



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Refinement of the Jensen integral inequality

This is the Published version of the following publication

Dragomir, Sever S, Khan, MA and Abathun, A (2016) Refinement of the Jensen integral inequality. *Open Mathematics*, 14 (1). 221 - 228.

The publisher's official version can be found at
<https://www.degruyter.com/view/j/math.2016.14.issue-1/math-2016-0020/math-2016-0020.xml>

Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/33400/>

Research Article

Silvestru Sever Dragomir*, Muhammad Adil Khan, and Addisalem Abathun

Refinement of the Jensen integral inequality

DOI 10.1515/math-2016-0020

Received September 7, 2015; accepted December 28, 2015.

Abstract: In this paper we give a refinement of Jensen's integral inequality and its generalization for linear functionals. We also present some applications in Information Theory.

Keywords: Convex functions, Jensen's inequality, f -divergences

MSC: 26D15, 94A17

1 Introduction

Let C be a convex subset of the linear space X and f be a convex function on C . If $\mathbf{p} = (p_1, \dots, p_n)$ is probability sequence and $\mathbf{x} = (x_1, \dots, x_n) \in C^n$, then

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

is well known in the literature as Jensen's inequality.

The Lebesgue integral version of the Jensen inequality is given below:

Theorem 1.1. Let (Ω, Λ, μ) be a measure space with $0 < \mu(\Omega) < \infty$ and let $\phi : I \rightarrow \mathbb{R}$ be a convex function defined on an open interval I in \mathbb{R} . If $f : \Omega \rightarrow I$ is such that $f, \phi \circ f \in L(\Omega, \Lambda, \mu)$, then

$$\phi\left(\frac{1}{\mu(\Omega)} \int_{\Omega} f d\mu\right) \leq \frac{1}{\mu(\Omega)} \int_{\Omega} \phi(f) d\mu. \quad (2)$$

In case when ϕ is strictly convex on I one has equality in (2) if and only if f is constant almost everywhere on Ω .

The Jensen inequality for convex functions plays a crucial role in the Theory of Inequalities due to the fact that other inequalities such as the arithmetic mean-geometric mean inequality, the Hölder and Minkowski inequalities, the Ky Fan inequality etc. can be obtained as particular cases of it.

There is an extensive literature devoted to Jensen's inequality concerning different generalizations, refinements, counterparts and converse results, see, for example [1–9].

In this paper we give a refinement of Jensen's integral inequality and its generalization for linear functionals. We also present some applications in Information Theory for example for Kullback-Leibler, total variation and Karl Pearson χ^2 -divergences etc.

*Corresponding Author: **Silvestru Sever Dragomir**: School of Computer Science and Mathematics, Victoria University of Technology PO Box 14428, MCMC 8001 Victoria, Australia, E-mail: sever.dragomir@vu.edu.au

Muhammad Adil Khan: Department of Mathematics, University of Peshawar, Peshawar, Pakistan, E-mail: adilswati@gmail.com

Addisalem Abathun: Department of Mathematics, Adis Ababa University, PO Box 1176, Ethiopia, E-mail: addisalem.abathun@aau.edu.et

2 Main results

Let (Ω, Λ, μ) be a measure space with $0 < \mu(\Omega) < \infty$ and $L(\Omega, \Lambda, \mu) = \{f : \Omega \rightarrow \mathbb{R} : f \text{ is } \mu\text{-measurable and } \int_{\Omega} f(t)d\mu(t) < \infty\}$ be a Lebesgue space. Consider the set $\mathfrak{S} = \{\omega \in \Lambda : \mu(\omega) \neq 0 \text{ and } \mu(\bar{\omega}) = \mu(\Omega \setminus \omega) \neq 0\}$ and $\phi : (a, b) \rightarrow \mathbb{R}$ be a convex function defined on an open interval (a, b) . If $f : \Omega \rightarrow (a, b)$ is such that $f, \phi \circ f \in L(\Omega, \Lambda, \mu)$, then for any set $\omega \in \mathfrak{S}$, define the functional as

$$F(\phi, f; \omega) = \frac{\mu(\omega)}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\omega)} \int_{\omega} f d\mu \right) + \frac{\mu(\bar{\omega})}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} f d\mu \right). \quad (3)$$

We give the following refinement of Jensen's inequality.

Theorem 2.1. *Let (Ω, Λ, μ) be a measure space with $0 < \mu(\Omega) < \infty$ and let $\phi : (a, b) \rightarrow \mathbb{R}$ be a convex function defined on an open interval (a, b) . If $f : \Omega \rightarrow (a, b)$ is such that $f, \phi \circ f \in L(\Omega, \Lambda, \mu)$, then for any set $\omega \in \mathfrak{S}$ we have*

$$\phi \left(\frac{1}{\mu(\Omega)} \int_{\Omega} f d\mu \right) \leq F(\phi, f; \omega) \leq \frac{1}{\mu(\Omega)} \int_{\Omega} \phi(f) d\mu. \quad (4)$$

Proof. As for any $\omega \in \mathfrak{S}$ we have

$$\phi \left(\frac{1}{\mu(\Omega)} \int_{\Omega} f d\mu \right) = \phi \left[\frac{\mu(\omega)}{\mu(\Omega)} \left(\frac{1}{\mu(\omega)} \int_{\omega} f d\mu \right) + \frac{\mu(\bar{\omega})}{\mu(\Omega)} \left(\frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} f d\mu \right) \right].$$

Therefore by the convexity of the function ϕ we get

$$\phi \left(\frac{1}{\mu(\Omega)} \int_{\Omega} f d\mu \right) \leq \frac{\mu(\omega)}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\omega)} \int_{\omega} f d\mu \right) + \frac{\mu(\bar{\omega})}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} f d\mu \right) = F(\phi, f; \omega), \quad (5)$$

Also for any $\omega \in \mathfrak{S}$ and by the Jensen inequality we have

$$\begin{aligned} F(\phi, f; \omega) &= \frac{\mu(\omega)}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\omega)} \int_{\omega} f d\mu \right) + \frac{\mu(\bar{\omega})}{\mu(\Omega)} \phi \left(\frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} f d\mu \right) \\ &\leq \frac{1}{\mu(\Omega)} \int_{\omega} \phi(f) d\mu + \frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} \phi(f) d\mu \\ &= \frac{1}{\mu(\Omega)} \int_{\Omega} \phi(f) d\mu. \end{aligned} \quad (6)$$

From (5) and (6) we have (4). □

Remark 2.2. *We observe that the inequality (4) can be written in an equivalent form as*

$$\inf_{\omega \in \mathfrak{S}} F(\phi, f; \omega) \geq \phi \left(\frac{1}{\mu(\Omega)} \int_{\Omega} f d\mu \right)$$

and

$$\frac{1}{\mu(\Omega)} \int_{\Omega} \phi(f) d\mu \geq \sup_{\omega \in \mathfrak{S}} F(\phi, f; \omega).$$

Remark 2.3. *If $\emptyset, \Omega \in \mathfrak{S}$ and if we take $\omega = \emptyset$ or $\omega = \Omega$, then we have $F(\phi, f; \omega)$ is equal to the left hand side of (2). In this case (5) holds trivially.*

Particularly Riemann integral version can be given as:

Corollary 2.4. Let $\phi : [a, b] \rightarrow \mathbb{R}$ be a convex function defined on the interval $[a, b]$. If $f : [c, d] \rightarrow [a, b]$, $p : [c, d] \rightarrow \mathbb{R}^+$ are such that f , fp and $(\phi \circ f)p$ are all integrable on $[c, d]$, then we have

$$\inf_{x \in (c, d)} \left[\frac{x-c}{d-c} \phi \left(\frac{1}{x-c} \int_c^x p(t)f(t)dt \right) + \frac{d-x}{d-c} \phi \left(\frac{1}{d-x} \int_x^d p(t)f(t)dt \right) \right] \geq \phi \left(\frac{1}{d-c} \int_c^d p(t)f(t)dt \right),$$

$$\frac{1}{d-c} \int_c^d p(t)\phi(f(t))dt \geq \sup_{x \in [c, d]} \left[\frac{x-c}{d-c} \phi \left(\frac{1}{x-c} \int_c^x p(t)f(t)dt \right) + \frac{d-x}{d-c} \phi \left(\frac{1}{d-x} \int_x^d p(t)f(t)dt \right) \right].$$

As a simple consequence of Theorem 2.1 we can obtain refinement of Hermite-Hadamard inequality:

Corollary 2.5. If $\phi : [a, b] \rightarrow \mathbb{R}$ is a convex function defined on the interval $[a, b]$, then for any $[c, d] \subseteq [a, b]$ we have

$$\phi \left(\frac{d+c}{2} \right) \leq \inf_{x \in [c, d]} \left[\frac{x-c}{d-c} \phi \left(\frac{x+c}{2} \right) + \frac{d-x}{d-c} \phi \left(\frac{d+x}{2} \right) \right],$$

$$\frac{1}{d-c} \int_c^d \phi(t)dt \geq \sup_{x \in [c, d]} \left[\frac{x-c}{d-c} \phi \left(\frac{x+c}{2} \right) + \frac{d-x}{d-c} \phi \left(\frac{d+x}{2} \right) \right].$$

3 Further generalization

Let E be a nonempty set, \mathfrak{A} be an algebra of subsets of E , and L be a linear class of real-valued functions $f : E \rightarrow \mathbb{R}$ having the properties:

- L1 : $f, g \in L \Rightarrow (\alpha f + \beta g) \in L$ for all $\alpha, \beta \in \mathbb{R}$;
- L2 : $\mathbf{1} \in L$, i.e., if $f(t) = 1$ for all $t \in E$, then $f \in L$;
- L3 : $f \in L, E_1 \in \mathfrak{A} \Rightarrow f \cdot \chi_{E_1} \in L$,

where χ_{E_1} is the indicator function of E_1 . It follows from L_2, L_3 that $\chi_{E_1} \in L$ for every $E_1 \in \mathfrak{A}$.

A positive isotonic linear functional $A : L \rightarrow \mathbb{R}$ is a functional satisfying the following properties:

- A1 : $A(\alpha f + \beta g) = \alpha A(f) + \beta A(g)$ for $f, g \in L, \alpha, \beta \in \mathbb{R}$;
- A2 : $f \in L, f(t) \geq 0$ on $E \Rightarrow A(f) \geq 0$;

It follows from L_3 that for every $E_1 \in \mathfrak{A}$ such that $A(\chi_{E_1}) > 0$, the functional A_{E_1} is defined for a fixed positive isotonic linear functional A as $A_{E_1}(f) = \frac{A(f \cdot \chi_{E_1})}{A(\chi_{E_1})}$, for all $f \in L$, with $A(\mathbf{1}) = 1$. Furthermore, we observe that

$$A(\chi_{E_1}) + A(\chi_{E \setminus E_1}) = 1,$$

$$A(f) = A(f \cdot \chi_{E_1}) + A(f \cdot \chi_{E \setminus E_1}). \quad (7)$$

Jessen (see [10, p-47]) gave the following generalization of Jensen's inequality for convex functions.

Theorem 3.1. Let L satisfy L_1 and L_2 on a nonempty set E , and assume that $\phi : [a, b] \rightarrow \mathbb{R}$ be a continuous convex function. If A is linear positive functional with $A(\mathbf{1}) = 1$, then for all $f \in L$ such that $\phi(f) \in L$ we have $A(f) \in [a, b]$ and

$$\phi(A(f)) \leq A(\phi(f)); \quad (8)$$

The following refinement of (8) holds.

Theorem 3.2. Under the above assumptions, if $\phi : [a, b] \rightarrow \mathbb{R}$ is a continuous convex function, then

$$\phi(A(f)) \leq \overline{D}(A, f, \phi; E_1) \leq A(\phi(f)); \quad (9)$$

where

$$\overline{D}(A, f, \phi; E_1) = A(\chi_{E_1})\phi\left(\frac{A(f \cdot \chi_{E_1})}{A(\chi_{E_1})}\right) + A(\chi_{E \setminus E_1})\phi\left(\frac{A(f \cdot \chi_{E \setminus E_1})}{A(\chi_{E \setminus E_1})}\right)$$

for all non empty set $E_1 \in \mathfrak{A}$ such that $0 < A(\chi_{E_1}) < 1$

Proof. Since

$$\begin{aligned} \overline{D}(A, f, \phi; E_1) &= A(\chi_{E_1})\phi\left(\frac{A(f \cdot \chi_{E_1})}{A(\chi_{E_1})}\right) + A(\chi_{E \setminus E_1})\phi\left(\frac{A(f \cdot \chi_{E \setminus E_1})}{A(\chi_{E \setminus E_1})}\right) \\ \text{i. e. } \overline{D}(A, f, \phi; E_1) &= A(\chi_{E_1})\phi(A_{E_1}(f)) + A(\chi_{E \setminus E_1})\phi(A_{E \setminus E_1}(f)). \end{aligned}$$

Using the inequality (8) we obtain

$$\overline{D}(A, f, \phi; E_1) \leq A(\phi(f) \cdot \chi_{E_1}) + A(\phi(f) \cdot \chi_{E \setminus E_1}) = A(\phi(f)). \quad (10)$$

This proves the second inequality in (9).

The first inequality follows by using definition of convex function and identity (7). \square

4 Applications for Csiszár divergence measures

Let (Ω, Λ, μ) be a probability measure space. Consider the set of all density functions on μ to be $S := \{p | p : \Omega \rightarrow \mathbb{R}, p(s) > 0, \int_{\Omega} p(s) d\mu(s) = 1\}$.

Csiszár introduced the concept of f -divergence for a convex function $f : (0, \infty) \rightarrow (-\infty, \infty)$ (cf. [11] , see also [12]) by

$$I_f(q, p) = \int_{\Omega} p(s) f\left(\frac{q(s)}{p(s)}\right) d\mu(s), \quad p, q \in S.$$

By appropriately defining the convex function f , various divergences can be derived. We give some important f -divergences, playing a significant role in Information Theory and Statistics.

(i) The class of χ -divergences: The f -divergences, in this class, are generated by the family of functions

$$\begin{aligned} f_{\alpha}(u) &= |u - 1|^{\alpha} \quad u \geq 0 \text{ and } \alpha \geq 1. \\ I_{f_{\alpha}}(q, p) &= \int_{\Omega} p^{1-\alpha}(s) |q(s) - p(s)|^{\alpha} d\mu(s). \end{aligned}$$

For $\alpha = 1$, it gives the total variation distance.

$$V(q, p) = \int_{\Omega} |q(s) - p(s)| d\mu(s).$$

For $\alpha = 2$, it gives the Karl Pearson χ^2 -divergence,

$$I_{\chi^2}(q, p) = \int_{\Omega} \frac{[q(s) - p(s)]^2}{p(s)} d\mu(s).$$

(ii) α -order Renyi entropy : For $\alpha > 1$ let

$$f(t) = t^{\alpha}, \quad t > 0.$$

Then I_f gives α -order entropy

$$D_\alpha(q, p) = \int_{\Omega} q^\alpha(s) p^{1-\alpha}(s) d\mu(s).$$

(iii) Harmonic distance: Let

$$f(t) = -\frac{2t}{1+t}, \quad t > 0.$$

Then I_f gives Harmonic distance

$$D_H(q, p) = \int_{\Omega} \frac{2 p(s) q(s)}{p(s) + q(s)} d\mu(s).$$

(iv) Kullback-Leibler: Let

$$f(t) = t \log t, \quad t > 0.$$

Then f -divergence functional give rise to Kullback-Leibler distance [13]

$$D_{KL}(q, p) = \int_{\Omega} q(s) \log \left(\frac{q(s)}{p(s)} \right) d\mu(s).$$

One parametric generalization of the Kullback-Leibler [13] relative information was studied in a different way by Cressie and Read [14].

(v) Jeffreys divergence: Let

$$f(t) = (t - 1) \log t, \quad t > 0.$$

Then f -divergence functional give Jeffreys divergence

$$J(q, p) = \int_{\omega} (p(s) - q(s)) \ln \left(\frac{q(s)}{p(s)} \right) d\mu(s).$$

(vi) The Dichotomy class: This class is generated by the family of functions $g_\alpha : (0, \infty) \rightarrow \mathbb{R}$,

$$g_\alpha(u) = \begin{cases} u - 1 - \log u, & \alpha = 0 \\ \frac{1}{\alpha(1-\alpha)} [\alpha u + 1 - \alpha - u^\alpha], & \alpha \in \mathbb{R} \setminus \{0, 1\}; \\ 1 - u + u \log u, & \alpha = 1. \end{cases} \tag{11}$$

This class gives, for particular values of α , some important divergences. For instance, for $\alpha = \frac{1}{2}$ it provides a distance, namely, the Hellinger distance.

There are various other divergences in Information Theory and Statistics such as Arimoto-type divergences, Matushita’s divergence, Puri-Vincze divergences etc. (cf. [15], [16]) used in various problems in Information Theory and statistics. An application of Theorem 1.1 is the following result given by Csiszár and Korner (cf. [17]).

Theorem 4.1. *Let $f : [0, \infty) \rightarrow \mathbb{R}$ be a convex function and p, q be positive functions from S . Then the following inequality is valid,*

$$I_f(q, p) \geq f(1). \tag{12}$$

Theorem 4.2. *Let $f : [0, \infty] \rightarrow \mathbb{R}$ be a convex function, then for any p and q in S we have:*

$$I_f(q, p) \geq \mu(\omega) f \left(\frac{1}{\mu(\omega)} \int_{\omega} q(s) d\mu(s) \right) + \mu(\bar{\omega}) f \left(\frac{1}{\mu(\bar{\omega})} \int_{\bar{\omega}} q(s) d\mu(s) \right) \geq f(1). \tag{13}$$

Proof. By substituting $\phi(s) = f(s)$, $f(s) = \frac{q(s)}{p(s)}$ and $d\mu(s) = p(s)d\mu(s)$ in Theorem 2.1, we deduce (13). \square

Proposition 4.3. Let $p, q \in S$, then we have

$$V(q, p) \geq 2 \sup_{\omega \in \mathfrak{S}} \left| \int_{\omega} q(s) d\mu(s) - \mu(\omega) \right| (\geq 0). \tag{14}$$

Proof. By putting $f(x) = |x - 1|$ for all $x \geq 0$ in Theorem 4.2 we get (14). □

Proposition 4.4. For any $p, q \in S$,

$$I_{\chi^2}(q, p) \geq \sup_{\omega \in \mathfrak{S}} \left\{ \frac{\left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right)^2}{\mu(\omega)(1 - \mu(\omega))} \right\} \geq 4 \sup_{\omega \in \mathfrak{S}} \left\{ \left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right)^2 \right\} (\geq 0). \tag{15}$$

Proof. By making use of the function $f(x) = (t - 1)^2$ in Theorem 4.2 we get

$$\int_{\Omega} p(s) \left(\frac{q(s)}{p(s)} - 1 \right)^2 d\mu(s) \geq \sup_{\omega \in \mathfrak{S}} \left\{ \mu(\omega) \left(\frac{1}{\mu(\omega)} \int_{\omega} q(s) d\mu(s) - 1 \right)^2 + \mu(\bar{\omega}) \left(\frac{1}{\mu(\bar{\omega})} \int_{\omega} q(s) d\mu(s) - 1 \right)^2 \right\} (\geq 0)$$

$$\text{i.e. } \int_{\Omega} \frac{(q(s) - p(s))^2}{p(s)} d\mu(s) \geq \sup_{\omega \in \mathfrak{S}} \left\{ \frac{\left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right)^2}{\mu(\omega)(1 - \mu(\omega))} \right\} (\geq 0).$$

Since by Arithmetic-Geometric mean inequality we have

$$\mu(\omega)(1 - \mu(\omega)) \leq \frac{1}{4} [\mu(\omega) + (1 - \mu(\omega))]^2 = \frac{1}{4},$$

therefore

$$\frac{\left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right)^2}{\mu(\omega)(1 - \mu(\omega))} \geq 4 \left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right)^2 (\geq 0). \tag{□}$$

Proposition 4.5. For any $p, q \in S$, we have:

$$D_{KL}(q, p) \geq \ln \left[\left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right)^{1 - \int_{\omega} q(s) d\mu(s)} \cdot \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right)^{\int_{\omega} q(s) d\mu(s)} \right] (\geq 0). \tag{16}$$

Proof. By putting $f(t) = t \ln(t)$ in Theorem 4.2 one can get first inequality in (16).

To prove the second inequality, we utilize the inequality between the geometric mean and harmonic mean,

$$x^{\alpha} y^{1-\alpha} \geq \frac{1}{\frac{\alpha}{x} + \frac{1-\alpha}{y}}, \quad x, y, \alpha \in [0, 1],$$

we have for

$$x = \frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)}, y = \frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \text{ and } \alpha = \int_{\omega} q(s) d\mu(s) \quad \text{that}$$

$$\left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right)^{1 - \int_{\omega} q(s) d\mu(s)} \cdot \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right)^{\int_{\omega} q(s) d\mu(s)} \geq 1,$$

for any $\omega \in \mathfrak{S}$, which implies the second inequality in (16). □

Proposition 4.6. For any $p, q \in S$, we have:

$$\begin{aligned}
 J(q, p) &\geq \ln \left(\sup_{\omega \in \mathfrak{S}} \left\{ \left[\frac{(1 - \int_{\omega} q(s) d\mu(s)) \mu(\omega)}{(1 - \mu(\omega)) \int_{\omega} q(s) d\mu(s)} \right]^{(\mu(\omega) - \int_{\omega} q(s) d\mu(s))} \right\} \right) \\
 &\geq \sup_{\omega \in \mathfrak{S}} \left(\frac{(\mu(\omega) - \int_{\omega} q(s) d\mu(s))^2}{\int_{\omega} q(s) d\mu(s) + \mu(\omega) - 2 \int_{\omega} q(s) d\mu(s) \mu(\omega)} \right) \geq 0.
 \end{aligned} \tag{17}$$

Proof. By putting $f(x) = (x - 1) \ln(x)$, $x > 0$ in Theorem 4.2 we have

$$\begin{aligned}
 \int_{\omega} p(s) \left(\frac{q(s)}{p(s)} - 1 \right) \ln \left(\frac{q(s)}{p(s)} \right) d\mu(s) &\geq \sup_{\omega \in \mathfrak{S}} \left(\mu(\omega) \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} - 1 \right) \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right) \right. \\
 &\quad \left. + \mu(\bar{\omega}) \left(\frac{\int_{\bar{\omega}} q(s) d\mu(s)}{\mu(\bar{\omega})} - 1 \right) \ln \left(\frac{\int_{\bar{\omega}} q(s) d\mu(s)}{\mu(\bar{\omega})} \right) \right) \\
 &= \sup_{\omega \in \mathfrak{S}} \left(\left(\int_{\omega} q(s) d\mu(s) - \mu(\omega) \right) \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right) \right. \\
 &\quad \left. + \left(\int_{\bar{\omega}} q(s) d\mu(s) - \mu(\bar{\omega}) \right) \ln \left(\frac{\int_{\bar{\omega}} q(s) d\mu(s)}{\mu(\bar{\omega})} \right) \right)
 \end{aligned}$$

that is

$$\begin{aligned}
 J(q, p) &\geq \sup_{\omega \in \mathfrak{S}} \left(\left(\mu(\omega) - \int_{\omega} q(s) d\mu(s) \right) \ln \left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right) \right. \\
 &\quad \left. - \left(\mu(\omega) - \int_{\omega} q(s) d\mu(s) \right) \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right) \right)
 \end{aligned}$$

proving the first inequality in (17).

Utilizing the elementary inequality for positive numbers ,

$$\frac{\ln b - \ln a}{b - a} \geq \frac{2}{a + b}, \quad a, b > 0$$

we have

$$\begin{aligned}
 &\left(\mu(\omega) - \int_{\omega} q(s) d\mu(s) \right) \left[\ln \left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right) - \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right) \right] \\
 &= \left(\mu(\omega) - \int_{\omega} q(s) d\mu(s) \right) \frac{\ln \left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right) - \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right)}{\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} - \frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)}} \\
 &\quad \times \left[\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} - \frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right] \\
 &= \frac{(\mu(\omega) - \int_{\omega} q(s) d\mu(s))^2}{\mu(\omega)(1 - \mu(\omega))} \cdot \frac{\ln \left(\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} \right) - \ln \left(\frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)} \right)}{\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} - \frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)}} \\
 &\geq \frac{(\mu(\omega) - \int_{\omega} q(s) d\mu(s))^2}{\mu(\omega)(1 - \mu(\omega))} \cdot \frac{2}{\frac{1 - \int_{\omega} q(s) d\mu(s)}{1 - \mu(\omega)} + \frac{\int_{\omega} q(s) d\mu(s)}{\mu(\omega)}} \\
 &= \frac{2(\mu(\omega) - \int_{\omega} q(s) d\mu(s))^2}{\int_{\omega} q(s) d\mu(s) + \mu(\omega) - 2 \int_{\omega} q(s) d\mu(s) \mu(\omega)} \geq 0,
 \end{aligned}$$

for each $\omega \in \mathfrak{S}$, giving the second inequality in (17). □

Proposition 4.7. For any $p, q \in S$, we have:

$$D_\alpha(q, p) \geq \sup_{\omega \in \mathfrak{S}} \left[(\mu(\omega))^{1-\alpha} \left(\int_{\omega} q(s) d\mu(s) \right)^\alpha + (1 - \mu(\omega))^{1-\alpha} \left(1 - \int_{\omega} q(s) d\mu(s) \right)^\alpha \right] \geq 1. \quad (18)$$

Proof. By putting $f(x) = x^\alpha$ for $\alpha > 1$, $x > 0$, in Theorem 4.2 we get the required inequalities. \square

Acknowledgement: The authors express their sincere thanks to the referees for their careful reading of the manuscript and very helpful suggestions that improved the manuscript.

References

- [1] Adil Khan M., Anwar M., Jakšetić J., and Pečarić J., *On some improvements of the Jensen inequality with some applications*, J. Inequal. Appl. **2009** (2009), Article ID 323615, 15 pages.
- [2] Adil Khan M., Khan G. A., Ali T., Batbold T., and Kilicman A., *Further refinement of Jensen's type inequalities for the function defined on the rectangle*, Abstr. Appl. Anal. **2013** (2013), Article ID 214123, 1-8.
- [3] Adil Khan M., Khan G. A., Ali T., and Kilicman A., *On the refinement of Jensen's inequality*, Appl. Math. Comput. **262**(1) (2015), 128-135.
- [4] Beesack P. R. and Pečarić J., *On Jessen's inequality for convex functions*, J. Math. Anal. Appl., **110** (1985), 536-552.
- [5] Dragomir S. S., *A refinement of Jensen's inequality with applications for f -divergence measures*, Taiwanese J. Math., **14**(1) (2010), 153-164.
- [6] Dragomir S. S., *A new refinement of Jensen's inequality in linear spaces with applications*, Math. Comput. Modelling, **52** (2010), 1497-1505.
- [7] Dragomir S. S., *Some refinements of Jensen's inequality*, J. Math. Anal. Appl., **168**(2) (1992), 518-522.
- [8] Dragomir S. S., *A further improvement of Jensen's inequality*, Tamkang J. Math., **25**(1) (1994), 29-36.
- [9] Micić-Hot J., Pečarić J. and Jurica P., *Refined Jensen's operator inequality with condition on spectra*, Oper. Matrices, **7**(2) (2013), 293-308.
- [10] Pečarić J., Proschan F. and Tong Y. L., *Convex functions, Partial Orderings and Statistical Applications*, Academic Press, New York, 1992.
- [11] Csiszár I., *Information measures, Acritical survey*, Trans. 7th Prague Conf. on Info. Th., Volume B, Academia Prague, (1978), 73-86.
- [12] Pardo M. C. and Vajda I., *On asymptotic properties of information-theoretic divergences*, IEEE Trans. Inform. Theory, **49**(3) (2003), 1860-1868.
- [13] Kullback S. and Leibler R. A., *On information and sufficiency*, Ann. Math. Statist., **22** (1951), 79-86.
- [14] Cressie P. and Read T. R. C., *Multinomial goodness-of-fit tests*, J. Roy. Statist. Soc. Ser. B, **46** (1984), 440-464.
- [15] Kafka P., Österreicher F. and Vincze I., *On powers of f -divergences defining a distance*, Studia Sci. Math. Hungar. **26**(4) (1991), 415-422.
- [16] Liese F. and Vajda I., *Convex statistical distances*. With German, Teubner-Texte Zur Mathematik [Teubner Texts in mathematics], 95. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1987.
- [17] Csiszár I. and Korner J., *Information theory: Coding Theorem for Discrete Memoryless systems*, Academic Press, New York, (1981).