

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.

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

ただし,

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

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

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

となる.

(v) X と Y の一次と二次の積率は以下ようになる:

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

となる. これらより

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

を得る.

証明 (i) は明らか. (ii) を示すために,

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

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

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

となる.

(iii) X の周辺確率密度関数は

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

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

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

からわかる.

(v) X と Y の一次と二次の積率を求める:

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

となる. これらより

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

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,

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

となり,

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