Distributions with Fixed Marginals and Related Topics
IMS Lecture Notes – Monograph Series Vol. 28, 1996

Bounds for the Distribution of a Multivariate Sum

By Haijun Li*, Marco Scarsini, and Moshe Shaked*
Washington State University, Università d'Annunzio, and University of Arizona

Makarov (1981) and Frank, Nelsen and Schweizer (1987), and independently Rüschendorf (1982), have found upper and lower bounds for $P\{X+Y<t\}$ ($t \in \mathbb{R} = (-\infty, \infty)$), when the marginal distributions of X and Y are fixed, and they have proved that their bounds are sharp.

In this paper we find similar bounds when $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vectors rather than scalars. First we determine lower and upper bounds by generalizing the method of Frank, Nelsen and Schweizer and we show that the method can be used also to determine bounds for distributions of functions other than the sum. Then, by generalizing Rüschendorf's method, based on a theorem of Strassen, we prove that the bounds previously obtained are sharp. Finally we use the bounds to obtain inequalities for expectations of increasing and of Δ-monotone functions of $\boldsymbol{X}+\boldsymbol{Y}$.

1. Introduction.

Makarov (1981) and Frank, Nelsen and Schweizer (1987), and independently Rüschendorf (1982), have solved the following problem: Let X and Y be real-valued random variables with respective one-dimensional distribution functions F_1 and F_2, and let $\mathcal{F}_{F_1,F_2}$ be the Fréchet class of joint distributions with marginals F_1 and F_2. For all $t \in \mathbb{R}$ find the best bounds

$$L(t) := \inf_{\mathcal{F}_{F_1,F_2}} P\{X+Y<t\} \tag{1.1}$$

and

$$U(t) := \sup_{\mathcal{F}_{F_1,F_2}} P\{X+Y<t\}. \tag{1.2}$$

A review can be found in Section 2.2.5 of Rüschendorf (1991) or in Section 8 of Schweizer (1991); see also Remark 7.3.3 in Rachev (1991).

---

* Supported by NSF Grant DMS 9303891.

AMS 1991 Subject Classification: 60E05

Key words and phrases: Given marginals, Δ-monotone functions, capacities.

In this paper we extend the methods of Frank, Nelsen and Schweizer (1987), and of Rüschendorf (1982), to the multivariate case, that is, to the case in which F_1 and F_2 are n-dimensional. Whereas the extension of Frank, Nelsen and Schweizer's method provides bounds that are easy to compute, we cannot prove sharpness using their method. To show that the bounds are sharp we generalize Rüschendorf's idea, which in turn is based on a theorem of Strassen. The bounds can be used to establish inequalities for the expectations of increasing and of Δ-monotone functions of the sum of two random vectors.

Below, the terms "increasing" and "decreasing" stand, respectively, for "nondecreasing" and "nonincreasing." Also, whenever we study an expectation or an integral we implicitly assume it exists. For two elements $\boldsymbol{s} = (s_1, s_2, \ldots, s_n)$ and $\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$ in $\mathbb{R}^n$, the notation $\boldsymbol{s} \leq \boldsymbol{t}$ will mean $s_i \leq t_i$, $i = 1, 2, \ldots, n$, and the notation $\boldsymbol{s} < \boldsymbol{t}$ will mean $s_i < t_i$, $i = 1, 2, \ldots, n$. In this paper, when we refer to the distribution function F of a random vector $\boldsymbol{T} = (T_1, T_2, \ldots, T_n)$ we mean the function F defined by $F(\boldsymbol{t}) \equiv P\{\boldsymbol{T} < \boldsymbol{t}\}$. The corresponding survival function $\overline{F}$ is defined by $\overline{F}(\boldsymbol{t}) \equiv P\{\boldsymbol{T} \geq \boldsymbol{t}\}$.

1. Bounds on the Distribution Function of $\boldsymbol{X} + \boldsymbol{Y}$. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be two random vectors with respective marginal distributions F_1 and F_2, and some joint distribution F. Define

$$W(\boldsymbol{x}, \boldsymbol{y}) = \max\{F_1(\boldsymbol{x}) + F_2(\boldsymbol{y}) - 1, 0\}, \tag{2.1}$$

$$Z(\boldsymbol{x}, \boldsymbol{y}) = \min\{F_1(\boldsymbol{x}) + F_2(\boldsymbol{y}), 1\}. \tag{2.2}$$

In general W and Z are not distribution functions. When $n = 1$, W is the lower Fréchet bound of $\mathcal{F}_{F_1, F_2}$ (and therefore a distribution function), and Z is one minus the lower Fréchet bound for the class of joint survival functions $\overline{F}(\boldsymbol{x}, \boldsymbol{y}) = P\{\boldsymbol{X} \geq \boldsymbol{x}, \boldsymbol{Y} \geq \boldsymbol{y}\}$ with the above marginals. Also define

$$L(\boldsymbol{t}) = \sup_{\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{t}} W(\boldsymbol{u}, \boldsymbol{v}), \tag{2.3}$$

$$U(\boldsymbol{t}) = \inf_{\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{t}} Z(\boldsymbol{u}, \boldsymbol{v}). \tag{2.4}$$

THEOREM 2.1. *For every pair of random vectors $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ having distributions F_1 and F_2, respectively,*

$$L(\boldsymbol{t}) \leq P\{\boldsymbol{X} + \boldsymbol{Y} < \boldsymbol{t}\} \leq U(\boldsymbol{t}). \tag{2.5}$$

PROOF. Note that for all $\boldsymbol{t}, \boldsymbol{u} \in \mathbb{R}^n$

$$\begin{aligned} W(\boldsymbol{u}, \boldsymbol{t} - \boldsymbol{u}) &= \max\{F_1(\boldsymbol{u}) + F_2(\boldsymbol{t} - \boldsymbol{u}) - 1, 0\} \\ &\leq P\{\boldsymbol{X} < \boldsymbol{u}, \boldsymbol{Y} < \boldsymbol{t} - \boldsymbol{u}\} \leq P\{\boldsymbol{X} + \boldsymbol{Y} < \boldsymbol{t}\}, \end{aligned} \tag{2.6}$$

$$\begin{aligned} Z(\boldsymbol{u}, \boldsymbol{t}-\boldsymbol{u}) &= \min\{F_1(\boldsymbol{u}) + F_2(\boldsymbol{t}-\boldsymbol{u}), 1\} \\ &\geq 1 - P\{\boldsymbol{X} \geq \boldsymbol{u}, \boldsymbol{Y} \geq \boldsymbol{t}-\boldsymbol{u}\} \geq P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\}. \end{aligned} \tag{2.7}$$

Therefore

$$L(\boldsymbol{t}) = \sup_{\boldsymbol{u}} W(\boldsymbol{u}, \boldsymbol{t}-\boldsymbol{u}) \leq P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\}, \tag{2.8}$$

$$U(\boldsymbol{t}) = \inf_{\boldsymbol{u}} Z(\boldsymbol{u}, \boldsymbol{t}-\boldsymbol{u}) \geq P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\}, \tag{2.9}$$

for each $2n$-dimensional distribution F of $(\boldsymbol{X}, \boldsymbol{Y})$ with n-dimensional margins F_1 and F_2. ∎

The central point is that the bounds $L(\boldsymbol{t})$ and $U(\boldsymbol{t})$ in (2.5) depend only on the margins and *not* on the joint distribution. Note that these bounds are not necessarily distribution functions.

The bounds given in Theorem 2.1 are on the probability of a lower orthant $\{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u} < \boldsymbol{t}\}$. We present next the analogous bounds on the probabilities of upper orthants $\{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u} \geq \boldsymbol{t}\}$. These bounds are obtained in an entirely similar fashion by way of the survival functions $\overline{F}_1$ and $\overline{F}_2$. Define

$$\overline{w}(\boldsymbol{x}, \boldsymbol{y}) = \max\{\overline{F}_1(\boldsymbol{x}) + \overline{F}_2(\boldsymbol{y}) - 1, 0\}, \tag{2.10}$$

$$\overline{z}(\boldsymbol{x}, \boldsymbol{y}) = \min\{\overline{F}_1(\boldsymbol{x}) + \overline{F}_2(\boldsymbol{y}), 1\}. \tag{2.11}$$

In general $\overline{w}$ is not a survival function except when $n = 1$. Also define

$$\bar{l}(\boldsymbol{t}) = \sup_{\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{t}} \overline{w}(\boldsymbol{u}, \boldsymbol{v}), \tag{2.12}$$

$$\overline{u}(\boldsymbol{t}) = \inf_{\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{t}} \overline{z}(\boldsymbol{u}, \boldsymbol{v}). \tag{2.13}$$

THEOREM 2.2. *For every pair of random vectors* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ *having distributions* F_1 *and* F_2*, respectively,*

$$\bar{l}(\boldsymbol{t}) \leq P\{\boldsymbol{X}+\boldsymbol{Y} \geq \boldsymbol{t}\} \leq \overline{u}(\boldsymbol{t}).$$

REMARK 2.3. A comparison of Theorems 2.1 and 2.2 shows the complication that arises in the multivariate case, in contrast to the univariate case. When $n = 1$ the two theorems yield the same bounds, those of Frank, Nelsen and Schweizer (1987). However the bounds are different when $n \geq 2$ because the complement of a lower "orthant" is an upper "orthant" only when $n = 1$.

We now extend Theorem 2.1 to functions other than addition. Note that the last inequalities in (2.6) and (2.7) follow from two simple set inclusions: For any fixed $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$

$$\{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' < \boldsymbol{u}, \boldsymbol{v}' < \boldsymbol{v}\} \subseteq \{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' + \boldsymbol{v}' < \boldsymbol{u} + \boldsymbol{v}\}$$

and

$$\{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' \geq \boldsymbol{u}, \boldsymbol{v}' \geq \boldsymbol{v}\} \subseteq \{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' + \boldsymbol{v}' \geq \boldsymbol{u} + \boldsymbol{v}\}.$$

Thus, one may replace $\boldsymbol{X} + \boldsymbol{Y}$ in Theorem 2.1 by a more general function $g(\boldsymbol{X}, \boldsymbol{Y})$ provided that

$$\{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' < \boldsymbol{u}, \boldsymbol{v}' < \boldsymbol{v}\} \subseteq \{(\boldsymbol{u}', \boldsymbol{v}') : g(\boldsymbol{u}', \boldsymbol{v}') < g(\boldsymbol{u}, \boldsymbol{v})\}$$

and

$$\{(\boldsymbol{u}', \boldsymbol{v}') : \boldsymbol{u}' \geq \boldsymbol{u}, \boldsymbol{v}' \geq \boldsymbol{v}\} \subseteq \{(\boldsymbol{u}', \boldsymbol{v}') : g(\boldsymbol{u}', \boldsymbol{v}') \geq g(\boldsymbol{u}, \boldsymbol{v})\}.$$

Any increasing function $g : \mathbb{R}^{2n} \to \mathbb{R}$, such that

$$(\boldsymbol{u}, \boldsymbol{v}) < (\boldsymbol{u}', \boldsymbol{v}') \Longrightarrow g(\boldsymbol{u}, \boldsymbol{v}) < g(\boldsymbol{u}', \boldsymbol{v}'), \tag{2.15}$$

satisfies these conditions. Define now

$$L_g(\boldsymbol{t}) = \sup_{g(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{t}} W(\boldsymbol{u}, \boldsymbol{v}),$$
$$U_g(\boldsymbol{t}) = \inf_{g(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{t}} Z(\boldsymbol{u}, \boldsymbol{v}),$$

where W and Z are defined in (2.1) and (2.2). Also define

$$\bar{l}_g(\boldsymbol{t}) = \sup_{g(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{t}} \overline{w}(\boldsymbol{u}, \boldsymbol{v}),$$
$$\overline{u}(\boldsymbol{t}) = \inf_{g(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{t}} \overline{z}(\boldsymbol{u}, \boldsymbol{v}),$$

where $\overline{w}$ and $\overline{z}$ are defined in (2.10) and (2.11). Then we have the following result.

THEOREM 2.4. *Let $g : \mathbb{R}^{2n} \to \mathbb{R}$ be an increasing function that satisfies (2.15). For every pair of random vectors $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ having distributions F_1 and F_2, respectively,*

$$L_g(\boldsymbol{t}) \leq P\{g(\boldsymbol{X}, \boldsymbol{Y}) < \boldsymbol{t}\} \leq U_g(\boldsymbol{t})$$

and

$$\bar{l}_g(\boldsymbol{t}) \leq P\{g(\boldsymbol{X}, \boldsymbol{Y}) \geq \boldsymbol{t}\} \leq \overline{u}_g(\boldsymbol{t}).$$

3. Sharp Bounds. In this section we show that the bounds in Theorems 2.1 and 2.2 are sharp.

We start by stating a special case of Theorem 11 of Strassen (1965). Let S and T be Polish spaces and let P_1 be a probability measure on $(S, \mathrm{Bor}(S))$, P_2 be a probability measure on $(T, \mathrm{Bor}(T))$, and let $\mathcal{P}_{P_1,P_2}$ be the class of probability measures on $(S \times T, \mathrm{Bor}(S) \otimes \mathrm{Bor}(T))$ with marginals P_1 and P_2.

LEMMA 3.1. *Fix a closed set $A \subseteq S \times T$ and an $\epsilon \geq 0$. Then the following two statements are equivalent.*

(a) There exists a $\mu \in \mathcal{P}_{P_1,P_2}$ such that $\mu(A) \geq 1 - \epsilon$.

(b) For every open set $C \subseteq T$, $P_2(C) - P_1(\pi_1(A \cap (S \times C))) \leq \epsilon$, where π_1 denotes the projection on S.

Now, for any closed set $A \subseteq S \times T$, let $M(A) = \sup\{\mu(A) : \mu \in \mathcal{P}_{P_1,P_2}\}$ and $m(A) = \inf\{\mu(A) : \mu \in \mathcal{P}_{P_1,P_2}\}$, and let $\mathcal{O}_T$ denote the class of open subsets of T. The following result was stated without proof in Rüschendorf (1982) for the case $S = T = \mathbb{R}$.

LEMMA 3.2. *For any closed set $A \subseteq S \times T$ one has*

$$M(A) = 1 - \sup_{C \in \mathcal{O}_T} \{P_2(C) - P_1(\pi_1(A \cap (S \times C)))\}.$$

PROOF. We will establish the inequalities

$$M(A) \leq 1 - \sup_{C \in \mathcal{O}_T} \{P_2(C) - P_1(\pi_1(A \cap (S \times C)))\}, \tag{3.1}$$

and

$$M(A) \geq 1 - \sup_{C \in \mathcal{O}_T} \{P_2(C) - P_1(\pi_1(A \cap (S \times C)))\}. \tag{3.2}$$

Recall that $M(A) = \sup\{\mu(A) : \mu \in \mathcal{P}_{P_1,P_2}\}$. Thus, for every $\delta > 0$, there exists a $\mu \in \mathcal{P}_{P_1,P_2}$ such that

$$\mu(A) \geq M(A) - \delta = 1 - (\delta - M(A) + 1).$$

By Lemma 3.1, for every open set $C \subseteq T$,

$$P_2(C) - P_1(\pi_1(A \cap (S \times C))) \leq \delta - M(A) + 1.$$

Therefore

$$\sup_{C \in \mathcal{O}_T} \{P_2(C) - P_1(\pi_1(A \cap (S \times C)))\} \leq \delta - M(A) + 1.$$

Letting $\delta \to 0$ we obtain (3.1).

Now, for a fixed closed set $A \subseteq S \times T$, let

$$\theta(A) = \sup_{C \in \mathcal{O}_T} \{P_2(C) - P_1(\pi_1(A \cap (S \times C)))\}.$$

Then, for every open set $C \subseteq T$,

$$P_2(C) - P_1(\pi_1(A \cap (S \times C))) \leq \theta(A).$$

By Lemma 3.1 there exists a $\mu \in \mathcal{P}_{P_1,P_2}$ such that $\mu(A) \geq 1-\theta(A)$. Since, by definition, $M(A) \geq \mu(A)$, it follows that $M(A) \geq 1-\theta(A)$, which is (3.2). ∎

We are now ready to provide sharp bounds for $P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\}$ and for $P\{\boldsymbol{X}+\boldsymbol{Y} \leq \boldsymbol{t}\}$ along the lines of Rüschendorf (1982). For $\boldsymbol{t} \in \mathbb{R}^n$ let $A(\boldsymbol{t}) = \{(\boldsymbol{x},\boldsymbol{y}) : \boldsymbol{x}+\boldsymbol{y} < \boldsymbol{t}\}$ and $A(\boldsymbol{t}+) = \{(\boldsymbol{x},\boldsymbol{y}) : \boldsymbol{x}+\boldsymbol{y} \leq \boldsymbol{t}\}$. For any $\boldsymbol{t} \in \mathbb{R}^n$ and $C \subseteq \mathbb{R}^n$, let $\boldsymbol{t} - C = \{\boldsymbol{y} : \boldsymbol{y} = \boldsymbol{t} - \boldsymbol{x} \text{ for some } \boldsymbol{x} \in C\}$. Recall that a set $A \subseteq \mathbb{R}^n$ is called *upper* [*lower*](3.6) if $\boldsymbol{t} \in A$ and $\boldsymbol{s} \geq [\leq]\ \boldsymbol{t}$ imply that $\boldsymbol{s} \in A$. Let $\mathcal{O}$ and $\mathcal{U}$ be, respectively, the class of open subsets and the class of upper open subsets of $\mathbb{R}^n$. As in Section 1 we denote by $\mathcal{F}_{F_1,F_2}$ the class of joint distributions with marginals F_1 and F_2.

THEOREM 3.3. *For every pair of random vectors* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ *having distributions* F_1 *and* F_2, *respectively,*

$$\inf_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\} = \sup_{\boldsymbol{a} \in \mathbb{R}^n} \{P_2((-\infty,\boldsymbol{a})) - P_1((-\infty,\boldsymbol{t}-\boldsymbol{a})^c)\} \tag{3.3}$$

and

$$\sup_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X}+\boldsymbol{Y} \leq \boldsymbol{t}\} = \inf_{C \in \mathcal{U}} \{P_2(C^c) + P_1(\boldsymbol{t}-C)\}, \tag{3.4}$$

where P_1 *and* P_2 *are the probability measures associated with* F_1 *and* F_2, *respectively.*

PROOF. To prove (3.4), first note that

$$\sup_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X}+\boldsymbol{Y} \leq \boldsymbol{t}\} = M(A(\boldsymbol{t}+)).$$

By Lemma 3.2 we have

$$M(A(\boldsymbol{t}+)) = 1 - \sup_{C \in \mathcal{O}} \{P_2(C) - P_1(\pi_1(A(\boldsymbol{t}+) \cap (\mathbb{R}^n \times C)))\}.$$

For $C \subseteq \mathbb{R}^n$ define

$$C^*_{\boldsymbol{t}} := \pi_1(A(\boldsymbol{t}+) \cap (\mathbb{R}^n \times C)) = \{\boldsymbol{x} : \boldsymbol{x} \leq \boldsymbol{t} - \boldsymbol{y} \text{ for some } \boldsymbol{y} \in C\}$$

($C^*_{\boldsymbol{t}}$ is a lower set).

Note that given $\boldsymbol{t}$ and C, there exists an upper set V such that

$$C^*_{\boldsymbol{t}} = V^*_{\boldsymbol{t}}. \tag{3.5}$$

(To see it just take $V = \{\boldsymbol{y}' : \boldsymbol{y}' \geq \boldsymbol{y} \text{ for some } \boldsymbol{y} \in C\}$ and verify that (3.5) holds.) When C is an upper set then $C^*_{\boldsymbol{t}} = \boldsymbol{t} - C$.

Now, from (3.5) it follows that

$$\sup_{C\in\mathcal{O}}\{P_2(C)-P_1(\pi_1(A(\boldsymbol{t}+)\cap(\mathbb{R}^n\times C)))\}$$
$$=\sup_{C\in\mathcal{U}}\{P_2(C)-P_1(\pi_1(A(\boldsymbol{t}+)\cap(\mathbb{R}^n\times C)))\}.$$

Therefore

$$\begin{aligned}M(A(\boldsymbol{t}+)) &= 1-\sup_{C\in\mathcal{U}}\{P_2(C)-P_1(\pi_1(A(\boldsymbol{t}+)\cap(\mathbb{R}^n\times C)))\}\\ &= 1-\sup_{C\in\mathcal{U}}\{P_2(C)-P_1(C^*_{\boldsymbol{t}})\}\\ &= 1-\sup_{C\in\mathcal{U}}\{P_2(C)-P_1(\boldsymbol{t}-C)\}\\ &= \inf_{C\in\mathcal{U}}\{P_2(C^c)+P_1(\boldsymbol{t}-C)\},\end{aligned}$$

and this gives (3.4).

The proof of (3.3) is similar. First note that

$$\inf_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X}+\boldsymbol{Y}<\boldsymbol{t}\}=m(A(\boldsymbol{t})).$$

Since the set $(A(\boldsymbol{t}))^c$ is closed, Lemma 3.2 yields

$$M((A(\boldsymbol{t}))^c)=1-\sup_{C\in\mathcal{O}}\{P_2(C)-P_1(\pi_1((A(\boldsymbol{t}))^c\cap(\mathbb{R}^n\times C)))\}.$$

For $C\subseteq\mathbb{R}^n$ define

$$\widetilde{C_{\boldsymbol{t}}}:=\pi_1((A(\boldsymbol{t}))^c\cap(\mathbb{R}^n\times C))=\{\boldsymbol{x}:\boldsymbol{x}\not<\boldsymbol{t}-\boldsymbol{y}\ \text{ for some }\ \boldsymbol{y}\in C\}$$

($\widetilde{C_{\boldsymbol{t}}}$ is an upper set).

For every $\boldsymbol{t}$ and C, there exists a lower set V such that

$$\widetilde{C_{\boldsymbol{t}}}=\widetilde{V_{\boldsymbol{t}}}. \tag{3.6}$$

(To see it just take $V=\{\boldsymbol{y}':\boldsymbol{y}'\le\boldsymbol{y}\ \text{ for some }\ \boldsymbol{y}\in C\}$ and verify that (3.6) holds.) More than that, for every $\boldsymbol{t}$ and C, there exists a lower orthant Q such that

$$\widetilde{C_{\boldsymbol{t}}}=\widetilde{Q_{\boldsymbol{t}}}. \tag{3.7}$$

(To see it suppose that C is a lower set and take Q to be the smallest lower orthant containing C. Then, clearly, $\widetilde{C_{\boldsymbol{t}}}\subseteq\widetilde{Q_{\boldsymbol{t}}}$. But if $\boldsymbol{x}\in\widetilde{Q_{\boldsymbol{t}}}$ then $\boldsymbol{x}\not<\boldsymbol{t}-\boldsymbol{y}$ for some $\boldsymbol{y}\in Q$. Thus there exists a $\boldsymbol{y}\in Q$ such that $x_i+y_i\ge t_i$ for some $i\in\{1,2,\ldots,n\}$. Since C is a lower set, it follows from the definition of Q that there exists a point $\boldsymbol{y}'\in C$ such that $y'_i=y_i$ for that i. For $\boldsymbol{y}'$ we have $x_i+y'_i\ge t_i$, and therefore $\boldsymbol{x}\in\widetilde{C_{\boldsymbol{t}}}$, and (3.7) follows.)

When C is the lower orthant, $C = (-\infty, \boldsymbol{a})$, say, then

$$\widetilde{C_{\boldsymbol{t}}} = (-\infty, \boldsymbol{t} - \boldsymbol{a})^c.$$

From (3.7) we have

$$\begin{aligned}\sup_{C \in \mathcal{O}} &\{P_2(C) - P_1(\pi_1((A(\boldsymbol{t}))^c \cap (\mathbb{R}^n \times C)))\} \\ &= \sup_{\boldsymbol{a} \in \mathbb{R}^n} \{P_2((-\infty, \boldsymbol{a})) - P_1((-\infty, \boldsymbol{t} - \boldsymbol{a})^c)\}.\end{aligned}$$

Therefore

$$M((A(\boldsymbol{t}))^c) = 1 - \sup_{\boldsymbol{a} \in \mathbb{R}^n} \{P_2((-\infty, \boldsymbol{a})) - P_1((-\infty, \boldsymbol{t} - \boldsymbol{a})^c)\}.$$

Hence

$$\begin{aligned}m(A(\boldsymbol{t})) &= 1 - M((A(\boldsymbol{t}))^c) \\ &= \sup_{\boldsymbol{a} \in \mathbb{R}^n} \{P_2((-\infty, \boldsymbol{a})) - P_1((-\infty, \boldsymbol{t} - \boldsymbol{a})^c)\},\end{aligned}$$

and this gives (3.3). ∎

The same method yields bounds on probabilities of upper orthants. This is stated next. The proof is omitted since it is similar to the proof of Theorem 3.3. Let $\mathcal{L}$ be the class of lower open subsets of $\mathbb{R}^n$.

THEOREM 3.4. *For every pair of random vectors $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ having distributions F_1 and F_2, respectively,*

$$\inf_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X} + \boldsymbol{Y} > \boldsymbol{t}\} = \sup_{\boldsymbol{a} \in \mathbb{R}^n} \{P_2((\boldsymbol{a}, \infty)) - P_1((\boldsymbol{t} - \boldsymbol{a}, \infty)^c)\} \tag{3.8}$$

and

$$\sup_{\mathcal{F}_{F_1,F_2}} P\{\boldsymbol{X} + \boldsymbol{Y} \geq \boldsymbol{t}\} = \inf_{C \in \mathcal{L}} \{P_2(C^c) + P_1(\boldsymbol{t} - C)\} \tag{3.9}$$

where P_1 and P_2 are the probability measures associated with F_1 and F_2, respectively.

From Theorems 3.3 and 3.4 we can obtain the sharpness of the bounds given in Theorems 2.1 and 2.2.

COROLLARY 3.5. *Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be a pair of random vectors with distributions F_1 and F_2, respectively. If $\boldsymbol{X} + \boldsymbol{Y}$ has a continuous distribution then the bounds given in (2.5) and in (2.14) are sharp.*

PROOF. It is easy to verify that the lower bounds given in (3.3) and in (3.8) are the same as the lower bounds given in (2.5) and in (2.14), respectively.

In order to show that the upper bound given in (2.5) is sharp, first note that since the upper bound in (3.4) is sharp it follows that for any $\boldsymbol{t} \in \mathbb{R}^n$,

$$U(\boldsymbol{t}) \geq \inf_{C \in \mathcal{U}} \{P_1(C^c) + P_2(\boldsymbol{t} - C)\},$$

where $U(\boldsymbol{t})$ is given in (2.5), and, for convenience, we interchanged the indices 1 and 2 in (3.4). Now we show that, for any $\boldsymbol{t} \in \mathbb{R}^n$,

$$U(\boldsymbol{t}) \leq \inf_{C \in \mathcal{U}} \{P_1(C^c) + P_2(\boldsymbol{t} - C)\}. \tag{3.10}$$

Fix a $\boldsymbol{t} \in \mathbb{R}^n$ and an upper open set C. Let $\boldsymbol{u}_0$ be a boundary point of C (some of the coordinates of $\boldsymbol{u}_0$ may be infinite). Since C is an upper set it follows that $(-\infty, \boldsymbol{u}_0] \subseteq C^c$ and that $(\boldsymbol{u}_0, \infty) \subseteq C$. Hence $F_1(\boldsymbol{u}_0) \leq P_1(C^c)$ and $F_2(\boldsymbol{t} - \boldsymbol{u}_0) \leq P_2(\boldsymbol{t} - C)$. Therefore, for any open upper set C,

$$P_1(C^c) + P_2(\boldsymbol{t} - C) \geq F_1(\boldsymbol{u}_0) + F_2(\boldsymbol{t} - \boldsymbol{u}_0) \geq \inf_{\boldsymbol{u}} \{F_1(\boldsymbol{u}) + F_2(\boldsymbol{t} - \boldsymbol{u})\} \geq U(\boldsymbol{t}),$$

and (3.10) follows.

In a similar fashion it can be shown that the upper bound given in (2.14) is equal to the upper bound given in (3.9). ∎

The results given in Theorems 3.3 and 3.4 can be extended to sums of more than just two random vectors. We describe only the extension of (3.4); the other bounds can be extended similarly. The discussion below is an extension of Proposition 3(a) of Rüschendorf (1982).

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_k$ be random vectors with distributions $F_1, F_2, \ldots, F_k$, respectively. Let $\mathcal{F}_{F_1, F_2, \ldots, F_k}$ denote the class of joint distributions with margins $F_1, F_2, \ldots, F_k$. For $F \in \mathcal{F}_{F_1, F_2, \ldots, F_{k-1}}$ denote by G_F the distribution function of the sum of $k-1$ random vectors that have the joint distribution F. Then

$$\sup_{\mathcal{F}_{F_1, F_2, \ldots, F_k}} P\{\boldsymbol{X}_1 + \boldsymbol{X}_2 + \cdots + \boldsymbol{X}_k \leq \boldsymbol{t}\} = \sup_{\mathcal{F}_{F_1, F_2, \ldots, F_{k-1}}} \inf_{C \in \mathcal{U}} \{P_k(C^c) + Q_F(\boldsymbol{t} - C)\},$$

where P_k and Q_F are the probability measures associated with F_k and G_F, respectively. The proof of this is immediate from (3.4).

4. Inequalities for Expectations of Functions of $\boldsymbol{X} + \boldsymbol{Y}$. In this section we show how Theorems 2.1, 2.2, 3.3 and 3.4 yield bounds on expectations of the form $E[\phi(\boldsymbol{X} + \boldsymbol{Y})]$ for certain classes of functions ϕ.

First recall that for a real n-variate function ϕ, the multivariate difference operator Δ is defined by

$$\Delta_{\boldsymbol{s}}^{\boldsymbol{t}}\phi = \sum_{(\epsilon_1,\epsilon_2,\ldots,\epsilon_n)\in\{0,1\}^n} (-1)^{\Sigma_{i=1}^n \epsilon_i}\phi(\epsilon_1 s_1 + (1-\epsilon_1)t_1, \ldots, \epsilon_n s_n + (1-\epsilon_n)t_n),$$

where $\boldsymbol{s}$ and $\boldsymbol{t}$ are elements of $\mathbb{R}^n$. The function ϕ is called Δ-monotone if

$$\Delta_{\boldsymbol{s}}^{\boldsymbol{t}}\phi \geq 0 \quad \text{whenever} \quad \boldsymbol{s} \leq \boldsymbol{t}.$$

Let $\mathcal{M}$ be the set of all n-variate functions that are Δ-monotone in any k of their coordinates when the other $n-k$ coordinates are held fixed, $0 \leq k \leq n-1$. For example, every distribution function is a member of $\mathcal{M}$. Also, all functions ϕ of the form $\phi(t_1, t_2, \ldots, t_n) = \prod_{i=1}^n \phi_i(t_i)$ belong to $\mathcal{M}$ provided ϕ_i is a nonnegative increasing function, $i = 1, 2, \ldots, n$. An n-times differentiable function ϕ is in $\mathcal{M}$ if, and only if, $\frac{\partial^m}{\partial t_{i_1}\partial t_{i_2}\cdots\partial t_{i_m}}\phi(\boldsymbol{t}) \geq 0$, $\{i_1, i_2, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$. We will consider only these members of $\mathcal{M}$. We then write $a_\phi(\boldsymbol{t}) = \frac{\partial^n}{\partial t_1 \partial t_2 \cdots \partial t_n}\phi(\boldsymbol{t})$ which is well defined and nonnegative. Using Theorem 2.2 (or Theorem 3.4) and ideas of Cambanis, Simons and Stout (1976), Tchen (1980), Rüschendorf (1980) and Mosler (1984) we obtain the following bounds.

THEOREM 4.1. *Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be any random vectors with respective marginal distributions F_1 and F_2. Let $\phi \in \mathcal{M}$ be n-times differentiable, and assume that there exists a $b > -\infty$ such that $\lim_{t_i \to -\infty} \phi(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = b$, $i = 1, 2, \ldots, n$. Then*

$$b + \int_{\mathbb{R}^n} a_\phi(\boldsymbol{t})\bar{l}(\boldsymbol{t})\, d\boldsymbol{t} \leq E[\phi(\boldsymbol{X}+\boldsymbol{Y})] \leq b + \int_{\mathbb{R}^n} a_\phi(\boldsymbol{t})\overline{u}(\boldsymbol{t})\, d\boldsymbol{t}, \tag{4.1}$$

where $\bar{l}$ and $\overline{u}$ are defined in (2.12) and (2.13).

PROOF. Let K denote the distribution of $\boldsymbol{X}+\boldsymbol{Y}$. Writing $\phi(\boldsymbol{t}) = b + \int_{-\infty}^{\boldsymbol{t}} a_\phi(\boldsymbol{x})\, d\boldsymbol{x}$, and applying Fubini's theorem to interchange the order of integration, we obtain

$$\begin{aligned} E[\phi(\boldsymbol{X}+\boldsymbol{Y})] &= b + \int_{\mathbb{R}^n} a_\phi(\boldsymbol{t})\overline{K}(\boldsymbol{t})\, d\boldsymbol{t} \\ &\geq b + \int_{\mathbb{R}^n} a_\phi(\boldsymbol{t})\bar{l}(\boldsymbol{t})\, d\boldsymbol{t}, \end{aligned}$$

where the inequality follows from the fact that $\overline{K}(\boldsymbol{t}) \geq \bar{l}(\boldsymbol{t})$ (see (2.14)).

The proof of the other inequality is similar. ∎

When $\bar{l}$ and $\bar{u}$ happen to be survival functions, associated with the distribution functions l and u, then (4.1) can be written as

$$\int_{\mathbb{R}^n} \phi(\boldsymbol{t})\, dl(\boldsymbol{t}) \le E[\phi(\boldsymbol{X}+\boldsymbol{Y})] \le \int_{\mathbb{R}^n} \phi(\boldsymbol{t})\, du(\boldsymbol{t}),$$

or as

$$(-1)^n \int_{\mathbb{R}^n} \phi(\boldsymbol{t})\, d\bar{l}(\boldsymbol{t}) \le E[\phi(\boldsymbol{X}+\boldsymbol{Y})] \le (-1)^n \int_{\mathbb{R}^n} \phi(\boldsymbol{t})\, d\bar{u}(\boldsymbol{t}).$$

Inequality (4.1) essentially follows from Theorem 2.2 . In a similar fashion one can use Theorem 2.1 (or Theorem 3.3) to obtain the following result.

THEOREM 4.2. *Let* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be any random vectors with respective marginal distributions* F_1 *and* F_2. *Let* $g : \mathbb{R}^n \to \mathbb{R}$ *be n-times differentiable such that* $\phi_g \in \mathcal{M}$ *where* ϕ_g *is defined by* $\phi_g(\boldsymbol{t}) = g(-\boldsymbol{t})$, $\boldsymbol{t} \in \mathbb{R}^n$ *(for example, g satisfies these conditions if* $(-1)^m \frac{\partial^m}{\partial t_{i_1} \partial t_{i_2} \cdots \partial t_{i_m}} g(\boldsymbol{t}) \ge 0$, $\{i_1, i_2, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\}$*). Suppose that there exists a* $b > -\infty$ *such that* $\lim_{t_i \to \infty} g(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = b$, $i = 1, 2, \ldots, n$. *Then*

$$b + (-1)^n \int_{\mathbb{R}^n} a_g(\boldsymbol{t}) L(\boldsymbol{t})\, d\boldsymbol{t} \le E[g(\boldsymbol{X}+\boldsymbol{Y})] \le b + (-1)^n \int_{\mathbb{R}^n} a_g(\boldsymbol{t}) U(\boldsymbol{t})\, d\boldsymbol{t}, \quad (4.2)$$

where L *and* U *are defined in (2.3) and (2.4), and* a_g *denotes the nth partial derivative of* g.

PROOF. A computation similar to the one in the proof of Theorem 4.1 yields

$$\begin{aligned} E[g(\boldsymbol{X}+\boldsymbol{Y})] &= E[\phi_g(-(\boldsymbol{X}+\boldsymbol{Y}))] \\ &= b + \int_{\mathbb{R}^n} a_{\phi_g}(\boldsymbol{t}) P\{-(\boldsymbol{X}+\boldsymbol{Y}) \ge \boldsymbol{t}\}\, d\boldsymbol{t} \\ &= b + \int_{\mathbb{R}^n} a_{\phi_g}(\boldsymbol{t}) P\{\boldsymbol{X}+\boldsymbol{Y} \le -\boldsymbol{t}\}\, d\boldsymbol{t} \\ &= b + \int_{\mathbb{R}^n} a_{\phi_g}(-\boldsymbol{t}) P\{\boldsymbol{X}+\boldsymbol{Y} \le \boldsymbol{t}\}\, d\boldsymbol{t} \\ &= b + (-1)^n \int_{\mathbb{R}^n} a_g(\boldsymbol{t}) P\{\boldsymbol{X}+\boldsymbol{Y} < \boldsymbol{t}\}\, d\boldsymbol{t} \\ &\ge b + (-1)^n \int_{\mathbb{R}^n} a_g(\boldsymbol{t}) L(\boldsymbol{t})\, d\boldsymbol{t}, \end{aligned}$$

where the inequality follows from (2.5).

The other inequality can be proven in a similar fashion. ∎

Functions g of the kind discussed in Theorem 4.2 are studied in Mosler and Scarsini (1991).

Using Theorems 2.1, 2.2, 3.3 and 3.4 we can also obtain bounds on $E[\phi(\boldsymbol{X}+\boldsymbol{Y})]$ for functions ϕ which are increasing with respect to the componentwise order in $\mathbb{R}^n$ (rather than just Δ-monotone). We need the following definitions.

DEFINITION 4.3. *A set function* $M : \mathrm{Bor}(\mathbb{R}^n) \to [0,1]$ *is called a capacity if*

(i) $M(\emptyset)=0$,

(ii) $M(\mathbb{R}^n)=1$,

(iii) $A,B \in \mathrm{Bor}(\mathbb{R}^n),\ A \subseteq B \Longrightarrow M(A) \le M(B)$.

DEFINITION 4.4. *Given a measurable function* $\phi : \mathbb{R}^n \to \mathbb{R}$, *the lower Choquet integral of* ϕ *with respect to the capacity* M *is defined as*

$$\underline{\int} \phi\, dM = \int_0^\infty [1 - M(\{\boldsymbol{t} : \phi(\boldsymbol{t}) \le \alpha\})]\, d\alpha - \int_{-\infty}^0 M(\{\boldsymbol{t} : \phi(\boldsymbol{t}) \le \alpha\})\, d\alpha,$$

and the upper Choquet integral of ϕ *with respect to* M *is defined as*

$$\overline{\int} \phi\, dM = \int_0^\infty M(\{\boldsymbol{t} : \phi(\boldsymbol{t}) > \alpha\})\, d\alpha - \int_{-\infty}^0 [1 - M(\{\boldsymbol{t} : \phi(\boldsymbol{t}) > \alpha\})]\, d\alpha.$$

For definitions and properties of the Choquet integrals see, e.g., Choquet (1953–54), Gilboa (1989) and Denneberg (1994).

Now let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be as in Theorem 2.1 and denote the joint distribution of $\boldsymbol{X}+\boldsymbol{Y}$ by K. Then, from Theorem 2.1, we have that for all $\boldsymbol{t} \in \mathbb{R}^n$,

$$L(\boldsymbol{t}) \le K(\boldsymbol{t}) \le U(\boldsymbol{t}). \tag{4.3}$$

Note that L and U are increasing functions such that for $i = 1, 2, \ldots, n$,

$$\lim_{t_i \to -\infty} L(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = \lim_{t_i \to -\infty} U(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = 0,$$

and

$$\lim_{\boldsymbol{t} \to \infty} L(\boldsymbol{t}) = \lim_{\boldsymbol{t} \to \infty} U(\boldsymbol{t}) = 1.$$

Let P be the probability measure on $(\mathbb{R}^n, \mathrm{Bor}(\mathbb{R}^n))$ associated with K, let $\mathcal{Q}$ be the class of lower orthants

$$\mathcal{Q} = \{(-\infty, \boldsymbol{t}) : \boldsymbol{t} \in \mathbb{R}^n\},$$

and let $\mathcal{Q}^*$ be the class of complements of lower orthants

$$\mathcal{Q}^* = \{\mathbb{R}^n \setminus (-\infty, \boldsymbol{t}) : \boldsymbol{t} \in \mathbb{R}^n\}.$$

Clearly $\mathcal{Q}, \mathcal{Q}^* \subseteq \mathrm{Bor}(\mathbb{R}^n)$. For $B = (-\infty, \boldsymbol{t}) \in \mathcal{Q}$ define

$$M_*(B) = L(\boldsymbol{t}) \quad \text{and} \quad M^*(B) = U(\boldsymbol{t}).$$

Extend M_* and M^* to $\mathrm{Bor}(\mathbb{R}^n)$ as follows. For $A \in \mathrm{Bor}(\mathbb{R}^n)$,

$$\begin{aligned} M_*(A) &= \sup_{\substack{B \subseteq A \\ B \in \mathcal{Q}}} M_*(B) \quad \text{and} \\ M^*(A) &= \inf_{\substack{A \subseteq B \\ B \in \mathcal{Q}}} M^*(B). \end{aligned}$$

Clearly, M^* and M_* are capacities.

For any lower set A we have

$$\begin{aligned} M_*(A) &= \sup_{\substack{B \subseteq A \\ B \in \mathcal{Q}}} M_*(B) \\ &\leq \sup_{\substack{B \subseteq A \\ B \in \mathcal{Q}}} P(B) \\ &\leq P(A) \\ &\leq \inf_{\substack{A \subseteq B \\ B \in \mathcal{Q}}} P(B) \\ &\leq \inf_{\substack{A \subseteq B \\ B \in \mathcal{Q}}} M^*(B) = M^*(A), \end{aligned}$$

where the first and the last inequalities follow from (4.3). Therefore (Dyckerhoff and Mosler (1993), Scarsini (1992)), for all increasing functions $\phi : \mathbb{R}^n \to \mathbb{R}$,

$$\underline{\int} \phi \, dM^* \leq \underline{\int} \phi \, dP \leq \underline{\int} \phi \, dM_*. \tag{4.4}$$

Since $\underline{\int} \phi \, dP = \int \phi \, dP = E[\phi(\boldsymbol{X} + \boldsymbol{Y})]$, (4.4) provides bounds for $E[\phi(\boldsymbol{X} + \boldsymbol{Y})]$ for all increasing ϕ.

Similar bounds can be obtained by defining, for $B = (\mathbb{R}^n \setminus (-\infty, \boldsymbol{t})) \in \mathcal{Q}^*$,

$$N_*(B) = 1 - U(\boldsymbol{t}), \quad \text{and} \quad N^*(B) = 1 - L(\boldsymbol{t}).$$

Now extend N_* and N^* to $\mathrm{Bor}(\mathbb{R}^n)$ by defining

$$N_*(A) = \sup_{\substack{B \subseteq A \\ B \in \mathcal{Q}^*}} N_*(B) \quad \text{and} \quad N^*(A) = \inf_{\substack{A \subseteq B \\ B \in \mathcal{Q}^*}} N^*(B),$$

for $A \in \mathrm{Bor}(\mathbb{R}^n)$. Again N_* and N^* are capacities, and for all upper sets A, a similar argument yields

$$N_*(A) \le P(A) \le N^*(A),$$

which implies that for any increasing function ϕ

$$\overline{\int} \phi \, dN_* \le \overline{\int} \phi \, dP \le \overline{\int} \phi \, dN^*.$$

In summary, for all increasing ϕ,

$$\begin{aligned} \max\left(\underline{\int} \phi \, dM^*, \overline{\int} \phi \, dN_*\right) &\le E[\phi(\boldsymbol{X}+\boldsymbol{Y})] \\ &\le \min\left(\underline{\int} \phi \, dM_*, \overline{\int} \phi \, dN^*\right). \end{aligned} \tag{4.5}$$

Bounds on $E[\phi(\boldsymbol{X}+\boldsymbol{Y})]$, for all increasing ϕ, the analogues of those in (4.5), can be obtained in a similar fashion using Theorem 2.2 (rather than Theorem 2.1). We omit the details.

Acknowledgement. We thank two referees for a very careful reading of the paper. Their illuminating comments led to a significant improvement of the contents and of the presentation of this paper.

References

CAMBANIS, S., SIMONS, G. and STOUT, W. (1976). Inequalities for $\mathcal{E}k(X, Y)$, when the marginals are fixed. *Z. Wahrsch. verw. Gebiete* **36** 285–294.

CHOQUET, G. (1953-54). Theory of capacities. *Ann. Inst. Fourier* **5** 131–295.

DENNEBERG, D. (1994). *Lectures on non-additive measure and integral.* Kluwer Academic Publishers, Boston, Massachusetts.

DYCKERHOFF, R. and MOSLER, K. (1993). Stochastic dominance with nonadditive probabilities. *Z. Oper. Res.* **37** 231–256.

FRANK, M. J., NELSEN, R. B. and SCHWEIZER, B. (1987). Best-possible bounds for the distribution of a sum — a problem of Kolmogorov. *Probab. Theory Related Fields* **74** 199–211.

GILBOA, I. (1989). Duality in non-additive expected utility. *Ann. Oper. Res.* **19** 405–414.

MAKAROV, G. D. (1981). Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probab. Appl.* **26** 803–806.

MOSLER, K. C. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Sci.* **30** 1311–1322.

MOSLER, K. and SCARSINI, M. (1991). Some theory of stochastic dominance. In *Stochastic Orders and Decision Under Risk* (edited by K. Mosler and M. Scarsini). Institute of Mathematical Statistics Lecture Notes — Monograph Series, Hayward, California, 261–284.

RACHEV, S. T. (1991). *Probability Metrics and the Stability of Stochastic Models.* John Wiley & Sons, New York, New York.

RÜSCHENDORF, L. (1980). Inequalities for the expectation of Δ-monotone functions. *Z. Wahrsch. verw. Gebiete* **54** 341–349.

RÜSCHENDORF, L. (1982). Random variables with maximum sums. *Advances in Applied Probability* **14** 623–632.

RÜSCHENDORF, L. (1991). Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals* (edited by G. Dall'Aglio, S. Kotz and G. Salinetti). Kluwer Academic Publishers, Boston, Massachusetts, 151–187.

SCARSINI, M. (1992). Dominance conditions in non-additive expected utility theory. *J. Math. Econom.* **21** 173–184.

SCHWEIZER, B. (1991), Thirty years of copulas. In *Advances in Probability Distributions with Given Marginals* (edited by G. Dall'Aglio, S. Kotz and G. Salinetti). Kluwer Academic Publishers, Boston, Massachusetts, 13–50.

STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–438.

TCHEN, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab.* **8** 814–827.

DEPARTMENT OF PURE AND APPLIED MATHEMATICS
WASHINGTON STATE UNIVERSITY
PULLMAN, WA 99264-3113
lih@haijun.math.wsu.edu

DIPARTIMENTO DI SCIENZE
UNIVERSITÀ D'ANNUNZIO
VIALE PINDARO 42
I–65127 PESCARA, ITALY
scarsini@unich.it

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON AZ 85721
shaked@math.arizona.edu.