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

Sasha Gnedin

Queen Mary, University of London
Notes on random compositions. J.P. 7/19/02

Let \((Q_n)\) be a law on the set \(B_n\) of compositions of \(n\). Then we can consider:

- First class deletion property \((\text{FCDP})\)
- Last class \(\cdots\) \((\text{LCDP})\)
- Random \(\cdots\) \((\text{RCDP})\).

The last means removal of \(X_j\) from \((X_1, \ldots, X_k)\) for suitable random \(J\). Various forms of different strength depending on how much conditioning is allowed

\((X_j)\) or \((X_1X_2)\) etc.; \(J\) could be uniform on \([1,k]\) or size-based by the \(X_j\).

Note that these concepts work sense without assuming the \((Q_n)\) are consistent in any other way.

Let \(X_1, X_2, \ldots\) be i.i.d. with values in \(\mathbb{N}\),

Let \(S_k = X_1 + \ldots + X_k\). \(\text{Renewal at } n\)

\[ Q_n = \text{Law}(X_1, \ldots, X_{Kn} \mid S_k = n \text{ for some } k = 1, \ldots) \]

where \(Kn\) is the index \(k\) such that \(S_k = n\).

Then, by elementary renewal theory,

\((Q_n)\) is a banded law with the \(\text{FCDP} \& \text{LCDP}\)

In fact \(Q_n\) is an exchangeable cop of \(n\), and \(n\).

These compositions of \(n\) are of G-Fss type.

Probability function is \(k\):

\[ g(n_1, \ldots, n_k) = \prod_{i=1}^{k} P_{n_i} \]

\[ P(\text{renewal at } n) \]

Need to assume \(P_n > 0\) all \(n\), else may not be well defined. Call these \text{RENEWAL COMPS}.
Notes on random compositions. J.P. 7/19/02(1)

Let $(Q_n)$ be a law on the set $\mathcal{C}_n$ of compositions of $n$. Then we can consider:
- First class deletion property (FCDP)
- Last class (LCDP)
- Random (RCDP).

The last means removal of $X_1$ from $(X_1, \ldots, X_k)$ for suitable random $J$. Variance form of different structures depending on how many conditions is allowed.

$(X_1, \ldots, X_k)$ or $(X_1 X_k)$ etc., $J$ could be uniform in $\{1, k\}$ or size-based by the $X_i$.

Note: There are concepts under which assumptions the $(Q_n)$ are consistent in any other way.

Let $X_1, X_2, \ldots$ be i.i.d. with values in $\mathbb{N}$. Let $S_k = X_1 + \cdots + X_k$.

$$Q_n = \text{Law}\left( X_1, \ldots, X_k \mid S_k = n \text{ for some } k\right)$$

where $\{X_k\}$ is the index for which $S_k = n$.

Then, by elementary renewal theory,

$(Q_n)$ is exponentially the FCDP & LCDP.

In fact, $Q_n$ is an exchangeable
copy of $n$, and $n$.

These compositions of $n$ are of $G-$type.

Probability function is $g(n_1, \ldots, n_k) = \frac{k!}{i!} P_n$.

Need to assume $P_n > 0$ a.e., else may not be well-defined. Call these RENEWAL COMPOSE.

Sasha Gnedin  Random permutations and the 2-parameter PD
Let \( (Q_n) \) be a family on the set \( \mathcal{B}_n \) of compositions of \( n \). Then we can consider:

- First class deletion property \((\text{FCDP})\)
- Last class \((\text{LCDP})\)
- Random \((\text{RCDP})\).

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 \((X_1, X_2, \ldots)\) etc., \( J \) could be uniform on \([1,k]\)
or size-biased by the \( X_i \).

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

Let \( X_1, X_2, \ldots \) be i.i.d. with values in \( \mathbb{Z} \cap \{1,2,\ldots\} \).
Let \( S_k = X_1 + \cdots + X_k \).

\[
Q_n = \text{Law}(X_1, \ldots, X_k \mid S_k = n \text{ for some } k \in \mathbb{N})
\]

where \( k \) is the index \( k \) such that \( S_k = n \).

Then, by elementary renewal theory,
\( (Q_n) \) is \( \text{FCDP} \) and \( \text{LCDP} \).

In fact \( Q_n \) is an exchangeable cop of \( n \), and \( n \).

These compositions of \( n \) are of \( \mathcal{B}_n \) type.

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

Need to assume \( P(n) > 0 \) for all \( n \), else may not be well-defined. Call these \( \text{RENEWAL COMPOSITIONS} \).
• PD(\(\alpha, \theta\)) is a probability law for a sequence of random frequencies

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

obtained by arranging in decreasing order another sequence

\[ \tilde{P} = (\tilde{P}_1, \tilde{P}_2, \cdots) \]
The Pitman-Yor definition

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

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

obtained by arranging in decreasing order another sequence

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

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

where $W_i$'s independent, with $W_j \overset{\mathcal{L}}{=} \text{Beta}(1 - \alpha, \theta + \alpha j)$
• 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: 
  \( k \)th 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 \( \rho_{k+1} \) 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 \( \rho_1, \rho_2, \cdots \).
• Ranking algorithm to arrange $p_1, p_2, \cdots$ in SB order:
  $k$th 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 $\rho_{k+1}$ 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

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

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

• $k$ 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 \( p_1, p_2, \ldots \) in SB order:

\( k \)th iteration only deals with \( p_1, \ldots, p_k \). After \( 1, \ldots, k \) have been arranged as \( i_1, \ldots, i_k \) with \((q_1, \ldots, q_k) := (p_{i_1}, \ldots, p_{i_k})\) the relative rank \( \rho_{k+1} \) of \( k + 1 \) is determined by moving \( k + 1 \) left-to-right through \( i_1, \ldots, i_k \) until settling in position \( \rho_{k+1} = m \in \{1, \ldots, k + 1\} \) 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 \( 1, \ldots, k \) (resp. \( p_1, \ldots, p_k \)) 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 $\rho_{k+1}$ 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 $\rho_k$ 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.
• If $\tilde{P}_1$ is independent of $(\tilde{P}_2, \tilde{P}_3, \cdots)/(1 - \tilde{P}_1)$ then the stick-breaking factors $Y_j$ are independent and (excluding some trivial cases)

$$P\downarrow \overset{\mathcal{L}}{=} \text{PD}(\alpha, \theta) \text{ for some } \alpha, \theta.$$

– 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 \(P^*\) of the frequencies \(P_{j}\)'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 \(P^*\) of the frequencies \(P_j\)'s in accord with the natural ordering of jump-times, resp. component intervals.

- The arrangement problem concerns features of this induced order \(P^*\), characterization of PD and sub-families, as well as connection of \(P^*\) to the well-orders \(P_\downarrow\) and \(\tilde{P}\).
A combinatorial counterpart of the arrangement problem

- Recall that $\tilde{P}_j$ is the asymptotic frequency of the $j$th occupied table in the Dubins-Pitman Chinese Restaurant.
A combinatorial counterpart of the arrangement problem

- Recall that $\tilde{P}_j$ is the asymptotic frequency of the $j$th occupied table in the Dubins-Pitman Chinese Restaurant.
A combinatorial counterpart of the arrangement problem

- Recall that $\tilde{P}_j$ is the asymptotic frequency of the $j$th occupied table in the Dubins-Pitman Chinese Restaurant.
A combinatorial counterpart of the arrangement problem

- Recall that $\tilde{P}_j$ is the asymptotic frequency of the $j$th 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 $P^*$ has the structure of *composition* (ordered partition) $\Pi^*_n$ of integer $n$, with the CRP ‘table’ occupancy counts arranged in the corresponding order. The $\Pi^*_n$’s are consistent as $n$ varies.
Hence a $n$-sample from $P^*$ has the structure of *composition* (ordered partition) $\Pi^*_n$ of integer $n$, with the CRP ‘table’ occupancy counts arranged in the corresponding order. The $\Pi^*_n$’s are consistent as $n$ varies.
Hence a \( n \)-sample from \( P^* \) has the structure of \textit{composition} (ordered partition) \( \Pi^*_n \) of integer \( n \), with the CRP ‘table’ occupancy counts arranged in the corresponding order. The \( \Pi^*_n \)'s are consistent as \( n \) varies.

\[
Z, U_1, \ldots, U_n
\]

- 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
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 \(P^*\) is exchangeable, i.e. \(P_j\) (equivalently, \(\tilde{P}_j\)’s) are shuffled ‘uniformly at random’.
Let $Z$ be the $\alpha$-stable set (e.g., the zero set of BM for $\alpha = 1/2$), so $Z \cap [0, 1]$ is the range of $\alpha$-stable subordinator $(S_t, t \geq 0)$ ‘cut’ by passing level 1.

The SB pick $\tilde{P}_1$ 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 $n \rightarrow \infty$ regime) arrangement of parts occurs on the combinatorial level of composition $\Pi^*$ (P ’97).
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 \(\Pi^*_n\) have a (characteristic) first-part deletion property: given the first part is \(m\), deleting the part from \(\Pi^*_n\) yields a copy of \(\Pi^*_{n-m}\).
• 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 \textit{neutral-to-the right prior}.

• Compositions \(\Pi_n^*\) have a (characteristic) first-part deletion property: given the first part is \(m\), deleting the part from \(\Pi_n^*\) yields a copy of \(\Pi_{n-m}^*\).
The distribution of $\Pi^*_n$ has a product form

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

where (assuming zero drift)

$$q(n, m) := \frac{\Phi(n, m)}{\Phi(n)}$$

$$\Phi(\lambda) := \int_0^\infty (1 - e^{-\lambda x})\nu(dx),$$

$$\Phi(n, m) := \binom{n}{m} \int_0^\infty (1 - e^{-x})^m e^{-(n-m)x}\nu(dx)$$

and $\nu$ is the Lévy measure of the subordinator.
• The NTR/regenerative representation is an intrinsic property of unordered objects $P^\downarrow$/partition structure $(\Pi_n)$, by the virtue of

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

• Moreover, if (unordered) partitions $\Pi_n$ (derived by sampling from some frequencies $P^\downarrow$) have some kind of part-deletion property, then they can be represented by regenerative compositions $\Pi_n^*$. 
• The NTR/regenerative representation is an intrinsic property of unordered objects $P^\downarrow$ / partition structure ($\Pi_n$), by the virtue of

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

• Moreover, if (unordered) partitions $\Pi_n$ (derived by sampling from some frequencies $P^\downarrow$) have some kind of part-deletion property, then they can be represented by regenerative compositions $\Pi_n^*$. 
The regenerative representation of PD

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 $\nu_{\alpha, \theta}(x, \infty) = (1 - e^{-x})^{-\alpha} e^{-x\theta}$, $x \geq 0$
- which corresponds to

$$\Phi(\lambda) = \frac{\lambda^{1-\alpha} \Gamma(\lambda + \theta)}{\Gamma(\lambda + 1 - \alpha + \theta)} q(n, m) = \binom{n}{m} (1 - \alpha)^{m-1} (\theta + n - m)^{n-1} (\alpha + m)^{m\theta + n - m}$$

and agrees with the Ewens-Pitman sampling formula, e.g. $p^*(n) = (1 - \alpha)^{n-1} (1 + \theta)^{n-1}$. 

Sasha Gnedin

Random permutations and the 2-parameter PD
The regenerative representation of PD

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

- there is a Lévy measure with the upper tail

$$\nu_{\alpha, \theta}[x, \infty) = (1 - e^{-x})^{-\alpha} e^{-x\theta}, \quad x \geq 0$$
The regenerative representation of PD

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

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

- 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.

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

\( P^* \) is a size-biased permutation of \( P^\downarrow \), likewise \( \Pi_n^* \) has parts in the SB order

If a partition structure has the SB-part deletion property, it is Ewens’ – Kingman ’78
The $(\alpha, \alpha)$ Lévy measure

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

was introduced by Lamperti in '73. He observed that $e^{S_t}$ is a time-changed $\alpha$-stable subordinator starting at 1, with the range $1 + Z_\alpha$ (for $Z_\alpha$ $\alpha$-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 $\alpha$-stable bridge

- $P^*$ is an exchangeable ordering of $P_{\downarrow}$
- If the first part and the last part of regenerative $\Pi_n^*$ have the same distribution (for every $n$), then the partition structure is Pitman’s $(\alpha, \alpha)$. 
Subfamily PD($\alpha, 0$)

- The $(\alpha, 0)$ Lévy measure is

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

- Killed subordinator $(S_t)$ belongs to the family of Lamperti-stable subordinators (Chaumont and Caballero '06)

- $P^*$ has the last part $\tilde{P}_1$, and the other parts are in exchangeable order. Same for $\Pi^*_n$. 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 $
u[x, \infty) = e^{-x\theta}(1 - e^{-x})^{-\alpha}$
  
  (a) split $[0, \infty)$ at points $E_1 + \cdots + E_j$ of homogeneous Poisson($\theta$) process,
  (b) run a Lamperti-stable $(\alpha, 0)$ (killed) subordinator from the origin $(0, 0)$ until crossing $E_1$ at some time $\tau_1$,
  (c) start another copy of this $(\alpha, 0)$-subordinator from $(E_1, \tau_1)$ and run it until passing $E_1 + E_2$, etc.

- The construction agrees with PY ’97, Corollary to Proposition 21: For $0 \leq \alpha < 1, \ \theta \geq 0$, $P^\perp$ can be obtained by first splitting the unity according to PD$(0, \theta)$, then further fragmenting each piece by an independent copy of the $\alpha$-stable set.
Fix $\xi \in [0, \infty]$. For positive $p_1, \ldots, p_k$ with $s = p_1 + \cdots + p_k$, a $\xi$-biased pick is a random element $p_J$, where

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

• $\infty$-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 $\xi$-biased permutation of $p_1, \ldots, p_k$, denoted

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

This is only defined for $k < \infty$!
The arrangement problem for fixed $n$

Let $\Pi_n$ be the $(\alpha, \theta)$-partition of $n$, and let $\xi = \theta/\alpha$.

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

$\xi$-biased permutations are not consistent under restrictions . . . How to extend to $p_1, p_2, \ldots$ with infinitely many $p_j > 0$?
• For $\xi \in [0, \infty]$, define $\triangleleft_\xi$ as a random order on $\mathbb{N}$ with independent ranks $\rho_1, \rho_2, \ldots$ such that

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

• $\triangleleft_\xi$ restricts to $[k]$ as a permutation with distribution

$$
\frac{\xi^{\#\text{records}}}{(\xi)_k}
$$
A solution to the arrangement problem

- Key observation:

\[ \text{perm}_\xi(p_1, \ldots, p_k) \overset{\sim}{=} \triangleleft_\xi \circ \text{perm}_\infty(p_1, \ldots, p_k) \]

where the record-biased order \( \triangleleft_\xi \) and SBP \( \text{perm}_\infty \) are independent. The RHS makes sense for infinite \( p_1, p_2, \ldots \)

\[ \text{– P-Winkel ’09, G-Haulk-P ’09} \]

- Conclusion: For \( 0 \leq \alpha < 1, \theta \geq 0 \), the regenerative version \( \tilde{P}^* \) of \( \text{PD}(\alpha, \theta) \) can be obtained by arranging the SB frequencies \( \tilde{P} \) in the independent order \( \triangleleft_{\theta/\alpha} \).
• $Z \subset [0, \infty)$ is *selfsimilar* if $Z \overset{\mathcal{L}}{=} cZ$ for $c > 0$

• In particular, $Z \cap [0, 1]$ could be the range of $(e^{-(S_t+X)}, t \geq 0)$ for mean-$\mu$ subordinator $(S_t)$ and independent $X$ with distribution
\[ P(X \in dx) = \mu^{-1} \nu[x, \infty] dx. \]

• Then the induced composition structure $(\Pi_n^*)$ 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 $0 \leq \alpha < 1$, $\theta > -\alpha$, $\text{PD}(\alpha, \theta)$ has such representation where

$$1 - e^{-x} = \tilde{P}_1 \overset{\mathcal{L}}{=} \text{beta}(1 - \alpha, \theta + \alpha)$$

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 $\tilde{P}_1$ is in the last position, while other frequencies are arranged in the order inverse to the regenerative-$\text{PD}(\alpha, \theta + \alpha)$.

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-$\text{PD}(1/2, 1)$.
Intelligence + effort + persistence = excellence

\[ ? + ? + ? = ? \]
Intelligence + effort + persistence = excellence