

情報統計力学試験問題

定義.

サイコロ投げのように、起こりうる現象（1の目が出るなど）の集合を Ω とし、標本空間という。サイコロ投げなら、 $\Omega = \{1, 2, 3, 4, 5, 6\}$ である。ただし、ここで、1は1の目が出ることを表す。 Ω の部分集合 A を事象と呼ぶ。事象の集合 \mathcal{F} が次の性質を満たすとする。

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F}$ ならば、 $A^c \in \mathcal{F}$, ここで、 $A^c = \Omega - A$, つまり、 A の補集合
3. $A_i \in \mathcal{F}$ ($i = 1, 2, \dots$) ならば、 $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

また、 \mathcal{F} の要素 $A \in \mathcal{F}$ に対して次の性質を持つ確率 $P(A)$ が定義されているとする。

1. $P(\Omega) = 1$.
2. $0 \leq P(A) \leq 1$, $A \in \mathcal{F}$.
3. $A_i \in \mathcal{F}$ ($i = 1, 2, \dots$) が背反事象、つまり、 $A_i \cap A_j = \phi$, ($i \neq j$) ならば

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

必修問題

I. 事象 A, B について、 $P(A) > 0, P(B) > 0$ とする。 B が起こる条件のもとで A が起こる確率を、事象 B が与えられたときに A が起こる条件付き確率といい、 $P(A|B)$ と書く。これは次のようになる。

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

(1) ベイズの公式

$$P(A|B)P(B) = P(B|A)P(A) \tag{1}$$

を示せ。

(2) 次の等式を示せ。ただし、 $P(B) > 0, P(B^c) > 0$ とする。

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) \tag{2}$$

ヒント.

$$P(A \cap B) + P(A \cap B^c) = P(A \cap \Omega) = P(A).$$

II. 大数の法則、中心極限定理、大偏差原理について述べよ。

選択問題 以下の問題から 1 問選択して解答せよ。

III. ベイズの公式の応用。乳がん検診に関する確率の問題を考える。以下のような確率を定義する。

$P(\text{陽性})$: 検診で乳がんと判定される確率。 $P(\text{乳がん})$: 乳がんである確率。

$P(\text{陽性} | \text{乳がん})$: 乳がんであるときに、検診で乳がんであると判定される確率。

$P(\text{乳がん} | \text{陽性})$: 検診で乳がんであると判定されたとき、実際に乳がんである確率。

いま、 $P(\text{陽性} | \text{乳がん}) = 0.9$ とする。つまり、乳がんであるときに、検診で乳がんであると判定される確率が 9 割であるとする。また、実際の乳がんである人の割合は 0.8 パーセント、つまり、 $P(\text{乳がん}) = 0.008$ とする。さらに、乳がんでないのに陽性となる確率が、 $P(\text{陽性} | \text{乳がんでない}) = 0.07$ とする。

(1) $P(\text{陽性}) \simeq 0.08$ となることを示せ。

ヒント. $P(A^c) = 1 - P(A)$ と、式 (2) を用いよ。

(2) ベイズの公式 (1) を用いて、検診で陽性と判定され、実際に乳がんである確率が約 9 パーセント、つまり、 $P(\text{乳がん} | \text{陽性}) \simeq 0.09$ となる事を示せ。

IV. 統計力学と情報理論におけるエントロピーの関係。

k_B をボルツマン定数とする。絶対温度 T の熱浴とエネルギーのやりとりをする系を考える。この系のエネルギーが E である一つの状態にある確率は、カノニカル分布

$$P(E) = \frac{1}{Z} e^{-\beta E}, \quad Z = \int dE \Omega(E) e^{-\beta E}$$

で与えられる。ここで、 $\beta = 1/(k_B T)$ 、 $\Omega(E)$ はエネルギーが E であるときの状態密度、 Z は分配関数と呼ばれる。このとき、自由エネルギー F とエントロピー S は、

$$F = -k_B T \ln Z,$$

$$S = -\frac{\partial F}{\partial T}$$

で与えられる。一方、情報エントロピーは、

$$H_E = -\int dE \Omega(E) P(E) \log_2 P(E)$$

で与えられる。

$$H_E = \frac{1}{k_B \ln 2} S$$

となることを示せ。