Communications in Statistics—Theory and Methods, 36: 1705–1717, 2007
Copyright © Taylor & Francis Group, LLC
ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920601126159

Extreme Value Theory and Order Statistics

Extreme Value Distributions for the Skew-Symmetric Family of Distributions

SHENG-MAO CHANG[1] AND MARC G. GENTON[2]

[1]Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA
[2]Department of Statistics, Texas A&M University, College Station, Texas, USA

We derive the extreme value distribution of the skew-symmetric family, the probability density function of the latter being defined as twice the product of a symmetric density and a skewing function. We show that, under certain conditions on the skewing function, this extreme value distribution is the same as that for the symmetric density. We illustrate our results using various examples of skew-symmetric distributions as well as two data sets.

Keywords Flexible skew-symmetric; Generalized skew-normal; Heavy tails; Multimodality; Selection models; Skew-Cauchy; Skew-*t*.

Mathematics Subject Classification Primary 62H05; Secondary 62E10.

1. Introduction

1.1. *The Skew-Symmetric Family of Distributions*

Interest in the construction of rival models to the normal distribution can be traced back to the nineteenth century, see e.g., Edgeworth (1886) and Pearson (1893). In recent years, applications from the environmental, financial, and biomedical sciences, among others, have further shown that data sets following the normal law are more often the exception rather than the rule. As a consequence, there has been renewed interest in the construction of parametric classes of non normal distributions. Wang et al. (2004) recently defined skew-symmetric (*SS*) distributions by means of their probability density function (pdf), g say, given by

$$g(x) = 2f(x)\pi(x), \tag{1}$$

Received February 1, 2005; Accepted May 31, 2006
Address correspondence to Marc G. Genton, Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA; E-mail: genton@stat.tamu.edu

where $f : \mathbb{R} \to [0, +\infty)$ is a symmetric pdf, and $\pi : \mathbb{R} \to [0, 1]$ is a skewing function satisfying $\pi(x) + \pi(-x) = 1$. When $f = \phi$, the standard normal pdf, the distribution with pdf (1) is referred to as being generalized skew-normal by Genton and Loperfido (2005). When $\pi(x) = 1/2$, the pdf $g = f$ and is thus symmetric. When the skewing function π is not constant, g is skewed. Clearly, the cumulative distribution function (cdf) of a symmetric distribution can be used as a skewing function. However, because a skewing function is not required to be increasing, it need not necessarily be a cdf. Depending on the skewing function chosen, g can even be multimodal. This fact has been exploited by Ma and Genton (2004) to define flexible skew-symmetric (*FSS*) distributions, a particular sub-family of *SS* distributions, for which the skewing function is defined by

$$\pi_K(x) = H(P_K(x)) = H\big(\alpha_1 x + \alpha_3 x^3 + \alpha_5 x^5 + \cdots + \alpha_K x^K\big). \tag{2}$$

Here H is the cdf of a continuous symmetric distribution evaluated at an odd polynomial P_K of order K in the variable x, with coefficients $\alpha_1, \alpha_3, \alpha_5, \ldots, \alpha_K$. *SS* distributions, and in particular *FSS* distributions, can model skewness, heavy tails, and multimodality systematically; see Genton (2004) for further discussions. Figure 1 depicts the pdf of two flexible skew-normal (*FSN*) distributions obtained from (1) and (2) with $f = \phi$ and $H = \Phi$, the standard normal cdf. Many of the parametric alternatives to the normal distribution that have been proposed in the literature are particular cases of (1); for example, the skew-normal (Azzalini, 1985), skew-*t* (Azzalini and Capitanio, 2003; Branco and Dey, 2001), and skew-Cauchy (Arnold and Beaver, 2000) distributions.

Besides the appealing features of the *SS* family for parametric modeling of non normally distributed data, there is also a natural and important context in which *SS* distributions arise, namely as selection models (see, e.g., Copas and Li, 1997). Consider, for instance, a random variable X with symmetric pdf $f(x; \beta)$, where β is a vector of unknown parameters. In practice, there are many situations where a random sample from f might not be available, for instance if it is too difficult or too costly to obtain. If the pdf f is distorted by some multiplicative non negative

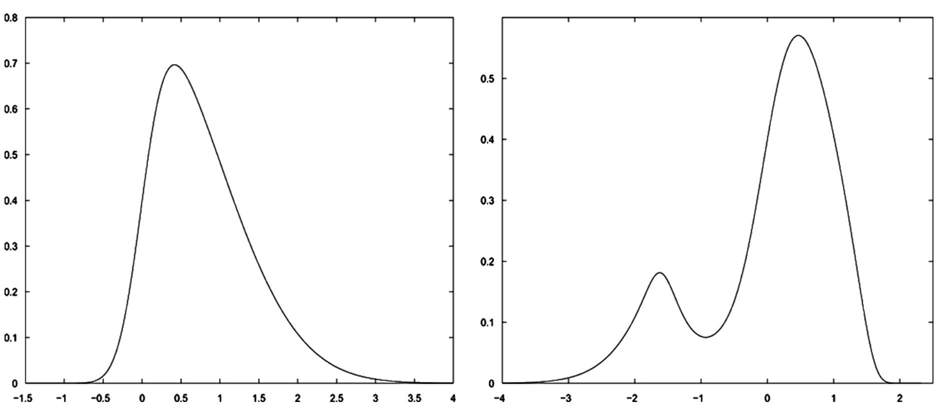

Figure 1. Two pdf's of *FSN* distributions. Left panel: $K = 1$ and $\alpha_1 = 4$ (a skew-normal pdf). Right panel: $K = 3$ and $\alpha_1 = 2, \alpha_3 = -1$ (a bimodal pdf).

weighting function $w(x; \beta, \alpha)$, where α denotes some vector of additional unknown parameters, then the observed data is a random sample from a distribution with pdf

$$g(x; \beta, \alpha) = f(x; \beta)\frac{w(x; \beta, \alpha)}{\mathrm{E}\{w(X; \beta, \alpha)\}}, \tag{3}$$

where g is said to be the pdf of a weighted distribution; see Rao (1985) and references therein. In particular, if the observed data are obtained only from a selected portion of the population of interest then (3) is called a selection model. When the weighting function satisfies $\mathrm{E}\{w(X; \beta, \alpha)\} = 1/2$, g is skew-symmetric of the form given by (1). For other selection schemes and discussions, we refer to Bayarri and DeGroot (1992), Arnold and Beaver (2002), Loperfido (2002), Genton (2004), Ma et al. (2005), and Genton (2005).

1.2. *Extreme Value Theory*

The Fisher–Tippett theorem (Fisher and Tippett, 1928) plays a key role in extreme value theory. It states that after suitable normalization (with location and scale normalizing constants c_n and $d_n > 0$, respectively), if the maximum (or minimum) of n independent and identically distributed random variables converges weakly to Q as $n \to \infty$, then Q is one of the following three families of extreme value distributions (EVD's) described by their cdf's:

(1) *Gumbel:* $\Lambda(x) = \exp(-e^{-x})$, $-\infty < x < \infty$,
(2) *Fréchet:* $\Phi_a(x) = \exp(-x^{-a})$, $x > 0$ and $a > 0$,
(3) *Weibull:* $\Psi_a(x) = \exp(-(-x)^a)$, $x \leq 0$ and $a > 0$.

Sketches of proofs, extensions, choices of normalizing constants, and applications, can be found in Embrechts et al. (1997), Kotz and Nadarajah (2000), and Coles (2001).

In applications, the problem is to identify the family to which Q belongs. Embrechts et al. (1997, Ch. 3), Resnick (1987, pp. 53–54), and Leadbetter et al. (1983, Ch. 1) summarize how this is done. Let $D(Q)$ denote the domain of attraction of Q, i.e., the collection of all those cdf's for which the distribution of the sample maximum converges to Q (where Q is one of the three EVD's: Λ, Φ_a, or Ψ_a). Following Resnick (1987, p. 53), define the tail ratio of a cdf G as

$$\lim_{x\to\infty} \frac{1 - G(tx)}{1 - G(x)}, \tag{4}$$

where $t > 1$. Then, $G \in D(\Phi_a)$ if and only if (4) converges to t^{-a} with $a > 0$. If $G \in D(\Lambda)$, then the limit (4) is equal to 0. Note that if G is continuous and defined on $\mathbb{R}$, then G never belongs to $D(\Psi_a)$. Our interest in this article is to study the tail behavior of G, a skew-symmetric distribution.

The remainder of the article is organized as follows. In Sec. 2, we describe the main results and their proofs. Basically, under certain conditions on the skewing function, the extreme value distribution of skew-symmetric distributions is the same as the extreme value distribution of the symmetric density. For instance, the extreme value distributions for the two pdf's depicted in Figure 1 are identical to that of the normal distribution. In Section 3, the extreme value distributions of several popular skew-symmetric distributions are examined closely: flexible and generalized skew-normal and three types of skew-t, including the skew-Cauchy. Finally, in Section 4, two real data sets are analyzed.

2. Extreme Value Distributions for SS Distributions

Consider a skew-symmetric distribution with pdf (1). Suppose we know that F, the cdf of a symmetric distribution with pdf f, belongs to one of the three EVD's. We are interested in the EVD of G, the cdf of the distribution with pdf g. In particular, we want to derive the conditions under which the EVD's of F and G are in the same family. We show next that these conditions are related to the skewing function π. Notice that EVD's do not exist for discrete distributions (Leadbetter et al., 1983, p. 117). Therefore, throughout this article we assume that the supports of F and π are the real line. In addition to the conditions on f and π mentioned after (1), we assume:

(A) π is continuous and there exists a constant $M > 0$ such that $\pi(x)$ is monotone for $x > M$;
(B) f and π have continuous second derivatives;
(C) there exists a constant $M^* > 0$ such that $g'(x) < 0$ for $x > M^*$.

Proposition 2.1. *Let f be a symmetric pdf with associated cdf F, and g the corresponding skew-symmetric pdf defined by* (1) *with associated cdf G and skewing function π. Assume* (A), (B), *and* (C). *If $F \in D(Q)$ and*

$$\lim_{x\to\infty} \pi(x) = \omega \in (0, 1],$$

then $G \in D(Q)$ also, where Q is either Λ (Gumbel) or Φ_a (Fréchet).

Proof. The proof of Proposition 2.1 is trivial if we are aware of Propositions 3.3.9 and 3.3.28 of Embrechts et al. (1997). The first of these states that if $F \in D(\Phi_a)$, i.e., $\lim_{x\to\infty} F^n(d_n x) = \Phi_a(x)$, then $\lim_{x\to\infty} G^n(d_n x) = \Phi_a(qx)$ for some $q > 0$ if and only if

$$\lim_{x\to\infty} \frac{1 - G(x)}{1 - F(x)} = q^{-a}. \tag{5}$$

The second proposition states that if $F \in D(\Lambda)$, i.e., $\lim_{x\to\infty} F^n(d_n x + c_n) = \Lambda(x)$, then $\lim_{x\to\infty} G^n(d_n x + c_n) = \Lambda(x + r)$ for some $r \in \mathbb{R}$ if and only if

$$\lim_{x\to\infty} \frac{1 - G(x)}{1 - F(x)} = e^{-r}. \tag{6}$$

Moreover, if $\lim_{x\to\infty} \pi(x) = \omega$ is in the interval (0,1], then by L'Hospital's rule:

$$\lim_{x\to\infty} \frac{1 - G(x)}{1 - F(x)} = \lim_{x\to\infty} \frac{G'(x)}{F'(x)} = \lim_{x\to\infty} \frac{2f(x)\pi(x)}{f(x)} = \lim_{x\to\infty} 2\pi(x) = [(2\omega)^{-1/a}]^{-a} = e^{\log(2\omega)}.$$

In words, if F is in $D(\Phi_a)$ then G satisfies (5) with $q = (2\omega)^{-1/a}$. On the other hand, if F is in $D(\Lambda)$ then G satisfies (6) with $r = -\log(2\omega)$. So the proof is complete. □

For distributions in the SS family it commonly occurs that $\omega = 0$, and a natural question to ask is what happens for this situation. For instance, take the skew-normal distribution, i.e., $f = \phi$, $H = \Phi$, and $K = 1$ in (1) and (2). If the shape parameter α_1 is negative, then $\omega = 0$ as $x \to \infty$. Therefore, we need a general procedure to show that $G \in D(\Lambda)$ or $G \in D(\Phi_a)$. The following criterion (Leadbetter et al., 1983, Theorem 1.6.1) gives a sufficient condition for the case $G \in D(\Lambda)$.

Proposition 2.2. *Assume* (A), (B), *and* (C). *Then,* $G \in D(\Lambda)$ *if*

$$\lim_{x\to\infty} \frac{[1 - G(x)]G''(x)}{[G'(x)]^2} = -1. \tag{7}$$

To summarize, for $F \in D(\Lambda)$, we have:

(1) if $\lim_{x\to\infty} \pi(x) = \omega \in (0, 1]$, then $G \in D(\Lambda)$;
(2) if not, then check Proposition 2.2.

Notice that if $F \in D(\Lambda)$ then G is never in $D(\Phi_a)$. Effectively, recall from the discussion in Sec. 1.2 that $F \in D(\Lambda)$ means its tail ratio is 0. Therefore, the tail ratio of G is always 0 because the limiting ratio $\pi(tx)/\pi(x)$ is always finite due to the bounded image of π and the monotonicity in its tail.

If, now, $F \in D(\Phi_a)$ and $\lim_{x\to\infty} \pi(x) = 0$, then we can check Proposition 2.3.

Proposition 2.3. *Assume* (A), (B), *and* (C). *Suppose* $F \in D(\Phi_a)$ *and* $\lim_{x\to\infty} \pi(x) = 0$. *If, for any* $t > 1$,

$$\lim_{x\to\infty} \frac{\pi(tx)}{\pi(x)} > 0, \tag{8}$$

then $G \in D(\Phi_a)$.

Proof. As mentioned in Sec. 1.2, if $F \in D(\Phi_a)$ then the tail ratio of F should be of the form t^{-a}, $a > 0$. So, the proof begins by considering the tail ratio of G. For any $t > 1$ we have, by L'Hospital's rule,

$$\lim_{x\to\infty} \frac{1 - G(tx)}{1 - G(x)} = \lim_{x\to\infty} \left[\frac{tf(tx)}{f(x)} \times \frac{\pi(tx)}{\pi(x)} \right].$$

Since the first term on the right-hand side converges to t^{-a}, we conclude that G is in $D(\Phi_a)$ if the limit of $\pi(tx)/\pi(x)$ is a constant. Again, due to the bounded image of π and the monotonicity of its tail, we are sure that this ratio cannot be infinity. Besides, for any $t > 1$, the parameter of the Fréchet distribution is $a^* = a - \log_t k$. Because $0 \leq \pi(x) \leq 1$ and the existence of $\lim_{x\to\infty} \pi(x)$, we can infer that $0 \leq k \leq 1$ and hence that $\log_t k < 0$. Therefore, $\lim_{x\to\infty} \frac{1-G(tx)}{1-G(x)} = t^{-a^*}$ and $a^* > 0$, i.e., $G \in D(\Phi_a)$. □

If condition (8) in Proposition 2.3 is not satisfied then we still have to check (7) in Proposition 2.2. Again, for $F \in D(\Phi_a)$, we have:

(1) if $\lim_{x\to\infty} \pi(x) = \omega \in (0, 1]$, then $G \in D(\Phi_a)$;
(2) if $\lim_{x\to\infty} \pi(x) = 0$ and $\lim_{x\to\infty} \pi(tx)/\pi(x) \neq 0$, then $G \in D(\Phi_a)$ for any $t > 1$;
(3) if $\lim_{x\to\infty} \pi(x) = 0$ and $\lim_{x\to\infty} \pi(tx)/\pi(x) = 0$, then check (7) in Proposition 2.2.

Three remarks are made to complete this section.

Remark 2.1. Propositions 3.3.9 and 3.3.28 of Embrechts et al. (1997) imply that if the limit of $(1-G)/(1-F)$ is a constant in $(0,1]$ then G is in the same EVD family as F. That is, we can generate a special family of *SS* distributions based on the corresponding symmetric distribution, such that G and F have the same EVD. However, if an *SS* distribution does not satisfy the conditions of Propositions 2.1–2.3, then we are unable to say anything specific concerning its EVD. Examples can be found in Sec. 3.

Remark 2.2. The use of the EVD of F when the sample size n is finite leads to an approximation, the quality of which is dictated by the von Mises form

$$1 - F(x) = c\exp\left\{\int_{-\infty}^{x} -\frac{1}{\mu(t)}dt\right\},$$

where $c > 0$ and μ is strictly positive and continuous. For example, the exponential distribution is exactly of the von Mises form but the normal distribution is only approximately of the von Mises form. Numerical studies show that a sample size of $n = 10$ is large enough for using the EVD of an exponential distribution, while $n = 100$ is too small in the case of a normal distribution.

Remark 2.3. If $F \in D(\Lambda)$, then the normalizing constants can be chosen as $c_n = F^{-1}(1-1/n)$ and $d_n(1-F(c_n))/F'(c_n)$. If $F \in D(\Phi_a)$, then $c_n = 0$ and $d_n = F^{-1}(1-1/n)$. Notice that the choices for c_n and d_n are not unique. For other choices see Leadbetter et al. (1983) and Resnick (1987). The optimal choices for the normalizing constants c_n and d_n, defined by the smallest departure from the von Mises form in Remark 2.2, were derived by Hall (1979) for the normal distribution.

3. Special Cases

In this section, we discuss the EVD's of several popular skew-symmetric distributions: flexible and generalized skew-normal, and three types of skew-t. We use Propositions 2.1–2.3 from the previous section to determine whether G belongs to $D(\Lambda)$ or $D(\Phi_a)$. In order to avoid any potential confusion, note that, throughout, Φ represents the standard normal cdf while Φ_a denotes the Fréchet EVD distribution with parameter a.

3.1. *Flexible and Generalized Skew-Normal Distributions*

Consider the flexible skew-normal (*FSN*) distribution defined by Ma and Genton (2004) through the pdf $g(x) = 2\phi(x)\Phi(P_K(x))$. It is well known that $\Phi \in D(\Lambda)$. Using Propositions 2.1 and 2.2, we show next that the cdf of the *FSN* distribution is also in $D(\Lambda)$.

Proposition 3.1. *Let $g(x) = 2\phi(x)\Phi(P_K(x))$ be the pdf of an FSN distribution with associated cdf G. Then, G is in $D(\Lambda)$.*

Proof. Suppose $\lim_{x\to\infty} P_K(x) \in (-\infty,\infty)$ or $\lim_{x\to\infty} P_K(x) = \infty$, i.e., $\lim_{x\to\infty}\pi(x) = \lim_{x\to\infty}\Phi(P_K(x)) = \omega \in (0,1]$. Applying Proposition 2.1 directly, we have $G \in D(\Lambda)$. Otherwise, when $\lim_{x\to\infty} P_K(x) = -\infty$, we need to use Proposition 2.2.

The following results are essential to the proof. First, when x is large, $1-\Phi(x) \approx \frac{\phi(x)}{x}$, $g(x) \approx -\frac{2}{P_K(x)}\phi(x)\phi(P_K(x))$ and $g'(x) \approx 2\big(\frac{x}{P_K(x)} + P_K'(x)\big)\phi(x)\phi(P_K(x))$. Also, $\lim_{x\to\infty} \frac{1-G(x)}{g'(x)} = 0$, $\lim_{x\to\infty} \frac{g'(x)}{g(x)P_K'(x)P_K(x)} = -1$ and $\lim_{x\to\infty}[1-G(x)]P_K'(x)P_K(x) = 0$. Thus,

$$
\begin{aligned}
\lim_{x\to\infty} \frac{[1-G(x)]g'(x)}{[g(x)]^2} &= \lim_{x\to\infty} \left(\frac{[1-G(x)]P_K'(x)P_K(x)}{g(x)} \times \frac{g'(x)}{g(x)P_K'(x)P_K(x)} \right) \\
&= \lim_{x\to\infty} \left(\frac{-g(x)P_K'(x)P_K(x)}{g'(x)} + \frac{[1-G(x)]/g'(x)}{[(P_K'(x))^2 + P_K''(x)P_K(x)]^{-1}} \right) \\
&\quad \times \lim_{x\to\infty} \frac{g'(x)}{g(x)P_K'(x)P_K(x)} \\
&= (1+0)\times(-1) = -1,
\end{aligned}
$$

where L'Hospital's rule is used at the second step. So the proof is complete. □

Notice that when $K=1$ we obtain the skew-normal (*SN*) distribution, which therefore belongs to $D(\Lambda)$; see also the discussion on quantiles of the *SN* distribution in Genton (2005, p. 190). For the generalized skew-normal (*GSN*) distributions defined by Genton and Loperfido (2005) through the pdf $g(x) = 2\phi(x)\pi(x)$, where π is a skewing function, we have the following result. If $\lim_{x\to\infty}\pi(x) = \omega \in (0,1]$ then, by Proposition 2.1, the *GSN* distribution is in $D(\Lambda)$. Otherwise, if $\lim_{x\to\infty}\pi(x) = 0$, no conclusion can be drawn about the EVD of the *GSN* distribution unless a parametric form of the skewing function π is specified.

3.2. *Skew-t Distributions*

Let $s(x;\nu)$ and $S(x;\nu)$ denote the pdf and cdf, respectively, of the t-distribution with ν degrees of freedom. The tail ratio of S is $t^{-\nu}$, i.e., $S \in D(\Phi_a)$ with $a=\nu$. There are three popular definitions of skew-t distributions. The Type I skew-t (*ST-I*) implicit from Azzalini (1985) with pdf

$$
g_I(x;\nu,\alpha) = 2s(x;\nu)S(\alpha x;\nu), \tag{9}
$$

where $\alpha \in \mathbb{R}$. When $\nu = 1$, g_I is the pdf of the skew-Cauchy distribution (Arnold and Beaver, 2000). The Type II skew-t (*ST-II*) proposed by Branco and Dey (2001), and in an equivalent form by Azzalini and Capitanio (2003), with pdf

$$
g_{II}(x;\nu,\alpha) = 2s(x;\nu)S\left(\alpha x\sqrt{\frac{1+\nu}{x^2+\nu}};\nu+1\right), \tag{10}
$$

and the Type III skew-t (*ST-III*) proposed by Jones (2001) with pdf

$$
g_{III}(x;\eta_a,\eta_b) = \Delta^{-1}_{\eta_a,\eta_b}\left(1+\frac{x}{[\eta_a+\eta_b+x^2]^{1/2}}\right)^{\eta_a+1/2}\left(1-\frac{x}{[\eta_a+\eta_b+x^2]^{1/2}}\right)^{\eta_b+1/2}, \tag{11}
$$

where $\eta_a, \eta_b > 0$, $\Delta_{\eta_a,\eta_b} = 2^{\eta_a+\eta_b-1}(\eta_a+\eta_b)^{1/2}B(\eta_a,\eta_b)$ and B is the beta function. Their EVD's are described in the next proposition.

Proposition 3.2. *The skew-t distributions with pdf's g_I, g_{II}, g_{III}, and associated cdf's G_I, G_{II}, G_{III} have the following EVD's:*

(i) *$G_I(x; v, \alpha)$ is in $D(\Phi_a)$ with $a = v$ when $\alpha \geq 0$ and $a = 2v$ when $\alpha < 0$;*
(ii) *$G_{II}(x; v, \alpha)$ is in $D(\Phi_a)$ with $a = v$ for all $\alpha \in \mathbb{R}$;*
(iii) *$G_{III}(x; \eta_a, \eta_b)$ is in $D(\Phi_a)$ with $a = 2\eta_a$ for all $\eta_a > 0$ and $\eta_b > 0$.*

Proof. (i) Applying Proposition 2.1, we know that G_I is in $D(\Phi_a)$ with $a = v$ if $\alpha \geq 0$. When $\alpha < 0$ we can also find that $\lim_{x\to\infty} S(\alpha tx; v)/S(\alpha x; v) = t^{-v}$. That is, by Proposition 2.3, $G_I \in D(\Phi_a)$ with $a = 2v$ for $\alpha < 0$.

(ii) Similarly, since the *ST-II*'s skewing function converges to $S(\alpha; v+1)$, a constant, as $x \to \infty$, we conclude that G_{II} is in $D(\Phi_a)$ and $a = v$ for all $\alpha \in \mathbb{R}$ by Proposition 2.1.

(iii) In order to relate the *ST-III* to results already obtained, its pdf can be rewritten as

$$
\begin{aligned}
g_{III}(x; \eta_a, \eta_b) &= \Delta^{-1}_{\eta_a,\eta_b}\left(1 + \frac{x^2}{\eta_a + \eta_b}\right)^{-\frac{\eta_a+\eta_b+1}{2}}\left(1 + \frac{x^2}{\eta_a + \eta_b}\right)^{-\frac{\eta_a-\eta_b}{2}} \\
&\quad \times \left(1 + \frac{x}{[\eta_a + \eta_b + x^2]^{1/2}}\right)^{\eta_a+1/2}\left(1 + \frac{x}{[\eta_a + \eta_b + x^2]^{1/2}}\right)^{-(\eta_b+1/2)} \\
&= \Delta^{-1}_{\eta_a,\eta_b} \times \left(1 + \frac{x^2}{\eta_a + \eta_b}\right)^{-\frac{\eta_a+\eta_b+1}{2}} \times \Pi^*(x, \eta_a, \eta_b),
\end{aligned}
$$

where

$$
\begin{aligned}
\Pi^*(x, \eta_a, \eta_b) &= \left(1 + \frac{x^2}{\eta_a + \eta_b}\right)^{-(\eta_a-\eta_b)/2} \times \left(1 + \frac{x}{[\eta_a + \eta_b + x^2]^{1/2}}\right)^{\eta_a-\eta_b} \\
&= \left\{\frac{[(\eta_a + \eta_b + x^2)^{1/2} + x](\eta_a + \eta_b)^{1/2}}{\eta_a + \eta_b + x^2}\right\}^{\eta_a-\eta_b}.
\end{aligned}
$$

Then, for any $t > 1$, the tail ratio of G_{III} is

$$
\begin{aligned}
\lim_{x\to\infty} \frac{1 - G_{III}(tx; \eta_a, \eta_b)}{1 - G_{III}(x; \eta_a, \eta_b)} &= \lim_{x\to\infty} \frac{t g_{III}(tx; \eta_a, \eta_b)}{g_{III}(x; \eta_a, \eta_b)} \\
&= \lim_{x\to\infty} \frac{t\left(1 + \frac{t^2x^2}{\eta_a+\eta_b}\right)^{-\frac{\eta_a+\eta_b+1}{2}}}{\left(1 + \frac{x^2}{\eta_a+\eta_b}\right)^{-\frac{\eta_a+\eta_b+1}{2}}} \times \frac{\Pi^*(tx, \eta_a, \eta_b)}{\Pi^*(x, \eta_a, \eta_b)} \\
&= t^{-(\eta_a+\eta_b)} \times t^{-(\eta_a-\eta_b)} = t^{-2\eta_a}.
\end{aligned}
$$

That is, G_{III} belongs to $D(\Phi_a)$ with $a = 2\eta_a$. □

From the perspective of its EVD, the *ST-II* is the simplest of the skew-*t* distributions since it has the same EVD for any $\alpha \in \mathbb{R}$. The EVD of the skew-Cauchy (*SC*) distribution introduced by Arnold and Beaver (2000) is that of an *ST-I* distribution with $v = 1$. The results of this section are summarized in Table 1.

Table 1
Extreme value distributions of certain skew-symmetric distributions

Family	Pdf	Domain of attraction	Condition
SN	$2\phi(x)\Phi(\alpha_1 x)$	$D(\Lambda)$	$\alpha_1 \in \mathbb{R}$
FSN	$2\phi(x)\Phi(P_K(x))$	$D(\Lambda)$	$\alpha_i \in \mathbb{R}$, $i = 1, 3, 5, \ldots, K$ odd
GSN	$2\phi(x)\pi(x)$	$D(\Lambda)$	$\lim_{x\to\infty} \pi(x) = \omega \in (0, 1]$
ST-I	$2s(x; \nu)S(\alpha x; \nu)$	$D(\Phi_a)$ $a = \nu$	$\alpha \geq 0$
		$a = 2\nu$	$\alpha < 0$
ST-II	$g_{II}(x; \alpha)$	$D(\Phi_a)$ $a = \nu$	$\alpha \in \mathbb{R}$
ST-III	$g_{III}(x; \eta_a, \eta_b)$	$D(\Phi_a)$ $a = 2\eta_a$	$\eta_a > 0$, $\eta_b > 0$

Remark 3.1. When $\alpha_1 = \alpha = 0$, the *SN*, *SC*, *ST-I*, and *ST-II* distributions reduce to the normal, Cauchy, and Student-t distributions, respectively. The *FSN* distribution reduces to the normal distribution if $\alpha_1 = \cdots = \alpha_K = 0$. Thus, the EVD's of these symmetric distributions are represented by the rows of Table 1 with the conditions $\alpha \in \mathbb{R}$ or $\alpha \geq 0$.

Remark 3.2. Consider an *FSS* distribution with pdf given by (1) and (2) with $H = F$, that is,

$$g(x) = 2f(x)F(P_K(x)). \tag{12}$$

Assume $F \in D(\Phi_a)$ with $a = r$, i.e., the tail ratio of the distribution F is t^{-r}. The tail ratio of G is, therefore,

$$\lim_{x\to\infty} \frac{1 - G(tx)}{1 - G(x)} = \lim_{x\to\infty} \frac{tf(tx)F(P_K(tx))}{f(x)F(P_K(x))} = \lim_{x\to\infty} \frac{tf(tx)}{f(x)} \times \frac{F(P_K(tx))}{F(P_K(x))}. \tag{13}$$

Trivially, if $\alpha_K > 0$ then the last term of (13) converges to 1. Otherwise, it is undefined. Hence, using L'Hospital's rule, we have

$$\begin{aligned} \lim_{x\to\infty} \frac{F(P_K(tx))}{F(P_K(x))} &= \lim_{x\to\infty} \frac{F(\alpha_K t^K x^K)}{F(\alpha_K x)} = \lim_{x\to\infty} \frac{f(\alpha_K t^K x^K)\alpha_K t^K K x^{K-1}}{f(\alpha_K x^K)\alpha_K K x^{K-1}} \\ &= \lim_{x\to\infty} \frac{t^K f(t^K x^*)}{f(x^*)} = t^{-Kr}, \end{aligned} \tag{14}$$

where $x^* = \alpha_K x^K$. Thus, combining (13) and (14), we conclude that G is in $D(\Phi_a)$ with $a = r(K + 1)$ when α_K is negative, and $a = r$ when α_K is positive.

Remark 3.3. All the skewed distributions defined so far have the same domain of attraction as their corresponding symmetric kernel distributions. However, this need not necessarily be the case. As an example, suppose we define a generalized skew-Cauchy (*GSC*) by the pdf

$$g(x) = 2c(x)\Phi(\alpha x), \tag{15}$$

where $c(x)$ is the standard Cauchy pdf whose cdf is in $D(\Phi_a)$ with $a = 1$, $\Phi(x)$ is the standard normal cdf, and $\alpha \in \mathbb{R}$. Considering the tail ratio of the GSC distribution with pdf (15) and cdf G, we have

$$\lim_{x\to\infty} \frac{1 - G(tx)}{1 - G(x)} = \lim_{x\to\infty} \frac{tg(tx)}{g(x)} = \lim_{x\to\infty} \frac{tc(tx)}{c(x)} \times \frac{\Phi(\alpha tx)}{\Phi(\alpha x)}, \tag{16}$$

for $t > 1$. Since the tail ratio of the Cauchy distribution is t^{-1}, (16) is equal to t^{-1} if α is non negative; otherwise it is equal to zero, i.e., G with negative α is not in $D(\Phi_a)$. According to Proposition 2.2, we can show that G is in $D(\Lambda)$. Now, when x is large and $\alpha < 0$, we have $G'(x) \approx -\frac{2}{\alpha x} c(x)\phi(\alpha x)$, $G''(x) \approx -\frac{2}{\alpha x}[c'(x) - \alpha^2 x c(x)]\phi(\alpha x)$ and $G'''(x) \approx -\frac{2}{\alpha x}[c''(x) - 2\alpha^2 x c'(x) + \alpha^3 x^2 c(x)]\phi(\alpha x)$. Also, $\lim_{x\to\infty} \frac{G''(x)}{xG'(x)} = -\alpha^2$ and $\lim_{x\to\infty} \frac{1-G(x)}{G''(x)} = \lim_{x\to\infty} \frac{c(x)}{\alpha^2 x c(x)} = 0$.

Thus,

$$\begin{aligned} \lim_{x\to\infty} \frac{[1 - G(x)]G''(x)}{[G(x)]^2} &= \lim_{x\to\infty} \frac{[1 - G(x)]x}{G'(x)} \times \frac{G''(x)}{xG'(x)} \\ &= \lim_{x\to\infty} \frac{[1 - G(x)] - xG'(x)}{G''(x)} \times \frac{G''(x)}{xG'(x)} \\ &= \left[0 - \left(-\frac{1}{\alpha^2}\right)\right] \times (-\alpha^2) = -1. \end{aligned}$$

So the condition of Proposition 2.2 is satisfied and hence G is in $D(\Lambda)$.

4. Applications

Our first illustrative example is based on a data set considered by Smith and Taylor (1987) consisting of the measured strengths of $n = 63$ pieces of 1.5 cm long glass fiber. A histogram of the data reveals negative skewness. This feature prompted Jones and Faddy (2003), as well as Azzalini and Capitanio (2003), to fit a skew-t distribution to the strength values. Ma and Genton (2004) used a flexible skew-t (FST) distribution in order to explore the possibility of multimodality. They concluded that a model with a polynomial of order $K = 1$ was sufficient for this data set, using a likelihood ratio test as well as other model selection criteria. All three studies revealed significant skewness and heavy tail behavior in the distribution of the data. We fitted the second form of the skew-t (ST-II) distribution via maximum likelihood, yielding the parameter estimates $\hat{\mu} = 1.749$, $\hat{\sigma} = 0.261$, $\hat{\alpha}_1 = -1.550$, and $\hat{\nu} = 2.734$.

Suppose now that we are interested in the distribution of the maximum strength over 1,000 randomly selected fibers. According to Sec. 3.2, the ST-II distribution is in $D(\Phi_a)$ with $a = \hat{\nu}$, using (10) as the pdf of the underlying distribution. From Remark 2.3, the normalizing constants are $c_n = 0$ and $d_n = 2.720$. Therefore, a reasonable estimate of the pth quantile of the distribution of the strongest fiber is the inverse function of the Fréchet distribution. More precisely, if we denote the strength of the ith fiber by X_i, then the pth quantile, q say, of the distribution of the strongest fiber is given by

$$p = \Pr\left(\max_i X_i \le q\right) = \Pr\left(\frac{\max_i X_i}{d_n} \le \frac{q}{d_n}\right) = \exp\left\{-\left(\frac{q}{d_n}\right)^{-a}\right\},$$

and hence, $\hat{q} = d_n(-\log p)^{-1/\hat{\nu}}$. For $p = 0.8$, 0.9, 0.95, and 0.99 we obtain $\hat{q} = 4.708$, 6.195, 8.061, and 14.632, respectively.

In our second example, we consider a data set from Cook and Weisberg (1994) referring to measurements made on $n = 202$ elite Australian athletes. Genton and Loperfido (2005) have argued that characteristics of the Australian adult population can be derived by means of a selection model (see Sec. 1), because only gifted athletes are included in the data set. Here, we consider the variable BMI (body mass index) which measures body fat based on height and weight. It also shows direct or indirect correlation with cardiovascular disease, high blood pressure, osteoarthritis, and diabetes. For adults (more than 20 years old), a person with a BMI larger than 25 is classified as being overweight (Garrow and Webster, 1985). We assume that BMI is normally distributed for the wider male and female populations, but is not necessarily so for the selected athletes.

If we believe the observed BMI's form random samples from the wider, normally distributed, male and female populations, then the fitted distributions are $N(23.904, 2.767^2)$ for the males and $N(21.989, 2.640^2)$ for the females. On the other hand, if we believe the observed BMI's have been obtained through a selection mechanism, then a skew-normal distribution may be more appropriate than a normal, see Ma et al. (2005). The skew-normal distributions fitted by maximum likelihood are $SN(20.743, 2.047^2, 3.623)$ for the males and $SN(19.228, 1.952^2, 2.129)$ for the females. Using the skew-normal fits, suppose we want to compute the probability of maximum BMI exceeding 25. Let MAX20, MAX50, and MAX100 denote the maximum over 20, 50, and 100 randomly sampled people, respectively. Further, denote the cdf of the fitted skew-normal distribution for males by G. Then, based on the results from Sec. 3.1, the probability that MAX20 is over the upper bound of 25 for male athletes is

$$\Pr\left(\max_i X_i > 25\right) = \Pr\left(\frac{\max_i X_i - c_n}{d_n} > \frac{25 - c_n}{d_n}\right) = 0.531,$$

with $c_n = G^{-1}(1 - 1/20)$ and $d_n = [1 - G(c_n)]/g(c_n)$. Table 2 presents this and the corresponding probabilities for the other sample sizes for males and females assumed to have been sampled from their wider populations or from their respective athletic sub-populations. These results appear to suggest that the current BMI classification based on an upper bound of 25 is confounded with gender, and thus that different BMI thresholds should be specified for the different sexes.

Table 2
Probability that at least one person is over the upper bound of 25 for BMI

	MAX20		MAX50		MAX100	
Assumed population	Male	Female	Male	Female	Male	Female
Athletic	0.531	0.093	0.855	0.170	0.985	0.283
Wider	0.335	0.065	0.610	0.106	0.855	0.170

Acknowledgments

The authors are most grateful to the guest editors and two referees for suggestions that improved the article. The work of Genton was partially supported by NSF grant DMS-0504896.

References

Arnold, B. C., Beaver, R. J. (2000). The skew-Cauchy distribution. *Statist. Probab. Lett.* 49:285–290.

Arnold, B. C., Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Test* 11:7–54.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist.* 12:171–178.

Azzalini, A., Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-*t* distribution. *J. Roy. Stat. Soc. Ser. B* 65:367–389.

Bayarri, M. J., DeGroot, M. (1992). A BAD view of weighted distributions and selection models. In: Bernardo, J. M., Berger, J. O., Dawid, A. P., Smith, A. F. M., eds. *Bayesian Statistics 4*. London: Oxford University Press, pp. 17–29.

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

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer.

Cook, R. D., Weisberg, S. (1994). *An Introduction to Regression Graphics*. New York: Wiley.

Copas, J. B., Li, H. G. (1997). Inference for non-random samples (with discussion). *J. Roy. Stat. Soc. Ser. B* 59:55–95.

Edgeworth, F. Y. (1886). The law of error and the elimination of chance. *Philosophical Mag.* 21:308–324.

Embrechts, P., Kluppelberg, C., Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. New York: Springer.

Fisher, R. A., Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Philosophical Soc.* 24:180–190.

Garrow, J. S., Webster, J. (1985). Quetelet's index (W/H2) as a measure of fatness. *Int. J. Obesity* 9:148–153.

Genton, M. G. (2004). *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Edited Volume. Boca Raton, FL: Chapman & Hall/CRC.

Genton, M. G. (2005). Discussion of "The skew-normal distribution and related multivariate families" by A. Azzalini. *Scand. J. Statist.* 32:189–198.

Genton, M. G., Loperfido, N. (2005). Generalized skew-elliptical distributions and their quadratic forms. *Ann. Inst. Statist. Math.* 57:389–401.

Hall, P. G. (1979). On the rate of convergence of normal extremes. *J. Appl. Probab.* 16:433–439.

Jones, M. C. (2001). Multivariate *t* and beta distributions associated with the multivariate *F* distribution. *Metrika* 54:215–231.

Jones, M. C., Faddy, M. J. (2003). A skew extension of the *t* distribution, with applications. *J. Roy. Stat. Soc. Ser. B* 65:159–174.

Kotz, S., Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. U.K: Imperial College Press.

Leadbetter, M. R., Lindgren, G., Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer.

Loperfido, N. (2002). Statistical implications of selectively reported inferential results. *Statist. Probab. Lett.* 56:13–22.

Ma, Y., Genton, M. G. (2004). A flexible class of skew-symmetric distributions. *Scand. J. Statist.* 31:459–468.

Ma, Y., Genton, M. G., Tsiatis, A. A. (2005). Locally efficient semiparametric estimators for generalized skew-elliptical distributions. *J. Amer. Statist. Assoc.* 100:980–989.

Pearson, K. (1893). Asymmetrical frequency curves. *Nature* 48:615–616.

Rao, C. R. (1985). Weighted distributions arising out of methods of ascertainment: What populations does a sample represent?. In: Atkinson, A. G., Fienberg, S. E., eds. *A Celebration of Statistics: The ISI Centenary Volume*. New York: Springer-Verlag, pp. 543–569.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. New York: Springer-Verlag.

Smith, R. L., Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *J. Roy. Statist. Soc. Ser. C* 36:358–369.

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