SOME INEQUALITIES FOR $f$–DIVERGENCE MEASURES
GENERATED BY $2n$–CONVEX FUNCTIONS

S.S. DRAGOMIR AND S. KOUMANDOS

Abstract. A double Jensen type inequality for $2n$–convex functions is obtained and applied to establish upper and lower bounds for the $f$–divergence measure in Information Theory. Some particular inequalities of interest are stated as well.

1. Introduction

Let $(\Omega, A, \mu)$ be a measure space satisfying $|A| > 2$ and $\mu$ a $\sigma$–finite measure on $\Omega$. Let $\mathcal{P}$ be the set of all probability measures on the measurable space $(\Omega, A)$ which are absolutely continuous with respect to $\mu$. 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 $\mu$. 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 $f : [0, \infty) \to (-\infty, \infty]$ 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 1.** Let $P, Q \in \mathcal{P}$. Then

$$I_f(Q, P) := \int_{\Omega} p(s) f\left[\frac{q(s)}{p(s)}\right] d\mu(s),$$

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

We observe that the integrand in (1.1) is undefined when $p(s) = 0$. We can overcome this problem by postulating for $f$ as above that

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

We recall now some important classes of $f$–divergences that play a key role in various problems in Information Theory and Statistics.

**A.** The class of $\chi$–divergences. The $f$–divergences in this class are generated by the family of functions $f_\alpha(u) := |u - 1|^{\alpha}$, $u \in [0, \infty)$, $\alpha \in [1, \infty)$. They have the form:

$$I_{f_\alpha}(Q, P) = \int_{\Omega} p^{1-\alpha}(s) |q(s) - p(s)|^{\alpha} d\mu(s).$$
From this class only the parameter $\alpha = 1$ provides a distance in the metric sense, namely, the \textit{total variation distance}
\begin{equation*}
V(Q, P) := \int_{\Omega} |q(s) - p(s)| \, d\mu(s).
\end{equation*}
The most prominent special case in this class is however the Karl Pearson $\chi^2$–divergence
\begin{equation*}
I_{\chi^2}(Q, P) = \int_{\Omega} \frac{(q(s) - p(s))^2}{p(s)} \, d\mu(s) = \int_{\Omega} q^2(s) \, d\mu(s) - 1.
\end{equation*}

\textbf{B. The Dichotomy class.} This class is generated by the family of functions $g_\alpha : [0, \infty) \to \mathbb{R}$ where
\begin{equation*}
g_\alpha(u) :=
\begin{cases}
  u - 1 - \log u & \text{for } \alpha = 0, \\
  \frac{1}{\alpha(\alpha - 1)} [\alpha u + 1 - \alpha - u^\alpha] & \text{for } \alpha \in \mathbb{R} \setminus \{0, 1\}, \\
  1 - u + u \log u & \text{for } \alpha = 1.
\end{cases}
\end{equation*}
Only the parameter $\alpha = \frac{1}{2}$, that is, $g_{1/2}(u) = 2(\sqrt{u} - 1)^2$ provides a genuine distance, namely, the \textit{Hellinger distance}
\begin{equation*}
H(Q, P) := \left[ \int_{\Omega} \left( \sqrt{q(s)} - \sqrt{p(s)} \right)^2 \, d\mu(s) \right]^{\frac{1}{2}}.
\end{equation*}
Another important divergence in this class is the \textit{Kullback-Leibler divergence} obtained for $\alpha = 1$ and given by
\begin{equation*}
KL(Q, P) := \int_{\Omega} q(s) \log \left( \frac{q(s)}{p(s)} \right) \, d\mu(s).
\end{equation*}
For other classes of $f$–divergence such as \textit{Matushita’s divergence}, \textit{Puri-Vincze divergences} or \textit{Arimoto-type divergences}, see [7], [12] and [11].

Now, for a continuous convex function $f : [0, \infty) \to \mathbb{R}$, consider the $\ast$–conjugate function
\begin{equation*}
f^*(u) := uf \left( \frac{1}{u} \right) \text{ if } u \in (0, \infty)
\end{equation*}
and
\begin{equation*}
f^*(0) := \lim_{u \to 0^+} f^*(u).
\end{equation*}
It is well known that if $f$ is continuous convex on $[0, \infty)$ then $f^*$ is the same. The following results contain the most basic properties of $f$–divergences (for their proof, we refer to Chapter 1 of [11]).

\textbf{Theorem 1} (Uniqueness and Symmetry Theorem). Let $f, f_1$ be continuous convex functions on $[0, \infty)$. We have
\begin{equation*}
I_{f_1}(Q, P) = I_f(Q, P)
\end{equation*}
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 $u \in [0, \infty)$.

\textbf{Theorem 2} (Range of Values Theorem). Let $f : [0, \infty) \to \mathbb{R}$ be a convex function on $[0, \infty)$. For any $P, Q \in \mathcal{P}$ we have the double inequality
\begin{equation*}
f(1) \leq I_f(Q, P) \leq f(0) + f^*(0).
\end{equation*}
Proof. We observe that, by Taylor’s representation theorem with integral remain-

\[ F(x) = \sum_{k=0}^{2n-1} \frac{(x - a)^k}{k!} F^{(k)}(a) + \frac{1}{(2n - 1)!} \int_a^x (x - t)^{2n-1} F^{(2n)}(t) \, dt. \]

Since \( F^{(2n)}(t) \geq 0 \) for any \( t \in \hat{I} \), then on denoting by 

\[ R(x) := \int_a^x (x - t)^{2n-1} F^{(2n)}(t) \, dt, \]
we observe that \( R(x) \geq 0 \) for any \( x \geq a \). Also, for \( x < a \) we can write that
\[
R(x) = -\int_x^a (x-t)^{2n-1} F^{(2n)}(t) \, dt = \int_x^a (t-x)^{2n-1} F^{(2n)}(t) \, dt \geq 0
\]
showing that in fact \( R(x) \geq 0 \) for each \( x, a \in I \). Therefore, by (2.2) we can state that
\[
(2.3) \quad F(x) \geq F(a) + \sum_{k=1}^{2n-1} \frac{(x-a)^k}{k!} F^{(k)}(a)
\]
for any \( x, a \in I \).

Now, if we choose in (2.3) \( x = f(s), s \in \Omega \) and \( a = \int_\Omega w(z) f(z) \, d\mu(z) \), then we get
\[
(2.4) \quad F(f(x)) \geq F\left( \int_\Omega w(z) f(z) \, d\mu(z) \right) + \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} (x-a)^k \cdot F^{(k)}(f(s)) + F(a) \geq F(x)
\]
for each \( s \in \Omega \). If we multiply this inequality by \( w(s) \geq 0 \) and integrate on \( \Omega \) over the positive measure \( \mu \) we deduce the first inequality in (2.1).

By changing the place of \( x \) with \( a \) in (2.3) we also have
\[
(2.5) \quad \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} (x-a)^k \cdot F^{(k)}(x) + F(a) \geq F(x)
\]
for each \( x, a \in I \).

Now, if in (2.5) we choose \( x = f(s), s \in \Omega \) and \( a = \int_\Omega w(s) f(s) \, d\mu(s) \), we obtain
\[
(2.6) \quad \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} \left( f(s) - \int_\Omega w(z) f(z) \, d\mu(z) \right)^k \cdot F^{(k)}(f(s)) + F\left( \int_\Omega w(z) f(z) \, d\mu(z) \right) \geq F(f(s))
\]
for any \( s \in \Omega \).

Finally, if we multiply (2.6) by \( w(s) \geq 0 \) and integrate on \( \Omega \) over the positive measure \( \mu \), we deduce the second part of the inequality (2.1) and the theorem is proved.

**Corollary 1.** If \( F \) is twice differentiable and convex and \( f, F \circ f, F' \circ f \in L_w(\Omega, \mu) \), then
\[
(2.7) \quad 0 \leq \int_\Omega w(s) F(f(s)) \, d\mu(s) - F\left( \int_\Omega w(s) f(s) \, d\mu(s) \right) \leq \int_\Omega w(s) f(s) F'(f(s)) \, d\mu(s)
\]
\[\quad \quad \quad \quad \quad \quad \quad - \int_\Omega w(s) f(s) \, d\mu(s) \cdot \int_\Omega w(s) F'(f(s)) \, d\mu(s) \, .
\]
A similar result has been obtained in [6].
Remark 1. The discrete case, i.e., where \( \Omega = \{1, \ldots, m\} \) and \( \mu \) is the discrete measure, is of interest and can be stated as:

\[
(2.8) \quad \sum_{k=1}^{2n-1} \frac{F^{(k)}(\bar{x}_p)}{k!} \cdot \sum_{i=1}^{m} p_i (x_i - \bar{x}_p)^k \\
\leq \sum_{i=1}^{m} p_i F(x_i) - F(\bar{x}_p) \\
\leq \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} \cdot \sum_{i=1}^{m} p_i (x_i - \bar{x}_p)^k F^{(k)}(x_i),
\]

where \( x_i \in \hat{I}, \ p_i \geq 0 \) with \( \sum_{i=1}^{n} p_i = 1 \) and \( \bar{x}_p := \sum_{i=1}^{m} p_i x_i \in \hat{I} \), while \( F : I \rightarrow \mathbb{R} \) is as in the statement of Theorem 3. We also notice that if \( F \) is differentiable and convex, then we can deduce from (2.8) the following reverse of the Jensen inequality:

\[
(2.9) \quad 0 \leq \sum_{i=1}^{n} p_i f(x_i) - f(\bar{x}_p) \\
\leq \sum_{i=1}^{n} p_i x_i f'(x_i) - \bar{x}_p \sum_{i=1}^{n} p_i f'(x_i)
\]

that was obtained by Dragomir and Ionescu in 1994, see [5].

Remark 2. We recall that a function \( f : (a, b) \rightarrow \mathbb{R} \) which has derivatives of all orders is said to be absolutely monotonic if \( f^{(n)}(t) \geq 0 \) for all \( t \in (a, b) \) and \( n = 0, 1, \ldots \). and \( f \) is called completely monotonic if \( (-1)^n f^{(n)}(t) \geq 0 \) for all \( t \in (a, b) \) and \( n = 0, 1, 2, \ldots \). It is therefore obvious that Theorem 3 can be applied for any absolutely monotonic or completely monotonic function \( F : I \rightarrow \mathbb{R} \) and any \( n \geq 1 \). However, the class of functions \( F \) for which Theorem 3 is valid is much larger. One can choose for instance 2n–differentiable functions \( g : (a, b) \rightarrow \mathbb{R} \) with the property that \( m := \inf_{t \in (a, b)} g^{(2n)}(t) > -\infty \) and consider the new function \( F : (a, b) \rightarrow \mathbb{R}, F(t) := g(t) - \frac{m}{(2n)!} t^2 \), which is 2n–differentiable and \( F^{(2n)}(t) = g^{(2n)}(t) - m \geq 0 \) for any \( t \in (a, b) \).

Remark 3. The discrete inequality (2.8) can be utilized to provide various inequalities for means.

For instance, if we choose \( F(t) = -\log t \), then

\[
F^{(k)}(t) = \frac{(-1)^k (k-1)!}{k}, \quad k \geq 1, \ t > 0
\]

and then for any \( x_i, p_i > 0, i \in \{1, \ldots, n\} \) with \( \sum_{i=1}^{n} p_i = 1 \) and \( \bar{x}_p := \sum_{i=1}^{n} p_i x_i \) and \( G(p; x) := \prod_{i=1}^{n} x_i^{p_i} \), we have:

\[
(2.10) \quad \sum_{k=1}^{2n-1} \frac{(-1)^k}{k} \sum_{i=1}^{m} p_i \left( \frac{x_i}{\bar{x}_p} - 1 \right)^k \leq \log \left[ \frac{\bar{x}_p}{G(p; x)} \right] \\
\leq \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k} \sum_{i=1}^{m} p_i \left( \frac{\bar{x}_p}{x_i} - 1 \right)^k,
\]

for any \( n \geq 1 \).
We remark that for \( n = 1 \), we obtained the simpler inequality:

\[
0 \leq \log \left[ \frac{x_p}{G(p; x)} \right] \leq \bar{x}_p \cdot \bar{h}_p - 1,
\]

where \( \bar{h}_p := \sum_{i=1}^{m} \frac{p_i}{x_i} \) is the harmonic mean of \( x_i \) with the weights \( p_i \). This is a known result, see for instance [5].

For \( F(t) = \exp(t), \ t \in \mathbb{R} \), we get from (2.8) the following result as well:

\[
\exp(\bar{x}_p) \leq \sum_{k=1}^{2n-1} \frac{(-1)^k}{k!} \sum_{i=1}^{m} p_i (x_i - \bar{x}_p)^k 
\]

\[
\leq \sum_{i=1}^{m} p_i \exp(x_i) - \exp(\bar{x}_p) 
\]

\[
\leq \sum_{k=1}^{2n-1} \frac{(-1)^{k+1}}{k!} \sum_{i=1}^{m} p_i (x_i - \bar{x}_p) \exp(x_i) .
\]

For \( n = 1 \), the inequality (2.12) produces the following particular case of interest:

\[
0 \leq \sum_{i=1}^{m} p_i \exp(x_i) - \exp(\bar{x}_p) 
\]

\[
= \sum_{i=1}^{m} p_i x_i \exp(x_i) - \bar{x} \cdot \sum_{i=1}^{m} p_i \exp(x_i) .
\]

One can state other particular inequalities by choosing elementary function for which \( F^{(2n)}(t) \geq 0 \) on the given interval. The details are omitted.

3. Inequalities for \( F \)-Divergences

We consider now a function \( F : [0, \infty) \to \mathbb{R} \) which has the \( 2n \)-derivative \( F^{(2n)} \) nonnegative on \( (0, \infty) \). If \( P, Q \) are two probability distributions as in the introduction, we can define the following divergences:

\[
I_{\chi^k}(Q, P) := \int_{\Omega} p(s) \left( \frac{q(s)}{p(s)} - 1 \right)^k d\mu(s)
\]

and

\[
I_{g_k}(Q, P) := \int_{\Omega} p(s) g_k \left( \frac{q(s)}{p(s)} \right) d\mu(s),
\]

where the function \( g_k \) is defined by

\[
g_k(t) := (t - 1)^k F^{(k)}(t), \quad t \in [0, \infty), \ k \geq 1
\]

and is arguably simpler than the function \( F \) which generates it. This indeed happens if \( F(t) = -\log t \) since the derivatives \( F^{(k)}(t) \) are in this case rational functions. The same applies if \( F(t) = \int_0^t u(\tau) d\tau \) and the integral cannot be represented by elementary functions.

The following result provides upper and lower bounds for the \( f \)-divergence

\[
I_F(Q, P) := \int_{\Omega} p(s) F \left( \frac{q(s)}{p(s)} \right) d\mu(s)
\]

in terms of the divergences introduced in equations (3.1) and (3.3) above.
Theorem 4. Let $F : [0, \infty) \to \mathbb{R}$ be a $2n$-differentiable function, $n \geq 1$ and such that $F^{(2n)}(t) \geq 0$ on $(0, \infty)$. Then

\begin{equation}
\sum_{k=1}^{2n-1} \frac{F^{(k)}(1)}{k!} I_{I_k^*}(Q, P) \leq I_F(Q, P) - F(1) \leq \sum_{k=1}^{2n-1} \frac{(-1)^k}{k!} I_{I_k^*}(Q, P).
\end{equation}

Proof. We apply Theorem 3 for the choices $w(s) = p(s), f(x) = \frac{q(s)}{p(s)}, s \in \Omega$ to get:

\begin{equation}
\sum_{k=1}^{2n-1} \frac{F^{(k)}(\int_\Omega q(s) d\mu(s))}{k!} \cdot \int_\Omega p(s) \left[ \frac{q(s)}{p(s)} - \int_\Omega q(z) d\mu(z) \right]^k d\mu(s)
\leq \int_\Omega p(s) F \left( \frac{q(s)}{p(s)} \right) d\mu(s) - F \left( \int_\Omega q(s) d\mu(s) \right)
\leq \sum_{k=1}^{2n-1} \frac{(-1)^k}{k!} \cdot \int_\Omega p(s) \left[ \frac{q(s)}{p(s)} - \int_\Omega q(z) d\mu(z) \right]^k F^{(k)} \left( \frac{q(s)}{p(s)} \right) d\mu(s)
\end{equation}

and since $\int_\Omega q(s) d\mu(s) = 1$, hence, with notations (3.1) and (3.3), we observe that (3.5) is exactly the desired inequality (3.4).

We observe that for $n = 1$, $I_{I_1^*}(Q, P) = 0$ and

\begin{align*}
I_{I_1^*}(Q, P) &= \int_\Omega p(s) \left( \frac{q(s)}{p(s)} - 1 \right) F' \left( \frac{q(s)}{p(s)} \right) d\mu(s) \\
&= \int_\Omega (q(s) - p(s)) F' \left( \frac{q(s)}{p(s)} \right) d\mu(s),
\end{align*}

therefore the following particular case may be stated:

Corollary 2. Assume that $F : [0, \infty) \to \mathbb{R}$ is continuous and twice differentiable on $[0, \infty)$. If $F$ is convex on $[0, \infty)$, then the following inequality can be stated:

\begin{equation}
0 \leq I_F(Q, P) - F(1) \leq \delta_F(Q, P),
\end{equation}

where

\begin{equation}
\delta_F(Q, P) := \int_\Omega (q(s) - p(s)) F' \left( \frac{q(s)}{p(s)} \right) d\mu(s)
\end{equation}

is the general $\delta$-divergence measure introduced in the recent paper [4] by the first author.

For the definition of $\delta$-divergence measures and some of its fundamental properties, see [4].

Remark 4. It is well known that the Gamma function

\[ \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0 \]

is logarithmic-convex on $(0, \infty)$. Therefore, one can consider the divergences generated by the convex function $\log \Gamma$, i.e.,

\[ I_{\log \Gamma}(Q, P) := \int_\Omega p(s) \log \left( \frac{q(s)}{p(s)} \right) d\mu(s). \]
If we denote by $\Psi(t) = \frac{d}{dt} \log \Gamma(t) = \frac{\Gamma'(t)}{\Gamma(t)}$, which is well known in the literature as the “psi” or “digamma function”, then utilizing Corollary 2, we can conveniently connect the $f$–divergence and $\log \Gamma$ with the $\delta$–divergence of $\Psi$ via an inequality. We have therefore the inequality:

$$0 \leq I_{\log \Gamma}(Q, P) \leq \delta_{\Psi}(Q, P),$$

for any $P, Q \in \mathcal{P}$, where, as above

$$\delta_{\Psi}(Q, P) := \int_{\Omega} \frac{(q(s) - p(s)) \Psi\left(\frac{q(s)}{p(s)}\right)}{\Psi(q(s))} \, d\mu(s).$$

If we consider now the zeta function $\zeta(x) := \sum_{n=0}^{\infty} \frac{1}{n^x}$, $x > 1$ then we have that $\zeta(x+1)$ is log-convex and if we denote by

$$m(t) := \frac{d}{dt} \log \zeta(t + 1) = \frac{\zeta'(t + 1)}{\zeta(t + 1)}, \quad t > 0,$$

then the following representation is well known

$$m(t) = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{t+1}}, \quad t > 1,$$

where $\Lambda(n)$ is the van Mangoldt function, i.e.,

$$\Lambda(n) := \begin{cases} 
\log p & \text{if } n = p^k \quad (p \text{ prime, } k \geq 1), \\
0 & \text{otherwise.}
\end{cases}$$

We can label $m(t)$ to be a “distance function” by following the same approach as that for gamma.

We can then introduce the $f$–divergence of $\log \zeta(x+1)$ and state, by utilizing Corollary 2 that

$$0 \leq I_{\log \zeta(x+1)}(Q, P) \leq \delta_{m}(Q, P),$$

for any $P, Q \in \mathcal{P}$, where $\delta_{m}(Q, P)$ is the $\delta$–divergence of $m$ defined above.

Similar results may be considered for other log-convex functions for which the log-derivatives are function of specific interest and generating $\delta$–divergences which can be easier to calculate/estimate. In the next section we give a different type of applications of $F$–divergence inequalities.

4. Applications to product measures

In this section we prove the following

**Proposition 1.** Let $a_n$ and $b_n$ be sequences of real numbers such that $|a_n| < 1$, $|b_n| < 1$. Then for all positive integers $n$ we have

$$\prod_{k=1}^{n} \frac{(1 + b_k)}{(1 + a_k)} \frac{(1 - b_k)}{(1 - a_k)} \frac{1}{2} \frac{1 - a_k}{2} \geq \exp \left[ 1 - \prod_{k=1}^{n} \left( 1 + \frac{(a_k - b_k)^2}{1 - b_k^2} \right) \right].$$

Before proving (4.1) let us state an application of it.
Corollary 3. Suppose that the sequences $a_n$ and $b_n$ are as above. If

$$\sum_{n=1}^{\infty} \frac{(a_n - b_n)^2}{1 - b_n^2} < \infty,$$

then the infinite product

$$\prod_{n=1}^{\infty} \left( \frac{1 + b_n}{1 + a_n} \right)^{\frac{1 + a_n}{2}} \left( \frac{1 - b_n}{1 - a_n} \right)^{\frac{1 - a_n}{2}},$$

converges.

Proof. It is clear that

$$0 < \left( \frac{1 + b_k}{1 + a_k} \right)^{\frac{1 + a_k}{2}} \left( \frac{1 - b_k}{1 - a_k} \right)^{\frac{1 - a_k}{2}} < 1.$$  

Then use condition (4.2) in combination with (4.1) to complete the proof.

Remark 5. We observe that the conditions

$$\sum_{n=1}^{\infty} (a_n - b_n)^2 < \infty \quad \text{and} \quad \lim \sup |b_n| < 1$$

imply (4.2).

We now give a proof of (4.1)

Proof. Let $r_n(x), n = 1, 2, \ldots$ be the Rademacher functions defined on the interval $[0, 1]$ by the relation $r_n(x) = \text{sign} \sin 2^n \pi x$. These functions form an orthonormal set in the Hilbert space $L^2([0, 1])$ and also

$$\int_0^1 r_n(x) dx = 0, \quad n = 1, 2, \ldots.$$  

The functions $r_n(x)$ are also independent random variables in the probability measure space $([0, 1], B, \lambda)$, where $\lambda$ is the Lebesgue measure on the $\sigma$-algebra $B$ of Borel subsets of $[0, 1]$. Let $a_n$ and $b_n$ be sequences of real numbers such that $|a_n| < 1, |b_n| < 1$. For every $n \in \mathbb{N}$ the relations

$$dP = \prod_{k=1}^{n} (1 + a_k r_k) d\lambda, \quad dQ = \prod_{k=1}^{n} (1 + b_k r_k) d\lambda,$$

define probability measures on $([0, 1], B, \lambda)$, which are absolutely continuous with respect to $\lambda$.

For $F(t) = -\log t$ we will calculate the $F$-divergence of the probability distributions $P$ and $Q$ defined in (4.5), viz.

$$I_F(Q, P) = -\int_0^1 \prod_{k=1}^{n} (1 + a_k r_k(x)) \log \left( \frac{\prod_{k=1}^{n} (1 + b_k r_k(x))}{\prod_{k=1}^{n} (1 + a_k r_k(x))} \right) dx$$

$$= -\int_0^1 \log \left( \frac{\prod_{k=1}^{n} (1 + b_k r_k(x))}{\prod_{k=1}^{n} (1 + a_k r_k(x))} \right) dP.$$  

We observe that

$$\log(1 + b_k r_k(x)) = \frac{1}{2} \log(1 - b_k^2) + \frac{1}{2} \log \left( \frac{1 + b_k}{1 - b_k} \right) r_k(x).$$
It is easy to see that
\[ \int_0^1 r_k(x) \, dP = a_k, \quad k = 1, 2, \ldots \]
Then by (4.7) we obtain
\[ \int_0^1 \sum_{k=1}^n \log(1 + b_k r_k(x)) \, dP = \frac{1}{2} \sum_{k=1}^n \log(1 - b_k^2) + \frac{1}{2} \sum_{k=1}^n a_k \log \frac{1 + b_k}{1 - b_k} \]
Hence, by (4.6) we find that
\[ -I_F(Q, P) = \frac{1}{2} \sum_{k=1}^n \log \frac{1 - b_k^2}{a_k^2} + \frac{1}{2} \sum_{k=1}^n a_k \log \frac{(1 + b_k)(1 - a_k)}{(1 - b_k)(1 + a_k)}, \]
and finally
\[ I_F(Q, P) = -\log \prod_{k=1}^n \left( \frac{1 + b_k}{1 + a_k} \right)^{\frac{1 + a_k}{2}} \left( \frac{1 - b_k}{1 - a_k} \right)^{\frac{1 - a_k}{2}}. \]
Now we set
\[ p(x) = \prod_{k=1}^n (1 + a_k r_k(x)), \quad q(x) = \prod_{k=1}^n (1 + b_k r_k(x)) \]
and calculate the \( \delta \)-divergence measure of \( P \) and \( Q \). We have
\[ \delta_F'(Q, P) = -\int_0^1 (q(x) - p(x)) \frac{p(x)}{q(x)} \, dx = -1 + \int_0^1 \frac{p(x)^2}{q(x)} \, dx \]
\[ = -1 + \sum_{k=1}^n \frac{(1 + a_k r_k(x))^2 (1 - b_k r_k(x))}{1 - b_k^2} \]
Combining (4.9) with (4.10) and using the inequality (3.6) we conclude the proof of (4.1).

5. Remarks

(1) Suppose that \( b_k = 0 \) for all \( k \in \mathbb{N} \). Then inequality (4.1) is equivalent to
\[ \prod_{k=1}^n (1 + a_k^2) \geq 1 + \frac{1}{2} \log \prod_{k=1}^n (1 + a_k)^{1+a_k} (1 - a_k)^{1-a_k}. \]
Inequality (5.1) has the following interpretation:
Let \( \mu \) be the Borel probability measure on \([0, 1]\) defined by
\[ d\mu = \lim_{n \to \infty} \prod_{k=1}^n (1 + a_k r_k(x)) \, dx, \]
where the limit is in the weak\( \ast \) sense. This measure is absolutely continuous with respect to the Lebesgue measure if and only if \( \sum_{n=1}^{\infty} a_n^2 < \infty \). Compare the paper [8].
Suppose now that $M$ is a Borel subset of the support of $\mu$ such that $\mu(M) > 0$. It follows from the results of [9] (see also [1]) that for the Hausdorff dimension of $M$, $\dim M$, one has

\[
\dim M = 1 - \limsup_n \frac{1}{n \log 4} \log \prod_{k=1}^n (1 + a_k)^{1 + a_k} (1 - a_k)^{1 - a_k}.
\]

It follows from this and (5.1) that in the case where $\sum_{n=1}^\infty a_n^2 < \infty$ for every Borel subset $M$ of the support of $\mu$ with $\mu(M) > 0$, we have $\dim M = 1$. Of course, this happens because in this case $\mu$ is absolutely continuous with respect to Lebesgue measure $\lambda$ and thus $\lambda(M) > 0$.

(2) Let the probability measure $\mu$ be defined by (5.2) and $\nu$ be a product measure of the same type, that is

\[
d\nu = \lim_{n \to \infty} \prod_{k=1}^n (1 + b_k r_k(x)) \, dx.
\]

In the case where $(a_n), (b_n) \in \ell^2$ we can calculate the $F$-divergence measure of the probability distributions $\mu, \nu$ for $F(t) = -\log t$. Indeed, a small adaptation of the proof of Proposition 1 yields

\[
I_F(\nu, \mu) = -\log \prod_{k=1}^\infty \left( \frac{1 + b_k}{1 + a_k} \right)^{1 + a_k} \left( \frac{1 - b_k}{1 - a_k} \right)^{1 - a_k}.
\]

Since the sequences $(a_n), (b_n) \in \ell^2$, both $\mu$ and $\nu$ is absolutely continuous with respect to the Lebesgue measure $\lambda$ and the condition (4.2) is satisfied therefore by Corollary 3, the infinite product above converges.

Finally, we note that we can obtain results analogous to (4.1), (5.1) and (5.5) using the more general product measures considered in [9] and [10].

REFERENCES


**SCHOOL OF COMPUTER SCIENCE AND MATHEMATICS, VICTORIA UNIVERSITY, PO BOX 14428, MCMC 8001, VIC, AUSTRALIA.**

*E-mail address*: sever.dragomir@vu.edu.au

*URL*: http://rgmia.vu.edu.au/dragomir

**DEPARTMENT OF MATHEMATICS AND STATISTICS, THE UNIVERSITY OF CYPRUS, P.O. BOX 20537, 1678 NICOSIA, CYPRUS**

*E-mail address*: skoumand@ucy.ac.cy

*URL*: http://www.mas.ucy.ac.cy/teachers/koumantos.htm