On the Asymptotic Statistics of the Number of Occurrences of Multiple Permutation Patterns

Svante Janson[1], Brian Nakamura[2], and Doron Zeilberger[3]

December 13, 2013

Abstract

We study statistical properties of the random variables $X_\sigma(\pi)$, the number of occurrences of the pattern σ in the permutation π. We present two contrasting approaches to this problem: traditional probability theory and the "less traditional" computational approach. Through the perspective of the first one, we prove that for any pair of patterns σ and τ, the random variables X_σ and X_τ are jointly asymptotically normal (when the permutation is chosen from S_n). From the other perspective, we develop algorithms that can show asymptotic normality and joint asymptotic normality (up to a point) and derive explicit formulas for quite a few moments and mixed moments empirically, yet rigorously. The computational approach can also be extended to the case where permutations are drawn from a set of pattern avoiders to produce many empirical moments and mixed moments. This data suggests that some random variables are not asymptotically normal in this setting.

1 Introduction

The primary area of interest in this article is the study of patterns in permutations. We will denote the set of length n permutations by $\mathcal{S}_n$. Let $a_1 a_2 \dots a_k$ be a sequence of k distinct real numbers. The *reduction* of this sequence, which is denoted by $\mathrm{red}(a_1 \dots a_k)$, is the length k permutation $\pi_1 \dots \pi_k \in \mathcal{S}_k$ such that order-relations are preserved (i.e., $\pi_i < \pi_j$ if and only if $a_i < a_j$ for every i and j). Given a (permutation) pattern $\tau \in \mathcal{S}_k$, we say that a permutation $\pi = \pi_1 \dots \pi_n \in \mathcal{S}_n$ *contains* the pattern τ if there exists $1 \le i_1 < i_2 < \dots < i_k \le n$ such that $\mathrm{red}(\pi_{i_1}\pi_{i_2}\dots\pi_{i_k}) = \tau$. Each such subsequence in π will be called an *occurrence* of the pattern τ. If π contains no such subsequence, it is said to *avoid* the pattern τ. Additionally, we will denote the number of occurrences of the pattern τ in permutation π by $N_\tau(\pi)$ (e.g., π avoids the pattern τ if and only if $N_\tau(\pi) = 0$).

For any pattern τ and integer $n \ge 0$, we define the set

$$\mathcal{S}_n(\tau) := \{\pi \in \mathcal{S}_n \ : \ \pi \text{ avoids the pattern } \tau\} \tag{1}$$

and also define $s_n(\tau) := |\mathcal{S}_n(\tau)|$. The patterns σ and τ are said to be *Wilf-equivalent* if $s_n(\sigma) = s_n(\tau)$ for all $n \ge 0$. We may also consider the more general set

$$\mathcal{S}_n(\tau, r) := \{\pi \in \mathcal{S}_n \ : \ \pi \text{ contains exactly } r \text{ occurrences of } \tau\}. \tag{2}$$

[1]Department of Mathematics, Uppsala University, Uppsala, Sweden. [svante.janson@math.uu.se]
[2]CCICADA/DIMACS, Rutgers University-New Brunswick, Piscataway, NJ, USA. [brian.nakamura@rutgers.edu]
[3]Mathematics Department, Rutgers University-New Brunswick, Piscataway, NJ, USA. [zeilberg@math.rutgers.edu]

We will analogously define $s_n(\tau, r) := |\mathcal{S}_n(\tau, r)|$.

A classical problem in this area is to find an enumeration for these sets or at the least, to study properties of the generating function encoding the enumerating sequence (for example, is it rational/algebraic/holonomic?). However, it is not even known if these generating functions are always holonomic. In general, the enumeration problem gets very difficult very quickly. Patterns up to length 3 are well-understood, but there are basic unresolved questions even for length 4 patterns. For example, it is known that there are three Wilf-equivalence classes for length 4 patterns: 1234, 1324, and 1342. While the enumeration problems have been solved for 1234 and 1342, no exact enumeration (or even asymptotics) is known for 1324.

A (probabilistic) variation of this problem was posed by Joshua Cooper [6]: Given two (permutation) patterns σ and τ, what is the expected number of copies of σ in a permutation chosen uniformly at random from $\mathcal{S}_n(\tau)$? We note that if the enumeration of $\mathcal{S}_n(\tau)$ is known, this question is equivalent to counting the total number of occurrences of σ in permutations from $\mathcal{S}_n(\tau)$, or put more precisely, to compute

$$T_n(\sigma, \tau) := \sum_{\pi \in \mathcal{S}_n(\tau)} N_\sigma(\pi). \tag{3}$$

Bóna first addressed the question for $\tau = 132$ when σ is either the increasing or decreasing permutation in [2]. He shows how to derive the generating functions for $T_n(12\dots k, 132)$ and $T_n(k\dots 21, 132)$, the total number of occurrences of $12\dots k$ in $\mathcal{S}_n(\tau)$ and occurrences of $k\dots 21$ in $\mathcal{S}_n(\tau)$, respectively. In [4], Bóna also shows that $T_n(213, 132) = T_n(231, 132) = T_n(312, 132)$ for all n and provides an explicit formula for them. Rudolph [13] also proves some conditions on when two patterns, say p and q, occur equally frequently in $\mathcal{S}_n(132)$ (i.e., $T_n(p, 132) = T_n(q, 132)$ for all n).

In [9], Homberger answers the analogous question when $\tau = 123$ and shows that there are three non-trivial cases to consider: $T_n(132, 123)$, $T_n(231, 123)$, and $T_n(321, 123)$. He finds generating functions and explicit formulas for each one.

We will consider a more general problem. Given the pattern τ, suppose that a permutation π is chosen uniformly at random from $\mathcal{S}_n(\tau)$. Given another pattern σ, we define the random variable $X_\sigma(\pi) := N_\sigma(\pi)$, the number of copies of σ in π. Observe that $T_n(\sigma, \tau) = \mathbb{E}[X_\sigma]$, the expected value of X_σ (i.e., the first moment of the random variable). The focus of this paper is to study higher moments for X_σ as well as mixed moments between two such random variables that count different patterns. We will consider the case where the permutation π is randomly chosen from $\mathcal{S}_n$ as well as some cases where π is chosen from $\mathcal{S}_n(\tau)$ (for various patterns τ).

In this paper, we approach the problem from two different angles. On one end, we will present (human-derived) results proving that the random variables are jointly asymptotically normal when the permutations are chosen at random from $\mathcal{S}_n$. Unfortunately, the techniques do not naturally extend to the scenario when the permutations are chosen from $\mathcal{S}_n(\tau)$. On the other end, we present a computational approach that can quickly and easily compute many empirical moments for the general case (permutations chosen from $\mathcal{S}_n(\tau)$). In addition, for the case where permutations are chosen from $\mathcal{S}_n$, the computational approach can rigorously produce closed-form formulas for quite a few moments and mixed moments of the random variables.

This paper is organized as follows. In Section 2, we review and outline the functional equations enumeration approach developed in [10, 11]. In Section 3, we derive both rigorous results and empirical values for higher order moments and mixed moments for various random variables X_σ. In Section 4, we show that the random variables are jointly asymptotically normal when the permutations are randomly chosen from $\mathcal{S}_n$. In Section 5, we conclude with some final remarks and observations.

2 Enumerating with functional equations

For various patterns τ, functional equations were derived for enumerating permutations with r occurrences of τ in [10, 11, 12]. These functional equations were then used to derive enumeration algorithms. We briefly review the relevant results here. The curious reader can see [10, 11, 12] for more details.

2.1 Functional equations for single patterns

Given a (fixed) pattern τ and non-negative integer n, we define the polynomial:

$$f_n(\tau;\ t) := \sum_{\pi \in \mathcal{S}_n} t^{N_\tau(\pi)}. \tag{4}$$

Recall that the coefficient of t^r is exactly $s_n(\tau, r)$. For certain patterns τ, a multi-variate polynomial $P_n(\tau;\ t;\ x_1, \ldots, x_n)$ was defined so that $P_n(\tau;\ t;\ 1, \ldots, 1) = f_n(\tau;\ t)$ and that functional equations could be derived for the P_n polynomial.

The pattern $\tau = 123$ was considered in [11, 12], and the polynomial P_n was defined to be:

$$P_n(123;\ t;\ x_1, \ldots, x_n) := \sum_{\pi \in \mathcal{S}_n} \left(t^{N_{123}(\pi)} \prod_{i=1}^{n} x_i^{|\{(a,b)\,:\,\pi_a = i < \pi_b,\ 1 \le a < b \le n\}|} \right). \tag{5}$$

It was shown that this P_n satisfies the functional equation:

Theorem 1. For the pattern $\tau = 123$,

$$P_n(123;\ t;\ x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^{n-i} \cdot P_{n-1}(123;\ t;\ x_1, \ldots, x_{i-1}, tx_{i+1}, \ldots, tx_n). \tag{FE123}$$

Since $P_1(123;\ t;\ x_1) = 1$, the functional equation can be used to recursively compute our desired quantity $P_n(123;\ t;\ 1, \ldots, 1) = f_n(123;\ t)$.

Similarly, in [10], the polynomial P_n was defined for the pattern $\tau = 132$ so that it satisfied the functional equation:

Theorem 2. For the pattern $\tau = 132$,

$$P_n(132;\ t;\ x_1, \ldots, x_n) = \sum_{i=1}^{n} x_1 x_2 \ldots x_{i-1} \cdot P_{n-1}(132;\ t;\ x_1, \ldots, x_{i-1}, tx_{i+1}, \ldots, tx_n). \tag{FE132}$$

Again $P_1(132;\ t;\ x_1) = 1$, so the functional equation can be used to recursively compute our desired quantity $P_n(132;\ t;\ 1, \ldots, 1) = f_n(132;\ t)$.

The same was also done for the pattern $\tau = 231$ in [10]. Although $f_n(231;\ t) = f_n(132;\ t)$, redeveloping the approach directly for the pattern 231 allows us to consider the patterns 132 and 231 simultaneously. For 231, the polynomial P_n was defined so that it satisfies the functional equation:

Theorem 3. For the pattern $\tau = 231$,

$$P_n(231;\ t;\ x_1, \ldots, x_n) = \sum_{i=1}^{n} x_1^0 x_2^1 \ldots x_i^{i-1} \cdot P_{n-1}(231;\ t;\ x_1, \ldots, x_{i-1}, tx_i x_{i+1}, x_{i+2}, \ldots, x_n). \tag{FE231}$$

We again have that $P_1(231;\ t;\ x_1) = 1$, so the functional equation can be used to recursively compute our desired quantity $P_n(231;\ t;\ 1, \ldots, 1) = f_n(231;\ t)$.

The approach for the pattern 123 was also extended to the pattern $\tau = 1234$ in [11]. The polynomial $P_n(1234;\ t;\ x_1, \ldots, x_n;\ y_1, \ldots, y_n)$ was defined so that $P_n(1234;\ t;\ 1 \text{ [n times]};\ 1 \text{ [n times]}) = f_n(1234;\ t)$ and in such a way that it satisfies the functional equation:

Theorem 4. For the pattern $\tau = 1234$,

$$P_n(1234;\ t;\ x_1, \ldots, x_n;\ y_1, \ldots, y_n) =$$

$$\sum_{i=1}^{n} y_i^{n-i} \cdot P_{n-1}(1234;\ t;\ x_1, \ldots, x_{i-1}, tx_{i+1}, \ldots, tx_n;\ y_1, \ldots, y_{i-1}, x_i y_{i+1}, \ldots, x_i y_n). \quad \text{(FE1234)}$$

Since $P_1(1234;\ t;\ x_1;\ y_1) = 1$, the functional equation can be used to recursively compute our desired quantity $P_n(1234;\ t;\ 1 \text{ [n times]};\ 1 \text{ [n times]}) = f_n(1234;\ t)$.

2.2 Merging functional equations for multiple patterns

It is also straight-forward to consider multiple patterns simultaneously if their corresponding functional equations are known, as shown in [10]. For example, suppose that we want to consider the two patterns $\sigma = 123$ and $\tau = 132$ simultaneously. We can extend the f_n polynomial in the natural way to:

$$f_n(\sigma, \tau;\ s, t) := \sum_{\pi \in \mathcal{S}_n} s^{N_\sigma(\pi)} t^{N_\tau(\pi)}. \quad (6)$$

In [10], the polynomial $P_n(123, 132;\ s, t;\ x_1, \ldots, x_n;\ y_1, \ldots, y_n)$ was defined so that

$$P_n(123, 132;\ s, t;\ 1 \text{ [n times]};\ 1 \text{ [n times]}) = f_n(123, 132;\ s, t). \quad (7)$$

The following functional equation was then derived:

Theorem 5. For the patterns $\sigma = 123$ and $\tau = 132$,

$$P_n(123, 132;\ s, t;\ x_1, \ldots, x_n;\ y_1, \ldots, y_n) =$$

$$\sum_{i=1}^{n} x_i^{n-i} \cdot y_1 y_2 \ldots y_{i-1} \cdot P_{n-1}(123, 132;\ s, t;\ x_1, \ldots, x_{i-1}, sx_{i+1}, \ldots, sx_n;\ y_1, \ldots, y_{i-1}, ty_{i+1}, \ldots, ty_n).$$

Observe that we combined the functional equations for the individual patterns 123 and 132 by re-labeling the x_i variables for 132 to y_i, merging the reductions in the P_{n-1} in the natural way, and multiplying the coefficient terms for the P_{n-1} within the summands. We again have that $P_1(123, 132;\ s, t;\ x_1;\ y_1) = 1$, so the functional equation can be used to recursively compute our desired quantity $P_n(123, 132;\ s, t;\ 1 \text{ [n times]};\ 1 \text{ [n times]}) = f_n(123, 132;\ s, t)$.

More generally, we can similarly extend $f_n(\tau;\ t)$ to k different patterns $\tau_1, \tau_2, \ldots, \tau_k$ and the corresponding variables $t_1, t_2, \ldots, t_k$ as:

$$f_n(\tau_1, \tau_2, \ldots, \tau_k;\ t_1, t_2, \ldots, t_k) := \sum_{\pi \in \mathcal{S}_n} t_1^{N_{\tau_1}(\pi)} t_2^{N_{\tau_2}(\pi)} \ldots t_k^{N_{\tau_k}(\pi)}. \quad (8)$$

The generalized polynomials P_n can be similarly defined and analogous functional equations can be derived.

For example, suppose that we want to consider all length three patterns simultaneously. We will consider the patterns in lexicographical order (i.e., $\tau_1 = 123,\ \tau_2 = 132,\ \ldots,\ \tau_6 = 321$). Our f_n polynomial now becomes:

$$f_n(123, 132, \ldots, 321;\ t_1, t_2, \ldots, t_6) := \sum_{\pi \in \mathcal{S}_n} t_1^{N_{123}(\pi)} t_2^{N_{132}(\pi)} \ldots t_6^{N_{321}(\pi)}. \quad (9)$$

For notational convenience, the polynomial $f_n(123, 132, \ldots, 321;\ t_1, t_2, \ldots, t_6)$ will be denoted by $f_n(\mathcal{S}_3;\ t_1, \ldots, t_6)$. In [10], we discuss how to extend this to the generalized polynomial P_n and derive analogous functional equations.

The previous polynomial could also be refined further to consider all length three patterns and the pattern 1234 simultaneously. We will again consider the length three patterns in lexicographical order. Our f_n polynomial now becomes:

$$f_n(1234, \mathcal{S}_3;\ s, t_1, t_2, \ldots, t_6) := \sum_{\pi \in \mathcal{S}_n} s^{N_{1234}(\pi)} t_1^{N_{123}(\pi)} t_2^{N_{132}(\pi)} \ldots t_6^{N_{321}(\pi)}. \tag{10}$$

Just like the previous case, this polynomial can be extended to the analogous generalized polynomial P_n and similar functional equations can be derived.

2.3 Adapting multi-pattern functional equations

The previously described f_n polynomials (and their corresponding generalized P_n polynomials and functional equations) can be easily specialized to consider a variety of scenarios. This allows us to quickly extract functional equations (and fast enumeration algorithms) in a number of cases.

The polynomial $f_n(\mathcal{S}_3;\ t_1, \ldots, t_6)$ (in Eq. 9) can be specialized to consider any subset of $\mathcal{S}_3$ by setting some t_i variables to 1. For example, $f_n(\mathcal{S}_3;\ t_1, t_2, 1, 1, 1, 1)$ would give us the polynomial tracking 123 and 132 simultaneously. Setting $t_i = 1$ for $3 \leq i \leq 6$ in the generalized polynomial P_n and its functional equation would reproduce Theorem 5. This approach actually allows us to quickly compute the bi-variate polynomial

$$f_n(\sigma, \tau;\ s, t) = \sum_{\pi \in \mathcal{S}_n} s^{N_\sigma(\pi)} t^{N_\tau(\pi)} \tag{11}$$

for any patterns $\sigma, \tau \in \mathcal{S}_3$ (with $\sigma \neq \tau$).

The polynomial $f_n(\mathcal{S}_3;\ t_1, \ldots, t_6)$ can actually be specialized in other ways. Suppose that we wanted to compute the bi-variate polynomial

$$\sum_{\pi \in \mathcal{S}_n(132)} s^{N_{123}(\pi)} t^{N_{321}(\pi)}. \tag{12}$$

Observe that this is exactly $f_n(\mathcal{S}_3;\ s, 0, 1, 1, 1, t)$. In other words, we may find the coefficient of t_2^0 in $f_n(\mathcal{S}_3;\ t_1, \ldots, t_6)$ and then set $t_3 = t_4 = t_5 = 1$ and $t_1 = s, t_6 = t$. The same approach can be used to compute the polynomial

$$\sum_{\pi \in \mathcal{S}_n(132)} s^{N_\sigma(\pi)} t^{N_\tau(\pi)}. \tag{13}$$

for any patterns $\sigma, \tau \in \mathcal{S}_3 \backslash \{132\}$ (with $\sigma \neq \tau$).

The analogous specialization can be done to quickly compute

$$\sum_{\pi \in \mathcal{S}_n(123)} s^{N_\sigma(\pi)} t^{N_\tau(\pi)}. \tag{14}$$

for any patterns $\sigma, \tau \in \mathcal{S}_3 \backslash \{123\}$ (with $\sigma \neq \tau$). In general, for any $p \in \mathcal{S}_3$, we can quickly compute

$$\sum_{\pi \in \mathcal{S}_n(p)} s^{N_\sigma(\pi)} t^{N_\tau(\pi)}. \tag{15}$$

for any patterns $\sigma, \tau \in \mathcal{S}_3 \backslash \{p\}$ (with $\sigma \neq \tau$).

We can also adapt the polynomial $f_n(1234, \mathcal{S}_3;\ s, t_1, t_2, \ldots, t_6)$ (from Eq. 10) similarly. In particular, we can quickly compute the polynomial

$$\sum_{\pi \in \mathcal{S}_n(1234)} s^{N_\sigma(\pi)} t^{N_\tau(\pi)}. \tag{16}$$

for any patterns $\sigma, \tau \in \mathcal{S}_3$ (with $\sigma \neq \tau$) by setting $s = 0$ (i.e. extracting the coefficient of s^0) and setting the appropriate t_i's to 1 in $f_n(1234, \mathcal{S}_3;\ s, t_1, t_2, \ldots, t_6)$.

The previously discussed functional equation approaches have been implemented in the Maple packages `PDSn`, `PDAV132`, `PDAV123`, and `PDAV1234`.

3 Computing moments for random permutations

3.1 Moments for random permutations from $\mathcal{S}_n$

The previously discussed functional equations approach allows us to compute both rigorous and empirical statistical properties on permutations.

For some fixed n and fixed pattern $\sigma \in \mathcal{S}_k$, suppose that a permutation $\pi \in \mathcal{S}_n$ is chosen uniformly at random. Let the random variable $X_\sigma(\pi)$ be the number of occurrences of the pattern σ in π. It is not hard to compute the expected value (i.e., the first moment of the random variable X): $\mathbb{E}[X] = \binom{n}{k}/k!$. More generally, it was shown in [16] that each of the higher moments of X is a polynomial in n. In particular, the r-th moment about the mean of X, which is $\mathbb{E}[(X - \mathbb{E}[X])^r]$, is a polynomial of degree $\lfloor r(k - 1/2) \rfloor$ for $r \geq 2$.[1]

For the patterns σ that were discussed in the previous section, the functional equations approach allows us to quickly compute $f_n(\sigma;\ t)$ for any desired n. Observe that $f_n(\sigma;\ t)/n!$ gives us the polynomial where the coefficient of t^i is the probability that a randomly chosen $\pi \in \mathcal{S}_n$ will have exactly i copies of σ. The important point is that we can (rigorously) find a closed-form expression (in n) for the higher order moments of X by computing sufficiently many terms to fit the polynomial.

For example, it was shown in [16] that the exact expression for the second moment (about the mean) of the random variable X_{123} (over $\mathcal{S}_n$) is:

$$\frac{n(n-1)(n-2)(39n^2 + 102n - 157)}{21600} \tag{17}$$

and that the third moment (about the mean) of the random variable X_{123} (over $\mathcal{S}_n$) is:

$$\frac{n(n-1)(n-2)(1437n^4 + 5592n^3 - 11277n^2 - 33990n + 34082)}{6350400} \tag{18}$$

Similarly, the exact expression for the second moment (about the mean) of the random variable X_{132} (over $\mathcal{S}_n$) is:

$$\frac{n(n-1)(n-2)(21n^2 + 78n + 77)}{21600} \tag{19}$$

and that the third moment (about the mean) of the random variable X_{132} (over $\mathcal{S}_n$) is:

$$\frac{n(n-1)(n-2)(129n^4 + 3705n^3 + 5355n^2 + 8655n + 11356)}{12700800} \tag{20}$$

We may also consider mixed moments for two patterns σ and τ. Suppose that a permutation π is chosen uniformly at random from $\mathcal{S}_n$, and again let the random variable $X_\sigma(\pi)$ be the number

[1] This corrects a minor inaccuracy in [16].

of occurrences of pattern σ in π (and equivalently for $X_\tau(\pi)$). It was also shown in [16] that the mixed moments of the random variables X_σ and X_τ (about their respective means) are also polynomials in n. This allows us to rigorously find closed-form expressions (in n) for the higher order mixed moments by computing enough terms to find the polynomial.

For example, the covariance of the two random variables X_{123} and X_{132} is:

$$\frac{n(n-1)(n-2)(18n^2-51n-109)}{21600} \tag{21}$$

while the covariance of the two random variables X_{123} and X_{312} is:

$$-\frac{n(n-1)(n-2)(39n^2-48n-7)}{43200} \tag{22}$$

and the covariance of the two random variables X_{123} and X_{321} is:

$$-\frac{n(n-1)(n-2)(9n^2+12n-92)}{5400} \tag{23}$$

Similar results for other random variables can be derived using the Maple packages available on the authors' website.

3.2 Moments for random permutations from $\mathcal{S}_n(\tau)$

There has been a flurry of recent activity studying occurrences of patterns in the set of permutations avoiding specific patterns. Many of the recent articles focus on counting the total number of occurrences of a pattern in $\mathcal{S}_n(132)$ or in $\mathcal{S}_n(123)$. Some examples (as previously mentioned) include [2, 4, 9, 13]. It is important to note that finding the total number of occurrences of pattern σ in the set $\mathcal{S}_n(\tau)$ is equivalent to picking a permutation uniformly at random from $\mathcal{S}_n(\tau)$ and finding the expected value $\mathbb{E}[X_\sigma]$ (assuming that the enumeration of $\mathcal{S}_n(\tau)$ is known).

In the previous section, we were able to rigorously derive closed-form expressions for moments of the random variable $X_\sigma(\pi)$ when the permutation π was randomly chosen from $\mathcal{S}_n$. While we currently cannot derive similar rigorous results for random permutations from $\mathcal{S}_n(\tau)$, we can still compute numerical moments for a variety of cases. Interestingly, a number of such random variables appear to *not* be asymptotically normal (as opposed to when $\pi \in \mathcal{S}_n$, where Miklós Bóna showed that such random variables are asymptotically normal [1], see also Section 4).

3.2.1 Permutations from $\mathcal{S}_{132}$

Suppose a permutation is chosen uniformly at random from $\mathcal{S}_n(132)$. Using the Maple packages that accompany this article, we can compute many empirical moments. The expected values of the random variables X_{123}, X_{312}, and X_{321} for $1 \le n \le 10$ can be found in Table 1.

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
123	0	0	0.200	0.714	1.619	2.970	4.809	7.171	10.083	13.570
312	0	0	0.200	0.786	1.929	3.790	6.513	10.244	15.115	21.253
321	0	0	0.200	0.929	2.595	5.667	10.653	18.097	28.572	42.672

Table 1: Expected values (first moments) of $X_{123}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(132)$.

The second moments (about the mean) of the random variables X_{123}, X_{312}, and X_{321} for $1 \le n \le 10$ can be found in Table 2.

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
123	0	0	0.160	1.204	4.617	12.757	28.933	57.463	103.720	174.140
312	0	0	0.160	1.026	3.733	10.213	23.392	47.403	87.787	151.710
321	0	0	0.160	1.352	6.003	19.101	49.313	110.180	221.360	409.960

Table 2: Second moments (about the mean) of $X_{123}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(132)$.

r-th moment	$n=15$	$n=16$	$n=17$	$n=18$	$n=19$	$n=20$
$r=3$	0.41867	0.42461	0.43073	0.43690	0.44303	0.44906
$r=4$	2.92652	2.95682	2.98412	3.00889	3.03152	3.05231
$r=5$	3.59958	3.69377	3.78619	3.87633	3.96389	4.04860
$r=6$	14.79293	15.24562	15.66679	16.06007	16.42853	16.77483

Table 3: r-th standardized moments for $X_{312}(\pi)$ for $3 \leq r \leq 6$, where π is chosen uniformly at random from $\mathcal{S}_n(132)$.

Data for the higher moments can be found on the authors websites. For example, the r-th standardized moments for X_{312} when $3 \leq r \leq 6$ and $15 \leq n \leq 20$ can be found in Table 3.

It is interesting to note that the random variable X_{312} does not appear to be asymptotically normal since the 3-rd and 5-th standard moments appear to be increasing (as opposed to going to 0 as a normal distribution would) and the 6-th moment appears to be larger than 15 (the value for a normal distribution).

This approach can also be used to consider the mixed (i, j) moments. For example, the mixed (i, j) moments of the random variables X_{123} and X_{321} for $3 \leq n \leq 10$ can be found in Table 4.

(i,j)	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
$(1,1)$	-0.040	-0.663	-3.392	-11.162	-28.714	-62.970	-123.370	-222.180
$(1,2)$	-0.024	-0.350	-1.445	-0.404	21.587	127.800	478.610	1417.300
$(2,1)$	-0.024	-0.644	-6.657	-38.272	-154.230	-491.000	-1322.000	-3140.400
$(2,2)$	0.011	1.288	33.666	382.200	2650.400	13264.000	52628.000	175500.000

Table 4: Mixed (i, j) moments of $X_{123}(\pi)$ and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(132)$.

Analogous data and outputs can be found on the authors websites.

3.2.2 Permutations from $\mathcal{S}_{123}$

Suppose a permutation is chosen uniformly at random from $\mathcal{S}_n(123)$. Using the Maple packages that accompany this article, we can compute many empirical moments. The expected values of the random variables X_{132}, X_{312}, and X_{321} for $1 \leq n \leq 10$ can be found in Table 5.

The second moments (about the mean) of the random variables X_{132}, X_{312}, and X_{321} for $1 \leq n \leq 10$ can be found in Table 6.

Data for the higher moments can be found on the authors websites. For example, the r-th standardized moments for X_{132} when $3 \leq r \leq 6$ and $15 \leq n \leq 20$ can be found in Table 7.

It is interesting to note that the random variable X_{132} does not appear to be asymptotically normal since the 3-rd and 5-th standard moments appear to be increasing (as opposed to going to

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
132	0	0	0.200	0.643	1.357	2.364	3.678	5.314	7.281	9.589
312	0	0	0.200	0.786	1.929	3.788	6.513	10.244	15.115	21.253
321	0	0	0.200	1.143	3.429	7.697	14.618	24.884	39.208	58.317

Table 5: Expected values (first moments) of $X_{132}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(123)$.

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
132	0	0	0.160	0.801	2.468	5.959	12.344	22.978	39.506	63.877
312	0	0	0.160	0.740	2.114	4.804	9.532	17.303	29.501	48.000
321	0	0	0.160	1.122	4.293	12.423	30.287	65.419	128.910	236.250

Table 6: Second moments (about the mean) of $X_{132}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(123)$.

r-th moment	$n=15$	$n=16$	$n=17$	$n=18$	$n=19$	$n=20$
$r=3$	1.53492	1.54020	1.54458	1.54823	1.55129	1.55385
$r=4$	6.28717	6.33967	6.38469	6.42356	6.45735	6.48687
$r=5$	23.59568	23.99423	24.34048	24.64315	24.90923	25.14433
$r=6$	108.90240	111.90699	114.55548	116.90184	118.99022	120.85698

Table 7: r-th standardized moments for $X_{132}(\pi)$ for $3 \le r \le 6$, where π is chosen uniformly at random from $\mathcal{S}_n(123)$.

0 as a normal distribution would), the 4-th moment appears to be larger than 3 (the value for a normal distribution), and the 6-th moment appears to be substantially larger than 15 (the value for a normal distribution).

This approach can also be used to consider the mixed (i, j) moments. For example, the mixed (i, j) moments of the random variables X_{132} and X_{312} for $3 \le n \le 10$ can be found in Table 8.

(i,j)	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
$(1,1)$	−0.040	−0.219	−0.641	−1.362	−2.332	−3.326	−3.890	−3.269
$(1,2)$	−0.024	−0.099	−0.039	0.841	3.917	11.254	25.372	48.890
$(2,1)$	−0.024	−0.386	−2.261	−8.566	−24.874	−60.099	−126.620	−239.570
$(2,2)$	0.011	0.551	6.309	39.592	172.880	592.420	1709.800	4350.100

Table 8: Mixed (i, j) moments of $X_{132}(\pi)$ and $X_{312}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(123)$.

Analogous data and outputs can be found on the authors websites.

3.2.3 Permutations from $\mathcal{S}_{1234}$

Suppose a permutation is chosen uniformly at random from $\mathcal{S}_n(1234)$. Using the Maple packages that accompany this article, we can compute many empirical moments. The expected values of the random variables X_{123}, X_{132}, X_{312}, and X_{321} for $1 \le n \le 10$ can be found in Table 9.

The second moments (about the mean) of the random variables X_{123}, X_{312}, and X_{321} for $1 \le n \le 10$ can be found in Table 10.

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
123	0	0	0.167	0.522	1.049	1.739	2.592	3.611	4.796	6.153
132	0	0	0.167	0.696	1.709	3.279	5.457	8.283	11.789	16.004
312	0	0	0.167	0.696	1.796	3.684	6.575	10.679	16.202	23.341
321	0	0	0.167	0.696	1.942	4.335	8.344	14.466	23.223	35.158

Table 9: Expected values (first moments) of $X_{123}(\pi)$, $X_{132}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(1234)$.

Pattern	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
123	0	0	0.139	0.510	1.172	2.236	3.863	6.257	9.654	14.324
132	0	0	0.139	0.820	2.828	7.332	15.959	30.863	54.767	91.002
312	0	0	0.139	0.820	2.667	6.524	13.484	24.911	42.468	68.157
321	0	0	0.139	0.994	3.764	10.566	24.936	52.338	100.740	181.280

Table 10: Second moments (about the mean) of $X_{123}(\pi)$, $X_{132}(\pi)$, $X_{312}(\pi)$, and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(1234)$.

Data for the higher moments can be found on the authors websites. For example, the r-th standardized moments for X_{123} when $3 \le r \le 6$ and $13 \le n \le 18$ can be found in Table 11.

r-th moment	$n=13$	$n=14$	$n=15$	$n=16$	$n=17$	$n=18$
$r=3$	1.14140	1.16076	1.17518	1.18585	1.19365	1.19926
$r=4$	5.14732	5.21356	5.26297	5.29971	5.32683	5.34656
$r=5$	16.61123	17.07925	17.43934	17.71522	17.92523	18.08348
$r=6$	74.59126	77.40043	79.60569	81.33022	82.67201	83.70841

Table 11: r-th standardized moments for $X_{123}(\pi)$ for $3 \le r \le 6$, where π is chosen uniformly at random from $\mathcal{S}_n(1234)$.

It is interesting to note that the random variable X_{123} does not appear to be asymptotically normal since the 3-rd and 5-th standard moments appear to be increasing (as opposed to going to 0 as a normal distribution would), the 4-th moment appears to be larger than 3 (the value for a normal distribution), and the 6-th moment appears to be substantially larger than 15 (the value for a normal distribution).

This approach can also be used to consider the mixed (i, j) moments. For example, the mixed (i, j) moments of the random variables X_{123} and X_{321} for $3 \le n \le 10$ can be found in Table 12.

(i,j)	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
$(1,1)$	-0.028	-0.363	-1.298	-3.258	-6.892	-13.121	-23.171	-38.611
$(1,2)$	-0.019	-0.266	-1.674	-5.958	-15.301	-31.716	-55.546	-82.648
$(2,1)$	-0.019	-0.166	-0.505	-1.531	-4.798	-13.664	-34.352	-77.387
$(2,2)$	0.007	0.386	4.969	33.937	159.600	593.990	1880.700	5274.100

Table 12: Mixed (i, j) moments of $X_{123}(\pi)$ and $X_{321}(\pi)$, where π is chosen uniformly at random from $\mathcal{S}_n(1234)$.

Analogous data and outputs can be found on the authors websites.

4 Joint asymptotic normality of multiple patterns

In this section we let π be a permutation chosen uniformly at random from $\mathcal{S}_n$ (without any condition) and we study the joint distribution of the random variables $X_{\sigma,n} := X_\sigma(\pi)$, the number of copies of σ in π, for different patterns $\sigma \in \mathcal{S}_* := \bigcup_{k=1}^{\infty} \mathcal{S}_k$. We consider asymptotics as $n \to \infty$ for (one or several) fixed σ.

Each $X_{\sigma,n}$ has an asymptotic normal distribution, as was shown by Bona [1] (see also [3]). We give another (perhaps simpler) proof of this; moreover, we extend the result to joint asymptotic normality for several patterns σ.

The asymptotic variances and covariances depend on the patterns in a slightly complicated way, so we begin with some definitions. For $k \geq 1$ and $1 \leq i \leq k$, define

$$g_{k,i}(x) := \binom{k-1}{i-1} x^{i-1}(1-x)^{k-i}. \tag{24}$$

For a permutation $\sigma \in S_k$, define

$$G_\sigma(x,y) := \frac{1}{(k-1)!^2}\left(\sum_{i=1}^{k} g_{k,i}(x) g_{k,\sigma(i)}(y) - \frac{1}{k}\right). \tag{25}$$

Let Z_σ, $\sigma \in \mathcal{S}_*$, be jointly normal random variables with $\mathbb{E}\, Z_\sigma = 0$ and (co)variances

$$\operatorname{Cov}(Z_\sigma, Z_\tau) = \Sigma_{\sigma,\tau} := \langle G_\sigma, G_\tau \rangle_{L^2([0,1]^2)} := \int_0^1 \int_0^1 G_\sigma(x,y) G_\tau(x,y)\, \mathrm{d}x\, \mathrm{d}y. \tag{26}$$

(Such normal random variables exist since the matrix $(\Sigma_{\sigma,\tau})_{\sigma,\tau}$ is non-negative definite. As is well known, the joint distribution is uniquely defined by the means and covariances.)

We denote the length of a permutation σ by $|\sigma|$, and let $\xrightarrow{\mathrm{d}}$ denote convergence in distribution of random variables.

Theorem 4.1. *For every pattern $\sigma \in \mathcal{S}_*$, as $n \to \infty$,*

$$\frac{X_{\sigma,n} - \mathbb{E}\, X_{\sigma,n}}{n^{|\sigma|-1/2}} = \frac{X_{\sigma,n} - \frac{1}{|\sigma|!}\binom{n}{|\sigma|}}{n^{|\sigma|-1/2}} \xrightarrow{\mathrm{d}} Z_\sigma. \tag{27}$$

Moreover, this holds jointly for any finite family of patterns σ. Furthermore, all (joint) moments converge; in particular, for any permutations σ, τ

$$\frac{\operatorname{Cov}(X_{\sigma,n}, X_{\tau,n})}{n^{|\sigma|+|\tau|-1}} \to \Sigma_{\sigma,\tau}. \tag{28}$$

Before giving the proof we give some comments. First, as noted above, if σ has length $|\sigma| = k$,

$$\mathbb{E}\, X_{\sigma,n} = \binom{n}{k}\frac{1}{k!} \sim \frac{1}{k!^2} n^k, \qquad \text{as } n \to \infty. \tag{29}$$

The asymptotic covariances $\Sigma_{\sigma,\tau}$ can be computed explicitly. By a beta integral,

$$\int_0^1 g_{k,i}(x)\, dx = \binom{k-1}{i-1}\frac{\Gamma(i)\Gamma(k-i+1)}{\Gamma(k+1)} = \frac{1}{k}, \tag{30}$$

and similarly, for any $k, \ell \geq 1$ and $1 \leq i \leq k$, $1 \leq j \leq \ell$,

$$\begin{aligned}\int_0^1 g_{k,i}(x) g_{\ell,j}(x)\, dx &= \binom{k-1}{i-1}\binom{\ell-1}{j-1}\frac{\Gamma(i+j-1)\Gamma(k+\ell-i-j+1)}{\Gamma(k+\ell)} \\ &= \frac{(k-1)!\,(\ell-1)!}{(k+\ell-1)!}\binom{i+j-2}{i-1}\binom{k+\ell-i-j}{k-i}.\end{aligned} \tag{31}$$

It follows from (30) that, if $|\sigma| = k$,

$$\int_0^1 \int_0^1 \sum_{i=1}^{k} g_{k,i}(x) g_{k,\sigma(i)}(y)\,\mathrm{d}x\,\mathrm{d}y = \frac{k}{k^2} = \frac{1}{k} \tag{32}$$

which implies, using (31) twice, if further $|\tau| = \ell$,

$$\begin{aligned}
\int_0^1 \int_0^1 &\Bigl(\sum_{i=1}^{k} g_{k,i}(x) g_{k,\sigma(i)}(y) - \frac{1}{k}\Bigr)\Bigl(\sum_{j=1}^{\ell} g_{\ell,j}(x) g_{\ell,\tau(j)}(y) - \frac{1}{\ell}\Bigr)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^1 \int_0^1 \sum_{i=1}^{k} g_{k,i}(x) g_{k,\sigma(i)}(y) \sum_{j=1}^{\ell} g_{\ell,j}(x) g_{\ell,\tau(j)}(y)\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{k\ell} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{\ell} \int_0^1 g_{k,i}(x) g_{\ell,j}(x)\,\mathrm{d}x \int_0^1 g_{k,\sigma(i)}(y) g_{\ell,\tau(j)}(y)\,\mathrm{d}y - \frac{1}{k\ell} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{\ell} \frac{(k-1)!^2\,(\ell-1)!^2}{(k+\ell-1)!^2} \binom{i+j-2}{i-1}\binom{k+\ell-i-j}{k-i}\binom{\sigma(i)+\tau(j)-2}{\sigma(i)-1}\binom{k+\ell-\sigma(i)-\tau(j)}{k-\sigma(i)} \\
&\qquad - \frac{1}{k\ell}.
\end{aligned}$$

Consequently, by (26) and (25), if $|\sigma| = k$ and $|\tau| = \ell$, then

$$\begin{aligned}
\Sigma_{\sigma,\tau} = \frac{1}{(k+\ell-1)!^2} \sum_{i=1}^{k}\sum_{j=1}^{\ell} \binom{i+j-2}{i-1}\binom{k+\ell-i-j}{k-i}\binom{\sigma(i)+\tau(j)-2}{\sigma(i)-1}\binom{k+\ell-\sigma(i)-\tau(j)}{k-\sigma(i)} \\
- \frac{1}{(k-1)!\,k!\,(\ell-1)!\,\ell!}.
\end{aligned} \tag{33}$$

Proof of Theorem 4.1. Let $U_1, \dots, U_n$ be independent and identically distributed (i.i.d.) random variables with a uniform distribution on $[0, 1]$. It is a standard trick that (by symmetry) the reduction $\mathrm{red}(U_1, \dots, U_n)$ is a uniformly random permutation in $\mathcal{S}_n$ (note that $U_1, \dots, U_n$ almost surely are distinct), so we can take this as our random π and obtain the representation, with $k = |\sigma|$,

$$X_{\sigma,n} = X_\sigma(\pi) = \sum_{i_1<\dots<i_k} \mathbf{1}\left[\mathrm{red}(U_{i_1}, \dots, U_{i_k}) = \sigma\right]. \tag{34}$$

This is an example of an asymmetric U-statistic, and (a rather simple instance of) the general theory in [14, Section 11.2] can be used to show the theorem. However, the details are a bit technical, in particular to calculate the asymptotic covariances, so we will instead use another, more symmetric representation. (See [14, Remark 11.21].)

Let $V_1, \dots, V_n$ be another sequence of i.i.d. random variables, uniformly distributed on [0,1] and independent of $U_1, \dots, U_n$. Let π' be the permutation that sorts these numbers such that $V_{\pi'(1)} < \dots < V_{\pi'(n)}$ and let π be the reduction of $U_{\pi'(1)}, \dots, U_{\pi'(n)}$. Then π is still uniformly random, and it is easy to see that

$$\begin{aligned}
X_{\sigma,n} = X_\sigma(\pi) &:= \sum_{i_1<\dots<i_k} \mathbf{1}\left[\mathrm{red}(U_{\pi'(i_1)}, \dots, U_{\pi'(i_k)}) = \sigma\right] \\
&= \sum_{j_1,\dots,j_k}^{*} \mathbf{1}\left[\mathrm{red}(U_{j_1}, \dots, U_{j_k}) = \sigma\right] \cdot \mathbf{1}\left[V_{j_1} < \dots < V_{j_k}\right],
\end{aligned} \tag{35}$$

where $\sum^*$ denotes summation over all distinct indices $j_1, \dots, j_k$. This representation, while in some ways more complicated that (34), has the great advantage that we sum over all ordered n-tuples of distinct indices; this is thus an example of a U-statistic, and we can apply the basic central limit theorem by Hoeffding [8, Theorem 7.1], see also [15] and [14, Section 11.1]. In order to compute the (co)variances, we follow the path of Hoeffding's proof.

The main idea of Hoeffding's proof of his central limit theorem is to use a projection. In our case we let $W_j := (U_j, V_j) \in [0,1]^2$ and write (35) as

$$X_{\sigma,n} = \sum_{j_1,\dots,j_k}^{*} f_\sigma(W_{j_1}, \dots, W_{j_k}), \tag{36}$$

for a certain (indicator) function f_σ. We then take the conditional expectation of $f_\sigma(W_1, \dots, W_k)$ given one of the variables W_i:

$$f_{\sigma;i}(x,y) := \mathbb{E}\bigl(f_\sigma(W_1, \dots, W_k) \mid W_i = (x,y)\bigr); \tag{37}$$

we also take the expectation

$$\mu := \mathbb{E}\, f_\sigma(W_1, \dots, W_k) = \mathbb{E}\, f_{\sigma;i}(W_i). \tag{38}$$

Hoeffding then shows that if we replace f_σ by $f'_\sigma(W_1, \dots, W_k) := \mu + \sum_{i=1}^k (f_{\sigma,i}(W_i) - \mu)$, then the resulting error for the sum in (36) will have variance $O(n^{2k-2})$, which is negligible with the normalization used in Theorem 4.1. Thus we can approximate $X_{\sigma,n} - \mathbb{E}\, X_{\sigma,n}$ by

$$\sum_{j_1,\dots,j_n}^{*} \sum_{i=1}^{k} \bigl(f_{\sigma;i}(W_{j_i}) - \mu\bigr) = \sum_{i=1}^{k} (n-1)^{\underline{k-1}} \sum_{j=1}^{n} \bigl(f_{\sigma;i}(W_j) - \mu\bigr) = (n-1)^{\underline{k-1}} \sum_{j=1}^{n} F_\sigma(W_j), \tag{39}$$

where $(n-1)^{\underline{k-1}} = (n-1)\cdots(n-k+1)$ and

$$F_\sigma(x,y) := \sum_{i=1}^{k} \bigl(f_{\sigma;i}(x,y) - \mu\bigr). \tag{40}$$

The asymptotic normality of $X_{\sigma,n}$ now follows by the standard central limit theorem for the i.i.d. random variables $F_\sigma(W_j)$, which yields $(X_{\sigma,n} - \mathbb{E}\, X_{\sigma,n})/n^{k-1/2} \xrightarrow{\mathrm{d}} N\bigl(0, \Sigma_{\sigma,\sigma}\bigr)$ where

$$\Sigma_{\sigma,\sigma} := \mathbb{E}\bigl(F_\sigma(W_1)^2\bigr) = \int_0^1 \int_0^1 F_\sigma(x,y)^2 \,\mathrm{d}x \,\mathrm{d}y. \tag{41}$$

Joint normality for several patterns σ (possibly of different lengths) follows in the same way, with the asymptotic covariances

$$\Sigma_{\sigma,\tau} := \mathbb{E}\bigl(F_\sigma(W_1) F_\tau(W_1)\bigr) = \int_0^1 \int_0^1 F_\sigma(x,y) F_\tau(x,y) \,\mathrm{d}x \,\mathrm{d}y. \tag{42}$$

It remains to compute the functions F_σ defined in (40) In order to do this, we see that from (37) and the definition of f_σ as an indicator function, cf. (35)–(36),

$$f_{\sigma;i}(x,y) = \mathbb{P}\bigl(\mathrm{red}(U_1, \dots, U_k) = \sigma \mid U_i = x\bigr)\, \mathbb{P}\bigl(V_1 < \dots < V_k \mid V_i = y\bigr). \tag{43}$$

For the second probability in (43) we require that $V_1, \dots, V_{i-1} < y$ and $V_{i+1}, \dots, V_k > y$, and furthermore that these two sets of variables are increasing; since the variables are independent and uniformly distributed, the probability is, recalling the notation (24),

$$\frac{y^{i-1}}{(i-1)!} \frac{(1-y)^{k-i}}{(k-i)!} = \frac{1}{(k-1)!} g_{k,i}(y). \tag{44}$$

Similarly, for the first probability in (43) we require that the $\sigma(i)$:th smallest of $U_1, \dots, U_k$ is x, and that the others come in the order specified by σ, and the probability of this is $(k-1)!^{-1} g_{k,\sigma(i)}(x)$. Consequently,

$$f_{\sigma;i}(x,y) = \frac{1}{(k-1)!^2}\, g_{k,\sigma(i)}(x) g_{k,i}(y). \tag{45}$$

Furthermore,

$$\mu := \mathbb{E}\, f_\sigma(W_1,\dots,W_k) = \mathbb{P}\big(\operatorname{red}(U_1,\dots,U_k)=\sigma\big)\,\mathbb{P}\big(V_1 < \dots < V_k\big) = \frac{1}{k!^2}. \tag{46}$$

It follows from (40), (45), (46) and (25) that $F_\sigma(x,y) = G_\sigma(y,x)$. Hence (41)–(42) agree with (26), and Hoeffding's theorem yields (27).

Hoeffding's theorem (and its proof sketched above) yields also the convergence (28) of the covariances. To see that moment convergence holds also for higher moments, let m be a positive integer. By (36),

$$\mathbb{E}\big(X_{\sigma,n} - \mathbb{E}\, X_{\sigma,n}\big)^m = \sum^*_{j_{11},\dots,j_{k1}} \cdots \sum^*_{j_{1m},\dots,j_{km}} \mathbb{E} \prod_{i=1}^m \big(f_\sigma(W_{j_{1i}},\dots,W_{j_{ki}}) - \mu\big) \tag{47}$$

where the expectation on the right-hand side vanishes unless each index set $\{j_{1i},\dots,j_{mi}\}$ contains at least one index shared by another such set. In this case, however, there are at most $mk - m/2$ distinct indices, and it follows that the moment (47) is a polynomial in n of degree at most $mk - m/2$. In particular, the normalized central moment $\mathbb{E}\big((X_{\sigma,n} - \mathbb{E}\, X_{\sigma,n})/n^{k-1/2}\big)^m = O(1)$. If m is an even integer, this implies, by standard results on uniform integrability, that all moments of lower order converge to the corresponding moments of the limit Z_σ, and the same holds for joint moments. Since m is arbitrary, this shows convergence of all moments.

□

Example 4.2. The case $k = 1$ is trivial, with $X_{1,n} = n$ deterministic. Indeed, (24)–(25) yield $g_{1,1}(x) = 1$ and $G_{1,1}(x,y) = 0$.

Example 4.3. The simplest non-trivial example is $k = 2$, where $X_{21}(\pi)$ is the *number of inversions* in π. The distribution of this random variable, for π uniformly at random in $\mathcal{S}_n$, is called the *Mahonian distribution*, and it is well-known that it is asymptotically normal, see e.g. [7, Section X.6]. (See [5] for the case of permutations of multi-sets; it would be interesting to obtain similar results for other patterns in multi-set permutations.) A simple calculation using (24)–(25) yields

$$G_{21,21}(x,y) = -2\big(x - \tfrac12\big)\big(y - \tfrac12\big) \tag{48}$$

and (26) or (33) yields $\Sigma_{21,21} = 1/36$. Hence Theorem 4.1 in this case yields the well-known

$$\frac{X_{21,n} - \frac12\binom{n}{2}}{n^{3/2}} \xrightarrow{\text{d}} N\big(0, 1/36\big). \tag{49}$$

Example 4.2 is the only case when the limit Z_σ in Theorem 4.1 vanishes, as we show next.

Theorem 4.4. *If $k > 1$, then $\Sigma_{\sigma,\sigma} > 0$ and thus Z_σ is non-degenerate, for every $\sigma \in \mathcal{S}_k$.*

Proof. By (24),

$$\sum_{i=1}^k g_{k,i}(x) = 1. \tag{50}$$

Hence (25) may be written, using Kronecker's delta $\delta_{i,j}$,

$$G_\sigma(x,y) := \frac{1}{(k-1)!^2}\sum_{i=1}^{k}\sum_{j=1}^{k}\Bigl(\delta_{j,\sigma(i)}-\frac{1}{k}\Bigr)g_{k,i}(x)g_{k,j}(y). \tag{51}$$

For a fixed k, the polynomials $g_{k,i}$, $1 \le i \le k$, are linearly independent (and form basis in the k-dimensional vector space of polynomials of degree $\le k-1$). Hence the k^2 tensor products $g_{k,i}(x)g_{k,j}(y)$ are linearly independent in $L^2([0,1]^2)$, and it follows from (51) and (26) that if $k \ge 2$, then G_σ is not identically 0 and thus $\Sigma_{\sigma,\sigma} = \iint G_\sigma(x,y)^2 > 0$. ☐

For a given k we have $k!$ patterns $\sigma \in \mathcal{S}_k$ and thus $k!$ limit variables Z_σ. We have just seen that (if $k > 1$) these are all non-degenerate; however, they are not linearly independent. For example, the sum $\sum_{\sigma\in\mathcal{S}_k} X_\sigma(\pi) = \binom{n}{k}$ for every π, so the sum is deterministic and it follows that $\sum_{\sigma\in\mathcal{S}_k} Z_\sigma = 0$. Many non-trivial linear combinations vanish too, as is seen by the following theorem.

Theorem 4.5. *Let $k \ge 1$. The $k!$ limit random variables Z_σ, $\sigma \in \mathcal{S}_k$, span a linear space of dimension $(k-1)^2$.*

Proof. By the definition (26), this linear space, V say, is isomorphic (and isometric for the appropriate L^2-norms) to the linear space V_1 spanned by the functions G_σ on $[0,1]^2$. Furthermore, by (51) and the comments after it, V_1 is isomorphic to the linear space V_2 of $k \times k$ matrices spanned by the matrices $A_\sigma := \bigl(\delta_{j,\sigma(i)} - \frac{1}{k}\bigr)_{ij=1}^{k}$. Let V_3 be the space of all $k \times k$ matrices with all row sums and column sums 0. Then each matrix $A_\sigma \in V_3$ and thus $V_2 \subseteq V_3$. Conversely, it is easily seen that each matrix in V_3 is a linear combination of matrices A_σ, for example using the well-known fact that every doubly stochastic matrix is a convex combination of permutation matrices. Hence $V_2 = V_3$. Finally, $\dim(V_3) = (k-1)^2$ since a matrix in V_3 is uniquely determined by its upper left corner $(k-1)\times(k-1)$ submatrix obtained by deleting the last row and column, and conversely this submatrix may be chosen arbitrarily. ☐

Example 4.6. There are 6 patterns of length $k = 3$. Taking them in lexicographic order 123, 132, 213, 231, 312, 321, and using Maple to calculate the covariance matrix of the limit variables Z_σ by (24)–(26), we find

$$\bigl(\operatorname{Cov}(Z_\sigma, Z_\tau)\bigr)_{\sigma,\tau\in\mathcal{S}_3} = \bigl(\Sigma_{\sigma,\tau}\bigr)_{\sigma,\tau\in\mathcal{S}_3} = \frac{1}{5!^2}\begin{pmatrix} 26 & 12 & 12 & -13 & -13 & -24 \\ 12 & 14 & -1 & -6 & -6 & -13 \\ 12 & -1 & 14 & -6 & -6 & -13 \\ -13 & -6 & -6 & 14 & -1 & 12 \\ -13 & -6 & -6 & -1 & 14 & 12 \\ -24 & -13 & -13 & 12 & 12 & 26 \end{pmatrix}. \tag{52}$$

We note that the asymptotic variances differ between different patterns; they are 13/7200 (for 123 and 321) or 7/7200 (for the other patterns).

The eigenvalues of the covariance matrix (52) are

$$\frac{3}{5!^2}\bigl(25, 5, 5, 1, 0, 0\bigr), \tag{53}$$

verifying that this matrix has rank 4 as given by Theorem 4.5. A choice of pairwise orthogonal eigenvectors (in the corresponding order) is

$$\begin{pmatrix} 2 \\ 1 \\ 1 \\ -1 \\ -1 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ -1 \\ -1 \\ -1 \\ -1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \tag{54}$$

Remark 4.7. The last eigenvector in (54) corresponds to the trivial fact mentioned above that the sum of all Z_σ vanishes. The fifth eigenvector, also with eigenvalue 0, says that

$$Z_{123} + Z_{231} + Z_{312} - Z_{132} - Z_{213} - Z_{321} = 0. \tag{55}$$

Let $Y(\pi)$ be the corresponding number

$$Y(\pi) := X_{123}(\pi) + X_{231}(\pi) + X_{312}(\pi) - X_{132}(\pi) - X_{213}(\pi) - X_{321}(\pi), \tag{56}$$

and let $Y_n := Y(\pi)$ with π chosen uniformly at random in $\mathcal{S}_n$. (Note that $Y(\pi)$ is the sum of the signs of the $\binom{n}{3}$ permutations $\operatorname{red}(\pi_{i_1}\pi_{i_2}\pi_{i_3})$.) Theorem 4.1 and (55) thus say that, as $n \to \infty$, $n^{-5/2}Y_n \xrightarrow{\mathrm{d}} 0$. However, in this case, the random variable Y_n does not vanish identically. (Take π as the identity permutation.) Using the same methods as in Section 3.1, we can show that

$$\operatorname{Var}(Y_n) = \frac{n^2(n-1)(n-2)}{18}. \tag{57}$$

In particular, we have that the leading term of $\operatorname{Var}(Y_n)$ is $\frac{1}{18}n^4$, i.e. of order n^{2k-2} instead of n^{2k-1} as in the cases when Theorem 4.1 yields a non-degenerate limit. In such cases, one can use a more advanced version of Hoeffding's argument above and show that there is an asymptotic distribution that can be represented as an (infinite) polynomial of degree 2 in normal random variables; this polynomial can further be diagonalized as a linear combination of squares of independent normal variables, see e.g. [15] and [14, Section 11.1]. In the present case this leads to

$$n^{-2}Y_n \xrightarrow{\mathrm{d}} Y^* = \sum_{\substack{\ell,m=-\infty \\ \ell,m\neq 0}}^{\infty} \frac{1}{2\pi^2\ell m}\bigl(\xi_{\ell,m}^2 - 1\bigr), \tag{58}$$

where $\xi_{\ell,m}$ are i.i.d. standard normal random variables. (We omit the details but note that the bilinear form in [14, Corollary 11.5(iii)] in this case after some calculation turns out to correspond to the convolution operator on $L^2(\mathbb{T}^2)$ given by convolution with $H(x,y) = \frac{1}{6}(2x-1)(2y-1)$ (where we identify the group $\mathbb{T}$ with $[0,1)$); hence its eigenvalues are the Fourier coefficients $\widehat{H}(\ell,m) = -1/(6\pi^2\ell m)$, which yields the coefficients in (58).) Note that, since $\operatorname{Var}(\xi_{\ell,m}^2) = 2$,

$$\operatorname{Var} Y^* = \sum_{\substack{\ell,m=-\infty \\ \ell,m\neq 0}}^{\infty} \frac{2}{4\pi^4\ell^2m^2} = \frac{1}{18}, \tag{59}$$

in accordance with the asymptotic formula $\operatorname{Var}(Y_n) \sim n^4/18$. Furthermore, the representation (58) of the limit Y yields its moment generating function as

$$\mathbb{E}\, e^{tY^*} = \prod_{\substack{\ell,m=-\infty \\ \ell,m\neq 0}}^{\infty} \Bigl(1 - \frac{2t}{2\pi^2\ell m}\Bigr)^{-1/2} = \prod_{\ell,m=1}^{\infty} \Bigl(1 - \frac{t^2}{\pi^4\ell^2m^2}\Bigr)^{-1} = \prod_{m=1}^{\infty} \frac{t/m\pi}{\sin(t/m\pi)}, \qquad |\Re t| < \pi^2. \tag{60}$$

This type of limit is typical of the degenerate cases that can occur for certain linear combinations of pattern counts. It is also possible to obtain higher degeneracies in special cases, with variance of still lower order and a limit that is a polynomial of higher degree in infinitely many normal variables; one example is to generalize (56) by taking, for any fixed $k \geq 3$, the sum of the signs of the $\binom{n}{k}$ patterns of length k occurring in π. It can be seen that for this example, $\mathrm{Var}(Y_n)$ is a polynomial in n of degree $k+1$ only (instead of the typical $2k-1$), because all higher order terms cancel in this highly symmetric example.

Example 4.8. There are 24 patterns of length $k = 4$. A calculation as in Example 4.6 of the covariance matrix yields a 24×24 matrix of rank $(4-1)^2 = 9$. The 9 non-zero eigenvalues are

$$\frac{8}{7!^2}\big(441, 147, 147, 49, 21, 21, 7, 7, 1\big). \tag{61}$$

Similarly, for $k = 5$ the covariance matrix is a 120×120 matrix with the $4^2 = 16$ non-zero eigenvalues

$$\frac{30}{9!^2}\big(7056, 3024, 3024, 1296, 756, 756, 324, 324, 84, 84, 81, 36, 36, 9, 9, 1\big). \tag{62}$$

The fact that the eigenvalues in (53), (61) and (62) all are simple rational numbers suggests that there is a general structure (valid for all k) for these eigenvalues, and presumably also for the corresponding eigenvectors; it would be interesting to know more about this.

5 Conclusion

In this article, we studied the moments and mixed moments of the random variables $X_\sigma(\pi)$ for a number of patterns σ, where π may be chosen from $\mathcal{S}_n$ or a pattern avoiding set $\mathcal{S}_n(\tau)$. In addition, we prove that for any two patterns, the corresponding random variables are joint asymptotically normal when the permutations are drawn from $\mathcal{S}_n$. The contrasting computational approach can compute a number of moments and mixed moments as well as derive (rigorous) formulas for the lower moments. We anticipate that this approach could be extended to provide an alternative proof to the joint asymptotic normality of multiple random variables, but we leave this as "future work".

In the setting where the permutations are chosen from the pattern avoiding set $\mathcal{S}_n(\tau)$ (for some fixed pattern τ), much less is known. Others have recently studied the total number of occurrences of a pattern in these sets, which is equivalent to the expected value (i.e., the first moment) of the random variable X_σ, generally for when both $\sigma, \tau \in \mathcal{S}_3$. Our approach allows us to quickly compute many empirical moments, far beyond the first moment. We expect that a more thorough analysis of these higher moments will uncover interesting properties and that in some cases, these higher moments will also have closed form formulas. In addition, the random variables for some patterns appear to *not* be asymptotically normal (whereas in the case where permutations are drawn from $\mathcal{S}_n$, they are asymptotically normal for every pattern [1]). It would be interesting to understand which patterns (if any) have corresponding random variables that are asymptotically normal when permutations are drawn from $\mathcal{S}_n(\tau)$.

References

[1] Miklós Bóna, The copies of any permutation pattern are asymptotically normal. `arXiv:0712.2792`.

[2] Miklós Bóna, The absence of a pattern and the occurrences of another. *Discrete Math. Theor. Comput. Sci.* **12** (2010), no. 2, 89–102.

[3] Miklós Bóna, On three different notions of monotone subsequences. *Permutation Patterns*, 89–114, London Math. Soc. Lecture Note Ser., 376, Cambridge Univ. Press, Cambridge, 2010.

[4] Miklós Bóna, Surprising symmetries in objects counted by Catalan numbers. *Electron. J. Combin.* **19** (2012), no. 1.

[5] E. Rodney Canfield, Svante Janson and Doron Zeilberger, The Mahonian probability distribution on words is asymptotically normal. *Adv. Appl. Math.* **46** (2011), no. 1–4, 109–124. Corrigendum: *ibid* **49** (2012), no. 1, 77.

[6] Joshua Cooper, Combinatorial Problems I like. *internet resource*, `http://www.math.sc.edu/%7Ecooper/combprob.html`.

[7] William Feller, *An Introduction to Probability Theory and Its Application*, volume I, third edition, Wiley, New York, 1968.

[8] Wassily Hoeffding, A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19**, (1948). 293–325.

[9] Cheyne Homberger, Expected patterns in permutation classes. *Electron. J. Combin.* **19**, (2012), no. 3.

[10] Brian Nakamura, Approaches for enumerating permutations with a prescribed number of occurrences of patterns. *Pure Math. Appl. (PU.M.A.)*, to appear.

[11] Brian Nakamura and Doron Zeilberger, Using Noonan-Zeilberger functional equations to enumerate (in polynomial time!) generalized Wilf classes. *Adv. in Appl. Math.* **50**, (2013), no. 3, 356–366.

[12] John Noonan and Doron Zeilberger, The enumeration of permutations with a prescribed number of "forbidden" patterns. *Adv. in Appl. Math.* **17**, (1996), no. 4, 381–407.

[13] Kate Rudolph, Pattern popularity in 132-avoiding permutations. *Electron. J. Combin.* **20**, (2013), no. 1.

[14] Svante Janson, *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, UK, 1997.

[15] H. Rubin and R. A. Vitale, Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** (1980), no. 1, 165–170.

[16] Doron Zeilberger, Symbolic moment calculus. I. Foundations and permutation pattern statistics. *Ann. Comb.* **8** (2004), no. 3, 369–378.