Statistics of Standard Young Tableaux Using the Greene-Nijenhuis-Wilf Algorithm

Aurora Hiveley & Lucy Martinez

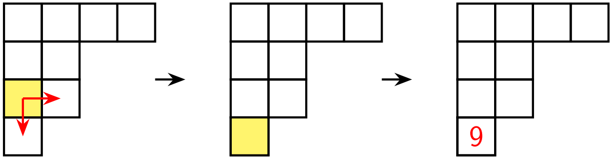
Department of Mathematics
Rutgers University

Experimental Mathematics Final Project Spring 2026

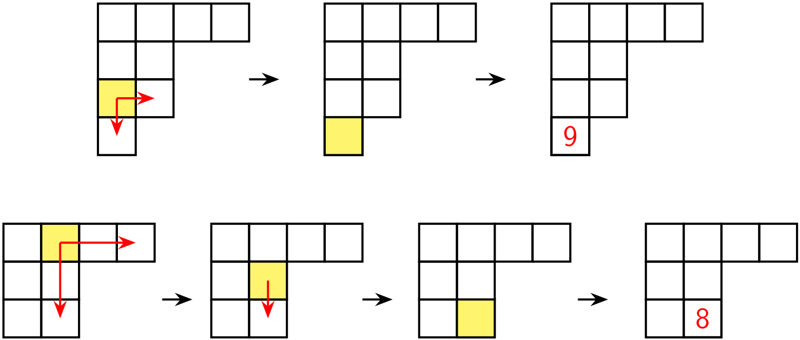
Greene-Nijenhuis-Wilf Algorithm

1. Goal: Calculate statistic $f : T \to \mathbb{R}$ across all tableau T
2. Hard to generate *all* tableau of a given shape
3. Idea: Generate one uniformly at random!
4. Procedure:
   1. Take shape L. Pick cell $[i, j]$ u.a.r.
   2. Identify hook of $[i, j]$. Recall: $H_{i,j} = \{[i', j'] \mid i' > i \wedge j' = j \text{ or } i' = i \wedge j' > j\}$
   3. Pick a new cell $[i', j'] \in H_{ij} \setminus [i, j]$
   4. Repeat until $H_{ij} = \emptyset$. Label that cell n.
   5. Repeat steps (1)-(4) for $L - [i, j]$ until all cells labeled.

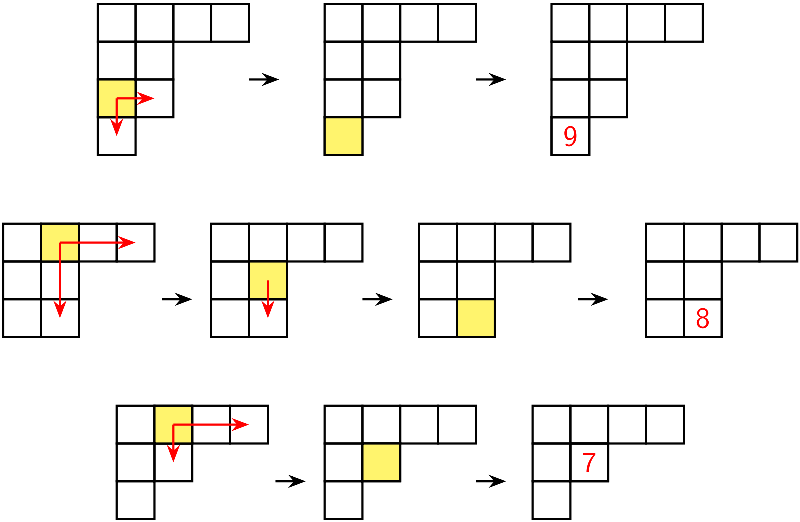
Algorithm Example



Algorithm Example



Algorithm Example



Bootstrapping

Goal: Approximate statistics over the space of all standard Young tableaux of a fixed shape using sampling via the Greene-Nijenhuis-Wilf algorithm.

Bootstrapping

Goal: Approximate statistics over the space of all standard Young tableaux of a fixed shape using sampling via the Greene-Nijenhuis-Wilf algorithm.

Since exact enumeration is not feasible for large shapes, we compute statistics on a finite random sample of Young tableaux and use bootstrap resampling to estimate the standard errors of these estimates.

Bootstrapping

Goal: Approximate statistics over the space of all standard Young tableaux of a fixed shape using sampling via the Greene-Nijenhuis-Wilf algorithm.

Since exact enumeration is not feasible for large shapes, we compute statistics on a finite random sample of Young tableaux and use bootstrap resampling to estimate the standard errors of these estimates.

Suppose that we record all the heights of N people in a list B, here $N = |B|$.

1. One bootstrap: create *one sample* by sampling list B with replacement. Call this sample list B_1 (B_1 is of length N).

Bootstrapping

Goal: Approximate statistics over the space of all standard Young tableaux of a fixed shape using sampling via the Greene-Nijenhuis-Wilf algorithm.

Since exact enumeration is not feasible for large shapes, we compute statistics on a finite random sample of Young tableaux and use bootstrap resampling to estimate the standard errors of these estimates.

Suppose that we record all the heights of N people in a list B, here $N = |B|$.

1. One bootstrap: create *one sample* by sampling list B with replacement. Call this sample list B_1 (B_1 is of length N).
2. Create N bootstraps: Create lists $B_1, B_2, \ldots, B_N$ where each list is a sample from the list B, and $|B_i| = N$ for all i.

Bootstrapping

Goal: Approximate statistics over the space of all standard Young tableaux of a fixed shape using sampling via the Greene-Nijenhuis-Wilf algorithm.

Since exact enumeration is not feasible for large shapes, we compute statistics on a finite random sample of Young tableaux and use bootstrap resampling to estimate the standard errors of these estimates.

Suppose that we record all the heights of N people in a list B, here $N = |B|$.

1. One bootstrap: create *one sample* by sampling list B with replacement. Call this sample list B_1 (B_1 is of length N).
2. Create N bootstraps: Create lists $B_1, B_2, \ldots, B_N$ where each list is a sample from the list B, and $|B_i| = N$ for all i.
3. For each bootstrap, we calculate the average height and get an estimated distribution of the average heights.

Maple Procedures

1. `RandSYT(L)`: inputs a shape `L`, outputs a SYT of shape `L` produced u.a.r. by Greene-Nijenhuis-Wilf algorithm
2. `GenerateManySYT(L,K)`: inputs tableau shape `L` and positive integer `K`, outputs a set of `K` SYT of shape `L` produced by Greene-Nijenhuis-Wilf algorithm, i.e., using `RandSYT(L)`.
   ↪ effectively gives us a bootstrap sample of size `K`
3. `BSsamp(L)`: inputs list `L`, outputs bootstrap sample of size `n := nops(L)`
4. `BSstat1(L,B,c,property)`: generates a sample space of SYT by using `GenerateManySYT(L,B)`, then generates each bootstrap sample using `BSsamp`. Calculates the average value of `property` across this sample for a cell `c`
   ↪ Properties: row sum, column sum, sum of all backwards hook lengths (number of cells in column to left and below `c`). *Note: cell not always needed*

Data: Row Sum of Second Row

$m \backslash n$	1	2	3	4	5	6
2	2	6.6225	13.368	22.5438	34.4313	48.6547
3	2	6.9453	14.6747	26.0495	39.7052	57.4629
4	2	7.4195	16.5274	28.8628	44.8041	64.4912
5	2	7.7569	17.1036	29.9127	48.4158	69.5084
6	2	7.6991	17.5705	31.5533	51.3487	74.6861
7	2	7.9652	18.237	33.9202	53.9115	79.7636
8	2	8.3673	18.6328	35.505	54.4733	81.6526
9	2	8.2149	19.6675	35.5543	56.4635	86.2543
10	2	8.365	19.8075	36.1206	60.37	88.0244

Data: Row Sum of Second Row

m \ n	1	2	3	4	5	6
2	2	6.6225	13.368	22.5438	34.4313	48.6547
3	2	6.9453	14.6747	26.0495	39.7052	57.4629
4	2	7.4195	16.5274	28.8628	44.8041	64.4912
5	2	7.7569	17.1036	29.9127	48.4158	69.5084
6	2	7.6991	17.5705	31.5533	51.3487	74.6861
7	2	7.9652	18.237	33.9202	53.9115	79.7636
8	2	8.3673	18.6328	35.505	54.4733	81.6526
9	2	8.2149	19.6675	35.5543	56.4635	86.2543
10	2	8.365	19.8075	36.1206	60.37	88.0244

Analysis: Row Sum of Second Row

The data for the average sum of the second row of the Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

Analysis: Row Sum of Second Row

The data for the average sum of the second row of the Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

Let $r \geq 1$ be an irrational number then $\mathcal{B} = \{\lfloor r \rfloor, \lfloor 2r \rfloor, \lfloor 3r \rfloor, \ldots\}$ is called a Beatty sequence.

Analysis: Row Sum of Second Row

The data for the average sum of the second row of the Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

Let $r \geq 1$ be an irrational number then $\mathcal{B} = \{\lfloor r \rfloor, \lfloor 2r \rfloor, \lfloor 3r \rfloor, \ldots\}$ is called a Beatty sequence.
Let $r = \frac{3+\sqrt{3}}{2}$, then the Beatty sequence (OEIS A054406) for such r is the following:

$$2, 4, 7, 9, 11, 14, 16, 18, 21, 23, 26, 28, \ldots$$

Analysis: Row Sum of Second Row

The data for the average sum of the second row of the Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

Let $r \geq 1$ be an irrational number then $\mathcal{B} = \{\lfloor r \rfloor, \lfloor 2r \rfloor, \lfloor 3r \rfloor, \ldots\}$ is called a Beatty sequence.

Let $r = \frac{3+\sqrt{3}}{2}$, then the Beatty sequence (OEIS A054406) for such r is the following:

$$2, 4, 7, 9, 11, 14, 16, 18, 21, 23, 26, 28, \ldots$$

Define the n-th partial sum of the Beatty sequence for $r = \frac{3+\sqrt{3}}{2}$. We obtain the following sequence (OEIS A194143):

$$2, 6, 13, 22, 33, 47, \ldots$$

Analysis: Row Sum of Second Row (2)

The data for the average sum of the second row of a Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

The n-th partial sum of the Beatty sequence for $r = \frac{3+\sqrt{3}}{2}$, we obtain the following sequence (OEIS A194143):

$$2, 6, 13, 22, 33, 47, \ldots$$

Analysis: Row Sum of Second Row (2)

The data for the average sum of the second row of a Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

The n-th partial sum of the Beatty sequence for $r = \frac{3+\sqrt{3}}{2}$, we obtain the following sequence (OEIS A194143):

$$2, 6, 13, 22, 33, 47, \ldots$$

If we divide the ratios of the average sums of the second row of a Young tableau of shape $L = (n, n)$ for $90 \leq n \leq 100$, we obtain

Analysis: Row Sum of Second Row (2)

The data for the average sum of the second row of a Young tableau of shape $L = (n, n)$ is given by

$$2, 6.6225, 13.368, 22.5438, 34.4313, 48.6547, \ldots$$

The n-th partial sum of the Beatty sequence for $r = \frac{3+\sqrt{3}}{2}$, we obtain the following sequence (OEIS A194143):

$$2, 6, 13, 22, 33, 47, \ldots$$

If we divide the ratios of the average sums of the second row of a Young tableau of shape $L = (n, n)$ for $90 \leq n \leq 100$, we obtain

$$0.9168057134, 0.9185373162, 0.9149356193, 0.9154768162,$$
$$0.9150371268, 0.9133809905, 0.9115493573, 0.9135424272,$$
$$0.9133154432, 0.9137544543, 0.9116848979$$

Data: Sums of "Backwards Hook" Lengths

$m \backslash n$	1	2	3	4	5	6
1	0	1	2	3	4	5
2	0	1.5258	2.9593	4.118	5.0659	6.3595
3	0	1.9268	3.1693	4.448	5.9111	7.1446
4	0	2.1643	3.4428	4.9423	6.2242	7.2443
5	0	2.1397	3.9404	4.9676	6.1988	7.4812
6	0	2.3325	4.0944	5.1536	6.6112	7.8091
7	0	2.2707	4.0357	5.4661	6.5983	7.9267
8	0	2.4478	3.9085	5.5521	6.822	8.0639
9	0	2.4913	3.9021	5.7733	6.7289	7.926
10	0	2.4139	4.6701	5.6252	6.8248	8.4535

Data: Standard Errors

★ Why? Bootstrap samples can sway statistics. Want to check for stability/replicability

- Row & Column Sums:
  - Similar increase in error as n increases (columns added)
  - Bigger increase in col. sums than row sums as m increases (rows added)
  - Similar errors when $n \approx m$, since similar numbers can be added to new row & col. of tableau
  - Each error is around $1\% - 2\%$ of the average
  - Conclusion: very stable!
- Backwards hooks:
  - Grow much slower as n and/or m increase
  - Each error is around $5\% - 7\%$ of the average
  - Conclusion: pretty stable! Slightly less than row/col. sums

Sources

- B. Efron *Bootstrap Methods: Another Look at the Jackknife*. The Annals of Statistics **7**, no. 1, 1-26 (1979).
- OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2023. Published electronically at `http://oeis.org`.
- C. Greene, A. Nijenhuis, and H. S. Wilf. *A Probabilistic Proof of a Formula for the Number of Young Tableaux of a Given Shape*. Advances in Mathematics **31**, 104-109 (1979).
- R. Richardson. *An Introduction to the Bootstrap Method*. Published electronically at `https://richardson.byu.edu/624/myl10/bootstrapping.pdf`.