北数教　第109回数学教育実践研究会　令和元年6月1日（土）
北海道大学理学部5号館203号室

因果関係を確率によって記述する

原因の確率からベイジアン・ネットワークへ

AIはどのようにして意思決定をするのか

AI（人工知能）というフレーズを耳にすることが多い。IBM「ワトソン」、NEC「the Wise」、Sony「Neural Network Console」等のAIソフトも普及して、予想外の因果関係が発見されて、医学や社会に影響を与えている。その基本的な数学理念は、ベイズの定理である。

松本睦郎（札幌啓成高等学校　講師）

Episode1 原因の確率

【例題１】早稲田大学入試問題

５回に１回の割合で帽子を忘れるくせのあるK君が、正月に、A, B, Cの３軒をこの順番に年始まわりをして自宅に帰ったとき、帽子を忘れてきたことに気が付いた。２番目のB宅に忘れてきた確率を求めよ。

〔解答例〕

「帽子を忘れた」(結果）が時間的に先行する「A宅, B宅, C宅」（原因）の確率を求める問題です。

X: ３軒のどこかに帽子を忘れる事象

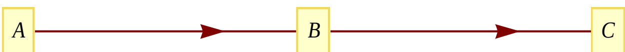
$\overline{X}$: : 帽子を忘れない事象

$$P(\overline{X}) = \left(\frac{4}{5}\right)^3, P(X) = 1 - \left(\frac{4}{5}\right)^3 = \frac{61}{125}$$

「Bで忘れる」⇔「Aで忘れない」×「Bで忘れる」

$$\mathrm{P}(B \cap X) = \frac{4}{5} \times \frac{1}{5} = \frac{4}{25}$$

$$P(B|X) = \frac{P(B \cap X)}{P(X)} = \frac{20}{61} \cdot\cdot\cdot (答)$$



（有向非巡回グラフ↑）

Episode 2 ベイズの定理

【例題２】天気予報問題

雨が降る確率0.4, 雨が降らない確率0.6

雨が降る時、窓を閉める確率0.7

雨が降る時、窓を閉めない確率0.3

雨が降らないとき窓を閉める確率0.05

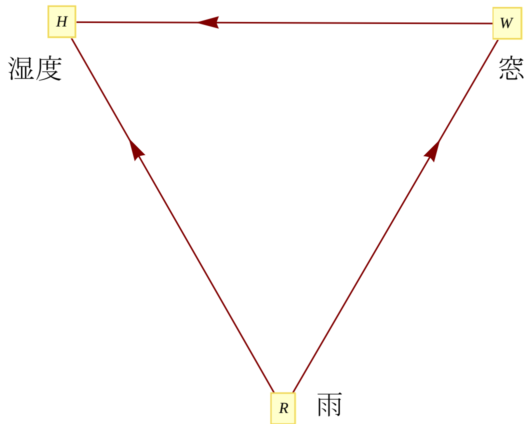
雨が降らないとき窓を閉めない確率0.95

観測結果（確率）

雨	窓	低湿度	高湿度
降る	閉める	0.1	0.9
降らない	閉める	0.2	0.8
降る	閉めない	0.35	0.65
降らない	閉めない	0.99	0.01

このとき、湿度が高いときの雨の降る確率？

〔解答例〕



R:雨が降る事象、W:窓を閉める事象

H:湿度が上がる事象

T：成立True　F：不成立False

$\mathrm{P}(W|R)$

W:窓 / R:雨	窓閉 WT	窓開 WF
雨降る RT 0.4	0.7	0.3
雨降らない RF 0.6	0.05	0.95

$\mathrm{P}(H|W,R)$

窓,雨（W,R）	高湿度　HT	低湿度　HF
T&T	0.9	0.1
T&F	0.8	0.2
F&T	0.65	0.35
F&F	0.01	0.99

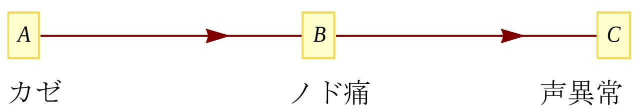
$$P(\text{雨}|\text{高湿度}) = P(RT|HT)$$

$$= \frac{P(R \cap H)}{P(HT)} = \frac{P(HT|RT)P(RT)}{P(HT)}$$

$$= \frac{0.9 \times 0.7 \times 0.4 + 0.65 \times 0.3 \times 0.4}{0.9 \times 0.7 \times 0.4 + 0.8 \times 0.05 \times 0.6 + 0.65 \times 0.3 \times 0.4 + 0.01 \times 0.6 \times 0.95}$$

Episode 3 ベイジアン・ネットワークへ

【例題３】医学的意思判断決定問題（単純な問題）



「A:カゼ」から「B:ノド痛」を引き起こし、「B:ノド痛」から「C:声異常」をもたらすものとする。

声異常のあるときのカゼである確率を求める。

観測結果（確率）

〔解答例〕　　　T：True　F：False

「カゼ」	$^1/_{10}$
「カゼ」から「ノド痛」へ	$^3/_5$
「カゼでない」から「ノド痛」へ	$^1/_5$
「ノド痛」から「声異常」へ	$^2/_3$
「ノド痛でない」から「声異常」へ	$^1/_4$

$$P(AT) = \frac{1}{10}, P(AF) = \frac{9}{10}$$

$$P(BT|AT) = \frac{3}{5}, P(BT|AF) = \frac{1}{5}$$

$$\frac{P(BT \cap AT)}{P(AT)} = \frac{3}{5}, \frac{P(BT \cap AF)}{P(AF)} = \frac{1}{5}$$

$$\therefore P(BT \cap AT) = \frac{3}{5} \times \frac{1}{10}, P(BT \cap AF) = \frac{1}{5} \times \frac{9}{10}$$

$$P(CT|BT) = \frac{2}{3}, P(CT|BF) = \frac{1}{4}$$

$$\frac{P(BT \cap CT)}{P(BT)} = \frac{2}{3}, \frac{P(BF \cap CT)}{P(BF)} = \frac{1}{4}$$

$$\therefore P(BT \cap CT) = \frac{2}{3}P(BT), P(BF \cap CT) = \frac{1}{4}P(BF)$$

声異常のあるときのカゼである確率は、

$$P(AT|CT) = \frac{P(AT \cap CT)}{P(CT)} \cdots (*)$$

$$P(CT) = P(BT \cap CT) + P(BF \cap CT)$$

$$= \frac{2}{3}P(BT) + \frac{1}{4}P(BF) \cdots ①$$

$$P(BT) = P(BT \cap AT) + P(BT \cap AF) = \frac{3+9}{50} = \frac{12}{50}$$

$$P(BF) = 1 - P(BT) = 1 - \frac{12}{50} = \frac{38}{50}$$

①より $P(CT) = \frac{2}{3} \times \frac{12}{50} + \frac{1}{4} \times \frac{38}{50} = \frac{7}{20} \cdots ②$

$$P(CT|AT) = \frac{P(CT \cap AT)}{P(AT)}$$

$$= P(BT|AT) \times P(CT|BT) + P(BF|AT) \times P(CT|BF)$$

$$= \frac{3}{5} \times \frac{2}{3} + \frac{2}{5} \times \frac{1}{4}$$

$$\frac{P(CT \cap AT)}{P(AT)} = \frac{3}{5} \times \frac{2}{3} + \frac{2}{5} \times \frac{1}{4}$$

	BT	BF
AT: $\frac{1}{10}$	$\frac{3}{5}$	$\frac{2}{5}$
AF: $\frac{9}{10}$	$\frac{1}{5}$	$\frac{4}{5}$

	CT	CF
BT	$\frac{2}{3}$	$\frac{1}{3}$
BF	$\frac{1}{4}$	$\frac{3}{4}$

$$P(AT \cap CT) = P(AT) \times \left(\frac{2}{5} + \frac{1}{10}\right) = \frac{1}{20}$$

(*)　②より

$$P(AT|CT) = \frac{P(AT \cap CT)}{P(CT)} = \frac{\frac{1}{20}}{\frac{7}{20}} = \frac{1}{7}$$

「声に異常がある」とき、「カゼ」が原因である確率は、14.3%である。

【例題４】医学的意思判断決定問題（複雑な問題）

「A:転移性腫瘍」は脳に転移して「C:脳腫瘍」になりあるいは、「B:高カルシウム血症」から「D:昏睡」を引き起こすことがある。また「C:脳腫瘍」は「E:激しい頭痛」をもたらす。ある受診者は、「D:昏睡」はないが「E:激しい頭痛」があるとき「A:転移性腫瘍」があるか？昏睡はないが激しい頭痛のとき、転移性腫瘍の確率を求める。

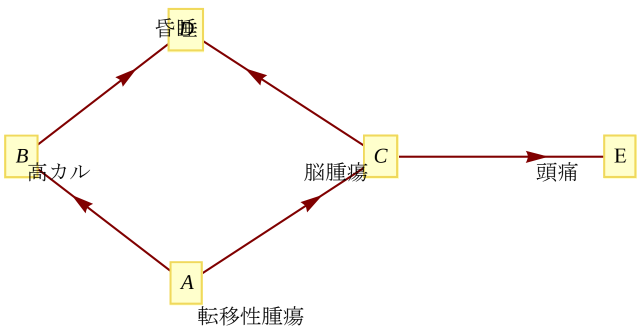
観測結果（確率）

「転移性腫瘍」	0.1
「転移性腫瘍」から「高カルシウム血症」へ	0.8
「転移性腫瘍でない」から「高カルシウム血症」へ	0.1
「転移性腫瘍」から「脳腫瘍」へ	0.3
「転移性腫瘍でない」から「脳腫瘍」へ	0.2
「高カルシウム血症」かつ「脳腫瘍」から「昏睡」へ	0.6
「高カルシウム血症でない」かつ「脳腫瘍」から「昏睡」へ	0.5
「高カルシウム血症」かつ「脳腫瘍でない」から「昏睡」へ	0.4
「高カルシウム血症でない」かつ「脳腫瘍でない」から「昏睡」へ	0.2
「脳腫瘍」から「頭痛」へ	0.9
「脳腫瘍でない」から「頭痛」へ	0.7

T：True　F：False

これらの確率は Big Data より求められたものとする。

〔解答例〕



	$B\cap C$ $(BT)\&(CT)$	$\overline{B}\cap C$ $(BF)\&(CT)$	$B\cap\overline{C}$ $(BT)\&(CF)$	$\overline{B}\cap\overline{C}$ $(BF)\&(CF)$
AT 0.1	0.8×0.3 0.24	0.2×0.3 0.06	0.8×0.7 0.56	0.2×0.7 0.14
AF 0.9	0.1×0.2 0.02	0.9×0.2 0.18	0.1×0.8 0.08	0.9×0.8 0.72

	$B\cap C$ $(BT)\&(CT)$	$\overline{B}\cap C$ $(BF)\&(CT)$	$B\cap\overline{C}$ $(BT)\&(CF)$	$\overline{B}\cap\overline{C}$ $(BF)\&(CF)$
DF	$1-0.6$ 0.4	$1-0.5$ 0.5	$1-0.4$ 0.6	$1-0.2$ 0.8
ET	0.9	0.9	0.7	0.7
DF&ET	0.4×0.9 0.36	0.5×0.9 0.45	0.6×0.7 0.42	0.8×0.7 0.56

「昏睡はないが頭痛がある」とき「転移性腫瘍」の確率を求める。

$$P(AT|DF\&ET)=\frac{P(A\cap\overline{D}\cap E)}{P(\overline{D}\cap E)}$$

ベイズの定理より

$$\mathrm{P}(B\cap C)=P(B\cap C|A)P(A)+P(B\cap C|\overline{A})P(\overline{A})$$
$$=0.24\times0.1+0.02\times0.9=0.042$$
$$\mathrm{P}(\overline{B}\cap C)=P(\overline{B}\cap C|A)P(A)+P(\overline{B}\cap C|\overline{A})P(\overline{A})$$
$$=0.06\times0.1+0.18\times0.9=0.168$$
$$\mathrm{P}(B\cap\overline{C})=P(B\cap\overline{C}|A)P(A)+P(B\cap\overline{C}|\overline{A})P(\overline{A})$$
$$=0.56\times0.1+0.08\times0.9=0.128$$
$$\mathrm{P}(\overline{B}\cap\overline{C})=P(\overline{B}\cap\overline{C}|A)P(A)+P(\overline{B}\cap\overline{C}|\overline{A})P(\overline{A})$$
$$=0.14\times0.1+0.72\times0.9=0.662$$

分母の$\mathrm{P}(\overline{D}\cap E)$　を求める。

$$\mathrm{P}(\overline{D}\cap E)$$
$$=P(\overline{D}\cap E|B\cap C)P(B\cap C)$$
$$+P(\overline{D}\cap E|\overline{B}\cap C)P(\overline{B}\cap C)$$
$$+P(\overline{D}\cap E|B\cap\overline{C})P(B\cap\overline{C})$$
$$+P(\overline{D}\cap E|\overline{B}\cap\overline{C})P(\overline{B}\cap\overline{C})$$
$$=0.36\times0.042+0.45\times0.168+0.42\times0.128$$
$$+0.56\times0.662=0.5152$$

分子の$\mathrm{P}(A\cap\overline{D}\cap E)$を求める。

表 AT の部分

$$\mathrm{P}(B\cap C)=P(B\cap C|A)P(A)=0.24\times0.1=0.024$$
$$\mathrm{P}(\overline{B}\cap C)=P(\overline{B}\cap C|A)P(A)=0.06\times0.1=0.006$$
$$\mathrm{P}(B\cap\overline{C})=P(B\cap\overline{C}|A)P(A)=0.56\times0.1=0.056$$
$$\mathrm{P}(\overline{B}\cap\overline{C})=P(\overline{B}\cap\overline{C}|A)P(A)=0.14\times0.1=0.014$$

$$\mathrm{P}(\overline{D}\cap E)=P(\overline{D}\cap E|B\cap C)P(B\cap C)$$
$$+P(\overline{D}\cap E|\overline{B}\cap C)P(\overline{B}\cap C)$$
$$+P(\overline{D}\cap E|B\cap \overline{C})P(B\cap \overline{C})$$
$$+P(\overline{D}\cap E|\overline{B}\cap \overline{C})P(\overline{B}\cap \overline{C})$$
$$=0.36\times 0.024+0.45\times 0.006+0.42\times 0.056$$
$$+0.56\times 0.00784=0.0427$$

$$P(AT|DF\&ET)=\frac{P(A\cap\overline{D}\cap E)}{P(\overline{D}\cap E)}=\frac{0.0427}{0.512}=0.0833984$$

「昏睡はないが頭痛がある」とき「転移性腫瘍」の確率は、8.33%となる。

Episode 4 病名のベイズ診断

「原因」「結果」の因果関係を「病名」「症状」として扱うと、ベイズの定理を活用して病名診断に活用することもできる。

病名を D1,D2、症状を S1,S2,S3 とする。３症状からベイズ診断をおこなう。ただし架空の診断である。

D1:肺気腫、D2:肺炎

S1:息苦しい。S2:せき込んで黄色のタンがでる。

S3:胸が苦しい。

$\mathrm{P}(Sj|Di)$

Di	$\mathrm{P}(Di)$	S1	S2	S3
D1	0.23	0.10	0.70	0.60
D2	0.77	0.80	0.20	0.50

症状（S1,S2,S3）の状態が（1,0,1）のとき、D1 の確率は、ベイズの定理より

$\mathrm{P}(D1|1,0,1)$

$$=\frac{0.23\times 0.1\times(1-0.7)\times 0.6}{0.23\times 0.1\times(1-0.7)\times 0.6+0.77\times 0.8\times(1-0.2)\times 0.5}$$

$$=0.01652$$

Episode 5 原子力潜水艦スコーピオン捜索

1968 年米国の原子力潜水艦スコーピオンが大西洋で行方不明になった事件があった。

沈没したとみられる潜水艦を発見するのに、ベイズの定理を用いた捜索方法を用いた。海図上を分割し沈没しているとみられる事前確率を経験から推定する。

その海域に沈没している事前確率をpとする。

沈没しているという条件下で、発見する確率をqとする。

A：沈没している事象。B：発見する事象。

$$\mathrm{P}(A)=p,\mathrm{P}(B|A)=q,P(\bar{B}|A)=1-q$$
$$\mathrm{P}(\bar{A})=1-p,P(B|\bar{A})=0,P(\bar{B}|\bar{A})=1$$

沈没していない海域では、発見する確率は 0 であり、沈没していない海域で、発見されない確率は 1 である。

その海域で、発見されないとき沈んでいる確率$\mathrm{P}(A|\bar{B})$を求める。

$$\mathrm{P}(A|\bar{B})=\frac{P(A\cap\bar{B})}{P(\bar{B})}\quad・・・①$$

$$\mathrm{P}(B|A)=q$$

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

$$\therefore \mathrm{P}(A\cap B)=p\times q\quad・・・②$$

$$P(\bar{B}|A)=1-q$$

$$\frac{P(A\cap\bar{B})}{P(A)}=1-q$$

$$\therefore \mathrm{P}(A\cap\bar{B})=p\times(1-q)\quad・・・③$$

①　の分子へ代入

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

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

$$\therefore \mathrm{P}(\bar{A}\cap\bar{B})=1-p\quad・・・④$$

③④より

$$\mathrm{P}(\bar{B})=P(A\cap\bar{B})+P(\bar{A}\cap\bar{B})=p(1-q)+1-p$$

①　の分母へ代入

$$\mathrm{P}(A|\bar{B})=\frac{P(A\cap\bar{B})}{P(\bar{B})}=\frac{p(1-q)}{p(1-q)+1-p}$$

指定された海域で、発見されないときに、沈没している確率になる。

（おわり）

【引用図書】

「入門ベイズ統計」松原望著

https://ja.wikipedia.org/wiki