CARD SHUFFLING

JUSTIN GUO

Abstract. We examine the most common model of card shuffling, the GSR-shuffle. We prove and analyze the meaning of the famous result that "seven riffle shuffles are sufficient to randomize a deck." Then, this proof is extended to a deck with repeated cards.

Contents

1. Introduction

The GSR-shuffle is a model of riffle shuffling developed by Gilbert and Shannon in 1955 and independently by Reeds in 1981. It is the most commonly used model in the literature on card shuffling problems because it comes from an intuitive notion of how riffle shuffling is done in practice. In 1992, P. Diaconis and D. Bayer published the now famous result that "seven riffle shuffles are sufficient to randomize a deck of cards." In Section 2, we reproduce that proof. In Section 3, we define two games to more closely consider how riffle shuffles lead to randomization. The results of the first game, Premo, are attributed to the same Diaconis and Bayer paper. The other game, New Age Solitaire, is attributed to A. van Zuylen and F. Schalekamp (2004). In Section 4, we examine an extension of the Section 2 results, attributed to M. Conger and D. Viswanath (2006), to decks in which cards can be repeated.

2. The GSR-Shuffle

The approach we will take in this section will be first to formalize the definition of the GSR-shuffle and then to derive a formula for the probability that a GSR-shuffle yields a certain ordering of the deck. Next, we need a tool to measure how much randomness is produced by a shuffle. One good choice is total variation distance, which measures the difference in distribution between two shuffles. Calculating how total variation distance from a uniform distribution depends on the number of shuffles performed leads to the general rule that $\frac{3}{2}\log_2(n)$ riffle shuffles sufficiently randomize a deck of n cards.

Date: September 13, 2013.

Definition 2.1. The *symmetric group*, S_n, is the set of all bijections from $\{1, 2, ..., n\}$ to $\{1, 2, ..., n\}$.

We can think of elements of S_n as permutations of $\{1, 2, ..., n\}$. Equivalently, we can also think about them as distributions of the deck, or in other words, ways of ordering a deck labelled $\{1, 2, ..., n\}$. For the purposes of the paper, we will require both interpretations.

In practice, a riffle shuffle consists of two steps. First, all the cards in the deck are separated into two piles. Then, the piles are interlaced together. We develop the model of the GSR-shuffle to reflect these actions.

Definitions 2.2. An *a-shuffle* is achieved by cutting a deck of n cards into a packets such that the number of cards in each packet is distributed according to a multinomial distribution. That is, the probability of having j_1 cards in the first packet, j_2 cards in the second packet and so on is given by

$$P(j_1, j_2, ..., j_a) = \frac{n!}{a^n j_1! j_2! ... j_a!}.$$

Then, a possible interlacement of those cards is chosen such that each interlacement has equal probability of being chosen. A *riffle shuffle* or a *GSR-shuffle* is a 2-shuffle.

A mathematically equivalent definition of the interlacement step approaches in a more familiar way. Take the first two packets and imagine holding one in the left hand and the other in the right hand. Cards are dropped from the left hand with probability $\frac{a}{a+b}$ and from the right hand with probability $\frac{b}{a+b}$ where a and b are the number of cards remaining in hand. Shuffle this combined packet in the same manner with the third packet, then with the fourth packet and so on.

Theorem 2.3. *An a-shuffle followed by a b-shuffle is equivalent to an ab-shuffle.*

Proof. We can also think of an a-shuffle geometrically as dropping n points at random onto the unit interval $[0, 1]$ and then mapping $x \to ax \pmod 1$. Consider the intervals $[0, \frac{1}{a}]$, $[\frac{1}{a}, \frac{2}{a}]$, ..., $[\frac{a-1}{a}, 1]$. The number of points in each interval is distributed according to the multinomial distribution. An a-shuffle takes the points in each of those intervals and maps them to [0,1], so the points become interlaced with each other. The proof of the theorem is trivial because we have the following identity.

$$b(ax (\text{mod } 1)) (\text{mod } 1) = abx (\text{mod } 1).$$

□

Definitions 2.4. Let π be a permutation in S_n. A *rising sequence* is a maximal subset consisting of successive face values displayed in order.

Example 2.5. One way of finding all the rising sequences is to proceed from left to right, searching for the numbers in order, starting with the lowest one. Once the end of the deck is reached, one rising sequence has been completed. Remove those numbers and repeat. For example, the permutation 1, 3, 6, 7, 2, 4, 5, 8 has three rising sequences, (1, 2), (3, 4, 5) and (6, 7, 8).

Definitions 2.6. The probability that a distribution of the deck $\pi \in S_n$ is the result of an a-shuffle of n cards is denoted $Q_a(\pi)$. The probability that π is the result of k riffle shuffles of n cards is denoted $R^k(\pi)$.

Theorem 2.7.

$$\textit{We have } Q_a(\pi) = \frac{\binom{a+n-r}{n}}{a^n}, \textit{ where } r \textit{ is the number of rising sequences in } \pi.$$

Proof. The goal is to count to the number of ways to cut the deck into a packets to produce those r rising sequences. Each of the r rising sequence must be formed from at least one packet since shuffling preserves the order of the cards within a packet. At least one cut must be placed in between each rising sequence, accounting for r cuts. There are $a-r$ cuts remaining to be placed. A deck of n cards create $n+1$ slots where a cut may be placed since there are $n-1$ slots between two neighboring cards in addition to the top and bottom. Using a stars and bars argument, there are $\binom{a+n-(d+1)}{n}$ ways to choose where to place the cuts. The denominator comes from the total number of a-shuffles, a^n. □

Corollary 2.8.

$$\textit{We have } R^k(\pi) = \frac{\binom{2^k+n-r}{n}}{2^{nk}}, \textit{ where } r \textit{ is the number of rising sequences in } \pi.$$

Proof. By Theorem 2.3, k riffle shuffles is equivalent to a 2^k shuffle. Then, the corollary follows from plugging into Theorem 2.7. □

Now that we have finished formalizing the model of the shuffle, we will now consider the question of how to measure the difference in randomness between two shuffles.

Definitions 2.9. Let Q be a probability distribution on S_n such that for all $\pi \in S_n$, the probability π is obtained is denoted by $Q(\pi)$. One probability distribution of importance is the *uniform probability distribution*, which is denoted by U. There are $n!$ possible permutations of n cards and each of these permutations occurs with equal probability so we have that for all π in S_n, $U(\pi) = \frac{1}{n!}$. Let Q_1 and Q_2 be probability distributions on S_n. The *variation distance* on A $\subseteq S_n$ is denoted as $\|Q_1(A) - Q_2(A)\|$.

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

The *total variation distance* is denoted as $\|Q_1 - Q_2\|$.

$$\|Q_1 - Q_2\| = \max_{A \in S_n} \|Q_1(A) - Q_2(A)\|.$$

Proposition 2.10. *We have* $0 \le \|Q_1 - Q_2\| \le 1$ *for all* Q_1, Q_2.

Proof. By the Triangle Inequality,

$$\frac{1}{2}\sum_{\pi \in A} |Q_1(\pi) - Q_2(\pi)| \le \frac{1}{2}\sum_{\pi \in A} |Q_1(\pi)| + |Q_2(\pi)| \le \frac{1}{2} \times 2 = 1.$$

□

Example 2.11. Often, inexperienced shufflers will reveal the bottom card of the deck while shuffling. We will consider the effect this has on the total variation distance from the uniform probability distribution U. Let Q be a shuffle such that 1 card is left in its original place and the other $n-1$ cards are distributed uniformly. This leads to $(n-1)!$ possible distributions of the deck, each with probability $\frac{1}{(n-1)!}$.

Since one card is fixed, there are also going to be distributions of the deck that are not possible to obtain. Since there are $n!$ total distributions, the shuffle Q results in $n! - (n-1)!$ distributions that have probability 0.

$$\|Q - U\| = \frac{1}{2}\left[(n-1)!\left(\frac{1}{(n-1)!} - \frac{1}{n!}\right) + (n! - (n-1)!)\left(\frac{1}{n!} - 0\right)\right] =$$

$$\frac{1}{2}\left[1 - \frac{1}{n} + 1 - \frac{1}{n}\right] = 1 - \frac{1}{n}.$$

This example shows that the total variation distance is a very unforgiving measure of randomness since fixing just one card leads to a value very close to 1. In general, a value below 0.5 is a strong result.

We are interested in investigating the behavior of $\|R^k - U\|$, which measures the distance between the distribution created by k riffle shuffles and the uniform distribution. Directly plugging the results of Corollary 2.8 into the definition, and adding over all possible distributions of the deck gives

$$\|R^k - U\| = \frac{1}{2}\sum_{j=1}^{n} A_{n,r}\left|\frac{\binom{n+2^k-j}{n}}{2^{kn}} - \frac{1}{n!}\right|$$

where $A_{n,r}$ is the number of permutations with j rising sequences. These are called the Eulerian numbers and they can be found using the following recursive formula.

$$A_{n,1} = 1.$$

$$A_{n,r} = r^n - \sum_{j=1}^{r-1}\binom{n+r-j}{n}A_{n,j}.$$

This is sufficient to solve the question at hand, but this is different from the approach used by P. Diaconis and D. Bayer, which directly tackles the problem. I will also present their proof.

Theorem 2.12. *Let R^k be the probability distribution obtained by riffle shuffling k times. Let $k = \frac{3}{2}\log_2 n + \theta$.*

$$\|R^k - U\| = 1 - 2\Phi\left(\frac{-2^{-}\theta}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{\frac{1}{4}}}\right) \text{ with } \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{-t^2}{2}}\,dt.$$

Proof. Let π be a permutation with $r = \frac{n}{2} + h$ rising sequences after k riffle shuffles where $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$. Let $k = log_2(n^{\frac{3}{2}}c)$ where c is a positive constant. First, we will need to do some work to get $R^k(\pi)$ into a more useful form.

$$\text{By Corollary 2.8, } R^k(\pi) = \frac{\binom{2^k+n-r}{n}}{2^{nk}}.$$

$$\text{Expanding the numerator gives } R^k(\pi) = \frac{1}{n!}\left(\frac{2^k+n-r}{2^k} \times ... \times \frac{2^k+1-r}{2^k}\right).$$

$$R^k(\pi) = \frac{1}{n!}\left(\left(1 + \frac{n-r}{2^k}\right) \times ... \times \left(1 + \frac{1-r}{2^k}\right)\right).$$

Plugging in $k = log_2(n^{\frac{3}{2}}c)$ gives us $R^k(\pi) = \dfrac{1}{n!}\left(\left(1 + \dfrac{n-r}{cn^{\frac{3}{2}}}\right) \times ... \times \left(1 + \dfrac{1-r}{cn^{\frac{3}{2}}}\right)\right).$

Rewriting the product and plugging in $r = \frac{n}{2} + h$ gives us

$$R^k(\pi) = \frac{1}{n!} \prod_{i=0}^{n-1} \left(1 + \frac{\frac{n}{2} - h - i}{cn^{\frac{3}{2}}}\right).$$

It is simpler to deal with a summation than with a product so we make use of the identity $a \times b = e^{\log a + \log b}$

$$R^k(\pi) = \frac{1}{n!} \exp\left(\sum_{i=0}^{n-1} \log\left(1 + \frac{\frac{n}{2} - h - i}{cn^{\frac{3}{2}}}\right)\right).$$

Using a Taylor expansion, we can use the bound $x - \frac{x^2}{2} + \frac{x^3}{3} - x^4 \leq \log(1+x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}$ for $-\frac{1}{2}<x<1$. Standard summation formulas give

$$\frac{1}{cn^{\frac{3}{2}}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right) = \frac{-h + \frac{1}{2}}{c\sqrt{n}}.$$

$$\frac{1}{2c^2n^3} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^2 = \frac{1}{24c^2} + \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right).$$

$$\frac{1}{3c^3n^{\frac{9}{2}}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^3 = O_c\left(\frac{h}{n^{\frac{3}{2}}}\right).$$

$$\frac{1}{c^4n^6} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h - i\right)^4 = O_c\left(\frac{1}{n}\right).$$

Combining gives $R^k(\pi) = \frac{1}{n!} \exp\left(\frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + O_c\left(\frac{h}{n}\right)\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right).$

Since our goal is to compare the probability distribution created by k riffle shuffles with the uniform distribution, we are only concerned with probabilities greater than $\frac{1}{n!}$ so we solve for the value of h' such that $R^k(\frac{n}{2} + h) \geq \frac{1}{n!} \Leftrightarrow h \leq h'$. The left side is true when the exponential term is greater than or equal to one, so we solve the equation

$$\frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + O_c\left(\frac{h}{n}\right)\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right) = 0.$$

This gives us $h' = \frac{-\sqrt{n}}{24c} + \frac{1}{12c^3} + B + O_c\left(\frac{1}{\sqrt{n}}\right)$, where $-1 \leq B \leq 1$.

To simplify the problem, we partition the relevant values of h into two zones, I_1 and I_2.

$$I_1 = \left\{\frac{-10n^{\frac{3}{4}}}{\sqrt{c}} \leq h \leq h'\right\}.$$

$$I_2 = \left\{\frac{-n}{2} \leq h < \frac{-10n^{\frac{3}{4}}}{\sqrt{c}}\right\}.$$

The problem we are trying to solve is the equation

$$\|R^k - U\| = \sum_{h=\frac{-n}{2}}^{h'} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right) - \frac{1}{n!}\right).$$

Partitioning the summation gives us:

$$\|R^k - U\| = \sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) + \sum_{I_2} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) - \frac{1}{n!}\sum_{\frac{-n}{2}<h\leq h*} A_{n,h}.$$

We will address each of these terms individually. For the first term, we will need to use the following results, which were proved by S. Tanny (1973), using the central limit theorem. Let $x_n = \frac{h}{\sqrt{\frac{n}{12}}}$.

$$\frac{A_{n,h}}{n!} = \frac{e^{\frac{-1}{2}x_n^2}}{\sqrt{\frac{2n\pi}{12}}}\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right).$$

$$\sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) = \frac{e^{\frac{-1}{24c^2}}}{\sqrt{\frac{2n\pi}{12}}}\sum_{I_1}\exp\left(\frac{-1}{2}\left(\frac{h}{\sqrt{\frac{n}{12}}}\right)^2 - \frac{h}{c\sqrt{n}} + O_c\left(\frac{1}{n^{\frac{1}{4}}}\right)\right)\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right).$$

$$\sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) = \frac{e^{\frac{-1}{24c^2}}}{\sqrt{2\pi}}\int_{-\infty}^{\frac{-1}{2c\sqrt{12}}} e^{\frac{-x^2c\sqrt{3}}{2-x}}\,dx\left(1 + O\left(\frac{1}{n^{\frac{1}{4}}}\right)\right).$$

$$\sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) = \Phi\left(\frac{1}{4c\sqrt{3}}\right)\left(1 + O\left(\frac{1}{n^{\frac{1}{4}}}\right)\right).$$

$$\sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) = \Phi\left(\frac{1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{n^{\frac{1}{4}}}\right).$$

Since $\Phi(x) = 1 - \Phi(-x), \sum_{I_1} A_{n,h}\left(R^k\left(\frac{n}{2}+h\right)\right) = 1 - \Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{n^{\frac{1}{4}}}\right).$

In $I_2, R^k\left(\frac{n}{2}+h\right) \leq R^k(1) \leq \frac{e^{\frac{\sqrt{n}}{2c}}}{n!}$. W. Feller (1951) proved the following result.

$$\sum_{I_2}\frac{A_{n,h}}{n!} \sim \frac{1}{10n^{\frac{1}{4}}\sqrt{2\pi}}\exp\left(\frac{-1}{2}\left(\frac{10\sqrt{12}n^{\frac{1}{4}}}{\sqrt{c}}\right)^2\right).$$

$$\sum_{I_2}\frac{A_{n,h}}{n!} = O\left(\frac{1}{n^{\frac{1}{4}}}\right).$$

To simplify the third term, we apply another result found in S. Tanny (1973).

$$\frac{1}{n!}\sum_{h=\frac{-n}{2}}^{h'} A_{n,h} = \Phi\left(\frac{-1}{4c\sqrt{3}}\right)\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

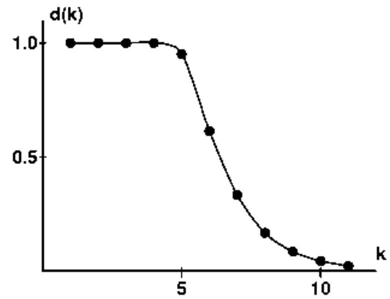
$$\frac{1}{n!}\sum_{h=\frac{-n}{2}}^{h'} A_{n,h} = \Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right).$$

Combining the three terms gives the desired result.

$$\|R^k - U\| = 1 - 2\Phi\left(\frac{-2^{-\theta}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{\frac{1}{4}}}\right).$$

□

This formula gives the following table of values and plot for $\|R^k - U\|$.



k	$\|R^k - U\|$
1	1
2	1
3	1
4	1
5	0.924
6	0.614
7	0.334
8	0.167

Recall that the number of shuffles $k = \frac{3}{2}log_2n + \theta$. As $\theta \to -\infty$, the total variation distance goes doubly exponentially fast to 1 and as $\theta \to \infty$, the total variation distance goes exponentially fast to 0. This makes k $= \frac{3}{2}\log_2 n$ a good general rule for the number of shuffles required to randomize a deck.

3. Exploring Different Notions of Randomness

The results of P. Diaconis and D. Bayer's paper reached the front page of the Science section of the New York Times, which proclaimed that "It takes just seven ordinary, imperfect shuffles to mix a deck of cards throughly." This statement alone does not tell the whole story. Obviously, seven riffle shuffles do not give a uniform probability distribution, but rather something similar in probabilities to one. We will attempt to examine the randomization process a bit more closely.

First, we examine our choice of total variation distance as the measure of randomness. It is, in some senses, intuitive because it adds over all possible distributions of the deck and it is quite convenient for mathematical purposes. Of course, we can consider other measures. We will mention just one, called separation distance, which measures the maximum distance between the two distributions.

Definition 3.1. The *separation distance* is denoted $s(k)$ and defined as

$$s(k) = \max_{\pi \in S_n} 1 - \frac{R^k(\pi)}{U(\pi)}.$$

Proposition 3.2. *It requires $2 \log_2 n$ shuffles to make the separation distance small.*

The proof of this statement can be found in D. Aldous and P. Diaconis's "Shuffling Cards and Stopping Times" (1986). It is not important to the purposes of this paper to offer the proof. Rather, we only note that changing how we measure randomness can potentially lead to a different shuffling rule.

The following two games use rising sequences to take advantage of how the GSR-shuffle is defined. They expose a lack of true randomness created by seven riffle shuffles.

1. **Premo**. P. Diaconis and D. Bayer accredit a magic trick, discovered by magicians Williams and Jordan, as the inspiration for the methodology presented in Section 2. In the trick, the magician uses a brand new deck of cards and gives it to a member of the audience and asks them to give it a cut followed by a riffle shuffle three times. Afterwards, the volunteer looks at the top card and places it somewhere inside the deck and returns the deck to the magician. The magician looks through the cards and is able to discern which card had been at the top.

Let us explain how the trick works. Three riffle shuffles are expected to generate 8 rising sequences, each of length about 6.5 cards. Adding a card from the top of the deck to the middle usually results in an additional rising sequence of 1 card which, to a trained eye, stands out. The cuts do not change the order of the cards if the deck is viewed as a cycle.

A computer program was written, which selected the most likely card to be at the top after a number of riffle shuffles. The results for the probability of success are displayed in the following table, where k is the number of riffle shuffles and m the number of guesses allowed to the computer.

m\k	2	3	4	5	6	7	8	9	10	∞
1	99.7	83.9	28.8	8.8	4.2	2.8	2.3	2.1	2.0	1.9
2	100	94.3	47.1	16.8	8.3	5.7	4.7	4.2	4.0	3.8
3	100	96.5	59.0	23.8	12.3	8.5	7.0	6.3	6.1	5.8
13	100	99.8	88.4	61.7	42.7	33.4	29.0	27.0	26.0	25.0
26	100	99.9	97.5	83.5	68.8	59.6	54.8	52.4	51.3	50.0

Given 26 guesses for a deck shuffled 8 times, the program succeeded 54.8% of the time. If the deck were truly random, then the success rate should be 50%.

2. **New Age Solitaire**. Go through a deck one card at a time, placing each card into the Yin pile, the Yang pile or a discard pile. The Yin pile contains, in order, the King through Ace of Hearts followed by the King through Ace of Clubs. The Yang pile consists of the Ace through King of Spades followed by the Ace through King of Diamonds. Once you have gone through the deck, turn over the discard pile and reiterate. You win if you finish the Yin pile prior to finishing the Yang pile.

In a brand new deck of cards, typically the order is the King through Ace of Hearts, followed by the King through Ace of Clubs, followed by the King through Ace of Diamonds, followed by the King through Ace of Spades. In other words, if we numbered the cards $1, 2, ..., 52$, the Yin pile corresponds to the ordering $1, 2, ..., 26$ and the Yang pile corresponds to the ordering $52, 51, ..., 27$. Of course, with a new deck, you always win since it requires just one go-through to finish the Yin pile and 26 go-throughs to finish the Yang pile since each go-through will put only one card on the Yang pile.

The intuition of the game is that rising sequences take cards from the top and bottom halves in roughly the same lengths. A rising sequence of Yin cards requires one go-through to be placed on the Yin pile but a rising sequence of Yang cards requires go-throughs equal to the length of the sequence to be placed on the Yang pile. It is demonstrated in A. van Zuylen and F. Schalekamp's "The Achilles Heel of the GSR-Shuffle" (2004) that the win probability for this game is

$$P(win) = \frac{1}{2} + \|R^k - U\|.$$

Note that this result assumes that we begin with a brand new deck. This result is surprising because it had generally been thought that a variation distance below 0.5 was already quite good. Seven riffle shuffles gives a win probability of 81%, far greater than the 50% for a truly random deck.

4. Decks with Repeated Cards

We will use common card games to motivate the question of randomizing a deck that contains repeated cards. For example, in Blackjack, the suits of the cards are irrelevant, so we can consider the deck in terms of only the card values.

Definitions 4.1. Label the deck $1^{n_1}, 2^{n_2}, ..., h^{n_h}$ where there are n_1 copies of 1, n_2 copies of 2 and so on. The total number of cards is $n_1 + n_2 + ... + n_h = n$. Given two orderings of the deck D_1, D_2, we define $\Pi(D_1, D_2)$ to be the set of all permutations in S_n that result in D_2 when applied to D_1. We define c_r be the number of permutations in $\Pi(D_1, D_2)$ with r rising sequences.

Example 4.2. If we have a deck with 4 cards that consists of 2 copies of a card labelled 1 and 2 copies of a card labelled 2, and let D_1 = 1,1,2,2. D_2 = 1,2,2,1. We want to consider all permutations that, when applied to D_1 result in D_2. This gives us the set $\Pi(D_1, D_2)$ = 1,4,3,2 or 1,4,2,3 or 4,1,2,3 or 4,1,3,2. There are 2, 1, 1 and 2 descents in each of those permutations, respectively. This gives us that $c_1 = 2$ and $c_2 = 2$.

Definition 4.3. The probability a riffle shuffle applied to D_1 results in D_2 is denoted $R(D_1, D_2)$.

Theorem 4.4.

$$\textit{We have } R(D_1, D_2) = \sum_{d=0}^{n-1} \frac{c_d \binom{2+n-r}{n}}{2^n}.$$

Proof. With a repeated deck, there can be multiple permutations that will yield D_2 when applied to D_1. Therefore, we add over all the elements of $\Pi(D_1, D_2)$. Theorem 2.7 gives us that the probability for one such element is $\frac{\binom{2+n-r}{n}}{2^n}$ where r is the number of rising sequences of that permutation. We have to account for the differences in the rising sequences of these permutations, so we multiply by the factor c_d and add over all values of d. □

The repeated card problem is reduced to finding the values of c_d given two distributions of the deck D_1 and D_2. This turns out to fall under the class of #P-complete problems. For a more complete treatment of this result, refer to M. Conger and D. Viswanath (2006). They also found an algorithm that solves for

c_d that is $O(n^{2d})$, but this is too computationally inefficient for large decks. One solution is to use a normal approximation for c_d which provides fairly good results.

To motivate the results of this repeated card model, we will use the following four games. D_1 refers to the initial state of the deck in each game.

Definitions 4.5.
1. **Blackjack.** This refers to games where suit is not considered. The deck we will work with is labelled $D_1 = 1^4, 2^4, ..., 13^4$.

2. **Red/Black.** This refers to games where only color is considered.
$D_1 = R^26, B^26$.

3. **Bridge.** This refers to games with four players, where each player receives a subset of the deck consisting of 13 cards. The goal is to randomize amongst the 13 card hands. $D_1 = N^{13}, E^{13}, S^{13}, W^{13}$. (Bridge players are labelled by the cardinal directions.)

4. **Alice/Bob.** This is the same as Bridge, but for two players. $D_1 = A^{26}, B^{26}$.

The following are the results for $\|R^k - U\|$ for each game, found by M. Conger and D. Viswanath in "Shuffling Cards for Blackjack, Bridge, and Other Card Games" (2006). The data from Section 2 is included for comparison.

	1	2	3	4	5	6	7	8
Normal Deck	1	1	1	1	0.924	0.614	0.334	0.167
Blackjack	1	1	1	0.481	0.215	0.105	0.052	0.026
Red/Black	0.580	0.360	0.208	0.105	0.052	0.026	0.013	0.07
Bridge	1	1	1	0.990	0.748	0.423	0.218	0.110
Alice/Bob	1	1	0.999	0.725	0.308	0.130	0.059	0.028

The results indicate savings of 3 shuffles if we are only interested in card value, 6 shuffles if we are only interested in color, 1 shuffle if we want to randomize the hands given to four players and 2 shuffles if we want to randomize the hands given to two players.

Acknowledgments. I would like to thank my mentor, Olga Turanova, for her guidance in writing this paper. I would also like to thank Peter May and the University of Chicago Math Department for running the REU program. Finally, I would like to acknowledge Professor Antonio Auffinger, whose lectures on card shuffling inspired my research.

References

[1] Anke van Zuylen and Frans Schalekamp. (2004) The Achilles Heel of the GSR-Shuffle, *Probability in the Engineering and Informational Sciences* **18**, 315-328.
[2] Brad Mann. (1995) How Many Times Should You Shuffle a Deck of Cards?, *Topics in Contemporary Probability and Its Applications* **15**, 1-33.
[3] D. Viswanath and Mark Conger. (2006)Shuffling Cards for Blackjack, Bridge, and Other Card Games. http://arxiv.org/pdf/math/0606031v1.pdf.
[4] Dave Bayer and Persi Diaconis. (1992) Trailing the Dovetail Shuffle to its Lair, *The Annals of Applied Probability* **2**, 294-313.
[5] David Aldous and Persi Diaconis. (1986) Shuffling cards and stopping times, *American Mathematical Monthly* **93**, 333-348.

[6] David Austin. How Many Times Do I Have to Shuffle This Deck? http://www.ams.org/samplings/feature-column/fcarc-shuffle.

[7] Jason Fulman (2002) Applications of Symmetric Functions to Cycle and Increasing Subsequence Structure after Shuffles, *Journal of Algebraic Combinatorics* **16**, 165-194. Jim Pitman, Michael McGrath and Persi Diaconis (1995) *Combinatorica* **15**, 11-29.

[8] Persi Diaconis (2003) Mathematical Developments from the Analysis of Riffle Shuffling, *Groups, Combinatorics and Geometry edited by Ivanov, Liebeck and Saxl. World Scientific*, 73-97.

[9] S. Tanny (1973) A Probabilistic Interpretation of Eulerian Numbers, *Duke Math Journal* **40**, 717-722.

[10] Steven P. Lalley (2000) On the Rate of Mixing for p-Shuffles, *The Annals of Applied Probability* **10**, 1302-1321.

[11] William Feller (1951) An Introduction to Probability and Its Application **2**, 2nd ed. Wiley, New York.