# 確率・統計 B

若木宏文

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

2017.10.6

**教科書**：「確率・統計の数学的基礎」(藤越，若木，柳原著)

**確率・統計 A の復習**

- 確率空間の定義と性質
- 確率変数の分布
- 期待値
- 特性関数
- 特性関数と分布

## 確率空間の定義と性質 (1/5)

### 定義 1 (確率空間)

$(\Omega, \mathcal{B}, \mathrm{P})$

- $\Omega$ : 標本空間.
  試行によって起こり得る結果全体からなる集合
- $\mathcal{B}$ : $\sigma$-集合体. 確率の定義域
- $\mathrm{P}$ : $\mathcal{B}$ 上で定義された確率測度

### 定義 2 ($\sigma$-集合体)

$\mathcal{B}$ が $\Omega$ 上の $\sigma$-集合体であるとは, $\mathcal{B} \subset \wp(\Omega)$ であって

(B1) $\Omega \in \mathcal{B}$

(B2) $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$

(B3) $A_1, A_2, \ldots \in \mathcal{B} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

## 確率空間の定義と性質 (2/5)

### $\sigma$-集合体の性質

$\mathcal{B}$ を $\Omega$ 上の $\sigma$-集合体とすると

(1) $\emptyset \in \mathcal{B}$.

(2) $A_1, \ldots, A_n \in \mathcal{B} \Rightarrow \bigcup_{i=1}^{n} A_i,\ \bigcap_{i=1}^{n} A_i \in \mathcal{B}$.

(3) $A_1, A_2, \ldots \in \mathcal{B} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$.

(4) $A_1, A_2, \ldots \in \mathcal{B} \Rightarrow \underline{\lim}_{n\to\infty} A_n,\ \overline{\lim}_{n\to\infty} A_n \in \mathcal{B}$.

## 確率空間の定義と性質 (3/5)

### 定義 3 (確率測度)

P が $(\Omega, \mathcal{B})$ 上の確率測度であるとは, P が $\mathcal{B}$ 上の実数値関数であって

(P1) 任意の $A \in \mathcal{B}$ に対して $0 \leq \mathrm{P}(A) \leq 1$.

(P2) $\mathrm{P}(\Omega) = 1$.

(P3) P は完全加法的である. すなわち, $A_1, A_2, \ldots \in \mathcal{B}$ で, $A_i \cap A_j = \emptyset \ (i \neq j)$ ならば

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i).$$

# 確率空間の定義と性質 (4/5)

## 確率測度の性質 1

(1) $\mathsf{P}(\emptyset) = 0$.

(2) $\mathsf{P}(A^c) = 1 - \mathsf{P}(A)$.

(3) $A \subset B \Rightarrow \mathsf{P}(A) \leq \mathsf{P}(B)$.

(4) $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$.

(5) $A_i \cap A_j = \emptyset \ (i \neq j) \quad \Rightarrow \mathsf{P}(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n \mathsf{P}(A_i)$.

(6) $\mathsf{P}(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mathsf{P}(A_i)$.

# 確率空間の定義と性質 (5/5)

## 確率測度の性質 2

(1) $\mathsf{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathsf{P}(A_n)$.

(2) $A_n \subset A_{n+1}\ (n = 1, 2, \ldots) \Rightarrow \mathsf{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n)$.

(3) $A_n \supset A_{n+1}\ (n = 1, 2, \ldots) \Rightarrow \mathsf{P}(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n)$.

(4) $\overline{\lim}_{n\to\infty} A_n = \underline{\lim}_{n\to\infty} A_n$
$\Rightarrow \mathsf{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n)$.

ただし

$$\overline{\lim}_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \quad \underline{\lim}_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

## 確率変数の分布 (1/8)

1 次元ボレル集合体 $\mathbb{B}_1$:
　$\mathbb{R}$ 上の, すべての開集合を含む最小の $\sigma$-集合体

### 定義 4 (確率変数)

$X$ が 確率空間 $(\Omega, \mathcal{B}, \mathsf{P})$ 上の確率変数であるとは, $X$ は $\Omega$ 上の実数値関数であって

$$\forall A \in \mathbb{B}_1 \ X^{-1}(A) \in \mathcal{B}$$

### 定義 5 (確率変数の分布)

$(\Omega, \mathcal{B}, \mathsf{P})$ 上の確率変数 $X$ に対して, 次で定義される $(\mathbb{R}, \mathbb{B}_1)$ 上の確率測度 $\mathsf{P}_X$ を $X$ によって導入される確率測度, あるいは $X$ の分布 という.

$$\mathsf{P}_X(A) = \mathsf{P}(X^{-1}(A)) \quad (A \in \mathbb{B}_1)$$

## 確率変数の分布 (2/8)

$n$ 次元ボレル集合体 $\mathbb{B}_n$:
　$\mathbb{R}^n$ 上の, すべての開集合を含む最小の $\sigma$-集合体

### 定義 6 ($n$ 次元確率変数 (確率ベクトル))

$\boldsymbol{X} = (X_1, \ldots, X_n)$ の各成分 $X_i$ が確率空間 $(\Omega, \mathcal{B}, \mathsf{P})$ 上の確率変数であるとき $\boldsymbol{X}$ は $n$ 次元確率変数 (確率ベクトル) であるという.

### 定義 7 ($n$ 次元確率変数の分布)

$(\Omega, \mathcal{B}, \mathsf{P})$ 上の $n$ 次元確率変数 $\boldsymbol{X}$ に対して, 次で定義される $(\mathbb{R}^n, \mathbb{B}_n)$ 上の確率測度 $\mathsf{P}_{\boldsymbol{X}}$ を $\boldsymbol{X}$ によって導入される 確率測度, あるいは $\boldsymbol{X}$ の分布 という.

$$\mathsf{P}_{\boldsymbol{X}}(A) = \mathsf{P}(\boldsymbol{X}^{-1}(A)) \quad (A \in \mathbb{B}_n)$$

# 確率変数の分布 (3/8)

分布関数 : $n$ 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_n)$ の分布関数とは

$$\begin{aligned} F_{\boldsymbol{X}}(x_1, \ldots, x_n) &= \mathsf{P}(X_1 \le x_1, \ldots, X_n \le x_n) \\ &= \mathsf{P}_{\boldsymbol{X}}((-\infty, x_1] \times \cdots \times (-\infty, x_n]) \end{aligned}$$

$F_{\boldsymbol{X}}$ と $\mathsf{P}_{\boldsymbol{X}}$ は 1 対 1 に対応する

# 確率変数の分布 (4/8)

## 定義 8 (連続型分布)

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\boldsymbol{X}}(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

あるいは

$$\mathrm{P}_{\boldsymbol{X}}(A) = \int_A f_{\boldsymbol{X}}(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

と表されるとき, $\boldsymbol{X}$, あるいは $\mathrm{P}_{\boldsymbol{X}}$ は連続型であるという. このとき, $f_{\boldsymbol{X}}$ を 確率密度関数 とよぶ.

連続型分布は, 確率密度関数によって一意に定められる.

## 確率変数の分布 (5/8)

### 例 1 (正規分布)

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)$$

によって定まる連続型分布を 平均 $\mu$, 分散 $\sigma^2$ の正規分布とよび $N(\mu, \sigma^2)$ と表す.

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp\Big(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big)$$

によって定まる $n$ 次元連続型分布を 平均ベクトル $\boldsymbol{\mu}$, 分散共分散行列 $\Sigma$ の $n$ 次元正規分布とよび $N_n(\boldsymbol{\mu}, \Sigma)$ と表す. ただし

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & & \vdots \end{pmatrix}$$

# 確率変数の分布 (6/8)

## 定義 9 (離散型分布)

$\mathbb{R}^n$ 内の点列 $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ と, 実数列 $p_1, p_2, \ldots$ が存在して

$$\mathsf{P}_{\boldsymbol{X}}(A) = \sum_{i:\boldsymbol{x}_i \in A} p_i$$

と表されるとき, $\boldsymbol{X}$, あるいは $\mathsf{P}_{\boldsymbol{X}}$ は離散型であるという.

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \mathsf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

を確率密度関数 (確率関数) とよぶ. このとき,

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \begin{cases} p_i & (x_1, \ldots, x_n) = \boldsymbol{x}_i \quad (i = 1, 2, \ldots) \\ 0 & \text{その他} \end{cases}$$

である.

# 確率変数の分布 (7/8)

## 例 2 (2 項分布)

$$f_X(x) = \begin{cases} {}_nC_x p^x (1-p)^{n-x} & x = 0, 1, \ldots, n \\ 0 & \text{その他} \end{cases}$$

によって定まる離散型分布を 繰り返し数 $n$, 成功確率 $p$ の 2 項分布とよび $B(n, p)$ と表す.

## 例 3 (3 項分布)

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{n!}{x!y!(n-x-y)!} p^x q^y (1-p-q)^{n-x-y} & \left( \begin{array}{c} x = 0, 1, \ldots \\ y = 0, 1, \ldots \\ x + y \leq n \end{array} \right. \\ 0 & \text{その他} \end{cases}$$

によって定まる離散型分布を 繰り返し数 $n$, 成功確率 $(p, q)$ の 3 項分布とよび $M_3(n, (p, q))$ と表す

# 確率変数の分布 (8/8)

## 確率変数の独立性

$(X_1, \ldots, X_n)$ の確率密度関数を $f_{\boldsymbol{X}}(x_1, \ldots, x_n)$,
$X_1, \ldots, X_n$ の周辺確率密度関数を, それぞれ $f_1(x_1), \ldots, f_n(x_n)$ とする.
$X_1, \ldots, X_n$ が互いに独立であるための必要十分条件は

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i)$$

# 期待値の定義と性質 (1/3)

## 定義 10 (期待値)

$\boldsymbol{X}$ : $n$ 次元確率変数
$g(\boldsymbol{x})$ : $\mathbb{R}^n$ 上の実数値 (ボレル可測) 関数
に対して

$$\mathrm{E}(g(\boldsymbol{X})) = \begin{cases} \int_{\mathbb{R}^n} g(\boldsymbol{x}) f_{\boldsymbol{X}}(x_1, \ldots, x_n) dx_1 \cdots dx_n & (\text{連続型の場合}) \\ \sum_{i=1}^{\infty} g(\boldsymbol{x}_i) f_{\boldsymbol{X}}(\boldsymbol{x}_i) & (\text{離散型の場合}) \end{cases}$$

を $g(\boldsymbol{X})$ の期待値とよぶ.

- $\mathrm{E}(X)$ : $X$ の平均
- $\mathrm{Var}(X) = \mathrm{E}\{(X - \mathrm{E}(X))^2\}$ : $X$ の分散
- $\mathrm{Cov}(X, Y) = \mathrm{E}\{(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))\}$ : $X$ と $Y$ の共分散

# 期待値の定義と性質 (2/3)

## 平均の性質

(1) $\mathsf{E}(aX) = a\mathsf{E}(X)$.

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

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

(4) $X$ と $Y$ が独立 $\Rightarrow \mathsf{E}\{g(X)h(Y)\} = \mathsf{E}\{g(X)\}\mathsf{E}\{h(Y)\}$

## 分散の性質

(1) $\mathsf{Var}(X+a) = \mathsf{Var}(X)$.

(2) $\mathsf{Var}(aX) = a^2\mathsf{Var}(X)$.

(3) $\mathsf{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}(X)^2$.

(4) $\mathsf{Var}(X+Y) = \mathsf{Var}(X) + \mathsf{Var}(Y) + 2\mathsf{Cov}(X,Y)$

(5) $X$ と $Y$ が独立 $\Rightarrow \mathsf{Var}(X+Y) = \mathsf{Var}(X) + \mathsf{Var}(Y)$

# 期待値の定義と性質 (3/3)

## 不等式

(1) (マルコフの不等式) $X$ を非負値確率変数とする．任意の正数 $a$ に対して

$$\mathsf{P}(X \geq a) \leq \mathsf{E}(X)/a.$$

(2) (チェビシェフの不等式) 確率変数 $X$ の平均, 分散を $\mathsf{E}(X) = \mu$, $\mathsf{Var}(X) = \sigma^2 < \infty$ とする．任意の正数 $k$ に対して

$$\mathsf{P}(|X - \mu| \geq k) \leq \sigma^2/k^2.$$

(3) (シュワルツの不等式)

$$\{\mathsf{Cov}(X, Y)\}^2 \leq \mathsf{Var}(X)\mathsf{Var}(Y)$$

# 特性関数 (1/4)

**定義 4.1 (特性関数)**

$X$ : 確率変数

$$\varphi_X(t) = \mathsf{E}[e^{itX}] = \begin{cases} \displaystyle\sum_x e^{itx} f_X(x) & \text{離散型} \\ \displaystyle\int_{-\infty}^{\infty} e^{itx} f_X(x)dx & \text{連続型} \end{cases}$$

# 特性関数 (2/4)

## 定理 4.1

(1) $\varphi_X(0) = 1$.

(2) $|\varphi_X(t)| \leq 1$.

(3) $\varphi_X(t)$ は一様連続である.

(4) $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$. ただし, $a$, $b$ は定数.

(5) $\mathrm{E}(|X|^n) < \infty$ ならば, $\varphi_X(t)$ は $C^n$ クラス (連続な $n$ 次導関数が存在) で

$$\left.\frac{d^n}{dt^n}\varphi_X(t)\right|_{t=0} = i^n \mathrm{E}(X^n).$$

# 特性関数 (3/4)

## 定義 4.2 (多次元分布の特性関数)

$\boldsymbol{X}$ を $p$ 次元の確率変数, $\boldsymbol{t} \in \mathbb{R}^p$ とし

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}, \quad \boldsymbol{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_p \end{pmatrix}$$

とする. このとき

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathsf{E}\{\exp(i\boldsymbol{t}'\boldsymbol{X})\} = \mathsf{E}\left\{\prod_{j=1}^{p} \exp(it_j X_j)\right\}$$

を $\boldsymbol{X}$ の特性関数という.

# 特性関数 (4/4)

**定理 4.2**

次の右辺の期待値が存在すれば等式が成り立つ.

$$\left.\frac{\partial^{n_1+\cdots+n_p}}{\partial t_1^{n_1}\cdots\partial t_p^{n_p}}\varphi_{\boldsymbol{X}}(\boldsymbol{t})\right|_{\boldsymbol{t}=\boldsymbol{0}} = i^{(n_1+\cdots+n_p)}\mathsf{E}(X_1^{n_1}\cdots X_p^{n_p}).$$

# 分布と特性関数 (1/2)

## 定理 4.3 (反転公式)

確率変数 $X$ の分布関数, 特性関数をそれぞれ, $F_X$, $\varphi_X$ とする. このとき, $F_X$ の連続点 $a, b$ $(a < b)$ において

$$F_X(b) - F_X(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)dt$$

が成り立つ.

## 定理 4.4 (一意性定理)

確率変数 $X_1$, $X_2$ の確率分布をそれぞれ $\mu_1$, $\mu_2$ とする. また, それらの特性関数をそれぞれ $\varphi_1$, $\varphi_2$ とする. このとき

$$\varphi_1 = \varphi_2 \Leftrightarrow \mu_1 = \mu_2$$

が成り立つ.

# 特性関数と分布 (2/2)

## 例 4.4 (再生性)

$X_1 + X_2$ の特性関数を求めることにより, 次の結果が得られる.

(1) $X_1 \sim B(m, p)$, $X_2 \sim B(n, p)$, $X_1$ と $X_2$ は互いに独立
$\Rightarrow X_1 + X_2 \sim B(m + n, p)$.

(2) $X_1 \sim p(\lambda_1)$, $X_2 \sim p(\lambda_2)$, $X_1$ と $X_2$ は互いに独立
$\Rightarrow X_1 + X_2 \sim p(\lambda_1 + \lambda_2)$.

(3) $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $X_1$ と $X_2$ は互いに独立
$\Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

## 注 4.2

$X_1$, $X_2$ が独立で, それぞれの分布が, ある分布族に含まれるとき, $X_1 + X_2$ の分布も同じ分布族に含まれるならば, その分布族は再生性を持つという.