symmetry

Article

Copulaesque Versions of the Skew-Normal and Skew-Student Distributions

Christopher Adcock

Sheffield University Management School, University College Dublin, Sheffield S10 1FL, UK;
c.j.adcock@sheffield.ac.uk

Abstract: A recent paper presents an extension of the skew-normal distribution which is a copula. Under this model, the standardized marginal distributions are standard normal. The copula itself depends on the familiar skewing construction based on the normal distribution function. This paper is concerned with two topics. First, the paper presents a number of extensions of the skew-normal copula. Notably these include a case in which the standardized marginal distributions are Student's t, with different degrees of freedom allowed for each margin. In this case the skewing function need not be the distribution function for Student's t, but can depend on certain of the special functions. Secondly, several multivariate versions of the skew-normal copula model are presented. The paper contains several illustrative examples.

Keywords: copula; multivariate distributions; skew-normal distributions; skew-Student distributions

JEL Classification: C18; G01; G10; G12

check for updates

Citation: Adcock, C.J. Copulaesque Versions of the Skew-Normal and Skew-Student Distributions. *Symmetry* **2021**, *13*, 815. https://doi.org/10.3390/sym13050815

Academic Editor: Nicola Maria Rinaldo Loperfido

Received: 14 April 2021
Accepted: 27 April 2021
Published: 6 May 2021

1. Introduction

The recent paper by [1] presents an extension of the skew-normal distribution which has subsequently been referred to as a copula. Under this model, the standardized marginal distributions are standard normal and some of the conditional distributions are skew-normal.The skew-normal distribution itself was introduced in two landmark papers [2] and [3]. These papers have led to a very substantial research effort by numerous authors over the last thirty five years. The result of these efforts include, but are certainly not limited to, numerous probability distributions, both univariate and multivariate, which are loosely referred to in the literature as skew-elliptical distributions. This term does not completely describe the rich features of these distributions, but for convenience will be used in this paper. Notable contributions to these developments include papers by [4–9] among numerous others.

The many multivariate distributions that may be referred to as skew-elliptical offer coherent probability models that are used in a wide variety of applications. However, they all share the feature that the factor which perturbs symmetry is applied as a multiplier to a distribution that is elliptically symmetric. Consequently the marginal distributions are all prescribed. For example, the multivariate skew-t distribution described in [5] leads to marginal distributions that are all univariate skew-t distributions with the same degrees of freedom. While such a restriction may be acceptable or even irrelevant for some purposes, it is nonetheless the case that some multivariate applications have marginal distributions that are different. As is very well known, there is a large literature based on copulas which derives from the original paper of [10]. The purpose of a copula is to separate the modeling of the dependence structure of set of variables from the analysis of the marginal distributions. By implication, the marginal distributions need not be the same; that is, they may differ by more than just scale and location.

In a recent paper concerned primarily with projection pursuit, ref [1] presents a trivariate distribution based on the skew-normal which is also in a general sense a copula.

The adjective general is used to refer to the fact that although the distribution does not satisfy the theoretical requirements laid down in [10] there is nonetheless separation between the modeling of the dependence structure and the treatment of the marginal distributions. The skew-normal and skew-Student copulas as they are termed in this paper are amenable to theoretical study, at least to some extent. Loperfido's model also leads naturally to various extensions, which are of theoretical interest and which have the potential to provide tools for empirical research.

The purpose of this expository paper is to describe some of the properties of these new distributions and to present a number of extensions. As described below, these distributions are tractable to some extent, but numerous results of interest must be computed numerically. It is of particular interest to note that the some of the developments described in the paper are dependent on Meijer's G function ([11]) and Fox's H-function ([12]) and that there are therefore computational and methodological issues to be resolved. The structure of the paper is as follows: Section 2 summarizes a basic bivariate skew-normal copula and presents some of its properties. The bivariate model of this section is used to illustrate the difference from a bivariate normal copula that does satisfy the conditions of in [10]. Section 3 presents a conditional version of this distribution that has close connections to the original skew-normal distribution. Section 4 presents an extended version of the distribution; that is a distribution that is analogous to the extended skew-normal. Section 5 extends the results to an n-variate skew-normal copula and Section 6 to Student versions. This distribution in particular allows the marginal distributions to have different degrees of freedom. Section 7 presents results for a different multivariate setup. In this, a vector of $m+n$ variables may be partitioned into components of length m and n each of which has a marginal multivariate normal distribution. The properties of this distribution are briefly discussed, including an outline of a multivariate Student version. Section 8 presents three numerical examples. As this is an expository paper, several of the sections also contain brief discussions of technical issues that are outstanding and which could be the subject of future work. The final section of the paper contains a short summary.

Many of the results in this paper require numerical integration. Example results are generally computed to four decimal places. The usual notation ϕ and Φ are used to denote the standard normal density and distributions functions, respectively. To avoid proliferation of subscripts the notation $f(.)$ is used indifferently to denote a density function. Other notation, if not defined explicitly in the text, is that in common use.

2. The Skew-Normal Distribution as a Copula

The skew-normal copula was introduced by [1]. In their paper, the distribution is presented ariate form. Its principles may be demonstrated by a bivariate form of the distribution with probability density function

$$f(x_1, x_2) = 2\phi(x_1)\phi(x_2)\Phi(\lambda x_1 x_2). \tag{1}$$

The scalar parameter $\lambda \in \mathbb{R}$ determines the extent of the dependence between X_1 and X_2. In Section 5, the bivariate distribution is extended to n variables. To accommodate this extension it is convenient to employ the notation $\mathbf{X} \sim SNC_n(\lambda)$, where $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$. As [1] shows, in the bivariate case the marginal distribution of each X_i is $N(0,1)$ and the conditional distribution of X_1 given that $X_2 = x_2$ is skew-normal with density function

$$f(x_1|X_2 = x_2) = 2\phi(x_1)\Phi(\lambda x_1 x_2). \tag{2}$$

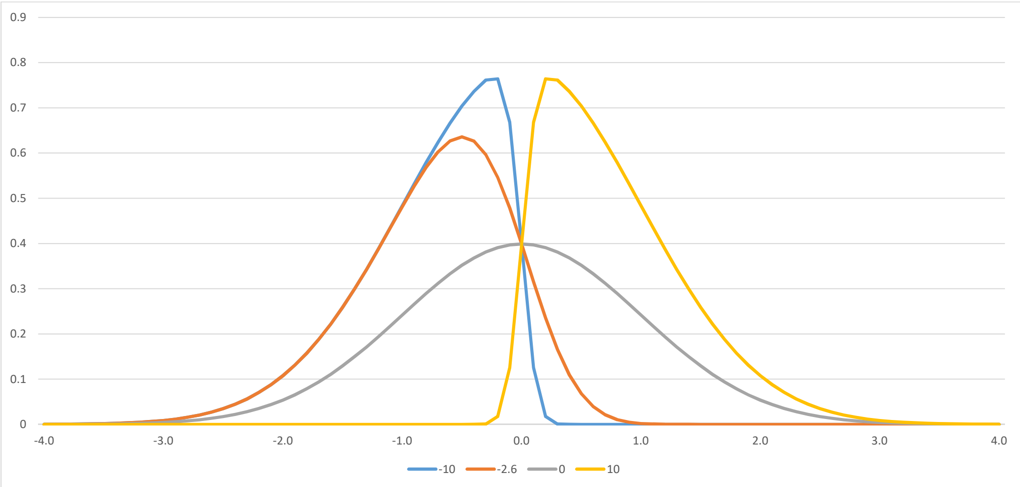
There is an analogous expression for the density function of the conditional distributions of X_2 given that $X_1 = x_1$. It may also be noted that X_2^2 is distributed independently of X_1 as a $\chi^2_{(1)}$ variable, with the analogous result for the distributions of X_2 and X_1^2. Hence the X_i^2, $i = 1, 2$ are independently distributed as $\chi^2_{(1)}$.

To illustrate the difference from the formal definition of a copula, consider the well-known Gaussian copula. In the bivariate case, dependence between X_1 and X_2 is described by the function

$$\Phi_{2,C}\left\{\Phi^{-1}(U_1), \Phi^{-1}(U_2)\right\},$$

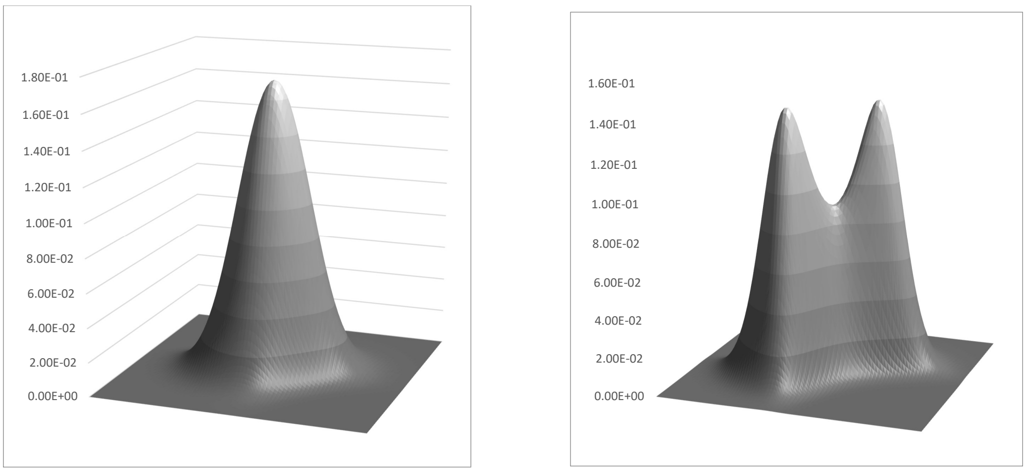
where $\Phi_{2,C}\{..\}$ denotes the distribution function of a bivariate normal distribution with zero means and correlation matrix $\boldsymbol{C}$ and the $U_{1,2}$ are each uniformly distributed on $[0,1]$. The formulation at Equation (2) uses only the standard univariate normal distribution function.

Sketches of the skew-normal density function for $\lambda \in \{-10.0, -2.6, 0.0, 10.0\}$ are shown in Figure 1. Sketches of the bivariate density function for $\lambda = 1.0$ and -5.0 are shown in Figure 2. The bi-modal nature of the density function when $\lambda = -5.0$ is noteworthy. Indeed it is straight forward to show that the density function is bi-modal if $|\lambda| > \sqrt{\pi/2}$, with the model values being at points $\pm(X, X)$ depending on λ. Examples of modal values are shown in Table 1 for a range of values of λ.



The Figure shows plots of the standardized skew-normal density function for λ = -10.0, -2.6, 0.0 & 10.

Figure 1. Skew-normal Density Functions; $\lambda = -10.0, -2.6, 0.0, 10.0$.



The Figures show plots of the bivariate skew-normal copula density function for - 4.0 ≤, x, y ≤ 4.0 for λ = 1.0 and λ = -5.0.

Figure 2. Bivariate Skew-normal Copula Density Functions; $\lambda = 1.0, -5.0$.

Table 1. Modal Values of the Bivariate Skew-normal Copula Distribution.

λ	Mode	λ	Mode
1.254	0.0231	35	0.2564
1.275	0.1296	50	0.2213
3	0.5422	75	0.1865
4	0.5258	10^3	0.1648
5	0.5029	10^3	0.0588
15	0.3565	10^4	0.0200
10	0.4112	10^5	0.0063
20	0.3299	10^6	0.0020

2.1. Cross-Moments

The covariance of X_1 and X_2, which also equals their correlation, is

$$cov(X_1, X_2) = 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 \phi(x_1)\phi(x_2)\Phi(\lambda x_1 x_2)dx_1 dx_2, \tag{3}$$

which is readily shown to be

$$cov(X_1, X_2) = 2\left(\lambda/\sqrt{2\pi}\right)\int_{-\infty}^{\infty}\left(x_2^2/\sqrt{1+\lambda^2 x_2^2}\right)\phi(x_2)dx_2, \tag{4}$$

or

$$cov(X_1, X_2) = \left(\lambda\sqrt{2/\pi}\right)\int_{0}^{\infty}\left(y/\sqrt{1+\lambda^2 y}\right)f(y)dy, \tag{5}$$

where $Y \sim \chi^2_{(1)}$. There is no reported analytic expression for this integral in general. However, it equals zero when $\lambda = 0$ and tends to $\pm 2/\pi$ as $\lambda \to \pm\infty$. Note that the integral in Equation (5) may also be expressed in terms of a Chi-squared distribution with 3 degrees of freedom. Table 2 shows values of the correlation for a range of positive values of λ.

Table 2. Correlation of the Bivariate Skew-normal Copula Distribution.

λ	Cov(=Cor)	λ	Cov(=Cor)
0.0010	0.0008	0.5000	0.3129
0.0020	0.0016	1.0000	0.4505
0.0040	0.0032	2.0000	0.5503
0.0100	0.0080	5.0000	0.6116
0.0200	0.0159	10.0000	0.6268
0.0500	0.0397	20.0000	0.6316
0.1000	0.0786	50.0000	0.6331
0.2000	0.1511	100.0000	0.6333

For higher order cross-moments the following results hold. If p and q are positive integers with $p + q$ odd, then

$$E\left(X_1^p X_2^q\right) = 0. \tag{6}$$

Note that there is no need to derive this result using integration by parts. Since p or q must be even, it is a consequence of one of the properties above, as is the result for the case where p and q are both even

$$E\left(X_1^p X_2^q\right) = 2^{\frac{p+q}{2}}\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{q+1}{2}\right)/\pi. \tag{7}$$

For p and q both odd, the expectation satisfies the recursion

$$E\left(X_1^p X_2^q\right) = (p-1)E\left(X_1^{p-2}X_2^q\right) + \lambda\sqrt{(2/\pi)}\int_{-\infty}^{\infty}\left\{x^{q+1}/\left(1+\lambda^2x^2\right)^{p/2}\right\}\phi(x)dx, \tag{8}$$

with

$$E\left(X_1 X_2^q\right) = \lambda\sqrt{(2/\pi)}\int_{-\infty}^{\infty}\left\{x^{q+1}/\left(1+\lambda^2x^2\right)^{1/2}\right\}\phi(x)dx. \tag{9}$$

In a recent paper, ref [13] show that cross-moments of this distribution may also be computed using a new extension of Stein's lemma, ref [14]. There is no particular advantage in using the new lemma for the bivariate distribution at Equation (1). The result is however employed in Section 5 which is concerned with the more general case of n variables. If $p = 2k+1$ and $q = 2l+1$, it is straightforward to show that as $\lambda \to \pm\infty$ the limiting value of $E(X^pY^q)$ is

$$lim_{\lambda\to\infty}E(X^pY^q) = \pm 2^{k+l+1}k!l!/\pi.$$

The limiting values of a selection of odd order cross moments $E\left(X_1^p X_2^q\right)$ are shown in Table 3 and a selection of moments corresponding to $\lambda = 1.0$ in Table 4.

Table 3. Cross Moments—Limiting Values.

Order	1	3	5	7	9
1	0.6366	1.2732	5.093	30.5577	244.462
3		2.5481	10.1891	61.1282	489.0004
5			40.7564	244.5129	1956.0016
7				1467.0776	11,736.0093
9					93,888.0744

Table 4. Cross Moments: $\lambda = 1.0$.

Order	1	3	5	7	9
1	0.4507	1.0808	4.5040	27.4555	221.5542
3	1.0808	2.4329	9.9094	59.7483	478.9353
5	4.5040	9.9094	40.1754	241.9398	1938.3608
7	27.4555	59.7483	241.9398	1456.7783	11,670.9677
9	221.5542	478.9353	1938.3608	11,670.9677	93,501.497
11	2227.8507	4794.0952	19,397.487	116,791.7771	935,671.459
13	26,837.3188	57,561.2949	232,862.2301	1,402,049.0507	11,232,438.0284

Although not specifically required for the determination of the normalization constant in Equation (1), the distribution of the product $Y = X_1X_2$ for the case where each $X_i, i = 1,2$

is independently distributed as $N(0,1)$ is of interest. When $X_i \sim N(0,\sigma_i^2)$ independently for $i = 1,2$, the distribution of Y has the density function

$$f(y) = K_0(|y|/\sigma_1\sigma_2)/\pi\sigma_1\sigma_2, \tag{10}$$

where $K_0(.)$ is the modified Bessel function of the second kind. See, for example, ref [15] for further details of this result and [16] (Chapter 9, Sections 6 to 8) for details of the function K itself.

3. A New Skew-Normal Type Distribution

As noted above in Section 2, the conditional distribution of X_1 given $X_2 = x_2$ is skew-normal with shape parameter λx_2. A new univariate distribution may be obtained by conditioning instead on $X_2 \leq y$. The resulting distribution of X_1 has the density function

$$f(x) = 2\phi(x)\left\{\int_{-\infty}^{y}\phi(s)\Phi(\lambda xs)ds\right\}/\Phi(y). \tag{11}$$

Straight forward integration gives the following results.

Proposition 1. *Let X have the distribution with density function given by Equation (11). The following results hold:*

1. *An alternative expression for the density function is*

$$f(x) = 2\phi(x)\Phi(\lambda yx) - 2\lambda x\phi(x)\left\{\int_{-\infty}^{y}\Phi(s)\phi(\lambda xs)ds\right\}/\Phi(y). \tag{12}$$

2. *A second alternative expression is*

$$f(x) = \phi(x) - 2x\phi(x)\xi_1(y)\int_0^{\lambda}\frac{\phi(\lambda xy)}{1+\lambda^2x^2}d\lambda, \tag{13}$$

where $\xi_1(x) = \phi(x)/\Phi(x)$.

3. X^2 *is distributed as* $\chi^2_{(1)}$.
4. *Odd order moments may be computed recursively from*

$$E(X^n) = (n-1)E\left(X^{n-2}\right) + \sqrt{2/\pi}\lambda\left\{E\left(Z^{n-1}\right)/\Phi(y)\right\}\int_{-\infty}^{y}\frac{s\phi(s)}{(1+\lambda^2s^2)^{n/2}}ds, \tag{14}$$

with $Z \sim N(0,1)$ *and*

$$E(X) = \left\{\sqrt{2/\pi}\lambda/\Phi(y)\right\}\int_{-\infty}^{y}\frac{s\phi(s)}{(1+\lambda^2s^2)^{1/2}}ds. \tag{15}$$

The expression at Equation (13) may be computed using the methods reported in [17]. This distribution possesses two interesting properties as $|y| \to \infty$.

Proposition 2. *Let X have the distribution with density function given by Equation (11). The following results hold:*

1. *For* $\lambda > 0$, *as* $y \to -\infty$ *the limiting form of the distribution of X is skew-normal with shape parameter* λy; *that is*

$$lim_{y\to-\infty}f(x) = 2\phi(x)\Phi(\lambda yx). \tag{16}$$

2. *For* $\lambda > 0$ *as* $y \to \infty$

$$lim_{y\to\infty}f(x) = \phi(x). \tag{17}$$

Part 1 of the proposition may be established using the well-known asymptotic formula for the standard normal integral ([16], page 932, Equation 26.2.12); that is

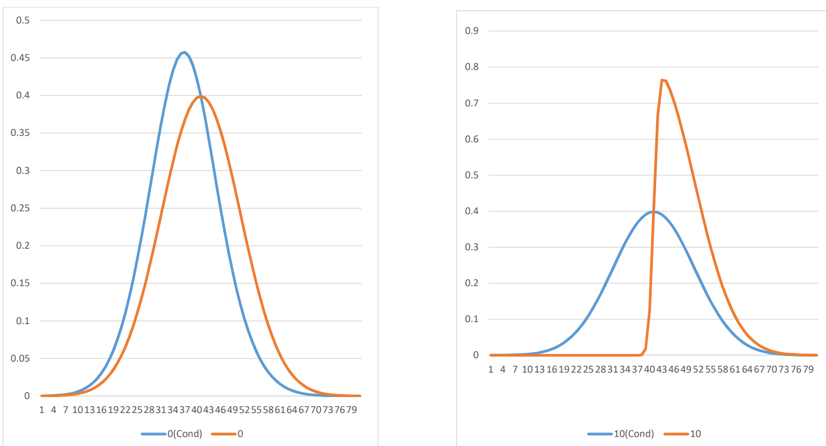
$$\Phi(x) \simeq \phi(x)/|x|; x << 0, \tag{18}$$

and integration by parts. In this case for $\lambda > 0$ and $y << 0$ the second term is negligible. There are analogous results for $\lambda < 0$. Equation (18) and the same assumption are used again in some of the results below. It is interesting to note that Part 1 of the propsition results in the same density function as that at Equation (2).

Numerical computations or asymptotic arguments are necessary in order to employ the distribution at Equation (11); for example to compute moments or critical values. Nonetheless, it is arguably a more flexible form than the familiar skew-normal shown at Equation (2). Examples of the density functions from Equations (2) and (11) are shown in Figure 3. In Table 5, $m1$ and $m3$ are the first and third moment about the origin, sk and ku are skewness and kurtosis, ssk and sku are the corresponding standardized values. The table shows a selection of moments for $\lambda = 1$ and a range of values of y from -10.0 to 10.0. For positive values of y, skewness and kurtosis rapidly tend to 0 and 3 , respectively. For $y < 0$ skewness is negative and there is excess kurtosis. As above, there are analogous results for $\lambda < 0$.

Table 5. Moments of New Skew-normal Type Distribution Function; $\lambda = 1$.

$\leq y$	$m1$	$m3$	sk	ssk	ku	sku
−10.0	−0.7940	−1.5957	−0.2148	−0.9563	0.5223	3.8241
−7.6	−0.7913	−1.5956	−0.2126	−0.9302	0.5304	3.7942
−5.1	−0.7839	−1.5950	−0.2068	−0.8640	0.5527	3.7198
−2.6	−0.7538	−1.5884	−0.1836	−0.6471	0.6514	3.4929
0.0	−0.4174	−1.0540	0.0528	0.0703	2.1945	3.2182
2.4	−0.0062	−0.0131	0.0054	0.0054	2.9999	3.0001
4.9	0.0000	0.0000	0.0000	0.0000	3.0000	3.0000
7.4	0.0000	0.0000	0.0000	0.0000	3.0000	3.0000
10.0	0.0000	0.0000	0.0000	0.0000	3.0000	3.0000



The density functions in blue show the new skew-normal type distribution function for values of λy = 0 and 10, respectively. For comparison, the density functions in orange show the corresponding skew-normal distributions.

Figure 3. Comparison of the Skew-normal and New Skew-normal Type Distributions; $\lambda y = 0.0$ and 10.0.

4. An Extended Skew-Normal Copula

The skew-normal and skew-Student distributions have extended forms. These arise naturally when conditional distributions are considered. They also arise as hidden truncation models. In standardized form, the univariate extended skew-normal distribution has the density function

$$f(x) = \phi(x)\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda x\right)/\Phi(\tau); \tau \in \mathbb{R}. \tag{19}$$

The extended skew-normal distribution may be derived by considering a standardized bivariate normal distribution of X and Y with correlation ρ. The distribution at Equation (19) is the density of X given that $Y \geq -\tau$, with $\lambda = \rho/\sqrt{1-\rho^2}$. The distribution may be denoted $ESN(\lambda, \tau)$. A bivariate extended skew-normal copula distribution that is analogous to that in Section 2 has the density function

$$f(x_1, x_2) = \phi(x_1)\phi(x_2)\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda x_1 x_2\right)/\Omega(\tau, \lambda), \tag{20}$$

where $\Omega(\tau, \lambda) > 0$ is the normalizing constant. Integration with respect to X_1 shows that

$$\Omega(\tau, \lambda) = \int_{-\infty}^{\infty} \phi(y)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 y^2}\right)dy, \tag{21}$$

which may be evaluated numerically. In addition, the marginal distribution of X is symmetric and has the density function.

$$f(x) = \phi(x)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 x^2}\right)/\Omega(\tau, \lambda). \tag{22}$$

Values of $\Omega(\tau, \lambda)$ computed to four decimal places for a selection of values of τ and λ are shown in Table 6. Note that for values of τ less than say -5 the values of Ω are small, implying the need for care with its computation. The shorthand notation $\mathbf{X} \sim ESNC_n(\tau, \lambda)$ which extends that defined in Section 2 will be used. The basic properties of the bivariate extended skew-normal copula distribution are as follows.

Table 6. Normalising Constant for the ESNC Distribution.

τ/λ	0.5	1	2.5	5	10	25
−15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
−10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
−5	0.0000	0.0002	0.0008	0.001	0.0011	0.0011
−2.5	0.0072	0.0105	0.0155	0.0168	0.0172	0.0173
−1	0.1550	0.1396	0.1108	0.1057	0.1048	0.1045
−0.5	0.3051	0.2898	0.2412	0.2138	0.2069	0.2052
0	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
0.5	0.6949	0.7102	0.7588	0.7862	0.7931	0.7948
1	0.8450	0.8604	0.8892	0.8943	0.8952	0.8955
2.5	0.9928	0.9895	0.9845	0.9832	0.9828	0.9827
5	1.0000	0.9998	0.9992	0.999	0.9989	0.9989
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

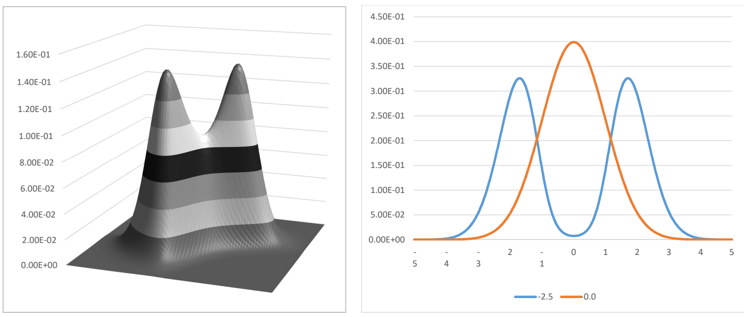
Proposition 3. *Let* $(X_1, X_2)^T \sim ESNC_2(\tau, \lambda)$, *let* X *denote either variable and define*

$$\tilde{\tau}_x = \frac{\tau\sqrt{1+\lambda^2}}{\sqrt{1+\lambda^2 x^2}}. \tag{23}$$

The following properties hold

1. $E(X) = 0$.
2. $var(X) = 1 - \left\{\lambda^2\tau\sqrt{1+\lambda^2}/\Omega(\tau,\lambda)\right\} \int_{-\infty}^{\infty} \frac{y^2}{(1+\lambda^2 y^2)^{3/2}} \phi(\tilde{\tau}_y)\phi(y)dy$.
3. $cov(X_1, X_2) = \{\lambda/\Omega(\tau,\lambda)\} \int_{-\infty}^{\infty} \frac{y^2}{\sqrt{1+\lambda^2 y^2}} \phi(\tilde{\tau}_y)\phi(y)dy$.
4. $X_1|X_2 = y \sim ESN(\tilde{\tau}_y, \lambda y)$.

Proof of this proposition is in the Appendix A. Table 7 shows standard errors for a range of values of τ and λ. These were computed using numerical integration. Table 8 shows the correlation coefficients for a range of parameter values. For $\lambda = 1.0$ Figure 4 shows (i) an example of the bivariate skew-normal density function for $\tau = -1.0$ and (ii) the marginal density for $\tau = -2.5$ and 0.0. Higher order cross moments may be computed recursively, but for non-zero values of τ the resulting integrals are two-dimensional. Depending on the values of λ and τ the extended version of the distribution is also bi-modal.



The left hand figure shows the bivariate ESNC density for λ = 1.0 and τ = -1.0. The right hand figure shows the marginal distributions also for λ = 1.0, but τ = -2.5 and 0.0.

Figure 4. Extended Skew-normal Copula Density Functions; $\lambda = 1.0$.

Table 7. Standard Errors for the ESNC Distribution.

τ/λ	0.5	1	2.5	5	10	25
−15	5.563804	4.654258	4.178148	4.089074	4.065563	4.058883
−10	4.441411	3.816603	3.472084	3.40682	3.38956	3.384654
−5	2.886818	2.724329	2.574579	2.54366	2.535375	2.533012
−2.5	1.610455	1.924303	1.967713	1.966597	1.965964	1.96576
−1	1.113491	1.249359	1.453953	1.498411	1.507775	1.510312
−0.5	1.041259	1.092314	1.208643	1.283663	1.304663	1.309794
0	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
0.5	0.981337	0.95979	0.923856	0.907661	0.903814	0.902917
1	0.977757	0.953395	0.927976	0.923458	0.922466	0.922193
2.5	0.994202	0.985502	0.977147	0.975165	0.974617	0.974459
5	0.999946	0.999327	0.997783	0.997263	0.997109	0.997064
10	1.000000	0.999999	0.999985	0.999976	0.999973	0.999972
15	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Table 8. Correlation for the ESNC Distribution.

τ/λ	0.5	1	2.5	5
-15	0.9840	0.9772	0.9718	0.9705
-10	0.9750	0.9662	0.9594	0.9578
-5	0.9419	0.9348	0.9273	0.9256
-2.5	0.8174	0.8745	0.8788	0.8786
-1	0.5440	0.7087	0.7914	0.8007
-0.5	0.4272	0.5909	0.7154	0.7417
0	0.3130	0.4507	0.5721	0.6125
0.5	0.2112	0.3122	0.3892	0.4034
1	0.1294	0.1975	0.2422	0.2492
2.5	0.0156	0.0355	0.0561	0.0611
5	0.0001	0.0015	0.0049	0.0060
10	0.0000	0.0000	0.0000	0.0001
15	0.0000	0.0000	0.0000	0.0000

The extended version of the distribution has a conditional distribution that is similar to that described in Proposition 2. As above, proof is in the Appendix A.

Proposition 4. *Let X_1 and X_2 have the distribution with density function given by Equation (20). The following results hold:*

1. *The normalizing constant $\Omega(\tau, \lambda)$ is given by*

$$\Omega(\tau, \lambda) = 1/2 + \tau\lambda^2\sqrt{1+\lambda^2}\int_{-\infty}^{\infty} \frac{y\Phi(y)\phi(\tilde{\tau}_y)}{(1+\lambda^2 y^2)^{3/2}} dy. \tag{24}$$

2. *The distribution function of X_2 is*

$$Pr(X_2 \le y) = \int_{-\infty}^{y} \phi(s)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 s^2}\right) ds/\Omega(\tau, \lambda). \tag{25}$$

3. *The density function of X_1 given that $X_2 \le y$ is*

$$f(x_1 | X_2 \le y) = \phi(x_1)\frac{\int_{-\infty}^{y}\phi(s)\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda x_1 s\right) ds}{\int_{-\infty}^{y}\phi(s)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 s^2}\right) ds} \tag{26}$$

4. *For $\lambda > 0$, as $y \to -\infty$ the limiting form of the distribution of X_1 has the density function*

$$lim_{y\to-\infty} f(x_1) = \phi(x_1)\Phi\left(\tilde{\tau}_y\sqrt{1+\tilde{\lambda}_y^2} + \tilde{\lambda}_y x_1\right)/\Phi(\tilde{\tau}_y), \tag{27}$$

where $\tilde{\tau}_y$ is as defined in Proposition 3 and $\tilde{\lambda}_y = \lambda y$; that is, $X_1 | X_2 \le y \sim ESN(\tilde{\tau}_y, \lambda y)$.

Note that as in Section 3 the distributions in Part 4 of both Propositions 3 and 4 are the same. There are analogous results for $\lambda < 0$ as $y \to \infty$.

5. Extension to n Variables

There is a self-evident extension for the distribution of an n-vector $\mathbf{X}$, with i-*th* element X_i, which has density function

$$f(\mathbf{x}) = 2\Pi_{i=1}^{n}\phi(x_i)\Phi(\lambda\Pi_{i=1}^{n}x_i), \tag{28}$$

and shorthand notation $\mathbf{X} \sim SNC_n(\lambda)$. Recall that [1] describes the case $n = 3$. This distribution has the following properties:

1. The marginal distribution of each X_i is $N(0,1)$.
2. The members of any subset of $\mathbf{X}$ of size $2,\ldots,n-1$ are independently distributed as $N(0,1)$.
3. The $X_i^2; i = 1,\ldots,n$ are independently distributed at $\chi^2_{(1)}$ variables.
4. If $\mathbf{X}$ is partitioned into two non-overlapping sets, $\{X_i\}; i = 1,\ldots,p$ and $\{X_j\}; j = p+1,\ldots,p+q$ with $p+q \leq n$ then (i) the $\{X_i^2\}$ are independently distributed as $\chi^2_{(1)}$ variables independently of (ii) the $\{X_j\}$, which are themselves independently distributed as $N(0,1)$.
5. The distribution of X_i given $X_j = x_j, j \neq i$ is skew-normal, with density function

$$f(x_i) = 2\phi(x_i)\Phi(\lambda\Pi_{i=1}^{n}x_i), \tag{29}$$

that is, with shape parameter $\lambda\Pi_{j\neq i}^{n}x_j$.
6. The distribution of $X_j, j \neq i$ given $X_i = x_i$ is a skew-normal copula, with shape parameter λx_i.

Other properties are briefly described in the rest of this section.

5.1. Cross-Moments

The first order cross-moment is

$$E(\Pi_{i=1}^{n}X_i) = \lambda\sqrt{2/\pi}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\left(\Pi_{i=1}^{n-1}x_i^2\phi(x_i)/\sqrt{1+\lambda^2\Pi_{i=1}^{n-1}x_i^2}\right)\Pi_{i=1}^{n-1}(dx_i). \tag{30}$$

As $|\lambda|$ increases the limiting value of Equation (30) is $\pm(2/\pi)^{n/2}$. There are expressions for higher order multivariate moments which are similar to those at Equations (7) and (8). If the $\{p_i\}; i = 1,\ldots,n$ are all even then property (4) implies that

$$E\left(\Pi_{i=1}^{n}X_i^{p_i}\right) = \Pi_{i=1}^{n}\Gamma\left(\frac{p_i+1}{2}\right)/2^{P/2}\Gamma(1/2)^n; P = \Sigma_{i=1}^{n}p_i. \tag{31}$$

If any p_i is even and any other is odd then the expectation at the left hand side of Equation (31) is zero. For the general multivariate case with all $\{p_i\}$ odd , higher order cross-moments may be computed in principle using a version of the extension to Stein's lemma reported in [13], as follows:

Proposition 5. ***Extension to Stein's lemma ([13])***

Let $\mathbf{X}^T = (X_1, X_2, \ldots, X_n)$ and let $g(\mathbf{X}) : \mathbb{R}^n \to \mathbb{R}$ be a function that is differentiable almost everywhere. Noting that

$$\Pi_{i=1}^{n}\phi(x_i) = \phi_n(\mathbf{x}),$$

is the density function of the standard multivariate normal distribution evaluated at $\mathbf{X} = \mathbf{x}$, the extension to Stein's lemma states that

$$cov\{\mathbf{X}, g(\mathbf{X})\} = E\{\nabla g(\mathbf{X})\} + 2\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}g(\mathbf{x})\nabla\Phi(\lambda\Pi_{i=1}^{n}x_i)\phi(\mathbf{x})d\mathbf{x},$$

where E denotes expectation taken over the skew-normal copula density function at Equation (28).

The gradient vector of $\Phi\left(\lambda \Pi_{i=1}^{n} x_i\right)$ is

$$\nabla\Phi(\lambda\Pi_{i=1}^{n}x_i) = \lambda \begin{bmatrix} \Pi_{j=2}^{n} X_j \\ \vdots \\ \Pi_{j\neq i} X_j \\ \vdots \\ \Pi_{j=1}^{n-1} X_j \end{bmatrix} \phi(\lambda\Pi_{i=1}^{n}x_i). \tag{32}$$

When

$$g(\mathbf{X}) = X_1^{(p_1-1)}\Pi_{i=2}^{n} X_i^{p_i}; p_i \in \mathbb{I}^+, \tag{33}$$

∇g is given by

$$\nabla g(\mathbf{X}) = \begin{bmatrix} (p_1-1)X_1^{p_1-2}\Pi_{j=2}^{n}X_j^{p_j} \\ \vdots \\ p_i X_1^{p_1-1} X_i^{p_i-1}\Pi_{j\neq 1,i}X_j^{p_j} \\ \vdots \\ p_n X_1^{p_1-1} X_n^{p_n-1}\Pi_{j=2}^{n-1}X_j^{p_j} \end{bmatrix}. \tag{34}$$

The first term of the vector $cov\{\mathbf{X}, g(\mathbf{X})\}$ recovers $E\left(\Pi_{i=1}^{n} X_i^{p_i}\right)$. Use of the lemma in Proposition 5 shows that

$$\begin{aligned} E\left(\Pi_{i=1}^{n} X_i^{p_i}\right) = (p_1-1)E\left(X_1^{p_1-2}\Pi_{i=2}^{n}X_i^{p_i}\right) + \\ 2\lambda\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_1^{p_1-1}\Pi_{i=2}^{n}x_i^{p_i+1}\Pi_{i=1}^{n}\{\phi(x_i)\}\phi(\lambda\Pi_{i=1}^{n}x_i)d\mathbf{x}. \end{aligned} \tag{35}$$

Note that the second term is equal to zero if p_1 is even. Otherwise it may be reduced to an integral in $n-1$ dimensions

$$\sqrt{\frac{2}{\pi}}\lambda\alpha_{p_1-1}\int_{0}^{\infty}\cdots\int_{0}^{\infty}\frac{\Pi_{i=2}^{n}y_i^{(p_i+1)/2}}{\left(1+\lambda^2\Pi_{i=2}^{n}y_i\right)^{p_1/2}}\Pi_{i=2}^{n}f(y_i)d\mathbf{x}$$

with

$$\alpha_{p_1-1} = \Gamma(p_1/2)2^{(p_1-1)/2}/\Gamma(1/2),$$

and the variables Y_i each independently distributed as $\chi^2_{(1)}$.

Note that the second and subsequent elements of $cov\{\mathbf{X}, g(\mathbf{X})\}$ will recover other cross-moments.

5.2. Skew Distributions Generated by Conditioning

Similar to the results in Section 3 a skewed distribution may be obtained by conditioning on $X_1 \leq y$. The distribution has the density function

$$f(x_i; i = 2,\ldots,n|X_1 \leq y) = \Pi_{i=2}^{n}\phi(x_i)\int_{-\infty}^{y}\phi(s)\Phi(\lambda s\Pi_{i=2}^{n}x_i)ds/\Phi(y). \tag{36}$$

Using arguments similar to those in Section 3, it then follows that for $\lambda > 0$ and as $y \to -\infty$ the distribution of the $n-1$ vector $\{X_i\}; i = 2,\ldots,n$ is also a skew-normal copula with shape parameter λy. Similarly the distribution of X_1 given $X_i \leq x_i; i = 2,\ldots,n$ with $X_i \to -\infty$ for at least one value of i is skew-normal with shape parameter $\lambda\Pi_{i=2}^{n}x_i$ There are analogous results for $\lambda < 0$ and $y \to \infty$.

5.3. Distribution Function and Related Computations

Details are omitted, but for all X_i substantially less than zero, an approximation to the distribution function is

$$P[\cap_{i=1}^{n}(X_i < x_i)] \simeq 2\Pi_{i=1}^{n}\Phi(x_i)\Phi\left(\lambda\Pi_{i=1}^{n}x_j\right), \tag{37}$$

from which VaR denoted X^* as before may be computed. Similar to the bivariate case defined at Equation (78), CVaR is defined in general as

$$E(\mathbf{X}|\mathbf{X} < X^*\mathbf{1}), \tag{38}$$

where $\mathbf{1}$ is an n-vector of ones. For a single variable, CVaR defined as

$$E(X_i|\mathbf{X} < X^*\mathbf{1}) \tag{39}$$

is approximately equal to VaR. As before, tail dependence is zero.

Similar to the bivariate case, the distribution of $Y = \Pi_{i=1}^{n}X_i$, when the X_i are independently distributed as $N\left(0, \sigma_i^2\right)$ has the density function

$$f(y) = G_{0,n}^{n,0}\left(x^2/2^n\sigma|0\right)/(2\pi)^{n/2}\sigma; \sigma = \Pi_{i=1}^{n}\sigma_i, \tag{40}$$

where $G(.)$ denotes Meijer's G-function. As above, see [15] for further details.

5.4. Extended Distributions

Similar to the results in Section 4, the skew-normal copula for n variables has an extended form. The multivariate extended skew-normal copula distribution has the density function

$$f(\mathbf{x}) = \Pi_{i=1}^{n}\phi(x_i)\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda\Pi_{i=1}^{n}x_i\right)/\Omega_n(\tau,\lambda), \tag{41}$$

where $\Omega_n(\tau,\lambda)$ is the normalizing constant. Integrating with respect to, say, X_1 shows that Ω is given by

$$\Omega_n(\tau,\lambda) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\Pi_{i=2}^{n}\phi(x_i)\Phi(\hat{\tau})\Pi_{i=2}^{n}dx_i; \hat{\tau} = \tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2\Pi_{i=2}^{n}x_i^2}. \tag{42}$$

In principle this may be reduced to a one-dimensional integral of the form

$$\Omega_n(\tau,\lambda) = \int_0^{\infty} f(s)\Phi(\hat{\tau}_s)ds; \hat{\tau}_s = \tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 s}, \tag{43}$$

where the scalar variable S is distributed as the product of $n-1$ independent variables each distributed as $\chi^2_{(1)}$. As already noted, the density function of S given by Fox's H-function, ref ([12]). The effect of non-zero τ is to induce dependence in the marginal distributions. The marginal distribution of $\bar{\mathbf{X}}^T = \left(X_1, \ldots, X_p\right); p < n$ has the symmetric density function

$$f(\bar{\mathbf{x}}) = \Pi_{i=1}^{p}\phi(x_i)\frac{\Omega_{n-p}\left(\tau_p^*, \lambda_p^*\right)}{\Omega_n(\tau,\lambda)}; \lambda_p^* = \lambda\Pi_{i=1}^{p}x_i, \tau_p^* = \tau\frac{\sqrt{1+\lambda^2}}{\sqrt{1+\lambda_p^{*2}}}, \tag{44}$$

Consequently the conditional distribution of $\hat{\mathbf{X}}^T = \left(X_{p+1}, \ldots, X_n\right); p < n$ given $\bar{\mathbf{X}} = \bar{\mathbf{x}}$ is $ESNC_{n-p}\left(\tau_p^*, \lambda_p^*\right)$. It is conjectured that, similar to the results in Propositions 2 and 4 that the conditional distribution of $\hat{\mathbf{X}}^T = \left(X_{p+1}, \ldots, X_n\right); p < n$ given $\bar{\mathbf{X}} << \mathbf{0}$ of the same type.

6. Skew-Student Copulas

Student's t distribution and its multivariate counterpart both arise as scale mixtures of the normal and multivariate normal distributions, respectively, as well as being sampling distributions in their own right. Similarly, the skew-Student distribution and its extended counterpart may be derived as scale mixtures. It is therefore natural to inquire whether there are parallel developments for the skew-normal copula distribution of Section 2 and subsequent sections. The potential attraction of such a development is the opportunity to have marginal Student's t distributions with differing degrees of freedom. In addition to the skew-Student distribution derived formally in [5], the earlier work in [18] suggests that more flexible constructions may also be contemplated. The first two sub-sections below therefore present two such approaches to skew-Student copulas. The third then describes a distributions that is derived as a scale mixture. In the interests of paper length, results are presented briefly, with further details available on request.

6.1. Skew-Student Copula—Case I

The first case has a density function given by

$$f(\mathbf{x}) = 2\Pi_{i=1}^{n} t_{\nu_i}(x_i) T_{\omega}(\lambda \Pi_{i=1}^{n} x_i), \tag{45}$$

where $t_\nu(.)$ and $T_\nu(.)$, respectively, denote the density and distribution functions of a Student's t variable with ν degrees of freedom. The univariate version of this distribution is referred to here as the linear skew-t. Allowing ω to increase without limit gives a distribution in which $T_\omega(.)$ is replaced by $\Phi(.)$. The properties of the distribution, which is denoted $SSC_{n,I}(\lambda; ^\circ, \omega)$, with $^{\circ\, T} = (\nu_1, \ldots \nu_n)$ are similar to those described in Sections 2 and 5, namely

1. The marginal distribution of each X_i is $t(\nu_i)$; that is, Student's t distribution with ν_i degrees of freedom.
2. The members of any subset of $\mathbf{X}$ of size $2, \ldots, n-1$ are independently distributed as $t(\nu_i)$.
3. The $X_i^2; i = 1, \ldots, n$ are independently distributed at $F(1, \nu_{(i)})$ variables.
4. If $\mathbf{X}$ is partitioned into two non-overlapping sets, $\{X_i\}; i = 1, \ldots, p$ and $\{X_j\}; j = p+1, \ldots, p+q$ with $p + q \leq n$ then (i) the $\{X_i^2\}$ are independently distributed as $F(1, \nu_{(i)})$ variables independently of (ii) the $\{X_j\}$, which are themselves independently distributed as t_{ν_i}.
5. The distribution of X_i given $X_j = x_j, j \neq i$ has the density function
$$f(x_i) = 2t_{\nu_i}(x_i) T_\omega \left(\lambda x_i \Pi_{j \neq i}^{n} x_j \right), \tag{46}$$
that is, a linear skew-t. There is a similar result for the conditional distribution of $X_j, j \neq i$ given $X_i = x_i$.

A conditional distribution is obtained in the same manner as that at Equation (11), namely by conditioning on $X_1 \leq y$. The resulting density function is

$$f(\mathbf{x}) = 2\Pi_{i=2}^{n} t_{\nu_i}(x_i) \int_{-\infty}^{y} t_{\nu_1}(s) T_\omega(\lambda s \Pi_{i=2}^{n} x_i) ds / T_{\nu_1}(y). \tag{47}$$

Cross-moments may be computed using recursions similar in principle to those in Section 2.1. Similar to the results in Section 3, for $\lambda > 0$ and $y << 0$ it may be shown that conditional on $X_1 \leq y << 0$ the remaining variables have an asymptotic skew-Student copula distribution of the same type as that at Equation (45), with shape parameter λy. Similarly the distribution of X_1 given $X_i \leq x_i; i = 2, \ldots, n$ with $x_i << 0$ for at least one value of i is linear skew-Student with shape parameter $\lambda \Pi_{i=2}^{n} x_i$. There are analogous results for $\lambda < 0$ and $y >> 0$.

The distribution defined at Equation (45) does not lead easily to an extended version. Although there are analytic expressions for the distribution of the difference of two independent Student's t variates, the expressions are complicated—see for example [19,20]. This remains a topic for future research.

6.2. Skew-Student Copula—Case II

A second skew-Student copula distribution, denoted $SSC_{n,II}(\lambda; {}^{\circ}, \omega)$, has the density function

$$f(\mathbf{x}) = 2\Pi_{i=1}^{n} t_{\nu_i}(x_i) T_{\omega+1}\left(\sqrt{\omega+1}\lambda\Pi_{i=1}^{n}\frac{x_i}{\sqrt{\nu_i + x_i^2}}\right) \tag{48}$$

When $n = 1$ and $\omega = \nu_1$ this is the skew-t distribution due originally to [5]. To distinguish it from the linear form above this is referred to as the Azzalini skew-t. The properties of the distribution at Equation (48) are essentially the same as those listed in Section 6.1. The asymptotic conditional distribution of $X_j; j = 2, \ldots, n$ given that $X_1 \leq x_1 << 0$ is $SSC_{n-1,II}(\hat{\lambda}; \tilde{N}, \omega)$ with

$$\tilde{N}^T = (\nu_2, \ldots, \nu_n); \hat{\lambda} = \frac{\lambda x_1}{\sqrt{\nu + x_1^2}}.$$

There is a similar result for the conditional distribution of X_1 given that $X_j \leq x_j << 0; j = 2, \ldots, n$.

6.3. Skew-Student Copula—Case III

In the third case, which is arguably a more realistic representation, conditional on n mixing variables $S_i = s_i$, collectively $\mathbf{S} = \mathbf{s}$, the joint density function of the variables X_i is

$$f(\mathbf{x}|\mathbf{S} = \mathbf{s}) = 2\Pi_{i=1}^{n} s_i^{1/2} \Pi_{i=1}^{n} \phi(x_i\sqrt{s_i})\Phi\left(\lambda\Pi_{i=1}^{n} s_i^{1/2} \Pi_{i=1}^{n} x_i\right), \tag{49}$$

with each S_i independently distributed as $\chi^2_{(\nu_i)}/\nu_i$. For this case the distribution of $\mathbf{X}$ has density function

$$f(\mathbf{x}) = 2\Pi_{i=1}^{n} t_{\nu_i}(x_i) M_{n,N}\left(\lambda\Pi_{i=1}^{n}\frac{x_i\sqrt{\nu_i+1}}{\sqrt{\nu_i + x_i^2}}\right), \tag{50}$$

where $M_{n,N}(x)$ denotes the distribution function, evaluated at x, of the variable $V = Z/W$ where $Z \sim N(0,1)$ independently of $W = \Pi_{i=1}^{n} U_i^{1/2}$ with U_i independently distributed as $\chi^2_{(\nu_i+1)}$. The density function corresponding to $M(.)$ is given by Fox's H-function, see [12]. It is interesting to note that the form of Equation (50) implies that the function M is the distribution of the variable V that is symmetric. The scale mixture also does not add great additional complications to the expressions for cross-moments, although, as above, numerical computation is required. The distribution shares properties 1 to 4 of the case I distribution reported in Section 6.1. This distribution does not appear in the literature and derivation of the distribution and density functions are future research tasks.

7. Multivariate Distributions

When $\mathbf{X}$ and $\mathbf{Y}$ are both n-vectors, a basic multivariate version of the SN-copula distribution has the density function

$$f(\mathbf{x}, \mathbf{y}) = 2\phi_n(\mathbf{x})\phi_n(\mathbf{y})\Phi\left(\mathbf{x}^T\mathbf{y}\right), \tag{51}$$

where $\phi_n(.)$ denotes the density function of the $N_n(\mathbf{0}, \mathbf{I})$ distribution. The marginal distributions of $\mathbf{X}$ and $\mathbf{Y}$ are each standard multivariate normal $N_n(\mathbf{0}, \mathbf{I})$. The conditional distribution of $\mathbf{X}$ given that $\mathbf{Y} = \mathbf{y}$ is multivariate skew-normal with density function

$$f(\mathbf{x}) = 2\phi_n(\mathbf{x})\Phi\left(\mathbf{x}^T\mathbf{y}\right). \tag{52}$$

This section of the paper briefly describes the basic properties of the distribution at Equation (51) and extensions thereof. Details of more advanced developments, such as conditional distributions similar in concept to those described in earlier sections are left as topics for future development. In private correspondence Loperfido, ref [21], has proposed an extension of the distribution at Equation (51). In this extension the scale matrices of $\mathbf{X}$ and $\mathbf{Y}$ are not restricted to be unit matrices and the location parameters are not restricted to be zero vectors. Using their notation, the joint distribution of the random vectors $\mathbf{X}$ and $\mathbf{Y}$ has the density function

$$f(\mathbf{x}, \mathbf{y}; \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\lambda}) = 2\tilde{\phi}_n(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\tilde{\phi}_n(\boldsymbol{y}; \boldsymbol{v}, \boldsymbol{\Omega})\Phi\left\{\sum_{i=1}^{n}\lambda_i(x_i - \mu_i)(y_i - v_i)\right\}, \tag{53}$$

where $\tilde{\phi}_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\tilde{\phi}_n(\boldsymbol{y}; \boldsymbol{v}, \boldsymbol{\Omega})$ are the density functions of vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ distributed, respectively, as $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $N_n(\boldsymbol{v}, \boldsymbol{\Omega})$ and where $\Phi(\cdot)$ is as defined above. In related correspondence, ref [22] notes that an alternative argument for Φ in Equation (51) is $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}$, where $\boldsymbol{A}$ is an $n \times m$ matrix and $\boldsymbol{y}$ is an m-vector. Setting $\boldsymbol{\mu}$ and $\boldsymbol{v}$ to $\mathbf{0}$, a vector of zeros of appropriate length, this leads to the density function

$$f(\boldsymbol{x}, \boldsymbol{y}; \mathbf{0}, \mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \boldsymbol{\lambda}) = 2\tilde{\phi}_n(\boldsymbol{x}; \boldsymbol{\Sigma})\tilde{\phi}_m(\boldsymbol{y}; \boldsymbol{\Omega})\Phi\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right). \tag{54}$$

In the same correspondence [21] considers a generalization of this multivariate distribution with density function

$$f(\boldsymbol{x}, \boldsymbol{y}) = 2h_n(\boldsymbol{x} - \boldsymbol{\mu})k_m(\boldsymbol{y} - \boldsymbol{v})\pi\left\{(\boldsymbol{y} - \boldsymbol{v})^T\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{\mu})\right\}, \tag{55}$$

where $h_n(\cdot)$ and $k_m(\cdot)$ are the density functions of n and m dimensional elliptically symmetric distributions. The function $\pi(\cdot)$ satisfies $0 \le \pi(-a) = 1 - \pi(a) \le 1$ for all $a \in \mathbb{R}$. The distribution at Equation (55) is an extension of the generalized skew-normal described in [7,23].

7.1. Marginal and Conditional Distributions—A Basic Example

To illustrate the properties of the marginal and conditional distributions consider the basic bivariate case. Let the density function of the four random variables X_1, X_2, Y_1 and Y_2 be

$$f(x_1, x_2, y_1, y_2) = 2\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\Phi(\lambda_1x_1y_1 + \lambda_2x_2y_2), \tag{56}$$

Integration with respect to X_2 gives

$$f(x_1, y_1, y_2) = 2\phi(x_1)\phi(y_1)\phi(y_2)\Phi\left\{\frac{\lambda_1x_1y_1}{\sqrt{1 + \lambda_2^2y_2^2}}\right\}. \tag{57}$$

Integration of Equation (57) with respect to X_1 gives the joint density of $Y_i; i = 1, 2$

$$f(y_1, y_2) = \phi(y_1)\phi(y_2), \tag{58}$$

as expected. Similarly, integration of Equation (57) with respect to Y_1 gives the marginal density joint density of X_1 and Y_2

$$f(x_1, y_2) = \phi(x_1)\phi(y_2). \tag{59}$$

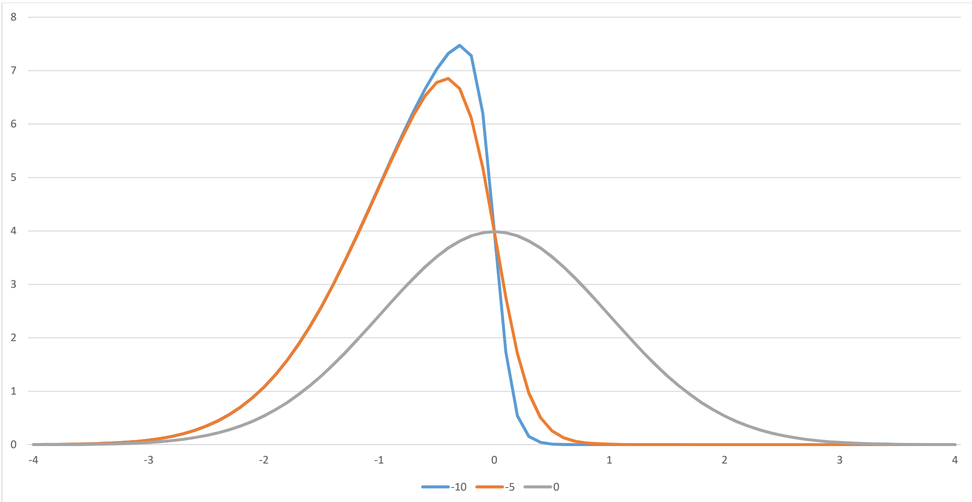
However, the marginal density joint density of X_1 and Y_1 is given by

$$f(x_1, y_1) = 2\phi(x_1)\phi(y_1)\int_{-\infty}^{\infty} \phi(y_2)\Phi\left\{\frac{\lambda_1 x_1 y_1}{\sqrt{1+\lambda_2^2 y_2^2}}\right\}dy_2. \tag{60}$$

To the best of my knowledge there is no closed form expression for the integral at Equation (60). The conditional distribution of X_1 given that $Y_1 = y_1$ has the PDF

$$f(x_1|Y_1 = y_1) = 2\phi(x_1)\int_{-\infty}^{\infty} \phi(y_2)\Phi\left\{\frac{\lambda_1 x_1 y_1}{\sqrt{1+\lambda_2^2 y_2^2}}\right\}dy_2. \tag{61}$$

Example densities for $\lambda_1 = \lambda_2 = 1$ and $Y_1 = -10, -5$ and 0 are shown in Figure 5. There is a corresponding expression for the PDF of Y_1 given that $X_1 = x_1$.



The figure shows conditional density functions that resemble the skew-normal for conditioning values equal to -10 and -5 for λ₁ = λ₂ =1. Conditioning on 0 gives the standard normal density.

Figure 5. Example Conditional Density Functions; $\lambda_{1,2} = 1$, $Y_1 = -10, -5, 0$.

7.2. Marginal and Conditional Distributions—General Results

For the distribution at Equation (54) suppose that $\boldsymbol{X}$ is partitioned into two components of length n_1 and $n_2 = n - n_1$ and that $\boldsymbol{A}$ and $\boldsymbol{\Sigma}$ are partitioned similarly. Thus

$$\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{A}_1\boldsymbol{x}_1 + \boldsymbol{y}^T\boldsymbol{A}_2\boldsymbol{x}_2, \tag{62}$$

and write $\tilde{\phi}_n(\boldsymbol{x};\boldsymbol{\Sigma})$ as

$$\tilde{\phi}_n(\boldsymbol{x};\boldsymbol{\Sigma}) = \tilde{\phi}_{n_1}(\boldsymbol{x}_1;\boldsymbol{\Sigma}_{11})\tilde{\phi}_{n_2}\left(\boldsymbol{x}_2 - \boldsymbol{B}\boldsymbol{x}_1;\boldsymbol{\Sigma}_{2|1}\right), \tag{63}$$

with $\boldsymbol{B} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}$ and $\boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$. The argument of $\Phi(.)$ may be written as

$$\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^T(\boldsymbol{A}_1 + \boldsymbol{A}_2\boldsymbol{B})\boldsymbol{x}_1 + \boldsymbol{y}^T\boldsymbol{A}_2(\boldsymbol{x}_2 - \boldsymbol{B}\boldsymbol{x}_1) \tag{64}$$

Integration over $\boldsymbol{x}_2$ yields the density of the joint distribution of $\boldsymbol{X}_1$ and $\boldsymbol{Y}$

$$f(\boldsymbol{x}_1, \boldsymbol{y}) = 2\tilde{\phi}_{n_1}(\boldsymbol{x}_1;\boldsymbol{\Sigma}_{11})\tilde{\phi}_m(\boldsymbol{y};\boldsymbol{\Omega})\Phi\left\{\frac{\boldsymbol{y}^T(\boldsymbol{A}_1 + \boldsymbol{A}_2\boldsymbol{B})\boldsymbol{x}_1}{\sqrt{1+\boldsymbol{y}^T\boldsymbol{A}_2\boldsymbol{\Sigma}_{2|1}\boldsymbol{A}_2^T\boldsymbol{y}}}\right\}. \tag{65}$$

Integration over $\boldsymbol{X}_1[\boldsymbol{Y}]$ recovers the normal distribution of $\boldsymbol{Y}[\boldsymbol{X}_1]$ as expected. The conditional distribution of $\boldsymbol{X}_1$ given $\boldsymbol{Y} = \boldsymbol{y}$ is skew-normal with shape parameter

$$\frac{(\boldsymbol{A}_1 + \boldsymbol{A}_2\boldsymbol{B})^T\boldsymbol{y}}{\sqrt{1 + \boldsymbol{y}^T\boldsymbol{A}_2\boldsymbol{\Sigma}_{2|1}\boldsymbol{A}_2^T\boldsymbol{y}}}.$$

There is no closed form expression for the joint (marginal) density of $\boldsymbol{X}_1$ and $\boldsymbol{Y}_i; i = 1$ or $= 2$, except in the very obscure special case for which $\boldsymbol{A}_2\boldsymbol{\Sigma}_{2|1}\boldsymbol{A}_2^T = \boldsymbol{0}$.

7.3. Extended Version

Following the approach used in Sections 4 and 5.4, an extended version of the distribution at Equation (54) has the density function

$$f(\boldsymbol{x}, \boldsymbol{y}) = 2\tilde{\phi}_n(\boldsymbol{x}; \boldsymbol{\Sigma})\tilde{\phi}_m(\boldsymbol{y}; \boldsymbol{\Omega})\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)/\Psi(\tau, \lambda), \tag{66}$$

where $\Psi(\tau, \lambda)$ is the normalizing constant. Integration with respect to $\boldsymbol{x}$ or $\boldsymbol{y}$ shows that this is given by the n-dimensional integral

$$\Psi(\tau, \lambda) = \int_{-\infty}^{\infty} \tilde{\phi}_n(\boldsymbol{x}; \boldsymbol{\Sigma})\Phi\left(\frac{\tau\sqrt{1+\lambda^2}}{\sqrt{1+\lambda^2\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}^T\boldsymbol{x}}}\right)d\boldsymbol{x},$$

or alternatively by the m-dimensional integral

$$\Psi(\tau, \lambda) = \int_{-\infty}^{\infty} \tilde{\phi}_m(\boldsymbol{y}; \boldsymbol{\Omega})\Phi\left(\frac{\tau\sqrt{1+\lambda^2}}{\sqrt{1+\lambda^2\boldsymbol{y}^T\boldsymbol{A}^T\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{y}}}\right)d\boldsymbol{y}.$$

7.4. Student Version

In the usual way, consider the distribution of $\boldsymbol{X}$ and $\boldsymbol{Y}$ conditional on $S_i = s_i; i = 1, 2$ where the S_i are independently distributed as $\chi^2_{(\nu_i)}/\nu_i$. The conditional density function is

$$f(\boldsymbol{x}, \boldsymbol{y}|S_i = s_i; i = 1, 2) = 2\tilde{\phi}_n(\boldsymbol{x}; \boldsymbol{\Sigma}/\sqrt{s_1})\tilde{\phi}_m(\boldsymbol{y}; \boldsymbol{\Omega}/\sqrt{s_2})\Phi\left(\lambda\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x}\sqrt{s_1 s_2}\right) \tag{67}$$

Standard manipulations give the following expression for the density function of $\boldsymbol{X}$ and $\boldsymbol{Y}$

$$f(\boldsymbol{x}, \boldsymbol{y}) = 2\tilde{t}_{\nu_1,n}(\boldsymbol{x}, \boldsymbol{\Sigma})\tilde{t}_{\nu_2,m}(\boldsymbol{y}, \boldsymbol{\Omega})E_{S^*}\left[\Phi\left\{\frac{\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\sqrt{s_1 s_2(\nu_1+1)(\nu_1+1)}}{\sqrt{\nu_2 + \boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}\sqrt{\nu_2 + \boldsymbol{y}^T\boldsymbol{\Omega}^{-1}\boldsymbol{y}}}\right\}\right], \tag{68}$$

where E_{S^*} denotes expectation over the distribution of the variables $S_i; i = 1, 2$ which are independently distributed as $\chi^2_{(\nu_i+1)}/(\nu_i + 1)$ variables and $\tilde{t}_{\nu,n}(\boldsymbol{x}, \boldsymbol{\Sigma})$ denotes the density function of an n-variate multivariate Student distribution with ν degrees of freedom , location parameter vector $\boldsymbol{0}$ and scale matrix $\boldsymbol{\Sigma}$. Note that integration over, say, S_1 reduces the right hand side of Equation (68) to

$$2\tilde{t}_{\nu_1,n}(\boldsymbol{x}, \boldsymbol{\Sigma})\tilde{t}_{\nu_2,m}(\boldsymbol{y}, \boldsymbol{\Omega})E_{S_2}\{T_{\nu_1+1}(R)\}, R = \frac{\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\sqrt{s_2(\nu_1+1)(\nu_1+1)}}{\sqrt{\nu_2 + \boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}\sqrt{\nu_2 + \boldsymbol{y}^T\boldsymbol{\Omega}^{-1}\boldsymbol{y}}}, \tag{69}$$

which may be computed numerically for given values of $\boldsymbol{x}$ and $\boldsymbol{y}$. An extended version of the distribution may be developed in the same way as in Section 7.3. The expression at (69) becomes

$$\tilde{t}_{\nu_1,n}(\boldsymbol{x}, \boldsymbol{\Sigma})\tilde{t}_{\nu_2,m}(\boldsymbol{y}, \boldsymbol{\Omega})E_{S_2}\left[T_{\nu_1+1}\left\{\tau\sqrt{1+\lambda^2} + R\right\}\right]/\Psi(\tau, \lambda), \tag{70}$$

Given the need for numerical computation indicated at Equations (68) and (69), an alternative skew-Student copula may be obtained by extending the distribution described in Section 6.1. The density function is

$$f(\boldsymbol{x},\boldsymbol{y}) = 2\tilde{t}_{\nu_1,n}(\boldsymbol{x},\boldsymbol{\Sigma})\tilde{t}_{\nu_2,m}(\boldsymbol{y},\boldsymbol{\Omega})T_\omega\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right).$$

An alternative version is

$$f(\boldsymbol{x},\boldsymbol{y}) = 2\tilde{t}_{\nu_1,n}(\boldsymbol{x},\boldsymbol{\Sigma})\tilde{t}_{\nu_2,m}(\boldsymbol{y},\boldsymbol{\Omega})T_\omega(Q); Q = \frac{\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\sqrt{(\nu_1+1)(\nu_1+1)}}{\sqrt{\nu_2+\boldsymbol{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{x}}\sqrt{\nu_2+\boldsymbol{y}^T\boldsymbol{\Omega}^{-1}\boldsymbol{y}}}.$$

Neither of these distributions, however, lead to easily tractable extended versions.

7.5. Stein's Lemma

This lemma is useful in Portfolio theory and for the computation of moments and cross moments. The treatment in this sub-section follows that in [13].

Let $g(\boldsymbol{X},\boldsymbol{Y})$ be a scalar valued function of $\boldsymbol{X}$ and $\boldsymbol{Y}$ subject to the usual regularity conditions and consider

$$E\{\boldsymbol{X}g(\boldsymbol{X},\boldsymbol{Y})\} = 2\int_{\boldsymbol{y}}\int_{\boldsymbol{x}} \boldsymbol{x}g(\boldsymbol{x},\boldsymbol{y})\phi_n(\boldsymbol{x};\boldsymbol{\Sigma})\phi_m(\boldsymbol{y};\boldsymbol{\Omega})\Phi\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)d\boldsymbol{x}d\boldsymbol{y}. \tag{71}$$

The right hand side of Equation (71) is

$$2E_Y\left[\Sigma E_X\left\{\nabla_{\boldsymbol{x}}g(\boldsymbol{x},\boldsymbol{y})\Phi\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right) + g(\boldsymbol{x},\boldsymbol{y})\lambda\boldsymbol{A}\boldsymbol{y}\phi\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)\right\}\right]. \tag{72}$$

This is

$$\Sigma\left[E\{\nabla_{\boldsymbol{x}}g(\boldsymbol{x},\boldsymbol{y})\} + 2E_Y\left\{\lambda\boldsymbol{A}^T\boldsymbol{y}\int_{\boldsymbol{x}} g(\boldsymbol{x},\boldsymbol{y})\phi_n(\boldsymbol{x};\boldsymbol{\Sigma})\phi\left(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)dx\right\}\right]. \tag{73}$$

The second term is

$$\sqrt{\frac{2}{\pi}}E_Y\left[\frac{\boldsymbol{A}\boldsymbol{y}}{\sqrt{1+\boldsymbol{y}^T\boldsymbol{A}^T\Sigma^{-1}\boldsymbol{A}\boldsymbol{y}}}\int_{\boldsymbol{x}} g(\boldsymbol{x},\boldsymbol{y})\phi_n\left\{\boldsymbol{x};\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{A}\boldsymbol{y}\boldsymbol{y}^T\boldsymbol{A}^T\right)^{-1}\right\}d\boldsymbol{x}\right] \tag{74}$$

with a similar expression for $E\{\boldsymbol{Y}g(.)\}$.

Example 1. *Bivariate case*

Let $n = m = 1$, $\boldsymbol{\Sigma} = \boldsymbol{\Omega} = 1$, $\boldsymbol{A} = 1$ and $g(X,Y) = Y$. Then

$$E\{Xg(X,Y)\} = E(XY) = cov(X,Y).$$

On using the lemma, $\nabla_{\boldsymbol{x}}g(x) = 0$ and the second term above is

$$\sqrt{\frac{2}{\pi}}E_Y\left[\frac{\lambda y}{\sqrt{1+y^2\lambda^2}}\int_{-\infty}^{\infty} y\phi_1\left\{x;\left(1+\lambda^2y^2\right)\right\}dx\right] = \sqrt{\frac{2}{\pi}}E_Y\left[\frac{\lambda y^2}{\sqrt{1+y^2\lambda^2}}\right],$$

which agrees with Equation (4).

Example 2. *General Case*

For the general case, $g(\boldsymbol{x},\boldsymbol{y}) = y_i$ for $i = 1,\ldots,m$. As above, $\nabla_{\boldsymbol{x}}g(x) = 0$ and the second term in the lemma is

$$\boldsymbol{\Sigma}\sqrt{\frac{2}{\pi}}E_Y\left[\frac{\boldsymbol{A}^T\boldsymbol{y}y_i}{\sqrt{1+\boldsymbol{y}^T\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{y}}}\right].$$

Hence, the cross covariance matrix is

$$E\left(\boldsymbol{X}\boldsymbol{Y}^T\right) = \boldsymbol{\Sigma}\boldsymbol{A}^T\sqrt{\frac{2}{\pi}}E_{\boldsymbol{Y}}\left[\frac{\boldsymbol{y}\boldsymbol{y}^T}{\sqrt{1+\boldsymbol{y}^T\boldsymbol{A}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{y}}}\right].$$

This expression must be computed numerically and it must equal

$$\sqrt{\frac{2}{\pi}}E_{\boldsymbol{X}}\left[\frac{\boldsymbol{x}\boldsymbol{x}^T}{\sqrt{1+\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{\Omega}^{-1}\boldsymbol{A}\boldsymbol{x}}}\right]\boldsymbol{A}\boldsymbol{\Omega}.$$

Note also that higher order cross moments, that is, $E\left(X_i^p Y_j^q\right)$*, may also be computed using Stein's lemma, albeit with numerical integration.*

Example 3. *Portfolio Selection*
For portfolio selection assume that $\boldsymbol{X}$ *denotes asset returns and that* $\boldsymbol{Y}$ *denotes sources of skewness in the conditional distribution of* $\boldsymbol{X}$ *given that* $\boldsymbol{Y}=\boldsymbol{y}$*. This model reflects an empirical feature of some markets, namely that skewness may be time varying. The return on a portfolio with weights* $\boldsymbol{w}$ *is* $\boldsymbol{w}^T\boldsymbol{X}$*. If the utility function is* $U(\boldsymbol{w}^T\boldsymbol{X})$ *the first order conditions for portfolio selection conditional on* $\boldsymbol{Y}=\boldsymbol{y}$ *contain the term*

$$E_{\boldsymbol{X}}\left\{\boldsymbol{X}U'\left(\boldsymbol{w}^T\boldsymbol{X}\right)\right\}$$

Hence $g(\boldsymbol{X}) = U'(\boldsymbol{w}^T\boldsymbol{X})$ *and* $\nabla g_{\boldsymbol{X}}(\boldsymbol{X}) = \boldsymbol{w}U''(\boldsymbol{w}^T\boldsymbol{X})$*. Stein's lemma yields*

$$\boldsymbol{\Sigma}\boldsymbol{w}E\left\{U''\left(\boldsymbol{w}^T\boldsymbol{X}\right)\right\} + 2\boldsymbol{\Sigma}E_{\boldsymbol{Y}}\left\{\lambda\boldsymbol{A}\boldsymbol{Y}\int_{\boldsymbol{x}}U'\left(\boldsymbol{w}^T\boldsymbol{x}\right)\phi_n(\boldsymbol{x};\boldsymbol{\Sigma})\phi\left(\lambda\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)d\boldsymbol{x}\right\}.$$

Assuming that the order of integration may be changed the second term is proportional to

$$\left\{\boldsymbol{A}\int_{\boldsymbol{y}}\boldsymbol{y}\phi_m(\boldsymbol{y};\boldsymbol{\Omega})\phi\left(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}\right)d\boldsymbol{y}\right\}.$$

which equals zero. Thus, portfolio selection results in a portfolio on the efficient frontier, as expected. Note that if expectations are taken over the conditional distribution of $\boldsymbol{X}$ *given* $\boldsymbol{Y}=\boldsymbol{y}$ *the result is the same as in [24] with appropriate changes of notation.*

8. Three Examples

This section of the paper contains three numerical examples, the purpose of which is to illustrate some aspects of the skew-normal copula and related distributions in action. The first presents results for the distribution function of the bivariate skew-normal copula of Section 2, focusing on asymptotic results for tail probability computations. Example two presents specimen estimation results for the bivariate skew-Student copula of Section 6.1. The final illustration has results for the multivariate skew-normal copula that is described in Section 7, specifically the distribution at Equation (53).

8.1. Bivariate Skew-Normal Copula Distribution Function and Related Computations

The distribution function corresponding to the density function at Equation (1) is

$$P(X_1 \le x_1, X_2 \le x_2) = 2\int_{-\infty}^{x_2}\int_{-\infty}^{x_1}\phi(s)\phi(t)\Phi(\lambda st)dsdt. \tag{75}$$

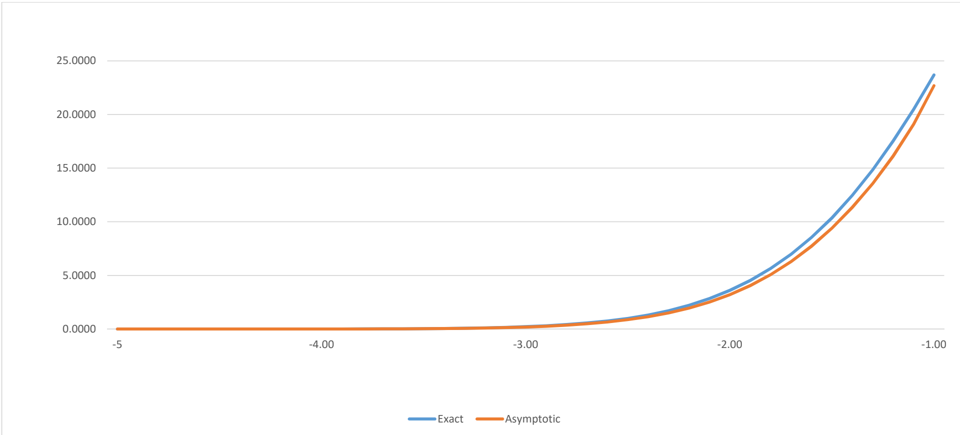
There is no analytic expression for this integral. For specified values of $x_{1,2}$ it may be computed numerically. For some applications, for example in finance, there is a requirement to compute the distribution function when both $X_{1,2}$ are substantially less than zero. For

$\lambda > 0$, $x_{1,2} << 0$ and ignoring power of $1/|x_{1,2}|$ greater than 2, the resulting asymptotic expression is

$$2\Phi(x_1)\Phi(x_2)\Phi(\lambda x_1 x_2) + 2\lambda \frac{\phi(x_1)\phi(x_2)\phi(\lambda x_1 x_2)}{x_1^2 x_2^2 (1 + x_1^2\lambda^2)(1 + x_2^2\lambda^2)} \left(x_1^2 + x_2^2 + \lambda^2 x_1^2 x_2^2 \right). \tag{76}$$

Figure 6 shows an example of the exact and approximate distribution function corresponding to $\lambda = 1$. The quantity tabulated is $\sqrt{P(X_1 \leq X^*, X_2 \leq X^*)}$ for values of X^* between -5.0 and -0.1. Similar results may be computed for $\lambda < 0$. It may be noted that there are combinations of values of λ, x_1 and x_2 under which the second term is negligible and $\Phi(\lambda x_1 x_2) \simeq 1$; that is

$$P(X_1 \leq x_1, X_2 \leq x_2) \simeq 2\phi(x_1)\phi(x_2).$$



The distribution function in blue is the square root of the probability that X and Y are both less than or equal to X*, for values of X* in the range [-5, -0.1], computed numerically. For comparison, the distribution function in orange shows the approximation given in Equation (11).

Figure 6. Exact and Approximate Left-hand Quadrant Probabilities; $\lambda = 1.0$.

The bivariate skew-normal copula distribution is not specifically suited for financial applications, but it may be used to compute Value at Risk, VaR henceforth. As the variables are standardized, VaR is a critical value in the left hand tail of the distribution X^* such that

$$P(X_1 < x^*, X_2 < x^*) \simeq 2\Phi(x^*)^2 \Phi\left(\lambda x^{*2}\right) = \alpha, \tag{77}$$

for a specified value of α that is small. Conditional Value at Risk, CVaR henceforth, is a related measure. For a single variable X, CVaR is defined as the expected value of X given that it is less than the VaR. For this distribution the properties listed above means that CVaR is the same as that based on the standard normal distribution. A bivariate version of CVaR is defined as

$$E(X_1, X_2 | X_1 \leq x^*, X_2 \leq x^*). \tag{78}$$

For X_1 alone this is

$$E(X_1 | X_1 \leq x^*, X_2 \leq x^*) = \alpha^{-1} 2 \int_{-\infty}^{x^*} \int_{-\infty}^{x^*} x_1 \phi(x_1)\phi(x_2)\Phi(\lambda x_1 x_2) dx_1 dx_2. \tag{79}$$

For $\lambda > 0$ and $x^* << 0$, similar arguments to those above lead to

$$E(X_1 | X_1 < x^*, X_2 < x^*) \simeq -\alpha^{-2}\phi(x^*) \left\{ P(X_2 \leq x^* | x^*) + \frac{2\lambda\phi(x^*)\phi(\lambda x^{*2})}{|x^*|(1 + x^{*2}\lambda^2)^2} \right\}, \tag{80}$$

where $P(X_2 \le x^*|x^*)$ denotes the distribution function of a standardized skew-normal distribution with shape parameter equal to λx^*. Using Equation (76) shows that tail dependence equals zero for the skew-normal copula distribution. A Selection of values of CVaR when $\lambda = 1$ is shown in Table 9 for critical values ranging from -9.5 to -0.5. The values shown in the second column were computed using numerical inegration. Those in the third column were computed using the asymptotic formula shown at Equation (80). It is suggested that the asymptotic formula leads to values that would be sufficiently accurate for practical purposes.

Table 9. Exact and Asymptotic Values of CVaR; $\lambda = 1$.

VaR (CV)	Computed Value	Asymptotic Value
−9.5	−9.5597	−9.6031
−9	−9.0664	−9.1085
−8.5	−8.5722	−8.6146
−8	−8.0787	−8.1214
−7.5	−7.5859	−7.6290
−7	−7.0942	−7.1375
−6.5	−6.6037	−6.6473
−6	−6.1146	−6.1585
−5.5	−5.6273	−5.6714
−5	−5.1423	−5.1865
−4.5	−4.6600	−4.7043
−4	−4.1813	−4.2256
−3.5	−3.7072	−3.7514
−3	−3.2392	−3.2831
−2.5	−2.7793	−2.8227
−2	−2.3307	−2.3732
−1.5	−1.8981	−1.9276
−1	−1.4976	−1.2545
−0.5	−1.1536	−0.2418

8.2. Bivariate Student t Copula

This example is based on the bivariate version of the skew-Student copula of Section 6.1 with $T_\omega(.)$ replaced by $\Phi(.)$. The density function at Equation (45) becomes

$$f(x_1, x_2) = 2t_{\nu_1}(x_1)t_{\nu_2}(x_2)\Phi(\lambda x_1 x_2). \tag{81}$$

First, note that for specified ranges of the shape parameter λ this distribution is bimodal. As in Section 2, bimodality requires that $|\lambda| > \sqrt{\pi/2}$. For the Student case the modal value of X and $\lambda > 0$) depends on the degrees of freedom and satisfies

$$\frac{\lambda\sqrt{\nu_1\nu_2}(1 + x^2/\nu_1)}{\sqrt{(\nu_1+1)(\nu_2+1)\Delta}}\xi_1\left(\frac{\lambda\sqrt{\nu_2(\nu_1+1)}x^2}{\sqrt{\nu_1(\nu_2+1)\Delta}}\right) = 1; \Delta = 1 + \frac{(\nu_2-\nu_1)x^2}{\nu_1(\nu_2+1)}.$$

Sets of computed values are shown in Table 10. The first column of the table shows values of λ. Panel 1 of the table shows the modal value for a set of cases in which the degrees of freedom are equal. Panel 2 shows corresponding values when $\nu_1 = 3$. For comparison purposes, column 2 of panel 2 shows the corresponding modal values for the bivariate skew-normal case.

As in Section 2 the correlation between the two variables is computed numerically. A selection of values is shown in Table 11. The first column of the table shows a range of values of λ. Panel 1 shows the computed correlation for a selection of cases for which $\nu_1 = \nu_2$. The last column shows the corresponding values for the skew-normal copula for comparison. The second panel shows a range of values when $\nu_1 = 3$. The results in the table indicate that the relationship between the shape parameter λ and the degrees of freedom is non-linear. Generally, however, correlation is reduced when the degrees of freedom are finite: a noteworthy difference from the bivariate Student distribution.

Table 10. Modal values under the bivariate Student copula.

Panel 1: Equal degrees of freedom						
	(3-3)	(5-5)	(7-7)	(10-10)	(20-20)	(50-50)
1.254	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.4827	0.5067	0.5169	0.5245	0.5333	0.5387
4	0.4857	0.5013	0.5082	0.5134	0.5195	0.5233
5	0.4714	0.4835	0.4889	0.4929	0.4979	0.5009
15	0.3442	0.3488	0.3509	0.3525	0.3545	0.3557
10	0.3943	0.4007	0.4035	0.4058	0.4085	0.4101
20	0.3101	0.3138	0.3155	0.3168	0.3183	0.3193
35	0.2498	0.2523	0.2534	0.2543	0.2553	0.2559
50	0.2162	0.2181	0.2190	0.2197	0.2205	0.2210
75	0.1826	0.1841	0.1847	0.1853	0.1859	0.1863
100	0.1616	0.1628	0.1633	0.1638	0.1643	0.1646
Panel 2: Unequal degrees of freedom						
	SNC	(3-5)	(3-7)	(3-10)	(3-20)	(3-50)
1.254	0.0231	0.0000	0.0000	0.0000	0.0000	0.0000
3	0.5422	0.4894	0.4922	0.4943	0.4966	0.498
4	0.5258	0.4865	0.4868	0.487	0.4871	0.4872
5	0.5029	0.4698	0.4689	0.4683	0.4675	0.4669
15	0.3565	0.3389	0.3365	0.3345	0.3323	0.3308
10	0.4112	0.3894	0.3871	0.3853	0.3832	0.3819
20	0.3200	0.3048	0.3024	0.3005	0.2982	0.2968
35	0.2564	0.2450	0.2428	0.2411	0.2390	0.2377
50	0.2213	0.2118	0.2098	0.2083	0.2064	0.2052
75	0.1865	0.1787	0.1770	0.1757	0.1740	0.1729
100	0.1648	0.1581	0.1565	0.1553	0.1538	0.1529

Table 11. Correlation under the Bivarate Student Copula.

Panel 1: Equal Degrees of Freedom						
	(3,3)	(7,7)	(10,10)	(20,20)	(50,50)	Normal
0.01	0.0211	0.0112	0.0100	0.0089	0.0083	0.0080
0.1	0.1369	0.1054	0.0962	0.0867	0.0817	0.0786
1	0.3484	0.4480	0.4527	0.4536	0.4523	0.4505
2	0.3799	0.5185	0.5322	0.5434	0.5482	0.5503
5	0.3972	0.5603	0.5798	0.5980	0.6071	0.6116
10	0.4009	0.5710	0.5919	0.6120	0.6223	0.6268
Panel 2: Unequal Degrees of Freedom						
	(3,3)	(3,5)	(3,7)	(3,10)	(3,20)	(3,50)
0.01	0.0211	0.0170	0.0157	0.0148	0.0140	0.0136
0.1	0.1369	0.1301	0.1250	0.1211	0.1167	0.1142
1	0.3484	0.3904	0.3980	0.4015	0.4040	0.4048
2	0.3799	0.4334	0.4452	0.4517	0.4573	0.4600
5	0.3972	0.4579	0.4724	0.4807	0.4885	0.4924
10	0.4009	0.4636	0.4787	0.4876	0.4959	0.5001

To illustrate parameter estimation, the inclusion of scale and location results in the density function

$$f(x_1, x_2) = 2t_{\nu_1}\{z_1/\sigma_1\}t_{\nu_2}\{z_2/\sigma_2\}\Phi(\lambda z_1 z_2)/\sigma_1\sigma_2; z_i = (x_i - \mu_i); i = 1, 2. \tag{82}$$

Note that this parameterization means that (i) estimators of scale and degrees of freedom are given by the analogous results for a univariate Student's t distribution, (ii) the estimator of shape depends only on the skewing function Φ and its argument and that (iii) only the estimators of location are complicated by the skewing function. Consequently, in the Fisher information [FI] matrix the non-zero off diagonal elements are in the cells corresponding to $\mu_{1,2}$ and λ and σ_i^2 and ν_i; $i = 1, 2$. Specifically, the FI matrix is

$$FI = \begin{bmatrix} l_{\mu_1\mu_1} & l_{\mu_1\mu_2} & 0 & 0 & l_{\mu_1\lambda} & 0 & 0 \\ l_{\mu_2\mu_1} & l_{\mu_2\mu_2} & 0 & 0 & l_{\mu_2\lambda} & 0 & 0 \\ 0 & 0 & l_{\sigma_1^2\sigma_1^2} & 0 & 0 & l_{\sigma_1^2\nu_1} & 0 \\ 0 & 0 & 0 & l_{\sigma_2^2\sigma_2^2} & 0 & 0 & l_{\sigma_2^2\nu_2} \\ l_{\lambda\mu_1} & l_{\lambda\mu_2} & 0 & 0 & l_{\lambda\lambda} & 0 & 0 \\ 0 & 0 & l_{\nu_1\sigma_1^2} & 0 & 0 & l_{\nu_1\nu_1} & 0 \\ 0 & 0 & 0 & l_{\nu_2\sigma_2^2} & 0 & 0 & l_{\nu_2\nu_2} \end{bmatrix},$$

where $l_{\theta\theta}$ denotes the expected value of the second derivative of the log-likelihood function with respect to a parameter θ. The non-zero elements of FI to be computed numerically using the distribution at Equation (82) are

$$l_{\mu_i\mu_i} = -(\nu_i + 1)/\sigma_i^2(\nu_i + 3) + \sigma_j^2\lambda^2 E\left\{z_j^2\xi_2(\lambda z_i z_j)\right\}; (i, j) = (1, 2), (2, 1).$$

$$l_{\mu_i\mu_j} = \lambda E\{\xi_1(\lambda z_i z_j)\} + \sigma_i\sigma_j\lambda^2 E\{z_i z_j \xi_2(\lambda z_i z_j)\},$$

$$l_{\mu_i\lambda} = -E\{z_j\xi_1(\lambda z_i z_j)\} + \lambda E\left\{z_i z_j^2\xi_2(\lambda z_i z_j)\right\},$$

where $\xi_1(x)$ is as defined in Proposition 1 and $\xi_2(x) = -\xi_1(x)\{x + \xi_1(x)\}$. The remaining elements corresponding to scale and degrees of freedom are the same as those for the FI matrix for Student's t, namely

$$l_{\sigma_i^2\sigma_i^2} = -\nu_i/2\sigma_i^4(\nu_i + 3); l_{\sigma_i^2\nu_i} = 1/\sigma_i^2(\nu_i + 1)(\nu_i + 3),$$

and

$$l_{\nu_i\nu_i} = \left[\Psi'\{(\nu_i+1)/2\} - \Psi'\{\nu_i/2\}\right]/4 - (\nu_i+5)/2\nu_i(\nu_i+1)(\nu_i+3),$$

where $\Psi'(\nu) = d^2 log\Gamma(\nu)/d\nu^2$. These are standard results, but may be found in [25]. The distributions described in this paper all possess the property that the shape parameter may take a value on the boundary of the parameter space. Nonetheless, the FI matrix is inverted and used in the usual way to provide an estimate of the variability of the estimated model parameters. ML estimators of the parameters may be computed using a Newton–Raphson type scheme.

The data set used consists of the weekly returns on 30 stocks that are constituents of the United States S&P500 index. As the example in this subsection and in the following are solely for purposes of illustration, the stocks are numbered. The example in this subsection presents results for 14 pairs of securities, namely stock 1 successively paired with stocks 2 through 15. The computations shown in Table 12 through Table 13 and 14 are based on 100 observations.A standard set of descriptive statistics for the 15 stocks is shown in Table 12.

Table 12. Descriptive Statistics, N = 100.

Stock	Mean	St.dev	Min	Max	Skewness	Kurtosis
1	−0.0025	0.0417	−0.1633	0.1125	−0.4161	4.6963
2	−0.0108	0.0968	−0.4700	0.2095	−1.1852	7.9222
3	−0.0010	0.0608	−0.2572	0.2010	−0.5750	8.0273
4	0.0073	0.0428	−0.1168	0.1468	−0.3540	4.3228
5	−0.0008	0.0349	−0.1456	0.0812	−0.9588	5.6347
6	−0.0020	0.0551	−0.2250	0.1707	−0.4596	5.9019
7	−0.0073	0.1030	−0.6183	0.4433	−1.8121	18.8279
8	−0.0003	0.0570	−0.2353	0.1347	−1.1859	7.2066
9	−0.0033	0.0640	−0.4110	0.2298	−1.8642	20.9143
10	0.0004	0.0332	−0.1520	0.1269	−1.2688	9.8876
11	−0.0059	0.0609	−0.1861	0.1891	0.2182	4.6922
12	−0.0066	0.0519	−0.2371	0.1556	−1.1826	9.3188
13	−0.0067	0.0781	−0.2880	0.1402	−0.6467	4.0701
14	−0.0031	0.0659	−0.3898	0.2074	−2.0377	15.2042
15	−0.0018	0.0428	−0.1304	0.127	−0.0926	4.6006

Table 13 shows estimated parameters for the 14 specified pairs of stocks, computed using the method of maximum likelihood [ML]. Panel 1 shows the estimated parameters. Note that the estimates of the degrees of freedom are shown truncated. Panel 2 shows estimates of parameter precision computed by inverting the FI matrix. The estimated values of the shape parameter λ are all positive, consistent with the stylized fact that stock returns are generally positively correlated. There are four other points to note. First, the estimate of the location parameter for stock 1 depends on the choice of the second stock; that is, it is affected by the presence of the skewing factor and the non-zero values of λ. Secondly, the magnitude of each estimated λ in panel 1 is less than the corresponding estimate of parameter precision in panel 2. This suggests that a test of the null hypothesis $H_0 : \lambda = 0$ would not be rejected against either a one- or two sided alternative. This suggestion, however, is not supported by corresponding likelihood ratio tests reported in Table 15, all of which would lead to rejection of H_0. Thirdly, the estimated values of λ are all less in magnitude than that required for the distribution to be bi-modal. Finally, it is of interest to inquire if the small estimated values λ have much effect on critical values [CVs]. For each of the fourteen pairs of stocks columns 2, 3 and 4 of Table 14 show computed critical values at probabilities of 0.002, 0.01 and 0.05. Columns 5 through 7 show a measure of the effect of the non-zero value of λ. This is computed as follows. When the degrees of freedom are equal, the CV is computed for Student's t distribution with the same degrees of freedom. The column entries show the percentage difference. For example, for stock pair 1–3 the Student's t CV corresponding to a probability of 0.001 is −7.1732. The CV under the model is −7.5351, a difference of abut 5%. When the degrees of freedom for a stock pair are different, the average is taken to compute the Student quantiles. For some applications the differences in the CVs might be regarded as negligible. For other applications they would not.

Table 13. Parameter Estimates under the Bivarate Student Copula.

Panel 1: Parameters							
	μ_1	μ_2	σ_1^2	σ_2^2	λ	ν_x	ν_y
1-2	−0.0039	−0.0097	0.0011	0.0043	0.0285	4	6
1-3	−0.0018	0.0032	0.0011	0.0009	0.0074	4	4
1-4	−0.0016	0.0111	0.0011	0.0010	0.0161	4	4
1-5	−0.0017	0.0015	0.0011	0.0007	0.0211	4	4
1-6	−0.0006	0.0013	0.0011	0.0014	0.0313	4	6
1-7	−0.0022	−0.0005	0.0011	0.0023	0.0067	4	4
1-8	−0.0011	0.0031	0.0011	0.0017	0.0608	4	4
1-9	−0.0035	−0.0006	0.0011	0.0008	0.0077	4	5
1-10	−0.0010	0.0044	0.0011	0.0004	0.0134	4	4
1-11	−0.0020	−0.0086	0.0011	0.0021	0.0325	4	5
1-12	−0.0057	−0.0053	0.0012	0.0010	0.0205	4	5
1-13	−0.0013	−0.0009	0.0011	0.0036	0.0081	4	4
1-14	−0.0034	0.0027	0.0011	0.0014	0.0189	4	4
1-15	−0.0006	−0.0008	0.0011	0.0011	0.0220	4	5
Panel 2: Estimated standard errors							
	μ_1	μ_2	σ_1^2	σ_2^2	λ	ν_x	ν_y
1-2	0.0040	0.0074	0.0003	0.0010	0.0747	1.66	3.45
1-3	0.0040	0.0036	0.0003	0.0002	0.0664	1.66	1.66
1-4	0.0040	0.0037	0.0003	0.0002	0.0664	1.66	1.66
1-5	0.0040	0.0032	0.0003	0.0002	0.0664	1.66	1.66
1-6	0.0040	0.0043	0.0003	0.0003	0.0747	1.66	3.45
1-7	0.0040	0.0057	0.0003	0.0006	0.0664	1.66	1.66
1-8	0.0040	0.0048	0.0003	0.0004	0.0664	1.66	1.66
1-9	0.0040	0.0032	0.0003	0.0002	0.0713	1.66	2.47
1-10	0.0040	0.0024	0.0003	0.0001	0.0664	1.66	1.66
1-11	0.0040	0.0052	0.0003	0.0005	0.0713	1.66	2.47
1-12	0.0040	0.0036	0.0003	0.0002	0.0713	1.66	2.47
1-13	0.0040	0.0071	0.0003	0.0009	0.0664	1.66	1.66
1-14	0.0040	0.0045	0.0003	0.0004	0.0664	1.66	1.66
1-15	0.0040	0.0038	0.0003	0.0003	0.0713	1.66	2.47

Table 14. Critical Values under the bivarate Student copula.

Pair	p = 0.001	p = 0.01	p = 0.05	%err (0.001)	%err (0.01)	%err (0.05)
1-2	−6.3921	−3.5784	−2.1139	8.4613	6.3429	4.9061
1-3	−7.5351	−3.8412	−2.1651	5.0453	2.5150	1.5594
1-4	−7.7466	−3.9225	−2.1976	7.9944	4.6864	3.086
1-5	−7.7792	−3.9714	−2.2139	8.4481	5.9893	3.8494
1-6	−6.4085	−3.5948	−2.1139	8.7405	6.8319	4.9061
1-7	−7.5025	−3.8249	−2.1651	4.5916	2.0807	1.5594
1-8	−7.8117	−4.1341	−2.3115	8.9018	10.3322	8.4292
1-9	−6.7250	−3.6073	−2.0977	4.7312	2.2758	1.5543
1-10	−7.6978	−3.9063	−2.1814	7.3139	4.2521	2.3227
1-11	−6.9876	−3.7714	−2.1633	8.8199	6.9281	4.7320
1-12	−6.9384	−3.7058	−2.1305	8.0533	5.0672	3.1431
1-13	−7.5676	−3.8412	−2.1651	5.499	2.5150	1.5594
1-14	−7.7629	−3.9551	−2.1976	8.2213	5.5550	3.0860
1-15	−6.9384	−3.7222	−2.1469	8.0533	5.5324	3.9376

Table 15. Likelihood ratio tests under the bivarate Student copula.

Pair	$logl_N$	$logl_{STC}$	$lrtest$	p-Value
1-2	268.3707	413.2626	289.7838	0.0000
1-3	314.8326	467.2746	304.8840	0.0000
1-4	349.8942	484.8219	269.8555	0.0000
1-5	370.2326	509.2042	277.9432	0.0000
1-6	324.6268	469.9988	290.7438	0.0000
1-7	262.1842	425.7775	327.1865	0.0000
1-8	321.3790	474.2957	305.8335	0.000
1-9	309.7183	481.3121	343.1876	0.0000
1-10	375.4270	524.3659	297.8779	0.0000
1-11	314.6729	455.9635	282.5812	0.0000
1-12	330.7399	483.9776	306.4756	0.0000
1-13	289.8306	416.5648	253.4685	0.0000
1-14	306.8000	464.5356	315.4711	0.0000
1-15	350.0291	489.5017	278.9452	0.0000

8.3. Multivariate Skew-Normal Copula

The final illustration uses the multivariate skew-normal copula. The parameterization of the distribution at Equation (53) facilitates estimation. The ML estimators of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ depend only on the sets of observations $\{x\}$ and $\{y\}$, respectively. The ML estimator of $\boldsymbol{\lambda}$ depends on $\xi_1(.)$. As above, ML parameter estimation requires only a simple Newton–

Raphson scheme. The FI matrix has a block structure. The $3n \times 3n$ matrix corresponding to $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$ is

$$FI = \begin{bmatrix} l_{\mu\mu^T} & l_{\mu\nu^T} & l_{\mu\lambda^T} \\ l_{\nu\mu^T} & l_{\nu\nu^T} & l_{\nu\lambda^T} \\ l_{\lambda\mu^T} & l_{\lambda\nu^T} & l_{\lambda\lambda^T} \end{bmatrix}. \tag{83}$$

With the following definitions

$$\boldsymbol{z}_1 = \boldsymbol{x} - \boldsymbol{\mu}, \boldsymbol{z}_2 = \boldsymbol{y} - \boldsymbol{\nu}, \boldsymbol{\Lambda} = diag(\lambda_1, \dots, \lambda_n), Q = \boldsymbol{z}_1 \boldsymbol{\Lambda} \boldsymbol{z}_1,$$

the sub-matrices in the FI matrix are

$$l_{\mu\mu^T} = -\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda} E\left\{\boldsymbol{z}_2 \boldsymbol{z}_2^T \xi_2(Q)\right\}, l_{\nu\nu^T} = -\boldsymbol{\Omega}^{-1} + \boldsymbol{\Lambda} E\left\{\boldsymbol{z}_1 \boldsymbol{z}_1^T \xi_2(Q)\right\},$$

$$l_{\mu\nu^T} = \boldsymbol{\Lambda} E\{\xi_1(Q)\} + \boldsymbol{\Lambda} E\left\{\boldsymbol{z}_2 \boldsymbol{z}_1^T \xi_2(Q)\right\}, l_{\lambda\lambda^T} = E\left[\left\{\left(\boldsymbol{z}_1 \boldsymbol{z}_1^T\right) \odot \left(\boldsymbol{z}_2 \boldsymbol{z}_2^T\right)\right\} \xi_2(Q)\right],$$

$$l_{\lambda\mu^T} = E\left[\left\{\left(\boldsymbol{\lambda} \boldsymbol{z}_1^T\right) \odot \left(\boldsymbol{z}_2 \boldsymbol{z}_1^T\right)\right\} \xi_2(Q)\right], l_{\lambda\mu^T} = E\left[\left\{\left(\boldsymbol{\lambda} \boldsymbol{z}_2^T\right) \odot \left(\boldsymbol{z}_1 \boldsymbol{z}_2^T\right)\right\} \xi_2(Q)\right],$$

with all expectations being computed numerically.

For this illustration, a set of 30 US $S\&P500$ stocks are divided into 2 groups each of 15 stocks. Returns are weekly as before and the data set has 500 observations. The standard set of descriptive statistics is shown in Table 16.

The ML estimators of location and shape are computed using a Newton–Raphson scheme, with the resulting estimates shown in Table 17. As in the previous section, the FI matrix is inverted to provide estimates of precision. Unlike the bivariate example, several of the shape parameters exhibit estimates that are more than twice the precision in absolute value, the estimate for the stock pair 1 and 16 being an example.

Table 16. Multivariate Skew-normal Copula—Descriptive Statistics, N = 500, p = 15.

Stock	Avg (1)	Vol (2)	Min (10)	Max (11)	Skew.s (7)	Kurt.s (8)
1	0.0013	0.0281	−0.1633	0.1125	−0.6504	7.0302
2	−0.0015	0.0645	−0.4700	0.2260	−0.9124	10.2368
3	0.0016	0.0377	−0.2572	0.2010	−0.8089	12.5767
4	0.0047	0.0346	−0.1179	0.1468	−0.3044	4.2946
5	0.0001	0.0252	−0.1456	0.0874	−0.7086	6.5132
6	0.0016	0.0469	−0.3430	0.1707	−0.8131	10.2547
7	0.0019	0.0537	−0.6183	0.4433	−2.9290	52.964
8	0.0009	0.0409	−0.2353	0.1347	−0.6603	7.1881
9	0.0001	0.0383	−0.4110	0.2298	−1.9705	34.2342
10	0.0008	0.0223	−0.152	0.1269	−1.0284	11.6669
11	0.0004	0.0428	−0.1861	0.1891	−0.2347	6.0851
12	−0.0011	0.0319	−0.2371	0.1556	−1.4995	15.5992
13	−0.0018	0.0436	−0.2880	0.1402	−1.0084	9.5305
14	0.0003	0.0492	−0.3898	0.2074	−1.2720	12.9359
15	−0.0003	0.0297	−0.1304	0.1270	−0.3297	6.0096
16	0.0016	0.0367	−0.2390	0.1358	−0.6832	7.7613

Table 16. *Cont.*

Stock	Avg (1)	Vol (2)	Min (10)	Max (11)	Skew.s (7)	Kurt.s (8)
17	0.0018	0.0219	−0.1421	0.1063	−0.7507	10.0922
18	0.0019	0.0264	−0.3156	0.1715	−2.9183	47.872
19	0.0018	0.0270	−0.1209	0.1007	−0.3455	5.7724
20	0.0042	0.0301	−0.1118	0.1342	−0.3669	5.6791
21	0.0032	0.0310	−0.2107	0.1495	−0.7606	8.9065
22	0.0017	0.0258	−0.1235	0.0909	−0.3944	5.3920
23	0.0016	0.0216	−0.1517	0.0888	−0.9443	9.9823
24	0.0044	0.0317	−0.3133	0.1586	−1.8778	25.3881
25	0.0014	0.0303	−0.3209	0.1208	−2.6446	30.2787
26	0.0012	0.024	−0.0841	0.0987	0.1249	4.4550
27	0.0018	0.0238	−0.1138	0.1098	−0.2549	6.8268
28	0.0025	0.0318	−0.1724	0.1529	−0.4396	9.2548
29	0.0019	0.0356	−0.2223	0.1114	−0.8660	7.3263
30	0.0004	0.0349	−0.3031	0.1918	−2.1926	22.9448

Table 17. Multivariate Skew-normal Copula—Location and Shape Parameter Estimates.

Pair	Locn-X	St.err-X	Locn-Y	St.err-Y	Shape	St.err
1-16	0.0013	0.0013	0.0016	0.0017	0.3870	0.1227
2-17	−0.0015	0.0029	0.0018	0.0010	−0.2022	0.1227
3-18	0.0016	0.0017	0.0020	0.0012	0.2279	0.1295
4-19	0.0047	0.0016	0.0018	0.0012	−0.0076	0.0815
5-20	0.0001	0.0011	0.0042	0.0014	0.1610	0.0811
6-21	0.0016	0.0021	0.0032	0.0014	0.4757	0.1431
7-22	0.0019	0.0024	0.0017	0.0012	0.1762	0.1378
8-23	0.0009	0.0018	0.0016	0.0010	−0.0085	0.0982
9-24	0.0001	0.0018	0.0044	0.0015	0.6524	0.1785
10-25	0.0008	0.0010	0.0015	0.0014	0.3382	0.1333
11-26	0.0004	0.0019	0.0012	0.0011	0.1169	0.0993
12-27	−0.0011	0.0015	0.0018	0.0011	0.1943	0.1192
13-28	−0.0018	0.0020	0.0025	0.0014	0.0699	0.1125
14-29	0.0003	0.0022	0.0019	0.0016	0.5011	0.1303
15-30	−0.0003	0.0014	0.0005	0.0016	0.7160	0.1488

A standard likelihood ratio test has a value of 259.3 and thus leads to rejection of the null hypothesis that all shape parameters equal zero.

9. Concluding Remarks

This paper paper reports the results of an investigation into the properties of a copula-like version of the skew-normal and skew-Student distributions. The distributions studied in the paper allow the marginal distributions to be either normal or Student's t with differing degrees of freedom. There are several conditional distributions that resemble the skew-normal or are closely related to it. Many of the required computations require numerical integration. The properties of some of the distributions studied depend on certain of the special functions, in particular the G and H functions. There are no explicit expressions available for the moment generating or characteristic functions, although moments and cross moments may be computed when they exist. The examples contained in the paper suggest that parameter estimation is straight forward.

The results show that study of marginal distributions may conceal the nature of a dependence structure and that furthermore there may be different such structures. For future research, there are a number of technical issues concerned with integration. There is also scope for more general results based on unified or generalized skew-elliptical distributions.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are grateful to the reviewers of the paper for comments which have led to an improved presentation of the original version.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Appendix A.1. Proof of Proposition 3

The marginal distribution of X is symmetric and has the density function at Equation (22)

$$f(x) = \phi(x)\Phi(\tilde{\tau}_x)/\Omega(\tau, \lambda); \tilde{\tau}_x = \tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 x^2}.$$

1. $E(X) = 0$ by symmetry.
2. $var(X)$ is computed using integration by parts, as follows, omitting the denominator $\Omega(.)$
$$var(X) \propto [-x\phi(x)\Phi(\tilde{\tau}_x)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \phi(x)\frac{d}{dx}\{x\Phi(\tilde{\tau}_x)\}dx.$$
The first term is zero. The second term is
$$\int_{-\infty}^{\infty} \phi(x)\Phi(\tilde{\tau}_x)dx - \left\{\lambda^2\tau\sqrt{1+\lambda^2}\right\}\int_{-\infty}^{\infty} \frac{x^2}{(1+\lambda^2 x^2)^{3/2}}\phi(\tilde{\tau}_x)\phi(x)dx.$$
On division by $\Omega(\tau, \lambda)$ the first term is equal to unity.
3. $cov(X, Y)$ is also computed using integration by parts. Omitting $\Omega(.)$ and integrating with respect to y gives
$$x\phi(x)\left\{\left[-\phi(y)\Phi\left(\tau\sqrt{1+\lambda^2}+\lambda xy\right)\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \lambda x\phi(y)\phi\left(\tau\sqrt{1+\lambda^2}+\lambda xy\right)dy\right\}.$$
The term in []s equals zero. On integration with respect to y the second term is
$$\lambda x\phi(\tilde{\tau}_x)/\sqrt{1+\lambda^2 x^2}.$$
Integration with respect to x gives the result.
4. The conditional distribution at Point 4 follows from dividing Equation (20) by (22).

Appendix A.2. Proof of Proposition 4

1. $\Omega(\tau, \lambda)$ is given by Equation (21)

$$\Omega(\tau, \lambda) = \int_{-\infty}^{\infty} \phi(y)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 y^2}\right)dy,$$

Integration by parts and noting tht

$$\left[\Phi(y)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 y^2}\right)\right]_{-\infty}^{\infty} = 1/2,$$

gives the result.

2. Follows directly from Equation (22).
3. Follows directly by dividing the joint density function at Equation (20) by Equation (22) and using standard methods.
4. The proof employs the asymptotic formula for the standard normal integral in [16] (page 932, Equation 26.2.12); that is

$$\Phi(x) \simeq \phi(x)/|x|; x << 0,$$

and assumes that higher order terms may be neglected as may integrals of the form

$$\int_{-\infty}^{x} \frac{\phi(s)}{s^p} ds \simeq \phi(x)/(-1)^p |x|^{p+1}; x << 0,$$

In Equation (26) the denominator is.

$$T_D = \int_{-\infty}^{y} \phi(s)\Phi\left(\tau\sqrt{1+\lambda^2}/\sqrt{1+\lambda^2 s^2}\right)ds.$$

Integrating by parts and using the definition at Equation (23) gives

$$T_D = \Phi(y)\Phi(\tilde{\tau}_y) - \tau\lambda\sqrt{1+\lambda^2}\int_{-\infty}^{y} \frac{\phi(s)\Phi(\tilde{\tau}_s)}{(1+\lambda^2 s^2)^{3/2}} ds \simeq \Phi(y)\Phi(\tilde{\tau}_y).$$

The numerator in (26) is

$$T_N = \phi(x_1)\int_{-\infty}^{y} \phi(s)\Phi\left(\tau\sqrt{1+\lambda^2} + \lambda x_1 s\right)ds$$

Integrating by parts again with $\tau' = \tau\sqrt{1+\lambda^2}$ gives

$$T_N = \phi(x_1)\phi(y)\Phi\left(\tau' + \lambda x_1 y\right) - \lambda x_1 \phi(x_1)\int_{-\infty}^{y} \Phi(s)\phi\left(\tau' + \lambda x_1 s\right)ds.$$

In the second term in T_N, $s \leq y << 0$, in which case the integrand

$$\Phi(s)\phi\left(\tau' + \lambda x_1 s\right) \simeq \frac{\phi(s)\phi\left(\tau' + \lambda x_1 s\right)}{|s|},$$

may be neglected. Noting that

$$\tau' = \tilde{\tau}_y\sqrt{1+\lambda^2 y^2},$$

completes the proof.

References

1. Loperfido, N. Skewness-based projection pursuit: A computational approach. *Comput. Stat. Data Anal.* **2018**, *120*, 42–57. [CrossRef]
2. Azzalini, A. A Class of Distributions which Includes The Normal Ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
3. Azzalini, A. Further Results on a Class of Distributions which Includes The Normal Ones. *Statistica* **1986**, *46*, 199–208.
4. Azzalini, A.; Dalla-Valle, A. The Multivariate Skew Normal Distribution. *Biometrika* **1996**, *83*, 715–726. [CrossRef]
5. Azzalini, A.; Capitanio, A. Distributions Generated by Perturbation of Symmetry With Emphasis on a Multivariate Skew t Distribution. *J. R. Stat. Soc. Ser. B* **2003**, *65*, 367–389. [CrossRef]
6. Sahu, S.K.; Dey, D.K.; Branco, M.D. A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models. *Can. J. Stat.* **2003**, *31*, 129–150. [CrossRef]
7. Genton, M.G.; Loperfido, N.M. Generalized skew-elliptical distributions and their quadratic forms. *Ann. Inst. Stat. Math.* **2005**, *57*, 389–401. [CrossRef]
8. Arellano-Valle, R.B.; Genton, M.G. Multivariate unified skew-elliptical distributions. *Chil. J. Stat.* **2010**, *926*, 17–33.
9. Bodnar, T.; Mazur, S.; Parolya, N. Central limit theorems for functionals of large sample covariance matrix and mean vector in matrix-variate location mixture of normal distributions. *Scand. J. Stat.* **2019**, *46*, 636–660. [CrossRef]
10. Sklar, A. Fonctions de Répartition à n Dimensions et Leurs Marges. *Publ. L'Institut Stat. L'Université Paris* **1959**, *8*, 229–231.
11. Meijer, C.S. Über Whittakersche bzw. Besselsche Funktionen und deren Produkte. *Nieuw Arch. Voor Wiskd.* **1936**, *18*, 10–39.
12. Fox, C. The G and H Functions as Symmetrical Fourier Kernels. *Trans. Am. Math. Soc.* **1961**, *98*, 395–429. [CrossRef]
13. Adcock, C.J.; Landsman, Z.; Sushi, T. Steins Lemma for generalized skew-elliptical random vectors. *Commun. Stat. Theory Methods* **2019**. [CrossRef]
14. Stein, C.M. Estimation of The Mean of a Multivariate Normal Distirbution. *Ann. Stat.* **1981**, *9*, 1135–1151. [CrossRef]
15. Gaunt, R.E. *Products of Normal, Beta and Gamma Random Variables: Stein Operators and Distributional Theory*; Working Paper; The University of Manchester: Manchester, Lancashire, UK, 2016.
16. Abramowitz, M.; Stegun, I. *Handbook of Mathematical Functions*; Dover: New York, NY, USA, 1965.
17. Owen, D.B. Tables for Computing Bivariate Normal Probabilities. *Ann. Math. Stat.* **1956**, *27,1075–1090*. [CrossRef]
18. Azzalini, A.; Capitanio, A. Statistical Applications of The Multivariate Skew Normal Distribution. *J. R. Stat. Soc. Ser. B* **1999**, *61*, 579–602. [CrossRef]
19. Berg, C.; Vignat, C. On the density of the sum of two independent Student t-random vectors. *Stat. Probab. Lett.* **2010**, *80*, 1043–1055. [CrossRef]
20. Walker, G.A.; Saw, J.G. The Distribution of Linear Combinations of t-Variables. *J. Am. Stat. Assoc.* **1978**, *73*, 876–878. [CrossRef]
21. Loperfido, N. *Private Correspondence*; Two Items; University of Urbino: Urbino, Italy, 2020.
22. Adcock, C.J. *Private Correspondence*; Two Items, March; Sheffield University Management School: Sheffield, UK, 2020.
23. Loperfido, N. Generalized Skew-Normal Distributions. In *Skew-Elliptical Distributions and Their Application: A Journey Beyond Normality*; Genton, M., Ed.; CRC/Chapman and Hall: Boca Raton, FL, USA, 2004; pp. 65–80.
24. Adcock, C.J. Extensions of Stein's Lemma for the Skew-Normal Distribution. *Commun. Stat. Theory Methods* **2007**, *36*, 1661–1671. [CrossRef]
25. Lange, K.L.; Little, R.J.A.; Taylor, J.M.G. Robust Statistical Modeling Using the t Distribution. *J. Am. Stat. Assoc.* **1989**, *85*, 881–896. [CrossRef]