Chapter 3

一維隨機變數及其機率分配

3.1 隨機變數

很多試驗中皆屬複雜的，且不易描述其樣本結果，因此在分析時會略感不便，故為了方便作統計分析，可將該試驗中我們關心的現象加以數值化。此種數值化之過程及表現，即為隨機變數（random variable或stochastic variable）的概念。

定義㈠：設 S 為試驗的樣本空間，則稱以樣本空間 S 為定義域，而將其對應到實數的函數為隨機變數，通常以大寫字母 X 、Y 或 Z 表示，即 $X: S \to R$，而隨機變數中每一個變量，皆表示樣本空間中之一種事件。

定義㈡：若隨機變數 X 的值域為可數集合或可數無限個，則稱 X 為間斷型隨機變數（discrete random variable）。若隨機變數 X 的值域為不可數集合，則稱 X 為連續型隨機變數（continuous random variable）。

例題 1

根據投擲一枚六面骰子一次的隨機試驗，下列何者定義為一隨機變數？

(A) X＝1，如出現偶數
　　2，如出現奇數
　　3，如出現1，2，3
　　4，如出現4，5，6

(B) X＝贏，如出現6
　　輸，如出現1，2，3，4，5

(C)$X=1$，如出現1，6
　　0，如出現2，3，4，5

(D)$X=1$，如出現1，2，3
　　2，如出現4
　　3，如出現5。 （93高考）

Sol：(C)；

因(A)、(B)錯；違反隨機變數之定義，在(D)中樣本點6無對應，故錯。

例題 2

投擲一公正硬幣直到正面出現為止。

(1)寫出此試驗的樣本空間。

(2)令隨機變數Y表示試驗中所經歷的次數，則Y的值域為何？

(3)寫出隨機變數Y之機率分配。

Sol：

(1)樣本空間$S=\{H,TH,TTH,TTTH,\ldots\}$

(2)因$Y(H)=1$，$Y(TH)=2$，$Y(TTH)=3$，$\cdots$

故知隨機變數Y的值域為$S_Y=\{1,2,3,4,\ldots\}$

(3)因$P(Y=y)=P(\underbrace{TTT\cdots TT}_{y-1}H)=\left(\frac{1}{2}\right)^{y-1}\cdot\frac{1}{2}$

故隨機變數Y之機率分配為

$$f(y)=P(Y=y)=\left(\frac{1}{2}\right)^{y-1}\cdot\left(\frac{1}{2}\right)\quad,\ y=1,2,3,\ldots$$

3.2 機率分配

統計學之主要目的是在母體中抽取一部分樣本資料，再利用此樣本訊息來推論母體之特性，但由於實際現象之母體分配常爲未知，故在統計學中常以機率分配來代替實際現象之母體分配，亦即爲該母體建立一個數學模式，此即爲機率分配之概念，常見之機率分配類型有：

㈠機率質量函數（probability mass function；p.m.f）。

㈡機率密度函數（probability density function；p.d.f）。

㈢累積分配函數（Cumulation Distribution Function；C.D.F）。

茲分述如下：

一、機率質量函數

定義㈢：設 X 爲一間斷隨機變數，若函數 $f(x)$ 滿足

1. $0 \le f(x) \le 1$ ，$\forall x \in R$
2. $\sum_{x=-\infty}^{\infty} f(x) = 1$

則稱函數 $f(x)$ 爲隨機變數 X 的機率質量函數或稱爲間斷機率分配，且 $f(x)$ 表示 $X = x$ 之機率，即 $f(x) = P(X = x)$ 。

例題3

試求 m 之值使得 $f(x)$ 為機率質量函數：

⑴ $f(x) = m(4x+2)$ ， $x = 1,2,3,4$

⑵ $f(x) = \dfrac{m}{(4x+2)}$ ， $x = 0,1,2$

Sol：

⑴因 $\sum_{x=-\infty}^{\infty} f(x) = 1$ ，故知

$$f(1)+f(2)+f(3)+f(4)=1$$

$$\Rightarrow 6m+10m+14m+18m=1$$

$$\Rightarrow 48m=1 \Rightarrow m=\frac{1}{48}$$

又 $m=\dfrac{1}{48}$ 時，$0 \le f(x) \le 1$ ， $x = 1,2,3,4$

所以 $m=\dfrac{1}{48}$ ，$f(x)$ 可為機率質量函數。

⑵因 $\sum_{x=-\infty}^{\infty} f(x) = 1$ ，故知

$$f(0)+f(1)+f(2)=1$$

$$\Rightarrow \frac{m}{2}+\frac{m}{6}+\frac{m}{10}=1 \Rightarrow m=\frac{30}{23}$$

又 $m=\frac{30}{23}$ 時，滿足 $0 \le f(x) \le 1$　　，$x=0,1,2$

所以 $m=\frac{30}{23}$ ，$f(x)$ 可為機率質量函數

二、機率密度函數

㈠機率密度函數之意義：

定義㈣：設隨機變數 X 爲一連續隨機變數，若函數 $f(x)$ 滿足

1. $f(x) \ge 0 \quad ,\forall x \in R$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

則稱 $f(x)$ 爲機率密度函數。

在定義㈣中的機率密度函數，事件 $a \le X \le b$ 發生機率是指在曲線下之面積（見圖3-1），此面積通常以積分表示，即 $\int_a^b f(x)dx$ ，讀作「曲線下面積 $y=f(x)$ ，當 x 從 a 到 b 整個區間」，此處之 $\int$ 稱爲積分（integral）之符號，a、b爲區間上下限的數字。

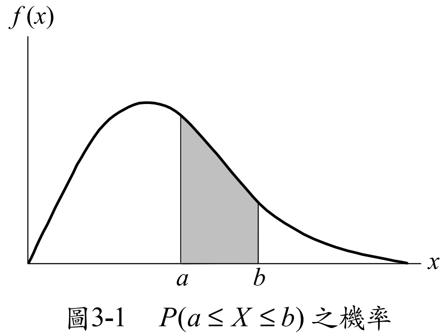

圖3-1　$P(a \le X \le b)$ 之機率