Received 30/08/11

SOME REVERSES OF THE JENSEN INEQUALITY WITH APPLICATIONS

S.S. DRAGOMIR[1,2]

Abstract. Two new reverses of the celebrated Jensen's inequality for convex functions in the general settings of the Lebesgue integral with applications for means, 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 $\mu-$measurable function $w : \Omega \to \mathbb{R}$, with $w(x) \geq 0$ for μ − a.e. (almost every) $x \in \Omega$, consider the Lebesgue space

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

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

If $f, g : \Omega \to \mathbb{R}$ are $\mu-$measurable functions and $f, g, fg \in L_w(\Omega, \mu)$, then we may consider the *Čebyšev functional*

$$T_w(f, g) := \int_\Omega wfg d\mu - \int_\Omega wf d\mu \int_\Omega wg d\mu. \tag{1.1}$$

The following result is known in the literature as the *Grüss inequality*

$$|T_w(f, g)| \leq \frac{1}{4}(\Gamma - \gamma)(\Delta - \delta), \tag{1.2}$$

provided

$$-\infty < \gamma \leq f(x) \leq \Gamma < \infty,\ \ -\infty < \delta \leq g(x) \leq \Delta < \infty \tag{1.3}$$

for μ − a.e. (almost every) $x \in \Omega$.

The constant $\frac{1}{4}$ is sharp in the sense that it cannot be replaced by a smaller constant.

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.

If we assume that $-\infty < \gamma \leq f(x) \leq \Gamma < \infty$ for $\mu-$a.e. $x \in \Omega$, then by the Grüss inequality for $g = f$ and by the Schwarz's integral inequality, we have

$$\int_\Omega w \left| f - \int_\Omega wf d\mu \right| d\mu \leq \left[\int_\Omega wf^2 d\mu - \left(\int_\Omega wf d\mu \right)^2 \right]^{\frac{1}{2}} \leq \frac{1}{2}(\Gamma - \gamma). \tag{1.4}$$

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 $\Phi : [m, M] \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on (m, M) and $f : \Omega \to [m, M]$ 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:*

$$\begin{aligned} 0 &\leq \int_\Omega w (\Phi \circ f) d\mu - \Phi\left(\int_\Omega wf d\mu\right) \\ &\leq \int_\Omega w (\Phi' \circ f) f d\mu - \int_\Omega w (\Phi' \circ f) d\mu \int_\Omega wf d\mu \\ &\leq \frac{1}{2}[\Phi'(M) - \Phi'(m)] \int_\Omega w \left| f - \int_\Omega wf d\mu \right| d\mu. \end{aligned} \tag{1.5}$$

For a generalization of the first inequality in (1.5) without the differentiability assumption and the derivative Φ' replaced with a selection φ from the subdifferential $\partial\Phi$, see the paper [27] by C.P. Niculescu.

If $\mu(\Omega) < \infty$ and $\Phi \circ f,\ f,\ \Phi' \circ f,\ (\Phi' \circ f) \cdot f \in L(\Omega, \mu)$, then we have the inequality:

$$\begin{aligned} 0 &\leq \frac{1}{\mu(\Omega)} \int_\Omega (\Phi \circ f) d\mu - \Phi\left(\frac{1}{\mu(\Omega)} \int_\Omega f d\mu\right) \\ &\leq \frac{1}{\mu(\Omega)} \int_\Omega (\Phi' \circ f) f d\mu - \frac{1}{\mu(\Omega)} \int_\Omega (\Phi' \circ f) d\mu \cdot \frac{1}{\mu(\Omega)} \int_\Omega f d\mu \\ &\leq \frac{1}{2}[\Phi'(M) - \Phi'(m)] \frac{1}{\mu(\Omega)} \int_\Omega \left| f - \frac{1}{\mu(\Omega)} \int_\Omega f d\mu \right| d\mu. \end{aligned} \tag{1.6}$$

The following discrete inequality is of interest as well.

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

$$\begin{aligned} 0 &\leq \sum_{i=1}^n w_i \Phi(x_i) - \Phi\left(\sum_{i=1}^n w_i x_i\right) \\ &\leq \sum_{i=1}^n w_i \Phi'(x_i) x_i - \sum_{i=1}^n w_i \Phi'(x_i) \sum_{i=1}^n w_i x_i \\ &\leq \frac{1}{2}[\Phi'(M) - \Phi'(m)] \sum_{i=1}^n w_i \left| x_i - \sum_{j=1}^n w_j x_j \right|. \end{aligned} \tag{1.7}$$

Remark 1. *We notice that the inequality between the first and the second term in (1.7) was proved in 1994 by Dragomir & Ionescu, see* [15].

On making use of the results (1.5) and (1.4), we can state the following string of reverse inequalities

$$
\begin{aligned}
0 &\leq \int_{\Omega} w\,(\Phi \circ f)\, d\mu - \Phi\left(\int_{\Omega} w f d\mu\right) \\
&\leq \int_{\Omega} w\,(\Phi' \circ f)\, f d\mu - \int_{\Omega} w\,(\Phi' \circ f)\, d\mu \int_{\Omega} w f d\mu \\
&\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \int_{\Omega} w \left| f - \int_{\Omega} w f d\mu \right| d\mu \\
&\leq \frac{1}{2}\left[\Phi'(M) - \Phi'(m)\right] \left[\int_{\Omega} w f^2 d\mu - \left(\int_{\Omega} w f d\mu\right)^2\right]^{\frac{1}{2}} \\
&\leq \frac{1}{4}\left[\Phi'(M) - \Phi'(m)\right](M - m),
\end{aligned} \tag{1.8}
$$

provided that $\Phi : [m, M] \subset \mathbb{R} \to \mathbb{R}$ is a differentiable convex function on (m, M) and $f : \Omega \to [m, M]$ 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. on Ω with $\int_{\Omega} w d\mu = 1$.

Remark 2. *We notice that the inequality between the first, second and last term from (1.8) was proved in the general case of positive linear functionals in 2001 by S.S. Dragomir in* [11].

Motivated by the above results, we establish in the current paper two new reverses of Jensen's integral inequality for a convex function. Some natural application for inequalities between means, reverses of Hölder's inequality and for the f-divergence measure that play an important role in information theory are given as well.

2. Reverse Inequalities

The following reverse of the Jensen's inequality holds:

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

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

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

$$
\begin{aligned}
0 &\leq \int_{\Omega} w\,(\Phi \circ f)\, d\mu - \Phi\left(\bar{f}_{\Omega,w}\right) \\
&\leq \frac{\left(M - \bar{f}_{\Omega,w}\right)\left(\bar{f}_{\Omega,w} - m\right)}{M - m} \sup_{t \in (m,M)} \Psi_{\Phi}(t; m, M) \\
&\leq \left(M - \bar{f}_{\Omega,w}\right)\left(\bar{f}_{\Omega,w} - m\right) \frac{\Phi'_{-}(M) - \Phi'_{+}(m)}{M - m} \\
&\leq \frac{1}{4}(M - m)\left[\Phi'_{-}(M) - \Phi'_{+}(m)\right],
\end{aligned} \tag{2.1}
$$

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

$$0 \le \int_\Omega w (\Phi \circ f) d\mu - \Phi \left(\bar{f}_{\Omega,w}\right) \le \frac{1}{4} (M - m) \Psi_\Phi \left(\bar{f}_{\Omega,w}; m, M\right) \le \frac{1}{4} (M - m) \left[\Phi'_- (M) - \Phi'_+ (m)\right], \tag{2.2}$$

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

Proof. By the convexity of Φ we have that

$$\begin{aligned} &\int_\Omega w (x) \Phi (f (x)) d\mu (x) - \Phi \left(\bar{f}_{\Omega,w}\right) \\ &= \int_\Omega w (x) \Phi \left[\frac{m (M - f (x)) + M (f (x) - m)}{M - m}\right] d\mu (x) \\ &- \Phi \left(\int_\Omega w (x) \left[\frac{m (M - f (x)) + M (f (x) - m)}{M - m}\right] d\mu (x)\right) \\ &\le \int_\Omega \frac{(M - f (x)) \Phi (m) + (f (x) - m) \Phi (M)}{M - m} w (x) d\mu (x) \\ &- \Phi \left(\frac{m \left(M - \bar{f}_{\Omega,w}\right) + M \left(\bar{f}_{\Omega,w} - m\right)}{M - m}\right) \\ &= \frac{\left(M - \bar{f}_{\Omega,w}\right) \Phi (m) + \left(\bar{f}_{\Omega,w} - m\right) \Phi (M)}{M - m} \\ &- \Phi \left(\frac{m \left(M - \bar{f}_{\Omega,w}\right) + M \left(\bar{f}_{\Omega,w} - m\right)}{M - m}\right) := B. \end{aligned} \tag{2.3}$$

By denoting

$$\Delta_\Phi (t; m, M) := \frac{(t - m) \Phi (M) + (M - t) \Phi (m)}{M - m} - \Phi (t), \quad t \in [m, M]$$

we have

$$\begin{aligned} \Delta_\Phi (t; m, M) &= \frac{(t - m) \Phi (M) + (M - t) \Phi (m) - (M - m) \Phi (t)}{M - m} \\ &= \frac{(t - m) \Phi (M) + (M - t) \Phi (m) - (M - t + t - m) \Phi (t)}{M - m} \\ &= \frac{(t - m) [\Phi (M) - \Phi (t)] - (M - t) [\Phi (t) - \Phi (m)]}{M - m} \\ &= \frac{(M - t) (t - m)}{M - m} \Psi_\Phi (t; m, M) \end{aligned} \tag{2.4}$$

for any $t \in (m, M)$.

Therefore we have the equality

$$B = \frac{\left(M - \bar{f}_{\Omega,w}\right) \left(\bar{f}_{\Omega,w} - m\right)}{M - m} \Psi_\Phi \left(\bar{f}_{\Omega,w}; m, M\right) \tag{2.5}$$

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

For $\bar{f}_{\Omega,w} = m$ or $\bar{f}_{\Omega,w} = M$ the inequality (2.1) is obvious. If $\bar{f}_{\Omega,w} \in (m, M)$, then

$$\begin{aligned}
\Psi_\Phi\left(\bar{f}_{\Omega,w}; m, M\right) &\leq \sup_{t\in(m,M)} \Psi_\Phi(t; m, M) \\
&= \sup_{t\in(m,M)} \left[\frac{\Phi(M) - \Phi(t)}{M - t} - \frac{\Phi(t) - \Phi(m)}{t - m}\right] \\
&\leq \sup_{t\in(m,M)} \left[\frac{\Phi(M) - \Phi(t)}{M - t}\right] + \sup_{t\in(m,M)} \left[-\frac{\Phi(t) - \Phi(m)}{t - m}\right] \\
&= \sup_{t\in(m,M)} \left[\frac{\Phi(M) - \Phi(t)}{M - t}\right] - \inf_{t\in(m,M)} \left[\frac{\Phi(t) - \Phi(m)}{t - m}\right] \\
&= \Phi'_-(M) - \Phi'_+(m)
\end{aligned}$$

which by (2.3) and (2.5) produces the desired result (2.1).

Since, obviously

$$\frac{\left(M - \bar{f}_{\Omega,w}\right)\left(\bar{f}_{\Omega,w} - m\right)}{M - m} \leq \frac{1}{4}(M - m),$$

then by (2.3) and (2.5) we deduce the first inequality (2.2). The second part is clear. □

Corollary 2. *Let $\Phi : I \to \mathbb{R}$ be a continuous convex function on the interval of real numbers I and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \mathring{I}$. If $x_i \in I$ and $p_i \geq 0$ for $i \in \{1, ..., n\}$ with $\sum_{i=1}^n p_i = 1$, then we have the inequalities*

$$\begin{aligned}
0 &\leq \sum_{i=1}^{n} p_i \Phi(x_i) - \Phi(\bar{x}_p) \\
&\leq (M - \bar{x}_p)(\bar{x}_p - m) \sup_{t\in(m,M)} \Psi_\Phi(t; m, M) \\
&\leq (M - \bar{x}_p)(\bar{x}_p - m) \frac{\Phi'_-(M) - \Phi'_+(m)}{M - m} \\
&\leq \frac{1}{4}(M - m)\left[\Phi'_-(M) - \Phi'_+(m)\right],
\end{aligned} \tag{2.6}$$

and

$$\begin{aligned}
0 &\leq \sum_{i=1}^{n} p_i \Phi(x_i) - \Phi(\bar{x}_p) \leq \frac{1}{4}(M - m)\Psi_\Phi(\bar{x}_p; m, M) \\
&\leq \frac{1}{4}(M - m)\left[\Phi'_-(M) - \Phi'_+(m)\right],
\end{aligned} \tag{2.7}$$

where $\bar{x}_p := \sum_{i=1}^n p_i x_i \in I$.

Remark 3. *Define the weighted arithmetic mean of the positive n-tuple $x = (x_1, ..., x_n)$ with the nonnegative weights $w = (w_1, ..., w_n)$ by*

$$A_n(w, x) := \frac{1}{W_n} \sum_{i=1}^{n} w_i x_i$$

where $W_n := \sum_{i=1}^{n} w_i > 0$ *and the weighted geometric mean of the same n-tuple, by*

$$G_n(w,x) := \left(\prod_{i=1}^{n} x_i^{w_i}\right)^{1/W_n}.$$

It is well know that the following arithmetic mean-geometric mean inequality holds true

$$A_n(w,x) \geq G_n(w,x).$$

Applying the inequality between the first and third term in (2.6) for the convex function $\Phi(t) = -\ln t, t > 0$ *we have*

$$1 \leq \frac{A_n(w,x)}{G_n(w,x)} \leq \exp\left[\frac{1}{Mm}(M - A_n(w,x))(A_n(w,x) - m)\right] \leq \exp\left[\frac{1}{4}\frac{(M-m)^2}{mM}\right], \tag{2.8}$$

provided that $0 < m \leq x_i \leq M < \infty$ *for* $i \in \{1, ..., n\}$.

Also, if we apply the inequality (2.7) for the same function Φ *we get that*

$$\begin{aligned} 1 &\leq \frac{A_n(w,x)}{G_n(w,x)} \\ &\leq \left[\left(\frac{M}{A_n(w,x)}\right)^{M-A_n(w,x)}\left(\frac{m}{A_n(w,x)}\right)^{A_n(w,x)-m}\right]^{\frac{1}{4}(M-m)} \\ &\leq \exp\left[\frac{1}{4}\frac{(M-m)^2}{mM}\right]. \end{aligned} \tag{2.9}$$

The following result also holds

Theorem 3. *With the assumptions of Theorem 2, we have the inequalities*

$$\begin{aligned} 0 &\leq \int_\Omega w(\Phi\circ f)\,d\mu(x) - \Phi\left(\bar{f}_{\Omega,w}\right) \\ &\leq 2\max\left\{\frac{M-\bar{f}_{\Omega,w}}{M-m}, \frac{\bar{f}_{\Omega,w}-m}{M-m}\right\}\left[\frac{\Phi(m)+\Phi(M)}{2} - \Phi\left(\frac{m+M}{2}\right)\right] \\ &\leq \frac{1}{2}\max\left\{M-\bar{f}_{\Omega,w}, \bar{f}_{\Omega,w}-m\right\}\left[\Phi'_-(M) - \Phi'_+(m)\right]. \end{aligned} \tag{2.10}$$

Proof. First of all, we recall the following result obtained by the author in [14] that provides a refinement and a reverse for the weighted Jensen's discrete inequality:

$$\begin{aligned} &n\min_{i\in\{1,\ldots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^{n}\Phi(x_i) - \Phi\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\right] \\ &\leq \frac{1}{P_n}\sum_{i=1}^{n}p_i\Phi(x_i) - \Phi\left(\frac{1}{P_n}\sum_{i=1}^{n}p_ix_i\right) \\ &n\max_{i\in\{1,\ldots,n\}}\{p_i\}\left[\frac{1}{n}\sum_{i=1}^{n}\Phi(x_i) - \Phi\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)\right], \end{aligned} \tag{2.11}$$

where $\Phi : C \to \mathbb{R}$ is a convex function defined on the convex subset C of the linear space X, $\{x_i\}_{i\in\{1,\dots,n\}}$ are vectors and $\{p_i\}_{i\in\{1,\dots,n\}}$ are nonnegative numbers with $P_n := \sum_{i=1}^n p_i > 0$.

For $n = 2$ we deduce from (2.11) that

$$\begin{aligned} &2\min\{t,1-t\}\left[\frac{\Phi(x)+\Phi(y)}{2}-\Phi\left(\frac{x+y}{2}\right)\right] \\ &\leq t\Phi(x)+(1-t)\Phi(y)-\Phi(tx+(1-t)y) \\ &\leq 2\max\{t,1-t\}\left[\frac{\Phi(x)+\Phi(y)}{2}-\Phi\left(\frac{x+y}{2}\right)\right] \end{aligned} \tag{2.12}$$

for any $x, y \in C$ and $t \in [0,1]$.

If we use the second inequality in (2.12) for the convex function $\Phi : I \to \mathbb{R}$ and $m, M \in \mathbb{R}$, $m < M$ with $[m, M] \subset \mathring{I}$, we have for $t = \frac{M-\bar{f}_{\Omega,w}}{M-m}$ that

$$\begin{aligned} &\frac{\left(M-\bar{f}_{\Omega,w}\right)\Phi(m)+\left(\bar{f}_{\Omega,w}-m\right)\Phi(M)}{M-m} \\ &-\Phi\left(\frac{m\left(M-\bar{f}_{\Omega,w}\right)+M\left(\bar{f}_{\Omega,w}-m\right)}{M-m}\right) \\ &\leq 2\max\left\{\frac{M-\bar{f}_{\Omega,w}}{M-m},\frac{\bar{f}_{\Omega,w}-m}{M-m}\right\} \\ &\times\left[\frac{\Phi(m)+\Phi(M)}{2}-\Phi\left(\frac{m+M}{2}\right)\right]. \end{aligned} \tag{2.13}$$

Utilizing the inequality (2.3) and (2.13) we deduce the first inequality in (2.10).

Since

$$\begin{aligned} &\frac{\frac{\Phi(m)+\Phi(M)}{2}-\Phi\left(\frac{m+M}{2}\right)}{M-m} \\ &=\frac{1}{4}\left[\frac{\Phi(M)-\Phi\left(\frac{m+M}{2}\right)}{M-\frac{m+M}{2}}-\frac{\Phi\left(\frac{m+M}{2}\right)-\Phi(m)}{\frac{m+M}{2}-m}\right] \end{aligned}$$

and, by the gradient inequality, we have that

$$\frac{\Phi(M)-\Phi\left(\frac{m+M}{2}\right)}{M-\frac{m+M}{2}}\leq\Phi'_{-}(M)$$

and

$$\frac{\Phi\left(\frac{m+M}{2}\right)-\Phi(m)}{\frac{m+M}{2}-m}\geq\Phi'_{+}(m),$$

then we get

$$\frac{\frac{\Phi(m)+\Phi(M)}{2}-\Phi\left(\frac{m+M}{2}\right)}{M-m}\leq\frac{1}{4}\left[\Phi'_{-}(M)-\Phi'_{+}(m)\right]. \tag{2.14}$$

On making use of (2.13) and (2.14) we deduce the last part of (2.10). □

Corollary 3. *With the assumptions in Corollary 2, we have the inequalities*

$$0 \leq \sum_{i=1}^{n} p_i \Phi(x_i) - \Phi(\bar{x}_p) \leq 2\max\left\{\frac{M-\bar{x}_p}{M-m}, \frac{\bar{x}_p - m}{M-m}\right\}\left[\frac{\Phi(m)+\Phi(M)}{2} - \Phi\left(\frac{m+M}{2}\right)\right] \leq \frac{1}{2}\max\{M-\bar{x}_p, \bar{x}_p - m\}\left[\Phi'_-(M) - \Phi'_+(m)\right]. \tag{2.15}$$

Remark 4. *Since, obviously,*

$$\frac{M-\bar{f}_{\Omega,w}}{M-m}, \frac{\bar{f}_{\Omega,w}-m}{M-m} \leq 1$$

then we obtain from the first inequality in (2.10) the simpler, however coarser inequality

$$0 \leq \int_\Omega w(\Phi \circ f)\, d\mu(x) - \Phi\left(\bar{f}_{\Omega,w}\right) \leq 2\left[\frac{\Phi(m)+\Phi(M)}{2} - \Phi\left(\frac{m+M}{2}\right)\right]. \tag{2.16}$$

We notice that the discrete version of this result, namely

$$0 \leq \sum_{i=1}^{n} p_i \Phi(x_i) - \Phi(\bar{x}_p) \leq 2\left[\frac{\Phi(m)+\Phi(M)}{2} - \Phi\left(\frac{m+M}{2}\right)\right] \tag{2.17}$$

was obtained in 2008 by S. Simic in [33].

Remark 5. *With the assumptions in Remark 3 we have the following reverse of the arithmetic mean-geometric mean inequality*

$$1 \leq \frac{A_n(w,x)}{G_n(w,x)} \leq \left(\frac{A(m,M)}{G(m,M)}\right)^{2\max\left\{\frac{M-A_n(w,x)}{M-m}, \frac{A_n(w,x)-m}{M-m}\right\}}, \tag{2.18}$$

where $A(m,M)$ is the arithmetic mean while $G(m,M)$ is the geometric mean of the positive numbers m and M.

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

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

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

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

Assume that $p>1$. If $h : \Omega \to \mathbb{R}$ is μ-measurable, satisfies the bounds

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

and is such that $h, |h|^p \in L_w(\Omega,\mu)$, for a μ−measurable function $w:\Omega\to\mathbb{R}$, with $w(x)\geq 0$ for μ − a.e. $x\in\Omega$ and $\int_\Omega w d\mu > 0$, then from (2.1) we have

$$
\begin{aligned}
0 &\leq \frac{\int_\Omega |h|^p w d\mu}{\int_\Omega w d\mu} - \left(\frac{\int_\Omega |h| w d\mu}{\int_\Omega w d\mu}\right)^p \\
&\leq \frac{\left(M-\overline{|h|}_{\Omega,w}\right)\left(\overline{|h|}_{\Omega,w}-m\right)}{M-m} B_p(m,M) \\
&\leq p\frac{M^{p-1}-m^{p-1}}{M-m}\left(M-\overline{|h|}_{\Omega,w}\right)\left(\overline{|h|}_{\Omega,w}-m\right) \\
&\leq \frac{1}{4}p(M-m)\left(M^{p-1}-m^{p-1}\right),
\end{aligned}
\tag{3.1}
$$

where $\overline{|h|}_{\Omega,w} := \frac{\int_\Omega |h| w d\mu}{\int_\Omega w d\mu} \in [m,M]$ and $\Psi_p(\cdot;m,M):(m,M)\to\mathbb{R}$ is defined by

$$
\Psi_p(t;m,M) = \frac{M^p - t^p}{M-t} - \frac{t^p - m^p}{t-m}
$$

while

$$
B_p(m,M) := \sup_{t\in(m,M)} \Psi_p(t;m,M). \tag{3.2}
$$

From (2.2) we also have the inequality

$$
\begin{aligned}
0 &\leq \frac{\int_\Omega |h|^p w d\mu}{\int_\Omega w d\mu} - \left(\frac{\int_\Omega |h| w d\mu}{\int_\Omega w d\mu}\right)^p \leq \frac{1}{4}(M-m)\Psi_p\left(\overline{|h|}_{\Omega,w};m,M\right) \\
&\leq \frac{1}{4}p(M-m)\left(M^{p-1}-m^{p-1}\right).
\end{aligned}
\tag{3.3}
$$

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

$$
\gamma \leq \frac{|f|}{|g|^{q-1}} \leq \Gamma \ \mu\text{-a.e on } \Omega,
$$

then we have

$$
\begin{aligned}
0 &\leq \frac{\int_\Omega |f|^p d\mu}{\int_\Omega |g|^q d\mu} - \left(\frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu}\right)^p \\
&\leq \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu}\right)\left(\frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu} - \gamma\right) \\
&\leq p\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\left(\Gamma - \frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu}\right)\left(\frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu} - \gamma\right) \\
&\leq \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right),
\end{aligned}
\tag{3.4}
$$

and

$$
\begin{aligned}
0 &\leq \frac{\int_\Omega |f|^p d\mu}{\int_\Omega |g|^q d\mu} - \left(\frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu}\right)^p \\
&\leq \frac{1}{4}(\Gamma-\gamma)\Psi_p\left(\frac{\int_\Omega |fg| d\mu}{\int_\Omega |g|^q d\mu};\gamma,\Gamma\right) \leq \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right),
\end{aligned}
\tag{3.5}
$$

where $B_p(\cdot,\cdot)$ and $\Psi_p(\cdot;\cdot,\cdot)$ are defined above.

Proof. The inequalities (3.4) and (3.5) follow from (3.1) and (3.3) by choosing

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

The details are omitted. □

Remark 6. *We observe that for $p=q=2$ we have $\Psi_2(t;\gamma,\Gamma)=\Gamma-\gamma=B_2(\gamma,\Gamma)$ and then from the first inequality in (3.4) we get the following reverse of the Cauchy-Bunyakovsky-Schwarz inequality:*

$$\begin{aligned}(3.6)\qquad &\int_\Omega |g|^2 d\mu \int_\Omega |f|^2 d\mu-\left(\int_\Omega |fg|\,d\mu\right)^2\\ &\le \left(\Gamma-\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2 d\mu}\right)\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^2 d\mu}-\gamma\right)\left(\int_\Omega |g|^2 d\mu\right)^2\\ &\le \frac{1}{4}(\Gamma-\gamma)^2\left(\int_\Omega |g|^2 d\mu\right)^2,\end{aligned}$$

provided that $f,g\in L_2(\Omega,\mu)$, and there exists the constants $\gamma,\Gamma>0$ such that

$$\gamma\le\frac{|f|}{|g|}\le\Gamma \ \mu\text{-a.e on } \Omega.$$

Corollary 4. *With the assumptions of Proposition 1 we have the following additive reverses of the Hölder inequality*

$$\begin{aligned}(3.7)\qquad 0&\le\left(\int_\Omega |f|^p d\mu\right)^{1/p}\left(\int_\Omega |g|^q d\mu\right)^{1/q}-\int_\Omega |fg|\,d\mu\\ &\le\left[\frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\right]^{1/p}\left(\Gamma-\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q d\mu}\right)^{1/p}\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q d\mu}-\gamma\right)^{1/p}\\ &\times\int_\Omega |g|^q d\mu\\ &\le p^{1/p}\left(\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\right)^{1/p}\left(\Gamma-\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q d\mu}\right)^{1/p}\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q d\mu}-\gamma\right)^{1/p}\\ &\times\int_\Omega |g|^q d\mu\\ &\le\frac{1}{4^{1/p}}p^{1/p}(\Gamma-\gamma)^{1/p}\left(\Gamma^{p-1}-\gamma^{p-1}\right)^{1/p}\int_\Omega |g|^q d\mu\end{aligned}$$

and

$$\begin{aligned}(3.8)\qquad 0&\le\left(\int_\Omega |f|^p d\mu\right)^{1/p}\left(\int_\Omega |g|^q d\mu\right)^{1/q}-\int_\Omega |fg|\,d\mu\\ &\le\frac{1}{4^{1/p}}(\Gamma-\gamma)^{1/p}\Psi_p^{1/p}\left(\frac{\int_\Omega |fg|\,d\mu}{\int_\Omega |g|^q d\mu};m,M\right)\int_\Omega |g|^q d\mu\\ &\le\frac{1}{4^{1/p}}p^{1/p}(\Gamma-\gamma)^{1/p}\left(\Gamma^{p-1}-\gamma^{p-1}\right)^{1/p}\int_\Omega |g|^q d\mu\end{aligned}$$

where $p>1$ and $\frac{1}{p}+\frac{1}{q}=1$.

Proof. By multiplying in (3.4) with $\left(\int_\Omega |g|^q \, d\mu\right)^p$ we have

$$\begin{aligned}
&\int_\Omega |f|^p \, d\mu \left(\int_\Omega |g|^q \, d\mu\right)^{p-1} - \left(\int_\Omega |fg| \, d\mu\right)^p \\
&\leq \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu}\right)\left(\frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu} - \gamma\right)\left(\int_\Omega |g|^q \, d\mu\right)^p \\
&\leq p\frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma}\left(\Gamma - \frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu}\right)\left(\frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu} - \gamma\right)\left(\int_\Omega |g|^q \, d\mu\right)^p \\
&\leq \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right)\left(\int_\Omega |g|^q \, d\mu\right)^p,
\end{aligned}$$

which is equivalent with

$$\begin{aligned}
&\int_\Omega |f|^p \, d\mu \left(\int_\Omega |g|^q \, d\mu\right)^{p-1} \\
&\leq \left(\int_\Omega |fg| \, d\mu\right)^p + \frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\left(\Gamma - \frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu}\right)\left(\frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu} - \gamma\right) \\
&\times \left(\int_\Omega |g|^q \, d\mu\right)^p \\
&\leq \left(\int_\Omega |fg| \, d\mu\right)^p + p\left(\Gamma - \frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu}\right)\left(\frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu} - \gamma\right) \\
&\times \left(\int_\Omega |g|^q \, d\mu\right)^p \frac{\Gamma^{p-1}-\gamma^{p-1}}{\Gamma-\gamma} \\
&\leq \left(\int_\Omega |fg| \, d\mu\right)^p + \frac{1}{4}p(\Gamma-\gamma)\left(\Gamma^{p-1}-\gamma^{p-1}\right)\left(\int_\Omega |g|^q \, d\mu\right)^p.
\end{aligned} \tag{3.9}$$

Taking the power $1/p$ with $p > 1$ and employing the following elementary inequality that state that for $p > 1$ and $\alpha, \beta > 0$,

$$(\alpha+\beta)^{1/p} \leq \alpha^{1/p} + \beta^{1/p}$$

we have from the first part of (3.9) that

$$\begin{aligned}
&\int_\Omega |f|^p \, d\mu \left(\int_\Omega |g|^q \, d\mu\right)^{1-\frac{1}{p}} \\
&\leq \int_\Omega |fg| \, d\mu + \left[\frac{B_p(\gamma,\Gamma)}{\Gamma-\gamma}\right]^{1/p}\left(\Gamma - \frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu}\right)^{1/p}\left(\frac{\int_\Omega |fg| \, d\mu}{\int_\Omega |g|^q \, d\mu} - \gamma\right)^{1/p} \\
&\times \int_\Omega |g|^q \, d\mu
\end{aligned} \tag{3.10}$$

and since $1 - \frac{1}{p} = \frac{1}{q}$ we get from (3.10) the first inequality in (3.7). The rest is obvious.

The inequality (3.8) can be proved in a similar manner, however the details are omitted. □

If $h : \Omega \to \mathbb{R}$ is μ-measurable, satisfies the bounds

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

and is such that $h,|h|^p\in L_w(\Omega,\mu)$, for a $\mu-$measurable function $w:\Omega\to\mathbb{R}$, with $w(x)\geq 0$ for $\mu-$ a.e. $x\in\Omega$ and $\int_\Omega w d\mu>0$, then from (2.10) we also have the inequality

$$0\leq\frac{\int_\Omega|h|^p wd\mu}{\int_\Omega wd\mu}-\left(\frac{\int_\Omega|h|\,wd\mu}{\int_\Omega wd\mu}\right)^p \tag{3.11}$$
$$\leq 2\left[\frac{m^p+M^p}{2}-\left(\frac{m+M}{2}\right)^p\right]\max\left\{\frac{M-\overline{|h|}_{\Omega,w}}{M-m},\frac{\overline{|h|}_{\Omega,w}-m}{M-m}\right\}$$
$$\leq\frac{1}{2}p\left(M^{p-1}-m^{p-1}\right)\max\left\{M-\overline{|h|}_{\Omega,w},\overline{|h|}_{\Omega,w}-m\right\}.$$

where, as above, $\overline{|h|}_{\Omega,w}:=\frac{\int_\Omega|h|wd\mu}{\int_\Omega wd\mu}\in[m,M]$.

From the inequality (3.11) we can state:

Proposition 2. *With the assumptions of Proposition 1 we have*

$$0\leq\frac{\int_\Omega|f|^p d\mu}{\int_\Omega|g|^q d\mu}-\left(\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}\right)^p \tag{3.12}$$
$$\leq 2\cdot\frac{\frac{\gamma^p+\Gamma^p}{2}-\left(\frac{\gamma+\Gamma}{2}\right)^p}{\Gamma-\gamma}\max\left\{\Gamma-\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu},\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}-\gamma\right\}$$
$$\leq\frac{1}{2}p\left(\Gamma^{p-1}-\gamma^{p-1}\right)\max\left\{\Gamma-\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu},\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}-\gamma\right\}.$$

Finally, the following additive reverse of the Hölder inequality can be stated as well:

Corollary 5. *With the assumptions of Proposition 1 we have*

$$0\leq\left(\int_\Omega|f|^p d\mu\right)^{1/p}\left(\int_\Omega|g|^q d\mu\right)^{1/q}-\int_\Omega|fg|\,d\mu \tag{3.13}$$
$$\leq 2^{1/p}\cdot\left(\frac{\frac{\gamma^p+\Gamma^p}{2}-\left(\frac{\gamma+\Gamma}{2}\right)^p}{\Gamma-\gamma}\right)^{1/p}$$
$$\times\max\left\{\left(\Gamma-\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}\right)^{1/p},\left(\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}-\gamma\right)^{1/p}\right\}\int_\Omega|g|^q d\mu$$
$$\leq\frac{1}{2^{1/p}}p^{1/p}\max\left\{\left(\Gamma-\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}\right)^{1/p},\left(\frac{\int_\Omega|fg|\,d\mu}{\int_\Omega|g|^q d\mu}-\gamma\right)^{1/p}\right\}$$
$$\times\left(\Gamma^{p-1}-\gamma^{p-1}\right)^{1/p}\int_\Omega|g|^q d\mu.$$

Remark 7. *As a simpler, however coarser inequality we have the following result:*

$$0\leq\left(\int_\Omega|f|^p d\mu\right)^{1/p}\left(\int_\Omega|g|^q d\mu\right)^{1/q}-\int_\Omega|fg|\,d\mu$$
$$\leq 2^{1/p}\cdot\left[\frac{\gamma^p+\Gamma^p}{2}-\left(\frac{\gamma+\Gamma}{2}\right)^p\right]^{1/p}\int_\Omega|g|^q d\mu,$$

where f and g are as above.

4. Applications for f-Divergence

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 [19], Kullback and Leibler [24], Rényi [30], Havrda and Charvat [17], Kapur [22], Sharma and Mittal [32], Burbea and Rao [4], Rao [29], Lin [25], Csiszár [7], Ali and Silvey [1], Vajda [39], Shioya and Da-te [34] and others (see for example [26] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [29], genetics [26], finance, economics, and political science [31], [37], [38], biology [28], the analysis of contingency tables [16], approximation of probability distributions [6], [23], signal processing [20], [21] 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 $\sigma-$finite measure μ are given. Consider the set of all probability densities on μ to be $\mathcal{P}:=\left\{p|p:\Omega\to\mathbb{R},\ p(x)\geq 0,\ \int_\Omega p(x)\,d\mu(x)=1\right\}$. The Kullback-Leibler divergence [24] is well known among the information divergences. It is defined as:

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

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 [18], χ^2-*divergence* D_{χ^2}, $\alpha-$*divergence* D_α, *Bhattacharyya distance* D_B [2], *Harmonic distance* D_{Ha}, *Jeffrey's distance* D_J [19], *triangular discrimination* D_Δ [36], etc... They are defined as follows:

$$D_v(p,q):=\int_\Omega |p(x)-q(x)|\,d\mu(x),\quad p,q\in\mathcal{P}; \tag{4.2}$$

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

$$D_{\chi^2}(p,q):=\int_\Omega p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2-1\right]d\mu(x),\quad p,q\in\mathcal{P}; \tag{4.4}$$

$$D_\alpha(p,q):=\frac{4}{1-\alpha^2}\left[1-\int_\Omega [p(x)]^{\frac{1-\alpha}{2}}[q(x)]^{\frac{1+\alpha}{2}}d\mu(x)\right],\quad p,q\in\mathcal{P}; \tag{4.5}$$

$$D_B(p,q):=\int_\Omega \sqrt{p(x)q(x)}d\mu(x),\quad p,q\in\mathcal{P}; \tag{4.6}$$

$$D_{Ha}(p,q):=\int_\Omega \frac{2p(x)q(x)}{p(x)+q(x)}d\mu(x),\quad p,q\in\mathcal{P}; \tag{4.7}$$

$$D_J(p,q):=\int_\Omega [p(x)-q(x)]\ln\left[\frac{p(x)}{q(x)}\right]d\mu(x),\quad p,q\in\mathcal{P}; \tag{4.8}$$

$$D_{\Delta}(p,q) := \int_{\Omega} \frac{[p(x) - q(x)]^2}{p(x) + q(x)} d\mu(x), \quad p, q \in \mathcal{P}. \tag{4.9}$$

For other divergence measures, see the paper [22] by Kapur or the book on line [35] by Taneja.

Csiszár $f-$divergence is defined as follows [8]

$$I_f(p,q) := \int_{\Omega} p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x), \quad p, q \in \mathcal{P}, \tag{4.10}$$

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. Most of the above distances (4.1) – (4.9), are particular instances of Csiszár $f-$divergence. There are also many others which are not in this class (see for example [35]). For the basic properties of Csiszár $f-$divergence see [8], [9] and [39].

The following result holds:

Proposition 3. *Let $f : (0,\infty) \to \mathbb{R}$ be a convex function with the property that $f(1) = 0$. Assume that $p, q \in \mathcal{P}$ and there exists the constants $0 < r < 1 < R < \infty$ such that*

$$r \leq \frac{q(x)}{p(x)} \leq R \text{ for } \mu\text{-a.e. } x \in \Omega. \tag{4.11}$$

Then we have the inequalities

$$\begin{aligned} I_f(p,q) &\leq \frac{(R-1)(1-r)}{R-r} \sup_{t\in(r,R)} \Psi_f(t;r,R) \\ &\leq (R-1)(1-r) \frac{f'_-(R) - f'_+(r)}{R-r} \\ &\leq \frac{1}{4}(R-r)\left[f'_-(R) - f'_+(r)\right], \end{aligned} \tag{4.12}$$

and $\Psi_f(\cdot;r,R) : (r,R) \to \mathbb{R}$ is defined by

$$\Psi_f(t;r,R) = \frac{f(R) - f(t)}{R-t} - \frac{f(t) - f(r)}{t-r}.$$

We also have the inequality

$$\begin{aligned} I_f(p,q) &\leq \frac{1}{4}(R-r) \frac{f(R)(1-r) + f(r)(R-1)}{(R-1)(1-r)} \\ &\leq \frac{1}{4}(R-r)\left[f'_-(R) - f'_+(r)\right]. \end{aligned} \tag{4.13}$$

The proof follows by Theorem 2 by choosing $w(x) = p(x), f(x) = \frac{q(x)}{p(x)}, m = r$ and $M = R$ and performing the required calculations. The details are omitted.

Utilising the same approach and Theorem 3 we can also state that:

Proposition 4. *With the assumptions of Proposition 3 we have*

$$\begin{aligned} I_f(p,q) &\leq 2\max\left\{\frac{R-1}{R-r}, \frac{1-r}{R-r}\right\}\left[\frac{f(r)+f(R)}{2} - f\left(\frac{r+R}{2}\right)\right] \\ &\leq \frac{1}{2}\max\{R-1, 1-r\}\left[f'_-(R) - f'_+(r)\right]. \end{aligned} \tag{4.14}$$

The above results can be utilized to obtain various inequalities for the divergence measures in Information Theory that are particular instances of f-divergence.

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 $f : (0,\infty) \to \mathbb{R}$, $f(t) = -\ln t$.

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

$$r \leq \frac{q(x)}{p(x)} \leq R \text{ for } \mu\text{-a.e. } x \in \Omega. \tag{4.15}$$

then we get from (4.12) that

$$D_{KL}(p,q) \leq \frac{(R-1)(1-r)}{rR}, \tag{4.16}$$

from (4.13) that

$$D_{KL}(p,q) \leq \frac{1}{4}(R-r) \ln \left[R^{-\frac{1}{R-1}} r^{-\frac{1}{1-r}} \right]$$

and from (4.14) that

$$\begin{aligned} D_{KL}(p,q) &\leq 2 \max \left\{ \frac{R-1}{R-r}, \frac{1-r}{R-r} \right\} \ln \left(\frac{A(r,R)}{G(r,R)} \right) \\ &\leq \frac{1}{2} \max \{R-1, 1-r\} \left(\frac{R-r}{rR} \right), \end{aligned} \tag{4.17}$$

where $A(r,R)$ is the arithmetic mean and $G(r,R)$ is the geometric mean of the positive numbers r and R.

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), Suplement, Art. 12. `[Online 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 and N.M. IONESCU, Some converse of Jensen's inequality and applications. *Rev. Anal. Numér. Théor. Approx.* **23** (1994), no. 1, 71–78.

[16] D.V. GOKHALE and S. KULLBACK, *Information in Contingency Tables*, New York, Marcel Decker, 1978.

[17] J.H. HAVRDA and F. CHARVAT, Quantification method classification process: concept of structural α-entropy, *Kybernetika*, **3** (1967), 30-35.

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

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

[20] 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.

[21] T. KAILATH, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Technology.,* Vol COM-15 (1967), 52-60.

[22] J.N. KAPUR, A comparative assessment of various measures of directed divergence, *Advances in Management Studies*, **3** (1984), 1-16.

[23] 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.

[24] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Annals Math. Statist.,* **22** (1951), 79-86.

[25] J. LIN, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.,* **37** (1) (1991), 145-151.

[26] M. MEI, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.

[27] C.P. NICULESCU, An extension of Chebyshev's inequality and its connection with Jensen's inequality. *J. Inequal. Appl.* **6** (2001), no. 4, 451–462.

[28] E.C. PIELOU, *Ecological Diversity,* Wiley, New York, 1975.

[29] C.R. RAO, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.

[30] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.

[31] A. SEN, *On Economic Inequality,* Oxford University Press, London 1973.

[32] B.D. SHARMA and D.P. MITTAL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.,* **2** (4)(1977), 122-132.

[33] S. SIMIC, On a global upper bound for Jensen's inequality, J. Math. Anal. Appl. 343(2008), 414-419.

[34] 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.

[35] I. J. TANEJA, *Generalised Information Measures and their Applications* (`http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html`).

[36] F. TOPSOE, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll., RGMIA,* **2** (1) (1999), 85-98.

[37] H. THEIL, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

[38] H. THEIL, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.

[39] I. VAJDA, *Theory of Statistical Inference and Information,* Dordrecht-Boston, Kluwer Academic Publishers, 1989.

[1]Mathematics, 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`

[2]School of Computational & Applied Mathematics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa