Review

- Probabilities

unconditional: $P(X)$
conditional: $P(Y|X)$
joint: $P(X,Y)$

- Conditional independence

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

- Rules

$$P(A,B,C,\ldots) = P(A)P(B|A)P(C|A,B)\ldots \quad \text{(Product rule)}$$

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \quad \text{(Bayes rule)}$$

$$P(X) = \sum_y P(X, Y=y) \quad \text{(Marginalization)}$$

Probabilistic Inference

Do probabilities capture common sense patterns of reasoning?

Examples: reasoning about —

1) multiple explanations of a single event
2) multiple events with a single explanation
3) intervening events

- Let's define some binary random variables

B = burglary? $\begin{cases} B=1 \Rightarrow \text{burglary occurred} \\ B=0 \Rightarrow \text{burglary did not occur} \end{cases}$

E = earthquake?

A = alarm?

- Joint distribution

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

- Domain knowledge

$$P(B=1) = 0.001$$

$P(E=1|B=0) = 0.002$
$P(E=1|B=1) = 0.002$ $\Big\}$ assumption of conditional independence (knowing B doesn't affect E)

$P(E=1|B) = 0.002 = P(E=1)$

B	E	$P(A=1 \mid B, E)$
0	0	0.001
0	1	0.29
1	0	0.94
1	1	0.95

1) Reasoning about multiple explanations

Want to compare:
① $P(B=1) = 0.001$
② $P(B=1|A=1)$
③ $P(B=1|A=1, E=1)$

② $$P(B=1|A=1) = \frac{\overbrace{P(A=1|B=1)}^{\alpha}\ \overbrace{P(B=1)}^{0.001}}{\underbrace{P(A=1)}_{\beta}} \qquad \text{(Bayes rule)}$$

$$\beta = P(A=1) = \sum_{b,e} P(A=1, B=b, E=e) \qquad \text{(marginalization)}$$

$$= \sum_{b,e} P(B=b)\, P(E=e \mid B=b)\, P(A=1 \mid B=b, E=e) \qquad \text{(product rule)}$$

$$= \sum_{b,e} P(B=b)\, P(E=e)\, P(A=1 \mid B=b, E=e) \qquad \text{(E,B conditionally independent)}$$

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

(expand sum)

use $P(B=0) = 1 - P(B=1)$
$P(E=0) = 1 - P(E=1)$

$$= \underline{0.00252} \qquad \text{(substitution)}$$

$$\alpha = P(A=1 \mid B=1)$$

$$= \sum_{e} P(A=1, E=e \mid B=1) \qquad \text{("conditionalized" marginalization)}$$

$$= \sum_{e} P(A=1 \mid E=e, B=1)\, P(E=e \mid B=1) \qquad \left(\begin{array}{c}\text{conditionalized product rule} \\ P(X, Y \mid Z) = P(X \mid Y, Z)P(Y \mid Z)\end{array}\right)$$

$$= \sum_{e} P(A=1 \mid E=e, B=1)\, P(E=e) \qquad \text{(conditional ind.)}$$

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

$$= (0.95 \times 0.002) + (0.94 \times 0.998) = \underline{0.94002}$$

$$\Rightarrow \quad P(B=1|A=1) = \frac{\overbrace{P(A=1|B=1)}^{\alpha}\,P(B=1)}{\underbrace{P(A=1)}_{\beta}}$$

$$= \frac{0.94002 \times 0.001}{0.00252} = \underline{0.37} \rightarrow ②$$

---

③ $P(B=1|A=1, E=1)$

$$= \frac{\overbrace{P(A=1|B=1, E=1)}^{0.95}\ \overbrace{P(B=1|E=1)}^{\alpha}}{\underbrace{P(A=1|E=1)}_{\beta}}$$ conditionalized Bayes rule

$$\alpha = P(B=1|E=1) = P(B=1) \qquad \text{(conditional independence)}$$

$$= \underline{0.001}$$

$$\beta = P(A=1|E=1)$$

$$= \sum_b P(A=1, B=b \,|\, E=1) \qquad \text{(conditionalized marginalization)}$$

$$= \sum_b P(A=1|B=b, E=1)\,P(B=b|\underbrace{\cancel{E=1}}_{\text{cond'l ind.}}) \qquad \text{(product rule)}$$

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

$$= 0.95 \times 0.001 + 0.29 \times 0.999 \approx \underline{0.29}$$

$$\Rightarrow \quad P(B=1|A=1, E=1) = \frac{0.95 \times 0.001}{0.29} = \underline{0.0033} \rightarrow ③$$

Now compare:

1. $P(B=1) = 0.001$
2. $P(B=1 \mid A=1) = 0.37$ ↑
3. $P(B=1 \mid A=1, E=1) = 0.0033$ ↓

(2) is $> P(B=1)$. It is expected because knowing alarm went off $(A=1)$, there is a higher chance that $B=1$.

(3) is $< P(B=1 \mid A=1)$. Here we also know that earthquake occurred, so knowing $A=1$ AND $E=1$, means there is a lower chance that $B=1$.

Earthquake "explains away" the alarm, weakening our belief in burglary.

→ Arises from multiple (causal) explanations of observed event.

---

2) <u>Multiple events with a common explanation</u>

Let's consider two more binary random variables

$J =$ John calls police?
$M =$ Mary calls police?

- Conditional independence

Already assumed: $P(B \mid E) = P(B)$
Also assume: $P(J \mid A) = P(J \mid A, B, E)$
$P(M \mid A) = P(M \mid A, B, E)$

Probability that Mary calls given alarm is same as prob. that Mary calls given information about alarm, burglary, earthquake.

In other words, prob. that Mary calls is conditionally independent of B, E given alarm.

- Joint distribution:

$$P(B,E,A,J,M) = P(B)\,P(E|B)\,P(A|B,E)\,P(J|A,B,E)\,P(M|A,B,E,J) \quad \text{(product rule)}$$

$P(E|B) \to P(E)$ — E, B cond. ind.

$P(J|A,B,E) \to P(J|A)$ — cond. ind. assumption

$P(M|A,B,E,J) \to P(M|A,B,E)$ — John calling doesn't affect Mary given A, B, E

$P(M|A,B,E) \to P(M|A)$ — cond. ind. assumption

$$P(B,E,A,J,M) = P(B)\,P(E)\,P(A|B,E)\,P(J|A)\,P(M|A)$$

- Conditional probs

$P(J=1|A=0) = 0.05$ $\qquad$ $P(M=1|A=0) = 0.01$

$P(J=1|A=1) = 0.9$ $\qquad$ $P(M=1|A=1) = 0.7$

Want to compare:

4 $\quad P(A=1) = 0.00252$ (from prev. ex.)

5 $\quad P(A=1|J=1)$

6 $\quad P(A=1|J=1, M=0)$

$$⑤ \quad P(A=1 \mid J=1) = \frac{\overset{0.9}{P(J=1 \mid A=1)}\ \overset{0.00252}{P(A=1)}}{P(J=1)}$$

$$P(J=1) = \sum_{b,e,a} P(J=1, B=b, E=e, A=a)$$

here we have 8 terms; not the smart way, but does give right answer

We can expand $P(J=1)$ as

$$P(J=1) = \sum_a P(J=1, A=a)$$

$$= \sum_a P(J=1 \mid A=a)\, P(A=a)$$

$$= P(J=1 \mid A=0) P(A=0) + P(J=1 \mid A=1) P(A=1)$$

$$= 0.05 \times (1-0.00252) + 0.9 \times 0.00252$$

$$= 0.0521$$

$$\Rightarrow P(A=1 \mid J=1) = \frac{P(J=1 \mid A=1) P(A=1)}{P(J=1)}$$

$$= \frac{0.9 \times 0.00252}{0.0521} = 0.0435 \rightarrow ⑤$$

$$⑥ \quad P(A=1 \mid J=1, M=0) = \frac{P(J=1, M=0 \mid A=1) P(A=1)}{P(J=1, M=0)}$$

(Bayes rule w/ multiple effects $P(X \mid Y,Z) = \frac{P(Y,Z \mid X) P(X)}{P(Y,Z)}$)

$$= \frac{P(J=1 \mid A=1) P(M=0 \mid A=1) P(A=1)}{P(J=1, M=0)}$$

(cond. ind. of J, M given A)

$$P(J=1, M=0) = \sum_a P(J=1, M=0, A=a)$$

$$= \sum_a P(A=a)P(J=1 \mid A=a)P(M=0 \mid A=a, \cancel{J=1})$$

$$= P(A=1)P(J=1 \mid A=1)P(M=0 \mid A=1) + P(A=0)P(J=1 \mid A=0)P(M=0 \mid A=0)$$

$$= 0.00252 \times 0.9 \times 0.3 + 0.99748 \times 0.05 \times 0.99$$

$$= 0.05$$

$$\Rightarrow P(A=1 \mid J=1, M=0) = \frac{P(J=1 \mid A=1)P(M=0 \mid A=1)P(A=1)}{P(J=1, M=0)}$$

$$= \frac{0.9 \times 0.3 \times 0.00252}{0.05} = 0.0136 \rightarrow ⑥$$

Now compare:

④ $P(A=1) = 0.00252$

⑤ $P(A=1 \mid J=1) = 0.0435 \uparrow$

⑥ $P(A=1 \mid J=1, M=0) = 0.0136 \downarrow$

⑤ is $> P(A=1)$ because knowing that John called police means there is a higher chance that alarm went off $(A=1)$

⑥ is $< P(A=1 \mid J=1)$ because knowing further that Mary didn't call police, our belief that alarm ringing $(A=1)$ will decrease.

⑥ is $> P(A=1)$ because, despite Mary not calling, John calling still increases our belief that alarm went off