A joint analysis of counts and severity with zero-inflated longitudinal data

Giovani L. Silva[1], *Elizabeth Juarez-Colunga*[2], *Charmaine Dean*[3]

[1]giovani.silva@tecnico.ulisboa.pt, CEAUL & Dep. Mathematics-IST, Universidade de Lisboa
[2]elizabeth.juarez-colunga@ucdenver.edu, Dep. Biostatistics and Informatics, University of Colorado at Denver
[3]scidean@uwo.ca, Dep. Statistical and Actuarial Sciences, University of Western Ontario

In longitudinal studies, one can sometimes observed the frequency and intensity of episodes in order to explain the recurrent occurrence of an event of interest, for example, earthquake sequences in a certain region combined with their magnitude levels may indicate long or short free-event periods studying the joint distribution of count and severity. This work is a joint approach for analysis of longitudinal data, as the severity of the event of interest is jointly observed with its occurrence, motivated by a clinical trial involving participants who were healthy menstruating women prior to hysterectomy/ovariectomy for benign disease (Prior *et al.*, 2007, *Clinical Science*, 112, 517-525).

That joint approach allows mixtures of discrete (counting of hotflush events) and categorical (severity of the events) response variables over time. Each response is related to individual-specific random effects, which may be correlated over time, through a generalized linear model (Dunson, 2003, *Journal of the American Statistical Association*, 98, 555-563). Because there is many zero counts in the motivating dataset, we adapted the proposed model to zero-inflated modelling by using different counting distributions.

We propose a Bayesian joint analysis of counts and severity with zero-inflated longitudinal data to formulate multivariate correlated models for a combination of binary, ordinal, discrete and continuous outcomes measuring the same underlying trend over time. Our models fall within the general framework of generalized linear latent and mixed models. But, the number of random effects makes this approach computationally very intensive. Following a Bayesian approach to inference, Markov chain Monte Carlo methods are used for computing some posterior quantities of interest, easily implemented using OpenBUGS (Spiegelhalter *et al.*, 2007, *OpenBUGS User Manual*, version 3.0.2).

Keywords: Repeated measurements, zero-inflated model, Bayesian statistics.