CS 452 – PROBABILISTIC GRAPHICAL MODELS

Unit # 2

RECAP

- Independence and Conditional independence
- Law of total probability
- Chain Rule
- Bayes Theorem
- Examples
- Exercises

CONDITIONAL INDEPENDENCE

Two events A and B are independent if knowing that A has happened does not say anything about B happening.

$$P(A\ B) = P(A)\ P(B)$$

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

Two events A and B are conditionally independent given a third event C precisely if the occurrence or non-occurrence of A and B are independent events in their conditional probability distribution given C.

$$P(A\ B \mid C) = P(A \mid C)\ P(B \mid C)$$

$$P(A \mid B\ C) = P(A \mid C)$$

EXAMPLE (SOURCE NEAPOLITAN, 2009)

Let S = Sex, H = Height and W = Wage

Case	Sex	Height (inches)	Wage ($)
1	female	64	30,000
2	female	64	30,000
3	female	64	40,000
4	female	64	40,000
5	female	68	30,000
6	female	68	40,000
7	male	64	40,000
8	male	64	50,000
9	male	68	40,000
10	male	68	50,000
11	male	70	40,000
12	male	70	50,000

EXAMPLE (CONT'D)

s	$P(s)$
female	$1/2$
male	$1/2$

h	$P(h)$
64	$1/2$
68	$1/3$
70	$1/6$

w	$P(w)$
30,000	$1/4$
40,000	$1/2$
50,000	$1/4$

The joint distribution of S and H is as follows:

s	h	$P(s,h)$
female	64	$1/3$
female	68	$1/6$
female	70	0
male	64	$1/6$
male	68	$1/6$
male	70	$1/6$

EXAMPLE (CONT'D) MARGINAL AND JOINT DISTRIBUTION

h / s	64	68	70	Distribution of S
female	$1/3$	$1/6$	0	$1/2$
male	$1/6$	$1/6$	$1/6$	$1/2$
Distribution of H	$1/2$	$1/3$	$1/6$	

EXERCISE 1 (SOURCE: NEAPOLITAN, 2009)

Are H and W independent?

Are H and W conditionally independent given S?

EXAMPLE (CONT'D)

P(h and w) = P(h) P(w)

P(h=64 and w=40,000) = P(h=64) P(w=40000)

3/12 = ½ * ½

¼ = ¼

P(hw|s) = P(h|s) P(w|s)

P(h=64 and w=40,000 |female) = P(h=64|female) P(w=40000|female)

2/6 = 4/6 * 3/6

1/3 = 1/3

Case	Sex	Height (inches)	Wage ($)
1	female	64	30,000
2	female	64	30,000
3	female	64	40,000
4	female	64	40,000
5	female	68	30,000
6	female	68	40,000
7	male	64	40,000
8	male	64	50,000
9	male	68	40,000
10	male	68	50,000
11	male	70	40,000
12	male	70	50,000

EXAMPLE (CONT'D)

Are H and W independent?

		P(h)	P(w)	P(h)P(w)	P(hw)
64	30000	½	¼	1/8	1/6
64	40000	½	½	¼	¼
64	50000	½	¼	1/8	1/12
68	30000	1/3	¼	1/12	1/12
68	40000	1/3	½	1/6	1/6
68	50000	1/3	¼	1/12	1/12
70	30000	1/6	¼	1/24	0
70	40000	1/6	½	1/12	1/12
70	50000	1/6	¼	1/24	1/12

Case	Sex	Height (inches)	Wage (\$)
1	female	64	30,000
2	female	64	30,000
3	female	64	40,000
4	female	64	40,000
5	female	68	30,000
6	female	68	40,000
7	male	64	40,000
8	male	64	50,000
9	male	68	40,000
10	male	68	50,000
11	male	70	40,000
12	male	70	50,000

EXAMPLE (CONT'D)

Are H and W conditionally independent given S?

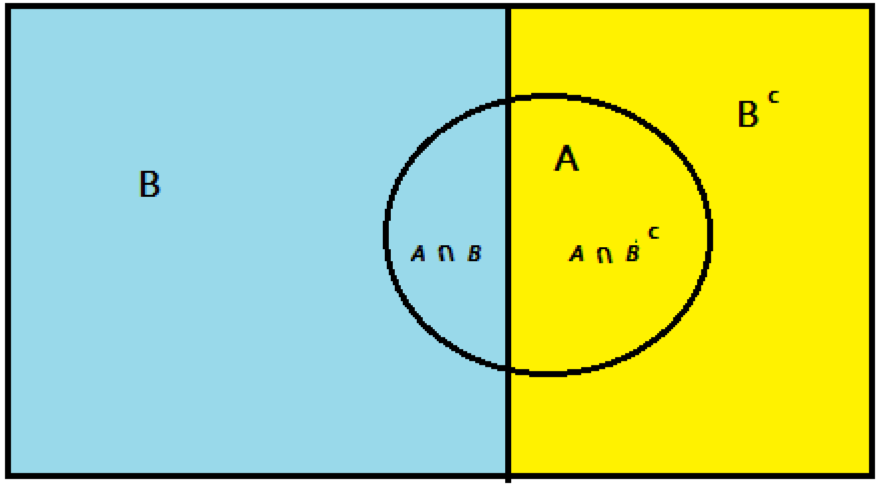
h	w	P(h | s)	P(w | s)	P(h | s)P(w | s)	P(hw | s)
64	30000	4/6	½	1/3	1/3
64	40000	4/6	½	1/3	1/3
64	50000	4/6	0	0	0
68	30000	2/6	½	1/6	1/6
68	40000	2/6	½	1/6	1/6
68	50000	2/6	0	0	0
70	30000	0	½	0	0
70	40000	0	½	0	0
70	50000	0	0	0	0

Case	Sex	Height (inches)	Wage ($)
1	female	64	30,000
2	female	64	30,000
3	female	64	40,000
4	female	64	40,000
5	female	68	30,000
6	female	68	40,000
7	male	64	40,000
8	male	64	50,000
9	male	68	40,000
10	male	68	50,000
11	male	70	40,000
12	male	70	50,000

LAW OF TOTAL PROBABILITY

- Law of Total Probability
  (aka “summing out” or marginalization):
  $P(a) = \Sigma_b P(a, b) = \Sigma_b P(a \mid b) P(b)$

  where B is any random variable



- Why is this useful?
  - given a joint distribution (e.g., P(a,b,c,d)) we can obtain any “marginal” probability (e.g., P(b)) by summing out the other variables, e.g.

CHAIN RULE

Given variables $X_1, X_2, \ldots.., X_n$, the chain rule is

$$P(x_1, x_2, \ldots.., x_n) = P(x_n \mid x_{n-1}, \ldots.., x_2, x_1) \ldots\ldots\ldots$$

$$P(x_2 \mid x_1)\, P(x_1)$$

CHAIN RULE

For events $A_1, \ldots, A_n$ whose intersection has not probability zero, the chain rule states

$$\begin{aligned}
\mathbb{P}\left(A_1 \cap A_2 \cap \ldots \cap A_n\right) &= \mathbb{P}\left(A_n \mid A_1 \cap \ldots \cap A_{n-1}\right) \mathbb{P}\left(A_1 \cap \ldots \cap A_{n-1}\right) \\
&= \mathbb{P}\left(A_n \mid A_1 \cap \ldots \cap A_{n-1}\right) \mathbb{P}\left(A_{n-1} \mid A_1 \cap \ldots \cap A_{n-2}\right) \mathbb{P}\left(A_1 \cap \ldots \cap A_{n-2}\right) \\
&= \mathbb{P}\left(A_n \mid A_1 \cap \ldots \cap A_{n-1}\right) \mathbb{P}\left(A_{n-1} \mid A_1 \cap \ldots \cap A_{n-2}\right) \cdot \ldots \cdot \mathbb{P}(A_3 \mid A_1 \cap A_2) \mathbb{P}(A_2 \mid A_1) \mathbb{P}(A_1) \\
&= \mathbb{P}(A_1) \mathbb{P}(A_2 \mid A_1) \mathbb{P}(A_3 \mid A_1 \cap A_2) \cdot \ldots \cdot \mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}) \\
&= \prod_{k=1}^{n} \mathbb{P}(A_k \mid A_1 \cap \cdots \cap A_{k-1}) \\
&= \prod_{k=1}^{n} \mathbb{P}\left(A_k \,\middle|\, \bigcap_{j=1}^{k-1} A_j\right).
\end{aligned}$$

EXAMPLE (CONT'D) CHAIN RULE

$$P(s, h, w) = P(w|h, s)P(h|s)P(s).$$

There are eight combinations of values of the three random variables. The table that follows shows that the equality holds for two of the combination.

s	h	w	$P(s, h, w)$	$P(w\|h, s)P(h\|s)P(s)$
female	64	30,000	$\frac{1}{6}$	$\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{6}$
female	64	40,000	$\frac{1}{12}$	$\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{12}$

It is left as an exercise to show that the equality holds for the other six combinations.

BAYES THEOREM

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

$$= \frac{P(B \mid A)\, P(A)}{P(B|A)P(A)+P(B|\neg A)P(\neg A)}$$

P(A) is the prior probability and P(A | B) is the posterior probability.

Suppose events A_1, A_2, ……., A_k are mutually exclusive and exhaustive; i.e., exactly one of the events must occur. Then for any event B:

$$P(A_i \mid B) = \frac{P(B \mid A_i)\, P(A_i)}{\sum P(B \mid A_i)\, P(A_i)}$$

EXAMPLE I

According to American Lung Association, 7% of the population has lung cancer. Of these people having lung disease, 90% are smokers; and of those not having lung disease, 25.3% are smokers.

Determine the probability that a randomly selected smoker has lung cancer.

EXAMPLE I SOLUTION

Let L = Lung Cancer, S = Smoker

Given that

- $P(L) = 0.07$
- $P(S \mid L) = 0.90$ $P(\sim S \mid L) = 0.10$
- $P(S \mid \sim L) = 0.253$ $P(\sim S \mid \sim L) = 0.747$

Find probability, $P(L \mid S)$

$$P(L \mid S) = \frac{P(S \cap L)}{P(S)} = \frac{P(S \mid L)P(L)}{P(S \mid L)P(L) + P(S \mid \sim L)P(\sim L)}$$

$$P(L \mid S) = \frac{0.9\ x\ 0.07}{0.9\ x\ 0.07 + 0.253\ x\ 0.93}$$

EXAMPLE II

Assume that about 1 in 1000 individuals in a given organization have committed a security violation.

Assume that the **sensitivity** of a routine screening polygraph is about 85%. That is, the probability that the polygraph report will indicate a concern is about 85% if the individual has committed a security violation.

Assume the **specificity** of the polygraph is about 80%. That is, if the individual has not committed a security violation, there is about an 80% chance that the polygraph report will not indicate a concern.

What is the posterior probability that an individual whose polygraph report indicates a concern has committed a security violation?

EXAMPLE II SOLUTION

Let

- S = Security Violation Committed,
- T = Test Positive

Given that

- P(S) = 0.001
- P(T | S) = 0.85 P(~T | S) = 0.15
- P(T | ~S) = 0.20 P(~T | ~S) = 0.80

Find probability, P(S | T)

$$P(S \mid T) = \frac{P(T \mid S)P(S)}{P(T \mid S)P(S) + P(T \mid {\sim}S)P({\sim}S)}$$

$$P(T|S) = \frac{0.85 \; x \; 0.001}{0.85 \; x \; 0.001 + 0.20 \; x \; 0.999}$$

EXERCISE 1 (SOURCE: NEAPOLITAN, 2009)

A forgetful nurse is supposed to give Mr. Nguyen a pill each day.

The probability that the nurse will forget to give the pill on a given day is 0.3.

If Mr. Nguyen receives the pill, the probability he will die is 0.1.

If he does not receive the pill, the probability he will die is 0.8.

Mr. Nguyen died today. Use Bayes' Theorem to compute the probability that the nurse forgot to give him the pill.

EXAMPLE 2

Suppose that Bob can decide to go to work by one of three modes of transportation, car, bus, or commuter train. Because of high traffic, if he decides to go by car, there is a 50% chance he will be late. If he goes by bus, which has special reserved lanes but is sometimes overcrowded, the probability of being late is only 20%. The commuter train is almost never late, with a probability of only 1%, but is more expensive than the bus.

Suppose that Bob is late one day, and his boss wishes to estimate the probability that he drove to work that day by car. Since he does not know which mode of transportation Bob usually uses, he gives a prior probability of 1/3 to each of the three possibilities. What is the boss' estimate of the probability that Bob drove to work?

EXERCISE 3 (SOURCE: NEAPOLITAN, 2009)

An oil well might be drilled on Professor Neapolitan's farm in Texas.

Based on what has happened on similar farms, we judge the probability of oil being present to be 0.5, the probability of only natural gas being present to be 0.2, and the probability of neither being present to be 0.3.

If oil is present, a geological test will give a positive result with probability 0.9; if only natural gas is present, it will give a positive result with probability 0.3; and if neither is present, the test will be positive with probability 0.1.

Suppose the test comes back positive. Use Bayes' Theorem to compute the probability that oil is present.

THANKS