

確率・統計 B

若木宏文

wakaki@hiroshima-u.ac.jp

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

2019.10.2

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 (3/5)

定義 3 (確率測度)

P が (Ω, \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 $0 \leq P(A) \leq 1$.

(P2) $P(\Omega) = 1$.

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

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

確率空間の定義と性質 (4/5)

確率測度の性質 1

- (1) $P(\emptyset) = 0$.
- (2) $P(A^c) = 1 - P(A)$.
- (3) $A \subset B \Rightarrow P(A) \leq P(B)$.
- (4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- (5) $A_i \cap A_j = \emptyset \ (i \neq j) \Rightarrow P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.
- (6) $P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$.

確率空間の定義と性質 (5/5)

確率測度の性質 2

- (1) $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.
- (2) $A_n \subset A_{n+1} \ (n = 1, 2, \dots) \Rightarrow P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- (3) $A_n \supset A_{n+1} \ (n = 1, 2, \dots) \Rightarrow P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- (4) $\overline{\lim}_{n \rightarrow \infty} A_n = \underline{\lim}_{n \rightarrow \infty} A_n$
 $\Rightarrow P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.

ただし

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

確率変数の分布 (1/9)

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 σ -集合体

定義 4 (確率変数)

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

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

定義 5 (確率変数の分布)

(Ω, \mathcal{B}, 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 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.

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

確率変数の分布 (2/9)

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 σ -集合体

定義 6 (n 次元確率変数 (確率ベクトル))

$\mathbf{X} = (X_1, \dots, X_n)$ の各成分 X_i が確率空間 (Ω, \mathcal{B}, P) 上の確率変数であるとき \mathbf{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)

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

$$P_{\mathbf{X}}(A) = P(\mathbf{X}^{-1}(A)) \quad (A \in \mathbb{B}_n)$$

確率変数の分布 (3/9)

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

$$\begin{aligned} F_{\mathbf{X}}(x_1, \dots, x_n) &= \mathbf{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= \mathbf{P}_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_n]) \end{aligned}$$

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

確率変数の分布 (4/9)

定義 8 (連続型分布)

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

あるいは

$$P_{\mathbf{X}}(A) = \int_A f_{\mathbf{X}}(x_1, \dots, x_n) dx_1 \cdots dx_n$$

と表されるとき, \mathbf{X} , あるいは $P_{\mathbf{X}}$ は連続型であるという. このとき, $f_{\mathbf{X}}$ 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 (5/9)

例 1 (正規分布)

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

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

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

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 Σ の n 次元正規分布とよび $N_n(\boldsymbol{\mu}, \Sigma)$ と表す. ただし

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$$

確率変数の分布 (6/9)

定義 9 (離散型分布)

\mathbb{R}^n 内の点列 x_1, x_2, \dots と, 実数列 p_1, p_2, \dots が存在して

$$P_{\mathbf{X}}(A) = \sum_{i: x_i \in A} p_i$$

と表されるとき, \mathbf{X} , あるいは $P_{\mathbf{X}}$ は離散型であるという.

$$f_{\mathbf{X}}(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)$$

を確率密度関数 (確率関数) とよぶ. このとき,

$$f_{\mathbf{X}}(x_1, \dots, x_n) = \begin{cases} p_i & (x_1, \dots, x_n) = x_i \quad (i = 1, 2, \dots) \\ 0 & \text{その他} \end{cases}$$

である.

確率変数の分布 (7/9)

例 2 (2 項分布)

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

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)

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

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 (8/9)

条件付き分布

X : m 次元確率変数, Y : n 次元確率変数

$f_{(X,Y)}(\mathbf{x}, \mathbf{y})$: 同時確率 (密度) 関数

$X = \mathbf{x}$ が与えられたときの Y の条件付き分布の確率 (密度) 関数は

$$f_{Y|X}(\mathbf{y} | \mathbf{x}) = \frac{f_{(X,Y)}(\mathbf{x}, \mathbf{y})}{f_X(\mathbf{x})},$$

$$f_X(\mathbf{x}) = \begin{cases} \int_{\mathbb{R}^n} f_{(X,Y)}(\mathbf{x}, \mathbf{y}) d\mathbf{y} & (\text{連続型}) \\ \sum_{\mathbf{y} \in D_Y} f_{(X,Y)}(\mathbf{x}, \mathbf{y}) & (\text{離散型}) \end{cases}$$

(D_Y : Y の取りうる点全体からなる集合)

確率変数の分布 (9/9)

確率変数の独立性

(X_1, \dots, X_n) の確率密度関数を $f_{\mathbf{X}}(x_1, \dots, x_n)$,
 X_1, \dots, X_n の周辺確率密度関数を, それぞれ $f_1(x_1), \dots, f_n(x_n)$
とする.

X_1, \dots, X_n が互いに独立であるための必要十分条件は

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

期待値の定義と性質 (1/4)

定義 10 (期待値)

\mathbf{X} : n 次元確率変数

$g(\mathbf{x})$: \mathbb{R}^n 上の実数値 (ボレル可測) 関数

に対して

$$E(g(\mathbf{X})) = \begin{cases} \int_{\mathbb{R}^n} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} & (\text{連続型の場合}) \\ \sum_{\mathbf{x} \in D_{\mathbf{X}}} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) & (\text{離散型の場合}) \end{cases}$$

を $g(\mathbf{X})$ の期待値とよぶ.

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

条件付き期待値 (2/4)

定義 11 (条件付き期待値)

\mathbf{X} : m 次元確率変数, \mathbf{Y} : n 次元確率変数

$g(\mathbf{y})$: \mathbb{R}^n 上の実数値 (ボレル可測) 関数

に対して

$$E(g(\mathbf{Y}) | \mathbf{X} = \mathbf{x}) = \begin{cases} \int_{\mathbb{R}^n} g(\mathbf{y}) f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} | \mathbf{x}) d\mathbf{y} & (\text{連続型の場合}) \\ \sum_{\mathbf{y} \in D_{\mathbf{Y}}} g(\mathbf{y}) f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} | \mathbf{x}) & (\text{離散型の場合}) \end{cases}$$

を $\mathbf{X} = \mathbf{x}$ が与えられたときの $g(\mathbf{Y})$ の期待値とよぶ.

$h(\mathbf{x}) = E(g(\mathbf{Y}) | \mathbf{X} = \mathbf{x})$ は \mathbf{x} の関数となるが, $h(\mathbf{X})$ を $E(g(\mathbf{Y}) | \mathbf{X})$ と表す. このとき

$$E[E(g(\mathbf{Y}) | \mathbf{X})] = E(g(\mathbf{Y}))$$

期待値の定義と性質 (3/4)

平均の性質

- (1) $E(aX) = aE(X)$.
- (2) $E(X + Y) = E(X) + E(Y)$.
- (3) $X \leq Y \Rightarrow E(X) \leq E(Y)$.
- (4) X と Y が独立 $\Rightarrow E\{g(X)h(Y)\} = E\{g(X)\}E\{h(Y)\}$

分散の性質

- (1) $\text{Var}(X + a) = \text{Var}(X)$.
- (2) $\text{Var}(aX) = a^2\text{Var}(X)$.
- (3) $\text{Var}(X) = E(X^2) - E(X)^2$.
- (4) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$
- (5) X と Y が独立 $\Rightarrow \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

期待値の定義と性質 (4/4)

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

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

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

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

- (3) (シュワルツの不等式)

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

特性関数 (1/4)

定義 4.1 (特性関数)

X : 確率変数

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

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

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

特性関数 (3/4)

定義 4.2 (多次元分布の特性関数)

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

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

とする. このとき

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

を \mathbf{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.

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

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, φ_X とする. このとき, F_X の連続点 a, b ($a < b$) において

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

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 μ_1, μ_2 とする. また, それらの特性関数をそれぞれ φ_1, φ_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.