



On the distribution of maximum of multivariate normal random vectors

DOI:

[10.1080/03610926.2018.1465088](https://doi.org/10.1080/03610926.2018.1465088)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Nadarajah, S., Afuecheta, E., & Chan, S. (2018). On the distribution of maximum of multivariate normal random vectors. *Communications in Statistics - Theory and Methods*. <https://doi.org/10.1080/03610926.2018.1465088>

Published in:

Communications in Statistics - Theory and Methods

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



On the distribution of maximum of multivariate normal random vectors

by

Saralees Nadarajah, Emmanuel Afuecheta

School of Mathematics, University of Manchester, Manchester M13 9PL, UK

Stephen Chan

Department of Mathematics and Statistics, American University of Sharjah, UAE

Abstract: Let (X_1, \dots, X_k) be a multivariate normal random vector. We derive explicit expressions for the cumulative distribution function, probability density function and the moments of $\max(X_1, \dots, X_k)$. Each expression involves *single* infinite sums of known special functions. Computational issues like accuracy, convergence, time and simulations are investigated.

Keywords: Maximum; Moments; Multivariate normal distribution

1 Introduction

Let (X_1, \dots, X_k) be a multivariate normal random vector with zero means, unit variances and correlation matrix

$$\mathbf{R} = \begin{pmatrix} \rho_{1,1} & \rho_{1,2} & \cdots & \rho_{1,k} \\ \rho_{2,1} & \rho_{2,2} & \cdots & \rho_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k,1} & \rho_{k,2} & \cdots & \rho_{k,k} \end{pmatrix}.$$

The distribution of $M = \max(X_1, \dots, X_k)$ or equivalently that of $\min(X_1, \dots, X_k)$ and their moments arise in many areas. Some recent examples are:

- Consider a network with n = number of nodes, m = number of links, N = number of nodes where customers reside, and L = number of all links. Assume that the speed of travel on each link is the inverse of a normal random variable and the speeds on different links are correlated. The server chooses the route to the node without knowing what is the speed on each link. According to Berman and Drezner [3], the time taken to reach all nodes is a maximum of correlated normal random variables. Berman and Drezner [3] computed its distribution by a numerical procedure.
- Kala *et al.* [14] developed two preliminary tests for discriminant analysis, that is, tests checking whether a given population say π_0 coincides with populations π_1, \dots, π_k or not. The test statistics involved are expressed in terms of the minimum of k correlated normal random variables. Kala *et al.* [14] proposed an approximation for the distribution of the minimum.
- Let y_{ijk} denote the average QTc response over recording replicates at the k th time interval of the j th subject receiving the i th treatment, where $i = 1$ indicates test drug, $i = 2$ indicates placebo, $i = 1, \dots, n$ and $k = 1, \dots, m$. Cheng *et al.* [6] considered the following model: $y_{ijk} = \mu_{ik} + e_{ij} + \epsilon_{ijk}$, where μ_{ik} is the mean of the i th treatment at the k th time interval, e_{ij} s are independent and normally distributed subject random effects with mean 0 and variance 1, and ϵ_{ijk} 's are independent and identically distributed normal random errors with mean 0 and variance σ^2 . Cheng *et al.* [6] showed that the test of $H_0 : \max(\mu_{11} - \mu_{21}, \dots, \mu_{1m} - \mu_{2m}) \geq 10$ versus $H_1 : \max(\mu_{11} - \mu_{21}, \dots, \mu_{1m} - \mu_{2m}) < 10$ involves the maximum of m correlated normal random variables. Cheng *et al.* [6] proposed an approximation for the distribution of maximum.
- Harper [11] [12] considered theoretical problems involving suprema of Gaussian processes and estimation of Pickands' constant. Both problems required the exact distribution of correlated normal random variables. Harper was able to establish only asymptotic results and approximations.

Some other recent examples are: probabilistic combinatorial optimizations (Bertsimas *et al.* [5]); statistical static time analysis (Eriksson *et al.* [9]).

In each of these examples, only approximations have been used to compute the distribution of maximum/minimum of k correlated normal random variables. If the approximations are not accurate enough serious consequences could result.

- In the example due to Berman and Drezner [3], if the time taken to reach all nodes is computed highly inaccurately that could lead to a meltdown of the network.
- In the example due to Kala *et al.* [14], if the test statistic is computed highly inaccurately that could lead to an incorrect discrimination between the given populations.
- In the example due to Cheng *et al.* [6], if the test statistic is computed highly inaccurately that could lead to an effective drug being labelled as being not good enough.

Hence, the exact computation of the distribution of maximum/minimum of k correlated normal random variables is important. It could avoid serious or even disastrous consequences from happening.

Afonja [1] derived explicit expressions for the first two moments of $M = \max(X_1, \dots, X_k)$ when (X_1, \dots, X_k) is a multivariate normal random vector. Apart from this work, we are aware of no explicit expressions for the moments of M . All of the work we are aware of have presented either approximations, asymptotes or bounds. Berman [4] and Deo [8] considered asymptotes. Cheng *et al.* [6] presented asymptotic approximations. Ross [17] and Chetverikov [7] gave bounds.

However, explicit expressions for the cumulative distribution function (cdf), probability density function (pdf) and the moments of M are known in the bivariate case, see Nadarajah and Kotz [16]. The aim of this paper is to derive explicit expressions for the cdf, pdf and moments of M for the multivariate case. Each of the derived expressions involves *single* infinite sums of known special functions. To the best of our knowledge, these results are new and original.

One may ask: why not compute the distribution of $\max(X_1, \dots, X_k)$ by simulation? Simulations give only approximations, even simulations based on samples of size one million give only approximations, see Figures 3 to 6 later. The exact expressions derived take less time to compute than simulating the distribution based on samples of size one million. The exact expressions can be computed to arbitrary precision by using algebraic manipulation packages like Mathematica. Hence, the use of simulation is computationally less accurate and computationally more expensive.

The derived expressions for the pdf, cdf and moments can have practical use. For example, those for the pdf can be used in maximum likelihood estimation. Those for the cdf can be used in quantile estimation and prediction. Those for moments can be used in moments estimation. Further practical use of the derived expressions is described in Section 3.

The contents of this paper are organized as follows. The explicit expressions for the cdf, pdf and moments of M are derived in Section 2. Some computational issues of the derived expressions are discussed in Section 3. The computer programs for Section 3 were written in Mathematica. The programs and their numerical outputs can be obtained from the corresponding author, email: mbbssn2@manchester.ac.uk Finally, some possible extensions of the results in this paper are discussed in Section 4.

2 Main results

Theorem 1 gives explicit expressions for the cdf and the pdf of M . Theorem 2 gives an explicit expression for the n th moment of M . Each explicit expression involves *single* infinite sums of terms involving the Lauricella function of type A (Exton [10]) and the Lauricella function of type B (Exton [10]). The former is defined by

$$F_A(b_1, \dots, b_n, c_1, \dots, c_n, a; x_1, \dots, x_n) = \sum_{m_1=0}^{\infty} \dots \sum_{m_n=0}^{\infty} \frac{(a)_{m_1+\dots+m_n} (b_1)_{m_1} \dots (b_n)_{m_n}}{(c_1)_{m_1} \dots (c_n)_{m_n}} \frac{x_1^{m_1} \dots x_n^{m_n}}{m_1! \dots m_n!},$$

where $(f)_k = f(f+1) \dots (f+k-1)$ denotes the ascending factorial. The latter is defined by

$$F_B(b_1, \dots, b_n, c_1, \dots, c_n, a; x_1, \dots, x_n) = \sum_{m_1=0}^{\infty} \dots \sum_{m_n=0}^{\infty} \frac{(b_1)_{m_1} \dots (b_n)_{m_n} (c_1)_{m_1} \dots (c_n)_{m_n}}{(a)_{m_1+\dots+m_n}} \frac{x_1^{m_1} \dots x_n^{m_n}}{m_1! \dots m_n!}.$$

In-built routines for computing these special functions are available in Mathematica.

Theorems 1 and 2 require five technical lemmas. Lemma 1 expresses

$$I(n, j, k, a_{i_1}, \dots, a_{i_q}) = \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \Phi^{k-j}(x) \text{He}_{a_{i_1}}(x) \cdots \text{He}_{a_{i_q}}(x) dx, \quad (1)$$

where $\Phi(\cdot)$ denotes the standard normal cdf and $\text{He}_a(\cdot)$ denotes the Hermite polynomial, as a finite sum of terms involving the Lauricella function of type A. Lemma 2 expresses

$$J(n, k, a_1, \dots, a_q) = \int_{-\infty}^{\infty} x^n \exp\left(-\frac{kx^2}{2}\right) \text{He}_{a_1}(x) \cdots \text{He}_{a_q}(x) dx \quad (2)$$

in terms of the Lauricella function of type B. Lemma 3 expresses

$$K(n, k) = \int_{-\infty}^{\infty} x^n \exp\left(-\frac{x^2}{2}\right) \Phi^{k-1}(x) dx \quad (3)$$

as a finite sum of terms involving the Lauricella function of type A. Lemma 4 states a known expression (see equation (45.52) in Kotz *et al.* [15]) for the joint pdf of the multivariate normal distribution. Lemma 5 derives an expression for the corresponding joint cdf.

Lemma 1 *The integral, (1), can be expressed as*

$$\begin{aligned} I(n, j, k, a_{i_1}, \dots, a_{i_q}) &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \left(\frac{2}{j}\right)^{\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2}} \\ &\cdot \sum_{n_{i_1}=0}^{\lfloor a_{i_1}/2 \rfloor} \cdots \sum_{n_{i_q}=0}^{\lfloor a_{i_q}/2 \rfloor} \frac{\left(-\frac{a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(-\frac{a_{i_q}}{2}\right)_{n_{i_q}} \left(\frac{1-a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(\frac{1-a_{i_q}}{2}\right)_{n_{i_q}}}{n_{i_1}! \cdots n_{i_q}!} \\ &\cdot (-j)^{n_{i_1}+\dots+n_{i_q}} \Gamma\left(\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2} - (n_{i_1}+\dots+n_{i_q})\right) \\ &\cdot F_A\left(\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2} - (n_{i_1}+\dots+n_{i_q}), \frac{1}{2}, \dots, \frac{1}{2}, \frac{3}{2}, \dots, \frac{3}{2}; -\frac{1}{j}, \dots, -\frac{1}{j}\right), \end{aligned}$$

where the first summation is over even values of $n+\ell+a_{i_1}+\dots+a_{i_q}$ and $\lfloor x \rfloor$ denotes the largest integer less than or equal to x .

Proof: One of the known series expansions for the Hermite polynomial is

$$\text{He}_a(x) = \sum_{i=0}^{\lfloor a/2 \rfloor} \frac{a!}{i!(a-2i)!} \left(-\frac{1}{2}\right)^i x^{a-2i},$$

which can be rewritten as

$$\text{He}_a(x) = \sum_{i=0}^{\infty} \frac{(-a)_{2i}}{i!} \left(-\frac{1}{2}\right)^i x^{a-2i} = \sum_{i=0}^{\infty} \frac{4^i \left(-\frac{a}{2}\right)_i \left(\frac{1-a}{2}\right)_i}{i!} \left(-\frac{1}{2}\right)^i x^{a-2i}. \quad (4)$$

The first equality in (4) follows because $(-a)_{2i} = a!/(a-2i)!$ for all $i \leq \lfloor a/2 \rfloor$ and $(-a)_{2i} = 0$ for all $i > \lfloor a/2 \rfloor$. The second equality in (4) follows from $(x)_{2n} = 4^n (x/2)_n ((1+x)/2)_n$.

Also one can write

$$\Phi(x) = \frac{1}{2} \left\{ 1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right) \right\},$$

where $\text{erf}(\cdot)$ denotes the error function defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt.$$

One of the known series expansions for the error function is

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{(2m+1)m!}. \quad (5)$$

Using (4) and (5), one can express (1) as

$$\begin{aligned} I(n, j, k, a_{i_1}, \dots, a_{i_q}) &= 2^{j-k} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \left\{1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right\}^{k-j} \operatorname{He}_{a_{i_1}}(x) \cdots \operatorname{He}_{a_{i_q}}(x) dx \\ &= 2^{j-k} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \operatorname{erf}^{\ell}\left(\frac{x}{\sqrt{2}}\right) \operatorname{He}_{a_{i_1}}(x) \cdots \operatorname{He}_{a_{i_q}}(x) dx \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\sqrt{\pi}}\right)^{\ell} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \left\{\sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{2^{m+1/2}(2m+1)m!}\right\}^{\ell} \operatorname{He}_{a_{i_1}}(x) \cdots \operatorname{He}_{a_{i_q}}(x) dx \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \\ &\quad \cdot \sum_{m_1=0}^{\infty} \cdots \sum_{m_{\ell}=0}^{\infty} \frac{(-1/2)^{m_1+\dots+m_{\ell}} x^{2(m_1+\dots+m_{\ell})+\ell}}{(2m_1+1)\cdots(2m_{\ell}+1)m_1!\cdots m_{\ell}!} \operatorname{He}_{a_{i_1}}(x) \cdots \operatorname{He}_{a_{i_q}}(x) dx \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{jx^2}{2}\right) \sum_{m_1=0}^{\infty} \cdots \sum_{m_{\ell}=0}^{\infty} \frac{(-1/2)^{m_1+\dots+m_{\ell}} x^{2(m_1+\dots+m_{\ell})+\ell}}{(2m_1+1)\cdots(2m_{\ell}+1)m_1!\cdots m_{\ell}!} \\ &\quad \cdot \sum_{n_{i_1}=0}^{[a_{i_1}/2]} \frac{4^{n_{i_1}} \left(-\frac{a_{i_1}}{2}\right)_{n_{i_1}} \left(\frac{1-a_{i_1}}{2}\right)_{n_{i_1}}}{n_{i_1}!} \left(-\frac{1}{2}\right)^{n_{i_1}} x^{a_{i_1}-2n_{i_1}} \sum_{n_{i_q}=0}^{[a_{i_q}/2]} \frac{4^{n_{i_q}} \left(-\frac{a_{i_q}}{2}\right)_{n_{i_q}} \left(\frac{1-a_{i_q}}{2}\right)_{n_{i_q}}}{n_{i_q}!} \left(-\frac{1}{2}\right)^{n_{i_q}} x^{a_{i_q}-2n_{i_q}} dx \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \sum_{m_1=0}^{\infty} \cdots \sum_{m_{\ell}=0}^{\infty} \frac{(-1/2)^{m_1+\dots+m_{\ell}}}{(2m_1+1)\cdots(2m_{\ell}+1)m_1!\cdots m_{\ell}!} \\ &\quad \cdot \sum_{n_{i_1}=0}^{[a_{i_1}/2]} \cdots \sum_{n_{i_q}=0}^{[a_{i_q}/2]} \frac{[a_{i_1}/2]! \cdots [a_{i_q}/2]! \left(-\frac{a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(-\frac{a_{i_q}}{2}\right)_{n_{i_q}} \left(\frac{1-a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(\frac{1-a_{i_q}}{2}\right)_{n_{i_q}}}{n_{i_1}! \cdots n_{i_q}!} (-2)^{n_{i_1}+\dots+n_{i_q}} \\ &\quad \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{jx^2}{2}\right) x^{n+2(m_1+\dots+m_{\ell})+\ell+a_{i_1}+\dots+a_{i_q}-2(n_{i_1}+\dots+n_{i_q})} dx \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \left(\frac{2}{j}\right)^{\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2}} \sum_{m_1=0}^{\infty} \cdots \sum_{m_{\ell}=0}^{\infty} \frac{(1/2)_{m_1} \cdots (1/2)_{m_{\ell}}}{(3/2)_{m_1} \cdots (3/2)_{m_{\ell}} m_1! \cdots m_{\ell}!} \\ &\quad \cdot \sum_{n_{i_1}=0}^{[a_{i_1}/2]} \cdots \sum_{n_{i_q}=0}^{[a_{i_q}/2]} \frac{[a_{i_1}/2]! \cdots [a_{i_q}/2]! \left(-\frac{a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(-\frac{a_{i_q}}{2}\right)_{n_{i_q}} \left(\frac{1-a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(\frac{1-a_{i_q}}{2}\right)_{n_{i_q}}}{n_{i_1}! \cdots n_{i_q}!} (-j)^{n_{i_1}+\dots+n_{i_q}-(m_1+\dots+m_{\ell})} \\ &\quad \cdot \Gamma\left(m_1+\dots+m_{\ell}-(n_{i_1}+\dots+n_{i_q})+\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2}\right) \\ &= 2^{j-k} \sum_{\ell=0}^{k-j} \binom{k-j}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \left(\frac{2}{j}\right)^{\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2}} \\ &\quad \cdot \sum_{n_{i_1}=0}^{[a_{i_1}/2]} \cdots \sum_{n_{i_q}=0}^{[a_{i_q}/2]} \frac{[a_{i_1}/2]! \cdots [a_{i_q}/2]! \left(-\frac{a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(-\frac{a_{i_q}}{2}\right)_{n_{i_q}} \left(\frac{1-a_{i_1}}{2}\right)_{n_{i_1}} \cdots \left(\frac{1-a_{i_q}}{2}\right)_{n_{i_q}}}{n_{i_1}! \cdots n_{i_q}!} \\ &\quad \cdot (-j)^{n_{i_1}+\dots+n_{i_q}} \Gamma\left(\frac{n+\ell+a_{i_1}+\dots+a_{i_q}+1}{2}-(n_{i_1}+\dots+n_{i_q})\right) \end{aligned}$$

$$\cdot \sum_{m_1=0}^{\infty} \cdots \sum_{m_\ell=0}^{\infty} \frac{(1/2)_{m_1} \cdots (1/2)_{m_\ell}}{(3/2)_{m_1} \cdots (3/2)_{m_\ell} m_1! \cdots m_\ell!} (-j)^{-(m_1+\cdots+m_\ell)} \\ \cdot \left(\frac{n+\ell+a_{i_1}+\cdots+a_{i_q}+1}{2} - (n_{i_1}+\cdots+n_{i_q}) \right)_{m_1+\cdots+m_\ell}.$$

The result now follows from the definition of the Lauricella function of type A. \square

Lemma 2 The integral, (2), can be expressed as

$$J(n, k, a_1, \dots, a_q) = \left(\frac{2}{k} \right)^{\frac{a_1+\cdots+a_q+n+1}{2}} \Gamma \left(\frac{a_1+\cdots+a_q+n+1}{2} \right) \\ \cdot F_B \left(-\frac{a_1}{2}, \dots, -\frac{a_q}{2}, \frac{1-a_1}{2}, \dots, \frac{1-a_q}{2}, \frac{1-n-a_1-\cdots-a_q}{2}; k, \dots, k \right)$$

for $n+a_1+\cdots+a_k$ even. If $n+a_1+\cdots+a_q$ is odd then $J(n, k, a_1, \dots, a_q) = 0$.

Proof: Using (4), one can express (2) as

$$J(n, k, a_1, \dots, a_q) = \sum_{i_1=0}^{\infty} \cdots \sum_{i_q=0}^{\infty} \frac{4^{i_1+\cdots+i_q} \left(-\frac{a_1}{2}\right)_{i_1} \cdots \left(-\frac{a_q}{2}\right)_{i_q} \left(\frac{1-a_1}{2}\right)_{i_1} \cdots \left(\frac{1-a_q}{2}\right)_{i_q}}{i_1! \cdots i_q!} \\ \cdot \left(-\frac{1}{2}\right)^{i_1+\cdots+i_q} \int_{-\infty}^{\infty} x^{a_1+\cdots+a_q-2(i_1+\cdots+i_q)+n} \exp\left(-\frac{kx^2}{2}\right) dx \\ = \sum_{i_1=0}^{\infty} \cdots \sum_{i_q=0}^{\infty} \frac{4^{i_1+\cdots+i_q} \left(-\frac{a_1}{2}\right)_{i_1} \cdots \left(-\frac{a_q}{2}\right)_{i_q} \left(\frac{1-a_1}{2}\right)_{i_1} \cdots \left(\frac{1-a_q}{2}\right)_{i_q}}{i_1! \cdots i_q!} \left(-\frac{1}{2}\right)^{i_1+\cdots+i_q} \left(\frac{2}{k}\right)^{\frac{a_1+\cdots+a_q-2(i_1+\cdots+i_q)+n+1}{2}} \\ \cdot \Gamma \left(\frac{a_1+\cdots+a_q-2(i_1+\cdots+i_q)+n+1}{2} \right) \\ = \left(\frac{2}{k}\right)^{\frac{a_1+\cdots+a_q+n+1}{2}} \Gamma \left(\frac{a_1+\cdots+a_q+n+1}{2} \right) \sum_{i_1=0}^{\infty} \cdots \sum_{i_q=0}^{\infty} \frac{\left(-\frac{a_1}{2}\right)_{i_1} \cdots \left(-\frac{a_q}{2}\right)_{i_q} \left(\frac{1-a_1}{2}\right)_{i_1} \cdots \left(\frac{1-a_q}{2}\right)_{i_q} k^{i_1+\cdots+i_q}}{\left(\frac{1-n-a_1-\cdots-a_q}{2}\right)_{i_1+\cdots+i_q} i_1! \cdots i_q!}.$$

The result now follows from the definition of the Lauricella function of type B. \square

Lemma 3 The integral, (3), can be expressed as

$$K(n, k) = 2^{\frac{n+3}{2}-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{4}{\pi}\right)^{\ell/2} \Gamma\left(\frac{\ell+n+1}{2}\right) F_A\left(\frac{\ell+n+1}{2}, \frac{1}{2}, \dots, \frac{1}{2}, \frac{3}{2}, \dots, \frac{3}{2}; -1, \dots, -1\right),$$

where the first summation is over all even values of $\ell+n$.

Proof: Using (5), one can express (3) as

$$K(n, k) = 2^{1-k} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{x^2}{2}\right) \left\{1 + \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right\}^{k-1} dx \\ = 2^{1-k} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{x^2}{2}\right) \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \operatorname{erf}^{\ell}\left(\frac{x}{\sqrt{2}}\right) dx \\ = 2^{1-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{2}{\sqrt{\pi}}\right)^{\ell} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{x^2}{2}\right) \left\{\sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{2^{m+1/2} (2m+1)m!}\right\}^{\ell} dx \\ = 2^{1-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \int_{-\infty}^{\infty} x^n \exp\left(-\frac{x^2}{2}\right) \sum_{m_1=0}^{\infty} \cdots \sum_{m_\ell=0}^{\infty} \frac{(-1/2)^{m_1+\cdots+m_\ell} x^{2(m_1+\cdots+m_\ell)+\ell}}{(2m_1+1)\cdots(2m_\ell+1)m_1!\cdots m_\ell!} dx$$

$$\begin{aligned}
&= 2^{1-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \sum_{m_1=0}^{\infty} \cdots \sum_{m_\ell=0}^{\infty} \frac{(-1/2)^{m_1+\cdots+m_\ell} (1/2)_{m_1} \cdots (1/2)_{m_\ell}}{(3/2)_{m_1} \cdots (3/2)_{m_\ell} m_1! \cdots m_\ell!} \\
&\quad \cdot \int_{-\infty}^{\infty} x^{2(m_1+\cdots+m_\ell)+\ell+n} \exp\left(-\frac{x^2}{2}\right) dx \\
&= 2^{1-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{2}{\pi}\right)^{\ell/2} \sum_{m_1=0}^{\infty} \cdots \sum_{m_\ell=0}^{\infty} \frac{(-1/2)^{m_1+\cdots+m_\ell} (1/2)_{m_1} \cdots (1/2)_{m_\ell}}{(3/2)_{m_1} \cdots (3/2)_{m_\ell} m_1! \cdots m_\ell!} 2^{m_1+\cdots+m_\ell+\frac{\ell+n+1}{2}} \\
&\quad \cdot \Gamma\left(m_1+\cdots+m_\ell+\frac{\ell+n+1}{2}\right) \\
&= 2^{\frac{n+3}{2}-k} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} \left(\frac{4}{\pi}\right)^{\ell/2} \Gamma\left(\frac{\ell+n+1}{2}\right) \\
&\quad \cdot \sum_{m_1=0}^{\infty} \cdots \sum_{m_\ell=0}^{\infty} \frac{(-1)^{m_1+\cdots+m_\ell} (1/2)_{m_1} \cdots (1/2)_{m_\ell}}{(3/2)_{m_1} \cdots (3/2)_{m_\ell} m_1! \cdots m_\ell!} \left(\frac{\ell+n+1}{2}\right)_{m_1+\cdots+m_\ell}.
\end{aligned}$$

The result now follows from the definition of the Lauricella function of type A. \square

Lemma 4 Let \mathbf{X} denote a $k \times 1$ normal random vector with zero means, unit variances and correlation matrix $\mathbf{R} = (\rho_{\alpha,\beta})$. Its joint pdf, say $\phi_k(\mathbf{x})$, can be expanded as

$$\phi_k(\mathbf{x}) = \left[\prod_{j=1}^k \phi(x_j) \right] \sum_{j=0}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha,\beta} \rho_{\alpha,\beta} \prod_{t=1}^k \text{He}_{i_t}(x_t), \quad (6)$$

where $\phi(\cdot)$ denotes the standard normal pdf, \sum_i^* is a sum over all possible sets of j $\rho_{\alpha,\beta}$'s including repeated values, C_i is the number of different permutations of the $\rho_{\alpha,\beta}$'s (that is, $j! \left(\prod_{\ell=1}^m j_\ell!\right)^{-1}$, where the same ρ is repeated j_1, \dots, j_m times with $j_1 + \cdots + j_m = j$ for $\alpha < \beta$), and i_t is the number of times t occurs among the suffices α, β in the i th term of \sum_i^* .

Proof: Equation (6) is equation (45.52) in Kotz *et al.* [15]. \square

Lemma 5 Let \mathbf{X} denote a $k \times 1$ normal random vector with zero means, unit variances and correlation matrix $\mathbf{R} = (\rho_{\alpha,\beta})$. Its joint cdf, say $\Phi_k(\mathbf{x})$, can be expanded as

$$\Phi_k(\mathbf{x}) = \prod_{j=1}^k \Phi(x_j) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha,\beta} \rho_{\alpha,\beta} \left[\prod_{\ell=1}^{k-m} \Phi(x_{s_\ell}) \right] \left\{ \prod_{\ell=1}^m [-\phi(x_{t_\ell}) \text{He}_{i_{t_\ell}-1}(x_{t_\ell})] \right\}, \quad (7)$$

where i_{t_1}, \dots, i_{t_m} are the positive values among i_1, \dots, i_k while $i_{s_1}, \dots, i_{s_{k-m}}$ are the zero values among i_1, \dots, i_k . The dependence of m on i, j, α, β is suppressed. The summations and products in (7) are as described in Lemma 4.

Proof: Equation (7) follows by integrating (6) term wise and using the facts $\text{He}_0(x) = 1$ and $\int \exp(-x^2/2) \text{He}_n(x) dx = -\exp(-x^2/2) \text{He}_{n-1}(x)$. \square

Theorem 1 Let \mathbf{X} denote a $k \times 1$ normal random vector with zero means, unit variances and correlation matrix $\mathbf{R} = (\rho_{\alpha,\beta})$. Let $M = \max(X_1, \dots, X_k)$. The cdf of M can be expressed as

$$F_M(x) = \Phi^k(x) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha,\beta} \rho_{\alpha,\beta} (-1)^m \Phi^{k-m}(x) \phi^m(x) \prod_{\ell=1}^m \text{He}_{i_{t_\ell}-1}(x). \quad (8)$$

The corresponding pdf of M can be expressed as

$$f_M(x) = k\Phi^{k-1}(x) \phi(x) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha,\beta} \rho_{\alpha,\beta} (-1)^m (k-m) \Phi^{k-m-1}(x) \phi^{m+1}(x) \prod_{\ell=1}^m \text{He}_{i_{t_\ell}-1}(x)$$

$$-\sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} (-1)^m \Phi^{k-m}(x) \phi^m(x) \sum_{p=1}^m \text{He}_{i_{t_p}}(x) \prod_{\ell \neq p} \text{He}_{i_{t_\ell}-1}(x). \quad (9)$$

The summations and products in (8) and (9) are as described in Lemmas 4 and 5.

Proof: Equation (8) follows by setting each element of \mathbf{x} in (7) to be equal to x . Equation (9) follows by differentiating (8) with respect to x and using the fact $\partial \text{He}_a(x)/\partial x = x \text{He}_a(x) - \text{He}_{a+1}(x)$. \square

Theorem 2 Let \mathbf{X} denote a $k \times 1$ normal random vector with zero means, unit variances and correlation matrix $\mathbf{R} = (\rho_{\alpha, \beta})$. Let $M = \max(X_1, \dots, X_k)$. The n th moment of M can be expressed as

$$\begin{aligned} \mathbb{E}(M^n) &= (2\pi)^{-\frac{1}{2}} k K(n, k) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} (-1)^m (k-m) (2\pi)^{-\frac{m+1}{2}} I(n, m+1, k, i_{t_1}-1, \dots, i_{t_m}-1) \\ &- \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} (-1)^m (2\pi)^{-\frac{m}{2}} \sum_{p=1}^m I(n, m, k, i_{t_1}-1, \dots, i_{t_{p-1}}-1, i_{t_p}, i_{t_{p+1}}-1, \dots, i_{t_m}-1) \end{aligned} \quad (10)$$

for $n \geq 1$, where the summations and products are as described in Lemmas 4 and 5. The $I(n, m+1, k, i_{t_1}-1, \dots, i_{t_m}-1)$ in the second term can be replaced by $J(n, m+1, i_{t_1}-1, \dots, i_{t_m}-1)$ if $m = k-1$. The $I(n, m, k, i_{t_1}-1, \dots, i_{t_{p-1}}-1, i_{t_p}, i_{t_{p+1}}-1, \dots, i_{t_m}-1)$ in the third term can be replaced by $J(n, m, i_{t_1}-1, \dots, i_{t_{p-1}}-1, i_{t_p}, i_{t_{p+1}}-1, \dots, i_{t_m}-1)$ if $m = k$.

Proof: Using (9), we can write

$$\begin{aligned} \mathbb{E}(M^n) &= k \int_{-\infty}^{\infty} x^n \Phi^{k-1}(x) \phi(x) dx + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} (-1)^m (k-m) \\ &\cdot \int_{-\infty}^{\infty} x^n \Phi^{k-m-1}(x) \phi^{m+1}(x) \prod_{\ell=1}^m \text{He}_{i_{t_\ell}-1}(x) dx \\ &- \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} (-1)^m \sum_{p=1}^m \int_{-\infty}^{\infty} x^n \Phi^{k-m}(x) \phi^m(x) \text{He}_{i_{t_p}}(x) \prod_{\ell \neq p} \text{He}_{i_{t_\ell}-1}(x) dx. \end{aligned}$$

The result now follows by applying Lemmas 1 to 3 to calculate the integrals. \square

At first sight Theorems 1 and 2 look complicated. But they can be interpreted to correspond to different probability structures of (X_1, \dots, X_k) . For example, (8) can be interpreted as follows: $\Phi^k(x)$ arises whenever the components of (X_1, \dots, X_k) are mutually independent; the terms containing $\Phi^{k-1}(x)\phi(x)$ arise whenever $(k-1)$ components of (X_1, \dots, X_k) are mutually independent and the remaining component is dependent on them; the terms containing $\Phi^{k-2}(x)\phi^2(x)$ arise whenever $(k-2)$ components of (X_1, \dots, X_k) are mutually independent and the remaining two components are dependent on them; and so on. To be more explicit, suppose a system has k components and (X_1, \dots, X_k) denote their lifetimes. Then the largest of the lifetimes can result if one of the following happens: the components work mutually independently (giving the term $\Phi^k(x)$); $(k-1)$ of the components work mutually independently and the remaining component is dependent on them (giving the terms containing $\Phi^{k-1}(x)\phi(x)$); $(k-2)$ of the components work mutually independently and the remaining two components are dependent on them (giving the terms containing $\Phi^{k-2}(x)\phi^2(x)$); and so on. Theorem 2 can be interpreted similarly.

3 Computational issues

Here, we compute some of the results in Section 2. In particular, we compute the pdf of M given by (9). We also compute the mean, variance, skewness and kurtosis of M given by (10). All of the computations were performed using Mathematica.

Throughout, we suppose that $\rho_{\alpha, \beta} = \rho$ for all $1 \leq \alpha \neq \beta \leq k$. In this case, \mathbf{R} is positive definite if and only if $\rho \in (-1/(k-1), 1)$. The values of ρ for our computations will be limited to this range although we shall also consider the case $\rho = 1$.

The results of Theorems 1 and 2 hold for any valid correlation structure, not just for $\rho_{\alpha, \beta} = \rho$ for all $1 \leq \alpha \neq \beta \leq k$. We suppose this correlation structure only for simplicity.

Both (9) and (10) are *single* infinite sums (the infinite summation being with respect to j) of known special functions. In our computations, we truncated the infinite sum with respect to j at twenty. This gave sufficient accuracy as explained later. The Lauricella function of type A and the Lauricella function of type B were computed using in-built routines in Mathematica.

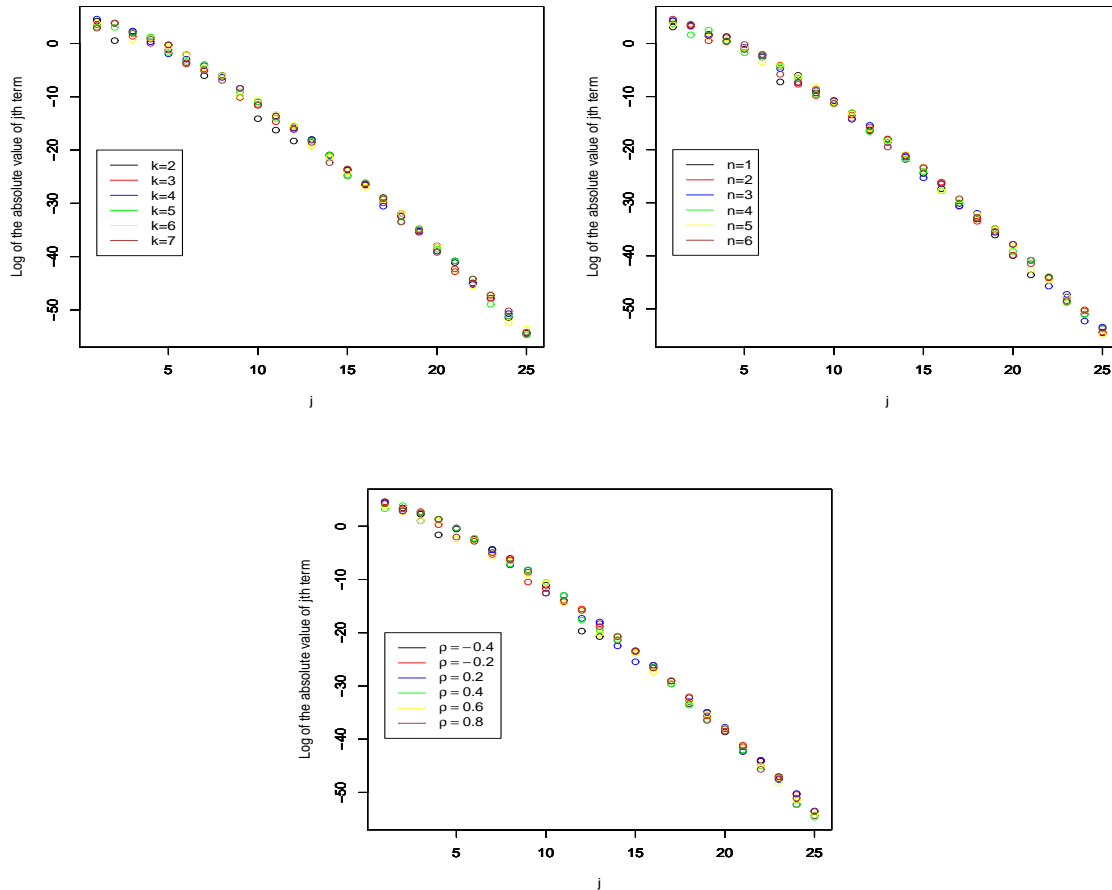


Figure 1: The logarithm of the absolute value of the j th term in (10) versus j for $k = 2, 3, 4, 5, 6, 7$, $n = 2$ and $\rho = 0.5$ (top left); The logarithm of the absolute value of the j th term in (10) versus j for $k = 3$, $n = 1, 2, 3, 4, 5, 6$ and $\rho = 0.5$ (top right); The logarithm of the absolute value of the j th term in (10) versus j for $k = 3$, $n = 2$ and $\rho = -0.4, -0.2, 0.2, 0.4, 0.6, 0.8$ (bottom left).

The logarithm of the absolute value of the j th term in (10) versus j for $k = 2, 3, 4, 5, 6, 7$ is plotted in Figure 1. The logarithm of the absolute value of the j th term in (10) versus j for $n = 1, 2, 3, 4, 5, 6$ is also plotted in Figure 1. The logarithm of the absolute value of the j th term in (10) versus j for $\rho = -0.4, -0.2, 0.2, 0.4, 0.6, 0.8$ is also plotted in Figure 1. The magnitude of the j th term decreases sharply with respect to j . The decreases appears to be $C/j!$ for large j . In particular, truncation at twenty can be expected to give good accuracy. However, there is no evidence of significant changes in the magnitude of the j th term with respect to n , k and ρ ; so, the point of truncation can be expected not to change significantly with respect to n , k and ρ . Also the sign of the j th term appeared to alternative with respect to j over the plotted parameters, although this is not evident from the figures. The observations were similar when the absolute value of the j th term in (8) was plotted versus j and when the absolute value of the j th term in (9) was plotted versus j .

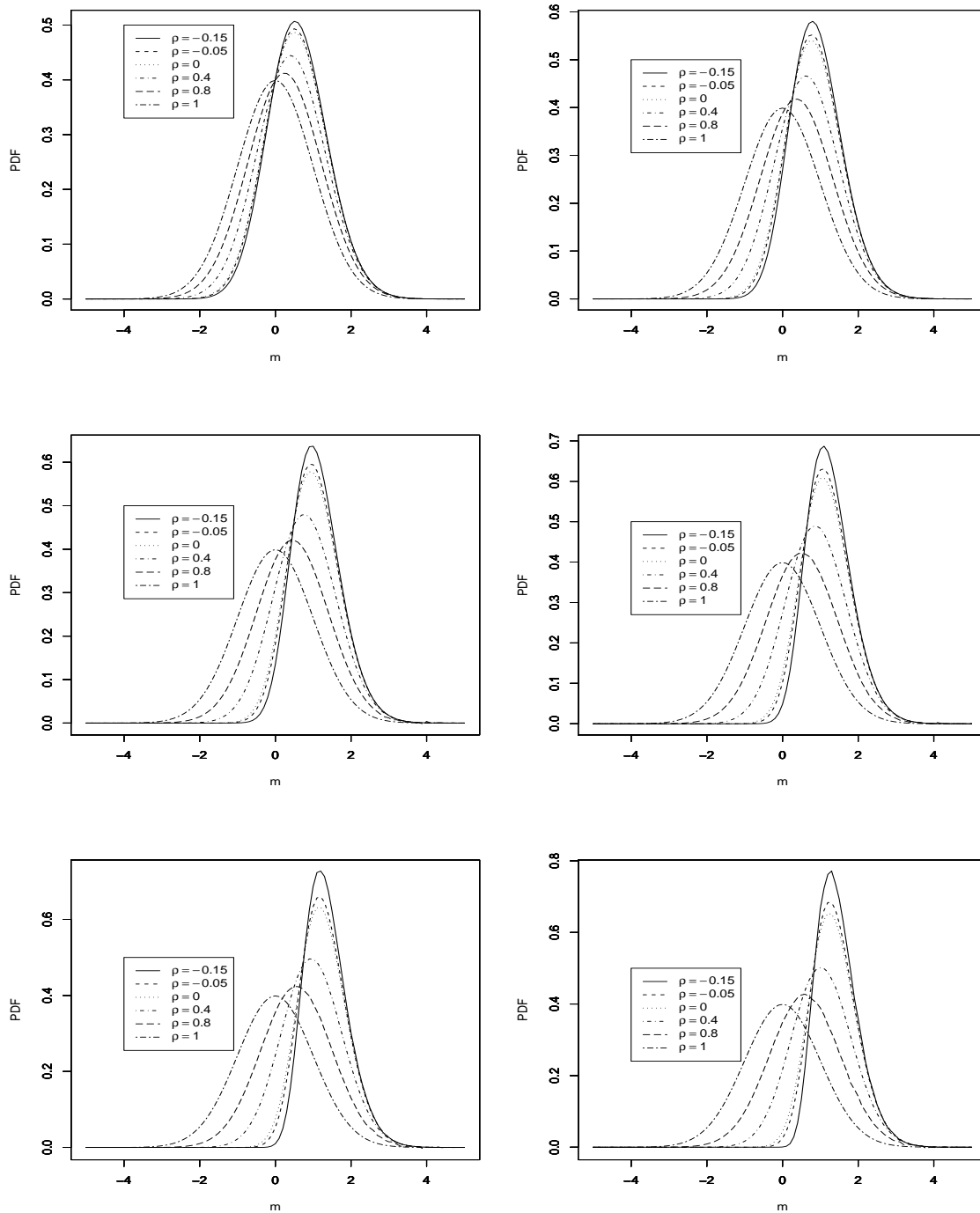


Figure 2: Pdfs of $M = \max(X_1, \dots, X_k)$ when $(X_1, \dots, X_k)^T$ is k -variate normal with zero means, unit variances and all correlations equal to ρ . Top left is $k = 2$, top right is $k = 3$, middle left is $k = 4$, middle right is $k = 5$, bottom left is $k = 6$ and bottom right is $k = 7$.

Plots of (9) for $k = 2, 3, 4, 5, 6, 7$ and for $\rho = -0.15, -0.05, 0, 0.4, 0.8, 1$ are shown in Figure 2. Note that the pdf of M is $f_M(x) = 2\Phi(\sqrt{1 - \rho x / \sqrt{1 + \rho}})\phi(x)$ when $k = 2$, see Nadarajah and Kotz [16]. The pdf of M is $f_M(x) = k\Phi^{k-1}(x)\phi(x)$ when $\rho = 0$. The pdf of M is the standard normal pdf when $\rho = 1$. The latter pdf was not computed using (9). But its proximity to the pdfs for $\rho = -0.15, -0.05, 0, 0.4, 0.8$ gives us confidence on mathematical

correctness of (9). The pdfs were also computed using the standard approach; that is, using

$$f_M(x) = \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \frac{\partial}{\partial x} \left[\int_{-\infty}^x \int_{-\infty}^x \cdots \int_{-\infty}^x \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) dx_k \cdots dx_2 dx_1 \right]. \quad (11)$$

These values were plotted on top of the values computed using (9). Since we see no difference, we have further confidence on mathematical correctness of (9).

The pdfs of M in Figure 2 appear symmetric for ρ close to one. They become more and more skewed with decreasing values of ρ . They also become more and more peaked with decreasing values of ρ . Their lower tails become lighter with decreasing values of ρ . Their upper tails become heavier with decreasing values of ρ .

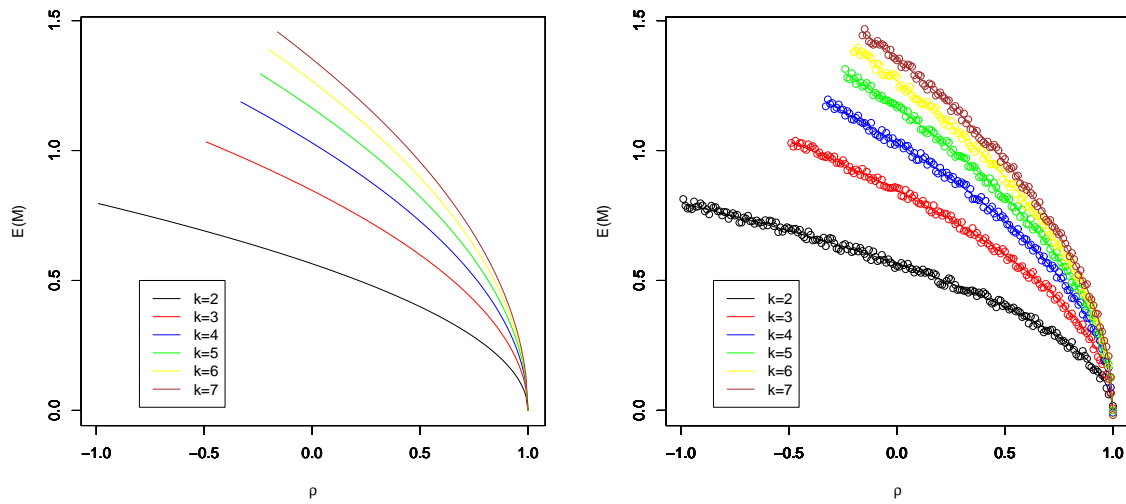


Figure 3: The mean of $\max(X_1, \dots, X_k)$ versus ρ and k when $(X_1, \dots, X_k)^T$ is k -variate normal with zero means, unit variances and all correlations equal to ρ . The left hand side plot is based on (10) and the right hand side plot is based on simulations.

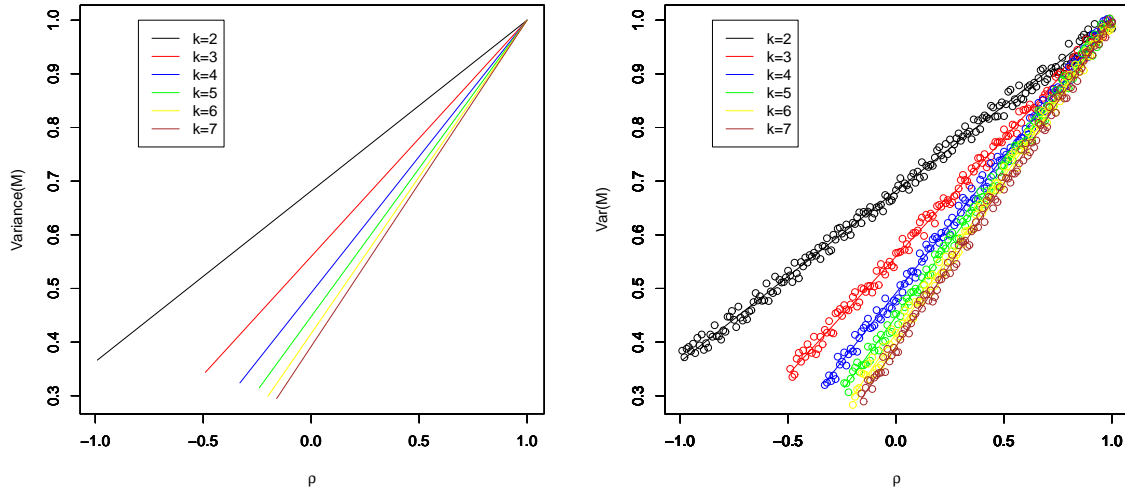


Figure 4: The variance of $\max(X_1, \dots, X_k)$ versus ρ and k when $(X_1, \dots, X_k)^T$ is k -variate normal with zero means, unit variances and all correlations equal to ρ . The left hand side plot is based on (10) and the right hand side plot is based on simulations.

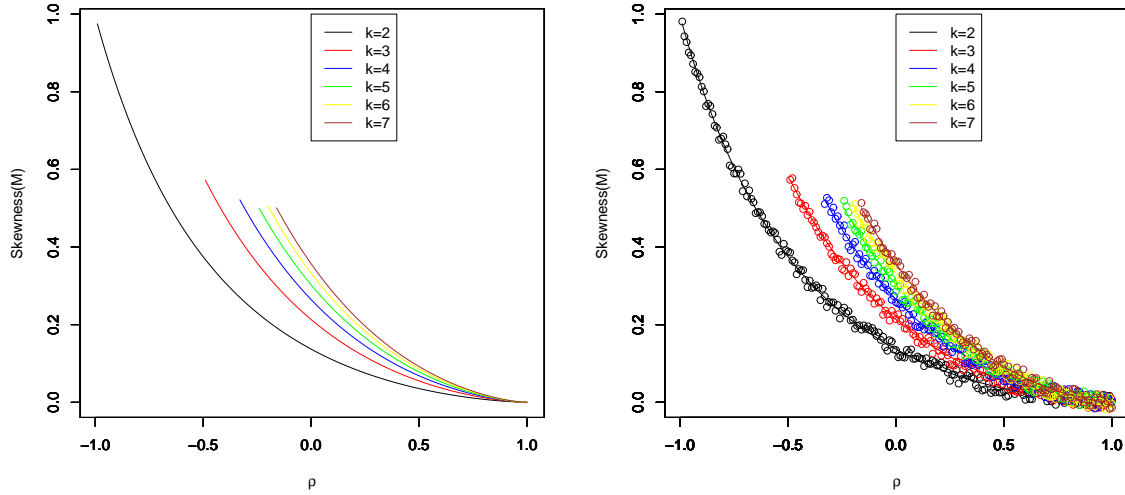


Figure 5: The skewness of $\max(X_1, \dots, X_k)$ versus ρ and k when $(X_1, \dots, X_k)^T$ is k -variate normal with zero means, unit variances and all correlations equal to ρ . The left hand side plot is based on (10) and the right hand side plot is based on simulations.

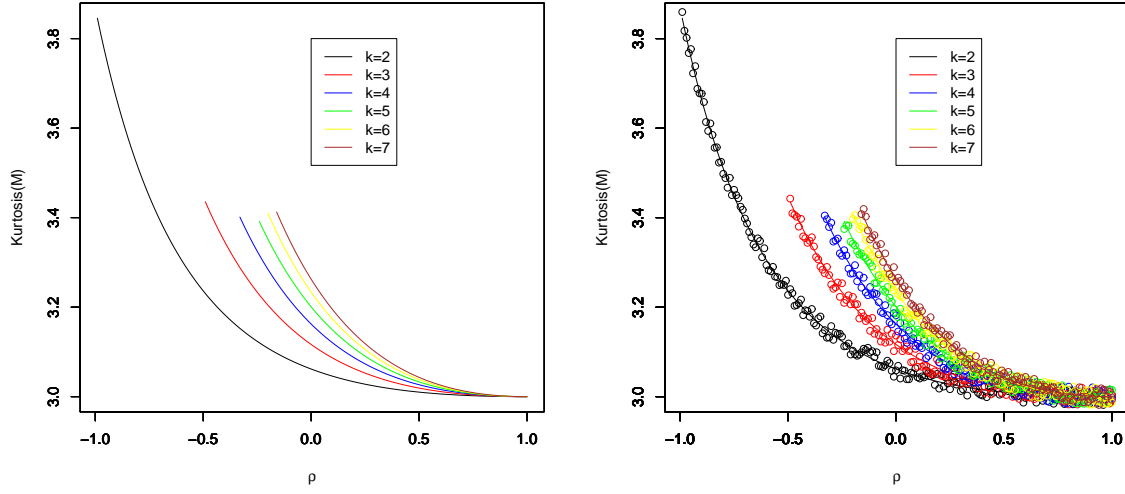


Figure 6: The kurtosis of $\max(X_1, \dots, X_k)$ versus ρ and k when $(X_1, \dots, X_k)^T$ is k -variate normal with zero means, unit variances and all correlations equal to ρ . The left hand side plot is based on (10) and the right hand side plot is based on simulations.

Plots of the mean of M versus $\rho \in (-1/(k-1), 1]$ for $k = 2, 3, 4, 5, 6, 7$ are shown in Figure 3. Plots of the variance of M versus $\rho \in (-1/(k-1), 1]$ for $k = 2, 3, 4, 5, 6, 7$ are shown in Figure 4. Plots of the skewness of M versus $\rho \in (-1/(k-1), 1]$ for $k = 2, 3, 4, 5, 6, 7$ are shown in Figure 5. Here, skewness of M was computed using the formula $E[(M - \mu)^3] / [\text{Var}(M)]^{3/2}$, where $\mu = E(M)$. Plots of the kurtosis of M versus $\rho \in (-1/(k-1), 1]$ for $k = 2, 3, 4, 5, 6, 7$ are shown in Figure 6. Here, kurtosis of M was computed using the formula $E[(M - \mu)^4] / [\text{Var}(M)]^2$. The mean, variance, skewness and kurtosis of M for $\rho < 1$ were computed using (10). The mean, variance, skewness and kurtosis of M for $\rho = 1$ were computed using the fact that the pdf of M is the standard normal pdf. Also shown in Figures 3 to 6 are the mean, variance, skewness and kurtosis based on a sample of size one million simulated from the multivariate normal distribution. The closeness of the simulated values as well as the values for $\rho = 1$ with those computed using (10) gives us confidence on mathematical correctness of the latter. The moments were also computed using the standard approach; that is, using

$$E(M^n) = \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \int_{-\infty}^{\infty} x^n \frac{\partial}{\partial x} \left[\int_{-\infty}^x \int_{-\infty}^x \cdots \int_{-\infty}^x \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) dx_k \cdots dx_2 dx_1 \right] dx. \quad (12)$$

These values were plotted on top of the values computed using (10). Since we see no difference, we have further confidence on mathematical correctness of (10).

Figure 3 shows that the mean of M is a monotonic decreasing function of ρ and a monotonic increasing function of k . All values of the mean are greater than zero, the mean of a standard normal distribution.

Figure 4 shows that the variance of M is a monotonic increasing function of ρ and a monotonic decreasing function of k . The increase appears linear. All values of the variance are less than one, the variance of a standard normal distribution.

To investigate the linearity, we fitted a linear regression model to each curve in Figure 4. We obtained the following fitted lines: $\text{var}(M) = 0.6816901 + 0.3183099\rho$ for $k = 2$, $\text{var}(M) = 0.5594672 + 0.4405328\rho$ for $k = 3$, $\text{var}(M) = 0.4917152 + 0.5082848\rho$ for $k = 4$, $\text{var}(M) = 0.4475341 + 0.5524659\rho$ for $k = 5$, $\text{var}(M) = 0.4159271 + 0.5840729\rho$ for $k = 6$, and $\text{var}(M) = 0.3919178 + 0.6080822\rho$ for $k = 7$. The residuals of the fitted regression lines were identically zero up to seven decimal places.

Figure 5 shows that the skewness of M is a monotonic decreasing function of ρ and a monotonic increasing function of k . All values of skewness are greater than zero, the skewness of a standard normal distribution.

Figure 6 shows that the kurtosis of M is a monotonic decreasing function of ρ and a monotonic increasing function of k . All values of kurtosis are greater than three, the kurtosis of a standard normal distribution.

We now compare the computational performances of (8), (9) and (10) versus the standard approach. We use central processing unit time as the criterion for assessing performance. Computational accuracy is not an issue since Mathematica like most other symbolic manipulation packages allows for arbitrary precision.

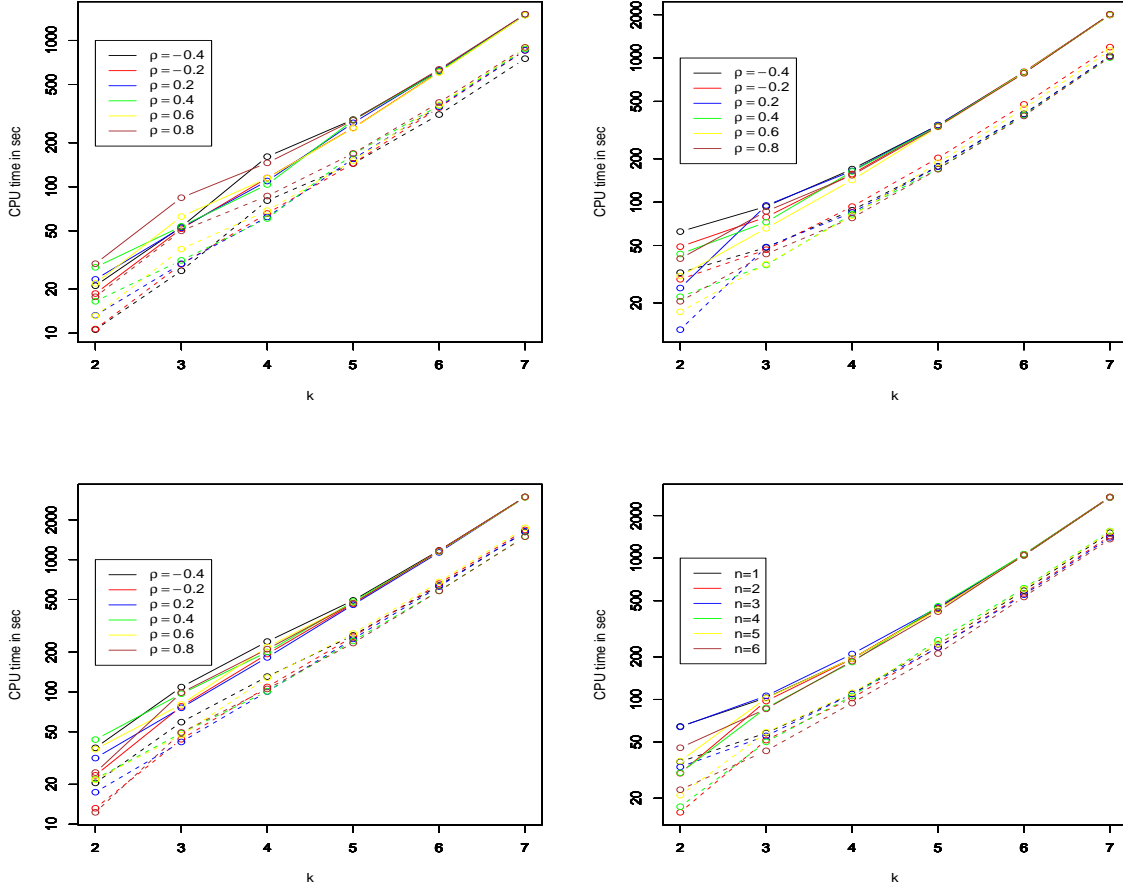


Figure 7: Top left: central processing times taken to compute the cdf of $\max(X_1, \dots, X_k)$ one hundred times using (13) (solid curves) and (8) (broken curves); Top right: central processing times taken to compute the pdf of $\max(X_1, \dots, X_k)$ one hundred times using (11) (solid curves) and (9) (broken curves); Bottom left: central processing times taken to compute the n th moment of $\max(X_1, \dots, X_k)$ one hundred times using (12) (solid curves) and (10) (broken curves) when $n = 3$; Bottom right: central processing times taken to compute the n th moment of $\max(X_1, \dots, X_k)$ one hundred times using (12) (solid curves) and (10) (broken curves) when $\rho = 0.2$. The y axes are in log scale.

Figure 7 compares the central processing unit times taken to compute the cdf of M using (8) and the standard approach given by

$$F_M(x) = \frac{1}{(2\pi)^{k/2} |\mathbf{R}|^{1/2}} \left[\int_{-\infty}^x \int_{-\infty}^x \cdots \int_{-\infty}^x \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) dx_k \cdots dx_2 dx_1 \right]. \quad (13)$$

Figure 7 also compares the central processing unit times taken to compute the pdf of M using (9) and the standard approach given by (11). Figure 7 also compares the central processing unit times taken to compute the n th moment of M using (10) and the standard approach given by (12). The central processing unit times are for one hundred computations of the cdf at $x = 1$, one hundred computations of the pdf at $x = 1$, and one hundred computations of the moments.

The central processing unit time taken to compute the pdf of M appears to increase exponentially with respect to k . The central processing unit time taken to compute the cdf of M also appears to increase exponentially with respect to k . The central processing unit time taken to compute the moments of M also appears to increase exponentially with respect to k . There appears to be no significant changes in the central processing unit time with respect to n or ρ . The central processing unit times for (8) as compared to (13) appear to have been nearly halved. Similarly, the central processing unit times for (9) as compared to (11) appear to have been nearly halved. Similarly, the central processing unit times for (10) as compared to (12) appear to have been nearly halved.

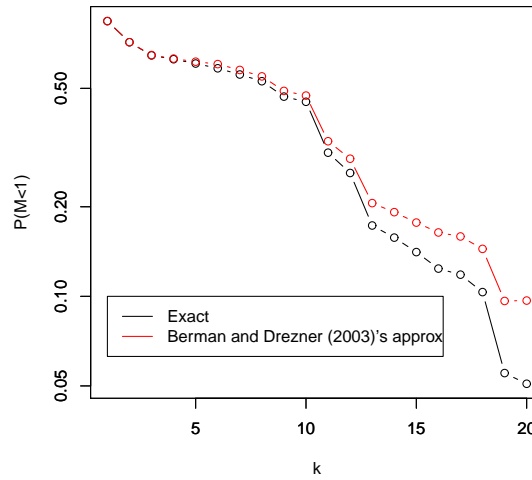


Figure 8: The cdf of $\max(X_1, \dots, X_k)$ computed using (8) and Berman and Drezner [3]'s numerical approximation. $\rho = 0.2$.

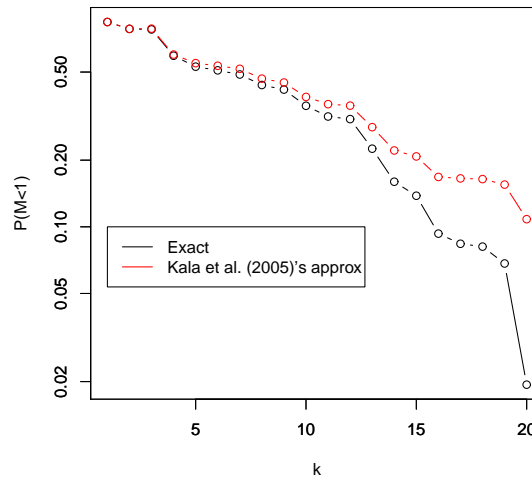


Figure 9: The cdf of $\max(X_1, \dots, X_k)$ computed using (8) and Kala *et al.* [14]'s numerical approximation. $\rho = 0.2$.

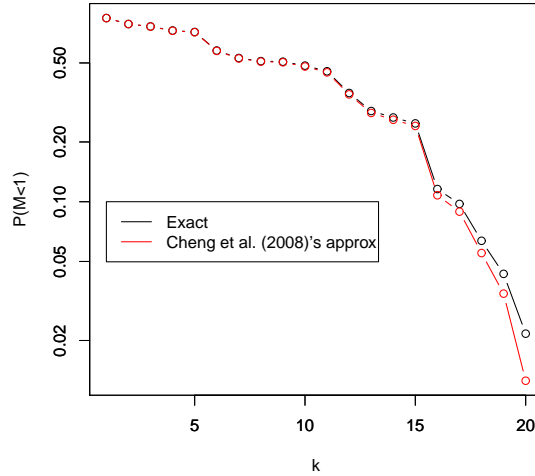


Figure 10: The cdf of $\max(X_1, \dots, X_k)$ computed using (8) and Cheng *et al.* [6]'s numerical approximation. $\rho = 0.2$.

We now compare (8) with known approximations for the cdf of M due to Berman and Drezner [3], Kala *et al.* [14] and Cheng *et al.* [6]. Figure 8 compares values of $F_M(1)$ computed using (8) and the numerical approximation due to Berman and Drezner [3]. Figure 9 compares values of $F_M(1)$ computed using (8) and the numerical approximation due to Kala *et al.* [14]. Figure 10 compares values of $F_M(1)$ computed using (8) and the numerical approximation due to Cheng *et al.* [6]. (8) was computed using Mathematica and hence its values are exact. We see that the numerical approximations become more and more inaccurate as k becomes greater. The values of $F_M(1)$ given by the numerical approximations appear significantly inaccurate when $k = 20$. They will become even more inaccurate for larger values of k .

There have been other exact expressions for the pdf and cdf of $\max(X_1, \dots, X_k)$. Arellano-Valle and Genton (2007, page 1889) expressed the pdf of $\max(X_1, \dots, X_k)$ as a $(k - 1)$ -variate integral. Corollary 3 in Jamalizadeh and Balakrishnan (2010) expressed the pdf and cdf of $\max(X_1, \dots, X_k)$ as $(k - 1)$ -variate and k -variate integrals, respectively. These expressions are not very efficient as their computations will become prohibitively expensive for large k . Our expressions for the cdf and pdf of $\max(X_1, \dots, X_k)$ given by (8) and (9), respectively, are more efficient. Both involve only a single infinite sum. The other summations involved are all finite sums.

The computational efficiency of (9) over the expressions in Arellano-Valle and Genton (2007) and Jamalizadeh and Balakrishnan (2010) is illustrated in Figure 11. (9) is more efficient than the expressions in Arellano-Valle and Genton (2007) and Jamalizadeh and Balakrishnan (2010) for every k . As expected, (9) becomes increasingly more efficient as k increases from 2 to 10.

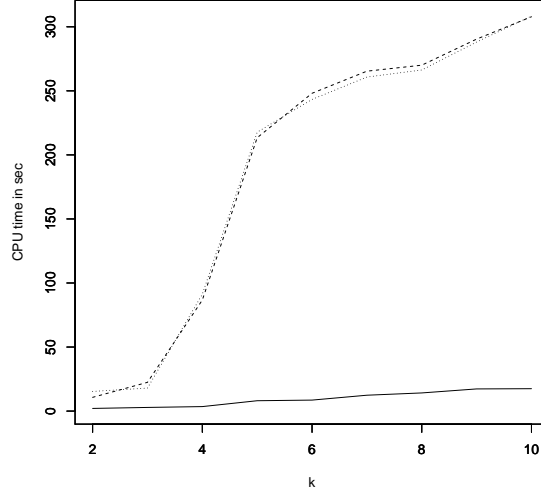


Figure 11: The central processing unit time taken for computing the pdf of M at 1 using (9) (solid curve), the expression in Arellano-Valle and Genton (2007) (curve of dots) and the expression in Jamalizadeh and Balakrishnan (2010) (curve of dashes). $\rho = 0.2$.

Neither Arellano-Valle and Genton (2007) nor Jamalizadeh and Balakrishnan (2010) give explicit expressions for moments of $\max(X_1, \dots, X_k)$. However, Remark 7 in Jamalizadeh and Balakrishnan (2010) does give an expression for the moment generating function of $\max(X_1, \dots, X_k)$, this expression involves a $(k-1)$ -variate integral. Our expression for the n th moment of $\max(X_1, \dots, X_k)$ given in (10) is a single infinite sum of known special functions. It can be computed efficiently for a wide range of values of k . Any expression for the moments of $\max(X_1, \dots, X_k)$ based on the results in Arellano-Valle and Genton (2007) and Jamalizadeh and Balakrishnan (2010) will involve a k -variate integral, so it is not likely to be efficient. The expression will be at least as inefficient as the expressions for the pdf given in Arellano-Valle and Genton (2007) and Jamalizadeh and Balakrishnan (2010). See Figure 11.

4 Extensions

An immediate extension of the results in Section 2 is to derive the distribution of $M = \max(X_1, \dots, X_k)$ when (X_1, \dots, X_k) is a multivariate normal random vector with means μ_i , variances σ_i^2 and correlation matrix $\mathbf{R} = (\rho_{\alpha, \beta})$. In this case, (6) generalizes to

$$\phi_k(\mathbf{x}) = \left[\prod_{j=1}^k \frac{1}{\sigma_j} \phi\left(\frac{x_j - \mu_j}{\sigma_j}\right) \right] \sum_{j=0}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \prod_{t=1}^k \text{He}_{i_t}\left(\frac{x_t - \mu_t}{\sigma_t}\right)$$

with the notation as described in Lemma 4. Also (7) generalizes to

$$\begin{aligned} \Phi_k(\mathbf{x}) &= \prod_{j=1}^k \Phi\left(\frac{x_j - \mu_j}{\sigma_j}\right) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\prod_{\ell=1}^{k-m} \Phi\left(\frac{x_{s_\ell} - \mu_{s_\ell}}{\sigma_{s_\ell}}\right) \right] \\ &\quad \cdot \left\{ \prod_{\ell=1}^m \left[-\phi\left(\frac{x_{t_\ell} - \mu_{t_\ell}}{\sigma_{t_\ell}}\right) \text{He}_{i_{t_\ell}-1}\left(\frac{x_{t_\ell} - \mu_{t_\ell}}{\sigma_{t_\ell}}\right) \right] \right\} \end{aligned}$$

with the notation as described in Lemma 5. So, the cdf of M becomes

$$F_M(x) = \prod_{j=1}^k \Phi\left(\frac{x - \mu_j}{\sigma_j}\right) + \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\prod_{\ell=1}^{k-m} \Phi\left(\frac{x - \mu_{s_\ell}}{\sigma_{s_\ell}}\right) \right]$$

$$\cdot \left\{ \prod_{\ell=1}^m \left[-\phi \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \text{He}_{i_{t_\ell}-1} \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \right] \right\}. \quad (14)$$

Differentiating (14) with respect to x , we obtain the pdf of M as

$$\begin{aligned} f_M(x) &= \sum_{p=1}^k \frac{1}{\sigma_p} \phi \left(\frac{x - \mu_p}{\sigma_p} \right) \prod_{j \neq p} \Phi \left(\frac{x - \mu_j}{\sigma_j} \right) \\ &+ \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\sum_{p=1}^{k-m} \frac{1}{\sigma_{s_p}} \phi \left(\frac{x - \mu_{s_p}}{\sigma_{s_p}} \right) \prod_{\ell \neq p} \Phi \left(\frac{x - \mu_{s_\ell}}{\sigma_{s_\ell}} \right) \right] \left\{ \prod_{\ell=1}^m \left[-\phi \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \text{He}_{i_{t_\ell}-1} \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \right] \right\} \\ &+ \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\prod_{\ell=1}^{k-m} \Phi \left(\frac{x - \mu_{s_\ell}}{\sigma_{s_\ell}} \right) \right] \\ &\quad \cdot \sum_{p=1}^m \frac{x - \mu_{t_p}}{\sigma_{t_p}^2} \phi \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \text{He}_{i_{t_p}-1} \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \prod_{\ell \neq p} \left[-\phi \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \text{He}_{i_{t_\ell}-1} \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \right] \\ &- \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\prod_{\ell=1}^{k-m} \Phi \left(\frac{x - \mu_{s_\ell}}{\sigma_{s_\ell}} \right) \right] \\ &\quad \cdot \sum_{p=1}^m \frac{x - \mu_{t_p}}{\sigma_{t_p}^2} \phi \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \text{He}_{i_{t_p}-1} \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \prod_{\ell \neq p} \left[-\phi \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \text{He}_{i_{t_\ell}-1} \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \right] \\ &+ \sum_{j=1}^{\infty} \frac{1}{j!} \sum_i^* C_i \prod_{\alpha, \beta} \rho_{\alpha, \beta} \left[\prod_{\ell=1}^{k-m} \Phi \left(\frac{x - \mu_{s_\ell}}{\sigma_{s_\ell}} \right) \right] \\ &\quad \cdot \sum_{p=1}^m \frac{1}{\sigma_{t_p}} \phi \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \text{He}_{i_{t_p}} \left(\frac{x - \mu_{t_p}}{\sigma_{t_p}} \right) \prod_{\ell \neq p} \left[-\phi \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \text{He}_{i_{t_\ell}-1} \left(\frac{x - \mu_{t_\ell}}{\sigma_{t_\ell}} \right) \right]. \end{aligned} \quad (15)$$

The n th moment of M can be calculated by integrating (15) term by term. But it does not appear possible to reduce the integrals to known special functions. For instance, even the integral

$$\int_{-\infty}^{\infty} x^n \phi \left(\frac{x - \mu_p}{\sigma_p} \right) \prod_{j \neq p} \Phi \left(\frac{x - \mu_j}{\sigma_j} \right) dx,$$

the integral corresponding to the first term in (15), does not appear to have a closed form expression. Hence, the derivation of the moments corresponding to (15) is an open problem.

Other future work are to

- find upper bounds or exact asymptotic behaviors for the truncation error of the infinite series in (8), (9) and (10). Such upper bounds will help to choose the point at which to truncate the infinite series.
- find the limiting distribution of $\max(X_1, \dots, X_k)$ as $k \rightarrow \infty$.
- determine if a normal distribution can be used to approximate the distribution of $\max(X_1, \dots, X_k)$.
- derive the exact distribution of $\max(X_1, \dots, X_k)$ when (X_1, \dots, X_k) follows non-normal distributions like the multivariate skew-normal, multivariate skew- t , multivariate Laplace, multivariate logistic, multivariate Kotz type, multivariate Pareto, multivariate gamma, multivariate extreme value and multivariate generalized hyperbolic distributions.

Acknowledgments

The authors would also like to thank the Editor and the referee for careful reading and comments which greatly improved the paper.

References

- [1] Afonja, B. (1972). The moments of the maximum of correlated normal and t variates. *Journal of the Royal Statistical Society, B*, **34**, 251-262.
- [2] Arellano-Valle, R. B. and Genton, M. G. (2007). On the exact distribution of linear combinations of order statistics from dependent random variables. *Journal of Multivariate Analysis*, **98**, 1876-1894.
- [3] Berman, O. and Drezner, Z. (2003). A probabilistic one-centre location problem on a network. *Journal of the Operational Research Society*, **54**, 871-877.
- [4] Berman, S. M. (1964). Limit theorems for the maximum term in stationary sequences. *Annals of Mathematical Statistics*, **35**, 502-516.
- [5] Bertsimas, D., Natarajan, K. and Teo, C. -P. (2004). Probabilistic combinatorial optimization: Moments, semidefinite programming, and asymptotic bounds. *SIAM Journal of Optimization*, **15**, 185-209.
- [6] Cheng, B., Chow, S. C., Burr, D. and Cosmatos, D. (2008). Statistical assessment of QT/QTc prolongation based on maximum of correlated normal random variables. *Journal of Biopharmaceutical Statistics*, **18**, 494-501.
- [7] Chetverikov, D. (2012). Adaptive test of conditional moment inequalities. *Working Paper CWP36/12*, The Institute for Fiscal Studies, Department of Economics, University College London, London.
- [8] Deo, C. M. (1972). Some limit theorems for maxima of absolute values of Gaussian sequences. *Sankhyā, A*, **34**, 289-292.
- [9] Eriksson, H., Larsson-Edefors, P. and Eckerbert, D. (2006). Toward architecture-based test-vector generation for timing verification of fast parallel multipliers. *IEEE Transactions on Very Large Scale Integration Systems*, **14**, 370-379.
- [10] Exton, H. (1978). *Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs*. Halsted Press, New York.
- [11] Harper, A. (2013). Bounds on the suprema of Gaussian processes, and omega results for the sum of a random multiplicative function. *Annals of Applied Probability*, **23**, 584-616.
- [12] Harper, A. (2015). Pickands' constant H_α does not equal $1/\Gamma(1/\alpha)$, for small α . *Technical Report*, University of Cambridge, Cambridge, UK.
- [13] Jamalizadeh, A. and Balakrishnan, N. (2010). Distributions of order statistics and linear combinations of order statistics from an elliptical distribution as mixtures of unified skew-elliptical distributions. *Journal of Multivariate Analysis*, **101**, 1412-1427.
- [14] Kala, R., Krzysko, M. and Wolynski, W. (2005). Two preliminary tests for discriminant analysis. *Communications in Statistics—Simulation and Computation*, **34**, 179-189.
- [15] Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000). *Continuous Multivariate Distributions: Models and Applications*. John Wiley and Sons, New York.
- [16] Nadarajah, S. and Kotz, S. (2008). Exact distribution of the max/min of two Gaussian random variables. *IEEE Transactions on Very Large Scale Integration Systems*, **16**, 210-212.
- [17] Ross, A. (2010). Computing bounds on the expected maximum of correlated normal variables. *Methodology and Computing in Applied Probability*, **12**, 111-138.