Journal of Inequalities in Pure and Applied Mathematics

A REFINEMENT OF JENSEN'S INEQUALITY

J. ROOIN

Department of Mathematics
Institute for Advanced Studies in Basic Sciences
Zanjan, Iran

EMail: rooin@iasbs.ac.ir

ISSN (electronic): 1443-5756
160-04

volume 6, issue 2, article 38, 2005.

Received 27 August, 2004; accepted 16 March, 2005.

Communicated by: C.E.M. Pearce

Abstract

We refine Jensen's inequality as

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

where $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ are two probability measure spaces, $\omega : X \times Y \to [0, \infty)$ is a weight function on $X \times Y$, I is an interval of the real line, $f \in L^1(\mu)$, $f(x) \in I$ for all $x \in X$ and φ is a real-valued convex function on I.

2000 Mathematics Subject Classification: Primary: 26D15, 28A35.
Key words: Product measure, Fubini's Theorem, Jensen's inequality.

Contents

1. Introduction

The classical integral form of Jensen's inequality states that

$$\varphi\left(\int_X f d\mu\right) \leq \int_X (\varphi \circ f) d\mu, \tag{1.1}$$

where $(X, \mathcal{A}, \mu)$ is a probability measure space, I is an interval of the real line, $f \in L^1(\mu)$, $f(x) \in I$ for all $x \in X$ and φ is a real-valued convex function on I; see e.g. [2, p. 202] or [4, p. 62]. Now suppose that $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ are two probability measure spaces. By a (separately) weight function on $X \times Y$ we mean a product- measurable mapping $\omega : X \times Y \to [0, \infty)$, see e.g. [4, p. 160], such that

$$\int_X \omega(x, y) d\mu(x) = 1 \qquad (\text{for each } y \text{ in } Y), \tag{1.2}$$

and

$$\int_Y \omega(x, y) d\lambda(y) = 1 \qquad (\text{for each } x \text{ in } X). \tag{1.3}$$

For example, if we take X and Y as 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 this paper, using a weight function ω, we refine Jensen's inequality (1.1) as in the following section. For some applications in the discrete case, see e.g. [3].

2. Refinement

In this section, using the terminologies of the introduction, we refine the integral form of Jensen's inequality (1.1) via a weight function ω.

Theorem 2.1. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two probability measure spaces and $\omega : X \times Y \to [0, \infty)$ be a weight function on $X \times Y$. If I is an interval of the real line, $f \in L^1(\mu)$, $f(x) \in I$ for all $x \in X$, and φ is a real convex function on I, then*

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

has meaning and we have

$$\varphi\left(\int_X f d\mu\right) \leq \int_Y \varphi\left(\int_X f(x)\omega(x,y)d\mu(x)\right) d\lambda(y) \leq \int_X (\varphi \circ f) d\mu. \tag{2.1}$$

Proof. The functions ω and $(x, y) \to f(x)$, and so

$$(x, y) \to f(x)\omega(x, y)$$

is product-measurable on $X \times Y$. Now since

$$\begin{aligned} \int_X \int_Y |f(x)|\omega(x,y)d\lambda(y)d\mu(x) &= \int_X |f(x)| \left(\int_Y \omega(x,y)d\lambda(y)\right) d\mu(x) \\ &= \int_X |f(x)|d\mu(x) = \|f\|_{L^1(\mu)} < \infty, \end{aligned} \tag{2.2}$$

by Fubini's theorem, the real-valued function $(x, y) \to f(x)\omega(x, y)$ on $X \times Y$ belongs to $L^1(\mu \times \lambda)$. Therefore for λ-almost all $y \in Y$, the function $x \to f(x)\omega(x, y)$ belongs to $L^1(\mu)$. Fix an arbitrary $\alpha \in I$. Define $F : Y \to \mathbb{R}$, by

$$F(y) = \int_X f(x)\omega(x, y)d\mu(x)$$

if the integral exists, and $F(y) = \alpha$ otherwise. By Fubini's theorem, we have $F \in L^1(\lambda)$. It is easy to show that $F(y) \in I$ $(y \in Y)$. So,

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

has meaning and is an extended real number belonging to $(-\infty, +\infty]$; see e.g. [4, p. 62]. Now, since $(x, y) \to f(x)\omega(x, y)$ belongs to $L^1(\mu \times \lambda)$, by (1.1) and Fubini's theorem, we have

$$\begin{aligned}\int_Y \varphi\left(\int_X f(x)\omega(x, y)d\mu(x)\right) d\lambda(y) &= \int_Y (\varphi \circ F)(y)d\lambda(y) \\ &\geq \varphi\left(\int_Y F(y)d\lambda(y)\right) \\ &= \varphi\left(\int_Y \int_X f(x)\omega(x, y)d\mu(x)d\lambda(y)\right) \\ &= \varphi\left(\int_X f(x)\left(\int_Y \omega(x, y)d\lambda(y)\right) d\mu(x)\right) \\ &= \varphi\left(\int_X f d\mu\right),\end{aligned}$$

and the left-hand side inequality (2.1) is obtained.

For the right-hand side inequality in (2.1), we consider two cases: If $\int_X (\varphi \circ f) d\mu = +\infty$, the assertion is trivial. Suppose then, $\varphi \circ f \in L^1(\mu)$. Take an arbitrary $y \in Y$ such that $x \to f(x)\omega(x,y)$ belongs to $L^1(\mu)$, and put

$$d\nu^y = \omega^y d\mu,$$

where

$$\omega^y(x) = \omega(x,y) \qquad (x \in X).$$

Trivially, $(X, \mathcal{A}, \nu^y)$ is a probability measure space, $f \in L^1(\nu^y)$ and

$$F(y) = \int_X f(x)\omega(x,y)d\mu(x) = \int_X f(x)d\nu^y(x).$$

Thus, by Jensen's inequality (1.1), we have

$$(\varphi \circ F)(y) = \varphi\left(\int_X f(x)d\nu^y(x)\right) \leq \int_X (\varphi \circ f)d\nu^y. \tag{2.3}$$

Since $\varphi \circ f \in L^1(\mu)$,

$$\begin{aligned} &\int_X \int_Y |(\varphi \circ f)(x)|\omega(x,y)d\lambda(y)d\mu(x) \\ &\qquad = \int_X |(\varphi \circ f)(x)|d\mu(x) \int_Y \omega(x,y)d\lambda(y) \\ &\qquad = \int_X |(\varphi \circ f)(x)|d\mu(x) < \infty, \end{aligned} \tag{2.4}$$

and so for λ-almost all $y \in Y$, the function $x \to (\varphi \circ f)(x)\omega(x,y)$ belongs to $L^1(\mu)$ and for these y's, we have

$$\int_X (\varphi \circ f)(x)\omega(x,y)d\mu(x) = \int_X (\varphi \circ f)(x)d\nu^y(x). \tag{2.5}$$

Thus, by (2.3) and (2.5), for λ-almost all $y \in Y$

$$(\varphi \circ F)(y) \leq \int_X (\varphi \circ f)(x)\omega(x,y)d\mu(x). \tag{2.6}$$

Denote temporarily the right-hand side of (2.6) by $\psi(y)$ (put $\psi(y) = 0$, if the integral does not exist). Since by (2.4), $\psi \in L^1(\lambda)$, from $(\varphi \circ F)^+ \leq \psi^+$ (λ-a.e.), we conclude that $\int_Y (\varphi \circ F)^+ d\lambda \leq \int_Y \psi^+ d\lambda < \infty$.

On the other hand, we know that $\int_Y (\varphi \circ F)^- d\lambda < \infty$. Thus $\varphi \circ F \in L^1(\lambda)$, and so by (2.6), (2.4) and Fubini's theorem,

$$\begin{aligned}
\int_Y \varphi\left(\int_X f(x)\omega(x,y)d\mu(x)\right) d\lambda(y) &= \int_Y (\varphi \circ F)(y)d\lambda(y) \\
&\leq \int_Y \psi(y)d\lambda(y) \\
&= \int_Y \int_X (\varphi \circ f)(x)\omega(x,y)d\mu(x)d\lambda(y) \\
&= \int_X (\varphi \circ f)(x)d\mu(x) \int_Y \omega(x,y)d\lambda(y) \\
&= \int_X (\varphi \circ f)d\mu.
\end{aligned}$$

This completes the proof. $\square$

Corollary 2.2. *If φ is a real convex function on a closed interval $[a, b]$, then we have Hermite-Hadamard inequalities [1]:*

$$\varphi\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b \varphi(t)dt \leq \frac{\varphi(a)+\varphi(b)}{2}. \tag{2.7}$$

Proof. Put $X = \{0, 1\}$ with $\mathcal{A} = 2^X$ and $\mu\{0\} = \mu\{1\} = \frac{1}{2}$, and $Y = [0, 1]$ with Lebesgue measure λ. Now, (2.7) follows from (2.1) by taking $\omega(0, y) = 2(1 - y)$, $\omega(1, y) = 2y$ $(0 \leq y \leq 1)$, $I = [a, b]$, $f(0) = a$, $f(1) = b$, and considering the change of variables $t = (1 - y)a + yb$. □

We conclude this paper by the following open problem:

Open problem. Characterize all weight functions. Actually, if $\omega(x, y)$ is a weight function, then $\theta(x, y) = \omega(x, y) - 1$ satisfies the following relations:

$$\int_X \theta(x, y)d\mu(x) = 0 \qquad (\text{for each } y \text{ in } Y), \tag{2.8}$$

$$\int_Y \theta(x, y)d\lambda(y) = 0 \qquad (\text{for each } x \text{ in } X). \tag{2.9}$$

So precisely, the weight functions are of the form $1 + \theta(x, y)$ with nonnegative values such that $\theta(x, y)$ is product-measurable and satisfies (2.8) and (2.9). Therefore, it is sufficient only to characterize these θ's.

References

[1] S.S. DRAGOMIR AND C.E.M. PEARCE, *Selected Topics on Hermite-Hadamard Inequalities and Applications*, RGMIA Monographs, Victoria University, 2000. [ONLINE: http://rgmia.vu.edu.au/monographs/]

[2] E. HEWITT AND K. STROMBERG, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.

[3] J. ROOIN, Some aspects of convex functions and their applications, *J. Ineq. Pure and Appl. Math.*, **2**(1) (2001), Art. 4. [ONLINE http://jipam.vu.edu.au/article.php?sid=120]

[4] W. RUDIN, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1974.