Methods for modeling data from multiple individuals

Implementation of SSFs in INLA and glmmTMB

Stefanie Muff, Johannes Signer, John Fieberg

ISEC2022, June 26, 2022

Recap – the model we want to fit

- Since conditional logistic regression is a special case of a multinomial model, and as such is likelihood-equivalent to a Poisson model with $\mathrm{E}(y_{ntj}) = \mu_{ntj}$ (McCullagh and Nelder 1989)[1], we can fit

$$\mu_{ntj} = \exp(\alpha_{nt} + \boldsymbol{\beta}^\top \boldsymbol{x}_{ntj} + \boldsymbol{u}_n^\top \boldsymbol{z}_{ntj}) \ , \quad y_{ntj} \sim \mathrm{Po}(\mu_{ntj}) \ ,$$

  with stratum-specific intercepts α_{nt}.
  Note: No global intercept (why?).

- The intercepts are treated as random effects assuming $\alpha_{nt} \sim \mathsf{N}(0, \sigma_\alpha^2)$ with large σ_α^2 (*e.g.*, 10^6).

[1] Chapter 6.4.2

The requirement on the software

Any software that fulfils the following two requirements can be used:

1. Capability to *fit mixed Poisson models*.
2. Possibility to *fix (not estimate) a random effect variance*.

- First, it looked most straightforward to use a Bayesian method: Fixing the prior on the variance is straightforward.
- But: of course this assumes that you are familiar with the rest of the Bayesian procedure.
- For some, it may be more staightforward to use a likelihood-based approach.

The Bayesian way: INLA

- Integrated nested Laplace approximations (Rue, Martino, and Chopin 2009).

Journal of the Royal Statistical Society
SERIES B Statistical Methodology

J. R. Statist. Soc. B (2009)
71, *Part* 2, *pp.* 319–392

Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations

Håvard Rue and Sara Martino

Norwegian University for Science and Technology, Trondheim, Norway

and Nicolas Chopin

Centre de Recherche en Economie et Statistique and Ecole Nationale de la Statistique et de l'Administration Economique, Paris, France

- Idea: Fast approximations instead of MCMC sampling. A lot of mathy details.
- Some resources: www.r-inla.org

Hierarchical Bayesian models with INLA

INLA is able to deal with latent Gaussian *hierarchical models*, consisting of three sub-models:

- Observation model: Encodes information about data $\boldsymbol{y}|\boldsymbol{v}, \boldsymbol{\theta}_1$
  $\rightarrow$ Typically the regression model.

Hierarchical Bayesian models with INLA

INLA is able to deal with latent Gaussian *hierarchical models*, consisting of three sub-models:

- Observation model: Encodes information about data $\boldsymbol{y}|\boldsymbol{v},\boldsymbol{\theta}_1$
  $\rightarrow$ Typically the regression model.
- Latent model: The unobserved process $\boldsymbol{v}|\boldsymbol{\theta}_2$
  $\rightarrow$ Model for the regression parameters and additional latent Gaussian variable(s).

Hierarchical Bayesian models with INLA

INLA is able to deal with latent Gaussian *hierarchical models*, consisting of three sub-models:

- Observation model: Encodes information about data $\boldsymbol{y}|\boldsymbol{v},\boldsymbol{\theta}_1$
  $\rightarrow$ Typically the regression model.

- Latent model: The unobserved process $\boldsymbol{v}|\boldsymbol{\theta}_2$
  $\rightarrow$ Model for the regression parameters and additional latent Gaussian variable(s).

- Hyperpriors: Models for the hyperparameters in the observation and latent processes $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2$
  $\rightarrow$ Variances and other non-Gaussian paramters.

Understanding the "hyperparameters" in INLA

- Simply speaking: Hyperparameters $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ are the *non-Gaussian unobserved variables*.
- Require integration in INLA. Despite smart choice of integration points, this is *time consuming*!
- Typically (in SSFs): The (co)variance parameters.
- No of variance parameters is therefore one of the most limiting factors for computational feasibility.

$\rightarrow$ *It is usually smart to carefully choose the variables that should be modeled with (animal-specific) random slopes.*

Pros and cons of INLA

Pros:

- Very flexible and modular toolset.
- Possible to combine with, *e.g.*, measurement error or spatially correlated random effects.
- Has most of the advantages Bayesian techniques have (priors, posteriors, joint modeling etc).
  $\rightarrow$ See the *coded session* for an example where we look at some posterior marginal distributions.

Cons:

- High threshold to get started, complex to use.
- Not very well documented.
- Slower than frequentist methods.

Bayesian SSFs using MCMC instead of INLA?

- Possible, but inefficient.
- We tried Stan, but the computation time was 1-2 orders of magnitues slower.

The frequentist way: glmmTMB

Brooks et al. (2017)

The R Journal: article published in 2017, volume 9:2

glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling

Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Mächler and Benjamin M. Bolker , *The R Journal* (2017) 9:2, pages 378-400.

- Very fast frequentist inference.
- Easy to fix the random intercept variance.

→ glmmTMB seems to be what most users of the Poisson-based SSF model prefer.

Brooks, Mollie E., Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Mächler, and Benjamin M. Bolker. 2017. "glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling." *The R Journal* 9: 378–400.

McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*. London: Chapman; Hall.

Rue, H, S Martino, and N Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations (with Discussion)." *Journal of the Royal Statistical Society. Series B (Methodological)* 71: 319–92.