確率・統計 A

若木宏文

wakaki@math.sci.hiroshima-u.ac.jp

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

2015.7.14

table of contents

事象と確率

確率空間, σ -集合体, 確率測度, 条件付き確率, 独立性

確率変数と分布

確率変数, 分布, 分布関数, 離散型・連続型, 独立性

平均と特性量

平均, 分散, 特性量, 条件付き分布, 条件付き期待値

特性関数

モーメントと微係数, 一意性定理

定義 1 (確率空間)

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

- Ω : 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
- $\mathcal{B}$: σ -集合体. 確率の定義域
- P : $\mathcal{B}$ 上で定義された確率測度

定義 1.1 (σ -集合体)

標本空間 Ω の部分集合の集まり $\mathcal{B}$ で次の条件 (B1), (B2), (B3) を満たすものを Ω 上の σ -集合体 (σ -field) または, σ -加法族 という.

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

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

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

σ -集合体の性質

定理 1.1

$\mathcal{B}$ を Ω の σ -集合体とする. このとき

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

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

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

(4) $A_1, A_2, \dots \in \mathcal{B} \Rightarrow \varliminf_{n \rightarrow \infty} A_n, \overline{\varliminf_{n \rightarrow \infty} A_n} \in \mathcal{B}$.

確率測度

定義 1.2

$(\Omega, \mathcal{B})$ を可測空間とする. $\mathcal{B}$ 上で定義された集合関数 P で次の条件 (P1), (P2), (P3) を満たすものを $(\Omega, \mathcal{B})$ 上の**確率測度 (probability measure)**, または, 単に**確率 (probability)** という.

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

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

(P3) (**完全加法的**または**可算加法性**) $A_1, A_2, \dots \in \mathcal{B}$ で, $A_i \cap A_j = \emptyset$ ($i \neq j$) ならば

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

確率測度の性質 1

定理 1.2

P を $(\Omega, \mathcal{B})$ 上の確率とする。このとき、次が成り立つ。

(1) $P(\emptyset) = 0$.

(2) $A \in \mathcal{B}$ に対し $P(A^c) = 1 - P(A)$.

(3) (単調性) $A, B \in \mathcal{B}, A \subset B \Rightarrow P(A) \leq P(B)$.

(4) (加法公式) $A, B \in \mathcal{B} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(5) (有限加法性) $A_i \in \mathcal{B}, i = 1, \dots, n, A_i \cap A_j = \emptyset (i \neq j)$

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

(6) (有限劣加法性) $A_i \in \mathcal{B}, i = 1, \dots, n$

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

確率測度の性質 2

定理 1.3

P を $(\Omega, \mathcal{B})$ 上の確率とし, $\{A_n\}$ を $\mathcal{B}$ に属する事象列とする. このとき, 次が成り立つ.

- (1) (劣加法性) $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.
- (2) $\{A_n\}$ が単調増加列のとき, $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- (3) $\{A_n\}$ が単調減少列のとき, $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- (4) (連続性) $\overline{\lim}_{n \rightarrow \infty} A_n = \underline{\lim}_{n \rightarrow \infty} A_n$ のとき

$$P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n).$$

ただし

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

定理 1.6

M を Ω の任意の部分集合族とする. このとき M を含む最小の σ -集合体が一意的に存在する. これを $\sigma[M]$ と書き, M から生成される Ω 上の σ -集合体という.

[証明]

$\mathcal{M}$ を M を含む Ω 上の σ -集合体 $\mathcal{B}$ の全体とし,

$$\mathcal{B}_0 = \bigcap_{\mathcal{B} \in \mathcal{M}} \mathcal{B}$$

と定義すると,

$$M \subset \mathcal{B}_0, \quad \mathcal{B}_0 \in \mathcal{M}, \quad \mathcal{B} \in \mathcal{M} \Rightarrow \mathcal{B}_0 \subset \mathcal{B}$$

が成り立つ.

ボレル集合

1次元ボレル集合体

$$\mathbb{B}_1 = \sigma[\{(a, b); -\infty < a < b < \infty\}]$$

$(a, b), [a, b], [a, b), (-\infty, b], (-\infty, b), [a, \infty), (a, \infty), \{a\}$ および、すべての開集合、閉集合はボレル集合である。

m 次元ボレル集合体

$$\mathbb{B}_m = \sigma[\{(a_1, b_1) \times \cdots \times (a_m, b_m); -\infty < a_k < b_k < \infty \ k = 1, \dots, m\}]$$

m -次元ボレル集合体は、 m 個の1次元ボレル集合体の直積、すなわち、

$$\begin{aligned} \mathbb{B}_m &= \mathbb{B}_1 \times \cdots \times \mathbb{B}_1 \quad (m \text{ 個}) \\ &= \sigma[\{A_1 \times \cdots \times A_m; A_i \in \mathbb{B}_1, i = 1, \dots, m\}] \end{aligned}$$

である。

条件付き確率の定義

定義 1.3

$(\Omega, \mathcal{B}, P)$ を確率空間, $A, B \in \mathcal{B}$ とし, $P(B) \neq 0$ とする. このとき

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

を事象 B が与えられたときの事象 A の条件付き確率という.

定理 1.7

$(\Omega, \mathcal{B}, P)$ を確率空間, $A, B \in \mathcal{B}$ とし, $P(B) \neq 0$ とする. B を固定し, $A \in \mathcal{B}$ に対して

$$P_B(A) = P(A|B)$$

とおく. このとき, P_B は $(\Omega, \mathcal{B})$ 上の確率である.

乗法公式

$$\begin{aligned}
 P(A_1 \cap A_2) &= P(A_1)P(A_2|A_1), \\
 P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\
 P(A_1 \cap A_2 \cap \cdots \cap A_n) \\
 &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\
 &\quad \times P(A_4|A_1 \cap A_2 \cap A_3) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}).
 \end{aligned}$$

定理 1.8 (ベイズの公式)

事象列 $\{B_i\}$ は互いに素で, $B_i \in \mathcal{B}, P(B_i) > 0, i = 1, 2, \dots$, かつ

$$\bigcup_{i=1}^{\infty} B_i = \Omega$$

とする. このとき, 任意の $A \in \mathcal{B}, P(A) > 0$ に対して

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{\infty} P(B_j)P(A|B_j)}.$$

n 個の事象の独立性

定義 1.5

n 個の事象 $A_1, \dots, A_n$ が独立であるとは, これらから任意に m 個 $A_{i_1}, A_{i_2}, \dots, A_{i_m}$ ($1 \leq i_1 < \dots < i_m \leq n$) をとるとき

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_m})$$

が満たされることをいう.

任意の事象族 $\{A_\lambda; \lambda \in \Lambda\}$ について, その任意の有限部分族が独立であるとき, 事象族 $\{A_\lambda; \lambda \in \Lambda\}$ は独立であるという.

定義 2.1

$(\Omega, \mathcal{B}, P)$ を確率空間とする. このとき, Ω 上の実数値関数 $X(\omega)$ が任意の実数 a に対して

$$\{\omega; X(\omega) \leq a\} \in \mathcal{B}$$

を満たすとき, 関数 $X(\omega)$ を $(\Omega, \mathcal{B}, P)$ または $(\Omega, \mathcal{B})$ の**確率変数**という.

定理 2.1

X を Ω 上の実数値関数とする. 次の条件 (1), (2) は同値である.

- (1) X が $(\Omega, \mathcal{B})$ 上の確率変数である.
- (2) 任意の $A \in \mathbb{B}_1$ に対して

$$X^{-1}(A) = \{\omega; X(\omega) \in A\} \in \mathcal{B}.$$

確率変数であるかどうかは $\mathcal{B}$ に依存している.

定理 2.2

X を $(\Omega, \mathcal{B})$ 上の確率変数とし, 任意の $A \in \mathbb{B}_1$ に対して

$$P_X(A) = P(X^{-1}(A))$$

とすると, P_X は $(\mathbb{R}, \mathbb{B}_1)$ 上の確率である.

定理 2.2 の P_X を X の分布 (測度) という.

$(\Omega, \mathcal{B}, P)$ 上で定義された k 個の確率変数 $X_1, \dots, X_k$ を成分とするベクトル $\mathbf{X} = (X_1, \dots, X_k)'$ を k 次元確率ベクトルと呼ぶ. $A \in \mathbb{B}_k$ に対して

$$P_{\mathbf{X}}(A) = P(\mathbf{X}^{-1}(A))$$

を, 確率ベクトル $\mathbf{X}$ の分布 よ呼ぶ.

k 次元確率変数 $\mathbf{X} = (X_1, \dots, X_k)'$ と, $\mathbf{x} = (x_1, \dots, x_k)' \in \mathbb{R}^k$ に対して

$$\begin{aligned} F(x_1, \dots, x_k) &= \mathbf{P}(X_1 \leq x_1, \dots, X_k \leq x_k) \\ &= \mathbf{P}(\bigcap_{i=1}^k \{\omega; X_i(\omega) \leq x_i\}) \end{aligned}$$

によって定義される関数 $F(\mathbf{x})$ を $\mathbf{X}$ (同時) 分布関数 という.
分布関数と分布は 1 対 1 に対応する

$$F(x_1, \infty, \dots, \infty) = \mathbf{P}(X_1 \leq x_1)$$

となるが, これを X_1 の (周辺) 分布関数 という. (他の成分も同様)

確率変数 X の分布関数を F とする.

定理 2.3

任意の実数 a, b ($a < b$) に対して

$$P(a < X \leq b) = F(b) - F(a)$$

定理 2.4

分布関数 $F(x)$ は次の性質をもつ

(F0) $0 \leq F(x) \leq 1$,

(F1) (単調性) $x_1 < x_2$ ならば $F(x_1) \leq F(x_2)$,

(F2) (右連続性) $F(x+0) = F(x)$,

(F3) $F(-\infty) = 0$, $F(\infty) = 1$.

$\mathbf{X} = (X_1, \dots, X_k)'$ の分布関数を F とする.

定理 2.5

任意の実数 a_i, b_i ($a_i < b_i$), $i = 1, \dots, k$ に対して

$$\begin{aligned} P(a_1 < X_1 \leq b_1, \dots, a_k < X_k \leq b_k) \\ = \Delta_{x_1}(a_1, b_1) \cdots \Delta_{x_k}(a_k, b_k) F(x_1, \dots, x_k) \end{aligned}$$

が成り立つ. ただし

$$\begin{aligned} \Delta_{x_i}(a_i, b_i) F(x_1, \dots, x_k) \\ = F(x_1, \dots, x_{i-1}, b_i, x_{i+1}, \dots, x_k) - F(x_1, \dots, x_{i-1}, a_i, x_{i+1}, \dots, x_k). \end{aligned}$$

とくに, $n = 2$ とすると

$$P(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2).$$

定理 2.6

k 次元確率変数 $\mathbf{X} = (X_1, \dots, X_k)'$ の分布関数 $F(x_1, \dots, x_k)$ は次の性質をもつ.

(F0) $0 \leq F(x_1, \dots, x_k) \leq 1.$

(F1) 任意の実数 a_i, b_i ($a_i < b_i$), $i = 1, \dots, k$ に対して

$$\Delta_{x_1}(a_1, b_1) \cdots \Delta_{x_k}(a_k, b_k) F(x_1, \dots, x_k) \geq 0$$

が成り立つ.

(F2) (右連続性) $F(x_1 + 0, \dots, x_k + 0) = F(x_1, \dots, x_k).$

(F3) $F(x_1, \dots, x_{i-1}, -\infty, x_{i+1}, \dots, x_k) = 0, F(\infty, \dots, \infty) = 1.$

多次元離散型分布と連続型分布

k 次元確率変数 $\mathbf{X} = (X_1, \dots, X_k)'$ 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, $\mathbf{X}$ は **離散型**確率変数, また, $\mathbf{X}$ の分布を **離散型分布**という.

離散型分布は, 確率関数

$$f(\mathbf{x}) = f(x_1, \dots, x_k) = P(X_1 = x_1, \dots, X_k = x_k), \quad \mathbf{x} \in \mathbb{R}^k$$

によって決定される.

k 次元確率変数 $(X_1, \dots, X_k)'$ 分布関数が

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

と表されるとき, $\mathbf{X} = (X_1, \dots, X_k)'$ は **連続型**確率変数, また, $\mathbf{X}$ の分布を **連続型分布**という.

$f(x_1, \dots, x_k)$ を $\mathbf{X}$ または, $\mathbf{X}$ の分布の**確率密度関数**という.

多次元確率関数の性質

確率関数 (離散型)

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

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

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

$$\text{同時分布関数: } F(z_1, \dots, z_k) = \sum_{x_{1j_1} \leq z_1} \cdots \sum_{x_{kj_k} \leq z_k} f(x_{1j_1}, \dots, x_{kj_k})$$

$$\text{周辺分布の確率関数: } f_{X_1}(x_1) = \sum_{j_2=1}^{\infty} \cdots \sum_{j_k=1}^{\infty} f(x_1, x_{2j_2}, \dots, x_{kj_k})$$

多次元確率密度関数の性質

確率密度関数 (連続型)

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

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

$$\text{同時分布関数: } F(x_1, \dots, x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(s_1, \dots, s_k) ds_1 \cdots ds_k$$

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

周辺分布の確率密度関数:

$$f_{X_1}(\mathbf{x}_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\mathbf{x}_1, x_2, \dots, x_k) dx_2 \cdots dx_k$$

連続型確率変数の関数の分布

$f_{X_1, X_2}(x_1, x_2)$: 2次元連続型確率変数 (X_1, X_2) の確率密度関数

$D = \{(x_1, x_2); f_{X_1, X_2}(x_1, x_2) > 0\}$ 上で

$$(y_1, y_2) = (u_1(x_1, x_2), u_2(x_1, x_2))$$

の関数の逆関数

$$(x_1, x_2) = (v_1(y_1, y_2), v_2(y_1, y_2))$$

が C^1 級であり、ヤコビ行列式

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} \neq 0$$

ならば $(Y_1 = u_1(X_1, X_2), Y_2 = u_2(X_1, X_2))$ の確率密度関数は

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(v_1(y_1, y_2), v_2(y_1, y_2)) |J|$$

定義 2.3

$X_1, \dots, X_n$ を $(\Omega, \mathcal{B}, P)$ 上の確率変数とする. 任意の区間 $I_1 = (a_1, b_1], \dots, I_n = (a_n, b_n]$ に対して, n 個の事象

$$\{\omega; X_1(\omega) \in I_1\}, \dots, \{\omega; X_n(\omega) \in I_n\}$$

が互いに独立であるとき, $X_1, \dots, X_n$ は互いに独立であるという. ただし, $-\infty \leq a_k < b_k < \infty$ ($k = 1, \dots, n$)

確率変数の独立性の必要十分条件

$F_{X_1, \dots, X_n}(x_1, \dots, x_n)$: $X_1, \dots, X_n$ の同時分布関数

$F_{X_i}(x_i)$: X_i の周辺分布関数 ($i = 1, \dots, n$)

$f_{X_1, \dots, X_n}(x_1, \dots, x_n)$: $X_1, \dots, X_n$ の同時確率 (密度) 関数

$f_{X_i}(x_i)$: X_i の周辺 (密度) 確率 ($i = 1, \dots, n$)

$\phi_{X_1, \dots, X_n}(t_1, \dots, t_n)$: $(X_1, \dots, X_n)'$ の特性関数

ϕ_{X_i} : X_i の特性関数 ($i = 1, \dots, n$)

このとき, 以下は同値

(1) $X_1, \dots, X_n$ は互いに独立

(2) $\forall x_i \in \mathbb{R}, (i = 1, \dots, n)$ に対して

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

(3) $\forall x_i \in \mathbb{R}, (f_{X_i}(x_i) > 0, i = 1, \dots, n)$ に対して

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

(4) $\forall t_i \in \mathbb{R}, (i = 1, \dots, n)$ に対して

$$\phi_{X_1, \dots, X_n}(t_1, \dots, t_n) = \phi_{X_1}(t_1) \cdots \phi_{X_n}(t_n)$$

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 のとりうる値の集合が可算集合 $D = \{x_1, x_2, \dots\}$ であるとき

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

を X の**平均**, あるいは**期待値**という.

このとき, $g(X)$ の期待値は

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

で与えられる.

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.

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

を X の **平均**, あるいは **期待値** という.

このとき, $g(X)$ の期待値は

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

で与えられる.

多変量の場合

定理 3.2

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

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

である.

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, $E(aX) = aE(X)$.
- (2) $E(X + Y) = E(X) + E(Y)$.
- (3) $X \leq Y \Rightarrow E(X) \leq E(Y)$.

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

$$E(XY) = E(X)E(Y).$$

定義 3.4

(分散, 歪度, 尖度)

(1) 分散 $\sigma^2 = \text{Var}(X) = E\{(X - \mu)^2\} = \alpha_2.$

(2) 歪度 $\beta_1 = E\left\{\left(\frac{X - \mu}{\sigma}\right)^3\right\} = \frac{\alpha_3}{\sigma^3}.$

(3) 尖度 $\beta_2 = E\left\{\left(\frac{X - \mu}{\sigma}\right)^4\right\} = \frac{\alpha_4}{\sigma^4}.$

定理 3.7 (分散の性質)

確率変数 X の分散に対して、次が成り立つ。ただし、 a は定数である。

- (1) $\text{Var}(X + a) = \text{Var}(X)$.
- (2) $\text{Var}(aX) = a^2\text{Var}(X)$.
- (3) $\text{Var}(X) = E(X^2) - E(X)^2$.

定理 3.8 (不等式)

- (1) (**Markov**) 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) (**Chevyshev**) 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

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

共分散と相関係数

定義 3.5

確率変数 X, Y に対して

$$\text{Cov}(X, Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$

を X と Y との**共分散**とよぶ。また

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

を X と Y との**相関係数**とよぶ。

$\rho(X, Y) = 0$ であるとき**無相関**という。

$\rho(X, Y)$ が正であるとき, X と Y の間には**正の相関**があるといい, 負であるときは**負の相関**があるという。

分散・共分散の基本性質

定理 3.9

X, Y, Z を確率変数, a を定数とする. このとき, 次が成り立つ.

- (1) $\text{Cov}(X, X) = \text{Var}(X)$.
- (2) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- (3) $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$.
- (4) $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$.
- (5) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.
- (6) X と Y が独立 $\Rightarrow \text{Cov}(X, Y) = 0$.
- (7) X と Y が有限な分散をもつとする. このとき

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

シュワルツの不等式

定理 3.10

確率変数 X, Y は有限な分散をもつとする. このとき

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

が成り立つ. したがって

$$|\rho(X, Y)| \leq 1.$$

等号が成り立つのは, 次の (1), (2), (3) のいずれかの場合である:

(1) $P(X = E(X)) = 1.$

(2) $P(Y = E(Y)) = 1.$

(3) 定数 a, b が存在して, $P(Y = a + bX) = 1.$

事象を与えたときの条件付き分布

$X : (\Omega, \mathcal{B}, P)$ 上の確率変数, $B \in \mathcal{B}$, $P(B) > 0$

$A \in \mathbb{B}_1$

事象 B を与えたときの, 事象 “ $X \in A$ ” の条件付き確率 :

$$P(X \in A | B) = \frac{P("X \in A" \cap B)}{P(B)}$$

B を固定して, $P_{X|B}(A) = P(X \in A | B)$ を A の関数とみると,
 $P_{X|B}$ は $(\mathbb{R}^1, \mathbb{B}_1)$ 上の確率となる.

$P_{X|B}$ を **B を与えたときの X の条件付き分布** とよぶ.

離散型分布の場合

(X, Y) : 2次元離散型確率変数,

$D = \{(x_i, y_j); i, j = 1, 2, \dots\}$: 取り得る値の集合

$f_{X,Y}(x, y)$: (X, Y) の確率関数, $f_Y(y)$: Y の周辺確率関数

確率関数

$$f_{X|Y}(x|y_j) \equiv \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)}$$

によって与えられる. 離散型分布を $Y = y_j$ を与えたときの X の条件付き分布とよぶ.

$f_{X|Y}(x|y_j)$ を $Y = y_j$ を与えたときの X の条件付き確率関数とよぶ.

$$E(X | Y = y_j) \equiv \sum_{i=1}^{\infty} x_i f_{X|Y}(x_i | y_j)$$

を $Y = y_j$ を与えたときの X の条件付き平均, あるいは条件付き期待値とよぶ.

連続型の場合

(X, Y) : 2次元連続型確率変数

$f_{X,Y}(x, y)$: (X, Y) の確率密度関数, $f_Y(y)$: Y の周辺確率密度関数

$f_Y(y) > 0$ である y に対して

$$f_{X|Y}(x|y) \equiv \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

によって定まる連続型分布を $Y = y$ を与えたときの X の条件付き分布, $f_{X|Y}(x|y)$ を $Y = y$ を与えたときの X の条件付き確率密度関数 とよぶ.

$$E(X | Y = y) \equiv \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

を $Y = y$ を与えたときの X の条件付き平均, あるいは条件付き期待値とよぶ.

条件付き平均の性質

定理 3.11

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

$f_{X|Y}(x|y) : Y = y$ を与えたときの X の条件付き確率 (密度) 関数

$$E(g(X)|Y = y) = \begin{cases} \sum_{j=1}^{\infty} g(x_j) f_{X|Y}(x_j|y), & \text{離散型のとき} \\ \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y) dx, & \text{連続型のとき} \end{cases}$$

次が成り立つ.

$g(y) = E(X|Y = y)$ とするとき, $g(Y)$ は確率変数となるが, これを $E(X|Y)$ と表す.

定理 3.12

$f_{X,Y}(x, y) : (X, Y)$ の同時確率密度関数 (確率関数) このとき

$$E\{E(X|Y)\} = E(X).$$

が成り立つ.

定義 4.1

X : 確率変数, $f(x)$: 確率密度関数, または確率関数

$$\begin{aligned} \varphi_X(t) &= \mathbf{E}(e^{itX}) \\ &= \begin{cases} \sum_x e^{itx} f(x), & \text{離散型} \\ \int_{-\infty}^{\infty} e^{itx} f(x) dx, & \text{連続型} \end{cases} \\ &= \begin{cases} \sum_x \cos(tx) f(x) + i \sum_x \sin(tx) f(x), & \text{離散型} \\ \int_{-\infty}^{\infty} \cos(tx) f(x) dx + i \int_{-\infty}^{\infty} \sin(tx) f(x) dx, & \text{連続型} \end{cases} \end{aligned}$$

を X の特性関数という。

定理 4.1

X の特性関数 $\varphi_X(t)$ について、次が成り立つ.

- (1) $\varphi_X(0) = 1$.
- (2) $|\varphi_X(t)| \leq 1$.
- (3) $\varphi_X(t)$ は一様連続である.
- (4) $\varphi_{cX+d}(t) = e^{itd}\varphi_X(ct)$. ただし, c, d は定数.
- (5) $E(|X|^n) < \infty$ ならば, $\varphi_X(t)$ は C^n 級 (連続な n 次導関数が存在) で

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

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

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

を $\mathbf{X}$ の**特性関数**という.

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。

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

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.