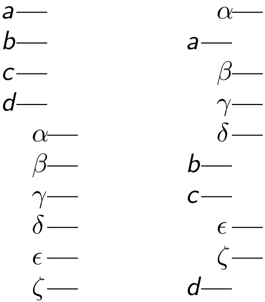
Lecture 37: A model for card shuffling

David Aldous

October 29, 2014

The usual scheme is called a **riffle shuffle**. [demo]



As with coin tossing, the point is that a human can't do exactly the same physical action each time. Unlike coin tossing, we need an explicit probability model for what a human does.

The model we use is called the GSR (Gilbert-Shannon-Reeds) model, and there are 3 equivalent ways to describe the model.

When dividing the deck, suppose a Binomial$(52, 1/2)$ *number of cards go to the left hand. And suppose that at each stage, the chance that the next drop is from the left or right hand is proportional to the number of cards remaining in that hand.*

Roughly speaking, this models a “quite good” card shuffler. This description is equivalent to

all possible riffle shuffles are equally likely.

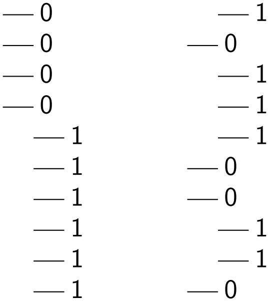
The question we study is

How many shuffles are needed for the deck to become “completely random”?

[board] connection with Markov chain theory – STAT 150.

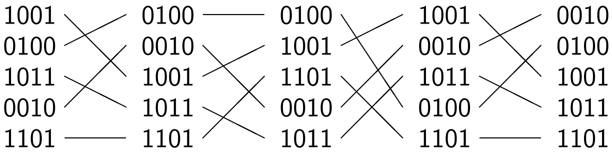
For the GSR model we can give an analysis which doesn't depend on any general theory.

We can record a particular shuffle as a sequence of 0s and 1s, in this case 1011100110



In the GSR model, every possible sequence of 0s and 1s is equally likely. So we can imagine (hypothetically, as math) a **reversed shuffle** in which we create IID random 0s and 1s as in the right diagram, and then get to the left diagram by pulling out the 0-cards (in order) and placing the pile of 0s on top of the remaining pile of 1s.

4 “shuffles of a 5-card deck



1001	0100	0100	1001	0010
0100	0010	1001	0010	0100
1011	1001	1101	1011	1001
0010	1011	0010	0100	1011
1101	1101	1011	1101	1101

Left-to-right represents 4 “radix sort” steps; right-to-left represents 4 riffle shuffles.

Reading left-to-right, if we start with cards named ABCDE and use random bits, the 4 “radix sort” steps get us to configuration DBACE.

conditional on all 5 of the 4-bit numbers being different, the right hand configuration in uniform random on all orderings of the deck.

But this works the same way right-to-left. We can implement the 4 riffle shuffles by using random bits which define a random integer: then

conditional on all these 5 integers being different, the 4 riffle shuffles get the deck into uniform random order.

Consider k shuffles of an n-card deck. We need to calculate

$$\mathbb{P}(\ n\ k\text{-bit numbers } \textbf{not} \text{ all different}).$$

But this is just the birthday problem with 2^k days, and the probability

$$\approx \binom{n}{2} 2^{-k}$$

when this is small. Fixing a large n, this becomes small when $k \approx 2 \log_2 n$, so this is a sufficient number of shuffles to mix an n-card deck.

Details of argument above in (undergrad-level) Aldous - Diaconis.

Bayer – Diaconis (*Trailing the dovetail shuffle to its lair*, 1992) gave a precise analysis (harder calculation by different method) for a 52-card deck;

number shuffles	k	4	5	6	7	8	9
non-uniformity	d(k)	1.000	0.924	0.614	0.334	0.167	0.085

Here $d(k)$ is "variation distance from uniformity". [board]

This work has entered popular science as "7 shuffles are enough" – try a Google search.

See also the book *Magic Tricks, Card Shuffling and Dynamic Computer memories* by S. Brent Morris.