SHUFFLING CARDS

GUHA SRIRAM

Abstract. In this paper, we will discuss the shuffling of playing cards and how it affects the randomness of the deck as a whole. To do this, we will be analyzing different kinds of shuffles, variation distances, strong uniform stopping rules, and other related ideas.

We will start with two simple, yet powerful combinatorial problems. They will prove to be useful in later sections.

1. The Birthday Paradox

Let us take n random people. What is the probability that they all have different birthdays? We have

$$p(n) = \prod_{i=1}^{n-1}(1 - \frac{i}{365}).$$

This is quite simple to see; If the first i people have different birthdays, the probability that the $(i+1)$-st person doesn't repeat a birthday is $1 - \frac{i}{365}$, because there are 365 - i birthdays left.

Remark 1.1. For the Birthday Paradox formula, we have $p(n) < \frac{1}{2}$ for $n = 23$, $p(n) \approx .01$ for $n = 70$, and *exactly* 0 for $n > 365$(Pigeonhole Principle). A closely related idea to the Birthday Paradox is its inverse; the probability of having two people with the *same* birthday out of n people. We have opposite probabilities; with the probability hitting .99 after only 70 people. One would think that you need far more people to satisfy a birthday repeat!

2. The Coupon Collector

Let's say we're collecting baseball cards. They are sold in envelopes where we cannot see which card we are getting, so there is no restriction on buying the same card multiple times. If there are n different cards, what is the expected number of pictures we have to buy until we get each one at least once? If we already have k distinct cards, the probability to get one we already have is $\frac{k}{n}$. Thus, the probability to need exactly s drawings for the next new card is

$$\left(\frac{k}{n}\right)^{s-1}\left(1 - \frac{k}{n}\right).$$

And so the expected number of times we have to buy for the next new card is

$$\sum_{s \geq 1}\left(\frac{k}{n}\right)^{s-1}(1 - \frac{k}{n})s,$$

where we are summing the values of s as our number of cards bought changes. We have a very nice way of calculating this; it goes as follows:

$$\sum_{s\geq 1}\left(\frac{k}{n}\right)^{s-1}\left(1-\frac{k}{n}\right)s$$

$$=\sum_{s\geq 1}\left(\frac{k}{n}\right)^{s-1}s-\sum_{s\geq 1}\left(\frac{k}{n}\right)^{s}s$$
$$=\sum_{s\geq 0}\left(\frac{k}{n}\right)^{s}(s+1)-\sum_{s\geq 0}\left(\frac{k}{n}\right)^{s}s$$
$$=\sum_{s\geq 0}\left(\frac{k}{n}\right)^{s}$$

We will now define a short lemma that will help us take this further.

Lemma 2.1. *The infinite geometric series*

$$Q=\frac{1}{q}+\frac{1}{q^2}+\frac{1}{q^3}+...$$

has $Q=$

$$\frac{1}{q-1}$$

Proof. We have the geometric series

$$Q=\frac{1}{q}+\frac{1}{q^2}+\frac{1}{q^3}+...$$

Assuming $q>1$, we have

$$qQ=1+\frac{1}{q}+\frac{1}{q^2}+...=1+Q$$

and thus,

$$Q=\frac{1}{q-`1}$$

■

Back to the coupon collector, we now have

$$\sum_{s\geq 0}\left(\frac{k}{n}\right)^{s}=\frac{1}{1-\frac{k}{n}},$$

as the expected number of buys for the next new card. Now, the expected number of buys until we have *each* of the n cards is

$$\sum_{k=0}^{n-1}\frac{1}{1-\frac{k}{n}}=\frac{n}{n}+\frac{n}{n-1}+...+\frac{n}{2}+\frac{n}{1}\approx n\log n$$

So, our answer to the coupon collector problem is that one needs roughly $n\log n$ purchases. Now, let us find the probability that we need more than $n\log n$ purchases. If we let V_n be the number of cards we need to buy(remember, $E[V_n]\approx n\log n$), for $n\geq 1$ and $c\geq 0$, the probability that we need more than $m=\lceil n\log n+cn\rceil$ purchases is

$$\text{Prob}[V_n > m] \leq e^{-c}$$

This can be seen if we let A_i be the event that card i is not purchased in the first m purchases, then

$$\text{Prob}[V_n > m] = \text{Prob}[\bigcup_i A_i] \leq \sum_i \text{Prob}[A_i] = n\left(1 - \frac{1}{n}\right)^m < ne^{\frac{-m}{n}} \leq e^{-c}$$

Let us now transition to a deck of cards. We will label them 1 through n, with the top card being 1 and the bottom one being n. From now on, we will call ϑ_n the set of all the permutations of the list 1, 2, 3,...n. We define *shuffling* to be applying certain *permutations* to the order of the cards. Hopefully, this permutation is completely arbitrary(i.e. we apply some random permutation $\pi \in \vartheta_n$, each of whose probability is $\frac{1}{n!}$). If we were able to do this, that would be quite the miracle and indeed, very ideal for card games. However, this isn't quite what happens. When we shuffle one time in real life, certain permutations occur with a higher probability than others, and after we shuffle more, we expect the deck to be as "close to random" as it can be.

3. Top-in-at-random shuffles

Though maybe not the most effective kind of shuffling, top-in-at-random shuffles can be quite good at generating randomness(assuming one has the patience to do it many, many times!). These shuffles are performed like this: take the top card from the deck, and insert it back in at a completely random place, n of which exist. Now, how do we exactly measure "randomness"? We do this using variation distance of probability distributions. Here are some examples of distributions: our starting distribution E, which is

$$\text{E(id)} = 1; \text{ (the probability of it being } \textit{itself} \text{ is 1)}$$
$$\text{E}(\pi) = 1; \text{ (the probability of it being any other random permutation is 0)}$$

And we also have the uniform distribution U, which is defined as

$$\text{U}(\pi) = \tfrac{1}{n!} \text{ for all } \pi \in \vartheta_n$$

Definition 3.1. The *variation distance* between two probability distributions Q_1 and Q_2 is

$$||Q_1 - Q_2|| := \frac{1}{2} \sum_{\pi \in \vartheta_n} |Q_1(\pi) - Q_2(\pi)|.$$

In the following equations, "being close to random" can be interpreted as "having small variation distance from the uniform distribution."The distance between the starting distribution and uniform distribution is quite the opposite:

$$||E - U|| = 1 - \frac{1}{n!}$$

And after 1 top-in-at-random shuffle:

$$||\text{Top} - \text{U}|| = 1 - \frac{1}{(n-1)!}$$

Not much better. Let us call the probability distribution on ϑ_n that we get from doing top-in-at-random shuffles k times as Top^k. We want to find out how the variation distance $d(k) := ||\text{Top}^k - \text{U}||$ goes to zero exponentially fast, as we can probably infer. We will do this using strong uniform stopping rules.

4. Strong uniform stopping rules

Imagine if a casino manager analyzed the shuffling of a deck of cards, and watches what permutations are applied. After a number of shuffles *depending* on the permutations they saw, they say "STOP!". In other words, they have some *stopping rule* that they use to end the shuffling.

Definition 4.1. The stopping rule is *strong uniform* if the following is true for all $k \geq 0$:

If the process is stopped after exactly k steps, then the resulting permutations of the deck have uniform distribution.

Let T be the number of shuffles until the stopping rule tells the manager to say "STOP!"; T is random and depends on the stopping rule. In addition, let the ordering of the deck after k shuffles be given by a random variable $X_k \in \vartheta_m$. So we have the stopping rule is random if for all values of k,

$$\text{Prob}[X_k = \pi | T = k] = \frac{1}{n!} \text{ for all } \pi \in \vartheta_n$$

This allows us to find stopping rules, as well as give us effective upper bounds on variation distances like $d(k) = ||\text{Top}^k - \text{U}||$. For example, a strong uniform stopping rule for top-in-at-random shuffles is

Stop shuffling after the original bottom card(nth card) is first inserted back into the deck.

We can see this easily with 5 cards:

1, 2, 3, 4, **5**
2, 3, 4, 1, **5**
3, 2, 4, 1, **5**
2, 4, 1, **5**, 3
4, 1, **5**, 3, 2
1, **5**, 3, 4, 2
...

As we can see, the ordering of the cards underneath the 5 is quite uniform. Now let T_i be the random variable which keeps track of the number of shuffles until, for the first time, i cards lie below the last card in the deck(labeled n). So we now have to determine the probability distribution of

$$T = T_1 + (T_2 - T_1) + ...(T_{n-1} - T_{n-2}) + (T - T_{n-1})$$

where each $T_i - T_{i-1}$ corresponds to the number of shuffles between when *i-1* cards lie beneath n and when i cards lie beneath n. Does this sound familiar? Yes! Each $T_i - T_{i-1}$ is also the number of purchases the coupon collector needs from the T_{i-1}-th card to the T_ith card. And because the coupon collector and the top-in-at-random shuffler perform the same kind of processes, we can say that the strong uniform stopping rule for top-in-at-random shuffling takes more than $k = \lceil n \log n + cn \rceil$ steps with probability:

$$\text{Prob}[T > k] \leq \text{e}^{-c}(\text{as we claimed for the coupon collector problem})$$

And this means that after $k = \lceil n \log n + cn \rceil$ top-in-at-random shuffles, we have

$$d(k) = ||\text{Top}^k - U|| \leq \text{e}^{-c}.$$

We will prove this using the following lemma.

Lemma 4.2. *Let $Q : \vartheta_n \to \mathbb{R}$ be any probability distribution that defines a shuffling process Q^k with a strong uniform stopping rule whose stopping time is T. Then for all $k \geq 0$,*

$$||Q^k - U|| \leq Prob[T > K].$$

Proof. If X is a random variable with values in ϑ_n, with probability distribution Q, then we write Q(S) for the probability that X takes a value in $S \subseteq \vartheta_n$. Thus we have Q(S) = Prob[X∈S], and in uniform distribution $Q = U$ we get

$$\mathrm{U}(S) = \mathrm{Prob}[\mathrm{X} \in S] = \frac{|S|}{\mathrm{n!}}$$

For every subset S $\subseteq \vartheta_n$, we have the probability after k steps our deck is ordered in a permutation in S as
$\mathrm{Q}^k(S) = \mathrm{Prob}[\mathrm{X}_k \in S]$

$$\begin{aligned}
&= \sum_{j \leq k} \mathrm{Prob}[\mathrm{X}_k \in S \wedge T = j] + \mathrm{Prob}[\mathrm{X}_k \in S \wedge T > k] \\
&= \sum_{j \leq k} \mathrm{U}(S)\mathrm{Prob}[T = j] + \mathrm{Prob}[\mathrm{X}_k \in S | T > k] \cdot \mathrm{Prob}[T > k] \\
&= \mathrm{U}(S)(1 - \mathrm{Prob}[T > k]) + \mathrm{Prob}[\mathrm{X}_k \in S | T > k] \cdot \mathrm{Prob}[T > k] \\
&= \mathrm{U}(S) + (\mathrm{Prob}[\mathrm{X}_k \in S | T > k] - \mathrm{U}(S)) \cdot \mathrm{Prob}[T > k].
\end{aligned}$$

So we get

$$|\mathrm{Q}^k(S) - \mathrm{U}(S)| \leq \mathrm{Prob}[T > k]$$

and because

$$\mathrm{Prob}[\mathrm{X}_k \in S | T > k] - \mathrm{U}(S)$$

is the difference of two probabilities, it has $| \cdot | \leq 1$. ■

So we have proved this upper bound for the number of top-in-at-random shuffles needed to get "close to random".

Theorem 4.3. *Let $c \geq 0$ and $k := \lceil n \log n + cn \rceil$. After performing k top-in-at-random shuffles on a deck of n cards, the variation distance from the uniform distribution satisfies*

$$d(k) := ||Top^k - U|| \leq e^{-c}.$$

This shows us the true inefficiency of top-in-at-random shuffles - with the bounds given by our theorem, we need more than $n \log n \approx 205$ shuffles until a standard deck of 52 cards reaches uniform distribution! So, let us move on to the one and only(and far superior) **riffle shuffle**.

5. Riffle shuffles

A *riffle shuffle* is performed by taking a deck of cards, splitting into two parts, and interleaving them in some way. This can be done by dropping cards from alternating parts in some irregular pattern, or "riffling" the parts together, holding a half-deck in each hand. Let's say the deck is split so that the top t cards are in one hand, and the rest of the deck(n-t cards) is in the left hand. We have $\binom{n}{t}$ distinct ways to interleave the two hands of cards, each of which generates a different permutation. However, we must remember that for each t, there is exactly 1 possibility that each card is inserted right back into the spot it was removed from(identity permutation). For example, if we took 26 cards off the top of the deck, the identity permutation would be if we inserted those 26 cards in the same order back on top of the deck. Now, let us look at the probability distribution of the function Rif on ϑ_n in two different ways.

1. Rif : $\vartheta_n \to \mathbb{R}$ is defined as

$$\mathrm{Rif}(\pi) = \begin{cases} \frac{n+1}{2^n} & \text{if } \pi = \mathrm{id} \\ \frac{1}{2^n} & \text{if } \pi \text{ has two separate, consecutive, increasing sequences} \\ 0 & \text{otherwise} \end{cases}$$

2. *Inverse shuffling* is when a subset of cards in the deck is taken out from the deck, and placed on top of the remaining cards. The order in both parts of the deck is not changed. Another way to think about reverse shuffling is assigning a "1" or a "0" to each card(with equal probability), and moving the cards labeled "0" to the top.
It's quite easy to see that these two descriptions are the same; whenever each and every single card labeled "0" is on top of each and every single card labeled "1", we have the identity permutation. Now, let us dive into the analysis of riffle shuffles.

Theorem 5.1. *After performing k riffle shuffles on a deck of n cards, the variation distance from a uniform distribution satisfies the inequality*

$$||Rif^k - U|| \leq 1 - \prod_{i=1}^{n-1}(1 - \frac{i}{2^k}).$$

Proof. Inverse shuffles correspond to the probability distribution $\overline{\mathrm{Rif}}(\pi) := \mathrm{Rif}(\pi^{-1})$ Because every permutation has its unique inverse, and the statement $\mathrm{U}(\pi) = \mathrm{U}(\pi^{-1})$ allow us to say

$$||\mathrm{Rif}^k - \mathrm{U}|| = ||\overline{\mathrm{Rif}^k} - \mathrm{U}||$$

(Reeds'inversion lemma)
So in every inverse riffle shuffle, every card gets assigned a digit 0 or 1, as we defined earlier. Let's say we inverse riffle shuffle k times, then each card has a string of length k 0s and 1s attached to it. Our stopping rule is :

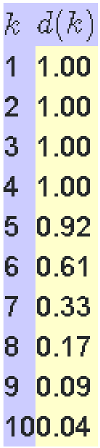
"Stop shuffling as soon as all cards have distinct strings attached to them" Since each assigned "bit" is completely random, as well as independent of previously assigned bits, this stopping rule is strong uniform! For example, if we take $n = 5$ cards, we need $T = 3$ inverse shuffles until every card has a different string. The time T taken by our stopping rule for inverse shuffles can actually be related to that of the birthday paradox! We will put two cards in the same "box" if they have the same k-long string. There are $K = 2^k$

boxes(1 for each possible string of 0s and 1s, we use 2^k in place of K in the birthday paradox formula). We want the probability that there is *more* than one card, so we have

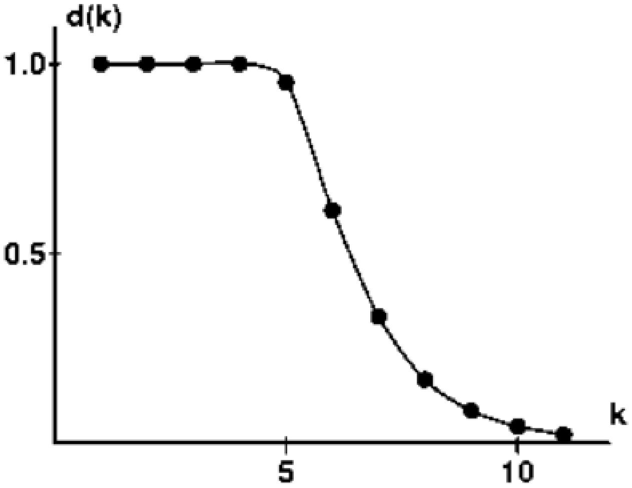
$$\text{Prob}[T > k] = 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right)$$

■

So for n = 52 cards, our upper bound from 5.1 gives us $d(12) \leq 0.28$ should be "random enough".



k	$d(k)$
1	1.00
2	1.00
3	1.00
4	1.00
5	0.92
6	0.61
7	0.33
8	0.17
9	0.09
10	0.04



The graph above gives us a good idea of what is called the *cutoff phenomena.* While we would expect *d(k)* to decrease exponentially, it instead suddenly cuts off at about $k = 5$! It has been found that 7 is a good stopping point, though the proof is beyond the scope of the paper. However, let us track the bottom card in the deck as we riffle shuffle. Say it's the 9 of hearts. To riffle shuffle, we split the deck into 2 parts. The 9 ♡ is on the bottom of one of the half-decks because it was on the bottom of the main deck. Every time we shuffle, the 9 ♡ has a $\frac{1}{2}$ chance of staying on the bottom, because the bottom card of the big deck has a $\frac{1}{2}$ chance of coming from each of the two smaller decks.

References

[1] Martin Aigner and Gunter M. Ziegler. *Proofs from the Book.* Springer-Verlag GmbH, Berlin, Germany, 1998.

[2] Dave Bayer and Persi Diaconis. *Trailing the Dovetail Shuffle to its Lair.* Columbia University and Harvard University, 1992.

[3] David Austin. *How Many Times do I Have to Shuffle this Deck.* American Mathematical Society. `https://rb.gy/htnq9a`