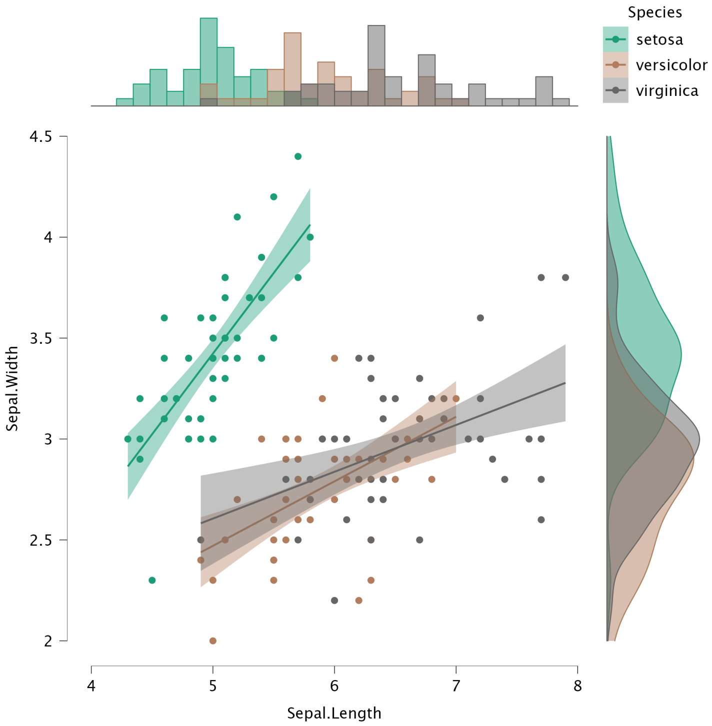


高等統計學

戴忠淵
chungyuandye@gmail.com

January 17, 2025

目錄

圖目錄

自序

我打研究室走過
那獨坐電腦前的容顏如苦瓜的糾結

靈感不來，長壽的煙霧不散
研究室如小小的寂寞的城
恰如商管的電梯向晚
鍵盤不響，彈菸的手指不歇
看動作片要把小小的窗扉緊掩

我達達的馬蹄是美麗的錯誤
我不是主人，是個過客

第 1 章　資料的統計量數

統計學是數學的一個分支，主要用來收集、整理、分析和解釋資料，已廣泛應用於自然科學、人文科學，以及近年來工商業和政府決策等領域。統計分析最主要可分為兩個部分：(Descriptive statistics) 和推論統計學 (Inferential statistics)，兩者雖是都利用資料進行分析，但所採用的分析方法和目的並不相同。描述統計學 (Descriptive statistics) 是以有組織的方式呈現資料的方法，其目的在於描述資料的特性並掌握資料的樣貌，讓人能夠輕易理解資料所傳達的訊息。推論統計學 (Inferential statistics) 則是將整理後的資料數據，利用機率模型將樣本資料推論至母體的方法，目的為估計母體參數或預測解釋自變項和依變項之間的關聯性。這兩種方法都可以被稱為應用統計學 (Applied statistics)。

在研究統計學時，我們必須先了解什麼是的『母體』(Population) 。所謂母體，乃指由具有特定共同特性的元素或個體所組成的一個群體，即研究人員所要研究觀察的對象的全體集合。母體可以是一個國家的人民、一個學校的學生，或是某工廠生產線所生產的商品。以市長選舉民意調查為例，其母體為具有高雄市市長選舉投票權的市民所組成的群體，而其組成的條件包括：(1) 我國國籍且年滿 20 歲、(2) 沒有受到監護宣告及 (3) 在該選舉區設籍繼續居住 4 個月以上並且在高雄市設籍繼續居住滿 4 個月以上。凡符合上述三個條件的市民即為合格選民，而符合上述三個條件的合格選民即稱為母體。

研究者蒐集母體資料的主要目的，即希望能對這些資料的特性進行分析，以獲得重要的資訊。例如高雄市長選舉民意調查的目的在於推估各市長候選人的支持率，支持率在此稱為『母體參數』(Population parameters) ，乃指研究問題中描述母體資料特性的統計量數，通常為一個未知數。因此，欲了解母體參數的特性，就必須利用樣本所計算出來的統計量，對未知的母體參數做估計或利用各種統計方法，進行推論。

然而，母體往往龐大無法全面調查，故在確保獲得的資料具有代表性的前提下，我們會從母體中抽取一定數量的樣本 (Sample) 來代表整個研究對象，此一過程稱為抽樣 (Sampling)。透過對樣本

資料進行敘述統計分析，我們不僅能了解母體的基本特性，還可進行相關的統計推論對母體參數進行有效的估計，並以此來解釋和預測母體的整體行為。因此，統計分析的相關流程可彙整如圖1.1。

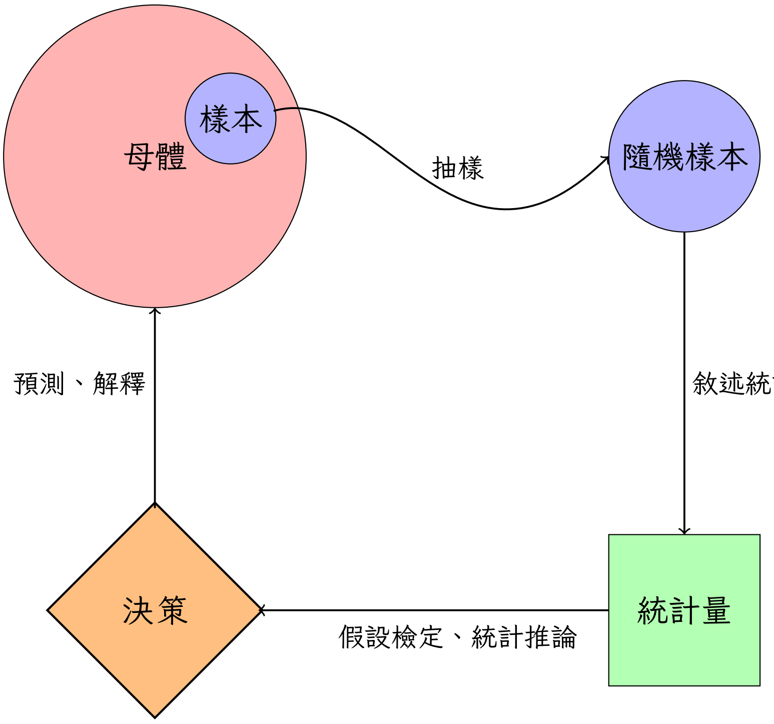


圖 1.1: 統計分析架構圖

欲了解母體的特性，就必須由樣本所計算出來的統計量級，來對未知的母體做估計或利用各種統計方法, 進行其推論。本章將介紹一些統計學上常用的量化數值來解釋資料的特徵，一般在統計學上常用以描述母體的量數有：

1. **母體參數**：是一量化的數值，用來描述母體的特性，一般來說是未知的，例如：μ, σ, η 等。
2. **統計量**：亦是一量化的數值，經過抽樣，而求出其特性質，例如：$\overline{X}, S, \widehat{p}$ 等。

1.1 集中趨勢量

集中趨勢量(Measures of Central Tendency)：資料向某一數值集中, 表示此種共同趨勢的量數，可做資料的中心代表值或兩個以上 (含) 群體間之比較或估計，推論母體的母體。以下介紹三種重要的集中趨勢量數：(1) 平均數、(2) 中位數及 (3) 眾數。

1.1.1 平均數

1. **算術平均數**：通常使用的平均數是指『算術平均數』

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

例 **1–1** 某統計諮詢顧問中心員工之薪資如下：

$$30, 30, 30, 84, 30, 30, 40, 30, 40, 45, 70, 38, 95$$

求此中心員工的平均薪資為多少？(單位：新台幣仟元)

解：

$$\overline{X} = \frac{\sum X_i}{n} = \frac{592}{13} = 45.538$$

2. **幾何平均數**：較適用於比例 (等比) 或變動率的數據

$$G = \sqrt[n]{X_1 X_2 \cdots X_n} = \sqrt[n]{\prod_{i}^{n} X_i}$$

其中，$X_i > 0$，否則無意義。

例 **1–2** 台泥公司最近四年之營業額為 3, 6, 12, 24(億元)，試求其每年之平均營業額

解：

$$G = \sqrt[4]{3 \times 6 \times 12 \times 24} = 8.49$$

3. **調和平均數**：較適用於比例，速率或物價之數據

$$H = \frac{1}{1/n \times \sum_{i=1}^{n} (1/X_i)} = \frac{n}{\sum_{i=1}^{n} (1/X_i)}$$

例 **1–3** 甲生返往兩地, 去時每小時平均 4 公里，回程每小時平均 5 公里，試求往返兩地之平均速率？

解：

$$H = \frac{1}{\frac{\frac{1}{4} + \frac{1}{5}}{2}} = \frac{40}{9}$$

4. **加權平均數**：在一組資料中按其重要性或所佔比例 (比率) 給予一個特殊的權數 (Weight) 後，所得之平均數，例如大專院校的學期成績。設 $w_1, w_2, \ldots, w_n$ 分別為 $X_1, X_2, \ldots, X_n$ 之權數則

$$\overline{X}_w = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

定理 1.1. 算術平均數，幾何平均數及調和平均數其間之關係如下：

1. 若 $X_i > 0, \forall i$，則 $\overline{X} \geq G \geq H$。

2. 當僅有任二正值時 *(*即 $X_1 > 0,\ X_2 > 0$*)*，則 $G = \sqrt{\overline{X}H}$。

證明*.*

1. 令 $0 \leq X_1 \leq X_2 \leq \cdots \leq X_n$，則

$$\begin{aligned}
\frac{X_1 + X_2 + \cdots + X_n}{n} &= \frac{X_n + (n-1)\left(X_1 + X_2 + \cdots + X_{n-1}\right)/(n-1)}{n} \\
&= \frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1} + \frac{X_n - \left(X_1 + X_2 + \cdots + X_{n-1}\right)/(n-1)}{n}
\end{aligned}$$

又由二項式定理得

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \geq a^n + na^{n-1}b$$

令 $a = \left(X_1 + X_2 + \cdots + X_{n-1}\right)/(n-1)$ 及 $b = \left[X_n - \left(X_1 + X_2 + \cdots + X_{n-1}\right)/(n-1)\right]/n$ 得

$$\begin{aligned}
\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)^n &\geq \left(\frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1}\right)^n \\
&\quad + n\left(\frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1}\right)^{n-1} \left(\frac{X_n - \frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1}}{n}\right) \\
&= X_n \times \left(\frac{X_1 + X_2 + \cdots + X_{n-1}}{n-1}\right)^{n-1} \\
&\geq X_n \times X_{n-1} \times \left(\frac{X_1 + X_2 + \cdots + X_{n-2}}{n-2}\right)^{n-2} \\
&\geq \cdots \\
&\geq X_n \times X_{n-1} \times \cdots \times X_1
\end{aligned}$$

故得 $\overline{X} \geq G$。另一方面，令 $Y_i = \frac{1}{X_i}$，則 $Y_n \leq Y_{n-1} \leq \cdots \leq Y_1$。由於 $\overline{X} \geq G$, 得

$$\begin{aligned}
\frac{Y_n + Y_{n-1} + \cdots + Y_1}{n} &\geq \sqrt[n]{Y_n Y_{n-1} \cdots Y_1} \\
&= \sqrt[n]{\frac{1}{X_1}\frac{1}{X_2}\cdots\frac{1}{X_n}} \\
&= \frac{1}{\sqrt[n]{X_1 X_2 \cdots X_n}}
\end{aligned}$$

所以，

$$\frac{1}{(1/X_1 + 1/X_2 + \cdots + 1/X_n)/n} \leq \sqrt[n]{X_1 X_2 \cdots X_n}$$

故得 $G \geq H$，得證

2.

$$\begin{aligned}
\sqrt{\overline{X}H} &= \sqrt{\frac{X_1 + X_2}{2} \times \frac{2}{1/X_1 + 1/X_2}} \\
&= \sqrt{\frac{X_1 + X_2}{2} \times \frac{2X_1X_2}{X_1 + X_2}} \\
&= \sqrt{X_1 X_2} = G
\end{aligned}$$

性質

1. $\sum_{i=1}^{n}(X_i - \overline{X}) = 0$
2. $\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \leq \sum_{i=1}^{n}(X_i - p)^2$，其中 p 表任何值。
3. $G_{\frac{X}{Y}} = G_X / G_Y$
4. 設有 k 組變量，其各組個數分別為 $n_1, n_2, \ldots, n_k$ 算術平均數分別為 $\overline{X}_1, \overline{X}_2, \overline{X}_3, \ldots, \overline{X}_k$，則全體平均分數為
$$\overline{X} = \frac{\sum_{i=1}^{k} n_i \overline{X}_i}{\sum_{i=1}^{k} n_i}$$

證明.

1.

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\overline{X}) &= \sum_{i=1}^{n}X_i-n\overline{X}\\ &= \sum_{i=1}^{n}X_i-n\frac{\sum_{i=1}^{n}X_i}{n}\\ &= 0\end{aligned}$$

2.

$$\frac{\partial\sum_{i=1}^{n}(X_i-p)^2}{\partial p}=\sum_{i=1}^{n}-2(X_i-p)$$

又

$$\frac{\partial\sum_{i=1}^{n}-2(X_i-p)}{\partial p}=\sum_{i=1}^{n}2=2n\geq 0$$

令 $\sum_{i=1}^{n}-2(X_i-p)=0$，所以當 $p=\overline{X}$ 時有極小值。故得

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\leq\sum_{i=1}^{n}(X_i-p)^2$$

3.

$$\begin{aligned}G_{\frac{X}{Y}} &= \sqrt[n]{\frac{X_1}{Y_1}\frac{X_2}{Y_2}\cdots\frac{X_n}{Y_n}}\\ &= \frac{\sqrt[n]{X_1X_2\ldots X_n}}{\sqrt[n]{Y_1Y_2\ldots Y_n}}\\ &= \frac{G_X}{G_Y}\end{aligned}$$

1.1.2 位數

1. 中位數：最中間的數，即其將資料群化分為兩部份，其符號為 **Me**。

$$\mathbf{Me}=\begin{cases}X_{\left(\frac{n+1}{2}\right)} & \text{，當 } n \text{ 為奇數時}\\ \dfrac{X_{\left(\frac{n}{2}\right)}+X_{\left(\frac{n}{2}+1\right)}}{2} & \text{，當 } n \text{ 為偶數時}\end{cases}$$

性質

(a) 中位數居於中間的一數，故不受極端值的影響。

(b) $\sum_{i=1}^{n}|X_i-\mathbf{Me}|\le\sum_{i=1}^{n}|X_i-p|$，其中 p 表任何值。

2. **四分位數**：將資料群 (處理過) 化分成四部分，即有三個分割點，則

$$Q_k=\begin{cases}X_{\left(\left[\frac{nk}{4}\right]+1\right)} & \text{，當 } kn/4 \text{ 不為整數時}\\ \dfrac{X_{\left(\frac{nk}{4}\right)}+X_{\left(\frac{nk}{4}+1\right)}}{2} & \text{，當 } kn/4 \text{ 為整數時}\end{cases}$$

其中 $k=1,2,3,$ $[\cdot]$ 為高斯符號

(a) 第一個分割點稱為第一四分位數 (下四分位數)，其符號為 Q_1。

(b) 第二個分割點稱為第二四分位數 (中位數)，其符號為 Q_2。

(c) 第三個分割點稱為第三四分位數 (上四分位數)，其符號為 Q_3。

3. **十分位數**：與前述兩種位數的概念是一樣的，名稱亦類似，故符號為 D_1, D_2, D_3, ..., D_9

(iv) 百分位數 Percentiles：同理，其符號為 P_1, P_2, ..., P_{99}; 即 $P_r,\ r=1,\ 2,\ 3,\ \ldots,99$

各位數之間的關係如下：

$$P_{50}=D_5=Q_2=\mathbf{Me},$$

$$P_{25}=Q_1,$$

$$P_{75}=Q_3,$$

$$P_{10}=D_1,$$

$$P_{20}=D_2,$$

$$\vdots$$

1.1.3 眾數

眾數：一資料群中出現次數最多者所對應的數值。

1. 眾數之符號常用 M_0。

2. 無任何一個數值出現超過一次以上，則眾數不存在。

3. 眾數可能不只一個值。

4. 如相鄰二數皆為眾數，則取平均值。

此外，平均數、中位數及眾數三者常有以下的關係：

$$\text{平均數} - \text{眾數} = 3(\text{平均數} - \text{中位數})$$

利用此一關係，可了解資料各集中趨勢量數的關係及求得眾數。

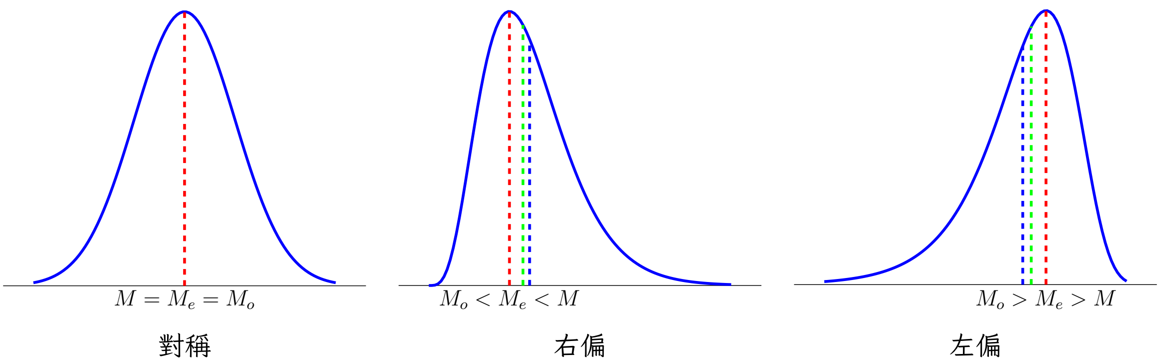


圖 1.2: 平均數、中位數及眾數關係圖

1.2 差異量數

討論資料的離散程度。雖然兩組資料具有相同的 $\overline{X}$ 或 Me 等時，並不代表這兩組資料群具有相同的分布情況。以下介紹三類的差異量數：

1. **非離中差異量數**：不以中心點為基準，而以兩極端點為基準，如全距。

$$R = X_{(n)} - X_{(1)}$$

2. **離中差異量數**：以中心點為基準

(a) 四分位差：以中位數 Me 為中心，

$$\frac{Q_3 - Q_1}{2}$$

(b) 平均差

i. 以中位數 Me 為中心的稱為**離中平均差**，

$$\text{M.D.} = \frac{\sum_{i=1}^{n} |X_i - Me|}{n}$$

ii. 以平均數 $\overline{X}$ 為中心的稱為**離均平均差**

$$\text{M.D.} = \frac{\sum_{i=1}^{n} \left|X_i - \overline{X}\right|}{n}$$

(c) 樣本變異數：以平均數 $\overline{X}$ 為中心

$$S^2 = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n-1}$$

3. **變異係數**：當我們用標準差來衡量資料的離散程度時，有時候很難從其表面的數據看出真實的情況。因為當資料群所含的數值愈大，它的 S^2 和 S 通常就愈大；相反地，數值愈小，即使離散程度大，往往它的 S^2 和 S 也不會太大，故比較兩組資料群的離散程度時，利用變異係數來衡量資料的離散情況。

$$\text{C.V.} = \frac{S}{\overline{X}} \times 100\%$$

例 **1–4**　己知某單位員工的平均身高為 160 公分, 身高的標準差為 8 公分; 又知他們的平均體重為 50 公斤, 體重的標準差為 4 公斤，試比較該單位的員工的身高與體重何者差異較大。

解：

$$\text{C.V.}_{\text{身高}} = \frac{8}{160} \times 100\% = 5\%,\ \text{C.V.}_{\text{體重}} = \frac{4}{50} \times 100\% = 8\%$$

1.3 平均數與標準差之應用

1. 柴比雪夫定理(Chebyshev's Theorem)：不論資料分布的型態如何，若母體平均數及母體變異數皆存在，則

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

2. 經驗法則：當資料群呈鍾型分配時，則有下列關係

   (a) 大約有 68% 的資料落入區間 $(\mu-\sigma\ ,\ \mu+\sigma)$

   (b) 大約有 95% 的資料落入區間 $(\mu-2\sigma\ ,\ \mu+2\sigma)$

   (c) 大約有 99.74% 的資料落入區間 $(\mu-3\sigma\ ,\ \mu+3\sigma)$

例 **1–5** 設有 100 人參加數學競試，其平均成績為 72 分，變異數為 36 分，則成績在 60 分至 84 分之間的人數至少有多少人？

解：

因為 60 分至 84 分之間，即 $72\pm2\times6$，所以得知 $k=2$。由柴比雪夫定理, 得至少有 $100\times75\%=75$ 人

3. 標準分數：將原始資料的數值轉化為標準值的方法

$$Z_i=\frac{X_i-\overline{X}}{S}$$

第 2 章 機率概論

統計推論中，應用機率來測度不確定性的大小，應用機率來導出樣本的抽樣分配，進而推論母體的性狀或作出有利的決策，故機率扮演極重要的角色。機率的理論是統計學的主要基礎。統計學之能成為一門獨立的科學，是有賴機率理論的發展。而集合的概念是研討機率理論與應用的基礎。

2.1 常用符號

1. 屬於（belong to）：$\in$
2. 子集（subset）：$\subseteq$
3. 空集合（null set）：ϕ
4. 聯集（union set）：$\cup$
5. 交集（intersection set）：$\cap$
6. 餘集（complement set）：A^c 或 A'

2.2 基本運算與性質

1. A, B 的聯集：$A \cup B = \{x \mid x \in A \text{ 或 } x \in B\}$
2. A, B 的交集：$A \cap B = \{x \mid x \in A \text{ 及 } x \in B\}$
3. A 的餘集：$A^c = \{x \mid x \in S \text{ 及 } x \notin A\}$
4. DeMorgan's laws：

(a) $(A \cap B)^c = A^c \cup B^c$

(b) $(A \cup B)^c = A^c \cap B^c$

2.3 名詞解釋

1. **集合**：由同類事物，符號或數字所成的聚集。常用 A, B, C, $\ldots$ 表示。

2. **元素**：構成集合的同類事物，符號或數字。常用 a, b, c, $\ldots$ 表示。

3. **隨機實驗**：是一種實驗，其結果為不能確定地加以預測而純屬機遇者。例：擲骰子試驗。

4. **樣本空間**：隨機實驗所有可能的結果所組成的集合。常用 S 表示。例：骰子的點數，$S = \{1, 2, 3, 4, 5, 6\}$。

   (a) 若包含有限個或可數但無限個樣本點，稱為離散樣本空間

   (b) 若包含一區間中的全部實數，稱為連續樣本空間

5. **事件**：樣本空間的部份集合。假設樣本空間中有幾個樣本點，即 $\#(S) = n$，則有 2^n 個可能的事件。

6. **空事件**：表示必定不發生的事件，符號固定用 ϕ。

   (a) 聯事件：$A \cup B$

   (b) 交事件：$A \cap B$

   (c) 餘事件：A^c 或 A'

   (d) 互斥事件：如果事件 A, B 為互斥，則 $A \cap B = \phi$

   (e) 完全事件：如果 $A_1 \cup A_2 \cup \cdots \cup A_n = S$

   (f) 母體與樣本皆為集合, 常用 A, B, C, $\ldots$ 表示之。

例 **2–1** 設母體 $A = a, b$，則有 $2^2 = 4$ 種可能的樣本選擇，即

$$A = \phi,\ \{a\},\ \{b\},\ \{a, b\}$$

例 **2–2** 設母體 $A=\{\clubsuit,\heartsuit,\spadesuit\}$，則有 $2^3=8$ 種可能的樣本選擇，即

$$A=\phi,\ \{\clubsuit\},\ \{\heartsuit\},\ \{\spadesuit\},\{\clubsuit,\heartsuit\},\ \{\clubsuit,\spadesuit\},\ \{\heartsuit,\spadesuit\},\ \{\clubsuit,\heartsuit,\spadesuit\}$$

2.4 機率

2.4.1 機率公理

1. 對任一事件 A, $\mathrm{P}(A)\geq 0$
2. $\mathrm{P}(S)=1$
3. 若 $A_i\cap A_j=\phi, \forall i\neq j$，則

$$\mathrm{P}(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1}^{n}\mathrm{P}(A_1)+\mathrm{P}(A_2)+\cdots+\mathrm{P}(A_n)$$

設每個樣本點出現的機率相等，則

$$\mathrm{P}(A)=\frac{\#(A)}{\#(S)},$$

其中 $\#(S)$ 表樣本空間的元素個數，$\#(A)$ 表事件 A 的元素個數。

例 **2–3** 投擲兩枚公正的骰子，求出現點數和為 7 的機率？

解：

$\#(S)=6^2=36$, 事件 $\mathrm{A}=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$. 故 $\mathrm{P}(A)=6/36=1/6$

例 **2–4** 設有 A, B, C 三台機器生產同一產品，若 A 的故障機會為 B 的兩倍，B 的故障機率為 C 的三倍，試求各台機器的故障機率為若干？

解：

令 $\mathrm{P}(C)=p$，則 $\mathrm{P}(B)=3p$. 又 $\mathrm{P}(A)=2\mathrm{P}(B)=6p$ 且 $p+3p+6p=1$, 得 $p=0.1$。故 $\mathrm{P}(A)=0.6$, $\mathrm{P}(B)=0.3$, $\mathrm{P}(C)=0.1$

例 **2–5** 6 個球隨機分到 3 個盒子裡，問每一個盒子都有球的機率為何？

解：

$P(\text{兩個盒子空})=\frac{\binom{3}{2}}{3^6}$

$P(\text{一個盒子空})=\frac{\binom{3}{1}(2^6-2)}{3^6}$

$P(\text{沒有盒子空})=1-P(\text{兩個盒子空})-P(\text{一個盒子空})=1-\frac{3}{729}-\frac{186}{729}=\frac{540}{729}=\frac{20}{27}$

性質

1. $\mathrm{P}(\phi)=0$
2. 任一事件 A, $0\le \mathrm{P}(A)\le 1$
3. $\mathrm{P}(A)=1-\mathrm{P}(A')$
4. 機率加法定理：設 A, B 為二事件，則

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

若 A, B 為互斥事件，則

$$\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)$$

5. 若 A, B, C 為任意三事件，則

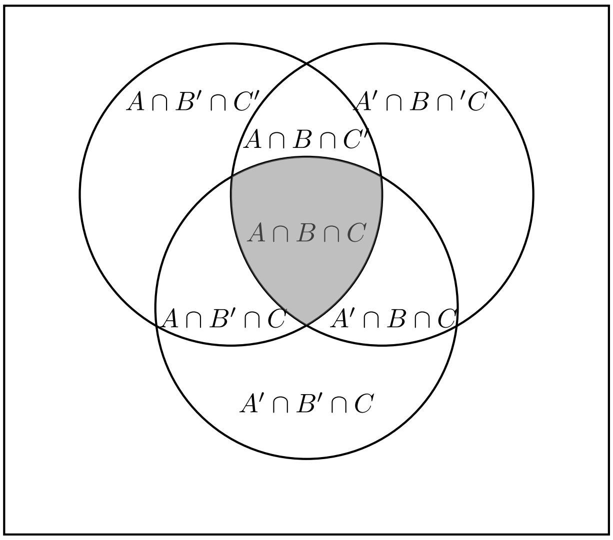
$$\mathrm{P}(A\cup B\cup C)=\mathrm{P}(A)+\mathrm{P}(B)+\mathrm{P}(C)-\mathrm{P}(A\cap B)-\mathrm{P}(A\cap C)-\mathrm{P}(B\cap C)+\mathrm{P}(A\cap B\cap C)$$

6. 若 A, B, C 為三互斥事件，則

$$\mathrm{P}(A\cup B\cup C)=\mathrm{P}(A)+\mathrm{P}(B)+\mathrm{P}(C)$$

7. 推廣至 n 個事件，則

$$\begin{aligned}
\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n) &= \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{i<j}^{n} \mathrm{P}(A_i \cap A_j) + \sum_{I<j<k}^{n} \mathrm{P}(A_I \cap A_j \cap A_k) \\
&\quad + \cdots + (-1)^{n+1} \times \mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n)
\end{aligned}$$



例 **2–6** 已知 $\mathrm{P}(A) = 0.4$，$\mathrm{P}(B) = 0.7$ 與 $\mathrm{P}(A \cap B) = 0.3$, 求

1. $\mathrm{P}(A')$
2. $\mathrm{P}(A \cup B)$
3. $\mathrm{P}(A \cap B')$
4. $\mathrm{P}(A' \cap B')$
5. $\mathrm{P}(A' \cup B')$

解：

1. $\mathrm{P}(A') = 1 - \mathrm{P}(A) = 1 - 0.4 = 0.6$
2. $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) = 0.4 + 0.7 - 0.3 = 0.8$
3. $\mathrm{P}(A \cap B') = \mathrm{P}(A) - \mathrm{P}(A \cap B) = 0.4 - 0.3 = 0.1$
4. $\mathrm{P}(A' \cap B') = \mathrm{P}((A \cup B)') = 1 - \mathrm{P}(A \cup B) = 1 - 0.8 = 0.2$
5. $\mathrm{P}(A' \cup B') = \mathrm{P}((A \cap B)') = 1 - \mathrm{P}(A \cap B) = 1 - 0.3 = 0.7$

2.4.2 條件機率

設 $\mathrm{P}(B)>0$，則在事件 B 已發生的條件下發生事件 A 的機率稱為事件 A 的條件機率，其定義為

$$\mathrm{P}(A\mid B)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}$$

在探討條件機率, 是將已發生的事件，視為一較小的新樣本空間，事件 A 的元素個數有多少在事件 B 發生。平常我們看的機率，亦可視為一種條件機率，即 $\mathrm{P}(A)=\mathrm{P}(A\mid S)$。

例 **2–7** 一袋中有 10 個黑球，6 個黃球，10 個白球，今自此袋中隨機取出一球，已知它不是白球，求取出的球為黃色球的機率？

解：

設 W' 表取出非白球的事件，Y 表取出黃球的事件，則

$$\mathrm{P}(Y\mid W')=\frac{\mathrm{P}(Y\cap W')}{\mathrm{P}(W')}=\frac{6/26}{16/26}=\frac{3}{8}$$

性質

1. $0\le \mathrm{P}(A\mid B)\le 1$
2. $\mathrm{P}(B\mid B)=1$
3. 若 $A_1,\, A_2,\, \cdots\, A_k$ 為互斥事件，即 $\mathrm{P}(A_i\cap A_j)=\phi,\, i\neq j$，且 $\mathrm{P}(B)>0$，則
$$\mathrm{P}(A_1\cup A_2\cup\cdots\cup A_k\mid B)=\sum_{i=1}^{k}\mathrm{P}(A_i\mid B)$$
4. 若 $\mathrm{P}(A)>0$ 且 $\mathrm{P}(B)>0$，則 $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B\mid A)=\mathrm{P}(B)\mathrm{P}(A\mid \mathrm{B})$
5. $\mathrm{P}(A_1\cap A_2\cap\cdots\cap A_k)=\mathrm{P}(A_1)\mathrm{P}(A_2\mid A_1)\mathrm{P}(A_3\mid A_1\cap A_2)\cdots\mathrm{P}(A_k\mid A_1\cap A_2\cap\cdots\cap A_{k-1})$

例 **2–8** 一學童在其左口袋裝有 5 粒藍色 BL 及 4 粒白色 WL 彈珠，在其右口袋中裝有 4 粒藍色 BR 及 5 粒白色 WR 彈珠。若該學童先由左口袋隨機取出一粒彈珠，放入右口袋，然後再由右口袋隨機取出一粒彈珠，問由右口袋取出一粒彈珠為藍色的機率？

解：

$$\begin{aligned}\mathrm{P}(BR) &= \mathrm{P}(BL\cap BR)+\mathrm{P}(WL\cap BR)\\ &= \mathrm{P}(BL)\mathrm{P}(BR\mid BL)+\mathrm{P}(WL)\mathrm{P}(BR\mid WL)\\ &= \frac{5}{9}\times\frac{5}{10}+\frac{4}{9}\times\frac{4}{10}\\ &= \frac{41}{90}\end{aligned}$$

例 **2–9** Suppose that $\mathrm{P}(A)=0.7$, $\mathrm{P}(B)=0.5$, and $\mathrm{P}\left[(A\cup B)'\right]=0.1$

1. Find $\mathrm{P}(A\cap B)$
2. Find $\mathrm{P}(A|B)$

解：

由於 $\mathrm{P}(A\cup B)=1-0.1=0.9$，故得

1.
$$\mathrm{P}(A\cap B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cup B)=0.7+0.5-0.9=0.3$$

2.
$$\mathrm{P}(A|B)=\frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}=\frac{0.3}{0.5}=0.6$$

2.4.3 獨立事件

若二事件 A，B 中任一事件發生並不影響另一事件發生的機率即

$$\mathrm{P}(A\mid B)=\mathrm{P}(A) \text{ 或 } \mathrm{P}(B\mid A)=\mathrm{P}(B),$$

稱為相依事件。

定理 2.1. 若 A 與 B 為獨立事件，則

1. A 與 B'

2. A' 與 B

3. A' 與 B'

亦各為獨立事件

證明.

1.

$$\begin{aligned}
\mathrm{P}\left(A \cap B'\right) &= \mathrm{P}\left(A\right)\mathrm{P}\left(B' \mid A\right) = \mathrm{P}\left(A\right)\left[1 - \mathrm{P}\left(B \mid A\right)\right] \\
&= \mathrm{P}\left(A\right)\left[1 - \mathrm{P}\left(B\right)\right] \\
&= \mathrm{P}\left(A\right)\mathrm{P}\left(B'\right)
\end{aligned}$$

2. 同理可證

3.

$$\mathrm{P}\left(A' \cap B'\right) = \mathrm{P}\left(A'\right)\mathrm{P}\left(B' \mid A'\right) = \mathrm{P}\left(A'\right)\left[1 - \mathrm{P}\left(B \mid A'\right)\right]$$

由 (2) 可知，A' 與 B 相獨立，所以 $\mathrm{P}(B \mid A') = \mathrm{P}(B)$ 故可推得

$$\begin{aligned}
\mathrm{P}\left(A' \cap B'\right) &= \mathrm{P}\left(A'\right)\left[1 - \mathrm{P}\left(B\right)\right] \\
&= \mathrm{P}\left(A'\right)\mathrm{P}\left(B'\right)
\end{aligned}$$

例 **2–10** 擲兩枚公正的骰子，其中一為紅色另一為白色，令 A 表紅骰子出現 4 點的事件，B 表二骰子點數和為奇數的事件，試問 A，B 二事件是否獨立？

解：

事件 $A = \{(4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6)\}$ 得 $\mathrm{P}(A) = 6/36 = 1/6$

事件 $B = \{(1,2)\ (1,4),\ (1,6),\ \cdots,(6,1),\ (6,3),\ (6,5)\}$ 得 $\mathrm{P}(B) = 18/36 = 1/2$

故 $A \cap B = \{(4,1),\ (4,3),\ (4,5)\}$ 得 $\mathrm{P}(A \cap B) = 3/36 = 1/12$。所以 $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$ 成立，故 A, B 獨立。

例 **2–11** 一罐中裝有四個同樣的小球, 分別註有 1, 2, 3, 4 號碼，今由罐中隨機取出一球，定義事件 $A=\{1,2\}$，事件 $B=\{1,3\}$, 事件 $C=\{1,4\}$，試問 A, B, C 是否獨立？

解：

$\mathrm{P}(A)=\mathrm{P}(B)=\mathrm{P}(C)=0.5$, 但 $A\cap B\cap C=1$, 得 $\mathrm{P}(A\cap B\cap C)=0.25$。所以 $\mathrm{P}(A\cap B\cap C)=0.25\neq 0.125=\mathrm{P}(A)\mathrm{P}(B)\mathrm{P}(C)$，故 A, B, C 為相依

注意：此題兩兩獨立，並不代表全部皆獨立。

例 **2–12** 假設 A 和 B 為兩個事件 (events)，如果 $A\cap B$ 的機率為 0，則 A 和 B 為獨立事件。

解：

$$A\text{和}B\text{為獨立事件}\Leftrightarrow \mathrm{P}\left(A\right)=\mathrm{P}\left(A|B\right)=\frac{\mathrm{P}\left(A\cap B\right)}{\mathrm{P}\left(B\right)}$$

所以 $\mathrm{P}\left(A\cap B\right)=0$ 表示 A 和 B 為互斥事件。

2.4.4 貝氏定理

定理 2.2. 設 $\{A_1,\ A_2,\cdots,A_k\}$ 為樣本空間 S 的一分割，且 A_i 的事前機率 $\mathrm{P}(A_i)>0,\ i=1,2,\cdots,k$。若 B 為 S 中的另一事件，且 $\mathrm{P}(B)>0$，則

$$\mathrm{P}(A_j\mid B)=\frac{\mathrm{P}(A_j)\times\mathrm{P}(B\mid A_j)}{\sum_{i=1}^{k}\mathrm{P}(A_i)\times\mathrm{P}(B\mid A_i)},\ i=1,2,\ldots,k$$

證明.

$$\begin{aligned}
\mathrm{P}\left(A_j|\,B\right) &= \frac{\mathrm{P}\left(A_j\cap B\right)}{\mathrm{P}\left(B\right)}\\
&= \frac{\mathrm{P}\left(A_j\cap B\right)}{\sum_{i=1}^{n}\mathrm{P}\left(A_i\cap B\right)}\\
&= \frac{\mathrm{P}\left(A_j\cap B\right)}{\mathrm{P}\left(A_1\cap B\right)+\mathrm{P}\left(A_2\cap B\right)+\cdots+\mathrm{P}\left(A_n\cap B\right)}\\
&= \frac{\mathrm{P}\left(A_j\right)\mathrm{P}\left(B|\,A_j\right)}{\mathrm{P}\left(A_1\right)\mathrm{P}\left(B|\,A_1\right)+\mathrm{P}\left(A_2\right)\mathrm{P}\left(B|\,A_2\right)+\cdots+\mathrm{P}\left(A_n\right)\mathrm{P}\left(B|\,A_n\right)}\\
&= \frac{\mathrm{P}\left(A_j\right)\mathrm{P}\left(B|\,A_j\right)}{\sum_{i=1}^{n}\mathrm{P}\left(A_i\right)\mathrm{P}\left(B|\,A_i\right)}
\end{aligned}$$

例 **2–13** 某工廠有兩部機器 A_1 和 A_2 生產相同的產品，已知機器 A_1 生產全部產品的 40%，機器 A_2 生產全部產品的 60%，又知機器 A_1 生產的不良率有 4%，機器 A_2 生產的不良率有 5%，試問

1. 由全部產品中任抽出一件，其為不良品的機率？
2. 已知其為不良品，由機器 A_1 產出的機率？

解：

1. $\mathrm{P}(B)=\mathrm{P}(A_1)\mathrm{P}(B \mid A_1)+\mathrm{P}(A_2)\mathrm{P}(B \mid A_2)=0.4\times 0.04+0.6\times 0.05=0.046$
2. $\mathrm{P}(A_1 \mid B)=\dfrac{\mathrm{P}(A_1)\mathrm{P}(B \mid A_1)}{\mathrm{P}(B)}=\dfrac{0.016}{0.046}=\dfrac{8}{23}$

例 **2–14** In a certain city, 30% of the people are republic, 50% are democrats, and 20% are independent. Records show that in a particular election, 65% of Republicans voted 82% of the democrats voted and 50% of independent voted. A person is chosen at random from this city and it is known that he/she did not vote, what is the probability that he/she is a democrat？

解：

$$\mathrm{P}(\text{民主黨員} \mid \text{沒有投票})=\frac{0.5\times 0.18}{0.3\times 0.35+0.5\times 0.18+0.2\times 0.5}=0.30508$$

例 **2–15** Bowl A contains two red chip; bowl B contains two white chips; and bowl C contains one red chip and one white chip. A bowl is selected at random (with equal probability), and one chip is taken at random from the bowl.

1. Compute the probability of selecting a white chip.
2. If the chip selected is white, compute the conditional probability that the other chip in the bowl is red.

解：

1.

$$\begin{aligned}
\mathrm{P}\left(\text{white}\right) &= \mathrm{P}\left(\text{white}\cap A\right)+\mathrm{P}\left(\text{white}\cap B\right)+\mathrm{P}\left(\text{white}\cap C\right)\\
&= \mathrm{P}\left(A\right)\mathrm{P}\left(\text{white}|A\right)+\mathrm{P}\left(B\right)\mathrm{P}\left(\text{white}|B\right)+\mathrm{P}\left(C\right)\mathrm{P}\left(\text{white}|C\right)\\
&= \frac{1}{3}\times 0+\frac{1}{3}\times 1+\frac{1}{3}\times\frac{1}{2}\\
&= \frac{1}{2}
\end{aligned}$$

2.

$$\mathrm{P}\left(C|\text{white}\right)=\frac{\mathrm{P}\left(\text{white}\cap C\right)}{\mathrm{P}\left(\text{white}\right)}=\frac{1/3\times 1/2}{1/2}=\frac{1}{3}$$

例 **2–16** 有 80% 的新進員工受過訓，而受過訓的員工有 88% 能符合需求，未受過訓的員工只有 30% 能符合需求，求一新進員工符合需求的機率？

解：

$$P\left(\text{新進員工} \cap \text{符合需求}\right) = 0.8 \times 0.88 + 0.2 \times 0.3 = 0.764$$

第 3 章　隨機變數與機率分配

3.1　隨機變數

隨機變數（Random Variable, 簡稱 r.v.）：在一隨機實驗中，以其樣本空間 S 中的每一元素為定義域, 經一實值函數 X，得唯一一個實數值 x，以集合 $\{x \mid x = X(S),\ \ s \in S\}$ 為值域，即 $X(s) = x$。一般而言, 隨機變數常以大寫英文字母表示 X，Y, Z, $\ldots$; 隨機變數的函數值或觀測值常以小寫字母書寫 x，y, z $\ldots$ 表示之。

例　**3–1**　擲一枚公正的銅板兩次，即 $S = \{\{正,\ 正\}, \{正,\ 反\}, \{反,\ 正\}, \{反,\ 反\}\}$, 令 X 表示正面的個數，即 $X(\{正, 正\}) = 2$, $X(\{正, 反\}) = X(\{反, 正\}) = 1$, $X(\{反, 反\}) = 0$。

3.2　機率分配

機率分配可分為

1. **離散型隨機變數**：可能值之個數為有限個或無限個但可計數，例如：不良品的個數、意外事件的次數。
2. **連續型隨機變數**：可能值之個數為無限且不可計數，例如：身高、體重、溫度、時間。

茲介紹如下：

1. **離散型隨機變數的機率分配**：

   定義：設 X 為一離散型隨機變數 (Discrete Random Varible)，若所有 $x \in X$，則 $f(x) = \mathrm{P}(X =$

x) 稱為 X 的機率質量函數（Probability Mass Function, 簡稱 p.m.f.）。X 的 p.m.f., $f(x)$ 為一實值函數，須滿足下列三性質：

(a) $f(x) \geq 0, \forall\, x \in R$

(b) $\sum_{x \in R} f(x) = 1$

(c) $\mathrm{P}(x \in A) = \sum_{x \in A} f(x)$

2. **連續型隨機變數的機率分配**

定義：設 X 為一連續型隨機變數 (Continuous Random Variable)，若所有 $x \in X$，則 $f(x) = \mathrm{P}(X = x)$ 稱為 X 的機率密度函數（Probability Density Function, 簡稱 p.d.f.）。X 的 p.d.f., $f(x)$ 為一實值函數，須滿足下列三性質：

(a) $f(x) \geq 0, -\infty < x < \infty$

(b) $\int_{-\infty}^{\infty} f(x)dx = 1$

(c) $\mathrm{P}(a \leq x \leq b) = \int_a^b f(x)dx$

(d) $f(x)$ 為一連續型 p.d.f.，並不代表事件 $\{X = x\}$ 的機率，即 $\mathrm{P}(X = x) = 0$。

(e) 連續型的機率，指函數曲線與該區間所構成的面積大小。

(f) $\mathrm{P}(a \leq x \leq b) = \mathrm{P}(a < x \leq b) = \mathrm{P}(a \leq x < b) = \mathrm{P}(a < x < b)$。

例 **3–2** 設下列各函數是 X 的 p.d.f.，試求常數 C 之值？

1. $f(x) = x/C,\ x = 1, 2, 3, 4$
2. $f(x) = Cx,\ x = 1, 2,\ \cdots 10$
3. $f(x) = C(x+1)^2,\ x = 0, 1, 2, 3$
4. $f(x) = x/C,\ x = 1, 2, 3, \cdots, n$

解：

1. $C = 1 + 2 + 3 + 4 = 10$
2. $C = \dfrac{1}{1 + 2 + 3 + \cdots + 10} = \dfrac{1}{55}$
3. $C = \dfrac{1}{(0+1)^2 + (1+1)^2 + (2+1)^2 + (3+1)^2} = \dfrac{1}{30}$
4. $C = 1 + 2 + 3 + \cdots + n = \dfrac{n\,(n+1)}{2}$

3.3 累積分配函數

累積分配函數其定義與符號為

$$F(x) = \mathrm{P}(X \le x) = \begin{cases} \sum_{t \le x} f(t) & ,\text{離散型} \\ \int_{-\infty}^{x} f(t)\,dt & ,\text{連續型} \end{cases}$$

性質

1. $0 \le F(x) \le 1$
2. $\lim\limits_{x \to -\infty} = 0,\ \lim\limits_{x \to \infty} = 1$
3. $F(x)$ 為一非遞減函數
   (a) 若 $f(x)$ 為一離散型分配，則 $F(x)$ 為一階梯函數
   (b) 若 $f(x)$ 為一連續型分配，則 $F(x)$ 為一連續函數
4. (a) 若 $f(x)$ 為一離散型分配，則 $\mathrm{P}(a < x \le b) = F(b) - F(a)$
   (b) 若 $f(x)$ 為一連續型分配，則 $\mathrm{P}(a \le x \le b) = F(b) - F(a)$

3.4 聯合機率密度函數

聯合機率密度函數為討論兩個以上 (含) 隨機變數的機率分配。設 $f(x, y)$ 為 X 與 Y 的聯合機率密度函數 (joint pdf)，則必須滿足下列三條件：

1. $0 \le f(x, y) \le 1,\ (x, y) \in \Re^2$
2. $\begin{cases} \sum\sum_{(x,y) \in R^2} f(x, y) = 1, & \text{離散分配} \\ \int\int_{(x,y) \in R^2} f(x, y)\,dxdy = 1, & \text{連續分配} \end{cases}$
3. $\mathrm{P}((x, y) \in A) = \begin{cases} \sum\sum_{t \le x} f(x, y), & A \in \Re^2 \\ \int\int_{-\infty}^{x} f(x, y)\,dxdy, & A \in \Re^2 \end{cases}$

3.5 邊際機率密度函數

設 $f(x,y)$ 為 r.v. X 和 Y 的聯合機率密度函數，則 X 與 Y 的邊際機率密度函數 (marginal pdf) 分別定義為

$$f_1(x) = \begin{cases} \sum_{y\in\Re^2} f(x,y) \\ \int_{y\in\Re^2} f(x,y)\,dy \end{cases}$$

及

$$f_2(y) = \begin{cases} \sum_{x\in\Re^2} f(x,y) \\ \int_{x\in\Re^2} f(x,y)\,dx \end{cases}$$

同理， $f_1(x)$ 必須滿足下列條件：

1. $0 \le f_1(x) \le 1,\ x \in R_1$
2. $\begin{cases} \sum_{x\in R_1} f_1(x) = 1 \\ \int_{x\in R_1} f_1(x)dx = 1 \end{cases}$

例 **3–3** 令 X 與 Y 的 joint pdf 定義為

$$f(x,y) = \frac{xy^2}{30},\ x = 1,\ 2,\ 3,\ y = 1,\ 2$$

試求 X 與 Y 之邊際機率密度函數？

解：

$$f_1(x) = \sum_{y=1}^{2} \frac{xy^2}{30} = \frac{x}{30} + \frac{4x}{30} = \frac{x}{6},\ x = 1,2,3$$

$$f_2(y) = \sum_{x=1}^{3} \frac{xy^2}{30} = \frac{y^2}{30} + \frac{2y^2}{30} + \frac{3y^2}{30} = \frac{y^2}{5},\ y = 1,2$$

例 **3–4** If X and Y are discrete random variables with joint probability density function,

$$f(x,y) = c\frac{2^{x+y}}{x!y!}, x = 0,1,2,\ldots; y = 0,1,2,\ldots,$$

and zero otherwise. Find the constant c？

解：

$$\sum_{x=0}^{\infty}\sum_{y=0}^{\infty} c\frac{2^{x+y}}{x!y!} = c\sum_{x=0}^{\infty}\sum_{y=0}^{\infty}\frac{1}{x!y!}2^x2^y = c\sum_{x=0}^{\infty}\frac{2^x}{x!}\sum_{y=0}^{\infty}\frac{2^y}{y!} = c\times e^2\times e^2 = ce^4$$

所以

$$c = e^{-4}$$

3.6 條件機率密度函數

設 $f(x,y)$ 為 r.v. X 和 Y 的聯合機率密度函數，其邊際 p.d.f. 分別為 $f_1(x)$ 與 $f_2(y)$，則在給定 $X=x$ 時，Y 的條件機率密度函數 (conditional pdf) 定義為：

$$f(y \mid X=x) = \frac{f(x,y)}{f_1(x)},\ \ f_1(x)>0,\ y\in R_2$$

同理，在給定 $Y=y$ 下之條件機率密度函數為

$$f(x \mid Y=y) = \frac{f(x,y)}{f_2(x)},\ \ f_2(x)>0,\ y\in R_1$$

性質

1. $f(y \mid x)\geq 0,\ f(x \mid y)\geq 0$
2. $\sum_{y\in R_2} f(y \mid x)=1,\ \sum_{x\in R_1} f(x \mid y)=1$
3. $\int_{-\infty}^{\infty} f(y \mid x)dy=1,\ \int_{-\infty}^{\infty} f(x \mid y)dx=1$

定義

r.v. X 與 Y 相互獨立（mutually independent），若 $f(x \mid y)=f_1(x),\ x\in R_1$ 或 $f(y \mid x)=f_2(y)$, $y\in R_2$ 或 $f(x,y)=f_1(x)f_2(y),\ (x,y)\in R^2$。否則，稱兩隨機變數 X, Y 為相依 (dependent)。

例 **3–5**　令 X 與 Y 的 joint pdf 定義為 $f(x,y)=xy^2/30,\ x=1,2,3,\ y=1,2$。試問 X 與 Y 是否獨立？

解：

由例 3-3可知

$$f_1(x) = \frac{x}{6}\,,\ \ x = 1, 2, 3$$

$$f_2(y) = \frac{y^2}{5}\,,\ \ y = 1, 2$$

又

$$f(x, y) = \frac{xy^2}{30} = \frac{x}{6} \times \frac{y^2}{5} = f_1(x)f_2(y),$$

所以 X 和 Y 相互獨立。

例 **3–6** Suppose that a point (X, Y) is chosen at random from the circle S defined as follows：

$$S : \{(x, y)(x-1)^2 + (y+2)^2 \leq 9\}.$$

Then find $P(Y > 0|X = 2) =$？

解：

$$P\left(Y > 0|X = 2\right) = \frac{\sqrt{9-1^2}-2}{2 \times \sqrt{9-1^2}} = 0.14645$$

例 **3–7** 淨重一公斤的乖乖桶中計有軟糖、果凍、餅乾等三種產品，假設 X 與 Y 分別代表乖乖桶中軟糖與果凍的重量，已知 X 與 Y 的聯合機率密度函數如下：

$$f(x, y) = 24xy, 0 < x < 1, 0 < y < 1, 0 < x + y < 1$$

1. 餅乾的重量超過 0.5 公斤的機率為何？
2. 若軟糖的重量為 0.75 公斤，試問果凍重量小於 0.1 公斤的機率為何？

解：

1.

$$\begin{aligned} P\left(\text{餅乾的重量超過 0.5 公斤}\right) &= P\left(\text{軟糖與果凍的重量小於 0.5 公斤}\right) \\ &= P\left(X + Y < 0.5\right) \\ &= \int_0^{0.5} \int_0^{0.5-x} 24xydydx \end{aligned}$$

$$
\begin{aligned}
&= \int_0^{0.5} 12x\,(0.5-x)^2\,dx \\
&= 0.0625
\end{aligned}
$$

2. X 的邊際機率密度函數為

$$
\begin{aligned}
f_1(x) &= \int_0^{1-x} 24xydy \\
&= 12x - 24x^2 + 12x^3
\end{aligned}
$$

所以在給定 $X = x$ 下，Y 的條件機率密度函數為

$$
f(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{24xy}{12x - 24x^2 + 12x^3}, 0 < y < 1 - x
$$

因此當 $X = 0.75$ 時，$Y < 0.1$ 的機率為

$$
P(Y < 0.1|X = 0.75) = \int_0^{0.1} 32ydy = 0.16
$$

3.7 期望值與變異數

3.7.1 期望值

1. 若單一隨機變數 X，其 p.d.f 為 $f(x)$，則 X 的期望值定義為

$$
\mu = \mathrm{E}X = \begin{cases} \sum_{x\in\Re} f(x), & \text{離散分配} \\ \int_{x\in\Re} xf(x)\,dx, & \text{連續分配} \end{cases}
$$

性質

(a) 設 a, b 為常數，則 $\mathrm{E}(aX + b) = aEX + b$

(b) 設 a, b 為常數，$g(X)$ 為 r.v. X 的函數，則 $\mathrm{E}[ag(X) + b] = aE[g(X)] + b$

(c) 設 a, b 為常數，$g(X)$ 與 $h(X)$ 皆為 r.v. X 的函數，則 $\mathrm{E}[ag(X) \pm bh(X)] = aE[g(X)] \pm bE[h(X)]$

(d) 當母體分配已知時，一般以樣本平均數 $\overline{X} = \sum_{i=1}^{n} X_i$ 做為母體平均數 μ 之估計式

2. 探討兩個 r.v.，其 joint pdf 為 $f(x,y)$，又設 $g(x,y)$ 為 X 和 Y 的函數則其期望值定義為

$$\mathrm{E}\left[g\left(x,y\right)\right]=\begin{cases}\sum_{x\in\Re_1}\sum_{y\in\Re_2}g\left(x,y\right)f\left(x,y\right) & ,\text{離散型}\\ \int_{x\in\Re_1}\int_{y\in\Re_2}g\left(x,y\right)f\left(x,y\right)dxdy & ,\text{連續型}\end{cases}$$

性質

(a) 設 a, b, c 為常數，則 $\mathrm{E}[aX\pm bY\pm c]=a\mathrm{E}X\pm b\mathrm{E}Y\pm c$

(b) 若 r.v. X 和 Y 相互獨立，則 $\mathrm{E}(XY)=\mathrm{E}X\mathrm{E}Y$，其逆不真

(c) 若 r.v. X 和 Y 相互獨立，則 $\mathrm{E}[g(X)h(Y)]=\mathrm{E}[g(X)]\mathrm{E}[h(Y)]$，其逆不真

證明.

(a)

$$\begin{aligned}\mathrm{E}\left(aX\pm bY\pm c\right) &= \sum_{x\in\Re_1}\sum_{y\in\Re_2}\left(ax\pm by\pm c\right)f\left(x,y\right)\\ &= \sum_{x\in\Re_1}\sum_{y\in\Re_2}axf\left(x,y\right)\pm\sum_{y\in\Re_2}\sum_{x\in\Re_1}byf\left(x,y\right)\pm\sum_{x\in\Re_1}\sum_{y\in\Re_2}cf\left(x,y\right)\\ &= \sum_{x\in\Re_1}axf_1\left(x\right)\pm\sum_{y\in\Re_2}byf_2\left(y\right)\pm c\\ &= a\mathrm{E}X\pm b\mathrm{E}Y\pm c\end{aligned}$$

(b) 因為 X 與 Y 獨立，所以 $f(x,y)=f_1(x)f_2(y)$。由定義可知

$$\begin{aligned}\mathrm{E}(XY) &= \sum_{x\in\Re_1}\sum_{y\in\Re_2}xyf\left(x,y\right)\\ &= \sum_{x\in\Re_1}\sum_{y\in\Re_2}xyf_2(y)f_1(x)\\ &= \sum_{x\in\Re_1}x\mathrm{E}Yf_1(x)\\ &= \mathrm{E}X\mathrm{E}Y\end{aligned}$$

(c)

$$\mathrm{E}[g(X)h(Y)] = \sum_{x\in\Re_1}\sum_{y\in\Re_2}g(x)h(y)f\left(x,y\right)$$

$$
\begin{aligned}
&= \sum_{x\in\Re_1}\sum_{y\in\Re_2} g(x)h(y)f_2(y)f_1(x) \\
&= \sum_{x\in\Re_1} g(x)\mathrm{E}[h(Y)]f_1(x) \\
&= \mathrm{E}[g(X)]\mathrm{E}[h(Y)]
\end{aligned}
$$

3. 若給定 $X=x$ 時，Y 的條件機率密度函數為 $f(y\mid x),\, y\in\Re_2$ 則定義 Y 的條件期望值為

$$
\mathrm{E}\left[Y\mid X=x\right]=\begin{cases}\sum_{y\in\Re_2} yf\left(y\mid x\right) & ,\text{離散型}\\ \int_{y\in\Re_2} yf\left(y|\,x\right)dy & ,\text{連續型}\end{cases}
$$

同理，可定義 X 的條件期望值，當給定 $Y=y$ 時為

$$
\mathrm{E}\left[X\mid Y=y\right]=\begin{cases}\sum_{x\in\Re_1} xf\left(x\mid y\right) & ,\text{離散型}\\ \int_{x\in\Re_1} xf\left(x\mid y\right)dx & ,\text{連續型}\end{cases}
$$

例 **3–8** 擲一枚公正骰子, 看點數得多少錢，換言之，出現一點給一元, 出現兩點給二元, 以此類推, 但玩一次 4 元，試問此人玩一次的期望值？

解：

令 X 表中獎金額，即

$$
\begin{aligned}
X &= \{(1-4),(2-4),(3-4),(4-4),(5-4),(6-4)\} \\
&= \{-3,\ -2,\ -1,\ 0,\ 1,\ 2\}
\end{aligned}
$$

則

$$
\mathrm{E}X=\sum_{x=1}^{6} xf\left(x\right)=\frac{-3\times 1}{6}+\frac{-2\times 1}{6}+\frac{-1\times 1}{6}+\frac{0\times 1}{6}+\frac{1\times 1}{6}+\frac{2\times 1}{6}=-0.5
$$

例 **3–9** 令 $f(x,y)$ 為隨機變數 X 與 Y 之聯合機率密度函數，若

$$
f(x,y)=\frac{x+y}{21},\ x=1,\,2,\,3\ \text{及}\ y=1,2
$$

求 $\mathrm{E}[Y\mid X=3]$？

解：

$$f_1(x) = \sum_{y=1}^{2} \frac{x+y}{21} = \frac{2x+3}{21},\ x = 1,\ 2,\ 3$$

$$f_2(y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{3y+6}{21},\ y = 1,\ 2$$

由定義

$$f(y \mid x) = \frac{\frac{x+y}{21}}{\frac{2x+3}{21}} = \frac{x+y}{2x+3}, x = 1,\ 2,\ 3 \text{ 及 } y = 1,\ 2$$

所以

$$f(y \mid x = 3) = \frac{3+y}{9},\ y = 1,\ 2$$

可得

$$\mathrm{E}[Y \mid X = 3] = \sum_{y=1}^{2} y \times \frac{3+y}{9} = 1 \times \frac{4}{9} + 2 \times \frac{5}{9} = \frac{14}{9}$$

例 **3–10** 有三個門，選擇第一個門後能在 2 小時後安全離開，選擇第二個門後會在 4 小時後回到原地，選擇第三個門後會在 5 小時後回到原地，選擇這三個門的機率分別為 1/6、1/3、1/2，求安全離開的時間期望值。

解：

令安全離開的時間期望值 $= \mu$，所以

$$\frac{1}{6} \times 2 + \frac{1}{3} \times (4 + \mu) + \frac{1}{6} \times (5 + \mu) = \mu$$

解方程式得 $\mu = 10$

3.7.2 變異數

1. 單一隨機變數 X, 其 p.d.f. 為 $f(x)$, $x \in R$，則定義 X 的變異數為

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X - \mu)^2] = \mathrm{E}X^2 - \mu^2$$

此外，定義標準差（Standard Deviation）為 $\sigma = \sqrt{\sigma^2}$

性質

(a) 設 a, b 皆為常數，則 $\text{Var}(aX+b) = a^2\text{Var}(X)$

(b) 設 $g(x)$ 為 X 的函數，則 $\text{Var}[g(X)] = \text{E}[g(X)]^2 - \text{E}^2[g(X)]$

(c) 當母體分配已知時，一般以樣本變異數 $S^2 = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2/n-1$ 做為母體平均數 σ^2 之估計式

證明.

(a)

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

2. 若給定 $X=x$ 時，Y 的條件機率密度函數為 $f(y \mid x),\, y \in \Re_2$ 則定義 Y 的條件期望值為

$$\text{Var}\left[Y \mid X=x\right] = \begin{cases} \sum_{y\in\Re_2}\left(y-\text{E}\left[Y \mid X=x\right]\right)^2 f\left(y \mid x\right), & \text{離散型} \\ \int_{y\in\Re_2}\left(y-\text{E}\left[Y \mid X=x\right]\right)^2 f\left(y \mid x\right)dy, & \text{連續型} \end{cases}$$

同理，可定義 X 的條件期望值，當給定 $Y=y$ 時為

$$\text{Var}\left[X \mid Y=y\right] = \begin{cases} \sum_{x\in\Re_1}\left(x-\text{E}\left[X \mid Y=y\right]\right)^2 f\left(x \mid y\right), & \text{離散型} \\ \int_{x\in\Re_1}\left(x-\text{E}\left[X \mid Y=y\right]\right)^2 f\left(x \mid y\right)dx, & \text{連續型} \end{cases}$$

3. 兩個隨機變數 X 和 Y, 其 joint pdf 為 $f(x,y),\, x\in\Re_1,\, y\in\Re_2$ 。設 $g(x,\ y)$ 為 X 和 Y 的函數，則定義其變異數為

$$\text{Var}[g(X,\ Y)] = \text{E}[g(X,\ Y) - \text{E}(g(X,Y))]^2$$

例 **3–11** 承例 **3.7** 求 $\text{Var}[Y \mid X=3]$？

解：

$$\mathrm{Var}[Y \mid X=3] = \sum_{y=1}^{2} \left(y - \frac{14}{9}\right)^2 \times \frac{3+y}{9} = \frac{25}{81} \times \frac{4}{9} + \frac{16}{81} \times \frac{5}{9} = \frac{20}{81}$$

例 **3–12** 假設隨機變數 X 和 Y 之聯合機率密度函數為

$$f(x,y) = 2,\ 0 \le x \le y \le 1$$

試求 $\mathrm{E}[Y \mid X]$ 及 $\mathrm{Var}[Y \mid X]$

解：

$$f_1(x) = \int_x^1 2dy = 2y\Big|_x^1 = 2(1-x)$$

所以

$$f(y \mid x) = \frac{2}{2(1-x)} = \frac{1}{1-x}$$

又

$$\mathrm{E}[Y \mid X] = \int_x^1 y\frac{1}{1-x}dy = \frac{1}{1-x} \times \frac{1}{2}y^2\Big|_x^1 = \frac{1+x}{2},$$

$$\mathrm{E}[Y^2 \mid X] = \int_x^1 y^2\frac{1}{1-x}dy = \frac{1}{1-x} \times \frac{1}{3}y^3\Big|_x^1 = \frac{1+x+x^2}{3}$$

因此

$$\begin{aligned}
\mathrm{Var}[Y \mid X] &= \mathrm{E}[\{Y - \mathrm{E}[Y \mid X]\}^2 \mid X] \\
&= \mathrm{E}[Y^2 + \{\mathrm{E}[Y \mid X]\}^2 - 2YE[Y \mid X] \mid X] \\
&= \mathrm{E}[Y^2 \mid X] + \mathrm{E}[\{\mathrm{E}[Y \mid X]\}^2 \mid X] - 2\mathrm{E}[YE[Y \mid X] \mid X] \\
&= \mathrm{E}[Y^2 \mid X] + \{\mathrm{E}[Y \mid X]\}^2 - 2\mathrm{E}[Y \mid X]\mathrm{E}[Y \mid X] \\
&= \mathrm{E}[Y^2 \mid X] - \{\mathrm{E}[Y \mid X]\}^2 \\
&= \frac{1+x+x^2}{3} - \left(\frac{1+x}{2}\right)^2 \\
&= \frac{1-2x+x^2}{12}
\end{aligned}$$

例 3–13 Let (X,Y) have joint density function $f(x,y)=8xy$, for $0<x<y<1$.

1. What is the conditional density function of X given $Y=y$？
2. What is the expected value and the variance of X given $Y=y$？

解：

1.

$$f_2(y)=\int_0^y 8xydx=4y^3$$

$$f(x|y)=\frac{8xy}{4y^3}=\frac{2x}{y^2}$$

2.

$$\mathrm{E}[X|Y=y]=\int_0^y x\frac{2x}{y^2}dx=\frac{2y}{3}$$

及

$$\mathrm{E}\left[X^2|Y=y\right]=\int_0^y x^2\frac{2x}{y^2}dx=\frac{y^2}{2}$$

所以

$$\mathrm{Var}[X|Y=y]=\frac{y^2}{2}-\left(\frac{2y}{3}\right)^2=\frac{y^2}{18}$$

例 3–14 Assume that you are given the following joint density function,

$$f(x,y)=\begin{cases} 8xy, & 0\le x\le 1, 0\le y\le x \\ 0, & \text{otherwise} \end{cases}$$

1. Please find $\mathrm{E}X$
2. Please find $\mathrm{Var}(X)$

解：

1.

$$f_x=\int_0^x 8xydy=8x\times\frac{1}{2}y^2\Big|_0^x=4x^3, 0\le x\le 1$$

$$\mathrm{E}X=\int_0^1 4x^4dx=\frac{4}{5}$$

2.

$$\mathrm{E}X^2 = \int_0^1 4x^5 dx = \frac{4}{6}$$

$$\mathrm{Var}\,(X) = \frac{4}{6} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}$$

例 **3–15** 若

$$f\,(x) = \frac{1}{\pi\,(1+x^2)}, -\infty < x < \infty$$

試求

1. $f(x)$ 為一機率密度函數
2. $\mathrm{E}X$ 及 $\mathrm{Var}(X)$ 不存在

解：

1.

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{1}{\pi\,(1+x^2)} dx &= \frac{2}{\pi} \int_0^{\infty} \frac{1}{1+x^2} dx \\
&= \frac{2}{\pi} \tan^{-1} x \Big|_0^{\infty} \\
&= \frac{2}{\pi} \times \frac{\pi}{2} \\
&= 1
\end{aligned}$$

又對所有 x 均有 $f(x) \geq 0$，所以 $f(x)$ 為一機率密度函數，且稱 $f(x)$ 為柯西分配(Chuchy Distribution)。

2.

$$\begin{aligned}
\mathrm{E}X &= \int_{-\infty}^{\infty} \frac{x}{\pi\,(1+x^2)} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2x}{1+x^2} dx \\
&= \frac{1}{2\pi} \ln\left|1+x^2\right| \Big|_{-\infty}^{\infty} \\
&= \infty
\end{aligned}$$

所以 $\mathrm{E}X$ 不存在。又

$$\mathrm{E}X^2 = \int_{-\infty}^{\infty} \frac{x^2}{\pi\,(1+x^2)} dx = \frac{2}{\pi} \int_0^{\infty} \frac{x^2}{1+x^2} dx$$

$$
\begin{aligned}
&= \frac{2}{\pi}\int_0^{\infty}\left(1-\frac{1}{1+x^2}\right)dx \\
&= \frac{2}{\pi}x\Big|_0^{\infty} - \frac{2}{\pi}\tan^{-1}x\Big|_0^{\infty} \\
&= \infty
\end{aligned}
$$

因此 $\mathrm{Var}(X)$ 亦不存在。

例 **3–16** X is a continuous type r.v with p.d.f $f(x)$, $0 < x < b <$. Prove that $\mathrm{E}(X) = \int_0^b [1-F(x)]dx$, where $F(x)$ is the distribution function of X.

解：

由部分積分法可得

$$
\begin{aligned}
\int_0^b [1-F(x)]\,dx &= x[1-F(x)]\Big|_0^b + \int_0^b xf(x)\,dx \\
&= \int_0^b xf(x)\,dx \\
&= \mathrm{E}X
\end{aligned}
$$

例 **3–17** 試證明下列二式成立

$$
\mathrm{EE}[X|Y] = \mathrm{E}X \text{ 及 } \mathrm{Var}(X) = \mathrm{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathrm{E}[X|Y]]
$$

證明.

1. 假設 X 和 Y 均為連續隨機變數，則

$$
\begin{aligned}
\mathrm{EE}[X|Y] &= \int_{-\infty}^{\infty} \mathrm{E}[X|Y]f_2(y)dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x \mid y)f_2(y)dxdy \\
&= \int_{-\infty}^{\infty} x\left\{\int_{-\infty}^{\infty} f(x,y)dy\right\}dx \\
&= \int_{-\infty}^{\infty} xf_1(x)dx \\
&= \mathrm{E}X
\end{aligned}
$$

2.

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}[X - \mathrm{E}[X \mid Y] + \mathrm{E}[X \mid Y] - \mathrm{E}X]^2 \\
&= \mathrm{E}[X - \mathrm{E}[X \mid Y]]^2 + \mathrm{E}[\mathrm{E}[X \mid Y] - \mathrm{E}X]^2 + 2\mathrm{E}\{[X - \mathrm{E}[X \mid Y]] \times [\mathrm{E}[X \mid Y] - \mathrm{E}X]\}
\end{aligned}
$$

由於在給定 $Y = y$ 的情況下，$\mathrm{E}[X \mid Y]$ 與 X 無關且 $\mathrm{E}X$ 為一常數，所以

$$
\begin{aligned}
\mathrm{E}\{[X - \mathrm{E}[X \mid Y]][\mathrm{E}[X \mid Y] - \mathrm{E}X]\} &= \{\mathrm{E}[X \mid Y] - \mathrm{E}X\}\mathrm{E}[X - \mathrm{E}[X \mid Y]] \\
&= \{\mathrm{E}[X \mid Y] - \mathrm{E}X\}[\mathrm{E}X - \mathrm{EE}[X \mid Y]] \\
&= 0
\end{aligned}
$$

所以

$$
\mathrm{Var}[X] = \mathrm{E}[X - \mathrm{E}[X \mid Y]]^2 + \mathrm{E}[\mathrm{E}[X \mid Y] - \mathrm{E}X]^2
$$

又 $\mathrm{E}[X - \mathrm{E}[X \mid Y]]^2 = \mathrm{Var}[X \mid Y]$ 與 X 無關且 $EE[X \mid Y] = \mathrm{E}X$，故得

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}[X|Y] + \mathrm{Var}[\mathrm{E}[X|Y]] \\
&= \mathrm{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathrm{E}[X|Y]]
\end{aligned}
$$

例 **3–18** Suppose that a point X is chosen in accordance with a uniform distribution on the interval $(0, 2)$. After the value $X = x$ has been observed, appoint Y is chosen. If Y is chosen in accordance with a uniform distribution on the interval $(0, x)$, what is the expected value of Y？

解：

$$
\mathrm{E}Y = EE\left[Y|X\right] = \mathrm{E}\left(\frac{0 + X}{2}\right) = \frac{\mathrm{E}X}{2} = \frac{1}{2}
$$

3.8 共變異數與相關係數

1. 共變異數：定義為

$$
\sigma_{xy} = \mathrm{Cov}(X,\ Y) = \mathrm{E}[(X - \mu_x)(Y - \mu_y)] = \mathrm{E}(XY) - \mu_x\mu_y
$$

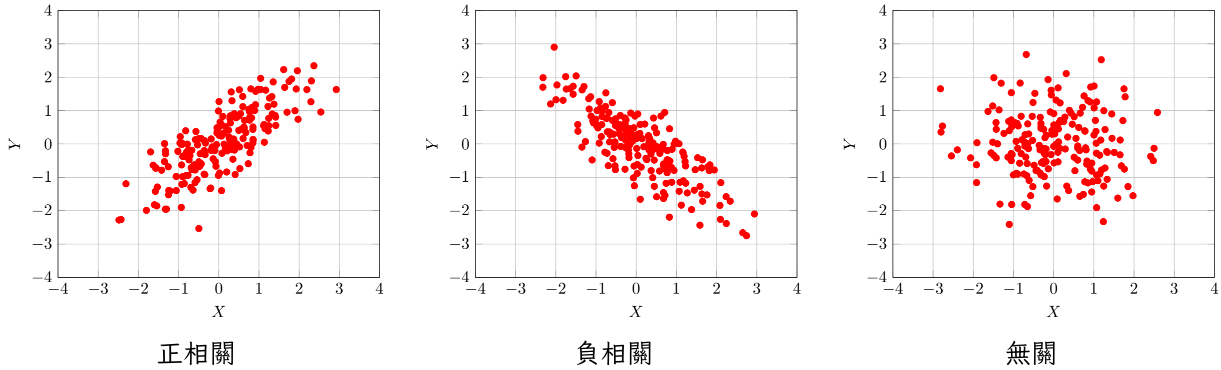


圖 3.1: 相關係數

2. 相關係數：衡量兩變數之間線性關係之強弱, 定義為

$$\rho_{xy} = \frac{\mathrm{Cov}(X,\ Y)}{\sigma_x \sigma_y},$$

其中 $0 < \sigma_x < \infty$ 及 $0 < \sigma_y < \infty$。

相關係數（Correlation coefficient）是用來衡量兩個變數之間的線性關係強弱的指標，通常以 ρ 表示，其值介於 -1 和 1 之間，數值的絕對值越接近 1 表示兩個變數之間的線性相關程度越高；反之，絕對值越接近 0 則表示兩個變數之間不存在線性相關。最重要的是相關係數只能用來描述兩個變數之間的關聯性質，並不能直接用來預測或解釋其中一個變數的變化。一般而言，相關程度的強度則可透過絕對值大小判斷。一般而言，$|r| \geq 0.7$ 被認為是高相關，$0.3 \leq |r| < 0.7$ 為中等相關，$|r| < 0.3$ 為弱相關。

根據圖3.1，當相關係數為正數時，表示當 X 增加時，Y 也會跟著增加；反之，當相關係數為負數時，表示當 X 增加時，Y 會下降。最後，當相關係數為零時，表示不管 X 如何變動，Y 都只是一個隨機亂數。

例 **3–19** 試證明

1. $-1 \leq \rho_{xy} \leq 1$
2. 若 $\mathrm{P}(Y = aX + b) = 1$，則當 $a > 0$ 時，$\rho = 1$; 反之，當 $a < 0$ 時，$\rho = -1$

證明.

1. 令 $h(t) = \mathrm{E}\left[t(X-\mu_x)+(Y-\mu_y)\right]^2$，則

$$
\begin{aligned}
h(t) &= t^2\mathrm{E}(X-\mu_x)^2 + \mathrm{E}(Y-\mu_y)^2 + 2tE(X-\mu_x)(Y-\mu_y) \\
&= t^2\sigma_x^2 + 2t\sigma_{xy} + \sigma_y^2 \\
&= \left(t\sigma_x + \frac{\sigma_{xy}}{\sigma_x}\right)^2 + \sigma_y^2 - \frac{\sigma_{xy}^2}{\sigma_x^2}
\end{aligned}
$$

因此，當 $t = -\frac{\sigma_{xy}}{\sigma_x^2} = -\rho\frac{\sigma_y}{\sigma_x}$ 時，$h(t)$ 有最小值 $\sigma_y^2 - \frac{\sigma_{xy}^2}{\sigma_x^2}$。又由於 $h(t) \geq 0$，即

$$\frac{\sigma_{xy}^2}{\sigma_x^2\sigma_y^2} \leq 1$$

所以

$$-1 \leq \rho_{xy} \leq 1$$

2. $\mathrm{P}(Y = aX+b) = 1$ 即 $Y = aX+b$, 所以

$$
\begin{aligned}
\rho_{xy} &= \frac{\sigma_{xy}}{\sqrt{\sigma_x^2\sigma_y^2}} \\
&= \frac{\mathrm{E}(X-\mu_x)(aX+b-a\mu_x-b)}{|a|\,\sigma_x\sigma_y} \\
&= \frac{aE(X-\mu_x)(X-\mu_x)}{|a|\,\sigma_x\sigma_y} \\
&= \frac{a}{|a|}\frac{\sigma_{xy}}{\sigma_x\sigma_y}
\end{aligned}
$$

所以，

$$\rho = \begin{cases} 1, & \text{當 } a > 0 \\ -1, & \text{當 } a < 0 \end{cases}$$

例 **3–20** Let X and Y be random variable, let μ_x (μ_y) and σ_x^2 (σ_y^2) be the mean and variance of X (Y), and let σ_{xy} be the covariance between X and Y.

1. Show that $\mathrm{E}\left(Y^2\right) = \sigma_y^2 + \mu_y^2$.
2. Show that $\mathrm{E}(XY) = \sigma_{xy} + \mu_x\mu_y$.
3. Show that $\sigma_{xy}^2 \leq \sigma_x^2\sigma_y^2$. [Hine：One possible way to prove this needs to use $\mathrm{Var}(aX+bY) = a^2\sigma_x^2 + 2ab\sigma_{xy} + b^2\sigma_y^2$.]

解：

1.

$$
\begin{aligned}
\sigma_y^2 &= \mathrm{E}\left(Y-\mu_y\right)^2=\mathrm{E}\left(Y^2-2\mu_y Y+\mu_y^2\right) \\
&= \mathrm{E}\left(Y^2\right)-2\mu_y \mathrm{E}Y+\mu_y^2 \\
&= \mathrm{E}\left(Y^2\right)-\mu_y^2
\end{aligned}
$$

故得 $\mathrm{E}\left(Y^2\right)=\sigma_y^2+\mu_y^2$

2.

$$
\begin{aligned}
\sigma_{xy} &= \mathrm{E}\left(X-\mu_x\right)\left(Y-\mu_y\right)=\mathrm{E}\left(XY-\mu_y X-\mu_x Y+\mu_x\mu_y\right) \\
&= \mathrm{E}\left(XY\right)-\mu_y \mathrm{E}X-\mu_x \mathrm{E}Y+\mu_x\mu_y \\
&= \mathrm{E}\left(XY\right)-\mu_x\mu_y
\end{aligned}
$$

3.

$$
\begin{aligned}
\mathrm{Var}\left(aX+bY\right) &= a^2\sigma_x^2+2ab\sigma_{xy}+b^2\sigma_y^2 \\
&= a^2\left(\sigma_x^2+2\times\sigma_x\times\frac{b\sigma_{xy}}{a\sigma_x}+\frac{b^2\sigma_{xy}^2}{a^2\sigma_x^2}\right)+b^2\sigma_y^2-\frac{b^2\sigma_{xy}^2}{\sigma_x^2} \\
&= a^2\left(\sigma_x^2+\frac{b\sigma_{xy}}{a\sigma_x}\right)^2+b^2\left(\sigma_y^2-\frac{\sigma_{xy}^2}{\sigma_x^2}\right) \\
&\geq 0
\end{aligned}
$$

故得 $\sigma_y^2-\frac{\sigma_{xy}^2}{\sigma_x^2}\geq 0\Rightarrow\sigma_{xy}^2\leq\sigma_x^2\sigma_y^2$.

性質

1. $\mathrm{Var}[g(X,\ Y)]=\mathrm{E}[g^2(X,Y)]-(\mathrm{E}[g(X,Y)])^2$
2. $\mathrm{Var}(X\pm Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)\pm 2\mathrm{Cov}(X,Y)$
3. 設 a,b 為常數，則 $\mathrm{Var}(aX\pm bY)=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)\pm 2ab\mathrm{Cov}(X,\ Y)$
4. 倘若 X 和 Y 為獨立，則 $\mathrm{Var}(aX\pm bY)=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)$
5. $\mathrm{Cov}(X,\ Y)=\mathrm{Cov}(Y,\ X)$
6. $\mathrm{Cov}(X,\ X)=\mathrm{Var}(X)=\sigma_x^2$

7. 設 a, b 為常數，則 $\mathrm{Cov}(aX,\ bY) = ab\mathrm{Cov}(X,\ Y)$

8. 若 X 和 Y 為獨立，則 $\mathrm{Cov}(X,\ Y) = 0$，其逆不眞

9. 相關係數為一純量，與單位無關

10. 相關係數僅能表示變數之間線性關係的強弱，並不能表示其它函數關係之強弱

例 **3–21** 已知 X 和 Y 的聯合機率密度函數如下表：試求

		Y		
		5	10	15
	10	0.1	0.2	0
X	15	0.1	0.2	0.1
	20	0	0.2	0.1

1. 分別得 X 與 Y 的期望值與變異數
2. 共變數 $\mathrm{Cov}(X,\ Y)$
3. 令 $W = 0.6X + 0.8Y$，求其平均值與變異數

解：

1. $\mu_x = \mathrm{E}X = \sum x f_1(x) = 15$
   $\sigma_x^2 = \mathrm{E}(X^2) - \mu_x^2 = (10^2 \times 0.3 + 15^2 \times 0.4 + 20^2 \times 0.3) - 15^2 = 15$
   $\mu_y = \mathrm{E}Y = \sum y f_2(y) = 10$
   $\sigma_y^2 = \mathrm{E}(Y^2) - \mu_y^2 = (5^2 \times 0.2 + 10^2 \times 0.6 + 15^2 \times 0.2) - 10^2 = 10$

2.

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathrm{E}(XY) - \mu_x\mu_y \\
&= 10 \times 5 \times 0.1 + 10 \times 10 \times 0.2 + \cdots + 20 \times 15 \times 0.1 - 15 \times 10 \\
&= 5
\end{aligned}$$

3.

$$\begin{aligned}
\mathrm{E}(W) &= \mathrm{E}(0.6X + 0.8Y) \\
&= 0.6 \times 15 + 0.8 \times 10
\end{aligned}$$

$$= \quad 17$$

$$\begin{aligned}
\mathrm{Var}(W) &= \mathrm{Var}(0.6X + 0.8Y) \\
&= 0.6^2\mathrm{Var}(X) + 0.8^2\mathrm{Var}(Y) + 2 \times 0.6 \times 0.8 \times \mathrm{Cov}(X,Y) \\
&= 0.36 \times 15 + 0.64 \times 10 + 2 \times 0.6 \times 0.8 \times 5 \\
&= 16.6
\end{aligned}$$

例 **3–22** $f(x,y) = \dfrac{1}{3}, (x,y) = (1,0),\ (0,1),\ (2,1)$，試求

1. ρ_{xy}
2. X 是否與 Y 獨立

解：

1.

$$\mathrm{E}X = 0 \times \frac{1}{3} + 1 \times \frac{1}{3} + 2 \times \frac{1}{3} = 1$$

$$\mathrm{E}Y = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} + 0 \times \frac{1}{3} = \frac{2}{3}$$

$$EXY = 0 \times 1 \times \frac{1}{3} + 1 \times 0 \times \frac{1}{3} + 2 \times 1 \times \frac{1}{3} = \frac{2}{3}$$

所以，相關係數

$$\rho_{xy} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x\sigma_y} = \frac{EXY - EXEY}{\sigma_x\sigma_y} = 0$$

2.

$$f(2,1) = \frac{1}{3} \neq \frac{1}{3} \times \frac{2}{3} = f_1(2) \times f_2(1)$$

所以，隨機變數 X 與 Y 並不獨立。

例 **3–23** 假設 X 與 Y 為兩隨機變數，其中相關係數為 ρ。若 X 與 Y 獨立，則 $\rho = 0$；反之，若 $\rho = 0$，則表示兩隨機變數獨立。

解：由上例可知，錯。

例 **3–24** Let X and Y denote two binary variables. Let $\alpha = \mathrm{P}(Y=1|X=1)/\mathrm{P}(Y=0|X=1)$, and $\beta = \mathrm{P}(Y=1|X=0)/\mathrm{P}(Y=0|X=0)$. Let $\theta = \alpha/\beta$.

1. Show that if X and Y are independent then $\theta = 1$.
2. Show that θ can be expressed as
$$\frac{\mathrm{P}(X=1|Y=1)\,\mathrm{P}(X=1|Y=0)}{\mathrm{P}(X=0|Y=1)\,\mathrm{P}(X=0|Y=0)}.$$
3. Are X and Y independent if $\theta = 1$.

解：

1. 若 X 與 Y 獨立，則
$$\alpha = \frac{\mathrm{P}(Y=1|X=1)}{\mathrm{P}(Y=0|X=1)} = \frac{\mathrm{P}(Y=1 \ \text{ and } X=1)/\mathrm{P}(X=1)}{\mathrm{P}(Y=0 \text{ and } X=1)/\mathrm{P}(X=1)} = \frac{\mathrm{P}(Y=1)}{\mathrm{P}(Y=0)}$$
及
$$\beta = \frac{\mathrm{P}(Y=1|X=0)}{\mathrm{P}(Y=0|X=0)} = \frac{\mathrm{P}(Y=1 \ \text{ and } X=0)/\mathrm{P}(X=0)}{\mathrm{P}(Y=0 \ \text{ and } X=0)/\mathrm{P}(X=0)} = \frac{\mathrm{P}(Y=1)}{\mathrm{P}(Y=0)}$$
所以
$$\theta = \frac{\alpha}{\beta} = \frac{\mathrm{P}(Y=1)/\mathrm{P}(Y=0)}{\mathrm{P}(Y=1)/\mathrm{P}(Y=0)} = 1$$

2.
$$\begin{aligned}
\theta &= \frac{\alpha}{\beta} \\
&= \frac{\mathrm{P}(Y=1 \ \text{ and } X=1)/\mathrm{P}(Y=0 \ \text{ and } X=1)}{\mathrm{P}(Y=1 \ \text{ and } X=0)/\mathrm{P}(Y=0 \ \text{ and } X=0)} \\
&= \frac{\mathrm{P}(Y=1 \ \text{ and } X=1)\,\mathrm{P}(Y=0 \ \text{ and } X=0)}{\mathrm{P}(Y=1 \ \text{ and } X=0)\,\mathrm{P}(Y=0 \ \text{ and } X=1)} \\
&= \frac{\mathrm{P}(X=1|Y=1)\,\mathrm{P}(Y=1)\,\mathrm{P}(X=0|Y=0)\,\mathrm{P}(Y=0)}{\mathrm{P}(X=0|Y=1)\,\mathrm{P}(Y=1)\,\mathrm{P}(X=1|Y=0)\,\mathrm{P}(Y=0)} \\
&= \frac{\mathrm{P}(X=1|Y=1)\,\mathrm{P}(X=0|Y=0)}{\mathrm{P}(X=0|Y=1)\,\mathrm{P}(X=1|Y=0)}
\end{aligned}$$

3. 假設 $\mathrm{P}(X=i \ \text{ and } Y=j) = P_{ij}$, $i=0,1$; $j=0,1$，則
$$\alpha = \frac{\mathrm{P}(Y=1|X=1)}{\mathrm{P}(Y=0|X=1)} = \frac{\mathrm{P}(X=1 \ \text{ and } Y=1)}{\mathrm{P}(X=1 \ \text{ and } Y=0)} = \frac{P_{11}}{P_{10}}$$

及

$$\beta = \frac{\mathrm{P}\left(Y=1|X=0\right)}{\mathrm{P}\left(Y=0|X=0\right)} = \frac{\mathrm{P}\left(X=0 \ \text{ and } Y=1\right)}{\mathrm{P}\left(X=0 \ \text{ and } Y=0\right)} = \frac{P_{01}}{P_{00}}$$

所以 $P_{11} = \alpha P_{10}$ 及 $P_{01} = \beta P_{00}$. 又 $P_{11} + P_{10} + P_{00} + P_{01} = 1$，故得 $(\alpha+1) P_{10} + (\beta+1) P_{00} = 1$。當 $\theta = 1$, $\alpha = \beta$ 且 $(\alpha+1)(P_{10} + P_{00}) = 1$。由於

$$\begin{aligned}
\mathrm{P}\left(X=1\right)\mathrm{P}\left(Y=1\right) &= (\alpha+1) P_{10} \times (\alpha P_{10} + \beta P_{00}) \\
&= \alpha P_{10} (\alpha+1)(P_{10} + P_{00}) \\
&= \alpha P_{10} \times 1 \\
&= \mathrm{P}\left(X=1 \ \text{ and } Y=1\right),
\end{aligned}$$

故可推得 X 與 Y 獨立

例 **3–25** 在資料分布的型態未知的情況下，若母體平均數及母體變異數皆存在，證明下式成立

1. 當 X 為大於零的隨機變數時，$\mathrm{P}\left(x > k\mu\right) \leq 1/k$
2. 當 X 為任意隨機變數時，$\mathrm{P}\left(\mid x-\mu \mid > k\sigma\right) \leq 1/k^2$

證明.

1.

$$\begin{aligned}
\mu &= \int_{x\in R} x f(x)\, dx \\
&= \int_{x\leq k\mu} x f(x)\, dx + \int_{x>k\mu} x f(x)\, dx \\
&\geq \int_{x>k\mu} x f(x)\, dx \\
&\geq \int_{x>k\mu} k\mu f(x)\, dx \\
&= k\mu \int_{x>k\mu} f(x)\, dx \\
&= k\mu \mathrm{P}\left(x > k\mu\right)
\end{aligned}$$

所以

$$\mathrm{P}\left(x > k\mu\right) \leq \frac{1}{k}$$

此不等式稱為馬可夫不等式(Markov’s Inequality)

2.

$$
\begin{aligned}
\sigma^2 &= \int_{x\in R} (x-\mu)^2 f(x)\,dx \\
&= \int_{|x-\mu|\le k\sigma} (x-\mu)^2 f(x)\,dx + \int_{|x-\mu|>k\sigma} (x-\mu)^2 f(x)\,dx \\
&\ge \int_{|x-\mu|>k\sigma} (x-\mu)^2 f(x)\,dx \\
&\ge \int_{|x-\mu|>k\sigma} k^2\sigma^2 f(x)\,dx \\
&= k^2\sigma^2 \int_{|x-\mu|>k\sigma} f(x)\,dx \\
&= k^2\sigma^2 \mathrm{P}\left(|x-\mu|>k\sigma\right)
\end{aligned}
$$

故得

$$
\mathrm{P}\left(|x-\mu|>k\sigma\right) \le \frac{1}{k^2}
$$

例 **3–26** If X is a random variable with mean 0 and finite σ^2, prove that

1. for any random variables Y, $[EXY]^2 \le \mathrm{E}X^2\mathrm{E}Y^2$
2. for any $a>0$, $\mathrm{P}(X>a) \le \sigma^2/[\sigma^2+a^2]$

證明.

1.

$$
\mathrm{E}X\left(Y-\mu_y\right) = EXY - \mu_y\mathrm{E}X = EXY
$$

又 $\mathrm{E}\left(Y-\mu_y\right)^2 \le \mathrm{E}Y^2$，故得

$$
1 \ge \rho^2 = \frac{(EXY)^2}{\mathrm{E}X^2\mathrm{E}\left(Y-\mu_y\right)^2} \ge \frac{(EXY)^2}{\mathrm{E}X^2\mathrm{E}Y^2}
$$

所以

$$
(EXY)^2 \le \mathrm{E}X^2\mathrm{E}Y^2
$$

2. 對所有的 $b>0$，由馬可夫不等式可得

$$
\mathrm{P}(X>a) = \mathrm{P}(X+b>a+b)
$$

$$
\begin{aligned}
&\leq \mathrm{P}\left[(X+b)^2>(a+b)^2\right]\\
&= \mathrm{P}\left[(X+b)^2>\frac{(a+b)^2}{\sigma^2+b^2}\left(\sigma^2+b^2\right)\right]\\
&\leq \frac{\sigma^2+b^2}{(a+b)^2}
\end{aligned}
$$

令 $b=\sigma^2/a$，可推得

$$
\mathrm{P}(X>a)\leq\frac{\sigma^2+\sigma^4/a^2}{(a+\sigma^2/a)^2}=\frac{\sigma^2}{a^2+\sigma^2}
$$

例 **3–27** Suppose a random chosen individual's verbal score X and quantitative score Y on a nationally administered attitude examination have joint pdf

$$
f(x,y)=\begin{cases}\frac{2}{5}(2x+3y), & 0\leq x\leq 1, 0\leq y\leq 1,\\ 0, & \text{otherwise.}\end{cases}
$$

1. Find the marginal distribution of X alone.
2. Find $\mathrm{E}X$ and $\mathrm{Var}(X)$.
3. Using Chebyshev's theorem, find the interval that contains at least $5/9$ of the values of X.
4. For the interval of X obtained in (3), find the true probability.
5. You are asked to provide a prediction T of the individual's total score $X+Y$. the error of prediction is the mean squared error $\mathrm{E}\left[(X+Y-T)^2\right]$. What value of T minimizes the error of prediction？

解：

1.

$$
f_1(x)=\int_0^1\frac{2}{5}(2x+3y)\,dy=\frac{4}{5}x+\frac{3}{5}
$$

2.

$$
\mathrm{E}X=\int_0^1\left(\frac{4}{5}x^2+\frac{3}{5}x\right)dx=0.5667, \mathrm{E}X^2=\int_0^1\left(\frac{4}{5}x^3+\frac{3}{5}x^2\right)dx=0.4
$$

所以，$\mathrm{Var}(X)=\mathrm{E}X^2-[\mathrm{E}X]^2=0.4-0.5667^2=0.0789$

3.

$$
\mathrm{P}(|X-\mu|\leq k\sigma)\geq 1-\frac{1}{k^2}=\frac{5}{9},\ 得k=\frac{3}{2}.
$$

故可推得

$$|X-0.5667| \le \frac{3}{2} \times \sqrt{0.0789},\ 即 0.1454 \le X \le 0.9980$$

4.

$$\mathrm{P}(0.1454 \le X \le 0.9980) = \int_{0.1454}^{0.9980} \frac{4}{5}x + \frac{3}{5}dx = 0.9015$$

5.

$$\begin{aligned} \mathrm{E}\left[(X+Y-T)^2\right] &= \mathrm{E}\left[X^2+Y^2+T^2+2XY-2XT-2YT\right] \\ &= \mathrm{E}X^2+\mathrm{E}Y^2+T^2+2EXY-2TEX-2TEY \end{aligned}$$

上式對 T 做微分得

$$\frac{dE\left[(X+Y-T)^2\right]}{dT} = 2T-2\mathrm{E}X-2\mathrm{E}Y=0,$$

故可得 $T=\mathrm{E}X+\mathrm{E}Y$。又

$$f_2(y) = \int_0^1 \frac{2}{5}(2x+3y)\,dx = \frac{6}{5}y+\frac{2}{5},\ \mathrm{E}Y = \int_0^1 \frac{6}{5}y^2+\frac{2}{5}ydy = 0.6.$$

所以 $T=0.5667+0.6=1.1667$

3.9 動差生成函數

在統計學中, 平均數及變異數是非常重要的母體特徵數。然而對於一些分配而言，例如常態分配，要藉由計算 $\mathrm{E}X$ 和 $\mathrm{E}X^2$ 求解變異數並不是一件容易的事，因此本節將介紹動差生成函數以幫助我們快速找出母體特徵數。

定義

若 X 為樣本空間分布於 S 的一隨機變數，其機率密度函數為 $f(x)$，若存在一個正數 h 使得 $\mathrm{E}(e^{tx})$ 存在且 $-h<t<h$，則稱此一函數

$$M(t) = \mathrm{E}\left(e^{tx}\right) = \begin{cases} \sum\limits_{x\in S} e^{tx}f(x), & \text{離散型} \\ \int_{x\in S} e^{tx}f(x)\,dx, & \text{連續型} \end{cases}$$

為隨機變數 X 的動差生成函數 (moment-generating function, 簡稱 m.g.f.)。

性質

1. $M^{(r)}(0) = \mathrm{E}(X^r)$
2. $\mathrm{E}X = M'(0)$ 及 $\mathrm{Var}(X) = M''(0) - [M'(0)]^2$
3. $M(t) = \displaystyle\sum_{r=0}^{\infty} \frac{M^{(r)}(0)}{r!} t^r$

證明.

假設 X 為離散型分配

1. 由定義可知

$$M'(t) = \frac{d}{dt} \sum_{x \in S} e^{tx} f(x) = \sum_{X \in S} x e^{tx} f(x)$$

$$M''(t) = \frac{d}{dt} M'(t) = \sum_{X \in S} x^2 e^{tx} f(x)$$

$$\vdots$$

$$M^{(r)}(t) = \sum_{X \in S} x^r e^{tx} f(x)$$

令 $t = 0$，則

$$M'(0) = \sum_{X \in S} x f(x),\ M''(0) = \sum_{X \in S} x^2 f(x), \ldots, M^{(r)}(0) = \sum_{X \in S} x^r f(x)$$

2. 由上述性質可知

$$M'(0) = \sum_{X \in S} x f(x) = \mathrm{E}X \text{ 及 } M''(0) = \sum_{X \in S} x^2 f(x) = \mathrm{E}X^2$$

所以

$$\mathrm{Var}(X) = \mathrm{E}X^2 - [\mathrm{E}X]^2 = M''(0) - [M'(0)]^2$$

3. 由 Maclaurin Series 可知

$$e^{tx} = \sum_{r=0}^{\infty} \frac{(tx)^r}{r!} = \sum_{r=0}^{\infty} \frac{x^r}{r!} t^r = 1 + tx + \frac{x^2}{2!} t^2 + \cdots$$

又 $M^{(r)}(0) = \sum_{X \in S} x^r f(x)$，所以

$$M(t) = \sum_{r=0}^{\infty} \frac{M^{(r)}(0)}{r!} t^r$$

例 **3–28** 假設隨機變數 X 之 m.g.f. 為

$$\frac{\frac{e^t}{2}}{1 - \frac{e^t}{2}},\ t < \ln 2$$

試求 X 之 p.d.f..

解：

因為

$$\begin{aligned} \frac{\frac{e^t}{2}}{1 - \frac{e^t}{2}} &= \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + \cdots \\ &= \sum_{r=1}^{\infty} \left(\frac{1}{2}\right)^r e^{rt} \end{aligned}$$

由定義可知

$$f(x) = \left(\frac{1}{2}\right)^x,\ x = 1, 2, \ldots$$

例 **3–29** 假設隨機變數 X 的 m.g.f. 為

$$M(t) = \frac{44}{120} + \frac{45}{120} e^t + \frac{20}{120} e^{2t} + \frac{10}{120} e^{3t} + \frac{1}{120} e^{5t}$$

求 EX 及 Var(X)。

解：

由定義可知 X 的機率分配如下

x	0	1	2	3	5
$f(x)$	$\frac{44}{120}$	$\frac{45}{120}$	$\frac{20}{120}$	$\frac{10}{120}$	$\frac{1}{120}$

所以

$$\begin{aligned} \mathrm{E}X &= 0 \times \frac{44}{120} + 1 \times \frac{45}{120} + 2 \times \frac{20}{120} + 3 \times \frac{10}{120} + 5 \times \frac{1}{120} \\ &= 1 \end{aligned}$$

及

$$\begin{aligned} \mathrm{E}X^2 &= 0^2 \times \frac{44}{120} + 1^2 \times \frac{45}{120} + 2^2 \times \frac{20}{120} + 3^2 \times \frac{10}{120} + 5^2 \times \frac{1}{120} \\ &= 2 \end{aligned}$$

故得

$$\mathrm{Var}(X) = 2 - 1 = 1$$

例 **3–30** Suppose X and Y are independent random variables having the uniform distribution on $1, 2, \ldots, N$. Compute the probability mass function of $Z = X + Y$.

解：

由於 $\mathrm{P}(X = x) = 1/N,\ x = 1, 2, \ldots, N$，所以 X 及 Y 的動差生成函數分別為

$$M_X(t) = \mathrm{E}\left(e^{tX}\right) = \sum_{x=1}^{N} \frac{1}{N} e^{tx},\ \text{及} M_Y(t) = \mathrm{E}\left(e^{tY}\right) = \sum_{y=1}^{N} \frac{1}{N} e^{ty}$$

因為 X 與 Y 獨立，所以

$$\begin{aligned} M_{X+Y}(t) &= \mathrm{E}\left(e^{t(X+Y)}\right) = \mathrm{E}\left(e^{tX}\right)\mathrm{E}\left(e^{tY}\right) \\ &= \sum_{x=1}^{N} \frac{1}{N} e^{tx} \sum_{y=1}^{N} \frac{1}{N} e^{ty} \\ &= \frac{1}{N^2}\left[\sum_{i=1}^{N} e^{tx} \sum_{y=1}^{N} e^{ty}\right] \\ &= \frac{1}{N^2}\left(e^t + e^{2t} + e^{3t} + \ldots + e^{Nt}\right)\left(e^t + e^{2t} + e^{3t} + \ldots + e^{Nt}\right) \end{aligned}$$

$$= \frac{1}{N^2}\left(e^t + e^{2t} + e^{3t} + \ldots + e^{Nt}\right)^2$$

故可推得

$$\mathrm{P}\left(Z = z\right) = \begin{cases} (z-1)/N^2, & 2 \leq z \leq N+1 \\ (2N-z+1)/N^2, & N+2 \leq z \leq 2N \end{cases}$$

第 4 章　常用之機率分配

4.1　離散型機率分配

1. **離散型均勻分配**（Uniform Distribution）：若離散隨機變數之分配具有下列機率密度函數：

$$f(x) = \frac{1}{N},\ x = 1,\ 2,\ \ldots,\ N,$$

則稱其為均勻分配。

性質

(a) $\mathrm{E}X = \frac{N+1}{2}$

(b) $\mathrm{Var}(X) = \frac{N^2-1}{12}$

證明.

(a)

$$\mathrm{E}X = \sum_{x=1}^{N} \frac{x}{N} = \frac{1}{N}\frac{N(N+1)}{2} = \frac{N+1}{2}$$

(b)

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X-\mu)^2 \\
&= \mathrm{E}X^2 - \mu^2 \\
&= \sum_{x=1}^{N} x^2 \frac{1}{N} - \left(\frac{N+1}{2}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{N}\frac{N(N+1)(2N+1)}{6} - \frac{N^2+2N+1}{4} \\
&= \frac{N^2-1}{12}
\end{aligned}
$$

2. **伯努力分配** (Bernoulli Distribution)：投擲銅板一枚，出現正面記為 1，反面記為 0，則稱此一實驗為 Bernoulli trial。若 X 表銅板出現正面次數，以 $X \sim Beronulli(p)$ 表示，其 p.d.f. 為

$$f(x) = p^x(1-p)^{1-x}\ ,\ \ x = 0,\ 1$$

性質

(a) $\mathrm{E}X = p$

(b) $\mathrm{Var}(X) = \mathrm{P}(1-p)$

(c) $M(t) = 1 - p + pe^t$

證明.

(a)

$$
\begin{aligned}
\mathrm{E}X &= 0 \times p^0(1-p)^{1-0} + 1 \times p^1(1-p)^{1-1} \\
&= p
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}X^2 &= 0^2 \times p^0(1-p)^{1-0} + 1^2 \times p^1(1-p)^{1-1} \\
&= p
\end{aligned}
$$

所以

$$\mathrm{Var}(X) = \mathrm{E}X^2 - \mu^2 = p - p^2 = \mathrm{P}(1-p)$$

(c)

$$M(t) = \sum_{x=0}^{1} p^x (1-p)^{1-x} e^{tx}$$

$$= \quad pe^t + 1 - p$$

3. **二項分配**（Binomial Distribution）：在有限母體中可分為成功和失敗情形下，樣本抽出後放回，考慮成功次數的隨機實驗，稱為二項分配，通常以 $X \sim B(n,\ p)$ 表示，其 p.d.f. 為

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}\ ,\ \ x = 0,\ 1,\ 2,\ \ldots, n$$

其中

(a) n 為重複獨立實驗次數

(b) p 為每次實驗之成功機率

(c) $\binom{n}{x}$ 為組合公式，其中 $\binom{n}{x} = \frac{n!}{x!\ (n-x)!}$

(d) $(a+b)^n = \sum\limits_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$

(e) 當 $n = 1$ 時稱為 Bernoulli 分配，即 $X \sim B(1,\ p)$, $f(x) = p^x (1-p)^{1-x}\ ,\ \ x = 0,\ 1$

性質

(a) $\mathrm{E}X = np$

(b) $\mathrm{Var}(X) = n\mathrm{P}(1-q)$

(c) $\mathrm{E}X\,(X-1)\,(X-1)\cdots(X-r+1) = n\,(n-1)\,(n-2)\cdots(n-r+1)\,p^r$

(d) $M(t) = \left(1 - p + pe^t\right)^n$

(e) 若 $X_i \overset{iid}{\sim} B\,(1, p),\ i = 1, 2, \ldots, n$，則 $Y = \sum_{i=1}^{n} X_i \sim B(n,\ p)$

(f) 若 $X \sim B(m,\ p),\ Y \sim B(n,\ p)$，則 $X + Y \sim B(m+n,\ p)$

證明.

(a)

$$\begin{aligned} \mathrm{E}X &= \sum_{x=0}^{n} x \binom{n}{x} p^x\,(1-p)^{n-x} \\ &= \sum_{x=0}^{n} \frac{n!}{x!\,(n-x)!} x p^x\,(1-p)^{n-x} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} (1-p)^{n-x} \\
&= n\mathrm{P}\,(p+1-p)^{n-1} \\
&= np
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}X\,(X-1) &= \sum_{x=0}^{n} x\,(x-1) \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!\,(n-x)!} p^x (1-p)^{n-x} \\
&= n\,(n-1)\,p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,(n-x)!} p^{x-2} (1-p)^{n-x} \\
&= n\,(n-1)\,p^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 \\
&= n^2p^2 - np^2 + np - n^2p^2 \\
&= n\mathrm{P}(1-p)
\end{aligned}
$$

(c)

$$
\begin{aligned}
&\mathrm{E}X\,(X-1)\,(X-1)\cdots(X-r+1) \\
= &\sum_{x=0}^{n} x\,(x-1)\,(x-2)\cdots(x-r+1) \binom{n}{x} p^x (1-p)^{n-x} \\
= &\sum_{x=0}^{n} x\,(x-1)\,(x-2)\cdots(x-r+1) \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
= &\sum_{x=0}^{n} \frac{n!}{(x-r)!\,(n-x)!} p^x (1-p)^{n-x} \\
= &\, n\,(n-1)\,(n-2)\cdots(n-r+1)\,p^r \sum_{x=r}^{n} \frac{(n-r)!}{(x-r)!\,(n-x)!} p^{x-r} (1-p)^{n-x} \\
= &\, n\,(n-1)\,(n-2)\cdots(n-r+1)\,p^r
\end{aligned}
$$

(d)

$$\begin{aligned} M(t) &= \sum_{x=0}^{n} p^x (1-p)^{n-x} e^{tx} \\ &= \sum_{x=0}^{n} \left(pe^t\right)^x (1-p)^{n-x} \\ &= \left(1-p+pe^t\right)^n \end{aligned}$$

(e) 由 **3.7.1 性質 c** (第37頁) 可知，

$$\begin{aligned} M_Y(t) &= \mathrm{E}\left(e^{ty}\right) \\ &= \mathrm{E}\left(e^{t\sum_{i=1}^{n} x_i}\right) \\ &= \mathrm{E}\left(e^{tx_1} e^{tx_2} e^{tx_3} \cdots e^{tx_n}\right) \\ &= \mathrm{E}\left(e^{tx_1}\right) \mathrm{E}\left(e^{tx_2}\right) \mathrm{E}\left(e^{tx_3}\right) \cdots \mathrm{E}\left(e^{tx_n}\right) \\ &= \left[1-p+pe^t\right]^n \end{aligned}$$

所以，由動差生成函數之定義可知 $Y \sim B(n,p)$

(f)

$$\begin{aligned} M_{X+Y}(t) &= \mathrm{E}\left(e^{t(x+y)}\right) \\ &= \mathrm{E}\left(e^{tx} e^{ty}\right) \\ &= \mathrm{E}\left(e^{tx}\right) \mathrm{E}\left(e^{ty}\right) \\ &= \left[1-p+pe^t\right]^n \left[1-p+pe^t\right]^m \\ &= \left[1-p+pe^t\right]^{n+m} \end{aligned}$$

所以，由動差生成函數之定義可知 $X+Y \sim B(n+m,p)$

例 **4–1** For positive integers $n > r$, show

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

證明.

$$\begin{aligned}\binom{n-1}{r-1}+\binom{n-1}{r} &= \frac{(n-1)!}{(r-1)!\,(n-r)!}+\frac{(n-1)!}{r!\,(n-1-r)!}\\ &= \frac{(n-1)!\times r+(n-1)!\times(n-r)}{r!\,(n-r)!}\\ &= \frac{(n-1)!\,[r+(n-r)]}{r!\,(n-r)!}=\frac{(n-1)!\times n}{r!\,(n-r)!}\\ &= \frac{n!}{r!\,(n-r)!}=\binom{n}{r}\end{aligned}$$

例 **4–2** 假設台北市約有 40% 的人喜歡棒球運動。現於市區中隨機訪問一人，試問：

(a) 此人喜歡棒球運動的期望值為何？變異數為何？

(b) 若隨機訪問 5 人，試問我們期望 5 人中喜歡棒球運動的期望值與變異數為何？有 2 個人喜歡棒球運動的機率為何？至少有 3 個人喜歡棒球運動的機率為何？

解：

(a) 令 X 表示喜歡棒球運動的人次，則 $X \sim B(1,\ 0.4)$
所以 $\mathrm{E}X = p = 0.4$, $\mathrm{Var}(X) = pq = 0.4 \times 0.6 = 0.24$

(b) 令 X 表示 5 人中喜歡棒球運動的人數，則 $X \sim B(5,\ 0.4)$

  i. $\mathrm{E}(X) = np = 5 \times 0.4 = 2$, $\mathrm{Var}(X) = npq = 5 \times 0.4 \times 0.6 = 1.2$

  ii. $\mathrm{P}(X = 2) = \binom{5}{2}(0.4)^2(0.6)^3 = 10 \times 0.03456 = 0.3456$

  iii. $\mathrm{P}(X \geq 3) = 1 - \mathrm{P}(X \leq 2) = 1 - \sum_{x=0}^{2}\binom{5}{x}(0.4)^x(0.6)^{5-x} = 1 - 0.68256 = 0.31744$

例 **4–3** 假設台北市吸香煙人口佔總人口的 30%，今隨機抽取 20 位市民作訪問：

(a) 20 人中只有 4 人吸香煙之機率？

(b) 20 人中有 10 人或 10 人以上吸香煙之機率？

(c) 20 人中僅有 2 人或 2 人以下吸香煙之機率？

解：

令 X 表示 20 人中吸香煙的人數，則 $X \sim B(20,\ 0.3)$

(a) $\mathrm{P}(X = 4) = \binom{20}{4}(0.3)^4(0.7)^{16} = 0.131$

(b) $\mathrm{P}(X \geq 10) = 1 - \mathrm{P}(X \leq 9) = 1 - 0.952 = 0.048$

(c) $\mathrm{P}(X \leq 2) = \sum_{x=0}^{2} \binom{20}{x} (0.3)^x (0.7)^{20-x} = 0.035$

例 4–4 To become a PhD candidate in th IM department of NTUST, a student has to pass an entrance exam and a qualify exam. Suppose that there are N students registered the entrance exam and each can pass the exam with probability p independently. Only those who past the entrance exam can take the qualify exam and each can pass the qualify exam with probability q independently. Let M be the number of students past the entrance exam and X be the number of PhD candidates among th N students.

(a) What is the joint pmf of X and M?

(b) What is the pmf of X?

(c) what is the expected value of X?

解：

(a)

$$\mathrm{P}\left(X = x|M = m\right) = \binom{m}{x} q^x (1-q)^{m-x} \text{ 及} \mathrm{P}\left(M = m\right) = \binom{N}{m} p^m (1-p)^{N-m}$$

所以

$$\mathrm{P}\left(X = x, M = m\right) = \binom{m}{x} q^x (1-q)^{m-x} \binom{N}{m} p^m (1-p)^{N-m}$$

(b)

$$\begin{aligned}
\mathrm{P}\left(X = x\right) &= \sum_{m=x}^{N} \binom{m}{x} q^x (1-q)^{m-x} \binom{N}{m} p^m (1-p)^{N-m} \\
&= \sum_{m=x}^{N} \frac{m!}{x!\,(m-x)!} \frac{N!}{m!\,(N-m)!} p^m (1-p)^{N-m} q^x (1-q)^{m-x} \\
&= \frac{1}{x!} \left[\sum_{m=x}^{N} \frac{N!}{(m-x)!\,(N-m)!} p^m (1-p)^{N-m} q^x (1-q)^{m-x} \right] \\
&= \frac{1}{x!} (pq)^x \left[\sum_{m=x}^{N} \frac{(N-x)!\,(N-x+1)\,(N-x+1)\cdots N}{(m-x)!\,(N-m)!} (1-p)^{N-m} (p-pq)^{m-x} \right] \\
&= \frac{(N-x+1)\,(N-x+2)\cdots N}{x!} (pq)^x \left[\sum_{y=0}^{N-x} \frac{(N-x)!}{y!\,(N-x-y)!} (1-p)^{N-x-y} (p-pq)^y \right] \\
&= \frac{(N-x+1)\,(N-x+2)\cdots N}{x!} (pq)^x \,(1-pq)^{N-x} \\
&= \binom{N}{x} (pq)^x \,(1-pq)^{N-x}
\end{aligned}$$

所以可求得 $X \sim B(N, pq)$

(c)

$$\mathrm{E}X = EE\left[X|M\right] = \mathrm{E}\left[Mq\right] = Npq$$

例 **4–5** 假設甲, 乙兩部機器, 每部機器每次可生產一個產品，而且甲的生產速率是乙的 3 倍; 甲生產的產品不良率為 1 %, 乙則為 3%。請問

(a) 產品的平均不良率為何？

(b) 任意選取兩的產品，這兩個產品都是良品的機率是多少？

(c) 隨機選取一個不良品，請問它是由機器甲所生產的機會是多少？

解：

(a)

$$\begin{aligned}
\mathrm{P}\left(\text{不良品}\right) &= \mathrm{P}\left(\text{甲} \cap \text{不良品}\right) + \mathrm{P}\left(\text{乙} \cap \text{不良品}\right) \\
&= \mathrm{P}\left(\text{甲}\right)\mathrm{P}\left(\text{不良品} \mid \text{甲}\right) + \mathrm{P}\left(\mathrm{B}\right)\mathrm{P}\left(\text{不良品} \mid \mathrm{A}\right) \\
&= \frac{3}{4} \times 0.01 + \frac{1}{4} \times 0.03 \\
&= 0.015
\end{aligned}$$

(b)

$$\begin{aligned}
\mathrm{P}\left(\text{良品}\right) &= \binom{2}{2} \times 0.015^0 \times 0.985^2 \\
&= 0.97023
\end{aligned}$$

(c)

$$\begin{aligned}
\mathrm{P}\left(\text{甲} \mid \text{不良品}\right) &= \frac{0.75 \times 0.01}{0.75 \times 0.01 + 0.25 \times 0.03} \\
&= 0.5
\end{aligned}$$

4. **三項分配** (Trinomial Distribution)：若有限母體中可分為 X, Y 及 Z 三種情形，其中 X, Y, Z 出現之機率分別為 p_1, p_2 及 p_3 且 $p_1 + p_2 + p_3 = 1$，樣本抽出後放回，考慮 X 與 Y 出現次數之隨機試驗稱為三項分配，其 p.d.f. 為

$$f\left(x, y\right) = \binom{n}{x, y} p_1^x p_2^y \left(1 - p_1 - p_2\right)^{n-x-y}, x = 0, 1, 2, \ldots, n; y = 0, 1, 2, \ldots, n \text{ 且} x + y \leq n$$

則

(a) $X \sim B(n, p_1)$

(b) 在給定 $X = x$ 條件下，$Y \sim B(n-x, p_2/(1-p_1))$

(c) $\rho_{xy} = -\sqrt{p_1 p_2 / (1-p_1)(1-p_2)}$

證明.

(a)

$$
\begin{aligned}
& \mathrm{P}(X=x) \\
= & \sum_{y=0}^{n-x} \binom{n}{x, y} p_1^x p_2^y (1-p_1-p_2)^{n-x-y} \\
= & \sum_{y=0}^{n-x} \frac{n!}{x! y! (n-x-y)!} p_1^x p_2^y (1-p_1-p_2)^{n-x-y} \\
= & \frac{n \times (n-1) \times (n-2) \times \cdots \times (n-x+1)}{x!} p_1^x \sum_{y=0}^{n-x} \frac{(n-x)!}{y! (n-x-y)!} p_2^y (1-p_1-p_2)^{n-x-y} \\
= & \frac{n \times (n-1) \times (n-2) \times \cdots \times (n-x+1)}{x!} p_1^x \sum_{y=0}^{n-x} \binom{n-x}{y} p_2^y (1-p_1-p_2)^{n-x-y} \\
= & \frac{n \times (n-1) \times (n-2) \times \cdots \times (n-x+1)}{x!} p_1^x [(1-p_1-p_2)+p_2]^{n-x} \\
= & \frac{n!}{x! (n-x)!} p_1^x (1-p_1)^{n-x}
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{P}(Y=y|X=x) &= \frac{[n!/x! y! (n-x-y)!] p_1^x p_2^y (1-p_1-p_2)^{n-x-y}}{[n!/x! (n-x)!] p_1^x (1-p_1)^{n-x}} \\
&= \frac{(n-x)!}{y! (n-x-y)!} \frac{p_2^y (1-p_1-p_2)^{n-x-y}}{(1-p_1)^{n-x}} \\
&= \frac{(n-x)!}{y! (n-x-y)!} \frac{p_2^y}{(1-p_1)^y} \frac{(1-p_1-p_2)^{n-x-y}}{(1-p_1)^{n-x-y}} \\
&= \frac{(n-x)!}{y! (n-x-y)!} \left(\frac{p_2}{1-p_1}\right)^y \left(\frac{1-p_1-p_2}{1-p_1}\right)^{n-x-y}
\end{aligned}
$$

(c) 因為

$$\begin{aligned}
EXY &= \sum_{x=0}^{n}\sum_{y=0}^{n-x} xy \times \frac{n!}{x!y!\,(n-x-y)!} p_1^x p_2^y \,(1-p_1-p_2)^{n-x-y} \\
&= \sum_{x=1}^{n-2}\sum_{y=1}^{n-x+1} \frac{(n-2)! \times (n-1) \times n}{(x-1)!\,(y-1)!\,(n-x-y)!} p_1 p_2 p_1^{x-1} p_2^{y-1} \,(1-p_1-p_2)^{n-x-y} \\
&= (n-1)\, np_1p_2 \sum_{x=1}^{n-2}\sum_{y=1}^{n-x+1} \binom{n-2}{x-1, y-1} p_1^{x-1} p_2^{y-1} \,(1-p_1-p_2)^{n-x-y} \\
&= (n-1) \times np_1p_2
\end{aligned}$$

所以

$$\begin{aligned}
\mathrm{Cov}\,(X,Y) &= (n-1) \times np_1p_2 - np_1 \times np_2 \\
&= n^2 p_1p_2 - np_1p_2 - n^2 p_1p_2 \\
&= -np_1p_2
\end{aligned}$$

由相關係數定義可知

$$\begin{aligned}
\rho_{xy} &= \frac{-np_1p_2}{\sqrt{np_1\,(1-p_1)}\sqrt{np_2\,(1-p_2)}} \\
&= -\sqrt{\frac{p_1p_2}{(1-p_1)\,(1-p_2)}}
\end{aligned}$$

5. **超幾何分配**(Hyper-Geometric Distribution):在有限母體中可分為成功和失敗情形下,樣本抽出後不放回,考慮成功次數的隨機實驗,稱為**超幾何實驗**,其 p.d.f. 為

$$f\,(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}},\ x = 0, 1, 2, \ldots, min(n, K), n-x \le N-K$$

常記為 $X \sim h(N,\ K,\ n)$,其中

N 為母體的大小,n 為抽出之樣本大小,K 為成功這類之大小,$N-K$ 為失敗這類之大小。

定理 4.1.

$$\binom{N}{n} = \sum_{x=0}^{n} \binom{N_1}{x}\binom{N_2}{n-x},$$

其中，$N = N_1 + N_2$ 及 $n \le min\{N_1, N_2\}$

證明.

由二項式定理可知，

$$\begin{aligned}
(1+y)^N &= (1+y)^{N_1} \times (1+y)^{N_2} \\
&= \left[\binom{N_1}{0}y^0 + \binom{N_1}{1}y^1 + \binom{N_1}{2}y^2 + \cdots + \binom{N_1}{n}y^n\right] \\
&\quad \times \left[\binom{N_2}{0}y^0 + \binom{N_2}{1}y^1 + \binom{N_2}{2}y^2 + \cdots + \binom{N_2}{n}y^n\right]
\end{aligned}$$

又上式左右兩式之 y^n 之係數必須相等，所以

$$\begin{aligned}
\binom{N}{n} &= \binom{N_1}{0}\binom{N_2}{n} + \binom{N_1}{1}\binom{N_2}{n-1} + \binom{N_1}{2}\binom{N_2}{n-2} + \cdots + \binom{N_1}{n}\binom{N_2}{0} \\
&= \sum_{x=0}^{n} \binom{N_1}{x}\binom{N_2}{n-x}
\end{aligned}$$

及

$$\sum_{x=0}^{n} \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} = 1$$

特性

(a) $\mathrm{E}X = n \times \dfrac{K}{N}$

(b) $\mathrm{Var}(X) = \dfrac{N-n}{N-1} \times n \times \dfrac{K}{N} \times \dfrac{N-K}{N}$，其中 $\frac{N-n}{N-1}$ 稱為有限母體校正因子

(c) 當 $\frac{n}{N} \le 0.05$ 時，可以二項分配做為超幾何分配之近似值

證明.

(a)

$$
\begin{aligned}
\mathrm{E}X &= \sum_{x=0}^{n} x\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=0}^{n} x\frac{\frac{K!}{x!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{N!}{n!(N-n)!}} \\
&= \sum_{x=1}^{n} \frac{\frac{K!}{(x-1)!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{N!}{n!(N-n)!}} \\
&= K\frac{n}{N}\sum_{x=1}^{n} \frac{\frac{(K-1)!}{(x-1)!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{(N-1)!}{(n-1)!(N-n)!}} \\
&= n\frac{K}{N}
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}X\left(X-1\right) &= \sum_{x=0}^{n} x\left(x-1\right)\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \\
&= \sum_{x=0}^{n} x\left(x-1\right)\frac{\frac{K!}{x!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{N!}{n!(N-n)!}} \\
&= \sum_{x=0}^{n} \frac{K\left(K-1\right)\frac{(K-2)!}{(x-2)!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{\mathcal{N}(N-1)}{n(n-1)}\frac{(N-2)!}{(n-2)!(N-n)!}} \\
&= \frac{nK\left(K-1\right)\left(n-1\right)}{\mathcal{N}\left(N-1\right)}\sum_{x=2}^{n} \frac{\frac{(K-2)!}{(x-2)!(K-x)!}\frac{(N-K)!}{(n-x)!(N-K-n+x)!}}{\frac{(N-2)!}{(n-2)!(N-n)!}} \\
&= \frac{nK\left(K-1\right)\left(n-1\right)}{N\left(N-1\right)}
\end{aligned}
$$

所以

$$
\begin{aligned}
\mathrm{Var}\left(X\right) &= \mathrm{E}X\left(X-1\right)+\mathrm{E}X-\left(\mathrm{E}X\right)^{2} \\
&= \frac{nK\left(K-1\right)\left(n-1\right)}{\mathcal{N}\left(N-1\right)}+n\frac{K}{N}-\left(n\frac{K}{N}\right)^{2} \\
&= n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}
\end{aligned}
$$

例 **4–6** 一袋內有 6 個紅球及 4 個白球，任取 3 個球，請問

(a) 抽中 2 個紅球之機率

(b) 至少一個紅球之機率

解：

(a)

$$\mathrm{P}\left(X=2\right)=\frac{\binom{6}{2}\binom{4}{1}}{\binom{10}{3}}=0.5$$

(b)

$$\mathrm{P}\left(X\geq 1\right)=1-\frac{\binom{6}{0}\binom{4}{3}}{\binom{10}{3}}=0.9667$$

例 **4–7** 一盒火星塞中有 10 個，其中有 2 個是不良品，自此盒中一把取出 3 個為一樣本，求

(a) 該 3 個中含有 1 個不良品機率

(b) 平均數及變異數

解：

(a) $X\sim h\left(N=10,K=2,n=3\right)$，所以

$$\mathrm{P}\left(X=1\right)=\frac{\binom{2}{1}\binom{8}{2}}{\binom{10}{3}}=\frac{7}{15}$$

(b)

$$\mu=3\times\frac{2}{10}=0.06,\sigma^2=\frac{10-3}{10-1}\times 3\times\frac{2}{10}\times\frac{10-2}{10}=\frac{28}{75}$$

例 **4–8** A lot of size $N=30$ contains five noconforming units. What is the probability that a sample of five units selected at a random contains exactly one noconforming unit？What is the probability that is contains one or more noconforming units？

解：

(a)

$$\mathrm{P}(X=1)=\frac{\binom{5}{1}\binom{25}{4}}{\binom{30}{5}}=0.44384$$

(b)

$$\mathrm{P}(X\geq 1)=1-\mathrm{P}(X=0)=1-\frac{\binom{5}{0}\binom{25}{5}}{\binom{30}{5}}=0.62717$$

例 **4–9** Three cards are drawn without replacement from 12 cards (Jacks, Queens and Kings)of an ordinary deck of 52 playing cards. Let X be the number of Kings selected and Y be the number of Jacks.

(a) Find the joint probability distribution of X and Y.

(b) Find $\mathrm{P}(X+Y\geq 2)$

(c) Find $\mathrm{E}X$

(d) Find the marginal distribution of Y.

(e) Find the conditional distribution of X, given $Y=1$.

解：

(a)

$$f(x,y)=\frac{\binom{4}{x}\binom{4}{y}\binom{4}{3-x-y}}{\binom{12}{3}},x=0,1,2,3;y=0,1,2,3;x+y\leq 3$$

或

		Y			
		0	1	2	3
	0	$\frac{1}{55}$	$\frac{6}{55}$	$\frac{6}{55}$	$\frac{1}{55}$
X	1	$\frac{6}{55}$	$\frac{16}{55}$	$\frac{6}{55}$	0
	2	$\frac{6}{55}$	$\frac{6}{55}$	0	0
	3	$\frac{1}{55}$	0	0	0

(b)

$$\begin{aligned}\mathrm{P}(X+Y\geq 2) &= 1-\mathrm{P}(X=0,Y=0)-\mathrm{P}(X=0,Y=1)-\mathrm{P}(X=1,Y=0)\\ &= 1-\frac{\binom{4}{0}\binom{4}{0}\binom{4}{3-0-0}}{\binom{12}{3}}-\frac{\binom{4}{1}\binom{4}{0}\binom{4}{3-1-0}}{\binom{12}{3}}-\frac{\binom{4}{0}\binom{4}{1}\binom{4}{3-0-1}}{\binom{12}{3}}\\ &= 0.7636\end{aligned}$$

(c)

$$\mathrm{E}X=0\times\frac{14}{55}+1\times\frac{28}{55}+2\times\frac{12}{55}+3\times\frac{1}{55}=1$$

(d)

Y	0	1	2	3
$f_2(y)$	$\frac{14}{55}$	$\frac{28}{55}$	$\frac{12}{55}$	$\frac{1}{55}$

(e)

x	0	1	2
$f(x\|y=1)$	$\frac{3}{14}$	$\frac{8}{14}$	$\frac{3}{14}$

例 **4–10** 有 A, B 兩箱，其中 A 箱中有 1 顆白球及 4 顆紅球，B 箱中有 5 顆紅球。今由 A 箱中抽取 3 顆球放入 B 箱，再由 B 箱中抽取 4 顆球放入 A 箱，試問白球落入 A 箱的機率為何？又白球落入 B 箱的機率為何？

解：

$$\begin{aligned} \mathrm{P}\left(\text{白球落入 A 箱}\right) &= \mathrm{P}\left(\text{ 白球從未落入 B 箱}\right) + \mathrm{P}\left(\text{白球由 A 箱放入 B 箱, 再由 B 箱放回 A 箱}\right) \\ &= \frac{\binom{1}{0}\binom{4}{3}}{\binom{5}{3}}\frac{\binom{8}{4}}{\binom{8}{4}} + \frac{\binom{1}{1}\binom{4}{2}}{\binom{5}{3}}\frac{\binom{1}{1}\binom{7}{3}}{\binom{8}{4}} = 0.7 \end{aligned}$$

及

$$\begin{aligned} \mathrm{P}\left(\text{白球落入 B 箱}\right) &= \mathrm{P}\left(\text{白球先由 A 箱放入 B 箱，並未由 B 箱放回 A 箱}\right) \\ &= \frac{\binom{1}{1}\binom{4}{2}}{\binom{5}{3}}\frac{\binom{1}{0}\binom{8}{4}}{\binom{9}{4}} = \frac{1}{3} \end{aligned}$$

例 **4–11** 假設樂透是由 1 到 42 的號碼中抽出 6 個號碼，請問開出的號碼都是奇數的機率為何？

解：

$$\mathrm{P}(6\text{ 個號碼皆為奇數}) = \frac{\binom{21}{6}\binom{21}{0}}{\binom{42}{6}}$$

例 **4–12** 全班男女生共 51 人，票選畢業旅行的目的地，每人限投一票，結果如下表。線以簡單隨機抽樣，抽出兩人，若這兩人都是女生，則這兩人都想去墾丁的機率為何？

	女	男
墾丁	10	10
澎湖	6	10
花東	9	6

解：

$$\frac{\binom{10}{2}\binom{15}{0}}{\binom{25}{2}} = 0.15$$

6. **卜瓦松分配**（Poisson Distribution）：若 X 表單位時間內出現 x 次之事件，若符合下列三項假設

   (a) 在任意兩個不重疊區間內發生次數之事件為相互獨立

   (b) 在充分小的區間 h 內恰巧出現一次的機率為 λh

   (c) 在充分小的區間內出現兩次以上的機率趨近於零

   則 X 有卜瓦松分配，以 $X \sim Poisson(\lambda)$ 表示，λ 為其參數，其 p.d.f. 為：

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!},\ x = 0, 1, 2\dots$$

   性質

   (a) $\mathrm{E}(X) = \lambda$

   (b) $\mathrm{Var}(X) = \lambda$

   (c) $\mathrm{E}X\,(X-1)\,(X-1)\cdots(X-r+1) = \lambda^r$

   (d) $M(t) = e^{\lambda\left(e^t-1\right)}$

   (e) 若 $X \sim Poisson(\lambda_1),\ Y \sim Poisson(\lambda_2)$ 且 X 和 Y 獨立，則 $X+Y \sim \mathrm{P}(\lambda_1+\lambda_2)$

   (f) 與二項分配之關係為

$$\lim_{\substack{n\to\infty \\ p\to 0}} \binom{n}{x} p^x(1-p)^{n-x} = \frac{\mu^x e^{-\mu}}{x!},\ \text{其中}\ \mu = np$$

   證明.

   (a)

$$\begin{aligned}
\mathrm{E}X &= \sum_{x=0}^{\infty} x\frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=1}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-1)!}
\end{aligned}$$

$$
\begin{aligned}
&= \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \\
&= \lambda
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}X(X-1) &= \sum_{x=1}^{\infty} x(x-1) \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=2}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-2)!} \\
&= \lambda^2
\end{aligned}
$$

所以 $\mathrm{Var}(X) = \mathrm{E}X^2 - (\mathrm{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

(c)

$$
\begin{aligned}
\mathrm{E}X(X-1)(X-1)\cdots(X-r+1) &= \sum_{x=0}^{\infty} x(x-1)(x-2)\cdots(x-r+1) \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{(x-r)!} = \lambda^r \sum_{x=r}^{\infty} \frac{\lambda^{x-r} e^{-\lambda}}{(x-r)!} \\
&= \lambda^r
\end{aligned}
$$

(d)

$$
\begin{aligned}
M(t) &= \sum_{x=0}^{n} \frac{\lambda^x e^{-\lambda}}{x!} e^{tx} \\
&= e^{-\lambda} \sum_{x=0}^{n} \frac{\left(\lambda e^t\right)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda\left(e^t - 1\right)}
\end{aligned}
$$

(e)

$$
\begin{aligned}
M_{X+Y}(t) &= \mathrm{E}\left(e^{t(x+y)}\right) \\
&= \mathrm{E}\left(e^{tx} e^{ty}\right) \\
&= \mathrm{E}\left(e^{tx}\right) \mathrm{E}\left(e^{ty}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= e^{\lambda_1\left(e^t-1\right)}e^{\lambda_2\left(e^t-1\right)} \\
&= e^{(\lambda_1+\lambda_2)\left(e^t-1\right)}
\end{aligned}
$$

所以，由動差生成函數之定義可知 $X+Y \sim Poisson\,(\lambda_1+\lambda_2)$

(f)

$$
\begin{aligned}
\binom{n}{x}p^x\,(1-p)^{n-x} &= \frac{(n-x+1)\,(n-x+2)\ldots(n-1)\,n}{x!}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x} \\
&= \frac{(n-x+1)\,(n-x+2)\ldots(n-1)\,n}{x!}\left(\frac{\mu}{n}\right)^x\left(\frac{n-\mu}{n}\right)^{-x}\left(1+\frac{-\mu}{n}\right)^n \\
&= \frac{(n-x+1)\,(n-x+2)\ldots(n-1)\,n}{(n-\mu)^x}\frac{\mu^x}{x!}\left(1+\frac{-\mu}{n}\right)^n \\
&= \frac{(n-x+1)}{(n-np)}\frac{(n-x+2)}{(n-np)}\ldots\frac{(n-1)}{(n-np)}\frac{n}{(n-np)}\frac{\mu^x}{x!}\left(1+\frac{-\mu}{n}\right)^n
\end{aligned}
$$

由於 $n\to\infty$ 且 $p\to 0$，因此

$$
\lim_{\substack{n\to\infty\\ p\to 0}}\frac{(n-x+1)}{(n-np)}\frac{(n-x+2)}{(n-np)}\ldots\frac{(n-1)}{(n-np)}\frac{n}{(n-np)}=1
$$

又 $\lim_{n\to\infty}\left(1+\frac{-\mu}{n}\right)^n=e^{-\mu}$，故得

$$
\lim_{\substack{n\to\infty\\ p\to 0}}\binom{n}{x}p^x(1-p)^{n-x}=\frac{\mu^x e^{-\mu}}{x!}
$$

例 **4–13** 若 $X \sim Poisson(\lambda)$，試求 $\mathrm{E}(1/(X+1))$

解：

$$
\begin{aligned}
\mathrm{E}\left(\frac{1}{X+1}\right) &= \sum_{x=0}^{\infty}\frac{1}{x+1}\frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=0}^{\infty}\frac{\lambda^x e^{-\lambda}}{(x+1)!}\frac{\lambda}{\lambda}=\frac{e^{-\lambda}}{\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x+1}}{(x+1)!} \\
&= \frac{e^{-\lambda}}{\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x+1}}{(x+1)!} \\
&= \frac{e^{-\lambda}}{\lambda}\left(e^{\lambda}-1\right)
\end{aligned}
$$

例 **4–14** 經銷商只有週日能進貨。依過去的經驗，平均每週賣 5 台冰箱，試求解下週供不應求之機率為何？

解：

假設 X 表下週之需求量，則

$$\begin{aligned}
\mathrm{P}(\text{供不應求}) &= \mathrm{P}(X > 5)\\
&= 1 - \mathrm{P}(X \le 5)\\
&= 1 - \left(\frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \cdots + \frac{5^5 e^{-5}}{5!}\right)\\
&= 0.384
\end{aligned}$$

例 **4–15** 根據主計處的資料顯示，全台灣大約有 2% 的成年人具有碩士以上學位。現在自全國成年人中，隨機抽取 100 位成年人，其中洽有 3 人擁有碩士以上學位的機率為何？

解：

令 X 代表擁有碩士以上學位的人數，所以 $X \sim B(100, 0.02)$，又 $p = 0.02 \le 0.05$, 故可以 Poisson 近似。所以

$$\mathrm{P}(x = 3) \approx \frac{e^{-2}2^3}{3!} = 0.1804$$

與眞實值 $\mathrm{P}(X = 3) = \binom{100}{3}0.02^3 0.98^{97} = 0.1823$ 相比較有些誤差, 但若 n 越大時，其結果與眞實值就越接近。

例 **4–16** Let X be the number of flaws on the surface of a randomly selected boiler of a certain type. Suppose X has a Poisson distribution when mean λ. However, the mean λ is unknown. Past experience indicates that the prior probabilities of λ are $P\{\lambda = 4\} = 0.3$, $P\{\lambda = 5\} = 0.5$, and $P\{\lambda = 6\} = 0.2$.

(a) Given that the true mean of X is 5, what is the probability that a randomly selected boiler has at most two flaws？

(b) Under the prior probabilities, what is the probability that a randomly selected boiler has exactly two flaws？

(c) Suppose that a boiler was randomly selected and found that it had two flaws. Using this information, revise the prior probabilities of λ.

解：

(a)

$$\mathrm{P}\left(X \leq 2|\lambda=5\right)=\sum_{x=0}^{2}\frac{5^{x}e^{-5}}{x!}=0.1247$$

(b)

$$\begin{aligned}
\mathrm{P}\left(X=2\right) &= \mathrm{P}\left(X=2\cap\lambda=4\right)+\mathrm{P}\left(X=2\cap\lambda=5\right)+\mathrm{P}\left(X=2\cap\lambda=6\right)\\
&= \mathrm{P}\left(\lambda=4\right)\mathrm{P}\left(X=2|\lambda=4\right)+\mathrm{P}\left(\lambda=5\right)\mathrm{P}\left(X=2|\lambda=5\right)\\
&\quad +\mathrm{P}\left(\lambda=6\right)\mathrm{P}\left(X=2|\lambda=6\right)\\
&= 0.3\times\frac{4^{2}e^{-4}}{2!}+0.5\times\frac{5^{2}e^{-5}}{2!}+0.2\times\frac{6^{2}e^{-6}}{2!}\\
&= 0.095
\end{aligned}$$

(c)

$$\begin{aligned}
\mathrm{P}\left(\lambda=4|X=2\right) &= \frac{\mathrm{P}\left(X=2\cap\lambda=4\right)}{\mathrm{P}\left(X=2\cap\lambda=4\right)+\mathrm{P}\left(X=2\cap\lambda=5\right)+\mathrm{P}\left(X=2\cap\lambda=6\right)}\\
&= \frac{0.3\times4^{2}e^{-4}/2!}{0.3\times4^{2}e^{-4}/2!+0.5\times5^{2}e^{-5}/2!+0.2\times6^{2}e^{-6}/2!}\\
&= 0.4627,
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}\left(\lambda=4|X=2\right) &= \frac{\mathrm{P}\left(X=2\cap\lambda=5\right)}{\mathrm{P}\left(X=2\cap\lambda=4\right)+\mathrm{P}\left(X=2\cap\lambda=5\right)+\mathrm{P}\left(X=2\cap\lambda=6\right)}\\
&= \frac{0.5\times5^{2}e^{-5}/2!}{0.3\times4^{2}e^{-4}/2!+0.5\times5^{2}e^{-5}/2!+0.2\times6^{2}e^{-6}/2!}\\
&= 0.4433,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{P}\left(\lambda=4|X=2\right) &= \frac{\mathrm{P}\left(X=2\cap\lambda=6\right)}{\mathrm{P}\left(X=2\cap\lambda=4\right)+\mathrm{P}\left(X=2\cap\lambda=5\right)+\mathrm{P}\left(X=2\cap\lambda=6\right)}\\
&= \frac{0.2\times6^{2}e^{-6}/2!}{0.3\times4^{2}e^{-4}/2!+0.5\times5^{2}e^{-5}/2!+0.2\times6^{2}e^{-6}/2!}\\
&= 0.0939
\end{aligned}$$

例 **4–17** $f(x)$ is a p.d.f, $f(x)>0$, given that $f\left(x\right)=4f\left(x-1\right)/x$, $x=1,2,3,\ldots$, find $f(x)$

解：

$f\left(x\right)=4f\left(x-1\right)/x,\ x=1,2,3,\ldots$，令 $f(0)=p$，則

$$f\left(0\right)=p,\ f\left(1\right)=\frac{4}{1}p,\ f\left(2\right)=\frac{4}{2}\frac{4}{1}p,...,\ f\left(x\right)=\frac{4}{x}\cdots\frac{4}{2}\frac{4}{1}p=\frac{4^{x}}{x!}p$$

又機率總和為 1，所以

$$\begin{aligned}\sum_{x=0}^{\infty}f\left(x\right) &= \sum_{x=0}^{\infty}\frac{4}{x}f\left(x-1\right)\\ &= \sum_{x=0}^{\infty}\frac{4^{x}p}{x!}\\ &= 1\end{aligned}$$

得 $p=e^{-4}$，由遞迴關係式可推得 $f(x)=4^{x}e^{-4}/x!$

7. **幾何分配**（Geometric Distribution）：在有限母體中可分為成功和失敗情形下，樣本抽出後放回，考慮第一次成功所需次數的隨機實驗，稱為幾何分配，通常以 $X\sim Geometric(p)$ 表示，其 p.d.f. 為

$$f(x)=\mathrm{P}(1-p)^{x-1},\ x=1,\ 2,\ 3,\ldots$$

其中

(a) x 為重複獨立實驗次數

(b) p 為每次實驗之成功機率

性質

(a) $\mathrm{P}(X>x)=(1-p)^{x}$

(b) $\mathrm{E}X=\dfrac{1}{p}$

(c) $\mathrm{Var}(X)=\dfrac{1-p}{p^{2}}$

(d) $M(t)=\dfrac{pe^{t}}{1-(1-p)e^{t}}$

(e) 具無記憶性即 $\mathrm{P}(X>x+y\mid X>x)=\mathrm{P}(X>y)$

證明. (a)

$$\mathrm{P}(X>x) = \mathrm{P}(1-p)^{x}+\mathrm{P}(1-p)^{x+1}+\mathrm{P}(1-p)^{x+1}+\cdots$$

$$
\begin{aligned}
&= \frac{\mathrm{P}(1-p)^x}{1-(1-p)} \\
&= (1-p)^x
\end{aligned}
$$

(b) 令

$$g(p) = \sum_{x=0}^{\infty} (1-p)^x = \frac{1}{p}$$

则

$$g'(p) = -\sum_{x=1}^{\infty} x(1-p)^{x-1} = -\frac{1}{p^2}$$

$$g''(p) = \sum_{x=1}^{\infty} x(x-1)(1-p)^{x-2} = \frac{2}{p^3}$$

所以,

$$
\begin{aligned}
\mathrm{E}X &= \sum_{x=1}^{\infty} x\mathrm{P}(1-p)^{x-1} \\
&= p\sum_{x=1}^{\infty} x(1-p)^{x-1} \\
&= p \times \frac{1}{p^2} \\
&= \frac{1}{p}
\end{aligned}
$$

(c)

$$
\begin{aligned}
\mathrm{E}X(X-1) &= \sum_{x=2}^{\infty} x(x-1)\mathrm{P}(1-p)^{x-1} \\
&= \mathrm{P}(1-p)\sum_{x=2}^{\infty} x(x-1)(1-p)^{x-2} \\
&= \mathrm{P}(1-p) \times \frac{2}{p^3} \\
&= \frac{2p-2p^2}{p^3}
\end{aligned}
$$

故得

$$\mathrm{Var}(X) = \mathrm{E}X(X-1) + \mathrm{E}X - (\mathrm{E}X)^2$$

$$
\begin{aligned}
&= \frac{2p-2p^2}{p^3}+\frac{1}{p}-\frac{1}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}
$$

(d)

$$
\begin{aligned}
M(t) &= \mathrm{E}\left(e^{tx}\right)=\sum_{x=1}^{\infty}\mathrm{P}(1-p)^{x-1}e^{tx} \\
&= \sum_{x=1}^{\infty}\mathrm{P}(1-p)^{x-1}\left(e^{t}\right)^{x-1}e^{t} \\
&= pe^{t}\sum_{x=1}^{\infty}\left((1-p)e^{t}\right)^{x-1} \\
&= \frac{pe^{t}}{1-(1-p)e^{t}}
\end{aligned}
$$

(e)

$$
\begin{aligned}
\mathrm{P}\left(X>x+y|X>x\right) &= \frac{\mathrm{P}\left(X>x+y\cap X>x\right)}{\mathrm{P}\left(X>x\right)}=\frac{\mathrm{P}\left(X>x+y\right)}{\mathrm{P}\left(X>x\right)} \\
&= \frac{\mathrm{P}\left(1-p\right)^{x+y}}{\mathrm{P}\left(1-p\right)^{x}} \\
&= (1-p)^{y} \\
&= \mathrm{P}\left(X>y\right)
\end{aligned}
$$

例 **4–18**

(a) 假設 X 表投擲一骰子直達點數 3 出現所需次數之事件，試求 $\mathrm{E}X$ 及 $\mathrm{Var}(X)$

(b) 假設 X 表投擲一對骰子至少出現一個點數 3 所需次數之事件，試求 $\mathrm{E}X$ 及 $\mathrm{Var}(X)$

(c) 假設 X 表同時投擲 n 個骰子至少出現一個點數 3 出現所需次數之事件，試求在 $\mathrm{E}X \le 2$ 之限制下所須之最所需之最小整數 n

解：

(a) $X \sim \mathrm{Geometric}(1/6)$，所以

$$
\mathrm{E}X=\frac{1}{p}=6,\ \mathrm{Var}(X)=\frac{1-p}{p^2}=30
$$

(b)

$$\mathrm{P}(\text{投擲一對骰子至少出現一個點數 } 3) = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}$$

所以 $X \sim \mathrm{Geometric}(11/36)$

$$\mathrm{E}X = \frac{1}{p} = \frac{36}{11},\ \mathrm{Var}(X) = \frac{1-p}{p^2} = \frac{900}{121}$$

(c)

$$\mathrm{P}(\text{投擲 } n \text{ 個骰子至少出現一個點數 } 3) = 1 - \left(\frac{5}{6}\right)^n$$

所以 $X \sim \mathrm{Geometric}(1-(5/6)^n)$，又

$$\frac{1}{1-(5/6)^n} \leq 2$$

得 $n = 4$

例 **4–19** A company puts six types of collectable into their product boxes, one in each box and in equal proportions. If a customer decides to collect all six of the collectable, what is the expected number of the product boxes that he or she should buy？

解：令 X_i 表第 i 個收藏物，所以 $X_i \sim Geometric(1-\frac{i-1}{6})$，所以

$$\begin{aligned} E\left(X_1+X_2+X_3+X_4+X_5+X_6\right) &= \frac{1}{1}+\frac{1}{5/6}+\frac{1}{4/6}+\frac{1}{3/6}+\frac{1}{2/6}+\frac{1}{1/6} \\ &= 14.7 \end{aligned}$$

8. **負二項分配**（Negative Binomial Distribution）：在有限母體中可分為成功和失敗情形下，樣本抽出後放回，考慮第 r 次成功所需次數的隨機實驗，稱為負二項分配，通常以 $X \sim NB(p,r)$ 表示，其 p.d.f. 為

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r},\ x = r,\ r+1,\ \ r+2, \ldots$$

其中

(a) x 為重複獨立實驗次數

(b) p 為每次實驗之成功機率

性質

(a) $\mathrm{E}X = \dfrac{r}{p}$

(b) $\mathrm{Var}(X) = \dfrac{r(1-p)}{p^2}$

(c) $M(t) = \left[\dfrac{pe^t}{1-(1-p)e^t}\right]^r$

(d) 當 $r = 1$ 時，則負二項分配與幾何分配相同。

證明.

(a)

$$
\begin{aligned}
\mathrm{E}X &= \sum_{x=r}^{\infty} x\binom{x-1}{r-1}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} x\frac{(x-1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{x(x-1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r}\frac{p}{p}\frac{r}{r} \\
&= \frac{r}{p}\sum_{x=r}^{\infty} \frac{x!}{r!(x-r)!}p^{r+1}(1-p)^{x-r}
\end{aligned}
$$

令 $w = x+1$，則上式可寫成

$$
\begin{aligned}
\mathrm{E}X &= \frac{r}{p}\sum_{w=r+1}^{\infty} \binom{w-1}{r+1-1}p^{r+1}(1-p)^{w-(r+1)} \\
&= \frac{r}{p}
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{E}X(X+1) &= \sum_{x=r}^{\infty} x(x+1)\binom{x-1}{r-1}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{x(x+1)(x-1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{x(x+1)(x-1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r}\frac{p^2}{p^2}\frac{r(r+1)}{r(r+1)} \\
&= \frac{r(r+1)}{p^2}\sum_{x=r}^{\infty} \frac{(x+1)!}{(r+1)!(x-r)!}p^{r+2}(1-p)^{x-r}
\end{aligned}
$$

令 $w=x+2$，則上式可寫成

$$\begin{aligned}\mathrm{E}X(X+1) &= \frac{r(r+1)}{p^2}\sum_{w=r+2}^{\infty}\binom{w-1}{r+2-1}p^{r+2}(1-p)^{w-(r+2)}\\ &= \frac{r(r+1)}{p^2}\end{aligned}$$

所以,

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}X(X+1)-\mathrm{E}X-(\mathrm{E}X)^2\\ &= \frac{r(r+1)}{p^2}-\frac{r}{p}-\frac{r^2}{p^2}\\ &= \frac{r+r^2-rp-r^2}{p^2}\\ &= \frac{r(1-p)}{p^2}\end{aligned}$$

(c)

$$\begin{aligned}M(t) &= \mathrm{E}\left(e^{tx}\right)=\sum_{x=1}^{\infty}\binom{x-1}{r-1}p^r(1-p)^{x-r}e^{tx}\\ &= \sum_{x=1}^{\infty}\binom{x-1}{r-1}p^r(1-p)^{x-r}\left(e^t\right)^{x-r}\left(e^t\right)^r\\ &= \left(pe^t\right)^r\sum_{x=1}^{\infty}\left[(1-p)e^t\right]^{x-r}\end{aligned}$$

又由 Maclaurin Series 可知

$$(1-w)^{-r}=1+\frac{r}{1!}w+\frac{r(r+1)}{2!}w^2+\frac{r(r+1)(r+2)}{3!}w^3+\cdots=\sum_{k=0}^{\infty}\binom{r+k-1}{r-1}w^k$$

令 $x=k+r$，所以

$$\begin{aligned}M(t) &= \left(pe^t\right)^r\sum_{x=1}^{\infty}\binom{x-1}{r-1}\left[(1-p)e^t\right]^{x-r}\\ &= \left(pe^t\right)^r\sum_{k=0}^{\infty}\binom{r+k-1}{r-1}\left[(1-p)e^t\right]^{k}\end{aligned}$$

$$= \left[\frac{pe^t}{1-(1-p)e^t}\right]^r$$

例 **4–20** 假設某籃球選手投籃命中之機率為 0.8，今假設每次投籃命中與否為相互獨立之事件。若 X 表該選手第十次命中所須之投籃次數，試求 X 之期望值及變異數？

解：

$X \sim NB(0.8, 10)$，所以

$$f(x) = \binom{x-1}{9} 0.8^{10} 0.2^{x-10},\ x = 10, 11, \ldots$$

$$\mathrm{E}X = \frac{r}{p} = \frac{10}{0.8} = 12.5,\ \mathrm{Var}(X) = \frac{r(1-p)}{p^2} = \frac{10 \times 0.2}{0.64} = 3.125$$

例 **4–21** 一 CD 播放機可放 6 片 CD，此播放機能夠隨機從 6 片 CD 中選出一片 CD，之後再從此片 CD 隨機選出一首歌，假設有 5 片 CD 為保羅麥卡尼的歌，另 1 片 CD 為比利喬的歌。此機器選歌直到比利喬的歌被播放之後，機器被關掉。假設直到比利喬的第 2 首歌被播放之後，機器被關掉，求最多 4 首歌被播放的機率。

解：令 X 表第二次撥放比利喬的歌曲所必須撥放的歌數，所以 $X \sim NB(\frac{1}{6}, 2)$

$$\begin{aligned} P(X \le 4) &= \sum_{x=2}^{4} \binom{x-1}{2-1} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{x-2} \\ &= 0.13194 \end{aligned}$$

4.2 連續型機率分配

1. **連續型均勻分配** (Uniform Distribution)：X 有均勻分配，以 $X \sim U(a, b)$ 表示之，其中 $0 \le a \le b < \infty$，其 p.d.f. 為

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

特性

(a) 定義域在 (a, b)

(b) 對稱

(c) 機率密度函數為一常數

性質

(a) $\mathrm{E}(X) = \dfrac{a+b}{2}$

(b) $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$

(c) $F(X) = \mathrm{P}(X \le x) = \dfrac{x-a}{b-a}$

證明.

(a)

$$
\begin{aligned}
\mathrm{E}X &= \int_a^b x\frac{x}{b-a}dx \\
&= \frac{1}{b-a} \times \frac{1}{2}x^2\Big|_a^b \\
&= \frac{1}{b-a} \times \frac{b^2-a^2}{2} \\
&= \frac{a+b}{2}
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{Var}(X) &= \int_a^b \frac{(x-\mathrm{E}X)^2}{b-a}dx \\
&= \int_a^b \frac{x^2}{b-a}dx - (\mathrm{E}X)^2 \\
&= \frac{1}{b-a} \times \frac{1}{3}x^3\Big|_a^b - \frac{a^2+2ab+b^2}{4} \\
&= \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} \\
&= \frac{a^2-2ab+b^2}{12} \\
&= \frac{(b-a)^2}{12}
\end{aligned}
$$

(c)

$$
\begin{aligned}
F(X) &= \int_a^x \frac{1}{b-a} dt \\
&= \left. \frac{t}{b-a} \right|_a^x \\
&= \frac{x-a}{b-a}
\end{aligned}
$$

例 **4–22** 火車站進站的人數呈現 Poisson 分配，期望值為 λt 而火車進站的時間呈現 Uniform 分配，時間為 $(0,T)$，這兩者互相獨立，求乘客進入火車的人數期望值及變異數。

解：$X|Y \sim Poisson\,(\lambda t),\, Y \sim U\,(0,T)$

$$
\mathrm{E}\,(X) = \mathrm{EE}\,(X|Y) = \mathrm{E}\lambda t = \lambda \times \frac{T}{2} = \frac{\lambda T}{2}
$$

$$
\begin{aligned}
\mathrm{Var}\,(X) &= \mathrm{E}\,(\mathrm{Var}\,(X|Y)) + \mathrm{Var}\,(\mathrm{E}\,(X|Y)) \\
&= \mathrm{E}\,(\lambda t) + \mathrm{Var}\,(\lambda t) \\
&= \frac{\lambda T}{2} + \frac{\lambda^2 T^2}{12}
\end{aligned}
$$

2. **指數分配** (Exponential Distribution)：X 有指數分配以 $X \sim \mathrm{E}X\mathrm{P}(\theta)$ 表示之，其中 θ 為母體參數，其 p.d.f. 為

$$
f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}},\, 0 \leq x < \infty
$$

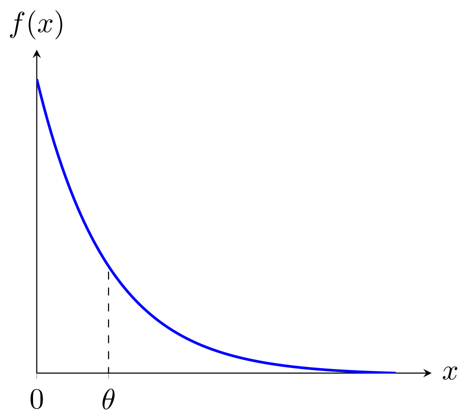


圖 4.1: 指數分配圖

特性

(a) 定義域在 $[0,\ \infty)$

(b) 不對稱, 為一右偏分配

(c) 參數 θ 不同，其圖形有所不同

性質

(a) $\mathrm{E}(X)=\theta$

(b) $\mathrm{Var}(X)=\theta^2$

(c) $M(t)=(1-\theta t)^{-1}$

(d) 具無記憶性，即 $\mathrm{P}(X>x+y \mid X>x)=\mathrm{P}(X>y)$

(e) 令 W 表事件第一次發生所需經過之時間，若單位時間內事件發生次數之機率分配服從 $\mathrm{Poisson}(\lambda)$，則 $f(w)=\lambda e^{-\lambda w},\ 0\le w<\infty$

證明.

(a)

$$\begin{aligned}\mathrm{E}X &= \int_0^\infty x\frac{1}{\theta}e^{-\frac{x}{\theta}}dx\\ &= -xe^{-\frac{x}{\theta}}\Big|_0^\infty+\int_0^\infty e^{-\frac{x}{\theta}}dx\\ &= -\theta e^{-\frac{x}{\theta}}\Big|_0^\infty\\ &= \theta\end{aligned}$$

(b)

$$\begin{aligned}\mathrm{E}X^2 &= \int_0^\infty x^2\frac{1}{\theta}e^{-\frac{x}{\theta}}dx\\ &= -x^2e^{-\frac{x}{\theta}}\Big|_0^\infty+2\int_0^\infty xe^{-\frac{x}{\theta}}dx\\ &= 2\theta^2\end{aligned}$$

所以

$$\mathrm{Var}(X)=\mathrm{E}X^2-(\mathrm{E}X)^2=\theta^2$$

(c)

$$\begin{aligned}
M(t) &= \int_0^\infty e^{tx}\frac{1}{\theta}e^{-\frac{x}{\theta}}dx \\
&= \int_0^\infty \frac{1}{\theta}e^{-\frac{(1-\theta t)x}{\theta}}dx \\
&= \frac{1}{(1-\theta t)}\int_0^\infty e^{-\frac{(1-\theta t)x}{\theta}}d\frac{(1-\theta t)x}{\theta} \\
&= \frac{1}{(1-\theta t)}\left(-e^{-\frac{(1-\theta t)x}{\theta}}\bigg|_0^\infty\right) \\
&= (1-\theta t)^{-1}
\end{aligned}$$

(d)

$$\begin{aligned}
\mathrm{P}(X > x+y \mid X > x) &= \frac{\mathrm{P}(X > x+y \cap X > x)}{\mathrm{P}(X > x)} \\
&= \frac{\mathrm{P}(X > x+y)}{\mathrm{P}(X > x)} \\
&= \frac{\int_{x+y}^\infty \frac{1}{\theta}e^{-\frac{s}{\theta}}ds}{\int_x^\infty \frac{1}{\theta}e^{-\frac{s}{\theta}}ds} \\
&= \frac{e^{-\frac{x+y}{\theta}}}{e^{-\frac{x}{\theta}}} \\
&= \mathrm{P}(X > y)
\end{aligned}$$

(e) 令 X 表單位時間內事件發生之次數，則

$$\begin{aligned}
F(w) &= \mathrm{P}(W < w) \\
&= 1 - \mathrm{P}(W \geq w) \\
&= 1 - \mathrm{P}(\text{在}W\text{時間內出現 0 次}) \\
&= 1 - \mathrm{P}(X = 0) \\
&= 1 - \frac{(\lambda w)^0 e^{-\lambda w}}{0!} \\
&= 1 - e^{-\lambda w}
\end{aligned}$$

所以,

$$f(w) = \frac{d}{dw}F(w) = \lambda e^{-\lambda w}$$

且指數分配之參數 $\theta = \frac{1}{\lambda}$

例 **4–23** A manufacturer of electronic calculators offers a one-year warranty. If the calculator fails for any reason during this period, it is replaced. The time to failure is well modeled by the following probability distribution:

$$f(x) = 0.125e^{-0.125x},\ x > 0.$$

What percentage of the calculators will fail within the warranty period？

解：

$$\mathrm{P}(X \le 1) = \int_0^1 0.125e^{-0.125x}dx = 1 - e^{-\frac{1}{8}}$$

例 **4–24** 設 $S = X_1 + X_2 + \cdots + X_N$，其中 X_i 彼此獨立且服從指數分配 $\mathrm{Exp}(\beta = 1/3)$，此外 N 服從卜瓦松分配 $Poisson(\lambda = 5)$，試求 S 的動差母函數？

解：

$$\begin{aligned}
\mathrm{E}(e^{ts}) &= \mathrm{EE}[\exp(tX_1 + tX_2 + \cdots + tX_N)\,|N] \\
&= \mathrm{E}[1/(1-\beta t)^N] \\
&= \sum_{N=0}^{\infty}\left(\frac{1}{1-\beta t}\right)^N \frac{\lambda^N e^{-\lambda}}{N!} \\
&= e^{-\lambda}\sum_{N=0}^{\infty}\frac{\left(\frac{\lambda}{1-\beta t}\right)^N}{N!} \\
&= \exp(-\lambda)\exp\left(\frac{\lambda}{1-\beta t}\right) \\
&= \exp\left(\frac{t\beta\lambda}{1-\beta t}\right) \\
&= \exp\left(\frac{5t/3}{1-t/3}\right)
\end{aligned}$$

3. **Gamma 分配**(Gamma Distribution)：令 W 表事件第 α 次發生所需經過之時間，若單位時間內事件發生次數之機率分配服從 Poisson(λ)，則

$$f(w) = \frac{\lambda(\lambda w)^{\alpha-1}}{(\alpha-1)!}e^{-\lambda w},\ 0 \le w < \infty$$

一般將其 p.d.f. 寫成

$$\frac{1}{\Gamma(\alpha)\theta^{\alpha}}w^{\alpha-1}e^{-\frac{w}{\theta}},\ 0\leq w<\infty$$

其中 $\theta=1/\lambda$, $\Gamma\left(\alpha\right)=\left(\alpha-1\right)!$

證明.

令 X 表單位時間內出現 x 次之事件，則

$$\begin{aligned}
F\left(w\right) &= \mathrm{P}\left(W<w\right)\\
&= 1-\mathrm{P}\left(W\geq w\right)\\
&= 1-\mathrm{P}\left(\text{在}W\text{時間內出現最多 }\alpha-1\text{ 次}\right)\\
&= 1-\mathrm{P}\left(X\leq\alpha-1\right)\\
&= 1-\sum_{x=0}^{\alpha-1}\frac{\left(\lambda w\right)^{x}e^{-\lambda w}}{x!}
\end{aligned}$$

所以,

$$\begin{aligned}
f\left(w\right) &= \frac{d}{dw}F\left(w\right)\\
&= \lambda e^{-\lambda w}-\lambda e^{-\lambda w}\sum_{x=1}^{\alpha-1}\left(\frac{x(\lambda w)^{x-1}\lambda}{x!}-\frac{\lambda(\lambda w)^{x}}{x!}\right)\\
&= \lambda e^{-\lambda w}-\lambda e^{-\lambda w}\sum_{x=1}^{\alpha-1}\left(\frac{(\lambda w)^{x-1}\lambda}{(x-1)!}-\frac{\lambda(\lambda w)^{x}}{x!}\right)\\
&= \lambda e^{-\lambda w}-\lambda e^{-\lambda w}\left\{1-\lambda w+\lambda w-\frac{(\lambda w)^{2}}{2!}+\frac{(\lambda w)^{2}}{2!}+\ldots+\frac{(\lambda w)^{\alpha-2}}{(\alpha-2)!}-\frac{(\lambda w)^{\alpha-1}}{(\alpha-1)!}\right\}\\
&= \lambda e^{-\lambda w}\frac{(\lambda w)^{\alpha-1}}{(\alpha-1)!}\\
&= \frac{1}{\Gamma(\alpha)\theta^{\alpha}}w^{\alpha-1}e^{-\frac{w}{\theta}}
\end{aligned}$$

定理 4.2.

$$\Gamma\left(\alpha\right)=\int_{0}^{\infty}x^{\alpha-1}e^{-x}dx,\ \alpha\text{為任一正整數。}$$

證明.

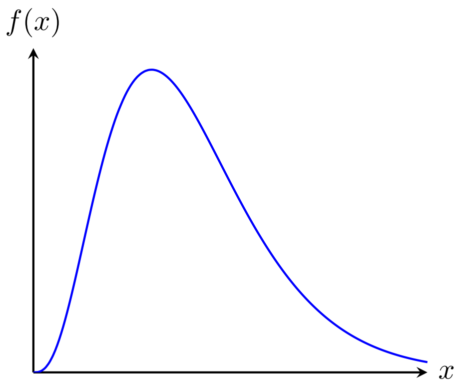


圖 4.2: Gamma 分配圖

由部分積分法可求得

$$
\begin{aligned}
\Gamma(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}dx \\
&= -x^{\alpha-1}e^{-x}\Big|_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx \\
&= -x^{\alpha-2}e^{-x}\Big|_0^\infty + (\alpha-1)(\alpha-2)\int_0^\infty x^{\alpha-2}e^{-x}dx \\
&\vdots \\
&= (\alpha-1)!\int_0^\infty e^{-x}dx \\
&= (\alpha-1)! = (\alpha-1)\Gamma(\alpha-1)
\end{aligned}
$$

所以,

$$
\begin{aligned}
\int_0^\infty \frac{1}{\theta^\alpha}w^{\alpha-1}e^{-\frac{w}{\theta}}dw &= \int_0^\infty \left(\frac{w}{\theta}\right)^{\alpha-1} e^{-\frac{w}{\theta}}d\frac{w}{\theta} \\
&= \Gamma(\alpha)
\end{aligned}
$$

及

$$
\int_0^\infty \frac{1}{\Gamma(\alpha)\theta^\alpha}w^{\alpha-1}e^{-\frac{w}{\theta}} = 1
$$

特性

(a) 定義域在 $[0,\ \infty)$

(b) 不對稱, 為一右偏分配

(c) 參數 θ 不同，其圖形有所不同

性質

(a) $EW = \alpha\theta$

(b) $\mathrm{Var}(W) = \alpha\theta^2$

(c) $M(t) = (1-\theta t)^{-\alpha}$

(d) 若 $X \sim \mathrm{Gamma}(\alpha_1, \theta)$, $Y \sim \mathrm{Gamma}(\alpha_2, \theta)$ 且 X 與 Y 相獨立，則 $X+Y \sim \mathrm{Gamma}(\alpha_1 + \alpha_2, \theta)$

證明.

(a)

$$
\begin{aligned}
EW &= \int_0^\infty \frac{w}{\Gamma(\alpha)\theta^\alpha} w^{\alpha-1} e^{-\frac{w}{\theta}} dw \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^\infty w^\alpha e^{-\frac{w}{\theta}} dw \\
&= \frac{\theta}{\Gamma(\alpha)} \int_0^\infty \left(\frac{w}{\theta}\right)^\alpha e^{-\frac{w}{\theta}} d\frac{w}{\theta} \\
&= \frac{\theta}{\Gamma(\alpha)} \times \Gamma(\alpha+1) = \alpha\theta
\end{aligned}
$$

(b)

$$
\begin{aligned}
EW^2 &= \int_0^\infty \frac{w^2}{\Gamma(\alpha)\theta^\alpha} w^{\alpha-1} e^{-\frac{w}{\theta}} dw \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^\infty w^{\alpha+1} e^{-\frac{w}{\theta}} dw \\
&= \frac{\theta^2}{\Gamma(\alpha)} \int_0^\infty \left(\frac{w}{\theta}\right)^{\alpha+1} e^{-\frac{w}{\theta}} d\frac{w}{\theta} \\
&= \frac{\theta^2}{\Gamma(\alpha)} \times \Gamma(\alpha+2) = \alpha(\alpha+1)\theta
\end{aligned}
$$

所以，

$$
\mathrm{Var}(W) = \alpha(\alpha+1)\theta^2 - \alpha^2\theta^2 = \alpha\theta^2
$$

(c)

$$
\begin{aligned}
M(t) &= \int_0^\infty \frac{e^{tw}}{\Gamma(\alpha)\theta^\alpha} w^{\alpha-1} e^{-\frac{w}{\theta}} dw \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^\infty w^{\alpha-1} e^{-\frac{(1-\theta t)w}{\theta}} dw \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \left(\frac{\theta}{(1-\theta t)}\right)^{\alpha-1} \int_0^\infty \left(\frac{(1-\theta t)w}{\theta}\right)^{\alpha-1} e^{-\frac{(1-\theta t)w}{\theta}} dw \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \left(\frac{\theta}{(1-\theta t)w}\right)^{\alpha-1} \frac{\theta}{(1-\theta t)} \int_0^\infty \left(\frac{(1-\theta t)}{\theta}\right)^{\alpha-1} e^{-\frac{(1-\theta t)w}{\theta}} d\frac{(1-\theta t)w}{\theta} \\
&= \frac{1}{\Gamma(\alpha)\theta^\alpha} \left(\frac{\theta}{(1-\theta t)w}\right)^{\alpha-1} \frac{\theta}{(1-\theta t)} \Gamma(\alpha) \\
&= (1-\theta t)^{-\alpha}
\end{aligned}
$$

(d)

$$
\begin{aligned}
M_{X+Y}(t) &= \mathrm{E}\left(e^{t(X+Y)}\right) = \mathrm{E}\left(e^{tX}\right)\mathrm{E}\left(e^{tY}\right) \\
&= (1-\theta t)^{-\alpha_1}(1-\theta t)^{-\alpha_2} \\
&= (1-\theta t)^{-(\alpha_1+\alpha_2)}
\end{aligned}
$$

所以，$X+Y \sim \text{Gamma}(\alpha_1+\alpha_2, \theta)$

例 **4–25** Suppose that a certain examination is to be taken by five students independently of one another, and that the number of minutes required by and particular student to complete the examination has an exponential distribution for which the mean is 80. Suppose that the examination begins at 9:00 A.M.

(a) Determine the probability that at least one of the students will complete the examination before 9:30 A.M.

(b) Determine the probability that no two students will complete the examination within 10 minutes of each others.

解：

(a) $X \sim \exp(80)$，所以

$$
\mathrm{P}(X \leq 30) = \int_0^{30} \frac{1}{80} e^{-\frac{x}{80}} dx = 0.3127
$$

(b) $Y \sim B(5, 0.3127)$，所以

$$\mathrm{P}(Y \geq 1) = 1 - \mathrm{P}(Y = 0) = 1 - \binom{5}{0}(0.3127)^0(1 - 0.3127)^5 = 0.84663$$

又

$$\lambda = \frac{1}{\theta} = \frac{1}{80}$$

$$\mathrm{P}(Z = 0) = \frac{\left(\frac{10}{80}\right)^0 e^{-\frac{10}{80}}}{0!} = 0.8825$$

4. **常態分配**（Normal Distribution）：X 有常態分配以 $X \sim \mathcal{N}(\mu, \sigma)$ 表示之，其中 μ 為母體期望值，σ 為母體變異數，其 p.d.f. 為

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

當期望值 $\mu = 0$ 且變異數 $\sigma = 1$ 時，稱為標準常態分配。

特性

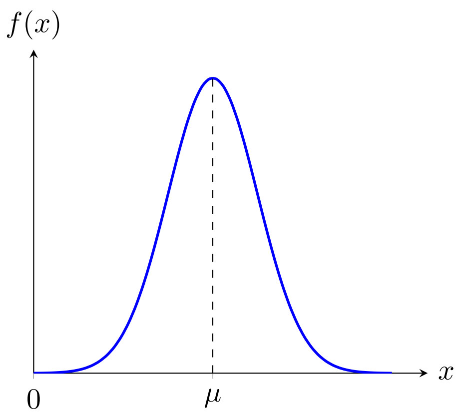


圖 4.3: 標準常態分配圖

(a) 以 $x = \mu$ 為中心線，左右對稱

(b) σ 的大小，表示圖形的高矮胖瘦

(c) X 軸 (即 $Y = 0$) 為水平漸近線

性質

(a) $\mathrm{E}(X) = \mu$

(b) $\mathrm{Var}(X) = \sigma^2$

(c) $M(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

(d) 若 $X \sim \mathcal{N}(\mu, \sigma^2)$，令 $Z = \frac{X-\mu}{\sigma}$ ，則 $Z \sim \mathcal{N}(0, 1)$，其中 Z 稱為標準常態分配，其平均數為 0，變異數為 1。

(e) 設有 n 個獨立的隨機變數 $X_1,\ X_2,\ \ldots, X_n$，其分配分別為 $\mathcal{N}(\mu_i,\ \sigma_i^2),\ i = 1, 2,\ \ldots, n$，令 $Y = \sum_{i=1}^{n} a_i X_i$ 且 a_i 為常數，則

$$Y \sim \mathcal{N}\left(\sum_{i=1}^{n} a_i \mu_i \ ,\ \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

(f) $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] dx,\ -\infty < x < \infty$

(g) $\mathrm{P}\left(Z > z_\alpha\right) = \mathrm{P}\left(Z < -z_\alpha\right) = \alpha$

特例

(a) 當 $a_i = 1, \forall\, i = 1, 2, \ldots, n$，則

$$Y = \sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\sum_{i=1}^{n} \mu_i,\ \sum_{i=1}^{n} \sigma_i^2\right)$$

(b) 當 $a_i = \frac{1}{n}, \forall\, i = 1, 2, \ldots, n$，則

$$Y = \sum_{i=1}^{n} \frac{X_i}{n} = \overline{X} \sim \mathcal{N}\left(\sum_{i=1}^{n} \frac{\mu_i}{n}\ ,\ \sum_{i=1}^{n} \frac{\sigma_i^2}{n^2}\right)$$

證明.

(a)

$$\begin{aligned}
\mathrm{E}X &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \\
&= \int_{-\infty}^{\infty} \frac{x-\mu+\mu}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \\
&= \int_{-\infty}^{\infty} \frac{x-\mu}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx
\end{aligned}$$

令 $y=x-\mu$，則上式可寫成

$$\begin{aligned}\mathrm{E}X &= \int_{-\infty}^{\infty}\frac{y}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{y^2}{2\sigma^2}\right]dy+\mu\\ &= \mu\end{aligned}$$

(b) 在證明 $\mathrm{E}\left(X-\mu\right)^2=\sigma^2$ 之前，我們首先證明 $\Gamma\left(1/2\right)=\sqrt{\pi}$

由 Gamma 函數可知

$$\Gamma\left(\frac{1}{2}\right)=\int_0^{\infty}x^{\frac{1}{2}-1}e^{-x}dx$$

令 $x=z^2$, Gamma 函數可改寫成

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) &= \int_0^{\infty}x^{\frac{1}{2}-1}e^{-x}dx\\ &= \int_0^{\infty}\frac{1}{z}e^{-z^2}2zdz\\ &= 2\int_0^{\infty}e^{-z^2}dz\end{aligned}$$

故可求得

$$\begin{aligned}\left[\Gamma\left(\frac{1}{2}\right)\right]^2 &= 4\int_0^{\infty}e^{-x^2}dx\int_0^{\infty}e^{-y^2}dy\\ &= 4\int_0^{\infty}\int_0^{\infty}e^{-\left(x^2+y^2\right)}dxdy\\ &= 4\int_0^{\frac{\pi}{2}}\int_0^{\infty}re^{-\left(r^2\right)}drd\theta\\ &= 4\int_0^{\frac{\pi}{2}}\frac{-1}{2}e^{-r^2}\bigg|_0^{\infty}d\theta\\ &= 4\int_0^{\frac{\pi}{2}}\frac{1}{2}d\theta\\ &= 4\times\frac{\pi}{4}=\pi\end{aligned}$$

所以 $\Gamma\left(1/2\right)=\sqrt{\pi}$，又

$$\begin{aligned}\mathrm{E}\left(X-\mu\right)^2 &= \int_{-\infty}^{\infty}\frac{\left(X-\mu\right)^2}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right]dx\\ &= \frac{\sqrt{2}}{\sqrt{\pi}}\int_0^{\infty}\frac{\left(X-\mu\right)^2}{\sigma}\exp\left[-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right]dx\end{aligned}$$

$$
\begin{aligned}
&= \frac{\sqrt{2}\sigma^2}{\sqrt{\pi}}\int_0^\infty \frac{(X-\mu)^2}{\sigma^2}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]d\frac{x-\mu}{\sigma} \\
&= \frac{4\sigma^2}{\sqrt{\pi}}\int_0^\infty \frac{(X-\mu)^2}{2\sigma^2}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]d\frac{x-\mu}{\sqrt{2}\sigma} \\
&= \frac{4\sigma^2}{\sqrt{\pi}}\int_0^\infty y^2 e^{-y^2}dy = \frac{4\sigma^2}{\sqrt{\pi}}\int_0^\infty z e^{-z}\frac{dz}{2\sqrt{z}} \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\int_0^\infty z^{\frac{1}{2}}e^{-z}dz = \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{3}{2}\right) \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\times\frac{1}{2}\times\Gamma\left(\frac{1}{2}\right) = \sigma^2
\end{aligned}
$$

(c)

$$
\begin{aligned}
M(t) &= \int_{-\infty}^\infty \exp[tx]\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{x^2-2\mu x+\mu^2-2tx\sigma^2}{2\sigma^2}\right]dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{x^2-2\left(\mu+\sigma^2 t\right)x+\mu^2}{2\sigma^2}\right]dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{\left[x-\left(\mu+\sigma^2 t\right)\right]^2+\mu^2-\left(\mu+\sigma^2 t\right)^2}{2\sigma^2}\right]dx \\
&= \exp\left[\frac{2\mu\sigma^2 t+\sigma^4 t^2}{2\sigma^2}\right]\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{\left[x-\left(\mu+\sigma^2 t\right)\right]^2}{2\sigma^2}\right\}dx \\
&= \exp\left[\mu t+\frac{\sigma^2 t^2}{2}\right]
\end{aligned}
$$

(d)

$$
\begin{aligned}
M_Z(t) &= \mathrm{E}\{\exp[tz]\} \\
&= \mathrm{E}\left\{\exp\left[t\frac{x-\mu}{\sigma}\right]\right\} \\
&= \mathrm{E}\left\{\exp\left[-\mu\frac{t}{\sigma}+x\frac{t}{\sigma}\right]\right\} \\
&= \exp\left[-\mu\frac{t}{\sigma}+\mu\frac{t}{\sigma}+\frac{\sigma^2\frac{t^2}{\sigma^2}}{2}\right] \\
&= \exp\left[\frac{t^2}{2}\right]
\end{aligned}
$$

所以，由動差生成函數可知 $Z \sim \mathcal{N}(0,1)$

(e)

$$
\begin{aligned}
M_Y(t) &= \mathrm{E}\{\exp[ty]\} = \mathrm{E}\left\{\exp\left[t\sum_{i=1}^{n} a_i x_i\right]\right\} \\
&= \mathrm{E}\{\exp[ta_1x_1]\exp[ta_2x_2]\cdots\exp[ta_nx_n]\} \\
&= \mathrm{E}\{\exp[ta_1x_1]\}\,\mathrm{E}\{\exp[ta_2x_2]\}\cdots\mathrm{E}\{\exp[ta_nx_n]\} \\
&= \exp\left[a_1\mu_1 t + \frac{a_1^2\sigma_1^2t^2}{2}\right]\exp\left[a_2\mu_2 t + \frac{a_2^2\sigma_2^2t^2}{2}\right]\cdots\exp\left[a_n\mu_n t + \frac{a_n^2\sigma_n^2t^2}{2}\right] \\
&= \exp\left[\sum_{i=1}^{n} a_i\mu_i t + \frac{\sum_{i=1}^{n} a_i^2\sigma_i^2t^2}{2}\right]
\end{aligned}
$$

所以，$Y \sim \mathcal{N}\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$

例 **4–26** 強生製藥廠製造一種抗胃癌新藥，假設每顆藥丸重量符合常態分配 $\mathcal{N}(0.3,\ 0.01^2)$，單位：公克。因為此藥含有某成份比例的珍貴藥材與劇毒，製造產品需要較嚴格的品管。強生製藥廠認為此種抗癌新藥的重量應為 0.3 ± 0.02 間才安全。試問此種抗癌新藥產品不被接受的比率為多少？

解：

令 r.v. X 表示每顆藥丸重量，所以 $X \sim \mathcal{N}(0.3,\ 0.01^2)$

安全重量範圍為 $0.3 \pm 0.02 = (0.28, 0.32)$，因此

$$
\begin{aligned}
\mathrm{P}(0.28 \le X \le 0.32) &= \mathrm{P}\left(\frac{0.28-0.3}{0.01} \le \frac{x-\mu}{\sigma} \le \frac{0.32-0.3}{0.01}\right) \\
&= \mathrm{P}(-2 \le Z \le 2) \\
&= 1 - 2\mathrm{P}(X \ge 2) = 0.9544
\end{aligned}
$$

所以，此種產品約有 4.56% 的比率不被接受。

例 **4–27** 某校招考新生 500 名，其報考學生人數為 3160 人，入學考試的滿分為 400 分。若計算得該項考試的平均分數為 126 分, 標準差為 64 分，且全部考試成績近似於常態分配，試問

(a) 考試成績為 300 分者約為第幾名？

(b) 最低錄取分數為多少分？

解：

(a) 令 r.v. X 表示考生成績，則 $X \sim \mathcal{N}(126,\ 64^2)$

$$\mathrm{P}\left(X \geq 300\right) = \mathrm{P}\left(Z \geq \frac{300-126}{64}\right) = \mathrm{P}\left(Z \geq 2.72\right) = 0.0033$$

因 $3160 \times 0.0033 = 10.428$，故成績 300 分的考生在此考試中約為第 10 名或第 11 名。

(b) 設最低錄取分數為 x_0，則

$$\mathrm{P}\left(X \geq x_0\right) = \mathrm{P}\left(Z \geq \frac{x_0-126}{64}\right) = \frac{500}{3160} = 0.1582$$

查表得 $\frac{x_0-126}{64} \approx 1.002$，故 $x_0 \approx 126 + 1.002 \times 64 = 190.13$
最低錄取分數大約為 190 分。

例 4–28 Let the k^{th} standardized moment of X be $\mathrm{E}\left\{\left(X-\mu\right)/\sigma\right\}^k$, where $\mu = \mathrm{E}X$ and σ is standard deviation of X.

(a) Explain in word the meaning of the first four standardized moments.

(b) What are the first four standardized moment for a normal random variable.

解：

(a) i. 為將隨機變數 X 經標準化轉換後之期望值。
ii. 為將隨機變數 X 經標準化轉換後之變異數。
iii. 為將隨機變數 X 經標準化轉換後之偏態係數，$EZ^3 > 0$ 則為右偏，$EZ^3 < 0$ 則為左偏，$EZ^3 = 0$ 則對稱分配。
iv. 為將隨機變數 X 經標準化轉換後之峰態係數，$EZ^4 > 3$ 則為高峽峰，$EZ^3 < 3$ 則為低闊峰，$EZ^4 = 3$ 則常態峰。

(b) 由於 $Z \sim \mathcal{N}(0,1)$，所以

$$\begin{aligned} M_Z\left(t\right) &= e^{\frac{t^2}{2}} \\ &= 1 + \frac{t^2}{2} + \frac{1}{2!} \times \left(\frac{t^2}{2}\right)^2 + \frac{1}{3!} \times \left(\frac{t^2}{2}\right)^3 + \frac{1}{4!} \times \left(\frac{t^2}{2}\right)^4 + \cdots \\ &= 1 + \frac{1}{2!}t^2 + \frac{3}{4!}t^4 + \cdots \end{aligned}$$

由 **4.9 性質 1** (第57頁) 可知，

$$M'\left(0\right) = 0,$$

$$
\begin{aligned}
M''(0) &= 1, \\
M^{(3)}(0) &= 0, \\
M^{(4)}(0) &= 3
\end{aligned}
$$

例 **4–29** Suppose that X is a random variable for which $\mathrm{E}X = 1$, $\mathrm{E}X^2 = 4$ and $\mathrm{E}X^3 = 10$. Find the value of the third central moment of X.

解：

由題意可知 $\sigma^2 = \mathrm{E}X^2 - \mu^2 = 4 - 1 = 3$，故

$$
\begin{aligned}
\mathrm{E}\left(\frac{X-\mu}{\sigma}\right)^3 &= \mathrm{E}\left(\frac{X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3}{\sigma^3}\right) \\
&= \frac{\mathrm{E}\left(X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3\right)}{\sigma^3} \\
&= \frac{\mathrm{E}X^3 - 3\mathrm{E}X^2 + 3\mathrm{E}X - 1}{3\sqrt{3}} \\
&= \frac{10 - 3 \times 4 + 3 - 1}{3\sqrt{3}} \\
&= 0
\end{aligned}
$$

例 **4–30** 若 $\ln X \sim \mathcal{N}(u, \sigma^2)$，試證明 $P(a < X < b) = \Phi\left(\frac{\ln b - \mu}{\sigma}\right) - \Phi\left(\frac{\ln a - \mu}{\sigma}\right)$。

解：

$$
\begin{aligned}
P(a < X < b) &= P(\ln a < \ln X < \ln b) \\
&= P\left(\frac{\ln a - \mu}{\sigma} < \frac{\ln X - \mu}{\sigma} < \frac{\ln b - \mu}{\sigma}\right) \\
&= P\left(\frac{\ln a - \mu}{\sigma} < Z < \frac{\ln b - \mu}{\sigma}\right) \\
&= \Phi\left(\frac{\ln b - \mu}{\sigma}\right) - \Phi\left(\frac{\ln a - \mu}{\sigma}\right)
\end{aligned}
$$

5. **卡方分配**（Chi-Square Distribution）：X 有卡方分配以 $X \sim \chi^2(r)$ 表示之，其中 r 稱為自由度（degree of freedom, 通常以 $d.f.$ 表示之），其 p.d.f. 為

$$
f(x) = \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}, \ 0 \le x < \infty
$$

其中 $\Gamma(r) = (r-1)!$, r 為一正整數。

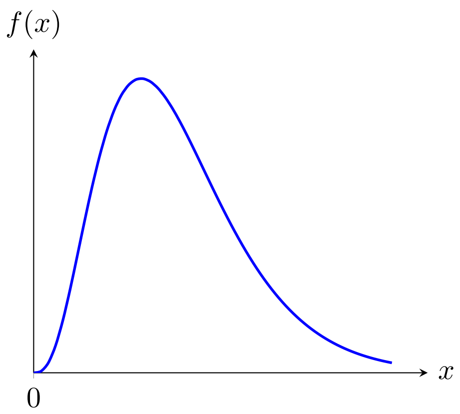


圖 4.4: 卡方分配圖

特性

(a) 定義域在 $[0,\ \infty)$

(b) 不對稱，為一右偏分配

(c) 自由度 r 不同，其圖形有所不同

(d) 為 $\alpha = \frac{r}{2}$ 及 $\theta = 2$ 之 Gamma 分配

性質

(a) $\mathrm{E}(X) = r$

(b) $\mathrm{Var}(X) = 2r$

(c) 若 $X \sim \chi^2(r_1),\ Y \sim \chi^2(r_2)$ 且 X 與 Y 相互獨立，則 $X + Y \sim \chi^2(r_1 + r_2)$

(d) 若隨機變數 $Z \sim \mathcal{N}(0,\ 1)$，則 $Z^2 \sim \chi^2(1)$

(e) 設 $X_i \overset{iid}{\sim} \mathcal{N}\left(\mu, \sigma^2\right)$，則

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

其中 $S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$ 且 $\overline{X}$ 與 S^2 獨立。

6. **Beta 分配**（Beta Distribution）：X 有 Beta 分配以 $X \sim Beta(\alpha, \beta)$ 表示之，其 p.d.f. 為

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1},\ 0<x<1, \alpha>0, \beta>0$$

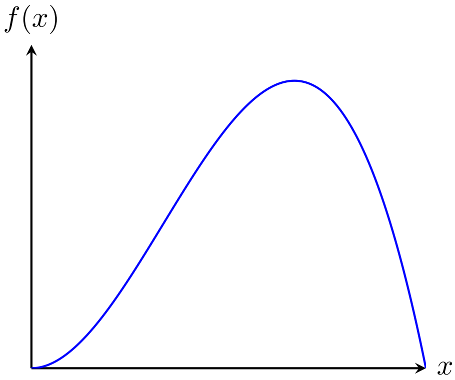


圖 4.5: Beta 分配圖

定理 4.3.

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx, \alpha>0, \beta>0$$

證明. 由分部積分法可知

$$\begin{aligned}
\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx &= -\frac{1}{\beta}x^{\alpha}(1-x)^{\beta}\Big|_0^1 + \int_0^1 \frac{\alpha}{\beta}(1-x)^{\beta}x^{\alpha-2}dx \\
&= \frac{\alpha-1}{\beta}\int_0^1 (1-x)^{\beta}x^{\alpha-2}dx \\
&= \frac{(\alpha-1)(\alpha-2)}{\beta(\beta+1)}\int_0^1 (1-x)^{\beta+1}x^{\alpha-3}dx \\
&\vdots \\
&= \frac{(\alpha-1)\cdots 2\times 1}{\beta(\beta+1)\cdots(\beta+\alpha-3)(\beta+\alpha-2)}\int_0^1 (1-x)^{(\alpha+\beta-2)}x^0 dx \\
&= \frac{(\alpha-1)\cdots 2\times 1}{\beta(\beta+1)\cdots(\beta+\alpha-2)(\beta+\alpha-2)}\frac{1}{\beta+\alpha-1} \\
&= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}
\end{aligned}$$

特性

(a) 定義域在 $(0,1)$

(b) α, β 不同，其圖形有所不同

性質

(a) $\mathrm{E}X = \alpha / (\alpha+\beta)$

(b) $\mathrm{Var}(X) = \alpha\beta / \left[(\alpha+\beta)^2 (\alpha+\beta+1)\right]$

(c) $\int_0^p \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx = \sum_{x=\alpha}^{\alpha+\beta-1} \binom{\alpha+\beta-1}{x} p^x (1-p)^{\alpha+\beta-1-x}$

證明.

(a)

$$\begin{aligned}
\mathrm{E}X &= \int_0^1 x \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+1)-1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)} \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

(b)

$$\begin{aligned}
\mathrm{E}X^2 &= \int_0^1 x^2 \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+2)-1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+2+\beta)} \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\end{aligned}$$

所以

$$\mathrm{Var}(X) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2$$

$$= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

(c) 由部分積分法可求得

$$\begin{aligned}
&\int_0^p \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \\
=\ & \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\left[-\frac{p^{\alpha-1}(1-p)^{\beta}}{\beta}+\frac{\alpha-1}{\beta}\int_0^p x^{\alpha-2}(1-x)^{\beta}\,dx\right] \\
=\ & \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\left[-\frac{p^{\alpha-1}(1-p)^{\beta}}{\beta}-\frac{(\alpha-1)p^{\alpha-2}(1-p)^{\beta+1}}{\beta(\beta+1)}\right. \\
& \left.+\frac{(\alpha-1)(\alpha-2)}{\beta(\beta+1)}\int_0^p x^{\alpha-3}(1-x)^{\beta+1}\,dx\right] \\
=\ & \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\left[-\frac{p^{\alpha-1}(1-p)^{\beta}}{\beta}-\frac{(\alpha-1)p^{\alpha-2}(1-p)^{\beta+1}}{\beta(\beta+1)}-\cdots\right. \\
& \left.-\frac{(\alpha-1)\cdots 2\times p^1\times(1-p)^{\alpha+\beta-2}}{\beta(\beta+1)\cdots(\beta+\alpha-2)}+\frac{(\alpha-1)\times\cdots\times 1}{\beta(\beta+1)\cdots(\beta+\alpha-2)}\int_0^p x^0(1-x)^{\alpha+\beta-2}\,dx\right] \\
=\ & 1-\binom{\alpha+\beta-1}{0}p^0(1-p)^{\alpha+\beta-1}-\binom{\alpha+\beta-1}{\alpha-1}p^{\alpha-1}(1-p)^{\beta} \\
& -\binom{\alpha+\beta-1}{\alpha-2}p^{\alpha-2}(1-p)^{\beta+1}-\cdots-\binom{\alpha+\beta-1}{1}p^1(1-p)^{\alpha+\beta-2} \\
=\ & 1-\sum_{x=0}^{\alpha-1}\binom{\alpha+\beta-1}{x}p^x(1-p)^{\alpha+\beta-1-x} \\
=\ & \sum_{x=\alpha}^{\alpha+\beta-1}\binom{\alpha+\beta-1}{x}p^x(1-p)^{\alpha+\beta-1-x}
\end{aligned}$$

例 **4–31** A quality control plans an assemble line involves sampling n finishing items per day and counting the number of defectives. If p denote the probability of observing a defective then Y has a binomial distribution, assuming that a large number of items are produced by the line. But p varies form day to day and is assumed to have a uniform distribution on the interval from 0 to 1.

1. Find the unconditional probability mass function of Y.
2. Find the expected value and variance of Y.

解：

1. $Y|P \sim B(n,p)$, $P \sim U(0,1)$

$$\begin{aligned} h(p,n) &= f(y|p)\,g(p) = \binom{n}{y} p^y (1-p)^{n-y} \times 1 \\ &= \binom{n}{y} p^y (1-p)^{n-y},\ y = 0,1,2,\ldots,n;\ 0<p<1 \end{aligned}$$

所以 Y 的邊際機率密度函數為

$$\begin{aligned} h_1(y) &= \int_0^1 \binom{n}{y} p^y (1-p)^{n-y}\,dp \\ &= \binom{n}{y} \int_0^1 p^{(y+1)-1} (1-p)^{(n-y+1)-1}\,dp \\ &= \frac{n!}{y!\,(n-y)!} \times \frac{\Gamma(y+1)\,\Gamma(n-y+1)}{\Gamma(y+1+n-y+1)} \\ &= \frac{n!}{y!\,(n-y)!} \times \frac{y!\,(n-y)!}{(n+1)!} \\ &= \frac{1}{n+1},\ y = 0,1,2,\ldots,n \end{aligned}$$

2.

$$\begin{aligned} \mathrm{E}Y &= \sum_{y=0}^{n} \frac{y}{n+1} = \frac{1}{n+1} \times (0+1+2+\cdots+n) \\ &= \frac{1}{n+1} \times \frac{n+1}{2} \\ &= \frac{1}{2} \end{aligned}$$

又

$$\begin{aligned} \mathrm{E}Y^2 &= \sum_{y=0}^{n} \frac{y^2}{n+1} = \frac{1}{n+1} \times \left(0^2+1^2+2^2+\cdots+n^2\right) \\ &= \frac{1}{n+1} \times \frac{n(n+1)(2n+1)}{6} \\ &= \frac{n(2n+1)}{6} \end{aligned}$$

所以

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{E}Y^2 - (\mathrm{E}Y)^2 \\ &= \frac{n(2n+1)}{6} - \frac{1}{4} \\ &= \frac{4n^2+2n-3}{12} \end{aligned}$$

4.3 混合分配

若隨機變數之機率分配為由連續型及離散型分配所共同組成，則稱該機率分配為混合分配 (Mixed Distribution)，即該分配在之 c.d.f. 在特定點並不連續。茲以下例說明之：

例 **4–32** 投擲公正銅板一枚，若銅板出現正面則該玩家可獲得 2 元，反之若銅板出現反面，則該玩家必須由 0 至 1 中隨機選出一數字並獲得該金額。若 X 表該玩家所獲得之金額，試求 X 之 c.d.f. 及變異數。

解：

當銅板出現反面時，則所獲得之金額服從 $U(0,1)$，所以

$$\mathrm{P}(X=x)=\mathrm{P}(\text{銅板出現反面})\times 1=\frac{1}{2}$$

可求得

$$F(x)=\frac{x}{2},\, 0<x<1$$

又當 $1\leq x<2$ 時，$\mathrm{P}(X=x)=0$，所以

$$F(x)=\frac{1}{2},\, 1\leq x<2$$

最後, 當銅板出現正面時，可求得

$$\mathrm{P}(X=2)=\mathrm{P}(\text{銅板出現正面})=\frac{1}{2}$$

因此，

$$F(x)=1, x=2$$

綜合以上結論，可得 X 之 p.d.f. 及 cdf 分別為

$$f\left(x\right)=\begin{cases}\frac{1}{2}, & 0<x<1\\ \frac{1}{2}, & x=2\\ 0, & \text{其他}\end{cases}$$

和

$$F(x)=\begin{cases}0, & x\le 0\\ \frac{x}{2}, & 0<x<1\\ \frac{1}{2}, & 1\le x<2\\ 1, & 2\le x\end{cases}$$

又

$$\mathrm{E}X=\int_0^1\frac{x}{2}dx+2\times\frac{1}{2}=\frac{1}{4}+1=\frac{5}{4},\ \mathrm{E}X^2=\int_0^1\frac{x^2}{2}dx+2^2\times\frac{1}{2}=\frac{1}{6}+2=\frac{13}{6}$$

因此,

$$\mathrm{Var}(X)=\mathrm{E}X^2-[\mathrm{E}X]^2=\frac{13}{6}-\frac{25}{16}=\frac{29}{48}$$

例 **4–33** 假設高速公路閘道管制器之紅燈及綠燈之時間長度分別為 60 秒及 40 秒，若 X 表駕駛人在紅燈期間到達等候綠燈通行所必須等候之時間長度，試求平均等候時間及變異數為何？

解：

$$f(x)=\begin{cases}\frac{40}{100}, & x=0\\ \frac{1}{100}, & 0<x<60\\ 0, & \text{其他}\end{cases}$$

所以

$$\mathrm{E}X=0\times\frac{40}{100}+\int_0^{60}\frac{x}{100}dx=18,$$

$$\mathrm{Var}(X)=\mathrm{E}X^2-[\mathrm{E}X]^2=0^2\times\frac{40}{100}+\int_0^{60}\frac{x^2}{100}dx-18^2=720-324=396$$

4.4 變數變換

定理 4.4. X 為一隨機變數，$F(x)$ 為隨機變數之分配函數。令 $Y=F(X)$，若 $F(x)$ 為一嚴格單調遞增函數，則

$$Y\sim U(0,1)$$

證明.

$$G(y) = \mathrm{P}(Y \le y) = \mathrm{P}(F(X) \le y)$$

由於 $F(x)$ 為一嚴格單調函數，

$$\begin{aligned} \mathrm{P}(F(X) \le y) &= \mathrm{P}(F^{-1}(F(X)) \le F^{-1}(y)) = \mathrm{P}(X \le F^{-1}(y)) = F(F^{-1}(y)) \\ &= y \end{aligned}$$

所以,

$$G(y) = y,\ \text{即 } G'(y) = g(y) = 1$$

故可推得 $Y \sim U(0,1)$

定理 4.5. X 為一隨機變數，$a \le x \le b,\ F(x)$ 為隨機變數之分配函數且 $F(x)$ 為一嚴格單調遞增函數。令 $Y = \mu(X)$ 為一對一之函數，若存在一反函數使得 $X = v(Y)$，則

$$g(y) = f(v(y))|v'(y)|$$

證明.

$$\begin{aligned} G(y) &= \mathrm{P}(Y \le y) = \mathrm{P}(\mu(X) \le y) = \mathrm{P}(X \le v(y)) \\ &= \int_{-\infty}^{v(y)} f(z)dz,\ \mu(a) \le y \le \mu(b) \end{aligned}$$

由微積分基本定理可知

$$G'(y) \ = \ g(y) = f[v(y)]v'(y)$$

又機率密度函數必須為正數，所以

$$g(y)=f[v(y)]|v'(y)|,\ \mu(a)\le y\le \mu(b)$$

例 **4–34** 假設 $X\sim\mathcal{N}(0,1)$, $-\infty\le z\le\infty$,

1. 令 $Y=e^X$，求 V 之 p.d.f.
2. 令 $Y=X^2$，求 V 之 p.d.f.

解：

1. 因為 $Y=e^X$，所以 $x=v(y)=\ln y$，故

$$v'(y)=\frac{dv(y)}{dy}=\frac{1}{y}$$

可推得

$$\begin{aligned} g(y) &= \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{(\ln y)^2}{2}\right]\frac{1}{y} \\ &= \frac{1}{y}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{(\ln y)^2}{2}\right],\ 0\le y\le\infty \end{aligned}$$

2.

$$\begin{aligned} G(v) &= \mathrm{P}(Y\le y)=\mathrm{P}\left(X^2\le y\right) \\ &= \mathrm{P}(-\sqrt{y}\le X\le\sqrt{y}) \\ &= \int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2}\right]dx,\ 0\le x\le\infty \end{aligned}$$

所以，由微積分基本定理得知

$$\begin{aligned} G'(y) &= g(y)=\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\sqrt{y}^2}{2}\right]\times\frac{1}{2\sqrt{y}}-\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{\left(-\sqrt{y}\right)^2}{2}\right]\times\frac{-1}{2\sqrt{y}} \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{y}}\exp\left[-\frac{y}{2}\right] \\ &= \frac{1}{\Gamma\left(\frac{1}{2}\right)2^{\frac{1}{2}}}y^{\frac{1}{2}-1}\exp\left[-\frac{y}{2}\right],\ 0\le y\le\infty \end{aligned}$$

故 $Y \sim \chi^2(1)$

例 **4–35** 令 Y 的動差母函數為 $M(t) = (1-4t)^{-1/2}$，試求 k 值使得 $\mathrm{P}(X < k) = 0.95$。

解：

$Y \sim Gamma(\frac{1}{2}, 4)$，所以

$$\mathrm{P}\left(Y < k\right) = \int_0^k \frac{1}{\Gamma\left(\frac{1}{2}\right) 4^{\frac{1}{2}}} y^{-\frac{1}{2}} e^{-\frac{y}{4}} dy$$

令 $y = 2x$，則 $dy = 2dx$, 上式可寫成

$$\begin{aligned}
\int_0^k \frac{1}{\Gamma\left(\frac{1}{2}\right) 4^{\frac{1}{2}}} y^{-\frac{1}{2}} e^{-\frac{y}{4}} dy &= \int_0^{\frac{k}{2}} \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{1}{2}} 2^{\frac{1}{2}}} \left(2x\right)^{-\frac{1}{2}} e^{-\frac{2x}{4}} d\left(2x\right) \\
&= \int_0^{\frac{k}{2}} \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{1}{2}}} x^{-\frac{1}{2}} e^{-\frac{x}{2}} dx \\
&= 0.95
\end{aligned}$$

由於

$$\frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{1}{2}}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}$$

為卡方自由度 1 之 p.d.f.，故可得 $0.5k = \chi^2_{0.05}\left(1\right)$。又 $Z^2 \sim \chi^2(1)$ 且

$$\mathrm{P}\left(|Z| > z\right) = \mathrm{P}\left(Z^2 > z^2\right) = \mathrm{P}\left(\chi^2\left(1\right) > z^2\right) = 0.05$$

故可推得

$$\frac{k}{2} = \chi^2_{0.05}\left(1\right) = z^2_{0.025} = 1.96^2,\ \text{即} k = 7.6832$$

例 **4–36** Let the independent random variables X_1 and X_2 be $\mathcal{N}\left(0, 1\right)$ and $\chi^2\left(1\right)$, respectively. Let $Y_1 = X_1/\sqrt{X_2}$ and $Y_2 = X_2$.

1. Find the joint pdf of Y_1 and Y_2.
2. Determine the marginal pdf of Y_1.

解：

1. 由於 X_1 與 X_2 獨立，所以其聯合機率密度函數為

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_1^2}{2}\right) \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{1}{2}}} x_2^{-\frac{1}{2}} \exp\left(-\frac{x_2}{2}\right)$$

又

$$|J| = \left\| \begin{array}{cc} \frac{\partial X_1}{\partial Y_1} & \frac{\partial X_1}{\partial Y_2} \\ \frac{\partial X_2}{\partial Y_1} & \frac{\partial X_2}{\partial Y_2} \end{array} \right\| = \left\| \begin{array}{cc} \sqrt{Y_2} & -\frac{Y_1}{2\sqrt{Y_2}} \\ 0 & 1 \end{array} \right\| = \sqrt{Y_2}$$

所以

$$\begin{aligned} g(y_1, y_2) &= \frac{1}{\sqrt{2}\sqrt{\pi}} \exp\left(-\frac{y_1^2 y_2}{2}\right) \frac{1}{\sqrt{2}\sqrt{\pi}} y_2^{-\frac{1}{2}} \exp\left(-\frac{y_2}{2}\right) |J| \\ &= \frac{1}{2\pi} \exp\left[-\frac{y_2\left(1+y_1^2\right)}{2}\right], -\infty < y_1 < \infty, 0 < y_1 < \infty \end{aligned}$$

2.

$$\begin{aligned} g_1(y_1) &= \int_0^\infty \frac{1}{2\pi} \exp\left[-\frac{y_2\left(1+y_1^2\right)}{2}\right] dy_2 \\ &= \frac{1}{2\pi} \frac{-2}{1+y_1^2} \exp\left[-\frac{y_2\left(1+y_1^2\right)}{2}\right]\Bigg|_0^\infty \\ &= \frac{1}{2\pi}\left[0 + \frac{2}{1+y_1^2}\right] \\ &= \frac{1}{\pi\left(1+y_1^2\right)}, -\infty < y_1 < \infty \end{aligned}$$

例 **4–37** 給定 $X_i \overset{iid}{\sim} \mathcal{N}(0,1)$ ，試求算 $Y = X_1/|X_2|$ 的機率分配為何？

解：令 $W = X_2^2$，則 $Y = X_1/\sqrt{W}$。證明同上

例 **4–38** Let X_1, X_2 be random variables distributed as normal $\mathcal{N}\left(0, \sigma^2\right)$ and let $T = X_1^2 + X_2^2$.

1. Find the distribution of T.
2. Find the mean and variance of T.

解：

1. 令 $Y=\left(X_1^2+X_2^2\right)/\sigma^2$，可推得

$$Y=\left(\frac{X_1-0}{\sigma}\right)^2+\left(\frac{X_2-0}{\sigma}\right)^2\sim\chi^2\left(2\right).$$

因此，Y 之 p.d.f. 為

$$f\left(y\right)=\frac{1}{2}\exp\left(-\frac{y}{2}\right),\ y>0.$$

又 $T=\sigma^2Y$ 及 $|J|=\sigma^2$，因此，T 之 p.d.f. 為

$$g\left(t\right)=\frac{1}{2\sigma^2}\exp\left(-\frac{t}{2\sigma^2}\right),\ t>0$$

即 $T\sim\exp\left(2\sigma^2\right)$.

2. 因為 $T\sim\exp\left(2\sigma^2\right)$，故 $\mathrm{E}\left(T\right)=2\sigma^2$ 及$\mathrm{Var}\left(T\right)=4\sigma^4$

例 **4–39** Let $X_1,\ X_2,\ X_3$ be independent with X_i having density $f\left(x_i\right)=\exp\left(-x_i\right),\ x_i>0$. Let $U_1=X_1+X_2+X_3$, $U_2=X_2/U_1$ and $U_3=X_3/U_1$. Find the joint density of U_1, U_2, and U_3.

解：

$$\begin{aligned}X_2&=U_1U_2\\X_3&=U_1U_3\\X_1&=U_1-X_2-X_3=U_1\left(1-U_2-U_3\right)\end{aligned}$$

$$\begin{aligned}J&=\left\|\begin{array}{ccc}1-U_2-U_3&-U_1&-U_1\\U_2&U_1&0\\U_3&0&U_1\end{array}\right\|=\left|U_1^2\left(1-U_2-U_3\right)+U_1^2U_3+U_1^2U_2\right|\\&=U_1^2\end{aligned}$$

$$\begin{aligned}g\left(u_1,u_2,u_3\right)&=\exp\left[-u_1\left(1-u_2-u_3\right)\right]\times\exp\left[-u_1u_2\right]\times\exp\left[-u_1u_3\right]\times J\\&=u_1^2\exp\left(-u_1\right),0<u_1<\infty,0<u_2<1,0<u_3<1,0<u_2+u_3<1\end{aligned}$$

4.5 二元常態分配

假設隨機變數 $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$, $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ 且 X 與 Y 之相關係數為 ρ。若

1. 在給定 $X = x$ 下，$\mathrm{E}[Y|X = x]$ 為 x 之線性函數

2. 在給定 $X = x$ 下，Y 之變異數齊一

則

$$\mathrm{E}\left(Y|X\right) = \mu_y + \rho\frac{\sigma_y^2}{\sigma_y^2}\left(x - \mu_x\right) \text{ 及} \mathrm{Var}\left(Y|X\right) = \left(1 - \rho^2\right)\sigma_y^2$$

證明.

令 $\mathrm{E}\left(Y|X\right) = a + bx$，由條件一

$$\int_{y \in R_2} yf\left(y|x\right)dy = \int_{y \in R_2} y\frac{f\left(x, y\right)}{f_1\left(x\right)}dy = a + bx$$

所以

$$\int_{y \in R_2} yf\left(x, y\right)dy = \left(a + bx\right)f_1\left(x\right)$$

左右同時對 x 做積分得

$$\int_{x \in R_1}\int_{y \in R_2} yf\left(x, y\right)dydx = \int_{x \in R_1}\left(a + bx\right)f_1\left(x\right)dx$$

故得

$$\mu_y = a + b\mu_x$$

又

$$\int_{y \in R_2} xyf\left(x, y\right)dy = \left(ax + bx^2\right)f_1\left(x\right)$$

左右同時對 x 做積分得

$$\int_{x \in R_1}\int_{y \in R_2} xyf\left(x, y\right)dydx = \int_{x \in R_1}\left(a + bx\right)f_1\left(x\right)dx$$

可得

$$\mathrm{E}XY = \rho\sigma_x\sigma_y + \mu_x\mu_y = a\mu_x + b\left(\mu_x^2 + \sigma_x^2\right) = a\mu_x + bEX^2$$

又

$$\rho\sigma_x\sigma_y + \mu_x\mu_y = \left(\mu_y - b\mu_x\right)\mu_x + b\left(\mu_x^2 + \sigma_x^2\right)$$

所以

$$b = \frac{\rho\sigma_x\sigma_y}{\sigma_x^2} = \rho\frac{\sigma_y}{\sigma_x} \text{ 及} a = \mu_y - b\mu_x = \mu_y - \rho\frac{\sigma_y}{\sigma_x}\mu_x$$

因此

$$\begin{aligned}
\mathrm{E}\left(Y|X\right) &= a + bX = \mu_y - \rho\frac{\sigma_y}{\sigma_x}\mu_x + \rho\frac{\sigma_y}{\sigma_x}x \\
&= \mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)
\end{aligned}$$

又 Y 之條件變異數為

$$\begin{aligned}
\mathrm{Var}\left(Y|X\right) &= \int_{y\in R_2}\left[y - \mu_y - \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)\right]^2 f\left(y|x\right)dy \\
&= \int_{y\in R_2}\left[y - \mu_y - \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)\right]^2 f\left(x,y\right)\int_{x\in R_1} f_1\left(x\right)dxdy \\
&= \int_{y\in R_2}\int_{x\in R_1}\left[y - \mu_y - \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)\right]^2 f\left(x,y\right)f_1\left(x\right)dxdy \\
&= \mathrm{E}\left[y - \mu_y - \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)\right]^2 \\
&= \mathrm{E}\left[y - \mu_y\right]^2 + \rho^2\frac{\sigma_y^2}{\sigma_x^2}\left[x - \mu_x\right]^2 - 2\rho\frac{\sigma_y}{\sigma_x}\mathrm{E}\left(x - \mu_x\right)\left(y - \mu_y\right) \\
&= \sigma_y^2 + \rho^2\frac{\sigma_y^2}{\sigma_x^2}\times\sigma_x^2 - 2\rho\frac{\sigma_y}{\sigma_x}\times\rho\sigma_x\sigma_y \\
&= \left(1 - \rho^2\right)\sigma_y^2
\end{aligned}$$

由上述推導可知，在給定 $X = x$ 下，Y 之條件機率密度函數為

$$Y|X = x \sim \mathcal{N}\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right), \left(1 - \rho^2\right)\sigma_y^2\right)$$

所以

$$f\left(y|x\right) = \frac{1}{\sqrt{2\pi\left(1-\rho^2\right)\sigma_y^2}}\exp\left\{\frac{-\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}\left(x-\mu_x\right)\right]^2}{2\left(1-\rho^2\right)\sigma_y^2}\right\},$$
$$-\infty < x < \infty, -\infty < y < \infty$$

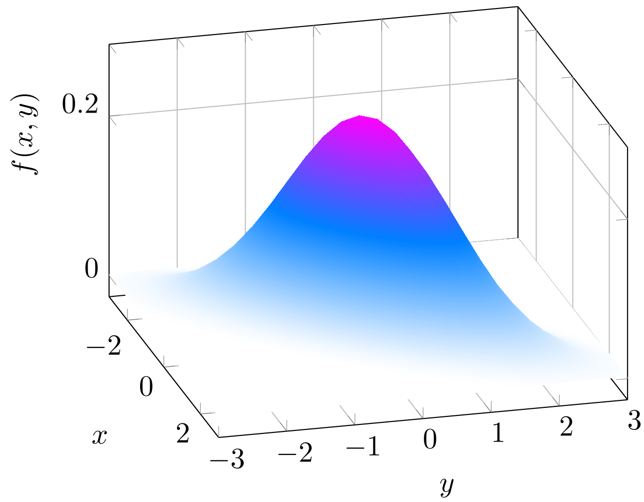


圖 4.6: 二元常態分配圖 ($\rho = 0.8$)

由於 $f(x,y) = f(y|x)f_1(x)$，所以 X 跟 Y 的聯合機率密度函數為

$$\begin{aligned} f\left(x,y\right) &= f\left(y|x\right)\times f_1\left(x\right) \\ &= \frac{1}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_y^2\sigma_x^2}}\exp\left\{\frac{-\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}\left(x-\mu_x\right)\right]^2}{2\left(1-\rho^2\right)\sigma_y^2}\right\}\exp\left\{\frac{-\left(x-\mu_x\right)^2}{2\sigma_x^2}\right\} \\ &= \frac{1}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_y^2\sigma_x^2}}\exp\left\{-\frac{\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}\left(x-\mu_x\right)\right]^2}{2\left(1-\rho^2\right)\sigma_y^2}-\frac{\left(x-\mu_x\right)^2}{2\sigma_x^2}\right\}, \end{aligned}$$

其中

$$\begin{aligned} & -\frac{\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}\left(x-\mu_x\right)\right]^2}{2\left(1-\rho^2\right)\sigma_y^2}-\frac{\left(x-\mu_x\right)^2}{2\sigma_x^2} \\ = & -\frac{\rho^2\sigma_y^2\left(x-\mu_x\right)^2+\sigma_x^2\left(y-\mu_y\right)^2-2\rho\sigma_x\sigma_y\left(x-\mu_x\right)\left(y-\mu_y\right)+\left(1-\rho^2\right)\sigma_y^2\left(x-\mu_x\right)^2}{2\left(1-\rho^2\right)\sigma_x^2\sigma_y^2} \end{aligned}$$

$$
\begin{aligned}
&= -\frac{\sigma_x^2\left(y-\mu_y\right)^2-2\rho\sigma_x\sigma_y\left(x-\mu_x\right)\left(y-\mu_y\right)+\sigma_y^2\left(x-\mu_x\right)^2}{2\left(1-\rho^2\right)\sigma_x^2\sigma_y^2} \\
&= -\frac{1}{2\left(1-\rho^2\right)}\left\{\frac{\sigma_x^2\left(y-\mu_y\right)^2}{\sigma_x^2\sigma_y^2}-2\frac{\rho\sigma_x\sigma_y\left(x-\mu_x\right)\left(y-\mu_y\right)}{\sigma_x^2\sigma_y^2}+\frac{\sigma_y^2\left(x-\mu_x\right)^2}{\sigma_x^2\sigma_y^2}\right\} \\
&= -\frac{1}{2\left(1-\rho^2\right)}\left\{\left(\frac{y-\mu_y}{\sigma_y}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)+\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right\}
\end{aligned}
$$

因此

$$
\begin{aligned}
f\left(x,y\right) &= \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_y}{\sigma_y}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)+\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]\right\}}{2\pi\sqrt{\left(1-\rho^2\right)\sigma_y^2\sigma_x^2}}, \\
&\quad -\infty<x<\infty,-\infty<y<\infty
\end{aligned}
$$

同理，我們亦可求得

$$
f\left(x|y\right) = \frac{1}{\sqrt{2\pi\left(1-\rho^2\right)\sigma_x^2}}\exp\left\{\frac{-\left[x-\mu_x-\rho\frac{\sigma_x}{\sigma_y}\left(y-\mu_y\right)\right]^2}{2\left(1-\rho^2\right)\sigma_x^2}\right\},-\infty<x<\infty,-\infty<y<\infty
$$

此外，亦可清楚看出當 $\rho=0,$

$$
\begin{aligned}
f\left(x,y\right) &= \frac{1}{2\pi\sqrt{\sigma_y^2\sigma_x^2}}\exp\left\{-\frac{1}{2}\left[\left(\frac{y-\mu_y}{\sigma_y}\right)^2+\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]\right\} \\
&= f_1\left(x\right)f_2\left(y\right),-\infty<x<\infty,-\infty<y<\infty,
\end{aligned}
$$

即 X 與 Y 相互獨立。

例 **4–40** 假設 $f(x,y)$ 為一二元常態分配之聯合機率密度函數，其中 $\mu_x=24.5,$ $\sigma_x^2=4.8^2,$ $\mu_y=-0.2,$ $\sigma_y^2=3.0^2$，及 $\rho=-0.032$，試求

1. $\mathrm{P}\left(1.3\leq Y\leq 5.8\right)$
2. $\mathrm{E}\left(Y|X=x\right)$
3. $\mathrm{E}\left(X|Y=y\right)$
4. $\mathrm{P}\left(1.3\leq Y\leq 5.8|X=18\right)$

解：

1.

$$\begin{aligned} \mathrm{P}(1.3 \le Y \le 5.8) &= \mathrm{P}\left(\frac{1.3+0.2}{3} \le Z \le \frac{5.8+0.2}{3}\right) \\ &= \mathrm{P}(0.5 \le Z \le 2) \\ &= 0.3085 - 0.0228 \\ &= 0.2857 \end{aligned}$$

2.

$$\begin{aligned} \mathrm{E}(Y|X=x) &= \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x) \\ &= -0.2 - .032 \times \frac{3}{4.8}(x-24.5) \\ &= 0.29 - 0.02x \end{aligned}$$

3.

$$\begin{aligned} \mathrm{E}(X|Y=y) &= \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y) \\ &= 24.5 - 0.032 \times \frac{4.8}{3}(y+0.2) \\ &= 24.4898 - 0.0512y \end{aligned}$$

4. 因為

$$\mathrm{E}(Y|X=18) = 0.29 - 0.02 \times 18 = -0.07,\ \mathrm{Var}(Y|X=18) = \left(1-0.32^2\right) \times 3^2 = 8.0784,$$

所以

$$\begin{aligned} \mathrm{P}(1.3 \le Y \le 5.8|X=18) &= \mathrm{P}\left(\frac{1.3+0.07}{\sqrt{8.0784}} < Z < \frac{5.8+0.07}{\sqrt{8.0784}}\right) \\ &= \mathrm{P}(0.48201 < Z < 2.0653) \\ &= 0.314875 \end{aligned}$$

例 4–41 Let $\mathbf{X} = (X_1, X_2, X_3)$ have a trivariate normal distribution with means 6 , 4 and 2. The covariance matrix is

$$\begin{bmatrix} 16 & 6 & 0 \\ 6 & 25 & 0 \\ 0 & 0 & 64 \end{bmatrix}.$$

Let $Y_1 = 2X_1 + 3X_2 + X_3 + 2$ and $Y_2 = 4X_1 + X_3 + 2$. Please find the joint probability density function of Y_1 and Y_2.

解：

$$\mu_{Y_1} = \mathrm{E}\left(2X_1 + 3X_2 + X_3 + 2\right) = 28 \text{ 及} \mu_{Y_8} = \mathrm{E}\left(4X_1 + X_3 + 2\right) = 28$$

又 Y_1, Y_2 的變異數及共變異數分別為

$$\mathrm{Var}\left(Y_1\right) = \mathrm{Var}\left(2X_1 + 3X_2 + X_3 + 2\right) = 20.616^2,\ \mathrm{Var}\left(Y_2\right) = \mathrm{Var}\left(4X_1 + X_3 + 2\right) = 17.889^2$$

與

$$\mathrm{Cov}\left(Y_1, Y_2\right) = 264\text{，所以} \rho_{Y_1 Y_2} = 0.7159$$

故二元常態分配之 p.d.f. 為

$$f\left(y_1, y_2\right) = \frac{\exp\left[-\frac{1}{2\times(1-0.7159^2)}\left\{\left(\frac{y_1-28}{20.616}\right)^2 - 2\times 0.7159 \times \left(\frac{y_1-28}{20.616}\right)\left(\frac{y_2-28}{17.889}\right) + \left(\frac{y_2-28}{17.889}\right)^2\right\}\right]}{2\pi \times \sqrt{20.616^2 \times 17.889^2 \times (1-0.7159^2)}}$$

例 **4–42** Suppose $X_1, X_2, \ldots, X_n$ are iid from a normal distribution $\mathcal{N}\left(\mu, \sigma^2\right)$.

1. Let $Y = \sum_{i=1}^n a_i X_i$ with $a_i \in R$, for all i. Calculate the mgf of Y.
2. Let $Z = \sum_{i=1}^n b_i X_i$ with $b_i \in R$, for all i. Find the necessary and sufficient conditions such that Y and Z are independent.
3. Compute the covariance of Y and Z. What kind of conclusion can you draw from this and previous parts.

解：

1. 由 **5.2.4** 性質 **e** 可知，

$$Y = \sum_{i=1}^n a_i X_i \sim \mathcal{N}\left(\sum_{i=1}^n a_i \mu, \sum_{i=1}^n a_i^2 \sigma^2\right), Z = \sum_{i=1}^n b_i X_i \sim \mathcal{N}\left(\sum_{i=1}^n b_i \mu, \sum_{i=1}^n b_i^2 \sigma^2\right)$$

所以

$$M_Y(t) = \exp\left(t\sum_{i=1}^{n} a_i\mu + \frac{t^2\sum_{i=1}^{n} a_i^2\sigma^2}{2}\right)$$

2.

$$Y + Z = \sum_{i=1}^{n}(a_i + b_i)X_i \sim \mathcal{N}\left(\sum_{i=1}^{n}(a_i+b_i)\mu, \sum_{i=1}^{n}\left(a_i^2 + b_i^2\right)\sigma^2\right)$$

若 Y 與 Z 獨立，則 Y 與 Z 之聯合動差生成函數為

$$\begin{aligned}
M(t,s) &= \mathrm{E}\left(e^{t\sum_{i=1}^{n} a_iX_i + s\sum_{i=1}^{n} b_iX_i}\right) \\
&= \mathrm{E}\left(e^{t\sum_{i=1}^{n} a_iX_i}\right)\mathrm{E}\left(e^{s\sum_{i=1}^{n} b_iX_i}\right) \\
&= \exp\left(t\sum_{i=1}^{n} a_i\mu + \frac{t^2\sum_{i=1}^{n} a_i^2\sigma^2}{2}\right)\exp\left(s\sum_{i=1}^{n} b_i\mu + \frac{s^2\sum_{i=1}^{n} b_i^2\sigma^2}{2}\right) \\
&= \exp\left(t\sum_{i=1}^{n} a_i\mu + s\sum_{i=1}^{n} b_i\mu + \frac{t^2\sum_{i=1}^{n} a_i^2\sigma^2}{2} + \frac{s^2\sum_{i=1}^{n} b_i^2\sigma^2}{2}\right) \\
&= \exp\left(t\sum_{i=1}^{n} a_i\mu + s\sum_{i=1}^{n} b_i\mu + \frac{t^2\sum_{i=1}^{n} a_i^2\sigma^2}{2} + \frac{s^2\sum_{i=1}^{n} b_i^2\sigma^2}{2}\right)
\end{aligned}$$

又 Y 與 Z 之聯合動差生成函數為

$$\begin{aligned}
\mathrm{E}\left(e^{(tY+sZ)}\right) &= \mathrm{E}\left(e^{\left(t\sum_{i=1}^{n} a_iX_i + s\sum_{i=1}^{n} b_iX_i\right)}\right) \\
&= \mathrm{E}\left(e^{\left(\sum_{i=1}^{n} ta_iX_i + \sum_{i=1}^{n} sb_iX_i\right)}\right) \\
&= \mathrm{E}\left(e^{\sum_{i=1}^{n}(ta_i+sb_i)X_i}\right) \\
&= \exp\left(\mu\sum_{i=1}^{n}(ta_i+sb_i) + \frac{\sum_{i=1}^{n}(ta_i+sb_i)^2\sigma^2}{2}\right)
\end{aligned}$$

所以 Y 與 Z 相獨立之充要條件為

$$\begin{aligned}
\sum_{i=1}^{n}(ta_i+sb_i)^2 &= \sum_{i=1}^{n}\left(t^2a_i^2 + s^2b_i^2 + 2a_ib_ist\right) \\
&= \sum_{i=1}^{n} t^2a_i^2 + \sum_{i=1}^{n} s^2b_i^2 + 2st\sum_{i=1}^{n} a_ib_i
\end{aligned}$$

即

$$\sum_{i=1}^{n} a_ib_i = 0$$

3.

$$\begin{aligned}
\operatorname{Cov}(Y,Z) &= \mathrm{E}YZ - \mathrm{E}Y \times \mathrm{E}Z \\
&= \mathrm{E}\left(\sum_{i=1}^{n} a_i X_i \times \sum_{i=1}^{n} b_i X_i\right) - \sum_{i=1}^{n} a_i \mu \times \sum_{i=1}^{n} b_i \mu \\
&= E\sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j X_i X_j\right) - \mu^2 \sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j\right) \\
&= \sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j \mathrm{E}X_i X_j\right) - \mu^2 \sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j\right) \\
&= \sum_{i=1}^{n}\left[a_i\left(b_i\sigma^2 + \mu^2 \sum_{j=1}^{n} b_j\right)\right] - \mu^2 \sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j\right) \\
&= \sum_{i=1}^{n} a_i b_i \sigma^2
\end{aligned}$$

由於 Y 與 Z 之聯合機率密度函數為二元常態分配，故當 Y 與 Z 之相關係數為 0 時，則 Y 與 Z 獨立。所以當

$$\sum_{i=1}^{n} a_i b_i = 0,$$

Y 與 Z 獨立。

例 **4–43** Let the 3×1 random vector $\mathbf{X} = (X_1, X_2, X_3)^{\mathrm{T}}$ follow a mutlivariate normal distribution,

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \mathbf{N}(\mu, \mathbf{\Sigma}),$$

where

$$\mu = \begin{bmatrix} 170 \\ 68 \\ 40 \end{bmatrix} \text{ and } \mathbf{\Sigma} = \begin{bmatrix} 400 & 64 & 128 \\ 64 & 16 & 0 \\ 128 & 0 & 256 \end{bmatrix}.$$

解：

1. Find the conditional distribution of X_1 given $X_2 = 72$.
2. Find the conditional distribution of X_1 given $X_2 = 72$ and $X_3 = 24$.

3.

$$
\begin{aligned}
\mathrm{E}\left(X_1|X_2=72\right) &= \mu_1+\rho\frac{\sigma_1}{\sigma_2}\left(x_2-\mu_2\right)\\
&= 170+\frac{64}{\sqrt{400\times 16}}\times\frac{20}{4}\times(72-68)\\
&= 186
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}\left(X_1|X_2=72\right) &= \sigma_1^2\left(1-\rho^2\right)=400\times\left(1-0.8^2\right)\\
&= 144
\end{aligned}
$$

故得 $X_1\sim\mathcal{N}(186,144)$

4.

$$
\begin{aligned}
\mathrm{E}\left(X_1|X_2=72,X_3=24\right) &= \mu_1+\frac{\begin{bmatrix}64 & 128\end{bmatrix}\begin{bmatrix}\frac{1}{16} & 0\\ 0 & \frac{1}{256}\end{bmatrix}\begin{bmatrix}-4\\ 16\end{bmatrix}}{20}\\
&= 170-\frac{8}{20}\\
&= 169.6
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}\left(X_1|X_2=72,X_3=24\right) &= 400-\begin{bmatrix}64 & 128\end{bmatrix}\begin{bmatrix}\frac{1}{16} & 0\\ 0 & \frac{1}{256}\end{bmatrix}\begin{bmatrix}64\\ 128\end{bmatrix}\\
&= 80
\end{aligned}
$$

故得 $X_1\sim\mathcal{N}(169.6,80)$

第 5 章　抽樣分配

統計資料主要的目的是利用統計量來估計母體中未知的參數，例如母體平均數及母體變異數，進而利用估計之結果來做預測或分析。

5.1　名詞解釋

1. **隨機樣本**：假設 $X_1, X_2, \ldots, X_n$ 為由母體 $f(x)$ 中抽出的 n 個隨機變數，若滿足

   (a) $X_1, X_2, \ldots, X_n$ 皆為獨立

   (b) $X_1, X_2, \ldots, X_n$ 的機率密度函數皆為 $f(x)$

   則稱 $X_1, X_2, \ldots, X_n$ 為抽自母體 $f(x)$ 的一組隨機樣本。

2. **統計量**：為隨機變數的函數，**並不包含未知的母體參數**，且本身亦為一隨機變數，其主要目的是用來估計母體參數。例如樣本平均數 $\overline{X} = \sum_{i=1}^{n} X_i/n$ 及樣本變異數 $S^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2/(n-1)$ 皆為統計量。

3. **抽樣分配**：為統計量的機率密度函數。

5.2　樣本平均數之抽樣分配

若 X_1, X_2, … X_n 為抽自母體 $f(x)$ 的一組樣本數為 n 的隨機樣本，令 $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$，則

$$\mu_{\overline{X}_n} = \mathrm{E}\left(\overline{X}_n\right) = \mu \text{ 及 } \sigma^2_{\overline{X}_n} = \mathrm{Var}\left(\overline{X}_n\right) = \frac{\sigma^2}{n}$$

證明.

1.

$$\mathrm{E}\left(\overline{X}\right)=\frac{1}{n}\mathrm{E}\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}X_i=\frac{1}{n}\sum_{i=1}^{n}\mu=\mu$$

2.

$$\begin{aligned}
\mathrm{Var}\left(\overline{X}\right) &= \mathrm{E}\left(\overline{X}-\mu\right)^2=\mathrm{E}\left[\frac{(X_1-\mu)+(X_2-\mu)+\cdots+(X_n-\mu)}{n}\right]^2 \\
&= \frac{1}{n^2}\Big\{\mathrm{E}\left(X_1-\mu\right)^2+\mathrm{E}\left(X_2-\mu\right)^2+\cdots \\
&\qquad +\mathrm{E}\left(X_n-\mu\right)^2+\sum_{i\neq j}\mathrm{E}\left(X_i-\mu\right)\left(X_j-\mu\right)\Big\} \\
&= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\mathrm{E}\left(X_i-\mu\right)^2+\sum_{i\neq j}\mathrm{E}\left(X_i-\mu\right)\left(X_j-\mu\right)\right\}
\end{aligned}$$

因為 $X_1, X_2, \ldots, X_n$ 相互獨立，所以 $\mathrm{Cov}(X_i, X_j)=0,\ \forall\, i\neq j$，故推得 $\mathrm{Var}\left(\overline{X}\right)=\frac{\sigma^2}{n}$。

例 **5–1** 假設 X_1, X_2 為抽自 $f(x)=0.25,\ X=0,\ 1,\ 2,\ 3$ 的一組隨機樣本，試求

1. 此一試驗之樣本空間
2. $\overline{X}$ 之抽樣分配
3. $\overline{X}$ 之平均數及變異數

解：

1. 此一試驗之樣本空間

$$S=\left\{\begin{array}{c}(0,0),(0,1),(0,2),(0,3),\\(1,0),(1,1),(1,2),(1,3),\\(2,0),(2,1),(2,2),(2,3),\\(3,0),(3,1),(3,2),(3,3)\end{array}\right\}$$

2. 由上可知 $\overline{X}$ 之樣本空間為 $S=\{\ 0,\ 0.5,\ 1,\ 1.5,\ 2,\ 2.5,\ 3\ \}$。因此，$\overline{X}$ 之抽樣分配為

$\overline{X}$	0	0.5	1	1.5	2	2.5	3
$f(\overline{x})$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

3. (a)

$$E(\overline{X}) = 0 \times \frac{1}{16} + 0.5 \times \frac{2}{16} + \cdots + 2.5 \times \frac{2}{16} + 3 \times \frac{1}{16} = 1.5$$

(b)

$$\mathrm{E}\left(\overline{X}^2\right) = 0^2 \times \frac{1}{16} + 0.5^2 \times \frac{2}{16} + \cdots + 2.5^2 \times \frac{2}{16} + 3^2 \times \frac{1}{16} = 2.875$$

所以，$\mathrm{Var}\left(\overline{X}\right) = 2.875 - 1.5^2 = 0.625$

(c)

$$\mu_{\overline{X}} = \frac{\sum_1^N X_i}{N} = \frac{1}{4}(0+1+2+3) = 1.5$$

$$\begin{aligned}
\sigma^2_{\overline{X}} &= \frac{\sum_{i=1}^N (X_i - \overline{X})^2}{N} \\
&= \frac{1}{4}\left\{(0-1.25)^2 + (1-1.25)^2 + (2-1.25)^2 + (3-1.25)^2\right\} \\
&= 1.25
\end{aligned}$$

所以，當試驗採放回方式實施時則

$$\mu_{\overline{X}} = \mu_X,\ \sigma_{\overline{X}} = \frac{\sigma^2_X}{n},$$

其中，N 為母體元素個數，n 為樣本個數。

例 5–2 以下是五名學生的統計學成績：30, 50, 60, 60, 80, 今自其中抽取兩名學生 X_1, X_2(不放回)，試求

1. 此一試驗之樣本空間
2. $\overline{X}$ 之抽樣分配
3. $\overline{X}$ 之平均數及變異數

解：

1. 由於本試驗之樣本空間為

$$S = \left\{ \begin{array}{c} (30, 50), (30, 60), (30, 60), (30, 80), \\ (50, 30), (50, 60), (50, 60), (50, 80), \\ (60, 30), (60, 50), (60, 60), (60, 80), \\ (60, 30), (60, 50), (60, 60), (60, 80), \\ (80, 30), (80, 50), (80, 60), (80, 60) \end{array} \right\}$$

2. 因此，$\overline{X}$ 之樣本空間為 $S = \{40, 45, 55, 60, 65, 70\}$。所以，$\overline{X}$ 之抽樣分配為

$\overline{X}$	40	45	55	60	65	70
$f(\overline{x})$	0.1	0.2	0.3	0.1	0.1	0.2

3.

$$\mathrm{E}\left(\overline{X}\right) = 40 \times 0.1 + 45 \times 0.2 + \cdots + 70 \times 0.2 = 56,$$

$$\mathrm{E}\left(\overline{X}^2\right) = 40^2 \times 0.1 + 45^2 \times 0.2 + \cdots + 70^2 \times 0.2 = 3235$$

所以 $\mathrm{Var}\left(\overline{X}\right) = 3235 - 56^2 = 99$

4.

$$\mu_X = \frac{\sum_1^N X_i}{N} = \frac{1}{5}(30 + 50 + 60 + 60 + 80) = 56$$

$$\begin{aligned} \sigma_X^2 &= \frac{\sum_{i=1}^N (X_i - \overline{X})^2}{N} \\ &= \frac{1}{5}\{(30-56)^2 + (50-56)^2 + (60-56)^2 + (60-56)^2 + (80-56)^2\} \\ &= 264 \end{aligned}$$

所以，當試驗採不放回方式實施時則

$$\mu_{\overline{X}} = \mu_X,\ \sigma_{\overline{X}} = \frac{N-n}{N-1} \times \frac{\sigma_X^2}{n}$$

其中，N 為母體元素個數，n 為樣本個數，$\frac{N-n}{N-1}$ 為有限母體校正數。

1. 單一母體平均數樣本平均數之抽樣分配

定理 5.1. $X_1,\ X_2,\ \ldots\ X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本，令 $\overline{X}=\sum_{i=1}^n X_i/n$，則

$$\mu_{\overline{X}_n}=\mathrm{E}\left(\overline{X}_n\right)=\mu \text{ 及 } \sigma^2_{\overline{X}_n}=\mathrm{Var}\left(\overline{X}_n\right)=\frac{\sigma^2}{n}$$

即 $\overline{X}_n\sim\mathcal{N}(\mu,\frac{\sigma^2}{n})$

證明.

因為 $M_X(t)=e^{\mu t+\frac{\sigma^2t^2}{2}}$，所以

$$\begin{aligned}
\mathrm{E}\left(e^{t\overline{X}}\right) &= \mathrm{E}\left\{\exp\left[t\frac{\sum_{i=1}^n X_i}{n}\right]\right\}\\
&= \mathrm{E}\left\{\exp\left[X_1\frac{t}{n}\right]\right\}\mathrm{E}\left\{\exp\left[X_2\frac{t}{n}\right]\right\}\cdots\mathrm{E}\left\{\exp\left[X_n\frac{t}{n}\right]\right\}\\
&= \exp\left[\mu\frac{t}{n}+\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\exp\left[\mu\frac{t}{n}+\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\cdots\exp\left[\mu\frac{t}{n}+\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\\
&= \exp\left\{\left[\mu\frac{t}{n}+\frac{\sigma^2}{2}\left(\frac{t}{n}\right)^2\right]\times n\right\}\\
&= \exp\left[\mu t+\frac{\sigma^2}{2}\frac{t^2}{n}\right]
\end{aligned}$$

故可推得 $\overline{X}\sim\mathcal{N}(\mu,\frac{\sigma^2}{n})$

2. 兩母體平均數樣本平均數差之抽樣分配

定理 5.2. $X_1,\ X_2,\ \ldots\ X_n$ 為抽自 $\mathcal{N}(\mu_x,\sigma_x^2)$ 的一組樣本數為 n_1 的隨機樣本，$Y_1,\ Y_2,\ \ldots\ Y_n$ 為抽自 $\mathcal{N}(\mu_y,\sigma_y^2)$ 的一組樣本數為 n_2 的隨機樣本。若 X 與 Y 為兩相獨立母體，令 $\overline{X}=\sum_{i=1}^n X_i/n$ 及 $\overline{Y}=\sum_{i=1}^n Y_i/n$，則

$$\mu_{\overline{X}_{n_1}-\overline{Y}_{n_2}}=\mathrm{E}\left(\overline{X}_{n_1}-\overline{Y}_{n_2}\right)=\mu_x-\mu_y \text{ 及 } \sigma^2_{\overline{X}_{n_1}-\overline{Y}_{n_2}}=\mathrm{Var}\left(\overline{X}_{n_1}-\overline{Y}_{n_2}\right)=\frac{\sigma_x^2}{n_1}+\frac{\sigma_y^2}{n_2}$$

即 $\overline{X}-\overline{Y}\sim\mathcal{N}\left(\mu_x-\mu_y,\sigma_x^2/n_1+\sigma_y^2/n_2\right)$

證明.

$$\mathrm{E}\left[e^{t\left(\overline{X}-\overline{Y}\right)}\right] = \mathrm{E}\left(e^{t\overline{X}}\right)\mathrm{E}\left(e^{-t\overline{Y}}\right)$$

$$
\begin{aligned}
&= \exp\left[\mu_x t + \frac{t^2}{2}\frac{\sigma_x^2}{n_1}\right]\exp\left[-u_y t + \frac{t^2}{2}\frac{\sigma_y^2}{n_2}\right] \\
&= \exp\left[(\mu_x - \mu_y)\,t + \frac{t^2}{2}\left(\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}\right)\right]
\end{aligned}
$$

故可推得 $\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \sigma_x^2/n_1 + \sigma_y^2/n_2\right)$

5.3 與常態分配有關的抽樣分配

5.3.1 t 分配

假設 $Z \sim \mathcal{N}(0,\ 1)$, $X \sim \chi^2(r)$ 且 X 與 Y 獨立，則統計量

$$
T = \frac{Z}{\sqrt{\dfrac{X}{r}}} \sim t(r)
$$

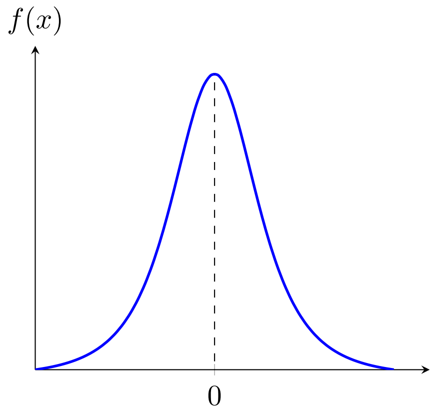


圖 5.1: t 分配圖

特性

1. 對 Y 軸對稱
2. 自由度 r 不同，其圖形不同
3. 與標準常態分配相似

4. 當自由度 $r > 30$ 時，近似標準常態分配

性質

1. $\mathrm{E}(T) = 0,\, r > 1$
2. $\mathrm{Var} = \dfrac{r}{r-2},\, r > 2$
3. 當自由度 $r = 1$ 時，為一科西分配。

證明.

1. 由 **4.7.1 性質 c** (第37頁) 可知

$$\begin{aligned}
\mathrm{E}\left(T\right) &= \mathrm{E}\left(\frac{Z}{\sqrt{X/r}}\right) = \mathrm{E}\left(Z\right) \times \mathrm{E}\left(\frac{\sqrt{r}}{\sqrt{X}}\right) \\
&= 0 \times \mathrm{E}\left(\frac{\sqrt{r}}{\sqrt{X}}\right) \\
&= 0
\end{aligned}$$

2.

$$\begin{aligned}
\mathrm{E}\left(T^2\right) &= \mathrm{E}\left(\frac{rZ^2}{X}\right) = r \times \mathrm{E}\left(Z^2\right) \times \mathrm{E}\left(\frac{1}{X}\right) \\
&= r \times \mathrm{E}\left(\frac{1}{X}\right)
\end{aligned}$$

又

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{X}\right) &= \int_0^\infty \frac{1}{x} \frac{1}{\Gamma\left(\alpha\right)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} dx \\
&= \int_0^\infty \frac{1}{\Gamma\left(\alpha\right)\theta^\alpha} x^{\alpha-2} e^{-\frac{x}{\theta}} dx \\
&= \frac{\theta^{\alpha-1}}{\Gamma\left(\alpha\right)\theta^\alpha} \int_0^\infty \left(\frac{x}{\theta}\right)^{\alpha-2} e^{-\frac{x}{\theta}} d\frac{x}{\theta} \\
&= \frac{1}{\Gamma\left(\alpha\right)\theta} \int_0^\infty y^{\alpha-2} e^{-y} dy \\
&= \frac{\Gamma\left(\alpha-1\right)}{\Gamma\left(\alpha\right)\theta} = \frac{1}{\left(\alpha-1\right)\theta}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\left(\frac{r}{2}-1\right)} \\
&= \frac{1}{r-2}
\end{aligned}
$$

所以

$$
\mathrm{Var}(T) = \frac{r}{r-2}
$$

3. 證明同 例 **4-36**。

定理 5.3. 若 $X_1,\ X_2,\ \ldots\ X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本，令

$$
\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \text{ 及 } S^2 = \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}
$$

則

1. $\overline{X}$ 與 S^2 獨立

2. $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

證明.

1.

$$
\begin{aligned}
f(x_1,x_2,\ldots,x_n) &= \frac{1}{(2\pi\sigma^2)^n}\exp\left[-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{(2\pi\sigma^2)^n}\exp\left[-\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{2\sigma^2}-\frac{n(\overline{x}-\mu)^2}{2\sigma^2}\right]
\end{aligned}
$$

令 $Y_1=\overline{X}$, $Y_2=X_2$, $Y_3=X_3$,..., $Y_n=X_n$, 得

$$
ny_1-y_2-y_3-\cdots-y_n=\sum_{i=1}^{n}x_i-x_2-x_3-\cdots-x_n=x_1
$$

及

$$
\begin{aligned}
J &= \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \cdots & \frac{\partial x_{n-1}}{\partial y_1} & \frac{\partial x_n}{\partial y_1} \\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_{n-1}}{\partial y_2} & \frac{\partial x_n}{\partial y_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial x_1}{\partial y_{n-1}} & \frac{\partial x_2}{\partial y_{n-1}} & \cdots & \frac{\partial x_{n-1}}{\partial y_{n-1}} & \frac{\partial x_n}{\partial y_{n-1}} \\ \frac{\partial x_1}{\partial y_n} & \frac{\partial x_2}{\partial y_n} & \cdots & \frac{\partial x_{n-1}}{\partial y_n} & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} n & 0 & \vdots & 0 & 0 \\ -1 & 1 & \vdots & 0 & 0 \\ \cdots & \cdots & \ddots & \cdots & \cdots \\ -1 & 0 & \vdots & 1 & 0 \\ -1 & 0 & \vdots & 0 & 1 \end{vmatrix} \\
&= n
\end{aligned}
$$

因此

$$
\begin{aligned}
& g\left(y_1, y_2, \ldots, y_n\right) \\
= & \frac{n}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}} \exp\left[-\frac{\left(ny_1 - y_2 - y_3 - \cdots - y_n - y_1\right)^2}{2\sigma^2} - \frac{\sum_{i=2}^{n}\left(y_i - y_1\right)^2}{2\sigma^2} - \frac{n\left(y_1 - \mu\right)^2}{2\sigma^2}\right] \\
= & \frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{n\left(y_1 - \mu\right)^2}{2\sigma^2}\right] \\
& \times \frac{\sqrt{n}}{\left(2\pi\sigma^2\right)^{\frac{n-1}{2}}} \exp\left[-\frac{\left(ny_1 - y_2 - y_3 - \cdots - y_n - y_1\right)^2}{2\sigma^2} - \frac{\sum_{i=2}^{n}\left(y_i - y_1\right)^2}{2\sigma^2}\right]
\end{aligned}
$$

所以在給定 $y_1 = \overline{x}$ 下，$y_1, y_2, \ldots, y_n$ 之條件機率密度函數為

$$
\begin{aligned}
& g\left(y_1, y_2, \ldots, y_n \middle| y_1 = \overline{x}\right) \\
= & \frac{\frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{n(\overline{x}-\mu)^2}{2\sigma^2}\right] \times \frac{\sqrt{n}}{\left(2\pi\sigma^2\right)^{\frac{n-1}{2}}} \exp\left[-\frac{(n\overline{x}-y_2-y_3-\cdots-y_n-\overline{x})^2}{2\sigma^2} - \frac{\sum_{i=2}^{n}(y_i-\overline{x})^2}{2\sigma^2}\right]}{\frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{n(\overline{x}-\mu)^2}{2\sigma^2}\right]} \\
= & \frac{\sqrt{n}}{\left(2\pi\sigma^2\right)^{\frac{n-1}{2}}} \exp\left[-\frac{\left(ny_1 - y_2 - y_3 - \cdots - y_n - \overline{x}\right)^2}{2\sigma^2} - \frac{\sum_{i=2}^{n}\left(y_i - \overline{x}\right)^2}{2\sigma^2}\right]
\end{aligned}
$$

由機率公理可知

$$
1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\sqrt{n}}{\left(2\pi\sigma^2\right)^{\frac{n-1}{2}}} e^{-\frac{q}{2\sigma^2}} dy_2 dy_3 \cdots dy_n
$$

又

$$
\mathrm{E}\left\{\exp\left[\frac{tQ}{\sigma^2}\right] \middle| y_1 = \overline{x}\right\}
$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\sqrt{n}}{(2\pi\sigma^2)^{\frac{n-1}{2}}}\exp\left[-\frac{q}{2\sigma^2}\right]\exp\left[\frac{tq}{\sigma^2}\right]dy_2dy_3\cdots dy_n\\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\sqrt{n}}{(2\pi\sigma^2)^{\frac{n-1}{2}}}\exp\left[-\frac{(1-2t)\,q}{2\sigma^2}\right]dy_2dy_3\cdots dy_n\\
&= \left(\frac{1}{1-2t}\right)^{\frac{n-1}{2}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{(1-2t)^{\frac{n-1}{2}}\sqrt{n}}{(2\pi\sigma^2)^{\frac{n-1}{2}}}\exp\left[-\frac{(1-2t)\,q}{2\sigma^2}\right]dy_2dy_3\cdots dy_n\\
&= \left(\frac{1}{1-2t}\right)^{\frac{n-1}{2}}
\end{aligned}
$$

故可求得在給定 $y_1=\overline{x}$ 下，Q/σ^2 為一 $\chi^2(n-1)$ 之隨機變數，即

$$\frac{Q}{\sigma^2}=\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{\sigma^2}=\frac{(n-1)\,S^2}{\sigma^2}\sim x^2\,(n-1)$$

且

$$g\,(y_1,y_2,\ldots,y_n)=h\,(y_1)\,k\left(\frac{(n-1)\,s^2}{\sigma^2}\right)$$

所以，

$$Y_1 \text{ 與}\frac{(n-1)\,S^2}{\sigma^2}\text{ 相獨立，換句話說，}\overline{X}\text{ 與}S^2\text{ 相獨立。}$$

2. 隨機變數 $(n-1)S^2/\sigma^2$ 之分配服從 $\chi^2(n-1)$ 之證明讀者可由前一小題得知，在此介紹另一證明方法, 惟此法必須在 $\overline{X}$ 與 S^2 相獨立之前提下才可成立。證明方法如下：

$$
\begin{aligned}
\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} &= \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}+\overline{X}-\mu\right)^2}{\sigma^2}\\
&= \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}+\frac{\sum_{i=1}^{n}\left(\overline{X}-\mu\right)^2}{\sigma^2}
\end{aligned}
$$

又

$$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2\,(n)$$

且

$$\frac{\sum_{i=1}^{n}\left(\overline{X}-\mu\right)^2}{\sigma^2}=\frac{n\left(\overline{X}-\mu\right)^2}{\sigma^2}=\left(\frac{\overline{X}-u}{\sigma/\sqrt{n}}\right)^2\sim\chi^2\,(1)$$

由於卡方分配具有可加性之性質，即 $\chi^2(r_1)+\chi^2(r_2)=\chi^2(r_1+r_2)$，所以可輕易求得

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}=\frac{(n-1)\,S^2}{\sigma^2}\sim\chi^2\,(n-1)$$

例 **5–3** Suppose $X_1, X_2, \ldots, X_n$ is a random sample from X, where $X\sim\mathcal{N}\left(\mu,\sigma^2\right)$. Find the probability density function of S^2, where $S^2=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2/\,(n-1)$.

解：

令 $Y=(n-1)\,S^2/\sigma^2$, 由定理 **6.3** 可知 $Y\sim\chi^2\,(n-1)$. 因此

$$f\,(y)=\frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}}y^{\frac{n-1}{2}-1}e^{-\frac{y}{2}}, 0<y<\infty$$

又

$$|J|=\left|\frac{dY}{dS^2}\right|=\frac{n-1}{\sigma^2}$$

所以

$$\begin{aligned}
g\left(s^2\right) &= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}}\left(\frac{(n-1)\,S^2}{\sigma^2}\right)^{\frac{n-1}{2}-1}e^{-\frac{(n-1)S^2}{2\sigma^2}}\times|J| \\
&= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}}\left(\frac{(n-1)\,S^2}{\sigma^2}\right)^{\frac{n-1}{2}-1}e^{-\frac{(n-1)S^2}{2\sigma^2}}\frac{n-1}{\sigma^2} \\
&= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}}\left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}-1}\left(s^2\right)^{\frac{n-1}{2}-1}e^{-\frac{(n-1)S^2}{2\sigma^2}}, 0<s^2<\infty
\end{aligned}$$

例 **5–4** 若 $X_i\overset{iid}{\sim}\mathcal{N}\left(\mu,\sigma^2\right)$, $i=1,2,\cdots,n$, 假設母體平均數 μ 已知且母體變異數 σ^2 未知時，令統計量

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}},$$

則 $T\sim t(n-1)$

證明.

因為 $X_i\overset{iid}{\sim}\mathcal{N}\left(\mu,\sigma^2\right)$，所以

$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim\mathcal{N}\,(0,\,1)$$

且

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

故得

$$\frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}}=\frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{S/\sigma}=\frac{\overline{X}-\mu}{S/\sqrt{n}}=T$$

又

$$\frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}}\sim t(n-1)$$

所以，

$$T\sim t(n-1)$$

例 **5–5** 米蘭寶石公司賣出藍寶石之平均價為 8.5 萬元/克拉，變異數未知之常態母體。上一週賣出 25 克拉藍寶石中，其價格標準差為 1.9 萬元/克拉。求此週賣出 25 克拉藍寶石中之平均價大於 9 萬元/克拉之機率？

解：

因為 $n=25<30$，又母體變異數未知，所以得知

$$\frac{\overline{X}-8.5}{\frac{1.9}{\sqrt{25}}}\sim t(24)$$

因此,

$$\begin{aligned}
\mathrm{P}\left(\overline{X}>9\right) &= \mathrm{P}\left(\frac{\overline{X}-8.5}{\frac{1.9}{\sqrt{25}}}>\frac{9-8.5}{\frac{1.9}{\sqrt{25}}}\right)\\
&= \mathrm{P}(T>1.312)\\
&= 0.103
\end{aligned}$$

5.3.2 $\mathcal{F}$ 分配

假設 $X_1 \sim \chi^2(r_1)$, $X_2 \sim \chi^2(r_2)$ 且 X_1, X_2 獨立，則統計量

$$F = \frac{X_1/r_1}{X_2/r_2} \sim \mathcal{F}(r_1,\, r_2)$$

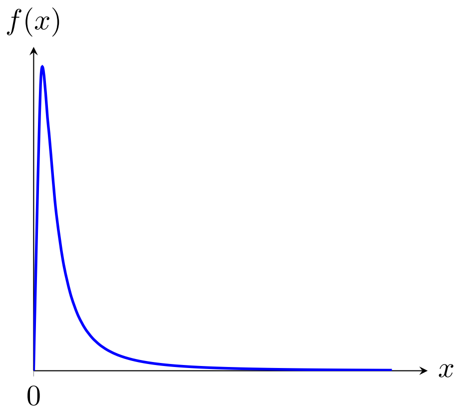


圖 5.2: F 分配圖

特性

1. 定義域為 $[0, \infty)$
2. 不對稱, 為一右偏分配
3. 與卡方分配相似

性質

1. $\mathrm{E}\,(F) = \frac{r_2}{r_2-2}$, $r_2 > 2$
2. $\mathrm{Var}\,(F) = 2\left(\frac{r_2}{r_2-2}\right)^2 \frac{r_2+r_1-2}{r_1(r_2-4)}$, $r_2 > 4$
3. $F_{1-\alpha}\,(r_1, r_2) = 1/F_\alpha\,(r_2, r_1)$
4. 若 $T \sim t\,(v)$，則 $T^2 \sim \mathcal{F}\,(1, v)$

證明.

1. 由 **4.7.1 性質 c** (第37頁) 可知

$$
\begin{aligned}
\mathrm{E}(F) &= \mathrm{E}\left(\frac{X_1/r_1}{X_2/r_2}\right) = \frac{r_2}{r_1} \times \mathrm{E}(X_1) \times \mathrm{E}\left(\frac{1}{X_2}\right) \\
&= \frac{r_2}{r_1} \times r_1 \times \frac{1}{r_2 - 2} \\
&= \frac{r_2}{r_2 - 2}
\end{aligned}
$$

2.

$$
\begin{aligned}
\mathrm{E}\left(F^2\right) &= \mathrm{E}\left(\frac{X_1^2/r_1^2}{X_2^2/r_2^2}\right) = \frac{r_2^2}{r_1^2} \times \mathrm{E}\left(X_1^2\right) \times \mathrm{E}\left(\frac{1}{X_2^2}\right) \\
&= \frac{r_2^2}{r_1^2} \times \left(r_1^2 + 2r_1\right) \times \mathrm{E}\left(\frac{1}{X_2^2}\right)
\end{aligned}
$$

又

$$
\begin{aligned}
\mathrm{E}\left(\frac{1}{X_2^2}\right) &= \int_0^\infty \frac{1}{x_2^2} \frac{1}{\Gamma(\alpha)\,\theta^\alpha} x_2^{\alpha-1} e^{-\frac{x_2}{\theta}} dx_2 \\
&= \int_0^\infty \frac{1}{\Gamma(\alpha)\,\theta^\alpha} x_2^{\alpha-3} e^{-\frac{x_2}{\theta}} dx_2 \\
&= \frac{\theta^{\alpha-2}}{\Gamma(\alpha)\,\theta^\alpha} \int_0^\infty \left(\frac{x_2}{\theta}\right)^{\alpha-3} e^{-\frac{x_2}{\theta}} d\frac{x_2}{\theta} \\
&= \frac{\theta^{\alpha-2}}{\Gamma(\alpha)\,\theta} \int_0^\infty y^{\alpha-3} e^{-y} dy \\
&= \frac{\Gamma(\alpha-2)}{\Gamma(\alpha)\,\theta^2} = \frac{1}{\theta^2(\alpha-1)(\alpha-2)} \\
&= \frac{1}{4\left(\frac{r_2}{2}-1\right)\left(\frac{r_2}{2}-2\right)} \\
&= \frac{1}{(r_2-2)(r_2-4)}
\end{aligned}
$$

所以

$$
\mathrm{E}\left(F^2\right) = \frac{r_2^2}{r_1^2} \times \frac{\left(r_1^2 + 2r_1\right)}{(r_2-2)(r_2-4)}
$$

故可推得

$$\begin{aligned}
\operatorname{Var}(F) &= \frac{r_2^2}{r_1^2}\frac{\left(r_1^2+2r_1\right)}{\left(r_2-2\right)\left(r_2-4\right)}-\left(\frac{r_2}{r_2-2}\right)^2 \\
&= \left(\frac{r_2}{r_2-2}\right)^2\left[\frac{1}{r_1^2}\frac{\left(r_2-2\right)\left(r_1^2+2r_1\right)}{\left(r_2-4\right)}-1\right] \\
&= \left(\frac{r_2}{r_2-2}\right)^2\frac{\left(r_2-2\right)\left(r_1^2+2r_1\right)-r_1^2\left(r_2-4\right)}{r_1^2\left(r_2-4\right)} \\
&= \left(\frac{r_2}{r_2-2}\right)^2\frac{r_1^2r_2+2r_1r_2-2r_1^2-4r_1-r_1^2r_2+4r_1^2}{r_1^2\left(r_2-4\right)} \\
&= 2\left(\frac{r_2}{r_2-2}\right)^2\frac{r_2+r_1-2}{r_1\left(r_2-4\right)}
\end{aligned}$$

3.

$$F\left(r_1,r_2\right)=\frac{X_1/r_1}{X_2/r_2}=\frac{1}{\frac{X_2/r_2}{X_1/r_1}}=\frac{1}{F\left(r_2,r_1\right)}$$

4.

$$T^2=\left(\frac{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\chi^2(v)}{v}}}\right)^2=\frac{Z^2}{\frac{\chi^2(v)}{v}}=\frac{\frac{Z^2}{1}}{\frac{\chi^2(v)}{v}}=\frac{\frac{\chi^2(1)}{1}}{\frac{\chi^2(v)}{v}}=F\left(1,v\right)$$

例 **5–6** 若 $F\sim\mathcal{F}(4,9)$, 求 c, d 常數使得 $\mathrm{P}(F\le c)=0.0025$ 且 $\mathrm{P}(F\le d)=0.05$

解：

$$c=F_{0.975}(4,9)=\frac{1}{F_{0.025}(9,4)}=\frac{1}{8.90}=0.1124$$

$$d=F_{0.95}(4,9)=\frac{1}{F_{0.05}(9,4)}=\frac{1}{6.00}=0.1667$$

5.3.3 大數法則

假設一隨機機率密度函數 $f(x)$，其平均數為 μ，變異數為 σ^2，且 $0<\sigma^2<\infty$, 令 ε 為任意微小正數，則

1. **弱大數法則(Weak Law of Large Number)**

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\overline{X}-\mu\right| > \varepsilon\right) = 0$$

證明.

由 Chebyshev's Theorem(第18頁) 可知

$$\mathrm{P}\left(\mid \overline{X}-\mu \mid > k\sigma_{\overline{X}}\right) \leq \frac{1}{k^2}$$

令 $\varepsilon = k\sigma_{\overline{X}}$，則

$$\lim_{n\to\infty} \mathrm{P}\left(\mid \overline{X}-\mu \mid > \varepsilon\right) \leq \frac{\sigma_{\overline{X}}^2}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} = 0$$

通常我們以 $\overline{X} \xrightarrow{p} \mu$ 表示。

2. **強大數法則 (Strong Law of Large Number)**

$$\mathrm{P}\left(\left|\lim_{n\to\infty} \overline{X}-\mu\right| = 0\right) = 1$$

定義 1. 假設 $X_1, X_2, \ldots, X_n$ 為一數列且其 *cdf* 為 F_n 之隨機變數，X 亦為一隨機變數且其 *cdf* 為 F，若在任意 F 的連續點上

$$\lim_{n\to\infty} F_n(x) = F(x),$$

則稱 X_n 在分配上收斂到 F, 通常以 $X_n \xrightarrow{d} X$ 表示。

定理 5.4. *Slutsky's Theorem*：假設 $X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n$ 為兩數列之隨機變數，且 $X_n \xrightarrow{d} X$ 及 $Y_n \xrightarrow{p} Y$. 則

1. $X_n + Y_n \xrightarrow{d} X + c$

2. $X_n Y_n \xrightarrow{d} cX$

3. 若 $c \neq 0$，則 $X_n/Y_n \xrightarrow{d} X/c$

5.3.4 中央極限定理

定理 5.5. 若隨機變數 $X_1, X_2, \ldots, X_n$ 為抽自任意母體的一組隨機樣本。若母體的期望值及變異數皆存在 $(\mu < \infty, \sigma^2 < \infty)$，則當樣本數趨近無窮大時 $(n \to \infty)$ 時，

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

證明.

$$\begin{aligned}
\mathrm{E}\left\{\exp\left[t\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right]\right\} &= \mathrm{E}\left\{\exp\left[t\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}\right]\right\} \\
&= \mathrm{E}\left\{\exp\left[\frac{X_i-\mu}{\sigma}\frac{t}{\sqrt{n}}\right]\right\}\mathrm{E}\left\{\exp\left[\frac{X_2-\mu}{\sigma}\frac{t}{\sqrt{n}}\right]\right\}\cdots\mathrm{E}\left\{\exp\left[\frac{X_n-\mu}{\sigma}\frac{t}{\sqrt{n}}\right]\right\} \\
&= M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)\cdots M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) \\
&= \left[M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)\right]^n
\end{aligned}$$

又

$$M'_{\frac{X_i-\mu}{\sigma}}(0) = \mathrm{E}\left(\frac{X_i-\mu}{\sigma}\right) = 0, M''_{\frac{X_i-\mu}{\sigma}}(0) = \mathrm{Var}\left(\frac{X_i-\mu}{\sigma}\right) + \left[\mathrm{E}\left(\frac{X_i-\mu}{\sigma}\right)\right]^2 = 1$$

由 **4.9** 性質 **3** 可知

$$\begin{aligned}
M_{\frac{X_i-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right) &= \sum_{r=1}^{\infty}\frac{M^{(r)}_{\frac{X_i-\mu}{\sigma}}(0)}{r!}\left(\frac{t}{\sqrt{n}}\right)^r \\
&\approx M_{\frac{X_i-\mu}{\sigma}}(0) + M'_{\frac{X_i-\mu}{\sigma}}(0)\frac{t}{\sqrt{n}} + \frac{1}{2!}M''_{\frac{X_i-\mu}{\sigma}}(0)\left(\frac{t}{\sqrt{n}}\right)^2 \\
&= 1 + \frac{t^2}{2n}
\end{aligned}$$

又

$$\lim_{n\to\infty}\left(1+\frac{t^2}{2n}\right)^n = \lim_{n\to\infty}\left(1+\frac{\frac{t^2}{2}}{n}\right)^n = e^{\frac{t^2}{2}}$$

故可推得當樣本數趨近於無窮大時，

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

一般來說，當 $n>30$(即統計學上所謂之大樣本) 時，上述定理即可成立。必須注意的是當 μ 或 σ^2 不存在時，此定理並不適用，例如柯西分配 (Cauchy Distribution，第44頁)。

例 **5–7** Let $Y_n \sim \chi^2(n)$. Find the limiting distribution of $(Y_n - n)/\sqrt{2n}$, as $n \to \infty$, using moment generating functions.

解：

$$
\begin{aligned}
M(t) &= \mathrm{E}\left\{\exp\left[t\frac{Y_n-n}{\sqrt{2n}}\right]\right\} = \mathrm{E}\left\{\exp\left[\frac{t}{\sqrt{2n}}Y_n\right]\exp\left[\frac{-\sqrt{n}t}{\sqrt{2}}\right]\right\} \\
&= \exp\left[\frac{-\sqrt{n}t}{\sqrt{2}}\right]\mathrm{E}\left\{\exp\left[\frac{t}{\sqrt{2n}}Y_n\right]\right\} \\
&= \exp\left[\frac{-\sqrt{n}t}{\sqrt{2}}\right]\left(1+\frac{\sqrt{2}t}{\sqrt{n}}\right)^{-\frac{n}{2}} = \exp\left[\frac{-\sqrt{n}t}{\sqrt{2}}\right]\exp\left[-\frac{n}{2}\ln\left(1+\frac{\sqrt{2}t}{\sqrt{n}}\right)\right] \\
&= \exp\left\{\frac{-\sqrt{n}t}{\sqrt{2}} - \frac{n}{2}\left[\frac{\sqrt{2}t}{\sqrt{n}} - \frac{2t^2}{2n} + \frac{1}{3}\left(\frac{\sqrt{2}t}{\sqrt{n}}\right)^3 + \cdots\right]\right\} \\
&\approx e^{\frac{t^2}{2}}
\end{aligned}
$$

所以

$$
\lim_{n\to\infty}\frac{Y_n-n}{\sqrt{2n}} \sim \mathcal{N}(0,1)
$$

例 **5–8** Let $X_i \sim \text{Bernoulli}(p)$, $i=1,2,\ldots$. Find the limiting distribution of

$$
\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}},
$$

as $n \to \infty$, using moment generating functions.

解：

$$
\begin{aligned}
M(t) &= \mathrm{E}\left\{\exp\left[t\left(\frac{\sum_{i=1}^n X_i - np}{\sqrt{n\mathrm{P}(1-p)}}\right)\right]\right\} \\
&= \mathrm{E}\left\{\exp\left[\tfrac{-np}{\sqrt{n\mathrm{P}(1-p)}}t\right]\exp\left[\frac{t\sum_{i=1}^n X}{\sqrt{n\mathrm{P}(1-p)}}\right]\right\} \\
&= \exp\left[\frac{-np}{\sqrt{n\mathrm{P}(1-p)}}t\right]\mathrm{E}\left\{\exp\left[\frac{t\sum_{i=1}^n X}{\sqrt{n\mathrm{P}(1-p)}}\right]\right\} \\
&= \exp\left[\frac{-np}{\sqrt{n\mathrm{P}(1-p)}}t\right]\left(1-p+p\exp\left[\frac{t}{\sqrt{n\mathrm{P}(1-p)}}\right]\right)^n
\end{aligned}
$$

$$
\begin{aligned}
&= \left\{(1-p)\exp\left[\frac{-pt}{\sqrt{n\mathrm{P}(1-p)}}\right]+p\exp\left[\frac{(1-p)\,t}{\sqrt{n\mathrm{P}(1-p)}}\right]\right\}^n \\
&\approx \left[(1-p)\left(1-\frac{pt}{\sqrt{n\mathrm{P}(1-p)}}+\frac{p^2t^2}{2n\mathrm{P}(1-p)}\right)+\mathrm{P}\left(1+\frac{(1-p)\,t}{\sqrt{n\mathrm{P}(1-p)}}+\frac{(1-p)^2t^2}{2n\mathrm{P}(1-p)}\right)\right]^n \\
&= \left[1+\frac{(1-p)\,p^2t^2}{2n\mathrm{P}(1-p)}+\frac{\mathrm{P}(1-p)^2t^2}{2n\mathrm{P}(1-p)}\right]^n=\left[1+\frac{p^2-p^3+p+p^3-2p^2}{2n\mathrm{P}(1-p)}t^2\right]^n \\
&= \left[1+\frac{t^2}{2n}\right]^n
\end{aligned}
$$

所以

$$
\lim_{n\to\infty}\frac{\sum_{i=1}^n X_i-np}{\sqrt{n\mathrm{P}(1-p)}}\sim\mathcal{N}(0,1)
$$

例 **5–9** 花蓮女中學生平均身高為 160 公分, 標準差為 9 公分; 今從中隨機抽取 36 人, 求其平均身高大於 160 公分而小於 162 公分的機率？

解：

因為 $n=36>30$，由中央極限定理可知

$$
\frac{\overline{X}-160}{9/\sqrt{36}}\sim\mathcal{N}(0,1)
$$

因此,

$$
\begin{aligned}
\mathrm{P}\left(160<\overline{X}<162\right) &= \mathrm{P}\left(\frac{160-160}{\sqrt{81/36}}<Z<\frac{162-160}{\sqrt{81/36}}\right) \\
&= \mathrm{P}(0<Z<1.333) \\
&= 0.5-0.0918 \\
&= 0.4082
\end{aligned}
$$

當 $X_i\sim B(1,p)$, $i=1,\ 2\ \ldots\ n$, 令 $Y=\sum_i^n X_i$，則當樣本數夠大時 (一般應用上，若 $np\geq 5$ 且 $n\mathrm{P}(1-p)\geq 5$ 即可)，由中央極限定理可知

$$
W=\frac{Y-np}{\sqrt{n\mathrm{P}(1-p)}}=\frac{\overline{X}-p}{\sqrt{\mathrm{P}(1-p)/n}}\sim\mathcal{N}(0,1)
$$

然而，由於二項分配為離散型分配，在使用上必須作離散母體校正。所以，以中央極限定理做為二項分配 $B(n, p)$ 之近似值其方法如下：

1.
$$\mathrm{P}(Y \leq a) = \mathrm{P}(Y < a + 0.5) \approx P\left(Z \leq \frac{(a+0.5)-\mu}{\sigma}\right)$$

2.
$$\mathrm{P}(Y < a) = \mathrm{P}(Y < a - 0.5) \approx P\left(Z \leq \frac{(a-0.5)-\mu}{\sigma}\right)$$

3.
$$\mathrm{P}(Y \geq b) = \mathrm{P}(Y > b + 0.5) \approx P\left(Z \geq \frac{(b-0.5)-\mu}{\sigma}\right)$$

4.
$$\mathrm{P}(Y > b) = \mathrm{P}(Y > b - 0.5) \approx P\left(Z \leq \frac{(b+0.5)-\mu}{\sigma}\right)$$

例 **5–10**　若 $X \sim B(100,\ 0.02)$，試以中央極限定理求解 $\mathrm{P}(X = 3)$

解：

$$\begin{aligned}
\mathrm{P}(X = 3) &= \mathrm{P}(X \leq 3) - \mathrm{P}(X < 3) \\
&= \mathrm{P}(X \leq 3) - \mathrm{P}(X \leq 2) \\
&\approx P\left(Z \leq \frac{(3+0.5)-2}{\sqrt{1.96}}\right) - P\left(Z \leq \frac{(2+0.5)-2}{\sqrt{1.96}}\right) \\
&= P\left(Z \leq \frac{1.5}{\sqrt{1.96}}\right) - P\left(Z \leq \frac{0.5}{\sqrt{1.96}}\right) \\
&= P(Z \leq 1.07) - \mathrm{P}(Z \leq 0.357) = 0.8577 - 0.64 \\
&= 0.2177
\end{aligned}$$

例 **5–11**　假設進入台灣大學網站的流量服從波爾生分配 (Poisson Distribution), 且平均流量為每 5 分鐘 20 人。如果在 1 分鐘內進入的人數超過 8 人, 便會產生網路壅塞的現象。假設你現在登入台灣大學的網站, 請以中央極限定理估算你會碰到網路壅塞的機率。

解：

$$\mathrm{P}(X>8)=\mathrm{P}(X>8.5)=\mathrm{P}\left(\frac{X-4}{\sqrt{4}}>\frac{8.5-4}{\sqrt{4}}\right)=\mathrm{P}(Z>2.25)=0.0122$$

5.4 順序統計量

定理 5.6. 若 $X_1,\ X_2,\ \ldots,\ X_n$ 為抽自母體 $f(x;\theta)$ 的一組樣本數為 n 的隨機樣本，若 Y_i 表是 X_1, $X_2,\ \ldots,\ X_n$ 中第 i 個小的值，$1\leq i\leq n$，若 $f(x;\theta)$ 為一連續型分配，則

$$f(y_i)=\frac{n!}{(i-1)!\,(n-i)!}\left[F(y_i)\right]^{i-1}\left[1-F(y_i)\right]^{n-i}f(y_i)$$

證明.

由於 $X_1,\ X_2,\ \ldots,\ X_n$ 為抽自母體 $f(x;\theta)$ 的一組樣本數為 n 的隨機樣本，所以 $Y_1,\ Y_2,\ \ldots,\ Y_n$ 之聯合機率密度函數為

$$g(y_1,y_2,\cdots y_{n-1},y_n)=n!f(y_1)f(y_2)\cdots f(y_{n-1})f(y_n),-\infty<y_1<y_2<\cdots<y_{n-1}<y_n<\infty$$

由邊際機率密度函數定義可知

$$\begin{aligned}h(y_i)&=\int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}}\int_{-\infty}^{y_i}\cdots\int_{-\infty}^{y_3}\int_{-\infty}^{y_2}n!f(y_1)f(y_2)\cdots f(y_{n-1})f(y_n)\,dy_1dy_2\cdots dy_{n-1}dy_n\\&=\int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}}\int_{-\infty}^{y_i}\cdots\int_{-\infty}^{y_3}n!f(y_2)\cdots f(y_{n-1})f(y_n)F(y_2)\,dy_2\cdots dy_{n-1}dy_n\end{aligned}$$

利用分部積分法可求得

$$\int_{-\infty}^{y_3}f(y_2)F(y_2)\,dy_2=\frac{1}{2}\left[F(y_2)\right]^2\Big|_{-\infty}^{y_3}=\frac{1}{2}\left[F(y_3)\right]^2$$

所以

$$f(y_i)=\int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}}\int_{-\infty}^{y_i}\cdots\int_{-\infty}^{y_3}n!f(y_2)\cdots f(y_{n-1})f(y_n)F(y_2)\,dy_2\cdots dy_{n-1}dy_n$$

$$
\begin{aligned}
&= \int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}}\int_{-\infty}^{y_i}\cdots\int_{-\infty}^{y_4} f(y_3)\cdots f(y_{n-1})\,f(y_n)\times\frac{n!}{2}\left[F(y_3)\right]^2 dy_3\cdots dy_{n-1}dy_n \\
&= \int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}}\int_{-\infty}^{y_i}\cdots\int_{-\infty}^{y_5} f(y_4)\cdots f(y_{n-1})\,f(y_n)\times\frac{n!}{2\times 3}\left[F(y_4)\right]^3 dy_4\cdots dy_{n-1}dy_n \\
&\quad\vdots \\
&= \int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}} f(y_i)\,f(y_{i+1})\cdots f(y_{n-1})\,f(y_n) \\
&\quad \times\frac{n!}{2\times 3\times\cdots(i-1)}\left[F(y_i)\right]^{i-1} dy_{i+1}\cdots dy_{n-1}dy_n \\
&= \frac{n!}{(i-1)!}\left[F(y_i)\right]^{i-1} f(y_i) \\
&\quad \times\int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}} f(y_i)\,f(y_{i+1})\cdots f(y_{n-1})\,f(y_n)\,dy_{i+1}\cdots dy_{n-1}dy_n
\end{aligned}
$$

又

$$
\begin{aligned}
&\int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}} f(y_i)\,f(y_{i+1})\cdots f(y_{n-1})\,f(y_n)\,dy_{i+1}\cdots dy_{n-1}dy_n \\
&= \int_{y_{n-1}}^{\infty}\int_{y_{n-2}}^{y_n}\cdots\int_{y_i}^{y_{i+3}} f(y_{i+2})\cdots f(y_{n-1})\,f(y_n)\times\left[F(y_{i+2})-F(y_i)\right] dy_{i+2}\cdots dy_{n-1}dy_n \\
&= \int_{y_i}^{\infty}\int_{y_i}^{y_n}\cdots\int_{y_i}^{y_{i+2}} f(y_{i+2})\cdots f(y_{n-1})\,f(y_n)\times\frac{1}{2}\left[F(y_{i+3})-F(y_i)\right]^2 dy_{i+3}\cdots dy_{n-1}dy_n \\
&\quad\vdots \\
&= \int_{y_i}^{\infty} f(y_n)\times\frac{1}{2\times 3\times\cdots\times(n-i-1)}\left[F(y_n)-F(y_i)\right]^{n-i-1} dy_n \\
&= \frac{1}{2\times 3\times\cdots\times(n-i-1)\times(n-i)}\left[F(y_n)-F(y_i)\right]^{n-i}\Big|_{y_i}^{\infty} \\
&= \frac{1}{2\times 3\times\cdots\times(n-i-1)\times(n-i)}\left[1-F(y_i)\right]^{n-i}
\end{aligned}
$$

所以

$$
\begin{aligned}
h(y_i) &= \frac{n!}{(i-1)!\,(n-i)!}\left[F(y_i)\right]^{i-1}\left[1-F(y_i)\right]^{n-i} f(y_i) \\
&= \frac{\Gamma(n+1)}{\Gamma(i)\,\Gamma(n-i+1)}\left[F(y_i)\right]^{i-1}\left[1-F(y_i)\right]^{n-i} f(y_i)
\end{aligned}
$$

故可知第 i 個順序統計量之 p.d.f. 為 $Beta(i,n-i+1)$，即 $Y_i\sim Beta(i,n-i+1)$。所以

$$
\mathrm{E}(Y_i)=\frac{i}{n+1}\ \text{及}\mathrm{Var}(Y_i)=\frac{i\times(n-i+1)}{(n+1)^2(n+2)}
$$

此外，定理5.6可推廣求得 Y_i 與 Y_j, $1 \le i \le j \le n$, 之聯合機率密度函數為

$$k(y_1, y_n) = \frac{n!}{(i-1)!(j-i-1)(n-j)!}[F(y_i)]^{i-1}[F(y_j)-F(y_i)]^{j-i-1}[1-F(y_i)]^{n-j}f(y_i)f(y_j)$$

例 **5–12** $X_i \sim U(0,\theta)$, $i=1,2,\ldots,n$，令 $Y_1 = \min\{X_1, X_2, \ldots, X_n\}$ 及 $Y_n = \max\{X_1, X_2, \ldots, X_n\}$，試求 Y_1 與 Y_n 之機率密度函數及聯合機率密度函數。

解：

$$\begin{aligned} g(y_1) &= \frac{n!}{(1-1)!(n-1)!} \times \left[\frac{y_1}{\theta}\right]^{1-1} \times \left[1-\frac{y_1}{\theta}\right]^{n-1} \times \frac{1}{\theta} \\ &= \frac{n(1-y_1)^{n-1}}{\theta^n}, 0 < y_1 < \theta, \end{aligned}$$

$$\begin{aligned} h(y_n) &= \frac{n!}{(n-1)!(n-n)!} \times \left[\frac{y_n}{\theta}\right]^{n-1} \times \left[1-\frac{y_n}{\theta}\right]^{n-n} \times \frac{1}{\theta} \\ &= \frac{ny_n^{n-1}}{\theta^n}, 0 < y_n < \theta \end{aligned}$$

及

$$\begin{aligned} k(y_1, y_n) &= \frac{n!}{(1-1)!(n-1-1)(n-n)!}\left[\frac{y_1}{\theta}\right]^{1-1} \times \left[\frac{y_n}{\theta}-\frac{y_1}{\theta}\right]^{n-1-1} \times \left[1-\frac{y_1}{\theta}\right]^{n-n} \times \frac{1}{\theta} \times \frac{1}{\theta} \\ &= \frac{n(n-1)(y_n-y_1)^{n-2}}{\theta^n}, 0 < y_1 < y_n < \theta \end{aligned}$$

例 **5–13** Let X_1, X_2, … be iid $U(0,\theta)$ r.v.'s. Let $Y_1 = \min(X_1, X_2, \ldots, X_n)$, find the limiting distribution of $Z_n = nY_1$ as $n \to \infty$.

解：由定理5.6可知 y_1 之機率密度函數為

$$\begin{aligned} f(y_1) &= \frac{n!}{(1-1)!(n-1)!}\left(\frac{y_1}{\theta}\right)^{1-1}\left(1-\frac{y_1}{\theta}\right)^{n-1}\frac{1}{\theta} \\ &= \frac{n(\theta-y_1)^{n-1}}{\theta^n}, 0 < y_1 < \theta. \end{aligned}$$

故可求得 Z_n 之累積機率密度函數為

$$\begin{aligned} G(z_n) &= \mathrm{P}(Z_n < z_n) = \mathrm{P}(nY_1 < z_n) = \mathrm{P}\left(Y_1 < \frac{z_n}{n}\right) \\ &= \int_0^{z_n/n} \frac{n(\theta - y_1)^{n-1}}{\theta^n} dy_1. \end{aligned}$$

因此由定義可知

$$g(z_n) = \frac{dG(z_n)}{dz_n} = \frac{n(\theta - z_n/n)^{n-1}}{\theta^n}\frac{1}{n} = \frac{(\theta - z_n/n)^{n-1}}{\theta^n},\ z_n > 0$$

所以，當 $n \to \infty$ 時，

$$\begin{aligned} \lim_{n\to\infty} g(z_n) &= \lim_{n\to\infty} \frac{(\theta - z_n/n)^{n-1}}{\theta^n} = \lim_{n\to\infty} \frac{1}{\theta}\left(1 - \frac{z_n}{\theta n}\right)^{n-1} \\ &= \lim_{n\to\infty} \frac{1}{\theta}\frac{1}{[1 - z_n/(\theta n)]}\left(1 - \frac{z_n}{\theta n}\right)^{n-1} \\ &= \lim_{n\to\infty} \frac{1}{\theta}\frac{1}{[1 - z_n/(\theta n)]} \lim_{n\to\infty}\left(1 - \frac{z_n}{\theta n}\right)^{n} \\ &= \frac{1}{\theta}\exp\left(-\frac{z_n}{\theta}\right),\ z_n > 0 \end{aligned}$$

例 **5–14** 某拍賣網站，有 20 家供應商參與競標，若每家廠商的定價 p 為一均等分配 $U(10, 20)$，20 家競標中以最低價格 P_l 為得標者，試求：

1. P_l 的機率密度函數
2. 期望值 $E(P_l)$，若參與競標家數由 20 家增為 40 家則 $E(P_l)$ 的期望值變化為何？

解：

1.

$$\begin{aligned} g(p_l) &= \frac{20!}{(1-1)!(20-1)!} \times \left(\frac{p_l - 10}{10}\right)^{1-1} \times \left(\frac{20 - p_l}{10}\right)^{20-1} \times \frac{1}{10} \\ &= \frac{2}{10^{19}}(20 - p_l)^{19}, 10 < p_l < 20 \end{aligned}$$

2.

$$E(P_l) = \int_{10}^{20} p_l \times \frac{2}{10^{19}}(20 - p_l)^{19} dp_l$$

$$
\begin{aligned}
&= \int_{10}^{20} (p_l - 10) \times \frac{2}{10^{19}} (20 - p_l)^{19} \, dp_l + \int_{10}^{20} 10 \times \frac{2}{10^{19}} (20 - p_l)^{19} \, dp_l \\
&= \frac{2}{10^{19}} \times \int_{10}^{20} (p_l - 10)(20 - p_l)^{19} \, dp_l + 10 \\
&= 200 \times \int_{10}^{20} \left(\frac{p_l - 10}{10} \right) \left(\frac{20 - p_l}{10} \right)^{19} d\frac{p_l - 10}{10} + 10 \\
&= 200 \times \int_{0}^{1} u (1 - u)^{19} \, du + 10 \\
&= 200 \times \int_{0}^{1} u^{2-1} (1 - u)^{20-1} \, du + 10 \\
&= 200 \times \frac{\Gamma(2)\Gamma(20)}{\Gamma(2+20)} + 10 \\
&= 200 \times \frac{19!}{21!} + 10 \\
&= \frac{200}{20 \times 21} + 10 \\
&= 10\frac{10}{21} = 10\frac{10}{\text{買家家數} + 1}
\end{aligned}
$$

當買家由 20 增加為 40 時，$E(P_l) = 10\frac{10}{41}$

例 **5–15** X_1, X_2, X_3, X_4, X_5 為取自區間 (0,1) 均勻分配母體的隨機樣本，求 X_1, X_2, X_3, X_4, X_5 中的最小值小於 0.5 的機率？

解：令 $Y_1 = \min\{X_1, X_2, X_3, X_4, X_5\}$，所以

$$
\begin{aligned}
g(y_1) &= \frac{5!}{(1-1)!(5-1)!} y_1^{1-1} (1 - y_1)^{5-1} \cdot 1 \\
&= 5(1 - y_1)^4, 0 < y_1 < 1
\end{aligned}
$$

故得

$$
P(Y_1 < 0.5) = \int_{0}^{0.5} 5(1 - y_1)^4 \, dy_1 = 0.96875
$$

例 **5–16** Find $P(Y_2 < \pi_{0.3} < Y_8)$, where Y_2 and Y_8 are the second and eighth order statistics of 16 continuous-type independent observations.

解： $Y_2 < \pi_{0.3} < Y_8$ 即比第三個十分位數小的個數至少兩個，最多不超過七個，故機率為

$$
\sum_{x=2}^{7} \binom{16}{x} 0.3^x 0.7^{16-x} = 0.89954
$$

第 6 章 估計

假設一隨機變數 X 的母體機率密度函數為 $f(x \mid \theta)$，其中 θ 為未知的參數，即 $X \sim f(x \mid \theta)$。估計之目的主要是為了解此未知參數 θ，故用 $\widehat{\theta}$ 表示之。一般估計可分為兩部分, 分別為**點估計**及**區間估計**。

6.1 點估計

所謂點估計乃是根據一組樣本資料, 並藉由一統計量之觀測值來作為欲估計母體參數之估計值。點估計常用的方法有：

1. 最大概似法（Maximum Likelihood Method）
2. 動差法（Moment Method）

而判斷估計式的優劣標準有以下四個原則：

1. **不偏性**：假設 $\widehat{\theta}$ 為 θ 的估計式，若 $\mathrm{E}\left(\widehat{\theta}\right) = \theta$ 則稱 $\widehat{\theta}$ 為 θ 的不偏估計式; 若 $\mathrm{E}\left(\hat{\theta}\right) \neq \theta$ 則稱 $\widehat{\theta}$ 為 θ 的偏誤估計式，且其偏誤 (bias)$=\mathrm{E}\left(\widehat{\theta} - \theta\right)$, 例：$\mathrm{E}\left(\overline{X}\right) = \mu$, $\mathrm{E}\left(S^2\right) = \sigma^2$。
2. **一致性**：當樣本增加時，估計式所產生之估計值接近真值之機率也相對增加, 而具備此一性質之估計式稱之為一致估計式，即對任意微小正數 ϵ，若 $\mathrm{E}\left(\widehat{\theta}\right) = \theta$ 且

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\widehat{\theta}_n - \theta\right| < \epsilon\right) = 1$$

或

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\widehat{\theta}_n - \theta\right| > \epsilon\right) = 0$$

或

$$\lim_{n\to\infty} \operatorname{Var}\left(\widehat{\theta}_n\right) = 0,$$

則稱 $\widehat{\theta}_n$ 為 θ 的一致估計式。

3. **有效性**：設 $\widehat{\theta}_1$ 與 $\widehat{\theta}_2$ 均為 θ 之不偏估計式，若 $\operatorname{Var}\left(\widehat{\theta}_1\right) \le \operatorname{Var}\left(\widehat{\theta}_2\right)$ 時，$\widehat{\theta}_1$ 相對有效，且 $\widehat{\theta}_1$ 相對於 $\widehat{\theta}_2$ 之效率為

$$\frac{\operatorname{Var}\left(\widehat{\theta}_2\right)}{\operatorname{Var}\left(\widehat{\theta}_1\right)}$$

定理 6.1. 設 $X_1, X_2, \ldots, X_n$ 為抽自母體 $f(\theta; x)$ 之一組樣本數為 n 之樣本，$\tau(\theta)$ 為 θ 之實數函數; $W = W(x)$ 為 $\tau(\theta)$ 之不偏估計式，若滿足下列條件：

(a) $S_x = \{\mathbf{X} = (X_1, X_2, \ldots, X_n) \mid f(\theta; x) > 0\}$

(b) $\frac{\partial}{\partial\theta} \int \cdots \int f(\theta; \mathbf{X})\, d\mathbf{X} = \int \cdots \int \frac{\partial}{\partial\theta} f(\theta; \mathbf{X})\, d\mathbf{X}$

(c) $\frac{\partial}{\partial\theta} \int \cdots \int W(x) f(\theta; \mathbf{X})\, d\mathbf{X} = \int \cdots \int W(x) \frac{\partial}{\partial\theta} f(\theta; \mathbf{X})\, d\mathbf{X}$

(d) $0 < \mathrm{E}\left[\frac{\partial}{\partial\theta} \ln f(\theta; \mathbf{X})\right]^2 < \infty$

則

$$\operatorname{Var}(W) \ge \frac{[\tau'(\theta)]^2}{\mathrm{E}\left[\left[\frac{\partial}{\partial\theta} \ln f(\theta; \mathbf{X})\right]^2\right]}$$

證明.

由機率密度函數之定義,

$$1 = \int \cdots \int f(\theta; \mathbf{X})\, d\mathbf{X}$$

左右對 θ 做微分得

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta} \int \cdots \int f(\theta; \mathbf{X})\, d\mathbf{X} = \int \cdots \int \frac{\partial}{\partial\theta} f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \int \cdots \int \frac{\partial}{\partial\theta} f(\theta; \mathbf{X}) \times \frac{1}{f(\theta; \mathbf{X})} \times f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \int \cdots \int \frac{\frac{\partial}{\partial\theta} f(\theta; \mathbf{X})}{f(\theta; \mathbf{X})} \times f(\theta; \mathbf{X})\, d\mathbf{X}
\end{aligned}$$

$$
\begin{aligned}
&= \int \cdots \int \frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X}) \times f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \mathrm{E}\left[\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right]
\end{aligned}
$$

又 $W = W(x)$ 為 $\tau(\theta)$ 之不偏估計式，所以

$$
\tau(\theta) = \int \cdots \int W(x) f(\theta; \mathbf{X})\, d\mathbf{X}
$$

左右對 θ 做微分得

$$
\begin{aligned}
\tau'(\theta) &= \int \cdots \int W(x) \frac{\partial}{\partial \theta} f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \int \cdots \int W(x) \frac{\partial}{\partial \theta} f(\theta; \mathbf{X}) \times \frac{1}{f(\theta; \mathbf{X})} \times f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \int \cdots \int W(x) \times \frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X}) \times f(\theta; \mathbf{X})\, d\mathbf{X} \\
&= \mathrm{E}\left[W(x) \frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right]
\end{aligned}
$$

又 $EXY = EXEY + \rho\sigma_X\sigma_Y$，所以

$$
\begin{aligned}
\tau'(\theta) &= \mathrm{E}[W]\,\mathrm{E}\left[\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right] + \rho\sqrt{\mathrm{Var}(W)}\sqrt{\mathrm{Var}\left(\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right)} \\
&= \rho\sqrt{\mathrm{Var}(W)}\sqrt{\mathrm{Var}\left(\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right)}
\end{aligned}
$$

因此，

$$
\begin{aligned}
[\tau'(\theta)]^2 &= \rho^2 \mathrm{Var}(W)\, \mathrm{Var}\left(\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right) \\
&\leq \mathrm{Var}(W)\, \mathrm{Var}\left(\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right)
\end{aligned}
$$

故

$$
\mathrm{Var}(W) \geq \frac{[\tau'(\theta)]^2}{\mathrm{Var}\left(\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right)} = \frac{[\tau'(\theta)]^2}{\mathrm{E}\left[\frac{\partial}{\partial \theta} \ln f(\theta; \mathbf{X})\right]^2}
$$

此一不等式稱為 **Rao-Cramer** 不等式。

註解

(a) Rao-Cramer 不等式對離散分配仍然符合。

(b) 定理中的 5 個條件稱為正規條件。

(c) 當 $X_1, X_2, \ldots, X_n$ 為來自樣本數為 n 之隨機樣本時，由於

$$\begin{aligned}
\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\right]^2 &= \mathrm{E}\left[\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f(\theta;x_i)\right]^2 = \sum_{i=1}^{n} E\left[\frac{\partial}{\partial\theta}\ln f(\theta;x_i)\right]^2 \\
&= nE\left[\frac{\partial}{\partial\theta}\ln f(\theta;x)\right]^2
\end{aligned}$$

及

$$\begin{aligned}
0 &= \frac{\partial}{\partial\theta}\int\cdots\int\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\times f(\theta;\mathbf{X})\,d\mathbf{X} \\
&= \int\cdots\int\frac{\partial}{\partial\theta}\left\{\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\times f(\theta;\mathbf{X})\right\}d\mathbf{X} \\
&= \int\cdots\int\left[\frac{\partial^2}{\partial\theta^2}\ln f(\theta;\mathbf{X})\times f(\theta;\mathbf{X})+\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\times\frac{\partial}{\partial\theta}f(\theta;\mathbf{X})\right]d\mathbf{X} \\
&= \int\cdots\int\left[\frac{\partial^2}{\partial\theta^2}\ln f(\theta;\mathbf{X})\times f(\theta;\mathbf{X})+\left(\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\right)^2\times f(\theta;\mathbf{X})\right]d\mathbf{X} \\
&= \mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln f(\theta;\mathbf{X})\right]+\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})\right]^2
\end{aligned}$$

所以，

$$\mathrm{Var}(W) \geq \frac{[\tau'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(\theta;x)\right]^2} = \frac{[\tau'(\theta)]^2}{-nE\left[\frac{\partial^2\ln f(\theta;x)}{\partial\theta^2}\right]},$$

其中

$$\frac{[\tau'(\theta)]^2}{-nE\left[\frac{\partial^2\ln f(\theta;x)}{\partial\theta^2}\right]}$$

稱為 Cramer-Rao Lower Bound。若 Var(W) 等於 Cramer-Rao Lower Bound，則估計式 W 必為一有效估計式。此外, W 也是在所有不偏估計式中具有最小變異數者, 簡稱為 UMVUE (Uniformly Minimum Variance Unbiased Estimator)。

定理 6.2. 設 $X_1, X_2, \ldots, X_n$ 為抽自母體 $f(\theta;x)$ 之一組樣本數為 n 之樣本，其中 $f(\theta;x)$ 滿足 *Rao-Cramer* 定理之充分條件。令 $f(\theta;X)$ 為 $f(\theta;x)$ 之概似函數且 $W(X)$ 為 $\tau(\theta)$ 之不偏估計式，若

$$\alpha(\theta)[W(X)-\tau(\theta)] = \frac{\partial}{\partial\theta}\ln f(\theta;\mathbf{X})$$

則 $W(X)$ 滿足 $CRLB$，其中 $\alpha(\theta)$ 為任意函數。

證明. 由於

$$\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)\right]=0,$$

所以

$$\begin{aligned}\rho^2 &= \frac{\left[\tau'\left(\theta\right)\right]^2}{\mathrm{Var}\left(W\right)\mathrm{Var}\left(\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)\right)}\\ &= \frac{\left[\tau'\left(\theta\right)\right]^2}{\mathrm{E}\left[W-\tau\left(\theta\right)\right]^2\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)\right]^2}\\ &\leq 1\end{aligned}$$

由例題 3-19可知，當 $\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)$ 為 $W\left(X\right)-\tau\left(\theta\right)$ 之線性組合時，ρ^2 有極大值 1，即

$$\alpha\left(\theta\right)\left[W\left(X\right)-\tau\left(\theta\right)\right]=\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)$$

故可求得

$$\begin{aligned}\mathrm{Var}\left(W\right) &= \frac{\left[\tau'\left(\theta\right)\right]^2}{\mathrm{E}\left[\frac{\partial}{\partial\theta}\ln f\left(\theta;\mathbf{X}\right)\right]^2}\\ &= \mathrm{CRLB}\end{aligned}$$

4. **充分性**：若一統計量能從樣本資料中, 對被估計的母體參數提供最多訊息, 則稱此一統計量為一充分統計量，及該估計式具備有充分性。

定理 6.3. 分解定理：假設 $X_1,\ X_2,\ \ldots,\ X_n$ 為抽自母體 $f(\theta;x)$ 之一組樣本數為 n 之隨機樣本，若隨機樣本之聯合機率密度函數 $f(\theta;x_1,x_2,\ldots,x_n)$ 可分解成

$$f(x_1,x_2,\cdots,x_n;\theta)=g\left(\mu(x_1,x_2,\ldots,x_n),\theta\right)h\left(x_1,x_2,\cdots,x_n\right),$$

其中 $h(x_1,x_2,\ldots,x_n)$ 與母體參數 θ 無關，則 $\mu(x_1,x_2,\ldots,x_n)$ 為 θ 的充分統計量。

例 **6–1** 設 $X_1,\ X_2,\ \ldots,\ X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本，若 σ^2 為已知，試證明樣本數 $\overline{X}$ 為 μ 一優良估計式。

證明.

由於 $\mathrm{E}\left(\overline{X}\right)=\mu$ 且 $\lim_{n\to\infty}\mathrm{Var}\left(\overline{X}\right)=\lim_{n\to\infty}\frac{\sigma^2}{n}=0$, 所以 $\overline{X}$ 具備不偏性及一致性。又聯合機率密度函數 $f\left(x_1,x_2,\ldots,x_n\right)$ 可寫成

$$\begin{aligned}
f\left(x_1,x_2,\ldots,x_n\right) &= \frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left[-\frac{\sum_{i=1}^{n}\left(x_i-\mu\right)^2}{2\sigma^2}\right] \\
&= \frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left[-\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}+\overline{x}-\mu\right)^2}{2\sigma^2}\right] \\
&= \frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left[-\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2+\left(\overline{x}-\mu\right)^2+2\left(x_i-\overline{x}\right)\left(\overline{x}-\mu\right)}{2\sigma^2}\right] \\
&= \underbrace{\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left[-\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{2\sigma^2}\right]}_{h(x_1,x_2,\ldots,x_n)}\overbrace{\exp\left[-\frac{n\left(\overline{x}-\mu\right)^2}{2\sigma^2}\right]}^{g(\mu(x_1,x_2,\ldots,x_n),\mu)}
\end{aligned}$$

所以由分解定理可知，$\overline{X}$ 為一充分統計量。最後，由 Rao-Cramer 不等式可知

$$\ln\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}=-\frac{1}{2}\ln\left(2\pi\sigma^2\right)-\frac{\left(x-\mu\right)^2}{2\sigma^2}$$

及

$$\frac{\partial^2\ln f\left(\mu;x\right)}{\partial\mu^2}=-\frac{1}{\sigma^2}$$

得, Cramer-Rao Lower Bound 為

$$\frac{\left[\frac{d}{d\mu}\mu\right]^2}{-nE\left[\frac{\partial^2\ln f(\mu;\mathbf{X})}{\partial\theta^2}\right]}=\frac{1}{\frac{n}{\sigma^2}}=\frac{\sigma^2}{n}$$

所以 $\mathrm{Var}(\overline{X})$ 等於 Cramer-Rao Lower Bound，即 $\mathrm{Var}(\overline{X})$ 為在所有不偏估計式中具有最小變異者。所以，$\overline{X}$ 同時具備充分性及有效性，故 $\overline{X}$ 為估計 μ 之一良好估計式。

例 **6–2** Let $X_1,X_2,\ldots,X_n$ be a random sample from a distribution with the density function

$$\frac{\theta}{\left(1+x\right)^{\theta+1}},0<\theta<\infty,0\leq x<\theta.$$

Find a sufficient statistic for θ.

解：

$$
\begin{aligned}
L(\mathbf{X},\theta) &= \frac{\theta^n}{\prod_{i=1}^{n}(1+x_i)^{\theta+1}} = \theta^n \exp\left(\ln \prod_{i=1}^{n}(1+x_i)^{\theta+1}\right) \\
&= \theta^n e^{-(\theta+1)\sum_{i=1}^{n}\ln(1+x_i)}
\end{aligned}
$$

由定理 **7.2**(分解定理) 可知，$\sum_{i=1}^{n}\ln\,(1+x_i)$ 為一充分統計量。

例 **6–3** 假設 $X_1, X_2, \ldots, X_n$ 為抽自 $f(x)$ 的一組樣本數為 n 的隨機樣本，令

$$T_1 = \frac{X_1+2X_2+X_3}{4},\ T_2 = \frac{1}{4}X_1 + \frac{X_2+X_3+\cdots+X_{n-1}}{2(n-2)} + \frac{1}{4}X_n,$$

$$T_3 = \frac{X_1+X_2+\cdots+X_n}{n},\ T_4 = \frac{2}{n(n+1)}(X_1+2X_2+3X_3+\cdots+nX_n)$$

試問以何者估計母體平均數 μ 較佳？

解：

1.

$$
\begin{aligned}
\mathrm{E}(T_1) &= \mathrm{E}\left(\frac{X_1+2X_2+X_3}{4}\right) \\
&= \frac{\mathrm{E}(X_1+2X_2+X_3)}{4} \\
&= \frac{\mu+2\mu+\mu}{4} \\
&= \mu
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(T_1) &= \mathrm{Var}\left(\frac{X_1+2X_2+X_3}{4}\right) \\
&= \frac{\mathrm{Var}(X_1+2X_2+X_3)}{16} \\
&= \frac{\sigma^2+4\sigma^2+\sigma^2}{16} \\
&= \frac{3}{8}\sigma^2
\end{aligned}
$$

2.

$$\begin{aligned}
\mathrm{E}\left(T_2\right) &= \mathrm{E}\left(\frac{1}{4}X_1 + \frac{X_2 + X_3 + \cdots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n\right) \\
&= \frac{1}{4}\mu + \frac{\mathrm{E}\left(X_2 + X_3 + \cdots + X_{n-1}\right)}{2(n-2)} + \frac{1}{4}\mu \\
&= \frac{1}{4}\mu + \frac{(n-2)}{2(n-2)}\mu + \frac{1}{4}\mu \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\left(T_2\right) &= \mathrm{Var}\left(\frac{1}{4}X_1 + \frac{X_2 + X_3 + \cdots + X_{n-1}}{2(n-2)} + \frac{1}{4}X_n\right) \\
&= \frac{1}{16}\sigma^2 + \frac{\mathrm{Var}\left(X_2 + X_3 + \cdots + X_{n-1}\right)}{4(n-2)^2} + \frac{1}{16}\sigma^2 \\
&= \frac{1}{16}\sigma^2 + \frac{(n-2)}{4(n-2)^2}\sigma^2 + \frac{1}{16}\sigma^2 \\
&= \frac{\sigma^2}{8} + \frac{\sigma^2}{4(n-2)}
\end{aligned}$$

3.

$$\mathrm{E}\left(T_3\right) = \mathrm{E}\left(\overline{X}\right) = \mu$$

$$\mathrm{Var}\left(T_3\right) = \mathrm{Var}\left(\overline{X}\right) = \sigma^2$$

4.

$$\begin{aligned}
\mathrm{E}\left(T_4\right) &= \mathrm{E}\left(\frac{2}{n(n+1)}(X_1 + 2X_2 + 3X_3 + \cdots + nX_n)\right) \\
&= \frac{2}{n(n+1)}\mathrm{E}\left((X_1 + 2X_2 + 3X_3 + \cdots + nX_n)\right) \\
&= \frac{2}{n(n+1)}(\mu + 2\mu + 3\mu + \cdots + n\mu) \\
&= \frac{2}{n(n+1)}\frac{n(n+1)}{2}\mu \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\left(T_4\right) &= \mathrm{Var}\left(\frac{2}{n(n+1)}(X_1 + 2X_2 + 3X_3 + \cdots + nX_n)\right) \\
&= \frac{4}{n^2(n+1)^2}\mathrm{Var}\left((X_1 + 2X_2 + 3X_3 + \cdots + nX_n)\right) \\
&= \frac{4}{n^2(n+1)^2}(\sigma^2 + 2^2\sigma^2 + 3^2\sigma^2 + \cdots + n^2\sigma^2)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{4}{n^2(n+1)^2}\frac{n(n+1)(2n+1)}{6}\sigma^2 \\
&= \frac{2(2n+1)}{3n(n+1)}\sigma^2
\end{aligned}
$$

由上述討論可發現

$$\lim_{n\to\infty} \mathrm{Var}\,(T_1) = \frac{3}{8}\sigma^2 \neq 0,$$

$$\lim_{n\to\infty} \mathrm{Var}\,(T_2) = \frac{\sigma^2}{8} \neq 0,$$

及

$$\lim_{n\to\infty} \mathrm{Var}\,(T_3) = \lim_{n\to\infty} \mathrm{Var}\,(T_4) = 0$$

因此, T_1 及 T_2 不具備一致性而 T_3 及 T_4 具備一致性。又由於

$$
\begin{aligned}
\mathrm{Var}\,(T_4) - \mathrm{Var}\,(T_3) &= \frac{2\,(2n+1)}{3n\,(n+1)}\sigma^2 - \frac{\sigma^2}{n} \\
&= \left(\frac{2\,(2n+1)}{3\,(n+1)} - 1\right)\frac{\sigma^2}{n} \\
&= \frac{n-1}{3\,(n+1)}\frac{\sigma^2}{n} \\
&\geq 0
\end{aligned}
$$

所以，T_3 為較佳之估計式。

例 **6–4** 假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma^2)$ 的一組樣本數為 n 的隨機樣本。試證明

$$\mathrm{E}\left(S^2\right) = \sigma^2 \text{ 及 } \mathrm{Var}\left(S^2\right) = \frac{2\sigma^4}{n-1}$$

證明.

1.

$$\mathrm{E}\left(S^2\right) = \mathrm{E}\left(\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}\right) = \frac{\sigma^2}{n-1}\mathrm{E}\left(\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{\sigma^2}\right)$$

由於

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}\sim\chi^2(n-1)$$

所以 $\mathrm{E}\left(S^2\right)=\sigma^2$

2.

$$\mathrm{Var}\left(S^2\right)=\mathrm{Var}\left(\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}\right)=\frac{\sigma^4}{(n-1)^2}\mathrm{Var}\left(\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}\right)$$

又

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sigma^2}\sim\chi^2(n-1)$$

所以,

$$\mathrm{Var}\left(S^2\right)=\frac{\sigma^4}{(n-1)^2}\times 2(n-1)=\frac{2\sigma^4}{n-1}$$

例 **6–5** 假設 $X_1, X_2, \ldots, X_n$ 為一組由 beta 分配 $Beta(\theta, 1)$, $\theta>0$ 所產生的隨機樣本。令 $T=-\sum_{i=1}^{n}\log X_i/n$

1. 試求 $\mathrm{E}(T)$ 及 $\mathrm{Var}(T)$。
2. 試求 $1/\theta$ 之 UMVUE, 並試問此 UMVUE 是否達到 CRLB。

解：

1. 令 $U=\log X$，所以

$$\begin{aligned}
M_U(t) &= \mathrm{E}\left(e^{tU}\right)=\mathrm{E}\left(e^{t\log X}\right)=\mathrm{E}\left(X^t\right)\\
&= \int_0^1\frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)}x^{\theta-1}(1-x)^0x^tdx\\
&= \theta\int_0^1 x^{\theta+t-1}dx\\
&= \frac{\theta}{\theta+t}
\end{aligned}$$

由於

$$M_U'(t) = -\frac{\theta}{(\theta+t)^2} \text{ 及} M_U''(t) = \frac{2\theta}{(\theta+t)^3}$$

故可求得

$$\mathrm{E}(U) = -\frac{1}{\theta} \text{ 及} \mathrm{Var}(U) = \frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$$

又

$$T = -\overline{U} = -\frac{\sum_{i=1}^n U_i}{n}$$

所以

$$\mathrm{E}(T) = -1 \times -\frac{1}{\theta} = \frac{1}{\theta}$$

及

$$\mathrm{Var}(T) = \frac{1}{n^2} \times \frac{n}{\theta^2} = \frac{1}{n\theta^2}$$

2.

$$\begin{aligned}
\frac{\partial \log f(X;\theta)}{\partial \theta} &= \frac{\partial}{\partial \theta}\sum_{i=1}^n \log\left[\theta x^{\theta-1}\right] \\
&= \frac{\partial}{\partial \theta}\left[n\log\theta + (\theta-1)\sum_{i=1}^n \log x_i\right] \\
&= \frac{n}{\theta} + \sum_{i=1}^n \log x_i \\
&= -n\left[-\frac{\sum_{i=1}^n \log x_i}{n} - \frac{1}{\theta}\right]
\end{aligned}$$

令 $\alpha(\theta) = -1/n$，由定理6.2可知，$-\sum_{i=1}^n \log X_i/n$ 為 $1/\theta$ 之 UMVUE 且滿足 CRLB，其中

$$\begin{aligned}
\text{CRLB} &= \frac{[\tau'(\theta)]^2}{-nE\left[\partial^2 \log f(x;\theta)/\partial\theta^2\right]} \\
&= \frac{\left[-1/\theta^2\right]^2}{-n \times \left[-1/\theta^2\right]} \\
&= \frac{1}{n\theta^2}
\end{aligned}$$

例 **6–6** 假設由母體 $\mathcal{N}(\mu, 16)$ 中隨機抽取二組獨立的隨機樣本 $X_1, X_2, \ldots, X_n$ 與 $Y_1, Y_2, \ldots, Y_m$, 請分別回下列獨立的兩小題：

1. 若 $\widehat{\mu} = a\overline{X} + (1-a)\overline{Y}$ 為 μ 的不偏估計式, $0 < a < 1$，求 a 之值使 $\widehat{\mu}$ 之變異數為最小。
2. 若估計式 $\widetilde{\mu}_1 = b\overline{X}\sum_{i=1}^n\left(X_i - \overline{X}\right)^2$, $\widetilde{\mu}_2 = c\overline{X}\sum_{i=1}^m\left(Y_i - \overline{Y}\right)^2$, $\widetilde{\mu}_3 = d\overline{Y}\sum_{i=1}^n\left(X_i - \overline{X}\right)^2$, 以及 $\widetilde{\mu}_4 =$

$e\overline{Y}\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2$ 皆為 μ 的不偏估計式，試求 b, c, d, e 值。

解：

1.
$$\operatorname{Var}\left(\widehat{\mu}_1\right)=\operatorname{Var}\left(a\overline{X}+(1-a)\overline{Y}\right)=a^2\times\frac{16}{n}+(1-a)^2\times\frac{16}{m}=f\left(a\right)$$

   對上式做一階微分並令等於 0 得

$$\frac{df\left(a\right)}{da}=2a\times\frac{16}{n}-2\left(1-a\right)\times\frac{16}{m}=0$$

   所以

$$a=\frac{n}{n+m}$$

2. 由定理 **6.3** 可知，$\overline{X}$ 與 S_x^2 獨立以及 $\overline{Y}$ 與 S_y^2 獨立，故可知 $\overline{X}$ 與 $\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$ 獨立以及 $\overline{Y}$ 與 $\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2$ 獨立, 又由題目可知 X 與 Y 獨立，所以

$$\begin{aligned}\mathrm{E}\left(\widetilde{\mu}_1\right) &= \mathrm{E}\left(b\overline{X}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)=bE\left(\overline{X}\right)\mathrm{E}\left(\left(n-1\right)S_x^2\right)\\ &= b\times\mu\times(n-1)\times 16=\mu\end{aligned}$$

   得

$$b=\frac{1}{16\left(n-1\right)}$$

$$\begin{aligned}\mathrm{E}\left(\widetilde{\mu}_2\right) &= \mathrm{E}\left(c\overline{X}\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2\right)=cE\left(\overline{X}\right)\mathrm{E}\left(\left(m-1\right)S_y^2\right)\\ &= c\times\mu\times(m-1)\times 16=\mu\end{aligned}$$

   得

$$c=\frac{1}{16\left(m-1\right)}$$

$$\begin{aligned}\mathrm{E}\left(\widetilde{\mu}_3\right) &= \mathrm{E}\left(d\overline{Y}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)=dE\left(\overline{Y}\right)\mathrm{E}\left(\left(n-1\right)S_x^2\right)\\ &= d\times\mu\times(n-1)\times 16=\mu\end{aligned}$$

   得

$$d=\frac{1}{16\left(n-1\right)}$$

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{\mu}_4\right) &= \mathrm{E}\left[e\overline{Y}\sum_{i=1}^{m}\left(Y_i-\overline{Y}\right)^2\right] = eE\left(\overline{Y}\right)\mathrm{E}\left[(m-1)\,S_y^2\right] \\
&= e\times\mu\times(m-1)\times 16=\mu
\end{aligned}
$$

得

$$
e=\frac{1}{16\,(m-1)}
$$

6.1.1 最大概似法

假設 X_1, X_2, ..., X_n 為抽自母體 $f(\theta;x)$ 的一組樣本數為 n 的隨機樣本，其中 θ 為未知母體參數，$f\left(\theta;x_1,x_2,\ldots,x_n\right)$ 為 X_1, X_2, ..., X_n 之聯合機率密度函數，則稱

$$
\begin{aligned}
L\left(\theta;x_1,x_2,\ldots,x_n\right) &= f\left(\theta;x_1,x_2,\ldots,x_n\right) \\
&= f(\theta;x_1)f(\theta;x_2)\cdots f(\theta;x_n)
\end{aligned}
$$

為概似函數。若能找到一 θ 之估計式，$\widehat{\theta}=\theta(x_1,x_2,\ldots,x_n)$，使得 $L\left(\theta;x_1,x_2,\ldots,x_n\right)$ 有最大值，則稱 $\widehat{\theta}$ 為 θ 之**最大概似估計式**(Maximum Likelihood estimator)，簡稱 MLE。

例 **6–7** X_1, X_2, ..., X_n 為抽自 $Berounlli(p)$ 的一組樣本數為 n 的隨機樣本，試求 p 之 MLE 為何？

解：

$$
\begin{aligned}
L(p;x) &= p^{x_1}(1-p)^{1-x_1}p^{x_2}(1-p)^{1-x_2}\ldots p^{x_n}(1-p)^{1-x_n} \\
&= p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}
\end{aligned}
$$

所以

$$
\mathrm{ln}L(p;x)=\left(\sum_{i=1}^{n}x_i\right)\mathrm{ln}p+\left(n-\sum_{i=1}^{n}x_i\right)\mathrm{ln}(1-p)
$$

今欲使得 $\mathrm{ln}L(p;x)$ 有最大值，即求解

$$
\frac{d\mathrm{ln}L(p;x)}{dp}=\frac{\sum_{i=1}^{n}x_i}{p}-\frac{n-\sum_{i=1}^{n}x_i}{1-p}=0
$$

求解上式得

$$
\widehat{p}=\frac{\sum_{i=1}^{n}x_i}{n}
$$

例 **6–8** $X_1, X_2, \ldots, X_n$ 為抽自 $Poisson(\lambda)$ 的一組樣本數為 n 的隨機樣本，試求 λ 之 MLE 為何？

解：

$$
\begin{aligned}
L(x;p) &= \frac{\lambda^{x_1}e^{-\lambda}}{x_1!}\frac{\lambda^{x_2}e^{-\lambda}}{x_2!}\cdots\frac{\lambda^{x_n}e^{-\lambda}}{x_n!} \\
&= \frac{\lambda^{\sum_{i=1}^{n} x_i}e^{-n\lambda}}{\prod\limits_{i=1}^{n} x_i}
\end{aligned}
$$

所以

$$
\ln L(x;\lambda) = \left(\sum_{i=1}^{n} x_i\right)\ln\lambda - n\lambda - \ln\prod_{i=1}^{n} x_i
$$

今欲使得 $\ln L(x;\lambda)$ 有最大值，即求解

$$
\frac{d\ln L(x;\lambda)}{d\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0
$$

求解上式得

$$
\widehat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}
$$

例 **6–9** The density for each obsearvation $y_i = 1, 2, \ldots, n$ is

$$
f\left(y_i|\theta\right) = \frac{e^{-\theta}\theta^{y_i}}{y_i!}
$$

1. Write down the log likelihood function.
2. Consider a random sample 5, 0, 1, 1, 0, 3, 2, 4, 1. Please use this data set to obtain the maximum likelihood estimate of θ.

解：

1.

$$
L\left(\theta|y_i\right) = \prod_{i=1}^{n}\frac{e^{-\theta}\theta^{y_i}}{y_i!} = \frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}
$$

$$
\log L\left(\theta|y_i\right) = \log\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!} = -n\theta + \sum_{i=1}^{n} y_i \log\theta - \sum_{i=1}^{n}\log y_i
$$

2. 由上例可知，$\widehat{\theta}=\sum_{i=1}^{n} y_i/n=(5+0+1+1+0+3+2+4+1)/9=1.8889$

例 **6–10** $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本，試求 μ 及 σ^2 之 MLE 為何？

解：

$$\begin{aligned} L(x;\mu,\sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right] \\ &= (2\pi)^{-\frac{n}{2}} \left(\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\right] \end{aligned}$$

所以

$$\ln L(x;\mu,\sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}$$

今欲使得 $\ln L(x;\lambda)$ 有最大值，即求解

$$\frac{\partial \ln L(x;\mu,\sigma^2)}{\partial\mu} = -2\sum_{i=1}^{n} \frac{(x_i-\mu)}{2\sigma^2} \times -1 = 0$$

及

$$\frac{\partial \ln L(x;\mu,\sigma^2)}{\partial\sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\left(\sigma^2\right)^2} \times -1 = 0$$

求解上式得

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} \text{ 及 } \widehat{\sigma}^2 = \frac{\sum_{i=1}^{n} (x_i-\overline{x})^2}{n}$$

又

$$\begin{aligned} \mathrm{E}\left(\widehat{\sigma}^2\right) &= \mathrm{E}\left[\frac{\sum_{i=1}^{n} (x_i-\overline{x})^2}{n}\right] \\ &= \mathrm{E}\left[\frac{(n-1)S^2}{n}\right] \\ &= \frac{n-1}{n}\sigma^2 \end{aligned}$$

所以, 由此範例可發現 $\widehat{\sigma}^2$ 並非為 σ^2 之不偏估計式。因此, 由最大概似法所推得之最大概似估計式並不完全是不偏估計式, 但可以由轉換而成為不偏估計式。

例 **6–11** $X_1, X_2, \ldots, X_n$ 為抽自 $U(0,\theta)$, $0<x<\theta<\infty$, 的一組樣本數為 n 的隨機樣本，試求 θ 之 MLE 為何？

解：

$$L(\theta;x) \;=\; \prod_{i=1}^{n}\frac{1}{\theta}=\frac{1}{\theta^n}$$

對 $\ln L(\theta;x)$ 做一階微分得

$$\frac{d\ln L(\theta;x)}{d\theta}=-\frac{n}{\theta^{n+1}}<0$$

所以，$L(\theta;x)$ 為 θ 之遞減函數，θ 愈大，則 $L(\theta;x)$ 愈小。又 $0<x<\theta<\infty$，所以 $\widehat{\theta}=\max\{x_1,x_2,\ldots,x_n\}$。

例 **6–12** Find the maximum likelihood estimate (MLE) which is based on a random sample X_1, X_2,..., X_n from the pdf $f(x;\theta_1,\theta_2)=1/(\theta_2-\theta_1)$; $\theta_1\le x\le\theta_2$, zero otherwise.

解：

$$L(x;\theta_1,\theta_2)=\left(\frac{1}{\theta_2-\theta_1}\right)^n$$

左右取對數可得

$$\ln L(x;\theta_1,\theta_2)=-n\ln(\theta_2-\theta_1)$$

又由於

$$\frac{\partial\ln L(x;\theta_1,\theta_2)}{\partial\theta_1}=\frac{n}{\theta_2-\theta_1}>0 \text{ 及} \frac{\partial\ln L(X)}{\partial\theta_2}=\frac{-n}{\theta_2-\theta_1}<0$$

故可知 $\ln L(x;\theta_1,\theta_2)$ 為 θ_2 之嚴格遞增函數及 $\ln L(x;\theta_1,\theta_2)$ 為 θ_1 之嚴格遞減函數，又 $\theta_1\le x\le\theta_2$，故可求得

$$\widehat{\theta}_1=\min\{X_1,X_2,\ldots,X_n\} \text{ 及} \widehat{\theta}_2=\max\{X_1,X_2,\ldots,X_n\}$$

最後，由定理 7.1 可知，若 $f(\theta;x_1,x_2,\ldots,x_n)$ 可寫成

$$f(\theta;x_1,x_2,\ldots,x_n)=g(\mu(x_1,x_2,\ldots,x_n),\theta)\,h(x_1,x_2,\ldots,x_n),$$

則 $\mu(x_1,x_2,\ldots,x_n)$ 即為 θ 之充分統計量。由上式可知，由於 $h(x_1,x_2,\ldots,x_n)$ 與 θ 無關，因此若有

一 $\widehat{\theta}$ 能使得 g 有最大值，則必定也可以使得 $L(\mu(x_1,x_2,\ldots,x_n),\theta)$ 有最大值。由概似函數之定義可知，$\widehat{\theta}$ 即為母體參數 θ 之 MLE。若 $\widehat{\theta}$ 為 $\mu(x_1,x_2,\ldots,x_n)$ 之一對一函數，則 $\widehat{\theta}$ 亦為一充分統計量。

此外，MLE 估計式具備不變性，一致性及漸近有效性，

1. **不變性**：若 $\widehat{\theta}$ 為 θ 之 MLE，則 $\tau\left(\widehat{\theta}\right)$ 為 $\tau\left(\theta\right)$ 之 MLE。
2. **一致性**：若 $\widehat{\theta}$ 為 θ 之 MLE，則

$$\lim_{n\to\infty}\mathrm{P}\left(\left|\widehat{\theta}-\theta\right|<\varepsilon\right)=1$$

由不變性可知

$$\lim_{n\to\infty}\mathrm{P}\left(\left|\tau\left(\widehat{\theta}\right)-\tau\left(\theta\right)\right|<\varepsilon\right)=1$$

3. **漸近有效性**：若 $\mathrm{E}(W)=\tau(\theta)$，且

$$\lim_{n\to\infty}\frac{\mathrm{Var}\left(W\right)}{\frac{\left[\tau'(\theta)\right]^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(\theta;x)\right]^2}}=\lim_{n\to\infty}\frac{\mathrm{Var}\left(W\right)}{\frac{\left[\tau'(\theta)\right]^2}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}=1$$

則稱 W 為 $\tau(\theta)$ 之一有效漸近估計式。故當 $n\to\infty$ 時，$\mathrm{Var}(W)$ 等於 Cramer-Rao Lower Bound。故可推得

$$\mathrm{Var}\left(W\right)\approx\frac{\left[\tau'\left(\theta\right)\right]^2}{\mathrm{E}\left[-\frac{\partial^2}{\partial\theta^2}\ln L\left(\mathbf{X};\theta\right)\right]}$$

又由最大概似估計法及泰勒展開式可知

$$\frac{\partial\ln L\left(\mathbf{X};\widehat{\theta}\right)}{\partial\theta}\approx\frac{\partial\ln L\left(\mathbf{X};\theta\right)}{\partial\theta}+\left(\widehat{\theta}-\theta\right)\frac{\partial^2\ln L\left(\mathbf{X};\theta\right)}{\partial\theta^2}\approx 0$$

化簡後可得

$$\left(\widehat{\theta}-\theta\right)=-\frac{\partial\ln L\left(\mathbf{X};\theta\right)}{\partial\theta}\bigg/\frac{\partial^2\ln L\left(\mathbf{X};\theta\right)}{\partial\theta^2}$$

由先前推導可知，$\widehat{\theta}$ 之變異數為

$$\frac{1}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L\left(\mathbf{X};\theta\right)\right]}$$

所以，若 $\widehat{\theta}$ 為 θ 之不偏估計式可得

$$\lim_{n\to\infty}\frac{\widehat{\theta}-\theta}{\sqrt{\frac{1}{-E\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}}\xrightarrow{d}\mathcal{N}(0,1)$$

又

$$\lim_{n\to\infty}\frac{\widehat{\theta}-\theta}{\sqrt{\frac{1}{-E\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}}=\lim_{n\to\infty}\frac{-\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\frac{\partial^2\ln L(\mathbf{X};\theta)}{\partial\theta^2}}}{\sqrt{\frac{1}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}}=\lim_{n\to\infty}\frac{\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}}{\frac{-\frac{\partial^2\ln L(\mathbf{X};\theta)}{\partial\theta^2}}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}},$$

其中

$$\mathrm{E}\left[\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}\right]=\frac{\mathrm{E}\left[\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}\right]}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}=0$$

且

$$\begin{aligned}\mathrm{Var}\left[\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}\right] &= \mathrm{E}\left[\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}\right]^2\\ &= \frac{\mathrm{E}\left[\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}\right]^2}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}=\frac{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}\\ &= 1\end{aligned}$$

所以

$$\lim_{n\to\infty}\frac{\frac{\partial\ln L(\mathbf{X};\theta)}{\partial\theta}}{\sqrt{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}}\xrightarrow{d}\mathcal{N}(0,1)$$

由 Slutsky's Theorem (第140頁) 可知，

$$\lim_{n\to\infty}\frac{-\frac{\partial^2\ln L(\mathbf{X};\theta)}{\partial\theta^2}}{-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]}=1$$

即

$$\lim_{n\to\infty}-\frac{\partial^2\ln L(\mathbf{X};\theta)}{\partial\theta^2}=\lim_{n\to\infty}-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]$$

或

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta) - \mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]\right| < \varepsilon\right) = 1$$

因此，$-\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)$ 可做為 $-\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]$ 之一致估計式。又 MLE 具不變性及一致性，故上述近似式可改寫為

$$\begin{aligned}\mathrm{Var}(W) &\approx \frac{[\tau'(\theta)]^2}{\mathrm{E}\left[-\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\right]} \\ &\approx \frac{[\tau'(\theta)]^2|_{\theta=\widehat{\theta}}}{-\frac{\partial^2}{\partial\theta^2}\ln L(\mathbf{X};\theta)\Big|_{\theta=\widehat{\theta}}}\end{aligned}$$

並稱此一近似式為 W 之漸近變異數(Asymptotic Variance)。

例 **6–13** Consider a random sample $\{X_1, X_2, \ldots, X_n\}$ from a distribution with density function

$$f(x|\alpha) = \frac{1}{2\alpha}\exp\left(-\frac{|x|}{\alpha}\right), -\infty < x < \infty.$$

1. Find the maximum likelihood estimator of α.
2. Find the asymptotic variance of MLE.

解：

1.

$$\prod_{i=1}^n f(x_i|\alpha) = \prod_{i=1}^n \frac{1}{2\alpha}\exp\left(-\frac{|x_i|}{\alpha}\right) = 2^{-n}\alpha^{-n}\exp\left(-\frac{\sum_{i=1}^n |x_i|}{\alpha}\right)$$

所以

$$\frac{d\ln L(\mathbf{X},\alpha)}{d\alpha} = -\frac{n}{\alpha} + \frac{\sum_{i=1}^n |x_i|}{\alpha^2} = 0$$

可求得

$$\widehat{\alpha} = \frac{\sum_{i=1}^n |x_i|}{n}$$

2.

$$\begin{aligned}\ln L(\mathbf{X},\alpha) &= -n\ln 2 - n\ln\alpha - \frac{\sum_{i=1}^n |x_i|}{\alpha} \\ &= -n\ln 2 - n\ln\alpha - \frac{n\widehat{\alpha}}{\alpha}\end{aligned}$$

上式對 α 做二階微分得

$$\left.\frac{\partial^2 \ln L(\mathbf{X},\alpha)}{\partial \alpha^2}\right|_{\alpha=\widehat{\alpha}} = \frac{n}{\widehat{\alpha}^2} - \frac{2n\widehat{\alpha}}{\widehat{\alpha}^3} = -\frac{n}{\widehat{\alpha}^2}$$

所以 $\widehat{\alpha}$ 之漸近變異數為

$$\operatorname{Var}(\widehat{\alpha}) \approx \frac{1}{-\left.\frac{\partial^2 \ln L(\mathbf{X},\alpha)}{\partial \alpha^2}\right|_{\alpha=\widehat{\alpha}}} = \frac{\widehat{\alpha}^2}{n}$$

例 **6–14** If gene frequencies are in equilibrium, the gene types AA, Aa and aa occur with probabilities $(1-\theta)^2$, $2\theta(1-\theta)$ and θ^2, respectively. Plato el al. (1964) published the following data on haptoglobin type in a sample of 200 people.

Haptoglobin type		
Hp1-1 (AA)	Hp1-2 (Aa)	Hp2-2 (aa)
14	52	134

1. Find the MLE of θ.
2. Find the asymptotic variance of MLE.
3. Find an approximate 95% confidence interval for θ.

解：

1. 令 y_1 表 AA 發生次數，y_2 表 Aa 發生次數，則概似函數為

$$L(\mathbf{X},\theta) = \left[(1-\theta)^2\right]^{y_1} \left[2\theta(1-\theta)\right]^{y_2} \left[\theta^2\right]^{n-y_1-y_2}$$

取對數得

$$\begin{aligned}\ln L(\mathbf{X},\theta) &= 2y_1 \ln(1-\theta) + y_2 \ln 2 + y_2 \ln\theta + y_2 \ln(1-\theta) + 2n - 2y_1 - 2y_2 \ln\theta \\ &= y_2 \ln 2 + (2y_1 + y_2)\ln(1-\theta) + (2n - 2y_1 - y_2)\ln\theta\end{aligned}$$

所以

$$\frac{d\ln L(\mathbf{X},\theta)}{d\theta}=\frac{2n-2y_1-y_2}{\theta}-\frac{2y_1+y_2}{1-\theta}=0$$

故可求得

$$\frac{2n-2y_1-y_2-2n\theta+2y_1\theta+y_2\theta-2y_1\theta-y_2\theta}{\theta(1-\theta)}=\frac{2n-2y_1-y_2-2n\theta}{\theta(1-\theta)}=0$$

得

$$\widehat{\theta}=\frac{2n-2y_1-y_2}{2n}$$

2.

$$\frac{d\ln L(\mathbf{X},\theta)}{d\theta}=\frac{2n\widehat{\theta}}{\theta}-\frac{2n\left(1-\widehat{\theta}\right)}{1-\theta}$$

及

$$\begin{aligned}\left.\frac{d^2\ln L(\mathbf{X},\theta)}{d\theta^2}\right|_{\theta=\widehat{\theta}} &= -\frac{2n\widehat{\theta}}{\theta^2}-\left.\frac{2n\left(1-\widehat{\theta}\right)}{(1-\theta)^2}\right|_{\theta=\widehat{\theta}}\\ &= -\frac{2n}{\widehat{\theta}}-\frac{2n}{1-\widehat{\theta}}\\ &= -\frac{2n}{\widehat{\theta}\left(1-\widehat{\theta}\right)}\end{aligned}$$

所以 $\widehat{\theta}$ 之漸近變異數為

$$\operatorname{Var}\left(\widehat{\theta}\right)\approx\frac{\widehat{\theta}\left(1-\widehat{\theta}\right)}{2n}$$

3.

$$\widehat{\theta}=\frac{2n-2y_1-y_2}{2n}=\frac{2\times200-2\times14-52}{2\times200}=0.8$$

及

$$\operatorname{Var}\left(\widehat{\theta}\right)\approx\frac{\widehat{\theta}\left(1-\widehat{\theta}\right)}{2n}=\frac{0.8\times0.2}{2\times200}=0.0004$$

所以 θ 之 95% 信賴區間為

$$\begin{aligned}&\left[\widehat{\theta}-z_{0.025}\sqrt{\operatorname{Var}\left(\widehat{\theta}\right)},\widehat{\theta}+z_{0.025}\sqrt{\operatorname{Var}\left(\widehat{\theta}\right)}\right]\\ =&\left[0.8-1.96\sqrt{0.0004},0.8+1.96\sqrt{0.0004}\right]\\ =&[0.7608,0.8392]\end{aligned}$$

6.1.2 動差法

$X_1, X_2, \ldots, X_n$ 為抽自母體 $f(\theta; x)$ 的一組樣本數為 n 的隨機樣本，θ 為未知母體參數，得樣本之 r 階動差, M_r, 為

$$M_r = \frac{\sum_{i=1}^{n} X_i^r}{n}$$

由於

$$\mathrm{E}\left(M_r\right) = \mathrm{E}\left(\frac{\sum_{i=1}^{x} x_i^r}{n}\right) = \sum_{i=1}^{n} \mathrm{E}\left(\frac{X_i^r}{n}\right) = \mathrm{E}\left(X^r\right)$$

可知樣本之 r 階動差為母體 r 階動差之不偏估計式, 利用此一觀念求得估計式之方法稱為動差法。

例 **6–15** $X_1, X_2, \ldots, X_n$ 為抽自 $\mathrm{Gamma}(\alpha, \theta)$ 的一組樣本數為 n 的隨機樣本，試以動差法求解 α 及 θ 之估計式。

解：

由於 $\mathrm{E}X = \alpha\theta$ 及 $\mathrm{Var}(X) = \alpha\theta^2$，所以

$$\mathrm{E}X = \alpha\theta = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x},$$

$$\mathrm{E}X^2 = \alpha\theta^2 + \alpha^2\theta^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} = v$$

同時求解上述二式可得

$$\widetilde{\theta} = \frac{v}{\overline{x}} - \overline{x} = \frac{\sum_{i=1}^{n} x_i^2}{n\overline{x}} - \overline{x}, \widetilde{\alpha} = \frac{\overline{x}^2}{v - \overline{x}^2} = \frac{n\overline{x}^2}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}$$

例 **6–16** $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma^2)$ 的一組樣本數為 n 的隨機樣本，試以動差法求解 μ 及 σ^2 之估計式？

解：

由於 $\mathrm{E}X=\mu$ 及 $\mathrm{E}X^2=\sigma^2+\mu^2$，所以

$$\mu=\frac{\sum_{i=1}^n x_i}{n},\ \sigma^2+\mu^2=\frac{\sum_{i=1}^n x_i^2}{n}$$

求解上式得

$$\widetilde{\mu}=\frac{\sum_{i=1}^n x_i}{n}\text{ 及 }\widetilde{\sigma}^2=\frac{\sum_{i=1}^n x_i^2-n\overline{x}^2}{n}=\frac{\sum_{i=1}^n\left(x_i-\overline{x}\right)^2}{n}$$

例 **6–17** Suppose X is a random sample from X, where X has density function

$$f\left(x|\theta\right)=\begin{cases} x/\theta\exp\left(-x^2/2\theta\right) & ,x>0\\ 0 & ,\text{otherwise}\end{cases}$$

where $\theta>0$.

1. Find the maximum likelihood estimate of θ
2. Find the moment estimate of θ

解：

1. $$L\left(\mathbf{X},\theta\right)=\prod_{i=1}^n\frac{x_i}{\theta}\exp\left(-\frac{x_i^2}{2\theta}\right)=\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2\theta}\right)\frac{\prod_{i=1}^n x_i}{\theta^n}$$

   取對數得

   $$\ln L\left(\mathbf{X},\theta\right)=-\frac{\sum_{i=1}^n x_i^2}{2\theta}+\ln\prod_{i=1}^n x_i-n\ln\theta$$

   對 $\ln L\left(\mathbf{X},\theta\right)$ 做一階微分

   $$\frac{d\ln L\left(\mathbf{X},\theta\right)}{d\theta}=\frac{\sum_{i=1}^n x_i^2}{2\theta^2}-\frac{n}{\theta}=0$$

   故可求得

   $$\widehat{\theta}=\frac{\sum_{i=1}^n x_i^2}{2n}$$

2. $$\mathrm{E}X=\int_0^\infty x\frac{x}{\theta}e^{-\frac{x^2}{2\theta}}dx=\int_0^\infty\frac{x^2}{\theta}e^{-\frac{x^2}{2\theta}}dx$$

   $$令\frac{x^2}{\theta}=t,可得dx=\frac{1}{2}\left(\theta t\right)^{-\frac{1}{2}}\theta dt$$

所以

$$
\begin{aligned}
\int_0^\infty \frac{x^2}{\theta} e^{-\frac{x^2}{2\theta}} dx &= \frac{\theta^{\frac{1}{2}}}{2} \int_0^\infty t^{\frac{1}{2}} e^{-\frac{t}{2}} dt \\
&= \frac{\theta^{\frac{1}{2}}}{2} \times 2^{\frac{1}{2}} \times 2 \times \int_0^\infty \left(\frac{t}{2}\right)^{\frac{1}{2}} e^{-\frac{t}{2}} d\left(\frac{t}{2}\right) \\
&= \sqrt{2\theta} \times \Gamma\left(\frac{3}{2}\right) = \sqrt{2\theta} \times \frac{\sqrt{\pi}}{2} \\
&= \sqrt{\frac{\pi\theta}{2}}
\end{aligned}
$$

$$
\frac{\sum_{i=1}^n x_i}{n} = \mathrm{E}X = \sqrt{\frac{\pi\theta}{2}}
$$

因此

$$
\widetilde{\theta} = \frac{2\overline{x}^2}{\pi}
$$

例 **6–18**

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from X, where X has density

$$
f(x|\theta) = \begin{cases} (1+\alpha x)/2, & -1 \le x \le 1, \ -1 \le \alpha \le 1 \\ 0, & \text{otherwise} \end{cases}
$$

1. Prove that $f(x)$ is a probability density function for and $-1 \le \alpha \le 1$.
2. Find the moment estimate of θ

解：

1. 由於 $-1 \le x \le 1$ 及 $-1 \le \alpha \le 1, -1 \le \alpha x \le 1$，所以

$$
0 \le \frac{1+\alpha x}{2} \le 1.
$$

又

$$
\int_{-1}^1 \frac{1+\alpha x}{2} dx = 1
$$

因此，由機率公理可知 $f(x)$ 為一機率密度函數。

2.

$$
\mathrm{E}X = \int_{-1}^1 \frac{x+\alpha x^2}{2} dx = \frac{\alpha}{3}
$$

所以

$$\overline{x}=\frac{\alpha}{3},\ \text{即}\ \widetilde{\alpha}=3\overline{x}$$

例 **6–19** Suppose that $X_1,\ldots,X_n$ is a random sample from a distribution for which the pdf $f(x|\lambda)$ is as follows:

$$f(x|\lambda)=\begin{cases}\lambda x^{\lambda-1}, & \text{for } 0<x<1\\ 0, & \text{otherwise}\end{cases}$$

Also, suppose that the value of λ is unknown $(\lambda>0)$. Find the maximum likelihood estimator (MLE) of λ.

解：

$$L(X,\lambda)=\lambda^n\Pi_{i=1}^n x_i^{\lambda-1}$$

$$\begin{aligned}\ln L(X,\lambda) &= n\ln\lambda+(\lambda-1)\ln\Pi_{i=1}^n x_i\\ &= n\ln\lambda+(\lambda-1)\sum_{i=1}^n \ln x_i\end{aligned}$$

$$\frac{d\ln L(X,\lambda)}{d\lambda}=\frac{n}{\lambda}+\sum_{i=1}^n\ln x_i=0$$

$$\widehat{\lambda}=-\frac{n}{\sum_{i=1}^n\ln x_i}$$

例 **6–20** 設 X_1, ...,X_n 為 iid $Beronulli(p)$，則 $Y=\sum_{i=1}^n X_i$ 是 $binomial(n,p)$。假設 p 的事前分配為 $Beta(\alpha,\beta)$，試求在給定 $Y=y$ 的情況下，p 的事後機率分配。

解：由聯合機率密度函數之定義可知

$$\begin{aligned}f(y,p) &= \binom{n}{y}p^y(1-p)^{n-y}\times\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}\\ &= \frac{n!}{y!(n-y)!}p^y(1-p)^{n-y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}\\ &= \frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{y+\alpha-1}(1-p)^{n-y+\beta-1}\end{aligned}$$

又 Y 的邊際機率密度函數為

$$\begin{aligned}f(y) &= \int_0^1\frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha+y-1}(1-p)^{n-y+\beta-1}dp\\ &= \frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+y)\Gamma(n-y+\beta)}{\Gamma(\alpha+\beta+n)}\end{aligned}$$

故可知

$$\begin{aligned}
f\left(p|y\right) &= \frac{\frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}}{\frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+y)\Gamma(n-y+\beta)}{\Gamma(\alpha+\beta+n)}}p^{y+\alpha-1}\left(1-p\right)^{n-y+\beta-1} \\
&= \frac{\Gamma\left(\alpha+\beta+n\right)}{\Gamma\left(\alpha+y\right)\Gamma\left(n-y+\beta\right)}p^{y+\alpha-1}\left(1-p\right)^{n-y+\beta-1}
\end{aligned}$$

得 $P|Y\sim\mathrm{Beta}\left(\alpha+y,n-y+\beta\right)$，其條件期望值為 $\mathrm{E}\left(P|Y\right)=\frac{\alpha+y}{\alpha+y+n-y+\beta}=\frac{\alpha+y}{\alpha+n+\beta}$。此一估計式稱為 p 的貝式估計量。

例 **6–21** 設 $X\sim\mathcal{N}(\theta,\sigma^2)$，且 θ 的事前分配為 $\mathcal{N}(a,b^2)$，試求在給定 $X=x$ 的情況下，θ 的事後機率分配。

解：

$$\begin{aligned}
f\left(x,\theta\right) &= \frac{\exp\left[-\frac{1}{2\sigma^2}\left(x-\theta\right)^2\right]}{\sqrt{2\pi\sigma^2}}\frac{\exp\left[-\frac{1}{2b^2}\left(\theta-a\right)^2\right]}{\sqrt{2\pi b^2}} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\times\frac{1}{\sqrt{2\pi b^2}}\times\exp\left[-\frac{1}{2\sigma^2}\left(x-\theta\right)^2-\frac{1}{2b^2}\left(\theta-a\right)^2\right] \\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\times\frac{1}{\sqrt{2\pi b^2}}\times\exp\left[-\frac{b^2\left(x-\theta\right)^2+\sigma^2\left(\theta-a\right)^2}{2\sigma^2b^2}\right]
\end{aligned}$$

其中

$$\begin{aligned}
b^2\left(x-\theta\right)^2+\left(\theta-a\right)^2 &= b^2x^2+b^2\theta^2-2xb^2\theta+\sigma^2\theta^2+\sigma^2a^2-2a\sigma^2\theta \\
&= \left(b^2+\sigma^2\right)\theta^2-2\left(xb^2+a\sigma^2\right)+\left(b^2x^2+\sigma^2a^2\right) \\
&= \left[\sqrt{b^2+\sigma^2}\theta-\frac{\left(xb^2+a\sigma^2\right)}{\sqrt{b^2+\sigma^2}}\right]^2+\left(b^2x^2+\sigma^2a^2\right)-\frac{\left(xb^2+a\sigma^2\right)^2}{b^2+\sigma^2} \\
&= \left(b^2+\sigma^2\right)\left[\theta-\frac{\left(xb^2+a\sigma^2\right)}{b^2+\sigma^2}\right]^2+\left(b^2x^2+\sigma^2a^2\right)-\frac{\left(xb^2+a\sigma^2\right)^2}{b^2+\sigma^2}
\end{aligned}$$

故上式可改寫為

$$\begin{aligned}
f\left(x,\theta\right) &= \frac{1}{\sqrt{2\pi\sigma^2}}\times\frac{1}{\sqrt{2\pi b^2}}\times\exp\left\{-\frac{b^2+\sigma^2}{2\sigma^2b^2}\left[\theta-\frac{\left(xb^2+a\sigma^2\right)}{b^2+\sigma^2}\right]^2\right\} \\
&\quad\times\exp\left[\frac{-\left(b^2x^2+\sigma^2a^2\right)+\frac{\left(xb^2+a\sigma^2\right)^2}{b^2+\sigma^2}}{2\sigma^2b^2}\right]
\end{aligned}$$

又 X 的邊際機率密度函數為

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \times \frac{1}{\sqrt{2\pi b^2}} \times \exp\left[\frac{-\left(b^2x^2+\sigma^2a^2\right)+\frac{\left(xb^2+a\sigma^2\right)^2}{b^2+\sigma^2}}{2\sigma^2b^2}\right] \times \sqrt{2\pi\frac{2\sigma^2b^2}{b^2+\sigma^2}}$$

因此可求得事後機率分配為

$$\begin{aligned} f(\theta|x) &= \frac{1}{\sqrt{2\pi\frac{2\sigma^2b^2}{b^2+\sigma^2}}} \times \exp\left\{-\frac{b^2+\sigma^2}{2\sigma^2b^2}\left[\theta-\frac{xb^2+a\sigma^2}{b^2+\sigma^2}\right]^2\right\} \\ &= \frac{1}{\sqrt{2\pi\frac{2\sigma^2b^2}{b^2+\sigma^2}}} \times \exp\left\{-\frac{1}{2\frac{\sigma^2b^2}{b^2+\sigma^2}}\left[\theta-\frac{xb^2+a\sigma^2}{b^2+\sigma^2}\right]^2\right\} \end{aligned}$$

故可推得 $\theta|x \sim N\left(\frac{xb^2+a\sigma^2}{b^2+\sigma^2}, \frac{\sigma^2b^2}{b^2+\sigma^2}\right)$。

6.2 區間估計

區間估計乃指根據觀察值 $X_1,\ X_2,\ldots,X_n$ 與運用抽樣分配及機率之原理, 來預測一未知參數 θ 在某種程度下可能所在的範圍的方法。以機率式表示：

$$\mathrm{P}(L \le \theta \le U) = 1-\alpha,$$

其中

1. $1-\alpha$ 稱為**信賴水準**（Confidence Level）
2. $(L,\ U)$ 稱為 θ 的 $(1-\alpha)\times 100\%$ 信賴區間
3. L 稱為信賴區間的信賴上限
4. U 稱為信賴區間的信賴上限

例 **6–22** Let X_1, X_2,..., X_n be a random sample from an exponenital distribution, $X_i \sim \exp(\theta)$.

1. If $\overline{x} = 17.9$ with $n = 50$, find a one-side lower 95% confidence limit for θ.
2. Find a one-side lower 95% confidence limit for $\mathrm{P}(X > t) = \exp(-t/\theta)$ where t is an arbitrary know value.

解：

1. 由於 $X_i \sim \exp(\theta)$，所以 $\mathrm{E}X = \theta$ 及 $\mathrm{Var}(X) = \theta^2$。又樣本數 $n = 50 \geq 30$ 為大樣本，由中央極限定理可知

$$\mathrm{P}\left(\frac{\overline{x}-\theta}{\theta/\sqrt{n}} > -Z_{0.05}\right) = 0.95$$

故

$$\mathrm{P}\left(\overline{x}-\theta > -\frac{Z_{0.05}}{\sqrt{n}}\theta\right) = \mathrm{P}\left(\theta < \frac{\overline{x}}{1-Z_{0.05}/\sqrt{n}}\right) = \mathrm{P}\left(\theta < 23.331\right) = 0.95$$

由信賴區間之定義可知，θ 之左尾 95% 信賴區間為 $[0, 23.331]$

2.
$$\mathrm{P}\left(\theta > 23.331\right) = \mathrm{P}\left(-\frac{t}{\theta} < -\frac{t}{23.331}\right) = \mathrm{P}\left(e^{-t/\theta} < e^{-t/23.331}\right) = 0.95$$

由信賴區間之定義可知，θ 之左尾 95% 信賴區間為 $[0, e^{-t/18.133}]$

6.2.1 母體變異數已知, 單一母體平均數之區間估計

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma^2)$ 的一組樣本數為 n 的隨機樣本。若母體變異數 σ^2 已知，則由定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}}\right) \\
&= \mathrm{P}\left(-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \overline{X}-\mu \leq z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) \\
&= \mathrm{P}\left(\overline{X}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right),
\end{aligned}$$

其中, $\frac{\alpha}{2} = \mathrm{P}\left(Z > z_{\frac{\alpha}{2}}\right)$。所以, $\left(\overline{x}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{x}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$ 為母體參數 μ 在信賴水準為 $(1-\alpha)\times 100\%$ 下眾多信賴區間中其中一組信賴區間。此信賴區間之上下限分別為

$$L = \overline{X}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \text{ 及 } U = \overline{X}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

除此之外, $\left(\overline{X}-z_{\frac{\alpha}{2}}\sigma/\sqrt{n}, \overline{X}+z_{\frac{\alpha}{2}}\sigma/\sqrt{n}\right)$ 為在信賴水準 $(1-\alpha)\times 100\%$ 下具有最小誤差界限之信賴區間。由定義可知

$$1-\alpha = \mathrm{P}\left(T_1 < \mu < T_2\right)$$

$$= \quad \mathrm{P}\left(\frac{\overline{X}-T_2}{\sigma/\sqrt{n}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{\overline{X}-T_1}{\sigma/\sqrt{n}}\right)$$

其中 T_1 及 T_2 分別為信賴上限及信賴下限。因此, 本問題即為求解下列最佳化問題

$$\min \quad T_2 - T_1$$

$$\text{subject} \quad \int_{\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}}^{\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1-\alpha$$

由 Lagrange 方法可將上述問題寫成

$$\min \quad L(T_1, T_2) = T_2 - T_1 - \lambda \Big\{ \int_{\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}}^{\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz - (1-\alpha) \Big\}$$

分別對 T_1, T_2 及 λ 做偏微分並令其等於零可得

$$\frac{\partial L}{\partial T_2} = 1 - \lambda \frac{1}{\sigma/\sqrt{n}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}\right)^2} = 0,$$

$$\frac{\partial L}{\partial T_1} = -1 + \lambda \frac{1}{\sigma/\sqrt{n}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}\right)^2} = 0,$$

及

$$\frac{\partial L}{\partial \lambda} = -\int_{\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}}^{\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + (1-\alpha) = 0$$

求解上述方程式可得

$$\left(\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}\right)^2 = \left(\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}\right)^2$$

及

$$\int_{\frac{\overline{X}-T_2}{\sigma/\sqrt{n}}}^{\frac{\overline{X}-T_1}{\sigma/\sqrt{n}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1-\alpha$$

故可推得

$$T_1 = \overline{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \text{ 及 } T_2 = \overline{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

所以母體參數 μ 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\overline{x}-z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{x}+z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

例 **6–23** 假設 $\overline{X}$ 為抽自 $\mathcal{N}(\mu,16)$ 樣本數為 5 的樣本平均數，試求 μ 之 90% 信賴區間。

解：

$$\left[\overline{x}-1.645\frac{4}{\sqrt{5}},\overline{x}+1.645\frac{4}{\sqrt{5}}\right]$$

6.2.2 母體變異數未知, 單一母體平均數之區間估計

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本。若母體變異數 σ^2 未知且 $n\leq 30$，則由定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-1)\leq \frac{\overline{X}-\mu}{S/\sqrt{n}}\leq t_{\frac{\alpha}{2}}(n-1)\right)\\
&= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\leq \overline{X}-\mu\leq t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)\\
&= \mathrm{P}\left(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\leq \mu\leq \overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)
\end{aligned}$$

所以, 母體參數 μ 在信賴水準為 $(1-\alpha)\times 100\%$ 下之信賴區間

$$\left[\overline{x}-t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}},\overline{x}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}(n-1)\right]$$

此外，當樣本為大樣本時，則由於 S^2 為母體變異數之優良估計式, 根據一致性原則可知，S^2 可直接作為母體變異數。因此，μ 在 $(1-\alpha)\%$ 下之信賴區間可改寫為

$$\left[\overline{x}-z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}},\overline{x}+z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$$

例 **6–24** 一個工程師要估計某種鋼條的平均強度, 假設該鋼條的強度為常態分配, 它做了四個試驗，得到強

度如下：

$$844,\ 847,\ 845,\ 844$$

計算該鋼條平均強度之 95% 信賴區間？

解：

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - n\overline{x}^2}{n-1} = \frac{2856106 - 2856100}{4-1} = 2$$

所以 μ 之 95% 信賴區間為

$$\left[845 - t_{0.025}(3)\sqrt{\frac{2}{4}}, 845 + t_{0.025}(3)\sqrt{\frac{2}{4}}\right] = [842.75, 847.25]$$

例 **6–25** 為了估計某名牌香煙的尼古丁量, 隨機檢驗該香煙 12 隻, 他們的尼古丁含量分別記錄如下：

$$18.2,\ 17.6,\ 19.0,\ 18.5,\ 18.1,\ 17.0$$

$$16.8,\ 17.5,\ 18.6,\ 17.9,\ 18.3,\ 16.5$$

單位：毫克。據以往經驗，發現該香煙一隻尼古丁含量一項都近似常態分配。

1. 求該香煙一隻的尼古丁平均含量的 90% 信賴區間
2. 假設該香煙一隻尼古丁含量不但近似常態分配，且標準差已知為 1.62 毫克, 是估計該香煙一支平均尼古丁含量, 並求該估計的 90% 信賴區間。

解：

$$\overline{x} = 17.8333,\ s^2 = 0.5933$$

1. 所以 μ 之 90% 信賴區間為

$$\left[17.8333 - t_{0.05}(11)\sqrt{\frac{0.5933}{12}}, 17.8333 + t_{0.05}(11)\sqrt{\frac{0.5933}{12}}\right] = [17.434, 18.232]$$

2. 因為標準差已知且資料服從常態分配, 所以 μ 之 90% 信賴區間為

$$\left[17.8333 - z_{0.05}\sqrt{\frac{0.5933}{12}}, 17.8333 + z_{0.05}\sqrt{\frac{0.5933}{12}}\right] = [17.468,\ 18.199]$$

6.2.3 母體變異數已知, 兩母體平均數差之區間估計

假設 $X_1, X_2, \ldots, X_{n_1}; Y_1, Y_2, \ldots, Y_{n_2}$ 為分別抽自 $\mathcal{N}(\mu_x, \sigma_x^2)$ 及 $\mathcal{N}(\mu_y, \sigma_y^2)$ 的兩組樣本數為 n_1 及 n_2 的隨機樣本，則 $\overline{X} - \overline{Y}$ 之抽樣分配可推得為

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}\right)$$

若母體變異數 σ_x^2 及 σ_y^2 已知，則由定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-z_{\frac{\alpha}{2}} \le \frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2/n_1 + \sigma_y^2/n_2}} \le z_{\frac{\alpha}{2}}\right) \\
&= \mathrm{P}\left(-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}} \le \overline{X} - \overline{Y} - (\mu_x - \mu_y) \le z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}\right) \\
&= \mathrm{P}\left(\overline{X} - \overline{Y} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n2}} \le \mu_x - \mu_y \le \overline{X} - \overline{Y} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n2}}\right)
\end{aligned}$$

所以，$\mu_x - \mu_y$ 之 $(1-\alpha) \times 100\%$ 信賴區間為

$$\left[\overline{x} - \overline{y} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n2}}, \overline{x} - \overline{y} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n2}}\right]$$

例 **6–26** X 品牌燈管之壽命服從 $\mathcal{N}(\mu_x, 784)$ 及 Y 品牌燈管之壽命服從 $\mathcal{N}(\mu_y, 627)$(單位：小時)，且兩品牌之燈管壽命長度為相互獨立。今由 X 品牌之燈管抽出 56 隻燈管得其樣本平均壽命為 937.4 小時，由 Y 品牌之燈管抽出 57 隻燈管得其樣本平均壽命為 988.9 小時，試求 $\mu_x - \mu_y$ 之 90% 之信賴區間。

解：

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_x - \mu_y, \frac{784}{56} + \frac{727}{57}\right)$$

所以，$\mu_x - \mu_y$ 之 90% 信賴區間為

$$\left[-51.5 - z_{0.05}\sqrt{\frac{784}{56} + \frac{627}{57}}, -51.5 + z_{0.05}\sqrt{\frac{784}{56} + \frac{627}{57}}\right] = [-59.725, -43.275]$$

6.2.4 母體變異數未知, 兩母體平均數差之區間估計

假設 $X_1, X_2, \ldots, X_{n_1}; Y_1, Y_2, \ldots, Y_{n_2}$ 為分別抽自 $\mathcal{N}(\mu_x, \sigma_x^2)$ 及 $\mathcal{N}(\mu_y, \sigma_y^2)$ 的兩組樣本數為 n_1 及 n_2 的隨機樣本, 其中 σ_x 及 σ_y 未知且 $n_1 < 30$ 及 $n_2 < 30$，則由定義可知

$$\frac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\dfrac{\sigma_x^2}{n_1}+\dfrac{\sigma_y^2}{n_2}}} \sim \mathcal{N}(0,1), \quad \frac{(n_1-1)\,S_x^2}{\sigma_x^2}+\frac{(n_2-1)\,S_y^2}{\sigma_2^2} \sim \chi^2(n_1+n_2-2)$$

因此，由 t 分配之定義可推得

$$\frac{\dfrac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\dfrac{\sigma_x^2}{n_1}+\dfrac{\sigma_y^2}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1-1)\,S_x^2}{\sigma_x^2}+\dfrac{(n_2-1)\,S_y^2}{\sigma_y^2}}{n_1+n_2-2}}} \sim t(n_1+n_2-2)$$

由於 σ_x^2 及 σ_y^2 未知，故分以下兩狀況討論

1. 兩變異數未知但相等

假設 $\sigma_x^2 = \sigma_y^2 = \sigma^2$, 上式可寫成

$$\begin{aligned}
\frac{\dfrac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\dfrac{\sigma_x^2}{n_1}+\dfrac{\sigma_y^2}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1-1)\,S_x^2}{\sigma_x^2}+\dfrac{(n_2-1)\,S_y^2}{\sigma_y^2}}{n_1+n_2-2}}}
&= \frac{\dfrac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}}}{\sqrt{\dfrac{\dfrac{(n_1-1)\,S_x^2}{\sigma^2}+\dfrac{(n_2-1)\,S_y^2}{\sigma^2}}{n_1+n_2-2}}} \\
&= \frac{\dfrac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sigma^2\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n2}}}}{\dfrac{1}{\sigma^2}\sqrt{\dfrac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2}}} \\
&= \frac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\dfrac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}
\end{aligned}$$

推得

$$\frac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{\sqrt{\frac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t\,(n_1+n_2-2)$$

因此，由定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n)\le \frac{\overline{X}-\overline{Y}-(\mu_x-\mu_y)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\le t_{\frac{\alpha}{2}}(n)\right)\\
&= \mathrm{P}\left(-S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}(n)\le \overline{X}-\overline{Y}-(\mu_x-\mu_y)\le S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}(n)\right)\\
&= \mathrm{P}\left(\overline{X}-\overline{Y}-S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}(n)\le \mu_x-\mu_y\le \overline{X}-\overline{Y}+S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}(n)\right),
\end{aligned}$$

其中

$$S_p^2=\frac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2},\ n=n_1+n_2-2$$

在此我們稱 S_p^2 為共同變異數 σ^2 之混和估計量（pooled estimator），且

$$\begin{aligned}
\mathrm{E}\left(S_p^2\right) &= \mathrm{E}\left(\frac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2}\right)\\
&= \frac{(n_1-1)\,\mathrm{E}\left(S_x^2\right)}{n_1+n_2-2}+\frac{(n_2-1)\,\mathrm{E}\left(S_y^2\right)}{n_1+n_2-2}\\
&= \frac{(n_1-1)\,\sigma^2}{n_1+n_2-2}+\frac{(n_2-1)\,\sigma^2}{n_1+n_2-2}\\
&= \sigma^2
\end{aligned}$$

所以，$\mu_x-\mu_y$ 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\overline{x}-\overline{y}-s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\,,\overline{x}-\overline{y}+s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\right]$$

2. 兩變異數未知並不相等

$\mu_x - \mu_y$ 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\overline{x}-\overline{y}-\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}t_{\frac{\alpha}{2}}\left(v\right),\overline{x}-\overline{y}+\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}t_{\frac{\alpha}{2}}\left(v\right)\right]$$

其中

$$v=\frac{\left(\dfrac{s_x^2}{n_1}+\dfrac{s_y^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_x^2}{n_1}\right)^2}{n_1-1}+\dfrac{\left(\dfrac{s_y^2}{n_2}\right)^2}{n_2-1}}$$

若 v 不為整數，則無條件捨去。

此外，當樣本皆為大樣本時，則由於 S_x^2 和 S_y^2 皆為母體變異數之優良估計式, 根據一致性原則可知，S_x^2 及 S_y^2 可直接作為母體變異數。因此，不論母體平均數是否相等, $\mu_x-\mu_y$ 在 $(1-\alpha)\%$ 下之信賴區間皆可改寫為

$$\left[\overline{x}-\overline{y}-z_{\frac{\alpha}{2}}\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}},\overline{x}-\overline{y}+z_{\frac{\alpha}{2}}\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}\right]$$

例 **6–27** 假設 A,B 兩班之統計期中考成績分別服從 $\mathcal{N}\left(\mu_x,\sigma^2\right)$ 及 $\mathcal{N}\left(\mu_y,\sigma^2\right)$，其中母體變異數 σ^2 未知。今由各班抽出一組隨機樣本, 得其樣本觀測值分別為 $\overline{x}$=81.31, s_x^2=60.76, n_1=9, $\overline{y}$=78.61, s_y^2=48.24, n_2 =15, 求 $\mu_x-\mu_y$ 之 95% 信賴區間為何？

解：

因為 $\sigma_x^2=\sigma_y^2=\sigma^2$，所以

$$s_p^2=\frac{(9-1)\times 60.76+(15-1)\times 48.24}{9+15-2}$$

故可求得 $\mu_x-\mu_y$ 之 95% 信賴區間為

$$\left[81.31-78.61-s_p\sqrt{\frac{1}{9}+\frac{1}{15}}t_{0.025}\left(22\right),81.31-78.61+s_p\sqrt{\frac{1}{9}+\frac{1}{15}}t_{0.025}\left(22\right)\right]=[-3.65,9.05]$$

例 **6–28** 將國民中學分為都市國民中學與鄉村國民中學二類, 各隨機抽取三年級學生 10 人作數學測驗，其

成績如下：

$$X(\text{都市國中})\ 73,\ 59,\ 69,\ 66,\ 55,\ 87,\ 89,\ 89,\ 84,\ 75$$

$$Y(\text{鄉村國中})\ 86,\ 60,\ 71,\ 85,\ 42,\ 34,\ 53,\ 60,\ 68,\ 92$$

試求都市國中與鄉村國中三年級學生數學平均成績之差的 95% 信賴區間, 假定其成績之分配為常態分配，而且二變異數相等。

解：

$$\overline{x}=74,\ \overline{y}=65.1,\ \sum_{i=1}^{10}(x-\overline{x})^2=1392.4,\ \sum_{i=1}^{10}(y-\overline{y})^2=3298.9$$

$$s_p^2=\frac{(n_1-1)s_x^2+(n_2-1)s_y^2}{n_1+n_2-2}=260.63$$

又 $t_{0.025}(18)=2.101$，所以 $\overline{X}-\overline{Y}$ 之 95% 信賴區間為

$$\left[(74.6-65.1)-2.101\sqrt{260.63}\sqrt{\frac{1}{10}+\frac{1}{10}}\right]=[-5.67,24.67]$$

6.2.5 兩相依母體平均數之區間估計

假設 $(X_1,Y_1),(X_2,Y_2),\ \ldots,\ (X_n,Y_n)$ 一樣本數為 n 之**相依樣本** (例如父子之身高, 減肥前減肥後之體重), 令 $d_i=X_i-Y_i,\ i=1,\ 2\ ,\ldots,,\ n.$ 假設 d_i 表兩隨機變數之差, μ_d 及 σ_d^2 分別表其平均數及變異數，則由 t 分配之定義可知

$$\frac{\overline{d}-\mu_d}{S_d/\sqrt{n}}\sim t(n-1)$$

因此，由信賴區間之定義可知

$$\begin{aligned}1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-1)\le\frac{\overline{d}-\mu_d}{S_d/\sqrt{n}}\le t_{\frac{\alpha}{2}}(n-1)\right)\\ &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-1)\frac{S_d}{\sqrt{n}}\le\overline{d}-\mu_d\le t_{\frac{\alpha}{2}}(n-1)\frac{S_d}{\sqrt{n}}\right)\\ &= \mathrm{P}\left(\overline{d}-t_{\frac{\alpha}{2}}(n-1)\frac{S_d}{\sqrt{n}}\le\mu_d\le\overline{d}+t_{\frac{\alpha}{2}}(n-1)\frac{S_d}{\sqrt{n}}\right)\end{aligned}$$

所以，μ_d 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[-t_{\frac{\alpha}{2}}\left(n-1\right)\frac{s_d}{\sqrt{n}}, t_{\frac{\alpha}{2}}\left(n-1\right)\frac{s_d}{\sqrt{n}}\right]$$

當樣本數 $n>30$ 時，則可寫成

$$\left[-z_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}, z_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}\right]$$

例 **6–29** 有一個修車廠的招牌生意是為各型汽車裝設省油裝置, 為了要了解各種省油裝置的裝設是否確實有省油功效, 隨機抽出了 8 輛有裝設這種省油裝置的汽車, 並記錄其裝設這種省油裝置前後行駛 100 公里的耗油量如下：

車輛編號	1	2	3	4	5	6	7	8
裝設前	3.2	4.6	3.6	5.3	6.2	3.2	3.6	4.5
裝設後	2.9	4.7	3.2	5.0	5.7	3.3	3.4	4.3

假設每輛車耗油量皆呈常態分配，求一輛汽車裝設這種省油裝置之前與之後 100 公里之平均耗油量的 90% 信賴區間。

解：

$$\overline{d}=0.2125,\ s_d^2=0.0470\ , n=8 \text{ 及 } t_{0.05}(7)=1.895$$

所以，μ_d 之 90% 信賴區間為

$$\left[0.2125-1.895\times\frac{0.2167}{\sqrt{8}}, 0.2125+1.895\times\frac{0.2167}{\sqrt{8}}\right]=[0.067, 0.3577]$$

6.2.6 母體變異數之區間估計

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu,\sigma^2)$ 的一組樣本數為 n 的隨機樣本。若母體變異數 σ^2 未知，由於 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，所以由信賴區間之定義可知

$$1-\alpha \;=\; \mathrm{P}\left(T_1\le\sigma^2\le T_2\right)$$

$$= \quad \mathrm{P}\left(\frac{(n-1)\,S^2}{T_2} \leq \frac{(n-1)\,S^2}{\sigma^2} \leq \frac{(n-1)\,S^2}{T_1}\right)$$

今要使得 σ 之信賴區間為最小即求解下列問題

$$\min \sqrt{T_2} - \sqrt{T_1}$$

$$\text{subject} \quad \int_{\frac{(n-1)S^2}{T_2}}^{\frac{(n-1)S^2}{T_1}} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx = 1-\alpha$$

由 Lagrange 方法可將上述問題寫成

$$\min \quad L = \sqrt{T_2} - \sqrt{T_1} - \lambda \left\{ \int_{\frac{(n-1)S^2}{T_2}}^{\frac{(n-1)S^2}{T_1}} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx - (1-\alpha) \right\}$$

分別對 T_1, T_2 及 λ 做偏微分並令其等於零可得

$$\frac{\partial L}{\partial T_2} = \frac{1}{2\sqrt{T_2}} - \lambda \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} \left(\frac{(n-1)\,S^2}{T_2}\right)^{\frac{n-1}{2}-1} e^{-\frac{\frac{(n-1)S^2}{T_2}}{2}} \frac{(n-1)\,S^2}{T_2^2} = 0$$

$$\frac{\partial L}{\partial T_1} = -\frac{1}{2\sqrt{T_1}} + \lambda \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} \left(\frac{(n-1)\,S^2}{T_1}\right)^{\frac{n-1}{2}-1} e^{-\frac{\frac{(n-1)S^2}{T_1}}{2}} \frac{(n-1)\,S^2}{T_1^2} = 0$$

$$\frac{\partial L}{\partial \lambda} = \int_{\frac{(n-1)S^2}{T_2}}^{\frac{(n-1)S^2}{T_1}} \frac{1}{\Gamma\left(\frac{n-1}{2}\right) 2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{-\frac{x}{2}} dx - (1-\alpha) = 0$$

求解前兩式可得最佳解之必要條件為

$$\frac{\left(\dfrac{(n-1)\,S^2}{T_2}\right)^{\dfrac{n-1}{2}-1} e^{-\dfrac{\dfrac{(n-1)\,S^2}{T_2}}{2}} \dfrac{(n-1)\,S^2}{T_2^2}}{\left(\dfrac{(n-1)\,S^2}{T_1}\right)^{\dfrac{n-1}{2}-1} e^{-\dfrac{\dfrac{(n-1)\,S^2}{T_1}}{2}} \dfrac{(n-1)\,S^2}{T_1^2}} = \frac{\sqrt{T_2}}{\sqrt{T_1}}$$

經化簡後得

$$e^{-\frac{b}{2}}b^{\frac{n}{2}} = e^{-\frac{a}{2}}a^{\frac{n}{2}},$$

其中

$$a = \frac{(n-1)\,S^2}{T_2}, \quad b = \frac{(n-1)\,S^2}{T_1}$$

然而, 以上述方法要求得 σ^2 之最佳信賴區間在技術上實太為繁瑣。因此，為求計算簡便, 我們通常以下列方法來求解母體參數，σ^2, 之信賴區間

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(\chi^2_{1-\frac{\alpha}{2}}(n-1) \le \frac{(n-1)\,S^2}{\sigma^2} \le \chi^2_{\frac{\alpha}{2}}(n-1)\right) \\
&= \mathrm{P}\left(\frac{1}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} \ge \frac{\sigma^2}{(n-1)\,S^2} \ge \frac{1}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) \\
&= \mathrm{P}\left(\frac{(n-1)\,S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \le \sigma^2 \le \frac{(n-1)\,S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)
\end{aligned}$$

所以，σ^2 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\frac{(n-1)\,s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}, \frac{(n-1)\,s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]$$

例 **6–30** 設由一常態母體中隨機抽出一大小為 20 的樣本，並計算得其樣本變異數 $s^2=40$，試以 σ^2 的 95% 信賴區間, 估計母變異數所在範圍。

解：

已知 $n=20$, $s^2=40$，又查表得 $\chi^2_{0.025}(19)=32.8523$ 及 $\chi^2_{0.975}(19)=8.9065$,，所以 σ^2 之 95% 信賴區間為

$$\left[\frac{19\times 40}{32.8523}, \frac{19\times 40}{8.9065}\right] = [23.1338, 85.3309]$$

6.2.7 兩母體變異數比例之區間估計

假設 $X_1, X_2, \ldots, X_{n_1}$; $Y_1, Y_2, \ldots, Y_{n_2}$ 為分別抽自 $\mathcal{N}(\mu_x, \sigma_x^2)$ 及 $\mathcal{N}(\mu_y, \sigma_y^2)$ 的兩組樣本數為 n_1 及 n_2 的隨機樣本, 其中 σ_x^2 及 σ_y^2 未知且 $n_1 < 30$ 及 $n_2 < 30$。由於

$$\frac{(n_1-1)S_x^2}{\sigma_x^2} \sim \chi^2(n_1-1),\ \frac{(n_2-1)S_y^2}{\sigma_x^2} \sim \chi^2(n_2-1)$$

由 F 分配及信賴區間之定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) < \frac{\dfrac{\frac{(n_1-1)S_x^2}{\sigma_x^2}}{n_1-1}}{\dfrac{\frac{(n_2-1)S_y^2}{\sigma_y^2}}{n_2-1}} < F_{\frac{\alpha}{2}}(n_1-1, n_2-1)\right) \\
&= \mathrm{P}\left(F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) > \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} > F_{\frac{\alpha}{2}}(n_1-1, n_2-1)\right) \\
&= \mathrm{P}\left(\frac{1}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)} < \frac{S_y^2}{S_x^2}\frac{\sigma_x^2}{\sigma_y^2} < \frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}\right) \\
&= \mathrm{P}\left(\frac{S_x^2/S_y^2}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)} < \frac{\sigma_x^2}{\sigma_y^2} < \frac{S_x^2/S_y^2}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}\right)
\end{aligned}$$

所以，兩母體變異數比例之 $(1-\alpha) \times 100\%$ 之信賴區間為

$$\left[\frac{s_x^2/s_y^2}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)}, \frac{s_x^2/s_y^2}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)}\right]$$

必須注意的是, 此一信賴區間所得到的結果方式與前一小節相同, 母體變異數比例之信賴區間長度並非最小, 僅是一計算上方便求解的一種方式。

例 **6–31** 設某系 A, B 班的統計學成績分別成常態分配 $\mathcal{N}\left(\mu_1, \sigma_1^2\right)$ 及 $\mathcal{N}\left(\mu_2, \sigma_2^2\right)$，其中 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ 皆為未知數。若隨機由 A 班抽出 16 位學生得其樣本平均數為 75.2，樣本變異數為 8.64; 由 B 班隨機抽出 10 位同學，其樣本平均數為 78.6，樣本變異數為 7.88(假設兩母體獨立)，試求 σ_1^2/σ_2^2 之 90% 信賴區間？

解：

$s_1^2/s_2^2 = 1.0964$, $F_{0.05}(15,9) = 3.01$ 及 $F_{0.05}(9,15) = 2.59$, 可求得

$$F_{0.95}(15,9) = \frac{1}{F_{0.05}(9,15)} = \frac{1}{3.01} = 0.3861$$

所以，σ_1^2/σ_2^2 之 90% 信賴區間

$$\left[\frac{1.0964}{3.01}, \frac{1.0964}{0.3861}\right] = [0.4233, 3.3004]$$

6.2.8 單一母體比例之區間估計

假設 X_1, X_2, ..., X_n 為抽自 Bernoulli(p) 的一組樣本數為 n 的隨機樣本。令 $\widehat{p} = \sum_{i=1}^{n} X_i/n$, 則當樣本數夠大時由中央極限定理可知

$$\frac{\widehat{p} - p}{\sqrt{\dfrac{\mathrm{P}(1-p)}{n}}} \sim \mathcal{N}(0,1)$$

所以由信賴區間之定義可知

$$\begin{aligned}
1-\alpha &\approx \mathrm{P}\left(-z_{\frac{\alpha}{2}} \le \frac{\widehat{p}-p}{\sqrt{\mathrm{P}(1-p)/n}} \le z_{\frac{\alpha}{2}}\right) \\
&= \mathrm{P}\left(-z_{\frac{\alpha}{2}}\sqrt{\mathrm{P}(1-p)/n} \le \widehat{p}-p \le z_{\frac{\alpha}{2}}\sqrt{\mathrm{P}(1-p)/n}\right) \\
&= \mathrm{P}\left(\widehat{p}-z_{\frac{\alpha}{2}}\sqrt{\mathrm{P}(1-p)/n} \le p \le \widehat{p}+z_{\frac{\alpha}{2}}\sqrt{\mathrm{P}(1-p)/n}\right)
\end{aligned}$$

由於母體參數 p 為一未知參數，因此無法由此法求得信賴區間。又由上式可知

$$\left|\frac{\widehat{p}-p}{\sqrt{\mathrm{P}(1-p)/n}}\right| \le z_{\frac{\alpha}{2}} \text{ 即 } \frac{(\widehat{p}-p)^2}{\mathrm{P}(1-p)/n} \le z_{\frac{\alpha}{2}}^2$$

故可推得

$$\left(1+\frac{z_{\frac{\alpha}{2}}^2}{n}\right)p^2 - \left(2\widehat{p}+\frac{z_{\frac{\alpha}{2}}^2}{n}\right)p + \widehat{p}^2 \le 0$$

經化簡後可改寫為

$$
\begin{aligned}
& \left[p-\frac{\left(2\widehat{p}+z_{\frac{\alpha}{2}}^2/n\right)+2z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{2\left(1+z_{\frac{\alpha}{2}}^2/n\right)}\right] \\
\times & \left[p+\frac{\left(2\widehat{p}+z_{\frac{\alpha}{2}}^2/n\right)+2z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{2\left(1+z_{\frac{\alpha}{2}}^2/n\right)}\right] \\
\leq & \ 0
\end{aligned}
$$

所以,

$$
\mathrm{P}\left(\frac{\widehat{p}+z_{\frac{\alpha}{2}}^2/2n-z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{1+z_{\frac{\alpha}{2}}^2/n}\leq p\leq\frac{\widehat{p}+z_{\frac{\alpha}{2}}^2/2n+z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{1+z_{\frac{\alpha}{2}}^2/n}\right)\approx 1-\alpha
$$

因此，p 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$
\left[\frac{\widehat{p}+z_{\frac{\alpha}{2}}^2/2n-z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{1+z_{\frac{\alpha}{2}}^2/n},\frac{\widehat{p}+z_{\frac{\alpha}{2}}^2/2n+z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n+z_{\frac{\alpha}{2}}^2/4n^2}}{2\left(1+z_{\frac{\alpha}{2}}^2/n\right)}\right]
$$

當 n 大時，由於 $z_{\frac{\alpha}{2}}^2/n\approx 0$，所以一般將 p 之 $(1-\alpha)\times 100\%$ 信賴區間寫成

$$
\left[\widehat{p}-z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}},\widehat{p}+z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}}\right]
$$

在此，讀者亦可利用 $\widehat{p}$ 為 p 之 MLE 之性質求得相同之信賴區間。由 **7.1.1**（第 170 頁）可知，$\widehat{p}$ 之漸近變異數為

$$
\frac{\left[\frac{d}{dp}p\right]^2\Big|_{p=\widehat{p}}}{-\frac{\partial^2}{\partial p^2}\ln L\left(\mathbf{X};p\right)\Big|_{p=\widehat{p}}}
$$

又

$$
\begin{aligned}
-\frac{\partial^2}{\partial p^2}\ln L\left(\mathbf{X};p\right) &= -\frac{\partial^2}{\partial p^2}\left[\sum_{i=1}^{n}x_i\ln p+\left(n-\sum_{i=1}^{n}x_i\right)\ln\left(1-p\right)\right] \\
&= \frac{n\widehat{p}}{\widehat{p}^2}+\frac{n\left(1-\widehat{p}\right)}{\left(1-\widehat{p}\right)^2}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{n}{\widehat{p}} + \frac{n}{1-\widehat{p}} \\
&= \frac{n}{\widehat{p}\left(1-\widehat{p}\right)}
\end{aligned}
$$

所以

$$
V\left(\widehat{p}\right) \approx \frac{\left[\frac{d}{dp}p\right]^2 |_{p=\widehat{p}}}{-\frac{\partial^2}{\partial p^2}\ln L\left(\mathbf{X};p\right)\Big|_{p=\widehat{p}}} = \frac{\widehat{p}\left(1-\widehat{p}\right)}{n}
$$

由於

$$
\lim_{n\to\infty}\frac{\widehat{p}-p}{\sqrt{\mathrm{P}\left(1-p\right)/n}} \xrightarrow{d} \mathcal{N}\left(0,1\right) \text{ 及 } \lim_{n\to\infty}\mathrm{P}\left(\left|\frac{\widehat{p}\left(1-\widehat{p}\right)}{n} - \frac{\mathrm{P}\left(1-p\right)}{n}\right| < \varepsilon\right) = 1
$$

由 Slutsky's Theorem 可知

$$
\lim_{n\to\infty}\frac{\sqrt{\mathrm{P}\left(1-p\right)/n}}{\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n}}\frac{\widehat{p}-p}{\sqrt{\mathrm{P}\left(1-p\right)/n}} = \lim_{n\to\infty}\frac{\widehat{p}-\widehat{p}_1}{\sqrt{\widehat{p}\left(1-\widehat{p}\right)/n}} \xrightarrow{d} \mathcal{N}\left(0,1\right)
$$

所以,

$$
\begin{aligned}
1-\alpha &\approx \mathrm{P}\left(-z_{\frac{\alpha}{2}} < \frac{\widehat{p}-p}{\sqrt{\mathrm{Var}\left(\widehat{p}\right)}} < z_{\frac{\alpha}{2}}\right) \\
&= \mathrm{P}\left(\widehat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}} < p < \widehat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}}\right)
\end{aligned}
$$

故 p 之 $(1-\alpha)\times 100\%$ 信賴區間寫成

$$
\left[\widehat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}}, \widehat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}}\right]
$$

以上母體比例 p 之區間估計討論之內容為在無限母體之假設下推導而得，若母體為有限母體時，則必須做有限母體校正，即

$$
\left[\widehat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}\frac{N-n}{N-1}}, \widehat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\left(1-\widehat{p}\right)}{n}\frac{N-n}{N-1}}\right],
$$

其中 N 為母體元素個數，n 為樣本數。

例 **6–32** 交通安全常識中一再強調機車騎士戴安全帽之重要性。由於南部氣候炎熱，所以機車騎士配戴安全帽之比例 p 成為我們關心的問題。今在路口隨機觀察 400 位機車騎士行車，其中 50 位有配戴安全帽。求 p 的 95% 信賴區間？

解：

$\widehat{p}=\frac{1}{8}=0.125$ 及 $z_{0.025}=1.96$，所以 p 之 95% 信賴區間為

$$\left[0.125-1.96\sqrt{\frac{0.125\times 0.875}{400}},0.125+1.96\sqrt{\frac{0.125\times 0.875}{400}}\right]=[0.0926,0.1574]$$

6.2.9 兩母體比例差之區間估計

假設 $X_1, X_2, ..., X_n; Y_1, Y_2, ..., Y_n$ 為分別抽自 $\text{Bernoulli}\,(p_1)$ 與 $\text{Bernoulli}\,(p_2)$ 的兩組樣本數為 n_1 與 n_2 的獨立隨機樣本。令 $\widehat{p}_1=\sum_{i=1}^{n}X_i/n_1$，與 $\widehat{p}_2=\sum_{i=1}^{n}Y_i/n_2$，所以 $\mathrm{E}\left(\widehat{p}_1-\widehat{p}_2\right)=p_1-p_2$, $\mathrm{Var}\left(p_1-p_2\right)=p_1\left(1-p_1\right)/n_1+p_2\left(1-p_2\right)/n_2$。當樣本數大時 $(n_1>30,\, n_2>30)$，由於 $\widehat{p}_1$ 及 $\widehat{p}_2$ 的漸近變異數分別為

$$\mathrm{Var}\left(\widehat{p}_1\right)\approx\frac{\widehat{p}_1\left(1-\widehat{p}_1\right)}{n_1}\text{ 及}\mathrm{Var}\left(\widehat{p}_2\right)\approx\frac{\widehat{p}_2\left(1-\widehat{p}_2\right)}{n_2}$$

由於

$$\lim_{n\to\infty}\frac{\widehat{p}_1-\widehat{p}_2-\left(p_1-p_2\right)}{\sqrt{p_1\left(1-p_1\right)/n_1+p_2\left(1-p_2\right)/n_2}}\xrightarrow{d}\mathcal{N}\left(0,1\right)$$

及

$$\lim_{n\to\infty}\mathrm{P}\left(\left|\frac{\widehat{p}_1\left(1-\widehat{p}_1\right)}{n_1}+\frac{\widehat{p}_2\left(1-\widehat{p}_2\right)}{n_2}-\frac{p_1\left(1-p_1\right)}{n_1}-\frac{p_2\left(1-p_2\right)}{n_2}\right|<\varepsilon\right)=1$$

由 Slutsky's Theorem 可知 $\widehat{p}_1-\widehat{p}_2$ 的抽樣分配

$$\lim_{n\to\infty}\frac{\widehat{p}_1-\widehat{p}_2-\left(p_1-p_2\right)}{\sqrt{\widehat{p}_1\left(1-\widehat{p}_1\right)/n_1+\widehat{p}_2\left(1-\widehat{p}_2\right)/n_2}}\xrightarrow{d}\mathcal{N}\left(0,1\right)$$

所以，利用與 **7.2.8** 相同的方式, 可得到 p_1-p_2 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\widehat{p}_1-\widehat{p}_2-z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_1\left(1-\widehat{p}_1\right)}{n_1}+\frac{\widehat{p}_2\left(1-\widehat{p}_2\right)}{n_2}},\widehat{p}_1-\widehat{p}_2+z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_1\left(1-\widehat{p}_1\right)}{n_1}+\frac{\widehat{p}_2\left(1-\widehat{p}_2\right)}{n_2}}\right]$$

例 **6–33** 在某大城市中隨機抽 120 位的成年男子，其中有 38 位沒有抽菸的習慣，又隨機抽出 72 位的成年女子，其中有 51 位沒有抽菸的習慣。試求估計此城市成年男子中與成年女子中有抽菸習慣的比利差之 95% 的信賴區間？

解：

$\widehat{p}_1 = \frac{82}{120} = 0.6833, \widehat{p}_1 = \frac{21}{72} = 0.2917$ 及 $z_{0.025} = 1.96$，又

$$\sqrt{\frac{0.6833\,(1-0.6833)}{120} + \frac{0.2917\,(1-0.2917)}{72}} = 0.0683$$

所以 $p_1 - p_2$ 之 95% 信賴區間為

$$[0.6833 - 0.2917 - 1.96 \times 0.0683, 0.6833 - 0.2917 + 1.96 \times 0.0683] = [0.2577, 0.5254]$$

6.3 樣本數大小之決定

6.3.1 母體平均數樣本數之決定

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma^2)$ 的一組樣本數為 n 的隨機樣本。若 $n \geq 30$，則不論母體變異數 σ^2 是否已知，由 **7.2.1** 及 **7.2.2** 可知母體平均數 μ 之 $(1-\alpha) \times 100\%$ 之信賴區間為

$$\left[\overline{x} - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}, \overline{x} + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right],$$

在此我們稱 $z_{\frac{\alpha}{2}} s/\sqrt{n}$ 為誤差界限。若要求誤差界限小於 ϵ，即

$$z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \epsilon,$$

經簡單的轉換之後，則可推導出在誤差界限為 ϵ 下，所需之最小樣本數為

$$n = \frac{z_{\frac{\alpha}{2}}^2 s^2}{\epsilon^2}$$

若母體變異數 σ^2 已知，則可直接寫成

$$n=\frac{z_{\frac{\alpha}{2}}^2\sigma^2}{\epsilon^2}$$

6.3.2 母體比例樣本數之決定

假設 $X_1, X_2, ..., X_n$ 為抽自 Bernoulli(p) 的一組樣本數為 n 的隨機樣本。令 $\widehat{p}=\sum_{i=1}^n X_i/n$，則由 **7.2.8** 可知，$p$ 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\widehat{p}-z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}},\widehat{p}+z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}\right]$$

所以，p 之誤差界限為 $z_{\frac{\alpha}{2}}\sqrt{\widehat{p}\,(1-\widehat{p})/n}$。若要求誤差界限小於 ϵ，即

$$z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}<\epsilon,$$

經簡單的轉換之後，則可推導出在誤差界限為 ϵ 下，所需之最小樣本數為

$$n=\frac{z_{\frac{\alpha}{2}}^2\widehat{p}\,(1-\widehat{p})}{\epsilon^2}$$

又 $\mathrm{P}(1-p)\leq 0.25$，所以

$$n=\frac{z_{\frac{\alpha}{2}}^2\widehat{p}\,(1-\widehat{p})}{\epsilon^2}\leq\frac{z_{\frac{\alpha}{2}}^2}{4\epsilon^2}$$

在一般應用上，由於在實驗之前並尚未決定樣本數 n 之大小，因此為求謹慎, 通常以在最大誤差界限等於 ϵ 為條件, 抽取最大的樣本數 n^*，其中

$$n^*=\frac{z_{\frac{\alpha}{2}}^2}{4\epsilon^2}$$

當 $\epsilon=0.03$ 及信賴水準 $\alpha=0.05$ 時，

$$n^*=\frac{1.96^2}{4\times 0.03^2}=1067.1$$

此外，當母體為有限母體時，則 p 之誤差界限為

$$\epsilon = z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}\frac{N-n}{N-1}}$$

經化簡, 上式可改寫為

$$n = \frac{z^2_{\frac{\alpha}{2}}\widehat{p}\,(1-\widehat{p})\,N}{\epsilon^2\,(N-1)+z^2_{\frac{\alpha}{2}}\widehat{p}\,(1-\widehat{p})} = \frac{z^2_{\frac{\alpha}{2}}N}{\dfrac{\epsilon^2\,(N-1)}{\widehat{p}\,(1-\widehat{p})}+z^2_{\frac{\alpha}{2}}}$$

又

$$\frac{\epsilon^2\,(N-1)}{\widehat{p}\,(1-\widehat{p})} \geq \frac{\epsilon^2\,(N-1)}{4}$$

所以在誤差界限為 ϵ 下，所需之最大樣本數為

$$n = \frac{z^2_{\frac{\alpha}{2}}N}{\dfrac{\epsilon^2\,(N-1)}{4}+z^2_{\frac{\alpha}{2}}}$$

例 **6–34** 海龍牌洗衣機想知道其產品在全國市場上的佔有率而舉辦抽樣抽查，希望樣本比例與全國占有率之誤差不大於 0.05，有 98% 的信賴度，問樣本應多大？

1. $\widehat{p} = 0.228$
2. $\widehat{p}$ 未知且無任何有關事前資料

解：

1. 已知 $\epsilon = 0.05$ 且經查表得 $z_{0.01} = 2.326$，所以

$$n = \frac{2.326^2 \times 0.772 \times 0.228}{0.05^2} = 380.9,\ \text{故取}n = 381$$

2.

$$n = \frac{z^2_{0.01}\widehat{p}\,(1-\widehat{p})}{\epsilon^2} \leq \frac{2.326^2}{4 \times 0.05^2} = 541.02\ \text{故取}n = 542$$

第 7 章 假設檢定

7.1 觀念介紹

7.1.1 名詞解釋

1. **假設檢定**：所謂假設檢定乃指先給予母體未知參數一個假設值，在利用實驗之結果或母體抽出的樣本，應用機率理論去判斷此一假設值對或是不對, 這種判斷的過程或這種方式的統計推論即稱為假設檢定 (Testing Hypotheses)。

2. **統計假設**：任何有關於描述母體的敘述皆稱為統計假設。

3. **虛無假設**：Null Hypotheses, 在假設檢定的術語中, 希望被否定的統計假設稱為虛無假設。

4. **對立假設**：Alternative Hypotheses, 在假設檢定的術語中, 否定虛無假設而被認為是對的統計假設稱為對立假設。

5. **簡單假設**：其意義乃是只在統計假設中只為一個數值的假設, 也就是說一個能完全地確定出機率分配的假設稱之。例如 $H_0\ \theta = \theta_0$ vs $H_1\ \theta = \theta_1$。

6. **複合假設**：在統計假設中，其假設不只包含一個數值，而是一個數值的集合。例如 $H_0 \theta = \theta_0$ vs $H_1 \theta \neq \theta_0$ 或 $H_0 \theta \geq \theta_0$ vs $H_1 \theta < \theta_0$。

7. **危險域**：危險域 C 為樣本空間一些點的集合, 導致否定虛無假設 H_0 之集合。

7.1.2 誤差之型態

利用樣本資料去估計母體參數時，由於抽樣可能產生偏差而可能產生錯誤決策的風險。

1. **型 I 誤差**：Type I Error, 當虛無假設 H_0 為真, 但檢定結果為拒絕 H_0 稱之。
2. **型 II 誤差**：Type II Error, 當虛無假設 H_0 不眞, 但檢定結果為不拒絕 H_0 稱之。
3. α **風險**：型 I 誤差發生之機率稱為 α 風險; 即 $\alpha = \mathrm{P}(\text{型 I 誤差}) = \mathrm{P}(\text{拒絕}H_0 \mid H_0\text{為眞})$。
4. β **風險**：型 II 誤差發生之機率稱為 β 風險; 即 $\beta = \mathrm{P}(\text{型 II 誤差}) = \mathrm{P}(\text{不拒絕}H_0 \mid H_0\text{為偽})$。
5. **顯著水準**：Level of significance，即型 I 誤差發生最大的機率值，以 α 表示。

		眞實狀況	
		H_0 為眞	H_0 為偽
決	接受 H_0	正確決策	型 II 誤差
策	棄卻 H_0	型 I 誤差	正確決策

6. **生產者風險**：即 α 風險, 當生產者產品不良率 $P \leq P_0$ 合乎要求時，因統計檢定後的結果而被判斷為不合格之機率。
7. **消費者風險**：即 β 風險, 當生產者的產品不良率不合格時，卻因統計檢定後的結果被判斷為合格的機率。
8. **檢定力函數**：**Power Function of Test**, 檢定虛無假設與對立假設的檢定力函數，乃指樣本點落入危險域 C 中的機率。亦即檢定力函數為否定虛無假設的機率, 而檢定力函數的圖形, 我們稱之為檢定力曲線，其函數型態表示為

$$K(\theta_0) = \mathrm{P}\left(\widehat{\theta} \in C \mid \theta = \theta_0\right)$$

9. **作業特性函數**：**Operating Characteristic**, 檢定虛無假設與對立假設的檢定力函數，乃指樣本點落入接受域 C' 中的機率。亦即檢定力函數為不否定虛無假設的機率, 而檢定力函數的圖形, 我們稱之為檢定力曲線。其函數型態表示為

$$OC(\theta_0) = \mathrm{P}\left(\widehat{\theta} \notin C \mid \theta = \theta_0\right)$$

10. **P-value**：以樣本觀測值為基準, 做為棄卻 H_0 為錯誤決策的機率。

綜合以上討論，可得以下結論

1. 型 I 誤差與型 II 誤差互有關聯, 其中之一的機率變大，則另一機率就變小。

2. 若危險域的範圍愈小，則型 I 誤差 α 的機率就愈小, 型 II 誤差 β 的機率就愈大; 相反地，若危險域的範圍愈大, 型 I 誤差的機率 α 就愈大, 型 II 誤差 β 的機率就愈小。

3. 增加樣本數 n 就可以使得 α 與 β 同時降低。

4. 顯著水準是參數值為虛無假設邊界點的 α 值。

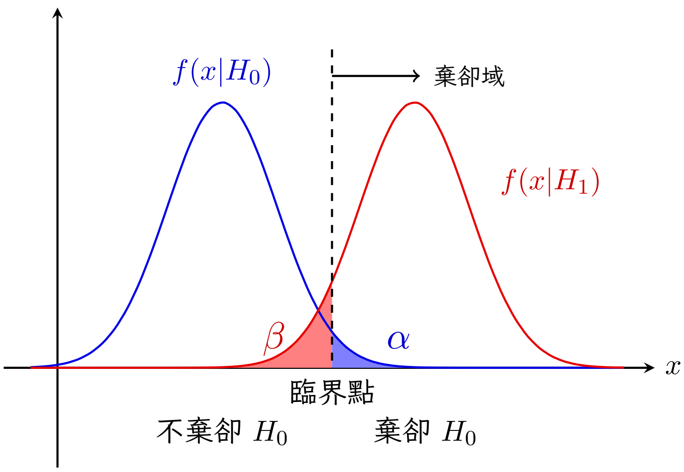


圖 7.1: 型 I 誤差與型 II 誤差關聯圖

例 **7–1** 某奶粉代理商認為台灣地區嬰兒奶粉消費戶所佔的比率為 0.4。為了檢定此一說法是否成立, 隨機抽取了 16 個家庭之戶長，若此 16 戶家庭中有 0 至 8 戶是嬰兒奶粉消費戶，則接受 $H_0: P = 0.4$; 否則, 接受 $H_1: P = 0.5$。試求：

1. 危險域？

2. 型 I 誤差？

3. 型 II 誤差？

解：

1. 令 X 表 16 戶家庭中該嬰兒奶粉之消費戶，所以，$X \sim B(16, 0.4)$。因此，危險域為

$$C = \{x \mid x \geq 9\}$$

2.

$$\alpha = \mathrm{P}(x \geq 9 \mid P = 0.4) = \sum_{x=9}^{16} \binom{16}{x} 0.4^x 0.6^{16-x} = 0.1423$$

3.

$$\beta = \mathrm{P}(x \leq 8 \mid P = 0.5) = \sum_{x=0}^{8} \binom{16}{x} 0.5^x 0.5^{16-x} = 0.5982$$

例 **7–2** 令樣本平均數 $\overline{X}$ 服從 $\mathcal{N}(\mu, 16/n)$, n 為樣本數，如果觀察值 $\overline{X} < c$, 我們就否定 H_0：$\mu \geq 20$ 而接受 H_1：$\mu < 20$。因此, 檢定力函數 $K(\mu) = \mathrm{P}(\overline{X} < c; \mu)$。試求滿足 $K(20) = 0.05$ 和 $K(19) = 0.90$ 的樣本數 n 及 c 之值。

解：

$$\begin{aligned} K(20) &= \mathrm{P}\left(\overline{X} < c \mid \mu = 20\right) \\ &= \mathrm{P}\left(\frac{\overline{X} - 20}{4/\sqrt{n}} < \frac{c - 20}{4/\sqrt{n}}\right) \\ &= \mathrm{P}\left(Z < \frac{c - 20}{4/\sqrt{n}}\right) = 0.05 \end{aligned}$$

$$\begin{aligned} K(19) &= \mathrm{P}\left(\overline{X} < c \mid \mu = 19\right) \\ &= \mathrm{P}\left(\frac{\overline{X} - 19}{4/\sqrt{n}} < \frac{c - 19}{4/\sqrt{n}}\right) \\ &= \mathrm{P}\left(Z < \frac{c - 19}{4/\sqrt{n}}\right) = 0.90 \end{aligned}$$

得下列二元一次方程組

$$\begin{cases} \dfrac{c - 20}{4/\sqrt{n}} = -1.645 \\ \dfrac{c - 19}{4/\sqrt{n}} = 1.28 \end{cases}$$

解上式得 $c = 19.4376$ 與 $n = 136.89$

例 **7–3** 假設隨機變數 X 的 p.d.f. 是

$$f(x, \theta) = \begin{cases} (1/\theta)\, e^{-x/\theta}, & 0 < x < \infty \\ 0, & \text{其他} \end{cases}$$

今有兩隨機變數 X_1 和 X_2, 並設定其棄卻域為

$$C=\{(X_1,X_2):8\leq X_1+X_2\leq\infty\}$$

試求當 H_0 是正確時; X_1 和 X_2 的聯合機率密度為何, 並計算型 I 誤差 (Type I Error) 的機率。

解：

1.
$$f\left(x_1,x_2\mid\theta=1\right)=e^{-(x_1+x_2)},\ 0<x_1<\infty,\ 0<x_2<\infty$$

2.
$$\begin{aligned}\alpha &= \mathrm{P}\left(8\leq X_1+X_2\leq\infty\mid\theta=1\right)\\ &= 1-\mathrm{P}\left(0\leq X_1+X_2\leq 8\mid\theta=1\right)\\ &= 1-\int_0^8\int_0^{8-x_2}e^{-(x_1+x_2)}dx_1dx_2\\ &= \frac{9}{e^8}\end{aligned}$$

7.1.3 檢定的型態

檢定之模式有以下三種

1. **雙尾檢定**：即雙尾對立假設的檢定問題，其統計假設為

$$\begin{cases}H_0:\theta=\theta_0\\ H_1:\theta\neq\theta_0\end{cases}$$

2. **右尾檢定**：即右尾對立假設的檢定問題，其統計假設為

$$\begin{cases}H_0:\theta=\theta_0\\ H_1:\theta>\theta_0\end{cases}\quad\text{或}\quad\begin{cases}H_0:\theta\leq\theta_0\\ H_1:\theta>\theta_0\end{cases}$$

3. **左尾檢定**：即左尾對立假設的檢定問題，其統計假設為

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases} \text{或} \begin{cases} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$$

虛無假設 H_0 與對立假設 H_1 之關係如下：

1. 均屬集合形式。
2. $\Theta_0 \cap \Theta_1 = \phi$, $\Theta_0 \cup \Theta_1 = \Theta$，其中 Θ_0 表 θ_0 之參數空間, Θ_1 表 θ_1 之參數空間及 Θ 表 θ 之參數空間。
3. 等號必存在 Θ_0。

7.1.4 檢定方法

檢定的原理乃將虛無假設 $H_0\theta = \theta_0$ 與經抽樣取得之樣本統計量 $\widehat{\theta}_0$ 比較，若 $|\theta_0 - \widehat{\theta}_0|$ 值愈小，即 H_0 為正確的決策之機愈大。因此, 在做相關決策時則必須有一比較之標準以決定棄卻 H_0 或接受 H_1, 常用的方法有以下兩種：

1. 利用檢定統計量與臨界值之比較，其決策法則為當檢定統計量落入拒絕區，則棄卻 H_0, 否則不棄卻 H_0。
2. 利用 P-Value 與顯著水準 α 值之比較, 決策法則為當 P-Value **小於** α 時，則棄卻 H_0, 否則不棄卻 H_0。

7.2 母體平均數之檢定

7.2.1 單一母體平均數之檢定

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma^2)$ 的一組樣本數為 n 的隨機樣本。若母體變異數 σ^2 已知，母體平均數之檢定方法介紹如下：

μ 之雙尾檢定

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

由於統計假設為要檢定 μ 是否等於 μ_0，因此若經抽樣所得的樣本平均數 $\overline{x}$ 與 μ_0 差異甚大則棄卻 H_0 虛無假設。由型 I 誤差之定義可知

$$\begin{aligned} \alpha &= \mathrm{P}\left(\overline{X} < c_1 \text{ 或 } \overline{X} > c_2 \,\middle|\, \mu = \mu_0\right) \\ &= \mathrm{P}\left(\overline{X} - \mu_0 < c_1 - \mu_0\right) + \mathrm{P}\left(\overline{X} - \mu_0 > c_2 - \mu_0\right) \\ &= \mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < \frac{c_1 - \mu_0}{\sigma/\sqrt{n}}\right) + \mathrm{P}\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{c_2 - \mu_0}{\sigma/\sqrt{n}}\right) \\ &= \mathrm{P}\left(Z < \frac{c_1 - \mu_0}{\sigma/\sqrt{n}}\right) + \mathrm{P}\left(Z > \frac{c_2 - \mu_0}{\sigma/\sqrt{n}}\right) \end{aligned}$$

由於為雙尾檢定，故採左右尾之機率均為 $\alpha/2$。由圖7.2可知，斜線部分卽為棄卻域。故可求得

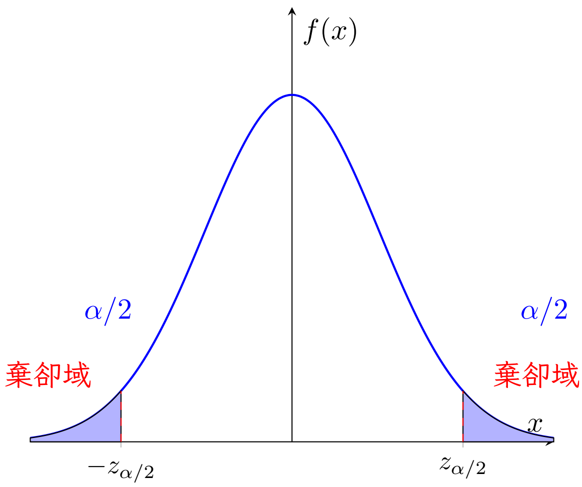


圖 7.2: 雙尾檢定圖

$$-z_{\frac{\alpha}{2}} = \frac{c_1 - \mu_0}{\sigma/\sqrt{n}} \text{ 及} z_{\frac{\alpha}{2}} = \frac{c_2 - \mu_0}{\sigma/\sqrt{n}}$$

或

$$c_1 = \mu_0 - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \text{ 及} c_2 = \mu_0 + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

所以，棄卻域為

$$C=\left\{\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|>z_{\frac{\alpha}{2}} \text{ 或 } |\overline{x}-\mu_0|>z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right\},$$

當

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|>z_{\frac{\alpha}{2}} \text{ 或 } |\overline{x}-\mu_0|>z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

則棄卻虛無假設 $H_0:\mu=\mu_0$, 否則不棄卻 $H_0:\mu=\mu_0$。

以上所介紹之檢定方法只在母體變異數 σ^2 已知的假設下成立，若母體變異數未知，且樣本數 $n<30$，則需將樣本變異數 S^2 做為 σ^2 之估計式。又

$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t\,(n-1)$$

所以，棄卻域 C 可改寫為

$$C=\left\{\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|>t_{\frac{\alpha}{2}}\,(n-1) \text{ 或 } |\overline{x}-\mu_0|>t_{\frac{\alpha}{2}}\,(n-1)\,\frac{s}{\sqrt{n}}\right\}$$

μ 之右尾檢定

$$\begin{cases} H_0:\mu=\mu_0 \\ H_1:\mu>\mu_0 \end{cases}$$

由於統計假設為要檢定 μ 是否大於 μ_0，因此若經抽樣所得的樣本平均數 $\overline{x}$ 有顯著的大於 μ_0 則棄卻 H_0 虛無假設。由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\overline{X}>c_1 \mid \mu=\mu_0\right)\\
&= \mathrm{P}\left(\overline{X}-\mu_0>c_1-\mu_0\right)\\
&= \mathrm{P}\left(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}>\frac{c_1-\mu_0}{\sigma/\sqrt{n}}\right)\\
&= \mathrm{P}\left(Z>\frac{c_1-\mu_0}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

由圖7.3可知，斜線部分即為棄卻域。故可求得

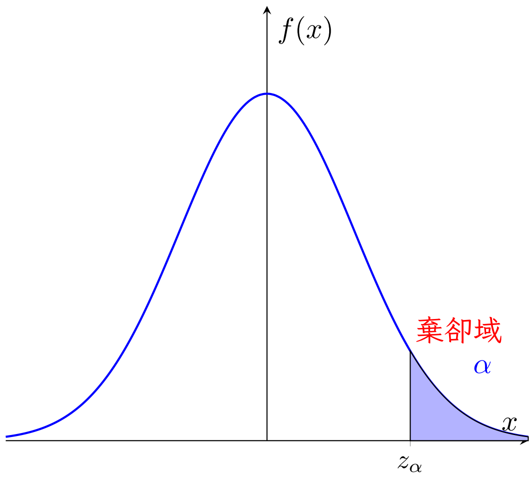


圖 7.3: 右尾檢定圖

$$z_{\frac{\alpha}{2}} = \frac{c_1 - \mu_0}{\sigma/\sqrt{n}} \text{ 或 } c_1 = \mu_0 + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

所以，當

$$\frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}} \text{ 或} \overline{x} - \mu_0 > z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

則棄卻虛無假設 $H_0 : \mu = \mu_0$, 否則不棄卻 $H_0 : \mu = \mu_0$。若母體變異數未知，且樣本數 $n < 30$，則棄卻域 C 可改寫為

$$C = \left\{ \frac{\overline{x} - \mu_0}{s/\sqrt{n}} > t_{\frac{\alpha}{2}} (n-1) \text{ 或} \overline{x} - \mu_0 > t_{\frac{\alpha}{2}} (n-1) \frac{s}{\sqrt{n}} \right\}$$

μ 之左尾檢定

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

由於統計假設為要檢定 μ 是否小於 μ_0，因此若經抽樣所得的樣本平均數 $\overline{x}$ 有顯著的小於 μ_0 則棄卻 H_0 虛無假設。由型 I 誤差之定義可知

$$\begin{aligned} \alpha &= \mathrm{P}\left(\overline{X} < c_1 \mid \mu = \mu_0\right) \\ &= \mathrm{P}\left(\overline{X} - \mu_0 < c_1 - \mu_0\right) \end{aligned}$$

$$
\begin{aligned}
&= \mathrm{P}\left(\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} < \frac{c_1-\mu_0}{\sigma/\sqrt{n}}\right) \\
&= \mathrm{P}\left(Z < \frac{c_1-\mu_0}{\sigma/\sqrt{n}}\right)
\end{aligned}
$$

由圖7.4可知，斜線部分卽為棄卻域。故可求得

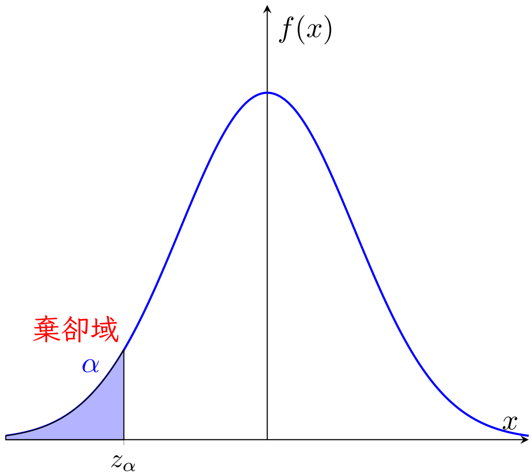


圖 7.4: 左尾檢定圖

$$
-z_{\frac{\alpha}{2}} = \frac{c_1-\mu_0}{\sigma/\sqrt{n}} \text{ 或 } c_1 = \mu_0 - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}
$$

所以，當

$$
\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} < -z_{\frac{\alpha}{2}} \text{ 或} \overline{x}-\mu_0 < -z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}
$$

則棄卻虛無假設 $H_0:\mu=\mu_0$, 否則不棄卻 $H_0:\mu=\mu_0$。若母體變異數未知，且樣本數 $n<30$，則棄卻域 C 可改寫為

$$
C = \left\{\frac{\overline{x}-\mu_0}{s/\sqrt{n}} < -t_{\frac{\alpha}{2}}(n-1) \text{ 或} \overline{x}-\mu_0 < -t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}\right\}
$$

例 **7–4** 某工廠的一項產品的壽命，歷年來一直維持為平均壽命 $\mu=900$ 小時，標準差 $\sigma=180$ 小時。今欲瞭解因生產原料的改變, 產品的平均壽命是否仍維持以往水準, 特由目前全部新產品中隨機抽出 200 件, 測得其樣本平均壽命為 935 小時，試依下列方法, 以 $\alpha=0.05$ 檢定 $H_0:\mu-900$(小時)

1. 以檢定統計量法, 進行檢定

2. 以 P-Value, 進行檢定

3. 以信賴區間, 進行檢定

解：

1. 因為

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{935-900}{180/\sqrt{200}}\right| = 2.749 > 1.96 = z_{0.025}$$

落入棄卻域，所以棄卻 H_0

2.

$$\begin{aligned} \text{P-Value} &= 2\text{P}\left(\overline{x} > 935 | \, \mu = 900\right) \\ &= 2\text{P}\left(Z > \frac{935-900}{180/\sqrt{200}}\right) = 2 \times \text{P}\left(Z > 2.749\right) \\ &= 2 \times 0.0031 = 0.0062 < 0.05 \end{aligned}$$

因為 P-Value 小於 0.05，所以棄卻 H_0

3.

$$\left[\overline{x} - z_{0.025} \times \frac{\sigma}{\sqrt{n}}, \overline{x} + z_{0.025} \times \frac{\sigma}{\sqrt{n}}\right] = [910.05, 959.95]$$

因為不包含 $\mu_0 = 900$，所以棄卻 H_0

例 **7–5** 為了控制品質, 品管工程師對於所生產罐裝調味料的重量是否都是 16 盎斯進行檢驗。因此, 從生產線上隨機抽取了 10 個成品測量它們的重量如下：

16.3, 16.2, 15.8, 15.4, 16.0, 15.6, 15.5, 16.1, 15.9, 16.1

在顯著水準 $\alpha = 0.05$ 之下，試問罐裝調味料裝填機是否需要做調整？

解：

$$\begin{cases} H_0 : \mu = 16 \\ H_1 : \mu \neq 16 \end{cases}$$

由於 $n = 10$, $\overline{x} = 15.890$, $s = 0.307$，所以棄卻域為

$$C = \left\{\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| > t_{0.025}\left(9\right)\right\}$$

又

$$\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| = \left|\frac{15.890-16}{0.307/\sqrt{10}}\right| = 1.13 < 2.262 = t_{0.025}\,(9)$$

所以不棄卻 H_0，即沒有充分的證據顯示所生產的罐裝調味料的重量會顯著異於 16 盎斯。因此，機器不需調整。

7.2.2 兩母體平均數差之檢定

假設 $X_1, X_2, \ldots, X_{n_1}; Y_1, Y_2, \ldots, Y_{n_2}$ 為分別抽自 $\mathcal{N}(\mu_x, \sigma_x^2)$ 及 $\mathcal{N}(\mu_y, \sigma_y^2)$ 的兩組樣本數為 n_1 及 n_2 的隨機樣本, $\mu_x - \mu_y$ 之檢定方法介紹如下：

$\mu_1 - \mu_2$ 之雙尾檢定且 σ_x^2, σ_y^2 已知

$$\begin{cases} H_0 : \mu_x - \mu_y = k \\ H_1 : \mu_x - \mu_y \neq k \end{cases}$$

由於統計假設為要檢定 $\mu_x - \mu_y$ 是否等於 k，因此若經抽樣所得的樣本平均數 $\overline{x} - \overline{y}$ 與 k 差異甚大則棄卻 H_0 虛無假設。由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\overline{X}-\overline{Y} < c_1 \text{ 及 } \overline{X}-\overline{Y} > c_2 \,\middle|\, \mu_x - \mu_y = k\right) \\
&= \mathrm{P}\left(\left(\overline{X}-\overline{Y}\right) - k < c_1 - k\right) + \mathrm{P}\left(\left(\overline{X}-\overline{Y}\right) - k > c_2 - k\right) \\
&= \mathrm{P}\left(\frac{\overline{X}-\overline{Y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}} < \frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right) + \mathrm{P}\left(\frac{\overline{X}-\overline{Y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}} > \frac{c_2-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right) \\
&= \mathrm{P}\left(Z < \frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right) + \mathrm{P}\left(Z > \frac{c_2-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right)
\end{aligned}$$

由於為雙尾檢定，故採左右尾之機率均為 $\frac{\alpha}{2}$。故可求得

$$-z_{\frac{\alpha}{2}} = \frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}} \text{ 及} z_{\frac{\alpha}{2}} = \frac{c_2-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}$$

或

$$c_1 = k - z_{\frac{\alpha}{2}}\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2} \text{ 及} c_2 = k + z_{\frac{\alpha}{2}}\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}$$

所以，可求得棄卻域為

$$C=\left\{\left|\frac{\overline{x}-\overline{y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right|>z_{\frac{\alpha}{2}} \text{ 或} \left|\overline{x}-\overline{y}-k\right|>z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1}+\frac{\sigma_y^2}{n_2}}\right\},$$

當

$$\left|\frac{\overline{x}-\overline{y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right|>z_{\frac{\alpha}{2}} \text{ 或} \left|\overline{x}-\overline{y}-k\right|>z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_x^2}{n_1}+\frac{\sigma_y^2}{n_2}}$$

時棄卻虛無假設 $H_0:\mu_x-\mu_y=k$, 否則不棄卻虛無假設。當母體變異數 σ_x^2 及 σ_y^2 未知時，若兩母體變異數相等 $(\sigma_x^2=\sigma_y^2=\sigma^2)$, 由 **7.2.4** 可知

$$\frac{\overline{x}-\overline{y}-k}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t\left(n_1+n_2-2\right)$$

所以，由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\overline{X}-\overline{Y}<c_1 \text{ 及 } \overline{X}-\overline{Y}>c_2 \,\middle|\, \mu_x-\mu_y=k\right)\\
&= \mathrm{P}\left(\overline{X}-\overline{Y}-k<c_1-k\right)+\mathrm{P}\left(\overline{X}-\overline{Y}-k>c_2-k\right)\\
&= \mathrm{P}\left(\frac{\overline{X}-\overline{Y}-k}{S_p\sqrt{1/n_1+1/n_2}}<\frac{c_1-k}{S_p\sqrt{1/n_1+1/n_2}}\right)+\mathrm{P}\left(\frac{\overline{X}-\overline{Y}-k}{S_p\sqrt{1/n_1+1/n_2}}>\frac{c_2-k}{S_p\sqrt{1/n_1+1/n_2}}\right)\\
&= \mathrm{P}\left(t<\frac{c_1-k}{S_p\sqrt{1/n_1+1/n_2}}\right)+\mathrm{P}\left(t>\frac{c_2-k}{S_p\sqrt{1/n_1+1/n_2}}\right)
\end{aligned}$$

由於為雙尾檢定，故可求得

$$-t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)=\frac{c_1-k}{S_p\sqrt{1/n_1+1/n_2}} \text{ 或} t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)=\frac{c_2-k}{S_p\sqrt{1/n_1+1/n_2}}$$

或

$$c_1=k-t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \quad c_2=k+t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},$$

其中

$$S_p^2=\frac{(n_1-1)\,S_x^2+(n_2-1)\,S_y^2}{n_1+n_2-2}$$

因此，棄卻域可寫成

$$C=\left\{\left|\frac{\overline{x}-\overline{y}-k}{s_p\sqrt{1/n_1+1/n_2}}\right|>t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\ \text{或}\,|\overline{x}-\overline{y}-k|>t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\}$$

所以，當

$$\left|\frac{\overline{x}-\overline{y}-k}{s_p\sqrt{1/n_1+1/n_2}}\right|>t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\ \text{或}\,|\overline{x}-\overline{y}-k|>t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\,S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},$$

時，棄卻虛無假設, 否則不棄卻虛無假設。

$\mu_1-\mu_2$ 之右尾檢定且 σ_x^2, σ_y^2 已知

$$\begin{cases} H_0:\mu_x-\mu_y=k\\ H_1:\mu_x-\mu_y>k\end{cases}$$

由於統計假設為要檢定 $\mu_x-\mu_y$ 是否大於 k，因此若經抽樣所得的樣本平均數 $\overline{x}-\overline{y}$ 有顯著大於 k 時則棄卻 H_0 虛無假設。由型 I 誤差之定義可知

$$\begin{aligned}\alpha &= \mathrm{P}\left(\overline{X}-\overline{Y}>c_1\mid\mu_x-\mu_y=k\right)\\ &= \mathrm{P}\left(\overline{X}-\overline{Y}-k>c_1-k\ \right)\\ &= \mathrm{P}\left(\frac{\overline{X}-\overline{Y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}>\frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right)\\ &= \mathrm{P}\left(Z>\frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\right)\end{aligned}$$

故可求得

$$z_\alpha=\frac{c_1-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}\ \text{或}c_1=k+z_\alpha\sqrt{\frac{\sigma_x^2}{n_1}+\frac{\sigma_y^2}{n_2}}$$

因此，棄卻域為

$$C=\left\{\frac{\overline{x}-\overline{y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}>z_\alpha\right\}$$

所以，當

$$\frac{\overline{x}-\overline{y}-k}{\sqrt{\sigma_x^2/n_1+\sigma_y^2/n_2}}>z_\alpha$$

時棄卻虛無假設 $H_0:\mu_x-\mu_y=k$, 否則不棄卻虛無假設。當母體變異數 σ_x^2 及 σ_y^2 未知時，若兩母體變異數相等 $(\sigma_x^2=\sigma_y^2=\sigma)$，則棄卻域可寫成

$$C=\left\{\frac{\overline{x}-\overline{y}-k}{s_p\sqrt{1/n_1+1/n_2}}>t_\alpha\left(n_1+n_2-2\right)\ \right\}$$

例 **7–6** LAST 是法律學校的入學檢定考試。今有兩所法律學校想要比較在這兩校登記的學生的 LAST 平均成績是否有顯著差異性？特於二校中分別抽出二獨立隨機樣本並計算得

A 校：$n_1=15,\overline{x}=680,s_x=84$

B 校：$n_2=21,\overline{y}=634,s_y=92$

使用顯著水準 $\alpha=0.05$ 檢定之, 假設兩校的母體變異數相等。

解：

$$\begin{cases}H_0:\mu_A=\mu_B\\H_1:\mu_A\neq\mu_B\end{cases}$$

$$s_p^2=\frac{(15-1)\times 84^2+(21-1)\times 92^2}{15+21-2}=7884.\,2$$

所以棄卻域為

$$C=\left\{\left|\frac{\overline{x}-\overline{y}}{s_p\sqrt{1/n_1+1/n_1}}\right|>t_{0.025}\left(34\right)=1.96\right\}$$

又

$$\left|\frac{\overline{x}-\overline{y}}{s_p\sqrt{1/n_1+1/n_1}}\right|=\left|\frac{680-634}{\sqrt{7884.2}\sqrt{1/15+1/21}}\right|=1.5324$$

並未落入棄卻域，所以不棄卻 H_0，即兩所法律學校登記的學生的 LAST 平均成績差不多。

例 **7–7** 續例 **8-5**，若現在甲校想要知道該校登記的學生的 LAST 平均成績是否會比乙校登記的學生的 LAST 平均成績高出 10 分以上？試檢定之。

解：

$$\begin{cases} H_0 : \mu_A - \mu_B = 10 \\ H_1 : \mu_A - \mu_B > 10 \end{cases}$$

$$s_p^2 = \frac{(15-1)\times 84^2 + (21-1)\times 92^2}{15+21-2} = 7884.2$$

所以棄卻域為

$$C = \left\{ \left| \frac{\overline{x} - \overline{y} - 10}{s_p\sqrt{1/n_1 + 1/n_2}} \right| > t_{0.05}(34) = 1.645 \right\}$$

又

$$\left| \frac{\overline{x} - \overline{y} - 10}{s_p\sqrt{1/n_1 + 1/n_2}} \right| = \left| \frac{680 - 634 - 10}{\sqrt{7884.2}\sqrt{1/15 + 1/21}} \right| = 1.1993$$

並未落入棄卻域，所以不棄卻 H_0，即在 $\alpha = 0.01$ 的水準之下, 甲校登記學生的 AST 平均成績並沒有比乙校登記學生的 LAST 平均成績高出 10 分以上。

例 **7–8** 某製造業的公司為了要瞭解哪一家廠商所提供的零件其品質比較好, 以便能夠決定到底要下訂單給甲廠商或是乙廠商。因此分別各從兩家供應商隨機抽取 16 個產品來作試驗, 得到甲供應商的樣本平均數 $\overline{X} = 300$，樣本標準差 $S_1 = 15$，而乙供應商的樣本平均數 $\overline{Y} = 276$，樣本標準差 $S_2 = 18$。再假設全體分配近似於常態分配的情況下，試以 $\alpha = 0.05$ 為顯著水準, 檢定兩家供應商所提供的產品品質是否有差異。(假設兩家廠商的變異數相等)

解：

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$$

$$C = \left\{ \left| \frac{\overline{x} - \overline{y}}{s_p\sqrt{1/16 + 1/16}} \right| > t_{0.025}(30) = 2.042 \right\}$$

其中

$$s_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{15\times 15^2 + 15\times 18^2}{30} = 274.5$$

所以

$$\left| \frac{300 - 276}{\sqrt{274.5}\sqrt{1/16 + 1/16}} \right| = 4.0972 \in C$$

故棄卻 H_0

7.3 母體變異數之檢定

7.3.1 單一母體變異數之檢定

假設 $X_1, X_2, \ldots, X_n$ 為抽自 $\mathcal{N}(\mu, \sigma_x^2)$ 的一組樣本數為 n 的隨機樣本。若母體變異數 σ^2 未知，則母體變異數檢定之方法介紹如下：

母體變異數 σ^2 之雙尾檢定

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

由於統計假設為要檢定 σ^2 是否等於 σ_0^2，因此若經抽樣所得的樣本變異數 S_x^2 與 σ_0^2 差異甚大則棄卻 H_0 虛無假設。又

$$\frac{(n-1)S_x^2}{\sigma^2} \sim \chi^2(n-1),$$

因此, 由型 I 誤差之定義可知

$$\begin{aligned} \alpha &= \mathrm{P}\left(S_x^2 \ < c_1 \text{ 或} S_x^2 > c_2 \middle| \sigma_x^2 = \sigma_0^2\right) \\ &= \mathrm{P}\left(\frac{(n-1)\,S_x^2}{\sigma_0^2} < \frac{(n-1)\,c_1}{\sigma_0^2}\right) + \mathrm{P}\left(\frac{(n-1)\,S_x^2}{\sigma_0^2} > \frac{(n-1)\,c_2}{\sigma_0^2}\right) \\ &= \mathrm{P}\left(\chi^2 < \frac{(n-1)\,c_1}{\sigma_0^2}\right) + \mathrm{P}\left(\chi^2 > \frac{(n-1)\,c_2}{\sigma_0^2}\right) \end{aligned}$$

由於為雙尾檢定，故採左右尾之機率均為 $\frac{\alpha}{2}$。故可求得

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \frac{(n-1)\,c_1}{\sigma_0^2} \text{ 及} \chi^2_{\frac{\alpha}{2}0}(n-1) = \frac{(n-1)\,c_2}{\sigma_0^2}$$

或

$$c_1 = \frac{\sigma_0^2}{n-1}\chi^2_{1-\frac{\alpha}{2}}(n-1) \text{ 及} c_2 = \frac{\sigma_0^2}{n-1}\chi^2_{\frac{\alpha}{2}}(n-1)$$

故棄卻域可寫成

$$C=\left\{s_x^2<\frac{\sigma_0^2}{n-1}\chi^2_{1-\frac{\alpha}{2}}(n-1)\text{ 或}s_x^2>\frac{\sigma_0^2}{n-1}\chi^2_{\frac{\alpha}{2}}(n-1)\right\}$$

所以，當

$$s_x^2<\frac{\sigma_0^2}{n-1}\chi^2_{1-\frac{\alpha}{2}}(n-1)\text{ 或}s_x^2>\frac{\sigma_0^2}{n-1}\chi^2_{\frac{\alpha}{2}}(n-1)$$

時棄卻虛無假設 $H_0:\sigma^2=\sigma_0^2$, 否則不棄卻虛無假設。

母體變異數 σ_x^2 之右尾檢定

$$\begin{cases} H_0:\sigma^2=\sigma_0^2 \\ H_1:\sigma^2>\sigma_0^2 \end{cases}$$

由於統計假設為要檢定 σ^2 是否大於 σ_0^2，所以若經抽樣所得的樣本變異數 S^2 有顯著的大於 σ_0^2，則棄卻 H_0 虛無假設。因此，由型 I 誤差之定義可知

$$\begin{aligned}\alpha &= \mathrm{P}\left(S_x^2>c_1\middle|\,\sigma_x^2=\sigma_0^2\right)\\ &= \mathrm{P}\left(\frac{(n-1)\,S_x^2}{\sigma_0^2}>\frac{(n-1)\,c_1}{\sigma_0^2}\right)\\ &= \mathrm{P}\left(\chi^2>\frac{(n-1)\,c_1}{\sigma_0^2}\right)\end{aligned}$$

故可求得

$$\chi^2_\alpha(n-1)=\frac{(n-1)\,c_1}{\sigma_0^2}\text{ 或}c_1=\frac{\sigma_0^2}{n-1}\chi^2_\alpha(n-1)$$

所以，棄卻域可寫成

$$C=\left\{s_x^2>\frac{\sigma_0^2}{n-1}\chi^2_\alpha(n-1)\right\}$$

例 **7–9** 自常態分配 $\mathcal{N}(\mu,\sigma^2)$ 之母體中隨機抽出 10 的樣本得一組觀測值為

10.5, 10.2, 9.6, 10.8, 9.3, 9.8, 9.9, 10.1, 9.5, 10.0

試以顯著水準 $\alpha=0.1$ 下，檢定其母體變異數是否為 $\sigma^2=0.09$？

解：

$$\begin{cases} H_0 : \sigma^2 = 0.09 \\ H_1 : \sigma^2 \neq 0.09 \end{cases}$$

$$C = \left\{ s_x^2 < \frac{\sigma_0^2}{n-1} \chi^2_{1-\frac{\alpha}{2}(9)} = 0.0325 \text{ 或} s_x^2 > \frac{\sigma_0^2}{n-1} \chi^2_{\frac{\alpha}{2}}(9) = 0.1691 \right\}$$

又

$$s_x^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1} = 0.209$$

落入棄卻域內，所以棄卻 H_0，即 $\sigma^2 \neq 0.09$。

7.3.2 兩母體變異數比例之檢定

假設 $X_1, X_2, \ldots, X_{n_1}; Y_1, Y_2, \ldots, Y_{n_2}$ 為分別抽自 $\mathcal{N}(\mu_x, \sigma_x^2)$ 及 $\mathcal{N}(\mu_y, \sigma_y^2)$ 的兩組樣本數為 n_1 及 n_2 的隨機樣本, 其中 σ_x^2 及 σ_y^2 未知且 $n_1 < 30$ 及 $n_2 < 30$。由 **7.2.7** 可知，

$$\frac{S_x^2}{S_y^2} \frac{\sigma_y^2}{\sigma_x^2} \sim \mathcal{F}(n_1 - 1, n_2 - 1)$$

所以，兩母體變異數比例之檢定方法介紹如下：

兩母體變異數比例 σ_x^2/σ_y^2 之雙尾檢定

$$\begin{cases} H_0 : \sigma_x^2/\sigma_y^2 = k \\ H_1 : \sigma_x^2/\sigma_y^2 \neq k \end{cases}$$

由於統計假設為要檢定 σ_x^2/σ_y^2 是否等於 k，因此若經抽樣所得的樣本變異數之比例 S_x^2/S_y^2 與 k 差異甚大則棄卻 H_0 虛無假設。因此，由型 I 誤差之定義可知

$$\begin{aligned} \alpha &= \mathrm{P}\left(\frac{S_x^2}{S_y^2} < c_1 \text{ 或} \frac{S_x^2}{S_y^2} > c_2 \middle| \frac{\sigma_x^2}{\sigma_y^2} = k \right) \\ &= \mathrm{P}\left(\frac{S_x^2}{S_y^2} \frac{1}{k} < \frac{c_1}{k} \right) + \mathrm{P}\left(\frac{S_x^2}{S_y^2} \frac{1}{k} > \frac{c_2}{k} \right) \end{aligned}$$

$$= \quad \mathrm{P}\left(F < \frac{c_1}{k}\right) + \mathrm{P}\left(F > \frac{c_2}{k}\right)$$

由於為雙尾檢定，故採左右尾之機率均為 $\frac{\alpha}{2}$。故可求得

$$F_{1-\frac{\alpha}{2}}\left(n_1-1, n_2-1\right) = \frac{c_1}{k} \text{ 及} F_{\frac{\alpha}{2}}\left(n_1-1, n_2-1\right) = \frac{c_2}{k}$$

或

$$c_1 = kF_{1-\frac{\alpha}{2}}\left(n_1-1, n_2-1\right) \text{ 及} c_2 = kF_{\frac{\alpha}{2}}\left(n_1-1, n_2-1\right)$$

故可求得棄卻域為

$$C = \left\{\frac{s_x^2}{s_y^2} < kF_{1-\frac{\alpha}{2}}\left(n_1-1, n_2-1\right) \text{ 或} \frac{s_x^2}{s_y^2} > kF_{\frac{\alpha}{2}}\left(n_1-1, n_2-1\right)\right\}$$

所以當

$$C = \left\{\frac{s_x^2}{s_y^2} < kF_{1-\frac{\alpha}{2}}\left(n_1-1, n_2-1\right) \text{ 或} \frac{s_x^2}{s_y^2} > kF_{\frac{\alpha}{2}}\left(n_1-1, n_2-1\right)\right\}$$

時，時棄卻虛無假設 $H_0 : \sigma_x^2/\sigma_y^2 = k$, 否則不棄卻虛無假設。

兩母體變異數比例 σ_x^2/σ_y^2 之右尾檢定

由於統計假設為要檢定 σ_x^2/σ_y^2 是否大於 k，所以若經抽樣所得的樣本變異數 σ_x^2/σ_y^2 有顯著的大於 k，則棄卻 H_0 虛無假設。因此, 由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\frac{S_x^2}{S_y^2} > c_1 \middle| \frac{\sigma_x^2}{\sigma_y^2} = k\right) \\
&= \mathrm{P}\left(\frac{S_x^2}{S_y^2}\frac{1}{k} > \frac{c_1}{k}\right) \\
&= \mathrm{P}\left(F > \frac{c_1}{k}\right)
\end{aligned}$$

故可求得

$$F_{\alpha}\left(n_1-1, n_2-1\right) = \frac{c_1}{k} \text{ 或} c_1 = kF_{\alpha}\left(n_1-1, n_2-1\right)$$

所以，棄卻域可寫成

$$C=\left\{\frac{s_x^2}{s_y^2}>kF_\alpha\left(n_1-1,n_2-1\right)\right\}$$

例 **7–10** 檢查兩台產製相同產品的機器所生產產品在規格上有否顯著性差異, 分別獨立地由兩台機器製品中隨機抽出二組樣本並計算得資料如下：

$$\text{A：}n_1=31,s_x^2=25$$
$$\text{B：}n_2=25,s_y^2=12$$

試以 $\alpha=0.1$ 檢定此兩台機器之穩定性是否一致？

解：

$$\begin{cases} H_0:\sigma_x^2=\sigma_y^2 \\ H_1:\sigma_x^2\neq\sigma_y^2 \end{cases}$$

棄卻域為

$$C=\left\{\frac{s_x^2}{s_y^2}<0.5291=F_{0.95}(n_1,n_2)\text{ 或}\frac{s_x^2}{s_y^2}=1.94=F_{0.05}(n_1,n_2)\right\}$$

又

$$\frac{s_x^2}{s_y^2}=\frac{25}{12}=2.083$$

落入棄卻域內，所以棄卻 H_0，即機器產品之穩性不一致。

7.4 母體比例之檢定

7.4.1 單一母體比例之檢定

假設 X_1, X_2, ..., X_n 為抽自 Bernoulli(p) 的一組樣本數為 n 的隨機樣本。令 $Y=\sum_{i=1}^n X_i/n$, 由 **7.2.8** 可知

$$\frac{Y-p}{\sqrt{\mathrm{P}\left(1-p\right)/n}}\backsim\mathcal{N}\left(0,1\right)$$

母體比例 p 之檢定方法介紹如下：

單一母體比例 p 之雙尾檢定

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

由於統計假設為要檢定 p 是否等於 p_0，因此若經抽樣所得的樣本比例 $\widehat{p} = \sum_{i=1}^{n} X_i/n$ 與 p_0 差異甚大則棄卻 H_0 虛無假設。因此，由型 I 誤差之定義可知

$$\begin{aligned} \alpha &= \mathrm{P}\left(\widehat{p} < c_1 \text{ 或} \widehat{p} > c_2 \middle| p = p_0\right) \\ &= \mathrm{P}\left(\frac{\widehat{p} - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} < \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right) + \mathrm{P}\left(\frac{\widehat{p} - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} > \frac{c_2 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right) \\ &= \mathrm{P}\left(Z < \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right) + \mathrm{P}\left(Z > \frac{c_2 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right) \end{aligned}$$

因為為雙尾檢定, 取兩邊機率相等可推得

$$-z_{\frac{\alpha}{2}} < \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} \text{ 及} z_{\frac{\alpha}{2}} > \frac{c_2 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}$$

或

$$c_1 = p_0 - z_{\frac{\alpha}{2}} \sqrt{\frac{p_0\left(1 - p_0\right)}{n}} \text{ 及} c_2 = p_0 + z_{\frac{\alpha}{2}} \sqrt{\frac{p_0\left(1 - p_0\right)}{n}}$$

所以，棄卻域可寫成

$$C = \left\{\widehat{p} < p_0 - z_{\frac{\alpha}{2}} \sqrt{\frac{p_0\left(1 - p_0\right)}{n}} \text{ 或} \widehat{p} > p_0 + z_{\frac{\alpha}{2}} \sqrt{\frac{p_0\left(1 - p_0\right)}{n}}\right\}$$

單一母體比例 p 之右尾檢定

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0 \end{cases}$$

由於統計假設為要檢定 p 是否大於 p_0，因此若經抽樣所得的樣本比例 $\widehat{p} = \sum_{i=1}^{n} X_i/n$ 有顯著的

大於 p_0 則棄卻 H_0 虛無假設。因此，由型 I 誤差之定義可知

$$
\begin{aligned}
\alpha &= \mathrm{P}\left(\widehat{p} > c_1 | \, p = p_0\right) \\
&= \mathrm{P}\left(\frac{\widehat{p} - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} > \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right) \\
&= \mathrm{P}\left(Z > \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}}\right)
\end{aligned}
$$

可求得

$$
z_{\frac{\alpha}{2}} = \frac{c_1 - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} \text{ 或} c_1 = p_0 + z_{\frac{\alpha}{2}}\sqrt{\frac{p_0\left(1 - p_0\right)}{n}}
$$

所以，棄卻域為

$$
C = \left\{\widehat{p} > p_0 + z_{\frac{\alpha}{2}}\sqrt{\frac{p_0\left(1 - p_0\right)}{n}}\right\}
$$

例 **7–11**　某班學生平日之到課率為 95%，在國慶日之次日，全體 155 名學生中有 14 名缺課，問該日之缺課率是否較平日高？($\alpha = 0.05$)

解：

假設 p 為學生曠課比例，則統計假設為

$$
\begin{cases}
H_0 : p = 0.05 \\
H_1 : p > 0.05
\end{cases}
$$

棄卻域為

$$
C = \left\{\frac{\widehat{p} - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} > z_{0.05} = 1.645\right\}
$$

又 $\widehat{p} = 14/155 = 0.0903$，所以

$$
\frac{\widehat{p} - p_0}{\sqrt{p_0\left(1 - p_0\right)/n}} = \frac{0.0903 - 0.05}{\sqrt{0.05\left(1 - 0.05\right)/155}} = 2.3021
$$

落入棄卻域內，故棄卻 H_0，即該日之缺課較平日高。

例 **7–12**　補習班宣稱其學生上榜率為 20% 以上。今隨機抽取 8 位學生, 發現只有一位考上, 請以 0.05 之顯著水準檢定該補習班宣稱是否事實。若隨機抽取 8 位學生均落榜, 請以 0.05 之顯著水準檢定該補習班宣稱是否

事實。此檢定有何問題又該如何補救？

解：

$$\begin{cases} H_0 : p = 0.2 \\ H_1 : p < 0.2 \end{cases}$$

1.

$$\begin{aligned} \text{P-Value} &= \mathrm{P}\left(X \le 1 | p = 0.2\right) \\ &= \sum_{x=0}^{1} \binom{8}{x} 0.2^x 0.8^{8-x} = 0.50332 \end{aligned}$$

所以不棄卻 H_0

2.

$$\begin{aligned} \text{P-Value} &= \mathrm{P}\left(X \le 0 | p = 0.2\right) \\ &= \binom{8}{0} 0.2^0 0.8^{8-0} = 0.16777 \end{aligned}$$

所以不棄卻 H_0

3. 樣本數太少, 應增加樣本數。

例 **7–13** It is claimed that the failure rate of some printed circuit is less than 0.03. A new design has been implemented. A sample of size 100 circuits from the new design are inspected and there are 2 defective. Does this evidence support the claim？Give your approach to analyze the problem？

解：

$$\begin{cases} H_0 : p = 0.03 \\ H_1 : p < 0.03 \end{cases}$$

$$\begin{aligned} P - \text{Value} &= \mathrm{P}\left(\widehat{P} < 0.02\right) = \mathrm{P}\left(Z < \frac{0.02 - 0.03}{\sqrt{0.03 \times 0.97/100}}\right) \\ &= \mathrm{P}\left(Z < -0.58621\right) \\ &= 0.2789 \end{aligned}$$

所以當顯著水準 $\alpha < 0.2789$ 時不棄卻 H_0, 反之當顯著水準 $\alpha \ge 0.2789$ 時棄卻 H_0。

7.4.2 兩母體比例之檢定

假設 $X_1, X_2, ..., X_n; Y_1, Y_2, ..., Y_n$ 為分別抽自 $\mathrm{Bernoulli}(p_x)$ 與 $\mathrm{Bernoulli}(p_y)$ 的兩組樣本數為 n_1 與 n_2 的獨立隨機樣本。令 $\widehat{p}_x = \sum_{i=1}^{n} X_i/n_1$，與 $\widehat{p}_y = \sum_{i=1}^{n} Y_i/n_2$，由 **7.2.9** 可知，

$$\frac{\widehat{p}_x - \widehat{p}_y - (p_x - p_y)}{\sqrt{p_x(1-p_x)/n_1 + p_y(1-p_y)/n_2}} \backsim \mathcal{N}(0,1)$$

所以，母體比例之檢定方法介紹如下;

兩母體比例 $p_x - p_y$ 之雙尾檢定

$$\begin{cases} H_0 : p_x - p_y = k \\ H_1 : p_x - p_y \neq k \end{cases}$$

由於統計假設為要檢定 $p_x - p_y$ 是否等於 k，因此若經抽樣所得的樣本比例 $\widehat{p}_x - \widehat{p_y}$ 與 k 相差甚大則棄卻 H_0 虛無假設。由 **7.2.9** 可知在 H_0 為真的情況下, $\widehat{p}_x - \widehat{p_y}$ 近似分配為

$$\lim_{n\to\infty} \frac{\widehat{p}_x - \widehat{p}_y - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}} \xrightarrow{d} \mathcal{N}(0,1)$$

因此，由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\widehat{p}_x - \widehat{p}_y < c_1 \text{ 或} \widehat{p}_x - \widehat{p}_y > c_2 \middle| p_x - p_y = k, k \neq 0\right) \\
&= \mathrm{P}\left(\frac{\widehat{p}_x - \widehat{p}_y - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}} < \frac{c_1 - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}}\right) \\
&\quad + \mathrm{P}\left(\frac{\widehat{p}_x - \widehat{p}_y - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}} > \frac{c_2 - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}}\right) \\
&= \mathrm{P}\left(Z < \frac{c_1 - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}}\right) + \mathrm{P}\left(Z > \frac{c_2 - k}{\sqrt{\widehat{p}_x(1-\widehat{p}_x)/n_1 + \widehat{p}_y(1-\widehat{p}_y)/n_2}}\right)
\end{aligned}$$

由於為雙尾檢定，故取兩邊機率相等，可求得

$$-z_{\frac{\alpha}{2}} = \frac{c_1 - k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1 + \widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}} \text{ 及} z_{\frac{\alpha}{2}} = \frac{c_2 - k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1 + \widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}}$$

或

$$c_1 = k - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n_1} + \frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}} \text{ 及} c_2 = k + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n} + \frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}}$$

所以，棄卻域可寫成

$$C = \left\{\widehat{p}_x - \widehat{p}_y < k - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n_1} + \frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}} \text{ 或} \widehat{p}_x - \widehat{p}_y > k + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n_1} + \frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}}\right\}$$

當 $k = 0$ 時，即欲檢定兩母體比例是否相等，由型 I 誤差之定義可知 $p_x = p_y = p$，因此必須先求得母體比例。由於參數 p 未知，在 H_0 為眞的情況下，由於

$$\sum_{i=1}^{n_1} X_i \sim B\left(n_1, p\right) \text{ 及} \sum_{i=1}^{n_2} Y_i \sim B\left(n_2, p\right)$$

故可由最大概似法求得 p 之 MLE 為

$$\overline{p} = \frac{\sum_{i=1}^{n_1} X_i + \sum_{i=1}^{n_2} Y_i}{n_1 + n_2}$$

又 MLE 具一致性，所以

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{\sum_{i=1}^{n_1} X_i + \sum_{i=1}^{n_2} Y_i}{n_1 + n_2} - p\right| < \varepsilon\right) = 1$$

且由中央極限定理可知

$$\lim_{n\to\infty} \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\mathrm{P}\left(1-p\right)/n_1 + \mathrm{P}\left(1-p\right)/n_2}} \sim \mathcal{N}\left(0,1\right)$$

因此由 Sultsky's Theorem 可得

$$\lim_{n\to\infty} \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\overline{p}\left(1-\overline{p}\right)/n_1 + \overline{p}\left(1-\overline{p}\right)/n_2}} \xrightarrow{d} \mathcal{N}\left(0,1\right)$$

因統計假設為檢定 $p_x - p_y$ 是否等於 0，因此若經抽樣所得的樣本比例 $\widehat{p}_x - \widehat{p}_y$ 與 0 相差甚大則棄卻 H_0 虛無假設。故由型 I 誤差之定義可知

$$
\begin{aligned}
\alpha &= \mathrm{P}\left(\widehat{p}_x - \widehat{p}_y < c_1 \text{ 或} \widehat{p}_x - \widehat{p}_y > c_2 \middle| p_x - p_y = 0\right) \\
&= \mathrm{P}\left(\frac{\widehat{p}_x - \widehat{p}_y}{\sqrt{\overline{p}\left(1-\overline{p}\right)/n_1 + \overline{p}\left(1-\overline{p}\right)/n_2}} < \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)/n_1 + \overline{p}\left(1-\overline{p}\right)/n_2}}\right) \\
&\quad + \mathrm{P}\left(\frac{\widehat{p}_x - \widehat{p}_y}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1 + 1/n_2}} > \frac{c_2}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1 + 1/n_2}}\right) \\
&= \mathrm{P}\left(Z < \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1 + 1/n_2}}\right) + \mathrm{P}\left(Z > \frac{c_2}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1 + 1/n_2}}\right)
\end{aligned}
$$

由於為雙尾檢定，故取兩邊機率相等，可求得

$$
-z_{\frac{\alpha}{2}} = \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1 + 1/n_2}} \text{ 及} z_{\frac{\alpha}{2}} = \frac{c_2}{\sqrt{\overline{p}\left(1-\overline{p}\right)\sqrt{1/n_1 + 1/n_2}}}
$$

或

$$
c_1 = -z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}} \text{ 及} c_2 = z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}
$$

所以，棄卻域可寫成

$$
C = \left\{\widehat{p}_x - \widehat{p}_y < -z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}} \text{ 或} \widehat{p}_x - \widehat{p}_y > z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1} + \frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}\right\}
$$

兩母體比例 $p_x - p_y$ 之右尾檢定

$$
\begin{cases}
H_0 : p_x - p_y = k \\
H_1 : p_x - p_y > k
\end{cases}
$$

由於統計假設為要檢定 $p_x - p_y$ 是否大於 k，因此若經抽樣所得的樣本比例 $\widehat{p}_x - \widehat{p}_y$ 有顯著大於 k 則棄卻 H_0 虛無假設。因此，由型 I 誤差之定義可知

$$
\alpha = \mathrm{P}\left(\widehat{p}_x - \widehat{p}_y > c_1 \middle| p_x - p_y = k, k \neq 0\right)
$$

$$
\begin{aligned}
&= \mathrm{P}\left(\frac{\widehat{p}_x-\widehat{p}_y-k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1+\widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}} > \frac{c_1-k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1+\widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}}\right) \\
&= \mathrm{P}\left(Z > \frac{c_1-k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1+\widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}}\right)
\end{aligned}
$$

可求得

$$
z_\alpha = \frac{c_1-k}{\sqrt{\widehat{p}_x\left(1-\widehat{p}_x\right)/n_1+\widehat{p}_y\left(1-\widehat{p}_y\right)/n_2}} \text{ 或} c_1 = k + z_\alpha\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n_1}+\frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}}
$$

所以，棄卻域為

$$
C = \left\{\widehat{p}_x-\widehat{p}_y > k + z_\alpha\sqrt{\frac{\widehat{p}_x\left(1-\widehat{p}_x\right)}{n_1}+\frac{\widehat{p}_y\left(1-\widehat{p}_y\right)}{n_2}}\right\}
$$

當 $k=0$ 時，則由型 I 誤差之定義可知

$$
\begin{aligned}
\alpha &= \mathrm{P}\left(\widehat{p}_x-\widehat{p}_y > c_1 \middle| p_x-p_y=0\right) \\
&= \mathrm{P}\left(\frac{\widehat{p}_x-\widehat{p}_y}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1+1/n_2}} > \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1+1/n_2}}\right) \\
&= \mathrm{P}\left(Z > \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1+1/n_2}}\right)
\end{aligned}
$$

故推得

$$
z_\alpha = \frac{c_1}{\sqrt{\overline{p}\left(1-\overline{p}\right)}\sqrt{1/n_1+1/n_2}} \text{ 或} c_1 = z_\alpha\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1}+\frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}
$$

因此，棄卻域為

$$
C = \left\{\widehat{p}_x-\widehat{p}_y > z_\alpha\sqrt{\frac{\overline{p}\left(1-\overline{p}\right)}{n_1}+\frac{\overline{p}\left(1-\overline{p}\right)}{n_2}}\right\}
$$

例 **7–14** 甲, 乙二名工人所製之產品檢查如下表, 問甲, 乙二名工人之技術是否有差異？($\alpha=0.05$)

	良品	不良品	合計
甲	480	120	600
乙	205	45	250

解：

令 $\widehat{p}_x, \widehat{p}_y$ 分別甲乙兩工人所製知不良品，則統計假設為

$$\begin{cases} H_0 : p_x - p_y = 0 \\ H_1 : p_x - p_y \neq 0 \end{cases}$$

又

$$\widehat{p}_x = \frac{120}{600} = 0.2, \widehat{p}_y = \frac{45}{250} = 0.18, \overline{p} = \frac{120+45}{600+250} = 0.1941$$

所以棄卻域為

$$C = \left\{ \left| \frac{\widehat{p}_x - \widehat{p}_y}{\sqrt{\overline{p}\,(1-\overline{p})\,(1/n_1 + 1/n_2)}} \right| > z_{0.025} = 1.96 \right\}$$

又

$$\left| \frac{\widehat{p}_x - \widehat{p}_y}{\sqrt{\overline{p}\,(1-\overline{p})\,(1/n_1 + 1/n_2)}} \right| = \left| \frac{0.2 - 0.18}{\sqrt{0.1941 \times (1 - 0.1941) \times (1/600 + 1/250)}} \right| = 0.6718$$

並未落入棄卻域內，故不棄卻 H_0，即無充分證據顯示甲, 乙二名工人之技術有顯著性差異。

第 8 章　變異數分析

統計資料常受多種因素的影響, 而使各個體的某種特徵發生變異, 實驗者經常會要比較超過兩個處理的問題，如：多個不同品種小麥的平均產量是否相同，多種不同廠牌的汽車平均每加侖可以行駛的理成數是否不同，而對這種影響因素所造成之變異的觀察與驗證的統計方法, 稱為變異數分析。

8.1　名詞解釋及基本假設

8.1.1　名詞解釋

1. **因子**：調查中研究某一獨立變數或為引起資料發生變動的原因。例如：價格對於銷售量之影響, 價格即為一因子。同樣地，研究四種不同電視節目對觀眾吸引力之影響，其中電視節目即為一因子。
2. **因子水準**：因子之特殊形式 (所包含的數目) 或為表示因子狀態之條件。例如上例中之價格 (X =50 元, 60 元, 70 元) 共有三個水準。
3. **一因子分類**：研究之對象只包含一個獨立變數或事項者。意即觀察值以一個標準為分類基礎稱之。
4. **多因子分類**：研究之對象只包含兩個或以上獨立變數或事項者。意即觀察值以多個標準為分類基礎稱之。
5. **屬質因子**：因子水準以屬質之方式來表示稱之，如廣告型態 (電視、報紙及雜誌廣告)。
6. **屬量因子**：因子水準可以用計量表示者，如價格 (X =50 元, 60 元, 70 元) , 溫度反應 ($X = 50°$, $60°$, $70°$)

8.1.2 基本假設

1. 每一因子水準所對應之機率分配皆服從常態分配。
2. 所有樣本都是隨機抽取而得，且彼此獨立。
3. 各常態母體之變異數皆相等。

8.2 一因子變異數分析

一因子變異數分析係研究之對象只包含一個獨立變數，而觀察此因素之不同對研究對象的影響是否有顯著差異，其數學模式如下：

$$X_{ij} = \mu_i + \varepsilon_{ij},\ \varepsilon_{ij} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right),$$

其中 X_{ij} 表第 i 個因子水準之第 j 的樣本觀測值，μ_i 表第 i 個母體之平均數，ε_{ij} 表隨機誤差項。假設實驗者由 k 個獨立常態母體中抽出 k 組隨機樣本，資料如下：

母體	U_1	U_2	$\cdots$	U_i	$\cdots$	U_k	
樣本	X_{11}	X_{21}	$\cdots$	X_{i1}	$\cdots$	X_{k1}	
	X_{12}	X_{22}	$\cdots$	X_{i2}		X_{k2}	
	$\vdots$	$\vdots$		$\vdots$		$\vdots$	
	$\vdots$	$\vdots$		$\vdots$		$\vdots$	
	X_{1n_1}	X_{2n_2}	$\cdots$	X_{in_i}	$\cdots$	X_{kn_k}	
總合	$T_{1.}$	$T_{2.}$	$\cdots$	$T_{i.}$	$\cdots$	$T_{k.}$	$T_{..}$
平均	$\overline{X}_{1.}$	$\overline{X}_{2.}$	$\cdots$	$\overline{X}_{i.}$	$\cdots$	$\overline{X}_{k.}$	$\overline{X}_{..}$

首先將 $N(= n_1 + n_2 + \cdots + n_k)$ 個樣本所求得之總變異分割得

$$\begin{aligned}
\text{SSTO} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{..}\right)^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.} + \overline{X}_{i.} - \overline{X}_{..}\right)^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.}\right)^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.}\right)\left(\overline{X}_{i.} - \overline{X}_{..}\right) + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{i.} - \overline{X}_{..}\right)^2,
\end{aligned}$$

其中

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{i.}\right)\left(\overline{X}_{i.}-\overline{X}_{..}\right) &= \sum_{i=1}^{k}\left[\left(\overline{X}_{i.}-\overline{X}_{..}\right)\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{i.}\right)\right] \\
&= \sum_{i=1}^{k}\left[\left(\overline{X}_{i.}-\overline{X}_{..}\right)\times\left(\sum_{j=1}^{n_i}X_{ij}-n_i\overline{X}_{i.}\right)\right] \\
&= \sum_{i=1}^{k}\left[\left(\overline{X}_{i.}-\overline{X}_{..}\right)\times 0\right] \\
&= 0
\end{aligned}$$

所以,

$$\begin{aligned}
\mathrm{SSTO} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{i.}+\overline{X}_{i.}-\overline{X}_{..}\right)^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{i.}\right)^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2 \\
&= \mathrm{SSE}+\mathrm{SSB}
\end{aligned}$$

由於上述的計算太過繁雜，總變異 (SSTO)，組間變異 (SSB) 及組內變異 (SSE) 可簡化成

$$\begin{aligned}
\mathrm{SSTO} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2+\sum_{i=1}^{k}\sum_{j=1}^{n_i}\overline{X}_{..}^2-2\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}\overline{X}_{..} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2+N\overline{X}_{..}^2-2\overline{X}_{..}\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2+N\overline{X}_{..}^2-2\overline{X}_{..}N\overline{X}_{..} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-N\overline{X}_{..}^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-N\left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}}{N}\right]^2 \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i}X_{ij}^2-\frac{T_{..}^2}{N},
\end{aligned}$$

$$\mathrm{SSB} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2 = \sum_{i=1}^{k}n_i\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2$$

$$\begin{aligned}
&= \sum_{i=1}^{k} n_i\left(\overline{X}_{i.}^2+\overline{X}_{..}^2-2\overline{X}_{i.}\overline{X}_{..}\right)=\sum_{i=1}^{k} n_i\overline{X}_{i.}^2+\sum_{i=1}^{k} n_i\overline{X}_{..}^2-2\sum_{i=1}^{k} n_i\overline{X}_{i.}\overline{X}_{..} \\
&= \sum_{i=1}^{k} n_i\overline{X}_{i.}^2-N\left[\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}}{N}\right]^2 \\
&= \sum_{i=1}^{k}\frac{T_{i.}^2}{n_i}-\frac{T_{..}^2}{N}
\end{aligned}$$

及

$$\begin{aligned}
\mathrm{SSE} &= \mathrm{SSTO}-\mathrm{SSB} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2-\frac{T_{..}^2}{N}-\sum_{i=1}^{k}\frac{T_{i.}^2}{n_i}+\frac{T_{..}^2}{N} \\
&= \sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2-\sum_{i=1}^{k}\frac{T_{i.}^2}{n_i} \\
&= \sum_{i=1}^{k}\left(n_i-1\right)S_i^2
\end{aligned}$$

又 $\frac{(n_i-1)S_i^2}{\sigma^2}\sim\chi^2\left(n_i-1\right)$ 由卡方之可加性可輕易推得

$$\frac{\sum_{i=1}^{k}\left(n_i-1\right)S_i^2}{\sigma^2}\sim\chi^2\left(N-k\right)$$

及

$$\begin{aligned}
\mathrm{E}\left(\mathrm{SSE}\right) &= \mathrm{E}\left(\sum_{i=1}^{k}\left(n_i-1\right)S_i^2\right)=\sum_{i=1}^{k}\left(n_i-1\right)\mathrm{E}\left(S_i^2\right) \\
&= \sum_{i=1}^{k}\left(n_i-1\right)\sigma^2 \\
&= \left(N-k\right)\sigma^2
\end{aligned}$$

因此，

$$\mathrm{E}\left(\widehat{\sigma}^2\right)=\mathrm{E}\left(\mathrm{MSE}\right)=\mathrm{E}\left(\frac{\mathrm{SSE}}{N-k}\right)=\sigma^2$$

所以 MSE 可做為母體變異數 σ^2 之估計式。又

$$\begin{aligned}
\mathrm{E}\left[\mathrm{SSB}\right] &= \mathrm{E}\left[\sum_{i=1}^{k} n_i \overline{X}_{i.}^2 - N\overline{X}_{..}^2\right] \\
&= \sum_{i=1}^{k} n_i \mathrm{E}\left[\overline{X}_{i.}^2\right] - N \times \mathrm{E}\left[\overline{X}_{..}^2\right] \\
&= \sum_{i=1}^{k} n_i \left[\mu_i^2 + \frac{\sigma^2}{n_i}\right] - N \times \left[\left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 + \frac{\sigma^2}{N}\right] \\
&= \sum_{i=1}^{k} n_i\mu_i^2 + \sum_{i=1}^{k} n_i \frac{\sigma^2}{n_i} - N \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 - \sigma^2 \\
&= (k-1)\,\sigma^2 + \sum_{i=1}^{k} n_i\mu_i^2 - N \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2
\end{aligned}$$

其中

$$\begin{aligned}
&\sum_{i=1}^{k} n_i\mu_i^2 - N \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 \\
= &\sum_{i=1}^{k}\sum_{j=1}^{n_i} \mu_i^2 - N \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 \\
= &\sum_{i=1}^{k}\sum_{j=1}^{n_i} \mu_i^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 - 2N \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right) \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right) \\
= &\sum_{i=1}^{k}\sum_{j=1}^{n_i} \mu_i^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 - 2\sum_{i=1}^{k} \left[n_i\mu_i \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)\right] \\
= &\sum_{i=1}^{k}\sum_{j=1}^{n_i} \mu_i^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2 - 2\sum_{i=1}^{k}\sum_{j=1}^{n_i} \left[\mu_i \times \left(\frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)\right] \\
= &\sum_{i=1}^{k}\sum_{j=1}^{n_i} \left[\mu_i - \frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right]^2 \\
= &\sum_{i=1}^{k} \left[n_i \times \left(\mu_i - \frac{\sum_{i=1}^{k} n_i\mu_i}{N}\right)^2\right]
\end{aligned}$$

所以，

$$\mathrm{E}\left(\mathrm{SSB}\right) = (k-1)\,\sigma^2 + \sum_{i=1}^{k}\left[n_i \times (\mu_i - \overline{\mu})^2\right] \geq (k-1)\sigma^2,\ \text{其中}\,\overline{\mu} = \frac{\sum_{i=1}^{k} n_i\mu_i}{N}$$

因此當 $\mu_1 = \mu_2 = \cdots = \mu_k = \mu$，所以

$$\mathrm{E}\,(\mathrm{MSB}) = \mathrm{E}\left(\frac{\mathrm{SSB}}{k-1}\right) = \sigma^2$$

否則

$$\mathrm{E}\,(\mathrm{MSB}) \geq \sigma^2$$

若統計假設為

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k = \mu \\ H_1 : \mu_i \text{ 不全相等} \end{cases}$$

當 H_0 為真時，由定理6.3可知 $\mathrm{SSB}/\sigma^2 \sim \chi^2(k-1)$ 及 $\mathrm{SSE}/\sigma^2 \sim \chi^2(N-k)$，則

$$F = \frac{\mathrm{MSB}}{\mathrm{MSE}} \sim \mathcal{F}\,(k-1, N-k)$$

且 $F = \mathrm{MSB}/\mathrm{MSE}$ 必接近 1, 反之若 μ_i 不全相等，則 $F = \mathrm{MSB}/\mathrm{MSE}$ 必遠大於 1，故檢定之棄卻域為

$$C = \left\{F = \frac{\mathrm{MSB}}{\mathrm{MSE}} > F_\alpha\,(k-1, N-k)\right\}$$

例 **8–1** Montgomery 認為人造纖維的強度可能會受纖維中含棉的比率所影響。在考慮的五種比率值中每一種皆取五個觀測值如下, 在顯著水準下, 利用 F 檢定來看由於使用含棉的比率不同而拉張強度是否有所差異。

含棉比率	拉張強度 (磅/平方吋)				
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	22

解：

$$\mathrm{SSTO} = \sum_{i=1}^{5}\sum_{j=1}^{5} X_{ij}^2 - \frac{\left[\sum_{i=1}^{5}\sum_{j=1}^{5} X_{ij}\right]^2}{N} = 6292 - \frac{376^2}{25} = 636.96$$

及

$$\mathrm{SSB} \quad = \quad \sum_{i=1}^{5} \frac{T_{i.}^2}{n_i} - \frac{\left[\sum_{i=1}^{5}\sum_{j=1}^{5} X_{ij}\right]^2}{N}$$

$$
\begin{aligned}
&= \left(\frac{49^2}{5}+\frac{77^2}{5}+\frac{88^2}{5}+\frac{108^2}{5}+\frac{54^2}{5}\right)-\frac{376^2}{25} \\
&= 475.76
\end{aligned}
$$

所以

$$\text{SSE} = \text{SSTO} - \text{SSB} = 636.96 - 475.76 = 161.2$$

ANOVA Table

變異來源	平方和	自由度	均方和	F 值
組間	475.76	4	118.94	14.757
組內	161.2	20	8.06	
總和	636.96	24		

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \\ H_1 : \mu_i \text{ 不全相等} \end{cases}$$

$$C = \left\{ F = \frac{\text{MSB}}{\text{MSE}} > F_{0.05}\left(4, 20\right) = 2.8661 \right\}$$

因為 $F = 14.757 > F_{0.05}(4, 20) = 2.8661 \in C$，所以棄卻 $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

例 **8–2** 一個研究把國內 134 位管理者依據一次面談與觀察的結果, 劃分分為 4 個群體, 有 62 位管理者在 A 群, 有 52 位管理者在 B 群, 有 7 位管理者在 C 群, 有 13 位管理者在 D 群, 並把各管理群的月薪加以比較（以萬元計)。A 群管理者之平均月薪為 7.87, B 群管理者平均薪資為 7.47, C 群管理者平均月薪為 5.14, D 群管理者平均月薪為 3.69, 假設總變異已知為 596.01。請建立 ANOVA TABLE 並檢定 4 個群體知管理者的母體平均數是否相等？

解：

ANOVA Table

變異來源	平方和	自由度	均方和	F 值
組間 B	221.34	3	73.780	25.6
組內	374.67	130	2.882	
總和	596.01	133		

$$\begin{cases} H_0 : \mu_A = \mu_B = \mu_C = \mu_D \\ H_1 : \mu_i \text{ 不全相等} \end{cases}$$

$$C = \left\{ F = \frac{\text{MSB}}{\text{MSE}} > F_{0.05}\left(3, 130\right) = 2.6049 \right\}$$

因為 $F = 25.6 > F_{0.05}(3, 130) \in C$，所以棄卻 $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$

8.3 二因子變異數分析

在一因子變異數分析中，因將其他條件固定，而只針對某一因子來解析，依其因子水準不同做隨機化之實驗配置，故其實驗結果比較狹窄，而同時對影響大之原因隨機化，故會使誤差變異加大，導致差異而使得檢定及推定的效率降低。因此要使得檢定效率提高，則必須探討多因子分析，其數學模式如下：

$$X_{ij} = \mu_{ij} + \varepsilon_{ij},\ \varepsilon_{ij} \sim \mathcal{N}\left(0, \sigma^2\right),$$

其中 X_{ij} 表列因子第 i 個因子水準之第 j 的樣本觀測值，$\mu_{i.}$ 表列因子第 i 個母體之平均數，$\mu_{.j}$ 表行因子第 j 個母體之平均數，ε_{ij} 表隨機誤差項, 資料如下：

		行因子						總和	平均
		1	2	$\cdots$	i	$\cdots$	R		
列因子	1	X_{11}	X_{21}	$\cdots$	X_{i1}	$\cdots$	X_{R1}	$T_{1.}$	$\overline{X}_{1.}$
	2	X_{12}	X_{22}	$\cdots$	X_{i2}		X_{R2}	$T_{2.}$	$\overline{X}_{2.}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
	C	X_{1C}	X_{2C}	$\cdots$	X_{iC}	$\cdots$	X_{RC}	$T_{R.}$	$\overline{X}_{R.}$
總和		$T_{.1}$	$T_{.2}$	$\cdots$	$T_{.i}$	$\cdots$	$T_{C.}$	$T_{..}$	
平均		$\overline{X}_{.1}$	$\overline{X}_{.2}$	$\cdots$	$\overline{X}_{.i}$	$\cdots$	$\overline{X}_{C.}$		$\overline{X}_{..}$

可將總變異分解如下：

$$\begin{aligned}
\text{SSTO} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij} - \overline{X}_{..}\right)^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\left[\left(\overline{X}_{i.} - \overline{X}_{..}\right) + \left(\overline{X}_{.j} - \overline{X}_{..}\right) + \left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..}\right)\right]^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.} - \overline{X}_{..}\right)^2 + \sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j} - \overline{X}_{..}\right)^2 + \sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..}\right)^2 \\
&\quad + 2\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.} - \overline{X}_{..}\right)\left(\overline{X}_{.j} - \overline{X}_{..}\right) + 2\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.} - \overline{X}_{..}\right)\left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..}\right) \\
&\quad + 2\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j} - \overline{X}_{..}\right)\left(X_{ij} - \overline{X}_{i.} - \overline{X}_{.j} + \overline{X}_{..}\right),
\end{aligned}$$

其中

$$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)\left(\overline{X}_{.j}-\overline{X}_{..}\right) = \sum_{i=1}^{r}\left[\left(\overline{X}_{i.}-\overline{X}_{..}\right)\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)\right]=0,$$

$$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right) = \sum_{i=1}^{r}\left[\left(\overline{X}_{i.}-\overline{X}_{..}\right)\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right)\right]=0$$

及

$$\begin{aligned}\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right) &= \sum_{j=1}^{c}\sum_{i=1}^{r}\left(\overline{X}_{.j}-\overline{X}_{..}\right)\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right)\\ &= \sum_{j=1}^{c}\left[\left(\overline{X}_{.j}-\overline{X}_{..}\right)\sum_{i=1}^{r}\left(X_{ij}-\overline{X}_{.j}-\overline{X}_{i.}+\overline{X}_{..}\right)\right]\\ &= 0\end{aligned}$$

所以總變異可寫成

$$\begin{aligned}\text{SSTO} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2+\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2+\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right)^2\\ &= \text{SSR}+\text{SSC}+\text{SSE}\end{aligned}$$

由於上述的計算太過繁雜, SSTO, SSR, SSC 及 SSE 可簡化成

$$\begin{aligned}\text{SSTO} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{..}\right)^2\\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}^2+\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{..}^2-2\sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}\overline{X}_{..}\\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}^2-\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{..}^2\\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}^2-\sum_{i=1}^{r}\sum_{j=1}^{c}\left[\frac{T_{..}}{rc}\right]^2\\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}^2-\frac{T_{..}^2}{rc},\end{aligned}$$

$$
\begin{aligned}
\mathrm{SSR} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{i.}^2+\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{..}^2-2\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{i.}\overline{X}_{..} \\
&= \sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{i.}^2-\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{..}^2 \\
&= \sum_{i=1}^{r}c\overline{X}_{i.}^2-\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{..}^2 \\
&= \sum_{i=1}^{r}c\left[\frac{\sum_{j=1}^{c}X_{ij}}{c}\right]^2-\frac{T_{..}^2}{rc} \\
&= \sum_{i=1}^{r}\frac{T_{i.}^2}{c}-\frac{T_{..}^2}{rc}
\end{aligned}
$$

及

$$
\begin{aligned}
\mathrm{SSC} &= \sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2=\sum_{j=1}^{c}\sum_{i=1}^{r}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2 \\
&= \sum_{j=1}^{c}\sum_{i=1}^{r}\overline{X}_{.j}^2+\sum_{j=1}^{c}\sum_{i=1}^{r}\overline{X}_{..}^2-2\sum_{j=1}^{c}\sum_{i=1}^{r}\overline{X}_{.j}\overline{X}_{..} \\
&= \sum_{j=1}^{c}\sum_{i=1}^{r}\overline{X}_{.j}^2-\sum_{j=1}^{c}\sum_{i=1}^{r}\overline{X}_{..}^2 \\
&= \sum_{j=1}^{c}\frac{T_{.j}^2}{r}-\frac{T_{..}^2}{rc}
\end{aligned}
$$

又

$$
\begin{aligned}
\mathrm{E}\left(\mathrm{SSE}\right) &= \mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right)^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left[\left(X_{ij}-\overline{X}_{i.}\right)-\left(\overline{X}_{.j}-\overline{X}_{..}\right)\right]^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)^2+\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2-2\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)\left(\overline{X}_{.j}-\overline{X}_{..}\right)\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)^2\right]+\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2\right]
\end{aligned}
$$

$$-2\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)\left(\overline{X}_{.j}-\overline{X}_{..}\right)\right]$$

其中

$$\begin{aligned}\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)^{2}\right] &= \sum_{i=1}^{r}\mathrm{E}\left[\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)^{2}\right]\\ &= \sum_{i=1}^{r}(c-1)\,\sigma^{2}\\ &= r\,(c-1)\,\sigma^{2},\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^{2}\right] &= \sum_{i=1}^{r}\mathrm{E}\left[\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^{2}\right]\\ &= \sum_{i=1}^{r}\frac{(c-1)\,\sigma^{2}}{r}\\ &= (c-1)\,\sigma^{2},\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)\left(\overline{X}_{.j}-\overline{X}_{..}\right) &= \sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}\overline{X}_{.j}-\sum_{i=1}^{r}\sum_{j=1}^{c}X_{ij}\overline{X}_{..}-\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{i.}\overline{X}_{.j}+\sum_{i=1}^{r}\sum_{j=1}^{c}\overline{X}_{i.}\overline{X}_{..}\\ &= \sum_{j=1}^{c}\sum_{i=1}^{r}X_{ij}\overline{X}_{.j}-rc\times\overline{X}_{..}^{2}-rc\times\overline{X}_{..}^{2}+rc\times\overline{X}_{..}^{2}\\ &= \sum_{j=1}^{c}r\overline{X}_{.j}^{2}-rc\times\overline{X}_{..}^{2},\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left[\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}\right)\left(\overline{X}_{.j}-\overline{X}_{..}\right)\right] &= \mathrm{E}\left[\sum_{j=1}^{c}r\overline{X}_{.j}^{2}-rc\times\overline{X}_{..}^{2}\right]\\ &= r\sum_{j=1}^{c}\mathrm{E}\left[\overline{X}_{.j}^{2}\right]-rc\times\frac{\sigma^{2}}{rc}\\ &= r\times\sum_{j=1}^{c}\frac{\sigma^{2}}{r}-\sigma^{2}\\ &= (c-1)\,\sigma^{2},\end{aligned}$$

所以可求得

$$
\begin{aligned}
\mathrm{E}(\mathrm{SSE}) &= r(c-1)\sigma^2+(c-1)\sigma^2-2(c-1)\sigma^2 \\
&= r(c-1)\sigma^2-(c-1)\sigma^2 \\
&= (rc-r-c+1)\sigma^2 \\
&= (r-1)(c-1)\sigma^2
\end{aligned}
$$

故可推得

$$
\mathrm{E}\left(\widehat{\sigma}^2\right)=\mathrm{E}(\mathrm{MSE})=\mathrm{E}\left(\frac{\mathrm{SSE}}{(r-1)(c-1)}\right)=\sigma^2
$$

所以 MSE 可做為母體變異數 σ^2 之估計式。又當 $\mu_{1.}=\mu_{2.}=\cdots=\mu_{r.}$ 及 $\mu_{.1}=\mu_{.2}=\cdots=\mu_{.c}$ 時，可推得

$$
\begin{aligned}
\frac{\mathrm{SSR}}{\sigma^2} &= \frac{\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2}{\sigma^2} \\
&= \sum_{i=1}^{r}\frac{c\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2}{\sigma^2} \\
&= \sum_{i=1}^{r}\frac{\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2}{\sigma^2/c}\sim\chi^2(r-1)
\end{aligned}
$$

及

$$
\begin{aligned}
\frac{\mathrm{SSC}}{\sigma^2} &= \frac{\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2}{\sigma^2} \\
&= \sum_{j=1}^{c}\frac{r\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2}{\sigma^2} \\
&= \sum_{j=1}^{c}\frac{\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2}{\sigma^2/r}\sim\chi^2(c-1)
\end{aligned}
$$

由於

$$
\frac{\mathrm{SSTO}}{\sigma^2}=\frac{\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{..}\right)^2}{\sigma^2}\sim\chi^2(rc-1)
$$

故可由卡方之可加性推得

$$\frac{\text{SSE}}{\sigma^2} \sim \chi^2\left((r-1)(c-1)\right)$$

因此，可將 ANOVA Table 編制如下：

ANOVA Table

變異來源	均方和	自由度	平方和	F 值
SSR	$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{i.}-\overline{X}_{..}\right)^2$	$r-1$	$\frac{\text{SSR}}{r-1}$	$F_a=\frac{\text{MSR}}{\text{MSE}}$
SSC	$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\overline{X}_{.j}-\overline{X}_{..}\right)^2$	$c-1$	$\frac{\text{SSC}}{c-1}$	$F_b=\frac{\text{MSC}}{\text{MSE}}$
SSE	$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{i.}-\overline{X}_{.j}+\overline{X}_{..}\right)^2$	$(r-1)(c-1)$	$\frac{\text{SSE}}{(r-1)(c-1)}$	
SSTO	$\sum_{i=1}^{r}\sum_{j=1}^{c}\left(X_{ij}-\overline{X}_{..}\right)^2$	$rc-1$		

利用 ANOVA Table 檢定下列假設

1.

$$\begin{cases} H_0: \mu_{1.}=\mu_{2.}=\cdots=\mu_{r.} \\ H_1: \mu_{1.},\mu_{2.},\cdots,\mu_{r.}\text{不全相等} \end{cases}$$

棄卻域為

$$C=\left\{F_a=\frac{\text{MSR}}{\text{MSE}}>F_\alpha\left(r-1,(r-1)(c-1)\right)\right\}$$

2.

$$\begin{cases} H_0: \mu_{.1}=\mu_{.2}=\cdots=\mu_{.c} \\ H_1: \mu_{.1},\mu_{.2},\cdots,\mu_{.c}\text{不全相等} \end{cases}$$

棄卻域為

$$C=\left\{F_b=\frac{\text{MSC}}{\text{MSE}}>F_\alpha\left(c-1,(r-1)(c-1)\right)\right\}$$

例 **8–3** 設有甲乙丙三種不同品種的稻米, 分別使用 A, B, C, D 四種不同的肥料。今隨機選擇面積相等條件相同的 12 塊田地做實驗, 得到其收穫量 (以千公斤計) 如下表：

		品種		
		甲	乙	丙
	A	8	3	7
肥	B	10	4	8
料	C	6	5	6
	D	8	4	7

取顯著水準 $\alpha = 0.05$，試分別檢定

1. 不同肥料
2. 不同品種

所得到的平均收穫量是否會有顯著差異？

解：

ANOVA Table

變異來源	平方和	自由度	均方和	F 值
列因子	4.6667	3	1.5556	1.2727
行因子	34.6667	2	17.3333	14.1818
殘差	7.3333	6	1.2222	
總和	46.6667	11		

1.

$$\begin{cases} H_0 : \mu_A = \mu_B = \mu_C = \mu_D \\ H_1 : \mu_A, \mu_B, \mu_C, \mu_D \text{ 不全相等} \end{cases}$$

$$C = \left\{ F = \frac{\text{MSC}}{\text{MSE}} > F_{0.05}(2,6) = 5.1432 \right\}$$

因為 $F = 14.1818 > F_{0.05}(2,6) = 5.1432 \in C$，所以棄卻 $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$

2.

$$\begin{cases} H_0 : \mu_{\text{甲}} = \mu_{\text{乙}} = \mu_{\text{丙}} \\ H_1 : \mu_{\text{甲}}, \mu_{\text{乙}}, \mu_{\text{丙}} \text{ 不全相等} \end{cases}$$

$$C = \left\{ F = \frac{\text{MSR}}{\text{MSE}} > F_{0.05}(3,6) = 4.7571 \right\}$$

因為 $F = 1.2727 < F_{0.05}(3,6) = 4.7571 \notin C$，所以不棄卻 $H_0 : \mu_{\text{甲}} = \mu_{\text{乙}} = \mu_{\text{丙}}$

第 9 章 簡單線性迴歸分析

迴歸分析是一種統計方法, 主要是利用一組獨立變數對某一應變數做預測。本章僅探討簡單線性迴歸模式，即利用一個獨立變數 (X) 對應變數 (Y) 做預測之模式。所謂模式, 乃指找出 X 與 Y 之間的函數關係，即 $Y = f(X)$。然而由於 Y 為一隨機變數，所以我們並不能正確的預測 Y 值是多少。因此我們僅能估計當 $X = x$ 時，隨機變數 Y 之期望值為何，其中實際觀測值與期望值之間的落差即為估計誤差值, 在此以 ε 表示之。由於我們所要建立之模式為簡單線性迴歸模式，即找出參數 β_0 與 β_1 使得

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,$$

其中 ε_i, $i = 1, 2, \ldots n$ 為相互獨立且服從 $\mathcal{N}(0, \sigma^2)$。由上式可知，在給定 $X_i = X_i$ 時，Y_i 之條件期望值及條件變異數分別為

$$\mathrm{E}\left[Y_i | X_i = x_i\right] = \beta_0 + \beta_1 x_i \text{ 及} \mathrm{Var}\left[Y_i | X_i = x_i\right] = \sigma^2,\ i = 1, 2, \ldots, n$$

9.1 最小平方法

由於我們所要建立的模式為

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,\ i = 1, 2, \ldots, n,$$

其中 β_0, β_1 及 σ^2 均為未知參數，在此我們首先以最大概似法估計之。在給定 $X = x$ 情況下, 隨機變數 Y 服從常態分配，平均數為 $\beta_0 + \beta_1 x$ 及變異數為 σ^2，所以其 p.d.f. 為

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right]$$

故概似函數為

$$L\left(\beta_0,\beta_1,\sigma^2\right)=f\left(x_1,x_2,\ldots,x_n\right)=\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left[-\frac{\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_i\right)^2}{2\sigma^2}\right]$$

兩邊同時取 ln 得

$$\ln L\left(\beta_0,\beta_1,\sigma^2\right)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_i\right)^2}{2\sigma^2}$$

所以，可得 $\ln L\left(\beta_0,\beta_1,\sigma^2\right)$ 有最大值之必要條件為

$$\frac{\partial\ln L\left(\beta_0,\beta_1,\sigma^2\right)}{\partial\beta_0}=-\frac{\sum_{i=1}^{n}2\left(y_i-\beta_0-\beta_1x_i\right)\times(-1)}{2\sigma^2}=0,$$

$$\frac{\partial\ln L\left(\beta_0,\beta_1,\sigma^2\right)}{\partial\beta_1}=-\frac{\sum_{i=1}^{n}2\left(y_i-\beta_0-\beta_1x_i\right)\times(-x_i)}{2\sigma^2}=0$$

及

$$\frac{\partial\ln L\left(\beta_0,\beta_1,\sigma^2\right)}{\partial\sigma^2}=-\frac{n}{2}\frac{1}{\sigma^2}+\frac{\sum_{i=1}^{n}\left(y_i-\beta_0-\beta_1x_i\right)^2}{2\sigma^4}=0$$

求解上述方程組可求得

$$\widehat{\beta}_1=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)y_i}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2},\ \widehat{\beta}_0=\overline{y}-\widehat{\beta}_1\overline{x}\text{ 及}\widehat{\sigma}^2_{MLE}=\frac{\sum_{i=1}^{n}\left(y_i-\widehat{\beta}_0-\widehat{\beta}_1x_i\right)^2}{n}$$

由於上述估計之方法必須在 $\varepsilon_i\overset{iid}{\sim}\mathcal{N}\left(0,\sigma^2\right)$，因此在使用上有所限制。以下介紹一般常用的方法，其概念為找出一組 $\widehat{\beta}_0$ 及 $\widehat{\beta}_1$ 使得誤差平方和為最小，而且不受 ε_i 分配限制，一般此種方法為**普通最小平方法**(Method of Ordinarily Least Square)，簡稱 OLS。

由圖9.1可知，誤差平方和為 $\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}\left(Y_i-\beta_0-\beta_1X_i\right)^2$。由於 ε, β_0, β_1 均未知，因此分別以 e_i, $\widehat{\beta}_0$ 及 $\widehat{\beta}_1$ 作為個別之估計式。所以目標函數可寫成

$$\min\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}\left(Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right)^2$$

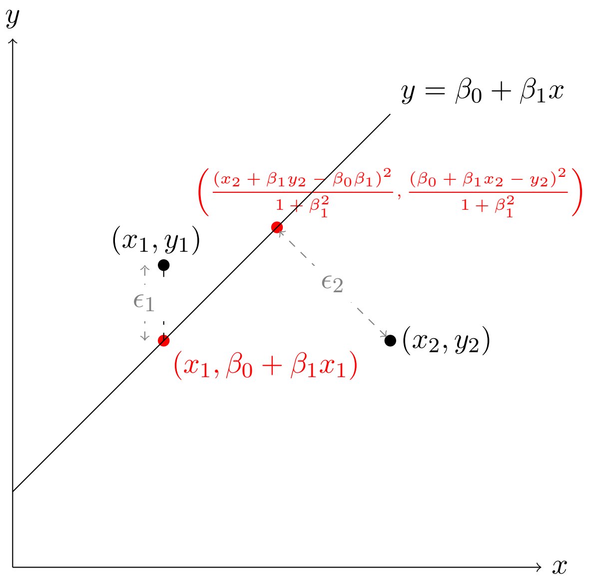


圖 9.1: 最小平方法示意圖

分別對 $\widehat{\beta}_0$ 及 $\widehat{\beta}_1$ 做微分，可得最小化之必要條件為

$$\frac{\partial \sum_{i=1}^n e_i^2}{\partial \widehat{\beta}_0} = \sum_{i=1}^n 2\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i\right) \times (-1) = 0$$

和

$$\frac{\partial \sum_{i=1}^n e_i^2}{\partial \widehat{\beta}_1} = \sum_{i=1}^n 2\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i\right) \times (-X_i) = 0$$

經化簡可得

$$n\widehat{\beta}_0 + \widehat{\beta}_1 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i$$

和

$$\widehat{\beta}_0 \sum_{i=1}^n X_i + \widehat{\beta}_1 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i$$

以矩陣方式可表示為

$$\begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix} \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{bmatrix}$$

求解二元一次方程組可得

$$\widehat{\beta}_0 = \frac{\sum_{i=1}^n Y_i \sum_{i=1}^n X_i^2 - \sum_{i=1}^n X_i Y_i \sum_{i=1}^n X_i}{n\sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2}$$

及

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{n\sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n\sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2} \\
&= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}
\end{aligned}$$

其中，$\widehat{\beta}_1$ 及 $\widehat{\beta}_0$ 亦可寫成

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2 \sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2}} \frac{\sqrt{\sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2}}{\sqrt{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}} \\
&= r_{xy} \times \sqrt{\frac{SS_y}{SS_x}}
\end{aligned}$$

及

$$\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1 \overline{X}$$

定理 9.1. 若 $Y_1, Y_2, \ldots, Y_n$ 為一組樣本數為 n 之隨機樣本，則

$$\operatorname{Cov}\left(\sum_{i=1}^n c_i Y_i, \sum_{i=1}^n d_i Y_i\right) = \sigma^2 \sum_{i=1}^n c_i d_i$$

證明.

$$\operatorname{Cov}\left(\sum_{i=1}^n c_i Y_i, \sum_{i=1}^n d_i Y_i\right)$$

$$
\begin{aligned}
&= \mathrm{E}\left(\sum_{i=1}^{n} c_i Y_i - \sum_{i=1}^{n} c_i \mu\right)\left(\sum_{i=1}^{n} d_i Y_i - \sum_{i=1}^{n} d_i \mu\right) \\
&= \mathrm{E}\left(\sum_{i=1}^{n} c_i \left(Y_i - \mu\right)\right)\left(\sum_{i=1}^{n} d_i \left(Y_i - \mu\right)\right) \\
&= \mathrm{E}\Big[c_1 d_1 \left(Y_1 - \mu\right)^2 + c_1 d_2 \left(Y_1 - \mu\right)\left(Y_2 - \mu\right) + \cdots + c_1 d_n \left(Y_1 - \mu\right)\left(Y_n - \mu\right) \\
&\quad + \cdots + c_n d_1 \left(Y_n - \mu\right)\left(Y_1 - \mu\right) + c_n d_2 \left(Y_1 - \mu\right)\left(Y_2 - \mu\right) + \cdots + c_n d_n \left(Y_n - \mu\right)^2\Big] \\
&= c_1 d_1 \mathrm{E}\left(Y_1 - \mu\right)^2 + c_1 d_1 \mathrm{E}\left(Y_1 - \mu\right)\left(Y_2 - \mu\right) + \cdots + c_1 d_n \mathrm{E}\left(Y_1 - \mu\right)\left(Y_n - \mu\right) \\
&\quad + \cdots + c_n d_1 \mathrm{E}\left(Y_n - \mu\right)\left(Y_1 - \mu\right) + c_n d_2 \mathrm{E}\left(Y_1 - \mu\right)\left(Y_2 - \mu\right) + \cdots + c_n d_n \mathrm{E}\left(Y_n - \mu\right)^2
\end{aligned}
$$

因為 $Y_i, Y_2, \ldots, Y_n$ 相互獨立, 所以對所有的 $i \neq j$ 可知

$$
\mathrm{E}\left(Y_i - \mu\right)\left(Y_j - \mu\right) = 0,
$$

故可推得

$$
\mathrm{Cov}\left(\sum_{i=1}^{n} c_i Y_i, \sum_{i=1}^{n} d_i Y_i\right) = \sigma^2 \sum_{i=1}^{n} c_i di
$$

由以上可發現以最小平方法所求得知迴歸估計式與以最大概似法所求得知迴歸估計式相等。又由於

$$
\begin{aligned}
\mathrm{E}\left(\widehat{\beta}_1\right) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\right) = \frac{\mathrm{E}\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right) Y_i\right]}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right) \mathrm{E}\left(Y_i\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)\left(\beta_0 + \beta_1 X_i\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= \sum_{i=1}^{n} \frac{\beta_0\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} + \sum_{i=1}^{n} \frac{\beta_1 X_i\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= \beta_1 \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} = \beta_1,
\end{aligned}
$$

$$
\mathrm{E}\left(\widehat{\beta}_0\right) = \mathrm{E}\left(\overline{Y} - \widehat{\beta}_1 \overline{X}\right)
$$

$$
\begin{aligned}
&= \beta_0 + \beta_1 \overline{X} - \beta_1 \overline{X} \\
&= \beta_0,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Var}\left(\widehat{\beta}_1\right) &= \operatorname{Var}\left(\frac{\sum_{i=1}^n \left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right) = \sum_{i=1}^n \operatorname{Var}\left(\frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i\right) \\
&= \sum_{i=1}^n \left[\frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right]^2 \operatorname{Var}\left(Y_i\right) \\
&= \sum_{i=1}^n \left[\frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right]^2 \sigma^2 \\
&= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\left[\sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right]^2} \sigma^2 = \frac{\sigma^2}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2},
\end{aligned}
$$

及

$$
\begin{aligned}
\operatorname{Var}\left(\widehat{\beta}_0\right) &= \operatorname{Var}\left(\overline{Y} - \widehat{\beta}_1 \overline{X}\right) = \operatorname{Var}\left(\sum_{i=1}^n \frac{1}{n} Y_i - \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \overline{X}\right) \\
&= \operatorname{Var}\left(\sum_{i=1}^n \frac{1}{n} Y_i - \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right) \overline{X}}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i\right) \\
&= \operatorname{Var}\left(\sum_{i=1}^n \left[\frac{1}{n} - \frac{\left(X_i - \overline{X}\right) \overline{X}}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right] Y_i\right) \\
&= \sum_{i=1}^n \left[\frac{1}{n} - \frac{\left(X_i - \overline{X}\right) \overline{X}}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right]^2 \sigma^2 \\
&= \sum_{i=1}^n \left[\frac{1}{n^2} + \frac{\left(X_i - \overline{X}\right)^2 \overline{X}^2}{\left[\sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right]^2} - \frac{2\left(X_i - \overline{X}\right) \overline{X}}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right] \sigma^2 \\
&= \frac{\sigma^2}{n} + \frac{\overline{X}^2 \sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\left[\sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right]^2} \sigma^2 - \frac{2\overline{X} \sum_{i=1}^n \left(X_i - \overline{X}\right)}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \sigma^2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\sigma^2}{n} + \frac{\overline{X}^2 \sigma^2}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&= \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2 + n\overline{X}^2}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \sigma^2 \\
&= \frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \sigma^2
\end{aligned}
$$

所以當 $\varepsilon_i \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$ 的假設成立時，可推得估計式 $\widehat{\beta}_0$ 和 $\widehat{\beta}_1$ 之機率分配分別為

$$
\widehat{\beta}_0 \sim \mathcal{N}\left(\beta_0, \frac{\sigma^2}{n} + \frac{\overline{X}^2 \sigma^2}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right)
$$

及

$$
\widehat{\beta}_1 \sim \mathcal{N}\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right)
$$

此外，由正規方程式 (Normal Equations), $\partial \sum_{i=1}^n e_i^2 / \partial \beta_0$ 及 $\partial \sum_{i=1}^n e_i^2 / \partial \beta_1$，可得下列性質：

1. $\sum_{i=1}^n e_i = 0$
2. $\sum_{i=1}^n X_i e_i = 0$
3. $\sum_{i=1}^n \widehat{Y}_i e_i = 0$

證明.

1. 由 $\partial \sum_{i=1}^n e_i^2 / \partial \beta_0$ 可知，

$$
\begin{aligned}
\frac{\partial \sum_{i=1}^n e_i^2}{\partial \beta_0} &= \sum_{i=1}^n 2\left(\widehat{Y}_i - \beta_0 - \beta_1 X_i\right) \times (-1) \\
&= -2 \sum_{i=1}^n e_i \\
&= 0
\end{aligned}
$$

所以，

$$\sum_{i=1}^{n} e_i = 0$$

2. 由 $\partial \sum_{i=1}^{n} e_i^2 / \partial \beta_1$ 可知，

$$\begin{aligned}
\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta_1} &= \sum_{i=1}^{n} 2\left(\widehat{Y}_i - \beta_0 - \beta_1 X_i\right) \times (-X_i) \\
&= -2 \sum_{i=1}^{n} X_i e_i \\
&= 0
\end{aligned}$$

所以，

$$\sum_{i=1}^{n} X_i e_i = 0$$

3.

$$\begin{aligned}
\sum_{i=1}^{n} \widehat{Y}_i e_i &= \sum_{i=1}^{n} \left(\widehat{\beta}_0 + \widehat{\beta}_1 X_i\right) e_i \\
&= \widehat{\beta}_0 \sum_{i=1}^{n} e_i + \widehat{\beta}_1 \sum_{i=1}^{n} X_i e_i \\
&= 0
\end{aligned}$$

由於, $\sum_{i=1}^{n} e_i = 0$, $\sum_{i=1}^{n} X_i e_i = 0$，故得

$$\sum_{i=1}^{n} \widehat{Y}_i e_i = 0$$

定理 9.2. 高斯馬可夫定理*(Guass-Makrov Theroem)*：以最小平方法所得之估計式必定為在所有線性不偏估計式中具有最小變異之估計式, 簡稱 BLUE*(Best Linear Unbiased Estimator)*。

證明.

1.

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right) Y_i}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} Y_i$$

$$
\begin{aligned}
&= \frac{\left(X_1-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}Y_1+\frac{\left(X_2-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}Y_2+\cdots+\frac{\left(X_n-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}Y_n \\
&= \sum_{i=1}^{n}k_iY_i,
\end{aligned}
$$

其中 $k_i=\left(X_i-\overline{X}\right)\Big/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$。所以, $\widehat{\beta}_1$ 為線性估計式。

2.

$$
\begin{aligned}
\mathrm{E}\left(\widehat{\beta}_1\right) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)Y_i}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)=\frac{\mathrm{E}\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)Y_i\right]}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&= \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\mathrm{E}\left(Y_i\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&= \sum_{i=1}^{n}\frac{\left(X_i-\overline{X}\right)\left(\beta_0+\beta_1X_i\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&= \sum_{i=1}^{n}\frac{\beta_0\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}+\sum_{i=1}^{n}\frac{\beta_1X_i\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&= \beta_1\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}=\beta_1
\end{aligned}
$$

所以，$\widehat{\beta}_1$ 為一不偏估計式。

3. 令 $\widetilde{\beta}_1=\sum_{i=1}^{n}c_iY_i$ 且 $\mathrm{E}\left(\widetilde{\beta}_1\right)=\beta_1$，則可推得下式

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{\beta}_1\right) &= \mathrm{E}\left[\sum_{i=1}^{n}c_iY_i\right]=\sum_{i=1}^{n}c_i\mathrm{E}\left(Y_i\right) \\
&= \sum_{i=1}^{n}c_i\left(\beta_0+\beta_1X_i\right) \\
&= \beta_0\sum_{i=1}^{n}c_i+\beta_1\sum_{i=1}^{n}c_iX_i
\end{aligned}
$$

由於 $\widetilde{\beta}_1$ 必須為一不偏估計式，故可推得

$$\sum_{i=1}^{n} c_i = 0 \text{ 且} \sum_{i=1}^{n} c_i X_i = 1$$

令 $c_i = k_i + d_i$，其中 d_i 為任意數，則

$$\begin{aligned}
\operatorname{Var}\left(\widetilde{\beta}_1\right) &= \operatorname{Var}\left(\sum_{i=1}^{n} c_i Y_i\right) = \sum_{i=1}^{n} \operatorname{Var}\left(c_i Y_i\right) \\
&= \sum_{i=1}^{n} c_i^2 \sigma^2 \\
&= \sum_{i=1}^{n} \left(k_i^2 + d_i^2 + 2k_i d_i\right) \sigma^2
\end{aligned}$$

上式中

$$\begin{aligned}
\sum_{i=1}^{n} k_i d_i &= \sum_{i=1}^{n} k_i c_i - \sum_{i=1}^{n} k_i^2 \\
&= \sum_{i=1}^{n} \frac{c_i\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} - \sum_{i=1}^{n}\left[\frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\right]^2 \\
&= \frac{\sum_{i=1}^{n} c_i X_i - \overline{X}\sum_{i=1}^{n} c_i}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} - \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right]^2} \\
&= \frac{1}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} - \frac{1}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= 0
\end{aligned}$$

所以

$$\begin{aligned}
\operatorname{Var}\left(\widetilde{\beta}_1\right) &= \sum_{i=1}^{n} k_i^2 \sigma^2 + \sum_{i=1}^{n} d_i^2 \sigma^2 \\
&= \frac{\sigma^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} + \sum_{i=1}^{n} d_i^2 \sigma^2 \\
&\geq \operatorname{Var}\left(\widehat{\beta}_1\right)
\end{aligned}$$

因此，$\widehat{\beta}_1$ 為在所有線性不偏估計式中具有最小變異數之估計式。

又最大概似法所求得之母體變異數估計值為 $\widehat{\sigma}^2 = \sum_{i=1}^{n}\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i\right)^2/n$, 由於此估計式是否為不偏估計式仍然未定，因此不能貿然使用它來作為母體變異數之估計式。為求得母體變異數之不偏估計式, 首先可將總變異來源 SSTO 分解如下：

$$
\begin{aligned}
\mathrm{SSTO} &= \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 = \sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i + \widehat{Y}_i - \overline{Y}\right)^2 \\
&= \sum_{i=1}^{n}\left(Y_i - \widehat{Y}\right)^2 + \sum_{i=1}^{n}\left(\widehat{Y} - \overline{Y}\right)^2 + 2\sum_{i=1}^{n}\left(Y_i - \widehat{Y}\right)\left(\widehat{Y} - \overline{Y}\right)
\end{aligned}
$$

其中,

$$
\begin{aligned}
\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)\left(\widehat{Y}_i - \overline{Y}\right) &= \sum_{i=1}^{n}\left(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i\right)\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_i - \overline{Y}\right) \\
&= \sum_{i=1}^{n}\left(Y_i - \overline{Y} + \widehat{\beta}_1\overline{X} - \widehat{\beta}_1 X_i\right)\left(\overline{Y} - \widehat{\beta}_1\overline{X} + \widehat{\beta}_1 X_i - \overline{Y}\right) \\
&= -\widehat{\beta}_1\sum_{i=1}^{n}\left(Y_i - \overline{Y} + \widehat{\beta}_1\overline{X} - \widehat{\beta}_1 X_i\right)\left(\overline{X} - X_i\right) \\
&= \widehat{\beta}_1\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right) - \widehat{\beta}_1^2\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \\
&= \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)\left(X_i - \overline{X}\right) \\
&\quad - \left[\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\right]^2\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \\
&= \frac{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)\right]^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} - \frac{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)\right]^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \\
&= 0
\end{aligned}
$$

由於依變數 Y 的變動並非完全不可預測，在我們知道迴歸模式中依變數 Y 有一部分可透過確定的線性估計模型來解釋，故可將總變異來源 SSTO 分解為：

$$
\mathrm{SSTO} = \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2
$$

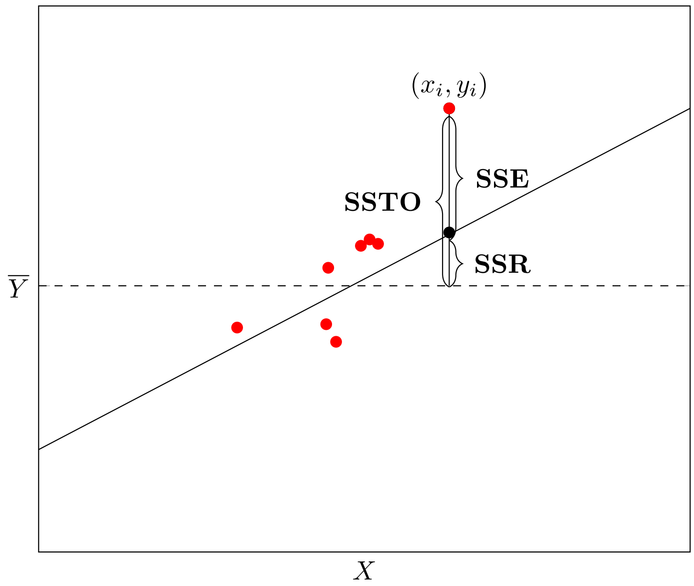


圖 9.2: 總變異分解圖

$$
\begin{aligned}
&= \sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i + \widehat{Y}_i - \overline{Y}\right)^2 \\
&= \sum_{i=1}^{n}\left(Y_i - \widehat{Y}\right)^2 + \sum_{i=1}^{n}\left(\widehat{Y} - \overline{Y}\right)^2 \\
&= \mathrm{SSE} + \mathrm{SSR}
\end{aligned}
$$

其中 SSR（Sum of squares regression）為透過解釋變數 X 建立的線性模型解釋的變異，故此一部分稱為可解釋變異；另一部分，SSE（Sum of squares error）則是不可透過上述線性模型解釋的隨機誤差，稱為不可解釋變異，說明如圖9.2。

所以 SSTO 可由可解釋變異 (SSR) 及不可解釋變異 (SSE) 所組成, 又 SSR 可寫成

$$
\begin{aligned}
\mathrm{SSR} &= \sum_{i=1}^{n}\left(\widehat{Y}_i - \overline{Y}\right)^2 = \sum_{i=1}^{n}\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_i - \overline{Y}\right)^2 = \sum_{i=1}^{n}\left(\overline{Y} - \widehat{\beta}_1\overline{X} + \widehat{\beta}_1 X_i - \overline{Y}\right)^2 \\
&= \sum_{i=1}^{n}\left(-\widehat{\beta}_1\overline{X} + \widehat{\beta}_1 X_i\right)^2 = \sum_{i=1}^{n}\left(-\widehat{\beta}_1\overline{X} + \widehat{\beta}_1 X_i\right)^2 \\
&= \widehat{\beta}_1^2 \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)\right]^2}{\sum_{i=1}^{n}\left(X_i - \overline{x}\right)^2}
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}}\right]^2 \times \sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2 \\
&= r_{xy}^2 \times \mathrm{SSTO}
\end{aligned}
$$

因此,

$$
\begin{aligned}
\mathrm{SSE} &= \sum_{i=1}^{n}\left(Y_i-\widehat{Y}\right)^2 \\
&= \sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2 - \widehat{\beta}_1^2\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \\
&= \mathrm{SSTO}-\mathrm{SSR}
\end{aligned}
$$

屬眞正的殘差，故可以利用它來估計母體變異數 σ^2。SSR 很大時，表示說依變數大部分的變異都可以經由所建立的迴歸模式解釋，即表示我們所篩選的解釋變數 X 對 Y 具預測力。相反的，若 SSR 很小，即表示所篩選的解釋變數眞的對 Y 並無預測力。最後，將各項變異來源編製成下列表格。如果 F 比值大，即 $F=\frac{MSB}{MSW}>F_{\alpha}(1,n-2)$，表示資料變易來源主要來自於研究者所篩選的解釋變數所導致，便傾向棄卻虛無假設 $H_0:\beta_1=0$。

變異數分析表

變異來源	平方和	自由度	均方和	F 值
SSR	$\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2$	1	MSR= $\frac{\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2}{k}$	$\mathrm{F}=\frac{\mathrm{MSR}}{\mathrm{MSE}}$
SSE	$\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2$	$n-2$	MSE= $\frac{\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2}{n-1-k}$	
SST	$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2$	$n-1$		

又

$$
\begin{aligned}
\mathrm{E}\left(\mathrm{SSE}\right) &= \mathrm{E}\left[\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2\right] = \sum_{i=1}^{n}\mathrm{E}\left[Y_i-\widehat{Y}_i\right]^2 = \sum_{i=1}^{n}\mathrm{E}\left[Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right]^2 \\
&= \sum_{i=1}^{n}\mathrm{Var}\left(Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right)+\sum_{i=1}^{n}\left\{\mathrm{E}\left[Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right]\right\}^2 \\
&= \sum_{i=1}^{n}\mathrm{Var}\left(Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \operatorname{Var}(Y_i) + \operatorname{Var}\left(\widehat{\beta}_0\right) + X_i^2 \operatorname{Var}\left(\widehat{\beta}_1\right) \\
&\quad -2\operatorname{Cov}\left(Y_i, \widehat{\beta}_0\right) - 2X_i \operatorname{Cov}\left(Y_i, \widehat{\beta}_1\right) + 2X_i \operatorname{Cov}\left(\widehat{\beta}_0, \widehat{\beta}_1\right)
\end{aligned}
$$

其中

$$
\begin{aligned}
\operatorname{Cov}\left(Y_i, \widehat{\beta}_0\right) &= \mathrm{E}\left(Y_i - \mu_Y\right)\left(\overline{Y} - \widehat{\beta}_1 \overline{X} - \left(\mu_y - \beta_1 \overline{X}\right)\right) \\
&= \mathrm{E}\left[Y_i - \mu_y\right]\left[\sum_{i=1}^n \frac{1}{n}\left(Y_i - \mu_y\right) - \sum_{i=1}^n \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\left(Y_i - \mu_y\right)\right] \\
&= \mathrm{E}\left[Y_i - \mu_y\right]\left[\sum_{i=1}^n \left(\frac{1}{n} - \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right)\left(Y_i - \mu_y\right)\right] \\
&= \left(\frac{1}{n} - \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right)\sigma^2,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}\left(Y_i, \widehat{\beta}_1\right) &= \mathrm{E}\left(Y_i - \mu_y\right)\left(\widehat{\beta}_1 - \beta_1\right) \\
&= \mathrm{E}\left[Y_i - \mu_y\right]\left[\sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\left(Y_i - \mu_y\right)\right] \\
&= \mathrm{E}\left[Y_i - \mu_y\right]\left[\sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\left(Y_i - \mu_y\right)\right] \\
&= \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\sigma^2
\end{aligned}
$$

及

$$
\begin{aligned}
\operatorname{Cov}\left(\widehat{\beta}_0, \widehat{\beta}_1\right) &= \mathrm{E}\left(\widehat{\beta}_0 - \beta_0\right)\left(\widehat{\beta}_1 - \beta_1\right) \\
&= \mathrm{E}\left[\sum_{i=1}^n \frac{\left(Y_i - \mu_y\right)}{n} - \sum_{i=1}^n \frac{\overline{X}\left(X_i - \overline{X}\right)\left(Y_i - \mu_y\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right]\left[\sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)\left(Y_i - \mu_y\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{E}\left[\sum_{i=1}^{n}\left(\frac{1}{n}-\frac{\overline{X}\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)(Y_i-\mu_y)\right]\left[\sum_{i=1}^{n}\frac{\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}(Y_i-\mu_y)\right] \\
&= \sigma^2\sum_{i=1}^{n}\left(\frac{1}{n}-\frac{\overline{X}\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)\times\left(\frac{\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right) \\
&= \sigma^2\sum_{i=1}^{n}\left(\frac{\left(X_i-\overline{X}\right)}{n\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}-\frac{\overline{X}\left(X_i-\overline{X}\right)^2}{\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)^2}\right) \\
&= -\sigma^2\frac{\overline{X}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\left(\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)^2} \\
&= -\sigma^2\frac{\overline{X}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}
\end{aligned}
$$

所以，$\mathrm{Var}\left(Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right)$ 可改寫成

$$
\begin{aligned}
\mathrm{Var}\left(Y_i-\widehat{\beta}_0-\widehat{\beta}_1X_i\right) &= \sigma^2+\frac{\sigma^2\sum_{i=1}^{n}X_i^2}{n\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}+\frac{\sigma^2X_i^2}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&\quad -2\left(\frac{\sigma^2}{n}-\frac{\sigma^2\overline{X}\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\right)-\frac{2\sigma^2X_i\left(X_i-\overline{X}\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}-\frac{2\sigma^2X_i\overline{X}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2} \\
&= \frac{n-2}{n}\sigma^2+\frac{\frac{\sum_{i=1}^{n}X_i^2}{n}+X_i^2+2\overline{X}\left(X_i-\overline{X}\right)-2X_i\left(X_i-\overline{X}\right)-2X_i\overline{X}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\sigma^2 \\
&= \frac{n-2}{n}\sigma^2+\frac{\frac{\sum_{i=1}^{n}X_i^2}{n}-X_i^2-2\overline{X}^2+2\overline{X}X_i}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\sigma^2
\end{aligned}
$$

故可推得

$$
\begin{aligned}
\mathrm{E}\left(\mathrm{SSE}\right) &= \sum_{i=1}^{n}\left\{\frac{n-2}{n}\sigma^2+\frac{\frac{\sum_{i=1}^{n}X_i^2}{n}-2\overline{X}^2-X_j^2+2\overline{X}X_j}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\sigma^2\right\} \\
&= (n-2)\,\sigma^2+\frac{\sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{n}X_i^2}{n}-2\overline{X}^2-X_j^2+2\overline{X}X_j\right)}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\sigma^2
\end{aligned}
$$

$$
\begin{aligned}
&= (n-2)\,\sigma^2 + \frac{\sum_{i=1}^n X_i^2 - 2n\overline{X}^2 - \sum_{j=1}^n X_j^2 + 2\overline{X}\sum_{j=1}^n X_j}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\sigma^2 \\
&= (n-2)\,\sigma^2 + \frac{\sum_{i=1}^n X_i^2 - 2n\overline{X}^2 - \sum_{j=1}^n X_j^2 + 2n\overline{X}^2}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\sigma^2 \\
&= (n-2)\,\sigma^2
\end{aligned}
$$

及

$$\mathrm{E}\,(\mathrm{MSE}) = \mathrm{E}\left(\frac{\mathrm{SSE}}{n-2}\right) = \sigma^2$$

因此，MSE 為母體變異數 σ^2 之不偏估計式，故以 MSE 作為母體變異數之估計值。由此可發現, 以最大概似法所求得之母體變異數估計式並非為不偏估計式。

由於 $\mathrm{SSTO}/\sigma^2 \sim \chi^2(n-1)$，且卡方分配具可加性, 又 $\sum_{i=1}^n \left(Y_i - \widehat{Y}\right)^2$ 與 $\sum_{i=1}^n \left(\widehat{Y} - \overline{Y}\right)^2$ 相獨立 (在此僅承認此一性質, 證明省略)，故可知 SSE/σ^2 及 SSR/σ^2 亦為卡方分配。因為 $\mathrm{E}(\mathrm{SSE}/\sigma^2) = n-2$，故可推得 $\mathrm{SSE}/\sigma^2 \sim \chi^2(n-2)$。因此由卡方分配之可加性可知 $\mathrm{SSR}/\sigma^2 \sim \chi^2(1)$。

例 **9–1** Consider the simple linear regression model $Y = \beta_0 + \beta_1 X + \varepsilon$, with $\mathrm{E}\,(\varepsilon) = 0$, $\mathrm{Var}\,(\varepsilon) = \sigma^2$ and the errors ε are uncorrelated.

1. Show that $\mathrm{Cov}\left(\widehat{\beta}_0, \widehat{\beta}_1\right) = -\overline{X}\sigma^2 / \sum_{i=1}^n \left(X_i - \overline{X}\right)^2$.
2. Show that $\mathrm{Cov}\left(\widehat{\beta}_1, \overline{Y}\right) = 0$

解：

由定理 **10-1** 可知

1.

$$
\begin{aligned}
\mathrm{Cov}\left(\widehat{\beta}_0, \widehat{\beta}_1\right) &= \mathrm{Cov}\left(\overline{Y} - \sum_{i=1}^n \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i, \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i\right) \\
&= \mathrm{Cov}\left(\sum_{i=1}^n \left[\frac{1}{n} - \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right] Y_i, \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i\right) \\
&= \sigma^2 \sum_{i=1}^n \left[\frac{1}{n} - \frac{\overline{X}\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}\right] \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&= \sigma^2 \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{n\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} - \sigma^2 \sum_{i=1}^n \frac{\overline{X}\left(X_i - \overline{X}\right)^2}{\left[\sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right]^2}
\end{aligned}
$$

$$= \quad -\sigma^2 \frac{\overline{X}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}{\left[\sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right]^2} = -\frac{\sigma^2 \overline{X}}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}$$

2.

$$\begin{aligned}
\operatorname{Cov}\left(\widehat{\beta}_1, \overline{Y}\right) &= \operatorname{Cov}\left(\sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2} Y_i, \sum_{i=1}^n \frac{1}{n} Y_i\right) \\
&= \sigma^2 \sum_{i=1}^n \frac{\left(X_i - \overline{X}\right)}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2} = \sigma^2 \frac{\sum_{i=1}^n \left(X_i - \overline{X}\right)}{n \sum_{i=1}^n \left(X_i - \overline{X}\right)^2} \\
&= 0
\end{aligned}$$

9.2 β_0 與 β_1 之統計推論

由於 $\sum_{i=1}^n \left(\widehat{Y}_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i\right)^2 / \sigma^2 \sim \chi^2(n-2)$，因此由 **7.2.6** 可知，

$$\begin{aligned}
1 - \alpha &= \mathrm{P}\left(\chi^2_{1-\frac{\alpha}{2}}(n-2) < \frac{(n-2) \times \mathrm{MSE}}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-2)\right) \\
&= \mathrm{P}\left(\frac{1}{\chi^2_{\frac{\alpha}{2}}(n-2)} < \frac{\sigma^2}{(n-2) \times \mathrm{MSE}} < \frac{1}{\chi^2_{1-\frac{\alpha}{2}}(n-2)}\right) \\
&= \mathrm{P}\left(\frac{(n-2) \times \mathrm{MSE}}{\chi^2_{\frac{\alpha}{2}}(n-2)} < \sigma^2 < \frac{(n-2) \times \mathrm{MSE}}{\chi^2_{1-\frac{\alpha}{2}}(n-2)}\right)
\end{aligned}$$

所以母體變異數 σ^2 之 $(1-\alpha) \times 100\%$ 之信賴區間為

$$\left[\frac{(n-2) \times \mathrm{MSE}}{\chi^2_{\frac{\alpha}{2}}(n-2)}, \frac{(n-2) \times \mathrm{MSE}}{\chi^2_{1-\frac{\alpha}{2}}(n-2)}\right]$$

又 $\widehat{\beta}_1 \sim \mathcal{N}\left(\beta_1, \sigma^2 / \sum_{i=1}^n \left(X_i - \overline{X}\right)^2\right)$, $\mathrm{SSE}/\sigma^2 \sim \chi^2(n-2)$, 及 σ^2 未知，因此由 t 分配之定義可知

$$\frac{\frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\sigma^2 / \sum_{i=1}^n (X_i - \overline{x})^2}}}{\sqrt{\frac{\mathrm{SSE}/\sigma^2}{n-2}}} = \frac{\widehat{\beta}_1 - \beta_1}{\sqrt{\mathrm{MSE}/\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}} \sim t\,(n-2)$$

故由 **7.2.2** 可求得

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-2) < \frac{\widehat{\beta}_1-\beta_1}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} < t_{\frac{\alpha}{2}}(n-2)\right) \\
&= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} < \widehat{\beta}_1-\beta_1 < t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\right) \\
&= \mathrm{P}\left(\widehat{\beta}_1-t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} < \beta_1 < \widehat{\beta}_1+t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\right)
\end{aligned}$$

因此，β_1 之 $(1-\alpha)\times 100\%$ 信賴區間為

$$\left[\widehat{\beta}_1-t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}, \widehat{\beta}_1+t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}\right]$$

又由 **8.2** 可知，當統計假設為雙尾檢定時，即

$$\begin{cases} H_0: \beta_1 = k \\ H_1: \beta_1 \neq k \end{cases}$$

由型 I 誤差之定義可知

$$\begin{aligned}
\alpha &= \mathrm{P}\left(\widehat{\beta}_1 < c_1 \text{ 或} \widehat{\beta}_1 > c_2 \middle| \beta_1 = k\right) \\
&= \mathrm{P}\left(\frac{\widehat{\beta}_1-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} < \frac{c_1-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\right) \\
&\quad +\mathrm{P}\left(\frac{\widehat{\beta}_1-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} > \frac{c_2-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\right)
\end{aligned}$$

由於為雙尾檢定，故取左右兩邊機率相等可得

$$-t_{\frac{\alpha}{2}}(n-2) = \frac{c_1-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)}} \text{ 及} t_{\frac{\alpha}{2}}(n-2) = \frac{c_2-k}{\sqrt{\mathrm{MSE}/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)}}$$

故可求得棄卻域為

$$C=\left\{\widehat{\beta}_1<k-t_{\frac{\alpha}{2}}\left(n-2\right)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)}}\text{ 或}\widehat{\beta}_1>k+t_{\frac{\alpha}{2}}\left(n-2\right)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)}}\right\}$$

當 $k=0$ 時，

$$\begin{aligned}\frac{\widehat{\beta}_1-0}{\sqrt{\sigma^2/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}} &= r\times\frac{\sqrt{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\times\frac{1}{\sqrt{\sigma^2/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\\ &= r\times\sqrt{\frac{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}{\sigma^2}}\end{aligned}$$

及

$$\sqrt{\frac{\mathrm{SSE}/\sigma^2}{n-2}}=\sqrt{\frac{\left(1-r^2\right)\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}{\left(n-2\right)\sigma^2}}=\sqrt{\frac{1-r^2}{n-2}}\times\sqrt{\frac{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}{\sigma^2}}$$

所以,

$$\begin{aligned}\frac{\widehat{\beta}_1-0}{\sqrt{\sigma^2/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}}\Bigg/\sqrt{\frac{\mathrm{SSE}/\sigma^2}{n-2}} &= r\times\sqrt{\frac{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}{\sigma^2}}\times\sqrt{\frac{n-2}{1-r^2}}\times\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}}\\ &= \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\end{aligned}$$

故可知統計量

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\sim t\left(n-2\right)$$

所以可推得下列結論：檢定 β_1 是否為 0 與檢定相關係數是否為 0 所得結論相同，即

$$\left\{\begin{array}{l}H_0:\beta_1=0\\ H_1:\beta_1\neq 0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}H_0:\rho=0\\ H_1:\rho\neq 0\end{array}\right.$$

因此，由型 I 誤差之定義可知

$$1-\alpha=\mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-2)<\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}<t_{\frac{\alpha}{2}}(n-2)\right)$$

故當

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}<-t_{\frac{\alpha}{2}}(n-2)\ \text{或}\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}>t_{\frac{\alpha}{2}}(n-2)\ \text{時},$$

則棄卻虛無假設 $H_0:\rho=0$。

例 **9–2**　為決定廣告費與銷貨額之關係, 乃從事一項研究，所得樣本如下：

廣告費：X	40	25	20	30	40	25	20	50	25	50
銷貨額：Y	52	48	40	48	49	40	42	56	36	51

單位：萬元，若假設條件皆符合簡單迴歸模式,

1. 求銷貨額對廣告費的迴歸方程式為何？
2. 試求 β_1 之 95% 信賴區間？
3. 以顯著水準為 0.05，試檢定 β_1 是否為 0？
4. 以顯著水準為 0.05，試檢定 ρ 是否為 0？

解：

1.

$$\widehat{\beta}_1=\frac{\sum_{i=1}^n(x_i-\overline{x})\,y_i}{\sum_{i=1}^n(x_i-\overline{x})^2}=\frac{\sum_{i=1}^n x_iy_i-\sum_{i=1}^n x_i\sum_{i=1}^n y_i/n}{\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2/n}$$

$$\sum_{i=1}^n x_i=325,\sum_{i=1}^n y_i=462,\sum_{i=1}^n x_iy_i=15570,\sum_{i=1}^n x_i^2=11775,\sum_{i=1}^n y_i^2=21710$$

$$\widehat{\beta}_1=\frac{15570-325\times 462/10}{11775-325^2/10}=0.4577\ \text{及}\widehat{\beta}_0=\frac{462}{10}-0.4577\times\frac{325}{10}=31.325$$

故可求得迴歸估計式為

$$\widehat{y}=31.397+0.4626x$$

2.

$$\text{SSTO}\quad=\quad\sum_{i=1}^n(y_i-\overline{y})^2=\sum_{i=1}^n y_i^2-\frac{\left(\sum_{i=1}^n y_i\right)^2}{n}$$

$$
\begin{aligned}
&= 21710-\frac{462^2}{10}\\
&= 365.6
\end{aligned}
$$

及

$$
\begin{aligned}
\mathrm{SSR} &= \widehat{\beta}_1^2\sum_{i=1}^{n}(x_i-\overline{x})^2=\widehat{\beta}_1^2\times\left[\sum_{i=1}^{n}x_i^2-n\overline{x}^2\right]\\
&= 0.4577^2\times\left[11775-10\times\left(\frac{325}{10}\right)^2\right]\\
&= 254.01
\end{aligned}
$$

可求得

$$
\mathrm{SSE}=\mathrm{SSTO}-\mathrm{SSR}=365.6-245.01=111.59\text{ 及 }\mathrm{MSE}{=}\frac{111.59}{10-2}=13.949
$$

$$
\begin{aligned}
&\left[\widehat{\beta}_1-t_{0.025}(8)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}},\widehat{\beta}_1+t_{0.025}(8)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}\right]\\
=&\left[0.4577-2.306\times\sqrt{\frac{13.949}{11775-325^2/10}},0.4577+2.306\times\sqrt{\frac{13.949}{11775-325^2/10}}\right]\\
=&\ [0.2104,0.7050]
\end{aligned}
$$

3.

$$
\begin{cases}
H_0:\beta_1=0\\
H_1:\beta_1\neq 0
\end{cases}
$$

$$
C=\left\{\widehat{\beta}_1<-t_{0.025}(8)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}=-0.2473\text{ 或}\widehat{\beta}_1>t_{0.025}(8)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}=0.2473\right\}
$$

又 $\widehat{\beta}_1=0.4626\in C$，所以棄卻 $H_0:\beta=0$ 的虛無假設。

4.

$$
\begin{cases}
H_0:\rho=0\\
H_1:\rho\neq 0
\end{cases}
$$

$$
r=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2\sum_{i=1}^{n}(y_i-\overline{x})^2}}=\frac{15570-\frac{325\times 462}{10}}{\sqrt{\left(11775-\frac{325^2}{10}\right)\left(21710-\frac{462^2}{10}\right)}}=0.8336
$$

故可求得棄卻域為

$$
C=\left\{\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}<-t_{0.025}(8)=-2.306\text{ 或}\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}>t_{0.025}(8)=2.306\right\}
$$

又

$$\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}=\frac{0.8336\sqrt{10-2}}{\sqrt{1-0.8336^2}}=4.2685\in C,$$

所以棄卻 $H_0:\beta=0$ 的虛無假設。

9.3 $\mu_{y|x}$ 之信賴區間及預測區間之估計

由之前討論可知，$\widehat{\beta}_0$ 及 $\widehat{\beta}_1$ 分別為 β_0 及 β_0 之不偏估計式，因此以 $\widehat{Y}$ 作為 $\mu_{y|x}=\beta_0+\beta_1X$ 之估計值。又由於

$$\begin{aligned}
\mathrm{Var}\left(\widehat{Y}\right) &= \mathrm{Var}\left(\widehat{\beta}_0\right)+X^2\mathrm{Var}\left(\widehat{\beta}_1\right)+2XCov\left(\widehat{\beta}_0,\widehat{\beta}_1\right)\\
&= \frac{\sigma^2}{n}+\frac{\overline{X}^2\sigma^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}+\frac{X^2\sigma^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}-\frac{2X\overline{X}\sigma^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}\\
&= \left(\frac{1}{n}+\frac{\left(X-\overline{X}\right)^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}\right)\sigma^2
\end{aligned}$$

故可得

$$\widehat{Y}\sim\mathcal{N}\left[\beta_0+\beta_1X,\left(\frac{1}{n}+\frac{\left(X-\overline{X}\right)^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}\right)\sigma^2\right]$$

由於母體變異數 σ^2 未知，因此由定義可知

$$\frac{\widehat{Y}_i-\mu_{y|x}}{\sqrt{\mathrm{MSE}\left(\frac{1}{n}+\frac{\left(X_i-\overline{X}\right)^2}{\sum_{i=1}^n\left(X_i-\overline{X}\right)^2}\right)}}\sim t\left(n-2\right)$$

由信賴區間之定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}\left(n-2\right)<\frac{\widehat{Y}-\mu_{y|x}}{\sqrt{\mathrm{MSE}\left(1/n+\left(X-\overline{X}\right)^2\Big/\sum_{i=1}^n\left(X_i-\overline{X}\right)^2\right)}}<t_{\frac{\alpha}{2}}\left(n-2\right)\right)\\
&= \mathrm{P}\left(-t_{\frac{\alpha}{2}}\left(n-2\right)\widehat{\sigma}_{y|x}<\widehat{Y}-\mu_{y|x}<t_{\frac{\alpha}{2}}\left(n-2\right)\widehat{\sigma}_{y|x}\right)\\
&= \mathrm{P}\left(\widehat{Y}-t_{\frac{\alpha}{2}}\left(n-2\right)\widehat{\sigma}_{y|x}<\mu_{y|x}<\widehat{Y}+t_{\frac{\alpha}{2}}\left(n-2\right)\widehat{\sigma}_{y|x}\right),
\end{aligned}$$

其中 $\widehat{\sigma}_{y|x}^2 = \sqrt{\text{MSE}\left(1/n + \left(X - \overline{X}\right)^2 \Big/ \sum_{i=1}^n \left(X - \overline{X}\right)^2\right)}$。所以, 在給定 $X = X_0$ 的條件下，母體平均數，Y, 之 $100 \times (1-\alpha)\%$ 信賴區間為

$$\left[\widehat{\beta}_0 + \widehat{\beta}_1 x_0 - t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\text{MSE}}{n} + \frac{(x_0 - \overline{x})^2 \,\text{MSE}}{\sum_{i=1}^n (x_i - \overline{x})^2}}, \widehat{\beta}_0 + \widehat{\beta}_1 x_0 + t_{\frac{\alpha}{2}}(n-2)\sqrt{\frac{\text{MSE}}{n} + \frac{(x_0 - \overline{x})^2 \,\text{MSE}}{\sum_{i=1}^n (x_i - \overline{x})^2}}\right]$$

例 **9–3** 承例 **10-1**，試求當廣告費用為 45 萬元時，銷貨量期望值之 95% 信賴區間為何？

解：

$$\sqrt{\frac{\text{MSE}}{n} + \frac{(x_0 - \overline{x})^2 \,\text{MSE}}{\sum_{i=1}^n (x_i - \overline{x})^2}} = \sqrt{\frac{13.949}{10} + \frac{(45 - 32.5)^2 \times 13.949}{11775 - \frac{325^2}{10}}} = 1.7867$$

所以在給定廣告費為 45 萬元的條件下，銷貨量期望值之 95% 信賴區間為

$$[31.325 + 0.4577 \times 45 - 2.306 \times 1.7867, 31.325 + 0.4577 \times 45 + 2.306 \times 1.7867] = [47.801, 56.042]$$

當給定 $X = X_{n+1}$ 的條件下，則由簡單迴歸模式之定義可知

$$Y_{n+1} = \beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1}$$

若隨機變數 $W = \widehat{Y}_{n+1} - Y_{n+1}$，由於

$$\begin{aligned}
\text{E}(W) &= \text{E}\left(\widehat{Y}_{n+1} - Y_{n+1}\right) \\
&= \text{E}\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{n+1} - \beta_0 - \beta_1 X_{n+1} - \varepsilon_{n+1}\right) \\
&= 0
\end{aligned}$$

及

$$\begin{aligned}
\text{Var}\left(\widehat{Y}_{n+1} - Y_{n+1}\right) &= \text{Var}\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{n+1} - \beta_0 - \beta_1 X_{n+1} - \varepsilon_{n+1}\right) \\
&= \text{Var}\left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{n+1}\right) + \text{Var}(\varepsilon_{n+1}) \\
&= \frac{\sigma^2}{n} + \frac{\sigma^2 \left(X_i - \overline{X}\right)^2}{\sum_{i=1}^n \left(X_{n+1} - \overline{X}\right)^2} + \sigma^2
\end{aligned}$$

故可求得

$$\frac{\widehat{Y}_{n+1}-Y_{n+1}}{\sqrt{\mathrm{MSE}\left(1+1/n+\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\Big/\left(\left(X_i-\overline{X}\right)^2\right)\right)}}\sim t_{n-2}$$

由信賴區間之定義可知

$$\begin{aligned}
1-\alpha &= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-2)<\frac{\widehat{Y}_{n+1}-Y_{n+1}}{\sqrt{\mathrm{MSE}\left(1+1/n+\left(X_i-\overline{X}\right)^2\Big/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)}}<t_{\frac{\alpha}{2}}(n-2)\right)\\
&= \mathrm{P}\left(-t_{\frac{\alpha}{2}}(n-2)\,\widehat{\sigma}_{\widehat{Y}_{n+1}-Y_{n+1}}<\widehat{Y}_{n+1}-Y_{n+1}<t_{\frac{\alpha}{2}}(n-2)\,\widehat{\sigma}_{\widehat{Y}_{n+1}-Y_{n+1}}\right)\\
&= \mathrm{P}\left(\widehat{Y}_{n+1}-t_{\frac{\alpha}{2}}(n-2)\,\widehat{\sigma}_{\widehat{Y}_{n+1}-Y_{n+1}}<Y_{n+1}<\widehat{Y}_{n+1}+t_{\frac{\alpha}{2}}(n-2)\,\widehat{\sigma}_{\widehat{Y}_{n+1}-Y_{n+1}}\right)
\end{aligned}$$

其中 $\widehat{\sigma}_{\widehat{Y}_{n+1}-Y_{n+1}}=\sqrt{\mathrm{MSE}\left(1+1/n+\left(X_{n+1}-\overline{X}\right)^2\Big/\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right)}$。故可求得當 $X_{n+1}=x_{n+1}$ 時，Y_{n+1} 之 $100\times(1-\alpha)\%$ 預測區間為

$$\left[\widehat{\beta}_0+\widehat{\beta}_1x_{n+1}-t_{\frac{\alpha}{2}}(n-2)\sqrt{\mathrm{MSE}\left(1+\frac{1}{n}+\frac{(x_{n+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)},\right.$$
$$\left.\widehat{\beta}_0+\widehat{\beta}_1x_{n+1}+t_{\frac{\alpha}{2}}(n-2)\sqrt{\mathrm{MSE}\left(1+\frac{1}{n}+\frac{(x_{n+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)}\right]$$

例 **9–4**　承例 **10-1**，試求當廣告費用為 45 萬元時，銷貨量之 95% 預測區間為何？

解：

$$\sqrt{\mathrm{MSE}+\frac{\mathrm{MSE}}{n}+\frac{(x_0-\overline{x})^2\,\mathrm{MSE}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}=\sqrt{13.949+\frac{13.949}{10}+\frac{(45-32.5)^2\times13.949}{11775-\frac{325^2}{10}}}=4.1402$$

所以在給定廣告費為 45 萬元的條件下，銷貨量之 95% 預測區間為

$$[31.325+0.4577\times45-2.306\times4.1402,31.325+0.4577\times45+2.306\times4.1402]=[42.374,61.469]$$

簡單迴歸研究中是研究獨立變數 X 與反應變數 Y 之關係，故很自然的我們會想到一個問題就是 X 對 Y 的影響為何？而變數 X 對 Y 之解釋能力貢獻為何？在前面我們以了解可利用是否為 0 的檢定來探討 X 對 Y 是否有影響，今我們可以利用更普遍化之方法，變異數分析 (ANOVA), 來探討且

此方法對未來複迴歸分析中有相當助益。由 **10.1** 可知總變異 SSTO 可分解成 SSR(可解釋變異) 及 SSE(不可解釋變異)，即

$$\begin{aligned}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2 &= \text{SSR}+\text{SSE}\\ &= \sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2+\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2\end{aligned}$$

由上式可知，獨立變數 X 對反應變數 Y 的變異解釋能力為

$$R^2=\frac{\text{SSR}}{\text{SSTO}}$$

一般將 R^2 稱為判定係數(Coefficient of Determination)。此外，我們也可由 R^2 之定義推導出下列結果

$$\begin{aligned}R^2 &= \frac{\text{SSR}}{\text{SSTO}}=\frac{\widehat{\beta}_1^2\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}\\ &= \frac{\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)\right]^2}{\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]^2}\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}\\ &= \frac{\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)\right]^2}{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}\\ &= \left[\frac{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2}}\right]^2\\ &= r_{xy}^2\end{aligned}$$

故可知判定係數即為樣本相關係數的平方。最後將各項變異來源編製成下列表格

ANOVA Table

變異來源	平方和	自由度	均方和	F 值
SSR	$\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2$	1	MSR= $\frac{\sum_{i=1}^{n}\left(\widehat{Y}_i-\overline{Y}\right)^2}{1}$	F = $\frac{\text{MSR}}{\text{MSE}}$
SSE	$\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2$	$n-2$	MSE= $\frac{\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2}{n-2}$	
SSTO	$\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2$	$n-1$		

由於 $t^2 = F$，因此利用上表可用來檢定下列問題

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

棄卻域為

$$C = \left\{ F = \frac{\text{MSR}}{\text{MSE}} > F_\alpha(1, n-2) \right\}$$

例 **9–5** 承例 **10-1**，試以 F 檢定法檢定在 95% 信賴水準下 β_1 是否為 0？

解：

ANOVA Table

變異來源	平方和	自由度	均方和	F 值
迴歸	245.01	1	245.01	18.209
殘差	111.59	8	13.95	
總和	365.6	9		

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

棄卻域為

$$C = \left\{ F = \frac{\text{MSR}}{\text{MSE}} > F_\alpha(1, 8) = 5.32 \right\}$$

由於 $F = 18.209 \in C$，所以棄卻 $H_0 : \beta_1 = 0$ 的虛無假設。

例 **9–6** 某工管人員研究握力 (X) 與抬舉重量 (Y) 之間的關係; 他找了 20 者蒐集資料, 並將結果整理如下：

$$n = 20,\ \sum_{i=1}^{n} x = 550,\ \sum_{i=1}^{n} y = 1620,\ \sum_{i=1}^{n} x^2 = 16080,\ \sum_{i=1}^{n} y^2 = 135000,\ \sum_{i=1}^{n} xy = 45900$$

1. 試求 X 與 Y 間的相關係數 r; 並檢定是否顯著 $(\alpha = 0.05)$
2. 試求 X 與 Y 間的線性迴歸; 並檢定其顯著性 $(\alpha = 0.05)$
3. 某人握力為 30，試求抬舉重量的 95% 預測區間

解：

1.

$$\widehat{\beta}_1=\frac{\sum_{i=1}^n xy-\sum_{i=1}^n x\sum_{i=1}^n y/\,n}{\sum_{i=1}^n x^2-\left(\sum_{i=1}^n x\right)^2\Big/\,n}=\frac{45900-550\times 1620/20}{16080-550^2/20}=1.4136$$

所以

$$\mathrm{SSR}=\widehat{\beta}_1^2\sum_{i=1}^n (x_i-\overline{x})^2=1.4136^2\times\left(16080-550^2/20\right)=1908.3,$$

$$\mathrm{SSTO}=\sum_{i=1}^n (y_i-\overline{y})^2=135000-1620^2/20=3780$$

得

$$R^2=\frac{\mathrm{SSR}}{\mathrm{SSTO}}=\frac{\widehat{\beta}_1^2\sum_{i=1}^n (x_i-\overline{x})^2}{\sum_{i=1}^n (y_i-\overline{y})^2}=\frac{1.4136^2\times\left(16080-550^2/20\right)}{135000-1620^2/20}=0.5049$$

由於 $\widehat{\beta}>0$，所以 $r=\sqrt{0.5049}=0.7106$

$$\begin{cases} H_0:\rho=0\\ H_1:\rho\neq 0\end{cases}$$

$$C=\left\{\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}<-t_{0.025}\left(18\right)=-2.101 \text{ or } \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}>t_{0.025}\left(18\right)=2.101\right\}$$

$$\text{由於}\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}=3.4925\in C\text{，所以棄卻}H_0$$

2. $\widehat{Y}_i=\widehat{\beta}_0+\widehat{\beta}_1X_i=\overline{y}-\widehat{\beta}_1\overline{X}+\widehat{\beta}_1X_i=42.126+1.4136X_i$

$$\begin{cases} H_0:\beta_1=0\\ H_1:\beta_1\neq 0\end{cases}$$

又

$$\mathrm{MSE}=\frac{\left(1-R^2\right)\times\mathrm{SSTO}}{n-2}=\frac{(1-0.5049)\times 3780}{18}=103.97$$

所以棄卻域為

$$C=\left\{\widehat{\beta}_1<-t_{0.025}\left(18\right)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^n (x_i-\overline{x})^2}}=-0.693\text{ 或}\widehat{\beta}_1>t_{0.025}\left(18\right)\sqrt{\frac{\mathrm{MSE}}{\sum_{i=1}^n (x_i-\overline{x})^2}}=0.693\right\}$$

$$\text{由於}\widehat{\beta}_1=1.4136\in C\text{，所以棄卻}H_0$$

3. $\widehat{Y}_{X=30} = 42.126 + 1.4136X_i = 84.534$，所以 $\widehat{Y}_{X=30}$ 之 95% 預測區間為

$$\begin{aligned}
&\left[\widehat{Y}_{X=30} - t_{0.025}(18)\sqrt{\mathrm{MSE}}\sqrt{1+\frac{1}{n}+\frac{(x_{i+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}},\right.\\
&\left.\widehat{Y}_{X=30} + t_{0.025}(18)\sqrt{\mathrm{MSE}}\sqrt{1+\frac{1}{n}+\frac{(x_{i+1}-\overline{x})^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}}\right]\\
=\ &[84.534 - 21.423, 84.534 + 21.423]\\
=\ &[62.51, 106.55]
\end{aligned}$$

例 **9–7** Support that the simple linear regression model $Y = \beta_0 + \beta_1 X + \varepsilon$ describes the following data

X	6	2	4	1	5	6	3
Y	33	10	22	6	29	30	18

1. Compute the least squares point estimates of $\widehat{\beta}_0$ and $\widehat{\beta}_1$.
2. Calculate SSE.

1.

$$\sum_{i=1}^{7} x_i = 27,\ \sum_{i=1}^{7} x_i^2 = 127,\ \sum_{i=1}^{7} y_i = 148,\ \sum_{i=1}^{7} y_i^2 = 3774,\ \sum_{i=1}^{7} x_i y_i = 691$$

可求得

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{7}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{7}(x_i-\overline{x})^2} = \frac{691 - 7(27/7)(148/7)}{127 - 7(27/7)^2} = 5.2563,\ \widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x} = 0.8697$$

2.

$$\mathrm{SSTO} = \sum_{i=1}^{7}(y_i-\overline{y})^2 = 644.86,\ \mathrm{SSR} = \widehat{\beta}_1^2\sum_{i=1}^{7}(x_i-\overline{x})^2 = 631.51$$

所以

$$\mathrm{SSE} = \mathrm{SSTO} - \mathrm{SSR} = 644.86 - 631.51 = 13.347$$

例 **9–8** 欲知理工及管理學院大學畢業生第一年工作的平均薪資 (Y) 是否與畢業的學院別有關, 設三虛擬變數：D_1, D_2, D_3; 其中

$$D_1 = \begin{cases} 1, & \text{理學院} \\ 0, & \text{其他} \end{cases},\ D_2 = \begin{cases} 1, & \text{工學院} \\ 0, & \text{其他} \end{cases},\ D_3 = \begin{cases} 1, & \text{管理學院} \\ 0, & \text{其他} \end{cases}$$

已知 Y 對 D_1, D_2 的迴歸估計式為

$$\widehat{Y} = 69 + 45D_1 + 18D_2$$

試求

1. Y 對 D_1, D_3 的迴歸估計式
2. Y 對 D_2, D_3 的迴歸估計式

解：

由 $\widehat{Y} = 69 + 45D_1 + 18D_2$ 可知

$$\mu_{\text{理學院}} = 69 + 45 = 114, \mu_{\text{工學院}} = 69 + 18 = 87, \mu_{\text{管理學院}} = 69$$

所以

1. Y 對 D_1, D_3 的迴歸估計式為

$$\widehat{Y} = 87 + (114 - 87)D_1 + (69 - 87)D3 = 87 + 27D_1 - 18D_3$$

2. Y 對 D_2, D_3 的迴歸估計式為

$$\widehat{Y} = 114 - (87 - 114)D_2 + (69 - 114)D3 = 114 - 27D_2 - 45D_3$$

例 **9–9** 考慮下列簡單迴歸模型：$Y_i = \alpha + \beta X_i + u_i$. 其估計式之一為

$$\widetilde{\beta} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{Y_i - Y_{i-1}}{X_i - X_{i-1}},$$

另一個為

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

1. 解釋估計式 $\widetilde{\beta}$ 的涵義
2. $\widetilde{\beta}$ 是否為未知參數 β 的一致估計量
3. 哪一個估計是較為有效率？$\widetilde{\beta}$ 或是 β？

解：

1. $\widetilde{\beta}$ 為 $n-1$ 個斜率之算術平均數。

2.

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{\beta}\right) &= \mathrm{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}\frac{Y_i-Y_{i-1}}{X_i-X_{i-1}}\right] \\
&= \frac{1}{n-1}\mathrm{E}\left[\sum_{i=1}^{n}\frac{Y_i-Y_{i-1}}{X_i-X_{i-1}}\right] \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\frac{\mathrm{E}Y_i-\mathrm{E}Y_{i-1}}{X_i-X_{i-1}} \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\frac{\alpha+\beta X_i-\alpha-\beta X_{i-1}}{X_i-X_{i-1}} \\
&= \frac{\beta}{n-1}\sum_{i=1}^{n}\frac{X_i-X_{i-1}}{X_i-X_{i-1}} \\
&= \beta
\end{aligned}
$$

所以 $\widetilde{\beta}$ 亦為 β 之線性不偏估計式。又

$$
\begin{aligned}
\mathrm{Var}\left(\widetilde{\beta}\right) &= \mathrm{Var}\left(\frac{1}{n-1}\sum_{i=1}^{n}\frac{Y_i-Y_{i-1}}{X_i-X_{i-1}}\right) = \frac{1}{(n-1)^2}\mathrm{Var}\left(\sum_{i=1}^{n}\frac{Y_i-Y_{i-1}}{X_i-X_{i-1}}\right) \\
&= \frac{1}{(n-1)^2}\sum_{i=1}^{n}\frac{\mathrm{Var}\left(Y_i-Y_{i-1}\right)}{\left(X_i-X_{i-1}\right)^2} \\
&= \frac{1}{(n-1)^2}\frac{2\sigma^2}{\sum_{i=1}^{n}\left(X_i-X_{i-1}\right)^2}
\end{aligned}
$$

且

$$
\lim_{n\to\infty}\frac{1}{(n-1)^2}\frac{2\sigma^2}{\sum_{i=1}^{n}\left(X_i-X_{i-1}\right)^2}=0
$$

所以 $\widetilde{\beta}$ 為 β 的一致估計量。

3. 根據高斯馬可夫定理 (第250頁), 以最小平方法所求得之估計式為在所有不偏估計式中具有最小變異者，因此 $\widehat{\beta}$ 為較有效率之估計式。

例 9–10 For the case of the first order autocorrelation:

$$
\begin{cases} y_t=\beta_1+\beta_2X_t+e_t \\ e_t=\rho e_{t-1}+v_t \end{cases} \quad \text{where } -1<\rho<1 \text{ and } v_t \text{ is white noise.}
$$

Prove that $\mathrm{Var}\left(e_t\right)=\sigma_{v_t}^2/\left(1-\rho^2\right)$

$$
e_t=\rho e_{t-1}+v_t
$$

解：

$$\begin{aligned}
\mathrm{Var}(e_t) &= \mathrm{Var}(\rho e_{t-1}+v_t)\\
&= \rho^2\mathrm{Var}(e_{t-1})+\mathrm{Var}(v_t)\\
&= \rho^2\left[\rho^2\mathrm{Var}(e_{t-2})+\mathrm{Var}(v_{t-1})\right]+\mathrm{Var}(v_t)\\
&= \rho^4\mathrm{Var}(e_{t-2})+\rho^2\mathrm{Var}(v_{t-1})+\mathrm{Var}(v_t)\\
&= \rho^6\mathrm{Var}(e_{t-3})+\rho^4\mathrm{Var}(e_{t-2})+\rho^2\mathrm{Var}(v_{t-1})+\mathrm{Var}(v_t)\\
&\vdots\\
&= \mathrm{Var}(v_t)+\rho^2\mathrm{Var}(v_{t-1})+\rho^4\mathrm{Var}(e_{t-2})+\rho^6\mathrm{Var}(e_{t-3})+\cdots\\
&= \mathrm{Var}(v_t)\left[1+\rho^2+\rho^4+\rho^6+\cdots\right]\\
&= \frac{\sigma_{v_t}^2}{1-\rho^2}
\end{aligned}$$

例 **9–11** Show that

$$r_s = 1-\frac{6\sum_{i=1}^n d_i^2}{n\left(n^2-1\right)},$$

where $d_i=$ the difference between two ranks of x_i and y_i, $i=1,2,\ldots,n$. (i.e. $d_i=R(x_i)-R(y_i)$)

證明.

$$\begin{aligned}
r_s &= \frac{\sum_{i=1}^n\left[R(x_i)-\sum_{i=1}^n R(x_i)/n\right]\left[R(y_i)-\sum_{i=1}^n R(y_i)/n\right]}{\sqrt{\sum_{i=1}^n\left[R(x_i)-\sum_{i=1}^n R(x_i)/n\right]\sum_{i=1}^n\left[R(y_i)-\sum_{i=1}^n R(y_i)/n\right]}}\\
&= \frac{\sum_{i=1}^n R(x_i)R(y_i)-n\left[(n+1)/2\right]^2}{\sqrt{\left[\sum_{i=1}^n R(x_i)^2-n\left[(n+1)/2\right]^2\right]\left[\sum_{i=1}^n R(x_i)^2-n\left[(n+1)/2\right]^2\right]}}\\
&= \frac{\sum_{i=1}^n R(x_i)R(y_i)-n\left[(n+1)/2\right]^2}{n(n+1)(2n+1)/6-n\left[(n+1)/2\right]^2}
\end{aligned}$$

由於

$$\begin{aligned}
\sum_{i=1}^n d_i^2 &= \sum_{i=1}^n\left[R(x_i)-R(y_i)\right]^2\\
&= \sum_{i=1}^n R(x_i)^2+\sum_{i=1}^n R(y_i)^2-2\sum_{i=1}^n R(x_i)R(y_i)\\
&= \frac{2n(n+1)(2n+1)}{6}-2\sum_{i=1}^n R(x_i)R(y_i)
\end{aligned}$$

代入可得

$$\begin{aligned}
r_s &= \frac{\sum_{i=1}^{n} R(x_i) R(y_i) - n\left[(n+1)/2\right]^2}{n(n+1)(2n+1)/6 - n\left[(n+1)/2\right]^2} \\
&= \frac{n(n+1)(2n+1)/6 - \sum_{i=1}^{n} d_i^2/2 - n\left[(n+1)/2\right]^2}{n(n+1)(2n+1)/6 - n\left[(n+1)/2\right]^2} \\
&= 1 - \frac{\sum_{i=1}^{n} d_i^2/2}{n(n+1)(2n+1)/6 - n\left[(n+1)/2\right]^2} \\
&= 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}
\end{aligned}$$

第 10 章　多變量常態分配

10.1　多變量常態分配

定理 10.1. 若 $\mathbf{A}$ 為一 $p \times p$ 之對稱及正定矩陣, $\mathbf{Z}$ 為 $p \times 1$ 之隨機向量，則

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}\right] dz_1 \cdots dz_p = (2\pi)^{\frac{p}{2}} |\mathbf{A}|^{\frac{1}{2}}$$

證明.

若 $\mathbf{A}$ 為一對稱及正定，則必存在一矩陣 $\mathbf{P}$ 使得 $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \mathbf{I}$. 令 $\mathbf{y} = \mathbf{P}^{\mathrm{T}}\mathbf{z}$，則

$$J = \left|\frac{\partial \mathbf{z}}{\partial \mathbf{y}}\right| = |\mathbf{P}|$$

由於 $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \mathbf{I}$ 及 $|\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}| = |\mathbf{P}^{\mathrm{T}}|\,|\mathbf{A}|\,|\mathbf{P}| = 1$，故得 $|\mathbf{A}|^{-\frac{1}{2}} = |\mathbf{P}|$. 所以

$$\begin{aligned}
& \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}\right] dz_1 \cdots dz_p \\
= \; & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}\mathbf{y}\right] \times |\mathbf{A}|^{\frac{1}{2}}\, dy_1 \cdots dy_p \\
= \; & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{y}\right] \times |\mathbf{A}|^{\frac{1}{2}}\, dy_1 \cdots dy_p \\
= \; & |\mathbf{A}|^{-\frac{1}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\frac{\sum_{i=1}^{p} y_i^2}{2}\right] dy_1 \cdots dy_p \\
= \; & (2\pi)^{\frac{p}{2}} |\mathbf{A}|^{-\frac{1}{2}}
\end{aligned}$$

因此

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{|\mathbf{A}|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left[-\frac{1}{2}\mathbf{z}^{\mathrm{T}}\mathbf{A}\mathbf{z}\right] dz_1 \cdots dz_p = 1$$

多變量常態分配（Mutlivariate Normal Distribution）：$\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \cdots \ \mathbf{X}_p]^{\mathrm{T}}$ 有多變量常態分配以 $\mathbf{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ 表示之，其中 μ 為母體期望值, $\mathbf{\Sigma}$ 為母體共變異數矩陣, 機率密度函數為

$$f(x_1, x_2, \cdots, x_p) = \frac{1}{(2\pi)^{\frac{p}{2}} |\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu)^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{x} - \mu)\right],$$

其中

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

性質

1. $M_{\mathbf{X}}(t) = \exp\left(\mathbf{t}^{\mathrm{T}}\mu + \frac{1}{2}\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{t}\right)$

2. 若 $\mathbf{X} \sim \mathbf{N}(\mu, \mathbf{\Sigma})$，令 $\mathbf{Z} = \mathbf{\Sigma}^{-\frac{1}{2}} (\mathbf{X} - \mu)$，則 $\mathbf{Z} \sim \mathbf{N}(\mathbf{0}, \mathbf{I})$

3. 若 $\mathbf{Y} = \mathbf{C}^{\mathrm{T}}\mathbf{X}$，則 $\mathbf{Y} \sim \mathbf{N}(\mathbf{C}^{\mathrm{T}}\mu,\ \mathbf{C}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{C})$

4. 若 $\mathbf{Y} = \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{X} & \mathbf{B}^{\mathrm{T}}\mathbf{X} \end{bmatrix}^{\mathrm{T}}$，其中 $\mathbf{A}$ 及 $\mathbf{B}$ 分別為 $p \times 1$ 之矩陣，則

$$\mathbf{Y} \sim \mathbf{N}\left(\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mu \\ \mathbf{B}^{\mathrm{T}}\mu \end{bmatrix}, \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \end{bmatrix}\right)$$

5. 若 $\mathbf{X}^{(1)} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \cdots \ \mathbf{X}_k]^{\mathrm{T}} \sim \mathbf{N}(\mu_1, \mathbf{\Sigma}_{11})$; $\mathbf{X}^{(2)} = [\mathbf{X}_k \ \mathbf{X}_{k+1} \ \cdots \ \mathbf{X}_p]^{\mathrm{T}} \sim \mathbf{N}(\mu_2, \mathbf{\Sigma}_{22})$

6. 在給定 $\mathbf{X}^{(2)} = [\mathbf{X}_k \ \mathbf{X}_{k+1} \ \cdots \ \mathbf{X}_p]^{\mathrm{T}}$ 情況下，

$$\mathbf{X}^{(1)} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \cdots \ \mathbf{X}_k]^{\mathrm{T}} \sim \mathbf{N}\left(\mu_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \mu_2), \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\right)$$

7. $(\mathbf{X} - \mu)^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{X} - \mu) \sim \chi^2(p)$

8. 相關係數矩陣 $\mathbf{R}=\mathbf{D}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{D}$，其中

$$\mathbf{R}=\begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix},\ \mathbf{D}=\begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_p} \end{bmatrix}$$

證明.

1. 由定義可知

$$M(t)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left(\mathbf{t}^{\mathrm{T}}\mathbf{x}\right)\frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right]dx_1\cdots dx_p$$

其中

$$\begin{aligned}
& (\mathbf{x}-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)-2\mathbf{t}^{\mathrm{T}}\mathbf{x} \\
= & \mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}-\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}+\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-2\mathbf{t}^{\mathrm{T}}\mathbf{x} \\
= & \mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}-\mathbf{2x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-2\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\mathbf{t}+\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu \\
= & \mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}-\mathbf{2x}^{\mathrm{T}}\mathbf{\Sigma}^{-\mathbf{1}}(\mu+\mathbf{\Sigma}\mathbf{t})+(\mu+\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-\mathbf{1}}(\mu+\mathbf{\Sigma}\mathbf{t}) \\
& +\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-(\mu+\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-\mathbf{1}}(\mu+\mathbf{\Sigma}\mathbf{t}) \\
= & (\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})+\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-(\mu+\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-\mathbf{1}}(\mu+\mathbf{\Sigma}\mathbf{t}) \\
= & (\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})+\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mu-\mu^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\mathbf{t} \\
& -\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{\Sigma}^{-1}\mu-\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{\Sigma}^{-1}\mathbf{\Sigma}\mathbf{t} \\
= & (\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})-2\mu^{\mathrm{T}}\mathbf{t}-\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{t}
\end{aligned}$$

所以

$$\begin{aligned}
& M(t) \\
= & \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\exp\left[-\frac{1}{2}(\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu-\mathbf{\Sigma}\mathbf{t})+\mu^{\mathrm{T}}\mathbf{t}+\frac{1}{2}\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{t}\right]}{(2\pi)^{\frac{p}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}}dx_1\cdots dx_p
\end{aligned}$$

$$
\begin{aligned}
&= \exp\left(\mu^{\mathrm{T}}\mathbf{t}+\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\right)\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}-\mu-\boldsymbol{\Sigma}\mathbf{t}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\mu-\boldsymbol{\Sigma}\mathbf{t}\right)\right]}{(2\pi)^{\frac{p}{2}}\left|\boldsymbol{\Sigma}\right|^{\frac{1}{2}}}dx_1\cdots dx_p \\
&= \exp\left(\mu^{\mathrm{T}}t+\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\right)
\end{aligned}
$$

2.

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{x}-\mu\right)\right)\right] &= \exp\left(-\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu\right)\mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{x}\right)\right] \\
&= \exp\left(-\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu\right)\exp\left(\mu^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{t}+\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{t}\right) \\
&= \exp\left(\frac{1}{2}\mathbf{t}^{\mathrm{T}}\mathbf{t}\right)
\end{aligned}
$$

所以由動差生成函數可知，$\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{x}-\mu\right)\sim\mathbf{N}\left(\mathbf{0},\mathbf{I}\right)$

3.

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(ty\right)\right] &= \mathrm{E}\left[\exp\left(t\mathbf{C}^{\mathrm{T}}\mathbf{x}\right)\right] \\
&= \exp\left(\mu^{\mathrm{T}}t\mathbf{C}+\frac{1}{2}t\mathbf{C}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{C}t\right) \\
&= \exp\left(t\mathbf{C}^{\mathrm{T}}\mu+\frac{t^2\mathbf{C}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{C}}{2}\right)
\end{aligned}
$$

所以，$Y\sim\mathcal{N}\left(\mathbf{C}^{\mathrm{T}}\mu,\mathbf{C}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{C}\right)$

4. 假設 $\mathbf{t}=\begin{bmatrix} t_1 & t_2 \end{bmatrix}^{\mathrm{T}}$，則 $\mathbf{Y}$ 之動差生成函數為

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\mathbf{Y}\right)\right] &= E\left[\exp\left(\begin{bmatrix} t_1 & t_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{X} \\ \mathbf{B}^{\mathrm{T}}\mathbf{X} \end{bmatrix}\right)\right] \\
&= \mathrm{E}\left[\exp\left(t_1\mathbf{A}^{\mathrm{T}}\mathbf{X}+t_2\mathbf{B}^{\mathrm{T}}\mathbf{X}\right)\right] \\
&= \mathrm{E}\left[\exp\left[\left(t_1\mathbf{A}^{\mathrm{T}}+t_2\mathbf{B}^{\mathrm{T}}\right)\mathbf{X}\right]\right] \\
&= \exp\left[\left(t_1\mathbf{A}^{\mathrm{T}}+t_2\mathbf{B}^{\mathrm{T}}\right)\mu+\frac{1}{2}\left(t_1\mathbf{A}^{\mathrm{T}}+t_2\mathbf{B}^{\mathrm{T}}\right)\boldsymbol{\Sigma}\left(t_1\mathbf{A}+t_2\mathbf{B}\right)\right] \\
&= \exp\left\{\begin{bmatrix} t_1 & t_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mu \\ \mathbf{B}^{\mathrm{T}}\mu \end{bmatrix}+\frac{1}{2}\begin{bmatrix} t_1 & t_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{T}} \end{bmatrix}\boldsymbol{\Sigma}\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}\right\}
\end{aligned}
$$

$$= \exp\left\{\begin{bmatrix} t_1 & t_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mu \\ \mathbf{B}^{\mathrm{T}}\mu \end{bmatrix}+\frac{1}{2}\begin{bmatrix} t_1 & t_2 \end{bmatrix}\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \end{bmatrix}\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}\right\}$$

所以

$$\mathbf{Y} = \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{X} \\ \mathbf{B}^{\mathrm{T}}\mathbf{X} \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mu \\ \mathbf{B}^{\mathrm{T}}\mu \end{bmatrix}, \begin{bmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{A}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \\ \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{A} & \mathbf{B}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{B} \end{bmatrix}\right)$$

5. 令 $\mathbf{t}_1 = [t_1\ t_2\ \cdots\ t_k\ ]^{\mathrm{T}}$ 及 $\mathbf{t}_2 = [t_{k+1}\ t_{k+2}\cdots t_p]^{\mathrm{T}}$，故得 $\mathbf{t} = [\mathbf{t}_1\ \mathbf{t}_2]^{\mathrm{T}}$。首先令 $\mathbf{t}_2 = \mathbf{0}$，則

$$\begin{aligned}
\mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\mathbf{x}\right)\right] &= \exp\left[\mu^{\mathrm{T}}\mathbf{t} - \mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{t}\right] \\
&= \exp\left[\begin{bmatrix} \mu_1^{\mathrm{T}} & \mu_2^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} \mathbf{t}_1 \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{t}_1^{\mathrm{T}} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}\begin{bmatrix} \mathbf{t}_1 \\ \mathbf{0} \end{bmatrix}\right] \\
&= \exp\left[\mu_1^{\mathrm{T}}\mathbf{t}_1 + \begin{bmatrix} \mathbf{t}_1^{\mathrm{T}}\mathbf{\Sigma}_{11} & \mathbf{t}_1^{\mathrm{T}}\mathbf{\Sigma}_{12} \end{bmatrix}\begin{bmatrix} \mathbf{t}_1 \\ \mathbf{0} \end{bmatrix}\right] \\
&= \exp\left[\mu_1^{\mathrm{T}}\mathbf{t}_1 + \mathbf{t}_1^{\mathrm{T}}\mathbf{\Sigma}_{11}\mathbf{t}_1\right]
\end{aligned}$$

所以 $\mathbf{X}^{(1)} \sim \mathbf{N}\left(\mu_1, \mathbf{\Sigma}_{11}\right)$; 同理可證 $\mathbf{X}^{(2)} \sim \mathbf{N}\left(\mu_2, \mathbf{\Sigma}_{22}\right)$。

6. 由條件分配之定義

$$\begin{aligned}
&\frac{f\left(x_1, x_2, \cdots, x_k\right)}{g\left(x_{k+1}, x_{k+2}, \cdots, x_p\right)} \\
= &\frac{\frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(\mathbf{x}-\mu\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\mathbf{x}-\mu\right)\right)}{\frac{1}{(2\pi)^{\frac{p-k}{2}}|\mathbf{\Sigma}_{22}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}}\mathbf{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right)} \\
= &(2\pi)^{\frac{-k}{2}}\frac{|\mathbf{\Sigma}_{22}|^{\frac{1}{2}}}{|\mathbf{\Sigma}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\mu\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\mathbf{x}-\mu\right)+\frac{1}{2}\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}}\mathbf{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right]
\end{aligned}$$

由於

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

假設

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix}$$

由於

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}=\begin{bmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{W}_{11} & \mathbf{W}_{12}\\ \mathbf{W}_{21} & \mathbf{W}_{22}\end{bmatrix}=\begin{bmatrix}\mathbf{I}_{11} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{22}\end{bmatrix}$$

故可求得

$$\begin{cases}\boldsymbol{\Sigma}_{11}\mathbf{W}_{11}+\boldsymbol{\Sigma}_{12}\mathbf{W}_{21}=\mathbf{I}_{11}\\ \boldsymbol{\Sigma}_{11}\mathbf{W}_{12}+\boldsymbol{\Sigma}_{12}\mathbf{W}_{22}=\mathbf{0}\\ \boldsymbol{\Sigma}_{21}\mathbf{W}_{11}+\boldsymbol{\Sigma}_{22}\mathbf{W}_{21}=\mathbf{0}\\ \boldsymbol{\Sigma}_{21}\mathbf{W}_{12}+\boldsymbol{\Sigma}_{22}\mathbf{W}_{22}=\mathbf{I}_{22}\end{cases}$$

由第三式可求得 $\mathbf{W}_{21}=-\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{W}_{11}$, 代入第一式可得 $\mathbf{W}_{11}=\left(\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1}$。由於 $\mathbf{W}_{21}=\mathbf{W}_{12}^{\mathrm{T}}$，所以 $\mathbf{W}_{22}=\boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{I}_{22}+\boldsymbol{\Sigma}_{21}\mathbf{W}_{11}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{22}^{-1}\right)=\boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{W}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$。故

$$\boldsymbol{\Sigma}^{-1}=\begin{bmatrix}\mathbf{W}_{11} & -\mathbf{W}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{W}_{11} & \boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{W}_{11}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\end{bmatrix},$$

其中 $\mathbf{W}_{11}=\left(\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)^{-1}$。又

$$\begin{aligned}|\boldsymbol{\Sigma}| &= \begin{vmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{vmatrix}\\ &= \begin{vmatrix}\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{2}\boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{12}\\ \mathbf{0} & \boldsymbol{\Sigma}_{22}\end{vmatrix}\\ &= \begin{vmatrix}\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{2}\boldsymbol{\Sigma}_{21} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Sigma}_{22}\end{vmatrix}\\ &= |\boldsymbol{\Sigma}_{22}|\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{2}\boldsymbol{\Sigma}_{21}\right|\\ &= |\boldsymbol{\Sigma}_{22}|\left|\mathbf{W}_{11}^{-1}\right|\end{aligned}$$

及

$$\begin{aligned}&(\mathbf{x}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mu)-(\mathbf{x}_2-\mu_2)^{\mathrm{T}}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\mu_2)\\ =\ &(\mathbf{x}_1-\mu_1)^{\mathrm{T}}\mathbf{W}_{11}(\mathbf{x}_1-\mu_1)-(\mathbf{x}_2-\mu_2)^{\mathrm{T}}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{W}_{11}(\mathbf{x}_1-\mu_1)\end{aligned}$$

$$
\begin{aligned}
& -\left(\mathbf{x}_1-\mu_1\right)^{\mathrm{T}} \mathbf{W}_{11} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)-\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right) \\
& +\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} \mathbf{W}_{11} \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{22}^{-1}\right)\left(\mathbf{x}_2-\mu_2\right) \\
= & \left(\mathbf{x}_1-\mu_1\right)^{\mathrm{T}} \mathbf{W}_{11}\left(\mathbf{x}_1-\mu_1\right)-\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} \mathbf{W}_{11}\left(\mathbf{x}_1-\mu_1\right) \\
& -\left(\mathbf{x}_1-\mu_1\right)^{\mathrm{T}} \mathbf{W}_{11} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)+\left(\mathbf{x}_2-\mu_2\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} \mathbf{W}_{11} \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right) \\
= & \left[\left(\mathbf{x}_1-\mu_1\right)-\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right]^{\mathrm{T}} \mathbf{W}_{11}\left[\left(\mathbf{x}_1-\mu_1\right)-\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right]
\end{aligned}
$$

所以

$$
\begin{aligned}
& f\left(x_1, x_2, \cdots, x_k \mid x_{k+1}, x_{k+2}, \cdots, x_p\right) \\
= & \frac{\exp \left\{-\frac{1}{2}\left[\left(\mathbf{x}_1-\mu_1\right)-\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right]^{\mathrm{T}} \mathbf{W}_{11}\left[\left(\mathbf{x}_1-\mu_1\right)-\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right)\right]\right\}}{(2 \pi)^{\frac{k}{2}}\left|\mathbf{W}_{11}^{-1}\right|^{\frac{1}{2}}}
\end{aligned}
$$

即

$$
\mathbf{X}_1 \mid \mathbf{X}_2=\mathbf{x}_2 \sim \mathbf{N}\left(\mu_1+\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}\left(\mathbf{x}_2-\mu_2\right), \boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}\right)
$$

7. 由於 $\boldsymbol{\Sigma}$ 為一對稱及正定，則必存在一矩陣 $\mathbf{P}$ 使得

$$
\mathbf{P}^{\mathrm{T}} \boldsymbol{\Sigma} \mathbf{P}=\boldsymbol{\Lambda} \text{ 及} \mathbf{P}^{\mathrm{T}} \mathbf{P}=\mathbf{I},
$$

其中 $\mathbf{P}=\left[\begin{array}{llll}\mathbf{P}_1 & \mathbf{P}_2 & \cdots & \mathbf{P}_p\end{array}\right]$。令 $\mathbf{Y}=\mathbf{P}^{\mathrm{T}}(\mathbf{X}-\mu)$，則

$$
\mathrm{E}\left(\mathbf{Y}_i\right)=\mathrm{E}\left[\mathbf{P}^{\mathrm{T}}(\mathbf{X}-\mu)\right]=\mathbf{P}^{\mathrm{T}} \mathrm{E}(\mathbf{X}-\mu)=\mathbf{0}
$$

及

$$
\begin{aligned}
\operatorname{Var}\left(\mathbf{Y}_i\right) & =\mathrm{E}\left[\mathbf{Y} \mathbf{Y}^{\mathrm{T}}\right]=E\left[\mathbf{P}_i^{\mathrm{T}}(\mathbf{X}-\mu)(\mathbf{X}-\mu)^{\mathrm{T}} \mathbf{P}_i\right] \\
& =\mathrm{E}\left[\mathbf{P}_i^{\mathrm{T}}(\mathbf{X}-\mu)(\mathbf{X}-\mu)^{\mathrm{T}} \mathbf{P}_i\right] \\
& =\mathbf{P}_i^{\mathrm{T}} \boldsymbol{\Sigma} \mathbf{P}_i \\
& =\lambda_i
\end{aligned}
$$

所以

$$\begin{aligned}(\mathbf{X}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mu) &= (\mathbf{X}-\mu)^{\mathrm{T}}\mathbf{P}\boldsymbol{\Lambda}^{-1}\mathbf{P}^{\mathrm{T}}(\mathbf{X}-\mu)=\mathbf{Y}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\mathbf{Y} \\ &= \sum_{i=1}^{p}\frac{Y_i^2}{\lambda_i}=\sum_{i=1}^{p}\left(\frac{Y_i-0}{\sqrt{\lambda_i}}\right)^2=\sum_{i=1}^{p}Z_i^2\end{aligned}$$

由卡方分配之可加性可知 $(\mathbf{X}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mu)\sim\chi^2(p)$

8.

$$\begin{aligned}\mathbf{D}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{D} &= \begin{bmatrix}\frac{1}{\sigma_1} & 0 & \cdots & 0\\ 0 & \frac{1}{\sigma_2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sigma_p}\end{bmatrix}\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp}\end{bmatrix}\begin{bmatrix}\frac{1}{\sigma_1} & 0 & \cdots & 0\\ 0 & \frac{1}{\sigma_2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sigma_p}\end{bmatrix} \\ &= \begin{bmatrix}\frac{\sigma_{11}}{\sigma_1} & \frac{\sigma_{12}}{\sigma_1} & \cdots & \frac{\sigma_{1p}}{\sigma_1}\\ \frac{\sigma_{21}}{\sigma_2} & \frac{\sigma_{22}}{\sigma_2} & \cdots & \frac{\sigma_{2p}}{\sigma_2}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\sigma_{p1}}{\sigma_p} & \frac{\sigma_{p2}}{\sigma_p} & \cdots & \frac{\sigma_{pp}}{\sigma_p}\end{bmatrix}\begin{bmatrix}\frac{1}{\sigma_1} & 0 & \cdots & 0\\ 0 & \frac{1}{\sigma_2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{\sigma_p}\end{bmatrix} \\ &= \begin{bmatrix}\frac{\sigma_{11}}{\sigma_1\sigma_1} & \frac{\sigma_{12}}{\sigma_1\sigma_2} & \cdots & \frac{\sigma_{1p}}{\sigma_1\sigma_p}\\ \frac{\sigma_{21}}{\sigma_2\sigma_1} & \frac{\sigma_{22}}{\sigma_2\sigma_2} & \cdots & \frac{\sigma_{2p}}{\sigma_2\sigma_p}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\sigma_{p1}}{\sigma_p\sigma_1} & \frac{\sigma_{p2}}{\sigma_p\sigma_2} & \cdots & \frac{\sigma_{pp}}{\sigma_p\sigma_p}\end{bmatrix} \\ &= \begin{bmatrix}1 & r_{12} & \cdots & r_{1p}\\ r_{21} & 1 & \cdots & r_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ r_{p1} & r_{p2} & \cdots & 1\end{bmatrix}\end{aligned}$$

例 **10–1** $\mathbf{X}=\begin{bmatrix}X_1 & X_2 & X_3\end{bmatrix}^{\mathrm{T}}$ 服從多變量常態分配且其平均數及共變異矩陣分別為 $\mu=\begin{bmatrix}6 & 4 & 2\end{bmatrix}$ 及

$$\boldsymbol{\Sigma} = \begin{bmatrix} 16 & 6 & 0 \\ 6 & 25 & 0 \\ 0 & 0 & 64 \end{bmatrix}$$

若 $Y_1 = 2X_1 + 3X_2 + X_3$ 及 $Y_2 = 4X_1 + X_3$. 試求 $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}^{\mathrm{T}}$ 之機率密度函數。

解：

令 $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}^{\mathrm{T}}$，所以

$$Y_1 = 2X_1 + 3X_2 + X_3 = \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

及

$$Y_2 = 4X_1 + X_3 = \begin{bmatrix} 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

由 **11.1 性質 (4)** 可知

$$\mu_y = \begin{bmatrix} \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 26 \\ 26 \end{bmatrix}$$

及

$$\boldsymbol{\Sigma}_y = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 16 & 6 & 0 \\ 6 & 25 & 0 \\ 0 & 0 & 64 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 425 & 264 \\ 264 & 320 \end{bmatrix}$$

因此

$$\mathbf{Y} \sim \mathbf{N}\left(\begin{bmatrix} 26 \\ 26 \end{bmatrix}, \begin{bmatrix} 425 & 264 \\ 264 & 320 \end{bmatrix} \right)$$

其機率密度函數為

$$\begin{aligned} f(y_1,y_2) &= \frac{\exp\left\{-\frac{1}{2}\begin{bmatrix} y_1-26 & y_2-26 \end{bmatrix}^{\mathrm{T}}\begin{bmatrix} 425 & 264 \\ 264 & 320 \end{bmatrix}^{-1}\begin{bmatrix} y_1-26 \\ y_2-26 \end{bmatrix}\right\}}{(2\pi)^{\frac{3}{2}}\begin{vmatrix} 425 & 264 \\ 264 & 320 \end{vmatrix}^{\frac{1}{2}}}, \\ -\infty &< y_1<\infty,\ -\infty<y_2<\infty \end{aligned}$$

10.2 μ 及 Σ 之估計

定理 10.2. 若 $\mathbf{\Sigma}$ 與 $\mathbf{B}$ 為一 $p\times p$ 的對稱正定矩陣，則對任一正數 b 使得

$$\frac{1}{|\mathbf{\Sigma}|^b}\exp\left[-\operatorname{trace}\left(\mathbf{\Sigma}^{-1}\mathbf{B}\right)\right]\le\frac{1}{|\mathbf{B}|^b}(2b)^{pb}\exp[-pb]$$

證明.

由於 $\mathbf{\Sigma}$ 與 $\mathbf{B}$ 均為 $p\times p$ 的對稱正定矩陣，所以 $\operatorname{trace}(\mathbf{\Sigma}^{-1}\mathbf{B})=\operatorname{trace}\left(\mathbf{B}^{\frac{1}{2}}\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\right)=\operatorname{trace}\left(\mathbf{P\Lambda P}^{\mathrm{T}}\right)=\operatorname{trace}(\mathbf{\Lambda})=\sum_{i=1}^{p}\lambda_i$，其中 λ_i 為 $\mathbf{B}^{\frac{1}{2}}\mathbf{\Sigma}\mathbf{B}^{\frac{1}{2}}$ 之特徵根。又

$$\left|\mathbf{B}^{\frac{1}{2}}\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\right|=\left|\mathbf{B}^{\frac{1}{2}}\right|\left|\mathbf{\Sigma}^{-1}\right|\left|\mathbf{B}^{\frac{1}{2}}\right|=|\mathbf{B}|^{\frac{1}{2}}|\mathbf{\Sigma}|^{-1}|\mathbf{B}|^{\frac{1}{2}}=\frac{|\mathbf{B}|}{|\mathbf{\Sigma}|}$$

可求得

$$\frac{1}{|\mathbf{\Sigma}|}=\frac{\left|\mathbf{B}^{\frac{1}{2}}\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\right|}{|\mathbf{B}|}$$

因此

$$\begin{aligned} \frac{1}{|\mathbf{\Sigma}|^b}\exp\left[-\frac{\operatorname{trace}(\mathbf{\Sigma}^{-1}\mathbf{B})}{2}\right] &= \frac{\left|\mathbf{B}^{\frac{1}{2}}\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\right|^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right] \\ &= \frac{\left|\mathbf{P\Lambda P}^{\mathrm{T}}\right|^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right] \\ &= \frac{\left|\mathbf{\Lambda P P}^{\mathrm{T}}\right|^b}{|\mathbf{B}|^b}\exp\left[\frac{-\sum_{i=1}^{p}\lambda_i}{2}\right]=\frac{|\mathbf{\Lambda}|^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{\left|\prod_{i=1}^{n}\lambda_i\right|^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right] \\
&= \frac{\prod_{i=1}^{n}\lambda_i^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right]
\end{aligned}
$$

由於 $f(x)=x^b e^{-x/2}$ 及 $f'(x)=bx^{b-1}e^{-x}-\frac{1}{2}x^b e^{-x}=\left(b-\frac{x}{2}\right)x^{b-1}e^{-x}$，因此對於所有的 $x>0$, 函數 $f(x)$ 之最大值必發生在 $x=2b$ 時，所以

$$
\begin{aligned}
\frac{1}{|\boldsymbol{\Sigma}|^b}\exp\left[-\text{trace}\left(\frac{\boldsymbol{\Sigma}^{-1}\mathbf{B}}{2}\right)\right] &= \frac{\prod_{i=1}^{n}\lambda_i^b}{|\mathbf{B}|^b}\exp\left[-\frac{\sum_{i=1}^{p}\lambda_i}{2}\right] \\
&\le \frac{\prod_{i=1}^{n}(2b)^b}{|\mathbf{B}|^b}\exp(-bp) \\
&= \frac{1}{|\mathbf{B}|^b}(2b)^{bp}\exp(-bp)
\end{aligned}
$$

故當

$$
\frac{1}{|\boldsymbol{\Sigma}|}=\frac{(2b)^p}{|\mathbf{B}|}=\frac{1}{\left|\frac{1}{2b}\mathbf{B}\right|}
$$

時，$\frac{1}{|\boldsymbol{\Sigma}|^b}\exp\left[-\text{trace}\left(\boldsymbol{\Sigma}^{-1}\mathbf{B}\right)\right]$ 有最大值，即等式發生於

$$
\boldsymbol{\Sigma}=\frac{1}{2b}\mathbf{B}
$$

假設 $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_n$ 為抽自多變量常態分配之一組樣本數為 n 的隨機樣本, 其母體平均數與共變異矩陣分別為 μ 及 $\boldsymbol{\Sigma}$ ，故機率密度函數為

$$
f(\mathbf{X})=\frac{1}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)\right]
$$

因此其概似函數 (Likelihood function) 為

$$
\begin{aligned}
L(\mathbf{X},\mu,\boldsymbol{\Sigma}) &= \prod_{i=1}^{n}\frac{1}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)\right] \\
&= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)\right]
\end{aligned}
$$

由於

$$
\begin{aligned}
\sum_{i=1}^{n}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu) &= \sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}}+\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}}+\overline{\mathbf{x}}-\mu) \\
&= \sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})+\sum_{i=1}^{n}(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu) \\
&\quad +\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu)+\sum_{i=1}^{n}(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}}) \\
&= \sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})+\sum_{i=1}^{n}(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu) \\
&\quad +2(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}}) \\
&= \sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})+\sum_{i=1}^{n}(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu) \\
&= \sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})+n(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu)
\end{aligned}
$$

所以概似函數可改寫為

$$
\begin{aligned}
L(\mathbf{X},\mu,\boldsymbol{\Sigma}) &= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)\right] \\
&= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})-\frac{1}{2}n(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu)\right]
\end{aligned}
$$

由於 $n(\overline{\mathbf{x}}-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\mu)\geq 0$, 欲使得 $L(\mathbf{X},\mu,\boldsymbol{\Sigma})$ 有極大值，則下次必須成立

$$\mu=\overline{\mathbf{x}}$$

當 $\mu=\overline{\mathbf{x}}$ 時

$$
\begin{aligned}
L(\mathbf{X},\mu,\boldsymbol{\Sigma}) &= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})\right] \\
&= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\mathrm{trace}\left[(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\overline{\mathbf{x}})\right]\right] \\
&= \frac{1}{(2\pi)^{\frac{np}{2}}|\boldsymbol{\Sigma}|^{\frac{n}{2}}}\exp\left[-\frac{1}{2}\mathrm{trace}\left[\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\right]\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{1}{(2\pi)^{\frac{np}{2}}\left|\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\right|^{\frac{n}{2}}}\left(2\times\frac{n}{2}\right)^{\frac{np}{2}}\exp\left(-\frac{np}{2}\right)\\
&= \frac{1}{(2\pi)^{\frac{np}{2}}\left|\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}\right|^{\frac{n}{2}}}n^{\frac{np}{2}}\exp\left(-\frac{np}{2}\right)
\end{aligned}
$$

由定理 **11.2** 可知，

$$
\boldsymbol{\Sigma}=\frac{1}{2b}\mathbf{B}=\frac{1}{2\times\frac{n}{2}}\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}=\frac{\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}}{n}
$$

此外, 我們亦可利用矩陣之微分求得 μ 及 $\boldsymbol{\Sigma}$ 之 MLE, 首先將概似函數左右取對數得

$$
\ln L(\mathbf{X},\mu,\boldsymbol{\Sigma})=-\frac{np}{2}\ln(2\pi)-\frac{n}{2}\ln|\boldsymbol{\Sigma}|-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)
$$

則上式最大值的必要條件為

$$
\frac{\partial\ln L(\mathbf{X},\mu,\boldsymbol{\Sigma})}{\partial\mu}=\sum_{i=1}^{n}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i-\mu)=\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)=0
$$

及

$$
\frac{\partial\ln L(\mathbf{X},\mu,\boldsymbol{\Sigma})}{\partial\boldsymbol{\Sigma}}=-\frac{n}{2}\boldsymbol{\Sigma}^{-1}+\frac{1}{2}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i-\mu)^{\mathrm{T}}(\mathbf{x}_i-\mu)\boldsymbol{\Sigma}^{-1}=0
$$

由第一式可得

$$
\widehat{\mu}=\frac{\sum_{i=1}^{n}\mathbf{x}_i}{n}=\overline{\mathbf{x}}
$$

將 $\widehat{\mu}$ 代入第二式得

$$
\widehat{\boldsymbol{\Sigma}}=\frac{\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}(\mathbf{x}_i-\overline{\mathbf{x}})}{n}=\frac{n-1}{n}\frac{\sum_{i=1}^{n}(\mathbf{x}_i-\overline{\mathbf{x}})^{\mathrm{T}}(\mathbf{x}_i-\overline{\mathbf{x}})}{n-1}=\frac{n-1}{n}\mathbf{S},
$$

其中 $\overline{\mathbf{x}}$ 為樣本平均數向量, $\mathbf{S}$ 為樣本變異數矩陣。又 $\overline{\mathbf{x}}$ 的動差生成函數為

$$
\begin{aligned}
\mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\overline{\mathbf{X}}\right)\right] &= \mathrm{E}\left[\exp\left(\mathbf{t}^{\mathrm{T}}\frac{\sum_{i=1}^{n}\mathbf{X}_i}{n}\right)\right]=\mathrm{E}\left[\exp\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{t}^{\mathrm{T}}\mathbf{X}_i\right)\right]\\
&= \mathrm{E}\left[\exp\left(\frac{\mathbf{t}^{\mathrm{T}}}{n}\mathbf{X}\right)\right]^{n}
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\exp\left\{\frac{\mathbf{t}^{\mathrm{T}}}{n}\mu+\frac{\mathbf{t}^{\mathrm{T}}}{n}\mathbf{\Sigma}\frac{\mathbf{t}}{n}\right\}\right]^n \\
&= \left[\exp\left\{\frac{1}{n}\mathbf{t}^{\mathrm{T}}\mu+\frac{1}{n^2}\mathbf{t}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{t}\right\}\right]^n \\
&= \exp\left\{\mathbf{t}^{\mathrm{T}}\mu+\mathbf{t}^{\mathrm{T}}\frac{\mathbf{\Sigma}}{n}\mathbf{t}\right\}
\end{aligned}
$$

所以由上式可知，

$$
\overline{\mathbf{X}}\sim \mathbf{N}\left(\mu,\frac{\mathbf{\Sigma}}{n}\right)
$$

此外,

$$
\begin{aligned}
\sum_{i=1}^{n}(\mathbf{X}_i-\mu)(\mathbf{X}_i-\mu)^{\mathrm{T}} &= \sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}+\overline{\mathbf{X}}-\mu\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}+\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}} \\
&= \sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)^{\mathrm{T}}+\sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}} \\
&\quad +\sum_{i=1}^{n}\left(\overline{\mathbf{X}}-\mu\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)^{\mathrm{T}}+\sum_{i=1}^{n}\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}} \\
&= \sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)^{\mathrm{T}}+\sum_{i=1}^{n}\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}} \\
&= \sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)^{\mathrm{T}}+n\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}}
\end{aligned}
$$

所以樣本共變異矩陣之期望值為

$$
\begin{aligned}
\mathrm{E}(\mathbf{S}) &= \frac{\mathbf{E}\left[\sum_{i=1}^{n}\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)\left(\mathbf{X}_i-\overline{\mathbf{X}}\right)^{\mathrm{T}}\right]}{n-1} \\
&= \frac{\mathrm{E}\left[\sum_{i=1}^{n}(\mathbf{X}_i-\mu)(\mathbf{X}_i-\mu)^{\mathrm{T}}-n\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}}\right]}{n-1} \\
&= \frac{\mathrm{E}\left[\sum_{i=1}^{n}(\mathbf{X}_i-\mu)(\mathbf{X}_i-\mu)^{\mathrm{T}}\right]-\mathrm{E}\left[n\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^{\mathrm{T}}\right]}{n-1} \\
&= \frac{\sum_{i=1}^{n}\mathrm{E}\left[(\mathbf{X}_i-\mu)(\mathbf{X}_i-\mu)^{\mathrm{T}}\right]-n\times\mathbf{\Sigma}/n}{n-1} \\
&= \frac{\sum_{i=1}^{n}\mathbf{\Sigma}-\mathbf{\Sigma}}{n-1} \\
&= \mathbf{\Sigma}
\end{aligned}
$$

由以上推導可知，$\overline{\mathbf{X}}$ 與 $\mathbf{S}$ 分別為 μ 及 $\boldsymbol{\Sigma}$ 之不偏估計式。

10.3 矩陣不等式及應用

定理 10.3. 若對所有 x 均有 $f(x) = ax^2 + bx + c > 0,\ a > 0$，則 $b^2 - 4ac \le 0$

證明.

$f'(x) = 2ax + b$ 及 $f''(x) = 2a > 0$，故函數 $f(x)$ 有極小值且極小值為

$$a \times \left(-\frac{b}{2a}\right)^2 + b \times \left(-\frac{b}{2a}\right) + c = -\frac{b^2 - 4ac}{4a}$$

由於 $f(x) > 0$，所以 $b^2 - 4ac \le 0$

定理 10.4. 若 $\mathbf{a}$, $\mathbf{b}$ 為任意兩向量，則

$$\left(\mathbf{a}^{\mathrm{T}}\mathbf{b}\right)^2 \le \left(\mathbf{a}^{\mathrm{T}}\mathbf{a}\right)\left(\mathbf{b}^{\mathrm{T}}\mathbf{b}\right)$$

證明.

令 $\mathbf{a} = x\mathbf{b}$，則

$$(\mathbf{a} - x\mathbf{b})^{\mathrm{T}}(\mathbf{a} - x\mathbf{b}) = \left(\mathbf{a}^{\mathrm{T}}\mathbf{a}\right) - 2x\mathbf{a}^{\mathrm{T}}\mathbf{b} + x^2\mathbf{b}^{\mathrm{T}}\mathbf{b} \ge 0$$

由定理10.4得 $(\mathbf{a} - x\mathbf{b})^{\mathrm{T}}(\mathbf{a} - x\mathbf{b})$ 大於 0 的必要條件為

$$\left(2\mathbf{a}^{\mathrm{T}}\mathbf{b}\right)^2 - 4\left(\mathbf{b}^{\mathrm{T}}\mathbf{b}\right)\left(\mathbf{a}^{\mathrm{T}}\mathbf{a}\right) \le 0$$

化簡得 $(\mathbf{a}^{\mathrm{T}}\mathbf{b})^2 - (\mathbf{b}^{\mathrm{T}}\mathbf{b})(\mathbf{a}^{\mathrm{T}}\mathbf{a}) \le 0$

定理 10.5. 若 $\mathbf{a}$, $\mathbf{b}$ 為 $p \times 1$ 之任意兩向量, $\mathbf{B}$ 為一 $p \times p$ 之正定矩陣，則

$$\left(\mathbf{a}^{\mathrm{T}}\mathbf{b}\right)^2 \le \left(\mathbf{a}^{\mathrm{T}}\mathbf{B}\mathbf{a}\right)\left(\mathbf{b}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{b}\right)$$

證明.

令 $\mathbf{x} = \mathbf{B}^{\frac{1}{2}}\mathbf{a}$ 及 $\mathbf{y} = \mathbf{B}^{-\frac{1}{2}}\mathbf{b}$，由定理 **11.4** 可得 $\left(\mathbf{a}^{\mathrm{T}}\mathbf{b}\right)^2 \leq \left(\mathbf{a}^{\mathrm{T}}\mathbf{B}\mathbf{a}\right)\left(\mathbf{b}^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{b}\right)$.

定理 10.6. 假設 $E\mathbf{X} = \mu$ 及 $\mathrm{Cov}(\mathbf{X}) = \mathbf{\Sigma}$ 且 $\mathbf{\Sigma}$ 為一正定矩陣，其特徵值 $\lambda_1 > \lambda_2 > \cdots > \lambda_p$，則當 $\mathbf{a}$ 為 λ_1 所對應之單位特徵向量 $\mathbf{e}_1$ 時，

$$\max_{\mathbf{a}\neq\mathbf{0}} \frac{\mathbf{a}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \lambda_1$$

當 $\mathbf{a}$ 為 λ_p 所對應之單位特徵向量 $\mathbf{e}_p$ 時，

$$\min_{\mathbf{a}\neq\mathbf{0}} \frac{\mathbf{a}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \lambda_p$$

證明.

令 $\mathbf{y} = \mathbf{P}^T\mathbf{a}$，則

$$\frac{\mathbf{a}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \frac{\mathbf{a}^{\mathrm{T}}\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{\mathrm{T}}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{P}\mathbf{P}^{\mathrm{T}}\mathbf{a}} = \frac{\mathbf{y}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{y}}{\mathbf{y}^{\mathrm{T}}\mathbf{y}} = \frac{\sum_{i=1}^{p}\lambda_i y_i^2}{\sum_{i=1}^{p} y_i^2} \leq \frac{\lambda_1\sum_{i=1}^{p} y_i^2}{\sum_{i=1}^{p} y_i^2} = \lambda_1$$

當 $\mathbf{a} = \mathbf{e}_1$ 時，

$$\mathbf{y} = \mathbf{P}^T\mathbf{a} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_p^T \end{bmatrix} \mathbf{e}_1 = \begin{bmatrix} \mathbf{e}_1^T\mathbf{e}_1 \\ \mathbf{e}_2^T\mathbf{e}_1 \\ \vdots \\ \mathbf{e}_p^T\mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

所以

$$\frac{\mathbf{a}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \frac{\mathbf{e}_1^T\mathbf{\Sigma}\mathbf{e}_1}{\mathbf{e}_1^T\mathbf{e}_1} = \lambda_1$$

同理可證當 $\mathbf{a}$ 等於 $\mathbf{e}_p$ 時

$$\min_{\mathbf{a}\neq\mathbf{0}} \frac{\mathbf{a}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}} = \lambda_p$$

定理 10.7. 假設 $E\mathbf{X}=\mu$ 及 $\mathrm{Cov}(\mathbf{X})=\mathbf{\Sigma}$ 且 $\mathbf{\Sigma}$ 為一正定矩陣，其特徵值 $\lambda_1>\lambda_2>\cdots>\lambda_p$。令

$$\mathbf{Y}=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}=\begin{bmatrix} \mathbf{e}_1^T\mathbf{X} \\ \mathbf{e}_2^T\mathbf{X} \\ \vdots \\ \mathbf{e}_p^T\mathbf{X} \end{bmatrix}=\mathbf{P}^T\mathbf{X},$$

$\mathbf{e}_i$ 為 λ_i 所對應之單位特徵向量，則

1. $\mathrm{Cov}(\mathbf{Y})=\mathbf{\Lambda}$

2. $\mathrm{trace}(\mathbf{\Sigma})=\sum_{i=1}^p\lambda_i$

3. $r_{Y_i,X_j}=\mathbf{e}_{ij}\sqrt{\lambda_i}/\sigma_{X_i}$，其中 $\mathbf{e}_{ij}$ 表第 i 個單位特徵向量之第 j 行的元素

4. $\mathbf{Y}$ 與 $\mathbf{X}$ 之相關係數矩陣為

$$\mathbf{R}_{\mathbf{YX}}=\begin{bmatrix} \mathbf{I} & \mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{D}^{-\frac{1}{2}} \\ \mathbf{D}^{-\frac{1}{2}}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{P}^{\mathrm{T}} & \mathbf{I} \end{bmatrix},$$

其中

$$\mathbf{D}=\begin{bmatrix} \frac{1}{\sigma_{X_1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_{X_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_{X_p}} \end{bmatrix}$$

證明.

1. 由 **11.1** 性質 **(2)** 可知，$\mathrm{Cov}\left(\mathbf{Y}\right)=\mathrm{Cov}\left(\mathbf{P}^{\mathrm{T}}\mathbf{X}\right)=\mathbf{P}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{P}=\mathbf{\Lambda}$

2. $\mathrm{trace}\left(\mathbf{\Sigma}\right)=\mathrm{trace}\left(\mathbf{P}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{P}\right)=\mathrm{trace}\left(\mathbf{\Lambda}\mathbf{P}\mathbf{P}^{\mathrm{T}}\right)=\mathrm{trace}\left(\mathbf{\Lambda}\right)=\sum_{i=1}^p\lambda_i$

3. 令

$$d = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \text{第} j \text{ 行}$$

所以

$$\begin{aligned}
\operatorname{Cov}(Y_i, X_j) &= \operatorname{Cov}\left(\mathbf{e}_i^{\mathrm{T}}\mathbf{X}, \mathbf{d}^{\mathrm{T}}\mathbf{X}\right) = \mathbf{e}_i^{\mathrm{T}}\mathbf{\Sigma}\mathbf{d} \\
&= \mathbf{e}_i^{\mathrm{T}} \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_i & \cdots & \mathbf{e}_p \end{bmatrix} \Lambda \begin{bmatrix} \mathbf{e}_{11} & \mathbf{e}_{21} & & \mathbf{e}_{p1} \\ \mathbf{e}_{12} & \mathbf{e}_{21} & & \mathbf{e}_{p1} \\ \vdots & & & \\ \mathbf{e}_{1p} & \mathbf{e}_{2p} & & \mathbf{e}_{pp} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & \cdots & \lambda_i & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{e}_{1j} \\ \vdots \\ \mathbf{e}_{ij} \\ \vdots \\ \mathbf{e}_{pj} \end{bmatrix} \\
&= \mathbf{e}_{ij}\lambda_i
\end{aligned}$$

及

$$r_{Y_i,X_j} = \frac{\operatorname{Cov}(Y_i, X_j)}{\sigma_{Y_i}\sigma_{X_j}} = \frac{\mathbf{e}_{ij}\lambda_i}{\sqrt{\lambda_i}\sigma_{X_j}} = \frac{\mathbf{e}_{ij}\sqrt{\lambda_i}}{\sigma_{X_j}}$$

4. 由於 $r_{Y_iY_i} = 1$, $r_{X_iX_i} = 1$ 及 $r_{Y_i,X_j} = \mathbf{e}_{ij}\sqrt{\lambda_i}/\sigma_{X_j}$，又 Y_i 與 Y_j 相互獨立，故可推得

$$\begin{bmatrix} \mathbf{I} & \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{P}\mathbf{D}^{-\frac{1}{2}} \\ \mathbf{D}^{-\frac{1}{2}}\mathbf{P}^{\mathrm{T}}\mathbf{\Lambda}^{\frac{1}{2}} & \mathbf{R} \end{bmatrix}$$

例 **10–2** 假設 $\mathbf{X} \sim \mathbf{N}(\mu, \mathbf{\Sigma})$，其中

$$\mathbf{\Sigma} = \begin{bmatrix} 1.00000 & 0.93982 & 0.28877 & 0.4216 & 0.34119 \\ 0.94002 & 1.00000 & 0.25384 & 0.42315 & 0.38246 \\ 0.28862 & 0.25384 & 1.00000 & 0.69297 & 0.89671 \\ 0.42153 & 0.42315 & 0.69297 & 1.00000 & 0.70774 \\ 0.34107 & 0.38246 & 0.89671 & 0.70774 & 1.00000 \end{bmatrix}$$

令

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^T\mathbf{X} \\ \mathbf{e}_2^T\mathbf{X} \\ \vdots \\ \mathbf{e}_p^T\mathbf{X} \end{bmatrix} = \mathbf{P}^T\mathbf{X}$$

試求 $\mathbf{Y}$ 與 $\mathbf{X}$ 之相關矩陣。

解：

首先求得 $\mathbf{\Sigma}$ 之特徵值及特徵向量如下

$$\mathbf{\Lambda} = \begin{bmatrix} 3.14810 & 0 & 0 & 0 & 0 \\ 0 & 1.35198 & 0 & 0 & 0 \\ 0 & 0 & 0.34127 & 0 & 0 \\ 0 & 0 & 0 & 0.12183 & 0 \\ 0 & 0 & 0 & 0 & 0.03686 \end{bmatrix}$$

及

$$\mathbf{P} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{e}_4 & \mathbf{e}_5 \end{bmatrix} = \begin{bmatrix} 0.41032 & 0.56714 & 0.08486 & 0.40784 & -0.58004 \\ 0.41193 & 0.56756 & 0.09014 & -0.33782 & 0.62160 \\ 0.45685 & -0.43421 & 0.34411 & 0.59229 & 0.36542 \\ 0.46840 & -0.20619 & -0.85905 & -0.01106 & -0.00371 \\ 0.48358 & -0.35444 & 0.35819 & -0.60713 & -0.37897 \end{bmatrix}$$

由於 $\sigma_{X_i}=1,\, i=1,2,3,4,5$，所以 $\mathbf{D}=\mathbf{I}$ 及

$$
\begin{aligned}
\mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}} \quad &= \begin{bmatrix}
0.41032 & 0.56714 & 0.08486 & 0.40784 & -0.58004 \\
0.41193 & 0.56756 & 0.09014 & -0.33782 & 0.62160 \\
0.45685 & -0.43421 & 0.34411 & 0.59229 & 0.36542 \\
0.46840 & -0.20619 & -0.85905 & -0.01106 & -0.00371 \\
0.48358 & -0.35444 & 0.35819 & -0.60713 & -0.37897
\end{bmatrix} \\
&\times \begin{bmatrix}
\sqrt{3.148063} & 0 & 0 & 0 & 0 \\
0 & \sqrt{1.351978} & 0 & 0 & 0 \\
0 & 0 & \sqrt{0.341269} & 0 & 0 \\
0 & 0 & 0 & \sqrt{0.121834} & 0 \\
0 & 0 & 0 & 0 & \sqrt{0.036856}
\end{bmatrix} \\
&= \begin{bmatrix}
0.72802 & 0.65944 & 0.04957 & 0.14236 & -0.11136 \\
0.73088 & 0.65993 & 0.05266 & -0.11792 & 0.11933 \\
0.81058 & -0.50488 & 0.20102 & 0.20674 & 0.07015 \\
0.83107 & -0.23975 & -0.50184 & -0.00386 & -0.07122 \\
0.85801 & -0.41212 & 0.20925 & -0.21192 & -0.07275
\end{bmatrix}
\end{aligned}
$$

所以，$\mathbf{Y}$ 與 $\mathbf{X}$ 之相關矩陣 $\mathbf{R_{XY}}$ 為

$$
\begin{bmatrix}
1.00000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.72802 & 0.65944 & 0.04957 & 0.14236 & -0.11136 \\
0.0000 & 1.00000 & 0.0000 & 0.0000 & 0.0000 & 0.73088 & 0.65993 & 0.05266 & -0.11792 & 0.11933 \\
0.0000 & 0.0000 & 1.00000 & 0.0000 & 0.0000 & 0.81058 & -0.50488 & 0.20102 & 0.20674 & 0.07015 \\
0.0000 & 0.0000 & 0.0000 & 1.00000 & 0.0000 & 0.83107 & -0.23975 & -0.50184 & -0.00386 & -0.07122 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.00000 & 0.85801 & -0.41212 & 0.20925 & -0.21192 & -0.07275 \\
0.72802 & 0.73088 & 0.81058 & 0.83107 & 0.85801 & 1.00000 & 0.93982 & 0.28877 & 0.4216 & 0.34119 \\
0.65944 & 0.65993 & -0.50488 & -0.23975 & -0.41212 & 0.94002 & 1.00000 & 0.25384 & 0.42315 & 0.38246 \\
0.04957 & 0.05266 & 0.20102 & -0.50184 & 0.20925 & 0.28862 & 0.25384 & 1.00000 & 0.69297 & 0.89671 \\
0.14236 & -0.11792 & 0.20674 & -0.00386 & -0.21192 & 0.42153 & 0.42315 & 0.69297 & 1.00000 & 0.70774 \\
-0.11136 & 0.11933 & 0.07015 & -0.07122 & -0.07275 & 0.34107 & 0.38246 & 0.89671 & 0.70774 & 1.00000
\end{bmatrix}
$$

表 10.1: 標準常態分配表; $\mathrm{P}(Z < z)$

z	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

第 10 章 版權宣告

戴忠淵

海洋大學航運管理系

January 17, 2025

chungyuandye@gmail.com

第 10 章 索引