13/14 Probability and Inference

Recap:

概念:
- sample space
- probability mode
- events
- random variable, domains
- (probability) statements
- probability distribution
- joint distribution ← on multi variables
- full joint probability distribution → all combinatio

New: probability inference: BG knowledge (prior) + Evidence (observed)

- Prior (unconditional) prob
- posterior (conditional) prob

$$p(a|b) = \frac{p(a \wedge b)}{p(b)} \text{ when } p(b) > 0$$

变形: $p(a \wedge b) = p(a|b) \cdot \underbrace{p(b)}_{\text{可以 normalize 成 } \alpha}$

Marginalization: (边缘化)

$$p(x) = \sum_{y \in Y} p(x|y)$$

$$\text{或} = \sum_{y_1 \in Y_1} \sum_{y_2 \in Y_2} p(x|y_1, y_2)$$

probability inference (single variable):

$$P(X|e) = \alpha p(X, e) = \alpha \sum_{y \in Y} p(X, e, y)$$

- x: asked
- e: observed
- y: hidden

解释:

for each e :
    for each x_i :
        for each combinations of y :

sum up $p(x_i, e, y)$

Result:

$$p(X|e) = \langle p(x_i, e, y) \rangle$$

复杂度: time: $O(m^n)$ space: $O(m^n)$

n = # syptoms = # evidence variables

m = # domain of query variables, if binary 2^n

Bayesian Reasoning

product rule: $p(a|b)\,p(b) = p(b|a)\,p(a)$

$$p(b|a) = \frac{p(a|b)\,p(b)}{p(a)} = \alpha\, p(a|b)\,p(b)$$

causal inference: $p(\text{effect}|\text{cause})$

diagnosis inference: $p(\text{cause}|\text{effect})$ (记这个)

Combining evidence

$\downarrow$

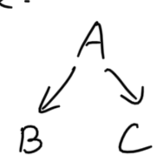
n evidence variables $\Rightarrow$ $\boxed{2^n}$ combinations

Independence
$$\begin{cases} p(a|b) = p(a) \\ p(a \wedge b) = p(a)\,p(b) \end{cases}$$

* can compute $n \times m$ prob from $n+m$ prob

Conditional Independence:

(a, b CI)



given A or ¬A: B and C are CI

$$\begin{cases} p(B|A,C) = p(B|A) \\ p(B,C|A) = p(B|A) \cdot p(C|A) \end{cases}$$

Complexity: $\boxed{O(n)}$

more likely to be assumption than absolute independence

Note:

query variable: X : $\text{Domain}(X) = \{x_1, x_2, \dots x_m\}$
evidence variable: E : $\{E_1, \dots, E_k\}$
observations : e $\{e_1, \dots e_i\}$ s.t. $E_i = e_i$
unobserved variables: Y : $\{Y_1, \dots Y_l\}$
$\text{Domain}(Y_i) = \{y_i, \dots y_{i,n_i}\}$.

△ Independence assumptions reduce number of probs required to represent a full joint distribution

(by factoring joint distribution)

作业题:

13.1 证 $p(a \mid b \wedge a) = 1$

$$p(a \mid b \wedge a) = \frac{p(a \wedge (b \wedge a))}{p(b \wedge a)} = \frac{p(a, b)}{p(a, b)} = 1$$

13.3 prove it or give counterexample

a. if $P(a \mid b, c) = P(b \mid a, c)$, then $P(a \mid c) = P(b \mid c)$

$$\frac{p(a, b, c)}{p(b, c)} = \frac{p(a, b, c)}{p(a, c)} \qquad \frac{p(a, c)}{p(c)} = \frac{p(b, c)}{p(c)}$$

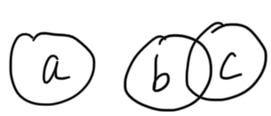
$$p(b, c) = p(a, c) \longrightarrow p(a, c) = p(b, c)$$

b. if $P(a \mid b, c) = p(a)$, then $P(b \mid c) = p(b)$

a 和 $b \wedge c$ 独立 $\qquad$ b, c 独立

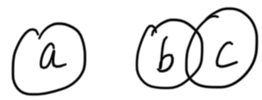
$p(b, c) = p(b) \cdot p(c)$

反例: b, c 不独立



c. 证 $\underbrace{P(a|b) = P(a)}_{a,b\text{独立}}$, then $\underbrace{P(a|b,c) = P(a|c)}_{b,c\text{独立} + a,b\text{独立}}$.

反例: b, c 不独立.



13.10 (a)

BAR / BAR / BAR	20	$(\frac{1}{4})^3 = \frac{1}{64}$
BELL / BELL / BELL	15	$(\frac{1}{4})^3$
L / L / L	5	$(\frac{1}{4})^3$
C / C / C	3	$(\frac{1}{4})^3$
C / C / ?	2	$(\frac{1}{4})^2 - (\frac{1}{4})^3 = \frac{3}{64}$
C / ? / ?	1	$\frac{1}{4} - (\frac{1}{4})^3 - \frac{3}{64} = \frac{12}{64}$

$$E(p) = (20+15+5+3) \times \frac{1}{64} + 2 \times \frac{3}{64} + 1 \times \frac{12}{64} = \frac{61}{64}$$

(b) $p(win) = 4 \times \frac{1}{64} + \frac{3}{64} + \frac{12}{64} = \frac{19}{64}$