Notes on Unravelling the Shuffle

Anav Sood

November 2019

1 Motivation and Groundwork

1.1 Introduction

These notes originally started out as a final project for Professor Persi Diaconis' class Mathematics and Statistics of Gambling. Much of Persi's well known work pertains to quantifying how the randomness of shuffling a deck of cards. It is useful to simply thing of a deck of N cards prior to shuffling as an ordered list $(1, \ldots, N)$, where ith card is represented value of the ith index. In this case the ith card is represented by the value i. Shuffling a deck of cards is simply the act of permuting the list. After shuffling, the value i will represent the $\sigma(i)$th card. A particular method of shuffling will induce a particular probability distribution on the order of the deck of cards post-shuffle. We want to get an idea of how close to uniformly random this probability distribution is. Note, we always consider $N \geq 2$.

There are multiple ways to measure the randomness of a shuffle of a deck of N cards, one common method being to measure the total variation distance between the measure that is uniform over all $N!$ permutations and the measure that is induced by shuffling. In this document, we will attempt to gauge the randomness of a shuffle by playing a guessing game. We attempt to guess the order of the cards after a single shuffle supposing that we knew their original order and the (possibly stochastic) manner in which they were shuffled. In the game, we guess the top card of the shuffled deck, and then the card is revealed to us. We want to try and maximize the number of cards we guess properly on expectation. Letting S represent the number of cards we guess correctly and I_i be the indicator that we guess the ith card correctly, we see that for an N card deck

$$\mathbf{E}[S] = \mathbf{E}[\sum_{i=1}^{N} I_i] = \sum_{i=1}^{N} \mathbf{E}[I_i] = \sum_{i=1}^{n} \mathbf{P}(I_i = 1)$$

We will call the guessing strategy that maximizes the above value **optimal**.

Example 1.1. Suppose that a deck of N cards is truly randomly shuffled, so each permutation of the cards is equally likely. Then, post shuffle, the best guess we can have for the ith card is simply any card that hasn't been revealed. We'll get the 1st card right with probability $\frac{1}{N}$, the second with $\frac{1}{N-1}$ etc. Thus the best we can do in this case is

$$\mathbf{E}[S] = \sum_{i=1}^{N} \frac{1}{i}$$

For a 52 card deck $\mathbf{E}[S] \approx 4.538$. Note that for large n, $\mathbf{E}[S] \sim \log n$

Example 1.2. Suppose we shuffle a deck of N cards simply by reversing the order of the deck. Then we can guess the order of the deck in reverse so $\mathbf{E}[S]$.

In class, Persi defined the overhead shuffle and posed the question of finding the optimal guessing strategy for a deck of cards after one overhead shuffle (without knowing himself). We find the provably optimal strategy and give a closed form for the exected number of cards guessed correctly in our first section. He also asked what the optimal guessing strategy is for undoing k riffle shuffles. The problem of proving that McGrath's guessing strategy is optimal was provided as the eighth open problem in Diaconis' "Mathematical Developments from the Analysis of Riffle Shuffling". We show that McGrath's strategy is in fact optimal in the case that $k = 1$, and provide a brief comment regarding extending this result to $k > 1$.

1.2 Setting up a Framework

Let C_i represent the value of the ith card after a single shuffle (so C_1 will be the value of the first card we attempt to guess), and define $C^{(i)} := (C_1, \ldots, C_i)$. We will let $\{C^{(i)}\} := \{C_1, \ldots, C_i\}$ represent the set of values. The following Theorem will guide our future work.

Theorem 1.3. The guessing strategy that guesses $g_i \in \arg\max_j \mathbf{P}(C_i = j|C^{(i-1)})$ for the ith card is optimal. Any strategy that guesses possibly random g_i' such that the conditional support of g_i' given $C^{(j-1)}$ doesn't have all support in $\arg\max_j \mathbf{P}(C_i = j|C^{(i-1)})$ is strictly sub-optimal

Proof. We first note that under some guessing strategy that guesses g_i for the ith card.

$$\begin{aligned}
\mathbf{E}[S] &= \sum_{i=1}^{N} \mathbf{E}[I_i] && (1)\\
&= \sum_{i=1}^{N} \mathbf{E}[\mathbf{E}[I_i|C^{(i-1)}]] && (2)\\
&= \sum_{i=1}^{N} \mathbf{E}[\mathbf{P}(I_i = 1|C^{(i-1)})] && (3)\\
&= \sum_{i=1}^{N} \mathbf{E}[\mathbf{P}(C_i = g_i|C^{(i-1)})] && (4)
\end{aligned}$$

Then consider the strategy that guesses g_i given above and some other strategy that guesses (possibly stoachstic) g_i'. Then for every i

$$\begin{aligned}
\mathbf{P}(C_i = g_i'|C^{(i-1)}) &= \sum_{k=1}^{N} \mathbf{P}(C_i = g_i', g_i' = k|C^{(i-1)})]\\
&= \sum_{k=1}^{N} \mathbf{P}(C_i = g_i'|g_i' = k, C^{(i-1)})\mathbf{P}(g_i' = k|C^{(i-1)})\\
&= \sum_{k=1}^{N} \mathbf{P}(C_i = k|C^{(i-1)})\mathbf{P}(g_i' = k|C^{(i-1)})\\
&\leq \mathbf{P}(C_i = g_i|C^{(i-1)})
\end{aligned}$$

where last inequality follows as $\sum_{k=1}^{N} \mathbf{P}(C_i = k|C^{(i-1)})\mathbf{P}(g_i' = k|C^{(i-1)})$ is a convex combination of values all $\leq \mathbf{P}(C_i = g_i|C^{(i-1)})$. Note that if the convex combination puts positive weight on any value strictly smaller than $\mathbf{P}(C_i = g_i|C^{(i-1)})$ then the above inequality will be strict. Using this we see that

$$\sum_i \mathbf{E}[\mathbf{P}(C_i = g_i|C^{(i-1)})] \geq \sum_i \mathbf{E}[\mathbf{P}(C_i = g_i'|C^{(i-1)})]$$

where the inequality is strict exactly in the situation stated by the theorem. □

Essentially, this theorem tells us exactly that the intuitive result that the strategy that maximizes the probability that the ith guess is correct given the values of the previously revealed $i-1$ cards is optimal.

2 The Overhead Shuffle

2.1 Definition

The overhead shuffle is a common, simple shuffling technique involving grabbing a stack of cards from the top of the deck and placing the stack, in order, at the bottom. As one can imagine, since the stacks are left in order, the shuffle is not an ideal randomizing agent. We can define the shuffle as follows.

Definition 2.1. Suppose we have a N card deck where we label the bottom card as the 1st card and the top card as the Nth card. Consider $X_1, \ldots, X_{N-1}$ to be $N-1$ i.i.d Bernoulli random variables with parameter $\theta \in [0,1]$, so $\mathbf{P}(X_i = 1) = \theta$ and $\mathbf{P}(X_i = 0) = 1-\theta$. Let $P = \sum_{i=1}^{N-1} X_i$. Suppose $i_1 > \cdots > i_P$ are indices such that $X_{i_k} = 1$. We take all the cards with index $> i_1$ and make them our first (lowest) stack, we take all the cards with index $> i_2$ and make it our second stack, so on and so forth. The remaining cards are made into our last stack.

Experiments shows that a reasonable θ value which mimics human performance is 0.2. Note if $\theta = 1$ the deck is simply put in reverse order and if $\theta = 0$ no changes occur. These edge cases are deterministic, so for our analyses we will assume $\theta \in (0,1)$. We provide an example of a overhead shuffle with a 10-card deck below.

10	$X_9 = 0$		2	———
9	$X_8 = 1$		1	———
8	$X_7 = 0$		5	———
7	$X_6 = 0$		4	———
6	$X_5 = 1$	$\rightarrow$	3	———
5	$X_4 = 0$		8	———
4	$X_3 = 0$		7	———
3	$X_2 = 1$		6	———
2	$X_1 = 0$		10	———
1			9	———

2.2 Developing the Optimal Guessing Strategy

To develop the optimal guessing strategy, we first focus on guessing the top card. It is easy to see that

$$\mathbf{P}(C_1 = k) = \begin{cases} \theta & \text{if } k = 1 \\ \theta(1-\theta)^{k-1} & \text{if } 1 < k < N \\ (1-\theta)^{N-1} & \text{if } k = N \end{cases}$$

We note that then by Theorem 1.3, our guess should be $g = \arg\max_g \mathbf{P}(C_1 = g)$. Since $\theta > \theta(1-\theta)^k$ for any k, it only ever makes sense to guess the top card N or the bottom card 1. Should we have $\theta \geq (1-\theta)^{N-1}$ then we pick 1, otherwise we pick N. This gives the optimal decision for the top card. Note that in this situation, we understand that we're going to be guessing the top card of a some stack, and we're trying to determine what the most likely such top card is.

Prior to continuing, we note that the comparison between θ and $(1-\theta)^n$ for integer values of n will come up quite frequently. We would like to know for which integer values of n it is the case that $\theta \geq (1-\theta)^n$. It should be clear that this is the case for $n \geq \lceil \frac{ln(\theta)}{ln(1-\theta)} \rceil := r(\theta)$ and not for $n < r(\theta)$. Note for $\theta \in (0,1)$, $r(\theta) \in \{1, 2, 3 \ldots, \}$

Now, suppose our first guess was incorrect and some card with value $i > 1$ is revealed. Then we know i was the top card of a stack of size more than one. Thus, the next card must be $i-1$ (the next card in the stack) with probability 1. In fact, we can continue guessing $i-2$, $i-3$, $\ldots$, 1 with certainty until we deplete the cards in that stack. In the case that $i = 1$, then we know the next card will be the top of another stack, and we repeat a reasoning process similar to when we guessed the first card. We can write this result more generally for guessing the jth card having seen the values of $C_1, \ldots, C_{j-1}$.

Strategy 2.2. Suppose we are guessing the value of the jst card. Let $m := \max(\{C^{(j-1)}\})$ and notationally take $[n] = \{1, \ldots, n\}$. Note that if you have seen $j-1$ cards then the largest value must be at least $j-1$, so $m \geq j-1$. We have two cases

- Case One: Suppose that $m > j-1$. Then guess $\max([m] \setminus \{C^{(j-1)}\})$.
- Case Two: Suppose that $m = j-1$. Then guess j if $N - j \geq r(\theta)$ and N otherwise.

We claim that this strategy is optimal

Theorem 2.3. Strategy 2.2 is optimal for a single overhead shuffle.

Proof. Consider the conditional distribution $\mathbf{P}(C_j | C^{(j-1)})$ for any j.If the $C^{(j-1)}$ such that Case One of Strategy 2.2, then we are in a situation where we know C_j must be the next card in a known stack, so it will take the same value as our guess with probability 1. If the $C^{(j-1)}$ are such that Case Two of Strategy 2.2 applies and $j < N$, then the conditional distribution of C_j is given by.

$$\mathbf{P}(C_j = k | C^{(j-1)}) = \begin{cases} \theta & \text{if } k = j \\ \theta(1-\theta)^{k-j} & \text{if } j < k < N \\ (1-\theta)^{N-j} & \text{if } k = N \end{cases}$$

From this it is clear that in this case C_j carries most of its support on j if $N - j \geq r(\theta)$ and N otherwise. In the case that we are in Case Two and $j = N$ then C_j carries all it's weight on N. Thus it is clear that our strategy is optimal by Theorem 1.3. □

2.3 Probabilistic Analysis

Note that in our strategy, we are effectively guaranteed to get any card that isn't the top of a stack correct, and we are mainly concerned with whether or not we correctly guess the tops of stacks. Suppose we let $P = \sum_{i=1}^{N-1} X_i$ as in Definition 2.1. Then $P+1$ is the total number of stacks in the shuffled deck. If we let T denote the number of tops of stacks we guess correctly, then we see that

$$S = N - (P+1) + T$$

All that remains is to find an expression for T. Suppose we're executing our strategy and at the jth step we're guessing the top of a stack. There are two cases.

- Case One: If $j \leq N - r(\theta)$ then we will guess j to be the top of the next stack. If $j = 1$ here (we are guessing the first card), we will be correct in guessing the top card of the stack if $X_1 = 1$. Otherwise we are in a situation where X_{j-1} is the top of a stack, so we will be correct if also $X_j = 1$ and wrong otherwise.
- Case Two: If $j > N - r(\theta)$ then we will guess N. In this case we can be correct exactly one time, and we will so long as the top card hasn't been revealed as the top of a stack by this point (the top card is garaunteed to be the top of a stack). If $N - r(\theta) < 1$ then we're guaranteed to guess the top card when it appears as we will guess it for the top of every stack. Otherwise, so long as it is not the case that $X_i = 0$ for all $i \in \{N - r(\theta), \ldots, N-1\}$ then the top card will not be revealed by this point, and we will guess it when it appears.

With this, we can write an expression for S with some casework

$$S = \begin{cases} N - (1 + \sum_{i=1}^{N-1} X_i) + 1 & N - r(\theta) < 1 \\ N - (1 + \sum_{i=1}^{N-1} X_i) + 1 - \prod_{i=1}^{N-1}(1 - X_i) + X_1 & N - r(\theta) = 1 \\ N - (1 + \sum_{i=1}^{N-1} X_i) + 1 - \prod_{i=N-r(\theta)}^{N-1}(1 - X_i) + X_1 + \sum_{i=2}^{N-r(\theta)} X_i X_{i-1} & N - r(\theta) > 1 \end{cases}$$

Defining $\tau(N, \theta) := N - r(\theta)$ to represent the threshold after which we guess the top card to start the next stack rather than the next card in line, we see that

$$S = \begin{cases} N - \sum_{i=1}^{N-1} X_i & \tau(N,\theta) < 1 \\ N - \sum_{i=2}^{N-1} X_i - \prod_{i=1}^{N-1}(1 - X_i) & \tau(N,\theta) = 1 \\ N - \sum_{i=2}^{N-1} X_i - \prod_{i=\tau(N,\theta)}^{N-1}(1 - X_i) + \sum_{i=2}^{\tau(N,\theta)} X_i X_{i-1} & \tau(N,\theta) > 1 \end{cases}$$

Computing the expectations of the above gives the following result

Theorem 2.4. Using Strategy 2.2, the expected number of cards we guess correctly is given by

$$\mathbf{E}[S] = \begin{cases} N - (N-1)\theta & \tau(N,\theta) < 1 \\ N - (N-2)\theta - (1-\theta)^{N-1} & \tau(N,\theta) = 1 \\ N - (N-2)\theta - (1-\theta)^{N-\tau(N,\theta)} + (\tau(N,\theta) - 1)\theta^2 & \tau(N,\theta) > 1 \end{cases}$$

3 The Riffle Shuffle

3.1 Definition

The riffle shuffle is likely the most common shuffling technique. This time we order the cards so that 1 is the top card and N is the bottom card. The shuffle is defined as follows.

Definition 3.1. Suppose we have a N card deck. We cut the deck beneath the ith card with probability

$$\mathbf{P}(K = i) = \frac{\binom{N}{i}}{2^N}$$

and $K = 0, N$ corresponds to not cutting the deck. We now have two stacks, one of size A and one of size B (where the size of a stack can be 0 if $K = 0, N$). We drop cards one at a time from each stack until we exhaust the deck. The probability of dropping from a stack is proportional to its size. In particular

$$\mathbf{P}(\text{drop from stack of size } A) = \frac{A}{A+B}$$

,

$$\mathbf{P}(\text{drop from stack of size } \ B) = \frac{B}{A+B}$$

3.2 Developing an Optimal Guessing Strategy

Note that one of the two piles always has the top card of the unshuffled deck as its top card. Suppose we are trying to guess the value of C_1.

Lemma 3.2. $\mathbf{P}(C_1 = 1) = \frac{1}{2} + \frac{1}{2^N}$ and $\mathbf{P}(C_1 = j) < \frac{1}{2}$ for $j > 1$

Proof. First we show $\mathbf{P}(C_1 = j) < \frac{1}{2}$ for $j > 1$. We always have

$$\begin{aligned}
\mathbf{P}(C_1 = j) &= \sum_{i=0}^{N} \mathbf{P}(C_1 = j, K = i) \\
&= \mathbf{P}(C_1 = j, K = j-1) \\
&= \mathbf{P}(C_1 = j | K = j-1)\mathbf{P}(K = j-1) \\
&< \mathbf{P}(K = j-1) \leq \frac{1}{2}
\end{aligned}$$

Now we consider $\mathbf{P}(C_1 = 1)$. Suppose N is even. Then for any $i \in \{1, \ldots, \frac{N}{2} - 1\}$ there exactly some $j \in \{\frac{N}{2} + 1, \ldots, N-1\}$ such that $K = i$ and $K = j$ are equal probability events where the cut is such that the two piles have non-empty size A and B, and the top card is in the bigger pile in one case, and smaller in the other. By symmetry then we know that if $\mathbf{P}(C_1 = 1 | K = i) = p$ then $\mathbf{P}(C_1 = 1 | K = i) = 1 - p$, so

$$\mathbf{P}(C_1 = 1, K = i) + \mathbf{P}(C_1 = 1, K = j) = p\mathbf{P}(K = i) + (1-p)\mathbf{P}(K = i) = \mathbf{P}(K = i)$$

Note also by symmetry that

$$\mathbf{P}(C_1 = 1, K = \frac{N}{2}) = \mathbf{P}(C_1 = 1 | K = \frac{N}{2})\mathbf{P}(K = \frac{N}{2}) = \frac{1}{2}\mathbf{P}(K = \frac{N}{2})$$

Noting then that K is binomial and symmetric about its mean of $\frac{N}{2}$, we see that

$$\begin{aligned}
\mathbf{P}(C_1 = 1) &= \sum_{i=0}^{N} \mathbf{P}(C_1 = 1, K = i) \\
&= \mathbf{P}(C_1 = 1, K = 0) + \sum_{i=1}^{\frac{N}{2}-1} \mathbf{P}(K = i) + \frac{1}{2}\mathbf{P}(K = \frac{N}{2})\mathbf{P}(C_1 = 1, K = N) \\
&= \sum_{i=0}^{\frac{N}{2}-1} \mathbf{P}(K = i) + \frac{1}{2}\mathbf{P}(K = \frac{N}{2}) + \mathbf{P}(K = N) \\
&= \frac{1}{2} + \frac{1}{2^N}
\end{aligned}$$

In the case N is odd, we similarly find

$$\begin{aligned}
\mathbf{P}(C_1 = 1) &= \sum_{i=0}^{N} \mathbf{P}(C_1 = 1, K = i) \\
&= \sum_{i=0}^{\lfloor \frac{N}{2} \rfloor} \mathbf{P}(K = i) + \mathbf{P}(K = N) \\
&= \frac{1}{2} + \frac{1}{2^N}
\end{aligned}$$

as desired. □

Now, let's go about creating a guessing strategy. Lemma 3.2 tells us the optimal strategy must guess the top card first. What do we do after? We investigate ths shuffle further.

Lemma 3.3. Consider a deck of cards of size N that undergoes one riffle shuffle. If it is known that the deck was cut with $K = i$, there are exactly $\binom{N}{i}$ orderings $c^{(N)}$ that have positive probability of occurring after one riffle shuffle.

Proof. Splitting the deck at $K = i$ leaves us with one pile of size i and another of size $N - i$. To characterize the valid (positive probability) orderings, we simply note that there are $\binom{N}{i}$ ways to choose the indices at which cards from the pile of size i appear in $c^{(N)}$. Because there is only one order in which these cards can appear relative to one another for such an outcome to be valid, there are exactly then $\binom{N}{i}$ positive probability orderings $c^{(N)}$. □

We prove a stronger result below.

Lemma 3.4. Consider a deck of cards of size N that undergoes on riffle shuffle. The conditional distribution of $C^{(N)}$ (the order of the deck post shuffle) given $K = i$ is uniform over its support and each individual ordering with positive probability has probability $\binom{N}{i}^{-1}$.

Proof. The result is obvious for $K = 0$ and $K = N$ as in these cases only one ordering has non-zero probability. Otherwise, suppose the $K = i$ so that the deck is divided into piles of positive size i

and $N-i$. Consider the case that $c^{(N)} = (1, \ldots, N)$, which simply corresponds to dropping all the cards from pile of size i, followed by all those from the pile of size $N-i$. This has probability

$$\mathbf{P}(C^{(N)} = c^{(N)} | K = i) \prod_{k=0}^{i-1} \frac{i-k}{N-k} = \binom{N}{i}^{-1}$$

Now consider some ${c^{(N)}}'$ such that we still finish dropping cards from the pile of size i first, but at some points we may drop cards from the pile of size $N-i$. Then the probability of this event will have a similar form to above, but will have additional terms in the product. Consider how these additional terms impact the numerator and denominator above. If we drop s cards from the pile of size $N-i$, we will see exactly a additional $(N-i)(N-i-1)(N-i-s+1)$ in the numerator, and the denominator will have s more terms of the form

$$\frac{1}{N-i}\frac{1}{N-i-1}\cdots\frac{1}{N-i-s+1}$$

These terms will cancel to yield the same result. In the case that we finish dropping cards from the deck of size $N-i$ first, the same proof shows that the probability comes out to $\binom{N}{N-i}^{-1} = \binom{N}{i}^{-1}$ and we are done. □

Corollary 3.5. If the deck is cut at $K = i$, then all $\binom{N}{i}$ orderings of the deck post-shuffle with positive probability have probability $\frac{1}{2^N}$.

Proof. Any valid ordering $c^{(N)}$ (one with positive probability jointly with $K = i$) after a split $K = i$ is such that $\mathbf{P}(C^{(N)} = c^{(N)} | K = i) = \binom{N}{i}^{-1}$ from Theorem 3.4. Thus

$$\mathbf{P}(C^{(N)} = c^{(N)}, K = i) = \mathbf{P}(C^{(N)} = c^{(n)} | K = i)\mathbf{P}(K = i) = \frac{1}{2^N}$$

□

We now offer an optimal guessing strategy for guessing the order of a deck after one riffle shuffle.

Strategy 3.6. First guess $C_1 = 1$. Then use the following case work for sequential guesses.

- Case One: Suppose $\max(\{C^{(k)}\}) = k$ for all $k \leq j-1$. Then guess j.
- Case Two: Suppose $\max(\{C^{(j-1)}\}) \neq j-1, N$. Let $m = \max\{C^{(j-1)}\}$. Then in the case that $|[m] \setminus \{C^{(j-1}\}| \geq |[N] \setminus [m]]|$ guess $\min([m] \setminus \{C^{(j-1}\})$ and otherwise guess $m+1$. Note in this case we know deterministically where the deck was split, so we know the next card will either be the next card from the stack containing the top card or the next card from the other stack. We are simply picking the card which is the next card of the stack which has more cards yet to be revealed.
- Case Three: Suppose $\max(\{C^{(j-1)}\}) = j-1$ but there is some $k < j-1$ such that $\max(C^k) \neq k$. Then guess j.
- Case Four: Suppose $\max(\{C^{(j-1)}\}) = N$. Then guess $\min([N] \setminus \{C^{(j-1)}\})$.

Theorem 3.7. Strategy 3.6 for is optimal for guessing the order of a deck after one riffle shuffle.

Proof. From Lemma 3.2 we know guessing $C_1 = 1$ is the optimal decision. We prove that for the following guesses, in each case we make the optimal decision. We start with Case Four and work our way to Case One. In each part of the proof, Theorem 1.3 guarantees we are making the optimal decision.

- Case Four: In this case, we have exhausted pile not containing the top card, so $P(C_j = g|C^{(j-1)}) = 1$ for $g = \min([N] \setminus \{C^{(j-1)}\})$.
- Case Three: In this case, we have exhausted the pile containing the top card, so $P(C_j = j|C^{(j-1)}) = 1$.
- Case Two: Again take $m := \max(\{C^{(j-1)}\}$. In this case, we have two possible guesses, the next card in the pile with the top card $g_1 = \min([m] \setminus \{C^{(j-1)}\})$ or the next card in the other pile $g_2 = m + 1$. This follows as if $g \notin \{g_1, g_2\}$, then clearly $\mathbf{P}(C_j = g|C^{(j-1)}) = 0$.

  Letting $s = \max([m] \setminus \{C^{(j-1)}\})$, we know that then that the deck was certainly split such that $K = s$. We can then compute

$$\begin{aligned}
\mathbf{P}(C_j = g_1|C^{(j-1)}) &= \sum_{k=0}^{N} \mathbf{P}(C_j = g_1, K = k|C^{(j-1)}) \\
&= \mathbf{P}(C_j = g_1, K = s|C^{(j)}) \\
&= \mathbf{P}(C_j = g_1|C^{(j-1)}, K = s)\mathbf{P}(K = s|C^{(j)}) \\
&= \mathbf{P}(C_j = g_1|C^{(j-1)}, K = s) \\
&= \frac{\mathbf{P}(C_j = g_1, C^{(j-1)}|K = s)}{\mathbf{P}(C^{(j-1)}|K = i)}
\end{aligned}$$

  Using Theorem 3.5 we can treat this as a combinatorics problem. The ratio in the final line is the conditional ratio of how many orderings start with $C^{(j-1)}$ and have $C_j = g_1$ to how many orderings start with $C^{(j-1)}$ given that we know $K = s$. Since we split at $K = s$, the number of cards in the pile with the top stack is s. In $C^{(j-1)}$ there are $g_1 - 1$ such cards, meaning after placing g_1 at slot j we have $N - j$ slots to place the remaining $s - g_1$ cards from this pile. Thus following the reasoning in the proof of Lemma 3.3, we see that

$$\mathbf{P}(C_j = g_1|C^{(j-1)}) = \frac{\binom{N-j}{s-g_1}}{\binom{N-j+1}{s-g_1+1}} = \frac{s - g_1 + 1}{N - j + 1}$$

  Similar computation gives that

$$\mathbf{P}(C_j = g_2|C^{(j-1)}) = \frac{\binom{N-j}{s-g_1+1}}{\binom{N-j+1}{s-g_1+1}} = \frac{N - j - s + g_1}{N - j + 1}$$

  The quantity $s - g_1 + 1$ is exactly how many cards from the pile with the top card have yet to be revealed at the time of guessing, and the quantity $N - j - s + g_1$ is exactly how many cards from the other pile have yet to be revealed at the time of guessing. This implies the result.

- Case One: In this case we exactly have $C^{(j-1)} = (1, \ldots, j-1)$. The any guess $g \geq j$ has positive probability. Prior to proceeding, we will compute the conditional distribution of K given $C^{(j-1)}$ (among other things). Using Lemma 3.4 we know that if $i \leq j-1$ then

$$\mathbf{P}(C^{(j-1)}|K=i) = \binom{N}{i}^{-1}$$

On the other hand, if $i \geq j-1$ then

$$\mathbf{P}(C^{(j-1)}|K=i) = \binom{N-j+1}{i-j+1}\binom{N}{i}^{-1}$$

From this we can also compute

$$\begin{aligned}\mathbf{P}(C^{(j-1)}) &= \sum_{k=0}^{N} \mathbf{P}(C^j = c^{(j-1)}|K=k)\mathbf{P}(K=k) \\ &= 2^{-N}\left(j-1+\sum_{i=j}^{N}\binom{N-j+1}{i-j+1}\right) \\ &= 2^{-N}(j-2+2^{N-j+1})\end{aligned}$$

Thus

$$\mathbf{P}(K=i|C^{(j-1)}) = \mathbf{P}(C^{(j-1)}|K=i)\frac{\mathbf{P}(K=i)}{\mathbf{P}(C^{(j-1)})} = \begin{cases}\frac{1}{j-2+2^{N-j+1}} & i \leq j-1 \\ \frac{\binom{N-j+1}{i-j+1}}{j-2+2^{N-j+1}} & i \geq j-1\end{cases}$$

Using these computations, we break into subcases

  - Subcase One: Suppose we consider $g = j$. Then noting that $\mathbf{P}(C_j = j|K=i, C^{(j-1)}) = 1$ for $i \leq j-1$ we have

$$\begin{aligned}\mathbf{P}(C_j = j|C^{(j-1)}) &= \mathbf{P}(C^{(j)} = (1,\ldots,j)|C^{(j-1)}) \\ &= \mathbf{P}(C^{(j-1)}|C^{(j)} = (1,\ldots,j))\frac{\mathbf{P}(C^{(j)} = (1,\ldots,j))}{\mathbf{P}(C^{(j-1)})} \\ &= \frac{j-1+2^{N-j}}{j-2+2^{N-j+1}}\end{aligned}$$

  - Subcase Two: Suppose we consider $g > j$. Then, in the case that $C_j = g$, we know for a

fact that $K = g - 1$. Then

$$\begin{aligned}
\mathbf{P}(C_j = g | C^{(j-1)}) &= \sum_{i=0}^{N} \mathbf{P}(C_j = g, K = i | C^{(j-1)}) \\
&= \mathbf{P}(C_j = g, K = g-1 | C^{(j-1)}) \\
&= \mathbf{P}(C_j = g | K = g-1, C^{(j-1)}) \mathbf{P}(K = g-1 | C^{(j-1)}) \\
&= \frac{\binom{N-j}{g-j}}{\binom{N-j+1}{g-j}} \frac{\binom{N-j+1}{g-j}}{j-2+2^{N-j+1}} \\
&= \frac{\binom{N-j}{g-j}}{j-2+2^{N-j+1}}
\end{aligned}$$

Since $j - 1 + 2^{N-j} > \binom{N-j}{g-j}$ it is clearly optimal to pick j.

□

The strategy above can be more concisely written as

Strategy 3.8. First guess $C_1 = 1$. Then use the following case work for sequential guesses.

- Case One: Suppose $\max(\{C^{(j-1)}\}) \neq j-1, N$. Let $m = \max\{C^{(j-1)}\}$. Then in the case that $|[m] \setminus \{C^{(j-1}\}| \geq |[N] \setminus [m]]|$ guess $\min([m] \setminus \{C^{(j-1}\})$ and otherwise guess $m+1$.
- Case Three: Suppose $\max(\{C^{(j-1)}\}) = j-1$. Then guess j.
- Case Four: Suppose $\max(\{C^{(j-1)}\}) = N$. Then guess $\min([N] \setminus \{C^{(j-1)}\})$.

We now offer what McGrath's strategy would be for guessing the order of the cards after one riffle shuffle.

Strategy 3.9. Guess $C_1 = 1$. Then continue guessing $C_j = j$ until you are incorrect. At the point that you are incorrect, you know exactly where the deck has been cut, and thus know the remaining number of cards in each stack. Guess the next card in the stack which has the larger number of remaining cards.

Note this exactly lines up with our derived optimal strategy. The above work implies the following result.

Theorem 3.10. McGrath's strategy (Strategy 3.9) is optimal for guessing the order of a deck of cards after one riffle shuffle.

3.3 Future Work

We have answered the question for one riffle shuffle. How can we extend to k? In fact, k riffle shuffles amount to splitting the cards into > 2 piles and then performing a similar procedure of dropping cards from each pile with probability proportional to the size of the pile. Next steps are to generalize the above methodology to this case.