



HAL
open science

A note on skew-elliptical distributions and linear functions of order statistics

Nicola Loperfido

► **To cite this version:**

Nicola Loperfido. A note on skew-elliptical distributions and linear functions of order statistics. *Statistics and Probability Letters*, 2009, 78 (18), pp.3184. 10.1016/j.spl.2008.06.004 . hal-00510972

HAL Id: hal-00510972

<https://hal.science/hal-00510972>

Submitted on 23 Aug 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

A note on skew-elliptical distributions and linear functions of order statistics

Nicola Loperfido

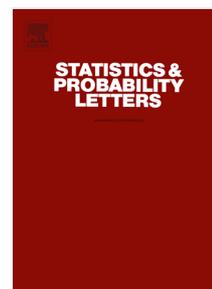
PII: S0167-7152(08)00295-2
DOI: [10.1016/j.spl.2008.06.004](https://doi.org/10.1016/j.spl.2008.06.004)
Reference: STAPRO 5108

To appear in: *Statistics and Probability Letters*

Received date: 13 December 2007
Revised date: 6 June 2008
Accepted date: 6 June 2008

Please cite this article as: Loperfido, N., A note on skew-elliptical distributions and linear functions of order statistics. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.06.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



A Note on Skew-Elliptical Distributions and Linear Functions of Order Statistics

Nicola Loperfido

Dipartimento di Economia e Metodi Quantitativi

Università di Urbino "Carlo Bo", Via Saffi 42, 61033 Urbino (PU), ITALY

Abstract

Linear functions of order statistics from bivariate, exchangeable, continuous and elliptical random vectors have a skew-elliptical distribution. This result generalizes previous ones on order statistics and skewed distributions.

Key words: Exchangeability, order statistics, skew-elliptical distribution, skew-normal distribution.

1 Introduction

Nagarajah (1982) obtained the distribution of a linear combination of order statistics from a bivariate, exchangeable and normal random vector. The same distribution independently appeared in biostatistics (Roberts, 1966), econometrics (Aigner *et al.*, 1977) and time series (Andel *et al.*, 1984). Nowadays it is known as skew-normal distribution, and it was given its actual name by Azzalini (1985). The probability density function of the skew-normal distribution $SN(\xi, \psi, \lambda)$ is

$$f(y; \xi, \psi, \lambda) = \frac{2}{\psi} \phi\left(\frac{y - \xi}{\psi}\right) \Phi\left(\lambda \frac{y - \xi}{\psi}\right), \quad (1)$$

where $\xi, y, \lambda \in \mathbb{R}$, $\psi \in \mathbb{R}^+$, ϕ and Φ are the pdf and cdf of a standard normal distribution, respectively.

Apparently, the connection between order statistics and skewed distributions was first noticed by Loperfido (2002), while independently rediscovering a result in Roberts (1966), and later generalized by Azzalini & Capitanio (2003).

¹Address for Correspondence: Telephone 0039-0722-331360. Fax: 0039-0722-305-550. E-mail: nicola.loperfido@uniurb.it

²This research was supported by the Ministero dell'Università, dell'Istruzione e della Ricerca with grant PRIN No. 2006132978.

Other connections between order statistics and generalizations of the normal distribution are dealt with in Crocetta & Loperfido (2005) and Arellano-Valle & Genton (2008).

Bivariate and exchangeable random vectors often occur when modelling observations taken from both sides of the same individual. Indeed, “Observations between related measurements, such as with eyes, ears, siblings, etc., possess intrinsic symmetries that may be relevant for assessing an underlying physiological process” (Olkin & Viana, 2000).

In vision research, for example, measures of visual acuity are often made in right (Y_R) and left (Y_L) eye, together with one or more covariates X . Ordered acuity measurements $Y_1 = \min(Y_R, Y_L)$ and $Y_2 = \max(Y_R, Y_L)$ are required to determine the person’s total vision impairment, defined as $(3Y_1 + Y_2)/4$ (Rubin *et al.*, 1984). Olkin & Viana (1995) and Viana (1998) modelled the joint distribution of (Y_R, Y_L, X) via normality and exchangeability assumptions, obtaining first and second moments of (Y_1, Y_2, X) , and hence of $(hY_1 + kY_2, X)$, where $h, k \in \mathbb{R}$.

The paper generalizes results in Nagarajah (1982) and Azzalini & Capitanio (2003). It is structured as follows. Section 2 reviews some basic concepts and notations related to elliptical and skew-elliptical distributions. Section 3 contains the main result, its proof and some corollaries.

2 Skew-elliptical distributions

A random vector is spherically distributed (spherical, for short) if its distribution is invariant with respect to orthogonal transformations and it is elliptically distributed (elliptical, for short) if it can be represented as an affine transformation of a spherical random vector (Fang *et al.*, 1990). The probability density function of an elliptical and p -dimensional random vector \mathbf{x} , when it exists, can be represented as follows:

$$f(\mathbf{x}; \boldsymbol{\xi}, \Omega) = \frac{1}{\sqrt{|\Omega|}} g\left[(\mathbf{x} - \boldsymbol{\xi})^T \Omega^{-1} (\mathbf{x} - \boldsymbol{\xi})\right], \quad (2)$$

where $\boldsymbol{\xi} \in \mathbb{R}^p$, $\Omega \in \mathbb{R}^p \times \mathbb{R}^p$ is a symmetric, positive definite $p \times p$ matrix and g is a function from \mathbb{R}^+ to \mathbb{R}^+ satisfying

$$\int_0^{+\infty} y^{p/2-1} g(y) dy = \frac{\Gamma(p/2)}{\sqrt{\pi^p}}. \quad (3)$$

The vector $\boldsymbol{\xi}$ and the matrix Ω are commonly referred to as the location vector (parameter) and the scale matrix (parameter). They are not in general equal to the mean vector and the covariance matrix, respectively. The function g is said to be the generator of the corresponding distribution, which is denoted by $E_p(\boldsymbol{\xi}, \Omega, g)$. Branco & Dey (2001) introduced the class of skew-elliptical distributions, which generalizes elliptical ones and it is defined via the following

conditioning argument. Let the joint distribution of the random variable Y and the random vector \mathbf{x} be

$$\begin{pmatrix} Y \\ \mathbf{x} \end{pmatrix} \sim E_{p+1} \left[\begin{pmatrix} 0 \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix}, g \right], \quad (4)$$

where Ω is a symmetric, positive definite $p \times p$ matrix and $\boldsymbol{\delta} \in \mathbb{R}^p$ satisfies the constraint $\boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta} < 1$. Then the distribution of $\mathbf{x}|Y > 0$ is said to be skew-elliptical with location parameter $\boldsymbol{\xi}$, scale parameter Ω , shape parameter $\boldsymbol{\alpha} = \Omega^{-1} \boldsymbol{\delta} / \sqrt{1 - \boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta}}$ and generating function g , denoted by $SE_p(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g)$. The probability density function of $SE_p(\boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g)$ has the following form:

$$f(\mathbf{z}; \boldsymbol{\xi}, \Omega, \boldsymbol{\alpha}, g) = \frac{2}{\sqrt{|\Omega|}} \int_{-\infty}^{\boldsymbol{\alpha}^T(\mathbf{z}-\boldsymbol{\xi})} g \left[r^2 + (\mathbf{y} - \boldsymbol{\xi})^T \Omega^{-1} (\mathbf{y} - \boldsymbol{\xi}) \right] dr. \quad (5)$$

Branco & Dey (2001) obtain several other representations of skew-elliptical densities which are not of interest in the present context and are therefore omitted. Notice that $SE_1(\xi, \psi^2, \lambda/\psi, g)$, where $g(x) = (2\pi)^{-1} \exp(-x/2)$, corresponds to the skew-normal distribution $SN(\xi, \psi, \lambda)$. Azzalini & Capitanio (2003) showed that the maximum component of a bivariate, continuous, elliptical and exchangeable random vector has a skew-elliptical distribution.

3 Main result

The following theorem generalizes results in Nagarajah (1982) as well as in Azzalini & Capitanio (2003), since it deals both with elliptical distributions and linear combinations of order statistics.

Theorem 1: *Let X and Y be two exchangeable, elliptical and continuous random variables:*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim E_2 \left[\begin{pmatrix} \xi \\ \xi \end{pmatrix}, \psi^2 \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix}, g \right], \quad (6)$$

where $\xi \in \mathbb{R}$, $\psi \in \mathbb{R}^+$ and $|\omega| < 1$. Then for any h, k with $h \neq -k$ the distribution of $h \min(X, Y) + k \max(X, Y)$ is

$$SE_1 \left[\xi(h+k), \psi^2(h^2 + k^2 + 2\omega hk), \frac{1}{\psi} \frac{k-h}{|k+h|} \sqrt{\frac{1-\omega}{1+\omega}}, g \right]. \quad (7)$$

Proof: Let the random variables Q, W and U be defined as follows:

$$Q = h \min(X, Y) + k \max(X, Y); \quad W = hX + kY; \quad U = \frac{Y - X}{\sqrt{2\psi^2(1-\omega)}}. \quad (8)$$

The random variable X (Y) equals $\min(X, Y)$ if X (Y) is smaller than Y (X). A similar argument holds for $\max(X, Y)$. Then for an arbitrary q the probability $P(Q \leq q)$ equals

$$P(hX + kY \leq q; X < Y) + P(hY + kX \leq q; Y < X). \quad (9)$$

By assumption, X and Y are exchangeable and continuous random variables. Hence $P(X < Y) = P(Y < X) = 0.5$ and

$$P(hX + kY \leq q | X < Y) = P(hY + kX \leq q | Y < X). \quad (10)$$

These equations, together with the definitions of W and U , imply that

$$P(Q \leq q) = P(hX + kY \leq q | X < Y) = P(W \leq q | U > 0). \quad (11)$$

We shall now consider the joint distribution of W and U . The class of elliptical distributions is closed with respect to linear transformations, so that the joint distribution of W and U is

$$E_2 \left\{ \left[\begin{array}{c} \xi(h+k) \\ 0 \end{array} \right], \left[\begin{array}{cc} \psi^2(h^2 + k^2 + 2\omega hk) & \psi(k-h)\sqrt{\frac{1-\omega}{2}} \\ \psi(k-h)\sqrt{\frac{1-\omega}{2}} & 1 \end{array} \right], g \right\}. \quad (12)$$

Results in Branco & Dey (2001) imply that the distribution of $W | U > 0$ is univariate skew-elliptical with location parameter $\xi(h+k)$, scale parameter $\psi^2(h^2 + k^2 + 2\omega hk)$, generating function g and shape parameter

$$\begin{aligned} & \frac{[\psi(k-h)\sqrt{(1-\omega)/2}] / [\psi^2(h^2 + k^2 + 2\omega hk)]}{\sqrt{1 - [\psi(k-h)\sqrt{(1-\omega)/2}]^2 / [\psi^2(h^2 + k^2 + 2\omega hk)]}} \\ &= \frac{1}{\psi} \frac{(k-h)\sqrt{1-\omega}}{\sqrt{2(h^2 + k^2 + 2\omega hk) - (k-h)^2(1-\omega)}} \\ &= \frac{1}{\psi} \frac{(k-h)\sqrt{1-\omega}}{\sqrt{h^2 + k^2 + 2\omega hk + 2hk + \omega k^2 + \omega h^2}} \\ &= \frac{1}{\psi} \frac{(k-h)\sqrt{1-\omega}}{\sqrt{(h+k)^2(1+\omega)}} = \frac{1}{\psi} \frac{k-h}{|k+h|} \sqrt{\frac{1-\omega}{1+\omega}}. \end{aligned} \quad (13)$$

From (11) we know that Q and $W | U > 0$ are identically distributed and this completes the proof.

In order to show the connection between the above theorem and previous literature, corollaries 1 and 2 restate results in Nagarajah (1982) and Azzalini & Capitanio (2003).

Corollary 1 Let X and Y be two random variables whose joint distribution is bivariate normal with $E(X) = E(Y) = \xi$, $V(X) = V(Y) = \psi^2$ and $C(X, Y) = \omega\psi^2$. Then for any two constants h and $k \neq -h$ the distribution of $h \min(X, Y) + k \max(X, Y)$ is

$$SN \left[\xi(h+k), \psi^2(h^2 + k^2 + 2\omega hk), \frac{k-h}{|k+h|} \sqrt{\frac{1-\omega}{1+\omega}} \right]. \quad (14)$$

Corollary 2 Let X and Y be two exchangeable, elliptical and continuous random variables:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim E_2 \left[\begin{pmatrix} \xi \\ \xi \end{pmatrix}, \psi^2 \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix}, g \right]. \quad (15)$$

Then $\max(X, Y) \sim SE_1 \left[\xi, \psi^2, \psi^{-1} \sqrt{(1-\omega)/(1+\omega)}, g \right]$.

Acknowledgements

The author would like to thank an anonymous referee for her or his helpful comments, which greatly improved the quality of the presentation.

References

- Aigner, D. J., Lovell, C. A. K. & Schmidt, P. (1977), Formulation and estimation of stochastic frontier production function model, *J. Econometrics* **12**, 21-37.
- Andel, J., Netuka, I. & Zvára, K. (1984), On threshold autoregressive processes, *Kybernetika* **20**, 89-106.
- Arellano-Valle, R. B. & Genton, M.G. (2008). On the Exact Distribution of the Maximum of Absolutely Continuous Dependent Random Variables. *Statist. & Prob. Lett.* **78**, 27-35.
- Azzalini, A. (1985), A class of distributions which includes the normal ones, *Scand. J. Statist.* **12**, 171-178.
- Azzalini, A. & Capitanio, A. (2003), Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution, *J. R. Statist. Soc. B* **65**, 367-389.
- Branco, M.D. & Dey, D.K. (2001), A general class of skew-elliptical distributions, *J. Multivariate Anal.* **79**, 99-113.
- Crocetta, C. & Loperfido, N. (2005). The Exact Sampling Distribution of L-statistics. *Metron* **63**, 213-223.
- Fang, K.T., Kotz, S. & Ng, K.W (1990), *Symmetric Multivariate and Related Distributions*, (Chapman and Hall, Boca Raton, FL).

Loperfido N. (2002), Statistical implications of selectively reported inferential results, *Statist. & Prob. Lett.* **56**, 13-22.

Nagarajah, H. N. (1982), A note on linear functions of ordered correlated normal random variables, *Biometrika* **69**, 284-285.

Olkin I. & Viana M. (1995), Correlation analysis of extreme observations from a multivariate normal distribution, *J. Am. Statist. Assoc.* **90**, 1373-1379.

Olkin I. & Viana M. (2000), Symmetrically dependent models arising in visual assessment data, *Biometrics* **56**, 1188-1191.

Roberts C. (1966), A correlation model useful in the study of twins, *J. Am. Statist. Assoc.* **61**, 1184-1190.

Rubin, G.S., Munoz, B., Fried, L.P. & West, S. (1984), Monocular vs binocular visual acuity as measures of vision impairment, *Vision Science and Its Applications-OSA Technical Digest Series* **1**, 328-331.

Viana M. (1998), Linear combinations of ordered symmetric observations with application to visual acuity, in: N. Balakrishnan & C. R. Rao, eds, *Handbook of Statistics* **17**, (Elsevier Science, Amsterdam) pp. 513-524.