二〇〇五年台灣資訊科技、數學科普、娛樂數學融入數學教學研討會

Hello Kitty 的數學

游森棚教授*
高雄大學應用數學系

August 25, 2005

Hello Kitty 磁鐵蒐集是前陣子的全民熱, 不論以行銷學, 以社會學, 以心理學的角度來看, 都值得探討. 但是其實這其中有很多有趣的數學呢! 而數學的立論基礎是最結實的. 這篇短文我們來談談 Hello Kitty 中的數學.

1 蒐集一整套

Hello-Kitty 磁鐵'按照年代'一共有 31 張, 後來追加 3 張'隱藏版', 最後又出現一個'Hello Kitty 遊台灣' 7 張, 所以全套一共是 41 張. 只要消費 77 元就可以得到一張磁鐵. 但是拆開之前不知道這一次拿到的是否已經有了, 爲了蒐集一整套各不相同的磁鐵, 就要不斷的消費. 相當多的商品採取這樣的行銷策略, 比如 "扭蛋", "魔法牌", "口袋怪獸卡片".

大家最關心的是, 到底要花多少元可以蒐集到全套?

當然, 在運氣極端好的狀況之下, 只要買 41 次就可以蒐集到全套, 雖然這顯然是很不可能的事. 在運氣極端背的狀況之下, 買 100000 次也蒐集不到全套, 這看來也是很不可能的事. (在其他商品有可能會發生. 比如 "青眼白龍", 或 "被封印的黑暗大法師"...). 因此我們要先相信廠商, 假設 7-11 製造每一種磁鐵的數量都一樣, 然後所有這些磁鐵很隨機地散佈在全省各個角落. 意思是說, 每一次拿到不同磁鐵的機率是一樣的. 我們假設一套有 n 個.

*'數學科普' 單元演講內容, 台灣省教師研習中心.

因此我們問, **平均**要花多少錢可以蒐集到全套?

令 X 為蒐集到一套時所要買的次數. 我們要算 X 的期望值 $E(X)$ (“期望值” 就是“平均”). 每買一次, 如果拿到的是之前沒有的磁鐵, 我們就說這次的購買 “成功”. 因此成功了 n 次就是已經蒐集了一套了. 令 $X_i, 0 \le i \le n-1$ 表示從第 i 次“成功”之後, 一直到第 $i+1$ 次“成功”時所要購買的次數. 比如說, X_5 表示從已經有五個不同的磁鐵開始算起, 直到拿到第六個不同的磁鐵時一共要買的次數. 顯然, 蒐集一套要買多少次$=$ 蒐集到第一個要買多少次 $+$ 蒐集到第二個要買多少次 $+\ldots+$ 蒐集到最後一個要買多少次, 就是說

$$X = X_0 + X_1 + X_2 + \ldots + X_{n-1}.$$

我們要求 $E(X)$. 但 $E(X)$ 是線性的 (即, 加減可以拆開), 因此

$$E(X) = E(X_0) + E(X_1) + E(X_2) + \ldots + E(X_{n-1}),$$

(這個式子意思是說, “平均蒐集一套要買多少次” $=$ “平均蒐集到第一個要買多少次 $+$ 平均蒐集到第二個要買多少次 $+\ldots+$ 平均蒐集到最後一個要買多少次”).

所以只要把每一個 $E(X_i)$ 算出來再相加就可以. 用例子來說明. 比如要算 $E(X_5)$: 此時手上有 5 個不同的了, 要算平均要買多少次可以拿到第 6 個. 這樣想: 因為每一個磁鐵出現的機率是相同的, 因此每次有 $\frac{5}{n}$ 的機率會拿到舊的, $\frac{n-5}{n}$ 的機率會拿到新的. 換句話說,

買 n 次, 平均可以拿到 5 的舊的, $n-5$ 個新的.

再換句話說,

每買 1 次, 平均可以拿到 $\frac{n-5}{n}$ 個新的.

再換句話說,

每拿到 1 個新的平均要買 $\frac{n}{n-5}$ 次.

因此 $E(X_5) = \frac{n}{n-5}$. 同理,

$$E(X_i) = \frac{n}{n-i}.$$

因此

$$E(X) = E(X_0) + E(X_1) + E(X_2) \ldots + E(X_{n-1})$$

$$
\begin{aligned}
&= \frac{n}{n-0}+\frac{n}{n-1}+\frac{n}{n-2}+\ldots+\frac{n}{n-(n-1)} \\
&= n(\frac{1}{n}+\frac{1}{n-1}+\frac{1}{n-2}+\ldots+\frac{1}{1}) \\
&= n(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}) \\
&:= nH_n.
\end{aligned}
$$

所以我們算出來, 如果一套有 n 個, 在每個磁鐵出現的機率都相同的條件下, 平均我們要購買 $n(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n})$ 次可以蒐集到整套. 這是有趣的結果呢!

2 分析

我們到底算出什麼? 現在來分析一下. 已經知道在運氣最好的情形下, 蒐集一套 41 個 Hello Kitty 要花掉 $77 \times 41 = 3157$ 元. 但是我們剛剛算出, “平均” 要花 $77 \times 41(1+\frac{1}{2}+\ldots+\frac{1}{41})$ 元才能蒐集到一套. 這到底是多少? 利用 Maple 軟體打入

$$77 \star \text{evalf}(41 \star \text{sum}(\text{'1/k'}, \text{'k'}=1..41));$$

得到 13584.36. 這是說, 平均要花 13584 元才能蒐集到一套!!

假設全台灣有 50000 個人在蒐集 Hello Kitty, 平均一共消費 $50000 \times 13584 = 679,200,000$ 元, 六億七千九百萬. 因此報紙上說 “Hello Kitty 帶來十億商機” 絕非誇張之詞. 現在我們有數學的佐證了, 可以看到不管當初是誤打誤撞或是精心設計, 這個企畫實在是太聰明了. (事實上, 根據媒體報導, 短短三個月內一共送出了四千萬個磁鐵. 老師們可以算算看消費額是多少.)

我們不一定要用 Maple 來算. 大一微積分裡有一個重要的定理,

$$H_n \sim \ln n.$$

因此利用對數表是可以估計的.

3 明智? 不明智?

因為 7-11 總公司給各分店的磁鐵是 500 個包裝成一包, 理論上裡面會有完整的一整套. 500 個全買要 $500 \times 77 = 38500$ 元. 因此前陣子有媒體這樣報導:

“這位消費者爲了確保自己能湊到全套完整版的 Hello Kitty 磁鐵, 經過仔細精算後, 一口氣花 下 38500 元買酒及高單價預購商品, 就能買到一整盒裡面共有 500 枚的 Hello Kitty 磁鐵. Hello Kitty 果然能令人爲她瘋狂!”

你覺得呢? 我們來想一想. 既然平均花 13584 元能蒐集到一套, 那這位消費者花了 38500 是非常不智的. 但是眞的是這樣嗎? 我們只是算出 “平均” 花 13584 元能蒐集到一套, 但是這位消費者 “保證” 可以蒐集到一套. 萬一我們運氣非常背呢?

因此除了算購買次數 X 的 “平均” , 我們還要算標準差. 高等數學上我們通常算標準差的平方, 稱爲**變異數**, 記號是 $Var(X)$. 變異數也是線性的, 所以 $Var(X) = \sum_{i=0}^{n-1} Var(X_i)$. 令 $p_i := \frac{n-i}{n}$ 爲做 X_i 時成功的機率. 利用變異數的定義 $Var(X) := E(X^2) - E(X)^2$ 可得 $Var(X_i) = \frac{1-p_i}{p_i^2}$. 經過一番計算後可得

$$\begin{aligned} Var(X) &= \sum_{i=0}^{n-1} Var(X_i) = \sum_{i=0}^{n-1} \frac{1-p_i}{p_i^2} \\ &= n^2 \sum_{i=1}^{n} \frac{1}{i^2} - nH_n \end{aligned}$$

當 $n = 41$ 時, 用 Maple 算出 $Var(X) = 2548.2$, 因此標準差是 $\sqrt{Var(X)} = 50.48$. 根據 $50-68-95$ 法則, 往上兩個標準差, 右端剩下的只有 $< 2.5\%$ 沒有涵蓋. 意思是: 雖然原來‘平均’買 176 次可以蒐集到一套, 可是如果沒蒐集到, 再多買 101 次, 都還蒐集不到一套的機率 $< 2.5\%$. 而多買的這 101 次要花掉 $77 \times 101 = 7777$ 元, 連同一開始的平均花 13584 元, 總共也才 21361 元. 這已經是很倒楣的情況了. 如果運氣壞到還蒐集不到一套, 再往上買一個標準差, 這時候是 $50-68-95-99.7$, 表示花了 $21361 + 77 \times 50.5 = 25250$ 元, 有 $> 99.8\%$ 的機會可以蒐集到一套.

如果還蒐集不到, 那... 不是運氣太差, 是運氣太好, 應該去買樂透. 所以呢, 那位花了 38500 元的先生/小姐絕對沒有‘仔細精算’, 以數學的觀點來看, 他至少浪費了一萬多元.

4 難以蒐集的磁鐵

如果我是廠商, 我大概會讓某些磁鐵出現的機率少一點. (這就是 “集五個字送轎車” 的原理, 第五個字怎麼都集不到). 現在假設有 n 種不同的磁鐵, 第 i 種出現的機率是 p_i, 當然 $\sum p_i = 1$. 那我們求 $E(X)$, 即平均要買多少次可以蒐集到一套?

我們已經從高中數學跨到高等數學了. 這是在機率論上一個非常有名的問題, 稱爲 *The coupon collecting problem.* 任何一本大學機率論的課本裡都可以找到. 答案是

$$E(X) = \int_0^\infty (1 - \prod_{i=1}^n (1 - e^{-p_i t}))dt.$$

當然, 標準差也是可以算的, 在此節省篇幅從略.

5 最新發展

理論上, 上面已經解決了所有的問題, 但是數學家之所以有趣就是永遠好奇, 永遠可以挖出一些新想法.

多出來的給別人, 這是正常的人性. 所以有數學家問了一個非常有趣的問題: 如果家裡有兩個小孩, 哥哥每次有重複的磁鐵就順手給弟弟 (從頭到尾只有哥哥購買), 請問, 當哥哥已經蒐集到一套時, 弟弟平均還缺多少個不同的磁鐵? 比如 Hello Kitty, 在看下去之前, 你能估計一下當哥哥蒐集到一套的那一刻, 弟弟還缺多少張不同的磁鐵? 十張? 十五張?

答案非常漂亮: 假設有 n 種不同的磁鐵, 則在哥哥蒐集到一套的那一刻, 弟弟還缺

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$$

張. 數學內在的美麗眞是不可思議的. $n = 41$ 時這個結果是 4.3, 就是說, 平均只差不到五張就已經可以集到第二套了! 這實在相當違反直覺.

數學家把問題再一般化, 現在假設有 n 種不同的磁鐵, 第 i 種出現的機率是 p_i, 而且家裡有 k 個小孩, 規則同上: 只有老大一直買磁鐵, 一有重複的就給老二, 而老二一有重複的就給老三, 如此下去. 則在老大蒐集到一套的那一刻, 其他兄弟各缺多少張?

我們已經跨入數學研究的前線, 這個問題稱爲 Collector's brotherhood problem, 一直到 2001 年才由 數學家 Foata, Han, Lass, Zeilberger 等人解決! 有興趣的老師可以參考他們的論文 [1, 2].

再接下去呢? 我想我會考慮一下如何讓某些磁鐵出現的機率小一點又不會小到讓消費者絕望, 這應該牽涉到一點賽局理論了.

6 後記: 數學科普

數學科普是非常不容易的, 因爲如果 (作者/讀者) 不花時間去 (解釋/瞭解) 數學的符號和論證, 就只剩下"講故事". 而講故事可能就喪失了 數學之所以是數學, 以及數學之所以迷人的最大原因: 理解並欣賞藉由數學獨特的語言 (符號) 來探索論證的過程, 以及發展理論的美麗. 如果大家能從這篇短文中體會到 "大學數學" 距離日常的教學並不會那麼遙遠, 而且體會到數學之美, 那 Hello Kitty 就更可愛了.

References

[1] D. Foata, G-N. Han and B. Lass, Les nombres hyperharmoniques et la fratrie du collectionneur de vignettes, *Séminaire Lotharingien de Combinatoire* **B47a** (2001), 20 pp.

[2] D. Foata, D. Zeilberger, The Collector's Brotherhood Problem Using the Newman-Shepp Symbolic Method , *Algebra Universalis* **49** (2003), 387-395.