7 統計量から誘導される分布

n 変数 $x_1, x_2, \cdots, x_n$ の関数 $y = g(x_1, x_2, \cdots, x_n)$ において，$x_1, x_2, \cdots, x_n$ を確率変数 $X_1, X_2, \cdots, X_n$ で置き換えた $Y = g(X_1, X_2, \cdots, X_n)$ を考える．このとき，Y を $X_1, X_2, \cdots, X_n$ の統計量 (statistic) という．

また，確率変数 $X_1, X_2, \cdots, X_n$ を大きさの順に並び替えた $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ を順序統計量 (order statistics) という．

確率変数の和，差，積，平均や分散はいずれも統計量である．また，順序統計量から得られる最大値，最小値，中央値，モードなどもまた統計量である．

1 最大値，最小値と順序統計量の分布

確率変数 X と Y は互いに独立で，確率分布関数 $F_X(x)$ と $F_Y(y)$ に従うものとする．このとき，(X, Y) の最大値，最小値の分布は

$$P\left(\max(X, Y) \leq z\right) = P\left(X \leq z,\ Y \leq z\right) = F_X(z)\ F_Y(z)$$

$$P\left(\min(X, Y) > z\right) = \left(1 - F_X(z)\right)\ \left(1 - F_Y(z)\right)$$

で与えられる．

確率変数 $X_i,\ (i = 1, 2, \cdots, n)$ は互いに独立で同一な確率分布関数 $F_X(x)$ に従うものとする．このとき，最大値，最小値の分布は

$$P\left(\max(X_1, X_2, \cdots, X_n) \leq z\right) = \left(F_X(z)\right)^n$$

$$P\left(\min(X_1, X_2, \cdots, X_n) > z\right) = \left(1 - F_X(z)\right)^n$$

で与えられる．

確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立で同一な確率密度関数 $p_X(x)$ に従うものとする．これらを小さい順に並び替えた順序統計量 $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ の同時確率密度関数は

$$p_{\mathbf{X}}(\mathbf{x}) = n!\ p_X(x_1)\ p_X(x_2)\ \cdots\ p_X(x_n), \qquad (x_1 \leq x_2 \leq \cdots \leq x_n)$$

となる．ただし，$\mathbf{x} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}} \in \Re^n$ である．

2　変数変換

2.1　確率変数の変数変換

一般に実数値関数 $y = g(x)$ の微分が可能で単調関数

$$\frac{dg(x)}{dx} > 0, \qquad (\text{または} \quad < 0)$$

であるとき，その逆関数 $x = g^{-1}(y)$ が存在し，

$$\frac{dg^{-1}(y)}{dy} = \left(\frac{dg(x)}{dx}\right)^{-1}\Big|_{x=g^{-1}(y)}$$

となる．

確率変数 X およびその統計量 $Y = g(X)$ を考える．このとき，それぞれの確率密度関数を $p_X(x)$ と $p_Y(y)$ で表すと

$$p_Y(y) = p_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right| = p_X(x)\left|\frac{dx}{dy}\right|$$

の関係が導かれる．ここで，$|x|$ は x の絶対値を表す関数である．

証明　確率変数 X およびその統計量 $Y = g(X)$ の確率分布関数を $F_X(x)$ と $F_Y(y)$ で表し，関数 $y = g(x)$ が単調増加関数であると仮定すると

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$$
$$= \int_{g(x)\leq y} p_X(x)dx = P(X \leq x) = F_X(x)$$

したがって，

$$p_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X(x)}{dy} = \frac{dF_X(x)}{dx}\frac{dx}{dy} = p_X(x)\frac{dx}{dy}$$

となる．また，関数 $y = g(x)$ が単調減少関数であると仮定すると，先と同様にして

$$F_Y(y) = P(X > x) = 1 - F_X(x)$$

したがって，

$$p_Y(y) = \frac{d(1 - F_X(x))}{dy} = -p_X(x)\frac{dx}{dy}$$

となる．

2.2　確率ベクトルの変数変換

より一般に，ベクトル $\mathbf{x} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}}$ と $\mathbf{y} = [y_1, y_2, \cdots, y_n]^{\mathrm{T}}$ がベクトル値関数

$$\begin{cases} y_1 = g_1(x_1, x_2, \cdots, x_n) \\ y_2 = g_2(x_1, x_2, \cdots, x_n) \\ \qquad \vdots \\ y_n = g_n(x_1, x_2, \cdots, x_n) \end{cases}$$

により 1 対 1 変換 $\mathbf{y} = \mathbf{g}(\mathbf{x}),\ (i = 1, 2, \cdots, n)$ である場合を考える．ただし，$\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), g_2(\mathbf{x}), \cdots, g_n(\mathbf{x})]^{\mathrm{T}}$ である．そのヤコビアン (Jacobian)$[J(\mathbf{g})]$ が

$$J(\mathbf{g}) = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} = \begin{vmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial g_n(\mathbf{x})}{\partial x_n} \end{vmatrix} \neq 0$$

であるとき，その逆関数 $\mathbf{x} = \mathbf{g}^{-1}(\mathbf{y})$ が存在し，

$$J(\mathbf{g}^{-1}) = \frac{\partial \mathbf{g}^{-1}(\mathbf{y})}{\partial \mathbf{y}} = \begin{vmatrix} \frac{\partial g_1^{-1}(\mathbf{y})}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}(\mathbf{y})}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}(\mathbf{y})}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}(\mathbf{y})}{\partial y_n} \end{vmatrix} = \left(\frac{dg(\mathbf{x})}{d\mathbf{x}} \right)^{-1} |_{\mathbf{x}=\mathbf{g}^{-1}(\mathbf{y})}$$

となる．

確率ベクトル $\mathbf{X}$ およびその統計量 $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ を考える．このとき，それぞれの同時確率密度関数を $p_{\mathbf{X}}(\mathbf{x})$ と $p_{\mathbf{Y}}(\mathbf{y})$ で表すと

$$p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{X}}(\mathbf{g}^{-1}(\mathbf{y})) \left| J(\mathbf{g}^{-1}) \right|$$

の関係が導かれる．ここで，$|x|$ は x の絶対値を表す関数である．

証明　確率ベクトル $\mathbf{X}$ と $\mathbf{Y}$ の標本空間は n 次ユークリッド空間 $\Re^n$ である．いま，$\Re^n$ の部分空間 $\mathbf{A}$ に $\mathbf{Y}$ が属する事象を $\mathbf{Y} \in \mathbf{A}$ で表す．また，このような事象 $\mathbf{Y} \in \mathbf{A}$ へ移るような $\mathbf{X}$ の集合全体を $\mathbf{B}$ で表す．このとき，

$$P(\mathbf{Y} \in \mathbf{A}) = \int_{y_1 \in A} \cdots \int_{y_n \in A} p_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = P(\mathbf{X} \in \mathbf{B})$$

である．ところで，

$$P(\mathbf{X} \in \mathbf{B}) = \int_{x_1 \in B} \cdots \int_{x_n \in B} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

であることから，n 重積分の変数変換と同様にして，

$$P(\mathbf{X} \in \mathbf{B}) = \int_{y_1 \in A} \cdots \int_{y_n \in A} p_{\mathbf{X}}(\mathbf{g}^{-1}(\mathbf{y})) |J(\mathbf{g}^{-1})| d\mathbf{y}$$

を得る．よって，

$$p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{X}}\left(\mathbf{g}^{-1}(\mathbf{y})\right)|J(\mathbf{g}^{-1})|$$

となる．

3　確率変数の和の分布

3.1　平均，分散

n 個の確率変数 $X_i,\ (i = 1, 2, \cdots, n)$ を線形変換することにより得られる確率変数 Y_n

$$Y_n = \sum_{i=1}^{n} c_i X_i$$

が従う分布の平均と分散は以下のようになる．

$$\mathrm{E}[Y_n] = \sum_{i=1}^{n} c_i \mathrm{E}[X_i]$$

$$\mathrm{V}[Y_n] = \sum_{i=1}^{n} c_i \mathrm{V}[X_i] + 2\sum_{i<j} c_i c_j \mathrm{C}[X_i, X_j]$$

証明

$$\mathrm{E}[Y_n] = \int \sum_{i=1}^{n} c_i x_i dF_X(x) = \sum_{i=1}^{n} c_i \int x_i dF_X(x) = \sum_{i=1}^{n} c_i \mathrm{E}[X_i]$$

$$\begin{aligned}
\mathrm{V}[Y_n] &= \mathrm{E}\left[\left(\sum_{i=1}^{n} c_i X_i - \mathrm{E}[\sum_{i=1}^{n} c_i X_i]\right)^2\right] \\
&= \mathrm{E}\left[\left\{\sum_{i=1}^{n} c_i \left(X_i - \mathrm{E}[X_i]\right)\right\}^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n} c_i^2 \left(X_i - \mathrm{E}[X_i]\right)^2 + 2\sum_{i<j} c_i c_j \left(X_i - \mathrm{E}[X_i]\right)\left(X_j - \mathrm{E}[X_j]\right)\right] \\
&= \sum_{i=1}^{n} c_i^2 \mathrm{E}\left[\left(X_i - \mathrm{E}[X_i]\right)^2\right] + 2\sum_{i<j} c_i c_j \mathrm{E}\left[\left(X_i - \mathrm{E}[X_i]\right)\left(X_j - \mathrm{E}[X_j]\right)\right] \\
&= \sum_{i=1}^{n} c_i \mathrm{V}[X_i] + 2\sum_{i<j} c_i c_j \mathrm{C}[X_i, X_j]
\end{aligned}$$

その結果，平均 $\mathrm{E}[X_i]=m$，分散 $\mathrm{V}[X_i]=\sigma^2<\infty$ をもつ母集団から抽出した大きさ n の確率変数 X_i，$(i=1,2,\cdots,n)$ から導かれる標本平均 (sample mean) は

$$\bar{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

で定義される．標本平均 $\bar{X}_n$ の平均と分散は，$c_i=\frac{1}{n}$ とすることにより

$$\mathrm{E}\left[\bar{X}_n\right]=\frac{1}{n}\left(\mathrm{E}[X_1]+\mathrm{E}[X_2]+\cdots+\mathrm{E}[X_n]\right)=m$$

$$\mathrm{V}\left[\bar{X}_n\right]=\frac{1}{n^2}\left(\mathrm{V}[X_1]+\mathrm{V}[X_2]+\cdots+\mathrm{V}[X_n]\right)=\frac{\sigma^2}{n}$$

である．

確率変数 X と Y が独立であるとき，$Z=X+Y$ の確率母関数（離散変数の場合）と積率母関数は，

$$G_Z(t)=G_{X+Y}(t)=\mathrm{E}\left[t^{X+Y}\right]=\mathrm{E}\left[t^Xt^Y\right]=\mathrm{E}\left[t^X\right]\mathrm{E}\left[t^Y\right]=G_X(t)G_Y(t)$$

$$M_Z(t)=M_{X+Y}(t)=\mathrm{E}\left[e^{t(X+Y)}\right]=\mathrm{E}\left[e^{tX}e^{tY}\right]=\mathrm{E}\left[e^{tX}\right]\mathrm{E}\left[e^{tY}\right]=M_X(t)M_Y(t)$$

となる．

n 個の確率変数 X_i，$(i=1,2,\cdots,n)$ が独立で同一の分布に従うとき，その分布の確率母関数（離散変数の場合）を $G_X(t)$ と積率母関数を $M_X(t)$ とする．このとき，これら n 個の和 $S_n=X_1+X_2+\cdots+X_n$ に関して

$$G_{S_n}(t)=\left(G_X(t)\right)^n$$

$$M_{S_n}(t)=\left(M_X(t)\right)^n$$

となる．

3.2　分布の畳込みと再生性

確率変数 X と Y が独立であり，その確率分布関数を $F_X(x)$ と $F_Y(y)$ とする．このとき，$Z=X+Y$ の確率分布関数を $F_Z(z)$ とすれば，

$$F_Z(z)=\int_{-\infty}^{\infty}F_X(z-y)dF_Y(y)=\int_{-\infty}^{\infty}F_Y(z-x)dF_X(x)$$

となる．これを $F_Z(z)=F_X*F_Y(z)$ で表し，畳み込み (convolution) という．

証明 確率変数 X と Y が独立であるから，

$$F_Z(z) = P(X+Y \leq z) = \int\int_{x+y\leq z} dF_X(x)dF_Y(y)$$
$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-y} dF_X(x)\right)dF_Y(y) = \int_{-\infty}^{\infty} F_X(z-y)dF_Y(y)$$

となる．同様に先に y について積分することも可能である．

確率変数 X と Y が独立で同一な確率分布に従っているものとし，その確率分布関数を $F_X(x)$ と $F_Y(y)$ とする．このとき，$Z = X+Y$ の確率分布関数を $F_Z(z)$ とし

$$F_X, F_Y \in F \Longrightarrow F_Z(z) = F_X * F_Y(z) \in F$$

が満たされるならば，F は再生性 (reproduction) をもつという．

分布の集合のことを分布族 F という．普通は，確率分布関数や確率密度関数の関数型は共通で，母数によって特徴づけられた分布の集合

$$F = \{p_X(x|\theta) : \theta \in \Theta\}$$

のことである．再生性をもつ分布族では，分布型がわかっているので母数のみを決定すればよいことがわかる．このため，確率変数の和の分布を求めるが非常に容易となる．再生性をもつ分布族の一部を以下に示す．

二項分布 $B(n_1, p) + B(n_2, p) \stackrel{d}{=} B(n_1+n_2, p)$

二項分布に従う確率変数 X_1 と X_2 の積率母関数はそれぞれ，$M_{X_1}(t) = \{pt+(1-p)\}^{n_1}$ と $M_{X_2}(t) = \{pt+(1-p)\}^{n_2}$ となる．このとき，$S_2 = X_1+X_2$ の積率母関数は

$$M_{S_2}(t) = M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = \{pt+(1-p)\}^{n_1+n_2}$$

となる．このことから，X_1+X_2 の分布が二項分布 $[B(n_1+n_2, p)]$ であることがわかる．

ポアソン分布 $Po(\lambda_1) + Po(\lambda_2) \stackrel{d}{=} Po(\lambda_1+\lambda_2)$

ポアソン分布に従う確率変数 X_1 と X_2 の積率母関数はそれぞれ，$M_{X_1}(t) = e^{\lambda_1(e^t-1)}$ と $M_{X_2}(t) = e^{\lambda_2(e^t-1)}$ となる．このとき，$S_2 = X_1+X_2$ の積率母関数は

$$M_{S_2}(t) = M_{X_1+X_2}(t) = M_{X_1}(t)M_{X_2}(t) = e^{(\lambda_1+\lambda_2)(e^t-1)}$$

となる．このことから，X_1+X_2 の分布がポアソン分布 $[Po(\lambda_1+\lambda_2)]$ であることがわかる．

負の二項分布（あるいは幾何分布）

n 個の確率変数 X_i, $(i = 1, 2, \cdots, n)$ が独立で同一の幾何分布 $[G(p)]$ に従うとき，これら n 個の和 $S_n = X_1 + X_2 + \cdots + X_n$ の分布は負の二項分布 $[NB(n, p)]$ に従う．S_n の積率母関数が

$$M_{S_n}(t) = M_{X_1+X_2+\cdots+X_n}(t) = M_{X_i}^n(t) = \left\{\frac{p}{1-(1-p)t}\right\}^n$$

となることからわかる．

正規分布　$N(m_1, \sigma_1^2) + N(m_2, \sigma_2^2) \stackrel{d}{=} N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

正規分布に従う確率変数 X_1 と X_2 の積率母関数はそれぞれ，$M_{X_1}(t) = e^{m_1 t \frac{\sigma_1^2 t^2}{2}}$ と $M_{X_2}(t) = e^{m_2 t \frac{\sigma_2^2 t^2}{2}}$ となる．このとき，$S_2 = X_1 + X_2$ の積率母関数は

$$M_{S_2}(t) = M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t) = e^{(m_1+m_2)t \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$$

となる．このことから，$X_1 + X_2$ の分布が正規分布 $[N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)]$ であることがわかる．

アーラン分布（あるいは指数分布）

n 個の確率変数 X_i, $(i = 1, 2, \cdots, n)$ が独立で同一の指数分布 $[Ex(\lambda)]$ に従うとき，これら n 個の和 $S_n = X_1 + X_2 + \cdots + X_n$ の分布はアーラン分布 $[Er(n, \frac{1}{\lambda})]$ に従う．S_n の積率母関数が

$$M_{S_n}(t) = M_{X_1+X_2+\cdots+X_n}(t) = M_{X_i}^n(t) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

となることからわかる．

ガンマ分布　$Ga(\alpha_1, \beta) + Ga(\alpha_2, \beta) \stackrel{d}{=} Ga(\alpha_1 + \alpha_2, \beta)$

ガンマ分布に従う確率変数 X_1 と X_2 の積率母関数はそれぞれ，$M_{X_1}(t) = (1 - \beta t)^{-\alpha_1}$ と $M_{X_2}(t) = (1 - \beta t)^{-\alpha_2}$ となる．このとき，$S_2 = X_1 + X_2$ の積率母関数は

$$M_{S_2}(t) = M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t) = (1 - \beta t)^{-(\alpha_1+\alpha_2)}$$

となる．このことから，$X_1 + X_2$ の分布がガンマ分布 $[Ga(\alpha_1 + \alpha_2, \beta)]$ であることがわかる．

4　問題 7

7.1　確率変数 X は一様分布 $[U(0,1)]$ に従うものとする．これから導かれる

$$Y = -2\log X$$

が従う分布は指数分布 $[E_x(\frac{1}{2})]$，すなわち，自由度 2 の χ^2 分布 $[\chi^2_2]$ となることを示せ．

7.2　確率変数 X と Y は独立な確率変数で，ともに標準正規分布 $[N(0,1)]$ に従うものとする．これから導かれる

$$C = \frac{X}{Y}$$

が従う分布はコーシー分布 $[Cy(0,1)]$ となることを示せ．

7.3　確率変数 X と Y は独立な確率変数で，それぞれガンマ分布 $[Ga(\alpha,\sigma)]$ とガンマ分布 $[Ga(\beta,\sigma)]$ に従うものとする．これから導かれる

$$Z = \frac{X}{X+Y}$$

が従う分布を求め，何分布となるかを述べよ．

7.4　(二次元正規分布)　確率変数 X_1 と X_2 が独立な確率変数で，ともに標準正規分布 $[N(0,1)]$ に従うものとする．これから導かれる

$$Y_1 = aX_1 + bX_2, \qquad Y_2 = cX_1 + dX_2$$

について以下の問いに答えよ．ただし，$a,\ b,\ c,\ d$ は定数であり，$ad-bc \neq 0$ である．また，既に前問で使用した記号は用いてよい．

(1)　$Y_1,\ Y_2$ の期待値 $m_1,\ m_2$，ならびに分散 $\sigma_1^2,\ \sigma_2^2$ を求めよ．

(2)　$Y_1,\ Y_2$ の共分散 σ_{12}^2，ならびに相関係数 ρ を求めよ．

(3)　$Y_1,\ Y_2$ の同時確率密度関数 $p_{Y_1Y_2}(y_1,y_2)$ を求めよ．

7.5　(ショックモデル)　システム S_1, S_2 の寿命を確率変数 X_1, X_2 で表す．これらは互いに独立で指数分布 $[E_x(\lambda_1)]$ と指数分布 $[E_x(\lambda_2)]$ に従うものとする．さらに，システムはショックを受けると故障することがある．ショックの発生間隔 Y は指数分布 $[E_x(\tau)]$ に従うものとする．システムの寿命が尽きるか，それまでにショックで故障する場合の時間は

$$X'_1 = \min(X_1, Y), \qquad X'_2 = \min(X_2, Y)$$

で与えられる．以下の問いに答えよ．

(1)　(X_1, X_2) の同時確率分布関数を求めよ．

(2)　X_1 と X_2 の相関係数 $\rho(X_1, X_2)$ を求めよ．

(3)　S_1, S_2 が直列に結合されているシステム全体の寿命を確率変数 Z とする．確率変数 Z が従う確率分布関数，ならびにその平均を求めよ．

(4)　S_1, S_2 が並列に結合されている場合はどうなるか．