インテリジェントシステム

#5 不確定性の取り扱い uncertainty

信州大学工学部電子情報システム工学科
丸山稔

不確定性の取り扱い：確率モデル

現実世界の問題：部分的観測可能性や非決定性などなど⇒不確定性の取り扱いが必要
⇒ 確率論を用いる

確率論基礎

・全ての起こりうる”世界”の集合（全ての起こりうることがらの集合）：標本空間（sample space）
⇒ Ωで表す
例：２回のサイコロ投げ→ 6x6 = 36通りの可能な結果(1,1),(1,2), …, (6,6)

・$\omega \in \Omega$ ：標本点…１つの特定の”世界”…主としてここでは離散的な場合を想定

・確率モデル
各$\omega \in \Omega$に対して確率P(ω)を対応付ける

$$\forall \omega : 0 \leq P(\omega) \leq 1, \quad \sum_{\omega \in \Omega} P(\omega) = 1$$

・事象（event）：起こりうる世界（ことがら）の集合
例：２回のサイコロ投げの目の総和が11…Total=11
⇒ logicにおける命題（proposition）に相当…命題が成り立つ世界の集合
任意の命題ϕについて

$$P(\phi) = \sum_{\omega \in \phi} P(\omega)$$

確率論基礎…続き

・確率論の公理

(a) 全ての事象Aについて $0 \leq P(A) \leq 1$

(b) $P(\Omega) = 1$

(c) 互いに排反である可算個の事象 $A_1, A_2, \cdots$に対して

$$P(A_1 \vee A_2 \vee \cdots) = \sum_i P(A_i)$$

・P(Total=11)など⇒条件なしの確率、事前確率（prior probability）

・何らかの情報が与えられている場合←evidence

例：１回目のサイコロの目（Die1）が５のときの２回の目の和の確率など

⇒ 条件付き確率（conditional probability）、事後確率（posterior probability）

P(Total = 11 | Die1 = 5)

・条件付き確率に関しては以下が成立

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

$$\Downarrow$$

$$P(a \wedge b) = P(a|b)P(b)$$ 乗法定理（product rule）と呼ぶ

確率論基礎…さらに続き

- 確率論における変数：確率変数（random variables）... この資料では時々r.v.と略記する
  命題：A = trueなる形式の命題⇒aと略記、A=falseは¬aと略記
  変数を示すためには大文字、値を示すためには小文字を用いる

- 以下が成立する

$$P(\neg a) = 1 - P(a)$$

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$ （inclusion-exclusion）

- 確率論では連言（conjunction）を示すために**,**（comma,カンマ、コンマ）を用いる
  例： $P(cavity|\neg toothache \wedge teen) \Rightarrow P(cavity|\neg toothache, teen)$

- この授業では、あるr.v.に関する全ての値の確率について表したいとき太字（bold）を用いる

P(Weather=sun) = 0.6
P(Weather=rain) = 0.1
P(Weather=cloud) = 0.29
P(Weather=snow) = 0.01

→ **P**(Weather) = <0.6, 0.1, 0.29, 0.01>

確率論基礎…まだまだ続き

・連続値を持つr.v.の場合⇒確率密度関数（pdf, probability distribution function）

$$P(x) = \lim_{dx \to 0} \frac{P(x \leq X \leq x + dx)}{dx}$$

・複数のr.v.に関する確率分布⇒同時確率分布（joint probability distribution）

P(X,Y)

P(X,Y) = **P**(X|Y)**P**(Y)

確率推論

・観測されたevidence
⇒ その条件の下で、クエリ（query, query propositions）の事後確率を求める

・全変数の同時確率：KB（知識ベース）に相当⇒回答を生成

探り針で歯を調べ虫歯を見つける
異常の有無？

例：歯科…全て２値（boolean）：Toothache, Cavity, Catch

全変数の同時確率：2 x 2 x 2 の配列

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

$$P(cavity \vee toothache) = 0.28$$

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

単一または特定のr.v. に関する分布の計算

単一⇒周辺確率（marginal probability）

→ 他の変数に関しては全てのケースについて可算して消去

周辺化（marginalization, summing out）

e.x. P(cavity) … Toothache, Catchに関しては全てのケースについて可算

P(cavity) = 0.2

周辺化：
$$\boldsymbol{P}(\boldsymbol{Y}) = \sum_{\boldsymbol{z}} \boldsymbol{P}(\boldsymbol{Y}, \boldsymbol{Z} = \boldsymbol{z}) \quad \Longrightarrow \quad \boldsymbol{P}(Cavity) = < 0.2,\ 0.8 >$$

product ruleを用いると：

$$\boldsymbol{P}(\boldsymbol{Y}) = \sum_{\boldsymbol{z}} \boldsymbol{P}(\boldsymbol{Y}|\boldsymbol{z})\boldsymbol{P}(\boldsymbol{z})$$

確率推論：

何を行う？⇒多くの場合、evidenceが与えられた条件下での他の変数の条件付き確率の計算

$$P(cavity|toothache) = \frac{P(cavity \land toothache)}{P(toothache)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$
$$\propto P(cavity \land toothache)$$

$$\begin{aligned}\boldsymbol{P}(Cavity|toothache) &= \alpha \boldsymbol{P}(Cavity, toothache) = \alpha \sum_{Catch} \boldsymbol{P}(Cavity, toothache, Catch) \\ &= \alpha(\boldsymbol{P}(Cavity, toothache, catch) + \boldsymbol{P}(Cavity, toothache, \neg catch)) \\ &= \alpha < 0.12, 0.08 > = < 0.6, 0.4 >\end{aligned}$$

単一のクエリ X の場合の一般形

E : 観測値が与えられた確率変数のリスト（evidence variables）

e : 観測された値のリスト

Y : 未観測の確率変数⇒全ての可能な値の組合せについて可算（して消去）

$$\boldsymbol{P}(X|\boldsymbol{e}) = \alpha \boldsymbol{P}(X, \boldsymbol{e}) = \alpha \sum_{\boldsymbol{y}} \boldsymbol{P}(X, \boldsymbol{e}, \boldsymbol{y})$$

同時確率の表から抽出可

独立性

例題で考えてみる…Toothache, Catch, CavityにWeatherを追加⇒表のサイズは2x2x2x4=32

$$P(toothache, catch, cavity, cloud) = P(cloud|toothache, catch, cavity)P(toothache, catch, cavity)$$

天候が歯の状態に左右されることはないから：

$$P(cloud|toothache, catch, cavity) = P(cloud)$$

$$\Downarrow$$

$$P(toothache, catch, cavity, cloud) = P(cloud)P(toothache, catch, cavity)$$

r.v. X,Yが独立であるとは：

$$\boldsymbol{P}(X|Y) = \boldsymbol{P}(X),\ or\ \boldsymbol{P}(Y|X) = \boldsymbol{P}(Y),\ or\ \boldsymbol{P}(X, Y) = \boldsymbol{P}(X)\boldsymbol{P}(Y)$$

独立性を利用⇒独立な部分集合の確率の積に分解⇒必要な表のサイズ縮小

Bayesの定理とその利用：

product ruleより $P(a \wedge b) = P(a, b) = P(a|b)P(b) = P(b|a)P(a)$

$$P(b|a) = \frac{P(a|b)P(b)}{P(a)}$$

多値r.v.の場合は： $\boldsymbol{P}(Y|X) = \dfrac{\boldsymbol{P}(X|Y)\boldsymbol{P}(Y)}{\boldsymbol{P}(X)}$

…さらにbackground evidence **e** を加えると： $\boldsymbol{P}(Y|X, \boldsymbol{e}) = \dfrac{\boldsymbol{P}(X|Y, \boldsymbol{e})\boldsymbol{P}(Y|\boldsymbol{e})}{\boldsymbol{P}(X|\boldsymbol{e})}$

Bayesの定理を用いると

$$P(cause \mid effect) = \frac{P(effect \mid cause)P(cause)}{P(effect)}$$

$\boldsymbol{P}(Y|X) = \alpha \boldsymbol{P}(X|Y)\boldsymbol{P}(Y)$ ⇒ 正規化により $\boldsymbol{P}(Y|X)$ を導出できる

例：

$$\boldsymbol{P}(Cavity|toothache \wedge catch) = \alpha \boldsymbol{P}(toothache \wedge catch|Cavity)\boldsymbol{P}(Cavity)$$

Bayesの定理とその利用：

条件付き確率の計算：観測できるevidenceが多数（n）の場合⇒2^n 通りの組み合わせ

条件付き独立性を利用

$$\boldsymbol{P}(X, Y|Z) = \boldsymbol{P}(X|Z)\boldsymbol{P}(Y|Z)$$

または　$\boldsymbol{P}(X|Y,Z) = \boldsymbol{P}(X|Z),\ \boldsymbol{P}(Y|X,Z) = \boldsymbol{P}(Y|Z)$

$$\boldsymbol{P}(toothache \wedge catch|Cavity) = \boldsymbol{P}(toothache|Cavity)\boldsymbol{P}(catch|Cavity)$$

$$\boldsymbol{P}(Cavity|toothache \wedge catch) = \alpha\boldsymbol{P}(toothache|Cavity)\boldsymbol{P}(catch|Cavity)\boldsymbol{P}(Cavity)$$

表のサイズを縮小可：7 ⇒ 5

Bayesの定理（例題）

ある病気に罹っているか検査する方式についての確率推論を考えてみる

・この検査方式は陽性（+:positive）か陰性（-:negative）か出力するものとする
　p(+|D) ：病気（D）に罹っている人が検査を受けて陽性となる確率
　p(- |D) ：病気（D）に罹っている人が検査を受けて陰性となる確率

$$p(-|D) = 1 - p(+|D)$$

・p(+|H) : 健康な人（病気Dに罹っていない人）が検査を受けて陽性となる確率

・p(D) :検査を受ける人（検査対象者）が病気（D）に罹っている確率
・p(H) :検査を受ける人（検査対象者）が健康である（病気Dに罹っていない確率）

$$p(H) = 1 - p(D)$$

このとき検査で陽性となっている人が実際に病気Dに罹っている確率は？
⇒ p(D|+) … これはBayesの定理を用いて計算してみる

…. 続き

検査で陽性となっている人が実際に病気Dに罹っている確率は？
⇒ p(D|+) … これはBayesの定理より

$$p(D|+) = \frac{p(+|D)p(D)}{p(+)} = \frac{p(+|D)p(D)}{p(+|D)p(D) + p(+|H)p(H)}$$

この検査方式は極めて優秀で p(+|D) = 0.98, p(+|H) = 0.01 であったとする。
p(D) = 0.001のときは？（i.e. p(H) = 0.999）

$$p(D|+) = \frac{0.98 * 0.001}{0.98 * 0.001 + 0.01 * 0.999} \approx 0.0893$$

なんらかの手段で検査対象者を絞り、その対象者については
p(D) = 0.1とできたときは？検査方式の能力は変わらないけれども…

$$p(D|+) = \frac{0.98 * 0.1}{0.98 * 0.1 + 0.01 * 0.9} \approx 0.916$$

Naive Bayesモデル：

ある原因（cause）により、数多くの効果（effects）が発生→観測可
観測データから原因を推定したい
⇒ 全変数による同時確率は、各効果変数の条件付き独立性を仮定すると…

$$\boldsymbol{P}(\mathrm{Cause}, \mathrm{Effect}_1, \cdots, \mathrm{Effect}_n) = \boldsymbol{P}(Cause)\Pi_{i=1}^{n}\boldsymbol{P}(\mathrm{Effect}_i|\mathrm{Cause})$$

単純ベイズモデル（Naive Bayes）

効果r.v.のうち、観測される集合**E**=**e**, その他**Y**（観測されないr.v.）とすると
条件付き独立性の下で…

$$\boldsymbol{P}(\mathrm{Cause}|\boldsymbol{e}) = \alpha\sum_{\boldsymbol{y}}\boldsymbol{P}(\mathrm{Cause}, \boldsymbol{e}, \boldsymbol{y})$$

$$\begin{aligned}\boldsymbol{P}(\mathrm{Cause}|\boldsymbol{e}) &= \alpha\sum_{\boldsymbol{y}}\boldsymbol{P}(\mathrm{Cause}, \boldsymbol{e}, \boldsymbol{y}) = \alpha\sum_{\boldsymbol{y}}\boldsymbol{P}(\mathrm{Cause})\boldsymbol{P}(\boldsymbol{y}|\mathrm{Cause})(\Pi_i\boldsymbol{P}(\boldsymbol{e}_j|\mathrm{Cause}))\\ &= \alpha\boldsymbol{P}(\mathrm{Cause})\Pi_i\boldsymbol{P}(\boldsymbol{e}_j|\mathrm{Cause})(\sum_{\boldsymbol{y}}\boldsymbol{P}(\boldsymbol{y}|\mathrm{Cause}))\\ &= \alpha\boldsymbol{P}(\mathrm{Cause})\Pi_i\boldsymbol{P}(\boldsymbol{e}_j|\mathrm{Cause})\end{aligned}$$