Shuffling Cards & Stopping Time

Team members: Boxiao, Wenjie, Cindy

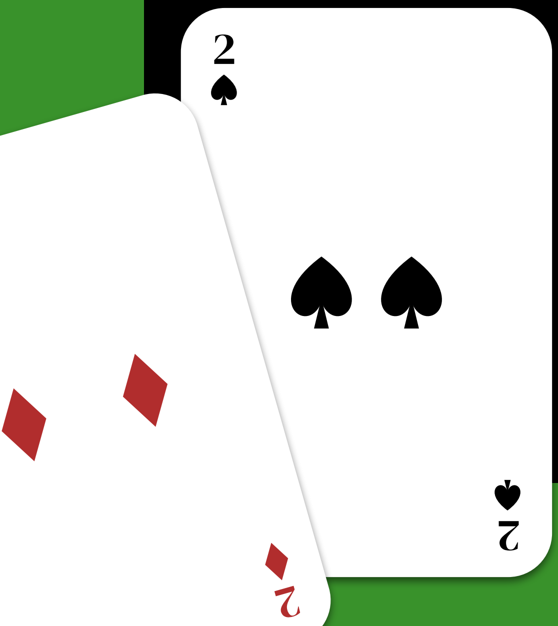
The Markov Chain in Cards

Let (X_t) be a Markov Chain representing the permutation of cards at time *t* of shuffles.

Consider an n-card deck...

The state space S is the set of permutation of n cards, so the size is n!.

The index set I represents discrete time I = {0, 1, 2, 3, ...}.



How we define the Random?

A shuffle is one step of this ergodic Markov Chain.
We expect each state can transfer to the same numbers of other states, so the limiting distribution is uniform on the set of permutations.
For permutation(state) i, $\pi_i = 1/n!$.

Theoretically, the limiting distribution is achieved as the number of shuffles tends to infinity. However, what we want to find is how many shuffles are needed to bring the deck sufficiently close to random, which means a measurement of how far a Markov chain is from its limiting distribution.
We used **Total Variation Distance**!

Total Variation Distance

How many steps of the chain are close enough to its limiting distribution?

Total Variation Distance

Let P be the transition matrix of an ergodic Markov chain with a limiting distribution π. The total variation distance at time n is

$$v(n) = \max_{i \in S} \max_{A \subseteq S} |P(X_n \in A | X_0 = i) - \pi_A|$$

The maximum absolute difference over all end states A and all starting states between Probability of being in state A at time n and π_A.

If $v(n) = 0$, then the chain is in stationarity at time n. Over time, $v(n) \to 0$, as $n \to \infty$. It can be shown that the definition is equivalent to

$$v(n) = \max_i \frac{1}{2} \sum_j |P_{ij}^n - \pi_j|$$

Example 5.9 from Textbook

$$P = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & \quad 2 \end{array} \\ \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array}.$$

Solution

Recall that

$$P^n = \frac{1}{p+q} \begin{pmatrix} q + p(1-p-q)^n & p - p(1-p-q)^n \\ q - q(1-p-q)^n & p + q(1-p-q)^n \end{pmatrix}$$

and

$$\pi = \left(\frac{q}{p+q} \quad \frac{p}{p+q} \right).$$

For $i = 1$,

$$\sum_{j=1}^{2} |P^n_{1j} - \pi_j| = \frac{2p}{p+q}(1-p-q)^n.$$

For $i = 2$,

$$\sum_{j=1}^{2} |P^n_{2j} = \pi_j| = \frac{2q}{p+q}(1-p-q)^n.$$

Example 5.9 from Textbook

Thus,

$$v(n) = \frac{\max(p, q)}{p + q}(1 - p - q)^n.$$

Convergence to stationarity, as measured by total variation distance, occurs exponentially fast, with the rate of convergence governed by $1 - p - q$. Note that the second largest eigenvalue of $\boldsymbol{P}$ is $\lambda_2 = 1 - p - q$.

01

Top In Add Random Shuffle

Consider the top card is removed and inserted into the random n position for each time.

The next position in {1, n} from the bottom it moved to is uniform distributed with *Unif* {1, n}.

Consider the bottom card (marked as card B) at time t=0.

Let (T_i) represents the waiting time of shuffles that a card inserted below the initial bottom card.

The card B stays at the bottom until the first time T_1 a card is inserted below it. T_1 takes about n shuffles.

$$E[T_1] = n$$

As the shuffles continue, at time T_2 the second card is inserted below the card B, taking about n/2 shuffles.

$$E[T_2 - T_1] = n/2$$

The two cards under card B are equally likely to be in any kinds of permutations, relative order low-high or high-low.

Similarly, at time T_i, $i \in \{1, n\}$, the i-th card is inserted under the card B with a random position. T_i takes about n/i more shuffles.

$$E[T_i - T_{i-1}] = n/i$$

Consider at time T_{n-1} the original bottom card B comes up to the top. By an inductive argument, all (n-1)! arrangements of cards below the card B are equally likely.

When the card B is inserted at random at time $T_n = T_{n-1} + 1$, all n! possible arrangements of the deck are equally likely.

The waiting time T_n is about

$$E(T) = n + \frac{n}{2} + \frac{n}{3} + \ldots + \frac{n}{n} = n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right) \approx n \log n$$

The T_n is known as **Strong Stationary Time.**

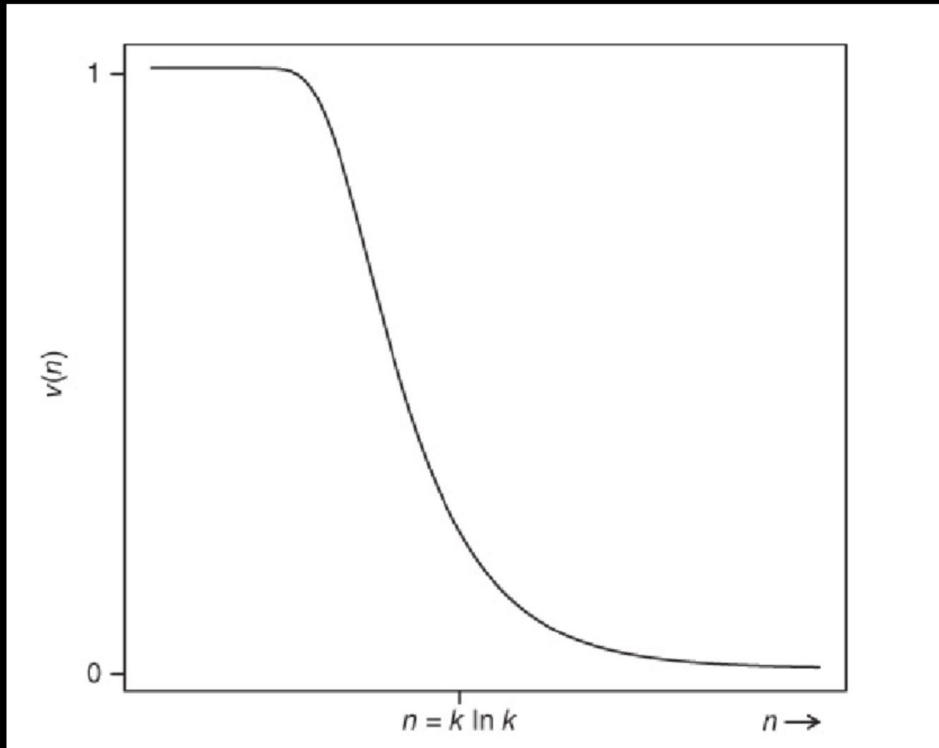
Strong Stationary Time

Strong Stationary Time

A strong stationary time for an ergodic Markov chain $X_0, X_1, ...$ is a stopping time T such that

$$P(X_n = j, T = n) = \pi_j P(T = n), \text{ for all states } j \text{ and } n \geq 0$$

The Cut-off Phenomenon



Strong Stationary Time and Total Variation Distance Lemma

For $n > 0$,

$$v(n) \leq P(T > n)$$

Applying the Lemma to TtR Shuffle

Strong Stationary Time and Total Variation Distance Lemma

For $n > 0$,

$$v(n) \leq \underline{P(T > n)}$$

There are still some cards that have not been inserted below the original bottom card

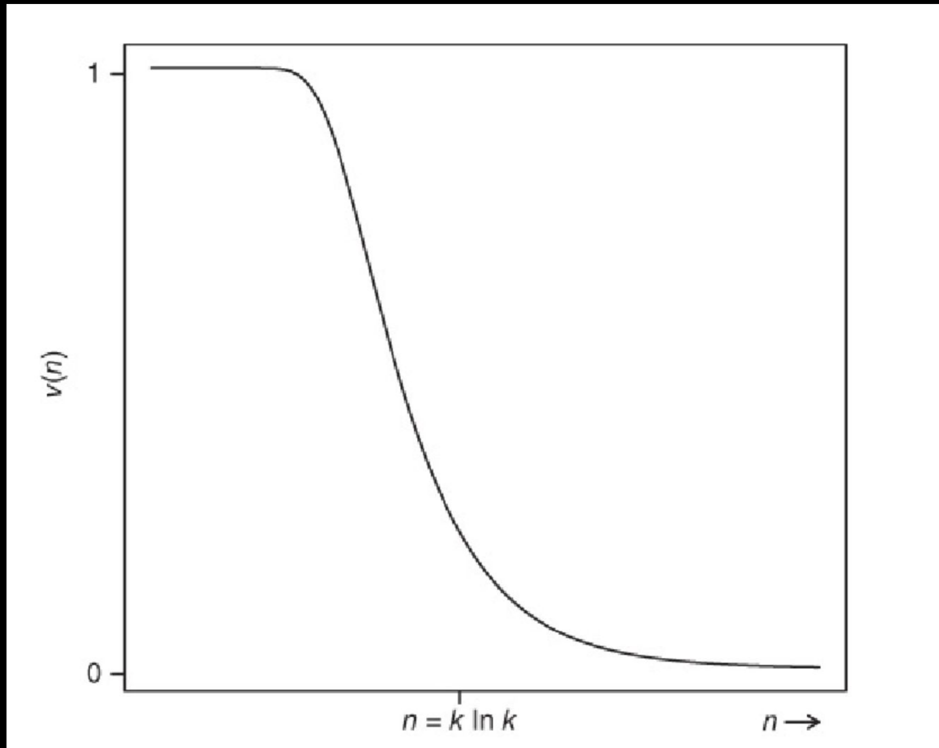
Applying the Lemma to TtR Shuffle

$$v(n) \leq P(T > n) = P(C > n) \leq ke^{-n/k}.$$

To make total variation distance small, find a value of n that makes the righthand side of Equation (5.5) small. For $c > 0$, let $n = k \ln k + ck$. Then,

$$v(n) = v(k \ln k + ck) \leq ke^{-(\ln k + c)} = e^{-c} \approx 0.$$

The Cut-off Phenomenon



Strong Stationary Time and Total Variation Distance Lemma

For $n > 0$,

$$v(n) \leq P(T > n)$$

02

Riffle Shuffle

Binomial Distribution Model

When you perform a riffle shuffle, you essentially decide the position where the deck will be split into two piles. Each card has a 1/2 probability of ending up in either pile. Thus, the number of cards in one pile follows a binomial distribution:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Applying to shuffle:

$$\binom{n}{k} \left(\frac{1}{2}\right)^n$$

Total variation distance

Total variation distance is a measure used in statistics to quantify the difference between two probability distributions. For two discrete probability distributions P and Q on the same probability space, the TVD is defined as:

$$\mathrm{TVD}(P, Q) = \frac{1}{2} \sum_{x} |P(x) - Q(x)|$$

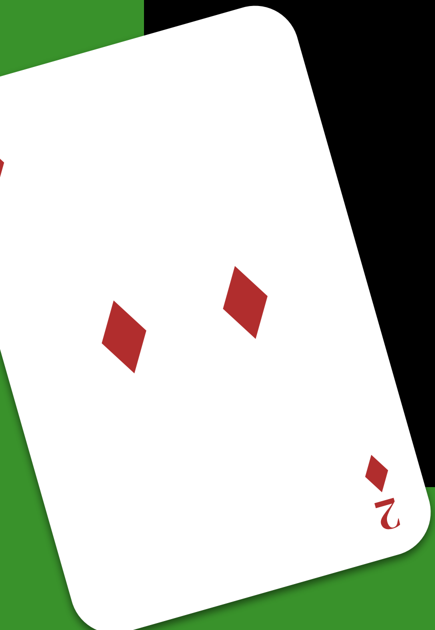
This measurement reflects the maximum possible difference between the probabilities that the two distributions assign to the same event.

Application to Card Shuffling

$$\mathrm{TVD}(P_k, U)$$

P_k represents the probability distribution of the deck's arrangement after k shuffles, and U represents the uniform distribution where every permutation of the deck is equally likely.

We want P_k to approach U, meaning the deck becomes increasingly randomized until the TVD is sufficiently small, indicating that no particular permutation is significantly more likely than any other.



Derivation of TVD in Shuffling

Bayer and Diaconis give a thorough analysis of the total variation distance for the resulting Markov chain. They show that about (3⁄2)log2(k) shuffles are necessary and sufficient to make total variation distance small.

*** The mixing time is approximately 7 to 8 times! ***

Cut-off phenomenon

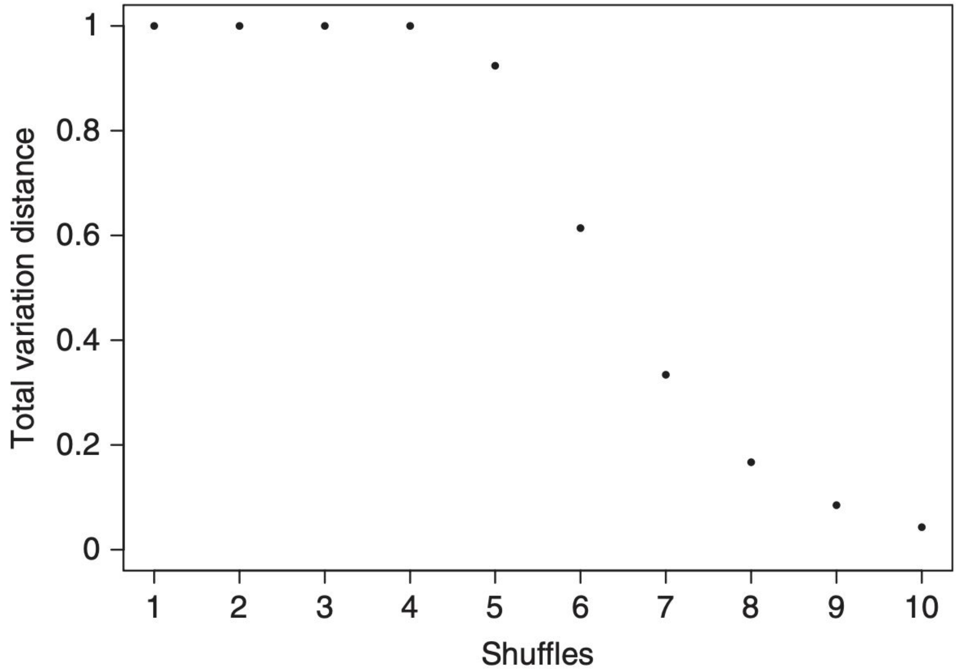


Figure 5.14 Total variation distance for riffle shuffling.

TABLE 5.2 Total Variation Distance for *n* Shuffles of 52 Cards

n	1	2	3	4	5	6	7	8	9	10
$v(n)$	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043

Source: Data from Bayer and Diaconis (1992).

Resources

Robert P. Dobrow, Introduction to stochastic processes with R(2016), Chapter 5.6

Aldous and Diaconis, Shuffling Cards and Stopping Times, 1986

Thanks for your listening!