The skew-normal distribution and related multivariate families

Adelchi Azzalini

Università di Padova

NordStat 2004: Jyväskylä, 6–10 June 2004

Overview

- Background & motivation
- Skew-normal distribution (scalar/multivariate case, extensions)
- More general families
  - skew-elliptical families, and some special cases
  - a general formulation
  - ‘semi-parametric’ families
- Connection with other areas
  - selective sampling
  - stochastic frontier models
  - forms of robust likelihood
  - applications in mathematical geology, finance, etc.
- Some open problems

Some general remarks

Motivation

- Context:
  parametric families of pdf's for continuous variates
- Want to build flexible parametric classes
  - extend normal family, in scalar and multivariate case
  - construct families of "arbitrary" flexibility
  - bridge between parametric and non-parametric approach
- normal class is central to multivariate data analysis,
  . . . but
  "who has ever seen a multivariate normal sample?"
  [Barnett (1979)]

An old history

- (bivariate) Pearson family
- marginal transformation approach, including:
  - Johnson family
  - $g-$ and $h-$transformations
- Edgeworth expansions
- mixtures and compound
- some approaches are specific of $d-$dimensional case:
  - copulae
  - specification via conditional distributions
- others...

[Reviews: Pretorious (1930), Johnson & Kotz (1970–72), Hutchinson & Lay (1990), Joe (1997), Patil et al. (1976), Kotz et al. (2000→), ...]

Desiderata

- high flexibility in shape
  with wide range for indices of skewness and kurtosis
- mathematically tractable:
  - simple form of density, moments, cdf, etc.
  - closed under marginalization and conditioning
  - retain nice properties of original family
  - parameters should be directly linked to some aspects of pdf
- association with ‘physical’ mechanism(s)
  - to support statistical modeling
  - for random number generation
- allow easy and reliable inference

Expanding the Normal family: an approach

A simple yet useful result

Lemma (1D case)

Suppose f_0 and G' are pdf's on $\mathbb{R}$ symmetric about 0, then

$$f(z) = 2\, f_0(z)\, G\{w(z)\}$$

is a density for any odd function $w(\cdot)$.

Proof. If $Y \sim f_0$ and $X \sim G'$ are independent,

$$\begin{aligned}\tfrac{1}{2} &= \mathbb{P}\{X - w(Y) \le 0\} = \mathbb{E}_Y\{\mathbb{P}\{X - w(Y) \le 0 | Y = z\}\} \\ &= \int_{\mathbb{R}} G\{w(z)\}\, f_0(z)\, \mathrm{d}z.\end{aligned}$$

[Azzalini (SJS, 1985) for case $w(z) = \alpha\, z$]

The skew-normal distribution ($d = 1$)

Use lemma to ‘skew’ the N(0,1) density:

$$\phi(z;\alpha) = 2\, \underbrace{\phi(z)}_{f_0(z)}\ \underbrace{\Phi(\alpha\, z)}_{G\{w(z)\}}$$

for $z,\ \alpha \in \mathbb{R}$.   Write

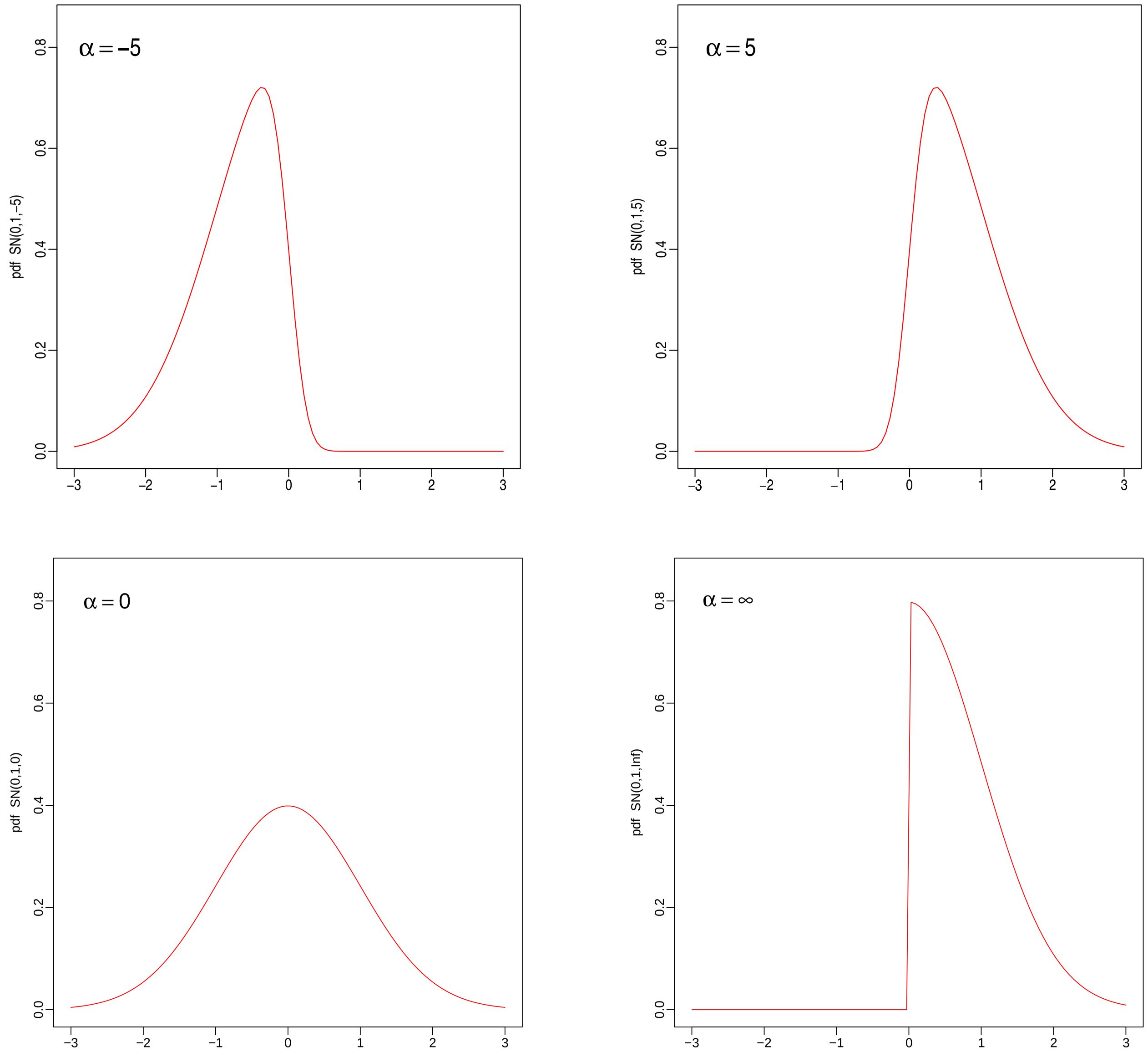
$$Z \sim \mathrm{SN}(\alpha)$$

Some basic facts:

- if $\alpha = 0$, obtain N(0,1)
- if $Z \sim SN(\alpha)$, then $-Z \sim SN(-\alpha)$
- $1 - \Phi(-z;\alpha) = \Phi(z;-\alpha)$ for its cdf

‘Extended’ version: $\phi(z)\, \Phi(\alpha_0 + \alpha\, z)/\Phi(\alpha_0/\sqrt{1+\alpha^2})$

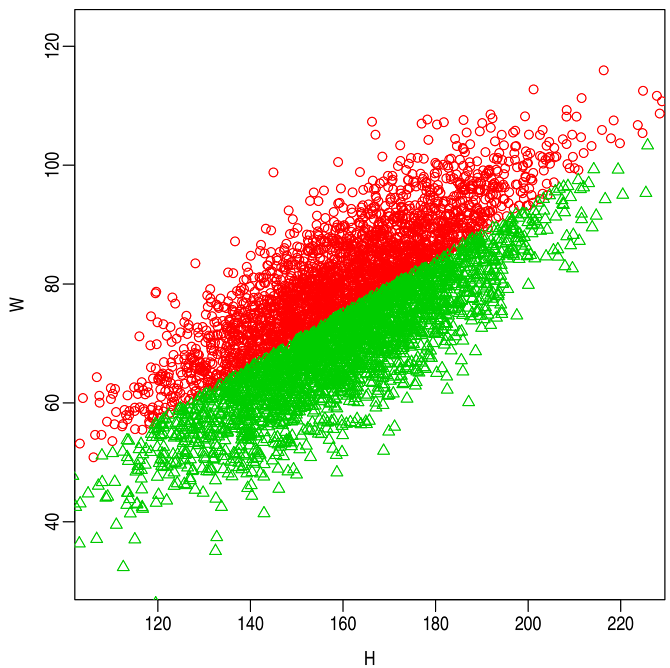
Some SN pdf's ($d = 1$)

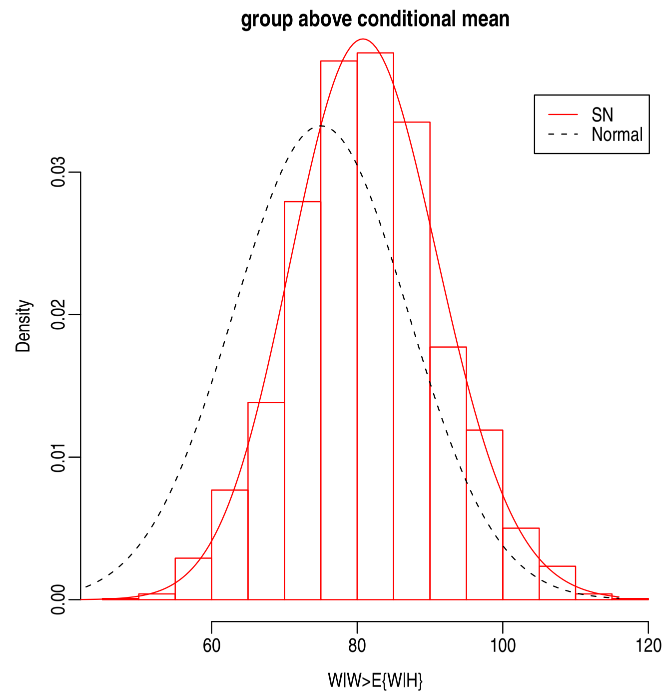


Forms of genesis

- Conditioning:
  if $(U_0, U_1) \sim N_2$ with N(0,1) marginals and correlation ρ, then
  $$U_0 | U_1 > 0 \sim \mathrm{SN}\left(\rho/\sqrt{1-\rho^2}\right)$$
  For $U_0 | U_1 > \tau \sim ESN$
- Sum of Normal+HN:
  if V_0, V_1 iid $N(0,1)$, then
  $$\delta\,|V_0| + \sqrt{1-\delta^2}\,V_1 \sim \mathrm{SN}\left(\delta/\sqrt{1-\delta^2}\right)$$
- Order statistics:
  $$\max(U_0, U_1) \sim \mathrm{SN}\left(\sqrt{(1-\rho)/(1+\rho)}\right)$$

A “constructive” example





$$\begin{pmatrix} H \\ W \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\right)$$

separate out group with W ‘above average’ (given H):

$$W|U>0, \quad U = W - \mathbb{E}\{W|H\}$$

$$W \sim N(\mu_2, V_{22})$$

$$(W|U>0) \sim \mathrm{SN}(\mu_2, V_{22}, \delta/\sqrt{1-\delta^2})$$

$$\delta = \sqrt{1-\rho^2}$$

Some nice properties

- $\mathbb{P}\{Z < z\} = \Phi(z) - 2\,T(z,\alpha)$, $\quad T$ is Owen's function
- $Z^2 \sim \chi^2_1$, independent of α
- hence even moments are the same of $N(0,1)$
- odd moments from N*HN representation
- or use $\mathbb{E}\left\{e^{tZ}\right\} = 2\,e^{t^2/2}\,\Phi(\delta t)$, $\quad \delta = \alpha/\sqrt{1+\alpha^2}$
- $SN(\alpha)^\star N(0,1) \sim$ (scaled) $\mathrm{SN}(\alpha/\sqrt{2+\alpha^2})$
- etc. . .

[Azzalini (1985), Henze (1986), Chiogna (1998)]

Early (independent) appearances

- Conditioning:
  Birnbaum (1950) obtains pdf of $U_1|U_0 > \tau$
  and some properties of first two moments.
  Motivation: selection process of personnel, in education, etc.
- Sum of Norm+HN:
  - in *Technom.* (1964), discussion with contributions by Weinstein, Lipow, Mantel, Wilkinson and Nelson
  - in Bayesian context, O'Hagan & Leonard (1976) obtain ESN from a two-stage prior construction
- Order statistics:
  Roberts (1966) obtains SN pdf, and its property

$$Z^2 \sim \chi_1^2$$

Other ways of skewing the normal pdf

- use previous lemma with $G \neq \Phi$,
  many options. . .
  [Nadarajah & Kotz (2003)]
- $$f(z) = \text{constant} \times \phi(z)\,\Phi(\alpha\, z)^n, \qquad n \in \mathbb{N}$$
  connnection with order statistics of N(0,1)
  [Balakrishnan (in discussion on Arnold & Beaver, 2002)]
- or take entirely different routes:
  two-pieces normal, etc.

Skewing non-normal distributions

$$f(z) = 2\, f_0(z)\, G(\alpha\, z), \qquad f_0(z) \neq \phi(z)$$

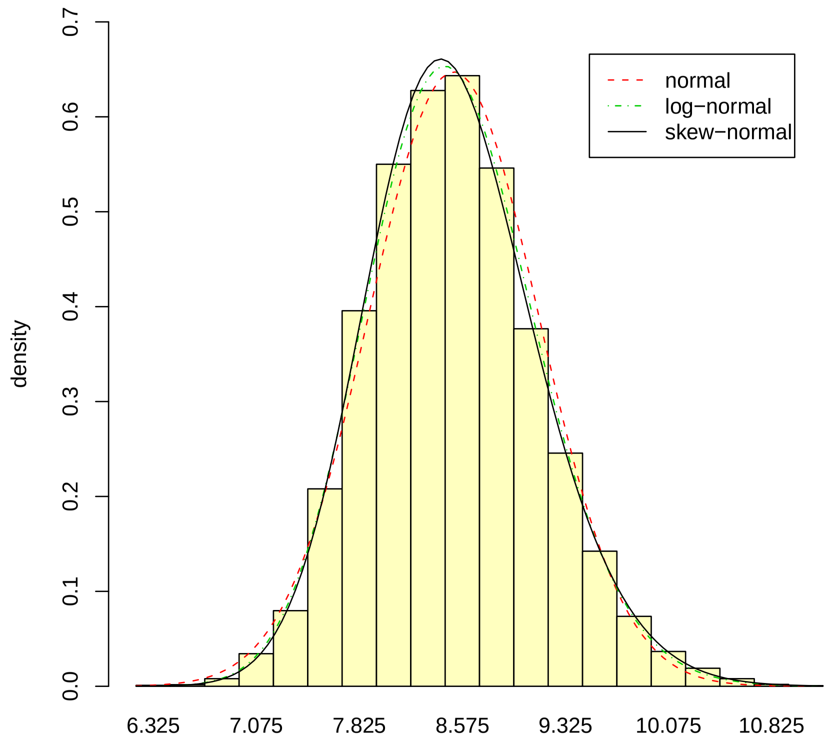
- f_0 exponential power dist'n (or GED or Normal$^{(r)}$) for heavy (or light) tails $\Rightarrow$ SEP

  [Azzalini (1986), DiCiccio & Monti (2004)]
- skew-Laplace [Balakrishnan and Ambalagaspitya (1994)]
- skew-Cauchy [Arnold & Beaver (2000b), include multivariate version]
- skew-logistic [Wahed & Ali (2000)]
- the wrapped SN on the circle (related to above form but not quite the same)

  [Pewsey (2000b, 2003)]

Statistical aspects

An old example: Johannsen's beans data



Distribution	X^2	d. f.	p-value
Normal	196.5	13	< 0.001
Log-normal	28.1	13	0.008
Edgeworth			
1st order	34.3	12	< 0.001
2nd order	14.9	11	0.19
Skew-normal	14.2	12	0.29

Note: data re-grouped from 20 into 16 bins

[Charlier (1931), Cramér (1946, p. 440–1)]

$$\hat{\gamma}_1 = 0.2809, \quad \hat{\gamma}_1\sqrt{n/6} = 12.56$$

Likelihood function is ‘regular’, but. . .

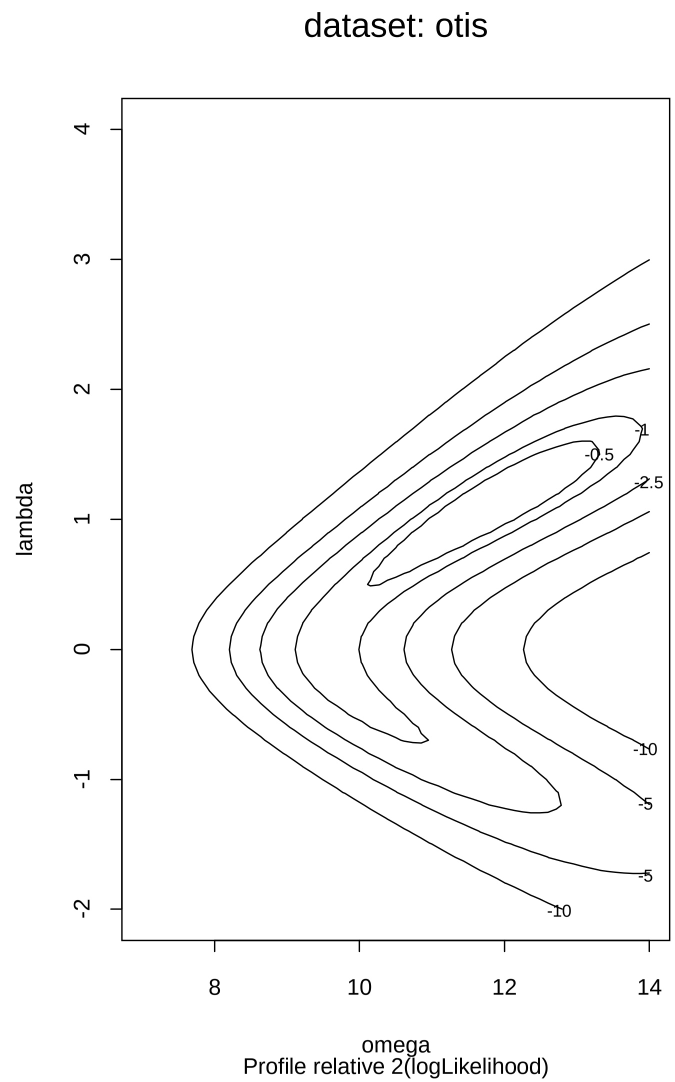
- at $\alpha = 0$:
  - ◇ always a stationarity point of $\log L$
  - ◇ Fisher expected information matrix becomes singular

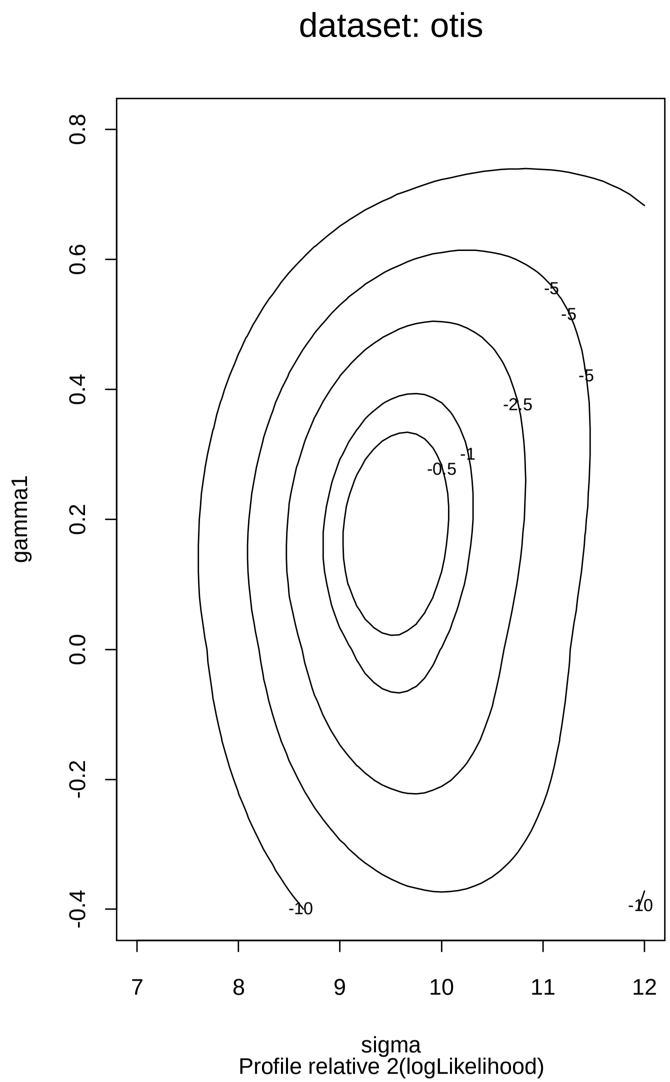
    [connection with non-standard cases of Rotnitzky et al. (2000)]

  Both can be avoided by suitable reparametrization
- MLE can occur at $\alpha = \pm\infty$ for ‘small n’
  some solutions proposed
- MLE can have more than one local maximum
- ‘small’ here is larger than elsewere

[Azzalini (1985), Liseo (1990), Azzalini & Capitanio (1999), Pewsey (2000a), Monti (2003), Sartori (2003+)]

An example of profile log-likelihood





use of ‘direct’ vs ‘centred’ parameterization

Bayesian inference

- ESN arises from two-stage prior

$$Y \sim N(\theta, \tau^2), \quad \theta \sim N(\mu, \sigma^2), \quad \mu \sim N_+(\mu_0, \sigma_0^2)$$

[O'Hagan & Leonard (1976)]

- Bayesian approach to estimation of parameters can avoid $\hat{\alpha} = \infty$.

[Liseo (1990), see also Liseo & Loperfido (2002, to appear)]

- linear Bayesian estimation for model

$$Y \sim N(\theta, 1), \quad \theta \sim \mathrm{SN}(\lambda)$$

[Mukhopadhyay & Vidakovic (1995)]

Multivariate SN distribution

Multivariate versions of 1D genesis

- Conditioning:

$$\begin{pmatrix} U_0 \\ U_1 \end{pmatrix} \sim N_{d+1}\left(0, \begin{pmatrix} 1 & \delta^\top \\ \delta & \bar{\Omega} \end{pmatrix}\right), \quad N(0,1) \text{ marginals}$$

$$(U_1 | U_0 > 0) \sim \mathrm{SN}_d(0, \bar{\Omega}, \alpha), \qquad \alpha = \text{function}(\delta, \bar{\Omega})$$

- Sum of Norm+HN:

$$V_0 \sim N(0,1), \quad V_1 \sim N_d(0, \Psi), \quad V_0 \perp\!\!\!\perp V_1$$

$$Z_j = \delta_j\,|V_0| + \sqrt{1-\delta_j^2}\,V_{1j}, \qquad j = 1, \ldots, d$$

$$Z \sim \mathrm{SN}_d(0, \bar{\Omega}, \alpha), \qquad (\bar{\Omega}, \alpha) = \text{function}(\delta, \Psi)$$

The two mechanisms generate the same class of pdf's

[Azzalini & Dalla Valle (1996)]

Multivariate SN distribution

‘Standard’ SN: write $Z \sim \mathrm{SN}_d(0, \bar{\Omega}, \alpha)$ if pdf at $z \in \mathbb{R}^d$ is

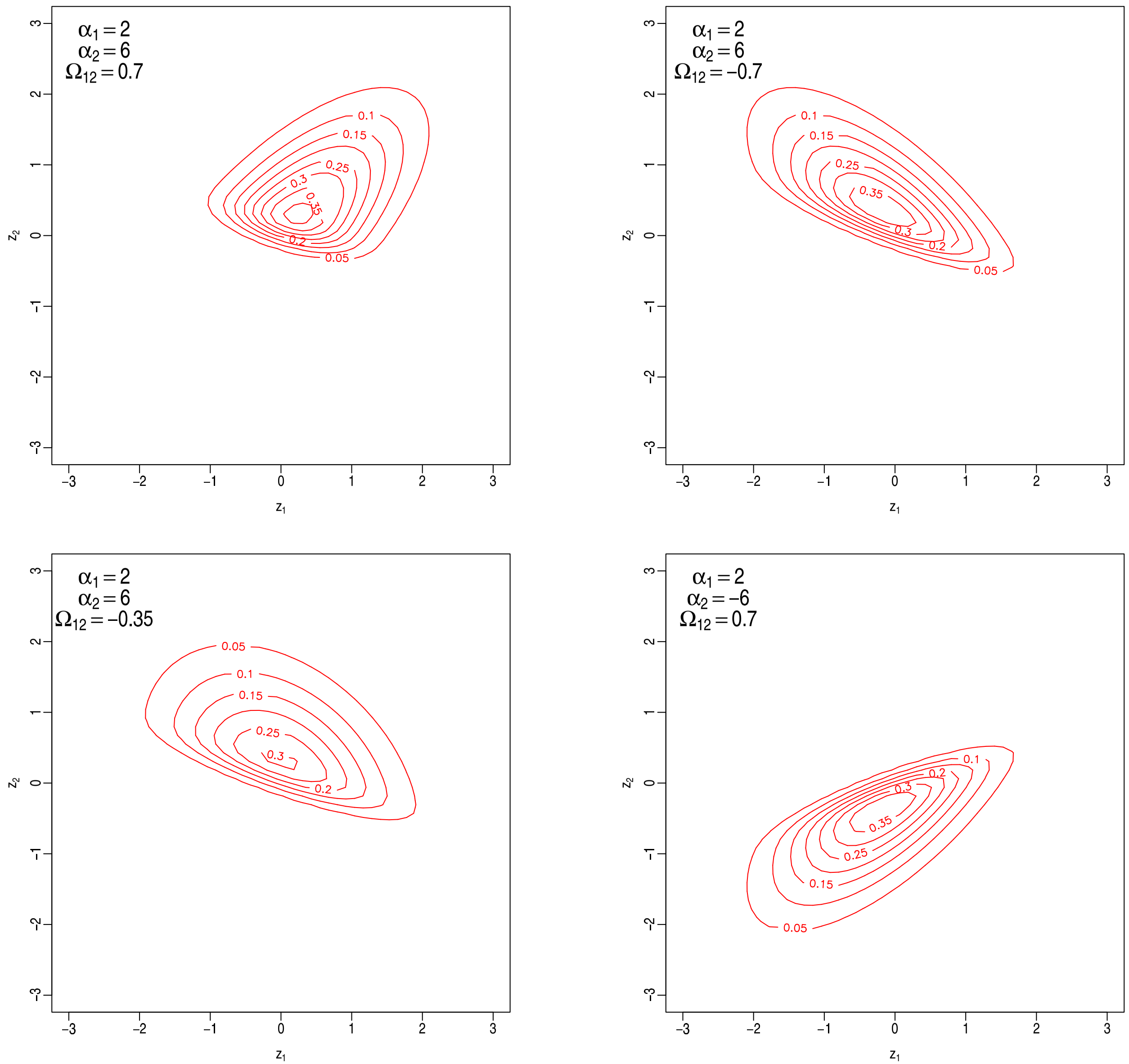
$$f(z) = 2\,\phi_d(z; \bar{\Omega})\,\Phi(\alpha^\top z)$$

$\alpha \in \mathbb{R}^d$. If $\alpha = 0$, get $N_d(0, \bar{\Omega})$.

Add location/scale parameters:

$$Y = \xi + \omega\,Z \sim \mathrm{SN}_d(\xi, \Omega, \alpha), \qquad \Omega = \omega\,\bar{\Omega}\,\omega$$

Some SN_2 pdf's



Some formal properties – I

- Parameters ξ, Ω, α are variation-independent
- Formal properties depend on

$$\delta = (1 + \alpha^\top \bar{\Omega} \alpha)^{-1/2} \bar{\Omega} \alpha$$

e. g.

$$\mathbb{E}\{Z\} = \sqrt{2/\pi}\, \delta = \mu, \qquad \text{var}\{Z\} = \bar{\Omega} - \mu\, \mu^\top$$

- moment generating function & cumulants:

$$\mathbb{E}\left\{e^{tZ}\right\} = 2\, \exp(\tfrac{1}{2} t^\top \bar{\Omega} t)\, \Phi(\delta^\top t)$$

$\Longrightarrow$ k-th order cumulant: constant $\times\, \delta_r\, \delta_s \cdots \delta_v$

$\Longrightarrow$ $\gamma_{1,d} \in (0, 0.9905), \qquad \gamma_{2,d} \in (0, 0.869)$

Some formal properties – II

- Random numbers: any of
  - ◇ $Z = \mathrm{sign}(U_0)\, U_1$
  - ◇ $Z = \delta\, |V_0|\, 1_d + (1-\delta^2)^{1/2}\, V_1$

  require a N_{d+1} variate
- cdf of SN_d computed via cdf of N_{d+1}, namely

$$\mathbb{P}\{Z < z\} = 2\,\mathbb{P}\{-U_0 < 0 \cap U_1 < z\}$$

- the class is closed under affine transformations

$$a + AY \sim \mathrm{SN}_p(a + A\xi, A\Omega A^\top, \alpha')$$

hence also closed under marginalisation

A characterisation property

A parallel of a well-known property of the Normal class.

Let X be a r. v. (in $\mathbb{R}^d$) with $\mathbb{E}\{X\} = \mu$ and var$\{X\} = \Sigma$.
Put $\Omega = \Sigma + \mu\,\mu^\top$. If

$$h^\top X \sim \mathrm{SN}(\alpha_h)$$

for any h such that $h^\top \Omega\, h = 1$, then $X \sim \mathrm{SN}_d$.

[Gupta & Huang (2002)]

Quadratic forms

- $(Y-\xi)^\top\Omega^{-1}(Y-\xi) \sim \chi^2_d$

  leads to Healy-type diagnostic plot
- if $X \sim N_d(0,\Omega)$ and $Y \sim \mathrm{SN}_d(0,\Omega,\alpha)$, then

$$X^\top A X \stackrel{d}{=} Y^\top A Y$$

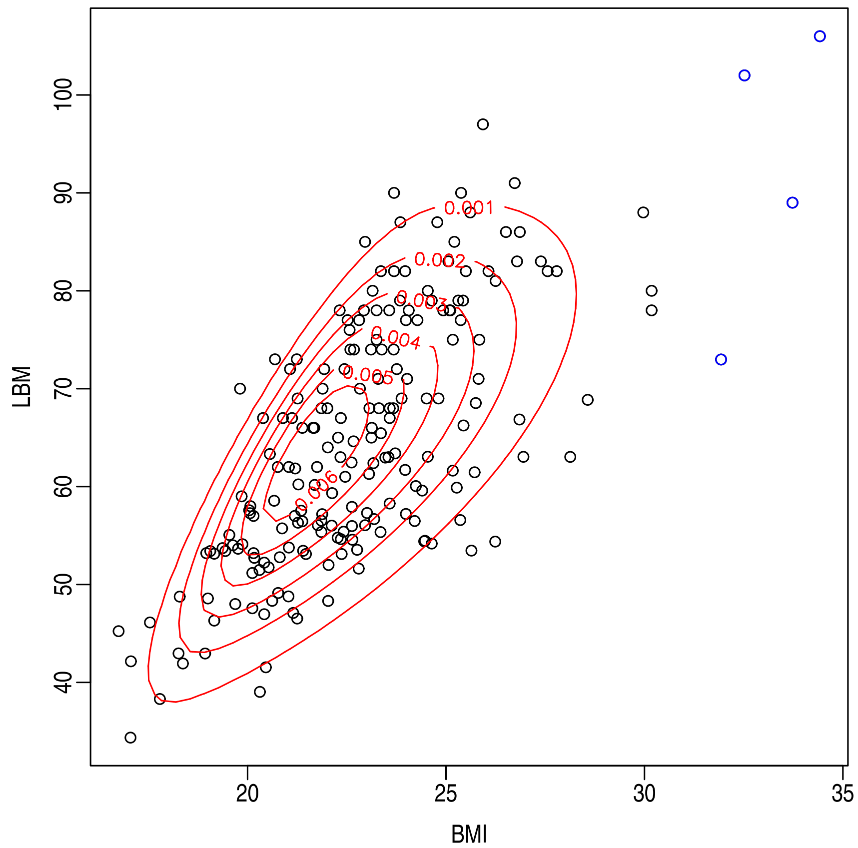
- more general: for an ‘even’ function $T(\cdot)$, such that $T(x)=T(-x)$ for all $x\in\mathbb{R}^d$,

$$T(X) \stackrel{d}{=} T(Y)$$

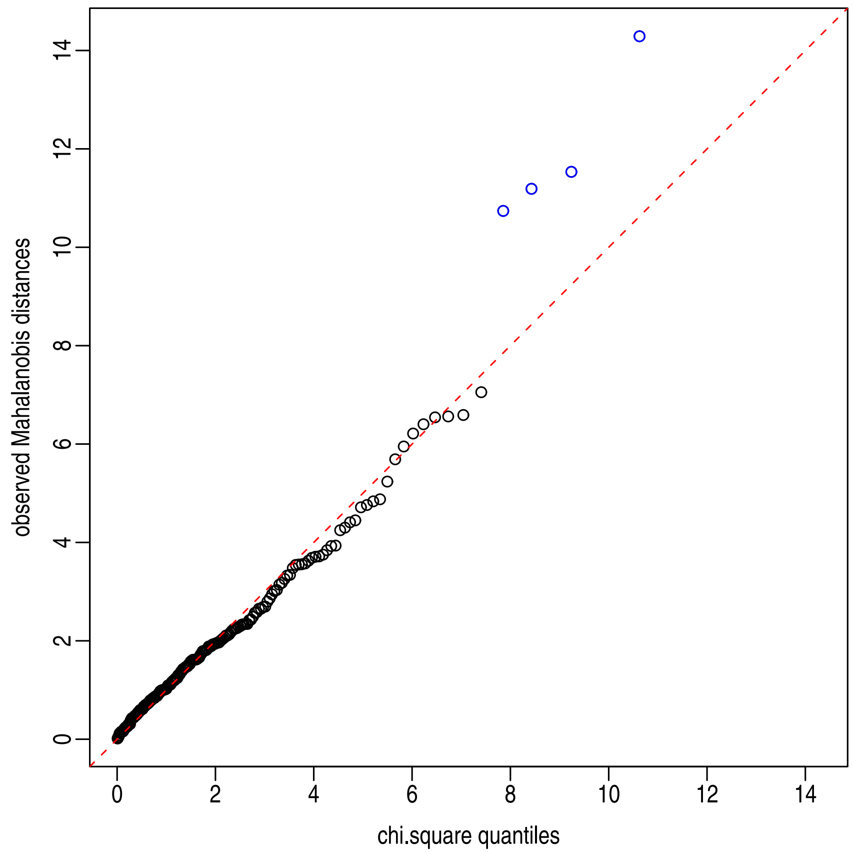
[Azzalini & Capitanio (1999, 2003), Genton et al. (2001), Loperfido (2001), Genton & Loperfido (to appear)]

A numerical example: AIS data

Australian Institute of Sport data ($n = 202$)



scatterplot and SN fit using MLE



Healy's QQ plot for SN fit

ESN and graphical models

'Extended' SN: $Z = (U_1 | U_0 > \tau)$ with pdf

$$\text{const} \times \phi_d(z; \bar\Omega)\, \Phi(\alpha_0 + \alpha^\top z), \qquad \alpha_0 = \tau\sqrt{1 + \alpha^\top \bar\Omega \alpha}$$

ESN class $\{Y = \xi + \omega\, Z\}$ is closed under conditioning
[Arnold & Beaver (2000a)]

Graphical models:

- can build a conditional independence graph for Y (with some restrictions)
- relationship between the conditional independence graphs of $\xi + \omega\, U \sim N_d(\xi, \Omega)$ and of $Y \sim \mathrm{SN}_d(\xi, \Omega, \alpha, \tau)$
- log-likelihood allows a parameter based factorization

[Capitanio, Azzalini & Stanghellini (SJS, 2003)]

Stochastic processes

- spatial processes with SN components,
  application to rainfall data,
  Bayesian inference via MCMC [Kim & Mallick (2004)]
- model for non-Gaussian noise in signals processing
  [Gualtierotti (2004)]

Multiple constraints

Consider $Z = (U_1 | U_0 > \tau)$ where $(U_0, U_1) \sim N_{h+d}$

Various cases:

- product of $\Phi(\cdot)$ terms [Azzalini (1985), Arnold & Beaver (2000a)]
- case $h = d$, involves a $\Phi_d(\cdot)$ term
  [Sahu, Dey & Branco (2004)]
- Hierarchical SN [Liseo & Loperfido (2003)]
- Fundamental SN [Arellano-Valle & Genton (to appear)]
- General SN

$$\phi_d(y; \xi, \Omega)\, \Phi_h(D(y-\xi); \nu, \Delta) / \Phi_h(D\xi; \nu, \Delta + D\Omega D^\top)$$

[mentioned in Gupta, González-Farías & Domínguez-Molina (2004), with focus on case $\nu = 0,\ \Delta = I_h$]

Perturbation of symmetric multivariate distributions

Extend basic result to $\mathbb{R}^d$

(central) symmetry about 0 in $\mathbb{R}^d$: $f_0(y) = f_0(-y)$

N.B. includes elliptically contoured densities

Lemma. Suppose f_0 and G' are pdf's symmetric about 0 in $\mathbb{R}^d$ and $\mathbb{R}$, respectively, then

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

is a density for any real-valued odd function $w(\cdot)$, i. e. $w(-x) = -w(x) \in \mathbb{R}$ for $x \in \mathbb{R}^d$.

Proof. Essentially the same of case $d = 1$.

Can be extended to h-dimensional G, and other directions.

[Azzalini & Capitanio (2003), Genton & Loperfido (to appear),
see also Arellano-Valle, del Pino & San Martín (2002) for another formulation]

A simple stochastic representation

Suppose $X \sim G'$, $Y \sim f_0$, independent. Then

$$Z = \begin{cases} Y & \text{if } X < w(Y) \\ -Y & \text{if } X > w(Y) \end{cases}$$

has density

$$f(z) = 2\, f_0(z)\, G\{w(z)\}.$$

Corollary. If $Y \sim f_0$ and $Z \sim f$, with $f(z) = 2\, f_0(z)\, G\{w(z)\}$, then

$$T(Y) \stackrel{d}{=} T(Z)$$

for any even function $T(\cdot)$.

Can obtain high flexibility – an example

Consider a symmetric Beta pdf, scaled in $(-1,1)^2$:

$$f_0(y) = \frac{(1-y_1^2)^{a-1}\,(1-y_2^2)^{b-1}}{4^{a+b-1}\,B(a,a)\,B(b,b)}, \qquad y=(y_1,y_2)\in(-1,1)^2,$$

perturbed by

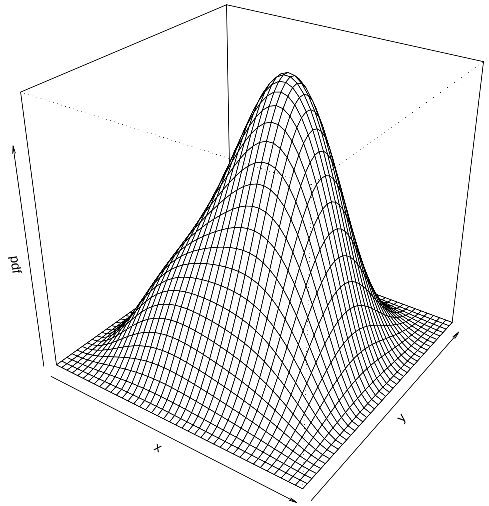
$$G(x) = \frac{e^x}{1+e^x}, \qquad w(y) = \frac{\sin(p_1 y_1 + p_2 y_2)}{1+\cos(q_1 y_1 + q_2 y_2)}$$

where p_1, p_2, q_1, q_2 are additional parameters.

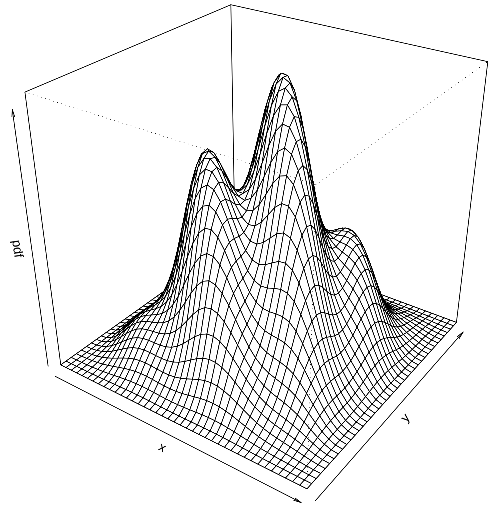
There is more than skewing in the ‘skew-symmetric’ class

Perturbed symmetric Beta pdf’s

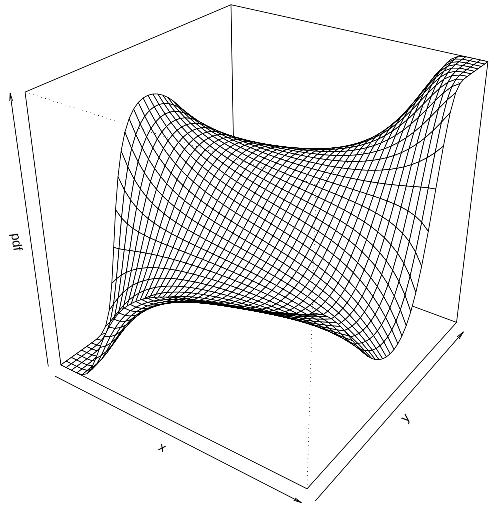
(a,b,p,q) = (2 , 3 , (3 , 3), (0 , 0))



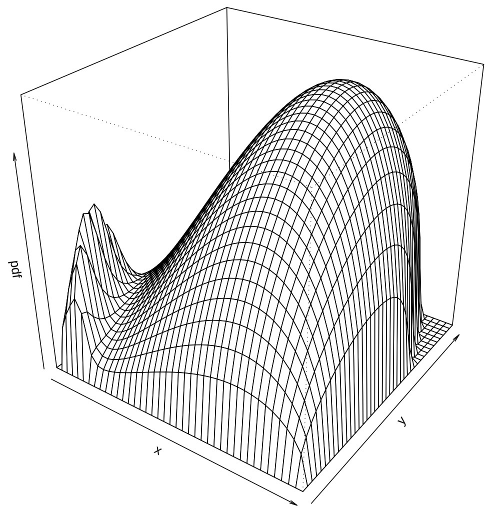
(a,b,p,q) = (2 , 3 , (8 , 8), (0 , 0))



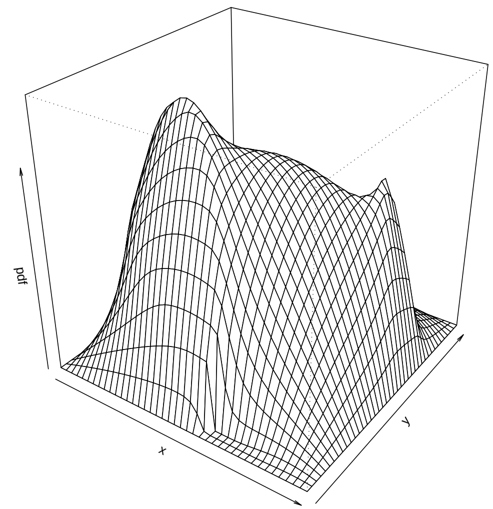
(a,b,p,q) = (3 , 1 , (−1 , 3), (2 , 1))



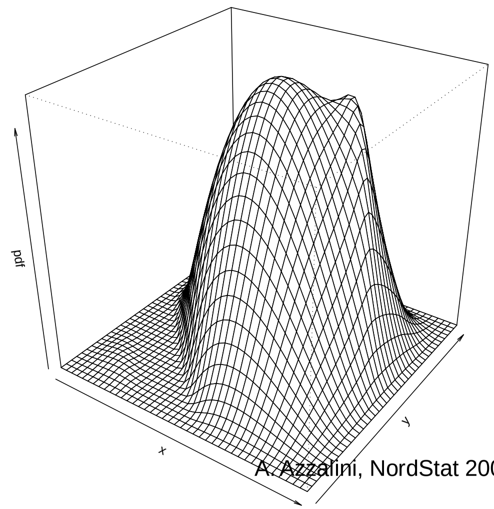
(a,b,p,q) = (3 , 1.5 , (3 , 1), (2.5 , 1))



(a,b,p,q) = (3 , 2 , (2 , −3), (2 , 4))



(a,b,p,q) = (3 , 3 , (1 , 1), (3 , 3))



Skew-ellipical densities

$$f_0(y) = \frac{c_d}{|\Omega|^{1/2}}\,\tilde{f}\left[(y-\xi)^\top\Omega^{-1}(y-\xi)\right]$$

- Route A: use previous lemma, with some $G(\cdot)$ and $w(\cdot)$ [Azzalini & Capitanio (1999)]
- Route B: apply a random scale factor to a SN variate [Branco & Dey (2001)]

Good news: for appropriate choice of the ingredients, 'essentially' the same thing. [Azzalini & Capitanio (2003)]

Analogies of SN and S-elliptical families

- the two stochastic representations for SN variates still hold, replacing normal with elliptical variates
  - ◇ via conditioning: $(U_1|U_0 > 0)$,
  - ◇ via addition: $\delta|V_0|\,1_d + \sqrt{1-\delta^2}\,V_1$
- for case $d = 2$, also representation via order statistics
- the distribution of quadratic forms is the same for skew-elliptical and elliptical variates with common symmetric component

[Azzalini & Capitanio (2003), Fang (2003)]

A noteworthy case: skew-t

If $Z \sim SN(0, \Omega, \alpha)$ and $V \sim \chi^2_\nu/\nu$, then

$$Y = \xi + V^{-1/2}\, Z$$

has pdf

$$f(y) = 2\, t_d(y; \nu)\, T_1\left(w(y); \nu + d\right)$$

where

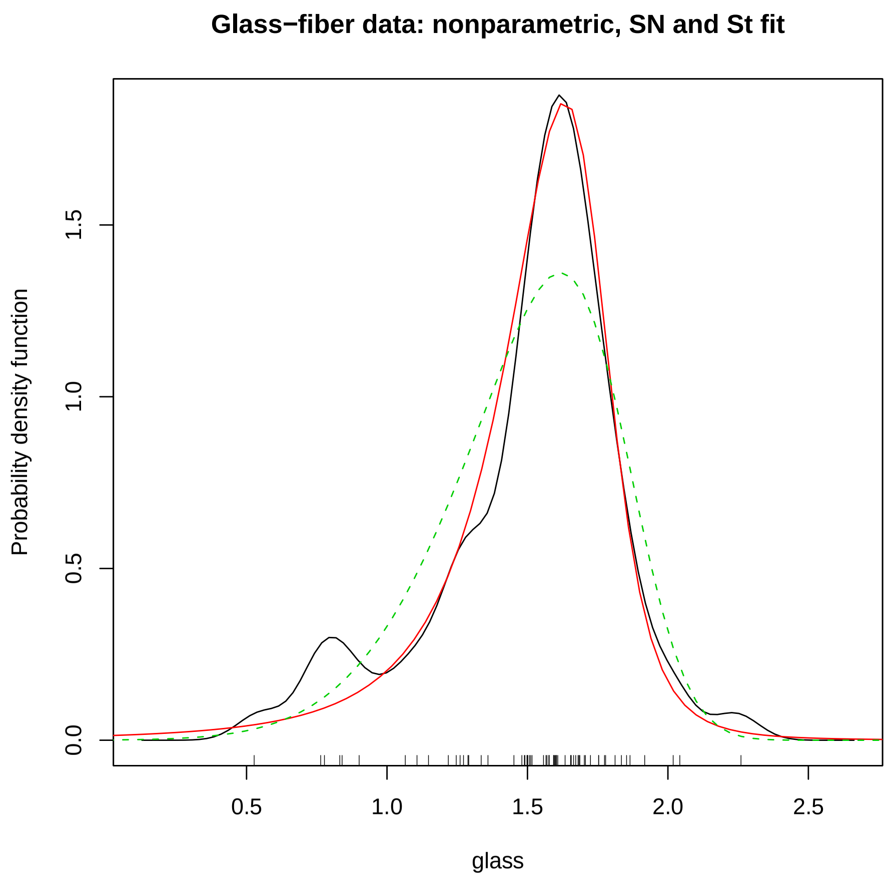
$$w(y) = \alpha^\top \omega\,(y - \xi) \left(\frac{d+\nu}{\|y-\xi\|^2_\Omega + \nu}\right)^{1/2}.$$

Unlimited range of indices of skewness and kurtosis.
Suitable for robust inference.

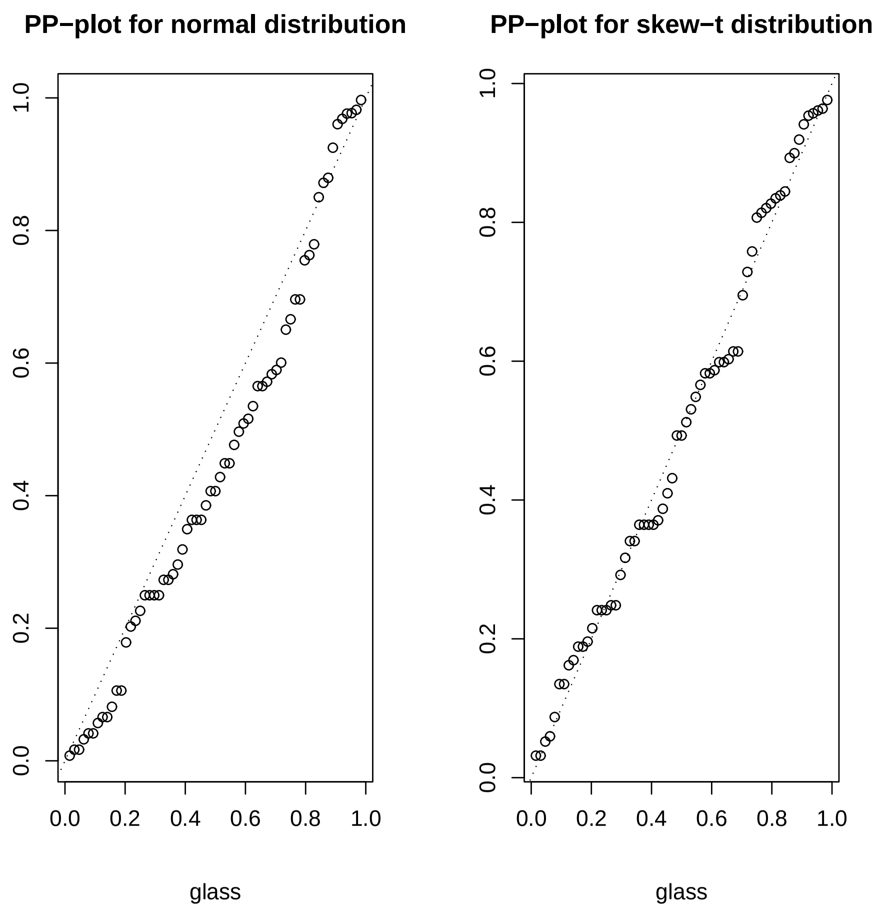
$(Y-\xi)^\top \Omega^{-1}(Y-\xi) \sim d \times F(d, \nu)$ for QQ- and PP-plot

Example: fiber glass strength

data on breaking strength of fiber glass ($n = 63$)



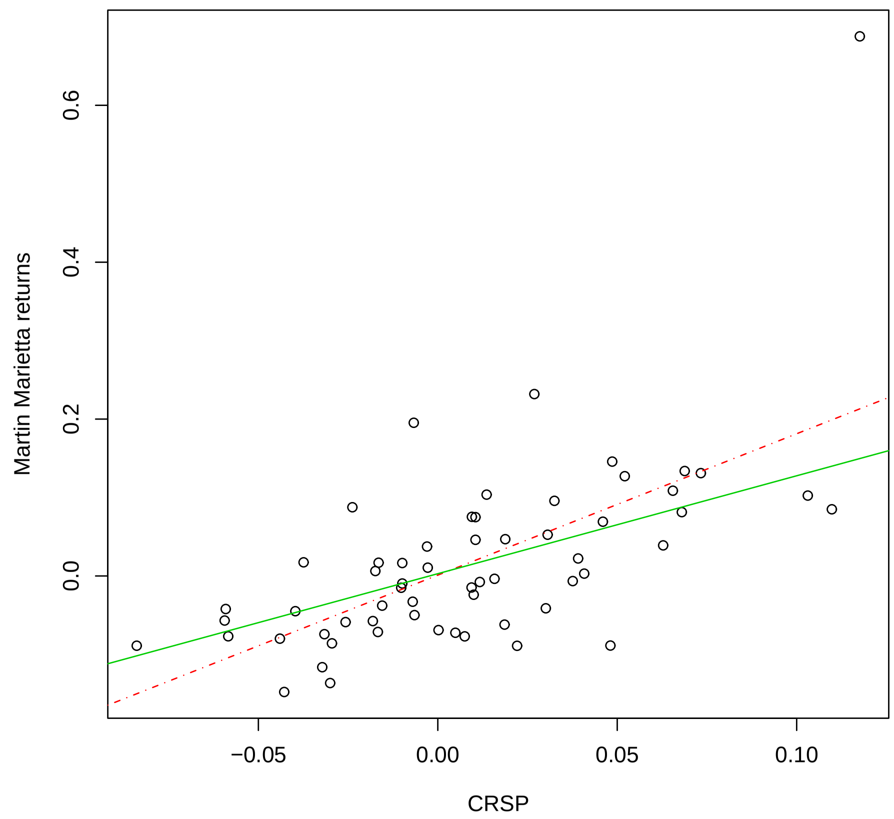
nonparametric estimate, SN fit, ST fit

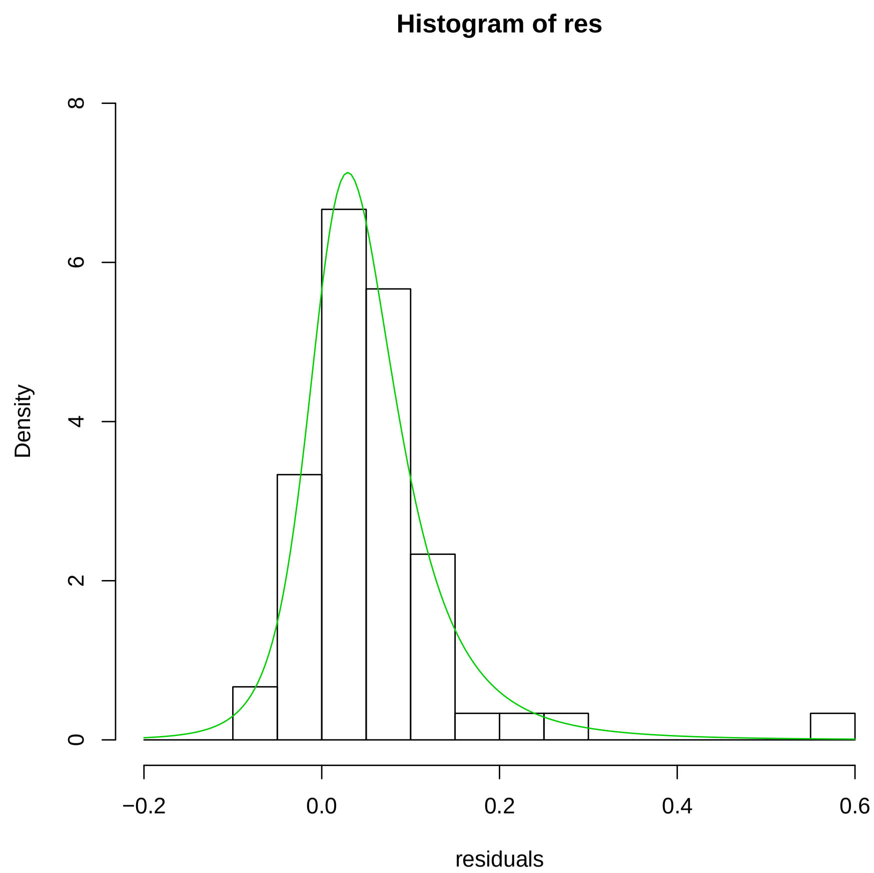


PP-plot for Normal model and ST model

Example: MM returns

MM excess returns vs CRSP index ($n = 60$)

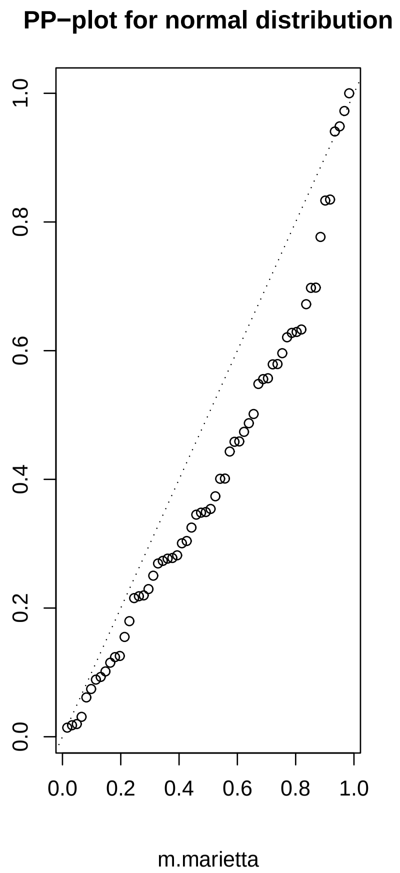


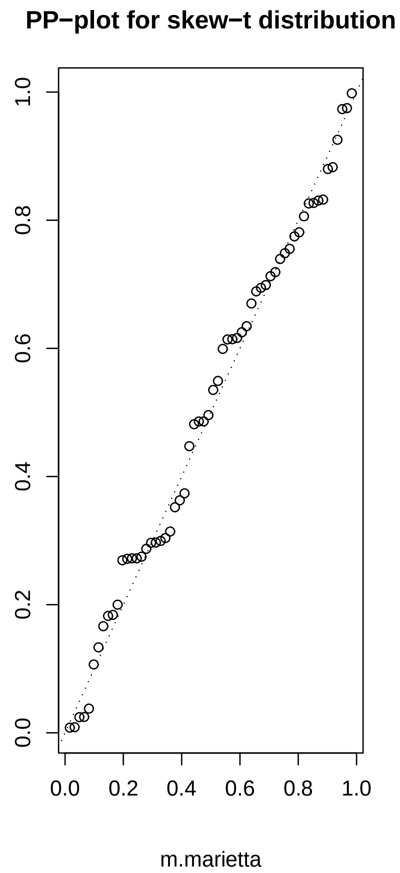


linear regression with Normal and ST fit

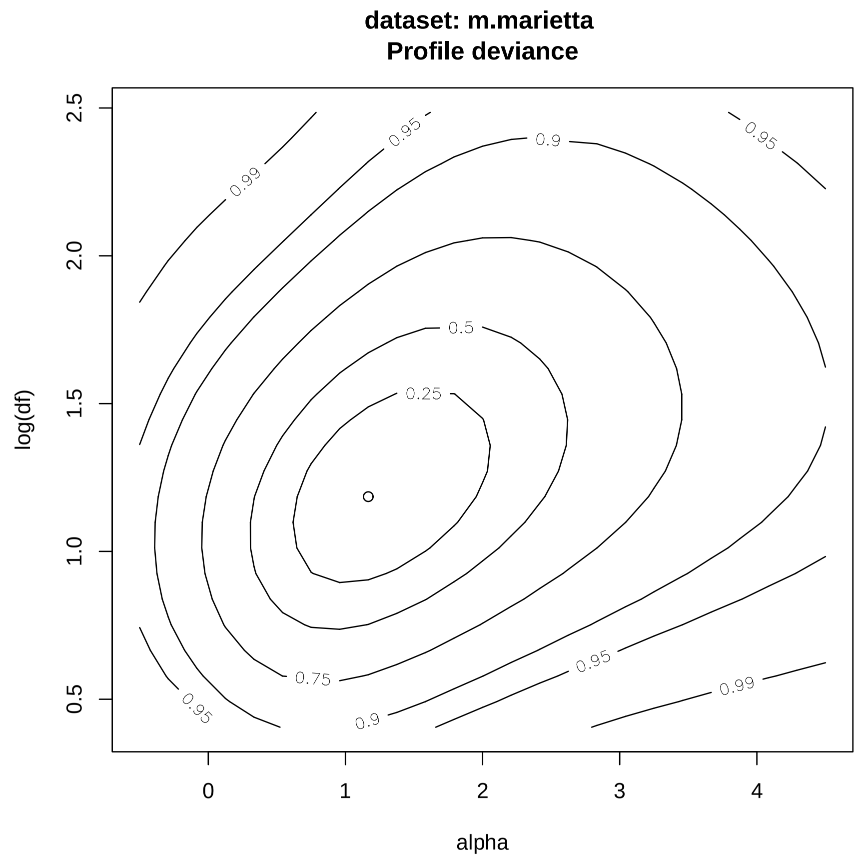
histogram of residuals and fitted ST dist'n

Example: MM returns





PP-plot for Normal model and ST model



profile $-2\log(\text{relative L})$ for $(\alpha, \log\nu)$

Building expansions of arbitrary pdf's

SkSy class: $f(z) = 2\, f_0(z)\, G\{w(z)\}$

- An arbitrary pdf $g(x)$ can be factorised as

$$g(x) = 2\, f_\xi(x-\xi)\, G\{w_\xi(x-\xi)\}$$

where f_ξ is symmetric around 0, and w_ξ is odd.
The factorisation depends on the chosen ξ.

- Let $P_K(\cdot)$ denote odd polynomial of degree K. The class

$$2\, f_0(x-\xi)\, G\{P_K(x-\xi)\}$$

is dense in SkSy-class (under some weak conditions).

[Wang, Boyer & Genton (2004), Ma & Genton (SJS, 2004)]

Some applications

Use as a mathematical tool

not really an 'application', still useful

- to approximate a distribution (see above)
- local departures from normality, local efficiency, . . .

  [Salvan (1986), Durio & Nikitin (2003), Muliere & Nikitin (2004)]
- to produce 'alternatives' in simulation studies

  [DiCiccio et al. (1997), Chu et al. (2001), etc.]
- 'skew link' for dicotomous response analysis

  [Chen, Dey & Qi (1999)]

Connected areas

- Selected (non-random) sampling and Heckman model

response eqn: $y = \beta^\top x + \sigma\varepsilon_1, \qquad \varepsilon_1 \sim N(0,1)$

selection eqn: $z = \gamma^\top x + \varepsilon_2 > 0, \qquad \varepsilon_2 \sim N(0,1)$

$\Longrightarrow \quad f(y|x, z > 0) =$ type ESN, $\delta = \text{corr}\{\varepsilon_1, \varepsilon_2\}$

[Copas & Li (1997)]

- stochastic frontier model

$$\begin{aligned} y_i &= f(x_i, \beta) - \sigma_1 |\eta_i| + \sigma_2\,\varepsilon_i \\ &= f(x_i, \beta) + \omega\left(-\delta|\eta_i| + \sqrt{1-\delta^2}\,\varepsilon_i\right) \end{aligned}$$

[Aigner, Lowell & Schmidt (1977),. . . , Tancredi (2003+)]

Compositional data

$$(x_1, \ldots, x_{d+1}), \quad x \in S_{d+1} \quad \left(\text{i.e. } x_j \in \mathbb{R}^+, \quad \sum_{j=1}^{d+1} x_j = 1\right)$$

- frequent sort of data in geology, among others areas
- popular approach: Aitchison ALR transformation to $\mathbb{R}^d$,

$$y_j = \log(x_j/x_{d+1}), \qquad (j = 1, \ldots, d),$$

  enjoyes a set of convenient formal properties
- then apply methods for multivariate Normal data to y_j's
- use of SN_d in place of N_d preserves closure of suitable transformations of x on S_{d+1}: $a \oplus (b \otimes x)$ and subcompositions

[Mateu-Figueras (Ph.D thesis, 2003), Aitchison, Mateu-Figueras & Ng (2003)]

Applications in finance

- multivariate SN fits well into portfolio selection theory
  use of skew-t allows fat tails

  [Adcock & Shutes (1999), Adcock & Mead (2003), Harvey et al. (to appear)]

- incorporating skewness/kurtosis in components of ARCH, GARCH and SV models

  [Goria (thesis, 1999), Pietrobon (thesis, 2003), De Luca & Loperfido (2004), Cappuccio et al. (2004), Liseo & Loperfido (to appear)]

nearly there...

Some open problems

- a multitude of proposals (emphasis on probability side, less about statistics; what about their properties, connections, suitability for data analysis, etc.?)
- various issues in inference:
  - avoid some peculiarities of MLE
  - alternatives to MLE
- extend various methods from Normal to non-Normal case
- can approximate any distribution, but
  - what is ‘best’ type of approximation?
  - how to get approximation with prescribed properties?
- is all of this really useful?

Resources

References, some papers, software, etc. at

`http://azzalini.stat.unipd.it/SN/`