Limit Laws for q-Hook Formulas

Sara Billey
University of Washington

Based on joint work with:
Joshua Swanson

arXiv:2010.12701
Slides:
`math.washington.edu/~billey/talks/northeastern.2021.pdf`

Northeastern Representation Theory and Related Topics Seminar
April 30, 2021

Outline

Motivating Example: q-enumeration of SYT's via major index

Generalized q-hook length formulas

Moduli space of limiting distributions for SSYTs and forests

Open Problems

Standard Young Tableaux

Defn. A *standard Young tableau* of shape λ is a bijective filling of λ such that every row is increasing from left to right and every column is increasing from top to bottom.

1	3	6	7	9
2	5	8		
4				

Important Fact. The standard Young tableaux of shape λ, denoted $\mathrm{SYT}(\lambda)$, index a basis of the irreducible S_n representation indexed by λ.

Counting Standard Young Tableaux

Hook Length Formula. (Frame-Robinson-Thrall, 1954)
If λ is a partition of n, then

$$\#SYT(\lambda) = \frac{n!}{\prod_{c \in \lambda} h_c}$$

where h_c is the *hook length* of the cell c, i.e. the number of cells directly to the right of c or below c, including c.

Example. Filling cells of $\lambda = (5, 3, 1) \vdash 9$ by hook lengths:

7	5	4	2	1
4	2	1		
1				

So, $\#SYT(5, 3, 1) = \frac{9!}{7 \cdot 5 \cdot 4 \cdot 2 \cdot 4 \cdot 2} = 162.$

Remark. Notable other proofs by Greene-Nijenhuis-Wilf '79 (probabilistic), Eriksson '93 (bijective), Krattenthaler '95 (bijective), Novelli -Pak -Stoyanovskii'97 (bijective), Bandlow'08,

q-Counting Standard Young Tableaux

Def. The *descent set* of a standard Young tableau T, denoted $D(T)$, is the set of positive integers i such that $i+1$ lies in a row strictly below the cell containing i in T.

The *major index* of T is the sum of its descents:

$$\operatorname{maj}(T) = \sum_{i \in D(T)} i.$$

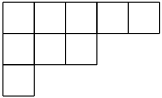
Example. The descent set of T is $D(T) = \{1, 3, 4, 7\}$ so $\operatorname{maj}(T) = 15$ for $T = \begin{array}{|c|c|c|c|c|}\hline 1 & 3 & 6 & 7 & 9 \\ \hline 2 & 4 & 8 \\ \cline{1-3} 5 \\ \cline{1-1}\end{array}$.

Def. The *major index generating function* for λ is

$$\operatorname{SYT}(\lambda)^{\operatorname{maj}}(q) := \sum_{T \in \operatorname{SYT}(\lambda)} q^{\operatorname{maj}(T)}$$

q-Counting Standard Young Tableaux

Example. $\lambda = (5, 3, 1)$



$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} =$

$$q^{23} + 2q^{22} + 4q^{21} + 5q^{20} + 8q^{19} + 10q^{18} + 13q^{17} + 14q^{16} + 16q^{15}$$

$$+16q^{14} + 16q^{13} + 14q^{12} + 13q^{11} + 10q^{10} + 8q^{9} + 5q^{8} + 4q^{7} + 2q^{6} + q^{5}$$

Note, at $q = 1$, we get back 162.

“Fast” Computation of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$

Thm.(Stanley's q-analog of the Hook Length Formula for $\lambda \vdash n$)

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \frac{q^{b(\lambda)}[n]_q!}{\prod_{c\in\lambda}[h_c]_q}$$

where

- $b(\lambda) := \sum(i-1)\lambda_i$
- h_c is the hook length of the cell c
- $[n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n-1}{q-1}$
- $[n]_q! := [n]_q[n-1]_q \cdots [1]_q$

“Fast” Computation of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$

Thm.(Stanley's q-analog of the Hook Length Formula for $\lambda \vdash n$)

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \frac{q^{b(\lambda)}[n]_q!}{\prod_{c \in \lambda}[h_c]_q}$$

where

- $b(\lambda) := \sum (i-1)\lambda_i$
- h_c is the hook length of the cell c
- $[n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n - 1}{q-1}$
- $[n]_q! := [n]_q[n-1]_q \cdots [1]_q$

The Trick. Each q-integer $[n]_q$ factors into a product of *cyclotomic polynomials* $\Phi_d(q)$,

$$[n]_q = 1 + q + \cdots + q^{n-1} = \prod_{d|n} \Phi_d(q).$$

Cancel all of the factors from the denominator of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ from the numerator, and then expand the remaining product.

Corollaries of Stanley's formula

Thm.(Stanley's q-analog of the Hook Length Formula for $\lambda \vdash n$)

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \frac{q^{b(\lambda)}[n]_q!}{\prod_{c\in\lambda}[h_c]_q}$$

Corollaries.

1. $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = q^{b(\lambda)-b(\lambda')}\,\mathrm{SYT}(\lambda')^{\mathrm{maj}}(q)$.
2. The coefficients of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ are symmetric.
3. There is a unique min-maj and max-maj tableau of shape λ.

Motivation for q-Counting Standard Young Tableaux

Thm.(Lusztig-Stanley 1979) Given a partition $\lambda \vdash n$, say

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum_{k \geq 0} b_{\lambda,k} q^k.$$

Then $b_{\lambda,k} := \#\{T \in \mathrm{SYT}(\lambda) : \mathrm{maj}(T) = k\}$ is the number of times the irreducible S_n module indexed by λ appears in the decomposition of the coinvariant algebra $\mathbb{Z}[x_1, x_2, \ldots, x_n]/I_+$ in the homogeneous component of degree k.

Key Questions for $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$

Recall $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum b_{\lambda,k} q^k$.

Existence Question. For which λ, k does $b_{\lambda,k} = 0$?

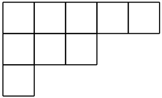
Distribution Question. What patterns do the coefficients in the list $(b_{\lambda,0}, b_{\lambda,1}, \ldots)$ exhibit?

Unimodality Question. For which λ, are the coefficients of $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ *unimodal*, meaning

$$b_{\lambda,0} \leq b_{\lambda,1} \leq \ldots \leq b_{\lambda,m} \geq b_{\lambda,m+1} \geq \ldots ?$$

q-Counting Standard Young Tableaux

Example. $\lambda = (5, 3, 1)$



$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum b_{\lambda,k} q^k =$$

$$q^{23} + 2q^{22} + 4q^{21} + 5q^{20} + 8q^{19} + 10q^{18} + 13q^{17} + 14q^{16} + 16q^{15}$$

$$+16q^{14} + 16q^{13} + 14q^{12} + 13q^{11} + 10q^{10} + 8q^{9} + 5q^{8} + 4q^{7} + 2q^{6} + q^{5}$$

Notation: (00000 1 2 4 5 8 10 13 14 16 16 16 14 13 10 8 5 4 2 1)

q-Counting Standard Young Tableaux

Examples. $(2,2) \vdash 4$: (0 0 1 0 1)

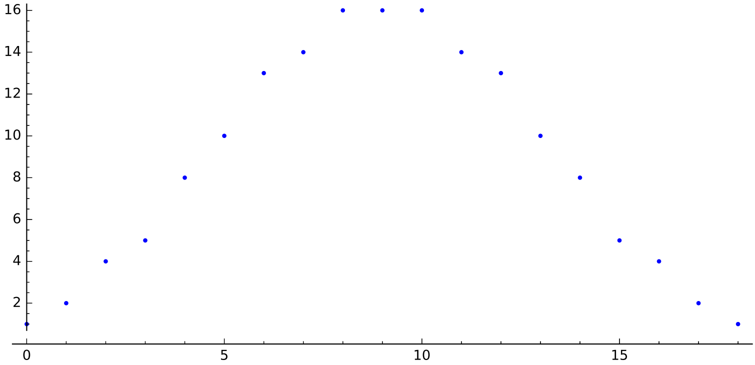
$(5,3,1)$: (00000 1 2 4 5 8 10 13 14 16 16 16 14 13 10 8 5 4 2 1)

$(6,4) \vdash 10$: (0 0 0 0 1 1 2 2 4 4 6 6 8 7 8 7 8 6 6 4 4 2 2 1 1)

$(6,6) \vdash 12$: (0 0 0 0 0 0 1 0 1 1 2 2 4 3 5 5 7 6 9 7 9 8 9 7 9 6 7 5 5 3 4 2 2 1 1 0 1)

$(11,5,3,1) \vdash 20$: (1 3 8 16 32 57 99 160 254 386 576 832 1184 1645 2255 3031 4027 5265 6811 8689 10979 13706 16959 20758 25200 30296 36143 42734 50163 58399 67523 77470 88305 99925 112370 125492 139307 153624 168431 183493 198778 214017 229161 243913 258222 271780 284542 296200 306733 315853 323571 329629 334085 336727 337662 336727 334085 329629 323571 315853 306733 296200 284542 271780 258222 243913 229161 214017 198778 183493 168431 153624 139307 125492 112370 99925 88305 77470 67523 58399 50163 42734 36143 30296 25200 20758 16959 13706 10979 8689 6811 5265 4027

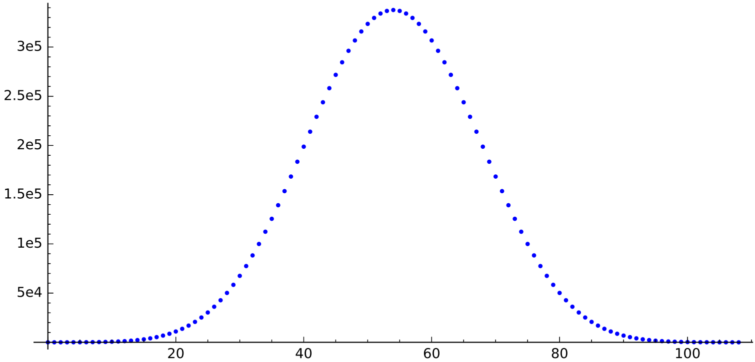
Visualizing Major Index Generating Functions



Visualizing the coefficients of $\mathrm{SYT}(5,3,1)^{\mathrm{maj}}(q)$:

$$(1,2,4,5,8,10,13,14,16,16,16,14,13,10,8,5,4,2,1)$$

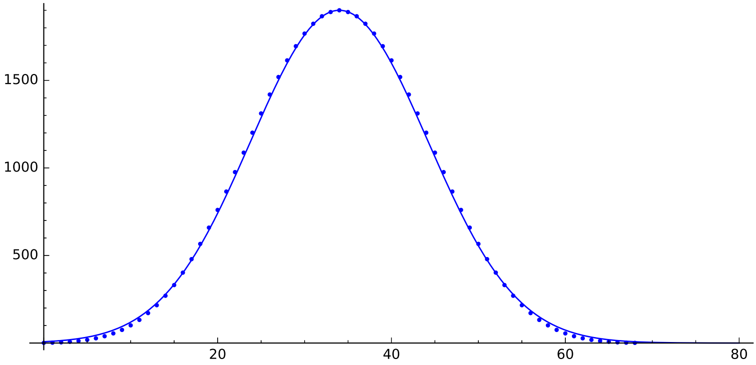
Visualizing Major Index Generating Functions



Visualizing the coefficients of $\mathrm{SYT}(11,5,3,1)^{\mathrm{maj}}(q)$.

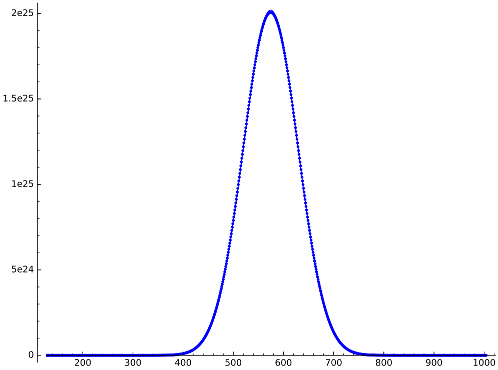
Question. What type of curve is that?

Visualizing Major Index Generating Functions



Visualizing the coefficients of $\mathrm{SYT}(10,6,1)^{\mathrm{maj}}(q)$ along with the Normal distribution with $\mu = 34$ and $\sigma^2 = 98$.

Visualizing Major Index Generating Functions



Visualizing the coefficients of $\mathrm{SYT}(8,8,7,6,5,5,5,2,2)^{\mathrm{maj}}(q)$ along with the corresponding normal distribution.

Existence Question: Classifying All Nonzero Fake Degrees

Recall $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum b_{\lambda,k} q^k$.

Existence Question. For which λ, k does $b_{\lambda,k} = 0$?

Existence Question: Classifying All Nonzero Fake Degrees

Recall $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)} = \sum b_{\lambda,k} q^k$.

Existence Question. For which λ, k does $b_{\lambda,k} = 0$?

Thm.(Billey-Konvalinka-Swanson, 2018)
For any partition λ which is not a rectangle,

$$\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) := \sum_{T \in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)}$$

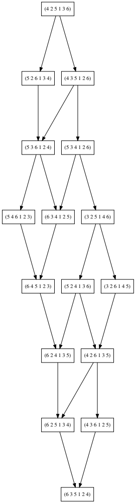
has no internal zeros. If λ is a rectangle with at least two rows and columns, $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$ has exactly two internal zeros, one at degree $b(\lambda) + 1$ and the other at degree $maxmaj(\lambda) - 1$.

Classifying All Nonzero Fake Degrees

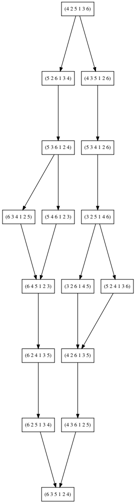
Proof Outline. We identify block and rotation rules on tableaux giving rise to two posets on $\mathrm{SYT}(\lambda)$– exceptional cases for rectangles which is ranked according to maj.

Note, these posets are different from those considered by Taskin, Poirier-Reutenuer, and Kosakowska-Schmidmeier-Thomas.

Strong and Weak Poset on $\mathrm{SYT}(3, 2, 1)$



Strong



Weak

Classifying All Nonzero Fake Degrees

Cor. The irreducible S_n-module indexed by λ appears in the decomposition of the degree k component of the coinvariant algebra if and only if $b_{\lambda,k} > 0$ as characterized above.

Similar results hold for all Shepard-Todd groups $G(m, d, n)$.

See arXiv:1809.07386 for more details.

Converting q-Enumeration to Discrete Probability

Distribution Question. What is the limiting distribution(s) for the coefficients in $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q)$?

From Combinatorics to Probability.
If $f(q) = a_0 + a_1 q + a_2 q^2 + \cdots + a_n q^n$ where a_i are nonnegative integers, then construct the random variable X_f with discrete probability distribution

$$\mathbb{P}(X_f = k) = \frac{a_k}{\sum_j a_j} = \frac{a_k}{f(1)}.$$

If f is part of a family of q-analog of an integer sequence, we can study the limiting distributions.

Converting q-Enumeration to Discrete Probability

Example. For $\mathrm{SYT}(\lambda)^{\mathrm{maj}}(q) = \sum b_{\lambda,k} q^k$, define the integer random variable $X_\lambda[\mathrm{maj}]$ with discrete probability distribution

$$\mathbb{P}(X_\lambda[\mathrm{maj}] = k) = \frac{b_{\lambda,k}}{|\mathrm{SYT}(\lambda)|}.$$

We claim the distribution of $X_\lambda[\mathrm{maj}]$ "usually" is approximately normal for most shapes λ. Let's make that precise!

Standardization

Thm.(Adin-Roichman, 2001)
For any partition λ, the mean and variance of $X_\lambda[\mathrm{maj}]$ are

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

and

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

Def. The *standardization* of $X_\lambda[\mathrm{maj}]$ is

$$X_\lambda^*[\mathrm{maj}] = \frac{X_\lambda[\mathrm{maj}] - \mu_\lambda}{\sigma_\lambda}.$$

So $X_\lambda^*[\mathrm{maj}]$ has mean 0 and variance 1 for any λ.

Asymptotic Normality

Def. Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables with standardized cumulative distribution functions $F_1(t), F_2(t), \ldots$. The sequence is *asymptotically normal* if

$$\forall t \in \mathbb{R}, \quad \lim_{n\to\infty} F_n(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} = \mathbb{P}(N < t)$$

where N is a Normal random variable with mean 0 and variance 1.

Asymptotic Normality

Def. Let $X_1, X_2, \ldots$ be a sequence of real-valued random variables with standardized cumulative distribution functions $F_1(t), F_2(t), \ldots$. The sequence is *asymptotically normal* if

$$\forall t \in \mathbb{R}, \quad \lim_{n\to\infty} F_n(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} = \mathbb{P}(N < t)$$

where N is a Normal random variable with mean 0 and variance 1.

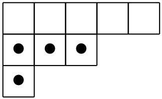
Question. In what way can a sequence of partitions approach infinity?

The Aft Statistic

Def. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$, let

$$\text{aft}(\lambda) := n - \max\{\lambda_1, k\}.$$

Example. $\lambda = (5, 3, 1)$ then $\text{aft}(\lambda) = 4$.



Look it up: Aft is now on FindStat as St001214

Distribution Question: From Combinatorics to Probability

Thm.(Billey-Konvalinka-Swanson, 2019)

Suppose $\lambda^{(1)}, \lambda^{(2)}, \ldots$ is a sequence of partitions, and let $X_N := X_{\lambda^{(N)}}[\mathrm{maj}]$ be the corresponding random variables for the maj statistic. Then, the sequence $X_1, X_2, \ldots$ is asymptotically normal if and only if $\mathrm{aft}(\lambda^{(N)}) \to \infty$ as $N \to \infty$.

Distribution Question: From Combinatorics to Probability

Thm.(Billey-Konvalinka-Swanson, 2019)

Suppose $\lambda^{(1)}, \lambda^{(2)}, \ldots$ is a sequence of partitions, and let $X_N := X_{\lambda^{(N)}}[\mathrm{maj}]$ be the corresponding random variables for the maj statistic. Then, the sequence $X_1, X_2, \ldots$ is asymptotically normal if and only if $\mathrm{aft}(\lambda^{(N)}) \to \infty$ as $N \to \infty$.

Question. What happens if $\mathrm{aft}(\lambda^{(N)})$ does not go to infinity as $N \to \infty$?

Distribution Question: From Combinatorics to Probability

Thm.(Billey-Konvalinka-Swanson, 2019)
Let $\lambda^{(1)}, \lambda^{(2)}, \ldots$ be a sequence of partitions. Then $(X_{\lambda^{(N)}}[\mathrm{maj}]^*)$ converges in distribution if and only if

(i) $\mathrm{aft}(\lambda^{(N)}) \to \infty$; or

(ii) $|\lambda^{(N)}| \to \infty$ and $\mathrm{aft}(\lambda^{(N)})$ is eventually constant; or

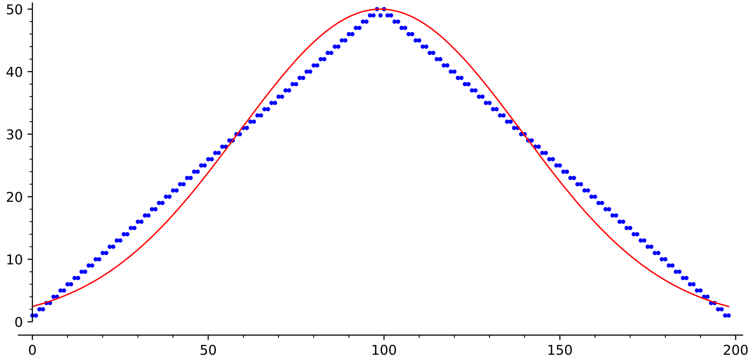
(iii) the distribution of $X^*_{\lambda^{(N)}}[\mathrm{maj}]$ is eventually constant.

The limit law is $\mathcal{N}(0,1)$ in case (i), $\mathcal{IH}^*_M$ in case (ii), and discrete in case (iii).

Here $\mathcal{IH}_M$ denotes the sum of M independent identically distributed uniform $[0,1]$ random variables, known as the Irwin–Hall distribution or the *uniform sum distribution*.

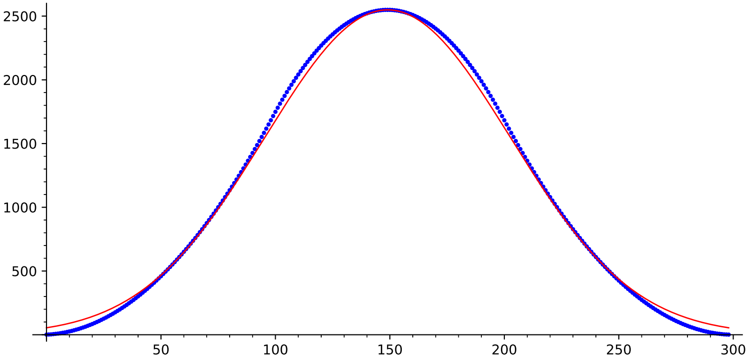
Distribution Question: From Combinatorics to Probability

Example. $\lambda = (100, 2)$ looks like the distribution of the sum of two independent uniform random variables on $[0, 1]$:



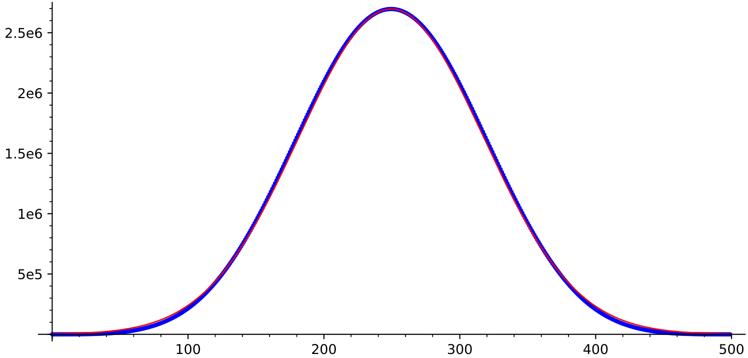
Distribution Question: From Combinatorics to Probability

Example. $\lambda = (100, 2, 1)$ looks like the distribution of the sum of three independent uniform random variables on $[0, 1]$:



Distribution Question: From Combinatorics to Probability

Example. $\lambda = (100, 3, 2)$ looks like the normal distribution, but not quite!



Proof ideas: Characterize the Moments and Cumulants

Definitions.

- For $d \in \mathbb{Z}_{\geq 0}$, the *dth moment*

$$\mu_d := \mathbb{E}[X^d]$$

- The *moment-generating function* of X is

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

- The *cumulants* $\kappa_1, \kappa_2, \ldots$ of X are defined to be the coefficients of the exponential generating function

$$K_X(t) := \sum_{d=1}^{\infty} \kappa_d \frac{t^d}{d!} := \log M_X(t) = \log \mathbb{E}[e^{tX}].$$

Nice Properties of Cumulants

1. *(Familiar Values)* The first two cumulants are $\kappa_1 = \mu$, and $\kappa_2 = \sigma^2$.
2. *(Shift Invariance)* The second and higher cumulants of X agree with those for $X - c$ for any $c \in \mathbb{R}$.
3. *(Homogeneity)* The dth cumulant of cX 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 and moments are determined by polynomials in the other sequence.

Examples of Cumulants and Moments

Example. Let $X = \mathcal{N}(\mu, \sigma^2)$ be the normal random variable with mean μ and variance σ^2. Then the cumulants are

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

and for $d > 1$,

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

.

Example. For a Poisson random variable X with mean μ, the cumulants are all $\kappa_d = \mu$, while the moments are $\mu_d = \sum_{i=1}^{d} \mu_i S_{i,d}$.

Cumulants for Major Index Generating Functions

Thm.(Billey-Konvalinka-Swanson, 2019)
Let $\lambda \vdash n$ and $d \in \mathbb{Z}_{>1}$. If κ_d^λ is the dth cumulant of $X_\lambda[\text{maj}]$, then

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

where $B_0, B_1, B_2, \ldots = 1, \frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, \ldots$ are the Bernoulli numbers (OEIS A164555 / OEIS A027642).

Remark. We use this theorem to prove that as aft approaches infinity the standardized cumulants for $d \geq 3$ all go to 0 proving the Asymptotic Normality Theorem.

Cumulants for Major Index Generating Functions

Thm.(Billey-Konvalinka-Swanson, 2019)
Let $\lambda \vdash n$ and $d \in \mathbb{Z}_{>1}$. If κ_d^λ is the dth cumulant of $X_\lambda[\mathrm{maj}]$, then

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

where $B_0, B_1, B_2, \ldots = 1, \frac{1}{2}, \frac{1}{6}, 0, -\frac{1}{30}, 0, \frac{1}{42}, 0, \ldots$ are the Bernoulli numbers (OEIS A164555 / OEIS A027642).

Remark. We use this theorem to prove that as aft approaches infinity the standardized cumulants for $d \geq 3$ all go to 0 proving the Asymptotic Normality Theorem.

Remark. Note, κ_2^λ is exactly the Adin-Roichman variance formula.

Cumulants of certain q-analogs

Thm.(Chen–Wang–Wang-2008 and Hwang–Zacharovas-2015)
Suppose $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ are multisets of positive integers such that

$$f(q) = \frac{\prod_{j=1}^m [a_j]_q}{\prod_{j=1}^m [b_j]_q} = \sum c_k q^k \in \mathbb{Z}_{\geq 0}[q]$$

.

Let X be a discrete random variable with $\mathbb{P}(X = k) = c_k / f(1)$.
Then the dth cumulant of X is

$$\kappa_d = \frac{B_d}{d} \sum_{j=1}^m (a_j^d - b_j^d)$$

where B_d is the dth Bernoulli number (with $B_1 = \frac{1}{2}$).

Example. This theorem applies to

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

Cyclotomic Generating Functions

Def. A polynomial $f(q)$ with nonnegative integer coefficients is a *cyclotomic generating function* provided it satisfies one of the following equivalent conditions:

(i) (Rational form.) There are multisets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of positive integers and $\alpha, \beta \in \mathbb{Z}_{\geq 0}$ such that

$$f(q) = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{1 - q^{a_j}}{1 - q^{b_j}}. \qquad (2)$$

Cyclotomic Generating Functions

Def. A polynomial $f(q)$ with nonnegative integer coefficients is a *cyclotomic generating function* provided it satisfies one of the following equivalent conditions:

(i) (Rational form.) There are multisets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of positive integers and $\alpha, \beta \in \mathbb{Z}_{\geq 0}$ such that

$$f(q) = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{1 - q^{a_j}}{1 - q^{b_j}}. \qquad (2)$$

(ii) (Cyclotomic form.) The polynomial $f(q)$ can be written as a non-negative integer times a product of cyclotomic polynomials and factors of q.

Cyclotomic Generating Functions

Def. A polynomial $f(q)$ with nonnegative integer coefficients is a *cyclotomic generating function* provided it satisfies one of the following equivalent conditions:

(i) (Rational form.) There are multisets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of positive integers and $\alpha, \beta \in \mathbb{Z}_{\geq 0}$ such that

$$f(q) = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} = \alpha q^{\beta} \cdot \prod_{j=1}^{m} \frac{1 - q^{a_j}}{1 - q^{b_j}}. \tag{2}$$

(ii) (Cyclotomic form.) The polynomial $f(q)$ can be written as a non-negative integer times a product of cyclotomic polynomials and factors of q.

(iii) (Complex form.) The complex roots of $f(q)$ are each either a root of unity or zero.

Cyclotomic Generating Functions

More examples of cyclotomic generating functions, aka q-hook length type formulas..

1. Stanley: $s_\lambda(1, q, q^2, \ldots, q^m)$.
2. Björner-Wachs: q-hook length formula for forests.
3. Macaulay: Hilbert series of polynomial quotients $k[x_1, \ldots, x_n]/(\theta_1, \theta_2, \ldots, \theta_n)$ where $deg(x_i) = b_i$, $deg(\theta_i) = a_i$, and $(\theta_1, \theta_2, \ldots, \theta_n)$ is a homogeneous system of parameters.
4. Chevalley: Length generating function restricted to minimum length coset representatives of a finite reflection group modulo a parabolic subgroup.
5. Iwahori-Matsumoto, Stembridge-Waugh, Zabrocki: Coxeter length generating function restricted to coset representatives of the extended affine Weyl group of type A_{n-1} mod translations by coroots. The associated statistic is $\mathrm{baj} - \mathrm{inv}$.

Cyclotomic Generating Functions

Remark. Corresponding with each cyclotomic generating function $f(q)$, there is a discrete random variable X_f supported on $\mathbb{Z}_{\geq 0}$ with probability generating function $f(q)/f(1)$ and higher cumulants for $d \geq 2$,

$$\kappa_d^f = \frac{B_d}{d} \sum_{j=1}^{m} (a_j^d - b_j^d).$$

Therefore, we can study asymptotics for interesting sequences of cyclotomic generating functions much like SYT.

Recent Progress based on joint work with Josh Swanson

1. MacMahon: q-counting plane partitions in box.
2. Stanley-Littlewood: $s_\lambda(1, q, q^2, \ldots, q^m)$.
3. Björner-Wachs: q-hook length formula for forests

MacMahon: *q*-counting plane partitions in box.

Let $\mathrm{PP}(a \times b \times c)$ be the set of all *plane partitions* that fit inside an $a \times b \times c$ box. Plane partitions can be represented by tableaux with decreasing rows and columns. The *size* of a plane partition is the sum of the numbers in the tableau.

MacMahon's Formula.

$$\sum_{T \in \mathrm{PP}(a \times b \times c)} q^{|T|} = \prod_{i=1}^{a} \prod_{j=1}^{b} \prod_{k=1}^{c} \frac{[i+j+k-1]_q}{[i+j+k-2]_q}.$$

MacMahon: q-counting plane partitions in box.

Let $\mathrm{PP}(a \times b \times c)$ be the set of all *plane partitions* that fit inside an $a \times b \times c$ box. Plane partitions can be represented by tableaux with decreasing rows and columns. The *size* of a plane partition is the sum of the numbers in the tableau.

MacMahon's Formula.

$$\sum_{T \in \mathrm{PP}(a\times b\times c)} q^{|T|} = \prod_{i=1}^{a} \prod_{j=1}^{b} \prod_{k=1}^{c} \frac{[i+j+k-1]_q}{[i+j+k-2]_q}.$$

MacMahon's Formula is a cyclotomic generating function. Let $\mathcal{X}_{a\times b\times c}[\mathrm{size}]^*$ the corresponding random variable.

Recent Progress based on joint work with Josh Swanson

Recall, $\mathcal{N}(0,1)$ is the standard normal distribution, and $\mathcal{IH}_M = \sum_{i=1}^{M} \mathcal{U}[0,1]$ is the Irwin-Hall distribution.

Theorem. Let a, b, c each be a sequence of positive integers.

(i) $\mathcal{X}_{a\times b\times c}[\text{size}]^* \Rightarrow \mathcal{N}(0,1)$ if and only if $\text{median}\{a,b,c\} \to \infty$.

(ii) $\mathcal{X}_{a\times b\times c}[\text{size}]^* \Rightarrow \mathcal{IH}_M$ if $ab \to M < \infty$ and $c \to \infty$.

The limit of the median value determines the limiting distribution for plane partitions, just like aft determined the limiting distribution for *SYTs*.

Moduli space of standardized distributions

Motivating Philosophy. By the Central Limit Theorem, $\lim_{M\to\infty}\mathcal{IH}_M^* \Rightarrow \mathcal{N}(0,1)$, so instead of parametrizing the Irwin-Hall distributions by $\{n \in \mathbb{Z}_{\geq 1}\}$, use the parameter space

$$\mathbf{P}_{\mathcal{IH}} := \left\{\frac{1}{n} : n \in \mathbb{Z}_{\geq 1}\right\} \subset \mathbb{R}$$

to get a related topological structure.

Moduli space of standardized distributions

Motivating Philosophy. By the Central Limit Theorem, $\lim_{M\to\infty} \mathcal{IH}_M^* \Rightarrow \mathcal{N}(0,1)$, so instead of parametrizing the Irwin-Hall distributions by $\{n \in \mathbb{Z}_{\geq 1}\}$, use the parameter space

$$\mathbf{P}_{\mathcal{IH}} := \left\{ \frac{1}{n} : n \in \mathbb{Z}_{\geq 1} \right\} \subset \mathbb{R}$$

to get a related topological structure.

Def. The *moduli space of Irwin-Hall distributions* is

$$\mathbf{M}_{\mathcal{IH}} := \{ \mathcal{IH}_M^* : M \in \mathbb{Z}_{\geq 0} \},$$

Endow $\mathbf{M}_{\mathcal{IH}}$ with the topology characterized by convergence in distribution using the Lévy metric.

Moduli space of standardized distributions

Conclusions.

1. $\overline{\mathbf{P}_{\mathcal{IH}}} = \mathbf{P}_{\mathcal{IH}} \sqcup \{0\}$.

2. $\overline{\mathbf{M}_{\mathcal{IH}}} = \mathbf{M}_{\mathcal{IH}} \cup \{\mathcal{N}(0,1)\}$.

3. The bijection $\overline{\mathbf{P}_{\mathcal{IH}}} \to \overline{\mathbf{M}_{\mathcal{IH}}}$ given by $\frac{1}{M+1} \mapsto \mathcal{IH}^*_M$ and $0 \mapsto \mathcal{N}(0,1)$ is a homeomorphism.

Moduli space of plane partition distributions

Def. The *moduli space of plane partition distributions* is

$$\mathbf{M}_{\mathrm{PP}} := \{\mathcal{X}_{a\times b\times c}[\mathrm{size}]^* : a, b, c \in \mathbb{Z}_{\geq 1}\}$$

with the topology characterized by convergence in distribution.

Corollary. In the Lévy metric,

$$\overline{\mathbf{M}_{\mathrm{PP}}} = \mathbf{M}_{\mathrm{PP}} \sqcup \overline{\mathbf{M}_{\mathcal{IH}}},$$

which is compact. The set of limit points of $\mathbf{M}_{\mathrm{PP}}$ is exactly $\overline{\mathbf{M}_{\mathcal{IH}}}$.

Moduli space of SYT distributions

Def. The *moduli space of* SYT *distributions* is

$$\mathbf{M}_{\mathsf{SYT}} := \{X_\lambda[\mathrm{maj}]^* : \lambda \in \mathrm{Par}, \#SYT(\lambda) > 1\}$$

with the topology characterized by convergence in distribution.

Corollary. In the Lévy metric,

$$\overline{\mathbf{M}_{\mathsf{SYT}}} = \mathbf{M}_{\mathsf{SYT}} \sqcup \overline{\mathbf{M}_{\mathcal{IH}}},$$

which is compact. The set of limit points of $\mathbf{M}_{\mathsf{SYT}}$ is exactly $\overline{\mathbf{M}_{\mathcal{IH}}}$.

Moduli space of SYT distributions

Def. The *moduli space of* SYT *distributions* is

$$\mathbf{M}_{\mathsf{SYT}} := \{X_\lambda[\mathrm{maj}]^* : \lambda \in \mathrm{Par}, \#SYT(\lambda) > 1\}$$

with the topology characterized by convergence in distribution.

Corollary. In the Lévy metric,

$$\overline{\mathbf{M}_{\mathsf{SYT}}} = \mathbf{M}_{\mathsf{SYT}} \sqcup \overline{\mathbf{M}_{\mathcal{IH}}},$$

which is compact. The set of limit points of $\mathbf{M}_{\mathsf{SYT}}$ is exactly $\overline{\mathbf{M}_{\mathcal{IH}}}$.

Semistandard tableaux and Schur functions

Defn. A *semistandard Young tableau* of shape λ is filling of λ such that every row is weakly increasing from left to right and every column is strictly increasing from top to bottom.

$$T = \begin{array}{|c|c|c|c|c|} \hline 1 & 3 & 3 & 3 & 3 \\ \hline 2 & 5 & 5 \\ \cline{1-3} 9 \\ \cline{1-1} \end{array} \qquad x^T = x_1 x_2 x_3^4 x_5^2 x_9 \qquad \mathrm{rank}(T) = 28$$

Associate a monomial to each semistandard tableau, $T \mapsto x^T = x_1^{\alpha_1} x_2^{\alpha_2} \cdots$ where α_i is the number of i's in T. Let $\mathrm{rank}(T) = \sum (i-1)\alpha_i$.

Def. The *Schur polynomial* indexed by λ on $(x_1, \ldots, x_m)$ is

$$s_\lambda(x_1, x_2, \ldots, x_m) = \sum x^T$$

summed over all semistandard Young tableaux of shape λ filled with numbers in $\{1, 2, \ldots, m\}$, denoted $\mathrm{SSYT}_{\leq m}(\lambda)$.

Semistandard tableaux and Schur functions

Stanley+Littlewood. The principle specialization of the Schur polynomial is a cyclotomic generating function

$$s_\lambda(1, q, q^2, \ldots, q^{m-1}) = \sum_{T \in \mathrm{SSYT}_{\leq m}(\lambda)} q^{\mathrm{rank}(T)}$$

$$= q^{b(\lambda)} \prod_{u \in \lambda} \frac{[m + c_u]_q}{[h_u]_q}$$

$$= q^{b(\lambda)} \prod_{1 \leq i < j \leq m} \frac{[\lambda_i - \lambda_j + j - i]_q}{[j - i]_q}$$

where $c_u = j - i$ is the *content* of cell $u = (i, j)$ and h_u is the hook length of u.

Moduli Space of SSYT Distributions

Def. Let $\mathcal{X}_{\lambda;m}[\mathrm{rank}]$ denote the random variable associated with the rank statistic on $\mathrm{SSYT}_{\leq m}(\lambda)$, sampled uniformly at random.

Def. The *moduli space of* SSYT *distributions* is

$$\mathbf{M}_{\mathrm{SSYT}} := \{\mathcal{X}_{\lambda;m}[\mathrm{rank}]^{*} : \lambda \in \mathrm{Par}, \ell(\lambda) \leq m\}$$

Moduli Space of SSYT Distributions

Def. Let $\mathcal{X}_{\lambda;m}[\text{rank}]$ denote the random variable associated with the rank statistic on $\text{SSYT}_{\le m}(\lambda)$, sampled uniformly at random.

Def. The *moduli space of* SSYT *distributions* is

$$\mathbf{M}_{\text{SSYT}} := \{\mathcal{X}_{\lambda;m}[\text{rank}]^* : \lambda \in \text{Par}, \ell(\lambda) \le m\}$$

Open Problem. Describe $\overline{\mathbf{M}_{\text{SSYT}}}$ in the Lévy metric. What are all possible limit points?

Toward Limit Laws of SSYT Distributions

Def. Given a finite multiset $\mathbf{t} = \{t_1 \geq t_2 \geq \cdots \geq t_m\}$ of non-negative real numbers, let

$$\mathcal{S}_{\mathbf{t}} := \sum_{t \in \mathbf{t}} \mathcal{U}\left[-\frac{t}{2}, \frac{t}{2}\right], \tag{3}$$

where we assume the summands are independent and $\mathcal{U}[a, b]$ denotes the continuous uniform distribution supported on $[a, b]$. We say $\mathcal{S}_{\mathbf{t}}$ is a *finite generalized uniform sum distribution*.

Example. If $\mathbf{t}$ consists of M copies of 1, then $\mathcal{S}_{\mathbf{t}} + \frac{M}{2}$ is the Irwin-Hall distribution $\mathcal{IH}_M$.

Distance Multisets

Def. The *distance multiset* of $\mathbf{t} = \{t_1 \geq t_2 \geq \cdots \geq t_m\}$ is the multiset

$$\Delta\mathbf{t} := \{t_i - t_j : 1 \leq i < j \leq m\}.$$

Theorem. Let λ be an infinite sequence of partitions with $\ell(\lambda) < m$ where $\lambda_1/m^3 \to \infty$. Let $\mathbf{t}(\lambda) = (t_1, \ldots, t_m) \in [0,1]^m$ be the finite multiset with $t_k := \frac{\lambda_k}{\lambda_1}$ for $1 \leq k \leq m$. Then $\mathcal{X}_{\lambda;m}[\text{rank}]^*$ converges in distribution if and only if the multisets $\Delta\mathbf{t}(\lambda)$ converge pointwise.

In that case, the limit distribution is $\mathcal{N}(0,1)$ if $m \to \infty$ and $\mathcal{S}_{\mathbf{d}}^*$ where $\Delta\mathbf{t}(\lambda) \to \mathbf{d}$ if m is bounded.

Moduli Space of Distance Distributions

Def. The *moduli space of distance distributions* is

$$\mathbf{M}_{\mathrm{DIST}} := \bigcup_{m \geq 2} \{\mathcal{S}^*_{\Delta \mathbf{t}} : \mathbf{t} = \{1 = t_1 \geq \cdots \geq t_m = 0\}\}$$

and its associated parameter space $\mathbf{P}_{\mathrm{DIST}}$ is a renormalized variation on $\left\{\Delta \mathbf{t} : \mathbf{t} = \{1 = t_1 \geq \cdots \geq t_m = 0\}\right\}$.

Conclusions/Thm.

1. $\overline{\mathbf{P}_{\mathrm{DIST}}} = \mathbf{P}_{\mathrm{DIST}} \sqcup \{\mathbf{0}\}$ where $\mathbf{0}$ is the infinite sequence of 0's.
2. $\overline{\mathbf{M}_{\mathrm{DIST}}} = \mathbf{M}_{\mathrm{DIST}} \sqcup \{\mathcal{N}(0,1)\}$.
3. The map $\overline{\mathbf{P}_{\mathrm{DIST}}} \to \overline{\mathbf{M}_{\mathrm{DIST}}}$ given by $\mathbf{d} \mapsto \mathcal{S}^*_{\mathbf{d}}$ and $\mathbf{0} \mapsto \mathcal{N}(0,1)$ is a homeomorphism between compact spaces.

Moduli Space of SSYT Distributions

Corollary. For any fixed $\epsilon > 0$, let

$$\mathbf{M}_{\epsilon\,\mathsf{SSYT}} := \{\mathcal{X}_{\lambda;m}[\text{rank}]^* : \ell(\lambda) < m \text{ and } \lambda_1/m^3 > (|\lambda|+m)^\epsilon\} \subset \mathbf{M}_{\mathsf{SSYT}}.$$

Then

$$\overline{\mathbf{M}_{\epsilon\,\mathsf{SSYT}}} = \mathbf{M}_{\epsilon\,\mathsf{SSYT}} \sqcup \overline{\mathbf{M}_{\mathsf{DIST}}},$$

which is compact. The set of limit points of $\mathbf{M}_{\epsilon\,\mathsf{SSYT}}$ is $\overline{\mathbf{M}_{\mathsf{DIST}}}$.

Moduli Space of SSYT Distributions

Corollary. For any fixed $\epsilon > 0$, let

$$\mathbf{M}_{\epsilon\,\mathrm{SSYT}} := \{\mathcal{X}_{\lambda;m}[\mathrm{rank}]^* : \ell(\lambda) < m \text{ and } \lambda_1/m^3 > (|\lambda|+m)^\epsilon\} \subset \mathbf{M}_{\mathrm{SSYT}}.$$

Then

$$\overline{\mathbf{M}_{\epsilon\,\mathrm{SSYT}}} = \mathbf{M}_{\epsilon\,\mathrm{SSYT}} \sqcup \overline{\mathbf{M}_{\mathrm{DIST}}},$$

which is compact. The set of limit points of $\mathbf{M}_{\epsilon\,\mathrm{SSYT}}$ is $\overline{\mathbf{M}_{\mathrm{DIST}}}$.

Corollary. For the moduli space of limit laws for Stanley's q-hook-content formula, we have shown

$$\mathbf{M}_{\mathrm{SSYT}} \cup \mathbf{M}_{\mathrm{DIST}} \cup \mathbf{M}_{\mathcal{IH}} \cup \{\mathcal{N}(0,1)\} \subset \overline{\mathbf{M}_{\mathrm{SSYT}}}.$$

Moduli Space of Generalized Sum Distributions

The limiting distributions q-hook length formulas for linear extensions of forests due to Björner–Wachs include all countably infinite generalized uniform sum distributions with finite variance, which is closely related to the 2-norm of the indexing multiset.

Theorem. The limit laws for all possible standardized general uniform sum distributions $\mathbf{M}_{\mathsf{SUMS}} : \{\mathcal{S}_{\mathbf{t}}^* : \mathbf{t} \in \widetilde{\ell}_2\}$ is exactly the *moduli space of DUSTPAN distributions,*

$$\overline{\mathbf{M}_{\mathsf{SUMS}}} = \mathbf{M}_{\mathsf{DUST}} := \{\mathcal{S}_{\mathbf{t}} + \mathcal{N}(0, \sigma^2) : |\mathbf{t}|_2^2/12 + \sigma^2 = 1\}.$$

Moduli Space of Generalized Sum Distributions

The limiting distributions q-hook length formulas for linear extensions of forests due to Björner–Wachs include all countably infinite generalized uniform sum distributions with finite variance, which is closely related to the 2-norm of the indexing multiset.

Theorem. The limit laws for all possible standardized general uniform sum distributions $\mathbf{M}_{\mathsf{SUMS}} : \{\mathcal{S}_{\mathbf{t}}^* : \mathbf{t} \in \widetilde{\ell_2}\}$ is exactly the *moduli space of DUSTPAN distributions,*

$$\overline{\mathbf{M}_{\mathsf{SUMS}}} = \mathbf{M}_{\mathsf{DUST}} := \{\mathcal{S}_{\mathbf{t}} + \mathcal{N}(0, \sigma^2) : |\mathbf{t}|_2^2/12 + \sigma^2 = 1\}.$$

The nomenclature DUSTPAN refers to a *distribution associated to a uniform sum for* $\mathbf{t}$ *plus an independent normal distribution.*

The moduli space of limit laws for q-hook formulas

Let $\mathbf{M}_{\text{Forest}}$ be the moduli space of standardized distributions associated to forests. We know $\mathbf{M}_{\text{Forest}} \cup \mathbf{M}_{\text{DUST}} \subset \overline{\mathbf{M}_{\text{Forest}}}$, implying there are an uncountable number of possible limit laws for distributions associated to forests.

Open Problem. Describe $\overline{\mathbf{M}_{\text{Forest}}}$ in the Lévy metric. What are all possible limit points?

Open Problem. Describe $\overline{\mathbf{M}_{\text{CGF}}}$ in the Lévy metric. What are all possible limit points? Is $\mathbf{M}_{\text{CGF}} \cup \mathbf{M}_{\text{DUST}}$ <u>the</u> moduli space of limit laws for q-hook formulas?

Conclusion

Many Thanks!