Extensions of recent combinatorial refinements of discrete and integral Jensen inequalities

LÁSZLÓ HORVÁTH

Abstract. The main purpose of this work is to present essential extensions of results in [7] and [8], and apply them to some special situations. Of particular interest is the refinement of the integral Jensen inequality for vector valued integrable functions. The applications related to four topics, namely $f$-divergences in information theory (an interesting refinement of the weighted geometric mean–arithmetic mean inequality is obtained as a consequence), norm inequalities, quasi-arithmetic means, Hölder’s and Minkowski’s inequalities.

Mathematics Subject Classification. 26D15, 26B25.

Keywords. Discrete and integral Jensen inequalities, Refinement, Finite and infinite permutations.

1. Introduction

The most important inequality regarding convex functions is Jensen’s inequality. A real function $f$ defined on a convex subset $C$ of a real vector space is called convex if it satisfies

$$f(\alpha v_1 + (1 - \alpha) v_2) \leq \alpha f(v_1) + (1 - \alpha) f(v_2)$$

for all $v_1, v_2 \in C$ and all $\alpha \in [0, 1]$.

Discrete and integral variants of Jensen’s inequality are formulated in the next two assertions.

**Theorem 1.1.** (discrete Jensen inequalities, see [12] and [14]) (a) Let $C$ be a convex subset of a real vector space $V$, and let $f : C \to \mathbb{R}$ be a convex function. If $p_1, \ldots, p_n$ are nonnegative numbers with $\sum_{i=1}^{n} p_i = 1$, and $v_1, \ldots, v_n \in C$, then

$$f \left( \sum_{i=1}^{n} p_i v_i \right) \leq \sum_{i=1}^{n} p_i f(v_i). \quad (1.1)$$
(b) Let $C$ be a closed convex subset of a real Banach space $V$, and let $f : C \to \mathbb{R}$ be a convex function. If $p_1, p_2, \ldots$ are nonnegative numbers with $\sum_{i=1}^{\infty} p_i = 1$, and $v_1, v_2, \ldots \in C$ such that the series $\sum_{i=1}^{\infty} p_i v_i$ and $\sum_{i=1}^{\infty} p_i f(v_i)$ are absolutely convergent, then
\[
f \left( \sum_{i=1}^{\infty} p_i v_i \right) \leq \sum_{i=1}^{\infty} p_i f(v_i).
\] (1.2)

Theorem 1.2. (integral Jensen inequalities, see [12]) Let $(X, A, \mu)$ be a probability space.

(a) Let $\varphi : X \to \mathbb{R}$ be an integrable function taking values in an interval $C \subset \mathbb{R}$. Then $\int_X \varphi d\mu$ lies in $C$. If $f$ is a convex function on $C$ such that $f \circ \varphi$ is $\mu$-integrable, then
\[
f \left( \int_X \varphi d\mu \right) \leq \int_X f \circ \varphi d\mu.
\] (1.3)

(b) Let $\varphi : X \to \mathbb{R}^d$ be an integrable function taking values in a closed convex set $C \subset \mathbb{R}^d$. Then $\int_X \varphi d\mu$ lies in $C$. If $f$ is a continuous convex function on $C$ such that $f \circ \varphi$ is $\mu$-integrable, then (1.3) also holds.

Remark 1.3. It follows from Theorem 1.2 (a) that Theorem 1.1 (b) can be generalized the case of $V = \mathbb{R}$: if $C \subset \mathbb{R}$ is an interval (not necessarily closed) and the other conditions of the statement are satisfied, then $\sum_{i=1}^{\infty} p_i v_i$ lies in $C$ and (1.2) holds.

The set of nonnegative integers and the set of positive integers will be denoted by $\mathbb{N}$ and $\mathbb{N}_+$, respectively.

Refining Jensen’s inequality is an intensively studied research area (see e.g. the book [6] and references therein). A multitude of applications (for other famous inequalities, in the theory of means, in information theory, etc.) underscores the importance of this topic. Recently, a lot of papers have been published dealing with the so called “cyclic refinements” of Jensen’s inequality (see e.g. the book [10] and the references therein). These papers mainly focus on the discrete Jensen inequality of the form (1.1), and they use $(j-1)$-cyclic permutations of the set $\{1, \ldots, n\}$ to the right $(j = 1, \ldots, k, \ 2 \leq k \leq n)$. The obtained results cannot be generalized for the discrete Jensen inequality containing infinite sums (1.2). It is not easy to give cyclic refinements of the integral Jensen inequality either. To overcome these difficulties, in the paper [7] we refined both (1.1) and (1.2) by using either permutations of the set $\{1, \ldots, n\}$ or bijections from $\mathbb{N}_+$ onto itself. The method can also be applied to the integral Jensen inequality (see [8]). The main purpose of this work is to
present essential extensions of results in [7] and [8], and apply them to some special situations. Of particular interest is the refinement of the integral Jensen inequality for vector valued integrable functions. The author could not find a similar result in the literature. The applications related to four topics, namely $f$-divergences (for discrete distributions supported on finite or countably infinite sets) in information theory (an interesting refinement of the weighted geometric mean - arithmetic mean inequality is obtained as a consequence), norm inequalities, quasi-arithmetic means, Hölder’s and Minkowski’s inequalities.

2. Preliminary results

We shall use the following result for double series (a more general form can be found e.g. in [16]).

**Theorem 2.1.** (Interchanging the order of summation) Let $(V, \|\cdot\|)$ be a real Banach space. Let $v(i, j) \in V$ $(i, j) \in \mathbb{N}_+ \times \mathbb{N}_+$. If any of the two sums

$$\sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} \|v(i, j)\| \right), \quad \sum_{j=1}^{\infty} \left( \sum_{i=1}^{\infty} \|v(i, j)\| \right)$$

is finite, then all of the series

$$\sum_{i=1}^{\infty} v(i, j), \quad \sum_{j=1}^{\infty} v(i, j)$$

and

$$\sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} v(i, j) \right), \quad \sum_{j=1}^{\infty} \left( \sum_{i=1}^{\infty} v(i, j) \right) \quad (2.1)$$

are absolutely convergent and the series in (2.1) have the same sum.

We need another result concerning the existence of special convex functions. The positive part $f^+$ and the negative part $f^-$ of a real valued function $f$ are defined in the usual way.

**Lemma 2.2.** Let $(X, \mathcal{A}, \mu)$ be a probability space, and let $\varphi : X \to \mathbb{R}^d$ be an integrable function taking values in a closed convex set $C \subset \mathbb{R}^d$. If $f$ is a continuous convex function on $C$ such that $f \circ \varphi$ is $\mu$-integrable, then there exists a convex function $g$ on $C$ such that $|f| \leq g$ and $g \circ \varphi$ is $\mu$-integrable too.

**Proof.** Along with the function $f$ the function $f^+$ is also convex. Since $f \circ \varphi$ is $\mu$-integrable, $f^+ \circ \varphi$ is also $\mu$-integrable.
Because \( f \) is a continuous convex function on \( C \), there is an affine function \( l : \mathbb{R}^d \to \mathbb{R} \), \( l(u) = h(u) + b \) (\( h : \mathbb{R}^d \to \mathbb{R} \) is a linear functional and \( b \in \mathbb{R} \)) for which

\[
f(u) \geq h(u) + b, \quad u \in C.
\]

Then

\[
f^-(u) = \max(-f(u), 0) \leq \max(-h(u) - b, 0) = l^-(u) \leq |l(u)|, \quad u \in C.
\]

Since the linear functional \( h \) is continuous and the function \( \varphi \) is \( \mu \)-integrable, \( |l| \circ \varphi \) is also \( \mu \)-integrable.

It is easy to check that \( |l| \) is convex. The sum of two convex function is also convex, and therefore \( g := f^+ + |l| \) is convex.

Since

\[
|f| = f^+ + f^- \leq f^+ + |l| = g,
\]

it follows from the above that \( g \) is an appropriate choice.

The proof is complete. \( \square \)

**Remark 2.3.** In the real case \((d = 1)\) the previous statement is valid without assuming the closedness of the convex interval \( C \) and the continuity of the convex function \( f \) (see Lemma 6 in [8]).

### 3. Main result

Let the set \( I \) denote either \( \{1, \ldots, n\} \) for some \( n \geq 1 \) or \( \mathbb{N}_+ \). We say that the numbers \((p_i)_{i \in I}\) represent a (positive) discrete probability distribution if \((p_i > 0) \ p_i \geq 0 \ (i \in I) \) and \( \sum_{i \in I} p_i = 1 \). A permutation \( \pi \) of \( I \) refers to a bijection from \( I \) onto itself.

We need the following hypotheses:

(H\(_1\)) Let the index set \( I \) denote either \( \{1, \ldots, n\} \) for some \( n \geq 1 \) or \( \mathbb{N}_+ \).

Let the index set \( J \) denote either \( \{1, \ldots, k\} \) for some \( k \geq 1 \) or \( \mathbb{N}_+ \).

(H\(_2\)) For each \( j \in J \) let \( \pi_j \) be a permutation of the set \( I \).

(H\(_3\)) Let \( \Lambda = (\lambda_{i,j})_{(i,j) \in I} \) be a matrix with positive entries satisfying

\[
\sum_{j \in J} \lambda_{\pi_i^{-1}(i), j} = 1, \quad i \in I.
\]

(H\(_{d1}\)) Let \((p_i)_{i \in I}\) represent a positive probability distribution.

(H\(_{d2}\)) Let \( C \) be a closed convex subset of a real Banach space \((V, \|\cdot\|)\), and \( f : C \to \mathbb{R} \) be a convex function.

The first result of this paper is the following refinement of the discrete Jensen inequality.
Theorem 3.1. Assume \((H_1 – H_3)\) and \((H_{d1} – H_{d2})\). If \((v_i)_{i \in I}\) is a sequence of vectors from \(C\) such that the series \(\sum_{i \in I} p_i v_i\) and \(\sum_{i \in I} p_i f (v_i)\) are absolutely convergent, then

\[
f \left( \sum_{i \in I} p_i v_i \right) \leq C_{dg} = C_{dg} (f, v, p, \Lambda, \pi) \tag{3.2}
\]

\[
:= \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right) f \left( \sum_{j \in J} \frac{\lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right) \leq \sum_{i \in I} p_i f (v_i) \tag{3.3}
\]

Proof. In this proof we don’t distinguish between finite and infinite sums, their common name is series. A series with a finite number of vectors has a finite evaluation, and hence it is obviously considered (absolutely) convergent.

We proceed in three steps.

(a) We first study the convergence of some series.

\((a_1)\) Since \(0 < \lambda_{i,j} \leq 1 (i \in I, j \in J)\) and the series \(\sum_{l \in I} p_l v_l\) is absolutely convergent,

\[
\sum_{l \in I} \lambda_{\pi_j^{-1}(l), j} p_l \|v_l\| \leq \sum_{l \in I} p_l \|v_l\| < \infty, \quad j \in J. \tag{3.4}
\]

By \((H_2)\), for a fixed \(j \in J\) the series

\[
\sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \|v_{\pi_j(i)}\|
\]

is a rearrangement of the first series in \((3.4)\), and therefore

\[
\sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \|v_{\pi_j(i)}\| = \sum_{l \in I} \lambda_{\pi_j^{-1}(l), j} p_l \|v_l\|, \quad j \in J. \tag{3.5}
\]

By using the absolute convergence of the series \(\sum_{l \in I} p_l v_l\) again, it follows from \((H_3)\) that

\[
\sum_{l \in I} \left( \sum_{j \in J} \lambda_{\pi_j^{-1}(l), j} p_l \|v_l\| \right) = \sum_{l \in I} p_l \|v_l\| < \infty. \tag{3.6}
\]

Then Theorem 2.1 implies that

\[
\sum_{j \in J} \left( \sum_{l \in I} \lambda_{\pi_j^{-1}(l), j} p_l \|v_l\| \right) < \infty,
\]

and thus \((3.5)\) and \((3.6)\) show that

\[
\sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \|v_{\pi_j(i)}\| \right) = \sum_{i \in I} p_i \|v_i\| < \infty.
\]
Now another application of Theorem 2.1 yields that the series
\[ \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \| v_{\pi_j(i)} \|, \quad i \in I \]
and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \| v_{\pi_j(i)} \| \right), \quad \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \| v_{\pi_j(i)} \| \right) \]
are (absolutely) convergent and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \| v_{\pi_j(i)} \| \right) = \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \| v_{\pi_j(i)} \| \right) = \sum_{l \in I} p_l \| v_l \| < \infty. \]

(a2) Based on what is proved in part (a1), we have from Theorem 2.1 that the series
\[ \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}, \quad i \in I \] (3.7)
and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)} \right), \quad \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)} \right) \]
are absolutely convergent and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)} \right) = \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)} \right) = \sum_{l \in I} p_l v_l. \] (3.8)

(a3) Choose a fixed vector \( u \in V \) such that \( \| u \| = 1 \), and let \( v_i := u \) (\( i \in I \)). Then by (H_d1), the series \( \sum_{l \in I} p_l v_l \) is absolutely convergent. It follows immediately from the establishments in (a1) that the series
\[ \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}, \quad i \in I \] (3.9)
and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right), \quad \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \right) \]
are (absolutely) convergent and
\[ \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right) = \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} \right) = 1. \] (3.10)
(a₄) Since the series \( \sum_{l \in I} p_l f(v_l) \) is absolutely convergent, we can prove by following the method used in (a₁) and (a₂) that the series
\[
\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) , \quad i \in I
\] (3.11)
and
\[
\sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) \right) , \quad \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) \right)
\]
are absolutely convergent and
\[
\sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) \right) = \sum_{j \in J} \left( \sum_{i \in I} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) \right) = \sum_{l \in I} p_l f(v_l) .
\] (3.12)

(b) In this part of the proof we show that inequality (3.3) holds. We have already proved that the series (3.7), (3.9) and (3.11) are absolutely convergent, and hence we now conclude from Theorem 1.1 (b) that
\[
f \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right) \leq \frac{1}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) , \quad i \in I.
\]
This implies by (3.12) that
\[
\sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} f(v_{\pi_j(i)}) \right)
\]
\[
\leq \sum_{i \in I} p_i f(v_i) .
\] (3.13)

(c) Finally, we prove inequality (3.2), and that the series
\[
\sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right)
\] (3.14)
is absolutely convergent.

(c₁) Since the series \( \sum_{i \in I} p_i f(v_i) \) is absolutely convergent,
\[
0 \leq \sum_{i \in I} p_i f^+(v_i) \leq \sum_{i \in I} p_i |f(v_i)| < \infty,
\]
and hence the convexity of \( f^+ \) implies that inequality (3.13) can be applied to \( f^+ \), and we get
\[
0 \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} \right) f^+ \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)} v_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right) < \infty .
\] (3.15)
(c2) Assume $f$ is bounded below, that is there exists $c < 0$ such that
\[ f(v) \geq c, \quad v \in C. \]

Then $f^- (v) \leq -c$ ($v \in C$), and therefore (3.10) yields that
\[ 0 \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v) \right) f^- \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v)}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right) \leq -c. \]

By using (3.15), we obtain that in this case the series (3.14) is absolutely convergent. Hence by (3.10), Theorem 1.1 (b) can be applied to (3.14), and we have from (3.8) that
\[
f \left( \sum_{i \in I} p_i v_i \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v) \right) f^- \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v)}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right), \tag{3.16}
\]
which is precisely (3.2).

(c3) If $f$ is not bounded below, define
\[ f_n : C \to \mathbb{R}, \quad f_n(v) = \max(f(v), -n), \quad n \in \mathbb{N}_+. \]

Since the series $\sum_{i \in I} p_i f(v_i)$ is absolutely convergent, and $|f| \geq |f_n|$ ($n \in \mathbb{N}_+$),
\[ 0 \leq \sum_{i \in I} p_i |f_n(v_i)| \leq \sum_{i \in I} p_i |f(v_i)| < \infty, \]
and thus the series $\sum_{i \in I} p_i f_n(v_i)$ ($n \in \mathbb{N}_+$) is also absolutely convergent. It is also true that $f_n$ is convex and bounded below ($n \in \mathbb{N}_+$). These show that (3.16) can be applied (replacing $f$ by $f_n$), and we obtain
\[ f_n \left( \sum_{i \in I} p_i v_i \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v) \right) f_n \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v)}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right), \quad n \in \mathbb{N}_+, \]
and the series on the right is absolutely convergent ($n \in \mathbb{N}_+$).

Furthermore, the sequence $(f_n)_{n \in \mathbb{N}_+}$ is decreasing and converges pointwise to $f$. Now it follows from Beppo Levi's theorem that
\[ f \left( \sum_{i \in I} p_i v_i \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v) \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}(v)}{\sum_{j \in J} \lambda_{i,j} p_{\pi_j(i)}} \right), \]
which gives (3.2) and that the series (3.14) is absolutely convergent.

The proof is complete. \[\square\]

Analyzing the proof of the previous theorem, we obtain the following extensions of it.
Corollary 3.2. Assume \((H_1–H_3,H_{d1})\), assume further that the index sets \(I\) and \(J\) are finite, and replace \((H_{d2})\) by the weaker condition

\((H_{df})\) Let \(C\) be a convex subset of a real vector space, and \(f : C \to \mathbb{R}\) be a convex function.

If \((v_i)_{i \in I}\) is a sequence of vectors from \(C\), then inequalities (3.2–3.3) are satisfied.

Proof. In this case we can work with finite sums, and hence Theorem 1.1 (a) can be used in the proof of Theorem 3.1. □

Corollary 3.3. Assume \((H_1–H_3,H_{d1})\), and assume further that

\((H_{dr})\) Let \(C\) be an arbitrary interval in \(\mathbb{R}\), and \(f : C \to \mathbb{R}\) be a convex function.

If \((v_i)_{i \in I}\) is a sequence of numbers from \(C\) such that the series \(\sum_{i \in I} p_i v_i\) and \(\sum_{i \in I} p_i f(v_i)\) are absolutely convergent, then inequalities (3.2–3.3) hold.

Proof. By referring to Remark 1.3, it follows immediately from Theorem 3.1. □

Remark 3.4. Assume \((H_1–H_2)\) and \((H_{d1}–H_{d2})\). Let \((\lambda_j)_{j \in J}\) represent a positive probability distribution and let the matrix \(\Lambda = (\lambda_{i,j})_{i \in I, j \in J}\) be defined by

\[ \lambda_{i,j} := \lambda_j, \quad i \in I, \quad j \in J, \]

that is all rows of \(\Lambda\) are the same, or otherwise every member of the \(j\)th column is \(\lambda_j\). Then (3.1) is obviously satisfied. It follows that Theorem 3.1 contains the main result in [7] (see Theorem 2.1 in [7]) as a special case.

To formulate the integral version of Theorem 3.1 we introduce the following new conditions:

\((H_{i1})\) Let \((X, \mathcal{A}, \mu)\) be a probability space.

\((H_{i2})\) Suppose we are given a sequence \(\mathcal{M}_I = (\mu_i)_{i \in I}\) of measures on \(\mathcal{A}\) with \(\mu_i(X) > 0\) for all \(i \in I\) and \(\sum_{i \in I} \mu_i = \mu\).

Now, we have the next refinement of the integral Jensen inequality.

Theorem 3.5. Assume \((H_1–H_3)\) and \((H_{i1}–H_{i2})\).

(a) Let \(C \subset \mathbb{R}\) be an interval, and \(f : C \to \mathbb{R}\) be a convex function. Let \(\varphi\) be a \(\mu\)-integrable function on \(X\) taking values in \(C\) such that \(f \circ \varphi\) is also...
\( \mu \)-integrable on \( X \). Then
\[
f \left( \int_X \varphi \, d\mu \right) \leq C_{ig} = C_{ig} \left( f, \varphi, \mu, \Lambda, \pi, \mathcal{M} \right)
\]
\[
:= \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X) \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} \int_X \varphi \, d\mu_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X)} \right)
\] (3.17)
\[
\leq \sum_{i \in I} \mu_i(X) f \left( \frac{1}{\mu_i(X)} \int_X \varphi \, d\mu_i \right) \leq \int_X f \circ \varphi \, d\mu.
\] (3.18)

(b) Let \( C \subset \mathbb{R}^d \) be a closed convex set, and \( f : C \to \mathbb{R} \) be a continuous convex function. If \( \varphi \) is a \( \mu \)-integrable function on \( X \) taking values in \( C \) such that \( f \circ \varphi \) is also \( \mu \)-integrable on \( X \), then inequalities (3.17–3.18) are satisfied too.

**Proof.** (a) By using Corollary 3.3, we can copy the proof of Theorem 8 in [8].

(b) The line of reasoning is similar to that of (a), but for the sake of completeness, we give the proof.

Define
\[
p_i := \mu_i(X) \quad \text{and} \quad v_i := \frac{1}{\mu_i(X)} \int_X \varphi \, d\mu_i, \quad i \in I.
\]

Then \( (p_i)_{i \in I} \) represents a positive probability distribution, and by Theorem 1.2 (b), \( v_i \in C \) (\( i \in I \)).

Now we show that the series \( \sum_{i \in I} p_i v_i \) and \( \sum_{i \in I} p_i f(v_i) \) are absolutely convergent.

Since \( \varphi \) is a \( \mu \)-integrable function on \( X \) and \( \sum_{i \in I} \mu_i = \mu \),
\[
\sum_{i \in I} p_i |v_i| = \sum_{i \in I} \left| \int_X \varphi \, d\mu_i \right| \leq \sum_{i \in I} \int_X |\varphi| \, d\mu_i = \int_X |\varphi| \, d\mu < \infty.
\]

Lemma 2.2 shows that there exists a convex function \( g \) on \( C \) such that
\[
|f| \leq g \quad \text{and} \quad g \circ \varphi \quad \text{is} \quad \mu \text{-integrable.} \tag{3.19}
\]

By applying Theorem 1.2 (b) again and using \( \sum_{i \in I} \mu_i = \mu \) we obtain that
\[
\sum_{i \in I} p_i |f(v_i)| = \sum_{i \in I} \mu_i(X) \left| f \left( \frac{1}{\mu_i(X)} \int_X \varphi \, d\mu_i \right) \right| 
\]
\[
\leq \sum_{i \in I} \mu_i(X) g \left( \frac{1}{\mu_i(X)} \int_X \varphi \, d\mu_i \right) \leq \sum_{i \in I} \int_X g \circ \varphi \, d\mu_i = \int_X g \circ \varphi \, d\mu < \infty.
\]
It follows that the conditions of Theorem 3.1 hold, and therefore an application of this result gives that
\[
\int_X \varphi d\mu = f \left( \sum_{i \in I} p_i f(v_i) \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X) \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X)}{\sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}} \right) \leq \sum_{i \in I} \mu_i \left( A_i \right) \left( \frac{1}{\mu(A_i)} \int_{A_i} \varphi d\mu \right) \leq \int_X f \circ \varphi d\mu. \tag{3.20}
\]

As a final step, we can apply Theorem 1.2 (b) in (3.20). The proof is complete. \[\square\]

A useful consequence of the previous theorem is the next result.

**Corollary 3.6.** Assume \(H_1 - H_3, H_{i^k}\) and \(H_{is}\) Suppose we are given a sequence \(\mathcal{S}_I = (A_i)_{i \in I}\) of pairwise disjoint sets \(A_i \in \mathcal{A}\) with \(\mu(A_i) > 0\) for all \(i \in I\) and \(\bigcup_{i \in I} A_i = X\).

(a) Let \(C \subset \mathbb{R}\) be an interval, and \(f : C \to \mathbb{R}\) be a convex function. Let \(\varphi\) be a \(\mu\)-integrable function on \(X\) taking values in \(C\) such that \(f \circ \varphi\) is also \(\mu\)-integrable on \(X\). Then
\[
f \left( \int_X \varphi d\mu \right) \leq C_{ig} = C_{ig} \left( f, \varphi, \mu, \Lambda, \pi, \mathcal{S}_I \right) \tag{3.21}
\]
\[
:= \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu(A_{\pi_j(i)}) \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} \int_{A_{\pi_j(i)}} \varphi d\mu}{\sum_{j \in J} \lambda_{i,j} \mu(A_{\pi_j(i)})} \right) \leq \sum_{i \in I} \mu(A_i) f \left( \frac{1}{\mu(A_i)} \int_{A_i} \varphi d\mu \right) \leq \int_X f \circ \varphi d\mu. \tag{3.22}
\]

(b) Let \(C \subset \mathbb{R}^d\) be a closed convex set, and \(f : C \to \mathbb{R}\) be a continuous convex function. If \(\varphi\) is a \(\mu\)-integrable function on \(X\) taking values in \(C\) such that \(f \circ \varphi\) is also \(\mu\)-integrable on \(X\), then inequalities (3.21–3.22) are satisfied too.

**Proof.** Let the measure \(\mu_i \ (i \in I)\) be defined on \(\mathcal{A}\) by
\[
\mu_i (A) := \mu(A \cap A_i), \quad A \in \mathcal{A},
\]
and then apply Theorem 3.5. The proof is complete. \[\square\]
Remark 3.7. Assume \((H_1 - H_2)\) and \((H_i_1 - H_i_2)\). Similarly to Remark 3.4 we can obtain that Theorem 3.5 is an essential generalization of the main result in [8] (see Theorem 8 in [8]).

4. Applications

We begin with some inequalities related to information theory.

The following notion was introduced by Csiszár in [3] and [4].

Definition 4.1. Let \( f : ]0, \infty[ \to ]0, \infty[ \) be a convex function, and let \( r := (r_1, \ldots, r_n) \in ]0, \infty[^n \) and \( q := (q_1, \ldots, q_n) \in ]0, \infty[^n \). The \( f \)-divergence functional is

\[
I_f(r, q) := \sum_{i=1}^{n} q_i f \left( \frac{r_i}{q_i} \right).
\]

Based on this concept, we now introduce a new functional which extends the \( f \)-divergence functional and its generalizations in [9] and [7].

Definition 4.2. Let the index set \( I \) denote either \( \{1, \ldots, n\} \) for some \( n \geq 1 \) or \( \mathbb{N}_+ \). Let \( C \) be a closed convex subset of a real Banach space \((V, \| \cdot \|)\), and \( f : C \to \mathbb{R} \) be a convex function. If \( w := (w_i)_{i \in I} \) is a sequence in \( V \) and \( q := (q_i)_{i \in I} \) is a sequence in \( ]0, \infty[ \) such that

\[
\frac{w_i}{q_i} \in C, \quad i \in I, \tag{4.1}
\]

and the series

\[
\sum_{i \in I} q_i, \quad \sum_{i \in I} q_i f \left( \frac{w_i}{q_i} \right), \quad \sum_{i \in I} w_i \tag{4.2}
\]

are absolutely convergent, then define

\[
I_f(w, q) := \sum_{i \in I} q_i f \left( \frac{w_i}{q_i} \right). \tag{4.3}
\]

Remark 4.3. (a) If \( I = \{1, \ldots, n\} \) for some \( n \geq 1 \), then the previous notion works even if \( V \) is a real vector space (see Definition 3.2 in [7]). In this case only the condition (4.1) is essential, the series (4.2) are finite sums.

(b) Divergences between probability distributions have been introduced to measure the difference between them. The classical divergences between discrete probability distributions, for example the \( f \)-divergence (especially, Kullback–Leibler divergence, Hellinger distance and total variation distance), Rényi divergence, Jensen–Shannon divergence, etc. (see [11] and [18]), are defined for distributions taking on a finite number of values. The functional (4.3) can be applied to distributions with countably infinite outcomes.
Proposition 4.4. Assume \((H_1-H_3)\) and \((H_{d1}-H_{d2})\). If \(w := (w_i)_{i \in I}\) is a sequence in \(V\) and \(q := (q_i)_{i \in I}\) is a sequence in \([0, \infty[\) such that (4.1) holds and the series (4.2) are absolutely convergent, then
\[
I_f(w, q) = \sum_{i \in I} q_i f \left( \frac{w_i}{q_i} \right)
\geq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)} \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} w_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)}} \right) \geq \left( \sum_{i \in I} q_i \right) \cdot f \left( \sum_{i \in I} \frac{w_i}{\sum_{i \in I} q_i} \right).
\]

Proof. Taking
\[
p_i := \sum_{i \in I} q_i, \quad v_i := \frac{w_i}{q_i}, \quad i \in I
\]
in Theorem 3.1 we obtain
\[
\sum_{i \in I} q_i f \left( \frac{w_i}{q_i} \right) = \left( \sum_{i \in I} q_i \right) \cdot \sum_{i \in I} \frac{q_i}{\sum_{i \in I} q_i} f \left( \frac{w_i}{q_i} \right)
\geq \left( \sum_{i \in I} q_i \right) \cdot \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)} \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} w_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)}} \right)
= \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)} \right) f \left( \frac{\sum_{j \in J} \lambda_{i,j} w_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} q_{\pi_j(i)}} \right) \geq \left( \sum_{i \in I} q_i \right) \cdot f \left( \sum_{i \in I} \frac{w_i}{\sum_{i \in I} q_i} \right).
\]
The proof is complete. \(\Box\)

We illustrate the previous result by an example.

Example. We consider two Poisson distributions with parameters \(\mu, \nu > 0\) (see Remark 4.3 (b)), that is
\[
q_i := \frac{\mu^i}{i!} e^{-\mu}, \quad w_i := \frac{\nu^i}{i!} e^{-\nu}, \quad i \in \mathbb{N},
\]
and let \(f_\alpha : [0, \infty[ \to \mathbb{R}, f_\alpha(t) = t^\alpha\), where \(\alpha \geq 1\). Then \(I := \mathbb{N}, f_\alpha\) is convex for all \(\alpha \geq 1\), and
\[
I_{f_\alpha}(\mu, \nu) = \sum_{i \in I} q_i f_\alpha \left( \frac{w_i}{q_i} \right) = \exp \left( -\mu (1 - \alpha) - \nu \alpha + \mu^{1-\alpha} \nu^\alpha \right).
\]
Let the index set \( J \) denote either \( \{0, \ldots, k\} \) for some \( k \geq 0 \) or \( \mathbb{N} \), and assume \((H_2-H_3)\) hold. We can see that the conditions of Proposition 4.4 are satisfied and therefore by applying it we obtain

\[
I_f(\mu, \nu) = \exp \left( -\mu (1 - \alpha) - \nu \alpha + \mu^{1-\alpha} \nu^\alpha \right)
\]

\[
\geq e^{-\mu(1-\alpha) - \nu \alpha} \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \frac{\mu^{\pi_j(i)}}{\pi_j(i)!} \right)^{1-\alpha} \left( \sum_{j \in J} \lambda_{i,j} \frac{\nu^{\pi_j(i)}}{\pi_j(i)!} \right)^\alpha \geq 1. \tag{4.5}
\]

If \( f_\alpha : [0, \infty) \to \mathbb{R}, f_\alpha(t) = t^\alpha \), where \( 0 < \alpha < 1 \), then \(-f_\alpha\) is convex for all \( 0 < \alpha < 1 \), and hence the reverse inequalities hold in (4.5).

Assume \( 0 < \alpha < 1 \). By taking the logarithm of each side of (4.5) with reverse inequalities, it follows that

\[
\mu^{1-\alpha} \nu^\alpha \leq \ln \left( \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \frac{\mu^{\pi_j(i)}}{\pi_j(i)!} \right)^{1-\alpha} \left( \sum_{j \in J} \lambda_{i,j} \frac{\nu^{\pi_j(i)}}{\pi_j(i)!} \right)^\alpha \right) \leq \mu (1 - \alpha) + \nu \alpha,
\]

which is an interesting refinement of the weighted geometric mean - arithmetic mean inequality.

Similarly, the reverse inequalities hold in (4.6) if \( \alpha \geq 1 \).

The next result is based on the well known fact that norms on real vector spaces are convex functions.

**Proposition 4.5.** Assume \((H_1-H_3,H_{d1})\), and assume \((V, \|\cdot\|)\) is a Banach space. If \((v_i)_{i \in I}\) is a sequence of vectors from \( V \) such that the series \( \sum_{i \in I} p_i v_i \) is absolutely convergent, and \( \alpha \geq 1 \), then

\[
\left\| \sum_{i \in I} p_i v_i \right\|^\alpha \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} p^{\pi_j(i)} \right)^\alpha \left\| \sum_{j \in J} \lambda_{i,j} p^{\pi_j(i)} v_{\pi_j(i)} \right\|^\alpha \leq \sum_{i \in I} p_i \|v_i\|^\alpha.
\]

**Proof.** Since the function \( f : V \to \mathbb{R} \) defined by \( f(v) = \|v\|^\alpha \) is convex, and the series \( \sum_{i \in I} p_i \|v_i\|^\alpha \) is convergent, the result follows immediately from Theorem 3.1. \( \square \)

The first application of refinements of the integral Jensen inequality concerns quasi-arithmetic means.

Let \( C \subset \mathbb{R} \) be an interval, and let \( q : C \to \mathbb{R} \) be a continuous and strictly monotone function. If \((X, \mathcal{A}, \mu)\) is a probability space, and \( \varphi : X \to C \) is a function such that \( q \circ \varphi \) is \( \mu \)-integrable on \( X \), then

\[
M_q(\varphi, \mu) := q^{-1} \left( \int_X q \circ \varphi \mu \right)
\]
is called the quasi-arithmetic mean (integral $q$-mean) of $\varphi$.

Now we introduce some new quasi-arithmetic means related to the formula $C_{ig}$.

**Definition 4.6.** Assume $(H_1-H_3)$ and $(H_{i_1}-H_{i_2})$. Let $C \subset \mathbb{R}$ be an interval, let $q, r : C \to \mathbb{R}$ be continuous and strictly monotone functions, and $\varphi : X \to C$ a $\mu$-integrable function for which $q \circ \varphi$ and $r \circ \varphi$ are also $\mu$-integrable functions. Then we define the following quasi-arithmetic mean of $\varphi$ with respect to $C_{ig}$:

$$M_{q,r}(\varphi, \mu, \Lambda, \pi, \mathcal{M}_I) := q^{-1} \left( \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X) \right) q \circ r^{-1} \left( \frac{\sum_{j \in J} \lambda_{i,j} \int_X r \circ \varphi d\mu_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X)} \right) \right).$$

In the next result the introduced means are compared.

**Proposition 4.7.** Assume the conditions in Definition 4.6 hold. If either $q \circ r^{-1}$ is convex and $q$ is strictly increasing, or $q \circ r^{-1}$ is concave and $q$ is strictly decreasing, then

$$M_r(\varphi, \mu) \leq M_{q,r}(\varphi, \mu, \Lambda, \pi, \mathcal{M}_I) \leq M_q(\varphi, \mu),$$

while if either $r \circ q^{-1}$ is convex and $r$ is strictly decreasing, or $r \circ q^{-1}$ is concave and $r$ is strictly increasing, then

$$M_q(\varphi, \mu) \leq M_{r,q}(\varphi, \mu, \Lambda, \pi, \mathcal{M}_I) \leq M_r(\varphi, \mu).$$

**Proof.** We consider only the case when $q \circ r^{-1}$ is convex and $q$ is strictly increasing. By applying Theorem 3.5 (a) with $f := q \circ r^{-1}$ and with $r \circ \varphi$ instead of $\varphi$, we obtain

$$q \circ r^{-1} \left( \int_X r \circ \varphi d\mu \right) \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X) \right) q \circ r^{-1} \left( \frac{\sum_{j \in J} \lambda_{i,j} \int_X r \circ \varphi d\mu_{\pi_j(i)}}{\sum_{j \in J} \lambda_{i,j} \mu_{\pi_j(i)}(X)} \right) \leq \int_X q \circ \varphi d\mu,$$

and this implies the result since $q^{-1}$ is strictly increasing.

The proof is complete. \hfill \Box

In the following result a refinement of Hölder’s inequality is given.

**Theorem 4.8.** Assume $(H_1-H_3)$. Let $(Y, \mathcal{B}, \nu)$ be a measure space, and let $g \in L^p(\nu)$ and $h \in L^q(\nu)$, where $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Suppose we are given
a sequence $\mathfrak{M}_I = (\nu_i)_{i \in I}$ of measures on $\mathcal{B}$ with $\nu_i(Y) > 0$ for all $i \in I$ and $\sum_{i \in I} \nu_i = \nu$. Then

$$\int_Y |gh| \, d\nu \leq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \int_Y |g|^p \, d\nu_{\pi_j(i)} \right)^{1/p} \left( \sum_{j \in J} \lambda_{i,j} \int_Y |h|^q \, d\nu_{\pi_j(i)} \right)^{1/q}$$

(4.7)

$$\leq \left( \int_Y |g|^p \, d\nu \right)^{1/p} \left( \int_Y |h|^q \, d\nu \right)^{1/q}.$$  

(4.8)

Proof. The set $\{ y \in Y \mid |g(y)| + |h(y)| \neq 0 \}$ is of $\sigma$-finite measure, and therefore we may suppose that $\nu$ is $\sigma$-finite.

If $\int_Y |h|^q \, d\nu = 0$, then the assertion is obvious, so it can also be supposed that $\int_Y |h|^q \, d\nu > 0$.

(a) We first consider the case $|h(y)| > 0$ for all $y \in Y$.

Define the measure $\mu_i$ on $\mathcal{B}$ having density $|h|^q / \int_Y |h|^q \, d\nu$ with respect to $\nu_i$ that is

$$\mu_i(B) := \frac{1}{\int_Y |h|^q \, d\nu} \cdot \int_B |h|^q \, d\nu_i, \quad B \in \mathcal{B}, \quad i \in I,$$

and let

$$\mu := \sum_{i \in I} \mu_i.$$

Then $\mu$ is a probability measure on $\mathcal{B}$, $\mu_i(Y) > 0$ ($i \in I$), and $g \in L^p(\nu)$ shows that the function

$$\varphi := |g|^p |h|^{-q}$$

is $\mu$-integrable. The function $f : [0, \infty] \to \mathbb{R}, \ f(t) = -t^{1/p}$ is strictly convex, and since $gh \in L^1(\nu)$, we have that

$$f \circ \varphi = |g| |h|^{-q/p}.$$
is also $\mu$-integrable. It can be seen that Theorem 3.5 (a) can be applied to the introduced functions $f$ and $\varphi$, and measures $\mu$ and $(\mu_i)_{i \in I}$, and therefore

$$\left( \int_Y \varphi \, d\mu \right)^{1/p} \geq \sum_{i \in I} \left( \sum_{j \in J} \lambda_{i,j} \mu \pi_j(i) (Y) \right)^{1 - 1/p} \left( \sum_{j \in J} \lambda_{i,j} \int_Y \varphi \, d\mu \pi_j(i) \right)^{1/p} \geq \int_Y \varphi^{1/p} \, d\mu. \tag{4.9}$$

By considering the relationship between $\mu$- and $\nu$- as well as $\mu_i$- and $\nu_i$-integrals ($i \in I$), (4.9) can be rewritten in the form

$$\left( \frac{1}{Y} \int |h|^q \, d\nu \int_Y |g|^p \, d\nu \right)^{1/p} \geq \sum_{i \in I} \left( \frac{1}{Y} \int |h|^q \, d\nu \int_Y |\lambda_{i,j}|^p \, d\nu \pi_j(i) \right)^{1/q} \left( \frac{1}{Y} \int |h|^q \, d\nu \int_Y |g|^p \, d\nu \pi_j(i) \right)^{1/p} \geq \frac{1}{Y} \int |h|^q \, d\nu \int_X |g| \, d\nu,$$

and this is equivalent to (4.7–4.8).

(b) Assume the measurable set $Z := \{ y \in Y \mid h(y) = 0 \}$ is not empty. Since $\nu$ is $\sigma$-finite, it is possible to choose a positive function $h_Z : Z \to [0, \infty[$ such that $h_Z \in L^q(\nu)$. Then the functions

$$h_{\varepsilon}(y) := \begin{cases} h(y), & y \in Y \setminus Z, \\ \varepsilon h_Z(y), & y \in Z \end{cases}, \quad \varepsilon > 0$$

belong to $L^q(\nu)$ too, and $h_{\varepsilon}(y) > 0$ for all $y \in Y$. Now apply part (a), and then let $\varepsilon$ tend to 0.

The proof is complete. \(\square\)

The final result is a refinement of Minkowski’s inequality.

**Theorem 4.9.** Assume $(H_1–H_3)$. Let $(Y, \mathcal{B}, \nu)$ be a measure space, and let $g$, $h \in L^p(\nu)$, where $p > 1$. Suppose we are given a sequence $\mathfrak{M}_I = (\nu_i)_{i \in I}$ of measures on $\mathcal{B}$ with $\nu_i(Y) > 0$ for all $i \in I$ and $\sum_{i \in I} \nu_i = \nu$. 
(a) If \( p \in ]-\infty, 0[ \cup [1, \infty[ \), then
\[
\left( \int_Y |g + h|^p \, d\nu \right)^{1/p} \leq \left( \int_Y (|g| + |h|)^p \, d\nu \right)^{1/p} \tag{4.10}
\]
\[
\leq \sum_{i \in I} \left( \lambda_{i,j} \int_Y |g|^p \, d\nu \right)^{1/p} + \left( \lambda_{i,j} \int_Y |h|^p \, d\nu \right)^{1/p} \tag{4.11}
\]
\[
\leq \left( \int_Y |g|^p \, d\nu \right)^{1/p} + \left( \int_Y |h|^p \, d\nu \right)^{1/p}. \tag{4.12}
\]

(b) If \( p \in ]0, 1[, \) then the reversed inequalities hold in \((4.10–4.12)\).

Proof. If \( \int_Y (|g| + |h|)^p \, d\nu = 0, \) then the inequalities are obvious, so it can be supposed that \( \int_Y (|g| + |h|)^p \, d\nu > 0. \)

It can also be supposed that \(|g(y)| + |h(y)| > 0 \) for all \( y \in Y. \) To justify this reduction, let

\[ \hat{Y} := \{ y \in Y \mid |g(y)| + |h(y)| > 0 \}. \]

Then \( \hat{Y} \) is measurable. Let \( \hat{\mathcal{B}} \) be the trace \( \sigma \)-algebra \( \hat{Y} \cap \mathcal{B}, \) and let \( \hat{\nu} \) be the restriction of \( \nu \) to \( \hat{\mathcal{B}}. \) It is easy to see that \((\hat{Y}, \hat{\mathcal{B}}, \hat{\nu})\) can be used instead of \((Y, \mathcal{B}, \nu).\)

(a) Define the measure \( \mu_i \) on \( \mathcal{B} \) having density \((|g| + |h|)^p / \int_Y (|g| + |h|)^p \, d\nu \)
with respect to \( \nu_i \) that is

\[ \mu_i(B) := \frac{1}{\int_Y (|g| + |h|)^p \, d\nu} \cdot \int_B (|g| + |h|)^p \, d\nu_i, \quad B \in \mathcal{B}, \quad i \in I, \]

and let

\[ \mu := \sum_{i \in I} \mu_i. \]

Then \( \mu \) is a probability measure on \( \mathcal{B} \) having density \((|g| + |h|)^p / \int_Y (|g| + |h|)^p \, d\nu \)
with respect to \( \nu, \mu_i(Y) > 0 \) \((i \in I), \) and since \( g, h \in L^p(\nu), \) the function

\[ \varphi := \left( \frac{|g|^p}{(|g| + |h|)^p}, \frac{|h|^p}{(|g| + |h|)^p} \right) \]

is \( \mu \)-integrable. It is not hard to check that the function

\[ f : [0, \infty[ \times [0, \infty[ \to \mathbb{R}, \quad f(u, v) = \left( u^{1/p} + v^{1/p} \right)^p \tag{4.13} \]

is strictly concave. Since
\[ f \circ \varphi = -1, \]
it is also \( \mu \)-integrable. It now follows that Theorem 3.5 (b) can be applied to the introduced functions \(-f \) and \( \varphi \), and measures \( \mu \) and \( (\mu_i)_{i \in I} \), and therefore
\[
\left( \left( \int_Y \frac{|g|^p}{(|g|+|h|)^p} \, d\mu \right)^{1/p} + \left( \int_Y \frac{|h|^p}{(|g|+|h|)^p} \, d\mu \right)^{1/p} \right)^p \geq 1.
\]
(4.14)

Concerning the relationship between \( \mu \)- and \( \nu \)- as well as \( \mu_i \)- and \( \nu_i \)-integrals \( (i \in I) \), we can obtain from (4.14–4.16) that
\[
\frac{1}{\int_Y (|g| + |h|)^p \, d\nu} \left( \left( \int_Y |g|^p \, d\nu \right)^{1/p} + \left( \int_Y |h|^p \, d\nu \right)^{1/p} \right)^p \geq 1.
\]
(4.15)

(b) We can prove it exactly as in (a) by using that the function \( f \) defined in (4.13) is strictly convex for all \( p \in ]0,1[ \).

The proof is complete. \( \square \)

Remark 4.10. There are a lot of papers dealing with refinements of Hölder’s and Minkowski’s inequalities (see, e.g., \([1,2,5,15]\) and \([17]\)). Our results are unique in some way due to the fact that the refinements contain different parameters: measures and matrices; and infinite sums can be used.

Acknowledgements

The research was supported by Hungarian National Foundations for Scientific Research Grant No. K120186, and Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015.
Funding Open access funding provided by University of Pannonia.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Abramovich, S., Pečarić, J., Varošanec, S.: Continuous sharpening of Hölder’s and Minkowski’s inequalities. Math. Inequal. Appl. 8, 179–190 (2005)
[2] Chao, C.Z., Cheung, W.S.: On Minkowski’s inequality and its applications. J. Inequal. Appl. 71, 5 (2011)
[3] Csiszár, I.: Information measures: a critical survey. In: Transactions 7th Prague Conference on Information Theory, Statistics Decision Function, Random Processes and 8th European Meeting of Statistics, vol. B, pp. 73–86, Academia Prague (1978)
[4] Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hungar. 2, 299–318 (1967)
[5] Horváth, G.: Refinements of Hölder’s Inequality. Acta Math. Hungar. 144, 110–118 (2014)
[6] Horváth, L., Khan, K.A., Pečarić, J.: Combinatorial Improvements of Jensen’s Inequality. Element, Zagreb (2014)
[7] Horváth, L.: New refinements of the discrete Jensen’s inequality generated by finite or infinite permutations. Aequat. Math. (2019). https://doi.org/10.1007/s00010-019-00696-z
[8] Horváth, L.: Refinements of the integral Jensen’s inequality generated by finite or infinite permutations. J. Inequal. Appl. 2021(12), 14 (2021)
[9] Horváth, L., Pečarić, D., Pečarić, J.: Estimations of $f$- and Rényi divergences by using a cyclic refinement of the Jensen’s inequality. Bull. Malays. Math. Sci. Soc. 42(3), 933–946 (2019)
[10] Butt, S.I., Horváth, L., Pečarić, D., Pečarić, J.: Cyclic Improvements of Jensen’s Inequalities. Cyclic Inequalities in Information Theory. Element, Zagreb (2020)
[11] Liese, F., Vajda, I.: Convex Statistical Distances (Teubner-Texte Zur Mathematik), vol. 95. Teubner, Leipzig (1987)
[12] Niculescu, C., Persson, L.E.: Convex Functions and Their Applications. A Contemporary Approach. Springer, Berlin (2006)
[13] Pavić, Z.: Refinements of Jensen’s inequality for infinite convex combinations. Turk. J. Inequal. 2(2), 44–53 (2018)
[14] Perlman, M.D.: Jensen’s inequality for a convex vector-valued function on an infinite-dimensional space. J. Multivar. Anal. 4, 52–65 (1974)
[15] Qiang, H., Zhicheng, H.: Generalizations of Hölder’s and some related inequalities. Comput. Math. Appl. 61, 392–396 (2011)
[16] Swartz, C.: Iterated series and the Hellinger–Toeplitz theorem. Publ. Mat. 36, 167–173 (1992)
[17] Tian, J.F., Hu, X.M.: Refinements of generalized Hölder’s inequality. J. Math. Inequal. 7, 701–710 (2013)
[18] Vajda, I.: Theory of Statistical Inference and Information. Kluwer, Dordrecht (1989)

László Horváth
Department of Mathematics
University of Pannonia
Egyetem u. 10
8200 Veszprém
Hungary
e-mail: lhorvath@almos.uni-pannon.hu

Received: December 3, 2020
Revised: May 30, 2021
Accepted: June 3, 2021