知識工学　第13回

二宮　崇

教科書と資料

- 教科書
  - Artificial Intelligence: A Modern Approach (3rd Edition): Stuart Russell, Peter Norvig (著), Prentice Hall, 2009

- この講義のウェブサイト

http://aiweb.cs.ehime-u.ac.jp/~ninomiya/ke/

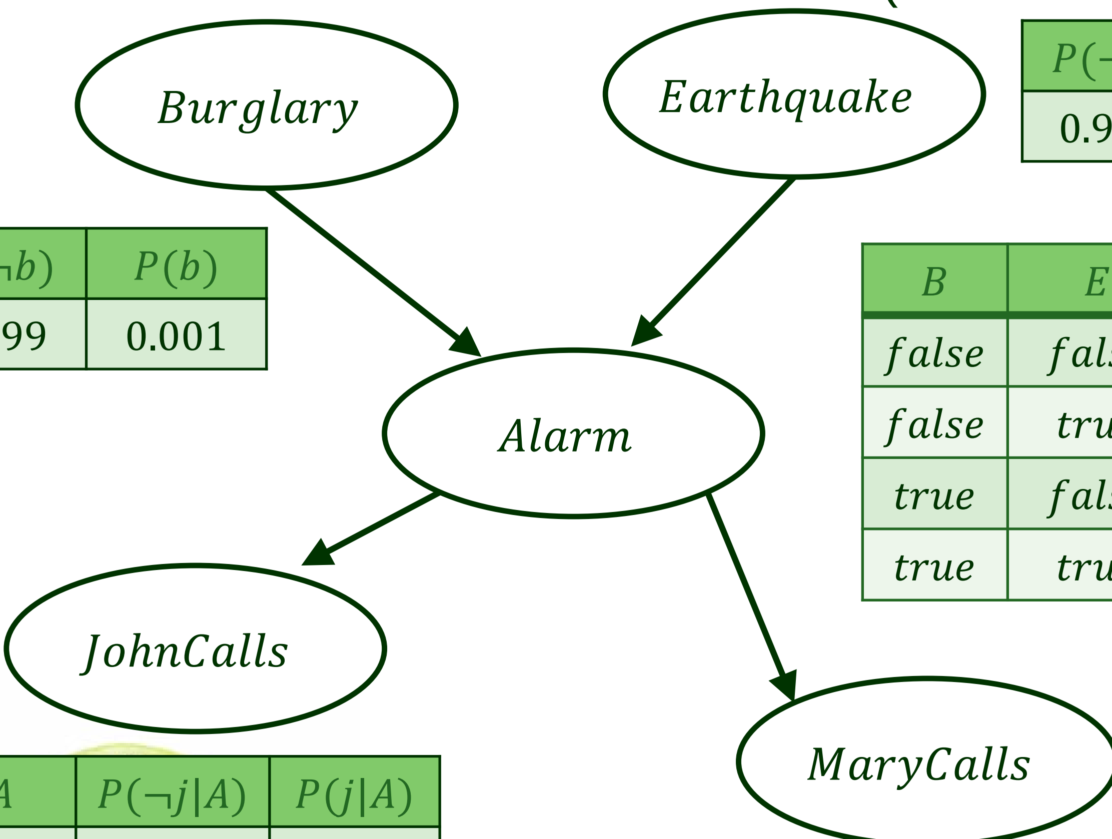
本日の講義内容

- ベイジアンネットの効率的な表現(§14.3)
  - 基準的な分布
  - 決定的なノード
  - Noisy-ORモデル
  - ハイブリッドベイジアンネット
  - 離散変数と連続変数に対する確率分布
- ベイジアンネットの厳密推論 (§14.4)
  - 列挙による推論 (§14.4.1)

ベイジアンネットの復習

- 防犯アラームの例

条件付き確率表
(Conditional Probability Table, CPT)



$P(\neg b)$	$P(b)$
0.999	0.001

$P(\neg e)$	$P(e)$
0.998	0.002

B	E	$P(\neg a \mid B,E)$	$P(a \mid B,E)$
$false$	$false$	0.999	0.001
$false$	$true$	0.71	0.29
$true$	$false$	0.06	0.94
$true$	$true$	0.05	0.95

A	$P(\neg j \mid A)$	$P(j \mid A)$
$false$	0.95	0.05
$true$	0.10	0.90

A	$P(\neg m \mid A)$	$P(m \mid A)$
$false$	0.99	0.01
$true$	0.30	0.70

ベイジアンネットの復習

- 完全結合分布の表現 (§14.2.1)
  - $P(X_1 = x_1 \land \cdots \land X_n = x_n)$を$P(x_1, \cdots, x_n)$と略記する
  - $P(x_1, \cdots, x_n) = \prod_{i=1}^{n} P(x_i | parents(x_i))$
  - ただし、$parents(x_i)$は$Parents(X_i)$に含まれる変数の具体的な値の組

例: アラームが鳴り(a)、しかし、泥棒(b)も入らず、地震(e)も起きずに、ジョン(j)とメアリー(m)が電話をする確率

$$P(j \land m \land a \land \neg b \land \neg e)$$
$$= P(j|a)P(m|a)P(a|\neg b \land \neg e)P(\neg b)P(\neg e)$$
$$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.000628$$

条件付き確率分布の効率的な表現 (§14.3)

- ベイジアンネットの問題点
  - 親の数をkとしたとき、CPTのサイズは $O(2^k)$ となってしまう

- 基準的な分布(canonical distribution)
  - 分布を決定するパターンといくつかのパラメータを与えることでCPTを再現
    - 決定的なノード
    - Noisy-ORモデル
    - ハイブリッドベイジアンネット

条件付き確率分布の効率的な表現 (§14.3)

- 決定的なノード
  - 親の値によって、子の値が不確実性なく論理的に与えられる。

- 例1

親ノード: $Canadian, US, Mexican$
子ノード: $NorthAmerican$

- 例2

親ノード: 複数の車のディーラーの価格
子ノード: ユーザーの購入価格(最低価格を常に選択)

- 例3

親ノード: 湖への流入水量, 流出量
子ノード: 湖の水位

条件付き確率分布の効率的な表現 (§14.3)

- Noisy-ORモデル
  - 不確実な論理和のモデル。ある親ノードが真であるとき子ノードを偽とする抑制力が独立に与えられる確率モデル。

条件付き確率分布の効率的な表現 (§14.3)

- Noisy-ORモデルの仮定
  - 仮定1) 原因が親によって全て列挙されている
  - 仮定2) ある親ノードが真であるとき子ノードを偽とする抑制力は他の親ノードに対し独立に与えられる
    - 各親ノードに対する中間ノードを定義し、中間ノードの論理和により子ノードの真偽を決定的に与えるモデルと等価

$$
\begin{array}{ccccc}
X_1 & X_2 & X_3 & \cdots & X_n \\
\downarrow & \downarrow & \downarrow & & \downarrow \\
Z_1 & Z_2 & Z_3 & \cdots & Z_n
\end{array}
\rightarrow Y
$$
(論理和による決定的ノード)

条件付き確率分布の効率的な表現 (§14.3)

- Noisy-ORモデルの例

親ノード: $Cold, Flu, Malaria$, 子ノード: $Fever$

$q_{malaria} = P(\neg fever | \neg cold, \neg flu, malaria) = 0.1$

$q_{flu} = P(\neg fever | \neg cold, flu, \neg malaria) = 0.2$

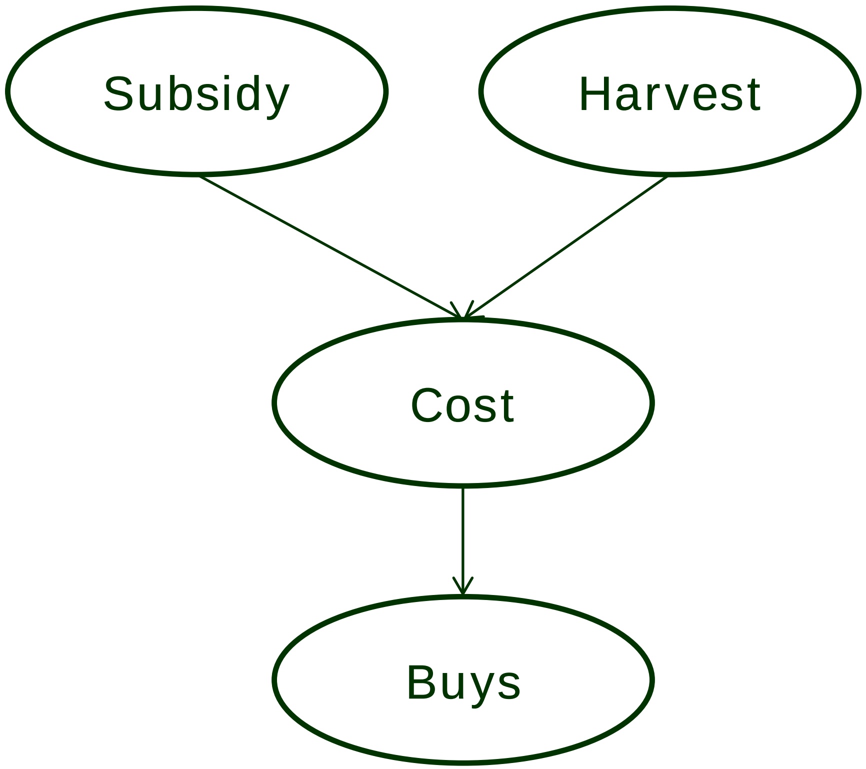
$q_{cold} = P(\neg fever | cold, \neg flu, \neg malaria) = 0.6$

$$P(\neg y | X_1, \cdots, X_n) = \prod_{\{j : X_j = true\}} q_j$$

$Cold$	Flu	$Malaria$	$P(\neg fever)$	$P(fever)$
0	0	0	1.0	0.0
0	0	1	$\mathbf{0.1}$	0.9
0	1	0	$\mathbf{0.2}$	0.8
0	1	1	$0.02 = 0.2 \times 0.1$	0.98
1	0	0	$\mathbf{0.6}$	0.4
1	0	1	$0.06 = 0.6 \times 0.1$	0.94
1	1	0	$0.12 = 0.6 \times 0.2$	0.88
1	1	1	$0.012 = 0.6 \times 0.2 \times 0.1$	0.988

条件付き確率分布の効率的な表現 (§14.3)

- ハイブリッドベイジアンネット
  - 離散変数と連続変数の両方を持つベイジアンネット
  - 論理確率変数
    - $Subsidy$(助成金)
    - $Buys$(購入)
  - 連続確率変数
    - $Harvest$ (収穫量)
    - $Cost$(価格)



条件付き確率分布の効率的な表現 (§14.3)

- 離散変数と連続変数に対する連続変数の確率分布
  - 線形ガウス分布による収穫量hに対する価格cの確率分布

$$P(c|h, subsidy) = N(a_t h + b_t, \sigma_t^2)(c) = \frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{c-(a_t h+b_t)}{\sigma_t}\right)^2}$$

$$P(c|h, \neg subsidy) = N\left(a_f h + b_f, \sigma_f^2\right)(c) = \frac{1}{\sigma_f\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{c-(a_f h+b_f)}{\sigma_f}\right)^2}$$

条件付き確率分布の効率的な表現 (§14.3)

- 連続変数に対する離散変数の確率分布
  - ある一定以上の値なら1に、以下なら0に近い確率分布
  - 正規分布の累積分布関数 (プロビット分布)
  - ロジスティック関数(ロジット分布)

条件付き確率分布の効率的な表現 (§14.3)

- 連続変数に対する離散変数の確率分布
  - 正規分布の累積分布関数 (プロビット分布)

$$\Phi(x) = \int_{-\infty}^{x} N(0,1)(x)dx$$

$$P(buys|Cost = c) = \Phi\left(\frac{-c+\mu}{\sigma}\right)$$

条件付き確率分布の効率的な表現 (§14.3)

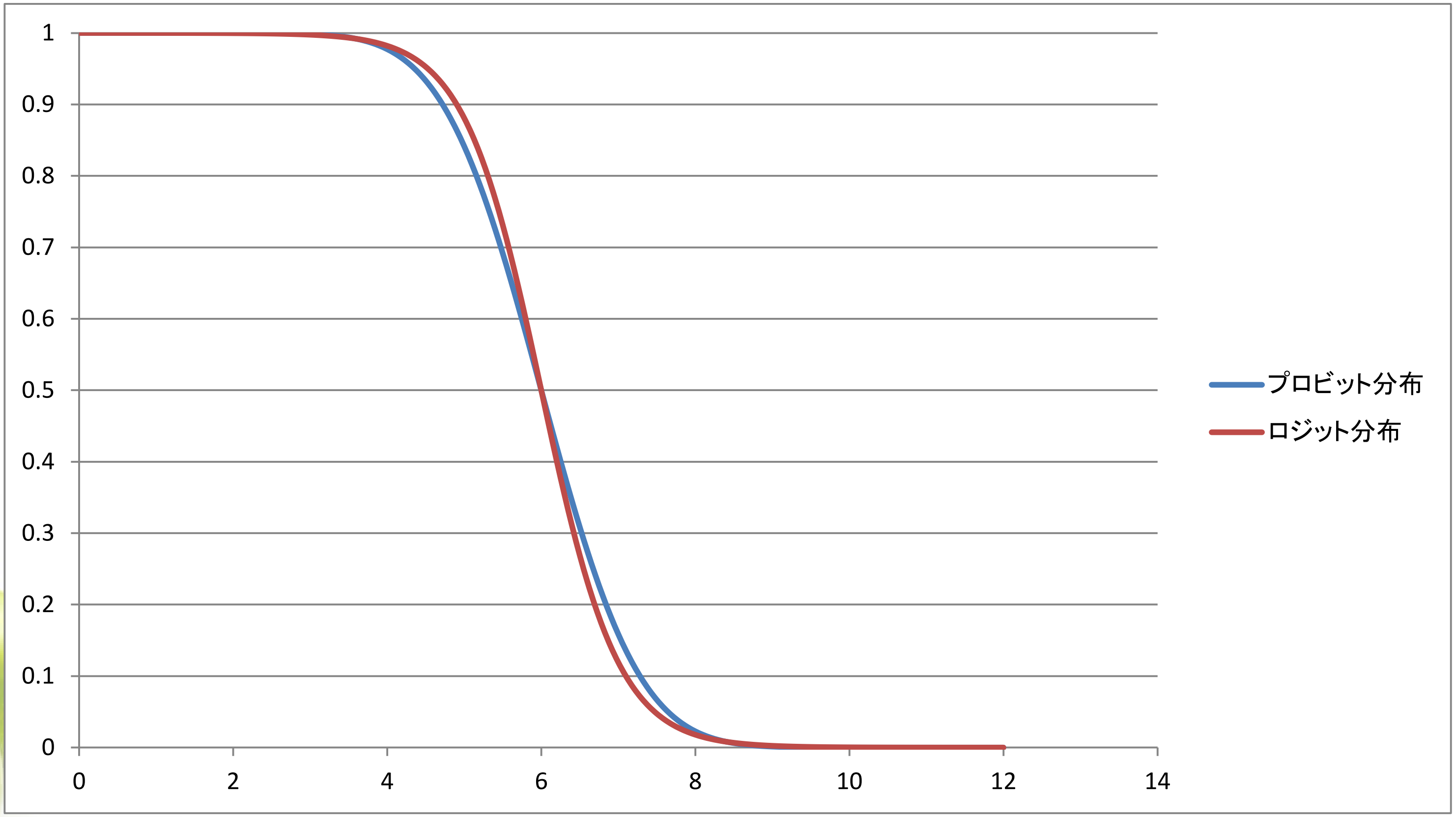
- 連続変数に対する離散変数の確率分布
  - ロジスティック関数(ロジット分布)

$$\Phi(x) = \frac{1}{1+e^{-x}}$$

$$P(buys|Cost = c) = \Phi\left(2\frac{-c+\mu}{\sigma}\right)$$

条件付き確率分布の効率的な表現 (§14.3)

- 連続変数に対する離散変数の確率分布
  - プロビット分布とロジット分布の違い



ベイジアンネットの厳密推論(§14.4)

- 推論
  - 特定の観測事象$e_1, \cdots, e_n$が与えられたもとで、質問変数Xの事後確率分布$P(X|e_1, \cdots, e_n)$を計算すること。
  - 観測事象$e_1, \cdots, e_n$は証拠変数の集合$E_1, \cdots, E_n$に値を割り当てたもの。
  - 全ての変数集合: $\{X\} \cup \{E_1, \cdots, E_n\} \cup \{Y_1, \cdots, Y_m\}$
  - $Y_1, \cdots, Y_m$: 隠れ変数(非証拠、非質問の変数集合)

例

$$P(Burglary|JohnCalls = true, MaryCalls = true)$$

$P(\neg b|j, m)$	$P(b|j, m)$
0.716	0.284

ベイジアンネットの厳密推論(§14.4)

- 列挙による推論 (§14.4.1)

$$P(X|e_1,\cdots,e_n) = \alpha P(X,e_1,\cdots,e_n) = \alpha \sum_{Y_1,\cdots,Y_m} P(X,e_1,\cdots,e_n,Y_1,\cdots,Y_m)$$

  - ベイジアンネットを使って$P(X,e_1,\cdots,e_n,Y_1,\cdots,Y_m)$を求めて$P(X|e_1,\cdots,e_n)$を計を計算すれば良い。

- 例

$$P(Burglary|JohnCalls = true, MaryCalls = true)$$

$$P(B|j,m) = \alpha P(B,j,m) = \alpha \sum_E \sum_A P(B,j,m,E,A)$$

- 時間計算量 $O(n2^n)$

ベイジアンネットの厳密推論(§14.4)

- 列挙による推論への工夫

$P(b|j,m) = \alpha \sum_E \sum_A P(b)P(E)P(A|b,E)P(j|A)P(m|A)$

$$= \alpha P(b) \sum_{E} P(E) \sum_{A} P(A|b,E)P(j|A)P(m|A)$$

$$= \alpha P(b)\{P(e)\big(P(a|b,e)P(j|a)P(m|a) + P(\neg a|b,e)P(j|\neg a)P(m|\neg a)\big)$$

ベイジアンネットの厳密推論(§14.4)

- 列挙による推論への工夫

同様に$P(\neg b|j,m)$を求める

$P(\neg b|j,m) = \alpha 0.00149191$

よって、$P(B|j,m)$は次のようになる。

$P(\neg b\|j,m)$	$P(b\|j,m)$
0.716	0.284

時間計算量: $O(2^n)$