J. of Inequal. & Appl., 2001, Vol. 6, pp. 451–462
Reprints available directly from the publisher
Photocopying permitted by license only

An Extension of Chebyshev's Inequality and its Connection with Jensen's Inequality*

CONSTANTIN P. NICULESCU†

Department of Mathematics, University of Craiova, Craiova 1100, Romania

(Received 24 November 1999; In final form 17 February 2000)

The aim of this paper is to show that Jensen's Inequality and an extension of Chebyshev's Inequality complement one another, so that they both can be formulated in a pairing form, including a second inequality, that provides an estimate for the classical one.

Keywords: Convex functions; Subdifferential; Mean value

Mathematics Subject Classifications 2000: Primary: 26A51, 26D15

1. INTRODUCTION

The well known fact that the derivative and the integral are inverse each other has a lot of interesting consequences, one of them being the duality between convexity and monotonicity. The purpose of the present paper is to relate on this basis two basic inequalities in Classical Analysis, precisely those due to Jensen and Chebyshev.

Both refer to mean values of integrable functions. Restricting ourselves to the case of finite measure spaces (X, Σ, μ), let us recall

* Partially supported by MEN Grant 39683/1998.
† e-mail: tempus@oltenia.ro

that the *mean value* of any μ-integrable function $f: X \to \mathbb{R}$ can be defined as

$$M(f) = \frac{1}{\mu(X)} \int_X f d\mu.$$

A useful remark is that the mean value of an integrable function belongs to any interval that includes its image; see [5], page 202.

In the particular case of an interval $[a, b]$, endowed with the Lebesgue measure, the two aforementioned inequalities reads as follows:

JENSEN'S INEQUALITY *Suppose that $f: [a, b] \to \mathbb{R}$ is an integrable function and φ is a convex function defined on an interval containing the image of f, such that $\varphi \circ f$ is integrable too. Then*

$$\varphi(M(f)) \leq M(\varphi \circ f).$$

CHEBYSHEV'S INEQUALITY *If $g, h: [a, b] \to \mathbb{R}$ are two nondecreasing functions then*

$$M(g)M(h) \leq M(gh).$$

Our goal is to show that Jensen's Inequality and an extension of Chebyshev's Inequality complement one another, so that they both can be formulated in a pairing form, including a second inequality, that provides an estimate for the classical one.

2. PRELIMINARIES

Before entering the details we shall need some preparation on the smoothness properties of the two type of functions involved: the convex and the nondecreasing ones.

Suppose that I is an interval (with interior Int I) and $f: \mathrm{I} \longrightarrow \mathbb{R}$ is a convex function. Then f is continuous on Int I and has finite left and right derivatives at each point of Int I. Moreover,

$$x < y \text{ in Int} I \Rightarrow D^{-}f(x) \leq D^{+}f(x) \leq \\ \leq D^{-}f(y) \leq D^{+}f(y) \tag{*}$$

which shows that both $D^- f$ and $D^+ f$ are nondecreasing on Int I. It is not difficult to prove that f must be differentiable except for at most countably many points. See [5], pp. 271–272. On the other hand, simple examples show that at the endpoints of I could appear problems even with the continuity. Notice that every discontinuous convex function $f: [a, b] \to \mathbb{R}$ comes from a continuous convex function $f_0 : [a, b] \to \mathbb{R}$, whose values at a and/or b were enlarged. In fact, any convex function f defined on an interval I is either monotonic or admits a point α such that f is nonincreasing on $(-\infty, \alpha] \cap I$ and nondecreasing on $[\alpha, \infty) \cup I$.

In the case of convex functions, the role of derivative is played by the *subdifferential*, a mathematical object which for a function $f: I \to \mathbb{R}$ is defined as the set ∂f of all functions $\varphi : I \to [-\infty, \infty]$ such that $\varphi(\operatorname{Int} I) \subset \mathbb{R}$ and

$$f(x) \geq f(a) + (x - a)\varphi(a), \quad (\forall)\ x, a \in I.$$

Geometrically, the subdifferential gives us the slopes of supporting lines for the graph of f.

The well known fact that a differentiable function is convex if (and only if) its derivative is nondecreasing has the following generalization in terms of subdifferential:

LEMMA 1 *Let I be an interval. The subdifferential of a function $f: I \to \mathbb{R}$ is non-empty if and only if f is convex. Moreover, if $\varphi \in \partial f$, then*

$$D^- f(x) \leq \varphi(x) \leq D^+ f(x)$$

for every $x \in$ Int I. Particularly, φ is a nondecreasing function.

Proof Necessity Suppose first that f is a convex function defined on an open interval I. We shall prove that $D^+ f \in \partial f$. For, let $x, a \in I$, with $x \geq a$. Then

$$\frac{f((1 - t)a + tx) - f(a)}{t} \leq f(x) - f(a)$$

for each $t \in (0, 1]$, which yields

$$f(x) \geq f(a) + D^+ f(a) \cdot (x - a).$$

If $x \leq a$, then a similar argument leads us to $f(x) \geq f(a) + D^{-}f(a) \cdot (x-a)$; or, $D^{-}f(a) \cdot (x-a) \geq D^{+}f(a) \cdot (x-a)$, because $x - a \leq 0$.

Analogously, we can argue that $D^{-}f \in \partial f$. Then from (∗) we infer that any $\varphi \in \partial f$ is necessarily nondecreasing.

If I is not open, say $I = [a, \infty)$, we can complete the definition of φ by letting $\varphi(a) = -\infty$.

Sufficiency Let $x, y \in I$, $x \neq y$, and let $t \in (0, 1)$. Then

$$\begin{aligned} f(x) \geq & f((1-t)x + ty) + \\ & + t(x-y) \cdot \varphi((1-t)x + ty) \\ f(y) \geq & f((1-t)x + ty) - \\ & - (1-t)(x-y) \cdot \varphi((1-t)x + ty). \end{aligned}$$

By multiplying the first inequality by $1 - t$, the second by t and then adding them side by side, we get

$$(1-t)f(x) + tf(x) \geq f((1-t)x + ty)$$

i.e., f is convex. ■

Let us consider now the case of nondecreasing functions. It is well known that each nondecreasing function $\varphi : I \to \mathbb{R}$ has at most countably many discontinuities (each of the first kind); less known is that φ admits primitives (also called *antiderivatives*). According to Dieudonné [2], a *primitive* of φ means any function $\Phi : I \to \mathbb{R}$ which is continuous on I, differentiable at each point of continuity of φ, and such that $\Phi' = \varphi$ at all those points. An example of a primitive of φ is

$$\Psi(x) = \int_a^x \varphi(t)dt, \quad x \in I,$$

a being arbitrarily fixed in I.

Because φ is nondecreasing, an easy computation shows that Ψ is a convex function. In fact, Ψ is continuous, so it suffices to show that

$$\Psi\left(\frac{x+y}{2}\right) \leq \frac{\Psi(x) + \Psi(y)}{2} \quad \text{for all } x, y \in I;$$

or, the last inequality is equivalent to

$$\int_{x}^{(x+y)/2} \varphi(t)dt \leq \int_{(x+y)/2}^{y} \varphi(t)dt,$$
$$\text{for all } x, y \in I, x \leq y$$

the later being clear because φ is nondecreasing.

On the other hand, by Denjoy–Bourbaki Theorem (The Generalized Mean Value Theorem) any two primitives of φ differ by a constant. See [2], 8.7.1. Consequently, all primitives of φ must be convex too!

3. THE MAIN RESULTS

We can now state our first result, the *complete form of Jensen's Inequality*:

THEOREM A *Let (X, Σ, μ) be a finite measure space and let $g: X \to \mathbb{R}$ be a μ-integrable function. If f is a convex function given on an interval I that includes the image of g and $\varphi \in \partial f$ is a function such that $\varphi \circ g$ and $g \cdot (\varphi \circ g)$ are integrable, then the following inequalities hold:*

$$0 \leq M(f \circ g) - f(M(g)) \leq M(g \cdot (\varphi \circ g)) - M(g)M(\varphi \circ g).$$

Proof The first inequality is that of Jensen, for which we give the following simple argument: If $M(g) \in \operatorname{Int} I$, then

$$f(g(x)) \geq f(M(g)) + (g(x) - M(g)) \cdot \varphi(M(g)) \quad \text{for all } x \in X$$

and the Jensen's inequality follows by integrating both sides over X. The case where $M(g)$ is an endpoint of I is straightforward because in that case $g = M(g)\, \mu$ a.e.

The second inequality can be obtained from

$$f(M(g)) \geq f(g(x)) + (M(g) - g(x)) \cdot \varphi(g(x)) \text{ for all } x \in X$$

by integrating both sides over X. ∎

COROLLARY 1 (See [3], for the case where f is a smooth convex function) *Let f be a convex function defined on an open interval I and let $\varphi \in \partial f$. Then*

$$0 \leq \sum_{k=1}^{n} \alpha_k f(x_k) - f\left(\sum_{k=1}^{n} \alpha_k x_k\right) \leq$$
$$\leq \sum_{k=1}^{n} \alpha_x x_k \varphi(x_k) - \left(\sum_{k=1}^{n} \alpha_k x_k\right)\left(\sum_{k=1}^{n} \alpha_k \varphi(x_k)\right)$$

for every $x_1, \ldots, x_n \in I$ and every $\alpha_1, \ldots, \alpha_n \in [0,1]$, with $\sum_{k=1}^{n} \alpha_k = 1$.

Corollary 1 above allows us to say something more even in the case of most familiar inequalities. Here are three examples, all involving concave functions; Theorem A and Corollary 1 both work in that case, simply by reversing the inequality signs. The first example concerns the sine function and improves on a well known inequality from Trigonometry: *If A, B, C are the angles of a triangle (expressed in radians) then*

$$0 \leq \frac{3\sqrt{3}}{2} - \sum \sin A \leq \sum \left(\frac{\pi}{3} - A\right) \cos A$$

i.e.,

$$\frac{3\sqrt{3}}{2} - \sum \left(\frac{\pi}{3} - A\right) \cos A \leq \sum \sin A \leq \frac{3\sqrt{3}}{2}.$$

To get a feeling of the lower estimate for $\sum \sin A$, just test the case of the triangle with angles $A = (\pi/2)$, $B = (\pi/3)$ and $C = (\pi/6)$!

Our second example concerns the function ln and improves on the *AM-GM Inequality*:

$$\sqrt[n]{x_1 \ldots x_n} \leq \frac{x_1 + \cdots + x_n}{n} \leq$$
$$\leq \sqrt[n]{x_1 \ldots x_n} \cdot \exp\left[\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)\left(\frac{1}{n}\sum_{k=1}^{n} \frac{1}{x_k}\right) - 1\right].$$

The exponent in the right hand side can be further evaluated *via* a classical inequality due to Schweitzer [7] (as strengthened by

Cîrtoaje [1]), so we can state the AM-GM Inequality as follows: *If* $0 < m \leq x_1, \ldots, x_n \leq M$, *then*

$$\frac{x_1 + \cdots + x_n}{n} \cdot \exp\left[1 - \frac{(M+m)^2}{4Mm} + \frac{[1+(-1)^{n-1}](M-m)^2}{8Mmn^2}\right] \leq$$
$$\leq \sqrt[n]{x_1 \ldots x_n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

Stirling's Formula suggests the possibility of further improvement, a subject which will be considered elsewhere.

Theorem A also allows us to estimate tricky integrals such as

$$I = \frac{4}{\pi}\int_0^{\pi/4} \ln(1 + \tan x)dx = \frac{1}{2}\ln 2 = 0.34657\ldots$$

By Jensen's Inequality,

$$\frac{4}{\pi}\int_0^{\pi/4} \ln(1 + \tan x)dx \leq \ln\left(\frac{4}{\pi}\int_0^{\pi/4} (1 + \tan x)dx\right)$$

which yields a pretty good upper bound for I because the difference between the two sides is (approximately) 1.8952×10^{-2}. Notice that an easy computation shows that

$$\int_0^{\pi/4} (1 + \tan x)dx = \frac{1}{4}\pi + \frac{1}{2}\ln 2,$$

Theorem A allows us to indicate a valuable lower bound for I, precisely,

$$I \geq \ln\left(\frac{4}{\pi}\int_0^{\pi/4} (1 + \tan x)dx\right) +$$
$$+ 1 - \frac{4}{\pi}\int_0^{\pi/4} (1 + \tan x)dx \cdot \frac{4}{\pi}\int_0^{\pi/4} (1 + \tan x)^{-1}dx.$$

In fact, Maple V4 shows that I exceeds the left hand side by 1.9679×10^{-2}.

Using the aforementioned duality between the convex functions and the nondecreasing ones, we can infer from Theorem A the following result:

THEOREM B (The extension of Chebyshev's Inequality) *Let* (X, Σ, μ) *be a finite measure space, let* $g\colon X \to \mathbb{R}$ *be a* μ*-integrable function and let* φ *be a nondecreasing function given on an interval that includes the image of* g *and such that* $\varphi \circ g$ *and* $g \cdot (\varphi \circ g)$ *are integrable functions. Then for every primitive* Φ *of* φ *such that* $\Phi \circ g$ *is integrable the following inequalities hold true*:

$$0 \leq M(\Phi \circ g) - \Phi(M(g)) \leq M(g \cdot (\varphi \circ g)) - M(g)M(\varphi \circ g).$$

In order to show how Theorem B yields Chebyshev's Inequality we have to consider two cases. The first one concerns the situation where $g\colon [a,b] \to \mathbb{R}$ is increasing and $h\colon [a,b] \to \mathbb{R}$ is nondecreasing. In that case we apply Theorem B to g and $\varphi = h \circ g^{-1}$. When both g and h are nondecreasing, we shall consider increasing perturbations of g, *e.g.*, $g + \varepsilon x$ for $\varepsilon > 0$. By the previous case,

$$M((g+\varepsilon x)h) \leq M(g+\varepsilon x)M(h)$$

for each $\varepsilon > 0$ and it remains to take the limit as $\varepsilon \to 0+$.

The following two inequalities are consequences of Theorem B:

$$\begin{aligned}&\int_0^1 \left(\sin\frac{1}{x}\right)\left(\sin\frac{1}{x} + \cos\sin\frac{1}{x}\right)dx - \left(\int_0^1 \sin\frac{1}{x}dx\right)\\ &\quad \cdot \left(\int_0^1 \left(\sin\frac{1}{x} + \cos\sin\frac{1}{x}\right)dx\right) > \int_0^1 \left(\frac{1}{2}\left(\sin\frac{1}{x}\right)^2 + \sin\sin\frac{1}{x}\right)dx\\ &\quad - \frac{1}{2}\left(\int_0^1 \sin\frac{1}{x}dx\right)^2 - \sin\left(\int_0^1 \sin\frac{1}{x}dx\right) > 0;\end{aligned} \tag{1}$$

$$\begin{aligned}&\int_0^1 \frac{\sin x}{x} e^{\sin x/x}dx - \left(\int_0^1 \frac{\sin x}{x}dx\right)\left(\int_0^1 e^{\sin x/x}dx\right) >\\ &\quad > \int_0^1 \exp\left(\frac{\sin x}{x}\right)dx - \exp\left(\int_0^1 \frac{\sin x}{x}dx\right) > 0.\end{aligned} \tag{2}$$

The first one corresponds to the case where $g(x) = \sin(1/x)$ and $\varphi(x) = x + \cos x$, while the second one to $g(x) = ((\sin x)/x)$ and $\varphi(x) = e^x$; the fact that the inequalities above are strict is straightforward. In both cases, the integrals involved cannot be computed exactly (*i.e.*, *via* elementary functions).

Using MAPLE V4, we can estimate the different integrals and obtain that (1) looks like

$$0.37673 > 0.16179 > 0$$

while (2) looks like $5.7577 \times 10^{-3} > 2.8917 \times 10^{-3} > 0$.

4. ANOTHER ESTIMATE OF JENSEN'S INEQUALITY

The following result complements Theorem A and yields a valuable upper estimate of Jensen's Inequality:

THEOREM C *Let $f:[a,b] \to \mathbb{R}$ be a continuous convex function, and let $[m_1, M_1], \ldots, [m_n, M_n]$ be compact subintervals of $[a,b]$. Given $\alpha_1, \ldots, \alpha_n \in [0,1]$ with $\sum_{k=1}^{n} \alpha_k = 1$, the function*

$$E(x_1, \ldots, x_n) = \sum_{k=1}^{n} \alpha_k f(x_k) - f\left(\sum_{k=1}^{n} \alpha_k x_k\right)$$

attains its supremum on $\Omega = [m_1, M_1] \times \cdots \times [m_n, M_n]$ at a boundary point (i.e., at a point of $\partial\Omega = \{m_1, M_1\} \times \cdots \times \{m_n, M_n\}$).

Moreover, the conclusion remains valid for every compact convex domain in $[a,b]^n$.

The proof of Theorem C depends upon the following refinement of Lagrange Mean Theorem:

LEMMA 2 *Let h: $[a,b] \to \mathbb{R}$ be a continuous function. Then there exists a point $c \in (a,b)$ such that*

$$\underline{D}h(c) \leq \frac{h(b)-h(a)}{b-a} \leq \overline{D}h(c).$$

Here the lower and respectively the upper derivative of h at c are defined by

$$\underline{D}h(c) = \liminf_{x \to c} \frac{h(x)-h(c)}{x-c} \quad \text{and} \quad \overline{D}h(c) = \limsup_{x \to c} \frac{h(x)-h(c)}{x-c}.$$

Proof As in the smooth case, we consider the function

$$H(x) = h(x) - \frac{h(b)-h(a)}{b-a}(x-a), \quad x \in [a,b].$$

Clearly, H is continuous and $H(a) = H(b)$. If H attains its supremum at $c \in (a, b)$, then $\underline{D}H(c) \leq 0 \leq \overline{D}H(c)$ and the conclusion of Lemma 2 is immediate. The same is true when H attains its infimum at an interior point of $[a, b]$. If both extrema are attained at the endpoints, then H is constant and the conclusion of Lemma 2 works for every c in (a, b). ■

Proof of Theorem C It suffices to show that

$$E(x_1, \ldots, x_k, \ldots, x_n) \leq \sup\{E(x_1, \ldots, m_k, \ldots, x_n), E(x_1, \ldots, M_k, \ldots, x_n)\}$$

for every $x_k \in [m_k, M_k]$, $k \in \{1, \ldots, n\}$.

By reductio ad absurdum, we may assume that

$$E(x_1, x_2, \ldots, x_n) > \sup\{E(m_1, x_2, \ldots, x_n), E(M_1, x_2, \ldots, x_n)\}.$$

for some $x_1, x_2, \ldots, x_n$ with $x_k \in [m_k, M_k]$ for each $k \in \{1, \ldots, n\}$.

Letting fixed $x_k \in [m_k, M_k]$ with $k \in \{1, \ldots, n\}$, we consider the function

$$h : [m_1, M_1] \to \mathbb{R}, \quad h(x) = E(x, x_2, \ldots, x_n).$$

According to Lemma 2, there exists a $\xi \in (m_1, x_1)$ such that $h(x_1) - h(m_1) \leq (x_1 - m_1)\overline{D}h(\xi)$. As $h(x_1) > h(m_1)$, it follows that $\overline{D}h(\xi) > 0$, equivalently,

$$\overline{D}f(\xi) > \overline{D}f(\alpha_1\xi + \alpha_2 x_2 + \cdots + \alpha_n x_n).$$

Or, $\overline{D}f$ is a nondecreasing function on (a, b) (actually $\overline{D}f = D^+f$), which leads to $\xi > \alpha_1\xi + \alpha_2 x_2 + \cdots + \alpha_n x_n$, *i.e.*, to

$$\xi > \frac{\alpha_2 x_2 + \cdots + \alpha_n x_n}{\alpha_2 + \cdots + \alpha_n}.$$

A new appeal to Lemma 2 (applied this time to $h \,|\, [x_1, M_1]$), yields an $\eta \in (x_1, M_1)$ such that

$$\eta < \frac{\alpha_2 x_2 + \cdots + \alpha_n x_n}{\alpha_2 + \cdots + \alpha_n}.$$

Or, the later contradicts the fact that $\xi < \eta$. ■

The following application of Theorem C is due to Khanin [6]: *Let* $p > 1$, $x_1, \ldots, x_n \in [0, M]$ *and* $\alpha_1, \ldots, \alpha_n \in [0, 1]$, *with* $\sum_{k=1}^{n} \alpha_k = 1$.

Then

$$\sum_{k=1}^{n} \alpha_k x_k^p \leq \left(\sum_{k=1}^{n} \alpha_k x_k \right)^p + (p-1)p^{p/(1-p)} M^p.$$

Particularly,

$$\frac{x_1^2 + \cdots + x_n^2}{n} \leq \frac{(x_1 + \cdots + x_n)^2}{n} + \frac{M^2}{4},$$

which represents an additive converse to Cauchy–Schwarz Inequality.

In fact, according to Theorem C, the function

$$E(x_1, \ldots, x_n) = \sum_{k=1}^{n} \alpha_k x_k^p - \left(\sum_{k=1}^{n} \alpha_k x_k \right)^p,$$

attains its supremum on $[0, M]^n$ at a boundary point *i.e.*, at a point whose coordinates are either 0 or M. Therefore

$$\begin{aligned} \sup E(x_1, \ldots, x_n) &\leq M^p \cdot \sup\{s - s^p; s \in [0,1]\} = \\ &= (p-1)p^{p/(1-p)} M^p. \end{aligned}$$

Another immediate consequence of Theorem C is the following fact, which improves on a special case of a classical inequality due to Hardy, Littlewood and Polya (*cf.* [4], p. 89): *Let* f: $[a, b] \to \mathbb{R}$ *be a continuous convex function. Then*

$$\frac{f(a) + f(b)}{2} - f\left(\frac{a+b}{2}\right) \geq \frac{f(c) + f(d)}{2} - f\left(\frac{c+d}{2}\right)$$

for every $a \leq c \leq d \leq b$; in [4], one restricts to the case where $a+b=c+d$. The problem of extending this statement for longer families of numbers is left open.

References

[1] Cîrtoaje, V. (1990). On some inequalities with restrictions, *Revista de matematica din Timisoara*, **1**, 3–7 (Romanian).

[2] Dieudonné, J. (1960). *Foundations of Modern Analysis*, Academic Press.

[3] Drimbe, M. O. (1996). A connection between the inequalities of Jensen and Chebyshev, *Gazeta Matematica*, **CI**(5–6), 270–273 (Romanian).

[4] Hardy, G. H., Littewood, J. E. and Polya, G., Inequalities, *Cambridge Mathematical Library*, 2nd edn., 1952, Reprinted 1988.
[5] Hewitt, E. and Stromberg, K. (1965). *Real and Abstract Analysis*, Springer-Verlag.
[6] Khanin, L. G. (1988). *Problem* M 1083, Kvant, **18**(1), p. 35 and Kvant, **18**(5), p. 35.
[7] Polya, G. and Szegö, G., *Aufgaben und Lehrsätze aus Analysis*, Vol. 1, Springer Verlag, 1925, English edition, Springer-Verlag, 1972.