Feller coupling of cycles of permutations and Poisson spacings in inhomogeneous Bernoulli trials

Joseph Najnudel*   Jim Pitman†

Abstract

Feller (1945) provided a coupling between the counts of cycles of various sizes in a uniform random permutation of $[n]$ and the spacings between successes in a sequence of n independent Bernoulli trials with success probability $1/n$ at the nth trial. Arratia, Barbour and Tavaré (1992) extended Feller's coupling, to associate cycles of random permutations governed by the Ewens (θ) distribution with spacings derived from independent Bernoulli trials with success probability $\theta/(n-1+\theta)$ at the nth trial, and to conclude that in an infinite sequence of such trials, the numbers of spacings of length ℓ are independent Poisson variables with means θ/ℓ. Ignatov (1978) first discovered this remarkable result in the uniform case $\theta = 1$, by constructing Bernoulli $(1/n)$ trials as the indicators of record values in a sequence of i.i.d. uniform $[0,1]$ variables. In the present article, the Poisson property of inhomogeneous Bernoulli spacings is explained by a variation of Ignatov's approach for a general $\theta > 0$. Moreover, our approach naturally provides random permutations of infinite sets whose cycle counts are exactly given by independent Poisson random variables.

Keywords: cycles; records; random permutations; Poisson process.
MSC2020 subject classifications: 60C05.
Submitted to ECP on July 23, 2019, final version accepted on September 22, 2020.

1 Introduction

In [6], Feller introduces a coupling between the cycle structure of a uniformly distributed random permutation of order n and the spacings between successes in a sequence of n independent Bernoulli variables of parameters $1/n, 1/(n-1), \ldots, 1/2, 1$. This coupling has been generalized to Ewens distributions for any parameter θ: a recent discussion on this topic, with references to further work is provided by Arratia, Barbour, and Tavaré [3], largely following their earlier work [1]. Their coupling, for a general positive integer n and $\theta > 0$, may be constructed as follows. Consider a sequence $(B_i(\theta))_{1 \le i \le n}$ of independent Bernoulli variables, $B_i(\theta)$ with parameters $\theta/(\theta + i - 1)$. Conditionally on $(B_i(\theta))_{1 \le i \le n}$, construct the random permutation σ of the set $[n] := \{1, 2, \ldots, n\}$, as follows. First, define $X_1 := 1$, and then, recursively for $2 \le i \le n$:

- If $B_{n+2-i}(\theta) = 1$, X_i is the smallest element of $[n]$, different from $X_1, \ldots, X_{i-1}$.
- Conditionally on the fact that $B_{n+2-i}(\theta) = 0$, and on the values of $X_1, \ldots, X_{i-1}$, the element X_i is uniformly distributed on $[n] \backslash \{X_1, \ldots, X_{i-1}\}$.

*School of Mathematics, University of Bristol, UK. E-mail: joseph.najnudel@bristol.ac.uk
†Statistics department, University of California, Berkeley, USA. E-mail: pitman@stat.berkeley.edu

Then, the cycle structure of the permutation σ is obtained by taking the subsequences of $(X_1, X_2, \ldots, X_n)$, in such a way that the value X_i is the start of a cycle if and only if $i = 1$ or $B_{n+2-i}(\theta) = 1$.

For example, suppose $n = 9$,

$$(B_1(\theta), B_2(\theta), \ldots, B_9(\theta)) = (1, 0, 1, 0, 0, 1, 1, 0, 0).$$

A possible realization of the X_i's is

$$(X_1, X_2, \ldots, X_9) = (1, 7, 3, 2, 4, 9, 5, 6, 8).$$

Since $B_3(\theta)$, $B_6(\theta)$ and $B_7(\theta)$ are equal to 1, we have cycles starting at X_8, X_5 and X_4, and then

$$\sigma = (173)(2)(495)(68).$$

Note that written in this fashion, each cycle starts with its minimal element, and the cycles are written in increasing order of their minimal elements.

To indicate the parameters n and θ used in this construction, let $\pi_{n,\theta}$ denote the random permutation σ of $[n]$ so constructed. Then $\pi_{n,\theta}$ follows the Ewens distribution

$$\mathbb{P}(\pi_{n,\theta} = \pi) = \frac{\theta^{K(\pi)}}{(\theta)_n} \text{ where } (\theta)_n := \theta(\theta + 1) \cdots (\theta + n - 1)$$

and $K(\pi)$ is the number of cycles of a permutation π of $[n]$. The proof of this fact is indicated in [3] and appeals to Feller's original coupling of $B_1(1), \ldots, B_n(1)$ to a uniform random permutation $\pi_{n,1}$ for $\theta = 1$, and a simple change of measure argument for $\theta \neq 1$.

The cycle structure of $\pi_{n,\theta}$ can be deduced from the spacings between the Bernoulli variables $B_i(\theta)$ which are equal to 1. More precisely, for $\ell \geq 1$, let us say that an ℓ-spacing occurs in a sequence $a_1, a_2, \ldots$ of 0s and 1s, starting at position $i - \ell$ and ending at position i, if

$$a_{i-\ell} \cdots a_i = 1\, 0^{\ell-1}\, 1$$

meaning that the string of length $\ell + 1$ is a 1 followed by $\ell - 1$ zeros followed by 1. If $C_{n,\ell}(\theta)$ is the number of ℓ-spacings in

$$B_1(\theta), \ldots, B_n(\theta), 1, 0, 0, 0, \ldots$$

then there is the equality

$$(C_{n,\ell}(\theta), 1 \leq \ell \leq n) = (K_\ell(\pi_{n,\theta}), 1 \leq \ell \leq n) \tag{1.1}$$

where $K_\ell(\pi_{n,\theta})$ is the number of cycles of length ℓ in the permutation $\pi_{n,\theta}$.

By regarding the sequence $(B_i(\theta))_{1 \leq i \leq n}$ as the first n terms of an infinite sequence $(B_i(\theta))_{i \geq 1}$ of independent Bernoulli variables, we get a coupling, on a single probability space, of the families of cycle lengths $(K_\ell(\pi_{n,\theta}))_{1 \leq \ell \leq n}$ for all values of n. We quickly deduce the following result by Arratia, Barbour and Tavaré, for which we provide a sketch of proof here for the reader's convenience in comparing with later arguments:

Theorem 1.1 (Arratia, Barbour and Tavaré ([1], [3])). *If $C_{\infty,\ell}(\theta)$ is the number of ℓ-spacings in the infinite sequence $(B_i(\theta))_{i \geq 1}$ of Bernoulli variables, $B_i(\theta)$ having parameter $\theta/(\theta + i - 1)$, then $(C_{\infty,\ell}(\theta))_{\ell \geq 1}$ is a sequence of independent Poisson(θ/ℓ) variables.*

Proof. If $L_n(\theta)$ is the position of the last 1 in $(B_i(\theta))_{1 \leq i \leq n}$ and $J_n(\theta) := n + 1 - L_n(\theta)$ is the last spacing in the finite n scheme, then

$$C_{n,\ell}(\theta) \leq C_{\infty,\ell}(\theta) + 1(J_n(\theta) = \ell) \qquad (1 \leq \ell \leq n) \tag{1.2}$$

with strict inequality iff there is an ℓ-spacing in the infinite sequence $(B_j(\theta))_{j \geq 1}$ starting at $j = i - \ell$ and ending at $j = i > n$. Now, this event and the event $\{J_n(\theta) = \ell\}$ have probability tending to zero when $n \to \infty$, so for fixed ℓ, $C_{n,\ell}(\theta) = C_{\infty,\ell}(\theta)$ with probability tending to one. On the other hand, by (1.1), for any fixed $K \geq 1$, $(C_{n,\ell}(\theta), 1 \leq \ell \leq K)$ tends in law to independent Poisson(θ/ℓ) variables. The two last facts together imply the theorem. □

Combining (1.1) and (1.2), we get a coupling of counts of small cycles $(K_\ell(\pi_{n,\theta}))_{1 \leq \ell \leq k}$ of a Ewens(θ) permutation to independent Poisson (θ/ℓ) counts $(C_{\infty,\ell}(\theta))_{1 \leq \ell \leq k}$, with a total variation error depending on k and θ which is easily bounded explicitly. This implies in particular that

$$(K_\ell(\pi_{n,\theta}), 1 \leq \ell \leq k) \overset{(d)}{\to} (C_{\infty,\ell}(\theta), 1 \leq \ell \leq k) \text{ as } n \to \infty \qquad (1.3)$$

for every fixed k, as well as estimates of total variation error in this approximation which are useful for $k = o(n)$: see [1, Theorems 1 and 3]. See also Sethuraman and Sethuraman [19] for a review of studies of the distribution of the numbers of ℓ-spacings in infinite sequences of independent Bernoulli trials with sequences of probabilities p_i other than the sequence $p_i = \theta/(\theta + i - 1)$ involved in this coupling with a sequence of Ewens(θ) permutations.

The coupling described above provides a way to define a sequence of Ewens(θ) random permutations $(\pi_{n,\theta})_{n \geq 1}$ whose cycle structures for different values of n are strongly related: from $\pi_{n,\theta}$ to $\pi_{n+1,\theta}$, either a single fixed point is added, or a single cycle of $\pi_{n,\theta}$ has its length increased by one. However, the coupling above does not uniquely define a joint distribution for $\pi_{n,\theta}$ and $\pi_{n+1,\theta}$, because it does not say how the content of the cycles of $\pi_{n+1,\theta}$ and $\pi_{n,\theta}$ are related.

In the particular case $\theta = 1$, when each $\pi_{n,1}$ is a uniform random element of the set $\mathfrak{S}_n$ of permutations of $[n]$, Ignatov [13] provides a nice construction which defines the joint law of $(\pi_{n,1})_{n \geq 1}$ in a unique way. Let $(U_i)_{i \geq 1}$ be a sequence of pairwise distinct elements of $[0, 1]$, with no smallest element. From this sequence, define the *lower record indices* $I_1 < I_2 < I_3 < \ldots$, as the set of indices I such that U_I is smaller than U_i for all $i < I$, the *lower indicators* $(B_i)_{i \geq 1}$, given by $B_i = 1$ if i is a lower record index and by $B_i = 0$ otherwise, and the *inter-record stretches* $(V_k)_{k \geq 1}$ given by:

$$V_k := (U_{I_k}, U_{I_k+1}, \ldots, U_{I_{k+1}-1}). \qquad (1.4)$$

We notice the following facts:

- the inter-record stretches are elements of the space $\cup_{\ell=1}^{\infty}[0, 1]^\ell$ of finite sequences in $[0, 1]$ with undetermined length;
- the first term of the stretch V_k is the k-th lower record value $R_k := U_{I_k}$;
- this first term R_k of V_k is the minimal term of the stretch V_k;
- the length of the stretch V_k is $I_{k+1} - I_k$, the k-th *inter-record spacing*.

We can then define, for all $n \geq 1$, a permutation π_n^U of $\{U_1, \ldots, U_n\}$ whose cycle structure is given by the inter-record stretches: more precisely, $\pi_n^U(U_{i-1}) = U_i$ for all $i \in \{2, \ldots, n\}$ which are not lower record indices, $\pi_n^U(U_{I_{k+1}-1}) = U_{I_k}$ if $k \geq 1$ is such that $I_{k+1} - 1 \leq n$, and $\pi_n^U(U_n) = U_{I_j}$ where I_j is the last lower record index such that $I_j \leq n$. The permutation π_n^U acts on the set $\{U_1, \ldots, U_n\}$: it induces a permutation π_n of $[n]$ if we rename the m-th smallest element of this set by m (for example, the permutation $0.2 \mapsto 0.9, 0.4 \mapsto 0.5, 0.5 \mapsto 0.2, 0.9 \mapsto 0.4$ induces the permutation $1 \mapsto 4, 2 \mapsto 3, 3 \mapsto 1, 4 \mapsto 2$). It is not difficult to check that the permutation π_n depends only on the relative order of $U_1, \ldots, U_n$, in a way which induces a bijective map from $\mathfrak{S}_n$ to itself. This

bijection was proposed by Rényi [17] in the early 60s, and called the "transformation fondamentale" in a paper by Foata and Schützenberger [8], in a more general setting of combinatorics on words. Diaconis and Pitman [5] exploited this bijection to obtain the convergence in distribution (1.3) in the case $\theta = 1$, with a total variation bound. This bound was sharpened and extended to the case of a general parameter $\theta > 0$ in [1], as indicated above. But this argument for general θ loses track of the full Poisson structure of the record process for $\theta = 1$.

Let us recall how this Poisson structure for $\theta = 1$ was first exposed by Ignatov [13]. If $(U_i)_{i \geq 1}$ is a sequence of i.i.d., uniform variables on $[0, 1]$, then for all $n \geq 1$, all possible orders of $U_1, \ldots, U_n$ occur with the same probability and then π_n is uniformly distributed on $\mathfrak{S}_n$. On the other hand, the lower record indicators $(B_i)_{i \geq 1}$ are independent, Bernoulli variables, B_i having parameter $1/i$. The link between the construction of π_n and the Feller coupling is the following: conditionally on $(B_i)_{i \geq 1}$, the distribution of π_n is uniform on the set of permutations whose lengths of the cycles, ordered by increasing lowest element, are equal to the successive spacings between the 1's in the sequence $(1, B_n, B_{n-1}, \ldots, B_1)$. One easily deduces the following result: given $(B_i)_{1 \leq i \leq n}$, the conditional distribution of π_n is the same as that given by the Feller coupling procedure using $B_i(1) = B_i$ for all i. The Poisson structure obtained by taking all the inter-record stretches together is quite remarkable:

Theorem 1.2. *If the variables $(U_i)_{i \geq 1}$ are i.i.d., uniform in $[0, 1]$, then the inter-record stretches $(V_k)_{k \geq 1}$ form a Poisson point process on $\cup_{\ell=1}^{\infty} [0, 1]^{\ell}$ with mean measure*

$$\mu(\bullet) = \sum_{\ell=1}^{\infty} \frac{P_{\ell}(\bullet)}{\ell} \tag{1.5}$$

where $P_{\ell}(\bullet)$ is the conditional distribution of $(U_1, \ldots, U_{\ell})$ given that $U_1 < U_i$ for every $1 < i \leq \ell$.

To illustrate the notation:

- $P_1(\bullet)$ is the uniform distribution of U_1 on $[0, 1]$,
- $P_2(\bullet)$ is uniform on $\{(u_1, u_2) : u_1 < u_2\} \subseteq [0, 1]^2$, that is the conditional distribution of (U_1, U_2) given the event $(U_1 < U_2)$ of probability $1/2$.
- $P_3(\bullet)$ is uniform on $\{(u_1, u_2, u_3) : u_1 < \min(u_2, u_3)\} \subseteq [0, 1]^3$, that is the conditional distribution of (U_1, U_2, U_3) given the event $(U_1 < \min(U_2, U_3))$ of probability $1/3$,

and so on.

Theorem 1.2 is a straightforward extension of the result of Ignatov [13] that $\{(R_k, I_{k+1} - I_k), k \geq 1\}$ is the collection of points of a Poisson point process on $(0, 1) \times \{1, 2, \ldots\}$ with mean number of points in $(s, 1] \times \{\ell\}$ equal to $(1-s)^{\ell}/\ell$. See Resnick [18, Proposition 4.1 (iv)] for a proof of this result using the basic spraying property of Poisson processes [18, Proposition 3.8]. The same spraying argument gives the stronger assertion of Theorem 1.2. For it is easily seen that given all the points $\{(R_k, I_{k+1} - I_k), k \geq 1\}$, for each particular k, the conditional distribution of the stretch V_k depends only on R_k and $I_{k+1} - I_k$, and given $R_k = r$ and $I_{k+1} - I_k = \ell$, the stretch V_k with initial term r and length ℓ has the distribution of $(U_1, \ldots, U_{\ell})$ given $U_1 = r$ and $r < U_i$ for all $1 < i \leq \ell$.

If we only consider the length of the inter-record stretches, we immediately deduce from Theorem 1.2 that the inter-record spacings $(I_{k+1} - I_k)_{k \geq 1}$, form a Poisson point process on the positive integers, with intensity $1/\ell$ at ℓ, meaning that the random variables

$$K_{\infty, \ell} := \sum_{k=1}^{\infty} 1(I_{k+1} - I_k = \ell) \tag{1.6}$$

are independent Poisson variables with means $1/\ell$. This results corresponds to the case $\theta = 1$ of Theorem 1.1, since the lower record indicators $(B_i)_{i \geq 1}$ are independent, Bernoulli variables, B_i having parameter $1/i$.

Regarded as a fact about inhomogeneous Bernoulli trials, this result is not at all obvious without a broader context involving additional randomization, such as Ignatov's context of record sequences, or the context of the Feller coupling for random permutations.

The link between the Feller coupling and the lower records of a sequence of random variables can be extended to the setting of Ewens distributed permutations with general parameter $\theta > 0$, by changing the distribution of the sequence $(U_i)_{i \geq 1}$. We will prove the following result:

Theorem 1.3. *For $\theta > 0$, let $\mathbb{P}_\theta$ be the probability measure on the set of infinite sequences in $[0, 1]$, endowed with its Borel σ-algebra, such that for $(U_i)_{i \geq 1}$ following the law $\mathbb{P}_\theta$:*

- *The first term U_1 is Beta distributed with parameters θ and 1.*
- *Conditionally on $(U_1, \ldots, U_n)$, for $\min(U_1, \ldots, U_n) = r$, the distribution of U_{n+1} is the mixture with weights r and $1 - r$ of the distribution of r times a Beta variable with parameters θ and 1, and the uniform distribution on $[r, 1]$, i.e.*

$$\frac{\mathbb{P}_\theta(U_{n+1} \in du \mid \min(U_1, \ldots, U_n) = r)}{du} = \theta \left(\frac{u}{r}\right)^{\theta - 1} 1(u < r) + 1(u \geq r) \quad (0 < u < 1).$$

In particular, no matter what $\theta > 0$, the conditional probability of a new lower record at time $n + 1$, given $(U_1, \ldots, U_n)$, is always $\min(U_1, \ldots, U_n)$.

Then, the following statements hold:

- *The finite dimensional distributions of $(U_1, \ldots, U_n)$ under $\mathbb{P}_\theta$ are absolutely continuous with respect to the Lebesgue measure on $[0, 1]^n$, with density*

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}_1}(u_1, \ldots, u_n) = \theta^{K_n} \min(u_1, \ldots, u_n)^{\theta - 1} \tag{1.7}$$

  where K_n is the number of lower records in the sequence $(u_1, \ldots, u_n)$. In particular, under $\mathbb{P}_1$, the variables $(U_i)_{i \geq 1}$ are i.i.d., uniform on $[0, 1]$.
- *If $(U_i)_{i \geq 1}$ follows the law $\mathbb{P}_\theta$, then this sequence has a.s. no smallest element, the U_i's are pairwise distinct, and the inter-record stretches $(V_k)_{k \geq 1}$ form a Poisson point process on $\cup_{\ell=1}^\infty [0, 1]^\ell$ with mean measure $\theta \mu(\bullet)$ for $\mu(\bullet)$ as in (1.5).*

The fact that $(U_i)_{i \geq 1}$ are i.i.d., uniform under $\mathbb{P}_1$ is a restatement of Ignatov's Theorem 1.2. The description of the law of $(U_1, \ldots, U_n)$ under $\mathbb{P}_\theta$ for general θ has already been indicated by Kerov and Tsilevich [15, Lemma 2], with upper rather than lower records, which exchanges U_i with $1 - U_i$ in the formulas. Kerov and Tsilevich have also associated random permutations to sequences following the distribution $\mathbb{P}_\theta$, and these permutations are distributed with respect to the Ewens measure of parameter θ. However, the construction of [15] does not coincide with the construction given in the present paper

The fact that $\mathbb{P}_\theta$ may also be described as in Theorem 1.3, by simply changing the mean intensity measure of the Poisson point process of inter-record stretches on $\cup_{\ell=1}^\infty [0, 1]^\ell$ by a scalar factor of θ, from $\mu(\bullet)$ under $\mathbb{P}_1$ to $\theta \mu(\bullet)$ under $\mathbb{P}_\theta$, does not seem to have been observed before.

The push forward of this result, from the Poisson point process of inter-record stretches to the Poisson point process of their lengths, gives the fact that the counting of the inter-record spacings $(I_{k+1} - I_k)_{k \geq 1}$ forms a Poisson point process on the positive integers, with intensity θ/ℓ at ℓ. The following corollary links Theorem 1.3 to the Feller coupling and explains how Theorem 1.3 implies Theorem 1.1.

Corollary 1.4. *For a sequence $(U_i)_{i\geq 1}$ of pairwise distinct elements of $[0,1]$, with no smallest element, let $(B_i)_{i\geq 1}$ be the corresponding lower record indicators, let π_n^U be the permutation of $\{U_1,\ldots,U_n\}$ whose cycle structure is given by the inter-record stretches, and let π_n be the corresponding permutation of $[n]$, with the same cycle structure. Then, for $(U_i)_{i\geq 1}$ governed by the law $\mathbb{P}_\theta$,*

- *The B_i are independent Bernoulli $(\theta/(i-1+\theta))$;*
- *Conditionally on all the B_i, the permutation π_n is uniformly distributed among the permutations whose cycle lengths, ordered by increasing lowest elements, are given by the successive spacings between 1's in the sequence $(1, B_n, B_{n-1}, \ldots, B_1)$;*
- *Given $(B_i)_{1\leq i\leq n}$, the conditional distribution of π_n is the same as that given by the Feller coupling procedure using $B_i(\theta) = B_i$ for all i;*
- *The unconditional distribution of π_n is Ewens with parameter θ.*

Another thing we can remark is that all the values of $\theta > 0$ can be coupled on a single probability space. Indeed, under $\mathbb{P}_\theta$, the family of inter-record stretches forms a Poisson point process of intensity $\theta\mu(\bullet)$, so can be constructed simultaneously for all θ by taking the points of a Poisson process of intensity equal to the product of Lebesgue measure on $\mathbb{R}_+$ by the measure $\mu(\bullet)$, and extracting the points for which the $\mathbb{R}_+$ coordinate is smaller than θ. Such a coupling provides a dynamic version of the Feller coupling, with the parameter θ of the Ewens measure as its time parameter. The path structure of this $\mathfrak{S}_n$-valued process $(\pi_{n,\theta}, \theta \geq 0)$ can be understood as follows. It may be constructed with right-continuous step function paths, in which each jump involves insertion of a new cycle of some length ℓ from 1 to n, corresponding to a Poisson point which is a sequence in some component $[0,1]^k$ of the sequence space with $k \geq \ell$, whose initial term is greater than the initial term of at least one sequence contributing to the current permutation of $[n]$. This insertion may delete some cycles, and/or shorten the final cycle, depending on the rank of the initial term of the new sequence relative to the initial terms associated with existing cycles. It does not seem easy to give a full probabilistic description of the dynamics of this $\mathfrak{S}_n$-valued process. In particular, it may not be Markovian, due to the latent initial terms of the sequential fragments which determine the order of the cycles. As the partition of n induced by $\pi_{n,\theta}$ is not necessarily refining as θ increases, this process is not the same as the evolution described by Gnedin and Pitman [10], in which partitions following the Ewens (θ) distribution are constructed for all values of $\theta > 0$ to be refining as θ increases.

Theorem 1.3 and Corollary 1.4 are proven in Section 2 of the present article. In Section 3, we use the measure $\mathbb{P}_\theta$ in order to construct some infinite random permutations, in a way which generalizes the Feller coupling. In Section 4, we provide a link between our construction and a result by Shepp and Lloyd on the cycle counts of permutations of random order.

2 Proof of Theorem 1.3 and Corollary 1.4

By induction on n, using the definition of $\mathbb{P}_\theta$, we see that the density at $(u_1,\ldots,u_n)$ of the law of $(U_1,\ldots,U_n)$ under $\mathbb{P}_\theta$ is given by

$$\theta u_1^{\theta-1} \prod_{2\leq j\leq n\,:\,u_j=\min(u_1,\ldots,u_j)} \left[\theta\left(\frac{\min(u_1,\ldots,u_j)}{\min(u_1,\ldots,u_{j-1})}\right)^{\theta-1}\right] = \theta^{K_n}(\min(u_1,\ldots,u_n))^{\theta-1},$$

which proves the first statement of the theorem. It is also clear from the definition that $(U_i)_{i\geq 1}$ has a.s. no smallest element and all elements pairwise distinct.

Now, for a given $s \in (0,1)$, $n \geq 0$ and $u_1, \ldots, u_n \in (s, 1)$, we easily check that

$$\frac{\mathbb{P}_\theta(U_1 \in du_1, \ldots, U_n \in du_n, U_{n+1} \leq s)}{du_1 \ldots du_n}$$

$$= \theta^{K_n} (\min(u_1, \ldots, u_n))^{\theta-1} \int_0^s \left[\theta \left(\frac{u}{\min(u_1, \ldots, u_n)} \right)^{\theta-1} \right] du$$

$$= \theta^{K_n} \int_0^s \theta u^{\theta-1} du = \theta^{K_n} s^\theta.$$

Let $(\mathcal{F}_i)_{i \geq 0}$ be the filtration generated by the variables $(U_i)_{i \geq 1}$, and let T_s be the first index i such that $U_i \leq s$: it is clear that T_s is a stopping time with respect to $(\mathcal{F}_i)_{i \geq 0}$. The equality above shows that for any event A_n which is $\mathcal{F}_n$-measurable,

$$\mathbb{P}_\theta(A_n, T_s = n+1) = s^{\theta-1} \mathbb{E}_{\mathbb{P}_1}[1(A_n, T_s = n+1) \theta^{K_n}]$$

where K_n is the number of lower records in the sequence $(U_1, \ldots, U_n)$. Now, let E be an event which is measurable with respect to the family of all inter-record stretches starting above the level s, and let L be the total length of these stretches. We can check that for $n \geq 0$, the intersection of E and the event $\{L = n\}$ can be written as the intersection of A_n and $T_s = n+1$ for some $\mathcal{F}_n$-measurable event A_n, which gives

$$\mathbb{P}_\theta(E, L = n) = s^{\theta-1} \mathbb{E}_{\mathbb{P}_1}[1(E, L = n) \theta^{N_s}]$$

where N_s is the number of inter-record stretches starting above the level s. Hence

$$\mathbb{P}_\theta(E) = s^{\theta-1} \mathbb{E}_{\mathbb{P}_1}[1(E) \theta^{N_s}],$$

which implies the following:

- The law of N_s under $\mathbb{P}_\theta$ has density $\theta^{N_s} s^{\theta-1}$ with respect to the law of N_s under $\mathbb{P}_1$.
- Conditionally on N_s, the set of inter-record stretches starting above the level s has the same law under $\mathbb{P}_\theta$ and under $\mathbb{P}_1$.

By Theorem 1.2, the law of N_s under $\mathbb{P}_1$ is the Poisson distribution with parameter

$$\sum_{\ell=1}^\infty \frac{P_\ell([s,1]^\ell)}{\ell} = \sum_{\ell=1}^\infty \frac{(1-s)^\ell}{\ell} = -\log s$$

and we deduce that the law of N_s under $\mathbb{P}_\theta$ is the Poisson distribution of parameter $-\theta \log s$. Moreover, conditionally on N_s, under $\mathbb{P}_1$, and then also under $\mathbb{P}_\theta$, the family of inter-record stretches starting above s has the same law as the family of elements of an i.i.d. sequence of variables which are distributed according to the probability measure:

$$-\frac{1}{\log s} \sum_\ell \frac{(P_\ell)_{|[s,1]^\ell}}{\ell}.$$

Hence, under $\mathbb{P}_\theta$, the inter-record stretches starting above s form a Poisson point process with intensity

$$\theta \sum_\ell \frac{(P_\ell)_{|[s,1]^\ell}}{\ell}.$$

Since $s \in (0,1)$ can be arbitrarily chosen, we get the second statement of Theorem 1.3.

For the corollaries, we use the following key property: the density on $\mathcal{F}_n$ of $\mathbb{P}_\theta$ with respect to $\mathbb{P}_1$ can be written as the product of a function of the relative order of $(U_1, \ldots, U_n)$,

i.e. θ^{K_n}, and a function of the order statistics of $(U_1, \ldots, U_n)$, i.e. $(\min(U_1, \ldots, U_n))^{\theta-1}$. Since the relative order and the order statistics of $(U_1, \ldots, U_n)$ are independent under $\mathbb{P}_1$, they remain independent under $\mathbb{P}_\theta$ for all $\theta > 0$. Moreover, under $\mathbb{P}_1$, the record indicators B_i are independent Bernoulli$(1/i)$ variables, and the change of measure on these variables when we go from $\mathbb{P}_1$ to $\mathbb{P}_\theta$ corresponds to a density factor proportional to

$$\theta^{K_n} = \prod_{i=1}^{n} \theta^{B_i}.$$

This easily implies that under $\mathbb{P}_\theta$, the record indicators are independent Bernoulli$(p_i(\theta))$, where

$$p_i(\theta) = \frac{\theta/i}{1 - 1/i + \theta/i} = \frac{\theta}{i - 1 + \theta}.$$

This gives the first item of Corollary 1.4. Moreover, conditionally on the lower record indicators and the order statistics of $(U_1, \ldots, U_n)$, all the possible relative orders of $(U_1, \ldots, U_n)$ have the same probability, because of the form of the density of $\mathbb{P}_\theta$ with respect to $\mathbb{P}_1$. This implies the second item of Corollary 1.4, since the permutation π_n is uniquely determined by the relative order of $(U_1, \ldots, U_n)$. The third item of Corollary 1.4 is a direct consequence of the two first items, and the last item is due to the classical properties of the Feller coupling.

3 Infinite permutations

From any sequence $(U_i)_{i \geq 1}$ of elements in $[0, 1]$, with distinct values and no smallest element, we have seen how to construct a permutation π_n^U of $\{U_1, \ldots, U_n\}$ and a permutation π_n of $[n]$ from the inter-record stretches. It is also possible to define a permutation π_∞^U of the infinite set $\{U_i, i \geq 1\}$, in such a way that the cycles are given by the set of all inter-record stretches, i.e. $\pi_\infty^U(U_{i-1}) = U_i$ for all $i \geq 2$ which are not lower record indices, and $\pi_\infty^U(U_{I_{k+1}-1}) = U_{I_k}$ for all $k \geq 1$. One easily checks that π_∞^U coincides with π_n^U on the set $\{U_1, \ldots, U_{n-1}\}$ for all $n \geq 1$.

The construction of π_∞^U can be seen as some kind of Feller coupling of infinite order, since the construction of π_n^U and π_n can be related with the Feller coupling of order n, as we have seen previously. However, we observe that contrary to the case of the permutation π_n which acts on the fixed set $[n]$, the infinite set on which π_∞^U acts is itself a random set. Moreover, we observe that the cycles of π_∞ appear in decreasing order of their smallest element, i.e. in the reverse order with respect to the usual description of the Feller coupling. If we look at the sequence of permutations $(\pi_n^U)_{n \geq 1}$ we get a coupling of permutations of different orders, which has the property noted in the analysis of [1, §3], that "the cycles are built and completed one by one, in contrast to the Chinese Restaurant Process", with reference to the alternative construction of cycle-consistent random permutations of $[n]$ discussed in [1, §2], and [16].

If the sequence of variables $(U_i)_{i \geq 1}$ is distributed like $\mathbb{P}_\theta$, then by Theorem 1.3, the cycle structure of π_∞^U is directly given by a Poisson point process on $\cup_{\ell=1}^{\infty}[0, 1]^{\ell}$ with mean measure $\theta\mu(\bullet)$. In particular, the number of cycles of different lengths ℓ is given by independent Poisson random variables of parameter θ/ℓ, which generalizes the case $\theta = 1$ studied by Ignatov.

If we consider, as at the end of the introduction, the dynamical version of our construction, where all the values of $\theta > 0$ are coupled together, then the evolution of the cycle structure of π_∞^U when θ varies is easy to describe in terms of Poisson processes, contrary to the case where we consider permutations of finite order. In particular, the set of cycles of the permutation corresponding to $\theta = \theta_1 + \theta_2$ has the same law as the

union of two independent sets of cycles, corresponding to the parameters $\theta = \theta_1$ and $\theta = \theta_2$.

4 Connection with work by Shepp and Lloyd

In this section, we connect Theorem 1.3 to a model for a random permutation π of a set of random size N, first introduced by the work of Shepp and Lloyd [20] on the distribution of the lengths of the longest and shortest cycles of a uniform random permutation. In the Shepp and Lloyd model, N is assigned the geometric (p) distribution $\mathbb{P}(N \geq n) = (1-p)^n$ for $n \geq 0$. In a following paper [4], Balakrishnan, Sankaranarayanan, and Suyambulingom extended the model of Shepp and Lloyd to a much more general model of random permutations π of a set of random size N. For a particular choice of parameters, which was not singled out for special discussion in [4], the model of [4] assigns N a negative binomial distribution, as in the following Corollary, and given $N = n$ the permutation π is governed by the Ewens(θ) distribution. See also [9], [12], [11] (Lemma 2.1). [21] (Theorem 2) and [14] for variants of this result with different interpretations, and further references.

Corollary 4.1 ([20], [4], [9], [12])**.** *Let $\theta > 0$ and $p \in (0, 1)$, and let $N(\theta, p)$ denote a random variable with the negative binomial (θ, p) distribution:*

$$\mathbb{P}(N(\theta, p) = n) = \frac{(\theta)_n}{n!}(1-p)^n p^\theta$$

for $n \geq 1$, which implies that $\mathbb{E}N(\theta, p) = \theta(1-p)/p$. Let π be a permutation of random order, such that conditionally given $N(\theta, p) = n$, π has order $[n]$ and is distributed according to the Ewens(θ) measure. Then the number of cycles of π of different orders ℓ are independent Poisson variables of parameter $(1-p)^\ell/\ell$.

Proof. Let us consider, under $\mathbb{P}_\theta$, the permutation π_n of random order, $n + 1$ being the first index such that $U_{n+1} < p$. If we condition on the value of this index and on the order statistics of $(U_1, \ldots, U_n)$, we get, from the expression of the density $d\mathbb{P}_\theta/d\mathbb{P}_1$, a permutation π_n following Ewens distribution of parameter θ. On the other hand, Theorem 1.3 implies that the number of cycles of different sizes in π_n are independent Poisson variables, the expectation of the number of ℓ-cycles being:

$$\theta\mu([p, 1]^\ell) = \theta(1-p)^\ell/\ell.$$

Hence, the corollary is proven if we show that the law of the size n of the permutation is negative binomial (θ, p). Since the cycle lengths form a Poisson process with intensity proportional to θ when p is fixed, the law of the size of the permutation in function of θ corresponds to the marginals of a Lévy process. It is also the case for the negative binomial distribution, so it is enough to check that for $\theta = 1$, n is geometrically distributed with parameter p. This fact is immediate since $n + 1$ is the first time when an i.i.d. sequence of uniform variables on $[0, 1]$ hits the interval $[0, p]$. $\square$

The proof above is related to the fact that if $K_\ell(\pi)$ is the number of ℓ-cycles of π,

$$N(\theta, p) = \sum_{\ell=1}^{\infty} \ell K_\ell(\pi) \tag{4.1}$$

is the canonical Lévy decomposition of the infinitely divisible distribution of $N(\theta, p)$ as a linear combination of independent Poisson variables. Compare with the discussion of

Feller [7, (2.17)] who gives the well-known probability generating function of the number K of cycles of a uniform random permutation π_n of $[n]$:

$$\mathbb{E}_1 \theta^K = \frac{(\theta)_n}{n!} \tag{4.2}$$

by use of his coupling with Bernoulli$(1/i)$ variables for $1 \leq i \leq n$. This comes immediately after discussion of the compound Poisson representation of the negative binomial distribution, but without the connection indicated in Corollary 4.1. This model for constructing a negative binomial variable from independent Poisson counts of cycles of a random permutation of random size is also not mentioned in the otherwise very comprehensive account [2] of models related to the Ewens sampling formula.

References

[1] R. Arratia, A. D. Barbour, and S. Tavaré. Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.*, 2(3):519–535, 1992. MR-1177897

[2] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003. MR-2032426

[3] R. Arratia, A. D. Barbour, and S. Tavaré. Exploiting the Feller coupling for the Ewens sampling formula. *Statist. Sci.*, 31(1):27–29, 2016. MR-3458588

[4] V. Balakrishnan, G. Sankaranarayanan, and C. Suyambulingom. Ordered cycle lengths in a random permutation. *Pacific J. Math.*, 36:603–613, 1971. MR-0391216

[5] P. Diaconis and J. Pitman. Permutations, record values and random measures. Unpublished lecture notes of a course at Dept. Statistics, Univ. California, Berkeley, 1986.

[6] W. Feller. The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.*, 51(11):800–832, 11 1945. MR-0013252

[7] W. Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968. MR-0228020

[8] D. Foata and M. P. Schützenberger. *Théorie géométrique des polynômes eulériens*. Lecture Notes in Mathematics, Vol. 138. Springer-Verlag, Berlin-New York, 1970. MR-0272642

[9] B. Fristedt. The structure of random partitions of large sets. Technical Report, Dept. of Mathematics, University of Minnesota, 1987.

[10] A. Gnedin and J. Pitman. Poisson representation of a Ewens fragmentation process. *Combin. Probab. Comput.*, 16(6):819–827, 2007. MR-2351686

[11] J. C. Hansen. A functional central limit theorem for the Ewens sampling formula. *J. Appl. Probab.*, 27(1):28–43, 1990. MR-1039182

[12] F. M. Hoppe. Faà di Bruno's formula and the distributions of random partitions in population genetics and physics. *Theoretical Population Biology*, 73(4):543–551, 2008.

[13] Z. Ignatov. Point processes generated by order statistics and their applications. In *Point processes and queuing problems (Colloq., Keszthely, 1978)*, volume 24 of *Colloq. Math. Soc. János Bolyai*, pages 109–116. North-Holland, Amsterdam-New York, 1981. MR-0617405

[14] P. Joyce and S. Tavaré. Cycles, permutations and the structure of the Yule process with immigration. *Stochastic Processes and their Applications*, 25:309–314, 1987. MR-0915145

[15] S. V. Kerov and N. V. Tsilevich. Stick breaking process generated by virtual permutations with Ewens distribution. *J. Math. Sciences*, 87(6):4082–4093, 1997. Translation of Zap. Nauchn. Semin. POMI, 223, 162–180 (1995). MR-1374319

[16] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. MR-2245368

[17] A. Rényi. Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.*, 8:7–13, 1962. MR-0286162

[18] S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987. MR-0900810

[19] J. Sethuraman and S. Sethuraman. Connections between Bernoulli strings and random permutations. In *The legacy of Alladi Ramakrishnan in the mathematical sciences*, pages 389–399. Springer, New York, 2010. MR-2744276

[20] L. A. Shepp and S. P. Lloyd. Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.*, 121:340–357, 1966. MR-0195117

[21] G. A. Watterson. The sampling theory of selectively neutral alleles. *Advances in Applied Probability*, 6(3):463–488, 1974. MR-0351504

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm