

# Selection models under generalized symmetry settings

Adelchi Azzalini

Received: 15 April 2010 / Revised: 20 December 2010  
© The Institute of Statistical Mathematics, Tokyo 2011

**Abstract** An active stream of literature has followed up the idea of skew-elliptical densities initiated by Azzalini and Capitanio (J. R. Stat. Soc. Ser. B 61:579–602, 1999). Their original formulation was based on a general lemma which is however of broader applicability than usually perceived. This note examines new directions of its use, and illustrates them with the construction of some probability distributions falling outside the family of the so-called skew-symmetric densities.

**Keywords** Central symmetry · Gamma distribution · Probability integral transform · Skew-normal distribution · Skew-symmetric distributions · Symmetric functions · Symmetry

## 1 Background and aims

A currently active stream of literature deals with  $d$ -dimensional continuous probability distributions such that their density function can be written in the form

$$f(z) = 2 f_0(z) G\{w(z)\}, \quad z \in \mathbb{R}^d, \quad (1)$$

where  $f_0$ , called the ‘base function’ in this paper, is a density function satisfying some form of symmetry condition, and  $G$  and  $w$  are functions whose requirements will be recalled shortly. The most prominent representative of this formulation is the

---

A. Azzalini (✉)  
Dipartimento di Scienze Statistiche, Università di Padova,  
via C. Battisti 241, 35121 Padova, Italy  
e-mail: azzalini@stat.unipd.it

skew-normal distribution whose density function at  $z$  is

$$2 \varphi_d(z; \Omega) \Phi(\alpha^\top z) \quad (2)$$

where  $\varphi_d$  denotes the  $d$ -dimensional normal density  $N_d(0, \Omega)$  and  $\Phi$  is the scalar standard normal distribution function (Azzalini and Dalla Valle 1996). In the general formulation  $\Omega$  is a covariance matrix, but for the present purposes we can restrict ourselves to the case where  $\Omega$  is a correlation matrix and  $\alpha$  is a vector of parameters which regulate the skewness of the distribution. Clearly, setting  $f_0(z) = \varphi_d(z; \Omega)$ ,  $G = \Phi$ ,  $w(z) = \alpha^\top z$  in (1) produces (2).

In fact, the chronological order of development of the two expressions above was the opposite. Near the end of a paper dedicated to the properties of distribution (2), Azzalini and Capitanio (1999) delineated a more general formulation, starting from the following result, whose proof is reproduced here because it is extremely simple, yet instructive.

**Proposition 1** *Denote by  $G$  the distribution function of a continuous random variable whose density function is symmetric about 0 and by  $Y = (Y_1, \dots, Y_d)^\top$  a continuous random variable with density function  $f_0$ . If the real-valued transform  $w(Y)$  has symmetric density about 0, then (1) is a  $d$ -dimensional density function.*

*Proof* If  $X \sim G$ , independent of  $Y$ , the distribution of  $X - w(Y)$  is symmetric about 0, implying that

$$\frac{1}{2} = \mathbb{P}\{X \leq w(Y)\} = \mathbb{E}_Y\{\mathbb{P}\{X \leq w(Y)|Y\}\} = \int_{\mathbb{R}^d} G\{w(y)\} f_0(y) \, dy. \quad (3)$$

□

The first formulation descending from the above proposition assumed  $f_0$  to be an elliptically contoured density centred at 0 and  $w(y)$  a linear function, leading to what was later called the family of skew-elliptical densities. Although the ensuing discussion mentioned that  $f_0$  does not need to be elliptical, the actual development of Azzalini and Capitanio (1999) focused on this case. The idea of skew-elliptical distributions has been followed up and expanded by several authors, including Branco and Dey (2001), Azzalini and Capitanio (2003), Genton and Loperfido (2005), and a number of contributors to the book edited by Genton (2004).

In these developments, it emerged that many results could hold replacing the assumption of elliptical distribution for  $f_0$  by the weaker assumption of central symmetry. From Serfling (2006), recall that a random variable  $Y$  is said to be centrally symmetric about 0 if  $-Y$  has the same distribution of  $Y$ ; if the density  $f_0$ , say, of  $Y$  exists, then  $f_0(y) = f_0(-y)$ . A quite general formulation is expressed in the following result presented by Azzalini and Capitanio (2003).

**Proposition 2** *Denote by  $f_0(\cdot)$  the density function of a  $d$ -dimensional continuous random variable which is centrally symmetric about 0, and by  $G$  a scalar distribution function such that  $G(-x) = 1 - G(x)$  for all real  $x$ . If  $w(z)$  is a function from  $\mathbb{R}^d$*

to  $\mathbb{R}$  such that  $w(-z) = -w(z)$  for all  $z \in \mathbb{R}^d$ , then (1) is a  $d$ -dimensional density function.

An essentially equivalent result has been obtained independently by Wang et al. (2004), with the component  $G\{w(y)\}$  in (1) replaced by a function  $\pi(y)$  which must satisfy the conditions  $0 \leq \pi(y) \leq 1$  and  $\pi(y) + \pi(-y) = 1$  for all  $y \in \mathbb{R}^d$ . The term skew-symmetric density has subsequently been adopted to denote this set of densities, or equivalently (1), when the base function  $f_0$  is centrally symmetric, to emphasize the broader settings which includes those with elliptical density as a subset.

An important result concerning these distributions is the existence of a stochastic representation for a random variable  $Z$  of this type given by

$$Z = \begin{cases} Y & \text{if } X \leq w(Y), \\ -Y & \text{otherwise,} \end{cases} \tag{4}$$

where  $X$  and  $Y$  are independent random variable with distribution function  $G$  and density function  $f_0$ , respectively, and the functions  $G$ ,  $f_0$  and  $w$  satisfy the conditions of Proposition 2. This representation is important not only for random number generation, but also for obtaining further theoretical conclusions. An especially important implication is a distributional invariance property stating that, for any function  $t(\cdot)$  from  $\mathbb{R}^d$  to  $\mathbb{R}^q$  such that  $t(y) = t(-y)$  for all  $y$ ,  $t(Z)$  and  $t(Y)$  have the same distribution, written as

$$t(Y) \stackrel{d}{=} t(Z). \tag{5}$$

Much work has been dedicated in recent years to distributions constructed via Proposition 2, or its counterpart based on the function  $\pi(x)$ . In addition to those already quoted, the review paper of Azzalini (2005) provides many other references. Since then, the list of contributions in this area has increased substantially, often leading to a very elaborate and in a sense very general constructions. These formulations are however based either on Proposition 2 or on some extension of it which still retains the key ingredient of applying a selection mechanism to a symmetric distribution, leading to a perturbation factor  $\pi(x)$  which in most cases satisfies  $\pi(x) + \pi(-x) = 1$ . Even the broad formulation of Arellano-Valle et al. (2006) effectively deals with the case where the base density  $f_0(x)$  is symmetric. The present contribution moves in a somewhat different direction, as explained next.

Notice that Proposition 2 is a special case of Proposition 1. Its role can be viewed in identifying some very simple conditions which ensure the fulfilment of those of Proposition 1. The aim of the present note is to examine situations which fall under the setting of Proposition 1 but not of Proposition 2. More specifically, we shall consider cases where  $f_0$  is not symmetric about 0, or  $w(\cdot)$  is not an odd function. As a side effect, this exploration leads to a deeper understanding of what has been developed so far in connection to Proposition 2, and it produces a more general formulation of representation (4) and of the invariance property (5).

Before tackling the specific target of this article, we note that the statement of Proposition 1 is still valid under somewhat weaker assumptions, as follows. We can relax

the assumption about absolute continuity of all distributions involved, and allow  $G$  or the distribution of  $w(Y)$  to be of discrete or of mixed type, provided the condition  $\mathbb{P}\{X - w(Y) \leq 0\} = \frac{1}{2}$  in (3) still holds. A sufficient condition to meet this requirement is that at least one of the random variables  $X$  and  $w(Y)$  is continuous.

In our development, we shall however work mostly with the original formulation of Proposition 1, with only a few remarks related to the above weaker assumptions.

## 2 Some general facts

In the framework of Proposition 1, the very argument of its proof justifies the representation of a random variable  $Z$  with the stated distribution as

$$Z = Y \quad \text{if } X \leq w(Y) \tag{6}$$

where the condition  $X \leq w(Y)$  is satisfied with probability  $\frac{1}{2}$ . It is seen that the density of  $Z$  is proportional to the integrand function of the last term in (3). There is an obvious sample selection mechanism which underlies the transformation of the distribution  $f_0$  into  $f$ .

A natural question is whether a representation similar to (4) can hold. The question is of theoretical interest, but also of practical relevance, since in random number generation the use of (4) in place of (6) avoids the rejection of half of the  $Y$  samples. Notice that the non-rejection of samples in (4) is achieved by exploiting the symmetry of the distribution  $f_0$ .

It appears that some additional conditions must be required to provide a solution to the above problem. One such set of conditions is as follows. If there exists an invertible transformation  $R(\cdot)$  such that, for all  $y \in \mathbb{R}^d$ ,

$$f_0(y) = f_0[R(y)], \quad |\det R'(y)| = 1, \quad w[R(y)] = -w(y), \tag{7}$$

where  $R'(y)$  denotes the Jacobian matrix of the partial derivatives, then

$$Z = \begin{cases} Y & \text{if } X \leq w(Y), \\ R^{-1}(Y) & \text{otherwise,} \end{cases} \tag{8}$$

has distribution (1). In fact the density function of  $Z$  at  $z$  is

$$\begin{aligned} f(z) &= f_0(z) G\{w(z)\} + f_0(R(z)) |\det R'(z)| [1 - G\{w(R(z))\}] \\ &= f_0(z) G\{w(z)\} + f_0(z) [1 - G\{-w(z)\}] \\ &= 2 f_0(z) G\{w(z)\} \end{aligned}$$

using (7) and  $G(-x) = 1 - G(x)$ . Under the weaker assumptions indicated at the end of Sect. 1, the latter equality may not hold for all  $x$ . However this may affect only a number of  $x$  values at most countable, and the density function  $f(\cdot)$  can be replaced by a regularized version without affecting the distribution.

According to the first two conditions in (7), the density  $f_0$  is required to behave according to a “generalized symmetry” with respect to  $R(\cdot)$ , that is  $R^{-1}(Y)$  must have the same density  $f_0$  of  $Y$ . In addition  $w(\cdot)$  must be an odd function in this generalized sense, as required by the third condition (7). In Proposition 2, in (4) and in (5), the transformation function is  $R(z) = -z = R^{-1}(z)$ .

If in (6) we reverse the sign of the inequality, this generates the dual variable of  $Z$ , having density function

$$2 f_0(y) [1 - G\{w(y)\}] = 2 f_0(y)G\{-w(y)\} = 2 f_0(y)G\{w(R(y))\}$$

where the last equality makes use of the third condition (7).

For the analog of the invariance property (5), consider a transformation  $t(\cdot)$  from  $\mathbb{R}^d$  to  $\mathbb{R}^q$  which is even in the adopted generalized sense, that is

$$t(z) = t(R^{-1}(z)) \tag{9}$$

for all  $z \in \mathbb{R}^d$ . From representation (8), it is then immediate that (5) holds.

The function  $R(\cdot)$  which satisfy (7) does not need to be unique. Typically, if  $R$  is one such function,  $R^{-1}$  is another one. In some cases, there are more than two functions  $R$  which satisfy (7), as we shall see in Sect. 3.1. *Vice versa*, in other cases  $R^{-1} = R$ , for instance when  $R(z) = -z$ , and we effectively have a single transformation.

In the one-dimensional case a general way to construct a random variable with symmetric distribution is via its integral transform. Specifically, if  $Y$  has distribution function  $F_0$ , then  $F_0(Y) - \frac{1}{2}$  is uniformly distributed on  $(-\frac{1}{2}, \frac{1}{2})$  and application of Proposition 1 provides the following conclusion.

**Proposition 3** *Assume that  $f_0$  is a probability density function with distribution function  $F_0$  on  $S_0 \subset \mathbb{R}$ ,  $G$  is a distribution function over  $\mathbb{R}$  which assigns a probability distribution symmetric about 0 and it is continuous except possibly at 0, and  $w_1$  is an odd function on  $(-\frac{1}{2}, \frac{1}{2})$ . Then*

$$2 f_0(x) G \left\{ w_1 \left[ F_0(x) - \frac{1}{2} \right] \right\}, \quad 2 f_0(x) G \left\{ w_1 \left[ \frac{1}{2} - F_0(x) \right] \right\}, \quad (x \in S_0 \subset \mathbb{R}), \tag{10}$$

are density functions on  $S_0$ .

The same result could however have been equally obtained from Proposition 2, although via a slightly more involved argument, as follows. Start by taking the base density function to be the uniform density in  $(-\frac{1}{2}, \frac{1}{2})$ ,  $w(z) = w_1(z)$ , and then shift the distribution to the interval  $(0, 1)$ ; we arrive at the density  $2 G \{w_1(x - \frac{1}{2})\}$ . Finally, transform this distribution via  $F_0^{-1}(\cdot)$ , to produce the first expression in (10). The second expression is obtained similarly using  $w(z) = w_1(-z)$ .

For  $d > 1$ , the probability integral transform is not so easily tractable as for  $d = 1$ . A manageable case occurs with  $d = 2$  and independent marginal components. If  $Y_1$  and  $Y_2$  are continuous independent variables with distribution functions  $F_1$  and  $F_2$ , respectively, then a simple computation gives

$$p(t) = \mathbb{P}\{F_1(Y_1) F_2(Y_2) \leq t\} = t(1 - \log t), \quad (0 < t < 1).$$

Recall that  $p[F_1(Y_1) F_2(Y_2)]$  is uniformly distributed in  $(0, 1)$ . If  $f_1 = F'_1$ ,  $f_2 = F'_2$  and  $w_1$  is an odd function, then from Proposition 1

$$2 f_1(y_1) f_2(y_2) G \left\{ w_1 \left( p[F_1(y_1) F_2(y_2)] - \frac{1}{2} \right) \right\}$$

is a proper probability density on the support given by the Cartesian product of the support sets of  $F_1$  and  $F_2$ .

This expression is of general validity for  $d = 2$ , but it has the disadvantage that the base function  $f_1(y_1)f_2(y_2)$  does not allow for dependence between the marginal components. In addition, in many cases, the distributions so produced are not very convenient to work with from a mathematical viewpoint. For these reasons, in the following section we examine other constructions, which are simpler to handle. Moreover they are more directly related to the formulation expressed by (7–9); hence they better serve the purpose of illustrating the underlying mechanism.

### 3 Some specific constructions

We shall now examine a few specific distributions to illustrate the formulation of the previous section. Given this aim, a complete study of the properties of these distributions is not attempted here. To ease exposition, we shall concentrate on the case  $d = 2$ .

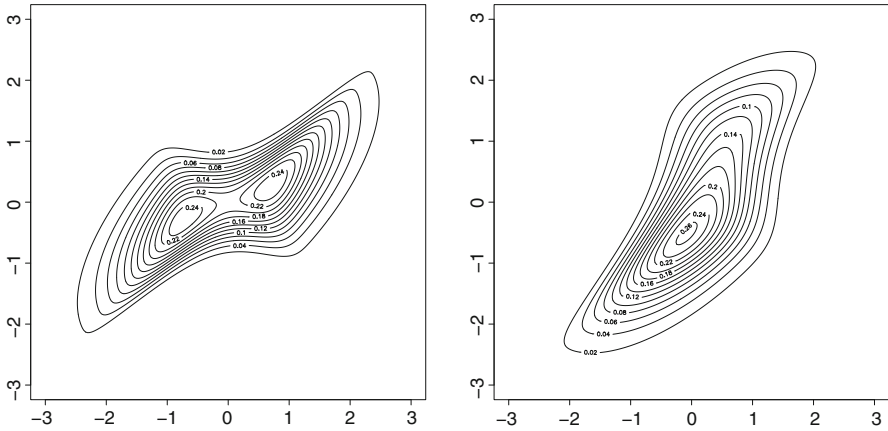
#### 3.1 Cases with $w(y)$ non-odd

To start with a simple example, consider the probability distribution on  $\mathbb{R}^2$  whose density function at  $z = (z_1, z_2)^\top$  is

$$f(z) = 2 \varphi_2(z; \Omega) \Phi \left\{ \alpha \left( z_1^2 - z_2^2 \right) \right\} \quad (11)$$

where  $\Omega$  is a  $2 \times 2$  correlation matrix with off-diagonal term  $\rho$ , and  $\alpha$  is a real parameter. This  $f(z)$  bears a superficial resemblance to the skew-normal density (2) for  $d = 2$ , but with the noticeable difference from the skew-normal distribution that (11) is not skew at all; in fact it is centrally symmetric. The effect of modification of the normal density when  $\alpha = 2$  and  $\rho = 2/3$  is shown graphically in the left panel of Fig. 1. This density is bimodal, but other choices of  $\alpha$  produce a unimodal density.

The fact that the appropriate normalization constant in (11) is 2 cannot however follow from Proposition 2 which requires an odd function  $w$ , while  $w(z) = \alpha(z_1^2 - z_2^2)$  is even. Equivalently,  $\pi(z) = \Phi\{\alpha(z_1^2 - z_2^2)\}$  does not satisfy the condition  $\pi(z) + \pi(-z) = 1$  required in the formulation of Wang et al. (2004). However, if  $Y = (Y_1, Y_2)^\top \sim N_2(0, \Omega)$ , it is true that  $\alpha(Y_1^2 - Y_2^2)$  has a symmetric distribution about 0, and so Proposition 1 can be applied to conclude that (11) integrates to 1. In this respect, it would be irrelevant to replace  $\Phi$  in (11) by some other distribution  $G$  with  $G'$  even.



**Fig. 1** Contour level plot of density function (11) when  $\alpha = 2$  and  $\rho = 2/3$  (left panel) and similar density when the last factor is  $\Phi\{z_1 - z_2 - (z_1^2 - z_2^2)\}$  (right panel)

Much more general forms than (11) can however be handled, taking into account the following statement, whose proof is trivial and omitted.

**Proposition 4** *If  $(X, Y)$  is a random variable on  $\mathbb{R}^2$  with density function  $h(x, y)$  which is a symmetric function of the variables, that is  $h(x, y) = h(y, x)$ , and  $w_2(x, y)$  is a real-valued function such that  $w_2(y, x) = -w_2(x, y)$ , then  $w_2(X, Y)$  has symmetric distribution around 0.*

A simple example of function  $w_2$  which fulfils the above requirement is

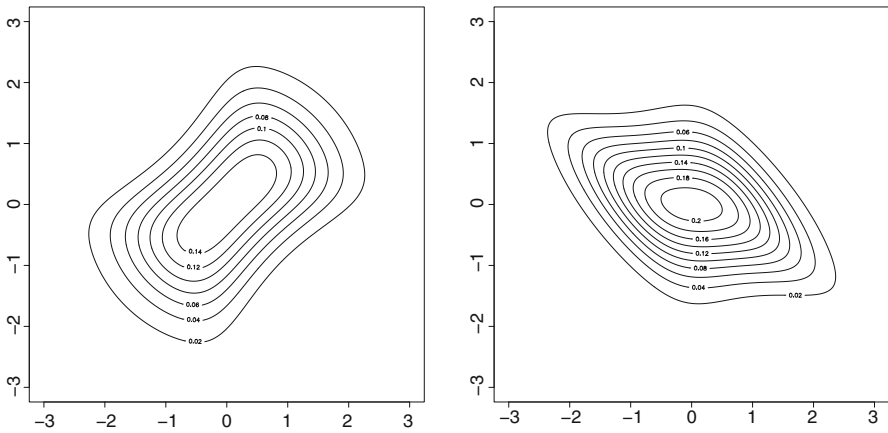
$$w(z) = w_2(z_1, z_2) = \alpha_1(z_1 - z_2) + \dots + \alpha_m(z_1^m - z_2^m) \tag{12}$$

where  $z = (z_1, z_2)^T \in \mathbb{R}^2$  and  $m$  is some natural number. The right-side panel of Figure 1 shows the contour level plot of the density obtained when the argument of  $\Phi$  in (11) is replaced by (12) with  $m = 2$ ,  $\alpha_1 = 1$ ,  $\alpha_2 = -1$ , and  $\rho = 2/3$  as before. Notice that in general (12) is neither odd nor even, if the coefficients  $\alpha_1, \dots, \alpha_m$  are unrestricted. It is even if all  $\alpha_j$ 's of odd order are 0, and it is odd in the dual case with only coefficients of odd order, which leads us back to the setting of Proposition 2. Combining Proposition 1 with Proposition 4, we can state the following corollary.

**Proposition 5** *If  $f_0(z)$  is a density function centrally symmetric over  $\mathbb{R}^2$  and  $f_0$  is a symmetric function of the components of  $z$ ,  $w_2$  is as in Proposition 4 and  $G$  is a continuous distribution function over  $\mathbb{R}$  which assigns a probability distribution symmetric about 0 and it is continuous except possibly at 0, then*

$$2 f_0(z) G\{w_2(z_1, z_2)\}, \quad z = (z_1, z_2)^T \in \mathbb{R}^2, \tag{13}$$

is a density function over  $\mathbb{R}^2$ .



**Fig. 2** Contour level plot of density function (15) when  $\alpha = \sqrt{\pi/2}$  (left panel) and density function given by the first expression in (17) when  $\rho = -2/3$  and  $\alpha = \sqrt{\pi/2}$  (right panel)

For simplicity, we now restrict ourselves to the case where  $f_0(z)$  in (13) is  $\varphi_2(z; \Omega)$  and  $w_2$  is as in Proposition 4, but the essence of the conclusions remains true for other centrally symmetric densities  $f_0$ , symmetric of the arguments. The stochastic representation (8) for a random variable  $Z = (Z_1, Z_2)^\top$  with such density function holds by choosing a transformation which swaps the two variables, that is

$$R(y) = R_0 y, \quad R_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = R_0^{-1}, \tag{14}$$

and consequently also (9) applies for suitable  $t$ 's. Among the many implications of this fact, the Mahalanobis distance  $Z^\top \Omega^{-1} Z$  is  $\chi_2^2$ . Another consequence is that, since  $t(y) = y_1 y_2 = y_2 y_1 = t(R_0 y)$ , then  $\mathbb{E}\{Z_1 Z_2\} = \mathbb{E}\{Y_1 Y_2\} = \rho$ . When central symmetry of  $f(z)$  holds, for instance when (12) is even, this implies that  $\mathbb{E}\{Z_1\} = \mathbb{E}\{Z_2\} = 0$ , and then  $\text{cov}\{Z_1, Z_2\} = \rho$ .

We now move to a different example, and consider the bivariate distribution studied by Arnold et al. (2002) whose density function at  $z = (z_1, z_2)^\top \in \mathbb{R}^2$  is

$$f(z) = 2 \varphi_2(z; I_2) \Phi(\alpha z_1 z_2), \tag{15}$$

which is centrally symmetric too, and it enjoys several interesting properties. If  $Z = (Z_1, Z_2)^\top$  is a random variable with this distribution, then each of  $Z_1, Z_2$  is marginally standard normal, and the distribution of each component conditional on the value  $u$ , say, of the other one is univariate skew-normal with parameter  $\alpha, u$ , that is of type (2) with  $d = 1$ . Other properties are presented in the quoted paper, including an expression for  $\text{corr}\{Z_1, Z_2\}$ . A graphical representation of this density is shown in the left-side panel of Fig. 2 for the case  $\alpha = \sqrt{\pi/2} \approx 1.253$  which has been proved by Arnold et al. (2002) to be the highest possible value producing a unimodal density; with larger  $\alpha$  there are two modes.

From the viewpoint of the present context, (15) has a function  $w(z) = \alpha z_1 z_2$  which is even, and again Proposition 2 does not apply. It is however true that  $w(Y)$  has a symmetric distribution around 0, if  $Y \sim N_2(0, I_2)$ , and Proposition 1 confirms that 2 is the appropriate normalization constant in (15). More importantly, this explains that the normalizing factor 2 in (15) is not occurring by accident, and it connects the distribution of Arnold et al. (2002) with an apparently unrelated type of construction.

To fulfil conditions (7) we can choose  $R(y)$  to be any of  $R_j y$  for  $j = 1, \dots, 4$  where

$$R_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad R_2 = -R_1, \quad R_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_4 = -R_3. \tag{16}$$

The first two of these matrices correspond to  $\pi/2$  and  $-\pi/2$  rotation, respectively, hence  $R_2 = R_1^\top = R_1^{-1}$  and  $R_4 = I_2$ . This provides a stochastic representation for  $Z$ , and so a mechanism for random number generation. Using (9) with  $t(z) = z_1^2 + z_2^2$ , we can state that  $Z_1^2 + Z_2^2 \sim \chi_2^2$ . Using  $t(z) = (z_1 z_2)^2$  we obtain that  $\mathbb{E}\{(Z_1 Z_2)^2\} = 1$ , which in turns implies  $\text{corr}\{Z_1^2, Z_2^2\} = 0$ , recalling that  $Z_1$  and  $Z_2$  have standardized marginals.

If in (15) we replace  $I_2$  by a correlation matrix  $\Omega$  with off-diagonal element  $\rho$ , then  $w(Y) = \alpha Y_1 Y_2$  does no longer give rise to a symmetric distribution around 0 if  $(Y_1, Y_2)^\top \sim N_2(0, \Omega)$ ; hence Proposition 1 does not apply. The symmetry condition is fulfilled instead by  $w(Y) = \alpha Y_1(Y_2 - \rho Y_1)$  and by the dual function  $w(Y) = \alpha Y_2(Y_1 - \rho Y_2)$ , since we have again the product of independent normal variables with 0 mean. It then follows that

$$2 \varphi_2(z; \Omega) \Phi\{\alpha z_1 (z_2 - \rho z_1)\}, \quad 2 \varphi_2(z; \Omega) \Phi\{\alpha z_2 (z_1 - \rho z_2)\} \tag{17}$$

are legitimate density functions on  $\mathbb{R}^2$ . The right-side panel of Fig. 2 displays the first of these densities for the case  $\rho = -2/3$  and  $\alpha = \sqrt{\pi/2}$  as before.

To search for a function  $R(\cdot)$  fulfilling (7), we start by imposing the first and the third condition indicated by (7). For each given point  $z_0 \in \mathbb{R}^2$ , it is required to find a point which lies on the ellipse representing the locus of all points having the same density of  $z_0$  and it also lies on the locus of all points having  $w(z)$  equal to  $-w(z_0)$ , that is to solve the equations

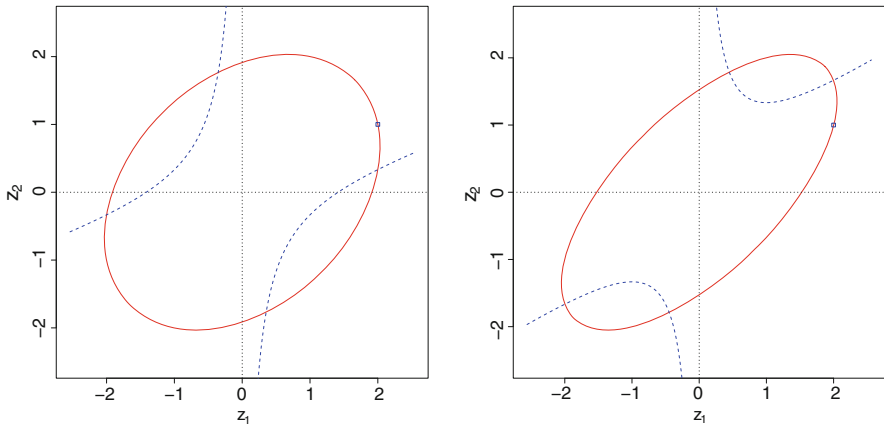
$$D(z) = D_0, \quad w(z) = -w_0$$

where

$$D(z) = z^\top \Omega^{-1} z = \frac{1}{1 - \rho^2} (z_1^2 - 2\rho z_1 z_2 + z_2^2), \quad w(z) = z_1 z_2 - \rho z_1^2$$

and  $D_0 = D(z_0)$ ,  $w_0 = w(z_0)$ . After simple algebra one arrives at a fourth degree equation for  $z_1$  with only even-order terms, or equivalently to the quadratic equation

$$(1 - \rho^2)u^2 - D_0(1 - \rho^2)u + w_0^2 = 0$$



**Fig. 3** Illustration of the possible choice of  $R(z_0)$  for the first of the densities (17) when  $z_0 = (2, 1)^\top$ ,  $\alpha = 1$  and  $\rho = 1/3$  (left panel) or  $\rho = 2/3$  (right panel). The ellipse represents the locus of points with the same density of  $z_0$ ; the dashed lines denote the loci of points with  $w(z)$  equal to  $-w(z_0)$

where  $u = z_1^2$ . The roots of this equation lead to four values for  $z_1 = \pm\sqrt{u}$  and four corresponding  $z_2 = (\rho z_1^2 - w_0)/z_1$ . There is a clear similarity with the earlier case where  $\rho = 0$  and four transformations associated to matrices (16), but now the four transformations  $R(\cdot)$  are non-linear. The construction is illustrated by Fig. 3 for the given point  $z_0 = (2, 1)^\top$ ,  $w(z) = z_1(z_2 - \rho z_1)$  and  $\rho$  equal either to  $1/3$  or to  $2/3$ ; the change of pattern in the two cases depends on the sign of  $w(z_0)$  as  $\rho$  varies.

The final step is to check whether the second condition in (7) holds. Some algebraic work not reported here shows that  $|\det R'| \neq 1$ . The implication is that, while (17) are valid density functions, a stochastic representation of type (8) does not hold for them. However (6) still holds.

### 3.2 Cases where $f_0$ is not a symmetric density

There are cases where  $Y$  is a continuous random variable with density function  $f_0$  which is not symmetric about 0 but there still are functions  $w(Y)$  with symmetric density about 0, so that Proposition 1 can be employed. Some constructions of this form have been sketched in Sect. 2.

If  $d = 2$ , a simple but quite general formulation of this type can be obtained when  $f_0(\cdot)$  and  $w(\cdot)$  satisfy the requirements for  $h$  and  $w_2$  in Proposition 4, respectively. In other words, we are asking that

$$f_0(y) = f_0(R_0 y), \quad w(R_0 y) = -w(y)$$

where  $R_0$  is as in (14). Under this setting, (8) holds with  $R(y) = R_0 y$ .

A simple illustrative example of this formulation can be started by choosing  $f_0$  to be the product of two Gamma densities, such that each marginal density is

$$f_1(x) = \frac{1}{\Gamma(\omega)} x^{\omega-1} e^{-x}$$

if  $x > 0$ , and 0 otherwise, for some  $\omega > 0$ . For  $f_0(y) = f_1(y_1) f_1(y_2)$  where  $y = (y_1, y_2)^\top$ , there is a natural line of reflection given by the identity line. Among the many feasible functions  $w$ , we choose the simple linear case  $w(y) = \alpha(y_1 - y_2)$  where  $\alpha$  is an arbitrary parameter. The density under consideration is then

$$f(z_1, z_2) = \frac{2}{\Gamma(\omega)^2} (z_1 z_2)^{\omega-1} \exp(-z_1 - z_2) G\{\alpha(z_1 - z_2)\} \tag{18}$$

for  $z_1, z_2 > 0$ , and 0 otherwise. Here  $G$  is as required in Proposition 1 with possibly a discontinuity at 0. If  $Y = (Y_1, Y_2)^\top$  has density  $f_0$ , then  $f(z_1, z_2)$  is the distribution of  $Y$  conditionally on the event  $\{X < \alpha(Y_1 - Y_2)\}$ , where  $X \sim G$  is independent of  $Y$ .

Various implications follow from (5) for a random variable  $Z = (Z_1, Z_2)^\top$  having density function (18). Some examples are

$$\mathbb{E}\{Z_1 + Z_2\} = \mathbb{E}\{Y_1 + Y_2\} = 2\omega, \quad \mathbb{E}\{Z_1 Z_2\} = \mathbb{E}\{Y_1 Y_2\} = \omega^2,$$

irrespective of  $G$ , since both  $t(y) = y_1 + y_2$  and  $t(y) = y_1 y_2$  satisfy (9).

To compute the marginal distributions of  $Z_1$  and  $Z_2$ , assume for the moment that  $\alpha > 0$ , and rewrite the event  $\{X < \alpha(Y_1 - Y_2)\}$  as  $\{T < Y_1\}$  where  $T = \alpha^{-1} X + Y_2$ . Denote by  $F_T$  the distribution function of  $T$ , which is the integral of the density

$$f_T(t) = \int_0^\infty f_1(y_2) g[\alpha(t - y_2)] \alpha \, dy_2 \tag{19}$$

where  $g = G'$ . It is simple to obtain the density functions of  $Z_1$  and  $Z_2$  which are

$$2 f_1(z_1) F_T(z_1), \quad 2 f_1(z_2) \{1 - F_T(z_2)\}, \quad (z_1 > 0, z_2 > 0), \tag{20}$$

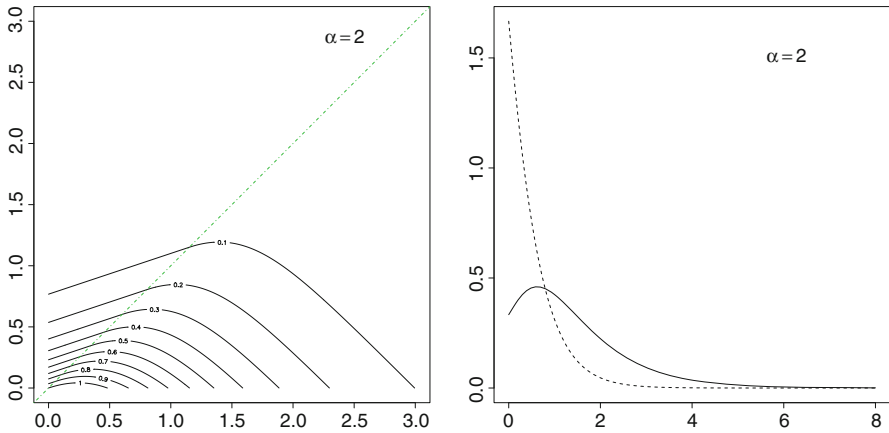
respectively, taking into account that  $\mathbb{P}\{T < Y_1\} = \frac{1}{2}$ . When  $\alpha < 0$ , the density function of  $T$  is still computed from (19), with  $\alpha$  replaced by its absolute value, since the distribution of  $\alpha^{-1} X$  does not depend on the sign of  $\alpha$ , and the above expressions of the densities of  $Z_1$  and  $Z_2$  are exchanged. Another way of looking at this aspect is to take into account (8) and to notice that reversing the sign of  $\alpha$  corresponds to swapping the distributions of the two components.

The explicit computation of (19) is feasible for a suitable form of  $g$ . A convenient option is to set  $g$  equal to the Laplace density

$$g(x) = \frac{1}{2} e^{-|x|}, \quad (x \in \mathbb{R}),$$

leading to

$$f_T(t) = \begin{cases} \frac{\alpha}{2(1+\alpha)^\omega} e^{\alpha t} & \text{if } t \leq 0, \\ \frac{\alpha}{2\Gamma(\omega)} (e^{-\alpha t} I_1 + e^{\alpha t} I_2) & \text{otherwise,} \end{cases}$$



**Fig. 4** Density function of  $Z$  in the exponential-Laplace case when  $w(y) = 2(y_1 - y_2)$  in the left panel, and marginal density of  $Z_1$  (continuous line) and  $Z_2$  (dashed line) in the right panel

where, for integer  $\omega$ ,

$$\begin{aligned}
 I_1 &= \int_0^t x^{\omega-1} e^{-x(1-\alpha)} dx \\
 &= \begin{cases} \frac{t^\omega}{\omega} & \text{if } \alpha = 1, \\ \Gamma(\omega) \left\{ \frac{1}{(1-\alpha)^\omega} - e^{-t(1-\alpha)} \sum_{k=0}^{\omega-1} \frac{t^k}{k! (1-\alpha)^{\omega-k}} \right\} & \text{if } \alpha \neq 1, \end{cases} \\
 I_2 &= \int_t^\infty x^{\omega-1} e^{-x(1+\alpha)} dx \\
 &= \Gamma(\omega) e^{-t(1+\alpha)} \sum_{k=0}^{\omega-1} \frac{t^k}{k! (1+\alpha)^{\omega-k}}.
 \end{aligned}$$

It is possible to produce a general expression of the distribution function  $F_T$ . However, for simplicity, we now restrict ourselves to the case  $\omega = 1$ , such that

$$F_T(t) = \begin{cases} \frac{1}{2(1+\alpha)} e^{\alpha t} & \text{if } t \leq 0, \\ 1 - e^{-t} \left( \frac{3}{4} + \frac{t}{2} \right) & \text{if } t > 0, \alpha = 1, \\ 1 - \frac{\alpha^2}{\alpha^2-1} e^{-t} + \frac{1}{2(\alpha-1)} e^{-\alpha t} & \text{if } t > 0, \alpha \neq 1, \end{cases}$$

when  $\alpha > 0$ . By inserting this expression of  $F_T(t)$  in (20), we obtain the marginal density functions of  $Z_1$  and  $Z_2$ . Figure 4 displays the bivariate density function of  $Z$  and the marginal density of  $Z_1$  and  $Z_2$  when  $\alpha = 2$ .

The moments of the marginal densities can be computed by direct integration, which for  $\alpha > 0$  and any  $r > 0$  gives

$$\mathbb{E}\{Z_1^r\} = \begin{cases} \Gamma(r + 1) \left(2 - \frac{r+4}{2^{r+2}}\right) & \text{if } \alpha = 1, \\ \Gamma(r + 1) \left(2 - \frac{1}{\alpha^2-1} \left[\frac{\alpha^2}{2^r} - \frac{1}{(1+\alpha)^r}\right]\right) & \text{if } \alpha > 0, \alpha \neq 1, \end{cases}$$

$$\mathbb{E}\{Z_2^r\} = 2 \Gamma(r + 1) - \mathbb{E}\{Z_1^r\}$$

while the moments for  $\alpha < 0$  are obtained by swapping the subscripts of  $Z_1$  and  $Z_2$  in the above expressions and taking the absolute value of  $\alpha$ . From marginal moments up to second order and from the fact obtained earlier that  $\mathbb{E}\{Z_1 Z_2\} = 1$ , we can compute the correlation. Numerical evaluation indicates that the correlation increases monotonically from 0 when  $\alpha = 0$  to  $1/\sqrt{5} \approx 0.447$  when  $\alpha = \pm\infty$ .

We close this section with some remarks on possible extensions. The above discussion has focused on the case with linear  $w(y) = \alpha(y_1 - y_2)$  but many other options are possible, such as  $w(y) = \alpha(y_1^2 - y_2^2)$  or  $w(y) = \sin[\alpha(y_1 - y_2)]$ . Another direction is to adopt a base density  $f_0$  with dependent components, since our choice of using independent components was only for mathematical simplicity. The overview of bivariate exponential distributions presented in Section 47.2 of [Kotz et al. \(2000\)](#) includes several distributions which meet this requirement. Among them, two especially important proposals are the Gumbel and the Marshall–Olkin bivariate exponential distributions, whose joint survival functions are

$$\exp\{-y_1 - y_2 - \lambda y_1 y_2\}, \quad \exp\{-y_1 - y_2 - \lambda \max(y_1, y_2)\}$$

respectively, for  $y_1, y_2 \geq 0$ , when scale factors are not included; here  $\lambda$  is a positive parameter, and  $\lambda < 1$  for the first case.

#### 4 Final remarks

We conclude with some remarks on cases with dimension  $d > 2$ , when the base function  $f_0$  is centrally symmetric. To ensure that (1) is a proper density we need that  $w(Y)$  is symmetrically distributed around 0 when  $Y$  has density  $f_0$ . The case when  $w(\cdot)$  is odd, which of course includes linear functions, is covered by Proposition 2, and symmetry of  $w(Y)$  holds in general.

For a non-odd  $w$  which is a function of two components of  $Y$  only, we can still make use of Proposition 5 with a  $d$ -dimensional  $f_0$ , by using the following argument. If  $f_0$  is centrally symmetric, the same is true for the distribution of any subset of its components. Hence a function  $w$  of the nominated components of the form required by Proposition 4 produces a symmetric  $w(Y)$ .

We are then left with the case where  $d > 2$  and  $w(Y)$  is a non-odd function of at least three components of  $Y$ . The characterization of the set of such functions  $w$  is a problem which does not appear amenable to a simple solution, and it deserves a separate study, given also its interest as a property of centrally symmetric distributions.

**Acknowledgments** I would like to thank Antonella Capitanio stimulating discussions on this problem and Giuliana Regoli for several insightful remarks on a preliminary version of the paper. My thanks also go to Ahad Jamalzadeh for drawing my attention to an incorrect quotation of the literature.

## References

- Arellano-Valle, R. B., Branco, M. D., Genton, M. G. (2006). A unified view on skewed distributions arising from selections. *Canadian Journal of Statistics*, *34*, 581–601.
- Arnold, B. C., Castillo, E., Sarabia, J. M. (2002). Conditionally specified multivariate skewed distributions. *Sankhyā, series A*, *64*, 206–226.
- Azzalini, A. (2005). The skew-normal distribution and related multivariate families (with discussion). *Scandinavian Journal of Statistics*, *32*, 159–188 (C/R 189–200).
- Azzalini, A., Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society, series B*, *61*, 579–602. Full version of the paper at [arXiv.org:0911.2093](https://arxiv.org/abs/0911.2093).
- Azzalini, A., Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew  $t$  distribution. *Journal of the Royal Statistical Society, series B*, *65*, 367–389. Full version of the paper at [arXiv.org:0911.2342](https://arxiv.org/abs/0911.2342).
- Azzalini, A., Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, *83*, 715–726.
- Branco, M. D., Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, *79*, 99–113.
- Genton, M. G., Ed. (2004). *Skew-elliptical Distributions and Their Applications: a Journey Beyond Normality*. Boca Raton: Chapman & Hall/CRC.
- Genton, M. G., Loperfido, N. (2005). Generalized skew-elliptical distributions and their quadratic forms. *Annals of the Institute Statistical Mathematics*, *57*, 389–401.
- Kotz, S., Balakrishnan, N., Johnson, N. L. (2000). *Continuous Multivariate Distributions Volume 1: Models and Applications*. (2nd ed). New York: J. Wiley & Sons.
- Serfling, R. (2006). Multivariate symmetry and asymmetry. In S. Kotz, N. Balakrishnan, C. B. Read, & B. Vidakovic (Eds.), *Encyclopedia of Statistical Sciences* (2nd ed, Vol. 8, pp. 5338–5345). New York: J. Wiley & Sons.
- Wang, J., Boyer, J., Genton, M.G. (2004). A skew-symmetric representation of multivariate distributions. *Statistica Sinica*, *14*, 1259–1270.