

Discussion of “Dependent Dirichlet Processes” by Steven N. MacEachern

Miguel de Carvalho*, Luis Gutiérrez Inostroza†, Cristian Capetillo
Constela† and Daniel Saavedra Morales†

We congratulate Professor MacEachern on this influential paper on the celebrated DDP (Dependent Dirichlet Process). The reviews of [Quintana et al. \(2022\)](#) and [Wade and Inácio \(2025\)](#) document its reach, yet extensions for extreme-valued responses remain relatively scarce, and such extensions would be of key interest for modeling risk and for extreme value theory ([Coles, 2001](#); [de Haan and Ferreira, 2006](#); [Resnick, 2007](#); [de Carvalho et al., 2026](#)). This context motivates and shapes the discussion that follows.

Bayesian nonparametric modeling of heavy tails requires an understanding of regular variation ([Bingham et al., 1987](#)) and related notions such as subexponentiality ([Nair et al., 2022](#)). For a distribution F with an unbounded right endpoint, its tail $\bar{F} = 1 - F$ is regularly varying with tail index $\alpha > 0$, written $\bar{F} \in \text{RV}_{-\alpha}$, if, for every $t > 0$,

$$\lim_{y \rightarrow \infty} \frac{\bar{F}(ty)}{\bar{F}(y)} = t^{-\alpha}. \quad (1)$$

For conditional tails, the limit is imposed at each fixed predictor value, allowing the tail index α to vary with the predictor ([Wang and Tsai, 2009](#)). Classical distributions such as the Pareto, Cauchy, and Burr satisfy (1), but which flexible random probability distributions share this property is much less clear in the literature.

To shed light on the tails of the DDP prior, we turn to a 1982 issue of *The Annals of Statistics*. [Doss and Sellke \(1982\)](#) showed that Dirichlet process (DP) draws a.s. have exponentially thinner tails than their centering measure (G_0). This implies that the tails of the DP prior centred around a Cauchy distribution are almost exponential—and hence much lighter than (1)—but G_0 has no mean ([Ghosal and van der Vaart, 2017](#), Example 4.24). By construction the DDP inherits this property at each fixed predictor value, implying that its tails are much lighter than those of the centering distribution, with important implications for modeling heavy tails and extremes. These claims concern the tail of the prior and by conjugacy the tails of the corresponding DP posterior. Interestingly, the tail of the posterior predictive distribution of $\text{DP}(\theta, G_0)$ is:

$$\bar{H}_{n+1}(y) = \Pr(Y_{n+1} > y \mid Y_1, \dots, Y_n) = \theta/(\theta + n) \bar{G}_0(y),$$

for $y > \max\{Y_1, \dots, Y_n\}$. Thus, if $\bar{G}_0 \in \text{RV}_{-\alpha}$, then $\bar{H}_{n+1} \in \text{RV}_{-\alpha}$, for all n ; but, simply retaining the correct tail index does not guarantee accurate probabilities for extremes. Say, if data are iid from $\bar{F}(y) = y^{-\alpha}$ ($y > 1$) and $\bar{G}_0(y) = \bar{F}(y)$, then the predictive

*School of Mathematics, University of Edinburgh, Edinburgh, United Kingdom, miguel.decarvalho@ed.ac.uk

†Department of Statistics, Faculty of Mathematics, Pontificia Universidad Católica de Chile, Santiago, Chile, lgutier@mat.uc.cl; cjcapetillo@uc.cl; dlsaavedra@uc.cl

has the correct tail index but assigns only $\theta/(\theta + n)$ of the true exceedance probability above M_n (e.g., if $\theta = 1$ and $n = 99$, a true probability of 1% becomes 0.01%).

Evidently, mixture models can generate heavy tails through the kernel, the mixing measure, or both. One could argue that a DDP can always be convolved with a heavy-tailed kernel, but the implications for tail behavior are considerably more subtle.

Two comments are in order regarding generating heavy tails through the kernel. [Li et al. \(2019\)](#) show that under mild conditions a fixed kernel and centring specification can yield Dirichlet process location–scale mixtures with a fixed tail index almost surely, even with heavy-tailed kernels. This precludes consistent estimation of the tail index. [Palacios Ramírez et al. \(2025\)](#) show that shape mixtures of Pareto-type kernels can have super-heavy tails when the centring admits tail indices arbitrarily close to zero.

The second route can generate heavy tails by mixing over kernel scales, even with light-tailed kernels. [Palacios Ramírez et al.](#) derive tail envelopes for normalised generalised gamma (NGG; [Lijoi et al., 2007](#)) mixing measures with positive discount via a theoretical characterization involving Lévy processes. Inspired by MacEachern’s paper, they then introduce predictor-dependent stick-breaking weights with shared scale atoms; for related dependent Pitman–Yor processes without addressing heavy tails, see [Leisen and Lijoi \(2011\)](#). A formal characterization of the above-mentioned tail envelopes is given by [Gil-Leyva et al. \(2026\)](#) via M-variation. A tail $\bar{F}$ is M-varying with index $-\alpha$, written $\bar{F} \in \text{MV}_{-\alpha}$, if there exist eventually positive measurable functions $\mathcal{L}, \mathcal{U} \in \text{RV}_{-\alpha}$ and constants $c, C, y_0 > 0$ such that

$$c\mathcal{L}(y) \leq \bar{F}(y) \leq C\mathcal{U}(y), \quad y \geq y_0, \quad \alpha > 0. \quad (2)$$

Regular variation implies M-variation, but (2) need not yield the ratio limit in (1). Using a simple class of transport maps, [Gil-Leyva et al.](#) show that a Pitman–Yor process on $\mathbb{R}_+$ with distribution function G , a centring measure with tail $\bar{G}_0 \in \text{RV}_{-\alpha_0}$ ($\alpha_0 > 0$), discount $D \in (0, 1)$ and strength parameter $\theta > -D$ satisfies $\bar{G} \in \text{MV}_{-\alpha_0/D}$.

Another elegant approach to modeling Pareto-type tails in a DDP context is through logarithmic transformations, by exploiting a folk result from extreme value theory that regularly varying tails become asymptotically exponential on the logarithmic scale. [Saavedra et al. \(2026\)](#) model $T(Y)$ using a linear DDP mixture of Laplace kernels with Bayesian variable selection, where:

$$T(y) = \begin{cases} \log(y - \ell_F), & \ell_F > -\infty, \quad y > \ell_F, \\ \text{arsinh}(y), & \ell_F = -\infty. \end{cases}$$

Here, $\ell_F = \inf\{y : P(Y \leq y) > 0\}$ is the lower endpoint of the response distribution. Back-transformation yields density estimates on the original response scale. Importantly, posterior consistency on the transformed scale carries over under this fixed back-transformation. Finally, while computationally appealing and well grounded, at this stage this approach offers no formal guarantees for extreme-value extrapolation.

Professor MacEachern’s framework offers flexibility across predictors. The open challenge is to develop flexible conditional models whose tails are compatible with the extrapolation principles of extreme value theory, enabling inference beyond the range of observed data. Which classes of predictor-dependent random probability measures, or models built from them, can meet this challenge?

References

- Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987). *Regular Variation*. Cambridge: Cambridge University Press.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. London: Springer.
- de Carvalho, M., Huser, R., Naveau, P., and Reich, B. J. (eds.) (2026). *Handbook of Statistics of Extremes*. Chapman & Hall/CRC.
- de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. New York: Springer.
- Doss, H. and Sellke, T. (1982). The Tails of Probabilities Chosen from a Dirichlet Prior. *The Annals of Statistics* 10, 1302–1305.
- Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge: Cambridge University Press.
- Gil-Leyva, M. F., Palacios Ramírez, V., and de Carvalho, M. (2026). On the Tails of Pitman–Yor Random Probability Measures: Transport Maps and Stick-Breaking Constructions. *Extremes*. Advance publication.
- Leisen, F. and Lijoi, A. (2011). Vectors of two-parameter Poisson–Dirichlet processes. *Journal of Multivariate Analysis* 102, 482–495.
- Li, C., Lin, L., and Dunson, D. B. (2019). On Posterior Consistency of Tail Index for Bayesian Kernel Mixture Models. *Bernoulli* 25, 1999–2028.
- Lijoi, A., Mena, R. H., and Prünster, I. (2007). Controlling the reinforcement in Bayesian non-parametric mixture models. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 69, 715–740.
- MacEachern, S. N. (2026). Dependent Dirichlet Processes (with Discussion). *Bayesian Analysis*. doi:10.1214/26-BA1610.
- Nair, J., Wierman, A., and Zwart, B. (2022). *The Fundamentals of Heavy Tails: Properties, Emergence, and Estimation*. Cambridge: Cambridge University Press.
- Palacios Ramírez, V., de Carvalho, M., and Gutiérrez, L. (2025). Heavy-Tailed NGG-Mixture Models. *Bayesian Analysis* 20, 1315–1343.
- Quintana, F. A., Müller, P., Jara, A., and MacEachern, S. N. (2022). The Dependent Dirichlet Process and Related Models. *Statistical Science* 37, 24–41.
- Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. New York: Springer.
- Saavedra, D., Capetillo, C., and Gutiérrez, L. (2026). Bayesian Model Selection for Heavy-Tailed Data via Transformations. Submitted.
- Wade, S. and Inácio, V. (2025). Bayesian Dependent Mixture Models: A Predictive Comparison and Survey. *Statistical Science* 40, 81–108.
- Wang, H. and Tsai, C.-L. (2009). Tail Index Regression. *Journal of the American Statistical Association* 104, 1233–1240.