

DUSTPAN distributions as limit laws for Mahonian statistics on forests

Joshua P. Swanson
University of California, San Diego (UCSD)

Based on joint work with
Sara Billey and Matjaž Konvalinka

Papers: [BS20], [BKS20] / arXiv: 2010.12701, 1905.00975
Slides: http://www.math.ucsd.edu/~jswanson/talks/2020_Ljubljana.pdf

Presented at University of Ljubljana Discrete Math Seminar
December 15th, 2020

Outline

- Background: Irwin-Hall Limits
- Hook Length Formulas
- DUSTPAN Distributions
- DUSTPAN Limits
- Further Directions

Background: Irwin-Hall Limits

Def Let $U[a,b]$ denote a uniform continuous random variable supported on $[a,b]$.

The M th Irwin-Hall distribution is

$$IH_M = U[0,1] + \dots + U[0,1]$$

M i.i.d. r.v.'s

Ex By CLT, $IH_M^* \xrightarrow{D} N(0,1)$, i.e. $\forall t \in \mathbb{R}$, $\lim_{n \rightarrow \infty} P[IH^* \leq t] = P[N(0,1) \leq t]$

Notation Write $X^* = \frac{X - \mu}{\sigma}$

for standardised r.v.

Background: Irwin-Hall Limits

Rem

Our motivation in [BKS20] was to study the distribution of the major index on standard Young tableaux, generalizing earlier work on the major index of permutations and words.

$T =$

1	2	4	7	9	12
3	6	10			
5	8	11			

$\in \text{SMT}(6,3,3)$

an integer partition

$$\text{Des}(T) = \{2, 4, 7, 9, 10\}$$

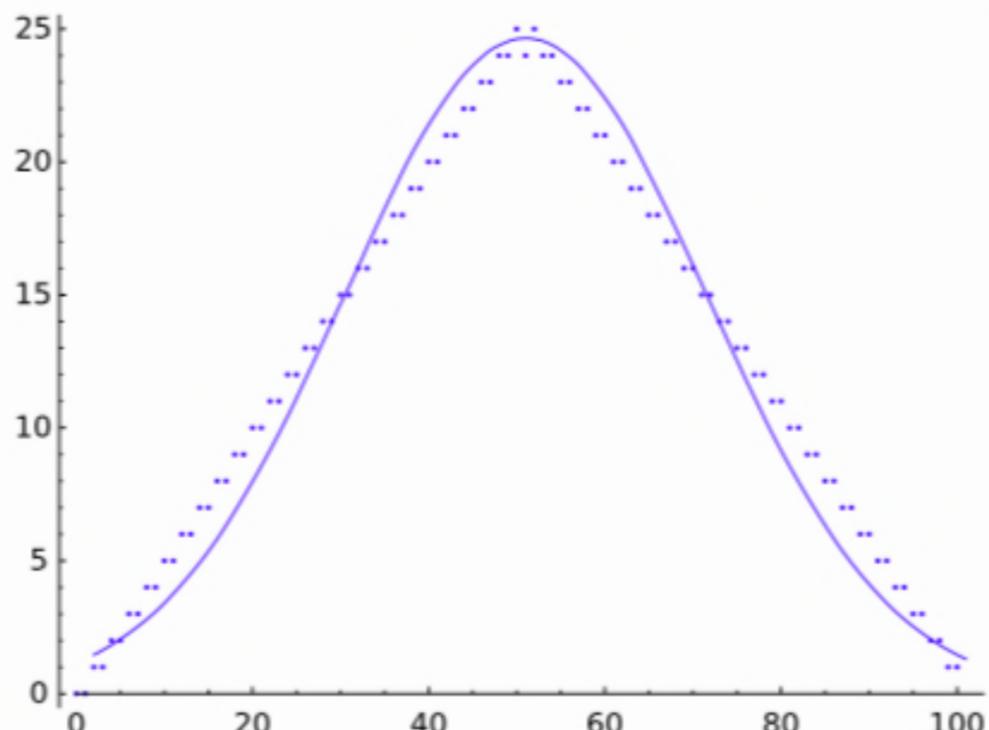
$$\text{des}(T) = |\text{Des}(T)| = 5$$

$$\text{maj}(T) = 2 + 4 + 7 + 9 + 10 = \boxed{32}$$

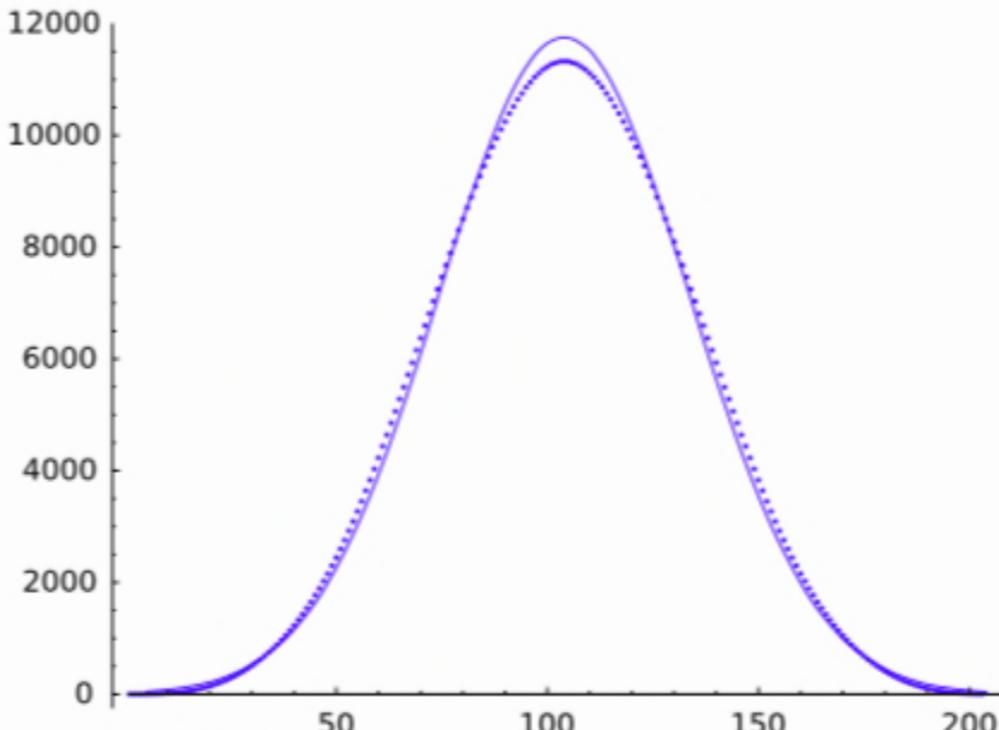
Background: Irwin-Hall Limits

Def] Each integer partition $\lambda \vdash n$ has a random variable $X_\lambda[\text{maj}]$

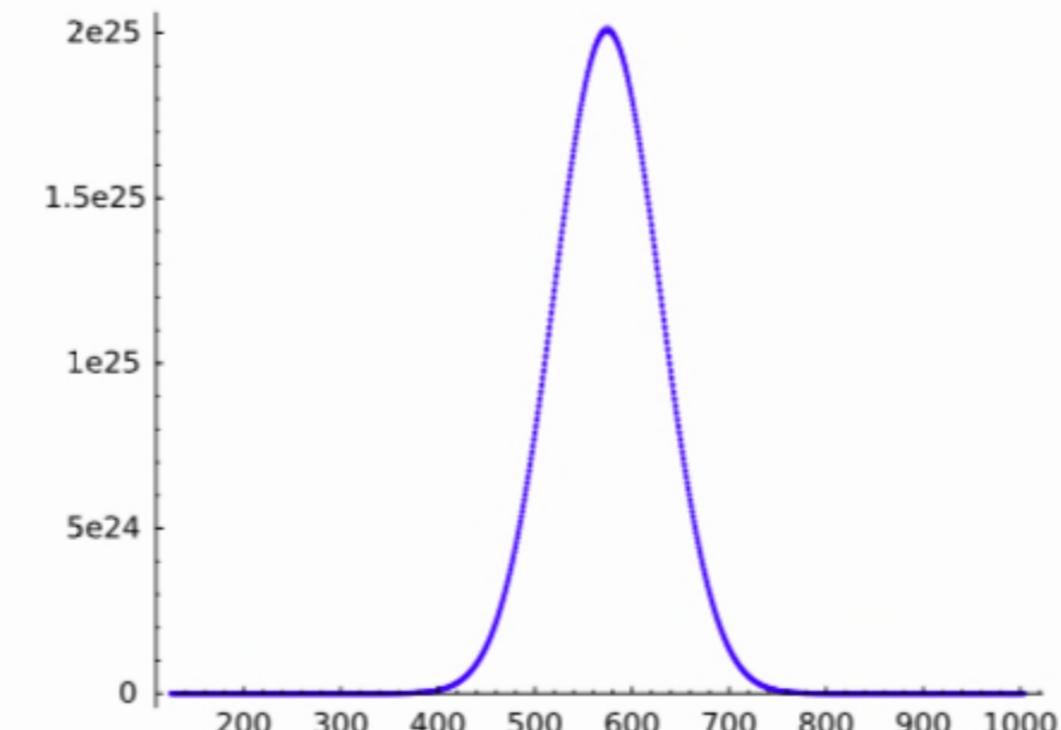
Ex] Distributions of $X_\lambda[\text{maj}]$:



(a) $\lambda = (50, 2)$, $\text{aft}(\lambda) = 2$



(b) $\lambda = (50, 3, 1)$, $\text{aft}(\lambda) = 4$



(c) $\lambda = (8, 8, 7, 6, 5, 5, 5, 2, 2)$, $\text{aft}(\lambda) = 39$

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.

Background: Irwin-Hall Limits

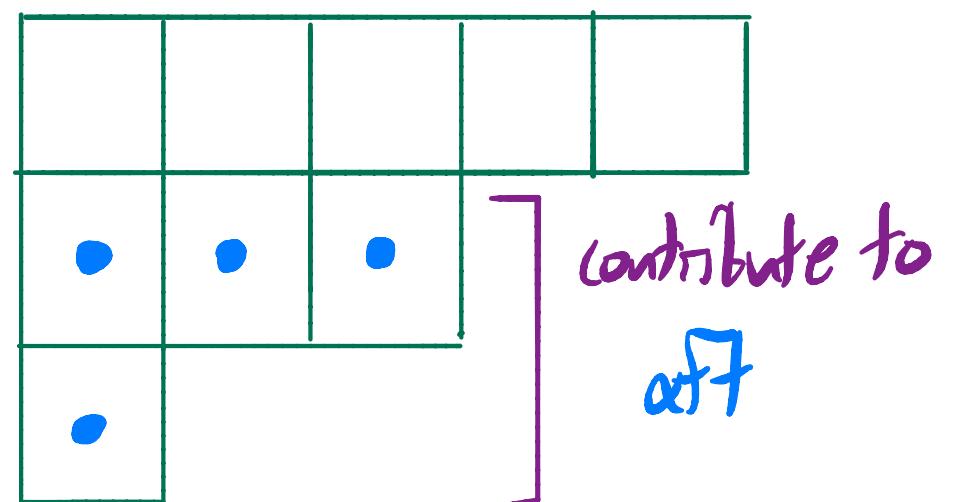
Note • $\chi_{(5,2)}[\text{maj}]$ "looks like" $M_2 = U[0,1] + U[0,1]$!

• $\chi_{(8,8,7,6,5,5,5,2,2)}[\text{maj}]$ "looks like" $N!$

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

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

Ex $\lambda = (5, 3, 1) \Rightarrow \text{aft}(\lambda) = 4$



Rem On FindStat
as ST001214;
"soon" on OEIS as
A338621

Background: Irwin-Hall Limits

Thm (Billey-Konvalinka-S. [BK20, Thm. 1.7])

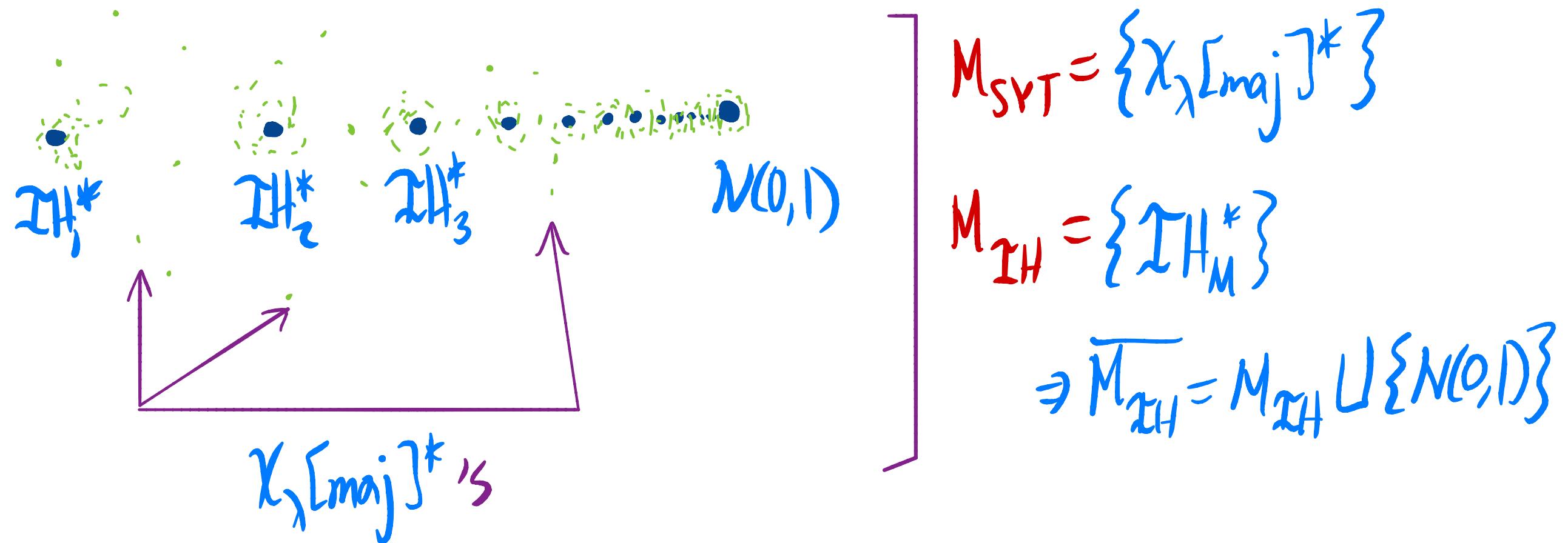
Suppose $\lambda^{(1)}, \lambda^{(2)}, \dots$ is a sequence of partitions. Then $\chi_{\lambda^{(1)}[\text{maj}]}^*, \chi_{\lambda^{(2)}[\text{maj}]}^*, \dots$ converges in distribution if and only if

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

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

Background: Irwin-Hall Limits

Idea The moduli space of SRT distributions under (say) the Lévy metric:



(or $\overline{M}_{SRT} = M_{SRT} \cup \overline{M}_{IH}$)

the set of limit points!

Hook Length Formulas

Rem] Proof in [BKS20, Thm. 1.7] relies on Stanley's
 q -hook length formula:

$$\sum_{\text{TSRT}(\lambda)} q^{\text{maj}(\text{t})} = q^{r(\lambda)} \frac{[n]_q!}{\prod_{c \in C} [h_c]_q}$$

ratio is key to
cumulant formula!
then method of moments

Q] What other combinatorial statistics arise as quotients of
 q -integers? (Cyclotomic generating functions)

Recall $[n]_q = \frac{1-q^n}{1-q}$
 $= 1 + q + \dots + q^{n-1}$

Hook Length Formulas

Thm The rank on semistandard tableaux of shape λ and entries $\leq m$ is

$$\sum_{T \in \text{SSRT}_{\leq m}(\lambda)} q^{\text{rank}(T)} = q^{r(\lambda)} \prod_{u \in \lambda} \frac{[m + c_u]_q}{[c_u]_q} = q^{r(\lambda)} \prod_{1 \leq i < j \leq m} \frac{[\lambda_i - \lambda_j + j - i]_q}{[j - i]_q}.$$

Stanley's q -hook-content formula q -Lkey dimension formula (type A)

Thm The size on plane partitions in an $a \times b \times c$ box is

$$\sum_{P \in \text{PP}(a \times b \times c)} q^{\text{size}(P)} = \prod_{i=1}^a \prod_{j=1}^b \prod_{k=1}^c \frac{[it+j+k-1]_q}{[it+j+k-2]_q} \quad \text{MacMahon}$$

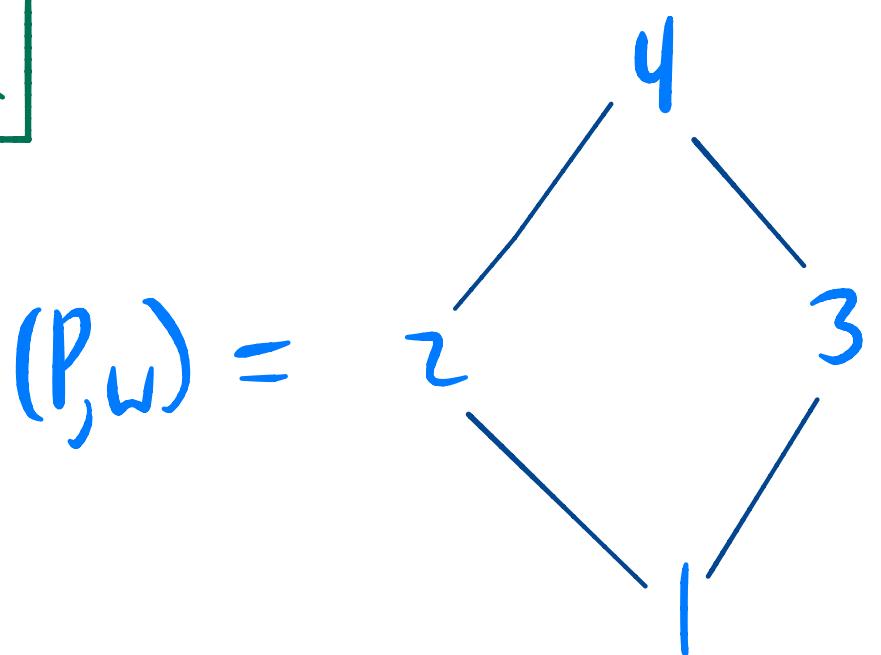
Hook Length Formulas

Def Let P be a forest viewed as a poset with roots as maximal elements. Fix an order-preserving bijection $w: P \rightarrow [n]$.

Let $\mathcal{L}(P, w) = \{\text{linear extensions of } P, \text{ viewed as permutations of } [n] \text{ via } w\}$.

The inversion number of $\pi \in S_n$ is $\text{inv}(\pi) = \#\{(i, j) : i < j \leq n, \pi(i) > \pi(j)\}$.

Ex



$$\mathcal{L}(P, w) = \{1234, 1324\} \Rightarrow \text{inv}'s = \{0, 1\}$$

Hook Length Formulas

Thm (Stanley, Björner-Wachs)

$$\sum_{\pi \in \mathcal{L}(P, w)} q^{\text{inv}(\pi)} = \frac{[n]_q!}{\prod_{u \in P} [h_u]_q}.$$

Q What does the moduli space of Forest distributions

$$M_{\text{Forest}} = \{ \chi_p^*[\text{inv}] \}$$

look like?

DUSTPAN Distributions

Def The rank of P is the length of a maximal chain.

Thm ("Generic" case) Given a sequence of forests P ,

$$\chi_{P[\text{inv}]} \xrightarrow{*} N(0, 1)$$

if

$$|P| \rightarrow \infty \quad \text{and} \quad \limsup \frac{\text{rank}(P)}{|P|} < 1.$$

Idea



DUSTPAN Distributions

Let $n = |P|$ and $r = \text{rank}(P)$.

Q Consider the set of rooted, unlabeled forests with n vertices, sampled uniformly at random. What is the expected value of the rank r ? How does r compare to n as $n \rightarrow \infty$?

Rem Boutin-Flajolet proved $E[r] \sim C\sqrt{n}$ for binary trees.

Typically $E[r] \sim D\sqrt{n}$ for ordered/labeled variations.

(Certainly $\limsup \frac{r}{n} < 1$ is "typical"!)

DUSTPAN Distributions

Rem the "degenerate" case $n \sim r$ hides a world of behavior!!

Def Let $\ell_2 = \{(t_1, t_2, \dots) : t_i \geq 0\}$ and

$$\tilde{\ell}_2 = \{(t_1, t_2, \dots) : t_1 \geq t_2 \geq \dots \geq 0\}$$

be sequence spaces.

Let $t \in \tilde{\ell}_2$. The corresponding generalized uniform random variable is

$$S_t = \sum_i U\left[-\frac{t_i}{2}, \frac{t_i}{2}\right]$$

works for $\text{supp}(t) > 0$ iff
 $\|t\|_2^2 = \sum_i t_i^2 < \infty$

DUSTPAN Distributions

Def A DUSTPAN Distribution is a distribution associated to a uniform sum for \underline{t} plus a normal distribution

$$\mathcal{S}_t + \mathcal{N}(0, \sigma^2)$$

where $t \in \tilde{\mathbb{L}}_2, \sigma \in \mathbb{R}_{\geq 0}$.

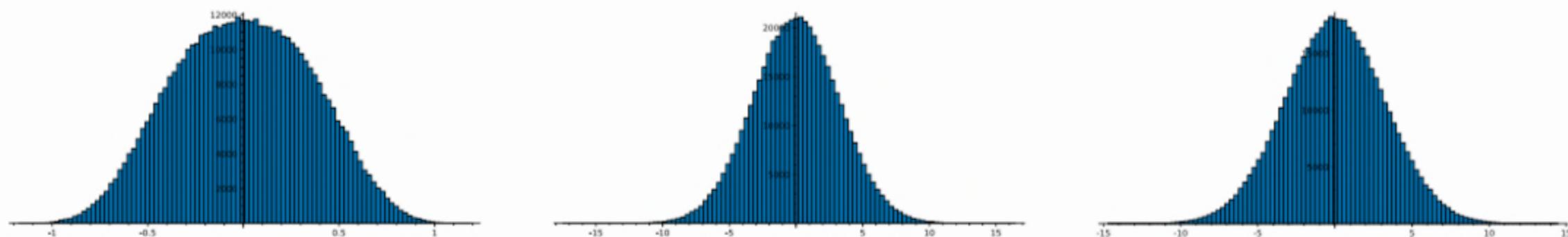


FIGURE 5. Histograms of the distributions \mathcal{S}_t , $\mathcal{N}(0, \sigma)$, and $\mathcal{S}_t + \mathcal{N}(0, \sigma)$ with $t = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8})$ and $\sigma \approx 3.22$.

DUSTPAN Distributions

Def] The moduli space of standardized DUSTPAN distributions is

$$M_{DUST} = \left\{ S_t + N(0, \sigma^2) : \frac{\|t\|_2^2}{12} + \sigma^2 \leq 1 \right\}.$$

The standardized DUSTPAN parameter space is

$$P_{DUST} = \left\{ t \in \ell_2 : \|t\|_2^2 \leq 12 \right\}. \quad \begin{array}{l} \text{metric space under} \\ \text{pointwise convergence} \end{array}$$

DUSTPAN Distributions

Thm | The map

$$\Phi: P_{\text{DUST}} \rightarrow M_{\text{DUST}}$$

where

$$\Phi(t) = S_t + N(0, \sigma^2) \quad \text{where } \sigma = \sqrt{1 - \|t\|_2^2} / \sqrt{2}$$

is a homeomorphism.

Cor | M_{DUST} is closed under convergence in distribution!

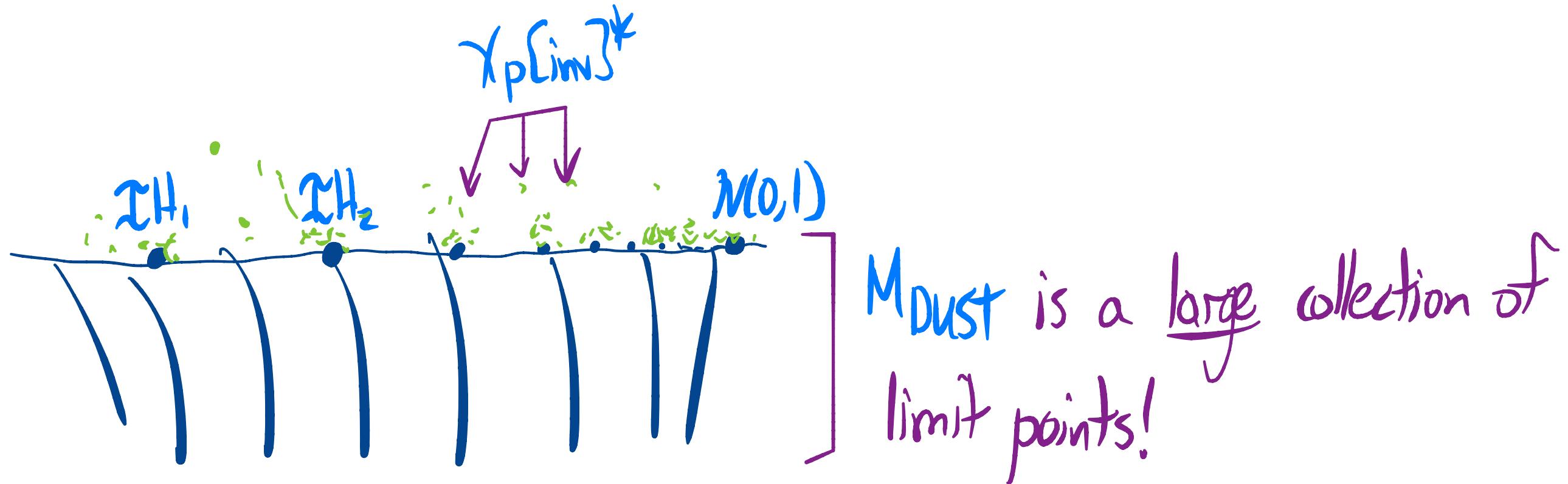
DUSTPAN Limits

Thm [BSZD, Thm. 1.13] Let P be an infinite sequence of standardized trees with $n - r = o(n^{1/2})$. Then $\chi_p^*[\text{inv}]$ converges in distribution if and only if the elevation multisets \hat{e} associated to P converge pointwise to some $\tilde{t} \in \tilde{\ell}_2$.

In that case, the limit distribution is $S_t + N(0, \sigma^2)$ where $\frac{\|t\|_2^2}{12} + \sigma^2 = 1$.

DUSTPAN Limits

Idea



Or For any fixed $\epsilon > 0$, let ϵTREE be the set of standardized trees for which $n - r < n^{\frac{1}{2} - \epsilon}$. Let $M_{\epsilon\text{TREE}} = \{X_p^*: P \in \epsilon\text{TREE}\} \subset M_{\text{Forest}}$. Then

$$\overline{M_{\epsilon\text{TREE}}} = M_{\epsilon\text{TREE}} \cup M_{\text{DUST}}.$$

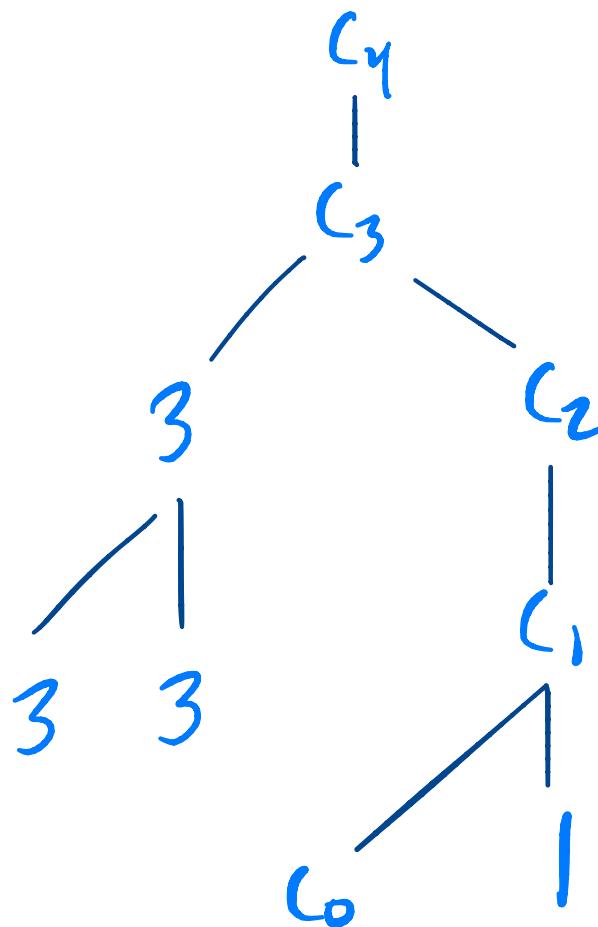
all limit points!

DUSTPAN Limits

"Standardized tree"? "Elevation multiset"?

Def A rooted tree is standardized if its root has at least two children.

Def The elevation multiset of P relative to a fixed maximal chain is:



$$e = \{1, 3, 3, 3\} = (3, 3, 3, 1, 0, 0, \dots)$$

Rem Normalize via

$$\hat{e} = \frac{\sqrt{2} \cdot e}{\|e\|_2}$$

Further Directions

Rem Have related results for...

- rank on $\text{SSYT}_{Sm}(\lambda)$: complicated subset of M_{DUSTPAN}
- size on $PP_{a \times b \times c}$: only M_{ITH}

Q • What about between $n-r=w(n^{1/12})$ and $n-r=o(n^{1/12})$ regimes for M_{Forest} ?

- What about M_{GF} ??
- More applications of DUSTPAN's?

