Ann Inst Stat Math (2012) 64:857–879
DOI 10.1007/s10463-011-0338-5

Some properties of skew-symmetric distributions

Adelchi Azzalini · Giuliana Regoli

Received: 22 December 2010 / Revised: 12 May 2011 / Published online: 10 September 2011

Abstract The family of skew-symmetric distributions is a wide set of probability density functions obtained by suitably combining a few components which can be quite freely selected provided some simple requirements are satisfied. Although intense recent work has produced several results for certain sub-families of this construction, much less is known in general terms. The present paper explores some questions within this framework and provides conditions for the above-mentioned components to ensure that the final distribution enjoys specific properties.

Keywords Central symmetry · Log-concavity · Peakedness · Quasi-concavity · Skew-symmetric distributions · Stochastic ordering · Strong unimodality · Unimodality

1 Introduction and motivation

1.1 Distributions generated by perturbation of symmetry

In recent years, there has been quite intense work on a broad class of absolutely continuous probability distributions which are generated starting from symmetric density functions and applying suitable forms of perturbation of the symmetry. The key representative of this formulation is the skew-normal distribution, whose density function

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

G. Regoli
Dipartimento di Matematica e Informatica, Università di Perugia, via Vanvitelli, 1, 06123 Perugia, Italy
e-mail: regoli@dipmat.unipg.it

in the scalar case is given by

$$f(x;\alpha) = 2\,\phi(x)\,\Phi(\alpha\,x), \qquad (x \in \mathbb{R}), \tag{1}$$

where $\phi(x)$ and $\Phi(x)$ denote the N(0,1) density function and distribution function, respectively, and α is an arbitrary real parameter. When $\alpha = 0$, (1) reduces to familiar $\mathrm{N}(0, 1)$ distribution; otherwise, an asymmetric distribution is obtained, with skewness having the same sign as α. The properties of (1) studied by Azzalini (1985) and other authors show a number of similarities with the normal distribution and support the adoption of the name 'skew-normal'.

The same sort of mechanism leading from the normal density function to (1) has also been applied to other symmetric distributions, including extensions to more elaborate forms of perturbation and constructions in a multivariate setting. Introductory accounts to this research area are provided by the book edited by Genton (2004) and the review paper of Azzalini (2005), to which readers are referred for a general overview.

For the aims of the present paper, we rely largely on the following lemma, presented by Azzalini and Capitanio (2003). This is very similar to a result developed independently by Wang et al. (2004); the precise interconnections between the two statements will be discussed in the course of the paper.

Before stating the result, let us recall that the concept of a symmetric density function has a single simple definition only in the univariate case, but different formulations exist in the multivariate case; see Serfling (2006) for an overview. In this paper, we adopt the concept of central symmetry which, in the case of continuous distribution on $\mathbb{R}^d$, requires a density function p which satisfies $p(x - x_0) = p(x_0 - x)$ for all $x \in \mathbb{R}^d$, for some centre of symmetry x_0.

Lemma 1 *Denote a d-dimensional probability density function centrally symmetric about* 0 *by f_0, a continuous distribution function on the real line by $G_0(\cdot)$, so that $g_0 = G_0'$ exists a.e. and is an even density function, and an odd real-valued function on $\mathbb{R}^d$ by w, so that $w(-x) = -w(x)$. Thus*

$$f(x) = 2\,f_0(x)\,G_0\{w(x)\}, \qquad (x \in \mathbb{R}^d), \tag{2}$$

is a density function.

It is useful to recall the essence of the proof, which is very simple. Denote a continuous random variable with density f_0 by X_0, and a univariate continuous random variable with distribution function G_0, independent of X_0, by U. It is easy to see that $w(X_0)$ has distribution symmetric about 0, and so the same is true for $T = U - w(X_0)$. We can then write

$$\frac{1}{2} = \mathbb{P}\{T \le 0\} = \mathbb{E}\{\mathbb{P}\{U \le w(X_0)|X_0\}\} = \int_{\mathbb{R}^d} G\{w(x)\}\,f_0(x)\,\mathrm{d}x \tag{3}$$

which states that (2) is a density function.

Lemma 1 provides a general mechanism for modifying initial symmetric 'base' density f_0 via the perturbation factor $G(x) = G_0\{w(x)\}$, whose components G_0 and

w can be chosen from a wide set of options. Clearly, the prominent case (1) can be obtained by setting $d = 1$, $f_0 = \phi$, $G_0 = \Phi$, $w(x) = \alpha\, x$ in (2). The term 'skew-symmetric' is often adopted for distributions of type (2).

Note that twice the integrand of the last term in (3) can be written as

$$f(x) = 2\, f_0(x)\, \mathbb{P}\{T < 0 | X_0 = x\} \tag{4}$$

and this provides an interpretation of (2) as the effect of a selection mechanism on a population with distribution f_0 via the condition $\{T < 0 | X_0 = x\}$. Although this representation is mentioned here in connection with the above-mentioned result by Azzalini and Capitanio (2003), it has been a persistent idea in this stream of literature since its very beginning. See Azzalini (1985, 1986) for early occurrences; more recent works focusing especially on this direction are those of Arnold and Beaver (2000), Arellano-Valle et al. (2002, 2006).

The above argument is also the basis for the important stochastic representation of variable X with density (2) as

$$X = \begin{cases} X_0 & \text{if} \quad U \le w(X_0), \\ -X_0 & \text{if} \quad U \ge w(X_0), \end{cases} \tag{5}$$

and (5) immediately supplies the following important corollary.

Proposition 1 (Perturbation invariance) *If random variable X_0 has density f_0 and X has density f, where f_0 and f satisfy the conditions required in Lemma 1, then the equality*

$$t(X) \overset{d}{=} t(X_0), \tag{6}$$

where '$\overset{d}{=}$' denotes equality in distribution, holds for any even q-dimensional function t on $\mathbb{R}^d$, irrespective, of factor $G(x) = G_0\{w(x)\}$.

1.2 A wealth of open questions

The intense research work devoted to distributions of type (2) has provided us with a wide collection of important results. However, many of these have been established for specific subclasses of (2). The most intensively studied instance is given by the skew-normal density which, in the case of $d = 1$, takes the form (1). Important results have been obtained also for other subclasses, especially when f_0 is Student's t density or Subbotin's density (also called exponential power distribution).

Much less is known in general terms, in the sense that there still is a relatively limited set of results which allows us to establish in advance, on the basis of qualitative properties of the components f_0, G_0, w of (2), what the formal properties of the resulting density function f will be. Results of this kind do exist, and Proposition 1 is the most prominent example, since it is both completely general and of paramount importance in associated distribution theory; several results on quadratic forms and even order moments follow from this property. Little is known about the distribution of non-even transformations. Among the limited results of the latter type, some general properties of odd moments of (2) have been presented by Umbach (2006, 2008). However, there are many other questions which arise quite naturally in connection with Lemma 1; the following is a non-exhaustive list.

- When $d = 1$, which assumptions on $G(x)$ ensure that the median of f is larger than 0? More generally, when can we say that the pth quantile of f is larger than the pth quantile of f_0? Obviously, 'larger' here may be replaced by 'smaller'.
- The even moments of f and those of f_0 coincide, because of (6). What can be said about the odd moments? For instance, is there an ordering of moments associated with some form of ordering of $G(x)$?
- If f_0 is unimodal, which are the additional assumptions on G_0 and w which ensure that f is still unimodal?
- When $d > 1$, a related but distinct question is whether high density regions of the type $C_u = \{x : f(x) > u\}$, for an arbitrary positive u, are convex.

Although the aim of the present paper is partly to tackle the above questions, at the same time we take a broader view, attempting to take a step forward in understanding the general properties of set of distributions (2). The latter target is the reason for the preliminary results of Sect. 2, which lead to a characterization result in Sect. 2.2 and provide the basis for the subsequent sections which deal with more specific results. In Sect. 3, we deal with the case $d = 1$ and tackle some of the questions listed above. Specifically, we obtain quite general results on stochastic ordering of skew-symmetric distributions with common base f_0, which imply orderings of quantiles and expected values of suitable transformations of the original variate. The final part of Sect. 3 concerns the uniqueness of the mode of density f. Section 4 deals with the general case of d, where various results are obtained. One of these is to establish the convexity of sets C_u for the more important subclass of the skew-elliptical family, provided the parent elliptical family enjoys a slightly more stringent property. We also examine the connection between the formulation of skew-elliptical densities of type (2) and those of Branco and Dey (2001) and prove the conjecture of Azzalini and Capitanio (2003) that the first formulation strictly includes the second one. Last, we give conditions for the log-concavity of skew-elliptical distributions not generated by the conditioning mechanism of Branco and Dey (2001).

One reviewer of this paper asked to examine the above sort of problems not only in connection with representations (2) and the corresponding one of Wang et al. (2004) but also with other types of related constructions. A case of special interest is based on a selection mechanism, leading to a form like (4) or similar. However, it is clear from (4) that, when the selection mechanism operates via a condition of type $(T < 0|X_0 = x)$, the ensuing distribution is of type (2), and we are already considering this situation. If the selection mechanism is modified, even simply to the form $(T < c|X_0 = x)$ for some arbitrary but fixed c, then the whole picture changes completely: the normalizing constant of (2) varies with c, f_0 and G, instead of being fixed at 2, stochastic representation (5) does not hold, at least in this form, and the property of perturbation invariance disappears. Given this radical modification of the context, its systematic exploration would correspondingly affect the present paper. Therefore, we do not pursue this direction, apart from a specific case discussed at the end of the paper.

2 Skew-symmetric densities with a common base

2.1 Preliminary facts

Clearly, f in (2) depends on G_0 and w only via perturbation function $G(x) = G_0\{w(x)\}$. The assumptions on G_0 and w in Lemma 1 ensure that

$$G(x) \geq 0, \qquad G(x) + G(-x) = 1, \qquad (x \in \mathbb{R}^d), \tag{7}$$

and, conversely, it is true that a function G satisfying these conditions ensures that

$$f(x) = 2\, f_0(x)\, G(x) \tag{8}$$

is a density function. So that (7)–(8) represent the formulation adopted by Wang et al. (2004) for their result essentially equivalent to Lemma 1.

Each of the formulations has its own advantages. As noted by Wang et al. (2004), the representation of $G(x)$ in the form $G(x) = G_0\{w(x)\}$ is not unique. In fact, given one such representation,

$$G(x) = G_*\{w_*(x)\}, \quad w_*(x) = G_*^{-1}[G_0\{w(x)\}]$$

is another one, for any strictly increasing distribution function G_* with even density function on $\mathbb{R}$.

Conversely, finding a function G fulfilling conditions (7) is immediate if we build it via expression $G(x) = G_0\{w(x)\}$; in fact, this is the usual way in the literature to select suitable G functions. Wang et al. (2004) also showed that the opposite fact holds: any function G satisfying (7) can be written in the form $G_0\{w(x)\}$, and this can be done in infinitely many ways. One choice of this representation which we find 'of minimal modification' is

$$\begin{aligned} G_0(t) &= \left(t + \tfrac{1}{2}\right) I_{(-1,1)}(2t) + I_{[1,+\infty)}(2t), \quad (t \in \mathbb{R}), \\ w(x) &= G(x) - \tfrac{1}{2}, \qquad\qquad\qquad\qquad\qquad\quad (x \in \mathbb{R}^d), \end{aligned} \tag{9}$$

where $I_A(x)$ denotes the indicator function of set A. More simply, this G_0 is the distribution function of a $U(-\frac{1}{2}, \frac{1}{2})$ variate.

Another important finding of Wang et al. (2004, Proposition 3) is that any positive density function f on $\mathbb{R}^d$ admits a representation of type (8), as indicated in their result, which we reproduce next, slightly modified as regards the arbitrariness of $G(x)$ outside the support of f_0. Here and in the following, we denote by $-A$ the set formed by reversing the sign of all elements of A, if A denotes a subset of a Euclidean space. If $A = -A$, we say that A is a symmetric set.

Proposition 2 *Let f be a density function with support $S \subseteq \mathbb{R}^d$. A representation of type (8) thus holds, with*

$$f_0(x) = \begin{cases} \frac{1}{2}\{f(x) + f(-x)\} & if \quad x \in S_0, \\ 0 & otherwise, \end{cases}$$
$$G(x) = \begin{cases} \dfrac{f(x)}{2 f_0(x)} & if \quad x \in S_0, \\ arbitrary & otherwise, \end{cases} \tag{10}$$

where $S_0 = (-S) \cup S$, and the arbitrary branch of G satisfies (7). In addition, f_0 is unique, and G is uniquely defined over S_0.

Let us now consider a density function with representation of type (8). We first introduce a property of cumulative distribution function F which is also of independent interest. We rewrite the first relation in (10) as

$$f(-x) = 2\, f_0(x) - f(x). \tag{11}$$

for any $x = (x_1, \ldots, x_d)$. If F_0 is the cumulative distribution function of f_0, then integration of (11) on $(-\infty, x_1] \times \cdots \times (-\infty, x_d]$ gives

$$\overline{F}(-x) = 2\, F_0(x) - F(x) \tag{12}$$

where $\overline{F}$ denotes the survival function, that is

$$\overline{F}(x) = \mathbb{P}\{X_1 \geq x_1, \ldots, X_d \geq x_d\}. \tag{13}$$

Equation (12) can be written as

$$\overline{F}(-x) + F(x) = F_0(x) + \overline{F_0}(-x)$$

and this in turn is equivalent to Proposition 1, as stated in Proposition 3 below.

2.2 A characterization

The five single statements composing the next proposition are known for the case $d = 1$, and some of them also for general d. The most important new fact is that they are equivalent, which therefore represents a characterization type of result.

More explicitly, while several papers have investigated implications of the assumption of a skew-symmetric representation, the result shown below states that some reverse implications also hold. The most notable of these is that, if perturbation invariance property (6) holds for all even $t(\cdot)$, this implies a skew-symmetric representation with common base f_0 for the underlying distributions.

Proposition 3 *Consider random variables $X = (X_1, \dots X_d)^\top$ and $Y = (Y_1, \dots Y_d)^\top$ with distribution functions F and H, and density functions f and h, respectively; denote by $\overline{F}$ and $\overline{H}$ the survival functions of F and H, defined as in (13). The following conditions then are equivalent:*

(a) *densities $f(x)$ and $h(x)$ admit a representation of type (8) with the same symmetric base density $f_0(x)$,*
(b) *$t(X) \stackrel{d}{=} t(Y)$, for any even q-dimensional function t on $\mathbb{R}^d$,*
(c) *$P(X \in A) = P(Y \in A)$, for any symmetric set $A \subset \mathbb{R}^d$,*
(d) *$F(x) + \overline{F}(-x) = H(x) + \overline{H}(-x)$,*
(e) *$f(x) + f(-x) = h(x) + h(-x)$,* (a.e.).

Proof (a)$\Rightarrow$(b) This follows from the perturbation invariance property of Proposition 1.

(b)$\Rightarrow$(c) Simply note that the indicator function of a symmetric set A is an even function.

(c)$\Rightarrow$(d) On setting

$$\begin{aligned}
A_+ &= \{s = (s_1, \dots, s_d) \in \mathbb{R}^d : s_j \le x_j, \forall j\},\\
A_- &= \{s = (s_1, \dots, s_d) \in \mathbb{R}^d : -s_j \le x_j, \forall j\} = -A_+,\\
A_\cup &= A_+ \cup A_-,\\
A_\cap &= A_+ \cap A_-,
\end{aligned}$$

both $A_\cup$, $A_\cap$ are symmetric sets; hence we obtain

$$\begin{aligned}
F(x) + \overline{F}(-x) &= P(X \in A_+) + P(X \in A_-),\\
&= P(X \in A_\cup) + P(X \in A_\cap).
\end{aligned}$$

(d)$\Rightarrow$(e) Taking the dth mixed derivative of (d), relationship (e) follows.

(e)$\Rightarrow$(a) This follows from the representation given in Proposition 2. $\square$

In the special case of $d = 1$, the above statements may be re-written in more directly interpretable expressions. Specifically, (12) leads to

$$1 - F(-x) = 2\,F_0(x) - F(x), \tag{14}$$

which will turn out to be useful later, and

$$F(x) - F(-x) = F_0(x) - F_0(-x).$$

In addition, when $d = 1$, conditions (c) and (d) in Proposition 3 may be replaced by the following more directly interpretable forms:

(c′) $|X| \stackrel{d}{=} |Y|$,
(d′) $F(x) - F(-x) = H(x) - H(-x)$,
the first of which appeared in Azzalini (1986) and the second is an immediate consequence.

3 Some results when $d = 1$

3.1 Stochastic ordering

For $d = 1$, we start by considering the problem of finding conditions under which two perturbations of same base density f_0 are stochastically ordered. Let us recall the standard concept of stochastic ordering of distributions functions: if F_1 and F_2 are univariate distributions functions and $F_1(s) \geq F_2(s)$ for all real s, then F_2 is stochastically larger than F_1. For background information on stochastic ordering, see Whitt (2006).

To exemplify the motivation in a basic case, it is natural to think that density functions $2\, f_0(x)\, G_0(\alpha x)$, under the conditions stated earlier, correspond to distributions "increasing with α", but this ordering seems to have been examined only for special cases, such as skew-normal distribution.

To proceed, we need to introduce a type of ordering on the set of functions which satisfy (7). This is much the same as the concept of peakedness order introduced by Birnbaum (1948) to compare the variability of distributions about some given points. However, the two concepts are distinct in the sense that Birnbaum's peakedness only applies to distribution functions, whereas the functions that we consider do not need to be probability distributions, although they must satisfy symmetry condition (7), not required for peakedness.

Definition 1 If G_1 and G_2 satisfy (7), we say that G_2 is greater than G_1 on the right, denoted $G_2 \geq_{GR} G_1$, if $G_2(x) \geq G_1(x)$ for all $x > 0$.

Of course it is equivalent to require that $G_2(x) \leq G_1(x)$ for all $x < 0$. Another equivalent condition is that

$$G_2(s) - G_2(r) \geq G_1(s) - G_1(r), \qquad (r < 0 < s).$$

If we now consider a fixed symmetric 'base' density f_0 and the perturbed distribution functions associated with G_1 and G_2, i.e.

$$F_k(x) = \int_{-\infty}^{x} 2\, f_0(u)\, G_k(u)\, \mathrm{d}u, \qquad (k = 1, 2), \tag{15}$$

ordering $G_2 \geq_{GR} G_1$ immediately implies stochastic ordering of F_1 and F_2. To see this, consider first $s \leq 0$; then $G_1(x) \geq G_2(x)$ for all $x \leq s$, and this clearly implies $F_1(s) \geq F_2(s)$. If $s > 0$, the same conclusion holds using (14) with $x = -s$. We have thus reached the following conclusion.

Proposition 4 *If G_1 and G_2 satisfy condition* (7)*, and $G_2 \geq_{GR} G_1$, then distribution functions* (15) *satisfy*

$$F_1(x) \geq F_2(x), \qquad (x \in \mathbb{R}). \tag{16}$$

In addition, if G_1 and G_2 are continuous and not identical, (16) *holds strictly for some x.*

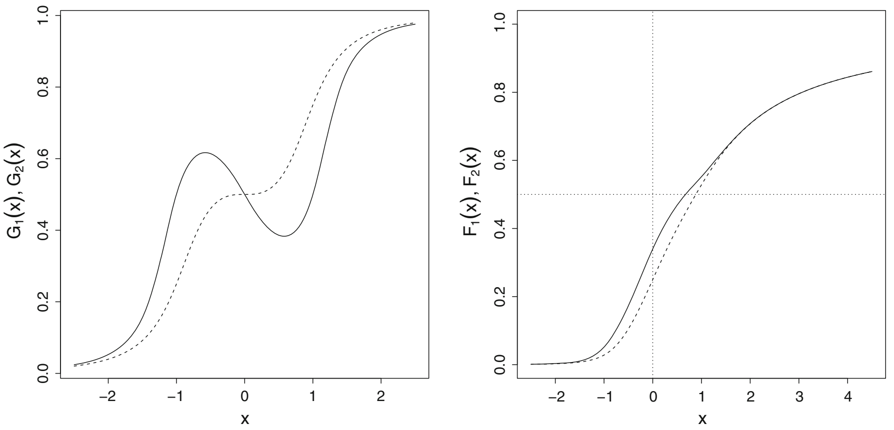

Fig. 1 Cauchy density function $f_0(x)$ perturbed by G_0 equal to Cauchy distribution, choosing $w_1(x) = x^3 - x$, and $w_2(x) = x^3$. *Left panel* G_1 (*continuous line*) and G_2 (*dashed line*); *right panel* F_1 (*continuous line*) and F_2 (*dashed line*)

An important special case of Proposition 4 arises when $G_1 \equiv \frac{1}{2}$, so that f_1 coincides with base density f_0. If $G_2(x) \geq \frac{1}{2}$ for all $x > 0$, then perturbed distribution F_2 is stochastically larger than F_0, the distribution function of base density f_0. The reverse happens when $\frac{1}{2} \geq_{GR} G_2$.

Another important special case occurs when $G_(x) = I_{[0,\infty)}(x)$ so that $G_2(x) \geq_{GR} G_1(x)$ for any $G_1(x)$ of type (7), and consequently ordering $F_1(x) \geq F_2(x)$ holds for all x.

Since G_0 is a monotonically increasing function, it may be easier to check the ordering of G_1 and G_2 via the ordering of the corresponding $w(x)$.

Proposition 5 *If $G_1 = G_0(w_1(x))$ and $G_2 = G_0(w_2(x))$, where G_0 is as in Lemma 1 and w_1 and w_2 are odd functions so that $w_2(x) \geq w_1(x)$ for all $x > 0$, then $G_2 \geq_{GR} G_1$ and (16) holds. In addition, if w_1 and w_2 are continuous and not identical, (16) holds strictly for some x.*

Figure 1 illustrates order $G_2 \geq_{GR} G_1$ and the stochastic order between the corresponding distribution functions $F_1(x) \geq F_2(x)$, as stated by Proposition 4. Here, f_0 is the Cauchy density, G_0 is the Cauchy distribution function, and two forms of $w(x)$ are considered, namely, $w_1(x) = x^3 - x$, $w_2(x) = x^3$. The two perturbation functions $G_1(x) = G_0(w_1(x))$ and $G_2(x) = G_0(w_2(x))$ are plotted in the left panel; the corresponding distribution functions $F_1(x)$ and $F_2(x)$ are shown in the right panel.

Stochastic ordering of the F_k translates immediately into a set of implications about the ordering of moments and quantiles of those F_k's. Specifically, if X_k is a random variable with distribution function F_k, for $k = 1, 2$, then the following statements hold.

– If $Q_k(p)$ denotes the pth quantile of X_k for any $0 < p < 1$, then

$$Q_1(p) \leq Q_2(p).$$

In those cases when (16) holds strictly for some x, the above inequality is strict for some p.

– For any non-decreasing function t such that expectations exist,

$$\mathbb{E}\{t(X_1)\} \le \mathbb{E}\{t(X_2)\} \tag{17}$$

and the inequality is strict if t is increasing and (16) is strict for some x.

A further specialized case occurs when $t(x) = x^{2n-1}$ in (17), for $n = 1, 2, \ldots$, which corresponds to the set of odd moments. Therefore, (17) extends one of the results of Umbach (2006) stating that

$$\mathbb{E}\{X_0^{2n-1}\} \le \mathbb{E}\{X_1^{2n-1}\} \le \mathbb{E}\{X_*^{2n-1}\}$$

where X_0 has density f_0 and X_* has density $2f_0$ on the positive axis, which corresponds to $G(x) = I_{[0,\infty)}(x)$ in (8), and the density of X_1 corresponds to a G_1 which is a distribution function.

It can be noted that, if $G_2 \ge_{GR} G_1 \ge_{GR} G_+ \equiv \frac{1}{2}$, then the variances of the corresponding variables X_k decrease with respect to $\ge_{GR}$, so that $\mathrm{var}\{X_2\} \le \mathrm{var}\{X_1\} \le \mathrm{var}\{X_0\}$; the reverse holds if $G_+ \ge_{GR} G_1 \ge_{GR} G_2$.

A simple but popular setting to which Proposition 5 applies is when $w(x) = \alpha x$, for some real α, leading to the following immediate implication.

Proposition 6 *If f_0 and G_0 are as in Lemma 1, then the set of densities*

$$f(x; \alpha) = 2\, f_0(x)\, G_0(\alpha x) \tag{18}$$

indexed by the real parameter α are associated with distribution functions which are stochastically ordered with α.

Note that, when α in (18) is positive, it has a direct interpretation as an inverse scale parameter for G_0, but it acts as a shape parameter for $f(x)$.

Another case of interest is given by

$$w(x) = \alpha x \sqrt{\frac{\nu + 1}{\nu + x^2}},$$

which occurs in connection with the skew Student's t distribution with ν degrees of freedom, studied by Azzalini and Capitanio (2003) and others, where f_0 and G_0 are of Student's t type with ν and $\nu + 1$ degrees of freedom, respectively. Because of Proposition 5, the distribution functions associated with (2) with this choice of $w(x)$ are stochastically ordered with respect to α, whether f_0 and G_0 correspond to a Student's t distribution or not.

3.2 On the uniqueness of the mode

To examine the problem of the uniqueness of the mode of f when $d = 1$, it is equivalent and more convenient to study $\log f$. If $f_0'(x)$ and $g(x) = G'(x)$ exist, then

$$\begin{aligned} h(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \log f(x) \\ &= \frac{f_0'(x)}{f_0(x)} + \frac{g(x)}{G(x)} \\ &= -h_0(x) + h_g(x), \text{ say.} \end{aligned}$$

The modes of f are a subset of the solutions of the equation

$$h_0(x) = h_g(x), \tag{19}$$

or they are on the extremes of the support. Since at least one mode always exists, we look for conditions which rule out the existence of other modes.

For the rest of this subsection, we assume that $G(x)$ is a strictly monotone function satisfying (7). Without loss of generality, we deal with the case that G is monotonically increasing; for decreasing functions, dual conclusions hold.

In the most common cases, f_0 is unimodal at 0, and hence non-decreasing for $x \le 0$. Therefore, product $f_0(x)\,G(x)$ is increasing, and no negative mode can exist. The same conclusion holds if f_0 is increasing and $G(x)$ is non-decreasing for $x \le 0$.

To ensure that there is at most one positive mode, some additional conditions are required. For simplicity of argument, we assume that f_0 and G have continuous derivatives everywhere on support S_0 of f_0; this means that we are concerned with the uniqueness of the solution of (19). A sufficient set of conditions for this uniqueness is that $h_0(x)$ is increasing and $g(x)$ is decreasing. These requirements imply that $h_g(0) > 0$ and h_g is decreasing, so that $0 = h_0(0) < h_g(0)$ and the two functions can cross at most once for $x > 0$. When S_0 is unbounded, a solution of (19) always exists, since $g \to 0$ and $h_g \to 0$ as $x \to \infty$. If S_0 is bounded, (19) may have no solution; in this case, $f(x)$ is increasing for all x and its mode occurs at the supremum of S_0. We summarize this discussion in the following statement.

Proposition 7 *If $G(x)$ in (8) is a increasing function and $f_0(x)$ is unimodal at* 0*, then no negative mode exists. If we assume that f_0 and G have continuous derivatives everywhere on support f_0, $G(x)$ is concave for $x > 0$ and $f_0(x)$ is log-concave, where at least one of these properties holds in a strict sense, then there is a unique positive mode of $f(x)$. If $G(x)$ is decreasing, similar statements hold, with reversed sign of the mode; the uniqueness of the negative mode requires that $G(x)$ is convex for $x < 0$.*

It should be recalled that the property of log-concavity of a univariate density function is equivalent to strong unimodality; see for instance Section 1.4 of Dharmadhikari and Joag-dev (1988).

To check the above conditions in specific instances, it is convenient to work with functions h_0 and g', if the latter exists. In the case of increasing $w(x)$, the uniqueness

Table 1 Some commonly used densities f_0 and associated components

Distribution	$f_0(x)$	$h_0(x)$	$h_0'(x)$
Standard normal	$\phi(x)$	x	1
Logistic	$\dfrac{e^x}{(1+e^x)^2}$	$\dfrac{e^x-1}{e^x+1}$	$\dfrac{2\,e^x}{(e^x+1)^2}$
Subbotin	$c_\nu \exp\left(-\dfrac{\lvert x\rvert^\nu}{\nu}\right)$	$\operatorname{sgn}(x)\,\lvert x\rvert^{\nu-1}$	$\operatorname{sgn}(x)\,(\nu-1)\lvert x\rvert^{\nu-2}$
Student's t_ν	$c_\nu\left(1+\dfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	$\dfrac{\nu+1}{\nu}\;\dfrac{x}{1+x^2/\nu}$	$\dfrac{(\nu+1)(\nu-x^2)}{(\nu+x^2)^2}$

of the mode is ensured if $g'(x) < 0$ for $x > 0$ and $h_0(x)$ is an increasing positive function. In the linear case $w(x) = \alpha\, x$, the log-concavity of $f_0(x)$ and unimodality of $g_0(x)$ at 0 suffice to ensure the unimodality of $f(x)$.

Table 1 lists some of the more commonly employed density functions f_0 and their associated functions h_0 and h_0'. For the first two distributions of Table 1, and for Subbotin's distribution when $\nu > 1$, h_0 is increasing. If we combine one of these three choices of f_0 with the distribution function of a symmetric density having a unique mode at 0, then the uniqueness of the mode of $f(x)$ follows. Clearly, the condition of unimodality of $g(x)$ holds if g_0 is unimodal at 0 and $w(x) = \alpha x$. The criterion of Proposition 7 does not apply for Student's distribution, since $h_0(x)$ is increasing only in the interval $(-\sqrt{\nu}, \sqrt{\nu})$. Hence, a second intersection with h_g cannot be ruled out, even if $g(x)$ is decreasing for all $x > 0$. However, for skew-t distribution, unimodality was established in the multivariate case by Capitanio (2008) and Jamalizadeh and Balakrishnan (2010), and it also follows as a corollary of a stronger result to be presented in Sect. 4.

The requirement of the differentiability of f_0 and G in Proposition 7 rules out a limited number of practically relevant cases. This is why we did not dwell on a specific discussion of less regular cases. One of the very few relevant distributions which are excluded occurs when f_0 is the Laplace density function. However, this case is included in the discussion of multivariate Subbotin distribution, developed in Sect. 4.3, when $\nu = 1$ and $d = 1$.

Although Proposition 7 only gives a set of sufficient conditions for unimodality, the condition that $g(x)$ is decreasing for $x > 0$ cannot be avoided completely. In other words, when f is represented in the form of (2), the sole condition of increasing $w(x)$ is not sufficient for unimodality. This fact is demonstrated by the simple case with $f_0 = \phi$, $G_0 = \Phi$, $w(x) = x^3$, key features of which are shown in Fig. 2. Since $w'(0) = 0$, then $g(0) = 0$; hence (19) has a solution in 0, but the left panel of Fig. 2 shows that there are two more intersections of h_0 and h_g for $x > 0$, one corresponding to an anti-mode and one to a second mode of $f(x)$, as shown in the right-hand panel.

This case falls under the setting examined by Ma and Genton (2004), who showed that, for $f_0 = \phi$, $G_0 = \Phi$, $w(x) = \alpha\, x + \beta\, x^3$, there are at most two modes. Some additional conditions may ensure unimodality: one such set of conditions is $\alpha, \beta > 0$ and $\alpha^3 > 6\beta$. To prove that they imply the unimodality of f, consider

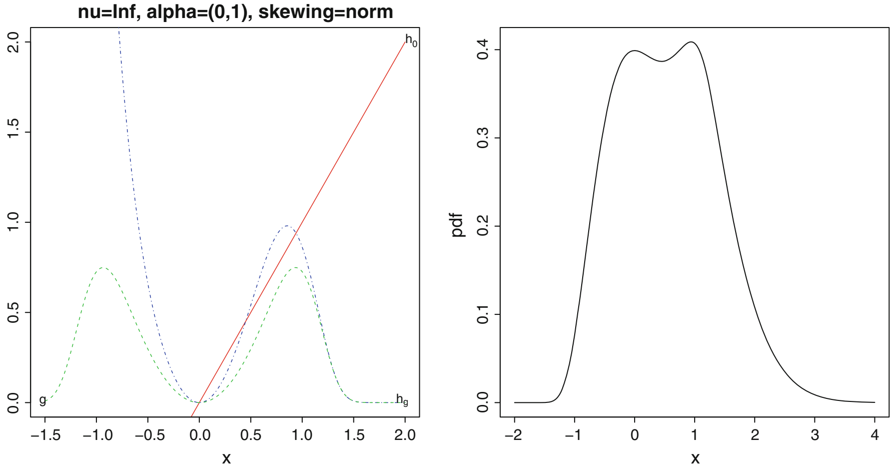

Fig. 2 Case with $f_0 = \phi, G_0 = \Phi, w(x) = x^3$. *Left* functions h_0 (*continuous line*), g (*dashed line*), h_g (*dot-dashed line*); *right* function $f(x)$

$$\frac{\mathrm{d}^2 \log \Phi(w(x))}{\mathrm{d}x^2} = -\frac{\phi(w(x))}{\Phi(w(x))^2} \times \left\{\Phi(w(x))[(\beta x^3 + \alpha x)(3\beta x^2 + \alpha)^2 - 6\beta\, x] + \phi(w(x))(3\beta x^2 + \alpha)^2\right\},$$

whose terms in curly brackets, except $-6\beta\, x$, are all positive for $x \geq 0$. Since $\alpha^3 > 6\beta$, then $(\alpha^3 - 6\beta)x$ is positive, so that this derivative is negative and $G_0(w(x))$ is log-concave for $x \geq 0$. For $x < 0$, we use this other argument: since G_0 is increasing and log-concave and $w(x)$ is concave in subset $x < 0$, composition $G_0(w(x))$ is log-concave in the subset $x < 0$; see Proposition 8 (iii) below. Since the above second derivative is continuous everywhere, $G_0(w(x))$ is log-concave everywhere.

4 Quasi-concave and unimodal densities in *d* dimensions

A real-valued function f defined on subset S of $\mathbb{R}^d$ is said to be quasi-concave if the sets of the form $C_u = \{x : f(x) \geq u\}$ are convex for all positive u. When $d = 1$, quasi-concavity coincides with the uniqueness of the maximum, provided a pole is regarded as a maximum point, but for $d > 1$ the two concepts separate. This motivates the following digression about concavity and related concepts, to develop some tools which will be used later on for our main target.

4.1 Concavity, quasi-concavity and unimodality

Let us first recall some standard concepts available, for instance, in Chapter 16 of Marshall and Olkin (1979). A real function f defined on a convex subset S of $\mathbb{R}^d$ is said to be concave if, for every x and $y \in S$ and $\theta \in (0, 1)$, we have

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta) f(y);$$

in this case, $-f$ is a convex function. A function f is said to be log-concave if $\log f$ is concave; that is, for every x and $y \in S$ and $\theta \in (0, 1)$ we have

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta}.$$

The terms 'strictly concave' and 'strictly log-concave' apply if the above inequalities hold in a strict sense for all $x \neq y$ and all θ.

Concave and log-concave functions defined on an open set are continuous. A twice differentiable function is also concave (strictly concave) if and only if its Hessian matrix is negative semi-definite (negative definite) everywhere on S.

The next proposition provides the concave and log-concave extension of classical composition properties for convex functions like that of statement (i) which can be found, for example, in Marshall and Olkin (1979, p. 451) together with its proof. The proofs of the other statements are very similar.

Proposition 8 *Let h be a real function defined on a convex set S, a subset of $\mathbb{R}^d$, and H a monotone real function defined on a convex subset of $\mathbb{R}$, so that composition $H(h)$ is defined on S. Then the following properties hold.*

(i) *If h is convex and H non-decreasing and convex, then $H(h)$ is convex. In addition, $H(h)$ is strictly convex if H is strictly convex, or if h is strictly convex and H strictly monotone.*
(ii) *If h is convex and H non-increasing and log-concave, then $H(h)$ is log-concave. Also, $H(h)$ is strictly log-concave if H is strictly log-concave, or if h is strictly convex and H strictly monotone. The same statements hold if the term log-concave is replaced by concave throughout.*
(iii) *If h is concave and H non-decreasing and log-concave, then $H(h)$ is log-concave. Also, $H(h)$ is strictly log-concave if H is strictly log-concave, or if h is strictly concave and H strictly monotone. The same statements hold if the term log-concave is replaced by concave throughout.*

We defined quasi-concavity by requiring convexity of all sets C_u. An equivalent condition is that, for every x and $y \in S \subseteq \mathbb{R}^d$ and $\theta \in (0, 1)$, we have

$$f(\theta x + (1-\theta)y) \geq \min\{f(x), f(y)\}.$$

Obviously, a function which is concave or log-concave is also quasi-concave. Similarly, both strict concavity and strict log-concavity imply strict quasi-concavity.

We now apply the above concepts to the case in which f represents a probability density function on a set $S \subseteq \mathbb{R}^d$. The concept of unimodality has a friendly formal definition in the univariate case, given for instance by Dharmadhikari and Joag-dev (1988, p. 2), but this has no direct equivalent in the multivariate case. Informally, we say that the term 'mode of a density' refers to a point at which the density takes a maximum value, either globally or locally. While a boring formal definition which allows for the non-uniqueness of the density function could be given, it is not really necessary for the main aims of the present paper, since the density functions with which we are concerned are so regular that their modes are either points of (local) maxima or poles.

The set of the modes of a quasi-concave density is a convex set. Also, if f is strictly quasi-concave, then the mode is unique. When this is so, we say that density f is *unimodal*, and that f is *c-unimodal* if the set of its modes is a convex set. If X is a random variable with density function f which is unimodal, we say that X is unimodal, with a slightly extended terminology. The same convention is adopted for log-concavity, quasi-concavity and other properties.

Another important concept is s-concavity, which helps to make the concept of quasi-concavity more tractable. A systematic discussion of s-concavity is given in Dharmadhikari and Joag-dev (1988); see specifically their Section 3.3, of which we now recall the main points. Given a real number $s \neq 0$, a density is said to be s-concave on S if

$$f(\theta x + (1-\theta)y) \geq \{\theta f(x)^s + (1-\theta) f(y)^s\}^{1/s},$$

for all $x,\ y \in S$ and all $\theta \in (0, 1)$.

Clearly, concavity corresponds to $s = 1$. A density f is s-concave with $s < 0$ if and only if f^s is convex; similarly, a density f is s-concave with $s > 0$ if and only if f^s is concave. If a function which is quasi-concave is said to be $(-\infty)$-concave, and a function which is log-concave is said to be 0-concave, then the class of sets of s-concave functions increases as s decreases; in other words, if f is s-concave, then it is r-concave for any $r < s$. Last, note that it is easy to adapt Proposition 8 to s-concave functions.

Closure with respect to marginalization of s-concave densities depends on the value of s and on the dimensions of the spaces, as indicated by the next proposition.

Proposition 9 *Let f be an s-concave density on a convex set S in $\mathbb{R}^{d+m}$, and f_d the marginal density of f on a d-dimensional subspace. If $s \geq -1/m$, then f_d is s_m-concave on the projection of the support of f, where $s_m = s/(1+ms)$, with the convention that, if $s = -1/m$, then $s_m = -\infty$. In addition, the marginal densities are strictly s_m-concave provided f is strictly s-concave or the set S is strictly convex.*

The first statement is essentially Theorem 3.21 of Dharmadhikari and Joag-dev (1988); the second statement follows by applying their argument for proving the first part also to the new context. Note that the above result includes the fact that the class of log-concave densities is closed with respect to marginalization.

4.2 Skew-elliptical distributions generated by conditioning

A d-dimensional random variable U is said to have elliptical density, with density generator function $\tilde{f}$, if its density f_U is of the form

$$f_U(y) = k_{d,\Omega}\, \tilde{f}(y^\top \Omega^{-1} y), \tag{20}$$

where Ω is a d-dimensional positive definite matrix, function $\tilde{f} : (0, +\infty) \to \mathbb{R}^+$ is such that $x^{d/2-1}\tilde{f}(x)$ has a finite integral on $(0, +\infty)$, and $k_{d,\Omega}$ is a suitable constant which depends on d and on Ω only via $\det(\Omega)$. In this case, we use the notation $U \sim \mathcal{E}_d(0, \Omega, \tilde{f})$.

Note that an elliptical density f is c-unimodal if and only if its density generator is non-increasing, and it is unimodal if and only if its density generator is decreasing. Then it turns out that f is c-unimodal if and only if it is quasi-concave, and unimodal if and only if it is strictly quasi-concave.

An initial formulation of skew-elliptical distribution was examined by Azzalini and Capitanio (1999) and was of type (2) with f_0 of elliptical class and $w(x)$ linear. Another formulation of skew-elliptical distribution was put forward by Branco and Dey (2001), the key points of which are now recalled. Consider a $(d+1)$-dimensional random variable

$$U = \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} \sim \mathcal{E}_{d+1}(0, \Omega_+, \tilde{f}), \qquad \text{where} \quad \Omega_+ = \begin{pmatrix} 1 & \delta^\top \\ \delta & \Omega \end{pmatrix} > 0, \tag{21}$$

and U_0 and U_1 have dimensions 1 and d, respectively. For our aims, there is no loss of generality in assuming that the diagonal elements of Ω_+ are all 1. Then a random variable $Z = (U_1 | U_0 > 0)$ is said to have a skew-elliptical distribution, and its density function at $u_1 \in \mathbb{R}^d$ is

$$f_Z(u_1) = 2 \int_0^{+\infty} k_{d+1,\Omega_+} \tilde{f}(u^\top \Omega_+^{-1} u)\, \mathrm{d}u_0 \tag{22}$$

where $u^\top = (u_0, u_1^\top)$. This construction arises as an extension of one of the mechanisms for generating skew-normal distribution to the case of elliptical densities, but study of connections with other densities of type (2) was not one of the aims of Branco and Dey (2001).

Consequently, one question examined by Azzalini and Capitanio (2003) was whether all distributions of type (22) are of type (2), with the requirement that f_0 is the density of an elliptical d-dimensional distribution. The conjecture was proven for a set of important cases, notably multivariate skew-normal and skew-t distributions, among others, but a general statement could not be reached. However, this general conclusion is quite simple to reach with representation (8), and recalling that Branco and Dey (2001) proved that (22) amy be written as

$$f_Z(y) = 2\; f_0(y)\, F_y(\alpha^\top y), \qquad\qquad (y \in \mathbb{R}^d), \tag{23}$$

where f_0 is the density of an elliptical d-dimensional distribution, and F_y is a cumulative distribution function of a symmetric univariate distribution, which depends on y only through $y^\top \Omega^{-1} y$. Since $F_y = F_{-y}$, it is immediate that $G(y) = F_y(\alpha^\top y)$ satisfies (7). Hence, (23) allows a representation of type (8) and, via (9), also of type (2).

Proposition 10 *Assume that random variable U in (21) is c-unimodal. If $\tilde{f}$ is log-concave, then the elliptical densities of U and U_1 and the skew-elliptical density of Z*

are log-concave. They are also strictly log-concave if U is unimodal or $\tilde{f}$ is strictly log-concave, or the support of $\tilde{f}$ is bounded.

Proof Function $h(u) = u^\top \Omega_+^{-1} u$ is strictly convex. Since U is c-unimodal, $\tilde{f}$ is non-increasing. It is also log-concave; therefore, $\tilde{f}(u^\top \Omega_+^{-1} u)$ is log-concave accordingly to Proposition 8 (ii). Both U and $(U|U_0 > 0)$, therefore, have log-concave densities. Since the marginals of a log-concave density are log-concave, the log-concavity of U_1 and Z holds from (22). Now, if U is unimodal, $\tilde{f}$ is decreasing and $\tilde{f}(u^\top \Omega_+^{-1} u)$ is strictly log-concave from Proposition 8 (i). If $\tilde{f}$ is strictly log-concave, then $\tilde{f}(u^\top \Omega_+^{-1} u)$ is strictly log-concave. Last, if the support of $\tilde{f}$ is bounded, then the support of U is strictly convex and, from Proposition 8 (i), also in this case $\tilde{f}(u^\top \Omega_+^{-1} u)$ is strictly log-concave. So, in all three cases, the strict log-concavity of U_1 and Z holds, accordingly to the final part of Proposition 9. □

Proposition 10 is a special case of the more general result which follows, but we keep it separate both because of the special role of log-concavity and because this arrangement allows more compact exposition of the combined discussion.

Proposition 11 *Assume that the random variable U in (21) is c-unimodal. If $\tilde{f}$ is s-concave, with $s \geq -1$, then U has s-concave density, whereas the elliptical density of U_1 and the skew-elliptical density of Z are s_1-concave, with $s_1 = s/(1+s)$. All conclusions also hold strictly if U is unimodal or $\tilde{f}$ is strictly s-concave, or if the support of $\tilde{f}$ is bounded.*

Proof Function $h(u) = u^\top \Omega_+^{-1} u$ is strictly convex. Since U is c-unimodal, $\tilde{f}$ is non-increasing and also s-concave. We now examine the properties of concavity separating cases $s < 0$ and $s > 0$; the case $s = 0$, which corresponds to log-concavity, has already been handled in Proposition 10. If $s < 0$, then $\tilde{f}^s$ is non-decreasing and convex. So $\tilde{f}^s(u^\top \Omega_+^{-1} u) = \{\tilde{f}(u^\top \Omega_+^{-1} u)\}^s$ is convex from Proposition 8 (i) and $\tilde{f}(u^\top \Omega_+^{-1} u)$ is s-concave. Instead, if $s > 0$ then $\tilde{f}^s$ is non-increasing and concave. So $\tilde{f}^s(u^\top \Omega_+^{-1} u) = \{\tilde{f}(u^\top \Omega_+^{-1} u)\}^s$ is concave from Proposition 8 (ii) and $\tilde{f}(u^\top \Omega_+^{-1} u)$ is s-concave. This means that both U and $(U|U_0 > 0)$ have s-concave densities. Now, the claim about the densities of U_1 and Z follows from Proposition 9 according to (22). The final statement follows the same type of argument used in the proof of Proposition 10. □

Note that, in the special case of a concave density generator, the support is bounded, and both the marginal density on $\mathbb{R}^d$ and the skew-symmetric density of Z are not necessarily concave. However, using Proposition 11 with $s = 1$, the strict $(1/2)$-concavity of their densities follows, and this fact implies strictly log-concavity.

The results of Propositions 10 and 11 allow us to handle several classes of distributions, of which we now describe the more noteworthy cases.

An important specific instance is the multivariate skew-normal density which may be represented by a conditioning method. For an expression of multivariate skew-normal density, see, for instance, (16) of Azzalini (2005). Since the density generator of the normal family, $\tilde{f}(x) = \exp(-x/2)$, is decreasing and log-concave, from Proposition 10 we obtain the log-concavity of the skew-normal family. However, this conclusion is a special case of a more general result on log-concavity of the SUN distribution obtained by Jamalizadeh and Balakrishnan (2010); see their Theorem 1.

The $(d+1)$-dimensional Pearson type II distributions for which $\tilde{f}(x) = (1-x)^\nu$, where $x \in (0, 1)$ and $\nu \geq 0$, satisfies the conditions of Proposition 11, being non-increasing and ν^{-1}-concave on a bounded support. So the skew-elliptical d-dimensional density is strictly $(\nu+1)^{-1}$-concave and therefore strictly log-concave. The density function of the skew-type II density function is given by (22) of Azzalini and Capitanio (2003).

Proposition 11 also holds for Pearson type VII distributions and in particular for the Student's distribution. In this case, the density generator is given by

$$\tilde{f}(x) = (1 + x/\nu)^{-M} \qquad (\nu > 0,\ (d+1)/2 < M). \tag{24}$$

and $M = (d+\nu+1)/2$ for Student's density. This generator is decreasing and s-concave with $s = -1/M$; $\tilde{f}(x)^{-1/M}$ is in fact convex. Since $s \geq -1$, then Proposition 11 applies and the skew-t is s_1-concave with $s_1 = -1/(M-1)$ and $s_1 = -2/(d+\nu-1)$ in Student's case. These densities are not log-concave, but they are still strictly quasi-concave. Hence, unimodality follows. For expressions of multivariate skew-type VII and skew-t density, see (21) and (26) of Azzalini and Capitanio (2003), respectively.

The above results establish not only the unimodality of the more appealing subset of the skew-elliptical family of distributions, i. e. those of type (22), but also the much stronger conclusion of quasi-concavity of these densities. It is intrinsic to the nature of skew-elliptical densities that their highest density regions are not of elliptical shape, but it is reassuring that they maintain qualitatively similar behaviour, in the sense that the convexity of these regions, C_u in our notation, holds as long as the parent $(d+1)$-dimensional elliptical density enjoys a qualitatively similar property but in a somewhat stronger variant, i. e. s-concavity with $s \geq -1$.

Note that there is no hope of extending Proposition 10 to quasi-concave densities, since a skew-symmetric generated by conditioning a quasi-concave density is not necessarily quasi-concave, as demonstrated by the following construction.

Example Consider $U = (U_0, U_1)^\top \sim \mathcal{E}_2(0, \Omega_+, \tilde{f})$, where

$$\tilde{f} = I_{(0,1)} + I_{(0,4^2)} \quad \text{and} \quad \Omega_+ = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix},$$

so that its density function is

$$f_U(x, y) = k\{I_{S_1}(x, y) + I_{S_4}(x, y)\}$$

where $S_j = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - xy \leq 3j^2/4\}$, $j = 1, 4$, and k is the normalizing constant given by $k = 1/(A_1 + A_4) \approx 0.0216$ where $A_j = \pi\sqrt{3}j^2/2$. Then both U_0, and U_1 have common support $[-4, 4]$ and density functions

$$f_{U_0}(x) = f_{U_1}(x) = k\left(\sqrt{3(1-x^2)}\, I_{(-1,1)}(x) + \sqrt{3(16-x^2)}\right).$$

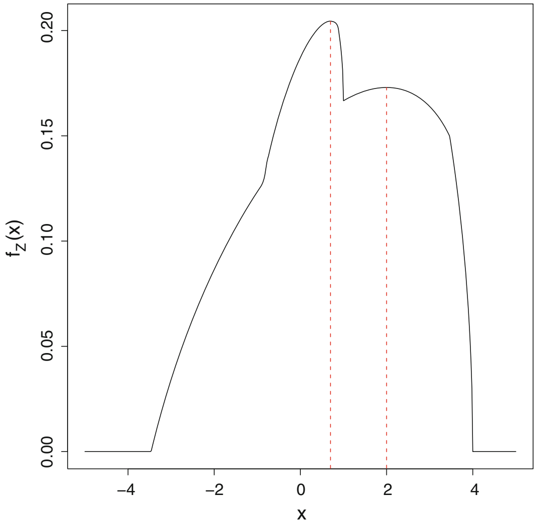

Fig. 3 Density function $f_Z(x)$, exhibiting lack of quasi-concavity, obtained by conditioning of a bivariate elliptical quasi-concave distribution

Because of (20) and (22), the density of $Z = (U_1|U_0 > 0)$ is given by

$$f_Z(y) = k\{f_1(y) + f_4(y)\}$$

where

$$f_j(y) = 2\int_0^{+\infty} I_{S_j}(x, y)\,\mathrm{d}x = \begin{cases} y + \sqrt{3(j^2 - y^2)} & if \quad -\sqrt{3}j/2 \le y \le \sqrt{3}j/2, \\ 2\sqrt{3(j^2 - y^2)} & if \quad \sqrt{3}j/2 \le y \le j, \\ 0 & otherwise, \end{cases}$$

for $j = 1, 4$, and it is shown in Fig. 3. The global maximum of f_Z occurs when $k(2y + \sqrt{3(16 - y^2)} + \sqrt{3(1 - x^2)})$ takes its maximum value, that is, at $y \approx 0.699$. When $y > 1$, $f_Z = kf_4$, and there is another local maximum at $y = 2$. Therefore, f_Z is not unimodal.

To conclude, although the density of U is quasi-concave, the skew-elliptical variable Z generated by conditioning is not.

4.3 Log-concavity of other families of distributions

There are several other families of distributions which belong to the area of interest of the stream of literature described at the beginning of this paper, but they are not included in the conditioning mechanism of elliptical distributions considered in Sect. 4.2. This section deals with the log-concavity of some of these other families, making use of the following immediate implication of Proposition 8.

Corollary 1 *If q_0 is a log-concave function defined on a convex set $S \subseteq \mathbb{R}^d$, and H and h are as in Proposition 8, either (ii) or (iii), then*

$$q(x) = q_0(x)\, H\{h(x)\}, \qquad (x \in S), \tag{25}$$

is log-concave on S.

Example The density function on the real line introduced by Subbotin (1923) has been variously denoted by subsequent authors as exponential power distribution, generalized error distribution and normal distribution of order ν. Its multivariate version is

$$f_\nu(x) = c_\nu \det(C)^{1/2} \exp\left(-\frac{(x^\top Cx)^{\nu/2}}{\nu}\right), \qquad (x \in \mathbb{R}^d),$$

where C is a symmetric positive definite matrix, ν a positive parameter and c_ν a normalization constant. For $\nu = 2$ and $\nu = 1$, f_ν lends the multivariate normal and multivariate Laplace densities, respectively.

We first need to show that f_ν is log-concave if $\nu \geq 1$. Consider $h(x) = (x^\top Cx)^{1/2}$, the Hessian matrix of which is

$$\frac{\partial^2 h(x)}{\partial x\, \partial x^\top} = h(x)^{-3}\left(x^\top CxC - Cxx^\top C\right) = h(x)^{-3} M, \text{ say.}$$

To show that this matrix is positive semi-definite, it is sufficient to prove this fact for matrix M, since $h(x) \geq 0$. For any $u \in \mathbb{R}^d$, write

$$u^\top Mu = (x^\top Cx)(u^\top Cu) - (u^\top Cx)(x^\top Cu) = \|\tilde{u}\|^2\, \|\tilde{x}\|^2 - (\tilde{u}^\top \tilde{x})^2$$

where $\tilde{u} = C^{1/2}u$ and $\tilde{x} = C^{1/2}x$ for any square root $C^{1/2}$ of C and, from the Cauchy–Schwarz inequality, we conclude that $u^\top Mu \geq 0$. Then h is convex. Next, write

$$-\log f_\nu(x) = \text{constant} + h(x)^\nu/\nu$$

and observe that, since t^ν is strictly convex for $t \geq 0$, $-\log f_\nu$ is convex for $\nu \geq 1$ and strictly convex for $\nu > 1$ from Proposition 8 (i). Hence f_ν is log-concave for $\nu \geq 1$ and strictly log-concave for $\nu > 1$.

We now introduce a skewed version of f_ν of type (2). If we aim at obtaining a density which fulfils the requirements of both Lemma 1 and Corollary 1, then $H = G_0$ is non-decreasing, while function $h = w$ must be odd and concave; hence, it must be linear. We then focus on the density function

$$f(x) = 2\, f_\nu(x)\, G_0(\alpha^\top x), \qquad (x \in \mathbb{R}^d), \tag{26}$$

where G_0 is a distribution function on $\mathbb{R}$, symmetric around 0.

Among the many options for G_0, a natural choice is to make G_0 equal to the distribution function of f_ν in the scalar case, which is

$$G_0(t) = \frac{1}{2}\left(1 + \operatorname{sgn}(t)\frac{\gamma(|t|^\nu/\nu, 1/\nu)}{\Gamma(1/\nu)}\right), \qquad t \in \mathbb{R},$$

where γ denotes the lower incomplete gamma function. This choice of G_0 was examined by Azzalini (1986) in the case of $d = 1$ of (26). That author showed that G_0 is strictly log-concave if $\nu > 1$, leading to the log-concavity of (26) when $d = 1$. The case of $\nu = 1$ which corresponds to the Laplace distribution function is easily handled by direct computation of the second derivative to show the strict log-concavity of G_0. Now, combining strict log-concavity of G_0 with log-concavity of f_ν, proven above, an application of Corollary 1 shows that (26) is strictly log-concave on $\mathbb{R}^d$, if $\nu \geq 1$.

Although (26) is of skew-elliptical type, it does not appear to be of the type generated by the conditioning mechanism of the $(d+1)$-dimensional elliptical variate considered in Sect. 4.2. In fact, the results of Kano (1994) show that the set of densities f_ν is not closed under marginalization, and this also affects the conditioning mechanism (22).

As an example of non-elliptical distribution, we can consider a d-fold product of Subbotin's univariate densities, which is

$$f_\nu^*(x) = \prod_{j=1}^{d} c_\nu \exp(-|x_j|^\nu/\nu), \qquad x = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

and this density can be used as a replacement of f_ν in (26). Since each factor of this product is log-concave, if $\nu \geq 1$, the same property holds for f_ν^*. Strict log-concavity also holds for $2 f_\nu^*(x) G_0(\alpha^\top x)$, if again we assume the strict log-concavity of G_0.

Example To illustrate the applicability of Corollary 1 to distributions outside the set of type (2), consider the so-called extended skew-normal density which, in the d-dimensional case, takes the form

$$f(x) = \phi_d(x; \Omega)\,\frac{\Phi(\alpha_0 + \alpha^\top x)}{\Phi(\tau)}, \qquad (x \in \mathbb{R}^d), \tag{27}$$

where $\tau \in \mathbb{R}$ and $\alpha_0 = \tau(\alpha^\top \Omega \alpha)^{1/2}$. Although this distribution does not quite fall under the umbrella of Lemma 1 unless $\tau = 0$, its constructive argument is closely related.

To show the log-concavity of (27), first recall the well-known fact that $\phi_d(x; \Omega)$ is strictly log-concave. Next, note that Φ is log-concave, as it follows by direct calculation of the second derivative of $\log \Phi$, if we take into account the fact that $-y\,\Phi(y) < \phi(y)$ for every $y \leq 0$. In addition, as Φ is strictly increasing and $\alpha_0 + \alpha^\top x$ is concave in a non-strict sense, Corollary 1 applies, allowing us to conclude that (27) is strictly log-concave.

Although this conclusion is a special case of the result of Jamalizadeh and Balakrishnan (2010) concerning the log-concavity of the SUN distribution, it was presented here because the above argument is different. The log-concavity of other related subclasses of distributions is reported by Gupta and Balakrishnan (2010).

Acknowledgments We would like to thank three referees for their comments which have led to a better presentation of the material. This research has been supported by MIUR, Italy, under grant scheme PRIN, project No. 2006132978.

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.

Arellano-Valle, R. B., del Pino, G., San Martín, E. (2002). Definition and probabilistic properties of skew-distributions. *Statistics & Probability Letters, 58*(2), 111–121.

Arnold, B. C., Beaver, R. J. (2000). Hidden truncation models. *Sankhyā, Series A, 62*(1), 22–35.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics, 12*, 171–178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica, XLVI*(2), 199–208.

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*(3), 579–602. Extended version at arXiv.org: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*(2), 367–389. Extended version at arXiv.org:0911.2342.

Birnbaum, Z. W. (1948). On random variables with comparable peakedness. *Annals of Mathematical Statistics, 19*, 76–81.

Branco, M. D., Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis, 79*(1), 99–113.

Capitanio, A. (2008). On the canonical form of scale mixtures of skew-normal distributions. Unpublished manuscript.

Dharmadhikari, S. W., Joag-dev, K. (1988). *Unimodality, convexity, and applications*. New York & London: Academic Press.

Genton, M. G. (Ed.) (2004). *Skew-elliptical distributions and their applications: A journey beyond normality*. Boca Raton, FL, USA: Chapman & Hall/CRC.

Gupta, R. C., Balakrishnan, N. (2010). Log-concavity and monotonicity of hazard and reversed hazard functions of univariate and multivariate skew-normal distributions. *Metrika* (to appear).

Jamalizadeh, A., Balakrishnan, N. (2010). Distributions of order statistics and linear combinations of order statistics from an elliptical distribution as mixtures of unified skew-elliptical distributions. *Journal of Multivariate Analysis, 101*, 1412–1427.

Kano, Y. (1994). Consistency property of elliptical probability density functions. *Journal of Multivariate Analysis, 51*, 139–147.

Ma, Y., Genton, M. G. (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics, 31*, 459–468.

Marshall, A. W., Olkin, I. (1979) *Inequalities: theory of majorization and its applications*. In *Mathematics in science and engineering* (Vol. 143). New York & London: Academic Press.

Serfling, R. (2006). Multivariate symmetry and asymmetry. In S. Kotz, N. Balakrishnan, C. B. Read, B. Vidakovic (Eds.), *Encyclopedia of statistical sciences* (Vol. 8, 2nd ed., pp. 5338–5345). New York: Wiley.

Subbotin, M. T. (1923). On the law of frequency of error. *Matematicheskii Sbornik, 31*, 296–301.

Umbach, D. (2006). Some moment relationships for skew-symmetric distributions. *Statistics & Probability Letters, 76*(5), 507–512.

Umbach, D. (2008). Some moment relationships for multivariate skew-symmetric distributions. *Statistics & Probability Letters, 78*(12), 1619–1623.

Wang, J., Boyer, J., Genton, M. G. (2004). A skew-symmetric representation of multivariate distributions. *Statistica Sinica, 14*, 1259–1270.

Whitt, W. (2006). Stochastic ordering. In S. Kotz, N. Balakrishnan, C. B. Read, B. Vidakovic (Eds.), *Encyclopedia of statistical sciences* (Vol. 13, 2nd ed., pp. 8260–8264). New York: Wiley.