{aligned}
(2.4) \quad & J(w_1 + w_2; \Phi, f) \\
&= \int_{\Omega} (w_1 + w_2) (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} (w_1 + w_2) f d\mu}{\int_{\Omega} (w_1 + w_2) d\mu} \right) \int_{\Omega} (w_1 + w_2) d\mu \\
&= \int_{\Omega} w_1 (\Phi \circ f) d\mu + \int_{\Omega} w_2 (\Phi \circ f) d\mu \\
&\quad - \Phi \left(\frac{\int_{\Omega} w_1 d\mu \cdot \frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} + \int_{\Omega} w_2 d\mu \cdot \frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu}}{\int_{\Omega} (w_1 + w_2) d\mu} \right) \int_{\Omega} (w_1 + w_2) d\mu \\
&\geq \int_{\Omega} w_1 (\Phi \circ f) d\mu + \int_{\Omega} w_2 (\Phi \circ f) d\mu \\
&\quad - \left[\frac{\int_{\Omega} w_1 d\mu}{\int_{\Omega} (w_1 + w_2) d\mu} \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) + \frac{\int_{\Omega} w_2 d\mu}{\int_{\Omega} (w_1 + w_2) d\mu} \Phi \left(\frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu} \right) \right] \\
&\quad \times \int_{\Omega} (w_1 + w_2) d\mu \\
&= \int_{\Omega} w_1 (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) \int_{\Omega} w_1 d\mu \\
&\quad + \int_{\Omega} w_2 (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu} \right) \int_{\Omega} w_2 d\mu \\
&= J(w_1; \Phi, f) + J(w_2; \Phi, f)
\end{aligned}$$

which proves the superadditivity property.

Now, if $w_2 \geq w_1 \geq 0$, then on applying the superadditivity property we have

$$\begin{aligned}
J(w_2; \Phi, f) &= J(w_1 + (w_2 - w_1); \Phi, f) \geq J(w_1; \Phi, f) + J(w_2 - w_1; \Phi, f) \\
&\geq J(w_1; \Phi, f)
\end{aligned}$$

since by the Jensen's inequality for the positive weights we have $J(w_2 - w_1; \Phi, f) \geq 0$. \square

The above theorem has a simple however interesting consequence that provides both a refinement and a reverse for the Jensen's integral inequality:

Corollary 2. *Let $w_i : \Omega \rightarrow \mathbb{R}$, with $w_i(x) \geq 0$ for μ -a.e. $x \in \Omega$, $\int_{\Omega} w_i d\mu > 0$, $i \in \{1, 2\}$ and there exists the nonnegative constants γ, Γ such that*

$$(2.5) \quad 0 \leq \gamma \leq \frac{w_2}{w_1} \leq \Gamma < \infty \quad \mu\text{-a.e. on } \Omega.$$

If $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_{w_1}(\Omega, \mu) \cap L_{w_2}(\Omega, \mu)$, then

$$\begin{aligned}
(2.6) \quad & 0 \leq \gamma \cdot \left[\int_{\Omega} w_1 (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) \int_{\Omega} w_1 d\mu \right] \\
& \leq \int_{\Omega} w_2 (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu} \right) \int_{\Omega} w_2 d\mu \\
& \leq \Gamma \cdot \left[\int_{\Omega} w_1 (\Phi \circ f) d\mu - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) \int_{\Omega} w_1 d\mu \right]
\end{aligned}$$

or, equivalently,

$$\begin{aligned}
 (2.7) \quad 0 &\leq \gamma \cdot \frac{\int_{\Omega} w_1 d\mu}{\int_{\Omega} w_2 d\mu} \left[\frac{\int_{\Omega} w_1 (\Phi \circ f) d\mu}{\int_{\Omega} w_1 d\mu} - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) \right] \\
 &\leq \frac{\int_{\Omega} w_2 (\Phi \circ f) d\mu}{\int_{\Omega} w_2 d\mu} - \Phi \left(\frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu} \right) \\
 &\leq \Gamma \cdot \frac{\int_{\Omega} w_1 d\mu}{\int_{\Omega} w_2 d\mu} \left[\frac{\int_{\Omega} w_1 (\Phi \circ f) d\mu}{\int_{\Omega} w_1 d\mu} - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) \right].
 \end{aligned}$$

Proof. From (2.5) we have $\gamma w_1 \leq w_2 \leq \Gamma w_1 < \infty$ μ -a.e. on Ω and by the monotonicity property (2.3) we get

$$(2.8) \quad J(\Gamma w_1; \Phi, f) \geq J(w_2; \Phi, f) \geq J(\gamma w_1; \Phi, f).$$

Since the functional is positive homogeneous, namely $J(\alpha w; \Phi, f) = \alpha J(w; \Phi, f)$, then we get from (2.8) the desired result (2.6). \square

Remark 2. Assume that $\mu(\Omega) < \infty$ and let $w : \Omega \rightarrow \mathbb{R}$, with $w(x) \geq 0$ for μ -a.e. $x \in \Omega$, $\int_{\Omega} w d\mu > 0$ and w is essentially bounded, i.e. $\text{ess inf}_{x \in \Omega} w(x)$ and $\text{ess sup}_{x \in \Omega} w(x)$ are finite.

If $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_w(\Omega, \mu) \cap L(\Omega, \mu)$, then

$$\begin{aligned}
 (2.9) \quad 0 &\leq \frac{\text{ess inf}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} \left[\frac{\int_{\Omega} (\Phi \circ f) d\mu}{\mu(\Omega)} - \Phi \left(\frac{\int_{\Omega} f d\mu}{\mu(\Omega)} \right) \right] \\
 &\leq \frac{\int_{\Omega} w (\Phi \circ f) d\mu}{\int_{\Omega} w d\mu} - \Phi \left(\frac{\int_{\Omega} w f d\mu}{\int_{\Omega} w d\mu} \right) \\
 &\leq \frac{\text{ess sup}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} \left[\frac{\int_{\Omega} (\Phi \circ f) d\mu}{\mu(\Omega)} - \Phi \left(\frac{\int_{\Omega} f d\mu}{\mu(\Omega)} \right) \right].
 \end{aligned}$$

This result can be used to provide the following result related to the Hermite-Hadamard inequality for convex functions that states that

$$\frac{1}{b-a} \int_a^b \Phi(t) dt \geq \Phi \left(\frac{a+b}{2} \right)$$

for any convex function $\Phi : [a, b] \rightarrow \mathbb{R}$.

Indeed, if $w : [a, b] \rightarrow [0, \infty)$ is Lebesgue integrable, then we have

$$\begin{aligned}
 (2.10) \quad 0 &\leq \frac{\text{ess inf}_{x \in [a,b]} w(x)}{\frac{1}{b-a} \int_a^b w(t) dt} \left[\frac{1}{b-a} \int_a^b \Phi(t) dt - \Phi \left(\frac{a+b}{2} \right) \right] \\
 &\leq \frac{\int_a^b w(t) \Phi(t) dt}{\int_a^b w(t) dt} - \Phi \left(\frac{\int_a^b w(t) t dt}{\int_a^b w(t) dt} \right) \\
 &\leq \frac{\text{ess sup}_{x \in [a,b]} w(x)}{\frac{1}{b-a} \int_a^b w(t) dt} \left[\frac{1}{b-a} \int_a^b \Phi(t) dt - \Phi \left(\frac{a+b}{2} \right) \right].
 \end{aligned}$$

Now we consider another functional depending on the weights

$$K(w; \Phi, f) := \frac{J(w; \Phi, f)}{\int_{\Omega} w d\mu} = \frac{\int_{\Omega} w (\Phi \circ f) d\mu}{\int_{\Omega} w d\mu} - \Phi \left(\frac{\int_{\Omega} w f d\mu}{\int_{\Omega} w d\mu} \right) \geq 0$$

and the composite functional

$$L(w; \Phi, f) := \left(\int_{\Omega} w d\mu \right) \ln [K(w; \Phi, f) + 1] \geq 0,$$

where $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I and $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_w(\Omega, \mu)$.

Theorem 5. *With the assumptions of Theorem 4, L is a superadditive and monotonic nondecreasing functional of weights.*

Proof. Let $w_i : \Omega \rightarrow \mathbb{R}$, with $w_i(x) \geq 0$ for μ -a.e. $x \in \Omega$ and $\int_{\Omega} w_i d\mu > 0, i \in \{1, 2\}$ such that $f, \Phi \circ f \in L_{w_1}(\Omega, \mu) \cap L_{w_2}(\Omega, \mu)$.

Utilising the superadditivity property of J we have

$$\begin{aligned} (2.11) \quad & L(w_1 + w_2; \Phi, f) \\ &= \left(\int_{\Omega} (w_1 + w_2) d\mu \right) \ln [K(w_1 + w_2; \Phi, f) + 1] \\ &= \left(\int_{\Omega} (w_1 + w_2) d\mu \right) \ln \left[\frac{J(w_1 + w_2; \Phi, f)}{\int_{\Omega} (w_1 + w_2) d\mu} + 1 \right] \\ &\geq \left(\int_{\Omega} (w_1 + w_2) d\mu \right) \ln \left[\frac{J(w_1; \Phi, f) + J(w_2; \Phi, f)}{\int_{\Omega} (w_1 + w_2) d\mu} + 1 \right] \\ &= \left(\int_{\Omega} (w_1 + w_2) d\mu \right) \\ &\quad \times \ln \left[\frac{\int_{\Omega} w_1 d\mu \cdot \frac{J(w_1; \Phi, f)}{\int_{\Omega} w_1 d\mu} + \int_{\Omega} w_2 d\mu \cdot \frac{J(w_2; \Phi, f)}{\int_{\Omega} w_2 d\mu}}{\int_{\Omega} (w_1 + w_2) d\mu} + 1 \right] \\ &= \left(\int_{\Omega} (w_1 + w_2) d\mu \right) \\ &\quad \times \ln \left[\frac{\int_{\Omega} w_1 d\mu \cdot \left(\frac{J(w_1; \Phi, f)}{\int_{\Omega} w_1 d\mu} + 1 \right) + \int_{\Omega} w_2 d\mu \cdot \left(\frac{J(w_2; \Phi, f)}{\int_{\Omega} w_2 d\mu} + 1 \right)}{\int_{\Omega} (w_1 + w_2) d\mu} \right] \\ &:= A. \end{aligned}$$

By the weighted arithmetic mean - geometric mean inequality we have

$$\begin{aligned} & \frac{\int_{\Omega} w_1 d\mu \cdot \left(\frac{J(w_1; \Phi, f)}{\int_{\Omega} w_1 d\mu} + 1 \right) + \int_{\Omega} w_2 d\mu \cdot \left(\frac{J(w_2; \Phi, f)}{\int_{\Omega} w_2 d\mu} + 1 \right)}{\int_{\Omega} (w_1 + w_2) d\mu} \\ & \geq \left(\frac{J(w_1; \Phi, f)}{\int_{\Omega} w_1 d\mu} + 1 \right)^{\frac{\int_{\Omega} w_1 d\mu}{\int_{\Omega} (w_1 + w_2) d\mu}} \left(\frac{J(w_2; \Phi, f)}{\int_{\Omega} w_2 d\mu} + 1 \right)^{\frac{\int_{\Omega} w_2 d\mu}{\int_{\Omega} (w_1 + w_2) d\mu}}, \end{aligned}$$

therefore, by taking the logarithm and utilizing the definition of the functional K , we get the inequality

$$\begin{aligned} (2.12) \quad & A \geq \left(\int_{\Omega} w_1 d\mu \right) \ln (K(w_1; \Phi, f) + 1) + \left(\int_{\Omega} w_2 d\mu \right) \ln (K(w_2; \Phi, f) + 1) \\ & = L(w_1; \Phi, f) + L(w_2; \Phi, f). \end{aligned}$$

Utilising (2.11) and (2.12) we deduce the superadditivity of the functional L as a function of weights.

Since $L(w; \Phi, f) \geq 0$ for any weight w and it is superadditive, by employing a similar argument to the one in the proof of Theorem 4 we conclude that it is also monotonic nondecreasing as a function of weights. \square

The following result provides another refinement and reverse of the Jensen inequality:

Corollary 3. *Let $w_i : \Omega \rightarrow \mathbb{R}$ with $w_i(x) \geq 0$ for μ -a.e. $x \in \Omega$, $\int_{\Omega} w_i d\mu > 0$, $i \in \{1, 2\}$ and there exists the nonnegative constants γ, Γ such that*

$$0 \leq \gamma \leq \frac{w_2}{w_1} \leq \Gamma < \infty \quad \mu\text{-a.e. on } \Omega.$$

If $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_{w_1}(\Omega, \mu) \cap L_{w_2}(\Omega, \mu)$, then

$$\begin{aligned} (2.13) \quad 0 &\leq \left[\frac{\int_{\Omega} w_1 (\Phi \circ f) d\mu}{\int_{\Omega} w_1 d\mu} - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) + 1 \right]^{\gamma \cdot \frac{(\int_{\Omega} w_1 d\mu)}{(\int_{\Omega} w_2 d\mu)}} - 1 \\ &\leq \frac{\int_{\Omega} w_2 (\Phi \circ f) d\mu}{\int_{\Omega} w_2 d\mu} - \Phi \left(\frac{\int_{\Omega} w_2 f d\mu}{\int_{\Omega} w_2 d\mu} \right) \\ &\leq \left[\frac{\int_{\Omega} w_1 (\Phi \circ f) d\mu}{\int_{\Omega} w_1 d\mu} - \Phi \left(\frac{\int_{\Omega} w_1 f d\mu}{\int_{\Omega} w_1 d\mu} \right) + 1 \right]^{\Gamma \cdot \frac{(\int_{\Omega} w_1 d\mu)}{(\int_{\Omega} w_2 d\mu)}} - 1. \end{aligned}$$

Proof. Since L is monotonic nondecreasing and positive homogeneous as a function of weights, we have

$$\gamma L(w_1; \Phi, f) \leq L(w_2; \Phi, f) \leq \Gamma L(w_1; \Phi, f)$$

which is equivalent with

$$\begin{aligned} [K(w_1; \Phi, f) + 1]^{\gamma (\int_{\Omega} w_1 d\mu)} &\leq [K(w_2; \Phi, f) + 1]^{(\int_{\Omega} w_2 d\mu)} \\ &\leq [K(w_1; \Phi, f) + 1]^{\Gamma (\int_{\Omega} w_1 d\mu)} \end{aligned}$$

which provides that

$$\begin{aligned} [K(w_1; \Phi, f) + 1]^{\gamma \frac{(\int_{\Omega} w_1 d\mu)}{(\int_{\Omega} w_2 d\mu)}} - 1 &\leq K(w_2; \Phi, f) \\ &\leq [K(w_1; \Phi, f) + 1]^{\Gamma \frac{(\int_{\Omega} w_1 d\mu)}{(\int_{\Omega} w_2 d\mu)}} - 1. \end{aligned}$$

\square

Remark 3. *Assume that $\mu(\Omega) < \infty$ and let $w : \Omega \rightarrow \mathbb{R}$, with $w(x) \geq 0$ for μ -a.e. $x \in \Omega$, $\int_{\Omega} w d\mu > 0$ and w is essentially bounded, i.e. $\text{ess inf}_{x \in \Omega} w(x)$ and $\text{ess sup}_{x \in \Omega} w(x)$ are finite.*

If $\Phi : I \rightarrow \mathbb{R}$ is a continuous convex function on the interval of real numbers I , $f : \Omega \rightarrow \mathbb{R}$ is μ -measurable and such that $f, \Phi \circ f \in L_w(\Omega, \mu) \cap L(\Omega, \mu)$, then

$$(2.14) \quad \begin{aligned} 0 &\leq \left[\frac{\int_{\Omega} (\Phi \circ f) d\mu}{\mu(\Omega)} - \Phi \left(\frac{\int_{\Omega} f d\mu}{\mu(\Omega)} \right) + 1 \right] \frac{\operatorname{ess\,inf}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} - 1 \\ &\leq \frac{\int_{\Omega} w (\Phi \circ f) d\mu}{\int_{\Omega} w d\mu} - \Phi \left(\frac{\int_{\Omega} w f d\mu}{\int_{\Omega} w d\mu} \right) \\ &\leq \left[\frac{\int_{\Omega} (\Phi \circ f) d\mu}{\mu(\Omega)} - \Phi \left(\frac{\int_{\Omega} f d\mu}{\mu(\Omega)} \right) + 1 \right] \frac{\operatorname{ess\,sup}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} - 1. \end{aligned}$$

In particular, if $w : [a, b] \rightarrow [0, \infty)$ is Lebesgue integrable, then we have the following result related to the Hermite-Hadamard inequality for the convex function $\Phi : [a, b] \rightarrow \mathbb{R}$

$$(2.15) \quad \begin{aligned} 0 &\leq \left[\frac{1}{b-a} \int_a^b \Phi(t) dt - \Phi \left(\frac{a+b}{2} \right) + 1 \right] \frac{\operatorname{ess\,inf}_{x \in [a,b]} w(x)}{\frac{1}{b-a} \int_a^b w(t) dt} - 1 \\ &\leq \frac{\int_a^b w(t) \Phi(t) dt}{\int_a^b w(t) dt} - \Phi \left(\frac{\int_a^b w(t) t dt}{\int_a^b w(t) dt} \right) \\ &\leq \left[\frac{1}{b-a} \int_a^b \Phi(t) dt - \Phi \left(\frac{a+b}{2} \right) + 1 \right] \frac{\operatorname{ess\,sup}_{x \in [a,b]} w(x)}{\frac{1}{b-a} \int_a^b w(t) dt} - 1. \end{aligned}$$

3. APPLICATIONS FOR THE HÖLDER INEQUALITY

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

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

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

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

Assume that $p > 1$. If $h : \Omega \rightarrow \mathbb{R}$ is μ -measurable, $\mu(\Omega) < \infty$, $|h|, |h|^p \in L_w(\Omega, \mu) \cap L(\Omega, \mu)$, then by (2.9) we have the bounds

$$(3.1) \quad \begin{aligned} 0 &\leq \frac{\operatorname{ess\,inf}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} \left[\frac{1}{\mu(\Omega)} \int_{\Omega} |h|^p d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} |h| d\mu \right)^p \right] \\ &\leq \frac{1}{\int_{\Omega} w d\mu} \int_{\Omega} w |h|^p d\mu - \left(\frac{1}{\int_{\Omega} w d\mu} \int_{\Omega} w |h| d\mu \right)^p \\ &\leq \frac{\operatorname{ess\,sup}_{x \in \Omega} w(x)}{\frac{1}{\mu(\Omega)} \int_{\Omega} w d\mu} \left[\frac{1}{\mu(\Omega)} \int_{\Omega} |h|^p d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} |h| d\mu \right)^p \right]. \end{aligned}$$

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

$$(3.2) \quad \delta \leq |g| \leq \Delta \text{ } \mu\text{-a.e on } \Omega,$$

then we have

$$\begin{aligned}
 (3.3) \quad 0 &\leq \frac{\delta^q}{\frac{1}{\mu(\Omega)} \int_{\Omega} |g|^q d\mu} \left[\frac{1}{\mu(\Omega)} \int_{\Omega} \frac{|f|^p}{|g|^q} d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} \frac{|f|}{|g|^{q-1}} d\mu \right)^p \right] \\
 &\leq \frac{\int_{\Omega} |f|^p d\mu}{\int_{\Omega} |g|^q d\mu} - \left(\frac{\int_{\Omega} |fg| d\mu}{\int_{\Omega} |g|^q d\mu} \right)^p \\
 &\leq \frac{\Delta^q}{\frac{1}{\mu(\Omega)} \int_{\Omega} |g|^q d\mu} \left[\frac{1}{\mu(\Omega)} \int_{\Omega} \frac{|f|^p}{|g|^q} d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} \frac{|f|}{|g|^{q-1}} d\mu \right)^p \right].
 \end{aligned}$$

Proof. The inequalities (3.3) follows from (3.1) by choosing

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

The details are omitted. \square

Remark 4. We observe that for $p = q = 2$ we have from (3.3) the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality

$$\begin{aligned}
 (3.4) \quad 0 &\leq \delta^2 \mu(\Omega) \left[\frac{1}{\mu(\Omega)} \int_{\Omega} \left| \frac{f}{g} \right|^2 d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} \left| \frac{f}{g} \right| d\mu \right)^2 \right] \int_{\Omega} |g|^2 d\mu \\
 &\leq \int_{\Omega} |g|^2 d\mu \int_{\Omega} |f|^2 d\mu - \left(\int_{\Omega} |fg| d\mu \right)^2 \\
 &\leq \Delta^2 \mu(\Omega) \left[\frac{1}{\mu(\Omega)} \int_{\Omega} \left| \frac{f}{g} \right|^2 d\mu - \left(\frac{1}{\mu(\Omega)} \int_{\Omega} \left| \frac{f}{g} \right| d\mu \right)^2 \right] \int_{\Omega} |g|^2 d\mu,
 \end{aligned}$$

provided that $f, g \in L_2(\Omega, \mu)$ and g satisfies the bounds (3.2).

Similar results can be stated by utilizing the inequality (2.13), however the details are not presented here.

4. APPLICATIONS FOR f -DIVERGENCE MEASURES

One of the important issues in many applications of Probability Theory is finding an appropriate measure of *distance* (or *difference* or *discrimination*) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [20], Kullback and Leibler [25], Rényi [31], Havrda and Charvat [18], Kapur [23], Sharma and Mittal [33], Burbea and Rao [4], Rao [30], Lin [26], Csiszár [7], Ali and Silvey [1], Vajda [40], Shioya and Da-te [35] and others (see for example [27] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [30], genetics [27], finance, economics, and political science [32], [38], [39], biology [29], the analysis of contingency tables [17], approximation of probability distributions [6], [24], signal processing [21], [22] and pattern recognition [3], [5]. A number of these measures of distance are specific cases of Csiszár f -divergence and so further exploration of this concept will have a flow on effect to other measures of distance and to areas in which they are applied.

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

Csiszár f -divergence is defined as follows [8]

$$(4.1) \quad I_f(p, q) := \int_{\Omega} p(x) f \left[\frac{q(x)}{p(x)} \right] d\mu(x), \quad p, q \in \mathcal{P},$$

where f is convex on $(0, \infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u = 1$. By appropriately defining this convex function, various divergences are derived.

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

$$(4.2) \quad D_{KL}(p, q) := \int_{\Omega} p(x) \ln \left[\frac{p(x)}{q(x)} \right] d\mu(x), \quad p, q \in \mathcal{P},$$

where \ln is to base e .

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are the: *variation distance* D_v , *Hellinger distance* D_H [19], χ^2 -*divergence* D_{χ^2} , α -*divergence* D_{α} , *Bhattacharyya distance* D_B [2], *Harmonic distance* $D_{H\alpha}$, *Jeffrey's distance* D_J [20], *triangular discrimination* D_{Δ} [37], etc... They are defined as follows:

$$(4.3) \quad D_v(p, q) := \int_{\Omega} |p(x) - q(x)| d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.4) \quad D_H(p, q) := \int_{\Omega} \left| \sqrt{p(x)} - \sqrt{q(x)} \right| d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.5) \quad D_{\chi^2}(p, q) := \int_{\Omega} p(x) \left[\left(\frac{q(x)}{p(x)} \right)^2 - 1 \right] d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.6) \quad D_{\alpha}(p, q) := \frac{4}{1 - \alpha^2} \left[1 - \int_{\Omega} [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x) \right], \quad p, q \in \mathcal{P};$$

$$(4.7) \quad D_B(p, q) := \int_{\Omega} \sqrt{p(x)q(x)} d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.8) \quad D_{H\alpha}(p, q) := \int_{\Omega} \frac{2p(x)q(x)}{p(x) + q(x)} d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.9) \quad D_J(p, q) := \int_{\Omega} [p(x) - q(x)] \ln \left[\frac{p(x)}{q(x)} \right] d\mu(x), \quad p, q \in \mathcal{P};$$

$$(4.10) \quad D_{\Delta}(p, q) := \int_{\Omega} \frac{[p(x) - q(x)]^2}{p(x) + q(x)} d\mu(x), \quad p, q \in \mathcal{P}.$$

For other divergence measures, see the paper [23] by Kapur or the book on line [36] by Taneja.

Most of the above distances (4.2) – (4.10), are particular instances of Csiszár f -divergence. There are also many others which are not in this class (see for example [36]). For the basic properties of Csiszár f -divergence see [8], [9] and [40].

The following result holds:

Proposition 2. *Let $f : (0, \infty) \rightarrow \mathbb{R}$ be a convex function with the property that $f(1) = 0$. Assume that $p, q \in \mathcal{P}$ and there exists the constants $0 < s < 1 < S < \infty$ such that*

$$(4.11) \quad s \leq \frac{p(x)}{q(x)} \leq S \text{ for } \mu\text{-a.e. } x \in \Omega.$$

Then we have the inequalities

$$(4.12) \quad \begin{aligned} s \left[I_{f(\frac{\cdot}{\cdot})}(q, p) - f(D_{\chi^2}(p, q) + 1) \right] \\ \leq I_f(p, q) \\ \leq S \left[I_{f(\frac{\cdot}{\cdot})}(q, p) - f(D_{\chi^2}(p, q) + 1) \right]. \end{aligned}$$

Proof. If we use the inequality (2.6) we get

$$(4.13) \quad \begin{aligned} s \left[\int_{\Omega} q f\left(\frac{q}{p}\right) d\mu - f\left(\int_{\Omega} \frac{q^2}{p} d\mu\right) \right] \\ \leq \int_{\Omega} p f\left(\frac{q}{p}\right) d\mu \\ \leq S \left[\int_{\Omega} q f\left(\frac{q}{p}\right) d\mu - f\left(\int_{\Omega} \frac{q^2}{p} d\mu\right) \right]. \end{aligned}$$

Since

$$\int_{\Omega} \frac{q^2}{p} d\mu = D_{\chi^2}(p, q) + 1$$

and

$$\int_{\Omega} q f\left(\frac{q}{p}\right) d\mu = I_{f(\frac{\cdot}{\cdot})}(q, p),$$

then from (4.13) we deduce the desired result (4.12). \square

Consider the Kullback-Leibler divergence

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

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

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

$$(4.14) \quad s \leq \frac{p(x)}{q(x)} \leq S \text{ for } \mu\text{-a.e. } x \in \Omega..$$

then we get from (4.12) that

$$(4.15) \quad \begin{aligned} s \left[\ln(D_{\chi^2}(p, q) + 1) - D_{KL}(q, p) \right] \\ \leq D_{KL}(p, q) \\ \leq S \left[\ln(D_{\chi^2}(p, q) + 1) - D_{KL}(q, p) \right]. \end{aligned}$$

Similar results for f -divergence measures can be stated by utilizing the inequality (2.13), however the details are not presented here.

REFERENCES

- [1] S.M. ALI and S.D. SILVEY, A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. Sec B*, **28** (1966), 131-142.
- [2] A. BHATTACHARYYA, On a measure of divergence between two statistical populations defined by their probability distributions, *Bull. Calcutta Math. Soc.*, **35** (1943), 99-109.
- [3] M. BETH BASSAT, f -entropies, probability of error and feature selection, *Inform. Control*, **39** (1978), 227-242.
- [4] I. BURBEA and C.R. RAO, On the convexity of some divergence measures based on entropy function, *IEEE Trans. Inf. Th.*, **28** (3) (1982), 489-495.
- [5] C.H. CHEN, *Statistical Pattern Recognition*, Rocelle Park, New York, Hoyderc Book Co., 1973.
- [6] C.K. CHOW and C.N. LIN, Approximating discrete probability distributions with dependence trees, *IEEE Trans. Inf. Th.*, **14** (3) (1968), 462-467.
- [7] I. CSISZÁR, Information-type measures of difference of probability distributions and indirect observations, *Studia Math. Hungarica*, **2** (1967), 299-318.
- [8] I. CSISZÁR, On topological properties of f -divergences, *Studia Math. Hungarica*, **2** (1967), 329-339.
- [9] I. CSISZÁR and J. KÖRNER, *Information Theory: Coding Theorem for Discrete Memoryless Systems*, Academic Press, New York, 1981.
- [10] S.S. DRAGOMIR, A converse result for Jensen's discrete inequality via Gruss' inequality and applications in information theory. *An. Univ. Oradea Fasc. Mat.* **7** (1999/2000), 178-189.
- [11] S.S. DRAGOMIR, On a reverse of Jessen's inequality for isotonic linear functionals, *J. Ineq. Pure & Appl. Math.*, **2**(2001), No. 3, Article 36.
- [12] S.S. DRAGOMIR, A Grüss type inequality for isotonic linear functionals and applications. *Demonstratio Math.* **36** (2003), no. 3, 551-562. Preprint RGMIA Res. Rep. Coll. **5**(2002), Supplement, Art. 12. [Online [http://rgmia.org/v5\(E\).php](http://rgmia.org/v5(E).php)].
- [13] S.S. DRAGOMIR, Bounds for the deviation of a function from the chord generated by its extremities. *Bull. Aust. Math. Soc.* **78** (2008), no. 2, 225-248.
- [14] S.S. DRAGOMIR, Bounds for the normalized Jensen functional, *Bull. Austral. Math. Soc.* **74**(3)(2006), 471-476.
- [15] S.S. DRAGOMIR, Some reverses of the Jensen inequality with applications, Preprint *RGMIA Res. Rep. Coll.* **14**(2011), Art. 72. [Online <http://rgmia.org/v14.php>].
- [16] S.S. DRAGOMIR and N.M. IONESCU, Some converse of Jensen's inequality and applications. *Rev. Anal. Numér. Théor. Approx.* **23** (1994), no. 1, 71-78.
- [17] D.V. GOKHALE and S. KULLBACK, *Information in Contingency Tables*, New York, Marcel Decker, 1978.
- [18] J.H. HAVRDA and F. CHARVAT, Quantification method classification process: concept of structural α -entropy, *Kybernetika*, **3** (1967), 30-35.
- [19] E. HELLINGER, Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen, *J. für reine und Angew. Math.*, **36** (1909), 210-271.
- [20] H. JEFFREYS, An invariant form for the prior probability in estimating problems, *Proc. Roy. Soc. London*, **186** A (1946), 453-461.
- [21] T.T. KADOTA and L.A. SHEPP, On the best finite set of linear observables for discriminating two Gaussian signals, *IEEE Trans. Inf. Th.*, **13** (1967), 288-294.
- [22] T. KAILATH, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Technology.*, Vol COM-15 (1967), 52-60.
- [23] J.N. KAPUR, A comparative assessment of various measures of directed divergence, *Advances in Management Studies*, **3** (1984), 1-16.
- [24] D. KAZAKOS and T. COTSIDAS, A decision theory approach to the approximation of discrete probability densities, *IEEE Trans. Perform. Anal. Machine Intell.*, **1** (1980), 61-67.
- [25] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Annals Math. Statist.*, **22** (1951), 79-86.
- [26] J. LIN, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.*, **37** (1) (1991), 145-151.
- [27] M. MEI, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.

- [28] C.P. NICULESCU, An extension of Chebyshev's inequality and its connection with Jensen's inequality. (English summary) *J. Inequal. Appl.* **6** (2001), no. 4, 451–462.
- [29] E.C. PIELOU, *Ecological Diversity*, Wiley, New York, 1975.
- [30] C.R. RAO, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.
- [31] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.
- [32] A. SEN, *On Economic Inequality*, Oxford University Press, London 1973.
- [33] B.D. SHARMA and D.P. MITTAL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.*, **2** (4)(1977), 122-132.
- [34] S. SIMIC, On a global upper bound for Jensen's inequality, *J. Math. Anal. Appl.* **343**(2008), 414-419.
- [35] 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.
- [36] I. J. TANEJA, *Generalised Information Measures and their Applications* (<http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html>).
- [37] F. TOPSOE, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll., RGMIA*, **2** (1) (1999), 85-98.
- [38] H. THEIL, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.
- [39] H. THEIL, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.
- [40] I. VAJDA, *Theory of Statistical Inference and Information*, Dordrecht-Boston, Kluwer Academic Publishers, 1989.

¹MATHEMATICS, SCHOOL 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`

²SCHOOL OF COMPUTATIONAL & APPLIED MATHEMATICS, UNIVERSITY OF THE WITWATERSRAND, PRIVATE BAG 3, JOHANNESBURG 2050, SOUTH AFRICA.