

情報領域演習第二 第5回 K演習 (確率論) 宿題

宿題1-1: ベイズの定理

ある病気の感染率は0.1%である。この病気の診断方法は、感染している患者を診断した時には90%の確率で「感染あり」との結果を得る。また感染していない患者を診断した時には10%の確率で「感染あり」という結果を得る。

1. 病気の感染の有無に関する周辺確率表を記せ。
2. 患者の感染の有無が判明しているときの、診断結果の「感染あり」「感染なし」の条件付き確率表を記せ。
3. 「診断結果が感染ありであった」という結果を知らされた。この患者が実際に感染している確率をベイズの定理を用いて求めよ。

宿題1-2: ベイズの定理

ある病気の感染率は0.1%である。この病気の診断方法Aは、感染している患者を診断した時の正答率が80%、感染していない患者を診断したときの正答率は60%である。なお正答率は、患者の真の状態を正しく当てる確率とする。

1. 病気の感染の有無に関する周辺確率表を記せ。
2. 患者の感染の有無が分かっているという条件の下で、方法Aの診断結果が「感染あり」、「感染なし」それぞれとなる条件付き確率表を記せ。
3. 「診断結果が感染ありであった」という結果を知らされた。この患者が実際に感染している確率をベイズの定理を用いて求めよ。
4. 別の診断方法Bは、感染している患者を診断した時の正答率が70%、感染していない患者を診断したときの正答率も70%である。診断方法Bで「診断結果が感染ありであった」という結果を知らされた患者が、実際に感染している確率をベイズの定理を用いて求めよ。

宿題 2 : 期待値の計算

1. 標本空間が $\{1, 2, 3\}$ の確率変数 X の確率関数が

$$p(1) = 1/3, p(2) = 1/6, p(3) = 1/2 \quad (1)$$

と与えられている。(a) まず累積分布関数 $F(x)$ を描け。(b) 次に X の平均と分散を求めよ。

2. 練習問題 2 が一回に 0.5 の費用がかかる投資のリターン (回収額) を表しているとする。この投資の平均リターンを求めよ。またこの投資のリスク (標準偏差) を求めよ。

3. 標本空間が $\{x; 0 \leq x \leq 2\}$ の確率変数 X の確率密度関数が

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ -x + 2 & 1 \leq x \leq 2 \end{cases} \quad (2)$$

と与えられている。(a) この確率分布の累積分布関数を描け。(b) 確率変数 X の平均と分散を求めよ。

宿題 3 - 1 : 確率不等式

1. 確率表

X	1	2	...	N	総和
確率	p_1	p_2	...	p_N	1

と期待値 μ と分散 σ^2 が与えられている。教科書の証明を参考に

$$\begin{aligned} \sigma^2 &= \mathbf{E} \left[(X - \mu)^2 \right] \\ &= \sum_{k=1}^N (k - \mu)^2 p_k \\ &= \dots \end{aligned} \quad (3)$$

から始め、不等式

$$\sigma^2 \geq a^2 \Pr[|X - \mu| \geq a] \quad (4)$$

を示して、チェビシェフの不等式を証明せよ。

宿題 3 - 2 : 確率不等式

- マルコフの不等式を用いて、平均年収が 500 万円の世代の中で、年収が 5000 万円以上の人の割合は最大でどれくらいとなるかを求めよ。
- マルコフの不等式を用いて、1 企業あたりの平均従業員数が 2000 人の国では、10 万人以上の従業員を抱える企業の割合が最大でどれくらいとなるかを求めよ。
- チェビシェフの不等式を用いて、年収の平均が 500 万円、標準偏差が 200 万円の世代の中で、年収が 5000 万円以上の人の割合は最大でどれくらいとなるかを求めよ。
- チェビシェフの不等式を用いて、1 企業あたりの従業員数の平均が 500 人、標準偏差が 200 人の国で、10 万人以上の従業員を抱える企業の割合が最大でどれくらいとなるかを求めよ。

宿題の提出について

提出場所は西5号館3階の「総合情報学専攻事務室」前のポストで、「K演習」というコーナーを見つけて、投函せよ。

提出期限は 月 日 () 時までとする。

なおこの演習では、提出された宿題は返却しない。その代わりに解答例を、翌日より <http://bit.ly/jinlu-teach> にて公開する。