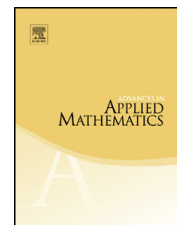




Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama


Asymptotic normality of the major index on standard tableaux

Sara C. Billey^{a,1}, Matjaž Konvalinka^{b,c,2}, Joshua P. Swanson^{d,*}^a Department of Mathematics, University of Washington, Seattle, WA 98195, USA^b Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, Ljubljana, Slovenia^c Institute for Mathematics, Physics and Mechanics, Jadranska 19, Ljubljana, Slovenia^d Department of Mathematics, University of California, San Diego (UCSD), La Jolla, CA 92093-0112, USA

ARTICLE INFO

Article history:

Received 2 May 2019

Received in revised form 1 November 2019

Accepted 5 November 2019

Available online xxxx

MSC:

05A16

60C05

60F05

Keywords:

Major index

Hook length

Tableaux

Asymptotic normality

ABSTRACT

We consider the distribution of the major index on standard tableaux of arbitrary straight shape and certain skew shapes. We use cumulants to classify all possible limit laws for any sequence of such shapes in terms of a simple auxiliary statistic, aft, generalizing earlier results of Canfield–Janson–Zeilberger, Chen–Wang–Wang, and others. These results can be interpreted as giving a very precise description of the distribution of irreducible representations in different degrees of coinvariant algebras of certain complex reflection groups. We conclude with some conjectures concerning unimodality, log-concavity, and local limit theorems.

© 2019 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail addresses: billey@math.washington.edu (S.C. Billey), matjaz.konvalinka@fmf.uni-lj.si (M. Konvalinka), jswanson@ucsd.edu (J.P. Swanson).¹ Partially supported by the Washington Research Foundation and DMS-1764012 from the National Science Foundation.² Partially supported by Research Project BI-US/16-17-042 of the Slovenian Research Agency and research core funding No. P1-0294.

1. Introduction

The study of permutation and partition statistics is a classic topic in enumerative combinatorics. The major index statistic on permutations was introduced a century ago by Percy MacMahon in his seminal works [27,28]. This statistic, denoted $\text{maj}(w)$, is defined to be the sum of the positions of the descents of the permutation $w = [w_1, w_2, \dots, w_n]$ in one-line notation. A descent is any position i such that $w_i > w_{i+1}$. At first glance, this function on permutations may be unintuitive, but it has inspired hundreds of papers and many generalizations; for example on Macdonald polynomials [21], posets [12], quasisymmetric functions [38], cyclic sieving [34,2], and bijective combinatorics [15,7].

The following central limit theorem for maj on the symmetric group S_n is well known and is an archetype for our results. Given a real-valued random variable \mathcal{X} , we let

$$\mathcal{X}^* := \frac{\mathcal{X} - \mu}{\sigma}$$

denote the corresponding normalized random variable with mean 0 and variance 1. Briefly, we say maj on S_n is *asymptotically normal* as $n \rightarrow \infty$ based on the following classical result. See Table 1 for further examples.

Theorem 1.1. [14] *Let $\mathcal{X}_n[\text{maj}]$ denote the major index random variable on S_n under the uniform distribution. Then, for all $t \in \mathbb{R}$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{X}_n[\text{maj}]^* \leq t] = \mathbb{P}[\mathcal{N} \leq t]$$

where \mathcal{N} is the standard normal random variable.

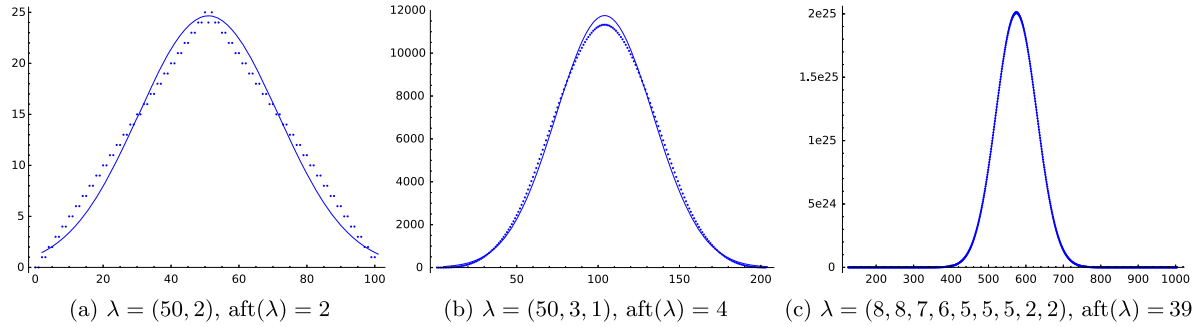
In this paper, we study the distribution of the major index statistic generalized to standard Young tableaux of straight and skew shapes. The properties we discuss here naturally generalize known properties of the major index distribution on permutations. They also have representation theoretic consequences in terms of coinvariant algebras of complex reflection groups. We will briefly introduce the main results. See Section 2 for more details on the background.

Let $\text{SYT}(\lambda)$ denote the set of all standard Young tableaux of partition shape λ . We say i is a *descent* in a standard tableau T if $i+1$ comes before i in the row reading word of T , read from bottom to top along rows in English notation. Equivalently, i is a descent in T if $i+1$ appears in a lower row in T . Let $\text{maj}(T)$ denote the *major index statistic* on $\text{SYT}(\lambda)$, which is again defined to be the sum of the descents of T . Fig. 1 shows

Table 1

Summary of some asymptotic normality results for combinatorial statistics. See [5, Ch. 3].

Statistic	Set	Generating function	References
# elements	subsets	$(1+q)^n$	classical
# parts	strict partitions	$\prod_{m=1}^{\infty} (1+xy^m)$	[13]
length/inversion number/major index	S_n	$[n]_q!$	[14], [19]
# cycles; # left-to-right minima	S_n	$\prod_{i=0}^{n-1} (q+i)$	[14], [19]
# descents	S_n	Eulerian polynomial $A_n(q)$	[10, pp. 150–154]
# descents	conjugacy classes in S_n	[17, Thm. 1]	[17,26]
# blocks	set partitions	$\sum_k S(n,k)q^k$	[22]
# valleys	Dyck paths	$\frac{1}{[n+1]_q} \binom{2n}{n}_q$	[8, Cor. 3.3]; [18, p. 255]
length/inversion number/major index	S_n/S_J , words type α	$\binom{n}{\alpha}_q$	see Theorem 3.17
major index	$\text{SYT}(\lambda)$	$q^{b(\lambda)} \frac{[n]_q!}{\prod_{c \in \lambda} [h_c]_q}$	Theorem 1.3

**Fig. 1.** Plots of $\#\{T \in \text{SYT}(\lambda) : \text{maj}(T) = k\}$ as a function of k for three partitions λ , overlaid with scaled Gaussian approximations using the same mean and variance.

some sample distributions for the major index on standard tableaux for three particular partition shapes. Note that Gaussian approximations fit the data well.

In Theorem 1.1, we simply let $n \rightarrow \infty$. For partitions, the shape λ may “go to infinity” in many different ways. The following statistic on partitions overcomes this difficulty.

Definition 1.2. Suppose λ is a partition. Let the *aft* of λ be

$$\text{aft}(\lambda) := |\lambda| - \max\{\lambda_1, \lambda'_1\}.$$

Intuitively, if the first row of λ is at least as long as the first column, then $\text{aft}(\lambda)$ is the number of cells *not* in the first row. This definition is strongly reminiscent of a *representation stability* result of Church and Farb [9, Thm. 7.1], which is proved with an analysis of the major index on standard tableaux.

Our first main result gives the analogue of Theorem 1.1 for maj on SYT(λ). In particular, it completely classifies which sequences of partition shapes give rise to asymptotically normal sequences of maj statistics on standard tableaux.

Theorem 1.3. *Suppose $\lambda^{(1)}, \lambda^{(2)}, \dots$ is a sequence of partitions, and let $\mathcal{X}_N = \mathcal{X}_{\lambda^{(N)}}[\text{maj}]$ be the corresponding random variables for the maj statistic on SYT($\lambda^{(N)}$). Then, the sequence $\mathcal{X}_1, \mathcal{X}_2, \dots$ is asymptotically normal if and only if $\text{aft}(\lambda^{(N)}) \rightarrow \infty$ as $N \rightarrow \infty$.*

Remark 1.4. In Section 5, we more generally consider maj on SYT($\underline{\lambda}$) where $\underline{\lambda}$ is a *block diagonal skew partition*. See [3, §2] for further representation-theoretic motivation and [3, Thm. 6.3] for the classification of the support of maj on SYT($\underline{\lambda}$).

The generalization of Theorem 1.3 to SYT($\underline{\lambda}$) is Theorem 5.8. Special cases of Theorem 5.8 include Canfield–Janson–Zeilberger’s main result in [35] classifying asymptotic normality for inv or maj on words (though see [6] for earlier, essentially equivalent results due to Diaconis [11]). The case of words generalizes Theorem 1.1. The $\lambda^{(N)} = (N, N)$ case of Theorem 1.3 also recovers the main result of Chen–Wang–Wang [8], giving asymptotic normality for q -Catalan coefficients.

Our proof of Theorem 1.3 relies on the *method of moments*, which requires useful descriptions of the moments of $\mathcal{X}_\lambda[\text{maj}]$. Adin–Roichman [1] gave exact formulas for the mean and variance of $\mathcal{X}_\lambda[\text{maj}]$ in terms of the *hook lengths* of λ . Their argument leverages the following q -analogue of the celebrated Frame–Robinson–Thrall Hook Length Formula [16, Thm. 1] (obtained by setting $q = 1$):

$$\text{SYT}(\lambda)^{\text{maj}}(q) := \sum_{T \in \text{SYT}(\lambda)} q^{\text{maj}(T)} = q^{b(\lambda)} \frac{[n]_q!}{\prod_{c \in \lambda} [h_c]_q}, \quad (1)$$

where h_c denotes the hook length of a cell c in λ and $b(\lambda) := \sum_{i \geq 1} (i-1)\lambda_i$. Equation (1) is due to Stanley [40, Cor. 7.21.5] and is strongly related to the stable principal specialization of Schur functions by the identity $s_\lambda(1, q, q^2, \dots) = \text{SYT}(\lambda)^{\text{maj}}(q) / \prod_{i=1}^{|\lambda|} (1 - q^i)$ [40, Prop. 7.19.11].

In fact, formulas for the d th moment μ_d^λ , d th central moment α_d^λ , and d th *cumulant* κ_d^λ of maj on SYT(λ) may be derived from (1). The most elegant of these formulas is for the cumulants, from which the moments and central moments are all easy to compute.

Theorem 1.5. *Let $\lambda \vdash n$ and $d \in \mathbb{Z}_{>1}$. We have*

$$\kappa_d^\lambda = \frac{B_d}{d} \left[\sum_{j=1}^n j^d - \sum_{c \in \lambda} h_c^d \right] \quad (2)$$

where $B_0, B_1, B_2, \dots = 1, \frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, \dots$ are the Bernoulli numbers.

See Theorem 2.9 for a generalization of (2) along with exact formulas for the moments and central moments. See Theorem 2.10 for some of the history of this formula.

Remark 1.6. For “most” partition shapes, one expects the term $\sum_{j=1}^n j^d$ in (2) to dominate $\sum_{c \in \lambda} h_c^d$, in which case asymptotic normality is quite straightforward. However, for some shapes there is a very large amount of cancellation in (2) and determining the limit law can be quite subtle.

While $\mathcal{X}_\lambda[\text{maj}]$ can be written as the sum of scaled indicator random variables $D_1, 2D_2, 3D_3, \dots, (n-1)D_{n-1}$ where D_i determines if there is a descent at position i , the D_i are not at all independent, so one may not simply apply standard central limit theorems. Interestingly, the D_i are identically distributed [40, Prop. 7.19.9]. The lack of independence of the D_i ’s likewise complicates related work by Fulman [17] and Kim–Lee [26] considering the limiting distribution of descents in certain classes of permutations.

The non-normal continuous limit laws for maj on $\text{SYT}(\lambda)$ turn out to be the *Irwin–Hall distributions* $\mathcal{IH}_M := \sum_{k=1}^M \mathcal{U}[0, 1]$, which are the sum of M i.i.d. continuous $[0, 1]$ random variables. The following result completely classifies all possible limit laws for maj on $\text{SYT}(\lambda)$ for any sequence of partition shapes. See Theorem 6.3 for the generalization to block diagonal skew shapes.

Theorem 1.7. *Let $\lambda^{(1)}, \lambda^{(2)}, \dots$ be a sequence of partitions. Then $(\mathcal{X}_{\lambda^{(N)}}[\text{maj}])^*$ converges in distribution if and only if*

- (i) $\text{aft}(\lambda^{(N)}) \rightarrow \infty$; or
- (ii) $|\lambda^{(N)}| \rightarrow \infty$ and $\text{aft}(\lambda^{(N)}) \rightarrow M < \infty$; or
- (iii) the distribution of $\mathcal{X}_{\lambda^{(N)}}^*[\text{maj}]$ is eventually constant.

The limit law is \mathcal{N} in case (i), \mathcal{IH}_M^ in case (ii), and discrete in case (iii).*

Case (iii) naturally leads to the question, when does $\mathcal{X}_\lambda^*[\text{maj}] = \mathcal{X}_\mu^*[\text{maj}]$? Such a description in terms of hook lengths is given in Theorem 7.1. Theorem 1.7 naturally raises several open questions and conjectures concerning unimodality, log-concavity, and local limit theorems, which are described in Section 8.

Example 1.8. We illustrate each possible limit in Theorem 1.7. For (i), let $\lambda^{(N)} := (N, \lfloor \ln N \rfloor)$, so that $\text{aft}(\lambda^{(N)}) = \lfloor \ln N \rfloor \rightarrow \infty$ and the distributions are asymptotically normal. For (ii), fix $M \in \mathbb{Z}_{\geq 0}$ and let $\lambda^{(N)} := (N+M, M)$, so that $\text{aft}(\lambda^{(N)}) = M$ is constant and the distributions converge to Σ_M^* . For (iii), let $\lambda^{(2N)} := (12, 12, 3, 3, 3, 2, 2, 1, 1)$ and $\lambda^{(2N+1)} := (15, 6, 6, 6, 4, 2)$, which have the same multisets of hook lengths despite not being transposes of each other, and consequently the same normalized maj distributions.

The rest of the paper is organized as follows. In Section 2, we give background focused on cumulants aimed at the combinatorial audience. In Section 3, we collect combinatorial background on permutations, tableaux, etc, aimed more at the probabilistic audience. In Section 4, we analyze $\text{maj} - \text{inv}$ on S_n as an introductory example. In Section 5, we classify when maj on $\text{SYT}(\underline{\lambda})$ is asymptotically normal. In Section 6, we determine the remaining continuous limit laws for maj on $\text{SYT}(\underline{\lambda})$. In Section 7, we characterize the possible discrete distributions for maj on $\text{SYT}(\lambda)$ in terms of hook lengths. Finally, Section 8 lists conjectures concerning unimodality, log-concavity, and local limit theorems.

2. Background on cumulants

In this section, we review some standard terminology and results on generating functions, random variables, and asymptotic normality, with a focus on *cumulants*. An excellent source for many further details in this area can be found in Canfield's Chapter 3 of [5].

2.1. Exponential generating functions

We now introduce our notation for exponential generating functions and the Bernoulli numbers, which will be used with cumulants shortly.

Definition 2.1. Given a rational sequence $(g_d)_{d=0}^\infty = (g_0, g_1, \dots)$, the corresponding *ordinary generating function* is

$$O_g(t) := \sum_{d \geq 0} g_d t^d$$

and the corresponding *exponential generating function* is

$$E_g(t) := \sum_{d \geq 0} g_d \frac{t^d}{d!}.$$

Conversely, any rational power series

$$F(t) = \sum_{d \geq 0} f_d t^d = \sum_{d \geq 0} d! f_d \frac{t^d}{d!}$$

is the ordinary generating function of the sequence $(f_d)_{d=0}^\infty = (f_0, f_1, \dots)$ and the exponential generating function of the sequence $(d! f_d)_{d=0}^\infty$. The exponential generating functions we will encounter will all have a positive radius of convergence.

It is easy to describe products, quotients and compositions of generating functions. We recall in particular a formula for compositions of exponential generating functions

for later use. Given two rational sequences $f = (f_d)_{d=0}^\infty$, $g = (g_d)_{d=0}^\infty$ such that $f_0 = 0$ and $g_0 = 1$, the composition of their exponential generating functions $E_g \circ E_f$ is again an exponential generating function for a rational sequence h , say $E_h(t) = E_g(E_f(t))$. For example, if $E_f(t) = \sum f_d t^d/d!$ and $E_g(t) = e^t$, so $g_i = 1$ for all i , then by [40, Cor. 5.1.6], the corresponding sequence $(h_d)_{d=0}^\infty$ is given by $h_0 = 1$ and, for $d \geq 1$,

$$h_d = \sum_{\pi \in \Pi_d} \prod_{b \in \pi} f_{|b|}, \quad (3)$$

where Π_d is the collection of all set partitions $\pi = \{b_1, b_2, \dots, b_k\}$ of $\{1, 2, \dots, d\}$. Collecting together S_d -orbits of Π_d in (3) quickly gives

$$h_d = \sum_{\lambda \vdash d} \frac{d!}{z_\lambda} \prod_i \frac{f_{\lambda_i}}{(\lambda_i - 1)!} \quad (4)$$

where if λ has m_i parts of length i , then $z_\lambda := 1^{m_1} 2^{m_2} \dots m_1! m_2! \dots$. A more computationally efficient, recursive approach to (3) is the formula [40, Prop. 5.1.7]

$$h_d = f_d + \sum_{m=1}^{d-1} \binom{d-1}{m-1} f_m h_{d-m}. \quad (5)$$

Example 2.2. The *Bernoulli numbers* $(B_d)_{d=0}^\infty$ are rational numbers determined by the exponential generating function $E_B(t) := t/(1 - e^{-t})$. The first few terms in the sequence are

$$\begin{aligned} B_0 = 1, \quad B_1 = \frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42}, \\ B_7 = 0, \quad B_8 = -\frac{1}{30}, \quad B_9 = 0, \quad B_{10} = \frac{5}{66}, \quad B_{11} = 0, \quad B_{12} = -\frac{691}{2730}. \end{aligned}$$

The *divided Bernoulli numbers* are given by $\frac{B_d}{d}$ for $d \geq 1$. Their exponential generating function $E_D(t)$ satisfies $1 + t \frac{d}{dt} E_D(t) = E_B(t)$, from which it follows that

$$E_D(t) := \sum_{d \geq 1} \frac{B_d}{d} \frac{t^d}{d!} = \log \left(\frac{e^t - 1}{t} \right).$$

We caution that a common alternate convention for Bernoulli numbers uses $B_1 = -\frac{1}{2}$ with all other entries the same, corresponding with the exponential generating function $t/(e^t - 1)$.

The Bernoulli numbers have many interesting properties; see [30,46] and [20, Section 6.5]. For example, they appear in the polynomial expansion of the sums of d th powers,

$$\sum_{k=1}^n k^d = \frac{1}{d+1} \sum_{k=0}^d \binom{d+1}{k} B_k n^{d+1-k}. \quad (6)$$

Compare the formula for sums of d th powers to the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ which can be evaluated at complex values $s \neq 1$ by analytic continuation. The divided Bernoulli numbers which appear in our formula (2) satisfy $\frac{B_d}{d} = -\zeta(1-d)$.

2.2. Probabilistic generating functions

We next review basic vocabulary and notation for moments and cumulants of random variables. All random variables we encounter will have moments of all orders. See [4] for more details.

Definition 2.3. Let \mathcal{X} be a real-valued random variable where either \mathcal{X} is continuous with probability density function $f: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ or \mathcal{X} is discrete with probability mass function $f: \mathbb{Z} \rightarrow \mathbb{R}_{\geq 0}$. The *cumulative distribution function* (CDF) of \mathcal{X} is given by

$$F(t) := \int_{-\infty}^t f(x) dx \quad \text{or} \quad F(t) := \sum_{k \leq t} f(k)$$

depending on whether \mathcal{X} is continuous or discrete. For any continuous real-valued function g , there is an associated random variable $g(\mathcal{X})$. The *expectation* of $g(\mathcal{X})$ is given by

$$\mathbb{E}[g(\mathcal{X})] := \int_{\mathbb{R}} g(x) f(x) dx \quad \text{or} \quad \mathbb{E}[g(\mathcal{X})] := \sum_{k=-\infty}^{\infty} g(k) f(k).$$

The *mean* and *variance* of \mathcal{X} are, respectively,

$$\mu := \mathbb{E}[\mathcal{X}] \quad \text{and} \quad \sigma^2 := \mathbb{E}[(\mathcal{X} - \mu)^2].$$

For $d \in \mathbb{Z}_{\geq 0}$, the d th *moment* and d th *central moment* of \mathcal{X} are, respectively,

$$\mu_d := \mathbb{E}[\mathcal{X}^d] \quad \text{and} \quad \alpha_d := \mathbb{E}[(\mathcal{X} - \mu)^d].$$

The *moment-generating function* of \mathcal{X} is

$$M_{\mathcal{X}}(t) := \mathbb{E}[e^{t\mathcal{X}}] = \sum_{d=0}^{\infty} \mu_d \frac{t^d}{d!},$$

which for us will always have a positive radius of convergence. The *characteristic function* of \mathcal{X} is

$$\phi_{\mathcal{X}}(t) := \mathbb{E}[e^{it\mathcal{X}}],$$

which exists for all $t \in \mathbb{R}$ and which is the Fourier transform of f , the density or mass function associated to \mathcal{X} .

Example 2.4. Let W be a finite set with an integer statistic $\text{stat}: W \rightarrow \mathbb{Z}_{\geq 0}$. We will use the notation

$$W^{\text{stat}}(q) := \sum_{w \in W} q^{\text{stat}(w)}$$

for the corresponding polynomial generating function. If $W^{\text{stat}}(q) = \sum c_k q^k$, define a random variable \mathcal{X} associated with $\text{stat}: W \rightarrow \mathbb{Z}_{\geq 0}$ sampled uniformly on W by $\mathbb{P}(\mathcal{X} = k) = c_k / \#W$. The *probability generating function* for \mathcal{X} is

$$\mathbb{E}[q^{\mathcal{X}}] = \frac{1}{\#W} W^{\text{stat}}(q) := \frac{1}{\#W} \sum_{w \in W} q^{\text{stat}(w)}.$$

Letting $q = e^t$, an easy computation shows that the moment-generating function and characteristic function of \mathcal{X} are

$$M_{\mathcal{X}}(t) = \frac{1}{\#W} W^{\text{stat}}(e^t) \quad \text{and} \quad \phi_{\mathcal{X}}(t) = \frac{1}{\#W} W^{\text{stat}}(e^{it}).$$

These expressions reveal an intimate connection between the study of generating functions of combinatorial statistics evaluated on the unit circle and the underlying probability distribution via the Laplace and Fourier transforms. In particular, the distribution determines the characteristic function and the moment-generating function, and conversely each of these determines the distribution.

Definition 2.5. The *cumulants* $\kappa_1, \kappa_2, \dots$ of \mathcal{X} are defined to be the coefficients of the exponential generating function

$$K_{\mathcal{X}}(t) := \sum_{d=1}^{\infty} \kappa_d \frac{t^d}{d!} := \log M_{\mathcal{X}}(t) = \log \mathbb{E}[e^{t\mathcal{X}}].$$

While cumulants of random variables may initially be less intuitive than moments, they lead to nicer formulas in many cases, including Theorem 1.5, and they often have more useful properties. See [32] for some history and applications. We will use the following properties of cumulants. The proofs are straightforward from the definitions.

1. (*Familiar Values*) The first three cumulants are $\kappa_1 = \mu$, $\kappa_2 = \sigma^2$, and $\kappa_3 = \alpha_3$. The higher cumulants typically differ from the moments and central moments.
2. (*Shift Invariance*) The second and higher cumulants of \mathcal{X} agree with those for $\mathcal{X} - c$ for $c \in \mathbb{R}$.

3. (*Homogeneity*) The d th cumulant of $c\mathcal{X}$ is $c^d\kappa_d$ for $c \in \mathbb{R}$.
4. (*Additivity*) The cumulants of the sum of *independent* random variables are the sums of the cumulants.
5. (*Polynomial Equivalence*) The cumulants, moments, and central moments are determined by polynomials in any one of these three sequences.

The polynomial equivalence property can be made explicit by the results in Section 2.1. Equation (5) allows us to express the d th moment of \mathcal{X} as a polynomial function of the first d cumulants of \mathcal{X} and vice versa via the recurrence

$$\mu_d = \kappa_d + \sum_{m=1}^{d-1} \binom{d-1}{m-1} \kappa_m \mu_{d-m}. \quad (7)$$

Using the shift invariance property of cumulants, the corresponding formula for the central moments in terms of the cumulants can be obtained from (7) by setting $\kappa_1 = 0$ and leaving the other cumulants alone. This gives, for $d > 1$,

$$\alpha_d = \kappa_d + \sum_{m=2}^{d-2} \binom{d-1}{m-1} \kappa_m \alpha_{d-m}. \quad (8)$$

For instance, at $d = 3$ we have

$$\mu_3 = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3.$$

Setting $\kappa_1 = 0$ yields $\alpha_3 = \kappa_3$ as mentioned above.

2.3. Cumulant formulas

Next we describe the cumulants of some well-known distributions and use one of them to deduce a result of Hwang–Zacharovas, which immediately yields Theorem 1.5 as a corollary.

Example 2.6. Let $\mathcal{X} = \mathcal{N}(\mu, \sigma^2)$ be the normal random variable with mean μ and variance σ^2 . The density function of \mathcal{X} is $f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. Taking the Fourier transform gives the characteristic function $\mathbb{E}[e^{it\mathcal{X}}] = \exp(i\mu t - \frac{1}{2}\sigma^2 t^2)$, so the moment-generating function is $\mathbb{E}[e^{t\mathcal{X}}] = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$ and the cumulants are

$$\kappa_d = \begin{cases} \mu & d = 1, \\ \sigma^2 & d = 2, \\ 0 & d \geq 3. \end{cases} \quad (9)$$

Using (4) to compute the central moments of \mathcal{X} from (9), we effectively set $\kappa_1 = 0$ and note that only $\lambda = (2, 2, \dots, 2) = (2^{d/2})$ contributes, in which case $\alpha_d = \kappa_2^{d/2} d! / (2^{d/2} (d/2)!)$. It follows that

$$\alpha_d = \begin{cases} 0 & \text{if } d \text{ is odd,} \\ \sigma^d (d-1)!! & \text{if } d \text{ is even.} \end{cases}$$

Example 2.7. Let $\mathcal{U} = \mathcal{U}[0, 1]$ be the continuous uniform random variable whose density takes the value 1 on the interval $[0, 1]$ and 0 otherwise. Then the moment generating function is $M_{\mathcal{U}}(t) = \int_0^1 e^{tx} dx = (e^t - 1)/t$, so the cumulant generating function $\log M_{\mathcal{U}}(t)$ coincides with the exponential generating function for the divided Bernoulli numbers from Section 2.1. That is, $\kappa_d^{\mathcal{U}} = B_d/d$ for $d \geq 1$.

Recall from Section 1, \mathcal{IH}_m is the *Irwin–Hall* distribution obtained by adding m independent, identically distributed $\mathcal{U}[0, 1]$ random variables. By Additivity, the d th cumulant of \mathcal{IH}_m is mB_d/d . More generally, let $\mathcal{S} := \sum_{k=1}^m \mathcal{U}[\alpha_k, \beta_k]$ be the sum of m independent uniform continuous random variables. Then the d th cumulant of \mathcal{S} for $d \geq 2$ is

$$\kappa_d^{\mathcal{S}} = \frac{B_d}{d} \sum_{k=1}^m (\beta_k - \alpha_k)^d \quad (10)$$

by the Homogeneity and Additivity Properties of cumulants.

Example 2.8. Let \mathcal{U}_n be the discrete uniform random variable supported on $\{0, 1, \dots, n-1\}$. The probability generating function for \mathcal{U}_n is $[n]_q/n := (q^n - 1)/(n(q - 1))$, so the cumulant generating function is

$$\log M_{\mathcal{U}_n}(t) = \log \left(\frac{e^{nt} - 1}{n(e^t - 1)} \right) = \log \left(\frac{e^{nt} - 1}{nt} \right) - \log \left(\frac{e^t - 1}{t} \right).$$

It follows that for $d \geq 1$, the divided Bernoulli numbers arise again in this context,

$$\kappa_d^{\mathcal{U}_n} = \frac{B_d}{d} (n^d - 1). \quad (11)$$

Product formulas for polynomials such as Stanley’s formula (1) give rise to explicit formulas for cumulants and moments according to the following theorem. The proof is immediate from Theorem 2.8 and the exponential generating function identity (4).

Theorem 2.9. Suppose $\{a_1, \dots, a_m\}$ and $\{b_1, \dots, b_m\}$ are multisets of positive integers such that

$$P(q) = \frac{\prod_{k=1}^m [a_k]_q}{\prod_{k=1}^m [b_k]_q} = \sum c_k q^k \in \mathbb{Z}_{\geq 0}[q],$$

so in particular each $c_k \in \mathbb{Z}_{\geq 0}$. Let \mathcal{X} be a discrete random variable with $\mathbb{P}[\mathcal{X} = k] = c_k/P(1)$. Then the d th cumulant of \mathcal{X} is

$$\kappa_d^{\mathcal{X}} = \frac{B_d}{d} \sum_{k=1}^m (a_k^d - b_k^d) \quad (12)$$

where B_d is the d th Bernoulli number (with $B_1 = \frac{1}{2}$). Moreover, the d th central moment of \mathcal{X} is

$$\alpha_d = \sum_{\substack{\lambda \vdash d \\ \text{has all parts even}}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{k=1}^m (a_k^d - b_k^d) \right], \quad (13)$$

and the d th moment of \mathcal{X} is

$$\mu_d = \sum_{\substack{\lambda \vdash d \\ \text{has all parts either} \\ \text{even or size 1}}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{k=1}^m (a_k^d - b_k^d) \right]. \quad (14)$$

Remark 2.10. Equation (12) appeared explicitly in the work of Hwang–Zacharovas [24, §4.1] building on the work of Chen–Wang–Wang [8, Thm. 3.1], who in turn used an argument going back at least to Sachkov [36, §1.3.1]. It was rediscovered experimentally through (14) by the present authors and also rediscovered by Thiel–Williams [45].

One frequently encounters polynomials of the form $q^\beta P(q)$ for some $\beta \in \mathbb{Z}_{\geq 0}$, as in (1). The formulas in Theorem 2.9 remain valid in this case except that one must add β to the expression for κ_1 and add β to each factor in the product in (14) for which $\lambda_i = 1$.

Remark 2.11. The generating function machinery used to construct the cumulants in (12) works whether or not the function $P(q)$ is polynomial. The corresponding κ_d 's are called *formal cumulants* in the literature.

2.4. Asymptotic normality

Asymptotic normality is a very old topic lying at the intersection of probability and combinatorics. For an introduction, we recommend Canfield's Chapter 3 in [5].

Definition 2.12. Let $\mathcal{X}_1, \mathcal{X}_2, \dots$ and \mathcal{X} be real-valued random variables with cumulative distribution functions F_1, F_2, \dots and F , respectively. We say $\mathcal{X}_1, \mathcal{X}_2, \dots$ *converges in distribution* to \mathcal{X} , written $\mathcal{X}_n \Rightarrow \mathcal{X}$, if for all $t \in \mathbb{R}$ at which F is continuous we have

$$\lim_{n \rightarrow \infty} F_n(t) = F(t).$$

Recall from the introduction that for a real-valued random variable \mathcal{X} with mean μ and variance $\sigma^2 > 0$, the corresponding *normalized random variable* is

$$\mathcal{X}^* := \frac{\mathcal{X} - \mu}{\sigma}.$$

Observe that \mathcal{X}^* has mean $\mu^* = 0$ and variance $\sigma^{*2} = 1$. The moments and central moments of \mathcal{X}^* agree for $d \geq 2$ and are given by

$$\mu_d^* = \alpha_d^* = \alpha_d / \sigma^d.$$

Similarly, the cumulants of \mathcal{X}^* are given by $\kappa_1^* = 0$, $\kappa_2^* = 1$, and $\kappa_d^* = \kappa_d / \sigma^d$ for $d \geq 2$.

Definition 2.13. Let $\mathcal{X}_1, \mathcal{X}_2, \dots$ be a sequence of real-valued random variables. We say the sequence is *asymptotically normal* if $\mathcal{X}_n^* \Rightarrow \mathcal{N}(0, 1)$.

The “original” asymptotic normality result is as follows. Let $2^{[n]}$ be the set of all subsets of $[n] := \{1, 2, \dots, n\}$. Let $\mathcal{X}_{2^{[n]}}[\text{size}]$ denote the random variable given by the cardinality, where $2^{[n]}$ is given the uniform distribution. This has the same distribution as the number of heads after n fair coin flips, so the probability generating function up to normalization is $(1 + q)^n$. The following result is credited to de Moivre and Laplace; see [5, Theorem 3.2.1] for further discussion.

Theorem 2.14 (*de Moivre–Laplace*). *The sequence $\mathcal{X}_{2^{[n]}}[\text{size}]$ is asymptotically normal.*

Asymptotic normality results for combinatorial statistics are plentiful. See Table 1 for more examples and further references.

2.5. The method of moments

We next describe two standard criteria for establishing asymptotic normality or more generally convergence in distribution of a sequence of random variables.

Theorem 2.15 (*Lévy’s Continuity Theorem, [4, Theorem 26.3]*). *A sequence $\mathcal{X}_1, \mathcal{X}_2, \dots$ of real-valued random variables converges in distribution to a real-valued random variable \mathcal{X} if and only if, for all $t \in \mathbb{R}$,*

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{it\mathcal{X}_n}] = \mathbb{E}[e^{it\mathcal{X}}].$$

Theorem 2.16 (*Frechét–Shohat Theorem, [4, Theorem 30.2]*). *Let $\mathcal{X}_1, \mathcal{X}_2, \dots$ be a sequence of real-valued random variables, and let \mathcal{X} be a real-valued random variable. Suppose the moments of \mathcal{X}_n and \mathcal{X} all exist and the moment generating functions all have a positive radius of convergence. If*

$$\lim_{n \rightarrow \infty} \mu_d^{\mathcal{X}_n} = \mu_d^{\mathcal{X}} \quad \forall d \in \mathbb{Z}_{\geq 1}, \quad (15)$$

then $\mathcal{X}_1, \mathcal{X}_2, \dots$ converges in distribution to \mathcal{X} .

By Theorem 2.15, we may test for asymptotic normality by checking if the normalized characteristic functions tend point-wise to the characteristic function of the standard normal. Likewise by Theorem 2.16 we may instead perform the check on the level of individual normalized moments, which is often referred to as the *method of moments*. By (7) we may further replace the moment condition (15) with the cumulant condition

$$\lim_{n \rightarrow \infty} \kappa_d^{\mathcal{X}_n} = \kappa_d^{\mathcal{X}}. \quad (16)$$

For instance, we have the following explicit criterion.

Corollary 2.17. *A sequence $\mathcal{X}_1, \mathcal{X}_2, \dots$ of real-valued random variables on finite sets is asymptotically normal if for all $d \geq 3$ we have*

$$\lim_{n \rightarrow \infty} \frac{\kappa_d^{\mathcal{X}_n}}{(\sigma^{\mathcal{X}_n})^d} = 0 \quad (17)$$

In fact, one may show a converse of the Frechét–Shohat theorem holds for quotients as in Theorem 2.9, though we will not have need of it here.

2.6. Local limit theorems

Asymptotic normality concerns cumulative distribution functions, so it gives estimates for the number of combinatorial objects with a large range of statistics. However, our original motivation was to count combinatorial objects with a given statistic. Estimates of this latter form are frequently referred to as *local limit theorems*. Here we review two motivating examples.

The present work was partly inspired by the following local limit theorem due to the third author with a uniform rather than normal limit law. For $\lambda \vdash n$, let $\text{maj}_n: \text{SYT}(\lambda) \rightarrow [n]$ be maj modulo n .

Theorem 2.18. [44, Theorem 1.9] *For $\lambda \vdash n$, let $X_\lambda[\text{maj}_n]$ denote the random variable maj_n on $\text{SYT}(\lambda)$. Suppose $\#\text{SYT}(\lambda) \geq n^5$. Then, for all $k \in [n]$,*

$$\left| \mathbb{P}[X_\lambda[\text{maj}_n] = k] - \frac{1}{n} \right| < \frac{1}{n^2}.$$

Further motivation was provided by the following analogue of Theorem 3.16.

Theorem 2.19. [35, Theorem 4.5] *There exists a positive constant c such that for every C , the following is true. Uniformly for all compositions $\alpha = (\alpha_1, \dots, \alpha_m)$ such that $\max_i \alpha_i \leq Ce^{cs(\alpha)}$ and all integers k ,*

$$\mathbb{P}[X_\alpha = k] = \frac{1}{\sigma\sqrt{2\pi}} \left(e^{-(k-\mu)^2/(2\sigma^2)} + O\left(\frac{1}{s(\alpha)}\right) \right),$$

where X_α denotes inversions on words of type α .

3. Combinatorial background

3.1. Combinatorial background for $\text{baj} - \text{inv}$ on S_n

Here we introduce the two most well-known permutation statistics, inv and maj , as well as one unusual permutation statistic, baj .

Definition 3.1. Let $\sigma \in S_n$ be a permutation of $\{1, \dots, n\}$. Set

$$\text{Inv}(\sigma) := \{(i, j) : i < j \text{ and } \sigma(i) > \sigma(j)\} \quad (\text{inversion set})$$

$$\text{inv}(\sigma) := |\text{Inv}(\sigma)| \quad (\text{inversion number, i.e. length})$$

$$\text{Des}(\sigma) := \{1 \leq i \leq n-1 : \sigma(i) > \sigma(i+1)\} \quad (\text{descent set})$$

$$\text{maj}(\sigma) := \sum_{i \in \text{Des}(\sigma)} i \quad (\text{major index}).$$

Following Zabrocki [47] for the nomenclature, we also set

$$\text{baj}(\sigma) := \sum_{i \in \text{Des}(\sigma)} i(n-i).$$

The equidistribution of inv and maj on S_n is due to MacMahon, who also first introduced maj . His proof gave the following generating function expression for both statistics.

Theorem 3.2 ([27, Art. 6]). *We have*

$$S_n^{\text{inv}}(q) = [n]_q! := \prod_{k=1}^{n-1} (1 + q + q^2 + \dots + q^k) = S_n^{\text{maj}}(q).$$

The statistic $\text{baj} - \text{inv}$ appeared in the context of extended affine Weyl groups and Hecke algebras in the work of Iwahori and Matsumoto in 1965 [25]. It is the Coxeter length function restricted to coset representatives of the extended affine Weyl group of type A_{n-1} mod translations by coroots. Stembridge and Waugh [43, Remarks 1.5 and 2.3] give a careful overview of this topic and further results. In particular, they prove the

following factorization formula for the generating function associated to $\text{baj} - \text{inv}$ on S_n . From this factorization, the corresponding cumulants can be read off from Theorem 2.9.

Theorem 3.3. [25,43] *We have*

$$S_n^{\text{baj} - \text{inv}}(q) := \sum_{\sigma \in S_n} q^{\text{baj}(\sigma) - \text{inv}(\sigma)} = n \prod_{i=1}^{n-1} \frac{[i(n-i)]_q}{[i]_q}. \quad (18)$$

Corollary 3.4. *The d th cumulant κ_d^n for $\text{baj} - \text{inv}$ on S_n is*

$$\kappa_d^n = \frac{B_d}{d} \left(\sum_{i=1}^{n-1} [i(n-i)]^d - i^d \right).$$

Remark 3.5. Indeed, (18) holds with S_n replaced by $\{\sigma \in S_n : \sigma(n) = k\}$ for any fixed $k = 1, \dots, n$ if the factor of n is deleted from the right-hand side. See [47] for a bijective proof of this generalization. In addition, [43, Thm. 1.1] gives another generalization of the product formula (18) to all crystallographic Coxeter groups.

3.2. Combinatorial background for maj on W_α and $\text{SYT}(\underline{\lambda})$

Here we review standard combinatorial notions related to words, tableaux, and their major index generating functions.

Definition 3.6. Given a word $w = w_1 w_2 \cdots w_n$ with letters $w_i \in \mathbb{Z}_{\geq 1}$, the *type* of w is the sequence $\alpha = (\alpha_1, \alpha_2, \dots)$ where α_i is the number of times i appears in w . Such a sequence α is a (weak) *composition* of n , written as $\alpha \models n$. Trailing 0's are often omitted when writing weak compositions, so $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ for some m . Note that a word of type $(1, 1, \dots, 1) \models n$ is a permutation in the symmetric group S_n written in one-line notation. Just as for permutations, the *inversion number* of w is

$$\text{inv}(w) := \#\{(i, j) : i < j, w_i > w_j\}.$$

The *descent set* of w is

$$\text{Des}(w) := \{0 < i < n : w_i > w_{i+1}\},$$

and the *major index* of w is

$$\text{maj}(w) := \sum_{i \in \text{Des}(w)} i.$$

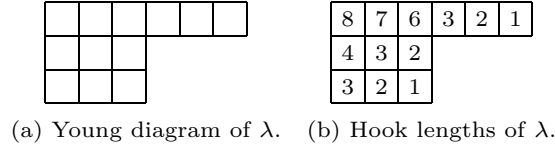


Fig. 2. Constructions related to the partition $\lambda = (6, 3, 3) \vdash 12$.

Definition 3.7. Let $\alpha = (\alpha_1, \dots, \alpha_m) \models n$. We use the following standard q -analogues:

$$\begin{aligned}
 [n]_q &:= 1 + q + \dots + q^{n-1} = \frac{q^n - 1}{q - 1}, & (q\text{-integer}) \\
 [n]_q! &:= [n]_q [n-1]_q \dots [1]_q, & (q\text{-factorial}) \\
 \binom{n}{k}_q &:= \frac{[n]_q!}{[k]_q! [n-k]_q!} \in \mathbb{Z}_{\geq 0}[q], & (q\text{-binomial}) \\
 \binom{n}{\alpha}_q &:= \frac{[n]_q!}{[\alpha_1]_q! \dots [\alpha_m]_q!} \in \mathbb{Z}_{\geq 0}[q] & (q\text{-multinomial}).
 \end{aligned}$$

Example 3.8. The identity statistic on the set $W = \{0, \dots, n-1\}$ has generating function $[n]_q$. The “sum” statistic on $W = \prod_{k=1}^n \{0, \dots, k-1\}$ has generating function $[n]_q!$.

For $\alpha \models n$, let W_α denote the words of type α . MacMahon’s classic result generalizing Theorem 3.2 in fact shows that maj and inv have the same distribution on W_α .

Theorem 3.9 ([27, Art. 6]). For each $\alpha \models n$,

$$W_\alpha^{\text{maj}}(q) = \binom{n}{\alpha}_q = W_\alpha^{\text{inv}}(q). \quad (19)$$

Definition 3.10. A composition $\lambda \models n$ such that $\lambda_1 \geq \lambda_2 \geq \dots$ is called a *partition* of n , written as $\lambda \vdash n$. The *size* of λ is $|\lambda| := n$ and the *length* $\ell(\lambda)$ of λ is the number of non-zero entries. The *Young diagram* of λ is the upper-left justified arrangement of unit squares called *cells* where the i th row from the top has λ_i cells following the English notation; see Fig. 2a. The *hook length* of a cell $c \in \lambda$ is the number h_c of cells in λ in the same row as c to the right of c and in the same column as c and below c , including c itself; see Fig. 2b. A *corner* of λ is any cell with hook length 1. A *bijective filling* of λ is any labeling of the cells of λ by the numbers $[n] = \{1, 2, \dots, n\}$.

Definition 3.11. A *skew partition* λ/ν is a pair of partitions (ν, λ) such that the Young diagram of ν is contained in the Young diagram of λ . The cells of λ/ν are the cells in the diagram of λ which are not in the diagram of ν , written $c \in \lambda/\nu$. We identify straight partitions λ with skew partitions λ/\emptyset where $\emptyset = (0, 0, \dots)$ is the empty partition. The *size* of λ/ν is $|\lambda/\nu| := |\lambda| - |\nu|$. The notions of bijective filling, hook lengths, and corners naturally extend to skew partitions as well.

Definition 3.12. Given a sequence of partitions $\underline{\lambda} = (\lambda^{(1)}, \dots, \lambda^{(m)})$, we identify the sequence with the *block diagonal skew partition* obtained by translating the Young diagrams of the $\lambda^{(i)}$ so that the rows and columns occupied by these components are

Many combinatorial statistics arise from sets indexed by more complicated objects than the positive integers, in which case one can “let $n \rightarrow \infty$ ” in many different ways. The following result due to Canfield, Janson, and Zeilberger illustrates a more interesting limit. Their result is characterized by the statistic $s(\alpha) := n - m$ where $\alpha = (\alpha_1, \dots, \alpha_\ell) \models n$ with $\max\{\alpha_i\} = m$.

Theorem 3.16. [35, Theorem 1.2] *Let $\alpha^{(1)}, \alpha^{(2)}, \dots$ be a sequence of compositions, possibly of differing lengths. Let \mathcal{X}_n be the inversion (or major index) statistic on words of type $\alpha^{(n)}$. Then $\mathcal{X}_1, \mathcal{X}_2, \dots$ is asymptotically normal if and only if*

$$s(\alpha^{(n)}) \rightarrow \infty.$$

Remark 3.17. Explorations equivalent to Theorem 3.16 appeared significantly earlier than [35] in other contexts, for instance [11, p. 127-128] and (in the two-letter case) [29]. See [6] for further discussion and references.

The cumulant formula for $\mathcal{X}_\lambda[\text{maj}]$, Theorem 1.5, follows immediately from Theorem 2.9 and Stanley’s formula (1). Adin and Roichman [1] had previously used (1) to compute the mean and variance of $\mathcal{X}_\lambda[\text{maj}]$ as

$$\mu = \frac{\binom{|\lambda|}{2} - b(\lambda') + b(\lambda)}{2} = b(\lambda) + \frac{1}{2} \left[\sum_{k=1}^{|\lambda|} k - \sum_{c \in \lambda} h_c \right],$$

and

$$\sigma^2 = \frac{1}{12} \left[\sum_{k=1}^{|\lambda|} k^2 - \sum_{c \in \lambda} h_c^2 \right].$$

The following common generalization of Stanley’s formula (1) and MacMahon’s formula, Theorem 3.9, is well known (e.g. see [42, (5.6)]). See [3, Thm. 2.15] for other applications.

Theorem 3.18. *Let $\underline{\lambda} = (\lambda^{(1)}, \dots, \lambda^{(m)})$ where $\lambda^{(i)} \vdash \alpha_i$ and $n = \alpha_1 + \dots + \alpha_m$. Then*

$$\text{SYT}(\underline{\lambda})^{\text{maj}}(q) = \binom{n}{\alpha_1, \dots, \alpha_m}_q \cdot \prod_{i=1}^m \text{SYT}(\lambda^{(i)})^{\text{maj}}(q). \quad (21)$$

Corollary 3.19. *Let κ_d^λ be the d th cumulant of maj on $\text{SYT}(\underline{\lambda})$ for $d > 1$. Then*

$$\kappa_d^\lambda = \frac{B_d}{d} \left(\sum_{k=1}^{|\underline{\lambda}|} k^d - \sum_{c \in \underline{\lambda}} h_c^d \right). \quad (22)$$

For general skew shapes, $\text{SYT}(\lambda/\nu)^{\text{maj}}(q)$ does not factor as a product of cyclotomic polynomials times q to a power. A “ q -Naruse” formula due to Morales–Pak–Panova, [31, (3.4)], gives an analogue of (1) involving a sum over “excited diagrams,” though the resulting sum has a single term precisely for the block diagonal skew partitions $\underline{\lambda}$.

4. Asymptotic normality for $\text{baj} - \text{inv}$ on S_n

We begin with a straightforward example which serves as a warmup and establishes some notation. See Section 3.1 for background. Asymptotic normality of $\text{baj} - \text{inv}$ on S_n follows from the cumulant formula in Corollary 3.4 by the following routine calculations. Recall that $a_n \sim b_n$ means that $\lim_{n \rightarrow \infty} a_n/b_n = 1$.

Lemma 4.1. *Fix $d \geq 1$. Then, as $n \rightarrow \infty$,*

$$\sum_{i=1}^{n-1} [i(n-i)]^d - i^d \sim n^{2d+1} \cdot \int_0^1 x^d (1-x)^d dx.$$

Proof. We have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^{n-1} [i(n-i)]^d - i^d}{n^{2d+1}} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{n-1} \left[\left(\frac{i}{n} \right)^d \left(1 - \frac{i}{n} \right)^d - \left(\frac{i}{n^2} \right)^d \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^d \left(1 - \frac{i}{n} \right)^d \\ &= \int_0^1 x^d (1-x)^d dx. \end{aligned}$$

Remark 4.2. The value of the integral in Lemma 4.1 is well known:

$$\int_0^1 x^d (1-x)^d dx = \frac{(d!)^2}{(2d+1)!} = \frac{1}{2d+1} \binom{2d}{d}^{-1}. \quad (23)$$

See [33, A002457] for a surprisingly large number of interpretations of the reciprocals of these values. Equation (23) is also a very special case of the Selberg integral formula [37], which has many interesting connections to algebraic combinatorics such as those in [39].

Corollary 4.3. *Fix $d \in \{1, 2, 4, 6, \dots\}$. Let κ_d^n be the d th cumulant of $\text{baj} - \text{inv}$ on S_n , and let κ_d^{n*} be the d th cumulant of the corresponding normalized random variable with mean 0 and variance 1. Then, uniformly for all n , we have*

$$|\kappa_d^{n*}| = \Theta(n^{1-d/2}). \quad (24)$$

That is, there are constants $c, C > 0$ depending only on d such that

$$cn^{1-d/2} \leq |\kappa_d^{n*}| \leq Cn^{1-d/2}.$$

Proof. It follows immediately from Corollary 3.4 and Lemma 4.1 that $|\kappa_d^n| = \Theta(n^{2d+1})$. Hence

$$|\kappa_d^{n*}| = |\kappa_d^n / (\kappa_2^n)^{d/2}| = \Theta(n^{2d+1-5d/2}) = \Theta(n^{1-d/2}).$$

Theorem 4.4. Let $\mathcal{X}_n = \mathcal{X}_{S_n}[\text{baj} - \text{inv}]$ be the random variable for the baj – inv statistic taken uniformly at random from S_n . Then, $\mathcal{X}_1, \mathcal{X}_2, \dots$ is asymptotically normal.

Proof. For fixed $d > 2$ even, we have $1 - d/2 < 0$, so by Corollary 4.3, $\kappa_d^{n*} \rightarrow 0$ as $n \rightarrow \infty$. The odd cumulants for $d > 2$ vanish since the odd Bernoulli numbers are 0. The result now follows from Corollary 2.17.

Remark 4.5. A key step in the above argument was to show that the variance σ_n^2 of baj – inv on S_n satisfies $\sigma_n^2 = \Theta(n^5)$. Indeed, the argument gives $\sigma_n^2 \sim n^5/360$. The weaker observation that $\sum_{i=1}^{n-1} [i(n-i)]^2$ is the dominant contribution to σ_n^2 is essentially enough to deduce asymptotic normality in this case. Our analysis of maj on standard tableaux includes non-normal limits, so more precise estimates like the above will become absolutely necessary. A straightforward modification of the above argument together with Theorem 3.2 also proves Theorem 1.1.

5. Asymptotic normality for maj on SYT(λ)

The main result of this section, Theorem 5.8, classifies the sequences of block diagonal skew partitions for which maj is asymptotically normal. We begin with a series of estimates for the differences $\sum_{k=1}^{|\lambda/\nu|} k^d - \sum_{c \in \lambda/\nu} h_c^d$, culminating in Corollary 5.7.

Definition 5.1. A reverse standard Young tableau of shape λ/ν is a bijective filling of λ/ν which strictly decreases along rows and columns. The set of reverse standard Young tableaux of shape λ/ν is denoted $\text{RSYT}(\lambda/\nu)$.

Lemma 5.2. Let $\lambda/\nu \vdash n$ and $T \in \text{RSYT}(\lambda/\nu)$. Then for all $c \in \lambda/\nu$,

$$T_c \geq h_c. \quad (25)$$

Furthermore, for any positive integer d ,

$$\sum_{j=1}^n j^d - \sum_{c \in \lambda/\nu} h_c^d = \sum_{c \in \lambda/\nu} (T_c^d - h_c^d) = \sum_{c \in \lambda/\nu} (T_c - h_c) \mathbf{h}_{d-1}(T_c, h_c), \quad (26)$$

where \mathbf{h}_{d-1} denotes the complete homogeneous symmetric function.

Proof. For (25), the entries in the hook of c form a subset of $\{1, 2, \dots, n\}$ of size h_c with maximum T_c , so $T_c \geq h_c$. Equation (26) follows immediately by rearranging the terms and factoring $(T_c^d - h_c^d) = (T_c - h_c) \sum_{k=0}^{d-1} T_c^{d-1-k} h_c^k$.

Lemma 5.3. *Let $\lambda/\nu \vdash n$ such that $\max_{c \in \lambda/\nu} h_c < 0.8n$. Let d be any positive integer. Then*

$$\frac{n^{d+1}}{26(d+1)} - 2(0.8)^d n^d < \sum_{j=1}^n j^d - \sum_{c \in \lambda/\nu} h_c^d < \frac{n^{d+1}}{d+1} + n^d.$$

Proof. Using Riemann sums for $\int_0^n x^d dx$, we obtain the bounds

$$\frac{n^{d+1}}{d+1} < \sum_{j=1}^n j^d < \frac{n^{d+1}}{d+1} + n^d \quad (27)$$

for all positive integers d, n . The upper bound in the lemma now follows immediately.

For the lower bound, label the cells of λ/ν by some $T \in \text{RSYT}(\lambda/\nu)$. By (25), $h_c \leq T_c$, and by assumption we have $h_c < 0.8n$ for all $c \in \lambda/\nu$. Considering the tighter of these two bounds on each summand and using (27) again, we have

$$\begin{aligned} \sum_{c \in \lambda/\nu} h_c^d &< \sum_{\substack{j \in [n] \\ j < 0.8n}} j^d + \sum_{\substack{j \in [n] \\ j \geq 0.8n}} (0.8n)^d \\ &< \frac{[0.8n]^{d+1}}{d+1} + [0.8n]^d + (n - [0.8n] + 1)(0.8n)^d \\ &\leq \frac{(0.8n)^{d+1}}{d+1} + 2(0.8n)^d + (0.2)(0.8)^d n^{d+1}. \end{aligned}$$

Consequently,

$$\begin{aligned} \sum_{j=1}^n j^d - \sum_{c \in \lambda/\nu} h_c^d &> \frac{n^{d+1}}{d+1} - \frac{(0.8n)^{d+1}}{d+1} - 2(0.8n)^d - (0.2)(0.8)^d n^{d+1} \\ &= \left(\frac{1}{d+1} (1 - (0.8)^{d+1}) - 0.2(0.8)^d \right) n^{d+1} - 2(0.8)^d n^d. \end{aligned}$$

It is easy to check that the coefficient on n^{d+1} is bounded below by $\frac{1}{26(d+1)}$ for all positive integers d . The result follows.

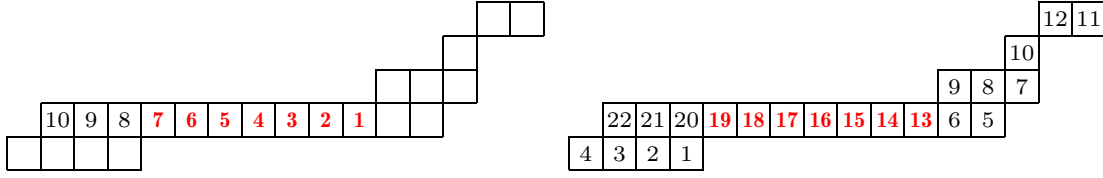


Fig. 5. On the left, the partially constructed $T \in \text{RSYT}(\lambda/\nu)$ after all the cells of R (in red and bold) have been filled. On the right, the final $T \in \text{RSYT}(\lambda/\nu)$. Here $\text{aft}(\lambda/\nu) = 10$.

Definition 5.4. Given any partition $\lambda/\nu \vdash n$, let the *aft* of λ/ν be the statistic

$$\text{aft}(\lambda/\nu) := n - \max_{c \in \lambda/\nu} \{\text{arm}(c), \text{leg}(c)\}$$

where $\text{arm}(c)$ is the number of cells in the same row as c to the right of c , including c itself, and $\text{leg}(c)$ is the number of cells in the same column as c below c , including c . When $\nu = \emptyset$, we have $\text{aft}(\lambda) = n - \max\{\lambda_1, \lambda'_1\}$ as above. When $\lambda/\nu = \underline{\lambda}$, we have $\text{aft}(\underline{\lambda}) = n - \max_i \{\lambda_1^{(i)}, \lambda_1^{(i)'}\}$. Note that $h_c = \text{arm}(c) + \text{leg}(c) - 1$.

Lemma 5.5. Let $\lambda/\nu \vdash n$ such that $\max_{c \in \lambda/\nu} h_c \geq 0.8n$, and let d be any positive integer. Furthermore, suppose $n \geq 10$. Then,

$$\text{aft}(\lambda/\nu) \frac{\lfloor 0.1n \rfloor^d}{d} \leq \sum_{j=1}^n j^d - \sum_{c \in \lambda/\nu} h_c^d \leq 2 \text{aft}(\lambda/\nu) (n^d + dn^{d-1}). \quad (28)$$

Proof. The result holds trivially if $\text{aft}(\lambda/\nu) = 0$ since in that case λ/ν is a single row or column, so assume $\text{aft}(\lambda/\nu) > 0$. Let $m \in \lambda/\nu$ have $h_m \geq 0.8n$, where we may assume m is the first cell in its row and column. For convenience, we may further assume by symmetry that $\text{arm}(m) \geq \text{leg}(m)$. Since $h_m \geq 0.8n$, it also follows that $\text{aft}(\lambda/\nu) = n - \text{arm}(m)$.

Now let R be the set of cells in the row of m , not including m itself, which are the only cells of λ/ν in their columns. Since λ/ν is a skew partition, R is connected. We claim that $\#R \geq 0.1n$. To prove the claim, we first observe that the hypothesis $h_m \geq 0.8n$ implies there are at most $n - h_m \leq 0.2n$ cells of λ/ν which could possibly be in the columns of the cells of the row of m not including m . Since $\text{arm}(m) \geq \text{leg}(m)$ and $\text{arm}(m) + \text{leg}(m) - 1 = h_m \geq 0.8n$, we have $\text{arm}(m) \geq 0.4n$. Hence no more than $0.2n$ of the $0.4n - 1$ cells in the row of m not including m can be excluded from R , so $\#R \geq 0.4n - 1 - 0.2n \geq 0.1n$ for $n \geq 10$.

Construct $T \in \text{RSYT}(\lambda/\nu)$ iteratively as follows; see Fig. 5 for an example. At each step of the iteration, we will first increment all existing labels by 1 and then label a new outer cell with 1. Begin by adding the cells of the row of m from left to right until the last cell of R has been added. Now add the remaining cells of λ/ν row by row starting at the topmost row and going from left to right. It is easy to see that the result respects the decreasing row and column conditions, so $T \in \text{RSYT}(\lambda/\nu)$.

Having established that $T_c - h_c \leq 2 \operatorname{aft}(\lambda/\nu)$, we now find by (26) and (27),

$$\begin{aligned} \sum_{k=1}^n k^d - \sum_{c \in \lambda/\nu} h_c^d &= \sum_{c \in \lambda/\nu} (T_c - h_c) \mathbf{h}_{d-1}(T_c, h_c) \\ &\leq \sum_{c \in \lambda/\nu} 2 \operatorname{aft}(\lambda/\nu) \mathbf{h}_{d-1}(T_c, T_c) \\ &= 2 \operatorname{aft}(\lambda/\nu) \sum_{j=1}^n d j^{d-1} \\ &< 2 \operatorname{aft}(\lambda/\nu) (n^d + d n^{d-1}). \end{aligned}$$

Corollary 5.6. *For fixed $d \in \mathbb{Z}_{\geq 1}$, uniformly for all skew shapes λ/ν ,*

$$\sum_{k=1}^{|\lambda/\nu|} k^d - \sum_{c \in \lambda/\nu} h_c^d = \Theta(\operatorname{aft}(\lambda/\nu) \cdot |\lambda/\nu|^d). \quad (29)$$

Proof. Let $n = |\lambda/\nu|$. When $\max_{c \in \lambda/\nu} h_c \geq 0.8n$, the result follows from Lemma 5.5. On the other hand, when $\max_{c \in \lambda/\nu} h_c < 0.8n$, then $n \geq \operatorname{aft}(\lambda/\nu) \geq 0.2n$, and the result follows from Lemma 5.3.

Corollary 5.7. *Fix d to be an even positive integer. Uniformly for all block diagonal skew shapes $\underline{\lambda}$, the absolute value of the normalized cumulant $|\kappa_d^{\underline{\lambda}*}|$ of $\mathcal{X}_{\underline{\lambda}}[\operatorname{maj}]$ is $\Theta(\operatorname{aft}(\underline{\lambda})^{1-d/2})$.*

Proof. For d even, by (22) and Corollary 5.6, we have

$$|\kappa_d^{\underline{\lambda}}| = \Theta(\operatorname{aft}(\underline{\lambda}) n^d),$$

where $n = |\underline{\lambda}|$. Consequently by the homogeneity of cumulants, we have

$$|\kappa_d^{\underline{\lambda}*}| = \left| \frac{\kappa_d^{\underline{\lambda}}}{(\kappa_2^{\underline{\lambda}})^{d/2}} \right| = \Theta \left(\frac{\operatorname{aft}(\underline{\lambda}) n^d}{\operatorname{aft}(\underline{\lambda})^{d/2} n^d} \right) = \Theta(\operatorname{aft}(\underline{\lambda})^{1-d/2}).$$

We now state and prove the generalization of Theorem 1.3 for the block diagonal skew shapes $\underline{\lambda}$ from Section 3.2.

Theorem 5.8. *Suppose $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \dots$ is a sequence of block diagonal skew partitions, and let $\mathcal{X}_N := \mathcal{X}_{\underline{\lambda}^{(N)}}[\operatorname{maj}]$ be the corresponding random variables for the maj statistic. Then, the sequence $\mathcal{X}_1, \mathcal{X}_2, \dots$ is asymptotically normal if and only if $\operatorname{aft}(\underline{\lambda}^{(N)}) \rightarrow \infty$ as $N \rightarrow \infty$.*

Proof. If $\text{aft}(\underline{\lambda}^{(N)}) \rightarrow \infty$, the result follows immediately from Corollary 2.17, Corollary 5.7, and the fact that the odd cumulants vanish. On the other hand, if $\text{aft}(\underline{\lambda}^{(N)}) \not\rightarrow \infty$, in the next section we will show that $\mathcal{X}_1^*, \mathcal{X}_2^*, \dots$ has a subsequence which converges to either a discrete or uniform-sum distribution, which in either case is non-normal.

Remark 5.9. Using work of Hwang–Zacharovas [24, Thm. 1.1], considering just the $d = 4$ case is sufficient to prove both directions of Theorem 5.8. However, the estimates we’ve given for κ_d^λ are strong enough to bound all the normalized cumulants simultaneously, and restricting to $d = 4$ (or even $d = 2$) does not simplify the argument.

6. Uniform sum limits for maj on SYT($\underline{\lambda}$)

The estimates from Section 5 apply when $\text{aft} \rightarrow \infty$. We next give an analogous estimate handling the case when aft is bounded, resulting in Theorem 6.2. We may then deduce Theorem 1.7 from the introduction and its generalization to block diagonal skew shapes, Theorem 6.3. Recall from Section 1 and Theorem 2.7 that \mathcal{IH}_M is the *Irwin–Hall* distribution obtained by adding M i.i.d. $\mathcal{U}[0, 1]$ random variables.

Lemma 6.1. Suppose $\lambda^{(N)}/\nu^{(N)} \vdash n_N$ is a sequence of skew partitions such that $\lim_{N \rightarrow \infty} n_N = \infty$ and

$$\lim_{N \rightarrow \infty} \text{aft}(\lambda^{(N)}/\nu^{(N)}) = M \in \mathbb{Z}_{\geq 0}. \quad (30)$$

Then for each fixed $d \in \mathbb{Z}_{\geq 1}$, we have

$$\lim_{N \rightarrow \infty} \frac{\sum_{k=1}^{n_N} k^d - \sum_{c \in \lambda^{(N)}/\nu^{(N)}} h_c^d}{M n_N^d} = 1. \quad (31)$$

Proof. Take N large enough so that $\text{aft}(\lambda^{(N)}/\nu^{(N)}) = M$ and $n_N \gg M$. Let $m \in \lambda^{(N)}/\nu^{(N)}$ be such that $\text{aft}(\lambda^{(N)}/\nu^{(N)}) = M = n_N - \text{arm}(m)$ so m is the first cell in its row and column, as in the proof of Lemma 5.5. Consider three regions of $\lambda^{(N)}/\nu^{(N)}$:

- (i) The rightmost $\text{arm}(m) - M = n_N - 2M$ cells in the row of m .
- (ii) The remaining leftmost M cells in the row of m .
- (iii) The remaining M cells in $\lambda^{(N)}/\nu^{(N)}$.

Construct $T \in \text{RSYT}(\lambda^{(N)}/\nu^{(N)})$ iteratively as in the proof of Lemma 5.5 as follows. First add cells in region (iii) row by row starting at the topmost row proceeding from left to right, stopping just before inserting the row of m . Next add the cells from region (ii) from left to right. Now add the remaining cells in region (iii) row by row starting at the row immediately below the row of m proceeding from left to right. Finally insert the cells from region (i) from left to right. It is easy to see that the cells in region (i) are

the lowest cells in their column, from which it follows that T indeed satisfies the column and row decreasing conditions.

We now consider the contributions of regions (i)-(iii) to the quotient

$$\frac{\sum_{k=1}^{n_N} k^d - \sum_{c \in \lambda^{(N)}/\nu^{(N)}} h_c^d}{Mn_N^d}.$$

Recall that $T_c - h_c$ can be interpreted as the number of times a cell inserted after cell c was not inserted in the same hook as c . It follows that $T_c - h_c = 0$ for region (i), leaving only contributions from the $2M$ cells in regions (ii) and (iii), a bounded sum. For region (ii), we have $T_c - h_c \leq M$, so that

$$T_c^d - h_c^d = (T_c - h_c) \mathbf{h}_{d-1}(T_c, h_c) \leq (2M)dn_N^{d-1}.$$

Dividing by Mn_N^d , cells in region (ii) contribute 0 to the sum in the limit. Finally, for region (iii), we find $1 \leq h_c \leq M+1$ and $n_N - 2M + 1 \leq T_c \leq n_N$, so that for each of the M cells c in region (iii),

$$(n_N - 2M + 1)^d - (M + 1)^d \leq T_c^d - h_c^d \leq n_N^d - 1^d.$$

Dividing by n_N^d , both bounds are asymptotic to 1 as $n_N \rightarrow \infty$. Adding up all M such contributions, the result follows.

Theorem 6.2. *Let $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \dots$ be a sequence of block diagonal skew partitions where $\lim_{N \rightarrow \infty} |\underline{\lambda}^{(N)}| = \infty$ and $\text{aft}(\underline{\lambda}^{(N)}) = M$ is constant. Let $\mathcal{X}_N := \mathcal{X}_{\underline{\lambda}^{(N)}}[\text{maj}]$ be the corresponding random variable for the maj statistic. Then $\mathcal{X}_1^*, \mathcal{X}_2^*, \dots$ converges in distribution to \mathcal{IH}_M^* .*

Proof. Using Equation (22) and Lemma 6.1, we have for $d \geq 2$ that

$$\begin{aligned} \lim_{N \rightarrow \infty} (\kappa_d^{\underline{\lambda}^{(N)}})^* &= \lim_{N \rightarrow \infty} \frac{\kappa_d^{\underline{\lambda}^{(N)}}}{(\kappa_d^{\underline{\lambda}^{(N)}})^{d/2}} \\ &= \lim_{N \rightarrow \infty} \frac{(B_d/d) \left(\sum_{k=1}^{n_N} k^d - \sum_{c \in \underline{\lambda}^{(N)}} h_c^d \right)}{(B_2/2)^{d/2} \left(\sum_{k=1}^{n_N} k^2 - \sum_{c \in \underline{\lambda}^{(N)}} h_c^2 \right)^{d/2}} \\ &= \lim_{N \rightarrow \infty} \frac{(B_d/d)}{(B_2/2)^{d/2}} \frac{Mn_N^d}{(Mn_N^2)^{d/2}} \\ &= \frac{(MB_d/d)}{(MB_2/2)^{d/2}}. \end{aligned}$$

From Theorem 2.7 and the homogeneity and additivity properties of cumulants, we have

$$\begin{aligned}
(\kappa_d^{\mathcal{IH}_M})^* &= \frac{\kappa_d^{\mathcal{IH}_M}}{(\kappa_2^{\mathcal{IH}_M})^{d/2}} \\
&= \frac{(MB_d/d)}{(MB_2/2)^{d/2}}.
\end{aligned}$$

The result now follows from Theorem 2.16 after converting moments to cumulants.

Theorem 6.3. *Let $\underline{\lambda}^{(1)}, \underline{\lambda}^{(2)}, \dots$ be a sequence of block diagonal skew partitions. Then the sequence $(\mathcal{X}_{\underline{\lambda}^{(N)}}[\text{maj}]^*)$ converges in distribution if and only if*

- (i) $\text{aft}(\underline{\lambda}^{(N)}) \rightarrow \infty$; or
- (ii) $|\underline{\lambda}^{(N)}| \rightarrow \infty$ and $\text{aft}(\underline{\lambda}^{(N)}) \rightarrow M < \infty$; or
- (iii) the distribution of $\mathcal{X}_{\underline{\lambda}^{(N)}}[\text{maj}]$ is eventually constant.

The limit law is \mathcal{N} in case (i), \mathcal{IH}_M^* in case (ii), and discrete in case (iii).

Proof. The backwards direction follows from Theorem 5.8 and Theorem 6.2. In the forwards direction, let $\underline{\lambda}^{(N)}$ be such a sequence where $(\mathcal{X}_{\underline{\lambda}^{(N)}}[\text{maj}]^*)$ converges in distribution. If $|\underline{\lambda}^{(N)}|$ is bounded, then there are only finitely many distinct $\underline{\lambda}^{(N)}$, forcing case (iii). If $|\underline{\lambda}^{(N)}|$ is unbounded, then we have subsequences satisfying either (i) or (ii) since the sequence converges in distribution, which from Theorem 5.8 and Theorem 6.2 gives convergence in distribution to \mathcal{N} or \mathcal{IH}_M^* , which are continuous, distinct distributions. The result follows.

From the Central Limit Theorem, we know the Irwin–Hall distribution \mathcal{IH}_M^* for M large closely resembles a normal distribution, so it will be quite rare for a plot of the coefficients of $\text{SYT}(\lambda)^{\text{maj}}(q)$ to look anything but normal. Since Irwin–Hall distributions are finitely supported, the difference between the two distributions is mainly in the tails. We note that even for $M = 5$, there is a close resemblance. See the plot in Fig. 7.

7. Discrete distributions for maj on SYT(λ)

We conclude by analyzing more carefully the discrete case of the limit law classification for maj on SYT(λ), Theorem 1.7. The result is Theorem 7.1, which lists several families of pairs of shapes λ and ν of differing sizes for which we nonetheless have $\#\text{SYT}(\lambda) = \#\text{SYT}(\nu)$.

A well-known corollary of (1) is that for partitions λ and ν of n , maj is equidistributed on SYT(λ) and SYT(ν) if and only if $b(\lambda) = b(\nu)$ and the multisets $\{h_c : c \in \lambda\}$ and $\{h_d : d \in \nu\}$ are equal. These hook multisets do not entirely characterize the partition—see [23]. The following theorem gives a similar result even if we consider the corresponding standardized random variables $\mathcal{X}_\lambda[\text{maj}]$ and $\mathcal{X}_\nu[\text{maj}]$.

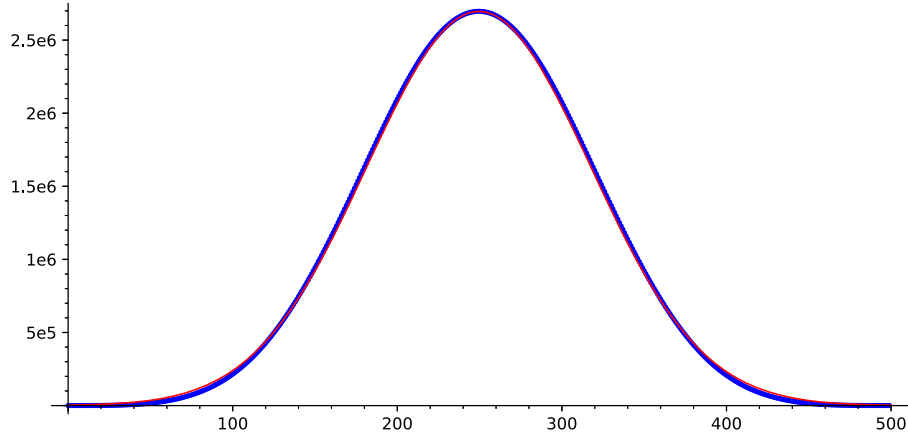


Fig. 7. Coefficients of $\text{SYT}(\lambda)^{\text{maj}}(q)$ for $\lambda = (100, 3, 2)$ where $\text{aft}(\lambda) = 5$ plotted in blue along with the corresponding normal distribution with the same mean and variance plotted in red. The difference is mostly in the tails.

Theorem 7.1. *Let λ and ν be partitions. Then $\mathcal{X}_\lambda[\text{maj}]^*$ and $\mathcal{X}_\nu[\text{maj}]^*$ have the same distribution if and only if*

- (i) *the multisets of hook lengths $\{h_c : c \in \lambda\}$ and $\{h_d : d \in \nu\}$ are equal; or*
- (ii) *the multisets $\{h_c : c \in \lambda\}$ and $\{|\lambda|\} \sqcup \{h_d : d \in \nu\}$ are equal; or*
- (iii) *λ and ν are each either a single row or column; or*
- (iv) *$\lambda, \nu \in \{(2, 1), (2, 2)\}$.*

Moreover, case (ii) occurs if and only if, up to transposing,

- (a) *$\lambda = (n)$ and $\nu = (n - 1)$ for $n \geq 2$; or*
- (b) *$\lambda = (r + 1, 1^{2r+2})$ and $\nu = (2^{r+1}, 1^r)$ for $r \geq 1$; or*
- (c) *$\lambda = (s, 1^{s+2})$ and $\nu = (s, s, 1)$ for $s \geq 4$; or*
- (d) *$\lambda = (3, 1^5)$ and $\nu = (3^2, 1)$, or $\lambda = (4, 1^6)$ and $\nu = (3^3, 1)$.*

Proof. Let $n := |\lambda|$ and $m := |\nu|$. Let $f^\lambda(q) = \frac{[n]_q!}{\prod_{c \in \lambda} [h_c]_q}$, which is a polynomial by (1) with constant coefficient 1. Let $f^\lambda = f^\lambda(1) = |\text{SYT}(\lambda)|$. Let f^ν and $f^\nu(q)$ be defined similarly.

In the backwards direction, if (i) holds, then $n = m$, both variances agree by Theorem 1.5, and $f^\lambda(q) = f^\nu(q)$, so $\mathcal{X}_\lambda[\text{maj}]^*$ and $\mathcal{X}_\nu[\text{maj}]^*$ have the same distribution. Similarly if (ii) holds $f^\lambda(q) = f^\nu(q)$, both variances agree, and $\mathcal{X}_\lambda[\text{maj}]^*$ and $\mathcal{X}_\nu[\text{maj}]^*$ have the same distribution again. Condition (iii) holds if and only if the distributions are concentrated at a single point. For (iv), we have $f^{(2,1)}(q) = 1 + q$ and $f^{(2,2)}(q) = 1 + q^2$, so the normalized distributions are clearly equal.

In the forwards direction, suppose $\mathcal{X}_\lambda[\text{maj}]^*$ and $\mathcal{X}_\nu[\text{maj}]^*$ have the same distribution. Since $f^\lambda(q)$ has constant coefficient 1, $\mathcal{X}_\lambda[\text{maj}]$ is concentrated at a single point if and only if $f^\lambda = 1$, which occurs if and only if λ is a single row or column which is covered by

case (iii). It is easy to see that $f^\lambda = 2$ if and only if $\lambda \in \{(2, 1), (2, 2)\}$ which is covered by case (iv).

Assume $f^\lambda, f^\nu > 2$. By [3, Thm. 1.1], it follows that $f^\lambda(q)$ and $f^\nu(q)$ each have two adjacent non-zero coefficients. Since $f^\lambda(q)$ and $f^\nu(q)$ each have constant term 1 and two adjacent non-zero coefficients, then it follows from the assumption $\mathcal{X}_\lambda[\text{maj}]^*$ and $\mathcal{X}_\nu[\text{maj}]^*$ have the same distribution that

$$f^\lambda(q) = \frac{[n]_q!}{\prod_{c \in \lambda} [h_c]_q} = \frac{[m]_q!}{\prod_{d \in \nu} [h_d]_q} = f^\nu(q). \quad (32)$$

Without loss of generality, we can assume $n \geq m$. If $n = m$, we have $\prod_{c \in \lambda} [h_c]_q = \prod_{d \in \nu} [h_d]_q$, from which it follows that the multisets of hook lengths are equal by considering multiplicities of zeros at all primitive roots of unity as in case (i).

From here on, assume $n > m$. The multiplicity of a zero of a primitive n th root of unity in (32) is 0 on the right, so from the left λ must have a hook of length n so it itself is a hook shape partition. Since λ is not a single row or column by the assumption $f^\lambda > 2$, we know λ does not have a cell with hook length $n - 1$. Consequently, the multiplicity of a zero at a primitive $(n - 1)$ th root of unity in (32) is 1 on the left, forcing $m = n - 1$ on the right. Thus (32) becomes

$$[m + 1]_q \prod_{d \in \nu} [h_d]_q = \prod_{c \in \lambda} [h_c]_q, \quad (33)$$

and as before the multiset condition (ii) must hold. This completes the proof of the first statement in the theorem.

For the second statement, suppose (ii) holds, so the multisets $\{h_c : c \in \lambda\}$ and $\{|\lambda|\} \sqcup \{h_d : d \in \nu\}$ are equal. Then, $m = n - 1$ and λ has a cell with hook length $|\lambda|$, so λ is a hook shape partition $(n - k, 1^k)$ for some $0 \leq k \leq n$, and

$$\{h_d : d \in \nu\} = [m - k] \sqcup [k]. \quad (34)$$

By transposing if necessary, we may assume $k \geq m - k$ is the maximum hook length in ν . If λ has one cell with hook length 1, then (a) holds. Otherwise, both λ and ν have precisely two cells with hook length 1, so ν is the union of two rectangles and not itself a rectangle. If ν were a hook, then it would have a hook length equal to m which would imply λ has a cell of hook length $m = n - 1$ contradicting the fact that λ has two outer corners. Thus ν is not itself a hook.

Transposing ν if necessary, we can assume its first two rows are equal, say $\nu_1 = \nu_2 = s$. If $\nu'_1 = \nu'_2$, one may check that the cell furthest from the origin in the intersection of the two rectangles forming ν would be the only cell of its hook length, and that moreover its two neighbors in the intersection would each have one larger hook length, contrary to (34). It follows that $\nu = (s^t, 1^r)$ where $r \geq 1$, $s \geq 2$, and $t \geq 2$. We now have several cases.

- If $s = 2$, the hook lengths of ν are $\{1, \dots, r, r+2, \dots, r+t+1, 1, \dots, t\}$. The “gap” between r and $r+2$ together with (34) forces $t = r+1$, so that $\nu = (2^{r+1}, 1^r)$ with $r \geq 1$. Here $k = r+t+1 = 2r+2$, resulting in case (b).
- If $s \geq 3$, the last two columns of ν already contain two cells with hook length 2. If $r > 1$, the first column would also have a cell with hook length 2, contradicting (34), so $r = 1$.
 - If $s = 3$, the hook lengths of ν are $\{1, \dots, t, 2, \dots, t+1, 1, 4, 5, \dots, t+3\}$. Because of the “gap” between $t+1$ and $t+3$, this is of the form in (34) if and only if $t = 2$ or $t = 3$, resulting in case (d).
 - Suppose $s > 3$. If $t \geq 3$, then the final three columns of ν contain three cells with hook length 3, contradicting (34), so $t = 2$. The hook lengths of ν are then $\{1, 1, 2, \dots, s-1, s+1, 2, 3, \dots, s, s+2\}$, which is already of the form (34), resulting in case (c).

The reverse implications from (a)-(d) to (ii) were verified in the course of the above argument.

Remark 7.2. The proof of Theorem 7.1 applies more generally to arbitrary scaling factors and translations of the distributions of $\mathcal{X}_\lambda[\text{maj}]$ and $X_\nu[\text{maj}]$, and not just those coming from means and variances.

8. Future work

We conjecture that almost all of the polynomials of the form $\text{SYT}(\lambda)^{\text{maj}}(q)$ are unimodal and log-concave. In this section, we discuss the deviations of each of these properties. In the rare cases where unimodality or log-concavity fails, it only seems to happen at the very beginning and end of the sequence of coefficients or near the middle coefficient.

Recall that a polynomial $P(q) = \sum_{i=0}^n c_i q^i$ is *unimodal* if

$$c_0 \leq c_1 \leq \dots \leq c_j \geq c_{j+1} \geq \dots \geq c_n$$

for some j , and $P(q)$ is *log-concave* if $c_i^2 \geq c_{i-1}c_{i+1}$ for all integers $0 < i < n$. A polynomial with nonnegative coefficients which is log-concave and has no internal zero coefficients is necessarily unimodal [41]. By [3], we know exactly where internal zeros occur so log-concavity would imply unimodality in these cases.

We say $P(q)$ is *nearly unimodal* if instead

$$c_0 \leq c_1 \leq \dots \leq c_j, c_{j+1} = c_j - 1 < c_{j+2} \leq \dots \leq c_{\lfloor \frac{n}{2} \rfloor}$$

for some j and $P(q)$ has symmetric coefficients. Also, a symmetric polynomial $P(q)$ is *nearly log-concave* if $c_i^2 \geq c_{i-1}c_{i+1}$ for all $1 < i < \lfloor \frac{n}{2} \rfloor$.

Conjecture 8.1. *The polynomial $\text{SYT}(\lambda)^{\text{maj}}(q)$ is unimodal if λ has at least 4 corners. If λ has 3 corners or fewer, then $\text{SYT}(\lambda)^{\text{maj}}(q)$ is unimodal except when λ or λ' is among the following partitions:*

1. Any partition of rectangle shape that has more than one row and column.
2. Any partition of the form $(k, 2)$ with $k \geq 4$ and k even.
3. Any partition of the form $(k, 4)$ with $k \geq 6$ and k even.
4. Any partition of the form $(k, 2, 1, 1)$ with $k \geq 2$ and k even.
5. Any partition of the form $(k, 2, 2)$ with $k \geq 6$.
6. Any partition on the list of 40 special exceptions:

$$\begin{aligned} & (3, 3, 2), (4, 2, 2), (4, 4, 2), (4, 4, 1, 1), (5, 3, 3), (7, 5), (6, 2, 1, 1, 1, 1), \\ & (5, 5, 2), (5, 5, 1, 1), (5, 3, 2, 2), (4, 4, 3, 1), (4, 4, 2, 2), (7, 3, 3), (8, 6), (6, 6, 2), \\ & (6, 6, 1, 1), (5, 5, 2, 2), (5, 3, 3, 3), (4, 4, 4, 2), (11, 5), (10, 6), (9, 7), (7, 7, 2), \\ & (7, 7, 1, 1), (6, 6, 4), (6, 6, 1, 1, 1, 1), (6, 5, 5), (5, 5, 3, 3), (12, 6), (11, 7), (10, 8), \\ & (15, 5), (14, 6), (11, 9), (16, 6), (12, 10), (18, 6), (14, 10), (20, 6), (22, 6). \end{aligned}$$

Conjecture 8.1 was checked for all partitions up to size $n = 50$. Each of the families $(k, 2)$, $(k, 4)$, or $(k, 2, 1, 1)$ have a relatively simple set of hook lengths so explicit formulas can be derived for the coefficients of $\text{SYT}(\lambda)^{\text{maj}}(q)$. We have found explicit proofs of near unimodality for each of these cases. They are related to known integer sequences [33, A266755] and [33, A008642] with nice generating functions. Furthermore, these families are all nearly unimodal as well as 20 of the special exceptions. All rectangles with at least 2 rows and columns are nearly unimodal for $30 \leq n \leq 100$. The only deviation occurs at $i = 1$ up to symmetry. We conjecture this trend also continues, hence the claim that all coefficients in $\text{SYT}(\lambda)^{\text{maj}}(q)$ are close to unimodal. The family $(k, 2, 2)$ is a bit further from being unimodal. The proof of the following result is omitted, but follows directly from a careful analysis of the hook lengths.

Proposition 8.2. *If $\lambda = (k, 2, 2)$ for any positive integer $k \geq 3$, then the maximal coefficient of $f^\lambda(q)$, say c_j , satisfies the equation $c_j = c_{j+1} + \text{floor}(k/6) + I(4 = (k \bmod 6))$ and $c_0 \leq c_1 \leq \dots \leq c_j$ and $j+1$ is the median nonzero coefficient. Here I is an indicator function which is 1 if true and 0 if false.*

Conjecture 8.3. *The polynomials $\text{SYT}(\lambda)^{\text{maj}}(q)$ are “nearly unimodal but not unimodal” for partitions λ or λ' in the following cases:*

1. Any partition of rectangle shape that has more than one row and column with more than 30 cells.
2. Any partition of the form $(k, 2)$ with $k \geq 4$ and k even.
3. Any partition of the form $(k, 4)$ with $k \geq 6$ and k even.

4. Any partition of the form $(k, 2, 1, 1)$ with $k \geq 2$ and k even.

Conjecture 8.3 was checked for all partitions of size up to $n = 100$. It also holds for the following 14 special exceptions:

$$(3, 3, 2), (4, 2, 2), (5, 3, 3), (7, 5), (6, 2, 1, 1, 1, 1), (5, 3, 2, 2), (4, 4, 3, 1), (7, 3, 3), (5, 3, 3, 3), \\ (11, 5), (6, 6, 1, 1, 1, 1), (6, 5, 5), (15, 5), (22, 6).$$

Log-concavity for the polynomials $\text{SYT}_\lambda^{\text{maj}}(q)$ appears to be harder to characterize. There are examples of partitions with even 5 corners which are not log-concave. For example $f^\lambda(q)$ for $\lambda = (9, 9, 7, 7, 5, 5, 3, 3, 2)$ is nearly log-concave but $c_1^2 = 4^2 = 16 < 17 = c_0 c_2$. The only deviation occurs at $i = 1$ up to symmetry. Thus, we summarize what we have observed in the following conjecture.

Conjecture 8.4. *The polynomials $\text{SYT}(\lambda)^{\text{maj}}(q)$ are almost always log-concave for partitions $\lambda \vdash n$ for large n (Fig. 8).*

This conjecture is based on the fact that the normal distribution is log-concave and the following evidence. The approximate probability that a uniformly chosen partition of n has the log-concave property $\mathbb{P}(\text{LC})$ and the corresponding probability for the nearly log-concave property $\mathbb{P}(\text{NLC})$ is given in Fig. 8:

n	30	40	50
$\mathbb{P}(\text{LC})$	0.6734475	0.7876426	0.8753587
$\mathbb{P}(\text{NLC})$	0.8003212	0.9204832	0.9688140

Fig. 8. Data supporting Conjecture 8.4.

By Theorem 1.3 and the conjectured claim that the coefficients of $\text{SYT}(\lambda)^{\text{maj}}(q)$ are unimodal or almost unimodal for large λ , one might hope that we could approximate the number of $T \in \text{SYT}(\lambda)$ with $\text{maj}(T) = k$ by the density function $f(k; \kappa_1^\lambda, \kappa_2^\lambda)$ for the normal distribution with mean κ_1^λ and variance κ_2^λ . We have the following conjectured bounds on such an approximation.

Conjecture 8.5. *Let $\lambda \vdash n$ be any partition. Uniformly for all n , for all integers k , we have*

$$|\mathbb{P}[X_\lambda[\text{maj}] = k] - f(k; \kappa_1^\lambda, \kappa_2^\lambda)| = O\left(\frac{1}{\sigma_\lambda \text{aft}(\lambda)}\right).$$

The conjecture has been verified for $25 < n \leq 50$ and $\text{aft}(\lambda) > 1$ with a constant of $1/9$, which is tight up to reasonable limits on computation in the sense that if it is changed to $1/10$ with the other constraints the same, it fails at $n = 50$.

Conjecture 8.6. *Asymptotic normality for general skew shapes and not just block diagonal skew shapes holds if and only if $\text{aft}(\lambda/\nu^{(N)}) \rightarrow \infty$ as $N \rightarrow \infty$, generalizing the result in Theorem 5.8.*

The argument in Section 5 proves that the “formal cumulants” associated with

$$\frac{[n]_q!}{\prod_{c \in \lambda/\mu} [h_c]_q}$$

exhibit asymptotic normality when $\text{aft}(\lambda/\mu) \rightarrow \infty$. However, this is only the first term in the general q -Naruse formula for $\text{SYT}(\lambda/\mu)^{\text{maj}}(q)$. One approach to Conjecture 8.6 would be to show the remaining terms are “appropriately negligible.”

Acknowledgments

We would like to thank Krzysztof Burdzy, Rodney Canfield, Persi Diaconis, Sergey Fomin, Pavel Galashin, Svante Janson, William McGovern, Andrew Ohana, Greta Panova, Mihael Perman, Martin Raič, Richard Stanley, Sheila Sundaram, Vasu Tewari, Lauren Williams, Alex Woo, and the anonymous referee for helpful discussions and suggestions related to this work.

References

- [1] Ron M. Adin, Yuval Roichman, Descent functions and random Young tableaux, *Combin. Probab. Comput.* 10 (3) (2001) 187–201.
- [2] Connor Ahlbach, Joshua P. Swanson, Refined cyclic sieving on words for the major index statistic, *European J. Combin.* 73 (2018) 37–60.
- [3] Sara C. Billey, Matjaž Konvalinka, Joshua P. Swanson, Tableaux posets and the fake degrees of coinvariant algebras, preprint, arXiv:1809.07386, Sep 2018.
- [4] Patrick Billingsley, *Probability and Measure*, third edition, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1995, A Wiley-Interscience Publication.
- [5] Miklós Bóna (Ed.), *Handbook of Enumerative Combinatorics*, Discrete Mathematics and Its Applications, Boca Raton, CRC Press, Boca Raton, FL, 2015.
- [6] E. Rodney Canfield, Svante Janson, Doron Zeilberger, Corrigendum to “The Mahonian probability distribution on words is asymptotically normal” [*Adv. in Appl. Math.* 46 (1–4) (2011) 109–124] [mr2794017], *Adv. in Appl. Math.* 49 (1) (2012) 77.
- [7] L. Carlitz, A combinatorial property of q -Eulerian numbers, *Amer. Math. Monthly* 82 (1975) 51–54.
- [8] William Y.C. Chen, Carol J. Wang, Larry X.W. Wang, The limiting distribution of the coefficients of the q -Catalan numbers, *Proc. Amer. Math. Soc.* 136 (11) (2008) 3759–3767.
- [9] Thomas Church, Benson Farb, Representation theory and homological stability, *Adv. Math.* 245 (2013) 250–314.
- [10] F.N. David, D.E. Barton, *Combinatorial Chance*, Hafner Publishing Co., New York, 1962.
- [11] Persi Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 11, Institute of Mathematical Statistics, Hayward, CA, 1988.
- [12] Richard Ehrenborg, Margaret Readdy, A poset view of the major index, *Adv. in Appl. Math.* 62 (2015) 1–14.
- [13] Paul Erdős, Joseph Lehner, The distribution of the number of summands in the partitions of a positive integer, *Duke Math. J.* 8 (1941) 335–345.
- [14] W. Feller, The fundamental limit theorems in probability, *Bull. Amer. Math. Soc.* 51 (1945) 800–832.

- [15] D. Foata, On the Netto inversion number of a sequence, *Proc. Amer. Math. Soc.* 19 (1968) 236.
- [16] J.S. Frame, G. de B. Robinson, R.M. Thrall, The hook graphs of the symmetric groups, *Canad. J. Math.* 6 (1954) 316–324.
- [17] Jason Fulman, The distribution of descents in fixed conjugacy classes of the symmetric groups, *J. Combin. Theory Ser. A* 84 (2) (1998) 171–180.
- [18] J. Fülring, J. Hofbauer, q -Catalan numbers, *J. Combin. Theory Ser. A* 40 (2) (1985) 248–264.
- [19] V.L. Goncharov, From the realm of combinatorics, *Izv. Akad. Nauk SSSR* 8 (1944) 3–48.
- [20] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics*, Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1989.
- [21] J. Haglund, M. Haiman, N. Loehr, A combinatorial formula for Macdonald polynomials, *J. Amer. Math. Soc.* 18 (3) (2005) 735–761.
- [22] L.H. Harper, Stirling behavior is asymptotically normal, *Ann. Math. Stat.* 38 (1967) 410–414.
- [23] Joan E. Herman, Fan R.K. Chung, Some results on hook lengths, *Discrete Math.* 20 (1) (1977/78) 33–40.
- [24] Hsien-Kuei Hwang, Vytas Zacharovas, Limit distribution of the coefficients of polynomials with only unit roots, *Random Structures Algorithms* 46 (4) (2015) 707–738.
- [25] N. Iwahori, H. Matsumoto, On some Bruhat decomposition and the structure of the Hecke rings of p -adic Chevalley groups, *Publ. Math. Inst. Hautes Études Sci.* 25 (1965) 5–48.
- [26] Gene B. Kim, Sangchul Lee, Central limit theorem for descents in conjugacy classes of S_n , *J. Combin. Theory Ser. A* 169 (2020) 105123, 13 pp.
- [27] P.A. MacMahon, The indices of permutations and the derivation therefrom of functions of a single variable associated with the permutations of any assemblage of objects, *Amer. J. Math.* 35 (3) (1913) 281–322.
- [28] P.A. MacMahon, Two applications of general theorems in combinatory analysis: (1) to the theory of inversions of permutations; (2) to the ascertainment of the numbers of terms in the development of a determinant which has amongst its elements an arbitrary number of zeros, *Proc. Lond. Math. Soc.* S2-15 (1) (1917) 314.
- [29] H.B. Mann, D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.* 18 (1947) 50–60.
- [30] Barry Mazur, Bernoulli numbers and the unity of mathematics: Bartlett lecture notes, Online. <http://www.math.harvard.edu/~mazur/papers/slides.Bartlett.pdf>, 2008.
- [31] Alejandro H. Morales, Igor Pak, Greta Panova, Hook formulas for skew shapes I. q -analogues and bijections, *J. Combin. Theory Ser. A* 154 (2018) 350–405.
- [32] Jonathan Novak, Piotr Śniady, What is ... a free cumulant? *Notices Amer. Math. Soc.* 58 (2) (2011) 300–301.
- [33] OEIS Foundation Inc., The on-line encyclopedia of integer sequences, Online. <http://oeis.org>, 2017.
- [34] V. Reiner, D. Stanton, D. White, The cyclic sieving phenomenon, *J. Combin. Theory Ser. A* 108 (1) (2004) 17–50.
- [35] E. Rodney Canfield, Svante Janson, Doron Zeilberger, The Mahonian probability distribution on words is asymptotically normal, *Adv. in Appl. Math.* 46 (1–4) (2011) 109–124.
- [36] Vladimir N. Sachkov, *Probabilistic Methods in Combinatorial Analysis*, Encyclopedia of Mathematics and Its Applications, vol. 56, Cambridge University Press, Cambridge, 1997. Translated from the Russian, Revised by the author.
- [37] Atle Selberg, Remarks on a multiple integral, *Norsk Mat. Tidsskr.* 26 (1944) 71–78.
- [38] John Shareshian, Michelle L. Wachs, Eulerian quasisymmetric functions, *Adv. Math.* 225 (6) (2010) 2921–2966.
- [39] Jang Soo Kim, Suho Oh, The Selberg integral and Young books, *J. Combin. Theory Ser. A* 145 (2017) 1–24.
- [40] R.P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999.
- [41] Richard P. Stanley, Log-concave and unimodal sequences in algebra, combinatorics, and geometry, in: *Graph Theory and Its Applications: East and West*, Jinan, 1986, in: *Ann. New York Acad. Sci.*, vol. 576, New York Acad. Sci., New York, 1989, pp. 500–535.
- [42] John R. Stembridge, On the eigenvalues of representations of reflection groups and wreath products, *Pacific J. Math.* 140 (2) (1989) 353–396.
- [43] John R. Stembridge, Debra J. Waugh, A Weyl group generating function that ought to be better known, *Indag. Math. (N.S.)* 9 (3) (1998) 451–457.
- [44] Joshua P. Swanson, On the existence of tableaux with given modular major index, *J. Algebraic Combin.* 1 (1) (2018) 3–21.

- [45] Marko Thiel, Nathan Williams, Strange expectations and the Winnie-the-Pooh problem, preprint, [arXiv:1811.02550](#), Nov 2018.
- [46] Wikipedia. Bernoulli number — Wikipedia, The Free Encyclopedia, 2017. [Online; accessed 21-July-2017].
- [47] Mike Zabrocki, A bijective proof of an unusual symmetric group generating function, preprint, [arXiv:math/0310301](#), 2003.