Received 25/06/15 Revised 03/09/15

REFINEMENT OF THE JENSEN INTEGRAL INEQUALITY

S. S. DRAGOMIR[1], M. ADIL KHAN[2], AND A. ABATHUN[3]

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.

1. INTRODUCTION

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

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

is well known in the literature as Jensen's inequality.
The Lebesgue intergral version of the Jensen's inequality is given below:

Theorem 1. *Let (Ω, Λ, μ) be a measure space with $0 < \mu(\Omega) < \infty$ and let $\phi : (a, b) \to \mathbb{R}$ be a convex function defined on an open interval (a, b). If $f : \Omega \to (a, b)$ is such that $f, \phi \circ f \in L(\Omega, \Lambda, \mu)$, then*

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

In case when ϕ is strictly convex on (a, b) 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.

2010 *Mathematics Subject Classification.* Primary 26D15; Secondary 94A17 .
Key words and phrases. Convex functions, Jensen's inequality, f- divergences.

In [8] Dragomir gave a refinement of Jensen's inequality:

$$\begin{aligned} f\left(\sum_{i=1}^{n} p_i x_i\right) &\leq \min_{j\in\{1,2,..,n\}} \left[(1-p_j) f\left(\frac{\sum_{i=1}^{n} p_i f(x_i) - p_j x_j}{1-p_j}\right) + p_j f(x_j)\right] \\ &\leq \frac{1}{n}\left[(1-p_j) f\left(\frac{\sum_{i=1}^{n} p_i f(x_i) - p_j x_j}{1-p_j}\right) + p_j f(x_j)\right] \\ &\leq \max_{j\in\{1,2,..,n\}} \left[(1-p_j) f\left(\frac{\sum_{i=1}^{n} p_i f(x_i) - p_j x_j}{1-p_j}\right) + p_j f(x_j)\right] \\ &\leq \sum_{i=1}^{n} p_i f(x_i) \end{aligned} \tag{3}$$

where f, x_i and p_i are as above.

In [9] the author has presented another refinement of Jensen's inequality:

$$\sum_{i=1}^{n} p_i f(x_i) \geq D(f, \mathbf{p}, \mathbf{x}, I) \geq f\Big(\sum_{i=1}^{n} p_i x_i\Big). \tag{4}$$

where I is any non-empty subset of $\{1, 2, \ldots, n\}$,

$$\bar{I} := \{1, 2, \ldots, n\} \setminus I,\ P_I = \sum_{i\in I} p_i,\ P_{\bar{I}} = \sum_{i\in \bar{I}} p_i,\ f,\ x_i \text{ and } p_i$$

are as above and

$$D(f, \mathbf{p}, \mathbf{x}, I) = P_I f\Big(\frac{1}{P_I}\sum_{i\in I} p_i x_i\Big) + P_{\bar{I}} f\Big(\frac{1}{P_{\bar{I}}}\sum_{i\in \bar{I}} p_i x_i\Big).$$

If $I = \{j\}$, then (4) implies the inequalities in (3).

There is an extensive literature devoted to Jensen's inequality concerning different generalizations, refinements, counterparts and converse results, see, for example [1]-[4], [10]-[12].

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.

2. Main results

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

$$F(\phi, f; \omega) = \frac{\mu(\omega)}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\omega)}\int_\omega f d\mu\right) + \frac{\mu(\overline{\omega})}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\overline{\omega})}\int_{\overline{\omega}} f d\mu\right). \tag{5}$$

We give the following refinement of Jensen's inequality.

Theorem 2. *Let* (Ω, Λ, μ) *be a measure space with* $0 < \mu(\Omega) < \infty$ *and let* $\phi : (a, b) \to \mathbb{R}$ *be a convex function defined on an open interval* (a, b)*. If* $f : \Omega \to (a, b)$ *is such that* $f, \phi \circ f \in L(\Omega, \Lambda, \mu)$*, then for any set* $\omega \in \mathfrak{S}$ *we have*

$$\phi\left(\frac{1}{\mu(\Omega)}\int_\Omega f d\mu\right) \leq \digamma(\phi, f; \omega) \leq \frac{1}{\mu(\Omega)}\int_\Omega \phi(f) d\mu. \tag{6}$$

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

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

Therefore by the convexity of the function ϕ we get

$$\begin{aligned}\phi\left(\frac{1}{\mu(\Omega)}\int_\Omega f d\mu\right) &\leq \frac{\mu(\omega)}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\omega)}\int_\omega f d\mu\right) + \frac{\mu(\overline{\omega})}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\overline{\omega})}\int_{\overline{\omega}} f d\mu\right)\\ &= \digamma(\phi, f; \omega),\end{aligned} \tag{7}$$

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

$$\begin{aligned}\digamma(\phi, f; \omega) &= \frac{\mu(\omega)}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\omega)}\int_\omega f d\mu\right) + \frac{\mu(\overline{\omega})}{\mu(\Omega)}\phi\left(\frac{1}{\mu(\overline{\omega})}\int_{\overline{\omega}} f d\mu\right)\\ &\leq \frac{1}{\mu(\Omega)}\int_\omega \phi(f) d\mu + \frac{1}{\mu(\Omega)}\int_{\overline{\omega}} \phi(f) d\mu\\ &= \frac{1}{\mu(\Omega)}\int_\Omega \phi(f) d\mu.\end{aligned} \tag{8}$$

From (7) and (8) we have (6). □

Remark 1. We observe that the inequality (6) can be written in an equivalent form as

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

and

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

Particularly Riemann integral version can be given as:

Corollary 1. *Let* $\phi : (a, b) \to \mathbb{R}$ *be a convex function defined on the interval* (a, b)*. If* $f : (c, d) \to (a, b), p : (c, d) \to \mathbb{R}^+$ *are such that* f, fp *and* $(\phi \circ f)p$ *are all integrable on* (c, d)*, then we have*

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

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

As a simple consequence of Theorem 2 we can obtain refinement of Hermite-Hadamard inequality:

Corollary 2. *If $\phi:(a,b)\to\mathbb{R}$ is a convex function defined on the interval (a,b), then for any $(c,d)\subseteq(a,b)$ we have*

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

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

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 $f:E\to\mathbb{R}$ 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 : $f\in L$, $E_1\in A\Rightarrow f.\chi_{E_1}\in L$,

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 postive isotonic linear functional $A:L\to\mathbb{R}$ 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_1 defined for all $f\in L$ as $A_1(f)=\frac{A(f.\chi_{E_1})}{A(\chi_{E_1})}$ is an isotonic linear functional with $A(\mathbf{1})=1$. Furthermore, we observe that

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

Jessen (see [17, p-47]) gave the following generalization of Jensen's inequality for convex functions.

Theorem 3. *Let L satisfy L_1, L_2 on a nonempty set E , and assume that $\phi:[a,b]\to\mathbb{R}$ be a continuous 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*

$$\phi(A(f))\leq A(\phi(f)); \tag{9}$$

The following refinement of (9) holds.

Theorem 4. *Under the above assumptions, if* $\phi : [a,b] \to \mathbb{R}$ *is a continuous convex function, then*

$$\phi(A(f)) \leq \overline{D}(A,f,\phi;E_1) \leq A(\phi(f)); \tag{10}$$

where

$$\overline{D}(A,f,\phi;E_1) = A(\chi_{E_1})\phi\left(\frac{A(f.\chi_{E_1})}{A(\chi_{E_1})}\right) + A(\chi_{E\setminus E_1})\phi\left(\frac{A(f.\chi_{E\setminus E_1})}{A(\chi_{E\setminus E_1})}\right) \tag{11}$$

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

Proof. The first inequality follows by using definition of convex function and the second follows by using (9) for $A_1(f)$ instead of $A(f)$
. □

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 $S := \{p | p : \Omega \to \mathbb{R}, p(s) > 0, \int_\Omega p(s) d\mu(s) = 1\}$.

Csiszár introduced the concept of $f-$divergence for a convex function $f : (0,\infty) \to (-\infty,\infty)$ (cf. [5] , see also [16]) by

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

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

$$f_\alpha(u) = \mid u-1 \mid^\alpha \quad u \geq 0 \ \ and \ \ \alpha \geq 1.$$

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

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

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

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

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

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

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

Then f-divergence functional give rise to Kullback-Leibler distance [14]

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

The one parametric generalization of the Kullback-Leibler [14] relative information studied in a different way by Cressie and Read [7].
(v) The Dichotomy class: This class is generated by the family of functions $g_\alpha : (0,\infty) \to \mathbb{R}$,

$$g_\alpha(u) = \begin{cases} u - 1 - \log\,u, & \alpha = 0 \\ \frac{1}{\alpha(1-\alpha)}\,[\,\alpha\,u + 1 - \alpha - u^\alpha\,], & \alpha \in \mathbb{R}\backslash\{0,1\}; \\ 1 - u + u\,\log\,u, & \alpha = 1. \end{cases} \tag{12}$$

This class gives, for particular values of α, some important divergences. For instance, for $\alpha = \frac{1}{2}$ provide 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. [13], [15]) used in various problems in Information Theory and statistics. An application of Theorem 1 is the following result given by Csiszár and Korner (cf. [6]).

Theorem 5. *Let $f : [0,\infty) \to \mathbb{R}$ be a convex function and p, q be positive functions from S. Then the following inequality is valid,*

$$I_f(q,p) \geq f(1). \tag{13}$$

Theorem 6. *Let $f : [0,\infty] \to \mathbb{R}$ be a n convex function, then for any p and $q \in S$ we have:*

$$I_f(q,p) \geq \mu(\omega) f\left(\frac{1}{\mu(\omega)}\int_\omega q(s)d\mu(s)\right) + \mu(\overline{\omega}) f\left(\frac{1}{\mu(\overline{\omega})}\int_{\overline{\omega}} q(s)d\mu(s)\right) \geq f(1). \tag{14}$$

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, we deduce (14). □

Proposition 1. *Let $p, q \in S$, then we have*

$$V(q,p) \geq 2\min_{\omega\in\mathfrak{S}}\left|\int_\omega q(s)d\mu(s) - \mu(\omega)\right|(\geq 0). \tag{15}$$

Proof. By putting $f(x) = |x-1|$ for all $x \geq 0$ in Theorem 6 we get (15). □

Proposition 2. *For any p, $q \in S$,*

$$\chi^2(q,p) \geq \max_{\omega \in \mathfrak{S}} \left\{ \frac{\left(\int_\omega q(s)d\mu(s) - \mu(\omega)\right)^2}{\mu(\omega)(1-\mu(\omega))} \right\}$$

$$\geq 4 \max_{\omega \in \mathfrak{S}} \left\{ \left(\int_\omega q(s)d\mu(s) - \mu(\omega)\right)^2 \right\} (\geq 0). \tag{16}$$

Proof. By making use of the function $f(x) = (t-1)^2$ in Theorem 6 we get

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

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

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

Since

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

therefore

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

□

Proposition 3. *For any p, $q \in S$, we have:*

$$KL(q,p) \geq \ln\left[\left(\frac{1 - \int_\omega q(s)d\mu(s)}{1-\mu(\omega)} \right)^{1-\int_\omega q(s)d\mu(s)} \cdot \left(\frac{\int_\omega q(s)d\mu(s)}{\mu(\omega)} \right)^{\int_\omega q(s)d\mu(s)} \right] (\geq 0). \tag{17}$$

Proof. By putting $f(t) = t\ln(t)$ in Theorem 6 one can get first inequality in (17).

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

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

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 (17). □

Proposition 4. *For any p, $q \in S$, we have:*

$$\begin{aligned} J(q,p) &\geq \ln\left(\max_{\omega\in\mathfrak{S}}\left\{\left[\frac{(1-\int_\omega q(s)d\mu(s))\mu(\omega)}{(1-\mu(\omega))\int_\omega q(s)d\mu(s)}\right]^{(\mu(\omega)-\int_\omega q(s)d\mu(s))}\right\}\right) \\ &\geq \max_{\omega\in\mathfrak{S}}\left(\frac{(\mu(\omega)-\int_\omega q(s)d\mu(s))^2}{\int_\omega q(s)d\mu(s)+\mu(\omega)-2\int_\omega q(s)d\mu(s)\mu(\omega)}\right) \geq 0. \end{aligned} \tag{18}$$

Proof. By putting $f(x)=(x-1)\ln(x)$, $x>0$ in Theorem 6 we have

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

that is

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

proving the first inequality in (18).
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

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

for each $\omega\in\mathfrak{S}$, giving the second inequality in (18). □

Proposition 5. *For any p, $q\in S$, we have:*

$$
\begin{aligned}
&D_\alpha(q,p)\\
&\geq\max_{\omega\in\mathfrak{S}}\left[(\mu(\omega))^{1-\alpha}\left(\int_\omega q(s)d\mu(s)\right)^\alpha+(1-\mu(\omega))^{1-\alpha}\left(1-\int_\omega q(s)d\mu(s)\right)^\alpha\right]\geq 1.
\end{aligned}
\tag{19}
$$

Proof. By putting $f(x)=x^\alpha$ for $\alpha>1$, $x>0$, in Theorem 6 we get the required inequalities. □

References

[1] M. Adil Khan, M. Anwar, J. Jakšetić and J. Pečarić, *On some improvements of the Jensen inequality with some applications*, J. Inequal. Appl. **2009** (2009), Article ID 323615, 15 pages.

[2] M. Adil. Khan, G. A. Khan, T. Ali, T. Batbold and A. Kilicman, *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] M. Adil. Khan, G. A. Khan, T. Ali and A. Kilicman, *On the refinement of Jensens inequality*, Appl. Math. Comput. **262**(1) (2015), 128-135.

[4] P. R. Beesack, Pečarić, *On Jessen's inequality for convex functions*, J. Math. Anal. Appl., **110** (1985), 536-552.

[5] I. Csiszár, Information measures, Acritical survey, *Trans. 7th Prague Conf. on Info. Th.*, Volume B, Academia Prague, (1978), 73-86.

[6] I. Csiszár and J. Korner, *Information theory: Coding Theorem for Dicsrete Memoryless systems,* Academic Press, New York, (1981).

[7] P. Cressie and T. R. C. Read*, Multinomial goodness-of-fit tests*, J. Roy. Statist. Soc. Ser. B, **46** (1984), 440-464.

[8] S. S. Dragomir, *A refinement of Jensen's inequality with applications for f-divergence measures*, Taiwanese J. Math., **14**(1) (2010), 153-164.
[9] S. S. Dragomir, *A new refinement of Jensen's inequality in linear spaces with applications*, Math. Comput. Modelling, **52** (2010), 1497-1505.
[10] S. S. Dragomir, *Some refinements of Jensen's inequality*, J. Math. Anal. Appl., **168**(2) (1992), 518–522.
[11] S. S. Dragomir, *A further improvement of Jensen's inequality*, Tamkang J. Math., 25(1) (1994), 29–36.
[12] J. Micić-Hot, J. Pečarić, P. Jurica, *Refined Jensen's operator inequality with condition on spectra*, Oper. Matrices, **7**(2) (2013), 293-308.
[13] P. Kafka, F. Österreicher and I. Vincze, *On powers of f-divergences defining a distance*, Studia Sci. Math. Hungar. **26**(4) (1991), 415-422.
[14] S. Kullback and R. A. Leiber, *On information and sufficency*, Ann. Math. Statist., **22** (1951), 79-86.
[15] F. Liese and I. Vajda, Convex statistical distances. With German, Teubner-Texte Zur Mathematika [Teubner Texts in mathematics], 95. BSB B. G. Teubner Verlagsgesellschaff, Leipzig, 1987.
[16] M. C. Pardo and I. Vajda, *On asymptotic properties of information-theoretic divergences*, IEEE Trans. Inform. Theory, **49**(3) (2003), 1860-1868.
[17] J. Pečarić, F. Proschan and Y. L. Tong, Convex functions, Partial Orderings and Statistical Applications, *Academic Press, New York*, 1992.

3 - School of Computer Science and Mathematics, Victoria University of Technology PO Box 14428, MCMC 8001 Victoria, Australia,
E-mail address: sever.dragomir@vu.edu.au

2 - Department of Mathematics, University of Peshawar, Peshawar, Pakistan
E-mail address: adilswati@gmail.com

3- department of mathematics Adis Ababa University pobox 1176, Ethiopia
E-mail address: addisalem.abathun@aau.edu.et