ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 8, 2004

Preservation of stochastic order in $\mathbb{R}^2$ under random rotations

RICHARD PINCUS

ABSTRACT. For a specified order relation in $\mathbb{R}^2$, it is shown that a shift parameter family of distributions remains stochastically ordered, if it is transformed by random orthogonal rotations. The result has applications in monotonicity problems of the roots of Wishart matrices.

1. Introduction

Let $\mathbf{x}_1, ..., \mathbf{x}_k$ denote independent normally distributed p-dimensional variables

$$\mathbf{x}_i \sim \mathsf{N}_p(\mathbf{0}, \mathbf{I}), \quad i = 1, ..., k.$$

The matrix

$$\mathbf{W} = \mathbf{x}_1\mathbf{x}_1' + \cdots + \mathbf{x}_k\mathbf{x}_k' \tag{1}$$

has a Wishart distribution with k degrees of freedom, shortly written $\mathbf{W} \sim \mathsf{W}_p(k, \mathbf{I})$. If $\mathbf{x}_1, ..., \mathbf{x}_k$ are not all centered at zero, but have mean vectors $\mathbf{m}_i$ instead, $i = 1, ..., k$, the random $p \times p$-matrix $\mathbf{W}$, the so-called "Sums of Squares of the Hypothesis", is said to have a non-central Wishart distribution $\mathsf{W}_p(k, \mathbf{I}; \mathbf{MM}')$. Its non-centrality parameter $\mathbf{MM}'$ is

$$\mathbf{MM}' = \mathbf{m}_1\mathbf{m}_1' + \cdots + \mathbf{m}_k\mathbf{m}_k'.$$

In Multivariate Analysis of Variance (MANOVA) the characteristic roots, or eigenvalues, of $\mathbf{W}$ play an important role. These roots remain unchanged if an orthogonal matrix $\mathbf{C}$ transforms the observations $\mathbf{x}_i$ and $\mathbf{W}$ to $\mathbf{Cx}_i$ and $\mathbf{CWC}'$, respectively, so one can assume without loss of generality that $\mathbf{MM}'$ has only non-negative diagonal entries. Consequently, we might assume the mean vectors $\mathbf{m}_i$ of $\mathbf{x}_i$ to have the form

$$\mathbf{m}_i' = (0, ..., \lambda_i^{1/2}, ...0), \quad i = 1, ..., p.$$

Received October 5, 2003.
2000 *Mathematics Subject Classification.* Primary 60E15; Secondary 62H15, 15A52.
Key words and phrases. Stochastic order, Wishart matrix, MANOVA.

Since many years the statistical question, if the roots of $\mathbf{W}$ are stochastically increasing under increasing roots λ_i of the non-centrality, remains open.

The term "stochastically increasing" means that for all increasing functions f of the p roots of $\mathbf{W}$ the expectation $\mathsf{E}f$ is increasing in the parameters $\lambda_1, ..., \lambda_p$, see Lehmann (1955). If f is used as a test statistic with rejection region $f \geq c_\alpha$, then stochastic increasing roots of $\mathbf{W}$ automatically imply monotonicity of the power of the respective test procedures.

In the simplest case $p = k = 2$ under fixed norm $\|\mathbf{x}_2\|$ one finds an orthogonal rotation $\mathbf{C}$, depending on $\mathbf{x}_2$, of course,which transforms $\mathbf{x}_2$ to a multiple of $\mathbf{e}_2$ (the 2-nd unit vector). Applying this rotation to $\mathbf{x}_1$ too, this would not change the eigenvalues of $\mathbf{W}$, but $\mathbf{W}$ is now represented by a sum

$$\mathbf{W} = \mathbf{y}_1\mathbf{y}_1' + \mathbf{e}_2\mathbf{e}_2'.$$

With $\mathbf{e}_2$ fixed, it remains to show that the random vector $\mathbf{y}_1 = \mathbf{C}_{\theta(\mathbf{x}_2)}\mathbf{x}_1$ is stochastically increasing in λ_1 with respect to a specific order relation. A simple analysis shows that the random rotation angle θ is distributed symmetrically and unimodal. Consequently its distribution can be represented as a mixture of uniform distributions. In Pincus (2000) a proof for stochastic monotonicity of the roots of a two-dimensional Wishart matrix was given that way. Since the two-dimensional normal distribution of $\mathbf{x}_1$ is a mixture of uniform distributions on lines, there is a motive to look at random rotations $\mathbf{C}_\theta\mathbf{u}$, where $\mathbf{u}$ stands for a one-dimensional uniformly distributed variable in $\mathbb{R}^2$ with a shift parameter t, see Section 3. By showing the stochastic monotonicity of this uniformly rotated family in the shift t in Section 4 we have a tool to generalize the stochastic monotonicity of the roots of (1) to certain non normal situations, and perhaps a tool for handling the case of general dimension p and number k.

Following Perlman and Olkin (1980) stochastic monotonicity of the roots of $\mathbf{W}$ carries over to stochastic monotonicity of the roots of the MANOVA-matrix $\mathbf{S}^{-1/2}\mathbf{W}\mathbf{S}^{-1/2}$, with $\mathbf{S}$ being the "Sums of Squares of Errors" matrix.

2. A special order in $\mathbb{R}^2_+$

In Euclidean spaces it is common to define an order relation between vectors componentwise, i.e. $\mathbf{x} \leq \mathbf{y}$ iff $x_i \leq y_i$ for all i. In this paper we use an alternative order relation $\leq_+$ in $\mathbb{R}^2_+$ which is defined as follows:

$$\mathbf{x} \leq_+ \mathbf{y} \qquad \text{iff} \quad x_2 \leq y_2 \text{ and } \frac{x_1}{x_2} \leq \frac{y_1}{y_2}. \tag{2}$$

In Figures 1 and 2 there are sketched the points $\mathbf{x}$ being larger or smaller than a given $\mathbf{x}_0$ in the context of the respective partial order.

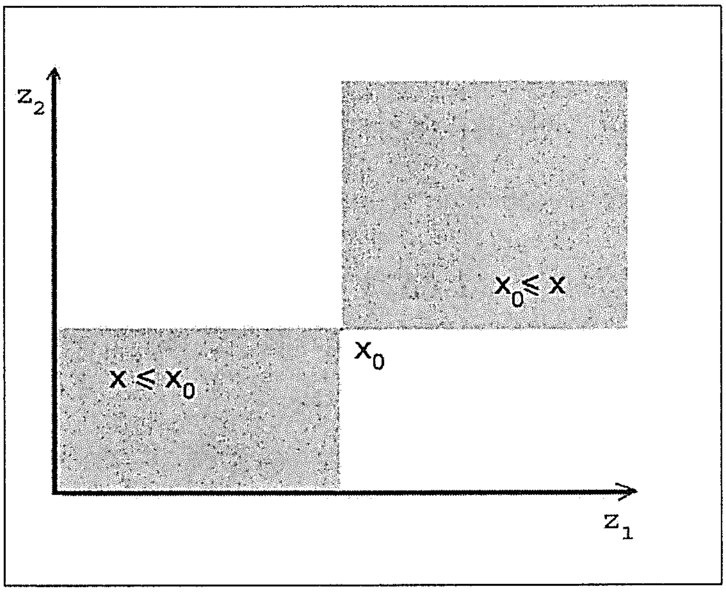

FIGURE 1. The order relation $\leq$ in $\mathbb{R}^2_+$

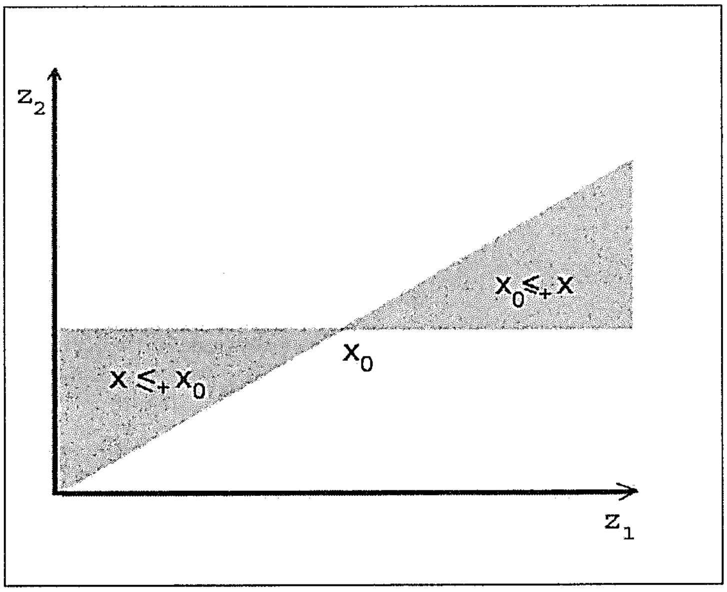

FIGURE 2. The order relation $\leq_+$in $\mathbb{R}^2_+$

Two equivalent definitions of the $\leq_+$ relation are:

$\mathbf{x} \leq_+ \mathbf{y}$ iff $\mathbf{y} = \beta\mathbf{x} + \delta\mathbf{e}_1$, and $0 \leq \delta, 1 \leq \beta$ ($\mathbf{e}_1$ is the first unit vector),

$\mathbf{x} \leq_+ \mathbf{y}$ iff $y_1 = \gamma x_1$ and $y_2 = \beta x_2$, and $1 \leq \beta \leq \gamma$.

The $\leq_+$ order easily extends in a symmetric way to $\mathbb{R}^2$ if we identify vectors with the same absolute components. Definition (2) simply changes to

$$\mathbf{x} \leq_+ \mathbf{y} \qquad \text{iff} \quad |x_2| \leq |y_2| \text{ and } \left|\frac{x_1}{x_2}\right| \leq \left|\frac{y_1}{y_2}\right|. \tag{3}$$

Two probability distributions Q and P are stochastically $\leq_+$-ordered, short $\mathsf{Q} \leq_+ \mathsf{P}$, iff

$$\int f \, \mathrm{d}\mathsf{Q} \leq \int f \, \mathrm{d}\mathsf{P}$$

for all measurable and bounded nondecreasing (in the $\leq_+$-sense) functions f, or equivalently

$$\mathsf{Q}(\mathcal{F}) \leq \mathsf{P}(\mathcal{F})$$

for all measurable and bounded nondecreasing (in the $\leq_+$-sense) sets $\mathcal{F}$, see Lehmann (1955). Obviously, if $X(\mathbf{z})$ is a family of random variables, fulfilling $\mathbf{z} \leq_+ X(\mathbf{z})$ a.s., and $\mathbf{z}$ is Q-distributed, then denoting the total distribution of X by P^X, we have $\mathsf{Q} \leq_+ \mathsf{P}^X$, see Strassen (1965). The family of distributions of $X(\mathbf{z})$ is called a stochastic kernel.

3. Random rotation of shifted uniform distributions

In this section we consider a family of one-dimensional uniform distributions V_t in $\mathbb{R}^2$ with (disjoint) supports $\{(t, z_2)', -a \leq z_2 \leq a\}$. Clearly $\mathsf{V}_t \leq_+ \mathsf{V}_{t'}$, if $0 < t < t'$. Our aim is to study the total distribution, Q_t say, of a random V_t-distributed vector $\mathbf{z}$, being randomly rotated by an orthogonal matrix $\mathbf{C}_\theta$, where

$$\mathbf{C}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

and the rotation angle θ is uniformly distributed between $-\alpha$ and α, $0 \leq \alpha \leq \frac{\pi}{2}$.

In Figure 3 there is sketched the basic position of the supports of two distributions V_t and $\mathsf{V}_{t'}$, $0 < t < t'$, before the rotations between $\pm\alpha$ are applied.

Only for illustration in Figure 4 a typical form of the support of the randomly rotated V_t is given. Note, that depending on the size of t, α and a, the point A, which is $(t, a)'$ rotated by α, may lie in the positive orthant. Also the second component of F, the rotation of $(t, a)'$ by $-\alpha$, might be negative. The proofs in the next section are, however, easy to adapt to these situations.

The distribution of the randomly rotated V_t, i.e. Q_t, has the density

$$q_t(u_1, u_2) = \frac{1}{4a\alpha(r^2 - t^2)^{1/2}}, \qquad \text{where} \quad r^2 = u_1^2 + u_2^2$$

on its support, multiplied by the factor 2 in the 'dark' region.

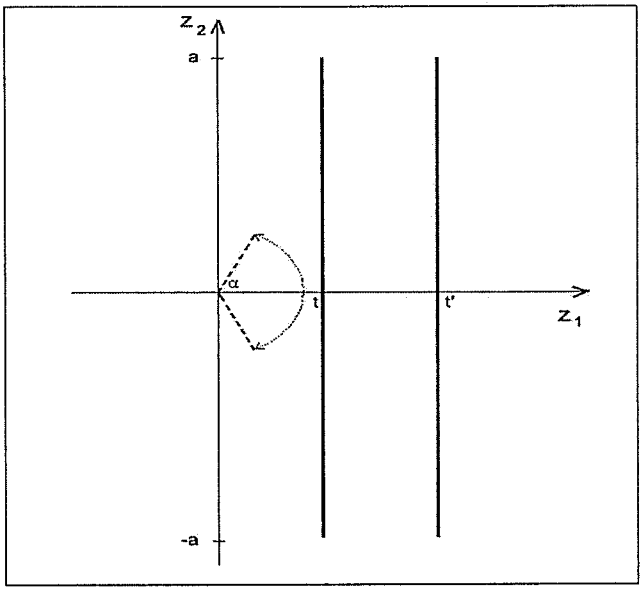

FIGURE 3. Supports of V_t and $V_{t'}$

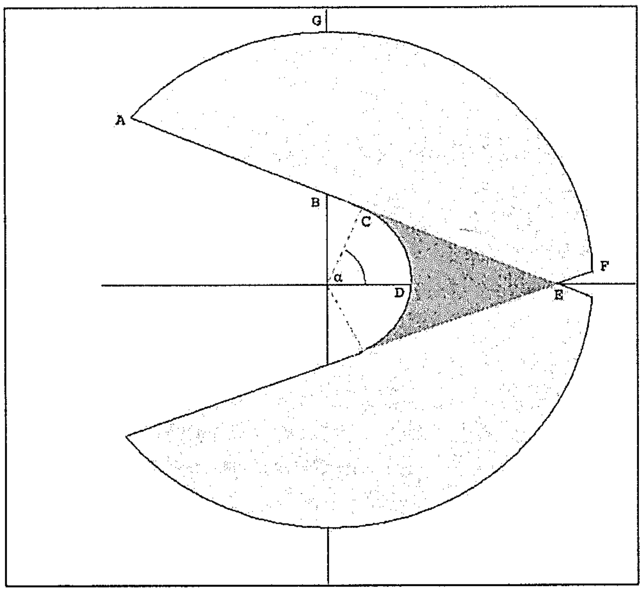

FIGURE 4. Support of a rotated V_t

Ultimately wishing to prove that this total distribution Q_t is stochastically increasing in t with respect to the $\leq_+$ order, we start with a monotonicity property of the marginals. $Q_t^{(1)}$ and $Q_t^{(2)}$ denote the distributions of $\left|\frac{u_1}{u_2}\right|$ and $|u_2|$, respectively, if Q_t is the distribution of the random vector (u_1, u_2).

Lemma 1. *If* $0 \le t < t'$, *then for the distributions of the marginals one has*

$$Q_t^{(1)} \le Q_{t'}^{(1)} \quad \textit{and} \quad Q_t^{(2)} \le Q_{t'}^{(2)}.$$

Proof. The probability of any randomly (between $-\alpha$ and α) rotated point $\mathbf{z}$ to lie in the double cone $\left\{(u_1, u_2) : \left|\frac{u_1}{u_2}\right| \ge c\right\}$ is an increasing function of $\left|\frac{z_1}{z_2}\right|$. Denoting this probability by $h_t(z)$ if $\mathbf{z}$ is a point (t, z) on the support of V_t, or by $h_{t'}(z)$ if $\mathbf{z}$ is a point (t', z) on the support of $V_{t'}$, we have $h_t(z) \le h_{t'}(z)$ and

$$Q_t(\left|\tfrac{u_1}{u_2}\right| \ge c) = \frac{1}{2a}\int_{-a}^{a} h_t(z)\ \mathrm{d}z \le \frac{1}{2a}\int_{-a}^{a} h_{t'}(z)\ \mathrm{d}z = Q_{t'}(\left|\tfrac{u_1}{u_2}\right| \ge c)\ .$$

In a similar way we prove $Q_t(|u_2| \ge c) \le Q_{t'}(|u_2| \ge c)$. Here we note that for a fixed rotation θ the probability, say $k_t(\theta)$, of the rotated support of V_t to lie outside the stripe $\{(u_1, u_2) : |u_2| \le c\}$ is increasing in t. Integration on θ gives

$$Q_t(|u_2| \ge c) = \frac{1}{2\alpha}\int_{-\alpha}^{\alpha} k_t(\theta)\ \mathrm{d}\theta \le \frac{1}{2\alpha}\int_{-\alpha}^{\alpha} k_{t'}(\theta)\ \mathrm{d}\theta = Q_{t'}(|u_2| \ge c).$$

□

Yet another monotonicity result will be useful in section 4. Let $\mathcal{A}_{\mathbf{x}}$ denote the set of all $\mathbf{y} \in \mathbb{R}^2$ with $\mathbf{x} \le_+ \mathbf{y}$.

Lemma 2. *If* $0 \le t < t'$, *then*

$$Q_t(\mathcal{A}_{\mathbf{x}}) \le Q_{t'}(\mathcal{A}_{\mathbf{x}})$$

holds for all $\mathbf{x}$.

Proof. We used in the proof of Lemma 1 that for a fixed rotation θ the probability $k_t(\theta)$, i.e. the length of the rotated support of V_t outside the stripe $\{(u_1, u_2) : |u_2| \le x_2\}$ was increasing in t. Elementary analytic geometry shows that the relative length, i.e. the conditional probability of that part which additionally is contained in the cone $\left|\frac{u_1}{u_2}\right| \ge \left|\frac{x_1}{x_2}\right|$, is also increasing in t. Consequently their product, which is the probability of $\mathcal{A}_{\mathbf{x}}$ under fixed θ is increasing in t. Integration on θ gives the desired result. □

4. Stochastic monotonicity of rotated uniform distributions

The stochastic monotonicity of its marginals does not mean that a multivariate distribution family is monotone in its specified sense.

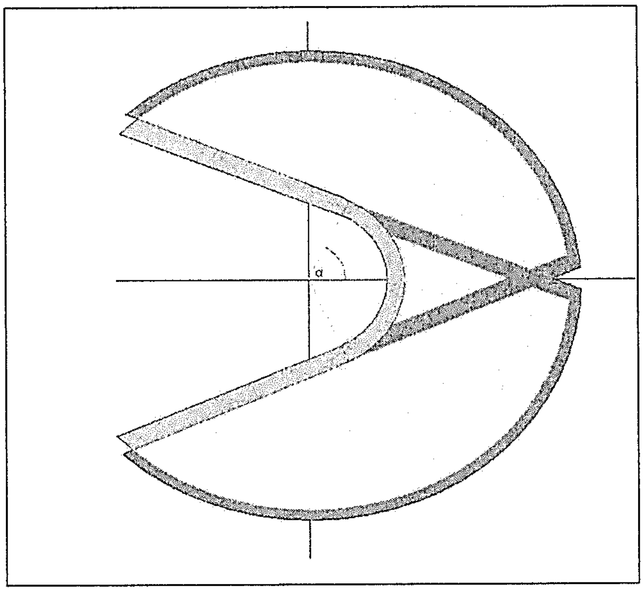

FIGURE 5. Comparison of Q_t and $Q_{t'}$

Since a direct proof for $Q_t \leq_+ Q_{t'}$ is not available, we will proceed with an infinitesimal argument. If t and t' are close enough, their supports intersect, as indicated in Figure 5.

Taking the difference $t' - t = \Delta$ small, in Figure 4 in the light area, the difference $q_{t'}(u_1, u_2) - q_t(u_1, u_2)$ is positive of order Δ, in the medium gray one $q_{t'}(u_1, u_2)$ is zero, while in the dark gray region $q_{t'}(u_1, u_2) - q_t(u_1, u_2)$ equals or approximates $q_{t'}(u_1, u_2)$.

For convenience we map all points to their counterparts (with the same absolute components) in the positive orthant, thus getting a situation sketched in Figure 6 . From now on by Q_t and $Q_{t+\Delta}$ and their densities we mean the induced probabilities on the positive orthant. The 'medium gray' stripe in Figure 5, which approximates the curve A-B-D and has thickness Δ will shortly be denoted as Δ-stripe.

Now we construct a stochastic kernel which transforms Q_t to $Q_{t+\Delta}$ in the following way:

If $q_{t+\Delta}(\mathbf{u}) \geq q_t(\mathbf{u})$, then $X(\mathbf{u}) = \mathbf{u}$ a.s., if $q_{t+\Delta}(\mathbf{u}) < q_t(\mathbf{u})$, which happens only in the Δ-stripe. Then with probability $\varkappa = q_{t+\Delta}(\mathbf{u})/q_t(\mathbf{u})$ we set $X(\mathbf{u}) = \mathbf{u}$, while with probability $1 - \varkappa$ the distribution of $X(\mathbf{u})$ is 'proportional' to $Q_{t+\Delta} - Q_t$ in an appropriately chosen region $\mathcal{R}_{\mathbf{u}}$. It remains to check if this can be done so that $\mathbf{u} \leq_+ X(\mathbf{u})$.

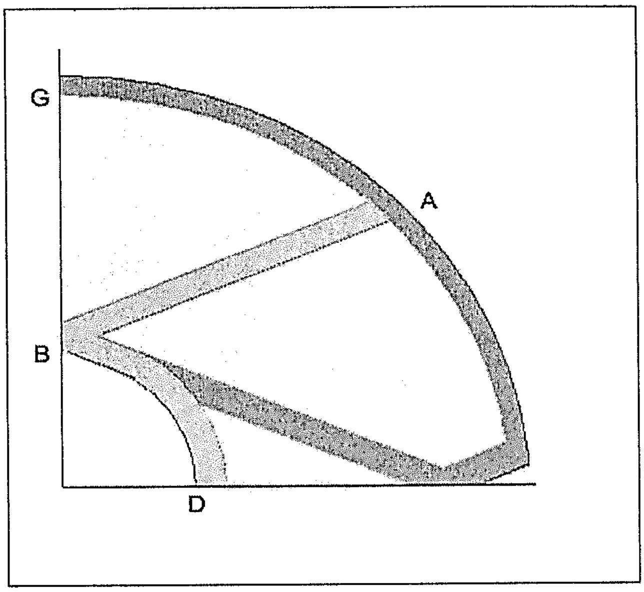

FIGURE 6. Difference between Q_t and Q_{t+dt}

(i) First we represent A-B-D by a parametrized curve $\mathbf{x}(s)$, where s denotes the second component of $\mathbf{x}$, so that $\mathbf{x}(0) = D$, $\mathbf{x}(b_2) = B$ and $\mathbf{x}(a_2) = A$. For any s, $b_2 \leq s \leq a_2$, we denote by $\widetilde{s}$ the value, for which $Q_t(\mathcal{A}_{\mathbf{x}(s)}) = Q_{t+\Delta}(\mathcal{A}_{\mathbf{x}(\widetilde{s})})$. By Lemma 2 we have $\widetilde{s} > s$, and therefore $\mathbf{x}(s) \leq_+ \mathbf{x}(\widetilde{s})$ and moreover $\mathbf{x}(s) \leq_+ \mathcal{A}_{\mathbf{x}(\widetilde{s})}$. Finally, note that $\mathbf{x}(\widetilde{s}) - \mathbf{x}(s)$ is of order $O(\Delta)$.

(ii) Now for any s, $0 \leq s < b_2$, we define $\widehat{s}$ to be the value, for which $Q_t(\mathcal{B}_{\mathbf{x}(s)}) = Q_{t+\Delta}(\mathcal{B}_{\mathbf{x}(\widehat{s})})$. Here $\mathcal{B}_{\mathbf{x}}$ denotes the stripe $\{(u_1, u_2) : 0 \leq u_2 \leq x_2\}$. By Lemma 1 we have $\widehat{s} > s$, and therefore $x_2(s) < x_2(\widehat{s})$. Moreover, $\mathbf{x}(\widehat{s}) - \mathbf{x}(s)$ is of order $O(\Delta)$.

(iii) The distribution of $X(\mathbf{u})$ has yet to be defined on the Δ-stripe. For a small fixed Δ and small $\mathrm{d}s = o(\Delta)$ we denote by $\mathcal{U}_s$ for $b_2 \leq s \leq a_2$ the intersection of the Δ-stripe and the convex cone with vertex at zero, spanned by $\mathbf{x}(s)$ and $\mathbf{x}(s+\mathrm{d}s)$. For every $\mathbf{u} \in \mathcal{U}_s$ the variable $X(\mathbf{u})$ equals $\mathbf{u}$ with probability $\varkappa = q_{t+\Delta}(\mathbf{u})/q_t(\mathbf{u})$, while with probability $1-\varkappa$ its distribution is the normalized difference $Q_{t+\Delta} - Q_t$ in the region $\mathcal{R}_s$, the intersection of $\mathcal{A}_{\mathbf{x}(\widetilde{s})} - \mathcal{A}_{\mathbf{x}(\widetilde{s}+d\widetilde{s})}$ and the complement of the Δ-stripe.

Let s^* denote $\widetilde{s}(b_2)$, and s_* be the value for which $\widehat{s}(s_*) = b_2$. For $0 \leq s \leq s_* < b_2$ we denote by $\mathcal{U}_s$ the intersection of the Δ-stripe and the set $\{(u_1, u_2) : s \leq u_2 < s + \mathrm{d}s\}$. For every $\mathbf{u} \in \mathcal{U}_s$ the variable $X(\mathbf{u})$ equals $\mathbf{u}$ with probability $\varkappa = q_{t+\Delta}(\mathbf{u})/q_t(\mathbf{u})$, while with probability $1 - \varkappa$

its distribution is the normalized difference $\mathrm{Q}_{t+\Delta} - \mathrm{Q}_t$ in the region $\mathcal{R}_s = \mathcal{B}_{\mathbf{x}(\widehat{s}+d\widehat{s})} - \mathcal{B}_{\mathbf{x}(\widehat{s})} = \{(u_1, u_2) : \widehat{s} \le u_2 \le \widehat{s} + \mathrm{d}\widehat{s}\}$.

$\mathcal{U}_*$ is the intersection of the Δ-stripe and the complements of $\mathcal{A}_{\mathbf{x}(s^*)}$ and $\mathcal{B}_{\mathbf{x}(s_*)}$. For every $\mathbf{u} \in \mathcal{U}_s$ the variable $X(\mathbf{u})$ equals $\mathbf{u}$ with probability $\varkappa = q_{t+\Delta}(\mathbf{u})/q_t(\mathbf{u})$, while with probability $1-\varkappa$ its distribution is the normalized difference $\mathrm{Q}_{t+\Delta} - \mathrm{Q}_t$ in the region $\mathcal{R}_*$, the intersection of the complements of $\mathcal{A}_{\mathbf{x}(s^*)}$, $\mathcal{B}_{\mathbf{x}(s_*)}$ and of the Δ-stripe.

(iv) Let now $s_0 < s_1 < \cdots < s_n$ be a decomposition of the intervall $[0, a_2]$ so that $s_0 = 0, s_n = a_2$ and $s_k = s_*, s_l = s^*$ for some k, l. Setting $s_{i+1} = s_i{+}\mathrm{d}s$, $i = 1, ..., n$, we get by the constructions in (iii) a stochastic kernel, defined on the whole support of Q_t. Remember that $X(\mathbf{u}) = \mathbf{u}$ a.s. outside the Δ-stripe.

It is evident by construction, that the stochastic kernel, applied on Q_t, induces $\mathrm{Q}_{t+\Delta}$ for fixed Δ. We can not, however, prove that $\mathbf{u} \le_+ X(\mathbf{u})$ a.s., since for fixed Δ there exist points $\mathbf{u} \in \mathcal{U}_s$, and $\mathbf{v} \in \mathcal{R}_s$, such that $\mathbf{u} \le_+ \mathbf{v}$ does not hold. Nevertheless we have

Theorem 1. *The distribution family Q_t of rotated uniform distributions is stochastically monotone increasing in the shift parameter t.*

Proof. (i) Let $\mathcal{F}$ be any nondecreasing set. For any fixed Δ the inequality

$$\mathrm{Q}_{t+\Delta}(\mathcal{F}) \ge \mathrm{Q}_t(\mathcal{F}) - \mathrm{Q}_t(\mathbf{u} \in \mathcal{F} : X(\mathbf{u}) \notin \mathcal{F})$$

holds. The set $\mathcal{Z} = \{\mathbf{u} \in \mathcal{F} : X(\mathbf{u}) \notin \mathcal{F}\}$ must be contained in the Δ-stripe. We consider at first the part $\mathcal{Z}_1$ between A and B, i.e. with second component u_2 between b_2 and a_2. Let f_* denote the infimum of the second component of all points in the intersection of $\mathcal{F}$ and the Δ-stripe. Because of the nondecreasing character of $\mathcal{F}$, the region 'right hand' of the Δ-stripe with second component $u_2 \ge f_*$ belongs to $\mathcal{F}$. By the construction of $X(\mathbf{u})$ and some analytic geometry we see that only points $\mathbf{u}$ in the stripe $\{f_* \le u_2 \le f_* + \gamma\Delta\}$ are candidates for $X(\mathbf{u}) \notin \mathcal{F}$, where γ depends on t and the angle α. This way we can estimate the area of $\mathcal{Z}_1$ by $\gamma\Delta^2$, and the probability $\mathrm{Q}_t(\mathcal{Z}_1)$ as $O(\Delta^2)$.

In a similar way we proceed for $\mathcal{Z}_2$ with the second component $0 \le u_2 \le s_*$, and get $\mathrm{Q}_t(\mathcal{Z}_2) = O(\Delta^2)$.

The set $\mathcal{Z}_3$ is contained in $\mathcal{U}_*$. But $\mathrm{Q}_t(\mathcal{U}_*)$ is seen to be of order $O(\Delta^2)$, so $\mathrm{Q}_t(\mathcal{Z}_3) = O(\Delta^2)$.

(ii) Assume that the family Q_t is not stochastically monotone increasing, then there would exist a nondecreasing set $\mathcal{F}$ and two values $0 \le t' < t''$ such that

$$\mathrm{Q}_{t'}(\mathcal{F}) > \mathrm{Q}_{t''}(\mathcal{F}).$$

Then there exists at least one t satisfying $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Q}_t(\mathcal{F}) < 0$, this however contradicts our result $\mathrm{Q}_{t+\Delta}(\mathcal{F}) \ge \mathrm{Q}_t(\mathcal{F}) - O(\Delta^2)$. □

References

Lehmann, E. L. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26**, 399–419.

Perlman, M. D. and Olkin, I. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8**, 1326–1341.

Pincus, R. (1997). Multidimensional stochastic ordering and orthogonal rotations. In: *Bulletin of the International Statistical Institute. Session 51st. Contributed Papers, Book 2*, Istanbul, 209–210.

Pincus, R. (2000). *Stochastic Order Relations for Multivariate Distribution Families* Habil.-Thesis. University Potsdam, Faculty of Science. (In German)

Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439.

Institute of Mathematics, University Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany

E-mail address: pincus@rz.uni-potsdam.de