Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 11, n. 1, 2025.

Trabalho apresentado no XLIII CNMAC, Centro de Convenções do Armação Resort - Porto de Galinhas - PE, 2024

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Adaptive Truncation of Infinite Sums: Applications to Statistics

Wellington J. Silva[1] Luiz M. Carvalho[2]
Escola de Matemática Aplicada – Fundação Getulio Vargas, Rio de Janeiro, RJ.
Guido A. Moreira[3]
Bavarian Nordic, Germany.

It is often the case in Statistics that one needs to compute sums of infinite series, especially in marginalising over discrete latent variables. This has become more relevant with the popularization of gradient-based techniques (e.g. Hamiltonian Monte Carlo) in the Bayesian inference context, for which discrete latent variables are hard or impossible to deal with. For many major infinite series, custom algorithms have been developed which exploit specific features of each problem. In contrast, here we employ basic results from the theory of infinite series to investigate general, problem-agnostic algorithms to approximate (truncate) infinite sums within an arbitrary tolerance $\varepsilon > 0$ and provide robust computational implementations with provable guarantees.

A major idea that will be explored in this paper is that of finding error-bounding pairs, and in what follows it will be convenient to establish Proposition 0.1, which is inspired by the results in [2].

Proposition 0.1 (**Bounding a convergent infinite series**). *Let $S := \sum_{k=0}^{\infty} a_k$ and $S_n := \sum_{k=0}^{n} a_k$. Under the assumptions that $(a_n)_{n\geq 0}$ is positive, decreasing and passes the ratio test, then for every $0 \leq n < \infty$ the following holds:*

$$S_n + a_{n+1}\left(\frac{1}{1-L}\right) < S < S_n + a_{n+1}\left(\frac{1}{1-\frac{a_{n+1}}{a_n}}\right), \tag{1}$$

*if $\frac{a_{n+1}}{a_n}$ **decreases** to L and*

$$S_n + a_{n+1}\left(\frac{1}{1-\frac{a_{n+1}}{a_n}}\right) < S < S_n + a_{n+1}\left(\frac{1}{1-L}\right), \tag{2}$$

*if $\frac{a_{n+1}}{a_n}$ **increases** to L.*

Here we show an application of the proposed techniques to the Conway-Maxwell Poisson distribution (COMP, [3]), which is a popular model for count data, mainly due to its ability to accommodate under- as well as over-dispersed data. For $\mu > 0$ and $\nu > 0$, the COMP probability mass function (p.m.f.) can be written as

$$\tilde{p}_{\mu,\nu}(n) = \frac{\mu^{\nu n}}{\tilde{Z}(\mu,\nu)(n!)^{\nu}},$$

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

 © 2025 SBMAC

where

$$\tilde{Z}(\mu,\nu) := \sum_{n=0}^{\infty} \left(\frac{\mu^n}{n!}\right)^{\nu} \tag{3}$$

is the normalising constant. The sum in (3) is not usually known in closed-form for most values of (μ, ν) and thus needs to be computed approximately.

Consider the situation where one has observed some independent and identically distributed (i.i.d.) data $\boldsymbol{y}$ assumed to come from a COMP distribution with parameters μ and ν and one would like to obtain a posterior distribution $p(\mu,\nu \mid \boldsymbol{y}) \propto f(\boldsymbol{y} \mid \mu,\nu)\pi(\mu,\nu)$. This Bayesian inference problem constitutes a so-called doubly-intractable problem, because neither the normalising constant of the posterior $p(\mu,\nu \mid \boldsymbol{y})$ nor that of the likelihood $f(\boldsymbol{y} \mid \mu,\nu)$ are known. Our approach fits within the so-called noisy algorithms, where the likelihood is replaced by a (noisy) estimate at every step of the MCMC [4]. An added benefit is that an approximation of the likelihood with controlled error will yield an algorithm where in principle one can make the approximation error negligible compared to Monte Carlo error.

In their Figure 5, [1] show that for some values of μ and ν the approximation will take many more than 1000 iterations. In Table 1, we leverage the techniques developed here to provide the exact numbers of iterations needed to achieve a certain tolerance ε using approaches 1 and 2, for the same parameter values considered by [1]. We show that many more iterations than one would normally set are needed in certain contexts, once more highlighting the value of having adaptive algorithms that bypass having to set K.

Table 1: **Numbers of iterations needed to approximate the normalising constant of the COMP**. We show the number n of iterations needed to obtain $|\tilde{Z}(\mu,\nu) - \sum_{x=0}^{n} \tilde{p}_{\mu,\nu}(x)| \leq \varepsilon$ for $\varepsilon = \delta$ and $\varepsilon = 10^6\delta$, where δ is machine precision (given in R by .Machine$double.eps). Results for the Sum-to-threshold and Error-bounding pair approaches are provided.

	$\varepsilon = 2.2 \times 10^{-10}$		$\varepsilon = 2.2 \times 10^{-16}$	
Parameters	Threshold	Error-bounding	Threshold	Error-bounding
$\mu = 10,\ \nu = 0.1$	136	138	186	188
$\mu = 100,\ \nu = 0.01$	1371	1481	1868	1963
$\mu = 100,\ \nu = 0.001$	13725	15661	18692	20410
$\mu = 10000,\ \nu = 0.0001$	137265	164853	186931	211670

References

[1] A. Benson and N. Friel. “Bayesian Inference, Model Selection and Likelihood Estimation using Fast Rejection Sampling: The Conway-Maxwell-Poisson Distribution”. In: **Bayesian Analysis** (2021).

[2] B. Braden. “Calculating sums of infinite series”. In: **The American mathematical monthly** 99.7 (1992), pp. 649–655.

[3] R.W. Conway and W. L. Maxwell. “A queuing model with state dependent service rates”. In: **Journal of Industrial Engineering** 12.2 (1962), pp. 132–136.

[4] P. Alquierand N. Friel, R. Everitt, and A. Boland. “Noisy Monte Carlo: Convergence of Markov chains with approximate transition kernels”. In: **Statistics and Computing** 26.1-2 (2016), pp. 29–47.

© 2025 SBMAC