# 確率・統計 A

若木宏文

wakaki@hiroshima-u.ac.jp
http://home.hiroshima-u.ac.jp/ wakaki/lecture/index.shtml

2018.7.3

# 目 次

# 2.5.3. 多次元の場合

# 多次元離散型分布と連続型分布

$k$ 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_k)'$ の取り得る値の集合 $D$ が, $\mathbb{R}^k$ の高々可算集合であるとき, $\boldsymbol{X}$ は 離散型確率変数, また, $\boldsymbol{X}$ の分布を離散型分布という.

離散型分布は, 確率関数

$$f(\boldsymbol{x}) = f(x_1, \ldots, x_k) = \mathrm{P}(X_1 = x_1, \ldots, X_k = x_k), \ \boldsymbol{x} \in \mathbb{R}^k$$

によって決定される.

$k$ 次元確率変数 $(X_1, \ldots, X_k)'$ 分布関数が

$$F(x_1, \ldots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(s_1, \ldots, s_k) ds_1 \cdots ds_k$$

と表されるとき,$\boldsymbol{X} = (X_1, \ldots, X_k)'$ は 連続型確率変数, また, $\boldsymbol{X}$ の分布を連続型分布という.
$f(x_1, \ldots, x_k)$ を $\boldsymbol{X}$ または, $\boldsymbol{X}$ の分布の確率密度関数という.

# 多次元確率関数の性質 (離散型)

**確率関数**

$X_i$ $(i=1,\dots,k)$ の取り得る値を $x_{i1}, x_{i2}, \dots$ とすると

(1) $f(x_{1j_1},\dots,x_{kj_k}) \geq 0,\ \ j_i = 1,2,\dots\ \ (i=1,\dots,k)$

(2) $\displaystyle\sum_{j_1=1}^{\infty}\cdots\sum_{j_k=1}^{\infty} f(x_{1j_1},\dots,x_{kj_k}) = 1$

**同時分布関数**：

$$F(z_1,\dots,z_k) = \sum_{x_{1j_1}\leq z_1}\cdots\sum_{x_{kj_k}\leq z_k} f(x_{1j_1},\dots,x_{kj_k})$$

# 多次元確率密度関数の性質 (連続型)

**確率密度関数**

(1) $f(x_1, \ldots, x_k) \geq 0, \ {}^{\forall}(x_1, \ldots, x_k) \in \mathbb{R}^k$

(2) $\displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_k) dx_1 \cdots dx_k = 1$

**同時分布関数**

$$F(x_1, \ldots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(s_1, \ldots, s_k) ds_1 \cdots ds_k$$

$$\left( \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1, \ldots, x_k) = f(x_1, \ldots, x_k) \right)$$

# 多項分布

1 回の試行で事象 $A_1, \ldots, A_{k+1}$ のいずれかが起こりうる.

$\mathsf{P}(A_j) = p_j \ (j = 1, \ldots, k)$
$\mathsf{P}(A_{k+1}) = p_{k+1} = 1 - (p_1 + \cdots + p_k)$

$X_j$ : $n$ 回の独立な試行で $A_j$ が起こる回数 $(j = 1, \ldots, k)$

$$f(x_1, \ldots, x_k) = \mathsf{P}(X_1 = x_1, \ldots, X_k = x_k)$$
$$= \frac{n!}{x_1! \cdots x_k! x_{k+1}!} p_1^{x_1} \ldots p_{k+1}^{x_{k+1}} \quad (x_{k+1} = n - (x_1 + \ldots + x_k))$$
$$0 \le p_j \ (j = 1, \ldots, k); p_1 + \cdots + p_k \le 1)$$
$$x_j \in \mathbb{Z}, z_j \ge 0 \ (j = 1, \ldots, k), x_1 + \cdots + x_k \le n\}$$

パラメータ $n, (p_1, \ldots, p_k)$ をもつ $(k+1)$ 項分布といい,
$M_{k+1}(n, (p_1, \ldots, p_k))$ と表す.

# 多変量正規分布

$(X_1, \ldots, X_k)'$ の確率密度関数が

$$f(x_1, \ldots, x_k) = (2\pi)^{-k/2} |\Sigma|^{-1/2} \exp\left\{ -\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})' \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right\}$$

と表せるとき，$k$ 次元正規分布に従うという．ここで

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kk} \end{pmatrix}.$$

パラメータ $\boldsymbol{\mu}$ は任意の定数，すなわち，$-\infty < \mu_i < \infty$ $(i = 1, \ldots, k)$ で，$\Sigma$ は正定値行列である．

## 定理 2.12

$(X_1, \ldots, X_k)'$ は離散型あるいは連続型.
$X_i,\ i = 1, \ldots, k$ の周辺確率密度関数を $f_i(x_i)$,
同時確率密度関数を $f(x_1, \ldots, x_k)$ とする.

このとき, $X_1, \ldots, X_k$ が独立であることと, 任意の
$(x_1, \ldots, x_k)' \in \{(x_1, \ldots, x_k)';\ f(x_1, \ldots, x_k) > 0\}$ に対して

$$f(x_1, \ldots, x_k) = f_1(x_1) \cdots f_k(x_k)$$

が成り立つことは同値である.

# 3.1. 平均の定義

# 離散型確率変数の平均

## 定義 3.1

確率変数 $X$ は離散型で, その確率関数を $f(x)$ とする.
$X$ のとりうる値が有限個で, 取り得る値の集合が
$D = \{x_1, \ldots, x_n\}$ の場合

$$\mathsf{E}(X) = \sum_{j=1}^{n} x_j \mathsf{P}(X = x_j) = \sum_{j=1}^{n} x_j f(x_j)$$

を $X$ の平均, あるいは期待値という.

# 離散型確率変数の平均(続き)

## 定義 3.1 (続き)

とりうる値の集合が可算集合 $D=\{x_1,x_2,\ldots\}$ の場合には, 級数

$$\mathsf{E}(X)=\sum_{j=1}^{\infty}x_j\mathsf{P}(X=x_j)=\sum_{j=1}^{\infty}x_jf(x_j)$$

が絶対収束するとき, $X$ の平均は存在するといい, その極限値を $X$ の平均, あるいは期待値という.

# 連続型確率変数の平均

## 定義 3.2

確率変数 $X$ は連続型で，その確率密度関数を $f(x)$ とする．

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

であるとき，$X$ の平均は存在するといい，

$$\mathsf{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$$

を $X$ の平均，あるいは期待値という．

# 平均の例

## 例 3.1

(1) $X \sim B(1, p)$ のとき, $\mathrm{E}(X) = p$

(2) $X \sim B(n, p)$ のとき, $\mathrm{E}(X) = np$

## 例 3.2

(1) $X \sim U(a, b]$ のとき, $E(X) = \dfrac{b-a}{2}$

(2) $X \sim N(\mu, \sigma^2)$ のとき, $\mathrm{E}(X) = \mu$.

(証明は板書で)

# メモ用紙

# メモ用紙

# 3.2. 基本的性質

# 確率変数の関数の期待値

## 定理 3.1

確率変数 $X$ が離散型のとき, $X$ のとりうる値を $x_j$, $j = 1, 2, \ldots$ とし, 確率関数を $f(x_j)$ とする. また, $X$ が連続型のとき, その確率密度関数を $f(x)$ とする. このとき

$$\mathrm{E}\{g(X)\} = \begin{cases} \displaystyle\sum_{j=1}^{\infty} g(x_j) f(x_j), & X \text{ が離散型のとき} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f(x) dx, & X \text{ が連続型のとき} \end{cases}$$

である.

(証明 (離散型の場合) は板書で)

## 注 3.3

$g(X)$ の分布を求めなくても, $X$ の分布から求められる.

# メモ用紙

## 例 3.3

$X$ は 一様分布 $U(0,1]$ に従うとする．このとき，$Y = X^2$ の平均を考える．定理 3.1 を利用すると

$$\mathsf{E}(X^2) = \int_0^1 x^2 \cdot 1 dx = \frac{1}{3}$$

一方，$Y$ の分布関数は $0 \le y \le 1$ のとき

$$F_Y(y) = \mathsf{P}(Y \le y) = \mathsf{P}(X \le \sqrt{y}) = \sqrt{y}$$

$$f_Y(y) = F_Y'(y) = \begin{cases} \dfrac{1}{2} y^{-1/2}, & 0 < x \le 1 \text{ のとき} \\ 0, & \text{その他のとき} \end{cases}$$

となるので

$$\mathsf{E}(Y) = \int_0^1 y \cdot \frac{1}{2} y^{-1/2} dy = \frac{1}{3}$$

## 例 3.4

$X$ : 連続型, $f(x)$ : 確率密度関数, $F(x)$ : 分布関数
$g$ : 微分可能な狭義の単調減少関数,

$$a = \inf\{g(X(\omega));\ \omega \in \Omega\},\ b = \sup\{g(X(\omega));\ \omega \in \Omega\}$$

$a < y < b$ のとき

$$F_Y(y) = \mathsf{P}(g(X) \le y) = \mathsf{P}(X \ge g^{-1}(y)) = 1 - F(g^{-1}(y))$$

$$\frac{d}{dy}F_Y(y) = -\frac{dg^{-1}(y)}{dy}f(g^{-1}(y)) = \left|\frac{dg^{-1}(y)}{dy}\right| f(g^{-1}(y))$$

$$\mathsf{E}(Y) = \int_a^b y \left|\frac{dg^{-1}(y)}{dy}\right| f(g^{-1}(y))dy$$

$x = g^{-1}(y)$ と変数変換すると

$$\mathsf{E}(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

(($x > g^{-1}(a)$ または $x < g^{-1}(b)$ で $f(x) = 0$) )

# 多変量の場合

## 定理 3.2

$n$ 次元確率ベクトル $\boldsymbol{X}$ が離散型のとき，そのとりうる値を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ とし確率関数を $f(\boldsymbol{x}_j)$ とする．また，$\boldsymbol{X}$ が連続型のとき，その確率密度関数を $f(\boldsymbol{x})$ とする．このとき

$$\mathrm{E}(g(\boldsymbol{X})) = \begin{cases} \displaystyle\sum_{j=1}^{\infty} g(\boldsymbol{x}_j) f(\boldsymbol{x}_j), & \boldsymbol{x} \text{ が離散型のとき} \\ \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}, & \boldsymbol{x} \text{ が連続型のとき} \end{cases}$$

である．

(離散型の場合の証明は定理 3.1 と同じ)

# 定理 3.2 の証明の準備

## 定理 3.3

$n$ 次元確率ベクトル $\boldsymbol{X}$ は連続型で, 確率密度関数 $f(x_1, \ldots, x_n)$ をもつとする. また, $\boldsymbol{X}$ のとりうる値は領域 $\mathcal{X} = \{\boldsymbol{x}; f(\boldsymbol{x}) > 0\} \subset \mathbb{R}^n$ であるとする. $\boldsymbol{x} = (x_1, \ldots, x_n)'$ から $\boldsymbol{y} = (y_1, \ldots, y_n)'$ への変換

$$\begin{cases} y_1 = u_1(\boldsymbol{x}) = u_1(x_1, \ldots, x_n) \\ \qquad\vdots \\ y_n = u_n(\boldsymbol{x}) = u_n(x_1, \ldots, x_n) \end{cases}$$

は領域 $\mathcal{X}$ を領域 $\mathcal{Y}$ に写し, 次の性質 $(1) \sim (3)$ をみたすとする.

## 定理 3.3 (続き)

(1) $1-1$ である.

(2) 逆変換

$$\boldsymbol{y} \to \boldsymbol{x} : \quad x_i = v_i(\boldsymbol{y}) = v_i(y_1, \ldots, y_n),\ i = 1, \ldots, n$$

は $C^1$ クラスである.

(3) ヤコビ行列式

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \cdots & \dfrac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial y_1} & \cdots & \dfrac{\partial x_n}{\partial y_n} \end{vmatrix}$$

の値が任意の $\boldsymbol{y} \in \mathcal{Y}$ に対してゼロでない.

## 定理 3.3 (続き)

この変換により定義される確率変数を

$$Y_i = u_i(\boldsymbol{X}) = u_i(X_1, \ldots, X_n), \quad i = 1, \ldots, n$$

とする．このとき，$\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$ の確率密度関数は

$$h(y_1, \ldots, y_n) = \begin{cases} f(v_1(\boldsymbol{y}), \ldots, v_n(\boldsymbol{y}))|J|, & \boldsymbol{y} \in \mathcal{Y} \\ 0, & \boldsymbol{y} \in \mathcal{Y}^c \end{cases}$$

となる．

# 定理 3.3 の証明

$\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$ の分布関数は

$$
\begin{aligned}
F_{Y_1,\ldots,Y_n}(a_1, \ldots, a_n) &= \mathsf{P}(Y_1 \le a_1, \ldots, Y_n \le a_n) \\
&= \mathsf{P}(u_1(\boldsymbol{x}) \le a_1, \ldots, u_n(\boldsymbol{x}) \le a_n)
\end{aligned}
$$

と表せる. $D_{\boldsymbol{a}} = \{\boldsymbol{x} \in \mathbb{R}^n;\ u_i(\boldsymbol{x}) \le a_i,\ i = 1, \ldots, n\}$ とおく.
変数変換 $y_i = u_i(\boldsymbol{x})$, $i = 1, \ldots, n$ を考えることによって

$$
\begin{aligned}
F_{Y_1,\ldots,Y_n}(a_1, \ldots, a_n) &= \int \cdots \int_{D_{\boldsymbol{a}}} f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
&= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} h(y_1, \ldots, y_n) dy_1 \cdots dy_n
\end{aligned}
$$

を得る. □

# 定理 3.2 の証明 (連続型)

$$Y = Y_1 = u_1(X_1, \ldots, X_n) = g(X_1, \ldots, X_n)$$

とし, (1)~(3) が成り立つような関数 $u_2(\boldsymbol{x}), \ldots, u_n(\boldsymbol{x})$ がとれる場合.

$Y = Y_1$ の確率密度関数は

$$h_1(y_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(y_1, \ldots, y_n) dy_2 \cdots dy_n$$

で与えられ, 変数変換 $y_i = u_i(\boldsymbol{x})$, $i = 1, \ldots, n$ を考えることによって

$$\mathrm{E}(Y_1) = \int_{-\infty}^{\infty} y_1 h_1(y_1) dy_1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y_1 h(y_1, \ldots, y_n) dy_1 \cdots dy_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}. \qquad \square$$

# 例

$X_1, X_2$ は独立で, $X_1 \sim N(2,1),\ X_2 \sim N(1,1)$

(1) $Y_1 = X_1 + X_2,\ Y_2 = X_1 - X_2$ とすると

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\boldsymbol{\mu}, \Sigma\right), \quad \boldsymbol{\mu} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

(2) $\mathsf{E}(X_1 + X_2) = 3$

(証明は板書で)

# メモ用紙

# メモ用紙

# 平均の線形性と単調性, 独立な確率変数の積の平均

### 定理 3.4

確率変数 $X, Y$ は平均をもつとする.

(1) 任意の定数 $a$ に対して, $\mathsf{E}(aX) = a\mathsf{E}(X)$.

(2) $\mathsf{E}(X+Y) = \mathsf{E}(X) + \mathsf{E}(Y)$.

(3) $X \leq Y \Rightarrow \mathsf{E}(X) \leq \mathsf{E}(Y)$.

(証明 (離散型の場合) を板書で)

### 定理 3.5

確率変数 $X, Y$ は平均をもち, 独立とする. このとき

$$\mathsf{E}(XY) = \mathsf{E}(X)\mathsf{E}(Y).$$

(証明 (離散型の場合) を板書で)

# メモ用紙

# メモ用紙

# メモ用紙

## 注 3.5

$g, h$ を $\mathbb{R}$ 上の任意のボレル関数とする．定理 3.2 を用いると定理 3.5 と同様に $X, Y$ が独立ならば

$$\mathsf{E}\{g(X)h(Y)\} = \mathsf{E}\{g(X)\}\mathsf{E}\{h(Y)\} \tag{1}$$

を示すことができる．逆に，任意のボレル関数 $g, h$ に対して (1) が成り立つならば，$g, h$ をそれぞれ，区間 $J_a = (a_1, a_2],\ J_b = (b_1, b_2]$ の定義関数ととることにより $\mathsf{P}(X \in J_a,\ Y \in J_b) = \mathsf{E}\{\mathrm{I}_{(a_1,a_2]}(X)\mathrm{I}_{(b_1,b_2]}(Y)\} = \mathsf{P}(X \in J_a)\mathsf{P}(Y \in J_b)$ が導かれるので，$X, Y$ は独立である．ここで，一般に集合 $S$ の部分集合 $A$ に対して，$A$ の定義関数とは

$$\mathrm{I}_A(s) = \begin{cases} 1, & s \in A \\ 0, & s \notin A \end{cases}$$

と定義される $S$ 上の関数を意味する．