

## 第三章

# 多維隨機變數

### 3.7 多維分佈

前幾節我們主要討論兩個變數之多維分佈。一般當然有 $n$ 維的隨機向量 $(X_1, \dots, X_n)$ ,  $n \geq 2$ 。我們以粗體的 $\mathbf{X}$ 表隨機向量 $(X_1, \dots, X_n)$ , 而以粗體的 $\mathbf{x}$ 表觀測值 $(x_1, \dots, x_n)$ 。

隨機向量 $\mathbf{X} = (X_1, \dots, X_n)$ 以 $R^n$ 之一子集為其樣本空間。若 $\mathbf{X}$ 為一離散型的隨機向量(樣本空間假設為可數的), 則其聯合p.d.f.為

$$\begin{aligned} f(\mathbf{x}) &= f(x_1, \dots, x_n) \\ &= P(X_1 = x_1, \dots, X_n = x_n), (x_1, \dots, x_n) \in R^n. \end{aligned}$$

因此對任意 $A \subset R^n$ ,

$$(7.1) \quad P(\mathbf{X} \in A) = \sum_{\mathbf{x} \in A} f(\mathbf{x}).$$

若 $\mathbf{X}$ 為一連續型的隨機向量, 則其聯合p.d.f.為一滿足下述條件的函數 $f(x_1, \dots, x_n)$ :

$$(7.2) \quad P(\mathbf{X} \in A) = \int \cdots \int_A f(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

## 2 第三章 多維隨機變數

令  $g(\mathbf{x}) = g(x_1, \dots, x_n)$  為一實值函數。則  $g(\mathbf{X})$  為一隨機變數，且對連續型及離散型，其期望值分別為

$$(7.3) \quad E(g(\mathbf{X})) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\mathbf{x}) f(\mathbf{x}) d\mathbf{x},$$

$$(7.4) \quad E(g(\mathbf{X})) = \sum_{\mathbf{x} \in R^n} g(\mathbf{x}) f(\mathbf{x}).$$

除了積分改為在  $R^n$  上，上述這些定義與二變數時的情況類似。另外，有時為了簡便，僅寫成“ $X_1, \dots, X_n$  之聯合 p.d.f.”，省去  $\mathbf{X}$ ，對  $X_1, \dots, X_n$  也可不用加括號。

$(X_1, \dots, X_n)$  之任一子集的邊際 p.d.f.，可將聯合 p.d.f. 對其餘變數之所有可能的值積分或相加而得。例如，首  $k$  個變數  $(X_1, \dots, X_k)$ ,  $1 \leq k < n$ ，之邊際 p.d.f. 為

$$(7.5) \quad f(x_1, \dots, x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_{k+1} \cdots dx_n,$$

或

$$(7.6) \quad f(x_1, \dots, x_k) = \sum_{(x_{k+1}, \dots, x_n) \in R^{n-k}} f(x_1, \dots, x_n).$$

條件分佈也可類似地定義。如

$$(7.7) \quad f(x_{k+1}, \dots, x_n | x_1, \dots, x_k) = \frac{f(x_1, \dots, x_n)}{f(x_1, \dots, x_k)}.$$

如果為了更清楚，可將上式左側寫成

$$f_{X_{k+1}, \dots, X_n | X_1, \dots, X_k}(x_{k+1}, \dots, x_n | x_1, \dots, x_k),$$

特別是當  $x_i$ 's 為給定數值的時候。

**例7.1** 設取整數值隨機變數  $X, Y, Z$  之聯合 p.d.f. 為

$$f(x, y, z) = \frac{1}{216}, 1 \leq x \leq 6, x+1 \leq y \leq x+6, y+1 \leq z \leq y+6.$$

試驗證  $f(x, y, z)$  爲一 p.d.f., 並求  $f_{Z|X,Y}(z|x, y)$  及  $f_{Z|Y}(z|y)$ 。

解. 因

$$\sum_{x=1}^6 \sum_{y=x+1}^{x+6} \sum_{z=y+1}^{y+6} f(x, y, z) = \sum_{x=1}^6 \sum_{y=x+1}^{x+6} \frac{6}{216} = \sum_{x=1}^6 \frac{6}{36} = 1,$$

故  $f(x, y, z)$  確爲一 p.d.f.。

其次

$$f_{X,Y}(x, y) = \sum_{z=y+1}^{y+6} f(x, y, z) = \frac{6}{216} = \frac{1}{36}, 1 \leq x \leq 6, x+1 \leq y \leq x+6。$$

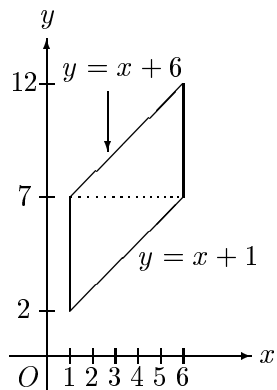
因此

$$f_{Z|X,Y}(z|x, y) = \frac{f(x, y, z)}{f_{X,Y}(x, y)} = \frac{1/216}{1/36} = \frac{1}{6},$$

$$y+1 \leq z \leq y+6, 1 \leq x \leq 6, x+1 \leq y \leq x+6。$$

又

$$f_{Y,Z}(y, z) = \begin{cases} \sum_{x=1}^{y-1} f(x, y, z) = \frac{y-1}{216}, & 2 \leq y \leq 7, y+1 \leq z \leq y+6, \\ \sum_{x=y-6}^6 f(x, y, z) = \frac{13-y}{216}, & 8 \leq y \leq 12, y+1 \leq z \leq y+6。 \end{cases}$$



因此

$$f_Y(y) = \begin{cases} \sum_{z=y+1}^{y+6} f_{Y,Z}(y, z) = \frac{6(y-1)}{216} = \frac{y-1}{36}, & 2 \leq y \leq 7, \\ \sum_{z=y+1}^{y+6} f_{Y,Z}(y, z) = \frac{6(13-y)}{216} = \frac{13-y}{36}, & 8 \leq y \leq 12. \end{cases}$$

故得

$$f_{Z|Y}(z|y) = \frac{f_{Y,Z}(y, z)}{f_Y(y)} = \frac{1}{6}, y+1 \leq z \leq y+6, 2 \leq y \leq 12.$$

比較 $f_{Z|X,Y}$ 與 $f_{Z|Y}$ ，可看出在給定 $Y = y$ 之下，多知道 $X$ 之值，並不影響 $Z$ 之p.d.f.。我們稱此為給定 $Y = y$ 之下， $Z$ 與 $X$ 為**條件獨立**(conditionally independent)。事實上可證明

$$(7.8) \quad f_{Z|X,Y}(z|x, y) = f_{Z|Y}(z|y)$$

與

$$(7.9) \quad f_{X,Z|Y}(x, z|y) = f_{X|Y}(x|y)f_{Z|Y}(z|y)$$

等價。即可以(7.8)或(7.9)式當做給定 $Y$ 之下， $Z$ 與 $X$ 為條件獨立之定義。

另外，

$$f_X(x) = \sum_{y=x+1}^{x+6} f_{X,Y}(x, y) = \frac{6}{36} = \frac{1}{6}.$$

因此 $X \sim \mathcal{D}\mathcal{U}(1, 6)$ 。讀者可以驗證 $Y \stackrel{d}{=} X_1 + X_2$ ，其中 $X_1, X_2$ 獨立且皆有 $\mathcal{D}\mathcal{U}(1, 6)$ 分佈。即 $X$ 相當於投擲一公正的骰子所得點數， $Y$ 相當於投擲二公正骰子之點數和。 $Z$ 的邊際分佈亦可求出，不過稍繁瑣些，在此略去。

**例7.2** 設 $X, Y, Z$ 之聯合p.d.f.為

$$f(x, y, z) = z^2 e^{-z(1+x+y)}, x, y, z \geq 0.$$

我們先證明 $f$ 的確為一p.d.f.:

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) dx dy dz \\
&= \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} z^2 e^{-z(1+x+y)} dx dy dz \\
&= \int_0^{\infty} \int_0^{\infty} -ze^{-z(1+x+y)} \Big|_0^{\infty} dy dz \\
&= \int_0^{\infty} \int_0^{\infty} ze^{-z(1+y)} dy dz \\
&= \int_0^{\infty} -e^{-z(1+y)} \Big|_0^{\infty} dz \\
&= \int_0^{\infty} e^{-z} dz = 1。
\end{aligned}$$

在上述過程中, 也附帶求出

$$f_Z(z) = e^{-z}, z \geq 0,$$

即  $Z$  有  $\mathcal{E}(1)$  分佈。因此在給定  $Z = z$  之下,  $X, Y$  之條件 p.d.f. 為

$$\begin{aligned}
f_{X,Y|Z}(x, y|z) &= \frac{f(x, y, z)}{f_Z(z)} = \frac{z^2 e^{-z(1+x+y)}}{e^{-z}} \\
&= z^2 e^{-z(x+y)}, x, y, z \geq 0。
\end{aligned}$$

由上式得

$$\begin{aligned}
f_{X|Z}(x|z) &= \int_0^{\infty} f_{X,Y|Z}(x, y|z) dy \\
&= \int_0^{\infty} z^2 e^{-z(x+y)} dy = ze^{-zx}, x \geq 0, z \geq 0。
\end{aligned}$$

同理有  $f_{Y|Z}(y|z) = ze^{-zy}, y \geq 0, z \geq 0$ 。因此

$$f_{X,Y|Z}(x, y|z) = f_{X|Z}(x|z)f_{Y|Z}(y|z)。$$

即在給定  $Z = z$  之下,  $X$  與  $Y$  為條件獨立, 且皆有  $\mathcal{E}(z)$  分佈。但若不給定  $Z = z$ , 則  $X$  與  $Y$  並不獨立:  $X, Y$  之聯合 p.d.f. 為

$$\begin{aligned}
f_{X,Y}(x, y) &= \int_0^{\infty} f(x, y, z) dz = \int_0^{\infty} z^2 e^{-z(1+x+y)} dz, \\
&= \frac{2}{(1+x+y)^3}, x, y \geq 0。
\end{aligned}$$

## 6 第三章 多維隨機變數

因  $f_{X,Y}(x, y)$  無法寫成  $g(x)h(y)$  的型式, 故  $X$  與  $Y$  不獨立。

附帶一提, 由  $f_{X,Y}(x, y)$  即得  $X$  之邊際 p.d.f. 為

$$f_X(x) = \int_0^\infty \frac{2}{(1+x+y)^3} dy = \frac{1}{(1+x)^2}, x \geq 0。$$

同理可得  $Y$  之邊際 p.d.f. 為

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

由此再度看出  $X$  與  $Y$  不獨立, 而且  $E(X), E(Y)$  皆不存在:

$$\begin{aligned} E(X) &= \int_0^\infty x f_X(x) dx = \int_0^\infty \frac{x}{(1+x)^2} dx = \infty, \\ E(Y) &= \int_0^\infty y f_Y(y) dy = \int_0^\infty \frac{y}{(1+y)^2} dy = \infty。 \end{aligned}$$

不過  $E(Z) = 1$  倒是存在的。

對於在給定  $Z$  之下,  $X$  與  $Y$  為條件獨立, 底下以另一方式來檢驗。首先  $X, Z$  之聯合 p.d.f. 為

$$\begin{aligned} f_{X,Z}(x, z) &= \int_0^\infty f(x, y, z) dy \\ &= \int_0^\infty z^2 e^{-z(1+x+y)} dy \\ &= z e^{-z(1+x)}, x, z \geq 0。 \end{aligned}$$

因此

$$f_{X|Z}(x|z) = \frac{f_{X,Z}(x, z)}{f_Z(z)} = \frac{z e^{-z(1+x)}}{e^{-z}} = z e^{-zx}, x, z \geq 0,$$

與之前所指出的在給定  $Z = z$  之下,  $X$  有  $\mathcal{E}(z)$  分佈吻合。同理可求出

$$f_{Y|Z}(y|z) = z e^{-zy}, y, z \geq 0。$$

故

$$f_{X,Y|Z}(x, y|z) = f_{X|Z}(x|z) f_{Y|Z}(y|z), x, y, z \geq 0。$$

讀者也可自行驗證  $f_{X|Y,Z}(x|y, z) = f_{X|Z}(x|z)$  (習題第1題)。

**多項分佈**(multinomial distribution)為二項分佈之一推廣,此為一重要的多維分佈。在此設 $n, k$ 為正整數,且常數 $p_1, \dots, p_k$ 滿足 $0 \leq p_i \leq 1, i = 1, \dots, k, \sum_{i=1}^k p_i = 1$ 。則隨機向量 $(X_1, \dots, X_k)$ 稱為有參數 $n, p_1, \dots, p_k$ 之**多項分佈**,以 $\mathcal{M}(n, p_1, \dots, p_k)$ 表之,若 $(X_1, \dots, X_k)$ 之聯合p.d.f.為

$$(7.10) \quad f(x_1, \dots, x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k},$$

其中 $x_1, \dots, x_k$ 為非負整數,且 $\sum_{i=1}^k x_i = n$ 。通常取 $0 < p_i < 1, i = 1, \dots, k$ 。

二項分佈可描述一有兩個不同的結果之試驗。有些試驗有超過兩個的結果,這時多項分佈便可派上用場。可看出二項分佈為多項分佈 $k = 2$ 之一特例:

$$f(x_1, x_2) = \frac{n!}{x_1! x_2!} p_1^{x_1} p_2^{x_2} = \binom{n}{x_1} p_1^{x_1} (1 - p_1)^{n - x_1}。$$

只是因 $k = 2$ 時,由於 $X_2 = n - X_1$ ,知道 $X_1$ 後, $X_2$ 便決定了,所以對 $k = 2$ 的情況,便僅簡單地以

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

表 $X_1$ 之p.d.f.。在第二章例2.11,將 $n$ 個球隨機地丟進 $n$ 個箱子中,若將箱子編號1至 $n$ ,且令 $X_1, \dots, X_n$ 表各箱中之球數,則 $(X_1, \dots, X_n)$ 便有 $\mathcal{M}(n, 1/n, \dots, 1/n)$ 分佈。底下再給一例。

**例7.3** 投擲一四面體10次,設四面為 $A, B, C, D$ ,其出現的機率分別為0.1, 0.2, 0.3及0.4。試求 $A, B, C, D$ 面分別出現1次,2次,3次及4次之機率。  
**解.** 令 $X_i$ 表第 $i$ 面出現的次數,  $i = 1, 2, 3, 4$ 。則 $(X_1, X_2, X_3, X_4)$ 有 $\mathcal{M}(10, 0.1, 0.2, 0.3, 0.4)$ 分佈。因此

$$\begin{aligned} & P(X_1 = 1, X_2 = 2, X_3 = 3, X_4 = 4) \\ &= \frac{10!}{1!2!3!4!} 0.1^1 0.2^2 0.3^3 0.4^4 \\ &= \frac{3,628,800}{288} \cdot 2.7648 \cdot 10^{-6} = 0.03483648。 \end{aligned}$$

在(7.10)式中,

$$(7.11) \quad \frac{n!}{x_1! \cdots x_k!} = \binom{n}{x_1, \dots, x_k}, \sum_{i=1}^k x_i = n, x_1, \dots, x_k \geq 0,$$

此即1.1節所定義的多項係數。多項係數就是 $n$ 個事物分成 $k$ 類, 第1類有 $x_1$ 個, 第2類有 $x_2$ 個,  $\dots$ , 第 $k$ 類有 $x_k$ 個, 總共之分法。而2.2.2節之二項式定理, 也可推廣至**多項式定理**(Multinomial theorem):

設 $n, k$ 為二正整數, 則對任意實數 $a_1, \dots, a_k$ ,

$$(7.12) \quad (a_1 + \cdots + a_k)^n = \sum \frac{n!}{x_1! \cdots x_k!} a_1^{x_1} \cdots a_k^{x_k},$$

其中和是對所有 $x_i$ 為非負整數, 且 $\sum_{i=1}^k x_i = n$ 而相加。有了多項式定理, 才保證(7.10)式所定義出的 $f$ 確為 $k$ 維的p.d.f.。

對多項分佈, 其邊際分佈, 及條件分佈, 直觀上便可以得到。說明如下:

考慮 $X_k$ , 若將第 $k$ 種結果視為成功, 其餘結果皆視為失敗, 則 $X_k$ 為 $n$ 個伯努力試驗成功的次數, 每次成功的機率皆為 $p_k$ 。因此,  $X_k$ 應有 $\mathcal{B}(n, p_k)$ 分佈。這點利用(7.6)式便可證明。對 $x_k \in \{0, 1, \dots, n\}$ , 我們必須將聯合p.d.f.對所有可能的 $(x_1, \dots, x_{k-1})$ 相加。即對所有非負整數的 $x_1, \dots, x_{k-1}$ , 且 $\sum_{i=1}^{k-1} x_i = n - x_k$ 相加。以 $B$ 表此集合。則

$$\begin{aligned} f(x_k) &= \sum_{(x_1, \dots, x_{k-1}) \in B} \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \\ &= \frac{n!}{x_k! (n - x_k)!} p_k^{x_k} \cdot \sum_{(x_1, \dots, x_{k-1}) \in B} \frac{(n - x_k)!}{x_1! \cdots x_{k-1}!} p_1^{x_1} \cdots p_{k-1}^{x_{k-1}}. \end{aligned}$$

因 $x_1 + \cdots + x_{k-1} = n - x_k$ , 且 $p_1 + \cdots + p_{k-1} = 1 - p_k$ , 利用多項式定理, 上述和 $= (p_1 + \cdots + p_{k-1})^{n-x_k} = (1 - p_k)^{n-x_k}$ 。故得

$$f(x_k) = \frac{n!}{x_k! (n - x_k)!} p_k^{x_k} (1 - p_k)^{n-x_k}.$$

因此 $X_k$ 有 $\mathcal{B}(n, p_k)$ 分佈。

同理可得證每一 $X_i$ 皆有 $\mathcal{B}(n, p_i)$ 分佈,  $1 \leq i \leq k$ 。

給定  $X_k = x_k$ , 則相當於有  $n - x_k$  次試驗, 而可能有  $k - 1$  個可能的結果,  $(X_1, \dots, X_{k-1})$  表  $n - x_k$  次試驗後, 各結果出現的次數。因此直觀上, 給定  $X_k = x_k$  之下,  $(X_1, \dots, X_{k-1})$  應仍有多項分佈。這是對的, 利用 (7.7) 式得

$$\begin{aligned} f(x_1, \dots, x_{k-1} | x_k) &= \frac{f(x_1, \dots, x_k)}{f(x_k)} \\ &= \frac{\frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}}{\frac{n!}{x_k!(n-x_k)!} p_k^{x_k} (1-p_k)^{n-x_k}} \\ &= \frac{(n-x_k)!}{x_1! \dots x_{k-1}!} \left(\frac{p_1}{1-p_k}\right)^{x_1} \dots \left(\frac{p_{k-1}}{1-p_k}\right)^{x_{k-1}}。 \end{aligned}$$

故  $X_1, \dots, X_{k-1} | X_k = x_k$  有  $\mathcal{M}(n - x_k, p_1/(1-p_k), \dots, p_{k-1}/(1-p_k))$ 。

事實上, 對任一  $(X_1, \dots, X_k)$  之一子集合, 在給定其餘  $X_i$  之值下, 皆有多項分佈。例如,  $X_1, \dots, X_i | X_{i+1} = x_{i+1}, \dots, X_k = x_k \sim \mathcal{M}(n - \sum_{j=i+1}^k x_j, p_1/(1 - \sum_{j=i+1}^k p_j), \dots, p_i/(1 - \sum_{j=i+1}^k p_j))$ 。

其次我們介紹隨機向量獨立的概念。

**定義 7.1** 設  $\mathbf{X}_1, \dots, \mathbf{X}_n$  為隨機向量, 以  $f(\mathbf{x}_1, \dots, \mathbf{x}_n)$  為其聯合 p.d.f., 且  $\mathbf{X}_i$  之邊際 p.d.f. 為  $f_{\mathbf{X}_i}(\mathbf{x}_i)$ ,  $i = 1, \dots, n$ 。則若對  $\forall(\mathbf{x}_1, \dots, \mathbf{x}_n)$ , 皆有

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = \prod_{i=1}^n f_{\mathbf{X}_i}(\mathbf{x}_i),$$

便稱  $\mathbf{X}_1, \dots, \mathbf{X}_n$  為相互獨立的隨機向量。若所有  $\mathbf{X}_i$  皆為一維, 則稱  $X_1, \dots, X_n$  為相互獨立的隨機變數。

如前, 相互獨立如果不會混淆的話, 有時只稱獨立。若  $X_1, \dots, X_n$  為相互獨立, 則知道其中幾個變數之值, 對其餘變數, 不能提供任何資訊。即利用定義 7.1, 可證明給定  $(X_{i_1}, \dots, X_{i_k})$ , 其中  $\{X_{i_1}, \dots, X_{i_k}\}$  為  $\{X_1, \dots, X_n\}$  之一子集, 其餘隨機變數之條件分佈, 與邊際分佈相同, 未做任何改變。

相互獨立導致每對獨立。即若  $X_1, \dots, X_n$  相互獨立, 則任二  $X_i, X_j, i \neq j$ , 亦獨立。但相互獨立比每對獨立還強, 即有可能每對獨立, 但不相互獨

立(見習題第15題)。亦可參考第一章例2.4, 事件的每對獨立, 卻不相互獨立的例子。

相互獨立的隨機變數, 有許多好的性質。底下幾個定理, 其證明均類似在3.2節及3.3節中, 兩個變數情況中對應的定理, 因此我們均略去證明。

**定理7.1** (定理2.2(ii)之推廣). 設 $X_1, \dots, X_n$ 為相互獨立的隨機變數, 且函數 $g_j(X_j), j = 1, \dots, n$ , 滿足 $E(g_j(X_j))$ 存在。則

$$(7.13) \quad E(g_1(X_1) \cdots g_n(X_n)) = \prod_{j=1}^n E(g_j(X_j))。$$

**定理7.2** (定理2.3之推廣).  $X_1, \dots, X_n$ 為相互獨立的隨機變數, 若且唯若

$$(7.14) \quad E(e^{it_1 X_1 + \cdots + it_n X_n}) = \prod_{j=1}^n E(e^{it_j X_j}), t_j \in R, j = 1, \dots, n。$$

**系理7.1** 設 $X_1, \dots, X_n$ 為相互獨立的隨機變數。令 $Z = X_1 + \cdots + X_n$ , 則

$$(7.15) \quad E(e^{itZ}) = \prod_{j=1}^n E(e^{itX_j}), t \in R。$$

特別地, 若 $X_1, \dots, X_n$ 有相同的分佈, 則

$$E(e^{itZ}) = (E(e^{itX_1}))^n, t \in R。$$

定理7.2及系理7.1, 亦有母函數, 拉普拉斯轉換及動差母函數之版本。

定理2.4可以推廣至 $n$ 個獨立隨機變數的情況。我們只給gamma分佈的結果。

**例7.4** 設 $X_1, \dots, X_n$ 為相互獨立的隨機變數, 且 $X_i$ 有 $\Gamma(\alpha_i, \beta)$ 分佈,  $i = 1, \dots, n$ 。因

$$E(e^{-sX_i}) = \frac{1}{(1 + \beta s)^{\alpha_i}}, s \geq 0,$$

故若令 $Z = X_1 + \cdots + X_n$ , 則

$$\begin{aligned}
E(e^{-sZ}) &= \prod_{i=1}^n E(e^{-sX_i}) = \prod_{i=1}^n \frac{1}{(1+\beta s)^{\alpha_i}} \\
&= \frac{1}{(1+\beta s)^{\sum_{i=1}^n \alpha_i}}, s \geq 0,
\end{aligned}$$

此為  $\Gamma(\sum_{i=1}^n \alpha_i, \beta)$  分佈之拉普拉斯轉換。故  $\sum_{i=1}^n X_i$  有  $\Gamma(\sum_{i=1}^n \alpha_i, \beta)$  分佈。即獨立且有 gamma 分佈之隨機變數，只要尺度參數皆相同，其和便仍有 gamma 分佈。

**例7.5** 設  $X_1, \dots, X_n$  為相互獨立的隨機變數，分別有動差母函數  $M_{X_i}(t)$ ,  $t \in R, i = 1, \dots, n$ 。又設  $a_1, \dots, a_n, b_1, \dots, b_n$  為常數。令  $Z = \sum_{i=1}^n (a_i X_i + b_i)$ 。則  $Z$  之動差母函數為

$$\begin{aligned}
M_Z(t) &= E(e^{tZ}) \\
&= E(e^{t \sum_{i=1}^n (a_i X_i + b_i)}) \\
&= e^{t \sum_{i=1}^n b_i} \prod_{i=1}^n E(e^{a_i t X_i}) \\
&= e^{t \sum_{i=1}^n b_i} \prod_{i=1}^n M_{X_i}(a_i t), t \in R.
\end{aligned}$$

利用上例，可得獨立的常態分佈，其線性組合之和仍為常態分佈。

**例7.6** 設  $X_1, \dots, X_n$  為相互獨立的隨機變數，分別有  $\mathcal{N}(\mu_i, \sigma_i^2)$  分佈， $i = 1, \dots, n$ 。令  $a_1, \dots, a_n, b_1, \dots, b_n$  為常數。則  $Z = \sum_{i=1}^n (a_i X_i + b_i)$  有  $\mathcal{N}(\sum_{i=1}^n (a_i \mu_i + b_i), \sum_{i=1}^n a_i^2 \sigma_i^2)$  分佈。

**證明.** 首先  $X_i$  之動差母函數為

$$M_{X_i}(t) = E(e^{tX_i}) = e^{\mu_i t + \sigma_i^2 t^2 / 2}, t \in R.$$

因此  $Z$  之動差母函數為

$$\begin{aligned}
M_Z(t) &= E(e^{tZ}) \\
&= e^{t \sum_{i=1}^n b_i} \prod_{i=1}^n M_{X_i}(a_i t)
\end{aligned}$$

$$\begin{aligned}
&= e^{t \sum_{i=1}^n b_i} \prod_{i=1}^n e^{\mu_i a_i t + \sigma_i^2 a_i^2 t^2 / 2} \\
&= e^{(\sum_{i=1}^n (a_i \mu_i + b_i))t + (\sum_{i=1}^n a_i^2 \sigma_i^2) t^2 / 2}, t \in R.
\end{aligned}$$

而此為  $\mathcal{N}(\sum_{i=1}^n (a_i \mu_i + b_i), \sum_{i=1}^n a_i^2 \sigma_i^2)$  之動差母函數。證畢。

底下為判定  $n$  個隨機向量相互獨立的一個定理，為定理 2.1 之推廣。

**定理 7.3** 設  $\mathbf{X}_1, \dots, \mathbf{X}_n$  為隨機向量。則  $\mathbf{X}_1, \dots, \mathbf{X}_n$  為相互獨立，若且唯若存在函數  $g_i(\mathbf{x}_i), i = 1, \dots, n$ ，使得  $(\mathbf{X}_1, \dots, \mathbf{X}_n)$  之聯合 p.d.f. 可寫成

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n) = g_1(\mathbf{x}_1) \cdots g_n(\mathbf{x}_n)。$$

底下為定理 3.2 之推廣。

**定理 7.4** 設  $\mathbf{X}_1, \dots, \mathbf{X}_n$  為獨立的隨機向量。令  $g_i(\mathbf{x}_i)$  為  $\mathbf{x}_i$  的函數， $i = 1, \dots, n$ 。則  $U_i = g_i(\mathbf{X}_i), i = 1, \dots, n$ ，為相互獨立。

$n$  維隨機向量的變數代換，只是二維隨機向量變數代換之一立即的推廣。例如，在做變換時，雅可彬為一  $n \times n$  矩陣之行列式的值。

底下我們給一定理 3.1 之推廣。

**例 7.7** 設  $X, Y, Z$  為相互獨立，分別有  $\Gamma(p, \lambda), \Gamma(q, \lambda)$  及  $\Gamma(r, \lambda)$  分佈。令

$$U = \frac{X}{X+Y}, \quad V = \frac{X+Y}{X+Y+Z}, \quad W = X+Y+Z。$$

則  $U, V, W$  相互獨立，且分別有  $Be(p, q), Be(p+q, r)$  及  $\Gamma(p+q+r, \lambda)$  分佈。

**證明.** 首先  $(X, Y, Z)$  之聯合 p.d.f. 為

$$f(x, y, z) = \frac{x^{p-1} e^{-x/\lambda}}{\Gamma(p) \lambda^p} \frac{y^{q-1} e^{-y/\lambda}}{\Gamma(q) \lambda^q} \frac{z^{r-1} e^{-z/\lambda}}{\Gamma(r) \lambda^r}, x, y, z > 0。$$

令

$$u = \frac{x}{x+y}, \quad v = \frac{x+y}{x+y+z}, \quad w = x+y+z, 0 < u, v < 1, w > 0,$$

解出

$$x = uvw,$$

$$y = -uvw + vw = vw(1 - u),$$

$$z = -vw + w = w(1 - v)。$$

則

$$J = \begin{vmatrix} vw & uw & uv \\ -vw & -uw + w & -uv + v \\ 0 & -w & -v + 1 \end{vmatrix} = vw^2。$$

故得

$$\begin{aligned} & f_{U,V,W}(u, v, w) \\ &= \frac{(uvw)^{p-1} e^{-uvw/p} (vw(1-u))^{q-1} e^{(uvw-vw)/\lambda} (w(1-v))^{r-1} e^{(vw-w)/\lambda}}{\Gamma(p)\Gamma(q)\Gamma(r)\lambda^{p+q+r}} \\ &= \frac{vw^2 \cdot u^{p-1}(1-u)^{q-1} v^{p+q-1} (1-v)^{r-1} w^{p+q+r-1} e^{-w/\lambda}}{\Gamma(p)\Gamma(q)\Gamma(r)\lambda^{p+q+r}} \\ &= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} u^{p-1} (1-u)^{q-1} \frac{\Gamma(p+q+r)}{\Gamma(p+q)\Gamma(r)} v^{p+q-1} (1-v)^{r-1} \\ &\quad \cdot \frac{w^{p+q+r-1} e^{-w/\lambda}}{\Gamma(p+q+r)\lambda^{p+q+r}}。 \end{aligned}$$

由此即知  $U, V, W$  獨立，且分別有  $\mathcal{B}e(p, q), \mathcal{B}e(p+q, r)$  及  $\Gamma(p+q+r, \lambda)$  分佈。

事實上此結果可推廣到  $n$  維的情況，見第五章例3.11。

底下為一有關隨機變數之和的期望值及變異數的定理。證明並不困難，因此留在習題中(第11題)。

**定理7.5** 設有隨機變數  $X_1, \dots, X_n$ 。

(i) 若  $E(X_i)$  皆存在， $i = 1, \dots, n$ ，則對任意常數  $a_1, \dots, a_n$ ，

$$(7.16) \quad E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)。$$

(ii) 若  $\text{Var}(X_i)$  皆存在,  $i = 1, \dots, n$ , 則對任意常數  $a_1, \dots, a_n$ ,

$$(7.17) \quad \text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i) + 2 \sum_{i < j} a_i a_j \text{Cov}(X_i, X_j);$$

若  $X_1, \dots, X_n$  為每對獨立, 則

$$(7.18) \quad \text{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \text{Var}(X_i)。$$

(iii) 若  $\text{Var}(X_i)$  及  $\text{Var}(Y_j)$  皆存在,  $i = 1, \dots, n, j = 1, \dots, m$ , 則對任意常數  $a_1, \dots, a_n, b_1, \dots, b_m$ ,

$$(7.19) \quad \text{Cov}\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \text{Cov}(X_i, Y_j)。$$

有些本來期望值或變異數不太容易計算的隨機變數, 若能表示成一些隨機變數的和, 則可利用定理7.5求之。我們給一些例子如下。

**例7.8** 在2.2.2節我們曾指出, 若  $Y_1, \dots, Y_n$  為有獨立的  $\text{Ber}(p)$  分佈之隨機變數, 則  $X = Y_1 + \dots + Y_n$  有  $\mathcal{B}(n, p)$  分佈。由於

$$E(Y_i) = p, \quad \text{Var}(Y_i) = p(1 - p), \quad i = 1, \dots, n,$$

利用(7.16)及(7.18)式, 可得

$$\begin{aligned} E(X) &= \sum_{i=1}^n E(Y_i) = np, \\ \text{Var}(X) &= \sum_{i=1}^n \text{Var}(Y_i) = np(1 - p)。 \end{aligned}$$

與2.2.2節所得相同, 但此處計算顯然容易多了。

**例7.9** 利用機器將  $n$  個人的薪資單放進  $n$  封貼有名字的信封, 由於作業錯誤, 變成隨機地將薪資單放進信封, 每個信封恰放進一份薪資單。令  $N$  表放對的信封數, 求  $E(N)$  及  $\text{Var}(N)$ 。

**解.** 首先 $N$ 之p.d.f.並不易給出。如果 $k$ 較接近 $n$ ,  $P(N = k)$ 尚還好寫出(見習題第19題), 但當 $k$ 離 $n$ 較遠,  $P(N = k)$ 就不是很好表示出來。幸好我們只是要求 $N$ 的期望值及變異數, 倒不用知道 $N$ 的p.d.f.。現先將 $N$ 表為 $n$ 個指示函數之和:

$$N = I_{A_1} + \cdots + I_{A_n}。$$

$I_{A_i} = 1$ 或 $0$ , 就視事件 $A_i$ 是否發生, 其中 $A_i$ 表第 $i$ 個信封放的是正確的薪資單之事件。注意 $A_1, \dots, A_n$ 既不互斥亦不獨立。但 $P(A_i) = 1/n$ , 因此 $E(I_{A_i}) = P(A_i) = 1/n, \forall i = 1, \dots, n$ 。由(7.16)式得

$$E(N) = \sum_{i=1}^n E(I_{A_i}) = n \cdot \frac{1}{n} = 1。$$

其次

$$E(N^2) = \sum_{i=1}^n E(I_{A_i}^2) + 2 \sum_{i < j} E(I_{A_i} I_{A_j})。$$

又 $E(I_{A_i}^2) = E(I_{A_i}) = 1/n$ , 且對 $i \neq j$ ,

$$E(I_{A_i} I_{A_j}) = P(A_i \cap A_j) = \frac{1}{n(n-1)}。$$

故

$$E(N^2) = n \cdot \frac{1}{n} + 2 \binom{n}{2} \frac{1}{n(n-1)} = 2。$$

因此

$$\text{Var}(N) = E(N^2) - (E(N))^2 = 2 - 1 = 1。$$

故在“亂放”的情況下, 平均會放對一封, 標準差亦不大, 也是1。

讀者也可試以(7.17)式求 $\text{Var}(N)$ (習題第20題)。

**例7.10** 將 $n$ 個球隨機地放進 $n$ 個箱子中。則空箱數 $N$ 之期望值為何?

**解.** 首先 $N$ 可表示為 $n$ 個指示函數之和:

$$N = I_1 + \cdots + I_n,$$

其中 $I_i = 1$ 或 $0$ , 就視第 $i$ 箱是否為空箱。則由(7.16)式, 得

$$E(N) = \sum_{i=1}^n E(I_i) = n \cdot P(\text{第 } i \text{ 箱為空箱})。$$

第  $i$  箱為空箱表  $n$  個球全要放進其餘的  $n-1$  箱中，故

$$P(\text{第 } i \text{ 箱為空箱}) = (1 - \frac{1}{n})^n。$$

因此

$$E(N) = n \cdot (1 - \frac{1}{n})^n。$$

例如，若  $n = 10$ ，則

$$E(N) = 10 \cdot (\frac{9}{10})^{10} \doteq 3.486(\text{箱})。$$

這就是第二章例2.11中，10個球隨機地放進10箱子中，空箱數之期望值。但此處的推導顯然容易多了。

空箱數佔全部箱數之比例  $N/n$  的期望值為

$$\frac{E(N)}{n} = (1 - \frac{1}{n})^n。$$

此即每一箱子會是空箱的機率。附帶一提，在第二章例2.11之後，我們提過每一箱的球數有  $\mathcal{B}(n, 1/n)$  分佈。 $(1 - 1/n)^n$  為此分佈取值0之機率。當  $n \rightarrow \infty$  時，此機率趨近至  $e^{-1} \doteq 0.367879$ 。此值剛好是  $\mathcal{P}(1)$  分佈取值0之機率。此點由波松近似亦可看出。

最後利用(7.17)式，可求出

$$\begin{aligned} (7.20) \quad \text{Var}(N) &= n \cdot ((1 - \frac{1}{n})^n - (1 - \frac{1}{n})^{2n}) \\ &\quad + n(n-1)((1 - \frac{2}{n})^n - (1 - \frac{1}{n})^{2n}), \end{aligned}$$

這部分留給讀者自行完成(習題第21題)。

**例7.11** 某食品公司為了促銷其產品，在每盒餅乾中均隨機地放一張風景卡片，且設總共有  $N$  張不同的卡片。

我們先求若每次只買一盒, 欲收集齊全 $N$ 張卡片所需買一盒數之期望值。可看出當已有 $n$ 張不同卡片時,  $0 \leq n < N$ , 欲得一張尚未有的卡片, 所需買一盒數 $X_n$ 有自1開始的幾何分佈, 且參數為 $(N-n)/N$ , 因此期望值為 $N/(N-n)$ 。故收集全部 $N$ 張卡片所需買一盒數 $X_0 + X_1 + \cdots + X_{N-1}$ 之期望值為

$$E(X_0 + X_1 + \cdots + X_{N-1}) = \sum_{n=0}^{N-1} \frac{N}{N-n} = N(1 + \frac{1}{2} + \cdots + \frac{1}{N})。$$

由微積分裡的結果知, 當 $N \rightarrow \infty$ 時,

$$(7.21) \quad 1 + \frac{1}{2} + \cdots + \frac{1}{N} - \log N \rightarrow \gamma,$$

其中 $\gamma$ 為**歐拉常數**(Euler's constant, Euler(1707-1783)為有史以來最多產的數學家, 其生平介紹見黃文璋(1999)第十一章), 其值約為0.5772。因此 $N$ 很大時, 所需買一盒數之期望值可以 $N(\log N + \gamma)$ 來估計。不過要注意 $N(1 + \frac{1}{2} + \cdots + \frac{1}{N})$ 與 $N(\log N + \gamma)$ 之差可能會很大。

其次, 若一次買很多盒, 譬如說 $T$ 盒, 則其中會包含多少張不同的卡片呢? 以 $Z$ 表此值, 求其期望值 $E(Z)$ 。為了方便, 引進隨機變數 $Y_1, \cdots, Y_N$ , 其中 $Y_i = 1$ , 若第 $i$ 張卡片在這 $T$ 盒裡, 否則 $Y_i = 0$ 。則

$$P(Y_i = 0) = (1 - \frac{1}{N})^T = 1 - P(Y_i = 1)。$$

因此

$$E(Y_i) = 1 - (1 - \frac{1}{N})^T。$$

故

$$E(Z) = E(\sum_{i=1}^N Y_i) = N(1 - (1 - \frac{1}{N})^T)。$$

明顯地上述期望值小於 $N$ 。而之前我們已推導出, 若一次只買一盒, 則當 $N$ 很大時, 欲收集全部 $N$ 張卡片, 所需盒數之期望值約為 $N \log N + N\gamma$ 。由此結果, 引導我們考慮 $T = N \log N$ 。當然此 $T$ 值並非整數, 不過我們只是想了解這時 $N(1 - (1 - 1/N)^T)$ 的大小。

令

$$y = \left(1 - \frac{1}{N}\right)^{N \log N}。$$

則

$$\begin{aligned} \log y &= N \log N \cdot \log\left(1 - \frac{1}{N}\right) \\ &= N \log N \cdot \left(-\frac{1}{N} - \frac{1}{2N^2} - \frac{1}{3N^3} - \cdots\right) \\ &= -\log N \cdot \left(1 + \frac{1}{2N} + \frac{1}{3N^2} + \cdots\right) \\ &< -\log N。 \end{aligned}$$

故得  $y < 1/N$ 。因此若一次買  $[N \log N] + 1$  盒, 其中  $[x]$  表小於或等於  $x$  之最大整數, 則所獲得不同卡片張數之期望值

$$\begin{aligned} &N\left(1 - \left(1 - \frac{1}{N}\right)^{[N \log N] + 1}\right) \\ &> N\left(1 - \left(1 - \frac{1}{N}\right)^{N \log N}\right) > N\left(1 - \frac{1}{N}\right) = N - 1。 \end{aligned}$$

由於最多只能得到  $N$  張不同的卡片, 故得此時期望值已算是很接近  $N$  了。

在2.2.5節我們曾介紹超幾何分佈, 但並未說明期望值及變異數是如何推導出來的。直接計算可能較繁瑣, 利用多變數倒是可求出來。見下例。

**例7.12** 設  $X$  有  $\mathcal{H}(N, D, n)$  分佈, 即  $X$  之 p.d.f. 為

$$(7.22) \quad f(x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}, \max\{0, n - N + D\} \leq x \leq \min\{n, D\}。$$

底下來求  $X$  之期望值及變異數。

仍以2.2.5節取球為例。令  $Y_i = 1$  或  $0$ , 依取出的第  $i$  個球是否為白球而定,  $i = 1, \dots, n$  (這些球的編號並不重要, 因每次取出後不放回, 所以可以一口氣取出  $n$  個, 再任意編號)。則每一  $Y_i$  皆有  $\text{Ber}(D/N)$  分佈, 且  $X = Y_1 + \cdots + Y_n$ 。因此  $E(Y_i) = D/N$ , 且  $\text{Var}(Y_i) = D(N-D)/N^2$ ,  $i = 1, \dots, n$ 。再利用(7.16)式, 便可求出  $X$  之期望值:

$$(7.23) \quad E(X) = \sum_{i=1}^n E(Y_i) = nE(Y_1) = nD/N。$$

$Y_1, \dots, Y_n$ 並不獨立, 我們利用(7.17)式來求 $\text{Var}(X)$ 。首先對 $i \neq j$ ,

$$\begin{aligned} \text{Cov}(Y_i, Y_j) &= E(Y_i Y_j) - E(Y_i)E(Y_j) \\ &= P(Y_i = 1, Y_j = 1) - (D/N)^2 \\ &= \frac{\binom{D}{2}}{\binom{N}{2}} - \left(\frac{D}{N}\right)^2 \\ &= -\frac{D(N-D)}{N^2(N-1)}, \end{aligned}$$

其中用到 $Y_i = 1$ 且 $Y_j = 1$ , 表第 $i$ 及第 $j$ 個球皆為白球, 此二球為 $N$ 個球中的任二球, 共有 $\binom{N}{2}$ 個可能, 若要皆為白球, 表此二球皆要屬於 $D$ 個白球中, 共有 $\binom{D}{2}$ 個可能, 因此 $Y_i = 1$ 且 $Y_j = 1$ 之機率為 $\binom{D}{2}/\binom{N}{2}$ (讀者可否想通為什麼共變異數為負)。由是得

$$\begin{aligned} (7.24) \quad \text{Var}(X) &= n\text{Var}(Y_1) + 2\binom{n}{2}\text{Cov}(Y_1, Y_2) \\ &= \frac{nD(N-1)}{N^2} - \frac{n(n-1)D(N-D)}{N^2(N-1)} \\ &= \frac{nD(N-D)(N-n)}{N^2(N-1)}。 \end{aligned}$$

其次考慮每次取出的球要放回, 且以 $Z$ 表 $n$ 次後共得之白球數。則 $Z$ 有 $\mathcal{B}(n, D/N)$ 分佈,  $E(Z) = nD/N$ 與 $E(X)$ 相等, 但此時 $\text{Var}(Z) = nD(N-D)/N^2$ 。故 $\text{Var}(X)/\text{Var}(Z) = (N-n)/(N-1)$ 。若樣本數 $n$ 與母體數 $N$ 之比值很小, 則二變數之比便很接近1。

若令 $p = D/N$ , 則 $X$ 有 $\mathcal{H}(N, Np, n)$ 分佈,  $Z$ 有 $\mathcal{B}(n, p)$ 分佈。由上推導知, 當 $n, p$ 固定且令 $N \rightarrow \infty$ , 則 $\text{Var}(X) \rightarrow np(1-p) = \text{Var}(Z)$ 。這是可以預期的。因當母體數 $N$ 很大,  $Y_1, \dots, Y_n$ 便近似獨立, 因此 $X$ 可以 $\mathcal{B}(n, p)$ 為其近似分佈。在做各種民意調查時, 選民總數 $N$ 往往很大, 且樣本數 $n$ 與 $N$ 相比很小, 而想估計選民中對某候選人之支持率 $p$ 。此時常近似地以較簡單的二項分佈 $\mathcal{B}(n, p)$ , 當做樣本中對該候選人之

支持人數 $X$ 的分佈，而 $X$ 原本是有 $\mathcal{H}(N, Np, n)$ 分佈的。事實上可以證明 $N \rightarrow \infty$ 時(且 $n, p$ 固定)

$$P(X = k) \rightarrow P(Z = k)。$$

當然若 $n$ 與 $N$ 相比並非很小，就不適合以 $\mathcal{B}(n, p)$ 分佈當做 $\mathcal{H}(N, Np, n)$ 分佈之近似。

對隨機向量 $\mathbf{X} = (X_1, \dots, X_n)$ ，令 $\boldsymbol{\mu}_{\mathbf{X}}$ 及 $\boldsymbol{\Sigma}_{\mathbf{X}}$ 分別表其期望向量(mean vector)及共變異數矩陣(covariance matrix)，也可僅以 $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ 表之，其中

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_n),$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix},$$

而

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

$$\sigma_i^2 = \text{Var}(X_i), i = 1, \dots, n,$$

$$\sigma_{ij} = \sigma_{ji} = \text{Cov}(X_i, X_j), i, j = 1, \dots, n。$$

令

$$\rho_{ij} = \rho(X_i, X_j), i, j = 1, \dots, n。$$

則

$$\sigma_{ij} = \text{Cov}(X_i, X_j) = \rho_{ij} \sigma_i \sigma_j, i, j = 1, \dots, n。$$

故 $\boldsymbol{\Sigma}$ 亦可寫成

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 & \cdots & \rho_{1n} \sigma_1 \sigma_n \\ \rho_{12} \sigma_1 \sigma_2 & \sigma_2^2 & \cdots & \rho_{2n} \sigma_2 \sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} \sigma_1 \sigma_n & \rho_{2n} \sigma_2 \sigma_n & \cdots & \sigma_n^2 \end{pmatrix}。$$

對一列向量  $\mathbf{a} = (a_1, \dots, a_n)$ ,

$$\mathbf{a}' = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

$\mathbf{a}'$  即為  $\mathbf{a}$  之轉置 (transpose), 為一行向量。則

$$\begin{aligned} & (\mathbf{X} - \boldsymbol{\mu})'(\mathbf{X} - \boldsymbol{\mu}) \\ &= \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_n - \mu_n \end{pmatrix} (X_1 - \mu_1, X_2 - \mu_2, \dots, X_n - \mu_n) \\ &= \begin{pmatrix} (X_1 - \mu_1)^2 & \cdot & \cdots & (X_1 - \mu_1)(X_n - \mu_n) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \cdots & (X_2 - \mu_2)(X_n - \mu_n) \\ \vdots & \vdots & \ddots & \vdots \\ (X_n - \mu_n)(X_1 - \mu_1) & \cdot & \cdots & (X_n - \mu_n)^2 \end{pmatrix}. \end{aligned}$$

上述  $\boldsymbol{\mu}$  及  $\boldsymbol{\Sigma}$  便可分別寫成

$$\boldsymbol{\mu} = E(\mathbf{X}),$$

$$\boldsymbol{\Sigma} = E((\mathbf{X} - \boldsymbol{\mu})'(\mathbf{X} - \boldsymbol{\mu})).$$

**例7.13** 設  $(X_1, X_2) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 。則

$$\begin{aligned} \boldsymbol{\mu} &= (\mu_1, \mu_2), \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \end{aligned}$$

**例7.14** 設  $(X_1, X_2, X_3)$  之共變異數矩陣為

$$\boldsymbol{\Sigma} = \begin{pmatrix} 2 & -0.5 & 0.2 \\ -0.5 & 6 & -0.4 \\ 0.2 & -0.4 & 1 \end{pmatrix}.$$

試求  $\text{Cov}(2X_1 + X_3, X_1 - X_2 - X_3)$ 。

解. 由定理7.5,

$$\begin{aligned} & \text{Cov}(2X_1 + X_3, X_1 - X_2 - X_3) \\ &= 2\text{Cov}(X_1, X_1) - 2\text{Cov}(X_1, X_2) - 2\text{Cov}(X_1, X_3) + \text{Cov}(X_3, X_1) \\ & \quad - \text{Cov}(X_3, X_2) - \text{Cov}(X_3, X_3) \\ &= 2 \cdot 2 - 2 \cdot (-0.5) - 0.2 - (-0.4) - 1 \\ &= 4 + 1 - 0.2 + 0.4 - 1 = 4.2。 \end{aligned}$$