Stick-breaking priors: The Pitman-Yor process and randomized generalized Gamma models.

Lancelot F. James (lancelot@ust.hk)

Hong Kong University of Science and Technology

Combinatorial Stochastic Processes
A conference in celebration of Jim Pitman's work
June 20, 2014

In around 2000/2001 we wrote the following two papers

1. Ishwaran, Hemant, and Lancelot F. James. "Gibbs sampling methods for stick-breaking priors." Journal of the American Statistical Association 96.453 (2001).
2. Ishwaran, Hemant, and Lancelot F. James. "Generalized weighted Chinese restaurant processes for species sampling mixture models." Statistica Sinica 13.4 (2003): 1211-1236.

1. A large part of this was based on the works of Jim, Marc Yor and co-authors, and also Albert Lo. Our intent was to show how such things could be practically implemented/used in rather complicated statistical models,
2. It is often said that this(the 2 papers mentioned) was the basis for a lot of subsequent work in the Bayesian Statistical Machine Learning literature
3. We understood very little about the broader CSP literature
4. We understood almost nothing in Pitman and Yor (1997)
5. This was done prior to real communication with Jim
6. The current work has its genesis in my first real conversations with Jim back in 2001

Outline

1. Background
   - Basic Definitions: mass Partitions, random cdf's, Pitman-Yor process, α-stable Poisson Kingman laws
   - Stick-breaking reps derived from species sampling models (size biased sampling)
   - What is explicitly known? Surprisingly not much

Outline

1. What is new
   - Randomized generalized gamma model with law $\mathrm{PG}(\alpha,\zeta) := \mathrm{PG}^{[0]}(\alpha,\zeta)$
   - Appears in Proof of Prop 21 Pitman-Yor (1997)
   - Stick-breaking results for the basic randomized generalized gamma model -
   - A much larger class -$\mathrm{PG}^{[\varrho]}(\alpha,\zeta)$ for $-\infty < \varrho < \infty$
     - some relations
     - Note: I have SB reps for all (ϱ,ζ) however requires a bit to explain,(omitted)
   - Stick-breaking reps for $\varrho = 1$

Consider the space of (mass partitions) ranked probabilities

$$\mathcal{P}_\infty = \{(Q_k) : Q_1 \geq Q_2 \geq \ldots \geq 0; \sum_{k=1}^{\infty} Q_k = 1\}$$

Let (P_k) denote a random sequence in $\mathcal{P}_\infty$ having some law $\mathbb{P}$. Independent of this, use a sequence of iid $\mathrm{Uniform}[0, 1]$ variables to form a random cumulative distribution function

$$F(y) = \sum_{k=1}^{\infty} P_k \mathbb{I}_{\{U_k \leq y\}} \stackrel{?}{=} \sum_{k=1}^{\infty} \tilde{P}_k \mathbb{I}_{\{U_k \leq y\}}$$

- One can call F a $\mathbb{P}$-bridge, which constitutes a prior law on the space of distribution functions
- Outside BNP/ML one often works directly with random $(P_k) \in \mathcal{P}_\infty$
- While in BNP/ML one works with F
- BNP=Bayesian non-parametric statisitcs, ML=Statistical Machine Learning.

In Bayesian Non-parametric Statistics and Bayesian statistical Machine Learning/Natural Language Processing the most popular choices for F are

1. $F := P_{0,\theta}$ is a Dirichlet Process: $(P_k) \sim \mathrm{PD}(0, \theta)$–Poisson Dirichlet$(\theta)$
2. $F := P_{\alpha,\theta}$ is a Pitman-Yor Process : $(P_k) \sim \mathrm{PD}(\alpha, \theta)$–2 parameter Poisson Dirichlet
3. $F := Q_{\alpha,\zeta}$ is a Normalized Generalized Gamma process (NGG). Note here ζ is a non-negative number.
4. Note" The name Pitman-Yor process was coined in the paper of Ishwaran, H., and L. F. James. "Gibbs sampling methods for stick-breaking priors." Journal of the American Statistical Association 96.453 (2001).

1. Applications often involve rather intricate modifications of these quantities which exploit the nice explicit properties of these cases. Many recent applications exploit the explicit stick-breaking representations of these processes.
2. In this sense the NGG lags behind as there is less-known about its structure. That is, prior to this work, there is no known explicit stick-breaking representation for this case for general α.

α-stable Poisson Kingman Priors: Distributions over $\mathcal{P}_\infty$

1. Let for $0 < \alpha < 1$, $(S_\alpha(y) : 0 \leq y < \infty)$ denote a α stable subordinator such that for each fixed y, $S_\alpha(y) \stackrel{d}{=} y^{1/\alpha} S_\alpha(1)$ where $S_\alpha(1)$ is a stable random variable with generally complicated density f_α and otherwise characterized by a remarkably simple Laplace transform

$$\mathbb{E}[\mathrm{e}^{-\lambda S_\alpha(1)}] = \mathrm{e}^{-\lambda^\alpha}$$

1. Set $T = S_\alpha(1) = \sum_{k=1}^{\infty} J_k$, where $J_1 > J_2 > \ldots$ represent the ranked jumps of the α stable subordinator.
2. Set $P_k = J_k/T$ for $k = 1, 2, \ldots$
3. Then the distribution of (P_k) is denoted as $\mathrm{PD}(\alpha, 0)$
4. The corresponding random distribution function for $0 \leq y \leq 1$

$$F(y) = \frac{S_\alpha(y)}{T} := \sum_{k=1}^{\infty} P_k \mathbb{I}_{\{U_k \leq y\}}$$

is sometimes referred to as a Pitman-Yor process with parameters $(\alpha, 0)$

Conditioning on T and then mixing

1. Condition (P_k), or equivalently F, on $T = t$
2. Then the conditional distribution of $(P_k)|T = t$ is denoted as $\mathrm{PD}(\alpha|t)$
3. Set $T = T_h$ where now T_h has density[can be discrete]

$$h(t)f_\alpha(t)$$

where $\mathbb{E}[h(S_\alpha(1))] = \mathbb{E}[h(T)] = 1.$

4. Then the law of (P_k), also F, is given by

$$PK_\alpha(h \cdot f_\alpha) = \int_0^\infty \mathrm{PD}(\alpha|t)h(t)f_\alpha(t)dt$$

That is a Poisson-Kingman distribution generated by an α-stable subordinator with mixing random variable T_h.

1. Basic idea, for all subordinators $S(y)$ in place of $S_\alpha(y)$, can be attributed to Kingman (1975) JRSSB. Later formalized and developed in Pitman (2002)
2. Choices of $h(t) = Ct^{-\theta}$, $\theta > -\alpha$ arise from a body of work of Pitman and Yor (1980s to 1990's) and co-authors related to investigations of the lengths of excursions of generalizations of Brownian Motion(BM)
3. BM corresponds to $\alpha = 1/2$ and $h(t) = 1$. Brownian Bridge Corresponds to $h(t) = Ct^{-1/2}$ and $\alpha = 1/2$
4. The Pitman-Yor process, $F := P_{\alpha,\theta}$, with parameters α, θ corresponds to $h(t) = Ct^{-\theta}$, $\theta > -\alpha$.
5. $(P_k) \sim PD(\alpha, \theta)$ is said to have a two-parameter Poisson Dirichlet Distribution

1. Taking limits $\mathrm{PD}(0, \theta)$ correspond to the Poisson Dirichlet distribution[see Kingman (poisson Process Book), Arratia Barbour Tavare: Logarithmic Combinatorial Analysis Book.]Connected to Ewens Sampling Formula and the Dirichlet Process which is now the distribution of F

Consider the following species sampling scheme(Generalized Residual Allocation Model) which can be read from Pitman (1996): For (U'_k) a sequence of iid $\mathrm{Uniform}[0,1]$ variables

1. Let $Y_i = F^{-1}(U'_i), i = 1, \ldots, n$ denote an exchangeable sample from F representing tags for up to n distinct species
2. Think Generalized Blackwell-MacQueen Schemes.
3. Since F is discrete there are K_n distinct species with iid tags $(\hat{U}_1, \ldots, \hat{U}_{K_n})$. This forms a partition of $[n] = \{1, \ldots, n\}$ by the relation $A_k = \{i : Y_i = \hat{U}_k\}$
4. A_1 represents the number of animals that have the same tag as the first observed animal. A_k have similar interpretations

1. Let $N_{k,n} = |A_k|$, then as $n \to \infty$, the limiting frequencies of the k-th species in order of appearance is almost surely

$$\lim_{n\to\infty} \frac{N_{k,n}}{n} = \tilde{P}_k := (1 - W_k) \prod_{l=1}^{k-1} W_l$$

2. The stick-breaking sequence $(\tilde{P}_k)$ coincides with the size biased re-arrangement of the ordered sequence (P_k) and

$$F(y) = \sum_{k=1}^{\infty} P_k \mathbb{I}_{\{U_k \le y\}} = \sum_{k=1}^{\infty} \tilde{P}_k \mathbb{I}_{\{U_k \le y\}}$$

3. $\tilde{P}_1 = 1 - W_1$, is the first size-biased pick and has the structural distribution of (P_l) or F.

Structural distributions and Deletion operations

1. $\tilde{P}_1 = 1 - W_1$, is the first size-biased pick and has the structural distribution of (P_l) or F.
2. Notice that the sequence below has total sum 1

$$(\tilde{P}_k/W_1)_{k\geq 2} = ((1-W_2), (1-W_\ell)\prod_{l=2}^{\ell-1} W_l; \ell = 3, \ldots)$$

   use these to form a bridge F_1 [This is the deletion operation]
3. There is the decomposition

$$F(y) = W_1 F_1(y) + (1-W_1)\mathbb{I}_{\{U_1 \leq y\}}$$

   [Recovering F from F_1 is the insertion operation]

1. In fact $(1 - W_k)$ has the structural distribution of a sequence (F_{k-1}) of cdfs formed by a process of deletion, constituting a Markov Chain.
2. Insertion is reflected in the decomposition

$$F_{k-1}(y) = W_k F_k(y) + (1 - W_k)\mathbb{I}_{\{U_k \le y\}}$$

Posterior distribution of F given 1 observation

1. Can set $Y_1 = U_1 = p$ then the posterior distribution of $F|Y_1 = p$ is equivalent to the law of the random cdf

$$W_1 F_1(y) + (1 - W_1)\mathbb{I}_{\{p \leq y\}}$$

Limitations:Independence is an exception

1. The W_k are independent if and only if $1 - W_k$ are $\mathrm{Beta}(1-\alpha, \theta + k\alpha)$ for $0 \leq \alpha < 1$ and $\theta > -\alpha$
2. That is the ranked frequencies $(P_k) \sim \mathrm{PD}(\alpha, \theta)$.
3. equivalently

$$F(y) := P_{\alpha,\theta}(y) = \sum_{k=1}^{\infty} \tilde{P}_k \mathbb{I}_{\{U_k \leq y\}}$$

   is a Pitman-Yor Process with parameters (α, θ).
4. The stick-breaking in the $\mathrm{PD}(\alpha, \theta)$ case for $\alpha \neq 0$ is due to Perman, Pitman and Yor (1992).
5. Griffiths, Engen, McCloskey, and Sethuraman in the $\mathrm{PD}(0, \theta)$ case

Generic Stick-breaking vs GRAM

1. Stick-breaking sometimes refers more loosely to the idea that one can represent a distribution function F in terms of some probability masses $(1-W_k)\prod_{l=1}^{k-1} W_l$ that are hopefully tractable. In this sense the weights can be taken iid from a variety of distributions on $[0,1]$ as long as a tail condition is satisfied
2. Note: However, if the W_k are independent with another distribution besides the Beta we described, they cannot arise from the species sampling scheme described above, hence have no immediate interpretation.
3. Note Interpretation is important for potential applications rather than mere simulation.
4. We are interested in very EXPLICIT descriptions of the weights in stick-breaking reps arising from such schemes. Hence the task becomes much more difficult.

What is Known beyond $\mathrm{PD}(\alpha, \theta)$: $1/2$-stable Poisson Kingman -Aldous and Pitman (1998)

1. Perman, Pitman and Yor (1992) provide general joint density formula for Poisson-Kingman models. This serves mostly as an essential guideline(blueprint) and one needs to do a bit of work to get explicit descriptions of relevant random variables.
2. Note even simple products of what should be recognizable random variables oftentimes produce nasty densities.
3. Moreover one has to be able to describe appropriately the dependence between variables.

1. Known cases that have been represented explicitly in terms of tractable Random Variables
   - $\mathrm{PD}(\alpha, \theta)$
   - $\mathrm{PD}(1/2|t)$ hence an SB for any $1/2$-stable PK model with some mixing variable T_h
   - For $\alpha = 1/2$, the corresponding size-biased arrangement of $(P_k)|T = t$ for all t, say $(\tilde{P}_k)|T = t$ is described explicitly in Corollary 5 of Aldous and Pitman (1998) *The Standard Additive Coalescent*. See also Pitman (2002) *Poisson Kingman Partitions* Proposition 14.
2. No results for other non-trivial α-stable PK models
3. We know there are no more cases where the W_k are independent.
4. No explicit results for PK models defined by other subordinators.

Let S_α denote a positive stable random variable with density f_α, and Laplace transform

$$\mathbb{E}[\mathrm{e}^{-\lambda S_\alpha}] = \mathrm{e}^{-\lambda^\alpha}$$

For a fixed non-negative number ζ, Let $T_0(\zeta) := \tau_\alpha(\zeta)/\zeta^{1/\alpha}$ denote the infinitely divisible random variable having density

$$\mathrm{e}^{-[\zeta^{1/\alpha}t-\zeta]}f_\alpha(t)$$

$(\tau_\alpha(\zeta y) : 0 \le y < \infty)$ is a generalized gamma subordinator and so is $\tau_{\alpha,\zeta}(y) := \tau_\alpha(\zeta y)/\zeta^{1/\alpha}$

Normalized Generalized Gamma process

For $0 \le y \le 1$, an $\mathrm{NGG}(\zeta)$ can be represented as

$$F(y) := Q_{\alpha,\zeta}(y) = \frac{\tau_\alpha(\zeta y)}{\tau_\alpha(\zeta)} = \frac{\tau_{\alpha,\zeta}(y)}{\tau_{\alpha,\zeta}(1)}$$

1. Note $T_0(\zeta) \stackrel{d}{=} \tau_{\alpha,\zeta}(1)$
2. Furthermore setting $S = T_0^{-\alpha}(\zeta)$ it follows from Pitman (2002) that

   $$n^{-\alpha} K_n \to S$$

   almost surely as $n \to \infty$.
3. That is to say S is the α-diversity which is identified in Pitman and Yor (1997)

Randomization

1. When ζ is fixed this GRAM model was studied by McCloskey (1965). Arises in many BNP/ML applications.
2. When ζ is a non-negative random variable this class for (P_k) becomes quite large and we refer to its law as $\mathrm{PG}(\alpha, \zeta)$.
3. This class is also due to Pitman and Yor (1997).
4. It appears in their proof of Prop 21
5. Prop 21 paraphrase: $\zeta := G_{\theta/\alpha} \sim \mathrm{Gamma}(\theta/\alpha, 1)$.

$$Q_{\alpha, G_{\theta/\alpha}} = P_{\alpha,\theta} \Leftrightarrow \mathrm{PG}(\alpha, G_{\theta/\alpha}) = \mathrm{PD}(\alpha, \theta),$$

$\theta \geq 0$

6. It is $\mathrm{PK}_\alpha(h \cdot f_\alpha)$ for $h(t) = \mathbb{E}[\mathrm{e}^{-[\zeta^{1/\alpha} t - \zeta]}]$

New: Stick-breaking Representation for $\mathrm{PG}(\alpha, \zeta)$

1. Let $(\beta^{(k)}_{1-\alpha,\alpha})$ denote a sequence of iid $\mathrm{Beta}(1-\alpha, \alpha)$. variables. Independent of this:
2. let (e_i) denote a sequence of iid exponential (1) variables and let $\tilde{G}_k = \sum_{i=1}^{k} \mathrm{e}_i$ denote partial sums for $k = 1, 2, \ldots$
3. Define a sequence of generally dependent random variables (R_k), taking values in $[0, 1]$, specified by

$$R_k = \left(\frac{\zeta + \tilde{G}_{k-1}}{\zeta + \tilde{G}_k} \right)^{1/\alpha}$$

$\tilde{G}_0 := 0.$

4. Then the size-biased frequencies $\tilde{P}_k = (1 - W_k) \prod_{i=1}^{k-1} W_i$ are such that

$$(1 - W_k) = \beta^{(k)}_{1-\alpha,\alpha}[1 - R_k]$$

Recovering $\mathrm{PD}(\alpha, \theta)$, $\theta \geq 0$

1. From Prop 21 Pitman-Yor (1997) set $\zeta \overset{d}{=} G_{\theta/\alpha}$ for $\theta > 0$, a gamma variable with shape θ/α. Then $\mathrm{PG}(\alpha, G_{\theta/\alpha}) = \mathrm{PD}(\alpha, \theta)$ for $\theta \geq 0$
2. For this choice of ζ, the (R_k) are independent $\mathrm{Beta}(\theta + (k-1)\alpha, 1)$

$$R_k = \left(\frac{G_{\theta/\alpha} + \tilde{G}_{k-1}}{G_{\theta/\alpha} + \tilde{G}_{k-1} + \mathrm{e}_k} \right)^{1/\alpha}$$

3. Note that for $\theta = 0$, $R_1 = 0$

1. It follows that for $\theta \geq 0$

$$1 - W_k = \beta^{(k)}_{1-\alpha,\alpha}[1 - R_k] = \beta^{(k)}_{1-\alpha,\alpha}\beta_{1,\theta+(k-1)\alpha} = \beta_{1-\alpha,\theta+k\alpha}$$

2. Note $\mathrm{PG}(\alpha, G_{\theta/\alpha})$ does not include $\mathrm{PD}(\alpha, \theta)$ for $-\alpha < \theta < 0$.
3. Also notice that $\mathrm{PG}(\alpha, G_{\theta/\alpha} + \mathrm{e}_1)$ never includes $\mathrm{PD}(\alpha, 0)$.
4. So we seek larger classes. Later.

Randomization arises naturally in the posterior and via deletion

1. Note that the cdf F_1 obtained by the deletion operation is for fixed or random ζ a

$$\mathrm{PG}(\alpha, \zeta + \mathrm{e}_1) = \mathrm{PG}(\alpha, \zeta / R_1^{\alpha}) - \mathrm{bridge}$$

2. Recall that $F(y) = W_1 F_1(y) + (1 - W_1)\mathbb{I}_{\{U_1 \leq y\}}$
3. F_1 is generally not independent of W_1 but that does not matter much.
4. More generally the sequence (F_{k-1}) are nested

$$\mathrm{PG}(\alpha, \zeta + \sum_{l=1}^{k-1} \mathrm{e}_l) = \mathrm{PG}(\alpha, \zeta / R_{k-1}^{\alpha})$$

bridges representing the states of a Markov Chain formed by deletion[set $R_0 = 1$]

1. Set $\zeta = G_{\theta/\alpha}$ to recover the $\mathrm{PD}(\alpha, \theta)$ cases
2. Note: There is a reversal by insertion. Need a coherent description of this operation that does not rely on independence. Independence only arises in the case of $\mathrm{PD}(\alpha, \theta)$.

How to interpret $(\zeta, (R_k))$? Another relation to Pitman and Yor (1997)

1. Suppose for a moment that $\zeta = \mathrm{e}_0$ an exponential (1) variable.
2. Now as in Pitman and Yor (1997) suppose $(V_k) \sim \mathrm{PD}(\alpha, 0)$
3. Then using the same notation as in PY(1997),

$$R_k = \frac{V_{k+1}}{V_k} = \frac{\Delta_{k+1}}{\Delta_k}$$

   where $\Delta_1 > \Delta_2 > \ldots$ are the ranked jumps of an α-stable subordinator. In particular $V_k := \Delta_k / \sum_{i=1}^{\infty} \Delta_i$.
4. Set the maximal jump $\Delta_1^{-\alpha} = \zeta$ then the conditional distribution of $(R_k)|\zeta$ appearing in the stick-breaking representation agrees with $(R_k)|\Delta_1^{-\alpha}$.

How to interpret $(\zeta, (R_k))$? Another relation to Pitman and Yor (1997)

1. From PY(97) Lemma 24 [(iii)], set for $k = 1, 2, \ldots$

$$S_k = \frac{\Delta_{k+1}}{\Delta_1} = \prod_{l=1}^{k} R_l \in [0, 1]$$

Then conditional on $\Delta_1 := \zeta^{-1/\alpha}$ the (S_k) are the ranked jumps of a subordinator specified by the Lévy density

$$\rho_\alpha(s|\zeta) = \frac{\zeta\alpha}{\Gamma(1-\alpha)} s^{-\alpha-1} \mathbb{I}_{\{0<s<1\}}$$

2. $\rho_\alpha(s|\zeta)$ should look (somewhat) familiar to the Indian Buffet process crowd.

1. When $(V_k) \sim \mathrm{PD}(\alpha, 0)$, $R_k \sim \beta_{k\alpha,1}$
2. $\Delta_1^{-\alpha} = \zeta := \mathrm{e}_0$
3. $1 - W_k := \beta^{(k)}_{1-\alpha,\alpha}[1 - R_k] \sim \beta_{1-\alpha,(k+1)\alpha}$ (independent)
4. Hence $(P_k) \sim \mathrm{PD}(\alpha, \alpha)$
5. See PY(97) Section 7 for related points.

In general

1. When $(V_k) \sim \mathrm{PD}(\alpha, 0)$, condition $(V_k)|\Delta_1^{-\alpha} = \zeta$
2. $R_k = V_{k+1}/V_k = \Delta_{k+1}/\Delta_k$
3. Form the sequence $1 - W_k = \beta^{(k)}_{1-\alpha,\alpha}[1 - R_k]$
4. Set $\tilde{P}_k = (1 - W_k) \prod_{j=1}^{k-1} W_j$
5. Randomize ζ according to the distribution of some non-negative variable.
6. Then the ranked re-arrangement of $(\tilde{P}_k)$, say (P_k), is $\mathrm{PG}(\alpha, \zeta)$
7. Note $\lim_{\Delta_1 \to \infty} R_k = \beta_{(k-1)\alpha,1}$

What other classes can we get explicit stick-breaking reps?

1. For $-\infty < \varrho < \infty$ Let $\mathrm{PG}^{[\varrho]}(\alpha, \zeta)$ be a class of α-stable Poisson Kingman priors so that when conditioned on ζ the mixing variable $T_h := T_\varrho(\zeta)$ has density of the form

$$\frac{t^{\varrho} e^{-[\zeta^{1/\alpha} t - \zeta]}}{\mathbb{E}[T_0^{\varrho}(\zeta)|\zeta]} f_\alpha(t) := h(t|\zeta) f_\alpha(t)$$

2. Note the collection of distributions say $\mathcal{D}_\varrho$, are considered for each fixed ϱ and indexed over all possible distributions for ζ. This creates an ordering in ϱ
3. $\mathcal{D}_{\varrho_1} \subset \mathcal{D}_{\varrho_2}$ for $\varrho_1 < \varrho_2$
4. $\mathrm{PG}^{[0]}(\alpha, \zeta) = \mathrm{PG}(\alpha, \zeta)$

$\mathrm{PG}^{[\varrho_2]}(\alpha, ?) = \mathrm{PG}^{[\varrho_1]}(\alpha, \zeta)$ for $\varrho_1 < \varrho_2$

1. For fixed or random ζ, consider the random variable

$$Y_{\varrho_2,\varrho_1}(\zeta) = \frac{G_{\varrho_2-\varrho_1}}{\zeta^{1/\alpha} T_{\varrho_1}(\zeta)}$$

2. Define $\zeta_{\varrho_2,\varrho_1}^{1/\alpha} := \zeta^{1/\alpha}(1 + Y_{\varrho_2,\varrho_1}(\zeta))$
3. Then $\mathrm{PG}^{[\varrho_2]}(\alpha, \zeta_{\varrho_2,\varrho_1}) = \mathrm{PG}^{[\varrho_1]}(\alpha, \zeta)$
4. $T_{\varrho_2}(\zeta_{\varrho_2,\varrho_1}) \overset{d}{=} T_{\varrho_1}(\zeta)$
5. Note $\zeta_{1,0} \overset{d}{=} \mathrm{e}_1 + \zeta = \zeta / R_1^{\alpha}$
6. Interested particularly in the cases where $\varrho_2 = m = 0, 1, 2, \ldots$

Representing $\mathrm{PD}(\alpha, \theta)$

1. $\mathrm{PG}^{[-\theta]}(\alpha, 0) = \mathrm{PD}(\alpha, \theta)$ for each fixed $\theta > -\alpha$.
2. This is the minimal order $-\theta$ representation
3. $\mathrm{PG}^{[\varrho]}(\alpha, ?) = \mathrm{PD}(\alpha, \theta)$ for each $\varrho \geq -\theta$

Representing $\mathrm{PD}(\alpha, \theta)$ in the m-class

1. For an integer $m = 1, 2. \ldots$ let K_m denote the number of blocks produced from a partition of $[m] = \{1, 2, \ldots, m\}$ following a $\mathrm{PD}(\alpha, \theta) - \mathrm{EPPF}$
2. Let $G_{\theta/\alpha + K_m}$ be such that given $K_m = \ell$ it is a Gamma$(\theta/\alpha + \ell)$ variable
3. Recall that $\mathrm{PG}^{[-\theta]}(\alpha, 0) = \mathrm{PD}(\alpha, \theta)$, then there is the distributional identity

$$\lim_{\zeta \to 0} \zeta_{m,-\theta}^{1/\alpha} = \frac{G_{\theta+m}}{S_{\alpha,\theta}} = G_{\theta/\alpha + K_m}^{1/\alpha}$$

$S_{\alpha,\theta}$ has density proportional to $t^{-\theta} f_\alpha(t)$. Note the last distributional equality in NOT obvious.

Representing $\mathrm{PD}(\alpha, \theta)$ in the m-class

1. For each $m = 1, 2, \ldots,$ and each fixed $\theta > -\alpha$,

$$\mathrm{PG}^{[m]}(\alpha, G_{\theta/\alpha + K_m}) = \mathrm{PD}(\alpha, \theta)$$

2. When $m = 0$ and $\theta \geq 0$ this is Pitman and Yor (1997) Prop 21

$\mathrm{PG}^{[1]}(\alpha, G_{\theta/\alpha+1}) = \mathrm{PD}(\alpha, \theta)$, $\theta > -\alpha$

1. When $m = 1$, $\lim_{\zeta \to 0} \zeta_{1,-\theta} = G_{\theta/\alpha+1}$, think the coagulation operation of Dong, Goldschmidt and Martin; and Bertoin and Goldschmidt when $\alpha \to 0$. [BDGM]
2. But this is encoded in the decomposition of the time spent positive of a generalized skewed Bessel Bridge (with skew parameter y) due to Pitman and Yor

$$P_{\alpha,\theta}(y) = \beta_{\theta+\alpha,1-\alpha} P_{\alpha,\theta+\alpha}(y) + (1 - \beta_{\theta+\alpha,1-\alpha}) \mathbb{I}_{\{U_1 \le y\}}$$

1. Which can be re-written in terms of a composition of two bridges (random cdfs).

$$P_{\alpha,\theta}(y) = P_{\alpha,\theta+1}(b_1(y)) := Q_{\alpha,G_{\frac{\theta+1}{\alpha}}}(b_1(y))$$

where $b_1(y) = \beta_{\frac{\theta+\alpha}{\alpha},\frac{1-\alpha}{\alpha}} y + (1 - \beta_{\frac{\theta+\alpha}{\alpha},\frac{1-\alpha}{\alpha}})\mathbb{I}_{\{U_1 \le y\}}$

$\mathrm{PG}^{[1]}(\alpha,\zeta)$ SB results

$F \sim \mathrm{PG}^{[1]}(\alpha,\zeta)$ then

$$F(y) = Q_{\alpha,\zeta+G_{\frac{1-\alpha}{\alpha}}}(\tilde{b}_1(y))$$

where

$$\tilde{b}_1(y) = \frac{\zeta}{\zeta + G_{\frac{1-\alpha}{\alpha}}}\mathbb{U}(y) + \frac{G_{\frac{1-\alpha}{\alpha}}}{\zeta + G_{\frac{1-\alpha}{\alpha}}}\mathbb{I}_{\{U_1 \le y\}}$$

(P_k) can be described in terms of the Coag operator of BDGM.

For its stick-breaking representation The (W_k) are generally dependent random variables represented as

$$W_k = [\beta^{(k)}_{1-\alpha,\alpha}[(1 - R_k)/R_k] + 1]^{-1}$$

Recovering $\mathrm{PD}(\alpha, \theta) = \mathrm{PG}^{[1]}(\alpha, G_{\theta/\alpha+1}), \theta > -\alpha$

1. In order to recover the stick-breaking representation for $\mathrm{PD}(\alpha, \theta)$ for all $\theta > -\alpha$, set $\zeta = G_{\theta/\alpha+1}$.
2. For this choice of ζ, the (R_k) are independent $\mathrm{Beta}(\theta + k\alpha, 1)$

$$R_k = \left(\frac{G_{\theta/\alpha+1} + \tilde{G}_{k-1}}{G_{\theta/\alpha+1} + \tilde{G}_{k-1} + \mathrm{e}_k} \right)^{1/\alpha}$$

Stick-breaking the general $\mathrm{PG}^{[\varrho]}(\alpha, \zeta)$ cases

1. I have, rather nice representations in the most general cases
2. They look in form like the case $m = 1$, except now one sees things reminiscent of the beta-binomial relationship between the structural W_1 and the size of the block containing 1. See Pitman (1995)
3. Furthermore the relation between variables playing the role of (R_k) is more complicated

General Strategy

1. Obtain results for $\mathrm{PG}^{[m]}(\alpha, \zeta)$ for any ζ and arbitrary m
2. Use the relation $\mathrm{PG}^{[m]}(\alpha, \zeta_{m,\varrho}) = \mathrm{PG}^{[\varrho]}(\alpha, \zeta)$ for any $m \geq \varrho$

Thank You
and
Thank You Jim!