第3章　多次元の確率変数

3.1　同時分布と周辺分布

$(\Omega, \mathcal{F}, \mathbb{P})$ 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，

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

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(3-1)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

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

で定める．各成分だけに注目した分布関数

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

をそれぞれの**周辺分布関数**とよぶ．

命題 3.1 (同時分布関数の性質)　(i)　すべての $(x, y) \in \mathbb{R}^2$ に対して，$0 \le F_{X,Y}(x, y) \le 1$.

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

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

証明　(i) 同時分布関数の定義と確率の定義からわかる．
(ii) $\{(X, Y) \in (-\infty, x_1] \times (-\infty, y_1]\} \subset \{(X, Y) \in (-\infty, x_2] \times (-\infty, y_2]\}$ に注意して，命題 1.4(vi) を用いればよい．
(iii) $\cap_{n=1}^{\infty}(-\infty, -n] = \emptyset$ に注意をして，命題 1.4(i) と (ix) を用いると

$$\begin{aligned} \lim_{n\to\infty} F_{X,Y}(x, -n) &= \lim_{n\to\infty} \mathbb{P}((X, Y) \in (-\infty, x] \times (-\infty, -n]) \\ &= \mathbb{P}((X, Y) \in \cap_{n=1}^{\infty}(-\infty, x] \times (-\infty, -n]) = \mathbb{P}(\emptyset) = 0 \end{aligned}$$

よりわかる．のこりも同様である．　□

命題 3.2 (同時分布関数と周辺分布関数の関係)

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

証明　命題 1.4(viii) に注意して

$$\begin{aligned}
\lim_{n\to\infty} F_{X,Y}(x, n) &= \lim_{n\to\infty} \mathbb{P}((X, Y) \in (-\infty, x] \times (-\infty, n]) \\
&= \mathbb{P}((X, Y) \in \cup_{n=1}^{\infty} (-\infty, x] \times (\infty, n]) \\
&= \mathbb{P}((X, Y) \in \cup_{n=1}^{\infty} (-\infty, x] \times \mathbb{R}) = \mathbb{P}(X \le x)
\end{aligned}$$

からわかる．□

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(3-2)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．
(1) X, Y は独立である．
(2) $F_{X,Y}(x, y) = F_X(x)F_Y(y)$．ただし，$x, y \in \mathbb{R}$ である．
証明　(1) ⇒ (2) は独立性の定義において，$A = (-\infty, x]$, $B = (-\infty, y]$ とすればわかる．逆については，略．□

3.1.1　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

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

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

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

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

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

$$F_{X,Y}(x, y) = \mathbb{P}\{(X, Y) \in (-\infty, x] \times (-\infty, y]\} = \sum_{(s, t): s\le x, t\le y, (s, y)\in S} f_{X,Y}(s, t)$$

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．

命題 3.3 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，$x \in S_x$ に対して，

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

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 $A \subset \mathbb{R}^2$ に対して，

$$\mathbb{P}((X, Y) \in A) = \int\int_A f_{X,Y}(x, y)\, dx\, dy$$

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．

命題 3.4 (同時確率密度関数の性質) (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 $(x, y) \in \mathbb{R}^2$ に対して，

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

(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．

証明 証明は明らか． □

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(3-4)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(3-5)，X と Y が独立であることが示せた． □

例 3.1 離散型確率ベクトル (X, Y) の同時確率関数が以下のように与えられているとする：

$$\begin{aligned} f_{X,Y}(0, 10) &= f_{X,Y}(0, 20) = \frac{2}{18}, \quad f_{X,Y}(1, 10) = f_{X,Y}(1, 30) = \frac{3}{18}, \\ f_{X,Y}(2, 20) &= \frac{4}{18}, \qquad f_{X,Y}(2, 30) = \frac{4}{18}. \end{aligned}$$

ただし，その他の $(x,\, y)$ では $f_{X,Y}(x,\, y) = 0$ である．X の周辺確率関数は

$$f_X(0) = \frac{4}{18},\, f_X(1) = \frac{6}{18},\, f_X(2) = \frac{8}{18}$$

となり，Y の周辺確率関数は

$$f_Y(10) = \frac{5}{18}, \qquad f_Y(20) = \frac{6}{18}, \qquad f_Y(30) = \frac{7}{18}$$

となる．よって，X と Y は独立でない．たとえば，

$$f_{X,Y}(0,\, 10) = \frac{2}{18} \neq \frac{4}{18} \times \frac{5}{18} = f_X(0)f_Y(10)$$

からわかる．

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，$g(X,\, Y)$ の期待値を定義することにする．また，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，$g(X,\, Y)$ 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 3.1 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 $h_1(x)$ と $h_2(y)$ 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，

$$\mathbb{E}[h_1(X)h_2(Y)] = \mathbb{E}[h_1(X)]\mathbb{E}[h_2(Y)]$$

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，

$$\begin{aligned}\mathbb{E}[h_1(X)h_2(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(x)h_2(y)f_{X,Y}(x, y)\,dx\,dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(x)h_2(y)f_X(x)f_Y(y)\,dx\,dy \\ &= \int_{-\infty}^{\infty} h_1(x)f_X(x)\left(\int_{-\infty}^{\infty} h_2(y)f_Y(y)\,dy\right)dx \\ &= \int_{-\infty}^{\infty} h_1(x)f_X(x)\,dx\int_{-\infty}^{\infty} h_2(y)f_Y(y)\,dy = \mathbb{E}[h_1(X)]\mathbb{E}[h_2(Y)]\end{aligned}$$

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 3.2 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 3.1 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 3.2 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．

3.2 条件付き分布と独立性

3.2.1 離散型確率変数の場合

定義 3.8 (X, Y) は離散型確率ベクトルとし，同時確率関数 $f_{X,Y}(x, y)$ および周辺確率関数 $f_X(x)$ と $f_Y(y)$ を持つとする．
(i) $\mathbb{P}(X = x) = f_X(x) > 0$ なる任意の x に対して，$X = x$ が与えられたときの Y の **条件付確率関数**を $f_{Y|X}(y|x)$ で記し，

$$f_{Y|X}(y|x) = \mathbb{P}(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

で定める．
(ii) $\mathbb{P}(Y = y) = f_Y(y) > 0$ なる任意の y に対して，$Y = y$ が与えられたときの X の **条件付確率関数**を $f_{X|Y}(x|y)$ で記し，

$$f_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

で定める．

注意 3.4 $f_{Y|X}(y|x)$ は確率関数であることに注意せよ．すなわち，各 x に対して，

- $f_{Y|X}(y|x) \geq 0,\ y \in \mathbb{R}$,
- $\sum_{y \in S_Y} f_{Y|X}(y|, x) = 1$

となっている．

例 3.2 離散型確率ベクトル (X, Y) の同時確率関数が以下のように与えられているとする：

$$\begin{aligned} f_{X,Y}(0, 10) &= f_{X,Y}(0, 20) = \frac{2}{18}, \quad f_{X,Y}(1, 10) = f_{X,Y}(1, 30) = \frac{3}{18}, \\ f_{X,Y}(1, 20) &= \frac{4}{18}, \qquad f_{X,Y}(2, 30) = \frac{4}{18}. \end{aligned}$$

ただし，その他の (x, y) では $f_{X,Y}(x, y) = 0$ である．$X = x,\ x = 0, 1, 2$ が与えられたときの Y の条件付確率関数を求めよう．そのために，X の周辺確率関数を求める：

$$\begin{aligned} f_X(0) &= f_{X,Y}(0, 10) + f_{X,Y}(0, 29) = \frac{4}{18}, \\ f_X(1) &= f_{X,Y}(1, 10) + f_{X,Y}(1, 20) + f_{X,Y}(1, 30) = \frac{10}{18}, \\ f_X(2) &= f_{X,Y}(2, 30) = \frac{4}{18}. \end{aligned}$$

$x = 0$ のとき，$y = 10, 20$ のとき $f_{X,Y}(0, y) > 0$ であるので，$y = 10, 20$ のとき $f_{Y|X}(y|0) > 0$ となり，

$$\begin{aligned} f_{Y|X}(10|0) &= \frac{f_{X,Y}(0, 10)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2}, \\ f_{Y|X}(20|0) &= \frac{f_{X,Y}(0, 20)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2} \end{aligned}$$

となる．したがって，$X = 0$ という情報から Y の条件付確率は $y = 10, 20$ にそれぞれ $1/2$ の確率を与える．

$x=1$ のとき，$y=10,\,20,\,30$ のとき $f_{Y|X}(y|1)>0$ となり，

$$\begin{aligned} f_{Y|X}(10|1) &= \frac{f_{X,Y}(1,\,10)}{f_X(1)} = \frac{\frac{3}{18}}{\frac{10}{18}} = \frac{3}{10}, \\ f_{Y|X}(20|1) &= \frac{f_{X,Y}(1,\,20)}{f_X(1)} = \frac{\frac{4}{18}}{\frac{10}{18}} = \frac{4}{10}, \\ f_{Y|X}(30|1) &= \frac{f_{X,Y}(1,\,30)}{f_X(1)} = \frac{\frac{3}{18}}{\frac{10}{18}} = \frac{3}{10}, \end{aligned}$$

となる．したがって，$X=1$ という情報から Y の条件付確率は $y=10,\,20,\,30$ にそれぞれ 3/10, 4/10, 3/10 の確率を与える．

$x=2$ のとき，$y=30$ のとき $f_{Y|X}(y|2)>0$ となり，

$$f_{Y|X}(30|2) = \frac{f_{X,Y}(2,\,30)}{f_X(2)} = \frac{\frac{4}{18}}{\frac{4}{18}} = 1$$

となる．したがって，$X=2$ という情報から $Y=30$ がわかる．

たとえば，

$$\begin{aligned} \mathbb{P}(Y>10|\,X=0) &= f_{Y|X}(20|0) = \frac{1}{2}, \\ \mathbb{P}(Y>10|\,X=1) &= f_{Y|X}(20|1) + f_{Y|X}(30|1) = \frac{7}{10}, \end{aligned}$$

となる．

3.2.2 連続型確率変数の場合

定義 3.9 $(X,\,Y)$ は連続型確率ベクトルとし，同時確率密度関数 $f_{X,Y}(x,\,y)$ および周辺確率密度関数 $f_X(x)$ と $f_Y(y)$ を持つとする．

(i) $f_X(x)>0$ なる任意の x に対して，$X=x$ が与えられたときの Y の **条件付確率密度関数**を $f_{Y|X}(y|x)$ で記し，

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,\,y)}{f_X(x)}$$

で定める．

(ii) $f_Y(y)>0$ なる任意の y に対して，$Y=y$ が与えられたときの X の **条件付確率密度関数**を $f_{X|Y}(x|y)$ で記し，

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,\,y)}{f_Y(y)}$$

で定める．

例 3.3 連続型確率ベクトル $(X,\,Y)$ は同時確率密度関数

$$f_{X,Y}(x,\,y) = \begin{cases} e^{-y}, & 0<x<y<\infty, \\ 0, & \text{その他} \end{cases}$$

を持つとする．$X=x$ が与えられたときの Y の条件付確率密度関数を求めるために，X の周辺確率密度関数を求めよう．$x\le 0$ の場合，すべての y に対して $f_{X,Y}(x,\,y)=0$ なので，$f_X(x)=0$ となる．$x>0$ の場合，$y>x$ ならば，$f_{X,Y}(x,\,y)>0$ なので，

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,\,y)\,dy = \int_x^{\infty} f_{X,Y}(x,\,y)\,dy = e^{-x}$$

となる．したがって

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{その他} \end{cases}$$

となる．これより，$X = x$ が与えられたときの Y の条件付確率密度関数は $x > 0$ の場合のみに定義される．各 $x > 0$ に対して

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \qquad y > x, \\ f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, \qquad y \le x \end{aligned}$$

となる．

3.2.3 独立性との関係

注意 3.5 もし，X と Y が独立ならば，x の値に関わらず

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

となる．

3.3 条件付き期待値

定義 3.10 $X = x$ が与えられたときの Y の条件付確率関数または条件付確率密度関数を $f_{Y|X}(y|,x)$ とする．$g:\mathbb{R}\to\mathbb{R}$ としたとき，$X = x$ が与えられたときの $g(Y)$ の条件付期待値を

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

で定める．ただし，条件付期待値は $\mathbb{E}[|g(Y)||\,x] < \infty$ のときに存在するものとする．

命題 3.5 (条件付期待値の性質) a_1, a_2, b を定数，$g_1:\mathbb{R}\to\mathbb{R}$, $g_2:\mathbb{R}\to\mathbb{R}$ とする．

(1) $\mathbb{E}[a_1 g_1(Y) + a_2 g_2(Y) + b|\,x] = a_1\mathbb{E}[g_1(Y)|\,x] + a_2\mathbb{E}[g_2(Y)|\,x] + b.$

(2) $g_1(y) \geq 0$ ならば，$\mathbb{E}[g_1(Y)|\,x] \geq 0.$

(3) $a_1 \leq g_1(y) \leq a_2$ ならば，$a_1 \leq \mathbb{E}[g_1(Y)|\,x] \leq a_2.$

(4) $\mathbb{E}[g_1(X)g_2(Y)|\,x] = g_1(x)\,\mathbb{E}[g_2(Y)|\,x].$

が成立する．ただし，上の条件付期待値はすべて存在するものとする．

証明 積分の性質からわかる． □

定義 3.11 X と Y を確率変数とし，$\mathbb{E}[Y^2] < \infty$ とする．$X = x$ が与えられたときの Y の条件付分散を

$$\mathbb{VAR}[Y|\,x] = \mathbb{E}[Y^2|\,x] - \{\mathbb{E}[Y|\,x]\}^2$$

で定義する．$v(x) := \mathbb{VAR}[Y|\,x]$ とおいてとき，$\mathbb{VAR}[Y|\,X] := v(X)$ で定める．以上の定義から $\mathbb{VAR}[Y|\,X] = \mathbb{E}[Y^2|\,X] - \{\mathbb{E}[Y|\,X]\}^2$ となることに注意する．

例 3.4 連続型確率ベクトル (X, Y) は同時確率密度関数

$$f_{X,Y}(x,\,y) = \begin{cases} e^{-y}, & 0 < x < y < \infty, \\ 0, & \text{その他} \end{cases}$$

を持つとき，$X = x$ が与えられたときの Y の条件付確率密度関数は $x > 0$ の場合のみに定義され，各 $x > 0$ に対して

$$\begin{aligned} f_{Y|X}(y|\,x) &= \frac{f_{X,Y}(x,\,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \qquad y > x, \\ f_{Y|X}(y|\,x) &= \frac{f_{X,Y}(x,\,y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, \qquad y \leq x \end{aligned}$$

であった．$X = x\,(x > 0)$ が与えられたときの Y の条件付期待値は

$$\mathbb{E}[Y|\,x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|\,x)\,dy = \int_x^{\infty} y e^{-(y-x)}\,dy = 1 + x$$

となる．同様に，$X = x\,(x > 0)$ が与えられたときの Y^2 の条件付期待値は

$$\mathbb{E}[Y^2|\,x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|\,x)\,dy = \int_x^{\infty} y^2 e^{-(y-x)}\,dy = \int_0^{\infty} (t+x)^2 e^{-t}\,dt = x^2 + 2x + 2$$

となる．したがって，$X = x\,(x > 0)$ が与えられたときの Y の条件付分散は

$$\mathbb{VAR}[Y|\,x] = \mathbb{E}[Y^2|\,x] - \{\mathbb{E}[Y|\,x]\}^2 = 1$$

となる．

$X = x$ が与えられたときの $g(Y)$ の条件付期待値（存在するならば）$\mathbb{E}[g(Y)|\, x]$ は x の関数であるので，$h(x) = \mathbb{E}[g(Y)|\, x]$ とおくとき，確率変数 $\mathbb{E}[g(Y)|\, X]$ を $\mathbb{E}[g(Y)|\, X] = h(X)$ で定めることにする．すなわち，$X = x$ のとき，確率変数 $\mathbb{E}[g(Y)|\, X]$ の値は $\mathbb{E}[g(Y)|\, x]$ である．記号の読み方であるが，$\mathbb{E}[g(Y)|\, X]$ に期待値の記号 $\mathbb{E}$ が使われているが，$\mathbb{E}[g(Y)|\, X]$ は X に依存する確率変数である．

定理 3.3 X と Y を確率変数とし，Y の期待値は存在するとする．このとき，

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|\, X]]$$

が成立する．

証明 X と Y が連続型確率変数の場合の証明を与える．$(X,\, Y)$ の同時確率密度関数を $f_{X,Y}(x,\, y)$ とする．$X = x$ が与えられたときの Y の条件付確率密度関数 $f_{Y|\, X}(y|\, x)$ および条件付期待値の定義から

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

となる． □

例 3.5 確率変数 X と Y は同時確率密度関数

$$f_{X,Y}(x,\, y) = \begin{cases} 2 & x + y < 1,\, x > 0,\, y > 0, \\ 0 & (\text{その他}), \end{cases}$$

を持つとする．このとき，X の周辺確率密度関数は，$0 < x < 1$ のとき，

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,\, y)\, dy = \int_0^{-x+1} 2\, dy = 2(1 - x)$$

となる．しがたって，X の周辺確率密度関数は

$$f_X(x) = \begin{cases} 2(1 - x) & 0 < x < 1, \\ 0 & \textit{(その他)}, \end{cases}$$

である．また，$0 < y < 1$ に対して

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,\, y)\, dx = \int_0^{-y+1} 2\, dx = 2(1 - y)$$

となることから Y の周辺確率密度関数は

$$f_Y(y) = \begin{cases} 2(1 - y) & 0 < y < 1, \\ 0 & \textit{(その他)}, \end{cases}$$

となる．これから X の二次までの積率を求めると

$$\begin{aligned}\mathbb{E}[Y] &= \int_{-\infty}^{\infty} x f_Y(y)\,dy = \int_0^1 y2(1-y)\,dy = \frac{1}{3},\\ \mathbb{E}[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y)\,dy = \int_0^1 y^2 2(1-y)\,dy = \frac{1}{6},\\ \mathbb{VAR}[Y] &= \mathbb{E}[Y^2] - \{\mathbb{E}[Y]\}^2 = \frac{1}{18}\end{aligned}$$

となる．$0 < x < 1$ としたとき，$X = x$ が与えられたときの Y の条件付確率密度関数は

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x} & 0 < y < 1-x, \\ 0 & \text{(その他)} \end{cases}$$

となる．したがって，$X = x\,(0 < x < 1)$ が与えられたときの Y の条件付き期待値は

$$\mathbb{E}[Y|\,x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|\,x)\,dy = \int_0^{1-x} y\frac{1}{1-x}\,dy = \frac{1}{1-x}\left[\frac{y^2}{2}\right]_0^{1-x} = \frac{1-x}{2}$$

となる．さらに，

$$\mathbb{E}[\mathbb{E}[Y|\,X]] = \int_{-\infty}^{\infty} \mathbb{E}[Y|\,x] f_X(x)\,dx = \int_0^1 \frac{1-x}{2} 2(1-x)\,dx = \frac{1}{3}$$

となる．

次に，$X = x\,(0 < x < 1)$ が与えられたときの Y の条件付き分散を求める．そのために，

$$\mathbb{E}[Y^2|\,x] = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|\,x)\,dy = \int_0^{1-x} y^2\frac{1}{1-x}\,dy = \frac{1}{1-x}\left[\frac{y^3}{3}\right]_0^{1-x} = \frac{(1-x)^2}{3}$$

となる．よって，$X = x\,(0 < x < 1)$ が与えられたときの Y の条件付き分散は

$$\mathbb{VAR}[Y|\,x] = \mathbb{E}[Y^2|\,x] - \{\mathbb{E}[X|\,x]\}^2 = \frac{(1-x)^2}{3} - \frac{(1-x)^2}{4} = \frac{(1-x)^2}{12}$$

となる．

定理 3.4 X, Y を確率変数とし，$\mathbb{E}[Y^2] < \infty$ とする．このとき，

$$\mathbb{VAR}[Y] = \mathbb{E}[\mathbb{VAR}[Y|\,X]] + \mathbb{VAR}[\mathbb{E}[Y|\,X]]$$

である．ただし，$h(x) = \mathbb{VAR}[Y|\,x]$ とおいたとき，$\mathbb{VAR}[Y|\,X] = h(X)$ で定めた．

証明 まず，

$$\begin{aligned}\mathbb{VAR}[Y] &= \mathbb{E}[\{Y - \mathbb{E}[Y]\}^2] = \mathbb{E}[\{Y - \mathbb{E}[Y|\,X] + \mathbb{E}[Y|\,X] - \mathbb{E}[Y]\}^2]\\ &= \mathbb{E}[\{Y - \mathbb{E}[Y|\,X]\}^2] + \mathbb{E}[\{\mathbb{E}[Y|\,X] - \mathbb{E}[Y]\}^2]\\ &\quad +2\mathbb{E}[\{Y - \mathbb{E}[Y|\,X]\}\{\mathbb{E}[Y|\,X] - \mathbb{E}[Y]\}]\end{aligned}$$

に注意する．上の式の最右辺の 3 項目は

$$\mathbb{E}[\{Y - \mathbb{E}[Y|\,X]\}\{\mathbb{E}[Y|\,X] - \mathbb{E}[Y]\}] = \mathbb{E}[\mathbb{E}[\{Y - \mathbb{E}[Y|\,X]\}\{\mathbb{E}[Y|\,X] - \mathbb{E}[Y]\}|\,X]]$$

となる．ここで，X が与えられたとき，$\mathbb{E}[Y|\,X]$ と $\mathbb{E}[Y]$ は定数であることに注意して，上の式の右辺の期待値の中を評価する：

$$\begin{aligned}\mathbb{E}[\{Y-\mathbb{E}[Y|\,X]\}\{\mathbb{E}[Y|\,X]-\mathbb{E}[Y]\}|\,X] &= \{\mathbb{E}[Y|\,X]-\mathbb{E}[Y]\}\mathbb{E}[Y-\mathbb{E}[Y|\,X]|\,X] \\ &= \{\mathbb{E}[Y|\,X]-\mathbb{E}[Y]\}\{\mathbb{E}[Y|\,X]-\mathbb{E}[Y|\,X]\} \\ &= 0\end{aligned}$$

となる．したがって，

$$\mathbb{E}[\{Y-\mathbb{E}[Y|\,X]\}\{\mathbb{E}[Y|\,X]-\mathbb{E}[Y]\}]=0$$

となり，定理は証明された． □

例 3.6 $X_1, X_2, \ldots, X_n, \ldots,$ は独立に同一分布に従う確率変数の列で

$$\mathbb{E}[X_1]=\mu_X, \qquad \mathbb{VAR}[X_1]=\sigma_X^2<\infty$$

とする．N は非負整数値確率変数で確率変数列 $\{X_i\}_{i=1}^{\infty}$ とは独立し，

$$\mathbb{E}[N]=\mu_N, \qquad \mathbb{VAR}[N]=\sigma_N<\infty$$

とする．このとき，N 個までの確率変数列 $\{X_i\}_{i=1}^{\infty}$ の部分和を

$$S_N=\sum_{i=1}^{N}X_i$$

と定義する．ただし，$S_0=0$ とする．定理 3.3 より

$$\mathbb{E}[S_N]=\mathbb{E}[\mathbb{E}[S_N|\,N]]$$

がわかる．しかし，

$$\mathbb{E}[S_N|\,n]=\mathbb{E}[X_1+X_2+\cdots+X_n]=n\mu_X$$

より $\mathbb{E}[S_N|\,N]=N\mu_X$ となる．これから

$$\mathbb{E}[S_N]=\mathbb{E}[N\mu_X]=\mathbb{E}[N]\mu_X=\mu_X\mu_N$$

がわかる．同様に，

$$\mathbb{VAR}[S_N|\,n]=\mathbb{VAR}[X_1+X_2+\cdots+X_n|\,n]=\sum_{i=1}^{n}\mathbb{VAR}[X_i|\,n]=n\mathbb{VAR}[X_1|\,n]=n\sigma_X^2$$

から $\mathbb{VAR}[S_N|\,N]=N\sigma_X^2$ がわかる．定理 3.4 から

$$\begin{aligned}\mathbb{VAR}[S_N] &= \mathbb{E}[\mathbb{VAR}[S_N|\,N]]+\mathbb{VAR}[\mathbb{E}[S_N|\,N]] \\ &= \sigma_X^2\mathbb{E}[N]+\mathbb{VAR}[N\mu_X] \\ &= \sigma_X^2\mathbb{E}[N]+\mu_X^2\mathbb{VAR}[N] \\ &= \sigma_X^2\mu_N+\mu_X^2\sigma_N^2\end{aligned}$$

となる．

N は母数 $p\,(0<p\leq 1)$ の幾何分布に従うとする:

$$f_N(n)=\begin{cases} p(1-p)^{n-1} & n=1,2,\ldots \\ 0 & \text{（その他）}\end{cases}$$

とする．したがって

$$\mathbb{E}[N] = \frac{1}{p}, \qquad \mathbb{VAR}[N] = \frac{1-p}{p^2}, \qquad m_N(t) = \frac{pe^t}{1-(1-p)e^t}$$

となる．これは

$$\begin{aligned} \mathbb{E}[e^{tN}] &= \sum_{n=1}^{\infty} e^{tn} p(1-p)^{n-1} = \sum_{n=0}^{\infty} pe^t \{(1-p)e^t\}^{n-1} \\ &= \frac{pe^t}{1-(1-p)e^t} \sum_{n=0}^{n} \{1-(1-p)e^t\}\{1-(1-(1-p)e^t)\}^{n-1} \\ &= \frac{pe^t}{1-(1-p)e^t} \end{aligned}$$

よりわかる．さらに，X_1 は母数 $\lambda\,(\lambda > 0)$ の指数分布に従うとする：

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & (\text{その他}) \end{cases}$$

とする．したがって，

$$\mathbb{E}[X] = \frac{1}{\lambda}, \qquad \mathbb{VAR}[X] = \frac{1}{\lambda^2}, \qquad m_X(t) = \frac{\lambda}{\lambda - t}, \qquad t < \lambda$$

である．

定理 3.3 から

$$\mathbb{E}[S_N] = \mu_N \mu_X = \frac{1}{p}\frac{1}{\lambda} = \frac{1}{p\lambda}$$

となる．定理 3.4 から

$$\mathbb{VAR}[S_N] = \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2 = \frac{1}{p}\frac{1}{\lambda} + \frac{1-p}{p^2}\frac{1}{\lambda^2} = \frac{1}{(p\lambda)^2}$$

がわかる．

S_n の分布を求めてみよう．そのために，

$$I\{S_N \le z\} = \begin{cases} 1 & (S_N \le z \text{ が起きたとき}), \\ 00 & (\text{その他}) \end{cases}$$

とおけば，$\mathbb{P}[S_N \le z] = \mathbb{E}[I\{S_N \le z\}]$ と $\mathbb{P}[S_n \le z |\, N = n] = \mathbb{E}[I\{S_N \le z\}|\, n]$ となることと S_n は母数 n, λ のガンマ分布

$$f_{S_n}(x) = \begin{cases} \frac{\lambda}{\Gamma(n)}(\lambda x)^{n-1}\lambda e^{-\lambda x} & x > 0 \\ 0 & (\text{その他}) \end{cases}$$

従うことに注意する．$z>0$ に対して

$$\begin{aligned}
\mathbb{P}[S_N \le z] &= \mathbb{E}[I\{S_N \le z\}] \\
&= \mathbb{E}[\mathbb{E}[I\{S_N \le z\}|\, N]] \\
&= \sum_{n=0}^{\infty} \mathbb{E}[I\{S_N \le z\}|\, n] f_N(n) \\
&= \sum_{n=1}^{\infty} \mathbb{E}[I\{S_N \le z\}|\, n] f_N(n) \\
&= \sum_{n=1}^{\infty} \mathbb{P}[S_N \le z|\, N=n] f_N(n) \\
&= \sum_{n=1}^{\infty} \left(\int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x}\, dx \right) p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \int_0^z \lambda p e^{-\lambda z} \frac{[\lambda x(1-p)]^{n-1}}{(n-1)!}\, dx \\
&= \int_0^{\infty} \lambda p e^{-\lambda z} \sum_{n=1}^{\infty} \left(\frac{[\lambda x(1-p)]^{n-1}}{(n-1)!} \right) dx \\
&= \int_0^{\infty} \lambda p e^{-\lambda z} e^{(1-p)\lambda x}\, dx \\
&= \lambda p \int_0^{\infty} e^{-\lambda p x}\, dx \\
&= 1 - e^{-\lambda p z}
\end{aligned}$$

となる．最後から 4 番目の等号における和の記号と積分記号の順序交換はつぎから保障されることがわかる：

$$\int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x}\, dx \le 1$$

から

$$\left| \sum_{n=1}^{\infty} \left(\int_0^z \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} \lambda e^{-\lambda x}\, dx \right) p(1-p)^{n-1} \right| \le \sum_{n=1}^{\infty} p(1-p)^{n-1} = 1$$

よりわかる．

したがって，S_N は母数 $p\lambda$ の指数分布に従うことがわかる．

3.4 共分散と相関係数

X と Y の 2 次までの積率としてそれぞれの平均と分散：

$$\mathbb{E}[X], \qquad \mathbb{E}[Y], \qquad \mathbb{VAR}[X], \qquad \mathbb{VAR}[Y]$$

に加えて，平均まわりの相互積率を X と Y の**共分散**といい，

$$\mathbb{COV}[X, Y] = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$$

と記すこと[3-6]にする．

定理 3.5 (共分散の公式)

$$\mathbb{COV}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

証明

$$\begin{aligned}\mathbb{COV}[X, Y] = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] &= \mathbb{E}[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

がわかる． □

定理 3.6 X と Y が独立ならば，

$$\mathbb{COV}[X, Y] = 0$$

である．

証明

$$\begin{aligned}\mathbb{COV}[X, Y] = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

がわかる． □

注意 3.6 定理 3.6 の逆は必ずしも成立しないことに注意せよ．

定理 3.7 a, b を定数として，$aX + bY$ の期待値と分散はつぎのようになる：

$$\begin{aligned}\mathbb{E}[aX + bY] &= a\mathbb{E}[X] + b\mathbb{E}[Y], \\ \mathbb{VAR}[aX + bY] &= a^2\mathbb{VAR}[X] + b^2\mathbb{VAR}[Y] + 2ab\mathbb{COV}[X, Y]\end{aligned}$$

となる．特に，X と Y が独立ならば，

$$\mathbb{VAR}[aX + bY] = a^2\mathbb{VAR}[X] + b^2\mathbb{VAR}[Y]$$

である．

証明 □

例 3.7 (X, Y) は同時確率密度関数

$$f_{X,Y} = \begin{cases} 1, & 0 < x < 1,\, x < y < x+1, \\ 0, & (\text{その他}), \end{cases}$$

を持つとする．このときの X と Y の共分散を求めよう．まず，X と Y の周辺確率密度関数を求めよう．$0 < x < 1$ のとき，$f_{X,Y}(x, y) > 0$ であることに注意すれば，$0 < x < 1$ のとき，

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy = \int_{x}^{x+1} dy = 1$$

となる．それ以外では $f_{X,Y}(x, y) = 0$ なので，$f_X(x) = 0$ となる．したがって

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & (\text{その他}), \end{cases}$$

となる．また，$0 < y < 2$ のとき，$f_{X,Y}(x, y) > 0$ となることに注意すれば，

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx \begin{cases} \int_0^y 1\, dx, & 0 < y < 1, \\ \int_{y-1}^1 1\, dx, & 1 \le y < 2, \end{cases} = \begin{cases} y, & 0 < y < 1, \\ 2 - y & 1 \le y < 2, \end{cases}$$

となる．したがって

$$f_Y(y) = \begin{cases} y, & 0 < y < 1, \\ 2 - y & 1 \le y < 2, \\ 0, & (\text{その他}) \end{cases}$$

となる．これらに注意すれば，

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x)\, dx \int_0^1 x\, dx = \frac{1}{2}, \\
\mathbb{E}[Y] &= \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^1 y^2\, dy + \int_1^2 y(2-y)\, dy = \frac{1}{3} + \frac{2}{3} = 1, \\
\mathbb{E}[XY] &= = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x, y)\, dx\, dy = \int_0^1 \int_x^{x+1} xy\, dy\, dx = \int_0^1 x \left[\frac{y^2}{2}\right]_x^{x+1} dx \\
&= \int_0^1 \frac{x(2x+1)}{2}\, dx = \left[\frac{x^3}{3} + \frac{x^2}{4}\right]_0^1 = \frac{7}{12}
\end{aligned}$$

となる．したがって，

$$\mathbb{COV}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{7}{12} - \frac{1}{2} = \frac{1}{12}$$

となる．

相関係数

X と Y の標準化

$$\frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{VAR}[X]}}, \qquad \frac{Y - \mathbb{E}[Y]}{\sqrt{\mathbb{VAR}[Y]}}$$

の積の平均を X と Y の**相関係数**といい，

$$\rho[X, Y] = \mathbb{E}\left[\frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{VAR}[X]}} \times \frac{Y - \mathbb{E}[Y]}{\sqrt{\mathbb{VAR}[Y]}}\right] = \frac{\mathbb{COV}[X, Y]}{\sqrt{\mathbb{VAR}[X]}\sqrt{\mathbb{VAR}[Y]}}$$

で記すことにする．

$\rho[X, Y] > 0$ のときには X と Y に**正の相関**，$\rho[X, Y] < 0$ のときには X と Y に**負の相関**があるといい，$\rho[X, Y] = 0$ のときには X と Y は**無相関**であるという．定義と定理 3.6 から，X と Y が独立ならば，X と Y は無相関である．また，注意 3.6 から X と Y は無相関であっても X と Y が必ずしも独立ではない．

定理 3.8 X と Y は 2 次の積率をもつ任意の確率変数とする．このとき，
(i) $-1 \le \rho[X, Y] \le 1$.
(ii) $|\rho[X, Y]| = 1$ となるための必要十分条件は定数 $a, b\,(a \ne 0)$ が存在して，

$$P(Y = aX + b) = 1$$

が成立することである．特に，$\rho[X, Y] = 1$ ならば，$a > 0$ となり，$\rho[X, Y] = -1$ ならば，$a < 0$ となる．

証明 (i) を示すために，

$$h(t) = \mathbb{E}[\{(X - \mu_X)t + (Y - \mu_Y)\}^2]$$

とする．ただし，$\mu_X = \mathbb{E}[X], \mu_Y = \mathbb{E}[Y]$ である．期待値の中を展開すれば，

$$\begin{aligned} h(t) &= t^2\mathbb{E}[(X - \mu_X)^2] + 2t\mathbb{E}[(X - \mu_X)(Y - \mu_Y)] + \mathbb{E}[(Y - \mu_Y)^2] \\ &= t^2\mathbb{VAR}[X] + 2\mathbb{COV}[X, Y] + \mathbb{VAR}[Y] \\ &= \mathbb{VAR}[X]\left(t + \frac{\mathbb{COV}[X, Y]}{\mathbb{VAR}[X]}\right)^2 + \frac{\mathbb{VAR}[X]\mathbb{VAR}[Y] - \{\mathbb{COV}[X, Y]\}^2}{\mathbb{VAR}[X]} \end{aligned}$$

となる．しかし，$h(t) \ge$ から

$$\mathbb{VAR}[X]\mathbb{VAR}[Y] - \{\mathbb{COV}[X, Y]\}^2 \ge 0$$

を得る．したがって，

$$\frac{\{\mathbb{COV}[X, Y]\}^2}{\mathbb{VAR}[X]\mathbb{VAR}[Y]} \le 1 \Longleftrightarrow \left|\frac{\mathbb{COV}[X, Y]}{\sqrt{\mathbb{VAR}[X]}\sqrt{\mathbb{VAR}[Y]}}\right| \le 1$$

となり，(i) は示せた．
(ii) 上の議論から $|\rho[X, Y]| = 1$ となるための必要十分条件は

$$\mathbb{VAR}[X]\mathbb{VAR}[Y] - \{\mathbb{COV}[X, Y]\}^2 = 0 \Longleftrightarrow t \text{ の二次方程式 } h(t) = \text{は単根である．}$$

である．この解を t_0 とする．しかし，$\mathbb{E}[(X - \mu_X)t_0 + (Y - \mu_Y)] = 0$ に注意すれば，

$$0 = h(t_0) = \mathbb{E}[\{(X - \mu_X)t_0 + (Y - \mu_Y)\}^2] = \mathbb{VAR}[(X - \mu_X)t_0 + (Y - \mu_Y)]$$

となり，分散の性質 (iii) から

$$P((X - \mu_X)t_0 + (Y - \mu_Y) = 0) = 1$$

となる．したがって，$a = -t_0, b = \mu_X t_0 + \mu_Y$ とおけば，$P(aX + b = Y) = 1$ が成立することがわかる．また，

$$h(t) = \mathbb{VAR}[X]\left(t - \frac{\mathbb{COV}[X, Y]}{\mathbb{VAR}[X]}\right)^2 = 0$$

に注意すれば，方程式

$$h(t) = \mathbb{VAR}[X]\left(t + \frac{\mathbb{COV}[X, Y]}{\mathbb{VAR}[X]}\right)^2 = 0$$

の解は

$$t = t_0 = -\frac{\mathbb{COV}[X, Y]}{\mathbb{VAR}[X]}$$

なので，

$$a = \frac{\mathbb{COV}[X, Y]}{\mathbb{VAR}[X]}$$

となる．したがって，a の符号と $\rho[X, Y]$ の符号は同じになる．よって，(ii) は示された． □

3.5　2 次元の確率変数の変換

(X, Y) を 2 次元の確率ベクトルとし，実数値関数 $g_1(x, y), g_2(x, y)$ によって定められる新たな確率ベクトル $(U, V) = (g_1(X, Y), g_2(X, Y))$ の分布を求めることを考える．$\mathbb{R}^2$ の任意の部分集合 B に対して，$\mathbb{R}^2$ の部分集合 A を

$$A = \{(x, y) \in \mathbb{R}^2 : (g_1(x, y), g_2(x, y)) \in B\}$$

で定めると

$$\mathbb{P}((U, V) \in B) = \mathbb{P}((X, Y) \in A)$$

となる．したがって，固定された g_1, g_2 に対して，(U, V) の分布は (X, Y) の分布のみに依存することがわかる．

3.5.1　離散型確率ベクトルの場合

(X, Y) が離散型確率ベクトルの場合をまず考えよう．(X, Y) の同時確率関数を $f_{X,Y}(x, y)$ とし，(U, V) の同時確率密度関数 $f_{U,V}(u, v)$ を求めよう．そのために，$S = \{(x, y) \in \mathbb{R}^2 : f_{X,Y}(x, y) > 0\}$ とし，

$$T = \{(u, v) \in \mathbb{R}^2 : \text{ある } (x, y) \in S \text{ に対して，} \quad u = g_1(x, y), v = g_2(x, y)\}$$

とする．すると $(u, v) \in T$ ならば，$f_{U,,V}(u, v) > 0$ となり，T は可算集合となることに注意する．任意の $(u, v) \in T$ に対して，

$$S_{uv} = \{(x, y) \in S : g_1(x, y) = u, g_2(x, y) = v\}$$

とおけば，

$$f_{U,V}(u, v) = \mathbb{P}(U = u, V = v) = \mathbb{P}((X, Y) \in S_{uv}) = \sum_{(x, y)\in S_{uv}} f_{X,Y}(x, y)$$

となる．

定理 3.9　X と Y は独立に平均 $\theta\,(\theta > 0)$ と $\lambda\,(\lambda > 0)$ のポアソン分布に従うとする．このとき，$X + Y$ は平均 $\theta + \lambda$ のポアソン分布に従う．

証明　仮定より (X, Y) の同時確率関数は

$$f_{X,Y}(x, y) = \frac{\theta^x e^{-\theta}}{x!}\frac{\lambda^y e^{-\lambda}}{y!}, \qquad x = 0, 1, 2, \ldots, y = 0, 1, 2, \ldots$$

となる．したがって，$S = \{0, 1, 2, \ldots\} \times \{0, 1, 2, \ldots\}$ である．

いま，$U = X + Y, V = Y$ とおく．すなわち，$g_1(x, y) = x + y, g_2(x, y) = y$ である．$y = v, x = u - v$ から $v = 0, 2, \ldots$ かつ $u - v = 0, 1, \ldots$ から $u = v, v + 1, v + 2, \ldots$ を得る．したがって，

$$T = \{v = 0, 1, 2, \ldots; u = v, v + 1, v + 2, \ldots\} = \{u = 0, 1, 2, \ldots; v = 0, 1, \ldots, u\}$$

となる．任意の $(u, v) \in T$ に対して，$S_{uv} = \{(u, u - v)\}$ となるので，

$$f_{U,V}(u, v) = f_{X,Y}(u, u - v) = \frac{\theta^{u-v} e^{-\theta}}{(u - v)!}\frac{\lambda^v e^{-\lambda}}{v!},$$

をえる．U の周辺確率関数を求めるために，固定した非負整数 u を考える．

$$\{U = u\} = \cup_{v: f_{U,V}(u,v)>0}\{U = u, V = v\}$$

と

$$\{v : f_{U,V}(u, v) > 0\} = \{0, 1, \ldots, u\}$$

となる．よって，

$$\begin{aligned} f_U(u) &= \mathbb{P}(U = u) = \mathbb{P}(\cup_{v=0}^{u}\{U = u, V = v\}) = \sum_{v=0}^{n}\mathbb{P}(U = u, V = v) = \sum_{v=0}^{n} f_{U,V}(u, v) \\ &= \sum_{v=0}^{n}\frac{\theta^{u-v}e^{-\theta}}{(u-v)!}\frac{\lambda^v e^{-\lambda}}{v!} = e^{-(\theta+\lambda)}\sum_{v=0}^{n}\frac{\theta^{u-v}}{(u-v)!}\frac{\lambda^v}{v!} \\ &= \frac{e^{-(\theta+\lambda)}}{u!}\sum_{v=0}^{n}\begin{pmatrix} u \\ v \end{pmatrix}\theta^{u-v}\lambda^v = \frac{e^{-(\theta+\lambda)}(\theta+\lambda)^u}{u!} \end{aligned}$$

となり，定理は証明された． □

3.5.2 連続型の場合

(X, Y) を連続型確率変数とし，同時確率密度関数 $f_{X,Y}(x, y)$ を持つとする．前節と同様，確率ベクトル (U, V) は $U = g_1(X, Y)$, $V = g_2(X, Y)$ で定義され，

$$\begin{aligned} S &= \{(x, y) \in \mathbb{R}^2 : f_{X,Y}(x, y) > 0\} \\ T &= \{(u, v) \in \mathbb{R}^2 : \text{ある } (x, y) \in S \text{ に対して，} \quad u = g_1(x, y), v = g_2(x, y)\} \end{aligned}$$

であったことを思い出そう．議論を簡単にするために，g_1, g_2 は一対一対応の変換と仮定し，逆変換を

$$x = h_1(u, v), \qquad y = h_2(u, v)$$

と書くことにする．さらに，

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

と(3-7)おく．ただし，

$$\frac{\partial x}{\partial u} = \frac{\partial h_1(u, v)}{\partial u}, \quad \frac{\partial x}{\partial v} = \frac{\partial h_1(u, v)}{\partial v}, \quad \frac{\partial y}{\partial u} = \frac{\partial h_2(u, v)}{\partial u}, \quad \frac{\partial x}{\partial v} = \frac{\partial h_2(u, v)}{\partial v},$$

である．ヤコビアン J は恒等的にゼロではないとする．このとき，

$$f_{U,V} = \begin{cases} f_{X,Y}(h_1(u, v), h_2(u, v))|J|, & (u, v) \in T \\ 0, & (u, v) \notin T \end{cases}$$

となる．これは，$(u, v) \in T$ に対して，重積分の変数変換の公式を使えば，

$$\begin{aligned} \frac{\partial^2}{\partial u \partial v}\mathbb{P}(U \le u, V \le v) &= \frac{\partial^2}{\partial u \partial v}\int\int_{g_1(x,y)\le u,\, g_2(x,v)\le v} f_{X,Y}(x, y)\,dx\,dy \\ &= \frac{\partial^2}{\partial u \partial v}\int_{-\infty}^{u}\int_{-\infty}^{v} f_{X,Y}(h_1(s, t), h_2(s, t))|J|\,ds\,dt \end{aligned}$$

となることからわかる．

例 3.8 X と Y は独立に標準正規分布に従っていると仮定し，変換

$$U = X + Y, \qquad V = X - Y$$

を考える．したがって，$g_1(x, y) = x + y$, $g_2(x, y) = x - y$ である．(X, Y) の同時確率密度関数は

$$f_{X,Y} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

であるので，$S = \mathbb{R}^2$ となる．また，g_1, g_2 は一対一対応で逆変換は

$$x = h_1(u, v) = \frac{u+v}{2}, \qquad y = h_2(u, v) = \frac{u-v}{2}$$

となる．したがって，$T = \mathbb{R}^2$ となる．さらに，

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

となるので，(U, V) の同時確率密度関数は

$$\begin{aligned} f_{U,V}(u, v) &= f_{X,Y}(h_1(u, v), h_2(u, v))|J| \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \times \frac{(u+v)^2}{4}\right) \exp\left(-\frac{1}{2} \times \frac{(u-v)^2}{4}\right) \frac{1}{2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-u^2/4}\right)\left(\frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-v^2/4}\right) \end{aligned}$$

となる．これと補題 3.1 から U と V は独立となる．

定理 3.10 X, Y は独立な確率変数とし，$g_1(x)$ は x のみの関数で $g_2(y)$ は y のみの関数とする．このとき，$U = g_1(X)$ と $V = g_2(Y)$ も独立である．

証明 U, V は連続型の場合を示すことにする．任意の $u \in \mathbb{R}$, $v \in \mathbb{R}$ に対して，

$$S_u = \{x \in \mathbb{R} : g_1(x) \le u\}, \qquad S_v = \{y \in \mathbb{R} : g_2(y) \le v\},$$

とおく．このとき，(U, V) の同時分布関数は

$$\begin{aligned} F_{U,V}(u, v) &= \mathbb{P}(U \le u, V \le v) \\ &= \mathbb{P}(X \in S_u, Y \in S_v) = \mathbb{P}(X \in S_u)\mathbb{P}(Y \in S_v) \end{aligned}$$

となる．これより (U, V) の同時確率密度関数は

$$f_{U,V}(u, v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u, v) = \left(\frac{d}{du}\mathbb{P}(X \in S_u)\right)\left(\frac{d}{dv}\mathbb{P}(Y \in S_v)\right)$$

となる．したがって，補題 3.1 から定理は証明された． □

3.6 多次元分布の代表的なモデル

3.6.1 二変量正規分布

定義 3.12 確率変数 (X, Y) がつぎの同時確率密度関数をもつとき，(X, Y) は母数 $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, $0 < \sigma_X < \infty$, $0 < \sigma_Y < \infty$, $-1 < \rho < 1$ の**二変量正規分布**に従うとする．

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}Q(x, y)\right\} \qquad -\infty < x < \infty,\ -\infty < y < \infty.$$

ただし，

$$Q(x, y) = \frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$$

である．

命題 3.6 (二変量正規分布の性質) (i) すべての x, y で $f_{X,Y}(x, y) \geq 0$ である．
(ii) $\int\int f_{X,Y}(x, y)\,dx\,dy = 1$ である．
(iii) X の周辺確率密度関数は

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$$

である．
(iv) (X, Y) の積率母関数は

$$M_{X,Y}(s, t) = \exp\left[s\mu_X + t\mu_Y + \frac{1}{2}(s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2)\right] \qquad -\infty < s, t < \infty,$$

となる．
(v) X と Y の一次と二次の積率は以下のようになる：

$$\mathbb{E}[X] \;=\; \mu_X, \qquad \mathbb{E}[X^2] = \sigma_X^2 + \mu_X^2, \qquad \mathbb{E}[XY] = \rho\sigma_X\sigma_Y + \mu_X\mu_Y$$

となる．これらより

$$\begin{aligned}\mathbb{VAR}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma_X^2, \\ \mathbb{COV}[X, Y] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \rho\sigma_X\sigma_Y\end{aligned}$$

を得る．

証明 (i) は明らか．(ii) を示すために，

$$u = \frac{x-\mu_X}{\sigma_X}, \qquad v = \frac{y-\mu_Y}{\sigma_Y}$$

とおく．すると

$$Q(x, y) = u^2 - 2\rho uv + v^2 = (u-\rho v)^2 + (1-\rho^2)v^2$$

となる．さらに，

$$w = \frac{u-\rho v}{\sqrt{1-\rho^2}}$$

とおけば

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,\,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left\{-\frac{(u-\rho v)^2}{2(1-\rho^2)} - \frac{v^2}{2}\right\} du\,dv$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) dw \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) dv$$
$$= 1$$

となる．

(iii) X の周辺確率密度関数は

$$f_X(x) = \int_{-\infty}^{\infty} f(x,\,y)\,dy$$
$$= \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2(1-\rho^2)}\left(v-\rho\frac{x-\mu_X}{\sigma_X}\right)^2\right] dv$$

となり，さらに

$$w = \frac{1}{\sqrt{1-\rho^2}}\left(v-\rho\frac{x-\mu_X}{\sigma_X}\right)$$

とおけば，

$$f_X(x) = \frac{1}{2\pi\sigma_X}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]\int_{-\infty}^{\infty} e^{-w^2/2}\,dw = \frac{1}{2\pi\sigma_X}\exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]$$

を得る．

(iv) $(X,\,Y)$ の積率母関数は

$$M_{X,Y}(s,\,t) = \mathbb{E}[\exp(sX+tY)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(sx+ty) f_{X,Y}(x,\,y)\,dx\,dy$$
$$= \exp\left[s\mu_X + t\mu_Y + \frac{1}{2}(s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2)\right]$$

であること示す．ただし，$-\infty < s,\,t < \infty$ である．

まず，

$$u = \frac{x-\mu_X}{\sigma_X}, \qquad v = \frac{y-\mu_Y}{\sigma_Y}$$

とおく．すると

$$m_{X,Y}(s,\,t) = \iint_{\mathbb{R}^2} \exp(sx+ty) f_{X,Y}(x,\,y)\,dx\,dy$$
$$= \exp\{\mu_X s + \mu_Y t\} \iint_{\mathbb{R}^2} \exp\{s\sigma_X u + t\sigma_Y v\} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{(u-\rho v)^2}{2(1-\rho^2)} - \frac{v^2}{2}\right\} du\,dv$$

となる．ここで

$$w = \frac{u-\rho v}{\sqrt{1-\rho^2}}$$

とおけば,

$$\begin{aligned}
&\exp\{-(\mu_X s+\mu_Y t)\}M_{X,Y}(s,\,t)\\
&=\iint_{\mathbb{R}^2}\exp\{s\sigma_X(\sqrt{1-\rho^2}w+\rho v)+t\sigma_Y v\}\frac{1}{2\pi}\exp\left\{-\frac{w^2}{2}-\frac{v^2}{2}\right\}dw\,dv\\
&=\frac{1}{2\pi}\iint_{\mathbb{R}^2}\exp\left\{-\frac{1}{2}\left(w-s\sigma_X\sqrt{1-\rho^2}\right)^2-\frac{1}{2}\left(v-(s\rho\sigma_X+t\sigma_Y)\right)^2\right\}dw\,dv\\
&\quad\times\frac{1}{2}\exp\left\{\left(s^2\sigma_X^2(1-\rho^2)+(s\rho\sigma_X+t\sigma_Y)^2\right)\right\}\\
&=\exp\left\{\frac{1}{2}(s^2\sigma_X^2+2st\rho\sigma_X\sigma_Y+t^2\sigma_Y^2)\right\}
\end{aligned}$$

よりわかる．最後の等号は

$$\frac{1}{2\pi}\iint_{\mathbb{R}^2}\exp\left\{-\frac{1}{2}\left(w-s\sigma_X\sqrt{1-\rho^2}\right)^2-\frac{1}{2}\left(v-(s\rho\sigma_X+t\sigma_Y)\right)^2\right\}dw\,dv=1$$

は左辺の被積分関数は平均 $N(s\sigma_X\sqrt{1-\rho^2},\,1)$ と $N(s\rho\sigma_X+t\sigma_Y,\,1)$ に独立に従うふたつの確率変数の同時確率密度関数であることに注意すればよい.

(v)　X と Y の一次と二次の積率を求める：

$$\begin{aligned}
\mathbb{E}[X] &= \frac{\partial}{\partial s}m_{X,Y}(s,\,t)\bigg|_{s=0}=\mu_X,\\
\mathbb{E}[X^2] &= \frac{\partial^2}{\partial s^2}m_{X,Y}(s,\,t)\bigg|_{s=0}=\sigma_X^2+\mu_X^2,\\
\mathbb{E}[XY] &= \frac{\partial^2}{\partial s\partial t}m_{X,Y}(s,\,t)\bigg|_{s=0,\,t=0}=\rho\sigma_X\sigma_Y+\mu_X\mu_Y
\end{aligned}$$

となる．これらより

$$\begin{aligned}
\mathbb{VAR}[X] &= \mathbb{E}[X^2]-(\mathbb{E}[X])^2=\sigma_X^2,\\
\mathbb{COV}[X,\,Y] &= \mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]=\rho\sigma_X\sigma_Y
\end{aligned}$$

を得る. □

注意 3.7 ここで

$$\Sigma=\begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}$$

とおけば,

$$Q(x,\,y)=(x-\mu_X,\,y-\mu_Y)\Sigma^{-1}\begin{pmatrix}x-\mu_X\\ y-\mu_Y\end{pmatrix},\qquad |\Sigma|=\sigma_X^2\sigma_Y^2(1-\rho^2)$$

となり,

$$f_{X,Y}(x,\,y)=\frac{1}{2\pi|\Sigma|^{1/2}}\exp\left[-\frac{1}{2}(x-\mu_X,\,y-\mu_Y)\Sigma^{-1}\begin{pmatrix}x-\mu_X\\ y-\mu_Y\end{pmatrix}\right]$$

3.7 確率不等式 2

定理 3.11 (ヘルダーの不等式) 正数 p, q は $1/p+1/q=1$ をみたすとする．$\mathbb{E}[|X|^p]<\infty, \mathbb{E}[|Y|^q]<\infty$ なる確率変数 X, Y を確率変数に対して，

$$\mathbb{E}[|XY|] \leq \{\mathbb{E}[|X|^p\}^{1/p}\{\mathbb{E}[|Y|^q\}^{1/q}$$

が成立する．ときに，$p=2, q=2$ のときの不等式

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[|X|^2]}\sqrt{\mathbb{E}[|Y|^2]}$$

をシュバルツの不等式という．

証明 まず，任意の正の数 a, b に対して，

$$\frac{1}{p}a^p + \frac{1}{b}b^q \geq ab \tag{3.3}$$

が成立することを示す．ただし，等号成立は $a^p = b^q$ の時に限る．b を固定して，

$$g(a) = \frac{1}{p}a^p + \frac{1}{q}b^q - ab$$

とおき，$g(a)$ を a に関して最小化する：

$$\frac{d}{da}g(a) = a^{p-1} - b = 0 \Longleftrightarrow b = a^{p-1}$$

となり，2 次の導関数を確認すれば，$a = b^{1/(p-1)}$ のとき，最小となる．したがって，

$$g(a) \geq g(b^{1/(p-1)}) = \frac{1}{p}b^{p/(p-1)} + \frac{1}{q}b^q - b^{1/(p-1)}b = \frac{1}{p}b^q + \frac{1}{q}b^q - b^q = 0$$

となる．最後から 2 番目の等号は $p/(p-1) = q$ よりわかる．

次に，

$$a = \frac{|X|}{(\mathbb{E}[|X|^p])^{1/p}}, \qquad b = \frac{|Y|}{(\mathbb{E}[|Y|^q])^{1/q}},$$

として，(3.3) を用いてば，

$$\frac{1}{p}\frac{|X|^p}{\mathbb{E}[|X|^p]} + \frac{1}{q}\frac{|Y|^q}{\mathbb{E}[|Y|^q]} \geq \frac{|XY|}{(\mathbb{E}[|X|^p])^{1/p}(\mathbb{E}[|Y|^q])^{1/q}}$$

を得る．この両辺の期待値を取れば，定理は証明された． □

例 3.9 X と Y を確率変数とし，$\mathbb{E}[X^2]<\infty, \mathbb{E}[Y^2]<\infty$ とする．このとき，$-\infty<\mu_X=\mathbb{E}[X], \mu_Y=\mathbb{E}[Y]<\infty$ がただちにわかる．さらに，Cauchy-Schwarz の不等式から

$$|\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]| \leq \mathbb{E}[|(X-\mu_X)(Y-\mu_Y)|] \leq \sqrt{\mathbb{E}[(X-\mu_X)^2\}}\sqrt{\mathbb{E}[(Y-\mu_Y)^2\}}$$

より

$$(\mathbb{COV}[X, Y])^2 \leq \mathbb{VAR}[X]\mathbb{VAR}[Y]$$

を得る．したがって，

$$|\rho[X, Y]| \leq 1 \Longleftrightarrow -1 \leq \rho[X, Y] \leq 1$$

を得る．