第3章　多次元の確率変数

3.1　同時分布と周辺分布

$(\Omega, \mathcal{F}, P)$ を確率空間とし，X, Y をこの確率空間上の確率変数とする．これらふたつの確率変数を組として考えた (X, Y) を2 **次元確率ベクトル**という．さらに，(X, Y) の分布を同時分布とよび，任意の $A, B \in \mathcal{B}(\mathbb{R})$ に対して，

$$P_{X,Y}(A \times B) = P\{\omega \in \Omega : X(\omega) \in A, Y(\omega) \in B\} = P(X \in A, Y \in B)$$

で定める．すなわち，$P_{X,Y}$ は確率ベクトル (X, Y) によって P より誘導された $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ 上[1]の確率測度である．

X, Y それぞれの分布 P_X, P_Y をそれぞれの**周辺分布**という．

定義 3.1　2 次元確率ベクトル (X, Y) の**同時分布関数**を

$$\begin{aligned} F_{X,Y}(x, y) &= P_{X,Y}((-\infty, x] \times (-\infty, y]) \\ &= P(X \le x, Y \le y) \\ &= P(\{\omega \in \Omega : X(\omega) \le x, Y(\omega) \le y\}), \qquad x, y \in \mathbb{R} \end{aligned}$$

で定める．各成分だけに注目した分布関数

$$F_X(x) = P_X((-\infty, x]) = P(X \le x), \qquad F_Y(y) = P_Y((-\infty, y]) = P(Y \le y)$$

をそれぞれの**周辺分布関数**とよぶ．

同時分布関数の性質

(1)　すべての $(x, y) \in \mathbb{R}^2$ に対して，$0 \le F_{X,Y}(x, y) \le 1$.

(2)　$x_1 < x_2, y_1 < y_2$ に対して，$F_{X,Y}(x_1, y_1) \le F_{X,Y}(x_2, y_2)$.

(3) $\lim_{x\to-\infty} F_{X,Y}(x, y) = 0$,　$\lim_{y\to-\infty} F_{X,Y}(x, y) = 0$,　$\lim_{x\to\infty, y\to\infty} F_{X,Y}(x, y) = 1$.

同時分布関数と周辺分布関数の関係

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x, y), \qquad F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y)$$

[1] $\mathcal{B}(\mathbb{R}^2) = \sigma(\{A \times B : A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{B}(\mathbb{R})\})$ である．

定義 3.2　2 つの確率変数 X, Y が独立であるとは，その同時分布 $P_{X,Y}$ が周辺分布 P_X, P_Y の積で表されることである：すなわち，任意の A, $B \in \mathcal{B}(\mathbb{R})$ に対して

$$P_{X,Y}(A \times B) = P_X(A)P_Y(B)$$

が成り立つこと[2]である．独立でないときを従属という．

注意 3.1　つぎは同値である．

(i)　X, Y は独立である．

(ii)　$F_{X,Y}(x, y) = F_X(x)F_Y(y)$．ただし，$x, y \in \mathbb{R}$ である．

3.1.1　離散同時確率関数

確率変数 X, Y はともに離散型であって，それぞれは高々可算個の点で値をとるとする．

定義 3.3 離散型確率変数 (X, Y) の**同時確率関数**とは，$\mathbb{R}^2$ 上の実数値関数 $f_{X,Y}(x, y)$ で

$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

をみたすものをいう．

$S = \{(x, y) \in \mathbb{R}^2 : f_{X,Y}(x, y) > 0\}$ とおけば，S は可算集合となる．さらに，$S_x = \{x \in \mathbb{R} : f_{X,Y}(x, y) > 0($ある $y \in \mathbb{R}\}$ と $S_y = \{y \in \mathbb{R} : f_{X,Y}(x, y) > 0($ある $x \in \mathbb{R}\}$ とする．このとき，同時確率関数は

(i) $f_{X,Y}(x, y) \geq 0$

(ii) $\sum_{(x,y)\in S} f_{X,Y}(x, y) = 1$

(iii) $\mathbb{R}^2$ の任意の部分集合[3] A に対して，$P((X, Y) \in A) = \sum_{(x,y)\in A\cap S} f_{X,Y}(x, y)$

定義 3.4 離散型確率変数 (X, Y) の同時分布関数とは，$\mathbb{R}^2$ 上の実数値関数 $F_{X,Y}(x, y)$ で

$$F_{X,Y}(x, y) = \sum_{(s,t):s\leq x,\, t\leq y,\, (s,y)\in S} f_{X,Y}(s, t)$$

で定義されるものをいう．

X と Y のそれぞれの確率関数を

$$f_X(x) = P(X = x), \qquad f_Y(x) = P(Y = y)$$

で定めることにする．同時確率関数に対して，f_X と f_Y を X と Y の周辺確率関数ということにする．

[2]これは $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ である．

[3]正確には，任意のボレル集合 $A \in \mathcal{B}(\mathbb{R}^2)$

定理 3.1 離散型確率変数 $(X,\,Y)$ は同時確率関数 $f_{X,Y}(x,\,y)$ を持つとする．このとき，

$$
\begin{aligned}
f_X(x) &= \sum_{y\in S_y} f_{X,Y}(x,\,y),\\
f_Y(y) &= \sum_{x\in S_x} f_{X,Y}(x,\,y)
\end{aligned}
$$

が成立する．

証明　f_X について示す．$A_x = \{(x,\,y)\in\mathbb{R}^2 : -\infty < y < \infty\}$ とおく．このとき，$s\in S_x$ に対して，

$$
\begin{aligned}
f_X(x) &= P(X=x)\\
&= P(X=x,\,-\infty<y<\infty)\\
&= P((X,\,Y)\in A_x)\\
&= \sum_{(x,\,y)\in A_x\cap S} f_{X,Y}(x,\,y)\\
&= \sum_{y\in S_y} f_{X,Y}(x,\,y)
\end{aligned}
$$

よりわかる．f_Y についても同様に示される． □

3.1.2　同時確率密度関数

定義 3.5 連続型確率ベクトル $(X,\,Y)$ とし，$F_{X,Y}(x,\,y)$ をその同時分布関数とする．$\mathbb{R}^2$ 上の実数値関数 $f_{X,Y}(x,\,y)$ ですべての $x\in\mathbb{R},\,y\in\mathbb{R}$ に対して，

$$
F_{X,Y}(x,\,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,\,t)\,ds\,dt
$$

をみたすものが存在するとき，$f_{X,Y}(x,\,y)$ を $(X,\,Y)$ の**同時確率密度関数**という．

<u>同時確率密度関数の性質</u>

(i) すべての $x\in\mathbb{R},\,y\in\mathbb{R}$ に対して，$f_{X,Y}(x,\,y)\geq 0$.

(ii) すべての $A\in\mathbb{R}^2$ に対して，$P((X,\,Y)\in A) = \int\int_A f_{X,Y}(x,\,y)\,dx\,dy$.

(iii) $F_{X,Y}(x,\,y)$ が同時確率密度関数を持つならば，$x\in\mathbb{R},\,y\in\mathbb{R}$ に対して，

$$
f_{X,Y}(x,\,y) = \frac{\partial^2}{\partial x\partial y}F_{X,Y}(x,\,y)
$$

となる．

注意 3.2 確率ベクトル (X, Y) が同時確率密度関数 $f_{X,Y}(x, y)$ を持つとき，X と Y の周辺確率密度関数は

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx$$

と表現できること[4]に注意せよ．

3.1.3 独立性

定義 3.6 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ をもつとする．このとき，X と Y が**独立**であるとは，すべての $x \in \mathbb{R}, y \in \mathbb{R}$ に対して

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

が成立することである．

補題 3.1 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ をもつとする．このとき，X と Y が**独立**であるとはための必要十分条件は，$\mathbb{R}$ 上で定義されたある関数 $g(x)$ と $h(y)$ が存在し，すべての $x \in \mathbb{R}, y \in \mathbb{R}$ に対して

$$f_{X,Y}(x, y) = g(x)h(y)$$

とかけることである．

証明 $\Rightarrow$ (必要条件) は $g(x) = f_X(x), h(y) = f_Y(y)$ とおけばよい．
$\Leftarrow$ (十分条件) は連続型についてのみ示すことにする．同時確率密度関数が $f_{X,Y}(x, y) = g(x)h(y)$ と表現されたとする．さらに，

$$\int_{-\infty}^{\infty} g(x)\,dx = c, \qquad \int_{-\infty}^{\infty} h(y)\,dy = d$$

とおくと定数 c と d は関係式

$$\begin{aligned} cd &= \left(\int_{-\infty}^{\infty} g(x)\,dx\right)\left(\int_{-\infty}^{\infty} h(y)\,dy\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx\,dy = 1 \end{aligned} \tag{3.1}$$

をみたす．さらに，

$$f_X(x) = \int_{-\infty}^{\infty} g(x)h(y)\,dy = g(x)d, \qquad f_Y(y) = \int_{-\infty}^{\infty} g(x)h(y)\,dx = h(y)c \tag{3.2}$$

となる．(3.1) と (3.2) から

$$f_{X,Y}(x, y) = g(x)h(y) = g(x)h(y)cd = f_X(x)f_Y(y)$$

となり[5]，X と Y が独立であることが示せた． □

[4]証明は定理 3.1 と同様に証明される．
[5]二番目の等号は $cd = 1$，三番目の等号は (3.2) よりわかる．

3.1.4 同時分布に関する期待値

定義 3.7 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ を持つとし，$g(x, y)$ を $\mathbb{R}^2$ 上の実数値関数とする．このとき，$g(X, Y)$ の期待値を

$$\mathbb{E}[g(X, Y)] = \begin{cases} \sum_{(x,y)\in S} g(x, y) f_{X,Y}(x, y), & \text{(離散型)} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, dx\, dy, & \text{(連続型)} \end{cases}$$

で定義する．ただし，離散型の場合は $\sum_{(x,y)\in S} |g(x, y)| f_{X,Y}(x, y) < \infty$ のとき，連続型の場合は $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |g(x, y)| f_{X,Y}(x, y)\, dx\, dy < \infty$ のとき，X の期待値を定義することにする．期待値が定義されるとき，$g(X, Y)$ の期待値が存在するという．

記法について

確率変数のベクトルや行列に対する期待値の作用を以下のように書くことにする．たとえば，確率ベクトル (X, Y) に対して，

$$\mathbb{E}(X, Y) = (\mathbb{E}(X), \mathbb{E}(Y))$$

などと書き，行列の成分が確率変数である確率行列に対しては，

$$\mathbb{E}\begin{bmatrix} X^2 & XY \\ XY & Y^2 \end{bmatrix} = \begin{bmatrix} \mathbb{E}[X^2] & \mathbb{E}[XY] \\ \mathbb{E}[XY] & \mathbb{E}[Y^2] \end{bmatrix}$$

である．

定理 3.2 X と Y は独立な確率変数とし，実数上で定義された実数値関数 $g(x)$ と $h(y)$ は x と y にのみにそれぞれ依存するものとする．このとき，

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)]$$

が成立する．ただし，それぞれの期待値は存在するものと仮定する．

証明 (X, Y) がともに連続型確率変数とし，同時確率密度関数 $f_{X,Y}(x, y)$ を持つ場合について証明する．独立性の定義を利用すれば，

$$\begin{aligned} \mathbb{E}[g(X)]\mathbb{E}[h(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y) f_{X,Y}(x, y)\, dx\, dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(x) f_X(x) f_Y(y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} h(x) f_X(x) \left(\int_{-\infty}^{\infty} g(y) f_Y(y)\, dy \right) dx \\ &= \int_{-\infty}^{\infty} h(x) f_X(x)\, dx \int_{-\infty}^{\infty} g(y) f_Y(y)\, dy = \mathbb{E}[g(X)]\mathbb{E}[h(Y)] \end{aligned}$$

より示せた．離散型の場合は積分記号を和の記号に直せたよい． □

定理 3.3 X と Y は独立な確率変数とし，それぞれは積率母関数 $M_X(t)$ と $M_Y(t)$ を持つとする．このとき，$Z = X + Y$ の積率母関数は

$$M_Z(t) = M_X(t)M_Y(t)$$

で与えられる．

証明 定理 3.2 から

$$M_Z(t) = \mathbb{E}[e^{tZ}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = M_X(t)M_Y(t)$$

がわかる． □

注意 3.3 X と Y は独立な確率変数とし，それぞれは $N(\mu_1, \sigma_2^2)$ と $N(\mu_2, \sigma_2^2)$ に従うとする．このとき，それぞれの積率母関数は

$$M_X(t) = \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t), \qquad M_Y(t) = \exp(\mu_2 t + \frac{1}{2}\sigma_2^2 t), \qquad t \in \mathbb{R}$$

となった．$Z = X + Y$ の積率母関数は定理 3.3 から

$$M_Z(t) = M_X(t)M_Y(t) = \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t)\exp(\mu_2 t + \frac{1}{2}\sigma_2^2 t) = \exp\{(\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2\}$$

となる．したがって，Z は $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ に従うことがわかる．