

## Practice#1：隨機與機率的初體驗

### 目的：

進行統計分析或推論時，『機率』(probability)是必要的基礎。而機率這門俱百年歷史的學問，來自從日常生活觀察而來的『隨機』(randomness)現象。人生處處充滿不確定性(uncertainty)，永遠不知道下一刻會發生什麼事，不知道明天的股市會漲還是跌，不知道阿娘的肚子懷的是弟弟還是妹妹？愈是不確定的事，人們愈是好奇的想早一點知道，甚至去控制或左右事情的發生。

聰明的科學家，發揮其敏銳的觀察能力，及無聊到極致的研究精神終於發現，隨機的現象可不是亂無章法到不可捉摸。相反的，從長期的觀察，『亂』中仍是有『序』的！本單元試圖讓同學重複『古代』科學家做的無聊實驗，利用電腦的高速計算能力，模擬一些隨機現象。正因為電腦的高速，費時又單調的實驗可以在瞬間完成。再也不必像十九世紀英國的統計學家卡爾·皮爾生(Karl Pearson)，神勇的擲一個銅板 24,000 次。只為了想知道擲出一枚銅板，它出現正面的機率是不是想像中的 0.5 呢？今天的你，會幹這種傻事嗎？

這個練習將開啟你對模擬(simulation)的認識，看看電腦如何以較快速、便宜的方式取代實際的經驗，去獲取對事情的認知與對未來的預測。或是提供理論分析的驗證。

### 練習：

1. 你怎麼告訴別人什麼叫做不確定性(uncertainty)？什麼是『隨機』(randomness)？什麼是『機率』(probability)？試著自己說明清楚。
2. 擲一個公正的銅板，會出現正面還是反面？沒有人可以很有把握的猜準。有人說：那是隨機的，不一定正面或反面，但是出現正面與反面的機會各半，都是 50%。或說：出現正反面的機率都是 0.5。這 0.5 或 50%從何而來？打從出娘胎開始，你什麼時候開始接受 0.5 這個答案呢？這個練習要模擬十八、九世紀的許多『鐵齒』學者，擲它個幾萬次銅板看看事情的真相？並進一步瞭解『隨機』與『機率』的意義。
  - MATLAB 有沒有擲銅板的指令？找找看？(Hint: random number generator)
  - 模擬擲 10 次銅板，出現正面的次數有幾次？比例等於 0.5 嗎？
  - 擲 100 次呢？200 次？500 次？、、、10,000 次出現正面的比例？
  - 記錄從擲 1 次、2 次到 10,000 次銅板，出現正面的比例。畫一張圖來呈現這 10,000 個比例。你觀察到什麼？
3. 有兩種刮刮樂彩券：A 與 B，賭注一樣。玩 A 彩券，刮中贏 10 元，贏的機率是 1/2。玩 B 彩券，刮中可贏 10,000 元，但是贏的機率只有 1/10。作為

一個賭徒(我說的是『賭徒』喔)，你玩哪一種呢？這是不是在考你「期望值」的計算呢？不要搖動你的筆桿，敲敲鍵盤，算出你的答案吧(p.292 [1])

### 觀察：

1. 利用電腦做機率的試驗時，特別要注意試驗的次數。一次試驗的結果，我們可以說是隨機的，但次數一多，便提供我們觀察隨機事件的行為模式。
2. 電腦永遠不能替代擲銅板的實際行為，只能模擬。也因此試驗結果的準確度，需仰賴這個模擬的精準度。MATLAB 的二項分配亂數產生器，模擬的好不好呢？其實透過實際的試驗與理論值的對照，你心裡大約有個準兒！但是使用不同軟體的亂數產生器做模擬時，還是要保持戒慎恐懼的心態。
3. 這個練習所模擬的資料，其行為都比較單純，因此資料的模擬也相對簡單。這不代表所有的模擬都是如此。有些資料的模擬需要很強的理論基礎作為依據、有些資料的模擬需要借助很多的假設。
4. 不管如何，模擬的目的在於對『不確定』的行為產生一次又一次的結果，就像重複轉動〔蒙地卡羅賭場〕裡的輪盤。這類的模擬一般稱為 Monte Carlo Simulation。

### 作業：

1. 練習 2 的圖及你的觀察。
2. 擲 6 個銅板都出現正面的機率與出現「正反正正反反」的機率孰大？印象中，連續出現相同結果的可能性「應該」很低。寫一支程式計算擲個銅板，至少三個正面連續出現的機率有多高？你必須決定作多少次的試驗，才能斷定這個機率值。如果你學過機率，這個由電腦跑出來的機率估計值，與真正的理論值相差多少？
3. 一對夫婦計畫生孩子，直到生出女兒為止，或者生到四個時停止，視那個狀況發生而定。請問他們可能擁有女兒的機率是多少？這是一個普通的機率問題，如果你的機率論學得還可以，大概可以算的出來。不過，如果算不出來也不必氣餒，只要你會寫程式，照樣給它算出來！(p.289, [1])
4. 同上題，採用這個計畫的夫婦，平均起來會有幾個孩子？(p.299,[1])這不是一種期望值的計算呢？

### 參考文獻：

David S. Moore, “統計，讓數字說話”，鄭惟厚 譯，天下文化。