

Why Pitman-Yor?

Generalizing the Dirichlet Process

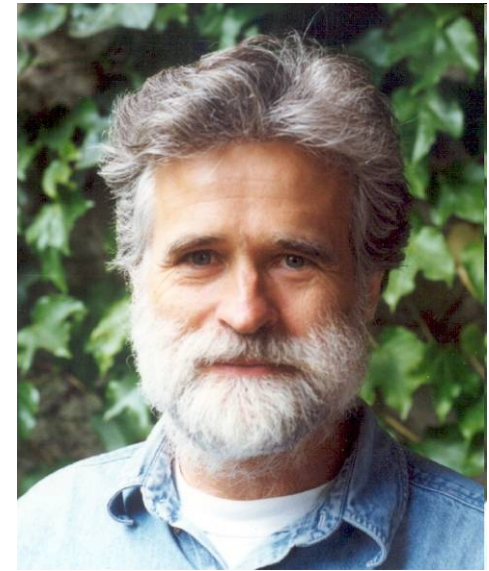
- Distribution on partitions leads to a generalized *Chinese restaurant process*
- Special cases of interest in probability: Markov chains, Brownian motion, ...

Power Law Distributions

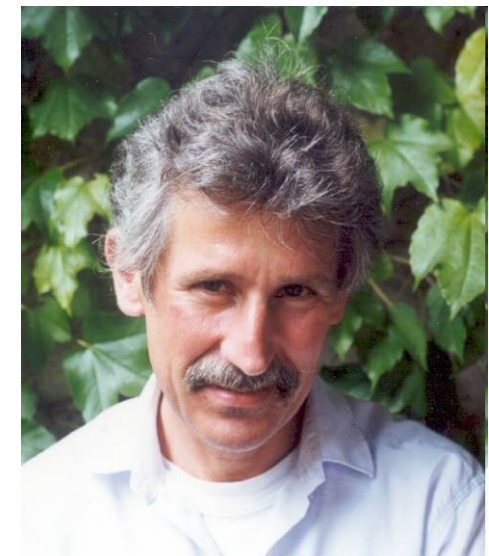
	DP	PY
Number of unique clusters in N observations	$\mathcal{O}(b \log N)$	Heaps' Law: $\mathcal{O}(bN^a)$
Size of sorted cluster weight k	$\mathcal{O}\left(\alpha_b \left(\frac{1+b}{b}\right)^{-k}\right)$	Zipf's Law: $\mathcal{O}\left(\alpha_{ab} k^{-\frac{1}{a}}\right)$

**Natural Language
Statistics**

Goldwater, Griffiths, & Johnson, 2005
Teh, 2006



Jim Pitman



Marc Yor