June 26, 2001

KENDALL'S TAU FOR ELLIPTICAL DISTRIBUTIONS*

FILIP LINDSKOG, ALEXANDER MCNEIL, AND UWE SCHMOCK

ABSTRACT. By using well known properties of elliptical distributions we show that the relation between Kendall's tau and the linear correlation coefficient for bivariate normal distributions holds more generally (subject to only slight modifications) for the class of elliptical distributions.

1. INTRODUCTION

It is well known, and easily demonstrated, that for the two-dimensional normal distribution with linear correlation coefficient ρ the relation

$$\tau = \frac{2}{\pi} \arcsin \varrho, \tag{1}$$

between Kendall's tau and the linear correlation coefficient holds (cf. [2, p. 290], where the calculations are traced back to publications of T. J. Stieltjes from 1889 and W. F. Sheppard from 1898). However, it does not seem to be at all well known that the elegant relationship (1) also holds (subject to only slight modifications) for all non-degenerate elliptical distributions, and this is the main result (Theorem 2) of this short communication.

The result is not only of theoretical interest; it is also extremely useful for statistical purposes. For example, it can be used to build a robust estimator of linear correlation for bivariate elliptically distributed data.

Many multivariate datasets encountered in practice, such as multivariate financial time series data, are not multivariate normally distributed but may plausibly be modelled by another member of the elliptical family with heavier tailed margins. In this situation it is well known that the standard estimator of correlation, based on normal assumptions and maximum-likelihood theory, is both inefficient and lacks robustness; many alternative covariance and correlation estimators have been proposed including M-estimators, estimators based on multivariate trimming and estimators based on variances of sums and differences of standardized variables (cf. [3] for an overview). Formula (1) provides an appealing bivariate method; we simply estimate Kendall's tau using the standard textbook estimator and invert the relationship to get the Kendall's tau transform estimate of ρ. Simulation studies suggest that this simple method performs better than most of its competitors, see Figure 1 and [5]. Note that, unlike almost all other methods of correlation estimation, the Kendall's tau transform method directly exploits the geometry of elliptical distributions and does not require us to estimate variances and covariances. This is advantageous when interest focusses explicitly on correlations, as it often does

---

1991 *Mathematics Subject Classification.* 60E05 (primary); 62H20 and 62J10 (secondary).

Key words and phrases. Robust estimation; linear correlation; Kendall's tau; elliptical distributions.

*Research supported by Credit Suisse, Swiss Re and UBS through RiskLab, Switzerland.

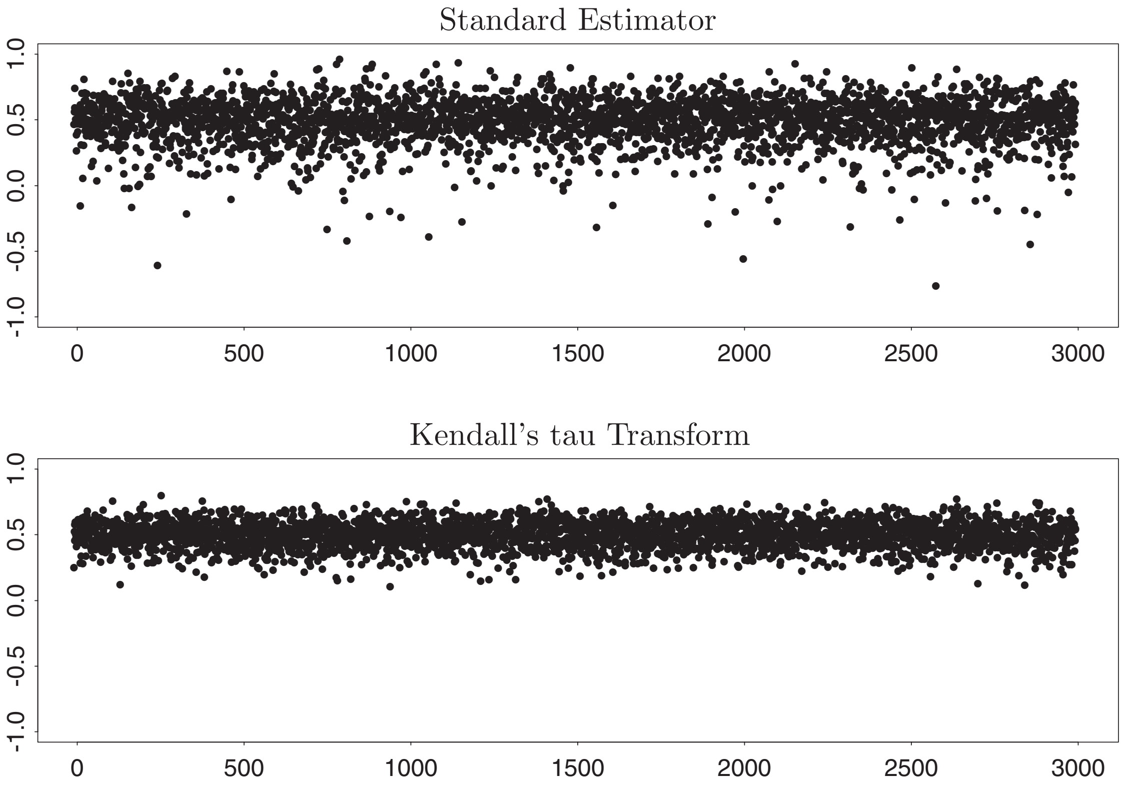

Figure 1. For 3000 independent samples of size 90 from a bivariate t_3-distribution with linear correlation 0.5, the upper graph shows the standard estimator and the lower graph Kendall's tau transform estimator for the linear correlation.

in financial derivative pricing applications. More generally the relationship can be used to calibrate the correlation matrices of higher dimensional elliptical distributions, although in some cases the matrix of pairwise correlations must be adjusted to ensure that the resulting matrix is positive definite; see [5] and [6] for details.

In Section 2 of this note we review the definition and some properties of elliptical distributions. The new result is stated in Section 3 and proved in Section 4.

2. Definitions and basic properties

All random variables mentioned in this paper are real or $\mathbb{R}^n$-valued and all matrices have real entries. Furthermore, all random variables mentioned are assumed to be defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition 1. If X is a p-dimensional random (column) vector and, for some vector $\mu \in \mathbb{R}^p$, some $p \times p$ nonnegative definite symmetric matrix Σ and some function $\phi : [0, \infty) \to \mathbb{R}$, the characteristic function $\varphi_{X-\mu}$ of $X - \mu$ is of the form $\varphi_{X-\mu}(t) = \phi(t^{\mathrm{T}}\Sigma t)$, we say that X has an *elliptical distribution* with parameters μ, Σ and ϕ, and we write $X \sim E_p(\mu, \Sigma, \phi)$.

When $p = 1$, the class of elliptical distributions coincides with the class of one-dimensional symmetric distributions.

Theorem 1. *$X \sim E_p(\mu, \Sigma, \phi)$ with* $\operatorname{rank}(\Sigma) = k$ *if and only if there exist a random variable $R \geq 0$ independent of U, a k-dimensional random vector uniformly* *distributed on the unit hypersphere* $\{z \in \mathbb{R}^k \,|\, z^{\mathrm{T}} z = 1\}$, *and a* $p \times k$ *matrix* A *with* $AA^{\mathrm{T}} = \Sigma$, *such that*

$$X \overset{\mathrm{d}}{=} \mu + RAU. \tag{2}$$

For the proof of Theorem 1 and details about the relation between R and ϕ, see Fang, Kotz and Ng (1987) [4] or Cambanis, Huang and Simons (1981) [1].

Remark 1. (a) Note that the representation (2) is not unique: if $\mathcal{O}$ is an orthogonal $k \times k$ matrix, then (2) also holds with $A' \triangleq A\mathcal{O}$ and $U' \triangleq \mathcal{O}^{\mathrm{T}} U$.

(b) Note that elliptical distributions with different parameters can be equal: if $X \sim E_p(\mu, \Sigma, \phi)$, then $X \sim E_p(\mu, c\Sigma, \phi_c)$ for every $c > 0$, where $\phi_c(s) \triangleq \phi(s/c)$ for all $s \geq 0$.

For $X = (X_1, \dots, X_p)^{\mathrm{T}} \sim E_p(\mu, \Sigma, \phi)$ with $\mathbb{P}\{X_i = \mu_i\} < 1$ and $\mathbb{P}\{X_j = \mu_j\} < 1$, we call $\varrho_{ij} \triangleq \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ the *linear correlation coefficient* for X_i and X_j. If $\operatorname{Var}(X_i)$ and $\operatorname{Var}(X_j)$ are finite, then $\varrho_{ij} = \operatorname{Cov}(X_i, X_j)/\sqrt{\operatorname{Var}(X_i)\operatorname{Var}(X_j)}$.

Definition 2. *Kendall's tau* for the random variables X_1, X_2 is defined as

$$\tau(X_1, X_2) \triangleq \mathbb{P}\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\} - \mathbb{P}\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\},$$

where $(\tilde{X}_1, \tilde{X}_2)$ is an independent copy of (X_1, X_2).

3. Main Results

The main result of this note is the following theorem; its proof is a combination of Lemmas 2 and 7 below. For the case of normal distributions, see also Lemma 6.

Theorem 2. *Let* $X \sim E_p(\mu, \Sigma, \phi)$. *If* $i, j \in \{1, \dots, p\}$ *satisfy* $\mathbb{P}\{X_i = \mu_i\} < 1$ *and* $\mathbb{P}\{X_j = \mu_j\} < 1$, *then*

$$\tau(X_i, X_j) = \left(1 - \sum_{x \in \mathbb{R}} (\mathbb{P}\{X_i = x\})^2\right) \frac{2}{\pi} \arcsin \varrho_{ij}, \tag{3}$$

where the sum extends over all atoms of the distribution of X_i. *If in addition* $\operatorname{rank}(\Sigma) \geq 2$, *then* (3) *simplifies to*

$$\tau(X_i, X_j) = \big(1 - (\mathbb{P}\{X_i = \mu_i\})^2\big) \frac{2}{\pi} \arcsin \varrho_{ij}, \tag{4}$$

which further simplifies to (1) *if* $\mathbb{P}\{X_i = \mu_i\} = 0$.

The following lemma states that linear combinations of independent elliptically distributed random vectors with the same dispersion matrix Σ (up to a positive constant, see Remark 1) remains elliptical. This lemma is of independent interest.

Lemma 1. *Let* $X \sim E_p(\mu, \Sigma, \phi)$ *and* $\tilde{X} \sim E_p(\tilde{\mu}, c\Sigma, \tilde{\phi})$ *for* $c > 0$ *be independent. Then for* $a, b \in \mathbb{R}$, $aX + b\tilde{X} \sim E_p(a\mu + b\tilde{\mu}, \Sigma, \phi^*)$ *with* $\phi^*(u) \triangleq \phi(a^2 u)\, \tilde{\phi}(b^2 c u)$.

Proof. For all $t \in \mathbb{R}^p$,

$$\begin{aligned} \varphi_{aX + b\tilde{X} - a\mu - b\tilde{\mu}}(t) &= \varphi_{a(X-\mu)}(t)\, \varphi_{b(\tilde{X}-\tilde{\mu})}(t) \\ &= \phi\big((at)^{\mathrm{T}} \Sigma (at)\big)\, \tilde{\phi}\big((bt)^{\mathrm{T}} (c\Sigma)(bt)\big) \\ &= \phi(a^2 t^{\mathrm{T}} \Sigma t)\, \tilde{\phi}(b^2 c t^{\mathrm{T}} \Sigma t). \end{aligned}$$

□

4. Proof of Theorem 2

The following lemma gives the relation between Kendall's tau and the linear correlation coefficient for elliptical random vectors of pairwise comonotonic or countermonotonic components. It proves Theorem 2 for the case $\operatorname{rank}(\Sigma) = 1$.

Lemma 2. *Let $X \sim E_p(\mu, \Sigma, \phi)$ with $\operatorname{rank}(\Sigma) = 1$. If $\mathbb{P}\{X_i = \mu_i\} < 1$, and $\mathbb{P}\{X_j = \mu_j\} < 1$, then*

$$\tau(X_i, X_j) = \Bigl(1 - \sum_{x\in\mathbb{R}} (\mathbb{P}\{X_i = x\})^2\Bigr) \frac{2}{\pi} \arcsin \varrho_{ij}. \tag{5}$$

Proof. Let $\tilde{X}$ be an independent copy of X. Let $X \stackrel{\mathrm{d}}{=} \mu + RAU$ and $\tilde{X} \stackrel{\mathrm{d}}{=} \mu + \tilde{R}A\tilde{U}$ be stochastic representations according to Theorem 1, where $(\tilde{R}, \tilde{U})$ denotes an independent copy of (R, U). In particular, A is a $p \times 1$ matrix and U is symmetric $\{1, -1\}$-valued. Furthermore, $\mathbb{P}\{X_i = \mu_i\} < 1$ and $\mathbb{P}\{X_j = \mu_j\} < 1$ imply $A_{i1} \neq 0$ and $A_{j1} \neq 0$. Therefore,

$$\varrho_{ij} = A_{i1}A_{j1}\Big/\sqrt{A_{i1}^2 A_{j1}^2} = \operatorname{sign}(A_{i1}A_{j1}) = \frac{2}{\pi} \arcsin \varrho_{ij}, \tag{6}$$

$(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) \stackrel{\mathrm{d}}{=} A_{i1}A_{j1}(RU - \tilde{R}\tilde{U})^2$ and

$$\mathbb{P}\{RU = \tilde{R}\tilde{U}\} = \sum_{x\in\mathbb{R}} (\mathbb{P}\{RU = x\})^2 = \sum_{x\in\mathbb{R}} (\mathbb{P}\{X_i = x\})^2. \tag{7}$$

If $A_{i1}A_{j1} > 0$, then by Definition 2

$$\tau(X_i, X_j) = \mathbb{P}\{(RU - \tilde{R}\tilde{U})^2 > 0\} = 1 - \mathbb{P}\{RU = \tilde{R}\tilde{U}\}$$

Using (6) and (7), the result (5) follows. If $A_{i1}A_{j1} < 0$, then

$$\tau(X_i, X_j) = -\,\mathbb{P}\{(RU - \tilde{R}\tilde{U})^2 > 0\}$$

and the result (5) follows in the same way. □

Lemma 3. *Let $X \sim E_p(\mu, \Sigma, \phi)$ with $\operatorname{rank}(\Sigma) = k \geq 2$ and let $\tilde{X}$ be an independent copy of X. If $\mathbb{P}\{X_i = \mu_i\} < 1$, then $\mathbb{P}\{X_i = \tilde{X}_i\} = (\mathbb{P}\{X_i = \mu_i\})^2$.*

Proof. Let $X \stackrel{\mathrm{d}}{=} \mu + RAU$ be a stochastic representation according to Theorem 1. Define $A_i \triangleq (A_{i1}, \ldots, A_{ik})$ and $a \triangleq A_i A_i^{\mathrm{T}}$. Since $\mathbb{P}\{X_i = \mu_i\} < 1$, the case $a = 0$ is excluded. By choosing an orthogonal $k \times k$ matrix $\mathcal{O}$ whose first column is A_i^{T}/a and using Remark 1(a) if necessary, we may assume that $A_i = (a, 0, \ldots, 0)$, hence $X_i \stackrel{\mathrm{d}}{=} \mu_i + aRU_1$. Note that U_1 is a continuous random variable because $k \geq 2$. Hence $\mathbb{P}\{aRU_1 = x\} = 0$ for all $x \in \mathbb{R} \setminus \{0\}$, and it follows that

$$\mathbb{P}\{X_i = \tilde{X}_i\} = \sum_{x\in\mathbb{R}} (\mathbb{P}\{X_i = x\})^2 = \sum_{x\in\mathbb{R}} (\mathbb{P}\{aRU_1 = x\})^2 = (\mathbb{P}\{X_i = \mu_i\})^2.$$
□

Lemma 4. *Let $X \sim E_p(\mu, \Sigma, \phi)$ with $\operatorname{rank}(\Sigma) = k \geq 2$, and let $\tilde{X}$ be an independent copy of X. If $\mathbb{P}\{X_i = \mu_i\} < 1$ and $\mathbb{P}\{X_j = \mu_j\} < 1$, then*

$$\tau(X_i, X_j) = 2\,\mathbb{P}\{(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) > 0\} - 1 + (\mathbb{P}\{X_i = \mu_i\})^2. \tag{8}$$

Proof. Since $Y \triangleq X - \tilde{X} \sim E_p(0, \Sigma, \phi^2)$ by Lemma 1, there exists a stochastic representation $Y = RAU$ according to Theorem 1. By Lemma 3, $\mathbb{P}\{Y_i = 0\} = (\mathbb{P}\{X_i = \mu_i\})^2 < 1$ and similarly $\mathbb{P}\{Y_j = 0\} < 1$. Define $A_i \triangleq (A_{i1}, \ldots, A_{ik})$ and $A_j \triangleq (A_{j1}, \ldots, A_{jk})$. With the same arguments as in the proof of Lemma 3, it follows that A_iU and A_jU are continuous random variables, which implies that $\mathbb{P}\{A_iU = 0\} = 0$ and $\mathbb{P}\{A_jU = 0\} = 0$. Therefore,

$$\mathbb{P}\{Y_iY_j = 0\} = \mathbb{P}\{R = 0\} = \mathbb{P}\{Y_i = 0\} = (\mathbb{P}\{X_i = \mu_i\})^2.$$

Since $\tau(X_i, X_j) = 2\,\mathbb{P}\{Y_iY_j > 0\} - 1 + \mathbb{P}\{Y_iY_j = 0\}$, the conclusion follows. □

Lemma 5. *Let* $X \sim E_p(0, \Sigma, \phi)$ *and* $\tilde{X} \sim E_p(0, c\Sigma, \tilde{\phi})$ *with* $\operatorname{rank}(\Sigma) \geq 2$ *and* $c > 0$*. If* $\mathbb{P}\{X_i = 0\} < 1$ *and* $\mathbb{P}\{\tilde{X}_i = 0\} < 1$*, then*

$$\mathbb{P}\{X_iX_j > 0\}(1 - \mathbb{P}\{\tilde{X}_i = 0\}) = \mathbb{P}\{\tilde{X}_i\tilde{X}_j > 0\}(1 - \mathbb{P}\{X_i = 0\}).$$

Proof. Take $X \stackrel{\mathrm{d}}{=} RAU$ according to Theorem 1 and set $W \triangleq AU$. Then

$$\begin{aligned}\mathbb{P}\{X_iX_j > 0\} &= \mathbb{P}\{RW_iRW_j > 0\} \\ &= \mathbb{P}\{RW_iRW_j > 0 \,|\, R > 0\}\,\mathbb{P}\{R > 0\} \\ &= \mathbb{P}\{W_iW_j > 0\}\,\mathbb{P}\{R > 0\}.\end{aligned}$$

Furthermore, $\tilde{X} \stackrel{\mathrm{d}}{=} \sqrt{c}\tilde{R}W$ according to Theorem 2, and a similar calculation shows

$$\mathbb{P}\{\tilde{X}_i\tilde{X}_j > 0\} = \mathbb{P}\{c\tilde{R}^2W_iW_j > 0 \,|\, \tilde{R} > 0\}\,\mathbb{P}\{\tilde{R} > 0\} = \mathbb{P}\{W_iW_j > 0\}\,\mathbb{P}\{\tilde{R} > 0\}.$$

As in the proof of Lemma 3, it follows that W_i has a continuous distribution. Therefore, $\mathbb{P}\{R > 0\} = 1 - \mathbb{P}\{X_i = 0\}$ and $\mathbb{P}\{\tilde{R} > 0\} = 1 - \mathbb{P}\{\tilde{X}_i = 0\}$, and Lemma 5 follows. □

Although the next result for normal distributions is well known, we give a proof for completeness of the exposition and for showing where the arcsin comes from.

Lemma 6. *Let* $X \sim \mathcal{N}_p(\mu, \Sigma)$*. If* $\mathbb{P}\{X_i = \mu_i\} < 1$ *and* $\mathbb{P}\{X_j = \mu_j\} < 1$*, then*

$$\tau(X_i, X_j) = 2\,\mathbb{P}\{(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) > 0\} - 1 = \frac{2}{\pi}\arcsin\varrho_{ij},$$

where $\tilde{X}$ *is an independent copy of* X*.*

Proof. Using $\sigma_i \triangleq \sqrt{\Sigma_{ii}} > 0$, $\sigma_j \triangleq \sqrt{\Sigma_{jj}} > 0$ and $\varrho_{ij} \triangleq \Sigma_{ij}/\sigma_i\sigma_j$, we have

$$\Sigma^{ij} = \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & \sigma_i\sigma_j\varrho_{ij} \\ \sigma_i\sigma_j\varrho_{ij} & \sigma_j^2 \end{pmatrix}.$$

Define $Y \triangleq X - \tilde{X}$ and note that $(Y_i, Y_j) \sim \mathcal{N}_2(0, 2\Sigma^{ij})$. Furthermore, $(Y_i, Y_j) \stackrel{\mathrm{d}}{=} \sqrt{2}(\sigma_iV\cos\varphi_{ij} + \sigma_iW\sin\varphi_{ij}, \sigma_jW)$, where $\varphi_{ij} \triangleq \arcsin\varrho_{ij} \in [-\pi/2, \pi/2]$ and (V, W) is standard normally distributed. By the radial symmetry of (Y_i, Y_j),

$$\begin{aligned}\tau(X_i, X_j) &= 2\,\mathbb{P}\{Y_iY_j > 0\} - 1 = 4\,\mathbb{P}\{Y_i > 0, Y_j > 0\} - 1 \\ &= 4\,\mathbb{P}\{V\cos\varphi_{ij} + W\sin\varphi_{ij} > 0,\, W > 0\} - 1.\end{aligned}$$

If Φ is uniformly distributed on $[-\pi, \pi)$, independent of $R \triangleq \sqrt{V^2 + W^2}$, then $(V, W) \stackrel{\mathrm{d}}{=} R(\cos\Phi, \sin\Phi)$ and

$$\begin{aligned}\tau(X_i, X_j) &= 4\,\mathbb{P}\big\{\cos\Phi\cos\varphi_{ij} + \sin\Phi\sin\varphi_{ij} > 0,\, \sin\Phi > 0\big\} - 1 \\ &= 4\,\mathbb{P}\big\{\Phi \in (\varphi_{ij} - \pi/2, \varphi_{ij} + \pi/2) \cap (0, \pi)\big\} - 1 = 4\,\frac{\varphi_{ij} + \pi/2}{2\pi} - 1,\end{aligned}$$

which simplifies to $(2/\pi)\arcsin\varrho_{ij}$. □

Lemma 7. *Let* $X \sim E_p(\mu, \Sigma, \phi)$ *with* $\operatorname{rank}(\Sigma) = k \geq 2$*. If* $\mathbb{P}\{X_i = \mu_i\} < 1$ *and* $\mathbb{P}\{X_j = \mu_j\} < 1$*, then*

$$\tau(X_i, X_j) = \bigl(1 - (\mathbb{P}\{X_i = \mu_i\})^2\bigr) \frac{2}{\pi} \arcsin \varrho_{ij}. \tag{9}$$

Proof. Let $\tilde{X}$ be an independent copy of X. By Lemma 4, we can use (8). By Lemmas 1 and 3, $X - \tilde{X} \sim E_p(0, \Sigma, \phi^*)$ with $\mathbb{P}\{X_i = \tilde{X}_i\} = (\mathbb{P}\{X_i = \mu_i\})^2 < 1$ and $\mathbb{P}\{X_j = \tilde{X}_j\} = (\mathbb{P}\{X_j = \mu_j\})^2 < 1$. If Z, $\tilde{Z} \sim \mathcal{N}_p(\mu, \Sigma/2)$ are independent, then $Z - \tilde{Z} \sim \mathcal{N}_p(0, \Sigma)$. By Lemma 5,

$$\mathbb{P}\{(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) > 0\} = \mathbb{P}\{(Z_i - \tilde{Z}_i)(Z_j - \tilde{Z}_j) > 0\}(1 - (\mathbb{P}\{X_i = \mu_i\})^2).$$

Substituting this into (8) and using Lemma 6, the result (9) follows. □

References

1. S. Cambanis, S. Huang, and G. Simons, *On the theory of elliptically contoured distributions*, J. Multivariate Anal. **11** (1981), 368–385.
2. H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.
3. S. J. Devlin, R. Gnanadesikan, and J. R. Kettenring, *Robust estimation of dispersion matrices and principal components*, J. Amer. Statist. Assoc. **76** (1981), 354–362.
4. K.-T. Fang, S. Kotz, and K.-W. Hg, *Symmetric Multivariate and Related Distributions*, Chapman & Hall, London, 1987.
5. F. Lindskog, *Linear correlation estimation*, Research report, RiskLab Switzerland, August 2000, http://www.risklab.ch/Papers.html#LCELindskog.
6. P. J. Rousseeuw and G. Molenberghs, *Transformation of non positive semidefinite correlation matrices*, Comm. Statist. A—Theory Methods **22** (1993), no. 4, 965–984.

(F. Lindskog) RiskLab, Departement Mathematik, ETH Zentrum, HG F 42.3, CH-8092 Zürich, Switzerland

E-mail address: `lindskog@math.ethz.ch`

URL: `http://www.math.ethz.ch/~lindskog/`

(A. McNeil) Departement Mathematik, ETH Zentrum, HG D 53.2, CH-8092 Zürich, Switzerland

E-mail address: `mcneil@math.ethz.ch`

URL: `http://www.math.ethz.ch/~mcneil/`

(U. Schmock) RiskLab, Departement Mathematik, ETH Zentrum, HG F 42.1, CH-8092 Zürich, Switzerland

E-mail address: `schmock@math.ethz.ch`

URL: `http://www.math.ethz.ch/~schmock/`