2020.4.7

情報統計力学 事前学習１

４月１３日 (金) 午後１時ー２時半

確率論についてのまとめです。例えば、テキスト　理系の数学９「確率と統計」などを参考にして下さい。

定義.

サイコロ投げのように、起こりうる現象（１の目がでるなど）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. $A_i \in \mathcal{F}$ $(i = 1, 2, \cdots)$ ならば、$\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

このような事象の集合 $\mathcal{F}$ を Ω の $\sigma-$ **集合族**と呼ぶ。

$\mathcal{F}$ 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. $A_i \in \mathcal{F}$ $(i = 1, 2, \cdots)$ が背反事象、つまり、$A_i \cap A_j = \phi,\ (i \neq j)$ ならば

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

3つの組 $(\Omega, \mathcal{F}, P)$ を**確率空間**あるいは**確率モデル**という。このとき、以下の性質が成り立つ。

1. $P(\phi) = 0$, ここで、ϕ は空集合。
2. $A, B \in \mathcal{F}$、かつ $A \subset B$ なら、$P(A) \leq P(B)$.
3. $A, B \in \mathcal{F}$ が互いに背反、つまり、$A \cap B = \phi$ なら、$P(A \cup B) = P(A) + P(B)$.
4. $A_1, A_n, \cdots, A_n \in \mathcal{F}$ が互いに背反、つまり、$i \neq j$ のとき、$A_i \cap A_j = \phi$ なら、$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.
5. $A \in \mathcal{F}$ なら、$P(A^c) = 1 - P(A)$.
6. $A_1, A_n, \cdots, A_n \in \mathcal{F}$ のとき、$P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$. (有限劣加法性）.
7. $A_1, A_n, \cdots, A_n, \cdots \in \mathcal{F}$ のとき、$P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$. (完全劣加法性）.

8. $A_1, A_n, \cdots, A_n \in \mathcal{F}$ のとき、$P(\cap_{i=1}^n A_i) \geq 1 - \sum_{i=1}^n P(A_i^c)$.

これらの証明を試みて下さい。

次の命題も成り立つ。

II. 確率 P の連続性。$A_1, A_n, \cdots, A_n, \cdots \in \mathcal{F}$ とする。

1. $\{A_n\}$ が単調増大、 つまり、$A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots$ なら、$P(\cup_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.
2. $\{A_n\}$ が単調減少、 つまり、$A_1 \supset A_2 \supset \cdots \supset A_n \supset A_{n+1} \supset \cdots$ なら、$P(\cap_{n=1}^\infty A_n) = \lim_{n\to\infty} P(A_n)$.

これらの証明も試みて下さい。

次回は、条件付き確率を定義し、機械学習を学ぶ際、最も重要な定理の一つ、**ベイズの定理**、を解説します。また、その応用として、乳がん検診で、乳がんと診断された場合、本当に乳がんである確率の計算を行います。思ったより、随分低い確率になるので、意外に思うと思います。