A Double-Weighted Refinement of Jensen's Inequality With Applications

Jamal Rooin
Department of Mathematics, Institute for Advanced Studies in Basic Sciences,
P.O. Box 45195-1159, Zanjan, Iran.
rooin@iasbs.ac.ir

Ali Morassaei
Department of Mathematics, Faculty of Sciences, University of Zanjan,
P.O. Box 45195-313, Zanjan, Iran
and
Education Office of Zanjan, Zanjan, Iran.
morassae@iasbs.ac.ir

Abstract

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two probability measure spaces, I an interval of the real line, $f \in L^1(\mu)$, $f(x) \in I$ for each $x \in X$, and φ a real-valued convex function on I. First we show that, if ω_0 and ω_1 are two appropriate weight functions on $X \times Y$, then

$$\varphi\left(\int_X f d\mu\right) \leq \int_Y A\left(\varphi; F_0(y), F_1(y)\right) d\lambda(y) \leq \int_X (\varphi \circ f) d\mu,$$

where A denotes the arithmetic mean of φ on the closed interval with end points $F_0(y)$ and $F_1(y)$, and for λ-almost all $y \in Y$'s

$$F_k(y) = \int_X f(x)\omega_k(x, y) d\mu(x) \qquad (k = 0, 1).$$

Then using this refinement, we give some nice applications in refining some important inequalities between means and the information inequality.

Key words and phrases: Jensen's inequality, Product measure, Weight function, Mean, Information Inequality.
2000 Mathematics Subject Classifications. 26D15, 26A51, 28A35.

1. INTRODUCTION

Throughout this paper, we assume that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ are two probability measure spaces, I is an interval of the real line, $f \in L^1(\mu)$, $f(x) \in I$ for each $x \in X$, and φ a real-valued convex function on I. The classical integral form of Jensen's inequality is as follows [2]:

$$\varphi\left(\int_X f d\mu\right) \leq \int_X (\varphi \circ f) d\mu. \tag{1}$$

We mean a weight function on $X \times Y$, an $\mathcal{A} \times \mathcal{B}$-measurable mapping $\omega : X \times Y \longrightarrow [0, +\infty)$; see e.g. [8], such that

$$\int_X \omega(x, y) d\mu(x) = 1 \qquad \text{for each } y \in Y,$$

and

$$\int_Y \omega(x, y) d\lambda(y) = 1 \qquad \text{for each } x \in X.$$

For example, if X and Y are the unit interval $[0, 1]$ with Lebesgue measure, then

$$\omega(x, y) = 1 + (\sin 2\pi x)(\sin 2\pi y)$$

is a weight function on $[0, 1] \times [0, 1]$.

In [4] the following refinement of Jensen's inequality (1) is achieved.

Theorem A.
With the above assumptions, if ω is a weight function on $X \times Y$, then

$$\int_Y \varphi \left(\int_X f(x)\omega(x, y) d\mu(x) \right) d\lambda(y)$$

is meaningful, and we have

$$\varphi \left(\int_X f d\mu \right) \leq \int_Y \varphi \left(\int_X f(x)\omega(x, y) d\mu(x) \right) d\lambda(y) \leq \int_X (\varphi \circ f) d\mu. \tag{2}$$

The following important inequality is a consequence of Theorem A:

Corollary B.
If φ is a real-valued convex function on a closed interval $[a, b]$, then we have Hermite-Hadamard inequality:

$$\varphi \left(\frac{a + b}{2} \right) \leq \frac{1}{b - a} \int_a^b \varphi(t) dt \leq \frac{\varphi(a) + \varphi(b)}{2}. \tag{3}$$

2. MAIN RESULTS

In this section, we extend Theorem A for two weight functions. According to the proof of Theorem A, for each weight function ω, the integral $\int_X f(x)\omega(x, y) d\mu(x)$ is existed and finite for λ-almost all $y \in Y$'s. Fix an $\alpha \in I$. Let ω_0 and ω_1 be two weight functions on $X \times Y$. There exists a measurable set E with $\lambda(E) = 0$ such that for each $y \in Y \setminus E$ the integrals

$$F_k(y) := \int_X f(x)\omega_k(x, y) d\mu(x) \qquad (k = 0, 1) \tag{4}$$

are existed and finite. We set on E, $F_k(y) = \alpha$ $(k = 0, 1)$. Clearly, for each $0 \leq t \leq 1$,

$$\omega_t(x, y) := (1 - t)\omega_0(x, y) + t\omega_1(x, y)$$

is a weight function on $X \times Y$, and the integral

$$\int_X f(x)\omega_t(x, y) d\mu(x) = (1 - t) \int_X f(x)\omega_0(x, y) d\mu(x) + t \int_X f(x)\omega_1(x, y) d\mu(x)$$

is existed and finite for each $y \in Y \setminus E$. Thus, if for each $0 \le t \le 1$, we set

$$F_t(y) := \begin{cases} \int_X f(x)\omega_t(x,y)d\mu(x) & ; y \in Y \setminus E, \\ \alpha & ; y \in E. \end{cases} \tag{5}$$

the function $F_t(y) = (1-t)F_0(y) + tF_1(y) \in I$ belongs to $L^1(\lambda)$; [4].
First we prove the following lemma.

Lemma 2.1
Let η be an arbitrary positive measure on X and $h = u - v : X \to (-\infty, +\infty]$, where $0 \le u \le +\infty$ and $0 \le v < +\infty$ are measurable functions on X. Now, if $\int_X v d\eta < \infty$, then $\int_X h d\eta$ is meaningful and

$$\int_X h d\eta = \int_X u d\eta - \int_X v d\eta. \tag{6}$$

Proof
Since $h^- \le v$, we have $\int_X h^- d\eta \le \int_X v d\eta < \infty$, and so $\int_X h d\eta$ is meaningful. Now, since $h^+ + v = h^- + u$, we have $\int_X h^+ d\eta + \int_X v d\eta = \int_X h^- d\eta + \int_X u d\eta$, and so

$$\int_X h d\eta = \int_X h^+ d\eta - \int_X h^- d\eta = \int_X u d\eta - \int_X v d\eta.$$

□

Theorem 2.2
With the above assumption, $\int_Y A(\varphi; F_0(y), F_1(y)) d\lambda(y)$ is meaningful and

$$\varphi\left(\int_X f d\mu\right) \le \int_Y A\left(\varphi; F_0(y), F_1(y)\right) d\lambda(y) \le \int_X (\varphi \circ f) d\mu, \tag{7}$$

where the arithmetic mean A for an arbitrary integrable function g over a closed interval with end points a and b, is defined by

$$A(g; a, b) = \frac{1}{b-a}\int_a^b g(x)dx. \tag{8}$$

Proof
By Theorem A, for each $0 \le t \le 1$, $\int_Y \varphi^-(F_t(y)) d\lambda(y) < \infty$ and

$$\varphi\left(\int_X f d\mu\right) \le \int_Y \varphi(F_t(y)) d\lambda(y) \le \int_X (\varphi \circ f) d\mu.$$

So,

$$\begin{aligned} \varphi\left(\int_X f d\mu\right) \quad + \quad & \int_Y \varphi^-\left(F_t(y)\right) d\lambda(y) \le \int_Y \varphi^+\left(F_t(y)\right) d\lambda(y) \\ \le \quad & \int_X (\varphi \circ f) d\mu + \int_Y \varphi^-\left(F_t(y)\right) d\lambda(y). \end{aligned} \tag{9}$$

On the other hand, the function $Y \times [0,1] \to I$ with

$$(y,t) \to \varphi(F_t(y)) = \varphi((1-t)F_0(y) + tF_1(y))$$

is product-measurable. So, integrating each side of (9) with respect to t on $[0,1]$, we have

$$\begin{aligned} \varphi\left(\int_X f d\mu\right) \quad + \quad & \int_0^1 \int_Y \varphi^-\left(F_t(y)\right) d\lambda(y) dt \le \int_0^1 \int_Y \varphi^+\left(F_t(y)\right) d\lambda(y) dt \\ \le \quad & \int_X (\varphi \circ f) d\mu + \int_0^1 \int_Y \varphi^-\left(F_t(y)\right) d\lambda(y) dt. \end{aligned} \tag{10}$$

Fix an $s_0 \in \operatorname{int} I$. There is a real number m, such that

$$\varphi(s) \geq m(s - s_0) + \varphi(s_0) \qquad (s \in I).$$

Thus,

$$\varphi^-(s) \leq |m(s - s_0) + \varphi(s_0)| \leq |m||s| + |m||s_0| + |\varphi(s_0)| \qquad (s \in I).$$

In particular, for each t and y, letting $s = F_t(y)$, we have

$$\begin{aligned} \varphi^-(F_t(y)) &\leq |m||F_t(y)| + |m||s_0| + |\varphi(s_0)| \\ &\leq |m|\left((1-t)|F_0(y)| + t|F_1(y)|\right) + |m||s_0| + |\varphi(s_0)|. \end{aligned}$$

So for each $y \in Y$,

$$\int_0^1 \varphi^-(F_t(y))dt \leq |m|\frac{|F_0(y)| + |F_1(y)|}{2} + |m||s_0| + |\varphi(s_0)| < \infty \qquad (0 \leq t \leq 1),$$

and

$$\int_Y \int_0^1 \varphi^-(F_t(y))dt d\lambda(y) \leq |m|||f||_1 + |m||s_0| + |\varphi(s_0)| < \infty,$$

because,

$$\begin{aligned} &\int_Y |F_k(y)|d\lambda(y) = \int_Y \left| \int_X f(x)\omega_k(x,y)d\mu(x) \right| d\lambda(y) \\ &\leq \int_Y \int_X |f(x)|\omega_k(x,y)d\mu(x)d\lambda(y) = \int_X |f(x)|d\mu(x) \int_Y \omega_k(x,y)d\lambda(y) \\ &= \int_X |f(x)|d\mu(x) = ||f||_1 \qquad (k = 0, 1). \end{aligned}$$

Therefore, applying Lemma 2.1 for the function $Y \to (-\infty, +\infty]$ with

$$y \to \int_0^1 \varphi(F_t(y))dt = \int_0^1 \varphi^+(F_t(y))dt - \int_0^1 \varphi^-(F_t(y))dt,$$

the integral $\int_Y \int_0^1 \varphi(F_t(y))dt d\lambda(y)$ is meaningful and

$$\int_Y \int_0^1 \varphi(F_t(y))dt d\lambda(y) = \int_Y \int_0^1 \varphi^+(F_t(y))dt d\lambda(y) - \int_Y \int_0^1 \varphi^-(F_t(y))dt d\lambda(y).$$

Now, changing the order of integrations, we conclude from (10) that

$$\varphi\left(\int_X f d\mu\right) \leq \int_Y \int_0^1 \varphi(F_t(y))dt d\lambda(y) \leq \int_X (\varphi \circ f)d\mu.$$

But,

$$\int_0^1 \varphi(F_t(y))dt = \int_0^1 \varphi((1-t)F_0(y) + tF_1(y))dt,$$

which with the change of the variable $u = (1-t)F_0(y) + tF_1(y)$, we have

$$\int_0^1 \varphi(F_t(y))dt = \frac{1}{F_1(y) - F_0(y)} \int_{F_0(y)}^{F_1(y)} \varphi(u)du = A(\varphi;\ F_0(y), F_1(y)).$$

This completes the proof. $\square$

Remark 2.3
In the discrete case, when m and n are two positive integers, and $x_i \in I$ $(1 \le i \le m)$, considering $X = \{1, \cdots, m\}$, $Y = \{1, \cdots, n\}$, $\mathcal{A} = 2^X$, $\mathcal{B} = 2^Y$, $\mu\{i\} = \mu_i$, $\lambda\{j\} = \lambda_j$, $f(i) = x_i$ $(i = 1, \cdots, m;\ j = 1, \cdots, n)$, we have [5]:

$$\varphi\left(\sum_{i=1}^{m}\mu_i x_i\right) \le \sum_{j=1}^{n}\lambda_j A(\varphi; F_0(j), F_1(j)) \le \sum_{i=1}^{m}\mu_i\varphi(x_i), \tag{11}$$

where

$$F_k(j) = \sum_{i=1}^{m}\omega_k(i,j)\mu_i x_i \qquad (k = 0, 1;\quad 1 \le j \le n).$$

In the following sections, we give some applications of (7) in refining of some important inequalities between means and the information inequality. Considering (11), these are extensions of the results obtained in [5], [6] and [7] previously.

3. APPLICATIONS TO INEQUALITIES BETWEEN MEANS

In this section, we give some refinements of several important inequalities between means, such as, AGM, Ky Fan and Sandor inequalities.

Theorem 3.1 (**First refinement of AGM inequality)**
With the above assumptions, if $f : X \to (0, +\infty)$ belongs to $L^1(\mu)$, then

$$\exp\left(\int_X \ln f d\mu\right) \le \exp\int_Y \ln I(F_0(y), F_1(y)) d\lambda(y) \le \int_X f d\mu, \tag{12}$$

where the identric mean I of each $a, b > 0$, is defined by

$$I(a,b) = \begin{cases} \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}} & ;\ a \ne b, \\ a & ; a = b. \end{cases} \tag{13}$$

This is a refinement of AGM inequality $\exp\left(\int_X \ln f d\mu\right) \le \int_X f d\mu$; see e.g. [8].

Proof
The function $\varphi(t) = -\ln t$ is convex on $(0, +\infty)$, and we have

$$A(\ln;\ a, b) = \ln I(a,b) \qquad (a, b > 0). \tag{14}$$

Now, the assertion follows from (7). $\square$

Theorem 3.2 (**Second refinement of AGM inequality)**
With the above assumptions, if $g : X \to (0, \infty)$ and $\ln g \in L^1(\mu)$, then

$$\exp\left(\int_X \ln g d\mu\right) \le \int_Y L\left(\exp(F_0(y)), \exp(F_1(y))\right) d\lambda(y) \le \int_X g d\mu, \tag{15}$$

where for λ-almost all $y \in Y$'s

$$F_k(y) = \int_X \ln g(x)\omega_k(x,y) d\mu(x) \qquad (k = 0, 1),$$

and the Logarithmic mean L of each $a, b > 0$, is defined by

$$L(a,b) = \begin{cases} \frac{b-a}{\ln b - \ln a} & ; a \neq b, \\ a & a = b. \end{cases} \tag{16}$$

Proof
The function $\varphi(t) = \exp(t)$ is convex on $\mathbb{R}$, and we have

$$A(\exp;\ a,b) = L(e^a, e^b) \qquad (a, b \in \mathbb{R}). \tag{17}$$

Now, the assertion follows from (7) by taking $f = \ln g$. □

Theorem 3.3 (**A refinement of Ky Fan's inequality**)
With the above assumptions, if $f : X \to (0, \frac{1}{2}]$ belongs to $L^1(\mu)$, then

$$\frac{1 - \int_X f d\mu}{\int_X f d\mu} \leq \exp \int_Y \ln \frac{I(1 - F_0(y), 1 - F_1(y))}{I(F_0(y), F_1(y))} d\lambda(y) \leq \exp \int_X \ln \frac{1-f}{f} d\mu, \tag{18}$$

which is a refinement of Ky Fan's inequality $\frac{1-\int_X f d\mu}{\int_X f d\mu} \leq \exp \int_X \ln \frac{1-f}{f} d\mu$; [1].

Proof
The function $\varphi : (0, \frac{1}{2}] \to \mathbb{R}$ with $\varphi(t) = \ln \frac{1-t}{t}$ is convex on $I = (0, \frac{1}{2}]$, and we have

$$A(\varphi; a, b) = \ln \frac{I(1-a, 1-b)}{I(a,b)} \qquad (0 < a, b < 1)\,.$$

Now, the assertion follows from (7). □

Theorem 3.4 (**A refinement of Sandor's inequality**)
With the above assumptions, if $f : X \to (0, \frac{1}{2}]$ belong to $L^1(\mu)$, then

$$\begin{aligned} \frac{1}{\int_X f d\mu} - \frac{1}{1 - \int_X f d\mu} \quad &\leq \quad \int_Y \left(\frac{1}{L(F_0(y), F_1(y))} - \frac{1}{L(1 - F_0(y), 1 - F_1(y))} \right) d\lambda(y) \\ &\leq \quad \int_X \frac{1}{f} d\mu - \int_X \frac{1}{1-f} d\mu, \end{aligned} \tag{19}$$

which is a refinement of Sandor's inequality $\frac{1}{\int_X f d\mu} - \frac{1}{1-\int_X f d\mu} \leq \int_X \frac{1}{f} d\mu - \int_X \frac{1}{1-f} d\mu$; see [9] and also [1].

Proof
The function $\varphi : (0, \frac{1}{2}] \to \mathbb{R}$ with $\varphi(t) = \frac{1}{t} - \frac{1}{1-t}$ is convex on $(0, \frac{1}{2}]$, and

$$A(\varphi;\ a, b) = \frac{1}{L(a,b)} - \frac{1}{L(1-a, 1-b)}, \qquad (0 < a, b < 1)\,.$$

Now, the assertion follows from (7). □

4. APPLICATIONS TO INFORMATION THEORY

Let $(X, \mathcal{A}, \eta)$ be a positive measure space and $p, q : X \to (0, +\infty)$ be $\mathcal{A}$-measurable and

$$\int_X p d\eta = \int_X q d\eta = 1.$$

The information inequality [3] states that

$$D(p||q) := \int_X p \ln \frac{p}{q} d\eta \geq 0. \tag{20}$$

In this section, we give some refinements of the Information Inequality (20).

Theorem 4.1
With the above assumptions, we have

$$D(p||q) \geq -\int_Y \ln I\left(F_0(y), F_1(y)\right) d\lambda(y) \geq 0, \tag{21}$$

where I is the identric mean (13), and for λ-almost all $y \in Y$'s,

$$F_k(y) = \int_X q(x)\omega_k(x,y)d\eta(x) \qquad (k = 0, 1).$$

Proof
The function $\varphi(t) = -\ln t$ is convex on $(0, +\infty)$, and so the assertion follows from (7) by considering $d\mu = pd\eta$, $f = \frac{q}{p}$ and (14). □

Theorem 4.2
With the above assumptions, we have

$$D(p||q) \geq \int_Y \ln \sqrt{(I(F_0(y))^2, (F_1(y))^2)^{\frac{F_0(y)+F_1(y)}{2}}} d\lambda(y) \geq 0 \tag{22}$$

where I is the identric mean (13), and for λ-almost all $y \in Y$'s,

$$F_k(y) = \int_X p(x)\omega_k(x,y)d\eta(x) \qquad (k = 0, 1).$$

Proof
The function $\varphi(t) = t \ln t$ is convex on $(0, +\infty)$ and it is easily seen that

$$A(\varphi; a, b) = \frac{a+b}{4} \ln I(a^2, b^2) \qquad (a, b > 0).$$

Now, the assertion follows from (7) by considering $d\mu = qd\eta$ and $f = \frac{p}{q}$. □

Theorem 4.3
With the above assumptions, if $p \ln \frac{p}{q} \in L^1(\eta)$, then

$$D(p||q) \geq -\ln \int_Y L\left(e^{F_0(y)}, e^{F_1(y)}\right) d\lambda(y) \geq 0,$$

where L is the logarithmic mean (16), and for λ-almost all $y \in Y$'s

$$F_k(y) = \int_X \ln \frac{q(x)}{p(x)} \omega_k(x,y)p(x)d\eta(x) \qquad (k = 0, 1).$$

Proof
The function $\varphi(t) = \exp(t)$ is convex on $\mathbb{R}$. Now, the assertion follows from (7) by taking $d\mu = pd\eta$, $f = \ln \frac{q}{p}$, and considering (17). □

REFERENCES

1. Alzer H. The Inequality of Ky Fan and related results. *Acta Applicandae Mathematicae* 1995; **38**:305-354.

2. Hardy GH, Littlewood JE, Pólya G. *Inequalities.* Cambridge University Press 1934.

3. Jones AG, Jones JM. *Information and coding theory.* Springer-Verlag 2000.

4. Rooin J. A refinement of Jensen's inequality. *Journal of Inequalities in Pure and Applied Mathematics* 2005; **6** Art. 38:1-4.

5. Rooin J. Some refinements of discrete Jensen's inequality and some of its applications. `http://arxiv.org/abs/math.NA/0610736`

6. Rooin J, Morassaei A. Some refinements of relative information inequality. *Creative Mathematics and Informatics* 2007; **16**:95-98.

7. Rooin J. Some applications of an important weighted refinement of discrete Jensen's inequality. *Submitted*.

8. Rudin W. *Real and complex analysis.* 3rd ed., McGraw-Hill 1974.

9. Sandor J. On an inequality of Ky Fan II. *International Journal of Mathematical Education in Science and Technology* 1991; **22**:326-328.