對機率要有信心

黃文璋
國立高雄大學應用數學系

1 前言

在高中數學學科中心網站的討論區看到一篇“機率或信心Q&A”的文章(以下稱為Q&A), 由題目即知該文試圖解釋機率與信心二名詞。

處在一個生活中已缺不了統計的時代, 具備足夠的統計知識, 似乎是必須的。這可能是近年來, 高中數學增加不少統計題材的主因。但我一直不贊成高中數學裡教太多統計。因要在一向講究精準及不變性的高中數學裡, 讓學生理解涉及隨機概念的題材, 先天便有其困難。尤其信賴區間, 在一般大學統計教科書中, 是放在較後面的章節。學生接觸信賴區間時, 已具備夠多的機率統計基礎。如今卻跳過一切預備知識, 堂而皇之地進入高中數學裡, 引發的困惑, 多過吸收到的統計知識, 也就不足為奇。有些認眞的高中教師, 想釐清一些概念, 但由於高中數學裡, 能用的工具極有限, 因此很多釐清的文章, 往往淪為郢書燕說。既未澄清任何概念, 反讓原本清晰的部分, 也模糊起來。本文便擬對Q&A, 略做修正及補充。

2 機率的意義

在Q&A中, 以下述問題開場:

> 在某一次教師研習的綜合座談中, 有老師提到以下的問題:
>
> > 投擲1只骰子(傳統6面骰字, 點數1,2,3,4,5,6,點數1,4為紅色, 其他點數為黑色), 擲出1 點的機率為何? 若已知擲出的顏色為紅色, 請問擲出1點的機率為何?
> >
> > 這個問題在學生沒有學過“信賴區間與信心水準”之前是沒有疑義的, 還沒有擲出骰子前, 擲出1點的機率是1/6; 在擲出的顏色為紅色的條件下, 擲出1點的機率為1/2 。但學生學過信賴區間之後會說, 不, 老師, 既然骰子已擲出, 則骰子是1點的機率不是1就是0, 不會是1/2或其他任何數字 。
> >
> > 結果, 學得越好的學生心中的疑惑卻越深 。我們該如何回應這個學生? 在條件機率的情境下, 應該使用「 機率」或「信心」?

Q&A作者如下回答:

> 擲一個公正骰子一次, 出現一點的機率= 1/6 。這是由大數法則得到的, 故以機率稱之 。擲一個公正骰子一次, 在已知出現紅色點數的條件下, 出現一點的機率= 1/2 。這是由大數法則得到的, 故仍以機率稱之 。

其中我們除了加上幾個標點外, 其餘一字未改 。

我們說投擲一個公正骰子, 此處“公正”一詞乃口語, 其含意為骰子各面出現的機率相同, 即皆為1/6 。這是骰子公正下的結果, 與“大數法則”並無關係 。就如若假設生男 、生女的機會相同, 那便表生男 、生女的機率各為1/2 。僅是一件事不同的講法而已 。說是由大數法則得到, 不但不是事實, 而且易讓人對機率生畏, 後果是令師生連對中文字的運用, 都將失去

信心。Q&A作者說“學得越好的學生心中的疑惑卻越深”, 豈不就是因常看到這些夾纏不清的講法所造成?

前述Q&A所提出的問題, 最後有句“在條件機率的情境下, 應該使用機率或信心?” 依題意, 在擲出骰子的顏色為紅色之條件下, 擲出1點的機率就是1/2, 毫無疑義。提到什麼在條件機率的“情境”下, 提到什麼“信心”, 只有令人更為困惑而已。本來無一物, 就不要惹塵埃。

現在來看“骰子已擲出, 則骰子是1點的機率不是1就是0, 不會是1/2或其他任何數字”, 此Q&A中所提學生的困擾。欲解此疑惑, 歸根究柢, 仍是要回到機率的意義。

事實上, 只要是隨機現象, 便能談機率。有人敲門, 是男是女? 你可說男女機率各1/2。門外那人的性別當然是確定的, 並不會忽男忽女。但由於你不知, 對你而言, 可能男可能女, 這樣的“未知”, 造成“你”可視門外者之性別為一隨機現象。而基於社會上男女約各半的認知, 在無其他資訊下, 1/2的機率便是這樣產生。此與Q&A中, 已知擲出的顏色為紅色(點數有1,4二可能), 則擲出1點的機率為1/2, 乃一樣的原理。但若你聽到門外傳來高跟鞋的走路聲, 說不定會說8成是女的。這裡便用到條件機率, 給定的條件是“敲門者穿高跟鞋”。當然, 若你以往的經驗, 是男生也有不少穿高跟鞋者, 因此認為敲門者乃女生的機率為0.6, 亦無不可。對同一隨機現象, 每個人認定的機率, 可以很主觀的。但要注意一點, 一旦門打開, 則門外那人, 不是男就是女。即這時敲門者為女生的機率, 將不是1就是0, 不能再說0.8或0.6了。機率裡自有一套邏輯, 而既然是邏輯, 就要合理。若明明已看到是女生(或男生), 還說女生的機率為0.8; 或若已假設骰子為公正, 卻認為投擲一次後, 會得到偶數的機率為0.7, 那就絕非在談機率了。

上述那類例子處處可見。投擲一銅板, 待落地後蓋起來, 要人猜正面或反面? 拿個蘋果放背後, 要人猜蘋果或橘子? 這些都是對有些人並非隨機現象, 但對某些人卻為隨機現象之例。又如考試前一天仍可猜題, 題目不是早已出好了嗎? 怎還有人在說“這題很可能會考”? 也是類似的道理。再給一例。大家聽過“汽車與山羊”的問題嗎? 此問題亦曾出現在電影“決勝21點” 中。有三扇門, 其中一扇門後有汽車, 另兩扇門後為

山羊。你選定一扇門後, 主持人打開另兩扇門之一, 問你是否要更改選擇。選擇要不要更改, 當然是依更改後得到汽車的機率是否會變大? 你看, 汽車在那一扇門之後, 也是早已確定, 但仍可談機率。汽車與山羊問題之詳細討論, 可參考黃文璋(2010)一文。

至於為什麼可以說骰子為公正, 即何以知道骰子各點數出現的機率皆為1/6? 這有幾種可能的想法, 牽涉到"機率"一詞幾種不同的意義, 可參考黃文璋(2011)一文。我們列出幾種常見的想法如下。第一種是因骰子有6個面, 基於相同的可能性(古典機率), 導致每個面出現的機率皆為1/6。第二種想法可稱為主觀的, 說不定就是覺得沒有道理那一面較易出現。第三種可能, 乃由過去多次投擲的經驗, 觀察到骰子各面出現的"相對頻率"(即某面出現的次數除以總投擲數)難分軒輊。還有一種可能(這是不少數學家採用的), 就是此為一假設。若骰子不是虛擬, 而是真有一骰子, 則不論基於上述那一種想法, 究竟是否可採信骰子各面出現的機率皆為1/6, 可以做一統計檢定。因此, 在上述第三種對機率的想法裡, 骰子各面出現的相對頻率, 是否夠接近到足以視為相等, 便可藉助統計檢定來判定。而也是在第三種想法裡, 某面出現的機率為1/6, 導致多次投擲後, 該面出現的相對頻率將接近1/6, 才用到大數法則。但要有隨機的概念, 即使是公正的骰子, 不論投擲再多次, 各面出現的相對頻率, 都很難相等。有些人以為投擲數夠多後, 會使各面出現的相對頻率"相等", 這完全是錯的。

另外, 必須一提的是, 在隨機世界中, 一切都是假設, 只看你接受那一個。骰子是否公正, 不論投擲再多次, 仍是天曉得。統計檢定裡, 依無罪推定的原則, 及在給定所能容忍之犯錯機率下, 有一套程序, 以判定該接受, 或拒絕(更保守的講法是"不能接受")那一假設。除非不是隨機現象(如銅板兩面皆為正), 否則不論證據再顯著(如投擲銅板100次都得正面), 其推論均可能犯錯。我們僅能以較好的統計方法, 減小犯錯機率。

3 信心水準

Q&A的第二部分如下:

中央極限定理

- 從平均數為μ, 標準差為σ的母群體中, 以簡單隨機抽樣法抽取n個樣本, 以$\overline{X}$表示樣本平均值, 當樣本數n足夠大時, 則$\overline{X}$的期望值為μ, 變異數為σ^2/n 。
- CLT: $\overline{X}$分配近似常態分配$\mathcal{N}(\mu, \sigma^2/n)$ 。
- 平均數標準化後, 分佈近似$\mathcal{N}(0,1)$的分佈, 即

$$\frac{\overline{X}-E(\overline{X})}{\sqrt{Var(\overline{X})}} = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1),$$

$$P\left(-1.96 < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < 1.96\right) \approx 0.95 \text{ 。} \qquad (*)$$

由$(*)$式子得到

$$\begin{aligned} & P\left(\mu - 1.96\sigma/\sqrt{n} < \overline{X} < \mu + 1.96\sigma/\sqrt{n}\right) \qquad (**) \\ & = P\left(\overline{X} - 1.96\sigma/\sqrt{n} < \mu < \overline{X} + 1.96\sigma/\sqrt{n}\right) \approx 0.95 \text{ 。} \end{aligned}$$

但是, 我們僅作一次抽樣, 得到樣本平均數$\overline{x}$, 標準差S, 95%信賴區間$[\overline{x}-1.96S, \overline{x}+1.96S]$, 此區間有包含到眞正的$\mu$值, 要不就是沒有包含到眞正的$\mu$值 。$(**)$式中, $\overline{X}$為隨機變數, 所以是重複做相同的抽樣很多次以後所得到機率, 故$(**)$的結果, 仍以機率表示 。

但是受限於人力、物力、財力, 我們僅作一次抽樣, 故稱$[\overline{x}-1.96S, \overline{x}+1.96S]$為$\mu$的95%信賴區間, 我們有95%信心說$\mu$會落在此區間內 。

此處, 我特別以μ、$\overline{X}$表示, 希望不要被p、$\widehat{p}$搞混, 但其意思是相同的 。

Q&A中缺乏足夠文字的鋪陳, 文字也屢有跳躍或不明確處。例如, 根本沒有p、$\widehat{p}$, 何以說不要搞混? 因此有時不太知道作者的意思。但再度,第一個黑點後那段敘述, 與中央極限定理沒有關係。此句並不嚴謹, 本擬設法修正此句, 但如前所述, 因Q&A的作者並未加以說明, 無法了解此句與全文之關聯。為避免誤導, 此句宜刪除, 所以就不費神去修改。

簡單隨機抽樣, 分取出後放回及取出後不放回兩種。若是前者, 則對所有的樣本數n, $\overline{X}$的期望值必為μ, 變異數必為σ^2/n, 不需n很大。若是後者, 如做民調時, 從母體中(設有N人)隨機抽取n個樣本(名字, 或電話)後不放回, 欲估計母體中對某一議題支持之比率, 則知道N, n及μ, 支持的人數X有超幾何分佈(hypergeometric distribution), 期望值為$n\mu$。因此支持的比率$\overline{X}(=X/n)$之期望值, 對每一樣本數n皆為μ。至於$\overline{X}$的變異數給在下式:

$$Var(\overline{X}) = \frac{N-n}{N-1}\frac{\sigma^2}{n},$$

稍複雜些。可看出只有當n與N相比很小時, $(N-n)/(N-1)$才會接近1。所以無論那一種簡單隨機抽樣, $\overline{X}$的期望值皆為μ。至於$\overline{X}$的變異數, 對取出後放回, 等於σ^2/n; 對取出後不放回, 只要n/N很小(而非n足夠大), 便接近σ^2/n。也就是都不需要用到中央極限定理。

必須一提的是, 中央極限定理通常是針對取出後放回的簡單隨機抽樣, 或更一般地, 針對獨立且有共同分佈之樣本。另外, 對於取出後不放回的簡單隨機抽樣, 在n與N相比很小下, 說n足夠大時(當然仍要比N小), $\overline{X}$的變異數"為"σ^2/n固然不對, 但"接近"或"近似"也都不是貼切的講法。因n夠大時, $\overline{X}$之變異數及σ^2/n, 二者皆近似於0。而二近似於0的數, 自然彼此近似。

第二個黑點後那句也宜刪除。中央極限定理, 並非在講$\overline{X}$的分佈, 乃是針對標準化後$\overline{X}$之分佈。對獨立且有共同分配之樣本, 當n很大時, $\overline{X}$之變異數σ^2/n近似於0。而一隨機變數之變異數很小, 表示它有很大的機率近似於常數。即此時$\overline{X}$近似於μ之機率很大。也就是對$\overline{X}$, 與它較相關的是大數法則, 並非中央極限定理。因此這句就不理它了, 直接刪掉。

在第三個黑點裡, 幾個式子大抵可接受。文字敘述中以樣本標準差S取代(**)中的σ也無妨, 仍為一近似的式子。只是第一個式子中有趨近的符號"$\rightarrow$", 因此宜在"即"之後, 加上"n趨近∞時", 或"$n \rightarrow \infty$時"。不過"所以是重複做相同的抽樣很多次以後所得到機率"這句一定得刪除, 完全畫蛇添足, 徒增困擾。$\overline{X}$為隨機變數, 是使μ落在區間有"機率"可言之原因, 與做多少次抽樣無關。再來"受限於人力、物力、財力, 我們僅作一次抽樣"也宜刪除。每次依同樣步驟抽樣, 不論得到再多的區間, 每一個皆是μ的95%信賴區間, 且95%永遠稱為"信心水準", 不會變成"機率"0.95。在估計時, 即使"人力、物力, 及財力" 皆不設限, 每次也只會(或說只想)得到一個信賴區間, 否則究竟要"信賴"那一個? 但在不受限制下, 取樣愈多(即n愈大), 所得的信賴區間長度將愈短。亦即對μ的估計將更準確, 這是花更多精力後的代價。

從估計的觀點, 區間較一個點更可"信賴"。想想醫生宣佈病人可再活1年, 跟說大約可再活6個月至3年間, 何者讓人覺得此醫生較科學, 更值得信賴? 對於估計, 譬如估計一銅板出現正面的機率μ, 在實際取樣前, μ有一機率值會落在信賴區間。如前述(**)式, 其中有一"機率"0.95。但我們一開始就不稱0.95"機率區間"。因就是為了估計才得到該區間, 接續便要取樣。而一旦得到一固定的信賴區間, 說此一特定的區間, 有0.95的機率包含μ, 難免讓人覺得不太妥當, 而又不易解釋清楚, 所以當初引進信賴區間概念之統計學家, 才自始便不稱0.95為機率。但μ的眞實值畢竟乃屬未知, 而數據會說話, 你還是覺得你的投擲結果, 總有值得"信賴"處。即除非已經知道μ值了, 否則有一定的"信心", μ會落在此區間。95%, 便用來描述該信心之大小, 稱為信心水準。至於Q&A的第一部分, 即原先大家所熟知的機率及條件機率的概念, 與信賴區間根本無關, 不會因引進信賴區間的概念後而改變。

只要是為隨機現象(即使對有些人並非隨機), 就能談機率。引進信賴區間及附隨的信心水準後, 原先的機率不會變成"信心"。對機率要有信心。

參考文獻

1. 黃文璋(2010). 機率應用不易 。數學傳播季刊, 34(1): 14-28 。

2. 黃文璋(2011). 認識機率 。數學傳播季刊, 35(2): 32-44 。