Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 12, n. 1, 2026.

Trabalho apresentado no XLIV CNMAC, Fundação Getulio Vargas - Rio de Janeiro - RJ, 2025.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Statistical Inference for Intractable Count Distributions

Wellington J. Silva[1] Luiz M. Carvalho[2]
School of Applied Mathematics - Getulio Vargas Foundation, Rio de Janeiro, RJ.

Markov chain Monte Carlo (MCMC) is routinely used to perform Bayesian inference for a wide range of complex models (see [1]). However, in Bayesian models with intractable normalising distributions — where the normalising constant lacks a closed-form expression — standard MCMC methods face significant challenges, with no definitive solution. To illustrate this, consider a posterior density of the usual form:

$$\pi(\theta \mid y) = \frac{p(y \mid \theta) \times \pi(\theta)}{p(y)} \qquad (1)$$
$$= \frac{f(y;\theta)}{Z(\theta)} \times \pi(\theta) \times \frac{1}{p(y)},$$

where the marginal likelihood is given by $p(y) = \int p(y \mid \theta)\pi(\theta)d\theta$ and $Z(\theta)$ does not have a closed form. This affects MCMC when we express the probability of accepting a proposed parameter from a proposal distribution $q(\theta' \mid \theta)$:

$$\alpha(\theta', \theta) = \min\left\{1, \frac{f(y;\theta')\pi(\theta')q(\theta' \mid \theta)}{f(y;\theta)\pi(\theta)q(\theta \mid \theta')} \times \frac{Z(\theta)}{Z(\theta')}\right\}. \qquad (2)$$

The last term, however, cannot be evaluated directly, then we are not able to use the classic MCMC. These constants often appear in likelihood functions, such as in the **Double Poisson distribution** a generalization of the Poisson distribution that introduces an extra dispersion parameter, allowing for overdispersion or underdispersion. Its probability mass function (p.m.f) is

$$p(y \mid \mu, \phi) = \frac{\exp(-y)y^y}{y!}\left(\frac{\exp(1)\mu}{y}\right)^{\phi y} \cdot \frac{1}{Z(\mu,\phi)}, \qquad (3)$$

where the normalising constant $Z(\mu, \phi)$ depends on both parameters (μ and ϕ). Specifically:

$$Z(\mu,\phi) = \sum_{y=0}^{\infty} \frac{\exp(-y)y^y}{y!}\left(\frac{\exp(1)\mu}{y}\right)^{\phi y}. \qquad (4)$$

This dependence on both μ and ϕ makes inference computationally challenging. Similarly, the **Conway-Maxwell Poisson (COM-Poisson) distribution** which the p.m.f is defined by

$$p(y \mid \lambda, \nu) = \frac{\lambda^y}{(y!)^\nu} \cdot \frac{1}{Z(\lambda,\nu)}, \qquad (5)$$

where $Z(\lambda, \nu) = \sum_{y=0}^{\infty} \frac{\lambda^y}{(y!)^\nu}$, note that for $\nu = 1$, the distribution reduces to a standard Poisson distribution. Or even the **Zipf distribution** which is a discrete power-law distribution where the probability mass function (p.m.f.) is given by

[1]wellington.71319@gmail.com
[2]luiz.fagundes@fgv.br

$$p(k \mid s, N) = \frac{k^{-s}}{Z(s,N)},\ k = 1, 2, \ldots, N \tag{6}$$

where, $Z(s,N) = \sum_{k=1}^{N} k^{-s}$ is the normalising constant (e.g., the Riemann zeta function when $N \to \infty$).

We investigate methods for evaluating the p.m.f of the above distributions in the context of MCMC methods. We discuss in [2] how to use the ratio test to approximate normalization constants with the desired error.

This method can be seen in Table 1, named Bounding pairs[3], where a simple example is given with different ways of evaluating the normalisation constant with guaranteed low error. In Table 1, we compare our method with the Threshold, which evaluates until the nth term is less than a defined error (but which is not guaranteed, although it is widely used). And Fixed, evaluating a fixed number of terms.

Table 1: **Bayesian analysis of inventory data [3] under Conway-Maxwell Poisson model using noisy algorithms**. We show the posterior mean and Bayesian credible interval (BCI) for μ and ν and the median number of iterations n needed to get an approximation within $\varepsilon = 2.2\times10^{-16}$ of the true normalizing constant. Results for the noisy algorithm with fixed K. We provide estimates of the Monte Carlo standard error (MCSE), effective sample size (ESS) and effective sample size per minute (ESS / minute).

		Posterior median (BCI)	Posterior sd	MCSE	ESS	ESS/minute
Threshold	μ	0.805 (0.533, 1.084)	0.140	0.002	3269	115443
	ν	0.127 (0.104, 0.150)	0.012	0.000	3258	114830
	n	80 (75, 86)	2.862	0.048	3618	127517
Bounding pairs	μ	0.799 (0.533, 1.063)	0.137	0.003	2902	58481
	ν	0.126 (0.105, 0.148)	0.011	0.000	2925	58944
	n	81 (76, 88)	2.944	0.051	3318	66868
Fixed K = 100	μ	0.801 (0.535, 1.069)	0.136	0.002	3794	140271
	ν	0.127 (0.105, 0.149)	0.011	0.000	3804	140622
Fixed K = 3300	μ	0.800 (0.532, 1.074)	0.138	0.002	2992	4113
	ν	0.126 (0.104, 0.150)	0.011	0.000	2969	4082

References

[1] S. Brooks, A. Gelman, G. Jones, and X. Meng, eds. **Handbook of Markov chain Monte Carlo**. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton: CRC Press, 2011. ISBN: 9781420079425. DOI: 10.1201/b10905.

[2] L. M. Carvalho, W. J. Silva, and G. A. Moreira. “Adaptive truncation of infinite sums: applications to Statistics”. In: **arXiv preprint arXiv:2202.06121** (Feb. 2022).

[3] G. Shmueli, T. P. Minka, J. B. Kadane, S. Borle, and P. Boatwright. “A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution”. In: **Journal of the Royal Statistical Society: Series C (Applied Statistics)** 54.1 (Jan. 2005), pp. 127–142. DOI: 10.1111/j.1467-9876.2005.00474.x.

[3] μ is such that $\mu = \lambda^{\nu}$, for this case.