

第三章

多維隨機變數

3.2 條件分佈及獨立性

當觀測兩個隨機變數 X 與 Y 時，兩個變數有時並非不相干。例如，令 X 表某位學生第一次考試的成績， Y 表第二次考試的成績。如果知道 $X = 100$ ，則 $Y = 0$ 的機率應是很低的。又如身高(X)與體重(Y)，如果知道 $X = 180$ (公分)，則 $Y \geq 60$ (公斤)的機率，應比 $X = 140$ (公分)時， $Y \geq 60$ (公斤)的機率高。上述這些例子顯示，有時知道 X 的值，是會提供一些關於 Y 的資訊。

設 (X, Y) 為離散型的隨機向量，因 $\{X = x\}$ 及 $\{Y = y\}$ 為二事件，故可以第一章(定義2.1)事件之條件機率的定義，來定義條件機率 $P(Y = y|X = x)$ 。即只要 $P(X = x) > 0$ ，則

$$(2.1) \quad P(Y = y|X = x) = \frac{P(Y = y, X = x)}{P(X = x)}。$$

上式也可以p.d.f.表示，我們給一定義如下。

定義2.1 設 (X, Y) 為離散型的隨機向量，以 $f(x, y)$ 為聯合p.d.f., $f_X(x)$, $f_Y(y)$ 為邊際p.d.f.。對任意 x , 只要 $f_X(x) > 0$, 在給定 $X = x$ 之下, Y 之條件p.d.f. $f(y|x)$ 為

$$(2.2) \quad f(y|x) = P(Y = y|X = x) = \frac{f(x, y)}{f_X(x)}。$$

對任意 y , 只要 $f_Y(y) > 0$, 在給定 $Y = y$ 之下, X 之條件p.d.f. $f(x|y)$ 為

$$(2.3) \quad f(x|y) = P(X = x|Y = y) = \frac{f(x, y)}{f_Y(y)}。$$

有時為了清楚隨機變數為何, 且避免混淆, 我們會以 $f_{Y|X}(y|x)$ 表在給定 $X = x$ 之下, Y 的條件p.d.f.。也常簡單地以 $Y|X$ 之分佈, 或 $Y|X$ 之p.d.f., 分別表條件分佈及條件p.d.f.。另外, 既然 $f(y|x)$ 被稱為一條件“p.d.f.”, 它必須滿足p.d.f.的要求。這當然沒問題。首先, 因 $f_X(x) > 0, f(x, y) \geq 0$, 故 $f(y|x) \geq 0$ 。其次

$$\sum_y f(y|x) = \sum_y \frac{f(x, y)}{f_X(x)} = \frac{\sum_y f(x, y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1。$$

故 $f(y|x)$ 的確是一p.d.f.。

例2.1 設 (X, Y) 之聯合p.d.f.如下:

$y \setminus x$	1	2	$f_Y(y)$
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$
2	$\frac{9}{16}$	$\frac{1}{16}$	$\frac{5}{8}$
$f_X(x)$	$\frac{11}{16}$	$\frac{5}{16}$	1

我們也求出邊際p.d.f. $f_X(x)$ 及 $f_Y(y)$, 一併填入上表。則

$$\begin{aligned} f_{Y|X}(1|1) &= \frac{f(1, 1)}{f_X(1)} = \frac{1/8}{11/16} = \frac{2}{11}, \\ f_{Y|X}(2|1) &= \frac{f(1, 2)}{f_X(1)} = \frac{9/16}{11/16} = \frac{9}{11}, \\ f_{Y|X}(1|2) &= \frac{f(2, 1)}{f_X(2)} = \frac{1/4}{5/16} = \frac{4}{5}, \\ f_{Y|X}(2|2) &= \frac{f(2, 2)}{f_X(2)} = \frac{1/16}{5/16} = \frac{1}{5}。 \end{aligned}$$

至於 $f_{X|Y}(x|y)$ 留給各位自行求出。可看出

$$\begin{aligned}\sum_y f_{Y|X}(y|1) &= f_{Y|X}(1|1) + f_{Y|X}(2|1) = \frac{2}{11} + \frac{9}{11} = 1, \\ \sum_y f_{Y|X}(y|2) &= f_{Y|X}(1|2) + f_{Y|X}(2|2) = \frac{4}{5} + \frac{1}{5} = 1.\end{aligned}$$

但

$$\begin{aligned}\sum_x f_{Y|X}(1|x) &= f_{Y|X}(1|1) + f_{Y|X}(1|2) = \frac{2}{11} + \frac{4}{5} \neq 1, \\ \sum_x f_{Y|X}(2|x) &= f_{Y|X}(2|1) + f_{Y|X}(2|2) = \frac{9}{11} + \frac{1}{5} \neq 1.\end{aligned}$$

上例顯示，雖 $\sum_y f_{Y|X}(y|x) = 1$ ，但 $\sum_x f_{Y|X}(y|x)$ 卻不一定等於1，其中原因應不難想通。

利用條件p.d.f.可求出各種條件機率。例如，

$$(2.4) \quad P(Y \leq y|X = x) = \sum_{v \leq y} f_{Y|X}(v|x).$$

若 (X, Y) 為連續型的隨機向量，則 $P(X = x) = 0, \forall x \in R$ ，因此無法利用事件的機率來定義給定 $X = x$ 之下， Y 的條件機率，不過經由機率理論的推導，參照離散型的隨機變數的情況，我們給出下述定義。

定義2.2 設 (X, Y) 為連續型的隨機向量，以 $f(x, y)$ 為聯合p.d.f.， $f_X(x)$ ， $f_Y(y)$ 為邊際p.d.f.。對任意 x ，只要 $f_X(x) > 0$ ，在給定 $X = x$ 之下， Y 之條件p.d.f.為

$$(2.5) \quad f(y|x) = \frac{f(x, y)}{f_X(x)}.$$

對任意 y ，只要 $f_Y(y) > 0$ ，在給定 $Y = y$ 之下， X 之條件p.d.f.為

$$(2.6) \quad f(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

如同離散型的情況，亦可證明 $f(y|x)$ 及 $f(x|y)$ 皆確為p.d.f.。

例2.2 承例1.3。可求出

$$\begin{aligned} f(y|x) &= \frac{f(x, y)}{f_X(x)} = \frac{24xy}{12x(1-x)^2} \\ &= \frac{2y}{(1-x)^2}, 0 \leq y \leq 1-x, 0 < x < 1。 \end{aligned}$$

同理

$$f(x|y) = \frac{2x}{(1-y)^2}, 0 \leq x \leq 1-y, 0 < y < 1。$$

注意因 $f_X(0) = f_Y(0) = 0$, 且 $f_X(1) = f_Y(1) = 0$, 故在 $f(y|x)$ 中, $x \neq 0, 1$, 在 $f(x|y)$ 中, $y \neq 0, 1$ 。又可驗證 $f(y|x)$ 及 $f(x|y)$ 都確為 p.d.f.。即

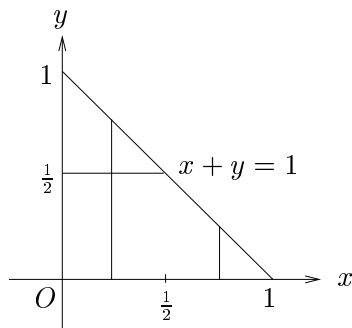
$$\int_0^{1-x} \frac{2y}{(1-x)^2} dy = \int_0^{1-y} \frac{2x}{(1-y)^2} dx = 1。$$

又 $P(Y \leq 1/2|x)$ 是否等於

$$\int_0^{1/2} f(y|x) dy = \int_0^{1/2} \frac{2y}{(1-x)^2} dy = \frac{1}{4(1-x)^2}?$$

要注意的是 y 的範圍要滿足 $0 \leq y \leq 1-x$, 而 $1/2$ 不一定小於 $1-x$ 。正確的答案為(參考附圖)

$$\begin{aligned} P(Y \leq \frac{1}{2}|x) &= \int_0^{\min(1/2, 1-x)} f(y|x) dy \\ &= \int_0^{\min(1/2, 1-x)} \frac{2y}{(1-x)^2} dy \\ &= \frac{(\min(1/2, 1-x))^2}{(1-x)^2} \\ &= \begin{cases} \frac{1}{4(1-x)^2}, & 0 < x \leq \frac{1}{2}, \\ \frac{(1-x)^2}{(1-x)^2} = 1, & \frac{1}{2} < x < 1。 \end{cases} \end{aligned}$$



條件 p.d.f. 亦可用來求 **條件期望值**。對一函數 g , 在給定 $X = x$ 之下, $g(Y)$ 之條件期望值, 以 $E(g(Y)|x)$ 表之, 定義為

$$(2.7) \quad E(g(Y)|x) = \begin{cases} \sum_y g(y) f(y|x) & , \text{離散型,} \\ \int_{-\infty}^{\infty} g(y) f(y|x) dy & , \text{連續型。} \end{cases}$$

條件期望值仍具有第一章定理6.1所列出的一般的期望值之特性。譬如說，由於 $E(Y|x)$ 為一 x 的函數，故當給定 $X = x$ ， $E(Y|x)$ 便非隨機了。因此(利用第一章(6.5)及(6.6)式)

$$(2.8) \quad E(E(Y|x)|x) = E(Y|x),$$

$$(2.9) \quad E(YE(Y|x)|x) = E(Y|x) \cdot E(Y|x) = (E(Y|x))^2。$$

又在誤差的平方之期望值最小下， $E(Y|X)$ 為在知道 X 之下， Y 之最佳預測值。這是第一章定理6.2之推廣，我們留在習題中讓各位自行完成。

例2.3 承例2.1。則

$$\begin{aligned} E(Y|X=1) &= \sum_y yf(y|1) = 1 \cdot \frac{2}{11} + 2 \cdot \frac{9}{11} = \frac{20}{11}, \\ E(Y|X=2) &= \sum_y yf(y|2) = 1 \cdot \frac{4}{5} + 2 \cdot \frac{1}{5} = \frac{6}{5}。 \end{aligned}$$

例2.4 承例2.2。

$$\begin{aligned} E(Y|X=x) &= \int_0^{1-x} yf(y|x)dy \\ &= \int_0^{1-x} \frac{2y^2}{(1-x)^2} dy \\ &= \frac{2(1-x)^3}{3(1-x)^2} \\ &= \frac{2}{3}(1-x), 0 < x < 1。 \end{aligned}$$

又

$$\begin{aligned} E(Y^2|X=x) &= \int_0^{1-x} y^2 \frac{2y}{(1-x)^2} dy \\ &= \frac{1}{(1-x)^2} \cdot \frac{1}{2}(1-x)^4 = \frac{1}{2}(1-x)^2, 0 < x < 1。 \end{aligned}$$

因此

$$\begin{aligned}
& E(2Y^2 + 3Y|X = x) \\
&= 2E(Y^2|X = x) + 3E(Y|X = x) \\
&= (1-x)^2 + 2(1-x) = (1-x)(3-x), 0 < x < 1。
\end{aligned}$$

在給定 $X = x$ 之下，除了可求條件期望值 $E(Y|X = x)$ ，亦可求**條件變異數** $\text{Var}(Y|X = x)$ ，或以 $\text{Var}(Y|x)$ 表之，其定義為

$$(2.10) \quad \text{Var}(Y|x) = E((Y - E(Y|x))^2|x)。$$

以上式來定義條件變異數是有道理的。變異數是量測一變數(此處為 Y)與其期望值(此處為 $E(Y|x)$)之離差平方(此處為 $(Y - E(Y|x))^2$)的期望值(此處為給定 $X = x$ 之下的期望值)，故得(2.10)式。由(2.10)式立即導致(證明留在習題第3題)

$$(2.11) \quad \text{Var}(Y|x) = E(Y^2|x) - (E(Y|x))^2。$$

例2.5 承例1.3, 2.2, 及2.4。可得

$$\begin{aligned}
\text{Var}(Y|x) &= E(Y^2|x) - (E(Y|x))^2 \\
&= \frac{1}{2}(1-x)^2 - \frac{4}{9}(1-x)^2 \\
&= \frac{1}{18}(1-x)^2, 0 < x < 1。
\end{aligned}$$

變異數與 x 有關。當 $x \rightarrow 1$ 時，變異數 $\rightarrow 0$ ，這是合理的，由例1.3中的圖形可看出，當 $x \rightarrow 1$ 時， $y \rightarrow 0$ ，變異愈來愈小，變異數自然也愈來愈小。甚至因 x 愈大時， y 的活動範圍愈小($0 \leq y \leq 1-x$)，變異也就愈小。事實上，我們所求出的 $\text{Var}(Y|x)$ 在 $0 < x < 1$ 間，的確為 x 的漸減函數。

例2.6 設 X, Y 之聯合p.d.f.為

$$f(x, y) = \frac{1}{64}(x + y), 0 \leq x, y \leq 4。$$

試求(i) $f_X(x)$, (ii) $f(y|x)$, (iii) $E(Y|x)$, 及(iv) $\text{Var}(Y|x)$, $0 \leq x \leq 4$ 。

解.(i) 依定義，即得

$$f_X(x) = \int_0^4 \frac{1}{64}(x+y)dy = \frac{1}{16}x + \frac{1}{8}, 0 \leq x \leq 4。$$

(ii) 利用(i)得

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{(x+y)/64}{x/16 + 1/8} = \frac{x+y}{4x+8}, 0 \leq y \leq 4, 0 \leq x \leq 4。$$

(iii) 利用(ii)得

$$\begin{aligned} E(Y|x) &= \int_0^4 y \cdot f(y|x)dy \\ &= \int_0^4 y \cdot \frac{x+y}{4x+8}dy \\ &= \frac{1}{4x+8} \int_0^4 (xy + y^2)dy \\ &= \frac{8x + 64/3}{4x+8} = \frac{6x+16}{3x+6}, 0 \leq x \leq 4。 \end{aligned}$$

(iv) 先求出

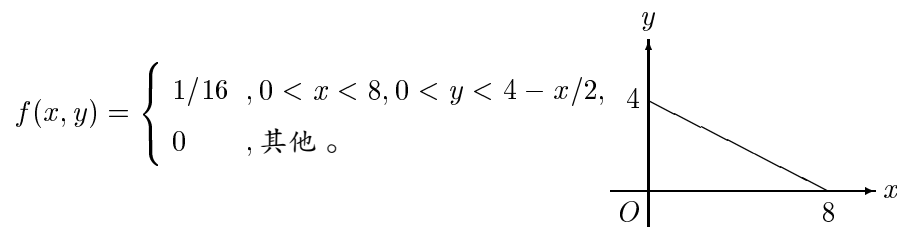
$$\begin{aligned} E(Y^2|x) &= \int_0^4 y^2 \cdot f(y|x)dy \\ &= \int_0^4 y^2 \cdot \frac{x+y}{4x+8}dy \\ &= \frac{1}{4x+8} \int_0^4 (xy^2 + y^3)dy \\ &= \frac{64x/3 + 64}{4x+8} = \frac{16x+48}{3x+6}, 0 \leq x \leq 4, \end{aligned}$$

因此

$$\begin{aligned} \text{Var}(Y|x) &= E(Y^2|x) - (E(Y|x))^2 \\ &= \frac{16x+48}{3x+6} - \left(\frac{6x+16}{3x+6}\right)^2 \\ &= \frac{12x^2 + 48x + 32}{(3x+6)^2}, 0 \leq x \leq 4。 \end{aligned}$$

例2.7 設 (X, Y) 在 xy 平面上, 以 $(0, 0)$, $(8, 0)$ 及 $(0, 4)$ 為三頂點之三角形內(不含三邊)均勻分佈。試求 $f(y|x)$, $E(Y|x)$, 及 $\text{Var}(Y|x)$ 。

解. 首先 (X, Y) 之聯合p.d.f.為



由此得

$$f_X(x) = \int_0^{4-x/2} \frac{1}{16} dy = \frac{4-x/2}{16} = \frac{8-x}{32}, 0 < x < 8。$$

則

$$f(y|x) = \frac{1/16}{(8-x)/32} = \frac{2}{8-x}, 0 < y < 4 - x/2, 0 < x < 8。$$

可看出給定 $X = x$, Y 在區間 $(0, 4 - x/2)$ 均勻分佈。最後, 利用對 $\forall a > 0$, 有 $\mathcal{U}(0, a)$ 分佈之隨機變數的期望值為 $a/2$, 變異數為 $a^2/12$, 即得

$$E(Y|x) = \frac{1}{2}\left(4 - \frac{x}{2}\right) = 2 - \frac{x}{4}, 0 < x < 8,$$

$$\text{Var}(Y|x) = \frac{1}{12}\left(4 - \frac{x}{2}\right)^2 = \frac{(8-x)^2}{48}, 0 < x < 8。$$

讀者亦可自行求出 $X|Y = y$ 有 $\mathcal{U}(0, 8 - 2y)$ 分佈, 且 $E(X|y) = 4 - y$, $\text{Var}(X|y) = (4 - y)^2/3, 0 < y < 4。$

註2.1 設 A 為平面上一區域, 且設 A 的面積(以 $S(A)$ 表之)存在且為正。則 (X, Y) 在 A 上均勻分佈, 表 (X, Y) 之聯合 p.d.f. 為

$$f(x, y) = \frac{1}{S(A)}, (x, y) \in A。$$

例2.8 設 (X, Y) 之聯合 p.d.f. 為

$$f(x, y) = \frac{1}{2}x^3e^{-x(1+y)}, x, y > 0。$$

試求 $f(y|x)$, $E(Y|x)$ 及 $\text{Var}(Y|x)$ 。

解. 首先

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

因此

$$f(y|x) = \frac{x^3 e^{-x(1+y)}/2}{x^2 e^{-x}/2} = x e^{-xy}, y > 0, x > 0。$$

即 $Y|X = x$ 有 $\mathcal{E}(x)$ 分佈。因此

$$E(Y|x) = \frac{1}{x}, \quad \text{Var}(Y|x) = \frac{1}{x^2}, x > 0。$$

讀者亦可自行求出 $f_Y(y) = 3(1+y)^{-4}$, $y > 0$ 。由此可得對 $\forall y > 0$, $X|Y = y$ 有 $\Gamma(4, (1+y)^{-1})$ 分佈, 且 $E(X|y) = 4/(1+y)$, $\text{Var}(X|y) = 4/(1+y)^2$, $y > 0$ 。

例2.9 設有一函數

$$f(x, y) = \begin{cases} 4xy - 2x - 2y + 2, & 0 < x, y < 1, \\ 0, & \text{其他。} \end{cases}$$

(i) 試證 f 爲一二維之 p.d.f.;

(ii) 假設 (X, Y) 以 f 爲聯合 p.d.f., 試求 $f(y|x)$, $E(Y|x)$ 及 $\text{Var}(Y|x)$ 。

解. (i) 對 $0 < x, y < 1$,

$$4xy - 2x - 2y + 2 = 2xy + 2(1-x)(1-y) > 0,$$

即 f 爲非負。又

$$\begin{aligned} \int_0^1 \int_0^1 f(x, y) dy dx &= \int_0^1 \int_0^1 (4xy - 2x - 2y + 2) dy dx \\ &= 1 - 1 - 1 + 2 = 1, \end{aligned}$$

故 f 爲一二維之 p.d.f.。

(ii) 首先

$$\begin{aligned} f_X(x) &= \int_0^1 (4xy - 2x - 2y + 2) dy \\ &= 2x - 2x - 1 + 2 = 1, 0 < x < 1。 \end{aligned}$$

因此

$$f(y|x) = 4xy - 2x - 2y + 2, 0 < y < 1, 0 < x < 1。$$

利用上述條件p.d.f., 即得

$$\begin{aligned} E(Y|x) &= \int_0^1 y(4xy - 2x - 2y + 2)dy = \frac{1}{3}(1+x), 0 < x < 1, \\ \text{Var}(Y|x) &= \int_0^1 y^2(4xy - 2x - 2y + 2)dy - \frac{1}{9}(1+x)^2 \\ &= \frac{1}{6}(1+2x) - \frac{1}{9}(1+x)^2 \\ &= \frac{1}{18}(1+2x-2x^2), 0 < x < 1. \end{aligned}$$

讀者可否看出 X, Y 之邊際分佈皆為 $\mathcal{U}(0, 1)$, 但 (X, Y) 並無二維的均勻分佈。又可否立即看出, 對本例, $f(x|y)$, $E(X|y)$ 及 $\text{Var}(X|y)$ 分別為何?

由例2.1及2.2知, 給定 $X = x$ 之下, Y 的分佈可能會隨 x 而變。所以在求條件分佈時, 我們得到一分佈族: 對每一 x 有一分佈。若我們想描述整個分佈族, 可以“ $Y|X$ ”(Y給定 X , 或說給定 X 之下 Y 之分佈)來表示。

例如, 若 X 為一取正整數值之隨機變數, 且在給定 $X = x$ 之下, Y 有 $\mathcal{B}(x, p)$ 分佈, 則可以說 $Y|X$ 有 $\mathcal{B}(X, p)$ 分佈。採用符號“ $Y|X$ ”, 或是分佈中有一參數為隨機變數, 則我們便描述出一條件分佈族。同樣地, $E(g(Y)|x)$ 亦為一 x 的函數, 而 $E(g(Y)|X)$ 則為一隨機變數, 其值依 X 之值而定。若 $X = x$, 則隨機變數 $E(g(Y)|X)$ 之值便為 $E(g(Y)|x)$ 。例如, 在例2.4中, 我們可以寫

$$E(Y|x) = \frac{2}{3}(1-x), \quad E(Y^2|x) = \frac{1}{2}(1-x)^2.$$

在前述幾個例子中, 給定 $X = x$, Y 之分佈均隨 x 而變。但有些時候, 給定 $X = x$, Y 之分佈並未改變, 亦即知道 X 之值, 對 Y 並未提供任何資訊。此時我們說 X 與 Y 獨立。

定義2.3 設二隨機變數 X, Y 之聯合p.d.f.為 $f(x, y)$, 邊際p.d.f.分別為 $f_X(x)$, $f_Y(y)$ 。若

$$(2.12) \quad f(x, y) = f_X(x)f_Y(y), \forall x, y \in R,$$

便稱 X 與 Y 獨立。

由定義2.1及2.2知, 對離散型或連續型的隨機向量 (X, Y) ,

$$(2.13) \quad f(y|x) = \frac{f(x, y)}{f_X(x)}, \text{ 只要 } f_X(x) > 0。$$

因此

$$(2.14) \quad f(x, y) = f(y|x)f_X(x), \text{ 只要 } f_X(x) > 0。$$

若 X 與 Y 獨立, 則(2.12)式成立, 因此

$$(2.15) \quad f(y|x) = f_Y(y), \text{ 只要 } f_X(x) > 0。$$

由上式又得對 $\forall A \subset R$, 只要 $f_X(x) > 0$, 則

$$P(Y \in A|x) = \int_A f(y|x)dy = \int_A f_Y(y)dy = P(Y \in A),$$

即知道 $X = x$ 之值, 對求 Y 的機率並無影響。

例2.10 在例2.1中, 因

$$f_X(1)f_Y(1) = \frac{11}{16} \cdot \frac{3}{8} \neq \frac{1}{8} = f(1, 1),$$

故 X 與 Y 不獨立。當然由 $f(y|x)$ 之值與 x 有關亦可判定 X 與 Y 不獨立。

例2.11 設 (X, Y) 有如例2.8之聯合p.d.f., 即

$$f(x, y) = \frac{1}{2}x^3e^{-x(1+y)}, x, y > 0。$$

試判斷 X 與 Y 是否獨立。

解. 首先由例2.8, X, Y 之邊際p.d.f.分別為

$$f_X(x) = \frac{1}{2}x^2e^{-x}, x > 0,$$

$$f_Y(y) = \frac{3}{(1+y)^4}, y > 0。$$

顯然

$$f(1, 1) = \frac{1}{2}e^{-2} \neq \frac{1}{2}e^{-1} \cdot \frac{3}{(1+1)^4} = f_X(1) \cdot f_Y(1)。$$

故 X 與 Y 不獨立。

要利用(2.12)式來決定 X 與 Y 是否獨立, 便須知 X 與 Y 之邊際p.d.f. $f_X(x)$ 及 $f_Y(y)$ 。利用下述定理, 有時可讓檢驗工作容易些。

定理2.1 設隨機向量 (X, Y) 以 $f(x, y)$ 為聯合p.d.f.。則 X 與 Y 獨立, 若且唯若存在函數 $g(x)$ 及 $h(y)$, 使得對 $\forall x, y \in R$,

$$(2.16) \quad f(x, y) = g(x)h(y)。$$

證明.若 X 與 Y 獨立, 依定義(2.12)式便成立。取 $g(x) = f_X(x)$, $h(y) = f_Y(y)$, 則(2.16)式成立。這便證明了“唯若”的部分。對“若”的部分, 我們先證連續型隨機變數的情況。

假設存在 $g(x)$ 及 $h(y)$, 使得(2.16)式成立。令

$$a = \int_{-\infty}^{\infty} g(x)dx, \quad b = \int_{-\infty}^{\infty} h(y)dy,$$

則

$$(2.17) \quad \begin{aligned} ab &= \int_{-\infty}^{\infty} g(x)dx \cdot \int_{-\infty}^{\infty} h(y)dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dx dy \\ &= 1, \end{aligned}$$

最後一等式成立, 是因 $f(x, y)$ 為一聯合p.d.f.。

另外, (2.16)式導致

$$(2.18) \quad f_X(x) = \int_{-\infty}^{\infty} g(x)h(y)dy = g(x)b,$$

$$(2.19) \quad f_Y(y) = \int_{-\infty}^{\infty} g(x)h(y)dx = h(y)a。$$

利用(2.17)、(2.18)及(2.19)式得

$$f(x, y) = g(x)h(y) = g(x)h(y)ab = f_X(x)f_Y(y)。$$

得證 X 與 Y 獨立。

至於離散型的情況, 只要將上述積分皆改為和即可。

下述系理為定理2.1之一立即的推論。

系理2.1 設隨機向量 (X, Y) 以 $F(x, y)$ 為其聯合分佈函數, 則 X 與 Y 獨立, 若且唯若

$$(2.20) \quad F(x, y) = F_X(x)F_Y(y), x, y \in R,$$

若且唯若存在函數 $G(x)$ 及 $H(y)$, 使得對 $\forall x, y \in R$,

$$(2.21) \quad F(x, y) = G(x)H(y)。$$

利用定理2.1, 立即可判定例2.11中的 X 與 Y 不獨立。底下再給一例。

例2.12 設 (X, Y) 之聯合p.d.f. 為

$$f(x, y) = \frac{1}{121,500} x^3 y^2 e^{-x/3-y/5}, x, y > 0。$$

則由定理2.1知, X 與 Y 獨立。又讀者宜練習由 $f(x, y)$ 看出 X, Y 分別有 $\Gamma(4, 3)$ 及 $\Gamma(3, 5)$ 分佈。

在利用獨立的定義, 或定理2.1, 以檢驗二隨機變數是否獨立, 要留意的是使 $f(x, y) > 0$ 之 (x, y) 的集合, 必須是 $\{(x, y) | x \in A, y \in B\}$ 的型式, 其中 $A = \{x | f_X(x) > 0\}, B = \{y | f_Y(y) > 0\}$, 即為 R^2 上一有“乘積”型式的區域, 通常以 $A \times B$ 表之。換句話說, x, y 的範圍要彼此無關。在例1.3中, 雖然 $f(x, y) = 24xy$, 表面上看起來, 可寫成 $g(x)h(y)$ 的型式, 但使 $f(x, y) > 0$ 之 x, y 的集合為 $\{0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1\}$, 為平面上—三角形, 並無 $A \times B$ 的型式。範圍如果是 $x^2 + y^2 \leq 4$ 當然也不行。更明確地說, 例1.3中之 $f(x, y)$ 應寫成

$$f(x, y) = 24xy I_{\{0 \leq x \leq 1, 0 \leq y \leq 1, x+y \leq 1\}}。$$

因其中指示函數的部分, 無法寫成一個 x 的函數乘上一個 y 的函數, 所以 X 與 Y 不獨立。

當 X 與 Y 獨立時, 有關 X 與 Y 的機率或期望值都會較好求。我們給一定理。

定理2.2 設 X, Y 為二獨立的隨機變數。則

(i) 對 $\forall A, B \subset R$,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B),$$

即事件 $\{X \in A\}$ 與 $\{Y \in B\}$ 相互獨立。

(ii) 設有函數 $g(x)$ 及 $h(y)$, 滿足 $E(g(X))$ 及 $E(h(Y))$ 皆存在。則

$$E(g(X)h(Y)) = E(g(X)) \cdot E(h(Y))。$$

證明.我們仍只證連續型隨機變數的情況。先證(ii)。由於 X 與 Y 獨立, $f(x, y) = f_X(x)f_Y(y)$, 故

$$\begin{aligned} E(g(X)h(Y)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f(x, y)dxdy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} h(y)f_Y(y)\left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right)dy \\ &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \int_{-\infty}^{\infty} h(y)f_Y(y)dy \\ &= E(g(X)) \cdot E(h(Y))。 \end{aligned}$$

得證(ii)。

其次對於(i)的部分, 可仿(ii)的證明, 或利用(ii)來證明。事實上(ii)是比(i)更一般的結果。

取

$$g(x) = I_A(x), \quad h(y) = I_B(y),$$

分別表 A, B 之指示函數。則 $g(x)h(y)$ 為 R^2 上之集合 C 之指示函數, 其中 $C = \{(x, y) | x \in A, y \in B\}$ 。又

$$E(g(X)) = P(X \in A), \quad E(h(Y)) = P(Y \in B),$$

$$E(g(X)h(Y)) = P((X, Y) \in C)。$$

故得

$$\begin{aligned}
 P(X \in A, Y \in B) &= P((X, Y) \in C) \\
 &= E(g(X)h(Y)) \\
 &= E(g(X))E(h(Y)) \\
 &= P(X \in A)P(Y \in B) \text{。}
 \end{aligned}$$

證畢。

例2.13 設 X, Y 為二獨立的隨機變數, X 有 $\mathcal{E}(2)$ 分佈, Y 有 $\mathcal{U}[0, 3]$ 分佈。則

$$\begin{aligned}
 P(X \geq 3, Y \geq 2) &= P(X \geq 3)P(Y \geq 2) \\
 &= \int_3^\infty 2e^{-2x} dx \cdot \int_2^3 \frac{1}{3} dy \\
 &= \frac{1}{3}e^{-6} \text{。}
 \end{aligned}$$

又

$$\begin{aligned}
 E(XY^2) &= E(X) \cdot E(Y^2) = \frac{1}{2} \cdot \int_0^3 \frac{1}{3} y^2 dy \\
 &= \frac{1}{2} \cdot 3 = \frac{3}{2} \text{。}
 \end{aligned}$$

例2.14 設 U_1, U_2 為二獨立的隨機變數, 且皆有 $\mathcal{U}(0, 1]$ 分佈。又設平面上有一隨機的矩形 A , 四個頂點之座標分別為 $(0, 0)$, $(U_1, 0)$, (U_1, U_2) 及 $(0, U_2)$ 。

則 A 之面積的期望值與中位數何者為大?

解. 首先面積 U_1U_2 之期望值為

$$E(U_1U_2) = E(U_1)E(U_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \text{。}$$

次求面積 U_1U_2 之中位數。對 $\forall 0 < y \leq 1$,

$$\begin{aligned}
 P(U_1U_2 \geq y) &= \int_y^1 \left(\int_{y/u_1}^1 du_2 \right) du_1 \\
 &= \int_y^1 \left(1 - \frac{y}{u_1} \right) du_1
 \end{aligned}$$

$$\begin{aligned}
&= (u_1 - y \log u_1) \Big|_y^1 \\
&= 1 - y(1 - \log y) \text{。}
\end{aligned}$$

而若 $y > 1$, 則 $P(U_1 U_2 \geq y) = 0$, 若 $y \leq 0$, 則 $P(U_1 U_2 \geq y) = 1$ 。由此即得

$$P(U_1 U_2 \geq \frac{1}{4}) = 1 - \frac{1}{4}(1 + \log 4) \doteq 0.403 \text{。}$$

故 $U_1 U_2$ 之中位數小於 $1/4$ 。即 $U_1 U_2$ 之中位數比期望值小。

例2.15 自區間 $(0, 1)$ 隨機地取兩數, 以 X 及 Y 表之。求最接近 X/Y 的整數恰為一偶數之機率。

解. 具有對稱性概念的人, 看到又是均勻分佈, 又是奇數、偶數, 會以為此機率應為 $1/2$ 。底下我們來看對不對。首先由題意知, X, Y 為二獨立的隨機變數, 且皆有 $\mathcal{U}(0, 1)$ 分佈。故 (X, Y) 之聯合 p.d.f. 為

$$f(x, y) = 1, 0 < x, y < 1 \text{。}$$

對任意 X 及 Y , 最接近 X/Y 的整數恰為偶數, 若且唯若 $X/Y < 0.5$, 或存在一正整數 n , 使得 $2n - 0.5 < X/Y < 2n + 0.5$ 。現因

$$P(X/Y < 0.5) = P(2X < Y) = \int_0^{0.5} \int_{2x}^1 dy dx = \frac{1}{4},$$

又

$$\begin{aligned}
&P(2n - 0.5 < X/Y < 2n + 0.5) \\
&= P\left(\frac{2X}{4n+1} < Y < \frac{2X}{4n-1}\right) \\
&= \int_0^1 \left(\frac{2x}{4n-1} - \frac{2x}{4n+1}\right) dx \\
&= \frac{1}{4n-1} - \frac{1}{4n+1},
\end{aligned}$$

故

$$P(\text{最接近 } X/Y \text{ 的整數恰為一偶數})$$

$$\begin{aligned}
&= \frac{1}{4} + \sum_{n=1}^{\infty} \left(\frac{1}{4n-1} - \frac{1}{4n+1} \right) \\
&= \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{5} + \frac{1}{7} - \frac{1}{9} + \cdots \right) \\
&= \frac{1}{4} + 1 - \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots \right) \\
&= \frac{1}{4} + 1 - \arctan 1 \\
&= \frac{1}{4} + 1 - \frac{\pi}{4} = \frac{5-\pi}{4} < \frac{1}{2}。
\end{aligned}$$

事實上我們所求即為 $X/Y \in S$ 之機率, 其中

$$S = (0, 0.5) \cup (1.5, 2] \cup (2, 2.5) \cup (3.5, 4] \cup (4, 4.5) \cup \cdots,$$

包含每一整數區間 $(i, i+1], i \geq 0$, 的一半部分。但此機率卻小於 $1/2$, 即使 X, Y 原來均有獨立的均勻分佈。

註2.2 此處用到

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots + \frac{(-1)^n}{2n+1}x^{2n+1} + \cdots, |x| \leq 1。$$

底下檢驗二隨機變數獨立的定理, 亦為定理2.1之一推論。我們略去其證明。

定理2.3 設有二隨機變數 X, Y 。則 X 與 Y 獨立, 若且唯若

$$(2.22) \quad \psi(s, t) = E(e^{isX+itY}) = E(e^{isX})E(e^{itY}), s, t \in R。$$

上式左側稱為 X, Y 之聯合特徵函數。定理2.3即是說 X 與 Y 獨立, 若且唯若其聯合特徵函數等於 X, Y 各自特徵函數之積。定理2.3也有母函數(針對取非負整數值之隨機變數), 拉普拉斯轉換(針對取非負整數值之隨機變數)及動差母函數之版本。只要這些轉換至少在一二維的開區間存在(母函數及拉普拉斯轉換沒問題, 動差母函數 $M(s, t) = E(e^{sX+tY})$ 有時只有在 $s = t = 0$ 才存在)。

定理2.3有下述一重要的推論。

系理2.2 設 X, Y 為二獨立的隨機變數。則

$$(2.23) \quad E(e^{it(X+Y)}) = E(e^{itX})E(e^{itY}), t \in R。$$

同樣地,系理2.2亦有關於其他三個轉換的版本。

例2.16 設 X, Y 為二獨立的隨機變數, 各有 $\mathcal{P}(\lambda)$ 及 $\mathcal{P}(\mu)$ 的分佈。則由系理2.2, $X + Y$ 之拉普拉斯轉換為

$$\begin{aligned} \phi(s, t) &= E(e^{-s(X+Y)}) = E(e^{-sX})E(e^{-sY}) \\ &= e^{-\lambda(1-e^{-s})}e^{-\mu(1-e^{-s})} \\ &= e^{-(\lambda+\mu)(1-e^{-s})}, s \geq 0。 \end{aligned}$$

可看出 $e^{-(\lambda+\mu)(1-e^{-s})}$ 對應 $\mathcal{P}(\lambda+\mu)$ 分佈之拉普拉斯轉換, 故 $X+Y$ 有 $\mathcal{P}(\lambda+\mu)$ 分佈。

利用適當的轉換, 可求出一些具有常見分佈(參數可能不同)之隨機變數和的分佈。我們列出如下, 證明則留在習題。

定理2.4 設 X_1, X_2 為二獨立的隨機變數。令 $S = X_1 + X_2$ 。則對 $i = 1, 2$,

- (i) 若 X_i 有 $\mathcal{P}(\lambda_i)$ 分佈, 則 S 有 $\mathcal{P}(\lambda_1 + \lambda_2)$ 分佈;
- (ii) 若 X_i 有 $\mathcal{B}(n_i, p)$ 分佈, 則 S 有 $\mathcal{B}(n_1 + n_2, p)$ 分佈;
- (iii) 若 X_i 有 $\Gamma(\alpha_i, \beta)$ 分佈, 則 S 有 $\Gamma(\alpha_1 + \alpha_2, \beta)$ 分佈;
- (iv) 若 X_i 有 $\chi_{n_i}^2$ 分佈, 則 S 有 $\chi_{n_1+n_2}^2$ 分佈;
- (v) 若 X_i 有 $\mathcal{N}(\mu_i, \sigma_i^2)$ 分佈, 則 S 有 $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ 分佈;
- (vi) 若 X_i 有自0開始之 $Ge(p)$ 分佈, 則 S 有 $\mathcal{NB}(2, p)$ 分佈;
- (vii) 若 X_i 有 $\mathcal{NB}(r_i, p)$ 分佈, 則 S 有 $\mathcal{NB}(r_1 + r_2, p)$ 分佈。

並非所有獨立的同一類分佈之隨機變數的和仍有同一類分佈(當然如果將幾何分佈視為負二項分佈, 則(vi)為(vii)之一特例)。定理2.4(vi)為一例。另外, 設 X_1, X_2 獨立。則當 X_1, X_2 皆有 $\mathcal{U}(0, 1)$ 分佈時, $X_1 + X_2$ 並無均勻分佈。讀者也可練習再找一些例子。

我們再給一有趣的例子。

例2.17 對一正整數 i ，假設任選一正整數是 i 的倍數之機率為 $1/i$ 。這個假設是合理的，原因則讓大家自行想想。利用

$$(2.24) \quad \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}。$$

試求任選二正整數互質的機率 p 。

解.任選的二正整數以 X, Y 表之，依題意 X 與 Y 應獨立。以 (X, Y) 表 X 與 Y 之最大公因數，互質就是 $(X, Y) = 1$ 。 $(X, Y) = i$ ，表 $X = iU, Y = iV$ ，其中 $(U, V) = 1$ 。 X, Y 皆要屬於 i 的倍數之集合 $\{i \cdot 1, i \cdot 2, \dots\}$ ，此機率為 $(1/i) \cdot (1/i) = 1/i^2$ ，又 $P((U, V) = 1) = p$ 。故 $P((X, Y) = i) = (1/i^2) \cdot p = p/i^2$ 。因此

$$1 = \sum_{i=1}^{\infty} P((X, Y) = i) = \sum_{i=1}^{\infty} \frac{p}{i^2}。$$

故得

$$p = \left(\sum_{i=1}^{\infty} \frac{1}{i^2} \right)^{-1} = \frac{6}{\pi^2} \doteq 0.6079。$$

即任選二正整數互質比不互質較可能發生。

我們再給兩個補充說明。首先，若(2.12)(或(2.16))式在一集合 A 不成立，但 $\int \int_A dx dy = 0$ ，則 X 與 Y 仍為獨立的隨機變數。例如，若

$$f(x, y) = \begin{cases} e^{-\lambda x - \mu y} & , x, y > 0, x \neq y, \\ e^{-\lambda x y} & , x, y > 0, x = y。 \end{cases}$$

令 $A = \{(x, y) | x, y > 0, x = y\}$ 。由於

$$\int \int_A dx dy = \int_0^{\infty} \int_y^y 1 dx dy = \int_0^{\infty} 0 dy = 0,$$

故 X 與 Y 獨立。

其次,若 X 與 Y 獨立, 由(2.15)式, 可得

$$E(Y|X) = E(Y),$$

條件期望值與給定的 X 無關。反之若上式成立, 是否導致 X 與 Y 獨立呢? 此答案我們留在第3.4節的習題第5題。

本節最後我們來介紹**偏斜常態分佈**(skew normal distribution)。此分佈族包含常態分佈, 且與常態分佈有一些相同或相似的性質。但此分佈由於有一**偏斜參數**(skewness parameter), 適合當做有**偏斜數據**(即p.d.f.並非左右對稱, 期望值也不等於中位數)的機率模型。

一隨機變數 Z , 若其p.d.f.為

$$(2.25) \quad f(z) = 2\phi(z)\Phi(\lambda z), -\infty < z < \infty,$$

其中 $\lambda \in R$, 且 $\phi(z)$, $\Phi(z)$ 分別為 $\mathcal{N}(0, 1)$ 分佈之p.d.f.及分佈函數, 便稱為有參數 λ 之**偏斜常態分佈**, 以 $\mathcal{SN}(\lambda)$ 表之。

圖2.1給出兩個 $\mathcal{SN}(\lambda)$ 分佈p.d.f.的圖形。可看出圖形是有偏斜的, 不像常態分佈的圖形為左右對稱。

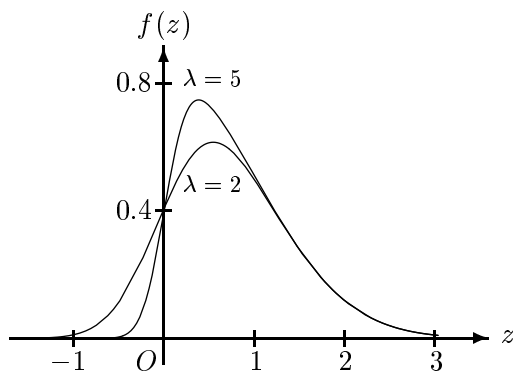


圖2.1 $\mathcal{SN}(2)$ 及 $\mathcal{SN}(5)$ p.d.f.之圖形

我們列出一些偏斜常態分佈簡單的性質如下。

- (i) $\mathcal{SN}(0) = \mathcal{N}(0, 1)$;
- (ii) 若 $Z \sim \mathcal{SN}(\lambda)$, 則 $-Z \sim \mathcal{SN}(-\lambda)$;
- (iii) 若 $Z \sim \mathcal{SN}(\lambda)$, 則 $Z^2 \sim \chi_1^2$;

(iv) 若 $Z \sim \mathcal{SN}(\lambda)$, 則 Z 之分佈函數為

$$\Phi(z; \lambda) = P(Z \leq z) = 2 \int_{-\infty}^z \int_{-\infty}^{\lambda t} \phi(t)\phi(u)du dt;$$

(v) $1 - \Phi(-z; \lambda) = \Phi(z; -\lambda)$;

(vi) $\Phi(z; 1) = (\Phi(z))^2$ 。

上述這些性質的證明並不困難, 因此留在習題中, 其他性質可參考 Azzalini(1985)。

由表面的形式, 一時可能不易看出(2.23)式的確定義出一p.d.f.。即使由性質(iv), 也不易看出

$$\lim_{z \rightarrow \infty} \Phi(z; \lambda) = 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\lambda t} \phi(t)\phi(u)du dt = 1。$$

事實上, 設 X, Y 為二獨立的隨機變數, 且皆有 $\mathcal{N}(0, 1)$ 分佈, 因 $X \stackrel{d}{=} (-X)$, $Y \stackrel{d}{=} (-Y)$, 故

$$-(X - \lambda Y) = (-X) - \lambda(-Y) \stackrel{d}{=} X - \lambda Y。$$

因此

$$P(X - \lambda Y < 0) = \frac{1}{2}。$$

又

$$\begin{aligned} \frac{1}{2} &= P(X - \lambda Y < 0) = P(X < \lambda Y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\lambda y} \phi(x)\phi(y)dx dy \\ &= \int_{-\infty}^{\infty} \phi(y)\Phi(\lambda y)dy。 \end{aligned}$$

由於 $2\phi(z)\Phi(\lambda z) \geq 0$, 且

$$\int_{-\infty}^{\infty} 2\phi(z)\Phi(\lambda z)dz = 1,$$

故 $2\phi(z)\Phi(\lambda z)$ 為一p.d.f.。習題裡我們會給出較一般的結果。底下先給一例。

例2.18 設 X, Y 為二獨立的隨機變數, 且皆有 $\mathcal{N}(0, 1)$ 分佈。令

$$Z = \begin{cases} X & , \text{若 } \lambda X > Y, \\ -X & , \text{若 } \lambda X \leq Y, \end{cases}$$

試證對 $\forall \lambda \in R$, Z 有 $\mathcal{SN}(\lambda)$ 分佈。

證明.我們只證 $\lambda > 0$ 的情況, $\lambda \leq 0$ 的情況可類似地證出。底下來求 Z 之分佈函數。

$$\begin{aligned} (2.26) \quad F(z) &= P(Z \leq z) \\ &= P(Z \leq z, \lambda X > Y) + P(Z \leq z, \lambda X \leq Y) \\ &= P(X \leq z, X > Y/\lambda) + P(-X \leq z, \lambda X \leq Y) \\ &= P(Y/\lambda < X \leq z, Y \leq \lambda z) + P(-z \leq X \leq Y/\lambda, Y \geq -\lambda z) \\ &= \int_{-\infty}^{\lambda z} \int_{y/\lambda}^z \phi(x)\phi(y)dx dy + \int_{-\lambda z}^{\infty} \int_{-z}^{y/\lambda} \phi(x)\phi(y)dx dy \\ &= \int_{-\infty}^{\lambda z} (\Phi(z) - \Phi(y/\lambda))\phi(y)dy + \int_{-\lambda z}^{\infty} (\Phi(y/\lambda) - \Phi(-z))\phi(y)dy \\ &= \Phi(z)\Phi(\lambda z) - \int_{-\infty}^{\lambda z} \Phi(y/\lambda)\phi(y)dy \\ &\quad + \int_{-\lambda z}^{\infty} \Phi(y/\lambda)\phi(y)dy - \Phi(-z)(1 - \Phi(-\lambda z))。 \end{aligned}$$

此處用到

$$P(X \leq z, X > Y/\lambda, Y > \lambda z) = 0,$$

$$P(-X \leq z, \lambda X \leq Y, Y < -\lambda z) = 0,$$

$$P(X \leq z, X > Y/\lambda, Y \leq \lambda z) = P(Y/\lambda < X \leq z, Y \leq \lambda z),$$

且

$$P(-X \leq z, \lambda X \leq Y, Y \geq -\lambda z) = P(-z \leq X \leq Y/\lambda, Y \geq -\lambda z)。$$

現利用

$$\phi(z) = \phi(-z), \quad \Phi(-z) = 1 - \Phi(z),$$

及微積分基本定理, 由(2.26)式即得

$$(2.27) \quad \frac{dF(z)}{dz} = 2\phi(z)\Phi(\lambda z), -\infty < z < \infty .$$

得證 Z 有 $\mathcal{SN}(\lambda)$ 分佈。