1 閒聊機率與統計⋯與其請問上帝是否擲骰子，不如研究上帝怎麼擲骰子

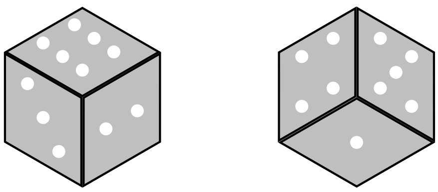
數學經文

頭腦是一切不真實的集合，它總是思考著讓所有的事情都完美。但人生幾時完美呢？是故，每當頭腦被大量使用，想做完美的事時，它總是站在選擇錯誤的那一面，因為它選擇了完美的冒牌貨或替代品⋯世故或是沒稜沒角，像球或圓一樣的圓滑。只有你相信生命只能美麗，沒辦法完美，那頭腦的選擇才會是正確的選擇。當你達到“不用選擇是唯一且最好的選擇”時，你就進入人生這粒骰子的最高境界了。

十賭九輸，十墓九空，雖然是兩句成語，但它們反映了人們貪婪的天性。所以很多遊戲，它們發生的機率並不如想像中那樣簡單。如何巧妙地發現它們，又科學地計算它們，確實是對人類的一大挑戰。

題目： 有甲、乙兩粒均勻的正四面體骰子，在甲骰子的四面分別寫上 $1, 2, 2, 3$ 四個數字；在乙骰子的四面分別寫上 $1, 3, 3, 5$ 四個數字。將甲、乙兩粒骰子同時投擲時，其點數和的機率分佈為何？（考慮壓在桌面上那一面的數字和）

丟一枚公正的銅板，不是正面就是反面朝上，而且其機率都一樣是 $\frac{1}{2}$；投擲一粒均勻的骰子（正立方體），也只會出現六種情況，點數 $1, 2, 3, 4, 5, 6$ 朝上的機率都一樣是 $\frac{1}{6}$。



這兩個結果可以從多次的實驗得到驗證，不需要求神問卜，請問上帝。日常生活中，類似這樣的機率活動有許多，如何得到各種

情況發生的機率分佈是人們很感興趣的一件事。與其求見上帝，問他這機率分佈是多少，不如善用手邊的數學知識，親手研究這機率分佈來得實際。

1.1 利用多項式的乘法運算擲骰子⋯均勻骰子

將甲、乙兩粒骰子同時投擲，那麼甲、乙兩粒骰子的點數和分佈次數可以由底下的多項式乘法看出

$$\begin{aligned}&\left(x^1+x^2+x^2+x^3\right)\left(x^1+x^3+x^3+x^5\right)\\&=\left(x^1+2x^2+x^3\right)\left(x^1+2x^3+x^5\right)\\&=x^2+2x^3+3x^4+4x^5+3x^6+2x^7+x^8.\end{aligned}$$

同時投擲甲、乙兩粒骰子的樣本空間有 $4\times4=16$ 種，而點數和的分佈次數可由上述多項式的係數決定。例如點數和 4 的情況有 3 種，即 $(\text{甲},\text{乙})=(1,3),(1,3),(3,1)$ 三種，它們分別對應到上述多項式的乘積展開式

$$x^1\cdot x^3+x^1\cdot x^3+x^3\cdot x^1=3x^4.$$

所以同時投擲甲、乙兩粒骰子的點數和機率表為

點數和	2	3	4	5	6	7	8
機率	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$\boxtimes$

練習 1 如果甲、乙兩粒均勻的正四面體骰子的四個面都是寫上 $1,2,3,4$ 四個數字，那麼將甲、乙兩粒骰子同時投擲時，其點數和的機率分佈為何？當你完成此問題時，請將答案與本章題目的答案比較一下，你的發現是什麼呢？對於你的發現，你可以用多項式的性質，講得更深入嗎？

練習 2 有甲、乙兩粒均勻的正六面體骰子，在甲骰子的六面分別寫上 $1,3,4,5,6,8$ 六個數字；在乙骰子的六面分別寫上 $1,2,2,3,3,4$ 六個數字。將甲、乙兩粒骰子同時投擲時，其點數和的機率分佈為何？

二項式定理是說 $(x+y)^n$ 的展開式可以寫成

$$(x+y)^n=\mathrm{C}_0^n x^n y^0+\mathrm{C}_1^n x^{n-1}y^1+\mathrm{C}_2^n x^{n-2}y^2+\cdots+\mathrm{C}_{n-1}^n x^1 y^{n-1}+\mathrm{C}_n^n x^0 y^n$$

的形式。這個恆等式的係數 $\mathrm{C}_k^n=\frac{n!}{(n-k)!k!}$。若將這些係數依 n 值的大小，由上而下；依 k 值的大小，從左至右排列，則可以排列成如下的楊輝三角：

$$\begin{array}{ccccccccc} &&&&1&&&& \\ &&&1&&1&&& \\ &&1&&2&&1&& \\ &1&&3&&3&&1& \\ 1&&4&&6&&4&&1 \end{array} \qquad \begin{array}{ccccccccc} &&&&\mathrm{C}_0^0&&&& \\ &&&\mathrm{C}_0^1&&\mathrm{C}_1^1&&& \\ &&\mathrm{C}_0^2&&\mathrm{C}_1^2&&\mathrm{C}_2^2&& \\ &\mathrm{C}_0^3&&\mathrm{C}_1^3&&\mathrm{C}_2^3&&\mathrm{C}_3^3& \\ \mathrm{C}_0^4&&\mathrm{C}_1^4&&\mathrm{C}_2^4&&\mathrm{C}_3^4&&\mathrm{C}_4^4 \end{array}$$

〈楊輝三角對照表〉

由二項式定理知道：$(x+y)^{n+1}$ 展開式中 $x^{(n+1)-r}y^r$ 項的係數為 C_r^{n+1}；而把 $(x+y)^{n+1}$ 寫成 $(x+y)^n(x+y)$ 的乘積，將 $(x+y)^n$ 展開並計算其與 $(x+y)$ 的乘積，得

$$\begin{aligned}&\left(\cdots+\mathrm{C}_r^n x^{n-r}y^r+\mathrm{C}_{r-1}^n x^{n+1-r}y^{r-1}+\cdots\right)(x+y)\\ &=\cdots+\left(\mathrm{C}_r^n+\mathrm{C}_{r-1}^n\right)x^{n+1-r}y^r+\cdots.\end{aligned}$$

因此，二項式定理的係數滿足

$$\mathrm{C}_r^{n+1}=\mathrm{C}_r^n+\mathrm{C}_{r-1}^n,$$

也就是說，楊輝三角上的任一個數等於其上方左、右兩數的和。

二項式定理（代數形式）與楊輝三角（幾何模型）被廣泛使用來解釋許多數學問題，接下來舉兩個例子作說明：

① （彈珠台）如果彈珠台是一個三角形，當彈珠從頂點下降一格時，有一半的機率向右走、另一半的機率會向左走，那麼該如何描述彈珠的落點分佈呢？楊輝三角就是一個很好的模型：將三角形的彈珠台想成楊輝三角，當你投下 1 粒彈珠時，楊輝三角的第 1 列的數字 1 就是它的分佈；當你投下 2 粒彈

珠時，楊輝三角的第 2 列的數字 $1,1$ 告訴你，它的分佈是左、右各 1 粒；當你投下 4 粒彈珠時，楊輝三角的第 3 列的數字 $1,2,1$ 告訴你，它的分佈是左、右各 1 粒，中間有 2 粒，也就是說，彈珠跑到左、右兩側的機率分別是 $\frac{1}{4}$，而落在正中央的機率是 $\frac{1}{2}$；依此類推，當你投下 16 粒彈珠時，楊輝三角的第 5 列的數字 $1,4,6,4,1$ 告訴你，應該有 6 粒彈珠會落在正中央。從這過程發現，楊輝三角是解釋彈珠台遊戲的最佳幾何模型。

② （分割線段）將一條一公尺長的線段，依 $1:3$ 的比例分割成兩條小線段，再將這兩條小線段，分別依 $1:3$ 的比例各分割成兩條小線段，依此類推，在五次分割之後，會製造出 32 條小線段。你如何知道這 32 條小線段的長度及同一長度的線段各幾條呢？拜二項式定理所賜，可以利用

$$\left(\frac{1}{4}+\frac{3}{4}\right)^5 = \mathrm{C}_0^5\frac{1}{4^5}+\mathrm{C}_1^5\frac{3^1}{4^5}+\mathrm{C}_2^5\frac{3^2}{4^5}+\mathrm{C}_3^5\frac{3^3}{4^5}+\mathrm{C}_4^5\frac{3^4}{4^5}+\mathrm{C}_5^5\frac{3^5}{4^5}$$

展開式中的每一項與該項係數來解釋這問題。將展開式中係數算出，得

$$\left(\frac{1}{4}+\frac{3}{4}\right)^5 = 1\times\frac{1}{4^5}+5\times\frac{3^1}{4^5}+10\times\frac{3^2}{4^5}+10\times\frac{3^3}{4^5}+5\times\frac{3^4}{4^5}+1\times\frac{3^5}{4^5}.$$

這個式子告訴我們：線段長度有

$$\frac{1}{4^5},\quad \frac{3^1}{4^5},\quad \frac{3^2}{4^5},\quad \frac{3^3}{4^5},\quad \frac{3^4}{4^5},\quad \frac{3^5}{4^5}$$

等六種，每種長度有

$$1,\quad 5,\quad 10,\quad 10,\quad 5,\quad 1$$

條，一共

$$32 = 1+5+10+10+5+1$$

條線段。由此發現，分割線段可透過二項式定理的代數形式闡釋。

機率分佈的一個重要性質就是所有事件的機率總和是 1，例如丟一枚公正的銅板，出現正面事件的機率是 $\frac{1}{2}$，而出現反面事件的機率也是 $\frac{1}{2}$，這兩個數字的和 $\frac{1}{2}+\frac{1}{2}$ 就是 1。投擲一粒均勻骰子也有同樣的性質，也就是說出現點數 $1,2,3,4,5$ 與 6 的機率都是 $\frac{1}{6}$，而它們的和 $\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}$ 也是 1。

如果你會利用對數性質 $\log_{10} a+\log_{10} b=\log_{10} ab$ 與 $\log_{10} 10=1$ 的話，你將發現底下十個介於 0 與 1 之間的數字和也是等於 1：

$$\log_{10}\left(1+\frac{1}{1}\right)=\log_{10}\frac{2}{1};$$
$$\log_{10}\left(1+\frac{1}{2}\right)=\log_{10}\frac{3}{2};$$
$$\log_{10}\left(1+\frac{1}{3}\right)=\log_{10}\frac{4}{3};$$
$$\log_{10}\left(1+\frac{1}{4}\right)=\log_{10}\frac{5}{4};$$
$$\log_{10}\left(1+\frac{1}{5}\right)=\log_{10}\frac{6}{5};$$
$$\log_{10}\left(1+\frac{1}{6}\right)=\log_{10}\frac{7}{6};$$
$$\log_{10}\left(1+\frac{1}{7}\right)=\log_{10}\frac{8}{7};$$
$$\log_{10}\left(1+\frac{1}{8}\right)=\log_{10}\frac{9}{8};$$
$$\log_{10}\left(1+\frac{1}{9}\right)=\log_{10}\frac{10}{9}.$$

班佛發現日常活動中，有許多以上述十個數字為機率分佈的機率模型。最有名，也最容易被誤解的一個模型就是“隨手寫下正整數”這個模型。如果請全班或全校的每位學生隨手寫下一個正整數，那麼這些正整數的首位數字（例如 2489 的首位數字就是 2）的分佈機率為何呢？很多人可能把這遊戲想成跟擲骰子一樣，認為首位數字是 1 的會佔 $\frac{1}{9}$，首位數字是 $2,3,4,\cdots,9$ 的也都各

佔 $\frac{1}{9}$。事實並不是這樣的，班佛研究發現，學生所寫正整數的首位數字之分佈並不均勻，首位數字是 $d(1 \le d \le 9)$ 的比率約為 $\log_{10}\left(1+\frac{1}{d}\right)$。

另一個有趣的發現：當妳翻閱圖書館的藏書時，你會發現每本書的每一頁被翻閱的次數不相同，如何辨識呢？很簡單，只需看書本側邊黑白的程度，越黑的哪幾頁肯定被閱讀最多次。如果你去統計的話，你將發現頁碼的首位數字是 $d(1 \le d \le 9)$ 的被翻閱的比例是 $\log_{10}\left(1+\frac{1}{d}\right)$。

像這兩個例子，當你研究與正整數相關的首位數字分佈，如果首位數字 $d(1 \le d \le 9)$ 的分佈比例為 $\log_{10}\left(1+\frac{1}{d}\right)$，就稱它符合班佛法則。接下來列舉一些符合班佛法則的事件：

① （存摺存款的首位數字）郵政總局如果統計每位存款人的存款錢數的首位數字，他可能會驚訝的發現，這個首位數字的分佈並不均勻，反而符合班佛法則。也就是說，一百多元，一千多元，一萬多元，十餘萬元，百餘萬元，千餘萬元，⋯等的存款戶約佔 $\log_{10} 2 = 0.3010$ 左右。

② （河流流過的區域面積）統計全世界所有河流所流經的區域面積，這面積的首位數字也會符合班佛法則。

③ （財務報表的首位數字）你的公司想作假帳，逃漏稅嗎？那你得懂班佛法則才有辦法得逞。公司每日進進出出的貨物很多，產生的數據自然龐大，國稅局如何監控公司是否數據造假，逃漏稅呢？這都要歸功於班佛法則，根據統計，如果公司正常運作下，那麼所產生的數據的首位數會符合班佛法則。國稅局為了節省時間與人力，只需查核該公司的數據的首位數字是否符合班佛法則，便可知道該公司的數據是否造假了。

④ （都市人口數字的首位數字）每個都市的人口數的第一位數字也會符合班佛法則。

⑤ （2^n 的首位數字）將數列 $\langle 2^n \rangle$ 的前一千項列出，統計這一千個數字之首位數，你將發現它們的分佈完全符合班佛法則的比例。不止一千項，列出一萬項或更多項，也都會符合班佛法則。

⑥（費氏數列的首位數字）費氏數列 $\langle f_n \rangle$ 定義為

$$f_1 = f_2 = 1, f_{n+2} = f_{n+1} + f_n (n \geq 1).$$

數學家猜測費氏數列 $\langle f_n \rangle$ 的首位數字符合班佛法則。

⑦（質數的首位數字）將質數依小到大排列 $2, 3, 5, 7, 11, \cdots$。數學家猜測它們的首位數字會符合班佛法則，但尚未被證實。

班佛發現的首位數字分佈現象不僅與人類選取數字的心裡相關，也與自然界的多種分佈一致。與班佛法則一樣出名的另一個法則是齊普夫法則，它是美國社會學家齊普夫發現的：齊普夫的學生很有耐心地去計算《哈姆雷特》這本書中所出現的字，並且依照它們的出現頻率做降冪排列。頻率最高的字是“the”，共出現 1087 次，緊跟在後面的則是“and”這個字。齊普夫發現，它們的分布狀況遵行了一個數學的金科玉律。當一個人寫書時，將整本書所出現的字，依照它們出現頻率做降冪排列，出現頻率第 n 高的字出現次數與出現頻率最高的字出現次數的比值，可以用數學公式

$$\frac{c}{n^k}$$

來表示，這裡的常數 c, k 與作者寫這本書時的心裡素質相關。

齊普夫法則不僅寫書適用，也適用在許許多多的日常生活中，網頁瀏覽人數統計分佈就是一個例子。如果將某特別類別的網頁每日瀏覽人數，依多到少做統計，那麼第 n 高的網頁每日瀏覽人數與最高的網頁每日瀏覽人數的比值，也會符合齊普夫法則。

例題 **1** 一位作家出版一本書，書中出現頻率最高的字一共出現 1000 次，第 5 與第 10 高的字的出現次數分別為 40 與 10 次。若該書符合齊普夫法則，則求齊普夫常數 c, k 的值。

〔解〕由齊普夫法則得到

$$\frac{c}{5^k} = \frac{40}{1000};$$
$$\frac{c}{10^k} = \frac{10}{1000}.$$

將兩式相除得到

$$2^k = 4 \Rightarrow k = 2.$$

將 $k = 2$ 代入第一式得

$$\frac{c}{5^2} = \frac{40}{1000} \Rightarrow c = 1.$$

故齊普夫常數 $c = 1, k = 2$。 ☒

練習 **3** 某特別類別的網頁每日瀏覽人數統計表符合齊普夫法則。若將每日瀏覽人數第 n 高的瀏覽人數記為 $f(n)$ 人，把 $\log_{10} n$ 當 x 座標，$\log_{10} f(n)$ 當 y 座標，則證明這些數對 (x, y) 座落在一條直線上。

練習 **4** 帕累托有一個法則說：在一個國家裡，年收入超過 m 元（美金）的人數佔全體人數的

$$\frac{c}{m^k},$$

這裡的常數 k, c 與這個國家有關。有一個國家年收入的中位數是 14400 元（美金），而且知道該國家帕累托常數 $k = 0.5$。

① 求帕累托常數 c 的值。

② 在該國家，年收入超過多少時，可以擠進高所得的族群。（註：年收入在前 20% 的人士算高所得族群）

1.4 利用夢境擲骰子⋯夢的奧秘

在丹增嘉措活佛所著的書《探索夢的奧秘》中，提到愛迪生如何利用夢境的故事：「我們剛睡醒的時候，夢境的回憶猶歷歷在目，這時尚未完全醒過來，稱為半睡半醒狀態。有人認為半睡半醒狀態是一種天才狀態，有很深刻的創造力，就像大發明家愛迪生發明電燈泡的創造力一樣。愛迪生很看重這種半睡半醒狀態，每次在潛心研究發明時，會運用自己的技巧，進入半睡半醒狀態。他會端坐在椅上，利用放鬆與冥想技巧進入寤寐之間的潛意識狀態。

愛迪生手握兩個鐵球，手心向下，然後舒服坐在椅子上，手肘靠近扶手，雙手下方的地面上放有鐵盤。當愛迪生入睡時，手就放鬆而自然張開，鐵球就自然掉落，碰擊鐵盤而發出聲響，吵醒自己，然後一再重複這個動作。」事實上，在夢學領域從事研究的學者指出："夢境約有百分之七十的內容是象徵和隱喻組成，百分之十五是實際記憶，最後的百分之十五是變形和偽裝。"

例題 **2** 從過去盜墓者間的資料得知，每盜十個墓，約有九個是空墓，早就被盜走陪葬品。試問：一位盜墓者至少應該盜幾個墓，才有過半的機會，能盜得陪葬品。

〔解〕令盜墓者盜 n 個墓時，會有過半的機會，盜得陪葬品。也就是說，盜 n 個墓時，都是空墓的機率少於一半，即

$$0.9^n < 0.5.$$

將兩邊取對數得到

$$n(\log_{10} 9 - 1) < -\log_{10} 2.$$

利用 $\log_{10} 9 = 2\log_{10} 3 = 2 \times 0.4771 = 0.9542, \log_{10} 2 = 0.3010$ 得到

$$0.0458n > 0.301 \Rightarrow n > 6.572.$$

故盜墓者至少應該盜 7 個墓，才有過半的機會，能盜得陪葬品。 $\boxtimes$

1.5 利用累積的經驗擲骰子⋯比賽制度的選擇

運動競賽或棋藝比賽常採取多場的比賽制度，如三戰兩勝，五戰三勝或者是七戰四勝等。運動選手或下棋者應該參加或採用那個比賽制度，對自己最有利呢？讓我們計算看看！

例題 **3** 你和你的好朋友經常打桌球，根據過去的經驗得知：在每一局中，你獲勝的機率為 p（$\frac{1}{2} < p < 1$）。今天你的好朋友想跟你來一場比賽，至於是採三戰兩勝或者是五戰三勝的比賽制度，由你決定。問：你應採取那一種比賽制度才有較高的勝算。

〔解〕設在三戰兩勝及五戰三勝中，贏得比賽的機率分別為 P_3 與 P_5。經由計算得到

$$
\begin{aligned}
P_3 =& p^2 + \left(\mathrm{C}_2^3 - 1\right) p^2(1-p) \\
=& p^2 + 2p^2(1-p); \\
P_5 =& p^3 + \left(\mathrm{C}_3^4 - \mathrm{C}_3^3\right) p^3(1-p) + \left(\mathrm{C}_3^5 - \mathrm{C}_3^4\right) p^3(1-p)^2 \\
=& p^3 + 3p^3(1-p) + 6p^3(1-p)^2.
\end{aligned}
$$

因為 $\frac{1}{2} < p < 1$，所以 $-1 < 1-2p < 0$ 及

$$
\begin{aligned}
P_3 - P_5 =& \left(p^2 + 2p^2(1-p)\right) - \left(p^3 + 3p^3(1-p) + 6p^3(1-p)^2\right) \\
=& 3p^2(1-p)^2(1-2p) < 0.
\end{aligned}
$$

故選五戰三勝贏得比賽的機率較高。 ⊠