Discussion on: *The role of skewed distributions in Bayesian inference: conjugacy, scalable approximations and asymptotics*, by D. Durante

Brunero Liseo

Sapienza Università di Roma

brunero.liseo@uniroma1.it

Statistical methods and models for complex data

Padova, September 21^{st}, 2022

Prologue

- SN family of densities and its generalizations have been very important in my careeer. I have been involved since 1987 for my dissertation thesis
- I am really happy that Daniele and his collaborators have uncovered the central role of the SUN family in several routinely used statistical models.

Key words in Durante's talk

- Unification
- Conjugacy
- Computation

Unification

The main message is:

Many different statistical models can be conveniently rephrased into a common likelihood framework.

$$p(y;\boldsymbol{\beta}) = p(y_1;\boldsymbol{\beta},\Sigma_1)p(y_0;\boldsymbol{\beta},\Sigma_0) \propto \varphi_{n_1}(y_1 - X_1\boldsymbol{\beta};\Sigma_1)\,\Phi_{n_0}(y_0 + X_0\boldsymbol{\beta};\Sigma_0)$$

with $y = (y_0', y_1')'$. It includes `probit`, `multivariate probit`, `multinomial probit`, `tobit` and so on ...

Luckily enough, this expression is the kernel of a SUN density! So many different procedures are merely particular cases of a more general one, both in a Bayesian and in a likelihood framework...

Conjugacy

This second key-word makes the Bayesian approach more convenient beacuse of the main theorem presented today

Theorem

If $\boldsymbol{\beta} \sim SUN_{p,q}(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\Delta}, \gamma, \boldsymbol{\Gamma})$ *and the likelihood is as in the previous slide, then*

$$\boldsymbol{\beta}|y \sim SUN_{p,q+n_0}(\boldsymbol{\xi}^*, \boldsymbol{\Omega}^*, \boldsymbol{\Delta}^*, \gamma^*, \boldsymbol{\Gamma}^*)$$

SUN and Conjugacy

- Conjugacy has been a key concept in the early development of Bayesian methods, until the mid 80's.
- It has kept a relevant role in the MCMC era, for example, in terms of
  - Conditional conjugacy in Gibbs-type algorithms
  - base measure of a Dirichlet Process Mixtures in BNP
  - Fast Variable Selection in Linear Models via the use of g-priors.

Conjugacy and SUN

The conjugacy property of the SUN family, extends the pros of conjugate analysis to a variety of statistical models and it allows, at least in principle, to

- Posterior sampling, provided that a good algorithm is available for SUN random vectors!
- Closed form expressions for
  - marginal posteriors
  - posterior moments (SUN vectors have Moment Generating Function)
  - normalizing constant $\Longrightarrow$ BF $\Longrightarrow$ model choice

Computation

The critical issues are

- the computation of the CDF

$$\Phi(x, \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

  of a multivariate Gaussian vector for large n ...
- sampling from truncated versions of $\Phi(x, \boldsymbol{\mu}, \boldsymbol{\Sigma})$

In general, the coniugacy operates in such a way that the sample size n will increase the n_0 quantity, i.e. the dimension of the cdf Φ.

Computation

- For the above reason, the methodology is in general more efficient in the $p >> n$ scenario.
- However in this scenario, the role of the prior density on the p-dimensional vector β becomes crucial and we should start learning how the elicit within the *SUN* family.

A further generalization

Research question:

Why the logistic model is out of this framework ?

The likelihood of a binary logistic regression is expressed in terms of the Logistic CDF which is a scale mixture of Gaussian CDF
Then, one can define a perturbed SUN family by replacing φ and Φ with scale mixtures of Gaussian pdf and CDF!

The density of a pSUN

Let $Y \sim pSUN_{d,m}(Q_V, \Theta, A, b, Q_W, \Omega, \xi)$. Then

$$f_Y(y) = \varphi_{\Omega,Q_W}(y-\xi)\frac{\Phi_{\Theta,Q_V}\left(A\,\mathrm{diag}^{-\frac{1}{2}}(\Omega)(y-\xi)+b\right)}{\Psi_{Q_V,\Theta,A,Q_W,\bar{\Omega}}(b)}, \qquad (1)$$

with

$$\varphi_{\Sigma,Q}(u) = \int_{\mathbb{R}^d}\prod_{i=1}^{d}\left(W_i^{-\frac{1}{2}}\right)\phi_\Sigma\left(\mathrm{diag}^{-\frac{1}{2}}(W)\,u\right)dQ(W),$$

$$\Phi_{\Sigma,Q_V}(u) = \int_{\mathbb{R}^d}\Phi_\Sigma\left(\mathrm{diag}^{-\frac{1}{2}}(V)\,u\right)dQ(V),$$

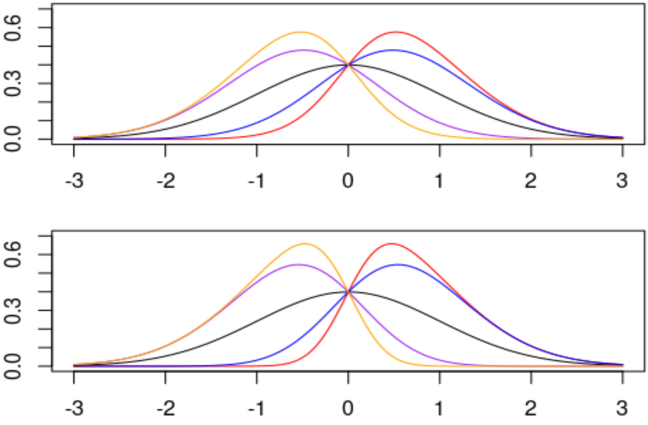
and

$$\Psi_{Q_V,\Theta,A,Q_W,\bar{\Omega}}(b) = \mathrm{P}(T - AZ \leq b)$$

$$T \sim \Phi_{\Theta,Q_V}(\cdot) \perp\!\!\!\perp Z \sim \Phi_{\bar{\Omega},Q_W}(\cdot)$$

Some pSUN densities

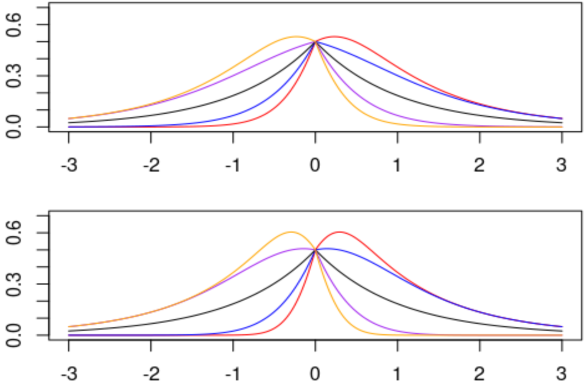
Logit: top: N(0,1); $V \sim LK(\cdot); W = 1, A = 3, b = 0$
$V \sim LK(\cdot); W = 1, A = 1.5, b = 0$



Probit: bottom: N(0,1); $V = W = 1, A = 3, b = 0$

Some pSUN densities

top: Lapl(0,1) ; $V \sim LK(\cdot); W \sim Exp(0.5), A = 3, b = 0;$
$V \sim LK(\cdot); W \sim Exp(0.5), A = 1.5, b = 0$



bottom: Lapl(0,1) $V = 1, W \sim Exp(0.5), A = 3, b = 0;$
$V = 1, W \sim Exp(0.5), A = 1.5, b = 0$

Sampling a pSUN

We adopted a Gibbs algorithm:

- Key aspect: one must be able to sample from the f.c.'s $W|Z$ and $V|T$.
- It is not always easy, and it depends on the specific values of $\Theta, \bar{\Omega}$, and the form of $Q_W(\cdot)$ and $Q_V(\cdot)$.
- Relatively simple in the most popular versions of the Bayesian binary regression, including the logistic case

For details, see
Onorati & L. (2022). *An extension of the Unified Skew-Normal family of distributions and application to Bayesian binary regression*. arXiv:2209.03474

Conclusions

- Congratulations to Adelchi for discovering the SUN family!
- Congratulations to Daniele & co. for boosting and popularizing its use!
- I look forward to see additional generalizations to multivariate logistic and other models!

Conclusions

- Congratulations to Adelchi for discovering the SUN family!
- Congratulations to Daniele & co. for boosting and popularizing its use!
- I look forward to see additional generalizations to multivariate logistic and other models!

THANKS!