sciendo

DOI 10.2478/bile-2023-0011

Biometrical Letters
Vol. 60 (2023), No. 2, 159-175

Fujiwara's inequality for synchronous functions and its consequences

Zdzisław Otachel

Department of Applied Mathematics and Computer Science, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland,
e-mail: zdzislaw.otachel@up.lublin.pl

Summary

The theory of inequalities is an essential tool in all fields of the theoretical and practical sciences, especially in statistics and biometrics. Fujiwara's inequality provides relationships between expected values of products of random variables. Special cases of this inequality include important classical inequalities such as the Cauchy–Schwarz and Chebyshev inequalities.
The purpose of this article is to prove Fujiwara's inequality in abstract probability spaces with more general assumptions for random variables than those associated with monotonicity as found in the literature. This newly introduced property of random variables will be called synchronicity. As a consequence, we obtain inequalities originated by Chebyshev, Hardy–Littlewood–Pólya and Jensen–Mercer under new conditions.
The results are illustrated by some examples constructed for specific probability measures and specific random variables. In biometrics, the obtained inequalities can be used wherever there is a need to compare the characteristics of experimental data based on means or moments, and the data can be modeled using functions that are synchronous or convex.

Key words: Fujiwara's inequality, Chebyshev's inequality, Hardy–Littlewood–Pólya inequality, Jensen–Mercer inequality, convex function, synchronous functions

1. Introduction

Let x, y, z, v be random variables and E, Cov be operations of expectation and covariance, respectively.

A version of Chebyshev's inequality states that

$$E(xy) \geqslant E(x)E(y), \text{ or equivalently, } Cov(x, y) \geqslant 0,$$

where x and y are similarly ordered random variables, and is related to the theory of positive dependence; see e.g. Shaked (1982).

In turn, the famous Cauchy–Schwarz inequality reads

$$|E(zv)|^2 \leqslant E(z^2)E(v^2).$$

Fujiwara (1918) (see also Mitrinović, 1970) proved an inequality that covers these both mentioned: if $y(t)/x(t)$ and $z(t)/v(t)$ are monotone in the same sense for $t \in (a,b)$, then

$$E(xz)E(yv) \leqslant E(yz)E(xv), \tag{1}$$

where $E(f) = \frac{1}{b-a}\int_a^b f(t)dt$.

Setting $x = v \equiv 1$, we get Chebyshev's inequality, while for $y = z$ and $x = v$ we obtain the Cauchy–Schwarz inequality (see e.g. Mitrinović, 1970, sec. 2.5–2.6 or Pečarić et al., 1992, sec. 7.1).

Toader (1992b) proved Fujiwara's inequality for arbitrary linear and positive functionals E under the following condition:

$$\begin{vmatrix} y(s) & y(t) \\ x(s) & x(t) \end{vmatrix} \begin{vmatrix} z(s) & z(t) \\ v(s) & v(t) \end{vmatrix} \geqslant 0, \text{ for all } s,t. \tag{2}$$

Note that for positive x and v, (2) is equivalent to

$$[y(s)/x(s) - y(t)/x(t)][z(s)/v(s) - z(t)/v(t)] \geqslant 0, \text{ for all } s,t, \tag{3}$$

i.e. y/x and z/v are similarly ordered.

A variant of Fujiwara's inequality for Lebesgue integrals on the unit interval is proved in Otachel (2011). Discrete versions of (1) were considered in Toader (1992a), Niezgoda (2006), and Otachel (2011). In fact (see Otachel, 2011, Corollary 4),

$$\sum_{i=1}^{\infty} x_i z_i \sum_{i=1}^{\infty} y_i v_i \leqslant \sum_{i=1}^{\infty} y_i z_i \sum_{i=1}^{\infty} x_i v_i \tag{4}$$

holds if there exists a scalar γ such that

$$\left[\frac{y_i}{x_i} - \gamma\right]\left[\frac{z_i}{v_i} - \frac{\sum_{k=1}^{\infty} x_i z_i}{\sum_{k=1}^{\infty} x_i v_i}\right] \geqslant 0, \text{ for } i = 1,2,\dots, \tag{5}$$

where $\overline{x} = (x_1,\dots), \overline{y} = (y_1,\dots) \in l^p$ and $\overline{z} = (z_1,\dots), \overline{v} = (v_1,\dots) \in l^q$, $p,q > 1$, $1/p + 1/q = 1$ with positive $\overline{x}$ and $\overline{v}$.

In the paper we give an extension of the above result for random variables with p-moments. Consequently, we obtain new variants of inequalities by Chebyshev, Hardy–Littlewood–Pólya (Hardy et al., 1929) and Jensen–Mercer (Mercer, 2003).

The organization of the paper is as follows. In section 2 we introduce a notion of synchronicity for real random variables and prove our variant of Fujiwara's inequality (Theorem 1). A related notion called synchronicity w.r.t. dual bases in Banach spaces was utilized in Otachel (2011). It is worth noting that the introduced notion of synchronicity is a wider notion than the relation of similar ordering for functions (Remark 1). As a consequence, we obtain Chebyshev's inequality for synchronous functions defined on a probability space (Proposition 1). It is known that the Chebyshev-type inequalities are key tools to produce variants of HLP inequalities for convex functions. This method was originated by Dragomir (2004) and continued in Niezgoda (2008), Barnett et al. (2009) and Otachel (2014). Such a version of an HLP inequality will be proved in section 3 (Theorem 2). It extends some recent results by Niezgoda (2008) and Barnett et al. (2009). Moreover, applications for particular convex functions are presented there (Example 3). Section 4 is devoted to the Jensen–Mercer inequality. We give there (Theorem 3) an extension of a result of Niezgoda (2009).

2. Synchronous functions and Fujiwara's inequality

Let μ be a probability measure on a space Ω. Unless otherwise stated, throughout the paper all functions and sets under considerations are μ-measurable. Here and subsequently, L^p, $p \geqslant 1$, denotes the space of all random variables with finite p-moments. As usual, L^∞ stands for the space of all functions on Ω which are essentially bounded. The symbol $E := E_\mu$ stands for the operation expected value $E_\mu(x) = \int_\Omega x d\mu$, $x \in L^1$, and I_Z is the index function of a set Z.

Two random variables f and g are said to be: synchronous (abbrev. $f \sim g$) if

$$f \sim g \equiv f(\omega)g(\omega) \geqslant 0, \quad \text{for all } \omega \in \Omega;$$

and similarly ordered if

$$[f(\omega) - f(\sigma)][g(\omega) - g(\sigma)] \geqslant 0, \quad \text{for all } \omega, \sigma \in \Omega.$$

Observe that

$$f \sim g \Rightarrow E(fg) \geqslant 0. \tag{6}$$

The following lemma brings together some properties of synchronous and similarly ordered functions.

Lemma 1.

i) *If f and g are similarly ordered and g is bounded, then there exists a scalar γ such that $f \sim g - \gamma$.*

ii) *$f \sim g$ if and only if there exist $T, R \subset \Omega,\ T \cup R = \Omega,\ T \cap R = \emptyset$, (i.e. T, R – complementary subsets of Ω) with $f(\varrho), g(\varrho) \leqslant 0 \leqslant f(\tau), g(\tau),\ \tau \in T,\ \varrho \in R$.*

iii) *Given three functions f, g, h, where g and h are similarly ordered, $g(\omega) = g(\sigma)$ implies $h(\omega) = h(\sigma)$ for $\omega, \sigma \in \Omega$ and h is bounded.*

If $f \sim g$, then there exists a scalar γ such that $f \sim h - \gamma$.

Proof.

i) Set $T = \{\omega : 0 \leqslant f(\omega)\}$ and let R be the complement of T. If $T \neq \emptyset$, then we define $\gamma = \inf\{g(\tau) : \tau \in T\}$ else $\gamma = \sup\{g(\omega) : \omega \in \Omega\}$. Clearly, $\gamma \leqslant g(\tau),\ \tau \in T$. We will show that $g(\varrho) \leqslant \gamma,\ \varrho \in R$. This is simply $f \sim g - \gamma$, because $f(\varrho) < 0 \leqslant f(\tau),\ \tau \in T, \varrho \in R$.

Suppose that there exists $\varrho \in R$ with $\gamma < g(\varrho)$. This forces the existence of $\tau \in T$ such that $\gamma < g(\tau) < g(\varrho)$, so $g(\tau) - g(\varrho) < 0$ for certain $\tau \in T,\ \varrho \in R$. On the other hand, $f(\varrho) < 0 \leqslant f(\tau)$. Hence $[f(\tau) - f(\varrho)][g(\tau) - g(\varrho)] < 0$, a contradiction.

ii) Set $T = \{\omega : f(\omega) > 0\} \cup \{\omega : f(\omega) = 0,\ g(\omega) \geqslant 0\}$ and let $R = \Omega \setminus T$, i.e. $R = \{\omega : f(\omega) < 0\} \cup \{\omega : f(\omega) \leqslant 0,\ g(\omega) < 0\}$. Clearly, $f(\varrho) \leqslant 0 \leqslant f(\tau),\ \tau \in T,\ \varrho \in R$.

If $f \sim g$, then $f(\omega) > 0$ implies $g(\omega) \geqslant 0$ and $f(\omega) < 0$ implies $g(\omega) \leqslant 0$. Therefore $g(\varrho) \leqslant 0 \leqslant g(\tau),\ \tau \in T,\ \varrho \in R$.

The converse implication is evident.

iii) By ii), if $f \sim g$, then there exist complementary subsets $T, R \subset \Omega$, such that $f(\varrho), g(\varrho) \leqslant 0 \leqslant f(\tau), g(\tau),\ \tau \in T,\ \varrho \in R$. In particular,

$g(\varrho) < g(\tau)$ for $\tau \in T, \varrho \in R$ with $g(\tau) \neq 0$ or $0 \neq g(\varrho)$. Hence $h(\varrho) \leqslant h(\tau)$, because g and h are similarly ordered. If $g(\tau) = 0 = g(\varrho)$, then $h(\tau) = h(\varrho)$, by the hypothesis. Consequently, $h(\varrho) \leqslant h(\tau)$ for all $\tau \in T$ and $\varrho \in R$. If $T \neq \emptyset$, then we define $\gamma = \inf\{h(\tau) : \tau \in T\}$, else $\gamma = \sup\{h(\omega) : \omega \in \Omega\}$. It is easily seen that $h(\varrho) \leqslant \gamma \leqslant h(\tau),\ \tau \in T,\ \varrho \in R$. Finally, we have $f(\tau) \geqslant 0,\ h(\tau) \geqslant \gamma,\ \tau \in T$ and $f(\varrho) \leqslant 0,\ h(\varrho) \leqslant \gamma,\ \varrho \in R$, or equivalently, $f(\omega)[h(\omega) - \gamma] \geqslant 0$ for all $\omega \in \Omega$. □

Now, we are in a position to give a variant of (1) that holds in abstract probability spaces.

Theorem 1. *(cf. Otachel, 2011, Proposition 2) Fix $x \in L^p$ and $v \in L^q$ with $E(xv) > 0$, where $p, q \geqslant 1$ and $1/p + 1/q = 1$. If for $y \in L^p$ and $z \in L^q$ there exists a scalar γ such that*

$$y - \gamma x \sim z - \lambda v,$$

where $\lambda = E(xz)/E(xv)$, then (1) holds for $x, y \in L^p$ and $z, v \in L^q$.

Proof. If $y - \gamma x \sim z - \lambda v$, then by (6) and the linearity of A we easily obtain

$$0 \leqslant E\left([y - \gamma x][z - \lambda v]\right) = \frac{1}{A(xv)}\left[E(yz)E(xv) - E(xz)E(yv)\right].$$

Since $E(xv) > 0$, inequality (1) holds. □

For positive random variables x and v, the above theorem can be immediately restated in the following interesting form.

Corollary 1. *(cf. Otachel, 2011, Corollary 4) Let $x, y \in L^p$ and $z, v \in L^q$ with positive x and v and $E(xv) > 0$, where $p, q \geqslant 1$ and $1/p + 1/q = 1$. Inequality (1) holds whenever*

$$\exists_{\gamma \in \mathbb{R}} :\ \left[\frac{y(\omega)}{x(\omega)} - \gamma\right]\left[\frac{z(\omega)}{v(\omega)} - \frac{E(xz)}{E(xv)}\right] \geqslant 0,\ \text{for all } \omega \in \Omega. \tag{7}$$

Now we will show that condition (7) is significantly weaker than (3). Let us start with an example.

Example 1. *Assume $\Omega = \bigcup_{k=1}^{m} B_k$ for even $m \geqslant 4$, where $B_k \cap B_l = \emptyset$ iff $k \neq l$. Define the probability measure μ on $\sigma(B_1, \ldots, B_m)$, i.e. on the σ-field generated by $B_1, \ldots, B_m$, as follows:*

$$\mu(B_k) = 1/m,\ k = 1, \ldots, m.$$

Set random variables x, v, z, y as follows

$$\begin{aligned} x &= v = I_\Omega \\ z &= \textstyle\sum_{k=1}^{m} k I_{B_k} \\ y &= \textstyle\sum_{i\in J_1} \left[\frac{3m}{2} + 1 - i\right] I_{B_i} + \sum_{j\in J_2} \left[\frac{m}{2} + 1 - j\right] I_{B_j}, \end{aligned}$$

where $J_1 = \{\frac{m}{2} + 1, \ldots, m\}$ and $J_2 = \{1, \ldots, \frac{m}{2}\}$. Write $T = \bigcup_{i\in J_1} B_i$ and $R = \bigcup_{j\in J_2} B_j$. Clearly, $\Omega = T \cup R$ and $T \cap R = \emptyset$ and

$$z(\varrho), y(\varrho) < \frac{E(xz)}{E(xv)} = \frac{m+1}{2} < z(\tau), y(\tau),\ \tau \in T,\ \varrho \in R.$$

Hence (7) is satisfied with $\gamma = E(xz)/E(xv)$, by Lemma 1 ii).

On the other hand, $y/x \equiv y$ and $z/v \equiv z$ are not similarly ordered, because

$$[y(s) - y(t)][z(s) - z(t)] < 0,\ s \in B_1, t \in B_{m/2}.$$

However, inequality (1) holds for x, y, z, v as defined above, by Corollary 1.

Remark 1. *Under the assumptions of Corollary 1, inequality (1) holds whenever y/x and z/v are similarly ordered, provided y/x is bounded. Moreover, (7) is weaker than (3).*

Proof. It is clear that y/x and z/v are similarly ordered if and only if y/x and $z/v - E(xz)/E(xv)$ are similarly ordered. Assume that y/x and z/v are similarly ordered, where y/x is bounded. By Lemma 1 i), condition (7) is satisfied and inequality (1) is a consequence of Corollary 1. Thus (3) implies (7), but not vice versa, as illustrated by Example 1. □

As was mentioned above, Chebyshev's inequality is a consequence of Fujiwara's inequality. In fact, applying Corollary 1 for $x = v \equiv 1$ gives

Proposition 1. *Let μ be a probability measure on a space Ω and $y \in L^p, z \in L^q$, where $p, q \geqslant 1$ and $1/p + 1/q = 1$. Chebyshev's inequality*

$$E(y)E(z) \leqslant E(yz)$$

holds whenever there exists a scalar γ such that

$$[y(\omega) - \gamma]\,[z(\omega) - E(z)] \geqslant 0, \text{ for all } \omega \in \Omega. \tag{8}$$

The conditions (7) and (8) originate from Niezgoda (2006), Corollary 4.1 and Example 4.2. He obtained them for the case of $\mathbb{R}^n$. Following Remark 1, condition (8) is satisfied by similarly ordered y and z if only y is bounded.

3. HLP inequality for synchronous functions

A classic result of the theory of majorization due to Hardy et al. (1929) (see also Marshall et al., 1979; Mitrinović, 1970, sec. 2.24; Pečarić et al., 1992) states that

$$E_\mu(f \circ y) \leqslant E_\mu(f \circ x) \tag{9}$$

if and only if y is majorized by x, where f is a convex function on $\mathbb{R}$, μ is the Lebesgue measure on the unit interval, and $x, y : \mathbb{R} \mapsto \mathbb{R}$ are bounded random variables. Recent contributions appear in Dragomir (2004), Niculescu et al. (2006), Niezgoda (2008), Barnett et al. (2009), Otachel (2014).

Now we recall some relevant results on convex functions; see Zălinescu (2002), Theorem 2.1.5. Let $f : I \mapsto \mathbb{R}$ be a convex function on an open interval $I \subset \mathbb{R}$. The function f is continuous on I. There exist finite $f'_-(t)$ and $f'_+(t)$, i.e. the left-hand and the right-hand derivative of first order, respectively, for all $t \in I$. If $t_1 < t_2$, then $f'_-(t_1) \leqslant f'_+(t_1) \leqslant f'_-(t_2) \leqslant f'_+(t_2)$. The subdifferential of f denoted by ∂f is the set of all functions $\phi : I \mapsto \mathbb{R}$ such that $f'_-(t) \leqslant \phi(t) \leqslant f'_+(t)$. Every $\phi \in \partial f$ is a nondecreasing function. If f is nondecreasing convex, then f'_- and f'_+ are both nonnegative, showing that ϕ is also nonnegative. Moreover, for every $s, t \in I$ and $\phi \in \partial f$

$$f(s) - f(t) \geqslant (s - t)\phi(t). \tag{10}$$

Chebyshev's inequality from Proposition 1 is crucial for the following variant of (9).

Theorem 2. *Let μ be a finite measure on Ω and $f : I \mapsto \mathbb{R}$ be a convex function on an open interval $I \subset \mathbb{R}$. Inequality (9) holds for bounded functions x and y on Ω with values in I if*

$$\exists_{\gamma \in \mathbb{R}} : \ [x(\omega) - y(\omega) - E(x - y)]\,[y(\omega) - \gamma] \geqslant 0, \ \text{for all } \omega \in \Omega \tag{11}$$

and

$$E(x) = E(y) \quad \textit{or} \quad \begin{cases} f \text{ is nondecreasing} \\ E(x) \geqslant E(y). \end{cases} \tag{12}$$

Proof. Fix $\phi \in \partial f$. Substituting $s = x(\omega), t = y(\omega)$ into (10) gives

$$(f \circ x)(\omega) - (f \circ y)(\omega) \geqslant [x(\omega) - y(\omega)](\phi \circ y)(\omega), \ \text{for } \omega \in \Omega. \tag{13}$$

Since f is continuous and x and y are bounded, $f \circ x, f \circ y, \phi \circ y$ are also bounded. Consequently, there exist finite expectations $E(f \circ x), E(f \circ y), E([x-y]\phi \circ y)$. Integrating (13), we obtain

$$E(f \circ x) - E(f \circ y) \geqslant E([x-y]\phi \circ y). \tag{14}$$

Since ϕ is nondecreasing, $y-\gamma$ and $\phi \circ y$ are similarly ordered for any scalar γ. Moreover, $y(\omega)-\gamma = y(\sigma)-\gamma$ implies $(\phi \circ y)(\omega) = (\phi \circ y)(\sigma)$ for $\omega, \sigma \in \Omega$. Hence, (11) gives

$$\exists_{\gamma' \in \mathbb{R}} : \ [x(\omega) - y(\omega) - E(x-y)]\,[(\phi \circ y)(\omega) - \gamma'] \geqslant 0, \ \text{for all } \omega \in \Omega,$$

by Lemma 1 iii). It follows that $E([x-y]\phi \circ y) \geqslant E(x-y)E(\phi \circ y)$, by Chebyshev's inequality (according to Proposition 1). Clearly, ϕ is nonnegative if f is nondecreasing. Hence, (12) forces $E([x-y]\phi \circ y) \geqslant 0$. Combining (14) and the last inequality gives (9). This is the desired result. □

In accordance with earlier considerations, note that

Remark 2. *Under assumptions as in the above theorem, (11) is satisfied if $x-y$ and y are similarly ordered.*

Example 2. *Assume $\Omega = \bigcup_{k=1}^{m} B_k$ for an integer m, where $B_k \cap B_l = \emptyset$ iff $k \neq l$.*

For given real m-tuples $\overline{x} = (x_1, \ldots, x_m)$, $\overline{y} = (y_1, \ldots, y_m)$ and $\overline{p} = (p_1, \ldots, p_m)$, where $\overline{p}$ is positive and $\sum_{j=1}^{m} p_j = 1$, we define the probability measure μ on $\sigma(B_1, \ldots, B_m)$ and random variables x and y on Ω as follows:

$$\begin{aligned} \mu(B_j) &= p_j, \ j = 1, \ldots, m, \\ x(\omega) &= \textstyle\sum_j x_j I_{B_j}(\omega), \\ y(\omega) &= \textstyle\sum_j y_j I_{B_j}(\omega). \end{aligned}$$

Then $E_\mu(x-y) = \sum_{j=1}^{m}(x_j - y_j)p_j$. Suppose that conditions (11) and (12) hold. By Lemma 1 ii) there exist complementary $T, R \subset \Omega$ such that

$$\begin{cases} x(\varrho) - y(\varrho) \leqslant \sum_{j=1}^{m}(x_j - y_j)p_j \leqslant x(\tau) - y(\tau) \\ y(\varrho) \leqslant \gamma \leqslant y(\tau) \end{cases} \quad \tau \in T, \varrho \in R. \tag{15}$$

Since x and y are $\sigma(B_1, \ldots, B_m)$-measurable, one can choose R and T to be random events. Thus, there exist $J_1, J_2 \subset \{1, \ldots, m\}$ with $J_1 \cup J_2 =$

$\{1,\dots,m\}$ and $J_1 \cap J_2 = \emptyset$ such that $T = \bigcup_{j\in J_1} B_j$ and $R = \bigcup_{j\in J_2} B_j$. Then (15) can be rewritten as

$$\begin{cases} x_i - y_i \leqslant \sum_{j=1}^{m}(x_j - y_j)p_j \leqslant x_j - y_j \\ y_i \leqslant \gamma \leqslant y_j \end{cases} \quad j \in J_1, i \in J_2.$$

By Theorem 2, inequality (9) holds and takes the form

$$E_\mu(f \circ y) = \sum_{k=1}^{m} p_k f(y_k) \leqslant \sum_{k=1}^{m} p_k f(x_k) = E_\mu(f \circ x).$$

This is simply Niezgoda's result (Niezgoda, 2008, Corollary 2.3).

Now, set $\Omega = [a,b] \subset \mathbb{R},\ -\infty < a < b < \infty$. Given $p, u : [a,b] \mapsto \mathbb{R}$, we assume that p is continuous nonnegative and u is bounded nondecreasing. Then the Lebesgue–Stieltjes integral $\int_a^b p(t)du(t)$ exists and is nonnegative finite. Let $\int_a^b p(t)du(t) > 0$. In this situation, one can define a nontrivial probability measure μ on Borel subsets $B \subset [a,b]$ as follows:

$$\mu(B) = \frac{\int_B p(t)du(t)}{\int_a^b p(t)du(t)}. \tag{16}$$

Evidently, continuous functions on Ω are μ-measurable. Thus, for any real continuous function x on Ω we have

$$E_\mu(x) = \frac{\int x(t)p(t)du(t)}{\int p(t)du(t)}. \tag{17}$$

In this situation, a simple application of Theorem 2 leads to Corollary 2.

Corollary 2. *Let $f : \mathbb{R} \mapsto \mathbb{R}$ be a convex function and $x, y, p, u : [a,b] \mapsto \mathbb{R}$, where x, y are continuous, p is nonnegative continuous and u is nondecreasing bounded with $\int p(t)du(t) > 0$.*

If (11) and (12) hold, then inequality (9) is satisfied, where μ is the probability measure defined by (16) and the expectation has the form (17).

In light of Remark 2, the above corollary extends Theorems 6–7 by Barnett et al. (2009). In the next example, we present further generalizations of those authors' results.

Example 3. *Let μ be again a probability measure on Ω and $I \subset \mathbb{R}$ be an open interval.*

First, assume $g : I \mapsto \mathbb{R}$ is a twice differentiable function with the property

$$\alpha \leqslant \frac{d^2 g(t)}{dt^2} \leqslant \beta,\ t \in I,$$

where α and β are fixed scalars. Define the functions $f_\alpha, f_\beta : I \mapsto \mathbb{R}$ as follows:

$$f_\alpha(t) = g(t) - \frac{1}{2}\alpha t^2,\ f_\beta(t) = \frac{1}{2}\beta t^2 - g(t),\ t \in I.$$

Since g is twice differentiable, so is f_α and f_β. Hence

$$\frac{d^2 f_\alpha(t)}{dt^2} = \frac{d^2 g(t)}{dt^2} - \alpha \geqslant 0,\ \frac{d^2 f_\beta(t)}{dt^2} = \beta - \frac{d^2 g(t)}{dt^2} \geqslant 0,\ t \in I,$$

showing that both f_α and f_β are convex functions. Let x and y be bounded functions defined on Ω with values in I for which $E_\mu(x) = E_\mu(y)$ and (11) holds. Applying Theorem 2 for f_α and f_β gives

$$E_\mu(f_\alpha \circ y) \leqslant E_\mu(f_\alpha \circ x) \text{ and } E_\mu(f_\beta \circ y) \leqslant E_\mu(f_\beta \circ x),$$

or equivalently,

$$E_\mu(\frac{1}{2}\alpha y^2 - g \circ y) \leqslant E_\mu(\frac{1}{2}\alpha x^2 - g \circ x) \text{ and } E_\mu(g \circ y - \frac{1}{2}\beta y^2) \leqslant E_\mu(g \circ x - \frac{1}{2}\alpha x^2).$$

Hence, it is easy to see that (cf. Proposition 1, Barnett et al., 2009)

$$\frac{1}{2}\alpha E_\mu(x^2 - y^2) \leqslant E_\mu(g \circ x) - E_\mu(g \circ y) \leqslant \frac{1}{2}\beta E_\mu(x^2 - y^2). \tag{18}$$

Since $E_\mu(x) = E_\mu(y)$ and (11) holds,

$$[x(\omega) - y(\omega)]\,[y(\omega) - \gamma] \geqslant 0,\ \omega \in \Omega,$$

for a certain scalar γ. By Lemma 1 ii), there exist complementary $T, R \subset \Omega$ such that

$$\begin{cases} x(\varrho) - y(\varrho) \leqslant 0 \leqslant x(\tau) - y(\tau) \\ y(\varrho) \leqslant \gamma \leqslant y(\tau) \end{cases} \quad \tau \in T, \varrho \in R.$$

It follows that

$$\begin{cases} x(\varrho) \leqslant \gamma \leqslant x(\tau) \\ y(\varrho) \leqslant \gamma \leqslant y(\tau) \end{cases} \quad \tau \in T, \varrho \in R.$$

Hence

$$x(\varrho) + y(\varrho) \leqslant 2\gamma \leqslant x(\tau) + y(\tau), \ \tau \in T, \varrho \in R.$$

Finally, we obtain

$$\begin{cases} x(\varrho) - y(\varrho) \leqslant 0 \leqslant x(\tau) - y(\tau) \\ x(\varrho) + y(\varrho) \leqslant 2\gamma \leqslant x(\tau) + y(\tau) \end{cases} \quad \tau \in T, \varrho \in R.$$

Applying Lemma 1 ii) once again gives

$$[x(\omega) - y(\omega)]\,[x(\omega) + y(\omega) - 2\gamma] \geqslant 0, \text{ for all } \omega \in \Omega.$$

By Proposition 1

$$E_\mu(x^2 - y^2) = E_\mu((x+y)(x-y)) \geqslant E_\mu(x+y)E_\mu(x-y) = 0. \quad (19)$$

Combining (18) and (19), for $\alpha \geqslant 0$ we obtain the following refinement of the HLP inequality (9):

$$E_\mu(g \circ x) - E_\mu(g \circ y) \geqslant \frac{1}{2}\alpha E_\mu(x^2 - y^2) \geqslant 0$$

(cf. Remark 4, Barnett et al., 2009).

Inequalities (19) are also valid if x and y are positive and bounded functions on Ω such that $E_\mu(x) \geqslant E_\mu(y)$ and (11) holds, while $g : (0, +\infty) \mapsto \mathbb{R}$ is twice differentiable with

$$\alpha t \leqslant \frac{dg(t)}{dt} \leqslant \beta t \text{ and } \alpha \leqslant \frac{d^2 g(t)}{dt^2} \leqslant \beta, \ t > 0.$$

This is a consequence of Theorem 2 applied to nondecreasing convex functions f_α and f_β.

Similarly, one can show the following result (cf. Proposition 2, Barnett et al., 2009).

Let $g : (0, +\infty) \mapsto \mathbb{R}$ be a twice differentiable function such that there exist scalars η and θ with

$$\eta \leqslant t \cdot \frac{d^2 g(t)}{dt^2} \leqslant \theta, \ t > 0.$$

If $E_\mu(x) = E_\mu(y)$ and (11) holds for positive and bounded functions x and y defined on Ω, then

$$\eta E_\mu(x \cdot \ln x - y \cdot \ln y) \leqslant E_\mu(g \circ x) - E_\mu(g \circ y) \leqslant \theta E_\mu(x \cdot \ln x - y \cdot \ln y).$$

It suffices to apply Theorem 2 for functions $f_\eta(t) = g(t) - \eta t \ln t$, $f_\theta(t) = \theta t \ln t - g(t)$, $t \in (0, +\infty)$ which are twice differentiable and convex.

4. The Jensen—Mercer inequality for synchronous functions

Let f be a convex function on an interval containing real numbers x_k, $(k = 1, \ldots, n)$ with assigned positive weights w_k. Then Jensen's inequality (see e.g. Pečarić et al., 1992, Theorem 2.1) states:

$$f\left(\frac{1}{W_n}\sum_{k=1}^{n} w_k x_k\right) \leqslant \frac{1}{W_n}\sum_{k=1}^{n} w_k f(x_k), \tag{20}$$

where $W_n = \sum_{k=1}^{n} w_k$. Let x be a bounded μ-measurable function defined on Ω. A version of (20) for a finite measure μ (cf. Pečarić et al., 1992, Theorem 2.3) reads:

If f is a convex function on an open interval containing the range of x, then

$$f(E_\mu(x)) \leqslant E_\mu(f \circ x). \tag{21}$$

Mercer (2003), in Theorem 1.2, proved the following variant of (20).

Let a, b, x_k be real numbers with $a \leqslant x_k \leqslant b$ $(k = 1, \ldots, n)$. If f is a convex function defined on an interval containing a and b, then

$$f\left(a + b - \frac{1}{W_n}\sum_{k=1}^{n} w_k x_k\right) \leqslant f(a) + f(b) - \frac{1}{W_n}\sum_{k=1}^{n} w_k f(x_k). \tag{22}$$

This result yielded many generalizations and applications; see e.g. Abramovich et al. (2005, 2006, 2008), Klaričić Bakula et al. (2006, 2007), Barić et al. (2010), Cheung et al. (2006), Gavrea (2004), Matković et al. (2006, 2007, 2008) and many others.

Niezgoda (2009) has studied the following generalization of (22):

$$\begin{aligned} f\left(\tfrac{1}{p_m}\left\{\sum_{j=1}^{m} p_j a_j \;-\; \sum_{j=1}^{m-1} p_j \sum_{i=1}^{n} w_i x_{ij}\right\}\right) \\ \leqslant \tfrac{1}{p_m}\left\{\sum_{j=1}^{m} p_j f(a_j) - \sum_{j=1}^{m-1} p_j \sum_{i=1}^{n} w_i f(x_{ij})\right\}, \end{aligned} \tag{23}$$

where $\overline{p} = (p_1, \ldots, p_m), \overline{w} = (w_1, \ldots, w_n)$ are positive tuples with $\sum_{i=1}^{n} w_i = 1$, $\overline{a} = (a_1, \ldots, a_m)$ is a real m-tuple, (x_{ij}) is a real $n \times m$ matrix and f is a convex function on an interval containing all of a_j and x_{ij}. He proved (Corollary 3.3, Niezgoda, 2009) that (23) holds if $\sum_{j=1}^{m}(a_j - x_{ij})p_j = 0$ and $(a_j - x_{ij})$ and (x_{ij}) are pairs of similarly separable m-tuples for fixed $i = 1, 2, \ldots, n$, i.e. there exist index sets J_1, J_2 dependent on i, with $J_1 \cup J_2 = \{1, \ldots, m\}$ such that

$$\begin{cases} a_l - x_{il} & \leqslant 0 \leqslant & a_j - x_{ij} \\ x_{il} & \leqslant & x_{ij} \end{cases} \quad j \in J_1, l \in J_2. \tag{24}$$

We will give a continuous version of (23) for spaces with probability measures. For this purpose let ν and μ be probability measures on spaces Σ and Ω, respectively. By $\nu \times \mu$ we denote the product probability measure on the Cartesian product $\Sigma \times \Omega$ of these spaces.

Theorem 3. *Let $x = x(\sigma, \omega)$ and $a = a(\omega)$ be bounded random variables with values in an open interval $I \subset \mathbb{R}$ defined on $\Sigma \times \Omega$ and Ω, respectively. Fix a convex function $f : \mathbb{R} \mapsto \mathbb{R}$ and $A, B \subset \Omega,\ A \cup B = \Omega,\ \mu(A \cap B) = 0$ with $\mu(B) > 0$. If C is the set of all $\sigma \in \Sigma$ such that*

$$E_\mu(a) = E_\mu(x) \text{ and } \exists_{\gamma \in \mathbb{R}} :\ [a(\omega) - x(\sigma, \omega)][x(\sigma, \omega) - \gamma] \geqslant 0, \text{ for all } \omega \in \Omega \tag{25}$$

then the inequality

$$\begin{array}{rcl} f\left(\frac{1}{\mu(B)}\left\{E_\mu(a) \right.\right. & - & \left.\left.\frac{1}{\nu(C)} E_{\nu\times\mu}(x \cdot I_{C\times A})\right\}\right) \\ & \leqslant & \frac{1}{\mu(B)}\left\{E_\mu(f \circ a) - \frac{1}{\nu(C)} E_{\nu\times\mu}(f \circ x \cdot I_{C\times A})\right\} \end{array} \tag{26}$$

holds, provided that $\nu(C) > 0$.

Proof. Fix $\sigma \in C$. Since $E_\mu(a) = E_\mu(x)$ i.e.

$$\int a(\omega) d\mu(\omega) = \int x(\sigma, \omega) d\mu(\omega),$$

$$f\left(\frac{1}{\mu(B)}\left\{\int a(\omega) d\mu(\omega) - \int_A x(\sigma, \omega) d\mu(\omega)\right\}\right) = f\left(\frac{1}{\mu(B)} \int_B x(\sigma, \omega) d\mu(\omega)\right).$$

By Jensen's inequality (21)

$$f\left(\frac{1}{\mu(B)} \int_B x(\sigma, \omega) d\mu(\omega)\right) \leqslant \frac{1}{\mu(B)} \int_B f \circ x(\sigma, \omega) d\mu(\omega).$$

Since (25), utilizing inequality (9) gives

$$\frac{1}{\mu(B)}\int_B f\circ x(\sigma,\omega)d\mu(\omega) \leqslant \frac{1}{\mu(B)}\left\{\int f\circ a(\omega)d\mu(\omega) - \int_A f\circ x(\sigma,\omega)d\mu(\omega)\right\},$$

by Theorem 2. Combining the above, we obtain

$$f\left(\frac{1}{\mu(B)}\left\{\int a(\omega)d\mu(\omega) - \int_A x(\sigma,\omega)d\mu(\omega)\right\}\right) \leqslant$$

$$\frac{1}{\mu(B)}\left\{\int f\circ a(\omega)d\mu(\omega) - \int_A f\circ x(\sigma,\omega)d\mu(\omega)\right\},\ \sigma\in C.$$

Now, multiplying the last inequality by $1/\nu(C)$ and integrating over C with respect to ν gives

$$\frac{1}{\nu(C)}\int_C f\left(\frac{1}{\mu(B)}\left\{\int a(\omega)d\mu(\omega) - \int_A x(\sigma,\omega)d\mu(\omega)\right\}\right)d\nu(\sigma) \leqslant$$

$$\frac{1}{\mu(B)}\left\{\int f\circ a(\omega)d\mu(\omega) - \frac{1}{\nu(C)}\int_C d\nu(\sigma)\int_A f\circ x(\sigma,\omega)d\mu(\omega)\right\}.$$

Applying Jensen's inequality (21) once again, we obtain

$$\frac{1}{\nu(C)}\int_C f\left(\frac{1}{\mu(B)}\left\{\int a(\omega)d\mu(\omega) - \int_A x(\sigma,\omega)d\mu(\omega)\right\}\right)d\nu(\sigma) \geqslant$$

$$f\left(\frac{1}{\mu(B)}\left\{\int a(\omega)d\mu(\omega) - \frac{1}{\nu(C)}\int_C d\nu(\sigma)\int_A x(\sigma,\omega)d\mu(\omega)\right\}\right).$$

Combining the last two inequalities gives

$$f\left(\frac{1}{\mu(B)}\left\{\int a(\omega)d\mu(\omega) - \frac{1}{\nu(C)}\int_C d\nu(\sigma)\int_A x(\sigma,\omega)d\mu(\omega)\right\}\right) \leqslant$$

$$\frac{1}{\mu(B)}\left\{\int f\circ a(\omega)d\mu(\omega) - \frac{1}{\nu(C)}\int_C d\nu(\sigma)\int_A f\circ x(\sigma,\omega)d\mu(\omega)\right\}.$$

Recall that $\int a(\omega)d\mu(\omega) = E_\mu(a)$ and $\int f\circ a(\omega)d\mu(\omega) = E_\mu(f\circ a)$. Fubini's Theorem ensures that $\int_C d\nu(\sigma)\int_A x(\sigma,\omega)d\mu(\omega) = \int_{C\times A} x d(\nu\times\mu) = E_{\nu\times\mu}(x\cdot I_{C\times A})$ and $\int_C d\nu(\sigma)\int_A f\circ x(\sigma,\omega)d\mu(\omega) = \int_{C\times A} f\circ x d(\nu\times\mu) = E_{\nu\times\mu}(f\circ x\cdot I_{C\times A})$. This completes the proof. □

The following example shows that under some simplifications, Theorem 3 reduces to Corollary 3.3 of Niezgoda (2009).

Example 4. *Let* $\Sigma = \bigcup_{i=1}^{n} A_i$ *and* $\Omega = \bigcup_{j=1}^{m} B_j$ *where* $A_i \cap A_k = \emptyset$ *iff* $i \neq k$ *and* $B_j \cap B_l = \emptyset$ *iff* $j \neq l$.

For given real tuples $\overline{a} = (a_1, \ldots, a_m), \overline{p} = (p_1, \ldots, p_m), \overline{w} = (w_1, \ldots, w_n)$ *and a real* $n \times m$ *matrix* (x_{ij}), *where* $\overline{p}$ *and* $\overline{w}$ *are positive with* $\sum_i p_i = 1 = \sum_i w_i$, *we define discrete probability measures* ν *and* μ *on* σ*-fields* $\sigma(A_1, \ldots, A_n)$ *and* $\sigma(B_1, \ldots, B_m)$, *respectively, and random variables on* Ω *and* $\Sigma \times \Omega$ *as follows:*

$$\begin{aligned} \nu(A_i) &= w_i,\ i = 1, \ldots, n, \\ \mu(B_j) &= p_j,\ j = 1, \ldots, m, \\ a(\omega) &= \textstyle\sum_j a_j I_{B_j}(\omega), \\ x(\sigma, \omega) &= \textstyle\sum_i \sum_j x_{ij} I_{A_i}(\sigma) I_{B_j}(\omega). \end{aligned}$$

Suppose that (25) is fulfilled for all $\sigma \in \Sigma$, *i.e. the set* C *defined in Theorem 3 is equal to* Σ. *In this situation,* $E_\mu(a) = E_\mu(x)$ *is equivalent to* $\sum_j a_j p_j = \sum_j x_{ij} p_j$ *for all* i *and there exist complementary* $T, R \subset \Omega$ *such that*

$$\begin{cases} a(\varrho) - x(\sigma, \varrho) \leqslant 0 \leqslant a(\tau) - x(\sigma, \tau) \\ x(\sigma, \varrho) \leqslant \gamma \leqslant x(\sigma, \tau) \end{cases} \quad \tau \in T, \varrho \in R, \tag{27}$$

where $\sigma \in \Sigma$ *is fixed. Since* a *and* $x(\sigma, \cdot)$ *are* $\sigma(B_1, \ldots, B_m)$*-measurable, one can choose* R *and* T *to be measurable, too. Thus, there exist* $J_1, J_2 \subset \{1, \ldots, m\}$ *with* $J_1 \cup J_2 = \{1, \ldots, m\}$ *and* $J_1 \cap J_2 = \emptyset$ *such that* $T = \bigcup_{j \in J_1} B_j$ *and* $R = \bigcup_{j \in J_2} B_j$. *Then (27) can be rewritten as*

$$\begin{cases} a_l - x_{il} \leqslant 0 \leqslant a_j - x_{ij} \\ x_{il} \leqslant \gamma \leqslant x_{ij} \end{cases} \quad j \in J_1, l \in J_2.$$

This is simply Niezgoda's condition (24).

Now, let $A = \bigcup_{l=1}^{m-1} B_l$ *and* $B = B_m$. *By Theorem 3 we obtain (26) in the form (23).* □

Acknowledgements

The author would like to express sincere thanks to the Editor-in-Chief for her patience and immediate contact and to one of the reviewers for a simple and kind review.

References

Abramovich S., Klaričić Bakula M., Matić M. and Pečarić J. (2005): A variant of Jensen-Steffensen's inequality and quasi-arithmetic means. J. Math. Anal. Appl. 307: 370–386.

Abramovich S. , Klaričić Bakula M. and Banić S. (2006): A variant of Jensen-Steffensen's inequality for convex and superquadratic functions. J. Inequal. Pure Appl. Math. 7(2): Article 70.

Abramovich S., Barić J., Pečarić J. (2008): A variant of Jessen's inequality of Mercer's type for superquadratic functions. J. Inequal. Pure Appl. Math. 9(3): Article 62.

Barić J., Matković A., Pečarić J. (2010) A variant of the Jensen-Mercer operator inequality for superquadratic functions. Mathematical and Computer Modelling. 51: 1230–1239.

Barnett N.S., Cerone P., Dragomir S.S. (2009) Majorisation inequalities for Stieltjes integrals. Applied Mathematics Letters. 22: 416–421.

Cheung W.S., Matković A. and Pečarić J. (2006) A variant of Jensen's inequality and generalized means. J. Inequal. Pure Appl. Math. 7(1): Article 10.

Dragomir S.S. (2004) Some majorisation type discrete inequalities for convex functions. Math. Inequal. Appl. 7(2): 207–216.

Fujiwara M. (1918) Ein von Brunn vermuteter Satz über konvexe Flächen und eine Verallgemeinerung der Schwarzschen und Tchebyscheffschen Ungleichungen für bestimmte Integrale. Tôhoku Math. J. 13: 228–235.

Gavrea I. (2004) Some considerations on the monotonicity property of power means. J. Inequal. Pure Appl. Math. 5 (4): Article 93.

Hardy G.H., Littlewood J.E. and Pólya G. (1928/29) Some simple inequalities satisfied by convex functions. Messenger Math. 58: 145–152.

Klaričić Bakula M., Matković A., Pečarić J. (2006) Variants of Čebyšev's inequality with applications. J. Inequal. Appl. Article no. 39692.

Klaričić Bakula M., Matković A., Pečarić J. (2007) On the Jensen-Steffensen inequality for generalized convex functions. Period. Math. Hungar. 55(1): 19–34.

Marshall A.W. and Olkin I. (1979) Theory of Majorization and Its Applications. Academic Press, New York.

Matković A., Pečarić J. and Perić I. (2006) A variant of Jensen's inequality of Mercer's type for operators with applications. Linear Algebra Appl. 418: 551–564.

Matković A. and Pečarić J. (2007) A variant of Jensen's inequality for convex functions of several variables. J. Math. Inequal. 1(1): 45–51.

Matković A., Pečarić J. and Perić I. (2008) Refinements of Jensen's inequality of Mercer's type for operator convex functions. Math. Inequal. Appl. 11(1): 113–126.

Mercer A.McD. (2003) A variant of Jensen's inequality, J. Inequal. Pure Appl. Math. 4(4): Article 73.

Mitrinović D.S. (1970) Analytic inequalities. Springer-Verlag, Berlin/New York: p. 40.

Niculescu C.P. , Popovici F. (2006) The extension of majorization inequalities within the frame of relative convexity. J. Inequal. Pure Appl. Math. 7(1): Article 27.

Niezgoda M. (2009) A generalization of Mercer's result on convex functions. Nonlinear Analysis. 71(7-8): 2771–2779.

Niezgoda M. (2008 Remarks on convex functions and separable sequences. Discrete Math. 308: 1765–1773.

Niezgoda M. (2006) Bifractional inequalities and convex cones. Discrete Math. 306: 231–243.

Otachel Z. (2011) Chebyshev type inequalities for synchronous vectors in Banach spaces. Math. Inequal. Appl. 14(2): 421—437.

Otachel Z. (2014) Synchronous sequences and inequalities for convex functions. Appl. Math. Comput. 247: 865–871.

Pečarić J.E., Proschan F., and Tong Y.L. (1992) Convex functions, partial orderings, and statistical applications. Mathematics in Science and Engineering, 87. Academic Press Inc.

Shaked M. (1982) A general theory of some positive dependence notions. J. Multivar. Anal. 12: 199–218.

Toader Gh. (1992) Integral and discrete inequalities. Rev. Anal. Numér. Théor. Approx. 21(1): 83–88.

Toader Gh. (1992) Fujiwara's inequality for functionals. Facta Univ. Ser. Math. Infor. 7: 43–48.

Zălinescu C. (2002) Convex analysis in general vector spaces. World Scientific.