SUPER-CLUSTERING OF CONSECUTIVE NUMBERS IN $p$-SHIFTED RANDOM PERMUTATIONS

ROSS G. PINSKY

Abstract. Let $A_{l,k}^{(n)} \subset S_n$ denote the event that the set of $l$ consecutive numbers \{k, k+1, \ldots, k+l-1\} appear in a set of $l$ consecutive positions. Let $p = \{p_j\}_{j=1}^{\infty}$ be a distribution on $\mathbb{N}$ with $p_j > 0$. Let $P_n$ denote the probability measure on $S_n$ corresponding to the $p$-shifted random permutation. Our main result, under the additional assumption that $\{p_j\}_{j=1}^{\infty}$ is non-increasing, is that

$$\lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{l,k}^{(n)}) = \left( \frac{k-1}{\prod_{j=1}^{\infty} j \sum_{i=1}^{p_i}} \right) \left( \prod_{j=1}^{\infty} \sum_{i=1}^{p_i} \right),$$

and that if $\lim_{n \to \infty} \min(k_n, n-k_n) = \infty$, then

$$\lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{l,k_n}^{(n)}) = \left( \prod_{j=1}^{\infty} \sum_{i=1}^{p_i} \right)^2.$$

In particular these limits are positive if and only if $\sum_{j=1}^{\infty} j p_j < \infty$.

We say that super-clustering occurs when the limits are positive. We also give a new characterization of the class of $p$-shifted probability distributions on $S_\infty$.

1. Introduction and Statement of Results

Let $S_n$ denote the set of permutations of $[n] := \{1, \ldots, n\}$. For each $j \in \{2, \ldots, n\}$, define the backward rank $I_{<j}(\sigma)$ of a permutation $\sigma \in S_n$ to be the number of inversions involving $j$ and a number less than $j$; that is,

$$I_{<j}(\sigma) = |\{1 \leq i < j : \sigma_i^{-1} < \sigma_j^{-1}\}|.$$

As is well-known, a permutation is uniquely determined by its backward ranks; more specifically, for each vector $(i_2, \ldots, i_n)$ satisfying $0 \leq i_j \leq j-1$, there exists a unique permutation in $\sigma \in S_n$ satisfying $I_{<j}(\sigma) = i_j, j = 2000$ Mathematics Subject Classification. 60C05, 05A05.

Key words and phrases. random permutation, $p$-shifted, backwards rank, clustering, inversion, Mallows distribution.
Similarly, let \( S_\infty \) denote the set of permutations of \( \mathbb{N} \). Then \( I_{<j}(\sigma) \) is defined as above, for \( \sigma \in S_\infty \) and all \( j \in \mathbb{N} \). It is still true that \( \sigma \) is uniquely defined by its backward ranks \( \{I_{<j}\}_{j=1}^\infty \), however not all sets of backward ranks lead to a permutation on \( S_\infty \); for example, there is no permutation \( \sigma \) satisfying \( I_{<j}(\sigma) = j - 1 \), for all \( j \in \mathbb{N} \).

The set of \( l \) consecutive numbers \( \{k, k+1, \ldots, k+l-1\} \subseteq [n] \) appears in a set of consecutive positions in the permutation \( \sigma \in S_n \) if there exists an \( m \) such that \( \{k, k+1, \ldots, k+l-1\} = \{\sigma_m, \sigma_{m+1}, \ldots, \sigma_{m+l-1}\} \). Let \( A_{l,k}^{(n)} \subseteq S_n \) denote the event that the set of \( l \) consecutive numbers \( \{k, k+1, \ldots, k+l-1\} \) appear in a set of \( l \) consecutive positions. In this paper, for a certain class of random permutations, we present a result that gives a quantitative expression of the inverse correlation between the tendency towards inversions and the tendency towards the clustering of consecutive numbers. We will consider sequences \( \{P_n\}_{n=1}^\infty \) of probability measures on the sequence of spaces \( \{S_n\}_{n=1}^\infty \) and will calculate explicitly the expression \( \lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{l,k}^{(n)}) \). When this limit is positive, we will say that the sequence of random permutations exhibits super-clustering.

We now introduce the class of random permutations we will study. A sequence \( \{P_n\}_{n=1}^\infty \) of probability measures on the sequence of spaces \( \{S_n\}_{n=1}^\infty \) will be obtained via a natural projection from a probability measure \( P \) on \( S_\infty \). The class of probability measures \( P \) on \( S_\infty \) will be the so-called \( p \)-shifted distributions, which we now describe.

Let \( p := \{p_j\}_{j=1}^\infty \) be a probability distribution on \( \mathbb{N} \) whose support is all of \( \mathbb{N} \); that is, \( p_j > 0 \) for all \( j \in \mathbb{N} \). Take a countably infinite sequence of independent samples from this distribution: \( n_1, n_2, \ldots \). Now construct a random permutation \( \Pi \in S_\infty \) as follows. Let \( \Pi_1 = n_1 \) and then for \( k \geq 2 \), let \( \Pi_k = \psi_k(n_k) \), where \( \psi_k \) is the increasing bijection from \( \mathbb{N} \) to \( \mathbb{N} - \{\Pi_1, \ldots, \Pi_{k-1}\} \). Thus, for example, if the sequence of samples \( \{n_j\}_{j=1}^\infty \) begins with 7, 3, 4, 3, 7, 2, 1, then the construction yields the permutation \( \Pi \) beginning with \( \Pi_1 = 7, \Pi_2 = 3, \Pi_3 = 5, \Pi_4 = 4, \Pi_5 = 11, \Pi_6 = 2, \Pi_7 = 1. \) The probability measure \( P \) on \( S_\infty \) is then the distribution of this random permutation \( \Pi \). We call \( P \) the \( p \)-shifted distribution and \( \Pi \) a \( p \)-shifted random permutation.
For $\sigma \in S_\infty$, write $\sigma = \sigma_1\sigma_2 \cdots$. For $n \in \mathbb{N}$, define $\text{proj}_n(\sigma) \in S_n$ to be the permutation obtained from $\sigma$ by deleting $\sigma_i$ for all $i$ satisfying $\sigma_i > n$. Thus, if $n = 4$ and $\sigma = 2539461 \cdots$, then $\text{proj}_n(\sigma) = 2341$.

Given the $p$-shifted random permutation $\Pi \in S_\infty$ that was constructed in the previous paragraph, define $\text{proj}_n$, as the distribution of the random permutation $\text{proj}_n(\Pi)$. Equivalently, given the probability measure $P$ on $S_\infty$ defined in the previous paragraph, define the probability measure $P_{\text{proj}_n}$ on $S_n$ by $P_{\text{proj}_n}(\sigma) = P(\text{proj}_{n-1}^{-1}(\sigma))$, $\sigma \in S_n$.

We will call $P_{\text{proj}_n}$ the $p$-shifted distribution on $S_n$ and $\text{proj}_n(\Pi)$ a $p$-shifted random permutation on $S_n$. Define $P_n = P_{\text{proj}_n}$. This gives us the sought after sequence $\{P_n\}_{n=1}^\infty$ of probability measures on the sequence of spaces $\{S_n\}_{n=1}^\infty$. We note that in the case that $p_j = 1 - q$, where $q \in (0, 1)$, the measure $P_n$ is the so-called Mallows distribution with parameter $q [3]$.

Remark. We assume in this paper that $p_j > 0$, for all $j$. In fact, the $p$-shifted random permutation can be constructed as long as $p_1 > 0$, with no positivity requirement on $p_j$, $j \geq 2$. The positivity requirement for all $j$ ensures that for all $n$, the support of the $p$-shifted measure $P_n$ is all of $S_n$.

It is known [3] that a random permutation under the $p$-shifted measure $P$ is strictly regenerative, where our definition of strictly regenerative is as follows. For a permutation $\pi = \pi_{a+1}\pi_{a+2} \cdots \pi_{a+m}$, of $\{a+1, a+2, \cdots, a+m\}$, define $\text{red}(\pi)$, the reduced permutation of $\pi$, to be the permutation in $S_m$ given by $\text{red}(\pi)_i = \pi_{a+i} - m$. We will call a random permutation $\Pi$ of $S_\infty$ strictly regenerative if almost surely there exist $0 = T_0 < T_1 < T_2 < \cdots$ such that $\Pi([T_j]) = [T_j]$, $j \geq 1$, and $\Pi([m]) \neq [m]$ if $m \notin \{T_1, T_2, \cdots\}$, and such that the random variables $\{T_k - T_{k-1}\}_{k=1}^\infty$ are IID and the random permutations $\{\text{red}(\Pi[T_k]-[T_{k-1}])\}_{k=1}^\infty$ are IID. The numbers $\{T_n\}_{m=1}^\infty$ are called the renewal or regeneration numbers. Our definition of strictly regenerative differs slightly from that in [3].

Let $u_n$ denote the probability that the $p$-shifted random permutation $\Pi$ has a renewal at the number $n$, that is, the probability that $\Pi([n]) = [n]$. It follows easily from the construction of the random permutation that

$$u_n = \prod_{j=1}^n \left( \sum_{i=1}^j p_i \right) = \prod_{j=1}^n (1 - \sum_{i=j+1}^\infty p_i).$$

(1.1)
Thus, \( u_n > 0 \), for all \( n \). (Note that this positivity, and the consequent aperiodicity of the renewal mechanism, does not require the positivity of all \( p_j \), but only of \( p_1 \).)

A strictly regenerative random permutation is called positive recurrent if 
\( ET_1 < \infty \). From standard renewal theory, it follows that

\[
\lim_{n \to \infty} u_n = \frac{1}{ET_1}.
\]

In [1] we showed that under the \( p \)-shifted probability measure \( P \), the random variables \( \{I_{<j}\}_{j=2}^{\infty} \) are independent, and that \( 1 + I_{<k} \) is distributed as \( \{p_j\}_{j=1}^{\infty} \), truncated at \( k \):

\[
P(I_{<k} = i) = \frac{p_{i+1}}{\sum_{j=1}^{k} p_j}, \text{ for } i = 0, \cdots, k - 1, \text{ and } k = 2, 3, \cdots.
\]

This offers an alternative way to construct a \( p \)-shifted random permutation in \( S_n \) or in \( S_\infty \). Let \( X \) be a random variable on \( \mathbb{Z}^+ \) whose distribution is characterized by \( 1 + X \) having the distribution \( \{p_j\}_{j=1}^{\infty} \); that is,

\[
P(X = j) = p_{j+1}, \text{ } j = 0, 1, \cdots.
\]

Let \( \{X_n\}_{n=2}^{\infty} \) be a sequence of independent random variables with the distribution of \( X_n \) being the distribution of \( X \) truncated at \( n - 1 \):

\[
P(X_n = i) = \frac{p_{i+1}}{\sum_{j=1}^{n} p_j}, \text{ } i = 0, 1, \cdots n - 1.
\]

To construct a \( p \)-shifted random permutation in \( S_n \), set the number 1 down on a horizontal line. Now inductively, if the numbers \( \{1, \cdots, j - 1\} \) have already been placed down on the line, where \( 2 \leq j \leq n \), then sample from \( X_j \) independently of everything that has already occurred, and place the number \( j \) on the line in the position for which there are \( X_j \) numbers to its right. Thus, for example, to create a \( p \)-shifted random permutation in \( S_4 \), if \( X_2, X_3, X_4 \) have been sampled independently as \( X_2 = 1 \), \( X_3 = 2 \) and \( X_4 = 0 \), then we obtain the permutation 3214. To obtain a \( p \)-shifted random permutation in \( S_\infty \), one just continues the above scenario indefinitely. Since

\[
EX = \sum_{j=1}^{\infty} P(X \geq j) = \sum_{j=1}^{\infty} \left( \sum_{i=j+1}^{\infty} p_i \right),
\]
it follows from (1.1) and (1.2) that the \( p \)-shifted random permutation is positive recurrent if and only if \( EX < \infty \), or equivalently, if and only if \( \sum_{n=1}^{\infty} np_n < \infty \).

Note that for the random permutation on \( S_n \) or \( S_{\infty} \) created in the previous paragraph, one has \( X_j = I_{<j} \) for all appropriate \( j \). The total number of inversions in a permutation \( \sigma \in S_n \) is given by \( I_n(\sigma) := \sum_{j=2}^{n} I_{<j}(\sigma) \). It follows from the construction in the above paragraph that the inversion statistic \( I_n \) satisfies the following weak law of large numbers as \( n \to \infty \):

\[
\mathbb{I}_n \text{ under } P_n \text{ converges in probability to } EX = \sum_{n=1}^{\infty} np_{n+1} \in (0, \infty].
\]

We now give our quantitative result relating the tendency toward inversions and the tendency toward clustering.

**Theorem 1.** Let \( A_{i,k}^{(n)} \subset S_n \) denote the event that the set of \( l \) consecutive numbers \( \{k, k+1, \ldots, k+l-1\} \) appear in a set of \( l \) consecutive positions. Let \( \{p_n\}_{n=1}^{\infty} \) be a probability distribution on \( \mathbb{N} \) with \( p_j > 0 \), for all \( j \in \mathbb{N} \). Also assume that the sequence \( \{p_n\}_{n=1}^{\infty} \) is non-increasing. Let \( \{P_n\}_{n=1}^{\infty} \) be the corresponding sequence of \( p \)-shifted distributions on the sequence of spaces \( \{S_n\}_{n=1}^{\infty} \). Then for all \( k \in \mathbb{N} \),

\[
\lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{i,k}^{(n)}) = (\prod_{j=1}^{k-1} \sum_{i=1}^{j} p_i)(\prod_{j=1}^{\infty} P(X \leq j-1)) = (\prod_{j=1}^{k-1} P(X \leq j-1)) (\prod_{j=1}^{\infty} P(X \leq j-1)).
\]

Also, if \( \lim_{n \to \infty} \min(k_n, n - k_n) = \infty \), then

\[
\lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{i,k,n}^{(n)}) = (\prod_{j=1}^{\infty} \sum_{i=1}^{j} p_i)^2 = (\prod_{j=1}^{\infty} P(X \leq j-1))^2.
\]

In particular, the limits in (1.6) and (1.7) are positive if and only if \( EX = \sum_{n=1}^{\infty} np_{n+1} < \infty \), or equivalently, if and only if the \( p \)-shifted random permutation is positive recurrent.

**Remark 1.** In light of (1.5), the theorem shows that super-clustering occurs if and only if the total inversion statistic \( I_n \) has linear rather than super-linear growth.
Remark 2. Note that \( \prod_{j=1}^{\infty} P(X \leq j - 1) \) and \( \prod_{j=1}^{k-1} P(X \leq j - 1) \) are decreasing with respect to stochastic dominance. Thus, if \( X^{(1)} \) and \( X^{(2)} \) satisfy (1.3) with \( p_{j+1} \) replaced respectively by \( p_{j+1}^{(1)} \) and \( p_{j+1}^{(2)} \), and if \( X^{(1)} \) stochastically dominates \( X^{(2)} \), that is, \( \sum_{j=n}^{\infty} p_j^{(1)} \geq \sum_{j=n}^{\infty} p_j^{(2)} \), for all \( n \in \mathbb{N} \), then the probability of super-clustering for the \( p^{(2)} \)-shifted random permutation is greater than for the \( p^{(1)} \)-shifted random permutation.

Remark 3. If one removes the requirement that the sequence \( \{p_j\}_{j=1}^{\infty} \) be non-increasing, then (1.6) and (1.7) hold with “\( \lim_{l \to \infty} \lim_{n \to \infty} \)” and “\( \lim_{l \to \infty} \lim_{n \to \infty} \)” replaced by “\( \lim_{l \to \infty} \liminf_{n \to \infty} \)” and “\( \lim_{l \to \infty} \liminf_{n \to \infty} \)”. This follows immediately from the proof of the theorem. Thus, for this more general case, the finiteness of \( EX \) is a sufficient condition for super-clustering.

Remark 4. The limits (1.6) and (1.7) were obtained in [1] for the case of the Mallows distribution; namely, \( p_j = \frac{1}{q^j} \). However they were expressed in a form that did not reveal the connection with the random variable \( X \).

Open Problem. It would be interesting to investigate the behavior of the statistic \( P_n(A_{l,k}^{(n)}) \) as \( n \to \infty \) and then \( l \to \infty \) in the case that the \( \{P_n\}_{n=1}^{\infty} \) are the so-called \( p \)-biased distributions; see [3] or [2] for the definition. Under these distributions, which are defined in a somewhat similar way to the \( p \)-shifted distributions, the backwards ranks \( I_{<j} \) are not independent.

We will also prove the following characterization of the class of positive recurrent \( p \)-shifted distributions, which might be of some independent interest. Recall that positive recurrence is equivalent to \( \sum_{n=1}^{\infty} np_n < \infty \).

Proposition 1. The class of positive recurrent \( p \)-shifted distributions, as \( p \) runs over all probability distributions \( \{p_j\}_{j=1}^{\infty} \) whose supports are all of \( \mathbb{N} \), and that satisfy \( \sum_{n=1}^{\infty} np_n < \infty \), coincides with the class of probability distributions \( P \) on \( S_{\infty} \) that satisfy the following three conditions:

i. The random variables \( \{I_{<j}\}_{j=2}^{\infty} \) are independent under \( P \);

ii. A random permutation under \( P \) is strictly regenerative with a positive recurrent renewal mechanism, and the probability \( u_1 \) of renewal at the number 1 is positive;

iii. For all \( n \in \mathbb{N} \), the support of \( P_{\text{proj}_n} \) is all of \( S_n \).
Remark. The proof of the proposition also shows that if one removes the requirement that the support of the distribution \( p \) is all of \( \mathbb{N} \), and only requires that \( p_1 > 0 \) (which in any case is necessary in order to implement the \( p \)-shifted construction), then the proposition holds with property (iii) deleted.

We prove Theorem 1 in section 2 and Proposition 1 in section 3.

2. Proof of Theorem 1

We note that the final statement of the theorem is almost immediate. Indeed, \( \prod_{j=1}^{\infty} P(X \leq j-1) = \prod_{j=1}^{\infty} (1-P(X \geq j)) \), and this infinite product is nonzero if and only if \( \sum_{j=1}^{\infty} P(X \geq j) < \infty \). However, \( \sum_{j=1}^{\infty} P(X \geq j) = E X \).

We now turn to the proofs of (1.6) and (1.7). We use the second method offered in this paper for constructing the \( p \)-shifted random permutation, as described after (1.3). Thus, we consider a sequence of independent random variables \( \{X_n\}_{n=2}^{\infty} \), with \( X_n \) distributed as in (1.4). For the proof, we will use the notation

\[
N_n = \sum_{i=1}^{n} p_i = P(X \leq n), \quad n \in \mathbb{N}, \text{ and } N_0 = 0,
\]

where \( X \) is as in (1.3). Note that \( N_n \) is the normalization constant on the right hand side of (1.4). Although \( P_n \) denotes the \( p \)-shifted probability measure on \( S_n \), we will also use this notation for probability when discussing events related to the random variables \( \{X_j\}_{j=2}^{n} \).

We begin with the proof of (1.6). Fix \( k \in \mathbb{N} \). Consider the event, which we denote by \( B_{l;k} \), that after the first \( k+l-1 \) positive integers have been placed down on the horizontal line, the set of \( l \) numbers \( \{k, k+1, \ldots, k+l-1\} \) appear in a set of \( l \) consecutive positions. Then \( B_{l;k} = \bigcup_{a=0}^{k-1} B_{l;k;a} \), where the events \( \{B_{l;k;a}\}_{a=0}^{k-1} \) are disjoint, with \( B_{l;k;a} \) being the event that the set of \( l \) numbers \( \{k, k+1, \ldots, k+l-1\} \) appear in a set of \( l \) consecutive positions and also that exactly \( a \) of the numbers in \( [k-1] \) are to the right of this set. We calculate \( P_n(B_{l;k;a}) \).

Suppose that we have already placed down on the horizontal line the numbers in \( [k-1] \). Their relative positions are irrelevant for our considerations.
Now we use $X_k$ to insert on the line the number $k$. Suppose that $X_k = a$, $a \in \{0, \cdots, k-1\}$. Then the number $k$ is inserted on the line in the position for which $a$ of the numbers in $[k-1]$ are to its right. Now in order for $k+1$ to be placed in a position adjacent to $k$, we need $X_{k+1} \in \{a, a+1\}$. (If $X_{k+1} = a$, then $k+1$ will appear directly to the right of $k$, while if $X_{k+1} = a+1$, then $k+1$ will appear directly to the left of $k$.) If this occurs, then $\{k, k+1\}$ are adjacent, and $a$ of the numbers in $[k-1]$ are to the right of $\{k, k+1\}$. Continuing in this vein, for $i \in \{1, \cdots, l-2\}$, given that the numbers $\{k, \cdots, k+i\}$ are adjacent to one another, and $a$ of the numbers in $[k-1]$ appear to the right of $\{k, \cdots, k+i\}$, then in order for $k+i+1$ to be placed so that $\{k, \cdots, k+i+1\}$ are all adjacent to one another (with $a$ of the numbers in $[k-1]$ appearing to the right of these numbers), we need $X_{k+i+1} \in \{a, \cdots, a+i+1\}$. We conclude then that 

$$P_n(B_{l;k;a}) = \prod_{j=0}^{l-1} P_n(X_{k+j} \in \{a, \cdots, a+j\}).$$

Using (1.4), we have

$$(2.1) \quad P_n(B_{l;k;a}) = \prod_{j=0}^{l-1} P_n(X_{k+j} \in \{a, \cdots, a+j\}) = \prod_{j=0}^{l-1} \frac{N_{a+j+1} - N_a}{N_{k+j}}.$$ 

We now consider the conditional probability, $P_n(A_{l;k}^{(n)} | B_{l;k;a})$, that is, the probability, given that $B_{l;k;a}$ has occurred, that the numbers $k+1, \cdots, n$ are inserted in such a way so as to preserve the mutual adjacency of the set $\{k, \cdots, k+l-1\}$. We will obtain lower and upper bounds on this conditional probability. However, first we note that it is clear from the construction that $P_n(A_{l;k}^{(n)} | B_{l;k;a})$ is decreasing in $n$. Thus, since $P_n(B_{l;k;a})$ is independent of $n$, it follows that $P_n(A_{l;k}^{(n)})$ is decreasing in $n$. Consequently $\lim_{n \to \infty} P_n(A_{l;k}^{(n)})$ exists.

We now turn to a lower bound on $P_n(A_{l;k}^{(n)} | B_{l;k;a})$. Our lower bound will be the probability of the event that all of the remaining numbers are inserted to the right of the set $\{k, \cdots, k+l-1\}$. This event is given by

$$\cap_{j=0}^{n-k-l} \{X_{k+l+j} \leq a+j\}.$$ 

Thus, we have

$$(2.2) \quad P_n(A_{l;k}^{(n)} | B_{l;k;a}) \geq P(\cap_{j=0}^{n-k-l} \{X_{k+l+j} \leq a+j\}) = \prod_{j=0}^{n-k-l} \frac{N_{a+j+1}}{N_{k+l+j}}.$$
Writing \( P_n(A_{t;k}^{(n)}) = \sum_{a=0}^{k-1} P_n(B_{t;k;a})P(A_{t;k}^{(n)} | B_{t;k;a}) \), (2.1) and (2.2) yield

\[
(2.3) \quad P_n(A_{t;k}^{(n)}) \geq \sum_{a=0}^{k-1} \left( \prod_{j=0}^{l-1} n_{a+j+1} - n_a \right) \left( \prod_{j=0}^{n-k-l} \frac{n_{a+j+1}}{n_{k+l+j}} \right).
\]

We have \( \prod_{j=0}^{n-k-l} \frac{n_{a+j+1}}{n_{k+l+j}} = \frac{n_{a+1} \cdots n_{k+l}}{n_{n-a-l+1} \cdots n_n} \). Using this along with the fact that \( \lim_{n \to \infty} N_n = 1 \) and the fact that the limit on the left hand side of (2.3) exists, we have

\[
(2.4) \quad \lim_{n \to \infty} P_n(A_{t;k}^{(n)}) \geq \sum_{a=0}^{k-1} \left( \prod_{j=0}^{l-1} \left( n_{a+j+1} - n_a \right) \right) \left( \prod_{i=a+1}^{k-l} n_i \right).
\]

We now let \( l \to \infty \) in (2.4). We only consider the term in the summation with \( a = 0 \), because it turns out that the terms with \( a \geq 1 \) converge to 0 as \( l \to \infty \). We obtain

\[
(2.5) \quad \lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{t;k}^{(n)}) \geq \left( \prod_{j=1}^{k-1} N_j \right) \left( \prod_{j=1}^{\infty} N_j \right) = \left( \prod_{j=1}^{k-1} P(X \leq j-1) \right) \left( \prod_{j=1}^{\infty} P(X \leq j-1) \right).
\]

Note that by the assumption that \( \{p_n\}_{n=1}^{\infty} \) is non-increasing, it follows that \( P(X \not\in \{j+1, \ldots, j+l-1\}) \) is increasing in \( j \). Also, note that \( P(X \not\in \{j+1, \ldots, j+l-1\}) > P(X_m \not\in \{j+1, \ldots, j+l-1\}) \) for \( j+l \leq m \). These facts will be used as we turn now to an upper bound on \( P_n(A_{t;k}^{(n)} | B_{t;k;a}) \), the conditional probability given \( B_{t;k;a} \) that the numbers \( k+l, \ldots, n \) are inserted in such a way so as to preserve the mutual adjacency of the set \( \{k, \ldots, k+l-1\} \). First the number \( k+l \) is inserted. The probability that its insertion preserves the mutual adjacency property of the set \( \{k, \ldots, k+l-1\} \) is \( P(X_{k+l} \not\in \{a+1, \ldots, a+l-1\}) \), which is less than \( P(X \not\in \{a+1, \ldots, a+l-1\}) \). If the insertion of \( k+l \) preserves the mutual adjacency, then either \( X_{k+l} \in \{0, \ldots, a\} \) or \( X_{k+l} \in \{a+l, \ldots, k+l-1\} \). If \( X_{k+l} \in \{0, \ldots, a\} \), then in order for the mutually adjacency to be preserved when the number \( k+l+1 \) is inserted, one needs \( P(X_{k+l+1} \not\in \{a+2, \ldots, a+l\}) \), while if \( X_{k+l} \in \{a+l, \ldots, k+l-1\} \), then one needs \( P(X_{k+l+1} \not\in \{a+1, \ldots, a+l-1\}) \). Either of these probabilities is less than \( P(X \not\in \{a+2, \ldots, a+l\}) \). Thus, an upper bound for the conditional.
probability given $B_{t;k:a}$ that the insertion of $k + l$ and $k + l + 1$ preserves the mutual adjacency is $P(X \not\in \{a + 1, \cdots, a + l - 1\})P(X \not\in \{a + 2, \cdots, a + l\})$.

Continuing in this vein, we conclude that

$$
P_n(A_{l;k}^{(n)} | B_{t;k:a}) \leq \prod_{j=1}^{n-k-l+1} P_n(X \not\in \{a, \cdots, a + j - 1\}) = \prod_{j=1}^{n-k-l+1} (1 - N_{a+j+l-1} + N_{a+j}).
$$

(2.6)

Using this upper bound, we obtain an upper bound on $P_n(A_{l;k}^{(n)})$. From (2.1) and (2.6), we have

$$
P_n(A_{l;k}^{(n)}) \leq \sum_{a=0}^{k-1} \left( \prod_{j=0}^{l-1} \frac{N_{a+j+1} - N_a}{N_{k+j}} \right) \left( \prod_{j=1}^{n-k-l+1} (1 - N_{a+j+l-1} + N_{a+j}) \right).
$$

(2.7)

Letting $n \to \infty$ and using the fact that the limit on the left hand side exists, we have

$$
\lim_{n \to \infty} P_n(A_{l;k}^{(n)}) \leq \sum_{a=0}^{k-1} \left( \prod_{j=0}^{l-1} \frac{N_{a+j+1} - N_a}{N_{k+j}} \right) \left( \prod_{j=1}^{\infty} (1 - N_{a+j+l-1} + N_{a+j}) \right).
$$

(2.8)

For $a \in \{1, \cdots, k - 1\}$, we have $\frac{N_{a+j+1} - N_a}{N_{k+j}} < 1 - N_a \in (0, 1)$. Therefore, when letting $l \to \infty$ in (2.8), a contribution will come from the right hand side only when $a = 0$. We obtain

$$
\lim_{l \to \infty} \lim_{n \to \infty} P_n(A_{l;k}^{(n)}) \leq \lim_{l \to \infty} \left( \prod_{j=0}^{l-1} \frac{N_{j+1}}{N_{k+j}} \right) \left( \prod_{j=1}^{\infty} (1 - N_{j+l-1} + N_j) \right) = \left( \prod_{j=1}^{k-1} N_j \right) \left( \prod_{j=1}^{\infty} N_j \right) = \left( \prod_{j=1}^{k-1} P(X \leq j - 1) \right) \left( \prod_{j=1}^{\infty} P(X \leq j - 1) \right).
$$

(2.9)

Now (1.6) follows from (2.5) and (2.9).

We now turn to the proof of (1.7). As with the proof of (1.6), the term with $a = 0$ will dominate. Thus, for the lower bound, using (2.1) and (2.2) with $k = k_n$ and ignoring the terms with $a \geq 1$, we have

$$
P_n(A_{l;k_n}^{(n)}) \geq \left( \prod_{j=0}^{l-1} \frac{N_{j+1}}{N_{k_n+j}} \right) \left( \prod_{j=0}^{n-k_n-l} \frac{N_{j+1}}{N_{k_n+j+l}} \right).
$$

(2.10)
Letting \( n \to \infty \) in (2.10) and using the assumption that \( \lim_{n \to \infty} \min(k_n, n - k_n) = \infty \), it follows that

\[
\liminf_{n \to \infty} P_n(A_{l; k_n}^{(n)}) \geq \left( \prod_{j=1}^{l} N_j \right) \left( \prod_{j=1}^{\infty} N_j \right).
\]

Now letting \( l \to \infty \) gives

\[
(2.11) \quad \lim_{l \to \infty} \liminf_{n \to \infty} P_n(A_{l; k_n}^{(n)}) \geq \left( \prod_{j=1}^{\infty} N_j \right)^2 = \left( \prod_{j=1}^{\infty} P(X \leq j - 1) \right)^2.
\]

For the upper bound, let \( k = k_n \) in (2.7). The second factor in the summand \( \left( \prod_{j=1}^{l-1} N_{a+j+1} - N_a \right) \left( \prod_{j=1}^{k_n-l+1} (1 - N_{a+j+l-1} + N_{a+j}) \right) \) is less than 1, while the first factor in the summand satisfies

\[
\prod_{j=0}^{l-1} \frac{N_{a+j+1} - N_a}{N_{k_n+j}} \leq \frac{N_{a+1} - N_a}{N_{k_n}} \leq \frac{p_{a+1}}{p_1},
\]

for \( a \in \{0, \cdots, k_n - 1\} \) and \( n \geq 1 \). Since \( \sum_{a=0}^{\infty} \frac{p_{a+1}}{p_1} < \infty \), the dominated convergence theorem and the assumption that \( \lim_{n \to \infty} \min(k_n, n - k_n) = \infty \) allow us to conclude upon letting \( n \to \infty \) in (2.7) with \( k = k_n \) that

\[
(2.12) \quad \limsup_{n \to \infty} P_n(A_{l; k_n}^{(n)}) \leq \sum_{a=0}^{\infty} \left( \prod_{j=0}^{l-1} (N_{a+j+1} - N_a) \right) \left( \prod_{j=1}^{\infty} (1 - N_{a+j+l-1} + N_{a+j}) \right).
\]

For \( a \geq 1 \), we have \( N_{a+j+1} - N_a \in (0, 1 - p_1) \). Consequently, when letting \( l \to \infty \) in (2.12), a contribution will come from the right hand side only when \( a = 0 \). We obtain

\[
(2.13) \quad \lim_{l \to \infty} \limsup_{n \to \infty} P_n(A_{l; k_n}^{(n)}) \leq \left( \prod_{j=1}^{\infty} N_j \right)^2 = \left( \prod_{j=1}^{\infty} P(X \leq j - 1) \right)^2.
\]

Now (1.7) follows from (2.11) and (2.13).

3. Proof of Proposition 1

It has already been noted that a \( p \)-shifted random permutation with \( p_1 > 0 \) and \( \sum_{n=1}^{\infty} np_n < \infty \) satisfies properties (i) and (ii) of the proposition. From the construction, it is clear that it also satisfies property (iii), if the support of the distribution \( p \) is all of \( \mathbb{N} \). Thus, we only need prove that if a probability distribution \( P \) on \( S_\infty \) satisfies the three properties stated in the
proposition, then it arises as a \( p \)-shifted permutation for some distribution \( \{p_j\}_{j=1}^{\infty} \) whose support is all of \( \mathbb{N} \) and that satisfies \( \sum_{n=1}^{\infty} np_n < \infty \).

Let \( \Pi \) denote the random permutation under \( P \). By property (ii), \( \Pi \) is strictly regenerative and the probability \( u_1 \) of renewal at the number 1 is positive. (From this it follows that the probability \( u_n \) of renewal at the number \( n \) is positive, for all \( n \). However, for this proof, we only need the fact that \( u_1 > 0 \).) The event that \( n \) is a renewal point, that is, the event \( \Pi([n]) = [n] \), can be written as \( \bigcap_{j=1}^{\infty} \{I_{<n+j} \leq j-1\} \). Thus, we have \( u_n = P(\bigcap_{j=1}^{\infty} \{I_{<n+j} \leq j-1\}) > 0 \). By property (i), this can be rewritten as

\[
(3.1) \quad u_n = \prod_{j=1}^{\infty} P(I_{<n+j} \leq j-1) = \prod_{j=1}^{\infty} (1 - P(I_{<n+j} \geq j)).
\]

Recall that the renewal times are labelled as \( \{T_n\}_{n=1}^{\infty} \). If \( n \) is a renewal point, say \( T_{k_0} = n \), then in order that the reduced permutation \( \text{red}(\Pi|_{[T_{k_0}+1]-[T_{k_0}]} \) have the same distribution as \( \Pi|_{[T_1]} \), we need

\[
(3.2) \quad \text{dist}(\{I_{<n+j}\}_{j=1}^{\infty}|\bigcap_{j=1}^{\infty} \{I_{<n+j} \leq j-1\}) = \text{dist}(\{I_{<j}\}_{j=1}^{\infty}).
\]

By property (i), the above reduces to

\[
(3.3) \quad \text{dist}(I_{<n+j}|I_{<n+j} \leq j-1) = \text{dist}(I_{<j}), \quad \text{for } j = 2, 3, \cdots \text{ and } n = 1, 2, \cdots.
\]

Now (3.3) for any particular \( n \) was obtained under the assumption that \( u_n > 0 \). By property (ii), we have \( u_1 > 0 \). Thus, (3.3) holds for \( n = 1 \). From this it follows that there exist nonnegative \( \{q_j\}_{j=1}^{\infty} \) with \( q_1 > 0 \) such that

\[
(3.4) \quad P(I_{<j} = i) = \frac{q_{i+1}}{\sum_{k=1}^{j} q_k}, \quad i = 0, 1, \cdots j-1 \text{ and } j = 2, 3, \cdots.
\]

We now show that \( \sum_{j=1}^{\infty} q_j < \infty \). Assume to the contrary. Then from (3.4) it follows that \( I_{<j} \) converges in probability to \( \infty \) as \( j \to \infty \). Thus \( \lim_{n \to \infty} P(I_{<n+j} \geq j) = 1 \), for all \( j \) and consequently

\[
\lim_{n \to \infty} \sum_{j=1}^{\infty} P(I_{<n+j} \geq j) = \infty.
\]

From this and (3.1), it follows that \( \lim_{n \to \infty} u_n = 0 \), which contradicts the assumption that the strictly regenerative random permutation is positive recurrent.
Since $\sum_{j=1}^{\infty} q_j < \infty$, without loss of generality we may assume that
$\sum_{j=1}^{\infty} q_j = 1$. Thus, from (3.4), we conclude that $P(I_{<j} = i) = q_{i+1}$, for
$i = 0, 1, \cdots, j - 1$ and $j = 2, 3, \cdots$. From this it follows that the measure $P$
is the $p$-shifted measure with $p$ distribution given by $\{q_j\}_{j=1}^{\infty}$. In order for
property (iii) to hold, it is necessary that $q_j > 0$, for all $j$. \qed

\section*{References}

[1] Pinsky, R. G., \textit{Clustering of consecutive numbers in permutations under a Mallows
distribution}, preprint.

[2] Pinsky, R. G., \textit{Comparing the inversion statistic for distribution-biased and
distribution-shifted permutations with the geometric and the GEM distributions},
preprint.

[3] Pitman, J. and Tang, W., \textit{Regenerative random permutations of integers}, Ann.
Probab., 47 (2019), 1378-1416.

Department of Mathematics, Technion—Israel Institute of Technology,
Haifa, 32000, Israel

Email address: pinsky@math.technion.ac.il

URL: http://www.math.technion.ac.il/~pinsky/