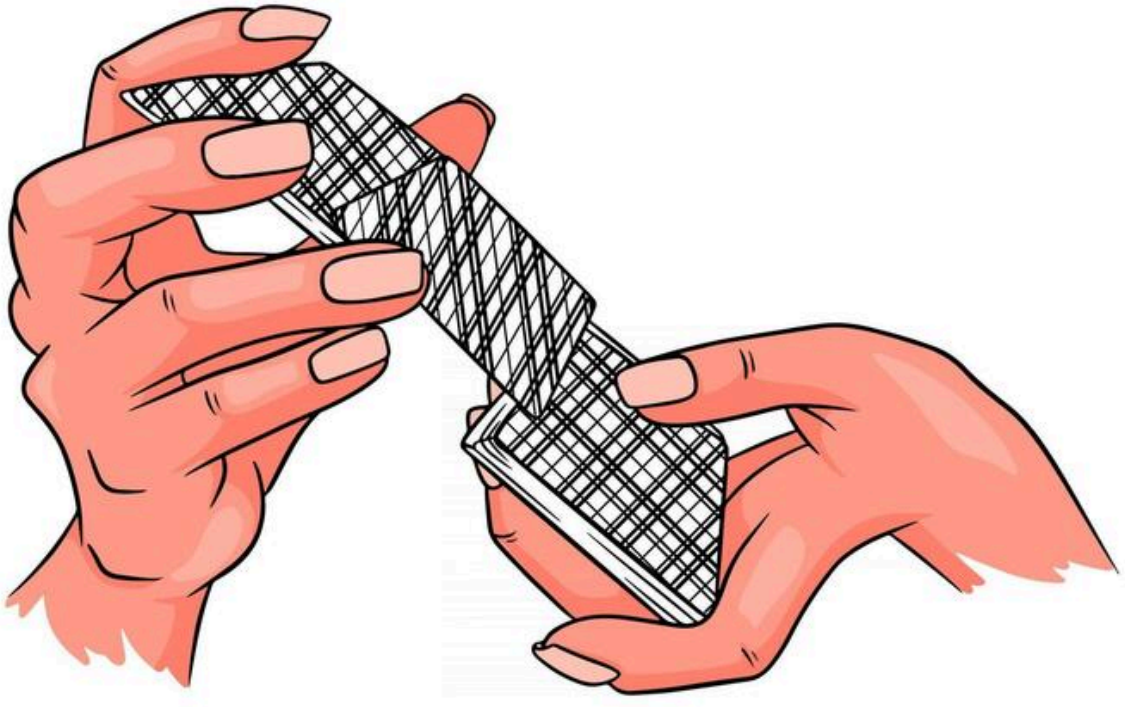
The Random shuffle

Collision and Shuffle Sufficiency in Riffle Shuffle



Group 4: Nour Romdhane, Katherine Orefice, Matteo Mainetti, Pablo Rosales, Elad Moshe

Madrid, November 2025

Hypothesis A

Probability of Getting the Same Order After Riffle Shuffling two Identically Ordered Decks

Problem Statement

The probability of obtaining the exact original order when shuffling a standard 52-card deck is often cited as 1/52!, which assumes a uniform distribution over all 52! possible permutations. This assumption does not hold for human shuffling. Empirical models show that typical riffle shuffles are biased: people tend to (i) cut near the middle, (ii) begin releasing cards from one hand, and (iii) drop cards in small clumps before switching hands. These behaviors concentrate probability mass on a small subset of permutations, so the effective distribution deviates substantially from uniform. Consequently, the theoretical 1/52! collision probability does not accurately describe real-world shuffling outcomes.

The goal of this project is to estimate the real-world probability that two human-style riffle shuffles, starting from the same factory order, produce the identical full-deck order.

The Pile Shuffle VS The Riffle Shuffle

For comparison, we took the example of the pile shuffle, which is considered one of the most basic and easy techniques for mixing cards. The main idea is for the dealer to divide the deck into a predetermined number of piles (for example, four, five, or seven) and then reassemble the piles. For instance, in a seven-pile shuffle, the first card goes to pile 1, the second to pile 2, … the eighth back to pile 1, and so on until every card is distributed. Hence, once the piles are assembled again, the deck order has altered in accordance with a systematic modular permutation pattern, following the selected order or a random one.
Research on card combinations reveals that given that pile shuffles only examine deterministic modular combinations, they are incredibly ineffective at producing randomness on their own. Persi Diaconis, an American mathematician, demonstrated that pile shuffles need numerous repetitions to give an approximate consistent randomness, even though the stacks are rearranged in a random order.

This study was conducted using the riffle shuffle instead of the pile and that's because the riffle shuffle has a deep mathematical foundation, as when the deck is split in half, cards fall alternately from each hand. Indeed, this results in a probabilistic process, whereas for the pile shuffle it's a deterministic one since it always follows the identical pattern. For instance, if we take seven piles and arrange them in an identical order, it just rearranges in a predictable manner rather than randomly.

Design

We recorded 75 single riffle shuffles of a standard 52-card deck. Every trial began from the same factory order and used the same procedure:

1. Cut size (K). The deck was cut roughly in half. We logged the exact cut: the number of cards in the right hand and, by complement, in the left hand (e.g., “28R–24L”). Observed cuts were near the middle (typical values in the mid-20s to ~30).
2. Starting hand. We noted which hand released cards first (Left or Right).
3. Interleaving sequence (run lengths). During the riffle, we recorded the *packets* (clumps) of consecutive cards dropped by each hand. Each packet was encoded as nL or nR, where n is the number of cards released in that clump and L/R indicates the hand (e.g., 1L, 2R). The resulting column labeled “Order of cards” is the full run-length sequence for that shuffle.

Example. For a cut of 28 (28 to the right, 24 to the left), we first mark the starting hand, then transcribe the packet sequence as it falls (e.g.,1L, 2R, 1L, 2R, 2L, …) until all 52 cards are interleaved.

Sample Results

Using 75 human riffle-shuffle trials (each starting from the same factory order), we estimated three probability mass functions (pmfs) that characterize typical behavior:

1) Cut-size pmf (K). Probability of cutting the deck at each position K (cards in the right hand; the left hand holds 52−K). The distribution is concentrated near the middle:

{20: 0.01, 22: 0.01, 23: 0.11, 24: 0.11, 25: 0.13, 26: 0.09, 27: 0.13, 28: 0.15, 29: 0.09, 30: 0.08, 31: 0.04, 32: 0.04}.

This confirms a strong tendency to cut approximately in half.

2) Start-hand pmf. Probability of which hand releases cards first:

{'L': 0.65, 'R': 0.35}.

Participants began from the left hand in roughly two-thirds of trials.

3) Run-length pmf (clump size). Probability that a hand drops a clump of t consecutive cards before switching:

{1: 0.45, 2: 0.31, 3: 0.13, 4: 0.06, 5: 0.04}.

Short clumps dominate; larger clumps are rare. For measurement consistency, all clumps exceeding five cards were grouped into a single 5+ category.

These empirical pmfs capture the core biases of human riffle shuffling (near-middle cuts, a preferred starting hand, and small clumps) and are used as inputs to our collision-probability model.

Calculating the Probability

Two independently riffled decks (starting from the same factory order) end in the same final order only if three components match:

1. Cut position: both cut at the same K (cards in one hand vs. 52−K in the other).
2. Starting hand: both begin releasing cards from the same side S∈{L,R}.
3. Clump sequence: both produce the same sequence of run-lengths (clump sizes) from alternating hands, in the same order, until all 52 cards are released.

Same Cut

Let pcut(K) be the probability that a single person cuts at position K.
Independently matching at K has probability:

$$pcut(K) \times pcut(K) = pcut(K)^2$$

Same Starting Hand

Let pstart(S) be the probability that a single person starts from side S∈{L,R}.
Independently matching the starting side has probability:

$$pstart(S) \times pstart(S) = pstart(S)^2$$

Same Clump Sequence

At each drop, a person emits a clump of size t∈{1,2,3,4,5+} with probability prun(t).
For one *step* of the shuffle, both choosing the same t occurs with probability:

$$\sum_{t} prun(t)^2$$

A shuffle consists of many steps; matching the entire run-length sequence is the product of those stepwise match probabilities. We denote the resulting factor, given cut K and starting side S, by:

$$Fhuman(K, 52-K, S)$$

Putting Everything Together

For a particular cut K and starting side S, the match probability is:

$$pcut(K)^2 \times pstart(S)^2 \times Fhuman(K, 52-K, S)$$

Because two shuffles could match at different (K,S) combinations, the overall collision probability sums over all possibilities:

$$\sum_{K}\sum_{S}(pcut(K))^2 \times (pstart(S))^2 \times Fhuman(K, 52-K, S)$$

Results

Evaluating the expression with the empirical pmfs from our experiment yields a collision probability of 4.969×10−13 (i.e., 0.0000000000004969), which is approximately 1 in 2.0×10^12 shuffles (about one in two trillion).

For comparison, under a uniform model over all 52! permutations, the collision probability for two independent shuffles would be 1/52! ≈ 1.93×10^−68, vastly smaller than the human-shuffle estimate.

Code Implementation

The pseudocode provided operationalizes the modeling framework described above.

```
function COLLISION_PROBABILITY(p_cut, p_start, p_run):
    total ← 0
    for each K in support(p_cut):
        for each S in {L, R}:
            w ← (p_cut[K])^2 * (p_start[S])^2
            F ← F_HUMAN(K, 52 - K, S, p_run)
            total ← total + w * F
    return total

# Probability that TWO independent shuffles match in clump sequence
# given cut (L_rem, R_rem) and starting side S.
function F_HUMAN(L_rem, R_rem, S, p_run):
    memo ← empty_map()
    return MATCH_PROB(L_rem, R_rem, S, memo, p_run)

function MATCH_PROB(L_rem, R_rem, turn, memo, p_run):
    # Base cases
    if L_rem == 0 and R_rem == 0:
        return 1.0                  # both piles exactly exhausted, sequences matched
    if L_rem < 0 or R_rem < 0:
        return 0.0                  # infeasible path

    key ← (L_rem, R_rem, turn)
    if key in memo: return memo[key]

    # Determine allowed clump sizes for the current hand
    if turn == 'L':
```

```
    max_allowed ← min(5, L_rem)
    # If the right pile is empty, the rest must be dropped from the left in one final clump
    if R_rem == 0: allowed ← { L_rem }       # exact finish
    else:          allowed ← {1,2,3,4,5} ∩ {1..max_allowed}
else:  # turn == 'R'
    max_allowed ← min(5, R_rem)
    if L_rem == 0: allowed ← { R_rem }
    else:          allowed ← {1,2,3,4,5} ∩ {1..max_allowed}

prob ← 0.0
for each t in allowed:
    step_match ← (p_run[t])^2              # both choose the same clump size t
    if turn == 'L':
        prob ← prob + step_match * MATCH_PROB(L_rem - t, R_rem, 'R', memo, p_run)
    else:
        prob ← prob + step_match * MATCH_PROB(L_rem, R_rem - t, 'L', memo, p_run)

memo[key] ← prob
return prob
```

Hypothesis B

How Many Riffle Shuffles Are Enough to Randomize a Deck?

Problem Statement

This section examines how many “Riffle” shuffles are required to randomize a deck of cards. Using combinatorial reasoning and probability theory, we model how the number of possible deck permutations grows with each shuffle and identify the practical threshold for randomness for an actual “human like” Riffle. This model is based both on basic combinatorics and on the “Gilbert–Shannon–Reeds model” which put the framework of measuring actual human like randomness in Riffling an $n - card\ deck$.

Theoretical Background

A deck of *n* cards can be arranged in *n!* possible orders.
For a 52-card deck, this equals approximately $52! = 8.07 \times 10^{67}$ permutations.

A “fully random” shuffle is one that can produce any of these permutations with (approximately) equal probability.

The riffle shuffle algorithm is as follows:

1. Split the deck into two piles (not necessarily equal).
2. Interleave the cards while maintaining the internal order within each pile

In theory, the split can potentially divide the n-size deck to n+1 divisions but in reality, even with best technique, the split happens around the middle of the deck in order to allow easy shuffling. This intuitive human behaviour reduces the overall possible permutations.

Also the interleaving process is also unevenly distributed in real life where usually interleaved groups will be relatively fractional to the overall splited deck.

Measuring Randomness

There are many ways to quantify actual “randomness” of a shuffling process. The most common way is that of Bayer - Diaconnis: the euclidean product of the actual possibility vector of permutation and the uniform possibility vector that assign each permutation with the possibility of $\frac{1}{n!}$. This can lead to a proper estimation and comparison between different shuffling processes.

Because of the sheer amount of different permutations, the computation of a direct permutation vector required is enormous O(n!). To lower that we could make the vectors

shorter by grouping “permutations families” with similar properties and still have good estimations. Such grouping can be done by numbering ascending orders → O(n), measuring coupling distance → O(n) and so on.

Coupling distance= measuring the difference between each 2 adjacent cards and summing it.
For exemple:
$12345 \rightarrow (2-1)+(3-2)+(4-3)+(5-4)=4$
$15243 \rightarrow (5-1)+(|2-5|)+(4-2)+(|3-4|)=10$

Simplified Example: 10-Card Deck

To make the concept tangible, consider a deck of ten cards.
Total possible orders $10! = 3,628,800$

Possible theoretical riffle outcomes: since each card can come from the left (L) or right (R) pile, there are $2^{10} = 1,024$ possible interleavings per shuffle.

thus , after one riffle:

$$\frac{1{,}024}{3{,}628{,}800} \approx 0.02\%$$

That is, even with a theoretical approach, we cover only 0.02% of the overall amount of permutations. That means the majority of permutations are not even possible to achieve with one shuffle.

Each new riffle ‘k’ roughly multiplies the number of accessible outcomes:

$$2^{10*k}\ permutations\ after\ k\ shuffles$$

- After two riffles: $2^{20} = 1,048,576 \ -> \ \approx 29\%\ coverage$
- After three riffles: $2^{30} = 1.07 \times 10^{9} >> 10!$

Since the rifle process is not uniform (some cuts are more likely than others), an extra shuffle is needed to even out the probabilities.

Result: about **four riffle shuffles** allow us to cover a significant amount of the total overall number of permutations of a 10-card deck.

Extension to a 52-Card Deck

Applying the same reasoning:

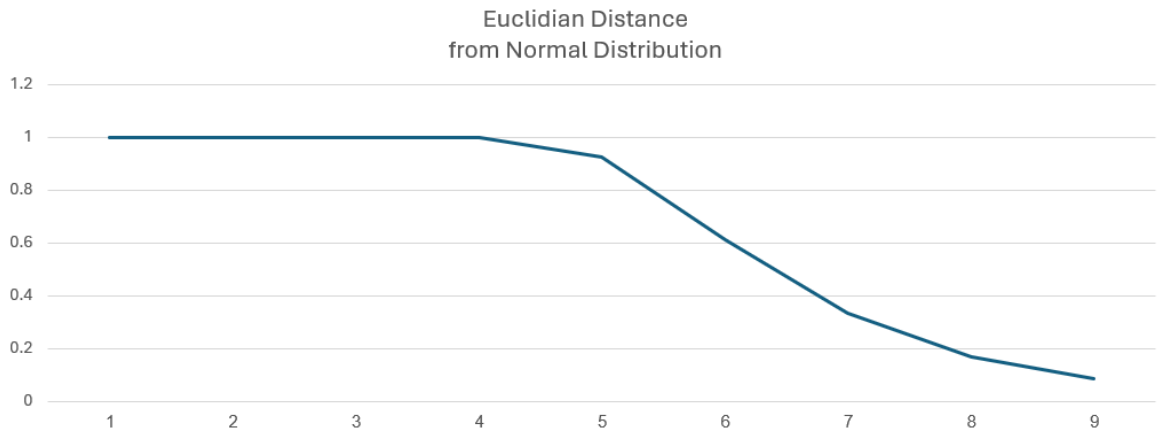
# of Riffles	Coverage $\approx 2^{52k} / 52!$	Comment
1	$\approx 10^{-52}$	No randomness
4	≈ 0.001% (≈ 1 in 20 000)	Low coverage
5	$\approx 2.2 \times 10^{10} >> 100\%$	Full combinatorial coverage

However, uniformity still matters: even when all permutations become theoretically reachable, not all are equally likely - as shown in the Theoretical Background. Empirical analysis by Persi Diaconis (Stanford University) shows that seven riffle shuffles make the distribution of deck orders statistically indistinguishable from uniform randomness.

The "Gilbert–Shannon–Reeds model" also uses the theoretical maximum number of ascending orders within a n-shuffled deck. It states that for a permutation on **n** cards with **r** rising sequences using **a**-riffle shuffles, **r** of the cuts are determined and **a-r** can be anywhere resulting in a probability of: $\frac{\binom{n+a-r}{n}}{a^n}$

That gives us the following table:

#shuffles	Euclidian Distance from Normal Dist.
1	1
2	1
3	1
4	1
5	0.924
6	0.614
7	0.334
8	0.167
9	0.085



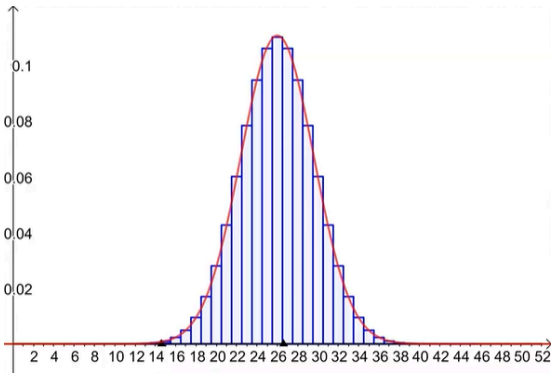
Appendix *for hypothesis B*

A. Counting the Outcomes of One Riffle

For a deck of n cards cut into piles of size k and $n-k$, the number of valid interleavings preserving internal order is:

$$\binom{n}{k}$$

Giving a theoretical distribution that looks like this:



Where in reality the variance could be even smaller taking into account human splitting.

Summing over all possible cuts:

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Hence a single riffle yields 2^n possible outcomes (e.g., $2^{10} = 1,024$ for a 10-card deck).

B. Coverage after *k* Riffles

An upper bound on the number of distinct outputs after *k* riffles 2^{nk}, giving

$$coverage(k) = \frac{2^{nk}}{n!}$$

For *n* = 10, coverage rises from 0.02% (1 riffle) to 29% (2 riffles), surpassing full coverage after 3 riffles.

C. Why Coverage > 1 Is Not Enough

The rifle model is non-uniform: cuts cluster near the middle and interleavings are biased. Randomness is measured by the total variation distance between the shuffle distribution and the uniform distribution on all *n!* permutations.
Research shows this distance falls sharply near seven riffle shuffles for a 52-card deck.

D. Practical Threshold

- Combinatorial coverage → reached by 5 riffles.
- Statistical uniformity → achieved by ≈ 7 riffles.

Therefore: ***Seven riffle shuffles are mathematically sufficient to randomize a standard 52-card deck.***