確率統計 復習

Kosuke Toda

目次

第 1 章

事象と確率

1.1　集合と事象

- 試行 (trials)
  - 実験や観測，調査の総称．
- 全事象 (total event)
  - 試行を行ったときに起こりうる全ての結果の集まり
- 母集団 (population)
  - 統計学では，全事象のことを試行の対象となる集団とみなす．
- 事象 (event)
  - ある性質を満たす結果の集まり

Ω をある集合 (全集合) とし，$\emptyset$ を空集合 (空事象) とする．可測集合の定義を以下で与える．

定義 1.1　Ω の部分集合を要素とする集合族 $\mathscr{A}$ が次の 3 条件を満たすとき，σ-加法族という．

1. $\Omega \in \mathscr{A}$, $\emptyset \in \mathscr{A}$.
2. $A \in \mathscr{A}$ ならば $A^{\mathrm{C}} \in \mathscr{A}$.
3. $A_n \in \mathscr{A}$ $(n = 1, 2, \ldots)$ ならば，$\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$.

また，$(\Omega, \mathscr{A})$ を可測空間，$\mathscr{A}$ の要素を可測集合という[*1]．

[*1] テキストでは，$\mathscr{A}$ (σ-加法族) を事象族，$\mathscr{A}$ の要素を事象と呼んでいる．

1.2　確率と確率空間

定義 1.2　事象族 $\mathscr{A}$ 上で定義された関数 P が次の 3 条件を満たすとき，確率 (または確率測度 (probability measure)) という．

(P1)　任意の事象 $A \in \mathscr{A}$ に対して，$0 \leq P(A) \leq 1$.

(P2)　$P(\emptyset) = 0,\ P(\Omega) = 1$.

(P3)　$A_n \in \mathscr{A}\,(n = 1, 2, \ldots)$ が互いに素 $(A_m \cap A_n = \emptyset\,(m \leq n))$ のとき[a]，

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

[a] $\sigma-$ 加法性 という．

このとき，$(\Omega, \mathscr{A}, P)$ を確率空間 (probability space) と呼ぶ．(P1)-(P3) から確率測度に関する基本的性質が導かれる:

1. $P(A^{\mathrm{C}}) = 1 - P(A)$
2. $A \subset B \Longrightarrow P(A) \leq P(B)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. $A_n \in \mathscr{A},\ A_n \supset A_{n+1}\,(n = 1, 2, \ldots) \Longrightarrow P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$
5. $A_n \in \mathscr{A},\ A_n \subset A_{n+1}\,(n = 1, 2, \ldots) \Longrightarrow P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$

1.3　事象の独立性と従属性

定義 1.3（独立性）　2 つの事象 A, B が独立 (independent) であるとは，

$$P(A \cap B) = P(A)P(B)$$

が成り立つことをいう．独立でないとき，従属 (dependent) であるという．

定義 1.4（条件付き確率）　2 つの事象 A, B があって，$P(A) > 0$ のとき，

$$P(B|A) \overset{\text{def}}{\coloneqq} \frac{P(A \cap B)}{P(A)}$$

を，A が与えられたときの B の条件付き確率という．

条件付き確率について，以下の性質が成り立つ．

1. $P(A \cap B) = P(A)P(B|A)$
2. A, B が独立のとき，$P(B|A) = P(B)$, $P(A|B) = P(A)$

定理 1.1（全確率の法則） $A_1, A_2 \ldots, A_n$ を互いに素な事象で，$\Omega = \bigcup_{i=1}^n A_i$ とする．$P(A_i) > 0\,(i = 1, 2, \ldots, n)$ のとき，任意の事象 B に対して，

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i) \tag{1.1}$$

が成立する．

証明 1（定理 1.1 の証明）

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (B \cap A_i)$$

および

$$(B \cap A_i) \cap (B \cap A_j) = \emptyset \ (i \neq j)$$

より，定義 1.2 から

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

が成立し，定義 1.4 より，(1.1) 式が成立する． □

事象 B が起こる前と後のことをそれぞれ事前と事後と呼ぶ．A_i に対して $P(A_i)$ は B が起こる前に与えられている確率であるから事前確率 (prior probability) といい，$P(A_i|B)$ は B が起こった後で与えられる確率であるから事後確率 (posterior probability) という．次のベイズの定理が成立する[*2]．

定理 1.2（ベイズの定理） $A_1, A_2, \ldots, A_n$ を互いに素な事象で，$\Omega = \bigcup_{i=1}^n A_i$ とする．$P(A_i) > 0\ (i = 1, 2, \ldots, n)$ のとき，任意の事象 B に対して

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)} \ \ (i = 1, 2, \ldots, n) \tag{1.2}$$

が成立する．

証明 2（定理 1.2 の証明） 定義 1.4 より，

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

条件付き確率の性質より，

$$P(A_i \cap B) = P(A_i)P(B|A_i).$$

[*2] ベイズの定理は，事前確率 $P(A_i)$, $P(B|A_i)$ $(i = 1, 2, \ldots, n)$ が求まれば，事後確率 $P(A_i|B)$ が導出できることを述べている．

定理 1.1 より，

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i) \tag{1.3}$$

であるから，(1.2) 式が成立する． □

第 2 章

確率変数と確率分布

2.1　母集団と標本

既知の無作為標本によって母集団の未知の性質を知ろうとすることを統計的推測 (statistical inference) という (図 2.1).

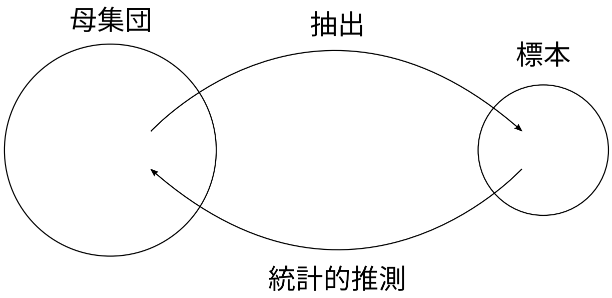

図 **2.1:** 統計的推測のイメージ

2.2　確率変数と確率分布

まず，離散型の確率変数とその分布を例に考える．X を 2 個のサイコロを投げたときに出た目の和を表す変数とする (表 2.1).

表 **2.1:** 確率変数の例

X の値	2	3	4	$\cdots$	12	計
確率	1/36	2/36	3/36	$\cdots$	1/36	1

この X のように，取りうる値に対して，確率が対応する変数を確率変数 (random variable) と呼ぶ．確率変数の取りうる値 (実現値) とその確率をペアにしたものを確率分布 (probability distribution) と呼ぶ.

一般に，確率変数 X の実現値を $x_1, x_2, \ldots, x_n, \ldots$，その確率を $p_1, p_2, \ldots, p_n, \ldots$ とすると，確率変数 X

の確率分布は，

$$P(X = x_k) = p_k \ \ (k = 1, 2, \ldots) \tag{2.1}$$

である．ただし，$\sum_{k=1}^{\infty} p_k = 1$, $p_k \geq 0$ $(k = 1, 2, \ldots)$ である．