An Analysis of Riffle Shuffles

Tong Chen

June 28, 2020

1 Definition & Analysis

How many shuffles are required to bring a deck of cards close to random? Before talking about how to shuffle, how to definite a suit is random?

Definition 1. *One suit is called* ***shuffled*** *if and only if the probability of every permutation is equal.*

And now we analyze the most commonly used method of shuffling cards called the ordinary riffle shuffle. This involves cutting the deck approximately in half, and interleaving the two halves together. In general, a permutation π of n cards made by a riffle shuffle will have exactly 2 rising sequences unless it is the identity. Conversely, any permutation of n cards with 1 or 2 rising sequences can be obtained by a physical riffle. Thus the mathematical definition of a *riffle shuffle* is "a permutation with 1 or 2 rising sequences". Suppose c cards are initially cut off the top. Then there are $\binom{n}{c}$ possible riffle shuffles(1 of which is the identity). Finally, the total number of possible riffle shuffles is

$$1 + \sum_{c=0}^{n} (\binom{n}{c} - 1) = 2^n - n$$

The following model for random riffle shuffle, suggested by Gilbert and Shannon(1955) and Reeds(1981), is mathematically tractable and qualitatively similar to shuffles done by simple card players.

Definition 2. *(1st description). Begin by choose an integer c from 0,1,...,n according to the binomial distribution* $P\{C = c\} = \frac{1}{2^n}\binom{n}{c}$*. Then, c cards are cut off and held in the left hand, and n-c cards are held in the right hand. The cards are dropped from a given hand with probability proportional to packet size. Thus, the chance that a card is first dropped from the*

left hand packet is c/n. If this happens, the chance that the next card is dropped from the left packet is (c-1)/(n-1).

There are two other descriptions of this shuffling mechanism that are useful.

Definition 3. *(2nd description). Cut an n card descriptions according to a binomial distribution. If c cards are cut off, pick one of the* $\binom{n}{c}$ *possible shuffles uniformly.*

Definition 4. *(3rd description). This generates* π^{-1} *with the correct probability. Label the back of each card with the result of an independent, fair coin flip as 0 or 1. Remove all cards labelled 0 and place them on top of the deck, keeping them in the same relative order.*

Lemma 5. *The three descriptions yield the same probability distribution.*

证明. The second and third descriptions are equivalent. Indeed, the binary labelling chooses a binomially distributed number of zeros, and contidional on this choice, all possible placements of the zeros are equally likely.

The first and second descriptions are equivalent. Suppose c cards have been cut off. For the first description, a given shuffle is specified by a sequence $D_1, D_2, \ldots, D_n$, where each D_i can be determined by multiplying the chance at each stage, is $c!(n-c)!/n!$. □

The argument to follow analyzes the repeated inverse shuffle. This has the same distance to uniform as repeated shuffling because of the following lemma.

Lemma 6. *Let G be a finite group,* $T : G \to G$ *one-to-one, and Q a probability on G. Then*

$$\|Q - U\| = \|QT^{-1} - U\|$$

where $QT^{-1}(g) = Q(T^{-1}(g))$ *is the probability induced by T.*

The result of repeated inverse shuffles of n cards can be recorded by forming a binary matrix with n rows. The first column records the zeros and ones that determine the first shuffle, and so on. The i-th row of the matrix is associated to the i-th card in the original ordering of the deck, recording in coordinate j the behavior of this card on the j-th shuffle.

Lemma 7. *(Reeds). Let T be the first time that the binary matrix formed from inverse shuffling has distrinct rows. Then T is a strong uniform time.*

证明. The matrix can be considered as formed by flipping a fair coin to fill out the i, j entry. At every stage, the rows are independent binary vectors. The joint distribution of the rows, conditional on being all distrinct, is invariant under permutations.

表 1: Total variation distance for m shufles of 52 cards

0	1	2	3	4
a 1101	c 0010	c 0010	f 1000	f 1000
b 1100	e 0110	d 1011	a 1101	b 1100
c 0010	a 1101	f 1000	b 1100	c 0010
d 1011	b 1100	e 0110	c 0010	e 0110
e 0110	d 1011	a 1101	d 1011	a 1101
f 1000	f 1000	b 1100	e 0110	d 1011

After the first inverse shuffle, all cards associated to binary vectors starting with 0 are above cards with binary vectors starting with 1. After two shuffles, cards associated with binary vectors starting (0, 0) are on top followed by cards associated to vectors beginning (1, 0), followed by (0, 1), followed by (1, 1) at the bottom of the deck.

Inductively, the inverse shuffles sort the binary vectors(from right to left) in lexicographic order. At time T the vectors are all distrinct, and all sorted. By permutation invariance, any of the n cards is equally likely to have been associated with the smallest row of the matrix (and so be on top). Similarly, at time T, all $n!$ orders are equally likely. □

To complete this analysis, the chance that $T > k$ must be computed. This is simply the probability that if n balls are dropped into 2^k boxes there are not two or more balls in a box. If the balls are thought of as people, and the boxes as birthdays, we have the familiar question of the birthday problem and its well-known answer. This yields:

Theorem 8. *For Q the Gilbert-Shannon-Reeds distribution defined in Lemma 5,*

$$\|Q^{*k} - U\| \leq P(T > k) = 1 - \prod_{i=1}^{n-1}(1 - \frac{1}{2^k})$$

Standard calculus shows that if $k = 2\log_2(n/c)$. In this sense, $2\log_2 n$ is the cut off point for this bound.

2 Anothor Idea

The idea here is to consider shuffling as inverse sorting. The argument works for any symmetric method of labelling the cards. We know that the worst case of sorting is $\Theta(n \log n)$. And the mininum operation for riffle shuffle is $\Theta(\log n)$.

According to Trailing the Dovetail Shuffle to Its Lair written by Dave Bayer and Persi Diaconis, we can get the conclusion: 6-8 times shuffling make the suit random. Here, if S_n is the symmetric group, U the uniform probability [so $U(\pi) = l/n!$] and Q^m the Gilbert-Shannon-Reeds probability after m shuffles, then the total variation distance is defined as

$$\|Q^m - U\| = \max_{A \subset S_n} |Q^m(A) - U(A)|$$

Table 2 gives the total variation distance for 52 cards. Table 1 shows that the total variation distance stays essentially at its maximum of 1 up to 5 shuffles, when it begins to decrease sharply by factors of 2 each time. This is an example of the cutoff phenomena described by Aldous and Diaconis (1986).

表 2: Total variation distance for m shufles of 52 cards

m	1	2	3	4	5	6	7	8	9	10
$\|Q^m - U\|$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043

3 Reference

1. Bayer D , Diaconis P . Trailing the Dovetail Shuffle to its Lair[J]. The Annals of Applied Probability, 1992, 2(2):294-313.

2. David, Aldous, Persi, et al. Shuffling Cards and Stopping Times[J]. The American Mathematical Monthly, 1986.