

第四章

隨機樣本及其他

4.1 基本性質

在上一章我們討論了多維隨機變數, 有一種特別的多維隨機變數, 是常用來當作機率模式的, 就是所謂**隨機樣本**(random sample)。我們先給一定義。

定義1.1 若隨機變數 $X_1, \dots, X_n$ 為相互獨立, 且 X_i 's 之邊際p.d.f. $f(x)$ 皆相同, 則稱 $X_1, \dots, X_n$ 為一組從**母體**(population) $f(x)$ 產生之樣本數為 n 之**隨機樣本**(a random sample of size n from the population $f(x)$)。此時亦稱 $X_1, \dots, X_n$ 為**獨立且有共同分佈** (independent and identically distributed, 簡稱i.i.d.)之隨機變數, 以 $f(x)$ 為其共同p.d.f.。

在上述定義中, 常省略“樣本數為 n 之”數字, 此因由 $X_1, \dots, X_n$ 已顯示有 n 個樣本。又p.d.f. $f(x)$ 可換成分佈函數 $F(x)$, 或明定那一分佈。如可以說 $X_1, \dots, X_n$ 為一組i.i.d.的隨機變數, 以 $\mathcal{N}(0, 1)$ 為其共同分佈。諸如, $X_1, \dots, X_n$ 為一組由 $\mathcal{P}(\lambda)$ 分佈所產生的隨機樣本, 也都是表示 $X_1, \dots, X_n$ 為i.i.d.的隨機變數, 以 $\mathcal{P}(\lambda)$ 分佈為其共同分佈。這些不太一樣的講法, 其實都是指同一回事。

另外,在上述定義中,我們提到**母體**。早期母體是指對與人有關的探討,所得到的**觀測值**(observation)。今日則不論對動物、機器等,只要是我們有興趣探討的事物,皆可以母體稱之。例如,若想探討猴子之壽命,則母體可取為全世界的猴子。因此所謂**母體**就是觀測值之全體所成的集合。而一組樣本就是母體之一子集合,母體中之一元素就是一個樣本。

由母體得到的一觀測值,為一隨機變數(如上述猴子的壽命),因此便伴隨著機率密度函數,及分佈函數。由於我們是要做統計分析,故雖原始的母體可能是猴子,或飛機等,但我們只關心一母體某一數值特性的機率結構,或說機率模式。而i.i.d.隨機變數,是一種特別、常見,且好處理的樣本。這是定義1.1產生的背景。至於母體與樣本空間之別為何,我們留給讀者自行討論。

由定義1.1,當 $X_1, \dots, X_n$ 為i.i.d.,以 $f(x)$ 為其共同p.d.f.,則 $X_1, \dots, X_n$ 之聯合p.d.f.為

$$(1.1) \quad f(x_1, \dots, x_n) = f(x_1) \cdots f(x_n) = \prod_{i=1}^n f(x_i)。$$

若 $f(x)$ 屬於某一**參數族**(parametric family),即 $f(x)$ 有 $f(x|\theta)$ 的型式, θ 屬於某一集合(在2.1節我們稱此時有一**分佈族**),則 $X_1, \dots, X_n$ 之聯合p.d.f.為

$$(1.2) \quad f(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f(x_i|\theta)。$$

有時在做某一統計的探討時,由過去的經驗,或由一些合理的假設,我們知道 $f(x|\theta)$ 的型式,但參數 θ 未知,則 $X_1, \dots, X_n$ 之聯合p.d.f.便有(1.2)式所給的型式,只是 θ 未知。

例1.1 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\lambda)$ 分佈所產生之隨機樣本。則 $X_1, \dots, X_n$ 之聯合p.d.f.為

$$f(x_1, \dots, x_n|\lambda) = \prod_{i=1}^n f(x_i|\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}。$$

$X_1, \dots, X_n$ 可以代表某房間內 n 個燈泡的壽命(單位可取為月)。則兩個月內燈泡皆不壞的機率為

$$\begin{aligned}
& P(X_1 > 2, \dots, X_n > 2) \\
&= P(X_1 > 2) \cdots P(X_n > 2) \\
&= (P(X_1 > 2))^n = (e^{-2\lambda})^n = e^{-2n\lambda}。
\end{aligned}$$

此處用到因 $X_1, \dots, X_n$ 為獨立, 所以對 $\forall x_1, \dots, x_n$,

$$P(X_1 > x_1, \dots, X_n > x_n) = P(X_1 > x_1) \cdots P(X_n > x_n),$$

又因 $X_1, \dots, X_n$ 有相同的分佈, 故

$$P(X_1 > 2) = \cdots = P(X_n > 2)。$$

4.5節之例5.5與本例相關, 可互相對照。

是否存在不是隨機樣本的樣本呢? 答案是肯定的。

自一有 N 張紙牌(分別寫數字1至 N)的袋子中隨機地取紙牌, 每次取一張, 每次取出後放回 (with replacement), 令 $X_1, \dots, X_n$ 表取 n 次後所得之結果。則顯然此為一組隨機樣本, 以 $\mathcal{D}\text{-}\mathcal{U}(1, N)$ 為共同的分佈, p.d.f. 為

$$(1.3) \quad f(x) = P(X = x) = \frac{1}{N}, x = 1, \dots, N。$$

但若每次取出後不放回, 則所得之樣本 $Y_1, \dots, Y_n$ 就不是 i.i.d. 了, 因此非隨機樣本。雖然 $Y_1, \dots, Y_n$ 之邊際分佈仍相同, 還是 $\mathcal{D}\text{-}\mathcal{U}(1, N)$ (習題第2題), 但 $Y_1, \dots, Y_n$ 並不獨立。事實上可看出

$$\begin{aligned}
(1.4) \quad P(Y_2 = y | Y_1 = x) &= \begin{cases} \frac{1}{N-1}, & y \neq x, \\ 0 & y = x, \end{cases} \\
&\neq P(Y_2 = y) = \frac{1}{N}, y = 1, \dots, N。
\end{aligned}$$

又對上述每次取出後不放回的模式, 樣本數 n 不能超過 N 。但對每次取出後放回的模式, 樣本數 n 可以任意大。前者母體數為有限, 後者母體數為無限。

自一有限的母體取樣，每次取出後不放回，稱為簡單隨機取樣 (simple random sampling)。這種取樣方式，在2.2.5節介紹超幾何分佈時，我們便已引進。諸如民意調查，便常以簡單隨機取樣的方式，自母體(如某市的選民)中，抽出要訪問的對象。要注意的是，這種樣本不是如定義1.1所描述的隨機樣本(見下例)。在實際情況中，若母體數 N 比樣本數 n 大很多，則 $X_1, \dots, X_n$ 便近似於獨立，亦即若視 $X_1, \dots, X_n$ 為獨立，所產生的誤差並不會太大(可參考第三章例7.12)。

例1.2 設母體為 $\{1, 2, \dots, 1,000\}$ ，即母體數 $N = 1,000$ 。隨機地取出10個樣本，每次取出後不放回。則所得之樣本 $X_1, \dots, X_{10}$ 皆大於200之機率為何？
解. 令 Y 表 $X_1, \dots, X_{10}$ 中大於200之個數。則因母體中有800個數大於200，故 Y 有 $\mathcal{H}(1,000, 800, 10)$ 分佈，因此

$$P(Y = 10) = \frac{\binom{800}{10} \binom{200}{0}}{\binom{1,000}{10}} = \frac{\binom{800}{10}}{\binom{1,000}{10}} \doteq 0.106164。$$

次若視 $X_1, \dots, X_{10}$ 為獨立，則

$$\begin{aligned} &P(X_1 > 200, \dots, X_{10} > 200) \\ &= P(X_1 > 200) \cdots P(X_{10} > 200) \\ &= (P(X > 200))^{10} = 0.8^{10} \doteq 0.107374。 \end{aligned}$$

誤差並不太大，相對誤差只有百分之一左右：

$$\frac{0.107374 - 0.106164}{0.107374} \doteq 0.0113 = 1.13\%。$$

但若將 $X_1, \dots, X_{10}$ 視為獨立，計算上顯然容易很多。

此後，我們凡是提到隨機樣本，除非特別聲明，否則均指符合定義1.1的條件，即樣本為i.i.d.。

取得隨機樣本後，我們便要將樣本加以整理及分析，這是我們下一章起要討論的。如1.3節所述，在數學裡我們喜歡(或習慣)處理函數。諸如隨機變數、分佈函數、變數代換等，都是與函數有關的內容。有一些與樣

本有關的函數我們特別有興趣。如**樣本平均** (sample mean 或說**樣本平均數**), 其定義為

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

及**樣本變異數** (sample variance), 其定義為

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2。$$

至於**樣本標準差** (sample standard deviation) 則為 $S_n = \sqrt{S_n^2}$ 。

有些統計學家喜愛以**樣本變異係數** (sample coefficient of variation) $S_n/\overline{X}_n$ 來量測散佈情況。為什麼標準差要除以樣本平均呢? 可以這樣想, 若考試平均成績為80分, 則標準差5分便不算太大; 但若平均成績為8分, 則標準差5分就算很大了。所以用 $S_n/\overline{X}_n$ 來度量散佈的大小似乎是合理的。當然這是對取正值隨機變數的解釋。若隨機變數可取正值及負值, 則 $\overline{X}_n$ 有可能為0(或很接近0), 此時 $S_n/\overline{X}_n$ 就不易解釋了。

如果樣本數 n 很清楚, 在不會混淆下, 有時可寫成 $\overline{X}$ 、 S^2 及 S 。又大家是否注意到, 既然是函數, 是否應寫成 $\overline{X}_n(X_1, \cdots, X_n)$, 及 $S_n^2(X_1, \cdots, X_n)$ 才較清楚? 習慣上我們省略掉 $(X_1, \cdots, X_n)$ 而只簡單地以 $\overline{X}_n, S_n^2$ 表示。又如同以小寫的 x 表示隨機變數 X 之觀測值, 通常也以小寫的 $\overline{x}_n, s_n^2$, 分別表 $\overline{X}_n$ 及 S_n^2 之觀測值。只是有時並不特別區分。例如, 可寫“得到 $\overline{X}_n = 3.2$ ”。

樣本平均就是隨機樣本的**算術平均** (arithmetic average), 因此大家可能較清楚。樣本變異數是度量樣本的變異, 分別將樣本減去樣本平均, 平方後相加, 那為什麼不除以 n 呢(相對於離散型分佈的變異數, 由於 $X_1, \cdots, X_n$ 可視為有相同的可能性發生, 所以每一 X_i 發生的機率應皆相同, 因此 $(X_i - \overline{X}_n)^2$ 似應各除以 n 再相加)? 這點我們稍後再說明。我們先導出一些關於樣本平均及樣本變異數之性質。

定理1.1 對 $x_1, \dots, x_n \in R$, 令 $\bar{x}_n = \sum_{i=1}^n x_i/n$, $s_n^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2 / (n - 1)$ 。則

- (i) $\min_a \sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2$;
(ii) $(n - 1)s_n^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2 = \sum_{i=1}^n x_i^2 - n\bar{x}_n^2$ 。

證明. (i) 首先

$$\begin{aligned} \sum_{i=1}^n (x_i - a)^2 &= \sum_{i=1}^n (x_i - \bar{x}_n + \bar{x}_n - a)^2 \\ &= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + 2 \sum_{i=1}^n (x_i - \bar{x}_n)(\bar{x}_n - a) + \sum_{i=1}^n (\bar{x}_n - a)^2 \\ &= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + n(\bar{x}_n - a)^2。 \end{aligned}$$

此處用到

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x}_n)(\bar{x}_n - a) &= (\bar{x}_n - a) \sum_{i=1}^n (x_i - \bar{x}_n) \\ &= (\bar{x}_n - a) \left(\sum_{i=1}^n x_i - n\bar{x}_n \right) = 0。 \end{aligned}$$

因 $(\bar{x}_n - a)^2 \geq 0$, 故若且唯若 $a = \bar{x}_n$ 時, $\sum_{i=1}^n (x_i - a)^2$ 達到極小值。得證(i)。

(ii) 左側經展開後得

$$\begin{aligned} \sum_{i=1}^n (x_i - \bar{x}_n)^2 &= \sum_{i=1}^n (x_i^2 - 2x_i\bar{x}_n + \bar{x}_n^2) \\ &= \sum_{i=1}^n x_i^2 - 2\bar{x}_n \sum_{i=1}^n x_i + n\bar{x}_n^2 \\ &= \sum_{i=1}^n x_i^2 - 2n\bar{x}_n^2 + n\bar{x}_n^2 \\ &= \sum_{i=1}^n x_i^2 - n\bar{x}_n^2。 \end{aligned}$$

證畢。

上述定理之(ii)不論在理論或計算上都很有用。有時表示成 $\sum_{i=1}^n x_i^2 - n\bar{x}_n^2$ 後, 對求 $(n-1)s_n^2$ 較容易。

隨機樣本由於是i.i.d.的隨機變數, 所以期望值及變異數都有簡潔的表示法。下述定理的證明留在習題中(第3題)。

定理1.2 設 $X_1, \dots, X_n$ 為一組隨機樣本, 函數 $g(x)$ 滿足 $E(g(X))$ 及 $\text{Var}(g(X))$ 皆存在。則

- (i) $E(\sum_{i=1}^n g(X_i)) = nE(g(X))$;
- (ii) $\text{Var}(\sum_{i=1}^n g(X_i)) = n\text{Var}(g(X))$ 。

特別地, 若取 $g(x) = x$, 則得

- (iii) $E(\sum_{i=1}^n X_i) = nE(X)$;
- (iv) $\text{Var}(\sum_{i=1}^n X_i) = n\text{Var}(X)$ 。

下述定理為定理1.2之應用。

定理1.3 設 $X_1, \dots, X_n$ 為一組隨機樣本, 且 $E(X) = \mu$, $\text{Var}(X) = \sigma^2$ 皆存在。則

- (i) $E(\bar{X}_n) = \mu$;
- (ii) $\text{Var}(\bar{X}_n) = \sigma^2/n$;
- (iii) $E(S_n^2) = \sigma^2$ 。

證明. (i) 在定理1.2 中, 取 $g(x) = x/n$, 則得

$$E(\bar{X}_n) = E\left(\sum_{i=1}^n \frac{X_i}{n}\right) = nE\left(\frac{X}{n}\right) = n \cdot \frac{1}{n}E(X) = E(X) = \mu。$$

(ii) 同理可得

$$\text{Var}(\bar{X}_n) = \text{Var}\left(\sum_{i=1}^n \frac{1}{n}X_i\right) = n\text{Var}\left(\frac{1}{n}X\right) = n \cdot \frac{1}{n^2}\text{Var}(X) = \frac{1}{n}\text{Var}(X)。$$

(iii) 利用定理1.1 (ii),

$$E(S_n^2) = E\left(\frac{1}{n-1}\left(\sum_{i=1}^n X_i^2 - n\bar{X}_n^2\right)\right)$$

$$\begin{aligned}
&= \frac{1}{n-1} (nE(X^2) - nE(\overline{X}_n^2)) \\
&= \frac{1}{n-1} (n(\mu^2 + \sigma^2) - n(\frac{\sigma^2}{n} + \mu^2)) = \sigma^2。
\end{aligned}$$

證畢。

底下為定理1.3之一立即的推論。

系理1.1 承定理1.3。令

$$(1.5) \quad Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}。$$

則 $E(Z_n) = 0$, $\text{Var}(Z_n) = 1$ 。

在上述二定理中，如同一般的習慣，我們常以 X 代表一組隨機樣本 $X_1, \dots, X_n$ 中的某一個。

樣本平均的期望值等於母體的期望值，這符合我們的直觀，也是我們之所以偏好樣本平均的原因。樣本平均的變異數則隨著樣本數之增加而漸減。至於樣本變異數，如前所述，初見之下除以 $n-1$ 可能不是那麼有道理。現在我們知道了，若照一般求平均的習慣而除以 n ，則 $E(S_n^2)$ 便不等於 σ^2 。

樣本平均由於是 i.i.d. 隨機變數的平均，所以它有一些好的性質。如 $\overline{X}_n$ 之動差母函數與 X 的動差母函數有下述關係：

$$\begin{aligned}
(1.6) \quad M_{\overline{X}_n}(t) &= E(e^{t\overline{X}_n}) = E(e^{t(X_1 + \dots + X_n)/n}) \\
&= (E(e^{(t/n)X}))^n = (M_X(t/n))^n。
\end{aligned}$$

例1.3 (i) 設 $X_1, \dots, X_n$ 為由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。則

$$M_{\overline{X}_n}(t) = (e^{\mu(t/n) + \sigma^2(t/n)^2/2})^n = e^{\mu t + (\sigma^2/n)t^2/2}。$$

因此 $\overline{X}_n$ 有 $\mathcal{N}(\mu, \sigma^2/n)$ 分佈。樣本平均仍有常態分佈，期望值不變，但變異數隨著 n 之增大而變小（見定理1.3）。由此即得 Z_n （定義見(1.5)式）有 $\mathcal{N}(0, 1)$ 分佈（見2.3.6節）。

(ii) 設 $X_1, \dots, X_n$ 爲由 $\Gamma(\alpha, \beta)$ 分佈所產生之隨機樣本。則

$$M_{\overline{X}_n}(t) = \left(\frac{1}{(1 - \beta t/n)^\alpha} \right)^n = \frac{1}{(1 - (\beta/n)t)^{n\alpha}}。$$

即 $\overline{X}_n \sim \Gamma(n\alpha, \beta/n)$ 。換句話說, $\overline{X}_n$ 仍有 gamma 分佈, 且仍是

$$\begin{aligned} E(\overline{X}_n) &= n\alpha \cdot \beta/n = \alpha\beta = E(X), \\ \text{Var}(\overline{X}_n) &= n\alpha \cdot (\beta/n)^2 = \alpha\beta^2/n = \text{Var}(X)/n。 \end{aligned}$$

對二獨立的隨機變數 X, Y , 如果其動差母函數(或母函數等轉換), 並無簡單的型式, 則仍可以變數代換的方法, 求和 $Z = X + Y$ 之分佈。如此便引出底下之褶積公式(convolution formula)。

若 X, Y 皆爲取整數值之隨機變數, 則 $Z = X + Y$ 亦然。且對每一整數 z ,

$$\begin{aligned} (1.7) \quad f_Z(z) &= P(Z = z) \\ &= \sum_{x=-\infty}^{\infty} P(X = x, Y = z - x) \\ &= \sum_{x=-\infty}^{\infty} P(X = x)P(Y = z - x) \\ &= \sum_{x=-\infty}^{\infty} f_X(x)f_Y(z - x)。 \end{aligned}$$

上式最後一個和, 便稱爲 f_X 與 f_Y 之褶積。故 Z 之 p.d.f. 爲 X, Y 之 p.d.f. 的褶積。注意 $f_Z(z)$ 亦可寫成

$$(1.8) \quad f_Z(z) = \sum_{y=-\infty}^{\infty} f_X(z - y)f_Y(y)。$$

例1.4 設 X 與 Y 獨立, 且分別有 $\mathcal{B}(n, p)$ 及 $\mathcal{B}(m, p)$ 分佈。則和 $Z = X + Y$ 之 p.d.f. 爲(注意: X, Y 皆只取非負整數值)

$$f_Z(z) = \sum_{x=-\infty}^{\infty} f_X(x)f_Y(z - x)$$

$$\begin{aligned}
&= \sum_{x=0}^z \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{z-x} p^{z-x} (1-p)^{m-z+x} \\
&= p^z (1-p)^{m+n-z} \sum_{x=0}^z \binom{n}{x} \binom{m}{z-x} \\
&= \binom{m+n}{z} p^z (1-p)^{m+n-z}。
\end{aligned}$$

即得 Z 有 $\mathcal{B}(m+n, p)$ 分佈, 與第三章定理2.4所得相同。

若 X, Y 皆為連續型的隨機變數, 則 $Z = X + Y$ 之 p.d.f. 為

$$\begin{aligned}
(1.9) \quad f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \\
&= \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy, z \in R。
\end{aligned}$$

上述積分亦稱為 f_X 及 f_Y 之褶積。至於 Z 之分佈函數則為

$$\begin{aligned}
(1.10) \quad F_Z(z) &= P(X + Y \leq z) \\
&= \int_{-\infty}^{\infty} f_X(x) F_Y(z-x) dx \\
&= \int_{-\infty}^{\infty} f_Y(y) F_X(z-y) dy, z \in R。
\end{aligned}$$

若 X, Y 皆為取正值之隨機變數, 則

$$\begin{aligned}
(1.11) \quad f_Z(z) &= \int_0^z f_X(x) f_Y(z-x) dx \\
&= \int_0^z f_X(z-y) f_Y(y) dy, z \in R,
\end{aligned}$$

且

$$\begin{aligned}
(1.12) \quad F_Z(z) &= \int_0^z f_X(x) F_Y(z-x) dx \\
&= \int_0^z f_Y(y) F_X(z-y) dy, z \in R。
\end{aligned}$$

一般而言, 若以 F_X 及 F_Y 分別表隨機變數 X 及 Y 之分佈函數, 則 $Z = X + Y$ 之分佈函數可寫成 $F_Z = F_X * F_Y$, 稱為 F_X 與 F_Y 之褶積。又以

F_2 表 $F * F$, 以 F_n 表 $F * F_{n-1}, n \geq 2$ 。 F_n 稱為 F 與其自身之 n 層褶積。 F_n 即為 n 個i.i.d.隨機變數, 以 F 為其共同分佈函數, 其和之分佈函數。至於p.d.f.亦有褶積符號。例如, $f_Z = f_X * f_Y$ 。

例1.5 設 X, Y 為獨立的隨機變數, 皆有 $\mathcal{E}(\lambda)$ 分佈。則和 $Z = X + Y$ 之p.d.f.為

$$\begin{aligned} f_Z(z) &= \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)} dx \\ &= \int_0^z \lambda^2 e^{-\lambda z} dx \\ &= \lambda^2 z e^{-\lambda z}, z > 0. \end{aligned}$$

即 Z 有 $\Gamma(2, \lambda^{-1})$ 分佈。

我們再來介紹幾個定義。首先對兩組隨機樣本 $X_1, \dots, X_n$ 及 $Y_1, \dots, Y_n$, 令

$$\begin{aligned} S_{XY} &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}), \\ R_{XY} &= \frac{S_{XY}}{S_1 \cdot S_2} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{((\sum_{i=1}^n (X_i - \bar{X})^2)(\sum_{i=1}^n (Y_i - \bar{Y})^2))^{1/2}}, \end{aligned}$$

分別表其樣本共變異數(sample covariance)及樣本相關係數(sample correlation coefficient), 後者亦可稱為樣本相關(sample correlation), 其中 S_1 及 S_2 分別為 $X_1, \dots, X_n$ 及 $Y_1, \dots, Y_n$ 之樣本標準差。可看出 $X_1, \dots, X_n$ 與 $X_1, \dots, X_n$ 之樣本共變異數, 即為 $X_1, \dots, X_n$ 之樣本變異數。 S_{XY} 及 R_{XY} 之定義, 可分別與隨機變數 X 與 Y 之共變異數及相關係數對照。

最後我們要指出, 一般而言, 在簡單隨機取樣裡(即取出後不放回), 對應的樣本平均之期望值及變異數, 與隨機取樣裡是不同的。我們以下例來說明。

例1.6 自整數 $1, \dots, N$ 中, 簡單隨機取樣 n 個數。設依序取出 $X_1, \dots, X_n$ 。令 $E(X) = \mu$, $\text{Var}(X) = \sigma^2$, $\bar{X}_n = (X_1 + \dots + X_n)/n$ 。試證

- (i) $X_1, \dots, X_n$ 之邊際分佈相同,
(ii) $\mu = \frac{N+1}{2}, \quad \sigma^2 = \frac{N^2-1}{12},$

$$E(\bar{X}_n) = \frac{N+1}{2}, \quad \text{Var}(\bar{X}_n) = \frac{(N+1)(N-n)}{12n} = \frac{N-n}{N-1} \frac{\sigma^2}{n} \circ$$

證明. (i) 由總機率法則,

$$\begin{aligned} P(X_{k+1} = x) &= \sum_{j_1, \dots, j_k \neq x} P(X_{k+1} = x | X_1 = j_1, \dots, X_k = j_k) \\ &\quad \cdot P(X_1 = j_1, \dots, X_k = j_k) \\ &= \sum_{j_1, \dots, j_k \neq x} \frac{1}{N-k} \cdot \frac{1}{N(N-1) \cdots (N-k+1)} \\ &= (N-1) \cdots (N-1-k+1) \cdot \frac{1}{N(N-1) \cdots (N-k)} \\ &= \frac{1}{N} = P(X_1 = x), k = 1, \dots, n-1, \end{aligned}$$

故得證 $X_1, \dots, X_n$ 有相同之邊際分佈。

(ii) 首先

$$\begin{aligned} \mu &= E(X) = \sum_{i=1}^N i \cdot \frac{1}{N} = \frac{N(N+1)}{2} \cdot \frac{1}{N} = \frac{N+1}{2}, \\ \sigma^2 &= \text{Var}(X) = E(X^2) - (E(X))^2 \\ &= \sum_{i=1}^N i^2 \cdot \frac{1}{N} - \left(\frac{N+1}{2}\right)^2 \\ &= \frac{N(N+1)(2N+1)}{6N} - \frac{(N+1)^2}{4} \\ &= \frac{N^2-1}{12} \circ \end{aligned}$$

故

$$E(\bar{X}_n) = E(X) = \frac{N+1}{2} \circ$$

其次利用第三章定理7.5,

$$\text{Var}(\bar{X}_n) = \frac{1}{n^2} (n \text{Var}(X) + n(n-1) \text{Cov}(X_1, X_2)) \circ$$

而

$$\begin{aligned}
 \text{Cov}(X_1, X_2) &= E(X_1 X_2) - (E(X))^2 \\
 &= \sum_{i \neq j} ij \cdot \frac{1}{N(N-1)} - \left(\frac{N+1}{2}\right)^2 \\
 &= \frac{1}{N(N-1)} \left((1 + \cdots + N)(1 + \cdots + N) - \sum_{i=1}^N i^2 \right) - \frac{(N+1)^2}{4} \\
 &= \frac{1}{N(N-1)} \left(\left(\frac{N(N+1)}{2}\right)^2 - \frac{N(N+1)(2N+1)}{6} \right) - \frac{(N+1)^2}{4} \\
 &= \frac{1}{N(N-1)} \frac{N(N+1)(3N(N+1) - 2(2N+1))}{12} - \frac{(N+1)^2}{4} \\
 &= \frac{(N+1)(N-1)(3N+2)}{12(N-1)} - \frac{(N+1)^2}{4} \\
 &= \frac{(N+1)(3N+2-3N-3)}{12} = -\frac{N+1}{12}。
 \end{aligned}$$

故得

$$\begin{aligned}
 \text{Var}(\bar{X}_n) &= \frac{1}{n^2} \left(\frac{n(N^2-1)}{12} - \frac{n(n-1)(N+1)}{12} \right) \\
 &= \frac{(N+1)(N-n)}{12n} = \frac{N-n}{N-1} \frac{\sigma^2}{n}。
 \end{aligned}$$

證畢。

如果是隨機取樣，則 $E(\bar{X}_n) = (N+1)/2$ ，而 $\text{Var}(\bar{X}_n) = \sigma^2/n$ 。故取出後放回與取出後不放回，樣本平均不變，但後者之樣本變異數較小，為前者之 $(N-n)/(N-1)$ 倍。減小是合理的，這點留給讀者自行討論。

一般民調裡的取樣，往往屬於取出後不放回。但若 n 與 N 相比很大，則 $(N-n)/(N-1)$ 很接近 1，此時 $\text{Var}(\bar{X}_n)$ 近似於取出後放回的 $\bar{X}_n$ 之變異數 σ^2/n 。

4.2 收斂的概念

在數學裡極限為一重要的概念。在機率論裡也是一樣，當樣本數無止盡的增大，一些關於樣本的量會有什麼極限的結果？雖然無限的樣本數，

只是一理論上的假設，但極限之下的結果(往往變得較簡單)，常可做為有限樣本數之下很好的近似值。

本節只簡單地介紹兩種收斂的型式。詳細有關收斂的討論，可見一般較高等的機率論的書，如Chung(1974)。

我們先看上一章柴比雪夫不等式之一些應用。

例2.1 某市長候選人欲估計選民對他的支持度 p ，也就是隨機地取一位選民會支持他之機率。現希望誤差不超過 $\varepsilon = 0.05$ 之機率至少是0.9。若利用柴比雪夫不等式，則至少要訪問多少選民？

解.欲建立一機率模式。令 $X_1, \dots, X_n$ 為i.i.d., 以 $Ber(p)$ 為共同分佈之隨機變數。 $X_i = 1$ 或0, 依第 i 個人是否支持該候選人而定。則樣本平均 $\overline{X}_n = \sum_{i=1}^n X_i/n$ 為此 n 個人中支持該候選人之比率。我們想求 n , 使得

$$(2.1) \quad P(|\overline{X}_n - p| < \varepsilon) \geq \delta.$$

因 $E(X) = p$, $\text{Var}(X) = p(1-p)$, 由定理1.3,

$$E(\overline{X}_n) = p, \quad \text{Var}(\overline{X}_n) = \frac{p(1-p)}{n}.$$

故第三章(8.22)式導致

$$(2.2) \quad P(|\overline{X}_n - p| < \varepsilon) \geq 1 - \frac{p(1-p)}{n\varepsilon^2}.$$

又因對任意 $0 \leq p \leq 1$, $p(1-p) \leq 1/4$, 故

$$(2.3) \quad P(|\overline{X}_n - p| < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}, \forall \varepsilon > 0.$$

因此只要

$$1 - \frac{1}{4n\varepsilon^2} \geq \delta,$$

則(2.1)式成立。

現取 $\varepsilon = 0.05$, $\delta = 0.9$, 則由

$$1 - \frac{1}{4n \cdot 0.05^2} \geq 0.9,$$

解出

$$n \geq \frac{1}{4 \cdot 0.1 \cdot 0.05^2} = 1,000。$$

即 n 至少要1,000。要留意的是, 此 n 值有可能過大。一方面是因由柴比雪夫不等式所得到的第三章(8.24)式並非很精確, 一方面以 $1/4$ 取代 $p(1-p)$ 也會造成一些誤差。稍後我們將利用中央極限定理, 得到一較佳的 n 之下界。

在給柴比雪夫不等式之另一應用前, 我們先給一定義。

首先對 $X_1, \dots, X_n$ 為一組隨機樣本, 以 F 為共同分佈函數。令

$$(2.4) \quad F_n(x) = \frac{X_i \leq x \text{ 之個數}}{n}, x \in R。$$

$F_n(x)$ 稱為此組樣本之**經驗分佈函數**(empirical distribution function), 為一階梯函數。這是一種用來估計分佈函數 F 的方法。

例2.2 設 $X_1, \dots, X_5$ 為一組隨機樣本。今得到觀測值3, 6, 5.5, 3, 4。則

$$F_5(x) = \begin{cases} 0 & , x < 3, \\ 0.4 & , 3 \leq x < 4, \\ 0.6 & , 4 \leq x < 5, \\ 0.8 & , 5.5 \leq x < 6, \\ 1 & , x \geq 6。 \end{cases}$$

例2.3 設 $F(x)$, $F_n(x)$ 分別為隨機樣本 $X_1, \dots, X_n$ 之分佈函數及經驗分佈函數。試證對 $\forall x \in R$,

$$(2.5) \quad P(|F_n(x) - F(x)| < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}, \forall \varepsilon > 0。$$

證明. 首先, 可看出對一固定的 $x \in R$, $F_n(x)$ 為 n 個i.i.d.伯努力隨機變數之樣本平均,

$$F_n(x) = \bar{I}_n = \frac{1}{n} \sum_{i=1}^n I_i,$$

其中指示函數

$$I_i = \begin{cases} 1, & X_i \leq x, \\ 0, & X_i > x. \end{cases}$$

又

$$p = E(I_i) = P(X_i \leq x) = F(x)。$$

故由(2.3)式, 對 $\forall x \in R$,

$$P(|F_n(x) - F(x)| < \varepsilon) = P(|\bar{I}_n - p| < \varepsilon) \geq 1 - \frac{1}{4n\varepsilon^2}, \varepsilon > 0。$$

證畢。

對同一分佈函數 F , 雖每次會有不同的觀測值, 因此會得到不同的 $F_n(x)$ 。但上例顯示只要 n 夠大, $F_n(x)$ 為 $F(x)$ 之一很好的估計。又對 $\forall \varepsilon > 0$, 當 $n \rightarrow \infty$ 時, 因 $4n\varepsilon^2 \rightarrow \infty$, 故得

$$(2.6) \quad \lim_{n \rightarrow \infty} P(|F_n(x) - F(x)| < \varepsilon) = 1。$$

根據**墨菲定律**(Murphy's law), 該錯的總是會錯(If something can go wrong it will)。有人認為這是於西元1917年出生在美國, 而服務於海軍的Edward A. Murphy最先說的。時至今日已有成千上萬則墨菲定律了, 有些甚至還有系理(corollary)呢。多半是一些人們的經驗法則, 或平常自我解嘲、諷刺或無可奈何時脫口說出的。如排隊時可能會覺得“另一條線總是較快”(The other line always moves faster); 欲借款時可能會認為“要獲得貸款須先證明不需要它”(In order to get a loan, you must prove you don't need it)。其他有趣的尚有“笑一下吧, 明天會更糟”(Smile, tomorrow will be worse)。讀者也可有自己的墨菲定律。如勸人照相時可說“現在照永遠比以後照好看”。就“該錯的總是會錯”此則而言, 從機率的角度的看是對的。若實驗中某事件發生的機率為正, 則持續重複此實驗, 該事件遲早會發生。事實上不僅於此, 設某事件發生之機率為 p , 若持續重複實驗, 則該事件發生之**相對頻率**(relative frequency)將接近 p 。此性質為機率

中之一極重要的定理—**大數法則**(Law of large numbers)之一特例。

註2.1 網路上可查到Murphy's law有許多則。但並未有太多字典列有此名詞。在Random House 書局出版的Webster's Unabridged Dictionary, Second Edition(2001)裡, 對Murphy's Law 之解釋為: The facetious proposition that if something can go wrong, it will. Also called Murphy's First Law; after a fictions Murphy, allegedly the name of a bungling mechanic in U.S. Navy educational Cartoons of the 1950's。Webster(韋伯)字典認為Murphy為虛構的人物。在The Oxford English Dictionary, Second Edition(1989) 有較多對於Murphy's law 的說明。

大數法則是說當樣本數夠多, 樣本平均會收斂至母體期望值。這個法則可讓大家心理舒服一點。也許人生運氣有好有壞, 但長期下來, 平均的表現, 與真正該有之表現(指期望值), 差異便不大。在伯努力試驗的情況下, 樣本平均即為事件發生之相對頻率(見例2.1), 而母體期望值即事件發生之機率。印證之前所述大數法則可應用至事件發生之相對頻率。

在例2.1中, 我們有一數列之i.i.d.的伯努力隨機變數, 成功機率為 p 。令 $\bar{X}_n$ 表首 n 項之平均, 此亦為首 n 次試驗成功數之比率。即

$$\bar{X}_1 = X_1, \quad \bar{X}_2 = \frac{X_1 + X_2}{2}, \quad \bar{X}_3 = \frac{X_1 + X_2 + X_3}{3}, \dots$$

(2.3)式指出, 只要取夠大的 n , 便可使 $\bar{X}_n$ 與 p 之差異不超過一很小的正數 ε 之機率任意接近1。要注意的是, 一般而言, $\bar{X}_n$ 並未剛好等於 p , 而且我們只是很有信心(而非絕對有把握) n 很大時, $\bar{X}_n$ 與 p 之差不超過一給定的值。

以上討論便引出了下述定義。

定義2.1 設 $X_1, X_2, \dots$ 為一數列之隨機變數, 若存在一隨機變數 X , 使得

$$(2.7) \quad \lim_{n \rightarrow \infty} P(|X_n - X| < \varepsilon) = 1, \forall \varepsilon > 0,$$

則稱 $\{X_n, n \geq 1\}$ **機率收斂**(converges in probability)至 X , 以 $X_n \xrightarrow[n \rightarrow \infty]{p} X$

表之。

(2.7)式等價於

$$\begin{aligned}
 (2.8) \quad & \lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0, \forall \varepsilon > 0 \\
 & \Leftrightarrow \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0, \forall \varepsilon > 0 \\
 & \Leftrightarrow \lim_{n \rightarrow \infty} P(|X_n - X| \leq \varepsilon) = 1, \forall \varepsilon > 0.
 \end{aligned}$$

注意在定義2.1中的隨機變數 $X_1, X_2, \dots$ 不一定要i.i.d.。又由例2.3知, 對 $\forall x \in R$, 隨機樣本之經驗分佈函數 $\{F_n(x), n \geq 1\}$, 機率收斂至 $F(x)$ 。

定理2.1 弱大數法則 (Weak law of large numbers). 設 $X_1, X_2, \dots$ 為i.i.d.

之隨機變數, 且 $E(X) = \mu$ 存在。則樣本平均 $\overline{X}_n \xrightarrow[n \rightarrow \infty]{p} \mu$ 。

證明. 本定理的證明此處略去。但當假設變異數 $E(X^2) = \sigma^2$ 亦存在, 則便可藉助第三章(8.20)式的柴比雪夫不等式, 證明便很容易: 由於 $E(\overline{X}_n) = \mu$, $\text{Var}(\overline{X}_n) = \sigma^2/n$, 故當 $n \rightarrow \infty$ 時, 對 $\forall \varepsilon > 0$,

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\overline{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0.$$

依定義知 $\overline{X}_n \xrightarrow[n \rightarrow \infty]{p} \mu$ 。

弱大數法則是一優美的結果, 而只要變異數存在, 證明便很簡單。此結果指出, 在很一般的條件下(期望值存在), 隨機樣本的樣本平均機率收斂至期望值。所以樣本平均當樣本數很大時, 在“某種意義下”與期望值很接近。更明確地說, 當 n 很大時, $\overline{X}_n$ 很接近 μ 的機率會很大。如果 X 的變異數存在, 那就看得更清楚:

$$\text{Var}(\overline{X}_n) = \frac{1}{n} \text{Var}(X) \xrightarrow[n \rightarrow \infty]{} 0.$$

但若期望值不存在, 則樣本平均的變異, 就不見得會隨著樣本數之增大而變得很小。見底下例2.4。

有弱大數法則，顯然也應有強大數法則(Strong law of large numbers)。的確有，不過本書不擬討論。另外，定理2.1的條件尚可再弱些，而 $\{X_n, n \geq 1\}$ 也可以不用假設獨立，此處我們也不多討論。

例2.4 設 $X_1, X_2, \dots$ 為i.i.d.，以 $\mathcal{C}(\theta, a)$ 為共同分佈。由第二章(3.39)式， $E(e^{itX}) = e^{-a|t|+it\theta}$ ， $t \in R$ 。因此

$$\begin{aligned} E(e^{it\bar{X}_n}) &= E(e^{i(t/n)\sum_{j=1}^n X_j}) \\ &= \prod_{j=1}^n E(e^{i(t/n)X_j}) \\ &= (E(e^{i(t/n)X}))^n \\ &= (e^{-a|t/n|+i(t/n)\theta})^n \\ &= e^{-a|t|+it\theta} \\ &= E(e^{itX})。 \end{aligned}$$

即對 $\forall n \geq 1$ ， $\bar{X}_n \stackrel{d}{=} X$ 。這是一特殊的現象，樣本平均與任一樣本的分佈仍相同！這是柯西分佈之一特性：採用樣本平均 $\bar{X}_n$ 與單一樣本 X 的“效果”一模一樣(什麼效果?)，不論樣本數為何。

例2.5 設 $X_1, X_2, \dots$ 為i.i.d.之隨機變數，且 $E(X) = \mu$ ， $\text{Var}(X) = \sigma^2$ 皆存在。則樣本變異數 S_n^2 是否會機率收斂至 σ^2 ？

解。利用柴比雪夫不等式，得 $\forall \varepsilon > 0$ ，

$$P(|S_n^2 - \sigma^2| \geq \varepsilon) \leq \frac{\text{Var}(S_n^2)}{\varepsilon^2}。$$

故 $S_n^2 \xrightarrow[n \rightarrow \infty]{p} \sigma^2$ 之一充分條件為 $\text{Var}(S_n^2) \xrightarrow[n \rightarrow \infty]{} 0$ 。當然這只是充分條件，不表此條件不成立時， S_n^2 便不機率收斂至 σ^2 。事實上即使沒有 $\text{Var}(S_n^2) \xrightarrow[n \rightarrow \infty]{} 0$ 的條件，此結果仍成立。我們留至第七章例5.3再證明。證明當然不能再藉助柴比雪夫不等式。

大數法則告訴我們，對一組隨機樣本 $X_1, \dots, X_n$ ，母體之期望值及變異數分別以 μ 及 σ^2 表之，則不論母體之分佈為何，當 $n \rightarrow \infty$ 時，樣本平均 $\bar{X}_n$ 皆機率收斂至 μ 。還有沒有其他結果呢？當 n 很大時，不論母體之分佈為何， $\bar{X}_n$ 的分佈是否會收斂至某分佈？

我們先給底下定義。這是我們要介紹的另一種收斂方式。如同定義2.1，其中 $X_1, \dots, X_n$ 亦不須假設為i.i.d.。

定義2.2 設 $X_1, X_2, \dots$ 為一數列之隨機變數， X_n 之分佈函數以 $F_{X_n}(x)$ 表之。若存在一隨機變數 X ，以 $F(x)$ 為分佈函數，使得

$$(2.9) \quad \lim_{n \rightarrow \infty} F_{X_n}(x) = F(x),$$

對函數 $F(\cdot)$ 之每個連續點 x 成立，則稱 $\{X_n, n \geq 1\}$ 分佈收斂(converges in distribution)至 X ，以 $X_n \xrightarrow[n \rightarrow \infty]{d} X$ 表之；或說 $\{X_n, n \geq 1\}$ 分佈收斂至以 F 為分佈函數之分佈。 F 又稱為 $\{X_n, n \geq 1\}$ 之極限分佈(limiting distribution)。

要注意的是，分佈收斂其實是分佈函數的收斂，而非隨機變數的收斂。此因可以有無限多個隨機變數有同一分佈函數 $F(x)$ 。所以在定義2.2中，當 $X_n \xrightarrow[n \rightarrow \infty]{d} X$ ，並不是說 X_n 與 X 有多接近，而是說它們的分佈函數會接近(隨著 n 之增大)。此點我們稍後會再說明。因此也常有下述這類寫法：

$$X_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1), \quad Y_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{P}(\lambda)。$$

附帶一提，類似地，有時會以 $P(\mathcal{N}(0, 1) \leq x)$ 表 $\mathcal{N}(0, 1)$ 分佈之分佈函數。其他分佈也一樣。

例2.6 設 X_n 有 $\mathcal{B}(n, p_n)$ 分佈， $n \geq 1$ ，且 $\lim_{n \rightarrow \infty} np_n = \lambda$ ， $0 < \lambda < \infty$ 。因

$$F_{X_n}(x) = \sum_{k=0}^{[x]} P(X_n = k),$$

而由第二章定理2.2,

$$\lim_{n \rightarrow \infty} P(X_n = k) = \frac{e^{-\lambda} \lambda^k}{k!}, k = 0, 1, \dots。$$

故對 $\forall x \in R$,

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{X_n}(x) &= \lim_{n \rightarrow \infty} \sum_{k=0}^{[x]} P(X_n = k) = \sum_{k=0}^{[x]} \lim_{n \rightarrow \infty} P(X_n = k) \\ &= \sum_{k=0}^{[x]} \frac{e^{-\lambda} \lambda^k}{k!} = F(x), k = 0, 1, \dots。 \end{aligned}$$

其中 F 為 $\mathcal{P}(\lambda)$ 之分佈函數, 又若 $i < 0$, 則定義 $\sum_{k=0}^i e^{-\lambda} \lambda^k / k! = 0$, 因此由定義2.2, $X_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{P}(\lambda)$ 。

例2.7 設 X_1 有 $\mathcal{N}(0, 1)$ 分佈, $X_2 = -X_1$, $X_3 = X_1$, $X_4 = -X_1, \dots$, 即 $X_{2n} = -X_1$, $X_{2n+1} = X_1, \forall n \geq 1$ 。顯然 X_n 皆有 $\mathcal{N}(0, 1)$ 分佈。取 $X = X_1$, 仍有 $\mathcal{N}(0, 1)$ 分佈。因 $\forall n \geq 1, F_{X_n}(x) = \Phi(x)$, 故 $X_n \xrightarrow[n \rightarrow \infty]{d} X$ 。但 n 再大時, X_n 的值與 X 都不會接近: 當 X 的值為5時, X_n 交替地取5及-5。

對期望值為 μ , 變異數為 σ^2 之隨機樣本, 令 Z_n 如(1.5)式, 則由例1.3, 當母體有常態分佈, Z_n 有 $\mathcal{N}(0, 1)$ 分佈, 故若 Z_n 有一共同的極限分佈存在, 此分佈必為常態。這就是著名的中央極限定理。

定理2.2 (中央極限定理). 設 $X_1, X_2, \dots$ 為i.i.d.之隨機變數, 且 $E(X) = \mu$, $\text{Var}(X) = \sigma^2$ 皆存在。令 $F_n(x)$ 表 $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ 之分佈函數。則

$$(2.10) \quad \lim_{n \rightarrow \infty} F_n(x) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du, x \in R。$$

即 $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ 分佈收斂至 $\mathcal{N}(0, 1)$ 分佈, 亦可以 $\sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow[n \rightarrow \infty]{d}$
 $\mathcal{N}(0, 1)$ 表之。

中央極限定理的證明我們略去。此定理指出適當地常態化後, i.i.d. 隨機變數的和, 即

$$(2.11) \quad \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \sum_{i=1}^n \frac{X_i - \mu}{\sigma\sqrt{n}},$$

會趨近至標準常態分佈。在很多實際的情況中, 通常樣本數 $n \geq 30$, 就已適用, 即誤差已不會太大。甚至可以有更一般的結果: 在較高等的機率論中, 有 $X_1, X_2, \dots$ 之分佈不一定要相同之中央極限定理的版本。

中央極限定理解釋了為什麼常態分佈可用來作為很多實際現象的模式。諸如身高、智商、量測的誤差, 都是很多(假設有 n 個)微小的獨立的量之和, 因此其總和 S_n , “差不多”可以常態分佈 $\mathcal{N}(n\mu, n\sigma^2)$ 當作近似分佈, 其中 μ, σ^2 分別為每一微小的量之期望值及變異數。

由於

$$(2.12) \quad \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}},$$

所以在定理2.2, 我們亦可說

$$(2.13) \quad \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1),$$

或

$$(2.14) \quad \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

例2.8 設 $X_1, X_2, \dots$ 為 i.i.d. 以 $\mathcal{P}(\lambda)$ 為其共同分佈。因 $E(X) = \text{Var}(X) = \lambda$, 故

$$\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

下例可與上例相對應。

例2.9 設 X_n 有 χ_n^2 分佈。底下我們證明

$$(2.15) \quad \frac{X_n}{n} \xrightarrow[n \rightarrow \infty]{p} 1,$$

$$(2.16) \quad \frac{X_n - n}{\sqrt{2n}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

此為 χ_n^2 分佈之性質。

首先由第三章定理2.4-(iv),

$$X_n \stackrel{d}{=} \sum_{i=1}^n Y_i,$$

其中 $Y_i, i = 1, \dots, n$, 為 i.i.d., 以 χ_1^2 為共同分佈。因 $E(Y_i) = 1, \text{Var}(Y_i) = 2$, 故由大數法則

$$\frac{X_n}{n} = \frac{\sum_{i=1}^n Y_i}{n} \xrightarrow[n \rightarrow \infty]{p} 1 = E(Y),$$

再由中央極限定理

$$\frac{X_n - n}{\sqrt{2n}} = \frac{\sum_{i=1}^n Y_i - nE(Y)}{\sqrt{n}\sqrt{\text{Var}(Y)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

對 $\forall 0 < p < 1$, 令 $\chi_{p,n}^2$ 為 χ_n^2 分佈之第 p 分位數。即 $\chi_{p,n}^2$ 滿足 $P(\chi_n^2 \leq \chi_{p,n}^2) = p$ 。則由(2.16)式得, 對 $\forall x > 0$,

$$(2.17) \quad \begin{aligned} P(\chi_n^2 \leq x) &= P\left(\frac{\chi_n^2 - n}{\sqrt{2n}} \leq \frac{x - n}{\sqrt{2n}}\right) \\ &\doteq P\left(\mathcal{N}(0, 1) \leq \frac{x - n}{\sqrt{2n}}\right) = \Phi\left(\frac{x - n}{\sqrt{2n}}\right)。 \end{aligned}$$

其次如前令 z_p 表 $\mathcal{N}(0, 1)$ 分佈之第 p 分位數。則上式導致 n 很大時,

$$(2.18) \quad \chi_{p,n}^2 \doteq n + z_p \sqrt{2n}。$$

χ_n^2 分佈之機率值即可藉助(2.17)或(2.18)式得到。例如,

$$P(\chi_{30}^2 \leq 40) \doteq \Phi\left(\frac{40 - 30}{\sqrt{2 \cdot 30}}\right) \doteq \Phi(1.291) \doteq 0.9017。$$

又

$$\chi_{30,0.95}^2 \doteq 30 + z_{0.95} \sqrt{2 \cdot 30} \doteq 30 + 1.645 \sqrt{60} \doteq 42.74。$$

由附表5, 此二值分別約為0.89及43.77, 差異不太大。

例2.10 設 $X_1, X_2, \dots$ 為 i.i.d. 以 $\text{Ber}(p)$ 為其共同分佈, $0 < p < 1$ 。因 $E(X) = p$, $\text{Var}(X) = p(1-p)$, 故

$$(2.19) \quad \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

由於 $\sum_{i=1}^n X_i$ 有 $\mathcal{B}(n, p)$ 分佈, 上式指出當樣本數 n 夠大時, 二項分佈可以常態分佈作為近似分佈。此結果又稱**棣美弗-拉普拉斯定理**(De Moivre-Laplace theorem), 這可說是機率論裡, 第一個關於極限的定理, 當然也是中央極限定理的一個特例。通常只要 np 及 $n(1-p)$ 皆大於 5, 此近似便夠精確。若 np 或 $n(1-p)$ 太小, 便不太適合以常態分佈來當做二項分佈之近似, 產生的誤差可能過大。實際應用時, 若採用**連續性的更正**(continuity correction)會更精確。即對任二非負整數 k 及 j ,

$$(2.20) \quad \begin{aligned} P(j \leq S_n \leq k) &= P(j - \frac{1}{2} \leq S_n \leq k + \frac{1}{2}) \\ &= P(\frac{j - np - 1/2}{\sqrt{np(1-p)}} \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{k - np + 1/2}{\sqrt{np(1-p)}}) \\ &\doteq \Phi(\frac{k - np + 1/2}{\sqrt{np(1-p)}}) - \Phi(\frac{j - np - 1/2}{\sqrt{np(1-p)}})。 \end{aligned}$$

當然如果 $1/2$ 與 $\sqrt{np(1-p)}$ 相比很小, 忽略 $1/2$ 便不致造成太大的影響。

例如, $n = 400, p = 1/2$, 則 $np = 200, \sqrt{np(1-p)} = 10$, 故得

$$\begin{aligned} &P(190 \leq \sum_{i=1}^{400} X_i \leq 210) \\ &= P(\frac{190 - 200 - 0.5}{10} \leq \frac{\sum_{i=1}^{400} X_i - 200}{10} \leq \frac{210 - 200 + 0.5}{10}) \\ &= P(-1.01 \leq \frac{\sum_{i=1}^{400} X_i - 200}{10} \leq 1.01) \\ &\doteq \Phi(1.01) - \Phi(-1.01) \\ &\doteq 0.6876。 \end{aligned}$$

上述推導顯示, 投擲一公正的銅板 400 次, 則中央極限定理告訴我們, 所得之正面數與期望次數 200 之差距, 不超出 10 之機率約為 0.6826。但若

想以二項分佈求

$$\sum_{i=190}^{210} \binom{400}{i} \left(\frac{1}{2}\right)^{400},$$

明顯地困難不少。

讀者也可自行計算, 投擲一公正的銅板, 要得到正面數有65個以上, 是很難發生的(機率約0.0013), 更不要說是得到95個正面了。

有時佐以其他極限的定理, 會使中央極限定理更有用。底下即著名的**史拉斯基定理**(Slutsky's theorem), 證明在此略去。

定理2.3 設 $X_n \xrightarrow[n \rightarrow \infty]{d} X$, $Y_n \xrightarrow[n \rightarrow \infty]{p} a$, a 為一常數。則

$$(i) X_n + Y_n \xrightarrow[n \rightarrow \infty]{d} X + a;$$

$$(ii) X_n - Y_n \xrightarrow[n \rightarrow \infty]{d} X - a;$$

$$(iii) X_n Y_n \xrightarrow[n \rightarrow \infty]{d} aX;$$

$$(iv) X_n / Y_n \xrightarrow[n \rightarrow \infty]{d} X/a, \text{ 只要 } a \neq 0。$$

例2.11 北銀樂透彩之中獎機率(指中頭獎至普獎)約為 $p = 0.02906$ (見第二章例2.13)。不過因投注者可自由選號, 所以每期實際中獎率均有波動。令 N 表某期賣出的張數, 則該期總共中獎數 X 有 $B(N, p)$ 分佈。一般而言 N 均很大, Np 及 $N(1-p)$ 皆大於5也不成問題, 因此由例2.10,

$$Z = \frac{X - Np}{\sqrt{Np(1-p)}}$$

有近似的 $\mathcal{N}(0, 1)$ 分佈。因此

$$P(|Z| \leq 2) = P\left(\left|\frac{X - Np}{\sqrt{Np(1-p)}}\right| \leq 2\right) \doteq 0.9544。$$

同理

$$P(|Z| \leq 3) \doteq 0.9974。$$

$|Z|$ 值要超過3之機率可說相當小,更不要說 $|Z|$ 值超過10了。現以第30期為例,該期 $N = 13,210,954$ 。而

$$\left| \frac{X - 13,210,954 \cdot 0.02906}{\sqrt{13,210,954 \cdot 0.02906 \cdot 0.97094}} \right| \leq 2$$

$$\Leftrightarrow 382,689.25 \leq X \leq 385,131.39。$$

即約有0.9544之機率,該期中獎數應介於382,689與385,131間。我們將該期各獎之實際中獎數列於表2.1,相加後得總中獎數為365,568,明顯地偏低。對應的 Z 值約為

$$Z = \frac{365,568 - 13,210,954 \cdot 0.02906}{\sqrt{13,210,954 \cdot 0.02906 \cdot 0.97094}}$$

$$\doteq -30.042。$$

對 $\mathcal{N}(0, 1)$ 分佈,會得到如此小的 Z 值,其機率是很低的。這顯示該期樂透彩填注之號碼不夠隨機(前30期 Z 值最大約為298,最小約為-238,但無一介於 ± 3 之間)。若是電腦選號較不會有此現象,但人腦“隨意”選號,則常有這種現象發生。此概念在第十章我們會再討論。

表2.1 第30期樂透彩各獎中獎數

頭獎	貳獎	參獎	肆獎	普獎
0	13	389	21,455	343,711

例2.12 假設進行某項化學實驗的量測,而得到30筆觀測值 $X_1, \dots, X_{30}$ 。設 $X_1, \dots, X_{30}$ 為i.i.d., 期望值 μ , 且依過去經驗, 變異數 $\sigma^2 = 6$ 。欲以樣本平均 $\overline{X}_{30}$ 來估計 μ , 試給出誤差超過1之機率的近似值。

解.由中央極限定理

$$P(|\overline{X}_{30} - \mu| > 1) = P\left(\left|\frac{\overline{X}_{30} - \mu}{\sigma/\sqrt{n}}\right| > \frac{1}{\sqrt{6}/\sqrt{30}}\right)$$

$$\doteq P(|Z| > 2.236) \doteq 0.02532。$$

雖然一次觀測之標準差為 $\sqrt{6} \doteq 2.45$, 但當樣本數為30, 樣本平均要偏離真實值一單位之機率並不大。這是為什麼很多時候, 以樣本平均來估計真實值。

如果是利用柴比雪夫不等式, 則得

$$P(|\bar{X}_{30} - \mu| > 1) \leq \frac{\sigma^2/30}{1^2} = \frac{6}{30} = 0.2。$$

雖然這是一“正確的”不等式, 而由中央極限定理所得到的 $P(|\bar{X}_{30} - \mu| > 1)$ 只是一近似值, 但由 0.02532 約只為 0.2 之八分之一倍, 從實用的觀點, 顯然會採前者。

例2.13 由定理2.3-(iii)(取 $Y_n \equiv a$)知, 在定理2.2中,

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow[n \rightarrow \infty]{d} \sigma\mathcal{N}(0, 1) = \mathcal{N}(0, \sigma^2)。$$

必須一提的定理還有不少。如**機率收斂會導致分佈收斂**, 即

$$(2.21) \quad X_n \xrightarrow[n \rightarrow \infty]{p} X \Rightarrow X_n \xrightarrow[n \rightarrow \infty]{d} X。$$

但若存在一常數 a , 使得 $P(X = a) = 1$, 則二者等價。即

$$(2.22) \quad X_n \xrightarrow[n \rightarrow \infty]{p} a \Leftrightarrow X_n \xrightarrow[n \rightarrow \infty]{d} a。$$

又設 g 為一連續函數, 則若 $X_n \xrightarrow[n \rightarrow \infty]{d} X$, 便有 $g(X_n) \xrightarrow[n \rightarrow \infty]{d} g(X)$ (習題第15題給出本結果之一些應用); 若 $X_n \xrightarrow[n \rightarrow \infty]{p} X$, 便有 $g(X_n) \xrightarrow[n \rightarrow \infty]{p} g(X)$ 。底下定理也常會用到。

定理2.4 設 b 為一常數, 而 $n \rightarrow \infty$ 時, $a_n \rightarrow \infty$, 且 $a_n(Z_n - b) \xrightarrow[n \rightarrow \infty]{d} X$ 。又設 g 為一可微的實函數, 且 g' 在 b 連續。則

$$(2.23) \quad a_n(g(Z_n) - g(b)) \xrightarrow[n \rightarrow \infty]{d} g'(b)X。$$

例2.14 在例2.5我們曾指出: 若 $\text{Var}(S_n^2) \rightarrow 0$, 則 $S_n^2 \xrightarrow[n \rightarrow \infty]{p} \sigma^2$ 。因此(習題第15題) $\sigma/S_n \xrightarrow[n \rightarrow \infty]{p} 1$ 。在中央極限定理裡, 若只知期望值 μ , 而不知變異

數 σ^2 之值, 則可以 S_n 取代 σ :

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} = \frac{\sigma}{S_n} \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

弱大數法則是說 n 很大時, $\bar{X}_n - \mu$ 會機率收斂至0。但若將 $\bar{X}_n - \mu$ 放大, 即乘上 $\sqrt{n}$, 便趨近至常態分佈。 $\sqrt{n}$ 是一適當的倍數, 若乘上位階(order)小於 $\sqrt{n}$ 者, 仍很小, 若乘上位階大於 $\sqrt{n}$ 者, 會使其過大(趨近至無限大)。如果以(2.6)式來看, $S_n - n\mu$ 除以 n 會機率收斂至0, 不再隨機了。但若除以 $\sqrt{n}$, 即得 $(S_n - n\mu)/\sqrt{n}$, 便趨近至一仍為隨機的分佈。

大數法則可用來支持**頻率對機率的解釋**(frequentistic interpretation of probability)。某事件發生的機率為 p 是什麼意思? 假設此事件是可以重複的觀測(見例2.1), 如反覆地投擲一銅板, 令

$$X_i = \begin{cases} 1, & \text{第} i \text{次出現正面,} \\ 0, & \text{其他。} \end{cases}$$

又設 $X_1, \dots, X_n$ 為i.i.d.之隨機變數, 令 $p = E(X)$ 。則

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

表出現正面的相對頻率。弱大數法則指出, n 很大時, 此相對頻率有相當大的機率會接近 p 。換句話說, 若投擲數夠多, $\bar{X}_n$ “差不多”可當做 p 之一很好的近似值。以這種頻率的觀點來解釋機率是很自然的。當然對不能重複觀測的事件, 如氣象局宣佈: 明日降雨機率為0.3, 就很難清楚解釋機率為0.3的意義了。

4.3 常態分佈之隨機樣本

由於常態分佈的重要性, 本節我們專門討論關於常態分佈之隨機樣本的一些性質。仍以 $\bar{X}_n, S_n^2$ 分別表樣本平均及樣本變異數。

定理3.1 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。則 $\bar{X}_n$ 與 S_n^2 獨立, 且 $\bar{X}_n$ 有 $\mathcal{N}(\mu, \sigma^2/n)$ 分佈, $(n-1)S_n^2/\sigma^2$ 有 χ_{n-1}^2 分佈。

證明. 我們以數學歸納法來證明。首先對 $\forall n \geq 1$, $\bar{X}_n$ 有 $\mathcal{N}(\mu, \sigma^2/n)$ 分佈是很明顯的。若 $n = 2$, 則 $\bar{X}_2 = (X_1 + X_2)/2$, 且經簡單的化簡可得 $S_2^2 = (X_1 - X_2)^2/2$ 。因 $(X_1 - X_2)/\sqrt{2}$ 有 $\mathcal{N}(0, \sigma^2)$ 分佈, 故 S_2^2/σ^2 有 χ_1^2 分佈。最後因 $\text{Cov}(X_1 + X_2, X_1 - X_2) = 0$, 故 $X_1 + X_2$ 與 $X_1 - X_2$ 獨立(亦見第三章例3.4), 因此 $\bar{X}_2$ 與 S_2^2 獨立。

假設對某 $n \geq 2$, $(n-1)S_n^2/\sigma^2$ 有 χ_{n-1}^2 分佈, 且 $\bar{X}_n$ 與 S_n^2 獨立, 我們要推導出此結果對樣本數為 $n+1$ 亦成立。我們有下述二關係式(見第4.1節習題第5題):

$$(3.1) \quad \bar{X}_{n+1} = (n\bar{X}_n + X_{n+1})/(n+1),$$

$$(3.2) \quad nS_{n+1}^2 = (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2。$$

因 $\bar{X}_n$ 與 X_{n+1} 獨立, 且分別有 $\mathcal{N}(\mu, \sigma^2/n)$ 及 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 故 $X_{n+1} - \bar{X}_n$ 有 $\mathcal{N}(0, ((n+1)/n)\sigma^2)$ 分佈, 因此 $(n/(n+1))(X_{n+1} - \bar{X}_n)^2$ 之分佈為一 $\mathcal{N}(0, \sigma^2)$ 隨機變數之平方。又因 $\bar{X}_n$ 與 S_n^2 獨立, 且 X_{n+1} 亦與 S_n^2 獨立, 故(3.2)式右側兩項獨立, 因此 nS_{n+1}^2/σ^2 有 χ_n^2 分佈。其次再度因 $\bar{X}_n$ 與 S_n^2 獨立, 故由(3.1)式可看出 $\bar{X}_{n+1}$ 與 S_n^2 獨立, 且因

$$\text{Cov}(n\bar{X}_n + X_{n+1}, X_{n+1} - \bar{X}_n) = \sigma^2 - n\sigma^2/n = 0,$$

故 $\bar{X}_{n+1}$ 與 $X_{n+1} - \bar{X}_n$ 亦獨立。由此即得 $\bar{X}_{n+1}$ 與 S_{n+1}^2 獨立。證畢。

在上述證明中, 我們主要用到常態分佈的幾個特性: 其一為若隨機變數 X 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 則 $(X - \mu)^2/\sigma^2$ 有 χ_1^2 分佈(見2.3.6節)。其二為設 X, Y 為二獨立的隨機變數, 且皆有常態分佈, 令 $Z_1 = a_1X + b_1Y$, $Z_2 = a_2X + b_2Y$, $a_1b_2 - a_2b_1 \neq 0$, 則 (Z_1, Z_2) 有二變數常態分佈(見3.6節之性質(vii))。其三為若 (U, V) 有二變數常態分佈, 則 U 與 V 獨立, 若且唯若 $\text{Cov}(U, V) = 0$ (見3.6節之性質(iv))。

關於 $\bar{X}_n$ 與 S_n^2 獨立之證法仍有不少, 不同的作者提供他們認為易於為初學者所接受的證法。在稍後4.4節我們將介紹另一藉由多變量常態隨機變數的證法。

註3.1 附帶一提, 對gamma分佈, $\bar{X}_n$ 與 S_n 間亦有一特殊關係。即設 $n \geq 3$, 且 $X_1, \dots, X_n$ 為一組由非負且非退化的分佈所產生之隨機樣本。則樣本變異係數 $S_n/\bar{X}_n$ 與 $\bar{X}_n$ 獨立, 若且唯若 $X_1, \dots, X_n$ 之共同分佈為gamma分佈。即 $S_n/\bar{X}_n$ 與 $\bar{X}_n$ 獨立, 為gamma分佈之一特性。

常態分佈與卡方分佈關係密切。利用獨立的卡方分佈之和, 仍為卡方分佈, 只是自由度要相加, 可得下述結果。

定理3.2 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。則

$$(3.3) \quad \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi_n^2。$$

下例為一大樣本下的結果。

例3.1 設有一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之樣本數為 n 之隨機樣本, 其樣本變異數以 S_n^2 表之。則先由定理3.1, 得

$$(3.4) \quad V_n = \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2。$$

再由例2.9, 得

$$(3.5) \quad \frac{V_n - (n-1)}{\sqrt{2(n-1)}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

將(3.5)式中之 V_n 以 $(n-1)S_n^2/\sigma^2$ 取代, 得

$$(3.6) \quad \frac{\sqrt{n-1}(S_n^2 - \sigma^2)}{\sigma^2\sqrt{2}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

(3.4)式給出 S_n^2 之確切的分佈: $S_n^2 \sim \sigma^2 \chi_{n-1}^2 / (n-1)$ 。(3.6)式則指出 n 較大時, S_n^2 的分佈近似於 $\mathcal{N}(\sigma^2, 2\sigma^4/(n-1))$ 。

利用定理2.4, 取 $Z_n = S_n^2$, $b = \sigma^2$, $a_n = \sqrt{n-1}/(\sigma^2\sqrt{2})$, $g(x) = \sqrt{x}$, 則 $g(Z_n) = S_n$, $g(b) = \sigma$, $g'(b) = 1/(2\sigma)$, 且得

$$(3.7) \quad a_n(g(Z_n) - g(b)) = \frac{\sqrt{n-1}(S_n - \sigma)}{\sigma^2\sqrt{2}} \xrightarrow[n \rightarrow \infty]{d} \frac{1}{2\sigma} \mathcal{N}(0, 1)。$$

即 n 較大時, S_n 的分佈近似於 $\mathcal{N}(\sigma, \sigma^2/(2(n-1)))$ 。

底下我們介紹在統計分析裡兩個很重要的分佈。首先是英國統計學家高賽特(Gosset, 1876-1937)在二十世紀初期, 以筆名Student最先引進的 **t 分佈**。此分佈又稱Student distribution, 或Student's t distribution。關於高賽特的一些事蹟, 可參考葉偉文譯(2001)第三章。

定義3.1 設 X, Y 為二獨立的隨機變數, 分別有 $\mathcal{N}(0, 1)$ 及 χ_n^2 分佈。令

$$(3.8) \quad T = X/\sqrt{Y/n},$$

則稱 T 有**自由度 n 之 t 分佈**(t distribution with k degrees of freedom), 以 $\mathcal{T}_n$ 表之。

我們求上述 T 之p.d.f.如下。首先 (X, Y) 之聯合p.d.f.為

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\Gamma(n/2)2^{n/2}} y^{n/2-1} e^{-y/2}, x \in R, y > 0。$$

若令 $S = Y$, 由

$$t = x/\sqrt{y/n}, \quad s = y,$$

得

$$x = t\sqrt{s}/\sqrt{n}, \quad y = s,$$

且

$$J = \begin{vmatrix} \sqrt{s}/\sqrt{n}, & t/(2\sqrt{s}\sqrt{n}) \\ 0 & 1 \end{vmatrix} = \frac{\sqrt{s}}{\sqrt{n}}。$$

故得

$$f_{T,S}(t, s) = \frac{\sqrt{s}}{\sqrt{n}\sqrt{2\pi}} e^{-t^2 s/(2n)} \frac{1}{\Gamma(n/2)2^{n/2}} s^{n/2-1} e^{-s/2}, t \in R, s > 0。$$

由上式便可求出 T 之邊際p.d.f.:

$$(3.9) \quad f(t) = \int_0^\infty f_{T,S}(t, s) ds$$

$$\begin{aligned}
&= \frac{1}{\sqrt{n\pi}\Gamma(n/2)2^{(n+1)/2}} \int_0^\infty s^{(n-1)/2} e^{-((1+t^2/n)/2)s} ds \\
&= \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} (1+t^2/n)^{-(n+1)/2}, t \in R.
\end{aligned}$$

此處用到

$$\begin{aligned}
(3.10) \quad &\int_0^\infty s^{(n-1)/2} e^{-((1+t^2/n)/2)s} ds \\
&= \Gamma((n+1)/2) 2^{(n+1)/2} (1+t^2/n)^{-(n+1)/2}.
\end{aligned}$$

這只要藉助gamma分佈之p.d.f.的積分爲1便可得證, 證明留在習題中(第1題)。

讀者是否注意到, 在 t 分佈裡, 若自由度 $n=1$, 則得到柯西分佈 $C(0, 1)$ 。即 $C(0, 1)$ 分佈爲 t 分佈之一特例。

所有 t 分佈之p.d.f.皆爲偶函數, 但 t 分佈之動差母函數並不存在, 它的動差也並非都存在。設 T 有 $\mathcal{T}_n$ 分佈, 則利用瑕積分收斂發散之判別法, 可驗證對整數 l , $E(T^l)$ 存在, 若且唯若, $-1 < l < n$ (即 $0 \leq l \leq n-1$)。注意因 T 可取負值, 故我們不考慮 l 不爲整數時的 $E(T^l)$ 。這是爲了避免混淆, 我們不考慮諸如 $(-16)^{1/2}$, 以及 $t < 0$ 且 l 不爲整數時的 t^l 。故得 T 的 $n-1$ 次動差存在。但 $\mathcal{T}_1$ 分佈的期望值及 $\mathcal{T}_2$ 分佈的變異數皆不存在。又

$$(3.11) \quad E(T) = 0, n \geq 2,$$

$$(3.12) \quad \text{Var}(T) = \frac{n}{n-2}, n \geq 3.$$

上二式可經由 T 之p.d.f.直接求得(習題第3題)。事實上 T 之所有動差(只要存在)皆可如下得到。

對奇數 $1 \leq 2r-1 \leq n-1$, 因 T 之p.d.f.爲偶函數, 故

$$(3.13) \quad E(T^{2r-1}) = 0.$$

對偶函數 $2 \leq 2r \leq n-1$, 首先利用(3.4)式, 及 X 與 Y 獨立, 得

$$(3.14) \quad E(T^{2r}) = E(X^{2r})E(Y^{-r})n^r.$$

又由第二章(3.27)式及(3.14)式, 得

$$E(X^{2r}) = \prod_{i=1}^r (2i-1),$$

$$E(Y^{-r}) = \frac{2^{-r} \Gamma(n/2 - r)}{\Gamma(n/2)}.$$

將上二式代入(3.14)式, 得

$$(3.15) \quad E(T^{2r}) = \frac{\prod_{i=1}^r (2i-1) \cdot (n/2)^r \cdot \Gamma(n/2 - r)}{\Gamma(n/2)}.$$

上式又可改寫為(習題第4題)

$$(3.16) \quad E(T^{2r}) = \frac{\Gamma(r+1/2)\Gamma(n/2-r)n^r}{\Gamma(1/2)\Gamma(n/2)}.$$

例3.2 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。則 $(\bar{X}_n - \mu)/(S_n/\sqrt{n})$ 有 $\mathcal{T}_{n-1}$ 分佈。

證明. 首先由定理3.1 知

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1), \quad (n-1)S_n^2/\sigma^2 \sim \chi_{n-1}^2,$$

且二者獨立。再由定義3.1,

$$\frac{(\bar{X}_n - \mu)/(\sigma/\sqrt{n})}{\sqrt{(n-1)S_n^2/(\sigma^2(n-1))}} = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \sim \mathcal{T}_{n-1}.$$

前面提過, $\mathcal{T}_1$ 分佈即 $\mathcal{C}(0, 1)$ 分佈, 故當 $n=2$, $(\bar{X}_2 - \mu)/(S_2/\sqrt{2})$ 有 $\mathcal{C}(0, 1)$ 分佈。

當 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本, 我們知道

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

有 $\mathcal{N}(0, 1)$ 分佈。若 σ 已知, μ 未知, 則若觀測到 $X_1, \dots, X_n$, 便有 $\bar{X}_n$ 之值, 如此可對 μ 做一些推論。但若 σ 也未知(這是很可能的), 例3.2告訴我們, 由於

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}$$

的分佈知道, 因此可以此為基礎(注意: 若 $X_1, \dots, X_n$ 知道, 則 S_n 也就可求出), 對 μ 做一些推論。也就是以 S_n 取代 σ (大家還記得吧, 在例2.14我們曾指出 $\sigma/S_n \xrightarrow[n \rightarrow \infty]{p} 1$), 而分佈也由 $\mathcal{N}(0, 1)$ 換成 $\mathcal{T}_{n-1}$ 。此結果在第八章及第九章等會屢屢用到。

靈敏的讀者是否已猜出, 當 n 很大時, $\mathcal{T}_{n-1}$ 分佈 ($\mathcal{T}_n$ 分佈也一樣) 應接近 $\mathcal{N}(0, 1)$ 分佈。這是對的。即

$$(3.17) \quad \mathcal{T}_{n-1} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)。$$

例2.14已指出這點。事實上由(3.8)式, 因 $Y/n \xrightarrow[n \rightarrow \infty]{p} 1$ (見(2.15)式), 因此 $\sqrt{Y/n} \xrightarrow[n \rightarrow \infty]{p} 1$ (4.2節習題第16題), 最後利用定理2.3(iv), 即得

$$T = \frac{X}{\sqrt{Y/n}} \xrightarrow[n \rightarrow \infty]{d} X \sim \mathcal{N}(0, 1)。$$

利用上述結果, 當 n 較大時, 可以 $\mathcal{N}(0, 1)$ 分佈當做 $\mathcal{T}_n$ 分佈之近似。對 $\forall 0 < p < 1$, 令 z_p 及 $t_{p,n}$ 分別表 $\mathcal{N}(0, 1)$ 分佈及 $\mathcal{T}_n$ 分佈之 p 分位數 ($P(\mathcal{T}_n \leq t_{p,n}) = p$)。即 n 很大時

$$(3.18) \quad t_{p,n} \doteq z_p。$$

以 $n = 30$ 為例。由附表4及6得 $t_{0.60,30} \doteq 0.256$, $t_{0.70,30} \doteq 0.530$, $t_{0.80,30} \doteq 0.854$, $t_{0.85,30} \doteq 1.055$, $t_{0.90,30} \doteq 1.310$, $t_{0.95,30} \doteq 1.697$, 而 $z_{0.60} \doteq 0.2533$, $z_{0.70} \doteq 0.5244$, $z_{0.80} \doteq 0.8418$, $z_{0.85} \doteq 1.0365$, $z_{0.90} \doteq 1.2814$, $z_{0.95} \doteq 1.645$ 。上述數據顯示 $t_{p,30}$ 與 z_p 算是很接近。圖3.1給出 $\mathcal{T}_1$, $\mathcal{T}_2$, $\mathcal{T}_5$ 及 $\mathcal{N}(0, 1)$ 分佈 p.d.f. 之圖形。

其次我們介紹 F 分佈。 F 分佈是為紀念著名的統計學家費雪(R.A. Fisher, 1890-1962, 現代統計學的創始者之一)而得名(F為Fisher的第一個字母)。是為求二變異數的比值而產生。

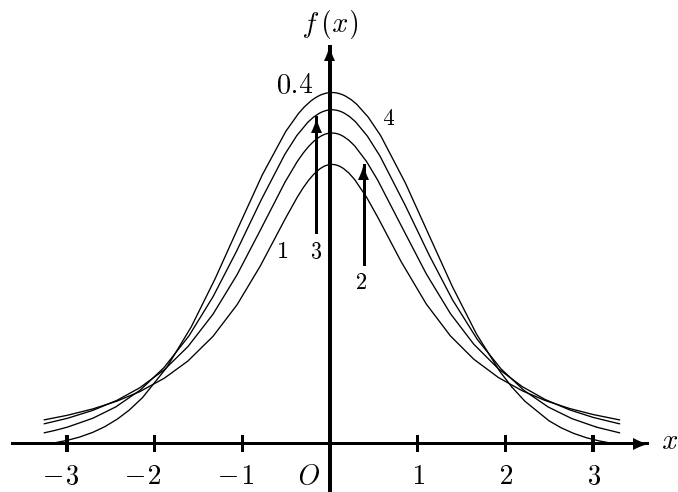


圖3.1 1,2,3,4分別對應 T_1 , T_2 , T_5 及 $\mathcal{N}(0,1)$ 分佈p.d.f.之圖形

定義3.2 設 U, V 為二獨立的隨機變數, 分別有 χ_n^2 及 χ_m^2 分佈。令

$$(3.19) \quad F = \frac{U/n}{V/m},$$

則 F 稱有**自由度 n 及 m 之 F 分佈**(F distribution with n and m degrees of freedom), 此分佈以 $\mathcal{F}_{n,m}$ 表之。

利用變數代換, 可得(習題第7題) F 之p.d.f.為

$$(3.20) \quad g(x) = \frac{(n/m)^{n/2} x^{(n-2)/2}}{B(n/2, m/2)(1 + (n/m)x)^{(n+m)/2}}, x > 0,$$

其中 $B(\cdot, \cdot)$ 為beta函數(定義見第二章(3.33)式)。又

$$(3.21) \quad E(U^r) = \frac{(m/n)^r \Gamma(n/2 + r) \Gamma(m/2 - r)}{\Gamma(n/2) \Gamma(m/2)}, -\frac{n}{2} < r < \frac{m}{2},$$

其中 r 並不須為整數。由此即得

$$(3.22) \quad E(F) = \frac{m}{m-2}, m \geq 3,$$

$$(3.23) \quad \text{Var}(F) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}, m \geq 5。$$

(3.21)式只要注意在定義3.2中, U 與 V 獨立, 因此

$$E(F^r) = \left(\frac{m}{n}\right)^r E(U^r) E(V^{-r}),$$

便可得證, 我們留在習題(第8題)中讓各位自行完成。當 X 有 $\mathcal{F}_{n,m}$ 分佈, $1/X$ 亦有 F 分佈, 只是二參數交換。又經過簡單的變換, F 分佈可轉換為beta分佈。此二性質皆留在習題(第7題)中。

底下給幾個 F 分佈p.d.f.之圖形。

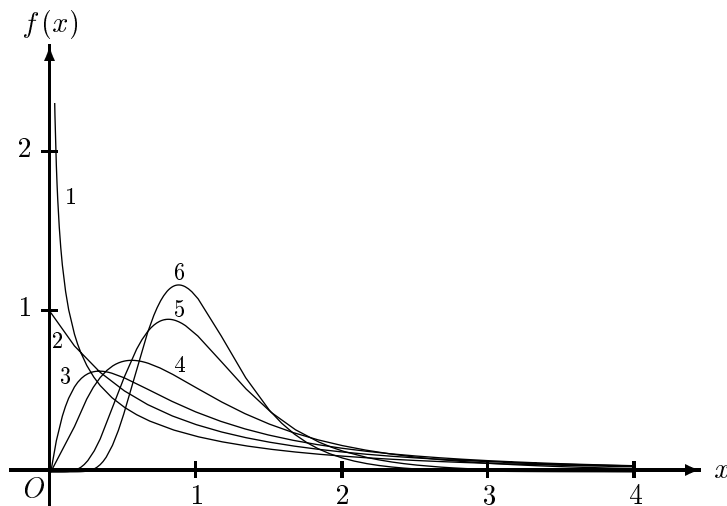


圖3.2 1,2,3,4,5,6分別對應 $\mathcal{F}_{1,6}$, $\mathcal{F}_{2,4}$, $\mathcal{F}_{4,4}$, $\mathcal{F}_{6,10}$, $\mathcal{F}_{16,24}$, $\mathcal{F}_{30,32}$ 分佈p.d.f.之圖形

例3.3 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。由定理3.1,

$$\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi_1^2, \quad \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2,$$

且二者獨立, 故由定義3.2,

$$\frac{(\bar{X}_n - \mu)^2/(\sigma^2/n)}{(n-1)S_n^2/(\sigma^2(n-1))} = \frac{(\bar{X}_n - \mu)^2}{S_n^2/n} \sim \mathcal{F}_{1,n-1}。$$

與例3.2比較, 即得 $\mathcal{T}_n^2 = \mathcal{F}_{1,n}$, 對 $n \geq 1$ 。故 $\mathcal{T}_n^2$ 之分佈為一特別的 F 分佈。

例3.4 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu_X, \sigma_X^2)$ 分佈所產生之隨機樣本, $Y_1, \dots, Y_m$ 為一組由 $\mathcal{N}(\mu_Y, \sigma_Y^2)$ 分佈所產生之隨機樣本。兩組樣本假設相互獨立。設以 S_X^2, S_Y^2 分別表 $\{X_i\}$ 及 $\{Y_i\}$ 之樣本變異數。則由定理3.1知, $(n-1)S_X^2/\sigma_X^2$ 有 χ_{n-1}^2 分佈, $(m-1)S_Y^2/\sigma_Y^2$ 有 χ_{m-1}^2 分佈。因此由定義3.2,

$$\frac{(n-1)S_X^2/((n-1)\sigma_X^2)}{(m-1)S_Y^2/((m-1)\sigma_Y^2)} = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim \mathcal{F}_{n-1, m-1}。$$

另外, 若 $m \rightarrow \infty$, 則仍由例2.5,

$$\frac{\sigma_Y^2}{S_Y^2} \xrightarrow[m \rightarrow \infty]{p} 1,$$

故由定理2.3,

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} = \frac{\sigma_Y^2}{S_Y^2} \frac{S_X^2}{\sigma_X^2} \xrightarrow[m \rightarrow \infty]{d} \frac{S_X^2}{\sigma_X^2} \sim \frac{\chi_{n-1}^2}{n-1}。$$

此亦顯示

$$(3.24) \quad \mathcal{F}_{n,m} \xrightarrow[m \rightarrow \infty]{d} \frac{\chi_n^2}{n}。$$

上述結果由 F 分佈之定義亦可看出: 由

$$F = \frac{U/n}{V/m},$$

$U \sim \chi_n^2, V \sim \chi_m^2$, 得(見(2.15)式)

$$\frac{V}{m} \xrightarrow[m \rightarrow \infty]{p} 1,$$

再由定理2.3(iv), 即得(3.24)式。同理可得

$$(3.25) \quad \mathcal{F}_{n,m} \xrightarrow[n \rightarrow \infty]{d} \frac{m}{\chi_m^2}。$$

現令 $f_{p,n,m}$ 滿足 $P(\mathcal{F}_{n,m} \leq f_{p,n,m}) = p$, $f_{p,n,m}$ 即為 $\mathcal{F}_{n,m}$ 分佈之第 p 分位數。則(3.24)及(3.25)式分別導致對 $0 < p < 1$,

$$(3.26) \quad f_{p,n,m} \doteq \frac{\chi_{p,n}^2}{n}, m \text{ 很大},$$

$$(3.27) \quad f_{p,n,m} \doteq \frac{m}{\chi_{1-p,m}^2}, n \text{ 很大},$$

其中如前 $\chi_{p,n}^2$ 為 χ_n^2 分佈之第 p 分位數。例如, 由附表7, 得對 $n = 1, \dots, 5$, $f_{0.9,n,60}$ 分別約為2.79, 2.39, 2.18, 2.04, 1.95。又由附表5, 得對 $n = 1, \dots, 5$, $\chi_{0.9,n}^2$ 分別約為2.71, 4.61, 6.25, 7.78, 9.24, 因此 $\chi_{0.9,n}^2/n$ 分別約為2.71, 2.305, 2.083, 1.945, 1.848。與對應的 $f_{0.9,n,60}$ 相比, 誤差並不太大。

要注意的是, 附表7並非如表面上看起來那麼不完整, 雖然只對一些較小的 α , 給出右端機率, 但利用若 X 有 $\mathcal{F}_{n,m}$ 分佈, 則 $1/X$ 有 $\mathcal{F}_{m,n}$ 分佈, 有些左端機率亦可得到。例如, 因

$$P(\mathcal{F}_{4,6} \geq 4.53) \doteq 0.05,$$

又 $1/4.53 \doteq 0.220$, 故

$$P(\mathcal{F}_{6,4} \leq 0.220) \doteq 0.05。$$

讀者必須熟悉利用附表來求與卡方分佈、 t 分佈及 F 分佈之機率有關的值。底下給一例。

例3.5 利用附表5-7可求出下述各值。

- (i) 設 $X \sim \chi_{15}^2$, 則 $P(7.26 \leq X \leq 22.31) \doteq 0.90 - 0.05 = 0.85$ 。
- (ii) 設 $X \sim \chi_{23}^2$, 且 $P(X \leq b) = 0.90$, 則 $b \doteq 32.01$ 。
- (iii) 設 $X \sim \chi_6^2$, 則 $P(X/(1+X) \geq 11/16) = P(X \geq 2.2) \doteq 1 - 0.1 = 0.90$ 。
- (iv) 設 $X \sim \mathcal{T}_{13}$, 則 $P(0.87 \leq X \leq 2.65) \doteq 0.99 - 0.80 = 0.19$ 。
- (v) 設 $X \sim \mathcal{T}_{26}$, 且 $P(X \leq b) = 0.60$, 則 $b \doteq 0.256$ 。
- (vi) 設 $X \sim \mathcal{T}_{23}$, 且 $P(|X| \geq b) = 0.02$, 則 $b \doteq 2.50$ 。
- (vii) 設 $X \sim \mathcal{F}_{7,12}$, 則 $P(2.91 \leq X \leq 8) \doteq 0.050 - 0.001 = 0.049$ 。
- (viii) 設 $X \sim \mathcal{F}_{20,8}$, 則 $P(1/X > 0.25) = P(X < 4) = 1 - P(X \geq 4) \doteq 1 - 0.025 = 0.975$ 。
- (xi) 設 $X \sim \mathcal{F}_{12,20}$, 則 $1/X \sim \mathcal{F}_{20,12}$, 則 $P(X \leq 1/3.07) = P(1/X \geq 3.07) = 0.025$ 。

本節最後給兩個應用的例子。

例3.6 某校學生要求圖書館晚上10時至12時亦開放。圖書館表示若該時段到館人數的平均值，真的超過30人便開放。經測試開放一個月，其中有一天是國定假日晚上不開放，得到29筆該時段到館人數：

30, 32, 29, 9, 25, 25, 32, 41, 49, 44, 37, 12, 13, 44, 31,
45, 55, 52, 16, 26, 39, 40, 46, 58, 38, 9, 15, 39, 46。

如何決定呢？這類問題第八章我們會詳加討論。此處以直觀來看怎樣做是可以接受的。一個辦法是假設每日於該時段到館人數有 $\mathcal{N}(\mu, \sigma^2)$ 分佈，我們想看在 $\mu = 30$ 的假設下(σ^2 仍設為未知)，會得到上述數據是否合理。要注意的是，人數其實不可能有常態分佈，此因人數必為整數，且不會取負值。不過我們只是以常態做為近似的分佈。

經求出樣本平均及樣本標準差分別約為

$$\bar{X}_{29} \doteq 33.69, \quad S_{29} \doteq 13.98,$$

由此得

$$T = \frac{\bar{X}_{29} - \mu}{S_{29}/\sqrt{n}} \doteq \frac{33.69 - 30}{13.98/\sqrt{29}} \doteq 1.42。$$

因 T 有 $\mathcal{T}_{28}$ 分佈，由附表6得

$$P(T \geq 1.313) \doteq 0.10, \quad P(T \geq 1.701) \doteq 0.05。$$

再利用內插法得

$$P(T \geq 1.42) \doteq 0.10 - \frac{1.42 - 1.313}{1.701 - 1.313} \cdot (0.10 - 0.05) \doteq 0.086$$

(事實上， $P(T \geq 1.42)$ 約為0.0833)。

所以在 $\mu = 30$ 之假設下，僅約有0.086的機率會得到如此大的 T 值。似乎有夠強的證據支持 $\mu > 30$ 。當然還要看主事者的態度。如果圖書館相當重視服務，可能便接受延長開放時間。如果圖書館覺得經費支出為主要的考慮，便可能尚無法被說服需延長開放時間。

例3.7 假設各學生之成績相互獨立，且在每班中學生之成績有相同的常態分佈，但在不同班之參數可能不同。今欲判定兩班學生(分別有31及36位學生)成績之變異是否相同。經計算出兩班成績之樣本變異數 $S_X^2 = 196.1$, $S_Y^2 = 690.0$ 。由例3.4, 若 $\sigma_X^2 = \sigma_Y^2$, 則

$$F = \frac{S_X^2}{S_Y^2} \sim \mathcal{F}_{30,25}。$$

先計算出 $S_X^2/S_Y^2 = 196.1/690.0 \doteq 0.284$ 。再由附表7得

$$P\left(\frac{1}{F_{30,25}} > 3.33\right) = P(F_{25,30} > 3.33) = 0.001,$$

即

$$P(F_{30,25} \leq \frac{1}{3.33}) \doteq P(F_{30,25} \leq 0.30) \doteq 0.001。$$

故得

$$P(S_X^2/S_Y^2 \leq 0.284) < P(F_{30,25} \leq 0.30) \doteq 0.001。$$

在 $\sigma_X^2 = \sigma_Y^2$ 之下，會得到如此小的 S_X^2/S_Y^2 之機率是很低的(小於0.001)。故可合理判定 $\sigma_X^2 \neq \sigma_Y^2$ 。

4.4 多變量常態分佈

在實際的問題中，單變數的常態分佈常出現，此因有中央極限定理，故若有許多獨立的隨機變數，則經常態化後，其和往往趨近至常態分佈。若我們有好幾個變數，則多維的中央極限定理便會導致**多維常態分佈**，或說**多變量常態分佈**。雖然我們在3.6節已介紹過二變數常態分佈，但由於多變量常態分佈之重要性，因此本節來討論多變量的情況。

在討論多變量常態分佈之前，我們先複習一些需要的**矩陣代數**(matrix algebra)之知識。至於更詳盡的關於機率及統計中所需之矩陣的知識，Graybill(1983)為一本很好的參考書。在本節中，我們所提到之向量及矩陣，其元素皆為實數。

設有變數 $x_1, \dots, x_n$, 並考慮下述變換:

$$(4.1) \quad \begin{aligned} y_1 &= a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n, \\ y_2 &= a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n, \\ &\vdots \\ y_n &= a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n, \end{aligned}$$

其中 $a_{ij}, 1 \leq i, j \leq n$, 為常數。若 $\sum_{i=1}^n y_i^2 = \sum_{i=1}^n x_i^2$, 則稱為一由 $x_1, \dots, x_n$ 至 $y_1, \dots, y_n$ 之**正交變換**(orthogonal transformation)。現對任一 $m \times n$ 的矩陣 $B = (b_{ij})$, 令 B' 表 B 之**轉置**(transpose), 即 $B' = (b_{ji})$ 。**列向量**(為 $1 \times n$ 的矩陣)之轉置, 就成為一 $n \times 1$ 的**行向量**。若令

$$\begin{aligned} \mathbf{x} &= (x_1, \dots, x_n), \\ \mathbf{y} &= (y_1, \dots, y_n), \\ \mathbf{A} &= (a_{ij})_{1 \leq i, j \leq n}, \end{aligned}$$

則(4.1)式之變換可以下述矩陣及向量來表示:

$$(4.2) \quad \mathbf{y} = \mathbf{x}\mathbf{A}。$$

因此(4.1)式定義出一正交變換, 若且唯若

$$\mathbf{x}\mathbf{x}' = \mathbf{y}\mathbf{y}' = \mathbf{x}\mathbf{A}\mathbf{A}'\mathbf{x}'。$$

即若且唯若 $\mathbf{A}\mathbf{A}' = \mathbf{I}_n$, 其中 $\mathbf{I}_n$ 為 $n \times n$ 之**單位矩陣**。 $n \times n$ 之矩陣 $\mathbf{A}$ 稱為一**正交矩陣**, 若其滿足

$$(4.3) \quad \mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}_n。$$

此時 $\mathbf{A}^{-1} = \mathbf{A}'$, 且 $\det \mathbf{A} = 1$ 。對一矩陣 $\mathbf{A}$ 我們以 $\det \mathbf{A}$ 或 $|\mathbf{A}|$ 表其**行列式**(determinant)。

設 $\mathbf{A}$ 為一 $m \times n$ 之矩陣, 以 $r(\mathbf{A})$ 表 $\mathbf{A}$ 之**秩**(rank), 則 $r(\mathbf{A}) \leq \min\{m, n\}$ 。而若 $r(\mathbf{A}) = \min\{m, n\}$, 則稱 $\mathbf{A}$ 為**非奇異的**(nonsingular)或**全秩**(full rank), 否則 $\mathbf{A}$ 稱為**奇異的**(singular)。

其次我們介紹**正定**(positive definite)矩陣。首先設 A 為一 $n \times n$ 之矩陣,則下述行列式

$$(4.4) \quad \det(A - \lambda I_n) = |A - \lambda I_n|,$$

其中 λ 為一**純量**(scalar), 為一 λ 之 n 次多項式, 稱為 A 之**特徵多項式**(characteristic polynomial)。而

$$(4.5) \quad |A - \lambda I_n| = 0$$

之 n 個根, 稱為 A 之**特徵根**(characteristic roots)或**固有值**(eigenvalues)。又 λ 為 A 之一**特徵根**, 若且唯若存在一向量 $x \neq 0$, 使得 $xA = \lambda x$, x 稱為**特徵向量**(characteristic vector)。

定義4.1 一 $n \times n$ 之**對稱矩陣** A (即若 $A = (a_{ij})$, 則 $a_{ij} = a_{ji}$)稱為**正定**、**負定**(negative definite)或**半正定**(positive semidefinite), 若且唯若對每一 n 維向量 $x \neq 0$, $xAx' > 0$, < 0 或 ≥ 0 。

定理4.1 一 $n \times n$ 實矩陣為正定、負定或半正定, 若其特徵根 $\lambda_1, \dots, \lambda_n$, 分別滿足 $\lambda_i > 0$, $\lambda_i < 0$, $\lambda_i \geq 0$, $\forall i = 1, \dots, n$ 。

另外, 一 $n \times n$ 之矩陣稱為**冪等**(idempotent), 若 $A^2 = A$ 。

定理4.2 (i) 設有一 $n \times n$ 之**對稱矩陣** A , 則存在一**正交矩陣** B , 使得 BAB' 為**對角矩陣**, 且對角之元素即為 A 之特徵根。而若加上 A 為正定, 則存在一非奇異矩陣 C , 使得 $CAC' = I_n$ 。

(ii) 設 B 為一 $n \times n$ 之矩陣且 $r(B) = r$ 。若 B 為冪等矩陣, 則 B 有 r 個非零特徵根且皆為1; 若 B 為對稱矩陣, 則 B 為冪等之充要條件為 B 有 r 個非零特徵根, 且皆為1。

定理4.3 設有變數 $x_1, \dots, x_n$ 及常數 $a_{11}, \dots, a_{n1}$, 且

$$(4.6) \quad y_1 = a_{11}x_1 + a_{21}x_2 + \dots + a_{n1}x_n \quad \circ$$

則只要 $\sum_{i=1}^n a_{i1}^2 = 1$, 必存在 a_{ij} , $i = 1, \dots, n$, $j = 2, \dots, n$, 使得若令 $y_i = \sum_{j=1}^n a_{ji}x_j$, $i = 2, \dots, n$, 則 $x_1, \dots, x_n$ 至 $y_1, \dots, y_n$ 爲一正交變換。即 $\mathbf{A} = (a_{ij})_{1 \leq i, j \leq n}$ 爲一正交矩陣。

現在我們可以定義多變量常態分佈了。

定義4.2 設 $X_1, \dots, X_n$ 爲 i.i.d. 有 $\mathcal{N}(0, 1)$ 分佈之隨機變數, $\mathbf{A} = (a_{ij})$ 爲一 $n \times m$ 之矩陣, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)$ 爲一列向量。令 $\mathbf{X} = (X_1, \dots, X_n)$, 則

$$(4.7) \quad \mathbf{Y} = (Y_1, \dots, Y_m) = \mathbf{X}\mathbf{A} + \boldsymbol{\mu}$$

稱爲有 m 維常態分佈, 或 m 變量常態分佈, 通稱多變量常態分佈。

由上述定義知, 若 $Y_1, \dots, Y_m$ 有多變量常態分佈, 則其任一子集仍有多變量常態分佈。 $\mathbf{Y}$ 之期望向量爲

$$E(\mathbf{Y}) = (E(Y_1), \dots, E(Y_m)) = E(\mathbf{X}\mathbf{A} + \boldsymbol{\mu}) = E(\mathbf{X})\mathbf{A} + \boldsymbol{\mu} = \boldsymbol{\mu}。$$

共變異數矩陣爲

$$\begin{aligned} \Sigma_{\mathbf{Y}} &= E((\mathbf{Y} - E(\mathbf{Y}))'(\mathbf{Y} - E(\mathbf{Y}))) \\ &= E((\mathbf{X}\mathbf{A})'(\mathbf{X}\mathbf{A})) \\ &= \mathbf{A}'E(\mathbf{X}'\mathbf{X})\mathbf{A} \\ &= \mathbf{A}'\mathbf{A}。 \end{aligned}$$

此處因 $X_1, \dots, X_n$ 爲 i.i.d. 之 $\mathcal{N}(0, 1)$ 隨機變數, 故若 $i = j$, 則 $E(X_i X_j) = 1$, 若 $i \neq j$, 則 $E(X_i X_j) = 0$ 。因此 $E(\mathbf{X}'\mathbf{X}) = \mathbf{I}_n$ 爲一單位矩陣。底下我們求 $\mathbf{Y}$ 之動差母函數 $M_{\mathbf{Y}}(\mathbf{t})$, 其中 $\mathbf{t} = (t_1, \dots, t_m) \in R^m$ 。依定義得

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= E(e^{t\mathbf{Y}'}) \\ &= E(e^{t(\mathbf{A}'\mathbf{X}' + \boldsymbol{\mu}')}) \\ &= e^{t\boldsymbol{\mu}'} E(e^{t\mathbf{A}'\mathbf{X}'})。 \end{aligned}$$

但

$$\mathbf{t}\mathbf{A}'\mathbf{X}' = \left(\sum_{j=1}^m t_j a_{1j}\right)X_1 + \cdots + \left(\sum_{j=1}^m t_j a_{nj}\right)X_n ,$$

故(仍利用 $X_1, \dots, X_n$ 爲 i.i.d. 之 $\mathcal{N}(0, 1)$ 隨機變數)

$$\begin{aligned} E(e^{\mathbf{t}\mathbf{A}'\mathbf{X}'}) &= M_{X_1}\left(\sum_{j=1}^m t_j a_{1j}\right) \cdots M_{X_n}\left(\sum_{j=1}^m t_j a_{nj}\right) \\ &= \exp\left\{\frac{1}{2}\left(\sum_{j=1}^m t_j a_{1j}\right)^2 + \cdots + \frac{1}{2}\left(\sum_{j=1}^m t_j a_{nj}\right)^2\right\} \\ &= \exp\left\{\frac{1}{2}\mathbf{t}\mathbf{A}'\mathbf{A}\mathbf{t}'\right\} . \end{aligned}$$

因此

$$(4.8) \quad M_Y(\mathbf{t}) = \exp\{\mathbf{t}\boldsymbol{\mu}' + \frac{1}{2}\mathbf{t}\boldsymbol{\Sigma}_Y\mathbf{t}'\} ,$$

其中如前

$$(4.9) \quad \boldsymbol{\Sigma}_Y = \mathbf{A}'\mathbf{A} .$$

可看出若知 $\mathbf{Y}$ 之動差母函數有(4.8)式的型式, 則 $\mathbf{Y}$ 之分佈可由期望向量 $\boldsymbol{\mu}$ 及共變數矩陣 $\boldsymbol{\Sigma}_Y$ 所完全決定, 此性質與單變數常態分佈類似。因此我們便以 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 表此分佈。

當 $\mathbf{A}$ 爲一 $n \times n$ 之非奇異矩陣, 即 $|\mathbf{A}| \neq 0$, 底下我們求 $\mathbf{Y}$ 之 p.d.f. f_Y 。首先 $\mathbf{X}$ 之 p.d.f. 可寫成

$$\begin{aligned} f_X(\mathbf{x}) &= (2\pi)^{-n/2} \exp\{-\frac{1}{2}(x_1^2 + \cdots + x_n^2)\} \\ &= (2\pi)^{-n/2} \exp\{-\frac{1}{2}\mathbf{x}\mathbf{x}'\} . \end{aligned}$$

由(4.7)式得

$$\mathbf{X} = (\mathbf{Y} - \boldsymbol{\mu})\mathbf{A}^{-1},$$

故

$$\begin{aligned} f_Y(\mathbf{y}) &= f_X((\mathbf{y} - \boldsymbol{\mu})\mathbf{A}^{-1})|\mathbf{A}^{-1}| \\ &= (2\pi)^{-n/2} \exp\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})\mathbf{A}^{-1}(\mathbf{A}^{-1})'(\mathbf{y} - \boldsymbol{\mu})'\}|\mathbf{A}|^{-1} \\ &= (2\pi)^{-n/2} \exp\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})\boldsymbol{\Sigma}_Y^{-1}(\mathbf{y} - \boldsymbol{\mu})'\}|\boldsymbol{\Sigma}_Y|^{-1/2}, \end{aligned}$$

此處用到

$$\mathbf{A}^{-1}(\mathbf{A}^{-1})' = (\mathbf{A}'\mathbf{A})^{-1} = \boldsymbol{\Sigma}_Y^{-1}, \quad |\mathbf{A}^{-1}| = |\mathbf{A}|^{-1} = |\boldsymbol{\Sigma}_Y|^{-1/2}.$$

因此 $\mathbf{Y}$ 之p.d.f.爲

$$(4.10) \quad f_Y(\mathbf{y}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}_Y|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})\boldsymbol{\Sigma}_Y^{-1}(\mathbf{y} - \boldsymbol{\mu})'\right\}, \mathbf{y} \in R^n.$$

一個有 n 維常態分佈之隨機變數，若其p.d.f.型式如(4.10)式，便稱有非奇異之 n 維常態分佈。亦即若 $\mathbf{Y} = (Y_1, \dots, Y_n)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈，且 $|\boldsymbol{\Sigma}| \neq 0$ ，則 $\mathbf{Y}$ 之p.d.f.可寫出如(4.10)式，其中 $\boldsymbol{\mu}$ 及 $\boldsymbol{\Sigma}$ 分別爲 $\mathbf{Y}$ 之期望向量及共變異數矩陣。即若 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$ ， $\boldsymbol{\Sigma} = (\sigma_{ij})$ 爲一 $n \times n$ 矩陣，則

$$E(Y_i) = \mu_i, i = 1, \dots, n,$$

$$\text{Var}(Y_i) = \sigma_{ii}, i = 1, \dots, n,$$

$$\text{Cov}(Y_i, Y_j) = \sigma_{ij}, i, j = 1, \dots, n.$$

我們再寫一較一般的定理如下，證明則留在習題。

定理4.4 設 $\mathbf{X} = (X_1, \dots, X_n)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈， $\mathbf{C}$ 爲一 $n \times m$ 矩陣， $\mathbf{V}$ 爲一 $1 \times m$ 之常數向量，令 $\mathbf{Y} = \mathbf{XC} + \mathbf{V}$ ，則 $\mathbf{Y}$ 亦有 m 維常態分佈，且 $E(\mathbf{Y}) = \boldsymbol{\mu}\mathbf{C} + \mathbf{V}$ ， $\boldsymbol{\Sigma}_Y = \mathbf{C}'\boldsymbol{\Sigma}\mathbf{C}$ 。特別地，若 $m = 1$ ，即 $\mathbf{C} = \mathbf{c}'$ 爲一 $n \times 1$ 之行向量，且設 $\mathbf{V} = 0$ ，則 $\mathbf{Y} = \mathbf{Xc}'$ （爲一 $X_1, \dots, X_n$ 之線性組合）有 $\mathcal{N}(\boldsymbol{\mu}\mathbf{c}', \mathbf{c}\boldsymbol{\Sigma}\mathbf{c}')$ 分佈。

例4.1 設 $\mathbf{X} = (X_1, X_2)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈。由定理4.4，此時 X_1, X_2 皆有常態分佈。設 X_1, X_2 分別有 $\mathcal{N}(\mu_1, \sigma_1^2)$ 及 $\mathcal{N}(\mu_2, \sigma_2^2)$ 分佈，且令 ρ 表 X_1 與 X_2 之相關係數，設 $|\rho| \neq 1$ 。則共變異數矩陣爲

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

且 $|\Sigma| = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$ 。反之，若先給共變異數矩陣，亦可得 X_1, X_2 之邊際分佈，及相關係數。因此可求出（見習題第1題） X_1, X_2 之聯合 p.d.f. 為

$$(4.11) f(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho}{\sigma_1\sigma_2}(x_1-\mu_1)(x_2-\mu_2) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right\},$$

與我們在第三章定義6.1所給相同。

例4.2 設 $\mathbf{X} = (X_1, X_2, X_3)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \Sigma)$ 分佈，其中

$$\boldsymbol{\mu} = (1, 2, 0), \quad \Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

試求 $\mathbf{X}$ 之 p.d.f.。

解. 首先

$$\Sigma^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/9 \end{pmatrix}, \quad |\Sigma| = 1 \cdot 4 \cdot 9 = 36.$$

又

$$(\mathbf{x} - \boldsymbol{\mu})\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})' = (x_1 - 1)^2 + \frac{1}{4}(x_2 - 2)^2 + \frac{1}{9}x_3^2.$$

因此 $\mathbf{X}$ 之 p.d.f. 為

$$\begin{aligned} f_{\mathbf{X}}(x_1, x_2, x_3) &= \frac{1}{(2\pi)^{3/2} \sqrt{36}} e^{-\frac{1}{2}((x_1-1)^2 + \frac{1}{4}(x_2-2)^2 + \frac{1}{9}x_3^2)} \\ &= \frac{1}{\sqrt{2\pi} \cdot 1} e^{-\frac{1}{2}(x_1-1)^2} \cdot \frac{1}{\sqrt{2\pi} \cdot 2} e^{-\frac{1}{2}(x_2-2)^2/4} \cdot \frac{1}{\sqrt{2\pi} \cdot 3} e^{-\frac{1}{2}x_3^2/9}. \end{aligned}$$

即 X_1, X_2, X_3 為相互獨立，分別有 $\mathcal{N}(1, 1)$, $\mathcal{N}(2, 4)$ 及 $\mathcal{N}(0, 9)$ 分佈。事實上，由 Σ 的形式，便已可知 X_1, X_2, X_3 獨立，見稍後的定理4.5。

例4.3 設 X_1, X_2, X_3 為 i.i.d. 之 $\mathcal{N}(\mu, \sigma^2)$ 隨機變數。令 $\mathbf{X} = (X_1, X_2, X_3)$,

$$\mathbf{Y} = (Y_1, Y_2, Y_3) = \mathbf{X}\mathbf{A}.$$

其中

$$\mathbf{A} = \begin{pmatrix} 1/3 & -1 & 0 \\ 1/3 & 0 & 1 \\ 1/3 & 1 & -1 \end{pmatrix}。$$

試決定 $\mathbf{Y}$ 之分佈。

解. 由定理4.4, $\mathbf{Y} \sim (\boldsymbol{\mu}\mathbf{A}, \mathbf{A}'\boldsymbol{\Sigma}\mathbf{A})$, 其中

$$\boldsymbol{\mu} = (\mu, \mu, \mu), \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix},$$

而

$$\begin{aligned} \boldsymbol{\mu}\mathbf{A} &= (\mu, 0, 0), \\ \mathbf{A}'\boldsymbol{\Sigma}\mathbf{A} &= \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} 1/3 & -1 & 0 \\ 1/3 & 0 & 1 \\ 1/3 & 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \sigma^2/3 & 0 & 0 \\ 0 & 2\sigma^2 & -\sigma^2 \\ 0 & -\sigma^2 & 2\sigma^2 \end{pmatrix}。 \end{aligned}$$

附帶一提, 由上述結果知, $Y_1 = (X_1 + X_2 + X_3)/3 = \bar{X}$ 之邊際分佈為 $\mathcal{N}(\mu, \sigma^2/3)$, 此證實定理3.1的結論。又 $Y_2 = -X_1 + X_3$, $Y_3 = X_2 + X_3$ 皆有 $\mathcal{N}(\mu, 2\sigma^2)$ 分佈, 此也與3.5節二變數常態分佈的性質(iii)吻合。又因 $\text{Cov}(Y_1, Y_2) = \text{Cov}(Y_1, Y_3) = 0$, 故由二變數常態分佈的性質(iv)知, Y_1 與 Y_2 , Y_1 與 Y_3 皆獨立。

底下二定理的證明皆留在習題中。

定理4.5 設 $\mathbf{X} = (X_1, \dots, X_n)$ 有 n 維常態分佈, 則 X_i 與 X_j 獨立, 其中 $1 \leq i \neq j \leq n$, 若且唯若 $\text{Cov}(X_i, X_j) = 0$ 。即若且唯若其共變異數矩陣 $\boldsymbol{\Sigma}$ 之 (i, j) (以及 (j, i))位置的元素為0。又 $X_1, \dots, X_n$ 相互獨立, 若且唯若 $\text{Cov}(X_i, X_j) = 0, \forall 1 \leq i, j \leq n$, 且 $i \neq j$ (見習題第3題)。即若且唯若 $\boldsymbol{\Sigma}$ 為對角矩陣。

定理4.6 設有 k 個 n 維之獨立的隨機向量 $\mathbf{X}_1, \dots, \mathbf{X}_k$, 分別有 $\mathcal{MN}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ 分佈, $j = 1, \dots, k$ 。令 $\mathbf{X} = \sum_{j=1}^k c_j \mathbf{X}_j$, 其中 $c_1, \dots, c_k$ 為常數, 且 $\sum_{j=1}^k c_j^2 \neq 0$, 則 $\mathbf{X}$ 有 $\mathcal{MN}(\sum_{j=1}^k c_j \boldsymbol{\mu}_j, \sum_{j=1}^k c_j^2 \boldsymbol{\Sigma}_j)$ 分佈。

設 $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{C}$ 為 $-n \times n$ 之對稱矩陣, 則 $\mathbf{x} \mathbf{C} \mathbf{x}'$ 稱為 $-\mathbf{x}$ 之二次形式(quadratic form)。關於二次形式, 我們有下述結果。

定理4.7 設 $\mathbf{X} = (X_1, \dots, X_n)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈, $|\boldsymbol{\Sigma}| \neq 0$, 且令 $Q = (\mathbf{X} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})'$, 則 Q 有 χ_n^2 分佈。

證明. 對 $\forall t \in R$,

$$\begin{aligned}
 (4.12) \quad M_Q(t) &= E(e^{tQ}) \\
 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\{t(\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})'\} (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \\
 &\quad \cdot \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})'\} dx_1 \cdots dx_n \\
 &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \\
 &\quad \cdot \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})' (1 - 2t)\} dx_1 \cdots dx_n \\
 &= (1 - 2t)^{-n/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (2\pi)^{-n/2} \left| \frac{\boldsymbol{\Sigma}}{1 - 2t} \right|^{-1/2} \\
 &\quad \cdot \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}) \left(\frac{\boldsymbol{\Sigma}}{1 - 2t} \right)^{-1} (\mathbf{x} - \boldsymbol{\mu})'\} dx_1 \cdots dx_n,
 \end{aligned}$$

此處用到

$$\left| \frac{\boldsymbol{\Sigma}}{1 - 2t} \right| = (1 - 2t)^{-n} |\boldsymbol{\Sigma}|.$$

(4.12)式最後一積分算子, 可視為 $-n$ 維 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/(1 - 2t))$ 分佈之p.d.f., 故其積分值為1。因此

$$M_Q(t) = (1 - 2t)^{-n/2},$$

但此為 χ_n^2 分佈之動差母函數。得證。

若 X_i 有 $\mathcal{N}(\mu_i, \sigma_i^2)$ 分佈, $i = 1, \dots, n$, 且 $X_1, \dots, X_n$ 相互獨立, 則易知此時

$$Q = \sum_{i=1}^n (X_i - \mu_i)^2 / \sigma_i^2$$

有 χ_n^2 分佈, 定理4.7 即此結果之推廣。我們再給此結果之另一推廣。

定理4.8 設 $X_1, \dots, X_n$ 為 i.i.d. 之 $\mathcal{N}(0, \sigma^2)$ 隨機變數, 且令 $\mathbf{X} = (X_1, \dots, X_n)$ 。又設 $\mathbf{C}$ 為一 $n \times n$ 之對稱非零矩陣, 令 $Q = \mathbf{X} \mathbf{C} \mathbf{X}' / \sigma^2$, 則 Q 有 χ_r^2 分佈, 若且唯若 $\mathbf{C}$ 為一幂等矩陣, 且 $r(\mathbf{C}) = r$ 。

註4.1 若 $\mathbf{C} = \mathbf{I}_n$, 則 $\mathbf{C}$ 為一幂等矩陣且 $r(\mathbf{C}) = n$ 。此時 $Q = \sum_{i=1}^n X_i^2 / \sigma^2$ 有 χ_n^2 分佈為一早已知道的結果。定理4.8 即為此結果之推廣。

證明. 不失一般性可設 $\sigma^2 = 1$ 。我們先求 Q 之動差母函數。對 $\forall t \in R$,

$$\begin{aligned} (4.13) \quad M(t) &= E(e^{t \mathbf{X} \mathbf{C} \mathbf{X}'}) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (2\pi)^{-n/2} \exp\{t \mathbf{x} \mathbf{C} \mathbf{x}' - \mathbf{x} \mathbf{x}' / 2\} dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (2\pi)^{-n/2} \exp\{-\frac{1}{2} \mathbf{x} (\mathbf{I}_n - 2t \mathbf{C}) \mathbf{x}'\} dx_1 \cdots dx_n. \end{aligned}$$

只要 $|t|$ 夠小, 譬如說 $|t| < h, h > 0$, 則 $\mathbf{I}_n - 2t \mathbf{C}$ 為一正定矩陣。又此時

$$(2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\{-\frac{1}{2} \mathbf{x} \Sigma^{-1} \mathbf{x}'\},$$

為一 $\mathcal{MN}(\mathbf{0}, \Sigma)$ 之 p.d.f., 其中 $\Sigma^{-1} = (\mathbf{I}_n - 2t \mathbf{C})$ 。因此

$$(4.14) \quad M(t) = |\Sigma|^{1/2} = |\mathbf{I}_n - 2t \mathbf{C}|^{-1/2}, |t| < h.$$

我們可將 $M(t)$ 以另一形式來表示。設 $a_1, \dots, a_n$ 為 $\mathbf{C}$ 之特徵根, 則由定理4.2 知, 存在一正交矩陣 $\mathbf{B}$, 使得

$$\mathbf{B} \mathbf{C} \mathbf{B}' = \begin{pmatrix} a_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & a_n \end{pmatrix}.$$

故

$$\mathbf{B}(\mathbf{I}_n - 2t\mathbf{C})\mathbf{B}' = \begin{pmatrix} 1 - 2ta_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 - 2ta_n \end{pmatrix} \circ$$

因此

$$|\mathbf{I}_n - 2t\mathbf{C}| = |\mathbf{B}(\mathbf{I}_n - 2t\mathbf{C})\mathbf{B}'| = \prod_{i=1}^n (1 - 2ta_i) \circ$$

故得

$$(4.15) \quad M(t) = \left(\prod_{i=1}^n (1 - 2ta_i) \right)^{-1/2}, |t| < h \circ$$

現設 $r(\mathbf{C}) = k$, $1 \leq k \leq n$, 且設 $a_1, \dots, a_k \neq 0$, $a_{k+1} = \dots = a_n = 0$, 則

$$(4.16) \quad M(t) = \left(\prod_{i=1}^k (1 - 2ta_i) \right)^{-1/2}, |t| < h \circ$$

若 $\mathbf{X}\mathbf{C}\mathbf{X}'$ 有 χ_r^2 分佈, 則

$$(4.17) \quad M(t) = (1 - 2t)^{-r/2} \circ$$

比較(4.16)及(4.17)式知

$$(4.18) \quad \prod_{i=1}^k (1 - 2ta_i) = (1 - 2t)^r, |t| < h \circ$$

由於(4.18)式左、右二多項式對無限多個 t 相等, 因此 $k = r$ 且 $a_1 = a_2 = \dots = a_r = 1$ 。故矩陣 $\mathbf{C}$ 為冪等矩陣, 且 $r(\mathbf{C}) = r$ 。

反之, 若 $\mathbf{C}$ 為冪等矩陣, 且 $r(\mathbf{C}) = r$, 則 $a_1 = \dots = a_r = 1$, 且 $a_{r+1} = \dots = a_n = 0$, 因此(4.17)式成立, 故 $\mathbf{X}\mathbf{C}\mathbf{X}'$ 有 χ_r^2 分佈。本定理證畢。

註4.2 在定理4.8中, 若 X_i 有 $\mathcal{N}(\mu, \sigma^2)$ 分佈, 則 Q 有非中心卡方分佈(non-central chi squared distribution), 且自由度為 r , 若且唯若 $\mathbf{C}^2 = \mathbf{C}$ 且 $r(\mathbf{C}) = r$ 。且此時非中心參數(noncentrality parameter)為 $\boldsymbol{\mu}\mathbf{C}\boldsymbol{\mu}'/\sigma^2$, 其中 $\boldsymbol{\mu} = (\mu, \dots, \mu)$ 。由於 $\boldsymbol{\mu}\mathbf{C}\boldsymbol{\mu}' = \mu^2 \sum_{i,j} c_{ij}$, 其中 $\mathbf{C} = (c_{ij})$, 故若 $\mu \neq 0$, 則 $\mathbf{C}^2 =$

$C, r(C) = r$, 且 $\sum_{i,j} c_{ij} = 0$, 為 Q 有 χ_r^2 之充要條件。

其他尚有許多關於二次形式之有趣的結果, 可參考 Hogg and Craig (1970) Chapter 12, 及 Roussas (1997) Chapter 19。至於有關多變量常態分佈較完整的討論可參考 Anderson (1984) Chapter 2。

最後, 對 $X_1, \dots, X_n$ 為 i.i.d. 之 $\mathcal{N}(\mu, \sigma^2)$ 隨機變數, 我們給樣本平均 $\bar{X}_n$ 與樣本變異數 S_n^2 獨立之另一證明。

由定理 4.3, 可造出一正交矩陣 $C = (c_{ij})$, 滿足 $CC' = I_n$, 且 C 之第一行為 $(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})'$ 。又令 $Y = XC$, 其中 $Y = (Y_1, \dots, Y_n)$, $X = (X_1, \dots, X_n)$ 。由定理 4.4 知 Y 有 $\mathcal{MN}(\mu C, C' \Sigma C)$ 分佈, 其中

$$\mu = (\mu, \dots, \mu),$$

$$\Sigma = \begin{pmatrix} \sigma^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma^2 \end{pmatrix} = \sigma^2 I_n。$$

顯然 $C' \Sigma C = \Sigma = \sigma^2 I_n$, 故 $Y_1, \dots, Y_n$ 仍為相互獨立。因 C 為正交矩陣, 且第一行為 $(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}})'$, 故 $\sum_{i=1}^n a_{ij} \frac{1}{\sqrt{n}} = 0, j = 2, \dots, n$ 。因此

$$E(Y_1) = \sqrt{n}\mu, \quad E(Y_j) = \mu \sum_{i=1}^n a_{ij} = 0, \quad j = 2, \dots, n。$$

即 $Y_1, \dots, Y_n$ 不但相互獨立, 且 Y_1 有 $\mathcal{N}(\sqrt{n}\mu, \sigma^2)$ 分佈, Y_j 有 $\mathcal{N}(0, \sigma^2)$ 分佈, $j = 2, \dots, n$ 。由此又得 $\sum_{j=2}^n Y_j^2 / \sigma^2$ 有 χ_{n-1}^2 分佈。又 $\bar{X}_n = Y_1 / \sqrt{n}$, 因此 $\bar{X}_n$ 有 $\mathcal{N}(\mu, \sigma^2/n)$ 分佈, 若能再證出

$$\sum_{j=2}^n Y_j^2 = \sum_{j=1}^n (X_j - \bar{X}_n)^2$$

便得證了。但

$$\begin{aligned} \sum_{j=1}^n (X_j - \bar{X}_n)^2 &= \sum_{j=1}^n X_j^2 - n\bar{X}_n^2 = \sum_{j=1}^n X_j^2 - Y_1^2 \\ &= \mathbf{X}\mathbf{X}' - Y_1^2 = \mathbf{X}CC'\mathbf{X}' - Y_1^2 = \mathbf{Y}\mathbf{Y}' - Y_1^2 \end{aligned}$$

$$= \sum_{j=2}^n Y_j^2。$$

證畢。

4.5 順序統計量

在一組隨機樣本中，有時候最大值、最小值或中間值，是我們較有興趣的。例如，過去五十年最大降雨量，幾次跑步最短的時間，台北市房價的中間值等，這些都涉及**順序統計量**(order statistics)。

定義5.1 設 $X_1, \dots, X_n$ 為一組隨機樣本，按照小至大排出，而得 $X_{(1)}, \dots, X_{(n)}$ ，便稱**順序統計量**。

順序統計量滿足 $X_{(1)} \leq \dots \leq X_{(n)}$ 。特別地，

$$X_{(1)} = \min_{1 \leq i \leq n} X_i, \quad X_{(n)} = \max_{1 \leq i \leq n} X_i。$$

另外，

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

稱為**樣本全距**(sample range)，或稱**全距統計量**(range statistic)，表最大值與最小值之差距。譬如說班上身高最高與最矮的差距。如果樣本數 n 是很清楚的，則為了簡便，也可僅以 $\min X, \max X$ 分別表 $X_{(1)}$ 及 $X_{(n)}$ 。又

$$\text{med}X = \begin{cases} X_{((n+1)/2)} & , \text{若 } n \text{ 是奇數,} \\ (X_{(n/2)} + X_{(n/2+1)})/2 & , \text{若 } n \text{ 是偶數,} \end{cases}$$

表**樣本中位數**(sample median)。當 n 為偶數時，有些書對任意介於 $X_{(n/2)}$ 與 $X_{(n/2+1)}$ 間的值，皆稱為樣本中位數。

在美國，職業球隊裡，例如職棒、職籃，極少數的球員薪水(或收入)很高，但大部分的球員薪水是很低的。報紙上，有時會刊出某位球員與球隊簽下新的合約，五年薪水一億美元。平均一年兩千萬美元，尚不包括廣告

等收入! 當球員抱怨薪水太低時, 老闆可能會想全隊平均年薪已有百萬美元了, 但球員會想有一半以上的球員年薪少於二十萬美元(職業球員由於球齡短以及會因受傷而結束球員生涯, 所以薪水是較其他行業高)。這兩個觀點都是正確的, 只是一計算平均, 一計算中位數。當討論到收入、價格等含有一些較極端的值(過高或過低), 中位數可能是一較合理的指標值。

若 $X_1, \dots, X_n$ 為一由連續型的母體所產生之隨機樣本, 則任二隨機變數 $X_i, X_j, i \neq j$, 會相等的機率為0, 因此 $P(X_{(1)} < \dots < X_{(n)}) = 1$ 。下述定理給出任一順序統計量之分佈。

定理5.1 設 $X_1, \dots, X_n$ 為由連續分佈函數 $F(x)$, 且p.d.f.為 $f(x)$, 所產生之隨機樣本, $X_{(1)}, \dots, X_{(n)}$ 表其順序統計量。則 $X_{(j)}$ 之p.d.f.為

$$(5.1) \quad f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} (F(x))^{j-1} (1-F(x))^{n-j} f(x)。$$

證明. 我們先求 $X_{(j)}$ 之分佈函數, 然後微分而得p.d.f.。事件 $X_{(j)} \leq x$, 等價於 $X_1, \dots, X_n$ 中至少有 j 個小於或等於 x 。因此

$$(5.2) \quad \begin{aligned} F_{X_{(j)}}(x) &= P(X_{(j)} \leq x) \\ &= \sum_{k=j}^n \binom{n}{k} (F(x))^k (1-F(x))^{n-k}。 \end{aligned}$$

將上式對 x 微分, 得

$$\begin{aligned} f_{X_{(j)}}(x) &= \sum_{k=j}^n \binom{n}{k} (k(F(x))^{k-1} (1-F(x))^{n-k} \\ &\quad - (n-k)(F(x))^k (1-F(x))^{n-k-1}) f(x) \\ &= \binom{n}{j} j (F(x))^{j-1} (1-F(x))^{n-j} f(x) \\ &\quad + \sum_{k=j+1}^n \binom{n}{k} k (F(x))^{k-1} (1-F(x))^{n-k} f(x) \\ &\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k) (F(x))^k (1-F(x))^{n-k-1} f(x)。 \end{aligned}$$

$$= \frac{n!}{(j-1)!(n-j)!} f(x)(F(x))^{j-1}(1-F(x))^{n-j}。$$

此處用到

$$\begin{aligned} & \sum_{k=j+1}^n \binom{n}{k} k(F(x))^{k-1}(1-F(x))^{n-k} f(x) \\ &= \sum_{k=j}^{n-1} \binom{n}{k+1} (k+1)(F(x))^k(1-F(x))^{n-k-1} f(x) \\ &= \sum_{k=j}^{n-1} \frac{n!(k+1)}{(k+1)!(n-k-1)!} (F(x))^k(1-F(x))^{n-k-1} f(x) \\ &= \sum_{k=j}^{n-1} \frac{n!(n-k)}{k!(n-k)!} (F(x))^k(1-F(x))^{n-k-1} f(x) \\ &= \sum_{k=j}^{n-1} \binom{n}{k} (n-k)(F(x))^k(1-F(x))^{n-k-1} f(x)。$$

證畢。

例5.1 設 $X_1, \dots, X_n$ 為由 $\mathcal{U}(0, 1)$ 分佈所產生隨機樣本。則 $F_X(x) = x, x \in (0, 1)$ 。利用定理5.1, 得

$$\begin{aligned} f_{X_{(j)}}(x) &= \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-j} \\ &= \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} x^{j-1}(1-x)^{n-j}, x \in (0, 1)。 \end{aligned}$$

因此 $X_{(j)}$ 有 $Be(j, n-j+1)$ 分佈。

例5.2 設 $X_1, \dots, X_n$ 為由分佈函數 F 所產生之隨機樣本。則

$$\begin{aligned} P(X_{(n)} \leq x) &= P(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\ &= \prod_{i=1}^n P(X_i \leq x) = (F(x))^n, x \in R。 \end{aligned}$$

當然在(5.2)式中令 $j = n$, 也會得到上式。若 F 之p.d.f.為 f , 則經由對 x 微分, 得

$$(5.3) \quad f_{X_{(n)}}(x) = n(F(x))^{n-1}f(x), x \in R。$$

與(5.1)式中令 $j = n$ 所得一致。

另外,

$$\begin{aligned}
 P(X_{(1)} \leq x) &= 1 - P(X_{(1)} > x) \\
 &= 1 - P(X_1 > x, X_2 > x, \dots, X_n > x) \\
 &= 1 - \prod_{i=1}^n P(X_i > x) \\
 &= 1 - (1 - F(x))^n, x \in R.
 \end{aligned}$$

經由對 x 微分, 得

$$(5.4) \quad f_{X_{(1)}}(x) = n(1 - F(x))^{n-1}f(x), x \in R,$$

仍與在(5.1)式中令 $j = 1$ 所得一致。

例如, 若 $X_1, \dots, X_n$ 之共同分佈為 $\mathcal{E}(\lambda)$, 則由(5.4)式, 即得 $X_{(1)} \sim \mathcal{E}(n\lambda)$, 仍有指數分佈, 只是參數改為 $n\lambda$ 。此為一有趣的結果。

定理5.2 設 $X_1, \dots, X_n$ 為由連續分佈函數 $F(x)$, 且p.d.f.為 $f(x)$, 所產生之隨機樣本, $X_{(1)}, \dots, X_{(n)}$ 表其順序統計量。則 $X_{(i)}, X_{(j)}, 1 \leq i < j \leq n$, 之聯合p.d.f.為

$$\begin{aligned}
 (5.5) \quad f_{X_{(i)}, X_{(j)}}(y_i, y_j) &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} (F(y_i))^{i-1} (F(y_j) - F(y_i))^{j-i-1} \\
 &\quad \cdot (1 - F(y_j))^{n-j} f(y_i) f(y_j), -\infty < y_i < y_j < \infty.
 \end{aligned}$$

上定理的證明我們略去了。三個或更多個順序統計量的聯合p.d.f.亦可求出。如設整數 $1 \leq k_1 < \dots < k_r \leq n$, 則 $X_{(k_1)}, \dots, X_{(k_r)}$ 之聯合p.d.f.為

$$\begin{aligned}
 (5.6) \quad f_{X_{(k_1)}, \dots, X_{(k_r)}}(y_{k_1}, \dots, y_{k_r}) &= \frac{n!}{(k_1-1)!(k_2-k_1-1)! \dots (n-k_r)!} (F(y_{k_1}))^{k_1-1} \\
 &\quad \cdot (F(y_{k_2}) - F(y_{k_1}))^{k_2-k_1-1} \dots (F(y_{k_r}) - F(y_{k_{r-1}}))^{k_r-k_{r-1}-1} \\
 &\quad \cdot (1 - F(y_{k_r}))^{n-k_r} f(y_{k_1}) f(y_{k_2}) \dots f(y_{k_r}), -\infty < y_{k_1} < \dots < y_{k_r} < \infty.
 \end{aligned}$$

$$\cdot (F(y_{k_2}) - F(y_{k_1}))^{k_2 - k_1 - 1} \cdots (1 - F(y_{k_r}))^{n - k_r} f(y_{k_1}) \cdots f(y_{k_r}),$$

$$-\infty < y_{k_1} < \cdots < y_{k_r} < \infty。$$

又 $X_{(1)}, \cdots, X_{(n)}$ 之聯合 p.d.f. 為

$$(5.7) \quad f_{X_{(1)}, \cdots, X_{(n)}}(y_1, \cdots, y_n) = n! f(y_1) \cdots f(y_n),$$

$$-\infty < y_1 < \cdots < y_n < \infty。$$

在(5.7)式中, $n!$ 的出現是很顯然的: 對任一組 $y_1 < \cdots < y_n$, 有 $n!$ 組 $x_1, \cdots, x_n$, 其順序統計量均對應 $y_1, \cdots, y_n$ 。

底下給幾個例子。

例5.3 設 X_1, X_2, X_3 為由 p.d.f. $f(x) = 2x, 0 < x < 1$, 所產生之隨機樣本。則 $X_{(1)}, X_{(2)}, X_{(3)}$ 之聯合 p.d.f. 為

$$f_{X_{(1)}, X_{(2)}, X_{(3)}}(y_1, y_2, y_3) = 3! \cdot (2y_1)(2y_2)(2y_3)$$

$$= 48y_1y_2y_3, 0 < y_1 < y_2 < y_3。$$

由此可得 $X_{(1)}$ 之邊際 p.d.f.

$$(5.8) \quad f_{X_{(1)}}(y_1) = \int_{y_1}^1 \int_{y_2}^1 48y_1y_2y_3 dy_3 dy_2$$

$$= 6y_1(1 - y_1^2)^2, 0 < y_1 < 1。$$

又 $X_{(1)}, X_{(3)}$ 之聯合 p.d.f. 為

$$(5.9) \quad f_{X_{(1)}, X_{(3)}}(y_1, y_3) = \int_{y_1}^{y_3} 48y_1y_2y_3 dy_2$$

$$= 24y_1y_3(y_3^2 - y_1^2), 0 < y_1 < y_3 < 1。$$

可驗證(5.8)及(5.9)式, 與利用(5.1)式及(5.5)式所得相同。

例5.4 設 $X_1, \cdots, X_n$ 為由 $\mathcal{U}(0, 1)$ 分佈所產生之隨機樣本。令全距 $R = X_{(n)} - X_{(1)}$, 半全距(midrange) $V = (X_{(1)} + X_{(n)})/2$ 。試求 R, V 之聯合

p.d.f., 邊際p.d.f., $E(R)$ 及 $E(V)$ 。

解. 首先

$$f_{X_{(1)}, X_{(n)}}(y_1, y_n) = n(n-1)(y_n - y_1)^{n-2}, 0 < y_1 < y_n < 1。$$

利用變數代換, 得

$$f_{R,V}(r, v) = n(n-1)r^{n-2}, 0 < r < 1, r/2 < v < 1 - r/2。$$

由此又得 R 之邊際p.d.f.爲

$$\begin{aligned} f_R(r) &= \int_{r/2}^{1-r/2} n(n-1)r^{n-2} dv \\ &= n(n-1)(1-r)r^{n-2}, 0 < r < 1。 \end{aligned}$$

可看出 R 有 $Be(n-1, 2)$ 分佈。又可求出

$$f_V(v) = \begin{cases} \int_0^{2v} n(n-1)r^{n-2} dr \\ \quad = n(2v)^{n-1}, 0 < v \leq 1/2, \\ \int_0^{2(1-v)} n(n-1)r^{n-2} dr \\ \quad = n(2(1-v))^{n-1}, 1/2 < v \leq 1。 \end{cases}$$

有了 R 及 V 之邊際p.d.f., 便可求出 $E(R)$ 及 $E(V)$ 。如

$$\begin{aligned} E(R) &= \int_0^1 n(n-1)(1-r)r^{n-2} \cdot r dr \\ &= n(n-1) \left(\int_0^1 r^{n-1} dr - \int_0^1 r^n dr \right) \\ &= n(n-1) \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= \frac{n-1}{n+1}。 \end{aligned}$$

但 $E(R)$ 亦可如下得到:

$$\begin{aligned} E(R) &= E(X_{(n)} - X_{(1)}) = E(X_{(n)}) - E(X_{(1)}) \\ &= \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1}, \end{aligned}$$

$$\begin{aligned} E(V) &= E\left(\frac{X_{(1)} + X_{(2)}}{2}\right) = \frac{1}{2}(E(X_{(1)}) + E(X_{(n)})) \\ &= \frac{1}{2}\left(\frac{1}{n+1} + \frac{n}{n+1}\right) = \frac{1}{2}。 \end{aligned}$$

此處用到 $X_{(1)} \sim \mathcal{B}e(1, n)$, $X_{(n)} \sim \mathcal{B}e(n, 1)$ (見例5.1)。

例5.5 某廠牌燈管宣稱可使用10,000小時, 某辦公室最近安裝40支該廠牌的燈管, 才使用1個月便壞了一支。這是否合理呢?

解. 爲了便於計算, 我們假設燈管的壽命有指數分佈。即設燈管壽命 $X_1, \dots, X_{40}$ 爲 i.i.d. 之 $\mathcal{E}(\lambda)$ 分佈, 期望值 $\lambda^{-1} = 10,000$ (小時)。又設每月上班25天, 每天開燈10小時。

對一特定的燈管, 使用1個月(250小時)內會壞的機率爲

$$P(X \leq 250) = 1 - e^{-250/10,000} = 1 - e^{-0.025} \doteq 1 - 0.9753 = 0.0247,$$

的確不大。但40支燈管1個月內至少壞一支的機率, 相當於最小順序統計量 $X_{(1)} \leq 250$ 之機率。此機率可利用例2.5, $X_{(1)}$ 有 $\mathcal{E}(40\lambda) = \mathcal{E}(1/250)$ 分佈, 或直接求:

$$\begin{aligned} P(X_{(1)} \leq 250) &= 1 - P(X_{(1)} > 250) \\ &= 1 - P(X_1 > 250, \dots, X_{40} > 250) \\ &= 1 - (P(X > 250))^{40} \\ &= (e^{-250/10,000})^{40} \\ &= 1 - e^{-1} \doteq 1 - 0.3679 = 0.6321。 \end{aligned}$$

便不算太小了。

至於5個月(1,250小時)內至少壞一支燈管的機率爲

$$\begin{aligned} P(X_{(1)} \leq 1,250) &= 1 - (P(X > 1,250))^{40} \\ &= 1 - (e^{-1,250/10,000})^{40} \\ &= 1 - e^{-5} \doteq 0.99326, \end{aligned}$$

非常接近1。但一特定的燈管, 5個月內壞的機率依然不大:

$$P(X \leq 1,250) = 1 - e^{1,250/10,000} = 1 - e^{-0.125} \doteq 1 - 0.8825 = 0.1175。$$

最後, 雖平均壽命為10,000小時, 但一燈管10年(30,000小時)後仍可能使用的可能性有多大呢? 即要求最大順序統計量 $X_{(40)}$ 要大於30,000之機率:

$$\begin{aligned} P(X_{(40)} > 30,000) &= 1 - (P(X \leq 30,000))^{40} \\ &= 1 - (1 - e^{30,000/10,000})^{40} \\ &= 1 - (1 - e^{-3})^{40} \doteq 1 - 0.1297 = 0.8703。 \end{aligned}$$

此機率不小。雖一特定的燈管使用10年以上的機率為 $e^{-3} \doteq 0.04979$ 很小。上述這些似令人驚訝的結果, 都是因燈管數較多所造成的。

底下我們來看一些關於順序統計量之極限結果。

例5.6 設 $X_1, \dots, X_n$ 為由 $\mathcal{U}[0, \theta]$ 分佈所產生之隨機樣本, $\theta > 0$ 。令 $Y_n = X_{(n)}$, $G_n(y) = P(Y_n \leq y)$ 。則對 $\forall 0 < y < \theta$,

$$G_n(y) = (P(X \leq y))^n = y^n,$$

且 $G_n(y) = 0$, 若 $y \leq 0$, $G_n(y) = 1$, 若 $y \geq \theta$ 。故得

$$\lim_{n \rightarrow \infty} G_n(y) = G(y) = \begin{cases} 0, & 0 \leq y < \theta, \\ 1, & y \geq \theta。 \end{cases}$$

$G(y)$ 對應常數隨機變數 $Y \equiv \theta$ 之分佈函數。即證出

$$Y_n \xrightarrow[n \rightarrow \infty]{d} \theta。$$

又由(2.22)式, 得

$$(5.10) \quad Y_n \xrightarrow[n \rightarrow \infty]{p} \theta。$$

也可依定義2.1, 直接證明(5.10)式。

例5.7 設 $X_1, \dots, X_n$ 為由 p.d.f. $f(x) = (1+x)^{-2}$, $x > 0$, 所產生之隨機樣本, $F(x) = 1 - (1+x)^{-1}$, $x > 0$ 。令 $Y_n = nX_{(1)}$, $G_n(y) = P(Y_n \leq y)$ 。則對 $\forall y > 0$,

$$\begin{aligned} G_n(y) &= P(nX_{(1)} \leq y) \\ &= 1 - P(X_{(1)} > y/n) \\ &= 1 - (1 + y/n)^{-n} \\ &\xrightarrow{n \rightarrow \infty} 1 - e^{-y}。 \end{aligned}$$

上述極限見第二章註2.3。又 $G_n(y) = 0$, $\forall y \leq 0$ 。故得

$$Y_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{E}(1)。$$

次令 $U_n = X_{(n)}/n$ 。則可證明 $U_n \xrightarrow{d} U$, 其中 $P(U \leq u) = e^{-1/u}$, $u > 0$ (習題第18題)。

4.6 敘述統計

在統計裡, 我們常需要**收集資料**。資料收集到後便要做**資料分析**(data analysis)。由於資料常是以數字的形式出現, 所以英文裡的data, 在中文裡, 除了稱為資料外, 也常稱為**數據**。所謂讓數字說話, 就是指對所收集到的資料, 整理並分析, 以提供有用的資訊。

對資料所做的初步整理, 是屬於**敘述統計**(descriptive statistics)的工作。本書主要介紹**推論統計**(inferential statistics)。不過敘述統計是一種較直觀的來了解數據的方式。諸如給出樣本平均及樣本變異數, 也都是屬於敘述統計的範疇。但以此二者, 不見得是最好的了解數據的方式, 往往也失之於簡略。適當地以圖形來展示資料, 可以讓我們對資料更了解。有時圖形可以告訴我們數據裡隱藏的故事, 引導我們進一步探討的步驟。所以統計圖形在資料分析的前置作業中, 扮演著重要的角色。

本節我們介紹幾個簡單的統計圖示法。圖示法很多, 並不限我們所介紹者。

4.6.1 點圖

如果資料數目並非很多，則可以點圖(dot plot, 或dot diagram)來了解資料。將數據以點表示在 x 軸上方，若有重複的值，則將點往上堆。例如，設有數據

28, 36, 37, 52, 36, 45, 39, 45, 38, 35, 36,

則其點圖如圖6.1。點圖可以讓我們對數據的散佈情況，有一粗略的概念。

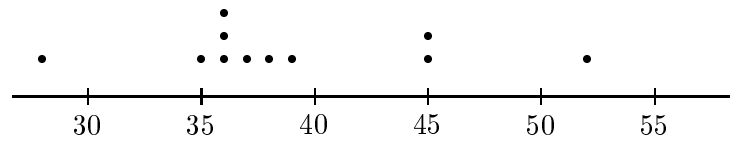


圖6.1 點圖

4.6.2 莖葉圖

仍以數據來說明。假設有下述數據：

9.5 10.8 8.8 11.2 10.2 10.3 10.2 11.3 10.0 8.8
10.7 9.9 11.4 9.8 10.5 9.8 9.9 10.9 8.1 10.5
10.6 8.2 8.6 9.2 9.9 10.0 11.0 9.2 10.7 10.9

莖葉圖(stem-and-leaf diagram)於選定適當的莖後，便可表示出數據。見圖6.2。

莖	葉
8	8 8 1 2 6
9	5 9 8 8 9 2 9 2
10	8 2 3 2 0 7 5 9 5 6 0 7 9
11	2 3 4 0

圖6.2 莖葉圖

點圖與莖葉圖其實未減縮任何資料，只是將資料以不同的方式呈現，以看得更清楚。

4.6.3 盒鬚圖

盒鬚圖(又稱**長鬚圖**，box-and-whisker plot，簡稱boxplot)，是一很有效的表示資料的方法，很多計算機程式中皆有此功能。盒鬚圖可用來了解資料的**偏斜性**(skewness)及**離群值**(outliers)。圖6.3為一典型的盒鬚圖。

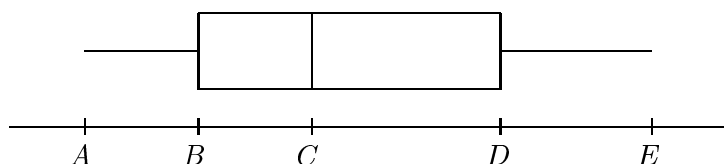


圖6.3 盒鬚圖

圖6.3中， A 稱為資料的最小值， E 稱為最大值， B 與 D 則分別為資料的下四分位數及上四分位數，因此圖中盒子(盒子之高度並不重要)包含資料的中間50%部分。又 C 為資料的中位數。圖6.3包含一個盒子，及二凸出來的鬚($\overline{AB}$ 及 $\overline{DE}$ 上方二線段)，這是此圖命名的由來。

樣本的中位數，上四分位數及下四分位數，以及一般的第 p 分位數，皆可與分佈的分位數對應。設有數據 $y_1, \dots, y_n$ ，將其按小至大排列，而得 $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$ 。關於**樣本第 p 分位數**(或說**第 p 樣本分位數**)，不同的書有不同的定義。有些書是定義為 $y_{([np]+1)}$ ，其中 $[\cdot]$ 為最大整數函數。此定義雖然簡單，但缺點為若 p 與 r 很接近，則第 p 分位數與第 r 分位數可能會相等。另外，亦可將樣本的第 p 分位數定義為

$$(6.1) \quad y_{([(n+1)p])} + ((n+1)p - [(n+1)p])(y_{([(n+1)p]+1)} - y_{([(n+1)p])}).$$

這其實是線性內插法的概念。我們以下例來說明。

例6.1 設有數據8, 20, 5, 15。求樣本之第0.37分位數。

解.先將樣本按小至大排列，而得 $y_{(1)} = 5$, $y_{(2)} = 8$, $y_{(3)} = 15$, $y_{(4)} =$

20。因 $n = 4$, $p = 0.37$, $[np] + 1 = [4 \cdot 0.37] + 1 = [1.48] + 1 = 2$ 。故以第一種定義所得之第0.37樣本分位數為 $y_{(2)} = 8$ 。又若 $p = 0.4$, 則 $[np] + 1 = 2$, 故第0.4樣本分位數亦為 $y_{(2)} = 8$ 。事實上從第0.25至第0.49樣本分位數皆為8。

現在來看第二種定義。四個樣本點, 將樣本的值域分成五個區間, 此四點便對應第0.2, 0.4, 0.6及0.8分位數。第0.37分位數介於第0.2及0.4分位數間。而 $y_{(1)} = 5$, $y_{(2)} = 8$, 因此第0.37樣本分位數為

$$5 + \frac{0.37 - 0.20}{0.40 - 0.20}(8 - 5) = 5 + 0.85 \cdot 3 = 7.55。$$

第二種定義的優點是不同的 p 值, 有不同的第 p 分位數。缺點是 p 不能很靠近0或1。此因若 $p < 1/(n+1)$, 則 $[(n+1)p] = 0$, 但我們並沒有 $y_{(0)}$ 的定義。又當 $p > n/(n+1)$, 則 $[(n+1)p] + 1 > n$, 因此 $y_{([(n+1)p]+1)}$ 也不存在。

當 $p = 1/2$, (6.1)式導致樣本中位數為

$$\begin{aligned} & y_{([(n+1)/2])} + \left(\frac{n+1}{2} - \left[\frac{n+1}{2}\right]\right)(y_{([(n+1)/2]+1}) - y_{([(n+1)/2]}) \\ &= \begin{cases} y_{((n+1)/2)} & , n \text{ 為奇數,} \\ (y_{(n/2)} + y_{(n/2+1)})/2 & , n \text{ 為偶數。} \end{cases} \end{aligned}$$

此與4.5節對樣本中位數所給的定義相符合。

底下我們對樣本第 p 分位數採第二種定義。如此樣本的第 $i/(n+1)$ 分位數, 即樣本的第 i 個順序統計量 $y_{(i)}$ (習題第4題)。

例6.2 瑞萊曾由不同的來源量測標準體積中氮(nitrogen)的重量(見Tukey (1977)), 而得到下述數據:

2.30143, 2.29816, 2.30182, 2.29890, 2.31017,
2.30986, 2.31010, 2.31001, 2.29889, 2.29940,
2.29849, 2.29889, 2.31024, 2.31030, 2.31028。

將數據表至小數第3位(四捨五入), 則圖6.4a, b分別為點圖及盒鬚圖。

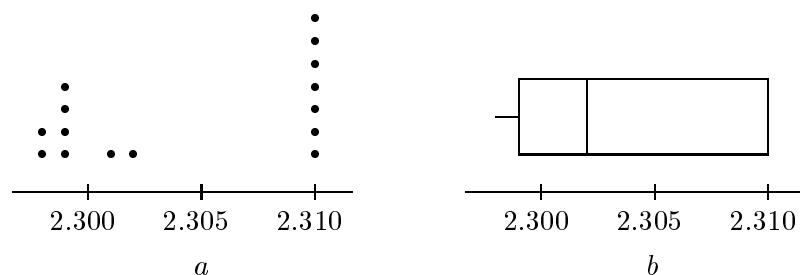


圖6.4 圖a及圖b分別為例6.2中之點圖及盒鬚圖

4.6.4 分位圖

利用樣本分位數來繪資料的圖，稱為**分位-分位圖**(quantile-quantile plot, 又稱Q-Q plot), 簡稱**Q-Q圖**。這是以圖形來檢驗數據是否來自某一分佈的方法。

以**常態分佈的Q-Q圖**(所謂normal quantile-quantile plot, 簡稱normal Q-Q plot, 有時也稱normal quantile plot)為例。在實際的應用裡, 我們常會想知道收集到的數據是否來自常態分佈的母體。在第十章我們會引進較客觀的檢驗方法, 此處我們看如何由圖形來做初步的檢驗。先選取一些適當的 p , 再以(6.1)式得到樣本之第 p 分位數 y_p , 然後在座標平面上繪出所有 (x_p, y_p) , 其中 x_p 為 $\mathcal{N}(0, 1)$ 分佈之 p 分位數。如果樣本是來自常態分佈的母體, 則其Q-Q圖近似於一條直線, 否則便不是。這是一種簡易地檢驗**常態性**(normality)的方法。

如果有 n 個樣本, 通常取 $p_i = i/(n+1)$, $i = 1, \dots, n$, 則 $y_{p_i} = y_{(i)}$, 且常態分佈的Q-Q圖, 即所有點 $(x_{p_i}, y_{(i)})$, $i = 1, \dots, n$, 其中 $x_{p_i} = \Phi^{-1}(p_i)$ 。

圖6.5a, b分別為20個來自 $\mathcal{E}(1)$ 及 $\mathcal{N}(0, 1)$ 分佈的樣本點之常態分佈的Q-Q圖, 二者似乎都不像是直線。但若各產生300個樣本點, 並繪於圖6.5c及d, 則圖6.5(d)就像一條直線了。這樣的以視覺來判定點是否在一直線上, 難免失之於主觀。但圖示法是一種初步的檢定。

例6.3 設有9個樣本點

2.33, 5.69, 0.39, -2.28, -2.14, 3.31, -0.72, 1.63, -3.91。

我們想看是否源自於常態分佈。

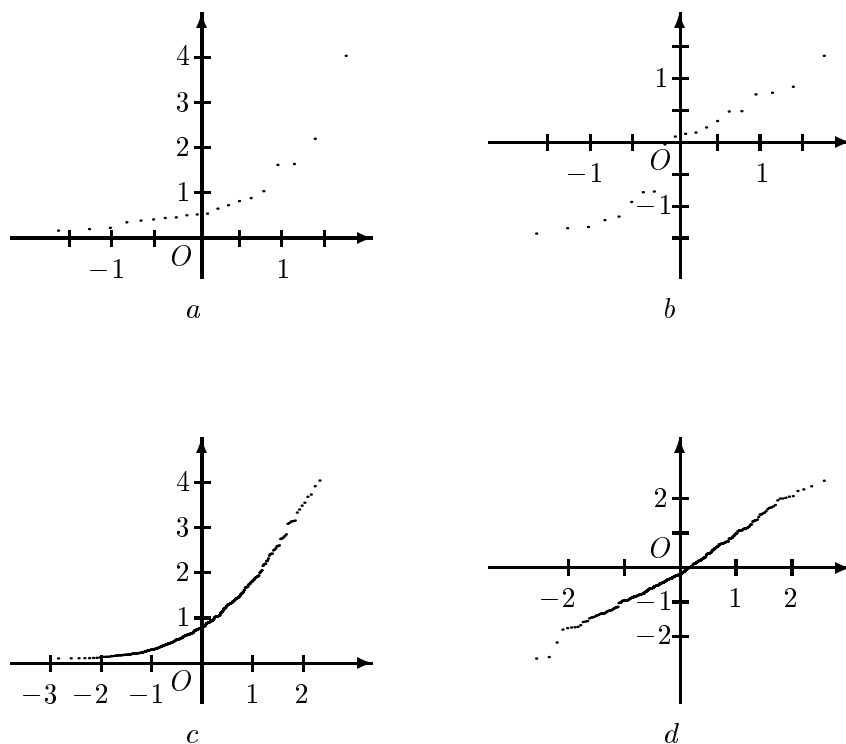


圖6.5 圖 a, b 分別為 $\mathcal{E}(1)$ 及 $\mathcal{N}(0, 1)$ 分佈20個樣本點之常態分佈的Q-Q圖,
圖 c, d 則為300個樣本點之常態分佈的Q-Q圖

由於樣本數 $n = 9$, 我們給出 $(x_i, y_i), i = 1, \dots, 9$, 之圖, 其中 x_i 為 $\mathcal{N}(0, 1)$ 分佈之第 $i/10$ 分位數, 即

$$P(Z \leq x_i) = \frac{i}{10}, i = 1, \dots, 9,$$

y_i 為樣本第 i 個順序統計量。如此得下述9點:

$$(-1.28, -3.91), (-0.84, -2.28), (-0.52, -2.14), (-0.25, -0.72), \\ (0, 0.39), (0.25, 1.63), (0.52, 2.33), (0.84, 3.31), (1.28, 5.69)。$$

圖6.6給出常態分佈之Q-Q圖, 大致是線性, 只有圖的右上方較偏離。這告訴我們說不定該進一步檢查資料是否登記有誤, 或是否資料無

誤，但有離群值存在。事實上上述數據為由 $\mathcal{N}(0, 9)$ 分佈所模擬產生，但將數據3.69改為5.69。若將5.69改回3.69，則 (x_i, y_i) , $i = 1, \dots, 9$ ，便大致落在直線 $y = 3x$ 上（對 $\mathcal{N}(\mu, \sigma^2)$ 分佈，其理想的Q-Q圖為直線 $y = \sigma x + \mu$ ）。

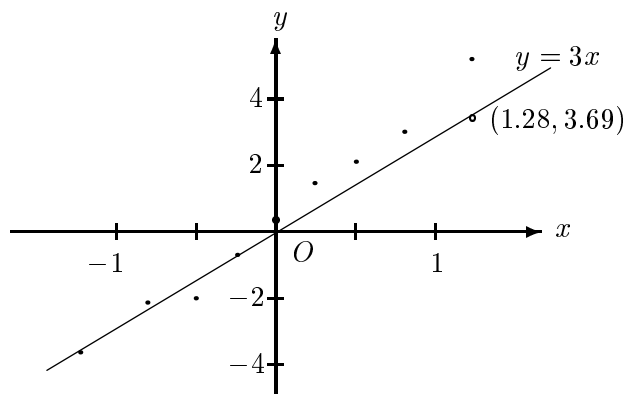


圖6.6 常態分佈之Q-Q圖

當有 n 個樣本點，有些書的常態分佈之Q-Q圖為繪點 $(x_{p_i}, y_{(i)})$, $i = 1, \dots, n$ ，其中 $p_i = (i - 1/2)/n$ ，而 x_{p_i} 為 $\Phi^{-1}(p_i)$ 。有些書則取 $p_i = (i - 3/8)/(n + 1/4)$ 。

如果是 $(p_i, y_{(i)})$, $i = 1, \dots, n$ ，的圖，便稱**分位圖**(quantile plot，不過分位-分位圖有時也簡單的稱為分位圖)。其中 p_i 則如前通常取為 $i/(n + 1)$ (或 $(i - 1/2)/n$ ，或 $(i - 3/8)/(n + 1/4)$)。與盒鬚圖相比，分位圖展示所有的觀測，及較多的分位數。

有Q-Q圖亦有**P-P圖**(probability-probability plot)。仍以常態分佈為例，點 $(p_i, \Phi(y_{(i)}))$, $i = 1, \dots, n$ ，之圖形，即為P-P圖。又在上述Q-Q圖及P-P圖的定義中，常態分佈可改為其他分佈。如對 χ_r^2 分佈，則得 χ_r^2 之Q-Q圖及P-P圖等。

圖6.7為 $\mathcal{E}(1)$ 及 $\mathcal{N}(0, 1)$ 分佈之P-P圖。

4.6.5 直方圖

機率密度函數的直方圖，在1.4節已引進。對於數據，直方圖也是常用的圖示法。一般可將數據按數值相近而分成若干群。再將水平軸分成若

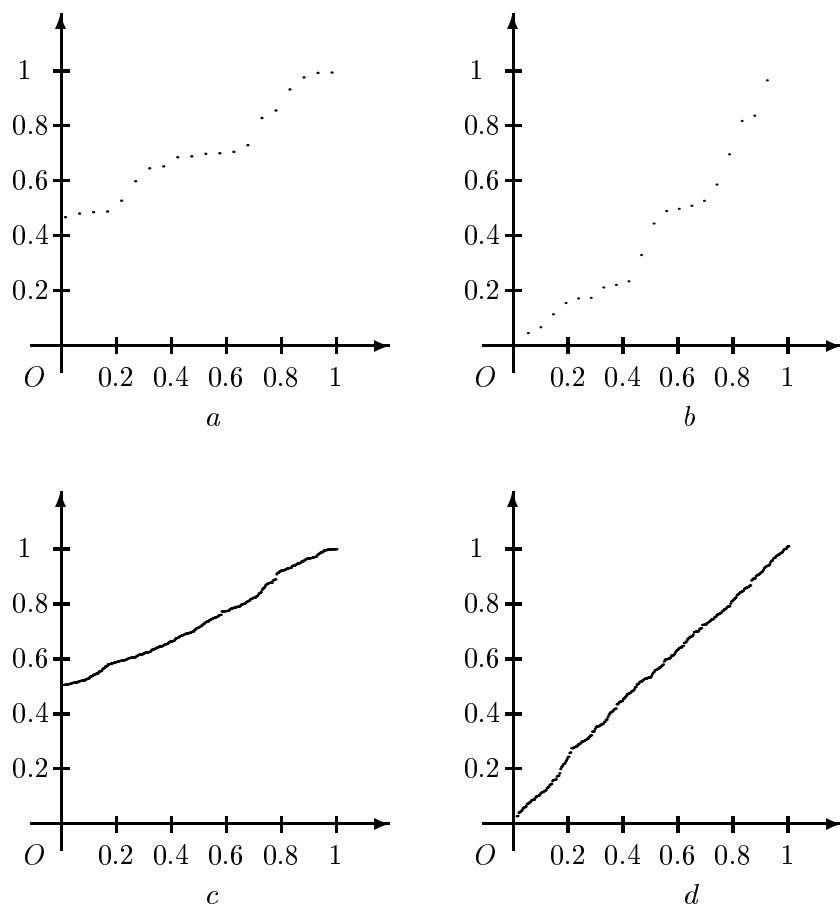


圖6.7 圖a,b分別為 $\mathcal{E}(1)$ 及 $\mathcal{N}(0,1)$ 分佈20個樣本點之常態分佈的P-P圖,
圖c,d則為300個樣本點之常態分佈的P-P圖

于區間，每一區間的之中心為每一群之中點，半徑為每一群數據寬度之半。然後在該區間繪一長方形，長方形的面積與該群數據的頻率(即個數)成正比。

要注意的是直方圖是以面積來反映頻率大小，而非高度。而且每一長方形之底部寬度也可不同。又在繪直方圖時，有時難以決定的是該分幾群？每群之寬度是否要相同？要知不同的分法，影響到平均值及變異數等。

例6.4 一袋中有紙牌分別為 $0, 1, \dots, 9$ 。每次隨機地取一張, 每次取出後放回, 共363次。 $0, 1, \dots, 9$ 分別得到的次數為

35, 44, 25, 40, 50, 26, 33, 38, 37, 35。

圖6.8給出其直方圖。

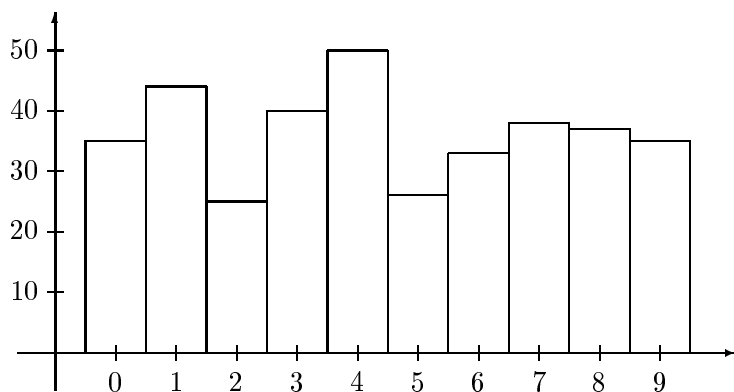


圖6.8 例6.4中之直方圖

習題

第4.1節

1. 試說明母體與樣本空間之分別。
2. 試證自一有 N 張紙牌(分別寫數字1至 N)的袋子中, 依序每次取出一張紙牌, 令 X_i 表第 i 次所得之點數 $1 \leq i \leq n$ 。則對任意 $n \leq N$, $X_1, \dots, X_n$ 之邊際分佈相同。
3. 試證定理1.2。
4. 試證系理1.1。
5. 設 $X_1, \dots, X_n$ 為一組隨機樣本。試證

$$(i) \overline{X}_{n+1} = (n\overline{X}_n + \overline{X}_{n+1})/(n+1);$$

$$(ii) nS_{n+1}^2 = (n-1)S_n^2 + (n/(n+1))(X_{n+1} - \overline{X}_n)^2。$$

6. 試以褶積公式重做第三章定理2.4((ii)除外)。
7. 在例1.6中, 試說明何以每次取出後不放回之 $\overline{X}_n$ 的期望值與每次取出後放回者相同, 但變異數較小。並說明在每次取出後不放回的情況下, $\text{Cov}(X_1, X_2) < 0$ 是否合理。
8. 設 $X_1, \dots, X_n$ 為一組由 $\Gamma(\alpha, \beta)$ 分佈所產生之隨機樣本。試證
 - (i) $E(\overline{X}_n^k) = (n\alpha + k - 1) \cdots (n\alpha + 1)n\alpha\beta^k/n^k, k \geq 1;$
 - (ii) $E(S_n^2) = \alpha\beta^2$ 。

第4.2節

1. 設 X_n 有 $P(n\lambda)$ 分佈, $n \geq 1, \lambda > 0$ 。試證 $X_n/n \xrightarrow[n \rightarrow \infty]{P} \lambda$, 且 $(X_n - n\lambda)/(\sqrt{n\lambda}) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, 1)$ 。
2. 投擲一骰子, 以獲得對點數6之出現機率 p 之資訊。試問需投擲多少次, 才會使對 p 之估計誤差小於0.01之機率至少是0.9。(解. 25,000次)
3. 某新產品電池壽命之期望值 μ 未知, 但根據過去之經驗可假設壽命之標準差不超過50小時。試問至少要測試多少電池, 才會使壽命之樣本平均與 μ 之差異不超過20小時之機率可達(i) 0.8, (ii) 0.9。(解. (i) 32個, (ii) 63個)
4. 某投資者欲得到每週投資報酬率之期望值 μ 的資訊。若知投資報酬率之變異數不超過0.0004, 則須觀測至少多少週之資料, 才能獲得對 μ 之估計誤差小於0.005之機率至少為0.8。(解. 80週)
5. 試 $X_1, \dots, X_8$ 之觀測值為3, 6, 4.5, 3, 8, 5, 6, 7。試給出經驗分佈函數, 並繪其圖。
6. 某政府單位對某項議題做民調。以1表極滿意, 2表滿意, 3表普通, 4表不滿意, 5表不清楚。在45次民調中, 得到下述平均數(求至小數第一位):

1.9, 1.6, 2.5, 2.3, 2.0, 2.5, 2.5, 1.4, 1.8, 3.1, 3.0, 2.3, 2.2, 1.8, 2.5,

1.9, 2.3, 2.1, 2.2, 2.2, 1.5, 2.7, 2.0, 1.8, 2.2, 2.1, 2.0, 2.5, 2.3, 2.2,
1.7, 1.9, 2.4, 2.2, 2.3, 2.2, 2.1, 2.7, 2.2, 2.1, 2.0, 2.1, 2.7, 2.5, 2.6。

試給出此組樣本之經驗分佈函數的圖形。

7. 投擲一公正的骰子100次, 試估計點數和會介於330與380間之機率。

8. 投擲一公正的骰子12,000次。

(i) 試估計點數6出現的次數介於1,800至2,200次之機率;

(ii) 若點數6出現2,500次, 你是否會懷疑此非一公正的骰子? 試說明理由。

9. 設 X_n 有 $\Gamma(n, \lambda)$ 分佈, $n \geq 1$ 。試證 $X_n/n \xrightarrow[n \rightarrow \infty]{p} \lambda$, 並利用中央極限定理, 估計 n 很大時, $P(X_n \leq x)$ 之值, $x \in R$ 。

10. 設 X 有 $\Gamma(100, 1)$ 分佈, 試利用上題求 $P(X \leq 111)$ 之近似值。(解. 0.8643)

11. 設 $X_1, \dots, X_{36}$ 為自 $Be(1, 2)$ 分佈所產生之隨機樣本。 $\bar{X}_{36}$ 為其樣本平均。試估計 $P(\bar{X}_{36} \leq 0.4)$ 。(解. 0.9551)

12. 設有一組樣本數為100之隨機樣本, 期望值為 μ , 變異數為 σ^2 。試估計 $\bar{X}_{100}$ 與 μ 差異不超過 $\sigma/10$ 之機率。(解. 0.68)

13. 某航空公司之某航班有200個座位, 但有215位預訂。假設每位預訂者相互獨立地有0.9的機率會按時搭乘。試估計該航班會有5位乘客無法搭乘之機率。(解. 0.0045)

14. 設 $X_1 \sim \mathcal{U}[0, 1]$, 令 $X_n = X_1^n$, $n \geq 1$ 。試證 $X_n \xrightarrow[n \rightarrow \infty]{p} 0$ 。

15. 設 $X_1, X_2, \dots$ 為一組取正值之隨機變數, 且 $X_n \xrightarrow[n \rightarrow \infty]{p} a$, $a \neq 0$ 。

(i) 令 $Y_n = \sqrt{X_n}$, $n \geq 1$, 試證 $Y_n \xrightarrow[n \rightarrow \infty]{p} \sqrt{a}$;

(ii) 令 $Z_n = a/X_n$, $n \geq 1$, 試證 $Z_n \xrightarrow[n \rightarrow \infty]{p} 1$;

(iii) 利用(i)及(ii), 試證在例2.5中, 若知 $S_n^2 \xrightarrow[n \rightarrow \infty]{p} \sigma^2$, $\sigma \neq 0$,

則 $\sigma/S_n \xrightarrow[n \rightarrow \infty]{p} 1$ 。

第4.3節

1. 試證(3.10)式。
2. 試證若 T 有 $\mathcal{T}_n$ 分佈, 則對 $l \in R$, $E(T^l)$ 存在, 若且唯若 $-1 < l < n$ 。
3. 試證(3.11)及(3.12)二式。
4. 試證(3.15)式可改寫為(3.16)式。
5. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{N}(\mu, \sigma^2)$ 分佈所產生之隨機樣本。試求 $-S_n^2$ 之函數, 以 $g(S_n^2)$ 表之, 使得 $E(g(S_n^2)) = \sigma$ 。
6. 設 X_1, X_2, X_3 為獨立的隨機變數, 且 X_i 有 $\mathcal{N}(i, i^2)$ 分佈, $i = 1, 2, 3$ 。試利用 X_1, X_2, X_3 的函數, 分別造出有如下的分佈。
(i) χ_3^2 , (ii) $\mathcal{T}_2$, (iii) $\mathcal{F}_{1,2}$ 。
7. 設 X 有 $\mathcal{F}_{n,m}$ 分佈。
(i) 試導出 X 之p.d.f.如(3.19)式;
(ii) 試證 $1/X$ 有 $\mathcal{F}_{m,n}$ 分佈;
(iii) 試證 $(n/m)X/(1 + (n/m)X)$ 有 $\mathcal{Be}(n/2, m/2)$ 分佈。
8. 試證(3.20)式, 並得到(3.21)及(3.22)二式。
9. 設 $X \sim \mathcal{F}_{2,2b}$, 其中 b 為一正整數。試證 $P(X > x) = (1 + x/b)^{-b}$, $\forall x > 0$ 。
10. 設 X_1 與 X_2 獨立, 且皆有 $\mathcal{N}(0, 25)$ 分佈。令 $D = \sqrt{X_1^2 + X_2^2}$ 。試求 $P(D \leq 12.25)$ 。(解. 約0.95)
11. 設 X 有 $\mathcal{N}(0, 1)$ 分佈。
(i) 試以附表4求 $P(X^2 \leq 3.84)$;
(ii) 試以附表5求 $P(X^2 \leq 3.84)$ 。
(解. (i) 0.95, (ii) 0.95)

12. 設 X 有 $\mathcal{F}_{6,6}$ 分佈, 試求 $P(X > 0.172)$ 。(解. 約0.975)
13. 設 $X_1, \dots, X_{16}$ 為一組由 $\mathcal{N}(0, 1)$ 分佈所產生之隨機樣本。令 $\bar{X}$ 表其樣本平均。試求
- (i) $P(\bar{X} < 0.5)$, (ii) $P(X_1 - X_2 \leq 2)$, (iii) $P(X_1 + X_2 \leq 2)$,
 (iv) $P(\sum_{i=1}^{16} X_i^2 \leq 32)$, (v) $P(\sum_{i=1}^{16} (X_i - \bar{X})^2 \leq 25)$ 。
- (解. (i) 0.9772, (ii) 0.921, (iii) 0.921, (iv) 0.988, (v) 0.95)
14. 設 $X_1, \dots, X_9$ 為一組由 $\mathcal{N}(6, 25)$ 分佈所產生之隨機樣本。試求(i) $P(3 < \bar{X}_9 < 7)$, (ii) $P(1.860 < 3(\bar{X}_9 - 6)/S_9)$, (iii) $P(S_9^2 \leq 31.9375)$ 。(解. (i) 0.6898, (ii) 0.05, (iii) 0.75)
15. 設 Z, V_1 及 V_2 相互獨立, 且 $Z \sim \mathcal{N}(0, 1)$, $V_1 \sim \chi_5^2$, $V_2 \sim \chi_9^2$ 。試求
- (i) $P(V_1 + V_2 \leq 8.6)$;
 (ii) $P(Z/\sqrt{V_1/5} \leq 2.015)$;
 (iii) $P(Z > 0.611\sqrt{V_2})$;
 (iv) $P(V_1/V_2 \leq 1.450)$;
 (v) 滿足 $P(V_1/(V_1 + V_2) \leq b) = 0.90$ 之 b 值。
- (解. (i) 0.144, (2) 0.95, (iii) 0.05, (iv) 0.90, (v) 0.592)
16. 設 $X_1, \dots, X_{13}$ 為一組由 $\mathcal{N}(\mu_X, \sigma_X^2)$ 分佈所產生之隨機樣本, $Y_1, \dots, Y_7$ 為一組由 $\mathcal{N}(\mu_Y, \sigma_Y^2)$ 分佈所產生之隨機樣本。且二組樣本相互獨立。又已知 $\sigma_X^2 = 10$, $\sigma_Y^2 = 12$ 。試求 $P(S_X^2/S_Y^2 > 3.33)$ 。(解. 約0.05)
17. 某地在西元2001年8月的21天中, 臭氧(Ozone)之指數為69, 70, 40, 35, 59, 40, 82, 48, 49, 35, 60, 59, 55, 46, 78, 77, 49, 60, 37, 38, 56。假設每日臭氧指數有常態分佈。試判斷是否有夠強的證據支持臭氧指數之平均值超過48。(解. $T = 2.0$, 證據夠強)
18. 承上題。在西元2002年8月的21天, 得到21筆臭氧指數: 27, 46, 47, 88, 44, 52, 46, 70, 69, 77, 41, 43, 75, 70, 37, 55, 56, 62, 28, 58, 69。試判斷在

兩年的同一段期間, 臭氣指數之變異數是否有差別。(解. $F = 0.796$, 沒有顯著差別)

19. 承例3.6。假設在二月28天中, 於同一時段到圖書館的人數為

53, 62, 39, 14, 25, 45, 42, 70, 53, 47, 18, 25, 46, 38,

52, 57, 38, 17, 25, 57, 43, 63, 51, 43, 17, 27, 68, 57。

試判斷這兩個不同月份中, 該時段到館人數之變異數是否顯著差別。(解. $F = 0.736$, 沒有顯著差別)

20. 設 $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $i = 1, \dots, n$, $Z_j \sim \mathcal{N}(0, 1)$, $j = 1, \dots, m$, 且這些隨機變數相互獨立。又令 $\bar{X}$, $\bar{Z}$ 分別表 $\{X_i\}$ 及 $\{Z_j\}$ 之樣本平均, S_Z 表 $\{Z_j\}$ 之樣本標準差。對下述隨機變數, 若為常見分佈則給出其分佈, 否則以不知回答。

(i) $X_1 - X_2$, (ii) $X_2 + 2X_3$, (iii) $(X_1 - X_2)/(\sigma S_Z \sqrt{2})$, (iv) $\sqrt{n}(\bar{X} - \mu)/(\sigma S_Z)$, (v) $Z_1^2 + Z_2^2$, (vi) $Z_1^2 - Z_2^2$, (vii) $Z_1/\sqrt{Z_2^2}$, (viii) Z_1^2/Z_2^2 , (ix) Z_1/Z_2 , (x) $\bar{X}/\bar{Z}$, (xi) $\sqrt{nm}(\bar{X} - \mu)/(\sigma(\sum_{j=1}^m Z_j^2)^{1/2})$, (xii) $m\bar{Z}^2$, (xiii) $\sum_{i=1}^n (X_i - \mu)^2/\sigma^2 + \sum_{j=1}^m (Z_j - \bar{Z})^2$, (xiv) $\bar{X}/\sigma^2 + \sum_{j=1}^m Z_j/m$, (xv) $(m-1)\sum_{i=1}^n (X_i - \bar{X})^2/((n-1)\sigma^2 \sum_{j=1}^m (Z_j - \bar{Z})^2)$ 。

(解. (i) $\mathcal{N}(0, 2\sigma^2)$, (ii) $\mathcal{N}(3\mu, 5\sigma^2)$, (iii) $\mathcal{T}_{m-1}$, (iv) $\mathcal{T}_{m-1}$, (v) χ_2^2 , (vi) 不知, (vii) $\mathcal{T}_1$, (viii) $\mathcal{F}_{1,1}$, (ix) $\mathcal{C}(0, 1)$, (x) 不知, (xi) $\mathcal{T}_m$, (xii) χ_1^2 , (xiii) χ_{n+m-1}^2 , (xiv) $\mathcal{N}(\mu/\sigma^2, (n\sigma^2)^{-1} + m^{-1})$, (xv) $\mathcal{F}_{n-1, m-1}$)

21. 試導出(3.25)及(3.27)式。

第4.4節

1. 在例4.1中試證(4.11)式成立。
2. 試證定理4.4。
3. 試證定理4.5。
4. 試證定理4.6。

5. 試證隨機變數 $X_1, \dots, X_n$ 之聯合分佈, 可由知道 $\sum_{j=1}^n \alpha_j X_j, \forall \alpha_j \in R$, 之分佈所決定。

6. 設 $\mathbf{X} = (X_1, X_2, X_3)$ 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈, 其中

$$\boldsymbol{\mu} = (0, 0, 1), \quad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}。$$

試求 $\mathbf{X}$ 之 p.d.f.。(解. $f_{\mathbf{X}}(x_1, x_2, x_3) = ((2\pi)^{3/2} \sqrt{15})^{-1} \exp\{(-1/2) \cdot (4/15)x_1^2 + (2/15)x_1x_2 + (4/15)x_2^2 + (x_3 - 1)^2\}$)

7. 設 (X, Y) 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈, 其中

$$\boldsymbol{\mu} = (84, 87), \quad \boldsymbol{\Sigma} = \begin{pmatrix} 49.2 & 40.8 \\ 40.8 & 56.3 \end{pmatrix}。$$

試求 $P(X \leq 78)$, $P(Y > 90)$ 及 $\rho(X, Y)$ 。(解. 0.195, 0.345, 0.78)

8. 設 (X, Y) 有 $\mathcal{MN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 分佈, 其中

$$\boldsymbol{\mu} = (7.4, 18.5), \quad \boldsymbol{\Sigma} = \begin{pmatrix} 0.000625 & -0.02844 \\ -0.02844 & 3.0625 \end{pmatrix}。$$

試求 $P(X > 7.45)$, $P(Y \leq 15)$, 及 $\rho(X, Y)$ 。(解. 0.023, 0.023, -0.65)

9. 設 X_1, X_2, X_3 相互獨立, 且 $X_1 \sim \mathcal{N}(0, 1)$, $X_2 \sim \mathcal{N}(4, 4)$, $X_3 \sim \mathcal{N}(-1, 2)$ 。令 $Y_1 = X_1 - X_2 + X_3$, $Y_2 = X_1 + X_3$ 。試求 Y_1, Y_2 之聯合分佈。(解. $\mathcal{N}(-5, -1, 7, 3, 3/\sqrt{21})$)

第4.5節

1. 在例5.3中, 試證利用(5.1)及(5.5)式所給之 $X_{(1)}$ 及 $(X_{(1)}, X_{(3)})$ 之 p.d.f. 如(5.8)及(5.9)式所給。

2. 設 X_1, X_2, X_3 為一組由 $\mathcal{W}(4, 2)$ 分佈所產生之隨機樣本。試求 $P(X_{(1)} > 0.5)$ 及 $P(X_{(3)} > 0.5)$ 。(解. $e^{-3}, 1 - (1 - e^{-1})^3$)

3. 自 $\{1, 2, 3, 4, 5\}$ 隨機地取三數, 每次取出後不放回。求取出之最小數等於1之機率。(解. $3/5$)

4. 設 $X_1, \dots, X_6$ 為一組由 $\mathcal{U}(0, 1)$ 分佈所產生之隨機樣本。令 $R = X_{(6)} - X_{(1)}$ 。試求 $P(R > 1/2)$ 。(解. $57/64$)
5. 在例5.4中, 試求 $E(X_{(1)}X_{(n)})$ 及 $E(2X_{(n)} - 3X_{(1)})$ 。
6. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x) = 1/x^2, 1 \leq x < \infty$, 所產生之隨機樣本。試分別給出 $X_{(1)}, X_{(n)}$ 及 $X_{(r)}$ 之 p.d.f., 其中 $r = (n+1)/2$, n 為奇數。(解. $g_{X_{(1)}}(y_1) = n/y_1^{n+1}, y_1 \geq 1, g_{X_{(n)}}(y_n) = n(y_n - 1)^{n-1}/y_n^{n+1}, y_n > 1, g_{X_{(r)}}(y_r) = (2r-1)!(y_r-1)^{r-1}/(((r-1)!)^2 y_r^{2r}), y_r > 1$)
7. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(1)$ 分佈所產生之隨機樣本。試分別給出 $X_{(1)}, X_{(n)}, R = X_{(n)} - X_{(1)}, (X_{(1)}, \dots, X_{(k)})$ 之 p.d.f., $1 \leq k \leq n$ 。(解. $g_{X_{(1)}}(y_1) = ne^{-ny}, y_1 > 0, g_{X_{(n)}}(y_n) = ne^{-y_n}(1 - e^{-y_n})^{n-1}, y_n > 0, g_R(r) = (n-1)e^{-r}(1 - e^{-r})^{n-2}, r > 0, g_{X_{(1)}, \dots, X_{(k)}}(y_1, \dots, y_k) = (n!/(n-k)!) \exp\{-\sum_{i=1}^k y_i - (n-k)y_k\}, 0 < y_1 < \dots < y_k$)
8. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(\alpha, \beta)$ 分佈所產生之隨機樣本, $\alpha < \beta$ 。令 $R = X_{(n)} - X_{(1)}$ 。試求
 - (i) $X_{(1)}, X_{(n)}$ 之聯合 p.d.f.;
 - (ii) R 之 p.d.f.;
 - (iii) $E(R)$ 。
9. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\lambda)$ 分佈所產生之隨機樣本。試證 $nX_{(1)}$ 與 X 有相同的分佈。
10. 設 X, Y 為二獨立的隨機變數, 且皆有 $\mathcal{N}(0, 1)$ 分佈。又令 $Z = \min\{X, Y\}$ 。試證 Z^2 有 χ_1^2 分佈, 即 $Z^2 \stackrel{d}{=} X^2$ 。
11. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{U}(0, \theta)$ 分佈所產生之隨機樣本, $\theta > 0$ 。試證 $X_{(1)}/X_{(n)}$ 與 $X_{(n)}$ 獨立。
12. 上題之結果可加以推廣。假設母體之 p.d.f. 為

$$f(x) = \frac{a}{\theta^a} x^{a-1}, 0 < x < \theta。$$

其中 $a > 0$ 為一常數。試證 $X_{(1)}/X_{(2)}, X_{(2)}/X_{(3)}, \dots, X_{(n-1)}/X_{(n)}$ 為相互獨立, 並分別求其分佈。

13. 在例5.5中, 試求(i) 1個月內, (ii) 5個月內, 燈管至少壞掉兩支之機率。
14. 在例5.5中, 求至少有兩支燈管用了10年以上之機率。
15. 在例5.6中, 試依定義2.1, 直接證明(5.10)式。
16. 在例5.6中, 令 $Z_n = X_{(1)}$ 。試證 $Z_n \xrightarrow[n \rightarrow \infty]{p} 0$ 。
17. 設 $X_1, \dots, X_n$ 為一組由 $\mathcal{E}(\lambda)$ 分佈所產生之隨機樣本。令 $Z_n = X_{(1)}$ 。試證 $Z_n \xrightarrow[n \rightarrow \infty]{p} 0$ 。
18. 在例5.7中, 試證 $U_n \xrightarrow[n \rightarrow \infty]{d} U$ 。
19. 設 $X_1, \dots, X_n$ 為一組由 p.d.f. $f(x) = x^{-2}, x \geq 1$, 所產生之隨機樣本。試求 $\{X_{(1)}^n, n \geq 1\}$ 之極限分佈。(解. $X_{(1)}^n \xrightarrow[n \rightarrow \infty]{d} X_1$)
20. 設 $X_1, \dots, X_n$ 為一組由分佈函數 $F(x) = (1 + e^{-x})^{-1}, x \in R$, 所產生之隨機樣本。令 $Y_n = X_{(n)} - \log n$ 。試求 $\{Y_n, n \geq 1\}$ 之極限分佈函數 $G(y)$ 。(解. $G(y) = \exp\{-\exp\{-y\}\}, y \in R$)

4.6節 習題

1. 對下述數據試給出點圖、莖葉圖及盒鬚圖:

7.2, 6.5, 11.3, 5.2, 8.4, 6.5, 10.9, 5.8, 7.9, 9.4, 6.5, 5.6,
8.6, 10.5, 7.8, 6.7, 10.3。

2. 對下述數據, 試給出其盒鬚圖:

6.8, 4.0, 6.3, 2.7, 1.7, 3.9, 6.7, 1.2, 0.7, 2.4, 3.0, 3.9, 4.5,
3.3, 0.9, 0.7, 2.2, 0.6, 0.8, 3.4, 1.4, 2.2, 2.5, 3.6, 3.9, 1.9,
3.8, 1.5, 5.8, 4.4, 2.1, 1.0, 1.2, 3.3, 1.5, 5.9, 2.4, 3.1, 1.5。

3. 某班14位學生，三次數理統計測驗成績分別為

測驗1 : 74, 85, 52, 87, 56, 61, 92, 57, 93, 90, 76, 78, 73, 51,

測驗2 : 66, 61, 76, 86, 72, 76, 70, 52, 85, 91, 90, 92, 88, 68,

測驗3 : 78, 90, 98, 93, 76, 80, 92, 62, 93, 90, 83, 97, 83, 82

試分別給其盒鬚圖，並指出哪一次測驗較佳，那一次測驗較差。(解. 測驗3較佳，測驗1與2之平均差不多，但測驗1成績較分散)

4. 試證樣本的第 $i/(n+1)$ 分位數，即樣本的第 i 個順序統計量。

5. 自 $\mathcal{U}[0, 1]$ 分佈模擬產生14個點：

0.46, 0.01, 0.67, 0.42, 0.78, 0.10, 0.41,

0.24, 0.53, 0.98, 0.99, 0.41, 0.80, 0.19。

試繪常態分佈之分位圖，並判斷是否像落在一直線上。(解. 並不易看出不落在一直線上，樣本數不夠多為一原因)

6. 設有16筆數據

6.70, 6.86, 6.86, 6.93, 6.74, 7.09, 6.94, 6.83,

6.81, 7.05, 7.02, 6.88, 6.63, 6.74, 6.86, 6.97。

試繪其常態分佈之分位圖，並判斷數據是否源自常態分佈。(解. 是)

7. 試對第二章表2.4，繪末一位數碼之直方圖。

參考文獻

1. 葉偉文譯(2001). 統計，改變了世界(Salsburg, D.原著: The Lady Tasting Tea)。天下遠見出版股份有限公司, 台北。
2. Anderson, T.W.(1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. John Wiley & Sons, New York.

3. Chung, K.L.(1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.
4. Graybill, F.A.(1983). *Matrices with Applications in Statistics*, 2nd ed. Wadsworth Publishing Company, Belmont, California.
5. Hogg, R.V. and Craig, A.T.(1970). *Introduction to Mathematical Statistics*, 3rd ed. Macmillan College Publishing Company, New York.
6. Roussas, G.G.(1997). *A Course in Mathematical Statistics*, 2nd ed. Academic Press, New York.
7. Tukey, J.W.(1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, Massachusetts.