Stein's Method: Expository Lectures and Applications
Institute of Mathematical Statistics
Lecture Notes – Monograph Series
Vol. 46 (2004) 66–74

4. Stein's method and non-reversible Markov chains

Jason Fulman

University of Pittsburgh

Abstract: Let $W(\pi)$ be either the number of descents or inversions of a permutation $\pi \in S_n$. Stein's method is applied to show that W satisfies a central limit theorem with error rate $n^{-1/2}$. The construction of an exchangeable pair (W, W') used in Stein's method is non-trivial and uses a non-reversible Markov chain.

4.1. Introduction

We begin by recalling two permutation statistics on the symmetric group S_n which are of interest to combinatorialists and statisticians. A good introduction to the combinatorial aspects of permutation statistics is Chapter 1 of Stanley [14], and a superb account of their applications to statistical problems is Chapter 6 of Diaconis [5].

The first statistic on S_n is $\mathrm{Des}(\pi)$, the number of descents of π. This is defined as the number of pairs $(i, i+1)$ with $1 \leq i \leq n-1$ such that $\pi(i) > \pi(i+1)$. Writing π in two-line form, this is the number of times the value of the permutation π decreases. (A more general definition of descents exists for Coxeter groups: the number of height one positive roots sent to negative roots by π). The number of permutations π in S_n with $k+1$ descents is also called the Eulerian number $A(n,k)$ and has been studied extensively [6], [8], [11]. Several proofs are known for the asymptotic $(n \to \infty)$ normality of $A(n,k)$. See for instance Diaconis and Pitman [6], Pitman [12], Bender [2], and Tanny [16]. A proof using the method of moments is also possible.

A second well-studied statistic on S_n is $\mathrm{Inv}(\pi)$, the number of inversions of π. In the statistics community this is called Kendall's tau. Inv is defined as the number of pairs (i,j) with $1 \leq i < j \leq n$ such that $\pi(i) > \pi(j)$. Writing π in two-line form, this is the number of pairs (i,j) whose values are out of order. $I(\pi)$ is also the length of π in terms of the standard generators $\{(i, i+1) : 1 \leq i \leq n-1\}$ for S_n. (For an arbitrary Coxeter group, $\mathrm{Inv}(\pi)$ is the number of positive roots sent to negative roots by π). Proofs of the asymptotic normality of $\mathrm{Inv}(\pi)$ for S_n can be found in Bender [2] and Chapter 6 of Diaconis [5].

The following definition generalizes both of these statistics. Let $M = (M_{i,j})$ be a real, anti-symmetric, $n * n$ matrix. Let X be the random variable on S_n defined by $X(\pi) = \sum_{i<j} M_{\pi(i),\pi(j)}$. Setting $M_{i,j} = -1$ if $j = i+1$, $M_{i,j} = 1$ if $j = i-1$, and $M_{i,j} = 0$ otherwise leads to $X(\pi) = 2\mathrm{Des}(\pi^{-1}) - (n-1)$. Setting $M_{i,j} = -1$ if $i < j$, $M_{i,j} = +1$ if $i > j$, and $M_{i,i} = 0$ leads to $X(\pi) = 2\mathrm{Inv}(\pi^{-1}) - \binom{n}{2}$. Define $W = \frac{X}{\sqrt{\mathrm{Var}(X)}}$, so that W has mean 0 and variance 1.

Charles Stein developed a method for bounding the sup norm between the distribution of a random variable and the standard normal distribution. His technique

Department of Mathematics, Thackeray 514, University of Pittsburgh, Pittsburgh, PA 15260, USA. e-mail: fulman@math.pitt.edu

has come to be known as Stein's method. Stein's book [15] and the papers in this volume are good references.

Let us recall some notation from probability theory. If Y, Z are random variables on a probability space (Ω, B, P), we let $E(Y)$ denote the expected value of Y and $E^Z(Y)$ the expected value of Y given Z, where both expectations are taken under P. In the case at hand, Ω is S_n, B is all subsets of S_n, and P is the uniform distribution. Call W, W' an exchangeable pair of random variables on S_n if $P(W = w_1, W' = w_2) = P(W = w_2, W' = w_1)$.

Theorem 4.1.1 is due to Rinott and Rotar [13].

Theorem 4.1.1 ([13]). *Let W, W' be an exchangeable pair of real random variables such that $E^W W' = (1-\lambda)W$ with $0 < \lambda < 1$. Suppose moreover that $|W' - W| \leq A$ for some constant A. Then for all real x,*

$$\left|P\{W \leq x\} - \Phi(x)\right| \leq \frac{12}{\lambda}\sqrt{Var(E^W(W'-W)^2)} + 48\frac{A^3}{\lambda} + 8\frac{A^2}{\sqrt{\lambda}}$$

where Φ is the standard normal distribution.

Theorem 4.1.1 will be used to prove Theorem 4.1.2.

Theorem 4.1.2. *Let $Des(\pi)$ and $Inv(\pi)$ be the number of descents and inversions of $\pi \in S_n$. Then for all real x,*

$$\left|P\left\{\frac{Des - \frac{n-1}{2}}{\sqrt{\frac{n+1}{12}}} \leq x\right\} - \Phi(x)\right| \leq \frac{C}{n^{\frac{1}{2}}}$$

$$\left|P\left\{\frac{Inv - \frac{\binom{n}{2}}{2}}{\sqrt{\frac{n(n-1)(2n+5)}{72}}} \leq x\right\} - \Phi(x)\right| \leq \frac{C}{n^{\frac{1}{2}}}$$

where C is a constant independent of n.

We remark that Theorem 4.1.2 is known by other proof techniques (see [6] for the case of descents and [3] for inversions). We recently learned that there is some overlap with results in [1], which gives bounds for permutation statistics using reversible Markov chains together with Bolthausen's variation of Stein's method.

Section 4.2 shows how, for $W =$ Des or $W =$ Inv, to construct an exchangeable pair (W, W') such that $E^W W' = (1-\frac{2}{n})W$. This step, which is usually the easy part of applying Stein's method, is non-trivial and uses a non-reversible Markov chain equivalent to the "move to front" chain. The only other example in the literature in which exchangeability was not obvious is the paper of Rinott and Rotar [13]. A connection with this work will be mentioned in Section 4.2. Section 4.3 develops bounds for the terms on the right-hand side of Theorem 4.1.1, and indicates why a somewhat weaker version of Theorem 4.1.1 due to Stein can only give $n^{-1/4}$ rates.

We remark that the move to front rule on the symmetric group is a very special case of a theory of random walk on the chambers of real hyperplane arrangements [4]. The corresponding Markov chains are non-reversible and have real eigenvalues. These nonreversible chains have recently been related to a reversible Markov chain on the set of irreducible representations of the symmetric group [9],[10].

4.2. Construction of an exchangeable pair (W, W')

This section constructs W' so that (W, W') is an exchangeable pair with nice properties. In most applications of Stein's method (e.g. the examples in Stein [15]), it is clear how to define W' and exchangeability comes for free. The situation here is more subtle.

This being said, define $W' = W'(\pi)$ as follows. Pick I uniformly at random between 1 and n and define π' as $(I, I+1, \dots, n)\pi$, where $(I, I+1, \dots, n)$ cycles by mapping $I \to I+1 \to \cdots \to n \to I$, and where permutation multiplication is from left to right. For example, suppose that $n = 7$ and $I = 3$. Then the permutation π which in 2-line form is:

$$\begin{array}{ccccccccc} i & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \pi(i) & : & 6 & 4 & 1 & 5 & 3 & 2 & 7 \end{array}$$

is transformed to:

$$\begin{array}{ccccccccc} i & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \pi'(i) & : & 6 & 4 & 5 & 3 & 2 & 7 & 1 \end{array}$$

In other words, one moves the number in position I in the second row of π to the end of this second row. Now define $W'(\pi) = W(\pi')$. Before discussing exchangeability, we prove Lemma 4.2.1, which was the motivation for the definition of W' and shows that one can take $\lambda = \frac{2}{n}$ in Theorem 4.1.1.

Lemma 4.2.1. *$E^W W' = (1 - \frac{2}{n})W$.*

Proof. Letting i be the value of the random variable I, ones sees from the definition of W' that:

$$\begin{aligned} E^{\pi}(W' - W) &= \frac{1}{\sqrt{\mathrm{Var}(X)}} \frac{1}{n} \sum_{i=1}^{n} \sum_{j:j>i} -2M_{\pi(i),\pi(j)} \\ &= \frac{1}{\sqrt{\mathrm{Var}(X)}} \frac{1}{n} \sum_{1 \le i < j \le n} -2M_{\pi(i),\pi(j)} \\ &= -\frac{2}{n} W. \end{aligned}$$

Since $E^{\pi}(W' - W)$ depends on π only through W, the lemma follows. □

Lemma 4.2.2 establishes a condition on $(M_{i,j})$ under which the pair (W, W') is exchangeable. This condition admittedly has limited scope, but as will be seen, holds for the cases of descents and inversions.

Lemma 4.2.2. *Given a subset S of $\{1, \dots, n\}$, for each $i \in S$ define $a_{i,S} = \sum_{j \in S: j>i} M_{i,j}$ and $b_{i,S} = \sum_{j \in S: j<i} M_{j,i}$. Suppose that for all subsets S of $\{1, \dots, n\}$, there is a bijection $\Theta : S \mapsto S$ satisfing the following conditions:*

1. *For each $i \in S$, $a_{i,S} - b_{i,S} = b_{\Theta(i),S} - a_{\Theta(i),S}$.*
2. *For each $i \in S$, there is a bijection $\Phi_i : S - \{i\} \mapsto S - \{\Theta(i)\}$ such that $M_{j,k} = M_{\Phi_i(j),\Phi_i(k)}$ for all $j, k \in S - \{i\}$.*

Then (W, W') is an exchangeable pair of random variables.

Proof. It will be shown that $P\{W=a,W'=b\}=P\{W=b,W'=a\}$. For this we prove the stronger claim that if $T=\{\pi\in S_n:\pi(j)=z_j \ for\ 1\le j\le I-1\}$, then

$$P\{W=a,W'=b|I,\pi\in T\}=P\{W=b,W'=a,|I,\pi\in T\}$$

In other words, assume that the value of I and the images of $\{1,\ldots,I-1\}$ under π are given. Let $S=\{\pi(I),\cdots,\pi(n)\}$ be as in the hypotheses of the lemma. Now define a bijection $\Lambda:T\mapsto T$ as follows:

1. $\Lambda(\pi)(j)=\pi(j)$ for $1\le j\le I-1$
2. $\Lambda(\pi)(I)=\Theta(\pi(I))$
3. $\Lambda(\pi)(j)=\Phi_{\pi(I)}(\pi(j))$ for $I+1\le j\le N$

We only show that $W(\pi)=W(\Lambda(\pi)')$, the argument that $W(\pi')=W(\Lambda(\pi))$ being similar. Since π and $\Lambda(\pi)'$ agree on $1,\ldots,I-1$, it is enough to show that

$$\begin{aligned}&\sum_{I<j\le n}M_{\pi(I),\pi(j)}+\sum_{I<i<j\le n}M_{\pi(i),\pi(j)}\\&\quad=\sum_{I\le i<j<n}M_{\Lambda(\pi)'(i),\Lambda(\pi)'(j)}+\sum_{I\le i<n}M_{\Lambda(\pi)'(i),\Lambda(\pi)'(n)}.\end{aligned}$$

Now observe that

$$\begin{aligned}\sum_{I\le i<j<n}M_{\Lambda(\pi)'(i),\Lambda(\pi)'(j)}&=\sum_{I<i<j\le n}M_{\Lambda(\pi)(i),\Lambda(\pi)(j)}\\&=\sum_{I<i<j\le n}M_{\Phi_{\pi(I)}(\pi(i)),\Phi_{\pi(I)}(\pi(j))}\\&=\sum_{I<i<j\le n}M_{\pi(i),\pi(j)}.\end{aligned}$$

The second equality is from the definition of $\Lambda(\pi)$ and the third equality is from condition 2 in the lemma. Also observe that

$$\begin{aligned}\sum_{I\le i<n}M_{\Lambda(\pi)'(i),\Lambda(\pi)'(n)}&=\sum_{I<j\le n}M_{\Lambda(\pi)(j),\Lambda(\pi)(I)}\\&=\sum_{I<j\le n}M_{\Lambda(\pi)(j),\Theta(\pi(I))}\\&=b_{\Theta(\pi(I)),S}-a_{\Theta(\pi(I)),S}\\&=a_{\pi(I),S}-b_{\pi(I),S}\\&=\sum_{I<j\le n}M_{\pi(I),\pi(j)}.\end{aligned}$$

The third equality holds because $\{\Lambda(\pi)(j):I<j\le n\}=S-\Theta(\pi(I))$. The fourth equality is from condition 1 in the lemma. □

Remarks.

1. Let us illustrate the proof of Lemma 4.2.2 by example for $X(\pi)=2\mathrm{Des}(\pi^{-1})-(n-1)$. Recall that here $M_{i,j}=-1$ if $j=i+1$, $M_{i,j}=1$ if $j=i-1$, and $M_{i,j}=0$ otherwise. Suppose that $I=3$ and $\pi(1)=6,\pi(2)=4$. Thus

$T = \{\pi \in S_n : \pi(1) = 6, \pi(2) = 4\}$. Note that $S = \{1,2,3,5,7\}$, because these are the images of $\pi(j)$ for $j \geq I = 3$. One observes that the bijection $\Theta : S \mapsto S$ defined by $\Theta(1) = 3$, $\Theta(2) = 2$, $\Theta(3) = 1$, $\Theta(5) = 5$, $\Theta(7) = 7$ satisfies condition 1 of Lemma 4.2.2 (in general, one defines Θ by reversing within each group of consecutive numbers in S). For each $i \in S$ it is also necessary to define bijections Φ_i such that condition 2 of Lemma 4.2.2 holds. This can be done by pairing the elements of $S - \{i\}$ and $S - \{\Theta(i)\}$ so as to preserve their relative order. For instance, $\Phi_1 : \{2,3,5,7\} \mapsto \{1,2,5,7\}$ is defined by $\Phi_1(2) = 1, \Phi_1(3) = 2, \Phi_1(5) = 5, \Phi_1(7) = 7$.

These choices determine the bijection $\Lambda : T \to T$ constructed in Lemma 4.2.2. For example,

i	:	1	2	3	4	5	6	7
$\pi(i)$	:	6	4	1	5	3	2	7

i	1	2	3	4	5	6	7
$\Lambda(\pi)(i)$	6	4	3	5	2	1	7

i	:	1	2	3	4	5	6	7
$\pi'(i)$	:	6	4	5	3	2	7	1

i	1	2	3	4	5	6	7
$\Lambda(\pi')(i)$	6	4	5	2	1	7	3

One checks that $X(\pi) = X(\Lambda(\pi')) = 0$ and $X(\pi') = X(\Lambda(\pi)) = 2$.

2. Let us illustrate the proof of Lemma 4.2.2 by example for $X(\pi) = 2\mathrm{Inv}(\pi^{-1}) - \binom{n}{2}$. Here $M_{i,j} = -1$ if $i < j$, $M_{i,j} = +1$ if $i > j$, and $M_{i,i} = 0$. As for the case of descents, suppose that $I = 3$ and $\pi(1) = 6, \pi(2) = 4$. Then $T = \{\pi \in S_n : \pi(1) = 6, \pi(2) = 4\}$ and $S = \{1,2,3,5,7\}$. The bijection $\Theta : S \mapsto S$ must be defined differently from the descent case so that condition 1 of Lemma 4.2.2 holds. It is easy to see that reversing the elements of S works. Thus $\Theta(1) = 7$, $\Theta(2) = 5$, $\Theta(3) = 3$, $\Theta(5) = 2$, and $\Theta(7) = 1$. Defining the maps Φ_i as in the descent case, condition 2 of Lemma 4.2.2 holds.

   These choices determine the bijection $\Lambda : T \to T$ constructed in Lemma 4.2.2. For example,

i	:	1	2	3	4	5	6	7
$\pi(i)$	:	6	4	1	5	3	2	7

i	1	2	3	4	5	6	7
$\Lambda(\pi)(i)$	6	4	7	3	2	1	5

i	:	1	2	3	4	5	6	7
$\pi'(i)$	:	6	4	5	3	2	7	1

i	1	2	3	4	5	6	7
$\Lambda(\pi')(i)$	6	4	3	2	1	5	7

   One checks that $X(\pi) = X(\Lambda(\pi')) = 1$ and $X(\pi') = X(\Lambda(\pi)) = 9$.

3. The above examples show that the pair (W, W') is exchangeable for descents and inversions. An interesting problem is to classify the matrices $(M_{i,j})$ such that the pair (W, W') is exchangeable. It would also be useful to construct exchangeable pairs (W, W') for other Coxeter groups.

4. Lemma 1.1 of [13] states the following. *Suppose that* $\{T^t\}$ *is a stationary, nonnegative, integer valued process satisfying* $T^{t+1} - T^t = +1, 0$ *or* -1. *Then* (T^t, T^{t+1}) *is an exchangeable pair.*

   For the case of descents, this gives an alternate proof that W, W' as we have defined them are an exchangeable pair, even though the underlying chain on permutations is not reversible. To see this, let R^0 be a uniformly distributed element of S_n; then given R^i, move to R^{i+1} according to the move random to end rule defined in the beginning of this section. This process is stationary. Defining T^t to be the number of descents of R^t, one sees that the conditions of the lemma hold.

It is interesting to note that Lemma 1.1 of [13] was applied there to study W equal to the number of ones in a random pick from the stationary distribution of the antivoter model. The antivoter chain is not reversible, but their lemma implies that if W' is the number of ones after a step from the antivoter chain, then (W, W') is an exchangeable pair.

4.3. Bounding the error terms

This section bounds the error terms on the right hand side of Theorem 4.1.1.

We start by computing the mean and variance of X and establishing a nice property of the pair (W, W'). For this it is helpful to define $A_i = \sum_{j>i} M_{i,j}$ and $B_i = \sum_{h<i} M_{h,i}$.

Lemma 4.3.1. *$E(X) = 0$ and $Var(X) = \frac{\sum_{i<j}(M_{i,j})^2 + \sum_{i=1}^n (A_i - B_i)^2}{3}$.*

Proof. Observe that the random variable X on S_n can be written as a sum of random variables $X_{i,j}$ on S_n. Defining a random variable $X_{i,j}$ on S_n by

$$X_{i,j}(\pi) = \begin{cases} M_{i,j} & \text{if } \pi^{-1}(i) < \pi^{-1}(j) \\ M_{j,i} & \text{if } \pi^{-1}(j) < \pi^{-1}(i) \end{cases}$$

one has that:

$$X(\pi) = \sum_{i<j} M_{\pi(i),\pi(j)} = \sum_{\pi^{-1}(i)<\pi^{-1}(j)} M_{i,j} = \sum_{i<j} X_{i,j}(\pi).$$

The mean of X is 0 since each $X_{i,j}$ has mean 0 and expectation is linear.

The variance of X is equal to $E[(\sum_{i<j} X_{i,j}(\pi))^2]$. The terms $E(X_{i,j}^2)$ contribute $(M_{i,j})^2$ each and thus $\sum_{i<j}(M_{i,j})^2$ in total. The terms $E(2X_{i,j}X_{k,l})$ vanish if i, j, k, l are distinct, by independence. Now consider what happens when two of these four indices are equal. Terms of the form $2E(X_{i,j}X_{i,l})$ contribute $\frac{2}{3}M_{i,j}M_{i,l}$ each. The sum of all such terms can be rewritten as $\frac{1}{3}[\sum_i A_i^2 - \sum_{i<j}(M_{i,j})^2]$. Similarly, terms of the form $2E(X_{i,l}X_{k,l})$ contribute $\frac{1}{3}[\sum_i B_i^2 - \sum_{i<j}(M_{i,j})^2]$. Finally, terms of the form $2E(X_{i,j}X_{j,k})$ contribute $-\frac{2}{3}M_{i,j}M_{j,k}$ each, and hence a total of $-\frac{2}{3}\sum_i A_i B_i$. The lemma follows. □

As a consequence of Lemma 4.3.1, one recovers the known facts that for a random permutation on n symbols, $\mathrm{Var}(\mathrm{Des}(\pi)) = \frac{n+1}{12}$ and $\mathrm{Var}(\mathrm{Inv}(\pi)) = \frac{n(n-1)(2n+5)}{72}$. Note that Lemma 4.3.1 has written $\mathrm{Var}(X)$ as a sum of positive quantities.

Lemma 4.3.2. *$E(W' - W)^2 = \frac{4}{n}$*

Proof.

$$\begin{aligned} E\big(W' - W\big)^2 &= E\big(E^W\big(W' - W\big)^2\big) \\ &= E\big(E^W\big(\big(W'\big)^2 + W^2 - 2WW'\big)\big) \\ &= E\big(\big(W'\big)^2 + E\big(W^2\big) - 2WE^W\big(W'\big)\big) \end{aligned}$$

$$
\begin{aligned}
&= 2\mathrm{Var}(W) - E\big(2WE^W(W')\big) \\
&= \frac{4}{n}\mathrm{Var}(W) \\
&= \frac{4}{n}.
\end{aligned}
$$

The fourth equality used the fact that W and W' have the same distribution. The fifth equality used Lemma 4.2.1. □

Lemma 4.3.3 establishes a well known inequality. For completeness, we include a proof.

Lemma 4.3.3. $E[E^W(W'-W)^2]^2 \le E[E^\pi(W'-W)^2]^2$.

Proof. Jensen's inequality says that if g is a convex function, and Z a random variable, then $g(E(Z)) \le E(g(Z))$. There is also a conditional version of Jensen's inequality (Section 4.1 of Durrett [7]) which says that if F is any σ subalgebra of B, then

$$E\big(g\big(E(Z|F)\big)\big) \le E\big(g(Z)\big).$$

The lemma follows by applying this inequality to the case $g(t) = t^2$, $Z = E^\pi(W'-W)^2$, B is all subsets of S_n, and F is the σ subalgebra of B generated by the level sets of W. □

Now we prove Theorem 4.1.2.

Proof of Theorem 4.1.2. We will apply Theorem 4.1.1. Note that the move random to end rule changes the number of descents by at most one. Hence the corresponding pair (W,W') satisfies $|W'-W| \le \frac{2}{\sqrt{\mathrm{Var}(X)}}$. Similarly the move random to end rule changes the number of inversions by at most $n-1$. Hence the corresponding pair (W,W') satisfies $|W'-W| \le \frac{2(n-1)}{\sqrt{\mathrm{Var}(X)}}$. Thus in both cases $|W'-W|$ is at most $An^{-1/2}$ for an absolute constant A. Also note by Lemma 4.2.1 that $E^W(W') = (1-\lambda)W$ with $\lambda = \frac{2}{n}$.

Thus by Theorem 4.1.1 the result will follow if it can be shown that $\mathrm{Var}(E^W(W'-W)^2) \le \frac{B}{n^3}$. Lemma 4.3.3 implies that $\mathrm{Var}(E^W(W'-W)^2) \le \mathrm{Var}(E^\pi(W'-W)^2)$. Hence we show that $\mathrm{Var}(E^\pi(W'-W)^2) \le \frac{B}{n^3}$.

Observe that

$$
\begin{aligned}
E^\pi\big(W'-W\big)^2 &= \frac{1}{\mathrm{Var}(X)}\frac{4}{n}\sum_{i=1}^n\left(\sum_{j>i} -M_{\pi(i),\pi(j)}\right)^2 \\
&= \frac{1}{\mathrm{Var}(X)}\frac{4}{n}\left(\sum_{i=1}^n\sum_{j>i}\big(M_{\pi(i),\pi(j)}\big)^2 + 2\sum_{i=1}^n\sum_{i<j_1<j_2\le n} M_{\pi(i),\pi(j_1)}M_{\pi(i),\pi(j_2)}\right).
\end{aligned}
$$

Since $\sum_{i=1}^n\sum_{j>i}(M_{\pi(i),\pi(j)})^2$ is independent of π, it follows that

$$
\begin{aligned}
\mathrm{Var}\big(E^\pi\big(W'-W\big)^2\big) = \frac{64}{\mathrm{Var}(X)^2n^2}\Bigg[&\sum_{1\le i<j_1<j_2\le n}\mathrm{Var}\big(M_{\pi(i),\pi(j_1)}M_{\pi(i),\pi(j_2)}\big) \\
&+ \sum_{\substack{i<j_1<j_2,k<l_1<l_2\\(i,j_1,j_2)\neq(k,l_1,l_2)}} Cov(M_{\pi(i),\pi(j_1)}M_{\pi(i),\pi(j_2)}, M_{\pi(k),\pi(l_1)}M_{\pi(k),\pi(l_2)})\Bigg]
\end{aligned}
$$

Let us analyze this bound for the case of descents (i.e. $M_{i,j} = -1$ if $j = i+1$, $M_{i,j} = 1$ if $j = i-1$, and $M_{i,j} = 0$ otherwise). We first study the summands and then divide by $\mathrm{Var}(X)^2 n^2$. The first summand has $O(n^3)$ terms, each contributing $O(n^{-2})$; hence it is $O(n)$. The covariance terms are also $O(n)$. To see this, first note that the covariance vanishes if $\{i, j_1, j_2\} \cap \{k, l_1, l_2\} = \emptyset$, so such terms can be ignored. Suppose that $i \neq k$. Then there are $O(n^5)$ terms each contributing $O(n^{-4})$. If $i = k$ there are subcases to consider based on which (if any) of elements of $\{j_1, j_2\}$ are equal to elements of $\{l_1, l_2\}$. It is straightforward to see that in all cases the contribution of the covariance term is $O(n)$. Since $\mathrm{Var}(X)$ is $\frac{n+1}{12}$, it follows as desired that $\mathrm{Var}(E^\pi(W'-W)^2) \leq \frac{B}{n^3}$.

The case of inversions is similar. The variance terms contribute at most $O(n^3)$ and the covariance terms at most order $O(n^5)$. Thus

$$\mathrm{Var}\big(E^\pi\big(W'-W\big)^2\big) \leq \frac{B_0 n^5}{\mathrm{Var}(X)^2 n^2} \leq \frac{B}{n^3}$$

where B_0, B are universal constants. □

To conclude the paper, we comment on the following result of Stein [15].

Theorem 4.3.1 (Stein). *Let W, W' be an exchangeable pair of real random variables such that $E^W W' = (1-\lambda)W$ with $0 < \lambda < 1$. Then for all real x,*

$$\big|P\{W \leq x\} - \Phi(x)\big| \leq 2\sqrt{E\left[1 - \frac{1}{2\lambda}E^W\big(W'-W\big)^2\right]^2} + (2\pi)^{-\frac{1}{4}}\sqrt{\frac{1}{\lambda}E\big|W'-W\big|^3}$$

where Φ is the standard normal distribution.

Applied to our exchangeable pair this would only yield bounds of order $n^{-1/4}$, since by Jensen's inequality $E|W'-W|^3 \geq (E(W'-W)^2)^{3/2} = \left(\frac{4}{n}\right)^{3/2}$.

Acknowledgments

The author thanks Persi Diaconis for introducing the author to Stein's method and urging him to find applications of it to permutation statistics. We also thank Y. Rinott for useful feedback on the 1997 version of this paper. This research was done under the support of the National Defense Science and Engineering Graduate Fellowship (grant no. DAAH04-93-G-0270) and the Alfred P. Sloan Foundation Dissertation Fellowship.

References

[1] Bai, Z., Chao, C. C., Liang, C., and Zhao, L., (1997). Error bounds in a central limit theorem of doubly indexed permutation matrices, *Annals of Stat.* **25**, 2210–2227.

MR1474091

[2] Bender, E., (1973). Central and local limit theorems applied to asymptotic enumeration, *J. Combin. Theory Ser. A* **15**, 91–111. MR375433

[3] Bickel, P. (1974). Edgeworth expansions in nonparametric statistics, *Annals of Stat.* **2**, 1–20.

MR350952

[4] Bidigare, P., Hanlon, P., and Rockmore, D., (1999). A combinatorial description of the spectrum for the Tsetlin library and its generalizations to hyperplane arrangements, *Duke Math J.* **99**, 135–174. MR1700744

[5] Diaconis, P., (1988). *Group representations in probability and statistics.* Institute of Mathematical Statistics Lecture Notes, Volume 11. MR964069

[6] Diaconis, P. and Pitman, J., (1991). Unpublished notes on descents.

[7] Durrett, R., (1991). *Probability: Theory and examples*, Brooks/Cole Publishing Company. MR1068527

[8] Foata, D. and Schutzenberger, M. P., (1970). Theorie geometriques des polynomes Euleriens. *Springer Lecture Notes in Math. 138.* MR272642

[9] Fulman, J., Stein's method and Plancherel measure of the symmetric group, to appear in *Trans. Amer. Math. Soc.*

[10] Fulman, J., Card shuffling and the decomposition of tensor products, to appear in *Pacific J. Math.*

[11] Knuth, D., (1973). *The art of computer programming. Volume 3. Sorting and searching.* Addison-Wesley Publishing Co.

[12] Pitman, J., (1997). Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combin. Theory Ser. A* **77**, 279–303.
MR1429082

[13] Rinott, Y. and Rotar, V., (1997). On coupling constructions with rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics, *Annals Appl. Probab.* **7**, 1080–1105. MR1484798

[14] Stanley, R., (1986). *Enumerative combinatorics, Volume 1.* Wadsworth and Brooks/Cole Mathematical Series.
MR847717

[15] Stein, C., (1986). *Approximate computation of expectations.* Institute of Mathematical Statistics Lecture Notes, Volume 7.
MR882007

[16] Tanny, S., (1973). A probabilistic interpretation of Eulerian numbers. *Duke Math J.* **40**, 717–722. MR340045