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

2014.5.12

目 次

1.7. 事象の独立性

事象の独立 (1/2)

2 つの事象 $A, B \in \mathcal{B}$ について

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B)$$

が成り立つとき, A と B は互いに独立であるという.

このとき, $0 < \mathsf{P}(B) < 1$ ならば

$$\mathsf{P}(A|B) = \mathsf{P}(A), \quad \mathsf{P}(A|B^c) = \mathsf{P}(A)$$

が成り立つ. $\Rightarrow$ A の条件付き確率は B の生起に無関係

同様に $0 < \mathsf{P}(A) < 1$

$\Rightarrow$ B の条件付き確率は A の生起に無関係

事象の独立 (2/2)

$$
\begin{aligned}
A \text{ と } B \text{ が独立} &\Leftrightarrow A \text{ と } B^c \text{が独立} \\
&\Leftrightarrow A^c \text{と } B \text{ が独立} \\
&\Leftrightarrow A^c \text{と } B^c \text{が独立}
\end{aligned}
$$

が成り立つ.

一般に $\mathsf{P}(A \cap B^c) = \mathsf{P}(A) - \mathsf{P}(A \cap B)$ であることを用いる.

(証明は板書で)

3 個の事象の独立性

定義 1.4

3 個の事象 A_1, A_2, A_3 が独立であるとは次の条件

$$\begin{aligned}
\mathsf{P}(A_1 \cap A_2 \cap A_3) &= \mathsf{P}(A_1)\mathsf{P}(A_2)\mathsf{P}(A_3) \\
\mathsf{P}(A_1 \cap A_2) &= \mathsf{P}(A_1)\mathsf{P}(A_2) \\
\mathsf{P}(A_2 \cap A_3) &= \mathsf{P}(A_2)\mathsf{P}(A_3) \\
\mathsf{P}(A_3 \cap A_1) &= \mathsf{P}(A_3)\mathsf{P}(A_1)
\end{aligned}$$

が満たされるときをいう.

例 1.14

コインを 2 回投げる試行
標本空間 $\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$
$A = \{(H,H),(H,T)\}$: 1 回目に表がでる事象
$B = \{(H,H),(T,H)\}$; 2 回目に表がでる事象
$C = \{(H,T),(T,H)\}$: 表と裏が 1 回づつ出る事象

$$\mathsf{P}(A \cap B) = \mathsf{P}(A \cap C) = \mathsf{P}(B \cap C) = \frac{1}{4}$$
$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = \frac{1}{2}$$

より, どの 2 つの事象も互いに独立

$$\mathsf{P}(A \cap B \cap C) = 0 \neq \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C) = \frac{1}{8}$$

n 個の事象の独立性

定義 1.5

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

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

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.

定理

定理 1.9

事象 $A_1, \cdot, A_n \in \mathcal{B}$ が独立のとき, その一部を余事象で置き換えて得られる事象族も独立である.
すなわち, $B_i\ (i = 1, \ldots, n)$ を A_i または A_i^c とするとき,
$B_1, \ldots, B_n$ は互いに独立である.

定理 1.10

事象 $A_1, \ldots, A_n$ が独立であるための必要十分条件は, $B_i = A_i$ または $B_i = A_i^c\ (i = 1, \ldots, n)$ のすべての組み合わせに対して

$$\mathsf{P}(B_1 \cap \cdots \cap B_n) = \mathsf{P}(B_1) \times \cdots \mathsf{P}(B_n)$$

が成り立つことである.

(証明は板書で)

第2章 確率変数と分布

2.1. 確率変数の定義

確率変数とは

試行の結果は必ずしも数値をとるとは限らない.
試行の各結果によって値が定まるある量 X に関心がある場合が多い.

例 1) 1 枚の銅貨を 2 回投げるとき, 表の出た回数 X に関心がある場合.
標本空間 $\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$ の各点に対して X は

$$X(H,H) = 2,\ X(H,T) = 1,\ X(T,H) = 1,\ X(T,T) = 0$$

例 2) ある試行による 事象 A の生起に関心がある場合

$$X(\omega) = \begin{cases} 1 & (\omega \in A) \\ 0 & (\omega \in A^c) \end{cases}$$

確率変数の定義

定義 2.1

$(\Omega, \mathcal{B}, \mathsf{P})$ 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 $(\Omega, \mathcal{B}, \mathsf{P})$ または $(\Omega, \mathcal{B})$ の確率変数 (random variable) という．

注意事項 1(1/2)

注 2.1

$\mathcal{B} = \wp(\Omega)$ の場合には, Ω 上の実数値関数 $X(\omega)$ は確率変数である. Ω が有限集合, あるいは, 可算集合の場合には通常 $\mathcal{B} = \wp(\Omega)$ と定めるので, すべての実数値関数は確率変数である.
一般の確率空間 $(\Omega, \mathcal{B}, \mathrm{P})$ では, 確率変数であるかどうかは $\mathcal{B}$ に依存している.

注意事項1(2/2)

注 2.2

銅貨を無限回投げる試行において, n 回目に表が出れば *1*, 裏が出れば *0* とする確率変数を X_n とする. このとき, 初めて表が出たときの銅貨を投げた回数 T は, 各 $\omega = (\omega_1, \ldots, \omega_n, \ldots)$ に対し

$$T(\omega) = \begin{cases} n, & X_j(\omega) = 0,\ j = 1, \ldots, n-1,\ X_n(\omega) = 1 \\ +\infty, & \text{その他} \end{cases}$$

と表され, $T(\omega) = +\infty$ となるのは毎回裏が出る場合に限られるから

$$\{\omega;\ T(\omega) = +\infty\} \subset \{\omega;\ X_1(\omega) = \cdots = X_n(\omega) = 0\}$$

であり, これらの事象の確率を考え $n \to \infty$ とすると

$$\mathsf{P}(\{\omega;\ T(\omega) = +\infty\}) \le \lim_{n\to\infty} 2^{-n} = 0$$

である. このように, 確率変数のとる値として, $+\infty$ を認めることによって統一的に議論することができる.

Ω 上の関数 $X(\omega)$ が実数, または, $+\infty, -\infty$ を値としてとる場合には, 定義 2.1 の条件に加えて

$$\{\omega;\ X(\omega) = +\infty\} \in \mathcal{B}, \quad \{\omega;\ X(\omega) = -\infty\} \in \mathcal{B}$$

を満たすものを確率変数とする.
しかし, 特に断らない限り, 確率変数といえば, 実数値をとるものとする.

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

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

(証明は板書で)

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

$$\mathsf{P}_X(A) = \mathsf{P}(X^{-1}(A))$$

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 確率変数 X によって誘導された確率分布 (測度) (induced probability distribution (measure)), あるいは, 単に 確率変数 X の分布 (distribution) という.

2.2. 分布関数

定義 2.2

X を $(\Omega, \mathcal{B}, \mathsf{P})$ 上の確率変数とする $\mathbb{R}$ 上で定義された関数

$$F_X(x) = \mathsf{P}(\{\omega;\ X(\omega) \leq x\})$$

を X の分布関数という.

以下, $F_X(x)$ を単に $F(x)$ と書く.

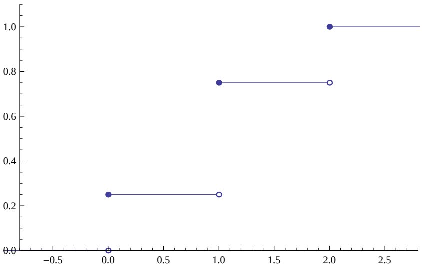
事象 $\{\omega;\ X(\omega) \leq x\}$ を単に, $''X \leq x''$ と書く. 確率 P, 確率分布 P_X と分布関数の関係は

$$F(x) = \mathsf{P}(X \leq x) = \mathsf{P}_X((-\infty, x])$$

例 2.1

(1) 銅貨を 2 回投げる試行で, 表が出た回数を X とする.

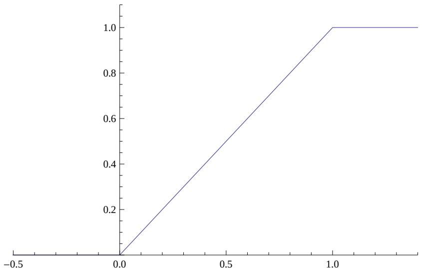
$$F(x) = \begin{cases} \mathrm{P}(\emptyset) = 0 & x < 0 \\ \mathrm{P}(\{(T,T)\}) = \frac{1}{4} & 0 \le x < 1 \\ \mathrm{P}(\{(T,T),(T,H),(H,T)\}) = \frac{3}{4} & 1 \le x < 2 \\ \mathrm{P}(\Omega) = 1 & 2 \le x \end{cases}$$



例 2.1 続き

(2) 区間 $(0, 1]$ からランダムに 1 つの実数値を選ぶ試行で選ばれた実数を X とする.

$$F(x) = \begin{cases} \mathsf{P}(\emptyset) = 0 & x \leq 0 \\ \mathsf{P}(0 < X \leq x) = x & 0 < x < 1 \\ \mathsf{P}(0 < X \leq 1) = 1 & 1 \leq x \end{cases}$$



定理 2.3

任意の実数 a, b $(a < b)$ に対して

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

(証明は板書で)

分布関数の性質

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

(証明は板書で)

注意事項 2

注 2.4 (分布と分布関数の 1 対 1 対応)

確率変数 X が与えられると, その分布関数が定義 2.2 で定義される. 逆に, 定理 2.4 を満たす関数 $F(x)$ が与えられると,

$$\mathsf{P}_X((a,b]) = F(b) - F(a)$$

となるような $(\mathbb{R}, \mathbb{B}_1)$ 上の分布 P_X が一意的に定義されることが知られている. (ルベーグ・スティルチェス測度)

2.3. 多次元確率ベクトルと分布

多次元データ (Multivariate data)

互いに関連 (相関) がある観測値の組が n 個あるようなデータセット

番号	国語	社会	数学	理科	英語
	テストの点数				
1	80	96	81	100	93
2	60	61	39	57	57
3	63	50	41	45	59
4	100	98	75	93	98
5	56	41	26	40	32
6	74	77	65	84	71
7	86	83	62	74	80
8	65	52	43	62	57
9	75	66	45	60	69
10	57	64	47	58	57

それぞれの個人の点数を観測ベクトル $\boldsymbol{x}_i$ とする.
ex) $\boldsymbol{x}_6 = (74, 77, 65, 84, 71)'$

観測行列 $\begin{bmatrix} \boldsymbol{x}_1' \\ \vdots \\ \boldsymbol{x}_n' \end{bmatrix}$ とする.

多次元確率変数

試行の結果に対して, 複数個の変数を同時に解析したい.

例. 全国学力試験受験者を無作為に選んで, その国語, 社会, 数学, 理科, 英語の得点を調べる.

$\Rightarrow$ Ω の各要素 (標本点) に対して, k 個の変数 $X_1, \ldots, X_k$ の値が定まる.

$(\Omega, \mathcal{B}, \mathrm{P})$ 上で定義された k 個の確率変数 $X_1, \ldots, X_k$ を成分とするベクトル $\boldsymbol{X} = (X_1, \ldots, X_k)'$ を k 次元確率ベクトル (k 次元確率変数) と呼ぶ.

多次元確率変数の分布関数

k 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_k)'$ を考える.
$\boldsymbol{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$ に対して

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

によって定義される関数 $F(\boldsymbol{x})$ を $\boldsymbol{X}$ の分布関数, あるいは同時分布関数, 結合分布関数 (joint distribution function) という.

1 次元確率変数の場合と同様に k 次元確率ベクトル $\boldsymbol{X}$ によって $(\mathbb{R}^k, \mathbb{B}_k)$ 上に

$$\begin{aligned} \mathsf{P}_{\boldsymbol{X}}(A_1 \times \cdots \times A_k) &\equiv \mathsf{P}(\boldsymbol{X}^{-1}(A_1 \times \cdots \times A_k)) \\ &= \mathsf{P}(\{\omega;\ X_1(\omega) \in A_1, \ldots, X_k(\omega) \in A_k\}) \end{aligned}$$

となる確率分布 $\mathsf{P}_{\boldsymbol{X}}$ が導入される.
これを k 次元確率ベクトル $\boldsymbol{X}$ の分布と呼ぶ.

例

例 2.2

$\mathbb{R}^2$ 上の区間 $\Omega = (0,1] \times (0,1]$ からランダムに点を選ぶ試行を考える．この試行によって選ばれた点を (X,Y) とする．このとき，(X,Y) の分布関数は

$$F(x,y) = \mathsf{P}(X \le x, Y \le y) = \begin{cases} 0, & x \le 0 \text{ or } y \le 0 \\ xy, & 0 < x \le 1,\ 0 < y \le 1 \\ x, & 0 < x \le 1,\ 1 < y \\ y, & 1 < x,\ 0 < y \le 1 \\ 1, & 1 < x,\ 1 < y \end{cases}$$

となる．

定理 2.5

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

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

が成り立つ. ただし

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

とくに, $n = 2$ とすると

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

(証明は板書で)

定理 2.6

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

(F0) $0 \le F(x_1, \ldots, x_k) \le 1$.

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

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

が成り立つ.

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

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

注 2.6

(F1)〜(F3) を満たす関数 $F(x_1, \ldots, x_k)$ は, *(F0)* も満たすことがわかるが, 逆に, *(F1)〜(F3)* を満たす関数 $F(x_1, \ldots, x_k)$ 分布関数となるような $(\mathbb{R}^k, \mathbb{B}_k)$ 上の分布が一意に存在する.

周辺分布

分布関数の定義において, i を除くすべての j について $x_j \to \infty$ とすると

$$\mathsf{P}(X_i \le x_i) = F(\infty, \ldots, x_i, \infty, \ldots, \infty)$$

が成り立ち, $F(\infty, \ldots, x_i, \infty, \ldots, \infty)$ は X_i の分布関数であることがわかる. これを 周辺分布関数 (marginal distribution function) という.

例 2.3

例 2.2 において, X の周辺分布関数は

$$F_X(x) = \mathsf{P}(X \le x) = \mathsf{P}(X \le x, Y < \infty) = \begin{cases} 0, & x \le 0 \\ x, & 0 < x \le 1 \\ 1, & 1 < x \end{cases}$$

で与えられる. Y の周辺分布関数も同様に求められる.