Chapter 1

Introduction

Convex functions have a significant role while dealing with optimization problems. Geometry of convex functions leads to many important inequalities which are frequently used to estimate and compare the values related to many physical problems in different branches of mathematics and physics. Entropies and divergences are widely studied in information theory. While dealing with many physical problems physicist have to deal with the structure involving higher dimension convexity. Therefore there are two important gaps first to estimate the entropies and divergences and second one is to study the inequalities for higher dimension problems. Therefore we estimate different entropies and divergences and secondly we generalize the related results for higher order convex functions.
It is of great interest for researchers to study inequalities of continuous data and arbitrary weights. For example integral version of Popoviciu's inequality are studied in the sequel.

1.1 Some Inequalities Involving Convex Functions

The first chapter contains: introduction to convex functions, various inequalities involving convex functions, refinement of these inequalities given by various researchers in recent years, the weighted version of Popoviciu's inequality, some notions from information theory containing entropies and divergences. These will be used frequently in the following chapters while obtaining main results.

1.1.1 J-Convex Functions

In 1905–1906, J. L. W. V. Jensen began the systematic study of convex functions (see [75, p.3]).

A function $\eta : I \to \mathbb{R}$ is said to be J-convex or mid-convex function or convex in Jensen sense on I if

$$\eta\left(\frac{u+v}{2}\right) \leq \frac{\eta(u)+\eta(v)}{2} \tag{1.1}$$

holds for all $u,v \in I$.

The $\eta(x) = x^2$ and $\eta(x) = |x|$ for all $x \in \mathbb{R}$ are the examples of J-convex functions.

1.1.2 Convex Functions

The notion of convex function is the generalization of J-convex function for the arbitrary weight $t \in [0,1]$. In [87, p. 1] the formal definition is given as follows.
Suppose X is a real vector space, $C \subset X$ is a convex set. A function $\eta : C \to \mathbb{R}$ is said to be convex if

$$\eta(\sigma u + (1-\sigma)v) \leq \sigma\eta(u) + (1-\sigma)\eta(v),$$

holds for all $u,v \in C$ and $\sigma \in [0,1]$.

The $\eta(x) = x^2$, $\eta(x) = |x|$, $-\log x$ and e^x for all $x \in \mathbb{R}$ are the examples of convex functions.

1.1.3 Operator Convex Functions

Let I be an interval of real numbers and $S(I)$ denotes the class of all self-adjoint bounded operators defined on complex Hilbert space H whose spectra are in I. Also, assume that $Sp(A)$ denotes the spectrum of a bounded operator A defined on H. An operator $A \in S(I)$ is said to be strictly positive if it is positive and invertible, or equivalently, $Sp(A) \subset [d_1, d_2]$ for $0 < d_1 < d_2$.

Let $\psi : I \to \mathbb{R}$ be a function defined on the interval I. Then ψ is said to be operator convex if ψ is continuous and

$$\psi(\zeta A_1 + (1-\zeta)A_2) \leq \zeta\psi(A_1) + (1-\zeta)\psi(A_2) \tag{1.2}$$

for all $A_1, A_2 \in S(I)$ and $\zeta \in [0,1]$. If the function $-\psi$ is operator convex on I, then ψ is said to be operator concave. The function ψ is said to be operator monotone on I if ψ is continuous on I and $A_1, A_2 \in S(I)$, $A_1 \leq A_2$ (i.e. $A_2 - A_1$ is positive operator), then $\psi(A_1) \leq \psi(A_2)$.

1.1.4 Discrete Jensen's Inequality

The Jensen inequality in discrete version [87, p. 43] generalizes the notion of convex function. Here the function operates on the convex combination of any finite number of points.

Suppose X is a real vector space, $C \subset X$ is a convex set, let $\psi : C \to \mathbb{R}$ be a convex function, $\zeta_1, \ldots, \zeta_n \in [0,1]$ are such that $\sum_{i=1}^{n} \zeta_i = 1$, and $y_1, \ldots, y_n \in C$, then

$$\psi\left(\sum_{\gamma=1}^{n} \zeta_\gamma y_\gamma\right) \leq \sum_{\gamma=1}^{n} \zeta_\gamma \psi(y_\gamma). \tag{1.3}$$

In the Jensen inequality, it is natural to ask the question that is it possible to relax the condition of nonnegative of ζ_γ $(\gamma = 1,2,\ldots,n)$ at the expense of restricting y_γ $(\gamma = 1,2,\ldots,n)$ more severely. The answer of this question was given by Steffensen [96]:

1.1.5 Discrete Jensen-Steffensen's Inequality

Let $\psi : I \to \mathbb{R}$ be a convex function defined on the interval I. Let $y_\gamma \in I$, $\zeta_\gamma \in \mathbb{R}$ $(\gamma = 1,2,\ldots,n)$ with $\bar{\zeta} = \sum_{\gamma=1}^{n} \zeta_\gamma$. If $y_1 \leq y_2 \leq \cdots \leq y_n$ or $y_1 \geq y_2 \geq \cdots \geq y_n$ and

$$0 \leq \sum_{\gamma=1}^{k} \zeta_\gamma \leq \sum_{\gamma=1}^{n} \zeta_\gamma \;\; k = 1,2,\ldots,n, \quad \sum_{\gamma=1}^{n} \zeta_\gamma > 0, \tag{1.4}$$

then

$$\psi\left(\frac{1}{\bar{\zeta}} \sum_{\gamma=1}^{n} \zeta_\gamma y_\gamma\right) \leq \frac{1}{\bar{\zeta}} \sum_{\gamma=1}^{n} \zeta_\gamma \psi(y_\gamma). \tag{1.5}$$

1.1.6 Integral Form of Jensen's Inequality

The integral form of Jensen's inequality [47] is defined as follows.
Let (X, A, μ) be a probability space, consider an integrable function $h : X \to I$. Also let $\psi : I \to \mathbb{R}$ be a convex function. Then

$$\psi\left(\int_X h\, d\mu\right) \leq \int_X \psi \circ h\, d\mu. \tag{1.6}$$

1.1.7 Integral Version of Jesnen-Steffensen's Inequality

Integral version of Jensen-Steffensen's inequality is given by:
Let I be an interval in $\mathbb{R}$ and $g, h : [a,b] \to \mathbb{R}$ are integrable functions such that $g(\rho) \in I$ for all $\rho \in [a,b]$. Also, assume that $\psi : I \to \mathbb{R}$ is convex function and $h(\psi \circ g)$ is integrable on $[a,b]$. If g is monotonic on $[a,b]$ and h satisfies

$$0 \leq \int_a^\lambda h(\rho)d\rho \leq \int_a^b h(\rho)d\rho, \; \lambda \in [a,b], \quad \int_a^b h(\rho)d\rho > 0, \tag{1.7}$$

then

$$\psi\left(\frac{\int_a^b g(\rho)h(\rho)d\rho}{\int_a^b h(\rho)d\rho}\right) \leq \frac{\int_a^b h(\rho)(\psi \circ g)(\rho)d\rho}{\int_a^b h(\rho)d\rho}. \tag{1.8}$$

This means convex function and Jensen-type inequalities are linked to each other. In fact definition of convex function involves inequality sign. Until now, inequalities have played a major role in convex function development. In mathematics the role of inequalities is very important, specially in approximation theory and analysis. The linear programming is based on inequalities. A number of mathematicians have a keen interest in the study of mathematical inequalities.

Jensen's inequality is the fundamental inequality for convex function. Many classical inequalities (for instance Minkowski's inequality, Hölder's inequality *etc.*) and other inequalities are the consequences of Jensen's inequality.

L. Horváth and J. Pečarić in [49] used a refinement of discrete Jensen's inequality to construct a new refinement of (1.6), which is a generalization of a result given in [25]. They also gave new monotone quasi arithmetic means.

In a last few decades, many researcher papers have appeared in literature concerning the refinement of discrete Jensen's inequality (see [47]). However the refinement of discrete Jensen's inequality has been studied more compared to the refinement of its integral version. The researchers used the refinements of (1.3) to construct new refinements of (1.6). For instance we can see the following results [88].
Suppose that f is a J-convex function on an interval J, $c_j \in J$, $j = 1,\ldots,n$. Then

$$\eta_{r,n} \geq \eta_{r-1,n} \quad r = 1,\ldots,n-1, \tag{1.9}$$

where

$$\eta_{r,n} = \eta_{k,n}(c_1,\ldots,c_n) := \frac{1}{\binom{n}{r}} \sum_{1\leq j_1<\ldots<j_r\leq n} \eta\left(\frac{1}{r}(c_{j_1}+\ldots+c_{j_r})\right).$$

For positive weights the above results are given in [84].
Suppose $\overline{\eta}$ is convex function defined on an interval J, $c_j \in J$ $(j = 1,...,n)$.

$$\overline{\eta}_{r,n}(c_1,\ldots,c_n,\sigma_1,\ldots,\sigma_n) := \frac{1}{\binom{n-1}{r-1}P_n} \sum_{1\leq j_1<\ldots<j_r\leq n} (\sigma_{j_1}+\ldots+\sigma_{j_r})\overline{\eta}\left(\frac{\sigma_{j_1}c_{j_1}+\ldots+\sigma_{j_r}c_{j_r}}{\sigma_{j_1}+\ldots+\sigma_{j_k}}\right)$$

where $(\sigma_1,\ldots,\sigma_n)$ is suppose to be a positive n-tuple with $\sum_{j=1}^{r}\sigma_j = P_r$, then

$$\overline{\eta}_{r,n}(c_1,\ldots,c_n,\sigma_1,\ldots,\sigma_n) \geq \overline{\eta}_{r+1,n}(c_1,\ldots,c_n,\sigma_1,\ldots,\sigma_n) \quad n = 1,\ldots,r-1. \tag{1.10}$$

J. Pečarić and D. Svrtan noted that by considering the expresssion

$$\tilde{\eta}_{r,n} = \frac{1}{\binom{n+r-1}{r-1}P_n} \sum_{1\leq j_1\leq\ldots\leq j_k\leq n} (\sigma_{j_1}+\ldots+\sigma_{j_r})\left(\frac{\sigma_{j_1}c_{j_1}+\ldots+\sigma_{j_r}c_{j_r}}{\sigma_{j_1}+\ldots+\sigma_{j_r}}\right)$$

we have the same results

$$\eta\left(\frac{1}{P_n}\sum_{i=1}^{n}\sigma_i c_i\right) \leq \ldots \leq \tilde{\eta}_{r+1,n} \leq \tilde{\eta}_{r,n} \leq \ldots \leq \tilde{\eta}_{1,n} = \frac{1}{P_n}\sum_{i=1}^{n}\sigma_i\eta(c_i). \tag{1.11}$$

The researchers have given the refinements of (1.3) by using different indexing sets (see [50, 48]). Like many other researchers L. Horváth and J. Pečarić gave a refinement of

(1.3) for convex functions (see [50]). They defined some essential notions to prove the refinement given as follows:

Let X be a set, let $P(X)$ and $|X|$ represent the power set and number of elements of set X respectively. Let $\mathbb{N} := \{0\} \bigcup \{1,2,\ldots\}$.
Suppose $q \geq 1$ and $r \geq 2$ are two fixed integers. Suppose

$$\nabla_r(q) := \{1,\ldots,q\}^r$$

Now let

$$F_{r,s} : \nabla_r(q) \to \nabla_{r-1}(q) \quad 1 \leq s \leq r,$$

$$F_r : \nabla_r(q) \to P\left(\nabla_{r-1}(q)\right),$$

and

$$T_r : P(\nabla_r(q)) \to P\left(\nabla_{r-1}(q)\right),$$

are functions defined by

$$F_{r,s}(i_1,\ldots,i_r) := (i_1,i_2,\ldots,i_{s-1},i_{s+1},\ldots,i_r) \quad 1 \leq s \leq r,$$

$$F_r(i_1,\ldots,i_r) := \bigcup_{s=1}^{r} \{F_{r,s}(i_1,\ldots,i_r)\},$$

and

$$T_r(I) = \begin{cases} \emptyset, & I = \emptyset; \\ \bigcup\limits_{(i_1,\ldots,i_r)\in I} F_r(i_1,\ldots,i_r), & I \neq \emptyset. \end{cases}$$

Next for all $i \in \{1,\ldots,q\}$ consider

$$\alpha_{r,i} : \{1,\ldots,q\}^r \to \mathbb{N},$$

defined by

$$\alpha_{r,i}(i_1,\ldots,i_r) \text{ is the number of occurences of } i_j \text{ in } (i_1,\ldots,i_r).$$

For each $I \in P\left(\nabla_r(q)\right)$ let

$$\alpha_{I,i} := \sum_{(i_1,\ldots,i_r)\in I} \alpha_{r,i}(i_1,\ldots,i_r) \quad 1 \leq i \leq q.$$

$(\mathbf{H_1})$: Let $n \in \{1,2,\ldots\}$ and $m \in \{2,3,\ldots\}$, suppose $I_m \subset \nabla_m(n)$ such that for all $i \in \{1,\ldots,n\}$

$$\alpha_{I_m,i} \geq 1. \tag{1.12}$$

Introduce the set $I_l \subset \nabla_l(n)(1 \leq l \leq m-1)$ inductively by

$$I_{l-1} := T_l(I_l) \quad m \geq l \geq 2.$$

Obviously the set $I_1$is $\{1,\ldots,n\}$, by (H_1) and this make certain that $\alpha_{I_1,i} = 1(1 \leq i \leq n)$. From (H_1) we have $\alpha_{I_l,i} \geq 1(1 \leq i \leq n, m-1 \geq l \geq 1)$.
For $m \geq l \geq 2$, and for any $(j_1,\ldots,j_{l-1}) \in I_{l-1}$, let

$$\mathscr{H}_{I_l}(j_1,\ldots,j_{l-1}) := \{((i_1,\ldots,i_l),k) \times \{1,\ldots,l\} | F_{l,k}(i_1,\ldots,i_l) = (j_1,\ldots,j_{l-1})\}.$$

With the help of these sets they defined the functions $\eta_{I_m,l} : I_l \to \mathbb{N}(m \geq l \geq 1)$ inductively by

$$\eta_{I_m,m}(i_1,\ldots,i_m) := 1 \quad (i_1,\ldots,i_m) \in I_m;$$
$$\eta_{I_m,l-1}(j_1,\ldots,j_{l-1}) := \sum_{((i_1,\ldots,i_l),k) \in \mathscr{H}_{I_l}(j_1,\ldots,j_{l-1})} \eta_{I_m,l}(i_1,\ldots,i_l).$$

They defined some special expressions for $1 \leq l \leq m$, as follows

$$\mathscr{A}_{m,l} = \mathscr{A}_{m,l}(I_m, x_1,\ldots,x_n, p_1,\ldots,p_n; f) := \frac{(m-1)!}{(l-1)!} \sum_{(i_1,\ldots,i_l) \in I_l} \eta_{I_m,l}(i_1,\ldots,i_l)$$
$$\left(\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}\right) f\left(\frac{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}} x_{i_j}}{\sum_{j=1}^{l} \frac{p_{i_j}}{\alpha_{I_m,i_j}}}\right) \tag{1.13}$$

and constructed the following new refinement of (1.3).

Theorem 1.1 *Assume (H_1), consider a convex function $f : I \to \mathbb{R}$. If $c_1,\ldots,c_n \in I$, $\sigma_1,\ldots,\sigma_n \in \mathbb{R}^+$ such that $\sum_{s=1}^{n} \sigma_s = 1$, then*

$$f\left(\sum_{s=1}^{n} \sigma_s c_s\right) \leq \mathscr{A}_{m,m} \leq \mathscr{A}_{m,m-1} \leq \ldots \leq \mathscr{A}_{m,2}$$
$$\leq \mathscr{A}_{m,1} = \sum_{s=1}^{n} \sigma_s f(c_s). \tag{1.14}$$

L. Horváth and J. Pečarić proved that (1.10) is the special case of Theorem 1.1.

In [25], I. Brnetić *et al*. gave the improvement of (1.6) as follows.
Suppose $\eta : I \to \mathbb{R}$ is a convex function, let $\mu : [a_1,a_2] \to I$ and $\chi : [a_1,a_2] \to \mathbb{R}^+$ be functions. Suppose $\sigma_1,\ldots,\sigma_n \in \mathbb{R}^+$ with $\sum_{i=1}^{n} \sigma_i = 1$, and

$$\bar{\chi} = \int_{a_1}^{a_2} \chi(t)dt,$$

then

$$\eta\left(\frac{1}{\bar{\chi}} \int_{a_1}^{a_2} \chi(t)\mu(t)dt\right) \leq \Delta_{n,n} \leq \ldots \leq \Delta_{k+1,n} \leq \Delta_{k,n} \ldots \leq \Delta_{1,n}$$

$$= \frac{1}{\overline{\chi}} \int_{a_1}^{a_2} \chi(t)\eta(\mu(t))dt,$$

where

$$\Delta_{k,n} = \frac{1}{\binom{n-1}{k-1}\overline{\chi}^k} \sum_{1 \le i_1 < \ldots < i_k \le n} \sum_{r=1}^{k} \sigma_{i_r} \int_{a_1}^{a_2} \cdots \int_{a_1}^{a_2} \left(\prod_{s=1}^{k} \chi(c_{i_s}) \right)$$
$$\times \eta \left(\frac{\sum_{j=1}^{k} \sigma_{i_j} \mu\left(c_{i_j}\right)}{\sum_{j=1}^{k} \sigma_{i_j}} \right) dc_{i_1} dc_{i_2} \ldots dc_{i_k}.$$

1.1.8 Jensen-Mercer's Inequality

In 2003 Mercer proved the following variant of Jensen's inequality, which is known as Jensen-Mercer's inequality.

Theorem 1.2 ([77]) *Let $\psi : [a,b] \to \mathbb{R}$ be a convex function and let $y_\gamma \in [a,b]$, $\zeta_\gamma \in \mathbb{R}^+$ be such that $\zeta = \sum_{\gamma=1}^{n} \zeta_\gamma$. Then*

$$\psi\left(a + b - \frac{1}{\zeta}\sum_{\gamma=1}^{n} \zeta_\gamma y_\gamma\right) \le \psi(a) + \psi(b) - \frac{1}{\zeta}\sum_{\gamma=1}^{n} \zeta_\gamma \psi(y_\gamma). \tag{1.15}$$

1.1.9 Variant of Jensen-Steffensen's Inequality

The following variant of Jensen-Steffensen's inequality has been given in [1].

Theorem 1.3 *Let $\psi : [a,b] \to \mathbb{R}$ be a convex function and $y_\gamma \in [a,b]$, $\zeta_\gamma \in \mathbb{R}, \zeta_\gamma \neq 0$ for $\gamma = 1,2,\ldots,n$ with $\zeta = \sum_{\gamma=1}^{n} \zeta_\gamma$. If $y_1 \le y_2 \le \cdots \le y_n$ or $y_1 \ge y_2 \ge \cdots \ge y_n$ and*

$$0 \le \sum_{\gamma=1}^{k} \zeta_\gamma \le \sum_{\gamma=1}^{n} \zeta_\gamma,\ k = 1,2,\ldots,n, \quad \sum_{\gamma=1}^{n} \zeta_\gamma > 0, \tag{1.16}$$

then

$$\psi\left(a + b - \frac{1}{\zeta}\sum_{\gamma=1}^{n} \zeta_\gamma y_\gamma\right) \le \psi(a) + \psi(b) - \frac{1}{\zeta}\sum_{\gamma=1}^{n} \zeta_\gamma \psi(y_\gamma). \tag{1.17}$$

1.1.10 Jensen's Inequality for Operator Convex Functions

The following Jensen's inequality for operator convex function has been given in [43].

Theorem 1.4 (JENSEN'S OPERATOR INEQUALITY) *Let $\psi : I \to \mathbb{R}$ be an operator convex function defined on the interval I. If $A_p \in S(I)$ and $\zeta_p > 0$ $(p = 1,\ldots,n)$ such that $\sum_{p=1}^{n} \zeta_p = 1$, then*

$$\psi\left(\sum_{p=1}^{n} \zeta_p A_p\right) \le \sum_{p=1}^{n} \zeta_p \psi(A_p). \tag{1.18}$$

1.1.11 Hermite-Hadamard Inequality

The following inequality proved by Hermite and Hadamard for convex function [40]. This inequality says that if the function $\psi : [a,b] \to \mathbb{R}$ is convex function then

$$\psi\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b \psi(x)dx \le \frac{\psi(a)+\psi(b)}{2}. \tag{1.19}$$

If ψ is concave function then the inequalities in (1.19) will hold in reverse directions.

1.1.12 Hölder Inequality

The discrete form of well-known Hölder inequality is given below:

Let $p,q > 1$ with $\frac{1}{p}+\frac{1}{q} = 1$ and $(a_1,a_2,\cdots,a_n)$, $(b_1,b_2,\cdots,b_n)$ be positive n-tuple. Then

$$\sum_{\mathtt{j}=1}^{n} a_{\mathtt{j}} b_{\mathtt{j}} \le \left(\sum_{\mathtt{j}=1}^{n} a_{\mathtt{j}}^{p}\right)^{\frac{1}{p}} \left(\sum_{\mathtt{j}=1}^{n} b_{\mathtt{j}}^{q}\right)^{\frac{1}{q}} \tag{1.20}$$

The integral form of Hölder inequality is given below:

Let $\mathtt{p},\mathtt{q} > 1$ such that $\frac{1}{\mathtt{p}}+\frac{1}{\mathtt{q}} = 1$ and $\mathtt{A},\mathtt{B} : [a,b] \to \mathbb{R}$ be integrable functions functions such that $|\mathtt{A}(\mathtt{z})|^{\mathtt{p}}$, $|\mathtt{B}(\mathtt{z})|^{\mathtt{q}}$ are also integrable on $[a,b]$. Then

$$\int_a^b |\mathtt{A}(\mathtt{z})\mathtt{B}(\mathtt{z})|\,d\mathtt{z} \le \left(\int_a^b |\mathtt{A}(\mathtt{z})|^{\mathtt{p}}\,d\mathtt{z}\right)^{\frac{1}{\mathtt{p}}} \left(\int_a^b |\mathtt{B}(\mathtt{z})|^{\mathtt{q}}\,d\mathtt{z}\right)^{\frac{1}{\mathtt{q}}}.$$

1.2 Power Means

In [75, p. 14] the power means are given as follows.

Suppose n is a natural number, let $(c_1,\ldots,c_n)$ and $(\sigma_1,\ldots,\sigma_n)$ belong to $(0,\infty)^n$ such that $P_n := \sum_{i=1}^{n}\sigma_i = 1$. The power mean (of order $s \in \mathbb{R}$) is defined by

$$P(c_1,\ldots,c_n;\sigma_1,\ldots,\sigma_n) = \begin{cases} \left(\frac{1}{P_n}\sum_{i=1}^{n} c_i^s\right)^{\frac{1}{s}}, & s \neq 0; \\ \left(\prod_{i=1}^{n} c^{\sigma_i}\right)^{\frac{1}{P_n}}, & s = 0. \end{cases} \tag{1.21}$$

For $c_i = \frac{1}{n}$, $i = 1,\ldots,n$, the power mean (1.21) is arithmetic mean, geometric mean and harmonic mean for $s = 1$, $s \to 0$ and $s = -1$ respectively.

The power means for the n-tuples of strictly positive operators $\mathbf{A} = (A_1,\ldots,A_n)$ with positive weights $\boldsymbol{\zeta} := (\zeta_1,\ldots,\zeta_n)$ of order $r \in \mathbb{R}\setminus\{0\}$ is defined by:

$$M_r(\mathbf{A};\boldsymbol{\zeta}) = \left(\frac{1}{\overline{\zeta}}\sum_{p=1}^{n}\zeta_p A_p^r\right)^{\frac{1}{r}}, \tag{1.22}$$

where $\overline{\zeta} := \sum_{p=1}^{n}\zeta_p$.

1.2.1 Quasi-Arithmetic Means

The importance of quasi-arithmetic means has been well understood at least since the 1930's and a number of writers have since then contributed to the characterization and to the study of their properties.

Consider a continuous function $\eta : I \to \mathbb{R}$ such that for all $u, v \in I$ if $u < v$ then $\eta(u) < \eta(v)$ (or if $u > v$ then $\eta(u) > \eta(v)$). Let $(\lambda_1, \ldots, \lambda_n) \in I^n$, also let $(\sigma_1, \ldots, \sigma_n) \in [0, \infty)^n$. Suppose $P_n := \sum_{i=1}^{n} p_i$. Then the quasi-arithmetic mean [75, p. 15] is

$$M_{\eta}^{[n]}(\boldsymbol{\lambda};\boldsymbol{\sigma}) = M_{\eta}(\lambda_1, \ldots, \lambda_n; \sigma_a \ldots, \sigma_n) = \eta^{-1}\left(\frac{1}{P_n}\sum_{i=1}^{n}\sigma_i\eta(\lambda_i)\right). \tag{1.23}$$

If $I = \mathbb{R}^+$ and $\eta(t) = t^p$, then (1.23) is a power mean.

In the current century, the Popoviciu inequality is studied by many authors (see [31, 29, 32, 30]).

The Popoviciu inequality for arbitrary non-negative weights given as follows (see [85]).

Let r and m are positive integers such that $m \geq 3$, $2 \leq r \leq m-1$, let $\eta : [a_1, a_2] \to \mathbb{R}$ be convex function, $(c_1, \ldots, c_m) \in [a_1, a_2]^m$ and $(\sigma_1, \ldots, \sigma_m)$ be non-negative m-tuple such that $\sum_{j=1}^{n} \sigma_j = 1$, then

$$\begin{aligned}\eta_{r,m}(c_1, \ldots, c_m; \sigma_1, \ldots, \sigma_m) \leq \frac{m-r}{m-1}\eta_{1,m}(c_1, \ldots, c_m; \sigma_1, \ldots, \sigma_m)\\ +\frac{r-1}{m-1}\eta_{m,m}(c_1, \ldots, c_m; \sigma_1, \ldots, \sigma_m),\end{aligned} \tag{1.24}$$

where

$$g_{r,m}(c_1, \ldots, c_m; \sigma_1, \ldots, \sigma_m) := \frac{1}{C_{r-1}^{m-1}} \sum_{1 \leq i_1 < \ldots < i_r \leq m} \left(\sum_{j=1}^{m} \sigma_{i_j}\right) \eta\left(\frac{\sum_{j=1}^{m} \sigma_{i_j} c_{i_j}}{\sum_{j=1}^{m} \sigma_{i_j}}\right).$$

Higher order convex function was introduced by T. Popoviciu (see [87, p. 15]). The inequalities involving higher order convex functions are used by physicists in higher dimensional problems. Many of the results that are true for convex functions are not true for higher order convex functions, this fact convince us to study the results involving higher order convexity (see [31]).

Let $\eta : I \to \mathbb{R}$ be a continuous strictly monotone function. Then the quasi arithmetic mean for operators is defined by

$$\tilde{M}_{\eta}(\mathbf{A}; \boldsymbol{\zeta}) = \eta^{-1}\left(\frac{1}{\sum_{p=1}^{n} \zeta_p} \sum_{p=1}^{n} \zeta_p \eta(A_p)\right), \tag{1.25}$$

where $A_p \in S(I)$ and $\zeta_p > 0$ for $p = 1, 2, \ldots, n$.

1.3 Divided Differences

The tools of divided difference are used to define the higher order convex functions. Divided difference is given in [87, p. 14] as follows.
Consider the function $\eta : [a_1, a_2] \to \mathbb{R}$. The r-th order divided difference for $r+1$ distinct points $u_0, u_1, \ldots, u_r \in [a_1, a_2]$ is defined by the following recursive formula

$$[u_i; \eta] = \eta(u_i) \;\; i = 0, 1, \ldots, r,$$

and

$$[u_0, u_1, \ldots, u_r; \eta] = \frac{[u_1, u_2, \ldots, u_r; \eta] - [u_0, u_1, \ldots, u_{r-1}; \eta]}{u_r - u_0}. \tag{1.26}$$

This is equivalent to

$$[u_0, u_1, \ldots, u_r; \eta] = \sum_{j=0}^{k} \frac{\eta(u_j)}{w'(u_j)},$$

where $w(u) = \prod_{j=0}^{k} (u - u_j)$. This definition may be extended to include the case in which some or all the points coincide. Namely, if all the points are same, then by taking limits in (1.26) we obtain

$$[\underbrace{u, u, \ldots, u}_{l-times}; \eta] = \frac{\eta^{(l-1)}(u)}{(l-1)!}, \tag{1.27}$$

where $\eta^{(l-1)}$ is supposed to exist.

1.4 Higher Order Convex Functions

A function $\eta : [a_1, a_2] \to \mathbb{R}$ is called r-convex function ($r \geq 0$) on $[a_1, a_2]$ if and only if

$$[u_0, u_1, \ldots, u_r; \eta] \geq 0 \tag{1.28}$$

for all $(r+1)$ distinct choices in $[a_1, a_2]$ (see [87, p. 14]).

The function η is r-concave on $[a_1, a_2]$ if inequality sign in (1.28) is reverse. The next result is useful to examine the convexity of a function [87, p. 16].

Theorem 1.5 *Suppose the $\eta^{(n)}$ exists where η is a real valued function. Then η is n-convex if and only if $\eta^{(n)}$ is non-negative.*

In recent years many researchers have generalized the inequalities for m-convex functions; like S. I. Butt *et al*. generalized the Popoviciu inequality for m-convex function using Taylor's formula, Lidstone polynomial, Montgomery identity, Fink's identity, Abel-Gonstcharoff interpolation and Hermite interpolating polynomial (see [31, 29, 32, 30, 33]). S. I. Butt *et al*. constructed the linear functional from these generalized Popoviciu type identities and using the inequalities for Cebysev functional and found some bounds for the generalized identities. Also they constructed Grüss and Ostrowski type inequalities. By using these new generalized Popoviciu type functionals they constructed new class of m-exponentially convex functions.

1.5 Information Divergence Measures and Entropies

Information theory is the science of information, which scientifically deals with the storage, quantification and communication of the information. Being an abstract entity information cannot be quantified easily. In 1948 Claude shannon in [93], presented the concept of information theory and introduced entropy as the fundamental measure of information in his first of the two fundamental and important theorems. The information can also be measured with the help of probability density function. Divergences are some important tools for measuring the difference between two probability density functions. A class of information divergence measures, which is one of the important divergence measures due to its compact behavior, is the Csiszár divergence [36, 37].

Let η be a positive function defined on $(0,\infty)$, suppose $\mathbf{p} := (p_1,\ldots,p_n)$ and $\mathbf{q} := (q_1,\ldots,q_n)$ are positive probability distributions. The Csiszár divergence (f-divergence) is defined as

$$I_f(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right). \tag{1.29}$$

In [52], L. Horv́ath, et al. gave the following generalization of (1.29):

Let $\eta : I \to \mathbb{R}$ be a function. Suppose $\mathbf{p} := (p_1,\ldots,p_n)$ is real and $\mathbf{q} := (q_1,\ldots,q_n)$ is positive n-tuple such that

$$\frac{p_j}{q_j} \in \mathbb{R}, \;\; j = 1,\ldots,n. \tag{1.30}$$

Then

$$\hat{I}_\eta(\mathbf{p},\mathbf{q}) := \sum_{j=1}^{n} q_j \eta\left(\frac{p_j}{q_j}\right). \tag{1.31}$$

They applied the cyclic refinement of Jensen's inequality [52] to $\hat{I}_f(\mathbf{p},\mathbf{q})$ in order to investigate the bounds for (1.31).