本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？
3. ベイズの定理の数式
4. 関連する数理統計

検査したら「陽性」だった！

感染症の検査をしたら「陽性」だった！
但し、検査キットの「精度は 80 %」

⇩

80 % の確度で感染症に罹った？

⇩

でも、
検査キットはあまり信頼できない…

本当に感染したのか？

緊急なので簡易的に検査したけど、
検査キットの「精度は 80 %」

今日のテーマ

「ベイズの定理」で調べると…

本当に感染している確率は
意外と低い！

実際には
3割程度

小林道正「デタラメにひそむ確率法則」岩波科学ライブラリー2012

精度80%の意味は？

「陽性」と判別されても、20% は間違い
→ 本当は感染していない（偽陽性）

病院に通っ
ている人々

本当に感染している人々を検査すると、
8割が「陽性」、2割が「陰性」と判別

街中の人々

感染の真偽が不明な人々を検査したら、
何割が本当に感染しているのか？

どちらの場合が深刻？

場合1

事象の発生回数（100人のデータ）

	本当に感染している	全く感染していない
検査で陽性と判別	80	0
検査で陰性と判別	20	0

↑
検査キットの
「精度は 80 %」

8割の確度で
感染している（深刻）！

場合2

事象の発生回数（100人のデータ）

	本当に感染している	全く感染していない
検査で陽性と判別	8	18
検査で陰性と判別	2	72

↑
検査キットの
「精度は 80 %」

感染率が1割なので
8割よりも低いかも？

今知りたいのは...

検査で「陽性」と
判別された人々のうち

本当に感染している
人々の割合

陽性と判別された条件で
本当に感染している割合

条件つき確率

場合2

事象の発生回数（100人のデータ）

	本当に感染している	全く感染していない
検査で陽性と判別	8	18
検査で陰性と判別	2	72

$$\frac{8}{8+18} = 0.308$$

実際には
3割程度

69.2 % は感染していない！

本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？
3. ベイズの定理の数式
4. 関連する数理統計

---

① 本当に感染した確率を計算したい

検査精度 と 感染率 で計算するには？

検査の精度は
- 感染している人々を陽性と判別 80 %（真の陽性）
- 感染していないのに陽性と判別 20 %（偽の陽性）

地域の感染率 10 %

③ 青文字の数字だけで計算する

④ ベイズの定理を使う

② 人数のデータは使わない

検査して陽性でも
69.2% は感染していない

本当に感染した確率は
30.8 %

事象の発生回数（??人のデータ）

	本当に感染している	全く感染していない
検査で陽性と判断	??	??
検査で陰性と判断	??	??

問題を 確率で表現 しよう

$P(B|A)$ ⇐ 感染している人々を陽性と判別 80 %

$P(B|A^c)$ ⇐ 感染していないのに陽性と判別 20 %

$P(A)$ ⇐ 地域で感染している人々の割合 10 %

ベイズの定理

$P(A|B)$

「陽性」と判別された人々の中で
真に感染している人々の割合

事象の表現

	感染している A	感染してない A^c
陽性と判別 B	$A \cap B$	$A^c \cap B$
陰性と判別 B^c	$A \cap B^c$	$A^c \cap B^c$

ベイズの定理

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

$P(B|A) = 0.8$

$P(B|A^c) = 0.2$

$P(A) = 0.1$

8割がた感染かと思ったら…

意外と低くて3割だった

$$\frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.2 \times 0.9}$$

$P(A|B) = 0.308$

「陽性」と判別された人が
真に感染している確率

「陽性」と判別されたけど
本当は感染していない確率
$P(A^c|B) = 0.602$

本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？　例題2
3. ベイズの定理の数式
4. 関連する数理統計

【問】 本当に感染している確率は？

簡易な検査キットで「陽性」となったけど、
本当は感染していないかも！？

必要な事前情報

- 簡易な検査キットがある。
- メーカーによる事前の実験によると、
  感染していないのに陽性であると
  間違って判別される確率は 20%
  正しく判定される確率は 80% とのこと。
- 今、住んでいる地域では、
  感染の割合は 10% である。

【解説】 確率で表現する

【知りたいこと】
・簡検査キットで「陽性」となったけど、
本当は感染していないかも!?

検査キットで「陽性」となった人々の中で、
↓
$$P(A^c|B) = ?$$
↑
実際には「陰性」であった人々の割合

【事前の情報】
・メーカーによる事前の実験によると
感染していないのに陽性であると
間違って判別される確率は 20%
正しく判定される確率は 80%

実際には「陰性」であった人々のなかで、
↓
$$P(B|A^c) = 0.2$$
↑
検査キットで「陽性」となった人々の割合

・今、住んでいる地域では、
感染の割合が 10% である。

$$P(A) = 0.1$$
↑
地域の実際の感染率
＝感染者数／地域の人口

ただし、
$$P(B|A) = 1 - P(B|A^c)$$ とする

【解説】 ベイズの定理を適用する

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

$$= \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.2 \times 0.9} = 0.31$$

$$P(B|A^c) = 0.2$$
$$P(A) = 0.1$$

排反

$$P(B|A) = 0.8$$
$$P(A^c) = 0.9$$

$$P(A^c|B) = 1 - P(A|B)$$
$$= 1 - 0.31$$
$$= 0.69$$

【知りたかったこと】
・簡検査キットで「陽性」となったけど、
本当は **69%の確率で** 感染していない！

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

本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？
3. ベイズの定理の数式
4. 関連する数理統計

例題3

ベイズの定理 で 出来ること

（既知）

本当に感染している人々の中で正しく「陽性」と判別される割合 → $P(B|A) = \alpha_P$

全く感染していない人々の中で誤って「陽性」と判別される割合 → $P(B|A^c) = \alpha_N$

これまでに検査した人々の中で本当に感染している人々の割合 → $P(A) = \beta$

ベイズの定理

$$\gamma = \frac{\alpha_P \beta}{\alpha_P \beta + \alpha_N (1-\beta)}$$

（未知）

「陽性」と判別された人々の中で本当に感染している人々の割合 ← $P(A|B) = \gamma$

「陽性」と判別された人々の中で本当は感染していない人々の割合 ← $P(A^c|B) = 1-\gamma$

簡易な検査キットは有効か？

検査キットの正しさ（既知）

★ 感染者を陽性と判別
$\alpha_P = 0.8$

★ 非感染者を陽性と誤認
$\alpha_N = 0.2$

場合1
症状があって、感染の
疑いがある人を検査した。
10人に1人が感染していた。
（β =0.1）

場合2
街中の出張検査所で
希望者を検査した。
1000人に1人が感染していた。
（β =0.01）

感染者の割合は？（未知）

$\gamma = 0.308$

陽性と判別ならば
30.8 %が感染

$\gamma = 0.0039$

陽性と判別されても
0.39 %しか感染してない

検査キットに必要な精度は？

検査結果を 90% 確信できる
→ 検査キットに必要な精度は？

確率で表現しよう↓

「陽性と判別された人が感染している割合」
が０．９となる場合について、 $P(A|B) = \gamma = 0.9$

「真の感染者が陽性と正しく判別される確率」
を求める $P(B|A) = \alpha_P = \alpha$

但し、

「非感染者が誤って陽性と判別される確率」 $P(B|A^c) = \alpha_N = 1 - \alpha$

「社会全体での感染者の割合（感染率）」 $P(A) = \beta = 0.01$

とする

高い精度が必要

ベイズの定理は、

$$\gamma = \frac{\alpha_P \beta}{\alpha_P \beta + \alpha_N (1-\beta)}$$

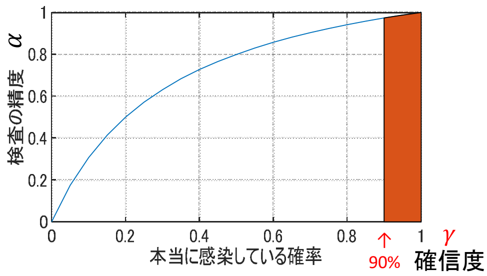
この例題では、

$$\gamma = \frac{\alpha \beta}{\alpha \beta + (1-\alpha)(1-\beta)}$$

αについて解くと、

$$\alpha = \frac{(1-\beta)\gamma}{(1-2\beta)\gamma + \beta}$$

確信度 90 % の精度を得るには
99.89 % の精度が必要！
（感染率 1 % のとき）



（感染 $\beta = 0.01$）

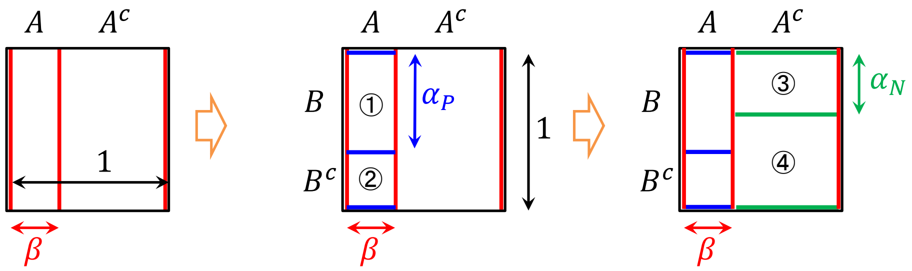
MATLAB

```
close all; clear all;
figure('Position',[10 10 300 150]);

pr=0.9; % 本当に感染している確率（必要な確信度）
bt=0.01; % 感染率

gm=0.0:0.05:1; % 本当に感染している確率
ap1=(1-bt)*gm;
ap2=(1-2*bt)*gm+bt;
ap =ap1./ap2;      % =検査の精度=特異度
ap=[0 ap]; gm=[0 gm];
plot(gm,ap); hold on;

gm=pr:0.05:1;
ap1=(1-bt)*gm;
ap2=(1-2*bt)*gm+bt;
ap =ap1./ap2;
area(gm,ap); hold on;
axis([0.0,1,0.0,1]); grid on;

xlabel('本当に感染している確率');
ylabel('検査の精度');

fprintf('感染率%4.2f 確信度%4.2f 必要精度%6.4f
¥n',bt,pr,ap(1));
```

本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？
3. ベイズの定理の数式
4. 関連する数理統計

予め分かっていること



$P(A)$
$$= \frac{\text{①}+\text{②}}{\text{①}+\text{②}+\text{③}+\text{④}}$$
$= \beta$

$P(B|A)$
$$= \frac{\text{①}}{\text{①}+\text{②}}$$
$= \alpha_P$

$P(B|A^c)$
$$= \frac{\text{③}}{\text{③}+\text{④}}$$
$= \alpha_N$

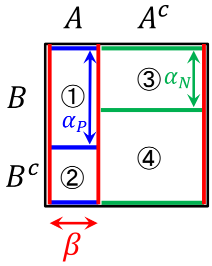
知りたいこと

分かっていること

$$P(A)=\frac{①+②}{①+②+③+④}=\beta$$

$$P(B|A)=\frac{①}{①+②}=\alpha_P$$

$$P(B|A^c)=\frac{③}{③+④}=\alpha_N$$

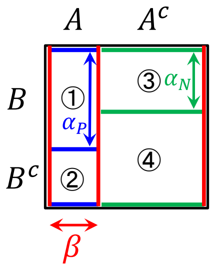


$$P(A|B)=\frac{①}{①+③}$$ を知りたい

ベイズの定理

$$P(A|B)=\frac{\alpha_P\beta}{\alpha_P\beta+\alpha_N(1-\beta)}$$

ベイズの定理に出来ること



ベイズの定理

$$P(A|B)=\frac{\alpha_P\beta}{\alpha_P\beta+\alpha_N(1-\beta)}$$

$$=\frac{\frac{①}{①+②}\frac{①+②}{①+②+③+④}}{\frac{①}{①+②}\frac{①+②}{①+②+③+④}+\frac{③}{③+④}\left(1-\frac{①+②}{①+②+③+④}\right)}$$

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

$$=\frac{①}{①+③}$$

$$P(A|B)=\frac{①}{①+③}$$ を計算できた

【事前の準備】
・事象Aが予め分かっているとする。
・事象Bの発生頻度を**正確に**調べておく。

事象A	事象B	【事後に推定できること】
・本当に感染した	・検査で陽性	・陽性と判別された人が本当に感染している確率
・商品を買った	・広告を見た	・広告を見た人が商品を買う確率
・桶屋が儲かる	・風が吹いた	・風が吹いたときに桶屋が儲かる確率
・翌日は晴れる	・前日が雨で北風	・前日が雨で北風のとき翌日に晴れる確率

本当に感染した確率？

1. 検査したら「陽性」だった！
2. 本当に感染したのか？
3. ベイズの定理の数式
4. 関連する数理統計

詳細

判別 の 評価指標

		真のラベル	
		事象Aが生じる A	事象Aが生じない A^c
判別の結果	事象Bが生じる B	True Positive ①	False Positive ③
	事象Bが生じない B^c	False Negative ②	True Negative ④

$$F値 = \frac{精度 \times 再現率}{精度 + 再現率} \times 2$$

$$正解率 = \frac{① + ④}{① + ② + ③ + ④}$$

$$精　度 = P(A\ |B\ ) = \frac{①}{① + ③}$$

$$再現率 = P(B\ |A\ ) = \frac{①}{① + ②}$$

$$特異度 = P(B^c|A^c) = \frac{④}{③ + ④}$$

$$真陽性率 = \frac{①}{① + ②}$$

$$偽陽性率 = \frac{③}{③ + ④}$$

再掲

条件つき確率とは？

事象の発生回数（データ）

	事象Aが生じる	事象Aが生じない
事象Bが生じる	$N(A \cap B)$	$N(A^c \cap B)$
事象Bが生じない	$N(A \cap B^c)$	$N(A^c \cap B^c)$

事象 A が生じる回数

$$N(A) = N(A \cap B) + N(A \cap B^c)$$

事象 A が生じない（A^cが生じる）回数

$$N(A^c) = N(A^c \cap B) + N(A^c \cap B^c)$$

事象 B が生じる条件の下で
事象 A が生じる
条件付き確率

$$P(A|B) = \frac{N(A \cap B)}{N(A \cap B) + N(A^c \cap B)}$$

事象 B がどうあっても
事象 A が生じる
条件無し確率

$$P(A) = \frac{N(A)}{N(A) + N(A^c)}$$

分母の違いに注意

再掲

確率 の 数理

事象の表現

	事象Aが生じる A	事象Aが生じない A^c
事象Bが生じる B	$A \cap B$	$A^c \cap B$
事象Bが生じない B^c	$A \cap B^c$	$A^c \cap B^c$

余事象の確率

$P(A^c) = 1 - P(A)$
$P(B^c) = 1 - P(B)$

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

条件付確率 $P(A \mid B)$, $P(B \mid A)$
結合確率 $P(A \cap B) = P(B \cap A)$

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

独立（AとBが）

$P(A|B) = P(A)$
$P(B|A) = P(B)$

$\therefore\ P(B)P(A) = P(A \cap B)$

ベイズの定理

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