Graduate Online Combinatorics Colloquium

Nadia Lafrenière

4/19/2023

How many times should ya shuffle a deck of cards?

Contents: ① Combinatorial lower bounds
② Upper bounds
③ Cutoff phenomenon
④ Take away

① Lower bounds : features and impossible permutations.

Shuffling method : random transpositions

To shuffle n cards:

- Pick two numbers a and b in $[n]$, at random (repetitions allowed)
- If $a \neq b$, swap what is at positions a and b.
- Repeat several times. How many?

Examples : permutations of $[20]$, generated with 20 transpositions.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \\ 17 & 18 & 19 & 1 & 4 & 12 & 2 & 20 & 8 & 10 & 11 & 9 & 5 & 14 & 6 & 7 & 16 & 13 & 3 & 15 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 \\ 11 & 2 & 1 & 3 & 4 & 10 & 13 & 15 & 5 & 9 & 6 & 20 & 17 & 8 & 14 & 16 & 7 & 18 & 19 & 12 \end{pmatrix}$$

Are n transpositions enough?

- If a number in $[n]$ has not been picked yet, it is for sure a fixed point of the permutation.
- k transpositions mean that $2 \cdot k$ entries may have been chosen, but we choose with repetitions.

- How many fixed points does a permutation typically have?
  - In average, 1 regardless of n.
  - The # of fixed points converge to a Poisson distribution (with parameter 1)

  => Very few fixed points.

- Reducing the number of fixed points by choosing all elements in $[n]$ within $2k$ draws (with repetitions)

  Coupon collector problem

  Time necessary to select all the n "coupons" (hockey cards?) if we draw with replacement : $n \cdot (\ln(n) + \gamma + \frac{1}{n} + O(\frac{1}{n^2}))$, in average.

  $\gamma \approx 0.577$

Therefore, if $2k \leq n \cdot \ln(n)$, we should expect to have fixed points.

Theorem (Diaconis, Shahshahani, 1984)

The mixing time of the random transposition shuffle is $\frac{1}{2} n \ln(n)$.

↳ time for the shuffle to be close to uniform.

Shuffling method : Top-to-random

To shuffle n cards:

- remove the top card
- insert it anywhere, uniformly at random (including on top)
- repeat several times. How many?

Remarks: - Except for the top card (that moves below), no card is going down.

- Until card $n-1$ reaches the top and is reinserted, $n-1$ and n are ordered (so $1/2$ of permutations are impossible). When is card $n-1$ on top, then inserted?

Dual shuffle: Random-to-top

Remarks: - As long as cards haven't been brought to top, they are in same relative order.

- All permutations are possible only if all cards but one have been selected => Coupon collector problem!

Theorem (Diaconis, Fill, Pitman, 1992)

It takes $n \ln(n) + n$ repetitions of the top-to-random shuffle to get close to the uniform distribution.

Shuffling method: Random-to-below

- Remove a card uniformly at random
- Reinsert it anywhere weakly below, uniformly at random
- Repeat several times. How many?

- Similar to top-to-random
- "Lazier" than top-to-random: it happens more often that no card move (e.g. if the bottom card is drawn)
- We can imitate the technique for top-to-random by placing a bookmark on top of the bottom card, and stop when the bookmark reaches the top.

Theorem (Grinberg, Lafrenière, 2023+)

It takes $n \ln(n) + n \ln(\ln(n)) + n$ repetitions of the random-to-below shuffle to get close to the uniform distribution.

Shuffling method: Riffle shuffle

- Cut the deck roughly in half
- Interlace the cards from each hand
- Repeat several times. How many?

"Feature" / Permutation statistic: A rising sequence in a permutation is a sequence $\sigma(i), \sigma(i)+1, \sigma(i)+2, \dots, \sigma(i)+j$ that appears from left to right in a permutation. Example: 5 4 6 2 1 9 3 7 8 $\mapsto \{1, 23, 4, 56\ 78, 9\}$

How many rising sequences are there in a permutation?

Remark: - A rising sequence of σ is a run in σ^{-1}

Example: $\sigma = 546219378$, $\sigma^{-1} = 5|4|7|2|1|3|8|9|6$

- Runs are separated with descents.

$\Rightarrow$ Avg. # of rising sequences = Avg # of runs = Avg # of descents + 1 = $\frac{n+1}{2}$.

How many rising sequences are there in a permutation obtained from 1 to n with k riffle shuffles?

- In 1 riffle shuffle, the cards in each hand are kept in the same relative order $\Rightarrow \leq 2$ rising sequences
- In k shuffles, there are $\leq 2^k$ rising sequences.
- Therefore, for all permutations to be likely, we need

$$2^k \geq n \Rightarrow k \geq \log_2(n).$$

Theorem (Bayer, Diaconis, 1992)

It takes $\frac{3}{2}\log_2(n)$ repetitions of the riffle shuffle to get close to uniform distribution.

Shuffling method: Random-to-random

- Pick a card, uniformly at random
- Move it anywhere in the deck, uniformly at random
- Repeat several times. How many?

We don't know the distinguishing feature of random-to-random.

Random transpositions ⟶ Fixed points

Top-to-random / Random-to-below ⟶ Order of the bottom cards

Riffle shuffle ⟶ Rising sequences

Random-to-random ⟶ ?

Fortunately, other techniques exist!

Theorem (Bernstein, Nestoridi, 2018)

It takes $\frac{3}{4}n\ln(n) - \frac{1}{4}n\ln(\ln(n))$ repetitions of the random-to-random shuffle to get close to the uniform distribution.

Upper bounds

Showing something is not random is hard, and showing that it is random can be even harder!

Several techniques exist, including strong stationary times and distances on probability distributions.

Stationary time: Time at which you know you have reached the uniform distribution

In top-to-random and random-to-below, the cards below the book mark can happen in any order, all with the same probability. This is a stationary time.

In general, stationary time are not known or very hard to keep track of for card shuffling.

Total variation distance: Given two probability distributions ν and μ on sample space Ω, the total variation distance is

$$\|\nu-\mu\|_{TV} = \frac{1}{2}\sum_{\omega\in\Omega} |\nu(\omega)-\mu(\omega)|.$$

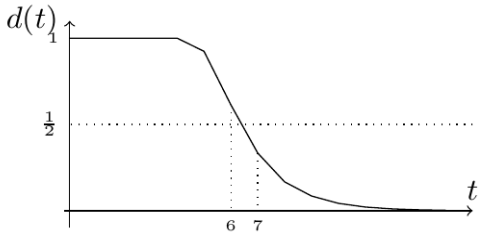
← Number ranging from 1 (very far apart) to 0 (the same)

For card shuffling, $\Omega = S_n$, $\mu(\omega)=\frac{1}{n!}$ for every permutation, and ν depends on the shuffle.

When the distance is below some treshold (say $\frac{1}{2}$), we say that the order of the deck is close to the uniform distribution.

Remarks: - it is often hard to compute
- there exists other metrics as well (e.g. the separation distance) that give other information.

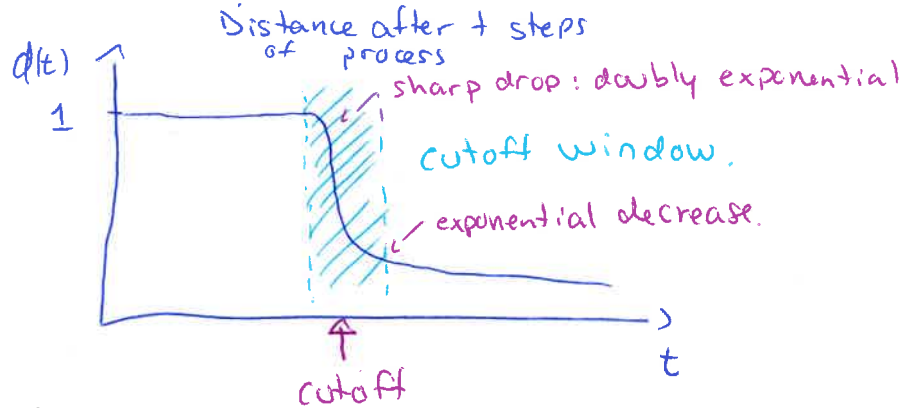
Total variation distance after k riffle shuffles



Cutoff

How do we know when the distance for two probability distributions is small enough?

Many processes exhibit the cutoff phenomenon: a sharp drop in the distance between the probability distribution of the process and its stationary distribution.



What processes have cutoff?

- Many "global" processes: riffle shuffle, random transpositions, top-to-random, ...
- Some Markov chains (some variants of random walks on graphs or Ehrenfest's urns; the drunkard problem) have no cutoff.
- It is a sensitive process: changing small factors (such as initial conditions) can break cutoff.
- Open problem: Can we predict cutoff without having to explicitly compute the distance between two probability distributions?

Take away (advice I received from Persi Diaconis).

- There are many things we can compute, but many are hard.
- Don't bother too much about the metric: if you do, you don't get anything done and only wonder about the metric.
- Different metrics yield different results and have different interpretations: they are all one part of the story.