機率統計入門

黃文璋

國立高雄大學統計學研究所

統計研習營

中央研究院統計科學研究所

104年7月1-3日

第二講

認識機率

- 1987年，印度傳奇數學家拉曼紐揚(S. Ramanujan, 1887-1920) 百年誕辰，有一系列的紀念活動。
- 著名統計學者勞氏(C. R. Rao, 1920-)，也應邀給了三場演講。

- 之後，印度統計學研究所(Indian Statistical Institute)基於勞氏的演講稿，於1989年，為他出版了統計與真理一書，1997年發行第二版。

在第一版的序文中，勞氏提到：

- 學生時代，我主修數學——一種從給定前提下演譯結果的邏輯。
(the logic of deducing consequences from giving premises)

- 後來我唸統計學——一種從經驗中學習的合理過程，及從給定的結果驗證前提的邏輯。
(a rational approach to learning from experience and the logic of identifying the premises given the consequences)

◆我已認識到數學及統計，在人類為提昇自然知識，及有效管理日常事務，所做的一切努力中，佔有重要性。

我相信：

- 在最終的分析中，所有知識皆為歷史。
- 在抽象的意義下，所有科學皆為數學。
- 在理性的世界裡，所有判斷皆為統計。

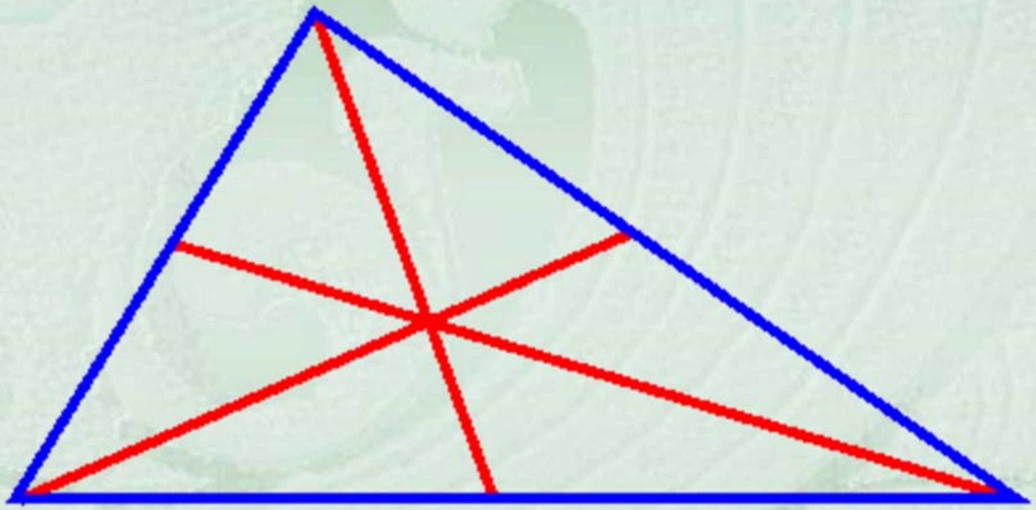
- 近年來，鑒於統計學的重要性，高中數學裡，逐漸加進統計的題材。
- 95學年開始的“普通高級中學數學課程綱要”中，新增的信賴區間與信心水準，卻帶給師生不小困擾。
- 此新加入的統計題材，由於需取樣，得到數據，使機率論裡隨機性的特質顯現出來。

- 雖有人認為機率與統計，這類數學所需的前置準備不多，因此提前教沒問題。
- 但隨機性的概念，在理解層次上，其實並不是那麼容易能掌握。
- 不少人認為，大學裡統計比微積分難教、難學。

- 微分、積分皆有物理意義。
- 投擲一公正骰子，所得點數之期望值為3.5？

宇宙的運轉，有必然性及隨機性。

數學中有很多必然性。



三角形重心、外心、垂心三點共線。

兩千三百年前的幾何原本中有許多這類的結果。

- 隨機性：
  - 銅板落地那一面朝上？
  - 氣象局預測莫拉克颱風降雨集中於北部地區。

- 對必然性的探討不論在天文、物理、化學或數學，歷史悠久，且成效不錯。
- 科學家樂意挑戰必然性的問題。

- 對隨機性的了解則很侷限：
  - 明天會不會下雨？
  - 颱風走向？
  - 下次地震是何時？

- 信賴區間的概念，是奈曼於1934年的演講中首度提出。
- 演講結束後，大會主席包雷(A. L. Bowley, 1869–1957)於致詞中提到：

我不很確定此信心不是一信心戲法
(I am not all sure that the confidence
is not a confidence trick)。

- 在95%信賴區間中，
  - 95%究竟是指什麼？
  - 是機率嗎？
  - 如果是，是什麼事件的機率？

◆統計裡有各種方法，但因處理隨機現象，少有倚天既出，誰與爭鋒的方法。

高中數學98課綱

- 附錄3.3"常態分布，信賴區間與信心水準的解讀"中說：

  高中程度的統計推論只做隨機變數期望值的估計，它的背後理論是中央極限定理。要介紹中央極限定理，就需引入常態分布。此部分僅做通識性的介紹，以活動方式建立學生對於中央極限定理的直觀。對一固定的信心水準，給出信賴區間公式，再讓學生以亂數表模擬或實驗投擲正面出現為p的銅板n次，代入信賴區間公式，以說明信心水準的意涵；並以此解讀，何以大多數學生所得的信賴區間都會涵蓋p？

- 何以信賴區間的概念不易理解？

  未能正確了解機率的涵義！

- 若輕忽機率概念的深度，由於其重要，便提早教機率統計，對國民善用統計，並無幫助。

2. 機率的意義

◆古典機率，基本假設是相同可能性。適用於擲銅板、骰子、玩撲克牌，…。

- 2009年，高爾夫球好手伍茲，參加別克公開賽。第一輪打完，落後領先者8桿，排名91。
- 前一場比賽，他首次提前被淘汰。伍茲不行了？

- 打完3輪後，伍茲躍居首位。
- 這時大家看法丕變，一致認為這座冠軍盃，幾乎是他的囊中物了。
- 因過去的紀錄，伍茲如能帶著54洞領先進入決賽圈，戰績35勝1敗。

機率值會變！

他贏了沒有？

阿根廷vs.瑞士 阿根廷出線機率63%

阿根廷想贏球打進八強得靠前鋒梅西，瑞士想要打敗阿根廷闖進八強，必須有效封鎖梅西，明天凌晨零時開踢的16強淘汰賽阿根廷對瑞士，梅西是頭號焦點。數據分析，阿根廷在這場比賽贏球的機率達63%。…

(聯合晚報 103年7月1日 李亦伸)

世足／美強歐弱 主場優勢求勝切

風水輪流轉，美強歐弱是本屆世足的特點之一，…，使得晉級16強的有一半是美洲球隊。不少評論指出，這是歐洲人對南美氣候「水土不服」所致，但德國隊教練勒夫則認為，美洲球隊因主場氣勢，贏球的慾望更強烈。根據歷史法則，只要是在美洲舉辦的世界盃，歐洲球隊奪冠的機率是零。…

(台灣醒報103年6月29日 陳正健)

◆以相對頻率來解釋機率，

在沒有更多資訊下，常被認為是一客觀的辦法。

◆某君看上一女孩，驚為天人，覺得這是他今生的新娘。

◆評估後信心滿滿，自認追上的機會有8成：

主觀機率！

- 上述三種常見對機率的解釋，大抵就是人們評估事件發生可能性之大小的幾種思維。
- 雖針對不同的情況，但常能交互著運用。

- 此三種對機率的解釋，已涵蓋不少實際生活中的情況。但數學家當然不會在此止步。
- 他們喜歡抽象化，及一般化。

以公理化的方式引進機率

- 公理化後，機率論便快速地發展，並成為數學中一重要的領域。

- 這歸功於俄國的科莫果洛夫(A. N. Kolmogorov, 1903-1987)，於1933年出版的小書機率論的基礎：

  機率論作為數學學科，可以而且應該從公理開始發展，就如同幾何、代數一樣。

- 從事一隨機實驗，或觀察一隨機現象，其所有可能的結果之集合，稱為一樣本空間(sample space)，通常以Ω表之，$\Omega \neq \varnothing$。

- Ω中的每一元素，便稱為一樣本，通常以ω表之。

  令$\mathcal{F}$表包含Ω之所有子集合之集合，每一$A \in \mathcal{F}$便稱為一事件(event)。

- 對每一事件A，給定一實數$P(A)$，稱為A之機率。$(\Omega,\mathcal{F},P)$就構成一機率空間。

- 在較高等的機率論的書裡，$\mathcal{F}$往往允許取成只是Ω之一些子集合之集合(但滿足若干條件)，而不須包含Ω所有子集合。
- 這樣放寬對$\mathcal{F}$之限制，不但使有些Ω之子集合不是事件，也讓一些後續的討論，如隨機變數(random variable)，更富彈性與變化。

- 在機率空間裡，P稱為機率測度，簡稱機率，它是定義在$\mathcal{F}$上之一實函數，且滿足

(i)$\forall A \in \mathcal{F}, P(A) \geq 0$;

(ii)若$A_n, n \geq 1$，為互斥事件，則

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n);$$

(iii)$P(\Omega) = 1$。

- 二事件A, B，若 $A \cap B = \varnothing$，便稱為互斥。

例. 令$\Omega=\{\omega_1,\omega_2,\cdots\}$為一可數的集合，$f$為一定義在$\Omega$之實函數，且滿足

$$f(\omega_i)\geq 0,\ \forall i\geq 1,\text{且}\sum_{i=1}^{\infty} f(\omega_i)=1,$$

$\mathcal{F}$為Ω之所有子集合之集合，再定義一$\mathcal{F}$上之函數P如下：

$$P(A)=\sum_{\omega\in A} f(\omega),\ A\in\mathcal{F}\ ,$$

則P為一機率測度。$(\Omega,\mathcal{F},P)$即構成一離散型的機率空間。

- f稱為該隨機變數之機率密度函數(簡稱p.d.f.)。

例. 設Ω為實數R上之一區間，f為Ω上之一實函數，且滿足

$$f(\omega) \geq 0, \forall \omega \in \Omega, \text{且} \int_{\Omega} f(\omega) d\omega = 1。$$

則由下式可定義出一機率測度

$$P((a,b]) = \int_a^b f(\omega) d\omega, \forall a,b \in \Omega, \text{且} a < b。$$

- 所謂隨機變數，就是一個定義在樣本空間Ω之實函數。若以X表一隨機變數，則此函數之定義域為Ω，對應域為實數集合R。

- 引進隨機變數，可能只是想將實驗結果數值化，如

$$X(\text{正面}) = 1,\ X(\text{反面}) = 0,$$

又如Ω表某校新生之集合，則對$\forall \omega \in \Omega$，可令

$$X(\omega) = \omega\text{之身高。}$$

- 隨機變數通常以英文大寫字母表示。

隨機變數就是函數，何以很多人難以理解？

◆對一隨機變數X，我們定義其分佈函數F為

$$F(x) = P(X \leq x) = P(\{\omega \mid X(\omega) \leq x\}), x \in R \text{。}$$

- 設有二隨機變數X, Y，定義在同一機率空間。對任二實數x, y，事件$\{X\leq x\}$與$\{Y\leq y\}$同時發生之機率，以$P(X\leq x, Y\leq y)$表之，其意義為

$$P(\{\omega | X(\omega) \leq x, \text{且} Y(\omega) \leq y\})\text{。}$$

以

$$F(x, y) = P(X \leq x, Y \leq y), x, y \in R,$$

當作X, Y之聯合分佈函數。

◆二隨機變數X及Y，若滿足

(2) $$P(X \le x, Y \le y) = P(X \le x)P(Y \le y), \forall x, y \in R,$$

便稱為獨立，否則稱為不獨立。

◆設有X,Y二隨機變數，機率密度函數分別為g及h，聯合機率密度函數為f。則X與Y獨立，若且唯若

(3) $$f(x, y) = g(x)h(y), x, y \in R。$$

◆也可將以上聯合分佈及獨立的概念推廣到n個變數的情況。

例. 設$\Omega=(0,1]$，$\mathcal{F}$為包含Ω上所有開區間之最小的$\sigma-$體，$P((0,a])=a, 0<a<1$，則得機率空間$(\Omega,\mathcal{F},P)$。

令$X(\omega)=\omega$，$Y(\omega)=1-\omega,\ \omega\in\Omega$。對$0<x\leq 1$,

$$P(X\leq x)=\mathrm{P}(\{\omega \mid \omega\in\Omega, X(\omega)\leq x\})=P((0,x])=x,$$
$$P(Y\leq x)=\mathrm{P}(\{\omega \mid \omega\in\Omega, Y(\omega)\leq x\})=P([1-x,1])=x,$$

故X, Y皆有$\mathcal{U}[0,1]$分佈。又對$0<x\leq 1, 0<\mathrm{y}\leq 1,$

$$\begin{aligned} P(X\leq x, Y\leq y) &= P((0,x]\cap[1-y,1]) \\ &= \begin{cases} 0, & x+y<1, \\ x+y-1, & x+y\geq 1。\end{cases} \end{aligned}$$

- 顯然，X與Y不獨立。事實上，因$X(\omega) + Y(\omega)=1$, $\forall\ \omega \in \Omega$，故$P(X+Y=1) = 1$。又$X$, Y之聯合機率密度函數不存在。

- 機率空間是最上游的，若給定機率空間，並在其上定義隨機變數，則可得分佈函數。但有時資訊不足，對一隨機變數，不知其從那裡來，需估計其分佈，或已知屬於那一分佈族，要估計其中的參數。二隨機變數間之關係，也可能需估計。
- 統計裡常在做估計。

3. 何處是機率天地

◆有法國牛頓之稱的拉普拉斯(P. S., M. de Laplace, 1749-1827)曾說：

> 這門源自考慮賭博中的機運之科學，必將成為人類知識中最重要的一部分，生活中最重要的問題中的大部分，都將只是機率的問題。

◆隨機的經驗源遠流長。

◆舊約聖經利未記第十六章：

為那兩隻羊拈鬮，一鬮歸與耶和華，
一鬮歸與阿撒瀉勒。

◆民數記(Numbers)第二十六章(The Second Census)，耶和華曉諭摩西：

還要拈鬮分地。

民數記第一章的英文標題為The Census。

◆在新約聖經，耶穌被釘死在十字架上後：

兵丁以拈鬮來分他的裏衣。

- 機率是針對隨機現象。
- 有些事雖非隨機，如

  胎兒性別、考前猜題、拿在背後的水果，…，

  仍可討論機率。
- 惠子對莊子說：

  子非魚，安知魚之樂？
- 但莊子可說：

  魚快樂之機率為…。

- 投擲銅板，那一面朝上？是否隨機？
- 一切按照神的旨意在進行？
- 物理學家的看法？

條件機率

- 在給定某事件B發生(有新的資訊)下，原先那一事件A發生的機率，有時會隨之而變。
- 何時不變？

  A與B獨立！
- 條件機率看來似乎不難，但實際應用時，並沒那麼容易。

在一篇名為機率與文字陷阱的文章

例1.

(i)好友有二小孩，已知老大是女孩。

問：老二亦是女孩之機率？

(ii)好友有二小孩，已知有一個女孩。

問：另一小孩亦是女孩之機率？

例2. (某高中數學競試)

甲投擲兩硬幣，乙則猜朝上的面相同或相異。乙正準備要猜，丙從旁經過，說：

有一正面。

問：乙該猜相同或相異？

- 例2與例1-(ii)情況相同否？

該文指出：

- 但再仔細想想，如果丙看到的是反面，那麼乙也要猜相異，而且猜對的機率也是2/3。
- 所以乙只要知道丙有說話，儘管乙不知丙說什麼，猜相異對的機率就是2/3。
- 那其實乙根本不需要丙幫忙，只要他猜的時候假想有一個丙走過來跟他說話，那猜相異猜對機率就比較大(因為不管丙說什麼都要猜相異)。
- 結論：投擲兩硬幣，朝上兩面相異之機率是2/3，因為一定會有正面或反面。

- 疑問：明顯有錯，因國中以來就教相同、相異機率皆為1/2。

- 最後，該文宣佈例1中的 (i)與 (ii)，及例2，其中的機率皆為1/2。
- 該文又給一例。

例3.

所有有兩個小孩且有女孩的家庭中，兩小孩皆為女孩的機率為何？

該文之結論

- 如果題目內有知道的知，或是有第三者當仲介給提示或條件，條件機率做出來都會錯。

- 反之，如果題目有強調所有的(如例3)，那麼每一情況發生的機率都相同，就可以放心的用條件機率。

- 搞了半天，是題目敘述有語病。這是文字陷阱。這題一開始的問法，應該是上題這種問法才對，只是不小心敘述錯誤，造成大問題。我想，大家往後解題，應該會多注意這種情況了。

說文解字

- 有時給的條件是以一段文字，描述一情境。如何將這段文字的內涵正確解讀，並不見得很容易。
- 若解讀錯誤，得到的條件機率當然也就不對。

分割

- 對一樣本空間Ω，事件$A_1, A_2, \cdots,$ 若滿足兩兩互斥，即

$$A_i \cap A_j = \varnothing, \forall i \neq j, \text{ 且} \bigcup_{i=1}^{\infty} A_i = \Omega,$$

便稱為Ω之一分割(partition) 。

- 有限分割 $A_1, A_2, \cdots, A_n, n \geq 2$ 。

◆定理.

設$A_1, A_2, \cdots$為樣本空間之一分割。

則對任一事件B，

$$P(B)=\sum_{i=1}^{\infty} P(B \mid A_i)P(A_i) \text{ 。}$$

◆貝氏定理.

設$A_1, A_2, \cdots$為樣本空間之一分割。

則對 $i \geq 1$，及事件B，只要$P(B)>0$，便有

$$P(A_i|B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B \mid A_j)P(A_j)}。$$

例4.

甲、乙、丙三囚犯，國王宣佈以抽籤決定釋放一位，處決另兩位。

他告訴獄卒那一位將被釋放，但要求獄卒不可先透露。

甲請獄卒透露那一位被釋放遭拒後，改問獄卒：

乙及丙中，那一位會被處決？

獄卒經一番思考，遂(誠實地)告訴甲：

乙會遭處決。

他認為這樣做並未違反國王規定：

乙、丙二人，至少有一會遭處決，這是大家都知道的，因此他並未提供甲任何有關甲會被釋放的有用資訊。

甲聽到獄卒說乙會被處決後很高興。原先他有1/3的機率遭釋放，現在因只剩他與丙了，所以機率提高至1/2。

◆獄卒與甲的分析，何者正確？

解.

令A, B, C分別表甲、乙、丙三人會被釋放的事件。

K表獄卒說乙會被處決的事件。

樣本空間

$$\Omega = A \cup B \cup C。$$

由假設

$$P(A) = P(B) = P(C) = \frac{1}{3}。$$

想求$P(A|K)$。

$$P(K)=P(K\mid A)P(A)+P(K\mid B)P(B)+P(K\mid C)P(C)$$

$$=\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}+1\cdot\frac{1}{3}$$

$$=\frac{1}{2}。$$

$$\Rightarrow P(A\mid K)=\frac{P(\text{獄卒說乙會被處決，且甲被釋放})}{P(K)}$$

$$=\frac{1/3\cdot 1/2}{1/2}$$

$$=\frac{1}{3},\ \text{機率沒有變大！}$$

獄卒所提供的資訊是否無用？

- 若丙偷聽到獄卒與甲的對話，則知他會被釋放的機率提高至2/3。
- 若乙偷聽到獄卒與甲的對話，則知他沒有活命機會。
- 乙、丙二人中，有一人被釋放之機率為2/3，若給定乙被處決，則丙便獨自擁有全部被釋放之機率2/3。

- 至於甲，被釋放之機率不會改變，還是1/3。
- 而三人被釋放之條件機率和：

$$P(A|K)+P(B|K)+P(C|K)=\frac{1}{3}+0+\frac{2}{3}=1$$ 。

- 此問題有時以其他形式出現，如汽車與山羊問題，稍後會再討論。

例5.

衛生局至高雄大學免費檢驗某疾病。假設檢驗結果有正、負兩種反應。

呈正反應，表可能有病，須至醫院進一步檢驗；

呈負反應，則衛生局便認為沒有問題。

衛生局宣稱檢驗之可靠度為90%，且平均每5,000人中，有一人患此病。

你是否願意接受檢驗？

解.

可靠度90 %的意義：

若無病，檢驗會呈負反應之機率為0.9；若有病，檢驗會正負反應之機率亦為0.9。

想知道：

檢驗呈正反應下，有病的機率，及檢驗呈負反應下，無病之機率。

以正表檢驗呈正反應，負表檢驗呈負反應。

則

$$P(\text{有病}|\text{正})=\frac{P(\text{正}|\text{有病})P(\text{有病})}{P(\text{正}|\text{有病})P(\text{有病})+P(\text{正}|\text{無病})P(\text{無病})}$$

$$=\frac{0.9\cdot 1/5000}{0.9\cdot 1/5000+0.1\cdot 4999/5000}$$

$$=\frac{9}{5008}$$

$$\approx 0.001797\text{。}$$

◆當檢驗呈負反應下，的確無病的機率：

$$P(\text{無病}\,|\,\text{負})=\frac{44991}{44992}\approx 0.999977773\text{。}$$

檢驗可靠度	90%	95%	99%	99.90%	99.99%
P(有病|正)	0.001797	0.003786	0.019419	0.166556	0.666689

表1. 患病比率1/5000，於不同檢驗可靠度下之P(有病|正)

患病比率	1/5000	1/500	1/50	1/5	1/2	4/5
P(有病|正)	9/5008	9/508	9/58	5/13	9/10	36/37

表2. 檢驗可靠度90%，於不同患病比率下之P(有病|正)

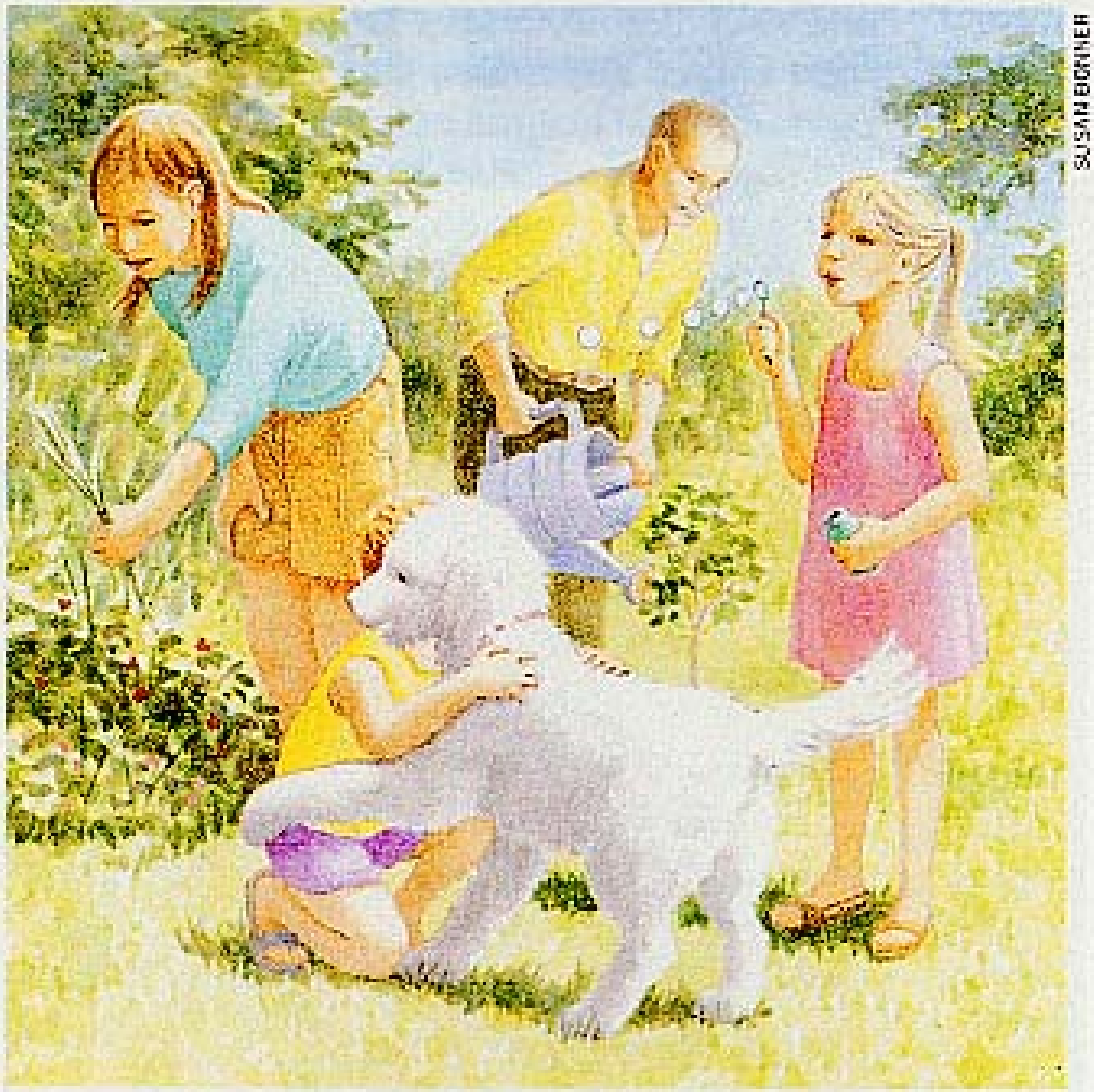
例6.

有一對夫妻剛搬進某社區，大家只知道他們有兩個小孩，並不知性別。

某日社區一管理員，見到此家之媽媽帶著家中一小孩在玩耍。

若該小孩是女孩，求此家兩小孩皆為女孩之機率。

此問題與下述問題等價嗎？



THE SMITH FAMILY

What is the probability that the kneeling child is a girl?

跪著的小孩是女孩之機率？

假設生男生女之機率皆為1/2。

解.

定義下述事件：

G_1：老大是女孩，

G_2：老二是女孩，

G ：媽媽帶著的小孩是女孩。

將女孩改為男孩，類似地定義B_1及B_2。

要求$P(G_1 \cap G_2 | G)$。

$$P(G_1 \bigcap G_2 \mid G) = \frac{P(G_1 \bigcap G_2 \bigcap G)}{P(G)} = \frac{P(G_1 \bigcap G_2)}{P(G)},$$

$$\begin{aligned} P(G) &= P(G \mid G_1 \bigcap G_2)P(G_1 \bigcap G_2) + P(G \mid G_1 \bigcap B_2)P(G_1 \bigcap B_2) \\ &\quad + P(G \mid B_1 \bigcap G_2)P(B_1 \bigcap G_2) + P(G \mid B_1 \bigcap B_2)P(B_1 \bigcap B_2) \\ &= \frac{1}{4} + \frac{1}{4}P(G \mid G_1 \bigcap B_2) + \frac{1}{4}P(G \mid B_1 \bigcap G_2) \end{aligned}$$ 。

用到

$$P(G \mid G_1 \cap G_2) = 1,\ P(G \mid B_1 \cap B_2) = 0,$$

且

$$P(G_1 \cap G_2) = P(G_1 \cap B_2) = P(B_1 \cap G_2) = P(B_1 \cap B_2) = \frac{1}{4}。$$

$$\Rightarrow P(G_1 \cap G_2 \mid G) = \frac{1/4}{1/4 + P(G \mid G_1 \cap B_2)/4 + P(G \mid B_1 \cap G_2)/4}$$

$$= \frac{1}{1 + P(G \mid G_1 \cap B_2) + P(G \mid B_1 \cap G_2)}。$$

此機率與$P(G|G_1 \cap B_2)$ 及$P(G|B_1 \cap G_2)$有關！

幾個特別的情況

<u>情況(i).</u> 假設不論兩小孩之性別為何，若只帶一小孩出門，媽媽會帶老大出門之機率為一定值p。

即設

$$P(G\,|\,G_1 \cap B_2) = p,\ P(G\,|\,B_1 \cap G_2) = 1 - p,$$

$$\Rightarrow P(G_1 \cap G_2\,|\,G) = \frac{1}{1 + p + (1 - p)} = \frac{1}{2}。$$

⇒原問題之答案為1/2，與p無關。

情況(ii). 假設當兩小孩之性別不同，則媽媽會帶女兒出門之機率為一定值q。

即設

$$P(G \mid G_1 \cap B_2) = P(G \mid B_1 \cap G_2) = q$$

$$\Rightarrow P(G_1 \cap G_2 \mid G) = \frac{1}{1+q+q} = \frac{1}{1+2q},$$

與q有關。

特例

- 設q=1。即若有兒子及女兒，媽媽一定帶女兒出門 ⇒ $P(G_1 \cap G_2 | G)$=1/3。

  此即例1-(ii)。

- 因

  看到該家媽媽帶女兒出門⇔該家至少有一女兒。

- q=1/2。即若有兒子及女兒，媽媽會帶女兒或兒子出門之機率各半 ⇒ $P(G_1 \cap G_2 | G)$=1/2。

- q=0 ⇒ $P(G_1 \cap G_2 | G)$=1。

- 除非有其他資訊，否則

  看到該家媽媽帶著一個女兒，

  不等價於

  該家至少有一女兒。

- 注意：情況(i)及情況(ii)並非樣本空間之一分割，兩情況也不互斥。
- 結論：由所給條件，無法解出本問題。
- 為什麼如此？

◆樣本空間Ω中共有8元素，包含所有型如(s_1, s_2, i)的樣本，

s_1：老大性別，

s_2：老二性別，

i ：見到媽媽帶著的小孩之排序。

◆欲知$P(\omega), \forall \omega \in \Omega$，光給$(s_1, s_2)$之機率不夠。

◆設生男生女的機率均為1/2，Ω中各元素的機率：

$$P(\{女,女,I\})=\frac{p_1}{4}, P(\{女,女,II\})=\frac{1-p_1}{4},$$

$$P(\{女,男,I\})=\frac{p_2}{4}, P(\{女,男,II\})=\frac{1-p_2}{4},$$

$$P(\{男,女,I\})=\frac{p_3}{4}, P(\{男,女,II\})=\frac{1-p_3}{4},$$

$$P(\{男,男,I\})=\frac{p_4}{4}, P(\{男,男,II\})=\frac{1-p_4}{4},$$

I, II 分別表媽媽帶著的小孩為老大或老二。

p_1 為給定兩個小孩皆為女孩下，I 發生之機率，p_2，p_3，及 p_4 依此類推。

因

$$G\bigcap G_1\bigcap B_2=\{(女,男,I)\},$$
$$G\bigcap B_1\bigcap G_2=\{(男,女,II)\},$$

$$\Rightarrow P(G\,|\,G_1\bigcap B_2)=\frac{P(\{女,男,I\})}{P(\{女,男\})}=\frac{p_2/4}{1/4}=p_2,$$

$$P(G\,|\,B_1\bigcap G_2)=\frac{P(\{男,女,II\})}{P(\{女,男\})}=\frac{(1-p_3)/4}{1/4}=1-p_3。$$

$$\Rightarrow P(G_1 \cap G_2|G) = \frac{1}{1+p_2+1-p_3} = \frac{1}{2+p_2-p_3}。$$

知道$p_2 - p_3$，才能得$P(G_1 \cap G_2|G)$。

- 當$p_2 = p_3 = p$ $\Rightarrow P(G_1 \cap G_2 \mid G) = \frac{1}{2}$,

  與情況(i)吻合。

- 當$p_2 = q$，$p_3 = 1-q$ $\Rightarrow P(G_1 \cap G_2 \mid G) = \frac{1}{1+2q}$。

  與情況(ii)吻合。

- 情況(i)及(ii)，均只為本一般情況之特例。

尚可有其他情境

(i)管理員問那位媽媽你有沒有女兒？
媽媽答有；

(ii)管理員問那位媽媽你老大是女兒嗎？
媽媽答是；

(iii)管理員見到那位媽媽帶兩小孩及一條狗在社區公園玩，其中有一女兒站著，另一小孩跪在地下，但被狗遮住，看不出性別。

- 分別對三情況，求兩小孩皆為女孩之機率。

回到一開始那三個例子

- 例1-(i)之答案顯然為1/2。
- 例1-(ii)，若無其他資訊，則假設有一個女孩，等價於家中至少有一女孩，仿例6-(ii)且q=1，可得另一小孩亦是女孩之機率，確為1/3。

- 對於例2，若假設當丙看到兩硬幣有1正面及1反面向上，便分別有1/2的機率說有一個正面及有一個反面，此對應例8中的情況(ii)，且$q = 1/2$。此時朝上的兩面相同及相異的機率皆為1/2。即在這種情況下，丙所提供的資訊沒用。
- 但若$q \neq 1/2$，則丙所提供的資訊就有用。可類似如例6的討論。

- 在例3，仍對應例6-(ii)且q=1，兩小孩皆為女孩之機率的確為1/3。
- 注意：例2中，

丙說有正面(假設若有一正面一反面，則有1/2之機率丙說有正面)

及

問丙有沒有正面？丙答有。

此二事件不同。

- 前者之機率為1/2，後者為3/4。

4. 解釋機率

- 投擲一骰子，點數1出現機率為0.1，是什麼意思？
- 數學系畢業生：

假設投擲n次，點數1出現a次，則相對頻率a/n與0.1之差的絕對值，會大於一給定正數(不管它多小)之機率，將隨著n的趨近至無限大，而趨近至0。即$\forall\varepsilon > 0$，

$$n \to \infty\text{時，}P\left(\left|\frac{a}{n} - 0.1\right| > \varepsilon\right) \to 0\text{。}$$

◆大數法則：

重複觀測，事件出現的相對頻率，會機率收斂至事件發生的機率。

疑惑

- 無限大是什麼？
- 什麼是趨近至無限大？
- 用機率來解釋機率？

- 事件出現的相對頻率，當觀測數夠大，必接近事件發生的機率？

- 事件只要機率為正，便都可能發生 ⇒ 不論觀測數再大，都不能排除很偏頗的事件發生。

- 統計學家跳出來，可以做一檢定，檢定點數1出現的機率是否真為0.1。
- 在點數1出現機率為0.1之假設下，觀測到的結果，若屬不尋常，便推翻原假設。
- 什麼是尋常？不尋常？

朱雀橋邊野草花，烏衣巷口夕陽斜，

舊時王謝堂前燕，飛入尋常百姓家。

唐 劉禹錫

烏衣巷

- 不尋常指發生的機率很小。
- 若觀測到不尋常的結果，則當初的假設就不宜接受。

- 在隨機的世界，究竟何者為真，常屬未知。
- 我們往往無法證明那件事是真實的。
- 都是假設，就看你接受那一個。

◆假設：

我願意為你赴湯蹈火。

做個假設檢定！

5. 信賴區間

- 以一區間來估計參數，並給出此區間會涵蓋該參數之機率。

- 此即區間估計，所得的區間，稱為信賴區間。

- 估計銅板出現正面之機率p。
- 取樣前，信賴區間為一隨機區間。
- 取樣後，得到一固定區間。
- 若知道p，則p會屬於該區間的機率，將不是1便是0，而不再是p了。但

若不知p呢？

◆打擊手揮棒前，可以說打出安打之機率為0.341。

◆打完不是安打就非安打，0.341已派不上用場了。

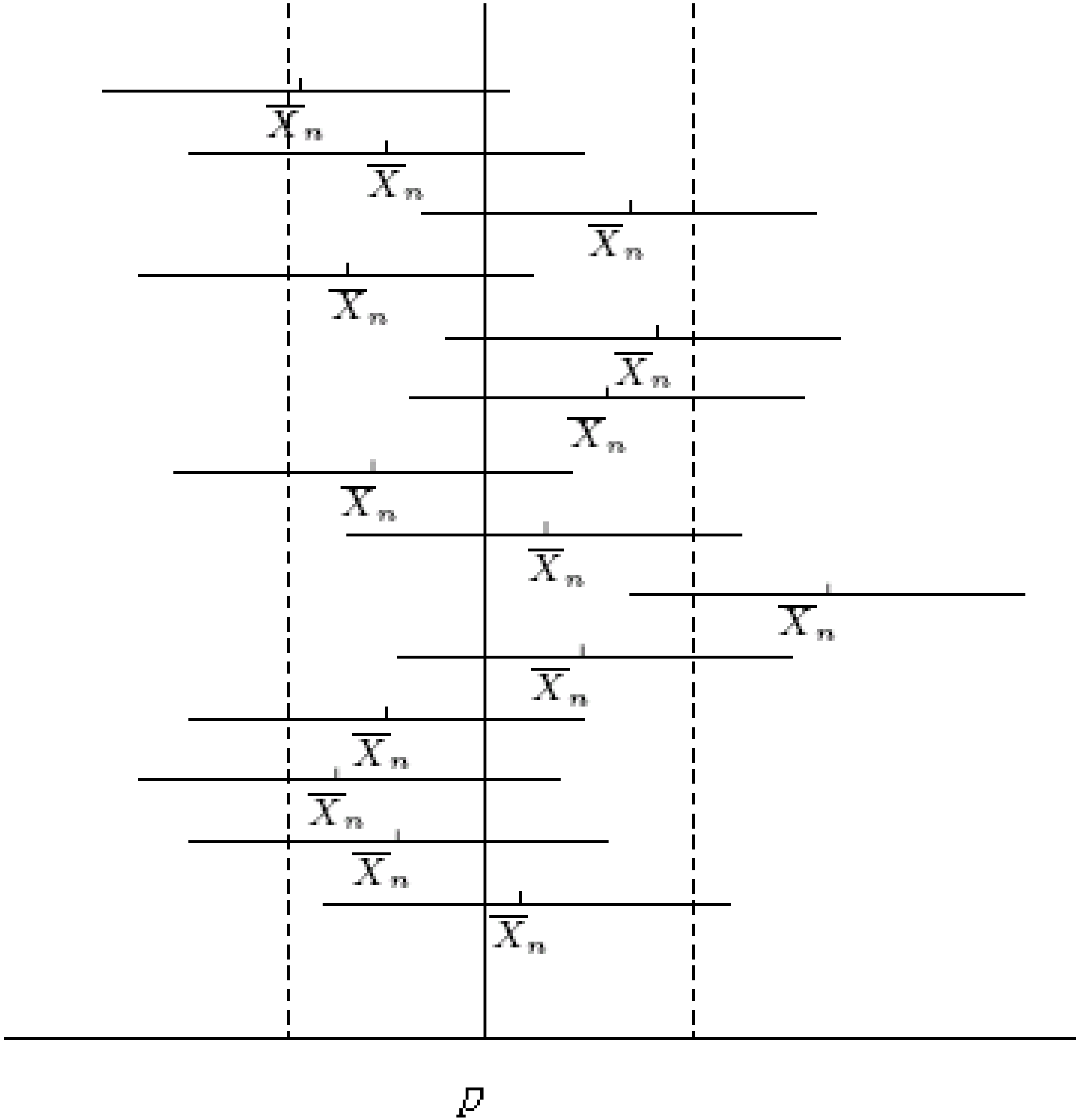
- 張無忌練成九陽真經的神功後，輕易鑽過洞穴，朱九齡卻被岩石卡住：

  這小子比我高大，他既能過去，我也必能過去，為什麼我會擠在這裡。當真豈有此理！

- 一旦發生，就不論機率大小，豈有此理又奈何？

- 課綱中說，可以亂數表模擬出現正面機率為p的銅板n次，以求得信賴區間。
- p乃事先設定，模擬所得之一固定區間，p有沒有落在其間，一看便知，如何能說該區間涵蓋p之機率為0.95？

- 那95%有何用？
- 0.95是一機率值，而機率值從來就不是只看一次的實驗結果。
- 大約可以這麼說，如果反覆實驗，而得到很多信賴區間，其中會包含p的信賴區間數，約佔全部區間數的95%。


$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$
$\overline{X}_n$

- 0.95的意義，乃如同對機率的解釋。
- 但對同一個p，如果全班40人，所得到40個95%信賴區間，其中包含p的個數未超過85%也不要太驚訝。
- 此機率約為0.01388。
- 雖不太大，但只要班級數夠多，便不難發生。

◆課綱說：

大多數學生所得的信賴區間都會涵蓋p，

並非精準的講法。

問：

一般而言，樣本數愈大，信賴區間的半徑愈小，因此所得信賴區間似乎便愈不容易包含p。對此你有何看法？

6. 情境解讀

- 機率與我們生活息息相關。若能善用機率，將有助於在隨機世界中，更精準的做決策。
- 但往往機率應用不易，得到的機率值，常被認為是錯的。
- 而且還眾說紛紜，各提出不同的機率值。
- 原因何在？

- 數學課程中，會遇到所謂應用題。
- 寫出數學式子後，就是解數學了。
- 這時便可拋開原先那段冗長的敘述。
- 但在機率裡，看似簡單的情境，因解讀不同，會導致南轅北轍的結論。


J
A
Jim Sturgess
Kate Bosworth
Laurence Fishburne
Kevin Spacey
21
FIVE STUDENTS WHO
CHANGED THE GAME FOREVER.
SPRING 2008

- 在決勝21點(21, 2008)中，

  有3扇門，其中1扇門後有汽車，另兩扇門後為山羊。你選擇第1扇門後，主持人打開第2扇，見到山羊。問你這時該不該換第3扇門？有位學生答

  Yes, because my chance of getting the car will increase from 33.33% to 66.67% by switching from door 1 to door 2.

◆正確的講法應該是：

若主持人事先知道汽車在那扇門後，則他會打開1扇門後是山羊的門，假設是第2扇。

這時若換選第3扇門，則得到汽車的機率，由1/3增加為2/3。

但隱含做一假設。即若第2、3扇門後皆為山羊，則隨機地打開第2或第3扇門。

- 若主持人事先不知汽車在那1扇門後，只是隨機地自第2及第3扇門中，挑一扇打開，且剛好門後是山羊，則便不用換，因換與不換，得到汽車的機率，皆為1/2。

◆可以有更一般的假設。當第2或第3扇門後皆是山羊，主持人分別以 q 及$1-q$ 的機率，打開第2或第3扇門。則換選第3扇門，得到汽車的機率成為

$$\frac{1}{1+q}$$ 。

- 此機率會受主持人是如何打開第2扇門的影響！
- 由於對 $0 \le q \le 1$ ，

$$\frac{1}{1+q} > \frac{1}{3} \ ,$$

所以換，仍是較好的選擇。

- 在例6中，如何將

  見到此家之媽媽，帶著家中一女孩，

  轉化為適當機率空間中的事件？
- 究竟如何帶小孩出門？
- 該事件並不等同於

  此家至少有一女孩！

- 另一常出現的例子：

  平面上有一單位圓，隨機地畫一條弦，求弦長大於此圓內接等邊三角形之邊長機率。

- 如何隨機畫弦？
- 由1至n的n個正整數中，隨機地取一數？
- 自區間[0,1]中隨機地取一數？

◆在處理機率問題時，情境要定義清楚。

◆即機率空間要明確給出，否則將導致各說各話。

◆有時情境較簡單，如

投擲一公正骰子，求點數大於4之機率。

- 當對情境有疑義時，就要如莊子在秋水篇講的，

請循其本，

把機率空間調出來。

- 有如政治或社會上，遇到有重大爭議時，就要祭出憲法，並由大法官解釋。
- 機率中一些獨特的概念，如條件機率，獨立性，及隨機取樣等，也是應用機率時，得謹慎留意的。

謝謝各位！