A GENERAL DIVERGENCE MEASURE FOR MONOTONIC
FUNCTIONS AND APPLICATIONS IN INFORMATION THEORY

S.S. DRAGOMIR

Abstract. A general divergence measure for monotonic functions is introduced. Its connections with the f-divergence for convex functions are explored. The main properties are pointed out.

1. Introduction

Let (X, A) be a measurable space satisfying |A| > 2 and μ be a σ-finite measure on (X, A). Let P be the set of all probability measures on (X, A) which are absolutely continuous with respect to μ. For P, Q ∈ P, let p = dp/dμ and q = dQ/dμ denote the Radon-Nikodym derivatives of P and Q with respect to μ.

Two probability measures P, Q ∈ P are said to be orthogonal and we denote this by Q ⊥ P if

\[ P(\{q = 0\}) = Q(\{p = 0\}) = 1. \]

Let f : [0, ∞) → (-∞, ∞] be a convex function that is continuous at 0, i.e.,

\[ f(0) = \lim_{u \downarrow 0} f(u). \]

In 1963, I. Csiszár [2] introduced the concept of f-divergence as follows.

Definition 1. Let P, Q ∈ P. Then

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

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

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

1.1. The Class of \( \chi^\alpha \)-Divergences. The f-divergences of this class, which is generated by the function \( \chi^\alpha \), \( \alpha \in [1, \infty) \), defined by

\[ \chi^\alpha(u) = |u - 1|^\alpha, \quad u \in [0, \infty) \]

have the form

\[ I_f(Q, P) = \int_X \left|\frac{q}{p} - 1\right|^\alpha d\mu = \int_X p^{1-\alpha} |q - p|^\alpha d\mu. \]

From this class only the parameter \( \alpha = 1 \) provides a distance in the topological sense, namely the total variation distance \( V(Q, P) = \int_X |q - p| d\mu \). The most prominent special case of this class is, however, Karl Pearson’s \( \chi^2 \)-divergence.

Date: 13 January, 2004.
2000 Mathematics Subject Classification. 94Axx, 26D15, 26D10.
Key words and phrases. f-divergence, Convexity, Divergence measures, Monotonic functions.
1.2. **Dichotomy Class.** From this class, generated by the function \( f_\alpha : [0, \infty) \to \mathbb{R} \)

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

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

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

Another important divergence is the *Kullback-Leibler divergence* obtained for \( \alpha = 1 \),

\[
   KL(Q, P) = \int_X q \ln \left( \frac{q}{p} \right) d\mu.
\]

1.3. **Matsushita’s Divergences.** The elements of this class, which is generated by the function \( \phi_\alpha \), \( \alpha \in (0, 1] \) given by

\[
   \phi_\alpha (u) := |1 - u^\alpha|^{\frac{1}{\alpha}}, \quad u \in [0, \infty),
\]

are prototypes of metric divergences, providing the distances \( [I_{\phi_\alpha}(Q, P)]^\alpha \).

1.4. **Puri-Vineze Divergences.** This class is generated by the functions \( \Phi_\alpha \), \( \alpha \in [1, \infty) \) given by

\[
   \Phi_\alpha (u) := \frac{|1 - u|^{\alpha}}{(u + 1)^{\alpha-1}}, \quad u \in [0, \infty).
\]

It has been shown in [3] that, this class provides the distances \( [I_{\Phi_\alpha}(Q, P)]^\frac{1}{\alpha} \).

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

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

It has been shown in [5] that, this class provides the distances \( [I_{\Psi_\alpha}(Q, P)]^{\min(\alpha, \frac{1}{\alpha})} \) for \( \alpha \in (0, \infty) \) and \( \frac{1}{2} V(Q, P) \) for \( \alpha = \infty \).

2. **Some Classes of Normalised Functions**

We denote by \( M^\#([0, \infty)) \) the class of *monotonic nondecreasing functions* defined on \([0, \infty)\) and by \( M^1([0, \infty)) \) the class of *measurable functions* on \([0, \infty)\). We also consider \( L_1([0, \infty)) \) the class of measurable functions \( g : [0, \infty) \to \mathbb{R} \) with the property that

\[
   g(t) \leq g(1) \leq g(s) \quad \text{for } 0 \leq t \leq 1 \leq s < \infty.
\]

It is obvious that

\[
   M^\#([0, \infty)) \subseteq L_1([0, \infty)),
\]
and the inclusion (2.2) is strict.

We say that a function \( f : [0, \infty) \to \mathbb{R} \) is normalised if \( f(1) = 0 \). We denote by \( \mathcal{M}_s([0, \infty)) \) the class of all normalised measurable functions defined on \([0, \infty)\).

We also need the following classes of functions

\[
\mathcal{C}_o([0, \infty)) := \{ f \in \mathcal{M}_s([0, \infty)) \mid f \text{ is continuous convex on } [0, \infty) \} ;
\]

\[
\mathcal{D}_0([0, \infty)) := \{ f \in \mathcal{M}_s([0, \infty)) \mid f(t) = (t - 1)g(t), \forall t \in [0, \infty), \ g \in \mathcal{M}_s^{\uparrow}([0, \infty)) \};
\]

\[
\mathcal{O}_0([0, \infty)) := \{ f \in \mathcal{M}_s([0, \infty)) \mid f(t) = (t - 1)g(t), \forall t \in [0, \infty), \ g \in L_1([0, \infty)) \} .
\]

From the definition of \( \mathcal{D}_0([0, \infty)) \) and \( \mathcal{O}_0([0, \infty)) \) and taking into account that the strict inclusion (2.2) holds, we deduce that

\[
\mathcal{D}_0([0, \infty)) \subsetneq \mathcal{O}_0([0, \infty)) ,
\]

and the inclusion is strict.

For the other two classes, we may state the following result.

**Lemma 1.** We have the strict inclusion

\[
\mathcal{C}_o([0, \infty)) \subsetneq \mathcal{D}_0([0, \infty)) .
\]

**Proof.** We will show that any continuous convex function \( f : [0, \infty) \to \mathbb{R} \) that is normalised may be represented as:

\[
f(t) = (t - 1)g(t) \text{ for any } t \in [0, \infty),
\]

where \( g \in \mathcal{M}_s^{\uparrow}([0, \infty)) \).

Now, let \( f \in \mathcal{C}_o([0, \infty)) \). For \( \lambda \in [D_- f(1), D_+ f(1)] \), define

\[
g_\lambda(t) := \begin{cases} 
\frac{f(t)}{t-1} & \text{if } t \in [0, 1) \cup (1, \infty) , \\
\lambda & \text{if } t = 1.
\end{cases}
\]

We use the following well known result [2, p. 111]:

If \( \Psi \) is convex on \((a, b)\) and \( a < s < t < u < b \), then

\[
\Psi(s, t) \leq \Psi(s, u) \leq \Psi(t, u) ,
\]

where

\[
\Psi(s, t) = \frac{\Psi(t) - \Psi(s)}{t - s} .
\]

If \( \Psi \) is strictly convex on \((a, b)\), equality will not occur in (2.6).

If we apply the above result for \( 0 < s < t < 1 \), then we can state

\[
\frac{f(s)}{s-1} \leq \frac{f(t)}{t-1} .
\]

Taking the limit over \( t \to 1, t < 1 \), we deduce

\[
\frac{f(s)}{s-1} \leq D_- f(1)
\]

showing that for \( 0 < t < 1 \), we have \( g_\lambda(t) \leq \lambda \).

Similarly, we may prove that for \( 1 < t < \infty, g_\lambda(t) \geq \lambda \). If we use the same result for \( 0 < t_1 < t_2 < 1 \), then we may write

\[
\frac{f(t_1)}{t_1 - 1} \leq \frac{f(t_2)}{t_2 - 1} ,
\]
which gives \( g_\Lambda (t_1) \leq g_\Lambda (t_2) \) for \( 0 < t_1 < t_2 < 1 \).

In a similar fashion we can prove that for \( 1 < t_1 < t_2 < \infty \), \( g_\Lambda (t_1) \leq g_\Lambda (t_2) \), and thus we may conclude that the function \( g_\Lambda \) is monotonic non-decreasing on the whole interval \([0, \infty)\).

If we consider now the function \( f(t) = (t - 1) e^{vt}, t \in [0, \infty) \), we observe that
\[
 f'(t) = (qt - 3) e^{vt}, \quad f''(t) = 8e^{vt} (2t - 1)
\]
which gives \( g_\Lambda \) for each \( t \in [0, \infty) \).

We have
\[
 I_f (Q, P) = I_f (Q, P),
\]
for any \( P, Q \in \mathcal{P} \) if and only if there exists a constant \( c \in \mathbb{R} \) such that
\[
 f_1 (u) = f (u) + c (u - 1),
\]
for any \( u \in [0, \infty) \);

(ii) We have
\[
 I_{f^*} (Q, P) = I_f (Q, P),
\]
for any \( P, Q \in \mathcal{P} \) if and only if there exists a constant \( d \in \mathbb{R} \) such that
\[
 f^* (u) = f (u) + d (c - 1),
\]
for any \( u \in [0, \infty) \).

Theorem 2 (Range of Values Theorem). Let \( f : [0, \infty) \to \mathbb{R} \) be a continuous convex function on \([0, \infty)\).

For any \( P, Q \in \mathcal{P} \), we have the double inequality
\[
 f (1) \leq I_f (Q, P) \leq f (0) + f^* (0). \tag{3.1}
\]

(i) If \( P = Q \), then the equality holds in the first part of (3.1).

If \( f \) is strictly convex at 1, then the equality holds in the first part of (3.1) if and only if \( P = Q \);
If \( Q \perp P \), then the equality holds in the second part of (3.1).
If \( f(0) + f^*(0) < \infty \), then equality holds in the second part of (3.1) if and only if \( Q \perp P \).

Define the function \( \tilde{f} : (0, \infty) \to \mathbb{R} \), \( \tilde{f}(u) = \frac{1}{2} (f(u) + f^*(u)) \). The following result is a refinement of the second inequality in Theorem 2 (see [3, Theorem 3]).

**Theorem 3.** Let \( f \in C(\mathbb{R}) \) with \( f(0) + f^*(0) < \infty \). Then

\[
(3.2) \quad 0 \leq I_f(Q, P) \leq \tilde{f}(0) V(Q, P)
\]

for any \( Q, P \in \mathcal{P} \).

4. A General Divergence Measure

If \( f : [0, \infty) \to \mathbb{R} \) is a general measurable function, then we may define the \( f \)-divergence in the same way, i.e., if \( P, Q \in \mathcal{P} \), then

\[
I_f(Q, P) = \int_{\mathbb{R}} p(x) f \left( \frac{q(x)}{p(x)} \right) d\mu(x).
\]

For a measurable function \( g : [0, \infty) \to \mathbb{R} \), we may also define the \( \delta \)-divergence by the formula

\[
\delta_g(Q, P) = \int_{\mathbb{R}} [q(x) - p(x)] g \left( \frac{q(x)}{p(x)} \right) d\mu(x).
\]

It is obvious that the \( \delta \)-divergence of a function \( g \) may be seen as the \( f \)-divergence of the function \( f \), where \( f(t) = (t - 1) g(t) \) for \( t \in [0, \infty) \).

If \( f \in C([0, \infty)) \) and since \( f(t) = (t - 1) g_\lambda(t) \), \( t \in [0, \infty) \), we have

\[
g_\lambda(t) := \begin{cases} 
\frac{f(t)}{t-1} & \text{if } t \in [0, 1) \cup (1, \infty), \\
\lambda & \text{if } t = 1;
\end{cases}
\]

and \( \lambda \in [D_- f(1), D_+ f(1)] \), shows that for any \( f \in C([0, \infty)) \) we have

\[
I_f(Q, P) = \delta g_\lambda(Q, P) \quad \text{for any } P, Q \in \mathcal{P},
\]

i.e., the \( f \)-divergence for any normalised continuous convex function \( f : [0, \infty) \to \mathbb{R} \) may be seen as the \( \delta \)-divergence of the function \( g_\lambda \) defined by (4.1).

In what follows, we point out some fundamental properties of the \( \delta \)-divergence.

**Theorem 4.** Let \( g : [0, \infty) \to \mathbb{R} \) be a measurable function on \([0, \infty)\) and \( P, Q \in \mathcal{P} \).
If there exists the constants \( m, M \) with

\[
-\infty < m \leq g \left( \frac{q(x)}{p(x)} \right) \leq M < \infty
\]

for \( \mu \)-a.e. \( x \in X \), then we have the inequality

\[
|\delta_g(Q, P)| \leq \frac{1}{2} (M - m) V(Q, P).
\]

**Proof.** We observe that the following identity holds true

\[
\delta_g(Q, P) = \int_{\mathbb{R}} [q(x) - p(x)] \left[ g \left( \frac{q(x)}{p(x)} \right) - \frac{m + M}{2} \right] d\mu(x)
\]
By (4.3), we deduce that
\[ \left| g \left[ \frac{q(x)}{p(x)} \right] - \frac{m + M}{2} \right| \leq \frac{1}{2} (M - m) \]
for \( \mu \)-a.e. \( x \in X \).

Taking the modulus in (4.5) we deduce
\[ |\delta g(Q, P)| \leq \int_X \left| q(x) - p(x) \right| g \left[ \frac{q(x)}{p(x)} \right] \left| \frac{m + M}{2} \right| d\mu(x) \]
\[ \leq \frac{1}{2} (M - m) \int_X |q(x) - p(x)| d\mu(x) \]
\[ = \frac{1}{2} (M - m) V(Q, P) \]
and the inequality (4.4) is proved.

The following corollary is a natural consequence of the above theorem.

**Corollary 1.** Let \( g : [0, \infty) \to \mathbb{R} \) be a measurable function on \( [0, \infty) \). If
\[ m := \text{ess inf}_{t \in [0, \infty)} g(t) > -\infty, \quad M := \text{ess sup}_{t \in [0, \infty)} g(t) < \infty, \]
then for any \( P, Q \in \mathcal{P} \), we have the inequality
\[ |\delta g(Q, P)| \leq \frac{1}{2} (M - m) V(Q, P). \]

**Remark 2.** We know that, if \( f : [0, \infty) \to \mathbb{R} \) is a normalised continuous convex function and if
\[ \lim_{t \to 0^+} f^*(t) = \lim_{u \to 0^+} \left[ uf \left( \frac{1}{u} \right) \right] =: f^*(0), \]
then we have the inequality [Theorem 2.3]
\[ I_f(Q, P) \leq \frac{f(0) + f^*(0)}{2} V(Q, P), \]
for any \( P, Q \in \mathcal{P} \). We can prove this inequality by the use of Corollary 1 as follows. We have
\[ I_f(Q, P) = \delta g_{\lambda}(Q, P), \]
where
\[ g_{\lambda}(t) := \begin{cases} \frac{f(t)}{t - 1} & \text{if } t \in [0, 1) \cup (1, \infty), \\ \lambda & \text{if } t = 1, \end{cases} \]
where \( \lambda \in [D_- f(1), D_+ f(1)] \) and \( g_{\lambda} \in \mathcal{M}^b([0, \infty)) \). We observe that for any \( t \in [0, \infty) \), we have
\[ g_{\lambda}(t) = \lim_{t \to 0^+} g_{\lambda}(t) = -f(0) = m > -\infty \]
and
\[ g_{\lambda}(t) \leq \lim_{t \to +\infty} g_{\lambda}(t) = \lim_{t \to +\infty} \frac{f(t)}{t - 1} = \lim_{u \to 0^+} \left[ \frac{f \left( \frac{1}{u} \right)}{1 - u} \right] \]
\[ = \lim_{u \to 0^+} \left[ \frac{uf \left( \frac{1}{u} \right)}{1 - u} \right] = f^*(0) = M < \infty. \]

Applying Corollary 1 for \( m = -f(0) \) and \( M = f^*(0) \), we deduce the desired inequality (4.7).
The following result also holds.

**Theorem 5.** Let \( g : [0, \infty) \to \mathbb{R} \) be a measurable function on \([0, \infty)\) and \( P, Q \in \mathcal{P} \). If there exists a constant \( K \) with \( K > 0 \) such that

\[
|g(t) - g(1)| \leq K|t - 1|^\alpha,
\]

for \( \mu \)-a.e. \( x \in X \), where \( \alpha \in (0, \infty) \) is a given number, then we have the inequality\(^{(4.9)}\)

\[
|\delta_g(Q,P)| \leq K I_{\chi^{\alpha+1}}(Q,P).
\]

**Proof.** We observe that the following identity holds true

\[
\delta_g(Q,P) = \int_X |q(x) - p(x)| \left( g \left( \frac{q(x)}{p(x)} \right) - g(1) \right) d\mu(x).
\]

Taking the modulus in \((4.10)\) and using the condition \((4.8)\), we have successively

\[
|\delta_g(Q,P)| \leq \int_X |q(x) - p(x)| \left| g \left( \frac{q(x)}{p(x)} \right) - g(1) \right| d\mu(x) \\
\leq K \int_X |p(x)|^{-\alpha} |q(x) - p(x)|^{\alpha+1} d\mu(x) \\
\leq K I_{\chi^{\alpha+1}}(Q,P)
\]

and the inequality \((4.9)\) is obtained. \( \blacksquare \)

The following corollary holds.

**Corollary 2.** Let \( g : [0, \infty) \to \mathbb{R} \) be a measurable function on \([0, \infty)\) with the property that there exists a constant \( K \) with the property that

\[
|g(t) - g(1)| \leq K|t - 1|^\alpha,
\]

for a.e. \( t \in [0, \infty) \), where \( \alpha > 0 \) is a given number. Then for any \( P, Q \in \mathcal{P} \), we have the inequality

\[
|\delta_g(Q,P)| \leq K I_{\chi^{\alpha+1}}(Q,P).
\]

**Remark 3.** If the function \( g : [0, \infty) \to \mathbb{R} \) is Hölder continuous with a constant \( H > 0 \) and \( \beta \in (0,1) \), i.e.,

\[
|g(t) - g(s)| \leq H|t - s|^\beta,
\]

for any \( t, s \in [0, \infty) \), then obviously \((4.7)\) holds with \( K = H \) and \( \alpha = \beta \).

If \( g : [0, \infty) \to \mathbb{R} \) is Lipschitzian with the constant \( L > 0 \), i.e.,

\[
|g(t) - g(s)| \leq L|t - s|,
\]

for any \( t, s \in [0, \infty) \), then

\[
|\delta_g(Q,P)| \leq K I_{\chi^2}(Q,P),
\]

for any \( P, Q \in \mathcal{P} \).

Finally, if \( g \) is locally absolutely continuous and the derivative \( g' : [0, \infty) \to \mathbb{R} \) is essentially bounded, i.e., \( \|g'\|_{0,\infty,\infty} := \text{ess} \sup_{t \in [0, \infty)} |g'(t)| < \infty \), then we have the inequality

\[
|\delta_g(Q,P)| \leq \|g'\|_{0,\infty,\infty} I_{\chi^2}(Q,P),
\]

for any \( P, Q \in \mathcal{P} \).

The following result concerning \( f \)-divergences for \( f \) convex functions holds.
Theorem 6. Let $f : [0, \infty] \to \mathbb{R}$ be a continuous convex function on $[0, \infty)$. If $\lambda \in [D_{-} f (1), D_{+} f (1)]$ ($\lambda = f' (1)$ if $f$ is differentiable at $t = 1$), and there exists a constant $K > 0$ and $\alpha > 0$ such that

\begin{equation}
| f (t) - \lambda (t - 1) | \leq K | t - 1 |^{\alpha + 1},
\end{equation}

for any $t \in [0, \infty)$, then we have the inequality

\begin{equation}
0 \leq I_{f} (Q, P) \leq K I_{\kappa}^{\alpha + 1} (Q, P),
\end{equation}

for any $P, Q \in P$.

Proof. We have

\begin{equation}
I_{f} (Q, P) = \int_{X} [q (x) - p (x)] g_{\lambda} \left[ \frac{p (x)}{q (x)} \right] d\mu (x) = \delta_{g_{\lambda}} (Q, P),
\end{equation}

where

\begin{equation}
g_{\lambda} (t) := \begin{cases}
\frac{f (t)}{t - 1} & \text{if } t \in [0, 1) \cup (1, \infty), \\
\lambda & \text{if } t = 1,
\end{cases}
\end{equation}

and $\lambda \in [D_{-} f (1), D_{+} f (1)]$.

Applying Corollary 2 for $g_{\lambda}$, we deduce the desired result. 

5. The Positivity of $\delta$–Divergence for $g \in \mathcal{M}^{\#} ([0, \infty))$

The following result holds.

Theorem 7. If $g \in \mathcal{M}^{\#} ([0, \infty))$, then $\delta_{g} (Q, P) \geq 0$ for any $P, Q \in P$.

Proof. We use the identity

\begin{equation}
\delta_{g} (Q, P) = \int_{X} [q (x) - p (x)] g \left[ \frac{q (x)}{p (x)} \right] d\mu (x)
\end{equation}

= \int_{X} p (x) \left[ \frac{q (x)}{p (x)} - 1 \right] g \left[ \frac{q (x)}{p (x)} \right] d\mu (x)
\end{equation}

= \frac{1}{2} \int_{X} \int_{X} p (x) p (y) \left[ \frac{q (x)}{p (x)} - \frac{q (y)}{p (y)} \right] g \left[ \frac{q (x)}{p (x)} \right] - g \left[ \frac{q (y)}{p (y)} \right] d\mu (x) d\mu (y).
\end{equation}

Since $g \in \mathcal{M}^{\#} ([0, \infty))$, then for any $t, s \in [0, \infty)$, we have

\begin{equation}
(t - s) (g (t) - g (s)) \geq 0
\end{equation}

giving that

\begin{equation}
\left[ \frac{q (x)}{p (x)} - \frac{q (y)}{p (y)} \right] \left[ g \left[ \frac{q (x)}{p (x)} \right] - g \left[ \frac{q (y)}{p (y)} \right] \right] \geq 0
\end{equation}

for any $x, y \in X$.

Using the representation (5.1), we deduce the desired result. 

The following corollary is a natural consequence of the above result.

Corollary 3. If $f \in \mathcal{D}_{0} ([0, \infty))$, then $I_{f} (Q, P) \geq 0$ for any $P, Q \in P$. 

Proof: If \( f \in D_0 ([0, \infty)) \), then there exists a \( g \in \mathcal{M}^+ ([0, \infty)) \) such that \( f(t) = (t-1) g(t) \) for any \( t \in [0, \infty) \). Then

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

\[
= \int_X p(x) \left[ \frac{q(x)}{p(x)} - 1 \right] g \left[ \frac{q(x)}{p(x)} \right] d\mu(x)
\]

\[
= \delta_g (Q, P) \geq 0,
\]

and the proof is completed. \( \blacksquare \)

In fact, the following improvement of Theorem 7 holds.

**Theorem 8.** If \( g \in \mathcal{M}^+ ([0, \infty)) \), then

\[
\delta_g (Q, P) \geq |\delta_{|g|} (Q, P)| \geq 0,
\]

for any \( P, Q \in \mathcal{P} \).

**Proof.** Since \( g \) is monotonic nondecreasing, we have

\[
\left[ \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right] \left[ g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right]
\]

\[
\geq \left( \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right) \left( g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right) d\mu(x) d\mu(y)
\]

for any \( x, y \in X \).

Multiplying (5.3) by \( p(x) p(y) \geq 0 \) and integrating on \( X^2 \), we deduce

\[
\int_X \left| \left| \int_X p(x) p(y) \left( \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right) \left( g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right) d\mu(x) d\mu(y) \right| \right|^2.
\]

Using the representation (5.4) and the same identity for \( |g| \), we deduce the desired inequality (5.2). \( \blacksquare \)

Before we point out other possible refinements for the positivity inequality \( \delta_g (Q, P) \geq 0 \), where \( g \in \mathcal{M}^+ ([0, \infty)) \), we need the following divergence measure as well:

\[
\tilde{\delta}_h (Q, P) := \int_X |g(x) - p(x)| h \left[ \frac{q(x)}{p(x)} \right] d\mu(x)
\]

which will be called the absolute \( \delta - divergence \) generated by the function \( h : [0, \infty) \to \mathbb{R} \) that is assumed to be measurable on \([0, \infty)\).

The following result holds.

**Theorem 9.** If \( g \in \mathcal{M}^+ ([0, \infty)) \), then

\[
\delta_g (Q, P) \geq \max \left\{ \left| \delta_g (Q, P) - V (Q, P) I_g (Q, P) \right|, \left| \tilde{\delta}_{|g|} (Q, P) - V (Q, P) I_{|g|} (Q, P) \right| \right\} \geq 0,
\]

for any \( P, Q \in \mathcal{P} \).
Proof. Since $g$ is monotonic, we have

$$\int X p(x) p(y) \left( \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right) \left( g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right) d\mu(x) d\mu(y)$$

for any $x, y \in X$.

If we multiply (5.5) by $p(x) p(y) \geq 0$ and integrate, we deduce

$$\int X p(x) p(y) \left( \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right) \left( g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right) d\mu(x) d\mu(y)$$

for any $x, y \in X$.

Now, observe that

$$\int X p(x) p(y) \left( \frac{q(x)}{p(x)} - \frac{q(y)}{p(y)} \right) \left( g \left[ \frac{q(x)}{p(x)} \right] - g \left[ \frac{q(y)}{p(y)} \right] \right) d\mu(x) d\mu(y)$$

and a similar identity holds for the quantity in the second branch of (5.6).

Finally, using the representation (5.1), we deduce the desired inequality (5.4).

6. The Positivity of $\delta$-Divergence for $g \in \mathcal{L}c_1 ([0, \infty))$

The following result extending the positivity of $\delta$-divergence for monotonic functions, holds.

**Theorem 10.** If $g \in \mathcal{L}c_1 ([0, \infty))$, then $\delta_g (Q, P) \geq 0$ for any $P, Q \in \mathcal{P}$. 

Proof. We use the identity

\begin{align}
\delta_g(Q, P) &= \int_X [q(x) - p(x)] g \left( \frac{q(x)}{p(x)} \right) d\mu(x) \\
&= \int_X p(x) \left( \frac{q(x)}{p(x)} - 1 \right) g \left( \frac{q(x)}{p(x)} \right) d\mu(x) \\
&= \int_X p(x) \left( \frac{q(x)}{p(x)} - 1 \right) \left[ g \left( \frac{q(x)}{p(x)} \right) - g(1) \right] d\mu(x).
\end{align}

Since \( g \in \mathcal{L}^1_0([0, \infty)) \), then for any \( t \in [0, \infty) \) we have

\[(t - 1) [g(t) - g(1)] \geq 0\]

giving that

\[\left( \frac{q(x)}{p(x)} - 1 \right) \left[ g \left( \frac{q(x)}{p(x)} \right) - g(1) \right] \geq 0\]

for any \( x \in X \).

Using the representation (6.1), we deduce the desired result.

\[\square\]

Corollary 4. If \( f \in \mathcal{O}_0([0, \infty)) \), then \( I_f(Q, P) \geq 0 \) for any \( P, Q \in \mathcal{P} \).

Proof. If \( f \in \mathcal{O}_0([0, \infty)) \), then there exists a \( g \in \mathcal{L}^1_0([0, \infty)) \) such that \( f(t) = (t - 1) g(t) \) for any \( t \in [0, \infty) \). Then

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

\[= \int_X p(x) \left( \frac{q(x)}{p(x)} - 1 \right) g \left( \frac{q(x)}{p(x)} \right) d\mu(x)\]

\[= \delta_g(Q, P) \geq 0,\]

and the proof is completed.

The following improvement of Theorem 10 holds.

Theorem 11. If \( g \in \mathcal{L}^1_0([0, \infty)) \), then

\[\delta_g(Q, P) \geq |\delta_{|g|}(Q, P)| \geq 0\]

for any \( P, Q \in \mathcal{P} \).

Proof. Since \( g \in \mathcal{L}^1_0([0, \infty)) \), we obviously have

\begin{align}
\left( \frac{q(x)}{p(x)} - 1 \right) \left[ g \left( \frac{q(x)}{p(x)} \right) - g(1) \right] \\
= \left| \left( \frac{q(x)}{p(x)} - 1 \right) \left( g \left( \frac{q(x)}{p(x)} \right) - g(1) \right) \right| \\
\geq \left| \left( \frac{q(x)}{p(x)} - 1 \right) \left( g \left( \frac{q(x)}{p(x)} \right) \right) - |g(1)| \right|.
\end{align}
Multiplying (6.3) by \(p(x) \geq 0\) and integrating on \(X\), we have
\[
\int_X p(x) \left[ \frac{q(x)}{p(x)} - 1 \right] \left[ g \left( \frac{q(x)}{p(x)} \right) - g(1) \right] d\mu(x) = \int_X p(x) \left[ \frac{q(x)}{p(x)} - 1 \right] \left( \left| g \left( \frac{q(x)}{p(x)} \right) \right| - \left| g(1) \right| \right) d\mu(x) \geq \left| \int_X p(x) \left( \frac{q(x)}{p(x)} - 1 \right) \left| g \left( \frac{q(x)}{p(x)} \right) \right| d\mu(x) \right| = |\delta|_{\|\|}(Q, P)|,
\]
and the inequality (6.2) is proved. \(\blacksquare\)

7. Bounds in Terms of the \(\chi^2\)-Divergence

The following result may be stated.

**Theorem 12.** Let \(g : [0, \infty] \to \mathbb{R}\) be a differentiable function such that there exists the constants \(\gamma, \Gamma \in \mathbb{R}\) with
\[
\gamma \leq g'(t) \leq \Gamma \quad \text{for any } t \in (0, \infty).
\]
Then we have the inequality
\[
\gamma D_{\chi^2}(Q, P) \leq \delta_g(Q, P) \leq \Gamma D_{\chi^2}(Q, P),
\]
for any \(P, Q \in \mathcal{P}\).

**Proof.** Consider the auxiliary function \(h_\gamma : [0, \infty] \to \mathbb{R}, h_\gamma(t) := g(t) - \gamma (t - 1)\). Obviously, \(h_\gamma\) is differentiable on \((0, \infty)\) and since, by (7.1),
\[
h'_\gamma(t) = g'(t) - \gamma \geq 0
\]
it follows that \(h_\gamma\) is monotonic nondecreasing on \([0, \infty)\).

Applying Theorem 7, we deduce
\[
\delta_{h_\gamma}(Q, P) \geq 0 \quad \text{for any } P, Q \in \mathcal{P}
\]
and since
\[
\delta_{h_\gamma}(Q, P) = \delta_{g-\gamma(\cdot-1)}(Q, P)
\]
\[
= \int_X \left[ q(x) - p(x) \right] \left[ g \left( \frac{q(x)}{p(x)} \right) - \gamma \left( \frac{q(x)}{p(x)} - 1 \right) \right] d\mu(x)
\]
\[
= \delta_g(Q, P) - \gamma D_{\chi^2}(Q, P),
\]
then the first inequality in (7.2) is proved.

The second inequality may be proven in a similar manner by using the auxiliary function \(h_\Gamma : [0, \infty] \to \mathbb{R}, h_\Gamma(t) := \Gamma(t - 1) - g(t)\). \(\blacksquare\)

The following corollary is a natural application of the above theorem.

**Corollary 5.** Let \(f : [0, \infty] \to \mathbb{R}\) be a differentiable convex function on \((0, \infty)\) with \(f(1) = 0\). If there exist the constants \(\gamma, \Gamma \in \mathbb{R}\) with the property that:
\[
\gamma (t - 1)^2 + f(t) \leq f'(t)(t - 1) \leq f(t) + \Gamma (t - 1)^2
\]
for any \(t \in (0, \infty)\), then we have the inequality:
\[
\gamma D_{\chi^2}(Q, P) \leq I_f(Q, P) \leq \Gamma D_{\chi^2}(Q, P)
\]
for any \(P, Q \in \mathcal{P}\).
Proof. We know that for any \( P, Q \in \mathcal{P} \), we have (see for example (4.2)):

\[ I_f (Q, P) = \delta_{g_f'(1)} (Q, P), \]

where

\[ g_f'(1) = \begin{cases} 
\frac{f(t)}{t} & \text{if } t \in [0,1) \cup (1,\infty), \\
 f'(1) & \text{if } t = 1.
\end{cases} \]

We observe that, by the hypothesis of the corollary, \( g_f'(1) \) is differentiable on \((0,\infty)\) and

\[ g_f'(1)'(t) = \frac{f'(t)(t-1) - f(t)}{(t-1)^2} \]

for any \( t \in (0,1) \cup (1,\infty) \).

Using (7.3), we deduce that

\[ \gamma \leq g_f'(1)'(t) \leq \Gamma \]

for \( t \in (0,\infty) \), and applying Theorem 12 above, for \( g = g_f'(1) \), we deduce the desired inequality (7.4).

8. Bounds in Terms of the J–Divergence

We recall that the Jeffreys divergence (or J–divergence for short) is defined as

\[ J(Q, P) := \int_X [q(x) - p(x)] \ln \left[ \frac{q(x)}{p(x)} \right] d\mu(x), \]

where \( P, Q \in \mathcal{P} \).

The following result holds.

**Theorem 13.** Let \( g : [0,\infty] \to \mathbb{R} \) be a differentiable function such that there exists the constants \( \phi, \Phi \in \mathbb{R} \) with

\[ \phi \leq tg'(t) \leq \Phi \quad \text{for any } t \in (0,\infty). \]

Then we have the inequality

\[ \phi J(Q, P) \leq \delta_g(Q, P) \leq \Phi J(Q, P), \]

for any \( P, Q \in \mathcal{P} \).

**Proof.** Consider the auxiliary function \( h_\phi : [0,\infty] \to \mathbb{R}, h_\phi(t) := g(t) - \phi \ln t. \) Obviously, \( h_\phi \) is differentiable on \((0,\infty)\) and, by (8.2),

\[ h_\phi'(t) = g'(t) - \frac{\phi}{t} = \frac{1}{t} [tg'(t) - \phi] \geq 0, \]

for any \( t \in (0,\infty) \), showing that the function is monotonic nondecreasing on \((0,\infty)\).

Applying Theorem 7, we deduce

\[ \delta_{h_\phi}(Q, P) \geq 0 \quad \text{for any } P, Q \in \mathcal{P} \]

and since

\[ \delta_{h_\phi}(Q, P) = \delta_{g - \phi \ln(\cdot)}(Q, P) \]

\[ = \int_X \left[ q(x) - p(x) \right] \left[ g \left[ \frac{q(x)}{p(x)} \right] - \phi \ln \left[ \frac{q(x)}{p(x)} \right] \right] d\mu(x) \]

\[ = \delta_g(Q, P) - \phi J(Q, P), \]
then the first inequality in (8.3) is proved.

The second inequality may be proven in a similar manner by using the auxiliary function \( h_\Phi : [0, \infty) \to \mathbb{R} \), \( h_\Phi (t) := \Phi \ln t - g(t) \).

The following corollary is a natural application of the above theorem.

**Corollary 6.** Let \( f : [0, \infty] \to \mathbb{R} \) be a differentiable convex function on \((0, \infty)\) with \( f(1) = 0 \). If there exist the constants \( \phi, \Phi \in \mathbb{R} \) with the property that:

\[
\phi (t - 1)^2 + tf(t) \leq t(t - 1)f'(t) \leq tf(t) + \Phi(t - 1)^2
\]

for any \( t \in (0, \infty) \), then we have the inequality:

\[
\phi J(Q, P) \leq I_f(Q, P) \leq \Phi J(Q, P)
\]

for any \( P, Q \in \mathcal{P} \).

The proof is similar to the one in Corollary 5 and we omit the details.

**References**


