高信頼システム 09

機械学習を活用した高信頼化技術

張山昌論

テストに関して

7月22日（月）13：00から

筆記用具，授業資料，電卓，自筆ノートのみ使用可.

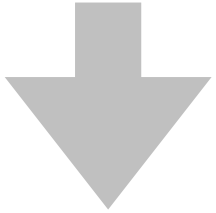
異常検出の問題点

事例が少ない

→ 通常の機械学習が有効に働きにくい

正常事例 1000件、異常事例 1件

→ 正常事例に合わせて学習されがち



統計的手法に基づく手法が有効

「ベイズ統計」

従来の統計学

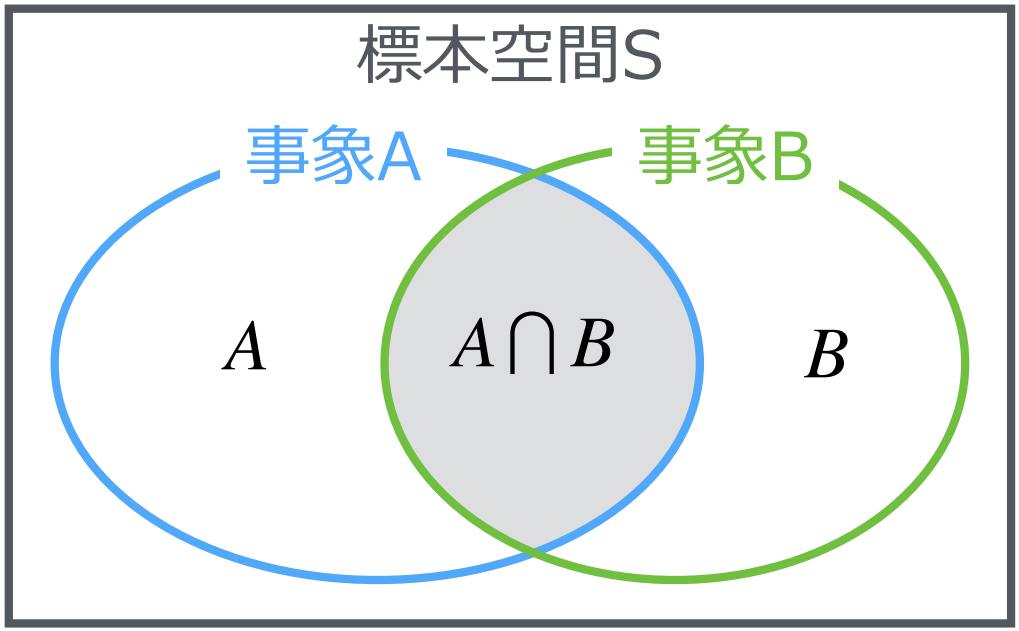
- 従来の統計学：「頻度論」
  - 大量のデータに基づく厳密な理論体系

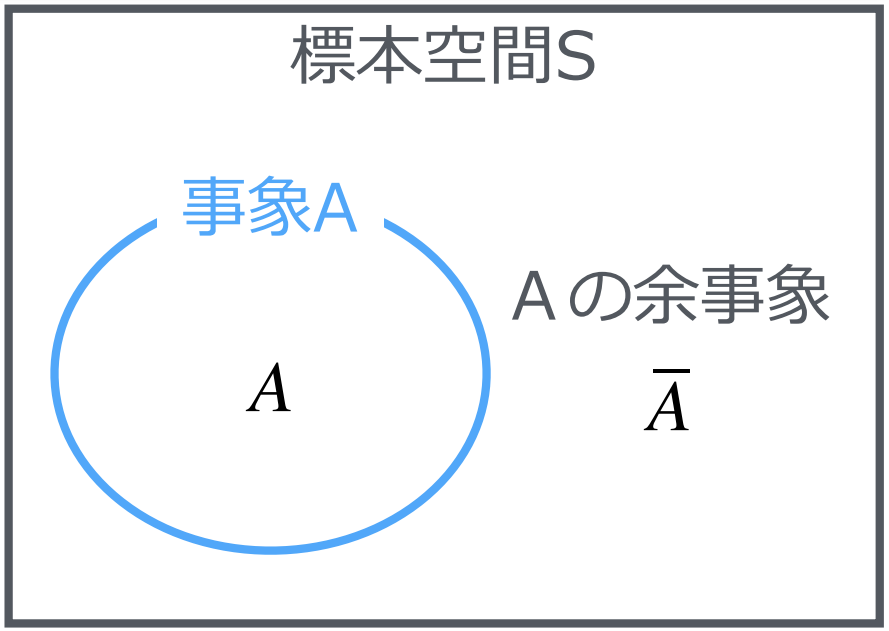
弱点

- 少ないデータのからの解析
- 時間の異なるデータの活用
- 経験を加味したデータ解析

「ベイズ統計」で解決

確率のおさらい：「同時確率」，「条件付き確率」





$P(A)$：事象Aの起こる確率

$P(A \cap B)$ ：AとBの同時確率

$P(A \mid B)$ ：Aが起こった前提でのBの生じる確率（条件付き確率）

同時確率と条件付き確率の違いを理解する！！

ベイズの定理

$$P(A \bigcap B) = P(B \bigcap A)$$

$$P(A \bigcap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$$

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

病気の検査の例題を通じて

検査で陽性反応が出た場合に，本当に病気である確率は？

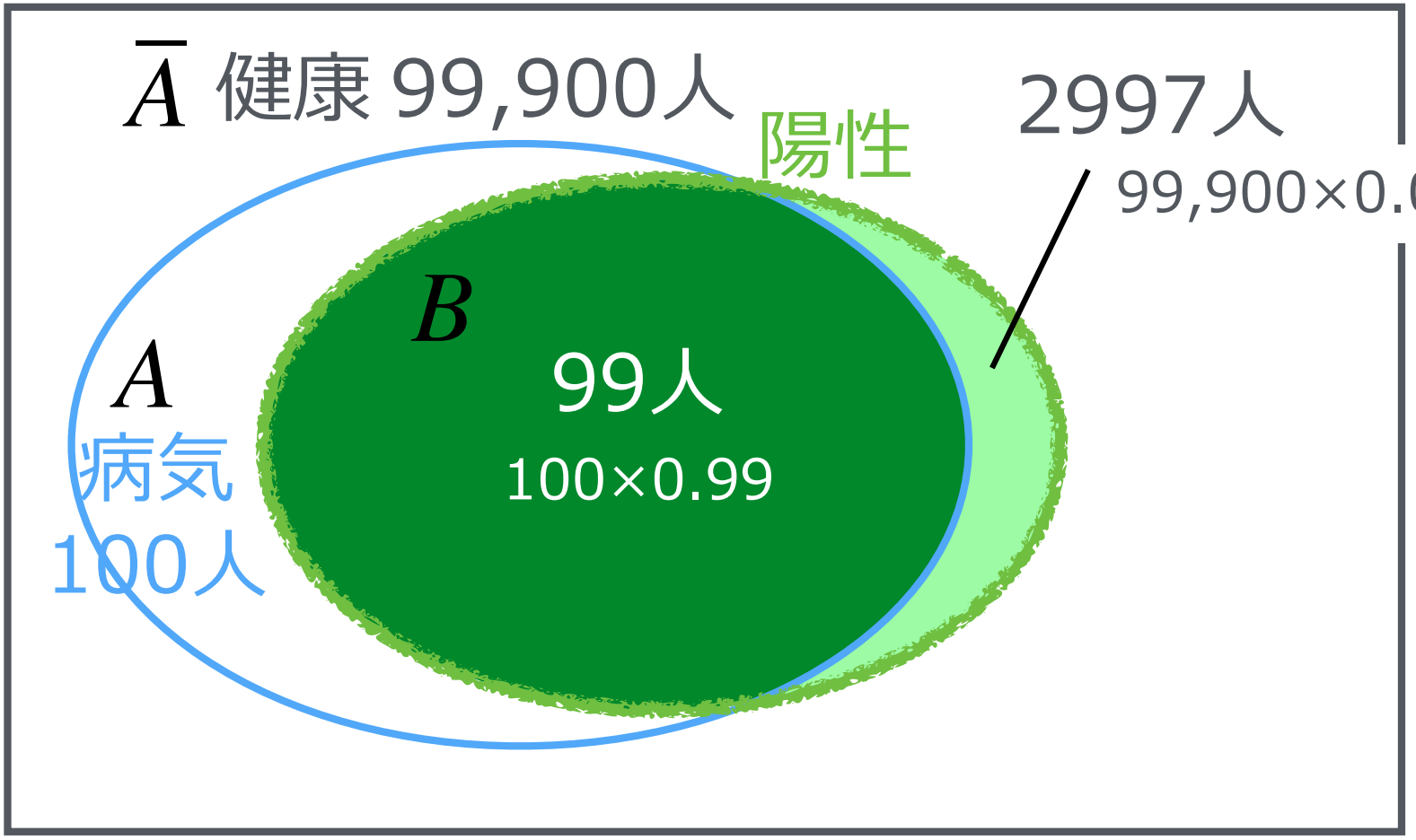
- 日本全体でその病気の患者の割合は0.1%である
- ある病気の検査方法は，
  - 病気の人が受けた場合，99%の人が陽性反応を示す
  - 健康な人が受けた場合，3%の人が陽性反応を示す

- 事象 A： 病気である
- 事象 $\overline{A}$： 健康である
  （事象Aの余事象）
- 事象 B： 陽性反応が出る

$$P(A) = 0.001$$
$$P(\overline{A}) = 0.999$$
$$P(B \mid A) = 0.99$$
$$P(B \mid \overline{A}) = 0.03$$

同時確率と条件付き確率

人口が10万人だとすると



$$P(B \mid A) = 0.99$$
$$P(B \mid \overline{A}) = 0.03$$

$$P(A \cap B) = \frac{99}{100{,}000} = 0.00099$$

病気の検査の例題を通じて

検査で陽性反応が出た場合に，本当に病気である確率は？

$$P(A \mid B)?$$

陽性反応が出た人のうち，病気である人の割合

人口が10万人だとすると

健康で陽性反応が出る人数は 2997人
病気で陽性反応が出る人数は 99人
陽性反応が出る人数の合計は 2997人+99人=3096人

$$P(A \mid B) = \frac{99}{3096} = 0.032$$

検査で陽性でも悲観することはない！！

ベイズの定理を使って確率を求める

問題から

$$P(A)=0.001 \quad P(\overline{A})=0.999 \quad P(B \mid \overline{A})=0.03$$

$$P(B \mid A)=0.99$$

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

$$=\frac{0.99\times 0.001}{0.99\times 0.001+0.03\times 0.999}=0.032$$

事前確率，事後確率，尤度，規格化定数

事後確率　尤度　事前確率

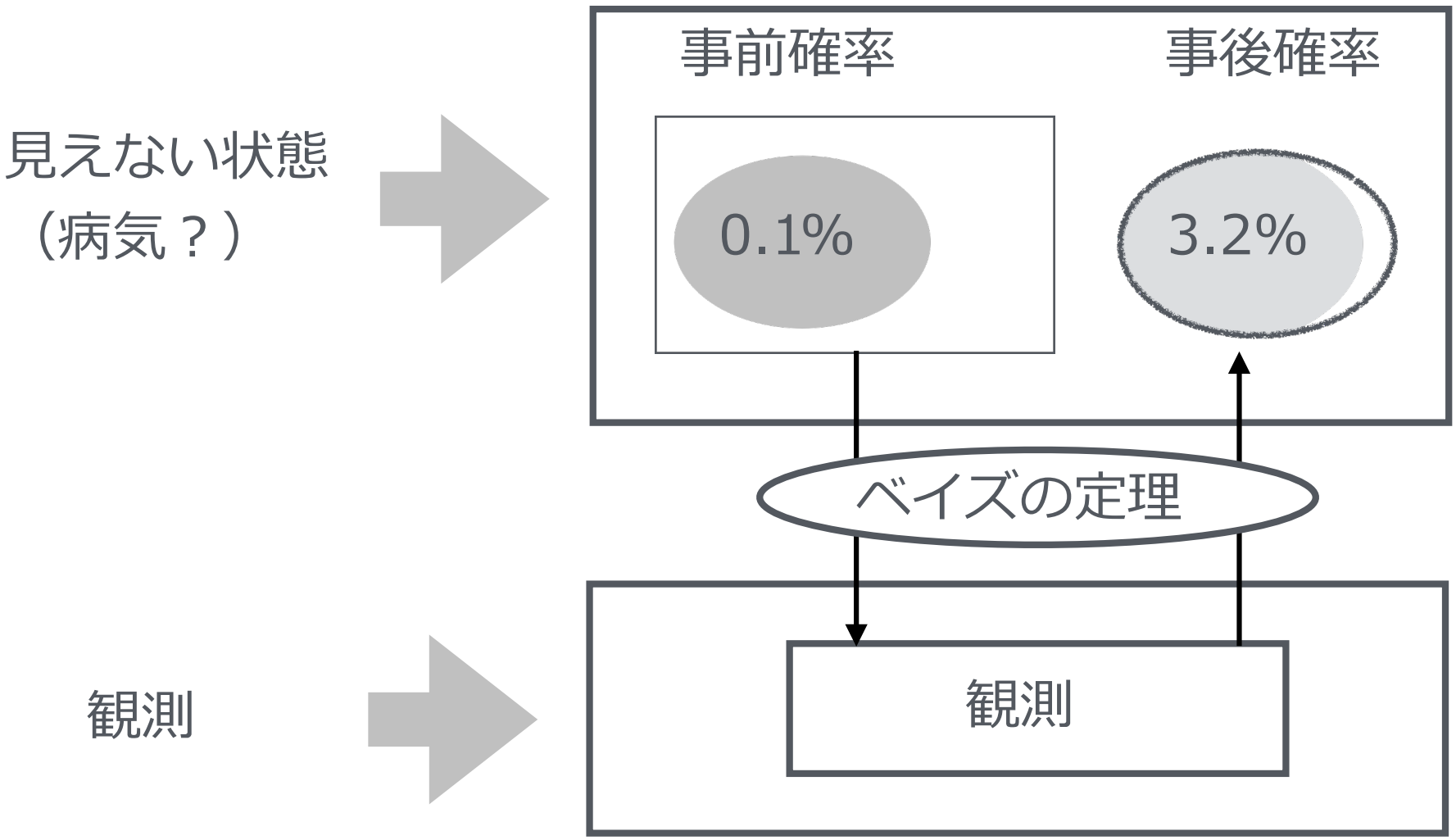
$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

規格化定数

（事後確率を0-1に収めるため）

P(A)をBという情報（観測結果）を得て，
確率を更新した，と考えることができる

ベイズの定理：観測により見えない状態を推測



別の検査を受けました

精密検査を受けるために別の検査で
陽性反応が出た場合に，本当に病気である確率は？

- 病気の人が受けた場合，98%の人が陽性反応を示す
- 健康な人が受けた場合，2%の人が陽性反応を示す

- 事象 A： 病気である
- 事象 $\overline{A}$： 健康である
  （事象Aの余事象）
- 事象 B： 陽性反応が出る

1回目の検査の事後確率を使う

$$P(A) = 0.032$$

$$P(\overline{A}) = 0.968$$

$$P(B \mid A) = 0.98$$

$$P(B \mid \overline{A}) = 0.02$$

ベイズの定理を使って確率を求める

問題から

$$P(A)=0.032 \quad P(\overline{A})=0.968 \quad P(B \mid \overline{A})=0.02$$

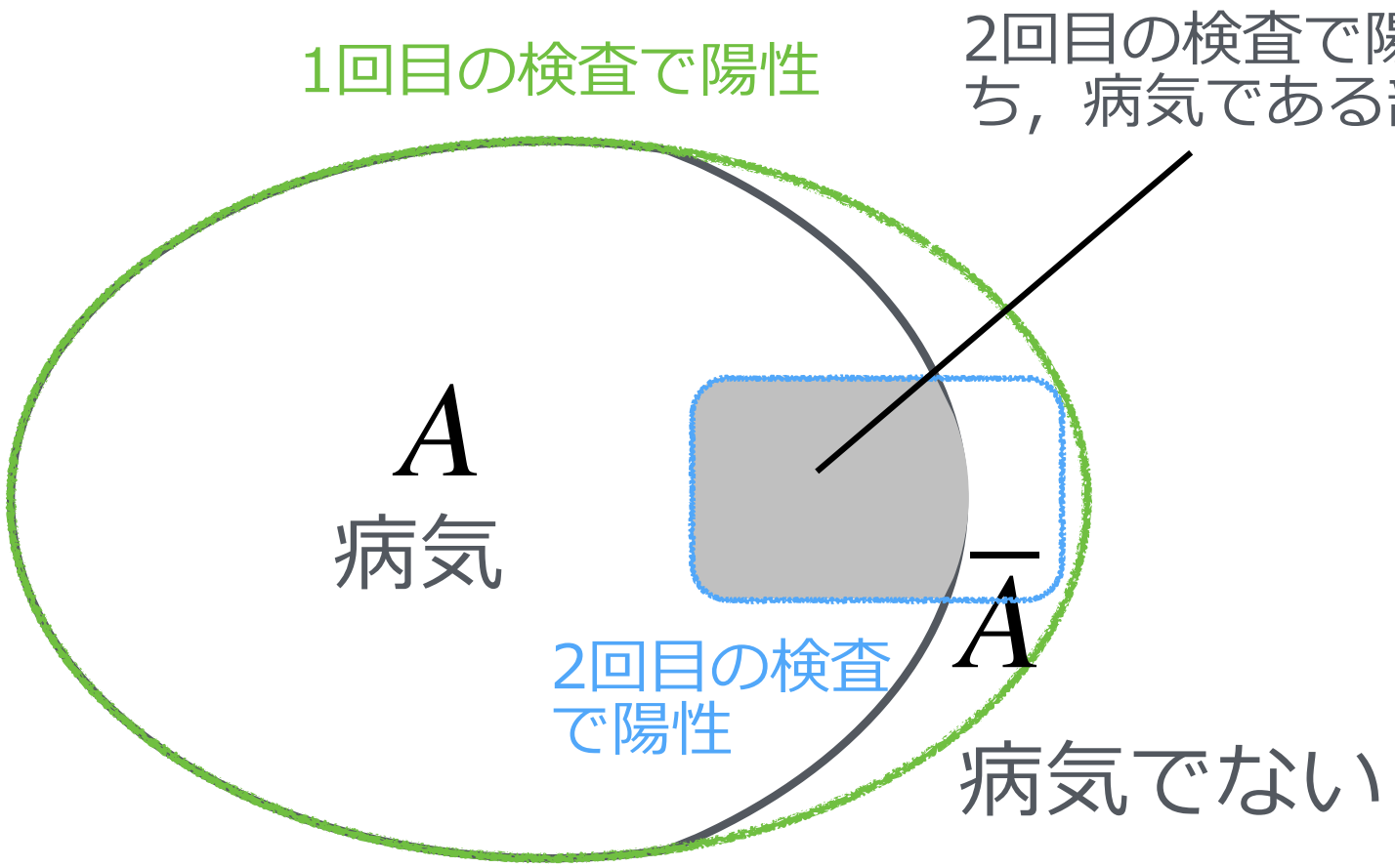
$$P(B \mid A)=0.98$$

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

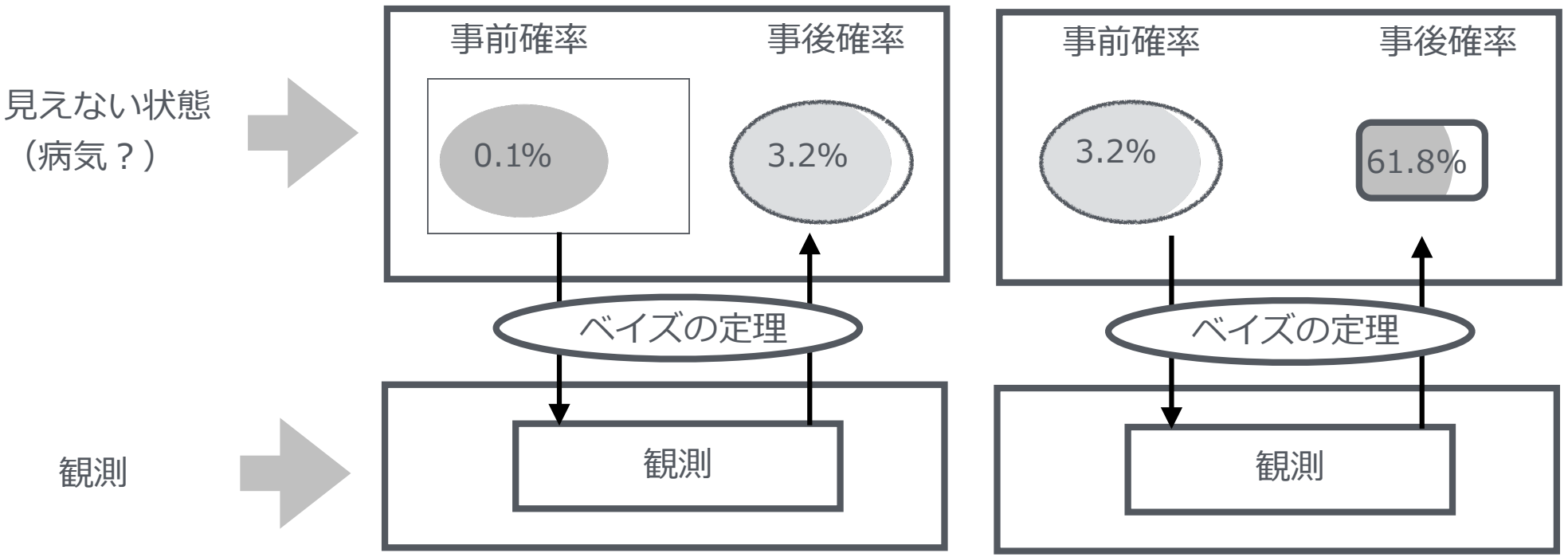
$$=\frac{0.98\times 0.032}{0.98\times 0.032+0.02\times 0.968}=0.618$$

かなり病気である確率が高くなってしまった．．．．

1回目の検査と2回目の検査の関係



観測を繰り返すことにより真の状態を明らかにする



ベイズ定理に基づく
スパムフィルタ

例

- メールA

  高信頼システムの授業では，**機械学習**を用いた高信頼化を勉強します．資料は下記のURLからダウンロード．．．．．

- メールB

  アイドルの**写真**を無料で差し上げます．1日以内に下記のURLからお申し込みください．

  ．．．．．

注目する単語と事前情報

迷惑メール/通常メールに”注目単語”が含まれる確率

検出語	迷惑メール	通常メール
アイドル (W_1)	0.75	0.125
無料 (W_2)	0.7	0.2
機械学習 (W_3)	0.2	0.5

インターネットでの迷惑メールと通常メールの比率は6：4とする

問題設定

あるメールを調べたところ，”アイドル”，”無料”，
”機械学習”という単語が検出された．このメールは
スパムメールとすべきか？ それとも通常メールとすべきか？

条件付き確率と事前確率

- メールがスパムメールである事象をA
- メールが単語W_iを含む事象をW_i

$$P(W_1 \mid A) = 0.75 \quad P(W_1 \mid \overline{A}) = 0.125$$
$$P(W_2 \mid A) = 0.7 \quad P(W_2 \mid \overline{A}) = 0.2$$
$$P(W_3 \mid A) = 0.2 \quad P(W_3 \mid \overline{A}) = 0.5$$

迷惑メールである事前確率 $P(A)=0.6$

通常メールである事前確率 $P(\overline{A})=0.4$

※ メールの単語間の相関を無視 → ナイーブベイズ

“アイドル” W_1に関するベイズの定理

メールから”アイドル”が検出された時，
そのメールがスパムメールである確率

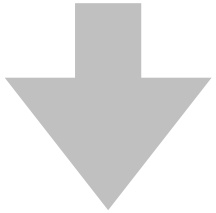
$$P(A \mid W_1) = \frac{P(W_1 \mid A)P(A)}{P(W_1)}$$

メールから”アイドル”が検出された時，
そのメールが通常メールである確率

$$P(\overline{A} \mid W_1) = \frac{P(W_1 \mid \overline{A})P(\overline{A})}{P(W_1)}$$

“アイドル” W_1に関するベイズの定理

$P(A \mid W_1)$ と $P(\overline{A} \mid W_1)$の大小関係がわかれば良い.



共通の分母は無視して良い

$$P(A \mid W_1) \propto P(W_1 \mid A)P(A)$$

$$P(\overline{A} \mid W_1) \propto P(W_1 \mid \overline{A})P(\overline{A})$$

事前確率としては，P(A)=0.6, P($\overline{A}$)=0.4を使って良い

$$P(A \mid W_1) \propto P(W_1 \mid A)0.6$$

$$P(\overline{A} \mid W_1) \propto P(W_1 \mid \overline{A})0.4$$

“無料” W_2に関するベイズの定理

$$P(A \mid W_2) \propto P(W_2 \mid A)P(A)$$
$$P(\bar{A} \mid W_2) \propto P(W_2 \mid \bar{A})P(\bar{A})$$

事前確率として何を使うか？

“アイドル” W_1に対する事後確率

$$P(A \mid W_2) \propto P(W_2 \mid A)P(A \mid W_1)$$
$$P(\bar{A} \mid W_2) \propto P(W_2 \mid \bar{A})P(\bar{A} \mid W_1)$$

“機械学習” W_3に関するベイズの定理

$$P(A \mid W_3) \propto P(W_3 \mid A)P(A)$$
$$P(\overline{A} \mid W_3) \propto P(W_3 \mid \overline{A})P(\overline{A})$$

事前確率として何を使うか？

“無料” W_2に対する事後確率

$$P(A \mid W_3) \propto P(W_3 \mid A)P(A \mid W_2)$$
$$P(\overline{A} \mid W_3) \propto P(W_3 \mid \overline{A})P(\overline{A} \mid W_2)$$

W_1，W_2，W_3に関するベイズ定理をまとめる

$$P(A \mid W_3, W_2, W_1) \propto P(W_3 \mid A) P(W_2 \mid A) P(W_1 \mid A) P(A)$$
$$P(\overline{A} \mid W_3, W_2, W_1) \propto P(W_3 \mid \overline{A}) P(W_2 \mid \overline{A}) P(W_1 \mid \overline{A}) P(\overline{A})$$

$$P(A \mid W_3, W_2, W_1) \propto 0.2 \times 0.7 \times 0.75 \times 0.6 = 0.063$$
$$P(\overline{A} \mid W_3, W_2, W_1) \propto 0.5 \times 0.2 \times 0.125 \times 0.4 = 0.005$$

スパムメールに分類すべき！

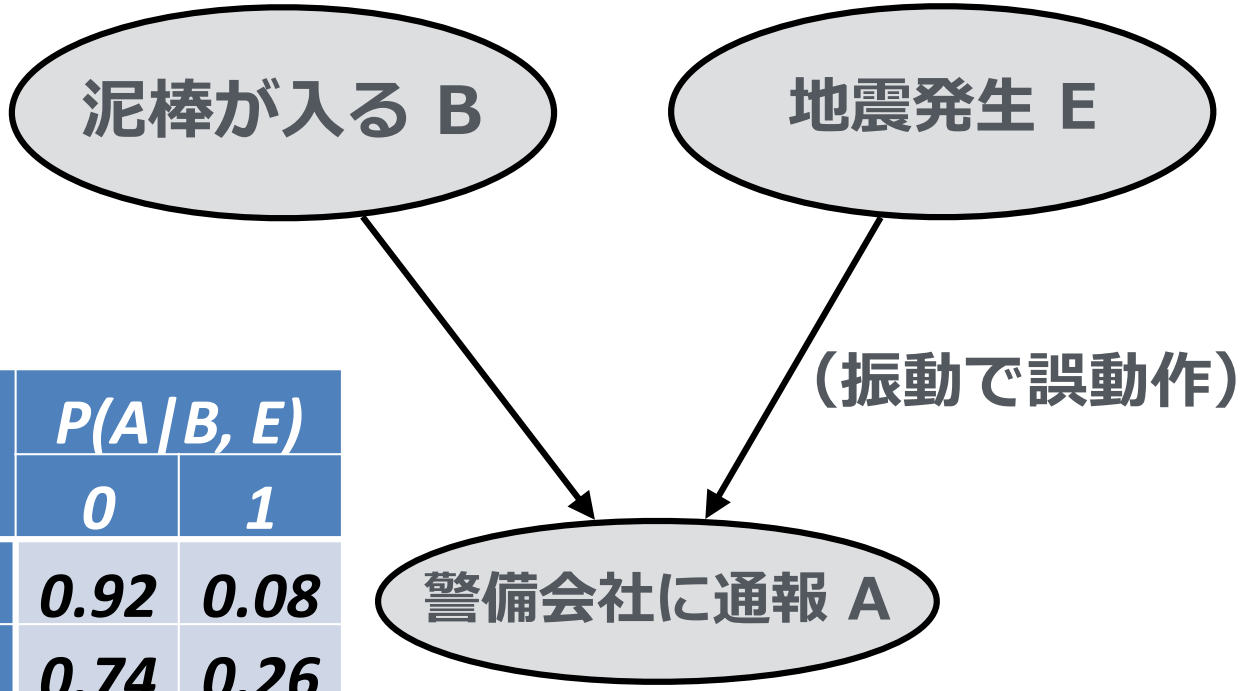
ベイジアンネットワーク

ベイジアンネットワーク

- ベイズの定理をネットワークに拡張
- 依存関係の大きさを条件付き確率で表す

例：空家の自動警備システム

B	P(B)
0	0.99
1	0.01



E	P(E)
0	0.98
1	0.02

B	E	P(A|B, E)	
		0	1
0	0	0.92	0.08
0	1	0.74	0.26
1	0	0.06	0.94
1	1	0.05	0.95

例題1

泥棒(Burglar)が入って，警備会社に通報が行く確率を求めよ．ただし，地震は同時に起こっていないとする．また，**事象B,Eは独立**であるとする．

$$P(B \cap E) = P(B)P(E)$$

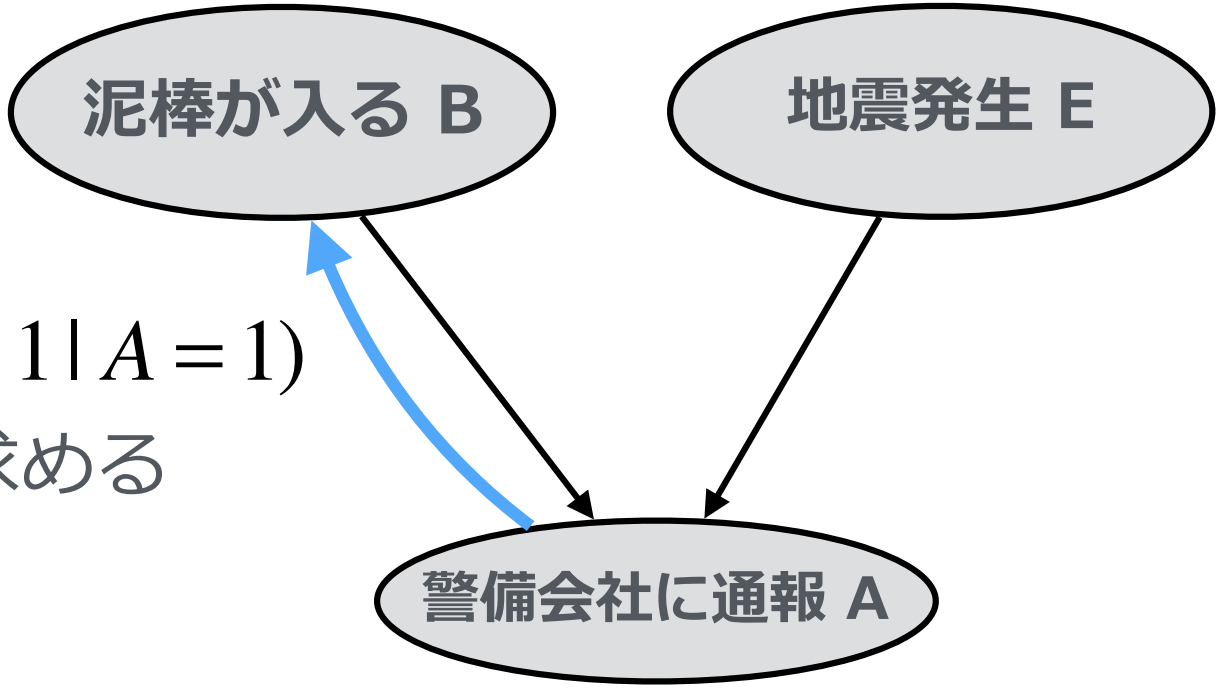
$$P(E=0, B=1, A=1) = P(A=1 \mid B=1, E=0)P(E=0)P(B=1)$$
$$= 0.94 \times 0.98 \times 0.01 = 0.009212$$

例題2

警備会社に通報（A）が入った時に，泥棒(B）が原因であった確率を求めよ

警備会社に通報（A）が入った → A＝1

泥棒が入った　→ B=1



逆確率

$P(B=1 \mid A=1)$

を求める

ベイズの定理より

$$P(B=1 \mid A=1) = \frac{P(A=1 \mid B=1)P(B=1)}{P(A=1)}$$

$P(A=1 \mid B=1)$

$$= P(A=1 \mid B=1, E=1)P(E=1) + P(A=1 \mid B=1, E=0)P(E=0)$$

$$= 0.95 \times 0.02 + 0.94 \times 0.98$$

$$= 0.9402$$

$$P(A=1)$$

$$= P(A=1,B=0,E=0)+P(A=1,B=0,E=1)$$

$$+P(A=1,B=1,E=0)+P(A=1,B=1,E=1)$$

$$= P(A=1\mid B=0,E=0)P(B=0,E=0)+P(A=1\mid B=0,E=1)P(B=0,E=1)$$

$$+P(A=1\mid B=1,E=0)P(B=1,E=0)+P(A=1\mid B=1,E=1)P(B=1,E=1)$$

$$= P(A=1\mid B=0,E=0)P(B=0)P(E=0)+P(A=1\mid B=0,E=1)P(B=0)P(E=1)$$

$$+P(A=1\mid B=1,E=0)P(B=1)P(E=0)+P(A=1\mid B=1,E=1)P(B=1)P(E=1)$$

$$= 0.08\times 0.99\times 0.98+0.26\times 0.99\times 0.02+0.94\times 0.01\times 0.98+0.95\times 0.01\times 0.02$$

$$=0.092166$$

したがって

$$P(B=1\mid A=1)=\frac{0.9402\times 0.01}{0.092166}=0.10$$

ベイジアンネットワークの展開

- 不確定な観測結果があっても推測可能
- 主観的・曖昧なデータを定量化できる

- センサーなどの観測結果から機器の故障原因の特定
- 医療：検査結果から病気の特定
- 人間の嗜好のモデル化→レコメンデーションなど