

第一章

基本概念

1.3 隨機變數

樣本空間中的元素，不一定是數字，可以是紅球、白球、正面、反面等。做民意調查，想了解民眾對某議題的看法，選項可能有極力贊成、贊成、不贊成、極不贊成，及沒意見等。我們一向比較喜歡處理數字的問題。對上述情況，可將紅球視為1，白球視為0；正面視為1，反面視為0；極力贊成視為5，贊成視為4，不贊成視為3，極不贊成視為2，沒意見視為1。也就是將實驗的結果數量化。

調查選民對某候選人支持與否，以1表支持，0表不支持。雖是數量化了，但若調查一千位，則樣本空間共有 $2^{1,000}$ 個元素，每一元素為一長串有1,000個的0或1的數字，其中第 i 個數字為1，表第 i 個人為支持，0表不支持。這麼大的樣本空間不但難以掌握，也顯得累贅。因我們可能只對其中共有多少人支持感到興趣。因此可令

$$X = \text{數字中1的個數}.$$

則對 X ，樣本空間成為 $\{0, 1, 2, \dots, 1,000\}$ ，減小許多。

2 第一章 基本概念

即使原先樣本空間中的元素就是數字, 我們可能會考慮它的某一函數. 例如, 若隨機地取一蘋果, 量測其重量, 則可取樣本空間 $\Omega = (0, \infty)$. 若蘋果每單位重量的價格為 a 元, 則將蘋果的重量 ω 對應至 $X = a\omega$ 而得其價格. 對於 X , 樣本空間仍可取為 $(0, \infty)$.

由於以上這些原因, **隨機變數**(random variable)的概念便自然地產生了. 所謂隨機變數, 就是一個定義在樣本空間 Ω 的實函數. 若以 X 表一隨機變數, 則此函數的定義域為 Ω , 對應域是實數 R , 值域則為 R 的一個子集合. 換句話說對 $\forall \omega \in \Omega$, $X(\omega) \in R$.

即使不論上述所提的原因, 在數學裡我們本來就常考慮各種函數, 如多項式、指數、對數及三角函數等. 我們喜歡討論函數的變化, 了解函數的性質. 就此觀點, 隨機變數的產生, 就毫不奇怪了. 對同一樣本空間, 往往可定義出無限多個隨機變數. 可能是因實際需要, 也可能只是純粹數學的目的.

其實在很多隨機試驗裡, 我們早就隱含地使用隨機變數了. 例如, 投擲一骰子二次, 令 X 表點數和; 投擲一銅板20次, 令 Y 表所得正面數等.

對於隨機變數 X , 可以想成原先觀測到 ω , 再轉換至 $X(\omega)$, 樣本空間由 Ω 轉為 R (或 R 的一個子集). 但也可想成所觀測到的結果就是 X . 譬如說, 前述民調的例子, 負責抽樣的人有原始資料, 但他整理好後, 交出來的統計資料, 是支持的人數 X . 對大部分的人而言, 觀察值就是 X , 樣本空間為 $\{0, 1, \dots, 1,000\}$. 由此可見隨機變數亦可當做一觀測值, 或說一隨機試驗的結果, 或說一樣本.

隨機變數的概念既然是如此卑之無甚高論, 它不過就是一個函數 (或僅是不同階段下的樣本), 但我們也經常聽到有人說他是在進了(統計)博士班之後, 才終於了解隨機變數的意義. 這又是為什麼呢?

首先以往我們對函數, 常以 f, g, h 或大寫的 F, G, H 等命名, 參數 (或稱變數), 則以 x, y, z , 等表示. 有時我們說有一函數 $f(x)$, 即使因函數較多, 取名為 $a(x), b(x)$ 等, 或藉助希臘字母 $\xi(x), \eta(x)$ 等, 較少見到以大寫的 X, Y, Z 當做函數的. 但對於隨機變數, 我們往往以大寫的 X, Y, Z 等命名. 且不像以往的函數, 變數常是附在一起的, 如 $f(x), g(y)$ 等. 隨機變數這

種函數，其變數並不常與函數(如 X)一起出現。例如，我們會說 $f(x) = 3$ ，但對隨機變數，往往說 $X = 3$ 。尤有進者，除了在一開始定義時，說隨機變數為一實函數，之後我們很少稱它為函數，連名稱都含有“變數”，因此不易使人與以往所學的函數連結在一起。

再看一個原因。隨機變數定義於樣本空間，樣本空間中的元素之出現情況是隨機的，以一機率函數來描述。例如，在數學裡解方程式是很平常的。我們說解 $f(x) = 3$ ，就是找出滿足 $f(x) = 3$ 的實數 x 。但對隨機變數 X ，我們較少說解 $X = 3$ 。 X 如果是由某一 Ω 映至實數 R ，是可以有一集合 $\{\omega | X(\omega) = 3\}$ （雖然如同方程式可能不易解，因此有時無法明確地表示出此集合），但我們通常不會就此打住，而會去求此集合之機率。甚至後來經過一番改造(下一節會說明)，我們常略去解 $X = 3$ 的過程，而直接給出 $X = 3$ 的機率(見底下例3.2)。換句話說，雖然本身是一函數，但一隨機變數的旁邊，另有跟機率有關的函數，指揮該隨機變數的行為。那些函數的味道，就比較接近我們以往所遇到的函數，而也採用一般熟悉的函數 f, g, h 及 F, G, H 等。這也可以解釋為什麼在隨機變數這一階段，我們不採慣用的函數符號。

第三個原因是，由於在樣本空間 Ω 上， σ -體 $\mathcal{F}$ 選取的不同，使得並非每一由 Ω 映至 R 的函數，皆可當做隨機變數。一般要求(尤其在較高等的機率論中)還要滿足

$$(3.1) \quad \text{對 } \forall x \in R, \{\omega | X(\omega) \leq x\} \in \mathcal{F}.$$

為什麼要滿足此條件，我們下一節再說明。如果 $\mathcal{F}$ 取得小些，則有些 Ω 上的實函數，就不一定是隨機變數了。問題因此變化多端，雖然增加挑戰性，對初學者而言，迷惘也隨之而生。幸好在統計學裡， $\mathcal{F}$ 常取得夠大，例如，對可數的 Ω ，我們通常取 $\mathcal{F}$ 為包含 Ω 之所有子集合所形成的集合，因此條件(3.1)自然成立。如此一來所有由 Ω 映至 R 的函數，便皆為隨機變數了。

其他還有使初學者不易了解隨機變數的概念之原因，我們陸續會解釋。

隨機變數常以大寫的英文字母表示，而它的觀測值則以對應的小寫字母表示。有時會說“隨機變數 X 取值 x ”，或說“求 $X = x$ 之機率”。

4 第一章 基本概念

定義一隨機變數後，其值域就是新的樣本空間，原來的機率函數(定義在原來樣本空間所產生之 σ -體)，是否仍適用此隨機變數？

假設樣本空間 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ ，且如一般地，取 $\mathcal{F}$ 為包含 Ω 之所有子集合所形成的 σ -體。定義一隨機變數 X ，其值域以 $\Omega_1 = \{x_1, x_2, \dots, x_n\}$ 表之。則會觀測到 $X = x_i$ ，若且唯若原試驗的結果為 $\omega \in \Omega$ ，使得 $X(\omega) = x_i$ 。即

$$(3.2) \quad P_X(X = x_i) = P(\{\omega | \omega \in \Omega \text{ 且 } X(\omega) = x_i\}).$$

要注意的是，比較好的寫法是 $P_X(\{X = x_i\})$ ，因 $\{X = x_i\}$ 才是一集合，不過為了簡便，我們常簡單地寫成 $P_X(X = x_i)$ 。(3.2)式之左側函數 P_X ，就是由原機率函數引發出來的定義在 Ω_1 上的機率函數， Ω_1 為新的機率空間，當然我們必須要驗證 P_X 的確滿足柯莫果洛夫公理，此部分留在習題中。雖然 P_X 與 P 為不同的函數，但因有(3.1)式，今後如果不會混淆，我們通常只寫 $P(X = x_i)$ ，且將 P 解釋為“機率”的意思。

例3.1 投擲一銅板3次，樣本空間 Ω 如例2.5所給，且設每一樣本之機率為 $1/8$ (因此這是一公正的銅板)。令 X 表共得之正面數。顯然 X 之值域為 $\{0, 1, 2, 3\}$ 。則

$$P_X(X = 1) = P(\{\text{正反反, 反正反, 反反正}\}) = \frac{3}{8}.$$

同理可得 $P_X(X = 0) = 1/8$, $P_X(X = 2) = 3/8$, $P_X(X = 3) = 1/8$ 。

次令 $Y = \text{正面數} - \text{反面數}$ 。則 Y 之值域為 $\{-3, -1, 1, 3\}$ 。且

$$P_Y(Y = -3) = P(\{\text{反反反}\}) = \frac{1}{8}.$$

同理可得 $P_Y(Y = -1) = P_Y(Y = 1) = 3/8$, $P_Y(Y = 3) = 1/8$ 。

例3.2 投擲一公正的銅板100次，則樣本空間 Ω 中共有 2^{100} 個元素。由於此銅板為公正，所以每個樣本點之機率均相同，皆為 $1/2^{100}$ 。令 X 表所得之正面數。欲求 $X = 10$ 之機率。由於每一樣本點皆為一串長達一百的“正、反”文

字, 本例不像上例可輕易列舉出對應 $X = 10$ 之原事件, 但利用排列組合的技巧, 可得

$$\begin{aligned} P_X(X = 10) &= \frac{\Omega \text{ 中有 10 個“正”之元素個數}}{\Omega \text{ 之元素個數}} \\ &= \frac{\binom{100}{10}}{2^{100}}. \end{aligned}$$

對 $i = 0, 1, \dots, 100$,

$$P_X(X = i) = \frac{\binom{100}{i}}{2^{100}}.$$

雖然我們並未展示出事件 $\{X = i\}$ 究竟為何, 但其機率卻可求出.

一般而言, 令 Ω_1 表隨機變數 X 所引發出來的樣本空間, 則對 $\forall A \subset \Omega_1$, 我們定義引發出來的機率函數

$$P_X(X \in A) = P(\{\omega | \omega \in \Omega, \text{ 且 } X(\omega) \in A\}).$$

例3.3 如果 Ω 本來就是實數的一個子集合, 則 $X(\omega) = \omega$ 為一很自然的隨機變數. 例如, 一袋中有 10 張紙牌, 分別寫數字 1 至 10. 隨機地取一張, 觀測所得之點數. 則樣本空間 $\Omega = \{1, 2, \dots, 10\}$. 可令 $X(\omega) = \omega$, 而得一隨機變數. 另一種講法是自袋中隨機地取一張紙牌, 並令 X 表所得之點數. 這種講法就是直接引進隨機變數.

其他如令 Y 表某袋中的紅球數, 令 Z 表某校戴眼鏡的學生人數, 都是直接引進隨機變數之例. 紅球、戴眼鏡本來皆非數字, 但因我們將觀測的結果數量化, 或者說考慮數值上的特性, 便產生了隨機變數.

對同一樣本空間可定義無數個隨機變數. 即使 $\Omega = \{0, 1\}$ 只有兩個元素, 對任二實數 a, b , 令 $X(0) = a, X(1) = b$, 便得一隨機變數.

常數也是一隨機變數. 即對一實數 c , 若 $X(\omega) = c, \forall \omega \in R$, 則 X 為一隨機變數. 有時以 $X \equiv c$ 表之. 這種只取一個值之隨機變數, 由於並不會變, 雖然仍視為隨機變數, 但稱之為**退化的**(degenerate).

例3.4 對例3.1中的 Ω , 依序令 $Z(\text{正正正}) = a_1, Z(\text{正正反}) = a_2, \dots$,

$Z(\text{反反反}) = a_8$, 其中 $a_1, a_2, \dots, a_8$ 皆為實數, 則 Z 為一隨機變數. 可看出若取 $a_1 = 3, a_2 = a_3 = a_4 = 2, a_5 = a_6 = a_7 = 1, a_8 = 0$, 即得例3.1中之 X ; 若取 $a_1 = 3, a_2 = a_3 = a_4 = 1, a_5 = a_6 = a_7 = -1, a_8 = -3$, 即得例3.1中的 Y . 任一定義於 Ω 上的隨機變數, 皆可經由適當地選取 $a_1, \dots, a_8$ 而得.

再給一特殊但常見之隨機變數.

例3.5 設 A 為樣本空間 Ω 之一子集. 令 I_A 表 A 之指示函數(indicator function), 即若 A 發生, 則 $I_A = 1$, 否則 $I_A = 0$. I_A 便是“指示”事件 A 是否發生. 因此

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

則 I_A 為一隨機變數, 它只取1與0兩個值, 且

$$P(I_A = 1) = P(A),$$

$$P(I_A = 0) = 1 - P(A).$$

利用指示函數, 可將每一事件對應一隨機變數.