

確からしさの考え方、確率

1 偶然な実験

統計学の理論展開を数学モデルとして取り扱うためには、確率の概念が必要となります。不確実性の数値化（モデル化）には、コイン投げ、さいころ投げを例として議論されますが、これには概念の定義に、(1) **再現性**（繰り返すことができる）と (2) **客観性**（多くの人が認める）をもつ対象を考えるからです。数学的には、「測度論」とよばれる面積や体積を定めるための道具をつかいます。公理的な確率論とよばれる由縁です。抽象的にモデル化し、つぎのように確率空間とよび、偶然の現象を数学的に定義します。

偶然によって起こると考えられる全ての結果全体 Ω を考えます。結果のあつまり、集合、を意味します。つぎにそのうちのいくつかの結果を事象とよび、これを集めたものを A と表します。厳密にはすこし厄介になるので、一般には大きさを測る、あるいは数え上げができるものと解釈します。実際には、 P という記号で、事象の大きさ（可能性）を数値で表現します。まず確率の大きさををはかる対象となる集合を事象とよび、全事象と空事象、和事象、積事象、余事象（補事象）を定義されているとします。**全事象** Ω は、すべての可能な結果全体で、当たり前のこと、この確率を 1 と定める。**空事象** $\emptyset$ は、起こり得ないことで、確率を 0 とします。**和事象** $A \cup B$ とは少なくともひとつ、いずれかが起こること、**積事象** AB , $A \cap B$ とは、すべてのものが同時に起こること、**余事象** $\bar{A}$, A^c とは、これが起こらない、否定を意味します。

確率の定義；3つの組 $(\Omega, \mathcal{A}, P)$ が確率空間とは加法性と連続性をみたす集合関数（集合から数値を対応）のことです。つぎを満たすとします；

(a) **基準化** $P(\emptyset) = 0, P(\Omega) = 1$

(b) **加法性** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(c) **連続性** $A_i, i = 1, 2, \dots$ が $A_1 \subset A_2 \subset \dots$, $\lim_i A_i = \bigcup_{i=1}^{\infty} A_i$ は

$$P(\lim_i A_i) = P(\bigcup_{i=1}^{\infty} A_i) = \lim_i P(A_i)$$

確率の基本定理：

(1) $A \subset B$ ならば、 $P(B) = P(A) + P(B \setminus A) = P(A) + P(B \cap A^c) \geq P(A)$

(2) $P(A^c) = 1 - P(A)$

(3) $P(\bigcup_{i=1}^n A_i) = \sum_i P(A_i) - \sum_{i < j} P(A_i A_j) + \sum_{i < j < k} P(A_i A_j A_k) \cdots + (-1)^{n+1} P(A_1 A_2 \cdots A_n)$, ここで $\sum_{i < j}$ は $i < j$ を満たすすべての (i, j) の組にわたる和をとる。他も同様。

(4) $A_i, i = 1, 2, \dots$ が互いに素、すなわち $A_i \cap A_j = \emptyset (i \neq j)$ ならば $P(\bigcup_i A_i) = \sum_i P(A_i)$

(5) de Morgan の法則より

$$P(\bigcup_i A_i) = 1 - P(\bigcap_i A_i^c)$$

少なくとも一つが起こる確率の計算は、1 からすべてがまったく起こらない確率を引けばよい。

$$P(\bigcap_i A_i) = 1 - P(\bigcup_i A_i^c)$$

2 条件付き確率と事象の独立

まず条件付き確率を定義する。ある事象が起こったときに、もう一つの事象にどの程度影響を及ぼしているかの尺度を表す。記号では $P(B|A)$ と表す。

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

ただし $P(A) > 0$ とする。事象 A は多少とも起こるものと前提条件をする。これを A が起こったときの B の起こる条件付き確率という。

$$(1) P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

$$(2) P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B), \quad P(A \cap B) > 0$$

(3) **ベイズの定理**; 事象列 $A_1, A_2, \dots, A_n$ を Ω の分割、つまり互いに素（どの2つをとっても積事象が空事象）であり、これらすべての和集合が全事象になるとき、つぎの条件付き確率についての関係式が成り立つ。

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_k P(B|A_k)P(A_k)}$$

事象 $A_k, k = 1, 2, \dots, n$ を原因と考え、結果の事象 B の条件付き確率がわかっていれば、結果からどの原因により起こっているかを考える。事前確率 $P(A_k), k = 1, 2, \dots, n$ から事後確率 $P(A_j|B)$ を計算する式を意味する。

独立事象の定義; 2つの事象の間について、独立とは一方の起こる可能性に影響を与えないこと。 $P(B|A) = P(B)$ を書き換えて

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

が成り立つとき。一般に $A_i, i = 1, 2, \dots, n$ が独立であるとは、

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P(A_2) \cdots P(A_n)$$

で n 個のすべての部分系（1組のかたまりと考えて）が独立であるとき、で定義する。

問 2.1

与えられた確率 $P(A) = 1/3, P(B) = 1/4, P(AB) = 1/6$ に対して、つぎの確率を計算しなさい。ただし $\bar{A} = A$ の補事象。

$$(i) P(\bar{A}) \quad (ii) P(\bar{A} \cup B) \quad (iii) P(A \cup \bar{B}) \quad (iv) P(A\bar{B}) \quad (v) P(\bar{A} \cup \bar{B})$$

問 2.2

事象列 $A_n, n \geq 1$ に対して

$$P(A_1 A_2 \cdots A_n) \geq \sum_{i=1}^n P(A_i) - (n-1) = 1 - \sum_{i=1}^n P(\bar{A}_i)$$

問 2.3

つぎを示せ。また一般化 (n 個) の事象列についてはどうなるか。

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC)$$

問 2.4

2つの事象 (A, B) が独立ならば、 $(A, \bar{B}), (\bar{A}, B), (\bar{A}, \bar{B})$ も独立となることを示せ。

問 2.5

つぎの等式を示せ。

$$P(AB) - P(A)P(B) = P(\bar{A})P(B) - P(\bar{A}B) = P(A)P(\bar{B}) - P(A\bar{B}) = P(\bar{A}\bar{B}) - P(\bar{A})P(\bar{B})$$

問 2.6

つぎの4つの命題は同値であることを示しなさい。

- (a) 事象 A と B は独立 (b) 事象 A^c と B は独立
(c) 事象 A と B^c は独立 (d) 事象 A^c と B^c は独立

問 2.7

もし $P(A|B) > P(A)$ ならば, $P(B|A) > P(B)$ を示せ。(ヒント: $\frac{P(A|B)}{P(B|A)} = \frac{P(A)}{P(B)}$ を示せ。)

問 2.8

事象列 $A_1, A_2, \dots, A_n$ が独立で $P(A_i) = p, i = 1, 2, \dots, n$ とおくとき,

- (a) 少なくとも一つの事象が起こる確率は?
(b) 少なくとも m 個の事象が起こる確率は?
(c) ちょうど m 個の事象が起こる確率は?

問 2.9

箱の中には $1, 2, \dots, n$ と数字の記されたボールが入っている。 r 個のボールをとりだすとき, 最大の数字が m となる確率を計算せよ。

- (a) もとに戻してからボールを取り出すとき (復元抽出)。
(b) もとには戻さずにボールを取り出すとき (非復元抽出)。