Skew-symmetric distributions in 45'
(that is, S^3=Skew-Symmetric Squeeze)

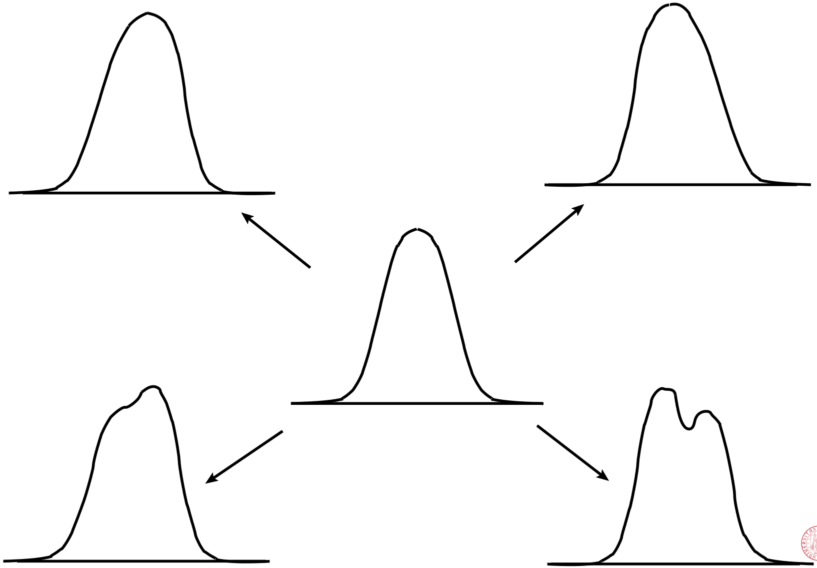
Adelchi Azzalini
University of Padua, Italy

BIRS at Banff, Canada, May 2013

Perturbation of symmetry: general aspects

Perturbation of symmetry: general aspects

Departure from normality



Perturbation of symmetry

- formalize the idea of perturbation of a symmetric 'base density'
- start with dimension $d = 1$
- 'perturbation' (or 'modulation') of symmetric pdf $f_0(x)$ as

$$f(x) = 2\, f_0(x) G_0\{w(x)\}$$

  where (1) $w(-x) = -w(x)$ and
  (2) G_0 is continuous cdf, $G_0(-x) = 1 - G_0(x)$
- Proof that integrates to 1, extraordinarily simple:
  if $T \sim G_0$ and $Z_0 \sim f_0$, independent, then

$$\begin{aligned} \tfrac{1}{2} &= \mathbb{P}\{T - w(Z_0) \leq 0\} = \mathbb{E}\{\mathbb{P}\{T \leq w(x) | Z_0 = x\}\} \\ &= \int_{\mathbb{R}} G_0\{w(x)\}\, f_0(x)\, \mathrm{d}x \qquad \text{QED} \end{aligned}$$

Perturbation of symmetry (ctd)

$$f(x) = 2\, f_0(x)\, \underbrace{G_0\{w(x)\}}_{G(x)}$$

- then

$$G(x) \ge 0, \qquad G(x) + G(-x) = 1$$

- any G of this form produces a valid pdf

$$f(x) = 2\, f_0(x)\, G(x)$$

- the two forms are essentially equivalent
- if $w(x) \equiv 0$, i.e. $G(x) \equiv \frac{1}{2}$, then $f = f_0$

Multivariate version

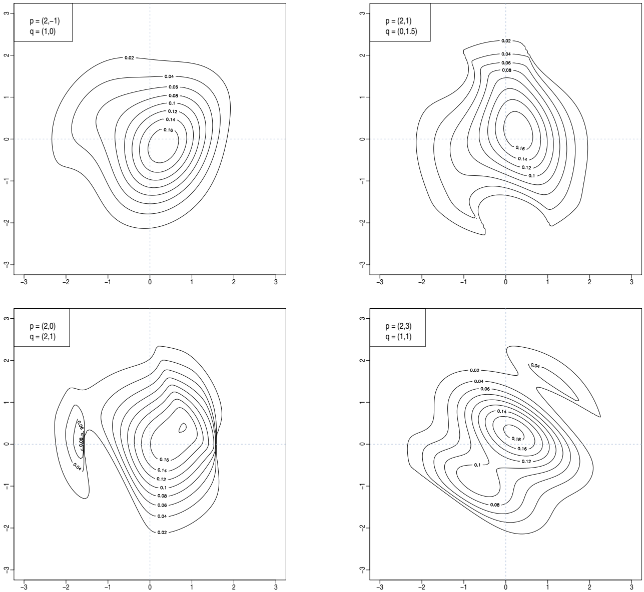
$$f(x) = 2\, f_0(x)\, \underbrace{G_0\{w(x)\}}_{G(x)} \qquad x \in \mathbb{R}^d$$

- $f_0(x) = f_0(-x)$ for $x \in \mathbb{R}^d$
- w is real-valued, with $w(-x) = -w(x)$
- the rest as before

Example with $d = 2$

- $f_0(x)$ is the $\mathrm{N}(0, I_2)$ density
- G_0 is standard logistic cdf
- $$w(x) = \frac{\sin(p_1\,x_1 + p_2\,x_2)}{1 + \cos(q_1\,x_1 + q_2\,x_2)}, \qquad x = (x_1, x_2) \in \mathbb{R}^2$$

Example with $d = 2$ (ctd)



Stochastic representations

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

- If $Z \sim f$, the argument of the proof indicates that

$$Z \stackrel{d}{=} (Z_0 | T \le w(Z_0))$$

- also

$$Z = S_{Z_0}\, Z_0, \qquad S_{Z_0} = \begin{cases} +1 & \text{w.p. } G(Z_0) \\ -1 & \text{w.p. } G(-Z_0) \end{cases}$$

Perturbation invariance

$$Z = S_{Z_0}\, Z_0, \qquad S_{Z_0} = \begin{cases} +1 & \text{w.p. } G(Z_0) \\ -1 & \text{w.p. } G(-Z_0) \end{cases}$$

- Corollary: property of perturbation invariance

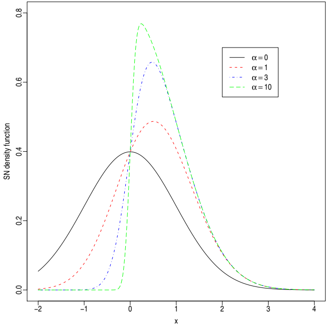
$$\text{for any even function } t(\cdot) \quad \Longrightarrow \quad t(Z) \overset{d}{=} t(Z_0)$$

- In the example, $\|Z\|^2 \sim \chi^2_2$
- Note: property holds for multi-valued functions $t(\cdot)$

A noteworthy case: the skew-normal distribution

A noteworthy case: the skew-normal distribution

The skew-normal distribution (SN), case $d = 1$



- $f(x) = 2\,\phi(x)\,\Phi(\alpha x), \qquad \alpha \in \mathbb{R}$
- $\alpha = 0$ leads back to usual Normal
- if $Z \sim \mathrm{SN}(\alpha)$, then $-Z \sim \mathrm{SN}(-\alpha)$
- $Z^2 \sim \chi^2_1$
- for practical work, add location and scale: $Y = \xi + \omega Z$, $\omega > 0$

The multivariate SN distribution

- 'Normalized' form (no location and scale):

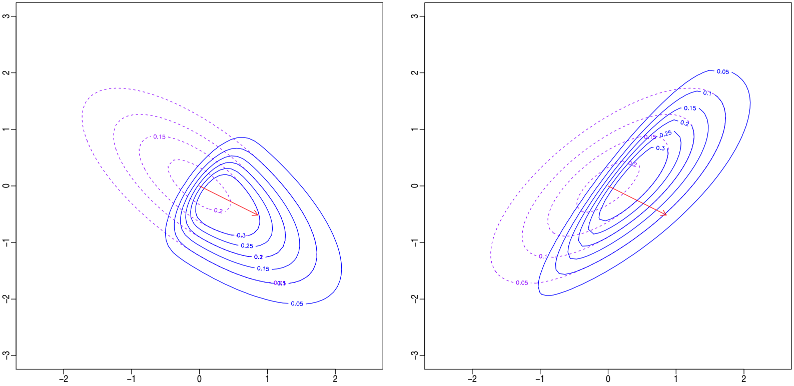
$$f(x) = 2\phi_d(x; \bar{\Omega})\, \Phi(\alpha^\top x), \qquad x \in \mathbb{R}^d$$

  for some correlation matrix $\bar{\Omega}$ and shape $\alpha \in \mathbb{R}^d$
- MGF: for an appropriate $\delta = \delta(\alpha, \bar{\Omega})$,

$$M(t) = 2\, \exp(\tfrac{1}{2} t^\top \bar{\Omega} t)\, \Phi(\delta^\top t)$$

- distribution of a quadratic form $Z^\top A Z$ as for $\mathrm{N}_d(0, \bar{\Omega})$
- for practical work, add location and scale: $Y = \xi + \omega Z$, where $\xi \in \mathbb{R}^d$ and $\omega = \mathrm{diag}(\omega_1, \ldots, \omega_d) > 0$

Multivariate SN density



Stochastic representations of SN

- representation by conditioning: can transform (Z_0, T) into

$$\begin{pmatrix} Z_0 \\ Z_1 \end{pmatrix} \sim N_{d+1}\left(0, \begin{pmatrix} \bar{\Omega} & \delta \\ \delta^\top & 1 \end{pmatrix}\right)$$

  and set

$$Z \stackrel{d}{=} (Z_0 | Z_1 > 0)$$

- additive representation: another manipulation leads to

$$Z = \left(I_d - \mathrm{diag}(\delta)^2\right)^{1/2} U_0 + \delta |U_1|$$

  for independent $U_0 \sim \mathrm{N}_d$ and $U_1 \sim \mathrm{N}(0,1)$
- representation via maxima/minima

Adjustable tails and skew-elliptical distributions

Adjustable tails and skew-elliptical distributions

Heavy and adjustable tails

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

- the mechanism can make tails thinner, but not thicker
- to handle heavy tails, start from base f_0 with heavy tails
- even better consider f_0 with adjustable tails

From EC to SEC

- Elliptically contoured (EC) densities: for a suitable $g(\cdot)$,

$$f_0(x) = \frac{k_d}{\det(\bar\Omega)^{1/2}}\, g(x^\top \bar\Omega^{-1} x), \qquad x \in \mathbb{R}^d$$

  denoted $EC_d(0, \bar\Omega, g)$

- A natural option for perturbation is

$$f(x) = 2\, f_0(x)\, G(x)$$

- ... but consider instead

$$\begin{pmatrix} Z_0 \\ Z_1 \end{pmatrix} \sim EC_{d+1}\left(0, \begin{pmatrix} \bar\Omega & \delta \\ \delta^\top & 1 \end{pmatrix}, g\right)$$

  followed by

$$Z \overset{d}{=} (Z_0 | Z_1 > 0), \quad \text{called SEC}$$

- the distribution of Z is of type $f(x)$. Note: not *vice versa*

A noteworthy case: the skew-t distribution (ST)

- Multivariate Student's t: genesis is

$$U/\sqrt{V}$$

where $U \sim \mathrm{N}_d$ and $V \sim \chi^2_\nu/\nu$ are independent

- Multivariate skew-t:

$$Z' = Z/\sqrt{V}$$

where $Z \sim \mathrm{SN}_d$ with shape α

- It is equivalent to start from $Y \sim SEC_{d+1}$ of Student's t type, and consider

$$Z' = (Y_{1:d} | Y_{d+1} > 0)$$

- Here α regulates skewness, ν regulates tail thickness

Further generalizations

Further generalizations

So-called 'extended' forms

- 'extended' form: non-odd $w(x)$,
- e.g. in SN case $w(x) = \alpha_0 + \alpha^\top x$,
- normalizing constant no longer 2,
  must be computed afresh for any case
- property of perturbation invariance vanishes
- in some cases, subject-matter motivation

Multiple latent variables/constraints

- start from $(d+m)$-dimensional variate (Z_0, Z_1) and consider

$$Z \stackrel{d}{=} (Z_0 | Z_1 \in C), \qquad C \subset \mathbb{R}^m$$

- density is

$$f(x) = f_0(x) \, \frac{\mathbb{P}\{Z_1 \in C | Z_0 = x\}}{\mathbb{P}\{Z_1 \in C\}}$$

- special focus on case where f_0 is symmetric
- extremely general in principle,
  but computation of the two probabilities often problematic
- beware of overparameterization

Statistical aspects

Statistical aspects

Statistics is harder than probability

- as a broad rule, the statistical side is less smooth than the probability side
- some formal issues (with proposed solutions)
- less formal but equally important issues
- Note: these are aspects with space for improvement, it does not mean we are helpless

Classical formal issues

- refer to parameter set (ξ, ω, α) or alike, for simplicity
- for SN (and some other cases) Info matrix singular at $\alpha = 0$;
  can be tackled via appropriate re-parameterization;
  proposals exist, but not unique
- for finite samples, $\mathbb{P}\{\text{MLE}(\alpha) = \infty\} > 0$
  can be avoided by penalized likelihood and/or prior;
  proposals exist, but not unique

Less formal issues but equally important

- what is the 'optimal' parameterization for inference?
  hassle-free *and* meaningful

- highly flexible distributions can be constructed:
  how much flexible can we be in practice?
  how to combine flexibility with meaningful parameterization?

General references

- M. G. Genton (2004), edited book
- A. Azzalini (2005, SJS) review paper + discussion with MGG
- A. Azzalini & A. Capitanio, forthcoming book