Random permutations and the two-parameter Poisson-Dirichlet distribution.

Sasha Gnedin

Queen Mary, University of London

Notes on random ~~compositions~~ left-regular compositions. J.P. 7/19/02 (1)

Let (Q_n) be a sequence of laws on the set $\mathcal{C}_n$ of compositions of n. Then we can consider

a First class deletion property (FCDP)

a Last class . . . . (LCDP)

a random . . . . (RCDP).

The last means removal of X_J from $(X_1, \dots, X_K)$ for suitable random J. Various forms of different strengths depending on how much conditioning is allowed (X_J, or (J, X_J) etc., J could be uniform on $[1, K]$ or size-biased by the X_j.

~~Let X~~ Note that these concepts make sense without assuming the (Q_n) are consistent in any other way.

Let $X_1, X_2, \dots$ be iid with values in $\{1, 2, \dots\}$.

Let $S_k = X_1 + \dots + X_k$.

$$Q_n = \text{Law}\left(X_1, \dots, X_{K_n} \,\middle|\, \overbrace{S_k = n \text{ for some } k = 1, 2, \dots}^{\text{Renewal at } n}\right)$$

where K_n is the index k such that $S_k = n$.

Then, by elementary renewal theory, (Q_n) ~~is ... exchangeable~~ has the FCDP & LCDP

In fact Q_n is an exchangeable comp of n, each n.

These compositions of n are of Gibbs type.

Probability function is

$$g(n_1, \dots, n_k) = \frac{\prod_{i=1}^{k} p_{n_i}}{P(\text{renewal at } n)}$$

Need to assume $p_n > 0$ all n, else may not be well defined. Call these RENEWAL COMPOSITIONS

Notes on random ~~left regenerative~~ compositions. J.P. 7/19/02 (1)

Let (Q_n) be a sequence of laws on the set $\mathcal{C}_n$ of compositions of n. Then we can consider

a First class deletion property (FCDP)

a Last class . . . . (LCDP)

a random . . . . (RCDP).

The last means removal of X_J from $(X_1, \cdots, X_K)$ for suitable random J. Various forms of different strengths depending on how much conditioning is allowed $(X_J$, or (J, X_J) etc, J could be uniform on $[1,K]$ or size-biased by the X_j.

~~Let X~~ Note that these concepts make sense without assuming the (Q_n) are consistent in any other way.

Let $X_1, X_2, \cdots$ be iid with values in $\{1,2,\cdots\}$.

Let $S_k = X_1 + \cdots + X_k$.

$$Q_n = \text{Law}(X_1, \cdots, X_{K_n} \mid \overbrace{S_k = n \text{ for some } k=1,2\cdots}^{\text{Renewal at } n})$$

where K_n is the index k such that $S_k = n$.

Then, by elementary renewal theory, (Q_n) ~~is ...~~ has the FCDP & LCDP

In fact Q_n is an exchangeable comp of n, each n.

These compositions of n are of Gibbs type.

Probability function is

$$g(n_1, \cdots, n_k) = \frac{\prod_{i=1}^{k} p_{n_i}}{P(\text{renewal at } n)}$$

Need to assume $p_n > 0$ all n, else may not be well defined. Call these RENEWAL COMPOSITION

Seems fine until ...

Should check it for α, α,

Compute the law of n_1 e.g. of the bridge, and see it sums to 1.

Maybe some ~~xxx~~ algebra mistake here.

Compare Proposition 8 of

Notes on random *left regenerative* compositions. J.P. 7/19/02 (1)

Let (Q_n) be a sequence of laws on the set $\mathcal{C}_n$ of compositions of n. Then we can consider

a First class deletion property (FCDP)

a Last class . . . . (LCDP)

a random . . . . (RCDP).

The last means removal of X_J from $(X_1, \ldots, X_K)$ for suitable random J. Various forms of different strengths depending on how much conditioning is allowed (X_J, or (J, X_J) etc., J could be uniform on $[1, K]$ or size-biased by the X_j.

~~Let X~~ Note that these concepts make sense without assuming the (Q_n) are consistent in any other way.

Let $X_1, X_2, \ldots$ be iid with values in $\{1, 2, \ldots\}$.

Let $S_k = X_1 + \cdots + X_k$.

~~Let~~

$$Q_n = \text{Law}\left(X_1, \ldots, X_{K_n} \,\middle|\, \overbrace{S_k = n \text{ for some } k = 1, 2 \ldots}^{\text{Renewal at } n}\right)$$

where K_n is the index k such that $S_k = n$.

Then, by elementary renewal theory, (Q_n) ~~is consistent~~ has the FCDP & LCDP

In fact Q_n is an exchangeable comp of n, each n.

These compositions of n are of Gibbs type.

Probability function is

$$g(n_1, \ldots, n_k) = \frac{\prod_{i=1}^{k} p_{n_i}}{P(\text{renewal at } n)}$$

Need to assume $p_n > 0$ all n, else may not be well defined. Call these RENEWAL COMPOSITIONS

Seems some mistake.

Should check it for α, α,

Compute the law of n_1 e.g. of the bridge, and see it sums to 1.

Maybe some ~~basic~~ algebra mistake here.

Compare Proposition 8 ~~7~~ of

- PD(α, θ) is a probability law for a sequence of random frequencies

$$\boldsymbol{P}^{\downarrow} = (P_1, P_2, \cdots), \quad \text{with} \quad P_1 > P_2 > \cdots > 0, \ \sum_j P_j = 1,$$

obtained by arranging in decreasing order another sequence

$$\widetilde{\boldsymbol{P}} = (\widetilde{P}_1, \widetilde{P}_2, \cdots)$$

The Pitman-Yor definition

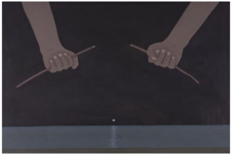
- $\mathrm{PD}(\alpha, \theta)$ is a probability law for a sequence of random frequencies

$$\boldsymbol{P}^{\downarrow} = (P_1, P_2, \cdots), \quad \text{with} \quad P_1 > P_2 > \cdots > 0, \ \sum_j P_j = 1,$$

obtained by arranging in decreasing order another sequence

$$\widetilde{\boldsymbol{P}} = (\widetilde{P}_1, \widetilde{P}_2, \cdots)$$

$$\widetilde{P}_j = W_j \prod_{i=1}^{j-1} (1 - W_j), \quad j = 1, 2, \ldots,$$



where W_i's independent, with $W_j \stackrel{\mathcal{L}}{=} \mathrm{Beta}(1 - \alpha, \theta + \alpha j)$

Two algorithms for size-biased ordering

- Conventional sampling without replacement algorithm:
  For $p_1, p_2, \ldots$ with $s = \sum p_j < \infty$, a size-biased pick $\tilde{p}_1 := p_J$ is defined by setting $\mathbb{P}(J = j) = p_j/s$. Removing J from $\mathbb{N}$, resp. p_J from $p_1, p_2, \cdots$, and iterating the SB-picking yields a SBP of $\mathbb{N}$, resp. of $p_1, p_2, \cdots$

- Ranking algorithm to arrange $p_1, p_2, \cdots$ in SB order: kth iteration only deals with $p_1, \cdots, p_k$. After $1, \cdots, k$ have been arranged as $i_1, \cdots, i_k$ with $(q_1, \cdots, q_k) := (p_{i_1}, \cdots, p_{i_k})$ the relative rank ρ_{k+1} of $k+1$ is determined by moving $k+1$ left-to-right through $i_1, \cdots, i_k$ until settling in position $\rho_{k+1} = m \in \{1, \cdots, k+1\}$ with odds

$$p_{k+1} : (q_m + \cdots + q_k).$$

The infinite SB order is defined by $\rho_1, \rho_2, \ldots$

- Ranking algorithm to arrange $p_1, p_2, \cdots$ in SB order: kth iteration only deals with $p_1, \cdots, p_k$. After $1, \cdots, k$ have been arranged as $i_1, \cdots, i_k$ with $(q_1, \cdots, q_k) := (p_{i_1}, \cdots, p_{i_k})$ the relative rank ρ_{k+1} of $k+1$ is determined by moving $k+1$ left-to-right through $i_1, \cdots, i_k$ until settling in position $\rho_{k+1} = m \in \{1, \cdots, k+1\}$ with odds

$$p_{k+1} : (q_m + \cdots + q_k).$$

  The infinite SB order is defined by $\rho_1, \rho_2, \ldots$
- k steps yield $1, \cdots, k$ (resp. $p_1, \cdots, p_k$) in size-biased order, showing that the finite orders are consistent under restrictions (cf also P-Tran '12).

- Ranking algorithm to arrange $p_1, p_2, \cdots$ in SB order: kth iteration only deals with $p_1, \cdots, p_k$. After $1, \cdots, k$ have been arranged as $i_1, \cdots, i_k$ with $(q_1, \cdots, q_k) := (p_{i_1}, \cdots, p_{i_k})$ the relative rank ρ_{k+1} of $k+1$ is determined by moving $k+1$ left-to-right through $i_1, \cdots, i_k$ until settling in position $\rho_{k+1} = m \in \{1, \cdots, k+1\}$ with odds

$$p_{k+1} : (q_m + \cdots + q_k).$$

  The infinite SB order is defined by $\rho_1, \rho_2, \ldots$
- k steps yield $1, \cdots, k$ (resp. $p_1, \cdots, p_k$) in size-biased order, showing that the finite orders are consistent under restrictions (cf also P-Tran '12).
- Works also if $\sum p_j = \infty$ although in this case the SB order is not a well-order.

- Ranking algorithm to arrange $p_1, p_2, \cdots$ in SB order: kth iteration only deals with $p_1, \cdots, p_k$. After $1, \cdots, k$ have been arranged as $i_1, \cdots, i_k$ with $(q_1, \cdots, q_k) := (p_{i_1}, \cdots, p_{i_k})$ the relative rank ρ_{k+1} of $k+1$ is determined by moving $k+1$ left-to-right through $i_1, \cdots, i_k$ until settling in position $\rho_{k+1} = m \in \{1, \cdots, k+1\}$ with odds

$$p_{k+1} : (q_m + \cdots + q_k).$$

  The infinite SB order is defined by $\rho_1, \rho_2, \ldots$
- k steps yield $1, \cdots, k$ (resp. $p_1, \cdots, p_k$) in size-biased order, showing that the finite orders are consistent under restrictions (cf also P-Tran '12).
- Works also if $\sum p_j = \infty$ although in this case the SB order is not a well-order.
- When $p_1 = p_2 = \cdots$ we have the ranks ρ_k independent, uniform on $[k] := \{1, \cdots, k\}$, and the resulting order is the *exchangeable* infinite order (Aldous '83), which restricts to $[k]$ as uniformly distributed permutation.

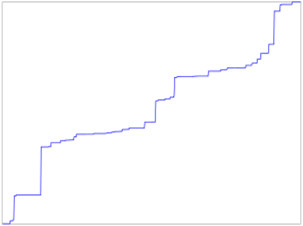
Characterisation of PD by SBP

- If $\widetilde{P}_1$ is independent of $(\widetilde{P}_2, \widetilde{P}_3, \cdots)/(1 - \widetilde{P}_1)$ then the stick-breaking factors Y_j are independent and (excluding some trivial cases) $\boldsymbol{P}^{\downarrow} \stackrel{\mathcal{L}}{=} \mathrm{PD}(\alpha, \theta)$ for some α, θ.

– McCloskey '65, P '96, G-Haulk-P '09

Ordered representations of PD involve

- either an increasing jump process (random c.d.f.) $(F_t, t \geq 0)$,

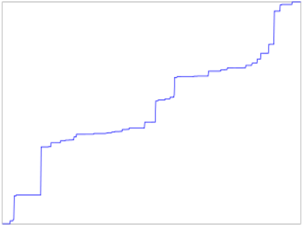


- or interval partition of $[0, 1]$ into components of $[0, 1] \setminus Z$, for Z a random measure-0 closed set.

Every such representation implies certain arrangement $\boldsymbol{P}^*$ of the frequencies $P_j^{\downarrow}$'s in accord with the natural ordering of jump-times, resp. component intervals.

Ordered representations of PD involve

- either an increasing jump process (random c.d.f.) $(F_t, t \geq 0)$,



- or interval partition of $[0,1]$ into components of $[0,1] \setminus Z$, for Z a random measure-0 closed set.

Every such representation implies certain arrangement $\boldsymbol{P}^*$ of the frequencies $P_j^\downarrow$'s in accord with the natural ordering of jump-times, resp. component intervals.

- The *arrangement problem* concerns features of this induced order $\boldsymbol{P}^*$, characterization of PD and sub-families, as well as connection of $\boldsymbol{P}^*$ to the well-orders $\boldsymbol{P}^\downarrow$ and $\widetilde{\boldsymbol{P}}$.

A combinatorial counterpart of the arrangement problem

- Recall that $\widetilde{P}_j$ is the asymptotic frequency of the jth occupied table in the Dubins-Pitman Chinese Restaurant

A combinatorial counterpart of the arrangement problem

- Recall that $\widetilde{P}_j$ is the asymptotic frequency of the jth occupied table in the Dubins-Pitman Chinese Restaurant

A combinatorial counterpart of the arrangement problem

- Recall that $\widetilde{P}_j$ is the asymptotic frequency of the jth occupied table in the Dubins-Pitman Chinese Restaurant

A combinatorial counterpart of the arrangement problem

- Recall that $\widetilde{P}_j$ is the asymptotic frequency of the jth occupied table in the Dubins-Pitman Chinese Restaurant

When the occupancy numbers are $n_1, \ldots, n_k,\ (n_1 + \cdots + n_k = n)$

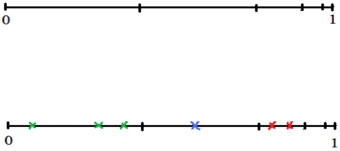
- sits at occupied table j with probability $\frac{n_j - \alpha}{n + \theta}$,
- occupies a new table with probability $\frac{\theta + k\alpha}{n + \theta}$.

- Hence a n-sample from $\boldsymbol{P}^*$ has the structure of *composition* (ordered partition) Π_n^* of integer n, with the CRP 'table' occupancy counts arranged in the corresponding order. The Π_n^*'s are consistent as n varies.

- Hence a n-sample from $\boldsymbol{P}^*$ has the structure of *composition* (ordered partition) Π_n^* of integer n, with the CRP 'table' occupancy counts arranged in the corresponding order. The Π_n^*'s are consistent as n varies.

- Hence a n-sample from $\boldsymbol{P}^*$ has the structure of *composition* (ordered partition) Π_n^* of integer n, with the CRP 'table' occupancy counts arranged in the corresponding order. The Π_n^*'s are consistent as n varies.

$Z, U_1, \cdots, U_n$



- sample uniform[0,1] points $U_1, \ldots, U_n$
- scan the gaps in Z in the left-to-right order
- record the sizes of clusters in each occupied gap

Subordinator ‘bridge’ representations of PD

- For $(S_t, t \geq 0)$ a subordinator with $S_0 = 0$ and tilted by manipulating the distribution of (T, S_T)

$$F_t = \frac{S_t}{S_T}, \quad 0 \leq t \leq T$$

- depending on choice of subordinator (gamma, stable, generalized gamma) some restricted range of $(\alpha, \theta) \in [0, 1) \times [0, \infty)$ may be covered
  –McCloskey ’65, Kingman ’75, Perman-PY ’92, PY ’97, P ’03

Subordinator 'bridge' representations of PD

- For $(S_t, t \geq 0)$ a subordinator with $S_0 = 0$ and tilted by manipulating the distribution of (T, S_T)

$$F_t = \frac{S_t}{S_T}, \quad 0 \leq t \leq T$$

- depending on choice of subordinator (gamma, stable, generalized gamma) some restricted range of $(\alpha, \theta) \in [0, 1) \times [0, \infty)$ may be covered
  –McCloskey '65, Kingman '75, Perman-PY '92, PY '97, P '03
- The induced order $\boldsymbol{P}^*$ is exchangeable, i.e. $P_j^{\downarrow}$ (equivalently, $\widetilde{P}_j$'s) are shuffled 'uniformly at random'.

The self-similar representation of PD($\alpha, 0$)

- Let Z be the α-stable set (e.g., the zero set of BM for $\alpha = 1/2$), so $Z \cap [0, 1]$ is the range of α-stable subordinator $(S_t, t \geq 0)$ 'cut' by passing level 1.
- The SB pick $\widetilde{P}_1$ is the size of the rightmost 'meander' interval, while all other frequencies occur in the exchangeable order (PY '96).
- *Exactly* the same (and not just in the $n \to \infty$ regime) arrangement of parts occurs on the combinatorial level of composition Π^* (P '97).

- For subordinator $(S_t, t \geq 0)$ (with $S_0 = 0$) the 'discrete' c.d.f.

$$F_t := 1 - e^{-S_t}$$

is known as a *neutral-to-the right prior*.

- Compositions Π_n^* have a (characteristic) first-part deletion property: given the first part is m, deleting the part from Π_n^* yields a copy of Π_{n-m}^*.

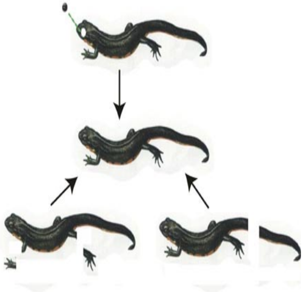
Regenerative compositions

- For subordinator $(S_t, t \geq 0)$ (with $S_0 = 0$) the ‘discrete’ c.d.f.

$$F_t := 1 - e^{-S_t}$$

is known as a *neutral-to-the right prior*.

- Compositions Π_n^* have a (characteristic) first-part deletion property: given the first part is m, deleting the part from Π_n^* yields a copy of Π_{n-m}^*.



- The distribution of Π_n^* has a product form

$$p^*(n_1, \dots, n_k) = \prod_{j=1}^{k} q(n_j + \cdots + n_k\,,\ n_j)$$

where (assuming zero drift)

$$\begin{aligned} q(n, m) : &= \frac{\Phi(n, m)}{\Phi(n)}, \\ \Phi(\lambda) &:= \int_0^\infty (1 - e^{-\lambda x})\nu(dx), \\ \Phi(n, m) &:= \binom{n}{m} \int_0^\infty (1 - e^{-x})^m e^{-(n-m)x} \nu(dx) \end{aligned}$$

and ν is the Lévy measure of the subordinator.

- The NTR/regenerative representation is an intrinsic property of unordered objects $\boldsymbol{P}^{\downarrow}$/partition structure (Π_n), by the virtue of

$$p^*(n) = \mathbb{E}[\widetilde{P}_1^{n-1}] = \mathbb{E}\left[\sum_j (P_j^{\downarrow})^n\right].$$

- Moreover, if (unordered) partitions Π_n (derived by sampling from some frequencies $\boldsymbol{P}^{\downarrow}$) have some kind of part-deletion property, then they can be represented by regenerative compositions Π_n^*.

- The NTR/regenerative representation is an intrinsic property of unordered objects $\boldsymbol{P}^{\downarrow}$/partition structure (Π_n), by the virtue of

$$p^*(n) = \mathbb{E}[\widetilde{P}_1^{n-1}] = \mathbb{E}\left[\sum_j (P_j^{\downarrow})^n\right].$$

- Moreover, if (unordered) partitions Π_n (derived by sampling from some frequencies $\boldsymbol{P}^{\downarrow}$) have some kind of part-deletion property, then they can be represented by regenerative compositions Π_n^*.

The regenerative representation of PD

The Redwoods Park Theorem: *for* $0 \leq \alpha < 1,\ \theta \geq 0$, *the PD*$(\alpha, \theta)$ *is regenerative:*

The Redwoods Park Theorem: *for* $0 \leq \alpha < 1,\ \theta \geq 0$, *the PD*$(\alpha, \theta)$ *is regenerative:*

- there is a Lévy measure with the upper tail

$$\nu_{\alpha,\theta}[x, \infty) = (1 - e^{-x})^{-\alpha} e^{-x\theta}, \qquad x \geq 0$$

The regenerative representation of PD

The Redwoods Park Theorem: *for* $0 \le \alpha < 1,\ \theta \ge 0$*, the PD*$(\alpha, \theta)$ *is regenerative:*

- there is a Lévy measure with the upper tail

$$\nu_{\alpha,\theta}[x, \infty) = (1 - e^{-x})^{-\alpha} e^{-x\theta}, \qquad x \ge 0$$

- which corresponds to

$$\Phi(\lambda) = \frac{\lambda\,\Gamma(1-\alpha)\Gamma(\lambda+\theta)}{\Gamma(\lambda+1-\alpha+\theta)}$$

$$q(n, m) = \binom{n}{m} \frac{(1-\alpha)_{m-1}}{(\theta+n-m)_{n-1}} \frac{(n-m)\alpha + m\theta}{n}$$

and agrees with the Ewens-Pitman sampling formula, e.g.

$$p^*(n) = \frac{(1-\alpha)_{n-1}}{(1+\theta)_{n-1}}.$$

Subfamily PD$(0, \theta)$

- For
$$\nu_{0,\theta}(dx) = \theta e^{-x\theta} dx, \quad x \geq 0$$
$(S_t, t \geq 0)$ is compound Poisson with exponential jumps
- The range of $F_t = 1 - e^{-S_t}$ is a Poisson point process with rate function $\theta(1-y)^{-1}, y \in [0,1)$, hence points obtainable by i.i.d. beta$(1, \theta)$ stick-breaking
- $\boldsymbol{P}^*$ is a size-biased permutation of $\boldsymbol{P}^{\downarrow}$, likewise Π_n^* has parts in the SB order
- If a partition structure has the SB-part deletion property, it is Ewens' – Kingman '78

Subfamily PD(α, α)

- The (α, α) Lévy measure

$$\nu_{\alpha,\alpha}(dx) = \frac{\alpha e^x}{(e^x - 1)^{\alpha+1}} dx, \qquad x \geq 0$$

  was introduced by Lamperti in '73. He observed that e^{S_t} is a time-changed α-stable subordinator starting at 1, with the range $1 + Z_\alpha$ (for Z_α α-stable regenerative set). Hence the range of $F_t = 1 - e^{-S_t}$ is $Z_\alpha/(Z_\alpha + 1)$, which is a known representation of the α-stable bridge
- $\boldsymbol{P}^*$ is an exchangeable ordering of $\boldsymbol{P}^\downarrow$
- If the first part and the last part of regenerative Π_n^* have the same distribution (for every n), then the partition structure is Pitman's (α, α).

- The $(\alpha, 0)$ Lévy measure is

$$\nu_{\alpha,0}(dx) = \frac{\alpha e^{\alpha x}}{(e^x - 1)^{\alpha+1}} dx + \delta_\infty(dx), \qquad x \geq 0$$

- Killed subordinator (S_t) belongs to the family of Lamperti-stable subordinators (Chaumont and Caballero '06)
- $\boldsymbol{P}^*$ has the last part $\widetilde{P}_1$, and the other parts are in exchangeable order. Same for Π_n^*. This kind of ordering is characteristic for $(\alpha, 0)$ among the regenerative structures.

Alternative proof of the regenerative representation for $\alpha, \theta > 0$

- To construct a path of subordinator with Lévy measure $\nu[x, \infty) = e^{-x\theta}(1 - e^{-x})^{-\alpha}$
  (a) split $[0, \infty)$ at points $E_1 + \cdots + E_j$ of homogeneous Poisson(θ) process,
  (b) run a Lamperti-stable $(\alpha, 0)$ (killed) subordinator from the origin $(0, 0)$ until crossing E_1 at some time τ_1,
  (c) start another copy of this $(\alpha, 0)$-subordinator from (E_1, τ_1) and run it until passing $E_1 + E_2$, etc.
- The construction agrees with PY '97, Corollary to Proposition 21: For $0 \le \alpha < 1,\ \theta \ge 0$, $\boldsymbol{P}^{\downarrow}$ can be obtained by first splitting the unity according to PD$(0, \theta)$, then further fragmenting each piece by an independent copy of the α-stable set.

Size-cosize biased permutations

- Fix $\xi \in [0, \infty]$. For positive $p_1, \ldots, p_k$ with $s = p_1 + \cdots + p_k$, a *ξ-biased pick* is a random element p_J, where

$$\mathbb{P}(J = j) = \frac{\xi p_j + (s - p_j)}{s(\xi + k - 1)}.$$

- ∞-biased pick is size-biased,
- 1-biased pick is uniformly random,
- 0-biased pick is cosize biased.

Removing J, p_J and iterating yields a *ξ-biased permutation* of $p_1, \ldots, p_k$, denoted

$$\mathrm{perm}_\xi(p_1, \ldots, p_n).$$

This is only defined for $k < \infty$!

Let $\Pi_n^{\downarrow}$ be the (α, θ)-partition of n, and let $\xi = \theta/\alpha$.

- partition structure $(\Pi_n^{\downarrow})$ has the ξ-biased pick deletion property, which is characteristic;
- Π_n^* is obtained from $(\Pi_n^{\downarrow})$ by arranging the parts in the ξ-biased order.

ξ-biased permutations are not consistent under restrictions ... How to extend to $p_1, p_2, \ldots$ with infinitely many $p_j > 0$?

Permutations/orders with biased record counts

- For $\xi \in [0, \infty]$, define $\lhd_\xi$ as a random order on $\mathbb{N}$ with independent ranks $\rho_1, \rho_2, \dots$ such that

$$\mathbb{P}(\rho_k = i) = \begin{cases} \frac{1}{k+\xi-1} & \text{for} \;\; 1 \le i \le k-1, \\ \frac{\xi}{k+\xi-1} & \text{for} \;\; i = k. \end{cases}$$

- $\lhd_\xi$ restricts to $[k]$ as a permutation with distribution

$$\frac{\xi^{\#\text{records}}}{(\xi)_k}$$

A solution to the arrangement problem

- Key observation:

$$\mathrm{perm}_\xi(p_1,\ldots,p_k) \stackrel{\mathcal{L}}{=} \lhd_\xi \circ \mathrm{perm}_\infty(p_1,\ldots,p_k)$$

  where the record-biased order $\lhd_\xi$ and SBP perm_∞ are independent. The RHS makes sense for infinite $p_1, p_2, \ldots$
  – P-Winkel '09, G-Haulk-P '09
- Conclusion: For $0 \le \alpha < 1, \theta \ge 0$, the regenerative version $\boldsymbol{P}^*$ of $\mathrm{PD}(\alpha,\theta)$ can be obtained by arranging the SB frequencies $\widetilde{\boldsymbol{P}}$ in the independent order $\lhd_{\theta/\alpha}$.

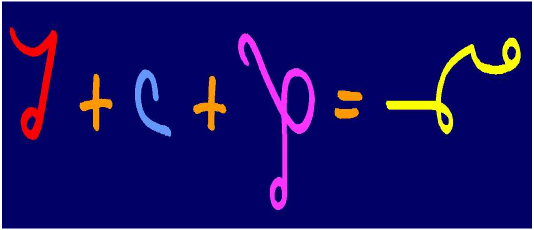
Selfsimiar, stationary regenerative representations of PD

- $Z \subset [0, \infty)$ is *selfsimilar* if $Z \stackrel{\mathcal{L}}{=} cZ$ for $c > 0$
- In particular, $Z \cap [0, 1]$ could be the range of $(e^{-(S_t+X)}, t \geq 0)$ for mean-μ subordinator (S_t) and independent X with distribution $\mathbb{P}(X \in dx) = \mu^{-1}\nu[x, \infty]dx$.
- Then the induced composition structure (Π_n^*) has a *last-bit* deletion property in the representation via binary code (e.g. 1010001 for (2,4,1)), and a version of the *last-part* deletion property.
  –Young '05, GP '05

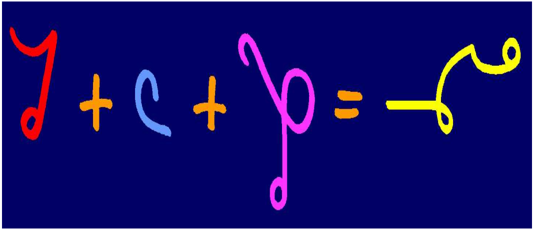
- For $0 \le \alpha < 1, \quad \theta > -\alpha$, PD$(\alpha, \theta)$ has such representation where

$$1 - e^{-X} = \widetilde{P}_1 \stackrel{\mathcal{L}}{=} \mathrm{beta}(1-\alpha, \theta+\alpha)$$

  and (S_t) is subordinator with
  $\nu[x, \infty] = e^{-x(\theta+\alpha)}(1-e^{-x})^{-\alpha}$.
  –PY '96, GP '05
- Then the SB pick $\widetilde{P}_1$ is in the last position, while other frequencies are frequencies are arranged in the order inverse to the regenerative-PD$(\alpha, \theta+\alpha)$.
- Example: the BM $(1/2, 0)$ interval partition is left-to-right regenerative (NTR), and at the same time right-to-left stationary-regenerative (hence self-similar).
- Example: The BB $(1/2, 1/2)$ interval partition is NTR, but also has a realization as a stationary version of regenerative-PD$(1/2, 1)$



$$? + ? + ? = ?$$



$? + ? + ? = ?$

Intelligence + effort + persistence = excellence