Contemporary Mathematics
http://ojs.wiserpub.com/index.php/CM/

UNIVERSAL WISER PUBLISHER

Research Article

Some Generalizations of Jensen's Inequality

Slavko Simić[1*], Bandar Almohsen[2]

[1]Mathematical Institute of the Serbian Academy of Sciences and Arts, 11000 Belgrade, Serbia
[2]College of Science, Department of Mathematics, King Saud University, Riyadh, Saudi Arabia
E-mail: ssimic@turing.mi.sanu.ac.rs

Received: 9 October 2020; **Revised:** 2 December 2020; **Accepted:** 8 December 2020

Abstract: In this article, we give some improvements and generalizations of the famous Jensen's and Jensen-Mercer inequalities for twice differentiable functions, where convexity property of the target function is not assumed in advance. They represent a refinement of these inequalities in the case of convex/concave functions with numerous applications in Theory of Means and Probability and Statistics.

***Keywords*:** Jensen's inequality, Jensen-Mercer inequality, twice differentiable functions, convex functions

MSC: 26D07(26D15)

1. Introduction

Recall that the Jensen functional $J_n(\mathbf{p}, \mathbf{x}; f)$ is defined on an interval $I \subseteq \mathbb{R}$ by

$$J_n(\mathbf{p}, \mathbf{x}; f) := \sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i),$$

where $f: I \to \mathbb{R}$, $\mathbf{x} = (x_1, x_2, ..., x_n) \in I^n$ and $\mathbf{p} = \{p_i\}_1^n$ is a positive weight sequence.

Let us now state the celebrated Jensen's inequality [1, 2], key to solve lots of variational problems, see e.g. [3, 4].

Jensen's inequality ([5]) If f is twice continuously differentiable function and $f'' \geq 0$ on an interval I, then f is convex on I and the inequality

$$0 \leq J_n(\mathbf{p}, \mathbf{x}; f)$$

holds for each $\mathbf{x} := (x_1, ..., x_n) \in I^n$ and any positive weight sequence $\mathbf{p} := \{p_i\}_1^n$ with $\sum_1^n p_i = 1$.

If $f'' \leq 0$ on I, then f is a concave function on I and

$$J_n(\mathbf{p}, \mathbf{x}; f) \leq 0.$$

Its counterpart is given by the following

DOI: https://doi.org/10.37256/cm.212021686

Jensen-Mercer inequality ([6]) Let ϕ : $[a, b] \subseteq \mathbb{R} \to \mathbb{R}$ be a convex function and $x_i \in [a, b]$, $i = 1, 2, ..., n$. Then

$$\phi(a+b-\sum_1^n p_i x_i) \le \phi(a)+\phi(b)-\sum_1^n p_i\phi(x_i). \quad (1)$$

Our first task in this paper is to find some global upper bounds for these inequalities. We prove the following.
Let f be a convex function on an interval I and $x_i \in [a, b] \subset I$. Then

$$0 \le \sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i) \le f(a)+f(b)-2f(\frac{a+b}{2});$$

$$0 \le f(a)+f(b)-\sum_1^n p_i f(x_i) - f(a+b-\sum_1^n p_i x_i) \le 2(f(a)+f(b)-2f(\frac{a+b}{2})).$$

Those bounds can be improved by the *characteristic* number $c(f)$ of the convex function f (cf. [7]), to the next

$$0 \le \sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i) \le c(f)[f(a)+f(b)-2f(\frac{a+b}{2})];$$

$$0 \le f(a)+f(b)-\sum_1^n p_i f(x_i) - f(a+b-\sum_1^n p_i x_i) \le (1+c(f))[f(a)+f(b)-2f(\frac{a+b}{2})].$$

As an example, we shall calculate characteristic number for the power function:

$$c(x^s)\begin{cases} 1, & s<0; \\ (1-s)s^{s/(1-s)}/(2^{1-s}-1), & 0<s<1; \\ (s-1)s^{s/(1-s)}/(1-2^{1-s}), & s>1. \end{cases}$$

Our second main task is to investigate the possibility of a form of Jensen's and Jensen-Mercer inequalities for functions which are not necessarily convex/concave on I.

The sole condition will be that the second derivative of the target function exists locally i.e., on a closed interval $E := [a, b] \subset I$. Since it is continuous on a closed interval, there exist numbers $m_f(E) = m(a, b; f) := \min_{t\in E} f''(t)$ and $M_f(E) = M(a, b; f) := \max_{t\in E} f''(t)$. Those numbers will play an important role in the sequel.

For instance, let $f \in C^{(2)}(E)$ and $x_i \in E$, $i = 1, 2, ..., n$. Then

$$\frac{1}{2}m_f(E)J_n(\mathbf{p},\mathbf{x};\, x^2) \le J_n(\mathbf{p},\mathbf{x};\, f) \le \frac{1}{2}M_f(E)J_n(\mathbf{p},\mathbf{x};\, x^2).$$

Note that this inequality represents an improvement of Jensen's inequality for convex functions since in this case we have $0 \le m_f(E) \le M_f(E)$.

2. Results and proofs

We firstly prove some global upper bounds for Jensen's and Jensen-Mercer inequalities. This will be done by an application of the following assertion from [8].

Lemma 2.1 Let h be a convex function on $E = [a, b]$ and, for some $x, y \in E$, $x + y = a + b$. Then

$$2h(\frac{a+b}{2}) \leq h(x) + h(y) \leq h(a) + h(b).$$

Theorem 2.2 Let f be a convex function on I and $\mathbf{x} \in [a, b]^n \subset I^n$. Then

$$0 \leq J_n(\mathbf{p}, \mathbf{x}; f) = \sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i) \leq f(a) + f(b) - 2f(\frac{a+b}{2}); \tag{2}$$

$$0 \leq f(a) + f(b) - \sum_1^n p_i f(x_i) - f(a + b - \sum_1^n p_i x_i) \leq 2[f(a) + f(b) - 2f(\frac{a+b}{2})], \tag{3}$$

independently of $\mathbf{p}$.

Proof. We obtain a simple proof of (2) directly from Jensen-Mercer inequality.

Namely, writing this inequality in the form

$$\sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i) \leq f(a) + (b) - (f(\sum_1^n p_i x_i) + f(a + b - \sum_1^n p_i x_i)),$$

the proof follows by Lemma 2.1.

For the proof of the assertion (3), note that if $x_i \in [a, b]$ then also $y_i := a + b - x_i \in [a, b]$. Hence, by (2) and Lemma 2.1, we get

$$f(a) + f(b) - 2f(\frac{a+b}{2}) \geq \sum_1^n p_i f(y_i) - f(\sum_1^n p_i y_i)$$

$$= \sum_1^n p_i f(a + b - x_i) - f(\sum_1^n p_i(a + b - x_i)) \geq \sum_1^n p_i[2f(\frac{a+b}{2}) - f(x_i)] - f(a + b) - \sum_1^n p_i x_i)$$

$$= f(a) + f(b) - \sum_1^n p_i f(x_i) - f(a + b - \sum_1^n p_i x_i) - [f(a) + f(b) - 2f(\frac{a+b}{2})],$$

and the proof is done.

Those bounds can be improved by the following

Theorem 2.3 Let h be a convex function on $E = [a, b]$ and $p, q > 0$; $p + q = 1$.

Then

$$\min\{p, q\}[h(a) + h(b) - 2h(\frac{a+b}{2})] \leq ph(a) + qh(b) - h(pa + qb) \leq \max\{p, q\}[h(a) + h(b) - 2h(\frac{a+b}{2})].$$

Proof. If $p = q(= 1/2)$ there is an identity in the above relations. Hence, assuming that $p > q$, we have

$$\max\{p, q\}[h(a) + h(b) - 2h(\frac{a+b}{2})] - [ph(a) + qh(b) - h(pa + qb)]$$

$$= (p-q)h(b) + h(pa+qb) - 2ph(\frac{a+b}{2}) = 2p[\frac{p-q}{2p}h(b) + \frac{1}{2p}h(pa+qb) - h(\frac{a+b}{2})]$$

$$\geq 2p[h(\frac{p-q}{2p}b + \frac{pa+qb}{2p}) - h(\frac{a+b}{2})] = 0.$$

Proof of the left-hand side inequality goes along the same lines.

Now it is not difficult to prove by induction the following

Theorem 2.4 We have

$$\min\{p_i\}[\sum_1^n h(x_i) - nh(\frac{1}{n}\sum_1^n x_i)] \leq J_n(\mathbf{p}, \mathbf{x};\ h) \leq \max\{p_i\}[\sum_1^n h(x_i) - nh(\frac{1}{n}\sum_1^n x_i)].$$

This theorem represents an improved variant of Corollary 2.4 from [9].

Another way to sharpen the global bounds in Theorem 2.2 is to use notion of the *characteristic* number $c(f)$ of a given convex function f. Namely, it is proved in ([7]) that there exists a number $c(f) \in [1/2, 1]$ depending only on f, such that

$$0 \leq \sum_1^n p_i f(x_i) - f(\sum_1^n p_i x_i) \leq c(f)[f(a) + f(b) - 2f(\frac{a+b}{2})]; \tag{4}$$

and, consequently,

$$0 \leq f(a) + f(b) - \sum_1^n p_i f(x_i) - f(a+b-\sum_1^n p_i x_i) \leq (1+c(f))[f(a)+f(b)-2f(\frac{a+b}{2})]. \tag{5}$$

The characteristic number $c(f)$ is defined as

$$c(f) := \sup_{p,q;a,b} \frac{pf(a) + qf(b) - f(pa+qb)}{f(a)+f(b) - 2f(\frac{a+b}{2})}.$$

By direct calculation we obtain

$$c(x^2) = \sup_{p,q;a,b} \frac{pa^2 + qb^2 - (pa+qb)^2}{a^2 + b^2 - 2(\frac{a+b}{2})^2} = \sup_{p,q} 2pq = 1/2.$$

We shall determine now the value of this constant for some classes of functions.

For this cause, recall the definitions of slowly varying and rapidly varying functions (cf. [10]).

Definition Let the function f be defined on $I := [a, +\infty)$.

It is said that f is slowly varying if $\lim_{x\to\infty} \frac{f(tx)}{f(x)} = 1$ for any $t > 0$.

If $\lim_{x\to\infty} \frac{f(tx)}{f(x)} = \infty$ for any $t > 1$, then f is a rapidly varying function.

Theorem 2.5 Let $f(a + x) := g_a(x)$ be a slowly or rapidly varying function. Then $c(f) = 1$.

Proof. Denote

$$H := \frac{pf(a)+qf(b)-f(pa+qb)}{f(a)+f(b)-2f(\frac{a+b}{2})} = \frac{pf(a)+qg_a(x)-g_a(qx)}{f(a)+g_a(x)-2g_a(\frac{x}{2})},$$

with $x = b - a$.

Since f is a convex function, so is $g_a(x)$. Hence $\lim_{x\to\infty} g_a(x)$ can be 0, c or $\pm\infty$.

In the first two cases, we obtain at once that $\lim_{x\to\infty} H = p$. Since $g_a(x)$ is also slowly varying, in the third case we get

$$\lim_{x\to\infty} H = \frac{pf(a)/g_a(x)+q-g_a(qx)/g_a(x)}{f(a)/g_a(x)+1-2g_a(\frac{x}{2})/g_a(x)} = \frac{q-1}{-1} = p.$$

As concerns the class of rapidly varying functions, note that $\lim_{x\to\infty} \frac{f(tx)}{f(x)} = 0$ for $0 < t < 1$, which can be easily proven by the change of variable $tx \to x$, $1/t \to t$.

Therefore, in this case we have

$$\lim_{x\to\infty} H = \frac{pf(a)/g_a(x)+q-g_a(qx)/g_a(x)}{f(a)/g_a(x)+1-2g_a(\frac{x}{2})/g_a(x)} = q.$$

Since p and q are arbitrary weights, we conclude that $c(f) = 1$ in both cases. For instance,

$$c(-\log x) = c(e^{-x}) = c(e^{x}) = c(x^{x}) = 1.$$

Our next contribution is an evaluation of the characteristic number for the power function.

Theorem 2.6 We have

$$c(x^s)\begin{cases} 1, & s < 0; \\ (1-s)s^{s/(1-s)}/(2^{1-s}-1), & 0 < s < 1; \\ (s-1)s^{s/(1-s)}/(1-2^{1-s}), & s > 1. \end{cases}$$

Proof. Main tool for the proof of this and similar theorems will be the following useful assertion from [11].

Lemma 2.7 For $-\infty < a < b < \infty$, let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, and be differentiable on (a, b), and let $g'(x) \neq 0$ on (a, b). If $f'(x)/g'(x)$ is increasing (deceasing) on (a, b), then so are

$$\frac{f(x)-f(a)}{g(x)-g(a)} \text{ and } \frac{f(b)-f(x)}{g(b)-g(x)}.$$

If $f'(x)/g'(x)$ is strictly monotone, then the monotonicity in the conclusion is also strict.

Let $a, p, q \in \mathbb{R}^+$, $p + q = 1$, $p \neq q$; $x \in (a, +\infty)$ and $s \in (0, 1) \cup (1, 2) \cup (2, +\infty)$.

Denote $f_1(x) = (q + pa/x)^{s-1}$; $g_1(x) = ((1 + a/x)/2)^{s-1}$.

Since

$$\frac{f_1'(x)}{g_1'(x)} = 2p\frac{(q+pa/x)^{s-2}}{((1+a/x)/2)^{s-2}} = 2p\left(\frac{pa+qx}{(a+x)/2}\right)^{s-2},$$

by Lemma (2.7) we conclude that the expression

$$\frac{f_1(x)-f_1(a)}{g_1(x)-g_1(a)}=\frac{(q+pa/x)^{s-1}-1}{((1+a/x)/2)^{s-1}-1}=\frac{x^{s-1}-(pa+qx)^{s-1}}{x^{s-1}-((a+x)/2)^{s-1}},$$

is monotone increasing for $q > p$, $s \in (2, +\infty)$ or $p > q$, $s \in (0, 1) \cup (1, 2)$ and monotone decreasing otherwise.

Denote now $f_2(x) = qx^s - (pa + qx)^s$; $g_2(x) = x^s - 2((a + x)/2)^s$.

Since

$$\frac{f_2'(x)}{g_2'(x)}=q\frac{x^{s-1}-(pa+qx)^{s-1}}{x^{s-1}-((a+x)/2)^{s-1}},$$

we conclude the same for

$$\frac{f_2(x)-f_2(a)}{g_2(x)-g_2(a)}=\frac{pa^s+qx^s-(pa+qx)^s}{a^s+x^s-2((a+x)/2)^s}:=H(x).$$

Hence, the maximum of $H(x)$ is attained at the endpoints of $(a, +\infty)$.

We have

$$\lim_{x\to a}H(x)=2pq;\quad \lim_{x\to+\infty}H(x)=\frac{q-q^s}{1-2^{1-s}}.$$

Because $\max_q(2pq) = 1/2$ is the least possible value of $c(f)$, we see that

$$c(x^s)=\max_q(q-q^s)/(1-2^{1-s}),$$

and the proof follows.

For $x \in (0, b)$, putting

$$f_1(x) = (p + qb/x)^{s-1},\ g_1(x) = ((1 + b/x)/2)^{s-1};$$

$$f_2(x) = px^s - (px + qb)^s;\ g_2(x) = x^s - 2((b + x)/2)^s,$$

and repeating the above procedure, we obtain the same result.

If $s < 0$, we have $\lim_{x\to\infty} x^s = 0$. Hence $c(x^s) = 1$ according to the previous theorem.

Remark 2.8 The described method can be applied for evaluation of the characteristic number of other convex functions.

For example, it can be proved that $c(x\log x) = (e\log 2)^{-1}$.

Our next achievement is the form of Jensen's and Jensen-Mercer inequalities for nonconvex functions.

Theorem 2.9 Let $g \in C^{(2)}(E)$ and $\mathbf{x} \in E := [a, b] \subset \mathbb{R}$.

Then

$$\frac{1}{2}m_f(E)J_n(\mathbf{p},\mathbf{x};x^2)\le J_n(\mathbf{p},\mathbf{x};g)\le\frac{1}{2}M_f(E)J_n(\mathbf{p},\mathbf{x};x^2).$$

where $m_f(E) := \min_{t\in E} g''(t)$ and $M_f(E) := \max_{t\in E} g''(t)$.

Proof. For a given $g \in C^{(2)}(E)$, define an auxiliary function f by $f(x) := g(x) - m_g(E)x^2/2$. Since $f''(x) = g''(x) - m_g(E) \ge 0$, we see that f is a convex functionon E. Therefore, applying Jensen's inequality, we obtain

$$0 \le J_n(\mathbf{p}, \mathbf{x}; f) = J_n(\mathbf{p}, \mathbf{x}; g) - \frac{1}{2} m_g(E) J_n(\mathbf{p}, \mathbf{x}; x^2).$$

On the other hand, taking the auxiliary function f as $f(x) = M_g(E)x^2/2 - g(x)$, we see that it is also convex on E. Applying Jensen's inequality again, we get

$$0 \le J_n(\mathbf{p}, \mathbf{x}; f) = \frac{1}{2} M_g(E) J_n(\mathbf{p}, \mathbf{x}; x^2) - J_n(\mathbf{p}, \mathbf{x}; g),$$

and the proof is done.

Another form is possible.

Theorem 2.10 Let $g \in C^{(2)}(E)$ and $\mathbf{x} \in E := [a, b] \subset \mathbb{R}$.

Then

$$g(a) + g(b) - 2g(\frac{a+b}{2}) + \frac{1}{4} M_g(E)[2J_n(\mathbf{p}, \mathbf{x}; x^2) - (b-a)^2]$$

$$\le J_n(\mathbf{p}, \mathbf{x}; g) \le$$

$$g(a) + g(b) - 2g(\frac{a+b}{2}) + \frac{1}{4} m_g(E)[2J_n(\mathbf{p}, \mathbf{x}; x^2) - (b-a)^2].$$

Proof. Applying the same auxiliary functions to the converse of Jensen's inequality (2), we obtain the desired result.

Two-sided improvement of Jensen's inequality is given by the next

Theorem 2.11 Let $f \in C^{(2)}(E)$ be a convex function and $\mathbf{x} \in E := [a, b] \subset \mathbb{R}$.

Then

$$\frac{m_f(E)}{m_f(E) + M_f(E)}[f(a) + f(b) - 2f(\frac{a+b}{2})] + \frac{m_f(E) M_f(E)}{m_f(E) + M_f(E)}(J_n(\mathbf{p}, \mathbf{x}; x^2) - \frac{1}{4}(b-a)^2)$$

$$\le J_n(\mathbf{p}, \mathbf{x}; f) \le$$

$$\frac{M_f(E)}{m_f(E) + M_f(E)}[f(a) + f(b) - 2f(\frac{a+b}{2})] + \frac{m_f(E) M_f(E)}{m_f(E) + M_f(E)}(J_n(\mathbf{p}, \mathbf{x}; x^2) - \frac{1}{4}(b-a)^2).$$

Proof. Adjusting the right-hand parts of Theorem 2.9 and Theorem 2.10, we obtain

$$J_n(\mathbf{p}, \mathbf{x}; f) \le \frac{M_f(E)}{m_f(E) + M_f(E)}[f(a) + f(b) - 2f(\frac{a+b}{2}) + \frac{1}{4} m_f(E)[2J_n(\mathbf{p}, \mathbf{x}; x^2) - (b-a)^2]]$$

$$+ \frac{m_f(E)}{m_f(E) + M_f(E)}[\frac{1}{2} M_f(E) J_n(\mathbf{p}, \mathbf{x}; x^2)]$$

$$= \frac{M_f(E)}{m_f(E) + M_f(E)}[f(a) + f(b) - 2f(\frac{a+b}{2})] + \frac{m_f(E) M_f(E)}{m_f(E) + M_f(E)}(J_n(\mathbf{p}, \mathbf{x}; x^2) - \frac{1}{4}(b-a)^2).$$

Similarly, adjusting left-hand sides we get

$$J_n(\mathbf{p},\mathbf{x};f) \ge \frac{m_f(E)}{m_f(E)+M_f(E)}[f(a)+f(b)-2f(\frac{a+b}{2})+\frac{1}{4}M_f(E)[2J_n(\mathbf{p},\mathbf{x};x^2)-(b-a)^2]]$$

$$+\frac{M_f(E)}{m_f(E)+M_f(E)}[\frac{1}{2}m_f(E)J_n(\mathbf{p},\mathbf{x};x^2)]$$

$$=\frac{m_f(E)}{m_f(E)+M_f(E)}[f(a)+f(b)-2f(\frac{a+b}{2})]+\frac{m_f(E)M_f(E)}{m_f(E)+M_f(E)}(J_n(\mathbf{p},\mathbf{x};x^2)-\frac{1}{4}(b-a)^2),$$

and the proof follows.

A simple consequence of the previous theorem is another converse of Jensen's inequality.

Corollary 2.12 Because $J_n(\mathbf{p}, \mathbf{x}; x^2) \le \frac{1}{4}(b-a)^2$, we obtain

$$J_n(\mathbf{p},\mathbf{x};f) \le \frac{M_f(E)}{m_f(E)+M_f(E)}[f(a)+f(b)-2f(\frac{a+b}{2})], \tag{6}$$

Remark 2.13 Since $\frac{M_f(E)}{m_f(E)+M_f(E)} \in [\frac{1}{2}, 1]$, it is interesting to compare this result with (4).

A non-convex variant of the Jensen-Mercer inequality follows.

Theorem 2.14 Let $g \in C^{(2)}(E)$ and $\mathbf{x} \in E := [a, b] \subset \mathbb{R}$.

Then

$$\frac{1}{2}m_g(E)[2(\sum_1^n p_i x_i - a)(b-\sum_1^n p_i x_i) - J_n(\mathbf{p},\mathbf{x};x^2)]$$

$$\le g(a)+g(b)-\sum_1^n p_i g(x_i) - g(a+b-\sum_1^n p_i x_i) \le$$

$$\frac{1}{2}M_g(E)[2(\sum_1^n p_i x_i - a)(b-\sum_1^n p_i x_i) - J_n(\mathbf{p},\mathbf{x};x^2)].$$

Proof. Applying Jensen-Mercer inequality

$$0 \le f(a)+f(b)-\sum_1^n p_i f(x_i) - f(a+b-\sum_1^n p_i x_i) := K_n(\mathbf{p},\mathbf{x};f)$$

to the convex function $f(x) = g(x) - \frac{1}{2}m_g(E)x^2$, we get

$$0 \le K_n(\mathbf{p},\mathbf{x};g) - \frac{1}{2}m_g(E)K_n(\mathbf{p},\mathbf{x};x^2)$$

$$= K_n(\mathbf{p},\mathbf{x};g) - \frac{1}{2}m_g(E)[a^2+b^2-(a+b-\sum_1^n p_i x_i)^2 - \sum_1^n p_i x_i^2]$$

$$= K_n(\mathbf{p},\mathbf{x};g) - \frac{1}{2}m_g(E)[-2ab + 2(a+b)\sum_1^n p_i x_i) - 2(\sum_1^n p_i x_i)^2 - (\sum_1^n p_i x_i^2 - (\sum_1^n p_i x_i)^2)]$$

$$= K_n(\mathbf{p},\mathbf{x};g) - \frac{1}{2}m_g(E)[2(\sum_1^n p_i x_i - a)(b - \sum_1^n p_i x_i) - J_n(\mathbf{p},\mathbf{x};x^2)].$$

Consequently, for the function $f(x) = \frac{1}{2}M_g(E)x^2 - g(x)$ we obtain

$$0 \le \frac{1}{2}M_g(E)[2(\sum_1^n p_i x_i - a)(b - \sum_1^n p_i x_i) - J_n(\mathbf{p},\mathbf{x};x^2)] - K_n(\mathbf{p},\mathbf{x};g),$$

and the proof is done.

3. Applications

General means Most known general means are

$$\mathcal{A}(\mathbf{w},\mathbf{x}) := \sum w_i x_i;$$

$$\mathcal{G}(\mathbf{w},\mathbf{x}) := \prod x_i^{w_i};$$

$$\mathcal{H}(\mathbf{w},\mathbf{x}) := (\sum w_i / x_i)^{-1},$$

i.e., arithmetic, geometric and harmonic mean, respectively.

Here $\mathbf{x} = \{x_i\}_1^n$ denotes an arbitrary sequence of positive numbers and $\mathbf{w} = \{w_i\}_1^n$ is a corresponding weight sequence.

The famous $\mathcal{A} - \mathcal{G} - \mathcal{H}$ inequality says that

$$0 \le \mathcal{H}(\mathbf{w}, \mathbf{x}) \le \mathcal{G}(\mathbf{w}, \mathbf{x}) \le \mathcal{A}(\mathbf{w}, \mathbf{x}).$$

It is proved in [2] that $1 \le \mathcal{A}/\mathcal{H} \le (a+b)^2/4ab$, whenever $\mathbf{x} \in [a, b] \subset \mathbb{R}^+$.

The same bounds hold for other $\mathcal{A} - \mathcal{G} - \mathcal{H}$ quotients.

Theorem 3.1 Let $\mathbf{x} \in [a, b] \subset \mathbb{R}^+$. Then

$$1 \le \frac{\mathcal{A}(\mathbf{w},\mathbf{x})}{\mathcal{H}(\mathbf{w},\mathbf{x})} \le \frac{(a+b)^2}{4ab};$$

$$1 \le \frac{\mathcal{A}(\mathbf{w},\mathbf{x})}{\mathcal{G}(\mathbf{w},\mathbf{x})} \le \frac{(a+b)^2}{4ab};$$

$$1 \le \frac{\mathcal{G}(\mathbf{w},\mathbf{x})}{\mathcal{H}(\mathbf{w},\mathbf{x})} \le \frac{(a+b)^2}{4ab};$$

Proof. Since $f(x) = -\log x$ is a convex function on $\mathbb{R}^+$, using Theorem 2.2 we get

$$\log(\sum w_i x_i) - \sum w_i \log x_i \le 2\log\frac{a+b}{2} - \log a - \log b,$$

that is,

$$\log[\frac{\mathcal{A}(\mathbf{w},\mathbf{x})}{\mathcal{G}(\mathbf{w},\mathbf{x})}] \le \log[\frac{(a+b)^2}{4ab}],$$

and the proof follows.

Finally,

$$1 \le \frac{\mathcal{G}(\mathbf{w},\mathbf{x})}{\mathcal{H}(\mathbf{w},\mathbf{x})} = \frac{\mathcal{A}(\mathbf{w},\mathbf{x})}{\mathcal{H}(\mathbf{w},\mathbf{x})} / \frac{\mathcal{A}(\mathbf{w},\mathbf{x})}{\mathcal{G}(\mathbf{w},\mathbf{x})} \le \frac{(a+b)^2}{4ab}.$$

Similar converses are valid for the $\mathcal{A} - \mathcal{G} - \mathcal{H}$ differences.

Theorem 3.2 Let $\mathbf{x} \in [a, b] \subset \mathbb{R}^+$. Then

$$0 \le \mathcal{A}(\mathbf{w},\mathbf{x}) - \mathcal{G}(\mathbf{w},\mathbf{x}) \le (\sqrt{b} - \sqrt{a})^2;$$

$$0 \le \mathcal{A}(\mathbf{w},\mathbf{x}) - \mathcal{H}(\mathbf{w},\mathbf{x}) \le (\sqrt{b} - \sqrt{a})^2;$$

$$0 \le \mathcal{G}(\mathbf{w},\mathbf{x}) - \mathcal{H}(\mathbf{w},\mathbf{x}) \le (\sqrt{b} - \sqrt{a})^2.$$

For example, taking $f(x) = e^x$ and applying Theorem 2.2, we obtain

$$\sum w_i e^{x_i} - e^{\sum w_i x_i} \le e^a + e^b - 2e^{\frac{a+b}{2}} = (e^{b/2} - e^{a/2})^2.$$

Now, change of variable $\mathbf{x} \to \log \mathbf{x}$; $a \to \log a$, $b \to \log b$ gives

$$\mathcal{A}(\mathbf{w},\mathbf{x}) - \mathcal{G}(\mathbf{w},\mathbf{x}) \le (\sqrt{b} - \sqrt{a})^2.$$

Rest of the proof is left to the reader.

Notion of $\mathcal{A} - \mathcal{G} - \mathcal{H}$ means is generalized by the power mean $\mathcal{P}_\alpha$ of order $\alpha \in \mathbb{R}$, defined as

$$\mathcal{P}_\alpha(\mathbf{x},\mathbf{w}) := (\sum w_i x_i^\alpha)^{1/\alpha}.$$

Hence,

$$\mathcal{P}_{-1}(\mathbf{x}, \mathbf{w}) = \mathcal{H}(\mathbf{x}, \mathbf{w}),\ \mathcal{P}_1(\mathbf{x}, \mathbf{w}) = \mathcal{A}(\mathbf{x}, \mathbf{w}),$$

and

$$\mathcal{P}_0(\mathbf{x},\mathbf{w}) = \lim_{\alpha \to 0} \mathcal{P}_\alpha(\mathbf{x},\mathbf{w}) = \mathcal{G}(\mathbf{x},\mathbf{w}).$$

It is well known ([5]) that power means are monotone increasing in α.

We give now an estimation of a difference of power means.

Theorem 3.3 For $0 < \alpha < 1$ and $\mathbf{x} \in [a, b]$, we have

$$0 \le \mathcal{A}(\mathbf{x},\mathbf{w}) - \mathcal{P}_\alpha(\mathbf{x},\mathbf{w}) \le 2(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} / (1-2^{\frac{\alpha-1}{\alpha}})[\frac{a+b}{2} - \left(\frac{a^\alpha + b^\alpha}{2}\right)^{1/\alpha}]. \tag{7}$$

For $\alpha > 1$, we have

$$0 \le \mathcal{P}_\alpha(\mathbf{x},\mathbf{w}) - \mathcal{A}(\mathbf{x},\mathbf{w}) \le 2(\alpha-1)\alpha^{\frac{\alpha}{1-\alpha}} / (2^{\frac{\alpha-1}{\alpha}} - 1)[\left(\frac{a^\alpha + b^\alpha}{2}\right)^{1/\alpha} - \frac{a+b}{2}]. \tag{8}$$

Proof. By Theorem 2.2 and (4), applied to the convex function $f(x) = x^\beta$, $\beta > 1$ with $c \le y_i \le d$, we have

$$0 \le \sum_1^n p_i y_i^\beta - (\sum_1^n p_i y_i)^\beta \le c(x^\beta)[c^\beta + d^\beta - 2(\frac{c+d}{2})^\beta].$$

The change of variable $y_i = x_i^{1/\beta}$ gives $a := c^\beta \le x_i \le d^\beta := b$ and

$$0 \le \sum_1^n p_i x_i - (\sum_1^n p_i x_i^{1/\beta})^\beta \le c(x^\beta)[a + b - 2(\frac{a^{1/\beta} + b^{1/\beta}}{2})^\beta].$$

Finally, the change of variable $\beta = 1/\alpha$, $0 < \alpha < 1$, gives the result.

The second part proof goes analogously, treating the convex function $f(x) = -x^\beta$, $0 < \beta < 1$.

A converse of Ky Fan inequality The most celebrated counterpart of $\mathcal{A} - \mathcal{G}$ inequality is the inequality of Ky Fan which says that

$$\frac{\sum_1^n w_i x_i}{\sum_1^n w_i (1-x_i)} \ge \frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}} \tag{9}$$

whenever $x_i \in (0, 1/2]$.

A converse of Ky Fan inequality is given in [12].

Theorem 3.4 If $0 < a \le x_i \le b \le 1/2$, then

$$\frac{\sum_1^n w_i x_i}{\sum_1^n w_i (1-x_i)} \le S(a,b) \frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}, \tag{10}$$

where

$$S(a,b) = \frac{(1-a)(1-b)(a+b)^2}{ab(2-a-b)^2}.$$

A two-sided improvement of this inequality is obtained by an application of Theorem 2.13.

Theorem 3.5 For $0 < a \le x_i \le b \le 1/2$, we have

$$\exp\left(\frac{1/2-b}{(b(1-b))^2}[\sum w_i x_i^2 - (\sum w_i x_i)^2]\right) \frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}$$

$$\le \frac{\sum_1^n w_i x_i}{\sum_1^n w_i (1-x_i)} \le$$

$$\exp\left(\frac{1/2-a}{(a(1-a))^2}[\sum w_i x_i^2-(\sum w_i x_i)^2]\right)\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}.$$

Proof. Let $f(x)=\log(\frac{1-x}{x})$. Since $f''(x)=\frac{1-2x}{(x(1-x))^2}$ and this function is decreasing on $E=(0, 1/2]$, we found that $m_f(E)=\frac{1-2b}{(b(1-b))^2}$, $M_f(E)=\frac{1-2a}{(a(1-a))^2}$.

Therefore, applying Theorem 2.9 we get

$$\frac{1}{2}m_f(E)J_n(\mathbf{p},\mathbf{x};x^2)\le\sum w_i\log\left(\frac{1-x_i}{x_i}\right)-\log\left(\frac{1-\sum w_i x_i}{\sum w_i x_i}\right)$$

$$=\log\left(\frac{\sum w_i x_i}{\sum w_i(1-x_i)}\right)-\log\left(\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}\right)\le\frac{1}{2}M_f(E)J_n(\mathbf{p},\mathbf{x};x^2),$$

and the proof follows.

It is of interest to find a form of Ky Fan inequality for $\mathbf{x}\in(0, 1)$. We shall give now two results of this kind in the special case $\mathbf{x}\in E:=[a, 1-a]$, $0<a<1/2$.

Theorem 3.6 If $\mathbf{x}\in E:=[a, 1-a]$, $0<a<1/2$, then

$$\frac{1}{T_n(a;\mathbf{w},\mathbf{x})}\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}\le\frac{\sum_1^n w_i x_i}{\sum_1^n w_i(1-x_i)}\le T_n(a;\mathbf{w},\mathbf{x})\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}, \tag{11}$$

where

$$T_n(a;\mathbf{w},\mathbf{x})=\exp\Big[\frac{1-2a}{2(a(1-a))^2}J_n(\mathbf{w},\mathbf{x};x^2)\Big].$$

Proof. Analogously to the previous reason, for $f(x)=\log(\frac{1-x}{x})$ we have $M_f(E)=\frac{1-2a}{(a(1-a))^2}=-m_f(E)$ and the proof is obtained by Theorem 2.9. Note that the function f is neither convex nor concave in this case.

Corollary 3.7 A weaker but more explicit variant of the above assertion is given in the next

Theorem 3.8 If $\mathbf{x}\in E:=[a, 1-a]$, $0<a<1/2$, then

$$\exp\Big[\frac{-(1-2a)^3}{8(a(1-a))^2}\Big]\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}\le\frac{\sum_1^n w_i x_i}{\sum_1^n w_i(1-x_i)}\le\exp\Big[\frac{(1-2a)^3}{8(a(1-a))^2}\Big]\frac{\prod_1^n x_i^{w_i}}{\prod_1^n (1-x_i)^{w_i}}.$$

Proof. Since $c(x^2)=1/2$, we obtain

$$J_n(\mathbf{w},\mathbf{x};x^2)\le\frac{1}{2}[a^2+b^2-2(\frac{a+b}{2})^2]=\frac{1}{4}(b-a)^2=\frac{1}{4}(1-2a)^2,$$

and the result follows from Theorem 3.6.

Applications in Probability Theory The Jensen's inequality has a great influence in Probability and Statistics. Here are some basic definitions.

If the generator of random variable X is discrete with probability mass function $x_1\to p_1, x_2\to p_2, ..., x_n\to p_n$, then the expected value EX is defined as

$$EX := \sum_1^n p_i x_i,$$

and the *variance* $Var(X)$ is

$$Var(X) := \sum_1^n p_i x_i^2 - (\sum_1^n p_i x_i)^2 = E(X^2) - (EX)^2 = E(X - EX)^2.$$

Also, the *moment of s-th order* is defined by

$$EX^s := \sum_1^n p_i x_i^s, \ s > 0.$$

Jensen's moment inequality says that

$$EX^s \geq (EX)^s, s > 1;$$

and

$$EX^s \leq (EX)^s, 0 < s < 1.$$

These inequalities follow from the Jensen's inequality applied to the convex functions $f(x) = -x^s$, $0 < s < 1$ and $f(x) = x^s$, $s > 1$. For example $Var(X) \geq 0$.

Our task in the sequel is to improve Jensen's moment inequality by an application of the results from this paper.

Theorem 3.9 For $a \leq X \leq b$, we have

$$\frac{1}{2}s(s-1)a^{s-2}Var(X) \leq E(X^s) - (EX)^s \leq \frac{1}{2}s(s-1)b^{s-2}Var(X), \ s > 2; \tag{12}$$

$$\frac{1}{2}s(s-1)b^{s-2}Var(X) \leq E(X^s) - (EX)^s \leq \frac{1}{2}s(s-1)a^{s-2}Var(X), \ 1 < s < 2; \tag{13}$$

$$\frac{1}{2}s(1-s)b^{s-2}Var(X) \leq (EX)^s - E(X^s) \leq \frac{1}{2}s(1-s)a^{s-2}Var(X), \ 0 < s < 1. \tag{14}$$

Proof. The proof follows by an application of Theorem 2.9.

Theorem 3.10 For $a \leq X \leq b$, we have

$$0 \leq (EX)^s - E(X^s) \leq (s-1)s^{s/(1-s)} / (1-2^{1-s})[a^s + b^s - 2(\frac{a+b}{2})^s], \ s > 1; \tag{15}$$

$$0 \leq E(X^s) - (EX)^s \leq (1-s)s^{s/(1-s)} / (2^{1-s} - 1)[2(\frac{a+b}{2})^s - (a^s + b^s)], \ 0 < s < 1. \tag{16}$$

Proof. Applying (4) and the result from Theorem 2.6 we obtain the proof.

Remark 3.11 Comparison of Theorem 3.9 and Theorem 3.10 is interesting. Although the left-hand side of Theorem 3.9 is evidently better than the left-hand side of Theorem 3.10, what can be said about their right-hand sides?

4. Conclusion

The celebrated Jensen's inequality for convex functions is applicable in many parts of Analysis, Probability and Statistics, Information Theory etc. Some important inequalities such as Cauchy's inequality, Hölder's inequality, Minkowski's inequality, Ky Fan inequality and Jensen-Mercer inequality are just special cases of Jensen's inequality.

In this article, we give several improvements and reverses of Jensen's and Jensen-Mercer inequalities. We also consider the form of these inequalities for twice differentiable functions which are not necessarily convex/concave on a given closed interval.

Finally, we demonstrate some applications of our results in Theory of Means and Probability Theory.

Acknowledgement

The author is grateful to the referees for their valuable comments.

References

[1] Dragomir SS. Some reverses of the Jensen inequality for functions of self-adjoint operators in Hilbert spaces. *Journal of Inequalities and Applications*. 2010; 15. Available from: https://doi.org/10.1155/2010/496821.

[2] Simic S. Another converse of Jensen's inequality. *RGMIA*. 2009; 12(2): 1-5. Available from: http://vuir.vu.edu.au/17742/1/simic.pdf.

[3] Ragusa MA, Tachikawa A. Partial regularity of the minimizers of quadratic functionals with VMO coefficients. *Journal of the London Mathematical Society*. 2005; 72: 609-620.

[4] Ragusa MA, Tachikawa A. Regularity for minimizers for functionals of double phase with variable exponents. *Advances in Nonlinear Analysis*. 2020; 710-728. Available from: doi: 10.1515/anona-2020-0022.

[5] Hardy GH, Littlewood JE, Polya G. *Inequalities*. Cambridge: Cambridge University Press; 1978.

[6] Mercer A. A variant of Jensen's inequality. *Journal of Inequalities in Pure and Applied Mathematics*. 2003; 4(4): Artcle ID 73. Available from: http://emis.maths.tcd.ie/journals/JIPAM/images/116_03/116_03.pdf.

[7] Simić S. On a converse of Jensen's discrete inequality. *Journal of Inequalities and Applications*. 2009; 6. Available from: https://doi.org/10.1155/2009/153080.

[8] Simić S. On a convexity property. *Kragujevac Journal of Mathematics*. 2016; 40(2): 166-171.

[9] Kian M. A characterization of mean values for Csiszars inequality and applications. *Indagationes Mathematicae*. 2014; 25: 505-515.

[10] Bingham NH, Goldie CM, Teugels JL. *Regular Variation*. Cambridge: Cambridge University Press; 1987.

[11] Anderson GD, Vamanamurthy MK, Vuorinen M. *Conformal Invariants, Inequalities and Quasiconformal Maps*. New York: John Wiley & Sons; 1997.

[12] Simić S. On a converse of Ky Fan inequality. *Kragujevac Journal of Mathematics*. 2010; 33: 95-99.