KYBERNETIKA — VOLUME *54* (2018), NUMBER 6, PAGES 1264–1283

MULTIVARIATE STOCHASTIC DOMINANCE FOR MULTIVARIATE NORMAL DISTRIBUTION

Barbora Petrová

Stochastic dominance is widely used in comparing two risks represented by random variables or random vectors. There are general approaches, based on knowledge of distributions, which are dedicated to identify stochastic dominance. These methods can be often simplified for specific distribution. This is the case of univariate normal distribution, for which the stochastic dominance rules have a very simple form. It is however not straightforward if these rules are also valid for multivariate normal distribution. We propose the stochastic dominance rules for multivariate normal distribution and provide a rigorous proof. In a computational experiment we employ these rules to test its efficiency comparing to other methods of stochastic dominance detection.

Keywords: multivariate stochastic dominance, multivariate normal distribution, stochastic dominance rules

Classification: 91B16, 91B28

1. INTRODUCTION

Stochastic dominance is a concept which has been widely used in probability theory, decision theory or stochastic optimization for decades. Univariate stochastic dominance is a partial order between random variables which is applied in situations where one decision represented by a random variable can be ranked as superior to another one for a broad class of decision-makers. It is based on shared preferences regarding sets of possible outcomes and their associated probabilities. Only limited knowledge of preferences is required for determining dominance. There exists an extensive theory concerning univariate stochastic dominance of different orders, we refer, for instance, to [10]. Univariate stochastic dominance can be extended into multiple dimension by introducing multivariate stochastic dominance which provides a partial order between random vectors. There is no consent between researches how to define this type of stochastic dominance. Some authors consider independence of marginals of random vectors or restrict a generator of multivariate stochastic dominance to very special classes of utility functions. For more details we refer, for instance, to papers [6, 9] or [12]. Another authors define a mapping which converts random vectors into random variables and consequently they can

DOI: 10.14736/kyb-2018-6-1264

define multivariate stochastic dominance between random vectors as univariate stochastic dominance between the corresponding random variables. This approach is adopted, for instance, by [4]. The most important work related to our paper is the concept by [8] which provides a very general approach to multivariate stochastic dominance stemming from utility theory. This concept was applied by many other researches in solving practical issues.

Stochastic dominance between random variables or random vectors can be detected by examining their distribution functions. However, this approach can be very demanding from the calculation perspective, especially when one deals with random vectors of high dimension. Situation becomes easier if we have information about family from which both distributions come from. In the univariate case, researches have developed stochastic dominance rules for dozens of distributions. These rules enable to avoid working with probability distributions and detect stochastic dominance just based on comparing parameters of the corresponding distributions. In the multivariate case, there has not been many paper focusing on this topic. In [7], we have already proposed stochastic dominance rules for general discrete distribution and continuous uniform distribution. This article aims to investigate stochastic dominance rules for multivariate normal distribution. The work is inspired by the rules valid for univariate normal distribution that has been known for decades. Theorem 1.1 summarizes the rule (for a proof we refer to [11]). There can be found more general results in the literature, see, for instance, [14, 15] or [16].

Theorem 1.1. (Levy [11]) Let X and Y be two normally distributed random variables, $X \sim N(\mu^x, \sigma^x)$ and $Y \sim N(\mu^y, \sigma^y)$. Then X stochastically dominates Y in the first order if and only if $\mu^x \geq \mu^y$ and $\sigma^x = \sigma^y$.

We are interested whether the rules stated in the theorem can be extended into multiple dimension. In other words, considering two random vectors with d-dimensional normal distributions $\boldsymbol{X} \sim N_d(\boldsymbol{\mu}^x, \Sigma^x)$ and $\boldsymbol{Y} \sim N_d(\boldsymbol{\mu}^y, \Sigma^y)$, can we say that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ if and only if $\boldsymbol{\mu}^x \geq \boldsymbol{\mu}^y$ and $\Sigma^x = \Sigma^y$? As a departing point to answer this question, we used a very simple simulation. We generated randomly realizations of two random vectors with 2-dimensional normal distributions with parameters set according to the desired stochastic dominance rules (see Figure 1). In the next step, we employed technique for detecting stochastic dominance between two discrete random vectors with equiprobable scenarios which was described in detail in [7]. We repeated this procedure for different parameters setting and found out that the suggested stochastic dominance rules are empirically valid.

The paper is organized as follows. Section 2 provides an overview on multivariate stochastic dominance. We recall basic definitions and formulate several statements that are relevant for construction of stochastic dominance rules for multivariate normal distribution. Section 3 describes basic properties of multivariate normal distribution that are relevant for our next work. The most important part of this paper is Section 4 in which we formulate stochastic dominance rules for multivariate normal distribution and provide a rigorous proof. In Section 5 we execute a simulation in which we employ the stochastic dominance rules to test their efficiency comparing to another method described in [7]. The secondly mentioned method uses sampling from distributions and

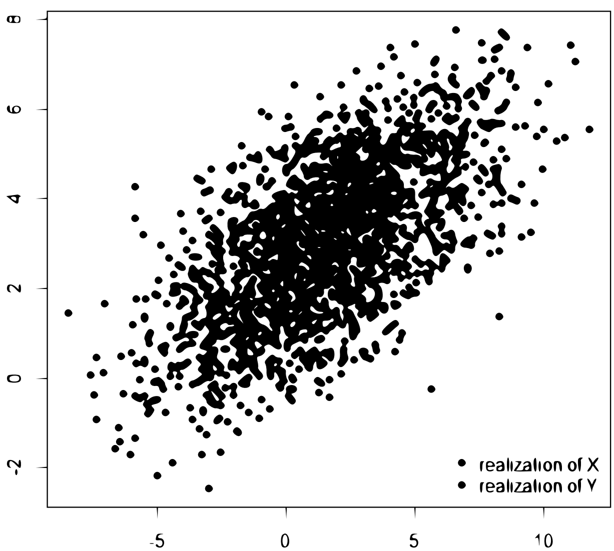

Fig. 1. Simulation of 1000 realizations of 2-dimensional random vectors $\boldsymbol{X} \sim N_2((2,4),\Sigma)$ and $\boldsymbol{Y} \sim N_2((1,2),\Sigma)$.

consequently compares the samples by an optimization procedure in order to detect stochastic dominance. Section 6 summarizes obtained results and concludes the paper.

2. MULTIVARIATE STOCHASTIC DOMINANCE

In the following section we provide a brief introduction to multivariate stochastic dominance that is relevant to our topic. For a general framework of multivariate stochastic dominance, we refer to [7]. In the whole text we employ the usual partial order relation on $\mathbb{R}^d$, that is for any $\boldsymbol{x} = (x_1, \dots, x_d)$ and $\boldsymbol{y} = (y_1, \dots, y_d)$ in $\mathbb{R}^d$, we say that $\boldsymbol{x} \leq \boldsymbol{y}$ if and only if $x_i \leq y_i$, $i = 1, \dots, d$. For defining multivariate stochastic dominance, the concept of upper sets is commonly used.

Definition 2.1. A closed subset $M \subset \mathbb{R}^d$ is called an upper set if for each $\boldsymbol{y} \in \mathbb{R}^d$ such that $\boldsymbol{y} \geq \boldsymbol{x}$ it holds that $\boldsymbol{y} \in M$ whenever $\boldsymbol{x} \in M$. We denote by $\mathcal{M}$ the set of all upper sets in $\mathbb{R}^d$.

The following definition of multivariate stochastic dominance is inspired by work of [8].

Definition 2.2. (Levhari et al. [8]) Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$, denoted as $\boldsymbol{X} \succeq \boldsymbol{Y}$, if for every upper set $M \in \mathcal{M}$ one has $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$.

Remark 2.3. In Definition 2.2 we use close upper sets instead of open upper sets as it is usual in defining one dimensional stochastic dominance. In the univariate case the condition given by Definition 2.2 reduces into $\mathbb{P}(X \geq x) \geq \mathbb{P}(Y \geq x)$ for every $x \in \mathbb{R}$.

The first order stochastic dominance is usually defined as $\mathbb{P}(X \leq x) \leq \mathbb{P}(Y \leq x)$ for every $x \in \mathbb{R}$.

In the next definition we introduce a different type of stochastic dominance which utilizes joint survival functions of considered random vectors. The joint survival function $\bar{F}$ of a random vector $\boldsymbol{X}$ is defined as $\bar{F}(\boldsymbol{m}) = \mathbb{P}(\boldsymbol{X} \geq \boldsymbol{m}) = \mathbb{P}(\boldsymbol{X} \in [m_1, \infty) \times \cdots \times [m_d, \infty))$ for $\boldsymbol{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$.

Definition 2.4. (Shaked and Shanthikumar [17]) Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors with joint survival functions $\bar{F}$ and $\bar{G}$. Then $\boldsymbol{X}$ orthantly dominates $\boldsymbol{Y}$, denoted as $\boldsymbol{X} \succeq_{ort} \boldsymbol{Y}$, if for each $\boldsymbol{m} \in \mathbb{R}^d$ one has $\bar{F}(\boldsymbol{m}) \geq \bar{G}(\boldsymbol{m})$.

When we consider random variables instead of random vectors, both above stated definitions of stochastic dominance are equivalent and we only talk about stochastic dominance. Indeed, in $\mathbb{R}$ all upper sets have the form of intervals $[m, \infty)$, $m \in \mathbb{R}$ and thus $\mathbb{P}(X \in [m, \infty)) = \mathbb{P}(X \geq m) = \bar{F}(m)$. In this case, we obtain a standard definition of the first order univariate stochastic dominance. The situation is more complicated in the multivariate case. While stochastic dominance always implies orthant dominance, the opposite implication is valid only under some restrictive assumptions. For more details about this topic we refer to [7].

There are different types of multivarite stochastic dominance in the literature. Authors often work with linear type of stochastic dominance. This framework can be found for instance in [4]. We state below the definition and a basic rule describing the relation to stochastic dominance defined in Definition 2.2 since we employ this concept in the proof of stochastic dominance rules in Section 4.

Definition 2.5. (Dentcheva and Ruszczýnski [4]) A random vector $\boldsymbol{X}$ dominates linearly a random vector $\boldsymbol{Y}$ in the first order, denoted as $\boldsymbol{X} \succeq_{(1)}^{lin} \boldsymbol{Y}$, if for all $\boldsymbol{c} \in \mathbb{R}_+^d$ one has $\boldsymbol{c}^T\boldsymbol{X} \succeq_1 \boldsymbol{c}^T\boldsymbol{Y}$, where $\succeq_1$ denotes the first order univariate stochastic dominance.

Theorem 2.6. (Műller and Stoyan [13]) If $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$, then $\boldsymbol{X}$ also linearly stochastically dominates $\boldsymbol{Y}$ in the first order.

Proof. We refer to [7] or [13]. $\square$

In the next theorem we clarify the relationship between multivariate stochastic dominance of random vectors and univariate stochastic dominance between coordinates of the considered vectors. The following statement will be very important when proving the stochastic dominance rules for multivariate normal distribution.

Theorem 2.7. (Műller and Stoyan [13]) Assume that the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ stochastically dominates the random vector $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$. Then the random variable X_i dominates in the first order the random variable Y_i for each $i = 1, \ldots, d$.

Proof. We refer to [7] or [13]. $\square$

In the final part of this section we will briefly focus on stochastic dominance between random vectors that have uniform discrete distribution on sets of equal cardinalities. As we have already mentioned in Section 1, we aim to compare efficiency of detecting stochastic dominance between two normally distributed random vectors by stochastic dominance rules with the method using sampling from both distributions. The later mentioned method treats the samples as realizations of random vectors with uniform discrete distribution on sets with the same cardinality.

The following theorem provides a necessary and sufficient condition for stochastic dominance of two random vectors with uniform discrete distribution on sets of the same cardinalities.

Theorem 2.8. (Armbruster and Luedtke [3]) Let $\boldsymbol{X}$ be a d-dimensional random vector with uniform distribution on the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ and let $\boldsymbol{Y}$ be a d-dimensional random vector with uniform distribution on the set $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$, $\boldsymbol{x}_i \in \mathbb{R}^d$ for all $i = 1, \ldots, m$ and $\boldsymbol{y}_j \in \mathbb{R}^d$ for all $j = 1, \ldots, m$. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ if and only if there exists a permutation $\Pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ such that $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$ for all $i = 1, \ldots, m$.

Proof. We refer to [7]. □

When dimension of random vectors is considerably high or if the number of realizations is high, seeking the right permutation can become complicated. Therefore we propose to use optimization. We note that each permutation can be identified with a permutation matrix which is a square binary matrix that has exactly one entry equal to 1 in each row and each column and zeros elsewhere. Denote by z_{ij} elements of a desired permutation matrix ($i, j = 1, \ldots, m$). Then the realization $\boldsymbol{x}_i$ can be assigned to the realization $\boldsymbol{y}_j$ if and only if $\boldsymbol{x}_i \geq \boldsymbol{y}_j$, i. e.

$$\begin{aligned} x_{ik} \geq y_{jk} \ \text{ for each } k = 1, \ldots, d \quad &\Longrightarrow \quad z_{ij} \in \{0, 1\}, \\ x_{ik} < y_{jk} \ \text{ for some } k = 1, \ldots, d \quad &\Longrightarrow \quad z_{ij} = 0. \end{aligned}$$

In other words, the decision variable z_{ij} can acquire values 0 or 1 if $\boldsymbol{x}_i \geq \boldsymbol{y}_j$. In case that $z_{ij} = 1$, the realization $\boldsymbol{x}_i$ is assigned to the realization $\boldsymbol{y}_j$. If $x_{ik} < y_{jk}$ the decision variable z_{ij} has to equal to 0 since the assignment is not possible.

The following optimization problem, inspired by [3], seeks for a desired permutation:

$$\text{(SD)} \quad \underset{z_{ij}}{\text{maximize}} \quad \sum_{i=1}^{m} \sum_{j=1}^{m} z_{ij}$$

$$\begin{aligned} \text{subject to} \quad & \sum_{i=1}^{m} z_{ij} = 1 && j = 1, \ldots, m, && (1) \\ & \sum_{j=1}^{m} z_{ij} = 1 && i = 1, \ldots, m, && (2) \\ & (x_{ik} - y_{jk}) z_{ij} \geq 0 && i, j = 1, \ldots, m, \quad k = 1, \ldots, d, && (3) \\ & z_{ij} \in \{0, 1\} && i, j = 1, \ldots, m. && \end{aligned}$$

Constraints (1) and (2) ensure that the decision variables z_{ij} constitute a permutation matrix. Constraint (3) ensures that z_{ij} can be positive, and thus $\boldsymbol{x}_i$ can be assigned to $\boldsymbol{y}_j$ if and only if $\boldsymbol{x}_i \geq \boldsymbol{y}_j$. We emphasize that the formulation of the objective function is not important, we only need to know if the set of feasible solutions is not empty. In this particular case, if the optimization problem has a feasible solution then the optimal value equals to m.

3. CHARACTERISTICS OF MULTIVARIATE NORMAL DISTRIBUTION

In this short section we state a definition of multivariate normal distribution mainly for notational purposes and recall basic properties of this distribution.

Definition 3.1. Let $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^T$ be a d-dimensional random vector with independent marginals such that each marginal has standard normal distribution, i. e. $Z_i \sim N(0, 1)$ for $i = 1, \ldots, d$. Let A be a $d \times n$ matrix with the range equal to n ($n \leq d$) and let $\boldsymbol{\mu} \in \mathbb{R}^d$ be a vector of constants. Then $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$ has multivariate normal distribution with parameters $\boldsymbol{\mu}$ and $\Sigma = AA^T$, i. e. $\boldsymbol{X} \sim N_d(\boldsymbol{\mu}, \Sigma)$.

The following theorems summarize characteristics of multivariate normal distribution that will be crucial for the proof of stochastic dominance rule formulated in the next part. Let us divide a d-dimensional random vector $\boldsymbol{X}$ with multivariate normal distribution $N_d(\boldsymbol{\mu}, \Sigma)$ as follows:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{4}$$

where $\boldsymbol{X}_1$ is a q-dimensional random vector, $\boldsymbol{X}_2$ is a $(d-q)$-dimensional random vector, $\boldsymbol{\mu}_1$ is a q-dimensional vector of constants, $\boldsymbol{\mu}_2$ is a $(d-q)$-dimensional vector of constants, and $\Sigma_{11}, \Sigma_{12}, \Sigma_{21}$ and Σ_{22} are matrices of constants of sizes $q \times q, (d-q) \times q, q \times (d-q)$ and $(d-q) \times (d-q)$.

Theorem 3.2. (Tong [18]) Let $\boldsymbol{X}$ be a d-dimensional random vector with multivariate normal distribution $N_d(\boldsymbol{\mu}, \Sigma)$. Consider a partition of the vector $\boldsymbol{X}$ as described by (4). Then

(1) q-dimensional random vector $\boldsymbol{X}_1$ has multivariate normal distribution with parameters $\boldsymbol{\mu}_1$ and Σ_{11}, i. e. $\boldsymbol{X}_1 \sim N_q(\boldsymbol{\mu}_1, \Sigma_{11})$,

(2) if the matrix Σ_{22} is regular, the conditional distribution of $\boldsymbol{X}_1$ given $\boldsymbol{X}_2 = \boldsymbol{x}_2$ is multivariate normal with parameters $\bar{\boldsymbol{\mu}}$ and $\bar{\Sigma}$, i. e. $\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2 \sim N_q(\bar{\boldsymbol{\mu}}, \bar{\Sigma})$, where $\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$ and $\bar{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

Theorem 3.3. (Tong [18]) Let $\boldsymbol{X} = (X_1, X_2)$ be a 2-dimensional random vector with bivariate normal distribution $N_2(\boldsymbol{\mu}, \Sigma)$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

Then the random variable $X_1 + X_2$ has normal distribution $N(\bar{\mu}, \bar{\sigma})$, where $\bar{\mu} = \mu_1 + \mu_2$ and $\bar{\sigma} = \sqrt{\sigma_{11}^2 + \sigma_{22}^2 + 2\sigma_{12}}$.

In the last definition we recall a special case of bivariate distribution. We assume a 2-dimensional random vector $\boldsymbol{X} = (X_1, X_2)$ for which both elements have standard normal distribution and their correlation equals to $\rho \in (-1,1)$, i.e. $X_1 \sim N_1(0,1)$, $X_2 \sim N_1(0,1)$ and $corr(X_1, X_2) = \rho$.

Definition 3.4. (Tong [18]) The random vector $\boldsymbol{X} = (X_1, X_2)$ is said to have standard bivariate normal distribution with correlation coefficient ρ if its joint probability density function is given by

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ -\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)} \right\}$$

where $\rho \in (-1,1)$.

4. MULTIVARIATE STOCHASTIC DOMINANCE FOR MULTIVARIATE NORMAL DISTRIBUTION

In this section we formulate stochastic dominance rules for multivariate normal distribution and provide a detailed proof. Our proof of the statement differs from that one in [13] by employing the concept of upper sets.

Theorem 4.1. (Műller and Stoyan [13]) Let $\boldsymbol{X}$ be a d-dimensional random vector with normal distribution $N_d(\boldsymbol{\mu}^X, \Sigma^X)$ and let $\boldsymbol{Y}$ be a d-dimensional random vector with normal distribution $N_d(\boldsymbol{\mu}^Y, \Sigma^Y)$. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ if and only if $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y$.

Remark 4.2. In the univariate case the above stated theorem is identical to Theorem 1.1 formulated in Section 1. When proving the statement in the univariate case, authors usually rely on the relation between cumulative distribution functions, or equivalently on the relation between survival functions. As it has been already indicated in Section 2, this approach is not possible in multiple dimension. The relation between joint survival functions corresponds to orthant stochastic dominance and this type of stochastic dominance does not imply stochastic dominance in a general setting. Therefore another technique has to be applied when proving the above formulated stochastic dominance rule.

Proof. We firstly assume that $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y = \Sigma$ and show that in this case $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$. The proof of the statement proceeds by induction on the number of dimensions d. For $d = 1$ we refer to Theorem 1.1. For $d = 2$ let us consider the random vectors $\boldsymbol{X} = (X_1, X_2)^T \sim N_2(\boldsymbol{\mu}^X, \Sigma)$ and $\boldsymbol{Y} = (Y_1, Y_2)^T \sim N_2(\boldsymbol{\mu}^Y, \Sigma)$ and an auxiliary random vector $\boldsymbol{Z} = (Z_1, Z_2)^T \sim N_2(\boldsymbol{\mu}^Z, \Sigma)$, where

$$\boldsymbol{\mu}^X = \begin{pmatrix} \mu_1^X \\ \mu_2^X \end{pmatrix}, \quad \boldsymbol{\mu}^Y = \begin{pmatrix} \mu_1^Y \\ \mu_2^Y \end{pmatrix}, \quad \boldsymbol{\mu}^Z = \begin{pmatrix} \mu_1^Y \\ \mu_2^X \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

Since stochastic dominance is transitive relation it suffices to show that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$ and $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$. In other words, we will prove that

$\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Z} \in M)$ and $\mathbb{P}(\boldsymbol{Z} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for an arbitrary upper set $M \in \mathbb{R}^2$. For each $x_2 \in \mathbb{R}$ we define a set $M_2(x_2) = \{x_1 : (x_1, x_2) \in M\}$. Similarly, we define a set $M_1(x_1) = \{x_2 : (x_1, x_2) \in M\}$. Both sets are upper sets in $\mathbb{R}$ since M is an upper set in $\mathbb{R}^2$. In fact both upper sets are intervals of type $(m_1, \infty]$, $m_1 \in \mathbb{R}$.

Let us start with showing that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$. According to Theorem 3.2, the second coordinate of each random vector is normally distributed, namely $X_2 \sim F_2 = N_1(\mu_2^X, \sigma_{22})$ and $Z_2 \sim H_2 = N_1(\mu_2^X, \sigma_{22})$. Take an arbitrary upper set $M \in \mathbb{R}^2$. We can derive the following inequality

$$\mathbb{P}(\boldsymbol{X} \in M) = \int_{\mathbb{R}} \mathbb{P}(X_1 \in M_2(x_2) \mid X_2 = x_2) \,\mathrm{d}F_2(x_2) = \int_{\mathbb{R}} \mathbb{P}(\bar{X}_1 \in M_2(x_2)) \,\mathrm{d}F_2(x_2),$$
$$\mathbb{P}(\boldsymbol{Z} \in M) = \int_{\mathbb{R}} \mathbb{P}(Z_1 \in M_2(x_2) \mid Z_2 = x_2) \,\mathrm{d}H_2(x_2) = \int_{\mathbb{R}} \mathbb{P}(\bar{Z}_1 \in M_2(x_2)) \,\mathrm{d}H_2(x_2),$$

where $\bar{X}_1 \sim N_1(\bar{\mu}_1^X, \bar{\sigma}_1)$ and $\bar{Z}_1 \sim N_1(\bar{\mu}_1^Z, \bar{\sigma}_1)$. The parameters of the distributions can be calculated as:

$$\begin{aligned}
\bar{\mu}_1^X &= \mu_1^X + \sigma_{12}\sigma_{22}^{-1}(x_2 - \mu_2^X),\\
\bar{\mu}_1^Z &= \mu_1^Y + \sigma_{12}\sigma_{22}^{-1}(x_2 - \mu_2^X),\\
\bar{\sigma}_1 &= \sigma_{11} - \sigma_{12}^2\sigma_{22}^{-1}.
\end{aligned}$$

Since the distributions F_2 and H_2 are both normal with the same parameters, we have

$$\int_{\mathbb{R}} \mathbb{P}(\bar{X}_1 \in M_2(x_2)) \,\mathrm{d}F_2(x_2) \geq \int_{\mathbb{R}} \mathbb{P}(\bar{Z}_1 \in M_2(x_2)) \,\mathrm{d}H_2(x_2)$$
$$\text{if } \ \mathbb{P}(\bar{X}_1 \in M_2(x_2)) \geq \mathbb{P}(\bar{Z}_1 \in M_2(x_2)) \ \text{ for all } M_2(x_2) \in \mathbb{R}.$$

We observe that the random variables $\bar{X}_1$ and $\bar{Z}_1$ have normal distributions with the same variance and $\bar{\mu}_1^X \geq \bar{\mu}_1^Z$. Therefore according to Theorem 1.1, $\bar{X}_1$ stochastically dominates $\bar{Z}_1$ in the first order. Since $M_2(x_2)$ is an upper set in $\mathbb{R}$, we must have $\mathbb{P}(\bar{X}_1 \in M_2(x_2)) \geq \mathbb{P}(\bar{Z}_1 \in M_2(x_2))$ due to stochastic dominance. We have thus shown that $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Z} \in M)$ for an arbitrary choice of upper set in $\mathbb{R}^2$ and therefore $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$.

In order to show that $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$, we proceed similarly as in the previous case. By Theorem 3.2, the first coordinate of each random vector is normally distributed, namely $Z_1 \sim H_1 = N_1(\mu_1^Y, \sigma_{11})$ and $Y_1 \sim G_1 = N_1(\mu_1^Y, \sigma_{11})$. We now have

$$\mathbb{P}(\boldsymbol{Z} \in M) = \int_{\mathbb{R}} \mathbb{P}(Z_2 \in M_1(x_1) \mid Z_1 = x_1) \,\mathrm{d}H_1(x_1) = \int_{\mathbb{R}} \mathbb{P}(\bar{Z}_2 \in M_1(x_1)) \,\mathrm{d}H_1(x_1),$$
$$\mathbb{P}(\boldsymbol{Y} \in M) = \int_{\mathbb{R}} \mathbb{P}(Y_2 \in M_1(x_1) \mid Y_1 = x_1) \,\mathrm{d}G_1(x_1) = \int_{\mathbb{R}} \mathbb{P}(\bar{Y}_2 \in M_1(x_1)) \,\mathrm{d}G_1(x_1),$$

where $\bar{Z}_2 \sim N_1(\bar{\mu}_2^Z, \bar{\sigma}_2)$ and $\bar{Y}_2 \sim N_1(\bar{\mu}_2^Y, \bar{\sigma}_2)$. The parameters of the distributions are given as:

$$\begin{aligned}
\bar{\mu}_2^Z &= \mu_2^X + \sigma_{12}\sigma_{11}^{-1}(x_1 - \mu_1^Y),\\
\bar{\mu}_2^Y &= \mu_2^Y + \sigma_{12}\sigma_{11}^{-1}(x_1 - \mu_1^Y),\\
\bar{\sigma}_2 &= \sigma_{22} - \sigma_{12}^2\sigma_{11}^{-1}.
\end{aligned}$$

Since the distributions H_1 and G_1 are both normal with the same parameters, we have

$$\int_{\mathbb{R}} \mathbb{P}(\bar{Z}_2 \in M_1(x_1))\,\mathrm{d}H_1(x_1) \geq \int_{\mathbb{R}} \mathbb{P}(\bar{Y}_2 \in M_1(x_1))\,\mathrm{d}G_1(x_1)$$
$$\text{if } \ \mathbb{P}(\bar{Z}_2 \in M_1(x_1)) \geq \mathbb{P}(\bar{Y}_2 \in M_1(x_1)) \ \text{ for all } M_1(x_1) \in \mathbb{R}.$$

The random variables $\bar{Z}_2$ and $\bar{Y}_2$ have normal distributions with the same variance and $\bar{\mu}_2^Z \geq \bar{\mu}_2^Y$ and therefore $\bar{Z}_2$ stochastically dominates $\bar{Y}_2$ in the first order. Since $M_1(x_1)$ is an upper set in $\mathbb{R}$, we must have $\mathbb{P}(\bar{Z}_2 \in M_1(x_1)) \geq \mathbb{P}(\bar{Y}_2 \in M_1(x_1))$. We conclude that $\mathbb{P}(\boldsymbol{Z} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for an arbitrary choice of upper set in $\mathbb{R}^2$ and therefore $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$.

We proceed the proof with the induction step. We assume that if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are two $(d-1)$-dimensional random vectors such that $\boldsymbol{X} \sim N_{d-1}(\boldsymbol{\mu}^X, \Sigma^X)$ and $\boldsymbol{Y} \sim N_{d-1}(\boldsymbol{\mu}^Y, \Sigma^Y)$, $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y$, then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$. Now let us consider d-dimensional random vectors $\boldsymbol{X} = (X_1, \ldots, X_d)^T \sim N_d(\boldsymbol{\mu}^X, \Sigma^X)$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_2)^T \sim N_d(\boldsymbol{\mu}^Y, \Sigma^Y)$ such that $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y = \Sigma$. We again define an auxiliary random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_d)^T \sim N_d(\boldsymbol{\mu}^Z, \Sigma)$. The parameters can be written down as

$$\boldsymbol{\mu}^X = \begin{pmatrix} \boldsymbol{\mu}_1^X \\ \mu_d^X \end{pmatrix} = \begin{pmatrix} \mu_1^X \\ \vdots \\ \mu_{d-1}^X \\ \mu_d^X \end{pmatrix}, \qquad \boldsymbol{\mu}^Y = \begin{pmatrix} \boldsymbol{\mu}_1^Y \\ \mu_d^Y \end{pmatrix} = \begin{pmatrix} \mu_1^Y \\ \vdots \\ \mu_{d-1}^Y \\ \mu_d^Y \end{pmatrix},$$

$$\boldsymbol{\mu}^Z = \begin{pmatrix} \boldsymbol{\mu}_1^Z \\ \mu_d^X \end{pmatrix} = \begin{pmatrix} \mu_1^X \\ \vdots \\ \mu_{d-1}^Y \\ \mu_d^X \end{pmatrix},$$

where $\boldsymbol{\mu}_1^X$, $\boldsymbol{\mu}_1^Y$ and $\boldsymbol{\mu}_1^Z$ are $(d-1)$-dimensional vectors of parameters, and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \sigma_{dd} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_{dd} \end{pmatrix},$$

where Σ_{11} corresponds to the $(d-1)$-dimensional left-top submatrix of Σ, Σ_{12} and Σ_{21} are the d-column, the d-th row respectively, of the matrix Σ. We employ transitivity of stochastic dominance and show that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$ and $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$. We will again prove that $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Z} \in M)$ and $\mathbb{P}(\boldsymbol{Z} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for an arbitrary upper set $M \in \mathbb{R}^d$. For each $x_i \in \mathbb{R}$, $i = 1, \ldots, d$, we define a set $M_i(x_i) = \{(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) : (x_1, \ldots, x_d) \in M\}$. This set is an upper set in $\mathbb{R}^{d-1}$, since M is an upper set in $\mathbb{R}^d$.

Firstly we show that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$. According to Theorem 3.2, the last coordinate of each random vector is normally distributed, i.e. $X_d \sim F_d = N_1(\mu_d^X, \sigma_{dd})$ and $Z_d \sim H_d = N_1(\mu_d^X, \sigma_{dd})$. Take an arbitrary upper set $M \in \mathbb{R}^d$. We can derive the following inequality

$$\begin{aligned}
\mathbb{P}(\boldsymbol{X} \in M) &= \int_{\mathbb{R}} \mathbb{P}((X_1, \dots, X_{d-1}) \in M_d(x_d) \mid X_d = x_d) \, \mathrm{d}F_d(x_d) \\
&= \int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{X}}_1 \in M_d(x_d)) \, \mathrm{d}F_d(x_d), \\
\mathbb{P}(\boldsymbol{Z} \in M) &= \int_{\mathbb{R}} \mathbb{P}((Z_1, \dots, Z_{d-1}) \in M_d(x_d) \mid Z_d = x_d) \, \mathrm{d}H_d(x_d) \\
&= \int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Z}}_1 \in M_d(x_d)) \, \mathrm{d}H_d(x_d),
\end{aligned}$$

where $\bar{\boldsymbol{X}}_1 \sim N_{d-1}(\bar{\boldsymbol{\mu}}_1^X, \bar{\Sigma}_1)$ and $\bar{\boldsymbol{Z}}_1 \sim N_{d-1}(\bar{\boldsymbol{\mu}}_1^Z, \bar{\Sigma}_1)$. By Theorem 3.2, the parameters of the distributions can be calculated as:

$$\begin{aligned}
\bar{\boldsymbol{\mu}}_1^X &= \boldsymbol{\mu}_1^X + \sigma_{dd}^{-1}(x_d - \mu_d^X)\Sigma_{12} \\
\bar{\boldsymbol{\mu}}_1^Z &= \boldsymbol{\mu}_1^Z + \sigma_{dd}^{-1}(x_d - \mu_d^X)\Sigma_{12} \\
\bar{\Sigma}_1 &= \Sigma_{11} - \sigma_{dd}^{-1}\Sigma_{12}\Sigma_{21}.
\end{aligned}$$

Since the distributions F_d and H_d are both normal with the same parameters, we have

$$\int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{X}}_1 \in M_d(x_d)) dF_d(x_d) \geq \int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Z}}_1 \in M_d(x_d)) \, \mathrm{d}H_d(x_d)$$
$$\text{if } \mathbb{P}(\bar{\boldsymbol{X}}_1 \in M_d(x_d)) \geq \mathbb{P}(\bar{\boldsymbol{Z}}_1 \in M_d(x_d)) \text{ for all } M_d(x_d) \in \mathbb{R}^{d-1}.$$

The random vectors $\bar{\boldsymbol{X}}_1$ and $\bar{\boldsymbol{Z}}_1$ have $(d-1)$-dimensional normal distribution with the same covariance matrix and $\bar{\boldsymbol{\mu}}_1^X \geq \bar{\boldsymbol{\mu}}_1^Z$. By assumption we must have that $\bar{\boldsymbol{X}}_1$ stochastically dominates $\bar{\boldsymbol{Z}}_1$. Since $M_d(x_d)$ is an upper in $\mathbb{R}^{d-1}$ we get $\mathbb{P}(\bar{\boldsymbol{X}}_1 \in M_d(x_d)) \geq \mathbb{P}(\bar{\boldsymbol{Z}}_1 \in M_d(x_d))$. We have shown that $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Z} \in M)$ for an arbitrary choice of upper set in $\mathbb{R}^d$ and therefore $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Z}$.

Now we need to show that $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$. The $(d-1)$-th coordinate of each random vector is normally distributed, i. e. $Z_{d-1} \sim H_{d-1} = N_1(\mu_{d-1}^Y, \sigma_{d-1,d-1})$ and $Y_{d-1} \sim G_{d-1} = N_1(\mu_{d-1}^Y, \sigma_{d-1,d-1})$. We can obtain the following inequality

$$\mathbb{P}(\boldsymbol{Z} \in M) = \int_{\mathbb{R}} \mathbb{P}((Z_1, \dots, Z_{d-2}, Z_d) \in M_{d-1}(x_{d-1}) \mid Z_{d-1} = x_{d-1}) \, \mathrm{d}H_{d-1}(x_{d-1}) =$$
$$\int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Z}}_2 \in M_{d-1}(x_{d-1})) \, \mathrm{d}H_{d-1}(x_{d-1}),$$
$$\mathbb{P}(\boldsymbol{Y} \in M) = \int_{\mathbb{R}} \mathbb{P}((Y_1, \dots, Y_{d-2}, Y_d) \in M_{d-1}(x_{d-1}) \mid Y_{d-1} = x_{d-1}) \, \mathrm{d}G_{d-1}(x_{d-1}) =$$
$$\int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Y}}_2 \in M_{d-1}(x_{d-1})) \, \mathrm{d}G_{d-1}(x_{d-1}),$$

where $\bar{\boldsymbol{Z}}_2 \sim N_{d-1}(\bar{\boldsymbol{\mu}}_2^Z, \bar{\Sigma}_2)$ and $\bar{\boldsymbol{Y}}_2 \sim N_{d-1}(\bar{\boldsymbol{\mu}}_2^Y, \bar{\Sigma}_2)$. Derivation of parameters $\bar{\boldsymbol{\mu}}_2^Z$, $\bar{\boldsymbol{\mu}}_2^Y$ and $\bar{\Sigma}_2$ is not straightforward as in the previous case. In order to obtain the formulas, we need to employ the parameters of the random vectors $\tilde{\boldsymbol{Z}}$ and $\tilde{\boldsymbol{Y}}$ which are created by switching the last two coordinates of the vectors $\boldsymbol{Z}$ and $\boldsymbol{Y}$, i. e. $\tilde{\boldsymbol{Z}} = (Z_1, \dots, Z_{d-2}, Z_d, Z_{d-1})$ and $\tilde{\boldsymbol{Y}} = (Y_1, \dots, Y_{d-2}, Y_d, Y_{d-1})$. Obviously, these vectors have normal distributions $\tilde{\boldsymbol{Z}} \sim N_d(\tilde{\boldsymbol{\mu}}^Z, \tilde{\Sigma})$, $\tilde{\boldsymbol{Y}} \sim N_d(\tilde{\boldsymbol{\mu}}^Y, \tilde{\Sigma})$, where

$$\tilde{\boldsymbol{\mu}}^Z = \begin{pmatrix} \tilde{\boldsymbol{\mu}}_1^X \\ \mu_{d-1}^Y \end{pmatrix} = \begin{pmatrix} \mu_1^X \\ \vdots \\ \mu_{d-2}^X \\ \mu_d^X \\ \mu_{d-1}^Y \end{pmatrix}, \qquad \tilde{\boldsymbol{\mu}}^Y = \begin{pmatrix} \tilde{\boldsymbol{\mu}}_1^Y \\ \mu_{d-1}^Y \end{pmatrix} = \begin{pmatrix} \mu_1^Y \\ \vdots \\ \mu_{d-2}^Y \\ \mu_d^Y \\ \mu_{d-1}^Y \end{pmatrix},$$

where $\tilde{\boldsymbol{\mu}}_1^X$ and $\tilde{\boldsymbol{\mu}}_1^Y$ are $(d-1)$-dimensional vectors of parameters, and

$$\tilde{\Sigma} = \begin{pmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\sigma}_{dd} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1,d-2} & \sigma_{1d} & \sigma_{1,d-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \sigma_{d-2,1} & \cdots & \sigma_{d-2,d-2} & \sigma_{d-2,d} & \sigma_{d-2,d-1} \\ \sigma_{d,1} & \cdots & \sigma_{d,d-2} & \sigma_{d,d} & \sigma_{d,d-1} \\ \sigma_{d-1,1} & \cdots & \sigma_{d-1,d-2} & \sigma_{d-1,d} & \sigma_{d-1,d-1} \end{pmatrix},$$

where $\tilde{\Sigma}_{11}$ corresponds to the $(d-1)$-dimensional left-top submatrix of $\tilde{\Sigma}$, $\tilde{\Sigma}_{12}$ and $\tilde{\Sigma}_{21}$ are the d-column, the d-th row respectively, of the matrix $\tilde{\Sigma}$. By Theorem 3.2, the parameters of the distributions can be calculated as:

$$\begin{aligned} \bar{\boldsymbol{\mu}}_2^Z &= \tilde{\boldsymbol{\mu}}_1^X + \sigma_{d-1,d-1}^{-1}(x_{d-1} - \mu_{d-1}^Y)\tilde{\Sigma}_{12}, \\ \bar{\boldsymbol{\mu}}_2^Y &= \tilde{\boldsymbol{\mu}}_1^Y + \sigma_{d-1,d-1}^{-1}(x_{d-1} - \mu_{d-1}^Y)\tilde{\Sigma}_{12}, \\ \bar{\Sigma}_2 &= \tilde{\Sigma}_{11} - \sigma_{d-1,d-1}^{-1}\tilde{\Sigma}_{12}\tilde{\Sigma}_{21}. \end{aligned}$$

Since the distributions H_{d-1} and G_{d-1} are both normal with the same parameters, then

$$\int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Z}}_2 \in M_{d-1}(x_{d-1}))dH_{d-1}(x_{d-1}) \geq \int_{\mathbb{R}} \mathbb{P}(\bar{\boldsymbol{Y}}_2 \in M_{d-1}(x_{d-1}))\,\mathrm{d}G_{d-1}(x_{d-1})$$
$$\text{if } \mathbb{P}(\bar{\boldsymbol{Z}}_2 \in M_{d-1}(x_{d-1})) \geq \mathbb{P}(\bar{\boldsymbol{Y}}_2 \in M_{d-1}(x_{d-1})) \text{ for all } M_{d-1}(x_{d-1}) \in \mathbb{R}^{d-1}.$$

The random vectors $\bar{\boldsymbol{Z}}_2$ and $\bar{\boldsymbol{Y}}_2$ have $(d-1)$-dimensional normal distribution with the same covariance matrix and $\bar{\boldsymbol{\mu}}_2^Z \geq \bar{\boldsymbol{\mu}}_2^Y$. By assumption we must have that $\bar{\boldsymbol{Z}}_2$ stochastically dominates $\bar{\boldsymbol{Y}}_2$. Since $M_{d-1}(x_{d-1})$ is an upper in $\mathbb{R}^{d-1}$ we get $\mathbb{P}(\bar{\boldsymbol{Z}}_2 \in M_{d-1}(x_{d-1})) \geq \mathbb{P}(\bar{\boldsymbol{Y}}_2 \in M_{d-1}(x_{d-1}))$. We have shown that $\mathbb{P}(\boldsymbol{Z} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for an arbitrary choice of upper set in $\mathbb{R}^d$ and therefore $\boldsymbol{Z}$ stochastically dominates $\boldsymbol{Y}$.

We now aim to show the opposite implication of the statement, i. e. let us assume two d-dimensional random vectors $\boldsymbol{X} \sim N_d(\boldsymbol{\mu}^X, \Sigma^X)$ and $\boldsymbol{Y} \sim N_d(\boldsymbol{\mu}^Y, \Sigma^Y)$. Assume that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$. We need to show that in this case $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y$. By Theorem 2.7, the random vector $\boldsymbol{X}$ stochastically dominates the random vector $\boldsymbol{Y}$ element-wise, i. e. X_i stochastically dominates Y_i for every $i = 1, \dots, d$ in the first order. Moreover, both random variables have normal distribution, namely $X_i \sim N(\mu_i^X, \sigma_{ii}^X)$ and $Y_i \sim N(\mu_i^Y, \sigma_{ii}^Y)$. Employing Theorem 1.1 we must have $\mu_i^X \geq \mu_i^Y$ and $\sigma_{ii}^X = \sigma_{ii}^Y$

for each $i = 1, \dots, d$. It remains to show that one also has $\sigma_{ij}^X = \sigma_{ij}^Y$ for arbitrary $i, j = 1, \dots, d$ such that $i \neq j$. In order to show this result, we use the concept of linear stochastic dominance given by Definition 2.5. Take arbitrary $i, j \in \{1, \dots, d\}, i \neq j$. If $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ then the random variable $X_i/2 + X_j/2$ stochastically dominates in the first order the random variable $Y_i/2 + Y_j/2$, or equivalently the random variable $X_i + X_j$ stochastically dominates in the first order the random variable $Y_i + Y_j$. According to Theorem 3.3, $X_i + X_j$ has normal distribution $N(\mu_i^X + \mu_j^X, ((\sigma_{ii}^X)^2 + (\sigma_{jj}^X)^2 + 2\sigma_{ij}^X)^{1/2})$ and $Y_i + Y_j$ has normal distribution $N(\mu_i^Y + \mu_j^Y, ((\sigma_{ii}^Y)^2 + (\sigma_{jj}^Y)^2 + 2\sigma_{ij}^Y)^{1/2})$. Since $X_i + X_j$ stochastically dominates $Y_i + Y_j$, we must have

$$(\sigma_{ii}^X)^2 + (\sigma_{jj}^X)^2 + 2\sigma_{ij}^X = (\sigma_{ii}^Y)^2 + (\sigma_{jj}^Y)^2 + 2\sigma_{ij}^Y.$$

We have already shown that $\sigma_{ii}^X = \sigma_{ii}^Y$. Therefore the previous equation is valid if and only if $\sigma_{ij}^X = \sigma_{ij}^Y$, which finishes the proof. □

Example 4.3. In this example we demonstrate that comparing joint survival functions of two normally distributed random vectors is not sufficient for claiming stochastic dominance. Let us consider two 2-dimensional random vectors $\boldsymbol{X} = (X_1, X_2)$ and $\boldsymbol{Y} = (Y_1, Y_2)$ with normal distributions with parameters given as:

$$\boldsymbol{\mu}^X = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{\mu}^Y = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \Sigma^X = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}, \qquad \Sigma^Y = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i. e. $\boldsymbol{X} \sim N_2(\boldsymbol{\mu}^X, \Sigma^X)$ and $\boldsymbol{Y} \sim N_2(\boldsymbol{\mu}^Y, \Sigma^Y)$. By Definition 3.4, the random vector $\boldsymbol{X}$ has standard bivariate normal distribution with the correlation coefficient $\rho^X = 1/2$ and the random vector $\boldsymbol{Y}$ has standard bivariate normal distribution with the correlation coefficient $\rho^Y = 0$. Let us denote by $\bar{F}$ the joint survival function of the random vector $\boldsymbol{X}$ and by $\bar{G}$ the joint survival function of the random vector $\boldsymbol{Y}$.

In the first part we prove that $\bar{F}(m_1, m_2) \geq \bar{G}(m_1, m_2)$ for every $\boldsymbol{m} = (m_1, m_2) \in \mathbb{R}^2$. Note that the distributions of both random vectors are standard bivariate normal and differ only in the correlation coefficient. Let us consider a random vector $\boldsymbol{Z} = (Z_1, Z_2)$ with standard bivariate normal distribution with the correlation coefficient ρ and denote its joint distribution function by Φ_ρ. Thus Φ_ρ at point $(m_1, m_2) \in \mathbb{R}^2$ is given as

$$\Phi_\rho(m_1, m_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{m_2}^{\infty} \int_{m_1}^{\infty} \exp\left\{-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right\} \mathrm{d}x_1 \, \mathrm{d}x_2.$$

Obviously, we have $\bar{F}(m_1, m_2) = \Phi_{1/2}(m_1, m_2)$ and $\bar{G}(m_1, m_2) = \Phi_0(m_1, m_2)$. Therefore to show the inequality $\bar{F}(m_1, m_2) \geq \bar{G}(m_1, m_2)$ for every $\boldsymbol{m} = (m_1, m_2) \in \mathbb{R}^2$, it is sufficient to prove that the function $\Phi_\rho(m_1, m_2)$ is a strictly increasing function with respect to the parameter ρ. Let us denote the integrand in the previous equation by $\theta(x_1, x_2, \rho)$. By direct calculation we have

$$\frac{\partial}{\partial \rho} \int_{m_1}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \, \theta(x_1, x_2, \rho) \, \mathrm{d}x_1 = \frac{1}{\sqrt{2\pi(1-\rho^2)^3}} \, (x_2 - \rho m_1) \exp\left\{-\frac{m_1^2 - 2\rho m_1 x_2 + x_2^2}{2(1-\rho^2)}\right\},$$

and therefore by direct calculation we get

$$\int_{m_2}^{\infty}\left(\frac{\partial}{\partial\rho}\int_{m_1}^{\infty}\frac{1}{2\pi\sqrt{1-\rho^2}}\,\theta(x_1,x_2,\rho)\,\mathrm{d}x_1\right)\mathrm{d}x_2 = \frac{1}{2\pi\sqrt{1-\rho^2}}\,\exp\left\{-\frac{m_1^2-2\rho m_1m_2+m_2^2}{2(1-\rho^2)}\right\} = \phi_\rho(m_1,m_2).$$

Because the function ϕ_ρ is a probability density function, we may reverse the order of integration and differentiation, concluding that for all (m_1, m_2)

$$\frac{\partial}{\partial\rho}\Phi_\rho(s,t) = \phi_\rho(s,t).$$

Because $\phi_\rho(s,t) > 0$ everywhere, Φ_ρ is strictly increasing in ρ. Now since $\rho^X > \rho^Y$, we necessarily must have that $\bar{F}(m_1,m_2) > \bar{G}(m_1,m_2)$ for every $(m_1,m_2) \in \mathbb{R}^2$.

By Theorem 4.1, $\boldsymbol{X}$ does not stochastically dominate $\boldsymbol{Y}$. Indeed, let us consider the following upper set $M = \{(x_1,x_2) : x_1,x_2 \in \mathbb{R}, x_1 \geq -1-x_2\}$. Then by direct calculation using the software Mathematica we get:

$$\mathbb{P}(\boldsymbol{X}\in M) = \int_{-\infty}^{\infty}\int_{-1-x_1}^{\infty}\frac{1}{2\pi}\,\exp\left\{-\frac{x_1^2+x_2^2}{2}\right\}\mathrm{d}x_2\mathrm{d}x_1 = 0.7181,$$
$$\mathbb{P}(\boldsymbol{Y}\in M) = \int_{-\infty}^{\infty}\int_{-1-x_1}^{\infty}\frac{1}{\sqrt{3\pi^2}}\,\exp\left\{-\frac{2(x_1^2-x_1x_2+x_2^2)}{3}\right\}\mathrm{d}x_2\mathrm{d}x_1 = 0.7603.$$

We found an upper set for which $\mathbb{P}(\boldsymbol{X}\in M) < \mathbb{P}(\boldsymbol{Y}\in M)$. An easy plot can be helpful when searching for such upper set for which $\mathbb{P}(\boldsymbol{X}\in M) < \mathbb{P}(\boldsymbol{Y}\in M)$, see Figure 2. In the first step one simulates two samples of the same size from both distributions. In the second step one needs to find an upper set which contains more realizations of the random vector $\boldsymbol{Y}$ compared to the number of realization of the random vector $\boldsymbol{X}$. Note that whereas the realizations of the random vector $\boldsymbol{Y}$ are evenly dispersed around its mean, the realizations of $\boldsymbol{X}$ are concentrated around the axis of the first and the third quadrant. In this case, the above stated choice for the set M is straightforward.

5. SIMULATION

In this part we carry out a simulation in which we consider two random vectors with 2-dimensional normal distributions $\boldsymbol{X}$ and $\boldsymbol{Y}$, i.e. $\boldsymbol{X} \sim N_2(\boldsymbol{\mu}^X, \Sigma^X)$ and $\boldsymbol{Y} \sim N_2(\boldsymbol{\mu}^Y, \Sigma^Y)$. The parameters are set in the spirit of Theorem 4.1 so as the random vector $\boldsymbol{X}$ stochastically dominates the random vector $\boldsymbol{Y}$, i.e. $\boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y$ and $\Sigma^X = \Sigma^Y = \Sigma$. We simulate m realizations from both distributions and based on these two samples we investigate whether the random vector $\boldsymbol{X}$ stochastically dominates the random vector $\boldsymbol{Y}$. We use two different methods for detection of stochastic dominance. The first one uses the optimization technique, namely we employ the optimization problem (SD) introduced in Section 2. The second one is a statistical approach which is based on testing hypotheses about the relation of parameters of both distributions. We aim to compare accuracy of these two methods and their behavior depending on the choice of parameters of the

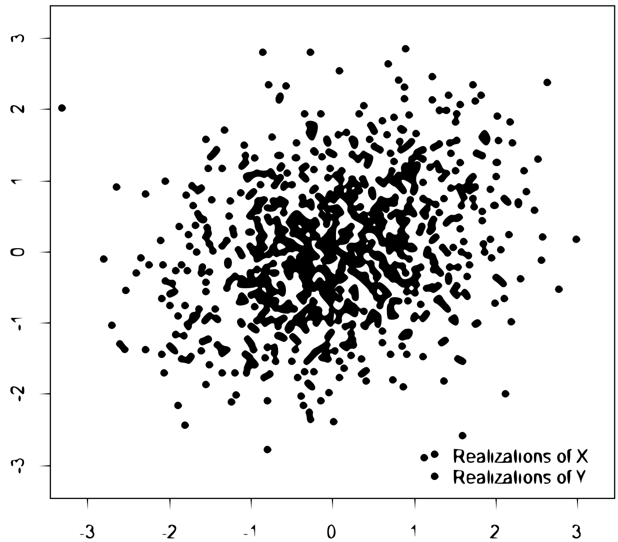

Fig. 2. Simulation of 500 realizations of 2-dimensional random vectors $\boldsymbol{X} \sim \Phi_{1/2}$ and $\boldsymbol{Y} \sim \Phi_0$.

distributions. For this reason we repeat the simulation for each choice of the parameters n times.

Whereas the optimization approach has been already explained in the previous part, the statistical approach deserves a detailed description. In this procedure, we firstly need to estimate parameters of both distributions. Let us denote by $\hat{\boldsymbol{\mu}}^X = (\hat{\mu}_1^X, \hat{\mu}_1^X)$ the maximum likelihood estimate of $\boldsymbol{\mu}^X = (\mu_1^X, \mu_2^X)$ and let $\hat{\Sigma}^X$ be the maximum likelihood estimate of Σ^X. Similarly, we denote by $\hat{\boldsymbol{\mu}}^Y = (\hat{\mu}_1^Y, \hat{\mu}_1^Y)$ the maximum likelihood estimate of $\boldsymbol{\mu}^Y = (\mu_1^Y, \mu_2^Y)$ and by $\hat{\Sigma}^Y$ be the maximum likelihood estimate of Σ^Y. In the simulation we aim to test the following null hypothesis H_0 against its alternative H_1:

$$H_0 : \boldsymbol{\mu}^X \geq \boldsymbol{\mu}^Y \text{ and } \Sigma^X = \Sigma^Y \qquad H_1 : \boldsymbol{\mu}^X < \boldsymbol{\mu}^Y \text{ or } \Sigma^X \neq \Sigma^Y.$$

Then by Theorem 4.1 the hypothesis that the random vector $\boldsymbol{X}$ stochastically dominates the random vector $\boldsymbol{Y}$ cannot be rejected if and only if the null hypothesis H_0 is not rejected. The null hypothesis can be equivalently written as a composite hypothesis $H_0 = \bigcap_{i=1}^3 H_0^i$, where

$$\begin{aligned} &H_0^1 : \mu_1^X \geq \mu_1^Y && H_1^1 : \mu_1^X < \mu_1^Y \\ &H_0^2 : \mu_2^X \geq \mu_2^Y && H_1^2 : \mu_2^X < \mu_2^Y \\ &H_0^3 : \Sigma^X = \Sigma^Y && H_1^3 : \Sigma^X \neq \Sigma^Y. \end{aligned}$$

In order to test the composite hypothesis, we employ Holm–Bonferroni method. It is a stepwise algorithm which is used to counteract the problem of multiple comparisons while controlling the family-wise error rate. It proceeds in the following steps.

1. Let us consider a family of hypotheses $H_0^1, \dots, H_0^I$. Calculate the corresponding p-values $p_1, \dots, p_I$.

2. Reorder the p-values into an increasing sequence $p_{(1)}, \dots, p_{(I)}$, i. e. $p_{(i)} \leq p_{(i+1)}$, $i = 1, \dots, I-1$. Let $H_0^{(1)}, \dots, H_0^{(I)}$ be the associated hypotheses.

3. For a given significance level α, let k be the minimal index such that $p_{(k)} > \frac{\alpha}{m+1-k}$.

4. Reject the null hypotheses $H_0^{(1)}, \dots, H_0^{(k-1)}$ and do not reject the null hypotheses $H_0^{(k)}, \dots, H_0^{(I)}$.

Note that in Holm–Bonferroni procedure, we first test $H_0^{(1)}$. If it is not rejected then the intersection of all null hypotheses $\bigcap_{i=1}^m H_0^{(i)}$ is not rejected either. For more details about this method we refer to [5].

For testing the hypotheses of inequality between elements of the means of the considered distributions $H_0^1 : \mu_1^X \geq \mu_1^Y$ and $H_0^2 : \mu_2^X \geq \mu_2^Y$, we employ the standard two sample one-sided t-test. For detailed description see for instance [2]. For testing the hypothesis of the equality of two covariance matrices $H_0^3 : \Sigma^X = \Sigma^Y$, we employ a version of Box's test, namely the likelihood-ratio test. This is the multivariate generalization of Bartletts test of homogeneity of variances. The test has the form:

$$-2\log\lambda = 2m\log|\hat{\Sigma}| - m\left(\log|\hat{\Sigma}^X| + \log|\hat{\Sigma}^Y|\right) = m\left(\log|\hat{\Sigma}^X_{-1}\hat{\Sigma}| + \log|\hat{\Sigma}^Y_{-1}\hat{\Sigma}|\right),$$

where $\hat{\Sigma}^X$ and $\hat{\Sigma}^Y$ are sample biased covariance matrices of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, and $\hat{\Sigma} = (\hat{\Sigma}^X + \hat{\Sigma}^y)/2$ is the maximum likelihood estimate of the common covariance matrix (under the null hypothesis). We emphasize that for derivation of the above stated test statistics we used the fact that both random vectors have the same number of realizations equal to m. The degrees of freedom of the asymptotic χ^2 distribution are $(d+1)/2$. For more details about the test statistics and its properties we refer to [1].

In the computational experiment we kept the means of both distribution fixed, namely $\boldsymbol{\mu}^X = (1,2)$ and $\boldsymbol{\mu}^Y = (0,1)$. The elements defining the common covariance matrix Σ were taken from the following sets:

$$\sigma_{11}, \sigma_{22} \in \{0.5, 1, 2, 3, 5, 10\}, \qquad \rho \in \{-0.8, -0.4, 0, 0.4, 0.8\},$$

where ρ is the correlation coefficient between elements of the considered random vectors, i. e. $\mathrm{cor}(X_1, X_2) = \mathrm{cor}(Y_1, Y_2) = \rho$ and thus $\sigma_{12} = \sigma_{21} = \rho\sqrt{\sigma_{11}\sigma_{22}}$. This means that we considered 180 different settings of parameters specifying two normal distributions. For each setting of parameters σ_{11}, σ_{22} and ρ we carried out 100 simulations with 100 realizations ($n = 100$, $m = 100$). In the second part of the experiment we fixed a choice of the distributions parameters and varied the size of the samples, we carried out the simulation for $m = \{50, 100, 200, 500, 1000, 2000\}$. The entire statistical approach was evaluated on the confidence level $\alpha = 0.95$.

Tables 1 and 2 show results of the simulations for different choices of the covariance matrices. Both tables display efficiency of detecting stochastic dominance between two random vectors for the case when stochastic dominance between these vectors is present,

		corr. parameter ρ				
σ_{11}	σ_{22}	-0.8	-0.4	0.0	0.4	0.8
0.5	0.5	0.99	1.00	0.98	1.00	1.00
0.5	1	0.93	0.88	0.88	0.88	0.83
0.5	2	0.58	0.40	0.47	0.49	0.50
0.5	3	0.20	0.20	0.22	0.29	0.21
0.5	5	0.11	0.04	0.10	0.12	0.07
0.5	10	0.01	0.02	0.02	0.05	0.02
1	0.5	0.85	0.86	0.90	0.84	0.88
1	1	0.80	0.77	0.65	0.73	0.75
1	2	0.32	0.19	0.19	0.25	0.44
1	3	0.14	0.12	0.12	0.14	0.21
1	5	0.03	0.03	0.06	0.06	0.08
1	10	0.02	0.02	0.02	0.02	0.02
2	0.5	0.49	0.39	0.49	0.46	0.51
2	1	0.41	0.23	0.22	0.26	0.40
2	2	0.09	0.05	0.01	0.09	0.14
2	3	0.02	0.00	0.03	0.01	0.10
2	5	0.00	0.01	0.00	0.00	0.04
2	10	0.00	0.00	0.01	0.00	0.01
3	0.5	0.24	0.18	0.23	0.23	0.29
3	1	0.21	0.13	0.10	0.15	0.18
3	2	0.01	0.00	0.01	0.03	0.10
3	3	0.00	0.00	0.02	0.03	0.07
3	5	0.00	0.00	0.00	0.01	0.03
3	10	0.00	0.00	0.00	0.00	0.00
5	0.5	0.07	0.05	0.09	0.11	0.08
5	1	0.07	0.03	0.04	0.06	0.02
5	2	0.00	0.00	0.00	0.00	0.03
5	3	0.00	0.00	0.00	0.00	0.04
5	5	0.00	0.00	0.00	0.00	0.01
5	10	0.00	0.00	0.00	0.00	0.00
10	0.5	0.05	0.01	0.02	0.03	0.03
10	1	0.03	0.03	0.01	0.05	0.02
10	2	0.00	0.00	0.00	0.00	0.03
10	3	0.00	0.00	0.00	0.00	0.00
10	5	0.00	0.00	0.00	0.00	0.00
10	10	0.00	0.00	0.00	0.00	0.00

Tab. 1. Simulation results for optimization approach (100 simulations with 100 realizations for each combination of parameters σ^X, σ^Y and ρ).

		corr. parameter ρ				
σ_{11}	σ_{22}	-0.8	-0.4	0.0	0.4	0.8
0.5	0.5	0.95	0.93	0.91	0.93	0.92
0.5	1	0.96	0.87	0.92	0.92	0.95
0.5	2	0.96	0.87	0.94	0.93	0.92
0.5	3	0.96	0.93	0.88	0.96	0.86
0.5	5	0.91	0.87	0.90	0.92	0.96
0.5	10	0.93	0.96	0.90	0.93	0.89
1	0.5	0.92	0.93	0.93	0.92	0.92
1	1	0.92	0.90	0.93	0.94	0.90
1	2	0.93	0.92	0.94	0.91	0.93
1	3	0.94	0.94	0.95	0.92	0.96
1	5	0.93	0.95	0.91	0.91	0.94
1	10	0.92	0.88	0.95	0.94	0.89
2	0.5	0.93	0.95	0.90	0.92	0.92
2	1	0.91	0.89	0.95	0.94	0.89
2	2	0.91	0.96	0.92	0.94	0.92
2	3	0.95	0.95	0.86	0.97	0.94
2	5	0.93	0.88	0.95	0.87	0.93
2	10	0.97	0.89	0.87	0.89	0.92
3	0.5	0.91	0.94	0.95	0.94	0.96
3	1	0.92	0.94	0.95	0.84	0.97
3	2	0.94	0.92	0.87	0.93	0.91
3	3	0.88	0.94	0.90	0.86	0.95
3	5	0.93	0.90	0.91	0.93	0.95
3	10	0.93	0.92	0.91	0.95	0.89
5	0.5	0.90	0.92	0.98	0.93	0.92
5	1	0.87	0.91	0.92	0.94	0.94
5	2	0.89	0.96	0.95	0.94	0.95
5	3	0.92	0.88	0.88	0.89	0.90
5	5	0.86	0.93	0.93	0.94	0.95
5	10	0.93	0.94	0.95	0.88	0.93
10	0.5	0.88	0.94	0.91	0.91	0.92
10	1	0.90	0.96	0.95	0.93	0.92
10	2	0.94	0.89	0.95	0.92	0.92
10	3	0.93	0.93	0.92	0.94	0.95
10	5	0.92	0.93	0.95	0.96	0.95
10	10	0.92	0.92	0.97	0.90	0.96

Tab. 2. Simulation results for statistical approach (100 simulations with 100 realizations for each combination of parameters σ^X, σ^Y and ρ).

	50	100	200	500	1000	2000
Optimization method	0.58	0.64	0.70	0.80	0.81	0.85
Statistical method	0.92	0.95	0.90	0.96	0.96	0.93

Tab. 3. Simulation results for various choice of sample size m (100 simulations with $\sigma_{11} = \sigma_{22} = 1$ and $\rho = 0.4$).

i. e. the tables provide ratio between the number of simulations for which we obtained positive test of stochastic dominance and the total number of simulations. Table 1 shows results for the optimization method, whereas Table 2 provides results for the statistical approach. We observe that efficiency of the statistical method is stable when varying the covariance matrix and the probability that the approach is able to detect stochastic dominance oscillate around 0.9. On the other hand, the optimization method is highly dependent of the choice of the parameters σ_{11}, σ_{22}. When these parameters have considerably low value, in comparison with the distance between the means of the random vectors, the probability of detecting stochastic dominance is nearly 1. This probability converges sharply to 0 as these parameters become higher. We do not observe any dependence of the efficiency in the correlation parameters. We can conclude that for small variances it is more confident to use the optimization method whereas for higher values the statistical method should be preferred. In practice, it is hard to find the right breakpoint for preferring one method against the other and one should always take into consideration the distance between the estimated means of the vectors and also the size of both samples. With increasing cardinality of the samples there is more flexibility in assigning the realizations of one vector to the other one which is executed by the optimization problem (SD) and therefore the optimization method becomes more efficient as shown in Table 3.

6. CONCLUSION

This paper focused on multivariate stochastic dominance for the case of random vectors with multivariate normal distribution. In the first part we provided a brief introduction to multivariate stochastic dominance in which we included definitions of three important types of multivariate stochastic dominance, namely strong, orthant and linear. In the main part of the work we formulated stochastic dominance rule and provided a detailed proof. We explained why one cannot use the same technique based on joint survival functions as it is usual when proving the rule in the univariate case. In the last part of this paper we executed a simulation in which we employed derived stochastic dominance rule in order to validate the efficiency of detecting stochastic dominance. We compared the method with another one based on an optimization problem. We examined behavior of both approaches depending on the setting of the covariance matrix and size of the simulated samples.

ACKNOWLEDGEMENT

I would like to express the gratitude to my supervisor doc. RNDr. Ing. Miloš Kopa, Ph.D. for his valuable comments to this article. The work was supported by Czech Science Foundation (grant 402/12/G097).

(Received November 28, 2017)

REFERENCES

[1] J. Aitchison: The Statistical Analysis of Compositional Data. The Blackburn Press 2003.

[2] J. Anděl: Základy matematické statistiky. Matfyzpress 2005.

[3] B. Armbruster and J. Luedtke: Models and formulations for multivariate dominance-constrained stochastic programs. IIE Transactions *47* (2015), 1–14. DOI:10.1080/0740817x.2014.889336

[4] D. Dentcheva and A. Ruszczýnski: Optimization with multivariate stochastic dominance constraints. Math. Programming *117* (2009), 111–127. DOI:10.1007/s10107-007-0165-x

[5] S. Holm: A simple sequentially rejective multiple test procedure. Scand. J. Statist. *6* (1979), 65–70.

[6] C. C. Huang, I. Vertinsky, and W. T. Ziemba: On multiperiod stochastic dominance. J. Financ. Quantitat. Anal. *13* (1978), 1–13. DOI:10.2307/2330516

[7] M. Kopa and B. Petrová: Strong and weak multivariate first-order stochastic dominance. SSRN Electron. J. *xx* (2017), xxx–xxx. DOI:10.2139/ssrn.3144058

[8] D. Levhari, J. Paroush, and B. Peleg: Efficiency analysis for multivariate distributions. Rev. Econom. Stud. *42* (1975), 87–91. DOI:10.2307/2296822

[9] H. Levy: Stochastic dominance, efficiency criteria and efficient portfolios: the multi-period case. Amer. Econom. Rev. *63* (1973), 986–994. DOI:10.1142/9789814417358_0018

[10] H. Levy: Stochastic Dominance: Investment Decision Making under Uncertainty. Springer 2006.

[11] H. Levy: Stochastic Dominance (Investment Decision Making Under Uncertainty). Springer 2016.

[12] H. Levy and J. Paroush: Multi-period stochastic dominance. J. Financ. Quantitat. Anal. *21* (1974), 428–435. DOI:10.1287/mnsc.21.4.428

[13] A. Műller and D. Stoyan: Comparison Methods for Stochastic Models and Risks. Wiley 2002.

[14] S. Ortobelli: The classification of parametric choices under uncertainty: analysis of the portfolio choice problem. Theory and Decision *51* (2001), 297–327. DOI:10.1023/a:1015511211848

[15] S. Ortobelli, T. Lando, F. Petronio, and T. Tichý: Asymptotic stochastic dominance rules for sums of iid random variables. J. Comput. Appl. Math. *300* (2016), 432–448. DOI:10.1016/j.cam.2015.12.017

[16] S. Ortobelli, T. Lando, F. Petronio, and T. Tichý: Asymptotic multivariate dominance: A financial application. Methodol. Comput. Appl. Probab. *18* (2016), 1097–1115. DOI:10.1007/s11009-016-9502-y

[17] M. Shaked and J. G. Shanthikumar: Stochastic Orders and their Applications. Academic Press 1993.

[18] Y. L. Tong: The Multivariate Normal Distribution. Springer Series in Statistics 1990. DOI:10.1007/978-1-4613-9655-0

Barbora Petrová, Charles University, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Praha 8. Czech Republic.

e-mail: petrova@karlin.mff.cuni.cz