Statistical Properties of Permutations and Standard Young Tableau reflected by the Robinson–Schensted Algorithm

Austin DeCicco, Jike Liu

MATH 640, Spring 2026

May 2026

Introduction

The Robinson–Schensted correspondence is a classical bijection between permutations $\pi \in S_n$ and pairs (P, Q) of standard Young tableaux of the same shape. It translates permutation statistics into tableau statistics.

In particular, Schensted's theorem states that the length of the longest increasing subsequence of π is the first row length of the common shape, while the length of the longest decreasing subsequence is the first column length [2].

Introduction: Maple Experiments

In this project, we study statistical properties of the Robinson–Schensted correspondence through Maple experiments. For each permutation π, we compute

$$\mathrm{RS}(\pi) = (P(\pi), Q(\pi)), \qquad \lambda(\pi) = \mathrm{shape}(P(\pi)) = \mathrm{shape}(Q(\pi)),$$

and investigate statistics of the shape $\lambda(\pi)$ and of the entries of $P(\pi)$ and $Q(\pi)$.

Introduction: Directions

We first test the Logan–Shepp limit shape theorem for random Young diagrams under Plancherel measure [4].

We then study the relationship between the first row and first column of the RS shape, equivalently the longest increasing and decreasing subsequence lengths of a random permutation.

Next, we investigate the location of a fixed entry m in the insertion and recording tableaux.

Finally, we introduce row and column sum vectors of standard Young tableaux.

Background on Robinson–Schensted Correspondence

The Robinson–Schensted correspondence is a bijection between permutations $\pi \in S_n$ and pairs (P, Q) of standard Young tableaux of the same shape. We write

$$\mathrm{RS}(\pi) = (P(\pi), Q(\pi)).$$

Here $P(\pi)$ is called the insertion tableau and $Q(\pi)$ is called the recording tableau. Both tableaux have the same shape, which we denote by

$$\lambda(\pi) = \mathrm{shape}(P(\pi)) = \mathrm{shape}(Q(\pi)).$$

Standard Young Tableau

A standard Young tableau is a filling of the boxes of a Young diagram with the numbers $1, 2, \ldots, n$, such that entries increase from left to right along each row and from top to bottom along each column. For example,

1	3	5
2	4	

is a standard Young tableau of shape $(3, 2)$.

In this paper, we often represent the shape of a tableau as a vector of row lengths. Thus the shape of the above tableau is $(3, 2)$.

Schensted's Theorem

Theorem (Schensted's theorem)

Let $\pi \in S_n$*, and let*

$$\mathrm{RS}(\pi) = (P(\pi), Q(\pi)).$$

If

$$\lambda = \mathrm{shape}(P(\pi)) = \mathrm{shape}(Q(\pi)),$$

then the length of the longest increasing subsequence of π *is equal to the length of the first row of* λ*, and the length of the longest decreasing subsequence of* π *is equal to the length of the first column of* λ*. In other words,*

$$\mathrm{LIS}(\pi) = \lambda_1, \qquad \mathrm{LDS}(\pi) = \lambda_1'.$$

Plancherel Measure

This theorem is the main reason that the Robinson–Schensted correspondence is useful for studying statistical properties of random permutations.

When π is chosen uniformly at random from S_n, the random shape $\lambda(\pi)$ is distributed according to Plancherel measure:

$$\mathbb{P}(\lambda) = \frac{(f^\lambda)^2}{n!},$$

where f^λ is the number of standard Young tableaux of shape λ.

This probability distribution is the background for the Logan–Shepp limit shape theorem discussed in the next section.

Logan-Shepp Limit Shape

Let $\pi \in S_n$ be a uniformly random permutation, and let

$$RS(\pi) = (P(\pi), Q(\pi)).$$

The two tableaux $P(\pi)$ and $Q(\pi)$ have the same shape, which we denote by $\lambda(\pi)$. As π varies uniformly over S_n, the random partition $\lambda(\pi) \vdash n$ is distributed according to the Plancherel measure.

The Logan–Shepp theorem says that after scaling both axes by $\sqrt{n}$, the random Young diagram concentrates around a deterministic curve.

Logan-Shepp Curve

In the coordinates used in our plots, this curve is given parametrically by

$$x(\theta) = \frac{2}{\pi}\left(\sin\theta - \theta\cos\theta\right),$$

and

$$y(\theta) = \frac{2}{\pi}\left(\sin\theta + (\pi-\theta)\cos\theta\right), \qquad 0 \le \theta \le \pi.$$

The curve starts at $(0, 2)$ and ends at $(2, 0)$. Thus the limit shape predicts that both the first row length and the first column length are asymptotically of order $2\sqrt{n}$:

$$\frac{\lambda_1}{\sqrt{n}} \longrightarrow 2, \qquad \frac{\lambda_1'}{\sqrt{n}} \longrightarrow 2.$$

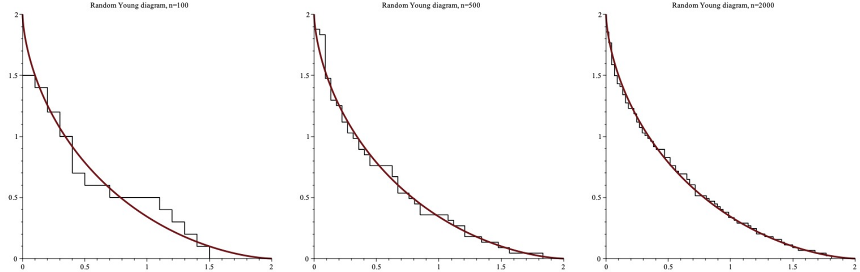


Figure: Scaled random Young diagrams compared with the Logan–Shepp limit shape for $n = 100, 500, 2000$.

Experimental Test

To test this experimentally, we implemented Maple procedures that generate random permutations, compute their Robinson–Schensted shapes, scale the resulting Young diagrams by $\sqrt{n}$, and plot them together with the Logan–Shepp curve.

The procedure `RandShape(n)` generates a random permutation in S_n and returns the shape of its insertion tableau.

The procedure `ScaledShapePoints(L)` converts a partition L into the boundary points of its scaled Young diagram.

Finally, `PlotLS(n)` displays the scaled random Young diagram together with the theoretical Logan–Shepp curve.

First Row and First Column

We also tested the convergence of the first row and first column numerically. For each n, we generated 1000 random permutations and computed the sample means of λ_1 and λ_1'.

n	K	$\mathbb{E}(\lambda_1)$	$\mathbb{E}(\lambda_1')$	$\mathbb{E}(\lambda_1)/\sqrt{n}$	$\mathbb{E}(\lambda_1')/\sqrt{n}$
100	1000	16.744	16.764	1.6744	1.6764
500	1000	40.307	40.197	1.8026	1.7977
1000	1000	58.278	58.247	1.8429	1.8419

The data show this trend: the normalized first row and first column averages increase from about 1.67 when $n = 100$ to about 1.84 when $n = 1000$.

Covariance and Correlation

Given a permutation $\pi \in S_n$, let

$$RS(\pi) = (P(\pi), Q(\pi)),$$

and let $\lambda = \text{shape}(P(\pi)) = \text{shape}(Q(\pi))$.

If

$$R_n = \lambda_1, \qquad C_n = \lambda_1' = \ell(\lambda),$$

then

$$R_n = \text{LIS}(\pi), \qquad C_n = \text{LDS}(\pi).$$

Covariance and Correlation: Question

A natural next question is to ask how the first row and first column are related to each other. Since the Young diagram has fixed area n, one might expect a competition between horizontal and vertical growth.

Recall that

$$\mathrm{Cov}(R_n, C_n) = \mathbb{E}[R_n C_n] - \mathbb{E}[R_n]\mathbb{E}[C_n],$$

and

$$\mathrm{Corr}(R_n, C_n) = \frac{\mathrm{Cov}(R_n, C_n)}{\sqrt{\mathrm{Var}(R_n)\,\mathrm{Var}(C_n)}}.$$

First Row / First Column Monte Carlo

The following table shows the output for $K = 1000$ samples.

n	K	$\mathbb{E}[R_n]$	$\mathbb{E}[C_n]$	$\mathrm{Var}(R_n)$	$\mathrm{Var}(C_n)$	$\mathrm{Cov}(R_n, C_n)$	$\mathrm{Corr}(R_n, C_n)$
100	1000	16.7180	16.6190	2.5805	2.6778	−0.6364	−0.2421
500	1000	40.1910	40.2380	5.2705	5.6394	−0.6685	−0.1226
1000	1000	58.2380	58.1340	7.0354	7.8600	−0.8889	−0.1195

In all three cases, the estimated covariance is negative. This supports the intuition that the first row and first column are negatively related for finite n. However, the estimated correlation becomes closer to zero as n increases.

Conjectures

Conjecture

For every $n \geq 2$,

$$\mathrm{Cov}(R_n, C_n) < 0.$$

Equivalently, for a uniformly random permutation $\pi \in S_n$, the lengths of the longest increasing and longest decreasing subsequences are negatively correlated.

Conjecture

As $n \to \infty$,

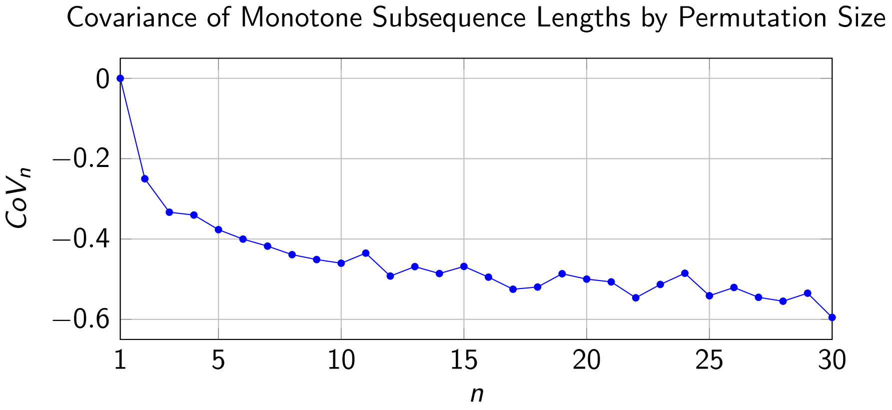
$$\mathrm{Corr}(R_n, C_n) \to 0.$$

Conjecture

$\exists k$ s.t,

$$\lim_{n\to\infty} \mathrm{Cov}(R_n, C_n) = k.$$

Covariance of Monotone Subsequence Lengths



A Experimental Row-Removal Heuristic: Procedures

Maple package `AvgShape.txt` contains procedures to calculate the average shape of a standard Young tableaux outputted by the Robinson–Schensted correspondence.

`SYTShape()` finds the shape of a standard Young tableaux and outputs it in vector representation.

We then pick with replacement to generate 3000 random observations of standard Young tableaux shapes for each n from 1 to 30.

`AvgnthLength(x)` and `senthLength(x)` find the average and standard error respectively of row $1 \leq x \leq 30$ for each n from 1 to 30.

Expanded Logic for Row 2

Cut-Paste: Cutting-out any row from a standard Young tableaux and pasting together all rows before and after in the same order always results in a valid standard Young tableaux after re-numerating by size of each element.

A somewhat natural sounding assumption follows that the length of row 2 in P is the length of the longest increasing subsequence in π that doesn't use any elements contained in the true longest increasing subsequence, however this is incorrect.

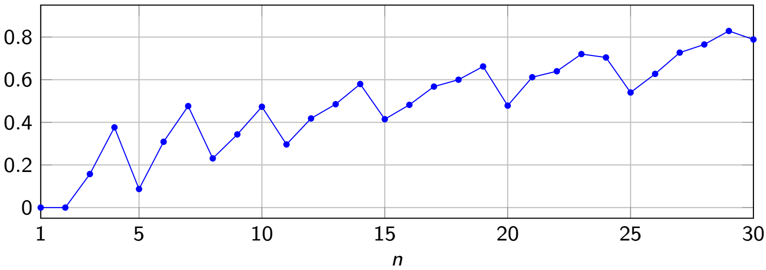
Proposition

Denote the amount of elements deleted from the standard Young tableaux by Cut-Paste to be k, then the resulting standard Young tableaux from the Cut-Paste algorithm starting with input P, of size n from the Robinson–Schensted correspondence does not have the same distribution as P of size $n - k$ generated from the correspondence.

Evidence of Proposition

TestProposition() rounds the size of the average first row in a standard Young tableaux generated by the Robinson–Schensted correspondence of size n to the closest integer k. It then compares the size of the average first row in a standard Young tableaux generated by the Robinson–Schensted correspondence of size $n - k$ of one of size n.

$E(\text{Row}_1 \text{ in } n - (\text{round}(E(\text{Row}_1) \text{ in } n) = k))$ - $E(\text{Row}_2 \text{ in } n)$



Location Statistics

Let $\pi \in S_n$ be a uniformly random permutation, and let

$$RS(\pi) = (P(\pi), Q(\pi))$$

be its Robinson–Schensted image. Here $P(\pi)$ is the insertion tableau and $Q(\pi)$ is the recording tableau. For $1 \leq m \leq n$, define

$$X^P_{n,m} = \text{the cell containing the value } m \text{ in } P(\pi),$$

and

$$X^Q_{n,m} = \text{the cell containing the label } m \text{ in } Q(\pi).$$

Although $X^P_{n,m}$ and $X^Q_{n,m}$ are generally different random variables, their marginal distributions are the same.

Marginal Probability

We first consider the marginal probability

$$p_m(i,j) = \mathbb{P}\left(X^P_{n,m} = (i,j)\right) = \mathbb{P}\left(X^Q_{n,m} = (i,j)\right).$$

The probability is independent of n once $n \geq m$.

Let $\lambda \vdash m$ be a partition of m, and let f^λ denote the number of standard Young tableaux of shape λ. For a cell $c = (i,j)$, let $\mathrm{Cor}(\lambda)$ denote the set of removable corners of λ. Then

$$p_m(i,j) = \frac{1}{m!} \sum_{\substack{\lambda \vdash m \\ (i,j) \in \mathrm{Cor}(\lambda)}} f^\lambda f^{\lambda \setminus (i,j)}.$$

Joint Law

For cells c and d, define

$$J_{n,m}(c,d) = \mathbb{P}\left(X^P_{n,m} = c,\ X^Q_{n,m} = d\right).$$

Unlike the marginal probability $p_m(c)$, the joint probability $J_{n,m}(c,d)$ generally depends on n.

For a partition $\lambda \vdash n$, define

$$A_{\lambda,m}(c) = \#\{T \in \mathrm{SYT}(\lambda) : T(c) = m\}.$$

Thus $A_{\lambda,m}(c)$ counts the number of standard Young tableaux of shape λ in which the entry m appears in cell c.

Joint Law Formula

Proposition

For $1 \leq m \leq n$, the joint law is given by

$$J_{n,m}(c,d) = \frac{1}{n!} \sum_{\lambda \vdash n} A_{\lambda,m}(c) A_{\lambda,m}(d).$$

Equivalently, the quantity $A_{\lambda,m}(c)$ can be expanded using skew tableaux:

$$A_{\lambda,m}(c) = \sum_{\substack{\mu \vdash m \\ \mu \subseteq \lambda \\ c \in \mathrm{Cor}(\mu)}} f^{\mu \setminus c} f^{\lambda/\mu},$$

where $f^{\lambda/\mu}$ denotes the number of standard Young tableaux of skew shape λ/μ.

Asymptotic Independence

The joint law therefore has an exact finite formula, but this formula depends on n. This leads to a natural asymptotic question: for fixed m and fixed cells c, d, do the two locations become independent as $n \to \infty$? In other words, do we have

$$J_{n,m}(c, d) \longrightarrow p_m(c)p_m(d)?$$

Conjecture (Asymptotic independence)

Fix m and cells c, d. Then

$$J_{n,m}(c, d) - p_m(c)p_m(d) = O_m\left(\frac{1}{n^2}\right) \qquad \text{as } n \to \infty.$$

Maple Verification

We verified these conjectures using Maple. The procedure computes the exact marginal probabilities $p_m(c)$ and $p_m(d)$, then computes the joint probability $J_{n,m}(c, d)$ for various values of n. For each choice of m, c, d, we examine the scaled difference

$$n^2 \left(J_{n,m}(c, d) - p_m(c) p_m(d) \right).$$

The conjecture predicts that this quantity should remain bounded as n grows.

These computations indicate that although the marginal distribution of the location of a distinguished entry is already well understood, the joint dependence between its locations in the insertion and recording tableaux contains additional structure.

Row and Column Sums: Preliminaries

Define a row sum of a standard Young tableaux as a vector with *i-th* entry equal to Σ Entries of Row_i. Similarly column sum as Σ Entries of Column_i.

The size of the input standard Young tableaux, n, into a row or column sum can simply be determined by summing the entries in the vector and setting that sum equal to $\frac{n(n+1)}{2}$.

Using the known size of n, we can look at the set of all partitions of n to construct a set of candidate shapes of the initial standard Young tableaux.

Questions on Uniqueness of Sum Vectors

Going back to the row or column sum and pairing it with the set of candidate shapes, we can show that for <u>some</u> vectors, $\exists!$ standard Young tableaux that generates it under the row or column sum algorithm.

I.e. any row or column sum vector that reads lexicographically corresponds to the unique standard Young tableaux that is just a single row $\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array}$ or column $\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array}$ respectively.

However, $\begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array}$ and $\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array}$ have the same row sum vector $(8, 7)$.

Some Natural Questions

Question

What proportion of standard Young tableaux have unique sum vectors?

Question

Which properties of a standard Young tableau determine a bijection with a sum vector?

Question

Does every standard Young tableau have a unique (row sum vector, column sum vector) pair?

Question 3 can quickly be rejected,

1	2	6
3	4	7
5	8	

1	3	5
2	4	8
6	7	

the smallest counterexample.

Further Questions

Question

For a fixed n, what is the maximum number of standard Young tableau described by the same sum vector?

Question

For a fixed n, what is the maximum number of standard Young tableau described by the same (row sum vector, column sum vector) pair?

Question

What proportion of standard Young tableaux have unique (row sum vector, column sum vector) pairs?

Question

Which properties of a standard Young tableau determine a bijection with a (row sum vector, column sum vector) pair?

Further Questions

Question

Do all groups of standard Young tableau described by the same sum vector have the same shape?

Question

Do all groups of standard Young tableau described by the same (row sum vector, column sum vector) pair have the same shape?

Questions 7.2, 7.7, 7.8, 7.9 are left unanswered for the remainder of this paper.

Procedures

Maple package `RSStatFinder.txt` contains procedures to calculate statistics and properties of row sum and column sum vectors of P's & Q's from the the Robinson–Schensted correspondence.

`MaxGroupBoth(n)` calculates by brute-force the maximum number of standard Young tableau of size n described by the same row sum vector OR column sum vector and returns the numbers as a pair.

`MaxGroupPair(n)` calculates by brute-force the maximum number of standard Young tableau of size n described by the same (row sum vector, column sum vector) pair.

`PropUniqueSYTPair(n)` calculates by brute-force the proportion of standard Young tableaux that have unique (row sum vector, column sum vector) pairs.

Largest Sum-Vector Fiber over standard Young tableaux of size n

We provide by exhaustive methods the partial sequences of answers for $n = i$ for i from 1 to 13 to Questions 7.4 and 7.5.

`MaxGroupBoth` =

$$1, 1, 1, 1, 2, 2, 4, 6, 8, 14, 22, 49, 97$$

`MaxGroupPair` =

$$1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 3, 3, 5$$

A search in OEIS yields no results for either sequence, making these likely novel integer sequences and the topic of Sum-Vector Fibers over standard Young tableaux of size n likely new and open.

Proportion of SYT with Unique Sum Vectors

n	Proportion unique by row sum	Proportion unique by column sum
1	1.000000	1.000000
2	1.000000	1.000000
3	1.000000	1.000000
4	1.000000	1.000000
5	0.923077	0.923077
6	0.815789	0.815789
7	0.698276	0.698276
8	0.566754	0.566754
9	0.449618	0.449618
10	0.347094	0.347094
11	0.260029	0.260029

Proportion of SYT with Unique (R, C) Pair

n	Proportion of SYTs with unique (R, C) pair
1	1.000000
2	1.000000
3	1.000000
4	1.000000
5	1.000000
6	1.000000
7	1.000000
8	0.994764
9	0.990840
10	0.984836
11	0.977252

Bibliography I

G. de B. Robinson, *On the representations of the symmetric group*, American Journal of Mathematics **60** (1938), no. 3, 745–760.

C. Schensted, *Longest increasing and decreasing subsequences*, Canadian Journal of Mathematics **13** (1961), 179–191.

D. E. Knuth, *Permutations, matrices, and generalized Young tableaux*, Pacific Journal of Mathematics **34** (1970), no. 3, 709–727.

B. F. Logan and L. A. Shepp, *A variational problem for random Young tableaux*, Advances in Mathematics **26** (1977), no. 2, 206–222.

A. M. Vershik and S. V. Kerov, *Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux*, Soviet Mathematics Doklady **18** (1977), 527–531.

W. Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, London Mathematical Society Student Texts, vol. 35, Cambridge University Press, Cambridge, 1997.

R. P. Stanley, *Enumerative Combinatorics, Volume 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999.

Bibliography II

J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, Journal of the American Mathematical Society **12** (1999), no. 4, 1119–1178.

B. D. McKay, J. Morse, and H. S. Wilf, *The distributions of the entries of Young tableaux*, Journal of Combinatorial Theory, Series A 97 (2002), no. 1, 117–128.

OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, Sequence A000085. `https://oeis.org/A000085`.