Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 78 (2008) 27–35
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

On the exact distribution of the maximum of absolutely continuous dependent random variables

Reinaldo B. Arellano-Valle[a,1], Marc G. Genton[b,c,*,2]

[a]Departamento de Estadística, Facultad de Matemática, Pontificia Universidad Católica de Chile, Santiago 22, Chile
[b]Department of Econometrics, University of Geneva, Bd du Pont-d'Arve 40, CH-1211 Geneva 4, Switzerland
[c]Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA

Received 27 June 2005; received in revised form 30 January 2007; accepted 10 April 2007
Available online 7 May 2007

Abstract

We derive the exact probability density function of the maximum of arbitrary absolutely continuous dependent random variables and of absolutely continuous exchangeable random variables. We show this density is related to the family of fundamental skew distributions. In particular, we examine the case where the random variables have an elliptically contoured distribution. We study some particular examples based on the multivariate normal and multivariate Student t distributions, and discuss numerical computation issues. We illustrate our results on a genetic selection problem and on an autoregressive time series model of order one.

Keywords: Elliptically contoured; Exchangeable; Kurtosis; Maximum; Fundamental skew distribution; Skewness; Time series

1. Introduction

Let $\mathbf{X} = (X_1, \ldots, X_n)^T$ be an absolutely continuous random vector, which will be assumed to have a probability density on $\mathbb{R}^n$. We are interested in the problem of finding the exact distribution of the maximum $X_{(n)} = \max_{1 \leqslant i \leqslant n}\{X_i\}$. The solution is well known in the case of independent and identically distributed (i.i.d.) random variables $X_1, \ldots, X_n$, see Gumbel (1958). The asymptotic distributions of $X_{(n)}$ have received considerable attention, both in the i.i.d. setting, see e.g. David (1981, Chapter 8) and references therein, and under some form of dependence such as m-dependence, see e.g. Watson (1954) and Ghosh (1972). An approximate formula for the distribution of $X_{(n)}$ in the case of normal dependent random variables has been derived by Greig (1967). The exact probability density function of $X_{(n)}$ when $\mathbf{X}$ is an exchangeable multivariate normal random vector, i.e. its covariance matrix is equicorrelated, has been considered by Tong (1990, p. 126), who proposed to derive this density as a location mixture of the distribution corresponding to the maximum of

*Corresponding author. Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA.
Fax: +1 979 845 3144.
E-mail addresses: reivalle@mat.puc.cl (R.B. Arellano-Valle), Marc.Genton@metri.unige.ch, genton@stat.tamu.edu (M.G. Genton).
[1]This research was partially supported by Grant FONDECYT 1040865/7050091/7060133/7070137-Chile.
[2]This research was partially supported by NSF Grant DMS-0504896.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2007.04.021

i.i.d. normal random variables. Thus, an integration problem needs to be solved in order to use Tong's result. We provide another exact form of this density and give also an extension to arbitrary dependent random variables, not necessarily normal and/or exchangeable. In particular, when the parent distribution of $\mathbf{X}$ is exchangeable, we show that the exact distribution of $X_{(n)}$ belongs to the family of fundamental skew distributions recently introduced by Arellano-Valle and Genton (2005). Specifically, let $\mathbf{X} = (\mathbf{Y}|\mathbf{Z}\geqslant\mathbf{0})$, where $\mathbf{Y} \in \mathbb{R}^k$ is a random vector with probability density function $f_{\mathbf{Y}}$, $\mathbf{Z} \in \mathbb{R}^m$ is a random vector, and the notation $\mathbf{Z}\geqslant\mathbf{0}$ is meant component-wise. Then Arellano-Valle and Genton (2005) say that $\mathbf{X}$ has a k-dimensional fundamental skew (FUS) distribution with probability density function:

$$f_{\mathbf{X}}(\mathbf{x}) = K_m^{-1} f_{\mathbf{Y}}(\mathbf{x}) Q_m(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^k, \tag{1}$$

where $Q_m(\mathbf{x}) = P(\mathbf{Z}\geqslant\mathbf{0}|\mathbf{Y} = \mathbf{x})$ and $K_m = \mathrm{E}(Q_m(\mathbf{Y})) = P(\mathbf{Z}\geqslant\mathbf{0})$. In particular, when $f_{\mathbf{Y}}$ is a symmetric probability density function (i.e. $f_{\mathbf{Y}}(-\mathbf{y}) = f_{\mathbf{Y}}(\mathbf{y})$ for all $\mathbf{y} \in \mathbb{R}^k$), (1) defines the fundamental skew-symmetric (FUSS) class of distributions. Note that K_m is a normalizing constant and the term Q_m may be interpreted as a function causing skewness in the density $f_{\mathbf{X}}$. Indeed, we consider $\mathbf{Y}$ conditionally on $\mathbf{Z}\geqslant\mathbf{0}$ and this selection mechanism induces skewness, see Arellano-Valle et al. (2006) for a unified view on skewed distributions resulting from selections.

Our motivation comes from a genetic selection problem in agricultural research, originally considered by Rawlings (1976) and Hill (1976, 1977), and more recently by Tong (1990, p. 129). To describe the problem briefly, suppose that an agricultural genetic selection project involves n animals, for example pigs, and the top performer is to be selected for breeding. Let $X_1, \ldots, X_n$ be the measurements of a certain biological or physical characteristic of the n animals, such as the body weights or back fats of the pigs. The animal with score $X_{(n)}$ is to be selected. If $X_1, \ldots, X_n$ are independent with mean μ, then the common mean of the observations of off-springs of the selected animal with score $X_{(n)}$ is $\mathrm{E}(X_{(n)})$, and therefore the expected gain in one generation is $\mathrm{E}(X_{(n)}) - \mu$. However, the assumption of independence is often not satisfied since the animals under selection are usually genetically related. This is the case, for example, when the pigs are from the same family and have the same parents. In this situation, a variance components model is generally assumed by geneticists, which means that $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ has an exchangeable multivariate normal distribution (Tong, 1990, p. 108) with a common mean μ, a common variance σ^2 and a common correlation coefficient $\rho \in [0, 1)$. In summary, the distribution of $\mathbf{X}$ is assumed to be multivariate normal $\mathrm{N}_n(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\})$, with $\rho \in [0, 1)$, $\mathbf{1}_n \in \mathbb{R}^n$ a vector of ones, and $I_n \in \mathbb{R}^{n\times n}$ the identity matrix. Our goal is to derive an explicit form for the probability density function of $X_{(n)}$ in the above context in order to study its shape with respect to ρ, but also in more general settings of dependence.

The structure of the paper is set up as follows. In Section 2, we derive our main results, namely the exact probability density function of $X_{(n)}$ for arbitrary absolutely continuous dependent random variables and for absolutely continuous exchangeable random variables. In particular, we examine the case where the random variables have an elliptically contoured distribution (Fang et al., 1990). In Section 3, we study some particular examples based on the multivariate normal and multivariate Student t distributions, and discuss numerical computation issues. Extensions of the results in this paper to the exact distribution of linear combinations of order statistics from dependent random variables can now be found in Arellano-Valle and Genton (2007).

2. Main results

2.1. Dependent random variables

Although the next result is straighforward to derive, its importance lies in the link between the exact distribution of the maximum and FUS distributions of the form (1).

Proposition 1. *Let $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ be an absolutely continuous random vector. The probability density function $f_{X_{(n)}}$ of $X_{(n)}$ is*

$$f_{X_{(n)}}(x) = \sum_{i=1}^{n} f_{X_i}(x) F_{\mathbf{X}_{-i}|X_i=x}(x\mathbf{1}_{n-1}), \quad x \in \mathbb{R},$$

where f_{X_i} is the marginal probability density function of X_i and $F_{\mathbf{X}_{-i}|X_i=x}$ is the conditional cumulative distribution function of $\mathbf{X}_{-i}=(X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)^{\mathrm{T}}$ given $X_i=x$.

Proof. Let $F_{X_{(n)}}$ be the cumulative distribution function of $X_{(n)}$. Then, for all x, we have

$$\begin{aligned}F_{X_{(n)}}(x) &= P(X_{(n)}\leqslant x)\\ &= \sum_{i=1}^{n} P(X_i\leqslant x|X_i-X_j\geqslant 0,\ \forall j\neq i)P(X_i-X_j\geqslant 0,\ \forall j\neq i)\\ &= \sum_{i=1}^{n} F_{X_i|X_i-X_j\geqslant 0,\ \forall j\neq i}(x)P(X_i-X_j\geqslant 0,\ \forall j\neq i),\end{aligned}$$

implying that

$$f_{X_{(n)}}(x)=\sum_{i=1}^{n} f_{X_i|X_i-X_j\geqslant 0,\ \forall j\neq i}(x)P(X_i-X_j\geqslant 0,\ \forall j\neq i).$$

Thus, the proof follows by noting from Bayes' theorem that

$$f_{X_i|X_i-X_j\geqslant 0,\ \forall j\neq i}(x)=f_{X_i}(x)\frac{P(X_i-X_j\geqslant 0,\ \forall j\neq i|X_i=x)}{P(X_i-X_j\geqslant 0,\ \forall j\neq i)}. \qquad \square$$

We consider next the class of elliptically contoured distributions for the random vector $\mathbf{X}$. Following Fang et al. (1990), we say that $\mathbf{X}\in\mathbb{R}^n$ has an elliptically contoured distribution with location vector $\boldsymbol{\mu}$, scale matrix Σ, and density generator $h^{(n)}$, if its probability density function has the form

$$f_n(\mathbf{x};\boldsymbol{\mu},\Sigma,h^{(n)})=|\Sigma|^{-1/2}h^{(n)}[(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})],\quad \mathbf{x}\in\mathbb{R}^n,$$

and we use the notation $\mathbf{X}\sim\mathrm{EC}_n(\boldsymbol{\mu},\Sigma,h^{(n)})$. Denote the cumulative distribution function of $\mathbf{X}$ by $F_n(\mathbf{x};\boldsymbol{\mu},\Sigma,h^{(n)})$. It is well known that elliptically contoured distributions are closed under marginalization and conditioning. In particular, if $\mathbf{X}\sim\mathrm{EC}_n(\boldsymbol{\mu},\Sigma,h^{(n)})$ and we consider, for a fixed i, the partition given by

$$\mathbf{X}=\begin{pmatrix}\mathbf{X}_{-i}\\ X_i\end{pmatrix},\quad \boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_{-i}\\ \mu_i\end{pmatrix},\quad \Sigma=\begin{pmatrix}\Sigma_{-i-i} & \boldsymbol{\Sigma}_{-ii}\\ \boldsymbol{\Sigma}_{i-i} & \Sigma_{ii}\end{pmatrix},$$

then $X_i\sim\mathrm{EC}_1(\mu_i,\Sigma_{ii},h^{(1)})$ with density $f_1(x;\mu_i,\Sigma_{ii},h^{(1)})=h^{(1)}(z_i^2)/\sqrt{\Sigma_{ii}}$, where $z_i=(x-\mu_i)/\sqrt{\Sigma_{ii}}$, and $(\mathbf{X}_{-i}|X_i=x)\sim\mathrm{EC}_{n-1}(\boldsymbol{\mu}_{-i.i}(x),\Sigma_{-i-i.i},h^{(n-1)}_{z_i^2})$, where

$$\boldsymbol{\mu}_{-i.i}(x)=\boldsymbol{\mu}_{-i}+(x-\mu_i)\boldsymbol{\Sigma}_{-ii}/\Sigma_{ii}\quad\text{and}\quad \Sigma_{-i-i.i}=\Sigma_{-i-i}-\boldsymbol{\Sigma}_{-ii}\boldsymbol{\Sigma}_{-ii}^{\mathrm{T}}/\Sigma_{ii}$$

are the conditional location and scale, respectively, and

$$h^{(n-1)}_{z_i^2}(u)=h^{(n)}(u+z_i^2)/h^{(1)}(z_i^2),\quad u\geqslant 0$$

is the conditional density generator. Thus, since

$$F_{\mathbf{X}_{-i}|X_i=x}(x\mathbf{1}_{n-1})=F_{n-1}(x\mathbf{1}_{n-1};\boldsymbol{\mu}_{-i.i}(x),\Sigma_{-i-i.i},h^{(n-1)}_{z_i^2}),$$

we have the following result based on Proposition 1.

Proposition 2. *Let $\mathbf{X}=(X_1,\ldots,X_n)^{\mathrm{T}}$ be a random vector with elliptically contoured distribution, $\mathbf{X}\sim\mathrm{EC}_n(\boldsymbol{\mu},\Sigma,h^{(n)})$. The probability density function $f_{X_{(n)}}$ of $X_{(n)}$ is*

$$f_{X_{(n)}}(x)=\sum_{i=1}^{n} f_1(x;\mu_i,\Sigma_{ii},h^{(1)})F_{n-1}(x\mathbf{1}_{n-1};\boldsymbol{\mu}_{-i.i}(x),\Sigma_{-i-i.i},h^{(n-1)}_{z_i^2}),\quad x\in\mathbb{R},$$

where $z_i=(x-\mu_i)/\sqrt{\Sigma_{ii}}$.

2.2. Exchangeable random variables

When the absolutely continuous random vector $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ is exchangeable, we have for all i that $f_{X_i} = f_{X_n}$ and $F_{\mathbf{X}_{-i}|X_i} = F_{\mathbf{X}_{-n}|X_n}$. Therefore, by Proposition 1, the computation of the probability density function of $X_{(n)}$ simplifies as follows.

Corollary 1. *Let $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ be an absolutely continuous exchangeable random vector. Then, the probability density function $f_{X_{(n)}}$ of $X_{(n)}$ is*

$$f_{X_{(n)}}(x) = nf_{X_n}(x)F_{\mathbf{X}_{-n}|X_n=x}(x\mathbf{1}_{n-1}), \quad x \in \mathbb{R}. \tag{2}$$

When $\mathbf{X} \sim \mathrm{EC}_n(\boldsymbol{\mu}, \Sigma, h^{(n)})$ is an exchangeable elliptically contoured random vector, $\boldsymbol{\mu} = \mu\mathbf{1}_n$ and $\Sigma = \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\}$, where $\rho \in [0, 1)$, implying that $\mu_i = \mu$, $\Sigma_{ii} = \sigma^2$, $\boldsymbol{\mu}_{-i} = \mu\mathbf{1}_{n-1}$, $\boldsymbol{\Sigma}_{-ii} = \sigma^2\rho\mathbf{1}_{n-1}$, $\Sigma_{-i-i} = \sigma^2\{(1-\rho)I_{n-1} + \rho\mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\mathrm{T}}\}$. Therefore, based on Proposition 2 and $z_i = z = (x-\mu)/\sigma$, $\boldsymbol{\mu}_{-i.i}(x) = (\mu + \sigma\rho z)\mathbf{1}_{n-1}$, $\Sigma_{-i-i.i} = \sigma^2(1-\rho)\{I_{n-1} + \rho\mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\mathrm{T}}\}$, we have the following result.

Corollary 2. *Let $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ be an exchangeable random vector with elliptically contoured distribution, $\mathbf{X} \sim \mathrm{EC}_n(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\}, h^{(n)})$, $\rho \in [0, 1)$. The probability density function $f_{X_{(n)}}$ of $X_{(n)}$ is*

$$f_{X_{(n)}}(x) = nf_1(x; \mu, \sigma^2, h^{(1)})F_{n-1}(\sqrt{1-\rho}z\mathbf{1}_{n-1}; \mathbf{0}, I_{n-1} + \rho\mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\mathrm{T}}, h_{z^2}^{(n-1)}), \quad x \in \mathbb{R},$$

where $z = (x-\mu)/\sigma$.

On the other hand, if $\Delta \in \mathbb{R}^{(n-1)\times n}$ denotes the difference matrix such that $\Delta\mathbf{X} = (X_2 - X_1, X_3 - X_2, \ldots, X_n - X_{n-1})^{\mathrm{T}}$, and $\Delta\mathbf{X} \geqslant \mathbf{0}$ means $X_2 - X_1 \geqslant 0, \ldots, X_n - X_{n-1} \geqslant 0$, i.e. $X_1 \leqslant X_2 \leqslant \cdots \leqslant X_n$, then under exchangeability we have

$$P(\Delta\mathbf{X} \geqslant \mathbf{0}) = \frac{1}{n!} \quad \text{and} \quad P(\Delta\mathbf{X} \geqslant \mathbf{0}|X_n = x) = \frac{F_{\mathbf{X}_{-n}}(x\mathbf{1}_{n-1})}{(n-1)!}, \quad x \in \mathbb{R}.$$

From the latter fact, (2) can be rewritten as

$$f_{X_{(n)}}(x) = n!f_{X_n}(x)P(\Delta\mathbf{X} \geqslant \mathbf{0}|X_n = x), \quad x \in \mathbb{R}, \tag{3}$$

which coincides with $f_{X_n|\Delta\mathbf{X}\geqslant\mathbf{0}}(x) = f_{X_n}(x)P(\Delta\mathbf{X} \geqslant \mathbf{0}|X_n = x)/P(\Delta\mathbf{X} \geqslant \mathbf{0})$, $x \in \mathbb{R}$, i.e. the probability density function of X_n given $\Delta\mathbf{X} \geqslant \mathbf{0}$. This implies the following result.

Corollary 3. *Let $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ be an absolutely continuous exchangeable random vector. Then,*

$$X_{(n)} \stackrel{\mathrm{d}}{=} (X_n|\Delta\mathbf{X} \geqslant \mathbf{0}). \tag{4}$$

The result in Corollary 3 means that the distribution of $X_{(n)}$ is intimately related to a specific selection mechanism when $X_1, \ldots, X_n$ are exchangeable random variables. Moreover, when the marginal density f_{X_n} is symmetric, the resulting density (3) is in the FUSS family (1) with $k = 1$ and $m = n-1$. The multiplicative factor $n!P(\Delta\mathbf{X} \geqslant \mathbf{0}|X_n = x)$ in the density (3) is causing skewness in $f_{X_{(n)}}$ although f_{X_n} might be symmetric. If f_{X_n} is not symmetric, the density (3) is still of the form given by the fundamental skew density (1). When $X_1, \ldots, X_n$ are not exchangeable random variables, it follows from the proof of Proposition 1 that the distribution of $X_{(n)}$ can be represented as a finite mixture of the FUS distributions resulting from $(X_i|X_i - X_j \geqslant 0, \ \forall j \neq i)$, $i = 1, \ldots, n$.

An extension of Corollary 3 to linear combinations of order statistics from a random vector with i.i.d. components has recently been derived by Crocetta and Loperfido (2005), who also established a link to FUSS distributions in that setting.

3. Examples

We examine closely the important cases of the multivariate normal and Student t distributions.

3.1. Multivariate normal distribution

The density generator of the multivariate normal distribution is simply given by $h^{(n)}(v) = (2\pi)^{-n/2}\exp(-v/2)$ and $h_a^{(k)}(v) = h^{(k)}(v)$ for all $k = 1, \ldots, n-1$, so that $f_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma, h^{(n)}) = \phi_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$ and $F_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma, h^{(n)}) = \Phi_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$, the multivariate normal probability density and cumulative distribution functions, respectively. Moreover, the conditional density generator $h_a^{(n)}$ is the same as $h^{(n)}$. Therefore, Proposition 2 and Corollary 2 immediately yield the following two important results.

Corollary 4. *Let* $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ *be a random vector with a multivariate normal distribution*, $\mathbf{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \Sigma)$. *The probability density function* $f_{X_{(n)}}$ *of* $X_{(n)}$ *is*

$$f_{X_{(n)}}(x) = \sum_{i=1}^{n} \phi_1(x; \mu_i, \Sigma_{ii})\Phi_{n-1}(x\mathbf{1}_{n-1}; \boldsymbol{\mu}_{-i.i}(x), \Sigma_{-i-i.i}), \quad x \in \mathbb{R}, \tag{5}$$

where ϕ_1 *is the marginal probability density function of* X_n.

Corollary 5. *Let* $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ *be an exchangeable random vector with a multivariate normal distribution*, $\mathbf{X} \sim \mathrm{N}_n(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\})$, $\rho \in [0, 1)$. *The probability density function* $f_{X_{(n)}}$ *of* $X_{(n)}$ *is*

$$f_{X_{(n)}}(x) = n\phi_1(x; \mu, \sigma^2)\Phi_{n-1}(\sqrt{1-\rho}\, z\mathbf{1}_{n-1}; \mathbf{0}, I_{n-1} + \rho\mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\mathrm{T}}), \quad x \in \mathbb{R}, \tag{6}$$

where $z = (x-\mu)/\sigma$ *and* ϕ_1 *is the marginal probability density function of* X_n.

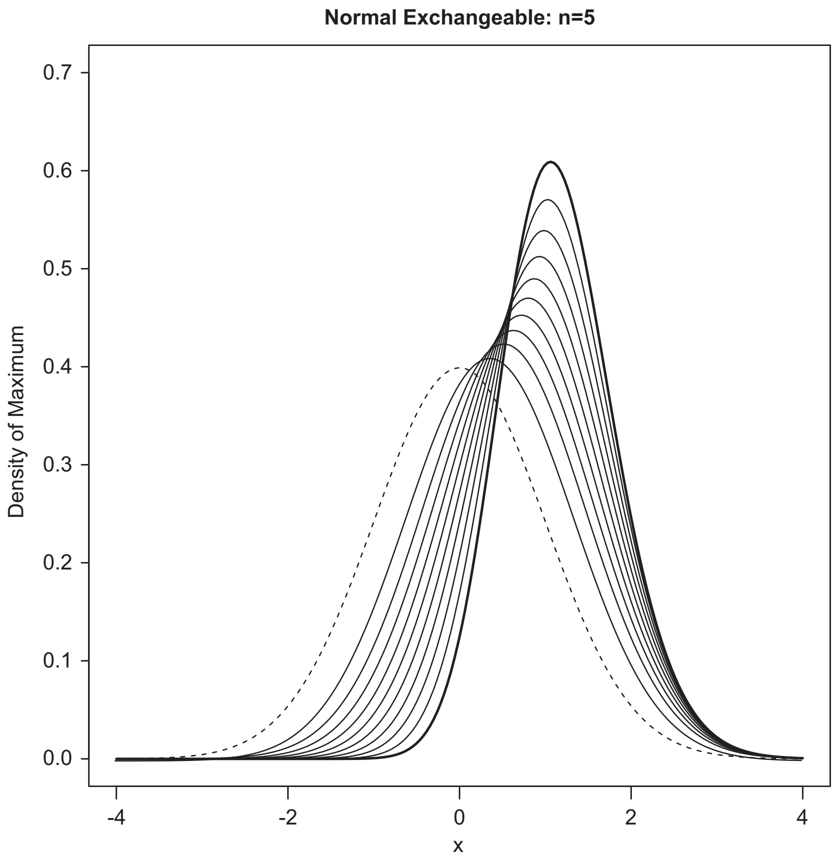

Fig. 1. Density of $X_{(n)}$ for a sample of size $n = 5$ based on an exchangeable standard multivariate normal distribution with correlation $\rho = 0, 0.1, \ldots, 0.9$. The bold curve is the density for $\rho = 0$. The dashed curve is the density for the limiting case $\rho = 1$, and thus also the marginal density of the sample.

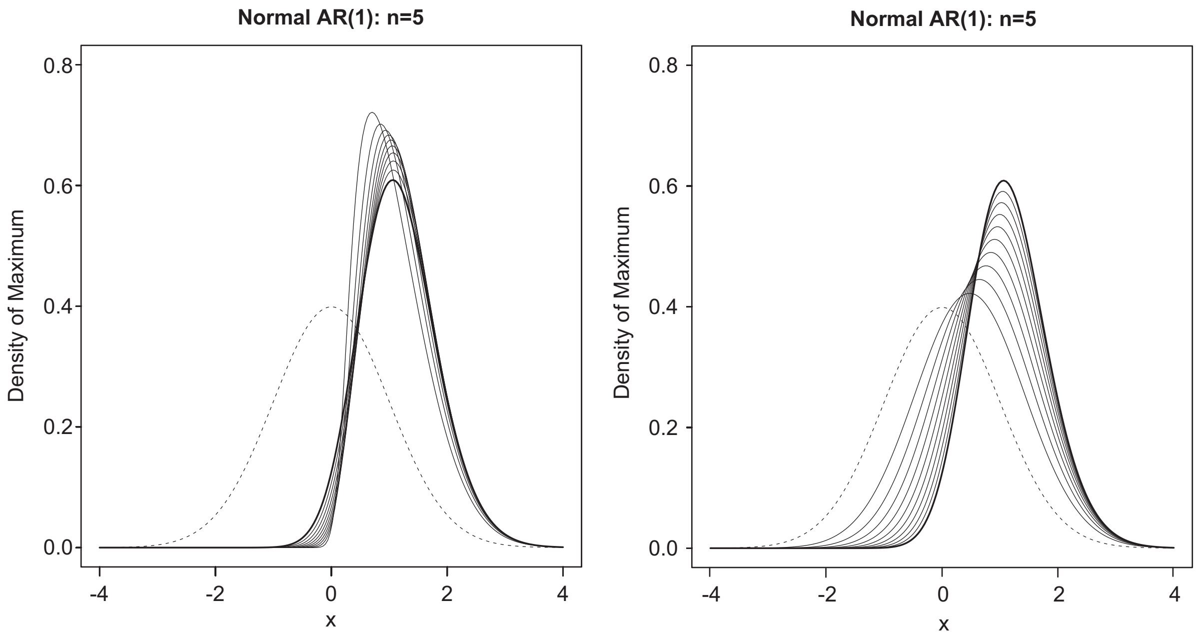

Fig. 2. Density of $X_{(n)}$ for a sample of size $n = 5$ based on a normal AR(1) time series with correlation $\rho = 0, -0.1, \ldots, -0.9$ (left panel) and with correlation $\rho = 0, 0.1, \ldots, 0.9$ (right panel). The bold curve is the density for $\rho = 0$. The dashed curve is the marginal density of the sample.

When $n = 2$, the density (6) reduces to $2\phi_1(x; \mu, \sigma^2)\Phi_1(x; \mu, \sigma^2\{(1+\rho)/(1-\rho)\})$, which is the density of a very particular FUSS distribution called skew-normal, see e.g. the book recently edited by Genton (2004) for a survey. This result has originally been given by Roberts (1966) and recently rediscovered by Loperfido (2002).

Algorithms for numerical evaluation of multivariate normal cumulative distribution functions have been studied by Genz (1992) and made available in the library `mvtnorm` of the statistical software `R` (R Development Core Team, 2004). We make use of this computing power to study the shape of the distribution of $X_{(n)}$ in two settings involving the multivariate normal distribution.

The first setting is the genetic selection problem described in the introduction. We set $n = 5$, $\mu = 0$, $\sigma^2 = 1$, and let $\rho = 0, 0.1, \ldots, 0.9$. Fig. 1 depicts the resulting plots based on the density (6) in Corollary 5. The bold curve is the density for $\rho = 0$. The dashed curve is the density for the limiting case $\rho = 1$, and thus also the marginal density of the sample. We see that the larger the correlation ρ, representing the heritability of the animals, the closer the density is to the standard normal density.

The second setting comes from time series analysis. We consider an autoregressive process of order one, denoted by AR(1), defined at time t by $X_t = \rho X_{t-1} + \varepsilon_t$, where ε_t are i.i.d. $N(0, 1-\rho^2)$ and $\rho \in (-1, 1)$. A sample $\mathbf{X} = (X_1, \ldots, X_n)^T$ from the AR(1) process has therefore a multivariate normal distribution $\mathbf{X} \sim N_n(\mathbf{0}, \Sigma)$, with $\Sigma_{ij} = \rho^{|i-j|}$, $i, j = 1, \ldots, n$. We set $n = 5$ and let $\rho = 0, \pm 0.1, \ldots, \pm 0.9$. Fig. 2 depicts the resulting plots based on the density (5) in Corollary 4, for $\rho \leqslant 0$ (left panel) and for $\rho \geqslant 0$ (right panel). The bold curve is the density for $\rho = 0$. The dashed curve is the marginal density of the sample. We see that the larger the positive correlation ρ, the closer the density is to the standard normal density.

Fig. 3 depicts the density of $X_{(n)}$ for a sample of size $n = 10, 20, 50, 100$ based on an exchangeable standard multivariate normal distribution with $\rho = 0.3$. The dashed curve is the marginal density of the sample.

3.2. Multivariate Student t distribution

The density generator of the Student t distribution with v degrees of freedom is $h^{(n)}(v) = c(n, v)v^{v/2}\{v + v\}^{-(n+v)/2}$, $v \geqslant 0$, with $c(n, v) = \Gamma[(n+v)/2]/(\Gamma[v/2]\pi^{n/2})$ and the conditional density generator is $h_a^{(n-1)}(v) = c(n-1, v+1)(v+a)^{(v+1)/2}\{v + a + v\}^{-(n+v)/2}$. Denote this distribution by $\text{Student}_n(\boldsymbol{\mu}, \Sigma, v)$ with

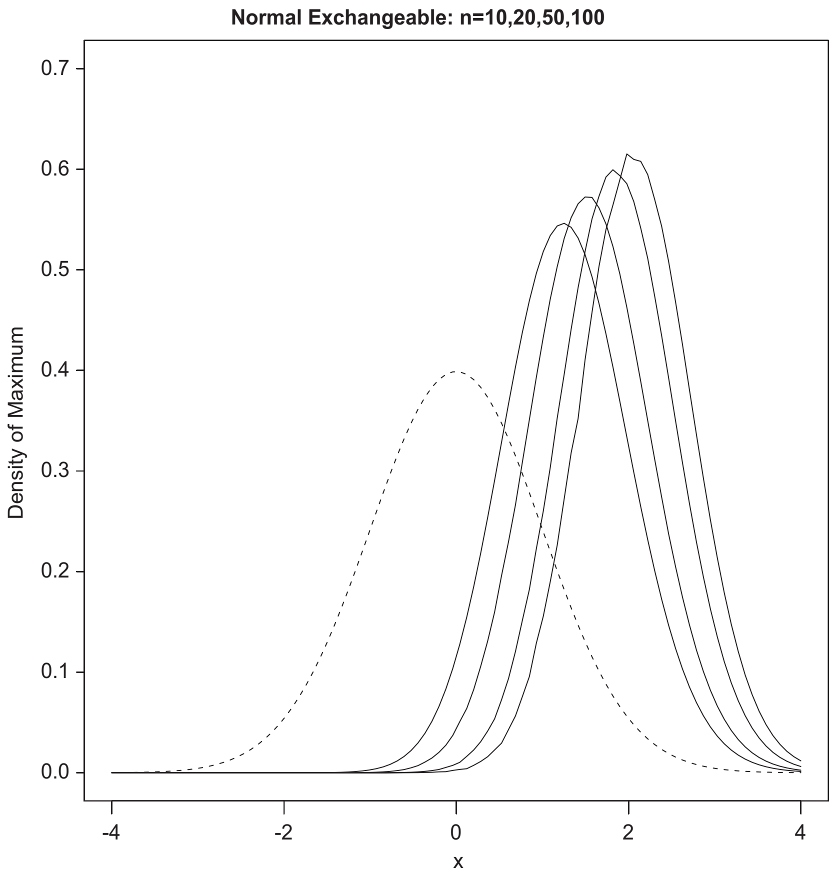

Fig. 3. Density of $X_{(n)}$ for a sample of size $n = 10, 20, 50, 100$ based on an exchangeable standard multivariate normal distribution with $\rho = 0.3$. The dashed curve is the marginal density of the sample.

probability density $t_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma, \nu)$ and cumulative distribution function $T_n(\mathbf{x}; \boldsymbol{\mu}, \Sigma, \nu)$. Proposition 2 and Corollary 2 yield the following results.

Corollary 6. *Let* $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ *be a random vector with a multivariate Student t distribution with* ν *degrees of freedom*, $\mathbf{X} \sim \mathrm{Student}_n(\boldsymbol{\mu}, \Sigma, \nu)$. *The probability density function* $f_{X_{(n)}}$ *of* $X_{(n)}$ *is*

$$f_{X_{(n)}}(x) = \sum_{i=1}^{n} t_1(x; \mu_i, \Sigma_{ii}, \nu) T_{n-1}\left(x\mathbf{1}_{n-1}; \boldsymbol{\mu}_{-i.i}(x), \frac{\nu + z_i^2}{\nu + 1} \Sigma_{-i-i.i}, \nu + 1 \right), \quad x \in \mathbb{R},$$

where t_1 *is the marginal probability density function of* X_n *and* $z_i^2 = (x - \mu_i)^2 / \Sigma_{ii}$.

Corollary 7. *Let* $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ *be an exchangeable random vector with a multivariate Student t distribution with* ν *degrees of freedom*, $\mathbf{X} \sim \mathrm{Student}_n(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\}, \nu)$, $\rho \in [0, 1)$. *The probability density function* $f_{X_{(n)}}$ *of* $X_{(n)}$ *is*

$$f_{X_{(n)}}(x) = nt_1(x; \mu, \sigma^2, \nu) T_{n-1}\left(\sqrt{1-\rho}\, z\mathbf{1}_{n-1}; \mathbf{0}, \frac{\nu + z^2}{\nu + 1}\{I_{n-1} + \rho\mathbf{1}_{n-1}\mathbf{1}_{n-1}^{\mathrm{T}}\}, \nu + 1 \right), \quad x \in \mathbb{R}, \tag{7}$$

where t_1 *is the marginal probability density function of* X_n *and* $z^2 = (x - \mu)^2 / \sigma^2$.

Algorithms for numerical evaluation of multivariate Student t cumulative distribution functions have been studied by Genz and Bretz (2002) and also made available in the statistical software R. Returning to the setting of the genetic selection problem described in the introduction, we investigate the shape of $X_{(n)}$ when $\mathbf{X} \sim \mathrm{Student}_n(\mu\mathbf{1}_n, \sigma^2\{(1-\rho)I_n + \rho\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\}, \nu)$. We set $n = 5$, $\mu = 0$, $\sigma^2 = 1$, $\nu = 3$, and let $\rho = 0, 0.1, \ldots, 0.9$. Fig. 4

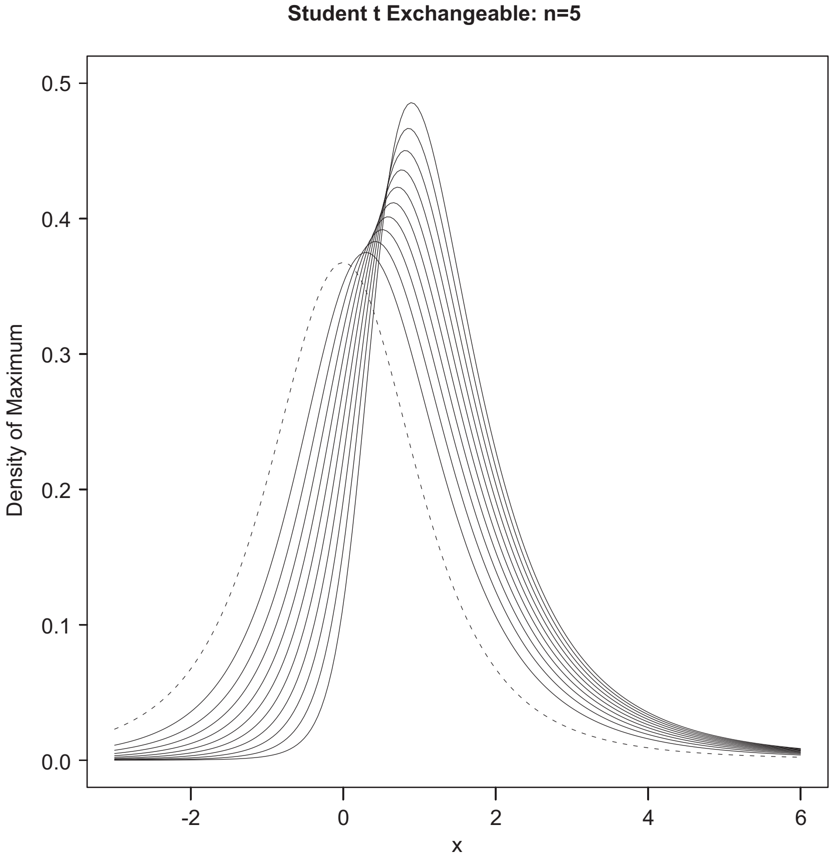

Fig. 4. Density of $X_{(n)}$ for a sample of size $n = 5$ based on an exchangeable standard multivariate Student t distribution with $\nu = 3$ degrees of freedom and correlation $\rho = 0, 0.1, \ldots, 0.9$. The bold curve is the density for $\rho = 0$. The dashed curve is the density for the limiting case $\rho = 1$, and thus also the marginal density of the sample.

depicts the resulting plots based on the density (7) in Corollary 7. The bold curve is the density for $\rho = 0$. The dashed curve is the density for the limiting case $\rho = 1$, and thus also the marginal density of the sample. We see that the larger the correlation ρ, representing the heritability of the animals, the closer the density is to the standard Student t density.

References

Arellano-Valle, R.B., Genton, M.G., 2005. On fundamental skew distributions. J. Multivariate Anal. 96, 93–116.

Arellano-Valle, R.B., Genton, M.G., 2007. On the exact distribution of linear combinations of order statistics from dependent random variables. J. Multivariate Anal. in press.

Arellano-Valle, R.B., Branco, M.D., Genton, M.G., 2006. A unified view on skewed distributions arising from selections. Canad. J. Statist. 34, 581–601.

Crocetta, C., Loperfido, N., 2005. The exact sampling distribution of L-statistics. Metron 63, 1–11.

David, H.A., 1981. Order Statistics. Wiley, New York.

Fang, K.-T., Kotz, S., Ng, K.-W., 1990. Symmetric multivariate and related distributions. Monographs on Statistics and Applied Probability, vol. 36. Chapman & Hall, Ltd., London.

Genton, M.G., 2004. Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Edited volume, Chapman & Hall/CRC Press, London, Boca Raton, FL.

Genz, A., 1992. Numerical computation of multivariate Normal probabilities. J. Comput. Graph Statist. 1, 141–149.

Genz, A., Bretz, F., 2002. Methods for the computation of multivariate t-probabilities. J. Comput. Graph. Statist. 11, 950–971.

Ghosh, M., 1972. Asymptotic properties of linear functions of order statistics for m-dependent random variables. Calcutta Statist. Assoc. Bull. 21, 181–192.

Greig, M., 1967. Extremes in a random assembly. Biometrika 54, 273–282.

Gumbel, E.J., 1958. Statistics of Extremes. Columbia University Press, New York.

Hill, W.G., 1976. Order statistics of correlated variables and implications in genetic selection programmes. Biometrics 32, 889–902.

Hill, W.G., 1977. Order statistics of correlated variables and implications in genetic selection programmes. II. Response to selection. Biometrics 33, 703–712.

Loperfido, N., 2002. Statistical implications of selectively reported inferential results. Statist. Probab. Lett. 56, 13–22.

R Development Core Team, 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL ⟨http://www.R-project.org⟩.

Rawlings, J.O., 1976. Order statistics for a special class of unequally correlated multinormal variates. Biometrics 32, 875–887.

Roberts, C., 1966. A correlation model useful in the study of twins. J. Amer. Statist. Assoc. 61, 1184–1190.

Tong, Y.L., 1990. The Multivariate Normal Distribution. Springer, New York.

Watson, G.S., 1954. Extreme values in samples from *m*-dependent stationary stochastic processes. Ann. Math. Statist. 25, 798–800.