NEW JENSEN-TYPE INEQUALITIES

CONSTANTIN P. NICULESCU AND CĂTĂLIN IRINEL SPIRIDON

Abstract. We develop a new framework for the Jensen-type inequalities that allows us to deal with functions not necessarily convex and Borel measures not necessarily positive.

It is well known the important role played by the classical inequality of Jensen in probability theory, economics, statistical physics, information theory etc. See [5] and [6]. In recent years, a number of authors have noticed the possibility to extend this inequality to the framework of functions that are mixed convex (in the sense of the existence of one inflection point). See [1], [2] and [4]. In all these papers one assumes that both the function and the measure under consideration verify certain conditions of symmetry. However the inequality of Jensen is much more general as shows the following simple remark. Suppose that K is a convex subset of the Euclidean space $\mathbb{R}^N$ carrying a Borel probability measure μ. Then every μ-integrable function $f : K \to \mathbb{R}$ that admits a supporting hyperplane at the barycenter of μ,

$$(B) \qquad b_\mu = \int_K x d\mu(x),$$

verifies the Jensen inequality

$$(J) \qquad f(b_\mu) \leq \int_K f(x) d\mu(x).$$

Indeed, the existence of a supporting hyperplane at b_μ is equivalent to the existence of an affine function $h(x) = \langle x, v \rangle + c$ such that

$$f(b_\mu) = h(b_\mu) \text{ and } f(x) \geq h(x) \text{ for all } x \in K.$$

Then

$$f(b_\mu) = h(b_\mu) = h\left(\int_K x d\mu(x)\right) = \int_K h(x) d\mu(x) \leq \int_K f(x) d\mu(x).$$

As is well known, the convexity assures the existence of a supporting hyperplane at each interior point. See [5], Theorem 3.7.1, p. 128. This explains why Jensen's inequality works nicely in that context. The aim of our paper is to extend the validity of Jensen's inequality outside mixed convexity and also outside the framework of Borel probability measures.

Date: March 14, 2012.

2000 *Mathematics Subject Classification.* Primary 26A51, 26 D15; Secondary 28A25.

Key words and phrases. Jensen's inequality, convex function, almost convex function, supporting hyperplane.

Published in J. Math. Anal. Appl. **401** (2013), No. 1, 343–348.

DOI: 10.1016/j.jmaa.2012.11.051 .

In order to make our approach easily understandable we will restrict ourselves to the case of functions of one real variable. However most of our results extends easily to higher dimensions, by replacing the usual intervals by N-dimensional intervals and symmetry with respect to a point by symmetry with respect to a hyperplane. See Example 3 below.

We start with the following version of the Jensen inequality for mixed convex functions that discards any assumption on the symmetry of the involved measure.

Theorem 1. *Suppose that f is a real-valued function defined on an interval $[a,b]$ and c is a point in $[a, \frac{a+b}{2}]$ such that:*
i) $f(c-x)+f(c+x)=2f(c)$ whenever $c \pm x \in [a,b]$;
ii) $f|_{[c,b]}$ is convex.
Then

$$f\left(b_\mu\right) \leq \int_a^b f(x)d\mu(x),$$

for every Borel probability measure μ on $[a,b]$ whose barycenter lies in the interval $[2c-a,b]$.

The last inequality works in the reverse direction when $f|_{[c,b]}$ is concave.

Proof. The case where $c=a$ is covered by the classical inequality of Jensen so we may assume that $c \in (a, \frac{a+b}{2})$. In this case the point $2c-a$ is interior to $[a,b]$. By our hypotheses, the barycenter b_μ lies in the interval $[2c-a,b]$. If $b_\mu = b$, then $\mu = \delta_b$ and the conclusion of Theorem 1 is clear. If b_μ is interior to $[a,b]$, we will denote by h the affine function joining the points $(a, f(a))$ and $(2c-a, f(2c-a))$ and we will consider the function

$$g(x) = \begin{cases} h(x) & \text{if } x \in [a, 2c-a] \\ f(x) & \text{if } x \in [2c-a, b]. \end{cases} \tag{1}$$

Clearly, g is convex and this fact motivates the existence of a support line ℓ of g at b_μ. See [5], Lemma 1.5.1, p. 30. Since $h \geq f$, then necessarily ℓ is a support line at b_μ also for f. By a remark above, this ends the proof. ∎

A useful consequence of Theorem 1 in the case of absolutely continuous measures is as follows:

Corollary 1. *Suppose that $f:[-b,b] \to \mathbb{R}$ is an odd function, whose restriction to $[0,b]$ is convex and $p:[-b,b] \to [0,\infty)$ is a nondecreasing function that does not vanish on $(-b/3,b]$. Then for every $a \in [-b/3,b)$,*

$$f\left(\frac{1}{\int_a^b p(x)dx}\int_a^b xp(x)dx\right) \leq \frac{1}{\int_a^b p(x)dx}\int_a^b f(x)p(x)dx.$$

Proof. The case where $a \geq 0$ is covered by the classical inequality of Jensen.

If $a < 0$, then

$$\begin{aligned}
\int_a^b (x+a)p(x)dx &\geq \int_a^{-3a} (x+a)p(x)dx \\
&= \int_a^{-a} (x+a)p(x)dx + \int_{-a}^{-3a} (x+a)p(x)dx \\
&\geq \int_a^{-a} (x+a)p(x)dx + p(-a)\int_{-a}^{-3a} (x+a)dx \\
&= \int_a^{-a} (x+a)p(x)dx - p(-a)\int_a^{-a} (x+a)dx \\
&= \int_a^{-a} (x+a)\left(p(x)-p(-a)\right)dx \geq 0,
\end{aligned}$$

and thus Theorem 1 applies. □

An inspection of the argument of Corollary 1 shows that the monotonicity hypothesis on p can be relaxed by asking only the integrability of p and the fact that $p(x) \leq p(-a) \leq p(y)$ for all x and y with $x \leq -a \leq y$. However, simple examples show that the restriction $a \in [-b/3, b]$ in Corollary 1 cannot be dropped.

Consider now the discrete version of Theorem 1.

Corollary 2. *Suppose that f is a real-valued function defined on an interval I that contains the origin such that $f|_{I\cap[0,\infty)}$ is a convex function and $f(-x) = -f(x)$ whenever x and $-x$ belong to I. Then for every family of points $a_1, ..., a_n$ of I and every family of weights $p_1, ..., p_n \in [0,\infty)$ such that $\sum_{k=1}^n p_k = 1$ and*

$$\sum_{k=1}^n p_k a_k + \min\{a_1, ..., a_n\} \geq 0,$$

we have

$$f\left(\sum_{k=1}^n p_k a_k\right) \leq \sum_{k=1}^n p_k f(a_k).$$

The conclusion of Corollary 2 can be considerably improved when all weights p_k are equal.

Corollary 3. *Suppose that f is a real-valued function defined on an interval I that contains the origin. If $f|_{I\cap[0,\infty)}$ is a convex function and $f(-x) = -f(x)$ whenever x and $-x$ belong to I, then for every family of points $a_1, ..., a_n$ of I such that*

$$\sum_{k=1}^n a_k + (n-2)\min\{a_1, ..., a_n\} \geq 0$$

we have

$$f\left(\frac{1}{n}\sum_{k=1}^n a_k\right) \leq \frac{1}{n}\sum_{k=1}^n f(a_k).$$

Proof. It suffices to consider the case where $a_1 \leq \cdots \leq a_n$ and $a_1 < 0$. According to our hypothesis, $\sum_{k=1}^n a_k > 0$ and

$$\frac{1}{n-1}\sum_{k=2}^n a_k \geq -a_1 \geq -a_2.$$

By Corollary 2,

$$f\left(\frac{1}{n-1}\sum_{k=2}^{n} a_k\right) \leq \frac{1}{n-1}\sum_{k=2}^{n} f(a_k),$$

and taking into account the function g given by the formula (1) we infer that

$$\begin{aligned} f\left(\frac{1}{n}\sum_{k=1}^{n} a_k\right) \leq g\left(\frac{1}{n}\sum_{k=1}^{n} a_k\right) &= g\left(\frac{1}{n}\cdot a_1 + \frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{k=2}^{n} a_k\right) \\ &\leq \frac{1}{n}g(a_1) + \frac{n-1}{n}g(\frac{1}{n-1}\sum_{k=2}^{n} a_k) \\ &= \frac{1}{n}f(a_1) + \frac{n-1}{n}f(\frac{1}{n-1}\sum_{k=2}^{n} a_k) \leq \frac{1}{n}\sum_{k=1}^{n} f(a_k). \end{aligned}$$

The proof is complete. □

A simple example illustrating Corollary 3 is

$$\tan\left(\frac{x+y+z}{3}\right) \leq \frac{\tan x + \tan y + \tan z}{3},$$

for every $x, y, z \in (-\pi/6, \pi/2)$ with $x + y + z + \min\{x, y, z\} \geq 0$.

Starting with the pioneering work of J. F. Steffensen [8], a great deal of research was done to extend the Jensen inequality outside the framework of probability measures. An account on the present state of art can be found in the monograph [5], Sections 4.1 and 4.2. We will recall here some basic facts for the convenience of the reader.

Definition 1. A *Steffensen-Popoviciu measure* is any real Borel measure μ on a compact convex set K such that $\mu(K) > 0$ and

$$\int_K f(x)\, d\mu(x) \geq 0 \quad \text{for every continuous convex function} \quad f : K \to \mathbb{R}_+.$$

In the case of intervals, a complete characterization of this class of measures is offered by the following result, independently due to T. Popoviciu [7], and A. M. Fink [3]:

Lemma 1. *Let μ be a real Borel measure on an interval $[a, b]$ with $\mu([a, b]) > 0$. Then μ is a Steffensen-Popoviciu measure if, and only if, it verifies the following condition of endpoints positivity,*

$$\int_a^t (t - x)\, d\mu(x) \geq 0 \text{ and } \int_t^b (x - t)\, d\mu(x) \geq 0$$

for every $t \in [a, b]$.

See [5], p. 179, for details.

Example 1. *A discrete measure $\mu = \sum_{k=1}^{n} p_k\, \delta_{x_k}$ (supported by the points $x_1 \leq ... \leq x_n$) is a Steffensen-Popoviciu measure if it verifies Steffensen's condition*

$$(2) \qquad \sum_{k=1}^{n} p_k > 0, \quad \text{and} \quad 0 \leq \sum_{k=1}^{m} p_k \leq \sum_{k=1}^{n} p_k, \quad \text{for every } m \in \{1, ..., n\}.$$

A concrete example is offered by the discrete measure $\frac{5}{9}\delta_{\frac{3a+b}{4}} - \frac{1}{9}\delta_{\frac{a+b}{2}} + \frac{5}{9}\delta_{\frac{a+3b}{4}}$.

Example 2. *According to Lemma 1,* $\left(\frac{x^2}{a^2}-\frac{1}{6}\right)dx$ *is an example of absolutely continuous measure on the interval* $[-a,a]$ *that is also a Steffensen-Popoviciu measure. As a consequence,* $\left(\frac{x^2}{a^2}-\frac{1}{6}\right)dxdy$ *provides an example of Steffensen-Popoviciu measure on any rectangle* $[-a,b]\times[c,d]$ *with* $0<a<b$ *and* $c<d$. *Indeed, if* $h:[-a,b]\times[c,d]\to\mathbb{R}$ *is a nonnegative convex function, then*

$$x\to\left(\int_c^d h(x,y)dy\right)$$

is also a nonnegative convex function, whence

$$\int_{-a}^{a}\int_c^d h(x,y)\left(\frac{x^2}{a^2}-\frac{1}{6}\right)dxdy\geq 0.$$

This yields

$$\int_{-a}^{b}\int_c^d h(x,y)\left(\frac{x^2}{a^2}-\frac{1}{6}\right)dxdy\geq 0,$$

and thus $\left(\frac{x^2}{a^2}-\frac{1}{6}\right)dxdy$ *is a Steffensen-Popoviciu measure on the rectangle* $[-a,b]\times[c,d]$. *Since this class of measures is closed under addition, we infer that*

$$\left(\frac{x^2}{a^2}+\frac{y^2}{c^2}-\frac{1}{3}\right)dxdy$$

is a Steffensen-Popoviciu measure on the rectangle $[-a,b]\times[-c,d]$, *whenever* $0<a<b$ *and* $0<c<d$.

The Steffensen-Popoviciu measures provide the natural framework for the Jensen inequality:

Theorem 2. *Suppose that* μ *is a Steffensen-Popoviciu measure on a compact convex set* K *(part of a locally convex separated space* E*). Then* μ *admits a barycenter* b_μ *and for every continuous convex function* f *on* K,

$$f(b_\mu)\leq\frac{1}{\mu(K)}\int_K f(x)\,d\mu(x).$$

For details, see [5], Theorem 4.2.1, pp. 184-185. When E is the Euclidean space $\mathbb{R}^N$, the barycenter b_μ is given by the formula (B) above.

It is worth to mention that the argument of Theorem 1 remains valid in the context of Steffensen-Popoviciu measures of total mass 1. Even more importantly, it can be adapted to the case of functions of two or more variables.

Example 3. *Suppose* $f:[-1,2]\times[-1,1]\to\mathbb{R}$ *is a function with the following two properties: i)* $f(-x,y)=-f(x,y)$ *for every* $x\in[-1,1]$ *and* $y\in[-1,1]$*; and ii)* $f|_{[0,2]\times[-1,1]}$ *is a convex function. According to Example 2,* $(x^2-\frac{1}{6})dxdy$ *is a Steffensen-Popoviciu measure on the rectangle* $[-1,2]\times[-1,1]$ *(of total mass* $\frac{10}{3}$ *and barycenter* $b_\mu=\left(\frac{7}{5},0\right)$*). Since* $\frac{7}{5}>1$, *an argument similar to that used in the proof of Theorem 1 leads us to the following Jensen-type inequality:*

$$f\left(\frac{7}{5},0\right)\leq\frac{3}{10}\iint_{[-1,2]\times[-1,1]}f(x,y)(x^2-\frac{1}{6})dxdy.$$

The above discussion leads us to the concept of almost convexity, whose 1-dimensional version is as follows:

Definition 2. *A real-valued function f defined on an interval I is left almost convex if it is integrable and there is a pair of interior points $c < d$ in I such that*

i) $f|_{[c,\infty)\cap I}$ *is convex; and*

ii) $f \geq h$ *on* $(-\infty, c]\cap I$, *where h is the affine function joining the points $(c, f(c))$ and $(d, f(d))$.*

The concept of right almost convexity can be introduced in a similar way.

The following result extends Theorem 1 above.

Theorem 3. *Suppose that $f : [a, b] \to \mathbb{R}$ is a left almost convex function and μ is a Steffensen-Popoviciu measure on $[a, b]$ with barycenter b_μ. If $c < d$ are interior points to $[a, b]$ as in Definition 2 such that*

i) $b_\mu \geq c$; *and*

ii) $\int_a^d ((f(x) - f(c))(d - c) - (f(d) - f(c))(x - c))\, d\mu \geq 0$,

then

$$f(b_\mu) \leq \frac{1}{\mu([a, b])} \int_a^b f(x)\, d\mu(x).$$

Proof. If h is the affine function joining the points $(c, f(c))$ and $(d, f(d))$, then the function

$$g(x) = \begin{cases} h(x) & \text{if } x \in [a, d] \\ f(x) & \text{if } x \in [d, b] \end{cases}$$

is convex and $g(b_\mu) = f(b_\mu)$. According to Theorem 2,

$$\begin{aligned} f(b_\mu) = g(b_\mu) &\leq \frac{1}{\mu([a, b])} \int_a^b g(x)\, d\mu(x) \\ &= \frac{1}{\mu([a, b])} \int_a^d h(x)\, d\mu(x) + \frac{1}{\mu([a, b])} \int_d^b f(x)\, d\mu(x) \\ &\leq \frac{1}{\mu([a, b])} \int_a^d f(x)\, d\mu(x) + \frac{1}{\mu([a, b])} \int_d^b f(x)\, d\mu(x) \\ &= \frac{1}{\mu([a, b])} \int_a^b f(x)\, d\mu(x) \end{aligned}$$

and the proof is complete. ∎

An illustration of Theorem 3 is offered by the following constrained optimization problem: Find

$$M = \max_{(x,y,z)\in\Omega} \left[\tan\left(\frac{2x - y + 3z}{4}\right) - \frac{2\tan x - \tan y + 3\tan z}{4}\right],$$

where Ω is the set of triplets $x \leq y \leq z$ in $[-\pi/3, \pi/3)$ such that $2x - y + 3z \geq 0$ and

$$\frac{\pi}{3\sqrt{3}} (2\tan x - \tan y + 3\tan z) \geq 2x - y + 3z. \tag{3}$$

The answer is $M = 0$. Indeed, according to (2), the measure $\frac{1}{2}\delta_x - \frac{1}{4}\delta_y + \frac{3}{4}\delta_z$ is Steffensen-Popoviciu on the interval $[-\pi/3, \pi/3]$ and the constraint (3) coincides with the inequality *ii)* in Theorem 3 (when applied to the tangent function for $a = -\pi/3$, $c = 0$ and $b = d = \pi/3$).

Of course, the phenomenon of almost convexity is present also in higher dimensions, at least for functions defined on N-dimensional intervals (or on other convex sets with a special geometry). It would be interesting in that context to prove a characterization of the Steffensen-Popoviciu measures (comparable to that offered by Lemma 1).

Acknowledgement. We acknowledge useful correspondence with Florin Popovici. This work is supported by a grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-II-ID-PCE-2011-3-0257.

References

[1] P. Czinder, A weighted Hermite-Hadamard type inequality for convex-concave symmetric functions, Publ. Math. Debrecen 68 (2006), 215-224.

[2] P. Czinder and Z. Páles, An extension of the Hermite-Hadamard inequality and an application for Gini and Stolarsky means, JIPAM 5 (2004), Issue 2, Article no. 42.

[3] A. M. Fink, A best possible Hadamard Inequality, Math. Inequal. Appl., 1 (1998), 223-230.

[4] A. Florea and C. P. Niculescu, A Hermite-Hadamard inequality for convex-concave symmetric functions, Bull. Math. Soc. Sci. Math. Roumanie 50 (98), No. 2, 2007, 149-156. 50 (98), No. 2, 2007, 149-156.

[5] C. P. Niculescu and L.-E. Persson, Convex Functions and their Applications. A Contemporary Approach. CMS Books in Mathematics vol. 23, Springer-Verlag, New York, 2006.

[6] J. E. Pečarić, F. Proschan and Y. L. Tong, Convex functions, partial orderings, and statistical applications, Academic Press, 1992.

[7] T. Popoviciu, Notes sur les fonctions convexes d'ordre superieur (IX), Bull. Math. Soc. Roum. Sci., 43 (1941), 85-141.

[8] J. F. Steffensen, On certain inequalities and methods of approximation, J. Inst. Actuaries 51 (1919), 274-297.

University of Craiova, Department of Mathematics, Street A. I. Cuza 13, Craiova RO-200585, Romania

E-mail address: cniculescu47@yahoo.com

University of Craiova, Department of Mathematics, Street A. I. Cuza 13, Craiova, RO-200585, Romania

E-mail address: catalin_gsep@yahoo.com