2017.8.1

情報統計力学 追試験問題

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，$\cup_{i=1}^{\infty} A_i \in \mathcal{F}$

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. $0 \le P(A) \le 1,\ A \in \mathcal{F}.$
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

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

必修問題

I. 以下を示せ。

1. 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。

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

   以下のベイズの公式を示せ。

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

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

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

3. 確率Pの連続性。$A_1, A_n, \cdots, A_n, \cdots \in \mathcal{F}$のとき、以下を示せ。

   (a) $\{A_n\}$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、$P(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$.

   (b) $\{A_n\}$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、$P(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$.

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

P(陽性 | 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(乳がん | 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(陽性 | 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(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(陽性 | 乳がんでない) $= 0.07$ とする。
(1) P(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(乳がん | 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。系が，体積 V の断熱容器に入れられており, 外部とのエネルギーのやりとりが行われていないとする。系のエネルギーは，E であり，その状態数を W とする。　系の状態を，$i, (i = 1, \cdots, W)$ で表すとき，状態 i にある確率は，ミクロカノニカル分布

$$P_i = \frac{1}{W}$$

で与えられる。このとき，

$$\begin{aligned} S &= k_{\mathrm{B}} T \ln W, \\ \frac{1}{T} &= \frac{\partial S}{\partial E}, \\ \frac{p}{T} &= \frac{\partial S}{\partial V}. \end{aligned}$$

である。一方，情報エントロピーは,

$$H_E = -\sum_{i=1}^{W} P_i \log_2 P_i$$

で与えられる。

$$H_E = \frac{1}{k_{\mathrm{B}} \ln 2} S$$

となることを示せ。