Asymptotic Normality of Statistics on Permutation Tableaux

Paweł Hitczenko and Svante Janson

Abstract. In this paper we use a probabilistic approach to derive the expressions for the characteristic functions of basic statistics defined on permutation tableaux. Since our expressions are exact, we can identify the distributions of basic statistics (like the number of unrestricted rows, the number of rows, and the number of 1s in the first row) exactly. In all three cases the distributions are known to be asymptotically normal after a suitable normalization. We also establish the asymptotic normality of the number of superfluous 1s. The latter results relies on a bijection between permutation tableaux and permutations and on a rather general sufficient condition for the central limit theorem for the sums of random variables in terms of dependency graph of the summands.

1. Introduction

Permutation tableaux are relatively new objects that are in bijection with permutations [**Bu, CoN, SW**]. They were introduced in the context of enumeration of the totally positive Grassmannian cells [**P, W**]. More recently, permutation tableaux generated additional research activity when they have been connected in [**Co, CoW, CoW1**] to a particle model in statistical physics called the Partially ASymmetric Exclusion Process (PASEP); see [**BrE, BrCPRE, DDM, DEHP, DS, S, USW**] for more information on PASEP.

A *permutation tableau* [**SW**] is a Ferrers diagram of a partition of a positive integer into non–negative parts whose boxes are filled with 0s and 1s according to the following rules:

(1) Each column of the diagram contains at least one 1.
(2) There is no 0 which has a 1 above it in the same column *and* a 1 to its left in the same row.

An example is given in Figure 1.

The size parameter of a permutation tableau is its *length* defined as the number of rows plus the number of columns of the tableau. For example, the tableau in Figure 1 has 6 rows and 7 columns so its length is 13.

2000 *Mathematics Subject Classification.* Primary 05E10, 60F05; Secondary 05A15, 60E10.
Key words and phrases. permutation tableau, central limit theorem.
The first author was supported in part by the NSA Grant #H98230-09-1-0062.
Research of the second author was partially done while visiting Institut Mittag-Leffler, Djursholm, Sweden.
Date: March 16, 2009.

0	0	1	0	0	1	1
0	0	1	0	1		
0	1	1	1	1		
0	0	0				
1						

FIGURE 1. Example of a permutation tableau

Different statistics on permutation tableaux were defined in [**CoW1, SW**]. We recall that a 0 in a permutation tableau is *restricted* if there is a 1 above it in the same column. A row is *unrestricted* if it does not contain a restricted 0. A 1 is *superfluous* if it has a 1 above itself in the same column. We will be interested in the number of unrestricted rows, the number of superfluous 1s, as well as the number of 1s in the top row, and the number of rows. For example the tableau in Figure 1 has three superfluous 1s, six rows, three 1s in the top row, and five unrestricted rows.

Most of the past research on statistics of permutation tableaux was based on bijections between permutation tableaux of length n and permutations of $[n] := \{1, \dots, n\}$ and using known properties of permutations. This was not always easy as sometimes it is not that easy to see into what parameter a given statistic is mapped by a bijection. In [**CoH**] a direct approach based on a probabilistic consideration was proposed. It enabled the authors to compute the expected values of these statistics in a simple and unified way. In this work we go one step further and we will concentrate on the limiting distributions of (properly normalized) statistics mentioned above. In most cases we will accomplish it by computing the probability generating functions of the quantities in question. As a matter of fact, since the approach proposed in [**CoH**] allows for exact (and not only asymptotic) computation, in most cases we will be able to identify the distribution of a given statistics exactly and not only asymptotically. As a consequence, we can see, in particular, that:

- the number of unrestricted rows in a random permutation tableaux of length n has the same distribution as the number of cycles, or the number of records, in a random permutation of $[n]$ (a record in a permutation $\sigma = (\sigma_i)$ is any σ_i such that $\sigma_i > \sigma_j$ for $j < i$);
- the number of rows has the same distribution as the number of descents in a random permutation counted by Eulerian numbers (a descent in (σ_i) is any pair (σ_i, σ_{i+1}) such that $\sigma_i > \sigma_{i+1}$);
- the number of 1s in the first row is, in distribution, one less than the number of unrestricted rows.

The first two results follow from the analysis of known bijections between permutations and permutation tableaux (see [**CoN, SW**]), while the third is a consequence of an involution on permutation tableaux presented in [**CoW1**]. However, neither the involution from [**CoW1**] nor any of the bijections from [**CoN, SW**] are very straightforward. In addition, the first two results given above required two different bijections in [**SW**]. So, while these bijections carry more information, they turn out to be rather cumbersome to work with. Looked at from this perspective, our approach provides a streamlined and unified derivation of the above results.

As for the number of superfluous 1s, the situation is a bit more complicated. Although we do derive its probability generating function, we do not deduce directly from this the asymptotic normality of its distribution. Instead, in this case we will rely on the fact that one of the two bijections between permutation tableaux and permutations described in [**SW**] sends the number of superfluous 1s to the number of generalized patterns 31−2 and we will prove the central limit theorem for the number of such patterns in random permutation of $[n]$ (an occurrence of a generalized pattern 31−2 in (σ_i) is any pair $(i,j), 1 < i < j$ such that $\sigma_{i-1} > \sigma_j > \sigma_i$). Our proof is based on a rather general sufficient condition for the central limit theorem developed in [**J**] and [**JŁR**, Section 6.1]. Bóna [**Bo**] has used the same method to show similar results on asymptotic normality of general permutation patterns, but as far as we know the central limit theorem for the number of a generalized pattern such as 31−2 is new. Furthermore, the number of permutation tableaux of length n with k superfluous 1s is equal to the number of permutations of $[n]$ with k crossings (a crossing [**Co**] in a permutation (σ_i) is a pair (i,j) such that $i < j \le \sigma_i < \sigma_j$ or $i > j > \sigma_i > \sigma_j$). It follows, therefore, that the number of crossings in a random permutation of $[n]$ is asymptotically normal.

2. Basic facts

Let $\mathcal{T}_n$ be the set of all permutation tableaux of length n. We denote the uniform probability measure on $\mathcal{T}_n$ by P_n and E_n will denote the expectation with respect to P_n. We denote by R_n, C_n, U_n, F_n, and S_n the random variables representing the numbers of rows, columns, unrestricted rows, 1s in the top row, and superfluous 1s, respectively, in a random tableau of length n. (We may allow $n = 0$; there is a single empty permutation tableau of length 0, and $U_0 = F_0 = R_0 = C_0 = S_0 = 0$. It is sometimes convenient to start inductions with $n = 0$, but for simplicity we often treat $n \ge 1$ only.)

If $\mathcal{F}$ is a σ–algebra then $\mathsf{E}(\,\cdot\,|\mathcal{F})$ denotes the conditional expectation given $\mathcal{F}$. If X is non-negative integer valued, for example one of the permutation tableaux statistics just defined, we let $g_X(z) = \mathsf{E}z^X$ be its probability generating function (in general defined at least for $|z| \le 1$; for the variables consider here, z can be any complex number). We will often omit the subscript X.

The arguments in [**CoH**] were based on a construction of tableaux of size k from tableaux of size $k-1$ by extending the length of the latter from its south-west (SW) corner either to the south (creating a new empty row) or to the west (creating and then filling a new column). Note that each permutation tableau in $\mathcal{T}_k$ is an extension of a unique permutation tableau in $\mathcal{T}_{k-1}$. We refer the reader to [**CoH**, Section 2] for details. We let M_k indicate the direction of the move at the kth step (i.e. when the length is increased from $k-1$ to k). We refer to M_k as the kth move and will write $M_k = \mathsf{S}$ or $M_k = \mathsf{W}$ to indicate its direction.

The following simple observations were crucial for the arguments in [**CoH**] and are crucial here as well (see [**CoH**] for some details that are omitted here).

- When extending a tableau by a W move, the new column has to be filled with 0s and 1s; each restricted row must be filled with 0, but the unrestricted rows can be filled arbitrarily except that there must be at least one 1. Hence, a tableau with U unrestricted rows can be extended by a W move in $2^U - 1$ ways; since there always is a unique S move, the total number of extensions is 2^U.

- We use T_n to denote a generic element of $\mathcal{T}_n$, and use T_{n-1} to denote the corresponding element of $\mathcal{T}_{n-1}$ (i.e., such that T_n is an extension of T_{n-1}); we let $U_n = U_n(T_n)$ be the number of unrestricted rows in T_n and $U_{n-1} = U_{n-1}(T_{n-1})$ the number of unrestricted rows in T_{n-1}, and similarly for the other variables that we study. In this way, P_n induces a probability measure (also denoted P_n) on $\mathcal{T}_{n-1}$, namely each element of $\mathcal{T}_{n-1}$ is assigned a measure that is proportional to the number of tableaux from $\mathcal{T}_n$ it generates when its length is increased from $n-1$ to n. Note that this differs from the uniform distribution P_{n-1} on $\mathcal{T}_{n-1}$. Since $|\mathcal{T}_n| = n!$ and a tableau T from $\mathcal{T}_{n-1}$ generates $2^{U_{n-1}(T)}$ tableaux of length n the relationship between these two measures is easy to find. One way to state this relationship is that if X is any random variable on $\mathcal{T}_{n-1}$ then
$$\mathsf{E}_n X = \frac{1}{n}\mathsf{E}_{n-1}\left(2^{U_{n-1}}X\right), \tag{2.1}$$
where integration on the left is with respect to the measure induced on $\mathcal{T}_{n-1}$ by P_n and integration on the right is with respect to P_{n-1}. Equivalently, if we let $\mathcal{F}_{n-1}$ be the σ-algebra on $\mathcal{T}_n$ generated by the mapping $T_n \mapsto T_{n-1}$, then $d\mathsf{P}_n/d\mathsf{P}_{n-1} = 2^{U_{n-1}}/n$ on $\mathcal{F}_{n-1}$.
- The sequence of the distributions of the number of unrestricted rows is given by $U_0 \equiv 0$, $U_1 \equiv 1$ and for $n \ge 1$ the conditional distribution of U_n given $\mathcal{F}_{n-1}$ (i.e., given T_{n-1}) is (under P_n)
$$\mathcal{L}(U_n \mid \mathcal{F}_{n-1}) = 1 + \mathrm{Bin}(U_{n-1}), \tag{2.2}$$
where $\mathrm{Bin}(m)$ denotes a binomial random variable with parameters m and $1/2$.

To illustrate how these facts are put together to work consider the number of unrestricted rows, U_n. For its probability generating function we have:
$$\begin{aligned} g_{U_n}(z) &= \mathsf{E}_n z^{U_n} = \mathsf{E}_n\mathsf{E}_n\left(z^{U_n} \mid \mathcal{F}_{n-1}\right) = \mathsf{E}_n\mathsf{E}_n\left(z^{1+\mathrm{Bin}(U_{n-1})} \mid \mathcal{F}_{n-1}\right) \\ &= z\mathsf{E}_n\Big(\frac{z+1}{2}\Big)^{U_{n-1}} = \frac{z}{n}\mathsf{E}_{n-1}\left(2^{U_{n-1}}\Big(\frac{z+1}{2}\Big)^{U_{n-1}}\right) \\ &= \frac{z}{n}\mathsf{E}_{n-1}\left(z+1\right)^{U_{n-1}}, \end{aligned} \tag{2.3}$$
where we have used (in that order) conditioning, (2.2), the obvious fact that $g_{\mathrm{Bin}(m)}(z) = \mathsf{E}z^{\mathrm{Bin}(m)} = \left(\frac{z+1}{2}\right)^m$, and (2.1). It follows by induction, and $g_{U_1}(z) = z$, that
$$g_{U_n}(z) = \frac{\Gamma(z+n)}{\Gamma(z)n!} = \binom{z+n}{n} = \prod_{j=0}^{n-1}\frac{z+j}{j+1} = \prod_{k=1}^{n}\Big(1 - \frac{1}{k} + \frac{z}{k}\Big). \tag{2.4}$$

The factor on the right-hand side is the probability generating function of a random variable that is 1 with probability $1/k$ and 0 with probability $1 - 1/k$. Since the product of probability generating functions corresponds to summing independent random variables, we obtain the following statement

THEOREM 2.1. *The number of unrestricted rows U_n is distributed like*
$$U_n \overset{\mathrm{d}}{=} \sum_{k=1}^{n} J_k,$$

where $J_1, J_2, \dots,$ are independent indicators with $\mathsf{P}(J_k = 1) = 1/k$. In particular, if $h_n = \sum_{k=1}^n k^{-1}$ and $h_n^{(2)} = \sum_{k=1}^n k^{-2}$ is the nth harmonic and generalized harmonic number, respectively, then

$$\mathsf{E}_n U_n = \sum_{k=1}^n \mathsf{P}(J_k) = h_n, \quad \mathsf{var}_n(U_n) = \sum_{k=1}^n \mathsf{var}(J_k) = \sum_{k=1}^n \frac{1}{k}\left(1 - \frac{1}{k}\right) = h_n - h_n^{(2)},$$

and

$$\frac{U_n - \ln n}{\sqrt{\ln n}} \xrightarrow{\mathrm{d}} N(0,1),$$

where $N(0,1)$ denotes a standard normal random variable.

Remark 2.2. It is seen from the above statement that the distribution of U_n coincides with the distribution of the number of cycles in a random permutation of $[n]$ (see for example [**F**, Chapter X, Section 6(b)] or [**R**, Chapter 4, Section 3]), or, equivalently, the number of RL minima in a random permutation (which is known to be equidistributed with the number of cycles). Indeed, a bijection between permutation tableaux and permutations described in [**CoN**] maps unrestricted rows onto RL minima in the corresponding permutation.

Remark 2.3. More generally, we may define on $\mathcal{T}_n$ permutation tableaux $T_k \in \mathcal{T}_k$ for every $k \le n$, such that T_{k+1} is an extension of T_k, and let $\mathcal{F}_k$ be the σ-algebra generated by T_k. It can be seen by induction, similar to (2.3), that a permutation tableau T_k of length k can be extended to $(n-k)!\,(n-k+1)^{U_k(T_k)}$ permutation tableaux in $\mathcal{T}_n$, and hence $d\mathsf{P}_n/d\mathsf{P}_k = (n-k+1)^{U_k}/\binom{n}{k}$ on $\mathcal{F}_k$, $1 \le k \le n$. Furthermore, (2.4) can be generalized to

$$\mathsf{E}_n(z^{U_n} \mid \mathcal{F}_k) = \frac{\Gamma(z+n-k)k!}{\Gamma(z)n!}(z+n-k)^{U_k}.$$

We can further study the sequence $(U_k)_{k=1}^n$ as a stochastic process under P_n; similar calculations show that this is an inhomogeneous Markov process with transitions given by $\mathcal{L}(U_{k+1} \mid \mathcal{F}_k) = 1 + \mathrm{Bin}(U_k, \frac{n-k}{n-k+1})$, which generalizes (2.2).

3. The number of 1s in the first row

Recall that F_n denotes the number of 1s in the top row of a random permutation tableau of length n. The following fact is a consequence of an involution on permutation tableaux given in [**CoW1**]. Since this involution is not too easy to describe we provide a more direct justification of (3.1) based on the arguments given in the preceding section.

Theorem 3.1. *For every $n \ge 1$, we have*

$$F_n \overset{\mathrm{d}}{=} U_n - 1. \tag{3.1}$$

In particular,

$$\mathsf{E}_n F_n = h_n - 1, \qquad \mathsf{var}_n(F_n) = h_n - h_n^{(2)},$$

and

$$\frac{F_n - \ln n}{\sqrt{\ln n}} \xrightarrow{\mathrm{d}} N(0,1).$$

To establish (3.1) it will be convenient to prove some auxiliary lemmas. Define G_k to be the position (counting only unrestricted rows) of the topmost 1 on the kth move, provided that this move is W. Then

$$F_n = \sum_{k=1}^{n} I_{G_k=1}.$$

Note that G_k is undefined if the kth move is S. While it is inconsequential in this section as we will be interested in the event $\{G_k = 1\}$, for the purpose of the subsequent sections it is convenient to set $G_k = U_{k-1} + 1$ if the kth move is S. We have

LEMMA 3.2. *For all $k \geq 2$,*

$$\mathsf{P}_k(G_k = 1 \mid \mathcal{F}_{k-1}) = \frac{1}{2}.$$

PROOF. A given tableau $T \in \mathcal{T}_{k-1}$ has $2^{U_{k-1}(T)}$ extensions. The number of extensions by a W move with a 1 in the topmost row (which always is unrestricted) is $2^{U_{k-1}(T)-1}$ since the other $U_{k-1}(T) - 1$ unrestricted rows can be filled arbitrarily. Hence the probability that $G_k = 1$ in a random extension of T is $2^{U_{k-1}(T)-1}/2^{U_{k-1}(T)} = 1/2$. □

LEMMA 3.3. *For any complex numbers z, w and every $k \geq 2$,*

$$\mathsf{E}_k(z^{I_{G_k}} w^{U_k} \mid \mathcal{F}_{k-1}) = w\frac{z+w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}}.$$

PROOF. Let $T \in \mathcal{T}_{k-1}$. If T is extended with $G_k = 1$, i.e., by a W move with 1 added in the topmost row, then we may add 0 or 1 arbitrarily to all other unrestricted rows, and the rows with 1 added remain unrestricted. Hence, conditionally given $G_k = 1$, the extension has $1 + \mathrm{Bin}(U_{k-1}(T) - 1)$ unrestricted rows. Recall from (2.2) that without further conditioning, the number of unrestricted rows is $1 + \mathrm{Bin}(U_{k-1}(T))$. Consequently, using also Lemma 3.2,

$$\begin{aligned}
\mathsf{E}_k(z^{I_{G_k}} w^{U_k} \mid \mathcal{F}_{k-1}) &= z\mathsf{E}_k(w^{U_k} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{U_k} I_{G_k\neq 1} \mid \mathcal{F}_{k-1}) \\
&= (z-1)\mathsf{E}_k(w^{U_k} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{U_k} \mid \mathcal{F}_{k-1}) \\
&= (z-1)\mathsf{E}_k(w^{1+\mathrm{Bin}(U_{k-1}-1)} I_{G_k=1} \mid \mathcal{F}_{k-1}) + \mathsf{E}_k(w^{1+\mathrm{Bin}(U_{k-1})} \mid \mathcal{F}_{k-1}) \\
&= (z-1)w\left(\frac{w+1}{2}\right)^{U_{k-1}-1} \mathsf{P}_k(G_k = 1 \mid \mathcal{F}_{k-1}) + w\left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= (z-1)\frac{w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}} + w\left(\frac{w+1}{2}\right)^{U_{k-1}} \\
&= w\frac{z+w}{w+1}\left(\frac{w+1}{2}\right)^{U_{k-1}}.
\end{aligned}$$
□

LEMMA 3.4. *The joint probability generating function of F_n and U_n is given by*

$$\mathsf{E}_n\big(z^{F_n} w^{U_n}\big) = w\frac{\Gamma(z+w+n-1)}{\Gamma(z+w)n!} = w\prod_{k=2}^{n}\frac{z+w+k-2}{k}.$$

Proof. By Lemma 3.3 and (2.1),

$$\begin{aligned}\mathsf{E}_n\bigl(z^{F_n}w^{U_n}\bigr) &= \mathsf{E}_n\mathsf{E}_n\bigl(z^{F_{n-1}+I_{G_n=1}}w^{U_n} \mid \mathcal{F}_{n-1}\bigr)\\ &= \mathsf{E}_n\left(z^{F_{n-1}}w\frac{z+w}{w+1}\Bigl(\frac{w+1}{2}\Bigr)^{U_{n-1}}\right)\\ &= w\frac{z+w}{w+1}\frac{1}{n}\mathsf{E}_{n-1}\left(z^{F_{n-1}}(w+1)^{U_{n-1}}\right),\end{aligned}$$

and the formula follows by induction. □

We now can complete the proof of Theorem 3.1.

Proof of Theorem 3.1. Taking $w = 1$ in Lemma 3.4 we obtain

$$\mathsf{E}_n z^{F_n} = \frac{\Gamma(z+n)}{\Gamma(z+1)n!},$$

which equals $\mathsf{E}_n z^{U_n-1}$ by (2.4). □

Note that we recover (2.4) by taking $z = 1$ in Lemma 3.4. We can also describe the joint distribution of F_n and U_n.

Theorem 3.5. *For every $n \geq 1$, the joint distribution of F_n and U_n is given by*

$$(F_n, U_n) \stackrel{\mathrm{d}}{=} \sum_{k=1}^{n}(J_k, I_k),$$

where the random vectors (J_k, I_k) are independent and $J_1 = 0$, $I_1 = 1$, and, for $k \geq 2$, $\mathsf{P}(I_k = 1,\ J_k = 0) = \mathsf{P}(I_k = 0,\ J_k = 1) = 1/k$, $\mathsf{P}(I_k = 0,\ J_k = 0) = 1 - 2/k$.

Proof. We have $\mathsf{E}(z^{J_k}w^{I_k}) = (k-2+z+w)/k$ if $k \geq 2$ and w if $k = 1$, and thus the joint probability generating function of the right-hand side equals the product in Lemma 3.4. □

Corollary 3.6. *The covariance of F_n and U_n is*

$$\mathsf{cov}_n(F_n, U_n) = \mathsf{E}_n(F_nU_n) - \mathsf{E}_nF_n\mathsf{E}_nU_n = -(h_n^{(2)} - 1).$$

Proof. We have $\mathsf{cov}(J_k, I_k) = -\mathsf{E}J_k\mathsf{E}I_k = -k^{-2}$ for $k \geq 2$, and $\mathsf{cov}(J_1, I_1) = 0$, and thus by Theorem 3.5 $\mathsf{cov}(F_n, U_n) = \sum_{k=2}^{n} k^{-2}$. □

It follows further easily from the central limit theorem for vector-valued random variables, e.g., with Lyapunov's condition, that the normal limits in Theorems 2.1 and 3.1 hold jointly, with the joint limit being a pair of two independent standard normal variables. We omit the details.

4. The number of rows

In this section we consider the number of rows R_n in a random permutation tableau of length n. Since the rows of a permutation tableau correspond to S steps in the process of its construction we can write

$$R_n = \sum_{k=1}^{n} I_{M_k=\mathsf{S}}.$$

LEMMA 4.1. *For $1 \le k \le n$ and any complex numbers z, w we have*

$$\mathsf{E}_k\big(z^{R_k}w^{U_k}\big) = \frac{w(z-1)}{k}\mathsf{E}_{k-1}\big(z^{R_{k-1}}w^{U_{k-1}}\big) + \frac{w}{k}\mathsf{E}_{k-1}\big(z^{R_{k-1}}(w+1)^{U_{k-1}}\big).$$

PROOF. By conditioning on $\mathcal{F}_{k-1}$ we get

$$\mathsf{E}_k(z^{R_k}w^{U_k}) = \mathsf{E}_k\left(z^{R_{k-1}}\mathsf{E}_k(z^{I_{M_k=\mathsf{S}}}w^{U_k} \mid \mathcal{F}_{k-1})\right). \tag{4.1}$$

Note that $M_k = \mathsf{S}$ if and only if $U_k = 1 + U_{k-1}$ and, conditionally on $\mathcal{F}_{k-1}$, this happens with probability $1/2^{U_{k-1}}$. Hence, using also Lemma 3.3 with $z = 1$ (or the argument in (2.3)),

$$\begin{aligned}
\mathsf{E}_k(z^{I_{M_k=\mathsf{S}}}w^{U_k} \mid \mathcal{F}_{k-1}) &= \mathsf{E}_k\big((z^{I_{M_k=\mathsf{S}}}-1)w^{U_k} \mid \mathcal{F}_{k-1}\big) + \mathsf{E}_k(w^{U_k} \mid \mathcal{F}_{k-1})\\
&= \mathsf{E}_k\big((z-1)I_{M_k=\mathsf{S}}w^{U_k} \mid \mathcal{F}_{k-1}\big) + w\left(\frac{w+1}{2}\right)^{U_{k-1}}\\
&= (z-1)\frac{w^{1+U_{k-1}}}{2^{U_{k-1}}} + w\left(\frac{w+1}{2}\right)^{U_{k-1}}\\
&= w(z-1)\left(\frac{w}{2}\right)^{U_{k-1}} + w\left(\frac{w+1}{2}\right)^{U_{k-1}}.
\end{aligned}$$

Putting this into (4.1) and applying (2.1) proves the lemma. □

We can now compute the probability generating function of R_n.

THEOREM 4.2. *We have*

$$g_{R_n}(z) = \frac{1}{n!}\sum_{r=1}^{n}\left\langle {n \atop r-1} \right\rangle z^r,$$

where $\left\langle {n \atop k} \right\rangle$ are the Eulerian numbers counting the number of permutations $\sigma = (\sigma_1, \dots, \sigma_n)$ of $[n]$ with k descents (recall that a descent is a pair (σ_j, σ_{j+1}) such that $\sigma_j > \sigma_{j+1}$). Thus R_n has the same distribution as 1 plus the number of descents in a random permutation of $[n]$.

PROOF. Consider $\mathsf{E}_n z^{R_n}$. Applying Lemma 4.1 k times and collecting the terms involving the expectation of the same expression we see that it is of the form

$$\mathsf{E}_n z^{R_n} = \frac{1}{n\cdots(n-k+1)}\sum_{m=0}^{k}c_{k,m}(z)\mathsf{E}_{n-k}\big(z^{R_{n-k}}(m+1)^{U_{n-k}}\big), \tag{4.2}$$

(for $0 \le k \le n$) with certain coefficient functions $c_{k,m} = c_{k,m}(z)$. Apply Lemma 4.1 again to each of the expectations in the sum to get

$$\begin{aligned}
\mathsf{E}_{n-k}\big(z^{R_{n-k}}(m+1)^{U_{n-k}}\big) = \frac{(m+1)(z-1)}{n-k}&\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&+ \frac{m+1}{n-k}\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+2)^{U_{n-k-1}}\big).
\end{aligned}$$

Putting that back in, we see that $(n\cdots(n-k))g_{R_n}(z)$ is equal to

$$\begin{aligned}
&\sum_{m=0}^{k}c_{k,m}(m+1)(z-1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&\qquad+ \sum_{m=0}^{k}c_{k,m}(m+1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+2)^{U_{n-k-1}}\big).
\end{aligned}$$

By rearranging the terms this is

$$\begin{aligned}
c_{k,0}&(z-1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}\big)\\
&\qquad+\sum_{m=1}^{k} c_{k,m}(m+1)(z-1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&+\sum_{m=0}^{k-1} c_{k,m}(m+1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+2)^{U_{n-k-1}}\big)\\
&\qquad+c_{k,k}(k+1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(k+2)^{U_{n-k-1}}\big)\\
&=c_{k,0}(z-1)\mathsf{E}_{n-k-1}z^{R_{n-k-1}}\\
&\quad+\sum_{m=1}^{k}\{c_{k,m}(m+1)(z-1)+mc_{k,m-1}\}\,\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big)\\
&\qquad+c_{k,k}(k+1)\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(k+2)^{U_{n-k-1}}\big)\\
&=\sum_{m=0}^{k+1} c_{k+1,m}\mathsf{E}_{n-k-1}\big(z^{R_{n-k-1}}(m+1)^{U_{n-k-1}}\big),
\end{aligned}$$

where the coefficients $c_{k,m}$ satisfy the following recurrence: $c_{k,m}=0$ unless $0\le m\le k$, $c_{0,0}=1$, and for $0\le m\le k$,

$$c_{k,m}=mc_{k-1,m-1}+(m+1)(z-1)c_{k-1,m}.$$

It follows by induction that $c_{k,m}(z)$ is a constant times $(z-1)^{k-m}$, so

$$c_{k,m}=a_{k,m}(z-1)^{k-m},$$

where $a_{0,0}=1$ and

$$a_{k,m}=\begin{cases} ma_{k-1,m-1}+(m+1)a_{k-1,m}, & \text{if } 0\le m\le k,\ k\ge 1;\\ 0, & \text{otherwise.}\end{cases} \tag{4.3}$$

If we now let $k=n-1$ in (4.2) and use $\mathsf{E}_1\big(z^{R_1}(m+1)^{U_1}\big)=z(m+1)$ then we get

$$g_{R_n}(z)=\frac{z}{n!}\sum_{m=0}^{n-1}(m+1)c_{n-1,m}=\frac{z}{n!}\sum_{m=0}^{n-1}(m+1)a_{n-1,m}(z-1)^{n-1-m}. \tag{4.4}$$

The recurrence (4.3) is solved by

$$a_{k,m}=m!\left\{{k+1 \atop m+1}\right\},$$

where $\left\{{k \atop m}\right\}$ is the number of partitions of the k element set into m non-empty subsets (recall the basic recurrence $\left\{{k+1 \atop m+1}\right\}=(m+1)\left\{{k \atop m+1}\right\}+\left\{{k \atop m}\right\}$). Hence, (4.4) becomes

$$\begin{aligned}
g_{R_n}(z)&=\frac{z}{n!}\sum_{m=0}^{n-1}(m+1)!\left\{{n \atop m+1}\right\}(z-1)^{n-(m+1)}\\
&=\frac{z}{n!}\sum_{m=0}^{n}m!\left\{{n \atop m}\right\}(z-1)^{n-m},
\end{aligned}$$

where in the last step we shifted the index by one and used $\left\{ {n \atop 0} \right\} = 0$ for $n \geq 1$. This completes the proof since for a complex z (see [**K**, Section 5.1.3])

$$\sum_{r=1}^{n} \left\langle {n \atop r-1} \right\rangle \frac{z^r}{n!} = \frac{z}{n!} \sum_{m=0}^{n} m! \left\{ {n \atop m} \right\} (z-1)^{n-m}. \qquad \square$$

Once the coefficients have been identified as Eulerian numbers we can use their known properties (see e.g. [**DB**, Chapter 10, pp. 150–154]) to obtain

Corollary 4.3. *The number of rows R_n in a random permutation tableau of length n satisfies*

$$\frac{R_n - (n+1)/2}{\sqrt{(n+1)/12}} \xrightarrow{d} N(0,1). \tag{4.5}$$

In fact, a local limit theorem holds as well (see e.g. [**CKSS**] or [**E**]).

Moreover, for the number of columns in a random permutation tableau, by definition, we have $C_n = n - R_n$. Further, let D_n denote the number of descents in a random permutation; thus $R_n \overset{\mathrm{d}}{=} D_n + 1$ by Theorem 4.2. Hence $C_n = n - R_n \overset{\mathrm{d}}{=} n - 1 - D_n = A_n$, the number of ascents in a random permutation. By symmetry, $A_n \overset{\mathrm{d}}{=} D_n \overset{\mathrm{d}}{=} R_n - 1$, and thus we obtain the following symmetry property:

Corollary 4.4. *The number C_n of columns in a random permutation tableau satisfies $C_n \overset{\mathrm{d}}{=} R_n - 1$.*

In particular,

$$\frac{C_n - (n-1)/2}{\sqrt{(n+1)/12}} \xrightarrow{d} N(0,1).$$

The fact that $R_n - 1 \overset{\mathrm{d}}{=} C_n \overset{\mathrm{d}}{=} D_n \overset{\mathrm{d}}{=} A_n$ follows also by the bijections described in [**CoN**].

Remark 4.5. The coefficients in (4.4) can be more explicitly written as

$$a_{n-1,n-r-1} = (n-r-1)! \sum_{1 \leq j_1 < \cdots < j_r \leq n-1} j_1(j_2-1)\cdots(j_r-(r-1)). \tag{4.6}$$

This can be seen by putting them in a Pascal type triangle

$$\begin{array}{ccccccccc}
 & & & & a_{0,0} & & & & \\
 & & & \swarrow & & \searrow & & & \\
 & & a_{1,0} & & & & a_{1,1} & & \\
 & \swarrow & & \searrow & & \swarrow & & \searrow & \\
 & & \cdots & & \cdots & & \cdots & & \\
 & & \cdots & & \cdots & & \cdots & & \\
\swarrow & & \searrow\ \swarrow & & \cdots & & \searrow\ \swarrow & & \searrow \\
a_{n-1,0} & & a_{n-1,1} & & \cdots & & a_{n-1,n-2} & & a_{n-1,n-1}
\end{array}$$

and observing that by (4.3), a move down (either SW or SE) from a coefficient $a_{k,m}$ has weight $m+1$. The value of a given coefficient at the bottom is obtained by summing over all possible paths leading to it from the root $a_{0,0}$ the products of weights corresponding to the moves along the path. Any path contributing to $a_{n-1,n-1-r}$, $0 \leq r \leq n-1$ has exactly r SW moves and if they are from levels

$j_1, \dots, j_r$ (with level 0 at the top and $n-1$ at the bottom), then the path has weight

$$(n-1-r)!(j_1+1)j_2(j_3-1)\cdots(j_r-(r-2)),$$

and (4.6) follows by renumbering the terms.

Expression (4.6) gives a direct way to get the moments; from (4.4) the probability generating function of $R_n - 1$ is

$$g_{R_n-1}(z) = \frac{1}{n!}\sum_{r=0}^{n-1}(n-r)a_{n-1,n-r-1}(z-1)^r,$$

and thus the rth factorial moment of $R_n - 1$ is

$$\mathsf{E}_n\big((R_n-1)(R_n-2)\cdots(R_n-r)\big) = \frac{d^r}{dz^r}g_{R_n-1}(z)\Big|_{z=1} = \frac{r!(n-r)}{n!}a_{n-1,n-r-1}$$
$$= \frac{r!\,(n-r)!}{n!}\sum_{1\le j_1<\cdots<j_r\le n-1} j_1(j_2-1)\cdots(j_r-(r-1)).$$

In particular,

$$\mathsf{E}_n R_n = 1 + \mathsf{E}_n(R_n-1) = 1 + \frac{1}{n}\sum_{j=1}^{n-1} j = 1 + \frac{n-1}{2} = \frac{n+1}{2},$$

and

$$\begin{aligned}\mathsf{var}(R_n) = \mathsf{var}(R_n-1) &= \mathsf{E}_n((R_n-1)(R_n-2)) + \mathsf{E}_n(R_n-1) - (\mathsf{E}_n(R_n-1))^2\\ &= \frac{2}{(n-1)n}\sum_{1\le j_1<j_2\le n-1} j_1(j_2-1) + \frac{n-1}{2} - \frac{(n-1)^2}{4}\\ &= \frac{(n-2)(3n-5)}{12} + \frac{n-1}{2} - \frac{(n-1)^2}{4} = \frac{n+1}{12},\end{aligned}$$

which explains the normalization in (4.5).

5. Probability generating function of the number of superfluous 1s

Let S_n be the number of superfluous 1s in a random tableau of length n. In [**CoH**] calculations based on (2.1) and (2.2) were used to show that

$$\mathsf{E}_n S_n = \frac{(n-1)(n-2)}{12}. \tag{5.1}$$

In this section we use the same approach to derive the expression for the probability generating function of S_n, see Proposition 5.3 or equation (5.5) below. Although the form of the probability generating function looks rather unwieldy, it can be used to compute (in practice low order) moments of S_n in a relatively straightforward, albeit a bit tedious way. We will illustrate it by deriving the exact expression for the variance of S_n (see Proposition 5.5) below. However, there does not seem to be an easy way to derive the central limit theorem for S_n directly from the probability generating function. For this reason, in the next section we will rely on a different approach to establish the CLT for S_n.

We write in this section

$$S_n = \sum_{k=1}^{n} V_k,$$

where V_k is the increase in the number of superfluous 1s when the length of a tableau is increased from $k-1$ to k. As in the preceding section let G_k be the row number of the topmost 1 (provided that the kth move is W and counting only unrestricted rows) and recall that it is set to be $U_{k-1}+1$ if the kth move is S. Then by the same argument as in Lemma 3.2

$$\mathsf{P}_k(G_k = j \mid \mathcal{F}_{k-1}) = \frac{1}{2^j}, \quad j = 1, \dots, U_{k-1},$$

and

$$\mathsf{P}_k(G_k = U_{k-1}+1 \mid \mathcal{F}_{k-1}) = \frac{1}{2^{U_{k-1}}}.$$

Moreover, the joint conditional distribution of (U_k, V_k) given $\mathcal{F}_{k-1}$ is

$$\mathcal{L}((U_k, V_k) \mid \mathcal{F}_{k-1}) = (G_k + \operatorname{Bin}(U_{k-1} - G_k), \operatorname{Bin}(U_{k-1} - G_k)), \tag{5.2}$$

in the sense that if we further condition on $G_k = m$, then the distribution equals that of $(m+X, X)$ with $X \sim \operatorname{Bin}(U_{k-1}-m)$ (thus the two occurences of $\operatorname{Bin}(U_{k-1} - G_k)$ in (5.2) signify the same random variable); further we interpret $\operatorname{Bin}(-1)$ as 0.

Lemma 5.1. *We have, for all complex z and w, and $k \geq 1$:*

$$\begin{aligned}\mathsf{E}_k(&z^{V_k} w^{U_k} \mid \mathcal{F}_{k-1}) \\ &= w \left\{ \frac{1}{1+w(z-1)} \left(\frac{zw+1}{2} \right)^{U_{k-1}} + \left(1 - \frac{1}{1+w(z-1)} \right) \left(\frac{w}{2} \right)^{U_{k-1}} \right\}.\end{aligned}$$

and

$$\begin{aligned}\mathsf{E}_k(z^{S_k} w^{U_k}) = \frac{w}{k} \Big\{ &\frac{1}{1+w(z-1)} \mathsf{E}_{k-1}\left(z^{S_{k-1}} (zw+1)^{U_{k-1}} \right) \\ &+ \Big(1 - \frac{1}{1+w(z-1)} \Big) \mathsf{E}_{k-1}\left(z^{S_{k-1}} w^{U_{k-1}} \right) \Big\}.\end{aligned}$$

Proof. Using (5.2) we have

$$\begin{aligned}\mathsf{E}_k(z^{V_k} w^{U_k} \mid \mathcal{F}_{k-1}) &= \mathsf{E}_k(z^{\operatorname{Bin}(U_{k-1}-G_k)} w^{G_k + \operatorname{Bin}(U_{k-1}-G_k)} \mid \mathcal{F}_{k-1}) \\ &= \sum_{m=1}^{U_{k-1}} \mathsf{E}_k\big((zw)^{\operatorname{Bin}(U_{k-1}-m)} w^m I(G_k = m) \mid \mathcal{F}_{k-1} \big) \\ &\qquad + w^{1+U_{k-1}} \mathsf{P}_k(G_k = U_{k-1}+1 \mid \mathcal{F}_{k-1}) \\ &= \sum_{m=1}^{U_{k-1}} \left(\frac{w}{2} \right)^m \mathsf{E}_k((zw)^{\operatorname{Bin}(U_{k-1}-m)} \mid \mathcal{F}_{k-1}) + w \left(\frac{w}{2} \right)^{U_{k-1}} \\ &= \sum_{m=1}^{U_{k-1}} \left(\frac{w}{2} \right)^m \left(\frac{zw+1}{2} \right)^{U_{k-1}-m} + w \left(\frac{w}{2} \right)^{U_{k-1}} \\ &= \sum_{\ell=0}^{U_{k-1}-1} \left(\frac{zw+1}{2} \right)^{\ell} \left(\frac{w}{2} \right)^{U_{k-1}-\ell} + w \left(\frac{w}{2} \right)^{U_{k-1}} \\ &= \left(\frac{w}{2} \right)^{U_{k-1}} \left\{ \frac{1}{(1+zw)/w - 1} \left(\left(\frac{zw+1}{w} \right)^{U_{k-1}} - 1 \right) + w \right\}\end{aligned}$$

$$= w\left\{\frac{1}{1+w(z-1)}\left(\frac{zw+1}{2}\right)^{U_{k-1}} + \left(1-\frac{1}{1+w(z-1)}\right)\left(\frac{w}{2}\right)^{U_{k-1}}\right\},$$

which is the first formula.

To prove the second write

$$\mathsf{E}_k\big(z^{S_k}w^{U_k}\big) = \mathsf{E}_k\mathsf{E}_k(z^{S_k}w^{U_k}\mid \mathcal{F}_{k-1}) = \mathsf{E}_k\big(z^{S_{k-1}}\mathsf{E}_k(z^{V_k}w^{U_k}\mid \mathcal{F}_{k-1})\big)$$

and use the first part and the usual reduction by (2.1) from P_k to P_{k-1}. □

For $\ell \ge 0$ set $b_\ell = b_\ell(z) = \sum_{j=0}^{\ell} z^j$ so that we have

$$zb_\ell + 1 = b_{\ell+1} \quad \text{and} \quad 1+(z-1)b_\ell = z^{\ell+1}. \tag{5.3}$$

If we substitute b_ℓ for w in Lemma 5.1 and use (5.3) we obtain the following.

Lemma 5.2. *For $0 \le k \le n-1$ and $\ell \ge 0$,*

$$\mathsf{E}_{n-k}\big(z^{S_{n-k}}b_\ell^{U_{n-k}}\big) = \frac{b_\ell}{n-k}\Big(z^{-\ell-1}\mathsf{E}_{n-k-1}\big(z^{S_{n-k-1}}b_{\ell+1}^{U_{n-k-1}}\big) + (1-z^{-\ell-1})\mathsf{E}_{n-k-1}\big(z^{S_{n-k-1}}b_\ell^{U_{n-k-1}}\big)\Big).$$

Proposition 5.3. *We have*

$$g_{S_n}(z) = \frac{1}{n!}\sum_{\ell=0}^{n-1} c_{n-1,\ell}(z)b_\ell(z),$$

where the coefficients $c_{m,\ell} = c_{m,\ell}(z)$ satisfy $c_{0,0}=1$ and for $m \ge 0$ the recurrence

$$c_{m+1,\ell} = c_{m,\ell}(1-z^{-\ell-1})b_\ell(z) + c_{m,\ell-1}z^{-\ell}b_{\ell-1}(z).$$

In particular, $c_{m,\ell}=0$ unless $0 \le \ell \le m$.

Proof. Skipping the dependence of b_ℓ's on z for notational convenience, we show first that for $0 \le k \le n-1$ we have

$$\mathsf{E}_n z^{S_n} = \mathsf{E}_n z^{S_n}b_0^{U_n} = \frac{1}{n\cdots(n-k+1)}\sum_{m=0}^{k} c_{k,m}\mathsf{E}_{n-k}z^{S_{n-k}}b_m^{U_{n-k}}. \tag{5.4}$$

Indeed, assume inductively (5.4) for some $k \ge 0$ ($k=0$ being trivial). Apply the previous lemma to each of the terms $\mathsf{E}_{n-k}z^{S_{n-k}}b_m^{U_{n-k}}$ to get

$$\mathsf{E}_n z^{S_n} = \frac{1}{n\cdots(n-k)}\sum_{m=0}^{k} c_{k,m}b_m\Big\{z^{-m-1}\mathsf{E}_{n-k-1}z^{S_{n-k-1}}b_{m+1}^{U_{n-k-1}} + (1-z^{-m-1})\mathsf{E}_{n-k-1}z^{S_{n-k-1}}b_m^{U_{n-k-1}}\Big\}.$$

Separating the sums, rearranging the terms, and collecting the coefficients in front of $\mathsf{E}_{n-k-1}z^{S_{n-k-1}}b_m^{U_{n-k-1}}$ we obtain that $\mathsf{E}_n z^{S_n}$ is equal to

$$\frac{1}{n\cdots(n-k)}\Big\{\sum_{m=0}^{k+1}\big(c_{k,m-1}b_{m-1}z^{-m} + c_{k,m}b_m(1-z^{-m-1})\big)\,\mathsf{E}_{n-k-1}z^{S_{n-k-1}}b_m^{U_{n-k-1}}\Big\}.$$

This completes the inductive proof of (5.4) with the coefficients $c_{k,m}$ given by the specified recurrence. Choosing $k=n-1$ in (5.4) and using the observation that

$$\mathsf{E}_1 z^{S_1}b_m^{U_1} = b_m,$$

(because $S_1 = 0$ and $U_1 = 1$) completes the proof. □

REMARK 5.4. Just as in the proof of Theorem 4.2, the coefficients $c_{k,m}$ can be put in a Pascal-type triangle

$$
\begin{array}{ccccccccc}
 & & & & c_{0,0} & & & & \\
 & & & \swarrow & & \searrow & & & \\
 & & c_{1,0} & & & & c_{1,1} & & \\
 & \swarrow & & \searrow & & \swarrow & & \searrow & \\
 & & \cdots & & \cdots & & \cdots & & \\
 & & \cdots & & \cdots & & \cdots & & \\
\swarrow & & \searrow\ \swarrow & & \cdots & & \searrow\ \swarrow & & \searrow \\
c_{n-1,0} & & c_{n-1,1} & & \cdots & & c_{n-1,n-2} & & c_{n-1,n-1}
\end{array}
$$

This time a SE move from a coefficient $c_{m,\ell}$ has weight $z^{-\ell-1}b_\ell(z)$ and a SW move has weight $(1-z^{-\ell-1})b_\ell(z)$. As was in the case of the number of rows, the value of a given coefficient at the bottom is obtained by summing over all possible paths leading to it from the root $c_{0,0}$ the products of weights corresponding to the moves along the path. This can be used to obtain an explicit expressions for the coefficients $(c_{n-1,m})$. Any path from $c_{0,0}$ to $c_{n-1,n-1-r}$, $0 \le r \le n-1$, has exactly r SW moves and $n-1-r$ SE moves. The total weight of SE moves (they are from $c_{\cdot,0}, c_{\cdot,1}, \dots, c_{\cdot,n-2-r}$) is

$$\prod_{j=0}^{n-2-r} z^{-j-1}b_j(z) = z^{-\binom{n-r}{2}} \prod_{j=0}^{n-r-2} b_j(z).$$

The SW moves may be from $c_{\cdot,\ell_1}, \dots, c_{\cdot,\ell_r}$, for some $0 \le \ell_1 \le \dots \le \ell_r \le n-1-r$. A SW move from $c_{\cdot,\ell_j}$ has weight $(1-z^{-\ell_j-1})b_{\ell_j}(z)$. Therefore,

$$c_{n-1,n-1-r}(z) = z^{-\binom{n-r}{2}} \prod_{k=0}^{n-r-2} b_k(z) \sum_{0\le\ell_1\le\dots\le\ell_r\le n-1-r} \prod_{j=1}^{r}(1-z^{-\ell_j-1})b_{\ell_j}(z).$$

This gives an explicit expression for the probability generating function of S_n, namely

$$g_{S_n}(z) = \frac{1}{n!}\sum_{r=0}^{n-1} z^{-\binom{n-r}{2}} \left(\prod_{k=0}^{n-r-1} b_k(z)\right) A_{n,r}(z), \tag{5.5}$$

where

$$A_{n,r}(z) = \sum_{0\le\ell_1\le\dots\le\ell_r\le n-1-r} \prod_{j=1}^{r}(1-z^{-\ell_j-1})b_{\ell_j}(z). \tag{5.6}$$

(Note that all negative powers cancel in (5.5) since $g_{S_n}(z)$ is a polynomial in z.) Although this expression looks quite complicated and it is not clear to us at the moment how to deduce the asymptotic normality of S_n from it, some information can be extracted from it. We will illustrate it by deriving an exact expression for the variance of S_n.

PROPOSITION 5.5. *For $n \ge 2$, the variance of S_n satisfies*

$$\mathsf{var}(S_n) = \frac{(n-2)(2n^2+11n-1)}{360}. \tag{5.7}$$

PROOF. We compute the second factorial moment of S_n:

$$\mathsf{E}_n S_n(S_n-1) = \frac{1}{n!}\frac{d^2}{dz^2}\Big(\sum_{k=0}^{n-1} c_{n-1,k} b_k\Big)\Big|_{z=1}. \tag{5.8}$$

Notice that every path contributing to $c_{n-1,r}$, for $0 \le r < n-3$ has at least three SW moves and so its weight will have at least three factors of the form $1 - z^{-j}$. Hence if it is differentiated twice and evaluated at $z = 1$ it will vanish. It follows that the sum on the right hand side of (5.8) reduces to the last three terms. For $k = n-3$ we have

$$\frac{1}{n!} c_{n-1,n-3} b_{n-3} = \frac{z^{-\binom{n-2}{2}}}{n!} \left(\prod_{k=0}^{n-4} b_k\right) b_{n-3} A_{n,2}(z) = G(z) A_{n,2}(z),$$

where we have set

$$G(z) := \frac{1}{n!} z^{-\binom{n-2}{2}} \prod_{k=0}^{n-3} b_k(z).$$

So, the second derivative of $G(z)A_{n,2}(z)$ is

$$G''(z)A_{n,2}(z) + 2G'(z)A'_{n,2}(z) + G(z)A''_{n,2}(z).$$

Since

$$A_{n,2}(z) = \sum_{1\le \ell \le m \le n-2} (1-z^{-\ell})(1-z^{-m}) b_{\ell-1}(z) b_{m-1}(z),$$

$A_{n,2}(1) = A'_{n,2}(1) = 0$ so we only need $G(1)A''_{n,2}(1)$. Now,

$$G(1) = \frac{1}{n!}\prod_{k=0}^{n-3}(k+1) = \frac{1}{n(n-1)}. \tag{5.9}$$

To compute $A''_{n,2}(1)$, writing

$$(1-z^{-\ell})(1-z^{-m}) = h_1(z), \quad b_{\ell-1}(z)b_{m-1}(z) = h_2(z),$$

we see that

$$h_1''(1)h_2(1) + 2h_1'(1)h_2'(1) + h_1(1)h_2''(1) = h_1''(1)h_2(1) = 2\ell^2 m^2.$$

Therefore,

$$\begin{aligned} A''_{n,2}(1) &= 2\sum_{m=1}^{n-2} m^2 \sum_{\ell=1}^{m} \ell^2 = \frac{1}{3}\sum_{m=1}^{n-2} m^3(m+1)(2m+1) \\ &= \frac{1}{180} n(n-1)(n-2)(5n-11)(2n-1)(2n-3), \end{aligned}$$

and combining this with (5.9) we obtain

$$\frac{1}{n!}\frac{d^2}{dz^2}\left(c_{n-1,n-3}b_{n-3}\right)\Big|_{z=1} = \frac{(n-2)(5n-11)(2n-1)(2n-3)}{180}.$$

We next handle

$$\begin{aligned} \frac{1}{n!} c_{n-1,n-1} b_{n-1} &= \frac{z^{-\binom{n}{2}}}{n!}\left(\prod_{k=0}^{n-2} b_k\right) b_{n-1} \\ &= z^{-\binom{n}{2}} \prod_{j=1}^{n-1} \frac{b_{j-1}}{j}. \end{aligned}$$

The last expression is the probability generating function of a sum of independent random variables $W_0, \dots, W_n$, where $W_0 \equiv -n(n-1)/2$ and for $1 \le j \le n$, W_j is a discrete uniform random variable on $\{0, \dots, j-1\}$. So, with $Y_n = \sum_{j=0}^n W_j$ we have

$$\frac{1}{n!}\frac{d^2}{dz^2}c_{n-1,n-1}b_{n-1}\Big|_{z=1} = \mathsf{E}Y_n(Y_n-1) = \sum_{j=1}^n \mathsf{var}(W_j) + (\mathsf{E}Y_n)^2 - \mathsf{E}Y_n.$$

Now,

$$\mathsf{E}W_j = \frac{j-1}{2}, \qquad 1 \le j \le n,$$

so that

$$\mathsf{E}Y_n = \frac{1}{2}\sum_{j=1}^n (j-1) - \frac{n(n-1)}{2} = -\frac{n(n-1)}{4}.$$

Furthermore, for $1 \le j \le n$

$$\mathsf{var}(W_j) = \frac{1}{j}\sum_{k=0}^{j-1} k^2 - \left(\frac{j-1}{2}\right)^2 = \frac{(j-1)(j+1)}{12}.$$

Therefore,

$$\sum_{j=1}^n \mathsf{var}(W_j) = \frac{n(n-1)(2n+5)}{72}.$$

Finally, putting the above together we get

$$\frac{1}{n!}\frac{d^2}{dz^2}c_{n-1,n-1}b_{n-1}\Big|_{z=1} = \frac{n(n-1)(2n+5)}{72} + \frac{n^2(n-1)^2}{16} + \frac{n(n-1)}{4}$$
$$= \frac{n(n-1)(9n^2-5n+46)}{144}.$$

It remains to handle $\frac{1}{n!}c_{n-1,n-2}b_{n-2}$. We write

$$c_{n-1,n-2}b_{n-2} = f(z)A_{n,1}(z),$$

where

$$f(z) = z^{-\binom{n-1}{2}}\prod_{k=0}^{n-2} b_k(z)$$

and, according to (5.6),

$$A_{n,1}(z) = \sum_{\ell=0}^{n-2}(1 - z^{-\ell-1})b_\ell(z).$$

Then the second derivative of $c_{n-1,n-2}b_{n-2}$ is

$$f''(z)A_{n,1}(z) + 2f'(z)A'_{n,1}(z) + f(z)A''_{n,1}(z).$$

At $z=1$ the first product vanishes. For the remaining two first notice that

$$\frac{1}{(n-1)!}f(z) = z^{-\binom{n-1}{2}}\prod_{j=1}^{n-1}\frac{b_{j-1}}{j},$$

is a generating function of a legitimate distribution function. Therefore, $f(1) = (n-1)!$ and $f'(1)/(n-1)!$ is the expected value of a random variable represented by $f(z)/(n-1)!$. Since this expected value is

$$-\binom{n-1}{2} + \sum_{j=1}^{n-1} \frac{1}{j} \sum_{k=0}^{j-1} k = -\frac{(n-1)(n-2)}{4},$$

we obtain

$$f'(1) = -\frac{(n-1)(n-2)(n-1)!}{4}.$$

It remains to compute the first two derivatives of $A_{n,1}(z)$ at $z = 1$.

$$A'_{n,1}(z) = \sum_{\ell=1}^{n-1} \left(b'_{\ell-1}(1 - z^{-\ell}) + b_{\ell-1}\ell z^{-\ell-1} \right),$$

so that

$$A'_{n,1}(1) = \sum_{\ell=1}^{n-1} \ell^2 = \frac{n(n-1)(2n-1)}{6}. \tag{5.10}$$

Also

$$A''_{n,1}(z) = \sum_{\ell=1}^{n-1} \left(b''_{\ell-1}(1 - z^{-\ell}) + 2b'_{\ell-1}\ell z^{-\ell-1} - b_{\ell-1}\ell(\ell+1) z^{-\ell-2} \right).$$

Since $b'_{\ell-1}(1) = (\ell-1)\ell/2$ we get

$$A''_{n,1}(1) = \sum_{\ell=1}^{n-1} \left((\ell-1)\ell^2 - \ell^2(\ell+1) \right) = -\frac{n(n-1)(2n-1)}{3}.$$

Hence,

$$\begin{aligned}
\frac{1}{n!}\frac{d^2}{dz^2} c_{n-1,n-2} b_{n-2}\Big|_{z=1} &= \frac{1}{n!}\left(2f'(1)A'_{n,1}(1) + f(1)A''_{n,1}(1)\right) \\
&= -\frac{(n-1)^2(n-2)(2n-1)}{12} - \frac{(n-1)(2n-1)}{3} \\
&= -\frac{(n-1)(2n-1)(n^2-3n+6)}{12}.
\end{aligned}$$

Combining all of these calculations gives

$$\begin{aligned}
\mathsf{E}_n S_n(S_n - 1) &= \frac{(n-2)(5n-11)(2n-1)(2n-3)}{180} \\
&\quad - \frac{(n-1)(2n-1)(n^2-3n+6)}{12} \\
&\quad + \frac{n(n-1)(9n^2-5n+46)}{144} \\
&= \frac{(n-2)(n-3)(5n^2-n-16)}{720}.
\end{aligned}$$

By the same argument

$$\begin{aligned}
\mathsf{E}_n S_n &= \frac{1}{n!}\frac{d}{dz}\Big(c_{n-1,n-1}b_{n-1} + c_{n-1,n-2}b_{n-2}\Big)\Big|_{z=1} \\
&= \mathsf{E}Y_n + \frac{1}{n!}\left(f'(1)A_{n,1}(1) + f(1)A'_{n,1}(1)\right)
\end{aligned}$$

Using $A_{n,1}(1) = 0$, $f(1) = (n-1)!$, (5.10), and the value of $\mathsf{E}Y_n$ we get

$$\mathsf{E}_n S_n = -\frac{n(n-1)}{4} + \frac{(n-1)(2n-1)}{6} = \frac{(n-1)(n-2)}{12},$$

which conforms to (5.1). Finally,

$$\begin{aligned}\mathsf{var}(S_n) &= \mathsf{E}S_n(S_n-1) - (\mathsf{E}S_n)^2 + \mathsf{E}S_n \\ &= \frac{(n-2)(n-3)(5n^2-n-16)}{720} - \left(\frac{(n-1)(n-2)}{12}\right)^2 + \frac{(n-1)(n-2)}{12} \\ &= \frac{(n-2)(2n^2+11n-1)}{360},\end{aligned}$$

which proves Proposition 5.5. □

6. Asymptotic normality of S_n

In this section we provide a self-contained proof of the following

THEOREM 6.1. *As $n \to \infty$ we have*

$$\frac{S_n - n^2/12}{\sqrt{n^3/180}} \xrightarrow{d} N(0,1). \tag{6.1}$$

PROOF. As we mentioned, the form of the probability generating function of S_n obtained in the previous section does not seem to be convenient to yield the central limit theorem. For this reason, we will rely on a bijective result of Steingrímsson and Williams [**SW**]. According to their result the number of superfluous 1s in a permutation tableau of length n is equidistributed with the number of occurrences of a generalized patterns 31−2 in a random permutation of $[n]$. (An occurrence of a generalized pattern 31−2 in a permutation σ is a pair $1 < i < j$ such that $\sigma_{i-1} > \sigma_j > \sigma_i$.) To analyze that quantity, it will be convenient to think of a random permutation as generated from a sample $X_1, \dots, X_n$ i.i.d. random variables each being uniform on $[0,1]$ (the permutation is obtained by reading off the ranks of $X_1, \dots, X_n$). If for $2 \le i < j \le n$ we let $I_{i,j} := I_{X_{i-1} > X_j > X_i}$ then

$$S_n = \sum_{2 \le i < j \le n} I_{i,j}.$$

Notice that from this representation we immediately recover (5.1) since

$$\mathsf{E}S_n = \mathsf{E}\sum_{2 \le i < j \le n} I_{i,j} = \binom{n-1}{2}\mathsf{P}(X_1 > X_3 > X_2) = \frac{1}{6}\binom{n-1}{2}.$$

Similarly, we can easily obtain the asymptotic value of the variance: we write

$$\mathsf{var}(S_n) = \sum_{\substack{i_1 < j_1 \\ i_2 < j_2}} \mathsf{cov}(I_{i_1,j_1}, I_{i_2,j_2}),$$

and note that if $\{i_1 - 1, i_1, j_1\} \cap \{i_2 - 1, i_2, j_2\} = \emptyset$ then I_{i_1,j_1} and I_{i_2,j_2} are independent and so their covariance vanishes. In the complementary case, the main contribution comes from the cases that contribute $\Theta(n^3)$ terms to the sum. We obtain

$$\begin{aligned}\mathsf{var}(S_n) \sim \frac{n^3}{3}\Big(&\mathsf{cov}(I_{2,3}, I_{2,4}) + \mathsf{cov}(I_{2,5}, I_{4,5}) + \mathsf{cov}(I_{2,4}, I_{3,5})\\ &+ \mathsf{cov}(I_{2,5}, I_{3,4}) + \mathsf{cov}(I_{2,3}, I_{4,5}) + \mathsf{cov}(I_{2,4}, I_{4,5})\Big),\end{aligned}$$

as all other cases contribute $O(n^2)$ terms to the sum. We calculate:

$$\mathsf{E}I_{2,3} \cap I_{2,4} = \mathsf{P}(X_1 > X_3 > X_2, X_1 > X_4 > X_2) = 2\mathsf{P}(X_1 > X_3 > X_4 > X_2) = \frac{2}{4!},$$

so that

$$\mathsf{cov}(I_{2,3}, I_{2,4}) = \frac{1}{12} - \left(\frac{1}{6}\right)^2.$$

Similarly,

$$\begin{aligned}\mathsf{E}I_{2,5} \cap I_{4,5} &= \mathsf{P}(X_1 > X_5 > X_2, X_3 > X_5 > X_4)\\ &= 4\mathsf{P}(X_1 > X_3 > X_5 > X_2 > X_4) = \frac{1}{30},\\ \mathsf{E}I_{2,4} \cap I_{3,5} &= \mathsf{P}(X_1 > X_4 > X_2, X_2 > X_5 > X_3) = \frac{1}{120},\\ \mathsf{E}I_{2,5} \cap I_{3,4} &= \mathsf{P}(X_1 > X_5 > X_2, X_2 > X_4 > X_3) = \frac{1}{120},\\ \mathsf{E}I_{2,3} \cap I_{4,5} &= \mathsf{P}(X_1 > X_3 > X_2, X_3 > X_5 > X_4)\\ &= 3\mathsf{P}(X_1 > X_3 > X_2 > X_5 > X_4) = \frac{1}{40},\\ \mathsf{E}I_{2,4} \cap I_{4,5} &= \mathsf{P}(X_1 > X_4 > X_2, X_3 > X_5 > X_4) = \frac{1}{40}.\end{aligned}$$

Hence,

$$\mathsf{var}(S_n) \sim \frac{n^3}{3}\cdot\left(\frac{1}{12} + \frac{1}{30} + \frac{1}{120} + \frac{1}{120} + \frac{1}{40} + \frac{1}{40} - \frac{6}{36}\right) = \frac{n^3}{180}. \tag{6.2}$$

Remark 6.2. The exact value of the variance could be obtained by computing the other terms.

Finally, to establish (6.1) we will rely on results presented in [**J**] and [**JŁR**, Section 6.1]; see [**Bo**] for a closely related theorem proved by the same method. See also [**E**] for a related simple proof of the asymptotic normality of the number of descents, cf. Corollary 4.3 above. We let $A = A_n := \{\alpha = (i, j) :\ 2 \le i < j \le n\}$. Then $S_n = \sum_{\alpha\in A} I_\alpha$. Recall that a dependency graph L for $\{I_\alpha :\ \alpha \in A\}$ is any graph whose vertex set is A and which has the property that if V_1, V_2 are two disjoint subsets of A such that L has no edges with one endpoint in V_1 and the other in V_2 then the families $\{I_\alpha :\ \alpha \in V_1\}$ and $\{I_\alpha :\ \alpha \in V_2\}$ are mutually independent. For our purposes it is enough to consider L defined by the following rule: we put an edge between $\alpha_1 = (i_1, j_1)$ and $\alpha_2 = (i_2, j_2)$ iff $\{i_1 - 1, i_1, j_1\} \cap \{i_2 - 1, i_2, j_2\} \ne \emptyset$. If $\alpha_1, \dots, \alpha_r \in A$ then the closed neighborhood of $\{\alpha_1, \dots, \alpha_r\}$ in L is defined by

$$\bar N_L(\alpha_1, \dots, \alpha_r) = \bigcup_{i=1}^r \{\beta \in A :\ \beta = \alpha_i \text{ or } \alpha_i\beta \in E(L)\},$$

where $E(L)$ denotes the edge set of L. Note that in our case for every fixed $r \ge 2$ $|\bar N_L(\alpha_1, \dots, \alpha_{r-1})| = O_r(n)$, where $O_r(\ \cdot\ )$ means that the constant may depend on r. Hence, trivially

$$\sum_{\alpha\in\bar N_L(\alpha_1,\dots,\alpha_{r-1})} \mathsf{E}\big(I_\alpha \big| I_{\alpha_1},\dots,I_{\alpha_{r-1}}\big) = O_r(n).$$

Since, cf. (5.1) for the exact value,

$$\sum_{\alpha\in A} \mathsf{E} I_\alpha \le |A| = O(n^2),$$

by [**JŁR**, Lemma 6.17] we conclude that the rth cumulant of S_n (defined by $\kappa_r(S_n) = i^{-k}\frac{d^k}{dt^k}\log\phi_{S_n}(0)$, where $\phi_X(t) = g_X(e^{it})$ is the characteristic function) satisfies

$$|\kappa_r(S_n)| = O_r(n^2\cdot n^{r-1}) = O_r(n^{r+1}).$$

Hence, for $r\ge 3$

$$\left|\kappa_r\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right)\right| = O_r(n^{r+1}n^{-\frac{3}{2}r}) = o(1),$$

as $n\to\infty$. Since $\kappa_1(X) = \mathsf{E}X$ and $\kappa_2(X) = \mathsf{var}(X)$ we have

$$\kappa_1\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right) \longrightarrow 0, \qquad \kappa_2\left(\frac{S_n - n^2/12}{\sqrt{n^3/180}}\right) \longrightarrow 1,$$

and the theorem follows by the cumulant convergence theorem (see e.g. [**JŁR**, Corollary 6.15]). $\square$

Remark 6.3. Our results can be used to draw conclusions about some other parameters. For example, let Y_n be the number of 1s in the random permutation tableaux T_n. Although we have not computed an explicit formula for the distribution of Y_n, we can easily obtain its asymptotic distribution. Namely, as $n\to\infty$ we have

$$\frac{Y_n - n^2/12}{\sqrt{n^3/180}} \xrightarrow{d} N(0,1).$$

This follows immediately from Theorem 6.1 upon noting that

$$Y_n = S_n + C_n = S_n + O(n).$$

It seems straightforward to prove a similar central limit theorem for Z_n, the number of 0s in the random permutation tableaux T_n, using further bijective results by [**SW**]; the main difference is that we need to consider several generalized patterns simultaneously and the joint distribution of their numbers of occurrences. We leave this to the reader.

References

[BaS] E. Babson, E. Steingrímsson, *Generalized permutation patterns and a classification of the Mahonian statistics*, Sém. Lothar. Combin., B44b (2000) 18 pp.

[Bo] M. Bóna, The copies of any permutation pattern are asymptotically normal, preprint 2007. arXiv:math.CO/0712.2792v1.

[BrCPRE] R. Brak, S. Corteel, R. Parviainen, A. Rechnitzer, J. Essam, *A combinatorial derivation of the PASEP stationary state*, Electron. J. Combin., **13** (2006) R108.

[BrE] R. Brak, J. Essam, *Asymmetric exclusion model and weighted lattice paths*, J. Phys. A 37, no. **14** (2004) 4183–4217.

[Bu] A. Burstein, *On some properties of permutation tableaux*, Ann. Combin. **11** (2008) 355–368.

[CKSS] L. Carlitz, D. S. Kurtz, R. Scoville, O. P. Stackelberg, *Asymptotic properties of Eulerian numbers*, Z. Wahrsch. Verw. Gebiete **23** (1972) 47–54.

[Cl] A. Claesson, *Generalized pattern avoidance*, European J. Combin. **22** (2001) 961-971.

[ClM] A. Claesson and T. Mansour, *Enumerating permutations avoiding a pair of Babson–Steingrímsson patterns*, Ars Combinatoria, **77** (2005) 17–31.

[Co] S. Corteel, *Crossings and alignments of permutations*, Adv. in Appl. Math. **38** (2007) 149-163.

[CoH] S. Corteel and P. Hitczenko, *Expected values of statistics on permutation tableaux*, in 2007 Conference on Analysis of Algorithms, Discrete Mathematics and Theoretical Computer Science AH, 325–340, 2007.

[CoN] S. Corteel and P. Nadeau, *Permutation tableaux and permutation descents*, Europ. J. Combin. **30** (2009) 295–310.

[CoW] S. Corteel, L. Williams, *Tableaux combinatorics for the asymmetric exclusion process*, Adv. in Appl. Math. **39**(3) (2007) 293–310.

[CoW1] S. Corteel, L. Williams, *A Markov chain on permutations which projects to the PASEP*, Int. Math. Res. Not. IMRN (2007) article ID rnm055.

[DB] F. N. David and D. E. Barton, *Combinatorial Chance*, London, Griffin,1962.

[DDM] B. Derrida, E. Domany, D. Mukamel, *An exact solution of a one-dimensional asymmetric exclusion model with open boundaries*, J. Stat. Phys. **69** (1992) 667-687.

[DEHP] B. Derrida, M. Evans, V. Hakim, V. Pasquier, *Exact solution of a 1D asymmetric exclusion model using a matrix formulation*, J. Phys. A: Math. Gen. **26** (1993) 1493–1517.

[DS] E. Duchi, G. Schaeffer, *A combinatorial approach to jumping particles*, J. Combin. Theory Ser. A **110** (2005) 1–29.

[E] C.-G. Esseen, On the application of the theory of probability to two combinatorial problems involving permutations. *Proceedings of the seventh conference on probability theory (Braşov, 1982)*, 137–147, VNU Sci. Press, Utrecht, 1985.

[F] W. Feller, *An Introduction to Probability Theory and Its Applications*. Vol. I. Third edition, John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[J] S. Janson, Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.* **16** (1988), 305–312.

[JŁR] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs*. Wiley, New York, 2000.

[K] D. N. Knuth, *The Art of Computer Programming, vol. 3*. Addison-Wesley, 1996.

[P] A. Postnikov, *Total positivity, Grassmannians, and networks*, preprint (2006). arXiv:math.CO/0609764.

[R] J. Riordan, *An Introduction to Combinatorial Analysis*. Wiley Publications in Mathematical Statistics John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1958.

[S] T. Sasamoto, *One-dimensional partially asymmetric simple exclusion process with open boundaries: orthogonal polynomials approach*, J. Phys. A: Math. Gen. **32** (1999) 7109–7131.

[SW] E. Steingrímsson, L. Williams, *Permutation tableaux and permutation patterns*, J. Combin. Theory, Ser. A **114** (2007) 211-234.

[USW] M. Uchiyama, T. Sasamoto, M. Wadati, *Asymmetric simple exclusion process with open boundaries and Askey-Wilson polynomials*, J. Phys. A: Math. Gen. **37** (2004) 4985–5002.

[W] L. Williams, *Enumeration of totally positive Grassmann cells*, Adv. Math, **190** (2005) 319-342.

Department of Mathematics, Drexel University, Philadelphia, Pennsylvania 19104
E-mail address: `phitczenko@math.drexel.edu`

Department of Mathematics, Uppsala University, Box 480, SE-751 06 Uppsala, Sweden
E-mail address: `svante.janson@math.uu.se`