FUDAN SUMMER SCHOOL INTRODUCTION TO AI

Probabilistic Reasoning and Decision Making

Day 2 – Examples of Probabilistic Reasoning

Announcements

- Work on homework 1 – Due next Monday

① belief as probability

② rules of probability

$$\begin{cases} \sum P(x_i) = 1 & \text{where } x_i \text{ are outcomes of } X \\ 0 \le P(X = x_i) \le 1 & \\ P(X = x_i \text{ or } X = x_j) = P(x_i) + P(x_j) & \end{cases}$$

x_i & x_j are mutually exclusive

③ $P(X = x_i \mid Y = y_j) = P(X = x_i) \Rightarrow X$ & Y are independent

$P(X = x_i \mid Y = y_j, Z = z_k) = P(X = x_i \mid \underline{Z = z_k})$

X & Y are conditionally indep given Z

Today: Probabilistic Inference
Do probabilities capture common sense reasoning?

1. Review of probability
2. Multiple explanations of a single event
3. Multiple events with a common explanation
4. Chain of linked events

More independence

$P(B=b_i \mid E=e_j) \overset{??}{=} P(B=b_i)$

Consider two events:

B = A burglar breaks into your apartment $\begin{cases} 1 : \text{happened} \\ 0 : \text{not} \end{cases}$

E = An earthquake occurs $\begin{cases} 1 : \text{happened} \\ 0 \end{cases}$

Are these events independent or dependent? (i.e. does knowing that one happened change your belief in the other?)

A. They are independent because knowing that one happened does not change your belief that the other happened.

B. They are dependent, because knowing that one happened changes your belief that the other happened.

Conditional dependence

B & E are marginally independant

$$P(B=1) = P(B=1|E=1) = P(B=1|E=0)$$

Now consider a third event:

A = Your alarm goes off

1: alarm went on

0: alarm didn't go on.

Which of the following relationships best models beliefs about the world?

A. $P(B=1|A=1) = P(B=1|A=1, E=1)$

B. $P(B=1|A=1) > P(B=1|A=1, E=1)$ ← explain away

C. $P(B=1|A=1) < P(B=1|A=1, E=1)$

B & E are conditionally dependant given A=1

Axioms of Conditional probability

marginal

$$\begin{cases} 0 \le P(x_i) \le 1 \\ \sum P(x_i) = 1 \end{cases}$$

$P(x_i \text{ or } x_j) = P(x_i) + P(x_j)$

- Which of the following axioms hold for conditional probabilities?

✓ A. $\cancel{\le 1}$ $P(X = x_i | Y = y_j) \geq 0$

$P(B=1 | \underline{A=1}) + P(B=0 | \underline{A=1}) = 1$

✓ B. $\sum_i P(X = x_i | Y = y_j) = 1$

✗ C. $\sum_j P(X = x_i | Y = y_j) = 1 \longrightarrow P(B=1 | A=1) + P(B=1 | A=0) \neq 1$

D. A and B only

E. A, B and C

conditional

$$\begin{cases} 0 \le P(x_i | y_j) \le 1 \\ \sum_i P(x_i | y_j) = 1 \\ P(x_i \text{ or } x_j | y_k) = P(x_i | y_k) + P(x_j | y_k) \end{cases}$$

$\ast P(B=1)$?

Joint Probability

P(T=hot) = P(T=hot, W=sunny) +
P(T=hot, W=cloudy)
= .5 + .05 = .55

$P(X = x_i, Y = y_j)$

"What is my belief that $X = x_i$ **and** that $Y = y_j$"

$T \in \{t_1 = hot, t_2 = cool\}$

W $\in \{w_1 = sunny, w_2 = cloudy\}$

P(T=hot, W=sunny)

P(T=hot)

A: .55

B: .8

C: Can't determine

D: None of above

Joint Probability Table

	T = hot	T = cool
W = sunny	0.5	0.3
W = cloudy	0.05	0.15

P(T=cool, W=cloudy)

Marginalization

Marginalize Y out (simplify)

$$P(X = x_i) = \sum_{j} P(\overbrace{X = x_i, Y = y_j}^{\text{joint}})$$

marginalize Y in (calculation)

$T \in \{t_1 = hot, t_2 = cool\}$

W $\in \{w_1 = sunny, w_2 = cloudy\}$

	T = hot	T = cool
W = sunny	0.5	0.3
W = cloudy	0.05	0.15

$$P(W = \text{sunny}) \xrightarrow{\text{marginalize}} P(W = \text{sunny}, T = \text{hot}) + P(W = \text{sunny}, T = \text{cool})$$

$= .5 + .3$

$= .8$

Review

Which of the following uses the correct notation?

A. Joint Probability: P(X,Y)
Conditional Probability: P(X|Y)
Unconditional Probability: P(X)

marginal probability
prior

B. Joint Probability: P(X|Y)
Conditional Probability: P(X,Y)
Unconditional Probability: P(X)

C. Joint Probability: P(X|Y)
Conditional Probability: P(X)
Unconditional Probability: P(X|Y)

D. Joint Probability: P(X,Y)
Conditional Probability: P(X)
Unconditional Probability: P(X|Y)

Product Rule:
relates joint and conditional probabilities

$P(T=hot \mid W=sunny) = \frac{P(hot, sunny)}{P(sunny)} = \frac{.5}{.8}$

A: $\frac{.5}{.55}$ B: $\frac{.5}{.8}$ C: $\frac{.5}{.85}$ D: None

conditional   joint

$$for\ all\ i,j\ P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_i)}$$

marginal

$P(X=x_i, Y=y_j) \Longleftrightarrow = P(X=x_i \mid Y=y_j)\, P(Y=y_j)$

$T \in \{t_1 = hot, t_2 = cool\}$

W $\in \{w_1 = sunny, w_2 = cloudy\}$

	T = hot	T = cool
W = sunny	0.5	0.3
W = cloudy	0.05	0.15

$P(T=hot \mid W=cloudy) = \frac{P(T=hot, W=cloudy)}{P(W=cloudy)} = \frac{.05}{.2} = \frac{1}{4}$

W = sunny	0.8
W = cloudy	0.2

Product Rule Generalization

$P(x, y) = P(x|y)P(y)$

$= P(y|x)P(x)$

$$P(A = a_i, B = b_j, C = c_k, D = d_l, \dots)$$
$$= P(A = a_i)P(B = b_j | A = a_i)P(C = c_k | A = a_i, B = b_j) \dots$$

$P(\underbrace{A=a_i, B=b_j}_{X}, \underbrace{C=C_k}_{Y}) = P(A=a_i, B=b_j | C=C_k) P(C=C_k)$

This decomposition often makes reasoning easier.

$= P(A=a_i | B=b_j, C=C_k) P(B=b_j | C=C_k)$

$P(C=C_k)$

$P(x, y, z) = P(z) P(y|z) P(x|y\ z)$

$= P(y) P(x|y) P(z|x, y)$

$= P(x) P(y|x) P(z|x, y)$

Shorthand

Implied Universality:

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

$\forall\ x_i, y_j$

$P(X=x_i, Y=y_j) = P(X=x_i | Y=y_j)P(Y=y_j)$

Implied Assignment:

random var

values

$$P(x,y,z) = P(X = x, Y = y, Z = z)$$

Bayes' Rule

$P(A=1 \mid E=1)$ easier

$\rightarrow P(E=1 \mid A=1)$ harder

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$ ✓

easier to get

prior

Take a moment to derive Bayes' Rule from the Product Rule.

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

normalization factor

goal: calculate a conditional prob.

Very hard when Y is the condition.

Bayes' Rule example

$P(T=1) \underset{\text{product}}{\overset{\text{Marginalize}}{=}} P(T=1, C=0) + P(T=1, C=1)$

$= P(T=1|C=0)\,P(C=0) + P(T=1|C=1)\,P(C=1)$

- Let's assume breast cancer affects 1% of all patients
- When a patient has cancer, a mammogram test comes back positive 80% of the time (true positive)
- When a patient does not have cancer, a mammogram test comes back positive 9% of the time (false positive)
- A patient has a positive mammogram test. What is the probability the patient has cancer?

- Let's assume breast cancer affects 1% of all patients
- When a patient has cancer, a mammogram test comes back positive 80% of the time (true positive)
- When a patient does not have cancer, a mammogram test comes back positive 9% of the time (false positive)
- A patient has a positive mammogram test. What is the probability the patient has cancer?

$C : \{ \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}, \quad T = \{ \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}$

$P(C=1) = 1\%$ $\quad P(T=1 \mid C=1) = .8, \quad P(T=1 \mid C=0) = .09$

$P(C=0) = .99$

add up to 1

$$\underbrace{P(C=1 \mid T=1)} \stackrel{\text{Bayes}}{=} \frac{\overbrace{P(T=1 \mid C=1)}^{\text{conditional}} \overbrace{P(C=1)}^{\text{prior}}}{\underbrace{P(T=1)}_{\text{normalization}}} = \frac{.8 * .01}{\boxed{P(T=1)}} = \frac{.008}{P(T=1)}$$

$$P(C=0 \mid T=1) \stackrel{\text{Bayes}}{=} \frac{P(T=1 \mid C=0) P(C=0)}{P(T=1)} = \frac{.09 * .99}{\boxed{P(T=1)}} = \frac{.0891}{P(T=1)}$$

$$\frac{.008}{P(T=1)} + \frac{.0891}{P(T=1)} = 1 \qquad \underline{P(T=1)} = .0971$$

$$P(C=1 \mid T=1) = \frac{.008}{.0971} = 8\%$$

Conditioning on Background Evidence

if true

- Product rule without background evidence: $P(X, Y) = P(Y|X)P(X)$
- Claim: $P(X, Y|E) = P(Y|X, E)P(X|E)$

Prove the conditional version of the product rule

$$P(X,Y|E) \overset{\text{product}}{=\!=} \frac{P(X,Y,E)}{P(E)} = \frac{\cancel{P(E)}\,P(X|E)\,P(Y|X,E)}{\cancel{P(E)}}$$

$$= P(Y|X,E)\,P(X|E) \quad \#$$

Conditioning on Background Evidence

- Product rule (conditional): $P(X, Y|E) = P(Y|X, E)P(X|E)$
- Bayes' rule (conditional):

$$P(X|Y, E) = \frac{P(Y|X, E)P(X|E)}{P(Y|E)}$$

- Marginalization (conditional):

$$P(X|E) = \sum_{y} P(X, Y = y|E)$$

Review

Complete each of the following rules (and give their conditionalized versions):

conditional

Product Rule:

$$P(A, B, C, D, \dots) = P(A)P(B|A)P(C|BA)P(D|ABC)\dots$$

$$P(A,B,C,D|E) = P(A|E)P(B|AE)P(C|ABE)P(D|ABCE)\dots$$

Bayes' Rule:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

$$P(X|Y,E) = \frac{P(Y|X,E)P(X|E)}{P(Y|E)}$$

Marginalization

$$P(X) = \sum_y P(X, Y = y)$$

$$P(X|E) = \sum_y P(X, Y=y|E)$$

A: always true

B: No

1 Rules of Probability

1.1 Marginalization

Which of the following statements are guaranteed to be true?

(a) $\sum_a P(A = a) = 1$ Yes

(b) $\sum_a P(A = a|B = b) = \cancel{P(B = b)}$ No

(c) $\sum_a P(A = a, B = b) = P(B = b)$ yes (marginalization rule)

(d) $\sum_b P(A = a|B = b) = P(A = a)$ No

$P(T=1|C=1) + P(T=1|C=0) \overset{\times ??}{=} P(T=1)$

$\overset{\checkmark}{=} P(T=1, C=1) + P(T=1, C=0)$

(e) $\sum_b P(A = a|B = b) = 1$ × No

(f) $\sum_a P(A = a|B = b) = 1$ ✓ yes

(g) $\sum_a P(A = a, B = b|C = c) = P(B = b|C = c)$ yes

1.2 Conditional Independence

If we know that A and B are conditionally independent given C, which of the following statements are guaranteed to be true?

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

(b) $P(A, B|C) = P(A, B)$

(c) $P(A, B|C) = P(A|C)P(B|C)$

(d) $P(A, B, C) = P(C)P(A|C)P(B|C)$

(e) $P(A|B, C) = P(A|B)$

(f) $P(B|A, C) = P(B|C)$

1.3 Product Rule

Use the product rule to rewrite $P(A, B, C)$ in 12 distinct ways. (For this problem, changing the order of variables and rearranging products do not count as "distinct" answers.)

$$P(ABC) = P(A)P(BC|A) = P(B)P(AC|B) = P(C)P(AB|C)$$

$$P(ABC) = P(A)P(B|A)P(C|AB) = P(A)P(C|A)P(B|AC)$$

$$= P(B) \checkmark \qquad\qquad = \checkmark$$

$$= P(C) \checkmark \qquad\qquad = \checkmark$$

$$P(ABC) = P(AB)P(C|AB)$$

$$= P(AC)P(B|AC)$$

$$= P(BC)P(A|BC)$$

1.4 Bayes' Rule

Suppose that on any given day there is a 40% chance that I will be in a bad mood if I read the news, and a 10% chance that I will be in a bad mood if I do not read the news. Suppose also that each day there is a 70% chance that I read the news. If I am in a bad mood, what is the probability that I have read the news that day?

Independence

- Marginal Independence

If X and Y are independent, which of the following is NOT true?

A. $P(X, Y) = P(X) + P(Y)$

B. $P(X|Y) = P(X)$

C. $P(Y|X) = P(Y)$

D. $P(X, Y) = P(X)P(Y)$

E. Not all of the choices are true

Education with AI

- https://www.youtube.com/watch?v=GQVIOchWIgw

Consider the following three events

- E: There was an earthquake
- B: A burglar broke into my apartment
- A: My alarm at my apartment went off

Assume that B and E are marginally independent

1. Multiple explanations for a single event

- Compare :
  - $P(B = 1)$
  - $P(B = 1|A = 1)$
  - $P(B = 1|A = 1, E = 1)$

Which of the following represents the most intuitive relationship between these probabilities?

$A.$ $P(B = 1) < P(B = 1|A = 1)$
$B.$ $P(B = 1) < P(B = 1|A = 1, E = 1)$
$C.$ $P(B = 1) > P(B = 1|A = 1)$
$D.$ $P(B = 1) > P(B = 1|A = 1, E = 1)$
E. More than one of these (which ones?)

1. Multiple explanations for a single event

- Compare :
  - $P(B = 1)$
  - $P(B = 1|A = 1)$
  - $P(B = 1|A = 1, E = 1)$

Which of the following represents the most intuitive relationship between these probabilities?

$A.\ P(B = 1|A = 1) < P(B = 1|A = 1, E = 1)$

$B.\ P(B = 1|A = 1) > P(B = 1|A = 1, E = 1)$

$C.\ P(B = 1|A = 1) = P(B = 1|A = 1, E = 1)$

1. Multiple explanations for a single event

- Compare :
  - $P(B = 1)$
  - $P(B = 1|A = 1)$
  - $P(B = 1|A = 1, E = 1)$

Intuitively, we expect:
$P(B = 1) < P(B = 1|A = 1, E = 1) < P(B = 1|A = 1)$

I.e. "I generally believe there is a very low chance that a burglar breaks in. But if I get notified that my alarm is going off, then I think the alarm is probably caused by the burglar and my belief that a burglar broke in goes up. However, if I then hear about an earthquake, I no longer am as fearful that a burglar broke in, but I'm still more worried than I would have been if the alarm hadn't gone off at all"

1. Multiple explanations for a single event

$P(B = 1)$
$P(B = 1|A = 1)$
$P(B = 1|A = 1, E = 1)$

Domain knowledge:

Assumption: B and E are marginally independent: $P(B|E) = P(B), P(E|B) = P(E)$

Prior probabilities:

$$P(B = 1) = 0.001 \quad P(B = 0) = 1 - 0.001 = 0.99$$
$$P(E = 1) = 0.005 \quad P(E = 0) = 1 - 0.005 = 0.95$$

Conditional probabilities (i.e. properties of the alarm):

B	E	$P(A = 1\|B, E)$	$P(A = 0\|B, E)$