

管理數學

第 8 版

王妙伶、陳獻清、黎煥中、廖珊慧 編著

CH06 機率論

目錄

6.1 緒 論

6.2 基本概念

6.3 機率測度方法

6.4 條件機率

6.5 獨立事件

6.6 貝氏定理

6.7 機率分配

6.8 分配函數

6.9 期望值與變異數

6.10 柴比雪夫不等式

6.11 常用離散分配

6.12 常態分配





6.1 緒論

人們常說「明日降雨機率80%」，而明天對今天而言，是否一定會下雨，通常不可預知，因此有大部分的人在出門前會帶著雨具。「統計學」即其中之一，透過蒐集、整理、分析、陳示等步驟分析統計資料，解釋統計資料，而能在不確定的情況下幫助人們擬定決策。其中「機率論」則是為衡量某一事件可能發生程度的一門學問，此理論乃統計學上推論統計之基礎。

在從事各種經濟的、生物的或社會的現象等關於結果不能確定的科學研究時，我們都必須用到機率(probability)的概念。





6.2 基本概念

在研究機率論之前，首先必須了解以下幾個名詞之意義，以為基礎：

出象(outcome) 一事件可能發生的結果。

隨機試驗 (random experiment) 一個試驗中有多個可能的出象，但其中只有一個會發生，而在試驗前並無法預知此結果，這種試驗過程，即稱為隨機試驗。

樣本空間 (sample space) 一個隨機試驗可能發生的結果所成之集合即稱為此試驗之樣本空間，通常以大寫英文字母 S 表示之。

樣本點(sample point) 樣本空間中的任一個元素皆稱為一個樣本點。



離散樣本空間(discrete sample space) 當樣本空間中樣本點的個數有限多個、或為可數個時，稱此樣本空間為有限或離散樣本空間。

連續樣本空間 (continuous sample space) 當樣本空間中樣本點的個數為無限多個時，稱之為無限或連續樣本空間。

事件(event) 樣本空間之子集合稱為事件，通常以大寫英文字母表示之。

簡單事件(simple event) 一事件中僅含有一元素則稱為簡單事件，反之則稱之為複合事件(compound event)。簡單事件通常以大寫英文字母 $E_1, E_2, \dots, E_n$ 表示之。

互斥事件(disjoint event) 當二事件無共同的元素時稱此二事件為互斥事件。





6.3 機率測度方法

機率論為測度一事件發生程度的工具，隨機試驗中之一事件發生的可能程度即稱為此事件發生的機率(probability)。



古典方法

事件大小

樣本空間大小

一離散樣本空間中，每一樣本點有相同發生的可能性，例如擲一骰子時，出現一點和其他點的機會相同，則此事件發生的機率應該是 $\frac{|A|}{|S|}$ 和 $\frac{|A|}{|S|}$ 之比值，亦即若 $|A|$ ， $|S|$ 分別表示事件 A 和樣本空間 S 之元素個數時， A 發生的機率可定義為
$$P(A) = \frac{|A|}{|S|}$$

若為一連續樣本空間則 $|A|$ 和 $|S|$ 將可定義為事件 A 及樣本空間 S 之大小。



樣本空間

相對次數方法

使用古典機率必須先將其樣本空間明確寫出來，但若某一試驗並無法確切地建立出其樣本空間時，便無法使用古典方法來測度機率。

總試驗次數的比

例

當連續進行此種試驗時，佔，即定義為此事件發生的機率，即若令 n_A 表事件 A 發生的次數，表示截至計算之同時總試驗次數則

$$P(A) = \frac{n_A}{n}$$

事件發生的次數



主觀方法

主觀方法乃根據個人對一事件主觀認定的看法來決定此事件發生的機率，即：

$P(A)$ = 相信事件 A 會發生的程度

公理 6.1

1. 任一事件的機率必大於等於0，但小於等於1。即
 $0 \leq P(A) \leq 1$
2. 令樣本空間為 S 則 $P(S) = 1$ 。
3. 若事件 $A_1, A_2, \dots, A_n$ 為樣本空間中之 n 個互斥事件，即對任意 $i \neq j, 1 \leq i, j \leq n, A_i \cap A_j = \Phi$ ，則
 $P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$



定理6.1 機率特性

1. $P(\phi)=0$ ， ϕ 表空集合。
2. A, B 為互斥事件則 $P(A \cap B)=0$ ， $P(A \cup B)=P(A)+P(B)$ 。
3. $P(\bar{A})=1-P(A)$ ， $\bar{A}$ 表示 A 之補集合。
4. 若 A, B, C 為樣本空間中之三事件，則事件 $A \cup B$ 之機率為

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

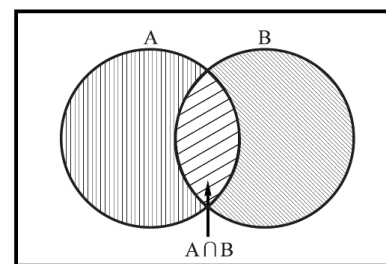


圖6.1
 $A \cup B$ 之文氏圖形

此性質可用文氏圖形(venn diagram)來解釋：

5. 若 $A_1, A_2, \dots, A_n$ 為樣本空間中之 n 個事件，則事件 $A_1 \cup A_2 \cup \dots \cup A_n$ 之機率為

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) = & \sum_{i=1}^n P(A_i) - \sum_{\substack{i < j}} \sum_j P(A_i \cap A_j) + \sum_{\substack{i < j < k}} \sum_j \sum_k P(A_i \cap A_j \cap A_k) \\ & + \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$



6.4 條件機率

已知B事件的發生，則A事件發生的機率即稱為條件機率 (conditional probability)，以 $P(A|B)$ 表示之，

而定義為：
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

同理， $P(B|A)$ 已知事件A發生後，事件B發生的機率 $= \frac{P(A \cap B)}{P(A)}$

綜合以上兩個條件機率之公式，可推演出以下之法則：

定理6.2 機率乘法法則

1. $P(A \cap B) = P(A)P(B | A) = P(B)P(A | B)$
2.
$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_k | A_1 \cap A_2 \cap \dots \cap A_{k-1})$$





6.5 獨立事件

當事件的發生與事件是否發生不受影響時，稱事件對事件為獨立，此時條件機率，故有以下之定義：

定義6.1

若 $P(A \cap B) = P(A)P(B)$ 則稱二事件A，B為獨立事件 (independent event)。反之，則稱之為相依事件(dependent event)。



定理6.3

三個以上之事件若彼此之間均能滿足獨立事件之定義時，方能稱這些事件彼此獨立。如：三事件A、B、C必須滿足：

1. $P(A \cap B) = P(A)P(B)$ 。
2. $P(A \cap C) = P(A)P(C)$ 。
3. $P(B \cap C) = P(B)P(C)$ 。
4. $P(A \cap B \cap C) = P(A)P(B)P(C)$ 。

此四個條件，才可稱三事件之獨立。

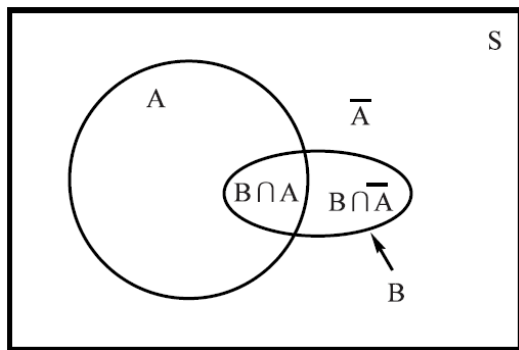




6.6 貝氏定理

任一樣本空間可以不同的標準化分成不同類型的子空間組合，例如：依事件 A 來做劃分，則樣本空間可分成 A 和 A 的補集合 $\bar{A}$ 兩個子空間，且 A 與 $\bar{A}$ 為互斥，因此對另一事件 B 而言，其發生可能來自子 A 空間之樣本與 $\bar{A}$ 子空間之樣本所組合而成，故事件 B 可以寫成： $B = (B \cap A) \cup (B \cap \bar{A})$

且 $B \cap A$ 與 $B \cap \bar{A}$ 為二互斥集合，其文氏圖形如圖6.3所示。



故事件 B 之機率可寫成：

$$P(B) = P(B \cap A) + P(B \cap \bar{A})$$

圖6.3 事件 B 對是 A 之分割文氏圖



定理6.2 總機率法則

令 $A_1, A_2, \dots, A_k$ 為樣本空間 S 中之 k 個互斥事件且滿足 $A_1 \cup A_2 \cup \dots \cup A_k = S$ ， B 為任一事件。則：

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_k)$$

利用定理6.2機率乘法法則與總機率法則，可將已知事件發生下事件發生的條件機率改寫為：

$$P(A_i | B) = \frac{P(A_i \cap B)}{P(B)}$$

此即有名
之貝氏定
理

$$= \frac{P(A_i \cap B)}{P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_k)} \quad (\text{總機率法則})$$

$$= \frac{P(A_i)P(B | A_i)}{P(A_1)P(B | A_1) + \dots + P(A_k)P(B | A_k)} \quad (\text{機率乘法法則})$$

$$= \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^k P(A_j)P(B | A_j)}$$

定理6.5 貝氏定理

若一樣本空間依某一標準分割為 $A_1, \dots, A_k$ 等個互斥之子空間，則已知另一事件發生下，事件之條件機率為：

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^k P(A_j)P(B | A_j)}$$

在貝氏定理中， $P(A_1), \dots, P(A_k)$ 等稱為 $A_1, \dots, A_k$ 等事件之事前機率(prior probability)，而 $P(A_i | B)$ 則稱為事件 A_i 之事後機率(posterior probability)。貝氏定理可應用於決策分析(decision analysis)上作為決策的工具，於第七章中將介紹。





6.7 機率分配

變
量

實數函
數

隨機變數(random variable)為定義在樣本空間之
通常以大寫英文字母 X 表示之，而以小寫字母 x 表示隨機變數 X 所有可能的值， x 即為一。

將隨機變數之各變量 x 的發生機率，以 $f(x)$ 表示之，
按照變量與其相對機率順序排列即稱為此隨機變數
 X 之 (probability distribution)。機率分配可
以透過

表格

式子

圖形

表達。

機
率
分
配



例 6.24

擲骰子之隨機試驗中期機率分配為何？

解

(1) 表格

x	1	2	3	4	5	6	和
f(x)	1/6	1/6	1/6	1/6	1/6	1/6	1

(2) 式子

$$f(x) = \begin{cases} \frac{1}{6} \\ 0 \end{cases}$$

$x = 1, 2, 3, 4, 5, 6$

其他

(3) 圖形

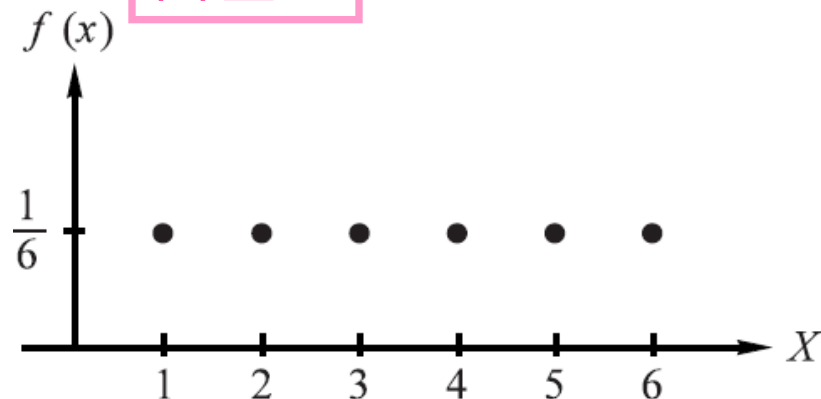


圖6.5 例6.24機率分配之圖示

當隨機變數之變量為有限多個或可數個時，此隨機變數稱為離散隨機變數(discrete random variable)，則其機率分配函數稱為不連續函數，稱之為機率結集函數(probability mass function)。

定義6.2

若離散隨機變數 X 的每一變量 x 均滿足以下條件：

1. $0 \leq f(x) \leq 1$ 。
2. $\sum_{\text{所有 } x} f(x) = 1$ 。

而變量 x 的機率可寫成 $P(X = x) = f(x)$

當隨機變量 X 之變量 x 為無限多個時，則稱 X 為連續隨機變數(continuous random variable)，其機率分配函數為連續函數，以機率密度函數(probability density function)命名之以茲區別。



定義6.3

若一函數 $f(x)$ 滿足

1. $f(x)$ 為 x 之連續函數。
2. $f(x) \geq 0, x \in (-\infty, \infty)$ 。
3. $\int_{-\infty}^{\infty} f(x) dx = 1$ 。

則稱此 $f(x)$ 為 X 之隨機密度函數。並定義事件 A 之機率為

$$P(A) = P(x \in A) = \int_{x \in A} f(x) dx$$





6.8 分配函數

定義6.4

若 $F(\cdot)$ 為隨機變數 x 之分配函數，則 $F(\cdot)$ 滿足以下之性質

1. 若 $a \geq b$ ，則 $F(a) \geq F(b)$ 。
2. 若 $F(-\infty) = 0$ ， $F(\infty) = 1$ 。
3. 任意一數 a ， $0 \leq F(a) \leq 1$ 。

離散隨機變數之分配函數寫為
$$F(a) = P(x \leq a) = \sum P(X=x)$$

連續隨機變數之分配函數寫成
$$F(a) = P(x \leq a) = \int_{-\infty}^a f(x) dx$$

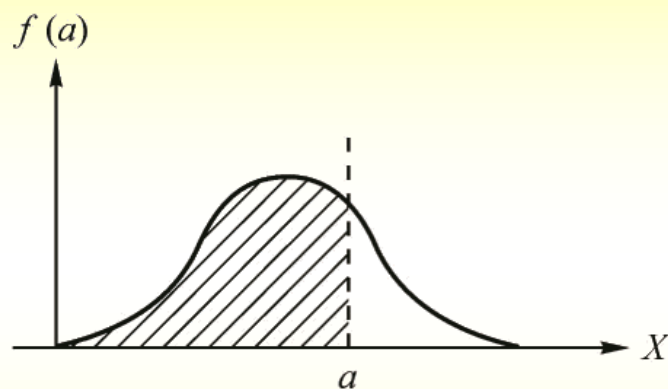


圖 6.7 陰影部分面積即 $F(a)$

亦即在連續變數，累積到 $x=a$ 之面積值即其分配函數值。





6.9 期望值與變異數

定義6.5

若 X 為一隨機變數，其機率函數為 $f(x)$ ，定義 X 之期望值(mean) $= \mu = E(X)$ 為

$$\mu = E(X) = \begin{cases} \sum_{\text{所有 } x} x f(x), & \text{若 } X \text{ 為離散隨機變數} \\ \int_{-\infty}^{\infty} x f(x) dx, & \text{若 } X \text{ 為連續隨機變數} \end{cases}$$

事實上，期望值即隨機變數之各可能值和其相對機率之加權平均值，換言之，期望值可視為母體之平均數，可用來測度其集中趨勢。

定理6.6

由期望值之定義可得以下之基本特性：

1. 常數值 a 之期望值為 a ，即 $E(a) = a$ 。
2. 若 a, b 為二常數， X, Y 為二隨機變數，則

$$E(aX + bY) = aE(X) + bE(Y)$$



定義6.6

隨機變數 X 之變異數(variance)為：

$$\sigma^2 = V_{ar}(X) = E(X - E(X))^2$$

$$= \begin{cases} \sum_{\text{所有 } x} (x - \mu)^2 f(x), & \text{若 } X \text{ 為離散隨機變數} \\ \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx, & \text{若 } X \text{ 為連續隨機變數} \end{cases}$$

變異數之平方根稱之為標準差(standard deviation)，亦即標準差 $\sigma = \sqrt{V_{ar}(X)}$ 。

由期望值之基本性質可得變異數之基本特性如下：

定理6.7

1. a 為一常數則 $V_{ar}(a) = 0$ 。
2. a 、 b 為二常數， X 、 Y 為二獨立隨機變數，則

$$V_{ar}(aX + bY) = a^2 V_{ar}(X) + b^2 V_{ar}(Y)$$

3. $V_{ar}(X) = E(X^2) - (E(X))^2$ 。





6.10 柴比雪夫不等式

定理6.8 柴比雪夫不等式

令 X 為一隨機變數其期望值為 μ ，變異數為 σ^2 ，則對任一正數 k ：

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

或

$$P(|X - \mu| > k\sigma) < \frac{1}{k^2}$$

隨機變數 X 落在平均數左右 k 倍標準差的範圍內的機率將至少等於 $1 - \frac{1}{k^2}$ 或 X 落在平均數左右 k 倍標準差外的機率將小於 $\frac{1}{k^2}$ 。





6.11 常用離散分配

● 均等分配 (uniform distribution)

「均等分配」指的是隨機變數變量之機率值均相等，其定義為：

定義6.7

若隨機變數 $\mathbf{X}$ 之所有可能變量 $\mathbf{x}_1, \dots, \mathbf{x}_k$ 有相等的機率，則稱之為均等分配(uniform distribution)，其機率函數記為

$$f(x) = \frac{1}{k}, x = x_1, x_2, \dots, x_k$$

定理6.9

均等分配之期望值為 $\mu = \frac{\sum_{i=1}^k x_i}{k}$ ，變異數為 $\frac{\sum_{i=1}^k (x_i - \mu)^2}{k}$ 。



● 二項分配 (binomial distribution)

隨機試驗之結果常可用二分法分成二種，如：成功、不成功或良品、不良品等，而 X 為離散隨機變數將此二種不同的結果分別映至0與1二數， $x = 1$ 代表試驗結果為「成功」或「良品」，其機率值為 p ，而 $x = 0$ ，則表示試驗結果為「不成功」或「不良品」，其機率值為 $1 - p$ ，則定義此隨機變數 X 服從白努力分配(Bernoulli distribution)，記為：

$$P(x = 1) = p$$

$$P(x = 0) = 1 - p$$

二項分配(binomial distribution)是由白努力分配衍生而來的。



定義6.8

當試行白努力實驗 n 次，每次試驗均獨立，每次實驗仍只分為成功、不成功二種結果，成功機率均 p 為，而不成功的機率均為 $1-p$ ，若定義隨機變數 X 為此實驗成功的次數，則稱隨機變數 X 符合二項分配，記作 $X \sim B(n, p)$ ，其機率函數為 $P(X=x) = f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ ，其中 $\binom{n}{x}$ 為 n 次中取 x 次組合之記號，其分配函數值可由附表一中查出。

定理6.10

二項分配之期望值為 $\mu = np$ ，變異數為 $np(1-p)$ 。





卜瓦松分配 (Poisson distribution)

卜瓦松分配(Poisson distribntion)為日常生活中常用到的離散機率分配，例如：某交通路口每週所發生的意外事故件數，某工廠每月機台當機台數等均為卜瓦松分配之應用。

定義6.9

於一段時間或一特定區域內進行實驗，記錄其成功或某一事件發生的次數，若 X 表示成功或發生的次數則其分配即服從卜瓦松分配，並具有以下之機率函數 $f(x) = \frac{e^{-\mu} \mu^x}{x!}$
 $x = 0, 1, \dots$ ，其中 μ 為此實驗成功或發生次數之期望值，以 $X \sim P(\mu)$ 表之。其分配函數值亦可由附表二中查出。



➤ 卜瓦松分配具有以下特質：

1. 特定時間或區域與另一段不重疊的時間或區域，成功或發生次數為互相獨立。例如：今、明二天某一十字路口車禍次數為互相獨立，不相關的。
2. 特定時間或區域實驗成功或發生次數的期望值已知為 μ ，則所觀察之另一段時間或區域之期望值會隨其時間長短或區域面積大小與已知之期望值之時間，區域大小成比例增減。
3. 在極短的時間或極小的區域內發生超過一次的機率接近0。



定理6.11

- 卜瓦松分配之平均數為 μ ，變異數亦為 μ 。
- 卜瓦松分配之另一功用為用來逼近二項分配，當所取的實驗次數 n 很大，而成功機率 p 很小時，卜瓦松分配可用來近似二項分配，一般而言，當 $n > 100$ ， $np < 5$ 時， $B(n, p)$ 之機率即可以 $P(\mu)$ 來取代，其中 $\mu = np$ 。





6.12 常態分配

常態分配(normal distribution)為連續性分配中最重要者，其定義如下：

定義6.10

若一連續隨機變數 X 之分配具有以下之機率密度函數，則稱之為常態分配，以 $X \sim N(\mu, \sigma)$ 表之。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$



定理6.12

常態分配之圖形類似一鐘形曲線，中間較高而左右兩尾漸漸向0靠近但不相交，且左右對稱其期望值 μ ，圖形如圖6.9所示。

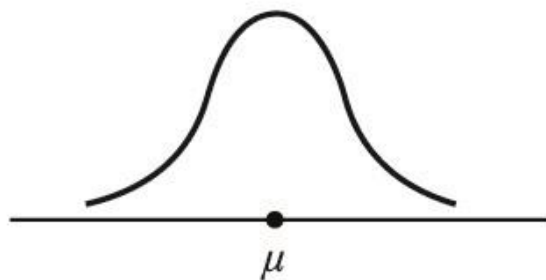


圖6.9 常態分配密度函數圖形

當 μ ， σ 不同時常態分配之圖形亦隨之改變， σ 越小則圖形的中峰越高，而 σ 值越大則圖形越趨扁平，在計算及應用上為求簡化，常常將之標準化，使之成為 $\mu = 0$ ， $\sigma = 1$ 之標準常態分配(standard normal distribution)。

