# A Monte Carlo EM for the Poisson Log-Normal Model 

Monday, 11 September 2023 14:40 (30 minutes)


#### Abstract

The Poisson log-normal (PLN) model is a generic model for the joint distribution of count data, accounting for covariates. It is also an incomplete data model. A classical way to achieve maximum likelihood inference for model parameters $\theta$ is to resort to the EM algorithm, which aims at maximizing, with respect to $\theta$, the conditional expectation, given the observed data $Y$, of the so-called complete $\log$-likelihood $\mathbb{E}\left[\log p_{\theta}(Y, Z) \mid\right.$ $Y]$. Unfortunately, the evaluation of $\mathbb{E}\left[\log p_{\theta}(Y, Z) \mid Y\right]$ is intractable in the case of the PLN model because the conditional distribution of the latent vector conditionally on the corresponding observed count vector has no closed form and none of its moments can be evaluated in an efficient manner.

Variational approaches have been studied to tackle this problem but lack from statistical guarantees. Indeed the resulting estimate $\widehat{\theta}_{V E M}$ does not enjoy the general properties of MLEs. In particular, the (asymptotic) variance of $\widehat{\theta}_{V E M}$ is unknown, so no test nor confidence interval can be derived easily from the variational inference. Starting from already available variational approximations, we define a first Monte Carlo EM algorithm to obtain maximum likelihood estimators of this model. We then extend this first algorithm to the case of a composite likelihood in order to be able to handle higher dimensional count data. Both methods are statically grounded and provide confidence region for model parameters.


## Keywords

Multivariate count data, Monte Carlo EM algorithm, composite likelihood

## Classification

Both methodology and application

Primary authors: STOEHR, Julien (Université Paris-Dauphine, PSL); Prof. ROBIN, Stéphane (Sorbonne Université)

Presenter: STOEHR, Julien (Université Paris-Dauphine, PSL)
Session Classification: INVITED SFdS on Bayesian Statistics

Track Classification: Other/special session/invited session

