横浜国立大学集中講義「現代物理学」
第3回　2021年11月22日

OKAYAMA
UNIVERSITY

ベイズ更新

岡山大学 異分野基礎科学研究所

大槻純也

RIIS

例題：迷惑メールフィルタ

OKAYAMA UNIVERSITY

今回も前回に続き、ベイズの定理の使い方を見ていきます。

やりたいこと：
送られてきたメールが迷惑メールかどうかを複数の条件から判定したい

- リンクが含まれている
- 件名に「アカウント」が含まれている
- 件名に「【重要】」と書かれている
- 件名が「〇mazonからの重要なお知らせ」
- 中国語
- …

今回も例題は
「完全独習・ベイズ統計学入門」小島寛之著　ダイヤモンド社より

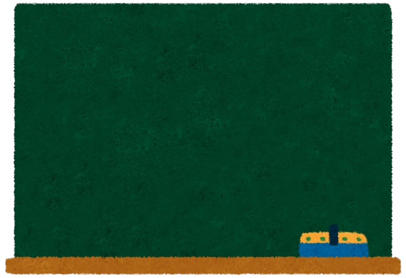
条件1

メールの文面にURLが書かれていたら迷惑メール

		$A=\mathrm{a}$	$\bar{\mathrm{a}}$
		URLありの確率	URLなしの確率
$X=\mathrm{x}$	迷惑メール	0.6	0.4
$\bar{\mathrm{x}}$	非迷惑メール	0.2	0.8

$p(A|X)$ がわかっている

$p(X|A)$ を知りたい



答え

ベイズの定理

$$p(X|A) = \frac{P(A|X)P(X)}{P(A)}$$

X: 迷惑メールかどうか（原因）
A: URLが書かれているか（結果）

X=x, A=aの場合に計算すると

$$p(\mathrm{x}|\mathrm{a}) = \frac{3}{4}$$

メールの文面にURLが書かれていた（*A*=a）時に
迷惑メール（*X*=x）の確率

結果の解釈

$$p(\mathrm{x}) = 50\%$$

条件1

$$p(\mathrm{x}|\mathrm{a}) = 75\%$$

条件1に合致したことにより
確率が上がった

条件2

メールの文面に「出会い」が含まれていたら迷惑メール

$$B = \mathrm{b} \qquad \bar{\mathrm{b}}$$

	「出会い」ありの確率	「出会い」なしの確率
$X = \mathrm{x}$ 迷惑メール	0.4	0.6
$\bar{\mathrm{x}}$ 非迷惑メール	0.05	0.95

$p(B|X)$ がわかっている

$p(X|\underline{A, B})$ を知りたい

条件が2つ付いている場合の
ベイズの定理を使う

条件が2つある場合

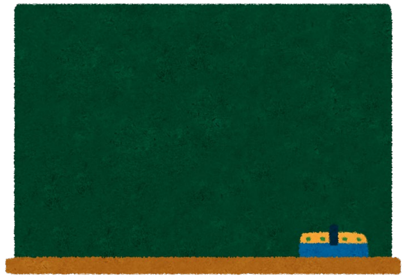
条件が2つある場合のベイズの定理

先に計算済み

$$p(X|A,B) = \frac{P(B|X,A)P(X|A)}{P(B|A)}$$

全てのPに条件Aを付ける
条件Aが確定した後の世界にいると考える

この式を使ってp(X|A, B)を計算



答え

$$p(\mathrm{x}|\mathrm{a}, \mathrm{b}) = \frac{24}{25}$$

メールの文面にURLが書かれていた（*A*=a）
&メールの文面に「出会い」が書かれていた（*B*=b）時に
迷惑メール（*X*=x）の確率

結果の解釈

$$p(\mathrm{x}) = 50\% \xrightarrow{\text{条件1}} p(\mathrm{x}|\mathrm{a}) = 75\% \xrightarrow{\text{条件2}} p(\mathrm{x}|\mathrm{a}, \mathrm{b}) = 96\%$$

$$\downarrow \text{条件3}$$

$$\vdots$$

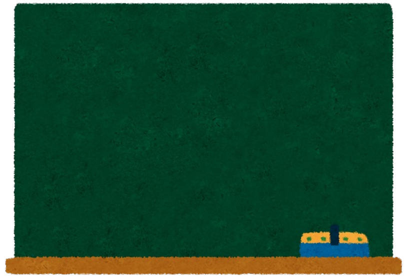
ベイズ更新の特長

- 条件をいくつでも追加できる
- 一旦条件を追加したら、その計算過程は忘れて良い
  （得られた事後確率を事前確率にする）
- 条件を追加する順番は結果に影響しない（次頁で証明）

条件の順番

条件を追加する順番は結果に影響しない

証明：
$p(X|A)$を先に計算して$p(X|A, B)$を計算した場合と
$p(X|B)$を先に計算して$p(X|A, B)$を計算した場合で
結果が変わらないことを示す



ベイズの定理のまとめ

OKAYAMA UNIVERSITY

ベイズの定理

事後確率
posterior probability

事前確率
prior probability

$$p(\boldsymbol{w}|\mathcal{D}) \propto p(\mathcal{D}|\boldsymbol{w})p(\boldsymbol{w})$$

尤度関数
likelihood function

モデルパラメータ： $\boldsymbol{w}$

観測データ： $\mathcal{D}$

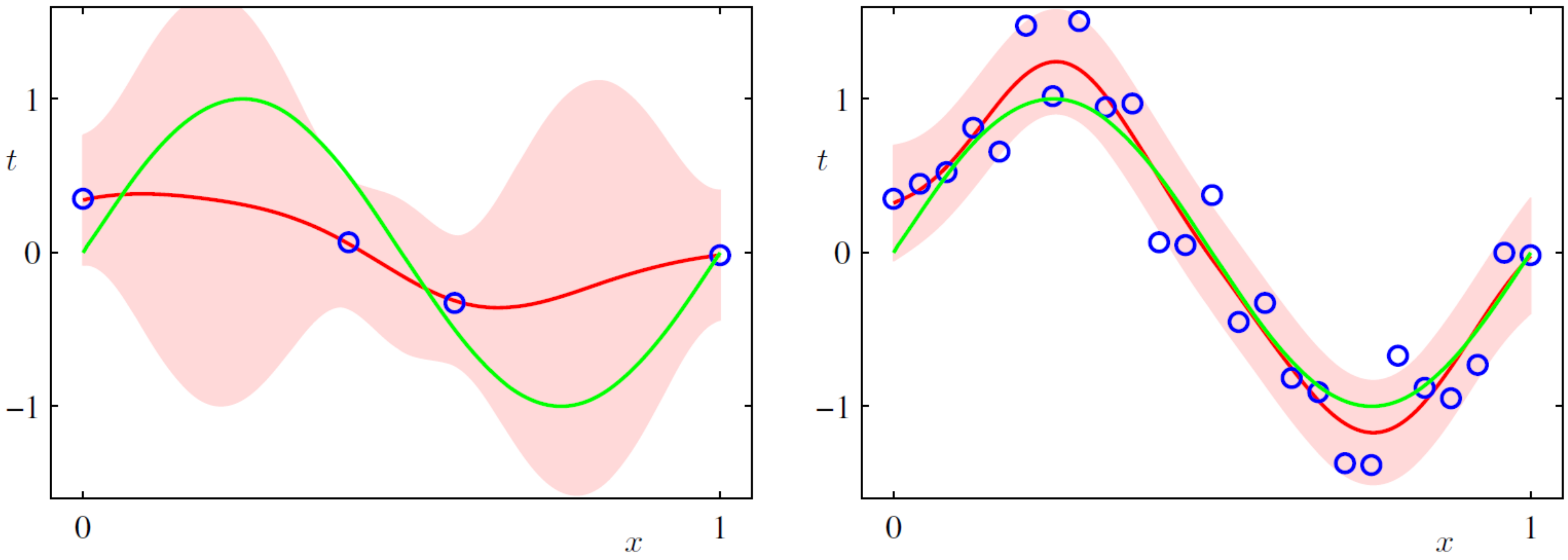
$\boldsymbol{w}$ を知りたい
そのために測定をして $\mathcal{D}$ を得る

$$p(\boldsymbol{w}) \longrightarrow p(\boldsymbol{w}|\mathcal{D})$$

測定により確率が更新される

ここまでの例では、p(D|w)は統計データとして与えられていた
物理で応用する場合は、モデルを仮定してp(D|w)を計算する
(モデル＝関数形、ハミルトニアン)

次回の予告：線形回帰



青丸：測定データ
緑：正解
赤：予測値の平均
ピンク：予測値の分散

PRML Fig.3.8

ベイズの定理を使ってモデル推定＆予測