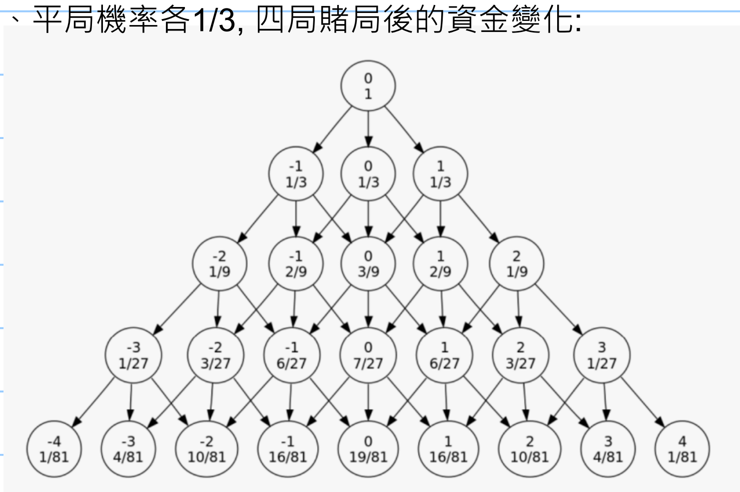
# Applications of Markov Chain

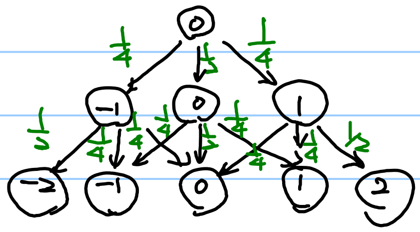
0.Probabilities and Matrices 機率與矩陣

輸、贏、平局機率各1/3, 四局賭局後的資金變化:




如果每次輸、贏、平局的機率都不相同，機率該如何計算? 使用轉移機率矩陣。

e.g. 兩次賭局



-2 :

-1 :

0 :

1 :

2 :

用矩陣 :

$$P_1 = \begin{array}{c c} & \begin{array}{ccc} -1 & 0 & 1 \end{array} \\ 0 & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix} \end{array}$$

$$P_2 = \begin{array}{c c} & \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} \\ \begin{array}{c} -1 \\ 0 \\ 1 \end{array} & \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \end{array}$$

$P_1P_2 =$

所有變化可表示成樹形圖的機率均可用矩陣計算。

# 1. Introduction to Markov Chain

## Andrei A. Markov (1856 – 1922)

- A stochastic process is a family of random variables $X_t$, where $t$ is a parameter running over a suitable index set $T$.

$T = \{0, 1, 2, \ldots\}$ 離散時間隨機過程

$T = [0, \infty)$ 連續 〃

$X_t$ 之可能值所成之集合稱為 state space (狀態空間) 連續 / 離散

Remark: 1. $X_t$ 有時亦寫做 $X(t)$

2. $T$ 也有其他選擇，視問題而定

- A *Markov process* $\{X_t\}$ is a stochastic process with the property that, given the value of $X_t$, the values of $X_s$ for $s > t$ are not influenced by the values of $X_u$ for $u < t$. In words,

$$P(X_s \in A \mid X_u,\ u \le t) = P(X_s \in A \mid X_t)$$

簡單來說，Markov process 的未來只和最後狀態有關，而和其他歷史紀錄無關

- A *discrete-time Markov chain* is a Markov process whose state space is a finite or countable set, and whose (time) index set is $T = (0, 1, 2, \ldots)$.

i.e.

$$\begin{aligned}\Pr\{X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i\} \\ = \Pr\{X_{n+1} = j \mid X_n = i\}\end{aligned}$$

for all time points $n$ and all states $i_0, \ldots, i_{n-1}, i, j$.

Remark: 1. Discrete-time Markov chain 通常简称 Markov chain

2. 一般情况下, Markov chain 之 state space 設为 $\{0,1,2,\dots\}$ 或 $\{0,1,2,\dots,N\}$
前者稱 infinite state space, 後者稱 finite state space
但其數字通常僅为状態編号而不具其他意义
如 $X_4=5$ 代表在時间 4 時, 隨机过程之状態为第 5 种状態

- $P_{ij}^{n,n+1} = \Pr\{X_{n+1}=j|X_n=i\}$. 稱为 one-step transition probability (一步轉移机率)

若其值与 n 無关, 则稱該 Markov chain 有 stationary (靜態) transition probability
此時因 $P_{ij}^{n,n+1}$ 与 n 無关, 故可改写为 $P_{ij}$, 代表由 i 下一步移至 j 之机率.

- 可將 $P_{ij}$ 整体表示成矩陣形式如下, 稱为 Markov matrix 或 transition probability matrix
符号記为 $\mathbf{P} = \|P_{ij}\|$ (或 $\mathbb{P}=[P_{ij}]$)。

$$\mathbf{P} = \begin{Vmatrix} P_{00} & P_{01} & P_{02} & P_{03} & \cdots \\ P_{10} & P_{11} & P_{12} & P_{13} & \cdots \\ P_{20} & P_{21} & P_{22} & P_{23} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ P_{i0} & P_{i1} & P_{i2} & P_{i3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}$$

(當 finite state space 時 $\mathbb{P}$ 即为有限矩陣)

- n步轉移機率:

令 $P_{ij}^{(n)}$ 为由 i 出發, n 步之後走到 j 之机率. 則 $P_{ij}^{(n)} = P^n$ 之 i,j 項.

Example: A Markov chain $\{X_n\}$ on the states 0, 1, 2 has the transition probability matrix

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{Vmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{Vmatrix} \end{array}.$$

(a) Compute the two-step transition matrix $P^2$.
(b) What is $\Pr\{X_3=1|X_1=0\}$?
(c) What is $\Pr\{X_3=1|X_0=0\}$?

HW1: A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}\begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left\|\begin{array}{ccc} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.1 & 0.4 \\ 0.5 & 0.2 & 0.3 \end{array}\right\| \end{array}$$

Find all $P(X_5 = i \mid X_0 = j)$ for $i, j = 0, 1, 2$ (用電腦算)

2. Simple Applications

Application 1: 自然機率--賭博性電玩

Application 2: Algorithmic Music Composition

USSC course: Generative Design 生成式設計

https://canvas.ucsc.edu/courses/26749 assignment 5

https://junshern.github.io/algorithmic-music-tutorial/

Application 3: Baseball run distribution analysis (非靜態馬可夫鏈)

Ursin, Daniel Joseph, "A Markov Model for Baseball with Applications" (2014). Theses and Dissertations . 964. https://dc.uwm.edu/etd/964

Bukiet, B., Harold, E., Palacios, J. A Markov chain approach to baseball Operations Research 45.1 (1997): 14-23

Note: 講義裡只有這部分應用是非靜態馬可夫鏈，其餘都是靜態。

應用: 球隊最佳棒次安排、運動彩券簽注....

每局壘上及出局狀況共24種可能性:


*Figure 1: State space of the Markov chain; states are labeled in (B,O) format.*

可能變化示意圖:

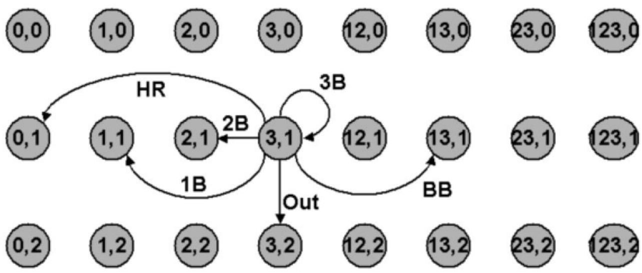



*Figure 2: Potential transitions from state (3,1) using simplified baseball event model.*

每位打者的Transition matrix: 25x25 矩陣

$$
\begin{array}{cc}
 & \begin{array}{cccc} \text{無人出局} & \text{1出局} & \text{2出局} & \text{單局結束} \end{array} \\
\mathrm{P} = & \begin{pmatrix} A_0 & B_0 & C_0 & D_0 \\ \mathbf{0} & A_1 & B_1 & E_1 \\ \mathbf{0} & \mathbf{0} & A_2 & F_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix} \begin{array}{l} \text{無人出局} \\ \text{1出局} \\ \text{2出局} \\ \text{單局結束} \end{array}
\end{array}
$$

為簡化問題，忽略下列可能: 1. 盜壘 2. 打者出局但壘上跑者進壘

A matrix: 打擊結果成功(含保送) 8x8

$$
\begin{array}{c}
\\ (O,0) \\ (O,1) \\ (O,2) \\ (O,3) \\ (O,12) \\ (O,13) \\ (O,23) \\ (O,123)
\end{array}
\begin{array}{c}
\begin{array}{cccccccc}
(O,0) & (O,1) & (O,2) & (O,3) & (O,12) & (O,13) & (O,23) & (O,123)
\end{array} \\
\begin{pmatrix}
p_{HR} & (p_{1B}+p_{BB}) & p_{2B} & p_{3B} & 0 & 0 & 0 & 0 \\
p_{HR} & 0 & 0 & p_{3B} & (p_{1B}+p_{BB}) & 0 & p_{2B} & 0 \\
p_{HR} & p_{1B} & p_{2B} & p_{3B} & p_{BB} & 0 & 0 & 0 \\
p_{HR} & p_{1B} & p_{2B} & p_{3B} & 0 & p_{BB} & 0 & 0 \\
p_{HR} & 0 & 0 & p_{3B} & p_{1B} & 0 & p_{2B} & p_{BB} \\
p_{HR} & 0 & 0 & p_{3B} & p_{1B} & 0 & p_{2B} & p_{BB} \\
p_{HR} & p_{1B} & p_{2B} & p_{3B} & 0 & 0 & 0 & p_{BB} \\
p_{HR} & 0 & 0 & p_{3B} & p_{1B} & 0 & p_{2B} & p_{BB}
\end{pmatrix}
\end{array}
$$

B matrix: 打擊結果增加一出局數 8x8

$$
B = p_{Out} \cdot I_8 = \begin{bmatrix}
p_{out} & 0 & \cdots & 0 \\
0 & p_{out} & 0 & \vdots \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & p_{out}
\end{bmatrix}_{8\times 8}
$$

C matrix: 打擊結果增加二出局數 8x8

$$
\begin{array}{c}
\\ (0,0) \\ (0,1) \\ (0,2) \\ (0,3) \\ (0,12) \\ (0,13) \\ (0,23) \\ (0,123)
\end{array}
\begin{array}{c}
\begin{array}{cccccccc}
(2,0) & (2,1) & (2,2) & (2,3) & (2,12) & (2,13) & (2,23) & (2,123)
\end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
p_{DP} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
p_{DP} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
p_{DP} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

D matrix: 無人出局==>三人出局

$$
D = p_{TP} \cdot < 0,0,0,0,1,1,1,1 >^T \approx
\begin{array}{c}
(O,0) \\ (O,1) \\ (O,2) \\ (O,3) \\ (O,12) \\ (O,13) \\ (O,23) \\ (O,123)
\end{array}
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ p_{TP} \\ p_{TP} \\ p_{TP} \\ p_{TP}
\end{bmatrix}
$$

E matrix: 一人出局==>三人出局

$$E = p_{DP} \cdot < 0,1,1,1,1,1,1,1 >^T$$

1

$(0,0)\quad(0,1)\quad(0,2)\quad(0,3)\quad(0,12)\quad(0,13)\quad(0,23)\quad(0,123)$

F matrix: 二人出局==>三人出局

$$F = p_{Out} \cdot < 1,1,1,1,1,1,1,1 >^T$$

2

$(0,0)\quad(0,1)\quad(0,2)\quad(0,3)\quad(0,12)\quad(0,13)\quad(0,23)\quad(0,123)$

紀錄得分:

設定Un矩陣紀錄第n次打擊後得分可能機率，假設每局得分最多20分(歷史經驗)，故單局得分的U矩陣為21x25 matrx, 其中列為得分數，行為當時壘上狀態。

Un計算公式為

$$\begin{aligned}\mathbf{U}_{n+1}(\text{row } j) &= \mathbf{U}_n(\text{row } j)\mathbf{P0} + \mathbf{U}_n(\text{row } j-1)\mathbf{P1} \\ &+ \mathbf{U}_n(\text{row } j-2)\mathbf{P2} + \mathbf{U}_n(\text{row } j-3)\mathbf{P3} \\ &+ \mathbf{U}_n(\text{row } j-4)\mathbf{P4}.\end{aligned}$$

其中P0,P1,P2,P3,P4 分別為P矩陣中該次打擊結果可得0,1,2,3,4分的機率
如 P4矩陣為25x25矩陣，在(8,1),(16,9),(24,17)三項為 $p_{HR}$ ,其餘為0.

$\underbrace{\mathbf{P}_1\mathbf{P}_2\mathbf{P}_3\ldots\mathbf{P}_9\mathbf{P}_1\mathbf{P}_2\ldots.}_{\text{共n項}}$ 之第一列即為n次打擊之後壘上狀態之機率分布

Un 之第25 行即為n次打擊前該局已經結束且得分0~20分的機率分布，由於有正機率第n次打擊時該局仍未結束，故上述0~20分機率總和小於1, 程式執行至球局為結束的機率小於 0.0001時結束。

若欲擴充至九局，則只需把Un 變成九倍 189x25 矩陣

分析結果:

季賽平均得分分布預測堪稱精準，但預測單局勝負不理想

TABLE 1. 2013 National League Run Production

| National League | | | |
|---|---|---|---|
| Team | Actual RPG | Simulated RPG | Difference |
| Washington | 4.056 | 4.275 | 0.219 |
| Atlanta | 4.253 | 4.123 | -0.13 |
| Miami | 3.173 | 3.035 | -0.138 |
| NY Mets | 3.827 | 3.662 | -0.165 |
| Philadelphia | 3.772 | 3.756 | -0.015 |
| Milwaukee | 3.957 | 3.704 | -0.252 |
| St. Louis | 4.84 | 4.051 | -0.788 |
| Pittsburgh | 3.92 | 3.867 | -0.053 |
| Cincinnati | 4.315 | 3.86 | -0.454 |
| Chicago Cubs | 3.698 | 3.041 | -0.657 |
| LA Dodgers | 4.012 | 4.391 | 0.378 |
| San Francisco | 3.889 | 3.966 | 0.077 |
| San Diego | 3.821 | 3.657 | -0.164 |
| Arizona | 4.235 | 3.911 | -0.324 |
| Colorado | 4.364 | 4.515 | 0.151 |
| AVERAGE | 4.009 | 3.854 | -0.154 |
| STDEV | 0.362 | 0.406 | 0.304 |

TABLE 2. 2013 American League Run Production

| American League | | | |
|---|---|---|---|
| Team | Actual RPG | Simulated RPG | Difference |
| Baltimore | 4.605 | 4.106 | -0.499 |
| NY Yankees | 4.019 | 3.469 | -0.55 |
| Toronto | 4.401 | 4.123 | -0.278 |
| Tampa Bay | 4.301 | 3.989 | -0.311 |
| Boston | 5.272 | 4.548 | -0.723 |
| Kansas City | 4.006 | 3.519 | -0.487 |
| Detroit | 4.92 | 4.761 | -0.159 |
| Cleveland | 4.605 | 3.796 | -0.809 |
| Chicago Sox | 3.722 | 3.527 | -0.195 |
| Minnesota | 3.796 | 3.675 | -0.121 |
| LA Angels | 4.531 | 4.091 | -0.44 |
| Oakland | 4.741 | 4.086 | -0.655 |
| Seattle | 3.858 | 3.907 | 0.049 |
| Houston | 3.772 | 3.529 | -0.243 |
| Texas | 4.485 | 3.821 | -0.663 |
| AVERAGE | 4.336 | 3.929 | -0.406 |
| STDEV | 0.448 | 0.365 | 0.241 |

FIGURE 3. AL - 2013 Baltimore and Boston Run Distributions. All run outcomes are discrete.

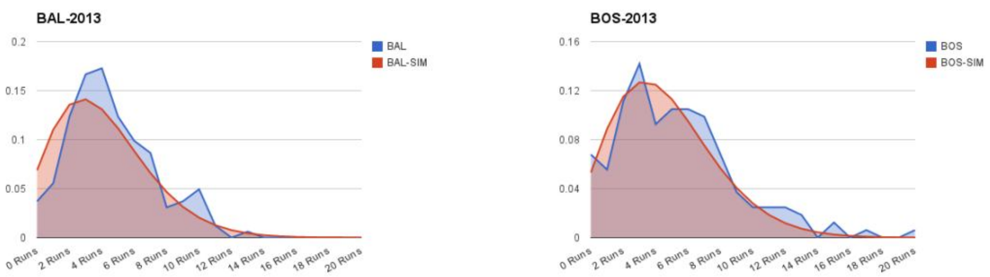



**Table I**

Best and Worst Orders for the 1989 Atlanta Braves (as calculated by the single and double-placement models and by full enumeration)

| | | Best Orders | | | Worst Order | | |
|---|---|---|---|---|---|---|---|
| Players | Scoring Index | Single Placement | Double Placement | True | Single Placement | Double Placement | True |
| Smith | 0.887 | Smith | Murphy | Perry | Pitcher | Pitcher | Pitcher |
| McDowell | 0.649 | McDowell | Perry | Smith | Murphy | Murphy | Treadway |
| Blauser | 0.484 | Perry | Treadway | McDowell | Treadway | Treadway | Perry |
| Treadway | 0.430 | Blauser | Smith | Blauser | Perry | Perry | Davis |
| Perry | 0.416 | Treadway | McDowell | Treadway | Davis | Davis | Thomas |
| Murphy | 0.411 | Murphy | Blauser | Murphy | Thomas | Thomas | Blauser |
| Thomas | 0.241 | Thomas | Thomas | Thomas | Blauser | Blauser | McDowell |
| Davis | 0.198 | Davis | Davis | Pitcher | McDowell | McDowell | Murphy |
| Pitcher | 0.078 | Pitcher | Pitcher | Davis | Smith | Smith | Smith |
| Expected Runs Per Game | | 3.523 | 3.539 | 3.557 | 3.264 | 3.264 | 3.260 |
| Win-Loss | | 85.9–76.1 | 86.1–75.9 | 86.5–75.5 | 76.1–85.9 | 75.9–86.1 | 75.5–86.5 |

**Table III**
Standings Predicted by a Non-identical Batter Model for the 1989 National League

| National League East | | |
|---|---|---|
| Team | Actual Games Won | Model Games Won |
| Cubs | 93 | 86.9 |
| Mets | 87 | 88.6 |
| Cardinals | 86 | 87.4 |
| Expos | 81 | 82.6 |
| Pirates | 74 | 78.4 |
| Phillies | 67 | 67.9 |

| National League West | | |
|---|---|---|
| Team | Actual Games Won | Model Games Won |
| Giants | 92 | 90.7 |
| Padres | 89 | 84.0 |
| Astros | 86 | 72.8 |
| Dodgers | 77 | 73.7 |
| Reds | 75 | 79.5 |
| Braves | 73 | 79.6 |

**Table V**
Expected Standings in the National League East in 1989 if Darryl Strawberry of the Mets Were Traded for Milt Thompson of the Cardinals

| National League East | | |
|---|---|---|
| Team | Pre-Trade Model Games Won | Post-Trade Model Games Won |
| Cubs | 86.9 | 87.0 |
| Mets | 88.6 | 86.9 |
| Cardinals | 87.4 | 88.9 |
| Expos | 82.6 | 82.6 |
| Pirates | 78.4 | 78.4 |
| Phillies | 67.9 | 67.9 |

## Application 4: 文字學

- *Random characters.*
  XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD
- *Sample from* $P^{(1)}$.
  OCRO HLI RGWR NMIELWIS EU LL NBBESEBYA TH EEI ALHENHTTPA OO BTTV
- *Sample from* $P^{(2)}$.
  ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE FUSO TIZIN ANDY TOBE SEACE CTISBE
- *Sample from* $P^{(3)}$.
  IN NO IST LAY WHEY CRATICT FROURE BERS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE
- *Sample from* $P^{(4)}$.
  THE GENERATED JOB PROVIDUAL BETTER TRAND THE DISPLAYED CODE ABOVERY UPONDULTS WELL THE CODERST IN THESTICAL IT TO HOCK BOTHE

3. First Step Analysis

分析有限時間內會停止的靜態馬可夫鍊在停止前的行為

以三個states的簡單例子說明

$$\mathbf{P} = \begin{array}{c|ccc|} & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 1 & \alpha & \beta & \gamma \\ 2 & 0 & 0 & 1 \end{array},$$

where $\alpha > 0, \beta > 0, \gamma > 0$, and $\alpha + \beta + \gamma = 1$.

馬可夫鍊 Xn最終停在0或2 (此時稱absorption)

Q: 1. 最終停在0,2機率各為多少?

2. 平均多久遊戲結束(到達0,2)?

假設X0=1 (其他情形結果顯而易見)

定義

$$T = \min\{n \geq 0; X_n = 0 \quad \text{or} \quad X_n = 2\}$$

則Q1,Q2可改寫為計算

1. $u = \Pr\{X_T = 0 | X_0 = 1\}$

2. $v = E[T | X_0 = 1].$

1、 $u = \Pr\{X_T = 0 | X_0 = 1\}$

$$= \sum_{k=0}^{2} \Pr\{X_T = 0 | X_0 = 1, X_1 = k\} \Pr\{X_1 = k | X_0 = 1\}$$

$$= \sum_{k=0}^{2} \Pr\{X_T = 0 | X_1 = k\} \Pr\{X_1 = k | X_0 = 1\}$$

(by the Markov property)

$$= 1(\alpha) + u(\beta) + 0(\gamma).$$

$$\Longrightarrow \quad u = \frac{\alpha}{1-\beta} = \frac{\alpha}{\alpha+\gamma}.$$

2. $v = 1 + \alpha(0) + \beta(v) + \gamma(0)$
$\quad = 1 + \beta v,$ $\qquad \Rightarrow \quad v = \dfrac{1}{1-\beta}.$

• 4 states :

$$\mathbf{P} = \begin{array}{c|cccc|} & 0 & 1 & 2 & 3 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 1 & P_{10} & P_{11} & P_{12} & P_{13} \\ 2 & P_{20} & P_{21} & P_{22} & P_{23} \\ 3 & 0 & 0 & 0 & 1 \end{array}.$$

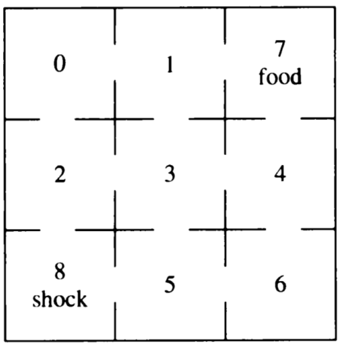
令

$$T = \min\{n \geq 0; X_n = 0 \quad \text{or} \quad X_n = 3\},$$
$$u_i = \Pr\{X_T = 0 | X_0 = i\} \quad \text{for } i = 1, 2,$$
$$v_i = E[T | X_0 = i] \quad \text{for } i = 1, 2.$$

$$\begin{cases} u_1 = P_{10} + P_{11}u_1 + P_{12}u_2. \\ u_2 = P_{20} + P_{21}u_1 + P_{22}u_2. \end{cases}$$ $\Rightarrow$ 解联立.可得 $u_1 . u_2$

$$\begin{cases} v_1 = 1 + P_{11}v_1 + P_{12}v_2, \\ v_2 = 1 + P_{21}v_1 + P_{22}v_2. \end{cases}$$ $\Rightarrow$ 解联立.可得 $v_1 . v_2$

HW: *A Maze* A white rat is put into the maze shown:




假設白老鼠被放置在房間3, 每個門被選擇的機率相同。得到食物或電擊則實驗結束。

Q: 1. 寫出transition probability matrix

2. 實驗結束時被電擊及得到食物的機率各為多少? (列出方程組並用電腦求解)

3. 平均需經過幾個房間實驗才能結束?

- Gambler's ruin problem (賭徒破產問題):

現考慮下列 transition probability matrix 之 Markov Chain：

$$
\mathbf{P} = \begin{array}{c|cccccc|}
 & 0 & 1 & 2 & 3 & \cdots & N \\
\hline
0 & 1 & 0 & 0 & 0 & \cdots & 0 \\
1 & q & 0 & p & 0 & \cdots & 0 \\
2 & 0 & q & 0 & p & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
N & 0 & 0 & 0 & 0 & \cdots & 1
\end{array}.
$$

定義 $T = \min\{n \geq 0; X_n = 0 \text{ or } X_n = N\}$.

此類 r.v. 稱為 hitting time，此處 hitting time 為第一次碰觸 0 或 N 之時間)

可証明最終必碰觸 0 或 N 而停滯於此。

- Q: 最終碰觸 0 或 N 之機率各為多少？ ← Gambler's ruin problem

令 $u_k = \Pr\{X_T = 0 | X_0 = k\}$.

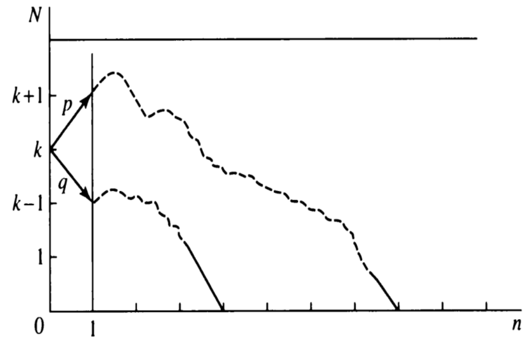



**Figure 3.4** First step analysis for the gambler's ruin problem.

由 1st step analysis：

$$u_k = pu_{k+1} + qu_{k-1}, \quad \text{for } k = 1, \ldots, N-1, \qquad \text{— (3.46)}$$

且已知 boundary condition

$$u_0 = 1, \quad u_N = 0.$$

以此解 $u_k$：

令 $x_k = u_k - u_{k-1}$ ， $k = 1, 2, \ldots, N$

Using $p + q = 1$ to write $u_k = (p+q)u_k = pu_k + qu_k$, —(*)

則 (3.46) —(*) 得

$$
\begin{aligned}
k = 1; &\quad 0 = p(u_2 - u_1) - q(u_1 - u_0) = px_2 - qx_1; \\
k = 2; &\quad 0 = p(u_3 - u_2) - q(u_2 - u_1) = px_3 - qx_2; \\
k = 3; &\quad 0 = p(u_4 - u_3) - q(u_3 - u_2) = px_4 - qx_3; \\
&\quad \vdots \\
k = N-1; &\quad 0 = p(u_N - u_{N-1}) - q(u_{N-1} - u_{N-2}) = px_N - qx_{N-1};
\end{aligned}
$$

$$
\Rightarrow \quad \begin{aligned}
x_2 &= (q/p)x_1, \\
x_3 &= (q/p)x_2 = (q/p)^2 x_1, \\
x_4 &= (q/p)x_3 = (q/p)^3 x_1 \\
&\vdots \\
x_k &= (q/p)x_{k-1} = (q/p)^{k-1} x_1, \\
&\vdots \\
x_N &= (q/p)x_{N-1} = (q/p)^{N-1} x_1.
\end{aligned}
$$

又由 $x_k$ 之定义，可知

$$
\begin{aligned}
u_k &= 1 + x_1 + x_2 + \cdots + x_k \\
&= 1 + x_1 + (q/p)x_1 + \cdots + (q/p)^{k-1} x_1 \\
&= 1 + [1 + (q/p) + \cdots + (q/p)^{k-1}] x_1,
\end{aligned} \tag{3.47}
$$

仍需解 $x_1$，但已知 $u_N = 0$，代入上式得

$$
0 = 1 + [1 + (q/p) + \cdots + (q/p)^{N-1}] x_1,
$$

$$
\Rightarrow \quad x_1 = -\frac{1}{1 + (q/p) + \cdots + (q/p)^{N-1}},
$$

代回 (3.47) 得

$$
u_k = 1 - \frac{1 + (q/p) + \cdots + (q/p)^{k-1}}{1 + (q/p) + \cdots + (q/p)^{N-1}}.
$$

$$
\left( 1 + (q/p) + \cdots + (q/p)^{k-1} = \begin{cases} k & \text{if } p = q = \frac{1}{2}, \\ \dfrac{1 - (q/p)^k}{1 - (q/p)} & \text{if } p \neq q, \end{cases} \right)
$$

故得解

$$
u_k = \begin{cases} 1 - (k/N) = (N-k)/N & \text{when } p = q = \dfrac{1}{2}, \\ 1 - \dfrac{1 - (q/p)^k}{1 - (q/p)^N} = \dfrac{(q/p)^k - (q/p)^N}{1 - (q/p)^N} & \text{when } p \neq q. \end{cases} \tag{3.48}
$$

$\Pr(X_T = N \mid X_0 = k) = 1 - u_k$

- 類似方法可計算 mean duration

$$v_i = E[T|X_0 = i].$$

由 1st step analysis

$$v_i = 1 + pv_{i+1} + qv_{i-1} \quad \text{for } i = 1, \ldots, N-1. \tag{3.50}$$

又已知 $v_0 = 0$, $v_N = 0$.

以下計算 $p = q = 1/2$ 之情況. 其他 p 值可用類似方法求得。

令 $x_k = v_k - v_{k-1}$ for $k = 1, \ldots, N$,

則

$$\begin{aligned}
k &= 1; \quad -1 = \frac{1}{2}(v_2 - v_1) - \frac{1}{2}(v_1 - v_0) = \frac{1}{2}x_2 - \frac{1}{2}x_1;\\
k &= 2; \quad -1 = \frac{1}{2}(v_3 - v_2) - \frac{1}{2}(v_2 - v_1) = \frac{1}{2}x_3 - \frac{1}{2}x_2;\\
k &= 3; \quad -1 = \frac{1}{2}(v_4 - v_3) - \frac{1}{2}(v_3 - v_2) = \frac{1}{2}x_4 - \frac{1}{2}x_3;\\
&\vdots\\
k &= N-1; \quad -1 = \frac{1}{2}(v_N - v_{N-1}) - \frac{1}{2}(v_{N-1} - v_{N-2}) = \frac{1}{2}x_N - \frac{1}{2}x_{N-1}.
\end{aligned}$$

$$\Rightarrow \quad x_k = x_1 - 2(k-1) \text{ for } k = 2, 3, \ldots, N.$$

$$\begin{aligned}
v_k &= x_1 + \cdots + x_k\\
&= kv_1 - 2[1 + 2 + \cdots + (k-1)] = kv_1 - k(k-1),
\end{aligned} \tag{3.51}$$

又 $v_N = 0$ to obtain $0 = Nv_1 - N(N-1)$ or $v_1 = (N-1)$. 代回上式得

$$v_k = k(N-k), \quad k = 0, 1, \ldots, N, \tag{3.52}$$

- Application: Cash Management 現金管理

現金支付未來支出，但現金太多即為閒滯（不能生利息、投資等）
希望能儘量滿足支付帳款，同時儘量減少現金量

令 $X_n$ 為時間 $n$ 之現金量.

假設 $$\Pr\{X_{n+1}=k\pm 1|X_n=k\}=\frac{1}{2}.$$ ← 將金額和時間離散化

財務長欲調整現金流. 令 $s$ and $\mathcal{S}$ 為滿足 $s<\mathcal{S}$ 之兩參數.
當現金為 0，賣國庫券將現金流補至 $s$.
: $\mathcal{S}$，買 : 減至 $s$

可看出 $X_n$ 之行為為一串統計上類似行為的 cycle 所組成

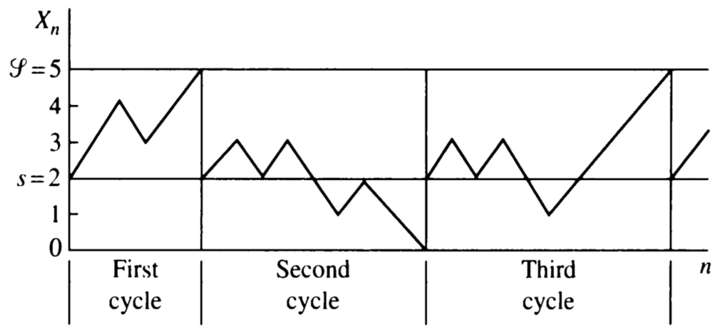



**Figure 3.5** Several typical cycles in a cash inventory model.

欲計算 ① mean length of a cycle
② mean total unit periods of cash on hand during a cycle
一個循環內現金*時間總和

③ 並用上述二量來衡量 long run performance of the model.

令 T = 第一次現金流為 0 或 $\mathcal{S}$ 之時間

① 令 $v_s=E[T|X_0=s]$

則由 (3.52)

$$v_s=s(\mathcal{S}-s). \tag{3.57}$$

② 令 $W_{sk}$ 為 $X_0=s$ 條件下，時間 T 之前，$X_n$ 經過狀態 k 之期望次數

則可由 First Step Analysis 算出（細節略）

$$W_{sk} = 2\left[\frac{s}{\mathcal{S}}(\mathcal{S}-k) - (s-k)^+\right]. \tag{3.58}$$

則

$$\begin{aligned} W_s &= \sum_{k=1}^{\mathcal{S}-1} kW_{sk} \\ &= 2\left\{\frac{s}{\mathcal{S}}\sum_{k=1}^{\mathcal{S}-1} k(\mathcal{S}-k) - \sum_{k=1}^{s-1} k(s-k)\right\} \\ &= 2\left\{\frac{s}{\mathcal{S}}\left[\frac{\mathcal{S}(\mathcal{S}-1)(\mathcal{S}+1)}{6}\right] - \frac{s(s-1)(s+1)}{6}\right\}^* \\ &= \frac{s}{3}\left[\mathcal{S}^2 - s^2\right]. \end{aligned} \tag{3.59}$$

$\rightarrow$ $\sum_{k=1}^{a-1} k(a-k) = \frac{1}{6}a(a+1)(a-1).$

即為 ② 所求

③ 可看出每個 cycle 為 i.i.d.,

令 K 為每次現金流調整至 s 所需之處理費.

$T_i$ : $i^{th}$ cycle 長度

$R_i$ : $i^{th}$ cycle 之機會成本

則前 n 個 cycle 之平均單位時間成本為

$$\text{Average cost} = \frac{nK + R_1 + \cdots + R_n}{T_1 + \cdots + T_n}.$$

令 $n \to \infty$. 則由大數法則可得

$$\text{Long run average cost} = \frac{K + E[R_i]}{E[T_i]}.$$

令 r 為每單位時間，單位金額現金流之機會成本. 則 $ER_i = rW_s$

又 $ET_i = \nu_s$ 代入上式得

$$\text{Long run average cost} = \frac{K + (1/3)rs\left(\mathcal{S}^2 - s^2\right)}{s(\mathcal{S}-s)}. \tag{3.60}$$

$$= \frac{K + (1/3)r\mathcal{S}^3 x\left(1-x^2\right)}{\mathcal{S}^2 x(1-x)},$$

其中 $x = s/\mathcal{S}$

由

$$\frac{\mathrm{d(average\ cost)}}{\mathrm{d}x} = 0 = -\frac{K(1-2x)}{\mathcal{S}^2 x^2 (1-x)^2} + \frac{1}{3} r\mathcal{S},$$

$$\frac{\mathrm{d(average\ cost)}}{\mathrm{d}\mathcal{S}} = 0 = -\frac{2K}{\mathcal{S}^3 x(1-x)} + \frac{r(1+x)}{3},$$

可得

$$x_{\mathrm{opt}} = \frac{1}{3} \quad \text{and} \quad \mathcal{S}_{\mathrm{opt}} = 3s_{\mathrm{opt}} = 3\sqrt[3]{\frac{3K}{4r}}.$$

- Matrix Form result of First Step Analysis

state $0, 1, \ldots, r-1$ transient

$r, r+1, \ldots, N$ absorbing

則 transition probability matrix 可寫做

$$\mathbf{P} = \begin{Vmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{Vmatrix},$$

$$U = (I-Q)^{-1} R$$

$$\vec{v} = \begin{pmatrix} v_0 \\ \vdots \\ v_{r-1} \end{pmatrix} = (I-Q)^{-1} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{r\times 1}$$

## 4. The long run behavior of Markov Chains

Theorem: 在適當條件之下，

(a) $\lim_{n\to\infty} P_{ii}^{(n)} = 1/$由 $i$ 出發，回到 $i$ 所需時間的期望值

(b) $\lim_{n\to\infty} P_{ji}^{(n)} = \lim_{n\to\infty} P_{ii}^{(n)}$ for any $i, j$

(c) 若令 $\pi_j = \lim_{n\to\infty} P_{ii}^{(n)}$，則 $\pi_j$ 為下列方程組的唯一解：

$$\begin{cases} \pi_j \geq 0 \\ \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j = 0, 1, 2, \dots \\ \sum_j \pi_j = 1 \end{cases}$$

$\pi_j$ 稱為 limiting distribution / stationary distribution / equilibrium distribution.

**Example** For the social class matrix

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left\| \begin{array}{ccc} 0.40 & 0.50 & 0.10 \\ 0.05 & 0.70 & 0.25 \\ 0.05 & 0.50 & 0.45 \end{array} \right\| \end{array},$$

the equations determining the limiting distribution $(\pi_0, \pi_1, \pi_2)$ are

$$0.40\pi_0 + 0.05\pi_1 + 0.05\pi_2 = \pi_0, \tag{4.8}$$
$$0.50\pi_0 + 0.70\pi_1 + 0.50\pi_2 = \pi_1, \tag{4.9}$$
$$0.10\pi_0 + 0.25\pi_1 + 0.45\pi_2 = \pi_2, \tag{4.10}$$
$$\pi_0 + \pi_1 + \pi_2 = 1. \tag{4.11}$$

$$\Rightarrow \quad -60\pi_0 + 5\pi_1 + 5\pi_2 = 0, \tag{4.12}$$
$$5\pi_0 - 3\pi_1 + 5\pi_2 = 0, \tag{4.13}$$
$$\pi_0 + \pi_1 + \pi_2 = 1. \tag{4.14}$$

Then, $\pi_0 = \frac{5}{65} = \frac{1}{13}$, $\pi_1 = \frac{5}{8}$, and then $\pi_2 = 1 - \pi_0 - \pi_1 = \frac{31}{104}$

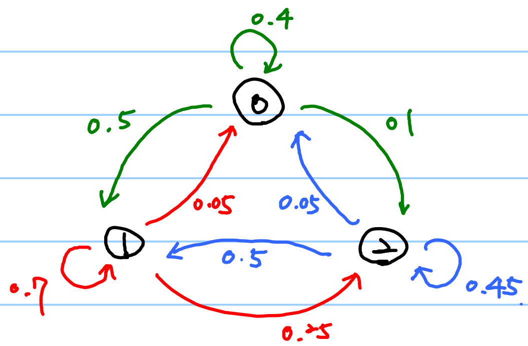
HW 3. A Markov chain $X_0, X_1, X_2, \ldots$ has the transition probability matrix

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left\| \begin{array}{ccc} 0.7 & 0.2 & 0.1 \\ 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \end{array} \right\| \end{array}.$$

Determine the limiting distribution. (須列式. 計算可用電腦輔助)

Remark: 1. $\pi_j$ 稱為 limiting distribution (極限分布) 是因 $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}$

2. : stationary distribution (靜態分布) 是因為若初始機率為 $\{\pi_j\}$. 則 $X_1, X_2, \ldots$ 之機率分布均為 $\pi_j$




$$\overbrace{\underbrace{0.50\pi_0}_{\text{0 轉至 1 的比例}} + \underbrace{0.70\pi_1}_{\text{1 留在 1 的比例}} + \underbrace{0.50\pi_2}_{\text{2 轉至 1 的比例}}}^{\text{下一步 state 1 之比例}} = \overbrace{\pi_1}^{\text{現在 state 1 之比例}}, \quad \leftarrow (4.9)$$

i.e. 下一步的比例 = 現在的比例.

3. $\pi_j$ 稱為 equilibrium distribution (穩態分布) 是因系統一旦達到該分布則系統進入平衡狀態.

4 由 Theorem (a), $\pi_j$ 可視為長期下來 state $j$ 發生的頻率

Example: 機器售票系統有2台電腦. 任何時間均只有一台運作, 一台備用.
假設運作中電腦每天掛掉之機率為 $p$.
掛掉之電腦維修需花2天時間, 且一次只能修一台電腦

Q: 希望知道什麼性質？

如何將 Markov Chain 應用在此問題上？

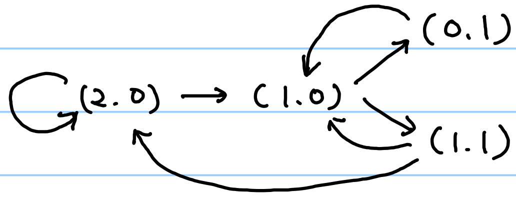
考慮 Markov Chain 之 state 為 $(x, y)$

其中 $x$ 為每日結束時可用之電腦數，

$$y = \begin{cases} 1 & \text{有壞掉電腦且已修理1天} \\ 0 & \text{o.w.} \end{cases}$$

是否已涵蓋所有狀態？

狀態變化圖：




Transition probability matrix 如下：

$$\begin{array}{ll} \text{To state} \rightarrow & (2,0) \quad (1,0) \quad (1,1) \quad (0,1) \\ \text{From state} \downarrow & \end{array}$$

$$\mathbf{P} = \begin{array}{c} (2,0) \\ (1,0) \\ (1,1) \\ (0,1) \end{array} \begin{Vmatrix} q & p & 0 & 0 \\ 0 & 0 & q & p \\ q & p & 0 & 0 \\ 0 & 1 & 0 & 0 \end{Vmatrix},$$

where $p + q = 1$.

希望知道長期兩台電腦均掛掉之機率(或時間比例)：$\pi_3$

解

$$\begin{cases} q\pi_0 \quad + q\pi_2 = \pi_0, \\ p\pi_0 \quad + p\pi_2 + \pi_3 = \pi_1, \\ \quad q\pi_1 = \pi_2, \\ \quad p\pi_1 = \pi_3 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1. \end{cases}$$

得

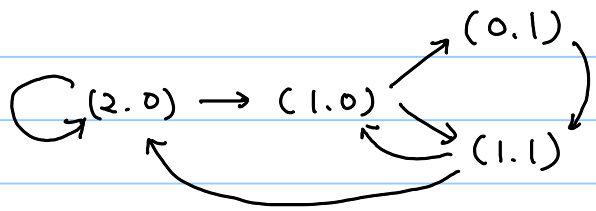
$$\pi_0 = \frac{q^2}{1+p^2}, \quad \pi_2 = \frac{qp}{1+p^2},$$

$$\pi_1 = \frac{p}{1+p^2}, \quad \pi_3 = \frac{p^2}{1+p^2}.$$

The availability is $R_1 = 1 - \pi_3 = 1/(1+p^2)$.

若欲增加 availability，除增加備用電腦數外，亦可考慮增加1名維修人員，
使該系統可同時維修2台電腦。

此時
狀態變化圖：




Transition probability matrix 變為：

$$
\begin{array}{ll}
 & \text{To state} \to \quad (2,0) \quad (1,0) \quad (1,1) \quad (0,1) \\
 & \text{From state} \downarrow
\end{array}
$$

$$
\mathbf{P} = \begin{array}{c} (2,0) \\ (1,0) \\ (1,1) \\ (0,1) \end{array}
\left\| \begin{array}{cccc}
q & p & 0 & 0 \\
0 & 0 & q & p \\
q & p & 0 & 0 \\
0 & 0 & 1 & 0
\end{array} \right\|,
$$

and the limiting distribution is

$$
\pi_0 = \frac{q}{1+p+p^2}, \quad \pi_2 = \frac{p}{1+p+p^2},
$$

$$
\pi_1 = \frac{p}{1+p+p^2}, \quad \pi_3 = \frac{p^2}{1+p+p^2}.
$$

Thus, availability has increased to $R_2 = 1 - \pi_3 = (1+p)/\left(1+p+p^2\right)$.

- Markov Chain Monte Carlo method（MCMC 馬可夫鏈蒙地卡羅方法）

# 蒙地卡羅方法 [編輯]

維基百科，自由的百科全書

**蒙特卡洛方法**（英語：Monte Carlo method），也稱**統計模擬方法**，是二十世紀四十年代中期由於科學技術的發展和電子計算機的發明，而被提出的一種以概率統計理論為指導的一類非常重要的數值計算方法。是指使用隨機數（或更常見的偽隨機數）來解決很多計算問題的方法。

20世紀40年代，在馮·諾伊曼，斯塔尼斯拉夫·烏拉姆和尼古拉斯·梅特羅波利斯在洛斯阿拉莫斯國家實驗室為核武器計劃工作時，發明了蒙特卡洛方法。因為烏拉姆的叔叔經常在蒙特卡洛賭場輸錢得名，而蒙特卡羅方法正是以概率為基礎的方法。

與它對應的是確定性算法。

蒙特卡洛方法在金融工程學，宏觀經濟學，生物醫學，計算物理學（如粒子輸運計算、量子熱力學計算、空氣動力學計算）等領域應用廣泛。[1]

## 蒙地卡羅方法的工作過程 [編輯]

在解決實際問題的時候應用蒙地卡羅方法主要有兩部分工作：

1. 用蒙地卡羅方法模擬某一過程時，需要產生各種概率分布的隨機變量。
2. 用統計方法把模型的數字特徵估計出來，從而得到實際問題的數值解。

## 蒙地卡羅方法在數學中的應用 [編輯]

通常蒙地卡羅方法通過構造符合一定規則的隨機數來解決數學上的各種問題。對於那些由於計算過於複雜而難以得到解析解或者根本沒有解析解的問題，蒙地卡羅方法是一種有效的求出數值解的方法。一般蒙地卡羅方法在數學中最常見的應用就是蒙地卡羅積分。下面是蒙特卡羅方法的兩個簡單應用：

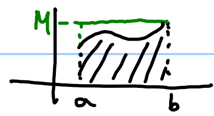
### 積分 [編輯]

欲計算 $\int_a^b f(x)dx$：

<u>方法1</u>： 令 $X_i \sim \text{Unif}(a,b)$， $i=1,2,\dots$ i.i.d.

則 $\frac{b-a}{n}\sum_{i=1}^{n} f(X_i) \longrightarrow \int_a^b f(x)\,dx$ as $n\to\infty$

故當 $n$ 夠大時，$\frac{b-a}{n}\sum_{i=1}^{n} f(X_i)$ 可為其近似值。



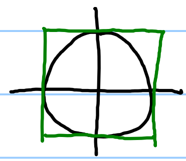

<u>方法2</u>： 假設 $0 \le f(x) \le M$ for $x \in (a,b)$.

令 $(X_i, Y_i) \sim \text{Unif}((a,b)\times(0,M))$， $i=1,2,\dots$ i.i.d.

則 $M\times(b-a)\times\frac{1}{n}\sum_{i=1}^{n} I_{\{Y_i \le f(X_i)\}} \longrightarrow \int_a^b f(x)dx$ as $n\to\infty$

圓周率： 令 $(X_i, Y_i) \sim \text{Unif}((-1,1)\times(-1,1))$ $i=1,2,\dots$ i.i.d.

則 $\frac{4}{n}\sum_{i=1}^{n} I_{\{X_i^2+Y_i^2 \le 1\}} \longrightarrow \pi$ as $n\to\infty$



<u>MCMC</u>：

$\mathbb{X}$ : random vector w/ values $x_j$, $j\ge1$.

欲計算 $\theta = \mathbb{E}[h(\mathbb{X})] = \sum_{j=1}^{\infty} h(x_j)\,P\{\mathbb{X}=x_j\}$

直接計算有困難時，通常用 simulation 來逼近 $\theta$.

## Monte Carlo simulation:

1. 用 random number 產生 i.i.d r. vectors $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$.

   $\underline{X}_i \overset{d}{=} \underline{X}$.

   則由 SLLN, $\lim\limits_{n\to\infty} \sum\limits_{i=1}^{n} \frac{h(\underline{X}_i)}{n} = \theta$

2. 通常產生 $\underline{X}_i \overset{d}{=} \underline{X}$ 無法做到.

   如下列情況:

   $P\{\underline{X} = x_j\} = c b_j \cdot \ j \geq 1,$

   其中 $b_j$ 已知, 但 C 無法計算 (可能 $\sum b_j$ 加太多項)

此時可造 MC $\underline{X}_1, \underline{X}_2, \ldots$ 使其 limiting prob 為 $P\{\underline{X} = x_j\}$. $j \geq 1$.

則 $\frac{\sum_{i=1}^{n} h(\underline{X}_i)}{n} \longrightarrow \theta$. as $n \to \infty$ (※)

Why (※) is true?

$$\frac{\sum_{i=1}^{n} h(\underline{X}_i)}{n} = \frac{\sum_j \left[h(j) \times (\underline{X}_1 \ldots \underline{X}_n \text{走到 } j \text{ 之次數})\right]}{n}$$

$$= \sum_j \left[h(j) \times \underbrace{\frac{(\underline{X}_1, \ldots, \underline{X}_n \text{走到 } j \text{ 之次數})}{n}}_{\pi_j}\right] \longrightarrow \sum_j h(j)\pi_j = E\, h(\underline{X})$$

## 如何造所要之 MC?

介紹兩種演算法: 1. Hastings - Metropolis Algorithm

2. Gibbs sampler

令 $b_j > 0$. $j = 1, 2, \ldots$ . $C = \frac{1}{\sum_j b_j}$

liming prob $\pi_j = c b_j$

Hastings-Metropolis Algorithm：找出一 time reversible MC.

1. 令 $Q=[q_{ij}]$ 為一任意之 irreducible Markov transition matrix
   令 $Y_i$ 為 r.v. w/ $P\{Y_i=j\}=q_{ij}$

2. 定义 $X_n$ 如下：
   (1) 若 $X_{n-1}=i$ 則產生一 r.v. $Y \stackrel{d}{=} Y_i$.
   (2) 若 $Y=j$. 定义

$$X_n = \begin{cases} j & \text{w/ prob } \alpha(i,j) \\ i & \text{w/ } " \quad 1-\alpha(i,j) \end{cases}$$

其中 $\alpha(i,j)=\min\{1, \dfrac{b_j q_{ji}}{b_i q_{ij}}\}$ ← 為何如此設定 $\alpha(i,j)$？
為讓 $X_n$ 之 limit distri 為 $\pi_j$

則 $\{X_n\}$ 為一 M.C. 且其 transition prob 為

$$P_{ij} = q_{ij}\,\alpha(i,j) \qquad \text{if } i \neq j$$

$$P_{ii} = q_{ii} + \sum_{k \neq i} q_{ik}(1-\alpha(i,k))$$

Note：1. 上述演算法完全不需要知道 c 之值.

2. $\alpha(i,j)$ 之設定為由下列推導所得： ← $\alpha(i,j)$ 為彈性調整項.
   欲令 $X_n$ 為 time reversible 且 $\pi_j$ 為其 limiting distribution.
   希望下列等式成立：

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i,j$$

(Why? 上式可推得 $\pi_i = \sum_j \pi_i P_{ij} = \sum_j \pi_j P_{ji}$)

i.e. $\pi_i q_{ij}\alpha(i,j) = \pi_j q_{ji}\alpha(j,i)$, $\forall i,j$

$$\Rightarrow \alpha(i,j) = \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\alpha(j,i)$$

$$\begin{cases} \text{若 } \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \le 1, & \text{則令 } \alpha(j,i)=1. \text{ 此時 } \alpha(i,j)=\dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \\ \text{若 } \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \ge 1 & \text{則令 } \alpha(i,j)=1 \end{cases}$$

為令 $\alpha(i,j)$ 為機率值. 故 $\alpha(i,j) \le 1$, $\forall i,j$

**Example**: 令 $\mathcal{S}=\{(x_1,x_2,\dots,x_n)$ 為 $(1,2,\dots,n)$ 之 permutation $: \sum_{k=1}^{n} k x_k > a\}$

欲產生 uniformly distributed random vector on $\mathcal{S}$

使用 Hastings - Metropolis 及 graph theory :

考慮一以 $\mathcal{S}$ 元素為其 vertices 之 graph

若 $\mathcal{S}$ 中之兩元素,其差異僅為兩位置對調,則稱此二元素(vertices)為 neighbors

令 $N(i)=\{i$ 之 neighbors$\}$, $i\in\mathcal{S}$

$$q_{ij} = \begin{cases} 1/\#N(i) , & \forall\ j\in N(i) \\ 0 , & o.w. \end{cases}$$

$$\alpha(i,j) = \min\left\{1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}}\right\} = \min\{1, q_{ji}/q_{ij}\} \quad (\because \pi_i \text{ uniform})$$

$$= \min\left\{1, \frac{\#N(i)}{\#N(j)}\right\}$$

**Remark**: $(1,\dots,n)$ 之 permutation 共有 $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ 種

但 $N(i)$ 之元素個數不超过 $\binom{n}{2}$

**Gibbs sampler**

令 $\vec{X}=(X_1,\dots,X_n)$ w/ prob $p(\vec{x}) = C\,g(\vec{x})$, discrete.

欲造 r.v. $\stackrel{d}{=} \vec{X}$.

**假設**: 可產生 r.v. $Y$ w/ prob

$$P\{Y=y\} = P\{X_i=y \mid X_j=x_j,\ j\neq i\}$$

$$= \frac{g(x_1,\dots x_{i-1},y,x_{i+1},\dots,x_n)}{\sum_z g(x_1,\dots x_{i-1},z,x_{i+1},\dots,x_n)}$$

i.e. $(\underbrace{X_1,X_2,\dots,X_{i-1}}_{\text{給定}}\ X_i,\ \underbrace{X_{i+1},\dots,X_n}_{\text{給定}})$

$\uparrow$ conditional distribution 有簡單型式

**方法**:

1. 假設現在狀態為 $(x_1,x_2,\dots,x_n)$

   在狀態向量中隨机(uniformly)擇一座標讓其變動

2. 若 1. 中选到 $i$, 則產生 r.v. $Y$ 如假設之 distribution.

3. 令 $y$ 為 $Y$ 產生之值, 則下一狀態為 $(x_1,x_2,\dots,x_{i-1},y,x_{i+1},\dots,x_n)$

上述方法可視為使用 Hastings-Metropolis，其中

$$q(\vec{x},\vec{y}) = \frac{1}{n} P\{X_i = y \mid X_j = x_j,\ j \neq i\} = \frac{P(\vec{y})}{n P\{X_j = x_j,\ j\neq i\}} \qquad \vec{y} = (x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)$$

$$\alpha(x,y) = \min\left\{ \frac{P(\vec{y})\, q(\vec{y},\vec{x})}{P(\vec{x})\, q(\vec{x},\vec{y})},\ 1 \right\}$$

$$= \min\left\{ \frac{P(\vec{y})\, P(\vec{x})}{P(\vec{x})\, P(\vec{y})},\ 1 \right\} = 1.$$ (i.e. 下一個 state 必為 $y$)

Remark: 1. Gibbs sampler 使用時機在單一座標改變之條件機率簡單時.

而 Hastings-Metropolis 不同處在於並不一定需要知道 $q(x)$ (or $b(j)$)

2. Gibbs sampler 在連續時亦可使用

Example 4.36 造出 r.v. $\overset{d}{=}$ Uniform on $\{(x_1,\dots,x_n) : x_i \text{ on unit circle},\ |x_i - x_j| > d\}$

Gibbs sampler:

1. 現在 state $(x_1,\dots,x_n)$
2. 隨機選出座標 $i$.
3. : $x$, 使得 $|x - x_j| > d$, $j \neq i$.
4. 下個 state $= (x_1,\dots,x_{i-1}, x, x_{i+1},\dots,x_n)$.

Application:

演講者：Wendy K. Tam Cho (Professor and Senior Research Scientist in National Center for Supercomputing Applications)

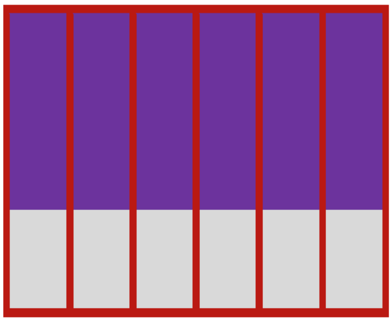
講　題：A Massively Parallel Evolutionary Multiple-Try Metropolis Markov Chain Monte Carlo Algorithm for Spatial State Space Traversal

時　間：2021年3月22日 (星期一) 14:00-15:00

地　點：臺灣大學天數館 202 室

https://pol.illinois.edu/directory/profile/wendycho

貓黨、狗黨在某州選民支持度各為66.6%與33.3%。該州劃分六選區，每選區選出一人。




左圖選區劃分貓黨全上，右圖選區劃分則貓狗黨各有4,2席。

<u>**目的**</u>: 起訴執政黨用不正當方式劃分選區，試圖以選區劃分方式影響選舉結果。說服審理法官現行劃分方式為刻意作為的不自然方式。

<u>**方法**</u>: 用均勻(或接近均勻)抽樣的方式抽樣n種選區劃分結果，每個被抽取出的劃分結果可計算出對應選舉結果，將n個選舉結果的機率分布圖畫出，最後說明當時的選舉結果在該機率分布圖的單邊5%(或更小)，亦即 reject H0: 目前選區劃分是在合理範圍

假設選區有五十個區域，劃分為五個選區，則在不考慮選區劃分規則下，共有

$$\frac{C_{10}^{50}\times C_{10}^{40}\times C_{10}^{30}\times C_{10}^{20}\times C_{10}^{10}}{5!}=\frac{\left(\frac{50!}{10!\times 10!\times 10!\times 10!\times 10!}\right)}{5!}=4.0279\times 10^{29}$$

種選區劃分。但滿足選區劃分規則的劃分方式只占上述的一小部分。故直接抽樣再去除不合規則的劃分就變得非常沒有效率而不可行。

解決方法: 使用Metropolis Algorithm 建立馬可夫模型做近似均勻分配的抽樣。

Algorithm:

1. 建立起使選區劃分
2. 隨機選取選區邊界，選定後將選區邊界相鄰區域所屬選區交換，並判斷交換後選區劃分是否合乎規則，若不合乎規則則重新選取其他邊界。
3. 重複2步驟N次，得到一個選區劃分樣本。
4. 重複2,3步驟，得到n個樣本。