2.6 変数変換

確率変数を $X_1, \ldots, X_p$ とし，その pdf を

$$f(x_1, \ldots, x_p) \quad x_1, \ldots, x_p \in \mathbb{R}$$

とする．いま，

$$y_i = y_i(x_1, \ldots, x_p) \quad (i = 1, \ldots, p)$$

と定義する．ここで，逆変換が存在することを仮定して，

$$x_i = x_i(y_1, \ldots, y_p) \quad (i = 1, \ldots, p)$$

とする．このとき，

$$Y_i = y_i(X_1, \ldots, X_p) \quad (i = 1, \ldots, p)$$

と定義すると確率変数 $Y_1, \ldots, Y_p$ の pdf は

$$g(y_1, \ldots, y_p) = f(x_1(y_1, \ldots, y_p), \ldots, x_p(y_1, \ldots, y_p))J(y_1, \ldots, y_p)$$

ただし，

$$J(y_1, \ldots, y_p) = \mathrm{mod} \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_p} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_p}{\partial y_1} & \cdots & \dfrac{\partial x_p}{\partial y_p} \end{vmatrix}$$

であり，**ヤコビアン**（Jacobian）という．ここに，$|\cdot|$ は行列式，mod は絶対値を表す．

例

確率変数 X_1, X_2 の pdf が

$$f(x_1, x_2) = \begin{cases} 4x_1x_2 & (0 < x_1 < 1, 0 < x_2 < 1) \\ 0 & (else) \end{cases}$$

ならば，

$$Y_1 = \frac{X_1}{X_2}, \quad Y_2 = X_1X_2$$

としたとき，Y_1, Y_2 の確率密度関数 $g(y_1, y_2)$ を求めよ．

逆関数を求めると，

$$x_1 = (y_1 y_2)^{\frac{1}{2}}, \quad x_2 = \left(\frac{y_2}{y_1}\right)^{\frac{1}{2}}$$

であるから，

$$\frac{\partial x_1}{\partial y_1} = \frac{1}{2}\left(\frac{y_2}{y_1}\right)^{1/2}, \quad \frac{\partial x_1}{\partial y_2} = \frac{1}{2}\left(\frac{y_1}{y_2}\right)^{1/2}$$
$$\frac{\partial x_2}{\partial y_1} = -\frac{1}{2}\left(\frac{y_2}{y_1^3}\right)^{1/2}, \quad \frac{\partial x_2}{\partial y_2} = \frac{1}{2}\left(\frac{1}{y_1 y_2}\right)^{1/2}$$

より，y_1 が正であることに注意して

$$J(y_1, y_2) = \frac{1}{2y_1} \ (\neq 0)$$

を得る．$x_1 x_2$ 平面の図形を

$$S = \{(x_1, x_2) | 0 < x_1 < 1, 0 < x_2 < 1\}$$

とおく．$y_1 = x_1/x_2$，$y_2 = x_1 x_2$ によって，図形 S が $y_1 y_2$ 平面の図形

$$T = \left\{(y_1, y_2) \mid y_1 > 0, y_2 > 0, y_1 y_2 < 1, \frac{y_2}{y_1} < 1\right\}$$

に 1 対 1 に対応するから，

$$g(y_1, y_2) = 2\frac{y_2}{y_1} \quad (y_1, y_2) \in T$$

である．

3 多変量正規分布

多変量正規分布を導入する前に，1 変量の正規分布を復習する．確率変数 X が平均 μ，分散 σ^2 の正規分布に従うとする．このとき，X の pdf は

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \quad -\infty < x < \infty$$

である．ここで

$$\int_{-\infty}^{\infty} f(x;\mu,\sigma^2)dx = 1$$

を示す．

変数変換 $y=(x-\mu)/\sigma$ を用いると

$$\int_{-\infty}^{\infty} f(x;\mu,\sigma^2)dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}y^2\right]dy$$

を得る．ここで，

$$I = \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}y^2\right]dy$$

とおくと

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}y^2\right]dy\right)\left(\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}t^2\right]dt\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[-\frac{y^2+t^2}{2}\right]dydt \end{aligned}$$

であり，極座標変換 $y=r\cos\theta$，$t=r\sin\theta$ のヤコビアンは，

$$J(r,\theta) = \boxed{\text{読者の演習}}$$

である．したがって，

$$\begin{aligned} I^2 &= \boxed{\text{読者の演習}} \\ &= \int_0^{2\pi}\int_0^{\infty} r\exp\left[-\frac{r^2}{2}\right]drd\theta \\ &= \boxed{\text{読者の演習}} \\ &= 2\pi \end{aligned}$$

より，$I=\sqrt{2\pi}$ を得る．以上から，

$$\int_{-\infty}^{\infty} f(x;\mu,\sigma^2)dx = \frac{1}{\sqrt{2\pi}}I = 1$$

である．

確率ベクトル（random vector）$\boldsymbol{X} = (X_1, \ldots, X_p)'$ が次の pdf をもつとき，$\boldsymbol{X}$ は**平均ベクトル**（mean vector）$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$，**共分散行列**（covariance matrix）$\boldsymbol{\Sigma} = (\sigma_{ij})_{p \times p}$ の多変量正規分布に従うという．任意の $\boldsymbol{x} = (x_1, \ldots, x_p)' \in \mathbb{R}^p$ に対して，

$$f(x_1, \ldots, x_p; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \tag{1}$$

後に詳細を説明するが，共分散行列 $\boldsymbol{\Sigma}$ はその名の通り対角成分に分散 $V(X_i) = \sigma_{ii}$，非対角成分に共分散 $Cov(X_i, X_j) = \sigma_{ij}$ を要素として持つ行列である．このことから，$\boldsymbol{\Sigma}$ が実対称行列であることがわかる．また，共分散行列 $\boldsymbol{\Sigma}$ は**正定値**（positive definite）であることを仮定する．

正定値

$n \times n$ 実対称行列 A が正定値であるとは，$\boldsymbol{0}$ でない全てのベクトル $\boldsymbol{x}$ に対して，二次形式 $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ が必ず正となることである．
正定値行列の同値条件

(i) $\boldsymbol{0}$ でないすべてのベクトル $\boldsymbol{x}$ に対して，$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$

(ii) $\boldsymbol{A}$ のすべての固有値 λ_i が $\lambda_i > 0$ を満たす

(iii) $\boldsymbol{A} = \boldsymbol{W}'\boldsymbol{W}$ となる正則行列 $\boldsymbol{W}$ が存在する

(1) が pdf の条件を満たすことを確認する．共分散行列 $\boldsymbol{\Sigma}$（対称行列）が正定値であることを仮定するので，

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \geq 0 \tag{2}$$

であるから，

$$0 \leq \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] \leq 1$$

であることが確認できる．また，正定値の同値条件 (iii) から

$$|\boldsymbol{\Sigma}| = \boxed{\text{読者の演習}} > 0$$

を得る．以上から，任意の $\boldsymbol{x} \in \mathbb{R}^p$ に対して $f(x_1, \ldots, x_p; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \geq 0$ である．

3.1 演習問題

問 1 確率変数 X_1, X_2 の pdf が

$$f(x_1, x_2) = \begin{cases} 4x_1x_2 & (0 < x_1 < 1, 0 < x_2 < 1) \\ 0 & (else) \end{cases}$$

ならば,

$$Y_1 = X_1 + X_2, \quad Y_2 = X_1 - X_2$$

としたとき，Y_1, Y_2 の pdf は $0 < y_1 + y_2 < 2$, $0 < y_1 - y_2 < 2$ に対して（ア）である．文中の（ア）に当てはまるものとして，次の ① ～ ④ のうちから適切なものを一つ選べ． $\boxed{\mathbf{1}}$

① $y_1 - y_2$　② $y_1^2 - y_2^2$　③ $(y_1 - y_2)/2$　④ $(y_1^2 - y_2^2)/2$

問 2 本文中の式 (2) における等号成立条件は,（イ）である．文中の（イ）に当てはまるものとして，次の ① ～ ④ のうちから適切なものを一つ選べ． $\boxed{\mathbf{2}}$

① $\boldsymbol{x} = \mathbf{0}$　② $\boldsymbol{x} = \mathbf{1}$　③ $\boldsymbol{x} = \boldsymbol{\mu}$　④ $\boldsymbol{x} = \boldsymbol{\mu}/2$

問 3 確率変数 X が平均 μ，分散 σ^2 の正規分布に従うとき，次の記述 I ～ III を考えた．

I. パラメータ μ は分布の期待値である．

II. パラメータ μ は分布のメジアンである．

III. パラメータ μ は分布のモードである．

記述 I～ III に関して，次の ① ～ ④ のうちから最も適切なものを一つ選べ． $\boxed{\mathbf{3}}$

① I のみ正しい．　② I と II のみ正しい．

③ I と III のみ正しい．　④ I と II と III はすべて正しい．