

第一章

基本概念

1.1 機率空間

統計的工作是對**隨機現象**(random phenomenon, 一事先不能預知結果的現象)做分析與推論。由於是處理隨機的模型, 其理論基礎便是**機率論**。而機率論乃是建立在**集合論**(set theory)的基礎。

一**隨機試驗**(random experiment), 其所有可能的**結果**(outcome)之集合, 稱為**樣本空間**(sample space), 通常以希臘字母 Ω 表之。機率空間中的一個**元素**, 便稱為一個**樣本**, 通常以 Ω 的小寫 ω 表之。

例如, 投擲一銅板一次, 觀察所得的結果。這便是一隨機試驗, 可能的結果有正面及反面。因此樣本空間為

$$\Omega = \{\text{正面}, \text{反面}\}。$$

又如觀測某城市一年之車禍次數, 則 Ω 可取成非負整數之集合, 即

$$\Omega = \{0, 1, 2, \dots\}。$$

如果是考慮某種電池的壽命, 則

$$\Omega = [0, \infty)。$$

2 第一章 基本概念

依照樣本空間元素的個數，可將其分為**可數的**(countable)及**不可數的**(uncountable)兩類。所謂可數的，是指有限，或無限但與正整數的集合有1-1(一對一)且映成的對應。樣本空間為一集合，反過來，任一非空的集合，皆可視為一樣本空間。另外，樣本空間中的每一元素稱為一**樣本點**(sample point)。

一旦樣本空間決定了，我們便可以考慮**事件**(event)。所謂事件就是試驗之一些可能的結果之集合。亦即樣本空間 Ω 的任一子集合皆為事件。例如，投擲一骰子一次，觀測所得之點數，則 $\Omega = \{1, 2, 3, 4, 5, 6\}$ ，且共有 $64(= 2^6)$ 個事件。空集合 $\emptyset$ 以及 Ω 本身，皆為事件。

在此對二集合 A, B ， $A \subset B$ 讀做 **A 包含於 B** ，即若 $x \in A$ ，則 $x \in B$ ，其中符號“ $\in$ ”表**屬於**。可以下述符號來說明：

$$“A \subset B” \Leftrightarrow “x \in A \Rightarrow x \in B”。$$

其中“ $\Rightarrow$ ”表**導致**，“ $\Leftrightarrow$ ”表**等價**(equivalent)。又

$$“A = B” \Leftrightarrow “A \subset B \text{ 且 } B \subset A”。$$

我們再定義**聯集**(union)、**交集**(intersection)、**餘集**(complement)及**差集**(difference)如下：

$$A \cup B = \{x | x \in A \text{ 或 } x \in B\},$$

$$A \cap B = \{x | x \in A \text{ 且 } x \in B\},$$

$$A^c = \{x | x \in \Omega \text{ 且 } x \notin A\},$$

$$A \setminus B = \{x | x \in A \text{ 且 } x \notin B\},$$

其中“ $\notin$ ”表**不屬於**。可看出

$$A^c = \Omega \setminus A, \quad A \setminus B = A \cap B^c。$$

例1.1 投擲一骰子一次，則 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。令事件 A 表得到偶數之事件， B 表得到點數不超過3之事件， C 表得到奇數之事件。即

$$A = \{2, 4, 6\}, \quad B = \{1, 2, 3\}, \quad C = \{1, 3, 5\}。$$

則

$$\begin{aligned} A \cup B &= \{1, 2, 3, 4, 6\}, \quad A \cap B = \{2\}, \quad A^c = \{1, 3, 5\} = C, \\ A \cap C &= \emptyset, \quad A \cup C = \Omega. \end{aligned}$$

由於事件就是集合，所以關於事件的很多性質，皆可對應集合中的性質。例如，對二事件 A, B ， $A \cup B$ 表 A, B 至少有一件發生之事件， $A \cap B$ 表 A, B 同時發生之事件， A^c 表 A 未發生之事件， $A \setminus B$ 表 A 發生但 B 未發生之事件。我們列出一些常見的性質，證明則略去。

定理1.1 設 A, B, C 為定義於某樣本空間的三事件。則下述各性質成立：

- (i) 交換律 $A \cup B = B \cup A$,
 $A \cap B = B \cap A$;
- (ii) 結合律 $A \cup (B \cap C) = (A \cup B) \cap C$,
 $A \cap (B \cup C) = (A \cap B) \cup C$;
- (iii) 分配律 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
- (iv) 棣莫根法則 $(A \cup B)^c = A^c \cap B^c$,
 $(A \cap B)^c = A^c \cup B^c$ 。

註1.1 棣莫根(De Morgan, 1806-1871)，為英國數學家。

我們再給可數個事件的聯集及交集之運算。設 $A_1, A_2, \dots$ 為定義於某樣本空間 Ω 的事件。則

$$\begin{aligned} \bigcup_{i=1}^{\infty} A_i &= \{x | x \in \Omega, \text{ 且對某一 } i \geq 1, x \in A_i\}, \\ \bigcap_{i=1}^{\infty} A_i &= \{x | x \in \Omega, \text{ 且對 } \forall i \geq 1, x \in A_i\}. \end{aligned}$$

其中符號“ $\forall$ ”表每一。棣莫根法則亦適用於可數個事件。即對事件 $A_1, A_2, \dots$,

4 第一章 基本概念

$$\begin{aligned} \left(\bigcup_{i=1}^{\infty} A_i\right)^c &= \bigcap_{i=1}^{\infty} A_i^c, \\ \left(\bigcap_{i=1}^{\infty} A_i\right)^c &= \bigcup_{i=1}^{\infty} A_i^c. \end{aligned}$$

例1.2 設 $\Omega = (0, 1]$, $A_i = [1/i, 1]$ 。則

$$\begin{aligned} \bigcup_{i=1}^{\infty} A_i &= \bigcup_{i=1}^{\infty} \left[\frac{1}{i}, 1\right] = \Omega, \\ \bigcap_{i=1}^{\infty} A_i &= \bigcap_{i=1}^{\infty} \left[\frac{1}{i}, 1\right] = \{1\}. \end{aligned}$$

一般設 T 為一**指標集合**(index set), 可數或不可數皆可, 則

$$\begin{aligned} \bigcup_{r \in T} A_r &= \{x | x \in \Omega, \text{ 且對某 } r \in T, x \in A_r\}, \\ \bigcap_{r \in T} A_r &= \{x | x \in \Omega, \text{ 且對 } \forall r \in T, x \in A_r\}. \end{aligned}$$

例1.3 取 T = 所有正有理數之集合, $A_r = (0, r]$ 。則

$$\bigcup_{r \in T} A_r = (0, \infty), \quad \bigcap_{r \in T} A_r = \emptyset.$$

在統計學裡, 我們會遇到的聯集及交集, 多半為可數個集合。我們再給幾個定義。

兩個事件 A, B , 若滿足 $A \cap B = \emptyset$, 便稱為**互斥**(disjoint, 又稱mutually exclusive); 事件 $A_1, A_2, \dots$ (有限個亦可), 若滿足 $A_i \cap A_j = \emptyset, \forall i \neq j$, 便稱為**每對互斥**(pairwise disjoint)。另外, 若 $A_1, A_2, \dots$ (有限個亦可)為每對互斥, 且 $\bigcup_{i=1}^{\infty} A_i = \Omega$, 則 $A_1, A_2, \dots$ 便形成 Ω 之一**分割**(partition)。

$\emptyset$ 與任一事件皆互斥, 而 Ω 除與 $\emptyset$ 互斥外, 與其他事件皆不互斥。

例1.4 取 $\Omega = (0, 1]$, $A_i = (1/(i+1), 1/i]$, $i \geq 1$ 。則可看出 $A_i, i \geq 1$, 為每對互斥, 且 $\bigcup_{i=1}^{\infty} A_i = \Omega$ 。故 $A_i, i \geq 1$, 形成 Ω 之一分割。

重複做一試驗若干次, 有些事件會出現較多次, 有些事件出現的次數則較少。例如, 撲克牌裡的梭哈遊戲, 52張牌, 每人發5張。則長期觀察後, 發現同花比四條出現的**相對頻率**高。這種事件出現的相對頻率, 可以用來解釋機率。機率較大的事件, 出現的相對頻率便較高。不過對機率的解釋, 並不只限於相對頻率的解釋, **主觀的解釋**也是一種。也就是視機率為對一事件之相信程度。

在數學裡給兩個數字0.3及0.5。不用去問0.3或0.5的含義, 而且毫無疑問的 $0.5 > 0.3$ 。但在機率裡, 如果說有兩個銅板 A 與 B , 出現正面的機率各為0.3及0.5, 所代表的意義究竟為何? 各丟一百次, 銅板 A 是否出現30次正面? 顯然不一定。那銅板 B 出現正面的次數是否會較銅板 A 多? 顯然也不一定。由 $0.5 > 0.3$, 並無法轉換成出現正面的次數較多。所以如何解釋機率, 並非一簡單的問題。對於機率, 不能用表面的數學來看它。

底下我們並不擬考慮如何解釋機率, 而要以**公理化的作法**(axiomatic approach), 來定義機率。這是數學的手法, 對初學者而言, 雖較抽象, 但從數學的角度而言, 卻是較簡單的。在數學上, 一個從實際上產生的概念, 一旦發展得夠完全了, 我們便會將其抽象化。例如, 當我們了解實數系統之後, 知道其基本特徵, 於是便列出一些**公理**(axiom), 然後說滿足這些公理的系統, 便是實數。至於這個被稱為實數的系統, 是否是用來計數, 或是是否真有一條數線, 我們就不太在乎了。

假設有一樣本空間 Ω , 對每一事件 A , 我們要給一介於0, 1間的數與其對應, 稱做 A 之**機率**, 並以 $P(A)$ 表之。本來將 P 的定義域取為 Ω 之所有子集合所形成之集合, 值域取為區間 $[0, 1]$, 應是很自然的作法。不過通常並非如此做, 稍後我們再說明, 先給一定義如下。

定義1.1 $\mathcal{F}$ 之一些(至少一個)子集合, 所形成的集合 $\mathcal{F}$, 若滿足下述條件, 便稱為一 **σ -體**(σ -field, 又稱 σ -algebra, 或Borel field):

- (i) 若 $A \in \mathcal{F}$, 則 $A^c \in \mathcal{F}$;
- (ii) 若 $A_i \in \mathcal{F}$, $i = 1, 2, \dots$, 則 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ 。

波瑞爾(Borel, 1871-1956), 為著名的法國數學家, 也是現代機率理論的創始者之一。由定義1.1立即得到 $\emptyset$ 及 Ω 皆屬於 $\mathcal{F}$ 。又若條件(i)及(ii)成立, 由**棣莫根法則**知, 若 $A_i \in \mathcal{F}$, $i = 1, 2, \dots$, 導致 $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ (習題第10題)。另外, 定義1.1亦導致, 若 $A_1, \dots, A_n \in \mathcal{F}$, 則有 $\bigcup_{i=1}^n A_i \in \mathcal{F}$, 此因只要將 $A_{n+1}, A_{n+2}, \dots$ 皆取為 Ω 即可。底下常有類似的情況: 對可數個事件成立的結果, 導致對有限個事件亦成立, 只要適當地選取 $A_{n+1}, A_{n+2}, \dots$ (如在稍後定義1.2, 要將 $A_{n+1}, A_{n+2}, \dots$ 皆取為 $\emptyset$)。σ-體中的每一元素亦皆為集合, 初學者往往會被此集合的集合之概念所迷惑, 其實倒不是那麼複雜。

例1.5 設 $\Omega = \{1, 2, 3\}$, 則可產生那些σ-體?

解. 首先 $\{\emptyset, \Omega\}$, 及

$$\{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$$

是兩個能最先寫出來的, 前者為最小的, 後者為最大的。所謂最小, 是指若有其他σ-體, 一定包含此σ-體; 而所謂最大, 是指若有其他σ-體, 一定包含於此σ-體。對有限集合, 最小的σ-體及最大的σ-體, 元素個數分別為最少及最多。

除了上述兩個之外, 尚有那些σ-體呢? 一σ-體如果包含 $\{1\}$, 則必包含其餘集 $\{2, 3\}$ 。可檢驗

$$\{\emptyset, \Omega, \{1\}, \{2, 3\}\}$$

為一σ-體。同理

$$\{\emptyset, \Omega, \{2\}, \{1, 3\}\},$$

$$\{\emptyset, \Omega, \{3\}, \{1, 2\}\},$$

亦皆為σ-體。如果包含 $\{1\}$ 及 $\{2\}$, 則必包含其聯集 $\{1, 2\}$, 因此餘集 $\{3\}$ 亦在其中。有 $\{1\}$, $\{2\}$ 及 $\{3\}$, 那就導致全部的子集都在其中了, 成為最大的σ-體。可檢驗上述5個σ-體, 就是所能產生之全部的σ-體。

例1.6 設 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。已知一 σ -體中有一元素 $\{1, 2\}$ ，試寫出最小的這種 σ -體。

解.首先說明為什麼是說找“最小”，因這種 σ -體有可能不唯一。例如， Ω 之所有子集合所形成的集合為一 σ -體，且當然包含 $\{1, 2\}$ 。

由於有 $\{1, 2\}$ ，因此餘集 $\{3, 4, 5, 6\}$ 亦須在其中。不難看出集合

$$\{\emptyset, \Omega, \{1, 2\}, \{3, 4, 5, 6\}\}$$

為一 σ -體，且是最小的這種 σ -體。

例1.7 在上例中，試求一包含 $\{1, 2\}$ 及 $\{3\}$ 之最小的 σ -體。

解.由於有 $\{1, 2\}$ 及 $\{3\}$ ，所以也有其聯集 $\{1, 2, 3\}$ ，因此餘集 $\{4, 5, 6\}$ 亦須在其中。再來是 $\{1, 2\}$ 與 $\{4, 5, 6\}$ 的聯集，及 $\{3\}$ 與 $\{4, 5, 6\}$ 的聯集，即 $\{1, 2, 4, 5, 6\}$ 與 $\{3, 4, 5, 6\}$ 皆須在其中。再加上 $\emptyset$ 及 Ω ，可檢驗包含此8個集合之集合，即為所求之 σ -體。

比較例1.6及1.7，可看出雖只是增加一個集合，最小的 σ -體中的元素個數卻可能增加不少。又由上述三個例子，對一有限集合，我們猜想其所形成之任一 σ -體，其元素個數皆為2的次方。這個猜想是對的，也不是特別難證明，不過此處我們略去此討論。

例1.8 設 $\Omega = R = (-\infty, \infty)$ ，令 $\mathcal{B}^1$ 表包含實數上所有開區間之最小的 σ -體。此最小的 σ -體 $\mathcal{B}^1$ 是如何產生？只要取所有這種 σ -體之交集，則由習題第11題知，此交集仍為一 σ -體，而當然就是最小的。任一集合 $A \in \mathcal{B}^1$ ，稱為一波瑞爾集合(Borel set)。由習題第12題知， $\mathcal{B}^1$ 亦包含諸如 $[a, b)$ ， $(a, b]$ 及 $[a, b]$ ， $a < b$ ，等區間。

現在我們可以定義機率函數了。

定義1.2 設 Ω 為一樣本空間， $\mathcal{F}$ 為 Ω 之一些子集合所形成之一 σ -體。則以 $\mathcal{F}$ 為定義域，且滿足下述條件的函數 P ，便稱為一機率函數(probability function):

- (i) $\forall A \in \mathcal{F}, P(A) \geq 0$;
- (ii) $P(\Omega) = 1$;
- (iii) 若 $A_1, A_2, \dots \in \mathcal{F}$, 且為每對互斥, 則

$$(1.1) \quad P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)。$$

在上述定義中, $(\Omega, \mathcal{F}, P)$ 便構成一**機率空間**(probability space)。而條件(i)-(iii)通常稱為**機率的公理**, 或**柯莫果洛夫公理**(Kolmogorov axioms)。柯莫果洛夫(Kolmogorov, 1903-1987), 被認為是二十世紀俄國最重要的數學家。他在西元1930年左右, 以集合論及**測度論**(measure theory), 經由建立一些公理, 開拓現代機率理論。機率論也因此堂而皇之的進入數學的領域, 並逐漸成為數學主要的一分支。

由定義1.2知, 對同一樣本空間 Ω , 可以有很多不同的機率函數 P , 而且定義域 $\mathcal{F}$ 也不一定相同。例如, 由例1.5-1.7, 對同一 Ω , σ -體可以有好多個。之前我們對樣本空間之每一子集合, 皆稱為事件, 在較高等的機率論裡, 只對所選定的 σ -體 $\mathcal{F}$ 中的元素才稱為事件, 而且只對 $\mathcal{F}$ 中的元素(即事件), 才給機率。對 $\forall A \in \mathcal{F}$, 以 $P(A)$ 表事件 A 發生之機率。 Ω 之子集合若不是 $\mathcal{F}$ 中的元素, 則不用知道(事實上也沒給)其機率。

機率空間是機率論的基礎。若只是處理一些簡單情況下的機率, 則有時可以不用太在意機率空間究竟是什麼, 但在處理一些較細膩的問題時, 所涉及的機率空間是什麼就可能很重要了。有時必須講清楚所欲討論問題之機率空間為何。相當於讓人了解你的活動範圍及遊戲規則為何。

例1.9 設 $\Omega = \{1, 2, 3, 4\}$, 取

$$\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2, 3, 4\}\}$$

為一 σ -體。再令

$$\begin{aligned} P(\emptyset) &= 0, & P(\Omega) &= 1, \\ P(\{1\}) &= 0.2, & P(\{2, 3, 4\}) &= 0.8。 \end{aligned}$$

可檢驗 P 為一機率函數, 因此 $(\Omega, \mathcal{F}, P)$ 便構成一機率空間。

在上例中, 我們並不知道 $\{2\}$ 之機率, 因 $\{2\} \notin \mathcal{F}$ 。 Ω 共有16個子集合, 其中有12個我們都未給機率值。這樣設計的想法是什麼?

從數學的角度來看, 這樣的設計是較一般的。即允許不取 Ω 之所有子集合所形成那個最大的 σ -體。允許因興趣(或需要), 只對 Ω 之某些子集合才給其機率。但又規定興趣要合理, 合理到你有興趣的子集合要構成一 σ -體。以例1.9為例, 投擲一四面體, 各面分別有數字1, 2, 3, 4。假設出現1你就贏否則就輸, 所以你會想知道1出現之機率, 因此可以不用關心2出現的機率, 或是諸如2或3(即 $\{2, 3\}$)出現的機率。但若不想知道輸(即 $\{2, 3, 4\}$ 出現)的機率, 便有些駝鳥了。這是在定義1.1裡, 我們要求若 $A \in \mathcal{F}$, 則 $A^c \in \mathcal{F}$ 的原因。定義1.1中的條件(ii)也可比照地解釋。

在大部分統計學的教科書中, 若樣本空間 Ω 為可數的, 且未特別聲明, 則 σ -體就是取成 Ω 之所有子集合所形成之集合。對於不可數的集合, 如實數集合, σ -體則取成包含所有開區間(open interval)之最小的 σ -體。

例1.10 投擲一銅板 n 次, 觀測所得之正面數。則 $\Omega = \{0, 1, 2, \dots, n\}$ 。令 $\mathcal{F}$ 表 Ω 之所有子集合所形成之集合, 共有 2^{n+1} 個元素。再令 $P(\{i\}) = p_i$, $i = 0, 1, \dots, n$, 其中 p_i 's 滿足 $\sum_{i=0}^n p_i = 1$ 。又對 $\forall A \in \mathcal{F}$, 令

$$P(A) = \sum_{i \in A} p_i。$$

則 $(\Omega, \mathcal{F}, P)$ 便構成一機率空間。

例1.11 投擲一銅板直到出現一正面才停止, 並記載總共之投擲數。則 $\Omega = \{1, 2, \dots\}$, 為一無限集合。取 $\mathcal{F}$ 為 Ω 之所有子集合所形成之集合。又設第一個正面在第 k 次投擲才出現之機率為 2^{-k} , $k \geq 1$ 。令 $f(k) = 2^{-k}$, $k \geq 1$, 則 $\sum_{k=1}^{\infty} f(k) = \sum_{k=1}^{\infty} 2^{-k} = 1$ 。對 $\forall A \in \mathcal{F}$, 令

$$P(A) = \sum_{k \in A} 2^{-k}。$$

則 $(\Omega, \mathcal{F}, P)$ 構成一機率空間。

例如, 若令 A 表投擲數為偶數之事件, 即 $A = \{2, 4, \dots\}$, 則

$$P(A) = \sum_{k=1}^{\infty} 2^{-2k} = \frac{1}{3}。$$

一般而言有下述二結果。

例1.12 令 $\Omega = \{\omega_1, \omega_2, \dots\}$ 為一可數的集合(因此也可以是有限集合), f 為一定義在 Ω 上的實函數, 滿足

$$f(\omega_i) \geq 0, i \geq 1, \text{ 且 } \sum_{i=1}^{\infty} f(\omega_i) = 1。$$

令 $\mathcal{F}$ 為 Ω 之所有子集合所形成之集合。再令

$$P(A) = \sum_{\omega \in A} f(\omega), A \in \mathcal{F},$$

則 $(\Omega, \mathcal{F}, P)$ 即構成一**離散型**的機率空間。

如果是投擲一公正的骰子, 則對應上例,

$$f(\omega_i) = \frac{1}{6}, \omega_i \in \Omega = \{1, 2, 3, 4, 5, 6\}。$$

例1.13 設 Ω 為實數 R 上之一區間(有限或無限均可), f 為 Ω 上之一實函數, 滿足

$$f(\omega) \geq 0, \omega \in \Omega, \text{ 且 } \int_{\Omega} f(\omega) d\omega = 1。$$

令 $\mathcal{F}$ 為包含 Ω 上所有開區間之最小的 σ -體。再令

$$P([a, b]) = \int_a^b f(\omega) d\omega, a, b \in \Omega, \text{ 且 } a < b。$$

則 $(\Omega, \mathcal{F}, P)$ 即構成一**連續型**的機率空間。

機率函數有一些基本性質, 我們列出如下, 證明則留在習題。

定理1.2 設 $(\Omega, \mathcal{F}, P)$ 為一機率空間。則

- (i) $P(\emptyset) = 0$;
- (ii) $P(A) \leq 1, A \in \mathcal{F}$;
- (iii) $P(A^c) = 1 - P(A), A \in \mathcal{F}$ 。

定理1.3 P 為一機率函數, A, B 為二事件。則

- (i) $P(B \setminus A) = P(B \cap A^c) = P(B) - P(A \cap B)$;
- (ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$;
- (iii) 若 $A \subset B$, 則 $P(B \setminus A) = P(B) - P(A)$, 且 $P(A) \leq P(B)$ 。

因 $P(A \cup B) \leq 1$, 故由上述定理, 即得一關於交集的不等式, 稱做**邦弗朗尼不等式**(Bonferroni's inequality):

$$(1.2) \quad P(A \cap B) \geq P(A) + P(B) - 1。$$

當二事件之交集的機率不易求時, 可藉此不等式給出該機率之一下限。

若 $P(B) = 1$, 由(1.2)式立即可得

$$P(A \cap B) = P(A)。$$

定理1.4 設有一機率空間 $(\Omega, \mathcal{F}, P)$ 。

- (i) 設事件 $C_1, C_2, \dots$ 構成 Ω 之一分割, 則

$$(1.3) \quad P(A) = \sum_{i=1}^{\infty} P(A \cap C_i), A \in \mathcal{F};$$

- (ii) 對任意事件 $A_1, A_2, \dots$,

$$(1.4) \quad P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)。$$

(1.4)式即為**波爾不等式**(Boole's inequality, Boole(1815-1864), 為英國數學家)。

若樣本空間 Ω 為一有限集合, 且 Ω 中之每一結果發生的**可能性相同**(所謂 equally likely), 則任一事件發生之機率, 只要計算該事件中元素的個數,

除以 Ω 中元素的個數即可。

例1.14 在一副52張的撲克牌(poker)之梭哈遊戲中, 在所分到的5張牌, 試分別求拿到四條, 一對與三條, 兩對, 三條, 一對, 同花, 順, 同花順之機率。在此所謂一對, 指拿到兩張一樣點數的牌, 外加三張不同點數的牌, 餘類推。

解. 機率空間 Ω 中有 $\binom{52}{5} = 2,598,960$ 個元素。假設牌洗得很好, 則可合理地假設每一手牌出現的機率為 $1/2,598,960$ 。因此(符號“ $\doteq$ ”表近似相等)

$$P(\text{四條}) = \frac{13 \cdot 48}{2,598,960} = \frac{624}{2,598,960} \doteq 0.00024,$$

$$P(\text{一對與三條}) = \frac{13 \cdot 12 \cdot \binom{4}{2} \cdot \binom{4}{3}}{2,598,960} = \frac{3,744}{2,598,960} \doteq 0.0014,$$

$$P(\text{兩對}) = \frac{\binom{13}{2} \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot 44}{2,598,960} = \frac{123,552}{2,598,960} \doteq 0.0475,$$

$$P(\text{三條}) = \frac{13 \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot 4^2}{2,598,960} = \frac{54,912}{2,598,960} \doteq 0.0211,$$

$$P(\text{一對}) = \frac{13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3}{2,598,960} = \frac{1,098,240}{2,598,960} \doteq 0.4226,$$

$$P(\text{同花}) = \frac{4 \cdot \binom{13}{5}}{2,598,960} = \frac{5,148}{2,598,960} \doteq 0.00198,$$

$$P(\text{順}) = \frac{9 \cdot 4^5}{2,598,960} = \frac{9,216}{2,598,960} \doteq 0.00355,$$

$$P(\text{同花順}) = \frac{4 \cdot 9}{2,598,960} = \frac{36}{2,598,960} \doteq 0.0000139。$$

將機率按小至大排出為同花順, 四條, 一對與三條, 同花, 順, 三條, 兩對, 一對。大英百科全書(Encyclopedia Britannica)說這是傳統的順序(traditional ranking)。我們現在知道, 這其實是可以計算出來的。在賭局中出現的機率愈小則便算愈贏。

上例就是所謂**排列**(permutation)與**組合**(combination)的題材, 大家在中學時代便學過了。在古典的機率模式中, 於丟骰子、玩撲克牌的遊戲裡, 往往是基於結果之出現有相同的可能性, 因此以排列組合的

手法來計算機率。在此，自一集合 Z 中取 k 個相異元素，所形成的一有序列 $(z_1, z_2, \dots, z_k)$ ，稱為 Z 之一 k 個元素的排列。而 Z 之一 k 個相異元素的子集 $\{z_1, z_2, \dots, z_k\}$ ，則稱為 Z 之一 k 個元素的組合。若 Z 中有 n 個元素，且 $k \leq n$ ，則 Z 之 k 個元素的排列有 $n(n-1)\cdots(n-k+1)$ 種， Z 之 k 個元素的組合有 $n(n-1)\cdots(n-k+1)/k!$ 種，其中

$$k! = k(k-1)\cdots 3 \cdot 2 \cdot 1, k \geq 1,$$

$$0! = 1。$$

我們通常採用記號

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}, n \geq 1, n \geq k \geq 0,$$

稱為**二項係數**(binomial coefficient)。又對非負整數 $n_1, n_2, \dots, n_k, k \geq 1$ ，且滿足 $\sum_{j=1}^k n_j = n$ ，

$$(1.5) \quad \binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!},$$

稱為**多項係數**(multinomial coefficient)。由(1.5)式得

$$\binom{n}{n_1} = \binom{n}{n_1, n_2}, n_1 + n_2 = n, n_1, n_2 \geq 0。$$

1.2 條件機率與獨立性

在上節中，我們在一樣本空間上定義機率。有時得到一些資訊，則根據所獲得的資訊，要修訂樣本空間，因而機率空間可能也會改變。這就是所謂**條件機率**(conditional probability)。

在數學裡不會有這種情況。給定某個數是2，它就一直是2。在機率裡，某事件的機率是有可能因情況而變。這本來是不奇怪的，但因大部分的人受數學的薰陶較久，而數學裡通常是處理“不變”的問題，所以在學習機率時，看到機率值居然會改變，便不易理解。

假設生男生女的機率各為 $1/2$ 。則隨機抽取一個學生會是男或女的機率也就大約是 $1/2$ 。但若知此學生是高雄女中的學生，則會是女生的機率

就是1了, 因高雄女中沒有收男學生。由於獲得資訊, 機率隨之而變, 其實是合理的, 否則就失去收集資訊的目的。受機率的訓練, 會使我們具有隨機的概念, 對事物的研判, 便能與時推移。

定義2.1 設 A, B 為樣本空間 Ω 中二事件, 且 $P(B) > 0$ 。則在給定 B 發生之下, A 之條件機率, 以 $P(A|B)$ 表之, 定義為

$$(2.1) \quad P(A|B) = \frac{P(A \cap B)}{P(B)}。$$

在條件機率的定義中, B 成為新的樣本空間: $P(B|B) = 1$ 。也就是原先的樣本空間 Ω 修正為 B 。所有事件發生之機率, 都要先將其針對與 B 的關係做修正。例如, 若 A 與 B 為互斥事件, 且 $P(B) > 0$, 則因 $P(A \cap B) = 0$, 故 $P(A|B) = 0$; 若 $P(A)$ 亦為正, 則此時亦有 $P(B|A) = 0$ 。

例2.1 在例1.14中, 我們看到在梭哈遊戲裡, 要拿到四條的機率很低。但若發了三張牌, 皆拿到 K , 則此時會拿到四條之機率為何?

解.令 B 表事件發三張牌皆為 K , A 表拿到四條的事件。則 $A \cap B = \{ \text{前三張皆為} K, \text{第四、五張有一張} K, \text{一張非} K \}$, 且

$$P(A \cap B) = \frac{\binom{4}{3}}{\binom{52}{3}} \cdot \frac{48}{\binom{49}{2}},$$

$$P(B) = \frac{\binom{4}{3}}{\binom{52}{3}}。$$

因此

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{48}{\binom{49}{2}} = \frac{2}{49}。$$

機率提高很多。

條件機率也可用來求非條件下的機率。由(2.1)式得

$$(2.2) \quad P(A \cap B) = P(A|B)P(B)。$$

故若知道 $P(A|B)$ 及 $P(B)$, 則可得到 $P(A \cap B)$ 。當然亦有

$$(2.3) \quad P(A \cap B) = P(B|A)P(A),$$

只要 $P(A) > 0$ 。結合(2.2)與(2.3)式, 得

$$(2.4) \quad P(A|B) = \frac{P(B|A)P(A)}{P(B)}。$$

今後即使不特別聲明, 上式要成立就隱含著 $P(A)$ 及 $P(B)$ 皆為正。

(2.4)式為**貝氏定理**(Bayes' Rule)之一特例, 這是英國牧師Thomas Bayes(1702-1761)首先提出, 因而命名。不過也有人認為法國的大數學家**拉普拉斯**(Laplace, 1749-1827)才是第一位明確給出此定理者, 所以應稱為**拉普拉斯公式**(Laplace's formula)。

定理2.1 設 $A_1, A_2, \dots$ 為某樣本空間之一分割。則對任一事件 B ,

$$(2.5) \quad P(B) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i)。$$

定理2.2 貝氏定理. 設 $A_1, A_2, \dots$ 為樣本空間之一分割, 則對任意 $i \geq 1$ 及任一事件 B , 只要 $P(B) > 0$,

$$(2.6) \quad P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}。$$

例2.2 有甲、乙、丙三囚犯, 國王以抽籤決定釋放其中一位, 處決另兩位。然後他告訴獄卒那一位將被釋放, 但要求獄卒不可先透露。甲先是要求獄卒告訴他那一位會被釋放, 遭到拒絕後, 改為問獄卒, 乙、丙中那一位會被處決。獄卒經過一番思考, 遂(誠實地)告訴甲, 乙會遭處決。他認為這樣做並未違反國王的規定, 原因為:

乙、丙二人, 至少有一會遭處決, 因此他並未提供甲有關甲是否會被釋放的有用資訊。

甲聽到獄卒說乙會被處決後很高興。原先他只知道自己有 $1/3$ 的機會遭釋放, 現因只剩他與丙了, 所以他會被釋放的機會提高至 $1/2$ 。

究竟獄卒與甲誰的分析，何者才是正確的呢？

解.令 A, B, C 分別表甲、乙、丙三人會被釋放的事件。如果我們考慮的結果是誰會被釋放，則 $A \cup B \cup C$ 就是樣本空間。由假設 $P(A) = P(B) = P(C) = 1/3$ 。令 K 表獄卒說乙會被處決的事件。必須先了解的是，若乙、丙皆會被處決，獄卒其實是任意挑一位告訴甲誰會被處決。而若乙將被釋放，獄卒只能告訴甲，丙會被處決，若丙將被釋放，獄卒只能告訴甲，乙會被處決。我們想求 $P(A|K)$ 。首先由貝氏定理：

$$\begin{aligned} P(K) &= P(K|A)P(A) + P(K|B)P(B) + P(K|C)P(C) \\ &= \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{1}{2}。 \end{aligned}$$

因此

$$\begin{aligned} P(A|K) &= \frac{P(\text{獄卒說乙會被處決, 且甲被釋放})}{P(K)} \\ &= \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{2}} = \frac{1}{3}。 \end{aligned}$$

換句話說，在獄卒告訴甲，乙會被處決後，甲會被釋放的機會(即 $P(A|K)$)仍維持為 $1/3$ 。那獄卒所提供的資訊是否毫無用處呢？那倒未必。假設丙偷聽到獄卒與甲的對話，則他知道他會被釋放的機會(即 $P(C|K)$)提高至 $2/3$ 。而若乙偷聽到獄卒與甲的對話，則便知沒有活命的機會了(即 $P(B|K) = 0$)。

最後， K 的機率為 $1/2$ ，直觀上是對的，這點讓各位自行想一想。

上例顯示對條件機率，必須要謹慎處理，否則極易犯錯。另外，我們也注意到原先 $P(A) + P(B) + P(C) = 1$ ，而 $P(A|K) + P(B|K) + P(C|K)$ 亦仍為 1 。這是對的，在習題裡(第18題)我們會給更一般的情況。

其次我們來看**獨立性**(independence)。在機率論裡，獨立是一很重要的概念。不過這種獨立的概念，是所謂**統計的獨立**(statistically independent)，或稱**隨機的獨立**(stochastically independent)，與日常生活裡的主權獨立，金雞獨立等的意義不同。獨立性也是數學裡難以找到與其對應之

一種概念。

定義2.2 二事件 A, B , 稱為**相互獨立**(mutually independent, 簡稱**獨立**), 若滿足

$$(2.7) \quad P(A \cap B) = P(A)P(B)。$$

上述定義是如何產生的呢? 在討論條件機率時, 有時事件 A 之發生, 對事件 B 之發生的機率沒有影響。即

$$(2.8) \quad P(B|A) = P(B)。$$

又由(2.4)式得

$$(2.9) \quad P(A|B) = \frac{P(B|A)P(A)}{P(B)} = P(A),$$

亦即此時 B 之發生對 A 也沒有影響。再由(2.2)式, 即得(2.7)式成立。在(2.8)及(2.9)二式中, 分別要有 $P(A)$, 及 $P(A)$ 與 $P(B)$ 為正的條件。不過(2.7)式對 $P(A)$ 或 $P(B)$ 為0(此時(2.7)式左、右均為0)仍成立。我們也就以(2.7)式當做 A, B 獨立的條件。當 $P(A) > 0$ 時, (2.7)與(2.8)二式等價, 但採用(2.7)式的好處是將二事件 A 與 B 對稱地對待, 且較易推廣到超過兩個事件獨立的情況(見底下定義2.3)。

當事件 A 與 B 獨立時, 由 B 之發生, 對事件 A 得不到任何**推論**(inference)。因此直觀上 A 與 B 獨立, 會導致 A 與 B^c 獨立。事實上有下述定理。

定理2.3 設 A 與 B 獨立, 則 A 與 B^c , A^c 與 B , A^c 與 B^c 皆獨立。

有些事件的獨立與否較易判斷。底下 Ω 表樣本空間, A, B 表事件, 讀者可練習判斷(i)至(vi)中, 所給的二事件, 那些是獨立。

- (i) A 與 $\emptyset$, (ii) A 與 Ω , (iii) A 與 A , (iv) A 與 A^c ,
- (v) A 與 B , 其中 $A \subset B$, (vi) A 與 B , 其中 $A \cap B = \emptyset$ 。

不過有時二事件並不易判斷是否為獨立。如在例2.2, 因

$$P(K \cap A) = \frac{1}{2} \cdot \frac{1}{3} = P(K)P(A),$$

故 K 與 A 獨立。但 K 與 B , K 與 C 皆不獨立。再給一例如下。

例2.3 投擲一公正的骰子兩次, 則樣本空間為

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), \dots, (6, 1), \dots, (6, 6)\},$$

共包含36個元素, 其中 (i, j) 表第一次出現點數 i , 第二次出現點數 j 。取 σ -體為包含 Ω 所有子集合之集合(共有 2^{36} 個元素)。令 A 表兩次投擲出現的點數相同, B 表兩次投擲點數和介於7與10間, C 表兩次投擲點數和為2, 7或8。則可求出

$$\begin{aligned} P(A) &= \frac{6}{36} = \frac{1}{6}, \\ P(B) &= \frac{6+5+4+3}{36} = \frac{1}{2}, \\ P(C) &= \frac{1+6+5}{36} = \frac{1}{3}. \end{aligned}$$

又

$$\begin{aligned} P(A \cap B \cap C) &= P(\text{出現}(4, 4)) = \frac{1}{36} = P(A)P(B)P(C), \\ P(A \cap B) &= P(\text{出現}(4, 4) \text{ 或 } (5, 5)) = \frac{2}{36} \neq P(A)P(B), \\ P(B \cap C) &= P(\text{點數和為} 7 \text{ 或 } 8) = \frac{6+5}{36} \neq P(B)P(C), \\ P(A \cap C) &= P(\text{出現}(1, 1) \text{ 或 } (4, 4)) = \frac{2}{36} = P(A)P(C). \end{aligned}$$

因此 A 與 B , B 與 C 均不獨立, 但 A 與 C 獨立。

本例不但顯示 $P(A \cap B \cap C) = P(A)P(B)P(C)$, 不一定導致 A, B, C 中任二獨立, 而且二事件獨立與否, 有時並不易看出。

定義2.3 對 $n \geq 2$, 事件 $A_1, A_2, \dots, A_n$, 若滿足 A_i 與 A_j 獨立, $\forall i \neq j$, 則為**每對獨立**, 或稱**兩兩獨立**(pairwise independent); 若對所有 $k \geq 2$,

及 $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, 滿足

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \prod_{j=1}^k P(A_{i_j}),$$

則為**相互獨立**(仍簡稱獨立)。

由定義2.3可得下述推論, 其證明留在習題中。

系理2.1 設 $A_1, A_2, \cdots, A_n$ 為相互獨立, 則對所有 $k \geq 2$, 及 $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ 為相互獨立; 特別地 $A_1, A_2, \cdots, A_n$ 為每對獨立。

由定義2.3得, 三事件 A, B, C 要相互獨立, 就須有

$$P(A \cap B) = P(A)P(B),$$

$$P(B \cap C) = P(B)P(C),$$

$$P(A \cap C) = P(A)P(C),$$

且

$$P(A \cap B \cap C) = P(A)P(B)P(C)。$$

即要驗證四個等式成立。一般而言, n 個事件要相互獨立, 就須有

$$\binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n} = (1+1)^n - \binom{n}{1} - \binom{n}{0} = 2^n - n - 1$$

個等式成立。

當 A, B, C 相互獨立時, 因 $P(B \cap C) = P(B)P(C)$, 故

$$P(A \cap (B \cap C)) = P(A)P(B \cap C),$$

即 A 與 $B \cap C$ 亦獨立。反之, 若 A 與 $B \cap C$ 不獨立, 則 A, B, C 便不相互獨立。

例2.4 設樣本空間 Ω 包含 a, b, c 三字母之六個排列, 及 $(a, a, a), (b, b, b), (c, c, c)$ 等共九個元素, 每一元素之機率均設為 $1/9$ 。令事件 A_k 表 Ω 中一元

素 (x, y, z) 之第 k 個位置為 a , $k = 1, 2, 3$, 如 $A_1 = \{(a, b, c), (a, c, b), (a, a, a)\}$ 。顯然

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3},$$

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{9}。$$

因此 A_1, A_2, A_3 為每對獨立。但因

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{9} \neq P(A_1)P(A_2)P(A_3),$$

故 A_1, A_2, A_3 不相互獨立。

事實上若 A_1, A_2 均發生, 則 A_3 必發生, 所以 A_3 與 $A_1 \cap A_2$ 不獨立, 因此 A_1, A_2, A_3 不相互獨立。

上例顯示雖相互獨立導致每對獨立, 但每對獨立並不一定導致相互獨立。

例2.5 考慮投擲一銅板三次的試驗。令“正正反”表頭兩次出現正面, 第三次出現反面, 餘類推。則樣本空間為

$$\Omega = \{\text{正正正}, \text{正正反}, \text{正反正}, \text{反正正}, \text{正反反}, \text{反正反}, \text{反反正}, \text{反反反}\}。$$

對每一樣本點給定機率 $1/8$ 。又令 A_i 表第 i 次投擲得到正面之事件。則

$$P(A_1) = P(\{\text{正正正}, \text{正正反}, \text{正反正}, \text{正反反}\}) = 4 \cdot \frac{1}{8} = \frac{1}{2}。$$

同理 $P(A_2) = P(A_3) = 1/2$ 。即每次投擲會得到正面之機率皆為 $1/2$ 。換句話說, 此為一公正的銅板。底下我們來驗證, 對此機率模型, 事件 A_1, A_2, A_3 為相互獨立:

$$P(A_1 \cap A_2 \cap A_3) = P(\{\text{正正正}\}) = \frac{1}{8} = P(A_1)P(A_2)P(A_3),$$

$$P(A_1 \cap A_2) = P(\{\text{正正正}, \text{正正反}\}) = \frac{2}{8} = P(A_1)P(A_2)。$$

同理可得 $P(A_1 \cap A_3) = P(A_1)P(A_3)$, $P(A_2 \cap A_3) = P(A_2)P(A_3)$ 。因此 A_1, A_2, A_3 相互獨立。

讀者可自行驗證，若欲使 $P(A_1) = P(A_2) = P(A_3) = 1/2$ ，且 A_1, A_2, A_3 相互獨立，則每一樣本點之機率皆須為 $1/8$ 。但若僅是 A_1, A_2, A_3 相互獨立，則 $P(A_1), P(A_2)$ 及 $P(A_3)$ 便不須皆為 $1/2$ ；而也可有 $P(A_1) = P(A_2) = P(A_3) = 1/2$ ，但 A_1, A_2, A_3 不相互獨立的情況（習題第27題）。

例2.6 假設生男與生女的機率均為 $1/2$ 。若一家庭中有兩個小孩，已知其中之一為男孩，試求另一為女孩之機率。

解.令事件 A, B 分別表兩個小孩為一男一女，及有一小孩為男孩之事件。則

$$P(A \cap B) = P(A) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

$$P(B) = \frac{3}{4}。$$

因此

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/2}{3/4} = 2/3。$$

對二事件 A, B ，有時我們以 $P(A, B)$ 表 $P(A \cap B)$ ，即以逗號“，”取代交集符號“ $\cap$ ”。

例2.7 在上例中，若亦知該男孩為老大，試回答同樣的問題。

解.令事件 C 表老大為男孩。則

$$P(A \cap C) = P(\text{老大為男孩, 老二為女孩}) = \frac{1}{4},$$

$$P(C) = \frac{1}{2}。$$

因此

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{1/4}{1/2} = \frac{1}{2}。$$

比較上二例，我們發現，知道一家庭中，兩個孩子之一為男孩，跟知道該男孩亦為老大，對另一小孩性別之判斷，會得到不同的結果。這是條件

機率奧妙之處：有時以為沒有用的資訊其實卻是有用的。總之，正確地利用所得之資訊是很重要的。

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。

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”。

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

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

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

$$\begin{aligned} P(I_A = 1) &= P(A), \\ P(I_A = 0) &= 1 - P(A). \end{aligned}$$

利用指示函數, 可將每一事件對應一隨機變數。

1.4 分佈函數及機率密度函數

對一隨機變數 X , 我們定義其**累積分佈函數**(cumulative distribution function, 縮寫為 c.d.f.), 通常簡稱為**分佈函數**, 縮寫為 d.f., 又稱**分布函數**, 或**分配函數**:

$$(4.1) \quad F_X(x) = P_X(X \leq x), x \in R.$$

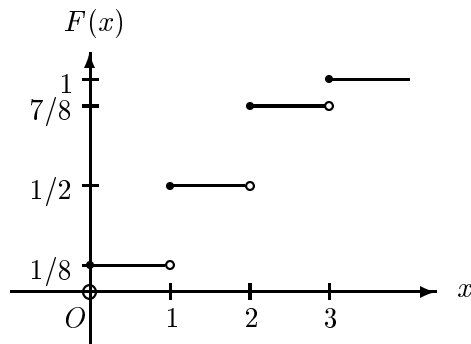
如果隨機變數 X 很清楚, $F_X(x)$ 也可只寫成 $F(x)$ 。

分佈函數是一定義域為實數, 對應域為區間 $[0, 1]$ 的函數。由機率空間, 隨機變數, 至分佈函數, 當只給分佈函數時, 源頭之機率空間為何, 可能是毫不清楚的。完全不同的機率空間, 是可能得到相同的分佈函數, 我們陸續會說明。

由於是以(4.1)式來定義分佈函數, 這可解釋為何在上一節裡, 我們曾提到, 須滿足(3.1)式才是隨機變數。否則若有一 $x \in R$, 使得 $\{X \leq x\} \notin \mathcal{F}$, 則便不知此 $\{X \leq x\}$ 之機率, 因此就得不到 $F_X(x)$ 了。

例4.1 投擲一公正的銅板3次, 令 X 表所得之正面數。則因 $P(X = 0) = 1/8$, $P(X = 1) = 3/8$, $P(X = 2) = 3/8$, $P(X = 3) = 1/8$, 故

$$F(x) = \begin{cases} 0 & , x < 0, \\ P(X = 0) = \frac{1}{8} & , 0 \leq x < 1, \\ P(X = 0) + P(X = 1) = \frac{1}{2} & , 1 \leq x < 2, \\ P(X = 0) + P(X = 1) + P(X = 2) = \frac{7}{8} & , 2 \leq x < 3, \\ \sum_{i=0}^3 P(X = i) = 1 & , x \geq 3. \end{cases}$$

圖4.1 為 F 的圖形(注意橫軸與縱軸之比例不同)。圖4.1 例4.1中 $F(x)$ 之圖形

我們說明一下圖4.1。分佈函數是定義在整個實數上，而非只定義在 X 取值的集合 $\Omega_1 = \{0, 1, 2, 3\}$ 。例如，

$$F(2.7) = P(X \leq 2.7) = P(X = 0, 1, 2) = P(X \leq 2) = \frac{7}{8}。$$

換句話說 $F(2.7) = F(2)$ 。事實上對 $2 \leq x < 3$ ， $F(x)$ 之值皆為 $7/8$ 。 F 的函數圖形，在 $\forall x_i \in \Omega_1$ ，有一**跳升**(jump)，跳升的高度即 $P(X = x_i)$ 。又 $x < 0$ 時， $F(x) = 0$ ，因 X 不會取負值。而 $x \geq 3$ 時， $F(x) = 1$ ，因 X 不會取大於3之值。也可看出 F 的圖形為一**階梯函數**(step function)，並且是**右連續**(right continuous)。即

$$(4.2) \quad \lim_{t \downarrow x} F(t) = F(x), x \in R,$$

其中 $t \downarrow x$ 表 t 由 x 的右側漸減趨近至 x 。又 $t \uparrow x$ 表 t 由 x 的左側漸增趨近至 x 。

退化的隨機變數，其分佈函數亦為階梯函數。即若 $X \equiv c$ ，則 $F(x) = 0$ ， $x < c$ ， $F(x) = 1$ ， $x \geq c$ 。反之若 X 有前述分佈函數，則知 $X \equiv c$ 。

我們將分佈函數具有的性質，列成一定理，證明則略去。

定理4.1 一函數 F 為一分佈函數，若且唯若下述諸條件成立：

- (i) $0 \leq F(x) \leq 1, \forall x \in R$;
- (ii) F 為一非漸減函數;
- (iii) F 為一右連續函數;
- (iv) $\lim_{x \rightarrow -\infty} F(x) = 0$, 且 $\lim_{x \rightarrow \infty} F(x) = 1$ 。

在上述定理中，條件(i)其實可以拿掉，因(ii)與(iv)便導致(i)成立。不過我們還是寫出來，使分佈函數的性質更清楚。

設隨機變數 X 之分佈函數為 F ，則對 $\forall a < b$ ，因 $\{X \leq a\} \subset \{X \leq b\}$ ，故由定理1.3(iii)，得

$$\begin{aligned} P(a < X \leq b) &= P(\{X \leq b\} \setminus \{X \leq a\}) = P(X \leq b) - P(X \leq a) \\ &= F(b) - F(a)。 \end{aligned}$$

即

$$(4.3) \quad P(a < X \leq b) = F(b) - F(a), a < b。$$

我們以 F 表示出一些與 X 有關之常見事件的機率如下：

$$P(X = a) = F(x)\text{-之圖形在 } x = a \text{ 跳升之高度,}$$

$$P(X > a) = 1 - F(a),$$

$$P(X \geq a) = 1 - F(a) + P(X = a),$$

$$P(X \leq b) = F(b),$$

$$P(X < b) = F(b) - P(X = b)。$$

又對 $a < b$,

$$P(a < X \leq b) = F(b) - F(a),$$

$$P(a < X < b) = F(b) - F(a) - P(X = b),$$

$$P(a \leq X \leq b) = F(b) - F(a) + P(X = a),$$

$$P(a \leq X < b) = F(b) - F(a) + P(X = a) - P(X = b)。$$

例4.2 投擲一出現正面的機率為 p 之銅板, $0 < p < 1$, 直到出現一正面才停止。令 X 表總共之投擲數。則

$$(4.4) \quad P(X = x) = p(1-p)^{x-1}, x \geq 1。$$

因此對 $x = 1, 2, \dots$,

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \sum_{i=1}^x P(X = i) = \sum_{i=1}^x p(1-p)^{i-1} = 1 - (1-p)^x。 \end{aligned}$$

F 之圖形仍為階梯函數, 在兩整數之間是水平的。一隨機變數, 若具有這種分佈函數, 稱為有**幾何分佈**(geometric distribution)。

有了 F 可得 X 落在任一區間之機率。如

$$\begin{aligned} P(3.1 < X \leq 7.2) &= P(3 < X \leq 7) = F(7) - F(3) \\ &= (1 - (1-p)^7) - (1 - (1-p)^3) \\ &= (1-p)^3 - (1-p)^7。 \end{aligned}$$

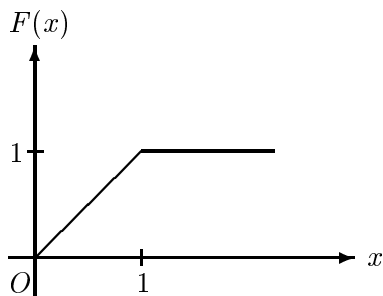
本例亦顯示, 對一分佈函數 F , 不一定有一 $x \in R$, 使得 $F(x) = 1$ 。可以猜測, 也不一定有一 $x \in R$, 使得 $F(x) = 0$ 。

我們亦給一分佈函數為連續的例子。

例4.3 設

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1。 \end{cases}$$

圖4.2給出 F 之圖形。 F 為一連續函數, 滿足定理4.1中之四個條件, 故為一分佈函數。

圖4.2 例4.3中 $F(x)$ 之圖形

底下再給一例。

例4.4 考慮

$$(4.5) \quad F(x) = \frac{1}{1 + e^{-x}}, x \in R.$$

其中 e 為自然對數的底, 其值約為2.71828。則因 $e^{-x} > 0, \forall x \in R$, 故 $1 + e^{-x} > 1$, 且 $0 < (1 + e^{-x})^{-1} < 1, \forall x \in R$, 即定理4.1中之條件(i)成立。其次因 $1 + e^{-x}$ 在 $x \in R$ 為漸減, 故其倒數 $(1 + e^{-x})^{-1}$ 為漸增, 即條件(ii)成立。又 F 為連續函數, 因此條件(iii)成立。最後因

$$\lim_{x \rightarrow \infty} e^{-x} = 0, \quad \lim_{x \rightarrow -\infty} e^{-x} = \infty,$$

故

$$\lim_{x \rightarrow \infty} F(x) = 1, \quad \lim_{x \rightarrow -\infty} F(x) = 0.$$

即得證 F 確為一分佈函數。這是標準的羅吉斯分佈(standard logistic distribution)。一般的形式見習題第17題。

我們繪出 $y = F(x)$ 之圖形於圖4.3。這是一典型連續型分佈函數之圖形: 以 $y = 1$ 及 $y = 0$ 為水平漸近線, 沒有垂直漸近線, 沒有相對極值, 但有反曲點(inflexion point), 本例圖形之反曲點為 $(0, 1/2)$ 。不過反曲點有時不存在, 如對下述二分佈函數

$$F_1(x) = \begin{cases} 0 & , x < 0, \\ 1 - e^{-x} & , x \geq 0, \end{cases}$$

$$F_2(x) = \begin{cases} 0 & , x < 0, \\ x & , 0 \leq x < 1, \\ 1 & , x \geq 1, \end{cases}$$

其圖形便皆無反曲點。此二者之圖形不夠平滑： F_1 在 $x = 0$ ， F_2 在 $x = 0, 1$ ，有較尖銳的點，造成反曲點不存在。

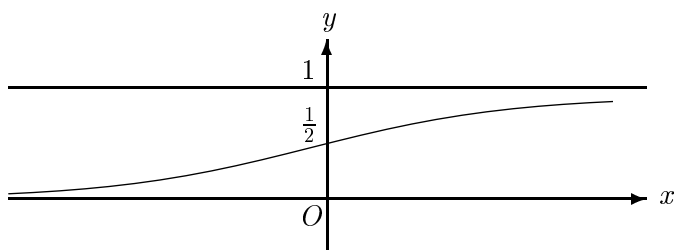


圖4.3 $y = F(x) = \frac{1}{1 + e^{-x}}$, $x \in R$, 之圖形

隨機變數 X 之分佈函數 F 若為連續函數，便稱 X 為**連續型**；若 F 為階梯函數，便稱 X 為**離散型**。也可稱分佈函數為連續型或離散型。不過也有既非離散亦非連續的隨機變數。例如考慮

$$F_1(x) = \begin{cases} \frac{1}{2} \frac{1}{1+e^{-x}} & , x < 0, \\ \frac{1}{2} + \frac{1}{2} \frac{1}{1+e^{-x}} & , x \geq 0. \end{cases}$$

則 F_1 之圖形並非階梯函數，且除了在 $x = 0$ 有一高度為 $1/2$ 之跳升(所以 $P(X = 0) = 1/2$)，其餘地方為連續。但可驗證 F_1 確為一分佈函數，只是既非連續型，亦非離散型。

離散型隨機變數，其分佈函數的圖形，有跳升之處的集合，必為可數的，可能是一有限集合，或可數的無限集合。

有時我們給一分佈函數“ $F(x), x > a$ ”，其中 a 為一常數，就表 $F(x) = 0, \forall x \leq a$ 。又若說有一分佈函數“ $F(x), a < x < b$ ”，就表 $F(x) = 0, x \leq a$ ，且 $F(x) = 1, x \geq b$ 。其他形式(如 $F(x), x < b$)也有類似的說明。

分佈函數可決定一隨機變數之機率分佈。

假設有一樣本空間 Ω ，機率函數為 P ， X 為一定義在 Ω 上之隨機變數，值域以 Ω_1 表之，分佈函數以 F 表之。則 Ω_1 表新的樣本空間， F 表新的機率函數。等於是將在 Ω 上的機率問題，轉為在 Ω_1 上的機率問題。不過要注意，不但在同一樣本空間，可以有不同的隨機變數，但分佈函數卻相同，在不同的樣本空間，亦可分別定義隨機變數，而有相同的分佈函數。

我們再給二隨機變數**分佈相同**(identically distributed)之定義。

定義4.1 二隨機變數 X, Y ，若滿足

$$(4.6) \quad P(X \leq x) = P(Y \leq x), x \in R,$$

則稱**分佈相同**，並以 $X \stackrel{d}{=} Y$ 表之。

若以 F, G 分別表 X 及 Y 之分佈函數。 X 與 Y 分佈相同，表 $F(x) = G(x)$ ， $\forall x \in R$ 。即 X 與 Y 有相同的分佈函數。分佈相同就是分佈函數相同，也可說有**共同的分佈**。但分佈相同與隨機變數相同是不一樣的，甚至二差異很大的隨機變數，可有相同的分佈函數。我們以下二例來說明。

例4.5 投擲一公正的銅板三次，令 X 表所得之正面數， Y 表所得之反面數。則 X 當然不等於 Y ($X \neq Y$)。因若 $X = 0$ 則 $Y = 3$ 。事實上，可看出 X 與 Y 二者永不相等，即 $X(\omega) \neq Y(\omega)$ ， $\forall \omega \in \Omega$ ，其中 Ω 如例2.5所給(如當 $\omega = \text{正正正}$ ，則 $X(\omega) = 3 \neq 0 = Y(\omega)$)。但因

$$P(X = i) = P(Y = i), i = 0, 1, 2, 3,$$

故 X 與 Y 有相同的分佈函數，即 $X \stackrel{d}{=} Y$ 。

例4.6 投擲一公正的銅板三次，令 Z 表所得之正面數。投擲一公正的骰子三次，令 W 表所得點數為偶數的次數。 Z 與 W 為兩個不相干的隨機變數，二者當然不等。

我們強調過隨機變數即函數。而二函數 f 與 g 相等的定義是

- (i) f 與 g 有相同的定義域,
- (ii) 對每一定義域中的 x , $f(x) = g(x)$ 。

現因 Z 的定義域為例2.5所給之 Ω , 共有8個元素, 而 W 之定義域為

$$\{111, 112, \dots, 666\},$$

共有216個元素, 其中 ijk 表第一次得 i , 第二次得 j , 第三次得 k , $1 \leq i, j, k \leq 6$, W 與 Z 之定義域不同, 所以不等。但明顯可看出

$$P(Z = i) = P(W = i), i = 0, 1, 2, 3,$$

故 $Z \stackrel{d}{=} W$ 。

在例4.5中, X 與 Y 雖為不等的隨機變數, 但二者的關係卻易知道, 即 $X + Y = 3$ 。我們也可求諸如 $X > Y$, $X = Y + 1$ 等事件之機率。但在例4.6中, $Z = 3$ 時, W 是多少, 是完全不知道的。而事件 $Z > W$, 或 $Z = W + 1$ 的機率, 在所給的資訊下, 也無法求得。至於要再給什麼資訊才能求出上述那些事件的機率? 我們留待第三章再說明。

雖 $X \stackrel{d}{=} Y$ 時不一定導致 $X = Y$, 但 $X = Y$ 卻必導致 $X \stackrel{d}{=} Y$ 。此因 $X = Y$ 表 $X(\omega) = Y(\omega)$, $\forall \omega \in \Omega$, 故

$$P(X \leq x) = P(\{\omega | X(\omega) \leq x\}) = P(\{\omega | Y(\omega) \leq x\}) = P(Y \leq x)。$$

隨機變數除了分佈相同, 亦有大小的概念, 見習題第24題。

對一隨機變數 X , 除了伴隨分佈函數 F_X , 亦有一與其關係密切的函數, 即**機率密度函數**(probability density function, 縮寫為p.d.f.), 或是**機率質量函數**(probability mass function, 縮寫為p.m.f.), 前者是對連續型, 後者是對離散型。不過為了簡便, 通常我們不分連續或離散, 皆說是機率密度函數。我們給定義如下。

設 X 為一隨機變數, 以 F_X 為其分佈函數(若不至混淆可只寫 F , 省略 X)。若 X 為離散型, 則下述函數 f_X (有時省略 X 只寫 f), 便稱為 X 之**機率密度函數**:

$$f_X(x) = P(X = x), x \in R。$$

由於離散型隨機變數，其分佈函數的圖形，有跳升之處的集合為可數的(證明在此略去)，而此集合就是使 $P(X = x)$ 為正之所有 x 的集合。設以 C 表此集合，則對 $\forall x \notin C$ ， $f_X(x) = 0$ 。因此離散型的機率密度函數，有時表示成 $f_X(x)$ ， $x \in C$ ，其中 C 為某一可數的集合，可能是 $\{1, 2, \dots, 10\}$ ，也可能是非負整數的集合 $\{0, 1, 2, \dots\}$ 。至於 $x \notin C$ 時的 $f_X(x)$ 往往不提，因其值皆為0。所以即是將此機率密度函數的定義域取成 C 。可看出對一實數上的集合 A ，

$$P(A) = P(X \in A) = \sum_{x \in A \cap C} f_X(x),$$

且

$$F_X(x) = \sum_{t \leq x, t \in C} f_X(t), x \in R。$$

例4.7 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} (8/15)(1/2)^x, & x = 0, 1, 2, 3, \\ 0 & , \text{其他}。 \end{cases}$$

則可求出其分佈函數為

$$F(x) = \begin{cases} 0 & , x < 0, \\ 8/15 & , 0 \leq x < 1, \\ 12/15 & , 1 \leq x < 2, \\ 14/15 & , 2 \leq x < 3, \\ 1 & , x \geq 3。 \end{cases}$$

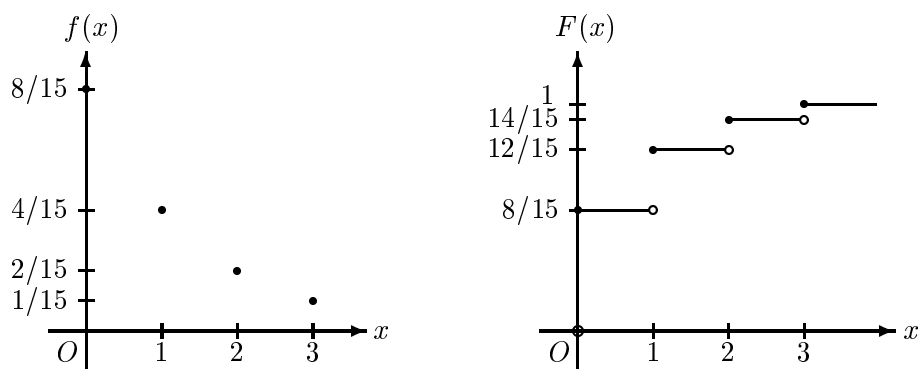
為一階梯函數。

又在求 $F(2.5)$ 時，可依定義得

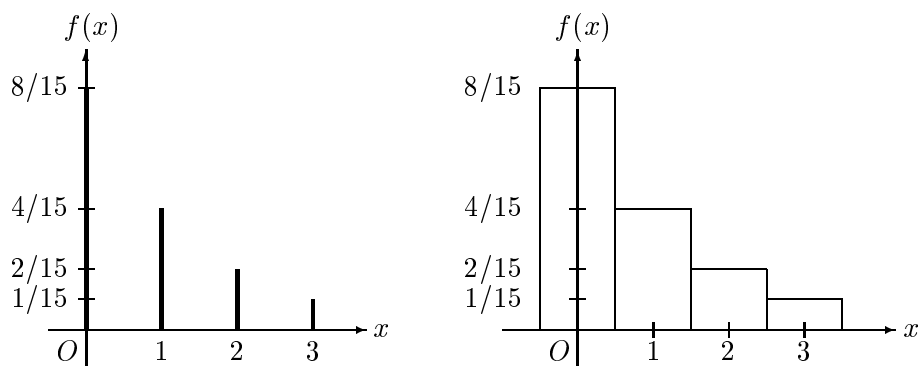
$$\begin{aligned} F(2.5) &= \sum_{x \leq 2.5} f(x) = f(0) + f(1) + f(2) \\ &= \frac{8}{15} + \frac{4}{15} + \frac{2}{15} = \frac{14}{15}, \end{aligned}$$

或利用已求出之 $F(x)$ ，因 $2.5 \in [2, 3)$ ，故 $F(2.5) = 14/15$ 。

圖4.4給出 $f(x)$ 及 $F(x)$ 之圖形。

圖4.4 例4.7中 $f(x)$ 及 $F(x)$ 之圖形

離散型機率密度函數之圖形，有時以線段表之。而對只取整數值之隨機變數，其機率密度函數的圖形，有時以一些小長方形表示，每一小長方形之底部中點為取值處，高度為該點之機率，因此小長方形面積和為1。這樣表示機率密度函數的圖形，稱為**直方圖**(histogram)，機率便對應直方圖之面積。例4.7中之 $f(x)$ 亦可繪成如圖4.5中之二圖，右圖即直方圖。

圖4.5 例4.7中 $f(x)$ 之圖形

其次若 X 為連續型，則其機率密度函數的定義為滿足下式之非負函數 f_X (稍後在例4.9會指出 f_X 並不唯一)：

$$(4.7) \quad F_X(x) = \int_{-\infty}^x f_X(t) dt, x \in R.$$

由定理4.1, 因

$$\lim_{x \rightarrow \infty} F_X(x) = 1,$$

故知 f_X 除了非負之外, 尚須滿足

$$(4.8) \quad \int_{-\infty}^{\infty} f_X(x) dx = 1。$$

即 $f_X(x)$ 之圖形與 x 軸間之面積為1。

對連續型的隨機變數 X , 由於 $P(X = x) = 0, \forall x \in R$, 即 X 取任一定值的機率為0。故在求 X 落在某一區間的機率時, 該區間是否包含端點是無關緊要的。即

$$\begin{aligned} P(a < X < b) &= P(a \leq X < b) = P(a \leq X \leq b) \\ &= P(a < X \leq b) = F_X(b) - F_X(a), \end{aligned}$$

$\forall a < b$ 。若以機率密度函數表示, 上述任一機率皆等於

$$\begin{aligned} P(a < X < b) &= \int_{-\infty}^b f_X(t) dt - \int_{-\infty}^a f_X(t) dt \\ &= \int_a^b f_X(t) dt。 \end{aligned}$$

例4.8 設隨機變數 X 之分佈函數為(見例4.4)

$$F(x) = \frac{1}{1 + e^{-x}}, x \in R。$$

此為一連續函數, 故

$$P(a < X < b) = F(b) - F(a) = \frac{1}{1 + e^{-b}} - \frac{1}{1 + e^{-a}}, a < b。$$

連續型隨機變數, 其機率密度函數, 也可有 $f(x), c \leq x \leq d$, 的型式, 即表 $f(x) = 0, \forall x > d$ 或 $x < c$, 即所對應的隨機變數, 取值不會大於 d 也不會小於 c 。

又由微積分基本定理(Fundamental Theorem of Calculus), 若 f_X 為連續函數, 則由(4.7)式,

$$(4.9) \quad F'_X(x) = f_X(x), x \in R。$$

可藉由上式得到此時之機率密度函數。又二隨機變數, 若有相同的機率密度函數, 則其分佈相同。

例4.9 設隨機變數 X 之分佈函數如例4.3所給。則若取

$$f(x) = \begin{cases} 0, & x \leq 0, \\ 1, & 0 < x < 1, \\ 0, & x \geq 1。 \end{cases}$$

(或簡單地寫 $f(x) = 1, 0 < x < 1$), 便滿足 $F(x) = \int_{-\infty}^x f(t)dt, x \in R$ 。但 $F(x)$ 在 $x = 0$ 及 1 並不可微($f(x)$ 在 $x = 0$ 及 1 也不連續)。在 $x = 0$ 及 1 , 無法利用微積分基本定理, 經由微分 F 而得到機率密度函數 f 。我們只能說 $F'(x) = f(x), \forall x \neq 0, 1$ 。圖4.6給出 $f(x)$ 之圖形。對連續型的隨機變數, 機率密度函數在有限個點的值改變, 是不用在意的, 因此前述 f 寫成

$$f(x) = 1, 0 \leq x < 1, \quad f(x) = 1, 0 < x \leq 1, \quad \text{或} \quad f(x) = 1, 0 \leq x \leq 1,$$

等皆可。這三個 f 亦皆滿足 $F(x) = \int_{-\infty}^x f(t)dt, x \in R$, 因此皆可當做對應 F 之機率密度函數。

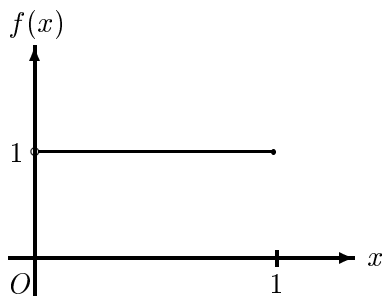


圖4.6 例4.9中 $f(x)$ 之圖形

例4.10 考慮如例4.4之標準的羅吉斯分佈，即

$$F_X(x) = \frac{1}{1 + e^{-x}}, x \in R。$$

則

$$(4.10) \quad f_X(x) = F'_X(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, x \in R。$$

$F_X(x)$ 及 $f_X(x)$ 也可分別寫成

$$F_X(x) = \frac{e^x}{1 + e^x}, \quad f_X(x) = \frac{e^x}{(1 + e^x)^2}, x \in R。$$

可驗證 $f_X(-x) = f_X(x)$ 。

一函數 f 若滿足 $f(-x) = f(x), \forall x \in R$ ，便稱為**偶函數**(even function)。偶函數之圖形對稱於 $x = 0$ 。更一般地，若存在一常數 a ，使得

$$(4.11) \quad f(a - x) = f(a + x), x \in R,$$

或等價地

$$(4.12) \quad f(x) = f(2a - x), x \in R,$$

便說 **f 對稱於 a** 。即若一隨機變數之p.d.f對稱於 a ，便稱其分佈對稱於 a 。當 $a = 0$ ，也稱此為**對稱的分佈**。也可以分佈函數來表示。分佈函數 F 若滿足

$$(4.13) \quad F(a - x) = 1 - F(a + x), x \in R,$$

或等價地

$$(4.14) \quad F(x) = 1 - F(2a - x), x \in R,$$

便稱分佈對稱於 a 。附帶一提，函數 f 若滿足 $f(-x) = -f(x), \forall x \in R$ ，便稱為**奇函數**(odd function)。例如， $f(x) = \sin x, x \in R$ ，為奇函數； $f(x) = \cos x, x \in R$ ，則為偶函數。

必須一提的是，連續型的分佈不一定有機率密度函數，此因一連續函數不一定可微。也就是有可能 F_X 為連續，但不存在函數 $f_X(x)$ ，使得(4.7)式成立。不過在本書中，對連續型的隨機變數，除非特別聲明，我們均假設其機率密度函數存在。

由分佈函數可得到機率密度函數，反之亦有下述定理，證明雖不難，但此處仍略去。

定理4.2 一函數 f 若滿足下述二條件，必為某一隨機變數之機率密度函數：

- (i) $f(x) \geq 0, \forall x \in R$,
- (ii) $\sum_x f(x) = 1$ (離散型)，或 $\int_{-\infty}^{\infty} f(x)dx = 1$ (連續型)。

機率密度函數是可以很容易得到的，只要 $f(x)$ 為非負且 $\sum_x f(x)$ (離散型)，或 $\int_{-\infty}^{\infty} f(x)dx$ 為正且有限，則經除以該和或積分的值，便立即可得一機率密度函數，見下例。

例4.11 設

$$f(i) = 3^{-i}, i = 1, 2, 3, \dots。$$

因

$$\sum_{i=1}^{\infty} f(i) = \frac{1}{2},$$

故

$$(4.15) \quad f_1(x) = 2 \cdot 3^{-x}, x = 1, 2, 3, \dots,$$

即成為一機率密度函數。

另外，設

$$g(x) = e^{-3x}, x > 0。$$

因

$$\int_0^{\infty} e^{-3x} dx = \frac{1}{3},$$

故

$$g_1(x) = 3e^{-3x}, x > 0,$$

亦成爲一機率密度函數。

對連續型的隨機變數, 其機率密度函數 $f(x)$ 與 x 軸間之面積等於1; 對離散型的隨機變數, 其機率密度函數之高度和等於1。我們較少去繪隨機變數的圖形(雖然此爲一函數), 但分佈函數及機率密度函數的圖形倒是常繪的。(4.10)及(4.15)式所給之兩個機率密度函數的圖形分別給在圖4.7及圖4.8。

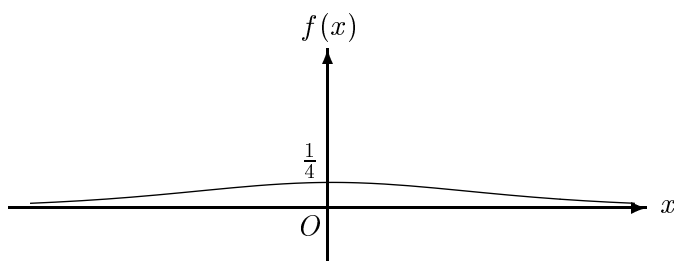


圖4.7 $f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, x \in R$, 之圖形

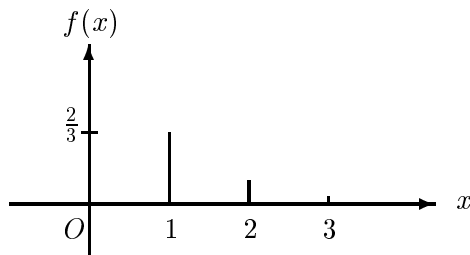


圖4.8 $f(x) = \frac{2}{3^x}, x = 1, 2, 3, \dots$, 之圖形

對一隨機變數 X , 以 $F(x)$ 爲分佈函數, $f(x)$ 爲機率密度函數, 有時可以 $X \sim F(x)$ 或 $X \sim f(x)$ 表之。至於若 $X \stackrel{d}{=} Y$, 也可以 $X \sim Y$ 表之。對一隨機變數, 給定其機率密度函數, 或分佈函數, 或其他可以唯一決定其機率如何分配的規則(函數也是一種規則), 皆稱爲**給定其分佈**。而所謂**決定其分佈**, 就是表示給出前三者之任一。譬如說指出有標準的羅吉斯分佈。

本節最後再給幾個例子。

如前所述，給一非負的函數 f ，只要其圖形與 x 軸間之面積為1，便是一機率密度函數。最簡單的機率密度函數，是在一有限的區間中為常數。即設有一區間 (a, b) ， $a < b$ ，且 f 在該區間為一常數。顯然此常數為 $(b - a)^{-1}$ 。即

$$f(x) = \frac{1}{b - a}, a < x < b。$$

具有這種機率密度函數的隨機變數，稱為在區間 (a, b) 有**均勻分佈**(uniform distribution)，以 $\mathcal{U}(a, b)$ 表之。例4.9中之 X 便有 $\mathcal{U}(0, 1)$ 分佈。同理亦有 $\mathcal{U}[a, b)$ ， $\mathcal{U}(a, b]$ ，及 $\mathcal{U}[a, b]$ 分佈。

還有那些較簡單的隨機變數？底下給一例。

例4.12 考慮如圖4.9之函數 f 的圖形，其中 $a < b$ ，且 $f(a) = 0$ 。欲使 f 之圖形與 x 軸間之面積為1(三角形面積等於底乘以高除以2)， $f(b)$ 應為 $2/(b - a)$ 。求出連接兩點 $(a, 0)$ 及 $(b, 2/(b - a))$ 之直線方程式 $y = 2(x - a)/(b - a)^2$ 。故

$$f(x) = \begin{cases} 2(x - a)/(b - a)^2, & a < x < b, \\ 0 & , \text{其他}, \end{cases}$$

為一機率密度函數。由此可求出分佈函數(習題第2題)

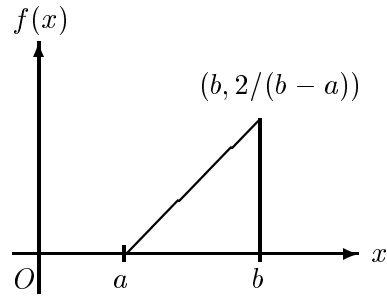
$$F(x) = \begin{cases} 0 & , x < a, \\ (x^2 - 2ax + a^2)/(b - a)^2, & a \leq x < b, \\ 1 & , x \geq b。 \end{cases}$$

例4.13 假設我們希望圖4.10中之圖形為一機率密度函數 f 的圖形。令 $f(8) = c$ 。利用兩個小三角形面積和須為1，得

$$\frac{1}{2}(3c + 2c) = 1。$$

解出 $c = 2/5$ ，即 $f(8) = 2/5$ 。連接 $(5, 0)$ 及 $(8, 2/5)$ 之直線方程式為

$$y = \frac{2}{15}(x - 5),$$

圖4.9 例4.12中 $f(x)$ 之圖形

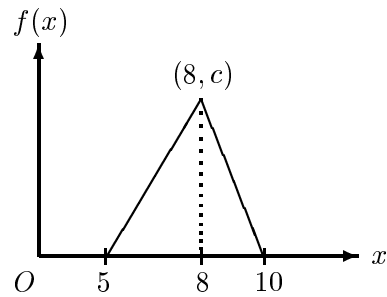
連接 $(8, 2/5)$ 及 $(10, 0)$ 之直線方程式為

$$y = -\frac{1}{5}(x - 10) = \frac{1}{5}(10 - x),$$

故

$$f(x) = \begin{cases} 0 & , x \leq 5, \\ \frac{2}{15}(x - 5) & , 5 < x < 8, \\ \frac{1}{5}(10 - x) & , 8 \leq x < 10, \\ 0 & , x \geq 10. \end{cases}$$

至於分佈函數，留給各位自行求出(習題第3題)。

圖4.10 例4.13中 $f(x)$ 之圖形

離散型隨機變數之機率密度函數代表機率，因此其值皆不能大於1。但連續型的隨機變數，其機率密度函數的值，卻可以大於1，見下例。

例4.14 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} 10(x - 7.7), & 7.7 \leq x < 8.1, \\ 40(8.2 - x), & 8.1 \leq x < 8.2, \\ 0 & , \text{其他}。 \end{cases}$$

圖4.11給出 $f(x)$ 之圖形。

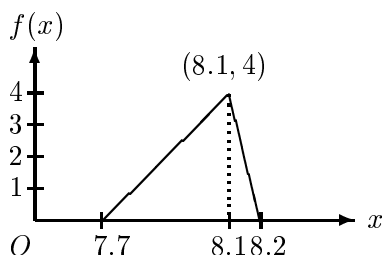


圖4.11 例4.14中 $f(x)$ 之圖形

不難驗證圖4.11中三角形的面積為1(習題第4題)。由例4.13的推導, 相信大家應已了解這類機率密度函數是如何得到的。又 $f(8.1) = 4 > 1$ 。可看出只要三角形的底部長度夠短, 則其高度可以超過任一給的正數。

底下來求 $X > 8$ 之機率。

$$\begin{aligned} P(X > 8) &= \int_8^{8.1} 10(x - 7.7)dx + \int_{8.1}^{8.2} 40(8.2 - x)dx \\ &= 10\left(\frac{1}{2}x^2 - 7.7x\right)\Bigg|_8^{8.1} + 40\left(8.2x - \frac{1}{2}x^2\right)\Bigg|_{8.1}^{8.2} \\ &= 0.35 + 0.2 = 0.55。 \end{aligned}$$

在習題第4題, 我們會要各位給出 X 之分佈函數。

對連續型的隨機變數, 不只機率密度函數 $f(x)$ 的值可大於1, f 也可不為有界(bounded)。所謂有界即存在一常數 $k > 0$, 使得 $|f(x)| < k$, $\forall x \in R$ 。例如, 令

$$(4.16) \quad f(x) = 2x^{-1/2}, 0 < x < 1。$$

則

$$\lim_{x \downarrow 0} f(x) = \infty,$$

因此 f 不為有界。但 f 為一機率密度函數(見習題第5題)。

我們給最後一例。

例4.15 設隨機變數 X 之機率密度函數為 $f(x) = e^{-x}$, $x > 0$ 。由於

$$\int_0^{\infty} e^{-x} dx = (-e^{-x}) \Big|_0^{\infty} = 1,$$

故 f 的確為一機率密度函數。又

$$P(2 \leq X \leq 7) = \int_2^7 e^{-x} dx = (-e^{-x}) \Big|_2^7 = e^{-2} - e^{-7},$$

$$P(-2 < X \leq 5) = \int_0^5 e^{-x} dx = (-e^{-x}) \Big|_0^5 = 1 - e^{-5}。$$

由於 X 不取非正的值, 所以 $P(-2 < X \leq 5)$ 與 $P(0 < X \leq 5)$ 相等。分佈函數亦可求出。即

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt \\ &= \begin{cases} 0 & , x \leq 0, \\ 1 - e^{-x} & , 0 < x < \infty. \end{cases} \end{aligned}$$

圖4.12中給出 f 及 F 之圖形。附帶一提, 如果前述機率密度函數改為 $g(x) = e^{-x}$, $x \geq 0$ 。雖 $g(0) = 1 \neq 0 = f(0)$, 但仍會得到相同的分佈函數 $F(x)$ 。

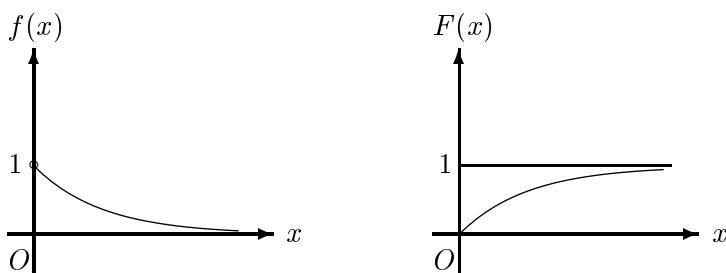


圖4.12 $f(x) = e^{-x}$, $x > 0$, 及 $F(x) = 1 - e^{-x}$, $x > 0$, 之圖形

1.5 變數代換

在很多試驗裡，對一隨機變數 X ，有時我們可能會對 X 之某一函數有興趣。譬如說量測得到圓的半徑 X ，但我們想知道圓的面積 $Y = \pi X^2$ 。即使不論實務上的需要，在1.3節我們已指出，在數學裡我們常會討論與一變數有關的函數之性質或行為。由樣本空間 Ω 的引出隨機變數即為一例。對一隨機變數 X ，有時我們會想知道它的一個函數 $Y = g(X)$ 的行為。 Y 便稱為 X 之變數代換(change of variable)。而由 X 至 $Y = g(X)$ ，也可視為一種隨機變數之轉換(transformation of a random variable)。 Y 仍為一隨機變數。通常 X 的分佈知道，而因 Y 的值是由 X 的值所決定，故 Y 的分佈，可由所給之 X 的分佈來決定。

對每一實數的子集合 A ，

$$(5.1) \quad P(Y \in A) = P(g(X) \in A) = P(X \in g^{-1}(A))。$$

在此

$$g^{-1}(A) = \{x | x \in R, g(x) \in A\}。$$

對於新的隨機變數 Y ，由(5.1)式，欲求 Y 落在某一集合 A 之機率，仍要回到關於 X 的機率。但若我們求出 Y 的分佈函數(或機率密度函數)，當然就可以不用理會 X 了。

若 X 為離散型的隨機變數，則 $Y = g(X)$ 的機率密度函數，只要留意由 X 至 Y 的變換是否為1-1(等價於 g 是否為嚴格單調函數)，則藉由(5.1)式，通常可很快地得到。底下為一典型的例子。

例5.1 設 X 為一離散型的隨機變數，且p.d.f.為

$$f_X(x) = \begin{cases} 0.15 & , x = 0, 3, \\ 0.20 & , x = 1, 2, \\ 0.30 & , x = 4, \\ 0 & , \text{其他。} \end{cases}$$

令 $Y = X^2$, $Z = (X - 2)^2$ 。我們分別來決定 Y 及 Z 之分佈。

首先 X 之可能的值為 0, 1, 2, 3, 4, Y 之可能的值則為 0, 1, 4, 9, 16。因 X 只取非負的值, 故由 X 至 Y 的變換為 1-1 (雖 $g(x) = x^2$, $x \in R$, 並非嚴格單調函數, 但在 $x \geq 0$ 處, g 為嚴格單調)。又

$$P(Y = y) = P(X^2 = y) = P(X = \sqrt{y}), y = 0, 1, 4, 9, 16。$$

故得 Y 之 p.d.f. 為

$$f_Y(y) = P(Y = y) = \begin{cases} 0.15, & y = 0, \\ 0.20, & y = 1, \\ 0.20, & y = 4, \\ 0.15, & y = 9, \\ 0.30, & y = 16, \\ 0, & \text{其他。} \end{cases}$$

可驗證 $\sum_y f_Y(y)$ 仍等於 1。

其次 Z 之可能的值為 0, 1, 4。 $X = 0$ 及 4, 皆對應 $Z = 4$; $X = 1$ 及 3, 皆對應 $Z = 1$; $X = 2$, 則對應 $Z = 0$ 。故由 X 至 Z , 並非 1-1 的變換。 Z 之 p.d.f. 為

$$f_Z(z) = P(Z = z) = \begin{cases} f_X(2) = 0.20, & z = 0, \\ f_X(1) + f_X(3) = 0.35, & z = 1, \\ f_X(0) + f_X(4) = 0.45, & z = 4。 \end{cases}$$

可驗證 $\sum_z f_Z(z)$ 仍等於 1。

底下再給一例。

例 5.2 設隨機變數 X 之 p.d.f. 為

$$f(x) = \frac{1}{31} 2^{4-x}, x = 0, 1, 2, 3, 4。$$

令 $Y = X^2 + 1$ 。此為由 X 至 Y 之 $1-1$ 變換, $X = \sqrt{Y-1}$ 。 Y 取值在 1, 2, 5, 10, 17。故 Y 之 p.d.f. 為

$$f_Y(y) = \frac{1}{31} 2^{4-\sqrt{y-1}}, y = 1, 2, 5, 10, 17。$$

其次令 $Z = g(X) = X^3 - 3X^2 + X + 2$ 。則因

$$g(0) = 2, \quad g(1) = 1, \quad g(2) = 0, \quad g(3) = 5, \quad g(4) = 22,$$

故此仍為由 X 至 Z 之 $1-1$ 變換。雖不易給出 $X = g^{-1}(Z)$ 之 g^{-1} 的型式, 但仍可寫出 Z 之 p.d.f. 如下:

$$\begin{aligned} f_Z(0) &= P(X = 2) = \frac{4}{31}, \\ f_Z(1) &= P(X = 1) = \frac{8}{31}, \\ f_Z(2) &= P(X = 0) = \frac{16}{31}, \\ f_Z(5) &= P(X = 3) = \frac{2}{31}, \\ f_Z(22) &= P(X = 4) = \frac{1}{31}。 \end{aligned}$$

要知所謂函數, 只要每一定義域中的元素, 在對應域中恰有一元素與其對應即可, 不一定要能給出一個適用定義域中每一元素的型式(如上述 X 或 Y 之 p.d.f. 之對應)。

其次我們給二連續型隨機變數的例子。

例5.3 設隨機變數 X 有 $\mathcal{U}(0,1)$ 分佈。即其 p.d.f. 為

$$f(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他,} \end{cases}$$

分佈函數為

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1。 \end{cases}$$

令 $Y = -\lambda^{-1} \log(1 - X)$, 其中 $\lambda > 0$ 為一常數。又除非特別標出底, 機率與統計裡出現的對數符號 $\log$, 均為自然對數, 即以 e 為底。此為一由 X 至 Y 之 $1-1$ 變換。由於 X 介於 $0, 1$ 之間, 故 Y 取值在 $(0, \infty)$ 。則對 $\forall y > 0$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(-\lambda^{-1} \log(1 - X) \leq y) \\ &= P(\log(1 - X) \geq -\lambda y) \\ &= P(X \leq 1 - e^{-\lambda y}) \\ &= 1 - e^{-\lambda y}, y > 0. \end{aligned}$$

因此 Y 之機率密度函數為

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \lambda e^{-\lambda y}, & y > 0, \\ 0 & , y \leq 0. \end{cases}$$

即 Y 有 **指數分佈** (exponential distribution), 參數為 λ , 以 $\mathcal{E}(\lambda)$ 表之。上節例 4.15 中之 X 便是有 $\mathcal{E}(1)$ 分佈。

例 5.4 設隨機變數 X 之 p.d.f. 為

$$f_X(x) = \frac{2}{\pi(1+x^2)}, x > 0.$$

令 $Y = 1/X$ 。則此為一由 X 至 Y 之 $1-1$ 變換, 且 Y 取值在 $(0, \infty)$ 。我們想決定 Y 之 p.d.f.。

對 $\forall y > 0$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(1/X \leq y) \\ &= P(X \geq 1/y) = P(X > 1/y) \\ &= 1 - P(X \leq 1/y) \\ &= 1 - F_X(1/y). \end{aligned}$$

因此

$$f_Y(y) = \frac{dF_Y(y)}{dy} = f_X\left(\frac{1}{y}\right) \cdot \frac{1}{y^2}$$

$$\begin{aligned}
&= \frac{2}{\pi(1+1/y^2)} \cdot \frac{1}{y^2} \\
&= \frac{2}{\pi(1+y^2)}, y > 0.
\end{aligned}$$

由於 X 與 Y 之p.d.f.相同, 故 X 與 Y 有相同的分佈, 即 $X \stackrel{d}{=} Y$ 。對此特別的分佈, 雖經過一倒數的變換, 所得之新隨機變數, 其分佈卻未改變。

在上述推導中, 由於只是要求 Y 之p.d.f., 所以我們並未將 $F_Y(y)$ 明確地寫出。事實上

$$\begin{aligned}
F_Y(y) &= 1 - F_X(1/y) \\
&= 1 - \int_0^{1/y} f_X(x) dx \\
&= 1 - \int_0^{1/y} \frac{2}{\pi(1+x^2)} dx \\
&= 1 - \frac{2}{\pi} \arctan x \Big|_0^{1/y} \\
&= 1 - \frac{2}{\pi} \arctan\left(\frac{1}{y}\right), y > 0.
\end{aligned}$$

將此 $F_Y(y)$ 對 y 微分, 便得到前述 $f_Y(y)$, $y > 0$ 。

在例5.4中, p.d.f.取值在 $x > 0$ 處。如果想利用此p.d.f.定義出一偶函數, 且仍為一p.d.f., 就要將一半的份量, 移至 $x < 0$ 之處, 如此才仍為一p.d.f.。又定義 $f(0)$, 使 f 成為一連續函數。

即可令

$$(5.2) \quad f(x) = \frac{1}{\pi(1+x^2)}, x \in R.$$

此必仍為一p.d.f.(習題第1題)。具有型如(5.2)式之p.d.f.者, 稱為有標準的**柯西分佈**(Cauchy distribution, Cauchy(1789-1857)為著名的法國數學家), 以 $C(0, 1)$ 表之。現設隨機變數 X 有 $C(0, 1)$ 分佈, 令 $Y = 1/X$ 。雖 X 之p.d.f.的定義域包含 $x = 0$, 因此 Y 之定義似有問題, 但因 $P(X = 0) = 0$, 所以此情況不用擔心。仿例5.4的作法, 仍可得到 Y 有 $C(0, 1)$ 分佈。

另外, 在例5.4的推導過程中, 有

$$f_Y(y) = f_X\left(\frac{1}{y}\right) \cdot \frac{1}{y^2}。$$

即 Y 之p.d.f., 可將 X 之p.d.f.中的參數 x , 以 $1/y$ 取代, 再乘上 $1/y$ 對 y 微分的絕對值。而由 $Y = g(X) = 1/X$, 解出

$$X = g^{-1}(Y) = \frac{1}{Y}。$$

故 $f_Y(y)$ 可寫成

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|。$$

上述這些討論與觀察, 便引出底下兩個一般的定理。

定理5.1 設隨機變數 X 之分佈函數為 F_X , 機率密度函數為 f_X 。令 $Y = g(X)$, Y 之分佈函數以 F_Y 表之。又令

$$(5.3) \quad \Omega_1 = \{x | f_X(x) > 0\},$$

$$(5.4) \quad \Omega_2 = \{y | \text{存在 } x \in \Omega_1, \text{ 使得 } y = g(x)\}。$$

(i) 若 g 在 Ω_1 為嚴格漸增函數, 則 $F_Y(y) = F_X(g^{-1}(y))$, $\forall y \in \Omega_2$;

(ii) 若 g 在 Ω_1 為嚴格漸減函數, 則 $F_Y(y) = 1 - F_X(g^{-1}(y)-)$, $\forall y \in \Omega_2$ 。

我們先說明一下。 Ω_1 表 X 之機率密度函數取正值處。這種集合稱為分佈之**支撐**(support)。例如, 對 $\mathcal{U}(0, 1)$ 分佈, 支撐為區間 $(0, 1)$, 對 $\mathcal{E}(\lambda)$ 分佈, 支撐為 $(0, \infty)$ 。另外對一函數 F ,

$$F(x-) = \lim_{t \uparrow x} F(t),$$

表 F 在 x 之**左極限值**。當然如果 F 在 x 為左連續, 則 $F(x-) = F(x)$ 。同理可定義**右極限** $F(x+)$ 。只是若 F 為分佈函數, 因其為右連續, 故 $F(x+) = F(x)$, $\forall x \in R$ 。附帶一提, 利用左極限的符號, 可得對 $a < b$,

$$P(a < X < b) = F(b-) - F(a),$$

$$P(a \leq X < b) = F(b-) - F(a-),$$

$$P(a \leq X \leq b) = F(b) - F(a-)。$$

定理5.1之證明

(i) 設 g 在 Ω_1 為嚴格漸增。則

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \leq g^{-1}(y)) \\ &= F_X(g^{-1}(y))。 \end{aligned}$$

(ii) 設 g 在 Ω_1 為嚴格漸減。則

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\ &= P(X \geq g^{-1}(y)) \\ &= 1 - F_X(g^{-1}(y)-)。 \end{aligned}$$

證畢。

定理5.1適用連續及離離散型的隨機變數。當 Y 之機率密度函數為連續，則利用定理5.1，可由微分分佈函數而得。我們列成一定理。

定理5.2 設隨機變數 X 之p.d.f.為 f_X 。令 $Y = g(X)$ ，其中 g 為一嚴格單調函數， Ω_1 及 Ω_2 分別定義在(5.3)及(5.4)式中。設 f_X 在 Ω_1 連續，且 g^{-1} 在 Ω_2 有一連續的導數，則 Y 之p.d.f.為

$$(5.5) \quad f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, y \in \Omega_2,$$

且 $f_Y(y) = 0, y \notin \Omega_2$ 。

例5.5 設 X 有 $\mathcal{E}(\lambda)$ 分佈。令 $Y = \sqrt{X}$ 。則 $\Omega_1 = \Omega_2 = (0, \infty)$ ，且 X 之p.d.f. $f_X(x) = \lambda e^{-\lambda x}, x > 0$ ，在 Ω_1 中連續； $g^{-1}(y) = y^2$ 在 Ω_2 之導數 $2y$ 為連續。故 Y 之p.d.f.為

$$f_Y(y) = f(y^2) \cdot 2y = 2\lambda y e^{-\lambda y^2}, y > 0。$$

Y 稱為具有參數 λ 之**瑞萊分佈**(Rayleigh distribution)。

例5.6 設隨機變數 X 有 $\mathcal{U}(0, 1)$ 分佈。令 $Y = X/(1-X)$ 。試決定 Y 之p.d.f.及分佈函數。

解.由 $y = x/(1-x)$, 知 $\Omega_2 = (0, \infty)$ 。又解出

$$x = 1 - \frac{1}{y+1},$$

故 X 至 Y 的變換為 $1-1$ 。又

$$\frac{dx}{dy} = \frac{1}{(y+1)^2}。$$

因此 Y 之p.d.f.為

$$f_Y(y) = \frac{1}{(y+1)^2}, y > 0。$$

由此又得

$$F_Y(y) = \int_0^y \frac{1}{(t+1)^2} dt = -\frac{1}{t+1} \Big|_0^y = 1 - \frac{1}{y+1} = \frac{y}{y+1}, y > 0。$$

當然 Y 之分佈函數亦可如下直接求出：

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X/(1-X) \leq y) \\ &= P(X \leq y/(y+1)) = \frac{y}{y+1}, y > 0, \end{aligned}$$

其中最後一等式成立，是因 $P(X \leq x) = x, \forall x \in (0, 1)$ 。求出 $F_Y(y)$ 後，經由對 y 微分，便得到 $f_Y(y)$ ，與前述作法所得相同。

在上例中， $\Omega_1 = (0, 1), \Omega_2 = (0, \infty)$ 。但有時為了簡便，我們不特別寫 Ω_1, Ω_2 。如只寫 $0 < x < 1, y > 0$ (或 $0 < y < \infty$)。又在做變數代換時，若一開始給的變換是 $X = h(Y)$ ，其中 h 為一個 $1-1$ 函數，則此 h 便相當於前述 g^{-1} 。因此

$$(5.6) \quad f_Y(y) = f_X(h(y)) \left| \frac{d}{dy} h(y) \right|, y \in \Omega_2,$$

其中 Ω_2 為所求出之 y 的支撐。

在做變數代換時，函數 g 有時不為嚴格單調，因此定理5.2就不適用。但往往會有 g 在一個個的集合中，分別為嚴格單調(想想諸如多項函數、正弦及餘弦函數的圖形)，因此在每一個集合中，皆可解出 $X = g^{-1}(Y)$ ，則仍有辦法求出 Y 之分佈。如果 g 不是在一個個的集合為嚴格單調的話，自然就不易求出 Y 的分佈了。先給一例子。

例5.7 設 X 為一連續型的隨機變數，令 $Y = X^2$ 。則對 $\forall y > 0$ ， Y 之分佈函數為

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X < -\sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}), \end{aligned}$$

此處用到 X 為連續型，所以 $P(X < -\sqrt{y}) = P(X \leq -\sqrt{y})$ 。

現若 X 之p.d.f.存在，則

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy} (F_X(\sqrt{y}) - F_X(-\sqrt{y})) \\ &= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})。 \end{aligned}$$

我們發現 Y 之p.d.f.為兩項之和，分別對應區間 $[0, \infty)$ 及 $(-\infty, 0)$ 。 $y = g(x) = x^2$ 在前者有反函數 $x = \sqrt{y}$ ，在後者有反函數 $x = -\sqrt{y}$ 。

定理5.1推廣如下：

假設 Ω_1 可寫成有限個不相交集之聯集，譬如說 $D_0, D_1, D_2, \dots, D_k$ ，其中 $P(X \in D_0) = 0$ ， D_0 可能是一空集合，使得 $y = g(x)$ 在由每一 D_i 至 Ω_2 皆為1-1之對應， $i = 1, 2, \dots, k$ 。則 $\forall y \in \Omega_2$ ，在 D_i 中恰有一 x_i 與 y 對應， $i = 1, 2, \dots, k$ 。因此即有 k 個反函數 $g_1^{-1}(y), \dots, g_k^{-1}(y)$ 。又設 $g_i^{-1}(y)$ 在 Ω_2 有一連續的導數。則 Y 之p.d.f.為

$$(5.7) \quad f_Y(y) = \sum_{i=1}^k f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right|, y \in \Omega_2,$$

且 $f_Y(y) = 0, y \notin \Omega_2$ 。

重做例5.7。 $g(x) = x^2$ 在 $[0, \infty)$ 有反函數 $g_1^{-1}(y) = \sqrt{y}$, 在 $(-\infty, 0)$ 有反函數 $g_2^{-1}(y) = -\sqrt{y}$ 。故由(5.7)式, 直接可得

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})。$$

對一分佈函數, 我們定義其**分位函數**(quantile function, 有時也稱**反函數**(inverse function), 但與一般反函數的定義不同)如下:

$$F^{-1}(x) = \inf\{t \mid t \in R \text{ 且 } F(t) \geq x\}, x \in (0, 1)。$$

若一隨機變數 X 以 F 為分佈函數, 則 F^{-1} 可同時稱為 X 及 F 之分位函數。即使 F 不一定是連續, 也不一定嚴格漸增, F^{-1} 必定存在。但若 F 為連續, 且嚴格漸增, 則 $F^{-1}(x)$ 等於滿足 $F(t) = x$ 之唯一的 t 。此時 F^{-1} 與一般反函數的意義相同。所以此處 F^{-1} 乃擴充反函數之定義。

對 $\forall p \in (0, 1)$, 令 $x_p = F^{-1}(p)$ 。則 x_p 稱為此分佈之**第 p 分位數**(p th-quantile), 或**第 $100p$ 百分位數**(100pth-percentile)。特別地 $x_{0.25}$ 稱為**下四分位數**(lower quartile), $x_{0.75}$ 稱為**上四分位數**(upper quartile)。有些書分位數的定義與我們略有不同, 不過這些本書不擬多討論。我們先給 F^{-1} 的一些基本性質, 證明則留在習題。

定理 5.3 設 F^{-1} 為分佈函數 F 之分位函數。

- (i) 對 $\forall x, t \in R, F^{-1}(x) \leq t$, 若且唯若 $x \leq F(t)$;
- (ii) F^{-1} 為非漸減且左連續的函數;
- (iii) 若 F 為連續函數, 則 $F(F^{-1}(x)) = x, \forall x \in (0, 1)$ 。

雖然 F 為連續函數時, $F(F^{-1}(x)) = x$, 但在 F 的連續點 x , $F^{-1}(F(x))$ 卻不一定等於 x , 在 F^{-1} 的連續點 x , $F(F^{-1}(x))$ 也不一定等於 x 。見底下例5.9。

底下給兩個例子。

例5.8 設隨機變數 X 有 $\mathcal{E}(1)$ 分佈。則 $F(x) = 1 - e^{-x}$, $x > 0$ 。 F 之圖形如圖4.12所給。由於在 $x > 0$ 處, F 為嚴格漸增, 故 F^{-1} 即 F 之反函數。經由解 $F(t) = 1 - e^{-t} = x$, 得 $t = F^{-1}(x) = -\log(1 - x)$ 。 F^{-1} 之圖形給在圖5.1。注意 $\lim_{x \rightarrow 1-} F^{-1}(x) = \infty$ 。讀者也可回頭看例5.3, 猜測該例中的 Y 是如何設出來的。

F 之0.5分位數為 $-\log(1/2) = \log 2$, 下四分位數為 $-\log(3/4) = \log(4/3)$, 上四分位數為 $-\log(1/4) = \log 4$ 。

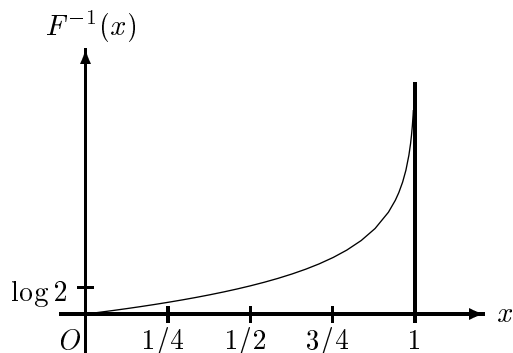


圖5.1 $F^{-1}(x) = -\log(1-x)$ 之圖形

例5.9 設有分佈函數

$$F(x) = \begin{cases} 0 & , x < 0, \\ 0.5x & , 0 \leq x < 1, \\ 0.5 & , 1 \leq x < 2, \\ 0.2x + 0.2 & , 2 \leq x < 4, \\ 1 & , x \geq 4. \end{cases}$$

則在 $x \in (0, 0.5]$, 解 $F(t) = 0.5t = x$, 得 $t = F^{-1}(x) = 2x$; 在 $x \in (0.5, 0.6]$, $t = F^{-1}(x) = 2$; 在 $x \in (0.6, 1]$, 解 $F(t) = 0.2t + 0.2 = x$, 得 $t = F^{-1}(x) = 5x - 1$ 。分位函數通常不定義在 $x = 0$ 及 $x = 1$, 不過在本例中, $F^{-1}(1) = 4$ 有定義。 F 之0.3分位數為0.6, 0.5分位數為1, 0.55分位數為2, 0.6分位數亦為2。 $F(t)$ 在 $t = 1.5$ 連續, 但

$$F^{-1}(F(1.5)) = F^{-1}(0.5) = 1 \neq 1.5。$$

又 $F^{-1}(x)$ 在 $x = 0.55$ 連續, 但

$$F(F^{-1}(0.55)) = F(2) = 0.6 \neq 0.55。$$

F 及 F^{-1} 之圖形如圖 5.2。

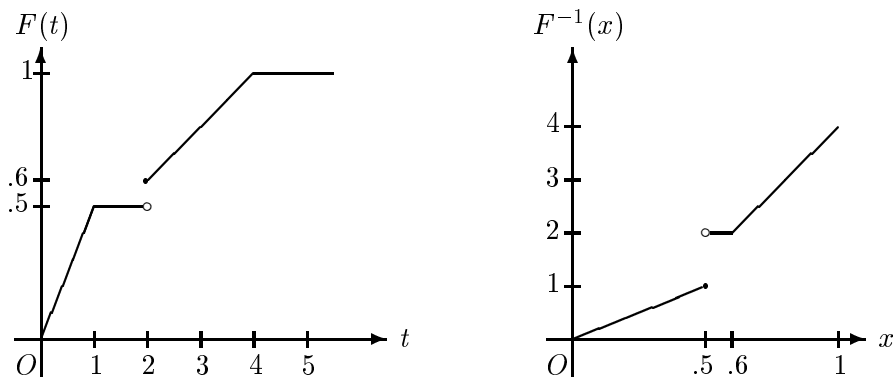


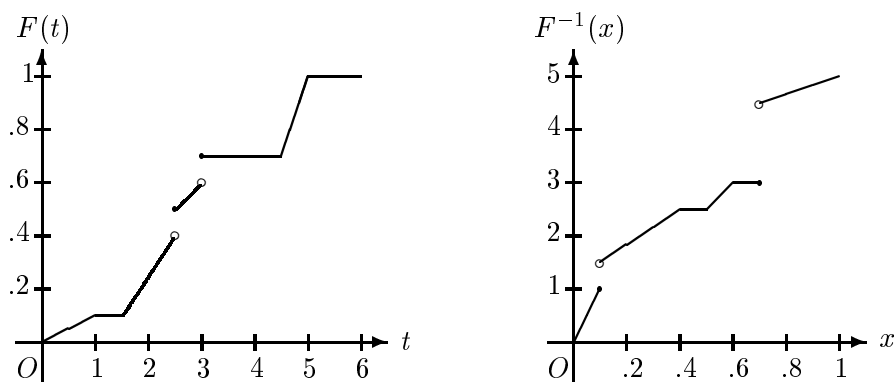
圖 5.2 例 5.9 中 F 及 F^{-1} 之圖形

由例 5.9 可看出, 當 F 之圖形在某區間為嚴格漸增, 則對應的 F^{-1} 亦為嚴格漸增; 當 F 之圖形在某區間為水平, 則對應的 F^{-1} 有跳升; 而當 F 有跳升時, 則對應一段水平的 F^{-1} 。讀者可想想, F^{-1} 之圖形, 與一般的反函數圖形有何差別。

如果隨機變數 X 之支撐包含負值, 則 F^{-1} 便亦包含負值 (見習題第 18 題)。底下圖 5.3 之左圖為一有較多段水平及跳升之 F 的圖形。讀者可練習看出其 F^{-1} 的圖形, 便如圖 5.3 中之右圖。

給一分佈函數 F , 利用其分位函數 F^{-1} 及一有 $\mathcal{U}(0, 1)$ 分佈的隨機變數, 可得一以 F 為分佈函數之隨機變數, 此過程即所謂 **分位轉換** (quantile transformation)。

定理 5.4 設 F 為一分佈函數, 且隨機變數 U 有 $\mathcal{U}(0, 1)$ 分佈。則 $X = F^{-1}(U)$ 以 F 為分佈函數。

圖5.3 F 及 F^{-1} 之圖形

證明. 首先 $F^{-1}(U)$ 仍為一隨機變數。利用定理5.3之(i), 對 $\forall t \in R$,

$$P(X \leq t) = P(F^{-1}(U) \leq t) = P(U \leq F(t)) = F(t)。$$

此處用到 $0 \leq F(t) \leq 1$, 而 U 有 $\mathcal{U}(0, 1)$ 分佈, 故最後一等式成立。證畢。

上述定理為許多模擬步驟之基礎。計算機裡有產生**隨機數**(random number)的功能, 可製造出獨立的有 $\mathcal{U}(0, 1)$ 分佈之數列。因此, 若將 F^{-1} 在前述所產生的數列取值, 便可得到以 F 為分佈函數之隨機數列。當然, 欲此法可行, 就須先能得 F^{-1} 之**明確的形式**(closed form), 或數值上的近似值。

當 F 為連續時(不一定有 p.d.f. 存在), 定理5.4有下述逆定理。

定理5.5 設隨機變數 X 之分佈函數 F 為連續, 則 $Y = F(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

證明. 對 $\forall x \in (0, 1)$,

$$\begin{aligned} P(Y \geq x) &= P(F(X) \geq x) = P(F^{-1}(x) \leq X) \\ &= 1 - F(F^{-1}(x)) = 1 - x, \end{aligned}$$

其中第一個等式用到定理5.3之(i), 第二個等式是利用假設 F 為連續(若 F 不一定連續, 則 $P(F(X) \geq x) = 1 - F(F^{-1}(x)-)$), 而最後一等式是用到定理5.3之(iii)。故得對 $\forall x \in (0, 1)$,

$$\begin{aligned}
P(Y \leq x) &= P(F(X) \leq x) = P(F(X) < x) \\
&= 1 - P(F(X) \geq x) = 1 - (1 - x) = x,
\end{aligned}$$

又顯然 $P(F(X) \leq x) = 0$, 當 $x \leq 0$; $P(F(X) \leq x) = 1$, 當 $x \geq 1$ 。故得證 $F(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

底下給二例。

例5.10 設隨機變數 X 之分佈函數為 $F(x) = x^3$, $0 < x < 1$ 。令 $Y = F(X) = X^3$ 。底下我們驗證 Y 有 $\mathcal{U}(0, 1)$ 分佈。

$$\begin{aligned}
P(Y \leq y) &= P(X^3 \leq y) \\
&= P(0 < X \leq y^{1/3}) \\
&= F(y^{1/3}) \\
&= (y^{1/3})^3 = y, 0 < y < 1。
\end{aligned}$$

故 Y 確實有 $\mathcal{U}(0, 1)$ 分佈。

例5.11 設隨機變數 X 有標準的羅吉斯分佈, 即分佈函數

$$F(x) = \frac{1}{1 + e^{-x}}, x \in \mathbb{R}。$$

底下驗證 $Y = F(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

$$\begin{aligned}
P(Y \leq y) &= P\left(\frac{1}{1 + e^{-X}} \leq y\right) \\
&= P(1 + e^{-X} \geq y^{-1}) \\
&= P(-X \geq \log(y^{-1} - 1)) \\
&= P(X \leq -\log(y^{-1} - 1)) \\
&= F(-\log(y^{-1} - 1)) \\
&= \frac{1}{1 + e^{-(-\log(y^{-1} - 1))}} \\
&= \frac{1}{1 + (y^{-1} - 1)} \\
&= y, 0 < y < 1。
\end{aligned}$$

得證。

1.6 期望值

對一隨機變數，我們常想粗略地知道其值究竟多大，**期望值**(expectation, 或稱expected value, mean)，就是常被拿來扮演這種以一單一的值，來代表一隨機現象中之變數大小的角色。

設隨機變數 X 之p.d.f.為 f ，則其期望值以 $E(X)$ 表示，定義為

$$(6.1) \quad E(X) = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx & , \text{連續型}, \\ \sum_x xf(x) & , \text{離散型}, \end{cases}$$

只要上述積分或和存在。此處 $\sum_x xf(x)$ 表對所有可能的 x (即 $f(x) \neq 0$)相加。通常以 μ_X (或只以 μ)表 $E(X)$ 。

假設有10個數: 1,1,2,2,2,3,3,3,4,4, 則其平均為

$$\frac{1}{10}(1+1+2+2+2+3+3+3+4+4) = 2.5。$$

由於數字中有重複: 有2個1, 3個2, 3個3, 2個4, 所以平均亦可表為

$$\frac{1}{10}(1 \cdot 2 + 2 \cdot 3 + 3 \cdot 3 + 4 \cdot 2) = 2.5,$$

所得相同。第二種作法是我們常採用的，特別是在數字較大時。又上式左側可改寫為

$$1 \cdot \frac{2}{10} + 2 \cdot \frac{3}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{2}{10},$$

其中 $2/10$, $3/10$, $3/10$, $2/10$, 可分別視為10個數字中, 1,2,3,4所佔的份量。所以第二種作法便是一種加權平均。隨機變數期望值的定義, 便是基於加權平均的想法。這點由離散型隨機變數期望值的定義, 特別可以看出來。

至於 X 之一函數 $g(X)$ 的期望值, 則定義為

$$(6.2) \quad E(g(X)) = \begin{cases} \int_{-\infty}^{\infty} g(x)f(x)dx, & \text{連續型,} \\ \sum_x g(x)f(x) & \text{離散型。} \end{cases}$$

亦可經由變數代換, 先求出 $Y = g(X)$ 之 p.d.f., 再利用 (6.1) 式而求得 $E(Y)$ 。

底下給幾個例子。

例6.1 設離散型隨機變數 X 之 p.d.f. 為

$$\begin{aligned} f(0) &= \frac{1}{16}, & f(1) &= \frac{2}{16}, & f(2) &= \frac{3}{16}, & f(3) &= \frac{4}{16}, \\ f(4) &= \frac{4}{16}, & f(5) &= \frac{1}{16}, & f(6) &= \frac{1}{16}. \end{aligned}$$

則

$$\begin{aligned} E(X) &= \sum_{x=0}^6 xf(x) \\ &= 0 \cdot \frac{1}{16} + 1 \cdot \frac{2}{16} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{4}{16} + 5 \cdot \frac{1}{16} + 6 \cdot \frac{1}{16} \\ &= \frac{47}{16}. \end{aligned}$$

又

$$\begin{aligned} E((X-3)^2) &= \sum_{x=0}^6 (x-3)^2 f(x) \\ &= 9 \cdot \frac{1}{16} + 4 \cdot \frac{2}{16} + 1 \cdot \frac{3}{16} + 0 \cdot \frac{4}{16} + 1 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16} + 9 \cdot \frac{1}{16} \\ &= \frac{37}{16}. \end{aligned}$$

在求 $E((X-3)^2)$ 時, 也可利用變數代換。令 $Y = (X-3)^2$ 。則 Y 取值 0, 1, 4, 9, 且

$$\begin{aligned} f_Y(0) &= P(X=3) = \frac{4}{16}, \\ f_Y(1) &= P(X=2) + P(X=4) = \frac{7}{16}, \\ f_Y(4) &= P(X=1) + P(X=5) = \frac{3}{16}, \\ f_Y(9) &= P(X=0) + P(X=6) = \frac{2}{16}. \end{aligned}$$

故

$$\begin{aligned} E((X-3)^2) &= E(Y) = \sum_y y f_Y(y) \\ &= 0 \cdot \frac{4}{16} + 1 \cdot \frac{7}{16} + 4 \cdot \frac{3}{16} + 9 \cdot \frac{2}{16} \\ &= \frac{37}{16}。 \end{aligned}$$

其中 $\sum$ 是對所有的 y 相加，即 $0, 1, 4, 9$ 。

例6.2 設隨機變數 X 之p.d.f.為

$$f(x) = 2x, x \in [0, 1]。$$

則

$$\begin{aligned} E(X) &= \int_0^1 x \cdot 2x dx = \int_0^1 2x^2 dx = \frac{2}{3}, \\ E(X^2) &= \int_0^1 x^2 \cdot 2x dx = \int_0^1 2x^3 dx = \frac{1}{2}。 \end{aligned}$$

另外，亦可利用變數代換。令 $Y = X^2$ ，先求出 Y 之p.d.f.為

$$g(y) = 2\sqrt{y} \cdot \frac{1}{2\sqrt{y}} = 1, y \in [0, 1]。$$

如此

$$E(X^2) = E(Y) = \int_0^1 y \cdot 1 dy = \frac{1}{2},$$

仍得到相同的結果。

例6.3 在微積分裡有下述結果：

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}。$$

故

$$(6.3) \quad f(i) = \frac{6}{\pi^2} \frac{1}{i^2}, i = 1, 2, \dots,$$

為一p.d.f.。但因

$$\sum_{i=1}^{\infty} \frac{1}{i} = \infty,$$

故若隨機變數 X 之p.d.f.如(6.3)式, 則 X 之期望值不存在。

另外, 因

$$\sum_{i=1}^{\infty} \frac{1}{i^4} = \frac{\pi^4}{90},$$

故

$$(6.4) \quad g(i) = \frac{90}{\pi^4} \frac{1}{i^4}, i = 1, 2, \dots,$$

為一p.d.f.。又設隨機變數 Y 之p.d.f.如(6.4)式所給。則

$$E(Y) = \frac{90}{\pi^4} \sum_{i=1}^{\infty} \frac{1}{i^3} < \infty,$$

$$E(Y^2) = \frac{90}{\pi^4} \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{90}{\pi^4} \frac{\pi^2}{6} = \frac{15}{\pi^2} < \infty。$$

即 $E(Y)$ 及 $E(Y^2)$ 皆存在。當然 $E(Y^3)$ 便不存在了。

例6.4 設隨機變數 X 有 $\mathcal{C}(0, 1)$ 分佈, 即其p.d.f.為

$$f(x) = \frac{1}{\pi(1+x^2)}, x \in R,$$

則 X 之期望值不存在。此因

$$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi(1+x^2)} dx,$$

為一瑕積分(improper integral)。而依定義, 此瑕積分存在, 若且唯若

$$\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx, \quad \int_{-\infty}^0 \frac{x}{\pi(1+x^2)} dx$$

皆要存在。但此二積分皆不存在:

$$\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx = \lim_{x \rightarrow \infty} \frac{1}{2\pi} \log(1+x^2) = \infty,$$

故 $E(X)$ 不存在。

要注意的是, 不可以為因 $x/(\pi(1+x^2))$ 為一奇函數, 所以在 $(0, \infty)$ 的積分與在 $(-\infty, 0)$ 的積分, 剛好正負消掉, 而得積分為0。要知道並沒有 $\infty - \infty = 0$ 的運算。 ∞ 並非一實數, 實數裡的一些運算對它並不一定適用。不過當一隨機變數 X 之p.d.f.為偶函數, 則 $E(X)$ 若存在必為0, 證明留在習題中。因此由於標準的羅吉斯分佈之p.d.f. f_X 為偶函數(見例4.10), 且

$$\begin{aligned}\int_0^\infty x f_X(x) dx &= \int_0^\infty x \cdot \frac{e^{-x}}{(1+e^{-x})^2} dx \leq \int_0^\infty x e^{-x} dx \\ &= - \int_0^\infty x d e^{-x} = -x e^{-x} \Big|_0^\infty + \int_0^\infty e^{-x} dx = 1 < \infty,\end{aligned}$$

故標準羅吉斯分佈之期望值存在且為0。

附帶一提, 由上述瑕積分存在的定義, 可看出 $E(X)$ 存在, 若且唯若 $E(|X|)$ 存在(見習題第29題), 有些書便以 $E(|X|)$ 存在, 當做期望值存在的定義。

期望值有不少性質。例如, 一**有界的**隨機變數, 期望值必存在。在此對一隨機變數 X , 若存在一常數 $M > 0$, 使得 $P(|X| \leq M) = 1$, 便稱該隨機變數為**有界的**。若 X 有 $\mathcal{U}(0, 1)$ 分佈, 便為有界, 有 $\mathcal{C}(0, 1)$ 分佈, 則非有界。我們再列出一些性質如下。

定理6.1 設 X 為一隨機變數, a, b, c 為常數, g_1, g_2 為二函數, 使得 $E(g_1(X)), E(g_2(X))$ 皆存在。則

- (i) $E(ag_1(X) + bg_2(X) + c) = aE(g_1(X)) + bE(g_2(X)) + c$;
- (ii) 若 $g_1(x) \geq 0, \forall x \in R$, 則 $E(g_1(X)) \geq 0$;
- (iii) 若 $g_1(x) \geq g_2(x), \forall x \in R$, 則 $E(g_1(X)) \geq E(g_2(X))$;
- (iv) 若 $a \leq g_1(x) \leq b, \forall x \in R$, 則 $a \leq E(g_1(X)) \leq b$ 。

由上定理之(i), 特別地, 只要 $E(X)$ 存在, 則

$$(6.5) \quad E(aX + b) = aE(X) + b,$$

$$(6.6) \quad E(b) = b。$$

即一常數之期望值等於該常數。由於 $E(X)$ 為一常數(已非隨機了), 故其期望值仍為本身。即

$$E(E(X)) = E(X)。$$

期望值這個名詞也許會令有些人感到困擾。投擲一公正的骰子一次, 令 X 表所得之點數。則 X 之p.d.f.為

$$f(x) = \frac{1}{6}, x = 1, 2, \dots, 6。$$

因此

$$E(X) = \sum_{i=1}^6 i \cdot \frac{1}{6} = \frac{21}{6} = 3.5。$$

投擲一骰子, 點數之期望值為3.5。點數為整數, 無論怎麼投擲都得不到3.5。那怎會期望得到3.5呢? 唐詩三百首裡, 元稹遣悲懷:

同穴窆冥何所望, 他生緣會更難期。

3.5的期望值, 可說較他生緣更難期了。

所以暫時不要去想期望值字面上的含義, 是不易想通的。不過期望值是一隨機變數之“最佳”常數的代表值。所謂最佳, 指的是誤差平方的期望值最小。見下定理。

定理6.2 設隨機變數 X 滿足 $E(X^2) < \infty$ 。則對所有 $a \in R$, $E((X - a)^2)$ 之極小值發生在 $a = E(X)$ 。

證明.由

$$E((X - a)^2) = E(X^2) - 2aE(X) + a^2,$$

將 $E(X^2) - 2aE(X) + a^2$ 視為 $-a$ 的二次式, 經由微分可得極小值發生在 $a = E(X)$ 。得證。

在3.8節我們會看到若 $E(X^2)$ 存在, 則導致 $E(X)$ 存在。

想以一常數 a 來代表一隨機變數 X , 很難說那一個 a 最好。誤差 $X - a$ 仍是隨機的。但在誤差的平方 $(X - a)^2$ 之期望值最小下, 上定理指出期望值 $E(X)$ 便是最好的選擇。

其次我們定義**動差**(moment)。設 n 為一整數，則若 $E(X^n)$ 存在，便稱此為隨機變數 X 之 **n 次動差**。而 $E((X - \mu)^n)$ 稱為 X 之 **n 次中心動差**(n th central moment)，其中 $\mu = E(X)$ 。特別地，二次中心動差稱為 X 之**變異數**(variance)，以 σ_X^2 , $\text{Var}(X)$ 或簡單地以 σ^2 表之，即

$$(6.7) \quad \text{Var}(X) = E((X - E(X))^2)。$$

$\text{Var}(X)$ 之正的平方根，以 σ_X 或 σ 表之，稱為 X 之**標準離差**(standard deviation)，或**標準差**、**均方差**。另外，如同期望值，對於 n 次動差仍有 $E(X^n)$ 存在，若且唯若 $E(|X|^n)$ 存在，其證明留在習題中。

如同期望值，標準離差此一名詞可能令人感到迷惑。這並不表示隨機變數 X 有較大的機率與期望值之差距不超過一個標準離差。常有人會有這樣的誤解。如見目前通行的某版本高中數學甲(上)(即自然組高三上學期)。我們給一例子如下，讀者不難造出更多的例子。

設 $P(X = 0) = 0.1$, $P(X = 2) = P(X = -2) = 0.45$ 。則

$$\mu = E(X) = 0,$$

$$\sigma^2 = \text{Var}(X) = E(X^2) = 0.1 \cdot 0^2 + 0.45 \cdot 2^2 + 0.45 \cdot (-2)^2 = 3.6。$$

因此 X 之標準離差 $\sigma = \sqrt{3.6} \doteq 1.8972 < 2$ 。但有0.9之機率， X 取值與 $\mu = 0$ 之差距超過 σ 。

在機率論裡我們引進了隨機變數的概念。例如，學生不再如小學的練習裡，每位體重40公斤，求25位學生的總體重，而是讓每位學生的體重有一機率分佈。但因覺得無法掌握隨機的量之大小，我們才想要有一代表值。由於有很多好的性質，期望值遂常被拿來當做隨機變數之代表值。期望值像是隨機變數分佈之一個核心，隨機變數的可能值，散佈在期望值的左右。其他亦常被拿來當做隨機變數之代表值的尚有**中位數**(median)及**眾數**(mode)等。眾數早期稱為**密集數**。在此一實數 m 若滿足

$$(6.8) \quad P(X \leq m) \geq \frac{1}{2}, \quad \text{且} \quad P(X \geq m) \geq \frac{1}{2},$$

便稱為隨機變數 X (或其分佈)之一**中位數**。依此定義中位數不一定唯一。當然如果 X 之分佈函數為連續且嚴格漸增，則中位數便唯一。此時

中位數為滿足 $F(x) = 0.5$ 之 x , 即 X 之第0.5分位數。有些書便將中位數定義為 X 之第0.5分位數。依其定義中位數便唯一存在了。若 X 之分佈對稱於 m , 則 X 之中位數為 m (習題第16題)。另外, 若 $f(a)$ 為 X 之機率密度函數的一絕對極大值, 便稱 a 為 X (或其分佈)之一眾數。對 X 為離散型的隨機變數, 其眾數表 X 最可能取的值。眾數也不一定唯一。中位數及眾數有時雖存在但卻不易求出(見習題第16題)。

例6.5 設 X 為一離散型的隨機變數, 且 $P(X = 0) = 1/4$, $P(X = 1) = 1/2$, $P(X = 2) = 1/4$ 。則 X 有唯一的中位數1:

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) = \frac{3}{4} \geq \frac{1}{2}, \\ P(X \geq 1) &= P(X = 1) + P(X = 2) = \frac{3}{4} \geq \frac{1}{2}. \end{aligned}$$

故1為一中位數。又任一 $a < 1$,

$$P(X \leq a) \leq P(X = 0) = \frac{1}{4} < \frac{1}{2},$$

任一 $a > 1$,

$$P(X \geq a) \leq P(X = 2) = \frac{1}{4} < \frac{1}{2},$$

皆非中位數。

至於 X 之眾數亦為1。

例6.6 設 X 為一離散型的隨機變數, 且 $P(X = 0) = 0.1$, $P(X = 1) = 0.4$, $P(X = 2) = 0.4$, $P(X = 3) = 0.1$ 。則任一 $a \in [1, 2]$ 皆為中位數, 而1,2皆為眾數。

例6.7 設隨機變數 X 之分佈函數為

$$F(x) = \begin{cases} 0 & , x < 0, \\ x^2 & , 0 \leq x < 1, \\ 1 & , x \geq 1. \end{cases}$$

令 $x^2 = 1/2$, 解出 $x = \sqrt{2}/2 \doteq 0.707$ 。即 X 之中位數為 $\sqrt{2}/2$ 。也可以說 F 之中位數為 $\sqrt{2}/2$ 。若取 X 之 p.d.f. 為 $f(x) = 2x, 0 \leq x \leq 1$, 則 X 之眾數為 1。若取 X 之 p.d.f. 為 $f(x) = 2x, 0 \leq x < 1$, 則眾數不存在。

我們常要藉助統計做各種決策。投資某項事業, 可能成功而獲利若干, 可能失敗而損失若干。那到底要不要投資呢? 不能只看到成功機率很大(說不定獲利很小), 也不能只看到獲利很大(說不定成功機率很低), 算一算所得(若損失所得便為負)之期望值, 是一簡單的決策依據。我們平常在判斷出門要不要帶雨傘, 也是先盤算一下期望損失(先估計下雨的機率, 下雨不帶傘的損失等)。期望值在做決策時, 常扮演重要的角色。

例6.8 某賭局的賭法如下: 投擲一公正的銅板, 若第 n 次投擲才首度出現正面, 則賭徒可獲得 1.95^n 元。試依期望所得, 決定賭徒每次賭之前, 該付多少元, 此賭局才是公平。

解. 不難看出第 n 次投擲才首度出現正面之機率為 $2^{-n}, n \geq 1$ 。令 X 表賭徒賭一次之所得, 則

$$\begin{aligned} E(X) &= \sum_{n=1}^{\infty} 1.95^n \cdot 2^{-n} = \sum_{n=1}^{\infty} \left(\frac{1.95}{2}\right)^n \\ &= \frac{1.95/2}{1 - 1.95/2} = 39。 \end{aligned}$$

故每次該付 39 元。

不論期望值是多好的一個隨機變數之代表值, 難免會偏離該隨機變數, 如何量測偏差究竟有多大呢? 這就是變異數與標準差的功能。

設有一隨機變數 X , 且設期望值 $E(X)$ 存在。則 X 與 $E(X)$ 之離差為 $X - E(X)$ 。此值有正有負, 但期望值

$$E(X - E(X)) = E(X) - E(X) = 0。$$

離差之期望值為 0, 此亦為期望值之一性質。但我們往往是對離差之絕對值較感興趣。此正如射飛鏢時, 有時上偏有時下偏, 總不能得意地說平均命中紅心。

以 $|X - E(X)|$ 表離差之絕對值，此量仍為一隨機變數，如何衡量其大小呢？取期望值為一辦法。也就是求量測離差之絕對值的“平均”。但如果你微積分學得還不錯的話，應知道有關絕對值函數之求和或積分，往往有些麻煩（想想求 $\int_{-3}^3 |\sin x - 0.3x| dx$ 的例子）。對於 $E(|X - E(X)|)$ ，要討論何時 $X \geq E(X)$ ，何時 $X < E(X)$ 。因此通常的作法，是考慮離差平方之期望值，而這就是**變異數**。由於是平方的期望值，像是距離的平方。因此若將變異數開平方，就像使其回到距離，**標準差**便產生了。

在統計裡是允許誤差的存在。我們不要求百分之百的精準，我們只希望誤差在可接受的範圍，而這誤差其實也是隨機的。因此我們才考慮“平均式”（取期望值）的誤差。大家常說，雖不中亦不遠矣。但不遠到底是多遠？標準差的功能，就是對誤差之大小，提供一參考值。如果說期望值是描述一隨機變數之分佈的“集中處”，標準差便是描述分佈對該集中處之偏離程度。標準差愈小（或等價地說變異數愈小），表隨機變數之分佈較集中在期望值附近；標準差愈大，表隨機變數之分佈，散佈的較開。人們常說命運，命就有點像是期望值，運則像是變異數。命也許不太好（期望值不大），但由於有變異的關係，使得有時好運會發生。

假設兩組學生參加測驗，每組各有三人。平均成績皆為60分，但第一組之成績為61, 60, 59；第二組之成績為90, 60, 30。第一組學生的程度很接近，而第二組學生的程度則有很大差異。光由平均成績，並無法了解兩組學生程度差異之情況。這也是對隨機變數，為何除了期望值，我們尚須給出其標準差。給了此二值，對其“中心”及“散佈範圍”，才有一較清晰的概念。

例6.9 假設有 A, B 二項投資。投資 A 於一年後，有 $1/2$ 的機率賺一倍的錢，有 $1/4$ 的機率賠一半的錢，有 $1/4$ 的機率不賺不賠。投資 B 一年後，有 $3/4$ 的機率賺一半的錢，有 $1/4$ 的機率不賺不賠。若放10,000元於投資 A ，則一年後之獲利 X ，為一離散型的隨機變數，p.d.f.為（略去“元”，底下同）

$$f_X(x) = \begin{cases} 1/2, & x = 10,000, \\ 1/4, & x = -5,000, \\ 1/4, & x = 0. \end{cases}$$

若放10,000於投資 B , 則一年後之獲利 Y , 亦為一離散型的隨機變數, p.d.f.為

$$f_Y(y) = \begin{cases} 3/4, & y = 5,000, \\ 1/4, & y = 0. \end{cases}$$

依上述條件, 可得

$$E(X) = \sum_x x f_X(x) = \frac{1}{2} \cdot 10,000 + \frac{1}{4} \cdot (-5,000) + \frac{1}{4} \cdot 0 = 3,750.$$

$$E(Y) = \sum_y y f_Y(y) = \frac{3}{4} \cdot 5,000 + \frac{1}{4} \cdot 0 = 3,750.$$

即投資 A 與投資 B 獲利之期望值相同。雖然有相同的期望值, 但 X 之分佈較 Y 之分佈, 散佈的較開, 所以我們猜測 X 之變異數較大。底下來驗證此猜測。

$$\begin{aligned} \text{Var}(X) &= \sum_x (x - 3,750)^2 f_X(x) \\ &= \frac{1}{2}(10,000 - 3,750)^2 + \frac{1}{4} \cdot (-5,000 - 3,750)^2 + \frac{1}{4}(0 - 3,750)^2 \\ &= 42,187,500. \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= \sum_y (y - 3,750)^2 f_Y(y) \\ &= \frac{3}{4}(5,000 - 3,750)^2 + \frac{1}{4}(0 - 3,750)^2 \\ &= 4,687,500. \end{aligned}$$

由於 $\text{Var}(X) > \text{Var}(Y)$, 表示投資 A 之變異較大。若喜愛冒險者, 說不定會採投資 A , 有時會有較高獲利。個性保守者, 可能會採投資 B , 穩紮穩打。

例6.10 某項食品是以機械填裝，每包標示16公斤。若容量不足，顧客可能會抱怨，若容量過多，公司也有損失。令 X 表每包之容量，假設 X 有p.d.f.

$$f(x) = 2.5, \quad 15.8 \leq x \leq 16.2,$$

即 X 有 $\mathcal{U}[15.8, 16.2]$ 分佈。則

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{15.8}^{16.2} 2.5x dx = 1.25x^2 \Big|_{15.8}^{16.2} = 16.0。$$

所以平均而言，機器填裝的容量是正確的。但

$$P(X < 16) = P(X > 16) = \frac{1}{2},$$

即各有一半的機率容量不足及容量超過。我們來求其變異數。

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} (x-16)^2 f(x) dx \\ &= \int_{15.8}^{16.2} 2.5(x-16)^2 dx = \frac{2.5}{3} (x-16)^3 \Big|_{15.8}^{16.2} = 0.0133。 \end{aligned}$$

標準差為 $\sqrt{\text{Var}(X)} \doteq 0.115$ 。

爲了降低變異，公司引進新機器，新的容量 Y 之p.d.f.爲

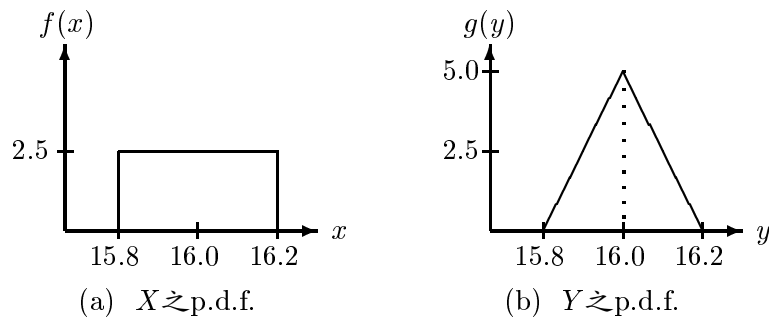
$$g(y) = \begin{cases} 25(x-16) + 5, & 15.8 \leq x < 16, \\ 25(16-x) + 5, & 16 \leq x < 16.2。 \end{cases}$$

則仍有(習題第1題), $E(Y) = 16$, 但 $\text{Var}(Y) = 1/150 \doteq 0.00\overline{6}$, 標準差 $\doteq 0.0816$ 。變異數下降一半以上，標準差則下降約29%。

圖6.1給出 X 及 Y p.d.f.之圖形。由圖形亦可看出 Y 之變異數較 X 之變異數小：較多的機率配置在 Y 之期望值的近傍。但在可能取的值爲區間 $[15.8, 16.2]$ ，且期望值爲16之隨機變數中，上述 X 倒並非變異數最大的隨機變數，此點做了習題第33題便可了解。

我們先給變異數之一些性質。

定理6.3 設 X 爲一隨機變數，則下述二式成立，其中 a, b 爲常數。

圖6.1 例6.10中 X 及 Y p.d.f.之圖形

$$(6.9) \quad \text{Var}(aX + b) = a^2 \text{Var}(X),$$

$$(6.10) \quad \text{Var}(X) = E(X^2) - (E(X))^2.$$

證明. 首先因 $E(aX + b) = aE(X) + b$, 由定義得

$$\begin{aligned} \text{Var}(aX + b) &= E((aX + b - (aE(X) + b))^2) \\ &= E((aX - aE(X))^2) \\ &= a^2 E((X - E(X))^2) \\ &= a^2 \text{Var}(X), \end{aligned}$$

其中第三個等號成立, 仍藉助(6.5)式。得證(6.9)式。

其次

$$\begin{aligned} \text{Var}(X) &= E((X - E(X))^2) \\ &= E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2(E(X))^2 + (E(X))^2 \\ &= E(X^2) - (E(X))^2, \end{aligned}$$

此處用到(6.6)式($E(X)$ 及 $(E(X))^2$ 皆為常數)。得證(6.10)式。

由於變異數一定不是負的(非負隨機變數 $(X - E(X))^2$ 之期望值必為非負, 見定理6.1之(ii)), 故由(6.10)式得

$$(6.11) \quad E(X^2) \geq (E(X))^2.$$

又由(6.9)式可看出(取 $a = 0$)常數的變異數為0。當然這點由(6.10)式亦可看出。即若存在一常數 b , 使得 $P(X = b) = 1$, 則因 $E(X^2) = b^2$, $E(X) = b$, 故 $\text{Var}(X) = 0$ 。這當然是合理的。既然是常數(可視為一退化的隨機變數), 就沒有變異, 每次觀測皆得到該常數。反之, 若 $\text{Var}(X) = 0$, 則存在常數 b , 使得 $P(X = b) = 1$ 。此部分的證明我們留在習題中。

例6.11 投擲一公正的骰子, 令 X 表所得之點數。試求 X 之變異數及標準差。

解.前面已求出 $E(X) = 3.5 = 7/2$ 。又

$$E(X^2) = \sum_{i=1}^6 \frac{1}{6} \cdot i^2 = \frac{91}{6}。$$

故

$$\begin{aligned}\text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \frac{91}{6} - \frac{49}{4} = \frac{35}{12},\end{aligned}$$

且 X 之標準差為 $\sqrt{35/12} \doteq 1.7078$ 。

例6.12 對一事件 A , 如例3.5, 令 I_A 表 A 之指示函數, I_A 為一只取0,1兩個值之隨機變數。且 $I_A = 1$, 若 A 發生, $I_A = 0$, 若 A 不發生。故

$$\begin{aligned}E(I_A) &= 1 \cdot P(I_A = 1) + 0 \cdot P(I_A = 0) = P(A)。 \\ \text{Var}(I_A) &= E(I_A^2) - (E(I_A))^2 \\ &= E(I_A) - (E(I_A))^2 \\ &= P(A) - (P(A))^2。 \end{aligned}$$

例6.13 承例6.2。由於 $E(X) = 2/3$, 依定義得

$$\begin{aligned}\text{Var}(X) &= E((X - E(X))^2) \\ &= \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x dx \\ &= \int_0^1 \left(2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x\right) dx = \frac{1}{18}。 \end{aligned}$$

另外，因已求出 $E(X^2) = 1/2$ ，利用(6.10)式得

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18},$$

所得相同。

例6.9及6.10中之變異數，亦皆可以(6.10)式求之，所得必相同。我們再給一例。

例6.14 設隨機變數 X 之 p.d.f. 為

$$f(x) = \frac{1}{x \log 2}, 1 \leq x \leq 2.$$

則

$$\begin{aligned} E(X) &= \int_1^2 x \cdot \frac{1}{x \log 2} dx = \frac{1}{\log 2} \int_1^2 1 dx = \frac{1}{\log 2} \doteq 1.443. \\ E(X^2) &= \int_1^2 x^2 \cdot \frac{1}{x \log 2} dx = \frac{1}{\log 2} \int_1^2 x dx = \frac{3}{2 \log 2}. \end{aligned}$$

故

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \frac{3}{2 \log 2} - \frac{1}{(\log 2)^2} \doteq 0.08267. \end{aligned}$$

標準差為 $\sqrt{\text{Var}(X)} \doteq 0.2875$ 。

附帶一提，因 X 之分佈函數

$$F(x) = \int_1^x f(t) dt = \frac{\log x}{\log 2}, 1 \leq x \leq 2,$$

解 $F(x) = 1/2$ ，得中位數為 $\sqrt{2}$ 。至於眾數則為 1。

對一隨機變數，於給出一次及二次動差(變異數是基於二次動差)後，對其分佈有了基本的了解。至於若繼續知道更多次的動差，對其分佈就愈了解。所謂分佈，顧名思義，就是一隨機變數的值之散佈情況。由於散佈著，我們才又想以一些基本的量來描述散佈的情況。這就是動差的功能。

另外, 亦可以一些**轉換**, 來描述隨機變數的分佈。我們介紹幾個常見且可以唯一決定隨機變數之分佈的轉換。設 X 為一隨機變數。

若 X 只取非負整數值, 且 $P(X = k) = a_k, k = 0, 1, \dots$, 我們常考慮其**母函數**(generating function, 或稱**生成函數**)

$$(6.12) \quad g(s) = E(s^X) = \sum_{k=0}^{\infty} a_k s^k, |s| \leq 1。$$

可看出 $g(1) = 1$ 。又

$$(6.13) \quad a_k = \frac{1}{k!} \frac{d^k g(s)}{ds^k}, k = 0, 1, \dots。$$

若 X 取值在一可數的集合 $\{\lambda_0, \lambda_1, \dots\}$ 。亦可類似地定義其母函數。

若 X 為非負, 則往往考慮其**拉普拉斯轉換**(Laplace transform)

$$(6.14) \quad \phi(s) = E(e^{-sX}), s \geq 0。$$

顯然 $\phi(0) = 1$ 。若 X 為連續型, 有p.d.f.f, 則

$$(6.15) \quad \phi(s) = \int_0^{\infty} e^{-sx} f(x) dx, s \geq 0。$$

若 X 為離散型, 取值在 $\{\lambda_0, \lambda_1, \dots\}$, 且 $P(X = \lambda_k) = p_k$, 則

$$(6.16) \quad \phi(s) = \sum_{k=0}^{\infty} p_k e^{-s\lambda_k}, s \geq 0。$$

特徵函數(characteristic function)則是對所有隨機變數均適用, 且必存在:

$$(6.17) \quad \psi(t) = E(e^{itX}), t \in R, i = \sqrt{-1}。$$

仍有 $\psi(0) = 1$ 。

上述這些轉換都具有一以貫之的功能, 能以一扼要的方式, 來包含關於隨機變數 X 分佈之所有資訊。由這些轉換, 也可得到 X 之各次動差。例如, 設 $\phi(s)$ 為 X 之拉普拉斯轉換。則只要 X 的 n 次動差存在,

$$(6.18) \quad E(X^n) = (-1)^n \left. \frac{d^n \phi(s)}{ds^n} \right|_{s=0} = (-1)^n \phi^{(n)}(0)。$$

又這些轉換只要存在, 彼此便有一簡單的關係。例如,

$$(6.19) \quad \phi(s) = g(e^{-s}), s \geq 0,$$

$$(6.20) \quad g(s) = \phi(-\log s), 0 < s < 1。$$

最後

$$(6.21) \quad M(t) = E(e^{tX})$$

稱為 X 之**動差母函數**(moment generating function)。不過動差母函數不一定存在, 有時只在 t 屬於某一範圍才存在, 但 $M(0) = 1$ 必存在。

動差母函數若存在, 亦可經由微分得到動差:

$$(6.22) \quad E(X^n) = \left. \frac{d^n M(t)}{dt^n} \right|_{t=0} = M^{(n)}(0)。$$

例6.15 設隨機變數 X 有 $\mathcal{E}(\lambda)$ 分佈, $\lambda > 0$ 。 X 之p.d.f.為 $f(x) = \lambda e^{-\lambda x}$, $x > 0$ 。則 X 之拉普拉斯轉換及動差母函數分別為

$$\begin{aligned} \phi(s) &= E(e^{-sX}) = \int_0^\infty e^{-sx} f(x) dx = \int_0^\infty e^{-sx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda+s)x} dx = -\frac{\lambda}{\lambda+s} e^{-(\lambda+s)x} \Big|_0^\infty \\ &= \frac{\lambda}{\lambda+s}, s \geq 0。 \end{aligned}$$

$$\begin{aligned} M(t) &= E(e^{tX}) = \int_0^\infty e^{tx} f(x) dx = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda-t)x} dx = -\frac{\lambda}{\lambda-t} e^{-(\lambda-t)x} \Big|_0^\infty \\ &= \begin{cases} \lambda/(\lambda-t), & t < \lambda, \\ \infty & , t \geq \lambda。 \end{cases} \end{aligned}$$

即 $s \geq 0$ 時, $\phi(s)$ 皆存在, 但 $M(t)$ 在 $t \geq \lambda$ 時則不存在。

因

$$\begin{aligned}\phi'(s) &= -\frac{\lambda}{(\lambda+s)^2}, \\ \phi''(s) &= \frac{2\lambda}{(\lambda+s)^3} \circ\end{aligned}$$

故

$$\begin{aligned}E(X) &= -\phi'(0) = \frac{1}{\lambda}, \\ E(X^2) &= \phi''(0) = \frac{2}{\lambda^2} \circ\end{aligned}$$

由此又得

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \circ$$

另外，若利用 $M(t)$ 亦可得到 $E(X)$ 及 $E(X^2)$:

$$M'(t) = \frac{\lambda}{(\lambda-t)^2}, \quad M''(t) = \frac{2\lambda}{(\lambda-t)^3} \circ$$

故

$$\begin{aligned}E(X) &= M'(0) = \frac{1}{\lambda}, \\ E(X^2) &= M''(0) = \frac{2}{\lambda^2} \circ\end{aligned}$$

結果相同。

習題

第 1.1 節

1. 試證 $A \setminus B = A \cap B^c$ 。
2. 對一樣本空間 Ω ，試證下述三集合皆為 σ -體。
 - (i) $C_1 = \{\emptyset, \Omega\}$;
 - (ii) $C_2 = \Omega$ 之所有子集合所形成之集合;
 - (iii) $C_3 = \{\emptyset, \Omega, A, A^c\}$ ，其中 $\emptyset \neq A \subset \Omega$ 。

3. 令 $\Omega = \{1, 2, 3, 4\}$, 試決定下述集合 C 是否為一 σ -體。

$$C = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \\ \{1, 2, 3\}, \{1, 3, 4\}, \{2, 3, 4\}, \Omega\}.$$

4. 設 $\Omega = \{1, 2, 3\}$, 試寫出包含 $\{1\}$ 及 $\{1, 3\}$ 之最小的 σ -體, 並問此 σ -體與已包含 $\{1\}$ 及 $\{3\}$ 之最小的 σ -體是否相同?
5. 設 $\Omega = \{1, 2, 3, 4\}$ 。試寫出包含 $\{2\}$ 及 $\{1, 4\}$ 之最小的 σ -體。
6. 設 $\Omega = \{1, 2, 3, 4, 5\}$ 。試寫出包含 $\{1\}$, $\{2, 3\}$ 及 $\{2, 3, 4\}$ 之最小的 σ -體。
7. 設 $\Omega = \{1, 2, 3, 4, 5, 6\}$, 已知一 σ -體中有三個元素 $\{1\}$, $\{1, 2\}$ 及 $\{3, 4\}$ 。試寫出最小的這種 σ -體。
8. 設 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。試問元素個數為 4 及 8 的 σ -體各有那些。
9. 設 Ω 中有 n 個元素, 試證 Ω 共有 2^n 個不同的子集合。
10. 試證 Ω 之一些(至少一個)子集合, 所形成的集合 $\mathcal{F}$ 為一 σ -體, 若且唯若
- (i) 若 $A \in \mathcal{F}$, 則 $A^c \in \mathcal{F}$;
- (ii) 若 $A_i \in \mathcal{F}$, $i = 1, 2, \dots$, 則 $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ 。
11. 設 T 為一指標集合, 且對 $\forall \alpha \in T$, $\mathcal{F}_\alpha$ 為一由 Ω 之子集合所形成之 σ -體。試證 $\bigcap_{\alpha \in T} \mathcal{F}_\alpha$ 仍為一 Ω 之子集合所形成之 σ -體。
12. 設 $a < b$ 。試證

$$(i) (a, b] = \bigcap_{n=1}^{\infty} (a, b + 1/n) = \bigcap_{n=1}^{\infty} (a, b + 1/n];$$

$$(ii) [a, b) = \bigcap_{n=1}^{\infty} (a - 1/n, b) = \bigcap_{n=1}^{\infty} [a - 1/n, b);$$

$$(iii) (a, b) = \bigcup_{n=1}^{\infty} (a, b - 1/n] = \bigcup_{n=1}^{\infty} (a, b - 1/n) = \bigcup_{n=1}^{\infty} [a + 1/n, b) = \bigcup_{n=1}^{\infty} (a + 1/n, b), \text{ 此處若 } b - 1/n < a, \text{ 則定義 } (a, b - 1/n] = (a, b - 1/n) = \emptyset;$$

$$(iv) \{a\} = \bigcap_{n=1}^{\infty} [a, a+1/n] = \bigcap_{n=1}^{\infty} [a, a+1/n] = \bigcap_{n=1}^{\infty} (a-1/n, a] = \bigcap_{n=1}^{\infty} [a-1/n, a];$$

$$(v) \emptyset = \bigcap_{n=1}^{\infty} (a, a+1/n) = \bigcap_{n=1}^{\infty} (a-1/n, a)。$$

13. 投擲一骰子一次，並觀測所得之點數。試給出兩個不同的機率空間。
14. 試證例1.10中的 $(\Omega, \mathcal{F}, P)$ 為一機率空間。
15. 試證例1.11中的 $(\Omega, \mathcal{F}, P)$ 為一機率空間。
16. 試證定理1.2。
17. 試證定理1.3。
18. 試證定理1.4。
19. 在例1.11中，將 $f(k)$ 改為 $2 \cdot 3^{-k}$, $k \geq 1$ 。試證仍可定義出一機率函數，並求此時之 $P(A)$ 。
20. 設 $P(A) = 1/3$, $P(B^c) = 1/4$ ，試問 A, B 可否為互斥事件。
21. 試在例1.7之 σ -體上，定義一機率函數，滿足 $P(\{1, 2\}) = 0.2$, $P(\{1, 2, 3\}) = 0.5$ 。
22. 設 $\Omega = \{1, 2, 3, 4\}$ ，已知 Ω 上一 σ -體有兩個元素 $\{1\}$ 及 $\{1, 2\}$ 。試寫出最小的這種 σ -體，並定義一機率密度函數，滿足 $P(\{1\}) = 0.1$, $P(\{1, 2\}) = 0.3$ 。
23. 試對第6題中之 σ -體，給出一機率函數，滿足 $P(\{1\}) = 0.1$, $P(\{1, 2, 3\}) = 0.3$, $P(\{1, 2, 3, 4\}) = 0.6$ 。
24. 試求方程式 $x + y + z = 10$ 之正整數解的個數，及非負整數解的個數。
25. 某家庭有10位成員，試求其生日皆相異的機率。
26. 設事件 $A_1 \subset A_2 \subset A_3$ ，且 $P(A_1) = 1/4$, $P(A_2) = 5/12$, $P(A_3) = 7/12$ 。試求下述各事件之機率： $A_1^c \cap A_2$, $A_1^c \cap A_3$, $A_2^c \cap A_3$, $A_1 \cap A_2^c \cap A_3^c$, $A_1^c \cap A_2^c \cap A_3^c$ 。

27. 試證對二事件 A, B , 若 $P(B) = 1$, 則 $P(A \cap B) = P(A)$, $P(A \cup B) = 1$, $P(B \setminus A) = 1 - P(A)$ 。
28. 設一袋中有20個球, 編號為1至20。依序取二球, 每次取出後不放回。令 X_1, X_2 表所得之編號。試求下述二事件之機率:
- (i) $X_1 + X_2 = 8$;
- (ii) $X_1 + X_2 \leq 5$ 。

第1.2節

1. 試證定理2.1。
2. 試證定理2.2。
3. 試證定理2.3。
4. 甲、乙兩人棋力相當, 在進行七戰四勝有獎金的比賽中, 下了三盤甲以2:1領先, 但因故無法繼續比賽。試問如何合理地分配獎金。
5. 投擲一公正的銅板10次, 已知得到6個正面。試求下述二條件機率:
 - (i) 第一次得到正面;
 - (ii) 首五次中得到3個正面。
6. 投擲二公正的骰子, 已知點數和為偶數。試求得到點數和為8之機率。
7. 假設有 A, B, C 三個盒子, 其中恰有一盒裝有獎品。某君任選一盒, 並宣佈他的選擇。然後主持人打開另兩盒中的某一盒, 並發現為一空盒。這時主持人問該君是否願意改變選擇。試分別對主持人事先知道及不知道獎品在那一盒, 回答該君是否須改變選擇。
8. 假設生男生女之機率均為 $1/2$ 。且設一家庭中有 n 個小孩之機率為 $(1/2) \cdot (2/3)^n$, $n \geq 1$ 。試求家庭中沒有男孩之機率。(解. $1/4$)
9. 試證

- (i) 若 A, B 為二互斥事件, 則 A 與 B 獨立, 若且唯若 $P(A) = 0$ 或 $P(B) = 0$;
- (ii) 若 $A \subset B$, 則除非 $P(A) = 0$ 或 $P(B) = 1$, 否則 A 與 B 不獨立;
- (iii) 若 A 與 A 獨立, 則 $P(A) = 0$ 或 1 。
10. 設 A 與 B 獨立, A 與 C 獨立, 且 $B \cap C = \emptyset$ 。
- (i) 試證 A 與 $B \cup C$ 獨立;
- (ii) 試舉一例說明若 $B \cap C \neq \emptyset$, 則 (i) 之結論便不一定成立。
11. 設一袋中有 5 個黑球, 3 個紅球及 2 個白球, 共 10 個球。隨機地取出 4 球, 每次取出後不放回。試求第一次取出黑球, 第二次取出紅球, 第三次取出白球且第四次取出黑球之機率。(解. $1/42$)
12. 設 $\Omega = \{1, 2, 3, 4\}$, 且 $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = 1/4$ 。令 $A_1 = \{1, 2\}$, $A_2 = \{1, 3\}$, $A_3 = \{1, 4\}$ 。試決定下述 (i), (ii) 中, 那些事件相互獨立。
- (i) A_1, A_2 ;
- (ii) A_1, A_2, A_3 。
13. 設 $P(A), P(B) > 0$, 且 $P(A|B) > P(A)$ 。試證 $P(B|A) > P(B)$ 。
14. 我們知道若 $P(B) > 0$, 且 A 與 B 獨立, 則 $P(A|B) = P(A)$ 。但若 A 與 B 不獨立, 是否必有 $P(A|B) > P(A)$? 若是則證明, 不是則舉例說明。
15. 試證
- (i) $P(A^c|B) = 1 - P(A|B)$;
- (ii) $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$ 。
16. 試舉例說明下述二結果不一定成立。
- (i) $P(A|B^c) = 1 - P(A|B)$;
- (ii) $P(C|A \cup B) = P(C|A) + P(C|B)$ 。

17. 設 $A \cap B = \emptyset$, 且 $P(A \cup B) > 0$ 。試以 $P(A)$ 及 $P(B)$ 來表示 $P(A|A \cup B)$ 及 $P(B|A \cup B)$ 。

18. 設 $A_1, A_2, \dots$ 為每對互斥, 且 $P(K) > 0$ 。試證

$$P\left(\bigcup_{i=1}^{\infty} A_i | K\right) = \sum_{i=1}^{\infty} P(A_i | K)。$$

特別地, 若 $A_1, A_2, \dots$ 為樣本空間 Ω 之一分割, 則 $P(\bigcup_{i=1}^{\infty} A_i | K) = 1$ 。

19. 設 $A_1, \dots, A_5$ 為 Ω 之一分割, A 為一事件, 且 $P(A_j) = j/15$, $P(A|A_j) = (5-j)/15$, $j = 1, \dots, 5$ 。試求 $P(A_j|A)$, $j = 1, \dots, 5$ 。

20. 設有事件 $A_1, \dots, A_n$, 且 $P(\bigcap_{j=1}^{n-1} A_j) > 0$ 。試證

$$\begin{aligned} P\left(\bigcap_{j=1}^n A_j\right) &= P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) \\ &\quad \cdot P(A_{n-1} | A_1 \cap \dots \cap A_{n-2}) \cdots P(A_2 | A_1) P(A_1)。 \end{aligned}$$

21. 試證系理2.1。

22. 設一家庭中有5個小孩, 已知其中有4個男孩。試求另一個亦為男孩之機率。

23. 自一副撲克牌中任取5張, 已知至少取中三張黑桃, 則五張皆是黑桃的機率為何?

24. 設一家庭中有 k 個小孩。令事件 A 表此家庭中有男孩及女孩, B 表最多有一女孩。試對 $k = 2, 3, 4$, 討論 A 與 B 是否獨立。

25. 你在某一宴會中交了一新朋友, 他告訴你他有兩個小孩。

(i) 若你問他是否有女兒, 他答有, 則兩個小孩皆為女兒之機率為何?

(ii) 若你問他是否有女兒, 他答有, 這時一個小孩跑過來, 他說這是我女兒, 則兩個女孩皆為女兒之機率為何? 試加以討論。

26. 試證在例2.5中, 若欲使 $P(A_1) = P(A_2) = P(A_3) = 1/2$, 且 A_1, A_2, A_3 相互獨立, 則每一樣本點之機率均須為 $1/8$ 。

27. 試對例2.5之 Ω , 給一機率函數,

- (i) 使得 $P(A_1)$, $P(A_2)$ 與 $P(A_3)$ 不皆為 $1/2$, 但 A_1, A_2, A_3 仍相互獨立;
- (ii) 使得 $P(A_1) = P(A_2) = P(A_3) = 1/2$, 但 A_1, A_2, A_3 不相互獨立。

第1.3節

1. 試證由(3.2)式所定義出的函數 P_X , 的確為一機率函數。
2. 隨機地投擲7個球進7個盒子。令 X 表有3個球之盒子數。試對所有可能的 x , 求 $P(X = x)$ 。
3. 投擲一公正的骰子一次, 並獲出現點數之二倍的賞金。令 X 表所得賞金。試求 $P(5 \leq X \leq 10)$ 。
4. 投擲一公正的骰子一次, 令 X 表所得之點數除以4之餘數。試求 X 之值域 Ω_1 , 並對 $\forall x \in \Omega_1$, 給出 $P_X(X = x)$ 。
5. 在例2.5中, 令 X = 正面數的平方。試求 X 之值域 Ω_1 。並對 $\forall x \in \Omega_1$, 給出 $P_X(X = x)$ 。

第1.4節

1. 試驗證例4.7中之分佈函數, 確如所給的形式。
2. 試驗證例4.12中之分佈函數, 確如所給的形式。
3. 試求出例4.13中之分佈函數。
4. 試求出例4.14中之分佈函數。並驗證圖4.11中之三角形面積為1。
5. 試證(4.12)式確實定義一機率密度函數, 並求對應的分佈函數 F 。
6. 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} (32/31)(1/2)^x, & x = 1, 2, 3, 4, 5, \\ 0 & \text{其他。} \end{cases}$$

- (i) 試求其分佈函數 $F(x)$;
- (ii) 試求 $P(X \geq 3)$, $P(2 \leq X \leq 4)$, 及 $P(X < 4)$ 。

7. 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} (1/4)(3/4)^x, & x = 0, 1, 2, \dots, \\ 0 & , \text{其他}。 \end{cases}$$

(i) 試求其分佈函數 $F(x)$;

(ii) 試求 $P(X > 3)$, $P(1 \leq X \leq 5)$, 及 $P(X < 3)$ 。

8. 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} 0 & , x < 0 \\ x & , 0 \leq x < 1, \\ 1/2 & , 1 \leq x < 2, \\ 0 & , x \geq 2。 \end{cases}$$

試求 X 之分佈函數 $F(x)$, 並分別繪 f 及 F 之圖形。

9. 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0 & , \text{其他}。 \end{cases}$$

(i) 試求 X 之分佈函數 $F(x)$;

(ii) 試求 $P(-1 < X \leq 1/2)$, 及 $P(1 \leq X \leq 3/4)$ 。

10. 設隨機變數 X 之機率密度函數為

$$f(x) = \begin{cases} x - 1, & 1 < x < 2, \\ 3 - x, & 2 \leq x < 3, \\ 0 & , \text{其他}。 \end{cases}$$

(i) 試求 X 之分佈函數 $F(x)$;

(ii) 試求 $P(X \leq 3/2)$, $P(3/2 \leq X \leq 5/2)$, 及 $P(X > 5/2)$ 。

11. 設某種燈泡的壽命 X (單位為月), 以 $f(x) = (1+x)^{-2}$, $x > 0$, 為其機率密度函數。試求壽命至少是3年之機率。

12. 試證下述函數爲一連續型隨機變數的機率密度函數。並繪 $f(x)$ 及其分佈函數 $F(x)$ 之圖形。

$$f(x) = \begin{cases} e^x/2, & x \leq 0, \\ 1/4, & 0 < x < 2, \\ 0, & \text{其他。} \end{cases}$$

13. 試證下述各函數皆爲分佈函數, 並繪其圖形。

- (i) $F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} x, x \in R;$
- (ii) $F(x) = e^{-e^{-x}}, x \in R;$
- (iii) $F(x) = 1 - e^{-x}, x > 0;$
- (iv) $F(x) = 1 - e^{-\lambda x^2}, x > 0$, 其中 $\lambda > 0$ 爲一常數;
- (v) $F(x) = 1 - 1/x^2, x \geq 1$ 。

14. 對上題各分佈函數, 分別求其機率密度函數並繪其圖形。

15. 試決定下述那些函數爲機率密度函數。

- (i) $f(x) = 5x^4, 0 < x < 1;$
- (ii) $f(x) = 2x, -1 \leq x \leq 2;$
- (iii) $f(x) = 1/2, -1 \leq x \leq 1;$
- (iv) $f(x) = x^{-1/2}, 1/4 \leq x \leq 1;$
- (v) $f(x) = x^{-2}, x \geq 1;$
- (vi) $f(x) = (1+x)^{-2}, x \geq 1$ 。

16. 試對下述各函數, 分別決定常數 c , 使其成爲一機率密度函數。

- (i) $f(x) = c \sin x, 0 < x < \pi/2;$
- (ii) $f(x) = ce^{-|x|}, x \in R;$
- (iii) $f(x) = cx^2e^{-x^3}, x > 0;$
- (iv) $f(x) = \frac{c}{x(x+1)}, x = 2, 3, 4, \dots;$

$$(v) f(x) = \begin{cases} 0 & , x \leq -1, \\ -cx & , -1 < x \leq 0, \\ ce^{-6x} & , x > 0. \end{cases}$$

17. 試證下述函數 F 為一分佈函數(即參數 α, β 之羅吉斯分佈):

$$F(x) = \frac{1}{1 + e^{-(\alpha x + \beta)}}, x \in R,$$

其中 $\alpha > 0, \beta \in R$ 為常數。試求其機率密度函數 f , 並證明 f 滿足 $f(x) = \alpha F(x)(1 - F(x))$ 。

18. 設 X 為一連續型的隨機變數, 以 f 及 F 分別為其機率密度函數及分佈函數。設 x_0 滿足 $F(x_0) < 1$ 。令

$$g(x) = \begin{cases} f(x)/(1 - F(x_0)) & , x \geq x_0, \\ 0 & , x < x_0. \end{cases}$$

試證 g 亦為一機率密度函數。

19. 投擲一公正的銅板三次, 令 U 表所得正面數的二倍減3, V 表所得正面數減反面數。試證 $U = V$ 。

20. 試給出二隨機變數 X, Y , 滿足 $X = -Y$, 且 $X \stackrel{d}{=} Y$ 。

21. 是否存在二隨機變數 X, Y , 滿足 $X = 1 + Y$, 且 $X \stackrel{d}{=} Y$?

22. 是否存在二隨機變數 X, Y , 滿足 $X = 1 - Y$, 且 $X \stackrel{d}{=} Y$?

23. 設隨機變數 X 之機率密度函數為一偶函數。試證

(i) $X \stackrel{d}{=} -X$;

(ii) 若 X 之分佈函數為 F , 則 $F(x) = 1 - F(-x), \forall x \in R$ 。

24. 試證若隨機變數 X 之機率密度函數對稱於 a , 則(4.13)式成立。

25. 觀測投擲一銅板 n 次所得之正面數, 其中 n 為一定值。試給出樣本空間 Ω , 及一個較一般的機率函數 P 。其次在 Ω 上定義出兩個不同的隨機變數, 並分別給出 X 及 Y 的機率密度函數 f_X, f_Y , 及其分佈函數 F_X, F_Y 。

26. 設 F 及 G 分別為隨機變數 X 及 Y 之分佈函數, 若 $F(x) \leq G(x), \forall x \in R$, 且對某 $x_0 \in R, F(x_0) < G(x_0)$, 則稱 X 比 Y 隨機較大(stochastically greater)。試證此時

$$P(X > x) \geq P(Y > x), \forall x \in R,$$

且對某 $x_0 \in R$,

$$P(X > x_0) > P(Y > x_0)。$$

第1.5節

1. 試證(5.2)式定義出一p.d.f., 並求其分佈函數。
2. 設隨機變數 X 有 $\mathcal{U}(0, 1)$ 分佈。令 $Y = \log(X/(1 - X))$ 。試證 Y 有標準的羅吉斯分佈。
3. 設隨機變數 X 有 $\mathcal{E}(1)$ 分佈。令 $Y = X^2$ 。試決定 Y 之p.d.f.及分佈函數。
4. 設隨機變數 X 有 $\mathcal{E}(1)$ 分佈。令 $Y = e^X$ 。試決定 Y 之p.d.f.及分佈函數。
5. 設隨機變數 X 有 $\mathcal{U}[-1, 1]$ 分佈。試決定 $Y = X^2$ 之p.d.f.及分佈函數。
6. 設隨機變數 X 之p.d.f.為 $f(x) = (1+x)^{-2}, x > 0$ 。試指出 $Y = X/(1+X)$ 有何分佈。
7. 分別對有下述所給p.d.f.之各隨機變數 X 及變換, 試求 Y 之p.d.f. f_Y 。
 - (i) $f_X(x) = 42x^5(1-x), 0 < x < 1, Y = X^3$;
 - (ii) $f_X(x) = 1, 0 < x < 1, Y = X^5$;
 - (iii) $f_X(x) = \frac{1}{\sigma}xe^{-x^2/(2\sigma)}, x > 0, \sigma > 0$ 為一常數, $Y = e^X$;
 - (iv) $f_X(x) = \frac{1}{3}(\frac{2}{3})^x, x = 0, 1, 2, \dots, Y = X/(X+1)$;
 - (v) $f_X(x) = \frac{1}{2}e^{-|x|}, x \in R, Y = |X|^3$;

$$(vi) f_X(x) = \frac{3}{8}(x+1)^2, -1 < x < 1, Y = 1 - X^2;$$

$$(vii) f_X(x) = \frac{2}{9}(x+1), -1 \leq x \leq 2, Y = X^2。$$

8. 設隨機變數 X 之p.d.f.為 f_X , 分佈函數為 F_X 。令 $Y = a + bX$, 其中 $a \in R, b > 0$ 。試分別以 F_X 及 f_X 表示 Y 之分佈函數 F_Y , 及p.d.f. f_Y 。
9. 利用上題, 當 X 有標準羅吉斯分佈時, 試求 Y 之分佈函數及p.d.f.。
10. 設隨機變數 X 有 $\mathcal{U}(a, b)$ 分佈, $a < b$, 令 $Y = \alpha + \beta X, \alpha \in R, \beta > 0$ 。試決定 Y 之分佈。
11. 設隨機變數 X 之p.d.f.為 f_X , 分佈函數為 F_X , 且 $F_X(0) = 0$ 。令 $Y = \sqrt{X}$ 。試分別以 F_X 及 f_X , 表示 Y 之分佈函數 F_Y , 及p.d.f. f_Y 。
12. 設隨機變數 X 之p.d.f. f_X 為一連續的偶函數, 分佈函數為 F_X 。令 $Y = X^2$ 。試分別以 F_X 及 f_X , 表示 Y 之分佈函數 F_Y , 及p.d.f. f_Y 。
13. 利用上題, 當 X 有 $\mathcal{C}(0, 1)$ 分佈, 試求 Y 之分佈函數及p.d.f.。
14. 設隨機變數 X 有 $\mathcal{C}(0, 1)$ 分佈。試證 $X \stackrel{d}{=} 1/X$ 。
15. 試證定理5.3。
16. 試分別對下述二分佈, 求其分佈函數之分位函數, 並繪分佈函數之圖。
(i) 標準的羅吉斯分佈, (ii) $\mathcal{C}(0, 1)$ 分佈。
17. 設有分佈函數

$$F(x) = \begin{cases} 0 & , x < 0, \\ 0.5x & , 0 \leq x < 1, \\ 0.5 & , 1 \leq x < 2, \\ 0.5x - 0.5 & , 2 \leq x < 3, \\ 1 & , x \geq 3. \end{cases}$$

試繪 F 及 F^{-1} 之圖形, 並驗證 $F(F^{-1}(x)) = x, \forall x \in (0, 1)$ 。

18. 試證下述 $F(x)$ 為一分佈函數, 並求 $F^{-1}(x)$ 。

$$F(x) = \begin{cases} e^x/2 & , x < 0, \\ 1/2 & , 0 \leq x < 1, \\ 1 - e^{1-x}/2 & , x \geq 1. \end{cases}$$

19. 設隨機變數 X 之分佈函數為 $F(x) = 1 - x^{-1}$, $x > 1$ 。試直接驗證 $F(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

20. 設隨機變數 X 有 $\mathcal{E}(\lambda)$ 分佈, $\lambda > 0$ 。令 F 表其分佈函數。試直接驗證 $F(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

21. 設隨機變數 X 之分佈函數為 $F(x) = 1 - e^{-\sqrt{x}}$, $x > 0$ 。令 $Y = F(X)$, 試證 Y 有 $\mathcal{U}(0, 1)$ 分佈。

22. 設隨機變數 X 之p.d.f.為 $f(x) = (x - 1)/2$, $1 < x < 3$ 。試求一函數 $g(x)$, 使得 $Y = g(X)$ 有 $\mathcal{U}(0, 1)$ 分佈。

第1.6節

1. 試驗證例6.10中, $g(y)$ 為一p.d.f., 且 $E(Y) = 16$, $\text{Var}(Y) = 1/150$ 。
2. 設一袋中有9個球, 編號分別為1至9。隨機地取一個球, 求所得數字之期望值及變異數。
3. 設一袋中有10張紙牌, 其中有一張有點數1, 兩張有點數2, 三張有點數3, 四張有點數4。隨機地取一張, 求所得點數之期望值及變異數。
4. 投擲二公正的骰子, 令 X 表較小的點數, Y 表較大的點數, Z 表點數和。試求 $E(X)$, $E(Y)$, 及 $E(Z)$ 。驗證 $E(Z)$ 是否等於 $E(X) + E(Y)$, 並解釋此現象是否合理。
5. 設隨機變數 X 之p.d.f.為 $f(x) = 6x(1 - x)$, $0 < x < 1$ 。試分別以例6.1中的兩個方法求 $E((X - 1)^2)$ 。
6. 設隨機變數 X 之p.d.f.為 $f(x) = \lambda x^{\lambda-1}$, $0 < x < 1$, $\lambda > 0$ 為一常數。試求 $E(X)$ 及 $E(1/X)$ 。

7. 設隨機變數 X 有 $\mathcal{U}(-2, 0)$ 分佈。試求 $E(1/X)$ 。
8. 設隨機變數 X 有 $\mathcal{U}[0, 3\pi]$ 分佈。試求 $E(\sin X)$ 。
9. 設隨機變數 X 有 $\mathcal{U}(-1, 1)$ 分佈。試求 $E(X^2)$ 、 $E(X^4)$ ，及 $\text{Var}(X^2)$ 。
10. 設隨機變數 X 有 $\mathcal{U}(0, 1)$ 分佈。試求 $\log X$ 之期望值、二次動差，及變異數。
11. 設隨機變數 X 之p.d.f.為 $f(x) = 4^{-1}(3/4)^x$, $x = 0, 1, \dots$ 。試求 X 之期望值、中位數及眾數。
12. 設 $f(x) = xe^{-x}$, $x \geq 0$ 。試證此為一p.d.f.。並求此分佈之期望值、變異數、中位數及眾數。
13. 設隨機變數 X 有 $\mathcal{E}(\lambda)$ 分佈, $\lambda > 0$ 。試求 $E(X)$, $\text{Var}(X)$, $E(e^X)$, X 之中位數及眾數。
14. 對有下述p.d.f.之隨機變數, 分別求其中位數及眾數。
 - (i) $f(x) = 3x^2$, $0 < x < 1$;
 - (ii) $f(x) = 1/(\pi(1+x^2))$, $x \in R$;
 - (iii) $f(x) = 2x^{-2}$, $1 \leq x < 2$;
 - (iv) $f(x) = e^{-x}/(1+e^{-x})^2$, $x \in R$ 。
15. 設隨機變數 X 之分佈函數為

$$F(x) = \begin{cases} 0 & , x < 0, \\ x/4 & , 0 \leq x < 1, \\ x^2/4 & , 1 \leq x < 2, \\ 1 & , x \geq 2. \end{cases}$$

試求 X 之期望值及中位數。

16. (i) 設隨機變數 X 之p.d.f.為一偶函數。試證此時 X 之中位數為0, 且若期望值存在便亦為0。
- (ii) 設隨機變數 X 之分佈對稱於 m 。試證此時 X 之中位數為 m , 且若期望值存在便亦為0。

17. 令 X 表投擲一公正的骰子所得之點數。試求 X 之母函數。
18. 設隨機變數 X 表投擲一公正的骰子所得之點數。試求 X 之母函數 $g(s)$ 。
19. 分別對下述各p.d.f., 求其動差母函數。
 - (i) $f(x) = 1/c, 0 < x < c$;
 - (ii) $f(x) = 2x/c^2, 0 < x < c$;
 - (iii) $f(x) = (e - 1)^{-1}e^x, 0 \leq x \leq 1$;
 - (iv) $f(x) = (e^2 + 1)^{-1}xe^x, 0 \leq x \leq 2$;
 - (v) $f(x) = (2\beta)^{-1}e^{-|x-\alpha|/\beta}, x \in R, \alpha \in R, \beta > 0$ 。
20. 設隨機變數 X 之p.d.f.為 $f(x) = \frac{1}{\sqrt{2\pi}x}e^{-(\log x)^2/2}, x > 0$ 。試利用 $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx = 1$, 證明 $E(X^r) = e^{r^2/2}, r = 0, 1, 2, \dots$ 。
21. 設隨機變數 X 之p.d.f.為 $f(x) = (1+x)/2, -1 < x < 1$ 。令 $Y = X^2$ 。試求 Y 之p.d.f., 及 $E(Y), \text{Var}(Y), E(e^{tY})$ 。
22. 設隨機變數 X 之動差母函數 $M(t) = e^{t^2/2}, t \in R$ 。試求 $E(X^k), k = 1, 2, 3, 4$ 。
23. 設隨機變數 X 之p.d.f.為一偶函數。試證
 - (i) $E(X)$ 若存在必為0;
 - (ii) $M_X(t)$ 若存在, 則滿足 $M_X(t) = M_X(-t)$ 。
24. 設隨機變數 X 之變異數為0。試證存在常數 b , 使得 $P(X = b) = 1$ 。
25. 對例6.8之賭法略做修改。令 N 為一固定正整數, 若第1次至第 N 次投擲皆未出現正面, 則一無所得。試問此時每次賭之前, 賭徒該付多少, 此賭局才是公平。
26. 對一隨機變數 X , 考慮轉換 $E(e^{-sX^2}), s \geq 0$ 。試問此轉換是否可唯一決定 X 之分佈。

27. 設 X 為一連續型的隨機變數, 以 F 為其分佈函數。試證

$$E(X) = \int_0^{\infty} (1 - F(x))dx - \int_{-\infty}^0 F(x)dx,$$

且當 $F(0) = 0$ 時,

$$E(X) = \int_0^{\infty} (1 - F(x))dx。$$

28. 設 X 為一取非負整數值之隨機變數。試證

$$E(X) = \sum_{k=1}^{\infty} P(X \geq k)。$$

29. 對一隨機變數 X , 試證對任一 $n \geq 1$, $E(X^n)$ 存在, 若且唯若 $E(|X|^n)$ 存在。但試舉一例來說明, 對函數 $g(x)$, $\int_{-\infty}^{\infty} g(x)dx$ 存在, 卻並不一定導致 $\int_{-\infty}^{\infty} |g(x)|dx$ 存在。

30. 試證標準羅吉斯分佈之任意次動差皆存在。

31. 設 X 之分佈函數為 $F(x) = 1 - e^{-\sqrt{x}}$, $x > 0$ 。試求 $E(X)$ 及 $\text{Var}(X)$ 。

32. 設有 50 個白球及 50 個紅球, 將此 100 個球分放進兩個袋子中後, 先任選一袋, 再自該袋中任取一球。試問

(i) 該如何放球, 才可以使選中白球的機率最大?

(ii) 可否同時使選中白球及紅球的機率皆最大?

(iii) 令 A 表選中白球之事件, I_A 表 A 之指示函數, 如何放球可使 $E(I_A)$ 最小((i)中即求 $E(I_A)$ 最大)?

33. 設有 X, Y, Z, W 四隨機變數, 可能取的值均為 0, 1, 2, p.d.f. 分別為

$$\begin{aligned} f_X(0) &= f_X(1) = f_X(2) = \frac{1}{3}, \\ f_Y(0) &= \frac{1}{4}, \quad f_Y(1) = \frac{1}{2}, \quad f_Y(2) = \frac{1}{4}, \\ f_Z(0) &= \frac{1}{2}, \quad f_Z(2) = \frac{1}{2}, \\ f_W(0) &= \frac{1}{3}, \quad f_W(2) = \frac{2}{3}。 \end{aligned}$$

試證 $E(X) = E(Y) = E(Z) = 1$, $E(W) = 4/3$, 且 $\text{Var}(Z) > \text{Var}(W) > \text{Var}(X) > \text{Var}(Y)$ 。說明變異數有此大小關係是否合理, 又問取值在 0,1,2 的隨機變數, 是否有變異數大於 $\text{Var}(Z)$ 者。