機　率　論

修訂三版 (網路版)

編著者：顏國勇

國立成功大學數學系

2011 年 9 月 28 日

初　版　序

科學文化的提昇, 不是一蹴可及, 必須付出很大的心力, 歷經很多年代, 才能看出一些端倪; 不是製造幾顆核彈, 發表幾篇論文或者得到某些大獎就代表某個國家如何先進。有二個因素是十分重要的, 其一是按步就班全心投入; 其二是針對當時的需要, 集思廣益, 努力解決。

台灣這些年來各方面的進步是有目共睹, 就因爲如此有很多人產生了更高期望, 恨不能明天一早, 我們就躋身列強之林, 國民所得超過日本, 科學技術趕上英美, 文學藝術不輸德法蘇俄。如果我們平心靜氣探討這些國家是如何達到今日的局面, 我們不難看出,『羅馬不是一天造成的』, 羅馬也不是天才的結晶, 雖然你我都不是天縱奇才, 但我們卻不應該放棄對國家社會的一分關愛和努力, 這就是我寫這本書的動機。

在成大數學系教了十多年的機率論, 用過一些英文教本, 有的很深, 有的容易, 但嚴格說來, 都不可一成不變的教下去, 理由有二, 其一我們的大學生所學的高中數學和大一微積分與美國大學生並不一致; 其二是機率論經過二百多年的演進有不少新東西, 這些新觀念不是一般理工大學生所能領必須有更進一步的工具才能登堂入室享受這些智識。職是之故, 機率論的教學就必須針對以上條件去搜集資料而編成教材; 十多年來, 本人的講義就在此一構想之下, 每年增增改改, 四年前又接下應數所的『高等機率論』課程, 因此深深体認, 大學機率論必須提昇到更高境界, 否則訓練出來的學生將難以勝任研究所的課程。

如今, 講義大致已經定型, 乃決心付印, 本書主要特點如下：

一、以公設化方法界定機率空間, 這是機率的基礎, 然後和古典機率銜接起來, 一方面可以避免要的詭論, 另一方面也不會過分空泛不著邊際;

二、以抽象積分方式界定期望值, 一般初等機率論書籍, 多半利用密度函數來界定期望值, 但因許多分配並無密度函數, 這種定義方法無異畫地自限。我們在第五章中採用以抽象積分方法界定期望值, 此一方法包含了傳統的定義, 也可適應進一步的研究和發展;

三、介紹特徵函數, **Laplace** 變換和 **Fourier** 變換在分析學中極具重要性, 應用于機率論更屬不可或缺, 但以前者所得之動差生成函數, 往往有存在性問題, 而以後者所得之特徵函數, 在第六章中我們可以確認它並沒有不存在的道理, 因此, 其重要性遠非動差生成函數可以比擬;

四、跟隨機率發展史, 介紹收斂理論, 許多學過機率的人對於中央極限定理和大數定律, 往往不能領略其意義和重要性, 本書特別由歷史的角度探討這兩個主題以使讀者有全盤的認識。

本書主要的對象爲大學理工管理學院之學生, 緊接初等微積分的步伐, 以一貫嚴格的數學精神寫成。對於部分讀者或許會稍感深奧, 其實, 清楚的界定, 邏輯的推演, 才是學習數學最佳的途徑。如有誤謬或未盡明白之處, 歡迎來函賜正。

最後, 本人要特別感謝成大數學系林宜禧、陳珍漢以及李育嘉三位教授, 在他們擔任系主任期間, 給予我講授『機率論』這門課的機會。當然更要感謝比利時魯汶大學數學所 **Ballieu** 教授, 由於他, 我才開始認識現代的機率論。

顏　國　勇

1987 年九月於成大數學系

網　路　版　序

本書經過一再修訂, 在 1998 年發行再版, 由於《機率論》屬於較冷門課程, 第一版發行以賠本收場, 第二版則勉強打平。但本於此書對於修課學生及其他機率入門者, 或有些許幫助, 當不計盈虧全力以赴。

不少人在退休之後, 投身志工, 回饋社會。筆者退休數年, 自認應以餘年對於社會略盡棉薄之力, 乃決定重編此書, 而成另類志工。退休前之教學過程中, 發現書中有若干錯誤或不週全之處, 必須修改, 然而原先以軟體 CTEX 編輯而成之檔案, 先天上受限於只能在 DOS 環境工作, 而發行 CTEX 之倚天公司無意將它升級爲 Window 版, 換言之, CTEX 已無法再使用, 筆者唯一的選擇乃改爲以軟體 CwTEX 排版系統重編此書。此二編輯軟體, 前者爲 Plain TEX, 後者則爲 LaTEX, 雖爲近親, 但語法差異頗大, 古稀之年重新學習一種新語法, 辛苦之程度可以想見, 經過數個月的努力, 克服困難, 終於完工。

近年來, 網際網路發展迅速, 本書之修訂三版, 不再以傳統印刷和大家見面, 僅以網路電子書面貌問世, 或可幫助更多各地使用中文之讀者。尚望各界先進不吝指教。

顏　國　勇

2011 年九月於台灣台南

目　錄

本 書 之 主 要 數 學 符 號

數系

$\mathbb{N}$: 自然數系 (不含 0)

$\mathbb{Z}$: 整數系

$\mathbb{Q}$: 有理數系

$\mathbb{R}$: 實數系

$\mathbb{R}_+$: 非負實數集合 (含 0)

$\mathbb{R}^* = \mathbb{R} \setminus \{0\}$

$\mathbb{R}_+^*$: 正實數集合

$\mathbb{C}$: 複數集合

機率符號 (數字表首次出現之頁數)

第一章

第二章

第三章

第四章

Chapter 1

機率概論

太陽下了山, 我們確信它明朝依舊會爬上來; 一棵落了葉的樹木, 我們卻不能肯定它是否會在明年春天再度萌出新芽; 同樣地, 一駕嶄新的飛機也沒有人能保證它的處女航一定平安無事; 我們可以繼續舉出更多的例子, 其中除了很少數的情形之外, 都充滿了不確定性, 固然我們可以感嘆世事無常, 但也可以進一步了解這種無常是否也具有一些『常理』, 嶄新的飛機之處女航百分之九十九以上是安全的, 一枚錢幣丟在桌上雖不一定是正面, 但拋擲了一千次, 總會有一半左右是正面, 機率論是研究這類問題的一個十分重要之基礎工具. 波蘭數學家 M. Fisz 說:『*機率論爲數學的一支, 其目的在於顯示及研究隨機事件之規則性.*』[1] 關於隨機事件一詞, 我們將於隨後詳細說明.

§ 1.1 隨機試驗

所謂**試驗**(experiment) 係指在某些固定條件, 可重複施行的一種程序, 而且可以由此觀察出某些結果者, 試驗通常分爲以下二類:

第一類是**命定試驗**(deterministic experiment), 就是在固定條件下, 我們所觀測之結果乃爲確定者. 例如, 純水在溫度 100°C 及氣壓 760 mm Hg 下, 其結果乃爲沸騰; 又例如在地球上手持一石頭, 手一鬆, 則石頭運動之方向必定垂直於地面.

第二類是**隨機試驗**(random experiment), 就是在固定條件下, 其觀測之結果不爲唯一, 且於試驗實施前不能肯定其結果者, (頂多知道可能結果之集合).

有一種原子筆, 在筆之上端一按, 則筆尖伸出, 再按則筆尖縮回, 這種試驗雖有“進”與“出”二種結果, 但每次試驗之前, 即已知道結果, 故爲命定隨機試驗, 而非隨機試驗, 本課程雖以探討隨機試驗爲主, 但將必然性 (即機率等於一) 之情況排除在外, 亦不適宜, 採取較爲寬廣之研究範圍實屬必要.

以下我們將先舉一些隨機試驗之例, 以加深讀者之印象, 爲方便計, 我們將以 $\mathcal{E}$ 表示隨機試驗.

[1] Probability theory is a part of Mathematics which is usefull in discovering and investigating the regular features of random events.

$\mathcal{E}_1$: 擲一骰子, 而觀察其 (頂面) 出現之點數;

$\mathcal{E}_2$: 擲一錢幣, 而觀察其出現爲正面或反面;

$\mathcal{E}_3$: 擲一錢幣四次, 而觀察其正反兩面之次序;

$\mathcal{E}_4$: 觀察某大醫院之產房, 每位產婦生產之嬰兒爲男或爲女或其他 (包括畸形或死胎); 而觀察其壽命爲若干小時;

$\mathcal{E}_5$: 某一日光燈工廠, 隨機取出一支燈管點亮之, 而觀察其壽命爲若干小時;

$\mathcal{E}_6$: 觀察某一地點, 每日最高及最低溫度;

$\mathcal{E}_7$: 在區間 $[0,1]$ 中, 隨機抽取一數.

定義 1.1

對於某一隨機試驗 $\mathcal{E}$, 所有可能結果 (我們所欲觀察者) 所成之集合稱爲一**樣本空間**(sample space), 並以 Ω 表之.

§ 1.2 機率論的歷史及古典機率

談到機率論, 我們很容易聯想到, 它與骰子、紙牌及硬幣等賭博遊戲有關, 但它之所以能成爲一門學問, 其中因素之一固然賭博有關, 然而如同許多學科一樣, 其發展之原動力, 經濟因素才是主因, 不少學者深信乃是資本家渴望從學術上求得比占星術更爲可靠的指導. 儘管有人指出, 在十五、六世紀一些有關機率問題的計算曾出現在義大利數學家 Cardano 等的論著之中, 但比較廣泛的機率問題解法, 則出現在法國數學家 B. Pascal 與 P. de Fermat 始於 1654 年著名的通信上. 隨後, 荷蘭數學家 C. Huygens (1629–1695) 發表了有關機率的第一本書《De Ratiociniis in Ludo Aleae》(機會遊戲之計算, 1657), 而眞正使機率成爲一門學問, 則始於瑞士的 Jakob Bernoulli [2] (1654–1705), 在他死後才出版的《Ars Conjectandi》(猜測的藝術, 1713), 以相當嚴格的數學寫出有關機率論中第一個極限定理 —— 大數定律 —— (詳見第七章), 由於 Bernoulli 所提出二項分配中, $\binom{n}{x}p^xq^{n-x}$ 的計算十分麻煩, 法國數學家 A. de Moivre(1667–1754) 在《Miscellanea Analytica Supplementum》(分析方法, 1730) 一文中, 首先嘗試以分析方法解決 $\binom{n}{x}p^xq^{n-x}$, 當 $p=1/2$ 時, 之近似值, 這是中央極限定理之起源. 獨立、條件機率及數學期望值都是 de Moivre 所創立的概念. 法國數學大師 P.-S. Laplace (1749–1827) 綜合各家研究成果而完成《 Théorie Analytique des Probabilités》(機率之解析論, 1812) 一書中, 更將 de Moivre 的中央極限定理推廣而適用於一般的 p 值. 此外, 他也將機率方法應用於誤差理論之中. 十九世紀的機率學者, 多半關注於機率的極限理論之上, 其中法國人 S. D. Poisson (1781–1840) 於 1837 發表了有關 Poisson 分配, 以另外一種方法解決二項分配近似值之問題.

[2] 英文名爲 James Bernoulli, 法文名爲 Jaques Bernoulli.

德國數學家 K. F. Gauss (1777–1855) 則利用常態分配函數觀念開創了誤差理論研究的新方向及最小平方法的基礎. 以法國數學家爲中心的機率論研究, 在十九世紀末葉開始轉移到俄國. P. L. Chebyshev (1821–1894), A. A. Lyapunov (1857–1918) 等致力於獨立隨機變數之和的極限問題, 也就是所謂大數定律以及中央極限定理的研究. 其中的 Markov 更創立了, 近代機率論中十分重要的 Markov 鏈.

微積分發明之後的數學界, 很多數學家將《分析學》做了十分穩固的扎根工作, 於是機率論便由有限個數的組合問題, 推廣而爲以線段之長短及面積之大小爲基礎的幾何機率論, 以分析的方法注入機率論的基礎, 帶來這門學問更爲嚴密的骨架, 然而一些似是而非的矛盾說法, 如 Bertrand 的詭論, 卻不時困擾著這門學問, 二十世紀初 E. Borel (1871–1956) 在測度論上的新理論, 促使俄國大數學家 A. Kolmogorov 於 1933 年宣布以公設法做爲機率論的基石, 奠定了機率論的數學基礎.

1920 年代以後, 除了有些數家如 P. Lévy 及 W. Feller 在中央極限定理有很好的成果, A. Kolmogorov 的強大數定律令人十分激賞之外, 部分學者則以 Markov 鏈開創機率論研究的新方向, 探討非獨立隨機變數之隨機過程. 二次大戰之後由於純粹數學在賦範空間 (normed space) 理論上的進展, 機率學者開始將這門學問帶進抽象的領域.

現在讓我們來看看什麼叫做機率.
Laplace 在《機率的解析論》中, 爲古典機率下了以下的定義:

設某隨機試驗之樣本空間 Ω 有 n 個元素, 且每一元素出現之機會均相等, 若 A 爲 Ω 之一子集, (又稱 A 爲一隨機事件), 則事件 A 發生之機率爲:

$$P(A) = \frac{\#A}{n} = \frac{\#A}{\#\Omega},$$

內 $\#A$ 表 A 之基數 (cardinal number of A).

在古典機率中, 以這種方法計算機率十分常見, 分數中的分母 $\#\Omega$ 又稱爲總方法數 (total number of possible outcomes), 分子則稱爲 A 發生之可能方法數 (number of ways event A can occur).

例 1. 一骰子連擲兩次, 試問二次點數不同之機率爲若干?

$\boxed{\text{解}}$ 此一試驗之樣本空間爲

$$\Omega = \{(a, b) \mid a, b \in \{1, 2, 3, 4, 5, 6\}\} = \{(1, 1), (1, 2), \cdots, (6, 6)\}.$$

顯然 $\#\Omega = 36$, 而事件“二次點數不同”應爲

$$A = \{(a, b) \in \Omega \mid a \neq b\} = \Omega \setminus \{(1, 1), (2, 2), \cdots, (6, 6)\}.$$

由於樣本空間 Ω 中每一序對出現之機會均爲相等, 故知二次點數不同之機率爲

$$P(A) = \frac{\#A}{\#\Omega} = \frac{30}{36} = \frac{5}{6}.$$

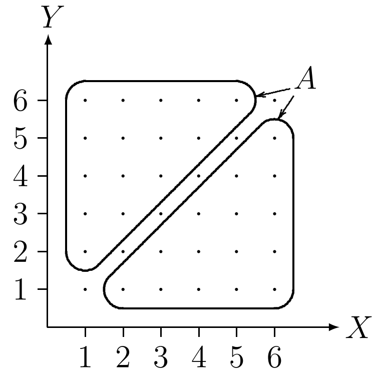


圖 1–1 □

例 2. (de Méré 的詭論) 法國人 de Méré 在很多次賭博中, 發現擲三個骰子得 11 點之機會較得 12 點爲高, 但 11 點之組成有 (以下每一組括號視爲一組合)

(1-4-6), (1-5-5), (2-3-6), (2-4-5), (3-3-5), (3-4-4)

等; 而 12 點之組成有

(1-5-6), (2-4-6), (2-5-5), (3-3-6), (3-4-5), (4-4-4)

等, 二者均有六種情形, 機率應爲相等, 他將此一矛盾情形就教於 B. Pascal. Pascal 以"排列"的關念解釋說:(4-4-4) 只有一種種可能就是第一、第二及第三個骰子都出現 4 點, 表爲 $(4, 4, 4)$. 但 (3-4-5) 則有以下六種可能 (排列)

$$(3, 4, 5),\ (3, 5, 4),\ (4, 3, 5),\ (4, 5, 3),\ (5, 3, 4),\ (5, 4, 3),$$

因此, 得 11 點之機率爲

$$P_1 = \frac{6+3+6+6+3+3}{6\times 6\times 6} = \frac{27}{216},$$

得 12 點之機率爲

$$P_2 = \frac{6+6+3+3+6+1}{6\times 6\times 6} = \frac{25}{216},$$

前者確較後者爲大, 此與實驗結果並無二致. □

註: 以抽樣方式討論機率問題是古典機率十分常見的方法, 我們知道排列組合共有四種: 重複排列、重複組合、不重複排列、不重複組合(後二種簡稱排列與組合, 參見附錄一), 其中重複組合通常不宜用以討論機率問題 , 最簡單的例子是丟二錢幣, 若以重複組合視之, 有『二正』、『一正一反』、『二反』三種情形, 但其中出現『二正』或『二反』之機率皆爲 1/4, 但出現『一正一反』之機率爲 1/2. de Méré 以重複組合處理機率問題, 擲三個骰子共有 $H_{6,3} = \binom{6+3-1}{3} = 56$ 種情形, 他提出的三骰共計 11 點和 12 點各有六種情形, 其誤謬乃由於以重複組合解決此類問題所產生.

在 Laplace 的定義中, 對於每一元素出現機會不完全相之情形並不適用, 因此有人略加推廣而得以下定義:

> 設某隨機試驗之樣本空間 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$, 若基本事件 $\{\omega_j\}$ 發生之機率為 p_j, $(p_1 + p_2 + \cdots + p_n = 1)$. 則事件 $A = \{\omega_{j_1}, \omega_{j_2}, \cdots, \omega_{j_k}\}$ 發生之機率為: $P(A) = p_{j_1} + p_{j_2} + \cdots + p_{j_k}$.

上述定義內, 每一基本事件 $\{\omega_j\}$ 之機率在事前既已確定, 因此稱為**事前機率**或**主觀機率** (a priori probability). 然而事前機率如何知悉卻為一大困難, 解決此一問題, Von Mises 乃於 1920 年利用 J. Bernoulli 的相對頻率觀念, 提出所謂試驗機率的觀念:

> 設某一隨機試驗重覆了 n 次, 而事件 A 發生之次數為 k, 則 $\frac{k}{n}$ 稱為事 A 之**相對頻率** (relative frequency). 如果相對頻率之極限 $\lim\limits_{n\to+\infty} \frac{k}{n}$ 存在, 則以此為事件 A 之**機率**.

上述觀念, 對於某些樣本空間為無限的情況, 可加以利用, 如:

例 3. 從自然數中, 隨機任取一數, 試求取出之數為 7 之倍數之機率為若干?

解 此試驗之樣本空間為$\Omega = \mathbb{N}$, 而事件 "取出之數為$\mathbb{Z}7\mathbb{Z}$之倍數" 為

$$A = \{7n \mid n \in \mathbb{N}\} = \{7, 14, 21, \cdots\},$$

先考慮自集合 $\{1, 2, \cdots, n\}$ 中隨機抽取一數, 由餘數定理知,

$$\exists k(n) \in \{0, 1, \cdots, n\}, \exists r \in \{0, 1, 2, 3, 4, 5, 6\},$$

使得 $n = 7k(n) + r$, 內 r 為餘數. 是以"得數為 7 之倍數"之相對頻率為

$$\frac{k(n)}{n} = \frac{\frac{n-r}{7}}{n} = \frac{1}{7}\Big(1 - \frac{r}{n}\Big),$$

則其極限顯然為

$$\lim_{n\to+\infty} \frac{k(n)}{n} = \lim_{n\to+\infty} \frac{1}{7}\Big(1 - \frac{r}{n}\Big) = \frac{1}{7}.$$

故從自然數中隨機任取一數而得 7 之倍數之機率為 1/7. □

多數的情形不像上例一樣, 可經由 n 趨近於正無限大而求得 A之發生之機率. 當然, 一般而言, 當 n 越大, 相對頻率與 A 發生之機率也越接近 (參閱第七章大數定律), 但究竟應大到什麼程度, 實無一定法則可資遵循.

由上述古典機率的觀念可獲得以下的性質:

1° 機率乃樣本空間子集合之函數. 意即 $P\colon \mathcal{P}(\Omega) \to \mathbb{R}$, 內 $\mathcal{P}(\Omega)$ 表 Ω 之冪集;

2° 一事件 (即 Ω 之子集) 之機率不得爲負, 不得大於 1 ;

3° 整個樣本空間 Ω 發生之機率爲 1 ;

4° 若 A 與 B 爲 Ω 之二互斥子集, 則 $P(A \cup B) = P(A) + P(B)$.

此外尙有其他很多性質, 但均可由以上四項一一推導而得, 於是我們可以公設化的方法規定 (古典模式之) 機率函數如下:

設 $P\colon \mathcal{P}(\Omega) \to [0, 1]$ 滿足:

(1) $P(\Omega) = 1$;

(2) $A, B \subset \Omega$ 且 $A \cap B = \varnothing$, 則 $P(A \cup B) = P(A) + P(B)$.

則稱 P 爲一**機率測度**, 而 $P(A)$ 則稱爲事件 A (發生) 之**機率**.

在上述定義中, Ω 之任一子集必須對應介於 0 與 1 之間之一實數, 爲此子集之機率, 但有時我們只關心樣本空間 Ω 中某些子集 (而非每一子集) 之機率.

例如, 從罹患肺癌之統計資料顯示:(以下數字均爲虛構)

男性抽菸者佔 70%,
女性抽菸者佔 12%,
男性不抽菸者佔 15%,

爲方便計, 令

$\Omega =$ 樣本空間 $=$ 全部罹患肺癌者集合,
$A = \Omega$ 中之男性抽菸者,
$B = \Omega$ 中之女性抽菸者,
$C = \Omega$ 中之男性不抽菸者,

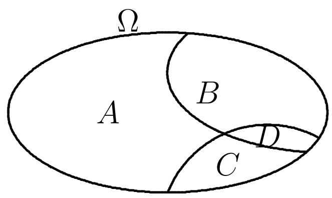


圖 1–2

則我們知道以下諸機率 (比率):

$$
\begin{aligned}
&P(A) = 0.7,\ P(B) = 0.12,\ P(C) = 0.15, \\
&P(A \cup B) = 0.7 + 0.12 = 0.82, (\Omega \text{ 中抽菸者之機率}) \\
&P(A \cup C) = 0.7 + 0.15 = 0.85, (\Omega \text{ 中男性之機率}) \\
&P(B \cup C) = 0.12 + 0.15 = 0.27, \\
&P(A \cup B \cup C) = 0.7 + 0.12 + 0.15 = 0.97, \\
&P(\Omega) = 1.
\end{aligned}
$$

雖然在原始資料中未曾列出 $D=$ “Ω 中之女性不抽菸者”之機率, 但因 A, B, C, D 爲兩兩互斥且 $\Omega = A\cup B\cup C\cup D$, 故

$$P(D)=1-P(A)-P(B)-P(C)=0.03.$$

此外, 我們也很容易求出 $P(A\cup D), P(B\cup D), P(C\cup D), P(A\cup B\cup D), P(B\cup C\cup D)$, $P(A\cup C\cup D)$ 等.

就以上資料而言, 我們不知道 (也不關心) 罹患肺癌之俄國人之機率爲多少. 這正說明: 機率測度 P 不需要也不能夠以 Ω 之冪集 $\mathcal{P}(\Omega)$ 爲其定義域, 我們僅需以

$$\{\varnothing,\, A,\, B,\, C,\, D,\, A\cup B,\, A\cup C,\, A\cup D,\, B\cup C,\, B\cup D,\, C\cup D,\, A\cup B\cup C,\\ A\cup B\cup D,\, A\cup C\cup D,\, B\cup C\cup D,\, \Omega\}$$

做爲 P 之定義域, 由此看來做爲 P 之定義域 $\mathcal{F}$ 須滿足以下條件:

1° $\mathcal{F}\subset\mathcal{P}(\Omega)$;

2° $\mathcal{F}\neq\varnothing$;

3° 若 $A\in\mathcal{F}$, 則 $\complement A\in\mathcal{F}$, (其中 $\complement A=\Omega\setminus A$ 乃 A 之餘集, 亦常寫爲 A^c) ;

4° $A,\, B\in\mathcal{F} \;\Rightarrow\; A\cup B\in\mathcal{F}$;

5° $A,\, B\in\mathcal{F} \;\Rightarrow\; A\cap B\in\mathcal{F}$;

但條件 5° 可由條件 3° 及 4° 蘊涵而得, 因爲

$$\begin{aligned} A,\, B\in\mathcal{F} &\;\Rightarrow\; \complement A,\, \complement B\in\mathcal{F}\\ &\;\Rightarrow\; \complement A\cup\complement B\in\mathcal{F}\\ &\;\Rightarrow\; A\cap B=\complement(\complement A\cup\complement B)\in\mathcal{F}. \end{aligned}$$

因此 5° 可以省略, 換言之, $\mathcal{F}$ 對於差集、聯集、交集均有封閉性. 我們規定如下:

定義 1.2

設 Ω 爲一非空集合, (暫不討論隨機試驗及其樣本空間), 其冪集合 $\mathcal{P}(\Omega)$ 之子集 $\mathcal{F}$ 若滿足以下三條件, 則稱其爲 Ω 之一**域**(field).

(1) $\mathcal{F}\neq\varnothing$;

(2) $A\in\mathcal{F} \;\Rightarrow\; \complement A\in\mathcal{F}$;

(3) $A,\, B\in\mathcal{F} \;\Rightarrow\; A\cup B\in\mathcal{F}$.

由於微積分的發明, 機率中很早就引進了幾何機率的概念, 幾何機率簡單的說: 就是以線的長短, 面積體積的大小來研究機率.

例 4. 自區間 [0, 1] 中隨機取一數, 試求所得之數出現下列集合之機率各若干?

$$A = \left[0,\ \frac{1}{2}\right],\ B = \left(\frac{1}{2},\ 1\right],\ C = [0,\ 1] \cap \mathbb{Q}.$$

$\boxed{\text{解}}$ (1) 由於樣本空間 $\Omega = [0,\ 1]$ 之長度爲 1, 而集合 A 乃爲事件 “ 所得之數出現於區間 $[0,\ \frac{1}{2}]$ ”, 我們應以 A 之長度, 即 $\frac{1}{2}$, 爲其機率.

(2) 至於 B, 我們以二種不同方法考慮.

法一: 由於 $A \cap B = \varnothing$, $A \cup B = \Omega$, 故應有

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

是以 $P(B) = 1 - \frac{1}{2} = \frac{1}{2}$.

法二: 我們直接研究 B 之長度, 由於區間 $(\frac{1}{2},\ 1]$ 可視爲序列

$$\left[\frac{1}{2} + \frac{1}{3},\ 1\right],\ \left[\frac{1}{2} + \frac{1}{4},\ 1\right],\ \cdots,\ \left[\frac{1}{2} + \frac{1}{n},\ 1\right], \cdots,$$

之極限, 而上述各閉區間之長度分別爲

$$\frac{1}{2} - \frac{1}{3},\ \frac{1}{2} - \frac{1}{4},\ \cdots,\ \frac{1}{2} - \frac{1}{n},\ \cdots,$$

此序列之極限顯然爲 $\frac{1}{2}$, 亦即我們可以 $\frac{1}{2}$ 做爲 B 之長度, 換言之, 所得之數出現在 B 之機率爲 $\frac{1}{2}$.

(3) 至於 $C = [0,\ 1] \cap \mathbb{Q}$, 有人認爲 C 之長度無法求出 (或無長度可言), 在十九世紀末, 法國數學家 Lebesgue 以較進步的觀念提出上述集合之測度 (measure) 觀念, (參閱 Royden [15]). 其中可證得零測度集合之可數聯集亦爲零測度; 已知單點集合爲零測度, 而 C 爲可數, 故 C 之長度爲 0, 即抽得之數出現在集合 C 之機率爲 0. □

由上例中我們發現一集合序列 $\{A_n\}_n$, 如果每一集合 A_n 皆有機率, 則其可數聯集 (及可數交集) 亦應有機率, 因此以“域” 做爲機率之定義域是不夠的. 此外在 Lebesgue 的觀念裡, 並非 $[0, 1]$ 之任一子集均爲可測 (measurable), 也就是說有些集合並沒有長度可言, 因此不能有機率的觀念, 本世紀初 E. Borel 以區間之可數聯集、可數交集、差集等研究實數系 $\mathbb{R}$ 的可測集合 (詳見第三節). 根據此一構想, 機率測度 P 之定義域顯然仍需加以修正; 這正是 Kolmogorov 在 1933 年所提公設化機率論的數學基礎.

在沒有介紹 Kolmogorov 公設化定義之前, 我們先看看幾何機率中著名的詭論之一 ——— Bertrand 的詭論. 這個詭論乃是說明, 解決幾何機率問題, 以直觀概念已不足以應付, 必須另行設法才行.

例 2. 設一等邊三角形, 邊長爲 a, 外接一圓, 於圓內隨機作一弦, 試問弦長 l 大於 a 之機率爲若干?(參閱圖 1–3).

$\boxed{解}$ (1) **平行法**: 令外接圓之半徑爲 r, 並設所做之弦均平行於 Y 軸, 若此弦之中點在 X 軸上之區間 $(-\frac{r}{2}, \frac{r}{2})$, 則弦長 l 大於 a, 故其機率爲 $\frac{1}{2}$.

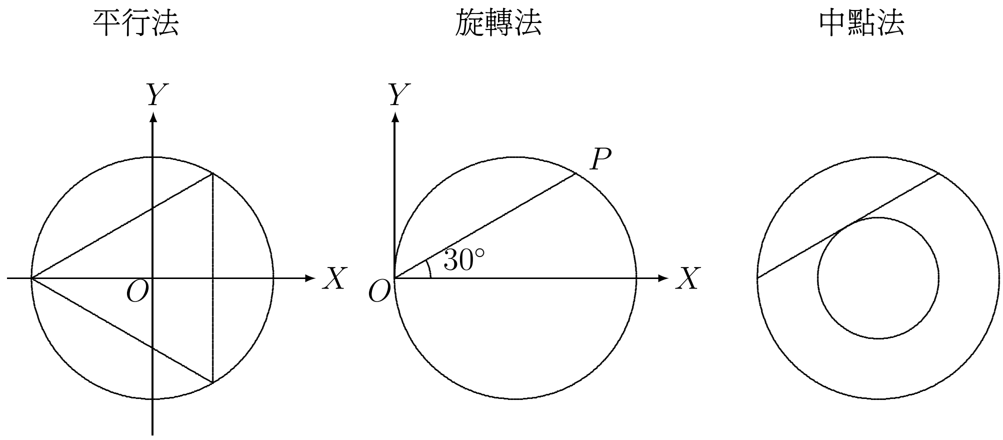


圖 1–3

(2) **旋轉法**: 設弦之一端固定於原點 O, 若弦 $\overline{OP}$ 與 X 軸之夾角 θ 小於 $30°$ 時 , 則弦長 l 大於 a, 故其機率爲 $\frac{1}{3}$.

(3) **中點法**: 做二同心圓, 外圓半徑爲 r, 內圓半徑爲 $\frac{r}{2}$, 若弦之中點落在內圓之內部, 則弦長 l 大於 a, 是以其 機率爲 $\frac{1}{4}$, (內圓面積爲全部面積的 $\frac{1}{4}$). □

§ 1.3 σ 域

在上節中我們業已發現以“域” 做爲機率之定義域是不夠的, 必須加以調整, 好使它對於集合之“可數運算”具封閉性, 此乃本節之目的.

定義 1.3

設 Ω 爲一非空集合. 冪集合 $\mathcal{P}(\Omega)$ 之子集 $\mathcal{F}$ 若滿足以下條件, 則稱其爲 Ω 上一 σ **域** (σ-field 亦稱 σ-ring 或 σ-algebra):

(F1) $\mathcal{F} \neq \varnothing$;

(F2) 若 $A \in \mathcal{F}$, 則 $\complement A \in \mathcal{F}$;

(F3) 若 $\forall j \in \mathbb{N}$, $A_j \in \mathcal{F}$, 則 $\bigcup_{j\in\mathbb{N}} A_j \in \mathcal{F}$; (即 $\mathcal{F}$ 對可數聯集具封閉性).

我們將先探討有關 σ 域的一些性質, 其次再舉些相關的範例.

定理 1.4

設 $\mathcal{F}$ 爲 Ω 上之一 σ 域, 則

(1) $\varnothing, \Omega \in \mathcal{F}$;

(2) 若 $A_1, \cdots, A_n \in \mathcal{F}$, 則 $\bigcup_{j=1}^{n} A_j \in \mathcal{F}$, 且 $\displaystyle\bigcap_{j=1}^{n} A_j \in \mathcal{F}$, (即 $\mathcal{F}$ 對於有限交集及有限聯集具封閉性) ;

(3) 若 $\forall j \in \mathbb{N}$, $A_j \in \mathcal{F}$, 則 $\displaystyle\bigcap_{j\in\mathbb{N}} A_j \in \mathcal{F}$, (即 $\mathcal{F}$ 對可數交集具封閉性) ;

(4) 若 $A, B \in \mathcal{F}$, 則 $A \setminus B \in \mathcal{F}$, (亦即 $\mathcal{F}$ 對差集具封閉性).

證 (1) 由 (F1) 知 $\mathcal{F}$ 中必有元素, 令 $A \in \mathcal{F}$. 又由 (F2) 知 $\complement A \in \mathcal{F}$, 今令

$$A_1 = A, \quad A_2 = A_3 = \cdots = \complement A,$$

則由 (F3) 知,

$$\Omega = A \cup \complement A = \bigcup_{j=1}^{+\infty} A_j \in \mathcal{F}.$$

此外空集合 $\varnothing$ 乃 Ω 之餘集, 是以 $\varnothing \in \mathcal{F}$.

(2) 令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$, 則由 (F3) 知

$$\bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{+\infty} A_j \in \mathcal{F},$$

其次

$$\begin{aligned} A_1, \cdots, A_n \in \mathcal{F} &\Rightarrow \complement A_1, \cdots, \complement A_n \in \mathcal{F} \\ &\Rightarrow \bigcup_{j=1}^{n} (\complement A_j) \in \mathcal{F}, \ (\text{由 } (1)) \\ &\Rightarrow \bigcap_{j=1}^{n} A_j = \complement\Big(\bigcup_{j=1}^{n} (\complement A_j)\Big) \in \mathcal{F}, (\text{De Morgan 定律.}) \end{aligned}$$

(3) 由 De Morgan 定律知,

$$\bigcap_{j\in\mathbb{N}} A_j = \complement\Big(\bigcup_{j\in\mathbb{N}} (\complement A_j)\Big),$$

仿 (2) 證明之後半部知其必屬於 $\mathcal{F}$.

(4) 因 $A \setminus B = A \cap \complement B$ 之故. □

例 1. $\mathcal{F}=\{\varnothing,\Omega\}$ 乃爲 Ω 上之一 σ 域.

例 2. $\mathcal{F}=\mathcal{P}(\Omega)$ 亦爲 Ω 上之一 σ 域.

例 3. 若 A 爲 Ω 之一非空眞子集, 則

$$\mathcal{F}=\{\varnothing,\ A,\ \complement A,\ \Omega\}$$

亦爲 Ω 上之一 σ 域.

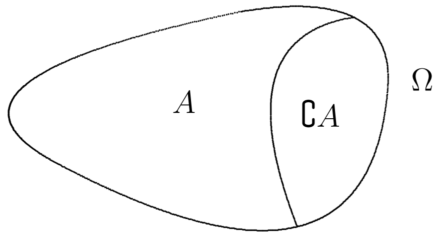


圖 1–4

例 4. 設 Ω 爲一不可數集合, 試證:

$$\mathcal{F}=\{A\subset\Omega \mid A \text{ 爲至多可數[3] 或 } \complement A \text{ 爲至多可數}\}$$

爲 Ω 上一 σ 域.

$\boxed{\text{證}}$ 因 $\mathcal{F}\subset\mathcal{P}(\Omega)$ 且滿足

(F1) $\mathcal{F}\neq\varnothing$, 此因 $\varnothing$ 顯爲 $\mathcal{F}$ 之元素.

(F2) 若 $A\in\mathcal{F}$, 則 A 爲至多可數或 $\complement A$ 爲至多可數, 是以 $\complement A\in\mathcal{F}$.

(F3) 設 $\forall j\in\mathbb{N}$, $A_j\in\mathcal{F}$,

1° 若每一 A_j 均爲至多可數, 則 $\bigcup_{j=1}^{+\infty}A_j$ 亦必爲至多可數, 故必屬於 $\mathcal{F}$.

2° 若存在 A_k 不爲至多可數, 即不可數, 則 $\complement A_k$ 必爲至多可數, 此時因

$$\complement\Big(\bigcup_{j\in\mathbb{N}}A_j\Big)=\bigcap_{j\in\mathbb{N}}\complement A_j\subset\complement A_k,$$

是以 $\complement\Big(\bigcup_{j\in\mathbb{N}}A_j\Big)$ 必爲至多可數, 故 $\bigcup_{j\in\mathbb{N}}A_j\in\mathcal{F}$. □

以上四例均爲 σ 域, 當然亦爲域; 但非所有的域均爲 σ 域.

例 5. 設 Ω 爲一可數集合, (例如 $\Omega=\mathbb{N}$), 則

$$\mathcal{F}=\{A\subset\Omega \mid A \text{ 爲有限或 } \complement A \text{ 爲有限 }\}$$

爲 Ω 上一域, 但非 σ 域, (讀者自證之). □

以下我們將研究如何將 $\mathcal{P}(\Omega)$ 之一子集 $\mathcal{C}$ 產生一 σ 域 ,

[3]至多可數 (at most denumerable) 係指有限或無限可數, 有些學者稱其爲 countable.

定理 1.5

設 $\mathcal{C}$ 爲 $\mathcal{P}(\Omega)$ 之一子集, 則

$$\mathcal{F} \underset{\text{def}}{=} \bigcap \{\mathcal{F}_j \mid \mathcal{F}_j \text{ 爲 } \Omega \text{ 上之 } \sigma \text{ 域且 } \mathcal{F}_j \supset \mathcal{C}\}$$

爲 Ω 上之一 σ 域, 顯然 $\mathcal{F}$ 爲包含 $\mathcal{C}$ 之最小 σ 域.

證 因爲

(F1) 各 $\mathcal{F}_j$ 中均有 $\varnothing$ 及 Ω 二元素, 故其交集 $\mathcal{F}$ 亦然, 是以 $\mathcal{F} \neq \varnothing$.

(F2) 由於

$$\begin{aligned} A \in \mathcal{F} &\Rightarrow (\forall j,\ A \in \mathcal{F}_j) \\ &\Rightarrow (\forall j,\ \complement A \in \mathcal{F}_j) \\ &\Rightarrow \complement A \in \bigcap_j \mathcal{F}_j = \mathcal{F}. \end{aligned}$$

(F3) 由於

$$\begin{aligned} A_1,\ A_2,\ \cdots,\ A_k,\ \cdots \in \mathcal{F} &\Rightarrow A_1, A_2, \cdots, A_k, \cdots \in \mathcal{F}_j,\ \forall j \\ &\Rightarrow \bigcup_{k=1}^{+\infty} A_k \in \mathcal{F}_j,\ \forall j \\ &\Rightarrow \bigcup_{k=1}^{+\infty} A_k \in \bigcap_j \mathcal{F}_j = \mathcal{F}. \end{aligned}$$ □

註: 二個以上之 σ 域之聯集, 則未必爲一 σ 域. (習題).

定義 1.6

設 $\mathcal{C} \subset \mathcal{P}(\Omega)$, 則上述 σ 域 $\mathcal{F}$ 稱爲 $\mathcal{C}$ **所產生之 σ 域** (σ-field generated by $\mathcal{C}$), 並記爲 $\sigma(\mathcal{C})$.

例 6. 若 A 爲 Ω 之一非空眞子集, 則 $\sigma(\{A\}) = \{\varnothing,\ A,\ \complement A,\ \Omega\}$ (參見例 3).

定義 1.7

設 $\mathcal{F}$ 爲 Ω 上 σ 一域, 則稱序對 $(\Omega, \mathcal{F})$ 爲一**可測空間**(measurable space).

註: 如果沒有特別聲明, 當樣本空間爲有限或可數時, 我們將取 $\mathcal{P}(\Omega)$ 爲其 σ 域, 當樣本空間爲不可數時, 則取較 $\mathcal{P}(\Omega)$ 爲小之 σ 域, 但將予以說明.

定義 1.8

設 $(\Omega, \mathcal{F})$ 爲一可測空間, A 爲 Ω 之一非空子集, 令

$$\mathcal{F}_A = \{B \cap A \mid B \in \mathcal{F}\},$$

(顯然 $\mathcal{F}_A$ 爲集合 A 上一 σ 域), 我們稱 $(A, \mathcal{F}_A)$ 爲 $(\Omega, \mathcal{F}_A)$ 之一**子可測空間**(measurable sub-space).

重要之可測空間

(1) **一維 Borel 域**

設 $\mathcal{C}_0 = \{I \mid I$ 爲 $\mathbb{R}$ 之子區間 $\} \subset \mathcal{P}(\mathbb{R})$, 則 $\mathcal{C}_0$ 所產生之 σ 域 $\sigma(\mathcal{C}_0)$ 稱爲**一維Borel域** (one-dimensional Borel field), 並以 $\mathcal{B}$ 表之.

區間、單元集合、有限集合、有限個區間之聯集、可數個區間之聯集 、以上各集合之餘集等等均爲 $\mathcal{B}$ 之元素, 我們稱其爲 Borel 可測集合, 與 Lebesgue 可測集合有微小差別, 此外 $\mathcal{B}$ 不等於 $\mathcal{P}(\mathbb{R})$, (參閱附錄二).

(2) **二維 Borel 域**

設 $\mathcal{C}_0 = \{I \times J \mid I, J$ 爲 $\mathbb{R}$ 之子區間 $\} \subset \mathcal{P}(\mathbb{R}^2)$, 則 $\mathcal{C}_0$ 所產生之 σ 域 $\sigma(\mathcal{C}_0)$ 稱爲**二維Borel域** (two-dimensional Borel field), 並以 $\mathcal{B}^2$ 表之.

在實變函數論中, 我們可以證明: $\mathbb{R}^2$ 之任一開集合皆可表爲可數個 $\mathcal{C}_0$ 元素之聯集[4], 因此, $\mathbb{R}^2$ 上之多邊形、圓形、開集合、閉集合及其他各類平面圖 (只要用筆畫得出來) 之有限或可數聯集、交集、餘集等均爲 $\mathcal{B}^2$ 之元素.

(3) **k 維 Borel 域**

設 $\mathcal{C}_0 = \{I_1 \times I_2 \times \cdots \times I_k \mid I_1, \cdots, I_k$ 爲 $\mathbb{R}$ 之子區間 $\} \subset \mathcal{P}(\mathbb{R}^k)$, 則 $\mathcal{C}_0$ 所產生之 σ 域 $\sigma(\mathcal{C}_0)$ 稱爲 **k 維Borel 域** (k-dimensional Borel field), 並以 $\mathcal{B}^k$ 表之.

[4] Wheeden & Zygmund, Measure and Integration. Theorem 1.11

定義 1.9

設 $\{A_n\}_n$ 爲一集合序列,

(1) 我們稱集合 $\bigcap_{n\geq 1}\bigcup_{k\geq n} A_k$ 爲序列 $\{A_n\}_n$ 之**上極限**(upper limit), 並且表爲 $\limsup_{n\to+\infty} A_n$, 或 $\overline{\lim}_{n\to+\infty}A_n$, (無誤解之可能時簡爲 $\limsup A_n$);

(2) 我們稱集合 $\bigcup_{n\geq 1}\bigcap_{k\geq n} A_k$ 爲序列 $\{A_n\}_n$ 之 **下極限**(lower limit), 並且表爲 $\liminf_{n\to+\infty} A_n$, 或 $\underline{\lim}_{n\to+\infty}A_n$, (亦可簡爲 $\liminf A_n$);

(3) 若 $\{A_n\}_n$ 之 上下極限相等, 則稱 $\{A_n\}_n$ 之極限存在, 並且稱此集合爲 $\{A_n\}_n$ 之 **極限**, 而記爲 $\lim_{n\to+\infty} A_n$, 或 $\lim A_n$, 亦即

$$\lim_{n\to+\infty} A_n = \limsup_{n\to+\infty} A_n = \liminf_{n\to+\infty} A_n.$$

(4) 若 $\{A_n\}_n$ 之 上下極限不相等, 則稱 $\{A_n\}_n$ 之極限不存在.

對於實數序列 $\{a_n\}_n$, 在《微積分》中, 我們討論了它的『極限』觀念, 而在《高等微積分》中, 更進一步研究其上、下極限: $\limsup_{n\to+\infty} a_n$, $\liminf_{n\to+\infty} a_n$. 機率論中則是模仿分析學之觀念而有所謂“集合之上、下極限”, 對於機率之理論探討頗爲重要. 然而上述定義, 對於初學者而言過份深奧難懂, 有加以簡化之必要, 以下定理除了幫助我們了解上、下極限之意義外, 尚可賴此解決一些問題.

定理 1.10

設 $\{A_n\}_n$ 爲一集合序列, 則

(1) $\limsup A_n = \{\omega \mid \omega\in A_n \text{ i.o.}\}$, (內 i.o. 係 infinitely often 之縮寫, $\omega\in A_n$ i.o. 表示 ω 屬於無限許多個 A_n .)

(2) $\liminf A_n = \{\omega \mid \omega\in A_n \text{ a.a.}\}$, (內 a.a. 係 almost all 之縮寫, $\omega\in A_n$ a.a. 表示 ω 屬於每一個 A_n 除了有限個.)

(3) $\liminf A_n \subset \limsup A_n$.

證 (1) 由定義知,

$$\begin{aligned}\omega \in \limsup A_n &\Leftrightarrow \omega \in \bigcap_{n\geq 1}\bigcup_{k\geq n} A_k\\ &\Leftrightarrow \forall n \in \mathbb{N},\ \omega \in \bigcup_{k\geq n} A_k\\ &\Leftrightarrow \forall n \in \mathbb{N},\ \exists k \geq n,\ \omega \in A_k\\ &\Leftrightarrow \omega \in A_n \text{ i.o.}\end{aligned}$$

(2) 由於,

$$\begin{aligned}\omega \in \liminf A_n &\Leftrightarrow \omega \in \bigcup_{n\geq 1}\bigcap_{k\geq n} A_k\\ &\Leftrightarrow \exists n \in \mathbb{N},\ \omega \in \bigcap_{k\geq n} A_k\\ &\Leftrightarrow \exists n \in \mathbb{N},\ \forall k \geq n,\ \omega \in A_k\\ &\Leftrightarrow \omega \in A_n \text{ a.a.}\end{aligned}$$

(3) 由 (1) 及 (2) 立即可得. □

例 7. 設 $A_1 = A_3 = \cdots = [0,\ 2]$, $A_2 = A_4 = \cdots = [1,\ 3]$, 試求 $\limsup A_n$ 及 $\liminf A_n$.

解 利用定理 1.10. 由於 $\forall x \in [0,\ 3]$, $x \in A_n$ i.o., 而當 $x \in [1,\ 2]$, $x \in A_n$, $\forall n \in \mathbb{N}$, 但當 $x \in [0,3] \setminus [1,2]$, $x \in A_n$ a.a. 顯然不眞, 是以

$$\limsup A_n = [0,\ 3], \quad \liminf A_n = [1,\ 2].$$

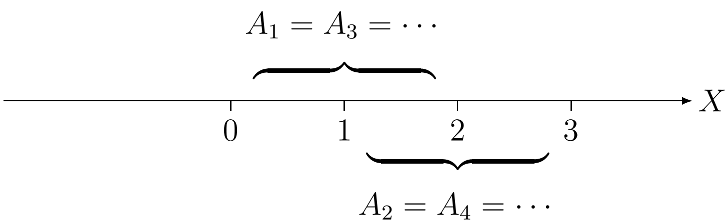


圖 1–5

定理 1.11

(1) 若 $\{A_n\}_n$ 爲一單調遞增序列, 即 $\forall n \in \mathbb{N}$, $A_n \subset A_{n+1}$, 則 $\lim A_n = \bigcup_{n\in\mathbb{N}} A_n$.

(2) 若 $\{A_n\}_n$ 爲一單調遞減序列, 即 $\forall n \in \mathbb{N}$, $A_{n+1} \subset A_n$, 則 $\lim A_n = \bigcap_{n\in\mathbb{N}} A_n$.

證 (1) $\{A_n\}_n$ 爲一單調遞增序列

$$\begin{aligned}
&\Rightarrow \bigcup_{k\geq n} A_k = \bigcup_{k\geq 1} A_k, \bigcap_{k\geq n} A_k = A_n \\
&\Rightarrow \bigcap_{n\geq 1}\bigcup_{k\geq n} A_k = \bigcup_{k\geq 1} A_k, \bigcup_{n\geq 1}\bigcap_{k\geq n} A_k = \bigcup_{n\geq 1} A_n \\
&\Rightarrow \limsup A_n = \bigcup_{k\geq 1} A_k = \bigcup_{n\geq 1} A_n = \liminf A_n \\
&\Rightarrow \lim A_n = \bigcup_{n\geq 1} A_n.
\end{aligned}$$

(2) 仿 (1), 讀者自證之. □

定理 1.12

設 $(\Omega, \mathcal{F})$ 爲一可測空間, 若 $\forall n \in \mathbb{N}$, $A_n \in \mathcal{F}$, 則 $\limsup A_n$ 及 $\liminf A_n$ 均屬於 $\mathcal{F}$.

證 易明. □

§ 1.4 機率空間

定義 1.13

設 $(\Omega, \mathcal{F})$ 爲一可測空間, 若$P\colon \mathcal{F} \to [0, 1]$ 滿足

1° $P(\Omega) = 1$,

2° P 具 σ 加法性, 即: 若 $\{A_j\}_{j\in\mathbb{N}}$ 爲 $\mathcal{F}$ 之一互斥序列, 則

$$P\Big(\sum_{j=1}^{+\infty} A_j\Big) = \sum_{j=1}^{+\infty} P(A_j). \quad \text{(見下註)}$$

則稱 P 爲 $(\Omega, \mathcal{F})$ 上之**機率測度**或**機率** (probability (measure)), 此時 $(\Omega, \mathcal{F}, P)$ 則稱其爲一**基本事件**(elementary event or simple event). **機率空間**(probability space); 而 $\mathcal{F}$ 之元素稱爲一**(隨機)事件**((random) event); 若單元集合 $\{\omega\}$ 爲一事件, 則稱其爲一**基本事件**(elementary event or simple event).

註: (1) 若 $A \cap B = \varnothing$, 則規定 $A + B = A \cup B$; 若 $A \cap B \neq \varnothing$, 則 $A + B$ 無意義;

(2) 若 J 爲有限或可數集合, $\{A_j\}_{j\in J}$ 中兩兩互斥, 則規定 $\sum_{j\in J} A_j = \bigcup_{j\in J} A_j$;

(3) 若存在 $i \neq j$ 使得 $A_i \cap A_j \neq \varnothing$, 則 $\sum_{j\in J} A_j$ 無意義.

我們稱這種聯集稱爲**互斥聯集**(disjoint union).

定理 1.14

> $P(\varnothing) = 0.$

證 設 $A_1 = A_2 = \cdots = \varnothing$, 則 $\{A_j\}_{j\in\mathbb{N}}$ 爲 $\mathcal{F}$ 之一互斥序列, 由 P 之 σ 加法性知

$$P(\varnothing) = P\Big(\sum_{j=1}^{+\infty} A_j\Big) = \sum_{j=1}^{+\infty} P(A_j) = \sum_{j=1}^{+\infty} P(\varnothing),$$

若 $P(\varnothing) > 0$, 則上述等式之右端等於 $+\infty$, 顯然矛盾, 是以 $P(\varnothing) = 0$. $\square$

定理 1.15

> 設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, 則
>
> (1) P 具有限加法性, 即: 若 $A_1, \cdots, A_n$ 爲兩兩互斥事件, 則其互斥聯集之機率爲
>
> $$P\Big(\sum_{j=1}^{n} A_j\Big) = \sum_{j=1}^{n} P(A_j).$$
>
> (2) 若 $A \in \mathcal{F}$, 則 $P(\complement A) = 1 - P(A)$.
>
> (3) 若 $A, B \in \mathcal{F}$ 且 $A \subset B$, 則 $P(A) \leq P(B)$.

證 (1) 令 $A_{n+1} = A_{n+2} = \cdots = \varnothing$, 則 $\{A_j\}_{j\in\mathbb{N}}$ 爲一互斥序列, 故

$$P\Big(\sum_{j=1}^{n} A_j\Big) = P\Big(\sum_{j=1}^{+\infty} A_j\Big) = \sum_{j=1}^{+\infty} P(A_j) = \sum_{j=1}^{n} P(A_j).$$

(2) 因爲 A 與 $\complement A$ 爲互斥, 由 (1) 知,

$$1 = P(\Omega) = P(A + \complement A) = P(A) + P(\complement A).$$

(3) 若 $A, B \in \mathcal{F}$ 且 $A \subset B$, 則
$P(A) \leq P(B)$. $\square$

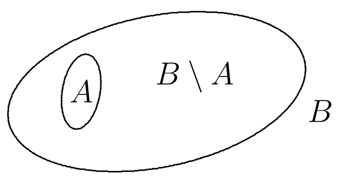


圖 1–6

註: 即使 A 爲 B 之眞子集, 亦未必有 $P(A) < P(B)$, 除非已知 $P(B \setminus A)$ 爲正.

我們將就本章第一節所舉隨機試驗之例, 表爲機率空間之型態, 亦即求出該試驗之樣本空間 Ω、σ 域 $\mathcal{F}$ 以及機率測度 P.

例 1. $\mathcal{E}_1$: 擲一骰子, 而觀察其 (頂面) 出現之點數; 則

$\Omega_1 = \{1,\, 2,\, 3,\, 4,\, 5,\, 6\}$;

$\mathcal{F}_1 = \mathcal{P}(\Omega_1)$; (因爲 Ω_1 爲有限);

$P\colon \mathcal{F}_1 \to [0,1] : P\{1\} = P\{2\} = P\{3\} = P\{4\} = P\{5\} = P\{6\} = \frac{1}{6}$.

說明: 我們假設此骰爲一公正之骰(fair die), 每面出現之機會相等, 故各面出現之機率均爲 1/6. 此外, $\mathcal{F}_1$ 內其他元素之機率雖未全部規定, 但可利用有限加法性而得之. 例如, $A = \{1,\, 3,\, 5\}$, $P(A)$ 乃出現奇數點之機率, 則

$$P(A) = P\{1\} + P\{3\} + P\{5\} = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

□

例 2. $\mathcal{E}_2$: 擲一錢幣, 而觀察其出現爲正面或反面; 則

$\Omega_2 = \{H,\, T\}$; 內 H 表正面 (head), T 表反面 (tail);

$\mathcal{F}_2 = \mathcal{P}(\Omega_2)$; (因爲 Ω_2 爲有限);

$P\colon \mathcal{F}_2 \to [0,1] : P\{H\} = P\{T\} = \frac{1}{2}$. □

例 3. $\mathcal{E}_3$: 擲一錢幣四次, 而觀察其正反兩面出現之次序; 則

$\Omega_3 = \{(a_1,\, a_2,\, a_3,\, a_4) \mid a_j \in \Omega_2\} = \{H,\, T\}^4$;

$\mathcal{F}_3 = \mathcal{P}(\Omega_3)$; (因爲 $\#\Omega_3 = 16$, 有限);

$P\colon \mathcal{F}_3 \to [0,1] : P\{(a_1,\, a_2,\, a_3,\, a_4)\} = \frac{1}{16}$. □

例 4. $\mathcal{E}_4$: 觀察某大醫院之產房, 每位產婦生產之嬰兒爲男或爲女或其他 (包括畸形或死胎); 則

$\Omega_4 = \{M,\, F,\, O\}$; M 爲男嬰, F爲女嬰, O爲其他 ;

$\mathcal{F}_4 = \mathcal{P}(\Omega_4)$;

$P\colon \mathcal{F}_4 \to [0,1] : P\{M\} = \frac{104}{205},\ P\{F\} = \frac{100}{205},\ P\{O\} = \frac{1}{205}$.

(以上均爲假設數字). □

例 5. $\mathcal{E}_5$: 某一日光燈工廠, 隨機取出一支燈管點亮之, 而觀察其壽命爲若干小時; (假設此工廠無限制繼續生產); 則

$\Omega_5 = \{t \mid t \geq 0\} = [0, +\infty)$;

$\mathcal{F}_5 = \{\Omega_5 \cap B \mid B \in \mathcal{B}\} = \{B \mid B \subset \Omega_5,\, B \in \mathcal{B}\}$ [5]

$P\colon \mathcal{F}_5 \to [0,1]$

[5] 此處之 Borel 域 $\mathcal{B}$, 若改爲 Lebesgue 可測集所成之集合則更爲完美, 但對未學過《實變函數論》之讀者, 恐太深奧, 二者差異不大.

我們不能仿照上述各例, 以規定各基本事件之機率爲手段而達到表達 P 之目的. 因爲每一基本事件之機率均爲零, 必須規定每一區間 I 之機率才可以, 在第二章中, 我們將利用積分方法以克服定義 P 之困難, 即 $P(I)=\int_I f(t)dt$, 至於 f 如何求得, 這與例 2 中 $P\{H\}$ 爲何等於 $\frac{1}{2}$ 相似, 但困難更多, 屬於數理統計的問題.) □

例 6. $\mathcal{E}_6$: 觀察某一地點, 每日最高及最低溫度;

$\Omega_6=\{(a,b)\mid m\le a\le b\le M\}$; 內 a,b 分別表某日之 最低及最高溫度;

$\mathcal{F}_6=\{\Omega_6\cap B\mid B\in\mathcal{B}^2\}$;

$P\colon \mathcal{F}_6\to[0,1]$

情形與例 5 相仿, 必須經過長時間觀方可獲得一個近似之分配. □

例 7. $\mathcal{E}_7$: 在區間 $[0,1]$ 中, 隨機抽取一數.

$\Omega_7=[0,1]$;

$\mathcal{F}_7=\{\Omega_7\cap B\mid B\in\mathcal{B}\}$;

$P\colon \mathcal{F}_7\to[0,1] : P(A)=A$ 之長度, (Lebesgue 之測度觀念);

例如, $A_1=[\frac{1}{2},1]$, 則 $P(A_1)=\frac{1}{2}$; $A_2=[0,1]\cap\mathbb{Q}$, 則 $P(A_2)=0$ 等等. □

定理 1.16 [**加法定理 (addition theorem)**]

(1) 若 $A_1,\,A_2\in\mathcal{F}$, 則 $P(A_1\cup A_2)=P(A_1)+P(A_2)-P(A_1\cap A_2)$.

(2) 若 $A_1,\,A_2,\,A_3\in\mathcal{F}$, 則

$$P(A_1\cup A_2\cup A_3)=\sum_{j=1}^{3}P(A_j)-\sum_{1\le j_1<j_2\le 3}P(A_{j_1}\cap A_{j_2})+P(A_1\cap A_2\cap A_3).$$

(3) 若 $A_1,\,A_2,\,\cdots,\,A_n\in\mathcal{F}$, 則

$$\begin{aligned}P\Big(\bigcup_{j=1}^{n}A_j\Big)=\sum_{j=1}^{n}P(A_j)&-\sum_{1\le j_1<j_2\le n}P(A_{j_1}\cap A_{j_2})\\&+\sum_{1\le j_1<j_2<j_3\le n}P(A_{j_1}\cap A_{j_2}\cap A_{j_3})\\&-\cdots+(-1)^{n+1}P(A_1\cap\cdots\cap A_n).\end{aligned}$$

證 (1) 因爲 $A_1 \cup A_2 = (A_1 \setminus A_2) + (A_2 \setminus A_1) + (A_1 \cap A_2)$, (參見圖 1–7), 故

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1 \setminus A_2) + P(A_2 \setminus A_1) + P(A_1 \cap A_2) \\ &= P(A_1 \setminus A_2) + P(A_1 \cap A_2) + P(A_2 \setminus A_1) + P(A_1 \cap A_2) \\ &\qquad - P(A_1 \cap A_2) \\ &= P(A_1) + P(A_2) - P(A_1 \cap A_2). \end{aligned}$$

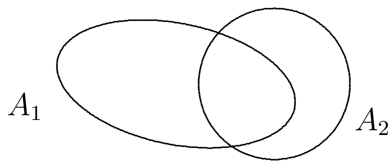


圖 1–7

(2) $P(A_1 \cup A_2 \cup A_3) = P(A_1 \cup (A_2 \cup A_3))$

$$\begin{aligned} &= P(A_1) + P(A_2 \cup A_3) - P(A_1 \cap (A_2 \cup A_3)) \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_2 \cap A_3) - P((A_1 \cap A_2) \cup (A_1 \cap A_3)) \\ &= \sum_{j=1}^{3} P(A_j) - P(A_2 \cap A_3) - (P(A_1 \cap A_2) \\ &\qquad + P(A_1 \cap A_3) - P(A_1 \cap A_2 \cap A_3)) \\ &= \sum_{j=1}^{3} P(A_j) - (P(A_2 \cap A_3) + P(A_1 \cap A_2) + P(A_1 \cap A_3)) \\ &\qquad + P(A_1 \cap A_2 \cap A_3). \end{aligned}$$

(3) 利用數學歸納法, 讀者自證之. □

例 8. 某電流自 P 流至 Q 經過 a, b, c 三個開關, 如圖所示. 已知 a 接通之機率爲 0.4; b 接通之機率爲 0.5; c 接通之機率爲 0.6; a 與 b 均接通之機率爲 0.2; b 與 c 均接通之機率爲 0.3; a 與 c 均接通之機率爲 0.24 ; a, b, c 均接通之機率爲 0.12. 試求:

(1) 電流自 P 流至 Q 之機率;

(2) 電流自 P 流至 Q 只經過一道開關之機率.

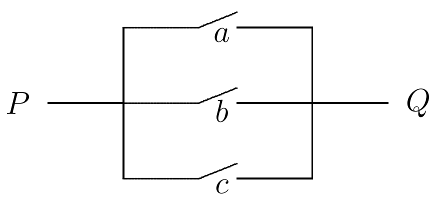


圖 1–8

解 設 A, B, C 分別表事件『電流自 P 經 a, b 或 c 流至 Q』.

(1) 電流自 P 流至 Q 之機率 $= P(A\cup B\cup C)$

$$\begin{aligned}&= P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)\\ &\qquad -P(A\cap C)+P(A\cap B\cap C)\\ &= 0.4+0.5+0.6-0.2-0.3-0.24+0.12=0.88.\end{aligned}$$

(2) 電流自 P 流至 Q 只經過一道開關之機率

$$\begin{aligned}&= P(A)+P(B)+P(C)\\ &\qquad -2(P(A\cap B)+P(B\cap C)+P(A\cap C))\\ &\qquad\qquad +3P(A\cap B\cap C)\\ &= 0.4+0.5+0.6-2(0.2+0.3+0.24)+3\times 0.12\\ &= 0.38.\end{aligned}$$ □

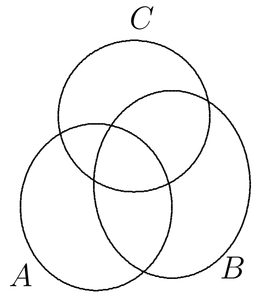


圖 1–9

例 9. 匹配問題(Matching problem)

設有 n 個學生, 每人帶一飯盒上學, 蒸飯後, 每人隨機取其一.

(1) 試求至少一人取得自已飯盒之機率?

(2) 當 $n\to\infty$ 時, 上述機率之極限爲若干?

(3) 試求恰有 k 人取得自已飯盒之機率, 內 $k<n-1$.

解 (1) 爲方便計, 我們先將學生編號爲 1 至 n 號; 第 k 號學生攜帶第 k 號飯盒上學, $k=1,\cdots,n$.

設 A_k 表事件『第 k 號學生取回第 k 號飯盒』, 則至少一人取得自已飯盒之機率爲 $P_n=P(A_1\cup A_2\cup\cdots\cup A_n)$. 利用加法定理知

$$\begin{aligned}P_n &= P\Big(\bigcup_{j=1}^{n}A_j\Big)\\ &= \sum_{j=1}^{n}P(A_j)-\sum_{1\le j_1<j_2\le n}P(A_{j_1}\cap A_{j_2})\\ &\qquad +\sum_{1\le j_1<j_2<j_3\le n}P(A_{j_1}\cap A_{j_2}\cap A_{j_3})-\cdots\\ &\qquad +(-1)^{n+1}P(A_1\cap\cdots\cap A_n). \qquad (*)\end{aligned}$$

其中

$$\begin{aligned}
P(A_j) &= \frac{(n-1)!}{n!} = \frac{1}{n},\\
P(A_{j_1}\cap A_{j_2}) &= \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},\\
P(A_{j_1}\cap A_{j_2}\cap A_{j_3}) &= \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)},\\
&\vdots\\
P(A_1\cap\cdots\cap A_n) &= \frac{1}{n!}.
\end{aligned}$$

代入 (*) 式中, 則

$$\begin{aligned}
P_n &= \binom{n}{1}\cdot\frac{1}{n} - \binom{n}{2}\cdot\frac{1}{n(n-1)} + \binom{n}{3}\cdot\frac{1}{n(n-1)(n-2)}\\
&\qquad + \cdots + (-1)^{n+1}\binom{n}{n}\cdot\frac{1}{n!}, \quad \text{(註)}\\
&= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1}\frac{1}{n!}.
\end{aligned}$$

$\Bigg($ 註: $\displaystyle\binom{n}{k} = \frac{n!}{k!(n-k)!}$ 二項係數也, 詳見附錄一定理 A.3. $\Bigg)$

(2) 上述 P_n 可化爲

$$P_n = 1 - \Big(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!}\Big) = 1 - \sum_{k=0}^{n}\frac{(-1)^k}{k!},$$

顯然 $\displaystyle\lim_{n\to+\infty} P_n = 1 - e^{-1} \approx 0.6321.$

(3) 設 Q_k 表 k 個學生, 無人取得自己飯盒之機率. 我們先求:『 n 個人中前 k 人拿到自己飯盒, 其餘皆未取得自己飯盒』之機率, 則由 (1) 知, 此一機率爲

$$\begin{aligned}
R_k &= \frac{1}{n}\cdot\frac{1}{n-1}\cdots\frac{1}{n-k+1}\cdot Q_k\\
&= \frac{(n-k)!}{n!}\cdot Q_k\\
&= \frac{1}{k!\binom{n}{k}}\cdot\Big(\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^k\frac{1}{k!}\Big).
\end{aligned}$$

是以恰有 k 人取得自己飯盒之機率乃 R_k 之 $\displaystyle\binom{n}{k}$ 倍, 即爲

$$\binom{n}{k}\cdot R_k = \frac{\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^k\frac{1}{k!}}{k!}. \qquad \square$$

例 10. 生日問題(Birthday Problem)

某一班級有 N 人 ($N \le 365$), 試問至少有二人生日相同之機率爲若干? (假設一年有 365 天且每一天做爲生日之機率皆爲 1/365).

解 N 人生日皆相異之機率爲

$$p_1(N) = \frac{P_{365,N}}{365^N} = \frac{365 \cdot 364 \cdots (365-N+1)}{365^N} = \prod_{j=0}^{N-1} \frac{365-j}{365}.$$

則至少有二人生日相同之機率爲

$$p_2(N) = 1 - p_1(N) = 1 - \prod_{j=0}^{N-1} \frac{365-j}{365}.$$

我們以電腦計算不同 N 時, 上述 p_1, p_2 諸值, 發現若一班有 60 位學生, 則幾可確定必有二人生日相同.

N	$p_1(N)$	$p_2(N)$
10	0.8831	0.1169
23	0.4927	0.5073
41	0.0968	0.9032
60	0.0059	0.9941

□

在本節結束前, 我們將介紹一個有關機率測度 P 的定理, 它在第二章以後將扮演十分重要的角色.

定理 1.17 [P **之連續性定理**]

> 設 $\{A_n\}_n$ 爲一事件序列, 且 $A = \lim_{n\to+\infty} A_n$, 則
>
> $$\lim_{n\to+\infty} P(A_n) = P(A) = P(\lim_{n\to+\infty} A_n).$$

證 我們將分三步驟證明之. 爲方便計, $\lim_{n\to+\infty}$ 將寫爲 $\lim$.

1° 先設 $\{A_n\}_n$ 爲一遞增序列, 即 $A_1 \subset A_2 \subset \cdots$, 令

$$A_0 = \varnothing,\ E_n = A_n \setminus A_{n-1},\ \forall n \in \mathbb{N},$$

則應有

- $$\begin{aligned} A_n &= (A_n \setminus A_{n-1}) + (A_{n-1} \setminus A_{n-2}) + \cdots + (A_2 \setminus A_1) + (A_1 \setminus A_0) \\ &= E_n + E_{n-1} + \cdots + E_1 = \sum_1^n E_j, \end{aligned}$$
- $$\sum_{j=1}^{+\infty} E_j = \bigcup_{n=1}^{+\infty} A_n = A.$$

故 $P(A_n)=\sum_{j=1}^{n}P(E_j)$,

$$\lim P(A_n)=\lim\sum_{j=1}^{n}P(E_j)=\sum_{j=1}^{+\infty}P(E_j)=P\Big(\sum_{j=1}^{+\infty}E_j\Big)=P(A).$$

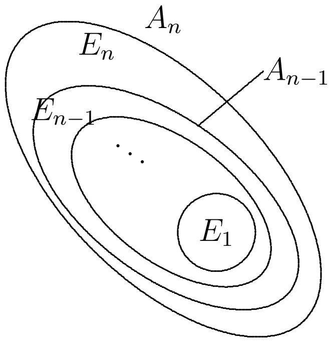


圖 1–10

2° 其次, 設 $\{A_n\}_n$ 爲一遞減序列, 即 $A_1\supset A_2\supset\cdots$, 則 $\{\complement A_n\}_n$ 爲一遞增序列, 且 $A=\lim A_n=\bigcap_{1}^{+\infty}A_n$, 故

$$\begin{aligned}\lim P(A_n)&=\lim(1-P(\complement A_n))\\&=1-\lim P(\complement A_n)\\&=1-P(\lim\complement A_n),\quad(\because\ 1^\circ)\\&=1-P\Big(\bigcup_{1}^{+\infty}\complement A_n\Big)\\&=1-P\Big(\complement\bigcap_{1}^{+\infty}A_n\Big)\\&=1-P(\complement A)=P(A).\end{aligned}$$

3° 最後, 設 $\{A_n\}_n$ 爲一般事件序列, 我們若能證得

$$P(A)\le\liminf P(A_n)\le\limsup P(A_n)\le P(A)$$

即可. 爲此令 $B_n=\bigcup_{k\ge n}A_k$, 則

- $\forall n\in\mathbb{N},\ B_n\supset A_n$;
- $\forall n\in\mathbb{N},\ B_n\supset B_{n+1}$; 即 $\{B_n\}_n$ 爲一遞減序列;
- $\lim_{n\to+\infty}B_n=\bigcap_{n\ge1}B_n=\bigcap_{n\ge1}\bigcup_{k\ge n}A_k=\limsup A_n=\lim A_n=A$.

故由 2° 得

$$P(A)=\lim P(B_n)=\limsup P(B_n)\ge\limsup P(A_n).$$

同理, 利用 1° 亦可證明 $P(A) \le \liminf P(A_n)$, 至於

$$\liminf P(A_n) \le \limsup P(A_n)$$

則顯然成立. □

§ 1.5 條件機率

例 1. 某校有學生 5,000ℤ人, 其中男生 4,000ℤ人, 女生 1,000ℤ人; 戴眼鏡者共有 1,500ℤ人, 爲方便計以 Ω、M、F、G 分別表全校學生、男生、女生及戴眼鏡者集合, 在樣本空間 Ω 中, 男生、女生及戴眼鏡者的比率 (即機率) 分別爲

$$P(M) = \frac{4,000}{5,000} = 0.8,\ P(F) = \frac{1,000}{5,000} = 0.2,\ P(G) = \frac{1,500}{5,000} = 0.3.$$

分析這類問題, 我們自然會遭遇到以下的問題:

(1) 女生中戴眼鏡者的比率 (機率) 爲若干?

(2) 戴眼鏡者中女生的比率 (機率) 爲若干?

對於問題 (1), 我們必須先知道女生中多少人戴眼鏡, 亦即 $\#(F \cap G) =?$ 假定是 240 人, 則女生中戴眼鏡者的比率 (機率) 爲

$$p_1 = \frac{\#(F \cap G)}{\#F} = \frac{240}{1,000} = 0.24.$$

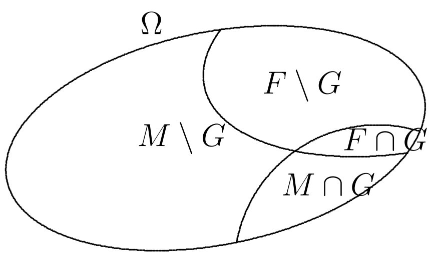


圖 1–11

對於問題 (2), 我們已知道女生中 240 人戴眼鏡, 是以戴眼鏡者中女生的比率 (機率) 爲

$$p_2 = \frac{\#(F \cap G)}{\#G} = \frac{240}{1,500} = 0.16.$$

在 p_1 中, 若分子分母分別除以 $\#\Omega = 5,000$, 則

$$p_1 = \frac{\#(F \cap G)}{\#F} = \frac{\#(F \cap G)/\#\Omega}{\#F/\#\Omega} = \frac{P(F \cap G)}{P(F)}.$$

□

這種機率, 我們稱其爲“給女生後, 戴眼鏡者之條件機率”, 於是我們乃有以下之定義:

定義 1.18

設 $(\Omega,\ \mathcal{F},\ P)$ 爲一機率空間, $A\in\mathcal{F},\ \ P(A)>0$, 則函數

$$P(\,\cdot\,|A)\colon \mathcal{F}\to[0,1]:P(B|A)=\frac{P(B\cap A)}{P(A)}$$

稱爲給 A **後之條件機率 (測度)**(conditional probability (measure) given A), 而 $P(B|A)$ 則稱爲給 A **後**, B **之條件機率** (conditional probability of B given A).

性質: 若 A, B 均爲事件, 且 $P(A)>0$, 則

$$P(A\cap B)=P(B|A)\cdot P(A).$$

定理 1.19

定義 1.18 中之函數 $P(\,\cdot\,|A)$ 爲 $(\Omega,\ \mathcal{F})$ 上一機率測度.

證 顯然, $\forall B\in\mathcal{F},\ \ P(B|A)\in[0,1]$, 此乃說明函數 $P(\,\cdot\,|A)$ 爲妥當定義 (well-defined). 此外,

1° $P(\Omega|A)=\dfrac{P(\Omega\cap A)}{P(A)}=1;$

2° 若 $\{A_j\}_j\in\mathbb{N}$ 爲一事件序列, 則

$$P\Big(\sum_{j=1}^{+\infty}A_j\ \Big|\ A\Big)=\frac{P\Big(A\bigcap\Big(\sum_1^{+\infty}A_j\Big)\Big)}{P(A)}=\frac{P\Big(\sum_1^{+\infty}(A\cap A_j)\Big)}{P(A)}$$

$$=\frac{\sum_1^{+\infty}P(A\cap A_j)}{P(A)}=\sum_{j=1}^{+\infty}P(A_j|A).$$

□

定理 1.20 [**乘法定理(multiplication theorem)**]

設 $(\Omega,\ \mathcal{F},\ P)$ 爲一機率空間, $A_1,\ A_2,\ \cdots,\ A_n\in\mathcal{F}$, 且 $P\Big(\bigcap_{j=1}^{n-1}A_j\Big)>0$, 則

$$P\Big(\bigcap_{j=1}^{n}A_j\Big)=P\Big(A_n\ \Big|\ \bigcap_{j=1}^{n-1}A_j\Big)\cdot P\Big(A_{n-1}\ \Big|\ \bigcap_{j=1}^{n-2}A_j\Big)\cdots P(A_2|A_1)\cdot P(A_1).$$

證 利用數學歸納法證明之. 當 $n=1$ 時, 原敘述顯然爲眞, 當 $n=2$ 時, 由定義 1.18 之後的“性質”知 $P(A_1\cap A_2)=P(A_2|A_1)P(A_1)$ 亦爲眞; 假設 $n=k$ 時原敘述爲眞, 往證 $n=k+1$ 時原敘述爲眞.

$$\begin{aligned}
P\Big(\bigcap_{j=1}^{k+1}A_j\Big) &= P\Big(\Big(\bigcap_{j=1}^{k}A_j\Big)\bigcap A_{k+1}\Big)\\
&= P\Big(A_{k+1}\,\Big|\,\bigcap_{j=1}^{k}A_j\Big)\cdot P\Big(\bigcap_{j=1}^{k}A_j\Big)\\
&= P\Big(A_{k+1}\,\Big|\,\bigcap_{j=1}^{k}A_j\Big)\cdot P\Big(A_k\,\Big|\,\bigcap_{j=1}^{k-1}A_j\Big)\cdots P(A_2|A_1)\cdot P(A_1). \qquad \square
\end{aligned}$$

或許有人會查問, 爲什麼定理 1.16 稱爲加法定理, 而定理Z1.20 稱爲乘法定理? 因爲集合論發展初期, 有人稱二集合之聯集 (union) 爲二集合之和 (sum); 而稱二集合之交集 (intersection) 爲二集合之積 (product); 由於定理 1.16 恰爲求事件 $A_1,\cdots,A_n$ 聯集之機率, 故稱爲加法定理, 而定理 1.20 則爲求事件 $A_1,\ \cdots,\ A_n$ 交集之機率, 故稱爲乘法定理.

例 2. 一盒子中置 12 球: 三藍四白五紅, 今以隨機不放回方式抽取四球, 試求:

(1) 第一球爲藍, 第二球爲白, 第三第四球均爲紅之機率;

(2) 前三球爲白, 第四球不爲白之機率.

解 設球之集合爲

$$T=\{藍_1,藍_2,藍_3,白_1,白_2,白_3,白_4,紅_1,紅_2,紅_3,紅_4,紅_5\}.$$

則此試驗之樣本空間爲

$$\Omega=\{(a,\ b,\ c,\ d)\mid a,\ b,\ c,\ d \text{ 相異} \in T\}.$$

(1) 設

$B_1=$ 事件 “第一球爲藍色” $=\{(a,\ b,\ c,\ d)\in\Omega\mid a$ 爲藍 $\}$;

$W_2=$ 事件 “第二球爲白色” $=\{(a,\ b,\ c,\ d)\in\Omega\mid b$ 爲白 $\}$;

$R_3=$ 事件 “第三球爲紅色” $=\{(a,\ b,\ c,\ d)\in\Omega\mid c$ 爲紅 $\}$;

$R_4=$ 事件 “第四球爲紅色” $=\{(a,\ b,\ c,\ d)\in\Omega\mid d$ 爲紅 $\}$;

則第一球爲藍, 第二球爲白, 第三第四球均爲紅之機率爲

$$\begin{aligned}
&P(B_1\cap W_2\cap R_3\cap R_4)\\
&\quad= P(R_4|B_1\cap W_2\cap R_3)P(R_3|B_1\cap W_2)P(W_2|B_1)P(B_1)\\
&\quad= \frac{4}{9}\cdot\frac{5}{10}\cdot\frac{4}{11}\cdot\frac{3}{12}\\
&\quad= \frac{2}{99}.
\end{aligned}$$

(2) 第四球不為白, 即為藍或為紅, 分別表為 B_4 及 R_4, 故所求機率為

$$
\begin{aligned}
&P(W_1 \cap W_2 \cap W_3 \cap (B_4 \cup R_4)) \\
&\qquad = P(B_4 \cup R_4 | W_1 \cap W_2 \cap W_3) P(W_3 | W_1 \cap W_2) P(W_2 | W_1) P(W_1) \\
&\qquad = \frac{8}{9} \cdot \frac{2}{10} \cdot \frac{3}{11} \cdot \frac{4}{12} \\
&\qquad = \frac{8}{495}.
\end{aligned}
$$

這一類問題若以樹圖 (tree diagram) 加以說明, 則一目了然.

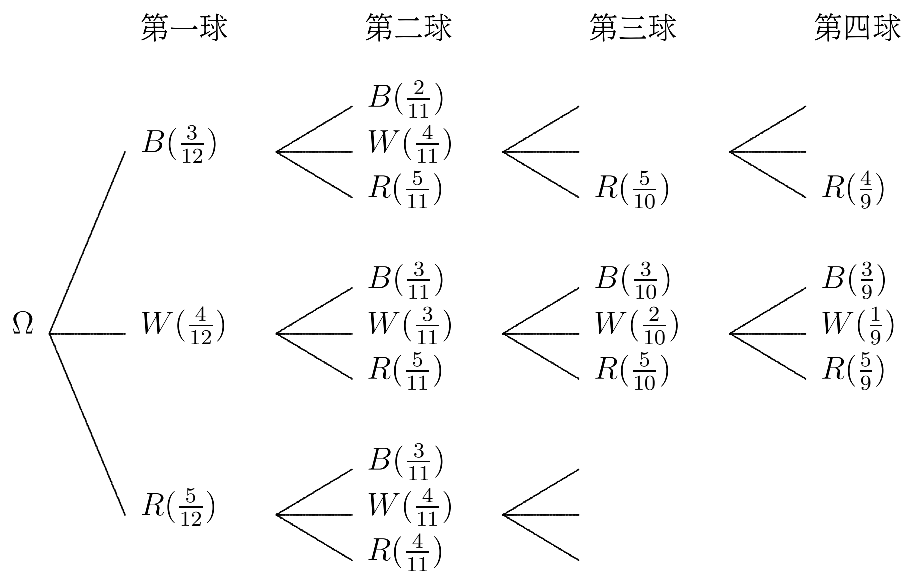


圖 1–12

上圖中之分數 $\frac{3}{12}$, $\frac{4}{12}$, $\frac{5}{12}$ 分別表第一球抽得藍、白或紅之機率; 第二欄之 $\frac{2}{11}$ 表條件機率, 第一球抽得藍球後, 第二球再抽得藍球之條件機率, 即 $P(B_2|B_1) = \frac{2}{11}$, 其餘同理. □

定理 1.21 [**全機率定理(Total probability theorem)**]

> 設 $(\Omega, \mathcal{F}, P)$ 為一機率空間, $\{A_j\}_j$ 為一至多可數互斥事件序列, 而且$\forall j$, $P(A_j) > 0$, 若事件 $B \subset \sum_j A_j$, 則
> $$P(B) = \sum_j P(B|A_j)P(A_j).$$

證 由於 $B = B \bigcap \sum_j A_j = \sum_j (B \cap A_j)$, (參閱圖 1–13), 故

$$P(B) = \sum_j P(B \cap A_j) = \sum_j P(B|A_j)P(A_j).$$

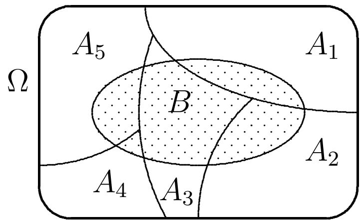


圖 1–13

□

例 3. (續例 2.) 試求抽出之第二球爲紅之機率爲若干?

解 設 B_1, W_1, R_1 分別表事件第一球爲藍色、白色、紅色, R_2 表事件第二球爲紅色, 顯然 $B_1 + W_1 + R_1 = \Omega \supset R_2$, 由全機率定理知

$$\begin{aligned}P(R_2) &= P(R_2|B_1)P(B_1) + P(R_2|W_1)P(W_1) + P(R_2|R_1)P(R_1)\\ &= \frac{5}{11}\cdot\frac{3}{12} + \frac{5}{11}\cdot\frac{4}{12} + \frac{4}{11}\cdot\frac{5}{12}, \text{(參考上例之樹圖)}\\ &= \frac{55}{132} = \frac{5}{12}.\end{aligned}$$

提醒讀者注意, 此一機率與 $P(R_1)$ 相等. 當然我們可以繼續追下去, 第三球爲紅之機率是否仍是 5/12 呢? 計算稍爲麻煩一點, 留給讀者自行思考. □

定理 1.22 [**Bayes 公式**]

設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, $\{A_j\}_j$ 爲一至多可數互斥事件序列, 而且$\forall j,\ P(A_j) > 0$, 若事件 $B \subset \sum_j A_j$, 則

$$P(B) = \sum_j P(B|A_j)P(A_j).$$

證 利用全機率定理,

$$\begin{aligned}P(A_k|B) &= \frac{P(A_k \cap B)}{P(B)}\\ &= \frac{P(B|A_k)P(A_k)}{\sum_j P(B|A_j)P(A_j)}.\end{aligned}$$ □

本定理的用意是: 如果諸條件機率 $P(B|A_j)$ 已經知道 (亦稱爲**事前機率** (a priori probability)), 則諸 $P(A_k|B)$ 均可一一求出,(稱爲**事後機率** (a posteriori probability)).

例 4. 甲乙丙三人在靶場射擊, 已知甲每射擊 50 發子彈時, 乙可射 53 發, 丙可射 60 發. 但命中靶心率卻爲: 甲 0.75, 乙 0.72, 丙 0.7. 今三人 同時對同一靶射擊, 試問

(1) 某子彈射中靶心之機率爲若干?

(2) 其中一發命中靶心之子彈爲乙所射擊之機率爲若干?

解

A_1 = 事件 "某子彈爲甲所射擊",

A_2 = 事件 "某子彈爲乙所射擊",

A_3 = 事件 "某子彈爲丙所射擊",

B = 事件 "某子彈射中靶心".

(1) 某子彈射中靶心之機率爲

$$\begin{aligned}P(B) &= \sum_{j=1}^{3} P(B|A_j)P(A_j)\\ &= 0.75 \times \frac{50}{163} + 0.72 \times \frac{53}{163} + 0.7 \times \frac{60}{163} = 0.722.\end{aligned}$$

(2) 其中一發命中靶心之子彈爲乙所射擊之機率爲

$$\begin{aligned}P(A_2|B) &= \frac{P(B|A_2)P(A_2)}{\sum_{j=1}^{3} P(B|A_j)P(A_j)}\\ &= \frac{0.72 \times \frac{53}{163}}{0.75 \times \frac{50}{163} + 0.72 \times \frac{53}{163} + 0.7 \times \frac{60}{163}}\\ &= 0.324.\end{aligned}$$

□

§ 1.6 獨立性

條件機率的概念啓示我們: 某一事件 A 之發生可能導致我們重新估另一事件 B 之機率, 亦即『B 發生之機率』與『給 A 後 B 之條件機率』可能不同; 但也 可能事件 A 之發生並不影響 B 之機率, 即 $P(B) = P(B|A)$. 例如在例 1 中, 戴眼鏡者之機率爲 $P(G) = 0.3$, 而女生中戴眼鏡之機率爲 $P(G|F) = 0.24$, 二者並不相等, 今假設該校女生中有 300 人戴眼鏡 (而非 240 人), 則

$$P(G|F) = \frac{\#(F \cap G)}{\#F} = \frac{300}{1000} = 0.3 = P(G).$$

此時

$$P(G|M) = \frac{\#(M \cap G)}{\#M} = \frac{1500 - 300}{4000} = 0.3 = P(G).$$

亦即, 男生中戴眼鏡之機率與全校戴眼鏡之機率亦爲相同. 此乃說明, 戴眼鏡者之機率與性別無關(independent).

由此一觀念, 我們將界定二事件之獨立性. 首先, 當 $P(B|A)=P(B)$ 時, 我們稱事件 A 與 B 爲獨立, 但如此定義有點小問題,

(1) 當 $A=\Omega,\ B=\varnothing$ 時, 顯然 $P(B|A)=0=P(B)$,

(2) 當 $A=\varnothing,\ B=\Omega$ 時, 則 $P(B|A)$ 無意義, 但 $P(B)=1$.

由 (1) 我們說 A 與 B 爲獨立, 即 Ω 與 $\varnothing$ 爲獨立; 但由 (2) 我們卻說 B 與 A 不爲獨立, 即 $\varnothing$ 與 Ω 不爲獨立.

爲解決此一缺點, 將 $P(B|A)=P(B)$ 改寫爲 $\dfrac{P(A\cap B)}{P(A)}=P(B)$, 移項之, 得 $P(A\cap B)=P(A)P(B)$, 而以之爲獨立之定義, 則因爲 A 與 B 之地位平等, 不致產生 A 與 B 爲獨立而 B 與 A 不爲獨立之困擾. 因此, 我們規定:

定義 1.23

事件 A 與事件 B 稱爲 (**隨機**) **獨立**((stochastically) independent), 並記爲 $A \perp\!\!\!\perp B$, 如果

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

二事件若不滿足上述條件, 則稱其爲**相依事件**(dependent events).

定理 1.24

若事件 A 與 B 爲獨立, 則 A 與 $\complement B, \complement A$ 與 $\complement B, \complement A$ 與 B 均爲獨立.

證 我們只證 A 與 $\complement B$ 爲獨立, 其餘同理. 由於 $A=(A\cap B)+(A\setminus B)$, 即有

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

是以

$$\begin{aligned}P(A\cap \complement B)&=P(A\setminus B)\\&=P(A)-P(A\cap B)\\&=P(A)-P(A)P(B),\quad (\because A\perp\!\!\!\perp B)\\&=P(A)(1-P(B))\\&=P(A)P(\complement B).\qquad\square\end{aligned}$$

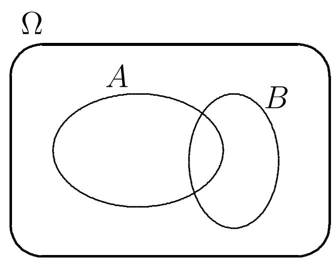


圖 1–14

初學機率之讀者, 往往希望能自圖形以了解『獨立』之概念, 我們在舉以下範例之前, 先聲明: 能以圖形表達之情況相對較少, 最佳方法仍是以『定義』去處理有關獨立之問題.

例 1. 一骰子連擲兩次, 若A表事件『第一次點數大於 3』, B表事件『第二次得 1 或 2 點』, 即

$$\begin{aligned}\Omega &= \{(a,\, b) \mid a,\, b \in \{1,\, 2,\, 3,\, 4,\, 5,\, 6\}\} \\ &= \{(1,\, 1),\, (1,\, 2),\, \cdots,\, (6,\, 6)\}, \\ A &= \{(a,b) \in \Omega \mid a > 3\}, \\ B &= \{(a,b) \in \Omega \mid b \in \{1,2\}\},\end{aligned}$$

則 $A \cap B = \{(4,1),(4,2),(5,1),(5,2),(6,1),(6,2)\}$, 顯然

$$P(A \cap B) = \frac{6}{36} = \frac{3}{6} \cdot \frac{2}{6} = P(A)P(B).$$

圖形如右. □

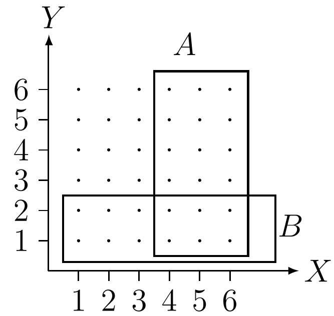


圖 1–15

由二事件之獨立性, 自然會聯想到三個事件 A、B 與 C 之獨立性. 除了應具有兩兩獨立之性質:

$$\begin{aligned}P(A \cap B) &= P(A)P(B), \\ P(A \cap C) &= P(A)P(C), \\ P(B \cap C) &= P(B)P(C),\end{aligned}$$

之外, $A \cap B$ 與 C, $A \cup B$ 與 C, $A \setminus B$ 與 C, $B \setminus A$ 與 C, $\cdots$ 等均應爲獨立, 如果以此爲定義, 顯然過於龐雜, 應用起來十分不便, 必須加以精簡. 精簡之道, 不在於捨棄某些條件, 而應找到最少之條件, 使能與原意相對等者. 我們發現

$$\begin{aligned}P(A \cap B) &= P(A)P(B), \\ P(A \cap C) &= P(A)P(C), \\ P(B \cap C) &= P(B)P(C), \\ P(A \cap B \cap C) &= P(A)P(B)P(C),\end{aligned}$$

等四條件, 即可證明事件 C 與下列諸事件均爲獨立:

$$A,\ B,\ A \cup B,\ A \cap B,\ A \setminus B,\ B \setminus A.$$

當然事件 A 與 B、C 諸集合運算亦爲獨立, 事件 B 與 A、C 諸集合之運算亦爲獨立, 因此, 我們們正式界定如下:

定義 1.25

我們稱事件 A, B, C 爲**獨立**, 如果滿足以下四條件:

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

我們知道, 當 $P(A\cap B)=P(A)P(B)$ 時, A 與 B 爲獨立, 因此有人希望上述四條件能精簡到只有 $P(A\cap B\cap C)=P(A)P(B)P(C)$, 這是錯誤的, 因爲此一條件不能蘊涵 A 與 B 爲獨立, 反例如下:

例 2. 設 $\Omega=\{1,2,3,4,5\}$, $\mathcal{F}=\mathcal{P}(\Omega)$,

$$P\{1\}=\frac{1}{8},\ P\{2\}=P\{3\}=P\{4\}=\frac{3}{16},\ P\{5\}=\frac{5}{16}.$$

若 $A=\{1,2,3\}$, $B=\{1,2,4\}$, $C=\{1,3,4\}$, 則雖有

$$\begin{aligned}P(A\cap B\cap C)&=P\{1\}=\frac{1}{8},\\ P(A)P(B)P(C)&=\left(\frac{1}{8}+\frac{3}{16}+\frac{3}{16}\right)^3\\ &=\frac{1}{8}\\ &=P(A\cap B\cap C).\end{aligned}$$

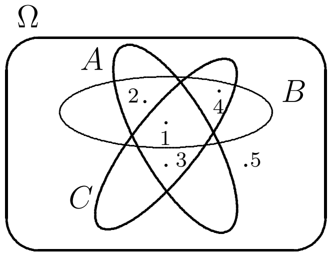


圖 1–16

但是

$$\begin{aligned}P(A\cap B)&=P\{1,2\}=\frac{5}{16},\\ P(A)P(B)&=\left(\frac{1}{8}+\frac{3}{16}+\frac{3}{16}\right)^2=\frac{1}{4},\end{aligned}$$

二者不等, 說明 A 與 B 並不爲獨立, 因此 A, B, C 亦不爲獨立. □

定義 1.26

(1) 設 $A_1, \cdots, A_n$ 爲 n 個事件, 若此 n 個事件中任何二個以上事件交集之機率等於各事件機率之積, 則稱 $A_1, \cdots, A_n$ 爲**獨立**. 更精確地說, $A_1, A_2, \cdots, A_n$ 應滿足

$$\forall k \in \{2, 3, \cdots, n\}, \forall j_1, \cdots, j_k \text{ 相異 } \in \{1, \cdots, n\},$$
$$P(A_{j_1} \cap \cdots \cap A_{j_k}) = P(A_{j_1}) \cdots P(A_{j_k}).$$

(2) 設 $\{A_n\}_{n\in\mathbb{N}}$ 爲一事件序列, 若 $\forall n \geq 2$, $A_1, \cdots, A_n$ 爲獨立, 則稱事件序列 $A_1, \cdots, A_n, \cdots$ 爲**獨立**.

如果我們討論二個以上隨機試驗之“結合”, 我們可以借用獨立性觀念而建造出一個新的試驗, 因而得一個新的機率空間. 此一討論可自二個機率空間開始推廣至 n 個乃至可數個, 最後到任意個機率空間之積, Loève [11] §1 有很精彩的討論, 但對初學者並不很容易. σ域是公設化機率論的重要基石, Chow & Teicher [2] 之第一章中加進了一些新東西 λ-class, π-class, monotone class 等一起討論. 關於獨立的觀念, 我們第六章中還要加以討論, 到時我們不再限於事件之間的獨立性.

第一章 習 題

隨機試驗

1-1 試問以下諸試驗, 何者爲命定, 何者爲隨機? 若爲隨機, 則寫出其樣本空間.

(1) 觀察某一十字路口之一紅綠燈, 當紅燈亮時, 有多少人闖越;

(2) 觀察收到的信件上有無郵票;

(3) 觀察農曆每月十五日有月亮的夜晚是否必爲月圓;

(4) 觀察每一列火車自台北至台南所費之時間;

(5) 觀察每次等公車之時間;

(6) 觀察每本書是否編上頁數.

1-2 試舉出兩個命定試驗(須非出自本書).

1-3 擲一錢幣, 若出現正面則再擲一次, 若爲反面則擲一骰子. 試寫出此一試驗之樣本空間及每一基本事件之機率.

1-4 某甲在其上班公司之一份年度報告中稱: 該公司 60 名新進人員中, 會游泳的 35 人, 會開汽車 24 人, 會使用電腦的 25 人, 會游泳又會開車的 12 人, 會游泳又會使用電腦的 10 人, 會開車又會使用電腦的 8 人, 三者皆會的 7 人. 這份報告呈上後, 某甲即被上司判定工作不力, 爲何?

1-5 區間 $[0, 3]$ 中隨機抽取二數.

(1) 寫出此試驗之樣本空間 ;

(2) 試求此二數之距離小於 1 之機率爲若干?

σ 域

1-6 設 $\Omega = \{1, 2, 3, 4\}$. 試列出 Ω 上三個不同之σ域.

1-7 設 A, B, C 爲 Ω 之三子集. $\mathcal{C} = \{A, A \cup B, A \cap C\}$. 試求$\sigma(\mathcal{C}) =$?

1-8 若 $\mathcal{F}_1$ 及 $\mathcal{F}_2$ 爲 Ω 上二 σ 域, 試問: $\mathcal{F}_1 \cap \mathcal{F}_2, \mathcal{F}_1 \cup \mathcal{F}_2$ 是否均爲 Ω 上之σ域 ? 證明或反證之.

1-9 設 $\mathcal{C} \subset \mathcal{D} \subset \mathcal{P}(\Omega)$, $\mathcal{D} \subset \sigma(\mathcal{C})$. 試證: $\sigma(\mathcal{C}) = \sigma(\mathcal{D})$.

1-10 設 $\mathcal{C} = \{(-\infty,\, x] \mid x \in \mathbb{R}\}$, 試證:

(1) $\forall a \in \mathbb{R},\ (-\infty,\, a) \in \sigma(\mathcal{C})$;

(2) $\sigma(\mathcal{C}) = \mathcal{B}$, (提示: 利用rm 1-9 題).

1-11 設 $A_n = \{x \in \mathbb{R} \mid 0 < x < 1/n\}$, 內 $n \in \mathbb{N}$. 試證: $\lim_{n\to+\infty} A_n = \varnothing$.

1-12 設 $A_n = [-1,\, \sin(n\pi/4)]$, 內 $n \in \mathbb{N}$. 試求: $\limsup A_n$ 及 $\liminf A_n$.

1-13 設 $\{A_n\}_n,\ \{B_n\}_n$ 爲二集合序列, 試證:

(1) $\limsup(A_n \cup B_n) = \limsup A_n \cup \limsup B_n$;

(2) $\limsup(A_n \cap B_n) \subset \limsup A_n \cap \limsup B_n$;

(3) 舉一列說明 (2) 小題之左右兩端未必相等.

1-14 設 $\{A_n\}_n$ 爲一集合序列, 若令

$$\sup_{k\geq n} A_k = \bigcup_{k\geq n} A_k, \quad \inf_{k\geq n} A_k = \bigcap_{k\geq n} A_k \ .$$

試證:

(1) $\limsup A_n = \lim_{n\to+\infty} \big(\sup_{k\geq n} A_k\big)$;

(2) $\liminf A_n = \lim_{n\to+\infty} \big(\inf_{k\geq n} A_k\big)$.

1-15 試證:

(1) $\complement(\limsup A_n) = \liminf \complement A_n$;

(2) $\complement(\liminf A_n) = \limsup \complement A_n$.

1-16 設 $A_{2n} = A,\ A_{2n-1} = B$, 內 $n \in \mathbb{N}$. 試求: $\limsup A_n$ 及 $\liminf A_n$. 並問 $\{A_n\}_n$ 極限是否存在?

1-17 設 Ω 爲一可數集合, (例如 $\Omega = \mathbb{N}$), 試證:

$$\mathcal{F} = \{A \subset \Omega \mid A \text{ 爲有限或 } \complement A \text{ 爲有限 }\}$$

爲 Ω 上一域, 但非一σ域.

機率空間

1-18 設 A 與 B 爲機率空間 $(\Omega,\, \mathcal{F},\, P)$ 之事件. 試證:

$$P(A \triangle B) = P(A) + P(B) - 2P(A \cap B).$$

內 $A \triangle B = (A \setminus B) + (B \setminus A)$ 稱爲 A 與 B 之 **對稱差**(symmetric difference).

1-19 設 $P_1, \cdots, P_n, \cdots$ 均為可測空間 $(\Omega, \mathcal{F})$ 上之機率測度.

(1) 若 $\alpha_j > 0$ 且 $\sum_{j=1}^n \alpha_j = 1$ 試證: $\displaystyle\sum_{j=1}^n \alpha_j P_j$ 亦為 $(\Omega, \mathcal{F})$ 上一機率測度;

(2) 若正項級數 $\displaystyle\sum_{n=1}^{+\infty} \alpha_n$ 之和為 1, 試證: $\displaystyle\sum_{n=1}^{+\infty} \alpha_n P_n$ 亦為 $(\Omega, \mathcal{F})$ 上一機率測度.

1-20 試將 Bertrand 之詭論表為三個機率空間.

1-21 設 $A_1, \cdots, A_n, \cdots$ 均為機率空間 $(\Omega, \mathcal{F}, P)$ 之事件, 試證:

(1) $P\Big(\displaystyle\bigcup_{j=1}^n A_j\Big) \le \sum_{j=1}^n P(A_j)$;

(2) $P(A_1 \cap A_2) \ge 1 - P(\complement A_1) - P(\complement A_2)$;

(3) $P\Big(\displaystyle\bigcup_{j=1}^{+\infty} A_j\Big) \le \sum_{j=1}^{+\infty} P(A_j)$.

1-22 設 $A_1, \cdots, A_j, \cdots$ 均為機率空間 $(\Omega, \mathcal{F}, P)$ 之事件, 且 $P(A_j) = 1,\ \forall j \in \mathbb{N}$, 試證: $P\Big(\displaystyle\bigcap_{j=1}^{+\infty} A_j\Big) = 1$.

條件機率與獨立

1-23 有一副紙牌 52 張, 隨機不放回抽取三張. 試問:

(1) 第二張為紅心之機率為若干?

(2) 第二張及第三張均為紅心之機率為若干?

1-24 某工廠有四部機器生產相同產品, 已知甲機器產量為乙機器之二倍, 乙為丙之 二倍, 丙為丁之二倍, 且甲、乙、丙及丁各有 $5\%, 4\%, 3\%, 2\%$ 之不良品, 今隨機抽取一件, 發現其為不良品, 試問此一產品為甲機器所生 產之機率為若干?

1-25 擲二骰, 設 A 表第一骰出現奇數, B 表第二骰出現偶數, C 表二骰之 和為偶數. 證明: $A \perp\!\!\!\perp B, B \perp\!\!\!\perp C, C \perp\!\!\!\perp A$, 但 A, B, C 不為 獨立.

1-26 擲二骰, 設 A 表第一骰出現偶數, B 表二骰之和為 4, C 表二骰點數 之差小於 3, 證明: $P(A \cap B \cap C) = P(A)P(B)P(C)$, 但 A, B, C 不 為獨立.

1-27 設 $A_1, \cdots, A_n$ 均為獨立之事件, 證明:

$$P\Big(\bigcup_{j=1}^n A_j\Big) = 1 - (1 - P(A_1)) \cdots (1 - P(A_n)).$$

1-28 設事件 A_1, A_2 分別與事件 B 為獨立.

(1) 若 A_1 與 A_2 為互斥, 證明: $A_1 \cup A_2$ 與 B 為獨立;

(2) 若 A_1 與 A_2 不為互斥, 證明: $A_1 \cup A_2$ 與 B 未必為獨立.

Chapter 2

隨機變數

§ 2.1 隨機變數

公設化之機率空間定義的主要目的, 乃利用嚴格的數學形成來探討隨機事件的規則性. 由於樣本空間內之元素通常僅以一可以識別之符號 (如 a, b, c 等), 很少將該元素之全部屬性均予表出, 因此, 我們很難迅速加以分類整理, 進行分析其中存在之規則. 是以有必要加以簡化, 針對我們所欲了解的性質加以分類, 其中最常用的分類乃是數量性的 (quantitative). 數量性資料除了有序數性之外, 尚有基數性, 數與數之間可以比較大小, 也可以施行四則運算, 增加我們對於事件規則的了解.

例如一錢幣連擲十次, 則其樣本空間為

$$\Omega = \{(a_1, \cdots, a_{10}) \mid \forall j \in \{1, 2, \cdots, 10\},\ a_j \in \{H, T\}\}.$$

Ω 內共有 $2^{10} = 1024$ 個元素, 我們若僅希望了解每一結果 $(a_1, \cdots, a_{10})$ 內正面出現之次數及其機率, 而不十分關心到底 H 與 T 是如何排列, 則可設函數

$$X : \Omega \to \mathbb{R} : X(a_1, \cdots, a_{10}) = \text{“}(a_1, \cdots, a_{10}) \text{ 中 } H \text{ 出現之次數 ”}.$$

此時, 顯然 X 之值域為 $A = \{0, 1, \cdots, 10\}$ (又稱為 X 之樣本空間). 至於 A 中各元素之機率, 我們將於稍後詳細討論. 這種自樣本空間映至實數系的函數, 通常乃為表示各元素之某一項特徵 (characteristic), 我們稱其為一隨機變數. 為避免可能的困擾, 我們將採取嚴格的數學形式做為隨機變數之正式定義.

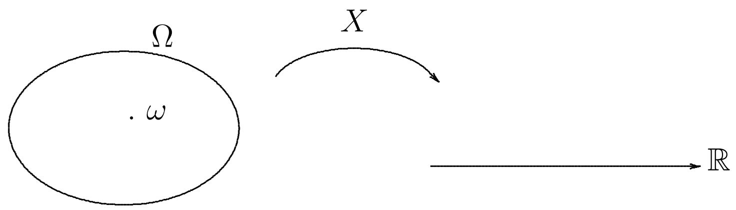


圖 2–1

定義 2.1

(1) 設 $(\Omega, \mathcal{F}, P)$ 為一機率空間, $\mathcal{B}$ 為一維 Borel 域. 我們稱 X 為 Ω 上一**隨機變數**(random variable), 簡為 r.v., 如果

$X\colon \Omega \to \mathbb{R}$, 且

$\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{F}$;

此時 X 之值域 $X(\Omega)$ 稱為 X 之**樣本空間**.

(2) 設 $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ 皆為可測空間, 我們稱 則稱 X 為 $(\mathcal{F}, \mathcal{F}')$**可測**(measurable), 如果

$X\colon \Omega \to \Omega'$, 且

$\forall B \in \mathcal{F}', X^{-1}(B) \in \mathcal{F}$.

隨機變數的原始意義, 只是將非數量性集合 Ω, 轉換為數量性集合 $\mathbb{R}$, 並不需其他條件, 但在上述定義中我們增加了一個很強的條件: $\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{F}$, 它究竟有何用意? 在系 2.4 之後, 我們再詳細說明. 如果讀者細心一點, 不難發現上述定義中的 (1) 是 (2) 的特例, 更清楚地說: 所謂隨機變數只是自 Ω 映至 $\mathbb{R}$ 的 $(\mathcal{F}, \mathcal{B})$ 可測函數. 隨機變數在機率論中之重要性正如實值實變函數在微積分中之份量.

在《機率論》中我們常需涉及集合論中之『像原』觀念, 在此不妨稍加回顧.

設 $f\colon A \to B$ 為一函數, 若 $T \subset B$, 則集合

$$f^{-1}(T) = \{x \in A \mid f(x) \in T\}$$

稱為 T 之 f 像原 (inverse image), 例如: $f\colon \mathbb{R} \to \mathbb{R} : f(x) = x^2$, 則

$$\begin{aligned} f^{-1}[1,2] &= \{x \in \mathbb{R} \mid x^2 \in [1,2]\} \\ &= [-\sqrt{2}, -1] \cup [1, \sqrt{2}]. \end{aligned}$$

如右圖所示.

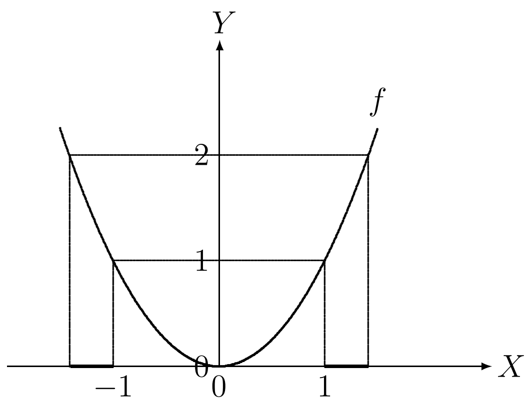


圖 2–2

關於像原, 以下性質常被引用, 利用像原之定義, 不難證得. 以下諸集合 S, T, T_j 皆為 B 之子集, J 為足碼集合, 有限或無限皆可.

(1) $f^{-1}(S \cup T) = f^{-1}(S) \cup f^{-1}(T)$;

(2) $f^{-1}(S \cap T) = f^{-1}(S) \cap f^{-1}(T)$;

(3) $f^{-1}(S \setminus T) = f^{-1}(S) \setminus f^{-1}(T)$;

(4) $f^{-1}\Big(\bigcup_{j\in J} T_j\Big) = \bigcup_{j\in J} f^{-1}(T_j)$;

(5) $f^{-1}\Big(\sum_{j\in J} T_j\Big) = \sum_{j\in J} f^{-1}(T_j)$;

(6) $f^{-1}\Big(\bigcap_{j\in J} T_j\Big) = \bigcap_{j\in J} f^{-1}(T_j)$.

例 1. 設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, 則常數隨機變數

$$X\colon \Omega \to \mathbb{R} : X(\omega) = c$$

乃一隨機變數.

證 對於 $B \in \mathcal{B}$, 若 $c \in B$, 則 $X^{-1}(B) = \Omega$. 若 $c \notin B$, 則 $X^{-1}(B) = \varnothing$. □

例 2. 設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, 若 A 爲一事件, 則函數

$$X = I_A\colon \Omega \to \mathbb{R} : I_A(\omega) = \begin{cases} 1, & \text{若 } \omega \in A, \\ 0, & \text{否則}. \end{cases}$$

乃一隨機變數. [在《微積分》中我們稱 I_A爲 A 之指標函數 (indicator function)[1]]

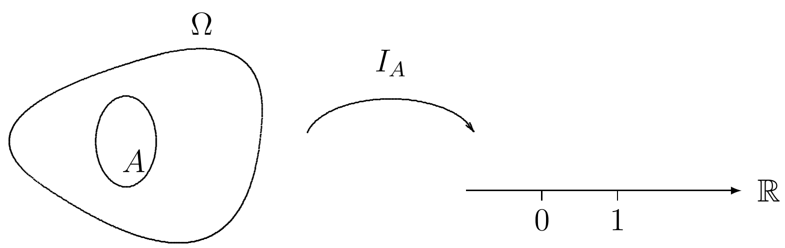


圖 2–3

證 對於 $B \in \mathcal{B}$,

1° 若 $0 \in B,\ 1 \notin B$, 則 $(I_A)^{-1}(B) = \complement A$;

2° 若 $1 \in B,\ 0 \notin B$, 則 $(I_A)^{-1}(B) = A$;

3° 若 $0, 1 \in B$, 則 $(I_A)^{-1}(B) = \Omega$;

4° 若 $0, 1 \notin B$, 則 $(I_A)^{-1}(B) = \varnothing$. □

---

[1]有些學者稱 I_A 爲 A 之特徵函數, 但機率論中“ 特徵函數 ” 一詞另有所指, 詳見第五章.

關於二隨機變數之間的對等關係, 我們將採取『更機率的』規定, 我們稱二隨機變數 X 與 Y 爲**殆必相等**(almost surely equal), 如果存在一機率爲零之事件 N 使得 $\forall\omega \in \Omega\setminus N,\ X(\omega)=Y(\omega)$, 並記爲 $X=Y$ a.s. 或 $X \underset{\text{a.s.}}{=} Y$, 同理, 我們亦可規定" 殆必小於 " $X<Y$ a.s., " 殆必非負 " $X \geq 0$ a.s. 等等, 凡是" 除了一機率爲零之事件外, 都成立的性質 "均可以 a.s. 觀念處理之.

定理 2.2

> 設 X 爲 $(\Omega,\ \mathcal{F},\ P)$ 上一隨機變數, 則函數
>
> $$P_X\colon \mathcal{B} \to [0,1] : P_X(B) = P(X^{-1}(B))$$
>
> 爲 $(\mathbb{R},\ \mathcal{B})$ 上一機率測度.(我們稱此一機率測度爲 X之**機率分配函數**(probability distribution function)).

$\boxed{\text{證}}$ 由於

1° $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$;

2° 若 $\{B_j\}_j$ 爲 $\mathcal{B}$ 上一互斥序列, 則

$$\begin{aligned}
P_X\left(\sum_{j=1}^{+\infty} B_j\right) &= P\left(X^{-1}\left(\sum_{j=1}^{+\infty} B_j\right)\right)\\
&= P\left(\sum_{j=1}^{+\infty} X^{-1}(B_j)\right), \quad \text{(像原性質)}\\
&= \sum_{j=1}^{+\infty} P(X^{-1}(B_j))\\
&= \sum_{j=1}^{+\infty} P_X(B_j).
\end{aligned}$$

$$\begin{tikzcd}
\Omega \arrow[rr, "X"] & & \mathbb{R} \\
\mathcal{F} \arrow[dr, "P"'] & & \mathcal{B} \arrow[ll, "X^{-1}"'] \arrow[dl, "P_X"] \\
& {[0,1]} &
\end{tikzcd}$$

(將像原 X^{-1} 視爲函數)

故知 P_X 爲 $(\mathbb{R},\ \mathcal{B})$ 上一機率測度. $\square$

定理 2.3

> 設 $X\colon \Omega \to \Omega', \mathcal{C} \subset \mathcal{P}(\Omega')$, 則 $\sigma(X^{-1}(\mathcal{C})) = X^{-1}(\sigma(\mathcal{C}))$, 其中
>
> $$X^{-1}(\mathcal{C}) = \{X^{-1}(B) \mid B \in \mathcal{C}\}.$$

$\boxed{\text{證}}$ 證明稍深, 參閱附錄三. $\square$

系 2.4

設 $X\colon \Omega \to \mathbb{R}$, 則 X 爲 Ω 上一隨機變數之充要條件爲

$$\forall a \in \mathbb{R}, \{\omega \in \Omega \mid X(\omega) < a\} \in \mathcal{F}.$$

證 1° 後者顯然爲前者之必要條件.

2° 次證, 若 $\forall a \in \mathbb{R}, \{\omega \mid X(\omega) < a\} \in \mathcal{F}$, 則 $\forall B \in \mathcal{B}, X^{-1}(B) \in \mathcal{F}$. 亦即證明 $X^{-1}(\mathcal{B}) \subset \mathcal{F}$. 爲此, 設 $\mathcal{C} = \{(-\infty, a) \mid a \in \mathbb{R}\}$, 而我們知道 $\sigma(\mathcal{C}) = \mathcal{B}$, 故

$$\begin{aligned} X^{-1}(\mathcal{B}) &= X^{-1}(\sigma(\mathcal{C})) \\ &= \sigma(X^{-1}(\mathcal{C})), \quad (\text{由定理 2.4 }) \\ &\subset \mathcal{F}, \ (\dagger) \end{aligned}$$

($\dagger$) 由原設 $\forall a \in \mathbb{R},\ \{\omega \mid X(\omega) < a\} \in \mathcal{F}$, 知所有 $X^{-1}(-\infty, a)$ 均爲σ 域 $\mathcal{F}$ 之元素. □

隨機變數的設計是將非數量性集合 Ω 轉換至 $\mathbb{R}$ 上. 自然我們會想到這些經由 X 映過來的點集之“機率”又是如何? 例如: Ω 表全世界的大學生, X 爲身高函數, 我們會問: 身高在 $[170, 175]$ (單位: 公分) 之機率爲若干? 更明確地說, $\mathbb{R}$ 上是否也可以建立一個機率空間? 首先, 選擇什麼集合做爲 $\mathbb{R}$ 上的 σ 域最爲恰當? 如果選擇 $\{\varnothing, \mathbb{R}\}$, 實在沒什麼用, 如果選擇 $\mathcal{P}(\mathbb{R})$, 由於其內存在實變函數論中所謂不可測集合 A, 我們不能測度其長度 (Lebesgue 觀念), 對於一般函數 f, 也不能求積分 $\int_A f$, 我們希望至少能討論 $\mathbb{R}$ 上某區間 I 上之機率爲若干, 因此, 其 σ 域最好是由所有區間之集合 $\{I \mid I \text{ 爲實數區間}\}$ 所產生者, 也就是一維 Borel 域 $\mathcal{B}$.(另一個選擇是 $\mathcal{M}$, 這是所有 Lebesgue 可測集合所成之 σ 域, 但在討論隨機變數之變換時會遭遇一些困難, 而它與 $\mathcal{B}$ 差異很微小, 因此我們仍選擇 $\mathcal{B}$). 最後我們利用定理 2.2 中的 P_X 做爲機率測度, 就可以得到 $\mathbb{R}$ 上的機率空間 $(\mathbb{R}, \mathcal{B}, P_X)$. 回頭來看隨機變數 X, 如果它不滿足 $\forall B \in \mathcal{B},\ X^{-1}(B) \in \mathcal{F}$, 那麼定理 2.2 中所界定的 P_X 就有問題, 因爲 $P_X(B) = P(X^{-1}(B))$ 中的 $X^{-1}(B)$ 未必屬於 $\mathcal{F}$, 其機率 $P(X^{-1}(B))$ 也就沒有意義. 這就是爲什麼定義 2.1 中的隨機變數 X 需要條件 “ $\forall B \in \mathcal{B},\ X^{-1}(B) \in \mathcal{F}$ ”.

在機率論中, 有部分符號與純粹數學常用符號相異, 而且經常使用, 希望讀者細心體會, 自然熟能生巧, 運用自如.

設 X 與 Y 爲 Ω 上之隨機變數, $a, b \in \mathbb{R}$, 我們規定

$$\begin{aligned}
&\{X \in B\} = X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\}, \\
&\{X \leq a\} = X^{-1}(-\infty,\, a] = \{\omega \in \Omega \mid X(\omega) \leq a\}, \\
&\{X < a\} = X^{-1}(-\infty,\, a) = \{\omega \in \Omega \mid X(\omega) < a\}, \\
&\{a \leq X \leq b\} = X^{-1}[a,\, b] = \{\omega \in \Omega \mid a \leq X(\omega) \leq b\}. \\
&\{X \leq a, Y \leq b\} = \{X \leq a\} \cap \{Y \leq b\},
\end{aligned}$$

其餘類推.

上面已詳細說明隨機變數定義中需要條件 $\forall B \in \mathcal{B},\ X^{-1}(B) \in \mathcal{F}$, 這種定義方法, 稱爲敍述性定義 (descriptive definition). 在往後有關期望值的討論中, 敍述性定義非但毫無幫助, 甚至常是一種阻礙. 因此, 有必要尋找其對等條件, 也就是所建構性定義 (constructive definition).

定義 2.5

設 $(\Omega,\, \mathcal{F},\, P)$ 爲一機率空間.

(1) 若 $A_1,\, A_2,\, \cdots,\, A_n$ 均爲事件, 則指標函數 $I_{A_1},\, I_{A_2},\, \cdots,\, I_{A_n}$ 之線性組合

$$X = \textstyle\sum_{j=1}^{n} x_j I_{A_j},$$

稱爲一**簡單隨機變數**(simple r.v.), 其中 $x_1,\, \cdots,\, x_n \in \mathbb{R}$.

(2) 設 X 爲 Ω 上一隨機變數, 我們稱函數

$$X^+ \colon \Omega \to \mathbb{R}_+ : X^+(\omega) = \max\{X(\omega),\, 0\}$$

爲 X 之**正部分**(positive part). 而函數

$$X^- \colon \Omega \to \mathbb{R}_+ : X^-(\omega) = \max\{-X(\omega),\, 0\}$$

則稱爲 X 之**負部分** (negative part), 顯然 $X = X^+ - X^-$.

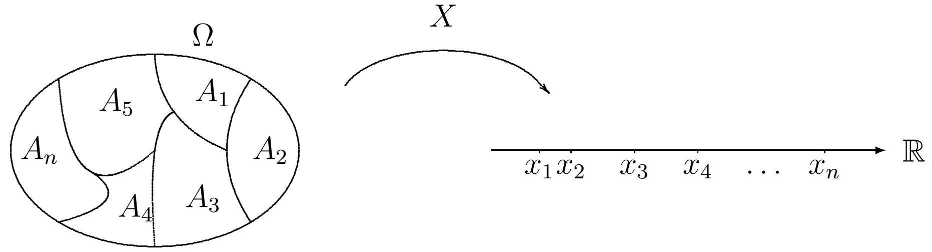


圖 2–4

如果隨機變數 X 之值域爲一有限集合, 我們可令

$$X(\Omega) = \{x_1,\, x_2,\, \cdots,\, x_n\},\ A_j = X^{-1}\{x_j\},\ \forall j \in \{1, \cdots, n\}$$

(參見圖 2–4), 則 X 可表爲指標函數 I_{A_j} 之線性組合 $X = \sum_{j=1}^{n} x_j I_{A_j}$, 因此 X 爲一簡單隨機變數; 反之, 若 X 爲一簡單隨機變數, 顯然其值域爲有限.

定理 2.6

設 $(\Omega,\, \mathcal{F},\, P)$ 爲一機率空間, 函數 $X\colon \Omega \to \mathbb{R}$, 則以下二敍述爲對等:

I. X 爲一非負隨機變數;

II. 存在一非負簡單隨機變數序列 $\{X_n\}_n$ 使得 $X_n \uparrow X$, 意即:

$$\forall n \in \mathbb{N},\ X_n \le X_{n+1} \text{ 且 } \forall \omega \in \Omega,\ \lim_{n\to+\infty} X_n(\omega) = X(\omega).$$

證 證明冗長, 參閱附錄四. □

本定理將在第四章中用以界定『期望值』.

定理 2.7

設 X 爲 $(\Omega,\, \mathcal{F},\, P)$ 上一隨機變數.

(1) 若函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲 $(\mathcal{B}, \mathcal{B})$ 可測 (亦稱 h 爲 Borel 可 測), 則合成函數 $h \circ X$ 爲 $(\Omega,\, \mathcal{F},\, P)$ 上一隨機變數;

(2) 若函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲連續, 則 $h \circ X$ 爲 $(\Omega,\, \mathcal{F},\, P)$ 上一隨機變數.

證 (1) 由於

$$\begin{aligned} B \in \mathcal{B} &\Rightarrow\ h^{-1}(B) \in \mathcal{B}, (\text{因 } h \text{ 爲可測}) \\ &\Rightarrow\ X^{-1}(h^{-1}(B)) \in \mathcal{F}, (\text{因 } X \text{ 爲 r.v. }) \\ &\Rightarrow\ (h \circ X)^{-1}(B) \in \mathcal{F}. \end{aligned}$$

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "h \circ X"'] & \mathbb{R} \arrow[d, "h"] \\ & \mathbb{R} \end{tikzcd}$$

是以 $h \circ X$ 爲 $(\Omega,\, \mathcal{F},\, P)$ 上一隨機變數.

(2) 令 $Y = h \circ X$, 利用系 2.4, 只需證明: $\forall a \in \mathbb{R}, Y^{-1}(-\infty, a) \in \mathcal{F}$; 而

$$
\begin{aligned}
(-\infty, a) \text{ 爲一開集 } &\Rightarrow\ h^{-1}(-\infty, a) \text{ 爲一開集, (因 } h \text{ 爲連續)} \\
&\Rightarrow\ h^{-1}(-\infty, a) = \bigcup_{j=1}^{+\infty} I_j, {}^{[\star]} \\
&\Rightarrow\ Y^{-1}(-\infty, a) = X^{-1}(h^{-1}(-\infty, a)) = X^{-1}\Big(\bigcup_{j=1}^{+\infty} I_j\Big) \\
&\qquad = \bigcup_{j=1}^{+\infty} X^{-1}(I_j) \in \mathcal{F}.
\end{aligned}
$$

[⋆] I_j 皆爲開區間, 參見 Wheeden & Zygmund, Measure and Integral, Theorem 1.10. □

系 2.8

> 設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, 則 $X + c$, cX, X^n, $[X]$ 及 $|X|^r$ 均爲隨機變數, 其中 c 爲一常數, n 爲一自然數, r 爲一正數, $[\ \ ]$ 爲高斯符號.

證

1° 令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = x + c$, 顯然 h 爲連續且 $h \circ X = X + c$;

2° 令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = cx$, 顯然 h 爲連續且 $h \circ X = cX$;

3° 令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = x^n$, 顯然 h 爲連續且 $h \circ X = X^n$;

4° 令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = [x]$, h 雖不爲連續但爲可測且 $h \circ X = [X]$;

5° 令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = |x|^r$, 顯然 h 爲連續且 $h \circ X = |X|^r$. □

如果 X 與 Y 均爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, 則 $X + Y$, $X - Y$, XY 及 X/Y 亦均爲隨機變數, (最後一個, 二隨機變數之商, 有點兒限制), 我們將在下一章中討論和證明.

§ 2.2 隨機變數之分配函數

隨機變數將機率空間 $(\Omega, \mathcal{F}, P)$ 轉換爲一量化機率空間 $(\mathbb{R}, \mathcal{B}, P_X)$ 後, 對於機率之計算及討論仍限於理論階段, 在機率分配函數 P_X 問世之前, 已有所謂『機率密度函數』及『累積分配函數』這兩個較爲容處理的觀念, 我們將用以計算機率、期望值及其他相關問題.

定義 2.9

設 X 為 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, 則函數

$$F_X : \mathbb{R} \to [0, 1] : F_X(x) = P\{X \le x\} = P_X(-\infty, x]$$

稱為 X 之(**累積**) **分配函數**((cumulative) distribution function) 簡為 c.d.f. 或 d.f., 若無誤解之可能時, F_X 亦常寫為 F.

定理 2.10

設 F 為隨機變數 X 之分配函數, 則

(1) F 為一不減函數 (non-decreasing function);

(2) F 為一右連續函數;

(3) $\lim\limits_{x\to-\infty} F(x) = 0$; $\lim\limits_{x\to+\infty} F(x) = 1$.

(有些學者寫為 $F(-\infty) = 0, F(+\infty) = 1$, 有妄用數學符號之虞).

證 (1) 若 $x_1, x_2 \in \mathbb{R}$ 且 $x_1 < x_2$, 則 $(-\infty, x_1] \subset (-\infty, x_2]$, 是以

$$F(x_1) = P_X(-\infty, x_1] \le P_X(-\infty, x_2] = F(x_2).$$

(2) 由於

$$\begin{aligned} F \text{ 為右連續} &\Leftrightarrow \forall p \in \mathbb{R}, \lim_{x\to p+} F(x) = F(p) \\ &\Leftrightarrow (\forall p \in \mathbb{R})(\forall \{x_n\}_n \downarrow p \;\Rightarrow\; \lim_{n\to+\infty} F(x_n) = F(p)). \end{aligned}$$

故只需證明最末敘述即可. 因為

$$\begin{aligned} \lim_{n\to+\infty} F(x_n) &= \lim_{n\to+\infty} P_X(-\infty, x_n] \\ &= P_X(\lim_{n\to+\infty}(-\infty, x_n]), (\text{因 } P_X \text{ 為一機率測度}) \\ &= P_X\Big(\bigcap_{n=1}^{+\infty}(-\infty, x_n]\Big), (\text{因 } \{x_n\}_n \downarrow p) \\ &= P_X((-\infty, p]) = F(p). \end{aligned}$$

是以 F 為一右連續函數.

(3) 由於

$$\lim_{x\to-\infty} F(x) = 0 \Leftrightarrow (\forall \{x_n\}_n \downarrow -\infty \;\Rightarrow\; \lim_{n\to+\infty} F(x_n) = 0).$$

但因

$$\begin{aligned}\lim_{n\to+\infty} F(x_n) &= \lim_{n\to+\infty} P_X(-\infty,\, x_n] = P_X(\lim_{n\to+\infty}(-\infty,\, x_n]),\\ &= P_X(\bigcap_{n=1}^{+\infty}(-\infty,\, x_n]) = P_X(\varnothing) = 0.\end{aligned}$$

是以證得 $\lim_{x\to-\infty} F(x) = 0$. 同理可證 $\lim_{x\to+\infty} F(x) = 1$. □

本定理之逆敘述亦爲眞, 亦即, 若函數 $F\colon \mathbb{R} \to [0,\,1]$ 滿足條件 (1) 不減, (2) 右連續, (3) $\lim\limits_{x\to-\infty} F(x) = 0$; $\lim\limits_{x\to+\infty} F(x) = 1$, 則存在一隨機變數 X 使得 F 爲 X 之分配函數. 證明係利用測度論觀念, 從略; 讀者可參閱 Laha & Rohagi [10], p.5.

爲使讀者易於掌握分配函數之性質, 我們先舉一範例.

例 1. 設 $F_j\colon \mathbb{R} \to [0,1], j = 1,2,3,4,$

$$F_1(x) = \begin{cases} 0, & 若\ x < 0, \\ x/2, & 若\ 0 \le x < 2, \\ 1, & 若\ 2 \le x; \end{cases}$$

$$F_2(x) = \begin{cases} 0, & 若\ x < 3, \\ 1, & 若\ x \ge 3; \end{cases}$$

$$F_3(x) = \begin{cases} 0, & 若\ x < 0, \\ 1 - a^x, & 若\ x \ge 0, \end{cases} \qquad (內\ 0 < a < 1;)$$

$$F_4(x) = \begin{cases} 0, & 若\ x < 0, \\ (1+x^2)/2, & 若\ 0 \le x < 1, \\ 1, & 若\ 1 \le x. \end{cases}$$

試繪出各 F_j 之圖形, 並問各 F_j 是否爲一累積分配函數?

解

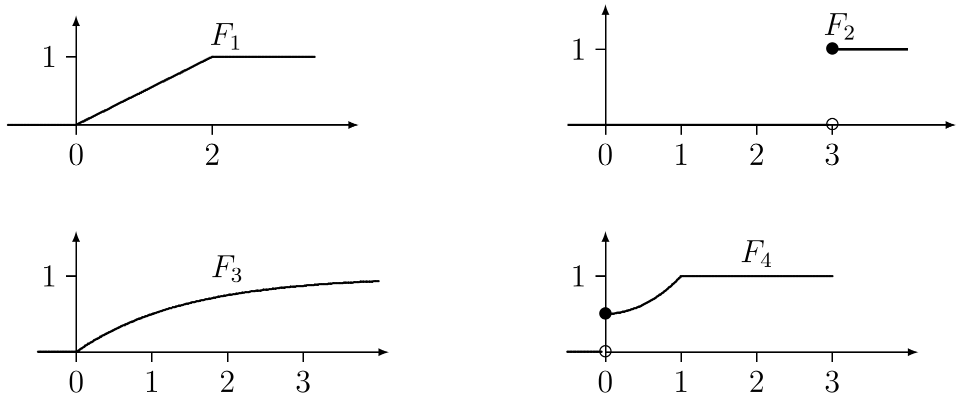


圖 2–5

此四函數皆滿足 (1) 不減, (2) 右連續, (3) $\lim\limits_{x\to-\infty} F(x) = 0$; $\lim\limits_{x\to+\infty} F(x) = 1$, 因此皆爲累積分配函數. □

定理 2.11

設 F 爲隨機變數 X 之分配函數. 則

(1) 若 $a, b \in \mathbb{R}$ 且 $a < b$, 則 $P\{a < X \leq b\} = F(b) - F(a)$.

(2) F 於任一點之左極限均存在, (但未必等於 F 在該點之值).

(3) 若 $p \in \mathbb{R}$, 則 $P\{X = p\} = F(p) - F(p^-)$.

證 (1)

$$\begin{aligned} P\{a < X \leq b\} &= P\{X \in (a, b]\} \\ &= P_X(a, b] \\ &= P_X(-\infty,\, b] - P_X(-\infty,\, a] \\ &= F(b) - F(a). \end{aligned}$$

(2) 因 F 爲一有界之不減函數, 故 $\forall p \in \mathbb{R}$,

$$F(p^-) = \sup\{F(x) \mid x < p\}.$$

(3) 設 $\forall \{x_n\}_n \uparrow p$, 則 $\{p\} = \bigcap\limits_{n=1}^{+\infty} (x_n, p] = \lim\limits_{n\to+\infty} (x_n, p]$, 是以

$$\begin{aligned} P\{X = p\} &= P\{X \in \{p\}\} \\ &= P_X\{p\} = P_X(\lim_{n\to+\infty} (x_n, p]) \\ &= \lim_{n\to+\infty} P_X(x_n, p] \\ &= \lim_{n\to+\infty} (F(p) - F(x_n)), (\text{由 } (1)) \\ &= F(p) - \lim_{n\to+\infty} F(x_n) \\ &= F(p) - F(p^-). \end{aligned}$$

隨機變數 X 之機率分配函數 P_X 與累積分配函數 F_X 所欲解決之問題完全一致, 二者均用以表現在實數系 $\mathbb{R}$ 上之機率測度而已. 在二十世紀之前, 測度觀念尚不完全清楚, 因此累積分配函數所扮演之角色十分重要, 但在二十世紀之後, P_X 已取代了 F_X 的一些功能.

§ 2.3 隨機變數之分類

由於計算機率的方式, 在傳統上有點機率及區間機率, 也就是所謂離散型及連續型之別, 我們將以更嚴密的方式界定如下:

定義 2.12

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, F 爲 X 之分配函數.

(1) 我們稱 X 爲**離散型**(discrete type), 如果存在一至多可數集合 $A \subset \mathbb{R}$ 使得 $P\{X \in A\} = 1$. 此時, 令函數

$$f_X\colon \mathbb{R} \to [0,1] : f_X(x) = P\{X = x\} = \begin{cases} P\{X = x\}, & \text{若 } x \in A, \\ 0, & \text{若 } x \in \mathbb{R} \setminus A, \end{cases}$$

而稱其爲 X 之(**機率**) **密度函數** (probability density function or probability mass function), 簡爲 p.d.f. 或 p.m.f. 若無誤解之可能, f_X 亦 常寫爲 f.

(2) 我們稱 X 爲**連續型**(continuous type), 如果其累積分配函數 F 爲一連續函數.

(3) 我們稱 X 爲**絕對連續型**(absolutely continuous type), 如果其累積分配函數 F 爲一絕對連續函數. 意即: $F\colon \mathbb{R} \to [0,\ 1]$ 滿足

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall a_1, b_1, \cdots, a_n, b_n \in \mathbb{R},\ a_1 < b_1 < a_2 \cdots < a_n < b_n,$$
$$\Big(\textstyle\sum_{j=1}^n (b_j - a_j) < \delta \Rightarrow \sum_{j=1}^n |F(b_j) - F(a_j)| < \epsilon \Big).$$

註: 當 X 爲離散型時, 我們稱上述的 f_X 爲機率密度函數, 然而此一名詞並不是被普遍接受的, 有些人稱其爲機率函數, 有些人稱其爲機率質量函數. 至於絕對連續型之隨機變數, 上述定義雖然漂亮, 對於未學過實變函數論的讀者而言實在太深, 稍後我們有更方便的對等條件以便使用.

定理 2.13

設 X 爲一離散型隨機變數, 即存在一有限或可數集合 $A \subset \mathbb{R}$ 使得 $P\{X \in A\} = 1$. 若 A 無聚點 (accumulation point), 則 X 之分配函數 F 爲一階梯函數.

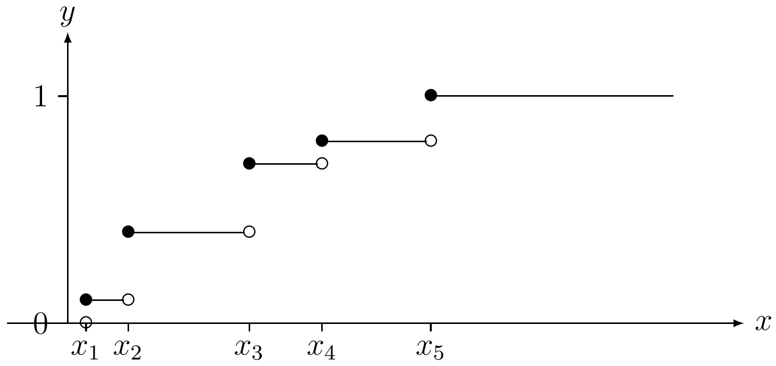


圖 2–6 離散型隨機變數之累積分配函數

證 只證 A 爲有限之情形, 餘仿之. 設 $A = \{x_1, \cdots, x_n\}$, $x_1 < \cdots < x_n$, 次令 $x_0 = -\infty$, $x_{n+1} = +\infty$, 則 $\forall t \in (x_{k-1}, x_k)$, $k = 1, \cdots, n+1$,

$$F(t) = P\{X \leq t\} = \begin{cases} 0, & \text{若 } k = 1, \\ f(x_1) + \cdots + f(x_{k-1}), & \text{若 } k > 1, \end{cases}$$

顯然 F 在每一開區間 (x_{k-1}, x_k) 上均爲常數函數, 故 F 爲一階梯函數. □

定理 2.14

隨機變數 X 爲連續型之充要條件爲 $\forall p \in \mathbb{R}$, $P\{X = p\} = 0$.

證 隨機變數 X 爲連續型

$\Leftrightarrow$ F 爲連續函數

$\Leftrightarrow \forall p \in \mathbb{R}$, F 爲左連續於 p, 且 F 爲右連續於 p

$\Leftrightarrow \forall p \in \mathbb{R}$, F 爲左連續於 p, (因 F 之右連續性恆眞)

$\Leftrightarrow \forall p \in \mathbb{R}$, $F(p) = F(p^-)$

$\Leftrightarrow \forall p \in \mathbb{R}$, $P\{X = p\} = 0$, (定理 2.11) □

定理 2.15

隨機變數 X 爲絕對連續型之充要條件爲存在一 Lebesgue 可積函數 $f_X \colon \mathbb{R} \to \mathbb{R}_+$ 滿足 $\forall B \in \mathcal{B}$, $P\{X \in B\} = \int_B f_X(x)\, dx$.

證 證明屬於實變函數論範疇, 有興趣讀者請參閱 Royden [15], §5.4. □

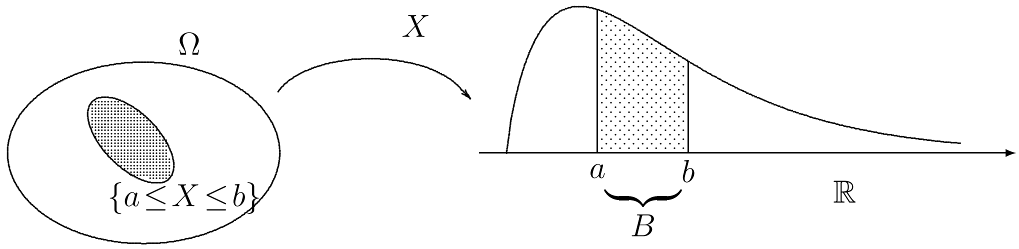


圖 2–7

定義 2.16

設 X 爲一絕對連續型隨機變數, 則定理 2.15 中之 Lebesgue 可 積函數 $f_X\colon \mathbb{R} \to \mathbb{R}_+$ 稱爲 X 之(**機率**) **密度函數**(probability density function). 簡爲 p.d.f. 若無誤解之可能, f_X 亦常寫爲 f.

定理 2.17

絕對連續型之隨機變數必爲連續型.

證

$$\begin{aligned} X \text{ 爲絕對連續型 } &\Leftrightarrow F \text{ 爲絕對連續函數} \\ &\Rightarrow F \text{ 爲連續函數, (由絕對連續之定義)} \\ &\Leftrightarrow X \text{ 爲連續型.} \end{aligned}$$ □

一絕對連續型之隨機變數必爲連續型, 但其逆不眞, 因爲存在連續之分配函數 F, 但其並非絕對連續, (這種函數係利用 Cantor 三進點集而定義得來. 詳見第六節奇異分配). 不過這種非絕對連續型之連續型隨機變數, 對於初學機率論者用途不大, 是以今後在沒有誤解之可能時, 我們常稱絕對連續型之隨機變數與其分配爲連續型.

定理 2.18

設 X 爲絕對連續型隨機變數, f 爲其密度函數, $g\colon \mathbb{R} \to \mathbb{R}_+$, 則 g 爲 X 之密度函數之充要條件爲 $f = g$ a.e.[2]

[2] a.e.=almost every (殆遍), 意即: 集合 $\{x \in \mathbb{R} \mid f(x) \neq g(x)\}$ 爲零測度.

證 1° 若 $f = g$ a.e., 則 $\forall B \in \mathcal{B}$,

$$\int_B g(t)dt = \int_B f(t)dt = P\{X \in B\}.$$

是以 g 亦爲 X 之密度函數.

2° 若 g 爲 X 之密度函數, 則 $\forall B \in \mathcal{B}$,

$$\int_B g(t)dt = P\{X \in B\} = \int_B f(t)dt.$$

利用 Lebesgue 積分性質, (見 Royden [16], §5.4.), 可證得: $f = g$ a.e. □

定理 2.19

設 F 爲絕對連續型隨機變數 X 之分配函數, $f\colon \mathbb{R} \to \mathbb{R}_+$, 則

$$(f \text{ 爲 } X \text{ 之密度函數}) \Leftrightarrow \left(\forall x \in \mathbb{R},\ F(x) = \int_{-\infty}^{x} f(t)\,dt\right).$$

證 1° 若 f 爲 X 之密度函數. 則 $\forall x \in \mathbb{R}$,

$$F(x) = P\{X \leq x\} = P\{X \in (-\infty, x]\} = \int_{-\infty}^{x} f(t)\,dt.$$

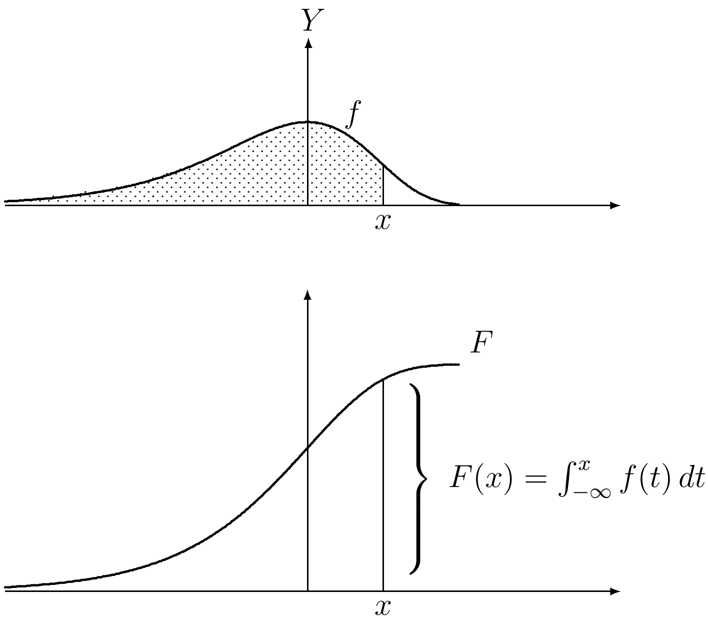


圖 2–8 由密度函數求分配函數

2° 因 X 爲絕對連續型, 故必存在一 Lebesgue 可積函數 $g\colon \mathbb{R} \to \mathbb{R}_+$ 滿足 $\forall B \in \mathcal{B},\ P\{X \in B\} = \int_B g(t)\,dt$. 亦即有

$$\forall x \in \mathbb{R}, F(x) = P\{X \le x\} = \int_{-\infty}^{x} g(t)\,dt,$$

故

$$\begin{aligned}\forall x \in \mathbb{R}, F(x) &= \int_{-\infty}^{x} f(t)\,dt \\ &\Rightarrow \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} g(t)\,dt, \forall x \in \mathbb{R} \\ &\Rightarrow \int_{-\infty}^{x} (f(t) - g(t))\,dt = 0, \forall x \in \mathbb{R} \\ &\Rightarrow f = g \text{ a.e. }, (\text{參閱 Royden [16] §5.3}) \\ &\Rightarrow f \text{ 爲 } X \text{ 之密度函數. (定理 2.18)} \square\end{aligned}$$

定理 2.20

設 F 爲絕對連續型隨機變數 X 之分配函數, 若令

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} F'(x), & \text{若 } x \in A, \\ 0, & \text{若 } x \notin A, \end{cases}$$

內 $A = \{x \in \mathbb{R} \mid F'(x) \text{ 存在}\}$. 則 f 爲 X 之密度函數.

證 因爲 F 爲絕對連續函數, 故由 Royden [16] §5.4 知, F 爲 a.e. 可微, 且

$$\forall x \in \mathbb{R},\ F(x) = \int_{-\infty}^{x} F'(t)\,dt,$$

是以

$$\forall x \in \mathbb{R},\ F(x) = \int_{-\infty}^{x} f(t)\,dt, (\text{因 } f = F' \text{ a.e.})$$

由定理 2.18 知函數 f 爲 X 之密度函數. $\square$

本定理之用途是: 當 F 爲一絕對連續型分配函數, 我們可利用上述公式求得密度函數. 許多初等機率論或統計學之書本常有以下的敍述: 若 F 爲絕對連續型隨機變數 X 之分配函數, 則 $f(x) = F'(x)$ 爲 X 之密度函數. 這是錯誤的, 因爲有很多 F 僅爲 a.e. 可微, 未必爲可微. 本章第五節的均勻分配就是最好的反例.

§ 2.4 常用離散型隨機變數

(一) 二項分配(binomial distribution)

一隨機試驗若恰有二種可能之結果, 我們稱其爲一**Bernoulli 試行**(trial). 例如某君在棒球場上之一次打擊, 其結果爲安打或非安打; 某甲之一次投籃, 只有『進』與『不進』; 我們丟一次銅板, 其結果只可能爲正面或反面. 更精確地說, 一袋子, 內白球與黑球各若干 (未必相等), 隨機抽取一球, 取得白球曰成功, 簡爲 S, 其機率爲 p; 取得黑球曰失敗, 簡爲 F, 其機率爲 $q = 1-p$. 若以機率空間模式表之, 則 $\Omega = \{S, F\},\ p = P\{S\},\ q = P\{F\}$, 如果函數 X 表成功之次數, 應有

$$X\colon \Omega \to \mathbb{R} : X(S) = 1,\ X(F) = 0,$$

$$P\{X = 1\} = P\{S\} = p,$$

$$P\{X = 0\} = P\{F\} = q = 1 - p.$$

其次, 若重複實行 Bernoulli 試行 n 次, 其中任何二試行均無關聯, 以袋子抽球爲例, 吾人以"放回"方式隨機抽取 n 個, 則

- 樣本空間爲 $\Omega = \{(a_1, a_2, \cdots, a_n) \mid \forall j \in \{1, \cdots, n\}, a_j \in \{S, F\}\}$.
- σ 域爲 $\mathcal{F} = \mathcal{P}(\Omega)$, (因爲 $\#\Omega = 2^n$ 有限) .
- 機率測度爲 $P\colon \mathcal{F} \to [0, 1] : P\{(a_1, \cdots, a_n)\} = p^x q^{n-x}$.

  (內 x 爲 $(a_1, a_2, \cdots, a_n)$ 中成功之次數).

令

$$X\colon \Omega \to \mathbb{R} : X(a_1, a_2, \cdots, a_n) = \text{“}(a_1, a_2, \cdots, a_n) \text{ 中成功之次數”},$$

則

(1) X 爲一隨機變數, (此因 $\mathcal{F} = \mathcal{P}(\Omega)$);

(2) X 爲離散型, 因若令 $A = X(\Omega) = \{0, 1, \cdots, n\}$, 則 $P\{X \in A\} = P(\Omega) = 1$. 此時, 函數

$$f\colon \mathbb{R} \to [0, 1] : f(x) = P\{X = x\} = \begin{cases} \dbinom{n}{x} p^x q^{n-x}, & \text{若 } x \in A, \\ 0, & \text{若 } x \in \mathbb{R} \setminus A, \end{cases}$$

乃爲 X 之機率密度函數. 其中, 當 $x \in A$ 時, $P\{X = x\} = \dbinom{n}{x} p^x q^{n-x}$ 之理由是: 由於事件

$$\{X = x\} = \{(a_1, a_2, \cdots, a_n) \in \Omega \mid \text{恰有 } x \text{ 個 } a_j \text{ 爲 } S, \text{其餘爲 } F\}$$

中之任一元素, 例如,

$$(\underbrace{S, \cdots, S}_{x\text{ 項}}, \underbrace{F, \cdots, F}_{n-x\text{ 項}})$$

發生之機率均為 $p^x(1-p)^{n-x}$, 而 $\{X = x\}$ 內共有 $\binom{n}{x} = \dfrac{n!}{x!(n-x)!}$ 種情形. 此種分配稱為**二項分配** 或 $B(n, p)$ **分配**, 並稱 X 為 **二項隨機變數**(binomial random variable); 常寫為 $X \sim B(n, p)$. 當 $X \sim B(1, p)$ 時, 亦稱 X 具 Bernoulli 分配, 稍早提及之指標函數 I_A 即具 $B(1, p)$ 分配, 其中 $p = P(A)$.

例 1.

設 $X \sim B(10, 0.27)$, 試求 X 之 d.f. 及 p.d.f. 並繪出其圖形.

$\boxed{\text{解}}$ 之密度函數為

$$f: \mathbb{R} \to [0, 1] : f(x) = \begin{cases} \binom{10}{x} 0.27^x 0.73^{10-x}, & \text{若 } x \in \{0, \cdots, 10\}, \\ 0, & \text{否則}. \end{cases}$$

我們可利用以下方法以求上述 p.d.f. 及 d.f. 之值.

(1) 利用計算器.
(2) 利用二項分配之表 (table); 許多機率論及統計學之書均附有二項 分配之累積分配函數之表, 當然 n 及 p 都有限制, 本例中 $p = 0.27$, 很多書本就找不到, 換言之, 此一方法不是萬靈丹.
(3) 利用電腦; 很多軟體提供極方便的計算並繪出相當美觀的圖形, 比較簡單的有 Math-CAD, Excel, Minitab 等, 另有統計專用的軟 體 , 如 SAS, SPSS 等使用起來十分方便而精確, 讀者不妨參考該軟體之手冊自可迎刃而解. (以下數值係利用 Excel 算出).

$F(x)$	$f(x)$
$F(0) = 0.0430$,	$f(0) = 0.0430$,
$F(1) = 0.2019$,	$f(1) = 0.1590$,
$F(2) = 0.4665$,	$f(2) = 0.2646$,
$F(3) = 0.7274$,	$f(3) = 0.2609$,
$F(4) = 0.8963$,	$f(4) = 0.1689$,
$F(5) = 0.9713$,	$f(5) = 0.0750$,
$F(6) = 0.9944$,	$f(6) = 0.0231$,
$F(7) = 0.9993$,	$f(7) = 0.0049$,
$F(8) = 0.9999$,	$f(8) = 0.0007$,
$F(9) = 1.0000$,	$f(9) = 0.0001$,
$F(10) = 1.0000$,	$f(10) = 0.0000$,

是以 X 之 p.d.f. 及 c.d.f. 圖形分別如下:

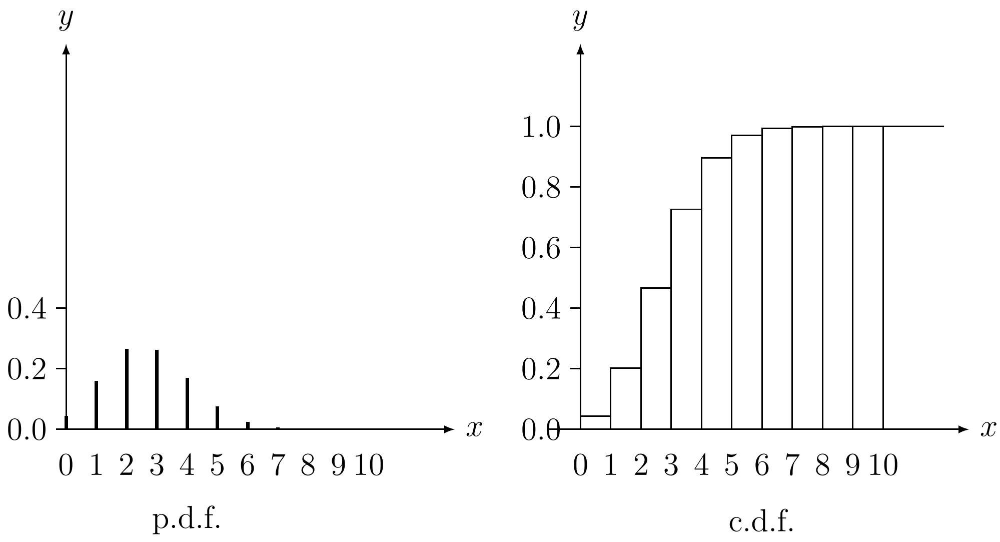


圖 2–9

例 2. 某君喜愛籃球, 聲稱其每投六球至少進一球之機率爲 0.9, 試求其每投一球之進球率爲若干? 並求投六球至少進三球之機率.

解 (1) 設每投一球之進球率爲 p, 投 $n=6$ 球之進球數爲隨機變數 X, 顯然 $X \sim B(6,p)$, 其密度函數爲

$$f\colon \mathbb{R} \to [0,\,1] : f(x) = P\{X=x\} = \begin{cases} \dbinom{6}{x} p^x q^{6-x}, & \text{若 } x \in \{0,1,2,3,4,5,6\}, \\ 0, & \text{否則}, \end{cases}$$

由題意知每投六球至少進一球之機率爲 0.9, 故

$$\begin{aligned} P\{X \geq 1\} = 0.9 &\Leftrightarrow 1 - P\{X \geq 1\} = 1 - 0.9 = 0.1 \\ &\Leftrightarrow \binom{6}{0} p^0 (1-p)^{6-0} = 0.1 \\ &\Leftrightarrow 1 - p = 0.6813 \Leftrightarrow p = 0.3187. \end{aligned}$$

是以每投一球之進球率爲 0.3187.

(2) 投六球至少進三球之機率爲

$$\begin{aligned} P\{X \geq 3\} &= 1 - [f(0) + f(1) + f(2)] \\ &= 1 - \Big[\binom{6}{0} 0.3187^0 0.6813^6 + \binom{6}{1} 0.3187^1 0.6813^5 \\ &\qquad + \binom{6}{2} 0.3187^2 0.6813^4 \Big] \\ &= 1 - (0.1000 + 0.2807 + 0.3283) = 0.2911. \end{aligned}$$

□

(二) **Poisson 分配**

1837 年 Poisson 發表了一篇有關機率的文章: Recherches sur la probabilité des jugements en matière criminelle et en matière civile (刑事案件與民事案件裁判之機率的研究), 首次提到 Poisson 分配, 此一分配適用於某一事件在較短的時間 (或較小之空間) 發生次數不多, 而其實驗次數又較大之情況下, 例如: 在觀察某一都市過去一年中之死亡車禍, 某日死亡人數爲 k 之機率 $P\{X = k\} =?$ $k = 0, 1, 2, \cdots$, 在該文中, Poisson 提到此一分配可應用於二項分配機率之近似值.

我們知道, 當 n 很大時, $\binom{n}{x}p^x q^{n-x}$ 的計算十分麻煩. 在十八、九世紀, 計算工具還十分原始, 對於如此麻煩的計算實在非想個辦法解決不可; 在十八世紀初葉, de Moivre 首先嘗試以分析的方法解決當 $p = \frac{1}{2}$ 時, $\binom{n}{x}p^x q^{n-x}$ 之近似值. 到了十九世紀初, Laplace 更將這種方法推廣到 $p \neq \frac{1}{2}$ 之情形, 而發展出所謂的常態分配, (我們將在下一節論及這種分配). Poisson 分配問世之後, 有人將它用來解決 $\binom{n}{x}p^x q^{n-x}$ 之近似值, 特別適用於 n 很大而 p 或 q 很小之時. 今日 Poisson 分配的面貌則如下述:

設$(\Omega,\, \mathcal{F},\, P)$ 爲一機率空間. 若函數 $X\colon \Omega \to \mathbb{R}$ 滿足

$$\forall x \in \mathbb{Z}_+ = \{0,\, 1,\, 2,\, \cdots\}, P\{X = x\} = e^{-\lambda}\frac{\lambda^x}{x!}, (\text{內常數 } \lambda > 0).$$

則稱 X 爲一 Poisson **隨機變數**或 $P(\lambda)$ **隨機變數**, 其中 λ 稱爲此分配之**參數**或**母數**(parameter). 由於

$$P\{X \in \mathbb{Z}_+\} = \sum_{x=0}^{+\infty} P\{X = x\} = \sum_{x=0}^{+\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

故知 X 爲離散型, 其密度函數顯然爲

$$f\colon \mathbb{R} \to [0,\, 1] : f(x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{x!}, & \text{若 } x \in \mathbb{Z}_+, \\ 0, & \text{否則.} \end{cases}$$

註: 利用《微積分》中 $e^x = \displaystyle\sum_{n=1}^{+\infty} \frac{x^n}{n!}$, $\forall x \in \mathbb{R}$, 我們可證得

$$\sum_{x=0}^{+\infty} f(x) = \sum_{x=0}^{+\infty} e^{-\lambda}\frac{\lambda^x}{x!} = 1.$$

例 3. 設 $X \sim P(8)$, 試求 $f_X(x)$, $x \in \{0, 1, \cdots, 18\}$, 之值. 並繪出 f_X 之圖形.

$\boxed{\text{解}}$ 利用計算器或電腦數學 (或統計) 軟體等可求得

x	$f_X(x)$	x	$f_X(x)$
0	0.0003	10	0.0993
1	0.0027	11	0.0722
2	0.0107	12	0.0481
3	0.0286	13	0.0296
4	0.0573	14	0.0169
5	0.0916	15	0.0090
6	0.1221	16	0.0045
7	0.1396	17	0.0021
8	0.1396	18	0.0009
9	0.1241		

其圖形如下:

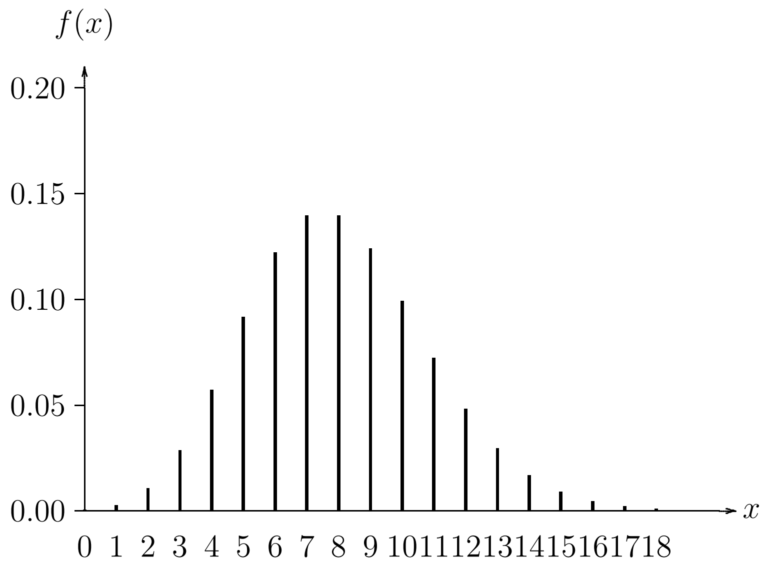


圖 2–10 Poisson 分配 $P(8)$ 之密度函數 □

例 4. 二次大戰期間, 英國人在德機一輪轟炸後, 統計倫敦市各區被炸之彈數, 在計算中彈 x 個之區間數時, 得到與 Poisson 分配極其相近之結果. 我們利用電腦之亂數模擬當年之轟炸, 在一塊 20×20 之正方形土地上, 投入 $N = 3000$ 個炸彈, 爲總區數爲 $M = 400$.

1° 其彈著點如圖 2–11 所示:

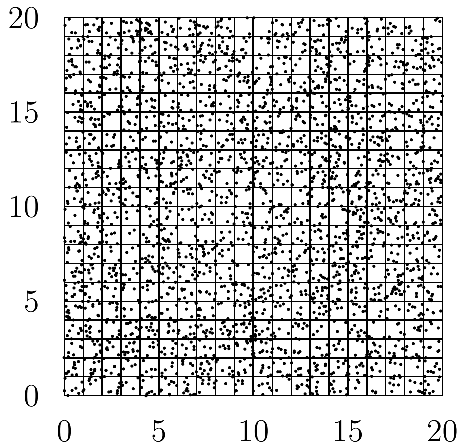


圖 2–11

j	$B(j)$	$g(j)$	$f(j)$
0	0	0	0.00055
1	1	0.0025	0.00415
2	7	0.0175	0.01556
3	17	0.0425	0.03889
4	30	0.075	0.07292
5	43	0.1075	0.1094
6	52	0.13	0.1367
7	54	0.135	0.1465
8	64	0.16	0.1373
9	42	0.105	0.1144
10	34	0.085	0.08583
11	22	0.055	0.05852
12	16	0.04	0.03658
13	9	0.0225	0.0211
14	7	0.0175	0.0113
15	2	0.005	0.00565
16	0	0	0.00265
17	0	0	0.00117
18	0	0	0.00049
19	0	0	0.00019

2° 統計各區之彈著數後, 設 $j = 0, 1, 2, \cdots$,

$B(j) =$ 被炸 j 個炸彈之區數 ;

$g(j) = \dfrac{B(j)}{M}$ (相對頻率);

$f(j) = e^{-\lambda}\dfrac{\lambda^j}{j!}$ (理論機率);

$\lambda = \dfrac{3000}{400} = 7.5$ (每區平均中彈數).

3° 比較右表中 $g(j)$ 與 $f(j)$ 諸值, 雖有若干差異, 這是由於 $g(j)$ 爲一實驗數值, 大體而言, 此一成果已顯示 Poisson 分配可應用於此類問題. □

此外, 有人以嚴格的公設法及微分方程式來探討 Poisson 分配, 而有所謂的 Poisson 過程. 這一部分屬於《隨機過程》課程之範疇, 我們不予探討. 有興趣者可參閱 Feller [5] 或 Ross [13] 之書.

(三) 超幾何分配(hypergeometric distribution)

有一袋子, 內置白球 m 個, 黑球 n 個, 以"不放回"方式隨機抽取 r 個, $(r \le m+n)$, 取得白球曰成功 (S), 取得黑球曰失敗 (F), 則

- 樣本空間爲 $\Omega = \{(b_1, \cdots, b_r) \mid b_j \in \{S, F\},\ \text{成功數} \le m,\ \text{失敗數} \le n\}$;
- σ 域爲 $\mathcal{F} = \mathcal{P}(\Omega)$, (因爲 $\#\Omega$ 爲有限);
- 機率測度爲 $P\colon \mathcal{F} \to [0,1]$ 而每一基本事件之機率則需仿照第一章第五節例 2 以樹圖討論之.

令 $X\colon \Omega \to \mathbb{R} : X(b_1, \cdots, b_r) =$ "$(b_1, \cdots, b_r)$ 中成功之次數", 則

(1) X 爲一隨機變數, (此因 $\mathcal{F} = \mathcal{P}(\Omega)$);

(2) X 爲離散型, 因若令

$$A = X(\Omega) = \{j \in \mathbb{Z}_+ \mid \max\{0,\, r-n\} \le j \le \min\{m,\, r\}\},$$

則 $P\{X \in A\} = P(\Omega) = 1$.

(3) X 之密度函數爲

$$f\colon \mathbb{R} \to [0,\,1] : f(x) = P\{X = x\} = \begin{cases} \dfrac{\dbinom{m}{x}\dbinom{n}{r-x}}{\dbinom{m+n}{r}}, & \text{若 } x \in A, \\ 0, & \text{若 } x \in \mathbb{R} \setminus A, \end{cases}$$

此種分配稱爲 **超幾何分配** 或 $H(m,\, n,\, r)$ 分配.

註: 由附錄一 A.4 定理, 我們知道

$$\forall m, n \in \mathbb{Z}_+, \forall r \in \{0, 1, \cdots, m+n\}, \ \binom{m+n}{r} = \sum_{j=0}^{r} \binom{m}{j}\binom{n}{r-j}.$$

是以

$$\sum_{x \in A} f(x) = \sum_{x=0}^{r} f(x) = \frac{1}{\binom{m+n}{r}} \sum_{x=0}^{r} \binom{m}{x}\binom{n}{r-x} = 1.,^{[\star]}$$

[⋆] 若 $k > \alpha$ 或 $k < 0$, 則 $\dbinom{\alpha}{k} = 0$.

例 5. 設 $X \sim H(18, 10, 15)$, 試求 $f_X(x),\ x \in \{5, 6, \cdots, 15\}$, 之值. 並繪出 f_X 之圖形.

解 計算超幾何分配之密度函數之値相當麻煩, 我們利用數學 (或統計) 軟體可求得 f_X 之諸值如下:

x	$f_X(x)$	x	$f_X(x)$
5	0.0000	11	0.1785
6	0.0050	12	0.0595
7	0.0382	13	0.0103
8	0.1402	14	0.0008
9	0.2727	15	0.0000
10	0.2945		

其圖形如下:

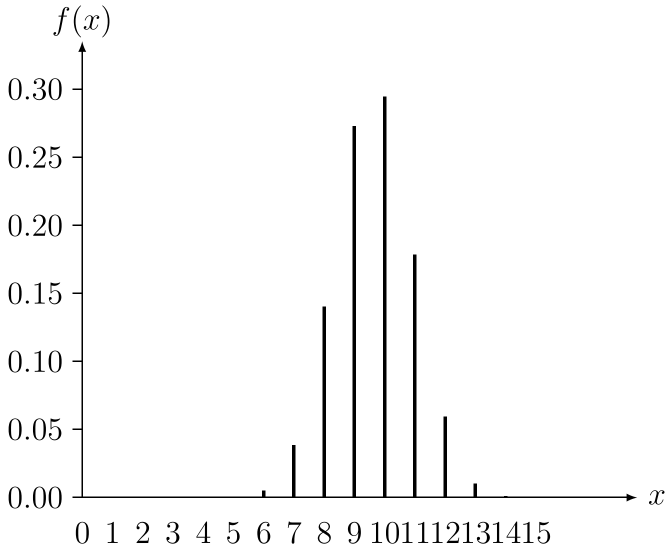


圖 2–12　超幾何分配 $H(18, 10, 15)$ 之密度函數 □

(四)　**負二項分配 (negative binomial distribution)**

有一袋子, 內置白球黑球各若干 (未必相等), 以“放回”方式隨機抽取袋中之球, 取得白球曰成功 (S), 取得黑球曰失敗 (F), 而每次白球被抽到之機率爲 p, 黑球之機率爲 $q = 1 - p$. 此一試驗在抽得白球數爲 r 個時停止. 則

- 樣本空間爲 $\Omega = \{\mathbf{s} \mid \mathbf{s}$ 爲取自 $\{S, F\}$ 之一有限序列, 內有 r 個 S, 最後爲 $S\}$;
- σ 域爲 $\mathcal{F} = \mathcal{P}(\Omega)$, (因爲 Ω 爲可數);
- 機率測度爲 $P\colon \mathcal{F} \to [0, 1] : P\{\mathbf{s}\} = p^r q^x$, (內 x 爲試驗結束時, 失敗之次數.)

令 $X\colon \Omega \to \mathbb{R} : X(\mathbf{s}) =$ (試驗結束時) $\mathbf{s}$ 中失敗之次數. 則

(1)　X 爲一隨機變數, (此因 $\mathcal{F} = \mathcal{P}(\Omega)$);

(2)　X 爲離散型, 因 X 之值域爲 $X(\Omega) = \mathbb{Z}_+ = \{0, 1, 2, \cdots\}$, 乃可數集合.

(3)　X 之密度函數爲

$$f\colon \mathbb{R} \to [0, 1] : f(x) = P\{X = x\} = \begin{cases} *, & \text{若 } x \in \mathbb{Z}_+, \\ 0, & \text{若 } x \in \mathbb{R} \setminus \mathbb{Z}_+, \end{cases}$$

內

$$* = \binom{r+x-1}{x} p^{r-1} q^x p = p^r \binom{r+x-1}{x} q^x.$$

(理由: 最後爲 S, 固定, 其餘 $r+x-1$ 個位子中 $(r-1)$ 個 S 以及 x 個 F, 共有 $\binom{r+x-1}{x}$ 種排列法.)

此種分配稱爲**負二項分配**或 $NB(r, p)$ **分配**.

註: 利用微積分之冪級數之理論, 我們可證明: $\forall \alpha \in \mathbb{R}$,

$$(1+x)^\alpha = \sum_{k=0}^{+\infty} \binom{\alpha}{k} x^k, \ \forall x \in (-1,1), \tag{2.1}$$

以 $-q$ 代 x, 以 $-r$ 代 α, 則因

$$\begin{aligned}\binom{-r}{k} &= \frac{(-r)(-r-1)\cdots(-r-k+1)}{k!} \\ &= (-1)^k \frac{(r+k-1)(r+k-2)\cdots(r+1)r}{k!} \\ &= (-1)^k \binom{r+k-1}{k},\end{aligned}$$

故 (1) 式可寫

$$p^{-r} = (1-q)^{-r} = \sum_{k=0}^{+\infty} \binom{r+k-1}{k} q^k,$$

移項之, 得

$$1 = \sum_{k=0}^{+\infty} p^r \binom{r+k-1}{k} q^k,$$

此級數之各項恰爲負二項分配之密度函數之值, 此乃負二項分配之由來.

例 6. 設 $X \sim NB(3, 0.32)$, 試求 $f_X(x)$, $x \in \{0, 1, \cdots, 17\}$, 之值. 並繪出 f_X 之圖形.

$\boxed{\text{解}}$ 計算負二項分配之密度函數之值相當麻煩, 我們利用數學 (或統計) 軟體可 求得 f_X 之諸值如下:

x	$f_X(x)$	x	$f_X(x)$
0	0.0328	9	0.1241
1	0.0668	10	0.0993
2	0.0909	11	0.0722
3	0.1030	12	0.0481
4	0.1051	13	0.0296
5	0.1000	14	0.0169
6	0.0907	15	0.0090
7	0.0793	16	0.0045
8	0.0674	17	0.0021

其圖形如下:

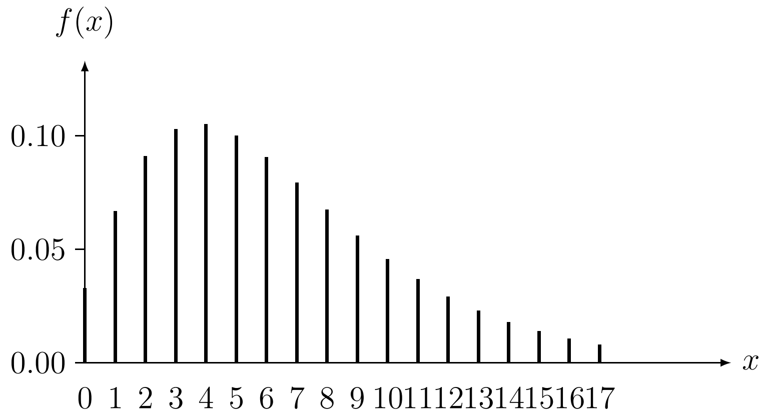


圖 2–13 負二項分配 $NB(3,\,0.32)$ 之密度函數

(五) **幾何分配(geometric distribution)**

$r=1$ 之負二項分配又稱爲**幾何分配**或 Pascal 分配, 即 X 之密度函數爲

$$f\colon \mathbb{R}\to[0,\,1]: f(x)=P\{X=x\}=\begin{cases} pq^x, & \text{若 } x\in\mathbb{Z}_+,\\ 0, & \text{若 } x\in\mathbb{R}\setminus\mathbb{Z}_+,\end{cases}$$

§ 2.5 常用連續型隨機變數

(一) **常態分配(normal distribution or Laplace-Gauss distribution)**

在上一章曾提及, 常態分配是 de Moivre 所創, 其目的乃爲解決二項分配中 $\binom{n}{x}p^xq^{n-x}$ 之近似值, 在機率論及統計學中, 此一分配乃“所有”分配中最爲重要者. 若函數

$$f\colon\mathbb{R}\to\mathbb{R}_+: f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big), \qquad (*)$$

(其中 $\mu\in\mathbb{R}$, $\sigma>0$), 爲隨機變數 X 之一密度函數, 則稱 X 具**常態分配**, 簡寫爲 $X\sim N(\mu,\sigma^2)$, 而 μ, σ^2 稱爲其**參數**或**母數**(parameters). 若 $\mu=0$, $\sigma=1$, 則稱 X 具**標準常態分配**(standard normal distribution).

做爲一離散型密度函數必須滿足以下二條件 (1) $f\geq 0$, (2) $\sum_{x\in A}f(x)=1$ 其中集合 $A=\{x\in\mathbb{R}\mid f(x)>0\}$; 做爲一連續型密度函數則須滿足 (1) $f\geq0$, (2) $\int_{\mathbb{R}}f=1$. 上面 $(*)$ 式所界定之函數顯然滿足 (1); 至於 (2), 先利用 $t=\frac{x-\mu}{\sigma}$ 變數代換, 使成爲

$$\begin{aligned} I&=\int_{-\infty}^{+\infty}f(x)dx\\ &=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)dx\\ &=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\exp\Big(-\frac{t^2}{2}\Big)\,dt=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}\exp\Big(-\frac{t^2}{2}\Big)\,dt. \end{aligned}$$

是以

$$I^2 = \frac{2}{\pi}\Big(\int_0^{+\infty} \exp\Big(-\frac{x^2}{2}\Big)dx\Big)\Big(\int_0^{+\infty} \exp\Big(-\frac{y^2}{2}\Big)dy\Big)$$
$$= \frac{2}{\pi}\int_0^{+\infty}\Big(\int_0^{+\infty} \exp\Big(-\frac{x^2+y^2}{2}\Big)dx\Big)dy.$$

利用二重積分之極坐標變換, 即令 $x = r\cos\theta,\ y = r\sin\theta$, 則

$$I^2 = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{+\infty} \exp\Big(-\frac{r^2}{2}\Big)\, r\, dr\, d\theta = \frac{2}{\pi}\int_0^{\pi/2} d\theta = 1.$$

是以 $I = 1$.

f 之圖形對稱於 $x = \mu$, 於點 μ 有最大値 $(\sigma\sqrt{2\pi})^{-1}$, 而於點 $\mu+\sigma$ 及點 $\mu-\sigma$ 有反曲點, σ 越大, f 越扁平, 表示資料越分散, 參閱圖 2–14.

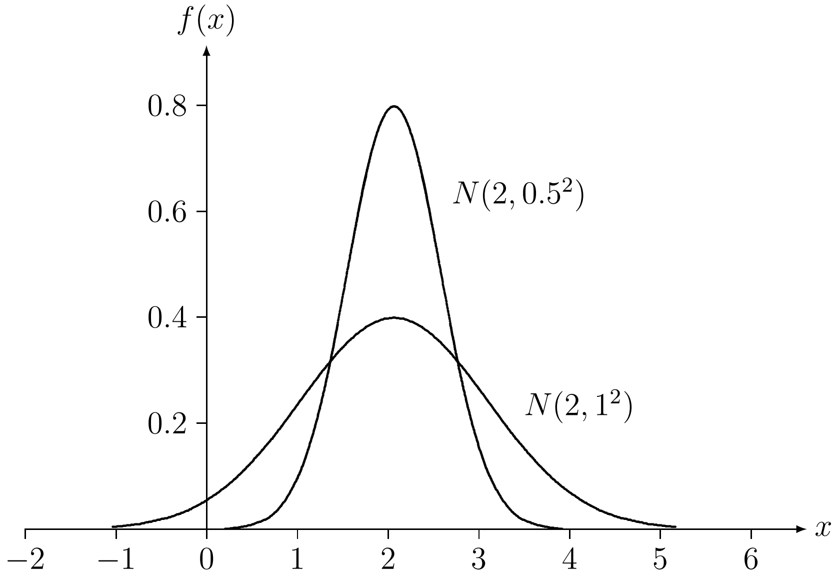


圖 2–14　常態分配之密度函數

(二)　**Gamma 分配**

若函數

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1}e^{-x/\beta}, & 若\ x > 0, \\ 0, & 若\ x \le 0, \end{cases}$$

(其中 $\alpha > 0,\ \beta > 0$), 爲隨機變數 X 之一密度函數, 則稱 X具 **Gamma 分配** 或 $X \sim G(\alpha, \beta)$, 而 α, β 稱爲此分配之**參數**或**母數**.

註:

(1) 在理論分析中, 我們已知 Gamma 函數爲

$$\Gamma(\alpha) = \int_0^{+\infty} y^{\alpha-1}e^{-y}dy, \forall \alpha > 0,$$

以及

1° $\forall \alpha > 1,\ \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$;

2° $\forall \alpha \in \mathbb{N},\ \Gamma(\alpha) = (\alpha-1)!$;

3° $\Gamma(\dfrac{1}{2}) = \sqrt{\pi},\ \Gamma(\dfrac{3}{2}) = \dfrac{\sqrt{\pi}}{2}$.

(2) 做爲一連續型密度函數須滿足 (i) $f \geq 0$ (顯然成立), (ii) $\int_{\mathbb{R}} f = 1$. 爲此，令 $y = \dfrac{x}{\beta}$ 變數代換之, 則 $dy = \dfrac{dx}{\beta}$,

$$\begin{aligned}\int_{-\infty}^{+\infty} f(x)\,dx &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x^{\alpha-1}e^{-x/\beta}dx \\ &= \frac{1}{\Gamma(\alpha)}\int_0^{+\infty} y^{\alpha-1}e^{-y}dy = \frac{1}{\Gamma(\alpha)}\Gamma(\alpha) = 1.\end{aligned}$$

Gamma 分配之密度函數圖形如下:

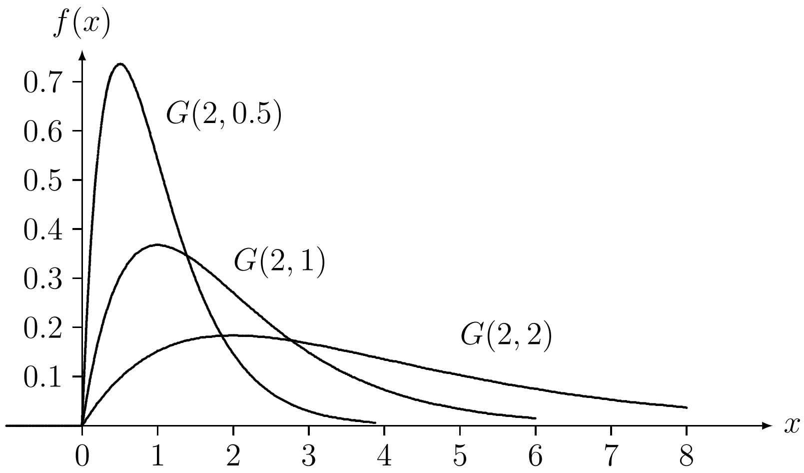


圖 2–15 Gamma 分配之密度函數

(三) χ^2 **分配**

在 $G(\alpha,\beta)$ 分配中, 若 $\alpha = \frac{r}{2},\ \beta = 2$, 其中 $r \in \mathbb{N}$, 即其密度函數爲

$$f : \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \frac{1}{\Gamma(r/2)2^{r/2}} x^{\frac{r}{2}-1}e^{-x/2}, & \text{若 } x > 0, \\ 0, & \text{若 } x \leq 0. \end{cases}$$

若 X 具有這種密度函數, 則稱 X 具 χ^2 **分配**(chi-square distribution) 或 $X \sim \chi^2_r$, 而 r 則稱爲此分配之**自由度** (the degrees of freedom). 在統計推論中, χ^2 分配用途甚廣. 其圖形如下:

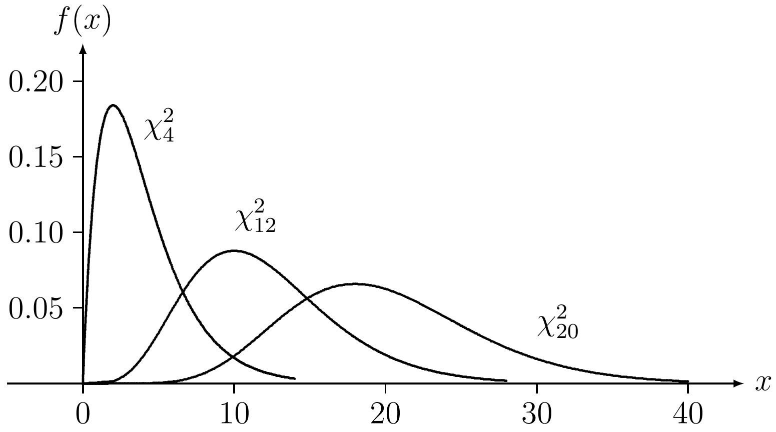


圖 2–16 χ^2 分配之密度函數

(四) **負指數分配 (negative exponential distribution)**

在 $G(\alpha,\, \beta)$ 分配中, 若 $\alpha = 1,\ \beta = 1/\lambda$, 其中 $\lambda > 0$, 則稱 X 具**負指數分配**, (有些學者稱 X 具指數分配) 或 $X \sim G(1,\, 1/\lambda)$, 即其密度函數為

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{若 } x > 0, \\ 0, & \text{若 } x \leq 0. \end{cases}$$

其圖形如下:

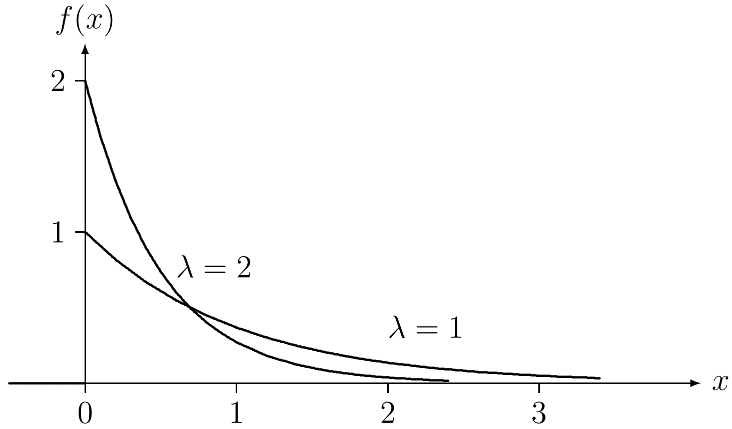


圖 2–17 負指數分配之密度函數

例 1. 設 X 表某一燈泡工廠所生產之日光燈之壽命 (單位: 小時), 已知 X 具負指數分配 $G(1,\, 700)$. (這表示日光燈之平均壽命為 700 小時, 詳第四章.)

(1) 試求 $P\{X > 400\}$;

(2) 試求 $P\{X > 500 \mid X \geq 100\}$.

$\boxed{\text{解}}$ 由原設知 X 之密度函數為

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \dfrac{1}{700} e^{-x/700}, & \text{若 } x > 0, \\ 0, & \text{若 } x \leq 0. \end{cases}$$

$$(1)\quad P\{X>400\}=\frac{1}{700}\int_{400}^{+\infty}e^{-x/700}dx=-e^{-x/700}\Big|_{400}^{+\infty}=e^{-4/7}.$$

$$\begin{aligned}(2)\quad P\{X>500\mid X\geq 100\}&=\frac{P\{X>500,\ X\geq 100\}}{P\{X\geq 100\}}\\&=\frac{P\{X>500\}}{P\{X\geq 100\}}=\frac{e^{-5/7}}{e^{-1/7}}=e^{-4/7}.\end{aligned}$$

此與第 (1) 題之機率相等. □

註: 更廣義地說, 若 X 具 $G(1,\,1/\lambda)$ 之負指數分配, 則

$$P\{X>t_0+t\mid X\geq t_0\}=P\{X>t\},\forall\, t_0,t\in(0,+\infty). \tag{1}$$

因爲

$$\begin{aligned}P\{X>t_0+t\mid X\geq t_0\}&=\frac{P\{X>t_0+t,\ X\geq t_0\}}{P\{X\geq t_0\}}\\&=\frac{P\{X>t_0+t\}}{P\{X\geq t_0\}}=\frac{\lambda\int_{t_0+t}^{+\infty}e^{-\lambda x}dx}{\lambda\int_{t_0}^{+\infty}e^{-\lambda x}dx}\\&=\frac{e^{-\lambda(t_0+t)}}{e^{-\lambda t_0}}=e^{-\lambda t}=P\{X>t\}.\end{aligned}$$

設 X 表一產品之壽命 (單位: 小時), 則

- $\{X=533\}$ 表使用 533 小時後此一產品即已損壞;
- $\{X>1200\}$ 表使用 1200 小時後此一產品仍未損壞;
- $P\{X\geq 1500\mid X>1200\}$ 表使用 1200 小時後此一產品仍未損壞之條件下, 再使用 300 小時以上之機率.

(1) 式表示, 在日光燈點了 t_0 小時之後, 仍然亮著, 則可再使用超過 t 小時之條件機率與一開始可使用超過 t 小時之機率是相等的. 這種性質說明 (具負指數分配之) 日光燈的壽命具**失憶性**(lack of memory property). 除了負指數分配之外, 幾何分配亦有類似性質, 若 $X\sim NB(1,p)$, 則

$$P\{X>n+m\mid X\geq n\}=P\{X>m\},\forall n,m\in\mathbb{Z}_+, \tag{2}$$

(習題). 讀者應注意: 此處之 n 與 m 必須爲非負整數, 而負指數分配則適用於任何正實數, 所有分配中只有負指數分配具有 (1) 性質, (參見附錄六).

(五) 均勻分配 (uniform distribution)

若函數 $f\colon \mathbb{R} \to \mathbb{R}_+$:

$$f(x) = \begin{cases} \dfrac{1}{\beta-\alpha}, & 若\ x \in [\alpha, \beta], \\ 0, & 若\ x \in \mathbb{R} \setminus [\alpha, \beta], \end{cases}$$

(內 $\alpha,\ \beta \in \mathbb{R},\ \alpha < \beta$), 爲隨機變數 X 之一密度函數, 則稱 X 具**均勻分配** 或 $X \sim U(\alpha, \beta)$, 而 α, β 則稱爲此分配之**參數**或**母數**. 其圖形如右:

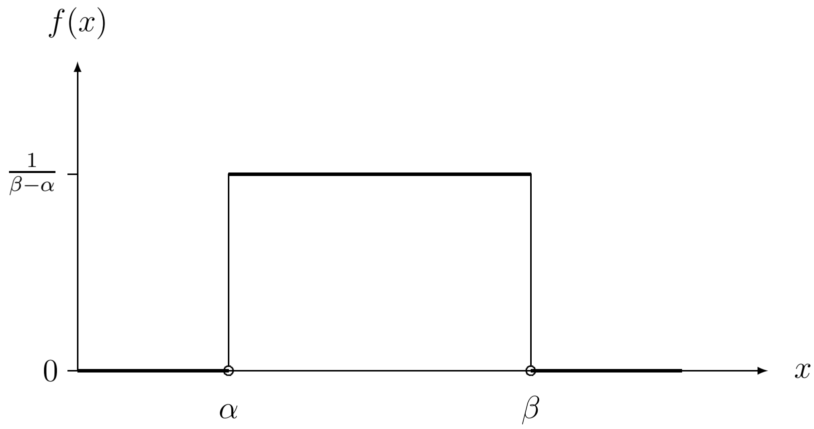


圖 2–18 均勻分配之密度函數

例 2. 某君自家中騎機車出發, 出門時將大門鑰匙放在口袋中, 騎了一段 300 公尺之路程後發現鑰匙不見了, 由於不知鑰匙在何處遺失, 故假設遺失位置爲 X, 顯然 $X \sim U(\alpha, \beta)$, 其中 $\alpha = 0,\ \beta = 300$, 故知鑰匙掉在前 100 公尺之機率爲

$$P\{0 \le X \le 100\} = \int_0^{100} \frac{1}{300-0}\, dx = \frac{1}{3}. \qquad \square$$

(六) **beta 分配**

若函數

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 若\ 0 < x < 1, \\ 0, & 否則. \end{cases}$$

(內 $\alpha > 0,\ \beta > 0$), 爲隨機變數 X 之一密度函數, 則稱 X 具 **Beta 分配**, 而 α, β 稱爲此分配之**參數**或**母數**. 其圖形如下:

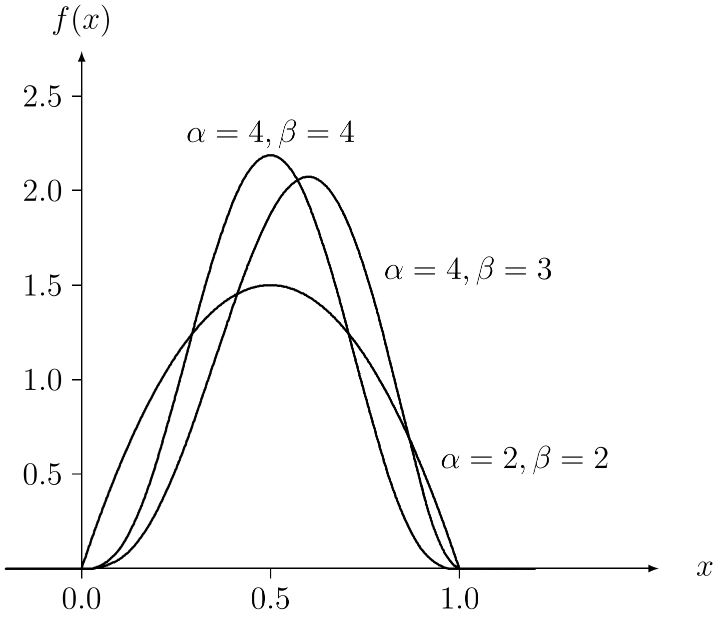


圖 2–19

註:

1° 若 $\alpha > 0$, $\beta > 0$, 則瑕積分 $\displaystyle\int_{0^+}^{1^-} x^{\alpha-1}(1-x)^{\beta-1}dx$ 稱爲 Beta 函數, 在理論分析中 (例如 Apostol [1], p.331), 可證上述瑕積分收斂於 $\dfrac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. 因之, 顯然有: (1) $f \geq 0$, 以及 (2) $\displaystyle\int_{\mathbb{R}} f = 1$.

2° 當 $\alpha = 1$, $\beta = 1$ 時, 此分配乃爲均勻分配.

(七) **Cauchy 分配**

若函數

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \frac{\sigma}{\pi} \cdot \frac{1}{\sigma^2 + (x-\mu)^2},$$

(內 $\mu \in \mathbb{R}$, $\sigma > 0$), 爲隨機變數 X 之一密度函數, 則稱 X 具 **Cauchy分配**或 $X \sim C(\mu, \sigma^2)$, 而 μ, σ 稱爲此分配之**參數**或**母數** .

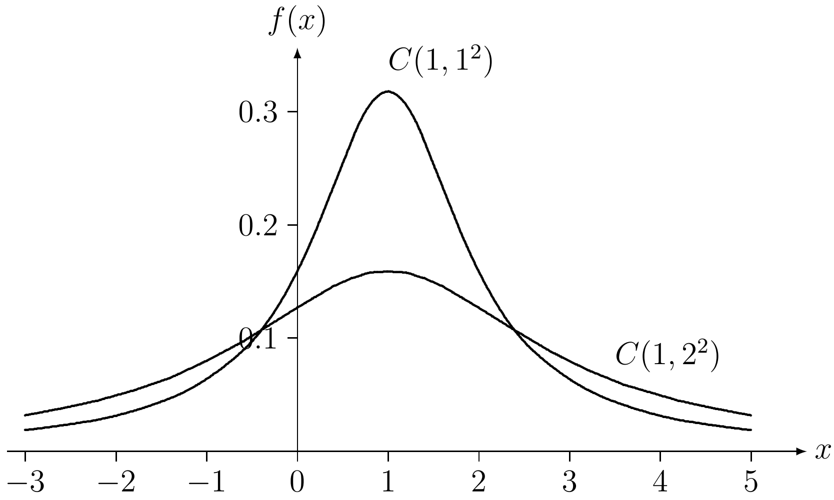


圖 2–20 Cauchy 分配之密度函數

上述函數 f 對稱於 $x = \mu$, 其圖形略似於 $N(\mu, \sigma^2)$, 但二者並不相同. 此外, 我們應證明 (1) $f \geq 0$ (易明), 以及 (2) $\displaystyle\int_{\mathbb{R}} f = 1$. 由於

$$
\begin{aligned}
I = \int_{\mathbb{R}} f(x)\,dx &= \frac{\sigma}{\pi}\int_{-\infty}^{+\infty} \frac{dx}{\sigma^2 + (x-\mu)^2} \\
&= \frac{1}{\pi\sigma}\int_{-\infty}^{+\infty} \frac{dx}{1 + (\frac{x-\mu}{\sigma})^2}, \\
&= \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{dy}{1+y^2}, (\text{令 } y = \frac{x-\mu}{\sigma}) \\
&= \frac{2}{\pi}\lim_{t\to+\infty}\int_0^t \frac{dy}{1+y^2} = \frac{2}{\pi}\lim_{t\to+\infty} \tan^{-1} y \,\Big|_0^t = \frac{2}{\pi}\cdot\frac{\pi}{2} = 1.
\end{aligned}
$$

例 3. 某雷射光源位於一東西向之直線牆壁之北方一公里 (如圖), 此光源以等角速度逆時針旋轉半周向南投射.

(a) 若牆線視爲 x 軸, 光源位於點 $(0, 1)$, 試問光線射於牆上之區間 $[-1, 1]$ 之機率爲何?

(b) 試問光線射於牆上之分配爲何?

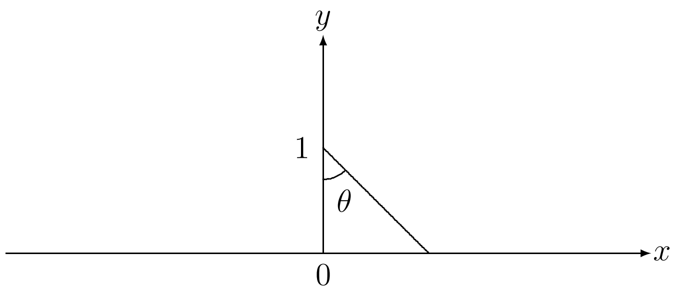


圖 2–21

解 (a) 設 θ 表 y 軸與光線之夾角 $(\in(-\frac{\pi}{2},\frac{\pi}{2}))$, 吾人可視光源 $\theta\sim U(-\frac{\pi}{2},\frac{\pi}{2})$, 光線射於區間 $[-1,1]$ 之機率爲

$$\begin{aligned}P&=P\{\tan^{-1}(-1)<\theta<\tan^{-1}1\}=\int_{\tan^{-1}(-1)}^{\tan^{-1}1}\frac{1}{\pi}\,d\theta\\&=\frac{\tan^{-1}1-\tan^{-1}(-1)}{\pi}=\frac{1}{2}.\end{aligned}$$

(b) 設 X 表光線射於牆上之點, 則 X 之分配函數爲

$$\begin{aligned}F_X(x)&=P\{X\le x\}\\&=P\{-\frac{\pi}{2}<\theta\le\tan^{-1}(x)\}\\&=\int_{-\pi/2}^{\tan^{-1}(x)}\frac{1}{\pi}\,d\theta\\&=\frac{\tan^{-1}(x)-\frac{\pi}{2}}{\pi}.\end{aligned}$$

是以 X 之密度函數爲

$$f\colon\mathbb{R}\to\mathbb{R}_+ : f(x)=F'(x)=\frac{1}{\pi(1+x^2)},$$

此爲 $C(0,1)$ 分配. □

(八) Laplace 分配

若函數 $f\colon\mathbb{R}\to\mathbb{R}_+ : f(x)=\dfrac{1}{2}e^{-|x|}$ 爲隨機變數 X 之一密度函數, 則稱 X 具 **Laplace 分配**. 其圖形如下:

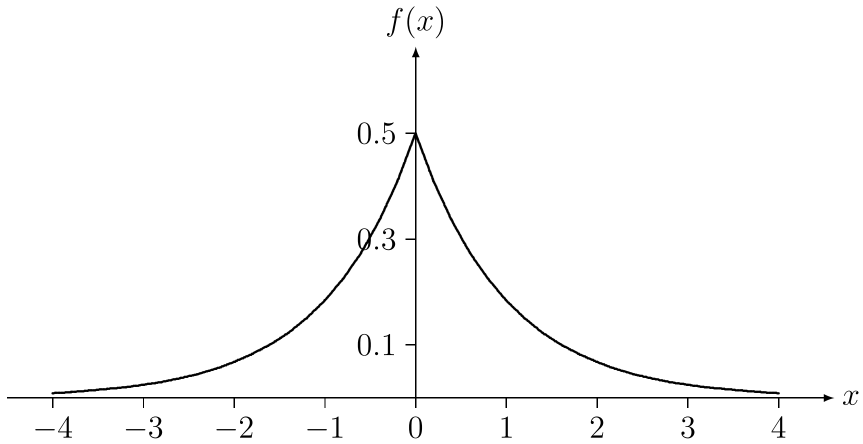


圖 2–22 Laplace 分配之密度函數

(九) 對數常態分配 (lognormal distribution)

若函數

$$f : \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \dfrac{1}{x\beta\sqrt{2\pi}} \exp\Big[-\dfrac{(\ln x - \ln\alpha)^2}{2\beta^2}\Big], & \text{若 } x > 0, \\ 0, & \text{若 } x \le 0, \end{cases}$$

(內 $\alpha > 0,\ \beta > 0$), 爲隨機變數 X 之一密度函數, 則稱 X 具**對數常態分配**, 而 α, β 稱爲此分配之**參數**或**母數**.

註: 若 X 具對數常態分配, 利用變數代換 (詳見第八節), 可證明隨機變數 $Y = \ln \circ X$具常態分配.

除上述介紹之各種分配, 我們尚可在某些專書上找到其他分配. 此外, 目前各項統計或數學之電腦軟體, 內建不少常用分配之累積分配函數及其反函數, 如 SAS, 若需要各分配之密度函數之值, 則可在 **MathCAD** 上直接求得, 使用前最好找一本該軟體之手册查閱一番, 或者在進入該軟體後, 利用線上求助以了解自己使用之函數應如何處理, 以 SAS 爲例, 其 PROBGAM(x,a) 表示 $G(a, 1)$ 之累積分配函數, 如果讀者需要 $G(\alpha, \beta)$ 之分配函數之值, 可利用本章稍後之隨機變數之變換去解決.

§ 2.6 奇異分配與混和型分配

我們知道絕對連續型隨機變數必爲連續型, 但連續型隨機變數則未必爲絕對連續型. 本節中, 我們將介紹一種不爲絕對連續之連續型隨機變數分配, 也就是所謂**奇異分配** (singular distribution). 它是利用 Cantor 三進點集及 Cantor-Lebesgue 函數設計而成的. 又, 本節稍難, 第一次閱讀時可省略之.

首先, 設

$$\begin{aligned}
I &= [0,\,1], \\
D_1 &= \Big(\frac{1}{3},\ \frac{2}{3}\Big), \\
D_2 &= D_{21} \cup D_{22}, \text{內} \\
&\quad D_{21} = \Big(\frac{1}{3^2},\ \frac{2}{3^2}\Big), \quad D_{22} = \Big(\frac{2}{3} + \frac{1}{3^2},\ \frac{2}{3} + \frac{2}{3^2}\Big), \\
D_3 &= D_{31} \cup D_{32} \cup D_{33} \cup D_{34}, \text{內} \\
&\quad D_{31} = \Big(\frac{1}{3^3},\ \frac{2}{3^3}\Big), \quad D_{32} = \Big(\frac{2}{3^2} + \frac{1}{3^3},\ \frac{2}{3^2} + \frac{2}{3^3}\Big), \\
&\quad D_{33} = \Big(\frac{2}{3} + \frac{1}{3^3},\ \frac{2}{3} + \frac{2}{3^3}\Big), \quad D_{34} = \Big(\frac{2}{3} + \frac{2}{3^2} + \frac{1}{3^3},\ \frac{2}{3} + \frac{2}{3^2} + \frac{2}{3^3}\Big), \\
&\ \vdots
\end{aligned}$$

$$
\begin{aligned}
D_n &= \bigcup_{r=1}^{2^{n-1}} D_{n,r}, \text{ 內} \\
&\quad D_{n1} = \Big(\frac{1}{3^n},\ \frac{2}{3^n}\Big), \\
&\quad \vdots \\
&\quad D_{n,2^{n-1}} = \Big(\frac{2}{3^2}+\frac{2}{3^3}+\cdots+\frac{1}{3^n},\ \frac{2}{3^2}+\frac{2}{3^3}+\cdots+\frac{2}{3^n}\Big), \\
&\vdots \\
D &= \bigcup_{n=1}^{+\infty} D_n.
\end{aligned}
$$

顯然 D 爲一開集, D_n 之測度 (即總長度) 爲 $\lambda(D_n)=2^{n-1}/3^n$, 是以 D 之測度爲

$$\lambda(D)=\sum_{n=1}^{+\infty}\lambda(D_n)=\sum_{n=1}^{+\infty}\frac{2^{n-1}}{3^n}=\frac{1}{3}\sum_{n=1}^{+\infty}\Big(\frac{2}{3}\Big)^{n-1}=1.$$

今令 $K=I\setminus D$, 此乃所謂 Cantor 三進點集 (Cantor's ternary set), 應有

$$\lambda(K)=\lambda(I)-\lambda(D)=1-1=0.$$

其次, 我們將界定所謂的 Cantor-Lebesgue 函數.

1° 令

$$G_1\colon I\to\mathbb{R}: G_1(x)=\begin{cases} 0, & \text{若 } x=0, \\ 1, & \text{若 } x=1, \\ 1/2, & \text{若 } x\in D_1, \\ *, & \text{否則.}\end{cases}$$

$*$: 以直線連之, (即點 (0, 0) 及 (1/3, 1/2) 連以直線, 點 (2/3, 1/2) 及 (1, 1) 連以直線.)

2° 令

$$G_2\colon I\to\mathbb{R}: G_2(x)=\begin{cases} 0, & \text{若 } x=0, \\ 1, & \text{若 } x=1, \\ 1/2, & \text{若 } x\in D_1, \\ 1/4, & \text{若 } x\in D_{21}, \\ 3/4, & \text{若 } x\in D_{22}, \\ *, & \text{否則.}\end{cases}$$

$*$: 以直線連之.

3° 仿之, $\forall n\in\mathbb{N}$, 令

$$G_n\colon I\to\mathbb{R}: G_n(x)=\begin{cases} 0, & \text{若 } x=0, \\ 1, & \text{若 } x=1, \\ r2^{-n}, & \text{若 } x\in D_{n,r},\ r=1,2,\cdots,2^{n-1}, \\ *, & \text{否則.}\end{cases}$$

$*$: 以直線連之. $G_1,\ G_2$ 之圖形如下:

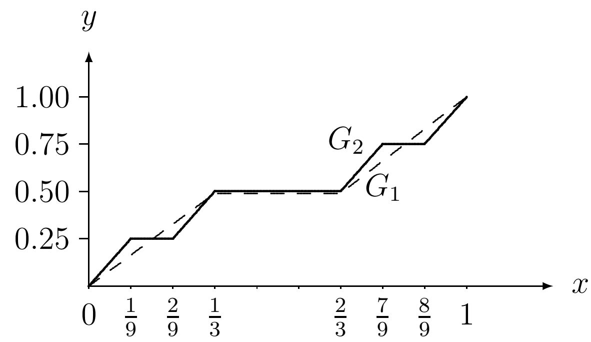


圖 2–23

由以上之設計, 顯然有

$$\text{“ }3^\circ\text{ 之定義”} \Rightarrow \begin{cases} |G_1 - G_2| < 1/2, \\ |G_2 - G_3| < 1/2^2, \\ \quad\vdots \\ |G_n - G_{n+1}| < 1/2^n, \\ \quad\vdots \end{cases}$$

$$\Rightarrow \sum_{n=1}^{+\infty} |G_n - G_{n+1}| < \sum_{n=1}^{+\infty} 1/2^n = 1$$

$$\Rightarrow \sum_{n=1}^{+\infty} (G_n - G_{n+1}) \text{ 均勻收斂於 } I \text{ 上, (利用 Weierstrass M-test)}$$

$$\Rightarrow \{s_n\}_n \text{ 均勻收斂於 } I \text{ 上,} \left(\text{內 } s_n = \sum_{j=1}^{n} (G_j - G_{j+1}) = G_1 - G_{n+1}\right)$$

$$\Rightarrow \{G_n\}_n \text{ 均勻收斂於 } I \text{ 上.}$$

令

$$G\colon I \to \mathbb{R} : G(x) = \lim_{n\to+\infty} G_n(x).$$

則由$\{G_n\}_n$ 之均勻收斂性, 知 G 亦爲連續, 顯然亦爲不減, 我們稱其爲 Cantor-Lebesgue 函數.

最後, 令

$$F\colon \mathbb{R} \to [0,1] : F(x) = \begin{cases} 0, & \text{若 } x < 0, \\ G(x), & \text{若 } 0 \le x \le 1, \\ 1, & \text{若 } 1 < x, \end{cases}$$

則函數 F 爲 (1) 不減, (2) 連續, (3)$\lim_{x\to-\infty} F(x) = 0, \lim_{x\to+\infty} F(x) = 1$; 故 F 爲一連續型分配函數, 又利用絕對連續之定義我們們可證明 F 並非絕對連續, 是以 F 爲一奇異分配函數; (另一種解釋如下: 函數 F 在 Cantor 集 K 上之各點皆不爲可微, 而在 $\mathbb{R} \setminus K$ 各點之導數皆爲零, 換言之, $F' = 0$ a.e., 因之 F 並無密度函數.) 此一奇異分配函數之圖形大致如下:

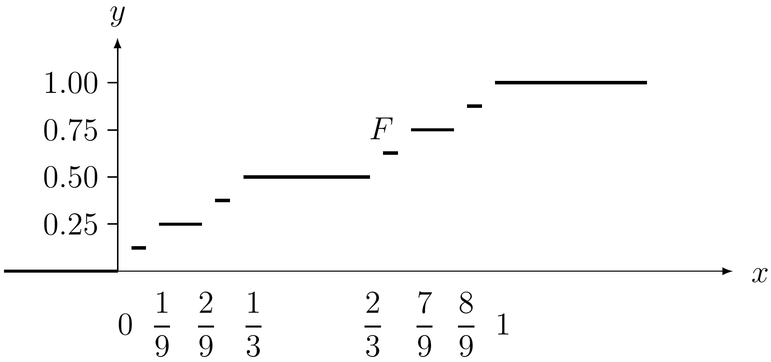


圖 2–24　奇異分配

隨機變數之分類, 實際上是依其累積分配函數而分類, 而分配函數除了離散型、絕對連續型及奇異連續型之外, 尚可將以上三型加以混合, 其方法是取一離散型分配函數 F_d, 一絕對連續型分配函數 F_a 以及一奇異分配函數 F_s 予以凸狀組合之, 即令 $\alpha,\ \beta,\ \gamma \in [0,\ 1]$, $\alpha+\beta+\gamma=1$, 則函數

$$F=\alpha F_d+\beta F_a+\gamma F_s$$

顯然滿足 (1) 不減, (2) 右連續, (3) $\lim_{x\to-\infty}F(x)=0, \lim_{x\to+\infty}F(x)=1$; 故 F 爲一分配函數.

一、若 $\alpha,\ \beta,\ \gamma$ 三者之二爲 0, 則 F 顯然爲離散型、絕對連續型或奇異型. 例如, $\beta=0$, $\gamma=0$, 則必 $\alpha=1$, 此時,

$$F=\alpha F_d+\beta F_a+\gamma F_s=F_d,$$

函數 F 顯然爲離散型.

二、若 $\alpha,\ \beta,\ \gamma$ 三者之一爲 0, 則 F 爲二型混合. 例如 F_d 爲 $B(1,\ p)$ 之分配函數, F_a 爲 $U(0,\ 1)$ 之分配函數, $\alpha=\beta=1/2$, 則

$$F=\alpha F_d+\beta F_a\colon \mathbb{R}\to[0,\ 1]: F(x)=\begin{cases} 0, & \text{若 } x<0,\\ (1-p+x)/2, & \text{若 } 0\le x<1,\\ 1, & \text{若 } x\ge 1,\end{cases}$$

其圖形如下:

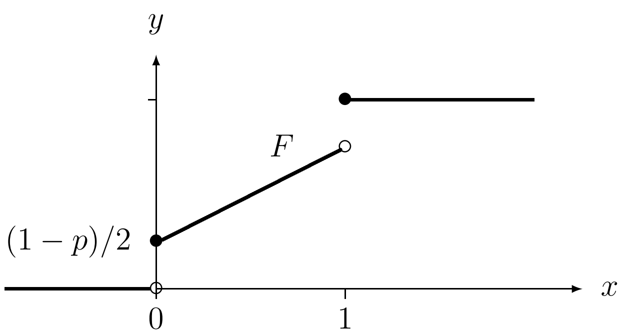


圖 2–25

三、若 α, β, γ 三者均不爲 0, 則 F 爲三型混合, 雖然實用價値不高, 但在機率的研究上具相當重要性.

§ 2.7 隨機變數變換之分配函數

設 X 爲一隨機變數, 傳統上將 X^2, $|X|$, $3X$, $\sin(X), \cdots$ 皆視爲 X 的變換, 其實所謂變換只不過是函數之合成 (composition of functions). 在本章的定理 2.7 中, 我們已知, 若函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲 Borel 可測, 則合成函數 $h \circ X$ 亦爲一隨機變數. 在機率論中, 了解 $h \circ X$ 之分配是一項十分重要的課題, 以下兩節我們將深入探討. 許多學者仍沿用古典數學合成函數之符號, 將 $h \circ X$ 寫爲 $h(X)$, 不過我們認爲 $h \circ X$ 較爲合理. 本節中我們研究以下兩個主題:

甲、若隨機變數 X 之機率分配函數 P_X 爲已知, 則 $Y = h \circ X$ 之機率分配函數 P_Y 能否以 P_X 表之?

乙、若隨機變數 X 之累積分配函數 F_X 爲已知, 則 $Y = h \circ X$ 之累積分配函數 F_Y 能否以 F_X 表之?

定理 2.21

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, P_X 爲其機率分配函 數, 函數 $h : \mathbb{R} \to \mathbb{R}$爲 Borel 可測, 則函數 $Y = h \circ X$ 之機率分配函數爲

$$P_Y\colon \mathcal{B} \to [0,\, 1] : P_Y(B) = P_X(h^{-1}(B)).$$

證 $\forall B \in \mathcal{B}$,

$$\begin{aligned} P_Y(B) &= P(Y^{-1}(B)) \\ &= P(h \circ X)^{-1}(B) \\ &= P(X^{-1}(h^{-1}(B))) \\ &= P_X(h^{-1}(B)). \end{aligned}$$

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "Y"'] & \mathbb{R} \arrow[d, "h"] \\ & \mathbb{R} \end{tikzcd} \qquad \begin{tikzcd} \mathcal{B} \arrow[r, "P_X"] & {[0,1]} \\ \mathcal{B} \arrow[u, "h^{-1}"] \arrow[ur, "P_Y"'] & \end{tikzcd}$$

換言之, P_Y 乃爲 h^{-1} 與 P_X 之合成, (注意:h^{-1} 在 此並非 h 之 反函數而是 h 之像原函數 (inverse image function)). $\square$

上述結果雖然簡單明瞭, 但並不實用. 例如, $X \sim N(0,\, 1)$, 則

$$P_X(B) = \frac{1}{\sqrt{2\pi}} \int_B \exp\Big(-\frac{x^2}{2}\Big) dx, \forall B \in \mathcal{B}.$$

令 $Y = X^2$, 即 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = x^2$, 由定理 2.21知

$$\begin{aligned} P_Y(B) &= P_X(h^{-1}(B)) \\ &= \frac{1}{\sqrt{2\pi}} \int_{h^{-1}(B)} \exp\Big(-\frac{x^2}{2}\Big) dx. \end{aligned}$$

由於 $h^{-1}(B) = \{x \in \mathbb{R} \mid h(x) = x^2 \in B\}$ 並不能更清楚的表示, 因此也看不出來 P_Y 是何種分配.

定理 2.22

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, F_X 爲其累積分配函數.

(1) 若函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲 Borel 可測, 則函數 $Y = h \circ X$ 之累積分配函數爲

$$F_Y\colon \mathbb{R} \to [0,\, 1] : F_Y(y) = P_X(h^{-1}(-\infty,\, y]);$$

(2) 若上述可測函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲遞增且爲蓋射, 則

$$F_Y\colon \mathbb{R} \to [0,\, 1] : F_Y(y) = F_X(h^{-1}(y)) = P\{X \le h^{-1}(y)\};$$

其中 h^{-1} 爲 h 之反函數;

(3) 若上述可測函數 $h\colon \mathbb{R} \to \mathbb{R}$ 爲遞減且爲蓋射, 則

$$F_Y\colon \mathbb{R} \to [0,\, 1] : F_Y(y) = 1 - F_X(h^{-1}(y)^-) = P\{X \ge h^{-1}(y)\}.$$

其中 h^{-1} 爲 h 之反函數, $F_X(h^{-1}(y)^-)$ 爲 F_X 在點 $h^{-1}(y)$ 之左極限.

證 (1) $\forall y \in \mathbb{R}$, 由定理 2.21 知

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} \\ &= P(Y^{-1}(-\infty,\, y]) \\ &= P_Y(-\infty,\, y] \\ &= P_X(h^{-1}(-\infty,\, y]). \end{aligned}$$

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "Y"'] & \mathbb{R} \arrow[r, "F_X"] \arrow[d, "h"] & {[0,1]} \\ & \mathbb{R} \arrow[ur, "F_Y"'] & \end{tikzcd}$$

(2)

$$\begin{aligned} F_Y(y) &= P\{Y \le y\} \\ &= P\{\omega \mid Y(\omega) \le y\} \\ &= P\{\omega \mid h(X(\omega)) \le y\} \\ &= P\{\omega \mid X(\omega) \le h^{-1}y\}, \quad (\text{因 } h \text{ 及其反函數 } h^{-1} \text{ 均爲遞增.}) \\ &= F_X(h^{-1}(y)). \end{aligned}$$

(3)

$$\begin{aligned}
F_Y(y) &= P\{Y \le y\} \\
&= P\{\omega \mid Y(\omega) \le y\} \\
&= P\{\omega \mid h(X(\omega)) \le y\} \\
&= P\{\omega \mid X(\omega) \ge h^{-1}y\}, \quad (\text{因 } h \text{ 及其反函數 } h^{-1} \text{ 均爲遞減.}) \\
&= 1 - P\{\omega \mid X(\omega) < h^{-1}y\} \\
&= 1 - F_X(h^{-1}(y)^-).
\end{aligned}$$

□

例 1. 設 $X \sim G(\alpha, \beta)$, 試求 $Y = kX$ 之分配, 其中 $k > 0$.

解

$$\begin{aligned}
F_Y(y) &= P\{Y \le y\} = P\{kX \le y\} \\
&= P\{X \le y/k\} = F_X(y/k) \\
&= \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} \displaystyle\int_0^{y/k} x^{\alpha-1} e^{-x/\beta} dx, & \text{若 } y > 0, \\ 0, & \text{若 } y \le 0, \end{cases}
\end{aligned}$$

利用定理 2.20 得 Y 之密度函數爲

$$\begin{aligned}
f_Y : \mathbb{R} \to \mathbb{R}_+ : f_Y(y) &= \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} (\dfrac{y}{k})^{\alpha-1} e^{-y/(k\beta)} \cdot \dfrac{1}{k}, & \text{若 } y > 0, \\ 0, & \text{若 } y \le 0, \end{cases} \\
&= \begin{cases} \dfrac{1}{\Gamma(\alpha)(k\beta)^\alpha} y^{\alpha-1} e^{-y/(k\beta)}, & \text{若 } y > 0, \\ 0, & \text{若 } y \le 0, \end{cases}
\end{aligned}$$

是以知 Y 具有 $G(\alpha, k\beta)$ 分配. □

定理 2.23

設隨機變數 X 具 $N(\mu, \sigma^2)$ 分配, 則 $Y = \dfrac{X-\mu}{\sigma}$ 具 $N(0, 1)$ 分配.

證 首先, Y 顯然爲一隨機變數. 因若令

$$h : \mathbb{R} \to \mathbb{R} : h(x) = \frac{x-\mu}{\sigma},$$

由於 h 爲一連續函數, 且 $h \circ X = (X-\mu)/\sigma = Y$, 由定理 2.7 知 Y 爲一隨機變數. 次證 Y 具 $N(0, 1)$ 分配. 設 Φ 爲 Y 之累積分配函數, 由於 h 爲遞增且爲蓋射, 而其反函數爲

$$h^{-1} : \mathbb{R} \to \mathbb{R} : h^{-1}(y) = \sigma y + \mu,$$

利用定理 2.22 之 (2), 知 $\forall y \in \mathbb{R}$,

$$\begin{aligned}\Phi(y) &= F_X(h^{-1}(y)) = F_X(\sigma y + \mu)\\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\sigma y+\mu}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)dx. \qquad (*)\end{aligned}$$

再利用微積分基本定理, 可求得 Y 之密度函數爲

$$f(y) = \Phi'(y) = \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{y^2}{2}\Big),\ \forall y \in \mathbb{R}.$$

是以知 Y 具 $N(0,\,1)$ 分配.

[註] 在此, 我們可不必大費周章利用定理 2.22 而求得 $(*)$ 式, 最常使用方法是:

$$\begin{aligned}\Phi(y) &= P\{Y \le y\} = P\Big\{\frac{X-\mu}{\sigma} \le y\Big\} = P\{X \le \sigma y + \mu\}\\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\sigma y+\mu}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)dx.\end{aligned}$$

□

在第四章中, 我們將證明 $N(\mu,\,\sigma^2)$ 中之 μ 爲 X 之期望值, 而 σ 爲 X 之標準差. 一隨機變數減去其期望值再除以非零之標準差稱爲**標準化** (standardized). 本定理之目的在於證明: 具有一般常態分配之隨機變數標準化之後, 則具有標準常態分配. 很多機率與統計學者將大寫希臘字母 Φ 專用以表示標準常態分配 $N(0,\,1)$ 之分配函數. 幾乎所有統計學教本之後都附有 $N(0,\,1)$ 之累積分配函數 Φ 之近似值表, 有些工程用掌上計算器也附有求出標準常態分配值之功能, 在電腦軟體中, 如 SAS, MathCAD 皆然. 我們也可以利用定理 2.23 進一步求得一般常態分配之機率.

例 2. 設 $X \sim N(1,\,4)$, 試求 $P\{-3 \le X \le 3\} =?$

解 由於不定積分

$$\int \frac{1}{2\sqrt{2\pi}}\exp\Big(-\frac{1}{2\cdot 2^2}(x-1)^2\Big)dx$$

無法解出, 是以本題不能利用微積分基本定理第二型求解, 而必須利用近似積分法或利用 (書本或電腦內) $N(0,\,1)$ 之表, 尤其後者更爲快速有效.

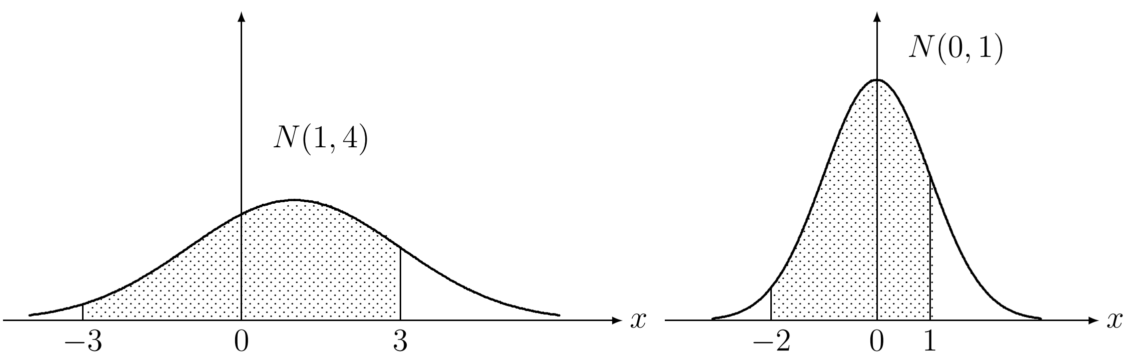


圖 2–26

$$\begin{aligned}
P\{-3 \le X \le 3\} &= P\Big\{\frac{-3-1}{2} \le \frac{X-1}{2} \le \frac{3-1}{2}\Big\} \\
&= P\Big\{-2 \le \frac{X-1}{2} \le 1\Big\} \\
&= P\Big\{\frac{X-1}{2} \le 1\Big\} - P\Big\{\frac{X-1}{2} \le -2\Big\} \\
&= \Phi(1) - \Phi(-2), (\because \tfrac{X-1}{2} \sim N(0,1)) \\
&= \Phi(1) + \Phi(2) - 1, \Big(\because\ \Phi(-x) = 1 - \Phi(x)\Big) \\
&= 0.841345 + 0.977250 - 1, \quad (\text{查表}) \\
&= 0.818595.
\end{aligned}$$ □

定理 2.24

(1) 若 $X \sim N(0,1)$, 則 $X^2 \sim \chi_1^2$;

(2) 若 $X \sim N(\mu, \sigma^2)$, 則 $\Big(\dfrac{X-\mu}{\sigma}\Big)^2 \sim \chi_1^2$.

證 (1) 令 $Y = X^2$, F_Y 爲 Y 之分配函數.

1° 若 $y < 0$, 則

$$F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\} = P(\varnothing) = 0;$$

2° 若 $y = 0$, 則

$$F_Y(y) = P\{Y \le y\} = P\{X^2 \le 0\} = P\{X = 0\} = 0;$$

3° 若 $y > 0$, 則

$$\begin{aligned}
F_Y(y) &= P\{X^2 \le y\} \\
&= P\{-\sqrt{y} \le X \le \sqrt{y}\} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{y}}^{\sqrt{y}} \exp\Big(-\frac{x^2}{2}\Big) dx \\
&= \frac{2}{\sqrt{2\pi}} \int_{0}^{\sqrt{y}} \exp\Big(-\frac{x^2}{2}\Big) dx.
\end{aligned}$$

是以 Y 之密度函數爲

$$\begin{aligned}
f_Y : \mathbb{R} \to \mathbb{R}_+ : f_Y(y) &= \begin{cases} F_Y'(y), & \text{若 } F_Y'(y) \text{ 存在}, \\ 0, & \text{否則}, \end{cases} \\
&= \begin{cases} \dfrac{1}{\sqrt{\pi}\sqrt{2}} y^{-1/2} e^{-y/2}, & \text{若 } y > 0, \\ 0, & \text{若 } y \le 0, \end{cases}
\end{aligned}$$

此乃 χ_1^2 之密度函數, 亦即 $Y \sim \chi_1^2$.

(2) 由定理 2.23 及 (1) 立即可得. □

§ 2.8 隨機變數變換之密度函數

上一節中, 我們討論了 $Y = h \circ X$ 之機率分配函數 P_Y 以及累積分配函數 F_Y 之求法, 對於初學機率的讀者而言, 這二者都遠比密度函數難以理解和掌握. 因此, 本節在實用上極具重要性.

$$
\begin{tikzcd}
\Omega \arrow[r, "X"] \arrow[dr, "Y"'] & \mathbb{R} \arrow[r, "f_X"] \arrow[d, "h"] & \mathbb{R}_+ \\
& \mathbb{R} \arrow[ur, "f_Y"'] &
\end{tikzcd}
$$

在本節中, 我們將分離散型及絕對連續型討論: 如果隨機變數 X 之密度函數 f_X 爲已知, 則 $Y = h \circ X$ 之密度函數 f_Y 爲何?

定理 2.25

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一離散型隨機變數, f_X 爲其機率密度函數. 集合 $A = \{x \in \mathbb{R} \mid f_X(x) > 0\}$; 次設 $h\colon \mathbb{R} \to \mathbb{R}$ 爲 Borel 可測, $Y = h \circ X$, 則

(1) 若 $B = h(A)$, 則 $P\{Y \in B\} = 1$, 因之 Y 爲離散型;

(2) Y 之密度函數爲

$$f_Y : \mathbb{R} \to [0,1] : f_Y(y) = \begin{cases} \displaystyle\sum_{x \in h^{-1}\{y\}} f_X(x), & \text{若 } y \in B, \\ 0, & \text{若 } y \in \mathbb{R} \setminus B \ ; \end{cases}$$

(3) 若 h 爲 A 映至 B 之一對射函數, 則 Y 之密度函數爲

$$f_Y : \mathbb{R} \to [0,1] : f_Y(y) = \begin{cases} f_X(h^{-1}(y)), & \text{若 } y \in B, \\ 0, & \text{若 } y \in \mathbb{R} \setminus B \ . \end{cases}$$

證 (1) 首先, 由於 A 爲至多可數, 故 B 亦然; 其次, 由於

$$\{Y \in B\} = Y^{-1}(B) = (h \circ X)^{-1}(B) = X^{-1}(h^{-1}(B))$$

是以

$$\begin{aligned} P\{Y \in B\} &= P\{X \in h^{-1}(B)\} \\ &\geq P\{X \in A\}, (\because\ A \subset h^{-1}(B)) \\ &= 1. \end{aligned}$$

(2) 由 (1) 知, 只需考慮 $\forall y \in B$,

$$\begin{aligned} f_Y(y) &= P\{Y = y\} = P\{Y \in \{y\}\} \\ &= P\{X \in h^{-1}\{y\}\} \\ &= \sum_{x \in h^{-1}\{y\}} f_X(x). \end{aligned}$$

(3) 當 h 爲 A 映至 B 之一對射函數時, $\forall y \in B$, $\;h^{-1}\{y\} = \{h^{-1}(y)\}$ 之故.
(注意: 上式左端之 h^{-1} 表 h 之像原, 而右端之 h^{-1} 則爲函數 h 之反函數). □

例 1. 設 $X \sim P(\lambda)$, 試求 $Y = X^2 + 2X - 3$ 之密度函數.

解 X 之密度函數爲

$$f_X : \mathbb{R} \to [0,1] : f_X(x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{x!}, & \text{若 } x \in A, \\ 0, & \text{若 } x \notin A, \end{cases}$$

內 $A = \mathbb{Z}_+ = \{0,\, 1,\, 2,\, \cdots\}$. 次設 $h(x) = x^2 + 2x - 3$, 則

(1) $B = h(A) = \{x^2 + 2x - 3 \mid x \in A\} = \{-3,\, 0,\, 5,\, 12,\, \cdots\}$;

(2) $h\colon A \to B$ 爲一對射. (此因 $H(x) = x^2 + 2x - 3$ 之導函數

$$H'(x) = 2x + 2 > 0,\ \forall x > -1,$$

故 H 於區間 $[-1,\, +\infty)$ 爲遞增, 是以 $h = H|_A$ 爲嵌射);

(3) 由於

$$y = h(x) = x^2 + 2x - 3 \;\Leftrightarrow x = -1 + \sqrt{y+4}$$
$$(x = -1 - \sqrt{y+4} \text{ 不合}, \quad \text{因 } x \geq 0),$$

是以 $h^{-1}(y) = -1 + \sqrt{y+4}$. 由定理 2.25 之 (3) 知, $Y = X^2 + 2X - 3 = h \circ X$ 之密度函數爲

$$f_Y : \mathbb{R} \to [0,1] : f_Y(y) = \begin{cases} f_X(h^{-1}(y)) = *, & \text{若 } y \in B, \\ \quad 0, & \text{若 } y \in \mathbb{R} \setminus B, \end{cases}$$

內 $* = e^{-\lambda}\dfrac{\lambda^{\sqrt{y+4}-1}}{(\sqrt{y+4}-1)!}$. □

定理 2.26 (連續型且變換爲一對射)

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一連續型隨機變數, f_X 爲其機率密度函數, 集合 $A = \{x \in \mathbb{R} \mid f_X(x) > 0\}$; 次設可測函數 $h\colon \mathbb{R} \to \mathbb{R}$ 滿足

(1) $h|_A\colon A \to B = h(A)$ 爲一對射;

(2) $h|_A$ 之反函數 $h^{-1}\colon B \to A$ 爲連續可微;

(3) $\forall y \in B,\ (h^{-1})'(y) \neq 0.$

則 $Y = h \circ X$ 之密度函數爲

$$f_Y : \mathbb{R} \to \mathbb{R}_+ : f_Y(y) = \begin{cases} f_X(h^{-1}(y)) \cdot |(h^{-1})'(y)|, & \text{若 } y \in B, \\ 0, & \text{若 } y \notin B, \end{cases}$$

證 我們只需證明: $\forall C \in \mathcal{B},\ P\{Y \in C\} = \displaystyle\int_C f_Y(y)dy$.

1° 若 $C \cap B = \varnothing$, 則

$$\begin{aligned} \text{左} &= P\{Y \in C\} \leq P\{Y \in \complement B\} = 1 - P\{Y \in B\} \\ &= 1 - P\{X \in A\} = 0, \\ \text{右} &= \int_C f_Y(y)dy = 0,\ (\because\ \forall y \in \complement B,\ f_Y(y) = 0.) \end{aligned}$$

2° 若 $C \subset B$, 則

$$\begin{aligned} \text{左} &= P\{Y \in C\} = P\{X \in h^{-1}(C)\} \\ &= \int_{h^{-1}(C)} f_X(x)dx \\ &= \int_C f_X(h^{-1}(y)) \cdot |(h^{-1})'(y)|dy, (*) \\ &= \int_C f_Y(y)dy = \text{右}. \end{aligned}$$

$(*)$ 註: 做變換 $x = h^{-1}(y)$, 參見 Apostol [1] Theorem 15.11.

3° 對於一般的 $C \in \mathcal{B}$, 由於 $C = (B \cap C) + (C \setminus B)$, 前者爲 B 之子集, 後者與 B 不相交, 利用 1° 及 2° 可得

$$\begin{aligned} \text{左} &= P\{Y \in C\} \\ &= P\{Y \in B \cap C\} + P\{Y \in C \setminus B\} \\ &= \int_{B \cap C} f_Y(y)dy + \int_{C \setminus B} f_Y(y)dy \\ &= \int_C f_Y(y)dy = \text{右}. \end{aligned}$$

□

例 2. 許多數學或統計軟體中關於 Gamma 分配, 只提供 $G(\alpha,1)$ 之分配函數及密度函數之值, 設 $X\sim G(\alpha,1)$, $Y\sim G(\alpha,\beta)$.

(a) 試以 F_X 表 F_Y;

(b) 試以 f_X 表 f_Y.

$\boxed{\text{解}}$ (a) 令 $Y=\beta X$, 則由第七節例 1 知, $Y\sim G(\alpha,\beta)$.

$$F_Y(y)=P\{Y\le y\}=P\{X\le\frac{y}{\beta}\}=F_X(\frac{y}{\beta}).$$

例如: $\alpha=5,\ \beta=3$, 則 $F_Y(15)=F_X(\frac{15}{3})=F_X(5)=0.5595.$

(b) 由 (a), 利用定理 2.20,

$$\begin{aligned} f_Y(y)&=\begin{cases} F_Y'(y), & \text{若 } F_Y'(y) \text{ 存在}, \\ 0, & \text{否則}, \end{cases} \\ &=\begin{cases} f_X(\frac{y}{\beta})\cdot\frac{1}{\beta}, & \text{若 } y>0, \\ 0, & \text{若 } y\le 0, \end{cases} \end{aligned}$$

例如: $\alpha=5,\ \beta=3$, 則

$$f_Y(15)=f_X(\frac{15}{3})\cdot\frac{1}{3}=0.0585. \qquad \square$$

註: 本例亦可利用定理 2.26 解之, 但過程稍繁.

例 3. 設X 具對數常態分配, 即其密度函數爲

$$f:\mathbb{R}\to\mathbb{R}_+:f(x)=\begin{cases}\dfrac{1}{x\beta\sqrt{2\pi}}\exp\Big[-\dfrac{(\ln x-\ln\alpha)^2}{2\beta^2}\Big], & \text{若 } x>0, \\ 0, & \text{若 } x\le 0,\end{cases}$$

內 $\alpha>0,\ \beta>0$, 試證: $Y=\ln\circ X$ 具常態分配.

$\boxed{\text{解}}$

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "Y"'] & \mathbb{R}_+^*\subset\mathbb{R} \arrow[r, "f_X"] \arrow[d, "h"] & \mathbb{R}_+ \\ & \mathbb{R} \arrow[ur, "f_Y"'] & \end{tikzcd}$$

令 $h\colon\mathbb{R}_+^*\to\mathbb{R}:h(x)=\ln x$, (註: $\mathbb{R}_+^*=(0,+\infty)$), 則

1° $Y=\ln\circ X=h\circ X$;

2° $B=h(A)=\{\ln x\mid x>0\}=\mathbb{R}$;

3° $h\colon A\to B$ 爲一對射, (此乃微積分中之定理);

$4°$ $h^{-1} = \ln^{-1} = \exp\colon B(=\mathbb{R}) \to A(=\mathbb{R}^*_+)$. 且

$$\forall y \in B,\ (h^{-1})'(y) = e^y.$$

是以 Y 之密度函數爲

$$\begin{aligned} f_Y\colon \mathbb{R} \to \mathbb{R}_+ : f_Y(y) &= f_X(h^{-1}(y)) \cdot |(h^{-1})'(y)| \\ &= f_X(e^y) \cdot e^y \\ &= \frac{1}{\beta\sqrt{2\pi}} \exp\left[-\frac{(y-\ln\alpha)^2}{2\beta^2}\right]. \end{aligned}$$

亦即 $Y \sim N(\ln\alpha, \beta^2)$. □

定理 2.27 (連續型且變換不爲一對射)

設 X 爲上一隨機變數, f_X 爲其密度函數, $A = \{x \in \mathbb{R} \mid f_X(x) > 0\}$; 次設可測函數 $h\colon \mathbb{R} \to \mathbb{R}$ 滿足

(1) $h_j = h|_{A_j}\colon A_j \to B_j$ 爲一對射; 內 $\displaystyle\sum_{j=1}^r A_j = A,\ B_j = h(A_j),\ B = \bigcup_{j=1}^r B_j$;

(2) $\forall j \in \{1, \cdots, r\},\ h_j^{-1}\colon B_j \to A_j$ 爲連續可微, 且 $\forall y \in B_j,\ (h_j^{-1})'(y) \neq 0$.

則 $Y = h \circ X$ 之密度函數爲

$$f_Y\colon \mathbb{R} \to \mathbb{R}_+ : f_Y(y) = \begin{cases} \sum_{j=1}^r I_{B_j}(y) f_X(h_j^{-1}(y)) \cdot |(h_j^{-1})'(y)|, & \text{若 } y \in B, \\ 0, & \text{若 } y \notin B, \end{cases}$$

其中 $B = \displaystyle\bigcup_{j=1}^r B_j$.

證 只需證明: $\forall D \in \mathcal{B},\ P\{Y \in D\} = \displaystyle\int_D f_Y(y)dy$. 我們若能證明:

$$\text{“ } D \subset B_1 \cap B_2,\ \text{且 } D \cap (B_3 \cup \cdots \cup B_r) = \varnothing \text{ ”}$$

之情形, 其餘皆相似. 爲此

$$\begin{aligned} P\{Y \in D\} &= P\{X \in h^{-1}(D)\} \\ &= P\{X \in C_1\} + P\{X \in C_2\}, \qquad (\dagger) \\ &= \int_{C_1} f_X(x)dx + \int_{C_2} f_X(x)dx = * \end{aligned}$$

(†) $C_1 = h^{-1}(D) \cap A_1,\ C_2 = h^{-1}(D) \cap A_2.$

對於上述第一積分, 做變換 $x = h_1^{-1}(y)$, 第二積分, 做變換 $x = h_2^{-1}(y)$, 則

$$\begin{aligned}
* &= \int_D f_X(h_1^{-1}(y)) \cdot |(h_1^{-1})'(y)|\, dy + \int_D f_X(h_2^{-1}(y)) \cdot |(h_2^{-1})'(y)|\, dy, \\
&= \int_D \sum_{j=1}^{2} f_X(h_j^{-1}(y)) \cdot |(h_j^{-1})'(y)|\, dy \\
&= \int_D f_Y(y)(y)\, dy.
\end{aligned}$$

□

例 4. 設 $X \sim N(0,\, 1)$, 試利用定理 2.27 證明: $Y = X^2$ 具 χ_1^2 分配.

$\boxed{\text{解}}$ 爲方便計, X 之密度函數可寫爲

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \exp\Bigl(\dfrac{-x^2}{2}\Bigr), & \text{若 } x \neq 0, \\ 0, & \text{若 } x = 0. \end{cases}$$

若令 $h\colon \mathbb{R} \to \mathbb{R} : h(x) = x^2$, 則

1° $h \circ X = X^2 = Y;$

2° $A = \mathbb{R}^*,\ B = h(A) = (0,\ +\infty);$

令 $A_1 = (0,\, +\infty),\ A_2 = (-\infty,\, 0),$ 則

$$A_1 + A_2 = A,\ B_1 = h(A_1) = B,\ B_2 = h(A_2) = B.$$

3° 顯然

$$h_1\colon A_1 \to B_1 : h_1(x) = x^2,$$
$$h_2\colon A_2 \to B_2 : h_2(x) = x^2,$$

均爲對射, 且

$$h_1^{-1}\colon B_1 \to A_1 : h_1^{-1}(y) = \sqrt{y},$$
$$h_2^{-1}\colon B_2 \to A_2 : h_2^{-1}(y) = -\sqrt{y}.$$

由定理 2.27 知, $Y = X^2$ 之密度函數爲

$$f_Y\colon \mathbb{R} \to \mathbb{R}_+ : f_Y(y) = \begin{cases} *, & \text{若 } y \in B, \\ 0, & \text{若 } y \notin B. \end{cases}$$

內

$$\begin{aligned}
* &= f_X(h_1^{-1}(y)) \cdot |(h_1^{-1})'(y)| + f_X(h_2^{-1}(y)) \cdot |(h_2^{-1})'(y)| \\
&= \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} \\
&= \frac{1}{\sqrt{\pi}2^{1/2}} y^{(1/2)-1} e^{-y/2}.
\end{aligned}$$

亦即 $Y \sim \chi_1^2$. □

第二章 習 題

隨機變數

2-1 若 X 爲有限集合 Ω 上一隨機變數. 試問 X 是否必爲一簡單隨機變數? 理由爲何?

2-2 設 $\Omega = [0,\ 2\pi] \subset \mathbb{R},\ X\colon \Omega \to \mathbb{R} : X(\omega) = \sin(\omega) + \cos(\omega)$. 試繪出 $X,\ X^+,\ X^-$ 及 $|X|$ 之圖形.

2-3 試證: $X^+ = X \cdot I_{\{X\geq 0\}},\ X^- = -X \cdot I_{\{X\leq 0\}}$, 其中 I 爲指標函數, 而 $X\colon \Omega \to \mathbb{R}$.

2-4 試證: 隨機變數間之 "殆必相等"爲一對等關係.

2-5 設 $X,\ Y,\ Z$ 爲 $(\Omega,\ \mathcal{F},\ P)$ 上隨機變數, 試證: $\forall \epsilon > 0$,

$$P\{|X-Y| \geq \epsilon\} \leq P\{|X-Z| \geq \epsilon/2\} + P\{|Y-Z| \geq \epsilon/2\}.$$

2-6 設 $\Omega,\ \Omega'$ 皆爲非空集合, $\mathcal{F}'$ 爲 Ω' 上一 σ 域, $X\colon \Omega \to \Omega'$, 試證:

$$X^{-1}(\mathcal{F}') \underset{\text{def}}{=} \{X^{-1}(B) \mid B \in \mathcal{F}'\}$$

爲 Ω 上一 σ 域.

分配函數及密度函數

2-7 設一骰連擲二次.

(1) 寫出機率空間 $(\Omega,\ \mathcal{F},\ P)$;

(2) X 表第一骰之點數減去第二骰之點數, 試求X之機率分配函數 P_X;

(3) 試求 X 之密度函數;

(4) 試求 $P\{|X-1| > 2\} = ?$

2-8 設X爲一隨機變數, $G(x) = P\{-x < X < x\}$.

(1) 試證: G 爲單調不減函數;

(2) 試問 G 是否爲右連續? 是否爲左連續? 理由?

2-9 某一餐廳之經驗: 有 1/4 預訂之桌位, 客人並未光臨. 今此餐廳有 60 個桌子, 卻接受 62 位客人之預約 (每人一桌). 若 X 表 62 人中光臨之人數, 試寫出 X 之密度函數, 並以之求此餐廳足以應付客人之機率.

2-10 一鏈條有 20 環節, 每一環節若以 1000 公斤拉之, 則斷裂之機率爲 $p = 0.01$. 試問以 1000 公斤拉此鏈條, 則其不斷裂之機率爲若干? 二個以上環節斷裂之機率爲若干?

2-11 烏龍飛機公司製造的飛機引擎每具損壞機率爲 p $(1/2 < p < 1)$, 若每個引擎是否損壞互不影響, 且一飛機若有超過半數之引擎正常運轉, 則可正常飛行, 試問五個引擎之飛機是否比三個引擎之飛機更能正常飛行?

2-12 設 $X \sim N(\mu,\ 0.16)$, 試求常數 c 使得 $P\{|X-\mu| \le c\} = 0.9$.

2-13 設 $X \sim N(\mu,\ \sigma^2)$, F爲 X 之分配函數. 試求 F 圖形之反曲點.

2-14 設 P_X 及 F_X 分別爲隨機變數 X 之機率分配函數及累積分配函數. 證明 :

(1) $P_{aX+b}\{x\} = P_X\left\{\dfrac{x-b}{a}\right\}$, 其中 $a \neq 0$.

(2) $F_{X^2}(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}) + P_X\{-\sqrt{y}\}$, 其中 $y > 0$.

2-15 設 $X \sim N(0,\ 1)$, k 爲介於 0 與 1 之間之常數, $a,\ b \in \mathbb{R},\ a < b$, 滿足

$$P\{a < X < b\} = k.$$

若 $L = b - a$, 試證: L 爲最小之充要條件爲 $a = -b$.

2-16 設 X 爲一隨機變數, 若 $m \in \mathbb{R}$ 滿足

$$P\{X \le m\} \ge 1/2,\ P\{X \ge m\} \ge 1/2,$$

則稱 m 爲 X 之**中位數**(median).

(1) 舉一例以說明一隨機變數之中位數未必爲唯一;

(2) 若 X 爲連續型, 其中位數是否必爲唯一? 理由爲何?

(3) 若 $X \sim N(0, 1)$, 試求隨機變數 $|X|$ 之中位數.

2-17 設 $X \sim NB(1, p)$, (幾何分配), 試證:

$$P\{X > n + m \mid X \ge n\} = P\{X > m\}, \forall n, m \in \mathbb{Z}_+.$$

隨機變數之變換

2-18 設 X 具 (平移之) 幾何分配, 其密度函數爲

$$f\colon \mathbb{R} \to [0,\ 1] : f(x) = \begin{cases} (1/2)^x, & \text{若 } x \in \mathbb{N}, \\ 0, & \text{否則.} \end{cases}$$

試求 $Y = \sin(\pi X/2)$ 之分配.

2-19 設 $X \sim N(2,\ 3)$, 函數

$$h\colon \mathbb{R} \to \mathbb{R} : h(x) = \begin{cases} [\sqrt[4]{x^2}], & 若\ |x| < 4, \\ 2, & 否則, \end{cases}$$

內 $[\cdot]$ 爲最大整數函數. 試求 $Y = h \circ X$ 之密度函數.

2-20 設絕對連續型隨機變數 X 之分配函數 $F\colon \mathbb{R} \to [0,\ 1]$ 爲一遞增函數. 試證: 隨機變數 $Y = F \circ X$ 具 $U(0,\ 1)$ 分配.

[註: 若 F 不爲一遞增時, 本題之結論亦眞, 唯證明較爲困難.]

2-21 設 $X \sim B(n,\ p)$, 次設

$$h\colon \{0,\ 1,\ 2, \cdots,\ n\} \to \mathbb{R} : h(x) = (-1)^x.$$

試證: $Y = h \circ X$ 之密度函數爲

$$f_Y\colon \mathbb{R} \to [0,\ 1] : f_Y(y) = \begin{cases} (1 + (2p-1)^n)/2, & 若\ y = 1, \\ (1 - (2p-1)^n)/2, & 若\ y = -1, \\ 0, & 否則. \end{cases}$$

2-22 設 X 具負指數分配 $G(1,\ 1/\lambda)$, c 爲一常數. 試求 $Y = X + c$ 之密度函數.

2-23 設 X 具常態分配 $N(1,\ 2), h\colon \mathbb{R} \to \mathbb{R} : h(x) = \begin{cases} x, & 若\ |x| < 3, \\ 0, & 若\ |x| \geq 3. \end{cases}$

(1) 試求 $Y = h \circ X$ 之分配函數;

(2) 將 Y 之分配函數化爲 $F = \alpha F_d + (1-\alpha) F_a$ 時, $\alpha =?$ $F_d =?$ $F_a =?$ (參見第六節).

2-24 設某大學男生之身高 $X \sim N(171, 5.2^2)$ (單位: 公分), 今只考慮身高在 170 以上之男生.

(1) 試求身高在 170 以上之男生佔全校男生之百分比;

(2) 試寫出其機率空間、隨機變數、分配函數及密度函數; [這種分配稱爲截尾分配 (truncated distribution)].

(3) 試求身高在 171 以下之機率爲若干?

Chapter 3

隨機向量

上一章中, 我們利用實值函數之方法, 將非數量性的樣本空間 Ω 予以數量化, 也就是所謂隨機變數. 由於有了分配函數以及密度函數, 使我們對於由隨機變數引進的數量性樣本空間上的機率分布有很清晰的了解. 當然我們自然會想到這種函數也可以是二維以上的. 例如 Ω 表某大學學生集合, 我們除了單獨研究學生之身高或體重之分配外, 是否也希望了解二者之間所存在的關聯. 最好的方法乃是利用所謂向量函數, 即令

$$(X, Y) : \Omega \to \mathbb{R}^2 : \omega \mapsto (X(\omega), Y(\omega)),$$

其中 $X(\omega)$ 表學生 ω 之身高, $Y(\omega)$ 表學生 ω 之體重. 這種函數有以下二個優點:

一、可提供單一因素 (身高或體重) 之分配;

二、可提供對二因素關聯之了解.

當然透過嚴格的數學形式, 予以界定並且探討其各項性質仍是必須的.

§ 3.1　隨機向量及其其分配

定義 3.1

設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間. 如果函數 $\mathbf{X}\colon \Omega \to \mathbb{R}^k$ 滿足

$$\forall B \in \mathcal{B}^k,\ \mathbf{X}^{-1}(B) \in \mathcal{F}$$

則稱 $\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一k **維隨機向量**(random vector).

註: (1) 由上述定義顯然可知, 一維隨機向量即隨機變數.

(2) 在微積分中, 我們知道, 由於 $\mathbf{X}(\omega) \in \mathbb{R}^k$, 故可表爲

$$\mathbf{X}(\omega) = (X_1(\omega), \cdots, X_k(\omega)),$$

其中 $X_1(\omega), \cdots, X_k(\omega)$ 均爲實數, 因此, 我們可以得到 k 個實值函數 $X_j\colon \Omega \to \mathbb{R}$, $1 \le j \le k$, 並稱其爲 $\mathbf{X}$ 之分量函數 (component functions), 通常寫爲 $\mathbf{X} = (X_1, \cdots, X_k)$.

定理 3.2

設 $\mathbf{X} = (X_1, \cdots, X_k)$ 爲 $(\Omega, \mathcal{F}, P)$ 上 爲一隨機向量.

(1) 若函數 $\mathbf{h}\colon \mathbb{R}^k \to \mathbb{R}^m$ 爲 $(\mathcal{B}^k, \mathcal{B}^m)$ 可測 (亦稱 $\mathbf{h}$ 爲 Borel 可測), 則合成函數 $\mathbf{h}\circ\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一 m 維隨機向量;

(2) 若函數 $\mathbf{h}\colon \mathbb{R}^k \to \mathbb{R}^m$ 爲連續, 則 $\mathbf{h}\circ\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機向量.

證 (1) 由於

$$\begin{aligned} B \in \mathcal{B}^m &\Rightarrow \mathbf{h}^{-1}(B) \in \mathcal{B}^k, \ (\text{因 } \mathbf{h} \text{ 爲可測}) \\ &\Rightarrow \mathbf{X}^{-1}(\mathbf{h}^{-1}(B)) \in \mathcal{F}, \ (\text{因 } \mathbf{X} \text{ 爲一隨機向量}) \\ &\Rightarrow (\mathbf{h}\circ\mathbf{X})^{-1}(B) \in \mathcal{F}. \end{aligned}$$

$$\begin{tikzcd} \Omega \arrow[r, "\mathbf{X}"] \arrow[dr, "\mathbf{Y}"'] & \mathbb{R}^k \arrow[d, "\mathbf{h}"] \\ & \mathbb{R}^m \end{tikzcd}$$

是以 $h\circ\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數.

(2) 連續函數必爲 Borel 可測之故. □

定理 3.3

設函數 $\mathbf{X} = (X_1, \cdots, X_k)\colon \Omega \to \mathbb{R}^k$, 則 $\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機向量之充要條件爲每一分量函數 X_j 均爲 $(\Omega, \mathcal{F}, P)$ 上之隨機變數.

證 (1) 若 $\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機向量, 因投影函數

$$P_j\colon \mathbb{R}^k \to \mathbb{R} : P_j(x_1, \cdots, x_k) = x_j, \ (\text{內 } j \in \{1, \cdots, k\})$$

$$\begin{tikzcd} \Omega \arrow[r, "\mathbf{X}"] \arrow[dr, "X_j"'] & \mathbb{R}^k \arrow[d, "P_j"] \\ & \mathbb{R} \end{tikzcd}$$

均爲連續, 且 $P_j\circ\mathbf{X} = X_j$, 故由定理 3.2 之 (2) 知, X_j 爲 $(\Omega, \mathcal{F}, P)$ 上之一隨機變數.

(2) 若 $X_1, \cdots, X_k$ 均爲 $(\Omega, \mathcal{F}, P)$ 上之 隨機變數, 先證:

$$I_1, \cdots, I_k \text{ 區間} \subset \mathbb{R}, \ \mathbf{X}^{-1}(I_1 \times \cdots \times I_k) \in \mathcal{F} \tag{$*$}$$

此因

$$\begin{aligned}\mathbf{X}^{-1}(I_1\times\cdots\times I_k) &= \{\omega\in\Omega\mid \mathbf{X}(\omega)\in I_1\times\cdots\times I_k\}\\ &= \{\omega\in\Omega\mid (X_1(\omega),\cdots,X_k(\omega))\in I_1\times\cdots\times I_k\}\\ &= \{\omega\in\Omega\mid X_1(\omega)\in I_1,\cdots,X_k(\omega)\in I_k\}\\ &= \bigcap_{j=1}^{k}\{\omega\in\Omega\mid X_j(\omega)\in I_j\}\\ &= \bigcap_{j=1}^{k}X_j^{-1}(I_j)\in\mathcal{F}.\end{aligned}$$

其次, 令

$$\mathcal{C}=\{I_1\times I_2\times\cdots\times I_k\mid I_1,\ \cdots,\ I_k \text{ 爲 } \mathbb{R} \text{ 之子區間 }\},$$

利用 $\mathcal{B}^k$ 之定義、定理 2.3 及 $(*)$ 之結果, 知

$$\mathbf{X}^{-1}(\mathcal{B}^k)=\mathbf{X}^{-1}(\sigma(\mathcal{C}))=\sigma(\mathbf{X}^{-1}(\mathcal{C}))\subset\mathcal{F}.$$

是以 $\mathbf{X}$ 爲 $(\Omega,\mathcal{F},P)$ 上一隨機向量. □

定理 3.4

設 X, Y 均爲 $(\Omega,\mathcal{F},P)$ 上隨機變數, 則 $X+Y$, $X-Y$, XY 均爲 $(\Omega,\mathcal{F},P)$ 上隨機變數.

證 由定理 3.2 知 (X,Y) 爲一隨機向量, 若令

$$h\colon\mathbb{R}^2\to\mathbb{R}: h(x,y)=x+y.$$

顯然 h 爲連續, 且 $X+Y=h\circ(X,Y)$, 是以 $X+Y$ 爲 $(\Omega,\mathcal{F},P)$ 上一隨機變數. $X-Y$, XY 之證同理. □

本定理告訴我們, 二隨機變數之和、差及積均爲隨機變數, 當然我們會問二隨機變數之商 X/Y 是否亦爲隨機變數? 茲討論如下:

一、如果分母 Y 沒有任何限制, 答案顯然是否定的, 例如: $Y(\omega)=0$, $\forall\omega\in\Omega$, 不論 X 是怎樣的隨機變數, 均使 X/Y 之定義域變成空集合, 當然不是個隨機變數.

二、如果 $Y(\omega)\neq 0,\forall\omega\in\Omega$, 仿照上述定理, 亦可證明 X/Y 爲 $(\Omega,\mathcal{F},P)$ 上隨機變數. 但此條件似乎太強, 因爲有許多連續型隨機變數, 如常態隨機變數, 其值域顯然包含 0; 我們若令集合 $N=\{Y=0\}\subset\Omega$; 只要 N 之機率爲零, (這並不表示 N 必爲空集合), 此時函數 X/Y 之定義域爲 $\Omega\setminus N$, 儘管它可能較 Ω 爲小, 但其機率仍爲 1, 只要稍爲修改隨機變數之定義, 此一困難即可予以克服, 本來規定 隨機變數 X 應滿足

(1) X 之定義域爲 Ω, (2) X 之值皆爲實數 (即有限), (3) X 爲可測.

基於實際需要, 我們將其修改爲：

(1) X 殆必界定於 Ω, (2) X 之值殆必爲有限, (3) X 爲殆必可測.

當然我們也以相同方式處理隨機向量之定義.

定理 3.5

> 設 $X,\ Y$ 均爲 $(\Omega,\ \mathcal{F},\ P)$ 上隨機變數, 且 $P\{Y \neq 0\} = 1$, 則 X/Y 爲一隨機變數.

證 如果我們令 $h\colon \mathbb{R} \times \mathbb{R}^* \to \mathbb{R} : h(x, y) = x/y$, 顯然 h 爲連續, 又令集合 $N = \{Y = 0\}$, 由原設知 $P(N) = 0$, 此時函數 X/Y 顯然滿足

(1) X/Y 殆必界定於 Ω, (因其定義域爲 $\Omega \setminus N$.)

(2) X/Y 之值殆必爲有限, (因 $\forall \omega \in \Omega \setminus N,\ Y(\omega) \in \mathbb{R}^*$.)

(3) X/Y 爲殆必可測, (因 $X/Y = h \circ (X,\ Y)$.)

是以 X/Y 爲一隨機變數. □

以下我們將仿照上一章, 將隨機向量予以分類; 爲便於討論, 我們稍加修改, 其實與上一章之定義並無矛盾之處.

定義 3.6

> 設 $\mathbf{X} = (X_1,\ \cdots,\ X_k)$ 爲 $(\Omega,\ \mathcal{F},\ P)$ 上一隨機向量.
>
> (1) 我們稱 $\mathbf{X}$ 爲**離散型**(discrete type), 如果存在一有限或可數集合 $A \in \mathbb{R}^k$ 使得 $P\{\mathbf{X} \in A\} = 1$. 此時, 令函數
>
> $$f_{\mathbf{X}}\colon \mathbb{R}^k \to [0,\ 1] : f_{\mathbf{X}}(\mathbf{x}) = P\{\mathbf{X} = \mathbf{x}\} = \begin{cases} P\{\mathbf{X} = \mathbf{x}\}, & \text{若 } \mathbf{x} \in A, \\ 0, & \text{若 } \mathbf{x} \in \mathbb{R}^k \setminus A, \end{cases}$$
>
> 而稱其爲 $\mathbf{X}$ 之(**機率**) **密度函數** (p.d.f. 或 p.m.f.) 若無誤解之可能, $f_{\mathbf{X}}$ 亦常寫爲 f.
>
> (2) 我們稱 $\mathbf{X}$ 爲**連續型**, 如果 $\forall \mathbf{x} \in \mathbb{R}^k,\ P\{\mathbf{X} = \mathbf{x}\} = 0$.
>
> (3) 我們稱 $\mathbf{X}$ 爲**絕對連續型**, 如果存在一 Lebesgue 可積函數 $f_{\mathbf{X}}\colon \mathbb{R}^k \to \mathbb{R}_+$ 滿足 $\forall B \in \mathcal{B}^k$,
>
> $$P\{\mathbf{X} \in B\} = \int_B f_{\mathbf{X}} = \int_B f_{\mathbf{X}}(x_1,\ \cdots,\ x_k)\, d(x_1,\ \cdots,\ x_k).$$
>
> 此時, 函數 $f_{\mathbf{X}}$ 稱爲 $\mathbf{X}$ 之(**機率**) **密度函數**.

如同隨機變數一樣, 絕對連續型之隨機向量必爲連續型, 但因非絕對連續型之連續型隨機向量用途不大, 故在無誤解之可能時, 常稱絕對連續型爲連續型.

定理 3.7

隨機向量 $\mathbf{X} = (X_1, \cdots, X_k)$ 爲離散型之充要條件爲 $X_1, \cdots, X_k$ 均爲離散型.

證 若 $\mathbf{X}$ 爲離散型, 則存在一至多可數集合 $A \in \mathbb{R}^k$ 使得 $P\{\mathbf{X} \in A\} = 1$, 令集合

$$A_1 = \{x_1 \in \mathbb{R} \mid \exists x_2, \cdots, x_k \in \mathbb{R}, \text{使得}\, (x_1, \cdots, x_k) \in A\}$$

顯然 A_1 爲至多可數, 且

$$\begin{aligned} P\{X_1 \in A_1\} &= P\{X_1 \in A_1,\ X_2 \in \mathbb{R}, \cdots, X_k \in \mathbb{R}\} \\ &\geq P\{(X_1, \cdots, X_k) \in A\} = 1 \end{aligned}$$

是以 X_1 爲離散型. $X_2, \cdots, X_k$ 同理可證之.

反之, 若 $X_1, \cdots, X_k$ 均爲離散型, 則存在有限或可數集合 $A_1, \cdots, A_k \subset \mathbb{R}$ 使得 $P\{X_1 \in A_1\} = 1,\ \cdots,\ P\{X_k \in A_k\} = 1$, 令 $A = A_1 \times \cdots \times A_k$, 則 A 顯然爲至多可數, 且因

$$\begin{aligned} P\{\mathbf{X} \in A\} &= P\Big\{(X_1, \cdots, X_k) \in A_1 \times \cdots \times A_k\Big\} \\ &= P\{X_1 \in A_1,\ \cdots,\ X_k \in A_k\} \\ &= P\Bigg(\bigcap_{j=1}^{k}\{X_j \in A_j\}\Bigg) \\ &= 1, (\text{參見習題1–22}). \end{aligned}$$

是以 $P\{\mathbf{X} \in A\} = 1$. 亦即 $\mathbf{X}$ 爲離散型. □

如果隨機向量 $\mathbf{X} = (X_1, \cdots, X_k)$ 爲絕對連續型, 我們可以證明每一分量 X_j 均爲絕對連續型 (詳第四節). 但若 $X_1, \cdots, X_k$ 均爲絕對連續型, 則隨機向量 $\mathbf{X} = (X_1, \cdots, X_k)$ 未必爲絕對連續型. 例如, $X = Y$ 具 $N(0, 1)$ 分配, 假定 (X, Y) 爲絕對連續型, 即存在 $f\colon \mathbb{R}^2 \to \mathbb{R}_+$ 滿足

$$\forall B \in \mathcal{B}^2, \iint_B f(x, y)\, dx\, dy = P\{(X, Y) \in B\}. \tag{1}$$

令 $L = \{(x, y) \in \mathbb{R}^2 \mid x = y\}$, 則

$$\begin{aligned} 1 = P\{X = Y\} &= P\{(X, Y) \in L\} \\ &\Rightarrow\ \iint_{\mathbb{R}^2 \setminus L} f(x, y)\, dx\, dy = P\{(X, Y) \in \mathbb{R}^2 \setminus L\} = 0 \\ &\Rightarrow\ f = 0 \text{ a. e.}, (\text{利用實變函數論知 } L \text{ 之 Lebesgue 測度爲 } 0) \\ &\Rightarrow\ P\{(X, Y) \in L\} = \iint_L f(x, y)\, dx\, dy = 0. \end{aligned}$$

但 $P\{(X, Y) \in L\} = 1$, 此一矛盾說明 (X, Y) 不為絕對連續型, 而為奇異連續型.

雖然 $X_1, \cdots, X_k$ 均為連續型時, $\mathbf{X} = (X_1, \cdots, X_k)$ 未必為絕對連續型. 但當 $X_1, \cdots, X_k$ 為獨立時 (詳見第六章), 則 $\mathbf{X} = (X_1, \cdots, X_k)$ 必為絕對連續型.

§ 3.2 常用隨機向量之分配

(一) 多項分配(Multinomial distribution)

有一袋子, 內置 k 種不同顏色之球各若干 (未必相等), 今以"放回"方式隨機抽取 n 個, 假設抽得 C_1 色球之機率為 $p_1, \cdots$, 抽得 C_k 色球之機率為 p_k, $(p_1 + \cdots + p_k = 1)$, 則此試驗之

- 樣本空間為 $\Omega = \{\omega = (a_1, a_2, \cdots, a_n) \mid \forall j,\ a_j \in \{C_1, \cdots, C_k\}\}$;
- σ 域為 $\mathcal{F} = \mathcal{P}(\Omega)$, (因為 $\#\Omega = k^n$ 有限);
- 機率測度為 $P\colon \mathcal{F} \to [0, 1] : P\{\omega\} = p_1^{x_1} \cdots p_k^{x_k}$. (內 ω 中有 x_1 個 $C_1, \cdots, x_k$ 個 C_k).

次設

$$\begin{aligned}
&X_1\colon \Omega \to \mathbb{R} : X_1(\omega) = \omega \text{ 中 } C_1 \text{ 色球數目},\\
&\qquad\vdots\\
&X_k\colon \Omega \to \mathbb{R} : X_k(\omega) = \omega \text{ 中 } C_k \text{ 色球數目},\\
&\mathbf{X} = (X_1, \cdots, X_k)\colon \Omega \to \mathbb{R}^k : \mathbf{X}(\omega) = (X_1(\omega), \cdots, X_k(\omega)),
\end{aligned}$$

則

(1) $\mathbf{X}$ 為一隨機向量, (此因 $\mathcal{F} = \mathcal{P}(\Omega)$) ;

(2) $\mathbf{X}$ 為離散型, 因為 $\mathbf{X}$ 之值域為

$$A = X(\Omega) = \{(x_1, \cdots, x_k) \mid x_j \in \mathbb{Z}_+, x_1 + \cdots + x_k = n\},$$

顯然為有限;

(3) $\mathbf{X}$ 之密度函數為

$$f\colon \mathbb{R}^k \to [0, 1] : f(\mathbf{x}) = P\{\mathbf{X} = \mathbf{x}\} = \begin{cases} *, & \text{若 } \mathbf{x} \in A,\\ 0, & \text{若 } \mathbf{x} \in \mathbb{R}^k \setminus A, \end{cases}$$

內 $\mathbf{x} = (x_1, \cdots, x_k), * = P\{X_1 = x_1,\ \cdots,\ X_k = x_k\} = \binom{n}{x_1, \cdots, x_k} p_1^{x_1} \cdots p_k^{x_k}$,

$\left(\text{理由仿二項分配, 其中組合符號 } \binom{n}{x_1, \cdots, x_k} = \dfrac{n!}{x_1! \cdots x_k!}, \text{ 參見本書附錄一}\right)$,

此種分配稱為**多項分配**. 記為 $\mathbf{X} \sim M_k(n, \mathbf{p})$.

(二) 二元常態分配(Bivariate Normal distribution)

若函數

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Big(-\frac{Q(x,y)}{2}\Big),$$

其中

$$Q(x,y) = \frac{1}{1-\rho^2}\Big[\Big(\frac{x-\mu_1}{\sigma_1}\Big)^2 - 2\rho\Big(\frac{x-\mu_1}{\sigma_1}\Big)\Big(\frac{y-\mu_2}{\sigma_2}\Big) + \Big(\frac{y-\mu_2}{\sigma_2}\Big)^2\Big],$$

$\mu_1,\ \mu_2 \in \mathbb{R},\ \sigma_1,\ \sigma_2 > 0,\ -1 < \rho < 1$, 爲隨機向量 $\mathbf{X} = (X_1,\ X_2)$ 之一密度函數，則稱 $\mathbf{X}$ 具**二元常態分配**. 而 $\mu_1,\ \mu_2,\ \sigma_1,\ \sigma_2,\ \rho$ 稱爲此分配之**參數**或**母數**. 若 $\mu_1 = 0,\ \mu_2 = 0,\ \sigma_1 = 1,\ \sigma_2 = 1,\ \rho = 0$, 則稱 $\mathbf{X}$ 具**二元標準常態分配**.

二元常態分配之密度函數之圖形, 狀如山形, 且於點 $(\mu_1,\ \mu_2)$ 有最大值. ρ 爲 X_1 與 X_2 之相關係數 (下一章), 當 $\rho = 0$ 且 $\sigma_1 = \sigma_2$ 時, 此山之等高線爲一圓, 否則爲一橢圓. 參見以下二圖,

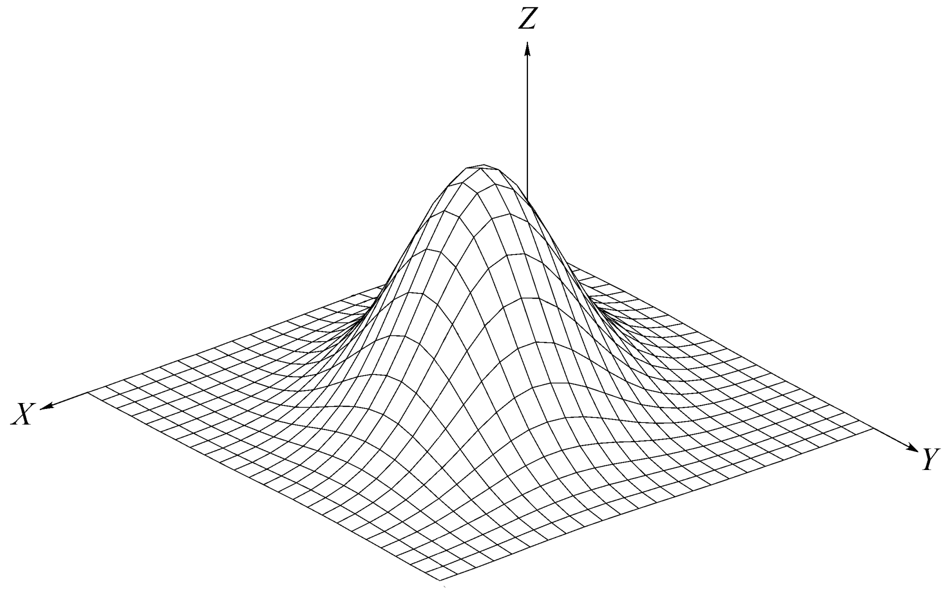


圖 3–1A　二元常態分配之密度函數, $\rho = 0.3$

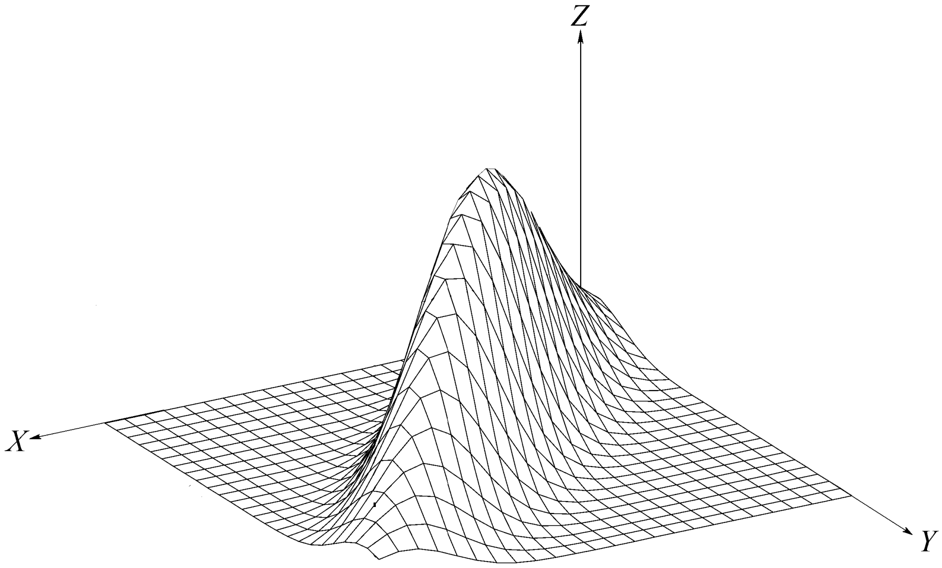


圖 3–1B　二元常態分配之密度函數, $\rho = 0.9$

(三)　**二元均勻分配(Bivariate uniform distribution)**

若函數

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,y) = \begin{cases} \dfrac{1}{c}, & \text{若 } (x,y) \in A, \\ 0, & \text{否則}, \end{cases}$$

爲隨機向量 (X, Y) 之一密度函數, 則稱 (X, Y) 於 A 上具**二元均勻分配**. 其中 $A \subset \mathbb{R}^2$, 其 Lebesgue 測度 (即面積) 爲 $c = \lambda(A) > 0$.

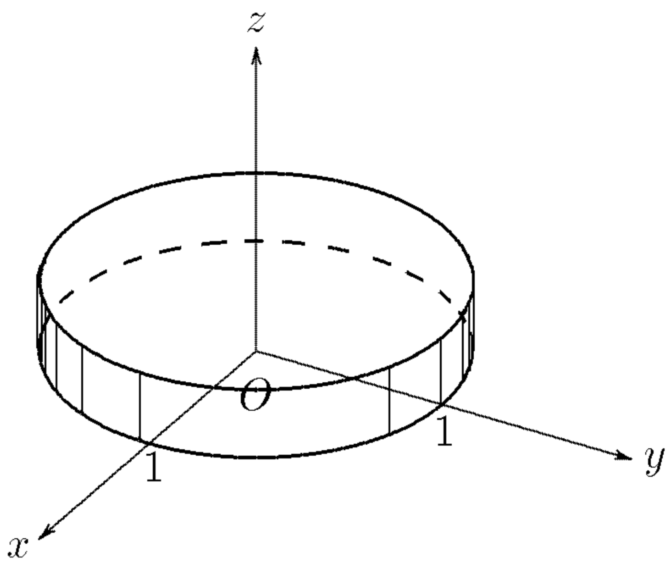


圖 3–2

例 1. 設 $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \le 1\}$, 顯然 $\lambda(A) = \pi$, 若函數

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \begin{cases} \dfrac{1}{\pi}, & \text{若 } (x, y) \in A, \\ 0, & \text{否則}, \end{cases}$$

為隨機向量 (X, Y) 之一密度函數, 則 (X, Y) 於 A 上具二元均勻分配. 參見圖 3-2.

§ 3.3 隨機向量之分配函數

定義 3.8

設 $\mathbf{X} = (X_1, \cdots, X_k)$ 為 $(\Omega, \mathcal{F}, P)$ 上一隨機向量, 則函數

(1) $P_{\mathbf{X}}\colon \mathcal{B}^k \to [0, 1] : P_{\mathbf{X}}(B) = P\{\mathbf{X} \in B\}$ 稱為 $\mathbf{X}$ 之**機率分配函數**;

(2) $$F_{\mathbf{X}}\colon \mathbb{R}^k \to [0, 1] : F_{\mathbf{X}}(x_1, \cdots, x_k) = P\{X_1 \le x_1, \cdots, X_k \le x_k,\} \\ = P_{\mathbf{X}}\big((-\infty, x_1] \times \cdots \times (-\infty, x_k]\big)$$

稱為 $\mathbf{X}$ 之(**累積**) **分配函數**(c.d.f. 或 d.f.), 若無誤解之可能時, $F_{\mathbf{X}}$ 亦常寫為 F.

例 1. 設 (X, Y) 之密度函數為 $f\colon \mathbb{R}^2 \to \mathbb{R}_+$:

$f(x,y)$	$y=0$	$y=1$	$y=3$	$y=5$
$x=-1$	1/12	1/6	1/12	0
$x=3$	1/6	1/12	0	1/12
$x=4$	0	1/12	1/6	1/12

[註: 上表意思是 $f(-1, 0) = 1/12, \cdots;$] 在其他點上, $f(x, y) = 0$, 試求 (X, Y) 之分配函數.

解

$F(x,y)$	$y<0$	$0 \le y < 1$	$1 \le y < 3$	$3 \le y < 5$	$5 \le y$
$x < -1$	0	0	0	0	0
$-1 \le x < 3$	0	1/12	3/12	4/12	4/12
$3 \le x < 4$	0	3/12	6/12	7/12	8/12
$4 \le x$	0	3/12	7/12	10/12	1

□

$\mathbf{X} = (X_1, \cdots, X_k)$ 之 (累積) 分配函數 $F_{\mathbf{X}}$ 亦常稱為 $X_1, \cdots, X_k$ 之 **聯合分配函數**(joint distribution function), 而寫為 $F_{X_1,\cdots,X_k}$. 同理, $\mathbf{X}$ 之密度函數 $f_{\mathbf{X}}$ 亦可稱為 $X_1, \cdots, X_k$ 之 **聯合密度函數**(joint density function), 而寫為 $f_{X_1,\cdots,X_k}$.

定理 3.9

設 F 爲隨機向量 $\mathbf{X} = (X,\ Y)$ 之分配函數, 則

(1) 若 $x_1,\ x_2,\ y_1,\ y_2 \in \mathbb{R}$ 且 $x_1 < x_2,\ y_1 < y_2$, 則

$$\begin{aligned} &P\{x_1 < X \leq x_2,\ y_1 < Y \leq y_2\} \\ &\qquad = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1). \end{aligned}$$

(2) F 對於變數 x 及 y 均分別爲右連續, 意即 $\forall p,\ \forall q \in \mathbb{R}$,

$$\lim_{x \to p^+} F(x,\ q) = F(p,\ q),$$
$$\lim_{y \to q^+} F(p,\ y) = F(p,\ q).$$

(3) $\displaystyle\lim_{x \to -\infty} F(x,\ q) = 0,\ \lim_{y \to -\infty} F(p,\ y) = 0,\ \lim_{\substack{x \to +\infty \\ y \to +\infty}} F(x,\ y) = 1.$

證 (1) 參考圖 3–3, 令

$A = (x_1,\ x_2] \times (y_1,\ y_2]$,

$B_1 = (-\infty,\ x_1] \times (-\infty,\ y_2]$,

$B_2 = (-\infty,\ x_2] \times (-\infty,\ y_1]$,

$C = B_1 \cap B_2 = (-\infty,\ x_1] \times (-\infty,\ y_1]$, (網點部分),

$R = A \cup B_1 \cup B_2 = (-\infty,\ x_2] \times (-\infty,\ y_2]$,

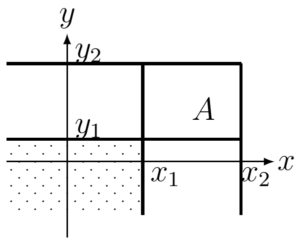


圖 3–3

則 $R = A + (B_1 \cup B_2)$, 是以

$$\begin{aligned} F(x_2, y_2) &= P_{\mathbf{X}}(R) = P_{\mathbf{X}}(A) + P_{\mathbf{X}}(B_1 \cup B_2) \\ &= P_{\mathbf{X}}(A) + P_{\mathbf{X}}(B_1) + P_{\mathbf{X}}(B_2) - P_{\mathbf{X}}(B_1 \cap B_2) \\ &= P_{\mathbf{X}}(A) + P_{\mathbf{X}}(B_1) + P_{\mathbf{X}}(B_2) - P_{\mathbf{X}}(C) \\ &= P_{\mathbf{X}}(A) + F(x_1,\ y_2) + F(x_2,\ y_1) - F(x_1,\ y_1). \end{aligned}$$

故 $F(x_2,\ y_2) - F(x_1,\ y_2) - F(x_2,\ y_1) + F(x_1,\ y_1) = P_{\mathbf{X}}(A)$.

(2) 及 (3) 仿定理 2.11, 讀者自證之. □

定理 3.10

設 $\mathbf{X}$ 爲絕對連續型隨機向量, f 爲其密度函數, $g\colon \mathbb{R}^k \to \mathbb{R}_+$, 則 g 爲 $\mathbf{X}$ 之密度函數之充要條件爲 $f = g$ a.e. (意即: 集合 $\{\mathbf{x} \in \mathbb{R}^k \mid f(\mathbf{x}) \neq g(\mathbf{x})\}$ 爲零測度.)

證 仿定理 2.19 之證. □

定理 3.11

設 F 爲絕對連續型隨機向量 $\mathbf{X}$ 之分配函數, $f\colon \mathbb{R}^k \to \mathbb{R}_+$, 則 f 爲 $\mathbf{X}$ 之密度函數之充要條件爲

$$\forall x_1, \cdots, \forall x_k \in \mathbb{R},\ F(x_1, \cdots, x_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} f(t_1, \cdots, t_k)\, dt_1 \cdots dt_k.$$

證 仿定理 2.20 之證. □

定理 3.12

設 F 爲絕對連續型隨機向量 $\mathbf{X}$ 之分配函數, 若令 $f\colon \mathbb{R}^k \to \mathbb{R}_+$:

$$f(x_1, \cdots, x_k) = \begin{cases} \dfrac{\partial^k F(x_1, \cdots, x_k)}{\partial x_1 \cdots \partial x_k}, & \text{若 } (x_1, \cdots, x_k) \in A, \\ 0, & \text{若 } (x_1, \cdots, x_k) \notin A, \end{cases}$$

內 $A = \left\{(x_1, \cdots, x_k) \in \mathbb{R}^k \ \middle|\ \dfrac{\partial^k F(x_1, \cdots, x_k)}{\partial x_1 \cdots \partial x_k} \text{ 存在} \right\}$, 則 f 爲 $\mathbf{X}$ 之密度函數.

證 仿定理 2.21 之證. □

多維隨機向量之分配函數及密度函數之間尙可導得多項性質, 惟對於初學者似嫌過分繁重, 故予從略.

§ 3.4 邊際分配

設 $\mathbf{X}$ 爲上一 k 維隨機向量. 本節中我們將探討: 如果 $\mathbf{X}$ 之密度函數 $f_{\mathbf{X}}$ 爲已知, 我們是否可以求得各隨機變數 X_j 之密度函數. 爲使讀者易於接受, 我們先舉一例.

例 1. 某班學生共 10 人, 其身高、體重之分組次數表下:

人數	164cm	166cm	170cm	172cm	176cm	合計
56kg	1	-	-	-	-	1
58kg	1	2	1	-	-	4
60kg	-	-	2	-	-	2
62kg	-	1	-	-	-	1
64kg	-	-	-	1	1	2
計	2	3	3	1	1	10

若以比率 (機率) 表示, 則上述各人數應分別乘以 1/10, 即

人數	164cm	166cm	170cm	172cm	176cm	合計
56kg	1/10	-	-	-	-	1/10
58kg	1/10	2/10	1/10	-	-	4/10
60kg	-	-	2/10	-	-	2/10
62kg	-	1/10	-	-	-	1/10
64kg	-	-	-	1/10	1/10	2/10
計	2/10	3/10	3/10	1/10	1/10	1

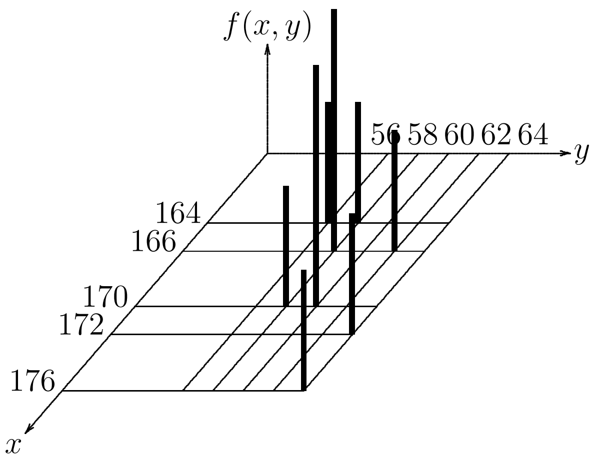


圖 3–4　X 與 Y 之聯合密度函數圖

今欲研究該班學生體重之分配情形, 則僅需將各體重之機率予以合計 (如第二表右欄), 即知該班學生體重爲 56, 58, 60, 62, 64 公斤各佔之比率爲若干. 換言之, 若 X 表身高變數, Y 表體重變數, $f_{X,Y}$ 爲 X 與 Y 之聯合密度函數 (如第二表所示), 則 Y 之密度函數應爲

$$f_Y : \mathbb{R} \to [0,\,1] : f_Y(y) = \begin{cases} 1/10, & \text{若 } y \in \{56,\,62\}, \\ 2/10, & \text{若 } y \in \{60,\,64\}, \\ 4/10, & \text{若 } y \in \{58\}, \\ 0, & \text{否則}, \end{cases}$$

此恰爲第二表右欄之數字, 是以我們有:

定理 3.13

設 $\mathbf{X} = (X_1, X_2)$ 為 $(\Omega, \mathcal{F}, P)$ 上為一離散型隨機向量, (即存在一至多可數集合 $A \subset \mathbb{R}^2$ 使得 $P\{\mathbf{X} \in A\} = 1$), $f_{\mathbf{X}}$ 為其密度函數, 次設

$$\begin{aligned} A_1 &= \{x_1 \in \mathbb{R} \mid \exists x_2 \in \mathbb{R},\ \text{使得 } (x_1, x_2) \in A\}, \\ B(x_1) &= \{x_2 \in \mathbb{R} \mid (x_1, x_2) \in A\}, \text{內 } x_1 \in A_1. \end{aligned}$$

則函數

$$f_1 \colon \mathbb{R} \to [0, 1] : f_1(x_1) = \begin{cases} \displaystyle\sum_{x_2 \in B(x_1)} f_{\mathbf{X}}(x_1, x_2), & \text{若 } x_1 \in A_1, \\ 0, & \text{否則}, \end{cases}$$

為隨機變數 X_1 之密度函數. (亦稱 f_1 為 X_1 之**邊際密度函數** (marginal p.d.f.), 相對於我們稱 $f_{\mathbf{X}} = f_{X_1, X_2}$ 為 X_1, X_2 之**聯合密度函數**(joint p.d.f.)).

證 由於

$$P\{X_1 \in A_1\} = P\{X_1 \in A_1,\ X_2 \in \mathbb{R}\} \geq P\{(X_1, X_2) \in A\} = 1.$$

顯然, 當 $x_1 \notin A_1$ 時, $P\{X_1 = x_1\} = 0 = f_1(x_1)$. 此外, 當 $x_1 \in A_1$ 時,

$$\begin{aligned} P\{X_1 = x_1\} &= P\{X_1 = x_1,\ X_2 \in \mathbb{R}\} \\ &= P\{X_1 = x_1,\ X_2 \in B(x_1)\} \\ &= \sum_{x_2 \in B(x_1)} f(x_1, x_2) = f_1(x_1). \end{aligned}$$

是以函數 f_1 為 X_1 之密度函數. □

當 $x_1 \in A_1$ 時, $f_1(x_1) = \displaystyle\sum_{x_2 \in B(x_1)} f_{\mathbf{X}}(x_1, x_2)$ 簡寫為 $f_1(x_1) = \displaystyle\sum_{x_2} f_{\mathbf{X}}(x_1, x_2)$. 同理, 函數

$$f_2 \colon \mathbb{R} \to [0, 1] : f_2(x_2) = \begin{cases} \displaystyle\sum_{x_1} f_{\mathbf{X}}(x_1, x_2), & \text{若 } x_2 \in A_2, \\ 0, & \text{否則}, \end{cases}$$

其中 $A_2 = \{x_2 \in \mathbb{R} \mid \exists x_1 \in \mathbb{R},\ \text{使得 } (x_1, x_2) \in A\}$, 則為隨機變數 X_2 之 (邊際) 密度函數. 此一觀念我們不難加以推廣, 設 $\mathbf{X} = (X_1, \cdots, X_k)$ 為 $(\Omega, \mathcal{F}, P)$ 上為一 k 維離散型隨機向量, $f_{\mathbf{X}}$ 為其密度函數, 則函數

$$f_1 \colon \mathbb{R} \to [0, 1] : f_1(x_1) = \begin{cases} \displaystyle\sum_{x_2} \cdots \sum_{x_k} f_{\mathbf{X}}(x_1, \cdots, x_k), & \text{若 } x_1 \in A_1, \\ 0, & \text{否則}, \end{cases}$$

內 $A_1 = \{x_1 \in \mathbb{R} \mid \exists x_2, \cdots, x_k \in \mathbb{R},\ (x_1, \cdots, x_k) \in A\}$, 爲隨機變數 X_1 之 (邊際) 密度函數. 而函數

$$f_{1,2}\colon \mathbb{R}^2 \to [0,\,1]:$$
$$f_{1,2}(x_1,\,x_2) = \begin{cases} \displaystyle\sum_{x_3} \cdots \sum_{x_k} f_{\mathbf{X}}(x_1,\,\cdots,\,x_k), & \text{若 } (x_1, x_2) \in A_{1,2}, \\ 0, & \text{否則}, \end{cases}$$

(內 $A_{1,2} = \{(x_1, x_2) \in \mathbb{R}^2 \mid \exists x_3, \cdots, x_k \in \mathbb{R}$ 使得 $(x_1, \cdots, x_k) \in A\})$, 爲隨機向量 (X_1, X_2) 之 (邊際) 密度函數. 其餘同理.

定理 3.14

設 $\mathbf{X} = (X_1,\, X_2)$ 爲 $(\Omega,\, \mathcal{F},\, P)$ 上爲一連續型隨機向量, $f_{\mathbf{X}}$ 爲其一密度函數, 則函數

$$f_1\colon \mathbb{R} \to \mathbb{R}_+ : f_1(x_1) = \int_{-\infty}^{+\infty} f_{\mathbf{X}}(x_1, x_2)\, dx_2,$$
$$f_2\colon \mathbb{R} \to \mathbb{R}_+ : f_2(x_2) = \int_{-\infty}^{+\infty} f_{\mathbf{X}}(x_1, x_2)\, dx_1,$$

分別爲 X_1 及 X_2 之密度函數. (亦稱爲 **邊際密度函數**).

證 設 F_1 爲 X_1 之分配函數, 往證: $\forall x_1 \in \mathbb{R}, F_1(x_1) = \displaystyle\int_{-\infty}^{x_1} f_1(t_1) dt_1$.

$$\begin{aligned} F_1(x_1) &= P\{X_1 \le x_1\} \\ &= P\{X_1 \le x_1,\ X_2 \in \mathbb{R}\} \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{+\infty} f_{\mathbf{X}}(t_1, t_2)\, dt_2\, dt_1 \\ &= \int_{-\infty}^{x_1} f_1(t_1)\, dt_1. \end{aligned}$$

由定理 2.19 知, f_1 爲 X_1 之密度函數. f_2 同理可證之. □

與離散型之情形相似, 我們也可推廣而得以下之邊際分配: 設 $f_{\mathbf{X}}$ 爲絕對連續型隨機向量 $\mathbf{X} = (X_1, \cdots, X_k)$ 之密度函數, 則

$$f_1\colon \mathbb{R} \to \mathbb{R}_+ : f_1(x_1) = \int \cdots \int_{\mathbb{R}^{k-1}} f_{\mathbf{X}}(x_1,\, \cdots,\, x_k)\, dx_2 \cdots dx_k$$

爲 X_1 之 (邊際) 密度函數. 而函數

$$f_{1,2}\colon \mathbb{R}^2 \to \mathbb{R}_+ : f_{1,2}(x_1, x_2) = \int \cdots \int_{\mathbb{R}^{k-2}} f_{\mathbf{X}}(x_1,\, \cdots,\, x_k)\, dx_3 \cdots dx_k$$

爲 $(X_1,\, X_2)$ 之 (邊際) 密度函數.

§ 3.5 條件分配

例 1. (續上節例 1) 在條件 58 公斤中, 我們欲了解身高分配之情況 (換言之, 體重不是 58 公斤者, 均不在考慮之列, 體重爲 58 公斤者按其身高分類之), 則應有

類別 \ 身高	164	166	170	合計
人　數	1	2	1	4
對全班比率 $f(x, 58)$	1/10	2/10	1/10	4/10 *
對 58kg 級比率 $f(x\|58)$	2/4	1/4	1/4	1

由『對全班比率』以求『對 58 級公斤級比率』時, 我們可以『合計比率 *』4/10 遍除各『對全班比率』, 即

$$f(x|58) = \frac{f(x, 58)}{4/10}, \text{內 } x \in \{164,\ 166,\ 170\}.$$

上表中最下面一列亦爲一種密度函數, 我們將稱其爲條件密度函數. □

定義 3.15

設 $\mathbf{X} = (X_1, X_2)$ 爲 $(\Omega, \mathcal{F}, P)$ 上爲一離散型隨機向量, $f_{\mathbf{X}}$ 爲其一密度函數; 次設 $P\{X_2 = x_2\} > 0$, 內 x_2 爲一 (固定) 實數, f_2 爲 X_2 之邊際密度函數, 則函數

$$\begin{aligned} f(\,\cdot\,|x_2)\colon \mathbb{R} \to [0,\ 1] : f(x_1|x_2) &= P\{X_1 = x_1|X_2 = x_2\} \\ &= \frac{P\{X_1 = x_1,\ X_2 = x_2\}}{P\{X_2 = x_2\}} \\ &= \frac{f_{\mathbf{X}}(x_1, x_2)}{f_2(x_2)} \end{aligned}$$

稱爲 **給 $X_2 = x_2$ 後, X_1 之條件密度函數**(conditional p.d.f.). 同理, 讀者可以自行界定給 $X_1 = x_1$ 後, X_2 之條件密度函數.

若 $\mathbf{X} = (X_1, X_2)$ 爲 $(\Omega, \mathcal{F}, P)$ 上爲一連續型之隨機向量, 如果我們規定 $f(x_1|x_2) = P\{X_1 = x_1|X_2 = x_2\}$, 則因 $P\{X_2 = x_2\} = 0$, 此一規定顯然是錯誤的. 爲此, 我們先考慮 $x_2 \in \mathbb{R}$ 使得

$$P\{x_2 \le X_2 \le x_2 + h\} > 0, \text{內 } \forall h > 0.$$

則給事件 $\{x_2 \le X_2 \le x_2 + h\}$ 後 X_1 之條件分配函數應爲

$$P\{X_1 \le x_1 \mid x_2 \le X_2 \le x_2 + h\} = \frac{\int_{x_2}^{x_2+h} \int_{-\infty}^{x_1} f_{\mathbf{X}}(t_1, t_2)\, dt_1\, dt_2}{\int_{x_2}^{x_2+h} f_2(t_2)\, dt_2},$$

若 f 及 f_2 均爲連續函數且 $f_2(x_2)>0$, 則上述機率之極限爲

$$\begin{aligned}\lim_{h\to 0^+} & P\{X_1\le x_1 \mid x_2 \le X_2\le x_2+h\}\\ &= \lim_{h\to 0^+}\frac{\int_{-\infty}^{x_1}\int_{x_2}^{x_2+h} f_{\mathbf{X}}(t_1,t_2)\,dt_2\,dt_1}{\int_{x_2}^{x_2+h} f_2(t_2)\,dt_2}\\ &= \frac{\int_{-\infty}^{x_1} f_{\mathbf{X}}(t_1,x_2)\,dt_1}{f_2(x_2)}, \text{(l'Hospital 規則)}\\ &= \int_{-\infty}^{x_1}\frac{f_{\mathbf{X}}(t_1,x_2)}{f_2(x_2)}\,dt_1.\end{aligned}$$

同理可證

$$\lim_{h\to 0^+} P\{X_1\le x_1 \mid x_2-h\le X_2\le x_2\} = \int_{-\infty}^{x_1}\frac{f_{\mathbf{X}}(t_1,x_2)}{f_2(x_2)}\,dt_1.$$

我們當可視此一極限爲給爲“給 $X_2=x_2$ 後, X_1 之條件分配函數”, 即

$$F(x_1 \mid X_2=x_2) = \int_{-\infty}^{x_1}\frac{f_{\mathbf{X}}(t_1,x_2)}{f_2(x_2)}\,dt_1.$$

利用定理 2.19 知, 函數 $\dfrac{f_{\mathbf{X}}(t_1,x_2)}{f_2(x_2)}$ 乃爲上述條件分配對應之密度函數. 因此, 我們界定如下:

定義 3.16

設 $\mathbf{X}=(X_1,\,X_2)$ 爲 $(\Omega,\,\mathcal{F},\,P)$ 上爲一連續型隨機向量, $f_{\mathbf{X}}$ 爲其一密度函數, 而 f_1 及 f_2 分別爲 X_1 及 X_2 之邊際密度函數.

(1) 若 $x_2\in\mathbb{R}$ 使得 $f_2(x_2)>0$, 則函數

$$f(\,\cdot\,|x_2)\colon \mathbb{R}\to\mathbb{R}_+ : f(x_1|x_2)=\frac{f_{\mathbf{X}}(x_1,x_2)}{f_2(x_2)}$$

稱爲**給 $X_2=x_2$ 後, X_1 之條件密度函數**.

(2) 若 $x_1\in\mathbb{R}$ 使得 $f_1(x_1)>0$, 則函數

$$f(\,\cdot\,|x_1)\colon \mathbb{R}\to\mathbb{R}_+ : f(x_2|x_1)=\frac{f_{\mathbf{X}}(x_1,x_2)}{f_1(x_1)}$$

稱爲 **給 $X_1=x_1$ 後, X_2 之條件密度函數**.

例 2. 設隨機向量 $(X,\,Y)$ 之密度函數爲 $f\colon\mathbb{R}^2\to[0,\,1]$:

$$f(x,y)=\begin{cases}\dfrac{2}{n(n+1)}, & \text{若 } (x,y)\in A,\\ 0, & \text{否則},\end{cases}$$

其中 $n\in\mathbb{N},\ A=\{(x,\,y)\in\mathbb{N}^2 \mid 1\le x\le n,\ 1\le y\le x\}$, 如圖 3–5.

(1) 檢驗 $\sum_{(x,y)\in A} f(x,y)$ 是否等於 1;

(2) 試求 X 及 Y 之邊際密度函數;

(3) 試求: 給 $X = x(\in \{1, \cdots, n\})$ 後 Y 之條件密度函數.

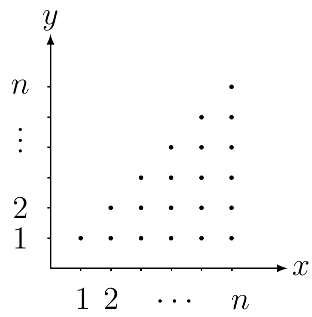


圖 3–5

解 (1) 由圖 3–5 知, A 之基數爲 $n(n+1)/2$, 是以

$$\sum_{(x,y)\in A} f(x,y) = \frac{n(n+1)}{2} \cdot \frac{2}{n(n+1)} = 1.$$

(2) X 之邊際密度函數爲 $f_X\colon \mathbb{R} \to [0, 1]$:

$$f_X(x) = \begin{cases} \displaystyle\sum_{y=1}^{x} f(x,y), & \text{若 } x \in \{1, \cdots, n\}, \\ 0, & \text{否則}, \end{cases}$$

$$= \begin{cases} \dfrac{2x}{n(n+1)}, & \text{若 } x \in \{1, \cdots, n\}, \\ 0, & \text{否則}, \end{cases}$$

Y 之邊際密度函數爲 $f_Y\colon \mathbb{R} \to [0, 1]$:

$$f_Y(y) = \begin{cases} \displaystyle\sum_{x=y}^{n} f(x,y), & \text{若 } y \in \{1, \cdots, n\}, \\ 0, & \text{否則}, \end{cases}$$

$$= \begin{cases} \dfrac{2(n-y+1)}{n(n+1)}, & \text{若 } y \in \{1, \cdots, n\}, \\ 0, & \text{否則}, \end{cases}$$

(3) 給 $X = x(\in \{1, \cdots, n\})$ 後 Y 之條件密度函數爲

$$f(\cdot|x)\colon \mathbb{R} \to [0, 1] : f(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} *, & \text{若 } y \in \{1, \cdots, x\}, \\ 0, & \text{否則}, \end{cases}$$

內 $* = \dfrac{\dfrac{2}{n(n+1)}}{\dfrac{2x}{n(n+1)}} = \dfrac{1}{x}.$ □

例 3. 設連續型隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,y) = \begin{cases} xe^{-x(1+y)}, & \text{若 } x>0,\ y>0, \\ 0, & \text{否則}, \end{cases}$$

(1) 檢驗 $\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy$ 是否等於 1;

(2) 試求 X 及 Y 之邊際密度函數;

(3) 試求: 給 $Y=y\,(>0)$ 後 X 之條件密度函數.

$\boxed{\text{解}}$ (1)

$$\begin{aligned}\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy &= \int_0^{+\infty}\int_0^{+\infty} xe^{-x(1+y)}dy\,dx \\ &= \int_0^{+\infty} xe^{-x}\Big\{\frac{e^{-xy}}{-x}\Big|_{y=0}^{y=+\infty}\Big\}dx \\ &= \int_0^{+\infty} e^{-x}dx = 1.\end{aligned}$$

(2) X 之邊際密度函數爲

$$f_X\colon \mathbb{R} \to \mathbb{R}_+ : f_X(x) = \int_{\mathbb{R}} f(x,y)dy = \begin{cases} *, & \text{若 } x>0, \\ 0, & \text{若 } x\le 0, \end{cases}$$

內 $\quad * = \displaystyle\int_0^{+\infty} xe^{-x(1+y)}dy = e^{-x},\quad$ (參見 (1) 之解).

Y 之邊際密度函數爲

$$f_Y\colon \mathbb{R} \to \mathbb{R}_+ : f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \begin{cases} **, & \text{若 } y>0, \\ 0, & \text{若 } y\le 0, \end{cases}$$

內

$$\begin{aligned} ** &= \int_0^{+\infty} xe^{-x(1+y)}\,dx \\ &= \Gamma(\alpha)\beta^\alpha \int_0^{+\infty} \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}\,dx, \quad (\text{內 } \alpha=2,\ \beta=(1+y)^{-1}) \\ &= \Gamma(\alpha)\beta^\alpha, \quad (\text{被積函數爲 } G(\alpha,\beta) \text{ 密度函數}) \\ &= \frac{1}{(1+y)^2}.\end{aligned}$$

(3) 給 $Y=y\,(>0)$ 後 X 之條件密度函數爲

$$f(\,\cdot\,|y)\colon \mathbb{R} \to \mathbb{R}_+ : f(x|y) = \begin{cases} ***, & \text{若 } x>0, \\ 0, & \text{若 } x\le 0, \end{cases}$$

內 $*** = \dfrac{f(x,y)}{f_Y(y)} = xe^{-x(1+y)}(1+y)^2.$ □

§ 3.6 隨機向量之變換

在上一章第七節, 我們介紹過, 如果一隨機變數 X 之分配爲已知, 則其變換 $Y = h \circ X$ 之機率分配函數、累積分配函數及密度函數亦可對應求出. 本節則進一步探討隨機向量之情形. 利用隨機變數之方法無法由已知 X 及 Y 分配求得 $X+Y$、$X-Y$ 及 XY 等之分配, 在本節中, 我們將可予以解決.

定理 3.17

設 $\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機向量, $P_{\mathbf{X}}$ 爲其機率分配函數, 函數 $\mathbf{h}\colon \mathbb{R}^k \to \mathbb{R}^m$ 爲 Borel 可測, 則函數 $\mathbf{Y} = \mathbf{h} \circ \mathbf{X}$ 之機率分配函數爲

$$P_{\mathbf{Y}}\colon \mathcal{B} \to [0,\, 1] : P_{\mathbf{Y}}(B) = P_{\mathbf{X}}(\mathbf{h}^{-1}(B)).$$

證 仿定理 2.21, 讀者自證之. □

註: 若 $\mathbf{X}$ 之累積分配函數 $F_{\mathbf{X}}$ 爲已知, 而欲求 $\mathbf{Y} = \mathbf{h} \circ \mathbf{X}$ 之分配函數, 則因

$$\begin{aligned} F_{\mathbf{Y}}(\mathbf{y}) &= F_{\mathbf{Y}}(y_1,\, \cdots,\, y_m) \\ &= P\{Y_1 \le y_1,\, \cdots,\, Y_m \le y_m\} \\ &= P\{\mathbf{Y} \in (-\infty,\, y_1] \times \cdots \times (-\infty,\, y_m]\} \\ &= P_{\mathbf{X}}\Big(\mathbf{h}^{-1}\Big((-\infty,\, y_1] \times \cdots \times (-\infty,\, y_m]\Big)\Big). \end{aligned}$$

除非集合 $\mathbf{h}^{-1}\Big((-\infty,\, y_1] \times \cdots \times (-\infty,\, y_m]\Big)$ 甚爲單純, 否則 $F_{\mathbf{Y}}(\mathbf{y})$ 不易表爲 $F_{\mathbf{X}}$ 之函數.

例 1. 設 $(X,\, Y)$ 於 $[0,\, 1] \times [0,\, 1]$ 上具均匀分配, 即其密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \begin{cases} 1, & \text{若 } 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{否則}. \end{cases}$$

試求: $Z = X + Y$ 之分配函數.

解 令函數 $h\colon \mathbb{R}^2 \to \mathbb{R} : h(x, y) = x + y$, 則 $Z = X + Y = h \circ (X,\, Y)$.$\forall z \in \mathbb{R}$,

$$\begin{aligned} F_Z(z) &= P\{Z \le z\} \\ &= P\{h \circ (X,\, Y) \in (-\infty,\, z]\} \\ &= P\{(X,\, Y) \in h^{-1}(-\infty,\, z]\}, \end{aligned}$$

其中

$$\begin{aligned} h^{-1}(-\infty,\, z] &= \{(x,\, y) \in \mathbb{R}^2 \mid h(x,\, y) \in (-\infty,\, z]\} \\ &= \{(x,\, y) \in \mathbb{R}^2 \mid x + y \le z\}, \end{aligned}$$

如圖 3–6 網點部分.

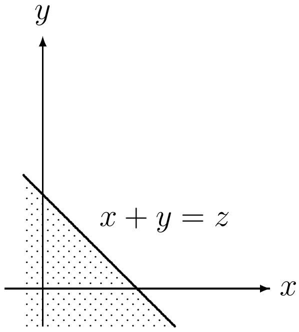


圖 3–6

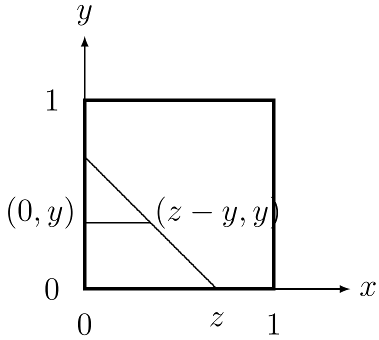


圖 3–7

(1) 若 $z<0$, 則因網點部分與 $[0,\,1]\times[0,\,1]$ 爲互斥, 故 $F_Z(z)=0$.

(2) 若 $0\le z\le 1$, 則因 $(X,\,Y)$ 在 $[0,\,1]\times[0,\,1]$ 上其均勻分配, 故

$$F_Z(z)=\int_0^z\int_0^{z-y}dx\,dy=\frac{z^2}{2}.\quad (\text{參考圖 3–6}).$$

(3) 若 $1<z\le 2$, 同理可得

$$\begin{aligned}F_Z(z)&=1-\int_{z-1}^{1}\int_{z-y}^{1}dx\,dy\\&=1-\int_{z-1}^{1}(1-z+y)dy=1-\frac{(2-z)^2}{2}.\end{aligned}$$

(4) 若 $z>2$, 則因網點部分包含 $[0,\,1]\times[0,\,1]$, 故 $F_Z(z)=1$.

總之, $Z=X+Y$ 之分配函數爲

$$F_Z\colon\mathbb{R}\to[0,\,1]:F_Z(z)=\begin{cases}0, & \text{若 } z<0,\\ \dfrac{z^2}{2}, & \text{若 } 0\le z\le 1,\\ 1-\dfrac{(2-z)^2}{2}, & \text{若 } 1<z\le 2,\\ 1, & \text{若 } 2<z.\end{cases}$$

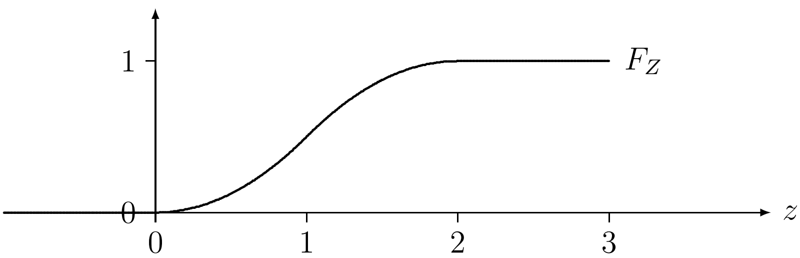


圖 3–8

爲了更了解 Z 之分配情形, 我們將 F_Z 予以微分而得 Z 之密度函數如下:

$$f_Z\colon \mathbb{R} \to \mathbb{R}_+ : f_Z(z) = \begin{cases} 0, & \text{若 } z < 0, \\ z, & \text{若 } 0 \le z \le 1, \\ 2 - z, & \text{若 } 1 < z \le 2, \\ 0, & \text{若 } 2 < z. \end{cases}$$

其圖形如下:

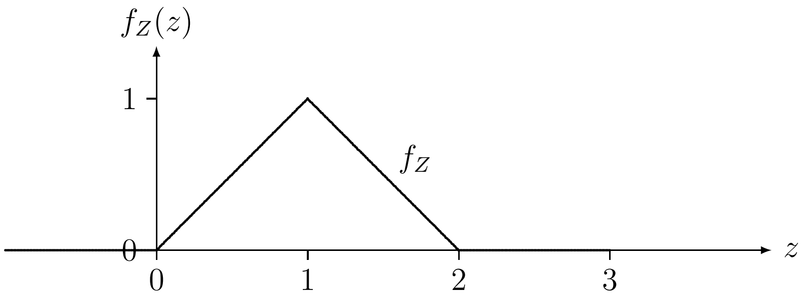


圖 3–9 □

利用上述結果, 我們可以迅速探討以下問題:

- 自區間 [0, 1] 中隨機抽取二數, 試問此二數之和大於 0.5 且小於 1.5 之機率爲若干? (習題 3.13).

定理 3.18 (離散型隨機向量之變換公式)

設 $\mathbf{X} = (X_1, \cdots, X_k)$ 爲 $(\Omega, \mathcal{F}, P)$ 上一離散型隨機向量, $f_{\mathbf{X}}$ 爲其密度函數, 集合 $A = \{\mathbf{x} \in \mathbb{R} \mid f_{\mathbf{X}}(\mathbf{x}) > 0\}$; 次設函數 $\mathbf{h}\colon \mathbb{R}^k \to \mathbb{R}^m$ 爲 Borel 可測, 隨機變數 $\mathbf{Y} = \mathbf{h} \circ \mathbf{X}$, 則

(1) $P\{\mathbf{Y} \in B\} = 1$, 內 $B = \mathbf{h}(A)$;

(2) $\mathbf{Y}$ 之密度函數爲

$$f_{\mathbf{Y}}\colon \mathbb{R}^m \to [0, 1] : f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} \displaystyle\sum_{\mathbf{x} \in \mathbf{h}^{-1}\{\mathbf{y}\}} f_{\mathbf{X}}(\mathbf{x}), & \text{若 } \mathbf{y} \in B, \\ 0, & \text{若 } \mathbf{y} \in \mathbb{R}^m \setminus B, \end{cases}$$

(3) 若 $\mathbf{h}$ 爲 A 映至 B 之一對射函數, 則當 $\mathbf{y} \in B$ 時, $f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{h}^{-1}(\mathbf{y}))$.

證 仿定理 2.25, 讀者自證之. □

例 2. 設 X_1 具 $B(n_1,\,p)$ 分配, 其密度函數爲 f_1, X_2 具 $B(n_2,\,p)$ 分配, 其密度函數爲 f_2. 次設 $\mathbf{X}=(X_1,\,X_2)$ 之密度函數爲

$$f_{\mathbf{X}}(x_1,\,x_2)=f_1(x_1)f_2(x_2);$$

(在第六章中, 我們將稱 X_1 與 X_2 爲獨立), 試證: 隨機變數 $Y_1=X_1+X_2$ 具 $B(n_1+n_2,\,p)$ 分配. 若設 $Y_2=X_2$, 並求: 給 $Y_1=y_1$ 後 Y_2 之條件分配.

$\boxed{\text{解}}$ 令 $\mathbf{Y}=(Y_1,\,Y_2)$,

1° 先求 $\mathbf{Y}$ 之密度函數. 由於 $\mathbf{X}=(X_1,\,X_2)$ 之密度函數爲

$$f_{\mathbf{X}}(x_1,\,x_2)=\begin{cases}\dbinom{n_1}{x_1}\dbinom{n_2}{x_2}p^{x_1+x_2}(1-p)^{n_1+n_2-(x_1+x_2)}, & \text{若 } (x_1,\,x_2)\in A,\\ \qquad 0, & \text{否則}.\end{cases}$$

內

$$\begin{aligned}A&=\Big\{(x_1,\,x_2)\,\Big|\,x_1\in\{0,\,1,\,\cdots,\,n_1\},\ x_2\in\{0,\,1,\,\cdots,\,n_2\}\Big\}\\ &=\{0,\,1,\,\cdots,\,n_1\}\times\{0,\,1,\,\cdots,\,n_2\}.\end{aligned}$$

令 $\mathbf{h}\colon\mathbb{R}^2\to\mathbb{R}^2:\mathbf{h}(x_1,\,x_2)=(x_1+x_2,\,x_2)$. 則

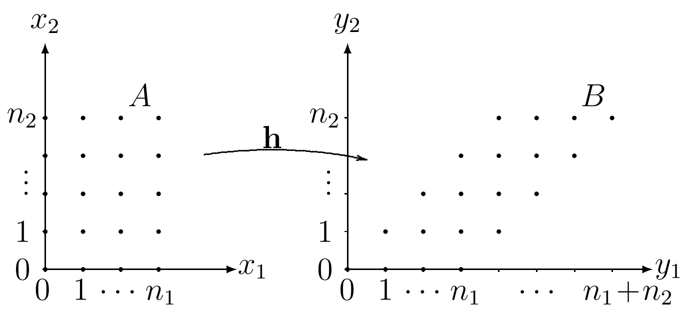


圖 3–10

(a) $B=\mathbf{h}(A)$.

$$\begin{aligned}&=\Big\{(y_1,\,y_2)\in\mathbb{Z}^2\,\Big|\,0\le y_1\le n_1+n_2,\ 0\le y_2\le n_2,\ 0\le y_1-y_2\le n_1\Big\}\\ &=\Big\{(y_1,\,y_2)\in\mathbb{Z}^2\,\Big|\,0\le y_1\le n_1+n_2,\max\{0,\,y_1-n_1\}\le y_2\le\min\{y_1,\,n_2\}\Big\}.\end{aligned}$$

(b) $\mathbf{h}$ 爲一嵌射, 此因

$$\begin{aligned}\mathbf{h}(x_1,\,x_2)=\mathbf{h}(x_1',\,x_2')&\Leftrightarrow(x_1+x_2,\,x_2)=(x_1'+x_2',\,x_2')\\ &\Leftrightarrow x_1+x_2=x_1'+x_2'\text{ 且 }x_2=x_2'\\ &\Leftrightarrow x_1=x_1'\text{ 且 }x_2=x_2'\\ &\Leftrightarrow(x_1,\,x_2)=(x_1',\,x_2').\end{aligned}$$

(c) 至於 $\mathbf{h}|_A$ 之反函數, 由於

$$
\begin{aligned}
(y_1,\, y_2) = \mathbf{h}(x_1,\, x_2) = (x_1 + x_2,\, x_2) \\
&\Leftrightarrow \begin{cases} y_1 = x_1 + x_2, \\ y_2 = x_2 \end{cases} \\
&\Leftrightarrow \begin{cases} x_1 = y_1 - x_2, \\ x_2 = y_2 \end{cases} \\
&\Leftrightarrow \begin{cases} x_1 = y_1 - y_2, \\ x_2 = y_2; \end{cases}
\end{aligned}
$$

知 $\mathbf{h}^{-1}\colon B \to A : \mathbf{h}^{-1}(y_1,\, y_2) = (y_1 - y_2,\, y_2)$.

由以上三點, 利用定理 3.18 知, $\mathbf{Y} = (Y_1,\, Y_2)$ 之 p.d.f. 爲

$$
f_{\mathbf{Y}}\colon \mathbb{R}^2 \to [0,\, 1] : f_{\mathbf{Y}}(y_1,\, y_2) = \begin{cases} *, & \text{若 } (y_1,\, y_2) \in B, \\ 0, & \text{若 } (y_1,\, y_2) \notin B, \end{cases}
$$

內 $* = f_{\mathbf{X}}(\mathbf{h}^{-1}(y_1,\, y_2)) = \dbinom{n_1}{y_1 - y_2}\dbinom{n_2}{y_2} p^{y_1}(1-p)^{n_1+n_2-y_1}$.

2° 次求 Y_1 之密度函數. $f_{Y_1}\colon \mathbb{R} \to [0,\, 1]$:

$$
f_{Y_1}(y_1) = \begin{cases} \displaystyle\sum_{y_2} f_{\mathbf{Y}}(y_1,\, y_2) = **, & \text{若 } y_1 \in \{0,\, 1,\, \cdots,\, n_1 + n_2\}, \\ 0, & \text{否則}, \end{cases}
$$

如果我們令 $u = \max\{0,\, y_1 - n_1\}$, $v = \min\{y_1,\, n_2\}$, 則上式中的

$$
\begin{aligned}
** &= \sum_{y_2 = u}^{v} \binom{n_1}{y_1 - y_2}\binom{n_2}{y_2} p^{y_1}(1-p)^{n_1+n_2-y_1} \\
&= \binom{n_1 + n_2}{y_1} p^{y_1}(1-p)^{n_1+n_2-y_1},
\end{aligned}
$$

即 Y_1 具 $B(n_1 + n_2,\, p)$ 分配.

3° 最後求“給 $Y_1 = y_1$ 後 Y_2 之條件分配.”

對於 $y_1 \in \{0,\, 1,\, \cdots,\, n_1 + n_2\}$,

- 若 y_2 不爲介於 $\max\{0,\, y_1 - n_1\}$ 及 $\min\{y_1,\, n_2\}$ 之整數, 顯然 $f(y_2|y_1) = 0$;
- 若 $\max\{0,\, y_1 - n_1\} \le y_2 \le \min\{y_1,\, n_2\}\}$,

$$
\begin{aligned}
f(y_2|y_1) = \frac{f_{\mathbf{Y}}(y_1,\, y_2)}{f_{Y_1}(y_1)} &= \frac{\dbinom{n_1}{y_1 - y_2}\dbinom{n_2}{y_2} p^{y_1}(1-p)^{n_1+n_2-y_1}}{\dbinom{n_1 + n_2}{y_1} p^{y_1}(1-p)^{n_1+n_2-y_1}} \\
&= \frac{\dbinom{n_1}{y_1 - y_2}\dbinom{n_2}{y_2}}{\dbinom{n_1 + n_2}{y_1}},
\end{aligned}
$$

亦即, 此條件分配爲一超幾何分配. □

定理 3.19 (連續型隨機向量之變換公式)

$\mathbf{X} = (X_1, \cdots, X_k)$ 爲 $(\Omega, \mathcal{F}, P)$ 上一連續型隨機向量, $f_{\mathbf{X}}$ 爲其密度函數, 集合 $A = \{\mathbf{x} \in \mathbb{R} \mid f_{\mathbf{X}}(\mathbf{x}) > 0\}$; 次設函數 $\mathbf{h}\colon \mathbb{R}^k \to \mathbb{R}^k$ 滿足

(1) $\mathbf{h}|_A\colon A \to B = \mathbf{h}(A)$ 爲一對射;

(2) $\mathbf{h}|_A$ 之反函數 $\mathbf{h}^{-1} = (g_1, \cdots, g_k)\colon B \to A$, 其內 $g_1, \cdots, g_k$ 皆爲連續可偏微分;

(3) $\mathbf{h}^{-1}$ 之 Jacobian

$$J_{\mathbf{h}^{-1}}(\mathbf{y}) = \det(D_i g_j(\mathbf{y}))_{\substack{1\le i\le k\\ 1\le j\le k}} \neq 0,\ \forall \mathbf{y} \in B.$$

則 $\mathbf{Y} = \mathbf{h} \circ \mathbf{X}$ 之密度函數爲

$$f_{\mathbf{Y}}\colon \mathbb{R}^k \to \mathbb{R}_+ : f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} f_{\mathbf{X}}(\mathbf{h}^{-1}(\mathbf{y})) \cdot |(J_{\mathbf{h}^{-1}})(\mathbf{y})|, & \text{若 } \mathbf{y} \in B, \\ 0, & \text{若 } \mathbf{y} \notin B. \end{cases}$$

證 仿定理 2.26, 讀者自證之. □

例 2. 設 $X_1 \sim U(\alpha, \beta)$, $X_2 \sim U(\alpha, \beta)$, (X_1, X_2) 之密度函數爲 X_1 及 X_2 密度函數之積 (在第六章中, 有此一條件時, 稱 X_1 與 X_2 爲獨立). 試求: $Y_1 = X_1 + X_2$ 之密度函數.

[**說明**] 提醒讀者注意: 在定理 3.19 之前提中, 規定 $\mathbf{h}$ 之定義域及對應域之維度 (dimension) 必須相同; 在此, 若令

$$h\colon \mathbb{R}^2 \to \mathbb{R} : h(x_1, x_2) = x_1 + x_2,$$

雖有 $Y_1 = X_1 + X_2 = h \circ (X_1, X_2)$, 但不能利用此一定理. 除非稍加修改, 以配合上述定理.

解 $\mathbf{X} = (X_1, X_2)$ 之密度函數爲

$$f_{\mathbf{X}}\colon \mathbb{R}^2 \to \mathbb{R}_+ : f_{\mathbf{X}}(x_1, x_2) = \begin{cases} \dfrac{1}{(\beta-\alpha)^2}, & \text{若 } (x_1, x_2) \in A, \\ 0, & \text{若 } (x_1, x_2) \notin A, \end{cases}$$

內 $A = [\alpha, \beta] \times [\alpha, \beta]$. 若令

$$\mathbf{h}\colon \mathbb{R}^2 \to \mathbb{R}^2 : \mathbf{h}(x_1, x_2) = (x_1 + x_2, x_2).$$

則

1° $\mathbf{h} \circ \mathbf{X} = (X_1 + X_2, X_2) = (Y_1, Y_2)$, (令 $Y_2 = X_2$);

2° $$\begin{aligned} B &= \mathbf{h}(A) \\ &= \{(x_1+x_2,\, x_2) \mid \alpha \le x_1 \le \beta,\ \alpha \le x_2 \le \beta\} \\ &= \{(x_1+x_2,\, x_2) \mid 2\alpha \le x_1+x_2 \le 2\beta,\ \alpha \le x_2 \le \beta,\ \alpha \le x_1 \le \beta\} \\ &\qquad\qquad\qquad (\text{令 } y_1 = x_1+x_2,\,, y_2 = x_2) \\ &= \{(y_1,\, y_2) \mid 2\alpha \le y_1 \le 2\beta,\ \alpha \le y_2 \le \beta,\ \alpha \le y_1-y_2 \le \beta\}. \end{aligned}$$

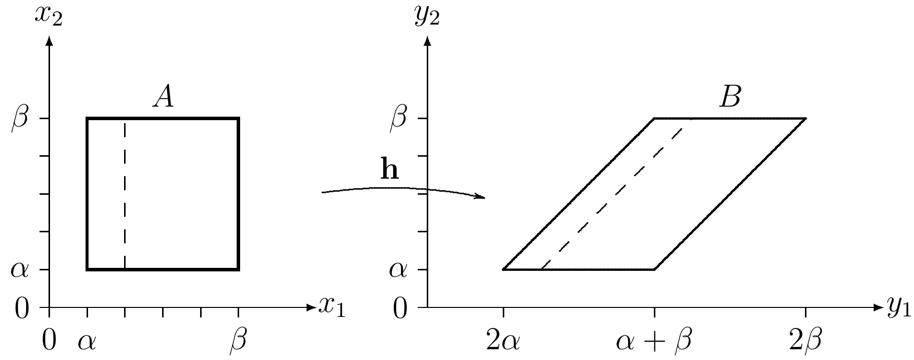


圖 3–11

3° $\mathbf{h}|_A\colon A \to B = \mathbf{h}(A)$ 顯然爲一對射, 因爲

$$\begin{aligned} \mathbf{h}(x_1,\, x_2) = \mathbf{h}(x_1', x_2') & \\ &\Leftrightarrow (x_1+x_2,\, x_2) = (x_1'+x_2',\, x_2') \\ &\Leftrightarrow \begin{cases} x_1+x_2 = x_1'+x_2', \\ x_2 = x_2' \end{cases} \\ &\Leftrightarrow x_1 = x_1' \text{ 且 } x_2 = x_2' \\ &\Leftrightarrow (x_1,\, x_2) = (x_1',\, x_2'). \end{aligned}$$

4° 至於 $\mathbf{h}|_A$ 之反函數, 由於

$$\begin{aligned} (y_1,\, y_2) = \mathbf{h}(x_1,\, x_2) = (x_1+x_2,\, x_2) & \\ &\Leftrightarrow \begin{cases} y_1 = x_1+x_2, \\ y_2 = x_2 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = y_1-x_2, \\ x_2 = y_2 \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = y_1-y_2, \\ x_2 = y_2; \end{cases} \end{aligned}$$

是以 $\mathbf{h}^{-1}\colon B \to A : \mathbf{h}^{-1}(y_1,\, y_2) = (y_1-y_2,\, y_2)$; 故其 Jacobian 爲

$$J_{\mathbf{h}^{-1}}(y_1,\, y_2) = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1 \ne 0,\ \forall (y_1,\, y_2) \in B,$$

由定理 3.19 知, $\mathbf{Y} = (Y_1, Y_2)$ 之 p.d.f. 爲

$$f_{\mathbf{Y}}\colon \mathbb{R}^2 \to \mathbb{R}_+ : f_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{1}{(\beta-\alpha)^2}, & \text{若 } (y_1, y_2) \in B, \\ 0, & \text{若 } (y_1, y_2) \notin B, \end{cases}$$

最後, 求 $Y_1 = X_1 + X_2$, 之邊際密度函數, 已知

$$f_{Y_1}\colon \mathbb{R} \to \mathbb{R}_+ : f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{\mathbf{Y}}(y_1, y_2)dy_2.$$

(i) 若 $y_1 < 2\alpha$, 則 $f_{\mathbf{Y}}(y_1, y_2) = 0$, 是以 $f_{Y_1}(y_1) = 0$.

(ii) 若 $2\alpha \le y_1 < \alpha+\beta$, 則

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{\alpha}^{y_1-\alpha} \frac{1}{(\beta-\alpha)^2} dy_2, \text{(參閱圖 3–10)} \\ &= \frac{y_1 - 2\alpha}{(\beta-\alpha)^2}; \end{aligned}$$

(iii) 若 $\alpha+\beta \le y_1 \le 2\beta$, 則

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{y_1-\beta}^{\beta} \frac{1}{(\beta-\alpha)^2} dy_2, \text{(參閱圖 3–10)} \\ &= \frac{2\beta - y_1}{(\beta-\alpha)^2}; \end{aligned}$$

(iv) 若 $2\beta \le y_1$, 則 $f_{\mathbf{Y}}(y_1, y_2) = 0$, 是以 $f_{Y_1}(y_1) = 0$.

故得

$$f_{Y_1}\colon \mathbb{R} \to \mathbb{R}_+ : f_{Y_1}(y_1) = \begin{cases} \dfrac{y_1 - 2\alpha}{(\beta-\alpha)^2}, & \text{若 } 2\alpha \le y_1 < \alpha+\beta, \\ \dfrac{2\beta - y_1}{(\beta-\alpha)^2}, & \text{若 } \alpha+\beta \le y_1 \le 2\beta, \\ 0, & \text{若 } y_1 \notin [\alpha, \beta]. \end{cases}$$

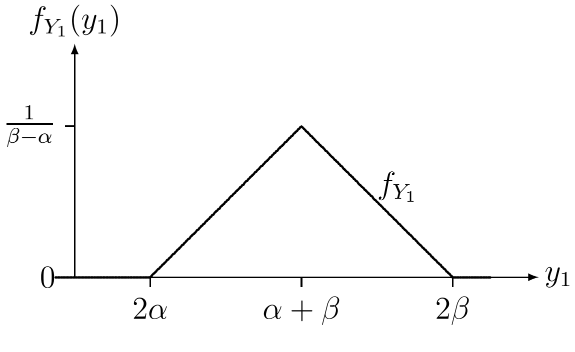


圖 3–12 □

仿照定理 2.27, 我們亦可研究不爲對射之連續型隨機向量之變換公式, 但此一公式相當龐大, 對於初學者並不適合, 因此不予列入.

§ 3.7 變換與捲積

在數學分析中, 對於二函數 $f, g\colon \mathbb{R} \to \mathbb{R}$, 規定其捲積爲

$$(f * g)\colon \mathbb{R} \to \mathbb{R} : (f * g)(t) = \int_{\mathbb{R}} f(t-y)g(y)\,dy,$$

(如果右端之瑕積分收斂), 此一觀念在機率論上常被用來解決二隨機變數之和的分配.

定義 3.20

設 f, g 分別爲連續型隨機變數 X 與 Y 之密度函數, 則函數

$$(f * g)\colon \mathbb{R} \to \mathbb{R} : (f * g)(t) = \int_{\mathbb{R}} f(t-y)g(y)\,dy$$

稱爲 f 與 g 之**捲積**(convolution).

定理 3.21

設 f, g 分別爲連續型隨機變數 X 與 Y 之密度函數, 若 $f_{(X,Y)}(x,\,y) = f(x)g(y)$, 則 f 與 g 之捲積乃 $X+Y$ 之密度函數.

證 設 $A(t) = \{(x,y) \mid x+y \le t\}$ (圖 3–12 網點部分), 則 (X,Y) 之分配函數爲

$$\begin{aligned}
F_{X+Y}(t) &= P\{X+Y \le t\} \\
&= \iint_{A(t)} f_{(X,Y)}(x,\,y)\,dx\,dy, \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{t-y} f_{(X,Y)}(x,\,y)\,dx\,dy \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{t-y} f(x)g(y)\,dx\,dy,
\end{aligned}$$

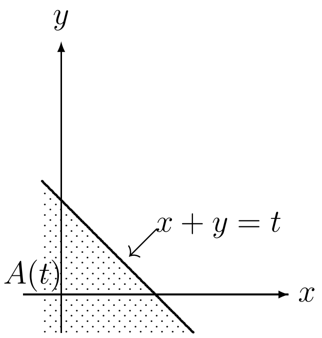


圖 3–12

利用定理 2.20,

$$\begin{aligned} f_{X+Y}(t) = \frac{\partial}{\partial t} F_{X+Y}(t) &= \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} \int_{-\infty}^{t-y} f(x)g(y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty} \frac{\partial}{\partial t} \int_{-\infty}^{t-y} f(x)g(y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty} f(t-y)g(y)\,dy = (f*g)(t), \end{aligned}$$

$\forall t \in \mathbb{R}$ a.e. $\square$

例 1. 設 X, Y 皆具 $G(1,\,1)$, 且 $f_{(X,Y)}(x,\,y) = f_X(x)f_Y(y)$, 試利用捲積方法以求 $X+Y$ 之密度函數.

$\boxed{\text{解}}$ 由原設知 $f_X(x) = f_Y(x) = \begin{cases} e^{-x}, & \text{若 } x > 0, \\ 0, & \text{若 } x \le 0, \end{cases}$ 由於

$$f_X(t-y)f_Y(y) > 0 \Leftrightarrow \Big(t-y>0 \ \wedge\ 0<y\Big) \Leftrightarrow 0<y<t.$$

是以 $\forall t > 0$,

$$\begin{aligned} f_{X+Y}(t) &= \int_{\mathbb{R}} f_X(t-y)f_Y(y)\,dy = \int_0^t f_X(t-y)f_Y(y)\,dy \\ &= \int_0^t e^{-t+y}e^{-y}\,dy = \int_0^t e^{-t}\,dy = te^{-t}. \end{aligned}$$

故

$$f_{X+Y}\colon \mathbb{R} \to \mathbb{R}_+ : f_{X+Y} = \begin{cases} te^{-t}, & \text{若 } t > 0, \\ 0, & \text{若 } t \le 0. \end{cases}$$

換言之, $X+Y \sim G(2,1)$. $\square$

就機率論本身而言, 隨機變數及隨機向量之變換之分配的研究並非十分重要; 然而由於機率論是統計學之基礎, 基於需要, 後者發展出很多隨機變數及隨機向量, 對於它們之分配的了解, 則必須利用各種變換公式, 我們還有另一種方法, 由於需要新的工具, 我們留待第六章再研究.

第三章 習 題

隨機向量及其分配

3-1 設 X 與 Y 均爲隨機變數. 試問: 下列函數何者爲隨機變數?

(1) $\sin\circ(X+Y)$; (2) e^{X+Y};

(3) $[XY]$, 內$[\,\cdot\,]$ 爲最大整數函數; (4) $\max\{X, Y\}$.

3-2 設 $\mathbf{X}$ 與 $\mathbf{Y}$ 均爲 k 維隨機向量. 試問: 下列函數何者爲隨機向量?

(1) $\mathbf{X}+\mathbf{Y}$;

(2) $\mathbf{X}\cdot\mathbf{Y}$: (此處指 $(\mathbf{X}\cdot\mathbf{Y})(\omega)=\mathbf{X}(\omega)$ 與 $\mathbf{Y}(\omega)$ 之內積).

3-3 試舉一例: Y 爲一隨機變數, 且 $Y\neq 0$, 但 $1/Y$ 不爲一隨機變數.

3-4 設連續型隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2\to\mathbb{R}_+ : f(x,y)=\begin{cases} cxy^2, & \text{若 } 0<x<y<1,\\ 0, & \text{否則}\end{cases}$$

(1) 試求常數 c ;

(2) 試求 $P\{Y\leq 1-X\}=$?

3-5 設 $\mathbf{X}=(X_1, \cdots, X_k)$ 具多項分配, 其密度函數如本章第二節, 隨機 變數 X 具 $B(n, p_1)$ 分配.

(1) 寫出 $k=2$ 時, $\mathbf{X}$ 之密度函數;

(2) 試問 $k=2$ 時, $\mathbf{X}$ 與 X 之密度函數有何異同?

(3) 當 $x_1\in\{0, 1, \cdots, n\}$ 時, $P\{X_1=x_1,\ X_2=n-x_1\}=P\{X=x_1\}$. 試解釋其所代表之意義爲何?

3-6 設 F 爲隨機向量 (X, Y) 之分配函數, 試證:

$$\lim_{x\to-\infty} F(x, y)=0, \quad \lim_{\substack{x\to+\infty\\ y\to+\infty}} F(x, y)=1.$$

3-7 設隨機變數 $X=Y$ 具 $U(0, 1)$ 分配, F 爲 (X, Y) 之分配函數.

(1) 證明: $\forall x, y\in\mathbb{R},\ F(x, y)=F_X(\min\{x, y\})$;

(2) 繪出集合 $\{(x, y)\mid F(x, y)=1/2\}$ 之圖形;

(3) 繪出 $F(x, y)=\frac{n}{5}$, 其中 $n=1, 2, 3, 4$, 各 等高線, 以及 $F(x, y)=0$, $F(x, y)=1$ 等二區域以說明 F 之圖形.

3-8 設 F 爲隨機向量 (X, Y) 之分配函數, 試證: $\forall x,\, y \in \mathbb{R}$,

$$F_X(x) + F_Y(y) - 1 \leq F(x,\, y) \leq \sqrt{F_X(x) \cdot F_Y(y)}.$$

3-9 設 F 爲隨機向量 (X, Y) 之分配函數, 且 $F(x,\, y) = G(x)H(y)$: 內 G 與 H 爲二實值之實變函數. 試證:

(1) $\lim_{x\to+\infty} G(x)$ 及 $\lim_{y\to+\infty} H(y)$ 均 存在;

(2) $\lim_{x\to+\infty} G(x) = 1/\lim_{y\to+\infty} H(y)$.

邊際分配及條件分配

3-10 一盒中置十球, 1 號至 10 號, 隨機抽取其一. 設 X 表第一次抽得之球 號, 其次將 1 號至 X 號球置於另一盒中, 隨機抽取其一, 設 Y 表第 二次得之球號.

(1) 寫出隨機向量 (X, Y) 之密度函數;

(2) 寫出 X 之密度函數;

(3) 寫出 Y 之密度函數;

(4) 試求: 給 $Y = 3$後, X 之條件密度函數.

3-11 設隨機變數 X 與 Y 之聯合密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,\, y) = \begin{cases} ce^{-x-y}, & \text{若 } 0 \leq x \leq y, \\ 0, & \text{否則.} \end{cases}$$

(1) 試求常數 c ;

(2) 試求隨機向量 (X, Y) 之分配函數;

(3) 試由 (2) 小題之分配函數, 利用定理 3.12 以求 (X, Y) 之密度函數. 此一函數是否等於原設之函數 f ?

(4) 試求 X 與 Y 之邊際密度函數;

(5) 設 $y > 0$, 試求 $Y = y$ 後, X 之條件密度函數.

3-12 設隨機變數 X 與 Y 之聯合密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,\, y) = \begin{cases} k(y^2 - x^2)e^{-y}, & \text{若 } -y \leq x \leq y,\ y > 0, \\ 0, & \text{否則.} \end{cases}$$

(1) 試求常數 k ;

(2) 設 $y > 0$, 試求 $Y = y$ 後, X 之條件密度函數.

隨機向量之變換

3-13 自區間 $[0, 1]$ 中隨機抽取二數, 試問此二數之和大於 0.5 且 小於 1.5 之 機率爲若干?

3-14 設離散型隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to [0,1] : f(x,y) = \begin{cases} (x+y^2)/42, & \text{若 } x \in \{1, 4\},\ y \in \{-1, 0, 1, 3\}, \\ 0, & \text{否則.} \end{cases}$$

試求下列各隨機變數之密度函數:

(1) $T = \max\{X, Y\}$;

(2) $U = XY$;

(3) $V = |Y| - X$.

3-15 設 $X \sim N(0, 1), Y \sim \chi^2_r, (X, Y)$ 之密度函數爲 X 及 Y 密度函數之積 (即 X 與 Y 爲獨立). 我們稱隨機變數

$$T = \frac{X}{\sqrt{Y/r}}$$

爲具自由度爲 r 之 t 分配.

(1) 試證: T 之密度函數爲

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \cdot \frac{1}{(1+(t^2/r))^{(r+1)/2}}.$$

(2) 試證: 當 $r \to +\infty$ 時, 上述密度函數 f 收斂於標準常態分配之機率密度函數, 即

$$\forall t \in \mathbb{R},\ \lim_{r\to+\infty} f(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right).$$

3-16 設 $X \sim \chi^2_{r_1}, Y \sim \chi^2_{r_2}, (X, Y)$ 之密度函數爲 X 及 Y 密度函數之積 (即 X 與 Y 爲獨立). 我們稱隨機變數 $F = \dfrac{X/r_1}{Y/r_2}$ 爲具自由度爲 r_1 及 r_2 之 F 分配.

(1) 試證: 隨機變數 $1/F$ 具自由度爲 r_2 及 r_1 之 F 分配;

(2) 若隨機變數 T 具 t 分配, 試證 T^2 具 F 分配; 此 時其自由度爲何?

(3) 若隨機變數 F 具 F 分配, 試證: 隨機變數 $X = \dfrac{1}{1+(r_1/r_2)F}$ 具 Beta 分配.

3-17 設連續型隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x,y) = \begin{cases} 1, & \text{若 } 0 < x < 1,\ 0 < y < 1, \\ 0, & \text{否則.} \end{cases}$$

(1) 試求隨機變數 $Z = XY$ 之密度函數:

(2) 試求隨機變數 $U = \max\{X, Y\}$ 之密度函數.

Chapter 4

期望值

儘管隨機變數之密度函數或分配函數, 已經能提供吾人有關其分布一個相當完整的形象, 但若要比較二個以隨機變數之分布, 則需要有更進一步的概念與方法. 最先被想到的, 自然是一分布之 “中心”, 通常是指 “期望值”, (期望值不存在時則以中位數代之), 其次則是 “變異數”, 此乃用以討論該分布之分散程度, 本章中以此二概念爲主軸, 探討各項有關性質及基本定理, 並進而爲更重要之機率理論做準備.

§ 4.1 隨機變數之期望值

期望值的最初觀念可能與賭博有關: 假設一賭博, 賭者可能 (自莊家) 得到 x_1 元, x_2 元, $\cdots$, x_r 元, 假定得到 x_1 元之機率爲 $f(x_1), \cdots$, 得到 x_r 元之機率爲 $f(x_r), (f(x_1)+\cdots+f(x_r)=1)$, 此一賭博總共進行 N 次, 其中 x_1 元出現 N_1 次, $\cdots$, x_r 元出現 N_r 次, $(N_1+\cdots+N_r=N)$, 則 N 次後賭者全部所得爲

$$S_N = x_1N_1 + \cdots + x_rN_r;$$

平均每次所得爲

$$\frac{S_N}{N} = x_1 \cdot \frac{N_1}{N} + \cdots + x_r \cdot \frac{N_r}{N};$$

令 $N \to +\infty$, 則我們可“期望”每次之所得平均爲

$$\lim_{N\to+\infty} \frac{S_N}{N} = x_1 f(x_1) + \cdots + x_r f(x_r) = \sum_{j=1}^{r} x_j f(x_j).$$

我們將利用隨機變數建構式之定義 (參考定理 2.6), 以界定期望值.

定義 4.1

設 X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數. 我們將界定 X 之 **期望值或數學期望值**(expected value or mathematical expectation) EX (亦寫作 $E(X)$ 或 $E[X]$) 如下:

1° 若 X 爲一簡單隨機變數, (參見定義 2.5), 即 $X = \sum_{j=1}^{n} x_j I_{A_j}$, 則規定

$$EX = \sum_{j=1}^{n} x_j P(A_j);$$

2° 若 X 爲一非負隨機變數, 則由定理 2.6 知, 存在 一非負簡單隨機變數序列 $\{X_n\}_n$ 使得 $X_n \uparrow X$; 我們規定

$$EX = \lim_{n \to +\infty} EX_n \ (\in [0, +\infty]);$$

3° 若 X 爲一般隨機變數, 則 $X = X^+ - X^-$. 如果 $E[X^+]$ 或 $E[X^-]$ 爲有限, 則稱 EX 存在且規定

$$EX = E[X^+] - E[X^-].$$

如果 $E[X^+]$ 及 $E[X^-]$ 均爲有限, 則稱 X 爲**可積**(integrable) 或 X 具**有限期望值**(finite expectation). 如果 $E[X^+] = E[X^-] = +\infty$, 則稱 X 之期望值不存在.

上述定義中之 2°, 與所選之序列 $\{X_n\}_n$ 無關, 我們將在附錄五中予以證明. 爲不使初學讀者感到過分深奧, 以下有關期望值之討論, 除非特別聲明, 均限制在可積之情形.

定理 4.2

設 X 爲一離散型隨機變數, 即存在 $\mathbb{R}$ 之一有限或可數子集合 A 使得 $P\{X \in A\} = 1$. 則 $EX = \sum_{x \in A} x f(x)$, 內 f 爲 X 之密度函數.

證 1° 若 A 爲有限, 即 X 爲一簡單隨機變數, 令 $A = \{x_1, \cdots, x_n\}$, 則 X 可寫爲 $X = \sum_{j=1}^{n} x_j I_{\{X = x_j\}}$, 由上述定義知,

$$EX = \sum_{j=1}^{n} x_j P\{X = x_j\} = \sum_{j=1}^{n} x_j f(x_j) = \sum_{x \in A} x f(x).$$

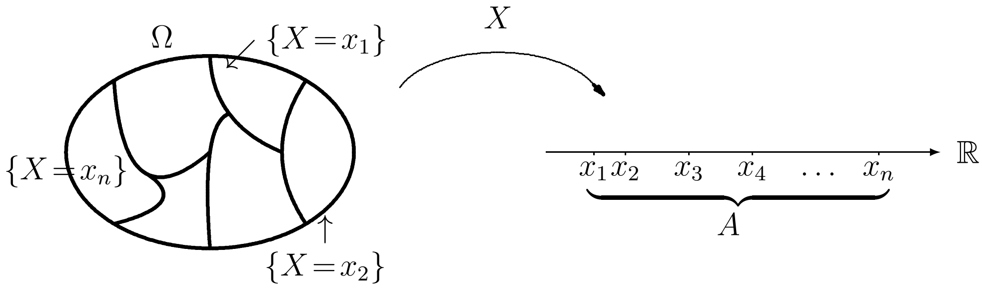


圖 4–1

2° 若集合 A 爲無限可數且 X 爲非負. 令 $A=\{x_1,\cdots,x_n,\cdots\}\subset\mathbb{R}_+$ 且 $X_n=\sum\limits_{j=1}^{n}x_jI_{\{X=x_j\}}$, 則 $0\le X_n\uparrow X$, 是以

$$\begin{aligned}EX&=\lim_{n\to+\infty}EX_n\\&=\lim_{n\to+\infty}\sum_{j=1}^{n}x_jP\{X=x_j\}\\&=\lim_{n\to+\infty}\sum_{j=1}^{n}x_jf(x_j)\\&=\sum_{j=1}^{+\infty}x_jf(x_j)=\sum_{x\in A}xf(x).\end{aligned}$$

3° 若集合 A 爲可數且 X 爲一般離散型. 令 $X=X^+-X^-$, $A_1=A\cap[0,+\infty)$, $A_2=A\cap(-\infty,0)$, 則

$$X^+=\max\{X,0\}=\sum_{x_j\in A_1}x_jI_{\{X=x_j\}},$$
$$X^-=\max\{-X,0\}=\sum_{x_j\in A_2}(-x_j)I_{\{X=x_j\}};$$

是以由 1° 及 2° 知,

$$E[X^+]=\sum_{x\in A_1}xf(x),\quad E[X^-]=\sum_{x\in A_2}(-x)f(x).$$

故

$$\begin{aligned}EX&=E[X^+]-E[X^-]\\&=\sum_{x\in A_1}xf(x)-\sum_{x\in A_2}(-x)f(x)=\sum_{x\in A}xf(x).\end{aligned}$$ □

在此特別提醒讀者, 如果 EX 存在且不爲可積, 本定理依然爲眞; 但一般之離散型隨機變數, 有可能 $E[X^+]=E[X^-]=+\infty$, 這種情況下, 本定理並不成立.

定理 4.3

設 $\mathbf{X}$ 爲 $(\Omega, \mathcal{F}, P)$ 上一 k 維隨機向量, $F_{\mathbf{X}}$ 爲 $\mathbf{X}$ 之分配函數, 而函數 $g\colon \mathbb{R}^k \to \mathbb{R}$ 爲 Borel 可測.

$$\begin{tikzcd} \Omega \arrow[r, "\mathbf{X}"] \arrow[dr, "g\circ\mathbf{X}"'] & \mathbb{R}^k \arrow[d, "g"] \\ & \mathbb{R} \end{tikzcd}$$

(1) 若隨機變數 $g\circ\mathbf{X}$ 爲可積, 則 $E(g\circ\mathbf{X}) = \displaystyle\int_{\mathbb{R}^k} g(\mathbf{x})dF_{\mathbf{X}}(\mathbf{x})$;

(2) 若 $\mathbf{X}$ 爲離散型, $f_{\mathbf{X}}$ 爲其密度函數, 則 $E(g\circ\mathbf{X}) = \displaystyle\sum_{\mathbf{x}\in A} g(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})$, 其中 $A = \{\mathbf{x}\in\mathbb{R}^k \mid f_{\mathbf{X}}(\mathbf{x}) > 0\}$;

(3) 若 $\mathbf{X}$ 爲絕對連續型, $f_{\mathbf{X}}$ 爲其密度函數, 則 $E(g\circ\mathbf{X}) = \displaystyle\int_{\mathbb{R}^k} g(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$.

證 (1) 須利用抽象積分之變數代換觀念才能證明, 超出本書之程度, 有興趣之讀者請參閱 Loève [11] p.168.

(2) 因 $\mathbf{X}$ 爲離散型, 存在一有限或可數集合 $A = \{\mathbf{x}_j \mid j\in J\} \subset \mathbb{R}^k$ 使得 $P\{\mathbf{X}\in A\} = 1$. 顯然 $g\circ\mathbf{X}$ 爲離散型, 此因 $\{\mathbf{X}\in A\} \subset \{g\circ\mathbf{X}\in g(A)\}$, 是以有 $P\{g\circ\mathbf{X}\in g(A)\} = 1$. 此外, 由於

$$g\circ\mathbf{X} = \sum_{j\in J} g(\mathbf{x}_j)I_{\{\mathbf{X}=\mathbf{x}_j\}},$$

由定理 4.2 知,

$$\begin{aligned} E(g\circ\mathbf{X}) &= \sum_{j\in J} g(\mathbf{x}_j)P\{\mathbf{X}=\mathbf{x}_j\} \\ &= \sum_{j\in J} g(\mathbf{x}_j)f_{\mathbf{X}}(\mathbf{x}_j) = \sum_{\mathbf{x}\in A} g(\mathbf{x})f_{\mathbf{X}}(\mathbf{x}). \end{aligned}$$

(3) 當 $k=1$ 時,

$$\begin{aligned} E(g\circ X) &= \int_{\mathbb{R}} g(x)dF_X(x) \\ &= \int_{\mathbb{R}} g(x)f_X(x)\,dx, \quad (\text{因 } F_X' = f_X \text{ a.e.}) \end{aligned}$$

當 $k\geq 1$ 時, 涉及多維 Lebesgue-Stieltjes 積分之變換, 有興趣之讀者 請參閱實變函數論書籍以證明之. □

定義 4.4

設 X 爲一可積之隨機變數.

(1) 若 $r > 0$, 則 $E|X|^r$ 稱爲 X 之 r **階絕對動差** (r-th absolute moment), 而 $E|X-EX|^r$ 稱爲 X 之 r **階絕對中央動差** (r-th absolute central moment).

(2) 若 $n \in \mathbb{N}$, 則 $E[X^n]$ 稱爲 X 之 n **階動差** (n-th moment), 而 $E[(X-EX)^n]$ 稱爲 X 之 n **階中央動差** (n-th central moment).

(3) X 之二階中央動差 $E[(X-EX)^2]$ 又稱爲 X 之**變異數** (variance), 並記爲 $\mathrm{Var}\ X$ 或 σ_X^2. 變異數之平方根 σ_X 又稱爲 X 之**標準差** (standard deviation).

註: (1) 若 X 爲離散型. 設 $A = \{x \in \mathbb{R} \mid f(x) > 0\}$, 則

$$\begin{aligned}
E|X|^r &= \sum_{x\in A} |x|^r f(x), \\
E|X-EX|^r &= \sum_{x\in A} |x-EX|^r f(x), \\
E[X^n] &= \sum_{x\in A} x^n f(x), \\
E[(X-EX)^n] &= \sum_{x\in A} (x-EX)^n f(x), \\
\mathrm{Var}\ X &= \sum_{x\in A} (x-EX)^2 f(x).
\end{aligned}$$

(2) 若 X 爲絕對連續型, 則

$$\begin{aligned}
E|X|^r &= \int_{\mathbb{R}} |x|^r f(x)\,dx, \\
E|X-EX|^r &= \int_{\mathbb{R}} |x-EX|^r f(x)\,dx, \\
E[X^n] &= \int_{\mathbb{R}} x^n f(x)\,dx, \\
E[(X-EX)^n] &= \int_{\mathbb{R}} (x-EX)^n f(x)\,dx, \\
\mathrm{Var}\ X &= \int_{\mathbb{R}} (x-EX)^2 f(x)\,dx.
\end{aligned}$$

§ 4.2 期望值之性質及定理

定理 4.5 (線性性質)

設 X, Y 均爲可積, $c \in \mathbb{R}$ (常數), 則
(1) $X = c \;\Rightarrow\; E[X] = c$;
(2) $E[cX] = cE[X]$;
(3) $X + Y$ 亦爲可積且 $E[X+Y] = E[X] + E[Y]$.

$\boxed{\text{證}}$ (1) 常數函數乃爲一簡單隨機變數, 故

$$E[X] = cP\{X = c\} = c \cdot 1 = c.$$

(2) 我們將分三步驟證明之:

1° 若 $X = \sum_{j=1}^{n} x_j I_{A_j}$ 爲一簡單隨機變數, 則 $cX = \sum_{j=1}^{n} cx_j I_{A_j}$ 亦爲一簡單隨機變數, 故

$$E[cX] = \sum_{j=1}^{n} cx_j P(A_j) = c\sum_{j=1}^{n} x_j P(A_j) = cE[X].$$

2° 若 $c \geq 0$, X 爲一非負隨機變數, 則存在一非負簡單隨機變數 序列 $\{X_n\}_n$ 使得 $X_n \uparrow X$; 此時, $cX_n \uparrow cX$, 是以

$$E[cX] = \lim_{n\to+\infty} E[cX_n] = \lim_{n\to+\infty} cE[X_n] = cE[X].$$

3° 若 X 爲一般隨機變數, 則 $X = X^+ - X^-$, 當 $c \geq 0$ 時, 則 $cX = (cX)^+ - (cX)^-$, 是以

$$\begin{aligned} E[cX] &= E[(cX)^+] - E[(cX)^-] \\ &= cE[X^+] - cE[X^-], \quad (\text{由 } 2°) \\ &= cE[X]. \end{aligned}$$

當 $c < 0$ 時, 先證: $E[-X] = -E[X]$, 此因

$$(-X)^+ = \max\{-X, 0\} = X^-,\; (-X)^- = \max\{X, 0\} = X^+,$$

是以

$$E[-X] = E[(-X)^+] - E[(-X)^-] = E[X^-] - E[X^+] = -E[X].$$

故得 $E[cX] = E[(-1)(-c)X] = (-1)(-c)E[X] = cE[X]$.

(3) 我們亦分三步驟證明之:

1° 若 X, Y 均爲簡單隨機變數, 設

$$X=\sum_{j=1}^{n}x_jI_{A_j},\ Y=\sum_{k=1}^{m}y_kI_{B_k},$$

則 $X+Y=\sum_{j=1}^{n}\sum_{k=1}^{m}(x_j+y_k)I_{A_j\cap B_k}$, 是以

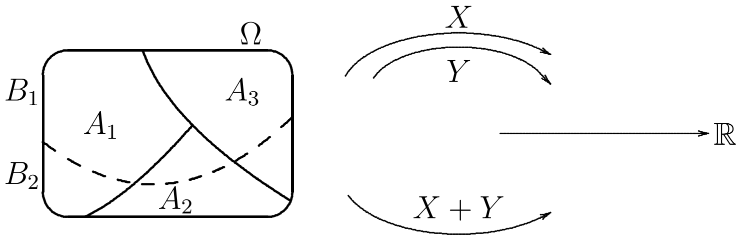


圖 4–1

$$\begin{aligned}E[X+Y]&=\sum_{j=1}^{n}\sum_{k=1}^{m}(x_j+y_k)P(A_j\cap B_k)\\&=\sum_{j=1}^{n}\sum_{k=1}^{m}x_jP(A_j\cap B_k)+\sum_{j=1}^{n}\sum_{k=1}^{m}y_kP(A_j\cap B_k)\\&=\sum_{j=1}^{n}x_jP(A_j)+\sum_{k=1}^{m}y_kP(B_k)\\&=E[X]+E[Y].\end{aligned}$$

2° 若 X, Y 均爲非負隨機變數, 則存在非負簡單隨機變數序列 $\{X_n\}_n$, $\{Y_n\}_n$ 使得 $X_n\uparrow X$; $Y_n\uparrow Y$; 此時, $\{X_n+Y_n\}_n$ 爲非負簡單隨機變數且 $X_n+Y_n\uparrow X+Y$, 是以

$$\begin{aligned}E[X+Y]&=\lim_{n\to+\infty}E[X_n+Y_n]\\&=\lim_{n\to+\infty}(E[X_n]+E[Y_n])=E[X]+E[Y].\end{aligned}$$

3° 若 X, Y 爲一般隨機變數, 則

$$\begin{aligned}E[X+Y]&=E[(X^+-X^-)+(Y^+-Y^-)]\\&=E[(X^++Y^+)-(X^-+Y^-)]\\&=E[X^++Y^+]-E[X^-+Y^-],\ ^{[\star]}\\&=E[X^+]+E[Y^+]-E[X^-]-E[Y^-]\\&=E[X]+E[Y].\end{aligned}$$

[⋆] 首先, 設 Z_1, Z_2 均爲非負可積隨機變數, 且 $Z=Z_1-Z_2$, 則因 $Z^+-Z^-=Z=Z_1-Z_2$, 移項後, 再利用 2° 可得

$$E[Z^+]+E[Z_2]=E[Z_1]+E[Z^-],$$

亦即 $E[Z_1]-E[Z_2]=E[Z^+]-E[Z^-]=E[Z]$, 最後以 X^++Y^+ 代上述之 Z_1, 以 X^-+Y^- 代上述之 Z_2, 即可. □

定理 4.6

(1) 若 $X \geq 0$ a.s., 則 $E[X] \geq 0$.

(2) 若 $X \geq Y$ a.s., 則 $E[X] \geq E[Y]$.

證 (1) 1° 若 $X = \sum_{j=1}^{n} x_j I_{A_j}$ 爲一簡單隨機變數, 令

$$S = X(\Omega) \cap [0, +\infty), T = X(\Omega) \cap (-\infty, 0),$$

由原設知 $P\{X \in T\} = 0$, 是以

$$E[X] = \sum_{x_j \in S} x_j P(A_j) + \sum_{x_j \in T} x_j P(A_j) = \sum_{x_j \in S} x_j P(A_j) \geq 0.$$

2° 若 $X \geq 0$ a.s., 則存在一殆必非負簡單隨機變 數序列 $\{X_n\}_n$ 使得 $X_n \uparrow X$; 由 1° 知, $\forall n, E[X_n] \geq 0$, 是以

$$E[X] = \lim_{n \to +\infty} E[X_n] \geq 0.$$

(2) $X \geq Y$ a.s. $\Rightarrow$ $X - Y \geq 0$ a.s. $\Rightarrow$ $0 \leq E[X - Y] = E[X] - E[Y]$.
(若 X 或 Y 不爲可積, 本性質仍爲眞, 但證明較爲麻煩.) □

定理 4.7

(1) 若 $X = 0$ a.s., 則 $E[X] = 0$.

(2) 若 $X = Y$ a.s., 則 $E[X] = E[Y]$.

證 (1) 利用前一定理,

$$\begin{aligned} X = 0 \text{ a.s.} \ &\Rightarrow \ (X \geq 0 \text{ a.s.} \ \wedge X \leq 0 \text{ a.s.}) \\ &\Rightarrow \ (EX \geq 0 \wedge EX \leq 0) \ \Rightarrow \ EX = 0. \end{aligned}$$

(2) $X = Y$ a.s. $\Rightarrow$ $X - Y = 0$ a.s. $\Rightarrow$ $0 = E[X - Y] = E[X] - E[Y]$. □

定理 4.8

X 爲可積之充要條件爲 $E|X| < +\infty$. 此時 $|EX| \leq E|X|$.

證 1° $X = X^+ - X^-, |X| = X^+ + X^-$, 是以

$$\begin{aligned} X \text{ 爲可積 } &\Leftrightarrow E[X^+] \text{ 及 } E[X^-] \text{ 皆爲有限} \\ &\Leftrightarrow E|X| = E[X^+] + E[X^-] < +\infty. \end{aligned}$$

2° 因爲 $|X| \geq \begin{cases} X, \\ -X, \end{cases}$ 故 $E|X| \geq \begin{cases} EX, \\ -EX, \end{cases}$ 是以 $E|X| \geq |EX|$. □

定理 4.9

(1) 設 $0 < t < r$, 若 $E|X|^r < +\infty$, 則 $E|X|^t < +\infty$.

(2) 設 $m, n \in \mathbb{N},\ m < n$, 若 X^n 爲可積, 則 X^m 亦爲可積.

證 (1) 因爲 $|X|^t \leq 1 + |X|^r$, 故

$$E[|X|^t] \leq 1 + E[|X|^r] < +\infty.$$

(2) 由於

$$\begin{aligned} X^n \text{ 爲可積 } &\Rightarrow |X^n| \text{ 爲可積}, (\because \text{ 定理 4.8}) \\ &\Rightarrow |X^m| \text{ 爲可積}, (\because (1)) \\ &\Rightarrow X^m \text{ 爲可積}, (\because \text{ 定理 4.8}) \end{aligned}$$ □

以下兩個定理, 對於機率論之理論發展極具重要性, 稍後某些定理之證明會用到. 限於授課時間, 我們不予證明, 建議讀者在實變函數論或高等機率論課程中再深入研究其證明.

定理 4.10 [**單調收斂定理 (Monotone Convergence Theorem)**]

若 $0 \leq X_n \uparrow X$, 則 $E[X_n] \uparrow E[X]$. (註: X_n 未必爲簡單隨機變數).

證 參閱 Loève [11], p.125. □

定理 4.11 [**Lebesgue 收斂定理 (Dominated Convergence Theorem)**]

若 $\{X_n\}_n$ 爲一隨機變數序列, $\lim_{n\to+\infty} X_n(\omega) = X(\omega), \forall \omega \in \Omega$, 若存在一可積之隨機變數 Y 使得 $\forall n \in \mathbb{N}, |X_n| \leq Y$ a.s., 則 X 爲可積, 且$\lim_{n\to+\infty} E[X_n] = E[X]$.

證 參閱 Loève [11], p.126. □

定理 4.12 (變異數之性質)

(1) 若 $X = c$ a.s., 則 $\mathrm{Var}\ X = 0$, (內c 爲一常數).

(2) $\mathrm{Var}\,(cX) = c^2\,\mathrm{Var}\ X$, (內 c 爲一常數).

(3) $\mathrm{Var}\,(X + k) = \mathrm{Var}\ X$, (內 k 爲一常數).

(4) $\mathrm{Var}\ X < +\infty$ 之充要條件爲 X^2 爲可積. 此時,

$$\mathrm{Var}\ X = E[X^2] - (EX)^2.$$

證 證明十分容易, 讀者自證之. □

§ 4.3 常用隨機變數之期望值及變異數

(I) 若隨機變數 $X \sim B(n, p)$, 則 $E[X] = np, \mathrm{Var}\ X = npq$.

證 X 之期望值爲

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x\frac{n!}{x!(n-x)!}p^x q^{n-x} \\
&= \sum_{x=1}^{n} \frac{n\cdot(n-1)!}{(x-1)!(n-x)!}p\cdot p^{x-1}q^{n-x} \\
&= np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!}p^{x-1}q^{(n-1)-(x-1)} \\
&= np\sum_{y=0}^{n-1} \frac{(n-1)!}{y!((n-1)-y)!}p^y q^{(n-1)-y}, \ (\text{令 } y = x-1) \\
&= np(p+q)^{n-1} = np.
\end{aligned}$$

仿上法, 可證 $E[X(X-1)] = n(n-1)p^2$, 是以

$$\begin{aligned}
\mathrm{Var}\ X &= E[X^2] - (EX)^2 \\
&= E[X(X-1)] + E[X] - (EX)^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= np(1-p) = npq.
\end{aligned}$$
□

(II) 若隨機變數 $X \sim P(\lambda)$, 則 $E[X] = \mathrm{Var}\ X = \lambda$.

證 X 之期望值爲

$$E[X]=\sum_{x=0}^{+\infty}x\frac{e^{-\lambda}\lambda^x}{x!}=\sum_{x=1}^{+\infty}\frac{e^{-\lambda}\cdot\lambda\cdot\lambda^{x-1}}{(x-1)!}=\lambda e^{-\lambda}\sum_{x=1}^{+\infty}\frac{\lambda^{x-1}}{(x-1)!}$$
$$=\lambda e^{-\lambda}\sum_{y=0}^{+\infty}\frac{\lambda^y}{y!},\ (\text{令 } y=x-1)$$
$$=\lambda e^{-\lambda}e^{\lambda}=\lambda.$$

仿上法, 可證 $E[X(X-1)]=\lambda^2$, 是以

$$\text{Var } X=E[X^2]-(EX)^2=E[X(X-1)]+E[X]-(EX)^2=\lambda.$$ □

(III) 若隨機變數 $X\sim NB(r,p)$, (負二項分配), 即其密度函數爲

$$f\colon\mathbb{R}\to[0,1]:f(x)=\begin{cases}p^r\dbinom{r+x-1}{x}q^x, & \text{若 } x\in\mathbb{Z}_+,\\ 0, & \text{否則;}\end{cases}$$

則 $E[X]=\dfrac{rq}{p}$, $\text{Var } X=\dfrac{rq}{p^2}$.

證 仿照二項分配方法, 讀者自證之 (習題). □

(IV) 若隨機變數 $X\sim H(m,n,r)$, (超幾何分配), 即其密度函 數爲

$$f\colon\mathbb{R}\to[0,\,1]:f(x)=\begin{cases}\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}, & \text{若 } x\in A,\\ 0, & \text{若 } x\in\mathbb{R}\setminus A,\end{cases}$$

內 $A=\{j\in\mathbb{Z}_+\mid\max\{0,\,r-n\}\le j\le\min\{m,\,r\}\}$. 則

$$E[X]=\frac{mr}{m+n},\text{Var } X=\frac{mnr(m+n-r)}{(m+n)^2(m+n-1)}.$$

證 仿照二項分配方法, 讀者自證之 (習題). □

(V) 若隨機變數 $X\sim N(0,1)$, 則 $\forall n\in\mathbb{N}$,

$$E[X^{2n-1}]=0,E[X^{2n}]=\frac{(2n)!}{2^n n!}.$$

證 為方便計, 令 $m_k = E[X^k]$, 內 $k \in \mathbb{Z}_+$, 則

$$m_{2n} = E[X^{2n}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2n} \exp\Big(-\frac{x^2}{2}\Big)\, dx$$
$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{+\infty} x^{2n} \exp\Big(-\frac{x^2}{2}\Big)\, dx.$$

利用分部積分法, 令 $u = x^{2n-1},\ dv = xe^{-x^2/2}dx$, 則

$$\begin{aligned}
m_{2n} &= \frac{2}{\sqrt{2\pi}}\Big[\Big[(-x^{2n-1}\exp\Big(-\frac{x^2}{2}\Big)\Big]_0^{+\infty} + (2n-1)\int_0^{+\infty} x^{2n-2}\exp\Big(-\frac{x^2}{2}\Big)\,dx\Big] \\
&= (2n-1)m_{2n-2} \\
&= (2n-1)(2n-3)m_{2n-4} \\
&= \cdots \\
&= (2n-1)(2n-3)\cdots 3\cdot 1\cdot m_0 \\
&= (2n-1)(2n-3)\cdots 3\cdot 1, (\text{因 } m_0 = E[X^0] = 1\ ) \\
&= \frac{(2n)!}{2^n n!}.
\end{aligned}$$

其次, 由於 X^{2n} 為可積, 由上節之定理 4.9 知 X^{2n-1} 亦為可積, 又因 $x^{2n-1}\exp\Big(-\frac{x^2}{2}\Big)$ 為一奇函數, 故

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} x^{2n-1}\exp\Big(-\frac{x^2}{2}\Big)\,dx = \frac{-1}{\sqrt{2\pi}}\int_{0}^{+\infty} x^{2n-1}\exp\Big(-\frac{x^2}{2}\Big)\,dx,$$

是以

$$E[X^{2n-1}] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x^{2n-1}\exp\Big(-\frac{x^2}{2}\Big)\,dx = 0. \qquad \square$$

(V') | 若隨機變數 $X \sim N(\mu, \sigma^2)$, 則 $EX = \mu,\ \mathrm{Var}\ X = \sigma^2.$

證 由定理 2.24 知, 隨機變數 $Y = \frac{X-\mu}{\sigma}$ 具 $N(0,1)$ 分配. 由 (V) 知, $EY = 0$, $\mathrm{Var}\ Y = 1$, 是以

$$\begin{aligned}
E[X] &= E\Big[\frac{X-\mu}{\sigma}\cdot\sigma + \mu\Big] = \sigma E[Y] + \mu = \mu\ ; \\
\mathrm{Var}\ X &= \mathrm{Var}\ \Big(\frac{X-\mu}{\sigma}\cdot\sigma + \mu\Big) \\
&= \mathrm{Var}\ \Big(\frac{X-\mu}{\sigma}\cdot\sigma\Big) = \sigma^2\,\mathrm{Var}\,(Y) = \sigma^2.
\end{aligned} \qquad \square$$

(VI) | 若隨機變數 $X \sim G(\alpha, \beta)$, 則 $E[X] = \alpha\beta, \mathrm{Var}\ X = \alpha\beta^2.$

證 X 之期望值為

$$\begin{aligned}
E[X] &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x\cdot x^{\alpha-1}e^{-x/\beta}dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x^{\alpha}e^{-x/\beta}dx \\
&= \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} \frac{1}{\Gamma(\alpha+1)\beta^{\alpha+1}}x^{\alpha}e^{-x/\beta}\,dx \\
&= \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^\alpha}, \text{(因被積函數為 } G(\alpha+1,\beta) \text{ 之 p.d.f.)} \\
&= \alpha\beta.
\end{aligned}$$

仿照上述方法, 可得

$$E[X^2] = \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x^{\alpha+1}e^{-x/\beta}dx = \frac{\Gamma(\alpha+2)\beta^{\alpha+2}}{\Gamma(\alpha)\beta^\alpha} = \alpha(\alpha+1)\beta^2.$$

故

$$\text{Var } X = E[X^2]-(EX)^2 = \alpha(\alpha+1)\beta^2-\alpha^2\beta^2 = \alpha\beta^2.$$ □

(VII) 若隨機變數 $X\sim\chi_r^2$, 則 $E[X]=r,\ \text{Var } X = 2r.$

證 因 χ^2 分配乃 $G(r/2,\,2)$ 分配之故. □

(VIII) 若隨機變數 X 具負指數分配, 參數為 λ, 則 $E[X]=1/\lambda, \text{Var } X = 1/\lambda^2.$

證 因參數為 λ之負指數分配乃 $G(1,\,1/\lambda)$ 分配之故. □

(IX) 若隨機變數 X 具 Cauchy 分配, 則 X 之期望值不存在.

證 只證 $\mu=0,\ \sigma=1$ 之情形, 一般情形讀者自證之, (習題). 由於

$$\begin{aligned}
E[X^+] &= E[\max\{X,\,0\}] \\
&= \int_{-\infty}^{+\infty}\max\{x,\,0\}f(x)\,dx \\
&= \int_{-\infty}^{0}\max\{x,\,0\}f(x)\,dx + \int_{0}^{+\infty}\max\{x,\,0\}f(x)\,dx \\
&= 0+\frac{1}{\pi}\int_0^{+\infty} x\cdot\frac{dx}{1+x^2} \\
&= \frac{1}{2\pi}\ln(1+x^2)\,\Big|_0^{+\infty} = +\infty;
\end{aligned}$$

$$
\begin{aligned}
E[X^-] &= E[\max\{-X,\,0\}] \\
&= \int_{-\infty}^{+\infty} \max\{-x,\,0\} f(x)\,dx \\
&= \int_{-\infty}^{0} \max\{-x,\,0\} f(x)\,dx + \int_{0}^{+\infty} \max\{-x,\,0\} f(x)\,dx \\
&= \frac{1}{\pi}\int_{-\infty}^{0} (-x)\cdot \frac{dx}{1+x^2} + 0 \\
&= -\frac{1}{2\pi}\ln(1+x^2)\,\Big|_{-\infty}^{0} = +\infty;
\end{aligned}
$$

故 X 之期望值不存在. □

我們知道上述 Cauchy 密度函數 $f(x) = \dfrac{1}{\pi}\cdot\dfrac{1}{1+x^2}$ 之圖形對稱於 Y 軸, 直觀的看來, $x=0$ 應該是此一分布的中心, 然而其期望值卻不存在. 這正說明: 期望值雖是一種中央趨勢之測度 (measure of central tendency), 但與直觀之『中央』, 並非完全一致. 對於不爲可積之隨機變數, 我們通常以中位數代表其中心.

§ 4.4 Chebyshev 不等式

以賭博方式來解釋最初的期望值觀念, 當然是十分傳神的. 與期望值具有同等重要性的標準差, 目的在於測度某一分配『分散』的程度, 似乎並不能由其定義直接顯示出它的含義. 本節中, 我們將介紹早期機率論中一個重要的定理 —— Chebyshev 不等式:

$$\forall k>0, P\{|X-EX|\ge k\sigma_X\} \le \frac{1}{k^2}.$$

此一不等式可以顯示出以標準差做爲分配函數分散程度的一種測度乃是十分合理的. 此外, 我們還要證明標準差爲 0 就是分配完全不分散的觀念.

定理 4.13 [**基本不等式 (Basic inequality)**]

設 Y 爲 $(\Omega,\,\mathcal{F},\,P)$ 上一非負隨機變數, 則

$$\forall c>0,\ P\{Y\ge c\} \le \frac{EY}{c}.$$

證 由於

$$Y = Y\,I_{\{Y\ge c\}} + Y\,I_{\{Y<c\}} \ge Y\,I_{\{Y\ge c\}} \ge c\,I_{\{Y\ge c\}},$$

是以

$$E[Y] \ge c\cdot E[I_{\{Y\ge c\}}] = c\cdot P\{Y\ge c\}.$$ □

註: 若 Y 不為可積, 此不等式亦為眞, 因其右端為 $+\infty$.

定理 4.14 [**Markov 不等式**]

> 設 X 為一可積隨機變數, $\mu = EX$, $r > 0$, 則
> $$\forall c > 0, P\{|X-\mu| \geq c\} \leq \frac{E|X-\mu|^r}{c^r}.$$

證 令 $Y = |X-\mu|^r$, 則 Y 為一非負隨機變數, 由基本不等式知, $\forall c > 0$,

$$\begin{aligned} P\{|X-\mu| \geq c\} &= P\{|X-\mu|^r \geq c^r\} \\ &= P\{Y \geq c^r\} \\ &\leq \frac{EY}{c^r} = \frac{E|X-\mu|^r}{c^r}. \end{aligned}$$ □

定理 4.15 [**Chebyshev 不等式**]

> 設 X 為一可積隨機變數, $\mu = EX, \sigma^2 = \mathrm{Var}\ X$, 則
> $$\forall c > 0, P\{|X-\mu| \geq c\} \leq \frac{\mathrm{Var}\ X}{c^2}. \tag{1}$$
> 再者, 若 $\sigma > 0$, 則
> $$\forall k > 0, P\{|X-\mu| \geq k\sigma\} \leq \frac{1}{k^2}. \tag{2}$$

證 在 Markov 不等式中, 令 $r = 2$ 即得 (1) 式; 至於 (2) 式, 係令 $c = k\sigma > 0$ 而得. □

對於初學者而言, 以標準差做為某一隨機變數之分配之『分散程度』的一種測度, 並不十分容易接受. Chebyshev 不等式中恰巧可提供我們, 對此一概念一個很好的解釋:

一、設 μ 及 σ 分別表某隨機變數 X 之期望值及標準差, 由 Chebyshev 不等式知,

$$P\{\mu - k\sigma < X < \mu + k\sigma\} = P\{|X-\mu| < k\sigma\} \geq 1 - \frac{1}{k^2}.$$

意思是 X 值在區間 $(\mu - k\sigma,\ \mu + k\sigma)$ 內之機率至少為 $1 - \frac{1}{k^2}$, 當 $k = 2$ 時, 上式表示 X 值在區間 $(\mu - 2\sigma,\ \mu + 2\sigma)$ 內之機率至少為 3/4. 當 $k = 3$ 時, 上式表示 X 值在區間

$(\mu-3\sigma,\, \mu+3\sigma)$ 內之機率至少爲 8/9, 參閱圖 4–2.

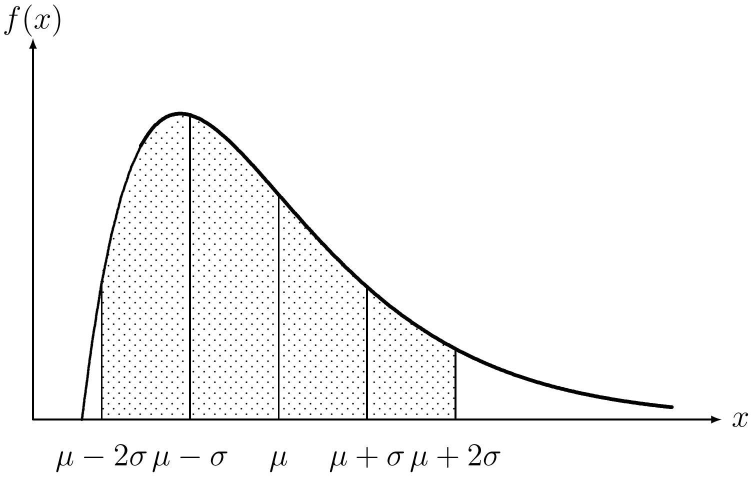


圖 4–2

二、設 μ_j 及 σ_j 分別表 $X_j, (j=1,\, 2)$, 之期望值及標準差, 由 Chebyshev 不等式知,

$$P\{|X_1-\mu_1|<k\sigma_1\} \geq 1-k^{-2}, \tag{1}$$

$$P\{|X_2-\mu_2|<k\sigma_2\} \geq 1-k^{-2}. \tag{2}$$

假定 $\mu_1=\mu_2=\mu, \sigma_1=1, \sigma_2=2\,(=2\sigma_1)$, 則

- (1) 式表示 X_1值在區間 $(\mu-3,\, \mu+3)$ 內之機率至少爲 8/9.
- (2) 式表示 X_2值在區間 $(\mu-6,\, \mu+6)$ 內之機率至少爲 8/9.

此乃說明有 8/9 以上之 X_1 值在較短區間 $(\mu-3,\, \mu+3)$ 內, 而有 8/9 以上之 X_2 值在較長區間 $(\mu-6,\, \mu+6)$ 內, 顯然 X_2 較為分散.

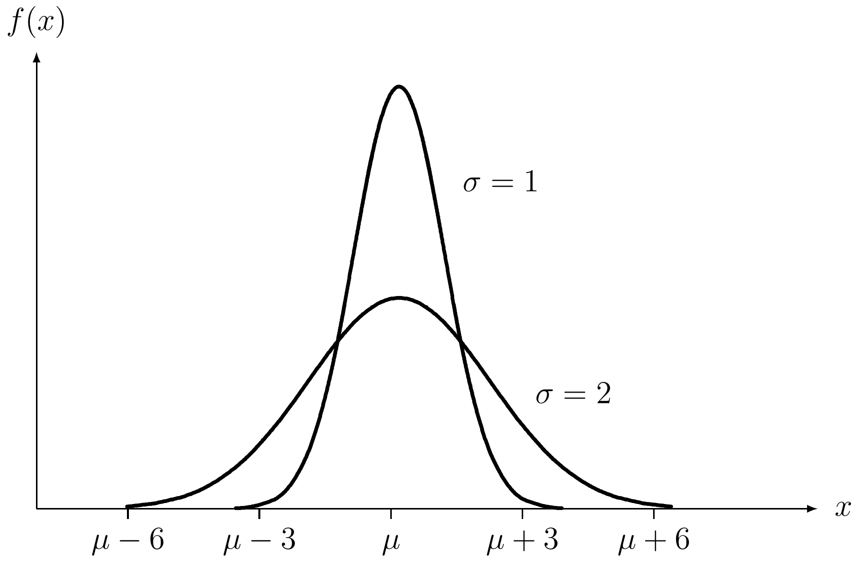


圖 4–3

三、在稍後之系 4.17中, 我們知道當標準差為 0 時, X 之值殆必集中在一點.

我們稱一隨機變數 X 為**退化的**(degenerate), 如果存在一常數 $c \in \mathbb{R}$, 使得 $P\{X = c\} = 1$, 亦即 $X = c$ a.s. 以下定理將討論 X 為退化之充要條件.

定理 4.16

> 設 X 為一隨機變數, 則以下三敍述為對等:
>
> (1) X 為退化的;
>
> (2) $\exists c \in \mathbb{R},\ \forall r > 0, E[|X - c|^r] = 0$;
>
> (3) $\exists c \in \mathbb{R},\ \exists r > 0$, 使得 $E[|X - c|^r] = 0$.

證 若 X 為退化的, 則存在 $c \in \mathbb{R}$ 使得 $X = c$ a.s. 亦即 $|X - c|^r = 0$ a.s., $\forall r > 0$, 顯然 (1) $\Rightarrow$ (2) 為眞. (2) $\Rightarrow$ (3) 顯而易見. 我們將證明 (3) $\Rightarrow$ (1); 由基本不等式知, $\forall \epsilon > 0$,

$$P\{|X - c| \geq \epsilon\} = P\{|X - c|^r \geq \epsilon^r\} \leq \frac{E|X - c|^r}{\epsilon^r} = 0,$$

以 $1/n$ 代上式中之 ϵ, 則得

$$P\{|X - c| \geq 1/n\} = 0, (內\ n \in \mathbb{N})$$

利用 P 之連續性,

$$\begin{aligned}
1 &= \lim_{n\to+\infty} P\{|X - c| < 1/n\} \\
&= P(\lim_{n\to+\infty} \{|X - c| < 1/n\}) \\
&= P\Big(\lim_{n\to+\infty} \Big\{c - \frac{1}{n} < X < c + \frac{1}{n}\Big\}\Big) \\
&= P\Big(\lim_{n\to+\infty} X^{-1}\Big(c - \frac{1}{n},\ c + \frac{1}{n}\Big)\Big) \\
&= P\Big(\bigcap_{n=1}^{+\infty} X^{-1}\Big(c - \frac{1}{n},\ c + \frac{1}{n}\Big)\Big), \text{ (集合序列為遞減)} \\
&= P\Big(X^{-1}\Big(\bigcap_{n=1}^{+\infty}\Big(c - \frac{1}{n}, c + \frac{1}{n}\Big)\Big)\Big) \\
&= P\{X^{-1}\{c\}\} = P\{X = c\}.
\end{aligned}$$

故知 X 為退化. □

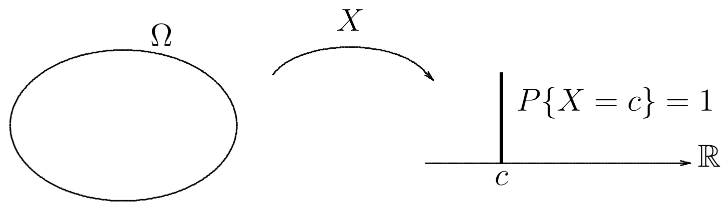


圖 4–4　退化隨機變數

系 4.17

> 隨機變數 X 爲退化之充要條件爲 $\mathrm{Var}\ X = 0$.

§ 4.5 Cauchy-Schwarz 不等式與協方差

本節中我們將研究二隨機變數間之相互關係. 我們是以 Cauchy-Schwarz 不等式做爲基礎, 進而研究協方差以及相關係數的概念.

定理 4.18 [**Cauchy-Schwarz 不等式**]

> 設 X 與 Y 均爲隨機變數. 若 $E[X^2] < +\infty$, $E[Y^2] < +\infty$, 則
>
> (1) $E|XY| < +\infty$;
>
> (2) $(E|XY|)^2 \le (E[X^2])(E[Y^2])$.

$\boxed{\text{證}}$ (1) $0 \le (|X| - |Y|)^2 = |X|^2 + |Y|^2 - 2|XY|$

$$\Rightarrow\ 2|XY| \le |X|^2 + |Y|^2$$
$$\Rightarrow\ 2E|XY| \le E|X|^2 + E|Y|^2 < +\infty.$$

(2) $0 \le (|X|t + |Y|)^2,\ \forall t \in \mathbb{R}$

$$\begin{aligned}
\Rightarrow\ & 0 \le E(|X|t + |Y|)^2 \\
& = E|X|^2\, t^2 + 2E|XY|\, t + E|Y|^2, \forall t \in \mathbb{R} \\
& \quad (\text{令 } A = E|X|^2, B = 2E|XY|, C = E|Y|^2) \\
\Rightarrow\ & 0 \le A\, t^2 + B\, t + C, \forall t \in \mathbb{R} \\
\Rightarrow\ & \triangle = B^2 - 4AC \le 0,\ (\text{二次方程式之判別式}) \\
\Rightarrow\ & 4(E|XY|)^2 - 4E|X|^2 \cdot E|Y|^2 \le 0 \\
\Rightarrow\ & (E|XY|)^2 \le (E[X^2])(E[Y^2]).
\end{aligned}$$

□

系 4.19

> X 與 Y 爲二隨機變數. 若 $EX = \mu_1$, $EY = \mu_2$, $\mathrm{Var}\ X = \sigma_1^2$, $\mathrm{Var}\ Y = \sigma_2^2$ 均爲有限, 則
>
> (1) $-\sigma_1\sigma_2 \le E(X - \mu_1)(Y - \mu_2) \le \sigma_1\sigma_2$;
>
> (2) $E(X - \mu_1)(Y - \mu_2) = \sigma_1\sigma_2 \Leftrightarrow P\{\sigma_1(Y - \mu_2) = \sigma_2(X - \mu_1)\} = 1$;
>
> (3) $E(X - \mu_1)(Y - \mu_2) = -\sigma_1\sigma_2 \Leftrightarrow P\{\sigma_1(Y - \mu_2) = -\sigma_2(X - \mu_1)\} = 1$.

證 若 $\sigma_1 = 0$ 或 $\sigma_2 = 0$, 即 $X = \mu_1$ a.s. 或 $Y = \mu_2$ a.s. 則上述三結論顯然均成立，故以下之證明設 σ_1 與 σ_2 均爲正.

(1) 令 $X_1 = \dfrac{X-\mu_1}{\sigma_1},\ Y_1 = \dfrac{Y-\mu_2}{\sigma_2}$,

$$\begin{aligned}
EX =&\mu_1,\ EY = \mu_2, \text{Var } X = \sigma_1^2,\ \text{Var } Y = \sigma_2^2\\
&\Rightarrow\ E[X_1] = 0,\ E[Y_1] = 0, \text{Var } X_1 = 1, \text{Var } Y_1 = 1\\
&\Rightarrow\ E[X_1] = 0,\ E[Y_1] = 0, E[X_1^2] = 1, E[Y_1^2] = 1\\
&\Rightarrow\ (E|X_1Y_1|)^2 \le (E[X_1^2])(E[Y_1^2]) = 1,\ (\because \text{Cauchy-Schwarz 不等式})\\
&\Rightarrow\ (E|(X-\mu_1)(Y-\mu_2)|)^2 \le \sigma_1^2\sigma_2^2\\
&\Rightarrow\ (E(X-\mu_1)(Y-\mu_2))^2 \le \sigma_1^2\sigma_2^2. \text{ (定理 4.8)}
\end{aligned}$$

(2) $$\begin{aligned}
P\{\sigma_1(Y-\mu_2) &= \sigma_2(X-\mu_1)\} = 1\\
&\Rightarrow\ X_1 = Y_1 \text{ a.s.}\\
&\Rightarrow\ E\Big[\frac{(X-\mu_1)(Y-\mu_2)}{\sigma_1\sigma_2}\Big] = E[X_1Y_1] = E[X_1^2] = 1\\
&\Rightarrow\ E(X-\mu_1)(Y-\mu_2) = \sigma_1\sigma_2.
\end{aligned}$$

反之,

$$\begin{aligned}
E(X-\mu_1)(Y-\mu_2) &= \sigma_1\sigma_2\\
&\Rightarrow\ E[X_1Y_1] = 1\\
&\Rightarrow\ E(X_1-Y_1)^2 = E[X_1^2] - 2E[X_1Y_1] + E[Y_1^2] = 0\\
&\Rightarrow\ \text{Var}\,(X_1-Y_1) = 0\\
&\Rightarrow\ X_1 = Y_1 \text{ a.s., (系 4.19)}\\
&\Rightarrow\ P\{\sigma_1(Y-\mu_2) = \sigma_2(X-\mu_1)\} = 1.
\end{aligned}$$

(3) 同理, 讀者自證之. □

定義 4.20

設 X 與 Y 爲二可積之隨機變數. 且 $EX = \mu_1, EY = \mu_2, \text{Var } X = \sigma_1^2, \text{Var } Y = \sigma_2^2$.

(1) 若 $E(X-\mu_1)(Y-\mu_2)$ 存在, 則稱其爲 X 與 Y 之**協方差**(covariance), 並以 $\text{Cov}(X, Y)$ 或 $C(X, Y)$ 表之.

(2) 若 $\sigma_1,\ \sigma_2$ 均爲有限正數, 則令

$$\rho(X, Y) = \rho_{X,Y} = \frac{C(X, Y)}{\sigma_1\sigma_2},$$

並稱其爲 X 與 Y 之**相關係數**(correlation coefficient).

定理 4.21 （協方差之性質）

(1) $C(X,\,Y) = E[XY] - (EX)(EY)$.

(2) $C(X,\,X) = \mathrm{Var}\ X$.

證 (1) 由協方差之定義,

$$\begin{aligned} C(X,\,Y) &= E(X-EX)(Y-EY) \\ &= E[XY - X(EY) - (EX)Y + (EX)(EY)] \\ &= E[XY] - (EX)(EY) - (EX)(EY) + (EX)(EY) \\ &= E[XY] - (EX)(EY). \end{aligned}$$

(2) 易明. □

定理 4.22 （相關係數之性質）

(1) $-1 \le \rho(X,\,Y) \le 1$. (若其存在).

(2) 設 $\mathrm{Var}\ X > 0, \mathrm{Var}\ Y > 0$, 則 $\rho(X,\,Y) = 1$ 之充要條件爲存在 $m > 0$ 及 $b \in \mathbb{R}$ 使得

$$P\{Y = mX + b\} = 1;$$

亦即 $(X,\,Y)$ 殆必位於一斜率爲正之直線上.

(3) 設 $\mathrm{Var}\ X > 0, \mathrm{Var}\ Y > 0$, 則 $\rho(X,\,Y) = -1$ 之充要條件爲存 在 $m < 0$ 及 $b \in \mathbb{R}$ 使得

$$P\{Y = mX + b\} = 1;$$

亦即 $(X,\,Y)$ 殆必位於一斜率爲負之直線上.

註: $P\{Y = mX + b\} = 1$ 表隨機向量 (X, Y) 之值落在直線 $y = mx + b$ 上之機率爲 1. 理由是: 設 $L = \{(x, y) \mid y = mx + b\}$, 則

$$\begin{aligned} \{Y = mX + b\} &= \{\omega \mid Y(\omega) = mX(\omega) + b\} \\ &= \{\omega \mid (X(\omega), Y(\omega)) \in L\} \\ &= \{(X, Y) \in L\}. \end{aligned}$$

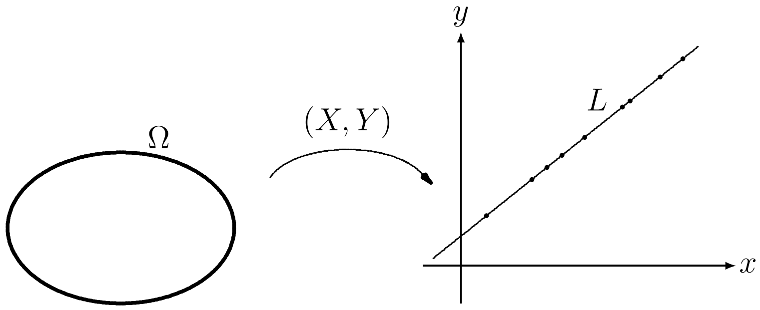


圖 4–5　直線相關

證 (1) 由 Cauchy-Schwarz 不等式定理之系之 (1), 立即可得.

(2) 1° 若 $\rho(X,\, Y)=1$, 則由 Cauchy-Schwarz 不等式定理之系之 (2) 知,

$$P\{\sigma_1(Y-\mu_2)=\sigma_2(X-\mu_1)\}=1,$$

亦即

$$P\Big\{Y=\mu_2+\frac{\sigma_2}{\sigma_1}(X-\mu_1)\Big\}=1.$$

只需令 $m=\dfrac{\sigma_2}{\sigma_1},\ b=\mu_2-\dfrac{\mu_1\sigma_2}{\sigma_1}$ 即可.

2° 若 $P\{Y=mX+b\}$=1, 即 $Y=mX+b$ a.s. 則

$$\begin{aligned}\mu_2&=EY=mEX+b=m\mu_1+b,\\ \sigma_2&=(\mathrm{Var}\ Y)^{1/2}=(\mathrm{Var}\,(mX+b))^{1/2}=m\sigma_1;\end{aligned}$$

是以

$$\begin{aligned}1&=P\{Y=mX+b\}\\ &=P\{Y=mX+\mu_2-m\mu_1\}\\ &=P\{Y-\mu_2=m(X-\mu_1)\}\\ &=P\{\sigma_1(Y-\mu_2)=\sigma_2(X-\mu_1)\};\end{aligned}$$

由 Cauchy-Schwarz 不等式定理之系之 (2) 知 $\rho(X,\, Y)=1$.

(3) 仿 (2) 讀者自證之. □

例 1. 試求第三章第四節例 1 中, 十位學生身高體重之相關係數.

解 首先, 我們推導有限樣本空間之相關係數公式. 設

$$\Omega=\{\omega_1,\cdots,\omega_n\},\ P\{\omega_j\}=1/n,\ \forall j=1,\cdots,n.$$

$$(X,\,Y)(\omega_j)=(x_j,\,y_j),\ (x_j,\,y_j\ \text{分別表學生}\ \omega_j\ \text{之身高及體重}).$$

則因 $X = \sum_{j=1}^{n} x_j I_{\{\omega_j\}}$, $Y = \sum_{j=1}^{n} y_j I_{\{\omega_j\}}, XY = \sum_{j=1}^{n} x_j y_j I_{\{\omega_j\}}$, 均爲簡單隨機變數, 故

$$\begin{aligned} C(X, Y) &= E[XY] - (EX)(EY) \\ &= \frac{1}{n}\sum_{j=1}^{n} x_j y_j - \frac{1}{n^2}\Big(\sum_{j=1}^{n} x_j\Big)\Big(\sum_{j=1}^{n} y_j\Big), \end{aligned}$$

$$\sigma_1^2 = \operatorname{Var} X = E[X^2] - (EX)^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \frac{1}{n^2}\Big(\sum_{j=1}^{n} x_j\Big)^2,$$

$$\sigma_2^2 = \operatorname{Var} Y = E[Y^2] - (EY)^2 = \frac{1}{n}\sum_{j=1}^{n} y_j^2 - \frac{1}{n^2}\Big(\sum_{j=1}^{n} y_j\Big)^2.$$

代入相關係數之定義公式之中, 得

$$\rho(X, Y) = \frac{n\sum x_j y_j - (\sum x_j)(\sum y_j)}{\sqrt{(n\sum x_j^2 - (\sum x_j)^2)(n\sum y_j^2 - (\sum y_j)^2)}}.$$

(此一公式與《統計學》中相關係數之公式相同). 在電腦未普及之前, 傳統之計算方法乃是將十位學生之資料寫入下表, 並予計算:

編號	身高 x_j	體重 y_j	x_j^2	y_j^2	$x_j y_j$
1	164	56	26896	3136	9184
2	164	58	26896	3364	9512
3	166	58	27556	3364	9628
4	166	58	27556	3364	9628
5	166	62	27556	3844	10292
6	170	58	28900	3364	9860
7	170	60	28900	3600	10200
8	170	60	28900	3600	10200
9	172	64	29584	4096	11008
10	176	64	30976	4096	11264
計	1684	598	283720	35828	100776

故身高 X 與體重 Y 之相關係數爲

$$\rho = \frac{10 \times 100776 - 1684 \times 598}{\sqrt{(10 \times 283720 - 1684^2)(10 \times 35828 - 598^2)}} = 0.764. \qquad \square$$

目前許多電腦軟體以及小計算器均附有計算相關係數之功能, 相當方便, 但讀者仍應了解相關係數最基本之運算與概念.

當二隨機變數 X 與 Y 之相關係數 $\rho(X, Y) = 0$ 時, 統計學者稱 X 與 Y 爲**不相關** (uncorrelated); 當 $\rho(X, Y) > 0$ 時, 稱 X 與 Y 爲**正相關** (positively correlated); 當 $\rho(X, Y) < 0$ 時, 稱 X 與 Y 爲**負相關** (negatively correlated).

§ 4.6 條件期望值及條件變異數

如果我們欲了解某一群體中, 身高 170 公分者之平均體重爲若干? 我們會先設法取得這些人之體重, 再平均之, 假定我們只有此一群體之聯合密度函數 $f(x, y)$, 我們又如何求得身高 170 公分者之平均體重? 由於我們已知 170 公分者體重之條件密度函數爲 $f(y|170)$, 利用離散型期望值公式可得

$$\text{身高 170 公分者之平均體重} = \sum_y y f(y|170).$$

設隨機變數 X 與 Y 均爲離散型或均爲連續型, f 爲 (X, Y) 之密度函數, f_1 及 f_2 分別爲 X 與 Y 之邊際密度函數. 在上一章中, 我們曾討過, 若 $f_1(x) > 0$, 則給 $X = x$ 後, Y 之條件密度函數爲

$$f(\cdot|x) \colon \mathbb{R} \to \mathbb{R}_+ : f(y|x) = \frac{f(x, y)}{f_1(x)}.$$

我們研究此一分配之期望值和變異數, 若存在, 應爲

$$E(Y|X = x) = E(Y|x) = \begin{cases} \displaystyle\sum_y y f(y|x), & \text{若 } (X, Y) \text{ 爲離散型}, \\ \displaystyle\int_{\mathbb{R}} y f(y|x) dy, & \text{若 } (X, Y) \text{ 爲連續型}, \end{cases}$$

$$\operatorname{Var}(Y|X = x) = \operatorname{Var}(Y|x) = \begin{cases} \displaystyle\sum_y (y - E(Y|x))^2 f(y|x), & \text{若 } (X, Y) \text{ 爲離散型}, \\ \displaystyle\int_{\mathbb{R}} (y - E(Y|x))^2 f(y|x) dy, & \text{若 } (X, Y) \text{ 爲連續型}, \end{cases}$$

並分別稱其爲給 $X = x$ 後, Y 之**條件期望值和條件變異數**. 同理, 我們也可界定: 給 $Y = y$ 後, X 之條件期望值和條件變異數.

若視 $E(Y|x) = m_2(x)$ 爲 x 之函數, 則此函數之圖形稱爲 **Y 對 X 之迴歸曲線**(regression curve of Y on X), 更明確地說: 設 $A_1 = \{x \in \mathbb{R} \mid f_1(x) > 0\}$, 則

$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[dr, "m_2 \circ X"'] & A_1 \subset \mathbb{R} \arrow[d, "m_2"] \\ & \mathbb{R} \end{tikzcd}$$

此外, X 與 m_2 之合成函數 $m_2 \circ X$ 又常寫爲 $E(Y|X)$, 此乃因爲 $\forall \omega \in \Omega$,

$$(m_2 \circ X)(\omega) = m_2(X(\omega)) = E(Y|X(\omega))$$

之故; 同理, 我們也可界定 X 對 Y 之迴歸曲線 $E(X|Y = y) = m_1(y)$, 以及 $E(X|Y) = m_1 \circ Y$. 提醒讀者注意: 此處之迴歸曲線與《統計學》中利用最小平方法所得迴歸曲線不同.

例 1. 試求第三章第四節例 1 中, 十位學生身高體重資料, 體重 Y 對身高 X 之迴歸曲線.

$\boxed{\text{解}}$ 由於樣本空間爲有限, 我們不妨列表計算較爲方便.

人 數		身 高 x (cm)				
		164	166	170	172	176
體	56	1	-	-	-	-
重	58	1	2	1	-	-
y	60	-	-	2	-	-
(kg)	62	-	1	-	-	-
	64	-	-	-	1	1
$m_2(x)$		57	$59\frac{1}{3}$	$59\frac{1}{3}$	64	64

有關 $m_2(x)$ 之計算, 其直觀意義爲: 本班中身高爲 x 之學生們的平均體重. 以 $x = 170$ 公分爲例, 其中一人爲 58 公斤, 二人爲 60 公斤, 其平均體重爲

$$\frac{58 + 60 + 60}{3} = 59\frac{1}{3} \text{ (公斤)}.$$

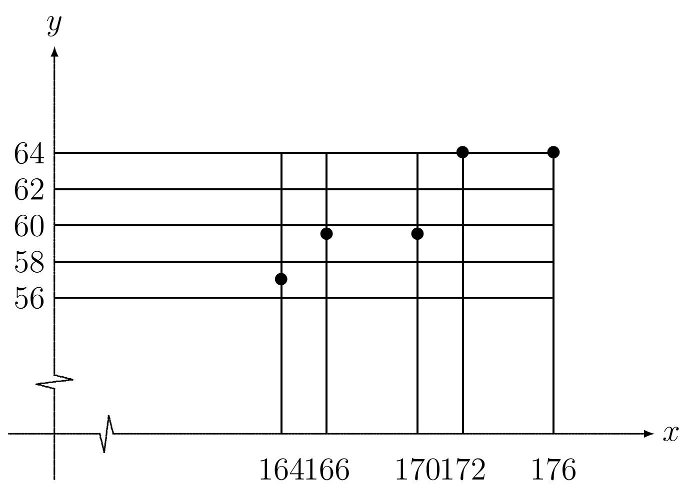


圖 4–6

若依定義求之, 則爲

$$m_2(170)=\sum yf(y|170)=58\times\frac{1}{3}+60\times\frac{2}{3}=59\frac{1}{3}\text{ (公斤)},$$

二者並無差異. 是以體重 Y 對身高 X 之迴歸曲線爲

$$m_2\colon A_1\to\mathbb{R}_+:m_2(x)=\begin{cases}57, & \text{若 } x=164,\\ 59\frac{1}{3}, & \text{若 } x\in\{166,\,170\},\\ 64, & \text{若 } x\in\{172,\,176\},\end{cases}$$

其中 $A_1=\{164,\,166,\,170,\,172,\,176\}$. 其圖形如上. □

例 2. 設連續型隨機向量 $(X,\,Y)$ 之密度函數爲

$$f\colon\mathbb{R}^2\to\mathbb{R}_+:f(x,\,y)=\begin{cases}2, & \text{若 } (x,\,y)\in A,\\ 0, & \text{否則},\end{cases}$$

內 $A=\{(x,\,y)\mid 0\le x\le y\le 1\}$. 試求:

(1) 給 $X=x$ 後, Y之條件期望值, 其中 $0\le x<1$; 並求 $E(Y|X)$;

(2) 給 $Y=y$ 後, X之條件期望值, 其中 $0<y\le 1$.

(3) EY 與 $E[E(Y|X)]$, 並檢驗二者是否相等.

解 (1) 由於 X 之邊際密度函數爲

$$f_1\colon\mathbb{R}\to\mathbb{R}_+:f_1(x)=\begin{cases}\displaystyle\int_x^1 2\,dx=2(1-x), & \text{若 } 0\le x<1,\\ 0, & \text{否則},\end{cases}$$

是以當 $x\in[0,\,1)$ 時, 由於給 $X=x$ 後, Y之條件密度函數爲

$$f(y|x)=\frac{f(x,\,y)}{f_1(x)}=\begin{cases}\dfrac{1}{1-x}, & \text{若 } y\in[x,\,1],\\ 0, & \text{否則}.\end{cases}$$

故給 $X=x$ 後, Y 之條件期望值爲

$$E(Y\mid X=x)=\int_x^1 yf(y|x)dy=\frac{1}{1-x}\int_x^1 y\,dy=\frac{1+x}{2},$$

此乃爲 Y 對 X 之迴歸曲線, 參見圖 4–7 之虛線. 此時,

$$E(Y|X)=\frac{1+X}{2}.$$

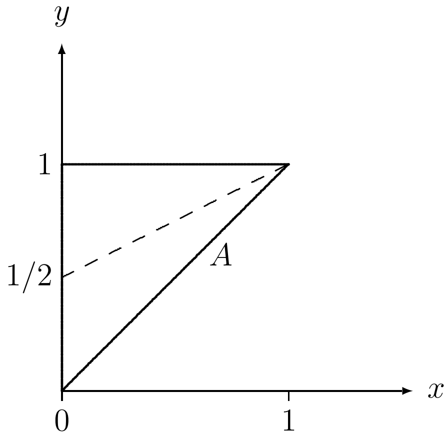


圖 4–7

(2) 由於 Y 之邊際密度函數爲

$$f_2\colon \mathbb{R} \to \mathbb{R}_+ : f_2(y) = \begin{cases} \int_0^y 2\,dy = 2y, & \text{若 } 0 < y \le 1, \\ 0, & \text{否則}, \end{cases}$$

是以當 $y \in (0,\,1]$ 時, 由於給 $Y = y$ 後, X之條件密度函數爲

$$f(x|y) = \frac{f(x,\,y)}{f_2(y)} = \begin{cases} \dfrac{1}{y}, & \text{若 } x \in [0,\,y], \\ 0, & \text{否則}. \end{cases}$$

是以, 給 $Y = y$ 後 X之條件期望值爲

$$E(X|Y = y) = \int_0^y x \cdot \frac{1}{y}\,dx = \frac{y}{2}.$$

(3) 由於 $EX = \displaystyle\int_0^1 f_1(x)\,dx = \frac{1}{3}$, 是以

$$EY = \int_{\mathbb{R}} y f_2(y)\,dy = \int_0^1 2y^2\,dy = \frac{2}{3};$$

$$E[E(Y|X)] = E\Big[\frac{1+X}{2}\Big] = \frac{1+EX}{2} = \frac{2}{3}.$$

二者相等. □

定理 4.23

設 $E[Y]$ 及函數 m_2 皆存在且爲有限, 則

$$E(E(Y|X)) = E(m_2 \circ X) = E(Y).$$

證 只證連續型. 離散型同理, 讀者自證之. 設 $A=\{(x,y)\mid f(x,y)>0\}$, $A_1=\{x\mid \exists y\in\mathbb{R}, f(x,y)>0\}$. 則

$$\begin{aligned}
E(E(Y|X)) &= E(m_2\circ X)\\
&= \int_{\mathbb{R}} m_2(x)f_1(x)\,dx = \int_{A_1} m_2(x)f_1(x)\,dx\\
&= \int_{A_1} E(Y|x)f_1(x)\,dx\\
&= \int_{A_1}\Big(\int_{\mathbb{R}} yf(y|x)dy\Big)f_1(x)\,dx = \iint_{A_1\times\mathbb{R}} yf(y|x)f_1(x)\,dy\,dx\\
&= \iint_{A_1\times\mathbb{R}} yf(x,\,y)\,dx\,dy = \iint_{\mathbb{R}^2} yf(x,\,y)\,dx\,dy\\
&= E(Y).
\end{aligned}$$

□

本定理之原始意義爲: 全班之體重平均 $E[Y]$ 等於『各種身高者體重平均值之平均』, $E(E(Y|X))$; 這種方法常見於, 較大群體計算總平均時使用. (例如, 已知各大學之人數及學生平均身高, 而欲求全國大學生之平均身高). 此外本定理還有一些應用, 我們只舉一個範例.

例 3. 某甲參加一賭博如下: 一袋中置 A, B, C 三球, 隨機抽取其一, 抽到 A 球時得 20 元, 結束; 抽到 B 球時得 30 元, 球放回再抽; 抽到 C 球時得 50 元, 球放回再抽. 試問此一賭博之賭資應爲多少方爲公平?

解 一公平之賭博其賭資應等於此賭博所得之期望值. 設 Y 表某甲賭一次所得之款項, X 表他第一次抽到之球, $X=1$ 表抽得 A 球等等. 則

$$\begin{aligned}
E(Y) &= E(E(Y|X))\\
&= \sum_{x=1}^{3} E(Y|X=x)P\{X=x\}\\
&= E(Y|X=1)P\{X=1\}+E(Y|X=2)P\{X=2\}\\
&\qquad\qquad + E(Y|X=3)P\{X=3\}\\
&= 20\cdot\frac{1}{3}+(30+E(Y))\cdot\frac{1}{3}+(50+E(Y))\cdot\frac{1}{3},
\end{aligned}$$

由於 $E(Y)$ 顯然爲有限, 移項之得 $E(Y)=100$, 是以此一賭博之賭資應爲 100 元方爲公平.

□

定理 4.24

設以下各期望值皆存在且為有限, 則

(1) $\mathrm{Var}\,(E(Y|X)) \le \mathrm{Var}\,Y$;

(2) $\mathrm{Var}\,(E(Y|X)) = \mathrm{Var}\,Y \Leftrightarrow P\{Y = m_2 \circ X\} = 1$.

證 (1) 1° 由於

$$\begin{aligned}\mathrm{Var}\,Y &= E(Y-\mu_2)^2, (\text{內 } \mu_2 = EY)\\ &= E[(Y-m_2\circ X)+(m_2\circ X-\mu_2)]^2\\ &= E(Y-m_2\circ X)^2 + E(m_2\circ X-\mu_2)^2\\ &\qquad + 2E(Y-m_2\circ X)(m_2\circ X-\mu_2)\\ &= E(Y-m_2\circ X)^2 + E(m_2\circ X-\mu_2)^2 + 0, \qquad (*)\end{aligned}$$

上式最後一項為 0 之理由如下:

$$\begin{aligned}&E(Y-m_2\circ X)(m_2\circ X-\mu_2)\\ &= E[Y\cdot(m_2\circ X)] - E(m_2\circ X)^2 - \mu_2 EY + \mu_2 E(m_2\circ X)\\ &= E[Y\cdot(m_2\circ X)] - E(m_2\circ X)^2, (\text{定理 } 4.23)\\ &= E[E(Y\cdot(m_2\circ X)|X)] - E(m_2\circ X)^2, (\text{定理 } 4.23)\\ &= E(m_2\circ X)^2 - E(m_2\circ X)^2, (\dagger\ \text{理由下詳})\\ &= 0.\end{aligned}$$

(†) 只證連續型, 離散型同理. 令 $X(\omega) = x$,

$$\begin{aligned}E(Y\cdot(m_2\circ X)|X(\omega)) &= \int_{\mathbb{R}} y\cdot m_2(x) f(y|x)dy\\ &= m_2(x)\int_{\mathbb{R}} y\cdot f(y|x)dy\\ &= m_2(x)\cdot E(Y|x)\\ &= (m_2(x))^2\\ &= (m_2\circ X)^2(\omega), \forall\omega\in\Omega,\end{aligned}$$

故 $E(Y\cdot(m_2\circ X) \mid X) = (m_2\circ X)^2$.

2° 由 $(*)$ 式顯然有

$$\begin{aligned}\mathrm{Var}\,Y &\ge E(m_2\circ X-\mu_2)^2\\ &= E[m_2\circ X - E(m_2\circ X)]^2, (\text{定理 } 4.23)\\ &= \mathrm{Var}\,(m_2\circ X)\\ &= \mathrm{Var}\,(E(Y|X)).\end{aligned}$$

(2) (1) 中之 $(*)$ 式知

$$\begin{aligned}\operatorname{Var} Y = \operatorname{Var}(E(Y|X)) &\Leftrightarrow E(Y - m_2 \circ X)^2 = 0 \\ &\Leftrightarrow Y = m_2 \circ X \text{ a.s.} \\ &\Leftrightarrow P\{Y = m_2 \circ X\} = 1.\end{aligned}$$ □

例 4. (續例 1)

(1) 由於

- $E(E(Y|X)) = E(m_2 \circ X) = 57 \times \dfrac{2}{10} + 59\dfrac{1}{3} \times \dfrac{6}{10} + 64 \times \dfrac{2}{10} = 59.8,$
- $EY = 56 \times \dfrac{1}{10} + 58 \times \dfrac{4}{10} + 60 \times \dfrac{2}{10} + 62 \times \dfrac{1}{10} + 64 \times \dfrac{2}{10} = 59.8.$

二者相等.

(2) 但,

- $$\begin{aligned}\operatorname{Var}(E(Y|X)) &= \operatorname{Var}(m_2 \circ X) \\ &= E(m_2 \circ X)^2 - (E(m_2 \circ X))^2 \\ &= 57^2 \times \frac{2}{10} + \left(59\frac{1}{3}\right)^2 \times \frac{6}{10} + 64^2 \times \frac{2}{10} - 59.8^2 \\ &= 5.22\ ,\end{aligned}$$
- $\operatorname{Var} Y = E[Y^2] - (EY)^2 = 3582.2 - 59.8^2 = 6.74\ ,$

後者顯然大於前者. □

限於篇幅以及考慮初學者接受的程度, 我們對於條件期望值觀念, 只介紹以上部分, 其實這些觀念非但可以推廣到一般分配 (未必爲連續型或離散型), 而且也不必限於二隨機變數間之情形. Laha & Rohagi [10] 第六章中有很深入的探討, 有興趣者請參閱之.

第四章　習　　題

4-1 若隨機變數 $X \sim NB(r,\, p)$, (負二項分配), 試證:

$$E[X] = \frac{rq}{p}, \mathrm{Var}\ X = \frac{rq}{p^2}.$$

4-2 若隨機變數 $X \sim H(m,\, n,\, r)$, (超幾何分配),

(1) 試證: $E[X] = \dfrac{mr}{m+n}, \mathrm{Var}\ X = \dfrac{mnr(m+n-r)}{(m+n)^2(m+n-1)}$.

(2) 若 n 爲 m 之函數 且 $\lim\limits_{m\to+\infty} \dfrac{m}{m+n} = p, (0 < p < 1)$. 則當 $m \to +\infty$ 時, 上述期望值與變異數之極限, 是否分別等於二項分配之期望值與變異數? (此乃說明: 當黑球白球很多時, 放回與不放回之抽樣其實幾乎沒有差異).

4-3 設函數

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \dfrac{x}{a^2}\exp\Big(-\dfrac{x^2}{2a^2}\Big), & \text{若 } x > 0, \\ 0, & \text{若 } x \le 0, \end{cases}$$

(1) 證明 f 爲一密度函數, $\forall a \neq 0$;

(2) 設 f 爲隨機變數 X 之密度函數, 試證:

$$E[X] = |a|\sqrt{\frac{\pi}{2}}, \mathrm{Var}\ X = \Big(2 - \frac{\pi}{2}\Big)a^2.$$

4-4 設隨機變數 X 具 Beta 分配. 試求 $E[X^n]$, 並進而求 X 之期望值與變 異數.

4-5 設 F 隨機變數 X 之分配函數. 若已知 $E[X]$ 存在, 且

$$\lim_{x\to-\infty} xF(x) = 0, \lim_{x\to+\infty} x(1-F(x)) = 0.$$

(1) 試證: $EX = \displaystyle\int_0^{+\infty} (1-F(x))\,dx - \int_{-\infty}^{0} F(x)\,dx$.

(2) 繪圖以說明上述公式之幾何意義.

4-6 設隨機變數 X 之密度函數爲

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \frac{1}{2}e^{-|x-a|}, \text{內 } a \in \mathbb{R}.$$

(1) 試證: X 爲可積, 並求 EX;

(2) 試求 k 值以使 $P\{|X-EX|<k\}=0.5$.

4-7 已知一箱中有 4 個白球和 6 個黑球. 以不放回方式抽三球, 規定每抽中一白球可得獎金 100 元, 抽中一黑球則須付出 40 元.

(1) 試求抽得 j $(j=0,1,2,3)$ 個白球之機率.

(2) 試問從箱中抽出 3 球之所得的期望值是多少?

4-8 設某高中首次模擬考數學成績近似於常態分配 $N(50,64)$. 試求高標之分數 (即前半成績之平均).

[提示] 設 f 爲 $N(50,64)$ 之密度函數, 則前半學生成績之密度函數爲

$$g\colon \mathbb{R}\to\mathbb{R}_+ : g(x)=\begin{cases} 2\,f(x), & \text{若 } x\ge\mu, \\ 0, & \text{否則.}\end{cases}$$

4-9 設 $X\sim\chi^2_r,\ k>-\frac{r}{2}$, 試證:

$$E[X^k]=2^k\frac{\Gamma(\frac{r}{2}+k)}{\Gamma(\frac{r}{2})}.$$

4-10 舉一例: X 爲離散型隨機變數, 但其期望值不存在, 即 $EX^+=EX^-=+\infty$.

[提示] $f(x)=\dfrac{c}{x^2}, \forall x\in\mathbb{Z}^*$.

4-11 設隨機變數 X^2 爲可積. 試證:

(1) $\forall c\in\mathbb{R}, E(X-c)^2=E(X-EX)^2+(EX-c)^2$;

(2) 函數 $\phi(c)=E(X-c)^2$ 於點 $c=EX$ 有最小値 ;

(3) 若 $P\{a\le X\le b\}=1$, 則 $\mathrm{Var}\ X\le\dfrac{(b-a)^2}{4}$, 內 $a,\,b\in\mathbb{R}$;

(4) 若 $P\{a\le X\le b\}=1$, 則

$$\mathrm{Var}\ X=\frac{(b-a)^2}{4}\Leftrightarrow P\{X=a\}=P\{X=b\}=\frac{1}{2}.$$

4-12 設 X 爲一隨機變數.

(1) 試證: $E[X^2]=0\ \Rightarrow\ \mathrm{Var}\ X=0$;

(2) 舉一反例以證明 (1) 之逆敍述 “ $\mathrm{Var}\ X=0\ \Rightarrow\ E[X^2]=0$ ” 不眞.

4-13 若隨機變數 X 具一般之 Cauchy 分配, 試證 X 之期望值不存在.

4-14 若隨機變數 X 具二項分配 $B(100,\,1/2)$, 試利用 Chebyshev 不等式以求 c 之最小值使得

$$P\Big\{\Big|\frac{X}{100}-\frac{1}{2}\Big|\le c\Big\}\ge 0.95\ .$$

4-15 若隨機變數 X 具常態分配 $N(\mu, \sigma^2)$, 試證:

$$\forall k > 0, P\{|X - \mu| \geq k\sigma\} < \frac{1}{k^2}.$$

即二者不可能相等.

4-16 試問是否存在隨機變數 $Y > 0$, 滿足

$$\forall c > 0,\ P\{Y \geq c\} = \frac{EY}{c}.$$

4-17 設隨機向量 (X, Y) 具二元常態分配. 試證此分配中之參數 ρ 爲 X 與 Y 之相關係數.

4-18 設 A 與 B 爲二事件, 試證: $C(I_A, I_B) = P(A \cap B) - P(A)P(B)$, 內 I_A, I_B 分別爲 A 與 B 之指標函數.

4-19 設隨機向量 (X, Y) 之密度函數爲 $f\colon \mathbb{R}^2 \to \mathbb{R}_+$:

$$f(x, y) = \begin{cases} 6xy(2 - x - y), & \text{若 } 0 < x < 1, 0 < y < 1, \\ 0, & \text{否則.} \end{cases}$$

試求 X 對 Y 之迴歸曲線 $E(X \mid Y = y)$, 並繪出其圖形.

4-20 (1) 設隨機向量 (X, Y) 爲絕對連續型 (或離散型), 若 $f_Y(y_0) > 0$, 試證:

$$\operatorname{Var}(X|Y = y_0) = E(X^2|Y = y_0) - (E(X|Y = y_0))^2;$$

(2) 若隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \begin{cases} \dfrac{12}{5}x^2(x + y), & \text{若 } 0 < x < 1, 0 < y < 1, \\ 0, & \text{否則.} \end{cases}$$

試求 $\operatorname{Var}(X|Y = y_0)$, 內 $0 < y_0 < 1$.

4-21 設隨機變數 X 與 Y 具相同之分配; $g\colon \mathbb{R} \to \mathbb{R}$ 爲 Borel 可測, 若 $E[g \circ X]$ 存在, 則 $E[g \circ Y]$ 存在且 $E[g \circ Y] = E[g \circ X]$.

4-22 擲一銅板, 正面發生之機率爲 p, 此一試驗在正面發生 (一次) 時停止. 設 N 表拋擲次數.

(1) 試利用條件期望值方法以求 N 之期望值爲若干?

(2) 先求 N 之密度函數, 再求其期望值.

4-23 設 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \begin{cases} 2, & \text{若 } 0 \leq x \leq y \leq 1, \\ 0, & \text{否則,} \end{cases}$$

試求 $E(X|Y = y) =?$ $\operatorname{Var}(Y|X = x) =?$ $\operatorname{Var}(X|Y = y) =?$

Chapter 5

特徵函數

對於一隨機變數或一隨機向量, 我們研究其分配函數時, 從最容易接受的密度函數開始, 進而累積分配函數, 最後更推展到以測度觀念來表示的機率分配函數. 其後的期望值觀念, 最初的目的, 在於測度隨機變數的集中趨勢及分散程度, 後來也發展出一系列的不等式, 其中以 Chebyshev 不等式最具實用價值也最爲易懂. 但在機率理論的發展上, 卻需要依賴另一項新的觀念, 由於十九世紀分析學上 Laplace 變換及 Fourier 變換的成就, 使得機率論也大受影響, 進而發展出一套極具效用的工具, 就是所謂動差生成函數及特徵函數, 前者所受之限制遠較後者爲多, 因此, 特徵函數在機率論中乃成爲一項不可或缺的研究工具, 並且獲得相當可觀的進展. 本章中, 我們將研究特徵函數之基本性質, 隨後兩章, 我們將利用此種函數以探討機率論中各項更重要之理論.

§ 5.1 複數隨機變數與其期望值

定義 5.1

設 X 與 Y 爲二 (實值) 隨機變數.

(1) 函數

$$Z = X + iY : \Omega \to \mathbb{C} : Z(\omega) = X(\omega) + iY(\omega)$$

稱爲一**複數值隨機變數**(complex-valued random variable);

(2) 若 X 與 Y 均爲可積, 則稱 $Z = X + iY$ 爲**可積**, 且規定 Z 之**期望值**爲 $E[Z] = E[X] + iE[Y]$;

(3) 若 X 或 Y 不爲可積, 則稱 $Z = X + iY$ 不爲可積或 Z 之期望值不存在.

(4) 若 Z 爲可積, 則稱 $\text{Var } Z = E|Z - EZ|^2$ 爲 Z 之**變異數**.

讀者是否已注意到, 我們規定複數值隨機變數之 $\text{Var } Z = E|Z - EZ|^2$, 而不是 $\text{Var } Z = E(Z - EZ)^2$, 理由是變異數表示 Z 值與其中心 $E[Z]$ 之距離之平方的平均值.

定理 5.2 (期望值之線性性質)

> 設 Z, Z_1 與 Z_2 均為可積之複數值隨機變數, $c \in \mathbb{C}$, 則
>
> (1) $E[cZ] = cE[Z]$;
>
> (2) $E[Z_1 + Z_2] = E[Z_1] + E[Z_2]$.

證 (1) 設 $Z = X + iY,\ c = \alpha + i\beta$, 其中 $\alpha,\ \beta \in \mathbb{R}$, 則

$$\begin{aligned}
E[cZ] &= E[(\alpha + i\beta)(X + iY)] \\
&= E[\alpha X - \beta Y + i(\alpha Y + \beta X)] \\
&= \alpha EX - \beta EY + i(\alpha EY + \beta EX) \\
&= (\alpha + i\beta)(EX + iEY) = cE[Z].
\end{aligned}$$

(2) 設 $Z_1 = X_1 + iY_1,\ Z_2 = X_2 + iY_2$, 則

$$\begin{aligned}
E[Z_1 + Z_2] &= E[(X_1 + iY_1) + (X_2 + iY_2)] \\
&= E[(X_1 + X_2) + i(Y_1 + Y_2)] \\
&= E[X_1 + X_2] + iE[Y_1 + Y_2] \\
&= E[X_1] + E[X_2] + iE[Y_1] + iE[Y_2] \\
&= E[Z_1] + E[Z_2].
\end{aligned}$$

□

定理 5.3

> (1) 若 Z 為一絕對可積之複數值隨機變數, 即 $E|Z| < +\infty$, 則 Z 為可積, 且 $|EZ| \le E|Z|$.
>
> (2) Z 為可積之充要條件為 $E|Z| < +\infty$.

證 (1) 由於

$$|Z| = |X + iY| = \sqrt{X^2 + Y^2} \ge \begin{cases} |X|, \\ |Y|. \end{cases}$$

故知 $E|X| < +\infty,\ E|Y| < +\infty$, 即 X 與 Y 均為可積, 由定義 5.1知 Z 為可積. 此外, 以極坐標表 Z, 則

$$Z = Re^{i\alpha},$$

內 $R = |Z|,\ \alpha = \arg(Z)$ 均為實值隨機變數; 若以極坐標 re^{it} 表 $E[Z]$, 則

$$re^{it} = E[Z] = E[Re^{i\alpha}],$$

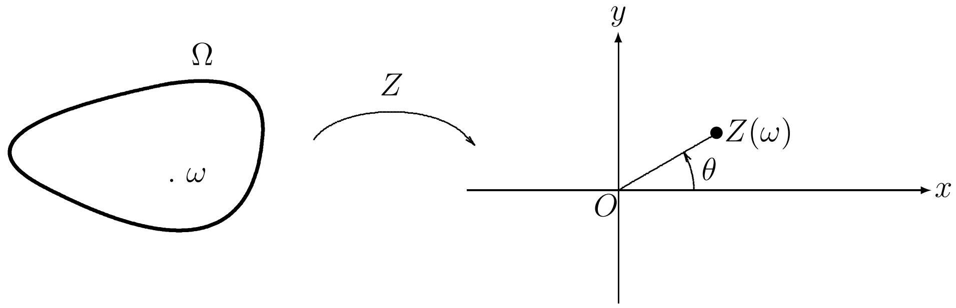


圖 5–1

是以

$$\begin{aligned}
|E[Z]| = r &= e^{-it}E[Re^{i\alpha}] \\
&= E[Re^{i(\alpha-t)}] \\
&= E[R\cos(\alpha-t) + iR\sin(\alpha-t)] \\
&= E[R\cos(\alpha-t)], (\text{虛數部分為 } 0) \\
&\leq E[R] = E|Z|.
\end{aligned}$$

(2) 由 (1) 知, 若 $E|Z| < +\infty$, 則 Z 為可積; 反之, 若 Z 為可積, 則因 $E|X| < +\infty$ 且 $E|Y| < +\infty$, 並且

$$|Z| = |X + iY| \leq |X| + |Y|,$$

故

$$E|Z| \leq E|X| + E|Y| < +\infty. \qquad \square$$

例 1. 設 $X \sim B(10,\, 0.3)$, $Y \sim N(1,\, 4)$, 試求複數隨機變數 $Z = X + iY$ 之期望值及變異數.

$\boxed{\text{解}}$ 顯然,

$$E[Z] = E[X] + iE[Y] = 3 + i.$$

至於 Z 之變異數,

$$\begin{aligned}
\text{Var } Z &= E|Z - E[Z]|^2 \\
&= E|X - EX + i(Y - EY)|^2 \\
&= E[|X - EX|^2] + E[|Y - EY|^2] \\
&= \text{Var } X + \text{Var } Y. \\
&= 10 \times 0.3 \times 0.7 + 4 = 6.1.
\end{aligned} \qquad \square$$

§ 5.2 隨機變數之特徵函數

定義 5.4

設 X 爲一實值隨機變數, F 爲其分配函數, 則函數

$$\begin{aligned}\phi_X : \mathbb{R} \to \mathbb{C} : \phi_X(t) &= E[e^{itX}] \\ &= E[\cos tX] + iE[\sin tX] \\ &= \int_{\mathbb{R}} e^{itx}\, dF(x)\end{aligned}$$

稱爲 X 之**特徵函數**(characteristic function), 簡爲 ch.f.

註: (1) 由 $|e^{ix}| = 1, \forall x \in \mathbb{R}$, 知 $|e^{itX}| = 1$, 由定理 5.3 知複數值隨機變數 e^{itX} 爲可積; 是以上述定義乃爲妥當 (well-defined).

(2) 若 X 爲絕對連續型, f 爲其之密度函數, 則 X 之特徵函數

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx}\, dF(x) = \int_{\mathbb{R}} e^{itx} f(x)\, dx$$

乃 f 之 Fourier 變換之變型, (關於 Fourier 變換, 請參閱 Apostol [1] p.326).

(3) 若無誤解之可能, ϕ_X 常簡爲 ϕ.

定理 5.5 (特徵函數之基本性質)

設 X 爲一實值隨機變數, 則

(1) $\phi_X(0) = 1$;

(2) $\forall t \in \mathbb{R}, |\phi_X(t)| \le 1$;

(3) 若 k 爲一實常數, 則 $\phi_{X+k}(t) = e^{itk}\phi_X(t)$;

(4) 若 c 爲一實常數, 則 $\phi_{cX}(t) = \phi_X(ct)$;

(5) $\phi_X(-t) = \overline{\phi_X(t)}$, (共軛複數) .

證 (1) $\phi_X(0) = E[e^{i\cdot 0\cdot X}] = E[e^0] = 1$.

(2) 因 $|e^{itX}| = 1$, 由定理 5.3 知

$$|\phi_X(t)| = |E[e^{itX}]| \le E|e^{itX}| = 1.$$

(此乃說明 ϕ_X 之值域必包含於複數平面之單位圓內.)

(3) $\phi_{X+k}(t) = E[e^{it(X+k)}] = E[e^{itk}e^{itX}] = e^{itk}\phi_X(t)$.

(4) $\phi_{cX}(t) = E[e^{itcX}] = \phi_X(ct)$.

(5) $\phi_X(-t) = E[e^{-itX}] = E[\cos(tX)] - iE[\sin(tX)] = \overline{\phi_X(t)}$. □

定理 5.6

任一 (實值) 隨機變數之特徵函數皆爲均勻連續.

證 往證: $\forall\epsilon > 0,\ \exists\delta > 0,\ (\forall t, u \in \mathbb{R})(|t-u| < \delta\ \Rightarrow\ |\phi_X(t) - \phi_X(u)| < \epsilon)$.
令 $u = t + h$, 由於

$$\begin{aligned}
|\phi_X(t+h) - \phi_X(t)| &= \Big|\int_{\mathbb{R}} e^{i(t+h)x}\, dF(x) - \int_{\mathbb{R}} e^{itx}\, dF(x)\Big| \\
&= \Big|\int_{\mathbb{R}} e^{itx}(e^{ihx} - 1)\, dF(x)\Big| \\
&\le \int_{\mathbb{R}} |e^{ihx} - 1|\, dF(x) \\
&= \int_{\mathbb{R}} 2|\sin(hx/2)|\, dF(x),\ (\dagger\ \text{理由下詳})
\end{aligned}$$

由於上述積分內之被積分小於或等於 2, 而 $\int_{\mathbb{R}} 2\, dF(x) = 2$, 利用 Lebesgue 受制收斂定理可得

$$\lim_{h\to 0}\int_{\mathbb{R}} 2|\sin(hx/2)|\, dF(x) = 2\int_{\mathbb{R}} \lim_{h\to 0}|\sin(hx/2)|\, dF(x) = 0,$$

是以 $\lim\limits_{h\to 0}|\phi_X(t+h) - \phi_X(t)| = 0$. 亦即,

$$\forall\epsilon > 0,\ \exists\delta > 0,\ (\forall h \in \mathbb{R})(|h| < \delta\ \Rightarrow\ |\phi_X(t+h) - \phi(t)| < \epsilon).$$

此乃對等於

$$\forall\epsilon > 0,\ \exists\delta > 0,\ (\forall t, u \in \mathbb{R})(|t-u| < \delta\ \Rightarrow\ |\phi_X(t) - \phi_X(u)| < \epsilon).$$

(†)
$$\begin{aligned}
|e^{i\alpha} - 1| &= |\cos\alpha + i\sin\alpha - 1| \\
&= ((\cos\alpha - 1)^2 + \sin^2\alpha)^{1/2} = (2 - 2\cos\alpha)^{1/2} \\
&= 2|\sin(\alpha/2)|,\ \forall\alpha \in \mathbb{R}.
\end{aligned}$$
□

定理 5.7

設 ϕ 爲隨機變數 X 之特徵函數. 若 X^n 爲可積, 內 n 爲一自然數, 則

$$\phi^{(n)}(0) = i^n E[X^n].$$

證 ϕ 之導函數為

$$\phi'(t) = \lim_{h\to 0} \frac{\phi(t+h)-\phi(t)}{h} = \lim_{h\to 0} \int_{\mathbb{R}} e^{itx} \frac{e^{ihx}-1}{h}\, dF(x)\ .$$

由於上式之被積函數之絕對值

$$\begin{aligned} \left| e^{itx} \frac{e^{ihx}-1}{h} \right| &= \left| \frac{2\sin(hx/2)}{h} \right|, \ (\text{見定理 5.6 證明之 } (\dagger)) \\ &\le \left| \frac{2}{h} \cdot \frac{hx}{2} \right| = |x|. \end{aligned}$$

而由原設知 $|x|$ 對於 F 為可積, 利用 Lebesgue 受制收斂定理,『極限』可與『積分』對調, 是以

$$\begin{aligned} \phi'(t) &= \int_{\mathbb{R}} \lim_{h\to 0} e^{itx} \frac{e^{ihx}-1}{h}\, dF(x) \\ &= i \int_{\mathbb{R}} x e^{itx}\, dF(x), \ (\text{l'Hospital 規則或導數定義}). \end{aligned}$$

同理,

$$\begin{aligned} \phi''(t) &= i^2 \int_{\mathbb{R}} x^2 e^{itx}\, dF(x), \\ &\ \vdots \\ \phi^{(n)}(t) &= i^n \int_{\mathbb{R}} x^n e^{itx}\, dF(x). \end{aligned}$$

是以 $\phi'(0) = iE[X],\ \phi''(0) = i^2E[X^2],\ \cdots,\ \phi^{(n)}(0) = i^nE[X^n]$. □

註: 本定理說明: 若"已知" $E[X^n]$ 存在且為有限, 則可由 ϕ 之 n 階導數求得 $E[X^n]$ 之值. 若僅知特徵函數 ϕ 而不知 $E[X^n]$ 是否為有限, 實際上亦無法求出 $E[X^n]$. 不過有一較為有利的定理: 若 $\phi^{(m)}(0)$ 為有限, 內 m 為一偶數, 則

$$\phi^{(n)}(0) = i^nE[X^n], \forall n \le m.$$

此一定理之證明較為困難, 需利用『對稱微分法』, 有興趣之讀者, 請參閱 Loève [11], p.212.

定理 5.8 [**Lévy** 反演公式 (Inversion formula)]

設 F 及 ϕ 分別為隨機變數 X 之分配函數及特徵函數, 則

$$F(b)-F(a) = \lim_{T\to+\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it} \phi(t)\, dt,$$

內 $a < b$ 均為 F 之連續點.

證 證明需要較多實變函數論之定理, 略之. 有興趣之讀者, 請參閱 Laha & Rohagi [10], p.144. □

定理 5.9 (**唯一性定理**)

設 X 與 Y 爲二隨機變數, 則 $F_X = F_Y$ 之充要條件爲 $\phi_X = \phi_Y$.

證 (1) 若 $F_X = F_Y$, 則 $\forall t \in \mathbb{R}$,

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx}\, dF_X(x) = \int_{\mathbb{R}} e^{itx}\, dF_Y(x) = \phi_Y(t).$$

(2) 若 $\phi_X = \phi_Y$, 令

$$C(F_X) = \{x \in \mathbb{R} \mid F_X \text{ 連續於點 } x\},$$
$$C(F_Y) = \{x \in \mathbb{R} \mid F_Y \text{ 連續於點 } x\}.$$

則由 Lévy 反演公式知, $\forall a, b \in C(F_X) \cap C(F_Y)$,

$$F_X(b) - F_X(a) = F_Y(b) - F_Y(a).$$

而 $C(F_X) \cap C(F_Y)$ 稠密於 $\mathbb{R}$, (參閱 Loève [11] p.178), 故可令 $a \to -\infty$, 得

$$F_X(b) = F_Y(b), \forall b \in C(F_X) \cap C(F_Y).$$

最後, $\forall x \in \mathbb{R}$, 自 $C(F_X) \cap C(F_Y)$ 中取序列 $\{b_n\}_n$ 使得 $b_n \downarrow x$, 則

$$\begin{aligned} F_X(x) &= \lim_{n \to +\infty} F_X(b_n), (\text{右連續性}) \\ &= \lim_{n \to +\infty} F_Y(b_n), \\ &= F_Y(x). \end{aligned}$$

□

本定理之目的在於說明: 我們固然可由密度函數或分配函數來分辨一隨機變數所具之分配, 但也可以由特徵函數來判定隨機變數之分配. 其次, 由本定理亦可看出分配函數集合與特徵函數集合具有一對射關係, 每一分配函數對應唯一之特徵函數, 反之亦然.

定理 5.10 (**Fourier 反演公式**)

設 X 爲一隨機變數, 若 $\int_{\mathbb{R}} |\phi_X(t)| dt < +\infty$, 則

(1) X 爲絕對連續型且其累積分配函數 F_X 爲可微;

(2) 函數 $f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}} e^{-itx} \phi_X(t)\, dt$ 爲 X 之一密度函數.

證 本定理需證: $\forall x \in \mathbb{R}$,

$$\lim_{h\to 0}\frac{F_X(x+h)-F_X(x)}{h}=\frac{1}{2\pi}\int_{-R} e^{-itx}\phi_X(t)\,dt.$$

其中需要利用一些實變函數論之定理及 Lévy 反演公式, 略之. 有興趣者, 請參閱 Laha & Rohagi [10], p.147. □

注意: 本定理說明, 如果 ϕ_X 爲絕對可積, 則 F_X 必爲可微, 因此知 X 爲絕對連續型, 且可利用反演公式, 由 ϕ_X 求出 X 之密度函數 f_X, 不過如果 X 爲絕對連續型, ϕ_X 未必爲絕對可積, 如 $X \sim U(\alpha,\beta)$, 本定理不再適用. 至於離散型隨機變數如何利用 ϕ_X 以求密度函數 f_X, 則需由下述定理來解決.

定理 5.11 （離散型反演公式）

設 X 爲一離散型隨機變數, 即存在至多可數集合 $A \subset \mathbb{R}$, 使得 $P\{X\in A\}=1$. 則

(1) $\forall x \in A$,

$$f_X(x)=\lim_{T\to+\infty}\frac{1}{2T}\int_{-T}^{T} e^{-itx}\phi_X(t)\,dt;$$

(2) 若 $A\subset\mathbb{Z}$, 則 $\forall x\in A$,

$$f_X(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itx}\phi_X(t)\,dt.$$

證 (1) 參閱 Chung [4], p.156.

(2) 由於 $\forall t\in\mathbb{R}$,

$$\phi_X(t)=E(e^{itX})=\sum_{j\in A}p_je^{itj},$$

內 $p_j=P\{X=j\}=f_X(j)$, 是以, x 固定 $\in A\subset\mathbb{Z}$,

$$e^{-itx}\phi_X(t)=\sum_{j\in A}p_je^{it(j-x)}=\sum_{j\in A\setminus\{x\}}p_je^{it(j-x)}+p_x,$$

故

$$\begin{aligned}\int_{-\pi}^{\pi} e^{-itx}\phi_X(t)\,dt &= \sum_{j\in A\setminus\{x\}}\int_{-\pi}^{\pi}p_je^{it(j-x)}dt+2\pi p_x\\ &=2\pi p_x,\ \ (\because j-x\neq 0, \text{故積分値爲 } 0)\\ &=2\pi f_X(x).\end{aligned}$$

□

例 1. 設 X 具 $B(n, p)$ 分配.

(1) 試求 X 之特徵函數 ϕ_X;

(2) 利用 ϕ_X 求 $E[X]$ 及Var X;

(3) 試以 ϕ_X 利用定理 5.11, 求 X 之密度函數.

$\boxed{解}$ (1) $\forall t \in \mathbb{R}$,

$$\begin{aligned}\phi_X(t) = E(e^{itX}) &= \sum_{x=0}^{n} e^{itx}\binom{n}{x}p^x q^{n-x} \\ &= \sum_{x=0}^{n}\binom{n}{x}(pe^{it})^x q^{n-x} = (pe^{it}+q)^n.\end{aligned}$$

利用電腦繪圖軟體, 我們可以繪出 ϕ_X 之值域的圖形如下:

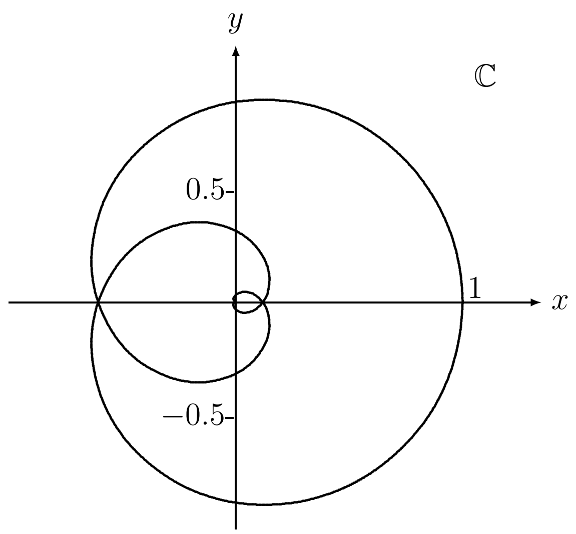


圖 5–2 $\phi(t) = (0.5\,e^{it} + 0.5)^{10}$ 之值域

(2) 由於

$$\phi'_X(t) = n(pe^{it}+q)^{n-1}ipe^{it},$$
$$\phi''_X(t) = n(n-1)(pe^{it}+q)^{n-2}i^2p^2e^{2it} + n(pe^{it}+q)^{n-1}i^2pe^{it}.$$

且已知 $E[X^2]$ 爲有限項之和, 必爲有限, 故

$$\begin{aligned}E[X] &= \frac{1}{i}\phi'_X(0) = \frac{1}{i}inp = np\ ; \\ E[X^2] &= \frac{1}{i^2}\phi''_X(0) = \frac{1}{i^2}[((n-1)np^2i^2 + i^2np] \\ &= n^2p^2 - np^2 + np\ .\end{aligned}$$

故

$$\text{Var } X = E[X^2] - (EX)^2 = -np^2 + np = npq\ .$$

(3) 由於 $A = \{x \in \mathbb{R} \mid f_X(x) > 0\} = \{0, 1, \cdots, n\} \subset \mathbb{Z}$, 利用離散型反演公式知, $\forall x \in A$,

$$\begin{aligned}
f_X(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itx}\phi_X(t)\,dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itx}\Big[\sum_{r=0}^{n}\binom{n}{r}(pe^{it})^r q^{n-r}\Big]\,dt \\
&= \frac{1}{2\pi}\sum_{r\in A\setminus\{x\}}\binom{n}{r}p^r q^{n-r}\int_{-\pi}^{\pi} e^{-it(x-r)}\,dt + \frac{1}{2\pi}\binom{n}{x}p^x q^{n-x}\int_{-\pi}^{\pi} dt \\
&= 0 + \binom{n}{x}p^x q^{n-x}, \text{ (上式第一項積分等於 0, 當 } x \neq r.)
\end{aligned}$$

□

例 2.

(1) 設 X 具 $N(0, 1)$ 分配. 試求 X 之特徵函數 ϕ_X ;

(2) 利用上述 ϕ_X 求 X 之密度函數;

(3) 設 Y 具 $N(\mu, \sigma^2)$ 分配. 試求 Y 之特徵函數 ϕ_Y ;

(4) 利用 ϕ_Y 求 EY 及Var Y .

解 (1) 為方便計, 令 $a = \dfrac{1}{\sqrt{2\pi}}$, $\phi = \phi_X$, 則 $\forall t \in \mathbb{R}$,

$$\begin{aligned}
\phi(t) &= E(e^{itX}) \\
&= a\int_{-\infty}^{+\infty} e^{itx}\exp\Big(-\frac{x^2}{2}\Big)\,dx \\
&= a\int_{-\infty}^{+\infty}\exp\Big(-\frac{x^2}{2}\Big)\cos tx\,dx + ai\int_{-\infty}^{+\infty}\exp\Big(-\frac{x^2}{2}\Big)\sin tx\,dx \\
&= a\int_{-\infty}^{+\infty}\exp\Big(-\frac{x^2}{2}\Big)\cos tx\,dx, \text{ (因第二項之被積分為奇函數)}
\end{aligned}$$

今將上述函數對 t 微分之, 則得

$$\begin{aligned}
\phi'(t) &= -a\int_{-\infty}^{+\infty} x\exp\Big(-\frac{x^2}{2}\Big)\sin tx\,dx \\
&= -at\int_{-\infty}^{+\infty}\exp\Big(-\frac{x^2}{2}\Big)\cos tx\,dx,
\end{aligned}$$

$$\Big(\text{分部積分法, 令 } u = \sin tx,\ dv = x\exp\Big(-\frac{x^2}{2}\Big)dx\Big),$$

亦即

$$\phi'(t) = -t\phi(t). \tag{$*$}$$

是以

$$\begin{aligned}(*)\ \Rightarrow\ & -t = \frac{\phi'(t)}{\phi(t)} = \frac{d}{dt}\ln\phi(t)\\ \Rightarrow\ & \ln\phi(t) = -\frac{t^2}{2} + c,\ (c\text{ 為一常數})\\ \Rightarrow\ & \phi(t) = K\exp\Big(-\frac{t^2}{2}\Big),\ (K\text{ 為一常數})\\ \Rightarrow\ & \phi(t) = \exp\Big(-\frac{t^2}{2}\Big),\ (\because\ \phi(0)=1)\end{aligned}$$

[註] 另法:

$$\begin{aligned}\phi(t) &= a\int_{-\infty}^{+\infty} e^{itx}\exp\Big(-\frac{x^2}{2}\Big)\,dx\\ &= a\int_{-\infty}^{+\infty}\exp\Big[-\frac{1}{2}(x^2-2itx+(it)^2)\Big]\exp\Big(\frac{1}{2}(it)^2\Big)\,dx\\ &\qquad\qquad(\text{再利用複數線積分之變數代換 } z=x-it,\ \ldots)\\ &= \exp\Big(-\frac{t^2}{2}\Big).\end{aligned}$$

(2) 由於

$$\int_{\mathbb{R}}|\phi_X(t)|dt = \int_{-\infty}^{+\infty}\exp\Big(-\frac{t^2}{2}\Big)dt = \sqrt{2\pi} < +\infty,$$

利用 Fourier 反演公式, $\forall x\in\mathbb{R}$,

$$f_X(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-itx}\phi_X(t)\,dt,$$

仿 (1) 之方法, 我們可證得

$$f_X(t) = \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{x^2}{2}\Big).$$

(3) $Y\sim N(\mu,\sigma^2)$, 設 $X=\dfrac{Y-\mu}{\sigma}\sim N(0,1)$, 是以

$$\begin{aligned}\phi_Y(t) &= \phi_{\sigma X+\mu}(t) = e^{it\mu}\phi_X(\sigma t)\\ &= \exp\Big(it\mu-\frac{1}{2}\sigma^2t^2\Big).\end{aligned}$$

(4) 由於

$$\begin{aligned}\phi'_Y(t) &= \exp\Big(it\mu-\frac{1}{2}\sigma^2t^2\Big)\cdot(i\mu-\sigma^2t)\\ \phi''_Y(t) &= \exp\Big(it\mu-\frac{1}{2}\sigma^2t^2\Big)\cdot(i\mu-\sigma^2t)^2-\sigma^2\exp\Big(it\mu-\frac{1}{2}\sigma^2t^2\Big)\end{aligned}$$

知 $\phi_Y''(0) = -\mu^2 - \sigma^2$ 爲有限, 由定理 5.7 之註, 吾人得知 EY 及 EY^2 皆存在且爲有限. 故

$$\begin{aligned} E[Y] &= \frac{1}{i}\phi_Y'(0) = \mu; \\ E[Y^2] &= \frac{1}{i^2}\phi_Y''(0) = \mu^2 + \sigma^2, \\ \text{Var } Y &= EY^2 - (EY)^2 = \sigma^2. \end{aligned}$$ □

除了特徵函數能生成各階動差之外, 還有二種函數, 其一爲動差生成函數, 另一爲階乘動差生成函數. 這三種函數均由高等分析學中之積分變換 (integral transformation) 改進而成, 以下我們將介紹這兩種函數. 但因其存在性並不穩定 (以下詳述), 是以重要性遠不如特徵函數.

定義 5.12

設 X 爲一實值隨機變數, 則實值之實變函數 $M_X(t) = E[e^{tX}]$ 稱爲 X 之**動差生成函數**(moment-generating function), 簡爲 m.g.f.

註: 若 $t \neq 0$, 則 m.g.f. $M_X(t) = E[e^{tX}]$ 未必爲有限, 是以 M_X 之定義域常因 X 之分配而改變. 此點有異於與特徵函數 $\forall t \in \mathbb{R},\ \phi_X(t) = E[e^{itX}]$ 均存在. (參見稍後之範例).

定理 5.13 (動差生成函數之性質)

設 X 爲一實值隨機變數, 則

(1) $M_X(0) = E[e^0] = 1$.

(2) 設 n 爲一自然數, M_X 之定義域包含原點之一鄰近且 X^n 爲 可積, 則 M_X 於原點爲 n 階可微且 $M_X^{(n)}(0) = E[X^n]$.

(3) $\forall t, M_{aX+b}(t) = e^{bt}M_X(at)$, 內 $a,\ b$ 爲常數.

證

(1) 易明.

(2) 仿定理 5.7 之證明, 讀者自證之.

(3) $M_{aX+b}(t) = E[e^{t(aX+b)}] = E[e^{bt}e^{atX}] = e^{bt}M_X(at)$. □

定義 5.14

設 X 爲一實值隨機變數, 則實值之實變函數 $\eta_X(t) = E[t^X]$ 稱爲 X 之**階乘動差生成函數**(factorial moment-generating function).

註: 若 $t \neq 1$, 則 $\eta_X(t) = E[t^X]$ 未必存在, 是以 η_X 之 定義域常因 X 之分配而改變.

定理 5.15 (階乘動差生成函數之性質)

設 X 爲一實值隨機變數, 則

(1) $\eta_X(1)=1$.

(2) 設 n 爲一自然數, 若 $X(X-1)\cdots(X-n+1)$ 爲可積, 則 η_X 於點 1 爲 n 階可微且 $\eta_X^{(n)}(1)=E[X(X-1)\cdots(X-n+1)]$.

證 (1) 易明.

(2) 仿定理 5.7 之證明, 讀者自證之. □

例 3. 設 X 具 $P(\lambda)$ 分配, 試求 ϕ_X, M_X 及 η_X.

解 (1) X 之特徵函數爲 $\phi_X\colon \mathbb{R}\to\mathbb{C}$:

$$\phi_X(t)=E[e^{itX}]=\sum_{x=0}^{+\infty}e^{itx}e^{-\lambda}\frac{\lambda^x}{x!}=e^{-\lambda}\sum_{x=0}^{+\infty}\frac{(\lambda e^{it})^x}{x!}=\exp(\lambda e^{it}-\lambda).$$

(2) X 之動差生成函數爲 $\forall t\in\mathbb{R}$,

$$M_X(t)=E[e^{tX}]=\sum_{x=0}^{+\infty}e^{tx}e^{-\lambda}\frac{\lambda^x}{x!}=e^{-\lambda}\sum_{x=0}^{+\infty}\frac{(\lambda e^{t})^x}{x!}=\exp(\lambda e^{t}-\lambda).$$

(3) X 之階乘動差生成函數爲 $\forall t\in(0,+\infty)$,

$$\eta_X(t)=E[t^X]=\sum_{x=0}^{+\infty}t^x e^{-\lambda}\frac{\lambda^x}{x!}=e^{-\lambda}\sum_{x=0}^{+\infty}\frac{(\lambda t)^x}{x!}=\exp(\lambda t-\lambda).$$ □

例 4. 設 X 具 $G(\alpha,\beta)$ 分配, 試求 ϕ_X 及 M_X.

解 (1) X 之特徵函數爲 $\phi_X\colon \mathbb{R}\to\mathbb{C}$:

$$\begin{aligned}\phi_X(t)&=E[e^{itX}]\\&=\frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty}e^{itx}x^{\alpha-1}e^{-x/\beta}dx\\&=\frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty}x^{\alpha-1}\exp\Big(-\frac{x(1-i\beta t)}{\beta}\Big)dx\\&=(1-i\beta t)^{-\alpha}\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}u^{\alpha-1}e^{-u}du,\ ^{[\star]}\\&=(1-i\beta t)^{-\alpha}.\end{aligned}$$

[⋆] 做變換 $u=\dfrac{x(1-i\beta t)}{\beta}, du=\dfrac{(1-i\beta t)\,dx}{\beta}$.

(2) 若 $t < 1/\beta$, 仿照 ϕ_X 我們可得

$$M_X(t) = E[e^{tX}] = (1-\beta t)^{-\alpha},\ \forall t < 1/\beta.$$

若 $t \geq 1/\beta$, 則因

$$\begin{aligned}
&\frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} e^{tx}x^{\alpha-1}e^{-x/\beta}dx \\
&\qquad = \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x^{\alpha-1}\exp\Big[x\Big(t-\frac{1}{\beta}\Big)\Big]dx \\
&\qquad \geq \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^{+\infty} x^{\alpha-1}dx \\
&\qquad = +\infty.
\end{aligned}$$

是以 $M_X\colon (-\infty,\, 1/\beta) \to \mathbb{R} : M_X(t) = (1-\beta t)^{-\alpha}$. □

註: 本例之之動差生成函數之定義域雖較 $\mathbb{R}$ 爲小, 但其已包含點 0 之一鄰近, 故仍可以之求 X 之諸動差.

例 5. 設 X 具 Laplace 分配, 即其密度函數爲 $f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \frac{1}{2}e^{-|x|}$, 試求其特徵函數 ϕ_X.

解 X 之特徵函數爲

$$\begin{aligned}
\phi_X(t) = E[e^{itX}] &= \frac{1}{2}\int_{-\infty}^{+\infty} e^{itx}e^{-|x|}\,dx \\
&= \frac{1}{2}\int_{-\infty}^{+\infty}(\cos tx + i\sin tx)\,e^{-|x|}\,dx \\
&= \frac{1}{2}\int_{-\infty}^{+\infty}(\cos tx)\,e^{-|x|}\,dx + \frac{i}{2}\int_{-\infty}^{+\infty}(\sin tx)\,e^{-|x|}\,dx \\
&= \int_0^{+\infty}(\cos tx)\,e^{-x}\,dx,\ \text{(第二項之被積函數爲奇函數)} \\
&= (-e^{-x}\cos tx)\Big|_{x=0}^{+\infty} - t\int_0^{+\infty}(\sin tx)\,e^{-x}\,dx,\ \text{(分部積分法)} \\
&= 1 - t^2\int_0^{+\infty}(\cos tx)\,e^{-x}\,dx,\ \text{(分部積分法)} \\
&= 1 - t^2\phi_X(t),
\end{aligned}$$

移項之, 得 $\phi_X(t) = \frac{1}{1+t^2}$. □

例 6. 設 X 具 Cauchy 分配 $C(0,1)$, 即其密度函數爲

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \frac{1}{\pi}\cdot\frac{1}{1+x^2}.$$

試求其特徵函數 ϕ_X.

解 直接由定義 $\phi_X(t) = E[e^{itX}]$ 以求 X 之特徵函數需利用微分方程, 稍爲困難些, 如果利用 Fourier 反演公式及上例, 則相當簡單. 爲方便 計, 設 Y 具 Laplace 分配, 則由上例知, $\phi_Y(t) = E[e^{itY}] = \frac{1}{1+t^2}$, 由於

$$\int_{\mathbb{R}} |\phi_Y(t)| dt = \pi < +\infty,$$

利用 Fourier 反演公式得

$$\frac{1}{2} e^{-|y|} = f_Y(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ity} \phi_Y(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ity}}{1+t^2} dt.$$

上式中, 先令 $t = -x$, 再令 $y = t$, 即得

$$\phi_X(t) = \int_{-\infty}^{+\infty} e^{itx} \cdot \frac{1}{\pi} \cdot \frac{dx}{1+x^2} = e^{-|t|}. \qquad \square$$

例 7. 設 X 之密度函數爲

$$f : \mathbb{R} \to [0,\, 1] : f(x) = \begin{cases} \dfrac{6}{\pi^2} \cdot \dfrac{1}{x^2}, & \text{若 } x \in \mathbb{N}, \\ 0, & \text{否則}. \end{cases}$$

(1) 試證: $\forall t > 0, M_X(t)$ 不存在;

(2) 試證: $\forall t > 1, \eta_X(t)$ 不存在.

證 在高等微積分中, 我們知道 $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, 是以函數 f 顯然爲一離散型之密度函數.

(1) 當 $t > 0$ 時, 因爲級數

$$\sum_{x=1}^{+\infty} e^{tx} f(x) = \frac{6}{\pi^2} \sum_{x=1}^{+\infty} \frac{e^{tx}}{x^2}$$

中之普遍項 $\dfrac{e^{tx}}{x^2}$ 趨近於 $+\infty$, 當 $x \to +\infty$. 因之, 級數爲發散, 是以 $M_X(t)$ 不存在.

(2) 同理, 當 $t > 1$ 時, 因爲級數

$$\sum_{x=1}^{+\infty} t^x f(x) = \frac{6}{\pi^2} \sum_{x=1}^{+\infty} \frac{t^x}{x^2}$$

中之普遍項 $\dfrac{t^x}{x^2}$ 趨近於 $+\infty$, 當 $x \to +\infty$. 因之, 級數爲發散, 故 $\eta_X(t)$ 不存在. $\square$

此一反例說明, 在討論機率之一般性理論時, 動差生成函數乃為『不可完全信賴』之工具. 此亦爲特徵函數在機率論諸積分變換中, 獨具重要性之理由.

§ 5.3 隨機向量之特徵函數

本節中, 我們將輕要介紹隨機向量之特徵函數及動差生成函數, 以及一些基本性質.

定義 5.16

設 $\mathbf{X} = (X_1, \cdots, X_k)$ 爲一 k 維隨機向量, 則函數

(1) $\phi_{\mathbf{X}} \colon \mathbb{R}^k \to \mathbb{C} : \phi_{\mathbf{X}}(\mathbf{t}) = E[\exp(it_1X_1 + \cdots + it_kX_k)]$ 稱爲 $\mathbf{X}$ 之**特徵函數**, 其中 $\mathbf{t} = (t_1, \cdots, t_k)$.

(2) $M_{\mathbf{X}}(\mathbf{t}) = E[\exp(t_1X_1 + \cdots + t_kX_k)]$ 稱爲 $\mathbf{X}$ 之 **動差生成函數**.

定理 5.17

設 $\mathbf{X}$ 爲一 k 維隨機向量.

(1) $\phi_{\mathbf{X}}(\mathbf{0}) = 1, M_{\mathbf{X}}(\mathbf{0}) = 1$.

(2) $\forall \mathbf{t} \in \mathbb{R}^k, |\phi_{\mathbf{X}}(\mathbf{t})| \leq 1$.

(3) 若 $c_1, \cdots, c_k \in \mathbb{R}, \mathbf{t} = (t_1, \cdots, t_k) \in \mathbb{R}^k$, 則

$$\phi_{X_1+c_1, \cdots, X_k+c_k}(\mathbf{t}) = \exp(it_1c_1 + \cdots + it_kc_k)\phi_{\mathbf{X}}(\mathbf{t});$$
$$\phi_{c_1X_1, \cdots, c_kX_k}(\mathbf{t}) = \phi_{\mathbf{X}}(c_1t_1 + \cdots + c_kt_k).$$

證 仿定理 5.5 證明即可. □

定理 5.18

設 ϕ 爲隨機向量 $\mathbf{X}$ 之特徵函數. 若$X_1^{n_1} \cdots X_k^{n_k}$ 爲可積, 內 $n_1, \cdots, n_k$ 均爲非負整數, 則

$$\frac{\partial^n}{\partial t_1^{n_1} \cdots \partial t_k^{n_k}} \phi(\mathbf{0}) = i^n E[X_1^{n_1} \cdots X_k^{n_k}],$$

其中 $n = n_1 + \cdots + n_k$.

證 仿定理 5.7 之證. □

Lévy 及 Fourier 反演公式亦可推廣到多維之情形, Laha & Rohagi [10], p.148 及其後有很詳細的說明, 我們將不予介紹.

例 1. 設 $\mathbf{X} = (X_1, \cdots, X_k)$ 具多項分配. 試求

(1) $\mathbf{X}$ 之特徵函數;

(2) $E[X_1 \cdots X_k]$.

$\boxed{\text{解}}$ (1) 我們知道 $\mathbf{X}$ 之密度函數爲 $f\colon \mathbb{R}^k \to [0,\,1]$:

$$f(x_1,\,\cdots,\,x_k) = \begin{cases} \dbinom{n}{x_1,\,\cdots,\,x_k} p_1^{x_1}\cdots p_k^{x_k}, & \text{若 } \mathbf{x} \in A, \\ 0, & \text{否則}. \end{cases}$$

內 $A = \{(x_1,\,\cdots,\,x_k) \mid x_j \in \mathbb{Z}_+, x_1 + \cdots + x_k = n\}$, 是以

$$\begin{aligned} \phi_{\mathbf{X}}(\mathbf{t}) &= E[\exp(it_1X_1 + \cdots + it_kX_k)] \\ &= \sum_{x\in A} \exp(it_1x_1 + \cdots + it_kx_k) f(x_1,\,\cdots,\,x_k) \\ &= \sum_{x\in A} \binom{n}{x_1,\,\cdots,\,x_k} (p_1e^{it_1})^{x_1} \cdots (p_ke^{it_k})^{x_k} \\ &= (p_1e^{it_1} + \cdots + p_ke^{it_k})^n. \end{aligned}$$

(2) 由於,

$$\begin{aligned} \frac{\partial \phi_{\mathbf{X}}(\mathbf{t})}{\partial t_1} &= n(p_1e^{it_1} + \cdots + p_ke^{it_k})^{n-1} ip_1e^{it_1}, \\ &\vdots \\ \frac{\partial^k \phi_{\mathbf{X}}(\mathbf{t})}{\partial t_k \cdots \partial t_1} &= n(n-1)\cdots(n-k+1) \\ &\qquad \cdot (p_1e^{it_1} + \cdots + p_ke^{it_k})^{n-k}\, i^k\, p_1 \cdots p_k\, e^{it_1+\cdots+it_k}, \end{aligned}$$

是以

$$\begin{aligned} E[X_1 \cdots X_k] &= \frac{1}{i^k} \cdot \frac{\partial^k \phi_{\mathbf{X}}(\mathbf{0})}{\partial t_k \cdots \partial t_1} \\ &= n(n-1)\cdots(n-k+1)p_1\cdots p_k. \end{aligned}$$

$\square$

第五章 習 題

5-1 設 X 與 Y 均爲實值隨機變數, $Z = X + iY$. 試證:

$$P\{|Z| > 1\} \geq P\{|X| > 1\} + P\{|Y| > 1\} - P\{|X| > 1, |Y| > 1\}.$$

5-2 設 X 爲一非負隨機變數, $m \in \mathbb{R}, Z = X + imX$. 試證: $|E[Z]| = E|Z|$.

5-3 試求下列各隨機變數 X 之特徵函數, 並以之求 $E[X]$ 及 $\operatorname{Var} X$.

(1) X 具均匀分配 $U(\alpha, \beta)$;

(2) X 具負二項分配;

(3) X 之密度函數爲

$$f\colon \mathbb{R} \to \mathbb{R}_+ : f(x) = \begin{cases} \lambda \exp(-\lambda(x-\alpha)), & \text{若 } x > \alpha, \\ 0, & \text{否則}, \end{cases}$$

內 $\lambda > 0, \alpha \in \mathbb{R}$.

5-4 若隨機變數 X 具一般之 Cauchy 分配 $C(\mu, \sigma^2)$, 試求其特徵函數, 並求此特徵函數在點 0 之左右導數.

5-5 設 X 之特徵函數爲

$$\phi\colon \mathbb{R} \to \mathbb{C} : \phi(t) = \begin{cases} 1-t, & \text{若 } 0 \leq t \leq 1, \\ 1+t, & \text{若 } -1 \leq t < 0, \\ 0, & \text{若 } |t| > 1. \end{cases}$$

試利用反演定理以求 X 之密度函數.

5-6 設 ϕ 爲隨機變數 X 之特徵函數, 試證以下三陳述爲對等:

(I) $\forall t \in \mathbb{R}, \phi(t) \in \mathbb{R}$;

(II) ϕ 爲一偶函數, 即 $\phi(-t) = \phi(t)$;

(III) X 與 $-X$ 之分配相同, (此時稱 X具對稱分配).

5-7 設 X 具常態分配 $N(\mu, \sigma^2)$, 試求 X 之動差生成函數.

5-8 設隨機變數 X 之期望值 $E[X] = 0$. 試證:

$$|\phi_X(t) - 1| \leq \frac{t^2 E[X^2]}{2}, \ \forall t \in \mathbb{R}.$$

5-9 設 X 之特徵函數爲 $\phi_X(t) = \exp(4(e^{it} - 1))$. 試求 $P\{1 \leq X \leq 5\} =?$

5-10 設隨機向量 (X, Y) 之特徵函數爲

$$\phi\colon \mathbb{R}^2 \to \mathbb{C} : \phi(t, u) = \left[\frac{1}{3}(e^{it+iu} + 1) + \frac{1}{6}(e^{it} + e^{iu})\right]^2.$$

若已知 $E[X]$, $\operatorname{Var} X$ 及 $C(X, Y)$ 均爲有限, 試求其值.

5-11 設 $X \sim N(0, 1)$, 直接利用特徵函數之定義以求 $Y = X^2$ 之特徵函數, 並以之說明 Y 具 χ_1^2 分配.

5-12 設 ϕ 爲隨機變數 X 之特徵函數, 若 ϕ 爲 n 階可微, 試利用 微積分中之 Taylor 定理證明:

$$\phi(t) = \phi(0) + \frac{\phi'(0)}{1!}t + \cdots + \frac{\phi^{(n)}(0)}{n!}t^n + o(t^n).$$

5-13 設絕對連續型隨機變數 X 之分配函數爲 F. 試利用特徵函數法證明: 隨機 變數 $Y = F \circ X$ 具 $U(0, 1)$ 分配.

[註] 請參照習題 2-20 題.

Chapter 6

隨機變數之獨立性

在第一章中, 我們曾以條件 $P(A \cap B) = P(A)P(B)$ 界定事件 A 與 B 之獨立性, 由於

$$A \perp\!\!\!\perp B \Leftrightarrow A \perp\!\!\!\perp \complement B \Leftrightarrow \complement A \perp\!\!\!\perp B \Leftrightarrow \complement A \perp\!\!\!\perp \complement B,$$

我們不難看出 A 與 B 爲獨立之充要條件爲 $\sigma\{A\}$ 中之每一事件與 $\sigma\{B\}$ 中之每一事件均爲獨立. 由這裡讓我們聯想到二"互不關聯"的隨機試驗, 是否能以嚴格的數學形式來表達? 它的關鍵在於由此二隨機試驗所產生的二隨機變數, 我們該如何界定其獨立性? 此一觀念影響了整個機率論發展史, 在 J. Bernoulli 之後的二百年中, 機率學者花費了很大的心血在研究獨立隨機變數之平均的"規則性", 也就是所謂的大數定律及中央極限定理, 我們將在下一章詳細討論. 本章中, 我們先界定所謂獨立性, 並探討一些重要性質, 以做爲下一章之準備.

§ 6.1 隨機獨立

定義 6.1

設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, $\mathcal{C}_1, \cdots, \mathcal{C}_k \subset \mathcal{F}$, 如果

$$\forall A_1 \in \mathcal{C}_1, \cdots, \forall A_k \in \mathcal{C}_k, \ A_1, \cdots, A_k \text{ 爲獨立},$$

則稱 $\mathcal{C}_1, \cdots, \mathcal{C}_k$ 爲(**隨機**) **獨立**.

定義 6.2

設 $X_1, \cdots, X_k$ 均爲 $(\Omega, \mathcal{F}, P)$ 之隨機變數. 我們稱 $X_1, \cdots, X_k$ 爲獨立, 如果 $X_1^{-1}(\mathcal{B}), \cdots, X_k^{-1}(\mathcal{B})$ 爲獨立, 內 $\mathcal{B}$ 爲一維 Borel 域.

一如本章一開始所說, 此一定義來自事件間之獨立性, 論述十分流暢, 含義也很清楚, 形式也是十分嚴格, 但使用起來卻非常困難; 因爲二隨機變數 X 與 Y 所界定之 σ 域 $X^{-1}(\mathcal{B})$ 與

$Y^{-1}(\mathcal{B})$ 何等複雜, 我們又如何能判定一 σ 域之每一事件與另一 σ 域之每一事件均爲獨立? 因此, 有加以改善的必要. 在本章中, 我們將先討論以下各陳述之對等情形, 其間之證明, 有的簡單, 有的則是十分困難.

(I) $X_1, \cdots, X_k$ 爲獨立;

(II) $\forall A_1 \in X_1^{-1}(\mathcal{B}), \cdots, \forall A_k \in X_k^{-1}(\mathcal{B}), P\Big(\bigcap_{j=1}^{k} A_j\Big) = \prod_{j=1}^{k} P(A_j)$;

(III) $\forall B_1, \cdots, \forall B_k \in \mathcal{B}, P\Big(\bigcap_{j=1}^{k}\{X_j \in B_j\}\Big) = \prod_{j=1}^{k} P\{X_j \in B_j\}$;

(IV) $\forall (x_1, \cdots, x_k) \in \mathbb{R}^k,\ F(x_1, \cdots, x_k) = F_1(x_1)\cdots F_k(x_k)$;

(V) $\forall (x_1, \cdots, x_k) \in \mathbb{R}^k,\ f(x_1, \cdots, x_k) = f_1(x_1)\cdots f_k(x_k)$;

(VI) $\forall (t_1, \cdots, t_k) \in \mathbb{R}^k,\ \phi(t_1, \cdots, t_k) = \phi_1(t_1)\cdots \phi_k(t_k)$.

其中 $F,\ f,\ \phi$ 分別爲 $\mathbf{X} = (X_1, \cdots, X_k)$ 之分配函數、密度函數及特徵函數 ; $F_j,\ f_j,\ \phi_j$ 則分別爲 X_j 之分配函數、密度函數及特徵函數. 我們知道密度函數僅在絕對連續型或離散型之隨機向量 (隨機變數) 才存在, 因此, 上述陳述 (V) 與其他陳述之對等需在上述二型之情況下方爲成立, 此外, 絕對連續型之密度函數爲殆遍唯一 (a.e. unique), 我們將在稍後詳細說明. 由於上述各陳述之對等性之證明十分困難, 許多機率論及統計學之書籍均予省略, 而以陳述 (V) 做爲隨機變數獨立之定義, 其優點爲簡單易用, 缺點則是定義之由來不明. 我們建議初學之讀者不妨大致瀏覽, 先學會如何使用, 再回頭詳細探討其由來.

定理 6.3

設 $X_1, \cdots, X_k$ 均爲 $(\Omega, \mathcal{F}, P)$ 之隨機變數, 則以下三陳述爲對等:

(I) $X_1, \cdots, X_k$ 爲獨立;

(II) $\forall A_1 \in X_1^{-1}(\mathcal{B}), \cdots, \forall A_k \in X_k^{-1}(\mathcal{B}),\ P\Big(\bigcap_{j=1}^{k} A_j\Big) = \prod_{j=1}^{k} P(A_j)$;

(III) $\forall B_1, \cdots, \forall B_k \in \mathcal{B},\ P\Big(\bigcap_{j=1}^{k}\{X_j \in B_j\}\Big) = \prod_{j=1}^{k} P\{X_j \in B_j\}$.

證 (I) $\Rightarrow$ (II). 由定義 6.1 立即可得.

(II) $\Rightarrow$ (I). 我們需證:

$$A_1 \in X_1^{-1}(\mathcal{B}), \cdots, A_k \in X_k^{-1}(\mathcal{B}) \ \Rightarrow\ A_1, \cdots, A_k \text{ 爲獨立.}$$

我們只證明: $P(A_1 \cap A_2) = P(A_1)P(A_2)$, 其餘部分同理. 在此, 令 $A_3' = \cdots = A_k' = \Omega$, 則因 $A_3' \in X_3^{-1}(\mathcal{B}), \cdots A_k' \in X_k^{-1}(\mathcal{B})$, 應有

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1 \cap A_2 \cap A_3' \cap \cdots \cap A_k') \\ &= P(A_1)P(A_2)P(A_3') \cdots P(A_k') \\ &= P(A_1)P(A_2). \end{aligned}$$

(II) $\Rightarrow$ (III). 因爲 $\{X_j \in B_j\} = X_j^{-1}(B_j) \in X_j^{-1}(\mathcal{B})$, 是以

$$P\Big(\bigcap_{j=1}^{k}\{X_j \in B_j\}\Big) = \prod_{j=1}^{k} P\{X_j \in B_j\};$$

(III) $\Rightarrow$ (II). 由於 $A_j \in X_j^{-1}(\mathcal{B})$, 知必存在 $B_j \in \mathcal{B}$ 使得 $A_j = X_j^{-1}(B_j)$. 故

$$\begin{aligned} P\Big(\bigcap_{j=1}^{k} A_j\Big) &= P\Big(\bigcap_{j=1}^{k} X_j^{-1}(B_j)\Big) = P\Big(\bigcap_{j=1}^{k}\{X_j \in B_j\}\Big) \\ &= \prod_{j=1}^{k} P\{X_j \in B_j\} = \prod_{j=1}^{k} P(A_j). \end{aligned}$$ □

在沒有繼續研究其他之對等關係之前, 我們將先探討隨機變數序列之獨立性, 在下一章中, 它將扮演十分重要之角色.

定義 6.4

設 $\{X_n\}_n$ 爲 $(\Omega, \mathcal{F}, P)$ 之一隨機變數序列. 若 $\forall n \in \mathbb{N}, X_1, \cdots, X_n$ 爲獨立, 則稱 $X_1, \cdots, X_n, \cdots$ 爲**獨立**.

定理 6.5

設 $X_1, \cdots, X_k$ 爲獨立之隨機變數, $g_j\colon \mathbb{R} \to \mathbb{R}$ 爲 Borel 可測函數, 其中 $j \in \{1, \cdots, k\}$. 則隨機變數 $g_1 \circ X_1, \cdots, g_k \circ X_k$ 爲獨立.

證 利用定理 6.3 之陳述 (III), 往證: 若 $B_1, \cdots, B_k \in \mathcal{B}$, 則

$$P\Big(\bigcap_{j=1}^{k}\{g_j \circ X_j \in B_j\}\Big) = \prod_{j=1}^{k} P\{g_j \circ X_j \in B_j\}.$$

此因

$$
\begin{aligned}
P\Big(\bigcap_{j=1}^{k}\{g_j\circ X_j\in B_j\}\Big) &= P\Big(\bigcap_{j=1}^{k}\{X_j\in g_j^{-1}(B_j)\}\Big),\ \text{(利用合成函數之像原)}\\
&= \prod_{j=1}^{k} P\{X_j\in g_j^{-1}(B_j)\},\ (\text{因 } X_1,\cdots,X_k \text{爲獨立})\\
&= \prod_{j=1}^{k} P\{g_j\circ X_j\in B_j\}. \qquad \square
\end{aligned}
$$

在古典數學中, 常將合成函數 $g_j\circ X_j$ 讀做 X_j 之變換 (transformation), 並寫爲 $g_j(X_j)$; 本定理乃是說明: 如果 $X_1,\cdots,X_k$ 為獨立, 則諸隨機變數之變換亦為獨立. 本定理可加以推廣而爲: 若 $X_1,\cdots,X_k$ 爲獨立, 則 $g(X_1,\cdots,X_n)$ 與 $h(X_{n+1},\cdots,X_k)$ 亦爲獨立, 其中 $1<n<k$; 當然我們還可將 $X_1,\cdots,X_k$ "切"爲兩"組"以上, 以下就是這樣的定理.

定理 6.6

設 $X_1,\cdots,X_k$ 爲獨立之隨機變數. 若

$$g_1\colon \mathbb{R}^{n_1}\to\mathbb{R},\ g_2\colon \mathbb{R}^{n_2-n_1}\to\mathbb{R},\cdots,g_{j+1}\colon \mathbb{R}^{k-n_j}\to\mathbb{R}$$

均爲 Borel 可測, 其中 $1\le n_1<n_2<\cdots<n_j<k$, 則隨機變數

$$g_1\circ(X_1,\cdots,X_{n_1}),\ g_2\circ(X_{n_1+1},\cdots,X_{n_2}),\cdots,\ g_{j+1}\circ(X_{n_j+1},\cdots,X_k)$$

亦爲獨立.

證 本定理證明要比定理 6.5 之證明困難得多, 而且需要一些有關 σ 域的新技巧. 有興趣者, 請參閱 Chow & Teicher [2], p.62. $\square$

在尋求獨立性之進一步的對等條件之前, 我們將先研究一下獨立性與期望值的一些關係, 這些關係對於陳述 (VI), 有關特徵函數與隨機變數之獨立性, 成爲相當重要的工具.

定理 6.7

設非負隨機變數 X 與 Y 爲獨立, 則 $E[XY]=EX\cdot EY$.

證 1° 若 X, Y 均爲簡單隨機變數, 其值均分別爲 $\{x_1, \cdots, x_n\}$, $\{y_1, \cdots, y_m\}$, 令 $A_j = \{X = x_j\}, \forall j = 1, \cdots, n$; $B_k = \{Y = y_k\}, \forall k = 1, \cdots, m$, 則

- $X = \sum_{j=1}^{n} x_j I_{A_j}$, $Y = \sum_{k=1}^{m} y_k I_{B_k}$;
- $XY = \sum_{j=1}^{n} \sum_{k=1}^{m} x_j y_k I_{A_j \cap B_k}$
- $\forall j$, $\forall k$, A_j 與 B_k 爲獨立;

是以

$$\begin{aligned} E[XY] &= \sum_{j=1}^{n} \sum_{k=1}^{m} x_j y_k P(A_j \cap B_k) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{m} x_j y_k P(A_j) P(B_k) \\ &= \sum_{j=1}^{n} x_j P(A_j) \cdot \sum_{k=1}^{m} y_k P(B_k) = E[X] \cdot E[Y]. \end{aligned}$$

2° 若 X, Y 均爲非負隨機變數, 由第二章之可測性定理知, 存在非負簡單隨機變數序列 $\{X_n\}_n, \{Y_n\}_n$ 使得

$$X_n \uparrow X, \quad Y_n \uparrow Y;$$

由於 X_n 爲 X 之一變換, (參見附錄四), Y_n 爲 Y 之一變換, 是以 $\forall n \in \mathbb{N}, X_n$ 與 Y_n 爲獨立. 此外, $\{X_n Y_n\}_n$ 爲非負簡單隨機變數且 $X_n Y_n \uparrow XY$, 是以

$$E[XY] = \lim_{n \to +\infty} E[X_n Y_n] = \lim_{n \to +\infty} E[X_n] \cdot E[Y_n] = E[X] \cdot E[Y]. \qquad \square$$

系 6.8

(1) 若 X 與 Y 爲獨立之可積隨機變數, 則 $E[XY] = EX \cdot EY$.

(2) 若 $X_1, \cdots, X_k$ 爲獨立之可積隨機變數, 則

$$E[X_1 \cdots X_k] = (EX_1) \cdots (EX_k).$$

證 (1) 令

$$g_1 : \mathbb{R} \to \mathbb{R}_+ : g_1(x) = \begin{cases} x, & \text{若 } x \geq 0, \\ 0, & \text{若 } x < 0, \end{cases}$$

$$g_2 : \mathbb{R} \to \mathbb{R}_+ : g_2(x) = \begin{cases} 0, & \text{若 } x \geq 0, \\ -x, & \text{若 } x < 0, \end{cases}$$

顯然 $X^+ = g_1 \circ X, X^- = g_2 \circ X,\ Y^+ = g_1 \circ Y,\ Y^- = g_2 \circ Y$, 由定理 6.5 知, X^+ 與 Y^+、X^+ 與 Y^-、X^- 與 Y^+、X^- 與 Y^- 等均爲獨立. 是以

$$\begin{aligned}
E[XY] &= E[(X^+ - X^-)(Y^+ - Y^-)] \\
&= E[X^+Y^+ - X^+Y^- - X^-Y^+ + X^-Y^-] \\
&= E[X^+Y^+] - E[X^+Y^-] - E[X^-Y^+] + E[X^-Y^-] \\
&\qquad\qquad\qquad\qquad (\text{因 } X,\ Y \text{ 均爲可積}) \\
&= E[X^+]E[Y^+] - E[X^+]E[Y^-] - E[X^-]E[Y^+] + E[X^-]E[Y^-] \\
&= (EX^+ - EX^-)(EY^+ - EY^-) \\
&= E[X] \cdot E[Y].
\end{aligned}$$

(2) 利用 (1) 及數學歸納法, 讀者自證之. □

系 6.9

> 若 X 與 Y 爲獨立之隨機變數, 則 X 與 Y 爲不相關, (如果相關係數存在).

證 因爲

$$\rho(X,Y) = \frac{C(X,Y)}{\sigma_X\sigma_Y} = \frac{E[XY] - (EX)(EY)}{\sigma_X\sigma_Y} = 0.$$ □

本系之逆陳述不爲眞, 亦即, 若 X 與 Y 為不相關, 則 X 與 Y 未必為獨立. 反例稍後再舉.

設 $X_1, \cdots, X_k$ 爲隨機變數, 我們知道, 不論其是否爲獨立, 其線性組合 $Z = \sum_{j=1}^k c_jX_j$ 之期望值 EZ 爲諸 EX_j 之線性組合 $\sum_{j=1}^k c_jE[X_j]$; 然而 Z 之變異數 Var Z 是否爲諸變異數 Var X_j 之線性組合? 答案是:『未必！』, 但當 $X_1, \cdots, X_k$ 爲獨立時, 答案則肯定, 但應注意線性組合之『係數』, 這就是 Bienaymé 公式.

定理 6.10 (**Bienaymé 定理**)

> 設隨機變數 $X_1, \cdots, X_k$ 之變異數分別爲 $\sigma_1^2 = \mathrm{Var}\ X_1, \cdots, \sigma_k^2 = \mathrm{Var}\ X_k$; 而 $c_1, \cdots, c_k \in \mathbb{R}$, 則
>
> (1) $\mathrm{Var}\left(\sum_{j=1}^k c_jX_j\right) = \sum_{j=1}^k c_j^2\sigma_j^2 + \sum_{i\neq j} c_ic_jC(X_i, X_j)$;
>
> (2) $X_1, \cdots, X_k$ 爲獨立
>
> $$\begin{aligned} &\Rightarrow X_i \text{ 與 } X_j \text{爲獨立}, \forall i, j \in \{1, \cdots, k\} \text{ 且 } i \neq j \\ &\Rightarrow \mathrm{Var}\left(\sum_{j=1}^k c_jX_j\right) = \sum_{j=1}^k c_j^2\,\mathrm{Var}\ X_j. \qquad (*) \end{aligned}$$
>
> $(*)$ 此一等式稱爲 **Bienaymé 公式**.

證 (1) 首先, 我們知道 $\sum_{i\neq j}$ 表示 $\displaystyle\sum_{j=1}^{k}\sum_{\substack{i=1\\ i\neq j}}^{k}$, 更明白地說應該是 $\displaystyle\sum_{(i,j)\in B}$ 其中

$$B=\{(i,j)\in\mathbb{N}^2 \mid 1\le i\le k, 1\le j\le k, \text{且 } i\neq j\},$$

由變異數之定義知

$$\begin{aligned}
\operatorname{Var}\Big(\sum_{j=1}^{k} c_jX_j\Big) &= E\Big[\sum_{j=1}^{k} c_jX_j - E\Big(\sum_{j=1}^{k} c_jX_j\Big)\Big]^2\\
&= E\Big[\sum_{j=1}^{k} c_j(X_j-EX_j)\Big]^2\\
&= E\Big[\sum_{j=1}^{k} c_j^2(X_j-EX_j)^2+\sum_{i\neq j} c_ic_j(X_i-EX_i)(X_j-EX_j)\Big]\\
&= \sum_{j=1}^{k} c_j^2\sigma_j^2+\sum_{i\neq j} c_ic_j\, C(X_i,X_j).
\end{aligned}$$

(2) 若 $X_1, \cdots, X_k$ 爲獨立, 由定義 6.2 知其中任二隨機 變數均爲獨立, 是以其協方差亦爲 0, Bienaymé 公式顯然成立. □

例 1. 某甲步行之速率每公里需 X 分鐘, 某乙步行之速率則需 Y分鐘, 已知 $EX=15$, $\operatorname{Var} X=1$, $EY=12$, $\operatorname{Var} Y=2$. 今甲自點 A 持一信件走到點 B 交給乙, 然後乙立即持此信走到點 C. 若 AB 之距離爲 2 公里, BC 之距離爲 3 公里, 試問此一信件在 56 分鐘至 76 分鐘之間送達點 C 之機率是否大於 0.7 ?

解 由題意知 X 與 Y 爲獨立之隨機變數, 由於 $AB=2$, $BC=3$, 故由點 A 送信至點 C 共需 $Z=2X+3Y$ 分鐘. 利用期望值之線性性質及 Bienaymé 公式知,

$$EZ=2EX+3EY=2\times15+3\times12=66,$$
$$\operatorname{Var} Z=2^2\operatorname{Var} X+3^2\operatorname{Var} Y=4\times1+9\times2=22.$$

是以

$$\begin{aligned}
P\{56<Z<76\} &= P\{-10<Z-66<10\}\\
&= P\{|Z-66|<10\}\\
&= 1-P\{|Z-66|\ge 10\}\\
&\ge 1-\frac{\operatorname{Var} Z}{10^2}, \text{ (Chebyshev 不等式)}\\
&= 1-\frac{22}{100}=0.78.
\end{aligned}$$

顯然此一信件在 56 分鐘 至 76 分鐘之間送達點 C 之機率大於 0.7 . □

提醒讀者注意: 在上例中, 我們利用 Chebyshev 不等式求隨機變數 Z 發生在某一區間之機率的『範圍』. 如果希望確知此一機率之『值』爲若干, 首先隨機變數 X與 Y 之分配須爲已知, 其次設法求得 Z 之分配, 我們將在本章第二節, 以特徵函數方法探討此一問題. 當然, 利用第三章隨機向量之變換也是一種可選擇之方法, 不過通常遠較特徵函數法麻煩.

以下, 我們將利用分配函數、密度函數及特徵函數以研究隨機變數獨立性之其他對等條件, 這對於今後處理獨立性問題大有幫助.

定理 6.11

設 X 與 Y 爲隨機變數, $F = F_{X,Y}$ 爲隨機向量 (X, Y) 之分配函數, 則 X 與 Y 爲獨立之充要條件爲 $F(x, y) = F_X(x)F_Y(y), \forall x,\ \forall y \in \mathbb{R}$.

證 若 X 與 Y 爲獨立. 則 $\forall x,\ \forall y \in \mathbb{R}$,

$$\begin{aligned} F(x, y) &= P\{X \leq x, Y \leq y\} \\ &= P\{X \leq x\}P\{Y \leq y\},\ (\text{定理 } 6.3) \\ &= F_X(x)F_Y(y). \end{aligned}$$

反之, 若 $F(x, y) = F_X(x)F_Y(y), \forall x,\ \forall y \in \mathbb{R}$. 我們需證明:

$$\forall B,\ C \in \mathcal{B}, P\{X \in B,\ Y \in C\} = P\{X \in B\}P\{Y \in C\}.$$

由於此一證明需要一些有關單調類 (monotone class) 的定理, 對於初學機率者 似嫌過深. 我們建議有興趣者參閱 Laha & Rohagi [10], §1.4. □

系 6.12

設 $X_1, \cdots, X_k$ 爲隨機變數, F 爲 $\mathbf{X} = (X_1, \cdots, X_k)$ 之分配函數, F_j 爲 X_j 之分配函數, 則以下二陳述爲對等.

(I) $X_1, \cdots, X_k$ 爲獨立;

(IV) $\forall (x_1, \cdots, x_k) \in \mathbb{R}^k,\ F(x_1, \cdots, x_k) = F_1(x_1) \cdots F_k(x_k)$.

證 (I) $\Rightarrow$ (IV). 仿定理 6.11即可, 至於其逆則更爲繁複. □

定理 6.13

設 $X_1, \cdots, X_k$ 爲隨機變數; f 爲 $\mathbf{X} = (X_1, \cdots, X_k)$ 之密度函數, f_j 爲 X_j 之密度函數, $\forall j = 1, \cdots, k$.

(1) 若 $\mathbf{X}$ 爲離散型, 則 $X_1, \cdots, X_k$ 爲獨立之充要條件爲

$$f(x_1, \cdots, x_k) = f_1(x_1) \cdots f_k(x_k), \forall (x_1, \cdots, x_k) \in \mathbb{R}^k;$$

(2) 若 $\mathbf{X}$ 爲絕對連續型, 則 $X_1, \cdots, X_k$ 爲獨立之充要條件爲

$$f(x_1, \cdots, x_k) = f_1(x_1) \cdots f_k(x_k), \forall (x_1, \cdots, x_k) \in \mathbb{R}^k \text{ a.e.}$$

說明: 本定理之目的, 在於證明本章開始時所提及陳述 (I) 與陳述 (V) 之對等性, 由於絕對連續型隨機變數之密度函數不爲唯一而爲殆遍唯一 (a.e. unique), 故離散連續二型之陳述有稍許不同, 提醒讀者小心.

證 爲使讀者易於了解, 我們將只證 $k = 2$ 之情形, 其餘同理.

(1) 若 X_1 與 X_2 爲離散型. 令 $B_1 = \{x_1\}, B_2 = \{x_2\}$, 則

$$\begin{aligned} X_1 \text{ 與 } X_2 \text{ 爲獨立} &\Rightarrow P\{X_1 \in B_1,\ X_2 \in B_2\} = P\{X_1 \in B_1\} P\{X_2 \in B_2\} \\ &\Rightarrow P\{X_1 = x_1,\ X_2 = x_2\} = P\{X_1 = x_1\} P\{X_2 = x_2\} \\ &\Rightarrow f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2). \end{aligned}$$

反之, 若 $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2),\ \forall (x_1, x_2) \in \mathbb{R}^2$, 則因

$$\begin{aligned} F(x_1, x_2) &= \sum_{t_1 \le x_1} \sum_{t_2 \le x_2} f(t_1, t_2) \\ &= \sum_{t_1 \le x_1} f_1(t_1) \sum_{t_2 \le x_2} f_2(t_2) = F_1(x_1) \cdot F_2(x_2). \end{aligned}$$

由定理 6.11 知 X_1 與 X_2 爲獨立.

(2) 若 X_1 與 X_2 爲絕對連續型, 則

$$\begin{aligned} X_1 \text{ 與 } X_2 \text{ 爲獨立} &\Rightarrow F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2), \text{ (定理 6.11)} \\ &\Rightarrow f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2) \quad \text{a.e.} \end{aligned}$$

(上式兩端分別對 x_1 及 x_2 偏微之.)

反之, 若 $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$ a.e., 則因

$$\begin{aligned} F(x_1, x_2) &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(t_1, t_2)\, dt_2\, dt_1 \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_1(t_1) f_2(t_2)\, dt_2\, dt_1 \\ &= \int_{-\infty}^{x_1} f_1(t_1)\, dt_1 \int_{-\infty}^{x_2} f_2(t_2)\, dt_2 = F_1(x_1) \cdot F_2(x_2), \end{aligned}$$

是以 X_1 與 X_2 爲獨立. □

系 6.14

設 X 與 Y 爲獨立之隨機變數, f, f_X, f_Y 分別爲 (X,Y)、X 及 Y 之密度函數, 集合

$$\begin{aligned} A &= \{(x,y) \in \mathbb{R}^2 \mid f(x,y)) > 0\}, \\ A_1 &= \{x \in \mathbb{R} \mid f_X(x) > 0\}, \\ A_2 &= \{y \in \mathbb{R} \mid f_Y(y) > 0\}. \end{aligned}$$

(1) 若 X 與 Y 爲離散型, 則 $A = A_1 \times A_2$;

(2) 若 X 與 Y 爲絕對連續型, 則 $A = A_1 \times A_2$ a.e.

證 我們只證 (1). 至於 (2), 留給讀者自行完成之. 由於

$$\begin{aligned} (x,y) \in A &\Leftrightarrow f(x,y) > 0 \\ &\Leftrightarrow f_X(x) f_Y(y) > 0 \\ &\Leftrightarrow f_X(x) > 0 \text{ 且 } f_Y(y) > 0 \\ &\Leftrightarrow x \in A_1 \text{ 且 } y \in A_2 \\ &\Leftrightarrow (x,y) \in A_1 \times A_2 \end{aligned}$$

是以 $A = A_1 \times A_2$. □

例 2. 設隨機向量 (X,Y) 之分配函數爲

$$F\colon \mathbb{R}^2 \to [0,\, 1] : F(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y)}, & \text{若 } x \geq 0,\ y \geq 0, \\ 0, & \text{否則.} \end{cases}$$

試問 X 與 Y 是否爲獨立 ?

解 我們將以兩種方法證明 X 與 Y 爲獨立.

法 1. 由於

$$\begin{aligned} F_X(x) &= P\{X \leq x\} = P\{X \leq x,\ Y \in \mathbb{R}\} = \lim_{y\to+\infty} P\{X \leq x,\ Y \leq y\} \\ &= \lim_{y\to+\infty} F(x,\,y) = \begin{cases} 1 - e^{-x}, & \text{若 } x \geq 0, \\ 0, & \text{若 } x < 0; \end{cases} \end{aligned}$$

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P\{Y \leq y,\ X \in \mathbb{R}\} = \lim_{x\to+\infty} P\{X \leq x,\ Y \leq y\} \\ &= \lim_{x\to+\infty} F(x,\,y) = \begin{cases} 1 - e^{-y}, & \text{若 } y \geq 0, \\ 0, & \text{若 } y < 0; \end{cases} \end{aligned}$$

是以

$$F_X(x)F_Y(y) = \begin{cases} (1-e^{-x})(1-e^{-y}), & 若\ x \geq 0\ 且\ y \geq 0, \\ 0, & 若\ x < 0\ 或\ y < 0, \end{cases}$$
$$= F(x,\, y), \forall x,\, y \in \mathbb{R}.$$

故知 X 與 Y 爲獨立.

法 2. 由於 $(X,\, Y)$ 之密度函數爲 $f\colon \mathbb{R}^2 \to \mathbb{R}_+$:

$$f(x,\, y) = \begin{cases} \dfrac{\partial^2 F(x,\, y)}{\partial x \partial y}, & 若偏導數存在, \\ 0\quad, & 否則, \end{cases}$$
$$= \begin{cases} e^{-(x+y)}, & 若\ \ x > 0\ 且\ y > 0, \\ 0\quad, & 若\ \ x \leq 0\ 或\ y \leq 0. \end{cases}$$

是以

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,\, y)dy = \begin{cases} e^{-x}, & 若\ x > 0, \\ 0, & 若\ x \leq 0; \end{cases}$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,\, y)dx = \begin{cases} e^{-y}, & 若\ y > 0, \\ 0, & 若\ y \leq 0; \end{cases}$$

顯然 $f(x,\, y) = f_X(x)f_Y(y),\ \forall (x, y) \in \mathbb{R}^2$, 故 X 與 Y 爲獨立. □

在本章稍早, 我們知道：若 X 與 Y 爲獨立, 則 X 與 Y 爲不相關; 我們也提過, 若 X 與 Y 爲不相關, 則 X 與 Y 未必爲獨立, 在此, 我們舉一反例以說明之.

例 3. 設隨機向量 $(X,\, Y)$ 之密度函數爲

$$f\colon \mathbb{R}^2 \to [0,\, 1] : f(x, y) = \begin{cases} 1/4, & 若\ (x,\, y) \in A, \\ 0, & 否則, \end{cases}$$

內 $A = \{(1,\, 0), (0,\, 1), (-1,\, 0), (0,\, -1)\}$. 證明 X 與 Y 爲不相關, 但 X 與 Y 亦不爲獨立.

解 $1°$ 由原設我們立即可得下表:

$f(x,\, y)$		$x=-1$	$x=0$	$x=1$	$f_Y(y)$
y	-1	$-$	$1/4$	$-$	$1/4$
	0	$1/4$	$-$	$1/4$	$1/2$
	1	$-$	$1/4$	$-$	$1/4$
$f_X(x)$		$1/4$	$1/2$	$1/4$	1

換言之, X 與 Y 之密度函數分別爲

$$f_X\colon \mathbb{R}\to[0,\,1]: f_X(x)=\begin{cases}1/4, & 若\ x\in\{-1,\,1\},\\ 1/2, & 若\ x=0,\\ 0, & 否則,\end{cases}$$

$$f_Y\colon \mathbb{R}\to[0,\,1]: f_Y(y)=\begin{cases}1/4, & 若\ y\in\{-1,\,1\},\\ 1/2, & 若\ y=0,\\ 0, & 否則,\end{cases}$$

由

$$f(1,\,1)=0\neq 1/16=f_X(1)f_Y(1),$$

知 X 與 Y 不爲獨立. (亦可利用系 6.14, 因爲

$$A_1=\{-1,\,0,\,1\}, A_2=\{-1,\,0,\,1\},$$
$$A_1\times A_2=\{(x,\,y)\mid x,\,y\in\{-1,\,0,\,1\}\}\neq A.)$$

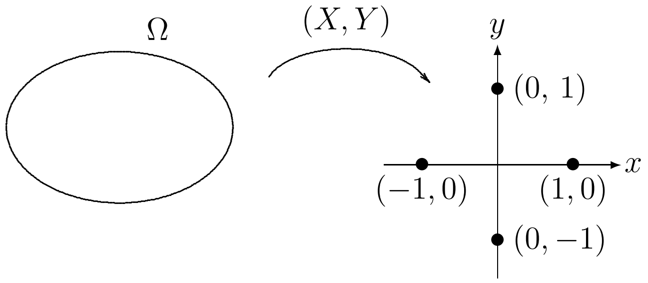


圖 6–1

2° 其次, 我們證明 X 與 Y 爲不相關. 此因

$$E[XY]=\sum_{(x,\,y)\in A} xyf(x,\,y)=0,$$

$$EX=\sum_{x\in A_1} xf_X(x)=0,$$

$$EY=\sum_{y\in A_2} yf_Y(y)=0,$$

是以

$$\rho(X,\,Y)=\frac{E[XY]-(EX)(EY)}{\sigma_X\sigma_Y}=0. \qquad \square$$

例 4. 設隨機向量 $(X,\,Y)$ 均勻分配於單位圓 $A=\{(x,\,y)\in\mathbb{R}^2\mid x^2+y^2\leq 1\}$ 上, 即其密度函數爲

$$f\colon \mathbb{R}^2\to\mathbb{R}_+: f(x,\,y)=\begin{cases}\dfrac{1}{\pi}, & 若\ (x,\,y)\in A,\\ \\ 0, & 否則.\end{cases}$$

試問 X 與 Y 是否爲獨立 ?

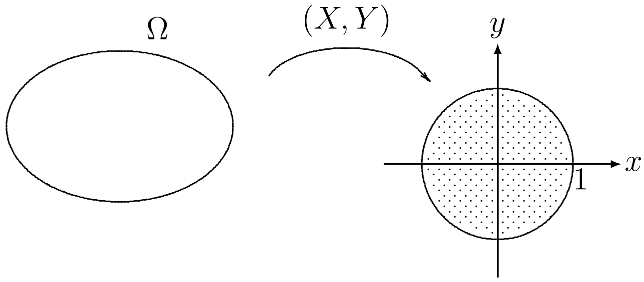


圖 6–2

解 因爲集合

$$A = \{(x,\, y) \in \mathbb{R}^2 \mid f(x,\, y)) > 0\} = \{(x,\, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\},$$

爲一圓, 不可能殆遍等於二實數集合之積, 由系 6.14 知, X 與 Y 不爲獨立. □

儘管對於一般分配函數, 不相關之二隨機變數未必爲獨立, 但若已知其聯合分配具二元常態分配, 不相關性與獨立性卻無差異, 我們將介紹於以下之定理.

定理 6.15

設隨機向量 $(X,\, Y)$ 具二元常態分配, 則 X 與 Y 爲獨立之充要條件爲 X 與 Y 爲不相關.

證 我們只需證明: 若X 與 Y 爲不相關, 則 X 與 Y 爲獨立. 因爲 $(X,\, Y)$ 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Big(-\frac{Q(x, y)}{2}\Big),$$

其中

$$Q(x, y) = \frac{1}{1-\rho^2}\Big[\Big(\frac{x-\mu_1}{\sigma_1}\Big)^2 - 2\rho\Big(\frac{x-\mu_1}{\sigma_1}\Big)\Big(\frac{y-\mu_2}{\sigma_2}\Big) + \Big(\frac{y-\mu_2}{\sigma_2}\Big)^2\Big],$$

我們可證明 X 與 Y 之相關係數 $\rho(X,\, Y) = \rho$, (習題 4.17), 是以若X 與 Y 爲不相關, 即 $\rho = \rho(X,\, Y) = 0$, 我們有

$$\begin{aligned} f(x,\, y) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\Big[-\frac{1}{2}\Big(\frac{x-\mu_1}{\sigma_1}\Big)^2 - \frac{1}{2}\Big(\frac{y-\mu_2}{\sigma_2}\Big)^2\Big]. \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} \exp\Big[-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\Big] \cdot \frac{1}{\sigma_2\sqrt{2\pi}} \exp\Big[-\frac{1}{2\sigma_2^2}(y-\mu_2)^2\Big] \end{aligned}$$

顯然 X 與 Y 之邊際密度函數分別爲

$$f_X\colon \mathbb{R} \to \mathbb{R}_+ : f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\Big[-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\Big],$$

$$f_Y\colon \mathbb{R} \to \mathbb{R}_+ : f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\Big[-\frac{1}{2\sigma_2^2}(y-\mu_2)^2\Big];$$

故得 $f(x,\, y) = f_X(x) f_Y(y)$. □

§ 6.2 獨立性與特徵函數之關係

在上一節中, 我們已經獲得獨立隨機變數之一些充要條件, 有些條件頗爲理論, 另一些則較易於掌握. 本節中, 首先, 我們要利用特徵函數, 求得獨立性之另一充要條件. 其次, 我們也要利用特徵函數方法探討有關獨立隨機變數之線性組合之分配問題, 這是繼第三章隨機向量變換的另一種方法, 實用價值很高.

定理 6.16

設 X 與 Y 爲隨機變數, $\phi = \phi_{(X,\,Y)}$ 爲隨機向量 $(X,\,Y)$ 之特徵函數, 則 X 與 Y 爲獨立之充要條件爲 $\phi(s,\,t) = \phi_X(s)\phi_Y(t), \forall s, \forall t \in \mathbb{R}$.

證 1° 若 X 與 Y 爲獨立, 則

$$\begin{aligned}\phi(s,\,t) &= E[e^{isX+itY}]\\ &= E[e^{isX}\cdot e^{itY}]\\ &= E[e^{isX}]\cdot E[e^{itY}],\ (\text{習題 6.10})\\ &= \phi_X(s)\phi_Y(t).\end{aligned}$$

2° 由於牽涉之技巧較爲繁複, 我們只大略介紹證明之步驟如下: 二變數之反演公式爲

$$\begin{aligned}&P\{a-h < X \le a+h, b-k < Y \le b+k\}\\ &\quad = \lim_{T\to+\infty}\frac{1}{\pi^2}\int_{-T}^{T}\int_{-T}^{T}\frac{\sin sh}{s}\cdot\frac{\sin tk}{t}\exp[-i(as+bt)]\phi(s,\,t)\,ds\,dt,\end{aligned}$$

其中 $(a-h,\,a+h]\times(b-k,\,b+k]\subset C(F_{X,Y})$, ($F_{X,Y}$ 之連續點集合). 是以, 當 $\phi(s,\,t) = \phi_X(s)\phi_Y(t)$ 時, 上式可分解爲二積分之積, 其一爲 ϕ_X 之反演公式, 另一則爲 ϕ_Y 之反演公式, 亦即

$$\begin{aligned}&P\{a-h < X \le a+h, b-k < Y \le b+k\}\\ &\qquad = P\{a-h < X \le a+h\}\cdot P\{b-k < Y \le b+k\},\end{aligned}$$

因此, 我們證得 $F_{X,Y}(x,\,y) = F_X(x)F_Y(y)$, 由定理 6.11 知, X 與 Y 爲獨立. □

系 6.17

設 $X_1, \cdots, X_k$ 爲隨機變數, ϕ 爲 $\mathbf{X} = (X_1, \cdots, X_k)$ 之特徵函數, ϕ_j 爲 X_j 之特徵函數, 則以下二陳述爲對等：

(I) $X_1, \cdots, X_k$ 爲獨立;

(VI) $\forall (t_1, \cdots, t_k) \in \mathbb{R}^k, \phi(t_1, \cdots, t_k) = \phi_1(t_1)\cdots\phi_k(t_k)$.

證 (I) $\Rightarrow$ (VI). 仿定理 6.16 即可, 至於其逆則更爲繁複. $\square$

定理 6.18

設 $X_1, \cdots, X_k$ 爲獨立之隨機變數, 則其線性組合 $X = \sum_{j=1}^{k} c_j X_j$ 之特徵函數爲

$$\phi_X(t) = \prod_{j=1}^{k} \phi_j(c_j t), \ \forall t \in \mathbb{R},$$

內 $c_j \in \mathbb{R}$, ϕ_j 爲 X_j 之特徵函數.

證 爲方便計, 以 $\sum$ 表 $\sum_{j=1}^{k}$, 以 $\prod$ 表 $\prod_{j=1}^{k}$.

$$\begin{aligned}
\phi_X(t) &= \phi_{\sum c_j X_j}(t) \\
&= E[\,\exp(it \textstyle\sum c_j X_j)\,] \\
&= E[\,\textstyle\prod \exp(itc_j X_j)\,] \\
&= \textstyle\prod E[\,\exp(itc_j X_j)\,], \ (\text{因 } X_1, \cdots, X_k \text{ 爲獨立}) \\
&= \textstyle\prod \phi_j(c_j t).
\end{aligned}$$

$\square$

系 6.19

獨立隨機變數之和的特徵函數等於各隨機變數之特徵函數之積.

定理 6.20 （二項分配之加法性定理）

設 $X_1 \sim B(n_1, p), \ \cdots, \ X_k \sim B(n_k, p)$, 且 $X_1, \cdots, X_k$ 爲獨立, 則其和 $X = \sum_{j=1}^{k} X_j$ 具 $B(n, p)$, 其中 $n = n_1 + \cdots + n_k$.

證 由於 $X_j \sim B(n_j, p)$, 已知其特徵函數爲

$$\phi_j(t) = (pe^{it} + q)^{n_j},$$

由定理 6.18 知,

$$\begin{aligned}
\phi_X(t) &= \prod_{j=1}^{k} (pe^{it} + q)^{n_j} \\
&= (pe^{it} + q)^{n_1 + \cdots + n_k} \\
&= (pe^{it} + q)^{n},
\end{aligned}$$

故知 X 具 $B(n, p)$ 分配. $\square$

定理 6.21 （**Poisson 分配之加法性定理**）

設 $X_1 \sim P(\lambda_1), \cdots, X_k \sim P(\lambda_k)$, 且 $X_1, \cdots, X_k$ 為獨立, 則其和 $X = \sum_{j=1}^{k} X_j$ 具 $P(\lambda)$, 其中 $\lambda = \lambda_1 + \cdots + \lambda_k$.

證 由於 $X_j \sim P(\lambda_)$, 已知其特徵函數為

$$\phi_j(t) = \exp(\lambda_j e^{it} - \lambda_j),$$

由定理 6.18 知,

$$\begin{aligned}\phi_X(t) &= \prod_{j=1}^{k} \exp(\lambda_j e^{it} - \lambda_j) \\ &= \exp\Big(\sum_{j=1}^{k} (\lambda_j e^{it} - \lambda_j)\Big) \\ &= \exp(\lambda e^{it} - \lambda),\end{aligned}$$

故知 X 具 $P(\lambda)$ 分配. □

定理 6.22 （**常態分配之加法性定理**）

設 $X_1 \sim N(\mu_1, \sigma_1^2), \cdots, X_k \sim N(\mu_k, \sigma_k^2)$, 且 $X_1, \cdots, X_k$ 為獨立, 則其線性組合 $X = \sum_{j=1}^{k} c_j X_j$ 具 $N(\mu, \sigma^2)$, 其中 $\mu = \sum_{j=1}^{k} c_j \mu_j,\ \sigma^2 = \sum_{j=1}^{k} c_j^2 \sigma_j^2$.

證 由於 X_j 具常態分配 $N(\mu_j, \sigma_j^2)$, 已知其特徵函數為

$$\phi_j(t) = \exp\left(it\mu_j - \frac{t^2\sigma_j^2}{2}\right),$$

由定理 6.18 知,

$$\begin{aligned}\phi_X(t) &= \prod_{j=1}^{k} \phi_j(c_j t) \\ &= \prod_{j=1}^{k} \exp\Big(itc_j\mu_j - \frac{t^2 c_j^2 \sigma_j^2}{2}\Big) \\ &= \exp\Big(\sum_{j=1}^{k}\Big(itc_j\mu_j - \frac{t^2}{2} c_j^2\sigma_j^2\Big)\Big) \\ &= \exp\Big(it\mu - \frac{t^2\sigma^2}{2}\Big).\end{aligned}$$

故知 X 具 $N(\mu, \sigma^2)$ 分配. □

本定理說明具常態分配之獨立隨機變數的線性組合亦具常態分配. 至於其期望值, 我們由期望值之線性性質知

$$EX = E\Big[\sum_{j=1}^{k} c_j X_j\Big] = \sum_{j=1}^{k} c_j EX_j = \sum_{j=1}^{k} c_j \mu_j,$$

(獨立性未用上). 而其變異數, 則因諸 X_j 爲獨立, 利用 Bienaymé 公式可得

$$\mathrm{Var}\ X = \sum_{j=1}^{k} \mathrm{Var}\,(c_j X_j) = \sum_{j=1}^{k} c_j^2\, \mathrm{Var}\ X_j = \sum_{j=1}^{k} c_j^2 \sigma_j^2.$$

此二者均與定理 6.22 之結果相同. 本定理有個很重要的系, 在機率與統計上常被提及.

系 6.23

設 $X_1, \cdots, X_n$ 爲獨立且均具常態分配 $N(\mu, \sigma^2)$, 則

(1) 隨機變數 $S_n = \sum_{j=1}^{n} X_j$ 具 $N(n\mu, n\sigma^2)$.

(2) 隨機變數 $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ 具 $N(\mu, \frac{\sigma^2}{n})$. ($\overline{X}$ 又稱爲 $X_1, \cdots, X_n$ 之**樣本平均**(sample mean)).

證 (1) 在本定理中令 $c_j = 1,\ \mu_j = \mu,\ \forall j$, 並且以 n 代定理中之 k 即可.

(2) 仿 (1), 但 $c_j = \frac{1}{n}$. □

我們在此特別聲明, 本系之 (2) 之含義爲：若某一母群體具常態分配, 則以放回方式(即獨立) 自其中抽取 n 個值 $X_1, \cdots, X_n$, 平均之, 則得 $\overline{X}$, 其分配亦為常態, 期望值不變, 但變異數則縮小為原來的 $1/n$. 此一性質在統計學上極具重要性.

例 1. 設 $X_1, \cdots, X_{20}, Y_1, \cdots, Y_{50}$ 爲獨立之隨機變數, 而諸 X_i 均具常態分配函數 $N(1, 3)$, 諸 Y_j 均具常態分配函數 $N(2, 5)$. 若 $\overline{X}$ 表諸 X_i 之樣本平均, $\overline{Y}$ 表諸 Y_j 之樣本平均, 試求 $P\{\overline{X} > \overline{Y}\} =?$

解 先證: $\overline{X} - \overline{Y} \sim N(-1, 1/4)$. 理由是:

1° 由系 6.23 知, $\overline{X}$ 具 $N(1, 3/20)$, $\overline{Y}$ 具 $N(2, 5/50)$, 故

$$E[\overline{X} - \overline{Y}] = E[\overline{X}] - E[\overline{Y}] = 1 - 2 = -1,$$

2° 由定理 6.6 知 $\overline{X}$ 與 $\overline{Y}$ 爲獨立, 故

$$\mathrm{Var}\,(\overline{X} - \overline{Y}) = \mathrm{Var}\ \overline{X} + \mathrm{Var}\ \overline{Y} = \frac{3}{20} + \frac{5}{50} = \frac{1}{4}.$$

所求之機率爲

$$
\begin{aligned}
P\{\overline{X} > \overline{Y}\} &= P\{\overline{X} - \overline{Y} > 0\} \\
&= P\Big\{\frac{\overline{X} - \overline{Y} - (-1)}{\sqrt{1/4}} > \frac{1}{\sqrt{1/4}}\Big\} \\
&= 1 - \Phi(2) \\
&= 1 - 0.977250, (\text{查表或計算機}) \\
&= 0.0022750. \qquad \square
\end{aligned}
$$

定理 6.24 （**Gamma 分配之加法性定理**）

> 設 $X_1 \sim G(\alpha_1, \beta), \cdots, X_k \sim G(\alpha_k, \beta)$, 且 $X_1, \cdots, X_k$ 爲獨立, 則其和 $X = \sum_{j=1}^k X_j$ 具 $G(\alpha, \beta)$ 分配, 其中 $\alpha = \alpha_1 + \cdots + \alpha_k$.

證 仿上, 因 X_j 具 Gamma 分配, 已知其特徵函數爲 $\phi_j(t) = (1 - i\beta t)^{-\alpha_j}$. $\square$

定理 6.25 （χ^2 **分配之加法性定理**）

> 設 $X_1 \sim \chi^2_{r_1}, \cdots, X_k \sim \chi^2_{r_k}$, 且 $X_1, \cdots, X_k$ 爲獨立, 則其和 $X = \sum_{j=1}^k X_j$ 具 χ^2_r 分配, 其中 $r = r_1 + \cdots + r_k$.

證 因 X_j 具 $\chi^2_{r_j} = G(r_j/2, \beta)$ 之故. $\square$

第六章 習 題

6-1 設 A 與 B 爲 $(\Omega, \mathcal{F}, P)$ 上二事件. 試證: A 與 B 爲獨立之充要條件爲指標函數 I_A 與 I_B 爲獨立.

6-2 (1) 設 X 與 Y 爲獨立之離散型隨機變數, 證明: 若 $P\{X = x\} > 0$, 則

$$\forall y \in \mathbb{R},\ f(y|x) = f_Y(y).$$

(2) 設 X 與 Y 爲獨立之絕對連續型隨機變數, 在什麼條件下仍有

$$\forall y \in \mathbb{R},\ f(y|x) = f_Y(y).$$

6-3 若 $X \sim B(1, p), Y \sim U(0, 1)$, 且二者爲獨立.

(1) 試求隨機向量 (X, Y) 之累積分配函數;

(2) 試問 (X, Y) 是否爲連續型? 是否爲絕對連續型 ?

6-4 設隨機向量 (X, Y) 之密度函數爲

$$f\colon \mathbb{R}^2 \to \mathbb{R}_+ : f(x, y) = \begin{cases} kxe^{-(x+y)}, & \text{若 } x > 0,\ y > 0, \\ 0, & \text{否則}, \end{cases}$$

(1) 試求常數 $k =?$

(2) 試問 X 與 Y 是否爲獨立 ?

6-5 設隨機變數 X 與 Y 爲獨立且均具相同之 Pascal 分配.

(1) 試求 $P\{X \geq x\} =?$ 內 $x \in \mathbb{Z}_+$;

(2) 試求 $Z = \min\{X, Y\}$ 之分配函數;

(3) 證明 Z 亦具 Pascal 分配.

6-6 試舉一例, 其中隨機變數 X 與 Y 均爲絕對連續型, X 與 Y 爲不相 關, 且 X 與 Y 不爲獨立.

6-7 證明: 若 (X, Y) 具多項分配, 則 X 與 Y 不爲獨立.

6-8 設隨機變數 X 與 Y 爲獨立且均具相同之 Cauchy 分配.

(1) 試求 $X + Y$ 之特徵函數;

(2) 試問 $(X + Y)/2$ 具有何種分配.

6-9 設隨機變數 X 與 Y 爲獨立且均具相同分配. 試證：$\phi_{X-Y}(t) = |\phi_X(t)|^2, \forall t \in \mathbb{R}$.

6-10 設 X 與 Y 爲獨立之隨機變數. 利用系 6.8 試證:

$$E[e^{isX}\cdot e^{itY}]=E[e^{isX}]\cdot E[e^{itY}].$$

6-11 設 $X_1\sim NB(r_1,\,p),\cdots,X_k\sim NB(r_k,\,p)$, 且 $X_1,\,\cdots,\,X_k$ 爲獨立, 試證其和 $X=\sum_{j=1}^k X_j$ 具 $NB(r,\,p)$, 其中 $r=r_1+\cdots+r_k$.

6-12 二位同學相約在中午 12:00 至 12:30 在數學系館門口見面. 假設二人於該時段到達之時間爲隨機且彼此獨立, 試問其中一人等候另一人之時間超過五分鐘之機率爲若干 ? (甲等乙或乙等甲.)

6-13 自區間 $[0,1]$ 上隨機取二數, 並將此二數標示於一單位長度之竹桿上, 然後於標示點將竹桿切斷, 試問此三段能圍成一三角形之機率爲若干 ?

6-14 設 $X,\,Y,\,Z$ 皆爲隨機變數, 若 X 與 Z 爲獨立且 Y 與 Z 爲獨立, 舉一反例以說明 $X+Y$ 與 Z 未必爲獨立.

Chapter 7

極限定理

§ 7.1　機率論發展初期的極限問題

機率的極限理論, 在機率論初步形成之際就已經出現. 1713 年 Bernoulli 在他遺作《猜測的藝術》中, 以樸拙的分析方法證明了機率論中的第一個極限定理：如果 $X_1, X_2, \cdots, X_n, \cdots$ 爲獨立且均具二項分配 $B(1, p)$ 之隨機變數序列, $S_n = X_1 + \cdots + X_n$, 則 $\forall \epsilon > 0$,

$$\lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| < \epsilon\Big\} = 1.$$

譬如投籃, X_j 表第 j 次投籃之結果, 進則爲 1, 不進則爲 0, 如此投了 n 個球 , S_n 表 n 次投籃之進球數, 而 S_n/n 則爲 n 次投籃之進球率, (也就是所謂的相對頻率). Bernoulli 的目的在於說明 S_n/n 與理論的進球率 p 十分接近.

機率中的第二個極限理論也和二項分配有關, 前面我們一再提及 $\binom{n}{x}p^x q^{n-x}$ 在計算上的困難, 1732 年 de Moivre 證明：當 $p = 1/2$ 時,

$$\binom{n}{x}p^x q^{n-x} \sim \frac{1}{\sqrt{2\pi npq}}\exp\Big[-\frac{1}{2}\Big(\frac{x-np}{\sqrt{npq}}\Big)^2\Big],$$

(即二者之比的極限爲 1, 當 $n \to +\infty$). 1801 年 Laplace 把這種結果推廣到 $0 < p < 1$ 之情形, 後來又進一步以積分方式解決此一問題, 亦即

$$\lim_{n\to+\infty} P\Big\{a < \frac{S_n - np}{\sqrt{npq}} \le b\Big\} = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\Big(-\frac{x^2}{2}\Big)dx.$$

常態分配於是誕生.

機率的第三個極限理論也是探討 $\binom{n}{x}p^x q^{n-x}$ 之近似值. 1837 年 Poisson 發現, 當 $np = \lambda$ 時, $\binom{n}{x}p^x q^{n-x}$ 之極限等於 $e^{-\lambda}\dfrac{\lambda^x}{x!}$, Poisson 分配於是問世. 以下我們將介紹這三個定理.

定理 7.1 [**Bernoulli 弱大數定律** (Weak law of large numbers)]

> 設 $X_1, X_2, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 分配之 隨機變數序列, 而且 $S_n = X_1 + \cdots + X_n$, 則 $\forall \epsilon > 0$,
>
> $$\lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big\} = 0.$$

證 由於 $S_n \sim B(n, p)$, 故 $E[S_n] = np, \mathrm{Var}\ S_n = npq$, 是以

$$\begin{aligned} P\Big\{\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big\} &= P\{|S_n - np| \geq n\epsilon\} \\ &\leq \frac{\mathrm{Var}\ S_n}{(n\epsilon)^2}, \ \text{(Chebyshev 不等式)} \\ &= \frac{pq}{n\epsilon^2}. \end{aligned}$$

當 $n \to +\infty$ 時, 顯然有

$$\lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big\} = 0. \qquad \square$$

註: 由於

$$\begin{aligned} \lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big\} = 0 &\Leftrightarrow \lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| < \epsilon\Big\} = 1 \\ &\Leftrightarrow \lim_{n\to+\infty} P\{p - \epsilon < \overline{X}_n < p + \epsilon\} = 1. \end{aligned}$$

本定理之意義乃爲: 不論 ϵ 多麼小, 當 n 很大時, 樣本平均 $\overline{X}_n$ 落在區間 $(p-\epsilon,\ p+\epsilon)$ 機率很接近 1.

de Moivre 的近似觀念

設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 分配, S_n 具 $B(n, p)$ 分配, 即其密度函數爲

$$f_n\colon \mathbb{R} \to [0,1] : f_n(j) = \begin{cases} \dbinom{n}{j} p^j q^{n-j}, & \text{若 } j \in \{0, 1, \cdots, n\}, \\ 0, & \text{否則;} \end{cases}$$

其期望值爲 $E[S_n] = np$, 其變異數爲 $\mathrm{Var}\ S_n = npq$. 令 S_n 之標準化爲 Y_n, 亦即

$$Y_n = \frac{S_n - np}{\sqrt{npq}}.$$

我們將先求 Y_n 之密度函數. 令

$$\begin{tikzcd} \Omega \arrow[r, "S_n"] \arrow[dr, "Y_n"'] & \mathbb{R} \arrow[r, "f_n"] \arrow[d, "h_n"] & {[0,1]} \\ & \mathbb{R} \arrow[ur, "g_n"'] & \end{tikzcd}$$

$$h_n\colon \mathbb{R} \to \mathbb{R} : h_n(j) = \frac{j - np}{\sqrt{npq}}.$$

(參見圖 7–1, 上方之圖先左移再縮小而成小方之圖),

則因

(1) $h_n|_A\colon A \to B$ 爲一對射, 其內

$$A = \{j \in \mathbb{R} \mid f_n(j) > 0\} = \{0,\, 1,\, \cdots,\, n\},$$
$$B = \{(j - np)/\sqrt{npq} \mid j \in A\}.$$

(2) $h_n|_A$ 之反函數爲

$$h_n^{-1}\colon B \to A : h_n^{-1}(y) = np + y\sqrt{npq}.$$

由定理 2.26 知, Y_n 之密度函數爲

$$g_n\colon \mathbb{R} \to [0,\, 1] : g_n(y) = \begin{cases} f_n(h_n^{-1}(y)), & \text{若 } y \in B, \\ 0, & \text{否則}, \end{cases}$$

內 $f_n(h_n^{-1}(y)) = \binom{n}{h_n^{-1}(y)} p^{h_n^{-1}(y)} q^{n-h_n^{-1}(y)}$, 爲方便計, 我們仍寫爲

$$g_n(y) = \binom{n}{j} p^j q^{n-j},$$

但別忘了 $j = h_n^{-1}(y) = np + y\sqrt{npq}$.

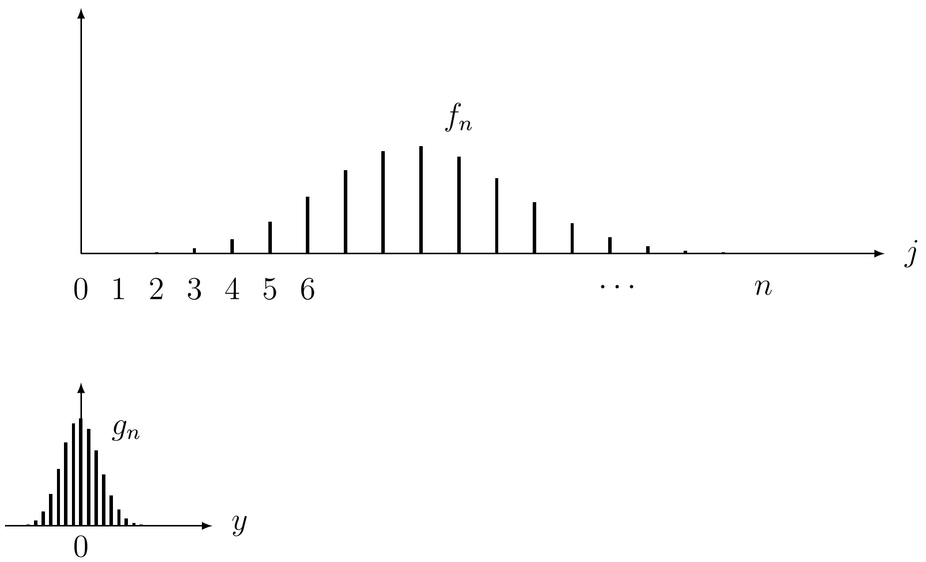


圖 7–1

de Moivre 的觀念是上述密度函數 g_n, 當 n 很大時, 與函數

$$\frac{1}{\sqrt{npq}\sqrt{2\pi}} \exp\Bigl(-\frac{y^2}{2}\Bigr)$$

十分接近, 經過一些學者的改進, 於是有了以下的面貌.

定理 7.2 （**de Moivre-Laplace 定理**）

設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 分配, $S_n = X_1 + \cdots + X_n$, $Y_n = \dfrac{S_n - np}{\sqrt{npq}}$, g_n 爲 Y_n 之密度函數. 當 $n \to +\infty$ 時, 函數

$$\frac{g_n(y)}{\dfrac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big(-\dfrac{y^2}{2}\Big)}$$

均勻收斂於 1 在 $\mathbb{R}$ 之任意緊緻區間 $I = [a, b]$.

證 1° 由 Stirling 公式知,

$$m! = \sqrt{2\pi m}\cdot m^m \cdot e^{-m}\cdot \exp(\theta_m), \text{內 } 0 < \theta_m < 1/(12m).$$

是以

$$\begin{aligned} g_n(y) &= \binom{n}{j} p^j q^{n-j} \\ &= \frac{n!}{j!\,k!} p^j q^k, (\text{令 } k = n - j) \\ &= \frac{\sqrt{2\pi n}\, n^n e^{-n}\exp(\theta_n)}{\sqrt{2\pi j}\, j^j e^{-j}\exp(\theta_j)\cdot\sqrt{2\pi k}\, k^k e^{-k}\exp(\theta_k)} p^j q^k \\ &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{jk}}\Big(\frac{np}{j}\Big)^j\Big(\frac{nq}{k}\Big)^k \exp(\theta_n - \theta_j - \theta_k). \end{aligned}$$

吾人可得

$$\frac{g_n(y)}{\dfrac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big(-\dfrac{y^2}{2}\Big)} = \frac{\sqrt{\dfrac{n}{jk}}}{\dfrac{1}{\sqrt{npq}}}\cdot\frac{\Big(\dfrac{np}{j}\Big)^j\Big(\dfrac{nq}{k}\Big)^k}{\exp\Big(\dfrac{-y^2}{2}\Big)}\cdot e^{\theta}.$$

內 $\theta = \theta_n - \theta_j - \theta_k$. 我們將證明每一部分都均勻收斂 於 1, 且因三者均爲有界, 故其乘積亦均勻收斂於 1 在區間 $[a, b]$ 上.

2° 第一部分之倒數

$$\begin{aligned} \frac{\dfrac{1}{\sqrt{npq}}}{\sqrt{\dfrac{n}{jk}}} &= \frac{\sqrt{jk}}{n\sqrt{pq}} \\ &= \frac{1}{\sqrt{pq}}\sqrt{\Big(\frac{np + y\sqrt{npq}}{n}\Big)\Big(\frac{nq - y\sqrt{npq}}{n}\Big)} \\ &= \frac{1}{\sqrt{pq}}\sqrt{\Big(p + y\sqrt{\frac{pq}{n}}\Big)\Big(q - y\sqrt{\frac{pq}{n}}\Big)}. \end{aligned}$$

由於 y 在緊緻區間上 $[a, b]$ 內, 上式顯然均勻收斂於 1.

3° 第二部分之分子

$$\begin{aligned}
\Big(\frac{np}{j}\Big)^j\Big(\frac{nq}{k}\Big)^k &= \exp\Big(j\ln\Big(\frac{np}{j}\Big)+k\ln\Big(\frac{nq}{k}\Big)\Big)\\
&= \exp\Big[-j\ln\Big(\frac{np+y\sqrt{npq}}{np}\Big)-k\ln\Big(\frac{nq-y\sqrt{npq}}{nq}\Big)\Big]\\
&= \exp\Big[-j\ln\Big(1+y\sqrt{\frac{q}{np}}\Big)-k\ln\Big(1-y\sqrt{\frac{p}{nq}}\Big)\Big]\\
&= \exp\Big[-(np+y\sqrt{npq})\Big(y\sqrt{\frac{q}{np}}-\frac{1}{2}\cdot\frac{qy^2}{np}+O\Big(\frac{1}{n^{3/2}}\Big)\Big)\\
&\qquad -(nq-y\sqrt{npq})\Big(-y\sqrt{\frac{p}{nq}}-\frac{1}{2}\cdot\frac{py^2}{nq}+O\Big(\frac{1}{n^{3/2}}\Big)\Big)\Big]
\end{aligned}$$

$\Big($因 $\ln(1+t)=t-\dfrac{t^2}{2}+\dfrac{t^3}{3}(1+c)^{-3}=t-\dfrac{t^2}{2}+O(t^3)$ [1] 其中 c 介於 0 與 t 之間之故,$\Big)$

$$= \exp\Big[-\frac{y^2}{2}+O\Big(\frac{1}{n^{1/2}}\Big)\Big],\ (\text{因 } y\in[a,b])$$

是以

$$\frac{\Big(\frac{np}{j}\Big)^j\Big(\frac{nq}{k}\Big)^k}{\exp\Big(-\dfrac{y^2}{2}\Big)}=\exp\Big[O\Big(\frac{1}{n^{1/2}}\Big)\Big]$$

顯然均勻收斂於 1.

4° 因 $\theta=\theta_n-\theta_j-\theta_k$, 故

$$|\theta|\le|\theta_n|+|\theta_j|+|\theta_k|<\frac{1}{12}\Big(\frac{1}{n}+\frac{1}{j}+\frac{1}{k}\Big).$$

而 $\lim\limits_{n\to+\infty} j=\lim\limits_{n\to+\infty} k=+\infty$, 是以 $\lim\limits_{n\to+\infty} e^\theta=1$. □

定理 7.3 (**Laplace 定理**)

設 $X_1,\cdots,X_n,\cdots$ 爲獨立且均具 $B(1,\,p)$ 分配, $S_n=X_1+\cdots+X_n$, $Y_n=\dfrac{S_n-np}{\sqrt{npq}}$, g_n 爲 Y_n 之密度函數. 當 $n\to+\infty$ 時, 函數

$$\frac{g_n(y)}{\displaystyle\int_{j-1/2}^{j+1/2} f(t)\,dt}$$

均勻收斂於 1 在 $\mathbb{R}$ 之任意緊緻區間 $I=[a,\,b]$ 上, 其中 $j=np+y\sqrt{npq}$, f 爲常態分配 $N(np,\,npq)$ 之密度函數.

---

[1] $f(t)=O(g(t))\Leftrightarrow(\exists K>0,$ 使得 $\forall t\in N_\delta(0)\ \Rightarrow\ |f(t)|<K|g(t)|,)$

證 由於 f 爲一連續函數, 由積分均值定理知存在 $\theta \in (-1/2, 1/2)$ 使得

$$\begin{aligned}\int_{j-1/2}^{j+1/2} f(t)\,dt &= f(j+\theta)\\ &= \frac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big[-\frac{1}{2}\frac{(j+\theta-np)^2}{npq}\Big]\\ &= \frac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big[-\frac{1}{2}\frac{(j-np)^2}{npq}-\frac{\theta(j-np)}{npq}-\frac{\theta^2}{2npq}\Big]\\ &= \frac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big(-\frac{y^2}{2}\Big)\exp\Big(-\frac{y\theta}{\sqrt{npq}}\Big)\exp\Big(-\frac{\theta^2}{2npq}\Big).\end{aligned}$$

是以

$$\frac{g_n(y)}{\displaystyle\int_{j-1/2}^{j+1/2} f(t)\,dt} = \frac{g_n(y)}{\frac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big(-\frac{y^2}{2}\Big)}\cdot\exp\Big(\frac{y\theta}{\sqrt{npq}}\Big)\exp\Big(\frac{\theta^2}{2npq}\Big),$$

利用 de Moivre-Laplace 定理知上述函數顯然均勻收斂於 1 於任意 y 之 區間 $I=[a, b]$ 上. □

如果我們不以均勻收斂觀念解釋, 而以較弱的點態收斂 (point-wise convergence) 觀念解釋, 則 Laplace 定理乃說明 ： 二項分配 $B(n,p)$ 之密度函數 $\binom{n}{j}p^jq^{n-j}$ 與常態分配函數 $N(np, npq)$ 之密度函數

$$f(j) = \frac{1}{\sqrt{npq}\sqrt{2\pi}}\exp\Big(-\frac{1}{2npq}(j-np)^2\Big)$$

二者十分接近, (當 n 很大時), 前者之值與後者在 $[j-\frac{1}{2}, j-\frac{1}{2}]$ 之面積比的極限爲 1. 參見圖 7–2, 斜線部分之面積乃 $\displaystyle\int_{j-1/2}^{j+1/2} f(t)\,dt$.

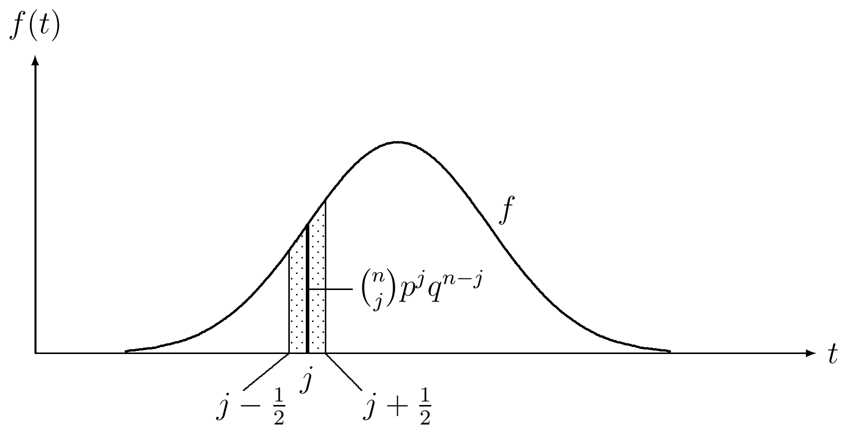


圖 7–2

系 7.4

設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1,\,p)$ 分配, $S_n = X_1 + \cdots + X_n$, 則

$$P\{\alpha \le S_n \le \beta\} \sim \Phi\Big(\frac{\beta + \frac{1}{2} - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{\alpha - \frac{1}{2} - np}{\sqrt{npq}}\Big),$$

內 $\alpha,\ \beta \in \{0, 1, \cdots, n\},\ \alpha \le \beta,\ \Phi$ 爲 $N(0,\,1)$ 之分配函數.

證 設隨機變數 Y 具 $N(np,\,npq), f$ 爲其密度函數, 即

$$\begin{aligned}
P\{\alpha \le S_n \le \beta\} &= \sum_{\alpha}^{\beta} \binom{n}{j} p^j q^{n-j} \\
&= \binom{n}{\alpha} p^{\alpha} q^{n-\alpha} + \binom{n}{\alpha+1} p^{\alpha+1} q^{n-\alpha-1} + \cdots + \binom{n}{\beta} p^{\beta} q^{n-\beta} \\
&\sim \int_{\alpha-1/2}^{\alpha+1/2} f(t)\,dt + \int_{\alpha+1/2}^{\alpha+3/2} f(t)\,dt + \cdots + \int_{\beta-1/2}^{\beta+1/2} f(t)\,dt \\
&= \int_{\alpha-1/2}^{\beta+1/2} f(t)\,dt \\
&= P\{\alpha - 1/2 \le Y \le \beta + 1/2\} \\
&= P\Big\{\frac{\alpha - \frac{1}{2} - np}{\sqrt{npq}} \le \frac{Y - np}{\sqrt{npq}} \le \frac{\beta + \frac{1}{2} - np}{\sqrt{npq}}\Big\} \\
&= \Phi\Big(\frac{\beta + \frac{1}{2} - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{\alpha - \frac{1}{2} - np}{\sqrt{npq}}\Big).
\end{aligned}$$

□

例 1. 將一公正之銅板連擲 100 次, 試問正面出現 30 次至 45 次之機率爲若干 ?

解 設 X_k 表第 k 次出現正面之值, 即若 $X_k = 1$ 表示第 k 次出現正面, 若 $X_k = 0$ 表示第 k 次出現反面. 則依題意可設 $X_1, \cdots, X_{100}$ 爲獨 立且均具 $B(1,\,p), p = 1/2$, 而求 $P\{30 \le S_{100} \le 45\} =?$

法 1. 利用計算機

$$P\{30 \le S_{100} \le 45\} = \sum_{x=30}^{45} \binom{100}{x} \Big(\frac{1}{2}\Big)^{100} = 0.184083.$$

(讀者不妨利用其他軟體計算核對一下).

法 2. 利用 Laplace 定理 : 由於 $n = 100,\ p = 1/2$, 故 $np = 50,\ npq = 25$. 是以

$$\begin{aligned}
P\{30 \le S_{100} \le 45\} &\approx \Phi\Big(\frac{45 + \frac{1}{2} - 50}{\sqrt{25}}\Big) - \Phi\Big(\frac{30 - \frac{1}{2} - 50}{\sqrt{25}}\Big) \\
&= \Phi(-0.9) - \Phi(-4.1) \\
&= 1 - \Phi(0.9) - 1 + \Phi(4.1) \\
&= 1 - 0.815940 - 1 + 0.999970 \\
&= 0.184030.
\end{aligned}$$

□

引理 7.5

設 $\{c_n\}_n$ 爲一複數序列且 $\lim\limits_{n\to+\infty} c_n = k \in \mathbb{C}$, 則

$$\lim_{n\to+\infty}\Big(1+\frac{c_n}{n}\Big)^n = e^k. \tag{$*$}$$

證 令 $A = \{n \in \mathbb{N} \mid c_n = 0\},\ B = \mathbb{N} \setminus A$.

1° 若 A 爲有限, 則存在 n_0 使得 $\forall n > n_0,\ c_n \neq 0$, 是以

$$\begin{aligned}
\lim_{n\to+\infty}\Big(1+\frac{c_n}{n}\Big)^n &= \lim_{n\to+\infty}\exp\Big(n\cdot \mathrm{Ln}\Big(1+\frac{c_n}{n}\Big)\Big)\\
&= \lim_{n\to+\infty}\exp\Big(c_n\cdot\frac{\mathrm{Ln}(1+\frac{c_n}{n})}{\frac{c_n}{n}}\Big)\\
&= \exp\Big(\lim_{n\to+\infty} c_n\cdot\frac{\mathrm{Ln}(1+\frac{c_n}{n})}{\frac{c_n}{n}}\Big)\\
&= \exp\Big(\lim_{n\to+\infty} c_n\cdot\lim_{m\to 0}\frac{\mathrm{Ln}(1+m)}{m}\Big), \quad \Big(\text{令 } m=\frac{c_n}{n}\Big)\\
&= \exp\Big(\lim_{n\to+\infty} c_n\cdot\frac{\frac{1}{1+0}}{1}\Big), \ \text{(l'Hospital 規則)}\\
&= e^k.
\end{aligned}$$

2° 若 A 爲無限, 則

$$k = \lim_{n\to+\infty} c_n = \lim_{n\in A, n\to+\infty} c_n = 0,$$

不論 B 是否爲無限, $(*)$ 式之兩端顯然皆等於 1. □

定理 7.6 (**Poisson 定理**)

設 $\forall n \in \mathbb{N}$, 隨機變數 X_n 具 $B(n, p_n)$, 其密度函數爲 f_n, 次設隨機變數 Y 具 $P(\lambda)$, 其密度函數爲 g. 若 $\lim\limits_{n\to+\infty} np_n = \lambda$, 則

$$\forall x \in \mathbb{R}, \lim_{n\to+\infty} f_n(x) = g(x).$$

證 1° 若 $x \in \mathbb{R}\setminus\mathbb{Z}_+$, 則 $\forall n \in \mathbb{N},\ f_n(x) = 0$, 是以

$$\lim_{n\to+\infty} f_n(x) = 0 = g(x).$$

$2°$ 若 $x \in \mathbb{Z}_+, \forall n \geq x$,

$$\begin{aligned}
f_n(x) &= \frac{n!}{x!(n-x)!} p_n^x (1-p_n)^{n-x} \\
&= \frac{n!}{x!(n-x)!} \Big(\frac{\lambda_n}{n}\Big)^x \Big(1-\frac{\lambda_n}{n}\Big)^{n-x}, \ (\text{令 } \lambda_n = np_n) \\
&= \frac{\lambda_n^x}{x!} \Big(1-\frac{\lambda_n}{n}\Big)^n \cdot 1 \cdot \Big(1-\frac{1}{n}\Big) \cdots \Big(1-\frac{x-1}{n}\Big) \cdot \frac{1}{\Big(1-\frac{\lambda_n}{n}\Big)^x},
\end{aligned}$$

由引理 7.5 知

$$\lim_{n\to+\infty} \Big(1-\frac{\lambda_n}{n}\Big)^n = e^{-\lambda},$$

是以

$$\lim_{n\to+\infty} f_n(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \cdot 1 \cdot 1^x \cdot \frac{1}{(1-0)^x} = \frac{\lambda^x}{x!} \cdot e^{-\lambda} = g(x).$$ □

註： 本定理之用途：當 n 很大而 p 很小, 令 $\lambda = np$, 則

$$\binom{n}{x} p^x q^{n-x} \approx e^{-\lambda} \cdot \frac{\lambda^x}{x!}, x \in \{0, 1, \cdots, n\}$$

簡言之, Poisson 分配可做爲二項分配之近似值.

例 2. 若 X 具 $B(100, 0.1)$, 試求 $P\{10 \leq X \leq 20\} =?$

$\boxed{\text{解}}$ $n = 100$ 之二項分配之表不易找到, 我們必須另行設法.

法 1. 利用計算機

$$P\{10 \leq X \leq 20\} = \sum_{x=10}^{20} \binom{100}{x} 0.1^x 0.9^{100-x} = 0.5479007.$$

法 2. 利用 Poisson 定理以求其近似值. 令隨機變數 Y 具 Poisson 分配$P(\lambda)$, 其中 $\lambda = 100 \times 0.1 = 10$; f 與 g 分別爲 X 與 Y 之密度函數. 則

$$\begin{aligned}
P\{10 \leq X \leq 20\} &= \sum_{x=10}^{20} f(x) \\
&\approx \sum_{x=10}^{20} g(x), \ (\text{Poisson 定理}) \\
&= P\{10 \leq Y \leq 20\} \\
&= P\{Y \leq 20\} - P\{Y \leq 9\} \\
&= 0.99841174 - 0.45792971, \ (\text{查表或利用電腦}) \\
&= 0.54058203.
\end{aligned}$$

法 3. 本題亦可利用 Laplace 定理以求其近似值, 讀者不妨參考例 1 自行求解之. □

§ 7.2 收斂性

上一節中, 我們知道某些隨機變數序列, 當 $n \to +\infty$ 時, 其分配可能收斂於某函數. 此外, 在高等微積分中, 我們知道一函數序列亦可能有簡單 (點態) 收斂性及均勻收斂性, 因此, 在機率論中乃發展出各種不同的收斂性. 我們將僅介紹其中較常用的部分, 有興趣者可參閱 Loève [11], Laha & Rohagi [10] 或 Chung [4], 當有更詳盡的介紹.

定義 7.7

設 $\{X_n\}_n$ 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數序列, X 爲 $(\Omega, \mathcal{F}, P)$ 上一隨機變數, F_n 及 F 分別爲 X_n 及 X 之分配函數.

(1) 我們稱序列 $\{X_n\}_n$ **殆必收斂於** X ($\{X_n\}_n$ converges a.s. to X), 並記爲 $X_n \xrightarrow{\text{a.s.}} X$, 如果存在一零機率之事件 N 使得

$$\forall \omega \in \Omega \setminus N, \lim_{n \to +\infty} X_n(\omega) = X(\omega).$$

(2) 我們稱 $\{X_n\}_n$ **機率收斂於** X ($\{X_n\}_n$ converges in probability to X), 並記爲 $X_n \xrightarrow{\text{P}} X$, 如果

$$\forall \epsilon > 0, \lim_{n \to +\infty} P\{|X_n - X| < \epsilon\} = 1.$$

(3) 我們稱 $\{X_n\}_n$ **分配收斂於** X ($\{X_n\}_n$ converges in distribution to X), 並記爲 $X_n \xrightarrow{\text{d}} X$, 如果

$$\forall x \in C(F), \quad \lim_{n \to +\infty} F_n(x) = F(x),$$

內 $C(F) = \{x \in \mathbb{R} \mid F \text{ 連續於點 } x\}$.

(4) 我們稱 $\{X_n\}_n$ L_r **收斂於** X ($\{X_n\}_n$ converges in L_r to X), 並記爲 $X_n \xrightarrow{L_r} X$, (其中 r 爲一 正數), 如果

$$\lim_{n \to +\infty} E|X_n - X|^r = 0.$$

(5) 如果 $\{X_n\}_n$ L_1 收斂於 X 我們亦稱 $\{X_n\}_n$ **平均收斂**於 X ($\{X_n\}_n$ converges in mean to X). 如果 $\{X_n\}_n$ L_2 收斂於 X 我們亦稱 $\{X_n\}_n$ **均方收斂**於 X ($\{X_n\}_n$ converges in quadratic mean to X).

註： 我們之所以稱 $\{X_n\}_n$ L_r 收斂於 X, 是因爲隨機變數集合

$$L_r = \{X \mid E|X|^r < +\infty\}$$

爲一向量空間, 當 $r > 1$ 時, 於 L_r 上我們界定其範數 (norm) 爲

$$\|\cdot\| \colon L_r \to \mathbb{R}_+ : \|X\|_r = (E|X|^r)^{1/r}.$$

如此一來, $X_n \xrightarrow{L_r} X$ 相當於在賦範空間 $(L_r, \|\cdot\|)$ 中, $\lim\limits_{n\to+\infty} X_n = X$.

例 1. 設 $\Omega = [0, 1], P$ 爲 Ω 上之 Lebesgue 測度, 若

$$X_n \colon \Omega \to \mathbb{R} : X_n(\omega) = \omega^n,$$
$$X \colon \Omega \to \mathbb{R} : X(\omega) = 0,$$

則顯然有 $\forall \omega \in \Omega \setminus \{1\}$, $\lim\limits_{n\to+\infty} X_n(\omega) = 0 = X(\omega)$. 因此, $X_n \xrightarrow{\text{a.s.}} X$. □

例 2. 在本章第一節 Bernoulli 弱大數定律中, 我們知道：若 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 分配, $S_n = X_1 + \cdots + X_n$, 則 $\forall \epsilon > 0$,

$$\lim_{n\to+\infty} P\Big\{\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big\} = 0.$$

換言之, 隨機變數序列 $\{S_n/n\}_n$ 機率收斂於退化之隨機變數 p. □

例 3. 設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $N(0, 1)$ 分配. 由常態分配加法性定理之系知 $\overline{X}_n = \dfrac{1}{n}\sum\limits_{j=1}^{n} X_j$ 具 $N(0, 1/n)$, 試證：$\{\overline{X}_n\}_n$ 分配收斂於 $X = 0$.

證 首先, X 之分配函數爲

$$F \colon \mathbb{R} \to [0, 1] : F(x) = \begin{cases} 0, & \text{若 } x < 0. \\ 1, & \text{若 } x \geq 0. \end{cases}$$

而 $\overline{X}_n$ 之分配函數爲

$$\begin{aligned} F_n(x) &= P\{\overline{X}_n \leq x\} \\ &= P\{\sqrt{n}\,\overline{X}_n \leq \sqrt{n}x\} \\ &= \Phi(\sqrt{n}x), \ (\text{因 } \sqrt{n}\,\overline{X}_n \sim N(0,1)). \end{aligned}$$

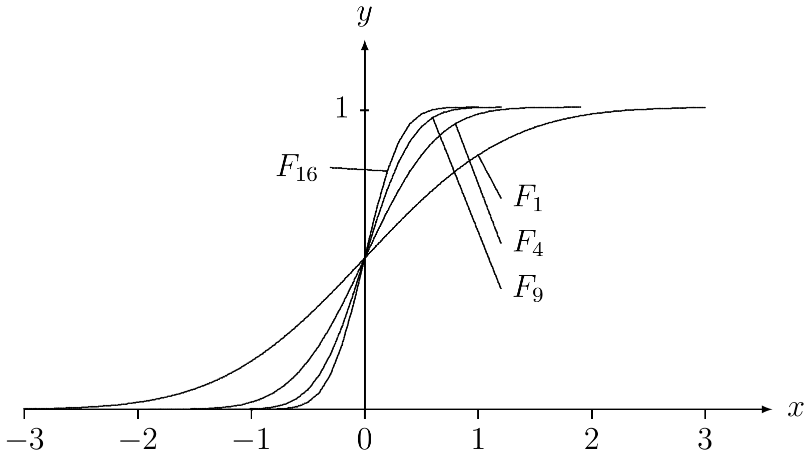


圖 7–3

故

$$\lim_{n\to+\infty} F_n(x) = \begin{cases} 0, & 若\ x < 0, \\ 1/2, & 若\ x = 0, \\ 1, & 若\ x > 0, \end{cases}$$

顯然 $\forall x \in C(F) = \mathbb{R}^*, \lim_{n\to+\infty} F_n(x) = F(x)$. 是以 $\overline{X}_n \xrightarrow{\mathrm{d}} X$. 參閱圖 7–3. □

例 4. 設 $\Omega = [0,\,1], P$ 爲 Ω 上之 Lebesgue 測度, 次設

$$\begin{cases} A_1 = (0,\ 1/2), \\ A_2 = (1/2,\ 1), \end{cases}$$

$$\begin{cases} A_3 = (0,\ 1/4), \\ A_4 = (1/4,\ 2/4), \\ A_5 = (2/4,\ 3/4), \\ A_6 = (3/4,\ 4/4), \end{cases}$$

$$\begin{cases} A_7 = (0,\ 1/8), \\ A_8 = (1/8,\ 2/8), \\ \quad\vdots \\ A_{14} = (7/8,\ 8/8), \end{cases}$$

$$A_{15} = (0,\ 1/16),$$
$$\vdots$$

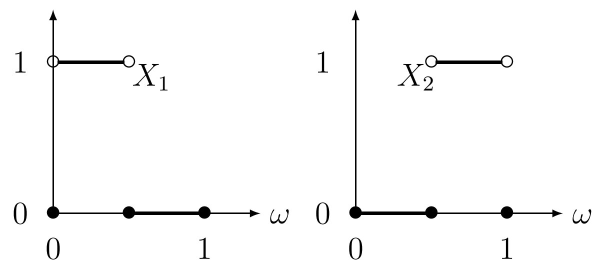


圖 7–4

再設隨機變數 $X_n = I_{A_n}, \forall n \in \mathbb{N}$, 則

$$E|X_n| = E[X_n] = E[I_{A_n}] = P(A_n) = 1/2^m,$$

如果 $n = 2^m + k - 2, (k \in \{1, \cdots, 2^m\})$. (參見下表),

m	k	n	$E[X_n]$
1	1	$2+1-2=1$	$1/2$
	2	$2+2-2=2$	$1/2$
2	1	$2^2+1-2=3$	$1/4$
	2	$2^2+2-2=4$	$1/4$
	3	$2^2+3-2=5$	$1/4$
	4	$2^2+4-2=6$	$1/4$
3	1	$2^3+1-2=7$	$1/8$
	2	$2^3+2-2=8$	$1/8$
	$\vdots$	$\vdots$	$\vdots$
	8	$2^3+8-2=14$	$1/8$
4	$\vdots$	$\vdots$	$1/16$

顯然 $\lim\limits_{n\to+\infty} E|X_n| = 0$. 是以 $\{X_n\}_n$ 平均收斂於 0. □

定理 7.8

在定義 7.7 中, 除分配收斂外, 其餘各收斂皆爲殆必唯一.

$\boxed{\text{證}}$ 我們將只證明殆必收斂之情形, 其餘部分讀者自證之. 若 $X_n \xrightarrow{\text{a.s.}} X$ 且 $X_n \xrightarrow{\text{a.s.}} Y$, 我們將證明 $X = Y$ a.s. 由於存在 零機率之事件 N_1, N_2 使得

$$\begin{cases} \forall\omega \in \Omega \setminus N_1, \lim\limits_{n\to+\infty} X_n(\omega) = X(\omega), \\ \forall\omega \in \Omega \setminus N_2, \lim\limits_{n\to+\infty} X_n(\omega) = Y(\omega). \end{cases}$$

只需令 $N = N_1 \cup N_2$, 則

$$\begin{aligned} \forall\omega \in \Omega \setminus N &\Rightarrow \omega \in ((\Omega \setminus N_1) \cap (\Omega \setminus N_2)) \\ &\Rightarrow \lim_{n\to+\infty} X_n(\omega) = X(\omega), \lim_{n\to+\infty} X_n(\omega) = Y(\omega) \\ &\Rightarrow X(\omega) = Y(\omega). \end{aligned}$$

是以 $X = Y$ a.s. □

部分初學之讀者對於殆必唯一 (a.s. unique) 與唯一 (unique) 感到困惑, 我們說明如下 :

一、在機率論中, 最常見之函數爲隨機變數 (及隨機向量), 對於二隨機變數之關係我們只關心其是否爲『殆必相等』, 而不問其是否相等, 因爲對於 $\{X \neq Y\} \neq \varnothing$ 但其機率爲 0, 以『函數』之觀點, 此二隨機變數並不相等, 但在『機率論』的角度, 不相等之事件 (集合) 之機率爲 0, 表示此一差異毫無意義, 完全可以忽略, 因此在機率論中, 我們只要求 "殆必如何" 就可以了. 同理, 當我們討論二隨機變數之大小, 也只關心是否『$X < Y$ a.s.』而不關心是否 $X < Y$.

二、集合論中,「 滿足條件 C 之函數爲唯一 」表示 :

$$\begin{cases} (1) & \text{存在一函數 } f \text{ 滿足條件 } C; \\ (2) & \text{若函數 } g \text{ 滿足條件 } C, \text{ 則 } f = g; \end{cases}$$

機率論中,「 滿足條件 C 之隨機變數爲殆必唯一 」表示 :

$$\begin{cases} (1) & \text{存在一隨機變數 } X \text{ 滿足條件 } C; \\ (2) & \text{若函數 } Y \text{ 滿足條件 } C, \text{ 則 } X = Y \text{ a.s.} \end{cases}$$

以下我們將探討各種收斂性之間的關係, 其中有些較強有些較弱.

定理 7.9

隨機變數序列$\{X_n\}_n$ 殆必收斂於 X 充要條件爲

$$\forall \epsilon > 0, \ \lim_{n \to +\infty} P\Big(\bigcup_{m=n}^{+\infty} \{|X_m - X| \geq \epsilon\} \Big) = 0.$$

證 設

$$\begin{aligned} A &= \{\omega \in \Omega \mid \lim_{n \to +\infty} X_n(\omega) = X(\omega)\} \\ &= \{\omega \in \Omega \mid \forall \epsilon > 0, \ \exists n \in \mathbb{N}, \ (\forall m)(m \geq n \ \Rightarrow \ |X_n(\omega) - X(\omega)| < \epsilon)\}, \end{aligned}$$

$D = \Omega \setminus A$, 則

$$\begin{aligned} D &= \{\omega \in \Omega \mid \exists \epsilon > 0, \ \forall n \in \mathbb{N}, \ (\exists m)(m \geq n \wedge |X_n(\omega) - X(\omega)| \geq \epsilon)\} \\ &= \bigcup_{\epsilon > 0} \bigcap_{n \geq 1} \bigcup_{m \geq n} \{|X_m - X| \geq \epsilon\} = \bigcup_{\epsilon > 0} B(\epsilon), \end{aligned}$$

內 $B(\epsilon)=\lim\limits_{n\to+\infty}\bigcup\limits_{m\geq n}\{|X_m-X|\geq\epsilon\}$, 是以

$$\begin{aligned}X_n\xrightarrow{\text{a.s.}}X &\Leftrightarrow P(A)=1\\ &\Leftrightarrow P(D)=0\\ &\Leftrightarrow \forall\epsilon>0, P(B(\epsilon))=0\\ &\Leftrightarrow \forall\epsilon>0, P\Big(\lim_{n\to+\infty}\bigcup_{m\geq n}\{|X_m-X|\geq\epsilon\}\Big)=0\\ &\Leftrightarrow \forall\epsilon>0, \lim_{n\to+\infty}P\Big(\bigcup_{m=n}^{+\infty}\{|X_m-X|\geq\epsilon\}\Big)=0. \qquad\square\end{aligned}$$

定理 7.10

設 $\{X_n\}_n$ 爲 $(\Omega,\,\mathcal{F},\,P)$ 上一隨機變數序列, X 爲 $(\Omega,\,\mathcal{F},\,P)$ 上一隨機變數, 則

(1) $X_n\xrightarrow{\text{a.s.}}X \;\Rightarrow\; X_n\xrightarrow{\text{P}}X$;

(2) $X_n\xrightarrow{L_r}X \;\Rightarrow\; X_n\xrightarrow{\text{P}}X$;

(3) $X_n\xrightarrow{\text{P}}X \;\Rightarrow\; X_n\xrightarrow{\text{d}}X$;

(4) 若 X 爲一退化隨機變數, 則 $X_n\xrightarrow{\text{P}}X\Leftrightarrow X_n\xrightarrow{\text{d}}X$.

證 (1) 由定理 7.9 立即可得 $X_n\xrightarrow{\text{a.s.}}X \;\Rightarrow\; X_n\xrightarrow{\text{P}}X$.

(2) 由第四章基本不等式知,

$$0\leq P\{|X_n-X|\geq\epsilon\}=P\{|X_n-X|^r\geq\epsilon^r\}\leq\frac{E|X_n-X|^r}{\epsilon^r},\forall\epsilon>0.$$

若 $\lim\limits_{n\to+\infty}E|X_n-X|^r=0$, 顯然有

$$\forall\epsilon>0,\ \lim_{n\to+\infty}P\{|X_n-X|\geq\epsilon\}=0.$$

亦即 $X_n\xrightarrow{L_r}X \;\Rightarrow\; X_n\xrightarrow{\text{P}}X$.

(3) 若 $X_n\xrightarrow{\text{P}}X$, 由於

$$\begin{aligned}\{|X_n-X|<\epsilon\}\cap\{X_n>x\}&=\{-\epsilon<X-X_n<\epsilon,\ X_n>x\}\\ &\subset\{X>X_n-\epsilon,\ X_n-\epsilon>x-\epsilon\}\\ &\subset\{X>x-\epsilon\},\end{aligned}$$

是以

$$\{X\leq x-\epsilon\}\subset\{|X_n-X|\geq\epsilon\}\cup\{X_n\leq x\},$$

故

$$\begin{aligned} F(x-\epsilon) &= P\{X \le x-\epsilon\} \\ &\le P\{|X_n - X| \ge \epsilon\} + P\{X_n \le x\} \\ &= P\{|X_n - X| \ge \epsilon\} + F_n(x), \end{aligned}$$

應有

$$\begin{aligned} F(x-\epsilon) &\le \liminf_{n\to+\infty} \ (P\{|X_n - X| \ge \epsilon\} + F_n(x)) \\ &= \liminf_{n\to+\infty} \ F_n(x). \end{aligned}$$

同理可證

$$\limsup_{n\to+\infty} \ F_n(x) \le F(x+\epsilon).$$

是以

$$F(x-\epsilon) \le \liminf_{n\to+\infty} \ F_n(x) \le \limsup_{n\to+\infty} \ F_n(x) \le F(x+\epsilon),$$

令 $\epsilon \to 0^+$, 因 $x \in C(F)$, 故 $\lim\limits_{n\to+\infty} F(x-\epsilon) = \lim\limits_{n\to+\infty} F(x+\epsilon) = F(x)$, 我們於是證得

$$\forall x \in C(F), \lim_{n\to+\infty} F_n(x) = F(x).$$

(4) 我們只需證明：若 X 爲一退化隨機變數, 則 $X_n \xrightarrow{\text{d}} X \ \Rightarrow \ X_n \xrightarrow{\text{P}} X$. 由於存在 $c \in \mathbb{R}$ 使得 $X = c$ a.s., X 之分配 函數爲

$$F\colon \mathbb{R} \to [0,\,1] : F(x) = \begin{cases} 0, & \text{若 } x < c, \\ 1, & \text{若 } x \ge c. \end{cases}$$

而且 $\forall \epsilon > 0,\ \forall n \in \mathbb{N}$,

$$\begin{aligned} P\{|X_n - X| < \epsilon\} &= P\{|X_n - c| < \epsilon\} \\ &= P\{c-\epsilon < X_n < c+\epsilon\} \\ &= P\{X_n < c+\epsilon\} - P\{X_n \le c-\epsilon\} \\ &\ge F_n(c+\epsilon/2) - F_n(c-\epsilon). \end{aligned}$$

又因

$$\lim_{n\to+\infty} (F_n(c+\epsilon/2) - F_n(c-\epsilon)) = F(c+\epsilon/2) - F(c-\epsilon) = 1.$$

故知

$$\forall \epsilon > 0, \quad \lim_{n\to+\infty} P\{|X_n - X| < \epsilon\} = 1.$$

□

例 1. 舉一例以說明陳述 “ $X_n \xrightarrow{L_r} X \ \Rightarrow \ X_n \xrightarrow{\text{a.s.}} X$” 不爲眞.

解 在例 4 中, 我們已知該隨機變數序列 $\{X_n\}_n$ 平均收斂於 0, 但因

$$
\begin{aligned}
\{\lim_{n\to+\infty} X_n = 0\} &= \{\omega \in \Omega \mid \lim_{n\to+\infty} X_n(\omega) = 0\} \\
&= \{k/2^m \mid m \in \mathbb{N}, k \in \{0, 1, \cdots, 2^m\}\}
\end{aligned}
$$

顯然爲可數, 是以

$$P\{\lim_{n\to+\infty} X_n = 0\} = 0.$$

因此 $X_n \xrightarrow{\text{a.s.}} 0$ 不成立. □

§ 7.3 中央極限定理

本章第一節中, 我們在跟隨機率論發展史的步伐時, 很清晰地發現, 爲了解決二項分配中 $\binom{n}{x}p^x q^{n-x}$ 之近似値, 每隔一段時日, 就有很精采的結果被發現, Laplace 定理便是其中之一, 此一結果利用第二節的觀念來闡釋, 則是

$$\frac{S_n - np}{\sqrt{npq}} \xrightarrow{\text{d}} Y \sim N(0,\, 1).$$

十九世紀末的機率與統計學者發現這種現象不僅存在於二項分配, 許多分配的樣本平均 $\overline{X}_n$ 經過標準化之後, 其分配都和 $N(0,\, 1)$ 很相似, 二十世紀初 Liapunov 已經能證明此一現象, 不過到了 1922 年 Lindeberg 才以更弱的條件敍述出所謂的中央極限定理, 稍後, Lévy 則利用特徵函數的方法予以證明, 其後 Feller 還有更進步更複雜的結果.

由於中央極限定理之證明的需要, 我們先介紹以下兩個準備定理.

定理 7.11 （**Lévy 連續性定理**）

設 $\forall n \in \mathbb{N}$, X_n 爲 $(\Omega,\, \mathcal{F},\, P)$ 上之隨機變數, ϕ_n 爲 X_n 之特徵函數.

(1) 若 $\{X_n\}_n$ 分配收斂於某隨機變數 X, 則 $\forall t \in \mathbb{R}$, $\lim\limits_{n\to+\infty} \phi_n(t) = \phi_X(t)$;

(2) 若 $\lim\limits_{n\to+\infty} \phi_n(t) = \phi(t)$, $\forall t \in \mathbb{R}$, 且 ϕ 連續於原點, 則存在一隨機變數 X 使得 $X_n \xrightarrow{\text{d}} X$ 且 $\phi = \phi_X$.

證 由於證明需要較深理論（Hally-Bray 定理）, 只好從略, 有興趣者 Loève [11] §13. □

引理 7.12 （**Polyà 引理**）

> 設 F 及 F_n 均爲分配函數, $\forall n \in \mathbb{N}$. 若 F 爲連續且 $\forall x \in \mathbb{R}$, $\lim\limits_{n\to+\infty} F_n(x) = F(x)$, 則 $\{F_n\}_n$ 均勻收斂於 F.

證 我們將證明：

$$\forall \epsilon > 0,\ \exists n_0 \in \mathbb{N},\ (\forall n)(n \geq n_0 \ \Rightarrow\ \forall x \in \mathbb{R},\ |F_n(x) - F(x)| < \epsilon).$$

1° 由分配函數之性質知

$$\forall \epsilon > 0,\ \exists \alpha, \beta \in \mathbb{R},\ \alpha < \beta \text{ 使得 } F(\alpha) < \epsilon/2,\ F(\beta) > 1 - \epsilon/2.$$

此外, F 連續於區間 $[\alpha,\, \beta]$ 上, 故必爲均勻連續, 亦即

$$\forall \epsilon > 0,\ \exists x_1, \cdots, x_r,\ \alpha = x_1 < \cdots < x_r = \beta \text{ 使得 } F(x_{j+1}) - F(x_j) < \epsilon/2.$$

2° 由原設知 $\forall j \in \{1, \cdots, r\}$, $\lim\limits_{n\to+\infty} F_n(x_j) = F(x_j)$. 故

$$\forall \epsilon > 0,\ \exists n_j \in \mathbb{N},\ (\forall n)(n \geq n_j \ \Rightarrow\ |F_n(x)_j - F(x_j)| < \epsilon/2).$$

令 $n_0 = \max\{n_1, \cdots, n_r\}$, 我們將證明：

$$\forall n \geq n_0 \ \Rightarrow\ \forall x \in \mathbb{R},\ |F_n(x) - F(x)| < \epsilon.$$

21 若 $x \in [\alpha,\, \beta]$, 則存在 $j \in \{1, \cdots, r-1\}$ 使得 $x \in [x_j,\, x_{j+1}]$. 故

$$\begin{aligned}
F_n(x) - F(x) &\leq F_n(x_{j+1}) - F(x) \\
&\leq F(x_{j+1}) + \frac{\epsilon}{2} - F(x) \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \\
F(x) - F_n(x) &= (F(x) - F(x_j)) + (F(x_j) - F_n(x)) \\
&\leq F(x_{j+1}) - F(x_j) + F_n(x_j) + \frac{\epsilon}{2} - F_n(x) \\
&< \frac{\epsilon}{2} + F_n(x) + \frac{\epsilon}{2} - F_n(x) = \epsilon.
\end{aligned}$$

亦即 $|F_n(x) - F(x)| < \epsilon$.

22 若 $x < \alpha$, 則

$$\begin{aligned}
F_n(x) - F(x) &\leq F_n(\alpha) \leq F(\alpha) + \frac{\epsilon}{2} < \epsilon; \\
F(x) - F_n(x) &\leq F(x) \leq F(\alpha) < \frac{\epsilon}{2}.
\end{aligned}$$

亦即 $|F_n(x) - F(x)| < \epsilon$.

23 若 $x > \beta$, 同理可證 $|F_n(x) - F(x)| < \epsilon$. $\square$

定理 7.13 [**Lindeberg-Lévy 中央極限定理 (Central limit theorem) 簡爲 CLT**]

> 設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配, $\mu = E[X_1]$, $\sigma^2 = \operatorname{Var} X_1 > 0$, 均爲有限. 次設 $\forall n \in \mathbb{N}$, G_n 爲隨機變數 $\dfrac{\overline{X}_n - \mu}{\sigma}\sqrt{n}$ 之分配函數, 內 $\overline{X}_n = \dfrac{1}{n}\sum_{j=1}^{n} X_j$.
> 則 $\{G_n\}_n$ 均勻收斂於 $N(0, 1)$ 之分配函數 Φ.

註： 若令 $S_n = \sum_{j=1}^{n} X_j$, 則因 S_n 之標準化與 $\overline{X}_n$ 之標準化相同, 本定理之意乃爲

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \xrightarrow{\mathrm{d}} N(0, 1),$$

但定理結論之均勻收斂性較上述簡單收斂性爲強.

證 設 ϕ_n 爲隨機變數 $\dfrac{\overline{X}_n - \mu}{\sigma}\sqrt{n}$ 之特徵函數. 我們若能證明:

$$\forall t \in \mathbb{R}, \ \lim_{n\to+\infty} \phi_n(t) = \exp\Big(-\frac{t^2}{2}\Big), \tag{1}$$

則由 Lévy 連續性定理知, $\dfrac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \xrightarrow{\mathrm{d}} N(0, 1)$, 亦即

$$\lim_{n\to+\infty} G_n(x) = \Phi(x), \ \forall x \in \mathbb{R}.$$

又由 Polyà 引理知, $\{G_n\}_n$ 均勻收斂於 Φ. 爲此, 令 Z_j 爲 X_j 之標準化, 即 $Z_j = \dfrac{X_j - \mu}{\sigma}$, 則

1° $Z_1, \cdots, Z_n, \cdots$ 爲獨立且均具相同之分配;

2° $E[Z_1] = 0, \ \operatorname{Var} Z_1 = 1 = E[Z_1^2]$;

3° $\dfrac{1}{\sqrt{n}}\sum_{j=1}^{n} Z_j = \dfrac{1}{\sqrt{n}}\sum_{j=1}^{n} \dfrac{X_j - \mu}{\sigma} = \dfrac{n\overline{X}_n - n\mu}{\sigma\sqrt{n}} = \dfrac{\overline{X}_n - \mu}{\sigma}\sqrt{n}.$

是以

$$\begin{aligned}
\phi_n(t) &= \phi_{(\sum_1^n Z_j)/\sqrt{n}}(t) = \phi_{\sum_1^n Z_j}\Big(\frac{t}{\sqrt{n}}\Big) \\
&= \textstyle\prod_{j=1}^{n} \phi_{Z_j}\Big(\frac{t}{\sqrt{n}}\Big), \ (\because Z_1, \cdots, Z_n \text{ 爲獨立}) \\
&= \Big(\phi_{Z_1}\Big(\frac{t}{\sqrt{n}}\Big)\Big)^n, \ (\because Z_1, \cdots, Z_n \text{ 具相同之分配}) \\
&= \Big(1 - \frac{t^2}{2n} + o\Big(\frac{t^2}{n}\Big)\Big)^n, \ (\text{見下面}^{[\star]}) \\
&= \Big(1 + \frac{1}{n}\Big(-\frac{t^2}{2} + n \cdot o\Big(\frac{t^2}{n}\Big)\Big)\Big)^n.
\end{aligned}$$

由於 $\forall t \in \mathbb{R}$,

$$\lim_{n\to+\infty}\Big(-\frac{t^2}{2}+n\cdot o\Big(\frac{t^2}{n}\Big)\Big)=\lim_{n\to+\infty}\Big(-\frac{t^2}{2}+\frac{o(\frac{t^2}{n})}{\frac{1}{n}}\Big)=-\frac{t^2}{2},$$

利用引理 7.5,

$$\lim_{n\to+\infty}\phi_n(t)=\exp\Big(-\frac{t^2}{2}\Big).$$

[⋆] 由於 $E[Z_1]=0,\ E[Z_1^2]=1$, 知 ϕ_{Z_1} 在原點爲二階可微, 利用微積分之 Taylor 定理, ϕ_{Z_1} 可展爲

$$\begin{aligned}\phi_{Z_1}\Big(\frac{t}{\sqrt{n}}\Big)&=\phi_{Z_1}(0)+\frac{t}{\sqrt{n}}\phi'_{Z_1}(0)+\frac{1}{2!}\Big(\frac{t}{\sqrt{n}}\Big)^2\phi''_{Z_1}(0)+o\Big(\frac{t^2}{n}\Big)\\&=1+\frac{t}{\sqrt{n}}\cdot i\cdot E[Z_1]+\frac{t^2}{2n}\cdot i^2\cdot E[Z_1^2]+o\Big(\frac{t^2}{n}\Big)\\&=1-\frac{t^2}{2n}+o\Big(\frac{t^2}{n}\Big).\end{aligned}$$

(參見習題 5-12). □

系 7.14

> 設 $\{X_n\}_n$ 爲獨立且均具相同之分配, $\mu=E[X_1],\ \sigma^2=\operatorname{Var}X_1>0$, 均爲有限. 次設 $\forall n\in\mathbb{N},\ S_n=\sum_{j=1}^{n}X_j$, 則 S_n 具漸近常態分配 $N(n\mu,\,n\sigma^2)$, 意即 : 若 $\{a_n\}_n$ 爲一實數序列, 則
>
> $$\lim_{n\to+\infty}(P\{S_n\le a_n\}-P\{Y_n\le a_n\})=0,$$
>
> 內 Y_n 爲具 $N(n\mu,\,n\sigma^2)$ 之隨機變數.

證 由於

$$\begin{aligned}&|P\{S_n\le a_n\}-P\{Y_n\le a_n\}|\\&\quad=\Big|P\Big\{\frac{S_n-n\mu}{\sqrt{n}\sigma}\le\frac{a_n-n\mu}{\sqrt{n}\sigma}\Big\}-P\Big\{\frac{Y_n-n\mu}{\sqrt{n}\sigma}\le\frac{a_n-n\mu}{\sqrt{n}\sigma}\Big\}\Big|\\&\quad=\Big|P\Big\{\frac{S_n-n\mu}{\sqrt{n}\sigma}\le a_n^*\Big\}-P\Big\{\frac{Y_n-n\mu}{\sqrt{n}\sigma}\le a_n^*\Big\}\Big|\\&\qquad\qquad\Big(\text{其中 } a_n^*=\frac{a_n-n\mu}{\sqrt{n}\sigma}\Big)\\&\quad=|G_n(a_n^*)-\Phi(a_n^*)|,\ (G_n\text{ 見中央極限定理})\\&\quad\le\sup_{x\in\mathbb{R}}|G_n(x)-\Phi(x)|.\end{aligned}$$

但由中央極限定理知 G_n 均勻收斂於 Φ, 亦即 $\sup_{x\in\mathbb{R}}|G_n(x)-\Phi(x)| \to 0$, 當 $n \to +\infty$, 故證得,

$$\lim_{n\to+\infty}(P\{S_n \le a_n\} - P\{Y_n \le a_n\}) = 0. \qquad \square$$

註： **1.** 本系之用途爲：當 n 很大時, 我們可用 $\Phi(a_n^*)$ 做爲 $P\{S_n \le a_n\}$ 之近似值. 當然, 若 $a_n < b_n$, 則因

$$P\{a_n < S_n \le b_n\} = P\{S_n \le b_n\} - P\{S_n \le a_n\}.$$

所以我們可以 $\Phi(b_n^*) - \Phi(a_n^*)$ 做爲上述機率之近似值, 其中

$$a_n^* = \frac{a_n - n\mu}{\sqrt{n}\sigma}, b_n^* = \frac{b_n - n\mu}{\sqrt{n}\sigma}.$$

2. 如果我們比較 Laplace 定理之系, 與註 1 之結果, 我們不難發現二者有些差異. 設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$, 由 Laplace 定理之系知, $\forall\alpha,\ \beta \in \{0, 1, \cdots, n\}$,

$$P\{\alpha \le S_n \le \beta\} \approx \Phi\Big(\frac{\beta + \frac{1}{2} - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{\alpha - \frac{1}{2} - np}{\sqrt{npq}}\Big),$$

由 CLT 之系, 則

$$\begin{aligned} P\{\alpha \le S_n \le \beta\} &= P\{\alpha - 1 < S_n \le \beta\} \\ &\approx \Phi\Big(\frac{\beta - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{\alpha - 1 - np}{\sqrt{npq}}\Big), \end{aligned}$$

二者雖然形不同, 但均無錯誤. 只是當 n 不是很大時, 則前者之近似效果較後者爲佳.

例 1. (參閱第一節例 1) 連擲一公正之銅板 100 次, 試以中央極限定理求正面出現 30 至 45 次之機率爲若干 ?

$\boxed{解}$ 由 CLT 之系知, $P\{30 \le S_n \le 45\} = P\{29 < S_n \le 45\}$ 之近似值爲

$$\Phi\Big(\frac{45 - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{29 - np}{\sqrt{npq}}\Big).$$

因 $n = 100, p = q = 1/2$, 故上述近似值爲

$$\begin{aligned} \Phi\Big(\frac{45 - 50}{\sqrt{25}}\Big) - \Phi\Big(\frac{29 - 50}{\sqrt{25}}\Big) &= \Phi(-1) - \Phi(-4.2) \\ &= 1 - \Phi(1) - 0.000030 \\ &= 0.158625. \end{aligned}$$

爲了便於比較, 我們將第一節例 1 之結果與上述結果並列於下:

$$\begin{cases} \text{以電算機求得 } P\{30 \le S_n \le 45\} = 0.184083, \\ \text{以 Laplace 定理求得之近似值} \quad = 0.184030, \\ \text{以中央極限定理求得之近似值} \quad = 0.158625. \end{cases}$$

如果 n 非常大 (例如 10000 以上), 則中央極限定理所得之近似值當不至太差. 總之, 欲求二項分配之近似值仍以 Laplace 定理爲佳, 有人稱 Laplace 定理之方法爲**連續性校正**(continuity correction). □

例 2. 自區間 [0, 1] 中隨機抽取 48 數. 試問這 48 個樣本之平均小於 0.45 之機率爲若干 ?

解 設 X_j 爲第 j 次抽取所得之數. 則本題可視爲 $X_1, \cdots, X_{48}$ 爲獨立且均具 $U(0,\, 1)$ 分配. 由題意, 應求 $P\{\overline{X} < 0.45\} =?$ 其中 $\overline{X} = (X_1 + \cdots + X_{48})/48$. 因 $EX_1 = 1/2$, $\text{Var } X_1 = 1/12$, 由中央極限定理知

$$\begin{aligned} P\{\overline{X} < 0.45\} &= P\Big\{\frac{\overline{X} - 0.5}{\sqrt{1/12}} \cdot \sqrt{48} < \frac{0.45 - 0.5}{\sqrt{1/12}} \cdot \sqrt{48}\Big\} \\ &\approx \Phi(-1.2) \\ &= 1 - \Phi(1.2) \\ &= 1 - 0.884930, (\text{查表或統計軟體}) \\ &= 0.115070. \end{aligned}$$

是以這 48 個樣本之平均小於 0.45 之機率爲 0.115070. □

例 3. 設 $X_1, \cdots, X_n$ 爲獨立且均具 Cauchy 分配 $C(0, 1)$. 試問其樣本平均 $\overline{X}_n$ 之絕對值小於 1 之機率爲若干 ?

解 由於 Cauchy 分配之期望值不存在, 不可利用中央極限定理以求 $P\{|\overline{X}_n| < 1\}$ 之近似值. 已知 $\phi_{X_1}(t) = e^{-|t|}$, (第五章第二節例 6), 故

$$\phi_{X_1+\cdots+X_n}(t) = (e^{-|t|})^n = \exp(-n|t|),$$

是以

$$\phi_{\overline{X}_n}(t) = \phi_{X_1+\cdots+X_n}\Big(\frac{t}{n}\Big) = \exp(-|t|).$$

此乃說明 $\overline{X}_n \sim C(0, 1)$. 故所求之機率

$$\begin{aligned} P\{|\overline{X}_n| < 1\} &= \frac{1}{\pi}\int_{-1}^{1} \frac{dx}{1+x^2} = \frac{1}{\pi}\tan^{-1} x \,\Big|_{-1}^{1} \\ &= \frac{1}{\pi}\Big(\frac{\pi}{4} + \frac{\pi}{4}\Big) = \frac{1}{2}. \end{aligned}$$

□

§ 7.4 大數定律

在機率論的歷史中, 研究的重點之一是二項分配的近似值, 當 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 時, 我們已知樣本平均 $\overline{X}_n$ 標準化之後與 $N(0, 1)$ 十分接近而發展出所謂中央極限定理. 另外一條研究路線則是, $\overline{X}_n$ 不加標準化而尋求其極限. 最早的定理是 Bernoulli 的弱大數定律, (詳見定理 7.1). 二十世紀初有了一些重要進展, 首先, E. Borel 發現這種收斂可以強些, 我們敍述如下：

定理 7.15 [**Borel 強大數定律 (Strong law of large numbers)**]

> 設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具 $B(1, p)$ 分配之隨機變數序列, 次設 $S_n = X_1 + \cdots + X_n, \overline{X}_n = S_n/n$, 則 $\overline{X}_n \xrightarrow{\text{a.s.}} p$.

證 參閱 Loève [11], p.19. □

Bernoulli 弱大數定律說：$B(1, p)$ 的樣本平均 $\overline{X}_n$ 與 p 之差小於 ϵ 之機率的極限等於 1, (其中 ϵ 爲任意正數), 即

$$\lim_{n\to+\infty} P\{|\overline{X}_n - p| < \epsilon\} = 1.$$

但 Borel 強大數定律則直截了當地說 $\overline{X}_n$ 殆必收斂於 p. (換句話說, 樣本平均的極限就是 p). 後來 Chebyshev 不等式出現了, 這種“大數”現象就從二項分配推廣到一般分配.

定理 7.16 [**Chebyshev 弱大數定律 (Weak law of large numbers)**]

> 設 $X_1, \cdots, X_n, \cdots$ 爲獨立之隨機變數序列, $S_n = X_1 + \cdots + X_n$, 若 $\{\text{Var } X_n\}_n$ 爲有界, 則
> $$\frac{S_n - ES_n}{n} \xrightarrow{\text{P}} 0.$$

證 利用 Chebyshev 不等式, $\forall \epsilon > 0$,

$$\begin{aligned} P\left\{\left|\frac{S_n - ES_n}{n}\right| \geq \epsilon\right\} &\leq \frac{E\left|\frac{S_n - ES_n}{n}\right|^2}{\epsilon^2} \\ &= \frac{\text{Var } S_n}{n^2\epsilon^2} \\ &= \frac{1}{n^2\epsilon^2}\sum_{j=1}^{n} \text{Var } X_j \\ &\leq \frac{c}{n\epsilon^2}, \quad (\because \text{ Var } X_j \leq c < +\infty) \end{aligned}$$

令 $n \to +\infty$, 則得所欲證. □

系 7.17

設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配, 若 $\mathrm{Var}\ X_1 < +\infty$, 則

$$\overline{X}_n \xrightarrow{\mathrm{P}} \mu = E[X_1].$$

內 $\overline{X}_n$ 爲 $X_1, \cdots, X_n$ 之樣本平均.

定理 7.18 (**Khintchine 弱大數定律**)

設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配, 若 $\mu = E[X_1]$ 爲有限, 則

$$\overline{X}_n \xrightarrow{\mathrm{P}} \mu.$$

證 由於常數隨機變數 μ 特徵函數爲 $e^{it\mu}$, 故只 需證明 $\overline{X}_n$ 之特徵函數 ϕ_n 收斂於 $e^{it\mu}$. 因爲

$$\phi_n(t) = \phi_{X_1+\cdots+X_n}\Big(\frac{t}{n}\Big) = \Big(\phi_{X_1}\Big(\frac{t}{n}\Big)\Big)^n,$$

而 ϕ_{X_1} 之一階 Maclaurin 展開式爲 (習題 5-12)

$$\phi_{X_1}(t) = 1 + it\mu + o(t),$$

是以

$$\phi_n(t) = \Big(1 + \frac{it\mu}{n} + o\Big(\frac{t}{n}\Big)\Big)^n.$$

仿照中央極限定理之證法, 我們可證得

$$\lim_{n\to+\infty} \phi_n(t) = e^{it\mu}. \qquad \square$$

如果我們仔細比較上述兩個弱大數定律 (WLLN) :

1° Khintchine's WLLN 並未較 Chebyshev's WLLN 爲廣, 因爲 Khintchine's WLLN 之前提爲 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配; 但 Chebyshev's WLLN 之前提並不需要各 X_n 具相同之分配。

2° Chebyshev's WLLN 亦未較 Khintchine's WLLN 爲廣, 因爲 Khintchine's WLLN 之前提不需要 $\{\mathrm{Var}\ X_n\}_n$ 爲有界.

3° 如果已知 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配, 則 Khintchine's WLLN 較 Chebyshev's WLLN 爲廣.

在這一切的發展中, 大家都期待 Khintchine 和 Borel 大數定律有個共同的推廣結果, 那就是

$$\overline{X}_n \xrightarrow{\text{a.s.}} \mu = EX_1.$$

此一願望終被 Kolmogorov 所實現. 由於證明所需要之技巧十分繁複而準備定理很多, 我們只能敘述這個世紀性的定理.

定理 7.19 [**Kolmogorov 強大數定律 (Strong law of large numbers)**]

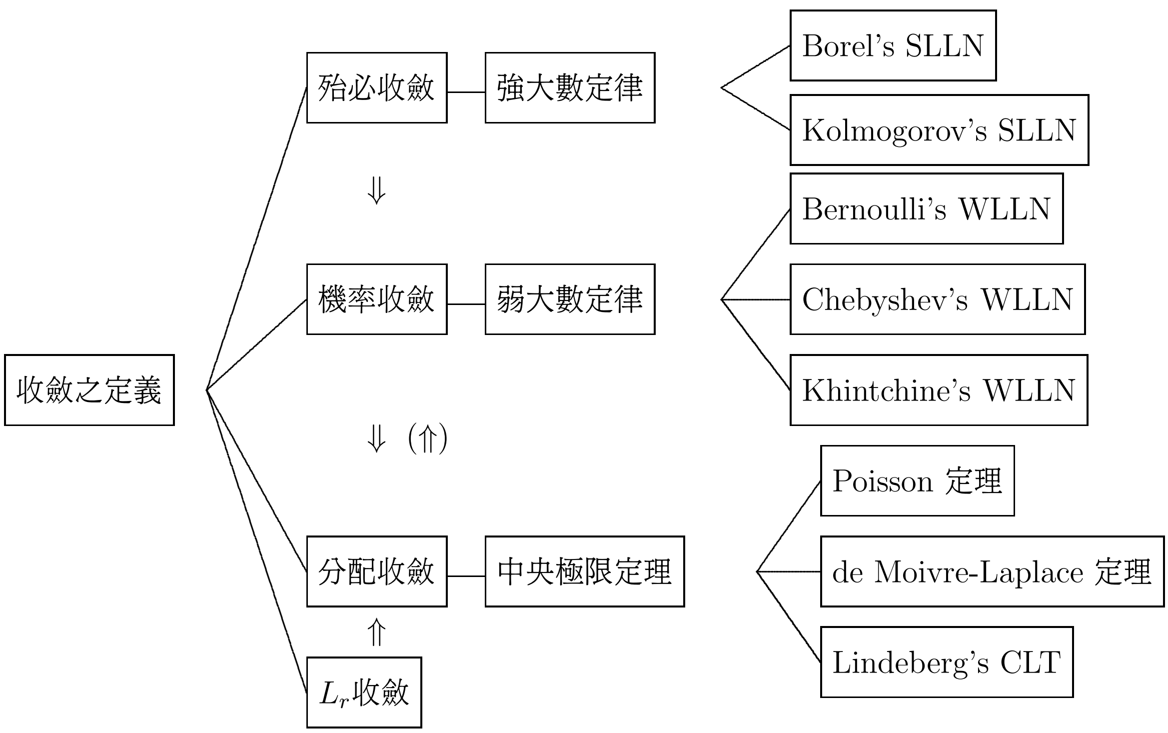
設 $X_1, \cdots, X_n, \cdots$ 爲獨立且均具相同之分配, 則

$$X_1 \text{ 爲可積 } \Leftrightarrow \overline{X}_n \xrightarrow{\text{a.s.}} \mu.$$

此時, $\mu = E[X_1]$.

證 參閱 Loève [11], §17. □

收斂之系統圖



第七章　習　　題

7-1 將一公正之骰連擲 30 次, S_{30} 表出現點數之總和, $A=\{90<S_{30}<120\}$.

(1) 試求 S_{30} 之期望值及變異數.

(2) 利用 Chebyshev 不等式, 試問 A 之機率至少爲若干 ?

(3) 利用中央極限定理, 試求 A 之機率之近似值;

(4) 除中央極限定理之外, 尚有什麼方法可求出 $P(A)$?

7-2 設 $(\Omega,\,\mathcal{F},\,P)$ 爲一機率空間, $\{A_n\}_n$ 爲一事件序列且 $\lim\limits_{n\to+\infty} A_n=A$, 試證 : $I_{A_n}\xrightarrow{\text{a.s.}} I_A$, (內 I_A 爲 A 之指標函數.)

7-3 設隨機變數 $X_n, n\in\mathbb{N}$, 之分配函數爲

$$F_n\colon \mathbb{R}\to[0,\,1] : F_n(x)=\begin{cases} 0, & \text{若 } x\le -n,\\ (x+n)/(2n), & \text{若 } -n<x<n,\\ 1, & \text{若 } n\le x.\end{cases}$$

(1) 試問 $\{F_n\}_n$ 之點態極限 (pointwise limit) $\lim\limits_{n\to+\infty}F_n$ 是否爲一分配函數 ?

(2) 試問 $\{X_n\}_n$ 是否分配收斂於某隨機變數 X, 若是, 寫出 X 之分配函數; 若否, 敍述其理.

7-4 (1) 若 X 具 $N(0,\,\sigma^2)$, 試求 $E|X|=$?

(2) 設 $X_1,\cdots,X_n,\cdots$ 爲獨立且均具相同之分配 $N(\mu,\,\sigma^2)$, 利用 (1) 之結果, 證明 $\overline{X}_n$ 平均收斂於 μ.

7-5 若 $X_n\xrightarrow{\text{P}}X$, 證明 : X 爲殆必唯一; 意即 : 若 $X_n\xrightarrow{\text{P}}X$ 且 $X_n\xrightarrow{\text{P}}Y$, 則 $X=Y$ a.s.

7-6 將一公正之錢幣連擲 100 次, S 表出現正面之總次數.

(1) 利用 Chebyshev 不等式以求最小之整數 N 使得

$$P\{|S-50|\le N\}\ge 0.95;$$

(2) 利用 Laplace 定理以求最小之整數 N 使得

$$P\{|S-50|\le N\}\ge 0.95.$$

7-7 自區間 $[0,\,1]$ 中隨機抽取 n 個數, 平均之, 欲使此平均數在區間 $[0.49,\,0.51]$ 之間的機率爲 0.95 以上, 試問 n 至少爲若干 ?

7-8 設 $\forall n \in \mathbb{N}$, 隨機變數 X_n 之密度函數爲

$$f_n\colon \mathbb{R} \to [0,\,1] : f_n(x) = \begin{cases} 1/n, & \text{若 } x \in \{1/n,\,2/n,\,\cdots,1\}, \\ 0, & \text{否則;} \end{cases}$$

次設 X 具 $U(0,\,1)$.

(1) 直接利用分配收斂之定義, 證明 : $X_n \xrightarrow{\text{d}} X$;

(2) 利用 Lévy 連續性定理, 證明 : $X_n \xrightarrow{\text{d}} X$.

7-9 若 $X_n \xrightarrow{\text{a.s.}} X$ 且 $Y_n \xrightarrow{\text{a.s.}} Y$, 證明 : $X_n + Y_n \xrightarrow{\text{a.s.}} X + Y$.

7-10 設 $A = [0,\,1], B = [0,\,2]$, 今以放回方式輪流自 A 與 B 中取數, 即第一次自 A 中取一數, 第二次自 B 中取一數, 第三次自 A 中取一 數, ...

(1) 試問這些樣本之平均 $\overline{X}_n$ 是否爲殆必收斂 ?

(2) 利用中央極限定理以求 $P\{0.7 < \overline{X}_n < 0.8\}$ 之近似值.

7-11 在中央極限定理之前提下, 若 a 爲一常數, 試證 :

$$\lim_{n\to+\infty} P\{S_n \le a\} = \begin{cases} 0, & \text{若 } \mu > 0, \\ 1/2, & \text{若 } \mu = 0, \\ 1, & \text{若 } \mu < 0. \end{cases}$$

7-12 試證 : $\displaystyle\lim_{n\to+\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}$.

[提示] 應用中央極限定理於 Poisson 分配, $X_j \sim P(1)$.

7-13 設 X_n 具 Poisson 分配 $P(n)$, 其中 $n \in \mathbb{N}$.

(1) 試證 : $\dfrac{X_n}{n} \xrightarrow{\text{P}} 1$;

(2) 試求序列 $\{a_n\}_n,\ \{b_n\}_n$ 以使 $\dfrac{X_n - a_n}{\sqrt{b_n}} \xrightarrow{\text{d}} N(0,1)$.

7-14 有一袋子, 內置十球, 分別標以 $0, 1, \cdots, 9$, 今以“放回”方式隨機抽 取 100 個, 試求其平均數 $\overline{X}$ 大於 4.75 之機率.

7-15 利用 Lévy 連續性定理證明 : 若 $X_n \sim B(n, p_n)$, $\displaystyle\lim_{n\to+\infty} np_n = \lambda$, $Y \sim P(\lambda)$, 則 $X_n \xrightarrow{\text{d}} Y$.

參 考 資 料

[1] Aposotol, T.M., Mathematical Analysis. 2nd ed., 1975.

[2] Chow, Y.S. & Teicher, H., Probability Theory, 1978.

[3] Chung, K.L., Elementary Probability Theory with Stochastical process, 1964.

[4] Chung, K.L., A Course in Probability Theory, 2nd ed., 1974.

[5] Feller W., An Introduction to Probability Theory and its Application. Vol I, II, 3rd ed., 1967.

[6] Fisz, M., Probability Theory and Mathematical Statistics, 3rd ed., 1963.

[7] Halmos, P., Measure Theory, 1950.

[8] Khazanie, R., Basic Probability Theory and its Applications, 1976.

[9] Kolomogorov, A.N., Grundbegriffe der Wahrschein-lichleitrechnung, 1933.

[10] Laha, R.G. & Rohagi, V.K., Probability Theory, 1970.

[11] Loève, M., Probability Theory I, II, 4th ed., 1977.

[12] Prohorov, Yu. V. & Rosanov, Yu. A., Probability Theory, 1969.

[13] Ross, S.M., Introduction to Probability Models, 5th ed., 1992.

[14] Ross, S.M., A first course in Probability.

[15] Roussas, G.G., A Course in Mathematical Statistics, 2nd ed., 1997.

[16] Royden, H.L. Real Analysis, 2nd ed., 1968.

[17] Shiryayev, A.N., Probability, 1984.

[18] 國立編譯館：數學名詞, 1981.

[19] 顏國勇：微積分學, 二版, 2000.8.

答 案 與 提 示

第 一 章

1-1 (1) 隨機; $\Omega = \{0, 1, 2, \cdots\}$;
(2) 隨機; $\Omega = \{$有, 無$\}$;
(3) 命定;
(4) 隨機; $\Omega = (0,\ +\infty)$, (單位：小時);
(5) 隨機; $\Omega = (0,\ +\infty)$, (單位：分);
(6) 命定.

1-3 $\Omega = \{(H,H), (H,T), (T,1), (T,2), (T,3), (T,4), (T,5), (T,6)\}$;
$P\{(H,H)\} = P\{(H,T)\} = \frac{1}{4}, P\{(T,n)\} = \frac{1}{12}$, $n \in \{1,2,3,4,5,6\}$.

1-4 三集合之聯集之人數應爲 60 人.

1-5 (1) $\Omega = [0,3] \times [0,3]$;
(2) $5/9$.

1-6 $\mathcal{F}_1 = \mathcal{P}(\Omega),\ \mathcal{F}_2 = \{\varnothing,\ \Omega\},\ \mathcal{F}_3 = \{\varnothing,\ \{1\},\ \{2,3,4\},\ \Omega\}$.

1-7 先將 Ω 化爲四個集合之互斥聯集. 再以之產生 σ 域.

1-8 交集爲一 σ 域, 但聯集則未必.

1-9 利用產生 σ 域之定義.

1-10 (1) 利用 $(-\infty,\ a - 1/n],\ n \in \mathbb{N}$, 聯集之.
(2) 利用 1-9 題.

1-11 利用矛盾證法.

1-12 $\limsup A_n = [-1, 1],\ \liminf A_n = \{-1\}$.

1-13 (1) 利用定理 1.10;
(2) 利用定理 1.10;
(3) 例如：$A_{2n-1} = B_{2n} = [0,1],\ A_{2n} = B_{2n-1} = (1,2]$.

1-14 利用定理 1.11;

1-15 利用上下極限之定義及 De Morgan 定律.

1-16 $\limsup A_n = A \cup B,\ \liminf A_n = A \cap B$. 若 $A = B$ 則極限存在.

1-17 證明非 σ 域時, 考慮奇數所成之集合.

1-18 利用 P 之有限加法性.

1-19 利用機率測度之定義.

1-20 (1) $\Omega_1 = [-r,\ r],\ \mathcal{F}_1 = \{B \cap \Omega_1 \mid B \in \mathcal{B}\},\ P_1 \colon \mathcal{F}_1 \to [0,1] : P_1(B) = \lambda(B)/(2r)$.

1-21 (1) 利用數學歸納法;

(2) 考慮 $A_1 \cap A_2$ 之餘集之機率;

(3) 利用 (1) 及P 之連續性.

1-22 利用$\mathbb{Z}$1.21(1) 題之觀念, 證明餘集之機率爲 0.

1-23 1/4; 1/17.

1-24 0.625.

1-25 因爲 $A \cap B \cap C = \varnothing$.

1-26 先求 $P(A \cap B)$, $P(A)$, $P(B)$.

1-27 考慮 $\cup A_j$ 之餘集的機率.

1-28 (1) 證明 $P(B \cap (A_1 \cup A_2)) = P(B)P(A_1 \cup A_2)$.

(2) 舉一反例.

第　二　章

2-1 是.

2-2 $\sin(\omega) + \cos(\omega) = \sqrt{2}\sin(\omega + \frac{\pi}{4})$.

2-3 利用 X^+ 及 X^- 之定義.

2-4 討論對等關係三條件.

2-5 考慮 $|X - Y| \leq |X - Z| + |Y - Z|$.

2-6 考慮 σ 域定義.

2-7 (1) $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, $P\{(a, b)\} = 1/36$;

(2) 考慮 $\{X = j\}$ 之基數;

(3) $f\colon \mathbb{R} \to [0, 1] : f(x) = \begin{cases} \dfrac{6 - |x|}{36}, & \text{若 } x \in X(\Omega), \\ 0, & \text{否則}, \end{cases}$

(4) 13/36.

2-8 (2) 左連續.

2-9 $1 - 62p^{61}(1 - p) - p^{62}$, $p = 0.75$.

2-10 $0.99^{20}, 1 - 0.99^{20} - 20 \times 0.01 \times 0.99^{19}$.

2-11 否.

2-12 0.658.

2-13 利用 F 之二階導函數.

2-14 (1) 利用 P_X 之定義;

(2) 利用 F_X 之定義.

2-15 將 L 化爲 a 之函數, 然後以導數方法證之.

2-16 (1) $X \sim B(1, 0.5)$;

(2) 未必;

(3) 0.674.

2-17 證明左右兩端等於 $(1-p)^{m+1}$.

2-18 $f_Y(-1)=2/15, f_Y(0)=1/3, f_Y(1)=8/15$.

2-19 $f_Y(0)=0.24022, f_Y(1)=0.63541, f_Y(2)=0.12457$.

2-20 分為 $y<0,\ 0\le y\le 1,\ 1\le y$ 三種情況, 以分配函數之定義討論之.

2-21 利用變換公式.

2-22 $f_Y(y)=\begin{cases}\lambda e^{-\lambda(y-c)}, & 若\ y>c,\\ 0, & 若\ y\le c.\end{cases}$

2-23 (1) $F_Y(y)=\begin{cases}0, & 若\ y<-3,\\ F_X(y)-0.002339, & 若\ -3\le y<0,\\ F_X(y)+0.07865, & 若\ 0\le y<3,\\ 1, & 若\ 3\le y.\end{cases}$

2-24 (1) 57.62%, (3) 0.1323.

第 三 章

3-1 全部是.

3-2 皆是.

3-3 $Y\sim B(1,1/2)$.

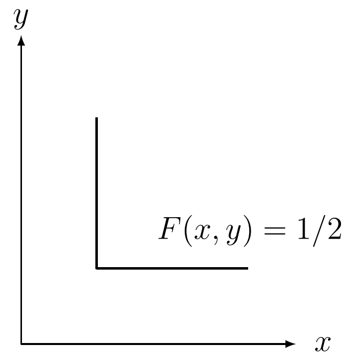
3-4 (1) 10; (2) 11/96.

3-5 意義相同但 $f_{\mathbf{X}}$ 定義於 $\mathbb{R}^2$ 上, f_X 定義於 $\mathbb{R}$ 上.

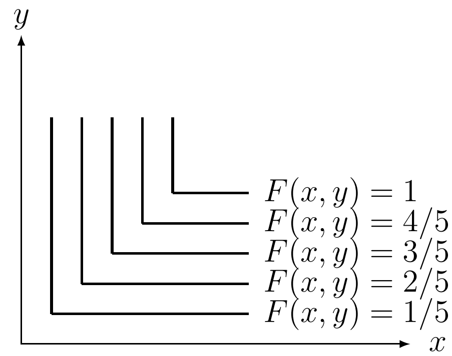
3-6 利用 P 之連續性.

3-7 (1) 考慮 $\Big(z\le x$ 且 $z\le y\Leftrightarrow z\le \min\{x,y\}\Big)$.

(2)



(3)



3-8 注意：$F(x,\,y)\le F_X(x), F(x,\,y)\le F_Y(y)$.

3-9 $G=0$ 及 $H=0$ 皆不可能.

3-10 (1) $f_{X,Y}:\mathbb{R}^2\to[0,1]$:

$$f_{X,Y}(x,y)=\begin{cases}\frac{1}{10x}, & 若\ (x,y)\in A,\\ 0, & 否則,\end{cases}$$

其中 $A=\{(x,y)\in\mathbb{N}^2\mid 1\le y\le x\le 10\}$.

(2) $f(x|3)=\begin{cases}0.6998/x, & 若\ x=3,4,\cdots,10,\\ 0, & 否則.\end{cases}$

3-11 (1) 2;

(2) 將定義域分爲三區：$0 \le x \le y$, $0 \le y \le x$ 及其餘, 再 討論 $F(x, y)$ 之值;

(4) $f_Y(y) = \begin{cases} 2e^{-y} - 2e^{-2y}, & \text{若 } y > 0, \\ 0, & \text{否則}; \end{cases}$

$f_X(x) = \begin{cases} 2e^{-2x}, & \text{若 } x > 0, \\ 0, & \text{否則}; \end{cases}$

(5) $f(x|y) = \begin{cases} e^{-x}(1 - e^{-y})^{-1}, & \text{若 } 0 \le x \le y, \\ 0, & \text{否則}. \end{cases}$

3-12 (1) 1/8 ;

(2) $f(x|y) = \begin{cases} 3(y^2 - x^2)(4y^3)^{-1}, & \text{若 } -y < x < y, \\ 0, & \text{否則}. \end{cases}$

3-13 3/4.

3-14 (1)

t	1	3	4	其他
$f_T(t)$	5/42	10/42	27/42	0

(2)

u	–4	–1	0	1	3	4	12	其他
$f_U(u)$	5/42	2/42	5/42	2/42	10/42	5/42	13/42	0

(3)

v	–4	–3	–1	0	2	其他
$f_V(v)$	4/42	10/42	14/42	4/42	10/42	0

3-15 利用 Stirling 公式 $\Gamma(x) \sim \sqrt{2\pi} \cdot e^{-x} x^{x-0.5}$.

3-17 (1) $f_Z(z) = \begin{cases} -\ln z, & \text{若 } 0 < z < 1, \\ 0, & \text{否則}. \end{cases}$

(2) $f_U(u) = \begin{cases} 2u, & \text{若 } 0 < u < 1, \\ 0, & \text{否則}. \end{cases}$

第　四　章

4-1 仿二項分配.

4-2 仿二項分配.

4-3 (1) 證明 $\int_{\mathbb{R}} f(x)\, dx = 1$.

(2) 利用 gamma 函數.

4-4 $E[X^n] = \dfrac{(n-\alpha-1)(n-\alpha-2)\cdots\alpha}{(n+\alpha+\beta-1)(n+\alpha+\beta-2)\cdots(\alpha+\beta)}$.

4-5 (1) 先將 EX 分爲 $\int_{-\infty}^{0} x\,dF(x) + \int_{0}^{+\infty} x\,dF(x)$, 再以分部積分討 論之.

4-6 (1) a; (2) $\ln 2$

4-7 (1) $P\{X = j\} = \dfrac{\binom{4}{j}\binom{6}{3-j}}{\binom{10}{3}}$; (2) 48.

4-8 $50 + \dfrac{16}{\sqrt{2\pi}}$.

4-9 仿第三節 $G(\alpha, \beta)$ 分配之期望值.

4-11 (3) 證明 $|X - \frac{a+b}{2}| \le \frac{b-a}{2}$, 再利用 (1);

(4) 甚難, 先證 $X \in \{a, b\}$ a.s.

4-12 (2) $X = 1$.

4-14 0.22.

4-15 令 Y 爲 X 之標準化, 先證 : $\int_{-\infty}^{+\infty} y^2 f(y)\,dy > 2\int_{k}^{+\infty} y^2 f(y)\,dy$, 再討論之.

4-16 不存在.

4-17 只需證明標準二元常態分配之情形.

4-18 提示 : $E[I_A] = P(A)$.

4-19 $E(X \mid Y = y) = \dfrac{2}{3} + \dfrac{1}{18y - 24}$.

4-20 (2) $(6 + 20y_0 + 15y_0^2)25^{-1}(3 + 4y_0)^{-2}$.

4-21 利用定理 4.3.

4-22 (1) $1/p$.

(2) $P\{N = x\} = q^{x-1}p,\ \forall x \in \mathbb{N}; EN = 1/p$.

第 五 章

5-1 考慮 $|Z| \le |X| + |Y|$.

5-2 左右兩端皆爲 $\sqrt{1 + m^2}\, E|X|$.

5-3 (1) $\phi(t) = \begin{cases} \dfrac{e^{it\beta} - e^{it\alpha}}{it(\beta - \alpha)}, & \text{若 } t \neq 0, \\ 1, & \text{若 } t = 0; \end{cases}$

(2) $p^r(1 - qe^{it})^{-r}$;

(3) $\dfrac{\lambda e^{it\alpha}}{\lambda - it}$.

5-4 $e^{it\mu - |\sigma t|},\ i\mu - \sigma,\ i\mu + \sigma$.

5-5 $f(x) = \begin{cases} \dfrac{1 - \cos x}{\pi x^2}, & \text{若 } x \neq 0, \\ 1/(2\pi), & \text{若 } x = 0. \end{cases}$

5-6 利用唯一性定理.

5-7 $\exp(t\mu + \frac{1}{2}\sigma^2 t^2)$.

5-8 先證 $|e^{it} - (1+it)| \le t^2/2$, 再以 tX 代 t, 取 E 即可.

5-9 0.7668.

5-10 $E[X] = 1$, $\text{Var } X = 0.5$, $C(X, Y) = 1/6$.

5-11 證明 $\phi_Y(t) = (1-2it)^{-1/2}$.

5-12 將 $\phi(t)$ 分爲實數部分與虛數部分, 再分別以 Taylor 定理討論之.

5-13 證明 $\phi_Y(t) = \begin{cases} \dfrac{e^{it}-1}{it}, & 若\ t \neq 0, \\ 1, & 若\ t = 0. \end{cases}$

第　六　章

6-1 [提示] $I_A^{-1}(\mathcal{B}) = \{\varnothing, \Omega, A, \complement A\}$.

6-2 (2) 在 $f_X(x) > 0$ 的情形下.

6-3 (1)

$F(x,y)$	$x<0$	$0 \le x < 1$	$1 \le x$
$y<0$	0	0	0
$0 \le y < 1$	0	qy	y
$1 \le y$	0	q	1

(2) 爲連續型, 但不爲絕對連續型.

6-4 (1) 1; (2) 是.

6-5 (1) q^x; (2) $F_Z(z) = \begin{cases} 1 - q^{2(z+1)}, & 若\ z \in \mathbb{Z}_+, \\ 0, & 否則. \end{cases}$

6-7 利用系 6.14.

6-8 (1) $\exp(2it\mu - 2\sigma|t|)$;

(2) 與 X 相同.

6-9 [提示] $z \in \mathbb{C}$ 時, $|z|^2 = z \cdot \bar{z}$;

6-10 [提示] e^{isX} 展爲 $\cos sX + i \sin sX$, e^{itY} 仿之.

6-11 利用 5-3 題 $\phi_j(t) = p^r(1-qe^{it})^{-r_j}$.

6-12 25/36.

6-13 1/4.

第　七　章

7-1 (1) 105, 87.5; (2) 0.6111; (3) 0.8784; (4) 利用電腦;

7-2 應分 $\omega \in A$ 及 $\omega \in \complement A$ 二情形討論之.

7-3 (1) $\lim\limits_{n\to+\infty} F_n(x) = \frac{1}{2}$, $\forall x \in \mathbb{R}$, 不一分配函數.

(2) 不爲分配函數收斂.

7-4 (1) $\sigma\sqrt{2/\pi}$;

(2) 先求 $E|\overline{X}_n - \mu|$.

7-5 考慮 $|X - Y| \le |X_n - X| + |X_n - Y|$.

7-6 (1) 22; (2) 10.

7-7 3202.

7-8 (1) 當 $0 < x < 1$ 時, $F_n(x) = \dfrac{[nx]}{n!}$; 取極限;

(2) 應分 $t \neq 0$ 及 $t = 0$ 二情形討論之.

7-9 取 $N = N_1 \cup N_2$, 內 $P(N_1) = P(N_2) = 0$.

7-10 (1) $\overline{X}_n \xrightarrow{\text{a.s.}} 0.75$;

(2) 0.7267.

7-13 (1) 利用 Chebyshev 不等式.

(2) $a_n = b_n = n$.

7-14 0.192.

7-15 令 $\lambda_n = np_n$ 於 $B(n, p_n)$ 之特徵函數內, 再利用引理 7.5.

附錄一　排列與組合

所謂**組合分析**(combinatorial analysis) 乃數學之一支, 其研究之界限並不十分清晰, 但討論之主題乃爲在某些特定規則下, 考慮某些事物排列組合的方法數目及其相關的機率問題.

本書中所需之組合分析乃爲最基本之部分, 即可重複及不重複之排列與組合之問題, 其應用對象多半爲有限之樣本空間.

定理 A.1

> 設某一工作 T 之完成係由二件互不關聯的子工作 T_1 及 T_2 而組成, 若 T_1、T_2 分別有 n_1、n_2 種方法以執行, 則完成 T 的方法數爲 $n_1 \times n_2$.

證 對於 T_1, 若我們執行其中之一種方法, 則由於 T_2 有 n_2 種方法, 故完成 T 有 n_2 種方法; 但 T_1 有 n_1 種方法以執行之, 是以總共有 $n_1 \times n_2$ 種方法以執行工作 T. □

系 A.2

> 設某一工作 T 之完成係由 k 件互不關聯的子工作 $T_1, T_2, \cdots, T_k$ 而組成, 若 T_j 有 n_j 種方法以執行, 內 $j \in \{1, 2, \cdots, k\}$, 則完成 T 的方法數爲 $\prod_{j=1}^{k} n_j$.

證 利用數學規納法及定理 A.1, 讀者自證之 . □

在討論機率問題時, 我們常遭遇到某一含有 n 個元素之集合 (母群體), 而欲自其中取出 k 個元素, 這種從有限 (或無限) 母群體抽取元素的過程稱爲**抽樣** (sampling); 如果抽樣時, 每一元素被抽取的機會相等, 而且事前並不知悉, 則稱此種抽樣爲**隨機抽樣**(random sampling), 此時, 被抽得之 k 個元素稱爲一大小爲 k 之**隨機樣本**(random sample of size k). 抽樣過程有二個主要方法, 其一爲每抽取一個, 記錄後放回, 再抽, 直至 k 個元素 (其中可能有相同者), 此即所謂“**放回之抽樣**”(sampling with replacement); 另一爲每次抽取一個, 不再放回, 繼續抽至 k 個爲止, 此即所謂“**不放回之抽樣**” (sampling without replacement).

不論放回與否, 我們若將所得之樣本, 依其抽出之次序加以排列, 則稱此種樣本爲**有序** (ordered), 一個有序樣本又稱爲一**排列**(arrangement), 如果抽得之樣本不考慮其先後之次序, 則稱爲**無序**(unordered), 一個無序樣本又稱爲一**組合**(combination), 因此, 就樣本型態而言, 有以下四種情形 :

(1) 放回之有序樣本, (常稱其爲重複排列);

(2) 放回之無序樣本, (常稱其爲重複組合) ;

(3) 不放回之有序樣本, (常稱其爲排列樣本) ;

(4) 不放回之無序樣本 , (常稱其爲組合樣本).

定理 A.3

設有 n 個球, 分別標以 1 至 n 號 (或 n 件相異物品), 自其中取出一大小爲 k 之樣本, 則

(1) 放回之有序樣本共有 n^k 種 ;

(2) 放回之無序樣本共有 $H_{n,k} = \binom{n+k-1}{k}$ 種 ;

(3) 不放回之有序樣本共有 $P_{n,k} = n(n-1)\cdots(n-k+1)$ 種 ;

(4) 不放回之無序樣本共有 $C_{n,k} = \binom{n}{k} = \dfrac{n!}{k!(n-k)!}$ 種.

證 (1) 抽取每一球均有 n 種不同方法, 故抽取 k 個球共有 n^k 種方法.

(2) 設「**甲方法**」爲
除原有 1 至 n 號球外, 另加入 $(k-1)$ 個球分別標以

$$n+1, n+2, \cdots, n+k-1$$

今由 $(n+k-1)$ 個球, 以不放回方式抽取一大小爲 k 之無序樣本, 若所得之球爲 1 至 n 號, 仍以原號稱之; 若抽得 $n+1$ 號, 則以該樣本中最小數再稱之,(此最小數必介於 1 與 n 之間); 若抽得 $n+2$ 號, 則以該樣本中次小數再稱之; (註 : 若 $n+1$ 號己被抽取, 先化爲最小數, 則 $n+2$ 號亦應化爲該最小數). 依此類推, 若抽得 $n+k-1$ 號, 則該樣本中其他各數中最大者再稱之, (亦介於 1 至 n).
令集合

$$A = \{(a_1, \cdots, a_k) \mid (a_1, \cdots, a_k) \text{ 爲甲方法之無序樣本}\};$$

$$B = \{(a_1, \cdots, a_k) \mid (a_1, \cdots, a_k) \text{ 爲放回之無序樣本}\};$$

顯然 $A \subset B$, 且 $B \subset A$, 是以

$$H_{n,k} = \#B = \#A = \binom{n+k-1}{k}.$$

(3) 抽取第一球有 n 種情形, 第二球爲 $(n-1)$ 種情形, $\cdots$, 第 k 球有 $(n-k+1)$ 種情形, 由定理一之系知共有 $n \cdot (n-1)\cdots(n-k+1)$ 種不同樣本.

(4) 由於不放回之緣故每一樣本內之元素均不相同, 排列之, 共有 $k!$ 個不同情況, 亦即 $P_{n,k} = C_{n,k} \times k!$; 移項之, 得

$$C_{n,k} = \frac{P_{n,k}}{k!} = \frac{n\cdot(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{(n-k)!k!}. \qquad \square$$

註：關於二項係數 $\binom{n}{k}$ 之符號, 常見的有 $C(n,k)$, $C_{n,k}$, C_n^k, C_n^k, ${}_kC_n$ 等, 但近年來, 絕大部分的數學家均採用 $\binom{n}{k}$ (讀爲 n choose (取) k). 此外, 二項係數亦可略加推廣如下：設 $a\in\mathbb{R}, k\in\mathbb{Z}$ 我們規定

$$\binom{a}{k} = \begin{cases} \dfrac{a\cdot(a-1)\cdots(a-k+1)}{1\cdot 2\cdots k}, & \text{若 } k\in\mathbb{N}, \\ 1, & \text{若 } k=0, \\ 0, & \text{若 } k\in-\mathbb{N}. \end{cases}$$

例如：$\binom{5}{2}=10, \binom{2}{5}=0, \binom{3}{0}=1, \binom{\pi}{0}=1,$

$$\binom{-e}{2} = \frac{(-e)(-e-1)}{2} = e\cdot\frac{(e+1)}{2}, \binom{\pi}{-1} = 0.$$

定理 A.4 (二項係數之性質)

(1) 若 $n\in\mathbb{Z}$ 且 $n\ge k\ge 0$, 則 $\binom{n}{k} = \binom{n}{n-k}$;

(2) 若 n 爲非負整數, 則 $2^n = \binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n}$;

(3) 若 k 爲非負整數, 則 $\binom{-1}{k} = (-1)^k, \binom{-2}{k} = (-1)^k(1+k)$;

(4) 若 $k\ge n$ 且二數皆爲非負整數, 則 $\binom{n}{k}=0$;

(5) $\forall a\in\mathbb{R}, \forall k\in\mathbb{Z}$, $\binom{a}{k-1}+\binom{a}{k} = \binom{a+1}{k}$.

(6) $\forall m,n\in\mathbb{Z}_+, \forall r\in\{0,1,\cdots,m+n\}$, $\binom{m+n}{r} = \sum_{j=0}^{r}\binom{m}{j}\binom{n}{r-j}$.

證 (1) 易明.

(2) 對於實數 x 與 y , 我們已知

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}.$$

令 $x=y=1$, 則得

$$2^n=(1+1)^n=\sum_{k=0}^{n}\binom{n}{k}=\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n}.$$

(3) 若 $k=0$ 二等式顯然成立.

若 $k\in\mathbb{N}$,

$$\begin{aligned}\binom{-1}{k}&=\frac{(-1)\cdot(-2)\cdots(-1-k+1)}{k!}=(-1)^k\\ \binom{-2}{k}&=\frac{(-2)\cdot(-3)\cdots(-2-k+1)}{k!}=(-1)^k\\ &=(-1)^k\frac{(k+1)!}{k!}=(-1)^k(1+k).\end{aligned}$$

(4) 易明.

(5) 若 $k\le 0$ 等式顯然成立, 故只需證 $k\in\mathbb{N}$ 之情形.

$$\begin{aligned}\binom{a}{k-1}+\binom{a}{k}&=\frac{a\cdot(a-1)\cdots(a-k+2)}{(k-1)!}+\frac{a\cdot(a-1)\cdots(a-k+1)}{k!}\\ &=\frac{a\cdot(a-1)\cdots(a-k+2)}{(k-1)!}\Big(1+\frac{a-k+1}{k}\Big)\\ &=\frac{(a+1)\cdot a\cdot(a-1)\cdots(a-k+2)}{k!}\\ &=\binom{a+1}{k}.\end{aligned}$$

(6) 由於

$$\begin{cases}(x+1)^m=\displaystyle\sum_{k=0}^{m}\binom{m}{k}x^k\\ (x+1)^n=\displaystyle\sum_{i=0}^{n}\binom{n}{i}x^i\end{cases}$$

$$\begin{aligned}\Rightarrow\ (x+1)^{m+n}&=\Big(\sum_{k=0}^{m}\binom{m}{k}x^k\Big)\Big(\sum_{i=0}^{n}\binom{n}{i}x^i\Big)\\ &=\Big(\sum_{k=0}^{+\infty}\binom{m}{k}x^k\Big)\Big(\sum_{i=0}^{+\infty}\binom{n}{i}x^i\Big),^{[\star]}\\ \Rightarrow\ \sum_{r=0}^{m+n}\binom{m+n}{r}x^r&=\sum_{r=0}^{+\infty}c_rx^r,\Big[\text{內 } c_r=\sum_{j=0}^{r}\binom{m}{j}\binom{n}{r-j}\Big]\\ \Rightarrow\ \sum_{r=0}^{m+n}\binom{m+n}{r}x^r&=\sum_{r=0}^{m+n}c_rx^r,(r>m+n\ \Rightarrow\ c_r=0)\\ \Rightarrow\ \binom{m+n}{r}&=\sum_{j=0}^{r}\binom{m}{j}\binom{n}{r-j},\ (\text{比較 } x^r \text{ 係數})\end{aligned}$$

[⋆] $k>m\ \Rightarrow\ \binom{m}{k}=0.$ □

定理 A.5

設有 m 個彼此相異之球, 欲放在 r 個盒子內.

(1) 若每盒不限數量, 則有 r^m 種不同放法;

(2) 若規定第 j 盒放 n_j 個, 內 $j=1,\cdots,r,\ n_1+\cdots+n_r=m$, 則有 $\binom{m}{n_1,\cdots,n_r}=\dfrac{m!}{n_1!n_2!\cdots n_r!}$ 種不同放法.

證 (1) 由於

第一球有 r 種不同放法,

第二球有 r 種不同放法,

$\vdots$

第 m 球有 r 種不同放法,

由定理 A.1 知, 共有 $N=r\cdots r=r^m$ 種不同放法.

(2) 由於

第一個盒子有 $\binom{m}{n_1}$ 種不同放法,

第二個盒子有 $\binom{m-n_1}{n_2}$ 種不同放法,

$\vdots$

第 r 個盒子有 $\binom{m-n_1-\cdots-n_{r-1}}{n_r}$ 種不同放法,

故有

$$
\begin{aligned}
N &= \binom{m}{n_1}\binom{m-n_1}{n_2}\cdots\binom{m-n_1-\cdots-n_{r-1}}{n_r} \\
&= \frac{m!}{n_1!n_2!\cdots n_r!}=\binom{m}{n_1,\cdots,n_r}
\end{aligned}
$$

種不同放法. □

定理 A.6

設有 m 個相同之球, 欲放置在 r 個盒子內.

(1) 若每盒不限數量, 則有 $\binom{m+r-1}{m}$ 種不同放法;

(2) 若每盒不限, 但不空, 則有 $\binom{m-1}{r-1}$ 種不同放法.

證 (1) 設 r 個盒子內可由 $r+1$ 支棍子 (含兩端) 所圍成, 將兩端棍子固定則共有 m 個 "◯"(球) 及 $r-1$ 支 "|"(棍子).

例如：$m=5, r=4$, 則以下是一些排法:

| ◯◯ | ◯◯ | ◯ | |,

(第一盒, 第二盒各二個, 第三盒一個, 第四盒不放),

| ◯◯ | ◯ ◯ ◯ | | |,

(第一盒放二個, 第二盒放三個, 第三盒及第四盒不放).

棍子與球共有 $m+r-1$ 之任何一種橫排列乃為所需之一種放法, 顯然有 $\binom{m+r-1}{m}$ 種不同放法.

(2) 先將球排為一橫列, m 個球有 $m-1$ 個間隔, 若欲分為 r 個部分且每一部分皆不為空, 則僅需將 $r-1$ 支棍子分別放入 $m-1$ 個間隔, 每一間隔至多放一支棍子, 故有 $\binom{m-1}{r-1}$ 種不同放法. □

註：本定理之 (1) 亦可利用定理 A.3 證明之. 雖不見更容易, 但至少可以了解二者之關係.

先將 r 個盒子編號為 1 至 r 號. 若將 r 個號碼以回方式自其中取出一大小為 m 之無序樣本. 且當樣本中第 j 號每出現一次, 則將一球放入第 j 號盒子. 例如, $m=5, r=4$, 一樣本為 $(1,1,3,4,4)$ 時, 則一號盒子放二球, 三號盒子放一球, 四號盒子放二球, 二號子無球, 即為

| ◯◯ | | ◯ | ◯◯ | .

這種放法的總數與 m 球放進 r 個盒子的方法數, 顯然相等; 前者, 由定理 A.3 知, 於 r 個號碼以放回方式自其中取出一大小為 m 之無序樣本, 則共有 $\binom{r+m-1}{m}$ 種方法.

定理 A.7

設有 m 個球大小相同, 其中 n_1 個球為第一色, n_2 個球為第二 色, $\cdots$, n_r 個球為第 r 色, $(n_1+\cdots+n_r=m)$, 則此 m 個球共有 $\binom{m}{n_1,\cdots,n_r}$ 種不同排列方法.

證 設排列方法數為 N 對於每一種固定排法, 若第一色之球視為相異 (例如加以編號), 則有 $n_1!$ 種不同排列法, $\cdots$, 若第 r 色之球視為相異, 則有 $n_r!$ 種不同排列法, 但 m 個球均相異之排列法有 $m!$ 種, 是以

$$m! = N\cdot n_1!\cdots n_r!,$$

故

$$N=\frac{m!}{n_1!\cdots n_r!}=\binom{m}{n_1,\cdots,n_r}.$$ □

附錄二

定理 B.1

設 $\mathcal{P}(\mathbb{R})$ 表 $\mathbb{R}$ 之冪集合, $\mathcal{B}$ 表一維 Borel field, $\mathcal{M}$ 表一維 Lebesgue 可測集合所成之集合, 則

(1) $\mathcal{M}$ 爲 $\mathcal{P}(\mathbb{R})$ 之眞子集；

(2) $\mathcal{B}$ 爲 $\mathcal{M}$ 之眞子集 .

證 (1) 換言之, 在 $\mathbb{R}$ 內存在某集合, 不爲 Lebesgue 可測; 很多實變函數論之書籍均有此一定理及詳細之證明, 如 Royden [15] 第三章或 Wheeden & Zygmund [17] 3.38 定理. 由於證明冗長, 我們只有簡介其方法, 希望深入了解之讀者, 請自行參閱上述資料.

1° 先規定：如果二實數 $x-y \in \mathbb{Q}$, 則稱 x 與 y 爲對等. 由此一對等關係可建立一商集合 $\mathbb{R}/\simeq$, 其元素爲對等集 $E_x = \{x+r \mid r \in \mathbb{Q}\}$.

2° 每一對等集中各取一元素, 而構成新集合 E.

3° 若E爲 Lebesgue 可測集合, 則必含有有理數, 此與 2° 中 E 之構成相矛盾. 故知 E 不爲 Lebesgue 可測.

(2) 設 K 爲 Cantor 三進點集, 集合 D_k, D 及函數 F 見第二章第六節, μ 爲 Lebesgue 測度. 次令

$$g\colon [0,\,1] \to \mathbb{R} : g(x) = F(x) + x.$$

則 g 爲遞增連續且 $g[0,\,1] = [0,\,2]$, 因

$$\mu(g(D)) = \sum_{k=1}^{+\infty} \mu(g(D_k)) = \sum_{k=1}^{+\infty} \mu(D_k) = 1$$

是以 $\mu(g(K)) = 2 - \mu(g(D)) = 1 > 0$. 取 A 爲 $g(K)$ 之 Lebesgue 不可測子集, 則 $B = g^{-1}(A) \subset K$, 必爲零測度, 故必 Lebesgue 可測, 但若 $B \in \mathcal{B}$, 則因 g 爲遞增連續, 是以必有 $A = g(B) \in \mathcal{B} \subset \mathcal{M}$, 顯然不可能. 即 $B \in \mathcal{M} \setminus \mathcal{B}$. □

附錄三

定理 C.1

函數 $X:\Omega\to\Omega'$, 集合 $\mathcal{C}'\subset\mathcal{P}(\Omega')$. 則 $\sigma(X^{-1}(\mathcal{C}'))=X^{-1}(\sigma(\mathcal{C}'))$.

證

1° 先證 : $\sigma(X^{-1}(\mathcal{C}'))\subset X^{-1}(\sigma(\mathcal{C}'))$.
由於 $\sigma(\mathcal{C}')$ 爲包含 $\mathcal{C}'$ 之最小 σ 域, 是以 $X^{-1}(\sigma(\mathcal{C}'))$ 爲 Ω 上一 σ 域, 且因其包含 $X^{-1}(\mathcal{C}')$, 故亦包含 $\sigma(X^{-1}(\mathcal{C}'))$.

2° 次證 : $X^{-1}(\sigma(\mathcal{C}'))\subset\sigma(X^{-1}(\mathcal{C}'))$. 令

$$\mathcal{F}'=\{A'\subset\Omega' \mid X^{-1}(A')\in\sigma(X^{-1}(\mathcal{C}'))\},$$

則

- $\mathcal{F}'$ 爲一 σ 域;
- $\sigma(\mathcal{C}')\subset\mathcal{F}'$. (因若 $A'\in\mathcal{C}'$, 則 $X^{-1}(A')\in X^{-1}(\mathcal{C}')\subset X^{-1}(\sigma(\mathcal{C}'))$.)

故由 $\mathcal{F}'$ 之定義知,

$$\begin{aligned}X^{-1}(\sigma(\mathcal{C}'))&\subset X^{-1}(\mathcal{F}')\\&\subset\sigma(X^{-1}(\mathcal{C}')).\end{aligned}$$

□

附錄四

定理 D.1 （可測性定理）

設 $(\Omega, \mathcal{F}, P)$ 爲一機率空間, 函數 $X\colon \Omega \to \mathbb{R}$, 則以下二敍述爲對等：

(I) X 爲一非負隨機變數;

(II) 存在一非負簡單隨機變數序列 $\{X_n\}_n$ 使得 $X_n \uparrow X$, 意即:

$$\forall n \in \mathbb{N},\ X_n \leq X_{n+1} \text{ 且 } \forall \omega \in \Omega,\ \lim_{n\to+\infty} X_n(\omega) = X(\omega).$$

證 (I) $\Rightarrow$ (II).

令 $\forall n \in \mathbb{N},\ g_n\colon [0,+\infty) \to [0,+\infty) : g_n(x) = \begin{cases} \dfrac{[2^n x]}{2^n}, & \text{若 } 0 \leq x < n, \\ n, & \text{若 } n \leq x. \end{cases}$

$X_n = g_n \circ X$, 則因 $g_n \leq g_{n+1},\ \forall n \in \mathbb{N}$, 且

$$\forall x \in [0,n],\ 0 \leq x - g_n(x) \leq \frac{1}{2^n},$$

知 $\{X_n\}_n$ 爲一非負簡單隨機變數序列, 且 $X_n \leq X_{n+1},\ \forall n$.

其次, $\forall \omega \in \Omega, \exists n_0 \in \mathbb{N}$ 使得 $X(\omega) < n_0$, 當 $n \geq n_0$,

$$\begin{aligned} g_n(X(\omega)) = \frac{[2^n X(\omega)]}{2^n} &\Rightarrow 0 \leq X(\omega) - g_n(X(\omega)) \leq \frac{1}{2^n} \\ &\Rightarrow 0 \leq X(\omega) - X_n(\omega) \leq \frac{1}{2^n} \\ &\Rightarrow \lim_{n\to+\infty} X_n(\omega) = X(\omega). \end{aligned}$$

(II) $\Rightarrow$ (I).

由於 $\{X_n\}_n \uparrow X$, 知 $X = \sup_{n\geq 1} X_n$, 是以 $\forall a \in \mathbb{R}$,

$$\{X \leq a\} = \{\sup_{n\geq 1} X_n \leq a\} = \{X_n \leq a,\ \forall n\} = \bigcap_{n\geq 1}\{X_n \leq a\} \in \mathcal{F}.$$

由 2.4 系知 X 爲一隨機變數. □

附錄五

定理 E.1

設 X 爲一非負隨機變數, 若非負簡單隨機變數序列 $\{X_n\}_n$ 及 $\{Y_n\}_n$ 均滿足 $X_n \uparrow X$ 及 $Y_n \uparrow X$; 則

$$\lim_{n\to+\infty} E[X_n] = \lim_{n\to+\infty} E[Y_n].$$

$\boxed{\text{證}}$ 1° 先證：若簡單非負隨機變數 $Y \le X$, 則 $EY \le \lim_{n\to+\infty} EX_n$. 爲此, 只需證明 $m = \min_{\omega\in\Omega} Y(\omega) > 0$ 之情形.

$\forall\epsilon \in (0, m)$, 令 $A_n = \{X_n > Y - \epsilon\}, B_n = \Omega \setminus A_n$, 則

$$\lim_{n\to+\infty} A_n = \bigcup_{n=1}^{+\infty} \{X_n > Y - \epsilon\} = \Omega, \lim_{n\to+\infty} B_n = \varnothing.$$

是以

$$\begin{aligned}
E[X_n] &\ge E[XI_{A_n}] \\
&\ge E[(Y-\epsilon)I_{A_n}] \\
&= E[YI_{A_n}] - \epsilon E[I_{A_n}] \\
&= E[Y] - E[YI_{B_n}] - \epsilon P(A_n),\ (\because I_{A_n} = 1 - I_{B_n}) \\
&\ge E[Y] - MP(B_n) - \epsilon, (M = \max Y(\Omega))
\end{aligned}$$

先令 $n \to +\infty$, 再令 $\epsilon \to 0$, 則可證得 $\lim_{n\to+\infty} E[X_n] \ge E[Y]$.

2° 由是

$$\begin{aligned}
Y_n \uparrow X \ &\Rightarrow\ Y_n \le X \\
&\Rightarrow\ E[Y_n] \le \lim_{n\to+\infty} E[X_n],\ \forall n,\ (\because 1°) \\
&\Rightarrow\ \lim_{n\to+\infty} E[Y_n] \le \lim_{n\to+\infty} E[X_n].
\end{aligned}$$

反之, 亦有 $\lim_{n\to+\infty} E[X_n] \le \lim_{n\to+\infty} E[Y_n]$. □

附錄六

定理 F.1

若隨機變數 X 具指數分配, 則其具失憶症, 即

$$P\{X > s+t \mid X \geq t\} = P\{X > s\}, \forall\, s,t \in (0,+\infty) \qquad (*)$$

反之, 若 X 具失憶症, 則 X 必具指數分配.

$\boxed{\text{證}}$ 只需證明：若 X 具失憶症, 則 X 具指數分配. 令

$$G(x) = 1 - F_X(x) = P\{X > x\},$$

則

$$\begin{aligned} X \text{ 具失憶症} &\Rightarrow P\{X > s+t\} = P\{X > s\} \cdot P\{X > t\}, \forall\, s,\ t > 0,\ [\because (*)] \\ &\Rightarrow G(s+t) = G(s)G(t), \forall\, s,t \geq 0 \\ &\Rightarrow \exists \lambda > 0 \text{ 使得 } \forall\, x \geq 0, G(x) = e^{-\lambda x}, \quad [\dagger] \\ &\Rightarrow \forall\, x \geq 0, F_X(x) = 1 - e^{-\lambda x} \\ &\Rightarrow X \sim G(1, \tfrac{1}{\lambda}). \end{aligned}$$

[†] 設 $G\colon [0,+\infty) \to [0,1]$ 爲右連續且 $G(s+t) = G(s)G(t)$. 則

- 令 $G(1) = a$, 則因 $a = G(1) = G(\frac{1}{n} + \cdots + \frac{1}{n}) = G^n(\frac{1}{n})$, 知 $G(\frac{1}{n}) = a^{1/n}$.
- $G(\frac{m}{n}) = G(\frac{1}{n} + \cdots + \frac{1}{n}) = G^m(\frac{1}{n}) = a^{m/n}$.
- 若 $x \geq 0$, 取 $\mathbb{Q}$ 內一序列 $\{r_k\}_k \downarrow x$, 利用右連續性得

$$G(x) = \lim_{k\to+\infty} G(r_k) = \lim_{k\to+\infty} a^{r_k} = a^x,$$

令 $\lambda = -\ln a$, 則因 $0 < a < 1$ (‡), 是以 $\lambda > 0$ 且

$$G(x) = a^x = (e^{-\lambda})^x = e^{-\lambda x}.$$

(‡) 若 $a = 1$, 則 $G(x) = 1$, 是以 $F_X(x) = 0$ 非分配函數也. 若 $a = 0$, 則 $\forall\, x \geq 0,\ G(x) = 0, F_X(x) = 1$, 使 $(*)$ 式

$$P\{X > s+t \mid X > t\} = P\{X > s\}$$

左端之分母爲 0, 無意義. $\square$

附錄七　標準常態分配表

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

漢英名詞索引

十七畫以上

英漢名詞索引