



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Inequalities for synchronous functions and applications

This is the Published version of the following publication

Dragomir, Sever S (2019) Inequalities for synchronous functions and applications. *Constructive Mathematical Analysis*, 2 (3). pp. 109-123. ISSN 2651-2939

The publisher's official version can be found at
<https://dergipark.org.tr/en/pub/cma/issue/45605/562166>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/43548/>

Inequalities for Synchronous Functions and Applications

SILVESTRU SEVER DRAGOMIR

ABSTRACT. Some inequalities for synchronous functions that are a mixture between Čebyšev's and Jensen's inequality are provided. Applications for f -divergence measure and some particular instances including Kullback-Leibler divergence, Jeffreys divergence and χ^2 -divergence are also given.

Keywords: Synchronous Functions, Lipschitzian functions, Čebyšev inequality, Jensen's inequality, f -divergence measure, Kullback-Leibler divergence, Jeffreys divergence measure, χ^2 -divergence.

2010 Mathematics Subject Classification: 26D15, 26D10, 94A17.

1. INTRODUCTION

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

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

For simplicity of notation we write everywhere in the sequel $\int_{\Omega} w d\nu$ instead of $\int_{\Omega} w(x) d\nu(x)$. Assume also that $\int_{\Omega} w d\nu = 1$. We have *Jensen's inequality*

$$(1.1) \quad \int_{\Omega} w(\Phi \circ f) d\nu \geq \Phi \left(\int_{\Omega} w f d\nu \right),$$

where $\Phi : [m, M] \rightarrow \mathbb{R}$ is a continuous convex function on the closed interval of real numbers $[m, M]$, $f : \Omega \rightarrow [m, M]$ is ν -measurable and such that $f, \Phi \circ f \in L_w(\Omega, \nu)$.

We say that the pair of measurable functions (f, g) are *synchronous* on Ω if

$$(1.2) \quad (f(x) - f(y))(g(x) - g(y)) \geq 0$$

for ν -a.e. $x, y \in \Omega$. If the inequality reverses in (1.2), the functions are called *asynchronous* on Ω . If (f, g) are synchronous on Ω and $f, g, fg \in L_w(\Omega, \nu)$ then the following inequality, that is known in the literature as *Čebyšev's Inequality*, holds

$$(1.3) \quad \int_{\Omega} w f g d\nu \geq \int_{\Omega} w f d\nu \int_{\Omega} w g d\nu,$$

where $w(x) \geq 0$ for ν -a.e. (almost every) $x \in \Omega$ and $\int_{\Omega} w d\nu = 1$.

In this paper we establish some inequalities for synchronous functions that are a mixture between Čebyšev's and Jensen's inequality. Applications for f -divergence measure and some particular instances including Kullback-Leibler divergence, Jeffreys divergence and χ^2 -divergence are also given.

Received: 9 May 2019; Accepted: 2 July 2019; Published Online: 3 July 2019

*Corresponding author: S. S. Dragomir; sever.dragomir@vu.edu.au

DOI: 10.33205/cma.562166

2. INEQUALITIES FOR SYNCHRONOUS FUNCTIONS

We have the following inequality for synchronous functions:

Theorem 2.1. *Let $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ be two synchronous functions on $[m, M]$ and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)(\Psi \circ g) \in L_w(\Omega, \nu)$, then*

$$(2.4) \quad \int_{\Omega} w (\Phi \circ g)(\Psi \circ g) d\nu + \Phi \left(\int_{\Omega} w g d\nu \right) \Psi \left(\int_{\Omega} w g d\nu \right) \\ \geq \Phi \left(\int_{\Omega} w g d\nu \right) \int_{\Omega} w (\Psi \circ g) d\nu + \Psi \left(\int_{\Omega} w g d\nu \right) \int_{\Omega} w (\Phi \circ g) d\nu.$$

If the functions (Φ, Ψ) are asynchronous, then the inequality in (2.4) reverses.

Proof. Since Φ, Ψ are synchronous on $[m, M]$ and $\int_{\Omega} w g d\nu \in [m, M]$, then we have

$$\left[\Phi(g(x)) - \Phi \left(\int_{\Omega} w g d\nu \right) \right] \left[\Psi(g(x)) - \Psi \left(\int_{\Omega} w g d\nu \right) \right] \geq 0$$

for ν -a.e. $x \in \Omega$.

This is equivalent to

$$(2.5) \quad \Phi(g(x))\Psi(g(x)) + \Phi \left(\int_{\Omega} w g d\nu \right) \Psi \left(\int_{\Omega} w g d\nu \right) \\ \geq \Phi \left(\int_{\Omega} w g d\nu \right) \Psi + \Psi \left(\int_{\Omega} w g d\nu \right) \Phi(g(x))$$

for ν -a.e. $x \in \Omega$.

Now, if we multiply (2.5) by $w \geq 0$ a.e. on Ω and integrate, we deduce the desired result (2.4). \square

Remark 2.1. *If the functions $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ have the same monotonicity (opposite monotonicity) on $[m, M]$, then they are synchronous (asynchronous) and the inequality (2.4) holds for any $g \in L_w(\Omega, \nu)$.*

If $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ are two synchronous functions on $[m, M]$, $x_i \in [m, M]$ and $w_i \geq 0$, $i \in \{1, \dots, n\}$ with $\sum_{i=1}^n w_i = 1$, then by applying the inequality (2.4) for the discrete counting measure, we have

$$(2.6) \quad \sum_{i=1}^n w_i \Phi(x_i) \Psi(x_i) + \Phi \left(\sum_{i=1}^n w_i x_i \right) \Psi \left(\sum_{i=1}^n w_i x_i \right) \\ \geq \Phi \left(\sum_{i=1}^n w_i x_i \right) \sum_{i=1}^n w_i \Psi(x_i) + \Psi \left(\sum_{i=1}^n w_i x_i \right) \sum_{i=1}^n w_i \Phi(x_i).$$

Example 2.1. *Let $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$.*

a). *If $p, q > 0$ (< 0) and $g : \Omega \rightarrow [0, \infty)$ is ν -measurable and such that $g, g^p, g^q, g^{p+q} \in L_w(\Omega, \nu)$, then*

$$(2.7) \quad \int_{\Omega} w g^{p+q} d\nu + \left(\int_{\Omega} w g d\nu \right)^p \left(\int_{\Omega} w g d\nu \right)^q \\ \geq \left(\int_{\Omega} w g d\nu \right)^p \int_{\Omega} w g^q d\nu + \left(\int_{\Omega} w g d\nu \right)^q \int_{\Omega} w g^p d\nu.$$

If $p > 0$ (< 0), and $q < 0$ (> 0) then the inequality (2.7) reverses.

b). If $\alpha, \beta > 0 (< 0)$ and $g : \Omega \rightarrow \mathbb{R}$ is ν -measurable and such that $g, \exp(\alpha g), \exp(\beta g), \exp((\alpha + \beta)g) \in L_w(\Omega, \nu)$, then

$$(2.8) \quad \int_{\Omega} w \exp((\alpha + \beta)g) d\nu + \exp\left(\left(\alpha + \beta\right) \int_{\Omega} wgd\nu\right) \geq \exp\left(\alpha \int_{\Omega} wgd\nu\right) \int_{\Omega} w \exp(\beta g) d\nu + \exp\left(\beta \int_{\Omega} wgd\nu\right) \int_{\Omega} w \exp(\alpha g) d\nu.$$

If $\alpha > 0 (< 0)$, and $\beta < 0 (> 0)$ then the inequality (2.8) reverses.

c). If $p > 0$ and $g : \Omega \rightarrow (0, \infty)$ is ν -measurable and such that $g, g^p, \ln g, g^p \ln g \in L_w(\Omega, \nu)$, then

$$(2.9) \quad \int_{\Omega} wg^p \ln g d\nu + \left(\int_{\Omega} wgd\nu\right)^p \ln\left(\int_{\Omega} wgd\nu\right) \geq \left(\int_{\Omega} wgd\nu\right)^p \int_{\Omega} w \ln g d\nu + \ln\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} wg^p d\nu.$$

If $p < 0$, then the inequality (2.9) reverses.

Corollary 2.1. Let $\Phi : [m, M] \rightarrow \mathbb{R}$ be a measurable function on $[m, M]$ and $w \geq 0$ a.e. on Ω and $\int_{\Omega} wd\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$, then

$$(2.10) \quad \frac{1}{2} \left[\int_{\Omega} w(\Phi \circ g)^2 d\nu + \Phi^2\left(\int_{\Omega} wgd\nu\right) \right] \geq \Phi\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} w(\Phi \circ g) d\nu.$$

We observe that the inequality (2.10) is of interest only if $\Phi\left(\int_{\Omega} wgd\nu\right) \neq 0$. In this case, by dividing with $\Phi^2\left(\int_{\Omega} wgd\nu\right) > 0$, we get

$$(2.11) \quad \frac{1}{2} \left[\frac{\int_{\Omega} w(\Phi \circ g)^2 d\nu}{\Phi^2\left(\int_{\Omega} wgd\nu\right)} + 1 \right] \geq \frac{\int_{\Omega} w(\Phi \circ g) d\nu}{\Phi\left(\int_{\Omega} wgd\nu\right)}.$$

Remark 2.2. Let $\Phi : [m, M] \rightarrow \mathbb{R}$ be a convex function on $[m, M]$ and $w \geq 0$ a.e. on Ω with $\int_{\Omega} wd\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$ and $\Phi\left(\int_{\Omega} wgd\nu\right) > 0$, then by (2.11) we have

$$(2.12) \quad \frac{1}{2} \left[\frac{\int_{\Omega} w(\Phi \circ g)^2 d\nu}{\Phi^2\left(\int_{\Omega} wgd\nu\right)} + 1 \right] \geq \frac{\int_{\Omega} w(\Phi \circ g) d\nu}{\Phi\left(\int_{\Omega} wgd\nu\right)} \geq 1.$$

This implies that

$$(2.13) \quad \frac{\int_{\Omega} w(\Phi \circ g)^2 d\nu}{\Phi^2\left(\int_{\Omega} wgd\nu\right)} \geq 1.$$

This inequality obviously holds for functions $\Phi : [m, M] \rightarrow \mathbb{R}$ that are square convex, namely Φ^2 is convex. There are examples of convex functions $\Phi : [m, M] \rightarrow \mathbb{R}$ for which Φ^2 is not convex and $\Phi\left(\int_{\Omega} wgd\nu\right) > 0$ holds. Indeed, if we consider $\Phi : [-k, k] \rightarrow \mathbb{R}, \Phi(t) = t^2 - 1$ for $k > 1$ then $\Phi^2(t) = (t^2 - 1)^2$ is convex on $\left[-k, -\frac{\sqrt{3}}{3}\right] \cup \left[\frac{\sqrt{3}}{3}, k\right]$ and concave on $\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$. Now, observe that for $g(t) = t, \Omega = [0, k], w(t) = \frac{1}{k}$ we have

$$\int_{\Omega} wgd\nu = \frac{1}{k} \int_0^k t dt = \frac{k}{2}$$

and

$$\Phi\left(\int_{\Omega} wgd\nu\right) = \Phi\left(\frac{k}{2}\right) = \frac{k^2}{4} - 1$$

which is positive for $k > 2$.

This shows that the Jensen's type inequality (2.13) holds for larger classes than the square convex functions, namely for convex functions Φ for which we have $\Phi\left(\int_{\Omega} wgd\nu\right) > 0$.

Corollary 2.2. Let $\Phi : [m, M] \rightarrow \mathbb{R}$ be a monotonic nondecreasing function on $[m, M]$ and $w \geq 0$ a.e. on Ω and $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, g(\Phi \circ g) \in L_w(\Omega, \nu)$, then

$$(2.14) \quad \int_{\Omega} wg(\Phi \circ g) d\nu \geq \int_{\Omega} wgd\nu \int_{\Omega} w(\Phi \circ g) d\nu.$$

Remark 2.3. We observe that, under the assumptions of Corollary 2.2 and if $g : \Omega \rightarrow [m, M]$ is convex and $\int_{\Omega} wgd\nu > 0$, then we get from (2.14) that

$$(2.15) \quad \frac{\int_{\Omega} wg(\Phi \circ g) d\nu}{\int_{\Omega} wgd\nu} \geq \int_{\Omega} w(\Phi \circ g) d\nu \geq \Phi\left(\int_{\Omega} wgd\nu\right).$$

Example 2.2. Let $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$.

a). If $p \geq 1$ and $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, g^p, g^{p+1} \in L_w(\Omega, \nu)$, then

$$(2.16) \quad \frac{\int_{\Omega} wg^{p+1} d\nu}{\int_{\Omega} wgd\nu} \geq \int_{\Omega} wg^p d\nu \geq \left(\int_{\Omega} wgd\nu\right)^p.$$

b). If $\alpha > 0$ and $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \exp(\alpha g), g \exp(\alpha g) \in L_w(\Omega, \nu)$, then

$$(2.17) \quad \frac{\int_{\Omega} wg \exp(\alpha g) d\nu}{\int_{\Omega} wgd\nu} \geq \int_{\Omega} w \exp(\alpha g) d\nu \geq \exp\left(\alpha \int_{\Omega} wgd\nu\right).$$

Corollary 2.3. Let $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ be two synchronous functions on $[m, M]$, Ψ also convex on $[m, M]$ and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)(\Psi \circ g) \in L_w(\Omega, \nu)$ and $\Phi\left(\int_{\Omega} wgd\nu\right) > 0$, then

$$(2.18) \quad \int_{\Omega} w(\Phi \circ g)(\Psi \circ g) d\nu \geq \Psi\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} w(\Phi \circ g) d\nu.$$

Proof. From (2.4) and Jensen's inequality for Ψ we have

$$\begin{aligned} & \int_{\Omega} w(\Phi \circ g)(\Psi \circ g) d\nu + \Phi\left(\int_{\Omega} wgd\nu\right) \Psi\left(\int_{\Omega} wgd\nu\right) \\ & \geq \Phi\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} w(\Psi \circ g) d\nu + \Psi\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} w(\Phi \circ g) d\nu \\ & \geq \Phi\left(\int_{\Omega} wgd\nu\right) \Psi\left(\int_{\Omega} wgd\nu\right) + \Psi\left(\int_{\Omega} wgd\nu\right) \int_{\Omega} w(\Phi \circ g) d\nu \end{aligned}$$

and the inequality (2.18) is obtained. \square

Let $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ be two synchronous functions on $[m, M]$, Ψ also convex on $[m, M]$. If $x_i \in [m, M]$ and $w_i \geq 0, i \in \{1, \dots, n\}$ with $\sum_{i=1}^n w_i = 1$, then by applying the inequality (2.18) for the discrete counting measure, we have

$$(2.19) \quad \sum_{i=1}^n w_i \Phi(x_i) \Psi(x_i) \geq \Psi\left(\sum_{i=1}^n w_i x_i\right) \sum_{i=1}^n w_i \Phi(x_i).$$

Example 2.3. Let $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$.

a). If $p > 0, q \geq 1$ and $g : \Omega \rightarrow [0, \infty)$ is ν -measurable and such that $g, g^p, g^q, g^{p+q} \in L_w(\Omega, \nu)$, then by (2.18) we have

$$(2.20) \quad \frac{\int_{\Omega} w g^{p+q} d\nu}{\int_{\Omega} w g^p d\nu} \geq \left(\int_{\Omega} w g d\nu \right)^q.$$

b). If $\alpha, \beta > 0$ and $g : \Omega \rightarrow \mathbb{R}$ is ν -measurable and such that $g, \exp(\beta g), \exp((\alpha + \beta)g) \in L_w(\Omega, \nu)$, then by (2.18) we have

$$(2.21) \quad \frac{\int_{\Omega} w \exp((\alpha + \beta)g) d\nu}{\int_{\Omega} w \exp(\beta g) d\nu} \geq \exp\left(\alpha \int_{\Omega} w g d\nu\right).$$

c). If $p \geq 1$ and $g : \Omega \rightarrow (0, \infty)$ is ν -measurable and such that $g, \ln g, g^p \ln g \in L_w(\Omega, \nu)$, then by (2.18) we have

$$(2.22) \quad \int_{\Omega} w g^p \ln g d\nu \geq \left(\int_{\Omega} w g d\nu \right)^p \int_{\Omega} w \ln g d\nu.$$

3. AN ASSOCIATED FUNCTIONAL

Let $\Phi, \Psi : I \rightarrow \mathbb{R}$ be two measurable functions on the interval I and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)(\Psi \circ g) \in L_w(\Omega, \nu)$, then we can consider the following functional

$$(3.23) \quad \begin{aligned} & \mathcal{F}(\Phi, \Psi; g, w) \\ & := \int_{\Omega} w (\Phi \circ g)(\Psi \circ g) d\nu + \Phi\left(\int_{\Omega} w g d\nu\right) \Psi\left(\int_{\Omega} w g d\nu\right) \\ & \quad - \Phi\left(\int_{\Omega} w g d\nu\right) \int_{\Omega} w (\Psi \circ g) d\nu - \Psi\left(\int_{\Omega} w g d\nu\right) \int_{\Omega} w (\Phi \circ g) d\nu. \end{aligned}$$

In particular, if $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$, we have

$$(3.24) \quad \begin{aligned} & \mathcal{F}(\Phi; g, w) \\ & := \int_{\Omega} w (\Phi \circ g)^2 d\nu + \Phi^2\left(\int_{\Omega} w g d\nu\right) - 2\Phi\left(\int_{\Omega} w g d\nu\right) \int_{\Omega} w (\Phi \circ g) d\nu \\ & \geq 0. \end{aligned}$$

Theorem 3.2. Let $\Phi, \Psi : I \rightarrow \mathbb{R}$ be two measurable functions on I and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2, (\Psi \circ g)^2 \in L_w(\Omega, \nu)$, then

$$(3.25) \quad \mathcal{F}^2(\Phi, \Psi; g, w) \leq \mathcal{F}(\Phi; g, w) \mathcal{F}(\Psi; g, w).$$

Proof. Observe that the following identity holds true

$$(3.26) \quad \begin{aligned} & \mathcal{F}(\Phi, \Psi; g, w) \\ & = \int_{\Omega} w(x) \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} w g d\nu\right) \right] \left[\Psi(g(x)) - \Psi\left(\int_{\Omega} w g d\nu\right) \right] d\nu(x). \end{aligned}$$

Using the Cauchy-Bunyakovsky-Schwarz integral inequality we have

$$\begin{aligned}
 (3.27) \quad & \left| \int_{\Omega} w(x) \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right] \left[\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right] d\nu(x) \right| \\
 & \leq \left(\int_{\Omega} w(x) \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right]^2 d\nu(x) \right)^{1/2} \\
 & \times \left(\int_{\Omega} w(x) \left[\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right]^2 d\nu(x) \right)^{1/2} \\
 & = \mathcal{F}^{1/2}(\Phi; g, w) \mathcal{F}^{1/2}(\Psi; g, w).
 \end{aligned}$$

On utilizing (3.26) and (3.27) we deduce the desired result (3.25). \square

For the functions $\Phi, \Psi : I \rightarrow \mathbb{R}$, the n -tuples of real numbers $x = (x_1, \dots, x_n) \in I^n$ and the probability distribution $w = (w_1, \dots, w_n)$ define the functionals

$$\begin{aligned}
 (3.28) \quad \mathcal{F}(\Phi, \Psi; x, w) & := \sum_{i=1}^n w_i \Phi(x_i) \Psi(x_i) + \Phi\left(\sum_{i=1}^n w_i x_i\right) \Psi\left(\sum_{i=1}^n w_i x_i\right) \\
 & - \Phi\left(\sum_{i=1}^n w_i x_i\right) \sum_{i=1}^n w_i \Psi(x_i) - \Psi\left(\sum_{i=1}^n w_i x_i\right) \sum_{i=1}^n w_i \Phi(x_i)
 \end{aligned}$$

and

$$(3.29) \quad \mathcal{F}(\Phi; x, w) := \sum_{i=1}^n w_i \Phi^2(x_i) + \Phi^2\left(\sum_{i=1}^n w_i x_i\right) - 2\Phi\left(\sum_{i=1}^n w_i x_i\right) \sum_{i=1}^n w_i \Phi(x_i).$$

From the inequality (3.25) we have

$$\mathcal{F}^2(\Phi, \Psi; x, w) \leq \mathcal{F}(\Phi; x, w) \mathcal{F}(\Psi; x, w).$$

Theorem 3.3. Let $\Phi : I \rightarrow \mathbb{R}$ be an L -Lipschitzian function on I , with $L > 0$, namely it satisfies the condition

$$|\Phi(t) - \Phi(s)| \leq L|t - s| \text{ for any } t, s \in I,$$

and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, g^2, \Phi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$, then

$$(3.30) \quad (0 \leq) \mathcal{F}^{1/2}(\Phi; g, w) \leq L\mathcal{D}(g, w),$$

where the dispersion $\mathcal{D}(g, w)$ is defined by

$$(3.31) \quad \mathcal{D}(g, w) := \left(\int_{\Omega} wg^2 d\nu - \left(\int_{\Omega} wgd\nu \right)^2 \right)^{1/2}.$$

Proof. By Lipschitz condition we have

$$\begin{aligned}
 \mathcal{F}(\Phi; g, w) &= \int_{\Omega} w(x) \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right]^2 d\nu(x) \\
 &\leq L^2 \int_{\Omega} w(x) \left(g(x) - \int_{\Omega} wgd\nu \right)^2 d\nu(x) \\
 &= L^2 \int_{\Omega} w(x) \left(g^2(x) - 2\left(\int_{\Omega} wgd\nu\right)g(x) + \left(\int_{\Omega} wgd\nu\right)^2 \right) d\nu(x) \\
 &= L^2 \left[\int_{\Omega} w(x)g^2(x) d\nu(x) - \left(\int_{\Omega} wgd\nu\right)^2 \right] \\
 &= L^2 \mathcal{D}^2(g, w).
 \end{aligned}$$

□

Corollary 3.4. Let $\Phi : [m, M] \rightarrow \mathbb{R}$ be an absolutely continuous function on $[m, M]$ with

$$(3.32) \quad \|\Phi'\|_{[m, M], \infty} := \operatorname{esssup}_{t \in [m, M]} |\Phi'(t)| < \infty$$

and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, g^2, \Phi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$, then

$$(3.33) \quad (0 \leq) \mathcal{F}^{1/2}(\Phi; g, w) \leq \|\Phi'\|_{[m, M], \infty} \mathcal{D}(g, w).$$

The proof follows by Theorem 3.3 on observing that for and $t, s \in [m, M]$ we have

$$|\Phi(t) - \Phi(s)| = \left| \int_s^t \Phi'(u) du \right| \leq |t - s| \|\Phi'\|_{[m, M], \infty}.$$

Corollary 3.5. Let $\Phi : I \rightarrow \mathbb{R}$ be an L -Lipschitzian function on I , with $L > 0$, and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow I$ is ν -measurable and there exists the constant $m, M \in I$ such that

$$(3.34) \quad m \leq g(x) \leq M \text{ for } \nu\text{-a.e. } x \in \Omega,$$

then $g, g^2, \Phi \circ g, (\Phi \circ g)^2 \in L_w(\Omega, \nu)$ and

$$(3.35) \quad (0 \leq) \mathcal{F}^{1/2}(\Phi; g, w) \leq \frac{1}{2} (M - m) L.$$

The proof follows by (3.30) and the Grüss inequality that states that

$$(3.36) \quad \mathcal{D}(g, w) \leq \frac{1}{2} (M - m)$$

provided that g satisfies the condition (3.34).

Corollary 3.6. Let $\Phi : I \rightarrow \mathbb{R}$ be Lipschitzian with constant $L > 0$, $\Psi : I \rightarrow \mathbb{R}$ be Lipschitzian with constant $K > 0$ and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2, (\Psi \circ g)^2 \in L_w(\Omega, \nu)$, then

$$(3.37) \quad |\mathcal{F}(\Phi, \Psi; g, w)| \leq LK \mathcal{D}^2(g, w).$$

Moreover, if $g : \Omega \rightarrow I$ is ν -measurable and there exists the constant $m, M \in I$ such that the condition (3.34) is satisfied, then

$$(3.38) \quad |\mathcal{F}(\Phi, \Psi; g, w)| \leq \frac{1}{4} (M - m)^2 LK.$$

The proof follows by (3.25), (3.30) and (3.35).

If $\Phi : I \rightarrow \mathbb{R}$ is Lipschitzian with constant $L > 0$, $\Psi : I \rightarrow \mathbb{R}$ is Lipschitzian with constant $K > 0$, the n -tuples of real numbers $x = (x_1, \dots, x_n) \in I^n$ then for any probability distribution $w = (w_1, \dots, w_n)$ we have by (3.37) that

$$(3.39) \quad |\mathcal{F}(\Phi, \Psi; x, w)| \leq LK \left(\sum_{i=1}^n w_i x_i^2 - \left(\sum_{i=1}^n w_i x_i \right)^2 \right).$$

If the interval I is closed, namely $I = [m, M]$ and $x = (x_1, \dots, x_n) \in [m, M]^n$ then by (3.38) we get the simpler upper bound:

$$(3.40) \quad |\mathcal{F}(\Phi, \Psi; x, w)| \leq \frac{1}{4} (M - m)^2 LK.$$

Consider the functional

$$(3.41) \quad \mathcal{F}_{p,q}(g, w) := \int_{\Omega} w g^{p+q} d\nu + \left(\int_{\Omega} w g d\nu \right)^p \left(\int_{\Omega} w g d\nu \right)^q \\ - \left(\int_{\Omega} w g d\nu \right)^p \int_{\Omega} w g^q d\nu - \left(\int_{\Omega} w g d\nu \right)^q \int_{\Omega} w g^p d\nu,$$

provided that $g > 0$, $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$, $g, g^p, g^q, g^{p+q} \in L_w(\Omega, \nu)$ and $p, q \in \mathbb{R} \setminus \{0\}$.

Assume that $g : \Omega \rightarrow [m, M] \subset (0, \infty)$ and for $p \neq 0$ define the constants

$$(3.42) \quad \Delta_p(m, M) := |p| \times \begin{cases} M^{p-1} & \text{if } p \geq 1, \\ m^{p-1} & \text{if } p < 1. \end{cases}$$

If we consider the function $\Phi : [m, M] \subset (0, \infty) \rightarrow (0, \infty)$, $\Phi(t) = t^p$ then $\Phi'(t) = p t^{p-1}$ and

$$\sup_{t \in [m, M]} |\Phi'(t)| = \Delta_p(m, M)$$

as defined by (3.42).

Proposition 3.1. *Let $g : \Omega \rightarrow [m, M] \subset (0, \infty)$ be ν -measurable and $p, q \in \mathbb{R} \setminus \{0\}$. Then we have the inequality*

$$(3.43) \quad |\mathcal{F}_{p,q}(g, w)| \leq \frac{1}{4} (M - m)^2 \Delta_p(m, M) \Delta_q(m, M).$$

The proof follows by Corollary 3.6 for the functions $\Phi(t) = t^p$ and $\Psi(t) = t^q$ for $p, q \in \mathbb{R} \setminus \{0\}$.

Consider now the functional

$$(3.44) \quad \mathcal{F}_{p,\ln}(g, w) := \int_{\Omega} w g^p \ln g d\nu + \left(\int_{\Omega} w g d\nu \right)^p \ln \left(\int_{\Omega} w g d\nu \right) \\ - \left(\int_{\Omega} w g d\nu \right)^p \int_{\Omega} w \ln g d\nu - \ln \left(\int_{\Omega} w g d\nu \right) \int_{\Omega} w g^p d\nu,$$

provided that $p > 0$ and $g : \Omega \rightarrow (0, \infty)$ is ν -measurable and such that $g, g^p, \ln g, g^p \ln g \in L_w(\Omega, \nu)$.

If we take the function $\Psi(t) = \ln t$, $t \in [m, M] \subset (0, \infty)$, then $\sup_{t \in [m, M]} |\Psi'(t)| = \frac{1}{m}$.

Using Corollary 3.6 for the functions $\Phi(t) = t^p$ and $\Psi(t) = \ln t$ for $p \in \mathbb{R} \setminus \{0\}$ we can state the following result as well:

Proposition 3.2. Let $g : \Omega \rightarrow [m, M] \subset (0, \infty)$ be ν -measurable and $p \in \mathbb{R} \setminus \{0\}$. Then we have the inequality

$$(3.45) \quad |\mathcal{F}_{p,\ln}(g, w)| \leq \frac{1}{4m} (M - m)^2 \Delta_p(m, M).$$

We have the following result:

Theorem 3.4. Let $\Phi, \Psi : I \rightarrow \mathbb{R}$ be two measurable functions such that there exists the real constants γ, Γ with

$$(3.46) \quad \gamma \leq \frac{\Phi(t) - \Phi(s)}{\Psi(t) - \Psi(s)} \leq \Lambda$$

for a.e. $t, s \in I$ with $t \neq s$. If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2, (\Psi \circ g)^2 \in L_w(\Omega, \nu)$, then we have the inequalities

$$(3.47) \quad \gamma \mathcal{F}(\Psi; g, w) \leq \mathcal{F}(\Phi, \Psi; g, w) \leq \Lambda \mathcal{F}(\Psi; g, w).$$

Proof. My multiplying (3.46) with $(\Psi(t) - \Psi(s))^2 \geq 0$ we get

$$\gamma (\Psi(t) - \Psi(s))^2 \leq [\Phi(t) - \Phi(s)] [\Psi(t) - \Psi(s)] \leq \Lambda (\Psi(t) - \Psi(s))^2$$

for a.e. $t, s \in I$.

This implies

$$(3.48) \quad \begin{aligned} & \gamma w(x) \left(\Psi(g(x)) - \Psi \left(\int_{\Omega} w g d\nu \right) \right)^2 \\ & \leq w(x) \left[\Phi(g(x)) - \Phi \left(\int_{\Omega} w g d\nu \right) \right] \left[\Psi(g(x)) - \Psi \left(\int_{\Omega} w g d\nu \right) \right] \\ & \leq \Lambda w(x) \left(\Psi(g(x)) - \Psi \left(\int_{\Omega} w g d\nu \right) \right)^2 \end{aligned}$$

for ν -a.e. $x \in \Omega$.

Integrating the inequality (3.48) on Ω and making use of the equality (3.26) we deduce the desired result (3.47). \square

Corollary 3.7. Let $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ be continuous on $[m, M]$ and differentiable on (m, M) . Assume that $\Psi'(t) \neq 0$ for any $t \in (m, M)$ and

$$\inf_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right) > -\infty, \quad \sup_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right) < \infty.$$

If $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2, (\Psi \circ g)^2 \in L_w(\Omega, \nu)$, then we have the inequalities

$$(3.49) \quad \begin{aligned} & \inf_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right) \mathcal{F}(\Psi; g, w) \leq \mathcal{F}(\Phi, \Psi; g, w) \\ & \leq \sup_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right) \mathcal{F}(\Psi; g, w). \end{aligned}$$

Proof. By Cauchy's mean value theorem, for any $t, s \in [m, M]$ with $t \neq s$ there exists a c between t and s such that

$$\frac{\Phi(t) - \Phi(s)}{\Psi(t) - \Psi(s)} = \frac{\Phi'(c)}{\Psi'(c)}.$$

Therefore, for any $t, s \in [m, M]$ with $t \neq s$ we have

$$\inf_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right) \leq \frac{\Phi(t) - \Phi(s)}{\Psi(t) - \Psi(s)} \leq \sup_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right).$$

By applying Theorem 3.4 for $\gamma = \inf_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right)$ and $\Gamma = \sup_{t \in (m, M)} \left(\frac{\Phi'(t)}{\Psi'(t)} \right)$ we get the desired result (3.49). \square

Remark 3.4. We observe that if $\Phi, \Psi : I \rightarrow \mathbb{R}$ are two measurable functions such that there exists the positive constant Θ with

$$(3.50) \quad \left| \frac{\Phi(t) - \Phi(s)}{\Psi(t) - \Psi(s)} \right| \leq \Theta$$

for a.e. $t, s \in I$ with $t \neq s$ and $g : \Omega \rightarrow I$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)^2, (\Psi \circ g)^2 \in L_w(\Omega, \nu)$, then we have the inequalities

$$(3.51) \quad |\mathcal{F}(\Phi, \Psi; g, w)| \leq \Theta \mathcal{F}(\Psi; g, w).$$

Moreover, if Φ, Ψ are as in Corollary 3.7, then we have

$$|\mathcal{F}(\Phi, \Psi; g, w)| \leq \sup_{t \in (m, M)} \left| \frac{\Phi'(t)}{\Psi'(t)} \right| \mathcal{F}(\Psi; g, w).$$

In the case of synchronous functions we can prove the following result as well:

Theorem 3.5. Let $\Phi, \Psi : [m, M] \rightarrow \mathbb{R}$ be two synchronous functions on $[m, M]$ and $w \geq 0$ a.e. on Ω with $\int_{\Omega} w d\nu = 1$. If $g : \Omega \rightarrow [m, M]$ is ν -measurable and such that $g, \Phi \circ g, \Psi \circ g, (\Phi \circ g)(\Psi \circ g), |\Phi \circ g|, |\Psi \circ g|, (|\Phi \circ g|)(|\Psi \circ g|) \in L_w(\Omega, \nu)$, then

$$(3.52) \quad \begin{aligned} \mathcal{F}(\Phi, \Psi; g, w) \\ \geq \max \{ |\mathcal{F}(|\Phi|, \Psi; g, w)|, |\mathcal{F}(\Phi, |\Psi|; g, w)|, |\mathcal{F}(|\Phi|, |\Psi|; g, w)| \} \geq 0. \end{aligned}$$

Proof. We use the continuity property of the modulus, namely

$$|a - b| \geq ||a| - |b||, \quad a, b \in \mathbb{R}.$$

Since Φ, Ψ are synchronous, then

$$(3.53) \quad \begin{aligned} & \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right] \left[\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right] \\ &= \left| \Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right| \left| \Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right| \\ &\geq \begin{cases} \left| |\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)| \right| \left| |\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)| \right| \\ \left| |\Phi(g(x)) - \Phi(\int_{\Omega} wgd\nu)| \right| \left| |\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)| \right| \\ \left| |\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)| \right| \left| |\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)| \right| \\ \left| (|\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \\ \left| (|\Phi(g(x)) - \Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \\ \left| (|\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \end{cases} \\ &= \begin{cases} \left| (|\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \\ \left| (|\Phi(g(x)) - \Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \\ \left| (|\Phi(g(x))| - |\Phi(\int_{\Omega} wgd\nu)|) (|\Psi(g(x))| - |\Psi(\int_{\Omega} wgd\nu)|) \right| \end{cases} \end{aligned}$$

for any $x \in \Omega$.

By using the identity (3.26) and the first branch in (3.53) we have

$$\begin{aligned}
 & \mathcal{F}(\Phi, \Psi; g, w) \\
 &= \int_{\Omega} w(x) \left[\Phi(g(x)) - \Phi\left(\int_{\Omega} wgd\nu\right) \right] \left[\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right] d\nu(x) \\
 &\geq \int_{\Omega} w(x) \left| \left(|\Phi(g(x))| - \left| \Phi\left(\int_{\Omega} wgd\nu\right) \right| \right) \left(\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right) \right| d\nu(x) \\
 &\geq \left| \int_{\Omega} w(x) \left(|\Phi(g(x))| - \left| \Phi\left(\int_{\Omega} wgd\nu\right) \right| \right) \left(\Psi(g(x)) - \Psi\left(\int_{\Omega} wgd\nu\right) \right) d\nu(x) \right| \\
 &= |\mathcal{F}(|\Phi|, \Psi; g, w)|,
 \end{aligned}$$

which proves the first part of (3.52).

The second and third part of (3.52) can be proved in a similar way and the details are omitted. \square

For the natural numbers $n, m \geq 1$ we consider the functions $\Phi(t) = t^{2n+1}$ and $\Psi(t) = t^{2m+1}$ for real numbers $t \in \mathbb{R}$. These functions are monotonic increasing on \mathbb{R} . If $g : \Omega \rightarrow \mathbb{R}$ is ν -measurable and such that $g, g^{2n+1}, g^{2m+1}, g^{2m+2n+2} \in L_w(\Omega, \nu)$, then by (3.52) we have the inequality

$$\begin{aligned}
 (3.54) \quad & \mathcal{F}\left((\cdot)^{2n+1}, (\cdot)^{2m+1}; g, w\right) \\
 & \geq \max \left\{ \left| \mathcal{F}\left(|\cdot|^{2n+1}, (\cdot)^{2m+1}; g, w\right) \right|, \right. \\
 & \left. \left| \mathcal{F}\left((\cdot)^{2n+1}, |\cdot|^{2m+1}; g, w\right) \right|, \left| \mathcal{F}\left(|\cdot|^{2n+1}, |\cdot|^{2m+1}; g, w\right) \right| \right\} (\geq 0).
 \end{aligned}$$

4. APPLICATIONS FOR f -DIVERGENCES

Let (X, \mathcal{A}) be a measurable space satisfying $|\mathcal{A}| > 2$ and μ be a σ -finite measure on (X, \mathcal{A}) . Let \mathcal{P} be the set of all probability measures on (X, \mathcal{A}) which are absolutely continuous with respect to μ . For $P, Q \in \mathcal{P}$, let $p = \frac{dP}{d\mu}$ and $q = \frac{dQ}{d\mu}$ denote the Radon-Nikodym derivatives of P and Q with respect to μ .

Two probability measures $P, Q \in \mathcal{P}$ are said to be *orthogonal* and we denote this by $Q \perp P$ if

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

Let $f : [0, \infty) \rightarrow (-\infty, \infty]$ be a convex function that is continuous at 0, i.e., $f(0) = \lim_{u \downarrow 0} f(u)$. In 1963, I. Csiszár [3] introduced the concept of f -divergence as follows.

Definition 4.1. Let $P, Q \in \mathcal{P}$. Then

$$(4.55) \quad I_f(Q, P) = \int_X p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x),$$

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

Remark 4.5. Observe that, the integrand in the formula (4.55) is undefined when $p(x) = 0$. The way to overcome this problem is to postulate for f as above that

$$(4.56) \quad 0f\left[\frac{q(x)}{0}\right] = q(x) \lim_{u \downarrow 0} \left[uf\left(\frac{1}{u}\right)\right], \quad x \in X.$$

We now give some examples of f -divergences that are well-known and often used in the literature (see also [2]).

For f continuous convex on $[0, \infty)$ we obtain the $*$ -conjugate function of f by

$$f^*(u) = uf\left(\frac{1}{u}\right), \quad u \in (0, \infty)$$

and

$$f^*(0) = \lim_{u \downarrow 0} f^*(u).$$

It is also known that if f is continuous convex on $[0, \infty)$ then so is f^* .

The following two theorems contain the most basic properties of f -divergences. For their proofs we refer the reader to Chapter 1 of [17] (see also [2]).

Theorem 4.6 (Uniqueness and Symmetry Theorem). *Let f, f_1 be continuous convex on $[0, \infty)$. We have*

$$I_{f_1}(Q, P) = I_f(Q, P),$$

for all $P, Q \in \mathcal{P}$ if and only if there exists a constant $c \in \mathbb{R}$ such that

$$f_1(u) = f(u) + c(u - 1),$$

for any $u \in [0, \infty)$.

Theorem 4.7 (Range of Values Theorem). *Let $f : [0, \infty) \rightarrow \mathbb{R}$ be a continuous convex function on $[0, \infty)$.*

For any $P, Q \in \mathcal{P}$, we have the double inequality

$$(4.57) \quad f(1) \leq I_f(Q, P) \leq f(0) + f^*(0).$$

(i) *If $P = Q$, then the equality holds in the first part of (4.57).*

If f is strictly convex at 1, then the equality holds in the first part of (4.57) if and only if $P = Q$;

(ii) *If $Q \perp P$, then the equality holds in the second part of (4.57).*

If $f(0) + f^(0) < \infty$, then equality holds in the second part of (4.57) if and only if $Q \perp P$.*

The following result is a refinement of the second inequality in Theorem 4.7 (see [2, Theorem 3]).

Theorem 4.8. *Let f be a continuous convex function on $[0, \infty)$ with $f(1) = 0$ (f is normalised) and $f(0) + f^*(0) < \infty$. Then*

$$(4.58) \quad 0 \leq I_f(Q, P) \leq \frac{1}{2} [f(0) + f^*(0)] V(Q, P)$$

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

For other inequalities for f -divergence see [1], [4]-[15].

The concept of f -divergence can be extended in a similar way for non-convex functions.

Theorem 4.9. *Let $f, h : [0, \infty) \rightarrow \mathbb{R}$ be synchronous and measurable on $[0, \infty)$. For any $P, Q \in \mathcal{P}$ we have*

$$(4.59) \quad I_{fh}(Q, P) \geq f(1) I_h(Q, P) + h(1) I_f(Q, P) - f(1) h(1).$$

Moreover, if f is normalised, then

$$(4.60) \quad I_{fh}(Q, P) \geq h(1) I_f(Q, P).$$

If both f and h are normalised, then

$$(4.61) \quad I_{fh}(Q, P) \geq 0.$$

Proof. If we write the inequality (2.4) for the synchronous functions $(\Phi, \Psi) = (f, h)$, $w = p$, $g = \frac{q}{p}$, $\Omega = X$ and $\nu = \mu$ we have

$$\begin{aligned} & \int_X pf\left(\frac{q}{p}\right)h\left(\frac{q}{p}\right)d\mu + f\left(\int_X qd\mu\right)h\left(\int_X qd\mu\right) \\ & \geq f\left(\int_X qd\mu\right)\int_X ph\left(\frac{q}{p}\right)d\mu + h\left(\int_X qd\mu\right)\int_X pf\left(\frac{q}{p}\right)d\mu \end{aligned}$$

that is equivalent to the desired result (4.59).

The rest is obvious. □

An important divergence in Information Theory is the *Kullback-Leibler divergence* obtained for the decreasing convex function $f(t) = -\ln t, t > 0$ and defined by

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

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

If $h : [0, \infty) \rightarrow \mathbb{R}$ is a decreasing function with $h(1) \geq 0$, then by (4.60) we have the inequality

$$(4.62) \quad I_{-h \ln(\cdot)}(Q, P) \geq h(1)KL(P, Q) \geq 0$$

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

In particular, we have the following inequalities

$$(4.63) \quad I_{-(\cdot)^p \ln(\cdot)}(Q, P) \geq KL(P, Q) \geq 0$$

and

$$(4.64) \quad I_{-\exp(-\alpha \cdot) \ln(\cdot)}(Q, P) \geq KL(P, Q) \exp(-\alpha) \geq 0$$

for $p, \alpha > 0$.

Theorem 4.10. *Let $f, h : [0, \infty) \rightarrow \mathbb{R}$ be Lipschitzian on $[0, \infty)$ with the constants L and K , respectively. For any $P, Q \in \mathcal{P}$ we then have*

$$(4.65) \quad |I_{fh}(Q, P) - f(1)I_h(Q, P) - h(1)I_f(Q, P) + f(1)h(1)| \leq KL\chi^2(Q, P)$$

where

$$\chi^2(Q, P) = \frac{1}{2} \int_X p \left(\frac{q}{p} - 1\right)^2 d\mu = \int_X \frac{q^2}{p} d\mu - 1$$

is Karl Pearson's χ^2 -divergence.

Moreover, if f is normalised, then

$$(4.66) \quad |I_{fh}(Q, P) - h(1)I_f(Q, P)| \leq KL\chi^2(Q, P).$$

If both f and h are normalised, then

$$(4.67) \quad |I_{fh}(Q, P)| \leq KL\chi^2(Q, P).$$

Proof. If we write the inequality (3.25) for the functions $(\Phi, \Psi) = (f, h)$, $w = p$, $g = \frac{q}{p}$, $\Omega = X$ and $\nu = \mu$ we have

$$(4.68) \quad \left| \int_X pf \left(\frac{q}{p} \right) h \left(\frac{q}{p} \right) d\mu + f \left(\int_X qd\mu \right) h \left(\int_X qd\mu \right) \right. \\ \left. - f \left(\int_X qd\mu \right) \int_X ph \left(\frac{q}{p} \right) d\mu - h \left(\int_X qd\mu \right) \int_X pf \left(\frac{q}{p} \right) d\mu \right| \\ \leq LK \left(\int_X \frac{q^2}{p} d\mu - 1 \right),$$

that is equivalent to the desired result (4.65).

The rest is obvious. \square

If some bounds for the likelihood ratio are known, then we can state the following results as well.

Theorem 4.11. *Let $P, Q \in \mathcal{P}$ such that for $0 < r < 1 < R$ we have*

$$(4.69) \quad r \leq \frac{q}{p} \leq R \text{ } \mu\text{-a.e. on } X.$$

If $f, h : [r, R] \rightarrow \mathbb{R}$ are Lipschitzian on $[r, R]$ with the constants L and K , then we have

$$(4.70) \quad |I_{fh}(Q, P) - f(1)I_h(Q, P) - h(1)I_f(Q, P) + f(1)h(1)| \\ \leq \frac{1}{4}(R-r)^2 KL.$$

Moreover, if f is normalised, then

$$(4.71) \quad |I_{fh}(Q, P) - h(1)I_f(Q, P)| \leq \frac{1}{4}(R-r)^2 KL.$$

If both f and h are normalised, then

$$(4.72) \quad |I_{fh}(Q, P)| \leq \frac{1}{4}(R-r)^2 KL.$$

If we consider the convex function $g(t) = (t-1)\ln t$, then this function generates the *Jeffreys divergence measure*

$$J(P, Q) := \int_X (p-q)(\ln p - \ln q) d\mu$$

where $P, Q \in \mathcal{P}$.

If we take $f(t) = t-1$, $h(t) = \ln t$ then f is Lipschitzian with the constant 1 and h is Lipschitzian with the constant $\frac{1}{r}$ on $[r, R]$ and by (4.72) we have

$$(4.73) \quad 0 \leq J(P, Q) \leq \frac{1}{4r}(R-r)^2$$

provided that $P, Q \in \mathcal{P}$ satisfy the condition (4.69).

The *Neyman Chi-square distance* is defined by

$$\chi_N^2(Q, P) := \frac{1}{2} \int_X \frac{(p-q)^2}{q} d\mu = \int_X \frac{p^2}{q} d\mu - 1 = \chi^2(P, Q)$$

and generated by the convex function $g(t) = \frac{(t-1)^2}{2t}$, $t > 0$.

Now, consider the functions $f(t) = \frac{1}{2}(t-1)^2$ and $h(t) = \frac{1}{t}$ defined on the interval $[r, R]$. Then $f'(t) = t-1$ and

$$\max_{t \in [r, R]} |f'(t)| = \max\{1-r, R-1\} = \frac{R-r}{2} + \left| \frac{r+R}{2} - 1 \right|.$$

Also $h'(t) = -\frac{1}{t^2}$ and

$$\max_{t \in [r, R]} |h'(t)| = \frac{1}{r^2}.$$

Then from (4.71) we have

$$(4.74) \quad \left| \chi_N^2(Q, P) - \chi^2(Q, P) \right| \leq \frac{1}{4} \left(\frac{R}{r} - 1 \right)^2 \left(\frac{R-r}{2} + \left| \frac{r+R}{2} - 1 \right| \right)$$

provided that $P, Q \in \mathcal{P}$ satisfy the condition (4.69).

Similar results may be obtained by utilizing (3.49), however the details are not presented here.

REFERENCES

- [1] P. Cerone and S. S. Dragomir, Approximation of the integral mean divergence and f -divergence via mean results. *Math. Comput. Modelling* **42** (2005), no. 1-2, 207–219.
- [2] P. Cerone, S. S. Dragomir and F. Österreicher, Bounds on extended f -divergences for a variety of classes, *Kybernetika* (Prague) **40** (2004), no. 6, 745–756. Preprint, *RGMLA Res. Rep. Coll.* **6**(2003), No.1, Article 5. [ONLINE: <http://rgmia.vu.edu.au/v6n1.html>].
- [3] I. Csiszár, Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. (German) *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8** (1963) 85–108.
- [4] S. S. Dragomir, Some inequalities for (m, M) -convex mappings and applications for the Csiszár Φ -divergence in information theory. *Math. J. Ibaraki Univ.* **33** (2001), 35–50.
- [5] S. S. Dragomir, Some inequalities for two Csiszár divergences and applications. *Mat. Bilten* No. **25** (2001), 73–90.
- [6] S. S. Dragomir, An upper bound for the Csiszár f -divergence in terms of the variational distance and applications. *Panamer. Math. J.* **12** (2002), no. 4, 43–54.
- [7] S. S. Dragomir, Upper and lower bounds for Csiszár f -divergence in terms of Hellinger discrimination and applications. *Nonlinear Anal. Forum* **7** (2002), no. 1, 1–13
- [8] S. S. Dragomir, Bounds for f -divergences under likelihood ratio constraints. *Appl. Math.* **48** (2003), no. 3, 205–223.
- [9] S. S. Dragomir, New inequalities for Csiszár divergence and applications. *Acta Math. Vietnam.* **28** (2003), no. 2, 123–134.
- [10] S. S. Dragomir, A generalized f -divergence for probability vectors and applications. *Panamer. Math. J.* **13** (2003), no. 4, 61–69.
- [11] S. S. Dragomir, Some inequalities for the Csiszár φ -divergence when φ is an L -Lipschitzian function and applications. *Ital. J. Pure Appl. Math.* No. **15** (2004), 57–76.
- [12] S. S. Dragomir, A converse inequality for the Csiszár Φ -divergence. *Tamsui Oxf. J. Math. Sci.* **20** (2004), no. 1, 35–53.
- [13] S. S. Dragomir, Some general divergence measures for probability distributions. *Acta Math. Hungar.* **109** (2005), no. 4, 331–345.
- [14] S. S. Dragomir, A refinement of Jensen's inequality with applications for f -divergence measures. *Taiwanese J. Math.* **14** (2010), no. 1, 153–164.
- [15] S. S. Dragomir, A generalization of f -divergence measure to convex functions defined on linear spaces. *Commun. Math. Anal.* **15** (2013), no. 2, 1–14.
- [16] H. Jeffreys, *Theory of Probability*, Oxford University Press, 1948, 2nd ed.
- [17] F. Liese and I. Vajda, *Convex Statistical Distances*, Teubner-Texte zur Mathematik, Band **95**, Leipzig, 1987.

MATHEMATICS, COLLEGE OF ENGINEERING & SCIENCE,
VICTORIA UNIVERSITY, PO BOX 14428,
MELBOURNE CITY, MC 8001, AUSTRALIA

SCHOOL OF COMPUTER SCIENCE & APPLIED MATHEMATICS,
UNIVERSITY OF THE WITWATERSRAND,
PRIVATE BAG 3, JOHANNESBURG 2050, SOUTH AFRICA.
E-mail address: sever.dragomir@vu.edu.au