# 第 6 章「標 本 分 布」

## １．確率変数の関数の分布

**定理１** 確率変数 $X$ が密度関数 $f(x)$ をもつとする．関数 $h(x)$ が微分可能で狭義単調関数のとき，$Y = h(X)$ の密度関数 $g(y)$ は，

$$\begin{aligned} g(y) &= f(x) \left| \frac{dx}{dy} \right| \\ &= f(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right| \end{aligned}$$

**（証明）** $h(x)$ が単調減少の場合に証明する．

$$G(y) = P(Y \leq y) = P(h(X) \leq y) = P(X \geq h^{-1}(y)) = \int_{h^{-1}(y)}^{\infty} f(x)\, dx$$

あとは，最左辺と最右辺を $y$ で微分すればよい．

**例：** $X = \log Y \sim \mathrm{N}(\mu, \sigma^2)$ のとき，$Y$ の密度関数（対数正規分布）を求めよ．
**（解）** $f(x) = \exp(-(x-\mu)^2/(2\sigma^2))/(\sqrt{2\pi}\,\sigma)$, $x = \log y$, $dx/dy = 1/y\ (> 0)$ であるから，

$$g(y) = f(\log y)\,\frac{1}{y} = \frac{1}{\sqrt{2\pi}\,\sigma}\,\frac{1}{y}\,\exp\left(-\frac{(\log y - \mu)^2}{2\sigma^2}\right)$$

**問：** 確率変数 $X$ が分布関数 $F(x)$，密度関数 $f(x)$ をもつとする．このとき，$Y = F(X)$ は一様分布 $U[0,1]$ に従うことを示せ．

**・単調でない場合** 確率変数 $X$ が $\mathrm{N}(0,\,1)$ に従うとき，$Y = X^2$ の密度関数 $g(y)$ は，次のように求めることができる．非負の $y$ に対して，

$$\begin{aligned} G(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= 2\,P(0 \leq X \leq \sqrt{y}) = 2\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}}\,\exp(-x^2/2)\,dx \end{aligned}$$

最初と最後の表現を $y$ で微分すれば，

$$g(y) = 2\,(\sqrt{y})'\,\frac{1}{\sqrt{2\pi}\,exp(-y/2)} = \frac{1}{\sqrt{2\pi}}\,y^{-1/2}\,e^{-y/2}$$

となる．これは，ガンマ分布 $G_A(1/2, 2)$ の密度関数（$= \chi^2(1)$ の密度関数）である．

## ２．確率変ベクトルの関数の分布

**定理２** $n$ 次元確率ベクトル $\mathbf{X}$ が密度関数 $f(\mathbf{x})$ をもつとする．また，関数 $\mathbf{y} = \mathbf{h}(\mathbf{x})$ は $n$ 次元ベクトル値関数で，$\mathbf{x}$ から $\mathbf{y}$ への変換のヤコビアン行列 $\partial\mathbf{x}/\partial\mathbf{y}$ が正則であるとする．このとき，確率ベクトル $\mathbf{Y} = \mathbf{h}(\mathbf{X})$ の密度関数 $g(\mathbf{y})$ は，

$$g(\mathbf{y}) = f(\mathbf{x})\,\left| det\left(\frac{\partial\mathbf{x}}{\partial\mathbf{y}}\right)\right|$$

で与えられる．

**例：** 2 つの確率変数 $X_1, X_2$ が互いに独立で，それぞれ一様分布 $U[0,1]$ に従うとき，次の確率変数 $Y_1, Y_2$ は互いに独立に標準正規分布 N(0, 1) に従う．（この変換を Box-Muller 変換という．）

$$Y_1 = \sqrt{-2\log X_1}\,\cos(2\pi X_2)$$
$$Y_2 = \sqrt{-2\log X_1}\,\sin(2\pi X_2)$$

**（解）** $X_1, X_2$ について解くと，

$$\begin{aligned} X_1 &= \exp\left(-\frac{Y_1^2+Y_2^2}{2}\right), \\ X_2 &= \frac{1}{2\pi}\tan^{-1}\frac{Y_2}{Y_1} \end{aligned}$$

となる．また，ヤコビアンは，

$$det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right) = \begin{vmatrix} -y_1\,\exp\left(-\frac{y_1^2+y_2^2}{2}\right) & -y_2\,\exp\left(-\frac{y_1^2+y_2^2}{2}\right) \\ \frac{1}{2\pi}\frac{-y_2}{y_1^2+y_2^2} & \frac{1}{2\pi}\frac{y_1}{y_1^2+y_2^2} \end{vmatrix} = -\frac{1}{2\pi}\exp\left(-\frac{y_1^2+y_2^2}{2}\right) \neq 0$$

したがって，

$$g(y_1,y_2) = f(x_1,x_2)\left|det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right)\right| = \frac{1}{2\pi}\exp\left(-\frac{y_1^2+y_2^2}{2}\right) = \phi(y_1)\,\phi(y_2)$$

となる．ここで，$\phi(z)$ は N(0, 1) の密度関数である．

**例：** $X_1$ と $X_2$ が互いに独立に N(0, 1) に従うとき，$X_1/X_2$ はコーシー分布に従うことを示せ．

**（解）** 次の変数変換

$$Y_1 = X_1/X2, \quad Y_2 = X_2 \qquad \Leftrightarrow \qquad X_1 = Y_1 Y_2, \quad X_2 = Y_2$$

を考えると，ヤコビアンは，

$$det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right) = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = y_2 \neq 0 \qquad \text{with probability 1}$$

となるから，

$$g(y_1,y_2) = f(x_1,x_2)\,|y_2| = \frac{1}{2\pi}\exp\left(-\frac{(y_1^2+1)}{2}\,y_2^2\right)|y_2|$$

を得る．したがって，$g(y_1,y_2)$ を $y_2$ で積分して次の結果を得る．

$$\begin{aligned} g(y_1) &= \int_{-\infty}^{\infty} g(y_1,y_2)\,dy_2 = 2\int_0^{\infty} g(y_1,y_2)\,dy_2 = \frac{1}{\pi}\int_0^{\infty} y_2\,\exp\left(-\frac{(y_1^2+1)}{2}\,y_2^2\right)dy_2 \\ &= \pi\left[\frac{-1}{y_1^2+1}\exp\left(-\frac{(y_1^2+1)}{2}\,y_2^2\right)\right]_0^{\infty} = \frac{1}{\pi}\frac{1}{y_1^2+1} \end{aligned}$$

**問 4** $X_1$ と $X_2$ が互いに独立にガンマ分布 $G_A(\alpha_1, \beta), G_A(\alpha_2, \beta)$ に従うとき，$Y = X_1/(X_1+X_2)$ はベータ分布 $B_E(\alpha_1, \alpha_2)$ に従うことを示せ.

## ３．正規分布から誘導される分布

**・$\chi^2$ 分布**

$Z_1, Z_2, \cdots, Z_n$ が互いに独立に N(0, 1) に従うとき，

$$Y = Z_1^2 + Z_2^2 + \cdots + Z_n^2$$

は，自由度 $n$ の $\chi^2$ 分布に従うといい，$Y \sim \chi^2(n)$ と書く．$Y$ の密度関数を求めるには，まず，$Z_1^2$ の積率母関数は,

$$\begin{aligned} M_{Z_1^2}(t) &= E(\exp(tZ_1^2)) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(1-2t)\,z^2\right) d\,z \\ &= (1-2t)^{-1/2} \end{aligned}$$

となる．したがって，$Y$ の積率母関数は,

$$\begin{aligned} M_Y(t) &= E(\exp(tY)) = E(\exp(t(Z_1^2 + Z_2^2 + \cdots + Z_n^2))) = E(\exp(tZ_1^2)) \cdots E(\exp(tZ_n^2)) \\ &= \prod_{j=1}^{n} E(\exp(tZ_j)) = (1-2t)^{-n/2} \end{aligned}$$

となるので，$Y$ はガンマ分布 $G_A(n/2, 2)$ に従うから，$Y$ の密度関数は

$$f(y) = \frac{1}{2\,\Gamma(n/2)} \left(\frac{y}{2}\right)^{n/2-1} \exp(-y/2)$$

$$E(Y) = n, \qquad V(Y) = 2\,n$$

**・t 分布**

$X$ と $Y$ は互いに独立に，それぞれ N(0, 1)，$\chi^2(n)$ に従うとき，

$$T = \frac{X}{\sqrt{Y/n}}$$

は，自由度 $n$ の $t$ 分布に従うといい，$T \sim t(n)$ と書く．$T$ の密度関数を求めるには，変数変換

$$t = x/\sqrt{y/n}, \quad u = y \qquad \Leftrightarrow \qquad x = t\,\sqrt{u/n}, \quad y = u$$

を考える．変換のヤコビアンは,

$$det\left(\frac{\partial(x,y)}{\partial(t,u)}\right) = \begin{vmatrix} \sqrt{u/n} & t/(2\sqrt{nu}) \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{u}{n}} \neq 0 \quad \text{ with probability 1}$$

となる．また,

$$f(x,y) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \frac{1}{2\Gamma(n/2)} \left(\frac{y}{2}\right)^{n/2-1} \exp(-y/2)$$

であるから,

$$
\begin{aligned}
g(t,u) &= f(x,y)\sqrt{u/n} \\
&= \frac{1}{\sqrt{2\pi n}\Gamma(n/2)\,2^{n/2}}\,u^{(n+1)/2-1}\,\exp\left(-\frac{u(t^2/n+1)}{2}\right)
\end{aligned}
$$

となる．これを $u$ に関して $[0,\infty)$ の範囲で積分すれば $T$ の密度関数

$$
g(t) = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\Gamma(n/2)}\,(t^2/n+1)^{-(n+1)/2}
$$

が得られる．

**(注 1)**　自由度 1 の $t$ 分布はコーシー分布である．
**(注 2)**　$E(T) = E(X/\sqrt{Y/n}) = \sqrt{n}\,E(X)\,E(Y^{-1/2})$．ここで,

$$
\begin{aligned}
E(Y^k) &= \int_0^\infty \frac{1}{2^{n/2}\,\Gamma(n/2)}\,y^{n/2+k-1}\,e^{-y/2}\,dy \\
&= \frac{2^k\,\Gamma(n/2+k)}{\Gamma(n/2)} \qquad (n/2+k>0)
\end{aligned}
$$

であるから，$E(Y^k)$ は $n>-2k$ のとき存在する．したがって，$E(Y^{-1/2})$ は $n>1$ のとき存在する．以上から，$E(T)$ は $n>1$ のとき存在して，0 である．
**(注 3)**　$V(T) = E(T^2) = E(X^2/(Y/n)) = n\,E(X^2)\,E(Y^{-1}) = n\,E(Y^{-1}) = n/(n-2)$ $(n>2)$

**・F 分布**

$X$ と $Y$ は互いに独立に，それぞれ $\chi^2(m)$，$\chi^2(n)$ に従うとき,

$$
Z = \frac{X/m}{Y/n}
$$

は，自由度 $(m,n)$ の $F$ 分布に従うといい，$Z \sim F(m,n)$ と書く．$F$ の密度関数を求めるには,変数変換

$$
z = \frac{x/m}{y/n},\quad w = y \qquad \Leftrightarrow \qquad x = \frac{mzw}{n},\quad y = w
$$

を考える．変換のヤコビアンは,

$$
det\left(\frac{\partial(x,y)}{\partial(z,w)}\right) = \begin{vmatrix} mw/n & mz/n \\ 0 & 1 \end{vmatrix} = \frac{mw}{n} > 0 \quad \text{with probability 1}
$$

となる．また,

$$
f(x,y) = \frac{1}{2\Gamma(m/2)}\left(\frac{x}{2}\right)^{m/2-1}\exp(-x/2) \times \frac{1}{2\Gamma(n/2)}\left(\frac{y}{2}\right)^{n/2-1}\exp(-y/2)
$$

であるから,

$$
g(z,w) = f(x,y)\times\frac{mw}{n} = f(mzw/n,w)\times\frac{mw}{n}
$$

を $w$ に関して $[0,\infty)$ の範囲で積分することにより，

$$g(z)=\frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)^{m/2}z^{m/2-1}\left(\frac{m}{n}\,z+1\right)^{-(m+n)/2}$$

**（注 1）**　$T\sim t(n)$ ならば，$T^2\sim F(1,n)$ である．
**（注 2）**　$E(Z)=(n/m)\,E(X)E(Y^{-1})=nE(Y^{-1})=n/(n-2)$ $(n>2)$ （$t$ 分布に関する**（注 2）**を参照のこと）
**（注 3）**　$E(Z^2)=(n/m)^2E(X^2)E(Y^{-2})=(n/m)^2\,m(m+2)E(Y^{-2})$．ここで，$E(Y^{-2})=1/((n-2)(n-4))$ となるので（$t$ 分布に関する**（注 2）**を参照のこと），

$$V(Z)=\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$$

## 4．標本平均と標本分散の独立性

以下，$X_1,X_2,\cdots,X_n$ は互いに独立に $\mathrm{N}(\mu,\sigma^2)$ に従うものとする．そして，

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\qquad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

を定義する．このとき，$\bar{X}$ と $S^2$ は互いに独立であることを証明する．まず

$$Y_i=\frac{X_i-\mu}{\sigma}\qquad(i=1,\cdots,n)$$

は，互いに独立に N(0, 1) に従う．次に，$Y=(Y_1,\cdots,Y_n)'$ から $Z=(Z_1,\cdots,Z_n)'$ への変換 $Z=HY$ を考える．ここで，

$$H=\begin{pmatrix} 1/\sqrt{n} & 1/\sqrt{n} & \cdots & \cdots & \cdots & 1/\sqrt{n}\\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & \cdots & \cdots & 0\\ 1/\sqrt{2\cdot 3} & 1/\sqrt{2\cdot 3} & -2/\sqrt{2\cdot 3} & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 1/\sqrt{(n-1)\cdot n} & \cdots & \cdots & \cdots & 1/\sqrt{(n-1)\cdot n} & -(n-1)/\sqrt{(n-1)\cdot n} \end{pmatrix}$$

は直交行列であり（この変換を **Helmert 変換**という），明らかに $Z\sim N(0,I_n)$ である．特に，

$$Z_1=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Y_i=\frac{\sqrt{n}\,(\bar{X}-\mu)}{\sigma}$$

また，$Y$ と $Z$ は互いに直交変換であるから，

$$Y'Y=Z'Z\qquad\Leftrightarrow\qquad Y_1^2+\cdots+Y_n^2=Z_1^2+\cdots+Z_n^2$$

が成り立つ．ここで，

$$\begin{aligned} Y_1^2+\cdots+Y_n^2 &= \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}=\sum_{i=1}^{n}\frac{((X_i-\bar{X})+(\bar{X}-\mu))^2}{\sigma^2}\\ &= \sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\sigma^2}+\frac{n(\bar{X}-\mu)^2}{\sigma^2}=Z_1^2+\frac{(n-1)\,S^2}{\sigma^2} \end{aligned}$$

であることから，

$$Z_1^2 + \cdots + Z_n^2 = Z_1^2 + \frac{(n-1)\,S^2}{\sigma^2}$$

となるので，結局，

$$\frac{(n-1)\,S^2}{\sigma^2} = Z_2^2 + \cdots + Z_n^2 \quad \sim \quad \chi^2(n-1) \qquad (*)$$

が示される．したがって，$\bar{X}$ は $Z_1$ にのみ依存し，他方，$S^2$ は $Z_2, \cdots, Z_n$ にのみ依存するので，$\bar{X}$ と $S^2$ は独立である．

**（注 1）** 上の $(*)$ から，$E((n-1)\,S^2/\sigma^2) = n-1$，したがって，$E(S^2) = \sigma^2$（不偏）を得る．
**（注 2）** 上の結果から次のことが成り立つ．

$$\frac{\sqrt{n}\,(\bar{X} - \mu)}{S} \quad \sim \quad t(n-1)$$