01/14/2014

Lecture Notes #3

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

$P(A,B,C\ldots) = P(A)\,P(B/A)\,P(C/A,B)\ldots$ (Product Rule)

$P(x/y) = P(y/x)\,P(x)\,P(y)$ (Bayes Rule)

$P(x) = \sum_{Y} P(x, Y=y)$ (marginalisation)

PROBABILISTIC INFERENCE

→ do probabilities capture commonsense 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? $\begin{cases} B=1 \Rightarrow \text{Burglary occured} \\ B=0 \Rightarrow \text{Burglary did not occur} \end{cases}$

E = earthquake?

A = alarm?

→ Joint distribution

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

→ Domain Knowledge:

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

$P(E=1|B=0) = 0.002$ } assumption of Conditional independence.
$P(E=1|B=1) = 0.002$ } (Knowing B, doesn't affect E)

→ $P(E=1|B) = 0.002$.

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

1) reasoning about multiple explanations.

$P(B=1) = 0.001 \rightarrow$ ①

$$P(B=1|A=1) = \frac{\overbrace{P(A=1|B=1)}^{\alpha}\ \overbrace{P(B=1)}^{0.001}}{\underbrace{P(A=1)}_{\beta}} \quad [\because \text{Bayes Rule}]$$

~~$\alpha\beta = P(A=1)$ $\beta = P(A=1)$~~

$$\beta = P(A=1) = \sum_{b,e} P(A=1, B=b, E=e) \quad [\text{Marginalisation}]$$

$$= \sum_{b,e} P(B=b)\ \underbrace{P(E=e|B=b)}_{E,B \text{ Conditionally independent}}\ P(A=1|B=b, E=e)$$

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

expanding, we get.

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

$= 0.00252$ // (after substitution)

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

~~P(A=0)~~

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

$$= \sum_e P(A=1, E=e / B=1) \quad \text{[marginalisation "Conditionalised"]}$$

$$= \sum_e P(A=1 / E=e, B=1)\, P(E=e / \cancel{B=1}) \quad \left[\begin{array}{l}\text{Conditionalised product rule} \\ P(XY/Z) = P(X/Y,Z)\,P(Y/Z)\end{array}\right]$$

$$= \sum_e P(A=1 / E=e, B=1)\, P(E=e) \quad \text{[E,B} \rightarrow \text{Conditional ind.]}$$

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

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

So, $$P(B=1 / A=1) = \frac{P(A=1 / B=1)\, P(B=1)}{P(A=1)}$$

$$= \frac{0.94002 \times 0.001}{0.00252}$$

$$P(B=1 / A=1) = 0.37 \rightarrow (2)$$

$$\boxed{\text{"Conditionalised" Bayes Rule} \quad P(X/Y,Z) = \frac{P(Y/X,Z)\, P(X/Z)}{P(Y/Z)}}$$

$$P(B=1 / A=1, E=1) = \frac{\overset{0.95}{P(A=1 / B=1, E=1)}\, \overset{\alpha}{P(B=1 / E=1)}}{\underset{\beta}{P(A=1 / E=1)}}$$

$$\alpha = P(B=1 / E=1) = P(B=1) \quad \text{[Conditional Independence]}$$

$$= 0.001$$

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

$$= \sum_b P(A=1, B=b / E=1) \quad \text{[Conditionalised marginalisation]}$$

$$= \sum_b P(A=1 / B=b, E=1)\, P(B=b / \cancel{E=1}) \quad \text{[product rule]}$$

"Conditional independence"

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

$$= 0.95 \times 0.001 + 0.29 \times 0.999 \approx 0.29$$

So, $P(B=1 / A=1, E=1) = \frac{0.95 \times 0.001}{0.29} = 0.0033$ $\rightarrow$ (3)

Comparing (1), (2) & (3) :

$P(B=1) = 0.001$

$P(B=1/A=1) = 0.37 \uparrow$ [This is $> P(B=1)$. It is ~~intuitive~~ expected because knowing alarm went off $(A=1)$, there are more chances that $B=1$]

$P(B=1/A=1, E=1) = 0.0033 \downarrow$ [This is $< P(B=1/A=1)$. Since here we also know that earthquake has occured, knowing $A=1$ $\rightarrow$ we have less chances for $B=1$ compared to only knowing $A=1$]

$\rightarrow$ earthquake "explains away" the alarm, weakening our belief in burglary.

$\rightarrow$ Arises from multiple (causal) explanations of observed event.

2) Multiple events with a common explanation:

Let's consider two more binary variables

J = John calls police ?

M = Mary calls police ?

$\rightarrow$ Conditional independence

we have already assumed – $P(B/E) = P(B)$

Also assume – $P(J/A) = P(J \mid A, B, E)$

$P(M/A) = P(M \mid A, B, E)$

which means

$\rightarrow$ Prob. 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:

$P(B,E,A,J,M) = P(B)\ \underbrace{P(E/B)}_{\substack{\downarrow \\ P(E) \\ E,B \text{ conditionally independent.}}}\ P(A/B,E)\ \underbrace{P(J/A,B,E)}_{\substack{\downarrow \text{ by assumption of Conditional independence} \\ P(J/A)}}\ \underbrace{P(M/A,B,E,J)}_{\substack{\downarrow \text{ JohnCalling doesn't affect Mary given } A,B,E \\ P(M/A,B,E) \\ \downarrow \text{ by conditional independence assumpti} \\ P(M/A)}}$ (Product rule)

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

→ Conditional probs.

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

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

Let's Compare i) $P(A=1) = 0.00252$ (from prev. example) → ④

ii) $P(A=1/J=1)$

iii) $P(A=1/J=1, M=0)$

ii) $P(A=1/J=1) = \dfrac{\overset{0.9}{P(J=1/A=1)}\ \overset{0.00252.}{P(A=1)}}{P(J=1)}$

$P(J=1) = \sum_{b,e,a} P(J=1, B=b, E=e, A=a)$ → here we have 8 terms. → not the smart way to do. although it gives the right ans.

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

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

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

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

$$= 0.05 \times (1-0.00252) + 0.9 \times \underbrace{0.00252}_{\text{from prev. example we have calculated } P(A=1)}$$

$$= 0.0521$$

So, $P(A=1 \mid J=1) = \dfrac{P(J=1 / A=1)\ P(A=1)}{P(J=1)}$

$$= \frac{0.9 \times 0.00252}{0.0521}$$

$P(A=1 / J=1) = 0.0435 \rightarrow$ (5)

iii) $P(A=1/J=1, M=0) = \dfrac{P(J=1, M=0 / A=1)\ P(A=1)}{P(J=1, M=0)}$

[Bayes rule with multiple ~~effec~~ effects.
$P(x/y,z) = \dfrac{P(y,z/x)\ P(x)}{P(y,z)}$]

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

[Conditional independence of J, M given A]

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

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

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

[Conditional independence given A; J, M are conditionally independent]

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

$= \cancel{0.0136}\ 0.00252 \times 0.9 \times 0.3 + 0.99748 \times 0.05 \times 0.99$

$= 0.05$

So, $P(A=1 \mid J=1, M=0) = \cancel{\dfrac{P(J=1, M=0/A=1)\ P(A=1)}{P(J=1, M=0)}}\ \dfrac{P(J=1/A=1)\ P(M=0/A=1)\ P(A=1)}{P(J=1, M=0)}$

$$= \frac{0.9 \times 0.3 \times 0.00252}{0.05}$$

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

Comparing ④, ⑤ & ⑥

$P(A=1) = 0.00252.$

$P(A=1/J=1) = 0.0521$ ↑ [This is $> P(A=1)$ because knowing that John called 'police, there is more that alarm went off $(A=1)$] -+

$P(A=1/J=1, M=0) = 0.0136$ ↓ [This is $< P(A=1/J=1)$ because knowing further that Mary didn't call police, oure belief on alarm ringing $(A=1)$ will decrease]