# §4 ベイズ推論　演習問題　解答

問題の難易度の目安【易】★☆☆　【基礎】★★☆　【標準】★★★

## 1 (★☆☆)(条件付き確率①)

(ジョーカーを除く)52 枚のトランプから無作為に 1 枚カードを選ぶとき,

$\boxed{\spadesuit}$ : スペードのカード,　$\boxed{\mathtt{J}}, \boxed{\mathtt{Q}}, \boxed{\mathtt{K}}$ : 絵札

とするとき,条件付き確率 $\mathbb{P}\left[\boxed{\mathtt{J}}, \boxed{\mathtt{Q}}, \boxed{\mathtt{K}} \mid \boxed{\spadesuit}\right]$ を求めよ.

**解** ジョーカーを除く 52 枚のトランプから無作為に 1 枚選ぶとき,スペードを引く確率は

$$\mathbb{P}\left[\boxed{\spadesuit}\right] = \frac{13}{52}.$$

次にスペードの絵札を引く確率は

$$\mathbb{P}\left[\boxed{\spadesuit} \cap \left(\boxed{\mathtt{J}}, \boxed{\mathtt{Q}}, \boxed{\mathtt{K}}\right)\right] = \frac{3}{52}.$$

したがって求める条件付き確率は

$$\mathbb{P}\left[\boxed{\mathtt{J}}, \boxed{\mathtt{Q}}, \boxed{\mathtt{K}} \mid \boxed{\spadesuit}\right] = \frac{\mathbb{P}\left[\boxed{\spadesuit} \cap \left(\boxed{\mathtt{J}}, \boxed{\mathtt{Q}}, \boxed{\mathtt{K}}\right)\right]}{\mathbb{P}\left[\boxed{\spadesuit}\right]} = \frac{\frac{3}{52}}{\frac{13}{52}} = \mathbf{\frac{3}{13}}.$$

∎

## 2 (★☆☆)(条件付き確率②)

5 回に 1 回の割合で帽子を忘れる癖のある K 君が,正月に A, B, C の 3 軒をこの順に年始回りして家に帰ったとき,帽子を忘れてきたことに気がついた.2 軒目の家 B に忘れてきた確率を求めよ.

**解** A, B, C に帽子を忘れてくる事象を,それぞれ $A$, $B$, $C$ として,家に帰ったときいずれかの場所で帽子を忘れてきた事象を $W$ とする.

$\mathbb{P}[W \cap A]$ は確率 1/5 で A に傘を忘れた確率であるから,$\mathbb{P}[W \cap A] = \frac{1}{5}$　$\cdots$①.

つづいて,$\mathbb{P}[W \cap B]$ は A に傘を忘れて来ず (確率 4/5),B で傘を忘れる確率であるから,$\mathbb{P}[W \cap B] = \frac{4}{5} \times \frac{1}{5} = \frac{4}{25}$　$\cdots$②.

最後に $\mathbb{P}[W \cap C]$ は A と B に傘を忘れて来ず (確率 $4/5 \times 4/5$),C で傘を忘れる確率であるから,$\mathbb{P}[W \cap C] = \frac{4}{5} \times \frac{4}{5}\frac{1}{5} = \frac{16}{125}$　$\cdots$③.

したがって，家に帰ったときいずれかの場所で帽子を忘れてきた確率 $\mathbb{P}[W]$ は，①–③より，

$$\mathbb{P}[W]=\mathbb{P}[W\cap A]+\mathbb{P}[W\cap B]+\mathbb{P}[W\cap C]=\frac{61}{125}\quad\cdots④.$$

求める確率は "2 軒目の家 B に忘れてきた確率" すなわち，家に帰ったときいずれかの場所で帽子を忘れてきた (事象 $W$)，という条件のもとで事象 $B$ が起こる条件付き確率 $\mathbb{P}[B\mid W]$ であるから，②, ④より

$$\mathbb{P}[B\mid W]=\frac{\mathbb{P}[W\cap B]}{\mathbb{P}[W]}=\frac{\frac{4}{25}}{\frac{61}{125}}=\boldsymbol{\frac{20}{61}}.$$

■

## 3 (★☆☆)(全確率の公式)

標本空間を $\Omega$ とし，事象 $A_1,\dots,A_n$ は互いに排反 (すなわち，$i\neq j$ ならば $A_i\cap A_j=\varnothing$) であり，$\displaystyle\bigcup_{i=1}^{n}A_i=\Omega$ を満たすとする．このとき任意の事象 $B$ に対して，

$$\mathbb{P}[B]=\sum_{i=1}^{n}\mathbb{P}[A_i]\mathbb{P}[B\mid A_i]$$

が成り立つことを示せ．ただし，$\mathbb{P}[B\mid A_i]$ は事象 $A_i$ の下での $B$ の条件付き確率を表す．

**解** $B=\displaystyle\bigcup_{i=1}^{n}(B\cap A_i)$ であり，$\{B\cap A_i\}$ も互いに排反であるから，$\mathbb{P}[B]=\sum_{i=1}^{n}\mathbb{P}[B\cap A_i]$．一方，条件付き確率の定義から，各 $i=1,\dots,n$ に対して $\mathbb{P}[B\cap A_i]=\mathbb{P}[A_i]\mathbb{P}[B\mid A_i]$ が成り立つ．したがって，$\mathbb{P}[B]=\sum_{i=1}^{n}\mathbb{P}[A_i]\mathbb{P}[B\mid A_i]$ となり，所望の式を得る． ■

## 4 (★☆☆)(ベイスの定理)

標本空間を $\Omega$ とし，事象 $A_1,\dots,A_n$ は互いに排反で $\displaystyle\bigcup_{i=1}^{n}A_i=\Omega$ を満たすとする．このとき任意の事象 $B$ に対して，

$$\mathbb{P}[A_k\mid B]=\frac{\mathbb{P}[A_k]\mathbb{P}[B\mid A_k]}{\sum_{i=1}^{n}\mathbb{P}[A_i]\mathbb{P}[B\mid A_i]},\qquad k=1,\dots,n$$

が成り立つことを示せ．

**解** 条件付き確率の定義を用いて，$\mathbb{P}[A_k \mid B] = \dfrac{\mathbb{P}[A_k \cap B]}{\mathbb{P}[B]} = \dfrac{\mathbb{P}[A_k]\mathbb{P}[B \mid A_k]}{\mathbb{P}[B]}$ であり，分母の $\mathbb{P}[B]$ に全確率の公式により得られる $\mathbb{P}[B] = \sum_{i=1}^{n} \mathbb{P}[A_i]\mathbb{P}[B \mid A_i]$ を代入すれば，証明が完了する． ■

**Remark** ベイズの定理より，結果から原因を推定する，すなわち，条件の役割を逆転した場合の条件付き確率を，逆転前の条件付き確率から計算することができる．

## 5 (★★☆)(ベイズの定理の応用①)

ある試験において，初めての受験者の合格率が 35%，2 回目以上の受験者の合格率が 70% であった．受験者比率は 1 回目の人が 40%，2 回目以上の人が 60% であった．これらのことから，合格者のうち初めての受験者の比率を求めよ．また，全体の合格率を求めよ．

**解** 事象 $A_1$, $A_2$, $G$ を次のように定める：

$A_1$ :　初めての受験者，　　$A_2$ :　2回目以上の受験者

$G$ :　合格者．

このとき条件から

$$\mathbb{P}[A_1] = 0.4, \quad \mathbb{P}[A_2] = 0.6, \quad \mathbb{P}[G \mid A_1] = 0.35, \quad \mathbb{P}[G \mid A_2] = 0.7 \quad \cdots ①,$$

ただし，$\mathbb{P}[G \mid A_1]$ =“初めての受験者の合格率”= 初めて受験する (事象 $A_1$) という条件の下で，事象 $G$ が起こる条件付き確率である．同様に $\mathbb{P}[G \mid A_2]$ も解釈される．求めるべきものは “合格者のうち初めての受験者の比率” すなわち，合格者になる (事象 $G$) という条件の下で初めての受験者 (事象 $A_1$) となる条件付き確率 $\mathbb{P}[A_1 \mid G]$ である．ゆえに，データ①およびベイズの定理により，

$$\begin{aligned}\mathbb{P}[A_1 \mid G] &= \frac{\mathbb{P}[A_1]\mathbb{P}[G \mid A_1]}{\mathbb{P}[A_1]\mathbb{P}[G \mid A_1] + \mathbb{P}[A_2]\mathbb{P}[G \mid A_2]} \\ &= \frac{0.4 \times 0.35}{0.4 \times 0.35 + 0.6 \times 0.7} = 0.25 = \mathbf{25}\%.\end{aligned}$$

一方，全体の合格率は全確率の公式により，

$$\begin{aligned}\mathbb{P}[G] &= \mathbb{P}[A_1]\mathbb{P}[G \mid A_1] + \mathbb{P}[A_2]\mathbb{P}[G \mid A_2] \\ &= 0.4 \times 0.35 + 0.6 \times 0.7 = 0.56 = \mathbf{56}\%.\end{aligned}$$

■

## 6 (★★☆)(ベイズの定理の応用②)

ある感染症 V の検査 T があり，検査 T は感染者の 75% に陽性反応，非感染者の 99% に陰性反応が出るものとする．また，被験者全体のうち，感染者であるという事象を $V$，非感染者である事象を $V^{\complement}$ とおく．また，被験者のうちで，検査 T によって陽性反応が出たという事象を $T_+$，陰性反応が出たという事象を $T_-$ とする．このとき，次の問いに答えよ．

(1) $\mathbb{P}[V] = p$ とおくとき，陽性反応が出た人が実際に感染者である (真に陽性) である確率 $f(p)$ を求めよ．

(2) $f(p)$ の挙動を調べることで，検査 T について分かることを簡潔に述べよ．特に，感染症が新型の場合における検査 T の実施方法について一つの考え方を述べよ．

(3) 陰性反応が出た人が実際に非感染者である (真に陰性) である確率を $g(p)$ を求めよ．

**解** (1)　検査 T は感染者の 75% に陽性反応，非感染者の 99% に陰性反応を示すから，

$$\mathbb{P}[T_+ \mid V] = 0.75, \quad \mathbb{P}\left[T_- \mid V^{\complement}\right] = 0.99 \quad \cdots①.$$

また，$\mathbb{P}\left[T_+ \mid V^{\complement}\right] + \underbrace{\mathbb{P}\left[T_- \mid V^{\complement}\right]}_{=0.99} = 1$ より，$\mathbb{P}(T_+ \mid V^{\complement}) = 0.01 \quad \cdots②$．ゆえに，ベイズの定理により

$$\begin{aligned} f(p) = \mathbb{P}[V \mid T_+] &= \frac{\mathbb{P}[V]\mathbb{P}[T_+ \mid V]}{\mathbb{P}[V]\mathbb{P}[T_+ \mid V] + \mathbb{P}[V^{\complement}]\mathbb{P}\left[T_+ \mid V^{\complement}\right]} \\ &= \frac{0.75p}{0.75p + 0.01(1-p)} = \boldsymbol{\frac{75p}{1+74p}}. \end{aligned}$$

(2)　$f'(p) = \frac{75}{(1+74p)^2} > 0$ であるから，$f(p)$ は単調増加．ゆえに，この感染症 $V$ が稀であれば，$p$ は限りなく小さいので，$f(p)$ の値は小さくなる，すなわち**陽性反応が出ても感染している確率は低い**．稀でなければ，$p$ は大きくなり，それに伴って $f(p)$ の値は大きくなるので，**陽性反応が出れば高確率で感染していることがわかる**．

感染症が新型の場合，検査で陽性反応が出ても実際に感染している確率は低く出ることから，検査対象をハイリスク群に絞って行うことが有効であると考えられる．

(3)　$\mathbb{P}\left[V^{\complement}\right] = 1 - \mathbb{P}[V] = 1 - p$ であるから，求める確率は，

$$\begin{aligned} g(p) = \mathbb{P}\left[V^{\complement} \mid T_-\right] &= \frac{\mathbb{P}\left[V^{\complement}\right]\mathbb{P}\left[T_- \mid V^{\complement}\right]}{\mathbb{P}\left[V^{\complement}\right]\mathbb{P}\left[T_- \mid V^{\complement}\right] + \mathbb{P}[V]\mathbb{P}[T_- \mid V]} \\ &= \frac{0.99(1-p)}{0.99(1-p) + 0.3p} = \boldsymbol{\frac{33(1-p)}{33-23p}}. \end{aligned}$$

∎

## 7 (★★☆)(ベイズの定理の応用③)

以下の表では，“タダ”や“アタリ”という単語を含むメールが迷惑メールか判定し，表としてまとめたデータである．1 行目，2 行目の“タダ”や“アタリ”という単語を含む場合には ∘ を入れ，そうでない場合には × を記入している．3 行目は，実際に迷惑メールかどうかを確認し，スパムであれば∘，そうでない場合には × を記入している．迷惑メールである確率を $\mathbb{P}[\mathrm{spam}]$ とする．

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| “タダ” | × | ∘ | × | × | ∘ | × | × | ∘ | × | × |
| “アタリ” | ∘ | × | ∘ | × | × | ∘ | × | × | × | ∘ |
| 迷惑メール (spam) | ∘ | ∘ | ∘ | × | × | × | × | ∘ | × | × |

“タダ”を含むメールである事象を free，“アタリ”を含むメールである事象を win とし，迷惑メールに対し free と win は互いに独立，すなわち

$$\mathbb{P}[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}]=\mathbb{P}[\texttt{free}\mid\mathrm{spam}]\,\mathbb{P}[\texttt{win}\mid\mathrm{spam}]$$

を仮定する．このとき，以下の問いに答えよ．

(1) $\mathbb{P}[\mathrm{spam}]$ および，迷惑メールでない確率 $\mathbb{P}\left[\mathrm{spam}^{\complement}\right]$ を求めよ．

(2) $\mathbb{P}[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}]$ および $\mathbb{P}\left[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}^{\complement}\right]$ を求めよ．

(3) “タダ”と“アタリ”を含む場合の迷惑メールの判定率を求めよ．

**解** (1)　与えられた表から $\mathbb{P}[\mathrm{spam}]=\frac{4}{10}=\boldsymbol{\frac{2}{5}}$，$\mathbb{P}[\mathrm{spam}^{\complement}]=1-\mathbb{P}[\mathrm{spam}]=\frac{6}{10}=\boldsymbol{\frac{3}{5}}$.

(2)　独立性から

$$\mathbb{P}[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}]=\mathbb{P}[\texttt{free}\mid\mathrm{spam}]\,\mathbb{P}[\texttt{win}\mid\mathrm{spam}]\quad\cdots①$$

$$\mathbb{P}\left[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}^{\complement}\right]=\mathbb{P}\left[\texttt{free}\mid\mathrm{spam}^{\complement}\right]\mathbb{P}\left[\texttt{win}\mid\mathrm{spam}^{\complement}\right]\quad\cdots②$$

である．

$$\mathbb{P}[\texttt{free}\mid\mathrm{spam}]=\frac{\mathbb{P}[\texttt{free}\cap\mathrm{spam}]}{\mathbb{P}[\mathrm{spam}]}=\frac{2/10}{4/10}=\frac{1}{2},$$

$$\mathbb{P}[\texttt{win}\mid\mathrm{spam}]=\frac{\mathbb{P}[\texttt{win}\cap\mathrm{spam}]}{\mathbb{P}[\mathrm{spam}]}=\frac{2/10}{4/10}=\frac{1}{2}$$

であるから，①より

$$\mathbb{P}[\texttt{free}\cap\texttt{win}\mid\mathrm{spam}]=\frac{1}{2}\times\frac{1}{2}=\boldsymbol{\frac{1}{4}}.$$

同様に，

$$\mathbb{P}\left[\texttt{free}\mid\mathrm{spam}^{\complement}\right]=\frac{\mathbb{P}[\texttt{free}\cap\mathrm{spam}^{\complement}]}{\mathbb{P}[\mathrm{spam}^{\complement}]}=\frac{1/10}{6/10}=\frac{1}{6},$$

$$\mathbb{P}\left[\texttt{win}\mid\mathrm{spam}^{\complement}\right]=\frac{\mathbb{P}[\texttt{win}\cap\mathrm{spam}^{\complement}]}{\mathbb{P}[\mathrm{spam}^{\complement}]}=\frac{2/10}{6/10}=\frac{1}{3}$$

であるから，②より

$$\mathbb{P}\left[\texttt{free} \cap \texttt{win} \mid \texttt{spam}^{\texttt{C}}\right] = \frac{1}{6} \times \frac{1}{3} = \frac{\mathbf{1}}{\mathbf{18}}.$$

(3)　“タダ” と “アタリ” を含む場合の迷惑メールの判定率 $\mathbb{P}[\texttt{spam} \mid \texttt{free} \cap \texttt{win}]$ は，ベイスの定理により

$$\begin{aligned}\mathbb{P}[\texttt{spam} \mid \texttt{free} \cap \texttt{win}] &= \frac{\mathbb{P}[\texttt{spam}]\mathbb{P}[\texttt{free} \cap \texttt{win} \mid \texttt{spam}]}{\mathbb{P}[\texttt{spam}]\mathbb{P}[\texttt{free} \cap \texttt{win} \mid \texttt{spam}] + \mathbb{P}[\texttt{spam}^{\texttt{C}}]\mathbb{P}\left[\texttt{free} \cap \texttt{win} \mid \texttt{spam}^{\texttt{C}}\right]} \\ &= \frac{\frac{4}{10} \times \frac{1}{4}}{\frac{4}{10} \times \frac{1}{4} + \frac{6}{10} \times \frac{1}{18}} = \frac{3}{4} = \mathbf{75}\%.\end{aligned}$$

∎