# Appendix to: "Efficient posterior sampling for Bayesian Poisson regression"

Laura D'Angelo*
Department of Economics, Management and Statistics,
University of Milano-Bicocca
and
Antonio Canale
Department of Statistical Sciences, University of Padova

September 5, 2022

## Appendix A: Derivation of the data augmentation scheme

In this section, we derive the data augmentation scheme based on Pólya-gamma random variables for the negative binomial model in Equation (2) of the paper, where we make explicit the mean parameter $\lambda_i = e^{x_i^T \beta}$

$$
\begin{aligned}
\tilde{f}_{r_i}(y_i \mid \beta) &= \binom{r_i + y_i - 1}{r_i - 1} \left( \frac{r_i}{r_i + e^{x_i^T \beta}} \right)^{r_i} \left( \frac{e^{x_i^T \beta}}{r_i + e^{x_i^T \beta}} \right)^{y_i} \\
&= \binom{r_i + y_i - 1}{r_i - 1} r_i^{r_i} \frac{(e^{x_i^T \beta})^{y_i}}{(r + e^{x_i^T \beta})^{r_i + y_i}} \\
&= \binom{r_i + y_i - 1}{r_i - 1} r_i^{r_i} \frac{r_i^{y_i}}{r_i^{r_i + y_i}} \frac{(e^{x_i^T \beta - \log r_i})^{y_i}}{(1 + e^{x_i^T \beta - \log r_i})^{r_i + y_i}} \\
&= \binom{r_i + y_i - 1}{r_i - 1} \frac{(e^{x_i^T \beta - \log r_i})^{y_i}}{(1 + e^{x_i^T \beta - \log r_i})^{r_i + y_i}}.
\end{aligned}
$$

From this form of the likelihood, it is immediate to obtain Equation (3) by simply adjusting the parameters of the original data augmentation of Polson et al. (2013).

*laura.dangelo@unimib.it

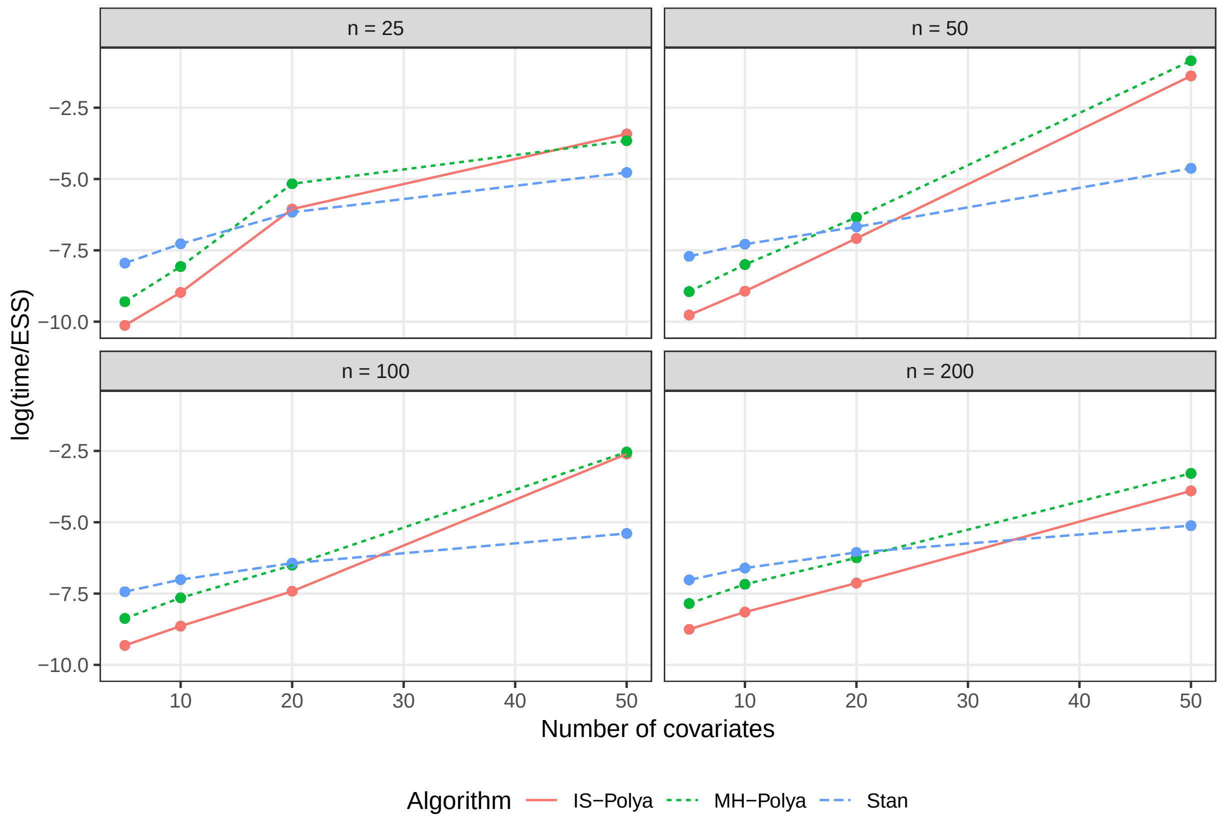


Figure 1: Time per independent sample (in logarithmic scale) for the three algorithms. For each combination of $n$ and $p$ the point is the median of the (log) time (in seconds) over the effective sample size using a Gaussian prior, over 50 replications.

# Appendix B: Simulation results for $p$ up to 50 covariates

In this section, we compare the performances of the proposed algorithms for a number of covariates $p = 5, 10, 20, 50$. Specifically, similarly to the main text, we analyze the logarithm of the time per independent sample for the proposed Metropolis-Hastings algorithm and importance sampler, and the Stan implementation of the HMC (Stan Development Team 2021). For each combination we only plot the median time per independent sample (instead of a boxplot, as in the paper), for graphical reasons and ease of interpretation.

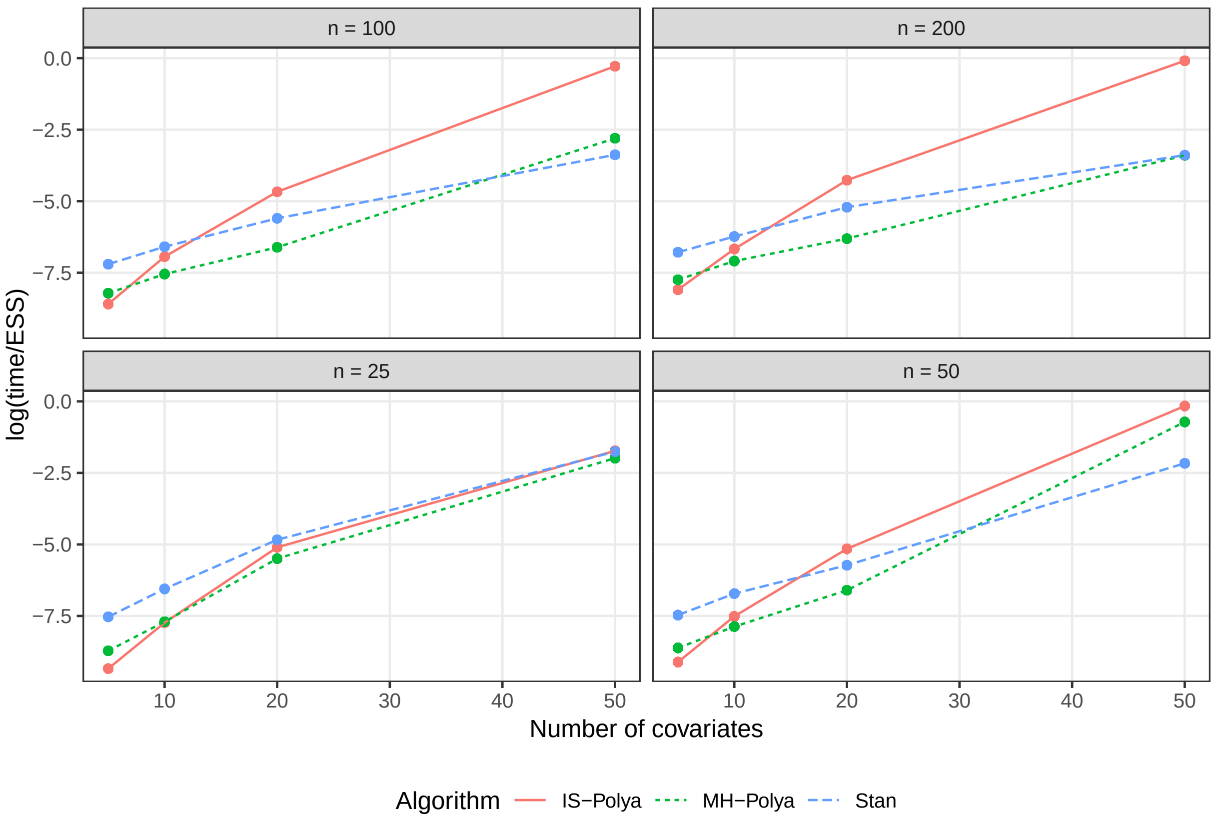


Figure 2: Time per independent sample (in logarithmic scale) for the three algorithms. For each combination of $n$ and $p$ the point is the median of the (log) time (in seconds) over the effective sample size using the horseshoe prior, over 50 replications.

# References


Polson, N. G., Scott, J. G. & Windle, J. (2013), 'Bayesian inference for logistic models using Pólya-gamma latent variables', *Journal of the American Statistical Association* **108**(504), 1339–1349.

Stan Development Team (2021), 'Stan modeling language users guide and reference manual', URL: `http://mc-stan.org/`.