Journal of Computational and Applied Mathematics 311 (2017) 1–10

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Letter to the editor

The exact density of the sum of independent skew normal random variables

Saralees Nadarajah *, Rui Li

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

ARTICLE INFO

Article history:
Received 29 May 2015
Received in revised form 10 December 2015

Keywords:
Convolution
Kampé de Fériet function
Normal distribution

ABSTRACT

The exact distribution of the sum of independent and non-identical skew normal random variables is derived. Its computational efficiency and a real data application are illustrated.

1. Introduction

The biggest weakness of the normal distribution is its inability to model skewness. Many skew versions of the normal distribution have been proposed in the literature. Perhaps the most popular and most widely used of these is the skew normal distribution due to Azzalini [1]. A random variable is said to have the skew normal distribution if its probability density function (pdf) is

$$f_X(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\frac{x-\xi}{\omega}\right) \tag{1}$$

for $-\infty < x < \infty$, $-\infty < \xi < \infty$, $\omega > 0$ and $-\infty < \alpha < \infty$, where $\phi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the pdf and cumulative distribution function (cdf) of a standard normal random variable. The cdf corresponding to (1) is

$$F(x) = \Phi(x) - 2T(x, \lambda)$$

for $-\infty < x < \infty$ and $-\infty < \lambda < \infty$, where $T(h, a)$ is Owen's T function [2] defined by

$$T(h, a) = \frac{1}{2\pi}\int_0^a \frac{1}{1+x^2}\exp\left[-\frac{h^2\left(1+x^2\right)}{2}\right]dx$$

for $-\infty < h < \infty$ and $-\infty < a < \infty$. The parameters ξ, ω and α are, respectively, the location, scale and shape parameters. Negative values for α give negatively skewed shapes for (1). Positive values for α give positively skewed shapes for (1). The shape of (1) for $\alpha = 0$ is symmetric. In fact, (1) for $\alpha = 0$ is the normal pdf with mean μ and standard deviation ω.

The properties of the skew normal distribution given by (1) have been studied by many authors. Some of the recent properties derived include: characteristic functions of scale mixtures [3]; log-concavity and monotonicity of hazard and reversed hazard functions [4]; measures of skewness [5]; procedures for small area estimation [6]; the Kullback–Leibler

* Corresponding author.
E-mail address: mbbsssn2@manchester.ac.uk (S. Nadarajah).

http://dx.doi.org/10.1016/j.cam.2016.06.032
0377-0427/

divergence measure [7]; asymptotic expansions for moments of the extremes [8]; rates of convergence of the extremes [9]; the distribution of quadratic forms [10]; identifiability of finite mixtures of skew normal distributions [11]; expressions for moments of order statistics and records [12]; sampling distributions [13]; procedures for tests in variance components [14]; bias corrected estimators [15].

Real life applications of the skew normal distribution given by (1) have also been studied by many authors. Some recent applications have included: analysis of student satisfaction towards university courses [16]; modelling of air pollution data [17]; modelling of bounded health scores [18]; modelling of psychiatric measures [19]; modelling of insurance claims [20]; asset pricing [21]; growth estimates of cardinalfish [22]; individual loss reserving [23]; detecting differential expression to microRNA data [24]; estimation of age-specific fertility rates [25]; an evaluation of the EU proposed farm income stabilisation tool [26]; robust portfolio estimation [27]; an empirical study on the energy intensity in China [28].

Excellent accounts of the theory and applications of the skew normal distribution can be found in [29,30].

Sums of skew normal random variables arise in different ways, see [29, Section 2.6] and [31]. Application areas of such sums include finance [32, Section 2.6], the cosmic evolution of faint satellite galaxies [33], Bayesian analysis [34] and data envelopment analysis [35].

To the best of our knowledge, the exact distribution of the sum of skew normal random variables has not been known to date in a closed form. The aim of this note is to derive closed form expressions for the probability density function (pdf) of the sum. The expressions involve the Kampé de Fériet function [36] defined by

$$F^{p_1,\ldots,p_k}_{q_1,\ldots,q_k}\left(\begin{matrix} a^{(1)}_1,\ldots,a^{(1)}_{p_1} \\ b^{(1)}_1,\ldots,b^{(1)}_{q_1} \end{matrix} \middle| \begin{matrix} a^{(2)}_1,\ldots,a^{(2)}_{p_2} \\ b^{(2)}_1,\ldots,b^{(2)}_{q_2} \end{matrix} \middle| \cdots \middle| \begin{matrix} a^{(k)}_1,\ldots,a^{(k)}_{p_k} \\ b^{(k)}_1,\ldots,b^{(k)}_{q_k} \end{matrix} \middle| x_1,\ldots,x_{k-1} \right)$$
$$= \sum_{j_1=0}^{\infty}\cdots\sum_{j_{k-1}=0}^{\infty} \frac{\prod_{j=1}^{p_1}\left(a^{(1)}_j\right)_{j_1+\cdots+j_{k-1}}}{\prod_{j=1}^{q_1}\left(b^{(1)}_j\right)_{j_1+\cdots+j_{k-1}}} \frac{\prod_{j=1}^{p_2}\left(a^{(2)}_j\right)_{j_1}}{\prod_{j=1}^{q_2}\left(b^{(2)}_j\right)_{j_1}} \cdots \frac{\prod_{j=1}^{p_k}\left(a^{(k)}_j\right)_{j_{k-1}}}{\prod_{j=1}^{q_k}\left(b^{(k)}_j\right)_{j_{k-1}}} \frac{x_1^{j_1}\cdots x_{k-1}^{j_{k-1}}}{j_1!\cdots j_{k-1}!},$$

where $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$ denotes the ascending factorial. As can be seen, $F^{p_1,\ldots,p_k}_{q_1,\ldots,q_k}(\cdots)$ is a $(k-1)$ fold infinite sum of $x_1^{j_1}\cdots x_{k-1}^{j_{k-1}}/(j_1!\cdots j_{k-1}!)$ over $j_1\geq 0,\ j_2\geq 0,\ldots,j_{k-1}\geq 0$. The numerator of the coefficient of $x_1^{j_1}\cdots x_{k-1}^{j_{k-1}}/(j_1!\cdots j_{k-1}!)$ is determined by $p_1,\ldots,p_k$ and the denominator of its coefficient is determined by $q_1,\ldots,q_k$. The numerator is a product of k terms, the first term being a product of p_1 ascending factorials, the second term being a product of p_2 ascending factorials, and so on. The denominator is a product of k terms, the first term being a product of q_1 ascending factorials, the second term being a product of q_2 ascending factorials, and so on. For example,

$$F^{1,0}_{0,1}\left(\begin{matrix} a \\ - \end{matrix}\middle|\begin{matrix} - \\ b \end{matrix}\middle| x\right) = \sum_{j_1=0}^{\infty}\frac{(a)_{j_1}}{(b)_{j_1}}\frac{x^{j_1}}{j_1!},$$

$$F^{1,1,0}_{0,1,1}\left(\begin{matrix} a \\ - \end{matrix}\middle|\begin{matrix} b \\ c \end{matrix}\middle|\begin{matrix} - \\ d \end{matrix}\middle| x_1,x_2\right) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\frac{(a)_{j_1+j_2}\,(b)_{j_1}}{(c)_{j_1}\,(d)_{j_2}}\frac{x_1^{j_1}x_2^{j_2}}{j_1!j_2!},$$

$$F^{1,1,1,0}_{0,1,1,1}\left(\begin{matrix} a \\ - \end{matrix}\middle|\begin{matrix} b \\ c \end{matrix}\middle|\begin{matrix} d \\ e \end{matrix}\middle|\begin{matrix} - \\ f \end{matrix}\middle| x_1,x_2,x_3\right) = \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\sum_{j_3=0}^{\infty}\frac{(a)_{j_1+j_2+j_3}\,(b)_{j_1}\,(d)_{j_2}}{(c)_{j_1}\,(e)_{j_2}\,(f)_{j_3}}\frac{x_1^{j_1}x_2^{j_2}x_3^{j_3}}{j_1!j_2!j_3!},$$

and so on.

In-built routines for computing the Kampé de Fériet function are available in packages like Maple, Matlab and Mathematica. See also [37, Section 1.3.2] and Exton and Krupnikov [38].

The closed form expressions for the pdf of the sum are derived in Section 2. Their computational efficiency is illustrated in Section 3. A real data application is illustrated in Section 4.

2. Main result

Let $X_j,\ j=1,2,\ldots,n$ be independent skew normal random variables specified by the pdfs

$$f_{X_j}(x) = \frac{2}{\omega_j}\phi\left(\frac{x-\xi_j}{\omega_j}\right)\Phi\left(\alpha_j\frac{x-\xi_j}{\omega_j}\right). \tag{2}$$

Theorem 2.1 derives the pdf of the sum of X_js. Its proof requires two technical results (Lemmas 2.1 and 2.2).

Lemma 2.1. *For $a>0$ and $-\infty<b<\infty$,*

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(\mathrm{i}bt-\frac{a^2t^2}{2}\right)dt = \frac{1}{a}\phi(ab),$$

where $\mathrm{i}=\sqrt{-1}$.