数学入門公開講座

マルコフ連鎖と混合時間

ー カード・シャッフルの数理 ー

熊谷　隆

(数理解析研究所)

http://www.kurims.kyoto-u.ac.jp/~kumagai/japanlist.html

2011年8月1日– 4日

0 Introduction

52枚のトランプのカードを切るとき、何回くらい切ればカードがよく混ざるか？

$\Rightarrow$ "In Shuffling Cards, 7 Is Winning Number." (New York Times 1990年1月9日)

P. Diaconis教授らの成果を報道

大体7回切ればよく、しかも7回前後で急に「よく混ざった」状態になる

（注）ここでいうシャッフルは、リフル・シャッフルと呼ばれる切り方。

（論文）[1] D. Bayer and P. Diaconis: Ann. Appl. Probab. **2** (1992), 294–313.

そもそも (Q) トランプのシャッフルを数学的にどのようにモデル化する？

(Q) 「よく混ざる」とはどういうこと？

1　シャッフリングの表現

（例1）1 2 3 4 5 6 という並びのカードを切った結果1 4 2 5 3 6 となったとする

$\pi(i)$ を、下（左）から i 番目のカードが移った場所とすると

$$\pi(1)=1,\ \pi(2)=3,\ \pi(3)=5,\ \pi(4)=2,\ \pi(5)=4,\ \pi(6)=6$$

となる。($\{1,2,\cdots,6\}$ から $\{1,2,\cdots,6\}$ への全単射（1対1対応）！）　これを

$$\pi=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 5 & 2 & 4 & 6 \end{pmatrix}$$

と書く。このとき、

$$\pi^{-1}=\begin{pmatrix} 1 & 3 & 5 & 2 & 4 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 5 & 3 & 6 \end{pmatrix}$$

$\pi^{-1}(i)$ は、切った後、下（左）から i 番目のカードがもともとあった場所 を表す。

（例2） ２回切る: $1\ 2\ 3\ 4\ 5\ 6 \Rightarrow 1\ 4\ 2\ 5\ 3\ 6 \Rightarrow 1\ 4\ 3\ 2\ 6\ 5$となったとする

（２回目の切り方は$1\ 2\ 3\ 4\ 5\ 6 \Rightarrow 1\ 2\ 5\ 3\ 6\ 4$） ２回目の切り方は

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 4 & 6 & 3 & 5 \end{pmatrix}$$

このとき、２回切った操作は以下のように計算できる。

$$\sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 5 & 2 & 4 & 6 \\ 1 & 4 & 3 & 2 & 6 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 2 & 6 & 5 \end{pmatrix}$$

なお、下（左）からi番目のカードがもともとあった場所を見るには

$$(\sigma \circ \pi)^{-1} = \pi^{-1} \circ \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 2 & 6 & 5 \end{pmatrix}$$

カードがd枚のとき、今具体的に見た操作をまとめなおすと、

d枚のカードを（一回）シャッフルする操作

カードを下から上に一列に並べ、下から順に$1, 2, \cdots, d$と番号付けする

$\Rightarrow$ カードをシャッフルすると（例えば）番号1のカードが下からi_1番目に移り、

2のカードが下からi_2番目に移り$\cdots$という風にカードの並び替えが行われる。

$\Rightarrow$ 数学的には、$\pi(1) = i_1, \pi(2) = i_2, \cdots, \pi(d) = i_d$なる全単射

$\pi : \{1, 2, \cdots, d\} \to \{1, 2, \cdots, d\}$（このような$\pi$を置換という）で表される。これを

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & d \\ i_1 & i_2 & \cdots & i_d \end{pmatrix} \tag{1}$$

と書く。$\{1, 2, \cdots, d\}$上の置換全体をd次対称群と呼び、$\mathcal{S}_d$と書く。（$\sharp \mathcal{S}_d = d!$）

$\mathcal{S}_d$についてもう少し

$\mathcal{S}_d$は（合成を積とする）群をなす。例えば$d=3$として

$$\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \text{ならば } \tau\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \neq \pi\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

ここで、$\tau \circ \pi$（πを施した後τを施すという操作）を$\tau\pi$と書いた。

- 結合律：　$(\pi_1\pi_2)\pi_3 = \pi_1(\pi_2\pi_3)$を満たす。
- 単位元（単位置換）：　どの文字も動かさない置換idは、$(id)\pi = \pi(id) = \pi$.
- 逆元（逆置換）：　πが(1)で与えられるとき、その逆元は$\pi^{-1} = \begin{pmatrix} i_1 & i_2 & \cdots & i_d \\ 1 & 2 & \cdots & d \end{pmatrix}$

カード・シャッフルに話を戻す

カードをシャッフルする時には、上の置換をランダムに行う、つまり何らかの確率分布に従って$\mathcal{S}_d$の元を選ぶという操作をしている。

例 **1.1** ランダム・カット：下から上に並んだd枚のカードをランダムに２つの山に分け、上の山をもう下の山の前に移動するという操作のこと。

（ランダム・カットの数学的表現）

1からdまでの間の数字kを等確率で選び（つまり、すべての$1 \leq i \leq d$について$P(k=i)=1/d$、一様分布に従って選ぶという）、以下の置換を行うこと。

$$\pi_k := \begin{pmatrix} 1 & 2 & \cdots & k & k+1 & \cdots & d \\ d-k+1 & d-k+2 & \cdots & d & 1 & \cdots & d-k \end{pmatrix} \tag{2}$$

つまり、$1 \leq i \leq d$の各kについて確率$1/d$で$\pi_k \in \mathcal{S}_d$という置換を選ぶ操作のこと。

（注） 数学的モデル化の際はある種の理想化をする。

– 今の場合、実際にトランプを切るとき常に同じ確率分布で切ることは人間業では不可能。分布が一様分布であることが妥当かどうかも、明らかではない。

例 **1.2** **トップ to ランダム**：

一番下のカードを取って、ランダムにカードの山の中に入れる操作のこと

(ランダム・カットの数学的表現)：ランダムな選択を、一様分布に従うものとすると

$1 \leq k \leq d$の各kについて確率$1/d$で以下の置換を選ぶ操作のこと。

$$k \geq 2 \text{ のとき } \tau_k := \begin{pmatrix} 1 & 2 & \cdots & k & k+1 & \cdots & d \\ k & 1 & \cdots & k-1 & k+1 & \cdots & d \end{pmatrix}, \quad k = 1 \text{ のとき単位置換} \tag{3}$$

例 **1.3** **ランダム to トップ**：　(*Cf.* *Amazon*のランキング、超整理法)

カードの山の中からランダムにカードを取って、一番下に入れる操作のこと

(ランダム・カットの数学的表現)：ランダムな選択を、一様分布に従うものとすると

$1 \leq k \leq d$の各kについて確率$1/d$で以下の置換を選ぶ操作のこと。

$$k \geq 2 \text{ のとき } \hat{\tau}_k := \begin{pmatrix} 1 & \cdots & k-1 & k & k+1 & \cdots & d \\ 2 & \cdots & k & 1 & k+1 & \cdots & d \end{pmatrix}, \quad k = 1 \text{ のとき単位置換} \tag{4}$$

例 1.4 ランダム・トランスポジション： d 枚のカードから、ランダムに２枚を選び（同じカードを選ぶことも許す)、その２枚を交換して戻す操作のこと

（ランダム・カットの数学的表現)：２枚のカードの選び方は、それぞれ一様分布に従い独立（一方の取り出し方が他方に影響しない）とすると...

- 確率 $1/d$ で単位置換（確率 $1/d$ で同じカードを選ぶ事になるから）
- 任意の $i < j$ なるペアについて、以下の置換を確率 $2/d^2$ で選ぶ（ペアとして (i, j) が選ばれるのは、i, j の順で選ばれる場合と j, i の順で選ばれる場合があるから）

$$\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & d \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & d \end{pmatrix}.$$

例 1.5 リフル・シャッフル：カードを二つの山に分け、パラパラ混ぜていく操作。

（リフル・シャッフルの数学的表現）順を追ってみて行く。

i) カードを二つの山に分ける：二項分布に従って分ける。つまり、$1 \leq k \leq d$ なる k について、山の大きさが k と $d-k$ になる確率が ${}_dC_k 2^{-d}$ であるとする。

ii)　次にパラパラパラと混ぜて行くステップ：

右手に k 枚、左手に $d-k$ 枚持っているとして、

$$\text{最初にパラッと机に落ちるカードが}\begin{cases}\text{右手のカードである確率は } k/d \\ \text{左手のカードである確率は } (d-k)/d\end{cases}$$

帰納的に考えて、右手に a 枚、左手に b 枚残っている段階では、

$$\text{次に落ちるカードが}\begin{cases}\text{右手のカードである確率は } a/(a+b) \\ \text{左手のカードである確率は } b/(a+b)\end{cases}$$

この操作を、両手のカードがすべて机に落ちるまで続けたとき、

- 起こりうる「二つの山の混ざり方」はすべて当確率で起こる。
- その確率は $k!(d-k)!/d! = \boxed{1/{}_dC_k}$ となる。

$\because$「最初に右手から、次に左手から$\cdots$」といった具合に起こりうるそれぞれの混ざり方の確率を計算すると、常に分母は $d\cdot(d-1)\cdots 1 = d!$ となり、分子は（かけ算の順番を無視すると）$k\cdot(k-1)\cdots 1\cdot(d-k)\cdot(d-k-1)\cdots 1 = k!(d-k)!$ となる。

iii) 結局混ざったカードの並び方はどういう確率分布をもつ？

iii)-1) シャッフルの結果、結局カードの順番が全く変わらない（つまり単位置換が選ばれた）確率は、

$$\sum_{k=0}^{d} P\Big(\text{一つ目の山(つまり左手にもつカード)が}k\text{枚}\Big)P\Big(\text{順番変わらず}\,\Big|\,\text{一つ目の山が}k\text{枚}\Big)$$

$$= \sum_{k=0}^{d} \frac{{}_dC_k}{2^d}\cdot\frac{1}{{}_dC_k} = \frac{d+1}{2^d}. \tag{5}$$

iii)-2) 単位置換ではないとすると、混ざったカードの順番はどうなっているか？

二つの山に分けたとき、それぞれの番号は$1,2,\cdots,k$と$k+1,\cdots,d$であり、これらが順に混じり合う（しかも単位置換にはならない）のだから、結局以下の$\mathcal{R}$の元に属する置換を選んだ事になる。

$$\mathcal{R} := \{\pi \in \mathcal{S}_d : 1 \leq l \leq d-1 \text{なる} l \text{があって、} \pi(1) < \cdots < \pi(l) > \pi(l+1), \quad \pi(l+1) < \cdots < \pi(d)\}$$

では、$\pi_l \in \mathcal{R}$はどんな確率で選ばれるか？（ここでπ_lのlは、$\mathcal{R}$の元の定義に出て来るlを表すものとする。）π_lが選ばれるには、はじめに二つの山に分けるときに一つ目の山がl枚でなくてはいけないので、求める確率は

$$P\Bigl(\text{一つ目の山 (つまり左手にもつカード) が}\, l\, \text{枚}\Bigr)P\Bigl(\pi_l \,\Big|\, \text{一つ目の山が}\, l\, \text{枚}\Bigr) = \frac{{}_dC_l}{2^d} \cdot \frac{1}{{}_dC_l} = \frac{1}{2^d}.$$

つまり$\mathcal{R}$の元はどれも等確率で選ばれた事になる！ □

実際にはカード・シャッフルは何回も繰り返して行うものである。

何回も繰り返す操作は数学的にどのように記述すればよいか？

（答え）$Y_1, Y_2, \cdots, Y_n$を、それぞれ$\mathcal{S}_d$に値を持つ独立な確率変数とする。各Y_iが、上の例で見た一回のカード・シャッフル（ある確率分布に従って$\mathcal{S}_d$の元を選ぶ操作）にあたる。このとき、

$$X_n := Y_n \cdot Y_{n-1} \cdots Y_1$$

とすると、X_nがn回のカード・シャッフルの結果に相当する。

ここで、X_nはX_{n-1}の情報とY_nの情報があれば決まることに注意しよう。

$$X_n := Y_n \cdot Y_{n-1} \cdots Y_1 \tag{6}$$

これで、カード・シャッフルを繰り返すという操作の数学的記述、つまり$\{X_i\}_{i=0,1,2,\cdots}$という確率変数の族によるランダムな時間発展ができた。(X_0は単位置換とする。)このような確率変数の族(ランダムな時間発展)を確率過程と呼ぶ。

特に、X_nがX_{n-1}の情報と、推移確率と呼ばれる確率分布のみから決まり、それ以前の過去の情報($X_1, \cdots, X_{n-2}$の情報)に依らないとき、マルコフ連鎖と呼ぶ。

(6) はマルコフ連鎖である(明日!)。

2　マルコフ連鎖と混合時間

Sを有限集合とし、その元の数をdとする。S上の確率測度全体を$\mathcal{P}(S)$と置く。

$d \times d$行列$P = (P(x,y))_{x,y\in S}$が以下を満たすとき、Pを確率行列という。

$$P(x,y) \geq 0, \quad \sum_{y\in S} P(x,y) = 1, \qquad x,y \in S$$

定理 2.1 $\mu \in \mathcal{P}(S)$と確率行列$P = (P(x,y))_{x,y\in S}$が与えられたとき、次の性質を持つSに値を取る確率過程$\{X_i\}_{i=0,1,2,\cdots}$が存在する。

1. X_0の分布がμに従う。(つまり任意の$x \in S$について$\mathbb{P}(X_0 = x) = \mu(x)$。)

2. 任意の$n \geq 0$と$x, y, x_0, \cdots, x_{n-1} \in S$について、以下が成り立つ。

$$\begin{aligned}\mathbb{P}(X_{n+1} = y | X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1}, X_n = x) &= \mathbb{P}(X_{n+1} = y | X_n = x) \\ &= P(x,y).\end{aligned}$$

このような確率過程を、S上の推移確率P、初期分布μのマルコフ連鎖と呼ぶ。

前節の例では、$S=\mathcal{S}_d$、初期分布は δ_{id}、つまり確率 1 で単位置換となる分布。

X_n は (6) で定まる。推移確率は

例 1.1 の場合、(2) で決まる π_k について $\mathcal{U}_1 := \{\pi_k : 1 \le k \le d\}$ と置くと、

$$\sigma, \tau \in \mathcal{S}_d \text{ に対して、 } P(\sigma, \tau) = \begin{cases} 1/d & \text{if} \quad \tau\sigma^{-1} \in \mathcal{U}_1 \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

例 1.2 の場合、(3) で決まる τ_k について $\{\tau_k : 1 \le k \le d\}$ を $\mathcal{U}_1$ と置き、

例 1.3 の場合、(4) で決まる $\hat{\tau}_k$ について $\{\hat{\tau}_k : 1 \le k \le d\}$ を $\mathcal{U}_1$ と置くことで、

(7) と同様に推移確率を定めるとよい。

例 1.4 の場合、$\mathcal{E} := \{(\sigma,\tau) \in \mathcal{S}_d \times \mathcal{S}_d : \tau\sigma^{-1}$が互換$\}$と置くと、

$$\sigma,\tau \in \mathcal{S}_d \text{ に対して } P(\sigma,\tau) = \begin{cases} 1/d & \text{if} \quad \sigma = \tau \\ 2/2^d & \text{if} \quad (\sigma,\tau) \in \mathcal{E} \\ 0 & \text{otherwise.} \end{cases}$$

なお、$\pi \in \mathcal{S}_d$が互換であるとは、ある$1 \leq i < j \leq d$について

$\pi(i) = j, \pi(j) = i$となり、$k \neq i, j$のとき$\pi(k) = k$となること。

$$\text{例 1.5 の場合、} \sigma,\tau \in \mathcal{S}_d \text{ に対して } P(\sigma,\tau) = \begin{cases} (d+1)/2^d & \text{if} \quad \sigma = \tau \\ 1/2^d & \text{if} \quad \tau\sigma^{-1} \in \mathcal{R} \\ 0 & \text{otherwise.} \end{cases}$$

定義 2.2 *i)* $\nu = (\nu(x))_{x \in S} \in \mathcal{P}(S)$ が以下を満たすとき、ν は定常分布であるという。

$$\sum_{y \in S} \nu(y) P(y, x) = \nu(x) \qquad \forall x \in S.$$

ii) $\nu = (\nu(x))_{x \in S} \in \mathcal{P}(S)$ が以下を満たすとき、ν は可逆分布であるという。

$$\nu(x) P(x, y) = \nu(y) P(y, x) \qquad \forall x, y \in S.$$

注意 2.3 *1.* 可逆分布ならば定常分布である。

2. 例 *1.1–1.5* は、$\mathcal{S}_d$ 上の一様分布を定常分布 として持つ（$\sum_{y \in S} P(y, x) = 1$ だから）。

例 *1.1,1.4* は、$\mathcal{S}_d$ 上の一様分布を可逆分布 として持つ（$P(x, y) = P(y, x)$ だから）。

定義 2.4 $\mu, \nu \in \mathcal{P}(S)$ に対して、これらの全変動距離を以下で定める。

$$\|\mu - \nu\|_{TV} = \max_{A \subset S} |\mu(A) - \nu(A)|.$$

命題 2.5 $\|\mu-\nu\|_{TV}=\frac{1}{2}\sum_{x\in S}|\mu(x)-\nu(x)|$.

証明：$B:=\{x\in S:\mu(x)>\nu(x)\}$ とおくと、任意の $A\subset S$ に対して

$$\begin{aligned}\mu(A)-\nu(A) &= \mu(A\cap B)-\nu(A\cap B)+\mu(A\cap B^c)-\nu(A\cap B^c)\\ &\leq \mu(A\cap B)-\nu(A\cap B)\leq \mu(B)-\nu(B)\end{aligned}$$

となる（∵ B の定義から $\mu(A\cap B^c)-\nu(A\cap B^c)\leq 0,\ \mu(A^c\cap B)-\nu(A^c\cap B)\geq 0$）。

μ と ν を入れ替えて同様の議論を行うと、

$$\nu(A)-\mu(A)\leq \nu(B^c)-\mu(B^c)\Big(=1-\mu(B^c)-1+\nu(B^c)=\mu(B)-\nu(B)\Big)$$

よって

$$|\mu(A)-\nu(A)|\leq \mu(B)-\nu(B)=\frac{1}{2}\Big(\mu(B)-\nu(B)+\nu(B^c)-\mu(B^c)\Big)=\frac{1}{2}\sum_{x\in S}|\mu(x)-\nu(x)|$$

を得る。A について左辺の最大値を取ると、$\leq$ は等号になり、結論を得る。 □

以下、推移確率 P に対してその n 乗である P^n の (x,y) 成分を $P^n(x,y)$ と表す。

任意の $x,y\in S$ に対し、ある $n>0$ が存在して $P^n(x,y)>0$ となるとき、マルコフ連鎖は既約であるという。マルコフ連鎖が既約とは、どの点から出発してもいつかはすべての点を（正の確率で）通ることである。

既約かつ非周期的なマルコフ連鎖について、次の収束定理が知られている。

定理 2.6 推移確率 P を持つマルコフ連鎖が既約かつ非周期的[1]とし、ν を P の定常分布とする。このとき $\theta\in(0,1)$ と $C>0$ が存在して、以下が成り立つ。

$$\max_{x\in S}\|P^n(x,\cdot)-\nu\|_{TV}\le C\theta^n, \qquad n\ge 1.$$

系 2.7 既約なマルコフ連鎖は定常分布を唯一つ持つ。

[1] $x\in S$ に対し、$\{n\ge 1 : P^n(x,x)>0\}$ の最大公約数を、マルコフ連鎖の x における周期という。任意の $x\in S$ の周期が 1 であるとき、マルコフ連鎖は非周期的であるという。なお、既約なマルコフ連鎖の周期は $x\in S$ の取り方によらず定まることが知られている。

このように、既約かつ非周期的なマルコフ連鎖は唯一つの定常分布νを持ち、

$$\boxed{\lim_{n\to\infty} P^n(x,y) = \nu(y), \qquad x,y \in S}$$

を満たす。混合時間は、この定常分布に“ほぼ収束”するのにかかる時間である。

定義 2.8 $d(n) = \max_{x\in S} \|P^n(x,\cdot) - \nu\|_{TV}$ とおく。このとき、$\varepsilon \in (0,1]$ に対して

$$t_{mix}(\varepsilon) = \inf\{n \geq 0 : d(n) \leq \varepsilon\}, \quad t_{mix} = t_{mix}(1/4)$$

と定め、t_{mix} をこのマルコフ連鎖の混合時間という。

（注）t_{mix} の定義に出て来る $1/4$ には特段の意味はない。($t_{mix}(\varepsilon)$ を用いても、ε に応じた定数倍の違いであることが知られている。)

混合時間が正確に求められるマルコフ連鎖は限定的$\Rightarrow$ 混合時間の評価が必要。

(1) 上からの評価：

確率分布や確率過程のカップリング

(X, Y)が、S上の二つの確率分布μ, νのカップリングであるとは、

- (X, Y)が一つの確率空間（その確率分布を$\tilde{P}$と書く）に定義された確率変数、
- X, Yの周辺分布がそれぞれμ, ν

  つまり$\tilde{P}(X = x) = \mu(x), \tilde{P}(Y = y) = \nu(y)$が任意の$x, y \in S$で成立。

二つのマルコフ連鎖X_n, Y_nに対し、$(\tilde{X}_n, \tilde{Y}_n)$が$X_n, Y_n$のカップリング、かつ$S \times S$上のマルコフ連鎖のとき、$(\tilde{X}_n, \tilde{Y}_n)$を$X_n, Y_n$のマルコビアン・カップリングという。

このとき、

$$\tau_{\text{couple}} := \min\{n \geq 0 : \tilde{X}_n = \tilde{Y}_n\}$$

（つまり初めて $\tilde{X}_n = \tilde{Y}_n$ となった時刻）を、カップリングタイムという。

（注）$n \geq \tau_{\text{couple}}$ では $\tilde{X}_n = \tilde{Y}_n$ となる（つまり一旦粒子が出会ったら、それ以降はくっついて動く）ものを考えることが多い。

命題 2.9 X_n, Y_n を、ともに推移確率が P であるマルコフ連鎖とし、$(\tilde{X}_n, \tilde{Y}_n)$ を X_n, Y_n のマルコビアン・カップリングとする。このとき

$$\|P^n(x,\cdot) - P^n(y,\cdot)\|_{TV} \leq \tilde{P}_{(x,y)}(\tau_{\text{couple}} > n) \tag{8}$$

が成り立つ（$\tilde{P}_{(x,y)}$ は、$(\tilde{X}_0, \tilde{Y}_0) = (x, y)$ を意味する）。特に以下が成り立つ。

$$d(n) \leq \max_{x,y \in S} \tilde{P}_{(x,y)}(\tau_{\text{couple}} > n). \tag{9}$$

証明： $(\tilde{X}_n, \tilde{Y}_n)$ が X_n, Y_n のカップリングだから

$$
\begin{aligned}
P^n(x,z) - P^n(y,z) &= \tilde{P}_{(x,y)}(\tilde{X}_n = z) - \tilde{P}_{(x,y)}(\tilde{Y}_n = z) \\
&= \tilde{P}_{(x,y)}(\tilde{X}_n = z, \tau_{\text{couple}} > n) + \tilde{P}_{(x,y)}(\tilde{X}_n = z, \tau_{\text{couple}} \leq n) \\
&\quad - \tilde{P}_{(x,y)}(\tilde{Y}_n = z, \tau_{\text{couple}} > n) - \tilde{P}_{(x,y)}(\tilde{Y}_n = z, \tau_{\text{couple}} \leq n) \\
&= \tilde{P}_{(x,y)}(\tilde{X}_n = z, \tau_{\text{couple}} > n) - \tilde{P}_{(x,y)}(\tilde{Y}_n = z, \tau_{\text{couple}} > n).
\end{aligned}
$$

よって

$$
|P^n(x,z) - P^n(y,z)| \leq \tilde{P}_{(x,y)}(\tilde{X}_n = z, \tau_{\text{couple}} > n) + \tilde{P}_{(x,y)}(\tilde{Y}_n = z, \tau_{\text{couple}} > n)
$$

となり、z について和を取り 2 で割ることにより命題2.5から (8) を得る。さらに

$$
|P^n(x,A) - \nu(A)| = |P^n(x,A)\sum_{y\in S}\nu(y) - \sum_{y\in S}\nu(y)P^n(y,A)| \leq \sum_{y\in S}\nu(y)|P^n(x,A) - P^n(y,A)|
$$

に注意して、(8) を x, y について max を取ることにより、(9) を得る。 $\square$

よって、混合時間の上からの評価を出すには、$\tilde{P}_{(x,y)}(\tau_{\mathrm{couple}} > n)$ を上から評価すればよいことが分かる。

(2) 下からの評価：　ここでは等周定数を用いた下からの評価を紹介する。

定義 2.10 S とその上のマルコフ連鎖 (P, ν) に対して、S の等周定数 Φ^* を

$$\Phi^* := \min_{D \subset S, \nu(D) \leq 1/2} \frac{\sum_{x \in D} \sum_{y \in D^c} \nu(x) P(x,y)}{\nu(D)}$$

で定義する。

命題 2.11 S 上のマルコフ連鎖 (P, ν) について、以下が成り立つ。

$$t_{mix} \geq \frac{1}{4\Phi^*}.$$

3　トランプは何回きればよいか？（カット・オフ現象）

いよいよ、リフル・シャッフルの混合時間を計算することにより、「7回すればよく混ざる」の真相に迫る。「よく混ざった状態」とは、数学的には定常状態を指すので、今の場合一様分布（つまり$\mathcal{S}_d$の各々の元が等確率で現れる分布、以下Uと書く）を指す。混ざり具合の尺度は、全変動距離を用い、初期分布は、確率1で単位置換を与える分布（以下δ_{id}と書く）とする。

すると、n回のシャッフルにおける「混ざり具合」は以下の量で表せる。

$$\|P^n(\delta_{id},\cdot)-U\|_{TV}=\frac{1}{2}\sum_{\pi\in\mathcal{S}_d}\left|P^n(\delta_{id},\pi)-\frac{1}{d!}\right|. \tag{10}$$

ここで、等号は命題2.5による。

$P^n(\delta_{id},\pi)$をどう計算するか？

そのため、多少天下り的だが$a\in\mathbb{N}$に対してa-シャッフルという「切り方」を定義する（リフル・シャッフルは、2-シャッフルに当たる）。