

## 52 班佛法則…不會丟擲骰子的上帝？

班佛法則是物理學家法蘭克·班佛所提出的一則數學法則，而且是與數字相關的法則。我是輾轉從中研院統計所的一位研究員那兒得知這法則的，這對從事數論研究的我來說，真是有點不好意思。這法則的故事是這樣的，在 1881 年，數學家西蒙·紐康觀察到數學家們所使用的對數表書籍裡面，前面的頁面老是比後面的頁面部份髒了很多，原來是翻閱前面頁數的人老是比翻閱後面頁數的人多的多。紐康提出了一則與數字出現有關的機率假設。在 1938 年，班佛根據紐康的假設，做了廣泛的資料蒐集，其中包括統計 335 條河流所流過的區域面積。在歸納這些資料後發現，財務報表中的數字，以「1」開頭的數字出現的機率為 30.1% 最大，而「2」出現的機率為 17.6%，依此遞減，到「9」出現的機率只剩下 4.6% 最小。

美國麻薩諸塞州的數字樂透遊戲是讓賭客猜一個即將搖出的四位數整數。麻州樂透天天開獎，完全猜中此四位數整數（數字次序必須相同）的賭客平分彩金，例如 2007 年 10 月 23 日至 27 日的開獎號碼與每人獲得彩金如下表：

日期	號碼	獎金
2007/10/27	2618	\$ 5417
2007/10/26	8984	\$ 5538
2007/10/25	2535	\$ 5382
2007/10/24	8862	\$ 5170
2007/10/23	8307	\$ 4762

想要瞭解麻州樂透要怎麼簽才會有最大獲利，就必須清楚人類對數字有哪些偏好。班佛法則告訴我們，第一位數字越小，人們對它的偏好越大。例如，可以做個實驗，叫每位學生寫出一個四位數的正整數，並統計這些正整數的第一位數字的分佈狀況。

---

如果麻州樂透每位數字出現的機會均等，而賭客也都依人類對數字的偏好簽賭，那麼長久下來簽首位數字為多少的賭客們獲利最大呢？

---

不僅麻州樂透以四位數當頭獎中獎號碼，世界彩金最高的西班牙「肥王」樂透也是以五位數當頭獎中獎號碼。肥王樂透是在每年的聖誕節前夕開出，2007年所開出的頭獎五碼為「06381」，彩金高達二十二億歐元（逾一千億台幣），這種傳統可追溯至1812年。

班佛法則是在談人類對首位數字的一種迷思，或者說，一種不可思議的現象。現在我們用對數函數來具體描述這個現象。給定數字  $d(d=1,2,3,4,5,6,7,8,9)$ ，紐康與班佛認為，自然界的很多種類的數據，如人們隨性所寫出的數，銀行帳簿的存款數據，河流的區域面積，都市的人口數據，公司的收支數據等，都會有底下的分佈比例：這些數據中，第一位數字為  $d$  所佔的比例約為

$$\log_{10}\left(1+\frac{1}{d}\right).$$

這個比例規則跟統計所用的單位或進位方法無關，也就是說，存款用新台幣或美元來表示，都符合班佛法則，河流的區域面積用平方英里或者平方公里也一樣；即使不用十進位，改用其它進位系統統計數據，班佛法則依然適用。

關於班佛法則，可以進一步延伸到首兩位數字的分佈，它的公式為：不作假帳的公司之財報數據中，首兩位數字為  $d_1d_2$  所佔的比例約為

$$\log_{10}\left(1+\frac{1}{d_1d_2}\right).$$

例如，數據為49...所佔的比例約為

$$\log_{10}\left(1+\frac{1}{49}\right)=\log_{10}\frac{50}{49}=2-\log_{10}2-2\log_{10}7\approx 0.88\%.$$

現在讓我們利用班佛法則來討論麻州樂透：根據法則，賭客簽注第一位數字為1的佔

$\log_{10}\left(1+\frac{1}{1}\right)\approx 0.3010$ ，約三成；簽注第一位數字為2的佔  $\log_{10}\left(1+\frac{1}{2}\right)\approx 0.1761$ ，約一成

八；簽注第一位數字為3的佔  $\log_{10}\left(1+\frac{1}{3}\right)\approx 0.1249$ ，約一成二；...；簽注第一位數字為

9 的佔  $\log_{10}\left(1+\frac{1}{9}\right) \approx 0.0458$ ，約零點四六成。顯然，簽注第一位數字為 9 的賭客最少，

但猜中的機會是一樣的。所以猜中的話會平分較多的彩金。

班佛法則被廣泛應用在政府查漏稅上，稽查單位只需檢查公司收支數據的第一位數字分佈是否符合班佛法則就可以認定公司是否作假帳。而統計數據的第一位數字分佈可以經由電腦解決。例如，我們國家所發生的總統府國務機要費懸案，有人就建議將所有單據統計其第一位數字，看是否符合班佛法則。可惜的是，班佛法則仍未被我國視為可用或可靠的法則，所以無法律效力。