One-day course
on symmetry-modulated distributions

Adelchi Azzalini
Università di Padova, Italia

Skewed world of data:
Workshop in honor of Reinaldo B. Arellano-Valle’s 65th birthday
October 2017
Pontificia Universidad Católica de Chile
Prólogos
Lots of distributions available, do we need more?

- Probability textbooks introduce ‘standard’ distributions
- Over the years many others have been introduced
- Classical work includes proposals by K. Pearson, Fechner, Edgeworth, Johnson, Burr, etc.
- Still the search keeps going.
- Two currently popular general approaches:
  (‘general’: allowing unlimited number of specific constructions)
  - copulae
  - symmetry-modulated distributions, AKA skew-symmetric distributions
- Question: why so much effort?
Illustration: QQ-normal probability plots from two samples

old days sample

today’s sample
Larger datasets require more accurate modelling

- The two datasets are sampled from the same distribution
- The visual message of normal QQ-plot is completely different
- although only the sample size has changed
- Only the larger sample could highlight non-normality
- Today larger and larger datasets are available
- More data is good, but also more challenging
- We need flexible tools for accurate modelling of large datasets
Multivariate datasets are increasingly more frequent

- data collection is more often multivariate, possibly highly so
- many above-quoted formulations are univariate
- special interest in developing flexible multivariate distributions
- ...flexible yet mathematically tractable
Our plan of work

A tutorial to symmetry-modulated distributions:

- introduce main concepts in the univariate case
- focus on key special cases
- extend concepts to the multivariate settings
- sketch of some extensions
- followed by practical work with R package ‘sn’
Básis \((d=1)\)
Skew-normal distribution – idea

Idea: start from a normal distribution and ‘perturb’ it. Perturbation, or modulation, is achieved by a selection mechanism.
Skew-normal distribution – compute density function

assume : \((X, W) \sim N_2(0, \Sigma), \quad \Sigma = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}\)

recall : \((W | X = x) \sim N(\delta x, 1 - \delta^2)\)

(density at \(x|W \geq 0\))

\[
\begin{align*}
\frac{1}{dx} \mathbb{P}\{X \in (x, x + dx)|W \geq 0\} &= \frac{1}{dx} \mathbb{P}\{X \in (x, x + dx) \cap W \geq 0\} \\
&= \frac{1}{dx} \mathbb{P}\{X \in (x, x + dx)\} \mathbb{P}\{W \geq 0\} \\
&= \frac{1}{dx} \mathbb{P}\{X \in (x, x + dx)\} \frac{1}{2} \\
&= 2 \varphi(x) \Phi(\alpha x), \quad \alpha = \frac{\delta}{\sqrt{1 - \delta^2}} \in \mathbb{R}
\end{align*}
\]

write : \(Z \equiv (X|W \geq 0) \sim \text{SN}(\alpha)\)
Skew-normal distribution – density function plots

\[ \alpha > 0: \text{positive asymmetry} \]
\[ \alpha = 0: \text{null asymmetry, i.e. } \mathcal{N}(0, 1) \]

\[ \alpha < 0: \text{negative asymmetry} \]
\[ \alpha = 0: \text{null asymmetry, i.e. } \mathcal{N}(0, 1) \]
Towards a general result, preliminaries

- let \((X, W) \sim \mathcal{N}_2(0, \Sigma)\) as before
- \(T = -(W - \delta X)/\sqrt{1 - \delta^2} \sim \mathcal{N}(0, 1)\)
- \(\text{cov}\{X, T\} = 0 \implies X \perp T\) (independent)
- \((W \geq 0)\) is algebraically equivalent to \((T \leq \alpha X)\)
- hence \(Z \equiv (X|W \geq 0) \equiv (X|T \leq \alpha X)\)
- Note the key ingredients here:
  \(X \perp T, X\) and \(T\) symmetric about 0, and so is \(T - \alpha X\)
A general result

**Lemma (Univariate version)**

If $f_0$ is PDF and $G_0$ a continuous CDF on $\mathbb{R}$, both symmetric about 0, then

$$f(x) = 2 f_0(x) G_0\{w(x)\}, \quad x \in \mathbb{R},$$

is a proper density function for any odd function $w$.

Proof. Denote $X \sim f_0$ and $T \sim G_0$, independent rv’s. The distribution of $T - w(X)$ is symmetric about 0. Then

$$\frac{1}{2} = \mathbb{P}\{T - w(X) \leq 0\} = \mathbb{E}_X \{\mathbb{P}\{T \leq w(x) | X = x\}\} = \int_{\mathbb{R}} G_0\{w(x)\} f_0(x) \, dx.$$
Some comments

- Above result allows to combine freely $f_0$, $G_0$ and $w$: a huge variety of constructions are possible.
- However, ‘possible’ does not automatically imply ‘useful’: need to select those which are worth of consideration.
- The result works also if the support is a subset of $\mathbb{R}$.
- The lemma allows a number of extensions: multivariate, non-odd $w$, discrete variables, etc. (Some of these extensions will be examined later).
- From the assumptions of the lemma, $G(x) = G_0\{w(x)\}$ satisfies
  \[ G(x) \geq 0, \quad G(x) + G(-x) = 1. \]
- Possible to formulate the result equivalently in terms of $G(x)$. 
Random number generation / stochastic representation

**Crude version**  Generate $X \sim f_0$ and $T \sim G_0$ independently and set

$$Z = \{X|T \leq w(X)\}$$

Drawback: reject sampled values with $T > w(X)$, half of them on average.

**Improved version**

$$Z = \begin{cases} 
X & \text{if } T \leq w(X) \\
-X & \text{otherwise}
\end{cases}$$

No rejection of sampled values
Perturbation invariance

- Recall stochastic representation

\[ Z = \begin{cases} 
  X & \text{if } T \leq w(X) \\
  -X & \text{otherwise}
\end{cases} \]

- then \(|Z|\) is distributed like \(|X|\), write \(|Z| \overset{d}{=} |X|\)

- more generally: \(t(Z) \overset{d}{=} t(X)\) for any even \(t(\cdot)\)
  \(\Rightarrow\) property of perturbation (or modulation) invariance

- Example: if \(Z \sim SN(\alpha)\), then \(Z^2 \sim \chi^2_1\)
<table>
<thead>
<tr>
<th>Prólogos</th>
<th>Básis ($d=1$)</th>
<th>Plus ($d = 1$)</th>
<th>Data</th>
<th>Básis ($d ≥ 1$)</th>
<th>Ultra</th>
<th>Ω</th>
</tr>
</thead>
</table>

**Plus ($d = 1$)**
More on SN: other stochastic representations

Representation by conditioning/selection
this was how we introduced the SN distribution

Additive representation
- If $U_0, U_1$ are independent $N(0, 1)$ variables, then
  \[ Z = \sqrt{1 - \delta^2} U_0 + \delta |U_1| \sim SN(\alpha) \]
  much used to develop EM-type algorithms

Representation via minima/maxima
- assume $(X, Y)$ is bivariate standard Normal with
  $\text{corr}\{X, Y\} = \rho$
- write $\alpha = \sqrt{(1 - \rho)/(1 + \rho)}$
- then $\max(X, Y) \sim SN(\alpha)$ and $\min(X, Y) \sim SN(-\alpha)$
More on SN: some formal properties

- Moment generating function has a simple expression:
  \[ M(t) = 2 \exp\left(\frac{1}{2}t^2\right) \Phi(\delta t) \]

  \[ \Rightarrow \text{can compute moments} \]
  e.g. \[ \mathbb{E}\{Z\} = \sqrt{\frac{2}{\pi}} \delta = \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1 + \alpha^2}} \]
  (only odd moments are necessary)

  \[ \Rightarrow \text{derive further properties} \]
  e.g. if \( Z \sim SN(\alpha) \perp U \sim N(0, 1) \), \( Z + U \sim \sqrt{2} \times SN(\tilde{\alpha}) \)

- Distribution function has a tractable expression
consider ratio of SN vs N tails:

\[
\text{ratio}(x) = \frac{2 \varphi(x) \Phi(\alpha x)}{\varphi(x)} \quad \text{as } x \to \pm \infty
\]

if \( \alpha > 0 \),

\[
\text{ratio}(x) = 2 \Phi(\alpha x) \rightarrow \begin{cases} 
2 & \text{if } x \to +\infty \\
0 & \text{if } x \to -\infty 
\end{cases}
\]

if \( \alpha < 0 \), just swap \( \pm \infty \)

Implication:

tails decay either at the same rate of \( N(0, 1) \) or faster

Same conclusion if SN density is replaced by another one like

\[
f(x) = 2 \varphi(x) G_0(\alpha x)
\]
Thick tails

- In many situations we need **thicker-than-normal tails** (occasionally need thinner-than-normal tails)
- This feature cannot be achieved by perturbation of $\mathcal{N}(0, 1)$
- We must start from a **baseline density** $f_0$ in
  \[
  f(x) = 2 \, f_0(x) \, G_0\{w(x)\}
  \]
  which already has thick tails
- Many possible options
- Preference for those where $f_0$ allows a **tail-regulation parameter**
Skew-$t$ (ST) distribution – genesis

- A good choice for $f_0$ is the Student's $t$ density: $t(x; \nu)$, $\nu > 0$
- Even then, still many possible options, such as the ‘linear form’
  
  $$2 \, t(x; \nu) \, T(\alpha \, x; \nu)$$

- There are strong reasons for picking up another option
- Recall origin of classical Student's $t$:

  $$Z \sim N(0, 1) \perp W_\nu \sim \chi^2_\nu \quad \Rightarrow \quad \frac{Z}{\sqrt{W_\nu/\nu}} \sim t(x; \nu)$$

- Use the the same construction with $Z \sim SN(\alpha)$
  
  $$\Rightarrow \quad \text{obtain the } \text{ST}(\alpha, \nu) \text{ distribution}$$

- Note: the link with the classical $t(x; \nu)$ is not the only reason
- Beware: in literature various other proposals named ‘skew-$t$’
Skew-\(t\) (ST) distribution – a closer look

- Algebraic work leads to \(ST(\alpha, \nu)\) density function:
  \[
  2 \, t(x; \nu) \, T\left(\alpha \sqrt{\frac{\nu + 1}{\nu + x^2}}; \nu + 1\right)
  \]

- \(Z \sim ST(\alpha, \nu) \implies Z^2 \sim F(1, \nu)\)
- \(m\)-th order moment exist if \(m < \nu\), like regular \(t\)
- explicit expressions available up to \(m = 4\)
  (if necessary, higher moments could be worked out)
- a very wide range of \(\gamma_1\) (skewness) and \(\gamma_2\) (kurtosis)
  \(-\infty < \gamma_1 < \infty, \quad 0 \leq \gamma_2 < \infty\) (but no \(\gamma_2 < 0\))
- widely flexible shape, well-suited for data fitting
  (when complemented with location and scale parameters)
- as \(\nu \to \infty\), density \(ST(\alpha, \nu) \to SN(\alpha)\)
Skew-\( t \) (ST) distribution – examples of density

\[ \alpha = 3 \]

\[ \nu = 0.5 \]

\[ \nu = 1 \]

\[ \nu = 2 \]

\[ \nu = 5 \]

\[ \nu = 20 \]
Data
Location and scale parameter

- Let $Z$ be a SN or ST or something-of-the-kind random variable.
- For applied work, introduce location and scale parameters:
  \[ Y = \xi + \omega Z, \quad \xi \in \mathbb{R}, \quad \omega \in \mathbb{R}^+ \]
- Correspondingly extend our notation to
  \[ Y \sim SN(\xi, \omega, \alpha) \text{ and } Y \sim ST(\xi, \omega, \alpha, \nu) \]
- Note: $\xi$ is not the mean, $\omega$ is not the standard deviation,
  (this is why we do not use classical $\mu, \sigma$ symbols)
Fitting a SN distribution

- Start from simple case of i.i.d. observations \( y = (y_1, \ldots, y_n) \)
- Log-likelihood for SN:

\[
\log L(\xi, \omega, \alpha) = \text{constant} - \frac{1}{2} n \log \omega - \frac{1}{2} \sum_i z_i^2 + \\
+ \sum_i \log \Phi(\alpha z_i)
\]

having set \( z_i = (y_i - \xi)/\omega \)

- In a regression model, location depends on covariates \( x_i \), typically in a linear form:

\[
\xi_i = x_i^\top \beta \quad x_i, \beta \in \mathbb{R}^p, \quad i = 1, \ldots, n
\]

- Log-likelihood \( \log L(\beta, \omega, \alpha) \) is as before, except that now

\[
z_i = (y_i - \xi_i)/\omega = (y_i - x_i^\top \beta)/\omega
\]
Illustration: fitting SN to phenols content in Barolo wine

\[ n = 59 \]
\[ \hat{\gamma}_1 = 0.8 \]
\[ \frac{\hat{\gamma}_1}{\text{std.err.}} = 2.5 \]
Illustration: graphical diagnostics of SN fitting

recall: \[ Z^2 = (Y - \xi)^2 / \omega^2 \sim \chi_1^2 \]

approx: \[ \hat{Z}^2 = (Y - \hat{\xi})^2 / \hat{\omega}^2 \sim \chi_1^2 \]

QQ-plot: \[ \hat{Z}^2_{(i)} \text{ vs } \chi_1^2 \text{ quantiles} \]

with ST: replace \( \chi_1^2 \) with \( F(1, \hat{\nu}) \)
SN log-likelihood: some unusual aspects

Two sort of noteworthy phenomena

- ‘Transient’ sort of occasional events
  Usually with small $n$, sporadic if $n$ beyond a few dozens
  Similar behaviour fairly common also with other models
    - multiple local maxima
    - $\max \log L$ occurs at $\alpha \to \pm \infty$

- ‘Persistent (but local)’ behaviour:
  that is, for all samples, but only at $\alpha = 0$
    - stationarity of $\log L$ at point $\alpha = 0$
    - correspondingly, singularity of the information matrix
SN log-likelihood: stationarity of log $L$ at $\alpha = 0$

deviance (LRT) : $D(\theta) = 2 \{\log L(\hat{\theta}) - \log L(\theta)\}$
profile deviance : $D(\theta) = 2 \{\log L(\hat{\theta}, \hat{\psi}) - \log L(\theta, \hat{\psi}(\theta))\}$
The twists of log $L$ at $\alpha = 0$ can be fixed by switching from ‘direct’ (DP) to ‘centred parameterization’ (DP)

Conceptually, we re-parameterize as

$$Y = \xi + \omega Z = \mu + \sigma Z_0$$

via the ‘centred variable’

$$Z_0 = (Z - \mathbb{E}\{Z\})/\text{std.dev.}(Z)$$

$\text{CP} = (\mu, \sigma, \gamma_1)$

In parallel, CP avoids singularity of the information matrix

Importantly, CP is easier to interpret than DP
SN log-likelihood: using CP with the Barolo data

deviance (LRT) : \[ D(\theta) = 2 \left\{ \log L(\hat{\theta}) - \log L(\theta) \right\} \]

profile deviance : \[ D(\theta) = 2 \left\{ \log L(\hat{\theta}, \hat{\psi}) - \log L(\theta, \hat{\psi}(\theta)) \right\} \]
ST log \( L \)

- With ST model **no stationarity** of log \( L \) at \( \alpha = 0 \)
- hence **no singularity** of information matrix at \( \alpha = 0 \)
- in fact, these issues are specific ‘only’ of \( \varphi \) baseline
- still CP useful for easier interpretability
Prólogos Básis ($d=1$)

Plus ($d = 1$)

Data

Básis ($d \geq 1$)

Ultra $\Omega$

Básis ($d \geq 1$)
Multivariate skew-normal distribution: genesis

- SN was constructed from bivariate standard normal \((X, W)\) as
  \[
  Z = (X \mid W \geq 0)
  \]

- Now start from \((d + 1)\)-dimensional Normal with std marginal
  \[
  \begin{pmatrix}
  X \\
  W
  \end{pmatrix}
  \sim
  N_{d+1}(0, \bar{\Sigma})
  \]

where \(\bar{\Sigma}\) is a correlation matrix

\[
\bar{\Sigma} = \begin{pmatrix}
\bar{\Omega} & \delta \\
\delta^\top & 1
\end{pmatrix}
\]

- and then use the same conditioning process: \(Z = (X \mid W \geq 0)\)
  except that now \(X\) is \(d\)-dimensional
Multivariate SN – illustration of genesis
Multivariate SN — basic formal facts

- If $Z = (X|W \geq 0)$, its density function turns out to be:
  \[
  2 \varphi_d(x; \bar{\Omega}) \Phi(\alpha^\top x), \quad x \in \mathbb{R}^d,
  \]
  where $\varphi_d(x; V)$ is $N_d(0, V)$ density and
  \[
  \alpha = \left(1 - \delta^\top \bar{\Omega}^{-1} \delta\right)^{-1/2} \bar{\Omega}^{-1} \delta \quad \in \mathbb{R}^d
  \]

- Moment generating function has a simple expression:
  \[
  M(t) = 2 \exp\left(\frac{1}{2} t^\top \bar{\Omega}^\top t\right) \Phi(\delta^\top t)
  \]
  \rightarrow can compute moments, e.g. $\mathbb{E}\{Z\} = \sqrt{2/\pi} \delta$
  \rightarrow derive further properties

- Additive representation extends to multivariate SN:
  \[
  Z = (I_d - \text{diag}(\delta)^2)^{1/2} U_0 + \delta |U_1|
  \]
  where $U_0 \sim N_d(0, \Psi) \perp U_1 \sim N(0, 1)$. 
Multivariate SN — include location and scale

- Start from \( Z = (Z_1, \ldots, Z_d)^\top \) with density \( 2 \varphi_d(x; \tilde{\Omega}) \Phi(\alpha^\top x) \)
- Introduce location and scale:

\[
\begin{pmatrix}
Y_1 \\
\vdots \\
Y_d
\end{pmatrix}
= \begin{pmatrix}
\xi_1 \\
\vdots \\
\xi_d
\end{pmatrix} + \begin{pmatrix}
\omega_1 & 0 \\
0 & \ddots \\
0 & \omega_d
\end{pmatrix} \begin{pmatrix}
Z_1 \\
\vdots \\
Z_d
\end{pmatrix}
\]

- Write more compactly

\[ Y = \xi + \omega Z \]

where \( \omega = \text{diag}(\omega_1, \ldots, \omega_d) \)
- Notation: \( Y \sim \text{SN}_d(\xi, \Omega, \alpha) \) where \( \Omega = \omega \tilde{\Omega} \omega \)
- Density at \( x \in \mathbb{R}^d \):

\[
2 \varphi_d(x - \xi; \Omega) \Phi(\alpha \omega^{-1}(x - \xi))
\]
Recall elliptical families

- Recall continuous **elliptically contoured (EC)** distributions
- Density constant on ellipsoids:
  \[
  f(x) = \frac{c_d}{(\det \Sigma)^{1/2}} g_d \left( (x - \mu)\top \Sigma^{-1} (x - \mu) \right), \quad x \in \mathbb{R}^d
  \]
- Notation: \( X \sim \text{EC}_d(\mu, \Sigma, g_d) \)
- Density is **centrally symmetric** about \( \mu \): \( f(x - \mu) = f(\mu - x) \)
- Extends the normal distribution which corresponds to
  \[
  g_d(u) = \exp(-u/2)
  \]
- The key aspect is that the EC family encompasses many others
- and it still preserves various properties of normal distribution:
  - family closed under marginalization
  - family closed under conditioning
  - conditional mean is linear function of the conditioning variables
- An interesting case is the multivariate Student’s t:
  \[
  g_d(u) = (1 + u/\nu)^{-(d+\nu)/2}
  \]
Skew-elliptical distributions

- Start from
  \[
  \begin{pmatrix} X \\ W \end{pmatrix} \sim \text{EC}_{d+1}(0, \bar{\Sigma}, g_{d+1})
  \]

- and apply the ‘usual’ conditioning (or selection) process:
  \[
  Z = (X|W > 0)
  \]

- Introduce location and scale: \( Y = \xi + \omega Z \)

- Terminology: \( Y \) and \( Z \) have skew-elliptical distribution (SEC)

- If \( (X, W) \) is normal, reproduce \( Y \sim \text{SN}_d(\xi, \Omega, \alpha) \)

- Another noteworthy case with \( (X, W) \sim t_{d+1}(0, \bar{\Sigma}, \nu) \):
  \[
  Y \sim \text{ST}_d(\xi, \Omega, \alpha, \nu)
  \]

- density of normalized r.v. \( Z \sim \text{ST}_d(0, \bar{\Omega}, \alpha, \nu) \):
  \[
  2 : t_d(z; \bar{\Omega}) T \left( \alpha^T z \sqrt{\frac{\nu + d}{\nu + z^T \bar{\Omega}^{-1} z}} ; \nu + d \right), \quad z \in \mathbb{R}^d
  \]
A general result

Lemma (Multivariate version)

If $f_0$ is a PDF on $\mathbb{R}^d$ and $G_0$ a continuous CDF on $\mathbb{R}$, both symmetric about 0, then

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

is a proper density function for any odd function $w(\cdot)$ on $\mathbb{R}^d$.

Proof: a simple extension of the univariate version.

Notes:
(1) $f_0$ symmetric on $\mathbb{R}^d$ means $f_0(x) = f_0(-x)$ for all $x \in \mathbb{R}^d$
(2) $w$ odd function on $\mathbb{R}^d$ means $w(-x) = -w(x)$ for all $x \in \mathbb{R}^d$. 
A general result — comments

- Both $\mathcal{SN}_d$ and $\mathcal{ST}_d$ have density like $f(x)$ in the lemma
- Can show that all SEC distributions have this structure with ‘baseline density’ $f_0$ of elliptical type
- But the lemma allows $f_0$ to be non-elliptical and $G_0$ can be unrelated to $f_0$, unlike in SEC’s
- This modulation process can produce all sort of shapes, even quite bizarre ones, not just ‘skew’
- Next plots illustrate this point using

$$f_0 = \varphi_2, \quad G_0 = \Phi$$

$$w(x_1, x_2) = a_1 x_1 + a_2 x_2 + a_3 x_1^3 + a_4 x_2^3 + a_5 x_1^2 x_2 + a_6 x_1 x_2^2$$
Examples of modulated bivariate normal densities
Some formal properties of the general construction

\[ f(x) = 2f_0(x) \, G_0\{w(x)\}, \quad x \in \mathbb{R}^d \]

**Stochastic representation** If \( X \sim f_0 \perp T \sim G_0 \), then

\[ Z = \begin{cases} 
X & \text{if } T \leq w(X) \\
-X & \text{otherwise} 
\end{cases} \] has density \( f(\cdot) \)

**Perturbation (or modulation) invariance** Now holds multivariate:

\[ t(Z) \overset{d}{=} t(X) \]

for any even \( t(x) \), mapping \( \mathbb{R}^d \to \mathbb{R}^q \)

**Examples** If \( Y \sim SN_d(\xi, \Omega, \alpha) \) and \( V \sim ST_d(\xi, \Omega, \alpha, \nu) \), then

\[ (Y - \xi)^\top \Omega^{-1}(Y - \xi) \sim \chi^2_d \]
\[ (V - \xi)^\top \Omega^{-1}(V - \xi) \sim d \times F(d, \nu) \]

These facts are useful for model diagnostics.
Prólogos Básis (d=1) Plus (d = 1) Data Básis (d ≥ 1) Ultra
Many additional developments

- Many forms of generalization exist
- The more tractable case is the extended SN and alike: start from \((X, W) \sim N_2\) and take \((X|W \geq c)\) with \(c \in \mathbb{R}\)
- Important extension: \(m\)-dimensional conditioning variable \(W\) relatively tractable in normal context (Closed SN) to some extent also tractable in EC class
- General selection mechanism: replace \((\cdots | W \geq 0)\) by \((\cdots | W \in C)\) with \(C \subset \mathbb{R}^m\)
  (For general \(C\), difficult to find normalizing constant)
Use in statistical methods and applied areas

Two intersecting levels of work:

- Extensions of standard statistical methods
- Application in diverse fields, often with suitable methodological adaption of existing techniques

Many domains:

- classical areas of statistical methods, such as longitudinal data, factor analysis, item response analysis, . . .
- much impact especially in model-based clustering
- flexible distributions provide a route to robustness
- much work in finance, theoretical and empirical
- but also in environmental risk, medical statistics, econometrics, income distribution, data confidentiality, insurance, industrial statistics and reliability, cell biology, forestry, et cetera . . .
Prólogos Básis ($d=1)$ Plus ($d=1)$ Data Básis ($d \geq 1)$ Ultra $\Omega$
Any future?

- Formidable work has been deployed, but still room for progress
- Extension of standard statistical methods for more flexible models, with applications
- Further advances possible in the study of flexible distributions
  (a personal view presented in more specialized topic session)
A complete list of references would take many pages. An absolutely minimal list is:

- MG Genton (2004), edited volume, C&H/CRC
- R software: https://cran.r-project.org/package=sn