条件付き確率とベイズの定理

条件付確率

事象Aが生じたという条件の下で、事象Bが生じる確率

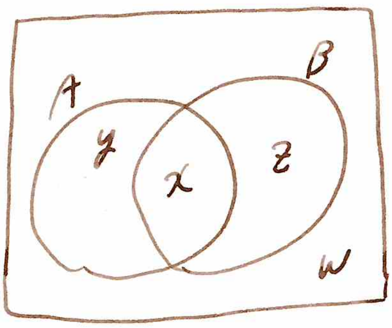
〜Aの下でのBの条件付確率 $P(B|A)$

例 コインを2回投げる例.

$$S = \{ \underbrace{(\text{オモテ}, \text{オモテ})}_{x}, \underbrace{(\text{オモテ}, \text{ウラ})}_{y}, \underbrace{(\text{ウラ}, \text{オモテ})}_{z}, \underbrace{(\text{ウラ}, \text{ウラ})}_{w} \}$$

$A = \{x, y\} \subset S$ 1回目にオモテが出るという事象

$B = \{x, z\} \subset S$ 2回目にオモテが出る 〃



$P(B|A) = \underbrace{\text{事象}A\text{が生じた}}_{(x \text{ or } y)}$という条件の下で、

$\underbrace{\text{事象}B\text{が生じる}}$確率 $= \dfrac{1}{2}$

(zが生じないことは既知なので xのみ考える.)

この例からわかることは、
(等確率の世界では)

$$P(B|A) = \frac{A \cap B\text{に含まれる根元事象の数}}{A\text{に含まれる根元事象の数}}$$

$$= \frac{\dfrac{A \cap B\text{に含まれる根元事象の数}}{S\text{に含まれる根元事象の数}}}{\dfrac{A\text{に含まれる根元事象の数}}{S\text{に含まれる根元事象の数}}}$$

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

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

Def.（条件付き確率）

事象 A が生じるという条件の下で
事象 B が生じる確率

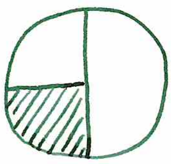
$$P(B|A) = \frac{P(A\cap B)}{P(A)} \quad (\text{ただし } P(A)>0)$$

the conditional probability of B given A

※等確率ではなく一般の場合．
上の式により条件付き確率を定義する。

Review

わり算…分母の量を新たな1単位として
分子の量を測り直す試み



← ケーキ $\frac{1}{4}$

⇩

← ケーキ $\frac{1}{2}$ コを全体（新たな単位量）とすると，元の $\frac{1}{4}$ は，その半分 $\left(\frac{1}{2}\right)$．

これが $\frac{1/4}{1/2} = \frac{1}{2}$ の意味である。

Thm (乗法公式)

$A, B \subset S$

$P(A) > 0$

$\Longrightarrow P(A \cap B) = P(B \mid A) P(A)$

$P(A) > 0,\ P(B) > 0$ ならば

$$\begin{aligned} P(A \cap B) &= P(B \mid A) P(A) \\ &= P(A \mid B) P(B) \end{aligned}$$

ベイズの定理の証明でポイントになる。

例.

S ある年にある地域で起きた事故の全体

$N(\subset S)$ 夜に発生した事故の全体

$Y(\subset S)$ 若者(26歳以下)の運転による事故の全体

・わかっていること

$P(N) = \frac{6}{10}$ 60%

$P(Y) = \frac{4}{10}$ 40%

$P(N \cap Y) = \frac{3}{10}$ 30%

(1) $\boxed{\text{夜に起きた事故にしめる若者の運転による事故の割合いは?}}$

$$P(Y|N) = \frac{P(Y \cap N)}{P(N)} = \frac{3/10}{6/10} = \frac{1}{2} \quad \underline{50\%}$$

$P(Y) = 40\%$ つまり、全体の中で若者による事故は40%だが、夜に限定すると、50%は若者による事故。

(2) 若者による事故のうち、夜に発生した事故の割合いは？

$$P(N|Y) = \frac{P(N \cap Y)}{P(Y)}$$

$$= \frac{3/10}{4/10} = \frac{3}{4} \qquad \therefore \underline{75\%}$$

$P(N) = 60\%$.

つまり全体の中で夜間発生する事故は6割。

しかし若者による事故に限定すると、75%が夜間に発生している。

※若者 = 25歳以下

夜間 = 18時〜4時

というように明確に範囲を定めておく必要がある。

例:

S ある会社の同期の全体

$A(\subset S)$ 3年以内にある資格を取得

$B(\subset S)$ 滋賀大学の出身者

・わかっていること

$P(A) = \frac{6}{10}$ 60%

$P(B) = \frac{4}{10}$ 40%

$P(A \cap B) = \frac{3}{10}$ 30%

(1) 3年以内に資格を取得した人の中で滋賀大出身者の割合は?

$P(B|A) = ?$

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

$$= \frac{3/10}{6/10} = \frac{1}{2}$$ 50%

$P(B) = 40\%$

つまり、滋賀大出身者は同期の中で40%。

しかし、資格を取得した人に限定すると

その半分(50%)を滋賀大出身者が占めている。

(2) 滋賀大卒業生における、資格取得者の割合いは？

$P(A|B) = ?$

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{3/10}{4/10} = \frac{3}{4}$$

$\underline{75\%}$

$P(A) = 60\%$

つまり3年以内に資格を取得するのは、

同期のうち60%。

しかし、$P(A|B) = 75\%$。

つまり滋賀大出身者に限定すると、75%は

3年以内にその資格を取得している。

Def.

$A_1, A_2, \cdots, A_n, B \subset S$

$\{A_1, \cdots, A_n\}$: <u>partition</u> of B

Bの<u>分割</u>

$\Longleftrightarrow$ ① $B = \bigcup_{i=1}^{n} A_i$

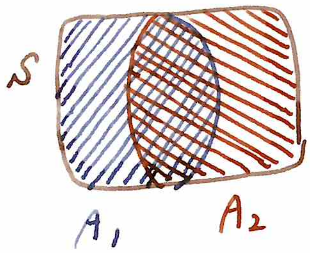
② $\{A_1, \cdots, A_n\}$: mutually exclusive

i.e. $A_i \cap A_j = \phi$

$(i, j = 1, 2, \cdots, n \,;\, i \neq j)$

$B = S$のケースで追加説明

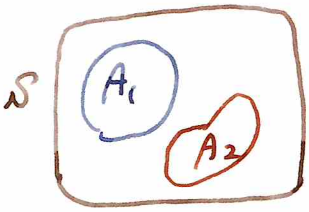
① だけではダメ.



$S = A_1 \cup A_2$ だが, これでは

$\{A_1, A_2\}$ は S の分割とは言わない。

② だけでもダメ.



$A_1 \cap A_2 = \phi$ だが, これでも

$\{A_1, A_2\}$ は S の分割とは言わない。

ex.

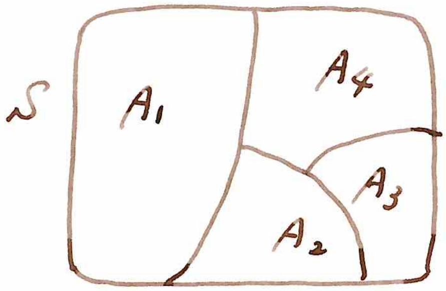
S 日本国内で流通しているバナナの集合.

$A_1 (\subset S)$ インド産

$A_2 (\subset S)$ 中国産

$A_3 (\subset S)$ フィリピン産

$A_4 (\subset S)$ その他



とすると. $\{A_1, A_2, A_3, A_4\}$: S の分割.

ex.

あるメーカー

S 扱っている部品の集合

A_i : 第 i 工場で生産され.納品されてきたもの

$(i = 1, 2, \cdots, 5)$

このとき.

$\{A_1, A_2, \cdots, A_5\}$: S の分割.

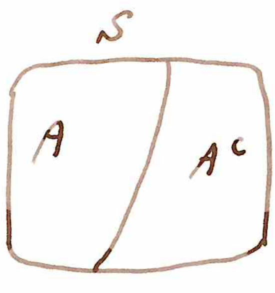
$A \subset S$

$\Rightarrow \{A, A^c\}$: S の分割

ex.

S : 社員の全体

A : 男性社員の全体



$\Rightarrow A^c$: 女性社員の全体

$\{A, A^c\}$: S の分割.

$A_1, A_2, \cdots$: partition of S

$B \subset S$

$\Rightarrow B \cap A_1, B \cap A_2, \cdots$:

partition of B

Proof

We will prove that

① $B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots$

② $(B \cap A_i) \cap (B \cap A_j) = \phi \quad (i, j = 1, 2, \cdots ; i \neq j)$.

Since $A_1, A_2, \cdots$ are a partition of S, we have

(1) $S = A_1 \cup A_2 \cup \cdots$

(2) $A_i \cap A_j = \phi \quad (i, j = 1, 2, \cdots ; i \neq j)$.

①
$$\begin{aligned} RHS &= (B \cap A_1) \cup (B \cap A_2) \cup \cdots \\ &= B \cap (A_1 \cup A_2 \cup \cdots) \\ &\overset{(1)}{=} B \cap S \\ &= B = LHS. \end{aligned}$$

② Let $i, j \in \mathbb{N} : i \neq j$.

It holds that

$$\begin{aligned} &(B \cap A_i) \cap (B \cap A_j) \\ &= B \cap A_i \cap A_j \subset A_i \cap A_j \overset{(2)}{=} \phi \end{aligned}$$

$\therefore (B \cap A_i) \cap (B \cap A_j) = \phi$.

//

ex

ある会社

S = 顧客全体の集合

A_1 = 10代の 〃

A_2 = 20代の 〃

⋮

A_5 = 50代の 〃

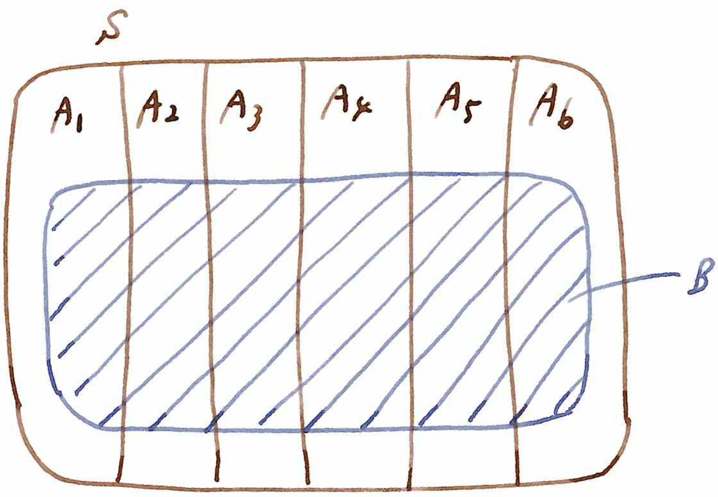
A_6 = 60代以上の 〃

このとき，$\{A_1, \cdots, A_6\}$: S の分割

$B\ (\subset S)$ 女性の顧客全体の集合

このとき，$\{B \cap A_1, B \cap A_2, \cdots, B \cap A_6\}$: B の分割

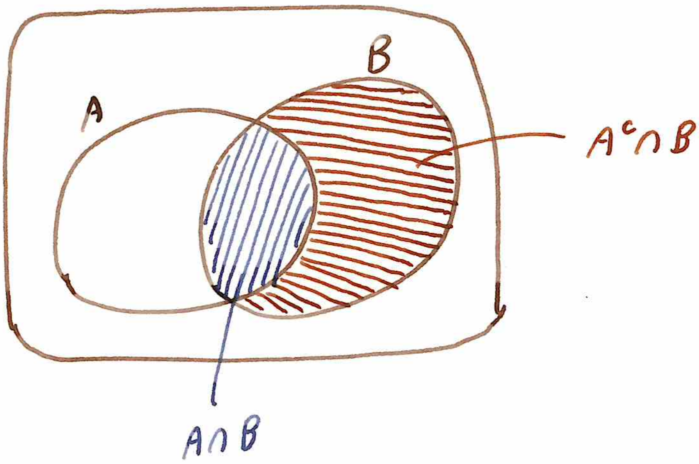
↖20代女性の顧客の全体



Cor

$A, B \subset S$

$\Rightarrow A\cap B,\ A^c\cap B$: partition of B



ex

$S = \{1, 2, 3, 4, 5, 6\}$

$A = \{2, 4, 6\}$ 偶数

$B = \{3, 4, 5, 6\}$ 3以上

Then,

$A\cap B = \{4, 6\}$ 3以上の偶数

$A^c\cap B = \{3, 5\}$ 3以上の奇数

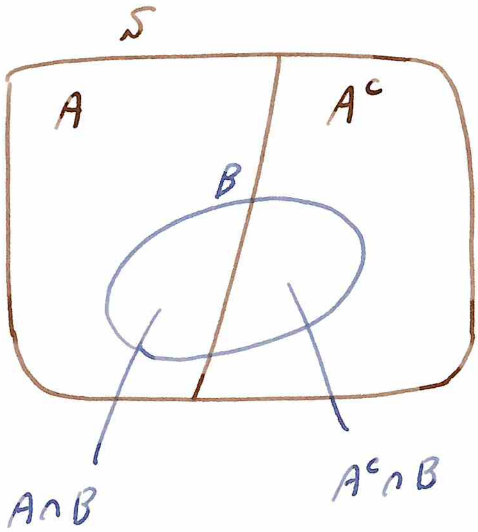
B の partition

ex

S : ある会社の大卒社員の全体

A : 文系学部出身者の全体

B : ある資格の取得者.



Th (partition theorem
全確率の公式)

$\{A_1, A_2, A_3\}$: partition of Ω
with $P(A_i) > 0$ $(i = 1, 2, 3)$

$B \subset \Omega$

$$\Longrightarrow P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$$

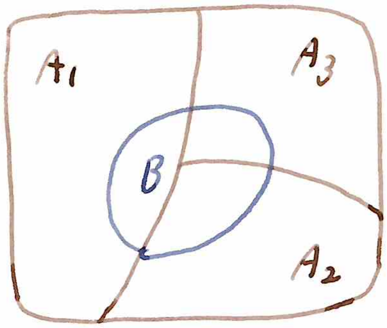
Proof

From the assumptions, we know that
$\{A_1 \cap B, A_2 \cap B, A_3 \cap B\}$ is a partition of B.

Thus, $P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B)$ 　乗法公式

$$= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3).$$

//



Th

$\{A_1, \cdots, A_n\}$: partition of S

$P(A_i) \in (0,1) \quad (i = 1, 2, \cdots, n)$

$B \subset S$

$$\Longrightarrow P(B) = \sum_{i=1}^{n} P(B \mid A_i) P(A_i)$$

Cor

$A, B \subset S$

$P(A) \in (0,1)$

$$\Longrightarrow P(B) = P(B \mid A) P(A) + P(B \mid A^c) P(A^c)$$

例

S 国内を流通しているバナナ

A_1 : インド産 (シェア40%)

A_2 : 中国産 (〃 30%)

A_3 : その他 (〃 30%)

$\longrightarrow P(A_1) = \frac{4}{10}$

$P(A_2) = \frac{3}{10}$

$P(A_3) = \frac{3}{10}$

・わかっていること。

インド産の 10% が不良品(虫つき)

中国産の 5% が 〃

その他の 15% が 〃

$\longrightarrow B$: 不良品 とすると

$P(B|A_1) = \frac{1}{10}$

$P(B|A_2) = \frac{1}{20}$

$P(B|A_3) = \frac{3}{20}$

(1) 全体の中で
インド産不良品，中国産不良品，その他の不良品
の割合いを求めよ。

$P(A_1 \cap B) = ?$ $P(A_2 \cap B) = ?$ $P(A_3 \cap B) = ?$

$P(A_1 \cap B) = P(B|A_1)P(A_1)$ ← 乗法公式

$= \frac{1}{10} \cdot \frac{4}{10} = \frac{4}{100}$ 4%

$P(A_2 \cap B) = P(B|A_2)P(A_2)$

$= \frac{1}{20} \cdot \frac{3}{10} = \frac{3}{200}$ 1.5%

$P(A_3 \cap B) = P(B|A_3)P(A_3)$

$= \frac{3}{20} \cdot \frac{3}{10} = \frac{9}{200}$ 4.5%

(2) 全体の中の不良品の割合いを求めよ

$P(B) = ?$

$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$

全確率の公式

$= \frac{4}{100} + \frac{3}{200} + \frac{9}{200}$

$= \frac{20}{200} = \frac{10}{100}$ 10%

例（乗法公式と全確率の公式の応用例）

S 社員全体の集合.

$A(\subset S)$ 男性社員

$A^c(\subset S)$ 女性 〃 → A と A^c は S の分割になっている。

$B(\subset S)$ ある資格の保有者の全体.

- 男性社員の全社員に占める割合いは 60%

→ $P(A)=\frac{6}{10}$, $P(A^c)=\frac{4}{10}$

- 男性社員の中でその資格をとっているのは 10%

→ $P(B|A)=\frac{1}{10}$

- 女性社員の中でのその資格取得者は 20%

→ $P(B|A^c)=\frac{2}{10}$

(1) 全社員の中で、その資格を取得している男性社員の割合いは？

$P(A\cap B)=?$

答 $P(A\cap B)=P(B|A)P(A)$ ← 乗法公式

$=\frac{6}{10}\cdot\frac{1}{10}=\frac{6}{100}$

6%

(2) 全社員の中で、その資格を取得している女性社員の割合いは？

$P(A^c \cap B) = ?$

答. $P(A^c \cap B) = P(B \mid A^c)\, P(A^c)$

$= \frac{2}{10} \cdot \frac{4}{10} = \frac{8}{100}$

8%

(3) 社員全体の中で その資格を取得しているのは何%か？

$P(B) = ?$

答.

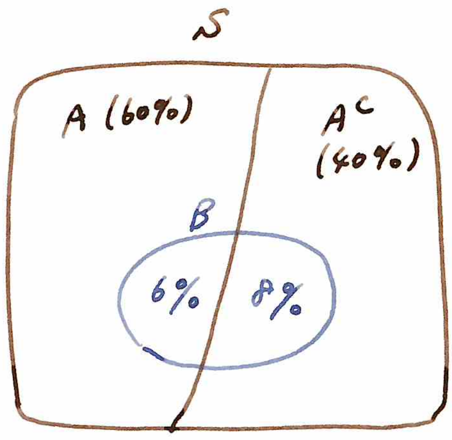
全確率の公式より

$P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c)$

$= \frac{6}{100} + \frac{8}{100}$

$= \frac{14}{100}$

14%



例

10本中3本の当りがあるくじを
A, Bの2人がこの順番でひく。
A(最初にひく人)とB(次にひく人)のどちらが有利か?

$\left.\begin{array}{l}\text{Aが当りをひくという事象を } A \\ \text{B 〃 〃 } B\end{array}\right)$ と書く。

- Aが当りをひく確率は?

$$P(A) = \underbrace{\frac{3}{10}}$$

- Bが当りをひく確率は?

(i) Aが当りをひく場合

$$P(B|A) = \frac{2}{9}$$

Aが当りをひくことは前提に
Bも当りをひく確率

(ii) Aが外れをひく場合 (その確率は $P(A^c) = \frac{7}{10}$)

$$P(B|A^c) = \frac{3}{9}$$

全確率の公式

以上より,

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

$$= \frac{2}{9}\cdot\frac{3}{10} + \frac{3}{9}\cdot\frac{7}{10}$$

$$= \frac{1}{30}(2+7) = \underbrace{\frac{3}{10}}$$

よって $P(A) = P(B)$.

Th (Bayes's theorem)

$A_1, A_2, \cdots, A_n (\subset S)$ partition of S

such that $P(A_k) \in (0,1)$ $(k=1,2,\cdots,n)$

$B \subset S$ with $P(B) \in (0,1]$

$$\Longrightarrow P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{k=1}^{n} P(B \mid A_k)P(A_k)}$$

$(i=1,2,\cdots,n)$

Proof

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)}$$

← 条件付確率の定義

乗法公式

$$= \frac{P(B \mid A_i)P(A_i)}{P(B)}$$

全確率の公式

$$= \frac{P(B \mid A_i)P(A_i)}{\sum_{k=1}^{n} P(B \mid A_k)P(A_k)}$$

//

※ $P(A_i \mid B)$ を $P(B \mid A_k)$ と $P(A_k)$ を用いて

表すことができている。

Cor

$A, B \subset S$

$P(A) \in (0,1)$

$P(B) \in (0,1]$

$$\Rightarrow P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B \mid A)\,P(A) + P(B \mid A^c)\,P(A^c)}$$

例.

住宅ローンの返済不履行(デフォルト)の状況

D: デフォルト

E: 事前の審査を優良でパス

以下のようなことがわかっている。

- 20人に1人がデフォルトする。
- デフォルトした人のうち審査で優良だったのは3割。
- デフォルトしなかった人のうち
  審査で優良だったのは8割。

このとき、審査を優良でパスしていながら、
デフォルトする確率はいくらか?

解

問題文より

$$P(D) = \frac{1}{20}$$

$$P(E|D) = \frac{3}{10}$$

$$P(E|D^c) = \frac{8}{10}$$

がわかっている。

求めたいのは $P(D|E)$ である。

ベイズの定理より

$$P(D|E) = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|D^c)P(D^c)}$$

$$= \frac{\frac{3}{10} \cdot \frac{1}{20}}{\frac{3}{10} \cdot \frac{1}{20} + \frac{8}{10} \cdot \frac{19}{20}}$$

$$= \frac{3}{3+152} = 0.019354\cdots$$

$\therefore$ 約1.94%

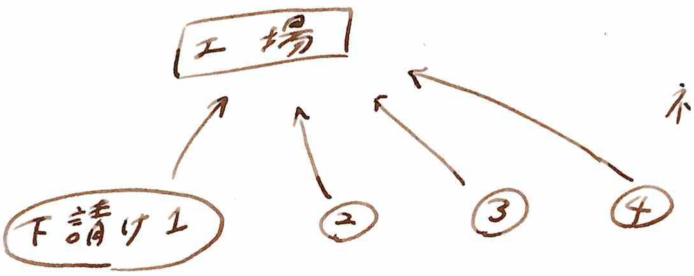
※ デフォルトするのは、全体では $\frac{1}{20} = 5\%$.

審査で優良だった人に限定すれば

約1.94% $(< 5\%)$

にまで下がる。

例



ネジを供給している。

各工場の納品数（月あたり），納品割合，欠陥率は下の通り。

工場	1	2	3	4	
納品数	2000	500	1000	1000	計 4500
納品割合	$4/9$	$1/9$	$2/9$	$2/9$	1
欠陥率	5%	20%	10%	10%	

ネジを1コ，ランダムに取る。

A_1 ： そのネジが工場1によって作られたものである。

A_2, A_3, A_4 ： 同様

$P(A_1) = 4/9$ ， $P(A_2) = 1/9$ ， $P(A_3) = 2/9$， $P(A_4) = 2/9$

D ： 選ばれたネジが欠陥品である。

$P(D|A_1) = 0.05$ ， $P(D|A_2) = 0.2$

$P(D|A_3) = 0.1$ ， $P(D|A_4) = 0.1$

(1) 取り出したネジが欠陥品である確率は？

$P(D) = ?$

$$P(D) = P(A_1)P(D|A_1) + P(A_2)P(D|A_2) + P(A_3)P(D|A_3) + P(A_4)P(D|A_4)$$

$$= \frac{4}{9}\cdot\frac{1}{20} + \frac{1}{9}\cdot\frac{1}{5} + \frac{2}{9}\cdot\frac{1}{10} + \frac{2}{9}\cdot\frac{1}{10}$$

$$= \frac{1}{45} + \frac{1}{45} + \frac{1}{45} + \frac{1}{45} = \frac{4}{45} \fallingdotseq 0.0888\cdots$$

全体の欠陥品率

$< P(D|A_2) = 0.2$

工場2の欠陥品率

(2) ネジが欠陥品であることが判明したとき、それが工場2製である確率は？

$P(A_2|D) = ?$

$$P(A_2|D) = \frac{P(A_2)P(D|A_2)}{P(A_1)P(D|A_1) + \cdots + P(A_4)P(D|A_4)}$$

$$= \frac{\frac{1}{45}}{\frac{4}{45}} = 0.25$$

25%

Remark

あるネジが工場2製であるという事前確率

$P(A_2) = 1/9$

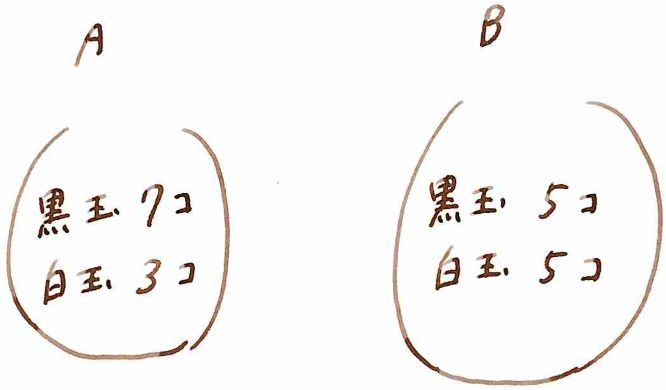
⇨ (そのネジが欠陥品であるという追加的な情報が得られる) ⇨

事後確率

$P(A_2|D) = \frac{1}{4}$

例

2つのツボがある。



← Bの方が白玉の比率が高い。

(1) A or B どちらかのツボをランダムに選択する。

$P(A) = P(B) = 1/2$

※ $A \cap B = \phi$ なので $A^c = B$

(2) 選んだツボから非復元抽出で2コ取り出す。

その結果，2つとも白玉であった。

$\boxed{\text{(1)で選んだツボがAであった確率は？}}$

(1)の段階だと，Aが選ばれていた確率は $1/2$.

(2)の情報が追加的に入ってくると，「Bが選ばれていた確率が高そうだ」と判断したくなる。逆に，Aが選ばれていた確率は下がりそう。

どれぐらい下がるか？

解答.

X = 取り出したボールに含まれる白玉の数

$$P(X=2\mid A) = \frac{3}{10}\times\frac{2}{9} = \frac{1}{15}$$

$$P(X=2\mid B) = \frac{5}{10}\times\frac{4}{9} = \frac{2}{9}$$

これより,

$$P(A\mid X=2) = \frac{P(X=2\mid A)P(A)}{P(X=2\mid A)P(A)+P(X=2\mid B)P(B)}$$

$$= \frac{\frac{1}{15}\times\frac{1}{2}}{\frac{1}{15}\times\frac{1}{2}+\frac{2}{9}\times\frac{1}{2}} = \frac{3}{3+10} = \frac{3}{13}$$

(2)′ 3つ取り出して3つとも白玉だった。

$$P(X=3\mid A) = \frac{3}{10}\cdot\frac{2}{9}\cdot\frac{1}{8} = \frac{1}{120}$$

$$P(X=3\mid B) = \frac{5}{10}\cdot\frac{4}{9}\cdot\frac{3}{8} = \frac{1}{12}$$

$$\therefore P(A\mid X=3) = \frac{P(X=3\mid A)P(A)}{P(X=3\mid A)P(A)+P(X=3\mid B)P(B)}$$

$$= \frac{\frac{1}{120}\cdot\frac{1}{2}}{\frac{1}{120}\cdot\frac{1}{2}+\frac{1}{12}\cdot\frac{1}{2}}$$

$$= \frac{1}{11}$$

条件付き確率とベイズの定理

問題1.

（1）事象Aが生じるという条件の下で事象Bが生じる条件付き確率の定義を数式で記述しなさい。

（2）乗法公式を数式で記述しなさい。

問題2. 会社の中で若手（40歳未満）は60% 、英語が得意な者（TOEIC800点以上）は20% 、若手でかつ英語の得意な者は10% である。以下の値を求めなさい。

（1）若手の中で、英語が得意な者は何割いるか？

（2）英語が得意な者に占める若手の割合はどれぐらいか？

問題3.（全確率の公式）

Sを標本空間、$\{A_1, A_2, A_3\}$をSの分割で$P(A_i) > 0$ $(i = 1,2,3)$ を満たすものとする。また、$B \subset S$とする。このとき、

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)$$

を証明しなさい。

問題4. ベイズの定理の主張を述べ、それを証明しなさい。

問題5. ある会社で大卒の社員は、文系学部出身が6割、理系学部出身が4割である。また、営業を経験しているのは文系学部出身者の中では50%、理系学部出身者の中では30%である。

（1）この会社の大卒社員の中で営業を経験している文系学部出身者は何%か。

（2）この会社の大卒社員の中で営業を経験している理系学部出身者は何%か。

（3）この会社の大卒社員全体の中で営業を経験しているのは何%か?

問題6. 自動車組み立て企業が、4つの企業からタイヤを供給されているとする。納品数や欠陥率は下表のとおり。下の問いに答えなさい。

タイヤ企業	1	2	3	4	計
納品数	2000	1000	2000	4000	9000
納品割合	2/9	1/9	2/9	4/9	1
欠陥率	10%	1%	5%	5%	

（1）全体の欠陥品率は何%か？

（2）あるタイヤが欠陥品であることが判明したとき、それが企業1製である確率は何%か？

問題7. レジュメの例などを参考にして、全確率の公式やベイズの定理を用いる問題を自分で考え、それに解答を与えよ。