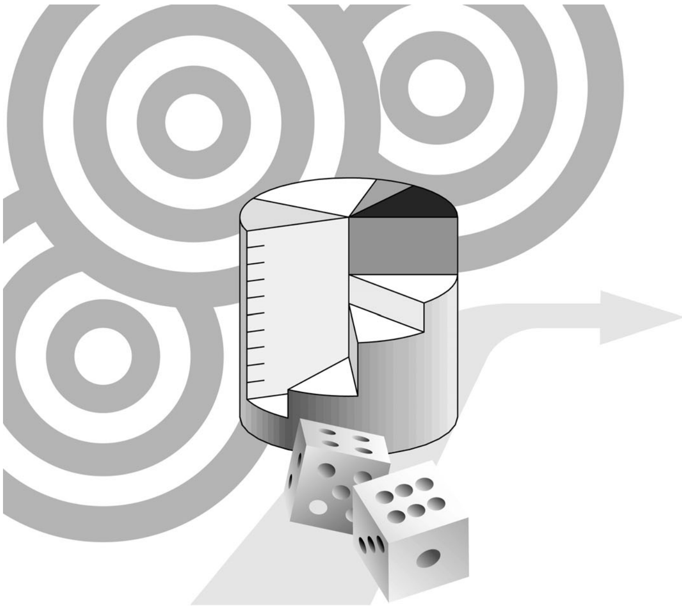

第 3 章

隨機變數與機率分佈

在鑽研隨機現象的特性下，我們已在第 2 章介紹過各種事件的相關陳述了，但這些事件的陳述是架構在較簡單的情形上，由於這些事件的關係可以相當複雜，所以架構一個有規範且一致的處理程序是需要的。其中，為了達成處理上的一致性，步驟之一便是將隨機實驗的所有可能結果都以一個實數表示，藉由這種觀念，每當執行隨機實驗時，直接用一個實數來代表一種結果，會比於直接利用結果本身的物理描述方便許多。舉個例子來說，當某個隨機實驗的結果成功或失敗時，則我們可恣意地指派數字 1 為「成功」之事件，而數字 0 指定給「失敗」之事件，因此其相對的樣本空間 (sample space) 將變成 $\{1,0\}$，取代原本的 { 成功、失敗 }，其中結果為「1」的意思就是實驗結果為「成功」。

經由這種程序，不但讓我們可以用實數所組成的新樣本空間，來取代原隨機實驗的樣本空間，更提供了在機率運算上的算數工具。另外，在科學與工程的領域裡，大部分的問題都在處理量值的測試。因此，對許多感興趣之隨機實驗而言，其相關的樣本空間均由實數所組成，如此一來就達成程序之目的，使隨機實驗一致化。進一步來說，建立在這種一致化觀念上，我們才能夠介紹變數 X ，代表某個隨機實驗，而其值即為隨機實驗的

結果，且以實數來取代之。上述所談，孕育出隨機變數的觀念，而更明確的說明將在往後詳加說明。

3.1 隨機變數

在機率空間下，考慮一個隨機實驗，其輸出為樣本空間 S 內的元素。為了替隨機變數架構出一個模型，假設每一個 s 都可對應一個 $X(s)$ 的實數值，且遵循某種規則，可發現 $X(s)$ 是一個定義在機率空間內的**實數點函數** (real-valued point function) (參考下列的定義 3.1)。

定義 3.1

點函數 $X(s)$ 稱之為**隨機變數** (random variable)，如果在下列條件成立：(a) $X(s)$ 是定義在其機率函數的隨機實驗中樣本空間 S 裡，且為有限實數點函數；(b) 對於所有的實數 x，集合 $\{s: X(s) \le x\}$ 為一個事件。 $X = X(s)$ 的關係代表樣本空間 S 裡的每一個元素 s，映至實數軸 $R^1 = (-\infty, \infty)$ 上的某一個點 X。

為了方便起見，對於在 s 與隨機變數 $X(s)$ 之間的符號關係略而不談。

在定義 3.1 裡的第二個條件又稱為**可量測性條件** (measurability condition)。這項條件確保了每一個 x，對於事件 $X \le x$ 的相關機率是有意義的。更廣泛來說，其確保該事件之任何有限或是可數無窮組合的機率也是有意義的。

為了進一步瞭解隨機變數在隨機現象裡所扮演的角色，再次考慮一個簡單的隨機實驗，其可能結果為成功或失敗，在此，用「1」代表成功的事件，「0」代表失敗的事件，並假設 X 是該實驗的隨機變數，則 X 只有兩種結果：1 和 0。此外，在這個例子裡，以下的陳述是相同的：

- 實驗結果為「成功」；
- 實驗結果是「1」；
- $X=1$。

如果隨機變數 X 被定義在有限或可數無窮個樣本點的樣本空間裡，則稱該隨機變數 X 為**離散隨機變數** (discrete random variable)。在這個例子裡，隨機變數 X 均為離散的數值，並且可以列舉出其所有可能的值。而對於不可數無窮個樣本點的樣本空間而言，其隨機變數 X 便稱為**連續隨機變數** (continuous random variable)，而該隨機變數的數值則分佈於一個或是多個連續區間上，且亦為實數值。之所以會特地區分，是因為這兩種隨機變數在機率值的給定上，有其不同的考量。在科學與工程領域之中，這兩種隨機變數都很重要，在以下的章節裡，我們將會做更深的探討。

往後，所有的隨機變數都將以大寫的英文字母表示之，例如：$X, Y, Z, \cdots$。而隨機變數 X 所對應的值則以小寫英文字母表示，例如：x, y, z 或是 $x_1, x_2, \cdots$。

以下將會探討一連串的隨機變數 $X_j, j=1,2,\cdots,n$，而在這些例子裡，我們都假設它們均定義在相同的機率空間裡。對這些隨機變數 $X_1, X_2, \cdots, X_n$而言，S 裡的每一個元素 s 都會對應到 n 維 **Euclidian 空間** (Euclidian space) R^n 裡的某一個點。注意在做 n 個隨機變數的分析時，可以將其考量成一個**隨機向量** (random vector)，而此隨機向量包含這 n 個隨機變數，且以粗體的大寫英文字母 $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \cdots$ 呈現。注意，在往後的內容，隨機向量的觀念將會大量地被採用。

3.2　機率分佈

機率分佈可呈現隨機變數的特性，也就是說，隨機變數的特性是由其樣本空間內所對應的機率值來呈現。用來表達離散型隨機變數的方法有兩

種，即**機率分佈函數** (probability distribution function, PDF) 與**機率質量函數** (probability mass function, pmf)。這兩種方法的任一種均可完整地描述隨機變數的特質。至於連續型隨機變數，其相關函數有機率分佈函數 (PDF)，以及**機率密度函數** (probability density function, pdf)，其中機率分佈函數的定義與離散型隨機變數的 PDF 是相同的，以上的定義將在下面一一介紹。

3.2.1 機率分佈函數

考慮一隨機實驗，其隨機變數為 X。在給定一個實數 x，觀察事件 $\{s: X(s) \le x\}$ 的機率；或是更簡單地說，想要瞭解 $P(X \le x)$ 為何？這個機率很明顯地與 x 的值有關。考慮下列函數

$$\boxed{F_X(x) = P(X \le x)} \tag{3.1}$$

並定義 (3.1) 式為隨機變數 X 的機率分佈函數，或簡稱為 X 的**分佈函數** (distribution function)。在 (3.1) 式裡，下標符號 X 代表其隨機變數。如果不會有所混淆的話，有時我們會省略這個下標符號。再強調一次，$F_X(x)$ 可以簡單地表示為 $P(A)$，代表事件 A 發生的機率，其中事件 A 亦即事件 $X \le x$，這些詳細的內容都已詳述在第 2 章。

而 PDF 本身也是一個機率，且為樣本空間 S 裡的一個子集合所對應的機率，其元素為小於或等於 x 的實數。當 x 遞增時，該子集合便囊括更多的實數線，故 PDF 的值也會遞增，直到機率值為 1。因此，隨機變數的 PDF 將隨著 x 的增加而遞增，故又稱為**累加分佈函數** (cumulative distribution function, CDF)。

根據上述的定義與討論，以下將探討 PDF 一些重要的特性：

- 無論是離散或是連續型隨機變數均存在 PDF，且值介於 0 和 1 之間。

- 它是一個非負值、對左邊連續 (continuous-to-the-left)、非遞減實數變數 x 的函數。此外，

$$F_X(-\infty)=0 \quad 且 \quad F_X(+\infty)=1 \tag{3.2}$$

- 如果 a 和 b 為兩個實數，且 $a>b$，則

$$P(a<X\le b)=F_X(b)-F_X(a) \tag{3.3}$$

由 (3.3) 式可直接得到下列式子

$$P(X\le b)=P(X\le a)+P(a\le X\le b)$$

從 (3.3) 式中，可發現隨機變數 X 值介於任意區間的機率，可用兩個 PDF 值的差值來表示。一般來說，任意區間的集合，其對應的的機率均可用 PDF 來表示。

例題 3.1

假設一離散隨機變數 X 的值有 $-1,1,2,3$，且其相對應的機率分別為 $\frac{1}{4},\frac{1}{8},\frac{1}{8},\frac{1}{2}$。則

$$F_X(x)=\begin{cases} 0, & \text{for} \quad x<-1 \\ \frac{1}{4}, & \text{for} \quad -1\le x<1 \\ \frac{3}{8}, & \text{for} \quad 1\le x<2 \\ \frac{1}{2}, & \text{for} \quad 2\le x<3 \\ 1, & \text{for} \quad x\ge 3 \end{cases}$$

如圖 3.1 所示，為離散隨機變數之 PDF，由 0 至 1 以階梯狀遞增。

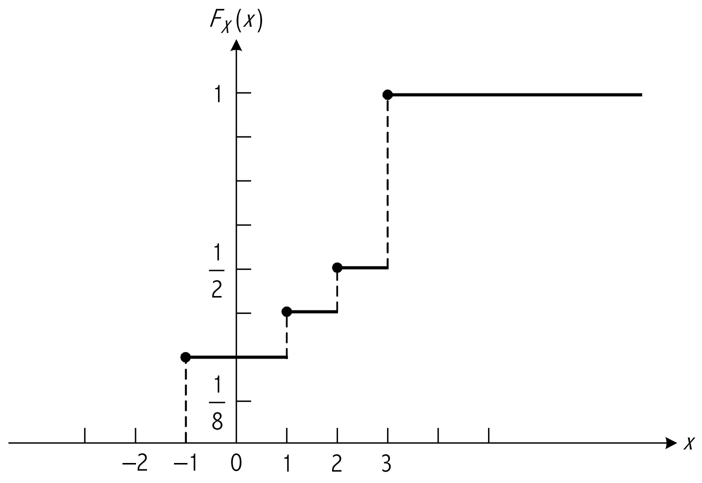

圖 3.1　例題 3.1 裡，離散隨機變數 X 的機率分佈函數 $F_X(x)$。

連續隨機變數則定義在實數線上，其數值為不可數。假設連續隨機變數在某特定值的時候，其機率為零，因此 PDF 不會有離散跳躍的問題。典型的連續隨機變數之 PDF 呈現於圖 3.2。可發現和離散隨機變數不一樣，

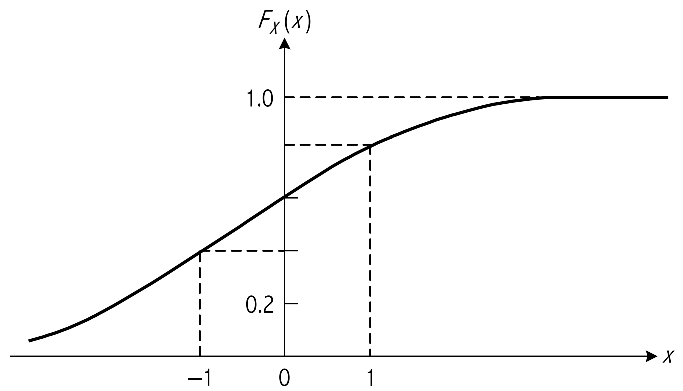

圖 3.2　連續隨機變數 X 的機率分佈函數 $F_X(x)$

連續隨機變數不會有跳躍或是不連續的情形發生。假設隨機變數 X 為某區間內的數值，則此機率亦可藉由 (3.3) 式求得，且只對連續隨機變數的機率才有意義。舉例來說，在圖 3.2 裡

$$P(-1 < X \le 1) = F_X(1) - F_X(-1) = 0.8 - 0.4 = 0.4$$

很明顯地，對於任意值 a ， $P(X = a) = 0$ 。

3.2.2　離散隨機變數之機率質量函數

假設 X 為離散隨機變數，其樣本空間為可數無窮個機率不為零的數值 $x_1, x_2, \cdots$ 所組成。讓 $P(X = x_i) = p(x_i), i = 1, 2, \cdots$ ，明顯地

$$\left.\begin{array}{l} 0 < p(x_i) \le 1, \quad \text{for all } i \\ \sum_i p(x_i) = 1 \end{array}\right\} \tag{3.4}$$

定義 3.2

定義 X 的**機率質量函數** (probability mass function, pmf) 為下列函數

$$\boxed{p_X(x) = P(X = x)} \tag{3.5}$$

其中，下標符號代表的是隨機變數 X 。

對於例題 3.1 中所定義的隨機變數而言，pmf 除了在 $x_i, i = 1, 2, \cdots$ 有值以外，其他的均為零，如圖 3.3 所示。

圖 3.3 為離散隨機變數的 pmf 之典型圖。對於所有連續隨機變數而言，在任意定點 x 處， $P(X = x) = 0$ ，因此連續隨機變數並不存在 pmf。如同機率分佈函數 $F_X(x), p_X(x)$ 亦可完全地描繪出隨機變數 X 的特性。這兩個函數之間的關係式如下：

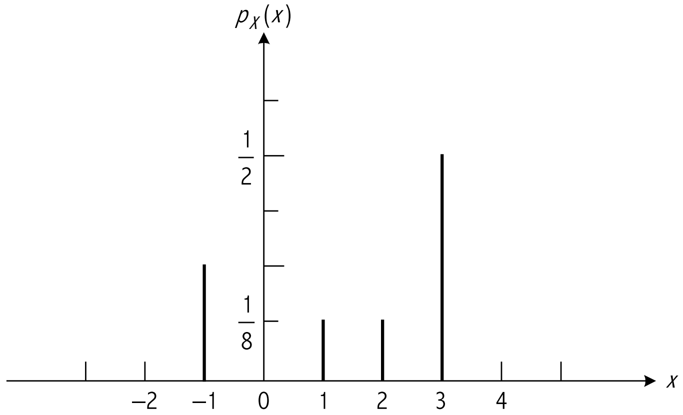

圖 3.3　例題 3.1 裡，離散隨機變數 X 的機率質量函數 $p_X(x)$。

$$\boxed{\begin{aligned} p_X(x_i) &= F_X(x_i) - F_X(x_{i-1}) \qquad (3.6) \\ F_X(x) &= \sum_{i=1}^{i:x_i \le x} p_X(x_i) \qquad (3.7) \end{aligned}}$$

其中假設 $x_1 < x_2 < \cdots$。

在 (3.7) 式裡，總和是針對所有滿足 $x_i \le x$ 的 i 值。因此，我們可以觀察到隨機變數之 PDF 與 pmf 帶有相同的資訊，且彼此互補。

從物理觀念來解析 PDF 及 pmf 的話，考慮單位質量的分佈介於實數線 $-\infty < x < \infty$ 上，當隨機變數的值為 x 時，其 PDF，即 $F_X(x)$，可被解釋為 x 該點以及所有小於 x 的質量總和。至於 pmf，代表的是該實數線上各個單位質量的分佈情形，且各個離散點 x_i 所佔有的質量，亦等同於 $p_X(x_i)$，$i = 1, 2, \cdots$。

例題 3.2

在大多數物理現象裡，常用的離散分佈為**二項式分佈** (binomial distribution)，而有關二項式分佈的探討將在第 6 章加以詳述，目前先利用其性

質，描繪出其 PDF 以及 pmf。

當離散隨機變數 X 為二項式分佈時，其機率質量函數如下所示：

$$p_X(k)=\binom{n}{k}p^k(1-p)^{n-k} \ , \ k=0,1,2,\ldots,n \tag{3.8}$$

其中 n 與 p 為此分佈的兩個參數，n 為一正整數，且 $0<p<1$。而多項式係數

$$\binom{n}{k}$$

被定義為

$$\binom{n}{k}=\frac{n!}{k!(n-k)!} \tag{3.9}$$

圖 3.4 是在 $n=10$, $p=0.2$ 下所描繪的，分別為 X 的 pmf 與 PDF。

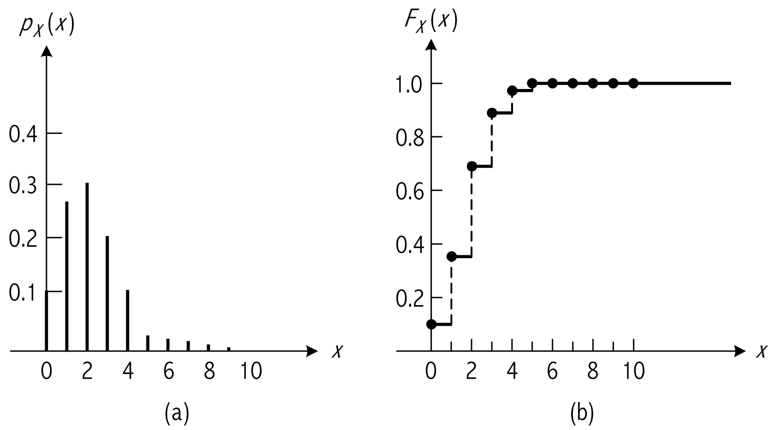

圖 3.4　例題 3.2 裡，離散隨機變數 X 的 (a) 機率質量函數 $p_X(x)$；(b) 機率分佈函數 $F_X(x)$。

3.2.3　連續隨機變數之機率密度函數

對連續隨機變數 X 的機率分佈函數 (PDF) $F_X(x)$ 而言，為 x 的連續函數，且對所有 x 其微分式

$$\boxed{f_X(x)=\frac{\mathrm{d}F_X(x)}{\mathrm{d}x}} \tag{3.10}$$

都存在。則函數 $f_X(x)$ 稱為 X *機率密度函數* (probability density function, pdf)，或簡稱為*密度函數* (density function)[1]。

由於 $F_X(x)$ 是個單調非遞減的函數，故

$$f_X(x)\geq 0 \quad \text{for all } x \tag{3.11}$$

而由 (3.10) 式，可推導出 $f_X(x)$ 的一些特性，這些特性包括

$$\boxed{F_X(x)=\int_{-\infty}^{x} f_X(u)\,\mathrm{d}u} \tag{3.12}$$

以及

$$\left.\begin{aligned}&\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x=1\\ &P(a<X\leq b)=F_X(b)-F_X(a)=\int_a^b f_X(x)\,\mathrm{d}x\end{aligned}\right\} \tag{3.13}$$

圖 3.5 提供一個 pdf 的例子。由 (3.13) 式可知，該曲線下的全部面積為 1，且從 a 至 b 間的斜影面積代表的是機率 $P(a<X\leq b)$。我們發現不論是 pdf 或是 PDF，均可以完整地描繪出連續隨機變數的特性。而由於離散隨機變數的 PDF 有不連續的地方，且在該不連續點亦無法微分，所以對離散型隨機變數而言，其 pdf 並不存在。

以質量分佈的觀念類推，連續隨機變數的 pdf 所扮演之角色，與離散隨

1 注意 PDF 與 pdf 大小寫的差異性，一個代表分佈函數 (distribution function)，而另一個代表密度函數 (density function)。

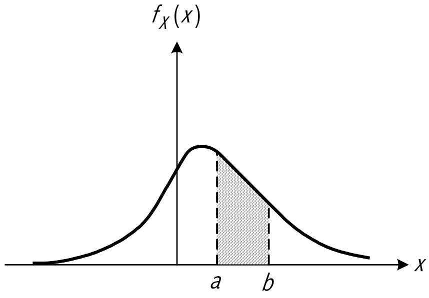

圖 3.5　機率密度函數 $f_X(x)$

機變數的 pmf 相同。因此函數 $f_X(x)$ 可被詮釋為質量密度（每單位長度的質量），但和離散隨機變數不一樣的是無法在每個離散點均配置相對質量。也因此，**密度函數** (density function) 這個名字較適用於 $f_X(x)$。

例題 3.3

考慮一隨機變數 X，其機率密度函數 (pdf) 如下 $(a>0)$：

$$f_X(x)=\begin{cases} a\mathrm{e}^{-ax}, & \text{for } x \ge 0 \\ 0, & \text{其他} \end{cases} \tag{3.14}$$

此隨機變數 X 又稱為**指數分佈** (exponential distribution)。我們可以簡單地驗證由 (3.11) 到 (3.13) 式均成立，而 X 的 pdf 繪於圖 3.6(a)，PDF 則繪於圖 3.6(b)。其中 X 的 PDF 是經由 (3.12) 式所求得，如下所示：

$$F_X(x)=\begin{cases} \int_{-\infty}^{x} f_X(u)\,\mathrm{d}u = 0, & \text{for } x<0 \\ 1-\mathrm{e}^{-ax}, & \text{for } x \ge 0 \end{cases} \tag{3.15}$$

利用 $f_X(x)$ 可計算一些機率值。機率 $P(0<X\le 1)$ 就等同於 $f_X(x)$ 在 $x=0$ 到 $x=1$ 底下的面積，如圖 3.6(a) 所示，可得

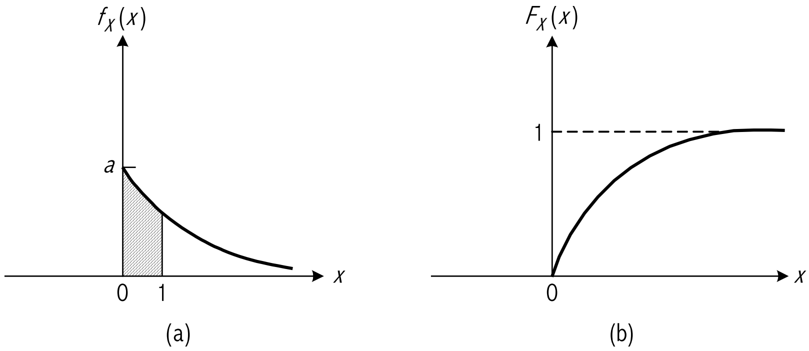

圖 3.6 例題 3.3 裡，隨機變數 X 的 (a) 機率密度函數 (pdf) $f_X(x)$；(b) 機率分佈函數 (PDF) $F_X(x)$。

$$P(0<X\le 1)=\int_0^1 f_X(x)\,\mathrm{d}x=1-\mathrm{e}^{-a}$$

機率 $P(X>3)$ 為 $f_X(x)$ 在 $x=3$ 一直往右延伸，底下的所有面積，即

$$P(X>3)=\int_3^\infty f_X(x)\,\mathrm{d}x=\mathrm{e}^{-3a}$$

而上述的機率值也可以利用 $F_X(x)$ 求得，如下所示：

$$P(0<X\le 1)=F_X(1)-F_X(0)=(1-\mathrm{e}^{-a})-0=1-\mathrm{e}^{-a}$$
$$P(X>3)=F_X(\infty)-F_X(3)=1-(1-\mathrm{e}^{-3a})=\mathrm{e}^{-3a}$$

值得注意的是，對連續型隨機變數而言，$P(0<X\le 1)$ 與 $P(0\le X\le 1)$ 為相同的機率值，這是由於 $P(X=0)=0$ 之緣故。

3.2.4 混合式分佈

在某些情況下，隨機變數會有些部分是離散，而有些部分為連續的情形。圖 3.7 之所呈現的即為一個例子，由圖中可發現隨機變數 X 在實數線

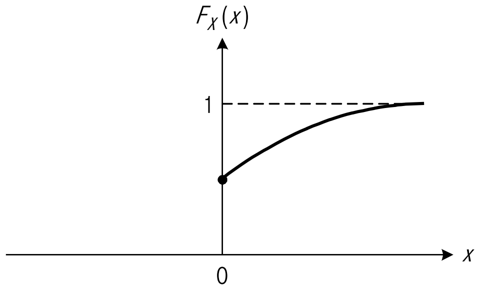

圖 3.7　混合型的機率分佈函數 $F_X(x)$

除了 $X=0$ 外，其餘均為連續分佈，其中 $P(X=0)$ 是個正值。舉例來說，考慮顧客在售票亭需等待的時間，並假設隨機變數 X 代表從開始排隊到買到票所需的時間。因此，X 在 $X\geq 0$ 時才會有值，這一點是很合理的。然而也有無需等待的情形發生，意即在 $X=0$ 時會有一機率值存在，也就是圖 3.7 裡的 $X=0$，為不連續的點。

嚴格來說，對於一個混合式的隨機變數而言，不論是 pmf 或 pdf，我們仍然可以將其分開，並使用各自函數的特性，以達計算上的簡便。假設 $f_X(x)$ 代表該混合分佈中的連續部分，由圖 3.7 可知，$f_X(x)$ 只在 x 為正值的部分才有值，且觀察出在 pdf 曲線下的所有面積不再是 1，而是 $1-P(X=0)$。

例題 3.4

打長途電話時，在一開始的三分鐘內，其費用較為便宜。讓隨機變數 X 代表打長途電話所花費的時間，則其 PDF 如下所示：

$$F_X(x)=\begin{cases}0, & \text{for } x<0\\ 1-e^{-x/3}, & \text{for } 0\leq x<3\\ 1-\frac{e^{-x/3}}{2}, & \text{for } x\geq 3\end{cases}$$

試計算 X 為以下情形時的機率：(a) 多於兩分鐘；(b) 介於二到六分鐘之間。

答：X 的 PDF 如圖 3.8 所示，可發現 X 為混合式的隨機變數。所求的機率可以藉由之前討論的方法，利用 PDF 而得之。因此，對於 (a) 小題

$$P(X>2)=1-P(X\le 2)=1-F_X(2)$$
$$=1-(1-\mathrm{e}^{-2/3})=\mathrm{e}^{-2/3}$$

對於 (b) 小題而言

$$P(2<X\le 6)=F_X(6)-F_X(2)$$
$$=\left(1-\frac{\mathrm{e}^{-2}}{2}\right)-(1-\mathrm{e}^{-2/3})=\mathrm{e}^{-2/3}-\frac{\mathrm{e}^{-2}}{2}$$

而圖 3.9 分別描繪出 X 的離散部分 $p_X(x)$；以及 X 的連續部分 $f_X(x)$。分別如下：

$$p_X(x)=\begin{cases}\dfrac{1}{2\mathrm{e}}, & \text{at } x=3\\ 0, & \text{其他}\end{cases}$$

以及

$$f_X(x)=\frac{\mathrm{d}F_X(x)}{\mathrm{d}x}=\begin{cases}0, & \text{for } x<0\\ \dfrac{1}{3}\mathrm{e}^{-x/3}, & \text{for } 0\le x<3\\ \dfrac{1}{6}\mathrm{e}^{-x/3}, & \text{for } x\ge 3\end{cases}$$

注意，由於是混合型的隨機變數，故 $f_X(x)$ 底下的面積不再是 1，而是

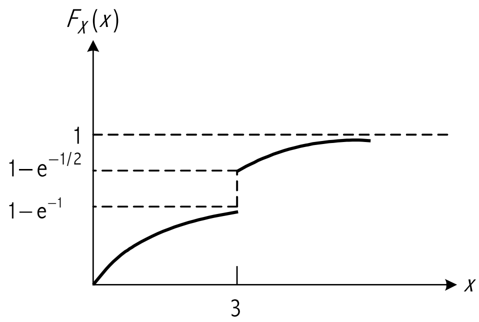

圖 3.8　例題 3.4 裡，X 的機率分佈函數 $F_X(x)$。

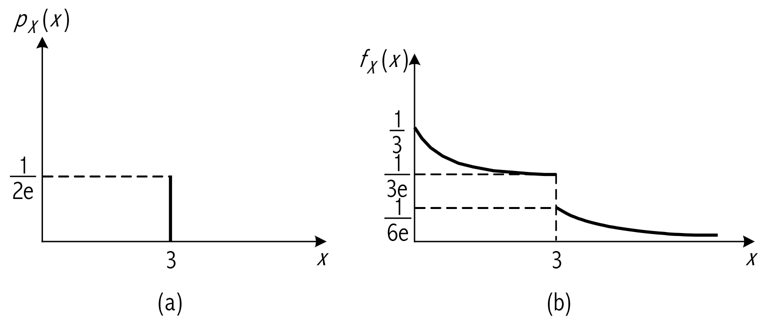

圖 3.9　例題 3.4 裡，隨機變數 X 的 (a) 部分機率質量函數 $p_X(x)$；(b) 部分機率密度函數 $f_X(x)$。

$$1-p_X(3)=1-\frac{1}{2e}$$

為瞭解出 $P(X>2)$ 與 $P(2<X\le 6)$，離散與連續的部分都要考慮，所以就 (a) 小題而言，可求得

$$\begin{aligned}P(X>2)&=\int_2^{\infty} f_X(x)\,dx+p_X(3)\\&=\frac{1}{3}\int_2^3 e^{-x/3}dx+\frac{1}{6}\int_3^{\infty} e^{-x/3}dx+\frac{1}{2e}\\&=e^{-2/3}\end{aligned}$$

而 (b) 小題為

$$\begin{aligned}P(2<X\le 6)&=\int_2^6 f_X(x)\,dx+p_X(3)\\&=\frac{1}{3}\int_2^3 e^{-x/3}dx+\frac{1}{6}\int_3^6 e^{-x/3}dx+\frac{1}{2e}\\&=e^{-2/3}-\frac{e^{-2}}{2}\end{aligned}$$

當然這些結果都等同直接使用 PDF 所求得的答案。

3.3 兩個或兩個以上的隨機變數

在許多的應用裡，通常會同時使用兩個或兩個以上的數值來描述某隨機實驗的結果。舉例來說，在固定的人群裡，探討每個人的體重與身高，或是在物理實驗下，研究溫度與壓力的變化，或是在某一年裡，調查某地區之月均溫的分佈。在這類型的情況裡，需同時探討兩個或兩個以上的隨機變數，而如何描述它們的聯合行為，則是我們所關注的焦點。

首先考慮兩個隨機變數 X 和 Y 的例子，跟之前單一隨機變數的處理程序一樣，來定義它們的**聯合機率分佈** (joint probability distribution)。此外，也可將隨機變數 X 和 Y 視為二維**隨機向量** (random vector) $\boldsymbol{Z}$ 的成員，而兩個隨機變數所對應的聯合機率分佈亦稱為**雙變數分佈** (bivariate distribution)。

因此，若推廣至兩個以上的隨機變數，則所對應之聯合機率分佈就叫做**多變數分佈** (multivariate distribution)。

3.3.1 聯合機率分佈函數

隨機變數 X 和 Y 的**聯合機率分佈函數** (joint probability distribution function, JPDF)，以 $F_{XY}(x,y)$ 表示。定義如下：對所有的 x 和 y

$$\boxed{F_{XY}(x,y) = P(X \le x \cap Y \le y)} \tag{3.16}$$

(3.16) 式為兩事件之交集的機率，亦即，隨機變數 X 和 Y 在二維 Euclidean 平面上，衍生出相對應的機率分佈。

再次利用質量分佈的觀念，讓 (x,y) 平面上分佈著各個點所對應的質量，則在任意區域 R 內的質量會等於 X 和 Y 在 R 內取值的機率，因此，$F_{XY}(x,y)$ 代表所有小於點 (x,y) 所對應的總質量。當 X 和 Y 均為離散型

隨機變數的情況下，(x, y) 平面上所有的**點質量** (point mass) 為有限個或可數無窮多個，而當兩者均為連續隨機變數時，質量將會在 (x, y) 平面上以連續的型式分佈。

很明顯地，聯合機率分佈函數 (JPDF) $F_{XY}(x, y)$ 是非負、非遞減、對**左邊連續** (continuous-to-the-left) 的函數。下列特性是根據定義所直接推論而得

$$\left.\begin{aligned} &F_{XY}(-\infty, -\infty) = F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0 \\ &F_{XY}(+\infty, +\infty) = 1 \\ &F_{XY}(x, +\infty) = F_X(x) \\ &F_{XY}(+\infty, y) = F_Y(y) \end{aligned}\right\} \tag{3.17}$$

舉例來說，觀察 (3.17) 式的第三項關係式。由於 $Y \le +\infty$ 是必定成立的事件，因此聯合事件 $X \le x \cap Y \le +\infty$ 等同於事件 $X \le x$，故

$$F_{XY}(x, +\infty) = P(X \le x \cap Y \le +\infty) = P(X \le x) = F_X(x)$$

同樣地，對於任意的 x_1, x_2, y_1 與 y_2，且 $x_1 < x_2$ 及 $y_1 < y_2$，我們可以證明機率 $P(x_1 < X \le x_2 \cap y_1 < Y \le y_2)$ 可用 $F_{XY}(x, y)$ 表示成

$$\begin{aligned} P(x_1 < X \le x_2 \cap y_1 < Y \le y_2) = &F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) \\ &- F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \end{aligned} \tag{3.18}$$

由此可知所有隨機變數 X 和 Y 相關的機率運算，都可以利用它們的 JPDF 來求得。

最後，注意到 (3.17) 式的最後兩個式子，個別隨機變數的分佈函數可以直接從它們的聯合分佈函數求得。但反過來是不正確的。在多隨機變數的議題裡，這些個別的分佈函數又稱為**邊界分佈函數** (marginal distribution function)。例如，$F_X(x)$ 即為隨機變數 X 的邊界分佈函數。

$F_{XY}(x, y)$ 的圖形大致上可利用 (3.17) 式來推斷。假設 X 和 Y 均是離散的情況下，其輪廓呈現不規則階梯狀，如圖 3.10 所示，由第三象限往第一象限的方向，從 0 遞增到 1。而當 X 和 Y 都是連續時，則 $F_{XY}(x, y)$ 將

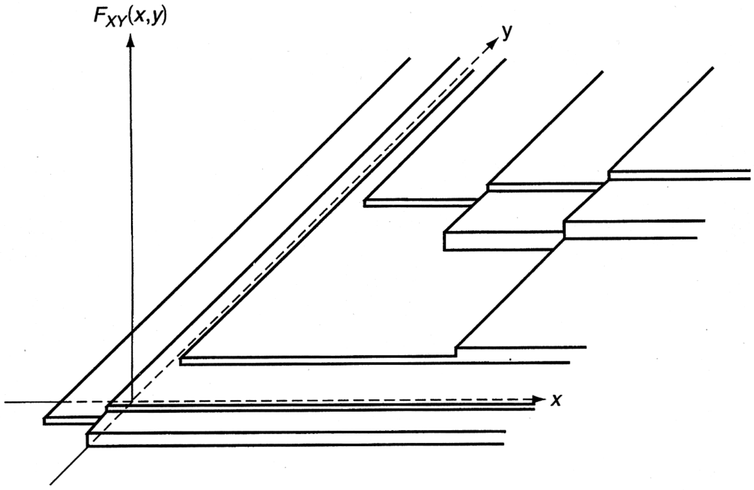

圖 3.10 當隨機變數 X 與 Y 為離散時，所對應的聯合機率分佈函數 $F_{XY}(x,y)$。

變成具有相同特徵的光滑平面。若其中一個隨機變數是離散的，其他均為連續型隨機變數，則造成在某一個方向上呈現階梯狀，而其他方向均為連續曲線的。

若考慮兩個以上的隨機變數，其聯合機率分佈函數也具有相同的性質。假設有 n 個隨機變數 $X_1, X_2, \cdots, X_n$，其 JPDF 定義如下：

$$\boxed{F_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n) = P(X_1 \le x_1 \cap X_2 \le x_2 \cap \cdots \cap X_n \le x_n)} \quad \textbf{(3.19)}$$

這些隨機變數分佈在 n 維 Euclidean 空間裡，同樣地，如同 (3.17) 與 (3.18) 式，n 個隨機變數也可推導出相對的特性。

根據之前提到的隨機向量，這種觀念可套用在有限的隨機變數 X_j, $j = 1, 2, \cdots, n$ 上，歸納為 n 維的隨機向量 $\boldsymbol{X}$，則 $\boldsymbol{X}$ 的 JPDF 與 (3.19) 式相同，但我們可以用更簡潔的形式來表達，即為 $F_{\boldsymbol{X}}(\boldsymbol{x})$，其中 $\boldsymbol{x}$ 也是一個向量，元素為 $x_1, x_2, \cdots, x_n$。

3.3.2　聯合機率質量函數

聯合機率質量函數 (joint probability mass function, jpmf) 能直接描述兩個或多個離散隨機變數之共同特性的函數。讓 X 和 Y 為兩個離散隨機變數，並假設其數值 (x_i, y_j) 為可數無窮多個，即 $i, j = 1, 2, \cdots$，且所對應的機率均不為零。則對所有的 i, j 而言，X 與 Y 的 jpmf 定義為

$$\boxed{p_{XY}(x, y) = P(X = x \cap Y = y)} \tag{3.20}$$

(3.20) 式只有在點 $(x_i, y_j), i, j = 1, 2, \cdots$ 才有值，且等同於聯合機率 $P(X = x_i \cap Y = y_j)$。由 (3.4)、(3.6) 以及 (3.7) 式，可推廣到多隨機變數的情形，如下所示：

$$\left.\begin{aligned} 0 < p_{XY}(x_i, y_j) \le 1 \\ \sum_i \sum_j p_{XY}(x_i, y_j) = 1 \\ \sum_i p_{XY}(x_i, y) = p_Y(y) \\ \sum_j p_{XY}(x, y_j) = p_X(x) \end{aligned}\right\} \tag{3.21}$$

其中 (3.21) 式的最後兩個關係式，即 $p_X(x)$ 和 $p_Y(y)$，又稱為邊界機率質量函數 (marginal probability mass function)。至此，可推得

$$\boxed{F_{XY}(x, y) = \sum_{i=1}^{i: x_i \le x} \sum_{j=1}^{j: y_j \le y} p_{XY}(x_i, y_j)} \tag{3.22}$$

例題 3.5

考慮一個二維隨機漫步 (random walk) 的問題。假設有一顆質點從原點出發，在平面上以單位距離的速度移動。每一步距離是一個單位且是往正的方向，假設往 x 軸方向移動的機率是 p，而往 y 軸的方向移動的機率

是 $q(p+q=1)$，且每一步之間是相互獨立的。試問此質點在移動五步後，其位置的機率分佈為何？

答：由於質點位置可以簡單地用兩個座標來表示，因此我們可以 $p_{XY}(x,y)$ 來表達質點位置的機率分佈，其中隨機變數 X 代表五步之後在 x 軸的位置，Y 則代表 y 軸。很明顯地，$p_{XY}(x,y)$ 只在滿足 $x+y=5$ 的點才有值，且 $x,y\geq 0$。假設每次移動互為獨立，由 3.3 節可知，$p_{XY}(5,0)$ 為質點移動五步之後在 $(5,0)$ 的機率，意即 $p_{XY}(5,0)$ 等於往 x 軸方向連續移動五步之機率的乘積。即

$$p_{XY}(5,0)=p^5$$

若考慮 $p_{XY}(4,1)$，共有五種不同方法可以到達此位置 (往 x 方向 4 步然後 y 方向 1 步；往 x 方向 3 步，y 方向 1 步，再 x 方向 1 步；依此類推)，且每種情形的機率均為 p^4q。因此可得到

$$p_{XY}(4,1)=5p^4q$$

同樣的，$p_{XY}(x,y)$ 其他值亦可求得如下：

$$p_{XY}(x,y)=\begin{cases}10p^3q^2, & \text{for } (x,y)=(3,2)\\ 10p^2q^3, & \text{for } (x,y)=(2,3)\\ 5pq^4, & \text{for } (x,y)=(1,4)\\ q^5, & \text{for } (x,y)=(0,5)\end{cases}$$

圖 3.11 為 $p=0.4, q=0.6$ 時所描繪的 jpmf $p_{XY}(x,y)$。如同 (3.21) 式的第二項，對所有 x 和 y 的而言，$p_{XY}(x,y)$ 的總和必須為 1。

根據 (3.21) 式的最後兩個式子，我們注意到 X 和 Y 的邊界機率質量函數分別為

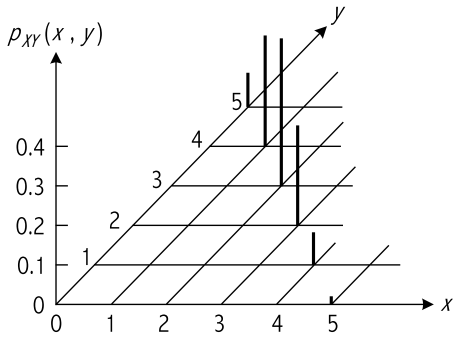

圖 3.11　例題 3.5，在 $p=0.4$ 與 $q=0.6$ 之下的聯合機率質量函數 $p_{XY}(x\,,y)$。

$$p_X(x)=\sum_j p_{XY}(x\,,y_j)=\begin{cases} q^5\,, & \text{for } x=0 \\ 5pq^4\,, & \text{for } x=1 \\ 10p^2q^3\,, & \text{for } x=2 \\ 10p^3q^2\,, & \text{for } x=3 \\ 5p^4q\,, & \text{for } x=4 \\ p^5\,, & \text{for } x=5 \end{cases}$$

以及

$$p_Y(y)=\sum_i p_{XY}(x_i\,,y)=\begin{cases} p^5\,, & \text{for } y=0 \\ 5p^4q\,, & \text{for } y=1 \\ 10p^3q^2\,, & \text{for } y=2 \\ 10p^2q^3\,, & \text{for } y=3 \\ 5pq^4\,, & \text{for } y=4 \\ q^5\,, & \text{for } y=5 \end{cases}$$

利用 (3.22) 式，亦可求得聯合機率分佈函數 $F_{XY}(x\,,y)$。如圖 3.12 所示，比起用三維空間的概念來表示，到不如直接對每個相隔區域內的機率值做加總。要注意的是，在圖 3.12 裡，在 $y=5$ 上所標示的值，即為邊界機率分佈函數 $F_X(x)$ 的值；同樣地，在 $x=5$ 上的值即為 $F_Y(y)$ 的值。

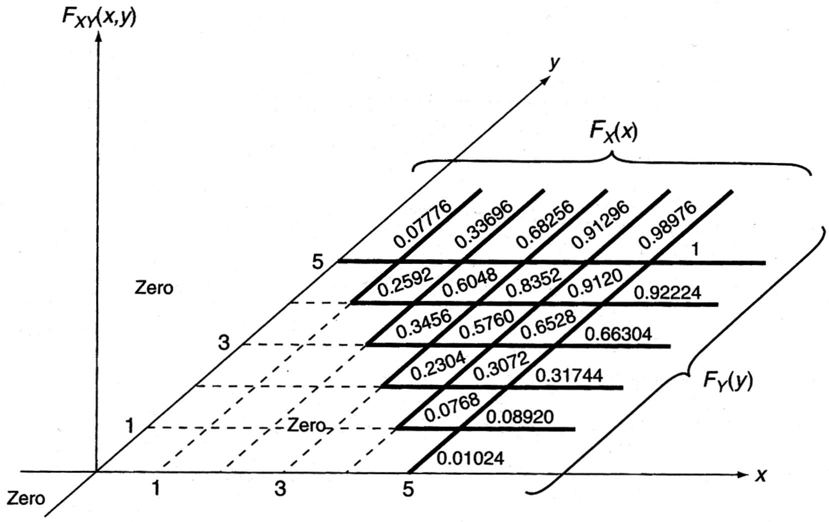

圖 3.12　例題 3.5 裡，在 $p=0.4$ 與 $q=0.6$ 之下的聯合機率分佈函數 $F_{XY}(x,y)$。

利用聯合機率質量函數 (jpmf)，我們可計算任意感興趣的機率值。對任何 X 與 Y 相關的事件，其機率值也可藉由先求得此事件所對應的數值，然後將所有數值的機率 $p_{XY}(x,y)$ 加總起來即可。在例題 3.5 中，假設我們想要知道 $X>Y$ 的機率，就可以由下得到

$$\begin{aligned} P(X>Y) &= P(X=5\cap Y=0)+P(X=4\cap Y=1)+P(X=3\cap Y=2) \\ &= 0.01024+0.0768+0.2304=0.31744 \end{aligned}$$

例題 3.6

回到例題 2.11 裡，讓隨機變數 X 代表暴風雨的強度，其值 1, 2 和 3 分別代表強度的低、中、高，而隨機變數 Y 代表暴風雨所帶來的降雨量，其中值 1 代表到達臨界值，值 2 則否。表 3.1 為 jpmf $p_{XY}(x,y)$ 的相關資訊。

為了求出降雨量到達臨界值的機率，我們將考慮滿足 $y=1$ 的所有可能