MONOGRAPHS IN INEQUALITIES 18

Cyclic Improvements of Jensen's Inequalities

Cyclic Inequalities in Information Theory

Saad Ihsan Butt, László Horváth, Đilda Pečarić and Josip Pečarić

Cyclic Improvements of Jensen's Inequalities

Cyclic Inequalities in Information Theory

Saad Ihsan Butt
COMSATS University Islamabad
Lahore Campus, Pakistan

László Horváth
Department of Mathematics
University of Pannonia
Veszprém, Hungary

Đilda Pečarić
Catholic University of Croatia
Zagreb, Croatia

Josip Pečarić
RUDN University
Moscow, Russia.

ELEMENT

Zagreb, 2020

MONOGRAPHS IN INEQUALITIES 18

Cyclic Improvements of Jensen's Inequalities

Cyclic Inequalities in Information Theory

Saad Ihsan Butt, László Horváth, Đilda Pečarić and Josip Pečarić

1st edition

Printed in Croatia.

A CIP catalogue record for this book is available from the National and University Library in Zagreb under 001060582.

ISBN 978-953-197-686-2

Preface

The most important inequality for convex functions is the Jensen's inequality. Other inequalities such as inequalities for means, the Hölder's and Minkowski's inequalities etc. can be obtained as particular cases of it, and it has many applications in different branches of mathematics. There are countless papers dealing with generalizations, refinements, and converse results of Jensen's inequality.

To give refinements of Jensen's inequality is an extensively investigated theme with numerous methods, results and applications. This book is mainly devoted to cyclic refinements (cyclic permutations are used to define the refining terms), and their applications in information theory. It contains the most recent research results of this promising topic.

The first chapter has a preparatory character. In the second chapter the basic cyclic refinements for the discrete and integral Jensen's inequalities are given. Among the many topics where Jensen's inequality finds application, mention should be made of information theory. Jensen's inequality plays a crucial role to obtain inequalities for divergences between probability distributions, which have been introduced to measure the difference between them. A lot of different type of divergences exist, for example the f-divergence (especially, Kullback-Leibler divergence, Hellinger distance and total variation distance), Rényi divergence, Jensen-Shannon divergence, etc. These important notions and the Zipf-Mandelbrot law (a special discrete probability distribution) are introduced in chapter three. The power of results in chapter two is also demonstrated in chapter three by obtaining refinements of inequalities for divergences. In chapter four cyclic refinements of Beck's inequality are given. This leads to some new refinements of the classical Hölder's and Minkowski's inequalities. Chapter five deals with cyclic refinements of operator Jensen's inequalities for convex and operator convex functions. In the next six chapters (sixth to eleventh) are devoted to extensions of cyclic refinements of Jensen's inequality via Taylor's formula, Fink's identity and Montgomery's identity, and by Lidstone interpolating polynomial, Abel-Gontscharoff interpolating polynomial and Hermite intepolating polynomial, respectively. The applicability of all the obtained results is demonstrated by means of information theory. In the last chapter, twelve, Levinson's type generalization of cyclic refinements of Jensen's inequality is given with applications.

Authors

Contents

Chapter 1

Introduction and Preliminaries

The notion of convexity plays an important role in different branches of mathematics.

Definition 1.1 *Let V be a real vector space.*

(a) A subset C of V is called convex, if for any two points $v_1, v_2 \in C$, the line segment between them also lies in C, that is $\lambda v_1 + (1-\lambda) v_2 \in C$ for all $\lambda \in [0,1]$.

(b) A function $f : C\,(\subset V) \to \mathbb{R}$ is called convex, if its domain C is a convex set, and for any two points $v_1, v_2 \in C$, and all $\lambda \in [0,1]$, we have that

$$f(\lambda v_1 + (1-\lambda) v_2) \le \lambda f(v_1) + (1-\lambda) f(v_2).$$

The most important inequality concerning convex functions is the Jensen's inequality, named after the Danish mathematician Johan Jensen. It was proven by Jensen in [49]. We emphasize the following two variants of Jensen's inequality:

Theorem A. *(discrete Jensen's inequality, see [36]) 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 $\sum\limits_{i=1}^{n} p_i = 1$, and $v_1, \ldots, v_n \in C$, then*

$$f\left(\sum_{i=1}^{n} p_i v_i\right) \le \sum_{i=1}^{n} p_i f(v_i) \tag{1.1}$$

holds. Particularly, we have

$$f\left(\frac{1}{n}\sum_{i=1}^{n} v_i\right) \le \frac{1}{n}\sum_{i=1}^{n} f(v_i). \tag{1.2}$$

Theorem B. *(integral Jensen's inequality, see [36]) Let g be an integrable function on a probability space* $(X, \mathscr{A}, \mu)$ *taking values in an interval* $I \subset \mathbb{R}$. *Then* $\int_X g d\mu$ *lies in I. If f is a convex function on I such that* $f \circ g$ *is integrable, then*

$$f\left(\int_X g d\mu\right) \leq \int_X f \circ g d\mu.$$

Various attempts have been made by many authors to refine either the discrete or the integral Jensen's inequality (see the book [36] and the references therein). A multitude of applications underscores the importance of refinements of different Jensen's inequalities.

The following result which provide the starting point for our discussion is from Brnetić at al. [12].

Theorem 1.1 *Suppose I is a real interval. If* $f : I \to \mathbb{R}$ *is a convex function, then for all* $t \in [0,1]$ *we have*

$$f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \leq \frac{1}{n}\sum_{i=1}^{n} f\left((1-t)x_i + t x_{i+1}\right) \leq \frac{\sum_{i=1}^{n} f(x_i)}{n},$$

where $x_i \in I$ $(1 \leq i \leq n)$ *and* $x_{n+1} = x_1$.

Recently, a lot of papers have been appeared dealing with generalizations of the previous theorem (see e.g. [13, 37]). The whole group of such results is now often known by the collective title "cyclic refinements". They find applications mainly in the theory of means and in information theory.

The title of this book indicates clearly the content of it. A synthesis of recent progress in the topic of cyclic refinements of different types of Jensen's inequalities is presented with the emphasis on their applications in information theory.

Let $2 \leq k \leq n$, and let $i \in \{1, \ldots, n\}$ and $j \in \{0, \ldots, k-1\}$. In further parts of this book $i + j$ always means $i + j - n$ in case of $i + j > n$.

While writing this book, S. I. Butt was supported by Higher Education Commission Pakistan under NRPU project 5327 and 7906. L. Horváth was supported by Hungarian National Foundations for Scientific Research Grant No. K101217 and Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015. J. Pečarić is supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No. 02.a03.21.0008)

Chapter 2

Cyclic Improvement of Jensen's Inequality

In this chapter we give cyclic refinements of Jensen's inequality and their applications.

2.1 A refinement of Jensen's inequality

We start with the special following Jensen's inequality

$$f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \leq \frac{\sum_{i=1}^{n} f(x_i)}{n}.$$

Throughout this section we are going to use some of the following hypotheses:

($\mathscr{H}_1$) Let $I \subset \mathbb{R}$ be an interval, $\mathbf{x} := (x_1, \ldots, x_n) \in I^n$, and $\lambda := (\lambda_1, \ldots, \lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for some k, $2 \leq k \leq n$.

($\mathscr{H}_2$) Let $f : I \to \mathbb{R}$ be a convex function.

($\mathscr{H}_3$) Let h, $g : I \to \mathbb{R}$ be continuous and strictly monotone functions.

Theorem 2.1 *Let* ($\mathscr{H}_1$), ($\mathscr{H}_2$) *be fulfilled. Then*

$$f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \leq \frac{1}{n} \sum_{i=1}^{n} f\left(\sum_{j=0}^{k-1} \lambda_{j+1} x_{i+j}\right) \leq \frac{\sum_{i=1}^{n} f(x_i)}{n}. \tag{2.1}$$

Proof. First, since f is convex, by Jensen's inequality we have

$$\sum_{i=1}^{n} f\left(\sum_{j=0}^{k-1} \lambda_{j+1} x_{i+j}\right) \leq \sum_{i=1}^{n} \sum_{j=0}^{k-1} \lambda_{j+1} f(x_{i+j})$$

$$= \sum_{i=1}^{n} f(x_i) \sum_{j=1}^{k} \lambda_j = \sum_{i=1}^{n} f(x_i).$$

On the other hand, since f is convex, by Jensen's inequality, we have

$$\frac{1}{n} \sum_{i=1}^{n} f\left(\sum_{j=0}^{k-1} \lambda_{j+1} x_{i+j}\right) \geq f\left(\frac{\sum_{i=1}^{n} \sum_{j=0}^{k-1} \lambda_{j+1} x_{i+j}}{n}\right)$$

$$= f\left(\frac{\sum_{i=1}^{n} x_i \sum_{j=1}^{k} \lambda_j}{n}\right) = f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right).$$

□

Theorem 2.1 is a generalization of Theorem 4 in [12].

2.1.1 Cyclic mixed symmetric means

Assume $(\mathscr{H}_1)$ for the positive n-tuple $\mathbf{x}$. We define the power means of order $r \in \mathbb{R}$ as follows:

$$M_r(x_i,...,x_{i+k-1};\lambda_1,...,\lambda_k) = \begin{cases} \left(\sum\limits_{j=0}^{k-1} \lambda_{j+1} x_{i+j}^r\right)^{\frac{1}{r}}; & r \neq 0, \\ \prod\limits_{j=0}^{k-1} x_{i+j}^{\lambda_{j+1}}; & r = 0, \end{cases}$$

and cyclic mixed symmetric means corresponding to (2.1) are

$$M_{r,s}(\mathbf{x},\lambda) := \begin{cases} \left(\frac{1}{n} \sum\limits_{i=1}^{n} M_r^s(x_i,...,x_{i+k-1};\lambda_1,...,\lambda_k)\right)^{\frac{1}{s}}; & s \neq 0, \\ \left(\prod\limits_{i=1}^{n} M_r(x_i,...,x_{i+k-1};\lambda_1,...,\lambda_k)\right)^{\frac{1}{n}}; & s = 0. \end{cases} \tag{2.2}$$

The standard power means of order $r \in \mathbb{R}$ for the positive n-tuple $\mathbf{x}$, are

$$M_r(x_1,...,x_n) = M_r(\mathbf{x}) := \begin{cases} \left(\frac{1}{n} \sum\limits_{i=1}^{n} x_i^r\right)^{\frac{1}{r}}; & r \neq 0, \\ \left(\prod\limits_{i=1}^{n} x_i\right)^{\frac{1}{n}}; & r = 0. \end{cases}$$

The bounds for cyclic mixed symmetric means are power means, as given in the following result.

Corollary 2.1 *Assume $(\mathscr{H}_1)$ for the positive n-tuple* $\mathbf{x}$. *Let $r,s \in \mathbb{R}$ such that $r \leq s$. Then*

$$M_r(\mathbf{x}) \leq M_{s,r}(\mathbf{x},\lambda) \leq M_s(\mathbf{x}). \tag{2.3}$$

Proof. Assume r, $s \neq 0$. To obtain (2.3), we apply Theorem 2.1, either for the function $f(x) = x^{\frac{s}{r}}$ $(x > 0)$ and the n-tuples $(x_1^r,\ldots,x_n^r)$ in (2.1) and then raising the power $\frac{1}{s}$, or $f(x) = x^{\frac{r}{s}}$ $(x > 0)$ and $(x_1^s,\ldots,x_n^s)$ and raising the power $\frac{1}{r}$.
When $r = 0$ or $s = 0$, we get the required results by taking limit. □

Special cases of Corollary 2.1 can be found in [11] (see Theorem 4 with Corollaries 4.1–4.4). Namely, the result of this theorem is an inequality (2.3) for $r = 0$, $s = 1$, $n = 3$ and $k = 3$.
Assume $(\mathscr{H}_1)$ and $(\mathscr{H}_3)$. Then we define the generalized means with respect to (2.1) as follows:

$$M_{g,h}(\mathbf{x},\lambda) := g^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}(g\circ h^{-1})(\sum_{j=0}^{k-1}\lambda_{j+1}h(x_{i+j}))\right).$$

Let $q : I \to \mathbb{R}$ be a continuous and strictly monotone function then the cyclic quasi-arithmetic means are given by

$$M_q(\mathbf{x}) := q^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}q(x_i)\right).$$

The relation among the generalized means and cyclic quasi-arithmetic means is given in the next corollary.

Corollary 2.2 *Assume $(\mathscr{H}_1)$ and $(\mathscr{H}_3)$. Then*

$$M_h(\mathbf{x}) \leq M_{g,h}(\mathbf{x},\lambda) \leq M_g(\mathbf{x}) \tag{2.4}$$

if either $g\circ h^{-1}$ is convex and g is strictly increasing or $g\circ h^{-1}$ is concave and g is strictly decreasing.

Proof. First, we can apply Theorem 2.1 to the function $g\circ h^{-1}$ and the n-tuples $(h(x_1),\ldots,h(x_n))$, then we can apply g^{-1} to the inequality coming from (2.1). This gives (2.4). □

For instance, if we put $g(x) = x$ and $h(x) = \ln x$ in Corollary 2.2 we obtain

$$M_0(x_1,...,x_n) \leq \frac{1}{n}\sum_{i=1}^{n}M_0(x_i,...,x_{i+k-1};\lambda_1,...,\lambda_k) \leq M_1(x_1,...,x_n).$$

which is a special case of Corollary 2.1 as well.

Remark 2.1 Under the conditions $(\mathscr{H}_1)$, we define

$$\Upsilon_1(f) = \Upsilon_1(\mathbf{x},\lambda,f) := \frac{1}{n}\sum_{i=1}^{n}f(x_i) - \frac{1}{n}\sum_{i=1}^{n}f(\sum_{j=0}^{k-1}\lambda_{j+1}x_{i+j}),$$

$$\Upsilon_2(f) = \Upsilon_2(\mathbf{x}, \lambda, f) := \frac{1}{n}\sum_{i=1}^{n} f(\sum_{j=0}^{k-1} \lambda_{j+1}x_{i+j}) - f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right),$$

where $f : I \to \mathbb{R}$ is a function and $2 \le k \le n$. The functionals $f \to \Upsilon_i(f)$ are linear, $i = 1,2$, and Theorem 2.1 implies that

$$\Upsilon_i(f) \ge 0, \quad i = 1,2$$

if $f : I \to \mathbb{R}$ is a convex function.

2.1.2 *m*-Exponential convexity

For log-convexity, exponential convexity and m-exponential convexity of the functionals obtained from the interpolations of the discrete Jensen's inequality, we refer [36] and references therein.

We apply the method given in [83], to prove the m-exponential convexity and exponential convexity of the functionals $f \to \Upsilon_i(f)$ for $i = 1,2$, together with the Lagrange type and Cauchy type mean value theorems.

Definition 2.1 *(see [83]) A function $g : I \to \mathbb{R}$ is called* m-exponentially convex *in the Jensen sense if*

$$\sum_{i,j=1}^{m} a_i a_j g\left(\frac{x_i + x_j}{2}\right) \ge 0$$

holds for every $a_i \in \mathbb{R}$ and every $x_i \in I$, $i = 1,2,\ldots,m$.

A function $g : I \to \mathbb{R}$ is m-exponentially convex *if it is* m-exponentially convex *in the Jensen sense and continuous on I.*

Note that 1-exponentially convex functions in the Jensen sense are in fact the nonnegative functions. Also, m-exponentially convex functions in the Jensen sense are n-exponentially convex in the Jensen sense for every $n \in \mathbb{N}$, $n \le m$.

Proposition 2.1 *If $g : I \to \mathbb{R}$ is an m-exponentially convex function, then for every $x_i \in I$, $i = 1,2,\ldots,m$ and for all $n \in \mathbb{N}$, $n \le m$ the matrix $\left[g\left(\frac{x_i+x_j}{2}\right)\right]_{i,j=1}^{n}$ is a positive semi-definite matrix. Particularly,*

$$\det\left[g\left(\frac{x_i + x_j}{2}\right)\right]_{i,j=1}^{n} \ge 0$$

for all $n \in \mathbb{N}$, $n = 1,2,\ldots,m$.

Definition 2.2 *A function $g : I \to \mathbb{R}$ is exponentially convex in the Jensen sense, if it is m-exponentially convex in the Jensen sense for all $m \in \mathbb{N}$.*

A function $g : I \to \mathbb{R}$ is exponentially convex if it is exponentially convex in the Jensen sense and continuous.

Remark 2.2 It is easy to see that a positive function $g : I \to \mathbb{R}$ is log-convex in the Jensen sense if and only if it is 2-exponentially convex in the Jensen sense, that is

$$a_1^2 g(x) + 2a_1a_2 g\left(\frac{x+y}{2}\right) + a_2^2 g(y) \geq 0$$

holds for every $a_1, a_2 \in \mathbb{R}$ and $x, y \in I$.

Similarly, if g is 2-exponentially convex, then g is log-convex. On the other hand, if g is log-convex and continuous, then g is 2-exponentially convex.

In sequel, we need the well known notion of "Divided difference".

Definition 2.3 *The second order divided difference of a function $g : I \to \mathbb{R}$ at mutually different points $y_0, y_1, y_2 \in I$ is defined recursively by*

$$[y_i; g] = g(y_i), \quad i = 0,1,2$$

$$[y_i, y_{i+1}; g] = \frac{g(y_{i+1}) - g(y_i)}{y_{i+1} - y_i}, \quad i = 0,1$$

$$[y_0, y_1, y_2; g] = \frac{[y_1, y_2; g] - [y_0, y_1; g]}{y_2 - y_0}. \tag{2.5}$$

Remark 2.3 The value $[y_0, y_1, y_2; g]$ is independent of the order of the points y_0, y_1, and y_2. By taking limits this definition may be extended to include the cases in which any two or all three points coincide as follows: for all y_0, y_1, $y_2 \in I$ such that $y_2 \neq y_0$

$$\lim_{y_1 \to y_0} [y_0, y_1, y_2; g] = [y_0, y_0, y_2; g] = \frac{g(y_2) - g(y_0) - g'(y_0)(y_2 - y_0)}{(y_2 - y_0)^2}$$

provided that g' exists, and furthermore, taking the limits $y_i \to y_0$, $i = 1,2$ in (2.5), we get

$$[y_0, y_0, y_0; g] = \lim_{y_i \to y_0} [y_0, y_1, y_2; g] = \frac{g''(y_0)}{2} \text{ for } i = 1,2$$

provided that g'' exist on I.

Now, we give the m-exponential convexity for the linear functionals $\Upsilon_i(f)$ $(i = 1,2)$.

Theorem 2.2 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t \mid t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \to [y_0, y_1, y_2; \phi_t]$ $(t \in J)$ is m-exponentially convex in the Jensen sense on I for every three mutually different points $y_0, y_1, y_2 \in I$. Let $\Upsilon_i(f)$ $(i = 1,2)$ be the linear functionals constructed in Remark 2.1. Then $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is an m-exponentially convex function in the Jensen sense on I for each $i = 1,2$. If the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is continuous for $i = 1,2$, then it is m-exponentially convex on I for $i = 1,2$.*

Proof. Fix $i = 1, 2$.

Let $t_k, t_l \in J$, $t_{kl} := \frac{t_k + t_l}{2}$ and $b_k, b_l \in \mathbb{R}$ for $k, l = 1, 2, ..., n$, and define the function ω on I by

$$\omega := \sum_{k,l=1}^{n} b_k b_l \phi_{t_{kl}}.$$

Since the function $t \to [y_0, y_1, y_2; \phi_t]$ $(t \in J)$ is m-exponentially convex in the Jensen sense, we have

$$[y_0, y_1, y_2; \omega] = \sum_{k,l=1}^{n} b_k b_l [y_0, y_1, y_2; \phi_{t_{kl}}] \geq 0.$$

Hence ω is a convex function on I. Therefore we have $\Upsilon_i(\omega) \geq 0$, which yields by the linearity of Υ_i, that

$$\sum_{k,l=1}^{n} b_k b_l \Upsilon_i(\phi_{t_{kl}}) \geq 0.$$

We conclude that the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is an m-exponentially convex function in the Jensen sense on I.
If the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is continuous on I, then it is m-exponentially convex on I by definition. □

As a consequence of the above theorem we can give the following corollaries.

Corollary 2.3 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t \mid t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \to [y_0, y_1, y_2; \phi_t]$ $(t \in J)$ is exponentially convex in the Jensen sense on I for every three mutually different points $y_0, y_1, y_2 \in I$. Let $\Upsilon_i(f)$ $(i = 1, 2)$ be the linear functionals constructed in Remark 2.1. Then $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is an exponentially convex function in the Jensen sense on I for $i = 1, 2$. If the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is continuous, then it is exponentially convex on I for $i = 1, 2$.*

Corollary 2.4 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t : t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \to [y_0, y_1, y_2; \phi_t]$ $(t \in J)$ is 2-exponentially convex in the Jensen sense on I for every three mutually different points $y_0, y_1, y_2 \in I$. Let $\Upsilon_i(f)$ $(i = 1, 2)$ be the linear functionals constructed in Remark 2.1. Then the following two statements hold for $i = 1, 2$:*

(i) If the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is positive and continuous, then it is 2-exponentially convex on I, and thus log-convex.

(ii) If the function $t \to \Upsilon_i(\phi_t)$ $(t \in J)$ is positive and differentiable, then for every $s, t, u, v \in J$, such that $s \leq u$ and $t \leq v$, we have

$$\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda) \leq \mathfrak{u}_{u,v}(\Upsilon_i, \Lambda) \tag{2.6}$$

where

$$\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda) := \begin{cases} \left(\frac{\Upsilon_i(\phi_s)}{\Upsilon_i(\phi_t)} \right)^{\frac{1}{s-t}}, & s \neq t, \\ \exp\left(\frac{\frac{d}{ds}\Upsilon_i(\phi_s)}{\Upsilon_i(\phi_s)} \right), & s = t \end{cases} \tag{2.7}$$

for $\phi_s, \phi_t \in \Lambda$.

Proof. Fix $i = 1,2$.

(i) The proof follows by Remark 2.2 and Theorem 2.2.

(ii) From the definition of a convex function ψ on I, we have the following inequality (see [82, page 2])

$$\frac{\psi(s) - \psi(t)}{s - t} \leq \frac{\psi(u) - \psi(v)}{u - v}, \tag{2.8}$$

$\forall s,t,u,v \in J$ such that $s \leq u$, $t \leq v$, $s \neq t$, $u \neq v$.
By (i), $s \to \Upsilon_i(\phi_s)$, $s \in J$ is log-convex, and hence (2.8) shows with $\psi(s) = \log \Upsilon_i(\phi_s)$, $s \in J$ that

$$\frac{\log \Upsilon_i(\phi_s) - \log \Upsilon_i(\phi_t)}{s-t} \leq \frac{\log \Upsilon_i(\phi_u) - \log \Upsilon_i(\phi_v)}{u-v} \tag{2.9}$$

for $s \leq u$, $t \leq v$, $s \neq t$, $u \neq v$, which is equivalent to (2.6). For $s = t$ or $u = v$ (2.6) follows from (2.9) by taking limit.

□

Remark 2.4 Note that the results from Theorem 2.2, Corollary 2.3, Corollary 2.4 are valid when two of the points $y_0, y_1, y_2 \in I$ coincide, say $y_1 = y_0$, for a family of differentiable functions ϕ_t such that the function $t \to [y_0, y_1, y_2; \phi_t]$ is m-exponentially convex in the Jensen sense (exponentially convex in the Jensen sense, log-convex in the Jensen sense), and moreover, they are are also valid when all three points coincide for a family of twice differentiable functions with the same property. The proofs can be obtained by recalling Remark 2.3 and suitable characterization of convexity.

The following result given in [35] is related to the first condition of Theorem 2.2.

Theorem 2.3 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t \mid t \in J\}$ is a family of twice differentiable functions defined on an interval $I \subset \mathbb{R}$ such that the function $t \mapsto \phi_t''(x)$ $(t \in J)$ is exponentially convex for every fixed $x \in I$. Then the function $t \mapsto [y_0, y_1, y_2; \phi_t]$ $(t \in J)$ is exponentially convex in the Jensen sense for any three points y_0, y_1, $y_2 \in I$.*

Remark 2.5 It comes from either the conditions of Theorem 2.3 or the proof of this theorem that the functions ϕ_t, $t \in J$ are convex.

2.1.3 Mean value theorems

Now we formulate mean value theorems of Lagrange and Cauchy type for the linear functionals $\Upsilon_i(f)$ $(i = 1,2)$ defined in Remark 2.1.

Theorem 2.4 *Let $\Upsilon_i(f)$ $(i = 1,2)$ be the linear functionals constructed in Remark 2.1 and $g \in C^2[a,b]$. Then there exists $\xi \in [a,b]$ such that*

$$\Upsilon_i(g) = \frac{1}{2} g''(\xi) \Upsilon_i(x^2); \quad i = 1,2.$$

Proof. Fix $i = 1,2$.

Since $g \in C^2[a,b]$, there exist the real numbers $m = \min_{x\in[a,b]} g''(x)$ and $M = \max_{x\in[a,b]} g''(x)$. It is easy to show that the functions ϕ_1 and ϕ_2 defined on $[a,b]$ by

$$\phi_1(x) = \frac{M}{2}x^2 - g(x),$$

and

$$\phi_2(x) = g(x) - \frac{m}{2}x^2,$$

are convex.
By applying the functional Υ_i to the functions ϕ_1 and ϕ_2, we have the properties of Υ_i that

$$\Upsilon_i\left(\frac{M}{2}x^2 - g(x)\right) \geq 0,$$

$$\Rightarrow \Upsilon_i(g) \leq \frac{M}{2}\Upsilon_i\left(x^2\right), \tag{2.10}$$

and

$$\Upsilon_i\left(g(x) - \frac{m}{2}x^2\right) \geqslant 0$$

$$\Rightarrow \frac{m}{2}\Upsilon_i\left(x^2\right) \leq \Upsilon_i(g). \tag{2.11}$$

From (2.10) and from (2.11), we get

$$\frac{m}{2}\Upsilon_i\left(x^2\right) \leq \Upsilon_i(g) \leq \frac{M}{2}\Upsilon_i\left(x^2\right).$$

If $\Upsilon_i\left(x^2\right) = 0$, then nothing to prove. If $\Upsilon_i\left(x^2\right) \neq 0$, then

$$m \leq \frac{2\Upsilon_i(g)}{\Upsilon_i(x^2)} \leq M.$$

Hence we have

$$\Upsilon_i(g) = \frac{1}{2}g''(\xi)\Upsilon_i\left(x^2\right).$$

□

Theorem 2.5 *Let $\Upsilon_i(f)$ $(i = 1,2)$ be the linear functionals constructed in Remark 2.1 and $g,h \in C^2[a,b]$. Then there exists $\xi \in [a,b]$ such that*

$$\frac{\Upsilon_i(g)}{\Upsilon_i(h)} = \frac{g''(\xi)}{h''(\xi)}; \quad i = 1,2,$$

provided that $\Upsilon_i(h) \neq 0$ $(i = 1,2)$.

Proof. Fix $i = 1,2$.

Define $L \in C^2[a,b]$ by
$$L := c_1 g - c_2 h,$$
where
$$c_1 := \Upsilon_i(h)$$
and
$$c_2 := \Upsilon_i(g).$$
Now using Theorem 2.4 for the function L, we have
$$\left(c_1 \frac{g''(\xi)}{2} - c_2 \frac{h''(\xi)}{2}\right) \Upsilon_i\left(x^2\right) = 0. \tag{2.12}$$
Since $\Upsilon_i(h) \neq 0$, Theorem 2.4 implies that $\Upsilon_i\left(x^2\right) \neq 0$, and therefore (2.12) gives
$$\frac{\Upsilon_i(g)}{\Upsilon_i(h)} = \frac{g''(\xi)}{h''(\xi)}.$$
□

2.1.4 Applications to Cauchy means

In this section we apply the results of previous sections to generate new Cauchy means. We mention that the functionals $\Upsilon_i(f)$, $i = 1,2$ defined in Remark 2.1 under the assumption $(\mathscr{H}_1)$, are linear on the vector space of real functions defined on the interval $I \subset \mathbb{R}$, and $\Upsilon_i(f) \geq 0$ for every convex function on I.

Example 2.1 Let $I = \mathbb{R}$ and consider the class of convex functions
$$\Lambda_1 := \{\phi_t : \mathbb{R} \to [0,\infty[\mid t \in \mathbb{R}\},$$
where
$$\phi_t(x) := \begin{cases} \frac{1}{t^2} e^{tx}; & t \neq 0, \\ \frac{1}{2} x^2; & t = 0. \end{cases}$$
Then $t \mapsto \phi_t''(x)$ $(t \in \mathbb{R})$ is exponentially convex for every fixed $x \in \mathbb{R}$ (see [47]), thus by Theorem 2.3, the function $t \mapsto [y_0, y_1, y_2; \phi_t]$, $t \in \mathbb{R}$ is exponentially convex in the Jensen sense for every three mutually different points $y_0, y_1, y_2 \in \mathbb{R}$.

Now fix $i = 1,2$. By applying Corollary 2.3 with $\Lambda = \Lambda_1$, we get the exponential convexity of $t \mapsto \Upsilon_i(\phi_t)$ $(t \in \mathbb{R})$ in the Jensen sense. This mapping is also differentiable, therefore exponentially convex, and the expression in (2.7) has the form
$$\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda_1) = \begin{cases} \left(\frac{\Upsilon_i(\phi_s)}{\Upsilon_i(\phi_t)}\right)^{\frac{1}{s-t}}; & s \neq t, \\ \exp\left(\frac{\Upsilon_i(id\,\phi_s)}{\Upsilon_i(\phi_s)} - \frac{2}{s}\right); & s = t \neq 0, \\ \exp\left(\frac{\Upsilon_i(id\,\phi_0)}{3\Upsilon_i(\phi_0)}\right); & s = t = 0, \end{cases}$$
where "*id*" means the identity function on $\mathbb{R}$.

From (2.6) we have the monotonicity of the functions $\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda_1)$ in both parameters s and t.

Suppose $\Upsilon_i(\phi_t) > 0$ $(t \in \mathbb{R})$, $a := min\{x_1,...,x_n\}$, $b := max\{x_1,...,x_n\}$, and let

$$\mathfrak{M}_{s,t}(\Upsilon_i,\Lambda_1) := \log \mathfrak{u}_{s,t}(\Upsilon_i,\Lambda_1); \quad s,t \in \mathbb{R}.$$

Then from Theorem 2.5 we have

$$a \leq \mathfrak{M}_{s,t}(\Upsilon_i,\Lambda_1) \leq b,$$

and thus $\mathfrak{M}_{s,t}(\Upsilon_i,\Lambda_1)$ $(s,t \in \mathbb{R})$ are means. The monotonicity of these means is followed by (2.6).

Example 2.2 Let $I =]0,\infty[$ and consider the class of convex functions

$$\Lambda_2 = \{\psi_t :]0,\infty[\to \mathbb{R} \mid t \in \mathbb{R}\},$$

where

$$\psi_t(x) := \begin{cases} \frac{x^t}{t(t-1)};\ t \neq 0,1, \\ -\log x;\ t = 0, \\ x\log x;\ t = 1. \end{cases}$$

Then $t \mapsto \psi_t''(x) = x^{t-2} = e^{(t-2)\log x}$ $(t \in \mathbb{R})$ is exponentially convex for every fixed $x \in]0,\infty[$.

Now fix $1 \leq i \leq 4$. By similar arguments as given in Example 2.1 we get the exponential convexity of $t \mapsto \Upsilon_i(\psi_t)$ $(t \in \mathbb{R})$ in the Jensen sense. This mapping is differentiable too, therefore exponentially convex. It is easy to calculate that (2.7) can be written as

$$\mathfrak{u}_{s,t}(\mathbf{x},\mathbf{p},\Upsilon_i,\Lambda_2) = \begin{cases} \left(\frac{\Upsilon_i(\psi_s)}{\Upsilon_i(\psi_t)}\right)^{\frac{1}{s-t}};\ s \neq t, \\ \exp\left(\frac{1-2s}{s(s-1)} - \frac{\Upsilon_i(\psi_s\psi_0)}{\Upsilon_i(\psi_s)}\right);\ s = t \neq 0,1, \\ \exp\left(1 - \frac{\Upsilon_i(\psi_0^2)}{2\Upsilon^i(\psi_0)}\right);\ s = t = 0, \\ \exp\left(-1 - \frac{\Upsilon_i(\psi_0\psi_1)}{2\Upsilon_i(\psi_1)}\right);\ s = t = 1. \end{cases}$$

Suppose $\Upsilon_i(\psi_t) > 0$ $(t \in \mathbb{R})$, and let $a := min\{x_1,...,x_n\}$, $b := max\{x_1,...,x_n\}$. By Theorem 2.5, we can check that

$$a \leq \mathfrak{u}_{s,t}(\mathbf{x},\mathbf{p},\Upsilon_i,\Lambda_2) \leq b; \quad s,t \in \mathbb{R}. \tag{2.13}$$

The means $\mathfrak{u}_{s,t}(\mathbf{x},\mathbf{p},\Upsilon_i,\Lambda_2)$ $(s,t \in \mathbb{R})$ are continuous, symmetric and monotone in both parameters (by use of (2.6)).

Let $s,t,r \in \mathbb{R}$ such that $r \neq 0$. By the substitutions $s \to \frac{s}{r}$, $t \to \frac{t}{r}$, $(x_1,\ldots,x_n) \to (x_1^r,\ldots,x_n^r)$ in (2.13), we get

$$\bar{a} \leq \mathfrak{u}_{s/r,t/r}(\mathbf{x}^r,\mathbf{p},\Upsilon^i,\Lambda_2) \leq \bar{b},$$

where $\bar{a} := min\{x_1^r,\ldots,x_n^r\}$ and $\bar{b} := max\{x_1^r,\ldots,x_n^r\}$. Thus new means can be defined with three parameters:

$$\mathfrak{u}_{s,t,r}(\mathbf{x},\lambda,\Upsilon_i,\Lambda_2) := \begin{cases} (\mathfrak{u}_{s/r,t/r}(\mathbf{x}^r,\lambda,\Upsilon_i,\Lambda_2))^{\frac{1}{r}}; & r \neq 0, \\ \mathfrak{u}_{s,t}(\log\mathbf{x},\lambda,\Upsilon_i,\Lambda_1); & r = 0, \end{cases}$$

where $\log\mathbf{x} = (\log x_1,...,\log x_n)$.

The monotonicity of these three parameter means is followed by the monotonicity and continuity of the two parameter means.

Example 2.3 Let $I =]0,\infty[$, and consider the class of convex functions

$$\Lambda_3 = \{\eta_t :]0,\infty[\rightarrow]0,\infty[\mid t \in]0,\infty[\},$$

where

$$\eta_t(x) := \begin{cases} \frac{t^{-x}}{\log^2 t}; & t \neq 1, \\ \frac{x^2}{2}; & t = 1. \end{cases}$$

$t \mapsto \psi_t''(x)$ $(t \in]0,\infty[)$ is exponentially convex for every fixed $x \in]0,\infty[$, being the restriction of the Laplace transform of a nonnegative function (see [47] or [89] page 210).

Now fix $1 \leq i \leq 4$. We can get the exponential convexity of $t \mapsto \Upsilon_i(\psi_t)$ $(t \in \mathbb{R})$ as in Example 2.1. For the class Λ_3, (2.7) has the form

$$\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda_3) = \begin{cases} \left(\frac{\Upsilon_i(\eta_s)}{\Upsilon_i(\eta_t)}\right)^{\frac{1}{s-t}}; & s \neq t, \\ \exp\left(-\frac{2}{s\log s} - \frac{\Upsilon_i(id\eta_s)}{s\Upsilon_i(\eta_s)}\right); & s = t \neq 1, \\ \exp\left(-\frac{\Upsilon_i(id\eta_1)}{3\Upsilon_i(\eta_1)}\right); & s = t = 1. \end{cases}$$

The monotonicity of $\mathfrak{u}_{s,t}(\Upsilon_i, \Lambda_3)$ $(s,t \in]0,\infty[)$ comes from (2.6).

Suppose $\Upsilon_i(\eta_t) > 0$ $(t \in]0,\infty[)$, and let $a := min\{x_1,\ldots,x_n\}$, $b := max\{x_1,\ldots,x_n\}$, and define

$$\mathfrak{M}_{s,t}(\Upsilon_i, \Lambda_3) := -L(s,t)\log \mathfrak{u}_{s,t}(\Upsilon_i, \Lambda_3), \quad s,t \in]0,\infty[,$$

where $L(s,t)$ is the well known logarithmic mean

$$L(s,t) := \begin{cases} \frac{s-t}{\log s - \log t}; & s \neq t, \\ t; & s = t. \end{cases}$$

From Theorem 2.5 we have

$$a \leq \mathfrak{M}_{s,t}(\Upsilon_i, \Lambda_3) \leq b, \quad s,t \in]0,\infty[,$$

and therefore we get means.

Example 2.4 Let $I =]0,\infty[$ and consider the class of convex functions

$$\Lambda_4 = \{\gamma_t :]0,\infty[\rightarrow]0,\infty[\mid t \in]0,\infty[\},$$

where

$$\gamma_t(x) := \frac{e^{-x\sqrt{t}}}{t}.$$

$t \mapsto \psi_t''(x) = e^{-x\sqrt{t}}$, $t \in]0,\infty[$ is exponentially convex for every fixed $x \in]0,\infty[$, being the restriction of the Laplace transform of a non-negative function (see [47] or [89] page 214).

Now fix $1 \leq i \leq 4$. As before $t \mapsto \Upsilon_i(\psi_t)$ $(t \in \mathbb{R})$ is exponentially convex and differentiable. For the class Λ_4, (2.7) becomes

$$\mathfrak{u}_{s,t}(\Upsilon_i,\Lambda_4) = \begin{cases} \left(\frac{\Upsilon_i(\gamma_s)}{\Upsilon_i(\gamma_t)}\right)^{\frac{1}{s-t}}; & s \neq t, \\ \exp\left(-\frac{1}{t} - \frac{\Upsilon_i(id\gamma_t)}{2\sqrt{t}\Upsilon_i(\gamma_t)}\right); & s = t, \end{cases}$$

where id means the identity function on $]0,\infty[$. The monotonicity of $\mathfrak{u}_{s,t}(\Upsilon_i,\Lambda_4)$ $(s,t \in]0,\infty[)$ is followed by (2.6).

Suppose $\Upsilon_i(\eta_t) > 0$ $(t \in]0,\infty[)$, let $a := min\{x_1,...,x_n\}$, $b := max\{x_1,...,x_n\}$, and define

$$\mathfrak{M}_{s,t}(\Upsilon_i,\Lambda_4) := -(\sqrt{s}+\sqrt{t})\log \mathfrak{u}_{s,t}(\Upsilon_i,\Lambda_4), \quad s,t \in]0,\infty[.$$

Then Theorem 2.5 yields that

$$a \leq \mathfrak{M}_{s,t}(\Upsilon_i,\Lambda_4) \leq b,$$

thus we have new means.

2.2 Cyclic refinements of the discrete and integral form of Jensen's inequality with applications

In this section we introduce new refinements both the discrete and the classical Jensen's inequality. First, we extend Theorem 2.1: the weighted version is given in real vector spaces. By using this result, we obtain new refinements of the classical Jensen's inequality. m-exponential convexity of some functionals coming from the new refinements are investigated. To apply our results we define some new mixed symmetric means, generalized means, and Caucy-means, and study their properties.

2.2.1 Cyclic refinements of the discrete and classical Jensen's inequalities

We say that the numbers $p_1,\ldots,p_n$ represent a (positive) discrete probability distribution if $(p_i > 0)$ $p_i \geq 0$ $(1 \leq i \leq n)$ and $\sum\limits_{i=1}^{n} p_i = 1$.

To refine the discrete Jensen's inequality, we need the following hypotheses:

(H_1) Let $2 \leq k \leq n$ be integers, and let $p_1,\ldots,p_n$ and $\lambda_1,\ldots,\lambda_k$ represent positive probability distributions.

(H_2) Let C be a convex subset of a real vector space V, and $f : C \to \mathbb{R}$ be a convex function.

Theorem 2.6 *Assume* (H_1) *and* (H_2). *If* $v_1,\ldots,v_n \in C$, *then*

$$f\left(\sum_{i=1}^{n} p_i v_i\right) \le C_{dis} = C_{dis}\left(f,\mathbf{v},\mathbf{p},\lambda\right) \tag{2.14}$$

$$:= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}v_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \le \sum_{i=1}^{n} p_i f\left(v_i\right)$$

where $i+j$ *means* $i+j-n$ *in case of* $i+j>n$.

Proof. By the discrete Jensen's inequality

$$C_{dis} \le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}f\left(v_{i+j}\right)\right) = \left(\sum_{i=1}^{n} p_i f\left(v_i\right)\right)\left(\sum_{j=1}^{k}\lambda_j\right) = \sum_{i=1}^{n} p_i f\left(v_i\right).$$

The left hand side inequality can be proved similarly. Since

$$\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) = 1,$$

the discrete Jensen's inequality implies that

$$C_{dis} \ge f\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}v_{i+j}\right)\right) = f\left(\sum_{i=1}^{n} p_i v_i\right).$$

The proof is complete. □

The previous result can be considered as the weighted form of Theorem 2.1.

To refine the classical Jensen's inequality, we first introduce some hypotheses and notations.

(H_3) Let $(X,\mathscr{B},\mu)$ be a probability space.

Let $l \ge 2$ be a fixed integer. The σ-algebra in X^l generated by the projection mappings $pr_m : X^l \to X$ $(m = 1,\ldots,l)$

$$pr_m\left(x_1,\ldots,x_l\right) := x_m$$

is denoted by $\mathscr{B}^l$. μ^l means the product measure on $\mathscr{B}^l$: this measure is uniquely (μ is σ-finite) specified by

$$\mu^l\left(B_1 \times \ldots \times B_l\right) := \mu\left(B_1\right)\ldots\mu\left(B_l\right), \quad B_m \in \mathscr{B}, \quad m = 1,\ldots,l.$$

(H_4) Let g be a μ-integrable function on X taking values in an interval $I \subset \mathbb{R}$.

(H_5) Let f be a convex function on I such that $f \circ g$ is μ-integrable on X.

Under the conditions (H_1) and (H_3-H_5) we define

$$C_{int} = C_{int}(f, g, \mu, \mathbf{p}, \lambda)$$
$$:= \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \int_{X^n} f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right) d\mu^n(x_1, \ldots, x_n), \quad (2.15)$$

and for $t \in [0,1]$

$$C_{par}(t) = C_{par}(t, f, g, \mu, \mathbf{p}, \lambda) := \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right)$$
$$\cdot \int_{X^n} f \left(t \frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} + (1-t) \int_X g d\mu \right) d\mu^n(x_1, \ldots, x_n), \quad (2.16)$$

where $i+j$ means $i+j-n$ in case of $i+j > n$.

Remark 2.6 It follows from Lemma 2.1 (b) in [40] that the integrals in (2.15) and (2.16) exist and finite.

First, the essential properties of the function C_{par} are given.

Theorem 2.7 *Assume (H_1) and (H_3-H_5). Then*
(a) C_{par} is convex and increasing.
(b)

$$C_{par}(0) = f\left(\int_X g d\mu \right), \quad C_{par}(1) = C_{int}.$$

(c) C_{par} is continuous on $[0,1[$.
(d) If f is continuous, then C_{par} is continuous on $[0,1]$.

Proof. (a) Convexity is invariant under affine maps, the integral is monotonic, and the sum of convex functions is also convex: these imply that C_{par} is convex on $[0,1]$.

By the classical Jensen's inequality

$$C_{par}(t) \geq \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right)$$
$$\cdot f \left(\int_{X^n} \left(t \frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} + (1-t) \int_X g d\mu \right) d\mu^n(x_1, \ldots, x_n) \right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) f \left(t \int_X g d\mu + (1-t) \int_X g d\mu \right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) f \left(\int_X g d\mu \right) = f \left(\int_X g d\mu \right) = C_{par}(0), \quad t \in [0,1].$$

Suppose $0 \leq t_1 < t_2 \leq 1$. The convexity of C_{par}, and $C_{par}(t) \geq C_{par}(0)$ $(t \in [0,1])$ imply that

$$\frac{C_{par}(t_2) - C_{par}(t_1)}{t_2 - t_1} \geq \frac{C_{par}(t_2) - C_{par}(0)}{t_2} \geq 0,$$

and thus

$$C_{par}(t_2) \geq C_{par}(t_1).$$

(b) These are obvious.

(c) It follows from (a).

(d) We have only to show that C_{par} is continuous at 1. To this end, it is enough to check that the functions

$$t \to \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \int_{X^n} f \left(t \frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} + (1-t) \int_X g d\mu \right) d\mu^n (x_1, \ldots, x_n),$$

$$t \in [0,1], \quad i = 1, \ldots, n$$

are all continuous at 1. To prove this, fix i from $\{1, \ldots, n\}$, and let (t_n) be a sequence from $[0,1]$ which converges to 1.

Since f is continuous

$$t_n \to f \left(t_n \frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} + (1-t_n) \int_X g d\mu \right)$$

$$\overset{n \to \infty}{\to} f \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j} g(x_{i+j})}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right), \quad (x_1, \ldots, x_n) \in X^n. \tag{2.17}$$

It comes from the discrete Jensen's inequality that

$$f\left(t\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}+(1-t)\int\limits_X gd\mu\right)$$
$$\le tf\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)+(1-t)f\left(\int\limits_X gd\mu\right)$$
$$\le \max\left(f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right),f\left(\int\limits_X gd\mu\right)\right) \tag{2.18}$$

for all $t\in[0,1]$ and $(x_1,\ldots,x_n)\in X^n$.

Choose a fixed interior point a of I. Since f is convex

$$f(t)\ge f(a)+f'_+(a)(z-a), \quad z\in I,$$

where $f'_+(a)$ means the right-hand derivative of f at a. It follows from this that

$$f\left(t\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}+(1-t)\int\limits_X gd\mu\right)$$
$$\ge f(a)+f'_+(a)\left(t\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}+(1-t)\int\limits_X gd\mu-a\right)$$
$$\ge f(a)-af'_+(a)+\min\left(f'_+(a)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g\left(x_{i+j}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}-a\right),f'_+(a)\int\limits_X gd\mu\right) \tag{2.19}$$

for all $t\in[0,1]$ and $(x_1,\ldots,x_n)\in X^n$.

The functions in (2.18) and (2.19) do not depend on t, and μ^n-integrable, and therefore Lebesgue's dominated convergence theorem and (2.17) imply that

$$C_{par}(t_n)\to C_{par}(1).$$

The proof is complete. □

We illustrate by a concrete example that C_{par} is not continuous at 1 in general.

Example 2.5 Let $k=n=2$, and $p_1=p_2=\lambda_1=\lambda_2=\frac{1}{2}$. We consider the measure space $\left([0,1],\mathscr{B},\frac{1}{2}\varepsilon_{1/2}+\frac{1}{2}\varepsilon_1\right)$, where $\mathscr{B}$ is the σ-algebra of Borel subsets of $[0,1]$, and $\varepsilon_{1/2}$ and ε_1 are the Dirac measures at $1/2$ and 1, respectively. Denote $\mu=\frac{1}{2}\varepsilon_{1/2}+\frac{1}{2}\varepsilon_1$. Define the functions $f,g:[0,1]\to\mathbb{R}$ by

$$g(x)=x, \quad f(x)=\begin{cases} x, \text{ if } 0\le x<1 \\ 2, \text{ if } x=1 \end{cases}.$$

In this case for every $t\in[0,1]$

$$\begin{aligned}
C_{par}(t) &= \frac{1}{2}\int\limits_{[0,1]^2} f\left(t\frac{x_1+x_2}{2}+(1-t)\int\limits_{[0,1]} x d\mu\right) d\mu^2(x_1,x_2) \\
&= \frac{1}{2}\int\limits_{[0,1]^2} f\left(t\frac{x_1+x_2}{2}+(1-t)\frac{3}{4}\right) d\mu^2(x_1,x_2) \\
&= \frac{1}{2}\left(\frac{1}{4}f\left(\frac{1}{2}t+(1-t)\frac{3}{4}\right)+\frac{2}{4}f\left(\frac{3}{4}\right)+\frac{1}{4}f\left(t+(1-t)\frac{3}{4}\right)\right).
\end{aligned}$$

It can be seen from this that

$$\lim_{t\to 1} C_{par}(t)=\frac{1}{2}\left(\frac{1}{4}\cdot\frac{1}{2}+\frac{2}{4}\cdot\frac{3}{4}+\frac{1}{4}\right)=\frac{3}{8},$$

while

$$C_{par}(1)=\frac{1}{2}\left(\frac{1}{4}\cdot\frac{1}{2}+\frac{2}{4}\cdot\frac{3}{4}+\frac{1}{4}\cdot 2\right)=\frac{1}{2}.$$

Now our second main result is the next:

Theorem 2.8 *Assume (H_1) and (H_3-H_5). Then*

$$f\left(\int\limits_X g d\mu\right)\le C_{par}(t)\le C_{int}\le\int\limits_X f\circ g d\mu, \quad t\in[0,1].$$

Proof. It follows from Theorem 2.7 that

$$f\left(\int\limits_X g d\mu\right)\le C_{par}(t)\le C_{int}, \quad t\in[0,1].$$

The discrete Jensen's inequality yields

$$\begin{aligned}
C_{int} &= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\int_{X^n} f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}g(x_{i+j})}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)d\mu^n(x_1,\ldots,x_n)\\
&\le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\int_{X^n}\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}f(g(x_{i+j}))}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}d\mu^n(x_1,\ldots,x_n)\\
&= \int_{X^n}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}f(g(x_{i+j}))\right)d\mu^n(x_1,\ldots,x_n)\\
&= \int_{X^n}\sum_{i=1}^{n}p_i f(g(x_i))\,d\mu^n(x_1,\ldots,x_n) = \int_X f\circ g\,d\mu.
\end{aligned}$$

The proof is complete. □

2.2.2 Applications to mixed symmetric means

Consider the following hypotheses for this section.
(M_1) Let $I \subset \mathbb{R}$ be an interval, $\mathbf{x} := (x_1,\ldots,x_n) \in I^n$ and let $p_1,\ldots,p_n$ and $\lambda_1,\ldots,\lambda_k$ represent positive probability distributions for $2 \le k \le n$.
(M_2) Let $f : I \to \mathbb{R}$ be a convex function.
(M_3) Let $\phi,\ \psi : I \to \mathbb{R}$ be continuous and strictly monotone functions.

Assume (M_1). Then we define the power means of order $r \in \mathbb{R}$ as follows:

$$M_r(x_i^{i+k-1};p_i^{i+k-1};\lambda_1^k) = M_r(x_i,\ldots,x_{i+k-1};p_i,\ldots,p_{i+k-1};\lambda_1,\ldots,\lambda_k)$$
$$:= \begin{cases} \left(\dfrac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}^r}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)^{\frac{1}{r}}; & r \ne 0,\\[2ex] \left(\prod\limits_{j=0}^{k-1}x_{i+j}^{\lambda_{j+1}p_{i+j}}\right)^{\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}}; & r = 0, \end{cases} \tag{2.20}$$

and weighted cyclic mixed symmetric means corresponding to C_{dis} are

$$M_{r,s}(\mathbf{x},\mathbf{p},\lambda) = \begin{cases} \left(\sum\limits_{i=1}^{n}\left(\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)M_r^s(x_i^{i+k-1};p_i^{i+k-1};\lambda_1^k)\right)^{\frac{1}{s}}; & s \ne 0,\\[2ex] \prod\limits_{i=1}^{n}\left(M_r(x_i^{i+k-1};p_i^{i+k-1};\lambda_1^k)\right)^{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}; & s = 0, \end{cases}$$

where $i+j$ means $i+j-n$ in case of $i+j > n$.

The standard power means of order $r \in \mathbb{R}$ for the positive n-tuple $\mathbf{x}$ and probability distributions $p_1, \ldots, p_n$, are

$$M_r(\mathbf{x},\mathbf{p}) = M_r(x_1,...,x_n;p_1,...,p_n) := \begin{cases} \left(\sum\limits_{i=1}^{n} p_i x_i^r\right)^{\frac{1}{r}}; & r \neq 0, \\ \left(\prod\limits_{i=1}^{n} x_i{}^{p_i}\right); & r = 0. \end{cases}$$

The bounds for weighted cyclic mixed symmetric means are power means, as given in the following result.

Corollary 2.5 *Assume (M_1) and $r,s \in \mathbb{R}$ such that $r \leq s$. Then*

$$M_r(\mathbf{x},\mathbf{p}) \leq M_{s,r}(\mathbf{x},\mathbf{p},\lambda) \leq M_s(\mathbf{x},\mathbf{p}). \tag{2.21}$$

Proof. Apply Theorem 2.6. □

Assume (M_1) and (M_3). Then we define the cyclic generalized means with respect to C_{dis} as follows:

$$M_{\phi,\psi}(\mathbf{x},\mathbf{p},\lambda) := \phi^{-1}\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\phi\circ\psi^{-1}\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\psi(x_{i+j})}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\right),$$

where $i+j$ means $i+j-n$ in case of $i+j>n$.

Let $q : I \to \mathbb{R}$ be a continuous and strictly monotone function then the standard quasi-arithmetic means of $x_1, \ldots, x_n$ for probability distribution $p_1, \ldots, p_n$ are given by

$$M_q(\mathbf{x},\mathbf{p}) = M_q(x_1,\ldots,x_n;p_1,\ldots,p_n) := q^{-1}\left(\sum_{i=1}^{n} p_i q(x_i)\right).$$

The relation among the cyclic generalized means and quasi-arithmetic means is given in the next corollary.

Corollary 2.6 *Assume (M_1) and (M_3). Then*

$$M_\psi(\mathbf{x},\mathbf{p}) \leq M_{\phi,\psi}(\mathbf{x},\mathbf{p},\lambda) \leq M_\phi(\mathbf{x},\mathbf{p}) \tag{2.22}$$

if either $\phi\circ\psi^{-1}$ is convex and ϕ is strictly increasing or $\phi\circ\psi^{-1}$ is concave and ϕ is strictly decreasing.

Proof. We apply Theorem 2.6. □

The unweighted versions of Corollaries 2.5 and 2.6 are given in [13].

Let $(X,\mathscr{A},\mu)$ be a measure space with $0<\mu(X)<\infty$, $r\in\mathbb{R}$, and $u:X\to\mathbb{R}$ be a positive measurable function such that u^r is μ-integrable, if $r\neq 0$, and $\log\circ u$ is μ-integrable, if $r=0$. Then the integral power means of order r are defined by (see [36]):

$$\widetilde{M}_r(u,\mu) := \begin{cases} \left(\frac{1}{\mu(X)}\int_X (u(x))^r d\mu(x)\right)^{\frac{1}{r}}, & r\neq 0, \\ \exp\left(\frac{1}{\mu(X)}\int_X \log(u(x))\, d\mu(x)\right), & r=0. \end{cases} \tag{2.23}$$

$(\widetilde{\mathscr{H}_6})$ Let $(X,\mathscr{A},\mu)$ be a probability space, and $u:X\to\mathbb{R}$ be a measurable function. Suppose probability distribution $(p_1,\ldots,p_n)$.

Under the conditions (H$_1$) and (H$_3$-H$_4$), we define the following cyclic mixed means corresponding to $C_{par}(t)$ for the class of positive μ-integrable functions g for which g^s (if $s\neq 0$) and $\log\circ g$ (if $s=0$) are also μ-integrable.

$$M_{r,s}(t,f,g,\mu,\mathbf{p},\lambda) :=$$

$$\begin{cases} \left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\cdot\int_{X^n}\left(t\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}(g(x_{i+j}))^s}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}+(1-t)\int_X g^s d\mu\right)^{\frac{r}{s}} d\mu^n(x_1,\ldots,x_n)\right)^{\frac{1}{r}} & ; r,s\neq 0, \\ \left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\cdot\int_{X^n}\left(\exp\left(\begin{array}{l}\log\left(\prod_{j=0}^{k-1}\left(g\left(x_{i+j}\right)\right)^{\lambda_{j+1}p_{i+j}}\right)^{\frac{t}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}} \\ +(1-t)\int_X\log(g(x))d\mu(x)\end{array}\right)\right)^r d\mu^n(x_1,\ldots,x_n)\right)^{\frac{1}{r}} & ; r\neq 0, \\ \exp\left(\frac{1}{s}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\cdot\int_{X^n}\log\left(\begin{array}{l}t\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}(g(x_{i+j}))^s}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}} \\ +(1-t)\int_X g^s d\mu\end{array}\right)d\mu^n(x_1,\ldots,x_n)\right) & ; s\neq 0, \\ \exp\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\cdot\int_{X^n}\left(\begin{array}{l}\log\left(\prod_{j=0}^{k-1}\left(g\left(x_{i+j}\right)\right)^{\lambda_{j+1}p_{i+j}}\right)^{\frac{t}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}} \\ +(1-t)\int_X\log(g(x))d\mu(x)\end{array}\right)d\mu^n(x_1,\ldots,x_n)\right) & ; s=0. \end{cases}$$

where $i+j$ means $i+j-n$ in case of $i+j>n$.

The cyclic mixed symmetric means corresponding to C_{int} are $M_{r,s}(0,f,g,\mu,\mathbf{p},\lambda)$.

Corollary 2.7 *Assume (M$_1$), (H$_3$) and (H$_4$). Let $r,s\in\mathbb{R}$ such that $r\leq s$ and suppose that g^s, g^r are μ-integrable functions for $r,s\neq 0$ and $\log\circ g$ is μ-integrable function if either $r=0$ or $s=0$. Then*

$$\widetilde{M}_r(\mathbf{x},\mathbf{p})\leq M_{r,s}(t,f,g,\mu,\mathbf{p},\lambda)\leq M_{r,s}(0,f,g,\mu,\mathbf{p},\lambda)\leq\widetilde{M}_s(\mathbf{x},\mathbf{p}). \tag{2.24}$$

Proof. Apply Theorem 2.8. $\square$

Assume (H_1), (H_3-H_4) and (M_3). Then we define the following quasi-arithmetic means with respect to C_{par} for μ-integrable functions $\phi \circ g$ and $\psi \circ g$.

$$M_{\phi,\psi}(t,f,g,\mu,\mathbf{p},\lambda) :=$$
$$\varphi^{-1}\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\int_{X^n}\phi\circ\psi^{-1}\left(\begin{array}{c} t\dfrac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\psi\circ g(x_{i+j})}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}} \\ +(1-t)\int_X \psi\circ g d\mu\end{array}\right)d\mu^n(x_1,\ldots,x_n)\right),$$

and standard quasi-arithmetic means are

$$\widetilde{M}_{\phi}(g,\mu) := \phi^{-1}\left(\int_X \phi\circ g d\mu\right).$$

The quasi- arithmetic mean related to C_{int} are $\widetilde{M}_{\psi,\phi}(0,f,g,\mu,\mathbf{p},\lambda)$.

Corollary 2.8 *Assume (H_1), (H_3-H_4) and (M_3). Suppose $\phi\circ g$ and $\psi\circ g$ are μ-integrable functions. If either $\phi\circ\psi^{-1}$ is convex and ϕ is increasing, or $\phi\circ\psi^{-1}$ is concave and ϕ is decreasing, then*

$$\widetilde{M}_{\psi}(g,\mu) \le \widetilde{M}_{\phi,\psi}(t,f,g,\mu,\mathbf{p},\lambda) \le \widetilde{M}_{\phi,\psi}(0,f,g,\mu,\mathbf{p},\lambda) \le \widetilde{M}_{\phi}(g,\mu),$$

while if either $\psi\circ\phi^{-1}$ is convex and ψ is decreasing, or $\psi\circ\phi^{-1}$ is concave and ψ is increasing, then

$$\widetilde{M}_{\phi}(g,\mu) \le \widetilde{M}_{\psi,\phi}(t,f,g,\mu,\mathbf{p},\lambda) \le \widetilde{M}_{\psi,\phi}(0,f,g,\mu,\mathbf{p},\lambda) \le \widetilde{M}_{\psi}(g,\mu).$$

Proof. Apply Theorem 2.8. □

Remark 2.7 Under the conditions (M_1), we define

$$J_1(f) = J_1(\mathbf{x},\mathbf{p},\lambda,f) := \sum_{i=1}^{n} p_i f(v_i) - C_{dis}(f,\mathbf{x},\mathbf{p},\lambda)$$

$$J_2(f) = J_2(\mathbf{x},\mathbf{p},\lambda,f) := C_{dis}(f,\mathbf{x},\mathbf{p},\lambda) - f\left(\sum_{i=1}^{n} p_i v_i\right)$$

where $f : I \to \mathbb{R}$ is a function. The functionals $f \to J_i(f)$ are linear, $i = 1,2$, and Theorem 2.6 imply that

$$J_i(f) \ge 0, \quad i = 1,2$$

if $f : I \to \mathbb{R}$ is a convex function.

Assume (H_1) and (H_3-H_5). Then we have the following more linear functionals

$$J_3(f) = J_3(f,g,\mu,\mathbf{p},\lambda) := \int_X f\circ g d\mu - C_{int}(f,g,\mu,\mathbf{p},\lambda) \ge 0,$$

$$J_4(f) = J_4(t,f,g,\mu,\mathbf{p},\lambda) := \int_X f\circ g d\mu - C_{par}(t,f,g,\mu,\mathbf{p},\lambda) \ge 0; \quad t\in[0,1],$$

$$J_5(f) = J_5(t,f,g,\mu,\mathbf{p},\lambda) := C_{int}(f,g,\mu,\mathbf{p},\lambda) - C_{par}(t,f,g,\mu,\mathbf{p},\lambda) \geq 0; \quad t \in [0,1],$$

$$J_6(f) = J_6(t,f,g,\mu,\mathbf{p},\lambda) := C_{par}(t,f,g,\mu,\mathbf{p},\lambda) - f\left(\int_X g d\mu\right) \geq 0; \quad t \in [0,1],$$

$$J_7(f) = J_6(f,g) := \int_X f \circ g d\mu - f\left(\int_X g d\mu\right).$$

The log-convexity, exponential convexity and m-exponential convexity and related results for $J_7(f)$ can be found in [36].

2.2.3 m-exponential convexity, mean value theorems and Cauchy means

We apply the method given in [83], to prove the m-exponential convexity and exponential convexity of the functionals $f \rightarrow J_i(f)$ for $i = 1,\ldots,6$, together with the Lagrange type and Cauchy type mean value theorems. The same method is used for Theorem 2.1. Hence the extension of Theorem 2.2 is as follows.

Theorem 2.9 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t \mid t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \rightarrow [y_0,y_1,y_2;\phi_t]$ $(t \in J)$ is m-exponentially convex in the Jensen sense on I for every three mutually different points $y_0,y_1,y_2 \in I$. Let $J_i(f)$ $(i = 1,\ldots,6)$ be the linear functionals constructed in Remark 2.7. Then $t \rightarrow J_i(\phi_t)$ $(t \in J)$ is an m-exponentially convex function in the Jensen sense on I for each $i = 1,\ldots,6$. If the function $t \rightarrow J_i(\phi_t)$ $(t \in J)$ is continuous for $i = 1,\ldots,6$, then it is m-exponentially convex on I for $i = 1,\ldots,6$.*

Proof. The proof is same as of 2.2. □

Similarly, the extensions for Corollary 2.3 and Corollary 2.4 are as follows.

Corollary 2.9 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t \mid t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \rightarrow [y_0,y_1,y_2;\phi_t]$ $(t \in J)$ is exponentially convex in the Jensen sense on I for every three mutually different points $y_0,y_1,y_2 \in I$. Let $J_i(f)$ $(i = 1,\ldots,6)$ be the linear functionals constructed in Remark 2.7. Then $t \rightarrow J_i(\phi_t)$ $(t \in J)$ is an exponentially convex function in the Jensen sense on I for $i = 1,\ldots,6$. If the function $t \rightarrow J_i(\phi_t)$ $(t \in J)$ is continuous, then it is exponentially convex on I for $i = 1,\ldots,6$.*

Corollary 2.10 *Assume $I \subset \mathbb{R}$ is an interval, and assume $\Lambda = \{\phi_t : t \in J\}$ is a family of functions defined on an interval $I \subset \mathbb{R}$, such that the function $t \rightarrow [y_0,y_1,y_2;\phi_t]$ $(t \in J)$ is 2-exponentially convex in the Jensen sense on I for every three mutually different points $y_0,y_1,y_2 \in I$. Let $J_i(f)$ $(i = 1,\ldots,6)$ be the linear functionals constructed in Remark 2.7. Then the following two statements hold for $i = 1,\ldots,6$:*

(i) *If the function* $t \to J_i(\phi_t)$ $(t \in J)$ *is positive and continuous, then it is* 2*-exponentially convex on I, and thus* log*-convex.*

(ii) *If the function* $t \to J_i(\phi_t)$ $(t \in J)$ *is positive and differentiable, then for every* $s,t,u,v \in J$*, such that* $s \le u$ *and* $t \le v$*, we have*

$$\mathfrak{u}_{s,t}(J_i,\Lambda) \le \mathfrak{u}_{u,v}(J_i,\Lambda) \tag{2.25}$$

where

$$\mathfrak{u}_{s,t}(J_i,\Lambda) := \begin{cases} \left(\frac{J_i(\phi_s)}{J_i(\phi_t)}\right)^{\frac{1}{s-t}}, & s \neq t, \\ \exp\left(\frac{\frac{d}{ds}J_i(\phi_s)}{J_i(\phi_s)}\right), & s = t \end{cases} \tag{2.26}$$

for $\phi_s, \phi_t \in \Lambda$.

The extensions of mean value theorems for the linear functionals $J_i(f)$ $(i = 1,...,6)$ are as follows.

Theorem 2.10 *Let* $J_i(f)$ $(i = 1,...,6)$ *be the linear functionals constructed in Remark 2.7 and* $g \in C^2[a,b]$*. Then there exists* $\xi \in [a,b]$ *such that*

$$J_i(g) = \frac{1}{2} g''(\xi) J_i\left(x^2\right); \quad i = 1,...,6.$$

Proof. The proof is same as of Theorem 2.4. □

Theorem 2.11 *Let* $J_i(f)$ $(i = 1,...,6)$ *be the linear functionals constructed in Remark 2.7 and* $g,h \in C^2[a,b]$*. Then there exists* $\xi \in [a,b]$ *such that*

$$\frac{J_i(g)}{J_i(h)} = \frac{g''(\xi)}{h''(\xi)}; \quad i = 1,...,6,$$

provided that $J_i(h) \neq 0$ $(i = 1,...,6)$.

Proof. The proof is same as of Theorem 2.5. □

By the application of Theorem 2.11, the Cauchy means constructed in Section 2.1.4 are generalized for two probability distributions $\mathbf{p}$ and λ.

Example 2.6 Under the settings of Example 6.1 of [13], we apply Corollary 2.9 to get the exponential convexity of $t \mapsto J_i(\phi_t)$ $(t \in \mathbb{R})$ and the monotone functions $\mathfrak{w}_{s,t}$ in (2.26) become

$$\mathfrak{w}_{s,t}(J_i,\Lambda_1) = \begin{cases} \left(\frac{J_i(\phi_s)}{J_i(\phi_t)}\right)^{\frac{1}{s-t}}; & s \neq t, \\ \exp\left(\frac{J_i(id\,\phi_s)}{J_i(\phi_s)} - \frac{2}{s}\right); & s = t \neq 0, \\ \exp\left(\frac{J_i(id\,\phi_0)}{3J_i(\phi_0)}\right); & s = t = 0, \end{cases}$$

for $i = 1,...,6$.

Suppose $J_i(\phi_t) > 0$ $(t \in \mathbb{R})$, $a := min\{x_1,...,x_n\}$, $b := max\{x_1,...,x_n\}$, and let

$$\mathfrak{M}_{s,t}(J_i,\Lambda_1) := \log \mathfrak{u}_{s,t}(J_i,\Lambda_1); \quad s,t \in \mathbb{R}.$$

Then from Theorem 2.11 we have

$$a \leq \mathfrak{M}_{s,t}(J_i,\Lambda_1) \leq b,$$

and thus $\mathfrak{M}_{s,t}(J_i,\Lambda_1)$ $(s,t \in \mathbb{R})$ are means. The monotonicity of these means is followed by (2.25).

Similarly, the Examples 2.1–2.4 can also be extended for $J_i(\psi_t)$ $(t \in \mathbb{R})$ $(i = 1,...,6)$.

2.3 Further Applications To Holder's Inequality

We say that the numbers $p_i \geq 0$ $(1 \leq i \leq n)$ and $\sum\limits_{i=1}^{n} p_i = P_n$.

Theorem 2.12 *Let $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$ and w_i, x_i, y_i, $i =, 1, 2, 3, ...$ be arbitrary sequences of poistive real numbers . Then under the assumptions of Theorem 2.6 the following inequalities hold:*

$$\left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} w_i y_i^q\right)^{\frac{1}{q}} \geq \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} w_{i+j} y_{i+j}^q\right)^{\frac{1}{q}} \left(\sum_{j=0}^{k-1} \lambda_{j+1} w_{i+j} x_{i+j}^p\right)^{\frac{1}{p}}$$
$$\geq \sum_{i=1}^{n} w_i x_i y_i. \tag{2.27}$$

where $i+j$ means $i+j-n$ in case of $i+j > n$.
If $0 < p < 1$ then inequalities sign are reversed in (2.27).

Proof. Consider the family of functions $f(x) = \frac{x^s}{s(s-1)}$ $s \neq 0,1$. Clearly $f''(x) = x^{s-2} > 0$ for all $x > 0$. Putting in weighted version of Theorem (2.14), we get

$$\frac{1}{s(1-s)} \left(\frac{1}{P_n} \sum_{i=1}^{n} p_i x_i\right)^s \geq \frac{1}{s(1-s)} \frac{1}{P_n} \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right) \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^s$$
$$\geq \frac{1}{s(1-s)} \frac{1}{P_n} \sum_{i=1}^{n} p_i \left(x_i^s\right) \tag{2.28}$$

Consider the substitutions $s=\frac{1}{p}$, $1-s=\frac{1}{q}$, $p_i=\frac{w_iy_i^q}{\sum\limits_{i=1}^{n}w_iy_i^q}$, $x_i=\frac{x_i^p}{y_i^q}$ in (2.28), we get

OR

$$pq\left(\frac{1}{\frac{\sum_{i=1}^{n}w_iy_i^q}{\sum_{i=1}^{n}w_iy_i^q}}\sum_{i=1}^{n}\frac{w_iy_i^q}{\sum_{i=1}^{n}w_iy_i^q}\frac{x_i^p}{y_i^q}\right)^{\frac{1}{p}}$$

$$\geq pq\frac{1}{\frac{1}{\frac{\sum_{i=1}^{n}w_iy_i^q}{\sum_{i=1}^{n}w_iy_i^q}}}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}\frac{w_{i+j}y_{i+j}^q}{\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{w_{i+j}y_{i+j}^q}{\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q}\frac{x_{i+j}^p}{y_{i+j}^q}}{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{w_{i+j}y_{i+j}^q}{\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q}}\right)^{\frac{1}{p}}$$

$$\geq pq\frac{1}{\frac{\sum_{i=1}^{n}w_iy_i^q}{\sum_{i=1}^{n}w_iy_i^q}}\sum_{i=1}^{n}\frac{w_iy_i^q}{\sum_{i=1}^{n}w_iy_i^q}\left(\frac{x_i^p}{y_i^q}\right)^{\frac{1}{p}} \tag{2.29}$$

Since $\sum\limits_{i=1}^{n}\sum\limits_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q=\sum\limits_{i=1}^{n}w_iy_i^q\sum\limits_{i=1}^{n}\lambda_i=\sum\limits_{i=1}^{n}w_iy_i^q$, we get

$$pq\left(\sum_{i=1}^{n}w_ix_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}w_iy_i^q\right)^{\frac{-1}{p}}\geq pq\frac{1}{\sum_{i=1}^{n}w_iy_i^q}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}x_{i+j}^p}{\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q}\right)^{\frac{1}{p}}$$

$$\geq pq\sum_{i=1}^{n}w_ix_iy_i\left(\sum_{i=1}^{n}w_iy_i^q\right)^{\frac{-1}{p}} \tag{2.30}$$

After simplification, we will get (2.27). □

Theorem 2.13 *Let $p>1$, $\frac{1}{p}+\frac{1}{q}=1$ and $w_i,x_i,y_i,\ i=,1,2,3,\ldots$ be arbitrary sequences of positive real numbers . Then under the assumptions of Theorem 2.6 the following inequalities hold:*

$$\sum_{i=1}^{n}w_ix_iy_i\leq\sum_{i=1}^{n}\left(w_iy_i^q\right)^{\frac{1}{q}}\sum_{i=1}^{n}\frac{\left(\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}x_{i+j}\right)}{\left(\sum_{j=0}^{k-1}\lambda_{j+1}w_{i+j}y_{i+j}^q\right)^{\frac{1}{q}}}$$

$$\leq\left(\sum_{i=1}^{n}w_ix_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}w_iy_i^q\right)^{\frac{1}{q}}. \tag{2.31}$$

where $i+j$ means $i+j-n$ in case of $i+j>n$.
The inequalities in (2.31) *are reversed for* $0<p<1$.

Proof. Consider the substitutions $f(x)=x^p$, $p_i=w_i y_i^q$, $x_i=x_i y_i^{1-q}$ in (2.14) and simplifying will give (2.31). $\square$

Chapter 3

Cyclic Improvements of Inequalities for Entropy of Zipf-Mandelbrot Law

The Jensen's inequality plays a crucial role to obtain inequalities for divergences between probability distributions. Divergences between probability distributions have been introduced to measure the difference between them. A lot of different type of divergences exist, for example the f-divergence (especially, Kullback–Leibler divergence, Hellinger distance and total variation distance), Rényi divergence, Jensen–Shannon divergence, etc. (see [63] and [91]). There are a lot of papers dealing with inequalities for divergences and entropies, see e.g. [32] and [88] and the references therein. The Jensen's inequality plays a crucial role some of these inequalities.

We first introduce some important definitions and results used for rest of this Chapter. The following notion was introduced by Csiszár in [19] and [18].

Definition 3.1 *Let $f :]0,\infty[\to]0,\infty[$ be a convex function, and let $\mathbf{p} := (p_1,\ldots,p_n)$ and $\mathbf{q} := (q_1,\ldots,q_n)$ be positive probability distributions. The f-divergence functional is*

$$I_f(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right).$$

It is possible to use nonnegative probability distributions in the f-divergence functional, by defining

$$f(0) := \lim_{t\to 0+} f(t); \quad 0f\left(\frac{0}{0}\right) := 0; \quad 0f\left(\frac{a}{0}\right) := \lim_{t\to 0+} tf\left(\frac{a}{t}\right), \quad a > 0.$$

Based on the previous definition, the following new functional was introduced in [38].

Definition 3.2 *Let $J \subset \mathbb{R}$ be an interval, and let $f : J \to \mathbb{R}$ be a function. Let $\mathbf{p} := (p_1, \ldots, p_n) \in \mathbb{R}^n$, and $\mathbf{q} := (q_1, \ldots, q_n) \in]0, \infty[^n$ such that*

$$\frac{p_i}{q_i} \in J, \quad i = 1, \ldots, n. \tag{3.1}$$

Then let

$$\tilde{I}_f(\mathbf{p}, \mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right).$$

Shannon entropy and the measures related to it are frequently applied in fields like population genetics, molecular ecology, information theory, dynamical systems and statistical physics(see [17, 61].

Definition 3.3 *The Shannon entropy of a positive probability distribution $\mathbf{p} := (p_1, \ldots, p_n)$ is defined by*

$$H(\mathbf{p}) := -\sum_{i=1}^{n} p_i \log(p_i).$$

One of the most famous distance functions used in information theory [14, 90], mathematical statistics [50, 92, 51] and signal processing [30, 60] is Kullback-Leibler distance. The **Kullback-Leibler** distance [58, 59] between the positive probability distributions $\mathbf{p} = (p_1, ..., p_n)$ and $\mathbf{q} = (q_1, ..., q_n)$ is defined by

Definition 3.4 *The Kullback-Leibler divergence between the positive probability distributions $\mathbf{p} := (p_1, \ldots, p_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ is defined by*

$$D(\mathbf{p} \| \mathbf{q}) := \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right).$$

Definition 3.5 *Zipf-Mandelbrot law is a discrete probability distribution depends on three parameters $N \in \{1, 2, \ldots\}$, $q \in [0, \infty[$ and $s > 0$, and it is defined by*

$$f(i; N, q, s) := \frac{1}{(i+q)^s H_{N,q,s}}, \quad i = 1, \ldots, N,$$

where

$$H_{N,q,s} := \sum_{k=1}^{N} \frac{1}{(k+q)^s}.$$

If $q = 0$, then Zipf–Mandelbrot law becomes Zipf's law.

Zipf's law is one of the basic laws in information science and bibliometrics. Zipf's law is concerning the frequency of words in the text. We count the number of times each word appears in the text. Words are ranked (r) according to the frequency of occurrence (f). The product of these two numbers is a constant: $r \cdot f = c$.

Apart from the use of this law in bibliometrics and information science, Zipf's law is frequently used in linguistics (see [22], p. 167). In economics and econometrics, this distribution is known as Pareto's law which analyze the distribution of the wealthiest members of the community (see [22], p. 125). These two laws are the same in the mathematical sense, they are only applied in a different context (see [26], p. 294).

The same type of distribution that we have in Zipf's and Pareto's law can be also found in other scientific disciplines, such as: physics, biology, earth and planetary sciences, computer science, demography and the social sciences. For example, the same type of distribution, which we also call the Power law, we can analyze the number of hits on web sites, the magnitude of earthquakes, diameter of moon craters, intensity of solar flares, intensity of wars, population of cities, and others (see [80]).

More general model introduced Benoit Mandelbrot (see [65]), by using arguments on the fractal structure of lexical trees.

The are also quite different interpretation of Zipf-Mandelbrot law in ecology, as it is pointed out in [79] (see also [29] and [93]).

3.1 Estimations of f- and Rényi divergences by using a cyclic refinement of the Jensen's inequality

In this section, we obtain inequalities for Rényi and Shannon entropies from cyclic refinements of Jensen's inequality results. Finally, some concrete cases are considered, by using Zipf-Mandelbrot law.

It is generally common to take log with base of 2 in the introduced notions, but in our investigations this is not essential.

3.1.1 Inequalities for Csiszár divergence and Shannon entropy

In the first result we apply Theorem 2.6 to $\tilde{I}_f(\mathbf{p},\mathbf{q})$.

Theorem 3.1 *Let $2 \le k \le n$ be integers, and let $\lambda := (\lambda_1,\ldots,\lambda_k)$ be a positive probability distribution. Let $J \subset \mathbb{R}$ be an interval, let $\mathbf{p} := (p_1,\ldots,p_n) \in \mathbb{R}^n$, and let $\mathbf{q} := (q_1,\ldots,q_n) \in]0,\infty[^n$ such that*

$$\frac{p_i}{q_i} \in J, \quad i = 1,\ldots,n.$$

(a) If $f : J \to \mathbb{R}$ is a convex function, then

$$\begin{aligned}\tilde{I}_f(\mathbf{p},\mathbf{q}) &= \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right)\\ &\ge \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) \ge f\left(\frac{\sum\limits_{i=1}^{n}p_i}{\sum\limits_{i=1}^{n}q_i}\right)\sum_{i=1}^{n}q_i. \qquad (3.2)\end{aligned}$$

If f is a concave function, then inequality signs in (3.2) are reversed.
(b) If $f : J \to \mathbb{R}$ is a function such that $x \to xf(x)$ $(x \in J)$ is convex, then

$$\tilde{I}_{id_J f}(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{n} p_i f\left(\frac{p_i}{q_i}\right)$$
$$\geq \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) \geq f\left(\frac{\sum_{i=1}^{n}p_i}{\sum_{i=1}^{n}q_i}\right)\sum_{i=1}^{n}p_i. \tag{3.3}$$

If $x \to xf(x)$ $(x \in J)$ is a concave function, then inequality signs in (3.3) are reversed. In all these inequalities $i+j$ means $i+j-n$ in case of $i+j>n$.

Proof. (a) By applying Theorem 2.6 with $C := J$, $f := f$,

$$p_i := \frac{q_i}{\sum_{i=1}^{n} q_i}, \quad v_i := \frac{p_i}{q_i}, \quad i = 1,\ldots,n$$

we have

$$\sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) = \left(\sum_{i=1}^{n} q_i\right)\cdot\sum_{i=1}^{n}\frac{q_i}{\sum_{i=1}^{n} q_i} f\left(\frac{p_i}{q_i}\right)$$
$$\geq \left(\sum_{i=1}^{n} q_i\right)\cdot\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}\frac{q_{i+j}}{\sum_{i=1}^{n} q_i}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{q_{i+j}}{\sum_{i=1}^{n} q_i}\frac{p_{i+j}}{q_{i+j}}}{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{q_{i+j}}{\sum_{i=1}^{n} q_i}}\right)$$
$$= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) \geq f\left(\frac{\sum_{i=1}^{n}p_i}{\sum_{i=1}^{n}q_i}\right)\sum_{i=1}^{n}q_i.$$

(b) We can prove similarly to (a), by using $f := id_J f$.
The proof is complete. □

Remark 3.1 (a) Csiszár and Körner classical inequality for the f-divergence functional is generalized and refined in (3.2).

(b) Other type of refinements are applied to the f-divergence functional in [23], [24] and [4].

(c) For example, the functions $x \to x\log_b(x)$ $(x>0,\ b>1)$ and $x \to x\arctan(x)$ $(x \in \mathbb{R})$ are convex.

We mention two special cases of the previous result.

The first case corresponds to the entropy of a discrete probability distribution.

Corollary 3.1 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$ be a positive probability distribution.*

(a) If $\mathbf{q} := (q_1, \ldots, q_n) \in \left]0, \infty\right[^n$, and the base of log *is greater than* 1, *then*

$$-\sum_{i=1}^{n} q_i \log (q_i) \leq -\sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \log \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \leq \log \left(\frac{n}{\sum\limits_{i=1}^{n} q_i} \right) \sum_{i=1}^{n} q_i. \quad (3.4)$$

If the base of log *is between* 0 *and* 1, *then inequality signs in (3.4) are reversed.*

(b) If $\mathbf{q} := (q_1, \ldots, q_n)$ is a positive probability distribution and the base of log *is greater than* 1, *then we have estimates for the Shannon entropy of* $\mathbf{q}$

$$H(\mathbf{q}) \leq -\sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \log \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \leq \log (n).$$

If the base of log *is between* 0 *and* 1, *then inequality signs in (3.4) are reversed.*

In all these inequalities $i+j$ means $i+j-n$ in case of $i+j > n$.

Proof. (a) It follows from Theorem 3.1 (a), by using $f := \log$ and $\mathbf{p} := (1, \ldots, 1)$.

(b) It is a special case of (a). □

The second case corresponds to the relative entropy or Kullback-Leibler divergence between two probability distributions.

Corollary 3.2 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$ be a positive probability distribution.*

(a) Let $\mathbf{p} := (p_1, \ldots, p_n) \in \left]0, \infty\right[^n$ and $\mathbf{q} := (q_1, \ldots, q_n) \in \left]0, \infty\right[^n$. If the base of log *is greater than* 1, *then*

$$\sum_{i=1}^{n} p_i \log \left(\frac{p_i}{q_i} \right) \geq \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \log \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) \geq \log \left(\frac{\sum\limits_{i=1}^{n} p_i}{\sum\limits_{i=1}^{n} q_i} \right) \sum_{i=1}^{n} p_i. \quad (3.5)$$

If the base of log *is between* 0 *and* 1, *then inequality signs in (3.5) are reversed.*

(b) If $\mathbf{p}$ *and* $\mathbf{q}$ *are positive probability distributions, and the base of* log *is greater than* 1, *then we have*

$$D(\mathbf{p} \| \mathbf{q}) \geq \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \log \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) \geq 0. \quad (3.6)$$

If the base of log *is between* 0 *and* 1, *then inequality signs in (3.6) are reversed.*

In all these inequalities $i+j$ means $i+j-n$ in case of $i+j > n$.

Proof. (a) We can apply Theorem 3.1 (b) to the function $f := \log$.

(b) It is a special case of (a). □

Remark 3.2 We can apply Theorem 3.1 to have similar inequalities for other distances between two probability distributions.

3.1.2 Inequalities for Rényi divergence and entropy

The Rényi divergence and entropy come from [85].

Definition 3.6 *Let $\mathbf{p} := (p_1, \ldots, p_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ be positive probability distributions, and let $\alpha \geq 0$, $\alpha \neq 1$.*

(a) The Rényi divergence of order α is defined by

$$D_\alpha(\mathbf{p}, \mathbf{q}) := \frac{1}{\alpha - 1} \log \left(\sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i} \right)^\alpha \right). \tag{3.7}$$

(b) The Rényi entropy of order α of $\mathbf{p}$ is defined by

$$H_\alpha(\mathbf{p}) := \frac{1}{1 - \alpha} \log \left(\sum_{i=1}^{n} p_i^\alpha \right). \tag{3.8}$$

The Rényi divergence and the Rényi entropy can also be extended to nonnegative probability distributions.

If $\alpha \to 1$ in (3.7), we have the Kullback-Leibler divergence, and if $\alpha \to 1$ in (3.8), then we have the Shannon entropy.

In the next two results inequalities can be found for the Rényi divergence.

Theorem 3.2 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$, $\mathbf{p} := (p_1, \ldots, p_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ be positive probability distributions.*

(a) If $0 \leq \alpha \leq \beta$, α, $\beta \neq 1$, and the base of log *is greater than* 1, *then*

$$D_\alpha(\mathbf{p}, \mathbf{q}) \leq \frac{1}{\beta - 1} \log \left(\sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \left(\frac{p_{i+j}}{q_{i+j}} \right)^{\alpha-1}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right)^{\frac{\beta-1}{\alpha-1}} \right) \tag{3.9}$$

$$\leq D_\beta(\mathbf{p}, \mathbf{q})$$

The reverse inequalities hold if the base of log *is between* 0 *and* 1.

(b) If $1<\beta$*, and the base of* $\log$ *is greater than* 1*, then*

$$\begin{aligned}
D_1(\mathbf{p},\mathbf{q}) &= D(\mathbf{p}\|\mathbf{q}) = \sum_{i=1}^{n} p_i \log\left(\frac{p_i}{q_i}\right) \\
&\leq \frac{1}{\beta-1}\log\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\exp\left(\frac{(\beta-1)\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\log\left(\frac{p_{i+j}}{q_{i+j}}\right)}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\right) \\
&\leq D_\beta(\mathbf{p},\mathbf{q}),
\end{aligned}$$

where the base of $\exp$ *is the same as the base of* $\log$*.*

The reverse inequalities hold if the base of $\log$ *is between* 0 *and* 1*.*

(c) If $0\leq\alpha<1$*, and the base of* $\log$ *is greater than* 1*, then*

$$D_\alpha(\mathbf{p},\mathbf{q}) \leq \frac{1}{\alpha-1}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\log\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \leq D_1(\mathbf{p},\mathbf{q})$$

The reverse inequalities hold if the base of $\log$ *is between* 0 *and* 1*.*

In all these inequalities $i+j$ *means* $i+j-n$ *in case of* $i+j>n$*.*

Proof. (a) By applying Theorem 2.6 with $C :=]0,\infty[$, $f :]0,\infty[\to \mathbb{R}$, $f(t) := t^{\frac{\beta-1}{\alpha-1}}$,

$$v_i := \left(\frac{p_i}{q_i}\right)^{\alpha-1}, \quad i=1,\ldots,n,$$

we have

$$\begin{aligned}
\left(\sum_{i=1}^{n} q_i\left(\frac{p_i}{q_i}\right)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} &= \left(\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right)^{\frac{\beta-1}{\alpha-1}} \\
&\leq \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)^{\frac{\beta-1}{\alpha-1}} \leq \sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\beta-1}
\end{aligned} \tag{3.10}$$

if either $0\leq\alpha<1<\beta$ or $1<\alpha\leq\beta$, and the reverse inequalities hold in (3.33) if $0\leq\alpha\leq\beta<1$. By raising the power $\frac{1}{\beta-1}$, we have from all these cases that

$$\left(\sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i}\right)^{\alpha}\right)^{\frac{1}{\alpha-1}} \leq \left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)^{\frac{\beta-1}{\alpha-1}}\right)^{\frac{1}{\beta-1}}$$

$$\leq \left(\sum_{i=1}^{n} p_i \left(\frac{p_i}{q_i}\right)^{\beta-1}\right)^{\frac{1}{\beta-1}} = \left(\sum_{i=1}^{n} q_i \left(\frac{p_i}{q_i}\right)^{\beta}\right)^{\frac{1}{\beta-1}}.$$

Since log is increasing if the base of log is greater than 1, it now follows (3.9).

If the base of log is between 0 and 1, then log is decreasing, and therefore inequality signs in (3.9) are reversed.

(b) and (c) When $\alpha = 1$ or $\beta = 1$, we have the result by taking limit.

The proof is complete. □

Theorem 3.3 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$, $\mathbf{p} := (p_1, \ldots, p_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ be positive probability distributions.*

If either $0 \leq \alpha < 1$ and the base of log *is greater than* 1*, or $1 < \alpha$ and the base of* log *is between* 0 *and* 1*, then*

$$\frac{1}{\sum\limits_{i=1}^{n} q_i \left(\frac{p_i}{q_i}\right)^{\alpha}} \sum_{i=1}^{n} p_i \left(\frac{p_i}{q_i}\right)^{\alpha-1} \log\left(\frac{p_i}{q_i}\right) \leq \frac{1}{(\alpha-1)\sum\limits_{i=1}^{n} p_i \left(\frac{p_i}{q_i}\right)^{\alpha-1}} \times$$

$$\times \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}\right)\log\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \leq D_{\alpha}(\mathbf{p},\mathbf{q})$$

$$\leq \frac{1}{\alpha-1}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\log\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \leq D_1(\mathbf{p},\mathbf{q}) \quad (3.11)$$

If either $0 \leq \alpha < 1$ and the base of log *is between* 0 *and* 1*, or $1 < \alpha$ and the base of* log *is greater than* 1*, then the reverse inequalities holds.*

In all these inequalities $i + j$ means $i + j - n$ in case of $i + j > n$.

Proof. We prove only the case when $0 \leq \alpha < 1$ and the base of log is greater than 1, the other cases can be proved similarly.

Since $\frac{1}{\alpha-1} < 0$ and the function log is concave, we have from Theorem 2.6 by choosing $C :=]0,\infty[$, $f := \log$,

$$v_i := \left(\frac{p_i}{q_i}\right)^{\alpha-1}, \quad i = 1, \ldots, n,$$

that

$$
\begin{aligned}
D_\alpha(\mathbf{p},\mathbf{q}) &= \frac{1}{\alpha-1}\log\left(\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) \\
&\leq \frac{1}{\alpha-1}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\log\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&\leq \frac{1}{\alpha-1}\sum_{i=1}^{n} p_i\log\left(\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) = \sum_{i=1}^{n} p_i\log\left(\frac{p_i}{q_i}\right) = D_1(\mathbf{p},\mathbf{q})
\end{aligned}
$$

and this gives the desired upper bound for $D_\alpha(\mathbf{p},\mathbf{q})$.

Since the base of log is greater than 1, the function $x \to x\log(x)$ $(x>0)$ is convex, and therefore $\frac{1}{1-\alpha}<0$ and Theorem 2.6 imply that

$$
\begin{aligned}
D_\alpha(\mathbf{p},\mathbf{q}) &:= \frac{1}{\alpha-1}\log\left(\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) \\
&= \frac{1}{(\alpha-1)\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}}\left(\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right)\log\left(\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) \\
&\geq \frac{1}{(\alpha-1)\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\times \\
&\quad\times\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\log\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&\frac{1}{(\alpha-1)\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}\right)\log\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\left(\frac{p_{i+j}}{q_{i+j}}\right)^{\alpha-1}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&\geq \frac{1}{(\alpha-1)\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}}\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\log\left(\left(\frac{p_i}{q_i}\right)^{\alpha-1}\right) \\
&= \frac{1}{\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}}\sum_{i=1}^{n} p_i\left(\frac{p_i}{q_i}\right)^{\alpha-1}\log\left(\frac{p_i}{q_i}\right)
\end{aligned}
$$

which gives the desired lower bound for $D_\alpha(\mathbf{p},\mathbf{q})$.

The proof is complete. □

Now, by using the previous theorems, some inequalities of Rényi entropy are obtained. Denote $\mathbf{\frac{1}{n}} := \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ be the discrete uniform distribution.

Corollary 3.3 *Let $2 \le k \le n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$ and $\mathbf{p} := (p_1, \ldots, p_n)$ be positive probability distributions.*

(a) If $0 \le \alpha \le \beta$, α, $\beta \ne 1$, and the base of log *is greater than* 1, *then*

$$H_\alpha(\mathbf{p}) \ge \frac{1}{1-\beta} \log \left(\sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}^{\alpha}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right)^{\frac{\beta-1}{\alpha-1}} \right) \ge H_\beta(\mathbf{p}).$$

The reverse inequalities hold if the base of log *is between* 0 *and* 1.

(b) If $1 < \beta$, and the base of log *is greater than* 1, *then*

$$\begin{aligned} H(\mathbf{p}) &= -\sum_{i=1}^{n} p_i \log(p_i) \\ &\ge \log(n) + \frac{1}{1-\beta} \log \left(\sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \exp \left(\frac{(\beta-1) \sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \log(n p_{i+j})}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right) \right) \\ &\ge H_\beta(\mathbf{p}), \end{aligned}$$

where the base of exp *is the same as the base of* log.

The reverse inequalities hold if the base of log *is between* 0 *and* 1.

(c) If $0 \le \alpha < 1$, and the base of log *is greater than* 1, *then*

$$H_\alpha(\mathbf{p}) \ge \frac{1}{1-\alpha} \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \log \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}^{\alpha}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right) \ge H(\mathbf{p})$$

The reverse inequalities hold if the base of log *is between* 0 *and* 1.

In all these inequalities $i+j$ means $i+j-n$ in case of $i+j>n$.

Proof. If $\mathbf{q} = \mathbf{\frac{1}{n}}$, then

$$D_\alpha\left(\mathbf{p}, \mathbf{\frac{1}{n}}\right) = \frac{1}{\alpha-1} \log \left(\sum_{i=1}^{n} n^{\alpha-1} p_i^{\alpha} \right) = \log(n) + \frac{1}{\alpha-1} \log \left(\sum_{i=1}^{n} p_i^{\alpha} \right),$$

and therefore

$$H_\alpha(\mathbf{p}) = \log(n) - D_\alpha\left(\mathbf{p}, \mathbf{\frac{1}{n}}\right). \tag{3.12}$$

(a) It follows from Theorem 3.2 and (3.12) that

$$
\begin{aligned}
H_\alpha(\mathbf{p}) &= \log(n) - D_\alpha\left(\mathbf{p}, \frac{\mathbf{1}}{\mathbf{n}}\right) \\
&\geq \log(n) - \frac{1}{\beta-1}\log\left(n^{\beta-1}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}^{\alpha}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)^{\frac{\beta-1}{\alpha-1}}\right) \\
&\geq \log(n) - D_\beta\left(\mathbf{p}, \frac{\mathbf{1}}{\mathbf{n}}\right) = H_\beta(\mathbf{p}).
\end{aligned}
$$

(b) and (c) can be proved similarly.

The proof is complete. □

Corollary 3.4 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1, \ldots, \lambda_k)$ and $\mathbf{p} := (p_1, \ldots, p_n)$ be positive probability distributions.*

If either $0 \leq \alpha < 1$ and the base of log *is greater than* 1*, or $1 < \alpha$ and the base of* log *is between* 0 *and* 1*, then*

$$
\begin{aligned}
&-\frac{1}{\sum\limits_{i=1}^{n}p_i^{\alpha}}\sum_{i=1}^{n}p_i^{\alpha}\log(p_i) \\
&\quad\geq \log(n) - \frac{1}{(\alpha-1)\sum\limits_{i=1}^{n}p_i^{\alpha}} \times \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}^{\alpha}\right)\log\left(n^{\alpha-1}\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}^{\alpha}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \geq H_\alpha(\mathbf{p}) \\
&\quad\geq \frac{1}{1-\alpha}\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\log\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}^{\alpha}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \geq H(\mathbf{p})
\end{aligned}
$$

If either $0 \leq \alpha < 1$ and the base of log *is between* 0 *and* 1*, or $1 < \alpha$ and the base of* log *is greater than* 1*, then the reverse inequalities holds.*

In all these inequalities $i+j$ means $i+j-n$ in case of $i+j > n$.

Proof. We can prove as Corollary 3.3, by using Theorem 3.3. □

We illustrate our results by using Zipf–Mandelbrot law.

3.1.3 Inequalities by using the Zipf-Mandelbrot law

Corollary 3.5 *Let* $\mathbf{p}$ *be the Zipf-Mandelbrot law as in Definition 3.5, let* $2 \leq k \leq N$ *be integers, and let* $\lambda := (\lambda_1, \ldots, \lambda_k)$ *be a probability distribution. By applying Corollary 3.3 (c), we have:*

If $0 \leq \alpha < 1$*, and the base of* log *is greater than* 1*, then*

$$
\begin{aligned}
H_\alpha(\mathbf{p}) &= \frac{1}{1-\alpha} \log \left(\frac{1}{H_{N,q,s}^{\alpha}} \sum_{i=1}^{N} \frac{1}{(i+q)^{\alpha s}} \right) \\
&\geq \frac{1}{1-\alpha} \sum_{i=1}^{n} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+q)^s H_{N,q,s}} \right) \log \left(\frac{1}{H_{N,q,s}^{\alpha-1}} \frac{\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+q)^{\alpha s}}}{\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+q)^{s}}} \right) \\
&\geq \frac{s}{H_{N,q,s}} \sum_{i=1}^{N} \frac{\log(i+q)}{(i+q)^s} + \log(H_{N,q,s}) = H(\mathbf{p})
\end{aligned}
$$

The reverse inequalities hold if the base of log *is between* 0 *and* 1*.*

In all these inequalities $i+j$ *means* $i+j-n$ *in case of* $i+j>n$*.*

Corollary 3.6 *Let* $\mathbf{p}_1$ *and* $\mathbf{p}_2$ *be the Zipf-Mandelbrot law with parameters* $N \in \{1,2,\ldots\}$*,* $q_1,\ q_2 \in [0,\infty[$ *and* $s_1,\ s_2 > 0$*, respectively, let* $2 \leq k \leq N$ *be integers, and let* $\lambda := (\lambda_1, \ldots, \lambda_k)$ *be a probability distribution. By applying Corollary 3.2 (b), we have:*

If the base of log *is greater than* 1*, then*

$$
\begin{aligned}
D(\mathbf{p}_1 \| \mathbf{p}_2) &= \sum_{i=1}^{N} \frac{1}{(i+q_1)^{s_1} H_{N,q_1,s_1}} \log \left(\frac{(i+q_2)^{s_2} H_{N,q_2,s_2}}{(i+q_1)^{s_1} H_{N,q_1,s_1}} \right) \\
&\geq \sum_{i=1}^{N} \left(\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+q_1)^{s_1} H_{N,q_1,s_1}} \right) \log \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+q_1)^{s_1} H_{N,q_1,s_1}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+q_2)^{s_2} H_{N,q_2,s_2}}} \right) \geq 0.
\end{aligned}
\tag{3.13}
$$

If the base of log *is between* 0 *and* 1*, then inequality signs in (3.13) are reversed.*

In all these inequalities $i+j$ *means* $i+j-n$ *in case of* $i+j>n$*.*

3.2 A refinement and an exact equality condition for the basic inequality of f-divergences

Measures of dissimilarity between probability measures play important role in probability theory, especially in information theory and in mathematical statistics. Many divergence measures for this purpose have been introduced and studied (see for example Vajda [91]). Among them f-divergences were introduced by Csiszár [19]-[18] and independently by Ali and Silvey [2]. Remarkable divergences can be found among f-divergences, such as the information divergence, the Pearson or χ^2-divergence, the Hellinger distance and total variational distance. There are a lot of papers dealing with f-divergence inequalities (see Dragomir [25], Dembo, Cover, and Thomas [21] and Sason and Verdú [88]). These inequalities are very useful and applicable in information theory.

One of the basic inequalities is (see Liese and Vajda [64])

$$D_f(P,Q) \geq f(1).$$

In this section we give a refinement and a precise equality condition for this inequality. Some applications for discrete distributions, for the Shannon entropy, and some examples are given.

3.2.1 Construction of the equality conditions and related results of classical integral Jensen's inequality

The classical Jensen's inequality is well known (see [34]).

Theorem 3.4 *Let g be an integrable function on a probability space $(Y, \mathscr{B}, \nu)$ taking values in an interval $I \subset \mathbb{R}$. Then $\int_Y g d\nu$ lies in I. If f is a convex function on I such that $f \circ g$ is ν-integrable, then*

$$f\left(\int_Y g d\nu\right) \leq \int_Y f \circ g d\nu. \tag{3.14}$$

The following approach to give a necessary and sufficient condition for equality in this inequality may be new. First, we introduce the next definition.

Definition 3.7 *Let $(Y, \mathscr{B}, \nu)$ be a probability space, and let g be a real measurable function defined almost everywhere on Y. We denote by $essint_\nu(g)$ the smallest interval in $\mathbb{R}$ for which*

$$\nu(g \in essint_\nu(g)) = 1.$$

Remark 3.3 (a) Obviously, the endpoints of $\mathrm{essint}_\nu(g)$ are the essential infimum ($\mathrm{essinf}_\nu(g)$) and the essential supremum of g, and either of them belong to $\mathrm{essint}_\nu(g)$ exactly if g takes this value with positive probability.

(b) It is easy to see that either $\mathrm{essint}_\nu(g) = \left\{ \int_Y g d\nu \right\}$ (in this case g is constant ν-a.e.) or $\int_Y g d\nu$ is an inner point of $\mathrm{essint}_\nu(g)$.

(c) The interval $\mathrm{essinf}_\nu(g)$ is connected with the essential range of g, but not the same set (for example, the essential range of g is always closed, and not an interval in general).

Lemma 3.1 *Assume the conditions of Theorem 3.4 are satisfied. Equality holds in (3.14) if and only if f is affine on $\mathrm{essint}_\nu(g)$.*

Proof. It is easy to see that the condition is sufficient for equality in (3.14).

Conversely, if $\mathrm{essint}_\nu(g)$ contains only one point, then it is trivial, so we can assume that $m := \int_Y g d\nu$ is an inner point of $\mathrm{essint}_\nu(g)$. Let

$$l : \mathbb{R} \to \mathbb{R}, \quad l(t) = f'_+(m)(t-m) + f(m).$$

If f is not affine on $\mathrm{essint}_\nu(g)$, then by the convexity of f, there is a point $t_1 \in \mathrm{essint}_\nu(g)$ such that $f(t_1) > l(t_1)$. Suppose $t_1 > m$ (the case $t_1 < m$ can be handled similarly). Since f is convex, $f(t) \geq l(t)$ $(t \in I)$ and $f(t) > l(t)$ $(t \in I,\ t \geq t_1)$. It follows by using $\nu(g > t_1) > 0$, that

$$\int_Y f \circ g d\nu = \int_{(g<t_1)} f \circ g d\nu + \int_{(g \geq t_1)} f \circ g d\nu \geq \int_{(g<t_1)} l \circ g d\nu + \int_{(g \geq t_1)} f \circ g d\nu > \int_Y l \circ g d\nu = f(m),$$

which is a contradiction.

The proof is complete. □

The next refinement of the Jensen's inequality can be found in Horváth [39].

Theorem 3.5 *Let $I \subset \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$ be a convex function. Let $(Y, \mathscr{B}, \nu)$ be a probability space, and let $g : Y \to I$ be a ν-integrable function such that $f \circ g$ is also ν-integrable. Suppose that $\alpha_1, \ldots, \alpha_n$ are nonnegative numbers with $\sum_{i=1}^{n} \alpha_i = 1$. Then*

(a)

$$f\left(\int_Y g d\nu\right) \leq \int_{Y^n} f\left(\sum_{i=1}^{n} \alpha_i g(x_i)\right) d\nu^n(x_1, \ldots, x_n) \leq \int_Y f \circ g d\nu.$$

(b)

$$\int_{Y^{n+1}} f\left(\frac{1}{n+1}\sum_{i=1}^{n+1} g(x_i)\right) d\nu^{n+1}(x_1, \ldots, x_{n+1})$$

$$\leq \int_{Y^n} f\left(\frac{1}{n}\sum_{i=1}^{n} g(x_i)\right) d\nu^n(x_1, \ldots, x_n) \leq \int_{Y^n} f\left(\sum_{i=1}^{n} \alpha_i g(x_i)\right) d\nu^n(x_1, \ldots, x_n).$$

By analyzing the proof of the previous result, it can be seen that the hypothesis "$f \circ g$ is ν-integrable" can be weaken.

Theorem 3.6 *Let $I \subset \mathbb{R}$ be an interval, and let $f : I \to \mathbb{R}$ be a convex function. Let $(Y, \mathscr{B}, \nu)$ be a probability space, and let $g : Y \to I$ be a ν-integrable function such that the integral $\int_Y f \circ g d\nu$ exists in $]-\infty, \infty]$. Suppose that $\alpha_1, \ldots, \alpha_n$ are nonnegative numbers with $\sum_{i=1}^{n} \alpha_i = 1$. Then the assertions of Theorem 3.5 remain true.*

We assume throughout that the probability measures P and Q are defined on a fixed measurable space $(X, \mathscr{A})$. It is also assumed that P and Q are absolutely continuous with respect to a σ-finite measure μ on $\mathscr{A}$. The densities (or Radon-Nikodym derivatives) of P and Q with respect to μ are denoted by p and q, respectively. These densities are μ-almost everywhere uniquely determined.

Let

$$F := \{f :]0, \infty[\to \mathbb{R} \mid f \text{ is convex}\},$$

and define for every $f \in F$ the function

$$f^* :]0, \infty[\to \mathbb{R}, \quad f^*(t) := tf\left(\frac{1}{t}\right).$$

If $f \in F$, then either f is monotonic or there exists a point $t_0 \in]0, \infty[$ such that f is decreasing on $]0, t_0[$. This implies that the limit

$$\lim_{t \to 0+} f(t)$$

exists in $]-\infty, \infty]$, and

$$f(0) := \lim_{t \to 0+} f(t)$$

extends f into a convex function on $[0, \infty[$. The extended function is continuous and has finite left and right derivatives at each point of $]0, \infty[$.

It is well known that for every $f \in F$ the function f^* also belongs to F, and therefore

$$f^*(0) := \lim_{t \to 0+} f^*(t) = \lim_{u \to \infty} \frac{f(u)}{u}.$$

We need the following simple property of functions belonging to F.

Lemma 3.2 *If $f \in F$, then $f^*(0) \geq f'_+(1)$. This inequality becomes an equality if and only if*

$$f(t) = f'_+(1)(t-1) + f(1), \quad t \geq 1. \tag{3.15}$$

Proof. Since f is convex,

$$f(t) \geq f'_+(1)(t-1) + f(1), \quad t \geq 1,$$

and therefore

$$f^*(0) = \lim_{t\to\infty} \frac{f(t)}{t} \geq f'_+(1).$$

If (3.15) is satisfied, then obviously $f^*(0) = f'_+(1)$.

If there exists $t_1 > 1$ such that $f'_+(t_1) > f'_+(1)$, then by the convexity of f,

$$f(t) \geq f'_+(t_1)(t-t_1) + f(t_1), \quad t \geq t_1,$$

and hence $f^*(0) > f'_+(1)$. It follows that $f^*(0) = f'_+(1)$ implies

$$f'_+(t) = f'_+(1), \quad t \geq t_1,$$

and this gives (3.15) (see [28] 1.6.2 Corollary 2).

The proof is complete. □

The next result prepares the notion of f-divergence of probability measures.

Lemma 3.3 *For every $f \in F$ the integral*

$$\int\limits_{(q>0)} q(\omega) f\left(\frac{p(\omega)}{q(\omega)}\right) d\mu(\omega)$$

exists and it belongs to the interval $]-\infty, \infty]$.

Proof. Since f is convex,

$$f(t) \geq f'_+(1)(t-1) + f(1), \quad t \geq 0.$$

This implies that for all $\omega \in (q>0)$

$$q(\omega) f\left(\frac{p(\omega)}{q(\omega)}\right) \geq h(\omega) := f'_+(1)(p(\omega) - q(\omega)) + f(1) q(\omega). \tag{3.16}$$

Elementary considerations show that the function h is μ-integrable over $(q>0)$, and this gives the result by (3.16).

The proof is complete. □

Now we introduce the notion of f-divergence.

Definition 3.8 *For every $f \in F$ we define the f-divergence of P and Q by*

$$D_f(P,Q) := \int\limits_X q(\omega) f\left(\frac{p(\omega)}{q(\omega)}\right) d\mu(\omega),$$

where the following conventions are used

$$0f\left(\frac{x}{0}\right) := xf^*(0) \text{ if } x>0, \quad 0f\left(\frac{0}{0}\right) = 0f^*(0) := 0. \tag{3.17}$$

Remark 3.4 (a) For every $f \in F$ the perspective $\hat{f} :]0,\infty[\times]0,\infty[\to \mathbb{R}$ of f is defined by

$$\hat{f}(x,y) := yf\left(\frac{x}{y}\right).$$

Then (see [86]) $\hat{f}$ is also a convex function. Vajda [91] proved that (3.17) is the unique rule leading to convex and lower semicontinuous extension of $\hat{f}$ to the set

$$\left\{(x,y) \in \mathbb{R}^2 \mid x,y \geq 0\right\}.$$

(b) Since $f^*(0) \in]-\infty,\infty]$, Lemma 3.3 shows that $D_f(P,Q)$ exists in $]-\infty,\infty]$ and

$$D_f(P,Q) = \int_{(q>0)} f\left(\frac{p(\omega)}{q(\omega)}\right) dQ(\omega) + f^*(0)P(q=0). \tag{3.18}$$

It follows that if P is absolutely continuous with respect to Q, then

$$D_f(P,Q) = \int_{(q>0)} f\left(\frac{p(\omega)}{q(\omega)}\right) dQ(\omega).$$

Various divergences in information theory and statistics are special cases of the f-divergence. We illustrate this by some examples.

(a) By choosing $f :]0,\infty[\to \mathbb{R}$, $f(t) = t\ln(t)$ in (3.18), the information divergence is obtained

$$I(P,Q) = \int_{(q>0)} p(\omega)\ln\left(\frac{p(\omega)}{q(\omega)}\right) d\mu(\omega) + \infty P(q=0). \tag{3.19}$$

(b) By choosing $f :]0,\infty[\to \mathbb{R}$, $f(t) = (t-1)^2$ in (3.18), the Pearson or χ^2-divergence is obtained

$$\chi^2(P,Q) = \int_{(q>0)} \frac{(p(\omega)-q(\omega))^2}{q(\omega)} d\mu(\omega) + \infty P(q=0). \tag{3.20}$$

(c) By choosing $f :]0,\infty[\to \mathbb{R}$, $f(t) = \left(\sqrt{t}-1\right)^2$ in (3.18), the Hellinger distance is obtained

$$H^2(P,Q) = \int_X \left(\sqrt{p(\omega)} - \sqrt{q(\omega)}\right)^2 d\mu(\omega). \tag{3.21}$$

(d) By choosing $f :]0,\infty[\to \mathbb{R}$, $f(t) = |t-1|$ in (3.18), the total variational distance is obtained

$$V(P,Q) = \int_X |p(\omega) - q(\omega)| \mu(\omega). \tag{3.22}$$

We need the following lemma.

Lemma 3.4 *Let* $t_0 := P(q > 0)$.

(a) For every $\varepsilon > 0$

$$Q\left(\frac{p}{q} < t_0 + \varepsilon,\ q > 0\right) > 0.$$

(b)

$$essinf_Q\left(\frac{p}{q}\right) \leq t_0$$

Proof. (a) Obviously,

$$Q\left(\frac{p}{q} < t_0 + \varepsilon,\ q > 0\right) = 1 - Q\left(\frac{p}{q} \geq t_0 + \varepsilon,\ q > 0\right).$$

The result follows from this, since

$$Q\left(\frac{p}{q} \geq t_0 + \varepsilon,\ q > 0\right) = \int_X q 1_{\left(\frac{p}{q} \geq t_0 + \varepsilon,\ q > 0\right)} d\mu \leq \int_{(q>0)} \frac{1}{t_0 + \varepsilon} p d\mu = \frac{t_0}{t_0 + \varepsilon} < 1.$$

(b) It comes from (a).

The proof is complete. □

The following result contains a key property of f-divergences. We give a simple proof which emphasizes the importance of the convexity of f, and give an exact equality condition.

Theorem 3.7 *(a) For every* $f \in F$

$$D_f(P, Q) \geq f(1). \tag{3.23}$$

(b) Assume $P(q = 0) = 0$. *Then equality holds in (3.23) if and only if* f *is affine on* $essint_Q\left(\frac{p}{q}\right)$.

(c) Assume $P(q = 0) > 0$. *Then equality holds in (3.23) if and only if* f *is affine on* $essint_Q\left(\frac{p}{q}\right) \cup [1, \infty[$.

Proof. (a) If $D_f(P, Q) = \infty$, then (3.23) is obvious.

If $D_f(P, Q) \in \mathbb{R}$, then the integral

$$\int_{(q>0)} f\left(\frac{p(\omega)}{q(\omega)}\right) dQ(\omega) \tag{3.24}$$

is finite, and therefore either $Q(p = 0) = 0$ or $Q(p = 0) > 0$ and $f(0)$ is finite. It follows that Jensen's inequality can be applied to this integral, and we have

$$D_f(P, Q) \geq f\left(\int_{(q>0)} p d\mu\right) + f^*(0) P(q = 0) \tag{3.25}$$

$$= f(P(q > 0)) + f^*(0) P(q = 0). \tag{3.26}$$

Let $t_0 := P(q > 0)$. By using Lemma 3.2, $t_0 \in [0,1]$, and the convexity of f, it follows from (3.26) that

$$D_f(P,Q) \geq f(t_0) + f'_+(1)(1-t_0) \tag{3.27}$$
$$\geq f(1) + f'_+(1)(t_0-1) + f'_+(1)(1-t_0) = f(1). \tag{3.28}$$

(b) If $D_f(P,Q) = f(1)$, then $D_f(P,Q)$ is finite.

Assume $P(q=0) = 0$. Then by (3.25) and (3.26), $D_f(P,Q) = f(1)$ is satisfied if and only if equality holds in the Jensen's inequality. Lemma 3.1 shows that this happens exactly if f is affine on $\operatorname{essint}_Q\left(\frac{p}{q}\right)$.

(c) Assume $P(q=0) > 0$. Then (3.25), (3.26), (3.27) and (3.28) yield that there must be equality in the Jensen's inequality, $f^*(0) = f'_+(1)$, and

$$f(t_0) = f(1) + f'_+(1)(t_0-1). \tag{3.29}$$

By Lemma 3.1 and Lemma 3.2, the first two equality conditions are satisfied exactly if f is affine on $\operatorname{essint}_Q\left(\frac{p}{q}\right) \cup [1,\infty[$.

Now assume that f is affine on $\operatorname{essint}_Q\left(\frac{p}{q}\right) \cup [1,\infty[$. In case of $t_0 > 0$, Lemma 3.4 (b) and the continuity of f at t_0 show that (3.29) also holds. In case of $t_0 = 0$, it is easy to see that $Q\left(\frac{p}{q} = 0\right) = 1$, and hence $0 \in \operatorname{essint}_Q\left(\frac{p}{q}\right)$ which implies (3.29) too.

The proof is complete. □

Remark 3.5 (a) Consider the subclass $F_1 \subset F$ such that $f \in F_1$ satisfies $f(1) = 0$. In this case inequality (3.23) has the usual form

$$D_f(P,Q) \geq 0.$$

(b) The usual equality condition is the next (see [64]): if f is strictly convex at 1, then $D_f(P,Q) = f(1)$ holds if and only if $P = Q$. Theorem 3.7 (b) and (c) give more precise conditions.

3.2.2 Refinements of basic inequality in f-divergences and related results

Suppose that $\alpha_1, \ldots, \alpha_n$ are nonnegative numbers with $\sum_{i=1}^{n} \alpha_i = 1$. Let

$$\mathscr{A}^n := \mathscr{A} \otimes \ldots \otimes \mathscr{A}, \quad \text{with } n \text{ factors},$$

and define the probability measures Q^n and R on $\mathscr{A}^n$ by

$$Q^n := Q \otimes \ldots \otimes Q, \quad \text{with } n \text{ factors},$$

and

$$R_\alpha := \sum_{i=1}^{n} \alpha_i Q \otimes \ldots \otimes Q \otimes \overset{i}{\check{P}} \otimes Q \otimes \ldots \otimes Q.$$

In case of $\alpha_i = \frac{1}{n}$ $(i = 1, \dots, n)$ the probability measure R_α will be denoted by R_n.

These measures are absolutely continuous with respect to μ^n on $\mathscr{A}^n$. The densities of R and Q^n with respect to μ^n are

$$\bigotimes_{i=1}^{n} q : X^n \to \mathbb{R}, \quad (\omega_1, \dots, \omega_n) \to \prod_{i=1}^{n} q(\omega_i),$$

and

$$(\omega_1, \dots, \omega_n) \to \sum_{i=1}^{n} \alpha_i q(\omega_1) \dots \overset{i}{\check{p}}(\omega_i) \dots q(\omega_n), \quad (\omega_1, \dots, \omega_n) \in X^n,$$

respectively.

It is easy to calculate that

$$R_\alpha \left(\bigotimes_{i=1}^{n} q = 0 \right) = 1 - R_\alpha \left(\bigotimes_{i=1}^{n} q > 0 \right) = 1 - R_\alpha \left((q > 0)^n \right)$$
$$= 1 - \sum_{i=1}^{n} \alpha_i Q(q > 0)^{n-1} P(q > 0) = 1 - P(q > 0) = P(q = 0).$$

It follows that for every $f \in F$

$$D_f(R_\alpha, Q^n) = \int_{(q>0)^n} f \left(\frac{\sum_{i=1}^{n} \alpha_i q(\omega_1) \dots p(\omega_i) \dots q(\omega_n)}{\prod_{i=1}^{n} q(\omega_i)} \right) dQ^n(\omega_1, \dots, \omega_n)$$
$$+ f^*(0) R_\alpha \left(\bigotimes_{i=1}^{n} q = 0 \right)$$
$$= \int_{(q>0)^n} f \left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i)}{q(\omega_i)} \right) dQ^n(\omega_1, \dots, \omega_n) + f^*(0) P(q = 0) \tag{3.30}$$
$$= \int_{(q>0)^n} \prod_{i=1}^{n} q(\omega_i) f \left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i)}{q(\omega_i)} \right) d\mu^n(\omega_1, \dots, \omega_n) + f^*(0) P(q = 0).$$

By applying Theorem 3.5, we obtain some refinements of the basic inequality 3.23.

Theorem 3.8 *Suppose that $\alpha_1, \dots, \alpha_n$ are nonnegative numbers with $\sum_{i=1}^{n} \alpha_i = 1$. If $f \in F$, then*

(a)

$$D_f(P, Q) \geq D_f(R_\alpha, Q^n) \geq D_f(R_n, Q^n) \geq f(1). \tag{3.31}$$

(b)

$$D_f(P, Q) = D_f\left(R_1, Q^1\right)$$
$$\geq \dots \geq D_f(R_m, Q^m) \geq D_f\left(R_{m+1}, Q^{m+1}\right) \geq \dots \geq f(1), \quad m \geq 1.$$

Proof. (a) The third inequality in (3.31) comes from Theorem 3.7.

So it remains to prove the first two inequalities in (3.31). By (3.18) and (3.30), it is enough to show that

$$\int\limits_{(q>0)} f\left(\frac{p(\omega)}{q(\omega)}\right) dQ(\omega) \geq \int\limits_{(q>0)^n} f\left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i)}{q(\omega_i)}\right) dQ^n(\omega_1,\ldots,\omega_n) \tag{3.32}$$

$$\geq \int\limits_{(q>0)^n} f\left(\frac{1}{n}\sum_{i=1}^{n} \frac{p(\omega_i)}{q(\omega_i)}\right) dQ^n(\omega_1,\ldots,\omega_n),$$

which is an immediate consequence of Theorem 3.6.

(b) We can proceed similarly as in (a).

The proof is complete. □

By considering the special f-divergences (3.19-3.22), we have after each other

(a) the information divergence

$$I(R_\alpha, Q^n) = \infty P(q=0)$$

$$+ \int\limits_{(q>0)^n} \sum_{i=1}^{n} \left(\alpha_i p(\omega_i) \prod_{\substack{j=1 \\ j\neq i}}^{n} q(\omega_j)\right) \ln\left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i)}{q(\omega_i)}\right) d\mu^n(\omega_1,\ldots,\omega_n),$$

(b) the Pearson divergence

$$\chi^2(R_\alpha, Q^n) = \int\limits_{(q>0)^n} \prod_{i=1}^{n} q(\omega_i) \left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i) - q(\omega_i)}{q(\omega_i)}\right)^2 d\mu^n(\omega_1,\ldots,\omega_n) + \infty P(q=0),$$

(c) the Hellinger distance

$$H^2(R_\alpha, Q^n) = \int\limits_{(q>0)^n} \prod_{i=1}^{n} q(\omega_i) \left(\left(\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i)}{q(\omega_i)}\right)^{1/2} - 1\right)^2 d\mu^n(\omega_1,\ldots,\omega_n),$$

(d) the total variational distance

$$V(R_\alpha, Q^n) = \int\limits_{(q>0)^n} \prod_{i=1}^{n} q(\omega_i) \left|\sum_{i=1}^{n} \alpha_i \frac{p(\omega_i) - q(\omega_i)}{q(\omega_i)}\right| d\mu^n(\omega_1,\ldots,\omega_n).$$

Now, we consider the special case, important in many applications, in which P and Q are discrete distributions.

Denote T either the set $\{1,\ldots,k\}$ with a fixed positive integer k, or the set $\{1,2,\ldots\}$. We say that P and Q are derived from the positive probability distributions $p := (p_i)_{i\in T}$ and $q := (q_i)_{i\in T}$, respectively, if p_i, $q_i > 0$ $(i \in T)$, and $\sum\limits_{i\in T} p_i = \sum\limits_{i\in T} q_i = 1$. In this case $X = T$, $\mathscr{A}$ is the power set of T, and μ is the counting measure on $\mathscr{A}$.

Corollary 3.7 *Suppose that $\alpha_1,\ldots,\alpha_n$ are nonnegative numbers with $\sum_{i=1}^{n}\alpha_i = 1$. Suppose also that P and Q are derived from the positive probability distributions $(p_i)_{i\in T}$ and $(q_i)_{i\in T}$, respectively. If $f \in F$, then*
(a)

$$D_f(P,Q) = \sum_{i\in T} q_i f\left(\frac{p_i}{q_i}\right) \geq \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} f\left(\sum_{j=1}^{n}\alpha_j \frac{p_{i_j}}{q_{i_j}}\right)$$
$$\geq \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} f\left(\frac{1}{n}\sum_{j=1}^{n}\frac{p_{i_j}}{q_{i_j}}\right) \geq f(1).$$

(b)

$$D_f(P,Q) \geq \ldots \geq \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} f\left(\frac{1}{n}\sum_{j=1}^{n}\frac{p_{i_j}}{q_{i_j}}\right)$$
$$\geq \sum_{(i_1,\ldots,i_{n+1})\in T^{n+1}} \prod_{j=1}^{n+1} q_{i_j} f\left(\frac{1}{n+1}\sum_{j=1}^{n+1}\frac{p_{i_j}}{q_{i_j}}\right) \geq \ldots \geq f(1), \quad n \geq 1.$$

Proof. This comes from Theorem 3.8 immediately. □

Finally, we give an example to illustrate the previous result. We consider only Corollary 3.7 (a).

Example 3.1 (a) By choosing $f:]0,\infty[\to \mathbb{R}$, $f(x) = -\ln(x)$ and $p_i = \frac{1}{k}$ $(i = 1,\ldots,k)$ in the previous corollary (in this case $T = \{1,\ldots,k\}$), we have

$$D_f(P,Q) = -\sum_{i=1}^{k} q_i \ln\left(\frac{1}{kq_i}\right) = \ln(k) + \sum_{i=1}^{k} q_i \ln(q_i)$$
$$\geq -\sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\frac{1}{k}\sum_{j=1}^{n}\frac{\alpha_j}{q_{i_j}}\right) = \ln(k) - \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\sum_{j=1}^{n}\frac{\alpha_j}{q_{i_j}}\right)$$
$$\geq -\sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\frac{1}{kn}\sum_{j=1}^{n}\frac{1}{q_{i_j}}\right) = \ln(kn) - \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\sum_{j=1}^{n}\frac{1}{q_{i_j}}\right) \geq 0.$$

It can be obtained from this some refinements of the classical upper estimation for the Shannon entropy

$$H(Q) := -\sum_{i=1}^{k} q_i \ln(q_i) \leq \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\sum_{j=1}^{n}\frac{\alpha_j}{q_{i_j}}\right)$$
$$\leq -\ln(n) + \sum_{(i_1,\ldots,i_n)\in T^n} \prod_{j=1}^{n} q_{i_j} \ln\left(\sum_{j=1}^{n}\frac{1}{q_{i_j}}\right) \leq \ln(k).$$

(b) If $f :]0,\infty[\to \mathbb{R}$, $f(x) = x\ln(x)$ in the previous corollary, then we have the following estimations for the information or Kullback–Leibler divergence:

$$I(P,Q) = \sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{q_i}\right) \geq \sum_{(i_1,\ldots,i_n)\in T^n} \left(\sum_{j=1}^{n} \alpha_j p_{i_j} \prod_{\substack{l=1\\ l\neq j}}^{n} q_{i_l}\right) \ln\left(\sum_{j=1}^{n} \alpha_j \frac{p_{i_j}}{q_{i_j}}\right)$$

$$\geq \frac{1}{n} \sum_{(i_1,\ldots,i_n)\in T^n} \left(\sum_{j=1}^{n} p_{i_j} \prod_{\substack{l=1\\ l\neq j}}^{n} q_{i_l}\right) \ln\left(\frac{1}{n}\sum_{j=1}^{n} \frac{p_{i_j}}{q_{i_j}}\right) \geq 0. \tag{3.33}$$

(c) The Zipf-Mandelbrot law (see Mandelbrot [65] and Zipf [95]) is a discrete probability distribution depends on three parameters $N \in \{1,2,\ldots\}$, $q \in [0,\infty[$ and $s > 0$, and it is defined by

$$f(i;N,q,s) := \frac{1}{(i+q)^s H_{N,q,s}}, \quad i = 1,\ldots,N,$$

where

$$H_{N,q,s} := \sum_{k=1}^{N} \frac{1}{(k+q)^s}.$$

Let P and Q be the Zipf-Mandelbrot law with parameters $N \in \{1,2,\ldots\}$, $q_1, q_2 \in [0,\infty[$ and $s_1, s_2 > 0$, respectively, and let $2 \leq k \leq N$ be an integer. It follows from the first part of (3.33) with $T = \{1,\ldots,N\}$ that

$$I(P,Q) = \sum_{i=1}^{N} \frac{1}{(i+q_1)^{s_1} H_{N,q_1,s_1}} \log\left(\frac{(i+q_2)^{s_2} H_{N,q_2,s_2}}{(i+q_1)^{s_1} H_{N,q_1,s_1}}\right)$$

$$\geq \sum_{(i_1,\ldots,i_N)\in T^n} \left(\sum_{j=1}^{n} \alpha_j \frac{1}{(i_j+q_1)^{s_1} H_{N,q_1,s_1}} \prod_{\substack{l=1\\ l\neq j}}^{n} \frac{1}{(i_l+q_2)^{s_2} H_{N,q_2,s_2}}\right)$$

$$\times \ln\left(\sum_{j=1}^{n} \alpha_j \frac{(i_j+q_2)^{s_2} H_{N,q_2,s_2}}{(i_j+q_1)^{s_1} H_{N,q_1,s_1}}\right) \geq 0.$$

This is another type of refinement for $I(P,Q)$ than it is given in [38].

Chapter 4

Cyclic Refinement of Beck's Inequalities

In the present Chapter 4, we refine the discrete Jensen's inequality for vectors by the idea recently given in [13]. As a consequence, we are able to refine the inequality of E. Beck [9] with the help of cyclic generalized mixed symmetric means. This leads to the refinements of the classical Hölder's and Minkowski's inequalities.

4.1 Introduction and preliminary results

In this subsection we first briefly summarise some results corresponding to Beck's inequalities (see [9]). We depend upon the papers [44]and [45].

Let $I \subset \mathbb{R}$ be an interval, let $h : I \to \mathbb{R}$ be a continuous and strictly monotone function, let $\mathbf{a} = (a_1, \ldots, a_n) \in I^n$, and let $\mathbf{p} = (p_1, \ldots, p_n)$ be a nonnegative n-tuple such that $\sum\limits_{i=1}^{n} p_i = 1$. The quasi-arithmetic h-mean of $\mathbf{a}$ with weights $\mathbf{p}$ is defined by

$$h_n(\mathbf{a};\mathbf{p}) = h_n(a_i; 1 \leq i \leq n; \mathbf{p}) = h(\mathbf{a};\mathbf{p};n) := h^{-1}\left(\sum_{i=1}^{n} p_i h(a_i)\right).$$

If $p_i = \frac{1}{n}$ $(1 \leq i \leq n)$, then $\mathbf{p}$ will be ignored from the previous notations.

First, we extend Beck's results (see [9]). The following hypothesis will be used:

(A$_1$) Let $L_t : I_t \to \mathbb{R}$ $(t = 1,...,m)$ and $N : I_N \to \mathbb{R}$ be continuous and strictly monotone functions whose domains are intervals in $\mathbb{R}$, and let $f : I_1 \times ... \times I_m \to I_N$ be a continuous function. Let $\mathbf{x}^{(1)},...,\mathbf{x}^{(m)} \in \mathbb{R}^n$ $(n \geq 2)$ such that $\mathbf{x}^{(t)} \in I_t^n$ for each $t = 1,\ldots,m$, and let $\mathbf{p} = (p_1,...,p_n)$ be a nonnegative n-tuple such that $\sum_{i=1}^n p_i = 1$. The next result is a simple consequence of the discrete Jensen's inequality (see Theorem A).

Theorem 4.1 *Assume (A$_1$). If N is an increasing function, then the inequality*

$$f\left(L_1(\mathbf{x}^{(1)};\mathbf{p};n),...,L_m(\mathbf{x}^{(m)};\mathbf{p};n)\right) \geq N^{-1}\left(\sum_{i=1}^n p_i N(f(x_i^{(1)},...,x_i^{(m)}))\right), \tag{4.1}$$

holds for all possible $\mathbf{x}^{(t)}$ $(t = 1,\ldots,m)$ and $\mathbf{p}$, if and only if the function H defined on $I_1 \times ... \times I_m$ by

$$H(t_1,...,t_m) := N\left(f\left(L_1^{-1}(t_1),...,L_m^{-1}(t_m)\right)\right) \tag{4.2}$$

is concave. The inequality in (4.1) is reversed for all possible $\mathbf{x}^{(t)}$ $(t = 1,\ldots,m)$ and $\mathbf{p}$, if and only if H is convex.

Proof. We replace the convex function f by $-H$ or H, and x_i by $L_t(x_i^{(t)})$ in (1.1) and then applying the increasing function N^{-1} we get the required results. □

Beck's original result was the special case of Theorem 4.1, where $m = 2$ and $I_1 = [k_1,k_2]$, $I_2 = [l_1,l_2]$ and $I_N = [n_1,n_2]$ (see [16], p. 249).

For simplicity, in the case $m = 2$ we use the following form of (A$_1$):

(A$_2$) Let $K : I_K \to \mathbb{R}$, $L : I_L \to \mathbb{R}$ and $N : I_N \to \mathbb{R}$ be continuous and strictly monotone functions whose domains are intervals in $\mathbb{R}$, and let $f : I_K \times I_L \to I_N$ be a continuous function. Let $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^n$ $(n \geq 2)$ such that $\mathbf{a} \in I_K^n$ and $\mathbf{b} \in I_L^n$, and let $\mathbf{p} = (p_1,...,p_n)$ be a nonnegative n-tuple such that $\sum_{i=1}^n p_i = 1$.

Then (4.1) has the form

$$f\left(K_n(\mathbf{a};\mathbf{p}),L_n(\mathbf{b};\mathbf{p})\right) \geq N_n(f(\mathbf{a},\mathbf{b});\mathbf{p}), \tag{4.3}$$

where $f(\mathbf{a},\mathbf{b})$ means $(f(a_1,b_1),...,f(a_n,b_n))$.

The following results are important special cases of Theorem 4.1, and generalize the corresponding results of Beck. The next hypothesis will be used:

(A$_3$) Let $K : I_K \to \mathbb{R}$, $L : I_L \to \mathbb{R}$ and $N : I_N \to \mathbb{R}$ be continuous and strictly monotone functions whose domains are intervals in $\mathbb{R}$ such that either $I_K + I_L \subset I_N$ and $f(x,y) = x+y$ $((x,y) \in I_K \times I_L)$ or I_K, $I_L \subset]0,\infty[$, $I_K \cdot I_L \subset I_N$ and $f(x,y) = xy$ $((x,y) \in I_K \times I_L)$. Assume further that the functions K, L and N are twice continuously differentiable on the interior of their domains, respectively. Let $\mathbf{a}$, $\mathbf{b} \in \mathbb{R}^n$ $(n \geq 2)$ be such that $\mathbf{a} \in I_K^n$ and $\mathbf{b} \in I_L^n$, and let $\mathbf{p} = (p_1,...,p_n)$ be a nonnegative n-tuple such that $\sum_{i=1}^n p_i = 1$.

The interior of a subset A of $\mathbb{R}$ is denoted by A°.

Corollary 4.1 *Assume (A$_3$) with $f(x,y) = x+y$ $((x,y) \in I_K \times I_L)$, and assume that K', L', N', K'', L'' and N'' are all positive. Introducing $E := \frac{K'}{K''}$, $F := \frac{L'}{L''}$, $G := \frac{N'}{N''}$, (4.3) holds for all possible $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{p}$ if and only if*

$$E(x) + F(y) \leq G(x+y), \quad (x,y) \in I_K^\circ \times I_L^\circ. \tag{4.4}$$

Corollary 4.2 *Assume (A_3) with $f(x,y) = xy$ $((x,y) \in I_K \times I_L)$. Suppose the functions $A(x) := \frac{K'(x)}{K'(x)+xK''(x)}$, $B(x) := \frac{L'(x)}{L'(x)+xL''(x)}$ and $C(x) := \frac{N'(x)}{N'(x)+xN''(x)}$ are defined on I_K°, I_L° and I_N°, respectively. Assume further that K', L', N', A, B and C are all positive. Then (4.3) holds for all possible* **a**, **b** *and* **p** *if and only if*

$$A(x) + B(y) \le C(xy), \quad (x,y) \in I_K^\circ \times I_L^\circ.$$

To prove these corollaries, similar arguments can be applied as in the analogous results of Beck. We just sketch the proof of Corollary 4.1.

Proof. By Theorem 4.1, it is enough to prove that the function

$$H : K(I_K) \times L(I_L) \to \mathbb{R}, \quad H(t,s) := N(K^{-1}(t) + L^{-1}(s))$$

is concave. Since H is continuous, and twice continuously differentiable on the interior $K(I_K^\circ) \times L(I_L^\circ)$ of its domain, we have to show that

$$h_1^2 \frac{\partial^2 H(t,s)}{\partial t^2} + 2h_1 h_2 \frac{\partial^2 H(t,s)}{\partial t \partial s} + h_2^2 \frac{\partial^2 H(t,s)}{\partial s^2} \le 0$$

for all $(t,s) \in K(I_K^\circ) \times L(I_L^\circ)$ and $(h_1,h_2) \in \mathbb{R}^2$. By computing the partial derivatives of H of order 2 at the points of $K(I_K^\circ) \times L(I_L^\circ)$, we have that this condition follows from (4.4). □

Interpolations of the discrete Jensen's inequality (1.2) given in [84] are used in [75] (see also [74], p. 195) to refine the inequality of Beck for a function of two variables. More general refinements of Beck' inequality can be found in [44, 45] (see also [36] Chapter 7) by applying refinements of the discrete Jensen's inequality (1.1) appeared in [41, 42, 43]. In this section we give some new refinements of Beck's inequality (4.1) by using the results in [13]. This leads to some new refinements of the classical Hölder's and Minkowski's inequalities.

4.2 Refinements when p is the discrete uniform distribution

For the sake of completeness we give the following refinement of the discrete Jensen's inequality which is a generalization of Theorem 2.1, and a special case of Theorem 2.6.

Theorem 4.2 *Assume U is a convex set in $\mathbb{R}^m$, $\mathbf{x}_1,\dots,\mathbf{x}_n \in U$, and $\lambda = (\lambda_1,\dots,\lambda_k)$ is a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for $2 \le k \le n$. Then*

$$f\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\right) \le S := \frac{1}{n}\sum_{i=1}^{n} f\left(\sum_{j=0}^{k-1}\lambda_{j+1}\mathbf{x}_{i+j}\right) \le \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i). \tag{4.5}$$

Assume (A_1) with $\mathbf{p} = \left(\frac{1}{n},\dots,\frac{1}{n}\right)$, and let $\lambda = (\lambda_1,\dots,\lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for $2 \le k \le n$. The cyclic mixed symmetric means relative to S are defined by:

$$M(L_1,...,L_m;\mathbf{x}^{(1)},...,\mathbf{x}^{(m)}) := N^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} N\left(f\left(L_1(\mathbf{x}^{(1)};k),...,L_m(\mathbf{x}^{(m)};k)\right)\right)\right) \quad (4.6)$$

$$L_t(\mathbf{x}^{(t)};k) = L_t^{-1}\left(\sum_{j=0}^{k-1}\lambda_{j+1}L_t(x_{i+j}^{(t)})\right); t = 1,\dots,m.$$

Now, we get an interpolation of (4.1) by the direct application of Theorem 4.2 as follows.

Theorem 4.3 *Assume (A_1) with $\mathbf{p} = \left(\frac{1}{n},\dots,\frac{1}{n}\right)$, and let $\lambda = (\lambda_1,\dots,\lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k}\lambda_i = 1$ for $2 \le k \le n$. If N is an increasing (decreasing) function, then the inequalities*

$$\begin{aligned} f\left(L_1(\mathbf{x}^{(1)};n),...,L_m(\mathbf{x}^{(m)};n)\right) &\le M(L_1,...,L_m;\mathbf{x}^{(1)},...,\mathbf{x}^{(m)}) \\ &\le N^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} N(f(x_i^{(1)},...,x_i^{(m)}))\right), \end{aligned} \quad (4.7)$$

hold for all possible $\mathbf{x}^{(t)}$ $(t = 1,\dots,m)$, if and only if the function H is defined in (4.2) *is convex (concave). If N is an increasing (decreasing) function, then the inequalities in (4.7) are reversed for all possible $\mathbf{x}^{(t)}$ $(t = 1,\dots,m)$, if and only if H is concave (convex).*

Proof. Suppose N is increasing and the function $H : L_1(I_1)\times ... \times L_m(I_m) \to \mathbb{R}$,

$$H(t_1,...,t_m) = N\left(f\left(L_1^{-1}(t_1),...,L_m^{-1}(t_m)\right)\right)$$

is convex. We can apply Theorem 4.2 to the function H and to the vectors $(L_1(x_i^{(1)}),\dots,L_m(x_i^{(m)}))$, $i = 1,\dots,n$. Then the first term in (4.5) gives

$$\begin{aligned} H\left(\frac{1}{n}\sum_{i=1}^{n}\left(L_1(x_i^{(1)}),\dots,L_m(x_i^{(m)})\right)\right) &= H\left(\frac{1}{n}\sum_{i=1}^{n}L_1(x_i^{(1)}),\dots,\frac{1}{n}\sum_{i=1}^{n}L_m(x_i^{(m)})\right) \\ &= N\left(f\left(L_1^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}L_1(x_i^{(1)})\right),...,L_m^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}L_m(x_i^{(m)})\right)\right)\right) \\ &= N\left(f\left(L_1(\mathbf{x}^{(1)};n),...,L_m(\mathbf{x}^{(m)};n)\right)\right). \end{aligned}$$

The last term in (4.5) is

$$\frac{1}{n}\sum_{i=1}^{n}H(L_1(x_i^{(1)}),\dots,L_m(x_i^{(m)})) = \frac{1}{n}\sum_{i=1}^{n}N\left(f\left(x_i^{(1)},\dots,x_i^{(m)}\right)\right),$$

and the middle term in (4.5) has the form

$$\frac{1}{n}\sum_{i=1}^{n} H\left(\sum_{j=0}^{k-1}\lambda_{j+1}\left(L_1(x_{i+j}^{(1)}),\ldots,L_m(x_{i+j}^{(m)})\right)\right)$$
$$=\frac{1}{n}\sum_{i=1}^{n} H\left(\sum_{j=0}^{k-1}\lambda_{j+1}L_1(x_{i+j}^{(1)}),\ldots,\sum_{j=0}^{k-1}\lambda_{j+1}L_m(x_{i+j}^{(m)})\right)$$
$$=\frac{1}{n}\sum_{i=1}^{n} N\left(f\left(L_1^{-1}\left(\sum_{j=0}^{k-1}\lambda_{j+1}L_1(x_{i+j}^{(1)})\right),\ldots,L_m^{-1}\left(\sum_{j=0}^{k-1}\lambda_{j+1}L_m(x_{i+j}^{(m)})\right)\right)\right)$$
$$=\frac{1}{n}\sum_{i=1}^{n} N\left(f\left(L_1(\mathbf{x}^{(1)};k),\ldots,L_m(\mathbf{x}^{(m)};k)\right)\right).$$

The inequalities (4.7) follow from these observations and Theorem 4.2 since N^{-1} is increasing.

The converse is obtained by Theorem 4.1. □

Assume (A_2) with $\mathbf{p}=\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$, and let $\lambda=(\lambda_1,\ldots,\lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k}\lambda_i=1$ for $2\le k\le n$. Then, for $m=2$, the reverse of (4.7) can be written as

$$f\left(K_n(\mathbf{a}),L_n(\mathbf{b})\right)\ge M(K,L;\mathbf{a},\mathbf{b})\ge N^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}N(f(a_i,b_i)).\right) \tag{4.8}$$

Example 4.1 Let $f(x)=xy$ and $N(x)=x$. Then $H(s,t)=K^{-1}(s)L^{-1}(t)$. If H is concave then (4.8) gives the following refinement of Hölder's inequality.

$$\frac{1}{n}\sum_{i=1}^{n}a_ib_i\le\frac{1}{n}\sum_{i=1}^{n}K(\mathbf{a};k)L(\mathbf{b};k)\le K_n(\mathbf{a})L_n(\mathbf{b}). \tag{4.9}$$

In particular, if $H(s,t)=s^{1/q}t^{1/r}$ so H is concave for $q,r>1$ and $q^{-1}+r^{-1}=1$; we get the following refinement of the classical Hölder's inequality for positive n-tuples **a** and **b**.

$$\sum_{i=1}^{n}a_ib_i\le\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}a_{i+j}^{q}\right)^{\frac{1}{q}}\left(\sum_{j=0}^{k-1}\lambda_{j+1}b_{i+j}^{r}\right)^{\frac{1}{r}}\le\left(\sum_{i=1}^{n}a_i^q\right)^{\frac{1}{q}}\left(\sum_{i=1}^{n}b_i^r\right)^{\frac{1}{r}}.$$

Example 4.2 If $H(s,t)=(s^{1/p}+t^{1/p})^p$ then H is concave for $p>1$, and (4.8) reduces to the following refinement of the classical Minkowski's inequality for positive n-tuples **a** and **b**.

$$\left(\sum_{i=1}^{n}(a_i+b_i)^p\right)^{\frac{1}{p}}\le\left(\sum_{i=1}^{n}\left(\left(\sum_{j=0}^{k-1}\lambda_{j+1}a_{i+j}^{p}\right)^{\frac{1}{p}}+\left(\sum_{j=0}^{k-1}\lambda_{j+1}b_{i+j}^{p}\right)^{\frac{1}{p}}\right)^{p}\right)^{\frac{1}{p}}$$
$$\le\left(\sum_{i=1}^{n}a_i^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}b_i^p\right)^{\frac{1}{p}}.$$

On the analogy of Corollary 4.1 and Corollary 4.2, we have the following consequences of Theorem 4.3.

Corollary 4.3 *Assume* (A_3) *with* $\mathbf{p} = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$*, and let* $\lambda = (\lambda_1, \dots, \lambda_k)$ *be a positive k-tuple such that* $\sum\limits_{i=1}^{k} \lambda_i = 1$ *for* $2 \le k \le n$*. Suppose* $f(x,y) = x + y$ $((x,y) \in I_K \times I_L)$*, and assume that* K'*,* L'*,* N'*,* K''*,* L'' *and* N'' *are all positive. Introducing* $E := \frac{K'}{K''}$*,* $F := \frac{L'}{L''}$*,* $G := \frac{N'}{N''}$*, (4.8) holds for all possible* $\mathbf{a}$ *and* $\mathbf{b}$ *if and only if*

$$E(x) + F(y) \le G(x+y), \quad (x,y) \in I_K^\circ \times I_L^\circ.$$

In this case

$$M(K,L;\mathbf{a},\mathbf{b}) = N^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} N\left(K(\mathbf{a};k) + L(\mathbf{b};k)\right)\right). \tag{4.10}$$

Corollary 4.4 *Assume* (A_3) *with* $\mathbf{p} = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$*, and let* $\lambda = (\lambda_1, \dots, \lambda_k)$ *be a positive k-tuple such that* $\sum\limits_{i=1}^{k} \lambda_i = 1$ *for* $2 \le k \le n$*, and* $f(x,y) = xy$ $((x,y) \in I_K \times I_L)$*. Suppose the functions* $A(x) := \frac{K'(x)}{K'(x) + xK''(x)}$*,* $B(x) := \frac{L'(x)}{L'(x) + xL''(x)}$ *and* $C(x) := \frac{N'(x)}{N'(x) + xN''(x)}$ *are defined on* I_K°*,* I_L° *and* I_N° *respectively. Assume further that* K'*,* L'*,* M'*,* A*,* B *and* C *are all positive. Then (4.8) holds for all possible* $\mathbf{a}$ *and* $\mathbf{b}$ *if and only if*

$$A(x) + B(y) \le C(xy), \quad (x,y) \in I_K^\circ \times I_L^\circ.$$

In this case

$$M(K,L;\mathbf{a},\mathbf{b}) = N^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} N\left(K(\mathbf{a};k)L(\mathbf{b};k)\right)\right). \tag{4.11}$$

4.3 Refinement of Minkowski's inequality

(A_4) Let I be an interval in $\mathbb{R}$, and let $M : I \to \mathbb{R}$ be a continuous and strictly monotone function, and let $\mathbf{x}_i \in I^m$ $(i = 1, \dots, n)$. Let $(\lambda_1, ..., \lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for $2 \le k \le n$. Let $\mathbf{w} = (w_1, ..., w_m)$ be a nonnegative m-tuple such that $\sum\limits_{i=1}^{m} w_i = 1$.

We give a refinement of the Minkowski's inequality by using Theorem 4.2.

Theorem 4.4 *Assume* (A_4)*, and assume that the quasi-arithmetic mean function*

$$\mathbf{x} \to M_m(\mathbf{x};\mathbf{w}) := M^{-1}\left(\sum_{i=1}^{m} w_i M(x_i)\right), \quad \mathbf{x} \in I^m$$

is convex. Then

$$M_m\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i;\mathbf{w}\right) \leq \frac{1}{n}\sum_{i=1}^{n}M_m\left(\sum_{j=0}^{k-1}\lambda_{j+1}\mathbf{x}_{i+j};\mathbf{w}\right) \leq \frac{1}{n}\sum_{r=1}^{n}M_m(\mathbf{x}_r;\mathbf{w}). \tag{4.12}$$

Proof. This is obtained by applying Theorem 4.2 to the function $M_m(\cdot;\mathbf{w})$ and to the vectors $\mathbf{x}_i$ $(i = 1,\ldots,n)$. □

The following necessary and sufficient condition for the quasi-arithmetic mean function to be convex is given in ([74], p. 197):

Theorem C. *If $M : [m_1, m_2] \to R$ has continuous derivatives of second order and it is strictly increasing and strictly convex, then the quasi-arithmetic mean function $M_m(\cdot;w)$ is convex if and only if M'/M'' is a concave function.*

(A_5) Let $M :]0,\infty[\to]0,\infty[$ be a continuous and strictly monotone function such that $\lim_{x\to 0} M(x) = \infty$ or $\lim_{x\to\infty} M(x) = \infty$. Let $\mathbf{x}_i \in I^m$ $(i = 1,\ldots,n)$, and let $(\lambda_1,...,\lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k}\lambda_i = 1$ for $2 \leq k \leq n$. Let $\mathbf{w} = (w_1,...,w_m)$ be positive m-tuple such that $w_i \geq 1$ $(i = 1,...,m)$.

Then we define

$$\widetilde{M}_m(\mathbf{x};\mathbf{w}) = M^{-1}\left(\sum_{i=1}^{m} w_i M(x_i)\right). \tag{4.13}$$

The following result is also given in ([74], page 197):

Theorem D. *If $M :]0,\infty[\to]0,\infty[$ has continuous derivatives of second order and it is strictly increasing and strictly convex, then $\widetilde{M}_m(\cdot;w)$ is a convex function if M/M' is a convex function.*

By using (4.13) we have

Theorem 4.5 *Assume (A_5). If the function*

$$\mathbf{x} \to \widetilde{M}_m(\mathbf{x};\mathbf{w}), \quad \mathbf{x} \in]0,\infty[^m$$

is convex, then Theorem 4.4 remains valid for $\widetilde{M}_m(\mathbf{x};\mathbf{w})$ instead of $M_m(\mathbf{x};\mathbf{w})$.

4.4 Weighted version of cyclic refinement of Beck's inequalities

In the pervious sections, a cyclic refinement (unweighted version) of the inequality of E. Beck [9] is presented. Now we generalize those results for positive weights. In the rest of the Chapter 4 we work out the weighted analogue of new refinement of Beck's inequality (4.1) by weighted cyclic mixed symmetric means as a consequence of the refinement developed in Section 2.1.1. This obviously, leads to refinements of weighted forms of discrete Hölder's and Minkowski's inequalities.

4.5 Weighted Refinement of Beck's Inequality

The following refinement of the discrete Jensen's inequality is a special case of Theorem 2.6. For the sake of completeness we give it.

Theorem 4.6 *Assume U is a convex set in $\mathbb{R}^m$, $\mathbf{x}_1,\ldots,\mathbf{x}_n \in U$. Let $2 \le k \le n$ be integers, and let $p_1,\ldots,p_n$ and $\lambda_1,\ldots,\lambda_k$ represent positive probability distributions. Then*

$$f\left(\sum_{i=1}^{n} p_i \mathbf{x}_i\right) \le$$

$$C_{dis} := \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\mathbf{x}_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \le \sum_{i=1}^{n} p_i f(\mathbf{x}_i). \tag{4.14}$$

Proof. Assume (A$_1$), and let $\lambda = (\lambda_1,\ldots,\lambda_k)$ be a positive k-tuple such that $\sum\limits_{i=1}^{k}\lambda_i = 1$ for $2 \le k \le n$. The weighted cyclic mixed symmetric means relative to C_{dis} are defined by:

$$\begin{aligned} &M(L_1,\ldots,L_m;\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(m)};\mathbf{p};\lambda) \\ &:= N^{-1}\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)N\left(f\left(L_1(\mathbf{x}^{(1)},\mathbf{p},\lambda;k),\ldots,L_m(\mathbf{x}^{(m)},\mathbf{p},\lambda;k)\right)\right)\right) \end{aligned} \tag{4.15}$$

$$L_t(\mathbf{x}^{(t)},\mathbf{p},\lambda;k) = L_t^{-1}\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}L_t(x_{i+j}^{(t)})}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right); \quad t = 1,\ldots,m.$$

□

Now, we get an interpolation of (4.1) by the direct application of Theorem 4.6 as follows.

Theorem 4.7 *Assume (A$_1$), and let $\lambda = (\lambda_1,\ldots,\lambda_k)$ be a positive k-tuple such that $\sum\limits_{i=1}^{k}\lambda_i = 1$ for $2 \le k \le n$. If N is an increasing (decreasing) function, then the inequalities*

$$\begin{aligned} &f\left(L_1(\mathbf{x}^{(1)},\mathbf{p};n),\ldots,L_m(\mathbf{x}^{(m)},\mathbf{p};n)\right) \le M(L_1,\ldots,L_m;\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(m)};\mathbf{p};\lambda) \\ &\le N^{-1}\left(\sum_{i=1}^{n} p_i N(f(x_i^{(1)},\ldots,x_i^{(m)}))\right), \end{aligned} \tag{4.16}$$

hold for all possible $\mathbf{x}^{(t)}$ $(t = 1,\ldots,m)$, $\mathbf{p}$ and λ if and only if the function H is defined in (4.2) is convex (concave). If N is an increasing (decreasing) function, then the inequalities in (4.16) are reversed for all possible $\mathbf{x}^{(t)}$ $(t = 1,\ldots,m)$, $\mathbf{p}$ and λ if and only if H is concave (convex).

Proof. We can apply Theorem 4.6 (the proof is almost the same as of Theorem 4.2. □

Assume (A_2), and let $\lambda = (\lambda_1, \ldots, \lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for $2 \le k \le n$. Then, for $m = 2$, the reverse of (4.16) can be written as

$$f\left(K_n(\mathbf{a},\mathbf{p}), L_n(\mathbf{b},\mathbf{p})\right) \ge M(K,L;\mathbf{a},\mathbf{b};\mathbf{p};\lambda) \ge N^{-1}\left(\sum_{i=1}^{n} p_i N(f(a_i,b_i))\right). \tag{4.17}$$

Example 4.3 Let $f(x) = xy$ and $N(x) = x$. Then $H(s,t) = K^{-1}(s)L^{-1}(t)$. If H is concave then (4.17) gives the following refinement of Hölder's inequality.

$$\sum_{i=1}^{n} p_i a_i b_i \le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right) K(\mathbf{a},\mathbf{p},\lambda;k) L(\mathbf{b},\mathbf{p},\lambda;k) \le K_n(\mathbf{a},\mathbf{p}) L_n(\mathbf{b},\mathbf{p}). \tag{4.18}$$

In particular, if $H(s,t) = s^{1/q} t^{1/r}$ so H is concave for $q, r > 1$ and $q^{-1} + r^{-1} = 1$; we get the following refinement of the classical Hölder's inequality for positive n-tuples $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{p}$.

$$\begin{aligned}
\sum_{i=1}^{n} p_i a_i b_i &\le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} a_{i+j}^{q}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^{\frac{1}{q}}\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} b_{i+j}^{r}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^{\frac{1}{r}} \\
&\le \left(\sum_{i=1}^{n} p_i a_i^{q}\right)^{\frac{1}{q}}\left(\sum_{i=1}^{n} p_i b_i^{r}\right)^{\frac{1}{r}}.
\end{aligned}$$

Example 4.4 If $H(s,t) = (s^{1/p} + t^{1/p})^p$ then H is concave for $p > 1$, and (4.17) reduces to the following refinement of the classical Minkowski's inequality for positive n-tuples $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{p}$.

$$\begin{aligned}
&\left(\sum_{i=1}^{n} p_i (a_i + b_i)^p\right)^{\frac{1}{p}} \\
&\le \left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right)\left(\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} a_{i+j}^{p}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^{\frac{1}{p}} + \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} b_{i+j}^{p}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^{\frac{1}{p}}\right)^{p}\right)^{\frac{1}{p}} \\
&\le \left(\sum_{i=1}^{n} p_i a_i^{p}\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} p_i b_i^{p}\right)^{\frac{1}{p}}.
\end{aligned}$$

On the analogy of Corollary 4.1 and Corollary 4.2, we have the following consequences of Theorem 4.7.

Corollary 4.5 *Assume (A_3), and let $(\lambda_1, \ldots, \lambda_k)$ be a positive k-tuple such that $\sum_{i=1}^{k} \lambda_i = 1$ for $2 \le k \le n$. Suppose $f(x,y) = x + y$ ($(x,y) \in I_K \times I_L$), and assume that K', L', N', K'', L'' and N'' are all positive. Introducing $E := \frac{K'}{K''}$, $F := \frac{L'}{L''}$, $G := \frac{N'}{N''}$, (4.17) holds for all possible* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{p}$ *if and only if*

$$E(x) + F(y) \le G(x+y), \quad (x,y) \in I_K^{\circ} \times I_L^{\circ}.$$

In this case

$$M(K,L;\mathbf{a},\mathbf{b};\mathbf{p};\lambda) = N^{-1}\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)N\left(K(\mathbf{a};k)+L(\mathbf{b};k)\right)\right). \tag{4.19}$$

Corollary 4.6 *Assume (A*$_3$*), and let* $(\lambda_1,\dots,\lambda_k)$ *be a positive k-tuple such that* $\sum_{i=1}^{k}\lambda_i = 1$ *for* $2 \le k \le n$. *Suppose* $f(x,y) = xy$ $((x,y) \in I_K \times I_L)$ *and functions* $A(x) := \frac{K'(x)}{K'(x)+xK''(x)}$, $B(x) := \frac{L'(x)}{L'(x)+xL''(x)}$ *and* $C(x) := \frac{N'(x)}{N'(x)+xN''(x)}$ *are defined on* I_K°, I_L° *and* I_N° *respectively. Assume further that* K', L', M', A, B *and* C *are all positive. Then (4.17) holds for all possible* **a** *and* **b** *if and only if*

$$A(x)+B(y) \le C(xy), \quad (x,y) \in I_K^\circ \times I_L^\circ.$$

In this case

$$M(K,L;\mathbf{a},\mathbf{b};\mathbf{p};\lambda) = N^{-1}\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)N\left(K(\mathbf{a};k)L(\mathbf{b};k)\right)\right). \tag{4.20}$$

4.6 Weighted Refinement of Minkowski's inequality

We give a refinement of weighted discrete Minkowski's inequality by using Theorem 4.6.

Theorem 4.8 *Assume (A*$_4$*) and let* $(p_1,\dots,p_n)$ *be a positive probability distribution. Assume that the quasi-arithmetic mean function*

$$\mathbf{x} \to M_m(\mathbf{x};\mathbf{w}) := M^{-1}\left(\sum_{i=1}^{m} w_i M(x_i)\right), \quad \mathbf{x} \in I^m$$

is convex. Then

$$\begin{aligned} M_m\left(\sum_{i=1}^{n} p_i\mathbf{x}_i;\mathbf{w}\right) &\le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)M_m\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\mathbf{x}_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}};\mathbf{w}\right) \\ &\le \sum_{r=1}^{n} p_r M_m(\mathbf{x}_r;\mathbf{w}). \end{aligned} \tag{4.21}$$

Proof. This is obtained by applying Theorem 4.6 to the function $M_m(\cdot;\mathbf{w})$ and to the vectors $\mathbf{x}_i$ $(i = 1,\dots,n)$. □

By using (4.13) we have

Theorem 4.9 *Assume (A*$_5$*) with* $(p_1,\dots,p_n)$ *and* $(\lambda_1,\dots,\lambda_k)$ *be positive probability distributions for* $2 \le k \le n$. *If the function*

$$\mathbf{x} \to \widetilde{M}_m(\mathbf{x};\mathbf{w}), \quad \mathbf{x} \in]0,\infty[^m$$

is convex, then Theorem 4.8 remains valid for $\widetilde{M}_m(\mathbf{x};\mathbf{w})$ *instead of* $M_m(\mathbf{x};\mathbf{w})$.

Chapter 5

Cyclic Refinements of the Different Versions of Operator Jensen's Inequality

Refinements of the operator Jensen's inequality for convex and operator convex functions are given by using cyclic refinements of the discrete Jensen's inequality. Similar refinements are fairly rare in the literature. Some applications of the results to norm inequalities, to the Hölder-McCarthy inequality and to generalized weighted power means for operators are presented. Refinements of operator versions of Jensen's inequality has been less extensively studied than refinements of the discrete or the integral form of Jensen's inequality. For some results, we refer to the papers Khosravi, Aujla, Dragomir and Moslehian [56], Niezgoda [81], Khan and Hanif [55], Kian and Moslehian [57], and the book Horváth, Khuram Ali Khan and Pečarić [36]. This chapter is based on the paper Horváth, Khuram Ali Khan and Pečarić [46].

5.1 Introduction

In the present Chapter 5 $(H, \langle\cdot,\cdot\rangle)$ will always mean a complex Hilbert space. The Banach algebra of all bounded linear operators on H will be denoted by $\mathscr{B}(H)$. We always understand the norm of an operator $A \in \mathscr{B}(H)$ as

$$\|A\| := \sup_{\|x\| \leq 1} \|Ax\| .$$

The operator I_H means the identity operator on H. The spectrum of an operator $A \in \mathscr{B}(H)$ is denoted by $\mathrm{Sp}(A)$. An operator $A \in \mathscr{B}(H)$ is called positive, if $\langle Ax, x\rangle \geq 0$ for every $x \in H$, or equivalently A is self-adjoint and $\mathrm{Sp}(A) \subset [0, \infty[$. An operator $A \in \mathscr{B}(H)$ is called strictly positive, if it is positive and invertible. For an interval $J \subset \mathbb{R}$, $S(J)$ means the class of all self-adjoint operators from $\mathscr{B}(H)$ whose spectra are contained in J.

Let $J \subset \mathbb{R}$ be an interval, and $f : J \to \mathbb{R}$ be a function. If f is continuous on J, and $A \in S(J)$, then $f(A)$ is defined by the symbolic calculus for self-adjoint operators (see Rudin [87]). The function f is said to be operator monotone (increasing on J) if f is continuous on J and A, $B \in S(J)$, $A \leq B$ (i.e. $A - B$ is a positive operator) imply $f(A) \leq f(B)$. The function f is called operator convex (on J) if f is continuous on J and

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B)$$

for all A, $B \in S(J)$ and for all $\lambda \in [0,1]$.

We say that the numbers $p_1, \ldots, p_n$ represent a (positive) discrete probability distribution if $(p_i > 0)$ $p_i \geq 0$ $(1 \leq i \leq n)$ and $\sum\limits_{i=1}^{n} p_i = 1$.

The following well known results are operator versions of Jensen's inequality:

Theorem 5.1 *Operator Jensen's inequality for convex functions (see Mond and Pečarić [76] and Furuta, Mićić, Pečarić and Seo [31]): Let $J \subset \mathbb{R}$ be an interval. Let $A_i \in S(J)$ and $x_i \in H$ $(i = 1, \ldots, n)$ with $\sum\limits_{i=1}^{n} \|x_i\|^2 = 1$. If $f : J \to \mathbb{R}$ is continuous and convex, then*

$$f\left(\sum_{i=1}^{n} \langle A_i x_i, x_i\rangle\right) \leq \sum_{i=1}^{n} \langle f(A_i) x_i, x_i\rangle . \tag{5.1}$$

Theorem 5.2 *Operator Jensen's inequality for operator convex functions (see Mond and Pečarić [77]): Let $J \subset \mathbb{R}$ be an interval, and K be a complex Hilbert space. Let $A_i \in S(J)$ $(i = 1, \ldots, n)$, $\Phi_i : \mathscr{B}(H) \to \mathscr{B}(K)$ $(i = 1, \ldots, n)$ be unital positive linear maps, and let $p_1, \ldots, p_n$ represent a discrete probability distribution. If $f : J \to \mathbb{R}$ is operator convex, then*

$$f\left(\sum_{i=1}^{n} p_i \Phi_i (A_i)\right) \leq \sum_{i=1}^{n} p_i \Phi_i (f(A_i)) . \tag{5.2}$$

A linear map $\Phi : \mathscr{B}(H) \to \mathscr{B}(K)$ is positive if $\Phi(A)$ is positive for all positive $A \in \mathscr{B}(H)$, and unital if $\Phi(I_H) = \Phi(I_K)$. Φ is called strictly positive if $\Phi(A)$ is strictly positive for all strictly positive $A \in \mathscr{B}(H)$.

5.2 Cyclic refinements of the operator Jensen's inequality for convex functions

In the present Chapter 5 we shall use the following convention: let $2 \leq k \leq n$ be integers, $i \in \{1,\ldots,n\}$ and $j \in \{0,\ldots,k-1\}$; if $i+j > n$, then $i+j$ means $i+j-n$.

Our first result a new refinement of the operator Jensen's inequality for convex functions:

Theorem 5.3 *Let $2 \leq k \leq n$ be integers, let $\mathbf{x} := (x_1,\ldots,x_n) \in H^n$ such that $x_i \neq 0$ $(i = 1,\ldots,n)$ and $\sum_{i=1}^{n} \|x_i\|^2 = 1$, and let $\lambda := (\lambda_1,\ldots,\lambda_k)$ represent a positive discrete probability distribution. Let $J \subset \mathbb{R}$ be an interval, $A_i \in S(J)$ $(i = 1,\ldots,n)$ and $\mathbf{A} := (A_1,\ldots,A_n)$. If $f : J \to \mathbb{R}$ is continuous and convex, then*

$$
\begin{aligned}
f\left(\sum_{i=1}^{n} \langle A_i x_i, x_i\rangle\right) &\leq D_c = D_c(f,\mathbf{A},\mathbf{x},\lambda) \\
&= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right) f\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2}\sum_{j=0}^{k-1}\lambda_{j+1}\left\langle A_{i+j}x_{i+j}, x_{i+j}\right\rangle\right) \\
&\leq \sum_{i=1}^{n}\langle f(A_i)x_i, x_i\rangle .
\end{aligned}
$$

Proof. Since

$$
\sum_{j=0}^{k-1}\left\|\frac{\sqrt{\lambda_{j+1}}x_{i+j}}{\left(\sum\limits_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right)^{1/2}}\right\|^2 = 1,
$$

the operator Jensen's inequality for convex functions yields

$$
\begin{aligned}
D_c &= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right)\cdot f\left(\sum_{j=0}^{k-1}\left\langle A_{i+j}\frac{\sqrt{\lambda_{j+1}}x_{i+j}}{\left(\sum\limits_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right)^{1/2}}, \frac{\sqrt{\lambda_{j+1}}x_{i+j}}{\left(\sum\limits_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right)^{1/2}}\right\rangle\right) \\
&\leq \sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}\left\langle f(A_{i+j})x_{i+j}, x_{i+j}\right\rangle = \left(\sum_{i=1}^{n}\langle f(A_i)x_i, x_i\rangle\right)\left(\sum_{j=1}^{k}\lambda_j\right) = \sum_{i=1}^{n}\langle f(A_i)x_i, x_i\rangle .
\end{aligned}
$$

Conversely, it is easy to check that

$$\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}\left\|x_{i+j}\right\|^2\right)=1,$$

and therefore the convexity of f implies

$$D_c \geq f\left(\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}\left\langle A_{i+j}x_{i+j},x_{i+j}\right\rangle\right)=f\left(\left(\sum_{i=1}^{n}\left\langle A_ix_i,x_i\right\rangle\right)\left(\sum_{j=1}^{k}\lambda_j\right)\right)$$

$$=f\left(\sum_{i=1}^{n}\left\langle A_ix_i,x_i\right\rangle\right).$$

The proof is complete. □

The following particular case is interesting.

Corollary 5.1 *Let $2\leq k\leq n$ be integers, let $x\in H$ with $\|x\|=1$, and let $\lambda_1,\dots,\lambda_k$ and $p_1,\dots,p_n$ represent positive discrete probability distributions. Let $J\subset\mathbb{R}$ be an interval, and $A_i\in S(J)$ $(i=1,\dots,n)$. If $f:J\to\mathbb{R}$ is continuous and convex, then*

(a)

$$f\left(\left\langle\sum_{i=1}^{n}p_iA_ix,x\right\rangle\right)\leq\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)f\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\left\langle\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}A_{i+j}x,x\right\rangle\right)$$

$$\leq\left\langle\sum_{i=1}^{n}p_if\left(A_i\right)x,x\right\rangle.$$

(b) In case of $A:=A_1=\ldots=A_n$

$$f\left(\left\langle Ax,x\right\rangle\right)\leq\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)f\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\left\langle\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}Ax,x\right\rangle\right)$$

$$\leq\left\langle f\left(A\right)x,x\right\rangle.$$

Proof. (a) Theorem 5.3 can be applied to the vectors $x_i:=\sqrt{p_i}x$ $(i=1,\dots,n)$.

(b) It is a special case of (a). □

Some norm inequalities can be obtained from Corollary 5.1 (a).

Corollary 5.2 *Let $2\leq k\leq n$ be integers, and let $\lambda_1,\dots,\lambda_k$ and $p_1,\dots,p_n$ represent positive discrete probability distributions. Let $J\subset[0,\infty[$ be an interval, and $A_i\in S(J)$ $(i=1,\dots,n)$. If $f:J\to\mathbb{R}$ is nonnegative, continuous, increasing and convex, then*

$$f\left(\left\|\sum_{i=1}^{n}p_iA_i\right\|\right)\leq\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)f\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\left\|\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}A_{i+j}\right\|\right)$$

$$\leq\left\|\sum_{i=1}^{n}p_if\left(A_i\right)\right\|.$$

Proof. If $A \in \mathscr{B}(H)$ is a positive operator, then $\|A\| = \sup\limits_{\|x\|=1} \langle Ax, x\rangle$. By using this, the continuity and the increase of f, and Corollary 5.1 (a), we have

$$
\begin{aligned}
f\left(\left\|\sum_{i=1}^{n} p_i A_i\right\|\right) &= f\left(\sup_{\|x\|=1}\left\langle \sum_{i=1}^{n} p_i A_i x, x\right\rangle\right) = \sup_{\|x\|=1} f\left(\left\langle \sum_{i=1}^{n} p_i A_i x, x\right\rangle\right) \\
&\le \sup_{\|x\|=1} \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\left\langle \sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}A_{i+j}x, x\right\rangle\right) \\
&\le \sup_{\|x\|=1}\left\langle \sum_{i=1}^{n} p_i f(A_i)x, x\right\rangle = \left\|\sum_{i=1}^{n} p_i f(A_i)\right\|.
\end{aligned}
$$

□

Remark 5.1 We consider now some special cases of Corollary 5.2. Let $2 \le k \le n$ be integers, and let $\lambda_1, \ldots, \lambda_k$ and $p_1, \ldots, p_n$ represent positive discrete probability distributions. Let $J \subset [0, \infty[$ be an interval, and $A_i \in S(J)$ $(i = 1, \ldots, n)$.

(a) For $\alpha \ge 1$

$$
\begin{aligned}
\left\|\sum_{i=1}^{n} p_i A_i\right\|^{\alpha} &\le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)^{1-\alpha}\left\|\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}A_{i+j}\right\|^{\alpha} \\
&\le \left\|\sum_{i=1}^{n} p_i A_i^{\alpha}\right\|,
\end{aligned} \tag{5.3}
$$

and for $0 < \alpha < 1$ the reverse inequalities hold. If the operators are strictly positive, (5.3) is also true for $\alpha < 0$.

(b) By choosing $f = \exp$, we have

$$
\begin{aligned}
\exp\left(\left\|\sum_{i=1}^{n} p_i A_i\right\|\right) &\le \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\exp\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\left\|\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}A_{i+j}\right\|\right) \\
&\le \left\|\sum_{i=1}^{n} p_i \exp(A_i)\right\|.
\end{aligned}
$$

From Corollary 5.1 (b) a refinement of the Hölder-McCarthy inequality (see [66]) is derived.

Corollary 5.3 *Let $2 \le k \le n$ be integers, let $x \in H$ with $\|x\| = 1$, and let $\lambda_1, \ldots, \lambda_k$ and $p_1, \ldots, p_n$ represent positive discrete probability distributions. Let $A \in \mathscr{B}(H)$ be a positive operator. Then* *(a) For every* $\alpha \geq 1$

$$\langle Ax,x\rangle^{\alpha} \leq \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)^{1-\alpha}\left\langle \sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}Ax,x\right\rangle^{\alpha} \leq \langle A^{\alpha}x,x\rangle . \tag{5.4}$$

(b) For every $0 < \alpha < 1$

$$\langle Ax,x\rangle^{\alpha} \geq \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)^{1-\alpha}\left\langle \sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}Ax,x\right\rangle^{\alpha} \geq \langle A^{\alpha}x,x\rangle .$$

(c) If A is strictly positive and $\alpha < 0$, then (5.4) also holds.

5.3 Cyclic refinements of the operator Jensen's inequality for operator convex functions

In the next result we obtain a new refinement for operator Jensen's inequality for operator convex functions.

Theorem 5.4 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1,\ldots,\lambda_k)$ and $\mathbf{p} := (p_1,\ldots,p_n)$ represent positive discrete probability distributions. Let $J \subset \mathbb{R}$ be an interval, $A_i \in S(J)$ $(i=1,\ldots,n)$ and $A := (A_1,\ldots,A_n)$. Let K be a complex Hilbert space, $\Phi_i : \mathscr{B}(H) \to \mathscr{B}(K)$ $(i=1,\ldots,n)$ be unital positive linear maps, and $\mathbf{\Phi} := (\Phi_1,\ldots,\Phi_n)$. If $f : J \to \mathbb{R}$ is operator convex, then*

$$f\left(\sum_{i=1}^{n}p_i\Phi_i(A_i)\right) \leq D_{oc} = D_{oc}(f,\mathbf{A},\mathbf{\Phi},\mathbf{p},\lambda)$$

$$:= \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\Phi_{i+j}(A_{i+j})}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \leq \sum_{i=1}^{n}p_i\Phi_i(f(A_i)).$$

Proof. The operator Jensen's inequality for operator convex functions shows that

$$D_{oc} \leq \sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\Phi_{i+j}(f(A_{i+j}))$$

$$= \left(\sum_{i=1}^{n}p_i\Phi_i(f(A_i))\right)\left(\sum_{j=1}^{k}\lambda_j\right) = \sum_{i=1}^{n}p_i\Phi_i(f(A_i)).$$

Since

$$\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) = 1,$$

we can apply the operator Jensen's inequality for operator convex functions again, and have

$$D_{oc} \geq f\left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\Phi_{i+j}\left(A_{i+j}\right)\right)\right) = f\left(\sum_{i=1}^{n}p_i\Phi_i\left(A_i\right)\right).$$

The proof is complete. □

In the following variant of the previous result the maps $\Phi_1,\ldots,\Phi_n$ are defined directly in terms of unitary operators.

Corollary 5.4 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1,\ldots,\lambda_k)$ and $\mathbf{p} := (p_1,\ldots,p_n)$ represent positive discrete probability distributions. Let $J \subset \mathbb{R}$ be an interval, $A_i \in S(J)$ $(i = 1,\ldots,n)$ and $A := (A_1,\ldots,A_n)$. Let $C_i \in \mathscr{B}(H)$ $(i = 1,\ldots,n)$ be unitary operators. If $f : J \to \mathbb{R}$ is operator convex, then*

$$f\left(\sum_{i=1}^{n}p_iC_i^*A_iC_i\right) \leq \sum_{i=1}^{n}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}C_{i+j}^*A_{i+j}C_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$
$$\leq \sum_{i=1}^{n}p_iC_i^*f\left(A_i\right)C_i.$$

Proof. For every $i = 1,\ldots,n$ the map $\Phi_i : \mathscr{B}(H) \to \mathscr{B}(H)$ defined by

$$\Phi_i\left(A\right) = C_i^*AC_i$$

is a unital positive linear map, and hence Theorem 5.4 can be applied. □

As an application, we present some monotonicity results for operator means.

Let $A_i \in \mathscr{B}(H)$ $(i = 1,\ldots,n)$ be strictly positive operators, $A := (A_1,\ldots,A_n)$, and let $\mathbf{p} := (p_1,\ldots,p_n)$ represent a positive discrete probability distribution. Let K be a complex Hilbert space, $\Phi_i : \mathscr{B}(H) \to \mathscr{B}(K)$ $(i = 1,\ldots,n)$ be unital strictly positive linear maps, and $\mathbf{\Phi} := (\Phi_1,\ldots,\Phi_n)$.The generalized weighted power mean of the operators A_i $(i = 1,\ldots,n)$ is defined by (see [73])

$$M_n^{[\alpha]}\left(\mathbf{A},\mathbf{\Phi},\mathbf{p}\right) = M_n^{[\alpha]}\left(A_1,\ldots,A_n;\Phi_1,\ldots,\Phi_n;p_1,\ldots,p_n\right)$$
$$:= \left(\sum_{i=1}^{n}p_i\Phi_i\left(A_i^{\alpha}\right)\right)^{1/\alpha}, \quad \alpha \in \mathbb{R}\setminus\{0\}.$$

Theorem 5.5 *Let $2 \leq k \leq n$ be integers, and let $\lambda := (\lambda_1,\ldots,\lambda_k)$ and $\mathbf{p} := (p_1,\ldots,p_n)$ represent positive discrete probability distributions. Let $A_i \in \mathscr{B}(H)$ $(i = 1,\ldots,n)$ be strictly positive operators, $A := (A_1,\ldots,A_n)$. Let K be a complex Hilbert space, $\Phi_i : \mathscr{B}(H) \to \mathscr{B}(K)$ $(i = 1,\ldots,n)$ be unital strictly positive linear maps, and $\mathbf{\Phi} := (\Phi_1,\ldots,\Phi_n)$.*

Then

$$\left(\sum_{i=1}^{n} p_i \Phi_i\left(A_i^{\alpha}\right)\right)^{1/\alpha} \leq M_n^{[\alpha,\beta]} = M_n^{[\alpha,\beta]}\left(\mathbf{A}, \boldsymbol{\Phi}, \mathbf{p}, \lambda\right)$$

$$:= \left(\sum_{i=1}^{n}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \Phi_{i+j}\left(A_{i+j}^{\alpha}\right)}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)^{\beta/\alpha}\right)^{1/\beta} \tag{5.5}$$

$$\leq \left(\sum_{i=1}^{n} p_i \Phi_i\left(A_i^{\beta}\right)\right)^{1/\beta},$$

if either $\alpha \leq \beta \leq -1$ *or* $1 \leq \beta \leq -\alpha$ *or* $1 \leq \beta$ *and* $\alpha \leq \beta \leq 2\alpha$.

The reverse inequalities hold in (5.5) if either $1 \leq \beta \leq \alpha$ *or* $-\alpha \leq \beta \leq -1$ *or* $\beta \leq -1$ *and* $2\alpha \leq \beta \leq \alpha$.

Proof. The following properties of the function $g :]0,\infty[\to \mathbb{R}$, $g(x) = x^r$ are well known (see [31]): it is operator convex if either $1 \leq r \leq 2$ or $-1 \leq r \leq 0$, and $-g$ is operator convex if $0 \leq r \leq 1$; g is operator monotone increasing if $0 \leq r \leq 1$ and operator monotone decreasing if $-1 \leq r \leq 0$.

By using these properties, Theorem 5.4 can be applied to the function $f :]0,\infty[\to \mathbb{R}$, $f(x) = x^{\beta/\alpha}$ and the operators A_i^{α} $(i = 1,\ldots,n)$.

The proof is complete. □

Remark 5.2 $M_n^{[\alpha,\beta]}$ can be considered as the mixed symmetric mean corresponding to D_{oc} in Theorem 5.4.

Chapter 6

Cyclic Refinements of Jensen's Inequality Via Taylor's Formula

In this Chapter, we give new extensions and improvements of cyclic refinements of Jensen's Inequality by Taylor's Formula with and without the effect of Green functions. As an application of our work we construct new entropic bounds for Shannon, Relative and Mandelbrot entropies. This chapter is based on the papers [67] and [68].

For $d = 1,\ldots,5$, consider the Green functions $G_d : [\alpha,\beta]\times[\alpha,\beta] \to \mathbb{R}$ defined as

$$G_1(x,r) = \begin{cases} \frac{(\beta-x)(\alpha-r)}{\beta-\alpha}, & \alpha \le r \le x; \\ \frac{(\beta-r)(\alpha-x)}{\beta-\alpha}, & x \le r \le \beta. \end{cases} \tag{6.1}$$

$$G_2(x,r) = \begin{cases} \alpha - r, & \alpha \le r \le x, \\ \alpha - x, & x \le r \le \beta. \end{cases} \tag{6.2}$$

$$G_3(x,r) = \begin{cases} x - \beta, & \alpha \le r \le x, \\ r - \beta, & x \le r \le \beta. \end{cases} \tag{6.3}$$

$$G_4(x,r) = \begin{cases} x - \alpha, & \alpha \le r \le x, \\ r - \alpha, & x \le r \le \beta. \end{cases} \tag{6.4}$$

$$G_5(x,r) = \begin{cases} \beta - r, & \alpha \le r \le x, \\ \beta - x, & x \le r \le \beta, \end{cases} \tag{6.5}$$

All these functions are convex and continuous w.r.t both x and r, and the following Lemma holds.

Lemma 6.1 *Suppose $f \in C^2[\alpha,\beta]$, then the following identities are valid:*

$$f(x) = \frac{\beta - x}{\beta - \alpha} f(\alpha) + \frac{x - \alpha}{\beta - \alpha} f(\beta) + \int_{\alpha}^{\beta} G_1(x,r) f''(r) dr. \tag{6.6}$$

$$f(x) = f(\alpha) + (x-\alpha) f'(\beta) + \int_{\alpha}^{\beta} G_2(x,r) f''(r) dr, \tag{6.7}$$

$$f(x) = f(\beta) + (\beta - x) f'(\alpha) + \int_{\alpha}^{\beta} G_3(x,r) f''(r) dr, \tag{6.8}$$

$$f(x) = f(\beta) - (\beta - \alpha) f'(\beta) + (x - \alpha) f'(\alpha) + \int_{\alpha}^{\beta} G_4(x,r) f''(r) dr, \tag{6.9}$$

$$f(x) = f(\alpha) + (\beta - \alpha) f'(\alpha) - (\beta - x) f'(\beta) + \int_{\alpha}^{\beta} G_5(x,r) f''(r) dr. \tag{6.10}$$

Proof. Consider the integral

$$\int_{\alpha}^{\beta} G_d(x,r) f''(r) dr = \int_{\alpha}^{x} G_d(x,r) f''(r) dr + \int_{x}^{\beta} G_d(x,r) f''(r) dr$$

Fix $d = 1, \ldots, 5$ and perform the integration for the specific value of the Green's function, we shall obtained identities (6.6)–(6.10) for $d = 1, \ldots, 5$. □

Remark 6.1 The Green's function $G_1(\cdot,\cdot)$ is called Lagrange Green's function (see [94]). The new Green functions $G_d(\cdot,\cdot),\ (d = 2,3,4,5)$, introduced by Pečarić et al. in [67]. The result (6.7) given in the previous Lemma represents a special case of the representation of the function using the so-called 'two-point right focal' interpolating polynomial in case when n = 2 and p = 0 (see [1]). Lemma 6.1 gives another proof of special case of Abel-Gontscharoff identity (6.7). G_4 and G_5 are new Green functions but results are not so simple as in other two cases.

6.1 Extensions of cyclic refinements of Jensen's inequality by Taylor's formula

For $f : [\alpha,\beta] \to \mathbb{R}$ where $f^{(n-1)}$ is absolutely continuous, the renowned Taylor's formula $\forall\ x \in [\alpha,\beta]$ at the point $\xi \in [\alpha,\beta]$ is

$$f(x) = \sum_{w=0}^{n-1} \frac{f^{(w)}(c)}{w!}(x-c)^w + \frac{1}{(n-1)!}\int_c^x f^{(n)}(\xi)(x-\xi)^{n-1}d\xi. \tag{6.11}$$

To start for real weights, we need the following assumptions for the cyclic Jensen's functionals defined in Chapter 2, Remark 2.7:

(A_1) For the linear functionals $J_u(\cdot)$ $(u = 1,2)$, suppose (H_1-H_2) are satisfied and
$\frac{\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j}x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j}} \in [\alpha,\beta]$ for $i = 1,\dots m$.

(A_2) For the linear functionals $J_u(\cdot)$ $(u = 3,\dots,6)$, suppose (H_3-H_5) are satisfied and
$\frac{\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j}g(x_{i+j})}{\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j}} \in [\alpha,\beta]$ for $i = 1,\dots m$.

Initially, we take cyclic refinements of Jensen's inequality in discrete as well as continuous version and form the following identities with real weights by using Taylor's formula.

Theorem 6.1 *Suppose $m,k \in \mathbb{N}$, $p_1,\dots,p_m$ and $\lambda_1,\dots,\lambda_k$ are real tuples for $2 \le k \le m$, such that $\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j} \neq 0$ for $u = 1,\dots m$ with $\sum_{i=1}^{m} p_i = 1$ and $\sum_{j=1}^{k} \lambda_j = 1$. Also let $x \in [\alpha,\beta] \subset \mathbb{R}$ and $\mathbf{x} \in [\alpha,\beta]^m$. Consider the function $f : [\alpha,\beta] \to \mathbb{R}$ such that $f^{(n-1)}$ is absolutely-continuous and G_v, $(v = 1,\dots,5)$ are same as given in* (6.1)–(6.5)*, respectively. Then for $(u = 1,\dots,6)$ along with the assumptions (A_1) and (A_2), we have the following generalized identities:*

(a)

$$J_u(f) = \sum_{w=1}^{n-1} \frac{f^{(w)}(\alpha)}{w!}J_u((x-\alpha)^w) + \frac{1}{(n-1)!}\int_\alpha^\beta f^{(n)}(\xi)J_u((x-\xi)_+^{n-1})d\xi, \tag{6.12}$$

(b)

$$J_u(f) = \sum_{w=1}^{n-1} \frac{(-1)^w f^{(w)}(\beta)}{w!}J_u((\beta-x)^w) - \frac{(-1)^{n-1}}{(n-1)!}\int_\alpha^\beta f^{(n)}(\xi)J_u((\xi-x)_+^{n-1})d\xi, \tag{6.13}$$

(c)

$$
\begin{aligned}
J_u(f) = \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right) J_u(x) + \int_{\alpha}^{\beta} J_u(G_1(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr \\
+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} f^{(n)}(\xi) \left(\int_{\xi}^{\beta} J_u(G_1(x,r))(r-\xi)^{n-3} dr \right) d\xi, \quad (6.14)
\end{aligned}
$$

$$
\begin{aligned}
J_u(f) = f'(\beta) J_u(x) + \int_{\alpha}^{\beta} J_u(G_2(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr \\
+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} f^{(n)}(\xi) \left(\int_{\xi}^{\beta} J_u(G_2(x,r))(r-\xi)^{n-3} dr \right) d\xi, \quad (6.15)
\end{aligned}
$$

$$
\begin{aligned}
J_u(f) = -f'(\alpha) J_u(x) + \int_{\alpha}^{\beta} J_u(G_3(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr \\
+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} f^{(n)}(\xi) \left(\int_{\xi}^{\beta} J_u(G_3(x,r))(r-\xi)^{n-3} dr \right) d\xi, \quad (6.16)
\end{aligned}
$$

$$
\begin{aligned}
J_u(f) = f'(\alpha) J_u(x) + \int_{\alpha}^{\beta} J_u(G_4(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr \\
+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} f^{(n)}(\xi) \left(\int_{\xi}^{\beta} J_u(G_4(x,r))(r-\xi)^{n-3} dr \right) d\xi, \quad (6.17)
\end{aligned}
$$

$$
\begin{aligned}
J_u(f) = f'(\beta) J_u(x) + \int_{\alpha}^{\beta} J_u(G_5(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr \\
+ \frac{1}{(n-3)!} \int_{\alpha}^{\beta} f^{(n)}(\xi) \left(\int_{\xi}^{\beta} J_u(G_5(x,r))(r-\xi)^{n-3} dr \right) d\xi. \quad (6.18)
\end{aligned}
$$

(d)

$$J_u(f) = \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right) J_u(x) + \int_\alpha^\beta J_u(G_1(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\beta)^{w-2}}{(w-2)!}\right) dr$$
$$- \frac{1}{(n-3)!} \int_\alpha^\beta f^{(n)}(\xi) \left(\int_\alpha^\xi J_u(G_1(x,r))(r-\xi)^{n-3} dr\right) d\xi, \quad (6.19)$$

$$J_u(f) = f'(\beta) J_u(x) + \int_\alpha^\beta J_u(G_2(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\beta)^{w-2}}{(w-2)!}\right) dr$$
$$- \frac{1}{(n-3)!} \int_\alpha^\beta f^{(n)}(\xi) \left(\int_\alpha^\xi J_u(G_2(x,r))(r-\xi)^{n-3} dr\right) d\xi, \quad (6.20)$$

$$J_u(f) = -f'(\alpha) J_u(x) + \int_\alpha^\beta J_u(G_3(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\beta)^{w-2}}{(w-2)!}\right) dr$$
$$- \frac{1}{(n-3)!} \int_\alpha^\beta f^{(n)}(\xi) \left(\int_\alpha^\xi J_u(G_3(x,r))(r-\xi)^{n-3} dr\right) d\xi, \quad (6.21)$$

$$J_u(f) = f'(\alpha) J_u(x) + \int_\alpha^\beta J_u(G_4(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\beta)^{w-2}}{(w-2)!}\right) dr$$
$$- \frac{1}{(n-3)!} \int_\alpha^\beta f^{(n)}(\xi) \left(\int_\alpha^\xi J_u(G_4(x,r))(r-\xi)^{n-3} dr\right) d\xi, \quad (6.22)$$

$$J_u(f) = f'(\beta) J_u(x) + \int_\alpha^\beta J_u(G_5(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\beta)^{w-2}}{(w-2)!}\right) dr$$
$$- \frac{1}{(n-3)!} \int_\alpha^\beta f^{(n)}(\xi) \left(\int_\alpha^\xi J_u(G_5(x,r))(r-\xi)^{n-3} dr\right) d\xi. \quad (6.23)$$

Proof. Fix $u = 1,\ldots,6$.

(a) Applying Taylor's formula (6.11) at point α, we get

$$f(x)=\sum_{w=0}^{n-1}\frac{f^{(w)}(\alpha)}{w!}(x-\alpha)^{w}+\frac{1}{(n-1)!}\int_{\alpha}^{\beta}f^{(n)}(\xi)(x-\xi)_{+}^{n-1}d\xi, \tag{6.24}$$

where $(x-\xi)_{+}$ is a real valued function defined as:

$$(x-\xi)_{+}=\begin{cases}(x-\xi), & \xi\leq x,\\ 0, & \xi>x.\end{cases}$$

Using (6.24) in cyclic Jensen type linear functionals $J_u(\cdot)$ and practicing constant property of the functional, we get (6.12).

(b) Applying Taylor's formula (6.11) at point β, we get

$$f(x)=\sum_{w=0}^{n-1}\frac{(-1)^{w}f^{(w)}(\beta)}{w!}(x-\beta)^{w}-\frac{(-1)^{n-1}}{(n-1)!}\int_{\alpha}^{\beta}f^{(n)}(\xi)(\xi-x)_{+}^{n-1}d\xi \tag{6.25}$$

and follow similar steps as above we get (6.13).

(c) For fix $v=1$, testing (6.6) in Jensen's type functional $J_u(\cdot)$ and employing the linearity of $J_u(\cdot)$, we have

$$\begin{aligned}
J_u(f)&=f(\alpha)J_u\left(\frac{\beta-x}{\beta-\alpha}\right)+f(\beta)J_u\left(\frac{x-\alpha}{\beta-\alpha}\right)+\int_{\alpha}^{\beta}J_u(G_1(x,r))f''(r)dr\\
&=f(\alpha)\frac{J_u(\beta-x)}{\beta-\alpha}+f(\beta)\frac{J_u(x-\alpha)}{\beta-\alpha}+\int_{\alpha}^{\beta}J_u(G_1(x,r))\phi''(r)dr\\
&=\frac{1}{\beta-\alpha}\Big(f(\alpha)J_u(\beta)-f(\alpha)J_u(x)+f(\beta)J_u(x)-f(\beta)J_u(\alpha)\Big)\\
&\quad+\int_{\alpha}^{\beta}J_u(G_1(x,r))f''(r)dr\\
&=\frac{1}{\beta-\alpha}\Big(f(\beta)J_u(x)-f(\alpha)J_u(x)\Big)+\int_{\alpha}^{\beta}J_u(G_1(x,r))f''(r)dr.
\end{aligned} \tag{6.26}$$

Differentiating (6.11) twice and put $c=\alpha$ or replacing n by $(n-2)$ or utilizing (6.11) on the function f'' at the point α , we get

$$f''(r)=\sum_{w=2}^{n-1}\frac{f^{(w)}(\alpha)}{(w-2)!}(r-\alpha)^{w-2}+\frac{1}{(n-3)!}\int_{\alpha}^{r}f^{(n)}(\xi)(r-\xi)^{n-3}d\xi. \tag{6.27}$$

Now, using (6.27) in (6.26), we get

$$J_u(f) = \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right) J_u(x) + \int_\alpha^\beta J_u(G_1(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right) dr$$
$$+ \frac{1}{(n-3)!} \int_\alpha^\beta J_u(G_1(x,r)) \left(\int_\alpha^r f^{(n)}(\xi)(r-\xi)^{n-3} d\xi\right) dr.$$

Now applying Fubini's Theorem on second term gives (6.14) respectively for $v = 1$ and $u = 1, \cdots, 6$. The cases for $v = 2, 3, 4, 5$, are treated analogously.

(d) Differentiating (6.11) twice and now taking $c = \beta$, we get

$$f''(r) = \sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)}{(w-2)!}(r-\beta)^{w-2} - \frac{1}{(n-3)!} \int_r^\beta f^{(n)}(\xi)(r-\xi)^{n-3} d\xi. \tag{6.28}$$

Analogously, putting (6.28) in (6.26) and using Fubini's Theorem gives (6.19) respectively for $v = 1$ and $u = 1, \cdots, 6$. The cases for $v = 2, 3, 4, 5$, are treated analogously.

□

We now give generalizations of cyclic Jensen type linear functionals in discrete and integral cases for real weights.

Theorem 6.2 *Consider f be n-convex function along with the suppositions of Theorem 6.1. Then the following results hold:*

(a) If for all $u = 1, \ldots, 6$,

$$J_u((x-\xi)_+^{n-1}) \geq 0, \quad \xi \in [\alpha, \beta] \tag{6.29}$$

holds, then we have

$$J_u(f) \geq \sum_{w=1}^{n-1} \frac{f^{(w)}(\alpha)}{w!} J_u((x-\alpha)^w) \tag{6.30}$$

for $u = 1, \ldots, 6$.

(b) If for all $u = 1, \ldots, 6$,

$$(-1)^{n-1} J_u((\xi-x)_+^{n-1}) \leq 0, \ \xi \in [\alpha, \beta] \tag{6.31}$$

holds, then we have

$$J_u(f) \geq \sum_{w=1}^{n-1} \frac{(-1)^w f^{(w)}(\beta)}{w!} J_u((\beta-x)^w) \tag{6.32}$$

for $u = 1, \ldots, 6$.

(c) If for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$

$$\int_{\xi}^{\beta} J_u(G_v(x,r))(r-\xi)^{n-3}dr \geq 0, \;\; \xi \in [\alpha,\beta] \tag{6.33}$$

holds, then we have

$$J_u(f) \geq \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right)J_u(x) + \int_{\alpha}^{\beta} J_u(G_1(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right)dr, \tag{6.34}$$

$$J_u(f) \geq f'(\beta)J_u(x) + \int_{\alpha}^{\beta} J_u(G_2(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right)dr, \tag{6.35}$$

$$J_u(f) \geq -f'(\alpha)J_u(x) + \int_{\alpha}^{\beta} J_u(G_3(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right)dr, \tag{6.36}$$

$$J_u(f) \geq f'(\alpha)J_u(x) + \int_{\alpha}^{\beta} J_u(G_4(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right)dr, \tag{6.37}$$

$$J_u(f) \geq f'(\beta)J_u(x) + \int_{\alpha}^{\beta} J_u(G_5(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!}\right)dr. \tag{6.38}$$

(d) If for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$

$$\int_{\alpha}^{\xi} J_u(G_v(x,r))(r-\xi)^{n-3}dr \leq 0, \;\; \xi \in [\alpha,\beta] \tag{6.39}$$

holds, then we have

$$J_u(f) \geq \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right)J_u(x) + \int_{\alpha}^{\beta} J_u(G_1(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!}\right)dr, \tag{6.40}$$

$$J_u(\psi) \geq f'(\beta)J_u(x) + \int_{\alpha}^{\beta} J_u(G_2(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!}\right)dr, \tag{6.41}$$

$$J_u(\psi) \geq -f'(\alpha)J_u(x) + \int_{\alpha}^{\beta} J_u(G_3(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!}\right)dr, \tag{6.42}$$

$$J_u(\psi) \geq f'(\alpha)J_u(x) + \int_{\alpha}^{\beta} J_u(G_4(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!} \right) dr, \tag{6.43}$$

$$J_u(\psi) \geq f'(\beta)J_u(x) + \int_{\alpha}^{\beta} J_u(G_5(x,r)) \left(\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!} \right) dr. \tag{6.44}$$

Proof. We begin with the proof of (a) and its assumed conditions. Fix $u = 1,\ldots,6$.
By our assumption $f^{(n-1)}$ is absolutely-continuous on $[\alpha,\beta]$ as a result $f^{(n)}$ exists almost everywhere. Moreover, f is suppose to be n-convex, so by definition of n-convex function (see [82], p. 16), $f^{(n)}(x) \geq 0$ for almost everywhere on $[\alpha,\beta]$. Therefore by applying Theorem 6.1, we get (6.30).
Similarly, rest of the inequalities can be proved. □

We now give the final results of this section:

Theorem 6.3 *If the suppositions of Theorem 6.1 be fulfilled with the conditions that* $p_1,\ldots,p_m$ *and* $\lambda_1,\ldots,\lambda_k$ *be non negative tuples for* $2 \leq k \leq m$*, in such a way that* $\sum_{i=1}^{m} p_i = 1$ *and* $\sum_{j=1}^{k} \lambda_j = 1$*. Then for n-convex function* $f : [\alpha,\beta] \to \mathbb{R}$ *, we have:*

(a) (6.30) *is valid when* $n \geq 3$*. Besides, for function*

$$F_1(x) := \sum_{w=1}^{n-1} \frac{f^{(w)}(\alpha)}{w!}(x-\alpha)^w. \tag{6.45}$$

to be convex, the right side of (6.30) *is non negative, means*

$$J_u(f) \geq 0, \qquad u = 1,\ldots,6. \tag{6.46}$$

(b) For n even (6.32) *holds. Furthermore, for function*

$$F_2(x) := \sum_{w=1}^{n-1} \frac{(-1)^w f^{(w)}(\beta)}{w!}(\beta - x)^w. \tag{6.47}$$

to be convex, the right hand side of (6.32) *is non negative, particularly* (6.46) *holds.*

(c) Inequalities (6.34)-(6.38) *hold for all* $n \geq 3$*. Moreover, let* (6.34)-(6.38) *are valid and*

$$\sum_{w=2}^{n-1} \frac{f^{(w)}(\alpha)(r-\alpha)^{w-2}}{(w-2)!} \geq 0 \tag{6.48}$$

then, we get (6.46) *for every* $u = 1,\ldots,6$ *and* $v = 1,\ldots,5$.

(d) *If $n =$ even, then* (6.40)-(6.44) *hold. Moreover, let* (6.40)-(6.44) *are valid and*

$$\sum_{w=2}^{n-1} \frac{f^{(w)}(\beta)(r-\beta)^{w-2}}{(w-2)!} \geq 0 \tag{6.49}$$

then, we get (6.46) *for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$.*

Proof.

(a) Fix $u = 1,\ldots,6$.
For $(n \geq 3)$, $x \mapsto ((x-t)_+)^{n-1}$ is convex function, so (6.29) holds by virtue of Remark 2.7 on account of given weights to be positive. Hence (6.30) is establish by taking into account Theorem 6.2. Moreover, the R.H.S. of (6.30) can be written in the functional form $J_u(F_1)$ for all $(u = 1,\ldots,6)$, after reorganizing this side. Employing Remark 2.7 the nonnegativity of R.H.S. of (6.30) is secure, especially (6.46) is establish.

(b) Similar to the proof of (a).

(c) Fix $u = 1,\ldots,6$.
We have assumed positive weights and for all $v = 1,\ldots,5$, $G_v(x,r)$ is convex. Thus by practicing Remark 2.7, $J_u(G_v(x,r)) \geq 0$. As f is n-convex, so by using Theorem 6.2 (c), we get (6.34)-(6.38). Moreover the linear function z is convex(concave), therefore considering the positive weights, Remark 2.7 gives $0 \leq J_u(z) \geq 0$ implies $J_u(z) = 0$. Finally using positivity of $J_u(G_v(z,r))$ and (6.48), (6.46) is obtained.

(d) Similar to the proof of (c).

□

6.2 Applications to information theory

Now as a consequences of Theorem 6.2 we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $u = 1$, from (6.30) and (6.32)with respect to n-convex function f in the explicit form:

$$\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \geq$$

$$\sum_{w=1}^{n-1}\frac{f^{(w)}(\alpha)}{w!}\left(\sum_{i=1}^{m} p_i(x_i-\alpha)^w - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}} - \alpha\right)^w\right). \tag{6.50}$$

$$\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right) \geq$$

$$\sum_{w=1}^{n-1} \frac{(-1)^w f^{(w)}(\beta)}{w!} \left(\sum_{i=1}^{m} p_i (\beta - x_i)^w - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \left(\beta - \frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right)^w \right). \tag{6.51}$$

Theorem 6.4 *Let $m,k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1, \ldots, \lambda_k$ be positive probability distributions. Let $\mathbf{p} := (p_1, \ldots, p_m) \in \mathbb{R}^m$, and $\mathbf{q} := (q_1, \ldots, q_m) \in (0, \infty)^m$ such that*

$$\frac{p_i}{q_i} \in [\alpha, \beta], \quad i = 1, \ldots, m.$$

Also let $f : [\alpha, \beta] \to \mathbb{R}$ be a function such that $f^{(n-1)}$ is absolutely-continuous and f is n-convex function. Then the following inequalities hold:

(a)

$$\widetilde{I}_f(\mathbf{p}, \mathbf{q}) \geq \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)$$

$$+ \sum_{w=1}^{n-1} \frac{f^{(w)}(\alpha)}{w!} \left(\sum_{i=1}^{m} q_i \left(\frac{p_i}{q_i} - \alpha \right)^w - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} - \alpha \right)^w \right). \tag{6.52}$$

(b)

$$\widetilde{I}_f(\mathbf{p}, \mathbf{q}) \geq \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) + \sum_{w=1}^{n-1} \frac{(-1)^w f^{(w)}(\beta)}{w!} \times$$

$$\times \left(\sum_{i=1}^{m} q_i \left(\beta - \frac{p_i}{q_i} \right)^w - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \left(\beta - \frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)^w \right). \tag{6.53}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i = 1, \ldots, m)$ in (6.50) and (6.51), we get (6.52) and (6.53) respectively. □

We explore two exceptional cases of the previous results.
First one is corresponding to the entropy of a discrete probability distribution.

Corollary 6.1 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distributions.*

(a) If $\mathbf{q} := (q_1,\dots,q_m) \in (0,\infty)^m$ and n is even, then

$$\sum_{i=1}^{m} q_i \ln q_i \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)$$
$$+\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \tag{6.54}$$

(b) If $\mathbf{q} := (q_1,\dots,q_m)$ is a positive probability distribution and n is even, then we get the bounds for Shannon entropy of $\mathbf{q}$.

$$H(\mathbf{q}) \le -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)$$
$$-\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \tag{6.55}$$

If n is odd, then (6.54) *and* (6.55) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$, and $\mathbf{p} := (1,1,\dots,1)$ in Theorem 6.4 (a), we get the required results.

(b) It is a specific case of (a). □

The second case is corresponding to the relative entropy also known as Kullback-Leibler divergence between the two probability distributions.

Corollary 6.2 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distributions.*

(a) If $\mathbf{q} := (q_1,\dots,q_m), \mathbf{p} := (p_1,\dots,p_m) \in (0,\infty)^m$ and n is even, then

$$\sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$
$$+\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m} q_i\left(\frac{p_i}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \tag{6.56}$$

(b) *If If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m)$ are positive probability distributions and n is even, then we have*

$$D(\mathbf{q}\parallel\mathbf{p}) \geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)+$$

$$\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \tag{6.57}$$

If n is odd, then (6.56) *and* (6.57) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$ in Theorem 6.4 (a), we get the desired results.

(b) It is particular case of (a).

□

Suppose $m \in \{1,2,\ldots\}$, $t \geq 0$, $s > 0$, then **Zipf-Mandelbrot entropy** is given as :

$$Z(H,t,s) = \frac{s}{H_{m,t,s}}\sum_{i=1}^{m}\frac{\ln(i+t)}{(i+t)^s}+\ln(H_{m,t,s}). \tag{6.58}$$

Consider

$$q_i = f(i;m,t,s) = \frac{1}{((i+t)^sH_{m,t,s})}. \tag{6.59}$$

Now we state our results involving entropy introduced by Mandelbrot Law:

Theorem 6.5 *Let $m,k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distributions and $\mathbf{q}$ be as defined in* (6.59) *by Zipf-Mandelbrot law with parameters $m \in \{1,2,\ldots\}$, $t \geq 0$, $s > 0$. For n is even, the following holds*

$$\begin{aligned}H(\mathbf{q}) =& Z(H,t,s) \leq -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}\right)\ln\left(\frac{1}{H_{m,t,s}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s)}\right)\\ &-\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m}\frac{1}{((i+t)^sH_{m,t,s})}\left(((i+t)^sH_{m,t,s})-\alpha\right)^w\right)\\ &+\sum_{w=1}^{n-1}\frac{(-1)^w}{w.(\alpha)^w}\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}}-\alpha\right)^w\right).\end{aligned} \tag{6.60}$$

If n is odd, then (6.60) *holds in reverse direction.*

Proof. Substituting this $q_i = \frac{1}{((i+t)^s H_{m,t,s})}$ in Corollary 6.1(b), we get the desired result. Since it is interesting to see that $\sum_{i=1}^{m} q_i = 1$. Moreover using above q_i in Shannon entropy (3.3), we get Mandelbrot entropy(6.58)

$$\begin{aligned}
H(q) = -q_i \ln q_i &= -\sum_{i=1}^{m} \frac{1}{((i+t)^s H_{m,t,s})} \ln \frac{1}{((i+t)^s H_{m,t,s})} \\
&= \frac{-1}{(H_{m,t,s})} \sum_{i=1}^{m} \frac{1}{(i+t)^s} \ln \frac{1}{(i+t)^s H_{m,t,s}} \\
&= \frac{-1}{(H_{m,t,s})} \sum_{i=1}^{m} \frac{1}{(i+t)^s} \Big(\ln(1) - s\ln(i+t) - \ln(H_{m,t,s}) \Big) \\
&= \frac{1}{(H_{m,t,s})} \sum_{i=1}^{m} \frac{1}{(i+t)^s} \Big(s\ln(i+t) + \ln(H_{m,t,s}) \Big) \\
&= \frac{s}{(H_{m,t,s})} \sum_{i=1}^{m} \frac{\ln(i+t)}{(i+t)^s} + \ln(H_{m,t,s}).
\end{aligned} \tag{6.61}$$

□

Corollary 6.3 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distributions and for $t_1,t_2 \in [0,\infty)$, $s_1,s_2 > 0$, let $H_{m,t_1,s_1} = \frac{1}{(k+t_1)^{s_1}}$ and $H_{m,t_2,s_2} = \frac{1}{(k+t_2)^{s_2}}$. Now using $q_i = \dfrac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}}$ and $p_i = \dfrac{1}{(i+t_2)^{s_2} H_{m,t_2,s_2}}$ in Corollary 6.2(b), with n is even, then the following holds*

$$\begin{aligned}
D(\mathbf{q} \parallel \mathbf{p}) &= \sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \ln\left(\frac{(i+t_2)^{s_2} H_{m,t_2,s_2}}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \right) \\
&\ge \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}} \right) \ln \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}} \right) \\
&\quad + \sum_{w=1}^{n-1} \frac{(-1)^w}{w.(\alpha)^w} \left(\sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \left(\frac{(i+t_1)^{s_1} H_{m,t_1,s_1}}{(i+t_2)^{s_2} H_{m,t_2,s_2}} - \alpha \right)^w \right) \\
&\quad - \sum_{w=1}^{n-1} \frac{(-1)^w}{w.(\alpha)^w} \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}} - \alpha \right)^w \right)
\end{aligned} \tag{6.62}$$

If n is odd, then (6.62) *holds in reverse direction.*

Remark 6.2 It is interesting to note in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon and Mandelbrot entropies using the other inequalities for n-convex functions constructed in Theorem 6.2 for discrete case of cyclic refinements of Jensen's inequality.

Chapter 7

Cyclic Refinements of Jensen's Inequality Via Fink's Identity

In this Chapter, we give new extensions and improvements of cyclic refinements of Jensen's Inequality by A. M. Fink's identity with and without the effect of Green functions. As an application of our work we construct new entropic bounds for Shannon, Relative and Mandelbrot entropies. This chapter is based on the paper Mehmood, Butt, Horváth and Pečarić [78].

The following theorem is proved by A. M. Fink in [27].

Theorem 7.1 *Let $\alpha,\beta \in \mathbb{R}$, $f:[\alpha,\beta]\to\mathbb{R}$, $n\geq 1$ and $f^{(n-1)}$ is absolutely continuous on $[\alpha,\beta]$. Then*

$$\begin{aligned} f(x) = \; & \frac{n}{\beta-\alpha}\int_{\alpha}^{\beta} f(t)\,dt \\ & -\sum_{w=1}^{n-1}\left(\frac{n-w}{w!}\right)\left(\frac{f^{(w-1)}(\alpha)(x-\alpha)^{w}-f^{(w-1)}(\beta)(x-\beta)^{w}}{\beta-\alpha}\right) \\ & +\frac{1}{(n-1)!(\beta-\alpha)}\int_{\alpha}^{\beta}(x-t)^{n-1}\mu^{[\alpha,\beta]}(t,x)f^{(n)}(t)\,dt, \end{aligned} \tag{7.1}$$

where

$$\mu^{[\alpha,\beta]}(t,x)=\begin{cases} t-\alpha, & \alpha\leq t\leq x\leq\beta, \\ t-\beta, & \alpha\leq x<t\leq\beta. \end{cases} \tag{7.2}$$

7.1 Generalization of cyclic refinements of Jensen's inequality for $n-$convex functions Via A. M. Fink's identity

First, we consider the discrete as well as integral version of cyclic refinements of Jensen's inequality and construct the following identities having real weights with the help of Fink's identity.

Theorem 7.2 *Suppose $m,k \in \mathbb{N}$, $p_1,\ldots,p_m$ and $\lambda_1,\ldots,\lambda_k$ are real tuples for $2 \leq k \leq m$, such that $\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j} \neq 0$ for $i = 1,\ldots,m$ with $\sum_{i=1}^{m} p_i = 1$ and $\sum_{j=1}^{k} \lambda_j = 1$. Also let $x \in [\alpha,\beta] \subset \mathbb{R}$ and $\mathbf{x} \in [\alpha,\beta]^m$. Consider the function $f : [\alpha,\beta] \to \mathbb{R}$ such that $f^{(n-1)}$ is absolutely-continuous, $\mu^{[\alpha,\beta]}(t,x)$ and G_v, $(v = 1,\ldots,5)$ are same as given in* (7.2) *and* (6.1)–(6.5)*, respectively. Then for $(u = 1,\ldots,6)$ along with the assumptions (A_1) and (A_2), we have the following generalized identities:*

(i)

$$J_u(f) = \sum_{w=1}^{n-l} \left(\frac{n-w}{w!(\beta-\alpha)}\right) \left(f^{(w-1)}(\beta) J_u((x-\beta)^w) - f^{(w-1)}(\alpha) J_u((x-\alpha)^w)\right)$$
$$+ \frac{1}{(n-1)!(\beta-\alpha)} \int_{\alpha}^{\beta} J_u((x-t)^{n-1}\mu^{[\alpha,\beta]}(t,x)) f^{(n)}(t)dt. \tag{7.3}$$

(ii)

$$J_u(f) = C_v(\alpha,\beta,f)J_u(x) + (n-2)\left(\frac{f^{(1)}(\beta) - f^{(1)}(\alpha)}{\beta-\alpha}\right) \int_{\alpha}^{\beta} J_u(G_v(x,r))dr$$
$$+ \frac{1}{(\beta-\alpha)} \int_{\alpha}^{\beta} J_u(G_v(x,r)) \times$$
$$\times \left(\sum_{w=1}^{n-3} \left(\frac{n-2-w}{w!}\right) \left(f^{(w+1)}(\beta)(r-\beta)^w - f^{(w+1)}(\alpha)(r-\alpha)^w\right)\right) dr$$
$$+ \frac{1}{(n-3)!(\beta-\alpha)} \int_{\alpha}^{\beta} f^{(n)}(t) \left(\int_{\alpha}^{\beta} J_u(G_v(x,r))(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)dr\right) dt \tag{7.4}$$

where

$$\begin{aligned} C_1(\alpha,\beta,f) &= \left(\frac{f(\beta)-f(\alpha)}{\beta-\alpha}\right) \\ C_2(\alpha,\beta,f) &= f'(\beta) = C_5(\alpha,\beta,f) \\ C_3(\alpha,\beta,f) &= f'(\alpha) = -C_4(\alpha,\beta,f). \end{aligned}$$

Proof.

(i) Fix $u = 1,\ldots,6$.

Using (7.1) in Jensen's type functional $J_u(\cdot)$ and using the linearity of $J_u(\cdot)$, we have

$$
\begin{aligned}
J_u(f) =& J_u\left(\frac{n}{\beta-\alpha}\int\limits_{\alpha}^{\beta} f(t)dt\right) \\
&+\sum_{w=1}^{n-l}\left(\frac{n-w}{w!(\beta-\alpha)}\right) f^{(w-1)}(\beta)J_u((x-\beta)^w) \\
&-\sum_{w=1}^{n-l}\left(\frac{n-w}{w!(\beta-\alpha)}\right) f^{(w-1)}(\alpha)J_u((x-\alpha)^w) \\
&+\frac{1}{(n-1)!(\beta-\alpha)}\int\limits_{\alpha}^{\beta} J_u((x-t)^{n-1}\mu^{[\alpha,\beta]}(t,x))f^{(n)}(t)dt.
\end{aligned}
$$

After simplification, we get (7.3).

(ii) Fix $u = 1,\ldots,6$. For fix $v = 1$, testing (6.6) in Jensen's type functional $J_u(\cdot)$ and using the linearity of $J_u(\cdot)$, we have

$$
J_u(f) = C_1(\alpha,\beta,f)J_u(x) + \int\limits_{\alpha}^{\beta} J_u(G_1(x,r))f''(r)dr. \tag{7.5}
$$

Differentiating (7.1), twice with respect variable r, we get

$$
\begin{aligned}
f''(r) =& \sum_{w=0}^{n-3}\left(\frac{n-2-w}{w!}\right)\left(\frac{f^{(w+1)}(\beta)(r-\beta)^w - f^{(w+1)}(\alpha)(r-\alpha)^w}{\beta-\alpha}\right) \\
&+\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)f^{(n)}(t)dt \\
=& \sum_{w=1}^{n-2}\left(\frac{n-1-w}{(w-1)!}\right)\left(\frac{f^{(w)}(\beta)(r-\beta)^{w-1} - f^{(w)}(\alpha)(r-\alpha)^{w-1}}{\beta-\alpha}\right) \\
&+\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)f^{(n)}(t)dt \\
=& (n-2)\left(\frac{f^{(1)}(\beta)-f^{(1)}(\alpha)}{\beta-\alpha}\right) \\
&+\sum_{w=2}^{n-2}\left(\frac{n-1-w}{(w-1)!}\right)\left(\frac{f^{(w)}(\beta)(r-\beta)^{w-1} - f^{(w)}(\alpha)(r-\alpha)^{w-1}}{\beta-\alpha}\right) \\
&+\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)f^{(n)}(t)dt.
\end{aligned} \tag{7.6}
$$

Using (7.6) in (7.5) and applying Fubini's Theorem in the last term we get (7.4) for $v = 1$.

Alternatively, we use formula (7.1) for the function f'' and replace n by $n-2$ $(n \geq 3)$, to get

$$\begin{aligned} f''(r) = {} & (n-2)\left(\frac{f^{(1)}(\beta)-f^{(1)}(\alpha)}{\beta-\alpha}\right) \\ & +\sum_{w=1}^{n-3}\left(\frac{n-2-w}{w!}\right)\left(\frac{f^{(w+1)}(\beta)(r-\beta)^{w}-f^{(w+1)}(\alpha)(r-\alpha)^{w}}{\beta-\alpha}\right) \\ & +\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)f^{(n)}(t)\,dt \qquad (7.7) \\ = {} & (n-2)\left(\frac{f^{(1)}(\beta)-f^{(1)}(\alpha)}{\beta-\alpha}\right) \\ & +\sum_{w=2}^{n-2}\left(\frac{n-1-w}{(w-1)!}\right)\left(\frac{f^{(w)}(\beta)(r-\beta)^{w-1}-f^{(w)}(\alpha)(r-\alpha)^{w-1}}{\beta-\alpha}\right) \\ & +\frac{1}{(n-3)!(\beta-\alpha)}\int_{\alpha}^{\beta}(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)f^{(n)}(t)\,dt. \qquad (7.8) \end{aligned}$$

Now using (7.7) in (7.5) and applying Fubini's Theorem in the last term we get (7.4) for $v = 1$. The cases for $v = 2,3,4,5$, are treated analogously.

□

In the following theorems we obtain generalizations of discrete and integral Jensen type linear functionals, with real weights for $n-$convex functions.

Theorem 7.3 *Let all the assumptions of Theorem 7.2 be satisfied. Also let f be $n-$convex function such that $f^{(n-1)}$ is absolutely continuous. Then we have the following two results:*

(i) If

$$J_u((x-t)^{n-1}\mu^{[\alpha,\beta]}(t,x)) \geq 0, \quad t \in [\alpha,\beta] \tag{7.9}$$

holds, then we have

$$J_u(f) \geq \sum_{w=1}^{n-l}\left(\frac{n-w}{w!(\beta-\alpha)}\right)\left(f^{(w-1)}(\beta)J_u((x-\beta)^{w})-f^{(w-1)}(\alpha)J_u((x-\alpha)^{w})\right) \tag{7.10}$$

for $u = 1,\ldots,6$.

(ii) If for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$

$$\int_{\alpha}^{\beta} J_u(G_v(x,r))(r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r)dr \geq 0, \ t \in [\alpha,\beta], \tag{7.11}$$

holds, then we have

$$J_u(f) \geq C_v(\alpha,\beta,f)J_u(x) + (n-2)\left(\frac{f^{(1)}(\beta)-f^{(1)}(\alpha)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_v(x,r))dr$$

$$+\frac{1}{(\beta-\alpha)}\int_\alpha^\beta J_u(G_v(x,r))\times$$

$$\times\left(\sum_{w=1}^{n-3}\left(\frac{n-2-w}{w!}\right)\left(f^{(w+1)}(\beta)(r-\beta)^w - f^{(w+1)}(\alpha)(r-\alpha)^w\right)\right)dr \tag{7.12}$$

for $u = 1,\ldots,6$.

Proof.

(i) Similar to that of Theorem 6.2. □

Now, we will state the final results of this section with the following theorem:

Theorem 7.4 *Let all the assumptions of Theorem 7.2 be satisfied in addition with the condition that* $p_1,\ldots,p_m$ *and* $\lambda_1,\ldots,\lambda_k$ *be non negative tuples for* $3 \leq k \leq n$*, such that* $\sum_{i=1}^m p_i = 1$, $\sum_{j=1}^k \lambda_j = 1$ *and consider* $f : [\alpha,\beta] \to \mathbb{R}$ *is* $n-$*convex function.*

(i) If n be even and $n > 3$, then (7.10) holds.

(ii) Let the inequality (7.10) be satisfied. If the function

$$F(x) := \sum_{w=1}^{n-1}\left(\frac{n-w}{w!(\beta-\alpha)}\right)\left(f^{(w-1)}(\beta)(x-\beta)^w - f^{(w-1)}(\alpha)(x-\alpha)^w\right). \tag{7.13}$$

is convex, the R.H.S. of (7.10) is non negative and we have inequality

$$J_u(f) \geq 0, \qquad u = 1,\ldots,6. \tag{7.14}$$

(iii) If n be even and $n > 3$, then (7.12) holds.

(iv) Let the inequality (7.12) be satisfied and

$$\sum_{w=0}^{n-3}\left(\frac{n-2-w}{w!}\right)\left(f^{(w+1)}(\beta)(r-\beta)^w - f^{(w+1)}(\alpha)(r-\alpha)^w\right) \geq 0. \tag{7.15}$$

Then we have (7.14) for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$.

Proof.

(i) Fix $i = 1,\ldots 6$.
For

$$\vartheta(x) := (x-t)^{n-1}\mu^{[\alpha,\beta]}(t,x) = \begin{cases} (x-t)^{n-1}(t-\alpha), & \alpha \leq t \leq x \leq \beta, \\ (x-t)^{n-1}(t-\beta), & \alpha \leq x < t \leq \beta, \end{cases}$$

we have,

$$\vartheta''(x) := \begin{cases} (n-1)(n-2)(x-t)^{n-3}(t-\alpha), & \alpha \le t \le x \le \beta, \\ (n-1)(n-2)(x-t)^{n-3}(t-\beta), & \alpha \le x < t \le \beta, \end{cases}$$

showing that ϑ is convex for even n, where $n > 3$. Since the weights are non negative, so by virtue of Remark 2.7, (7.9) holds for even n, where $n > 3$. Therefore following Theorem 7.3 (i), we can obtain (7.10).

(ii) Similar to the proof of Theorem 6.3 (a).

(iii) Fix $i = 1, \ldots 6$ and $j = 1, \ldots 5$.
Since Green's function $G_v(x,r)$ is convex and the weights are positive. So $J_u(G_v(x,r)) \geq 0$ by virtue of Remark 2.7. Also, since

$$\vartheta(r) := (r-t)^{n-3}\mu^{[\alpha,\beta]}(t,r) = \begin{cases} (r-t)^{n-3}(t-\alpha), & \alpha \le t \le r \le \beta, \\ (r-t)^{n-3}(t-\beta), & \alpha \le r < t \le \beta, \end{cases}$$

ϑ is positive for even n, where $n > 3$. So, (7.11) holds for even n. Now following Theorem 7.3 (ii), we can obtain (7.12).

(iv) Using (7.15) in (7.12), we get (7.14). □

7.2 Applications to information theory

Now as a consequence of Theorem 7.3 we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $u = 1$, from (7.10) with respect to n-convex function f in the explicit form:

$$\begin{aligned}
&\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \geq \sum_{w=1}^{n-l}\left(\frac{(n-w)f^{(w-1)}(\beta)}{w!(\beta-\alpha)}\right) \times \\
&\times\left(\sum_{i=1}^{m} p_i(x_i-\beta)^w - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}} - \beta\right)^w\right) \\
&- \sum_{w=1}^{n-l}\left(\frac{(n-w)f^{(w-1)}(\alpha)}{w!(\beta-\alpha)}\right) \\
&\times\left(\sum_{i=1}^{m} p_i(x_i-\alpha)^w - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}} - \alpha\right)^w\right). \qquad (7.16)
\end{aligned}$$

Theorem 7.5 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution. Let $\mathbf{p} := (p_1,\dots,p_m) \in \mathbb{R}^m$, and $\mathbf{q} := (q_1,\dots,q_m) \in (0,\infty)^m$ such that*

$$\frac{p_i}{q_i} \in [\alpha,\beta], \quad u = 1,\dots,m.$$

Also let $f : [\alpha,\beta] \to \mathbb{R}$ be a function such that $f^{(n-1)}$ is absolutely-continuous and f is n-convex function. Then the following inequalities hold:

$$\begin{aligned}
\widetilde{I}_f(\mathbf{p},\mathbf{q}) &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\\
&+\sum_{w=1}^{n-l}\left(\frac{(n-w)f^{(w-1)}(\beta)}{w!(\beta-\alpha)}\right)\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\beta\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^w\right)\\
&-\sum_{w=1}^{n-l}\left(\frac{(n-w)f^{(w-1)}(\alpha)}{w!(\beta-\alpha)}\right)\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right).
\end{aligned} \tag{7.17}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i = 1,\dots,m)$ in (7.16), we get (7.17). □

We explore two exceptional cases of the previous result.
First one is corresponding to the entropy of a discrete probability distribution.

Corollary 7.1 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\dots,q_m) \in (0,\infty)^m$ and n is even, then

$$\begin{aligned}
\sum_{i=1}^{m} q_i \ln q_i &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\\
&+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\times\\
&\times\left(\sum_{i=1}^{m}q_i\left(\frac{1}{q_i}-\beta\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^w\right)\\
&-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)}\right)\times\\
&\times\left(\sum_{i=1}^{m}q_i\left(\frac{1}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right).
\end{aligned} \tag{7.18}$$

(b) If $\mathbf{q} := (q_1,\dots,q_m)$ is a positive probability distribution and n is even, then we get the bounds for Shannon entropy of $\mathbf{q}$.

$$H(\mathbf{q}) \le -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)$$
$$-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)\,(\beta)^{w-1}\,(\beta-\alpha)}\right)\times$$
$$\left(\sum_{i=1}^{m}q_i\left(\frac{1}{q_i}-\beta\right)^{w}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{w}\right)$$
$$+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)\,(\alpha)^{w-1}\,(\beta-\alpha)}\right)\times$$
$$\left(\sum_{i=1}^{m}q_i\left(\frac{1}{q_i}-\alpha\right)^{w}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{w}\right). \tag{7.19}$$

If n is odd, then (7.18) *and* (7.19) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$, and $\mathbf{p} := (1,1,\dots,1)$ in Theorem 7.5, we get the required result.

(b) It is a special case of (a). □

The second case is corresponding to the relative entropy also known as Kullback-Leibler divergence between two probability distributions.

Corollary 7.2 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\dots,q_m), \mathbf{p} := (p_1,\dots,p_m) \in (0,\infty)^m$ and n is even, then

$$\sum_{i=1}^{m}q_i\ln\left(\frac{q_i}{p_i}\right) \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$
$$+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)\,(\beta)^{w-1}\,(\beta-\alpha)}\right)\times$$
$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\beta\right)^{w}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{w}\right)$$

$$-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)}\right)\times$$

$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \quad (7.20)$$

(b) *If If* $\mathbf{q}:=(q_1,\ldots,q_m),\mathbf{p}:=(p_1,\ldots,p_m)$ *are positive probability distributions and* n *is even, then we have*

$$D(\mathbf{q}\,\|\,\mathbf{p})\geq\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$

$$+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\times$$

$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\beta\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^w\right)$$

$$-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)}\right)\times$$

$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\alpha\right)^w-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^w\right). \quad (7.21)$$

If n *is odd, then* (7.20) *and* (7.21) *hold in reverse directions.*

Proof.

(a) Using $f(x):=-\ln x$ in Theorem 7.5 (a), we get the desired result.

(b) It is particular case of (a).

□

Now we state our results involving entropy introduced by Mandelbrot Law:

Theorem 7.6 *Let $m,k\in\mathbb{N}$ $(2\le k\le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution and $\mathbf{q}$ be as defined in (6.59) by Zipf-Mandelbrot law with parameters $m\in\{1,2,\ldots\}$, $t\ge 0$, $s>0$. For n is even, the following holds*

$$\begin{aligned}
H(\mathbf{q}) = Z(H,t,s) \le &-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s H_{m,t,s})}\right)\ln\left(\frac{1}{H_{m,t,s}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s)}\right)\\
&-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+t)^s H_{m,t,s})}\Big(((i+t)^s H_{m,t,s})-\beta\Big)^w\right)\\
&+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s H_{m,t,s})}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s H_{m,t,s})}}-\beta\right)^w\right)\\
&+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+t)^s H_{m,t,s})}\Big(((i+t)^s H_{m,t,s})-\alpha\Big)^w\right)\\
&-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)}\right)\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s H_{m,t,s})}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s H_{m,t,s})}}-\alpha\right)^w\right).
\end{aligned}\tag{7.22}$$

If n is odd, then (7.22) holds in reverse direction.

Proof. Similar to that of Theorem 6.5. □

Corollary 7.3 *Let $m,k\in\mathbb{N}$ $(2\le k\le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distributions and for $t_1,t_2\in[0,\infty)$, $s_1,s_2>0$, let $H_{m,t_1,s_1}=\sum_{k=1}^{m}\frac{1}{(k+t_1)^{s_1}}$ and $H_{m,t_2,s_2}=\sum_{k=1}^{m}\frac{1}{(k+t_2)^{s_2}}$. Now using $q_i=\dfrac{1}{(i+t_1)^{s_1}H_{m,t_1,s_1}}$ and $p_i=\dfrac{1}{(i+t_2)^{s_2}H_{m,t_2,s_2}}$ in Corollary 7.2(b), with n is even, then the following holds*

$$\begin{aligned}
D(\mathbf{q}\,\|\,\mathbf{p}) &= \sum_{i=1}^{m}\frac{1}{(i+t_1)^{s_1}H_{m,t_1,s_1}}\ln\left(\frac{(i+t_2)^{s_2}H_{m,t_2,s_2}}{(i+t_1)^{s_1}H_{m,t_1,s_1}}\right)\\
&\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+t_1)^{s_1}H_{m,t_1,s_1}}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+t_1)^{s_1}H_{m,t_1,s_1}}}{\sum_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+t_2)^{s_2}H_{m,t_2,s_2}}}\right)\\
&+\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\left(\sum_{i=1}^{m}\frac{1}{(i+t_1)^{s_1}H_{m,t_1,s_1}}\left(\frac{(i+t_1)^{s_1}H_{m,t_1,s_1}}{(i+t_2)^{s_2}H_{m,t_2,s_2}}-\beta\right)^w\right)\\
&-\sum_{w=1}^{n-l}\left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\beta)^{w-1}(\beta-\alpha)}\right)\times
\end{aligned}$$

$$
\times \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}} - \beta \right)^{w} \right)
$$

$$
- \sum_{w=1}^{n-l} \left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)} \right) \left(\sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \left(\frac{(i+t_1)^{s_1} H_{m,t_1,s_1}}{(i+t_2)^{s_2} H_{m,t_2,s_2}} - \alpha \right)^{w} \right)
$$

$$
+ \sum_{w=1}^{n-l} \left(\frac{(n-w)(-1)^{(w-1)}}{w(w-1)(\alpha)^{w-1}(\beta-\alpha)} \right) \times
$$

$$
\times \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}} - \alpha \right)^{w} \right). \tag{7.23}
$$

If n is odd, then (7.23) *holds in reverse direction.*

Remark 7.1 It is interesting to note in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon and Mandelbrot entropies using the other inequalities for n-convex functions constructed in Theorem 7.3 for discrete case of cyclic refinements of Jensen's inequality.

Chapter 8

Cyclic Refinements of Jensen's Inequality VIA Montgomery's Identity

In this Chapter, we give new extensions and improvements of cyclic refinements of Jensen's Inequality by Montgomery's identity with and without the effect of Green functions. As an application of our work we construct new entropic bounds for Shannon, Relative and Mandelbrot entropies. This chapter is based on the paper [69].

In order to obtain our main results in this chapter, we use the generalized Montgomery identity via Taylor's formula given in paper [3].

Theorem 8.1 *Let $n \in \mathbb{N}$, $f : I \to \mathbb{R}$ be such that $f^{(n-1)}$ is absolutely continuous, $I \subset \mathbb{R}$ an open interval, $\alpha, \beta \in I$, $\alpha < \beta$. Then the following identity holds*

$$\begin{aligned} f(x) =& \frac{1}{\beta-\alpha}\int_{\alpha}^{\beta} f(v)\,dv + \sum_{l=0}^{n-2} \frac{f^{(l+1)}(\alpha)}{l!\,(l+2)} \frac{(x-\alpha)^{l+2}}{\beta-\alpha} \\ & - \sum_{l=0}^{n-2} \frac{f^{(l+1)}(\beta)}{l!\,(l+2)} \frac{(x-\beta)^{l+2}}{\beta-\alpha} + \frac{1}{(n-1)!}\int_{\alpha}^{\beta} R_n(x,v)\, f^{(n)}(v)\,dv \end{aligned} \tag{8.1}$$

where

$$R_n(x,v)=\begin{cases}-\frac{(x-v)^n}{n(\beta-\alpha)}+\frac{x-\alpha}{\beta-\alpha}(x-v)^{n-1}, & \alpha\le v\le x,\\ -\frac{(x-v)^n}{n(\beta-\alpha)}+\frac{x-\beta}{\beta-\alpha}(x-v)^{n-1}, & x<v\le\beta.\end{cases} \tag{8.2}$$

In case $n=1$ the sum $\sum_{l=0}^{n-2}\cdots$ is empty, so identity (8.1) reduces to well-known Montgomery identity (see for instance [74])

$$f(x)=\frac{1}{\beta-\alpha}\int_\alpha^\beta f(v)\,dv+\int_\alpha^\beta P(x,s)f'(v)\,dv$$

where $P(x,v)$ is the Peano kernel, defined by

$$P(x,v)=\begin{cases}\frac{v-\alpha}{\beta-\alpha}, & \alpha\le v\le x,\\ \frac{v-\beta}{\beta-\alpha}, & x<v\le\beta.\end{cases}$$

Remark 8.1 It is important to note that for n even, $R_n\ge 0$ defined in (8.2).

8.1 Generalization of cyclic refinements of Jensen's inequality by Montgomery identity

First, we consider the discrete as well as continuous version of cyclic refinements of Jensen's inequality and construct the following identities having real weights utilizing Montgomery's identity.

Theorem 8.2 *Suppose $m,k\in\mathbb{N}$, $p_1,\ldots,p_m$ and $\lambda_1,\ldots,\lambda_k$ are real tuples for $2\le k\le m$, such that $\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\neq 0$ for $i=1,\ldots,m$ with $\sum_{i=1}^{m}p_i=1$ and $\sum_{j=1}^{k}\lambda_j=1$. Also let $x\in[\alpha,\beta]\subset\mathbb{R}$ and $\mathbf{x}\in[\alpha,\beta]^m$. Consider the function $f:[\alpha,\beta]\to\mathbb{R}$ such that $f^{(n-1)}$ is absolutely-continuous, $R_n(\cdot,\xi)$, G_v, $(v=1,\ldots,5)$ are the same as given in* (8.2) *and* (6.1)–(6.5)*, respectively. Then for $(u=1,\ldots,6)$ along with the assumptions (A_1) and (A_2), we have the following generalized identities:*

(a)

$$J_u(f)=\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{1}{l!(l+2)}\right)\left(f^{(l+1)}(\alpha)J_u((x-\alpha)^{l+2})-f^{(l+1)}(\beta)J_u((x-\beta)^{l+2})\right)$$
$$+\frac{1}{(n-1)!}\int_\alpha^\beta J_u(R_n(x,\xi))f^{(n)}(\xi)\,d\xi, \tag{8.3}$$

(b)

$$J_u(f) = C_v(\alpha,\beta,f)J_u(x) + \left(\frac{f'(\alpha)-f'(\beta)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_v(x,r))dr$$

$$+\frac{1}{\beta-\alpha}\int_\alpha^\beta J_u(G_v(x,r))\left(\sum_{l=2}^{n-1}\frac{l}{(l-1)!}\left(f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}\right)\right)dr$$

$$+\frac{1}{(n-3)!}\int_\alpha^\beta f^{(n)}(\xi)(\int_\alpha^\beta J_u(G_v(x,r))\tilde{R}_{n-2}(r,\xi)\,dr)d\xi, \tag{8.4}$$

where C_v, $(v=1,\ldots,5)$ is defined in (7.5),

$$\tilde{R}_{n-2}(r,\xi) = \begin{cases} \frac{1}{\beta-\alpha}\left[\frac{(r-\xi)^{n-2}}{(n-2)}+(r-\alpha)(r-\xi)^{n-3}\right], & \alpha\le\xi\le r,\\ \frac{1}{\beta-\alpha}\left[\frac{(r-\xi)^{n-2}}{(n-2)}+(r-\beta)(r-\xi)^{n-3}\right], & r<\xi\le\beta,\end{cases}$$

and

$$J_u(f) = C_v(\alpha,\beta,f)J_u(x) + \left(\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_v(x,r))dr$$

$$+\frac{1}{\beta-\alpha}\int_\alpha^\beta J_u(G_v(x,r))\left(\sum_{l=3}^{n-1}\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}}{(l-3)!(l-1)}\right)dr$$

$$+\frac{1}{(n-3)!}\int_\alpha^\beta f^{(n)}(\xi)\left(\int_\alpha^\beta J_u(G_v(x,r))R_{n-2}(r,\xi)dr\right)d\xi. \tag{8.5}$$

Proof.

(a) Fix $u=1,\ldots,6$.

Using Montgomery identity (8.1) in cyclic Jensen type linear functional $J_u(\cdot)$ and practicing its properties, we have

$$J_u(f) = J_u\left(\frac{1}{\beta-\alpha}\int_\alpha^\beta f(\xi)d\xi\right) + \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{1}{l!(l+2)}\right)f^{(l+1)}(\alpha)J_u((x-\alpha)^{l+2})$$

$$-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{1}{l!(l+2)}\right)f^{(l+1)}(\beta)J_u((x-\beta)^{l+2})$$

$$+\frac{1}{(n-1)!}\int_\alpha^\beta J_u(R_n(x,\xi))f^{(n)}(\xi)d\xi.$$

Using the constant property of the functionals, we get (8.3).

(b) Fix $u=1,\ldots,6$. For fix $v=1$, testing (6.6) in cyclic Jensen type functional $J_u(\cdot)$ and employing the linearity of $J_u(\cdot)$, we have

$$J_u(f) = C_1(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u(G_1(x,r))f''(r)dr. \tag{8.6}$$

Differentiating Montgomery identity (8.1) twice with respect to the first variable, we have

$$f''(r) = \frac{f'(\alpha)-f'(\beta)}{\beta-\alpha} + \sum_{l=2}^{n-1}\left(\frac{l}{(l-1)!}\right)\left(\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}}{\beta-\alpha}\right)$$
$$+\frac{1}{(n-3)!}\int_\alpha^\beta \tilde{R}_{n-2}(r,\xi)f^{(n)}(\xi)d\xi. \tag{8.7}$$

Using (8.7) in (8.6), we get

$$J_u(f) = C_1(\alpha,\beta,f)J_u(x) + \left(\frac{f'(\alpha)-f'(\beta)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_1(x,r))dr$$
$$+\sum_{l=2}^{n-1}\frac{l}{(l-1)!}\int_\alpha^\beta J_u(G_1(x,r))\left(\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}}{\beta-\alpha}\right)dr$$
$$+\frac{1}{(n-3)!}\int_\alpha^\beta J_u(G_1(x,r))\Big(\int_\alpha^\beta \tilde{R}_{n-2}(r,\xi)f^{(n)}(\xi)d\xi\Big)dr.$$

By executing Fubini's Theorem in the last term, we have (8.4) for $u=1,\cdots,6$, respectively.

Next, using formula (8.1) on the function f'', replacing n by $n-2$ $(n\geq 3)$ and rearranging the indices, we have

$$f''(r) = \left(\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\right)$$
$$+\sum_{l=3}^{n-1}\left(\frac{1}{(l-3)!(l-1)}\right)\left(\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}}{\beta-\alpha}\right) \tag{8.8}$$
$$+\frac{1}{(n-3)!}\int_\alpha^\beta R_{n-2}(r,\xi)f^{(n)}(\xi)d\xi.$$

Similarly, using (8.8) in (8.6) and employing Fubini's Theorem, we get (8.5) for $u=1,\cdots,6$, respectively.

The cases for $v=2,3,4,5$ are treated analogously. □

Now we obtain generalizations of discrete and integral cyclic Jensen type linear functionals, with real weights.

Theorem 8.3 *Under the suppositions of Theorem 8.2, let f be $n-$convex function. Then we conclude the following results:*

(a) If

$$J_u(R_n(x,\xi)) \geq 0, \quad \xi \in [\alpha,\beta] \tag{8.9}$$

holds, then we have

$$J_u(f) \geq \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{1}{l!(l+2)}\right)\Big(f^{(l+1)}(\alpha)J_u((x-\alpha)^{l+2}) - f^{(l+1)}(\beta)J_u((x-\beta)^{l+2})\Big) \tag{8.10}$$

for $u = 1,\ldots,6$.

(b) If for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$

$$\int_\alpha^\beta J_u(G_v(x,r))\tilde{R}_{n-2}(r,\xi)\,dr \geq 0, \ \ \xi \in [\alpha,\beta], \tag{8.11}$$

holds, then we have

$$J_u(f) \geq C_v(\alpha,\beta,f)J_u(x) + \left(\frac{f'(\alpha)-f'(\beta)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_v(x,r))dr$$
$$+\frac{1}{\beta-\alpha}\int_\alpha^\beta J_u(G_v(x,r))\times\left(\sum_{l=2}^{n-1}\frac{l}{(l-1)!}\left(f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}\right)\right)dr \tag{8.12}$$

for $u = 1,\ldots,6$, and if

$$\int_\alpha^\beta J_u(G_v(x,r))R_{n-2}(r,\xi)\,dr \geq 0, \ \ \xi \in [\alpha,\beta], \tag{8.13}$$

holds, then we have

$$J_u(f) \geq C_v(\alpha,\beta,f)J_u(x) + \left(\frac{f'(\beta)-f'(\alpha)}{\beta-\alpha}\right)\int_\alpha^\beta J_u(G_v(x,r))dr$$
$$+\frac{1}{\beta-\alpha}\int_\alpha^\beta J_u(G_v(x,r))\times\left(\sum_{l=3}^{n-1}\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1}-f^{(l)}(\beta)(r-\beta)^{l-1}}{(l-3)!(l-1)}\right)dr \tag{8.14}$$

for $u = 1,\ldots,6$.

Proof. Similar to that of Theorem 6.2. □

We will finish the present section by stating the following theorem:

Theorem 8.4 *If the assumptions of Theorem 8.2 be fulfilled with additional conditions that $p_1,\dots,p_m$ and $\lambda_1,\dots,\lambda_k$ be non negative tuples for $3 \le k \le n$, such that $\sum_{i=1}^{m} p_i = 1$, $\sum_{j=1}^{k} \lambda_j = 1$. Then for $f:[\alpha,\beta] \to \mathbb{R}$ being $n-$convex function, we conclude the following results:*

(a) For even $n \ge 4$, (8.10) holds.

(b) If inequality (8.10) is valid and the function

$$F(x) = \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{f^{(l+1)}(\alpha)(x-\alpha)^{l+2} - f^{(l+1)}(\beta)(x-\beta)^{l+2}}{l!(l+2)}\right) \tag{8.15}$$

is convex. Then the inequality

$$J_u(f) \ge 0, \qquad u = 1,\dots,6. \tag{8.16}$$

(c) For even $n \ge 4$, (8.12) and (8.14) holds.

(d) If the inequality (8.12) is true and

$$\sum_{l=1}^{n-1}\frac{l}{(l-1)!}\left(f^{(l)}(\alpha)(r-\alpha)^{l-1} - f^{(l)}(\beta)(r-\beta)^{l-1}\right) \ge 0; \quad \forall r \in [\alpha,\beta], \tag{8.17}$$

OR

(8.14) be satisfied and

$$f'(\beta) - f'(\alpha) + \sum_{l=3}^{n-1}\frac{f^{(l)}(\alpha)(r-\alpha)^{l-1} - f^{(l)}(\beta)(r-\beta)^{l-1}}{(l-3)!(l-1)} \ge 0; \quad \forall r \in [\alpha,\beta]. \tag{8.18}$$

Then we have (8.16) for all $u = 1,\dots,6$ and $v = 1,\dots,5$.

Proof.

(a) Fix $u = 1,\dots,6$.
Since

$$R_n(x,\xi) = \begin{cases} -\frac{(x-\xi)^n}{n(\beta-\alpha)} + \frac{x-\alpha}{\beta-\alpha}(x-\xi)^{n-1}, & \alpha \le \xi \le x \le \beta, \\ -\frac{(x-\xi)^n}{n(\beta-\alpha)} + \frac{x-\beta}{\beta-\alpha}(x-\xi)^{n-1}, & \alpha \le x < \xi \le \beta. \end{cases}$$

So $\frac{d}{dx}R_n(x,\xi) =$

$$\begin{cases} \frac{1}{\beta-\alpha}\left[-(x-\xi)^{n-1} + (x-\xi)^{n-1} + (n-1)(x-\alpha)(x-\xi)^{n-2}\right], & \xi \le x \le \beta, \\ \frac{1}{\beta-\alpha}\left[-(x-\xi)^{n-1} + (x-\xi)^{n-1} + (n-1)(x-\beta)(x-\xi)^{n-2}\right], & x < \xi \le \beta. \end{cases}$$

and

$$\frac{d^2}{dx^2} R_n(x,\xi) = \begin{cases} \frac{n-1}{\beta-\alpha}\left[(x-\xi)^{n-2} + (n-2)(x-\alpha)(x-\xi)^{n-3}\right], & \xi \le x \le \beta, \\ \frac{n-1}{\beta-\alpha}\left[(x-\xi)^{n-2} + (n-2)(x-\beta)(x-\xi)^{n-3}\right], & x < \xi \le \beta. \end{cases}$$

Applying second derivative test on $R_n(\cdot,\xi)$, it can be seen easily that it is convex for even $n \ge 4$. Since the weights are nonnegative, so by advantage of Remark 2.7, (8.9) holds. Pursuing Theorem 8.3 (a), (8.10) is evident.

(b) Similar to the proof of Theorem 6.3 (a).

(c) Fix $u = 1, \ldots 6$ and $v = 1, \ldots 5$.
Since, we have assumed nonnegative tuples and the Green's function $G_v(x,r)$ is convex for all $v = 1, \ldots 5$. Thus by practicing Remark 2.7, $J_u(G_v(x,r)) \ge 0$. Moreover $\tilde{R}_{n-2}(r,\xi) \ge 0$ and $R_{n-2}(r,\xi) \ge 0$ for $n = 4, 6, \ldots$, so (8.11) and (8.13) hold. As f is $n-$convex, hence by following Theorem 8.3 (ii), we obtain (8.12) and (8.14) respectively.

(d) Utilizing (8.17) in (8.12) and (8.18) in (8.14), (8.16) is established for all $u = 1, \ldots, 6$.

□

8.2 Applications to information theory

Now as a consequence of Theorem 8.3 we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $u = 1$, from (8.10) with respect to n-convex function f in the explicit form:

$$\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right)$$
$$\ge \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{f^{(l+1)}(\alpha)}{l!(l+2)}\right) \times$$
$$\times \left(\sum_{i=1}^{m} p_i (x_i-\alpha)^{l+2} - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} - \alpha\right)^{l+2}\right)$$

$$-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{f^{(l+1)}(\beta)}{l!(l+2)}\right)\times$$
$$\times\left(\sum_{i=1}^{m}p_i(x_i-\beta)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}-\beta\right)^{l+2}\right). \tag{8.19}$$

Theorem 8.5 *Let $m,k\in\mathbb{N}$ $(2\le k\le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution. Let $\mathbf{p}:=(p_1,\dots,p_m)\in\mathbb{R}^m$, and $\mathbf{q}:=(q_1,\dots,q_m)\in(0,\infty)^m$ such that*

$$\frac{p_i}{q_i}\in[\alpha,\beta],\quad i=1,\dots,m.$$

Also let $f:[\alpha,\beta]\to\mathbb{R}$ be a function such that $f^{(n-1)}$ is absolutely-continuous and f is n-convex function. Then the following inequalities hold:

$$\widetilde{I}_f(\mathbf{p},\mathbf{q})\ge\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)$$
$$+\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{f^{(l+1)}(\alpha)}{l!(l+2)}\right)\times$$
$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\alpha\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{l+2}\right)$$
$$-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{f^{(l+1)}(\beta)}{l!(l+2)}\right)\times$$
$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\beta\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{l+2}\right). \tag{8.20}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i=1,\dots,m)$ in (8.19), we get (8.20). □

We explore two exceptional cases of the previous results.
First one is corresponding to the entropy of a discrete probability distribution.

Corollary 8.1 *Let $m,k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m) \in (0,\infty)^m$ and n is even, then

$$\begin{aligned}
\sum_{i=1}^{m} q_i \ln q_i \geq & \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) \\
& + \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\times \\
& \times\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\alpha\right)^{l+2} - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{l+2}\right) \\
& - \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\times \\
& \times\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\beta\right)^{l+2} - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{l+2}\right).
\end{aligned} \tag{8.21}$$

(b) If $\mathbf{q} := (q_1,\ldots,q_m)$ is a positive probability distribution and n is even, then we get the bounds for Shannon entropy of $\mathbf{q}$.

$$\begin{aligned}
H(\mathbf{q}) \leq & -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) \\
& - \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\times \\
& \times\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\alpha\right)^{l+2} - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{l+2}\right) \\
& + \frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\times \\
& \times\left(\sum_{i=1}^{m} q_i\left(\frac{1}{q_i}-\beta\right)^{l+2} - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{l+2}\right).
\end{aligned} \tag{8.22}$$

If n is odd, then (8.21) *and* (8.22) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$, and $\mathbf{p} := (1,1,\ldots,1)$ in Theorem 8.5, we get the required result.

(b) It is a special case of (a). □

The second case is corresponding to the relative entropy also known as Kullback-Leibler divergence between two probability distributions.

Corollary 8.2 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m) \in (0,\infty)^m$ and n is even, then

$$\begin{aligned}
\sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&+\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\times \\
&\times\left(\sum_{i=1}^{m} q_i\left(\frac{p_i}{q_i}-\alpha\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{l+2}\right) \\
&-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\times \\
&\times\left(\sum_{i=1}^{m} q_i\left(\frac{p_i}{q_i}-\beta\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{l+2}\right).
\end{aligned} \tag{8.23}$$

(b) If If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m)$ are positive probability distributions and n is even, then we have

$$\begin{aligned}
D(\mathbf{q}\parallel\mathbf{p}) &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&+\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\times \\
&\times\left(\sum_{i=1}^{m} q_i\left(\frac{p_i}{q_i}-\alpha\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha\right)^{l+2}\right)
\end{aligned}$$

$$-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\times$$
$$\times\left(\sum_{i=1}^{m}q_i\left(\frac{p_i}{q_i}-\beta\right)^{l+2}-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\beta\right)^{l+2}\right). \tag{8.24}$$

If n is odd, then (8.23) *and* (8.24) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$ in Theorem 8.5 (a), we get the desired result.

(b) It is particular case of (a). □

Now we state our results involving entropy introduced by Mandelbrot Law:

Theorem 8.6 *Let* $m,k\in\mathbb{N}$ $(2\le k\le m)$*,* $\lambda_1,\ldots,\lambda_k$ *be positive probability distribution and* $\mathbf{q}$ *be as defined in* (6.59) *by Zipf-Mandelbrot law with parameters* $m\in\{1,2,\ldots\}$*,* $t\ge 0$*,* $s>0$*. For n is even, the following holds*

$$\begin{aligned}
H(\mathbf{q})=Z(H,t,s)\le&-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}\right)\ln\left(\frac{1}{H_{m,t,s}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s)}\right)\\
&-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+t)^sH_{m,t,s})}\left(((i+t)^sH_{m,t,s})-\alpha\right)^{l+2}\right)\\
&+\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\alpha)^{l+1}}\right)\times\\
&\times\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}}-\alpha\right)^{l+2}\right)\\
&+\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+t)^sH_{m,t,s})}\left(((i+t)^sH_{m,t,s})-\beta\right)^{l+2}\right)\\
&-\frac{1}{\beta-\alpha}\sum_{l=0}^{n-2}\left(\frac{(-1)^{(l+1)}}{(l+2)(\beta)^{l+1}}\right)\times\\
&\times\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}\right)\left(\frac{1}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^sH_{m,t,s})}}-\beta\right)^{l+2}\right).
\end{aligned}\tag{8.25}$$

If n is odd, then (8.25) *holds in reverse direction.*

Proof. Similar to that of Theorem 6.5. □

Corollary 8.3 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distributions and for $t_1,t_2 \in [0,\infty)$, $s_1,s_2 > 0$, let $H_{m,t_1,s_1} = \sum_{k=1}^{m} \frac{1}{(k+t_1)^{s_1}}$ and $H_{m,t_2,s_2} = \sum_{k=1}^{m} \frac{1}{(k+t_2)^{s_2}}$. Now using $q_i = \dfrac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}}$ and $p_i = \dfrac{1}{(i+t_2)^{s_2} H_{m,t_2,s_2}}$ in Corollary 8.2(b), with n is even, then the following holds*

$$
\begin{aligned}
D(\mathbf{q}\,\|\,\mathbf{p}) &= \sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \ln\left(\frac{(i+t_2)^{s_2} H_{m,t_2,s_2}}{(i+t_1)^{s_1} H_{m,t_1,s_1}}\right) \\
&\ge \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}\right) \ln\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}\right) \\
&\quad + \frac{1}{\beta-\alpha} \sum_{l=0}^{n-2} \left(\frac{(-1)^{(l+1)}}{(l+2)\,(\alpha)^{l+1}}\right) \left(\sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \left(\frac{(i+t_1)^{s_1} H_{m,t_1,s_1}}{(i+t_2)^{s_2} H_{m,t_2,s_2}} - \alpha\right)^{l+2}\right) \\
&\quad - \frac{1}{\beta-\alpha} \sum_{l=0}^{n-2} \left(\frac{(-1)^{(l+1)}}{(l+2)\,(\alpha)^{l+1}}\right) \times \\
&\quad \times \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}\right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}} - \alpha\right)^{l+2}\right) \\
&\quad - \frac{1}{\beta-\alpha} \sum_{l=0}^{n-2} \left(\frac{(-1)^{(l+1)}}{(l+2)\,(\beta)^{l+1}}\right) \left(\sum_{i=1}^{m} \frac{1}{(i+t_1)^{s_1} H_{m,t_1,s_1}} \left(\frac{(i+t_1)^{s_1} H_{m,t_1,s_1}}{(i+t_2)^{s_2} H_{m,t_2,s_2}} - \beta\right)^{l+2}\right) \\
&\quad + \frac{1}{\beta-\alpha} \sum_{l=0}^{n-2} \left(\frac{(-1)^{(l+1)}}{(l+2)\,(\beta)^{l+1}}\right) \times \\
&\quad \times \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}\right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_2)^{s_2} H_{m,t_2,s_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+t_1)^{s_1} H_{m,t_1,s_1}}} - \beta\right)^{l+2}\right).
\end{aligned} \tag{8.26}
$$

If n is odd, then (8.26) *holds in reverse direction.*

Remark 8.2 It is interesting to note in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon and Mandelbrot entropies using the other inequalities for n-convex functions constructed in Theorem 8.3 for discrete case of cyclic refinements of Jensen's inequality.

Chapter 9

Cyclic Refinements of Jensen's Inequality by Lidstone Interpolating Polynomial

In the present Chapter , we generalize Cyclic Refinements of Jensen's inequality for higher order convex functions using Lidstone interpolating polynomial with applications in information theory. This chapter is based on the paper [70].
To move on, we consider Lidstone series, a generalization of the Taylor series, approximating a given function in the neighborhood of two points instead of one by using the even derivatives. Such series have been studied by G. J. Lidstone (1929), H. Poritsky (1932), J. M. Wittaker (1934) and others (see [1, 6]). Widder proved the fundamental lemma:

Lemma 9.1 *[94] If $f \in C^{2n}[0,1]$, then*

$$f(x) = \sum_{l=0}^{n-1} \left[f^{(2l)}(0)P_l(1-x) + f^{(2l)}(1)P_l(x) \right] + \int_0^1 G_{(n)}(x,r) f^{(2n)}(r)\,dr,$$

where P_n is a Lidstone's polynomial of degree $(2n+1)$ defined by the relations

$$P_0(x) = x$$
$$P_n''(x) = P_{n-1}(x)$$
$$P_n(0) = P_n(1) = 0, \quad n \geq 1$$

and

$$G_{(1)}(x,r) = G(x,r) = \begin{cases} (x-1)r\,, & r \leq x, \\ (r-1)x, & x \leq r, \end{cases} \tag{9.1}$$

is homogeneous Green function of the differential operator $\frac{d^2}{dr^2}$ *on* $[0,1]$*, and with the successive iterates of* $G(x,r)$

$$G_{(n)}(x,r) = \int_0^1 G_{(1)}(x,s)G_{(n-1)}(s,r)ds, \quad n \geq 2. \tag{9.2}$$

The Lidstone's polynomial can be expressed in terms of $G_{(n)}(x,r)$ *as*

$$P_n(x) = \int_0^1 G_{(n)}(x,r)rdr. \tag{9.3}$$

9.1 Extensions of cyclic refinements of Jensen's inequality by Lidstone interpolating polynomial

We propose the following Lemma in which we construct the generalized identities having real weights utilizing Lidstone's interpolating polynomial and Green functions.

Lemma 9.2 *Suppose* $m,k \in \mathbb{N}$*,* $p_1,\ldots,p_m$ *and* $\lambda_1,\ldots,\lambda_k$ *are real tuples for* $2 \leq k \leq m$*, such that* $\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j} \neq 0$ *for* $i = 1,\ldots,m$ *with* $\sum_{i=1}^{m} p_i = 1$ *and* $\sum_{j=1}^{k} \lambda_j = 1$*. Also let* $x \in [\alpha,\beta] \subset \mathbb{R}$ *and* $\mathbf{x} \in [\alpha,\beta]^m$*. Consider the function* $f \in C^{2n}[\alpha,\beta]$*,* $G_{(n)}$ *and* G_v*,* $(v = 1,\ldots,5)$ *be the same as defined in* (9.2) *and* (6.1)–(6.5)*, respectively. Then for* $(u = 1,\ldots,6)$ *along with the assumptions* (A_1) *and* (A_2)*, we have the following generalized identities:*

(a) For $n \geq 1$

$$J_u(f) = \sum_{l=0}^{n-1}(\beta-\alpha)^{2l}\left[f^{(2l)}(\alpha)J_u\left(P_l\left(\frac{\beta-x}{\beta-\alpha}\right)\right) + f^{(2l)}(\beta)J_u\left(P_l\left(\frac{x-\alpha}{\beta-\alpha}\right)\right)\right]$$
$$+(\beta-\alpha)^{2n-1}\int_\alpha^\beta J_u\left(G_{(n)}\left(\frac{x-\alpha}{\beta-\alpha},\frac{r-\alpha}{\beta-\alpha}\right)\right)f^{(2n)}(r)dr. \tag{9.4}$$

(b) For $n \geq 2$

$$
\begin{aligned}
J_u(f) = C_v(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u\Big(G_v(x,r)\Big) \\
\times \left(\sum_{l=0}^{n-2}(\beta-\alpha)^{2l}\left[f^{(2l+2)}(\alpha)P_l\left(\frac{\beta-r}{\beta-\alpha}\right) + f^{(2l+2)}(\beta)P_l\left(\frac{r-\alpha}{\beta-\alpha}\right)\right]\right) dr \\
+ (\beta-\alpha)^{2n-3}\int_\alpha^\beta f^{(2n)}(s)\left(\int_\alpha^\beta J_u\Big(G_v(x,r)\Big)G_{(n-1)}\left(\frac{r-\alpha}{\beta-\alpha},\frac{s-\alpha}{\beta-\alpha}\right)dr\right)ds
\end{aligned}
\tag{9.5}
$$

where C_v, $(v = 1,\ldots,5)$ is defined in (7.5)*.*

Proof. Fix $u = 1,\ldots,6$.

(a) As $f \in C^{2n}([\alpha,\beta])$. By Widder's lemma we have

$$
\begin{aligned}
f(x) = \sum_{l=0}^{n-1}(\beta-\alpha)^{2l}\left[f^{(2l)}(\alpha)P_l\left(\frac{\beta-x}{\beta-\alpha}\right) + f^{(2l)}(\beta)P_l\left(\frac{x-\alpha}{\beta-\alpha}\right)\right] \\
+ (\beta-\alpha)^{2n-1}\int_\alpha^\beta G_{(n)}\left(\frac{x-\alpha}{\beta-\alpha},\frac{r-\alpha}{\beta-\alpha}\right)f^{(2n)}(r)dr.
\end{aligned}
\tag{9.6}
$$

Now employing our respective cyclic Jensen's functional $J_u(\cdot)$ on (9.6) and practicing its linearity, we get (9.4) for $u = 1,\ldots,6$.

(b) For fix $v = 2$, testing identity (6.7) in cyclic Jensen's functional $J_u(\cdot)$ and employing its properties, we have

$$
J_u(f) = C_2(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u(G_2(x,r))f''(r)dr. \tag{9.7}
$$

Using representation (9.6) for f'', we get

$$
\begin{aligned}
f''(r) = \sum_{l=0}^{n-2}(\beta-\alpha)^{2l}\left[f^{(2l+2)}(\alpha)P_l\left(\frac{\beta-r}{\beta-\alpha}\right) + f^{(2l+2)}(\beta)P_l\left(\frac{r-\alpha}{\beta-\alpha}\right)\right] \\
+ (\beta-\alpha)^{2n-3}\int_\alpha^\beta G_{(n-1)}\left(\frac{r-\alpha}{\beta-\alpha},\frac{s-\alpha}{\beta-\alpha}\right)f^{(2n)}(s)ds.
\end{aligned}
\tag{9.8}
$$

Now, using (9.8) in (9.7) and applying Fubini's theorm, we get (9.5) for $v = 2$ and $u = 1,\ldots,6$. The cases for $(v = 1,3,4,5)$ can be treated analogously. □

Now we obtain generalizations of discrete and integral cyclic Jensen's type linear functionals, with real weights for $2n-$convex functions.

Theorem 9.1 *Consider $f \in C^{2n}[\alpha,\beta]$ be such that f is $2n-$convex function along with the suppositions of Lemma 9.2. Then we conclude the following results:*

(a) If for all $u = 1,\ldots,6$,

$$J_u\left(G_{(n)}\left(\frac{x-\alpha}{\beta-\alpha},\frac{r-\alpha}{\beta-\alpha}\right)\right) \geq 0, \quad r \in [\alpha,\beta] \tag{9.9}$$

holds, then we have

$$J_u(f) \geq \sum_{l=0}^{n-1}(\beta-\alpha)^{2l}\left[f^{(2l)}(\alpha)J_u\left(P_l\left(\frac{\beta-x}{\beta-\alpha}\right)\right)+f^{(2l)}(\beta)J_u\left(P_l\left(\frac{x-\alpha}{\beta-\alpha}\right)\right)\right]. \tag{9.10}$$

(b) If for all $u = 1,\ldots,6$ and $(v = 1,\ldots,5)$

$$\int_\alpha^\beta J_u\Big(G_v(x,r)\Big)G_{(n-1)}\left(\frac{r-\alpha}{\beta-\alpha},\frac{s-\alpha}{\beta-\alpha}\right)dr \geq 0, \ r \in [\alpha,\beta] \tag{9.11}$$

holds, then we have

$$J_u(f) \geq C_v(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u\Big(G_v(x,r)\Big)$$
$$\times\left(\sum_{l=0}^{n-2}(\beta-\alpha)^{2l}\left[f^{(2l+2)}(\alpha)P_l\left(\frac{\beta-r}{\beta-\alpha}\right)+f^{(2l+2)}(\beta)P_l\left(\frac{r-\alpha}{\beta-\alpha}\right)\right]\right)dr. \tag{9.12}$$

Proof. We start with the proof of (a) and its assumed conditions. Fix $u = 1,\ldots,6$.
By our assumption $f \in C^{2n}[\alpha,\beta]$ and is $2n-$convex function, we have $f^{(2n)}(\cdot) \geq 0$ (see [82], p. 16). Therefore, by applying Lemma 9.2(a) and taking into account assumption (9.9) and $f^{(2n)} \geq 0$, we get (9.10).
In the similar passion, we can give the proof of (9.12). □

We will finish the present section by the following results:

Theorem 9.2 *If the assumptions of Lemma 9.2 be fulfilled with additional conditions that $p_1,\ldots,p_m$ and $\lambda_1,\ldots,\lambda_k$ be nonnegative tuples for $2 \leq k \leq m$, such that $\sum_{i=1}^{m} p_i = 1$ and $\sum_{j=1}^{k} \lambda_j = 1$. Then for $f : [\alpha,\beta] \to \mathbb{R}$ being $2n-$convex function, we conclude the following results:*

(a) (9.10) *is valid for odd* $n \geq 1$. *Besides, for function*

$$H(x) := \sum_{l=0}^{n-1} (\beta-\alpha)^{2l} \left[f^{(2l)}(\alpha) P_l\left(\frac{\beta-x}{\beta-\alpha}\right) + f^{(2l)}(\beta) P_l\left(\frac{x-\alpha}{\beta-\alpha}\right) \right] \tag{9.13}$$

to be convex, the right side of (9.10) *is nonnegative, means*

$$J_u(f) \geq 0, \qquad u = 1, \ldots, 6. \tag{9.14}$$

(b) For odd $n \geq 3$, (9.12) *holds. Moreover, let* (9.12) *is valid and*

$$\left(\sum_{l=0}^{n-2} (\beta-\alpha)^{2l} \left[f^{(2l+2)}(\alpha) P_l\left(\frac{\beta-r}{\beta-\alpha}\right) + f^{(2l+2)}(\beta) P_l\left(\frac{r-\alpha}{\beta-\alpha}\right) \right] \right) \geq 0, \tag{9.15}$$

then, we get (9.14) *for all* $u = 1, \ldots, 6$ *and* $(v = 1, \ldots, 5)$.

Proof.

(a) Fix $u = 1, \ldots, 6$.
From (9.2), we get $G_{(n)}(x,r) \leq 0$ for odd n and $G_{(n)}(x,r) \geq 0$ for even n. Moreover G_1 in (9.1) is convex and G_{n-1} is positive for odd n. Thus taking into account (9.2), $G_{(n)}$ is convex in first variable if n is odd. Therefore (9.9) holds by virtue of Remark 2.7 on account of given weights to be positive. Hence (9.10) is established by taking into account Theorem 9.1 (a). Moreover, the R.H.S. of (9.10) can be written in the functional form $J_u(H)$ for all $(i = 1, \ldots, 6)$ after reorganizing this side. Employing Remark 2.7 the nonnegativity of R.H.S. of (9.10) is secure, especially (9.14) is established.

(b) Fix $u = 1, \ldots, 6$.
For odd $n \geq 3$, G_{n-1} is positive. Also we have assumed positive weights and for all $(j = 1, \ldots, 5)$, $G_v(x,r)$ is convex. Thus by practicing Remark 2.7, $J_u\Big(G_v(x,r)\Big) \geq 0$ which together with positivity of G_{n-1} yields (9.11). As f is $2n-$convex, hence by following Theorem 9.1 (b), we obtain (9.12). Finally, taking into account the positivity of $J_u\Big(G_v(x,r)\Big)$ and (9.15), we get (9.14).

$\square$

9.2 Applications to information theory

Under the assumptions of Theorem 9.1 (a) to be valid, we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $u = 1$, from (9.10) with respect to $2n-$convex function f in the explicit form:

$$\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right)$$
$$\geq \sum_{l=0}^{n-1} (\beta-\alpha)^{2l} f^{(2l)}(\alpha) \times$$
$$\times \left(\sum_{i=1}^{m} p_i \cdot P_l \left(\frac{\beta - x_i}{\beta - \alpha} \right) - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) P_l \left(\frac{\beta - \frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}}{\beta - \alpha} \right) \right)$$
$$+ \sum_{l=0}^{n-1} (\beta-\alpha)^{2l} f^{(2l)}(\beta) \times$$
$$\times \left(\sum_{i=1}^{m} p_i \cdot P_l \left(\frac{x_i - \alpha}{\beta - \alpha} \right) - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) P_l \left(\frac{\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} - \alpha}{\beta - \alpha} \right) \right) \tag{9.16}$$

where P_n is a Lidstone's polynomial defined in Lemma 9.1.

Theorem 9.3 *Let $m, k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1, \ldots, \lambda_k$ be positive probability distribution. Let $\mathbf{p} := (p_1, \ldots, p_m) \in \mathbb{R}^m$ and $\mathbf{q} := (q_1, \ldots, q_m) \in (0, \infty)^m$ be such that*

$$\frac{p_i}{q_i} \in [\alpha, \beta], \quad i = 1, \ldots, m.$$

Also let $f \in C^{2n}[\alpha, \beta]$ such that f is $2n-$convex function. Then the following inequalities hold:

$$\begin{aligned}
\widetilde{I}_f(\mathbf{p},\mathbf{q}) &\geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\\
&+\sum_{l=0}^{n-1}(\beta-\alpha)^{2l}f^{(2l)}(\alpha)\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\beta-\frac{p_i}{q_i}}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\beta-\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}}{\beta-\alpha}\right)\right)\\
&+\sum_{l=0}^{n-1}(\beta-\alpha)^{2l}f^{(2l)}(\beta)\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\frac{p_i}{q_i}-\alpha}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha}{\beta-\alpha}\right)\right).
\end{aligned} \tag{9.17}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i=1,\ldots,m)$ in (9.16) , we get (9.17). $\square$

Remark 9.1 Under the assumptions of Theorem 9.2 (a) for $u=1$ to be fulfilled, (9.17) becomes

$$\widetilde{I}_f(\mathbf{p},\mathbf{q}) \geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right). \tag{9.18}$$

We now explore two exceptional cases of the previous result.
One corresponds to the entropy of a discrete probability distribution.

Corollary 9.1 *Let $m,k\in\mathbb{N}$ $(2\leq k\leq m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q}:=(q_1,\ldots,q_m)\in(0,\infty)^m$, then

$$\begin{aligned}
\sum_{i=1}^{m}q_i\ln q_i &\geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\\
&\quad+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\times\\
&\quad\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\beta-\frac{1}{q_i}}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\beta-\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}}{\beta-\alpha}\right)\right)
\end{aligned}$$

$$+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\times$$
$$\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\frac{1}{q_i}-\alpha}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha}{\beta-\alpha}\right)\right). \tag{9.19}$$

(b) *If* $\mathbf{q}:=(q_1,\ldots,q_m)$ *is a positive probability distribution, then we get the bounds for the Shannon entropy of* $\mathbf{q}$.

$$H(\mathbf{q})\leq-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)$$
$$-\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\times$$
$$\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\beta-\frac{1}{q_i}}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\beta-\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}}{\beta-\alpha}\right)\right)$$
$$-\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\times$$
$$\times\left(\sum_{i=1}^{m}q_i\cdot P_l\left(\frac{\frac{1}{q_i}-\alpha}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha}{\beta-\alpha}\right)\right). \tag{9.20}$$

Proof.

(a) Using $f(x):=-\ln x$, and $\mathbf{p}:=(1,1,\ldots,1)$ in Theorem 9.3, we get the required result.

(b) It is a special case of (a). □

Remark 9.2 Using Remark 9.1, (9.20) becomes

$$H(\mathbf{q})\leq-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right). \tag{9.21}$$

The second case corresponds to the Relative entropy or Kullback-Leibler divergence between two probability distributions. Some recent bounds for Relative entropy can be seen in [52, 38]. We propose the following results:

Corollary 9.2 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) *If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m) \in (0,\infty)^m$, then*

$$\begin{aligned}
\sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\\
&+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\times\\
&\times\left(\sum_{i=1}^{m} q_i\cdot P_l\left(\frac{\beta-\frac{p_i}{q_i}}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\beta-\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}}{\beta-\alpha}\right)\right)\\
&+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\times\\
&\times\left(\sum_{i=1}^{m} q_i\cdot P_l\left(\frac{\frac{p_i}{q_i}-\alpha}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha}{\beta-\alpha}\right)\right).
\end{aligned} \tag{9.22}$$

(b) *If If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m)$ are positive probability distributions, then we have*

$$\begin{aligned}
D(\mathbf{q}\,\|\,\mathbf{p}) &\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\\
&+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\times\\
&\times\left(\sum_{i=1}^{m} q_i\cdot P_l\left(\frac{\beta-\frac{p_i}{q_i}}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\beta-\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}}{\beta-\alpha}\right)\right)
\end{aligned}$$

$$+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\times$$
$$\times\left(\sum_{i=1}^{m} q_i\cdot P_l\left(\frac{\frac{p_i}{q_i}-\alpha}{\beta-\alpha}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)P_l\left(\frac{\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}-\alpha}{\beta-\alpha}\right)\right)\Bigg). \tag{9.23}$$

Proof.

(a) Using $f(x) := -\ln x$ in Theorem 9.3, we get the desired result.

(b) It is special case of (a).

□

Remark 9.3 Using Remark 9.1, (9.23) becomes

$$D(\mathbf{q}\parallel\mathbf{p})\geq\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right). \tag{9.24}$$

Some of the recent study regarding Zipf-Mandelbrot law can be seen in the listed references (see [48, 52, 53, 38]). Now we state our results involving entropy introduced by Mandelbrot Law by establishing the relationship with Shannon and Relative entropies:

Theorem 9.4 *Let $m,k\in\mathbb{N}$ $(2\leq k\leq m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution and $\mathbf{q}$ be as defined in* (6.59) *by Zipf-Mandelbrot law with parameters $m\in\{1,2,\dots\}$, $c\geq 0$, $d>0$. Then, the following holds*

$$H(\mathbf{q})=Z(H,c,d)\leq-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^dH_{c,d}^m)}\right)\ln\left(\frac{1}{H_{c,d}^m}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^d)}\right)$$
$$-\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\left(\sum_{i=1}^{m}\frac{1}{((i+c)^dH_{c,d}^m)}\cdot P_l\left(\frac{\beta-((i+c)^dH_{c,d}^m)}{\beta-\alpha}\right)\right.$$
$$\left.-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^dH_{c,d}^m)}\right)P_l\left(\frac{\beta-\sum_{j=0}^{k-1}\frac{((i+j+c)^dH_{c,d}^m)}{\lambda_{j+1}}}{\beta-\alpha}\right)\right)$$

$$-\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\left(\sum_{i=1}^{m}\frac{1}{((i+c)^{d}H_{c,d}^{m})}\cdot P_l\left(\frac{((i+c)^{d}H_{c,d}^{m})-\alpha}{\beta-\alpha}\right)\right.$$
$$\left.-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^{d}H_{c,d}^{m})}\right)P_l\left(\frac{\sum_{j=0}^{k-1}\frac{((i+j+c)^{d}H_{c,d}^{m})}{\lambda_{j+1}}-\alpha}{\beta-\alpha}\right)\right). \tag{9.25}$$

Proof. Similar to that of Theorem 6.5. Finally, substituting this $q_i=\frac{1}{((u+c)^{d}H_{c,d}^{m})}$ in Corollary 9.1(b), we get the desired result. □

The next result establish the relationship of Relative entropy with Mandelbrot entropy:

Corollary 9.3 *Let $m,k\in\mathbb{N}$ $(2\le k\le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distributions and for $c_1,c_2\in[0,\infty)$, $d_1,d_2>0$, let $H_{c_1,d_1}^{m}=\sum_{\sigma=1}^{m}\frac{1}{(\sigma+c_1)^{d_1}}$ and $H_{c_2,d_2}^{m}=\sum_{\sigma=1}^{m}\frac{1}{(\sigma+c_2)^{d_2}}$. Now by using $q_i=\dfrac{1}{(i+c_1)^{d_1}H_{c_1,d_1}^{m}}$ and $p_i=\dfrac{1}{(i+c_2)^{d_2}H_{c_2,d_2}^{m}}$ in Corollary 9.2(b), we obtain*

$$\begin{aligned}
D(\mathbf{q}\parallel\mathbf{p}) &= \sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H_{c_1,d_1}^{m}}\ln\left(\frac{(i+c_2)^{d_2}H_{c_2,d_2}^{m}}{(i+c_1)^{d_1}H_{c_1,d_1}^{m}}\right)\\
&= -Z(H,c_1,d_1)+\frac{d_2}{H_{c_1,d_1}^{m}}\sum_{i=1}^{m}\frac{\ln(i+c_2)}{(i+c_1)^{d_1}}+\ln\left(H_{c_2,d_2}^{m}\right)\\
&\ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{c_1,d_1}^{m}}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{c_1,d_1}^{m}}}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_2)^{d_2}H_{c_2,d_2}^{m}}}\right)\\
&\quad+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\alpha)^{2l}}\left(\sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H_{c_1,d_1}^{m}}\cdot P_l\left(\frac{\beta-\left(\frac{(i+c_1)^{d_1}H_{c_1,d_1}^{m}}{(i+c_2)^{d_2}H_{c_2,d_2}^{m}}\right)}{\beta-\alpha}\right)\right.\\
&\quad\left.-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{c_1,d_1}^{m}}\right)P_l\left(\frac{\beta-\frac{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_2)^{d_2}H_{c_2,d_2}^{m}}}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{c_1,d_1}^{m}}}}{\beta-\alpha}\right)\right)
\end{aligned}$$

$$+\sum_{l=0}^{n-1}\frac{(\beta-\alpha)^{2l}(2l-1)!}{(\beta)^{2l}}\left(\sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H^m_{c_1,d_1}}\cdot P_l\left(\frac{\left(\frac{(i+c_1)^{d_1}H^m_{c_1,d_1}}{(i+c_2)^{d_2}H^m_{c_2,d_2}}\right)-\alpha}{\beta-\alpha}\right)\right.$$

$$\left.-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}\right)P_l\left(\frac{\frac{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_2)^{d_2}H^m_{c_2,d_2}}}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}}-\alpha}{\beta-\alpha}\right)\right). \tag{9.26}$$

Remark 9.4 By using Remark 9.1, (9.25) and (9.28) becomes

$$H(\mathbf{q}) \;=\; Z(H,c,d) \;\leq\; -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^d H^m_{c,d})}\right)\ln\left(\frac{1}{H^m_{c,d}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^d)}\right). \tag{9.27}$$

$$\begin{aligned}
D(\mathbf{q}\,\|\,\mathbf{p}) &= \sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H^m_{c_1,d_1}}\ln\left(\frac{(i+c_2)^{d_2}H^m_{c_2,d_2}}{(i+c_1)^{d_1}H^m_{c_1,d_1}}\right)\\
&= -Z(H,c_1,d_1)+\frac{d_2}{H^m_{c_1,d_1}}\sum_{i=1}^{m}\frac{\ln(i+c_2)}{(i+c_1)^{d_1}}+\ln\left(H^m_{c_2,d_2}\right)\\
&\geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}}{\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_2)^{d_2}H^m_{c_2,d_2}}}\right).
\end{aligned} \tag{9.28}$$

Remark 9.5 It is interesting to note that, in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon, Relative and Mandelbrot entropies using the other inequalities for $2n-$convex functions constructed in Theorem 9.1 for discrete case of cyclic refinements of Jensen's inequality.

Chapter 10

Cyclic Refinements of Jensen's Inequality and Abel-Gontscharoff Interpolating Polynomial

In the present Chapter, we use Abel-Gontscharoff interpolating polynomial and prove many interesting results. This chapter is based on the paper[71].

Let $-\infty < \alpha < \beta < \infty$ and let $\alpha \le \xi_1 < \xi_2 < \cdots \xi_n \le \beta$ be the given points. For $f \in C^n[\alpha, \beta]$, Abel-Gontscharoff interpolating polynomial AP of degree $(n-1)$ satisfying Abel-Gontscharoff conditions

$$AP^{(\sigma)}(\xi_{\sigma+1}) = f^{(\sigma)}(\xi_{\sigma+1}), \quad 0 \le \sigma \le n-1$$

exists uniquely [20, 33]. This condition in particular includes two point right focal conditions.

$$\begin{aligned} &AP_{(2)}^{(\sigma)}(\xi_1) = f^{(\sigma)}(\xi_1), \quad 0 \le \sigma \le t \\ &AP_{(2)}^{(\sigma)}(\xi_2) = f^{(\sigma)}(\xi_2), \quad t+1 \le \sigma \le n-1, \quad \alpha \le \xi_1 < \xi_2 \le \beta. \end{aligned}$$

First we give representation of Abel-Gontscharoff interpolating polynomial:

Theorem 10.1 *[1] Abel-Gontscharoff interpolating polynomial AP of function f can be expressed as*

$$AP(x) = \sum_{\sigma=0}^{n-1} \Lambda_\sigma(x) f^{(\sigma)}(\xi_{\sigma+1}) \tag{10.1}$$

where $\Lambda_0(x) = 1$ and Λ_σ, $1 \leq \sigma \leq n-1$ is the unique polynomial of degree σ satisfying

$$\Lambda_\sigma^{(l)}(\xi_{l+1}) = 0, \quad 0 \leq l \leq \sigma - 1$$
$$\Lambda_\sigma^{(\sigma)}(\xi_{\sigma+1}) = 1$$

and it can be written as

$$\Lambda_\sigma(x) = \frac{1}{1!2!\cdots\sigma!}\begin{vmatrix} 1 & \xi_1 & \xi_1^2 & \ldots & \xi_1^{\sigma-1} & \xi_1^\sigma \\ 0 & 1 & 2\xi_2 & \ldots & (\sigma-1)\xi_2^{\sigma-2} & \sigma\xi_2^{\sigma-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \ldots & (\sigma-1)! & \sigma!\xi_\sigma \\ 1 & x & x^2 & \ldots & x^{\sigma-1} & x^\sigma \end{vmatrix}$$
$$= \int_{\xi_1}^{x}\int_{\xi_2}^{x_1}\int_{\xi_3}^{x_2}\cdots\int_{\xi_\sigma}^{x_{\sigma-1}} dx_\sigma dx_{\sigma-1}\cdots dx_1, \quad (x_0 = x). \tag{10.2}$$

In particular, we have

$$\Lambda_0(x) = 1$$
$$\Lambda_1(x) = x - \xi_1$$
$$\Lambda_2(x) = \frac{1}{2}[x^2 - 2\xi_2 x + \xi_1(2\xi_2 - \xi_1)].$$

Corollary 10.1 *The two point right focal interpolating polynomial $AP_{(2)}(x)$ of the function f can be written as*

$$AP_{(2)}(x) = \sum_{\sigma=0}^{t} \frac{(x-\xi_1)^\sigma}{\sigma!} f^{(\sigma)}(\xi_1)$$
$$+ \sum_{w=0}^{n-t-2}\left[\sum_{\sigma=0}^{w} \frac{(x-\xi_1)^{t+1+\sigma}(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!}\right] f^{(t+1+w)}(\xi_2) \tag{10.3}$$

The associate error $Error(x) = f(x) - AP(x)$ can be represented in terms of the Green function $AG(x,r;n)$ of the boundary value problem
$y^{(n)} = 0$, $y^{(\sigma)}(\xi_{\sigma+1}) = 0$, $0 \leq \sigma \leq n-1$
and appears as (see [1]):

$$AG(x,r;n) = \begin{cases} \displaystyle\sum_{\sigma=0}^{l-1} \frac{\Lambda_\sigma(x)}{(n-\sigma-1)!}(\xi_{\sigma+1} - r)^{(n-\sigma-1)}, & \xi_l \leq r \leq x, \\ \displaystyle-\sum_{\sigma=l}^{n-1} \frac{\Lambda_\sigma(x)}{(n-\sigma-1)!}(\xi_{\sigma+1} - r)^{(n-\sigma-1)}, & x \leq r \leq \xi_{l+1}, \\ & l = 0,1,\ldots,n(\xi_0 = \alpha, \xi_{n+1} = \beta). \end{cases} \tag{10.4}$$

Corresponding to the two point right focal conditions, Green function $AG_{(2)}(x,r;n)$ of the boundary value problem
$y^{(n)}=0,\ y^{(\sigma)}(\xi_1)=0,\ 0\le\sigma\le t\ ,\ y^{(\sigma)}(\xi_2)=0,\ t+1\le\sigma\le n-1$
is given by (see [1]):

$$AG_{(2)}(x,r;n)=\frac{1}{(n-1)!}\begin{cases}\sum\limits_{\sigma=0}^{t}\binom{n-1}{\sigma}(x-\xi_1)^{\sigma}(\xi_1-r)^{n-\sigma-1}, & \alpha\le r\le x,\\ -\sum\limits_{\sigma=t+1}^{n-1}\binom{n-1}{\sigma}(x-\xi_1)^{\sigma}(\xi_1-r)^{n-\sigma-1}, & x\le r\le\beta.\end{cases} \tag{10.5}$$

Further, for $\xi_1\le r,\ \ x\le\xi_2$ the following inequalities hold

$$(-1)^{n-t-1}\frac{\partial^{\sigma}AG_{(2)}(x,r;n)}{\partial x^{\sigma}}\ge 0,\quad 0\le\sigma\le t, \tag{10.6}$$

$$(-1)^{n-\sigma}\frac{\partial^{\sigma}AG_{(2)}(x,r;n)}{\partial x^{\sigma}}\ge 0,\quad t+1\le\sigma\le n-1. \tag{10.7}$$

Theorem 10.2 *Let $f\in C^n[\alpha,\beta]$, and let $AP(\cdot)$ be its Abel-Gontscharoff interpolating polynomial . Then*

$$f(x)=AP(x)+Error(x)=\sum_{\sigma=0}^{n-1}\Lambda_{\sigma}(x)f^{(\sigma)}(\xi_{\sigma+1})+\int_{\alpha}^{\beta}AG(x,r;n)f^{(n)}(r)dr \tag{10.8}$$

where $\Lambda(\cdot)$ is defined by (10.2) and $AG(x,r;n)$ is defined by (10.4).

Theorem 10.3 *Let $f\in C^n[\alpha,\beta]$, and let $AP_{(2)}(\cdot)$ be its two points right focal Abel-Gontscharoff interpolating polynomial . Then*

$$\begin{aligned}f(z)&=AP_{(2)}(x)+Error(x)\\&=\sum_{\sigma=0}^{t}\frac{(x-\xi_1)^{\sigma}}{\sigma!}f^{(\sigma)}(\xi_1)+\sum_{w=0}^{n-t-2}\left[\sum_{\sigma=0}^{w}\frac{(x-\xi_1)^{t+1+\sigma}(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!}\right]f^{(t+1+w)}(\xi_2)\\&\qquad+\int_{\alpha}^{\beta}AG_{(2)}(x,r;n)f^{(n)}(r)dr\end{aligned} \tag{10.9}$$

where $AG_{(2)}(x,r;n)$ is defined by (10.5).

10.1 Extensions of cyclic refinements of Jensen's inequality by Abel-Gontscharoff interpolation

We consider discrete as well as continuous version of cyclic refinements of Jensen's inequality and construct the generalized new identities having real weights utilizing Abel-Gontscharoff interpolating polynomial.

Theorem 10.4 *Suppose $m,k \in \mathbb{N}$, $p_1,\ldots,p_m$ and $\lambda_1,\ldots,\lambda_k$ are real tuples for $2 \leq k \leq m$, such that $\sum_{j=0}^{k-1} \lambda_{j+1}p_{i+j} \neq 0$ for $i = 1,\ldots,m$ with $\sum_{i=1}^{m} p_i = 1$ and $\sum_{j=1}^{k} \lambda_j = 1$. Also let $x \in [\alpha,\beta] \subset \mathbb{R}$ and $\mathbf{x} \in [\alpha,\beta]^m$. Assume $f \in C^n[\alpha,\beta]$ and consider interval with points $-\infty < \alpha \leq \xi_1 < \xi_2 < \cdots \xi_n \leq \beta < \infty$, $\Lambda(\cdot)$ is defined by (10.2) , $AG(\cdot,r;n)$ in (10.4) and G_v, $(v = 1,\ldots,5)$ be the Green functions defined in (6.1)–(6.5), respectively. Then for $u = 1,\ldots,6$ along with assumptions (A_1) and (A_2), we have the following generalized identities:*

(a) For $n \geq 1$

$$J_u(f(x)) = \sum_{\sigma=1}^{n-1} f^{(\sigma)}(\xi_{\sigma+1})J_u\Big(\Lambda_\sigma(x)\Big) + \int_\alpha^\beta J_u\Big(AG(x,r;n)\Big)f^{(n)}(r)dr. \tag{10.10}$$

(b) For $n \geq 3$

$$\begin{aligned}J_u(f(x)) = C_v(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u\Big(G_v(x,r)\Big)\sum_{\sigma=0}^{n-3} f^{(\sigma+2)}(\xi_{\sigma+1})\Lambda_\sigma(r)dr \\ + \int_\alpha^\beta\int_\alpha^\beta J_u\Big(G_v(x,r)\Big)AG(r,s;n-2)f^{(n)}(s)dsdr\end{aligned} \tag{10.11}$$

where C_v, $(v = 1,\ldots,5)$ is defined in (7.5).

Proof. Fix $u = 1,\ldots,6$.

(a) Applying cyclic Jensen's type linear functionals $J_u(\cdot)$ on (10.8) and practicing properties of the functional, we get (10.10).

(b) For fix $v = 5$, testing (6.10) in cyclic Jensen's type functional $J_u(\cdot)$ and employing the properties of $J_u(\cdot)$ along with the assumed conditions, we have

$$J_u(f) = C_5(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u(G_5(x,r))f''(r)dr. \tag{10.12}$$

By Theorem 10.2, $f''(r)$ can be expressed as:

$$f''(r) = \sum_{\sigma=0}^{n-3} \Lambda_\sigma(r) f^{(\sigma+2)}(\xi_{\sigma+1}) + \int_\alpha^\beta AG(r,s;n-2) f^{(n)}(s)ds. \tag{10.13}$$

Putting (10.13) in (10.12), we get (10.11) respectively for $v=5$ and $u=1,\ldots,6$. The cases for $v=1,2,3,4$ are treated analogously and are left for the reader interest.

□

Now we obtain extensions and improvements of discrete and integral cyclic Jensen's inequalities, with real weights.

Theorem 10.5 *Consider f be $n-$convex function along with the suppositions of Theorem 10.4. Then we conclude the following results:*

(a) If for all $u=1,\ldots,6$,

$$J_u\Big(AG(x,r;n)\Big) \geq 0, \quad r \in [\alpha,\beta] \tag{10.14}$$

holds, then we have

$$J_u(f(x)) \geq \sum_{\sigma=1}^{n-1} f^{(\sigma)}(\xi_{\sigma+1}) J_u\Big(\Lambda_\sigma(x)\Big) \tag{10.15}$$

for $u=1,\ldots,6$.

(b) If for all $u=1,\ldots,6$ and $v=1,\ldots,5$

$$\int_\alpha^\beta J_u\Big(G_v(x,r)\Big) AG(r,s;n-2)dr \geq 0, \qquad r \in [\alpha,\beta] \tag{10.16}$$

holds then

$$J_u(f(x)) \geq C_v(\alpha,\beta,f) J_u(x) + \int_\alpha^\beta J_u\Big(G_v(x,r)\Big) \sum_{\sigma=0}^{n-3} f^{(\sigma+2)}(\xi_{\sigma+1}) \Lambda_\sigma(r)dr \tag{10.17}$$

for $u=1,\ldots,6$.

Proof. Similar to that of Theorem 6.2. □

In the next corollary, we give Theorem 10.5 by considering two points right focal Abel-Gontscharoff interpolating polynomial:

Corollary 10.2 *Assume $f \in C^n[\alpha,\beta]$ on the interval with points $\alpha \leq \xi_1 < \xi_2 < \beta$ along with the suppositions of Theorem 10.4. Let $AG_{(2)}(x,r;n)$ be the Green function defined in* (10.5)*. If f be $n-$convex function, then we conclude the following results:*

(a) If for all $u = 1, \ldots, 6$,

$$J_u\Big(AG_{(2)}(x,r;n)\Big) \geq 0, \quad r \in [\alpha, \beta] \tag{10.18}$$

holds, then we have

$$\begin{aligned}J_u(f(x)) \geq & \sum_{\sigma=1}^{t} \frac{f^{(\sigma)}(\xi_1)}{\sigma!} J_u\Big((x-\xi_1)^{\sigma}\Big)\\ &+ \sum_{w=0}^{n-t-2}\left[\sum_{\sigma=0}^{w} \frac{f^{(t+1+w)}(\xi_2)(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!}\right] J_u\Big((x-\xi_1)^{t+1+\sigma}\Big)\end{aligned} \tag{10.19}$$

for $u = 1, \ldots, 6$.

(b) If for all $u = 1, \ldots, 6$ and $v = 1, \ldots, 5$

$$J_u\Big(G_v(x,r)\Big) \geq 0, \quad r \in [\alpha, \beta] \tag{10.20}$$

holds, provided that n is even and t is odd or n is odd and t is even, then

$$\begin{aligned}J_u(f(x)) \geq & C_v(\alpha,\beta,f)J_u(x) + \sum_{\sigma=0}^{t} \frac{f^{(\sigma+2)}(\xi_1)}{\sigma!} \int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)(r-\xi_1)^{\sigma} dr\\ &+ \sum_{w=0}^{n-t-4}\left[\sum_{\sigma=0}^{w} \frac{f^{(t+3+w)}(\xi_2)(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!}\right] \int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)(r-\xi_1)^{t+1+\sigma} dr\end{aligned} \tag{10.21}$$

for $u = 1, \ldots, 6$.

Proof. Fix $u = 1, \ldots, 6$.

(a) Applying cyclic Jensen's type linear functionals $J_u(\cdot)$ on (10.9) and practicing properties of the functional, we get

$$\begin{aligned}J_u(f(x)) = & \sum_{\sigma=1}^{t} \frac{f^{(\sigma)}(\xi_1)}{\sigma!} J_u\Big((x-\xi_1)^{\sigma}\Big)\\ &+ \sum_{w=0}^{n-t-2}\left[\sum_{\sigma=0}^{w} \frac{f^{(t+1+w)}(\xi_2)(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!}\right] J_u\Big((x-\xi_1)^{t+1+\sigma}\Big)\\ &+ \int_{\alpha}^{\beta} J_u\Big(AG_{(2)}(x,r;n)\Big) f^{(n)}(r) dr.\end{aligned} \tag{10.22}$$

Now by using (10.18) and $n-$convexity of the function f, we get (10.19).

(b) Fix $u = 1,\ldots,6$ and $v = 1,\ldots,5$.
By Theorem 10.4(b), we already proved

$$J_u(f) = C_v(\alpha,\beta,f)J_u(x) + \int_{\alpha}^{\beta} J_u(G_v(x,r))f''(r)dr. \tag{10.23}$$

By Theorem 10.3, $f''(r)$ can be expressed as:

$$\begin{aligned} f''(r) = & \sum_{\sigma=0}^{t} \frac{(r-\xi_1)^{\sigma}}{\sigma!} f^{(\sigma+2)}(\xi_1) \\ & + \sum_{w=0}^{n-t-4} \left[\sum_{\sigma=0}^{w} \frac{(r-\xi_1)^{t+1+\sigma}(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!} \right] f^{(t+3+w)}(\xi_2) \\ & + \int_{\alpha}^{\beta} AG_{(2)}(r,s;n-2)f^{(n)}(s)ds. \end{aligned} \tag{10.24}$$

Putting (10.24) in (10.23), we get the following identity

$$\begin{aligned} J_u(f(x)) = & C_v(\alpha,\beta,f)J_u(x) + \sum_{\sigma=0}^{t} \frac{f^{(\sigma+2)}(\xi_1)}{\sigma!} \int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)(r-\xi_1)^{\sigma}dr \\ & + \sum_{w=0}^{n-t-4} \left[\sum_{\sigma=0}^{w} \frac{f^{(t+3+w)}(\xi_2)(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!} \right] \int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)(r-\xi_1)^{t+1+\sigma}dr \\ & + \int_{\alpha}^{\beta}\int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)AG_{(2)}(r,s;n-2)f^{(n)}(s)dsdr. \end{aligned} \tag{10.25}$$

Now from (10.6), we have $(-1)^{n-t-3}AG_{(2)}(r,s;n-2) \geq 0$. Therefore utilizing our assumptions n is even and t is odd or n is odd and t is even, we get $AG_{(2)}(r,s;n-2) \geq 0$. Now employing (10.20) alongside with $n-$convexity of f yields (10.21).

□

We will finish the present section by the following generalizations of cyclic refinements of Jensen's inequalities by two points right focal Abel-Gontscharoff interpolating polynomial:

Theorem 10.6 *If the assumptions of Corollary 10.2 be fulfilled with additional conditions that $p_1,\ldots,p_m$ and $\lambda_1,\ldots,\lambda_k$ be non negative tuples for $2 \leq k \leq m$, such that $\sum_{i=1}^{m} p_i = 1$ and $\sum_{j=1}^{k} \lambda_j = 1$. Then for $f : [\alpha,\beta] \to \mathbb{R}$ being $n-$convex function, we conclude the following results:*

(a) (10.19) holds for the cases when n is even and t is odd or n is odd and t is even. If (10.19) is valid along with the function

$$\Gamma(x) := \sum_{\sigma=0}^{t} \frac{(x-\xi_1)^{\sigma}}{\sigma!} f^{(\sigma)}(\xi_1) + \sum_{w=0}^{n-t-2} \left[\sum_{\sigma=0}^{w} \frac{(x-\xi_1)^{t+1+\sigma}(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!} \right] f^{(t+1+w)}(\xi_2) \quad (10.26)$$

to be convex, the right side of (10.19) is non negative, means

$$J_u(f) \geq 0, \qquad u = 1, \ldots, 6. \quad (10.27)$$

(b) For $(n = even, t = odd)$ or $(n = odd, t = even)$, (10.21) holds. Further

$$\sum_{\sigma=0}^{t} \frac{(r-\xi_1)^{\sigma}}{\sigma!} f^{(\sigma+2)}(\xi_1) + \sum_{w=0}^{n-t-4} \left[\sum_{\sigma=0}^{w} \frac{(r-\xi_1)^{t+1+\sigma}(\xi_1-\xi_2)^{w-\sigma}}{(t+1+\sigma)!(w-\sigma)!} \right] f^{(t+3+w)}(\xi_2) \geq 0 \quad (10.28)$$

the right side of (10.21) is non negative, particularly (10.27) is established for all $u = 1, \ldots, 6$ and $v = 1, \ldots, 5$.

Proof.

(a) Fix $u = 1, \ldots, 6$.
Using (10.6), for $\xi_1 \leq r, \quad x \leq \xi_2$,

$$(-1)^{n-t-1} \frac{\partial^2 AG_{(2)}(x, r; n)}{\partial x^2} \geq 0 \quad (10.29)$$

ensures the convexity of $AG_{(2)}(x, r; n)$ w.r.t. first variable for the cases when n is even and t is odd or n is odd and t is even. So (10.18) holds by virtue of Remark 2.7 on account of given weights to be positive. Hence (10.19) is established by taking into account Corollary 10.2(a). Moreover, the R.H.S. of (10.19) can be written in the functional form $J_u(\Gamma)$ for all $(u = 1, \ldots, 6)$, after reorganizing this side. Employing Remark 2.7 the non negativity of R.H.S. of (10.19) is secured, especially (10.27) is established.

(b) Similar to the proof of Theorem 6.3 (c).

□

10.2 Applications to information theory

Now as a consequence of Theorem 10.5 we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $(u=1)$, from (10.15) with respect to $n-$convex function f in the explicit form:

$$\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}z_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$
$$\geq \sum_{\sigma=1}^{n-1} f^{(\sigma)}(\xi_{\sigma+1}) \times \left(\sum_{i=1}^{m} p_i\Lambda_\sigma(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right)\Lambda_\sigma\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\right), \tag{10.30}$$

where $\Lambda(\cdot)$ is defined by (10.2).

Theorem 10.7 *Let $m,k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution. Let $\mathbf{p} := (p_1,\ldots,p_m) \in \mathbb{R}^m$ and $\mathbf{q} := (q_1,\ldots,q_m) \in (0,\infty)^m$ such that*

$$\frac{p_i}{q_i} \in [\alpha,\beta], \quad u = 1,\ldots,m.$$

Also let $f \in C^n[\alpha,\beta]$ and consider interval with points $-\infty < \alpha \leq \xi_1 < \xi_2 < \cdots \xi_n \leq \beta < \infty$ such that f is $n-$convex function. Then the following inequality holds:

$$\widetilde{I}_f(\mathbf{p},\mathbf{q}) \geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)$$
$$+\sum_{\sigma=1}^{n-1} f^{(\sigma)}(\xi_{\sigma+1}) \times \left(\sum_{i=1}^{m} q_i\Lambda_\sigma\left(\frac{p_i}{q_i}\right) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\Lambda_\sigma\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{10.31}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i=1,\ldots,m)$ in (10.30) , we get (10.31). □

We now explore two exceptional cases of the previous result.
One corresponds to the entropy of a discrete probability distribution.

Corollary 10.3 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m) \in (0,\infty)^m$ and n is even , then

$$\sum_{i=1}^{m} q_i \ln q_i \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) + \left(\sum_{\sigma=1}^{n-1}\frac{(-1)^\sigma(\sigma-1)!}{(\xi_{\sigma+1})^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m} q_i\Lambda_\sigma\left(\frac{1}{q_i}\right) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\Lambda_\sigma\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{10.32}$$

(b) If $\mathbf{q} := (q_1,\ldots,q_m)$ is a positive probability distribution and n is even , then we get the bounds for the Shannon entropy of $\mathbf{q}$.

$$H(\mathbf{q}) \le -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) - \left(\sum_{\sigma=1}^{n-1}\frac{(-1)^\sigma(\sigma-1)!}{(\xi_{\sigma+1})^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m} q_i\Lambda_\sigma\left(\frac{1}{q_i}\right) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\Lambda_\sigma\left(\frac{1}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{10.33}$$

If n is odd, then (10.32) *and* (10.33) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$, and $\mathbf{p} := (1,1,\ldots,1)$ in Theorem 10.7, we get the required result.

(b) It is a special case of (a). □

The second case corresponds to the relative entropy or Kullback-Leibler divergence between two probability distributions.

Corollary 10.4 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m) \in (0,\infty)^m$ and n is even , then

$$\sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) + \left(\sum_{\sigma=1}^{n-1}\frac{(-1)^\sigma(\sigma-1)!}{(\xi_{\sigma+1})^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m} q_i\Lambda_\sigma\left(\frac{p_i}{q_i}\right) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\Lambda_\sigma\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{10.34}$$

(b) *If If* $\mathbf{q} := (q_1, \ldots, q_m), \mathbf{p} := (p_1, \ldots, p_m)$ *are positive probability distributions and* n *is even , then we have*

$$D(\mathbf{q} \parallel \mathbf{p}) \geq \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \ln \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} \right) + \left(\sum_{\sigma=1}^{n-1} \frac{(-1)^{\sigma}(\sigma-1)!}{(\xi_{\sigma+1})^{\sigma}} \right) \times$$
$$\times \left(\sum_{i=1}^{m} q_i \Lambda_{\sigma} \left(\frac{p_i}{q_i} \right) - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \Lambda_{\sigma} \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) \right). \tag{10.35}$$

If n *is odd, then* (10.34) *and* (10.35) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$ in Theorem 10.7, we get the desired result.

(b) It is special case of (a).

□

Now we state our results involving entropy introduced by Mandelbrot Law:

Theorem 10.8 *Let* $m, k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1, \ldots, \lambda_k$ *be positive probability distribution and* $\mathbf{q}$ *be as defined in* (6.59) *by Zipf-Mandelbrot law with parameters* $m \in \{1, 2, \ldots\}$, $c \geq 0$, $d > 0$. *If* n *is even , the following holds*

$$\begin{aligned} H(\mathbf{q}) &= Z(H, c, d) \\ &\leq -\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{((i+j+c)^d H_{m,c,d})} \right) \ln \left(\frac{1}{H_{m,c,d}} \sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{((i+j+c)^d)} \right) \\ &- \left(\sum_{\sigma=1}^{n-1} \frac{(-1)^{\sigma}(\sigma-1)!}{(\xi_{\sigma+1})^{\sigma}} \right) \left(\sum_{i=1}^{m} \frac{1}{((i+c)^d H_{m,c,d})} \Lambda_{\sigma} \left(((i+c)^d H_{m,c,d}) \right) \right) \\ &+ \left(\sum_{\sigma=1}^{n-1} \frac{(-1)^{\sigma}(\sigma-1)!}{(\xi_{\sigma+1})^{\sigma}} \right) \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{((i+j+c)^d H_{m,c,d})} \right) \Lambda_{\sigma} \left(\frac{1}{\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{((i+j+c)^d H_{m,c,d})}} \right) \right). \end{aligned} \tag{10.36}$$

If n *is odd, then* (10.36) *holds in reverse direction.*

Proof. Similar to that of Theorem 6.5. □

Corollary 10.5 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution and for $c_1,c_2 \in [0,\infty)$, $d_1,d_2 > 0$, let $H_{m,c_1,d_1} = \sum_{s=1}^{m} \frac{1}{(s+c_1)^{d_1}}$ and $H_{m,c_2,d_2} = \sum_{s=1}^{m} \frac{1}{(s+c_2)^{d_2}}$. Now using $q_i = \dfrac{1}{(i+c_1)^{d_1} H_{m,c_1,d_1}}$ and $p_i = \dfrac{1}{(i+c_2)^{d_2} H_{m,c_2,d_2}}$ in Corollary 10.4(b), with even n, we obtain*

$$\begin{aligned}
D(\mathbf{q} \parallel \mathbf{p}) &= \sum_{i=1}^{m} \frac{1}{(i+c_1)^{d_1} H_{m,c_1,d_1}} \ln\left(\frac{(i+c_2)^{d_2} H_{m,c_2,d_2}}{(i+c_1)^{d_1} H_{m,c_1,d_1}}\right) \\
&\ge \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+c_1)^{d_1} H_{m,c_1,d_1}}\right) \ln\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+c_1)^{d_1} H_{m,c_1,d_1}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+c_2)^{d_2} H_{m,c_2,d_2}}}\right) \\
&+ \left(\sum_{\sigma=1}^{n-1} \frac{(-1)^{\sigma}(\sigma-1)!}{(\xi_{\sigma+1})^{\sigma}}\right) \left(\sum_{i=1}^{m} \frac{1}{((i+c_1)^{d_1} H_{m,c_1,d_1})} \Lambda_{\sigma}\left(\frac{((i+c_1)^{d_1} H_{m,c_1,d_1})}{((i+c_2)^{d_2} H_{m,c_2,d_2})}\right)\right) \\
&- \left(\sum_{\sigma=1}^{n-1} \frac{(-1)^{\sigma}(\sigma-1)!}{(\xi_{\sigma+1})^{\sigma}}\right) \times \\
&\quad \times \left(\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \frac{\lambda_{j+1}}{(i+j+c_1)^{d_1} H_{m,c_1,d_1}}\right) \Lambda_{\sigma}\left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+c_2)^{d_2} H_{m,c_2,d_2}}}{\sum_{j=0}^{k-1} \lambda_{j+1} \frac{1}{(i+j+c_1)^{d_1} H_{m,c_1,d_1}}}\right)\right).
\end{aligned} \tag{10.37}$$

If n is odd, then (10.37) *holds in reverse direction.*

Remark 10.1 It is interesting to note that, in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon, Relative and Mandelbrot entropies using the other inequalities for $n-$convex functions constructed in Theorem 10.5 for discrete case of cyclic refinements of Jensen inequality.

Remark 10.2 We left for reader interest to construct upper bounds for Shannon, Relative and Mandelbrot entropies by considering two points right focal Abel-Gontscharoff interpolating polynomial in the above results.

Chapter 11

Cyclic Refinements of Jensen's Inequality by Hermite Intepolating Polynomial

In the present Chapter, we use Hermite interpolating polynomial to obtain generalizations of cyclic refinements of Jensen's inequality for n-convex functions. This chapter is based on the paper [72].

In what follows, we shall use the following conventions for the sake of simplicity Conditions(C), Hermite(H), Lagrange(L) and Taylor(T). Let $f \in C^n[\alpha, \beta]$ and consider interval with points $-\infty < \alpha = b_1 < b_2 \cdots < b_t = \beta < \infty$, $(t \geq 2)$. Then there exists a unique polynomial $\Lambda_H(\cdot)$ of degree $(n-1)$ satisfying any of the following axioms:

HC:$(1 \leq \omega \leq t,\ \sum_{\omega=1}^{t} s_\omega + t = n)$

$$\Lambda_H^{(\sigma)}(b_\omega) = f^{(\sigma)}(b_\omega);\ 0 \leq \sigma \leq s_\omega. \tag{11.1}$$

Further particular cases are :

LC:$(t = n,\ s_\omega = 0$ for all $\omega)$

$$\Lambda_L(b_\omega) = f(b_\omega), 1 \leq \omega \leq n,$$

Type$(\eta, n-\eta)$**C:** $(t = 2, 1 \leq \eta \leq n-1,\ s_1 = \eta - 1,\ s_2 = n - \eta - 1)$

$$\Lambda_{(\eta,n)}^{(\sigma)}(\alpha) = f^{(\sigma)}(\alpha), 0 \leq \sigma \leq \eta - 1,$$
$$\Lambda_{(\eta,n)}^{(\sigma)}(\beta) = f^{(\sigma)}(\beta), 0 \leq \sigma \leq n - \eta - 1,$$

Two-point TC: $(n = 2\eta, t = 2, s_1 = s_2 = \eta - 1)$

$$\Lambda_{2T}^{(\sigma)}(\alpha) = f^{(\sigma)}(\alpha), \Lambda_{2T}^{(\sigma)}(\beta) = f^{(\sigma)}(\beta), 0 \leq \sigma \leq \eta - 1.$$

The associated error $|E_H(z)|$ can be approximated by Green's function (Peano's Kernal) concerning boundary value problem for multiple points

$$z^{(n)}(x) = 0,\ z^{(\sigma)}(b_\omega) = 0,\ 0 \leq \sigma \leq s_\omega,\ 1 \leq \omega \leq t,$$

that is stated in the coming theorem:

Theorem 11.1 *[1] Let $-\infty < \alpha < \beta < \infty$ with $\alpha \leq b_1 < b_2 \cdots < b_t \leq \beta$, $(t \geq 2)$ be the given points, and $f \in C^n([\alpha, \beta])$. Then we have*

$$f(x) = \Lambda_H(x) + R_H(f, x) \tag{11.2}$$

where $\Lambda_H(x)$ represents Hermite polynomial, i.e.

$$\Lambda_H(x) = \sum_{\omega=1}^{t} \sum_{\sigma=0}^{s_\omega} H_{\sigma\omega}(x) f^{(\sigma)}(b_\omega);$$

where $H_{\sigma\omega}$ are Hermite basis given as

$$H_{\sigma\omega}(x) = \frac{1}{\sigma!} \frac{\vartheta(x)}{(x - b_\omega)^{s_\omega + 1 - \sigma}} \sum_{s=0}^{s_\omega - \sigma} \frac{1}{s!} \frac{d^s}{dx^s} \left(\frac{(x - b_\omega)^{s_\omega + 1}}{\vartheta(x)} \right) \Bigg|_{x = b_\omega} (x - b_\omega)^s, \tag{11.3}$$

with

$$\vartheta(x) = \prod_{\omega=1}^{t} (x - b_\omega)^{s_\omega + 1},$$

and remainder

$$R_H(f, x) = \int_{\alpha}^{\beta} G_{H,n}(x, r) f^{(n)}(r) dr$$

where $G_{H,n}(x, r)$ is defined by

$$G_{H,n}(x, r) = \begin{cases} \sum\limits_{\omega=1}^{l} \sum\limits_{\sigma=0}^{s_\omega} \frac{(b_\omega - r)^{n-\sigma-1}}{(n-\sigma-1)!} H_{\sigma\omega}(x);\ r \leq x, \\ -\sum\limits_{\omega=l+1}^{t} \sum\limits_{\sigma=0}^{s_\omega} \frac{(b_\omega - r)^{n-\sigma-1}}{(n-\sigma-1)!} H_{\sigma\omega}(x);\ r \geq x, \end{cases} \tag{11.4}$$

for all $b_l \leq r \leq b_{l+1}$; $l = 0, \ldots, t$ with $b_0 = \alpha$ and $b_{t+1} = \beta$.

Considering the particular cases for the Hermite conditions (**HC**), we have the following corollary:

Corollary 11.1 *For Lagrange conditions (**LC**), we have*

$$f(x) = \Lambda_L(x) + R_L(f, x) \tag{11.5}$$

where $\Lambda_L(x)$ indicates Lagrange interpolating polynomial i.e,

$$\Lambda_L(x) = \sum_{\omega=1}^{n} \prod_{\substack{s=1 \\ s \neq \omega}}^{n} \left(\frac{x - b_s}{b_\omega - b_s} \right) f(b_\omega)$$

with remainder $R_L(f,x)$

$$R_L(f,x)=\int_{\alpha}^{\beta} G_L(x,r)f^{(n)}(r)dr$$

with

$$G_L(x,r)=\frac{1}{(n-1)!}\begin{cases}\sum\limits_{\omega=1}^{l}(b_\omega-r)^{n-1}\prod\limits_{\substack{s=1\\ s\neq\omega}}^{n}\left(\frac{x-b_s}{b_\omega-b_s}\right), & r\leq x\\ -\sum\limits_{\omega=l+1}^{n}(b_\omega-r)^{n-1}\prod\limits_{\substack{s=1\\ s\neq\omega}}^{n}\left(\frac{x-b_s}{b_\omega-b_s}\right), & r\geq x\end{cases} \tag{11.6}$$

$b_l\leq r\leq b_{l+1}$, $l=1,2,...,n-1$ *with* $b_1=\alpha$ *and* $b_n=\beta$.

Considering type $(\eta,n-\eta)$ *conditions* (***Type***$(\eta,n-\eta)$***C***), *we get*

$$f(x)=\Lambda_{(\eta,n)}(x)+R_{(\eta,n)}(f,x) \tag{11.7}$$

where $\Lambda_{(\eta,n)}(x)$ *is* $(\eta,n-\eta)$ *interpolating polynomial, that is*

$$\Lambda_{(\eta,n)}(x)=\sum_{\sigma=0}^{\eta-1}\tau_\sigma(x)f^{(\sigma)}(\alpha)+\sum_{\sigma=0}^{n-\eta-1}\zeta_\sigma(x)f^{(\sigma)}(\beta),$$

with

$$\tau_\sigma(x)=\frac{1}{\sigma!}(x-\alpha)^\sigma\left(\frac{x-\beta}{\alpha-\beta}\right)^{n-\eta}\sum_{s=0}^{\eta-1-\sigma}\binom{n-\eta+s-1}{s}\left(\frac{x-\alpha}{\beta-\alpha}\right)^s \tag{11.8}$$

and

$$\zeta_\sigma(x)=\frac{1}{\sigma!}(x-\beta)^\sigma\left(\frac{x-\alpha}{\beta-\alpha}\right)^{\eta}\sum_{s=0}^{n-\eta-1-\sigma}\binom{\eta+s-1}{s}\left(\frac{x-\beta}{\alpha-\beta}\right)^s \tag{11.9}$$

along with the remainder $R_{(\eta,n)}(f,x)$, *given as*

$$R_{(\eta,n)}(f,x)=\int_{\alpha}^{\beta}G_{(\eta,n)}(x,r)f^{(n)}(r)dr$$

with

$$G_{(\eta,n)}(x,r)=\begin{cases}\sum\limits_{\omega=0}^{\eta-1}\left[\sum\limits_{p=0}^{\eta-1-\omega}\binom{n-\eta+p-1}{p}\left(\frac{x-\alpha}{\beta-\alpha}\right)^p\right]\times\\ \frac{(x-\alpha)^\omega(\alpha-r)^{n-\omega-1}}{\omega!(n-\omega-1)!}\left(\frac{\beta-x}{\beta-\alpha}\right)^{n-\eta}, & \alpha\leq r\leq x\leq\beta\\ -\sum\limits_{\sigma=0}^{n-\eta-1}\left[\sum\limits_{q=0}^{n-\eta-\sigma-1}\binom{\eta+q-1}{q}\left(\frac{\beta-x}{\beta-\alpha}\right)^q\right]\times\\ \frac{(x-\beta)^\sigma(\beta-r)^{n-\sigma-1}}{\sigma!(n-\sigma-1)!}\left(\frac{x-\alpha}{\beta-\alpha}\right)^{\eta}, & \alpha\leq x\leq r\leq\beta.\end{cases} \tag{11.10}$$

For Type Two-point Taylor conditions (***Two-point TC***), *from Theorem 11.1 we have*

$$f(x)=\Lambda_{2T}(x)+R_{2T}(f,x) \tag{11.11}$$

where $\Lambda_{2T}(x)$*is the two-point Taylor interpolating polynomial i.e,*

$$\Lambda_{2T}(x)=\sum_{\sigma=0}^{\eta-1}\sum_{s=0}^{\eta-1-\sigma}\binom{\eta+s-1}{s}\Big[\frac{(x-\alpha)^{\sigma}}{\sigma!}\Big(\frac{x-\beta}{\alpha-\beta}\Big)^{\eta}\Big(\frac{x-\alpha}{\beta-\alpha}\Big)^{s}f^{(\sigma)}(\alpha)$$
$$+\frac{(x-\beta)^{\sigma}}{\sigma!}\Big(\frac{x-\alpha}{\beta-\alpha}\Big)^{\eta}\Big(\frac{x-\beta}{\alpha-\beta}\Big)^{s}f^{(\sigma)}(\beta)\Big] \quad (11.12)$$

and the remainder $R_{2T}(f,x)$ *is given by*

$$R_{2T}(f,x)=\int_{\alpha}^{\beta}G_{2T}(x,r)f^{(n)}(r)dr$$

with

$$G_{2T}(x,r)=\begin{cases}\frac{(-1)^{\eta}}{(2\eta-1)!}p^{\eta}(x,r)\sum\limits_{\omega=0}^{\eta-1}\binom{\eta-1+\omega}{\omega}(x-r)^{\eta-1-\omega}q^{\omega}(x,r), & r\leq x;\\ \frac{(-1)^{\eta}}{(2\eta-1)!}q^{\eta}(x,r)\sum\limits_{\omega=0}^{\eta-1}\binom{\eta-1+\omega}{\omega}(r-x)^{\eta-1-\omega}p^{\omega}(x,r), & r\geq x;\end{cases} \quad (11.13)$$

where $p(x,r)=\frac{(r-\alpha)(\beta-x)}{\beta-\alpha}$, $q(x,r)=p(x,r),\forall x,r\in[\alpha,\beta]$.

(Beesack [10] and Levin [62]) characterize the non-negativity of Green's function $G_{H,n}(x,r)$:

Lemma 11.1 *(i)* $\frac{G_{H,n}(x,r)}{\vartheta(x)}>0,\ \ \mathrm{b}_1\leq \mathrm{x}\leq \mathrm{b}_t, \mathrm{b}_1\leq \mathrm{r}\leq \mathrm{b}_t.$

(ii) $G_{H,n}(x,r)\leq\frac{1}{(n-1)!(\beta-\alpha)}|\vartheta(x)|.$

(iii) $\int\limits_{\alpha}^{\beta}G_{H,n}(x,r)dr=\frac{\vartheta(x)}{n!}.$

11.1 Extensions of cyclic refinements of Jensen's inequality by Hermite interpolation

We start this section by considering the discrete as well as continuous version of cyclic refinements of Jensen's inequality and construct the generalized new identities having real weights utilizing Hermite interpolating polynomial.

Theorem 11.2 *Suppose* $m,k\in\mathbb{N}$, $p_1,\ldots,p_m$ *and* $\lambda_1,\ldots,\lambda_k$ *are real tuples for* $2\leq k\leq m$, *such that* $\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\neq 0$ *for* $i=1,\ldots,m$ *with* $\sum\limits_{i=1}^{m}p_i=1$ *and* $\sum\limits_{j=1}^{k}\lambda_j=1$. *Also let* $x\in[\alpha,\beta]\subset\mathbb{R}$ *and* $\mathbf{x}\in[\alpha,\beta]^m$. *Assume* $f\in C^n[\alpha,\beta]$ *and consider interval with points* $-\infty<\alpha=b_1<b_2\cdots<b_t=\beta<\infty$, $(t\geq 2)$, $H_{\sigma\omega}$ *in* (11.3) *are Hermite basis,* $G_{H,n}(x,r)$ *in* (11.4) *be the Hermite Green function and* G_v, $(v=1,\ldots,5)$ *be the Green functions defined in* (6.1)–(6.5), *respectively. Then for* $u=1,\ldots,6$ *along with assumptions* (A_1) *and* (A_2), *we have the following generalized identities:*

(a)

$$J_u(f(x)) = \sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma)}(b_\omega)J_u\Big(H_{\sigma\omega}(x)\Big) + \int_{\alpha}^{\beta} J_u\Big(G_{H,n}(x,r)\Big)f^{(n)}(r)dr. \quad (11.14)$$

(b)

$$J_u(f(x)) = C_v(\alpha,\beta,f)J_u(x) + \int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma+2)}(b_\omega)H_{\sigma\omega}(r)dr$$

$$+\int_{\alpha}^{\beta}\int_{\alpha}^{\beta} J_u\Big(G_v(x,r)\Big)G_{H,n-2}(r,\xi))f^{(n)}(\xi)d\xi dr \quad (11.15)$$

where C_v, $(v = 1,\ldots,5)$ is defined in (7.5).

Proof. Fix $u = 1,\ldots,6$.

(a) Applying cyclic Jensen's type linear functionals $J_u(\cdot)$ on (11.2) and practicing properties of the functional, we get (11.14).

(b) For fix $v = 2$, testing (6.7) in cyclic Jensen's type functional $J_u(\cdot)$ and employing the properties of $J_u(\cdot)$ along with the assumed conditions, we have

$$J_u(f) = C_2(\alpha,\beta,f)J_u(x) + \int_{\alpha}^{\beta} J_u\Big(G_2(x,r)\Big)f''(r)dr. \quad (11.16)$$

By Theorem 11.1, $f''(r)$ can be expressed as:

$$f''(r) = \sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} H_{\sigma\omega}(r)f^{(\sigma+2)}(b_\omega) + \int_{\alpha}^{\beta} G_{H,n-2}(r,\xi)f^{(n)}(\xi)d\xi. \quad (11.17)$$

Putting (11.17) in (11.16), we get (11.15) respectively for $v = 2$ and $u = 1,\ldots,6$. The cases for $v = 1,3,4,5$, are treated analogously and are left for the reader interest.

□

Now we obtain extensions and improvements of discrete and integral cyclic Jensen type linear functionals, with real weights.

Theorem 11.3 *Consider f be n-convex function along with the suppositions of Theorem 11.2. Then we conclude the following results:*

(a) If for all $u = 1,\ldots,6$,

$$J_u\Big(G_{H,n}(x,r)\Big) \geq 0, \quad r \in [\alpha,\beta] \quad (11.18)$$

holds, then we have

$$J_u(f(x)) \geq \sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma)}(b_\omega)J_u\Big(H_{\sigma\omega}(x)\Big) \quad (11.19)$$

for $u = 1,\ldots,6$.

(b) *If for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$*

$$J_u\Big(G_v(x,r)\Big) \geq 0, \;\; r \in [\alpha,\beta] \tag{11.20}$$

holds, provided that s_ω is odd for each $\omega = 2,3,4,\cdots,t$, then

$$J_u(f(x)) \geq C_v(\alpha,\beta,f)J_u(x) + \int_\alpha^\beta J_u\Big(G_v(x,r)\Big)\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma+2)}(b_\omega)H_{\sigma\omega}(r)dr. \tag{11.21}$$

for $u = 1,\ldots,6$.

(c) *If* (11.20) *holds for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$, provided that s_ω is odd for each $\omega = 2,3,4,\cdots,t-1$ and s_t is even then* (11.21) *holds in reverse direction for $u = 1,\ldots,6$.*

Proof.

(a) Fix $u = 1,\ldots,6$.
As the function $f \in C^n[\alpha,\beta]$ and assumed to be $n-$convex, therefore using the characterization of $n-$convex function $f^{(n)}(x) \geq 0$ for all $x \in [\alpha,\beta]$ (see [82], p. 16). Hence we can apply Theorem 11.2(a) to obtain (11.19).

(b) Fix $u = 1,\ldots,6$ and $v = 1,\ldots,5$.
As we have discussed in part(a) $f^{(n)}(x) \geq 0$ for all $x \in [\alpha,\beta]$. Also as it is given that s_ω is odd for each $\omega = 2,3,4,\cdots,t$, we have $\vartheta(r) = \prod_{\omega=1}^{t}(r-b_\omega)^{s_\omega+1} \geq 0$ for any $r \in [\alpha,\beta]$ therefore taking into account Lemma 11.1 (i) we have $G_{H,n-2}(r,\xi)) \geq 0$. Thus by applying Theorem 11.2(b) yields (11.21).

(c) If s_t is even then $(r-b_t)^{s_t+1} \leq 0$ for any $r \in [\alpha,\beta]$. Also clearly $(r-b_1)^{s_1+1} \geq 0$ for any $r \in [\alpha,\beta]$ and $\prod_{\omega=2}^{t-1}(r-b_\omega)^{s_\omega+1} \geq 0$ for $r \in [\alpha,\beta]$ if s_ω is odd for each $\omega = 2,3,4,\cdots,t-1$. Therefore summing this information $\vartheta(r) = \prod_{\omega=1}^{t}(r-b_\omega)^{s_\omega+1} \leq 0$ for any $r \in [\alpha,\beta]$ and taking into account Lemma 11.1 (i) we have $G_{H,n-2}(r,\xi)) \leq 0$. Thus by applying Theorem 11.2(b) gives (11.21) in reverse direction. □

By using **Type**$(\eta,n-\eta)$**C:** $(t = 2, 1 \leq \eta \leq n-1,\; s_1 = \eta - 1,\; s_2 = n-\eta-1)$ we give the following result:

Corollary 11.2 *Consider f be n-convex function along with the suppositions of Theorem 11.2. Let $G_{(\eta,n)}$ be the Green's function defined in* (11.10) *and τ_σ, ζ_σ be as defined* (11.8), (11.9) *respectively. Then we conclude the following results:*

(a) *If for all $u = 1,\ldots,6$,*

$$J_u\Big(G_{(\eta,n)}(z,r)\Big) \geq 0, \quad r \in [\alpha,\beta] \tag{11.22}$$

holds, then we have

$$J_u(f(z)) \geq \sum_{\sigma=0}^{\eta-1} f^{(\sigma)}(\alpha) J_u\Big(\tau_\sigma(z)\Big) + \sum_{\sigma=0}^{n-\eta-1} f^{(\sigma)}(\beta) J_u\Big(\zeta_\sigma(z)\Big) \tag{11.23}$$

for $u = 1, \ldots, 6$.

(b) *If* (11.20) *holds for all* $u = 1, \ldots, 6$ *and* $v = 1, \ldots, 5$, *provided that* $n - \eta$ *is even, then*

$$J_u(f(z)) \geq C_v(\alpha, \beta, f) J_u(x) + \int_\alpha^\beta J_u\Big(G_v(z,r)\Big) \left(\sum_{\sigma=0}^{\eta-1} f^{(\sigma+2)}(\alpha)\tau_\sigma(r)\Big) + \sum_{\sigma=0}^{n-\eta-1} f^{(\sigma+2)}(\beta)\zeta_\sigma(r) \right) dr. \tag{11.24}$$

for $u = 1, \ldots, 6$.

(c) *If* (11.20) *holds for all* $u = 1, \ldots, 6$ *and* $v = 1, \ldots, 5$, *provided that* $n - \eta$ *is odd then* (11.24) *holds in reverse direction for* $u = 1, \ldots, 6$.

By using **Two-point TC:** $(n = 2\eta, t = 2, s_1 = s_2 = \eta - 1)$ we give the following result:

Corollary 11.3 *Consider* f *be* n*-convex function along with the suppositions of Theorem 11.2. Let* G_{2T} *be the Green's function defined in* (11.13), *then we conclude the following results:*

(a) *If for all* $u = 1, \ldots, 6$,

$$J_u\Big(G_{2T}(z,r)\Big) \geq 0, \quad r \in [\alpha, \beta] \tag{11.25}$$

holds, then we have

$$\begin{aligned} J_u(f(z)) \geq & \sum_{\sigma=0}^{\eta-1} \sum_{s=0}^{\eta-1-\sigma} \binom{\eta+s-1}{s} \times \\ & \times \Bigg[f^{(\sigma)}(\alpha) J_u\Big(\frac{(z-\alpha)^\sigma}{\sigma!} \Big(\frac{z-\beta}{\alpha-\beta}\Big)^\eta \Big(\frac{z-\alpha}{\beta-\alpha}\Big)^s \Big) \\ & + f^{(\sigma)}(\beta) J_u\left(\frac{(z-\beta)^\sigma}{\sigma!} \Big(\frac{z-\alpha}{\beta-\alpha}\Big)^\eta \Big(\frac{z-\beta}{\alpha-\beta}\Big)^s \right) \Bigg] \end{aligned} \tag{11.26}$$

for $u = 1, \ldots, 6$.

(b) *If* (11.20) *holds for all* $u = 1, \ldots, 6$ *and* $v = 1, \ldots, 5$, *provided that* η *is even, then*

$$\begin{aligned} J_u(f(z)) \geq C_v(\alpha, \beta, f) J_u(x) + & \int_\alpha^\beta J_u\Big(G_v(z,r)\Big) \sum_{\sigma=0}^{\eta-1} \sum_{s=0}^{\eta-1-\sigma} \binom{\eta+s-1}{s} \times \\ \times \Big[& \frac{(r-\alpha)^\sigma}{\sigma!} \Big(\frac{r-\beta}{\alpha-\beta}\Big)^\eta \Big(\frac{r-\alpha}{\beta-\alpha}\Big)^s f^{(\sigma)}(\alpha) \\ & + \frac{(r-\beta)^\sigma}{\sigma!} \Big(\frac{r-\alpha}{\beta-\alpha}\Big)^\eta \Big(\frac{r-\beta}{\alpha-\beta}\Big)^s f^{(\sigma)}(\beta) \Big] dr. \end{aligned} \tag{11.27}$$

for $u = 1, \ldots, 6$.

(c) *If* (11.20) *holds for all* $u = 1,\ldots,6$ *and* $v = 1,\ldots,5$, *provided that* η *is odd then* (11.27) *holds in reverse direction for* $u = 1,\ldots,6$.

We will finish the present section by the following generalizations of cyclic refinements of Jensen inequalities:

Theorem 11.4 *If the assumptions of Theorem 11.2 be fulfilled with additional conditions that* $p_1,\ldots,p_m$ *and* $\lambda_1,\ldots,\lambda_k$ *be non negative tuples for* $2 \le k \le m$, *such that* $\sum\limits_{i=1}^{m} p_i = 1$ *and* $\sum\limits_{j=1}^{k} \lambda_j = 1$. *Then for* $f : [\alpha,\beta] \to \mathbb{R}$ *being n-convex function, we conclude the following results:*

(a) If (11.19) is valid along with the function

$$\Gamma(z) := \sum_{\omega=1}^{t} \sum_{\sigma=0}^{s_\omega} H_{\sigma\omega}(z) f^{(\sigma)}(b_\omega). \tag{11.28}$$

to be convex, the right side of (11.19) is non negative, means

$$J_u(f) \ge 0, \qquad u = 1,\ldots,6. \tag{11.29}$$

(b) If s_ω to be odd for each $\omega = 2,3,4,\cdots,t$, (11.21) holds. Further

$$\sum_{\omega=1}^{t} \sum_{\sigma=0}^{s_\omega} H_{\sigma\omega}(r) f^{(\sigma+2)}(b_\omega) \ge 0. \tag{11.30}$$

the right side of (11.21) is non negative, particularly (11.29) is establish for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$.

(c) Inequality (11.21) holds reversely if s_ω is odd for each $\omega = 2,3,4,\cdots,t-1$ and s_t is even. Moreover, let (11.30) holds in reverse direction then reverse of (11.29) holds for all $u = 1,\ldots,6$ and $v = 1,\ldots,5$.

Proof.

(a) Fix $u = 1,\ldots,6$.
As (11.19) is valid, the R.H.S. of (11.19) can be written in the functional form $J_u(\Gamma)$ for all $(u = 1,\ldots,6)$, after reorganizing this side. Employing Remark 2.7 the non-negativity of R.H.S. of (11.19) is secure, especially (11.29) is establish.

(b) Fix $u = 1,\ldots,6$.
We have assumed positive weights and for all $v = 1,\ldots,5$, $G_v(z,r)$ is convex. Thus by practicing Remark 2.7, $J_u(G_v(z,r)) \ge 0$. As f is n-convex and s_ω odd for each $\omega = 2,3,4,\cdots,t$, hence by following Theorem 11.3 (b), we obtain (11.21). Now taking into account the positivity of $J_u(G_v(z,r))$ and (11.28), we get (11.29).

(c) Similar to the proof of (b) □

Remark 11.1 We left for reader interest to give generalizations of cyclic refinements of Jensen inequalities by considering **Type**$(\eta, n-\eta)$**C** and **Two-point TC** instead of **HC** in Theorem 11.4.

11.2 Applications to information theory

Now as a consequence of Theorem 11.3 we consider the discrete extensions of cyclic refinements of Jensen's inequalities for $(u = 1)$, from (11.19) with respect to n-convex function f in the explicit form:

$$\begin{aligned}
&\sum_{i=1}^{m} p_i f(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right) \\
&\quad\geq \left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma)}(b_\omega)\right)\times \\
&\quad\times\left(\sum_{i=1}^{m} p_i H_{\sigma\omega}(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}\right) H_{\sigma\omega}\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)\right),
\end{aligned} \tag{11.31}$$

where $H_{\sigma\omega}$ are Hermite basis defined in (11.3).

Theorem 11.5 *Let $m,k \in \mathbb{N}$ $(2 \leq k \leq m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution. Let $\mathbf{p} := (p_1,\ldots,p_m) \in \mathbb{R}^m$, and $\mathbf{q} := (q_1,\ldots,q_m) \in (0,\infty)^m$ such that*

$$\frac{p_i}{q_i} \in [\alpha,\beta], \quad u = 1,\ldots,m.$$

Also let $f \in C^n[\alpha,\beta]$ and consider interval with points $-\infty < \alpha = b_1 < b_2 \cdots < b_t = \beta < \infty$, $(t \geq 2)$ such that f is n-convex function. Then the following inequalities hold:

$$\begin{aligned}
\widetilde{I}_f(\mathbf{p},\mathbf{q}) &\geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) \\
&\quad+\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega} f^{(\sigma)}(b_\omega)\right)\times \\
&\quad\times\left(\sum_{i=1}^{m} q_i H_{\sigma\omega}\left(\frac{p_i}{q_i}\right) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) H_{\sigma\omega}\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right).
\end{aligned} \tag{11.32}$$

Proof. Replacing p_i with q_i and x_i with $\frac{p_i}{q_i}$ for $(i = 1,\ldots,m)$ in (11.31), we get (11.32). □

We now explore two exceptional cases of the previous result.

One corresponds to the entropy of a discrete probability distribution.

Corollary 11.4 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m) \in (0,\infty)^m$ and n is even, then

$$\begin{aligned}\sum_{i=1}^{m} q_i \ln q_i \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\\ +\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^{\sigma}(\sigma-1)!}{(b_\omega)^{\sigma}}\right)\times\\ \times\left(\sum_{i=1}^{m} q_i H_{\sigma\omega}\left(\frac{1}{q_i}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)H_{\sigma\omega}\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right).\end{aligned} \tag{11.33}$$

(b) If $\mathbf{q} := (q_1,\ldots,q_m)$ is a positive probability distribution and n is even, then we get the bounds for the Shannon entropy of $\mathbf{q}$.

$$\begin{aligned}H(\mathbf{q}) \le -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\\ -\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^{\sigma}(\sigma-1)!}{(b_\omega)^{\sigma}}\right)\times\\ \times\left(\sum_{i=1}^{m} q_i H_{\sigma\omega}\left(\frac{1}{q_i}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)H_{\sigma\omega}\left(\frac{1}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right).\end{aligned} \tag{11.34}$$

If n is odd, then (11.33) *and* (11.34) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$, and $\mathbf{p} := (1,1,\ldots,1)$ in Theorem 11.5, we get the required result.

(b) It is a special case of (a). □

The second case corresponds to the relative entropy or Kullback-Leibler divergence between two probability distributions.

Corollary 11.5 *Let $m,k \in \mathbb{N}$ $(2 \le k \le m)$, $\lambda_1,\ldots,\lambda_k$ be positive probability distribution.*

(a) If $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m) \in (0,\infty)^m$ and n is even, then

$$\sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) \ge \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$

$$+\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m} q_i H_{\sigma\omega}\left(\frac{p_i}{q_i}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)H_{\sigma\omega}\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{11.35}$$

(b) *If If* $\mathbf{q} := (q_1,\ldots,q_m), \mathbf{p} := (p_1,\ldots,p_m)$ *are positive probability distributions and* n *is even, then we have*

$$D(\mathbf{q}\parallel\mathbf{p}) \geq \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}\right)$$
$$+\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m} q_i H_{\sigma\omega}\left(\frac{p_i}{q_i}\right)-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)H_{\sigma\omega}\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)\right). \tag{11.36}$$

If n *is odd, then* (11.35) *and* (11.36) *hold in reverse directions.*

Proof.

(a) Using $f(x) := -\ln x$ in Theorem 11.5, we get the desired result.

(b) It is special case of (a).

□

Theorem 11.6 *Let* $m,k\in\mathbb{N}$ $(2\leq k\leq m)$, $\lambda_1,\ldots,\lambda_k$ *be positive probability distribution and* $\mathbf{q}$ *be as defined in* (6.59) *by Zipf-Mandelbrot law with parameters* $m\in\{1,2,\ldots\}$, $c\geq 0$, $d>0$. *If* n *is even, we obtain*

$$S(\mathbf{q}) = Z(H,c,d)$$
$$\leq -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^d H_{m,c,d})}\right)\ln\left(\frac{1}{H_{m,c,d}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+t)^s)}\right)$$
$$-\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+c)^d H_{m,c,d})}H_{\sigma\omega}\left(((i+c)^d H_{m,c,d})\right)\right)$$

$$+\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^dH_{m,c,d})}\right)H_{\sigma\omega}\left(\frac{1}{\sum\limits_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^dH_{m,c,d})}}\right)\right). \quad (11.37)$$

If n is odd, then (11.37) *holds in reverse direction.*

Proof. Similar to that of Theorem 6.5. □

Corollary 11.6 *Let $m,k\in\mathbb{N}$ $(2\le k\le m)$, $\lambda_1,\dots,\lambda_k$ be positive probability distribution and for $c_1,c_2\in[0,\infty)$, $d_1,d_2>0$, let $H_{m,c_1,d_1}=\frac{1}{(i+c_1)^{d_1}}$ and $H_{m,c_2,d_2}=\frac{1}{(i+c_2)^{d_2}}$. Now using $q_i=\dfrac{1}{(i+c_1)^{d_1}H_{m,c_1,d_1}}$ and $p_i=\dfrac{1}{(i+c_2)^{d_2}H_{m,c_2,d_2}}$ in Corollary 11.5(b), with even n, we obtain*

$$D(\mathbf{q}\parallel\mathbf{p})=\sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H_{m,c_1,d_1}}\ln\left(\frac{(i+c_2)^{d_2}H_{m,c_2,d_2}}{(i+c_1)^{d_1}H_{m,c_1,d_1}}\right)$$
$$\ge\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{m,c_1,d_1}}\right)\ln\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+c_1)^{d_1}H_{m,c_1,d_1}}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+c_2)^{d_2}H_{m,c_2,d_2}}}\right)$$
$$+\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\left(\sum_{i=1}^{m}\frac{1}{((i+c_1)^{d_1}H_{m,c_1,d_1})}H_{\sigma\omega}\left(\frac{((i+c_2)_2^dH_{m,c_2,d_2})}{((i+c_1)^dH_{m,c_1,d_1})}\right)\right)$$
$$-\left(\sum_{\omega=1}^{t}\sum_{\sigma=0}^{s_\omega}\frac{(-1)^\sigma(\sigma-1)!}{(b_\omega)^\sigma}\right)\times$$
$$\times\left(\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H_{m,c_1,d_1}}\right)H_{\sigma\omega}\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+c_1)^{d_1}H_{m,c_1,d_1}}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}\frac{1}{(i+j+c_2)^{d_2}H_{m,c_2,d_2}}}\right)\right)$$
(11.38)

If n is odd, then (11.38) *holds in reverse direction.*

Remark 11.2 It is interesting to note that, in the similar passion we are able to construct different estimations of f-divergences along with their applications to Shannon and Mandelbrot entropies using the other inequalities for n-convex functions constructed in Theorem 11.3 for discrete case of cyclic refinements of Jensen inequality.

Remark 11.3 We left for reader interest to construct upper bounds for Shannon, Relative and Mandelbrot entropies by considering **Type**$(\eta,n-\eta)$**C** and **Two-point TC** instead of **HC** in the above results.

Chapter 12

Levinson's Type Generalization of Cyclic Refinements of Jensen's Inequality and Related Applications

We present Levinsons type generalizations of cyclic refinements of Jensen's inequality by employing recent class of functions that further characterize and extend the class of $3-$convex functions. We get monotonic cyclic Jensen's inequalities and particularly the renowned Jensen's inequality for $3-$convex function at a point $(f \in \kappa_1^c(I)))$. As an applications in information theory, we first introduce new Csiszár type cyclic divergence functional for $3-$convex functions and establish cyclic-Kullback-Leibler and Hellinger distnaces. We also obtained monotonic Shannon, Relative and Zipf-Mandelbrot entropies.
We start by recent class of functions that further characterize and extend the class of $3-$convex functions:

Definition 12.1 *[7] Let $f : I \to \mathbb{R}$ and $c \in I^0$, where I is an arbitrary interval(open, closed or semi-open in either direction) in $\mathbb{R}$ and I^0, is its interior. We say that $f \in \kappa_1^c(I)$ $(resp.\ f \in \kappa_2^c(I))$ that is f is 3-convex function at point c (respectively 3-concave function at point c) if there exists a constant A such that the function $F(x) = f(x) - \frac{A}{2}x^2$ is concave(resp. convex) on $I \bigcap (-\infty, c]$ and convex(resp. concave) on $I \bigcap [c, \infty)$.*

Remark 12.1 It is interesting to note that if $f : I \to \mathbb{R}$ is 3-convex (3-concave), then $f \in \kappa_1^c(I)$ $(resp.\ f \in \kappa_2^c(I))$ that is f is 3-convex function at point c for every $c \in I$. Moreover Pečarić et al. in [7] proved that $f \in \kappa_1^c(I)$ is the largest class of functions for which Levinson's inequality holds.

12.1 Cyclic refinements of Jensen's inequalities for 3-convex function at a point

We need the following assumptions to move on:
(H_1^*) Let $2 \le k \le m$, $2 \le l \le n$ be integers such that $q_1,\dots,q_m$ and $\lambda_1,\dots,\lambda_k$; $p_1,\dots,p_n$ and $\omega_1,\dots,\omega_l$ represent positive probability distributions.

Theorem 12.1 *Assume* (H_1^*)*. If* $x_1,\dots,x_m \in I^m;y_1,\dots,y_n \in I^n$ *with*

$$\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^2-\left(\sum_{i=1}^{m}q_ix_i\right)^2$$
$$=\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^2-\left(\sum_{r=1}^{n}p_ry_r\right)^2 \tag{12.1}$$

and also there exists $c \in I^o$ *such that*

$$\max_i\ x_i \le c \le \min_r\ y_r. \tag{12.2}$$

If $f \in K_1^c$*, then following inequality*

$$\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)-f\left(\sum_{i=1}^{m}q_ix_i\right)$$
$$\le\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)f\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)-f\left(\sum_{r=1}^{n}p_ry_r\right) \tag{12.3}$$

holds.

Proof. Since $f \in K_1^c(I)$, then we have a constant A such that $F(x) = f(x) - \frac{A}{2}x^2$ is concave on $I \bigcap (-\infty, c]$, so by the reverse of left inequality of (2.14) for $x_1, ..., x_m \in I \bigcap (-\infty, c]$, we have

$$0 \geq \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) F \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) - F \left(\sum_{i=1}^{m} q_i x_i \right)$$

$$= \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) - f \left(\sum_{i=1}^{m} q_i x_i \right)$$

$$- \frac{A}{2} \left[\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)^2 - \left(\sum_{i=1}^{m} q_i x_i \right)^2 \right] \tag{12.4}$$

Also, using the fact that $F(y) = f(y) - \frac{A}{2}y^2$ is convex on $I \bigcap [c, \infty)$, so by left inequality of (1) for $y_1, ..., y_n \in I \bigcap [c, \infty)$, we have

$$0 \leq \sum_{r=1}^{n} \left(\sum_{s=0}^{l-1} \omega_{s+1} p_{s+r} \right) F \left(\frac{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} y_{r+s}}{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s}} \right) - F \left(\sum_{r=1}^{n} p_r y_r \right)$$

$$= \sum_{r=1}^{n} \left(\sum_{s=0}^{l-1} \omega_{s+1} p_{s+r} \right) f \left(\frac{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} y_{r+s}}{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s}} \right) - f \left(\sum_{r=1}^{n} p_r y_r \right)$$

$$- \frac{A}{2} \left[\sum_{r=1}^{n} \left(\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} \right) \left(\frac{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} y_{r+s}}{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s}} \right)^2 - \left(\sum_{r=1}^{n} p_r y_r \right)^2 \right] \tag{12.5}$$

Rearranging above inequalities, we have

$$\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) f \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) - f \left(\sum_{i=1}^{m} q_i x_i \right)$$

$$-\frac{A}{2}\left[\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^{2}-\left(\sum_{i=1}^{m}q_ix_i\right)^{2}\right]$$

$$\leq 0 \leq$$

$$\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)f\left(\frac{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)-f\left(\sum_{r=1}^{n}p_ry_r\right)$$

$$-\frac{A}{2}\left[\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^{2}-\left(\sum_{r=1}^{n}p_ry_r\right)^{2}\right]$$

So,

$$\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)-f\left(\sum_{i=1}^{m}q_ix_i\right)$$

$$-\frac{A}{2}\left[\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^{2}-\left(\sum_{i=1}^{m}q_ix_i\right)^{2}\right]$$

$$\leq \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)f\left(\frac{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)-f\left(\sum_{r=1}^{n}p_ry_r\right)$$

$$-\frac{A}{2}\left[\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^{2}-\left(\sum_{r=1}^{n}p_ry_r\right)^{2}\right]$$

Using (12.1), we get (12.3). □

The next result can be obtained by using an appropriate inequality from (2.14) with same idea of proof as in above Theorem 12.1.

Theorem 12.2 *Assume* (H_1^*). *If* $x_1,\dots,x_m \in I^m; y_1,\dots,y_n \in I^n$ *with*

$$\sum_{i=1}^{m} q_i\,(x_i)^2 - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^2$$
$$= \sum_{r=1}^{n} p_r\,(y_r)^2 - \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^2 \tag{12.6}$$

and also (12.2) *holds.*

If $f \in K_1^c$, *then following inequality*

$$\sum_{i=1}^{m} q_i f\,(x_i) - \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)$$
$$\leq \sum_{r=1}^{n} p_r f\,(y_r) - \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right) f\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right) \tag{12.7}$$

holds.

In the next theorem, we state well known discrete Jensen inequality for this new class of functions, however the idea of the proof is similar to that adopted in Theorem 12.1:

Theorem 12.3 *Assume* (H_1^*). *If* $x_1,\dots,x_m \in I^m; y_1,\dots,y_n \in I^n$ *with*

$$\sum_{i=1}^{m} q_i(x_i)^2 - \left(\sum_{i=1}^{m} q_i x_i\right)^2 = \sum_{r=1}^{n} p_r(y_r)^2 - \left(\sum_{r=1}^{n} p_r y_r\right)^2 \tag{12.8}$$

and also (12.2) *holds. Now, if* $f \in \kappa_1^c(I)$, *then following inequality*

$$\sum_{i=1}^{m} q_i f(x_i) - f\left(\sum_{i=1}^{m} q_i x_i\right) \leq \sum_{r=1}^{n} p_r f(y_r) - f\left(\sum_{r=1}^{n} p_r y_r\right) \tag{12.9}$$

holds.

The next result weakens the assumption (12.1) of Theorem 12.1 for $f \in \kappa_1^c(I)$.

Theorem 12.4 *Assume* (H_1^*). *If* $x_1,\dots,x_m \in I^m; y_1,\dots,y_n \in I^n$ *with* (12.2) *holds and* $f \in \kappa_1^c(I)$ *for some* $c \in [max\ x_i, min\ y_r]$. *Now, if*
(*a*)
$$f_-''(max\ x_i) \geq 0$$
and
$$\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^2-\left(\sum_{i=1}^{m}q_ix_i\right)^2$$
$$\leq \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^2-\left(\sum_{r=1}^{n}p_ry_r\right)^2 \quad (12.10)$$
or
(*b*)
$$f_+''(min\ y_r) \leq 0$$
and
$$\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^2-\left(\sum_{i=1}^{m}q_ix_i\right)^2$$
$$\geq \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^2-\left(\sum_{r=1}^{n}p_ry_r\right)^2 \quad (12.11)$$
or
(*c*)
$$f_-''(max\ x_i) < 0 < f_+''(min\ y_r) \text{ and } f \text{ is } 3-convex,$$
then (12.3) *holds.*

Proof. As $f \in \kappa_1^c(I)$, then we have a constant A such that $F(x) = f(x) - \frac{A}{2}x^2$ is concave on $I\bigcap(-\infty,c]$ and $F(y) = f(y) - \frac{A}{2}y^2$ is convex on $I\bigcap[c,\infty)$. For $x_i \in I\bigcap(-\infty,c] (i = 1,\dots,m)$, we have (12.4) and for $y_r \in I\bigcap[c,\infty) (r = 1,\dots,n)$, we have (12.5).
Now, using (12.4) and (12.5), we get
$$\frac{A}{2}\left[\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^2-\left(\sum_{r=1}^{n}p_ry_r\right)^2\right]$$
$$-\frac{A}{2}\left[\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^2-\left(\sum_{i=1}^{m}q_ix_i\right)^2\right]$$

$$\leq \sum_{r=1}^{n} f\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right) f\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right) - f\left(\sum_{r=1}^{n}p_r y_r\right)$$

$$-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) - f\left(\sum_{i=1}^{m}q_i x_i\right)$$

Now, due to the concavity of $F(x)$ and convexity of $F(y)$, for every distinct point $\tilde{x_u} \in I\bigcap(-\infty, max\ x_i]$ and $\tilde{y_u} \in I\bigcap[min\ y_r, \infty)$, $u = 1,2,3$, we have

$$[\tilde{x_1}, \tilde{x_2}, \tilde{x_3}]f \leq A \leq [\tilde{y_1}, \tilde{y_2}, \tilde{y_3}]f$$

Letting $\tilde{x_u} \nearrow max\ x_i$ and $\tilde{y_u} \searrow min\ y_r$, we get (if exists)

$$f_{-}^{''}(max\ x_i) \leq A \leq f_{+}^{''}(min\ y_r)$$

Therefore, if the assumptions (a) or (b) holds, then

$$\frac{A}{2}\left[\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}p_{r+s}\right)\left(\frac{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}y_{r+s}}{\sum\limits_{s=0}^{l-1}\omega_{s+1}p_{r+s}}\right)^{2} - \left(\sum_{r=1}^{n}p_r y_r\right)^{2}\right.$$

$$\left. -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\left(\frac{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}x_{i+j}}{\sum\limits_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right)^{2} + \left(\sum_{i=1}^{m}q_i x_i\right)^{2}\right]$$

is positive and we conclude the result. If the assumption (c) holds, the $f_{-}^{''}$ is left continuous, $f_{+}^{''}$ is right continuous, they are both nondecreasing and $f_{-}^{''} \leq f_{+}^{''}$.
Therefore, there exists $\bar{c} \in [max\ x_i, min\ y_r]$, such that $f \in \kappa_1^{\bar{c}}(I)$ with associated constant $\tilde{A} = 0$ and we can again deduce the result. □

Now next results that weakens the assumption (12.6) of Theorem 12.2 and (12.8) of Theorem 12.3 respectively for $\in \kappa_1^c(I)$.

Theorem 12.5 *Assume* (H_1^*). *If* $x_1,\ldots,x_m \in I^m; y_1,\ldots,y_n \in I^n$ *with* (12.2) *holds and* $f \in \kappa_1^c(I)$ *for some* $c \in [max\ x_i, min\ y_r]$. *Now, if*
(a)

$$f_{-}^{''}(max\ x_i) \geq 0$$

and

$$\sum_{i=1}^{m} q_i (x_i)^2 - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)^2$$

$$\leq \sum_{r=1}^{n} p_r (y_r)^2 - \sum_{r=1}^{n} \left(\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} \right) \left(\frac{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} y_{r+s}}{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s}} \right)^2, \quad (12.12)$$

or
(b)

$$f_+^{''}(min\ y_r) \leq 0$$

and

$$\sum_{i=1}^{m} q_i (x_i)^2 - \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) \left(\frac{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} x_{i+j}}{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)^2$$

$$\geq \sum_{r=1}^{n} p_r (y_r)^2 - \sum_{r=1}^{n} \left(\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} \right) \left(\frac{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s} y_{r+s}}{\sum_{s=0}^{l-1} \omega_{s+1} p_{r+s}} \right)^2, \quad (12.13)$$

or
(c)

$$f_-^{''}(max\ x_i) < 0 < f_+^{''}(min\ y_r) \text{ and } f \text{ is } 3-convex,$$

then (12.7) *holds.*

Theorem 12.6 *Assume* (H_1^*). *If* $x_1,\ldots,x_m \in I^m; y_1,\ldots,y_n \in I^n$ *with* (12.2) *holds and* $f \in \kappa_1^c(I)$ *for some* $c \in [max\ x_i, min\ y_r]$. *Now, if*
(a)

$$f_-^{''}(max\ x_i) \geq 0$$

and

$$\sum_{i=1}^{m} q_i (x_i)^2 - \left(\sum_{i=1}^{m} q_i x_i \right)^2 \leq \sum_{r=1}^{n} p_r (y_r)^2 - \left(\sum_{r=1}^{n} p_r y_r \right)^2. \quad (12.14)$$

or
(b)

$$f_+^{''}(min\ y_r) \leq 0$$

and

$$\sum_{i=1}^{m} q_i (x_i)^2 - \left(\sum_{i=1}^{m} q_i x_i \right)^2 \geq \sum_{r=1}^{n} p_r (y_r)^2 - \left(\sum_{r=1}^{n} p_r y_r \right)^2. \quad (12.15)$$

or
(c)

$$f''_{-}(max\, x_i) < 0 < f''_{+}(min\, y_r) \text{ and } f \text{ is } 3-convex,$$

then (12.9) *holds.*

For the simplification of our next results, we need the following hypothesis:
(H_4^*) Let h and g be $\mu - integrable$ and $\nu - integrable$ functions respectively on X taking values in an interval $I \subset \mathbb{R}$.
(H_5^*) Let f be a convex function on I such that foh is $\mu - integrable$ and fog $\nu - integrable$ on X.

Theorem 12.7 *Assume* $(H_1^*),(H_3),(H_4^*)$ *and* (H_5^*). *If* $x_i \in I \bigcap (-\infty, c]$ *and* $y_r \in I \bigcap [c,\infty)$*for* $i = 1,...,m$ *and* $r = 1,...,n$ *such that*

$$C_{par}(\lambda, id^2, h, \mu, q, \alpha) - \left(\int_X h d\mu\right)^2 = C_{par}(\eta, id^2, g, \nu, p, \beta) - \left(\int_X g d\nu\right)^2 \tag{12.16}$$

and also there exist $c \in I^0$ *such that* (12.2) *holds. If* $f \in \kappa_1^c(I)$, *then following inequality*

$$C_{par}(\lambda, f, h, \mu, q, \alpha) - f\left(\int_X h d\mu\right) \leq C_{par}(\eta, f, g, \nu, p, \beta) - f\left(\int_X g d\nu\right) \tag{12.17}$$

holds.

Theorem 12.8 *Assume* $(H_1^*),(H_3),(H_4^*)$ *and* (H_5^*). *If* $x_i \in I \bigcap (-\infty, c]$ *and* $y_r \in I \bigcap [c,\infty)$*for* $i = 1,...,m$ *and* $r = 1,...,n$ *with*

$$C_{int}(id^2, h, \mu, q, \alpha) - \left(\int_X h d\mu\right)^2 = C_{int}(id^2, g, \nu, p, \beta) - \left(\int_X g d\nu\right)^2 \tag{12.18}$$

such that (12.2)*holds. Now if* $f \in \kappa_1^c(I)$, *then following inequality*

$$C_{int}(f, h, \mu, q, \alpha) - f\left(\int_X h d\mu\right) \leq C_{int}(f, g, \nu, p, \beta) - f\left(\int_X g d\nu\right) \tag{12.19}$$

.

Theorem 12.9 *Assume* $(H_3^*),(H_4^*)$ *and* (H_5^*). *If* $x \in I \bigcap (-\infty, c]$ *and* $y \in I \bigcap [c,\infty)$ *with*

$$\int_X h^2 d\mu - \left(\int_X h d\mu\right)^2 = \int_X g^2 d\nu - \left(\int_X g d\nu\right)^2 \tag{12.20}$$

such that (12.2)*holds. Now if* $f \in \kappa_1^c(I)$, *then following inequality*

$$\int_X foh d\mu - f\left(\int_X h d\mu\right) \leq \int_X fog d\nu - f\left(\int_X g d\nu\right) \tag{12.21}$$

.

Theorem 12.10 *Assume* $(H_1^*),(H_3),(H_4^*)$ *and* (H_5^*). *If* $x_i \in I\bigcap(-\infty,c]$ *and* $y_r \in I\bigcap[c,\infty)$*for* $i=1,...,m$ *and* $r=1,...,n$ *with*

$$\int_X h^2 d\mu - C_{par}(\lambda, id^2, h, \mu, q, \alpha) = \int_X g^2 d\nu - C_{par}(\eta, id^2, g, \nu, p, \beta) \tag{12.22}$$

such that (12.2)*holds. Now if* $f \in \kappa_1^c(I)$, *then following inequality*

$$\int_X f o h d\mu - C_{par}(\lambda, f, h, \mu, q, \alpha) \le \int_X f o g d\nu - C_{par}(\eta, f, g, \nu, p, \beta) \tag{12.23}$$

.

Theorem 12.11 *Assume* $(H_1^*),(H_3),(H_4^*)$ *and* (H_5^*). *If* $x_i \in I\bigcap(-\infty,c]$ *and* $y_r \in I\bigcap[c,\infty)$*for* $i=1,...,m$ *and* $r=1,...,n$ *with*

$$C_{int}(id^2, h, \mu, q, \alpha) - C_{par}(\lambda, id^2, h, \mu, q, \alpha) = C_{int}(\eta, id^2, g, \nu, p, \beta) - C_{par}(\eta, id^2, g, \nu, p, \beta) \tag{12.24}$$

such that (12.2)*holds. Now if* $f \in \kappa_1^c(I)$, *then following inequality*

$$C_{int}(f, h, \mu, q, \alpha) - C_{par}(\lambda, f, h, \mu, q, \alpha) \le C_{int}(f, g, \nu, p, \beta) - C_{par}(\eta, id^2, g, \nu, p, \beta) \tag{12.25}$$

.

Theorem 12.12 *Assume* $(H_1^*),(H_3),(H_4^*)$ *and* (H_5^*). *If* $x_i \in I\bigcap(-\infty,c]$ *and* $y_r \in I\bigcap[c,\infty)$*for* $i=1,...,m$ *and* $r=1,...,n$ *with*

$$\int_X h^2 d\mu - C_{int}(id^2, h, \mu, q, \alpha) = \int_X g^2 d\nu - C_{int}(id^2, g, \nu, p, \beta) \tag{12.26}$$

such that (12.2)*holds. Now if* $f \in \kappa_1^c(I)$, *then following inequality*

$$\int_X f o h d\mu - C_{int}(f, h, \mu, q, \alpha) \le \int_X f o g d\nu - C_{int}(f, g, \nu, p, \beta) \tag{12.27}$$

12.2 Applications In Information Theory

In fields like probability theory, mathematical statistics and information theory, measures of dissimilarity between probability distributions play a pivotal role. Various divergence measures have been introduced for this purpose. For instance, the f-divergence, some particular cases of which are Kullback-Leibler divergence, Jensen-Shannon divergence, etc. Entropies are used to quantify the uncertainty, diversity and randomness of a system. The idea is frequently used in several scientific disciplines.

In the current section we will work in discrete space, i.e, with discrete probability distributions. We first introduce some important definitions and results used for rest of this section.

On the basis of divergence functionals 3.1 and 3.2, we propose a new cyclic divergence functional for 3−convex functions as:

Definition 12.2 *Let $2 \leq k \leq m$ be integers such that $\lambda_1, ..., \lambda_k$ represent positive probability distributions and $f : I \to \mathbb{R}$ be a 3-convex function with I being an interval in $\mathbb{R}$. Also Let $\mathbf{p} := (p_1, \ldots, p_m) \in \mathbb{R}^m$ and $\mathbf{q} := (q_1, \ldots, q_m) \in]0, \infty[^m$ such that*

$$\frac{p_i}{q_i} \in I, \quad i = 1, \ldots, m,$$

$$\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \in I, \quad i + j (\mod m)$$

Then let

$$\tilde{I}_f(\mathbf{p}, \mathbf{q}; \lambda_{i+j}) = \sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j} \right) f \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) \tag{12.28}$$

It is interesting to see that, if we choose $j = 0$, then

$$\tilde{I}_f(\mathbf{p}, \mathbf{q}; \lambda_{\mathbf{i+j}}) = \tilde{I}_f(\mathbf{p}, \mathbf{q})$$

for f to be 3-convex function.

In the examples below we obtain, for suitable choices of the 3-convex function f, some of the best known distance functions used in mathematical statistics, information theory and signal processing between the positive probability distributions for our cyclic divergence functional.

Example 12.1 ***Cyclic-Kullback-Leibler Divergence***
Choosing $f(t) = -t \ln t$, then f is 3-convex. So

$$\begin{aligned}
\tilde{I}_{-t\ln t}(\mathbf{p}, \mathbf{q}; \lambda_{\mathbf{i+j}}) &= -\sum_{i=1}^{m} \left(\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j} \right) \ln \left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right) \\
&= -D(\mathbf{p} \parallel \mathbf{q}; \lambda_{\mathbf{i+j}}).
\end{aligned}$$

Moreover for $j = 0$, we get

$$\tilde{I}_{-t\ln t}(\mathbf{p}, \mathbf{q}) = -D(\mathbf{p} \parallel \mathbf{q}).$$

Example 12.2 ***Cyclic-Hellinger-Distance***
Choosing $f(t) = -\frac{1}{2}(1 - \sqrt{t})^2$, then f is 3-convex. So

$$\tilde{I}_{-\frac{1}{2}(1-\sqrt{t})^2}(\mathbf{p}, \mathbf{q}; \lambda_{\mathbf{i+j}}) = -\frac{1}{2} \sum_{i=1}^{m} \left(\sqrt{\sum_{j=0}^{k-1} \lambda_{j+1} p_{i+j}} - \sqrt{\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}} \right)^2$$

$$= -H^2(\mathbf{p} \parallel \mathbf{q}; \lambda_{\mathbf{i+j}}).$$

Moreover for $j = 0$, we get Hellinger-Distance [8]

$$\hat{I}_{-\frac{1}{2}(1-\sqrt{t})^2}(\mathbf{p},\mathbf{q}) = -H^2(\mathbf{p},\mathbf{q}) = -\frac{1}{2}\sum_{i=1}^{m}\left(\sqrt{p_i}-\sqrt{q_i}\right)^2.$$

Now we give application of our main results given in Theorem 12.5 and Theorem 12.6 by constructing monotonic divergence functionals for 3-convex functions.

Theorem 12.13 *Under the assumptions of Theorem 12.5 and Theorem 12.6 with* $\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m} \in I^m; \frac{\hat{p}_1}{\hat{q}_1},\ldots,\frac{\hat{p}_m}{\hat{q}_n} \in I^n$ *and*

$$\max_i \frac{p_i}{q_i} \le c \le \min_r \frac{\hat{p}_r}{\hat{q}_r} \tag{12.29}$$

for some $c \in [\max \frac{p_i}{q_i}, \min \frac{\hat{p}_r}{\hat{q}_r}]$.
(a) If $f''_{-}(\max \frac{p_i}{q_i}) < 0 < f''_{+}(\min \frac{\hat{p}_r}{\hat{q}_r})$ *and* f *is 3-convex, then*

$$f\left(\sum_{r=1}^{n}\hat{p}_r\right) - f\left(\sum_{i=1}^{m}p_i\right) \le \tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}}) - \tilde{I}_f(\mathbf{p},\mathbf{q}). \tag{12.30}$$

Moreover, if $\mathbf{p}$ *and* $\hat{\mathbf{p}}$ *are probability distributions then we get*

$$\tilde{I}_f(\mathbf{p},\mathbf{q}) \le \tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}}). \tag{12.31}$$

(b) If $f''_{-}(\max \frac{p_i}{q_i}) < 0 < f''_{+}(\min \frac{\hat{p}_r}{\hat{q}_r})$ *and* f *is 3-convex, then*

$$\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}\right) f\left(\frac{\sum_{s=0}^{l-1}\omega_{s+1}\hat{p}_{r+s}}{\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}}\right)$$
$$-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) f\left(\frac{\sum_{j=0}^{k-1}\lambda_{j+1}p_{i+j}}{\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}}\right) \le \tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}}) - \tilde{I}_f(\mathbf{p},\mathbf{q}). \tag{12.32}$$

OR

$$\tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}};\omega) - \tilde{I}_f(\mathbf{p},\mathbf{q};\lambda) \le \tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}}) - \tilde{I}_f(\mathbf{p},\mathbf{q}). \tag{12.33}$$

Moreover, if

$$0 \le \tilde{I}_f(\hat{\mathbf{p}},\hat{\mathbf{q}};\omega) - \tilde{I}_f(\mathbf{p},\mathbf{q};\lambda) \tag{12.34}$$

then we get (12.31).
Proof. *(a) Employing Theorem 12.6 by substituting* $x_i = \frac{p_i}{q_i}$, $p_r \to \hat{q}_r$ *and* $y_r = \frac{\hat{p}_r}{\hat{q}_r}$ *in* (12.9), *we get*

$$\sum_{i=1}^{m} q_i f\left(\frac{p_i}{q_i}\right) - f\left(\sum_{i=1}^{m}p_i\right) \le \sum_{r=1}^{n}\hat{q}_r f\left(\frac{\hat{p}_r}{\hat{q}_r}\right) - f\left(\sum_{r=1}^{n}\hat{p}_r\right) \tag{12.35}$$

thus (12.30) *holds. Using the fact that* $\sum_{i=1}^{m} p_i = 1 = \sum_{r=1}^{n} \hat{p}_r$, *we get* (12.31) *immediately.*
(b) *Using Theorem 12.5 by substituting* $x_i = \frac{p_i}{q_i}$, $p_r \to \hat{q}_r$, $p_{r+s} \to \hat{q}_{r+s}$ *and* $y_r = \frac{\hat{p}_r}{\hat{q}_r}$ *in* (12.7)*, we get* (12.33)*. Immediately using* (12.34) *gives* (12.31)*.*

□

Since we have obtained monotonic divergence functionals for 3-convex functions. It enabled us to investigate monotonicity of renowned distance functions used in mathematical statistics, information theory and signal processing.
We now present two significant applications of the results given in Theorem 12.13 as follows:

Corollary 12.1 *Consider the assumptions of Theorem 12.13.*
(a) *If* $\hat{\mathbf{q}} := (\hat{q}_1,\dots,\hat{q}_n) \in]0,\infty[^n$ *and* $\mathbf{q} := (q_1,\dots,q_m) \in]0,\infty[^m$, *then we get*

$$\ln\left(\frac{n}{m}\right) \leq -\sum_{r=1}^{n} \hat{q}_r \ln \hat{q}_r + \sum_{i=1}^{m} q_i \ln q_i. \tag{12.36}$$

(b) *If* $\hat{\mathbf{q}}, \mathbf{q}$ *are positive probability distributions with* $\frac{n}{m} \geq 1$, *then we get monotonic Shannon entropies*

$$H(\mathbf{q}) \leq H(\hat{\mathbf{q}}). \tag{12.37}$$

(c) *Let* $2 \leq k \leq m$, $2 \leq l \leq n$ *be integers such that* $\lambda_1,\dots,\lambda_k$ *and* $\omega_1,\dots,\omega_l$ *represent positive probability distributions. If* $\hat{\mathbf{q}} := (\hat{q}_1,\dots,\hat{q}_n) \in]0,\infty[^n$ *and* $\mathbf{q} := (q_1,\dots,q_m) \in]0,\infty[^m$, *then we get*

$$\begin{aligned} -\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}\right)\ln\left(\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}\right) + \sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right) \\ \leq -\sum_{r=1}^{n}\hat{q}_r\ln\hat{q}_r + \sum_{i=1}^{m} q_i \ln q_i. \end{aligned} \tag{12.38}$$

(d) *If* $\hat{\mathbf{q}}, \mathbf{q}$ *are positive probability distributions, then we get difference inequality for Shannon entropies*

$$H(\hat{\mathbf{q}}, \omega_{r+s}) - H(\mathbf{q}, \lambda_{i+j}) \leq H(\hat{\mathbf{q}}) - H(\mathbf{q}) \tag{12.39}$$

where

$$H(\hat{\mathbf{q}}, \omega_{r+s}) = -\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}\right)\ln\left(\sum_{s=0}^{l-1}\omega_{s+1}\hat{q}_{r+s}\right)$$

$$H(\mathbf{q}, \lambda_{i+j}) = -\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)\ln\left(\sum_{j=0}^{k-1}\lambda_{j+1}q_{i+j}\right)$$

are cyclic Shannon entropies.

Proof. (a) It follows from Theorem 12.13 (a) by choosing $f(x) = \log x$ and $\mathbf{p} = \hat{\mathbf{p}} = (1,1,\dots,1)$.

(b) It is a special case of (a).
(c) It follows from Theorem 12.13 (b) by choosing $f(x) = \log x$ and $\mathbf{p} = \hat{\mathbf{p}} = (1, 1, \ldots, 1)$.
(d) It is a special case of (c). □

The second application is about famous Kullback-Leibler divergence:

Corollary 12.2 *Consider the assumptions of Theorem 12.13.*
(a) If $\hat{\mathbf{p}} := (\hat{p}_1, ..., \hat{p}_n)$, $\hat{\mathbf{q}} := (\hat{q}_1, ..., \hat{q}_n) \in]0, \infty[^n$ and $\mathbf{p} := (p_1, ..., p_m)$, $\mathbf{q} := (q_1, ..., q_m) \in]0, \infty[^m$, then we get

$$\ln\left(\frac{\sum\limits_{r=1}^{n} \hat{p}_r}{\sum\limits_{i=1}^{m} p_i}\right) \leq \sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) - \sum_{r=1}^{n} \hat{q}_r \ln\left(\frac{\hat{q}_r}{\hat{p}_r}\right). \tag{12.40}$$

(b) If $\hat{\mathbf{p}}, \mathbf{p}, \hat{\mathbf{q}}, \mathbf{q}$ are positive probability distributions, then we get monotonic relative entropies or Kullback-Leibler divergences.

$$D(\hat{\mathbf{q}} \parallel \hat{\mathbf{p}}) \leq D(\mathbf{q} \parallel \mathbf{p}). \tag{12.41}$$

(c) Let $2 \leq k \leq m$, $2 \leq l \leq n$ be integers such that $\lambda_1, ..., \lambda_k$ and $\omega_1, ..., \omega_l$ represent positive probability distributions. If $\hat{\mathbf{p}} := (\hat{p}_1, ..., \hat{p}_n)$, $\hat{\mathbf{q}} := (\hat{q}_1, ..., \hat{q}_n) \in]0, \infty[^n$ and $\mathbf{p} := (p_1, ..., p_m)$, $\mathbf{q} := (q_1, ..., q_m) \in]0, \infty[^m$, then we get

$$\begin{aligned} &\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1} \lambda_{j+1} q_{i+j}\right) \ln\left(\frac{\sum\limits_{j=0}^{k-1} \lambda_{j+1} q_{i+j}}{\sum\limits_{j=0}^{k-1} \lambda_{j+1} p_{i+j}}\right) - \sum_{r=1}^{n}\left(\sum_{s=0}^{l-1} \omega_{s+1} \hat{q}_{r+s}\right) \ln\left(\frac{\sum\limits_{s=0}^{l-1} \omega_{s+1} \hat{q}_{r+s}}{\sum\limits_{s=0}^{l-1} \omega_{s+1} \hat{p}_{r+s}}\right) \\ &\leq \sum_{i=1}^{m} q_i \ln\left(\frac{q_i}{p_i}\right) - \sum_{r=1}^{n} \hat{q}_r \ln\left(\frac{\hat{q}_r}{\hat{p}_r}\right). \end{aligned} \tag{12.42}$$

(d) If $\hat{\mathbf{p}}, \mathbf{p}, \hat{\mathbf{q}}, \mathbf{q}$ are positive probability distributions, then we get difference inequality for Kullback-Leibler divergences

$$D(\mathbf{q} \parallel \mathbf{p}; \lambda_{\mathbf{i+j}}) - D(\hat{\mathbf{q}} \parallel \hat{\mathbf{p}}; \omega_{\mathbf{r+s}}) \leq D(\mathbf{q} \parallel \mathbf{p}) - D(\hat{\mathbf{q}} \parallel \hat{\mathbf{p}}) \tag{12.43}$$

where $D(\mathbf{q} \parallel \mathbf{p}; \lambda_{\mathbf{i+j}}), D(\hat{\mathbf{q}} \parallel \hat{\mathbf{p}}; \omega_{\mathbf{r+s}})$ are cyclic Kullback-Leibler divergences defined in Example 12.1.

Proof. (a) It follows from Theorem 12.13 (a) by choosing $f(x) = \log x$.
(b) It is a special case of (a).
(c) It follows from Theorem 12.13 (b) by choosing $f(x) = \log x$.
(d) It is a special case of (c). □

Remark 12.2 It is interesting to note that if we choose the 3-convex function $f(x) = -x\log(x)$ in Corollary 12.2 (b), then we get monotonicity as

$$D(\hat{\mathbf{p}} \parallel \hat{\mathbf{q}}) \leq D(\mathbf{p} \parallel \mathbf{q}). \tag{12.44}$$

By similar substitution in Corollary 12.2 (d), we get

$$D(\mathbf{p} \parallel \mathbf{q}; \lambda_{\mathbf{i+j}}) - D(\hat{\mathbf{p}} \parallel \hat{\mathbf{q}}; \omega_{\mathbf{r+s}}) \leq D(\mathbf{p} \parallel \mathbf{q}) - D(\hat{\mathbf{p}} \parallel \hat{\mathbf{q}}). \tag{12.45}$$

12.2.1 Monotonic inequalities Via Zipf-Mandelbrot law

Zipf's law is one of the basic laws in information science and is extensively applied in linguistics. For the rest of the section, let $m \in \{1,2,\ldots\}$, $c \geq 0$, $d > 0$, then **Zipf-Mandelbrot entropy** can be given as:

$$Z^m(H,c,d) = \frac{d}{H^m_{c,d}} \sum_{i=1}^{m} \frac{\ln(i+c)}{(i+c)^d} + \ln(H^m_{c,d}) \tag{12.46}$$

where

$$H^m_{c,d} = \sum_{u=1}^{m} \frac{1}{(u+c)^d}.$$

In the similar passion for $n \in \{1,2,\ldots\}$, $\hat{c} \geq 0$, $\hat{d} > 0$ one can define

$$Z^n(H,\hat{c},\hat{d}) = \frac{\hat{d}}{H^n_{\hat{c},\hat{d}}} \sum_{r=1}^{n} \frac{\ln(r+\hat{c})}{(r+\hat{c})^d} + \ln(H^n_{\hat{c},\hat{d}}) \tag{12.47}$$

where

$$H^n_{\hat{c},\hat{d}} = \sum_{v=1}^{n} \frac{1}{(v+\hat{c})^{\hat{d}}}.$$

Consider

$$q_i = f(i;m,c,d) = \frac{1}{(i+c)^d H^m_{c,d}} \tag{12.48}$$

$$\hat{q}_r = f(r;n,\hat{c},\hat{d}) = \frac{1}{(r+\hat{c})^{\hat{d}} H^n_{\hat{c},\hat{d}}} \tag{12.49}$$

where $\hat{q}_r, q_u$ are discrete probability distributions known as **Zipf-Mandelbrot** law. Application of Zipf-Mandelbrot law can be found in linguistics, information sciences and also is often applicable in ecological field studies. Some of the recent study regarding Zipf-Mandelbrot law can be seen in the listed references (see [48, 52, 53, 38]). Now we state our results involving entropy introduced by Mandelbrot Law by establishing the relationship with Shannon and Relative entropies:

Corollary 12.3 *Let* $\hat{\mathbf{q}}$, $\mathbf{q}$ *be as defined in* (12.48), (12.49) *by Zipf-Mandelbrot law with parameters* $m,n \in \{1,2,\ldots\}$, $\hat{c},c \geq 0$, $\hat{d},d > 0$. *Then, the following holds.*
(*a*) *If* $\frac{n}{m} \geq 1$, *then we get monotonic Zipf-Mandelbrot entropies*

$$H(\mathbf{q}) = Z^m(H,c,d) \leq Z^n(H,\hat{c},\hat{d}) = H(\hat{\mathbf{q}}). \tag{12.50}$$

(*b*) *Let* $2 \leq k \leq m$, $2 \leq l \leq n$ *be integers such that* $\lambda_1,\ldots,\lambda_k$ *and* $\omega_1,\ldots,\omega_l$ *represent positive probability distributions, then*

$$Z^n\left(H,\hat{c},\hat{d},\omega_{r+s}\right) - Z^m\left(H,c,d,\lambda_{i+j}\right) \leq Z^n(H,\hat{c},\hat{d}) - Z^m(H,c,d) \tag{12.51}$$

where

$$Z^n\left(H,\hat{c},\hat{d},\omega_{r+s}\right) = -\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1} \frac{\omega_{s+1}}{\left((r+s+\hat{c})^{\hat{d}} H^n_{\hat{c},\hat{d}}\right)}\right) \ln\left(\frac{1}{H^n_{\hat{c},\hat{d}}} \sum_{s=0}^{l-1} \frac{\omega_{s+1}}{\left((r+s+\hat{c})^{\hat{d}}\right)}\right)$$

$$Z^m\left(H,c,d,\lambda_{i+j}\right)=-\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{\left((i+j+c)^dH^m_{c,d}\right)}\right)\ln\left(\frac{1}{H^m_{c,d}}\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{((i+j+c)^d)}\right).$$

Proof. (a) One can see that for $q_i,\hat{q}_r$ be as defined in (12.48) and (12.49), $\sum\limits_{i=1}^{m} q_i = 1 = \sum\limits_{r=1}^{n}\hat{q}_r$. Therefore, using above $q_i,\hat{q}_r$ in Shannon entropy (3.3), we get Mandelbrot entropies (12.46) and (12.47) respectively. Consequently, by applying Corollary 12.1 (b), we get (12.50).
(b) By employing Corollary 12.1 (d) we get (12.51). □

Finally we will establish the nice connection of Relative entropy with Mandelbrot entropy:

Corollary 12.4 *Let* $\hat{\mathbf{q}}_1$, $\mathbf{q}_1$, $\hat{\mathbf{p}}_2$, $\mathbf{p}_2$ *be Zipf-Mandelbrot law with parameters* $m,n\in\{1,2,\ldots\}$, $\hat{c}_1,c_1,\hat{c}_2,c_2\geq 0$ *and* $\hat{d}_1,d_1,\hat{d}_2,d_2>0$. *Then, the following holds.*
(a) *Employing Corollary 12.2* (b), *we get*

$$\begin{aligned}
&-Z^n(H,\hat{c}_1,\hat{d}_1)+\frac{\hat{d}_2}{H^n_{\hat{c}_1,\hat{d}_1}}\sum_{r=1}^{n}\frac{\ln(r+\hat{c}_2)}{(r+\hat{c}_1)^{\hat{d}_1}}+\ln\left(H^n_{\hat{c}_2,\hat{d}_2}\right) \qquad (12.52)\\
&=\sum_{r=1}^{n}\frac{1}{(r+\hat{c}_1)^{\hat{d}_1}H^n_{\hat{c}_1,\hat{d}_1}}\ln\left(\frac{(r+\hat{c}_2)^{\hat{d}_2}H^n_{\hat{c}_2,\hat{d}_2}}{(r+\hat{c}_1)^{\hat{d}_1}H^n_{\hat{c}_1,\hat{d}_1}}\right)\\
&=D(\hat{\mathbf{q}}\parallel\hat{\mathbf{p}})\leq D(\mathbf{q}\parallel\mathbf{p})\\
&=\sum_{i=1}^{m}\frac{1}{(i+c_1)^{d_1}H^m_{c_1,d_1}}\ln\left(\frac{(i+c_2)^{d_2}H^m_{c_2,d_2}}{(i+c_1)^{d_1}H^m_{c_1,d_1}}\right)\\
&=-Z^m(H,c_1,d_1)+\frac{d_2}{H^m_{c_1,d_1}}\sum_{i=1}^{m}\frac{\ln(i+c_2)}{(i+c_1)^{d_1}}+\ln\left(H^m_{c_2,d_2}\right).
\end{aligned}$$

(b) *Let* $2\leq k\leq m$, $2\leq l\leq n$ *be integers such that* $\lambda_1,\ldots,\lambda_k$ *and* $\omega_1,\ldots,\omega_l$ *represent positive probability distributions, then by Corollary 12.2* (d), *we get*

$$\begin{aligned}
&\sum_{i=1}^{m}\left(\sum_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}\right)\ln\left(\frac{\sum\limits_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_1)^{d_1}H^m_{c_1,d_1}}}{\sum\limits_{j=0}^{k-1}\frac{\lambda_{j+1}}{(i+j+c_2)^{d_2}H^m_{c_2,d_2}}}\right)\\
&\quad-\sum_{r=1}^{n}\left(\sum_{s=0}^{l-1}\frac{\lambda_{j+1}}{(r+s+\hat{c}_1)^{\hat{d}_1}H^n_{\hat{c}_1,\hat{d}_1}}\right)\ln\left(\frac{\sum\limits_{s=0}^{l-1}\frac{\lambda_{s+1}}{(r+s+\hat{c}_1)^{\hat{d}_1}H^n_{\hat{c}_1,\hat{d}_1}}}{\sum\limits_{s=0}^{l-1}\frac{\omega_{s+1}}{(r+s+\hat{c}_2)^{\hat{d}_2}H^n_{\hat{c}_2,\hat{d}_2}}}\right)\\
&\quad\leq Z^n(H,\hat{c}_1,\hat{d}_1)-Z^m(H,c_1,d_1)\\
&\quad+\frac{d_2}{H^m_{c_1,d_1}}\sum_{i=1}^{m}\frac{\ln(i+c_2)}{(i+c_1)^{d_1}}+\ln\left(H^m_{c_2,d_2}\right)-\frac{\hat{d}_2}{H^n_{\hat{c}_1,\hat{d}_1}}\sum_{r=1}^{n}\frac{\ln(r+\hat{c}_2)}{(r+\hat{c}_1)^{\hat{d}_1}}+\ln\left(H^n_{\hat{c}_2,\hat{d}_2}\right). \qquad (12.53)
\end{aligned}$$

Bibliography

[1] R. P. Agarwal, P. J. Y. Wong, *Error Inequalities in Polynomial Interpolation and their Applications*, Kluwer Academic Publishers, Dordrecht, 1993.

[2] M. S. Ali, D. Silvey, *A general class of coefficients of divergence of one distribution from another*, J. Roy. Statist. Soc., ser. B 28 (1966), 131–140.

[3] A. A. Aljinović, J. Pečarić, A. Vukelić, *On some Ostrowski type inequalities via Montgomery identity and Taylor's formula II*, Tamkang Jour. Math., 36(4) (2005), 279–301.

[4] M. Anwar, S. Hussain, J. Pečarić, *Some inequalities for Csiszár-divergence measures*, Int. Journal of Math. Analysis, 3 (2009), 1295–1304.

[5] M. Anwar, J. Pečarić, *On log-convexity for differences of mixed symmetric means*, Mathematical Notes, Vol. 88, No. 6 (2010), 776–784.

[6] G. Aras-Gazić, V. Culjak, J. Pečarić, A. Vukelić, *Generalization of Jensen's inequality by Lidstone's polynomial and related results*, Math. Inequal. Appl., 16(4) (2013), 1243–1267.

[7] I. A. Baloch and J. Pečarić and M. Praljak,*Generalization of Levinson's inequality*, J. Math. Inequal., 9(2) (2015), 571–586.

[8] R. Beran, *Minimum Hellinger distance estimates for parametric models*, Ann. Statist., 5 (1977), 445–463.

[9] E. Beck, *Über Ungleichungen von der Form* $f\left(M_\phi(x,\alpha),M_\psi(y,\alpha)\right) \geq M_\chi(f(x,y),\alpha)$, Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz., No. 320–328 (1970), 1–14.

[10] P. R. Beesack, *On the Green's function of an N-point boundary value problem*, Pacific J. Math., 12 (1962), 801–812.

[11] M. Bencze, S. Wu, *New inequalities obtained by rearrangement*, Octogon Mathematical Magazine., 21(2) (2013), 593–612.

[12] I. Brnetić, C. E. M. Pearce, J. Pečarić, *Refinements of Jensen's inequality*, Tamkang J. Math., 31 (2000), 63–69.

[13] I. Brnetić, K. A. Khan, J. Pečarić, *Refinement of Jensen's inequality with applications to cyclic mixed symmetric means and Cauchy means*, J. Math. Inequal., 9(4) (2015), 1309–1321.

[14] I. Burbea, C. R. Rao, *On the convexity of some divergence measures based on entropy functions*, IEEE Transactions on Information Theory., 28 (1982), 489–495.

[15] P. S. Bullen, *Handbook of means and their inequalities*, Kluwer Academic Publisher, 2003, pp. 537.

[16] P. S. Bullen, D. S. Mitrinović, P. M. Vasić, *Means and their inequalities*, Kluwer Academic Publisher, 1988, pp. 459.

[17] A. Chao, L. Jost, T. C. Hsieh, K. H. Ma, W. B. Sherwin, et al., *Expected Shannon Entropy and Shannon Differentiation between Subpopulations for Neutral Genes under the Finite Island Model*, PLOS ONE, 10(6) (2015), 1–24.

[18] I. Csiszár, *Information-type measures of difference of probability distributions and indirect observations*, Studia Sci. Math. Hungar., 2 (1967), 299–318.

[19] I. Csiszár , *Information measures: A critical survey*, Trans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Random Processes and 8th European Meeting of Statist., Volume B, pp. 73–86, Academia Prague, 1978.

[20] P. J. Davis,*Interpolation and Approxiamtion*, Blaisdell Publishing Co., Boston, 1961.

[21] A. Dembo, T. M. Cover, J. A. Thomas, *Information theoretic inequalities*, IEEE Trans. Inf. Theory, 37 (1991), 1501–1518.

[22] V. Diodato, *Dictionary of Bibliometrics*, Haworth Press, New York, 1994.

[23] S. S. Dragomir, *A new refinement of Jensen's inequality in linear spaces with applications*, Math. Comput. Modelling, 2 (2010), 1497–1505.

[24] S. S. Dragomir, *A refinement of Jensen's inequality with applications for f-divergence measures*, Taiwanese J. Math., 14 (2010), 153–164.

[25] S. S. Dragomir (Ed.), *Inequalities for Csiszár f-Divergence in Information Theory*, RGMIA Monographs, Victoria University, 2000.
(online: http://ajmaa.org/RGMIA/monographs.php/)

[26] L. Egghe, R. Rousseau, *Introduction to Informetrics. Quantitative Methods in Library*, Documentation and Information Science, Elsevier Science Publishers, New York, 1990.

[27] A. M. Fink, *Bounds of the deviation of a function from its avereges*, Czechoslovak Math. J., 42(117) (1992), 289–310.

[28] T. M. Flett, *Differential Analysis*, Cambridge University Press, 1980.

[29] S. Frontier, *Diversity and structure in aquatic ecosystems*, Mar. Biol. Ann. Rev., 23 (1985), 253–312.

[30] B. R. Frieden, *Image enhancement and restoration*, Picture Processing and Digital Filtering (T. S. Huang, Editor), Springer-Verlag, Berlin, 1975.

[31] T. Furuta, J. Mićić, J. Pečarić, Y. Seo, *Mond-Pečarić, Method in Operator Inequalities, Inequalities for Bounded Selfadjoint Operators on a Hilbert Space*, Element, Zagreb, 2005.

[32] A. L. Gibbs , F. E. Su, *On choosing and bounding probability metrics*, Int. Statist. Rev., 70 (2002), 419–435.

[33] V. L. Gontscharoff, *Theory of Interpolation and Approxiamtion of Functions*, Gostekhizdat Moscow, 1954.

[34] E. Hewitt, K. R. Stromberg, *Real and Abstract Analysis*, Graduate Text in Mathematics 25, Springer-Verlag, Berlin-Heidelberg-New York, 1965.

[35] L. Horváth, K. A. Khan, J. Pečarić, *Further refinement of results about mixed symmetric means and Cauchy means*, Advances in Inequalities and Applications, 1 (2012), 12–32.

[36] L. Horváth, K. A. Khan, J. Pečarić, *Combinatorial Improvements of Jensen's Inequality, Monographs in Inequalities 8*, Element, Zagreb, 2014.

[37] L. Horváth , K. A. Khan, J. Pečarić, *Cyclic refinements of the discrete and integral form of Jensen's inequality with applications*, Analysis, Munich, 36(4) (2016), 253–263.

[38] L. Horváth, Đ. Pečarić, J. Pečarić, *Estimations of f - and Rényi Divergences by Using a Cyclic Refinement of the Jensen's Inequality*, Bull. Malays. Math. Sci. Soc. (2017). https://doi.org/10.1007/s40840-017-0526-4.

[39] L. Horváth, *Inequalities corresponding to the classical Jensen's inequality*, J. Math. Inequal., 3 (2009), 189–200.

[40] L. Horváth, *Inequalities corresponding to the classical Jensen's inequality*, J. Math. Inequal., 3(2) (2009), 189–200.

[41] L. Horváth, *A method to refine the discrete Jensen's inequality for convex and mid-convex functions*, Math. Comput. Mod., 54 (2011) 2451–2459.

[42] L. Horváth, *A parameter dependent refinement of the discrete Jensen's inequality for convex and mid-convex functions*, J. Ineq. and Appl., 2011 (2011), 14 pages.

[43] L. Horváth,J. Pečarić, *A refinement of the discrete Jensen's inequality*, Math. Ineq. Appl., 14(4) (2011), 777–791.

[44] L. Horváth, K. A. Khan,J. Pečarić, *Refinements of Hölder and Minkowski inequalities with weights*, Proc. A. Razmadze Math. Inst., 158 (2012), 33–56.

[45] L. Horváth, K. A. Khan, J. Pečarić, *New Refinements of Hölder and Minkowski inequalities with weights*, Proc. A. Razmadze Math. Inst., 161 (2013), 97–116.

[46] L. Horváth, K. A. Khan, J. Pečarić, *Cyclic refinements of the different versions of operator Jensen's inequality*, Electron. J. Linear Algebra, 31 (2016), 125–133.

[47] J. Jakšetic, J. Pečarić, *Exponential Convexity Method*, Journal of Convex Analysis, 20(1) (2013), 181–197.

[48] J. Jakšetic, Đ. Pečarić, J. Pečarić, *Some properties of Zipf-Mandelbrot law and Hurwitz ζ-function*, Math. Inequal. Appl., 21(2) (2018), 575–584.

[49] J. L. W. V. Jensen, *Sur les fonctions convexes et les iné galités entre les valeurs moyennes*, Acta Math., 30 (1906), 175–193.

[50] J. H. Justice, (editor), *Maximum Entropy and Bayssian Methods in Applied Statistics*, Cambridge University Press, Cambridge, 1986.

[51] J. N. Kapur, *On the roles of maximum entropy and minimum discrimination information principles in Statistics*, Technical Address of the 38th Annual Conference of the Indian Society of Agricultural Statistics, (1984), 1–44.

[52] M. A. Khan, Đ. Pečarić, J. Pečarić, *Bounds for Shannon and Zipf-mandelbrot entropies*, Mathematical Method in Applied Sciences, 40(18) (2017), 7316–7322.

[53] M. A. Khan, Đ. Pečarić, J. Pečarić, *On Zipf-Mandelbrot entropy*, Computational and Applied Mathematics, 346 (2019), 192–204.

[54] K. A. Khan, J. Pečarić, *Cyclic Refinement of Beck's inequalities*, Proc. A. Razmadze Math. Inst., 169 (2015), 83–92.

[55] M. K. Khan, M. Hanif, *On Some Refinements of Jensen's Inequality*, J. Approx. Theory, 93 (1998), 313–330.

[56] M. Khosravi, J. S. Aujla, S. S. Dragomir, M. S. Moslehian, *Refinements of Choi-Davis-Jensen's inequality*, Bull. Math. Anal. Appl, 3 (2011), 127–133.

[57] M. Kian, M. S. Moslehian, *Refinements of operator Jensen-Mercer inequality*, Electron. J. Linear Algebra, 26 (2013), 742–753.

[58] S. Kullback, *Information Theory and Statistics*, J. Wiley, New York, 1959.

[59] S. Kullback, R. A. Leibler, *On information and sufficiency*, Annals Math. Statist., 22 (1951), 79–86.

[60] R. M. Leahy, C. E. Goutis, *An optimal technique for constraint-based image restoration and mensuration*, IEEE Trans. on Acoustics, Speech and Signal Processing, 34 (1986), 16920–1642.

[61] A. Lesne, *Shannon entropy: a rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics*, Mathematical Structures in Computer Science, 24(3) (2014), 63 pages. doi:10.1017/S0960129512000783.

[62] A. Yu. Levin, Some problems bearing on the oscillation of solutions of linear differential equations, Soviet Math. Dokl., 4 (1963), 121–124.

[63] F. Liese, I. Vajda, *Convex Statistical Distances (Teubner-Texte Zur Mathematik)*, vol. 95. Leipzig, Germany: Teubner, 1987.

[64] F. Liese, I. Vajda, *On divergences and informations in statistics and information theory*, IEEE Trans. Inf. Theory, 52 (2006), 4394–4412.

[65] B. Mandelbrot, *An informational theory of the statistical structure of language*, In Jackson, W., editor, Communication Theory, pp. 486–502. Academic Press, New York, 1953.

[66] C. A. McCarthy cp. *Israel J. Math.*, 5 (1967), 249–271.

[67] N. Mehmood , R. P. Agarwal , S. I. Butt, J. Pečarić , *New generalizations of Popoviciu-type inequalities via new Green's functions and Montgomery identity*, J. Inequal. Appl., (2017), 2017:108.

[68] N. Mehmood , S. I. Butt, L. Horváth, J. Pečarić, *Several new cyclic Jensen type inequalities and their applications*, J. Inequal. Appl., (2019), 2019:240.

[69] N. Mehmood , S. I. Butt, J. Pečarić, *Generalization of cyclic refinements of Jensen inequality by montgomery identity and Green's function*, Asian-European Journal of Mathematics, 11(4) (2018), 2018:1850060.

[70] N. Mehmood , S. I. Butt, Đ. Pečarić, J. Pečarić, *Generalizations of Cyclic Refinements of Jensen's Inequality by Lidstone's Polynomial With Applications in Information Theory*, J. Math. Inequal., 14(1) (2020), 249–271.

[71] N. Mehmood , S. I. Butt, Đ. Pečarić, J. Pečarić, *New Bounds For Shannon, Relative And Mandelbrot Entropies Via Abel-Gontscharoff Interpolating Polynomial*, Math. Inequal. Appl., 22(4) (2019), 1283–1301.

[72] N. Mehmood , S. I. Butt, Đ. Pečarić, J. Pečarić, *New bounds for Shannon, Relative and Mandelbrot entropies via Hermite interpolating polynomial*, Demonstratio Mathematica, 51(1) (2018), 112–130.

[73] J. Mićić, J. Pečarić, *Order among power means of positive operators*, Sci. Math. Jpn., e-2009, 677–693.

[74] D. S. Mitrinović, J. E. Pečarić, A. M. Fink, *Inequalities for functions and their integrals and derivatives*, Kluwer Academic Publishers, Dordrecht, 1994.

[75] D. S. Mitrinović, J. Pečarić, *Unified treatment of some inequalities for mixed means*, Österreìch Akad. Wiss. Math.-Natur. Sitzungsber, 197 (1988), 391–397.

[76] B. Mond, J. Pečarić, *Convex inequalities in Hilbert space*, Houston J. Math., 19 (1993), 405–420.

[77] B. Mond, J. Pečarić, *Converses of Jensen's inequality for several operators*, Rev. Anal. Numér. Théor. Approx., 23 (1994), 179–183.

[78] N. Mehmood , S. I. Butt , G. Pečarić, J. Pečarić, *Generalization of cyclic refinements of Jensens inequality by Finks identity*, J. Inequal. Appl., (2018), 2018:51.

[79] D. Mouillot, A. Lepretre, *Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in*

community diversity, Environmental Monitoring and Assessment, Springer, 2000; 63: 279–295.

[80] M. E. J. Newman, *Power laws, Pareto distributions and Zipf's law*, URL: http://arxiv.org/pdf/cond-mat/0412004.pdf?origin=publication_detail (6.4.2014.)

[81] M. Niezgoda, *Choi-Davis-Jensen's inequality and generalized inverses of linear operators*, Electron. J. Linear Algebra, 26 (2013), 406–416.

[82] J. Pečarić, F. Proschan, Y. L. Tong, *Convex functions, partial orderings and statistical applications*, Academic Press, New York, 1992.

[83] J. Pečarić, J. Perić, *Improvement of the Giaccardi and the Petrović Inequality and Related Stolarsky Type Means*, An. Univ. Craiova Ser. Mat. Inform., 39(1) (2012), 65–75.

[84] J. Pečarić, V. Volenec, *Interpolation of the Jensen inequality with some applications*, Österreìch. Akad. Wiss. Math.-Natur. Kl. Sitzungsber, II 197, 8–10, (1988), 463–467.

[85] A. Rényi, *On measures of information and entropy, Proceedings of the fourth Berkeley Symposium on Mathematics*, Statistics and Probability, (1960), pp. 547–561.

[86] T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[87] W. Rudin, *Functional Analysis*, McGraw-Hill, Inc., 1991.

[88] I. Sason, S. Verdú, *f-Divergence Inequalities*, IEEE Trans. Inf. Theory, 62 (2016), 5973–6006.

[89] J.L. Schiff, *The Laplace transform. Theory and applications*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1999.

[90] C. E. Shannon, *A mathematical theory of communication*, Bull. Sept. Tech. J., 27 (1948), 370–423 and 623–656.

[91] I. Vajda, *Theory of Statistical Inference and Information*, Kluwer, Dordrecht, 1989.

[92] E. Wedrowska, *Application of Kullback-Leibler Relative Entropy for Studies on the Divergence of Household Expenditures Structures*, Olsztyn Economic Journal, 6 (2011), 133–142.

[93] J. B. Wilson, *Methods for fitting dominance/diversity curves*, J. Veget. Sci., 2 (1991), 35–46.

[94] D. V. Widder, *Completely convex function and Lidstone series*, Trans. Am. Math. Soc., 51 (1942), 387–398.

[95] G. K. Zipf, *Selected Studies of the Principle of Relative Frequency in Language Cambridge*, MA: Harvard University Press, 1932.

Index

Zagreb, April 2020