On Bivariate Transformation of Scale Distributions

M.C. Jones

The Open University, U.K.

ABSTRACT

Elsewhere, I have promoted the notion of (univariate continuous) "transformation of scale" (ToS) distributions which have densities of the form $2g(W^{-1}(x))$ where g is a symmetric distribution and W is a transformation function with a special property. Here, I develop particular bivariate (readily multivariate) ToS distributions. Univariate ToS distributions have a transformation of random variable relationship with Azzalini-type skew-symmetric distributions; the bivariate ToS distribution here arises from marginal variable transformation of a non-standard form of bivariate skew-symmetric distribution. Some examples are given, as are a few basic properties — unimodality, a covariance property, random variate generation — and connections with a particular bivariate inverse Gaussian distribution are pointed out.

KEY WORDS: *bivariate inverse Gaussian distribution; R-symmetry; sinh-arcsinh transformation; skew-symmetric distribution; two-piece distribution.*

---

1. INTRODUCTION

In Jones (2010, 2013), I developed a general technique for generating families of univariate continuous distributions from simple symmetric initial distributions (such as the Gaussian), which I call "transformation of scale" (ToS) distributions; see Section 2.1 below for brief details. A prime reason for developing such distributions is the introduction of skewness into distributions. After talks on the subject, the inevitable question is asked: "What about the multivariate case?"

At one level, there is an easy answer: any univariate distribution(s) can be generalised to the multivariate case by use of a generally applicable method, the most obvious and popular of which is the use of copulas (e.g. Nelsen, 2010). Elliptical copulas, that is copulas associated with the elliptical extension of the "simple symmetric initial distributions" above (such as the Gaussian copula), might be a natural choice in this case (Song, 2000, Demarta & McNeil, 2005). I have to admit that, since ToS distributions do not, in general, have closed form expressions for their distribution or quantile functions, transforming the marginals of the chosen copula to the desired ToS distributions is not so immediate. (But this appears to be no impediment to Smith, Gan & Kohn, 2012, in a closely related context.)

However, in offering this answer, I have already interpreted the question in one particular way (which is the one I will stick with throughout this chapter after this section). That is to say that a multivariate ToS distribution is one with univariate ToS marginals. This is an appropriate interpretation of the question when, as is often the case, the desire for introducing skewness into multivariate distributions is driven by observations of skewness in its constituent random variables, i.e. marginals.

If one is instead in another multivariate world where linear combinations of the original random variables are the main currency, then many more options — and questions about what you mean by a "multivariate ToS distribution" — open up (e.g. Ferreira & Steel, 2007).

Yet another multivariate world, of graphical modelling and the like, pieces together multivariate distributions via (univariate, potentially ToS) marginals and conditionals. And the importance of multivariate distributions with marginals all of a specific univariate type is also somewhat undermined in graphical models and elsewhere by the typical situation of a large set of random variables having a variety of different distributions and, indeed, being of a variety of different measurement types.

But usually, none of the above is what the questioner wishes to hear about. S/he thought s/he was asking: "What about the 'natural' extension to the multivariate case?" There often is such a thing. For example, skew-symmetric distributions of the classical "Azzalini-type" (Azzalini, 1985) arise via various underlying mechanisms, one of which is as the marginals of truncated distributions of higher dimension. This mechanism generalises immediately and 'naturally' to the multivariate case (Azzalini & Capitanio, 1999, Branco & Dey, 2001), in which it has a number of nice properties and is very popular. These 'natural' generalisations then actually explode in number (Arellano-Valle, Branco & Genton, 2006). They do not, however, except in important special cases, retain the form of the univariate skew-symmetric distribution as their marginal distributions (this is not a criticism, just a relevant fact).

ToS distributions, on the other hand, can be viewed in terms of transformations of certain skew-symmetric distributions (see below) and so have an alternative 'natural' multivariate generalisation: apply the transformations marginally to an appropriate multivariate initial distribution, thereby guaranteeing ToS marginals.

Note, however, that 'natural' does not necessarily equate to 'best' or even to 'useful'. It remains to understand the other consequences of the 'natural' extension and this is what I shall be doing, in a limited way, for the extension by marginal transformation of ToS distributions in this chapter.

Although the general discussion above is couched in multivariate terms, only the bivariate case — which incorporates many of the main difficulties — is considered in detail, for clarity and convenience, in the main body of the paper. Multivariate extension will just be mentioned briefly a further few times.

Univariate ToS distributions are explained briefly in Section 2.1. The bivariate ToS distributions of interest in this chapter result from applying the same transformations marginally to a bivariate skew-symmetric distribution of a certain, non-standard, form described in Section 2.2. The main general properties — and examples — of bivariate ToS distributions are given in Section 3; these include unimodality, a covariance property and random variate generation. A link is made, in Section 4, of a special case of the bivariate ToS construction to an existing bivariate inverse Gaussian distribution. Some brief closing remarks complete the chapter in Section 5.

2. BACKGROUND

2.1. Univariate Transformation of Scale Distributions

Univariate ToS distributions have densities of the form

$$f(x) = 2g\{W^{-1}(x)\}, \quad x \in \text{ support } \mathcal{S}, \tag{1}$$

where g is the density of a continuous distribution with support $\mathbb{R}$ which is symmetric about zero, and the increasing transformation function $W : \mathbb{R} \to \mathcal{S}$ satisfies

$$W(y) - W(-y) = y, \quad \text{for all } y \in \mathbb{R}. \tag{2}$$

Prime examples of this construction include:

- when $\mathcal{S} = \mathbb{R}^+$, the R-symmetric, or Cauchy-Schlömilch, distributions (which include a close relation of the inverse Gaussian), with densities

$$f(x) = 2g\left(x - \frac{\theta}{x}\right), \quad \theta > 0$$

(Mudholkar & Wang, 2007, Baker, 2008);

- when $\mathcal{S} = \mathbb{R}$, the long-standing two-piece distributions, with densities

$$f(x) = 2g\left\{\frac{2x}{1-\alpha}I(x<0) + \frac{2x}{1+\alpha}I(x \ge 0)\right\}, \quad -1 < \alpha < 1$$

(Fechner, 1897, Fernández & Steel, 1998, Mudholkar & Hutson, 2000);

- when $\mathcal{S} = \mathbb{R}$, 'sinh-arcsinh-type' distributions, with (rescaled) densities

$$f(x) = \frac{1}{\cosh\epsilon} g\left[\sinh\{\sinh^{-1}(x) - \epsilon\}\right], \quad \epsilon \in \mathbb{R}$$

(Jones, 2013).

Inverse Batschelet distributions on the circle (Jones & Pewsey, 2012) are also very closely related.

Now, transformation of scale, transformation of random variable and Azzalini-type skew-symmetric distributions come together in the following way (Jones, 2013). Let $\sim$ denote 'has distribution with density' and define $w(y) = W'(y) = dW(y)/dy$ where W is increasing and satisfies (2). Then,

$$Y \sim f_s \quad \Rightarrow \quad X = W(Y) \sim f \tag{3}$$

where

$$f_s(y) = 2g(y)w(y), \quad x \in \mathbb{R}, \tag{4}$$

and, from (2),

$$w(y) + w(-y) = 1, \quad \text{for all } y \in \mathbb{R}. \tag{5}$$

That is, the ToS random variable X arises by the transformation $X = W(Y)$ applied to a random variable with the Azzalini-type skew-symmetric distribution with density (4), in its general form with w satisfying (5) explored by Wang, Boyer & Genton (2004).

2.2. Bivariate Distributions With Marginals of the Form (4)

Similar to Azzalini & Capitanio (1999) and Branco and Dey (2001), in the (most usual) case where w is a distribution function, one of the simplest, canonical, forms of bivariate Azzalini-type skew-symmetric distribution has density

$$\check{f}(x,y) = 2g(x,y)w(a_1x + a_2y), \quad x,y \in \mathbb{R}^2; \tag{6}$$

this arises as the bivariate marginal of the trivariate distribution of $(X,Y) \sim g$ (where g is a bivariate distribution with an appropriate form of symmetry) and, independently, $Z \sim w'$, where Z is truncated to values of $z < a_1x + a_2y$. It is a version of the bivariate extension of (4) investigated by Wang, Boyer & Genton (2004).

The marginals of (6) have the desired skew-symmetric (in fact, skew-normal) form when g and w' are normal densities (Azzalini & Capitanio, 1999). But this occurs because of special properties of the normal distribution, and construction (6) does not have the desired form of skew-symmetric marginals in general. (The latter is essentially because this standard type of multivariate skew-symmetric distribution fits into the "multivariate world where linear combinations of the original random variables are the main currency" mentioned in Section 1.)

A simple version of the more general selection mechanism first put forward by Sahu, Dey & Branco (2003) — which uses as many extra "selection" variables as the dimension of interest — does provide the desired marginals. In the bivariate case, consider

$$\tilde{f}(x,y) = 4g(x,y)w_1(x)w_2(y), \quad x,y \in \mathbb{R}^2, \tag{7}$$

where w_1 and w_2 both satisfy (5) (they could, of course, be the same). When w_1 and w_2 are distribution functions, this arises as the bivariate marginal of the quadrivariate distribution of $(X,Y) \sim g$ and, independently, $Z_1 \sim w_1'$, $Z_2 \sim w_2'$, where Z_1 and Z_2 are truncated to values of $z_1 < x$ and $z_2 < y$, respectively. Note that (7) is not a special case of Wang, Boyer & Genton (2004). For future use, write $g_1(x)$, $g_2(y)$, $g(y|x)$ and $g(x|y)$ for the marginal and conditional densities of g.

I will take g to be sign-symmetric, which in the bivariate case means that

$$g(x,y) = g(x,-y) = g(-x,y) = g(x,y) \quad \text{for all } x,y \in \mathbb{R}^2. \tag{8}$$

This includes the usual spherical symmetry as a special case (e.g. Serfling, 2006); other examples include ℓ_q-spherical models (Osiewalski & Steel, 1993, Gupta & Song, 1997), especially its further special case of ℓ_1-spherical distributions, and an interesting case considered in Section 4. Clearly, the marginal and conditional distributions of sign-symmetric distributions are symmetric about zero and this is what makes the marginalisation work out as desired. For instance:

$$\int \tilde{f}(x,y)dy = 4g_1(x)w_1(x)\int g(y|x)w_2(y)dy = 2g_1(x)w_1(x),$$

using the fact underlying the simplicity of (4) that $\int h(y)w(y)dy = 1/2$ for any h symmetric about zero and w satisfying (5). Rather attractively, the conditional distributions of $\tilde{f}$ are also the conditional distributions of g skewed by the w functions e.g. $\tilde{f}(y|x) = 2g(y|x)w_2(y)$.

It is reasonable to note that distributions (7) might not show a great deal of dependence between X and Y in many cases. Indeed, in the normal-based case, (7) usually reduces to independent skew-symmetric distributions. In general, the local dependence function of $\tilde{f}$ (Holland & Wang, 1987) is the same as that of g.

Of course, construction (7) extends readily to the multivariate case, retaining similar properties for sign-symmetric g:

$$\tilde{f}(x_1, ..., x_d) = 2^d g(x_1, ..., x_d) \prod_{i=1}^{d} w_i(x_i), \quad x_1, ..., x_d \in \mathbb{R}^d.$$

3. BIVARIATE ToS DISTRIBUTIONS

Let $(U, V) \sim \tilde{f}$ and transform them marginally via $X = W_1(U)$, $Y = W_2(V)$, where $w_1(u) = W_1'(u)$ and $w_2(v) = W_2'(v)$, to obtain the bivariate transformation of scale distribution with density

$$\hat{f}(x, y) = 4g(W_1^{-1}(x), W_2^{-1}(y)), \quad x, y \in \mathcal{S}_1 \times \mathcal{S}_2, \tag{9}$$

where g is a sign-symmetric bivariate distribution as at (8). By construction, the marginal distributions of $\hat{f}$ are univariate ToS distributions:

$$\hat{f}_1(x) = 2g_1(W_1^{-1}(x)), \quad x \in \mathcal{S}_1, \quad \hat{f}_2(x) = 2g_2(W_2^{-1}(y)), \quad y \in \mathcal{S}_2;$$

and, collaterally, the conditional distributions of $\hat{f}$ are also univariate ToS distributions with appropriate conditioning, for example,

$$\hat{f}(y|x) = 2g(W_2^{-1}(y)|W_1^{-1}(x)), \quad x, y \in \mathcal{S}_1 \times \mathcal{S}_2.$$

If, as would usually be the case, g is unimodal, with mode at $(0, 0)$, then it is immediate that $\hat{f}$ given by (9) is also unimodal, with mode at $(W_1(0), W_2(0))$. Unimodality is not guaranteed in the same way for alternative constructions like (7).

Note that, as with (7), in practical versions of (9) one would introduce marginal location and scale parameters for each $\mathcal{S} = \mathbb{R}$ or just marginal scale parameters for each $\mathcal{S} = \mathbb{R}^+$, for example,

$$\frac{1}{\sigma_1\sigma_2}\hat{f}\left(\frac{x - \mu_1}{\sigma_1}, \frac{y - \mu_2}{\sigma_2}\right), \quad x, y \in \mathbb{R}^2.$$

It is not, however, appropriate to linearly transform ((7) or) (9) to introduce further correlation since that would destroy the marginal properties that I have been at pains to preserve.

Because of their relationship via marginal transformation, the bivariate ToS distribution with density (9) and the bivariate skew-symmetric distribution with density (7) share the same copula.

There is a multivariate version of density (9) which is, of course,

$$\hat{f}(x_1,...,x_d) = 2^d g(W_1^{-1}(x_1),...,W_d^{-1}(x_d)), \quad x_1,...,x_d \in \prod_{i=1}^{d} \mathcal{S}_i,$$

where g is a multivariate sign-symmetric density and each W_i satisfies (2).

3.1. Examples

A couple of examples of (9) are shown in Figure 1. In Figure 1(a) is displayed the contours of the bivariate R-symmetric Cauchy distribution — or bivariate Cauchy-Schlömilch-Cauchy distribution! — formed by applying the transformation of scale $W^{-1}(x) = x - (\theta/x)$ to both marginals of the bivariate Cauchy distribution, in the case $\theta_1 = \theta_2 = 1$. It has density

$$f_{\mathrm{a}}(x,y) = \frac{1}{\pi\left\{1 + \left(x - \frac{1}{x}\right)^2 + \left(y - \frac{1}{y}\right)^2\right\}^{3/2}}, \quad x,y \in \mathbb{R}^+ \times \mathbb{R}^+. \tag{10}$$

Figure 1(b) is based on the ℓ_1-symmetric t_4 distribution as given in Example 2.5 of Gupta & Song (1997) which has concentric diamond-shaped contours. I have applied ToS in the sinh-arcsinh-type form of Section 2.1 with $\epsilon = -0.5$ to effect a negatively skewed x-marginal, and left the y-scale alone, viz.,

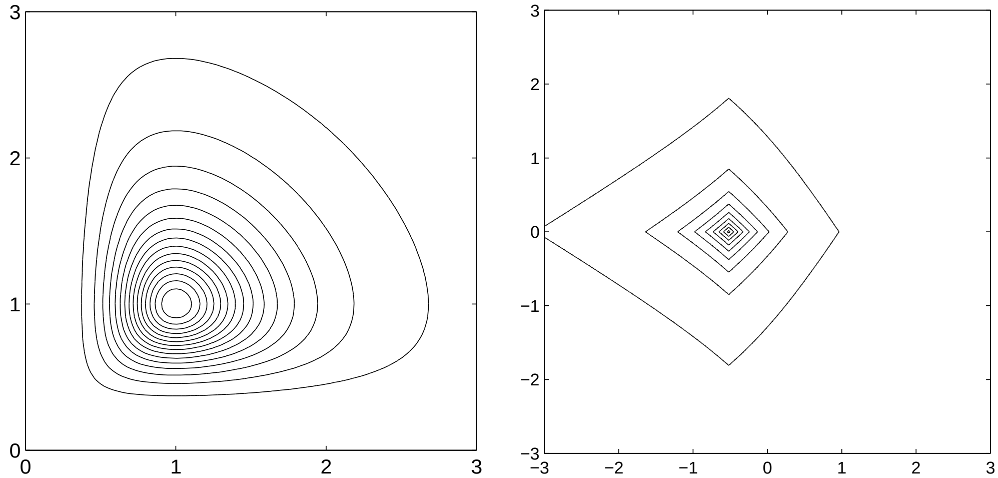
$$f_{\mathrm{b}}(x,y) = \frac{1}{2\cosh(\frac{1}{2})\left[1 + \left|\sinh\left\{\sinh^{-1}(x) - \frac{1}{2}\right\}\right| + |y|\right]^3}, \quad x,y \in \mathbb{R}^2. \tag{11}$$

A nice example of a bivariate two-piece ToS distribution with different two-piece marginals, based on the spherically symmetric t_6 distribution, can be found in Figure 1 of Bauwens & Laurent (2005). One can also use different transformations of scale in different marginals, including producing distributions with marginals on different supports.

3.2. A Covariance Property

Like in the univariate case in Jones (2010, 2013), write

$$s_i(x) = x - W_i^{-1}(x), \quad i = 1, 2.$$



(a) (b)

Figure 1: Contours increasing from 0.02 (a) in steps of 0.02 of the bivariate R-symmetric Cauchy density given at (10); (b) in steps of 0.05 of the bivariate sinh-arcsinh ToS t_4 density given by (11).

Then, formula (3.3) of Jones (2010) gives that

$$E\{s_1'(X)\} = E\{s_2'(Y)\} = -1.$$

In the bivariate case, the same applies conditionally; for example,

$$\begin{aligned} E\{1 - s_2'(Y)|X = x\} &= 2\int \{1 - s_2'(y)\} g(y - s_2(y)|W_1^{-1}(x))dy \\ &= 2\int g(z|W_1^{-1}(x))dz = 2 \end{aligned}$$

so that the bivariate ToS distribution has the constant regressions property

$$E\{s_2'(Y)|X = x\} = E\{s_1'(X)|Y = y\} = -1.$$

It follows immediately that

$$E\{s_1'(X)s_2'(Y)\} = E_X\left[s_1'(X)E\{s_2'(Y)|X\}\right] = -E\{s_1'(X)\} = 1$$

and hence that

$$\text{Cov}\{s_1'(X), s_2'(Y)\} = 0.$$

3.3. Random Variate Generation

As at the start of Section 3, $(X, Y) \sim \hat{f}$ given by (9) if $X = W_1(U)$, $Y = W_2(V)$ and $(U, V) \sim \tilde{f}$ given by (7). It remains to generate (U, V) from $\tilde{f}$. This can be achieved assuming that random variables (R, S) are available from the distribution with density g. Then,

$$(U, V) = \begin{cases} (R, S) & \text{with probability } w_1(R)w_2(S), \\ (-R, S) & \text{with probability } (1 - w_1(R))w_2(S), \\ (R, -S) & \text{with probability } w_1(R)(1 - w_2(S)), \\ (-R, -S) & \text{with probability } (1 - w_1(R))(1 - w_2(S)). \end{cases} \tag{12}$$

For each pair $(X, Y) \sim \hat{f}$, one therefore needs to generate three random variables, in addition to making the appropriate transformations: $(R, S) \sim g$ and a further independent $U(0, 1)$ random variable to make the test in (12). Note that there is no rejection of generated random variables.

4. BIVARIATE ToS DISTRIBUTIONS BASED ON A MIXTURE FORM FOR g

4.1. A Mixture Form for g

Sign-symmetric g also arises if one takes a half-and-half mixture of elliptical distributions in standard form with equal and opposite correlations, that is,

$$g_{\mathrm{m}}(x, y) = \frac{1}{2} \left\{ g_\rho(x, y) + g_{-\rho}(x, y) \right\}, \quad x, y, \in \mathbb{R}^2, \tag{13}$$

where $0 \leq \rho < 1$. While one could take g_ρ to be any standard form elliptical distribution, I will take g to be the standard normal distribution in this section for simplicity. This results, after a little manipulation, in the density

$$g_{\mathrm{n}}(x, y) = \frac{1}{\sqrt{1 - \rho^2}} \cosh\left(\frac{\rho x y}{1 - \rho^2} \right) \phi\left(\frac{x}{\sqrt{1 - \rho^2}} \right) \phi\left(\frac{y}{\sqrt{1 - \rho^2}} \right), \quad x, y \in \mathbb{R}^2 \tag{14}$$

where ϕ is the univariate standard normal density function. (This is a rescaling to marginals with variance 1 of the rather beautiful density formula

$$\sqrt{1 - \rho^2} \cosh(\rho x y) \, \phi(x) \, \phi(y), \quad x, y \in \mathbb{R}^2.)$$

See Kowalski (1973) for some properties of g_n and its natural extensions.

Two examples of density (14) are shown in Figure 2. The density, which can easily be shown to be unimodal, does not strongly exhibit the diagonal cross shape that one might have expected until ρ becomes quite large.

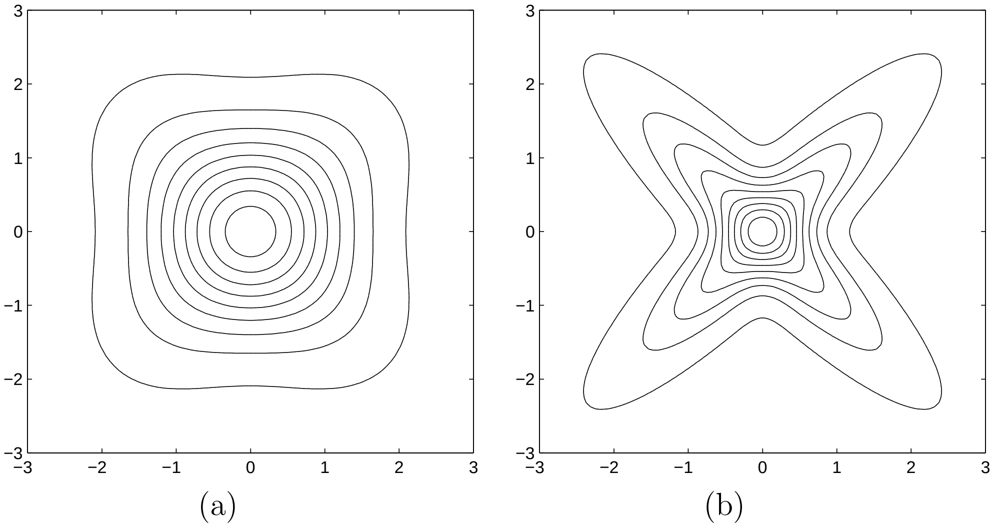


Figure 2: Contours of the bivariate normal mixture density given at (14) (a) increasing from 0.01 in steps of 0.02 when $\rho = 0.5$, (b) increasing from 0.01 in steps of 0.04 when $\rho = 0.9$.

4.2. Bivariate ToS Distributions Based on $g_{\rm n}$

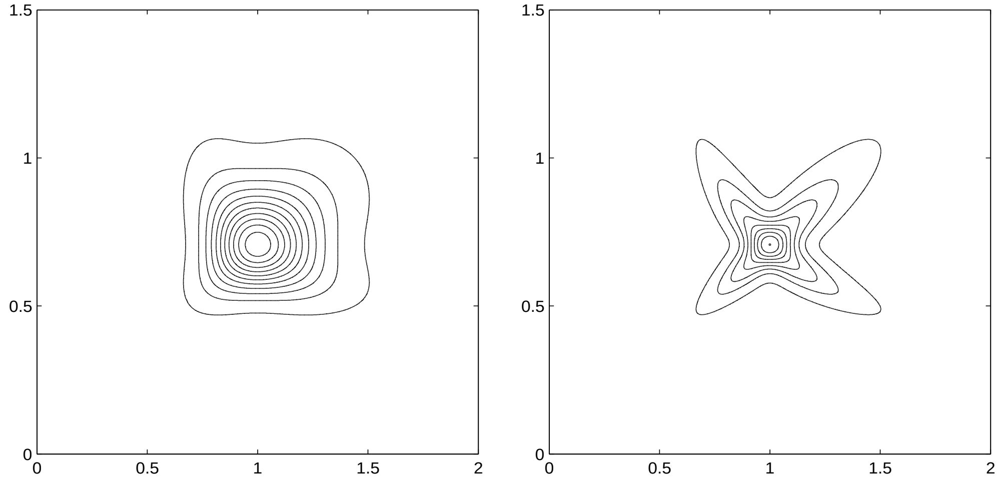
Bivariate ToS distributions now arise from (14) by introducing transformation of scale functions W_1 and W_2 as in (9) for this particular choice of $g_{\rm n}$, as exemplified in Section 3.1 for other choices of g.

I shall, however, confine myself to just two, R-symmetric, illustrations of such distributions here. They are shown in Figure 3. The transformations employed are again of the form $W_i^{-1}(x) = x - (\theta_i/x)$, $\theta_i > 0$, $i = 1, 2$, and scale parameters $\sigma_1, \sigma_2 > 0$ are also introduced. Their densities are therefore

$$\begin{aligned} g_r(x,y) &= \frac{4}{\sqrt{1-\rho^2}\,\sigma_1\sigma_2} \cosh\left\{\frac{\rho}{1-\rho^2}\left(\frac{x}{\sigma_1} - \frac{\sigma_1\theta_1}{x}\right)\left(\frac{y}{\sigma_2} - \frac{\sigma_2\theta_2}{y}\right)\right\} \\ &\times\ \phi\left\{\frac{1}{\sqrt{1-\rho^2}}\left(\frac{x}{\sigma_1} - \frac{\sigma_1\theta_1}{x}\right)\right\} \phi\left\{\frac{1}{\sqrt{1-\rho^2}}\left(\frac{y}{\sigma_2} - \frac{\sigma_2\theta_2}{y}\right)\right\}, \qquad (15) \end{aligned}$$

$x, y \in \mathbb{R}^+ \times \mathbb{R}^+$. Density (15) can be called a bivariate root reciprocal inverse Gaussian (RRIG) distribution, since its marginals are univariate RRIG distributions (Mudholkar & Wang, 2007). In Figure 3, $\theta_1 = \theta_2 = 8$, $\sigma_1 = 1/(2\sqrt{2})$ and $\sigma_2 = 1/4$; in Figure 3(a), $\rho = 0.5$, in Figure 3(b), $\rho = 0.9$. (The reason for these particular choices of θs and σs will become apparent in Section 4.3.)

Clearly, the bivariate RRIG distribution is the result of 'pulling' and thereby 'skewing' the normal mixture density onto its new support while retaining aspects of the same basic density shape (including unimodality). In the case of $\rho = 0.9$ in



(a) (b)

Figure 3: Contours of the bivariate root reciprocal inverse Gaussian density given at (15) when $\theta_1 = \theta_2 = 8$, $\sigma_1 = 1/(2\sqrt{2})$, $\sigma_2 = 1/4$, (a) increasing from 0.25 in steps of 0.75 when $\rho = 0.5$, (b) increasing from 0.5 in steps of 2 when $\rho = 0.9$.

Figure 3(b), the diagonal cross shape of Figure 2(b) persists. It is not clear to me whether or not this might sometimes a realistic shape for data clouds on $\mathbb{R}^+ \times \mathbb{R}^+$.

4.3. The Link to a Bivariate Inverse Gaussian Distribution

If $(X, Y) \sim g_r$ then $(T, Z) \sim g_{ig}$ when $T = 1/X^2$, $Z = 1/Y^2$ and g_{ig} is Kocherlakota's (1986) bivariate inverse Gaussian density

$$\begin{aligned} g_{ig}(t,z) &= \sqrt{\frac{\lambda_1\lambda_2}{t^3 z^3(1-\rho^2)}} \cosh\left\{\frac{\rho\sqrt{\lambda_1\lambda_2}}{1-\rho^2}\left(\frac{\sqrt{t}}{\mu_1}-\frac{1}{\sqrt{t}}\right)\left(\frac{\sqrt{z}}{\mu_2}-\frac{1}{\sqrt{z}}\right)\right\} \\ &\times \ \phi\left\{\sqrt{\frac{\lambda_1}{(1-\rho^2)}}\left(\frac{\sqrt{t}}{\mu_1}-\frac{1}{\sqrt{t}}\right)\right\}\phi\left\{\sqrt{\frac{\lambda_2}{(1-\rho^2)}}\left(\frac{\sqrt{z}}{\mu_2}-\frac{1}{\sqrt{z}}\right)\right\}, \qquad (16) \end{aligned}$$

$t, z \in \mathbb{R}^+ \times \mathbb{R}^+$. Here, I have used Kocherlakota's (1986) parametrisation in which

$$\mu_i = \lambda_i/\theta_i > 0 \quad \text{and} \quad \lambda_i = 1/\sigma_i^2 > 0, \qquad i = 1, 2.$$

This is the one (out of several) bivariate inverse Gaussian distributions singled out for inclusion in Balakrishnan & Lai's (2009) compendium of bivariate distributions and said by Pal & SenGupta (2000) to be an "important bivariate distribution ... having wide applications in modelling bivariate failure-time data" but, frankly, hardly cited or used elsewhere. The transformation link between (16) and (7) when g is g_n and

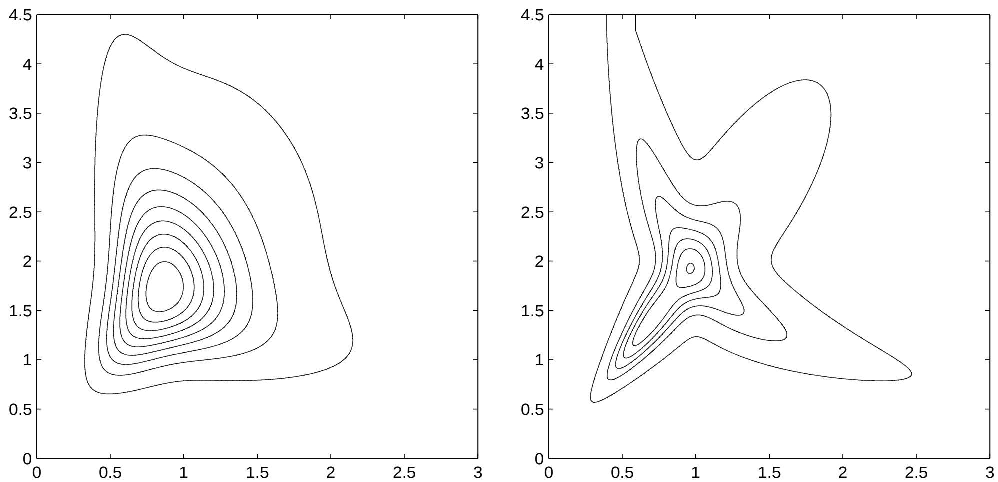


Figure 4: Contours of the bivariate inverse Gaussian density given at (16) when $\mu_1 = 1$, $\mu_2 = 2$, $\lambda_1 = 8$, $\lambda_2 = 16$, increasing from 0.02 (a) in steps of 0.1 when $\rho = 0.5$, (b) in steps of 0.25 when $\rho = 0.9$.

w_1 and w_2 are t_2 densities — the t_2 being the distribution underlying the Cauchy-Schlömilch transformation (Jones, 2013) — has also been recently pointed out by Joe, Seshadri & Arnold (2012).

The further transformation of the RRIG random variables underlying Figure 3 to obtain the corresponding bivariate IG distribution — with, of course, univariate inverse Gaussian marginals — results in the densities shown in Figure 4. The values of μ_i, λ_i used here correspond precisely to the values of θ_i, σ_i used in Figure 3. Unsurprisingly, somewhat similar shapes ensue. But they may not be unimodal (and I suspect Figure 4(b) of hiding a secondary mode on the main diagonal).

But here I have to admit that the densities in Figure 4(a) and 4(b) do not match with the supposedly same densities shown in Kocherlakota's Figure 1(b) and 1(c), respectively (whose choices of values for λ_i, μ_i, $i = 1, 2$, I have followed). Kocherlakota's contours are of a much more conventional 'pseudo-elliptical' shape than these. However, those shapes seem 'too good to be true' given their bivariate normal mixture provenance.

For bivariate R-symmetric/Cauchy–Schlömilch distributions, which include the bivariate RRIG distribution of Section 4.2, $s_i(x) = x - W_i^{-1}(x) = \theta_i/x$ and so $s_i'(x) = -\theta_i/x^2$, $i = 1, 2$. The covariance property of Section 3.2 then translates to

$$\mathrm{Cov}\left(\frac{1}{X^2}, \frac{1}{Y^2}\right) = 0,$$

for all values of the parameters. But this means that Kocherlakota's bivariate IG distribution has the property that

$$\text{Cov}(T, Z) = 0,$$

in particular, whatever the value of ρ. This example of zero covariance/correlation under dependence (when $\rho \neq 0$) was noted by Kocherlakota (1986) and Balakrishnan & Lai (2009). (It holds for any bivariate "square reciprocal R-symmetric distribution"!)

In addition, taking rescaling by a factor σ_2 into account in extension of Section 3.2 yields in general

$$E\left\{1 - s'\left(\frac{Y}{\sigma_2}\right)\big|X = x\right\} = 2$$

and in particular

$$E\left(\theta_2\sigma_2^2 Z\big|T = t\right) = 1$$

so that

$$E\left(Z|T = t\right) = \frac{1}{\theta_2\sigma_2^2} = \mu_2,$$

the corresponding constant regression property of the bivariate IG distribution also noted by Kocherlakota (1986) and Balakrishnan & Lai (2009).

5. CLOSING REMARKS

I have stopped there because of the danger of the chapter becoming dominated by what I suspect is just an interesting little sideshow, namely, Kocherlakota's bivariate inverse Gaussian distribution. It, and probably the rest of the work of Section 4, might not appeal greatly for applications because of the distorted cross shape contours evident for large ρ — though I could imagine this occurring occasionally by dint of a mechanism like that underlying this work. A formula for the multivariate extension of g_{ig} is given at (4.1) of Joe, Seshadri & Arnold (2012). With reference to the discussion in Section 1, I note that the bivariate inverse Gaussian distribution is, however, "natural", according to Kocherlakota (1986, p.1083)!

Away from the inverse Gaussian, it is the general work of Section 3 applied to more usually shaped sign-symmetric distributions, showing how to "pull them about" to achieve the desired marginal ToS distributions as well as to reasonably useful joint effect, that I think might have the greater potential.

REFERENCES

Arellano-Valle, R.B., Branco, M.D., and Genton, M.G. (2006), "A Unified View on Skewed Distributions Arising From Selections," *Canadian Journal of Statistics*, 34, 581–601.

Azzalini, A. (1985), "A Class of Distributions Which Includes the Normal Ones," *Scandinavian Journal of Statistics*, 12, 171–178.

Azzalini, A., and Capitanio, A. (1999), "Statistical Applications of the Multivariate Skew Normal Distribution," *Journal of the Royal Statistical Society Series B*, 61, 579–602.

Baker, R. (2008), "Probabilistic Applications of the Schlömilch Transformation," *Communications in Statistics – Theory and Methods*, 37, 2162–2176.

Balakrishnan, N., and Lai, C.D. (2009), *Continuous Bivariate Distributions*, second edition. New York: Springer.

Bauwens, L., and Laurent, S. (2005), "A New Class of Multivariate Skew Densities, With Application to Generalized Autoregressive Conditional Heteroscedasticity Models," *Journal of Business and Economic Statistics*, 23, 346–354.

Branco, M.D., and Dey, D.K. (2001), "A General Class of Multivariate Skew-Elliptical Distributions," *Journal of Multivariate Analysis*, 79, 99–113.

Demarta, S., and McNeil, A. (2005), "The t Copula and Related Copulas," *International Statistical Review*, 73, 111–129.

Fechner, G.T. (1897), *Kollectivmasslehre*. Leipzig: Engleman.

Fernández, C., and Steel, M.F.J. (1998), "On Bayesian Modeling of Fat Tails and Skewness," *Journal of the American Statistical Association*, 93, 359–371.

Ferreira, J.T.A.S., and Steel, M.F.J. (2007), "A New Class of Skewed Multivariate Distributions With Applications to Regression Analysis," *Statistica Sinica*, 17, 505–529.

Gupta, A.K., and Song, D. (1997), "L_p-Norm Spherical Distribution," *Journal of Statistical Planning and Inference*, 60, 241–260.

Holland, P.W., and Wang, Y.J. (1987), "Dependence Function for Continuous Bivariate Densities," *Communications in Statistics – Theory and Methods*, 16, 863–876.

Joe, H., Seshadri, V., and Arnold, B.C. (2012), "Multivariate Inverse Gaussian and Skew-Normal Densities," *Statistics and Probability Letters*, 82, 2244–2251.

Jones, M.C. (2010), "Distributions Generated by Transformations of Scale Using an Extended Schlömilch Transformation," *Sankhyā Series A*, 72, 359–375.

Jones, M.C. (2013), "Generating Distributions by Transformation of Scale," unpublished.

Jones, M.C., and Pewsey, A. (2012), "Inverse Batschelet Distributions for Circular Data," *Biometrics*, 68, 183–193.

Kocherlakota, S. (1986), "The Bivariate Inverse Gaussian Distribution: An Introduction," *Communications in Statistics – Theory and Methods*, 15, 1081–1112.

Kowalski, C.J. (1973), "Non-Normal Bivariate Distributions With Normal Marginals," *The American Statistician*, 27, 103–106.

Mudholkar, G.S., and Hutson, A. (2000), "The Epsilon-Skew-Normal Distribution for Analyzing Near-Normal Data," *Journal of Statistical Planning and Inference*, 83, 291–309.

Mudholkar, G.S., and Wang, H. (2007), "IG-Symmetry and R-Symmetry: Inter-Relations and Applications to the Inverse Gaussian Theory," *Journal of Statistical Planning and Inference*, 137, 3655–3671.

Nelsen, R.B. (2010), *An Introduction to Copulas*, second edition. New York: Springer.

Osiewalski, J., and Steel, M.F.J. (1993), "Robust Bayesian Inference in ℓ_q-Spherical Models," *Biometrika*, 80, 456–460.

Pal, C., and SenGupta, A. (2000), "Optimal Tests for No Contamination in Reliability Models," *Lifetime Data Analysis*, 6, 281–290.

Sahu, S.K., Dey, D.K., and Branco, M.D. (2003), "A New Class of Multivariate Skew Distributions With Applications to Bayesian Regression Models," *Canadian Journal of Statistics*, 31, 129–150.

Serfling, R.J. (2006), "Multivariate Symmetry and Asymmetry," in *Encyclopedia of Statistical Sciences*, second edition, vol. 8 (S. Kotz, N. Balakrishnan, C.B. Read, B. Vidakovic and N.L. Johnson, eds.), New York: Wiley, pp. 5338–5345.

Smith, M.S., Gan, Q., and Kohn, R.J. (2012), "Modelling Dependence Using Skew t Copulas: Bayesian Inference and Applications," *Journal of Applied Econometrics*, 27, 500–522.

Song, P. (2000), "Multivariate Dispersion Models Generated from Gaussian Copula," *Scandinavian Journal of Statistics*, 27, 305–320.

Wang, J.Z., Boyer, J., and Genton, M.G. (2004), "A Skew-Symmetric Representation of Multivariate Distributions," *Statistica Sinica*, 14, 1259–1270.