Lecture 9

Quantifying Uncertainty

9.1 Uncertainty

Probability theory is a way to deal with **uncertain information**. Probability represents the **degree of belief** not ~~the degree of truth.~~

Example 9.1 "*A probability of* 0.8 *for having cavity in a patient with toothache*" means "*We believe that there is an* 80% *chance that the patient with toothache has a cavity*"

Example 9.2 When we draw a card from a shuffled pack, we have a probability of 1/52 that the card is the ace of spades before we look at the card. Then, *it becomes either* 0 *or* 1 *after looking at the card.*

9.2 Logic and Uncertainty

In a dental diagnosis domain, we may have

$$Toothache \Rightarrow Cavity \vee GumProblem \vee Abscess \ldots$$

Using logic to cope with this domain fails for three main reasons:

- **Laziness**
  It requires a lot of efforts to generate all related rules and facts.
- **Theoretical ignorance**
  There is no complete theory for the domain.
- **Practical ignorance**
  Even we have all the rules, it may not be possible to obtain all the necessary facts.

9.3 Basic Probability Theory

set of all possible events

Let Ω be the sample space, and $\omega \in \Omega$ be a sample point or an atomic event.

$$0 \leq P(\omega) \leq 1 \text{ for every } \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1$$

Example 9.3 If we roll two fair distinct dice, there are 36 possible worlds: $(1,1), (1,2), \ldots, (6,6)$. The probability that both dice produce $(1,1)$ is

$$P((1,1)) = \frac{1}{36}$$

The probability associated with a proposition is computed from the sum of the probabilities of the worlds:

a set of events

$$\text{For any proposition } \phi, P(\phi) = \sum_{\omega \in \phi} P(\omega)$$

Exercise 9.1 When rolling two fair distinct dice, what is the probability that the total is 11.

proposition $\longrightarrow \{(5,6), (6,5)\}$

$$P(Total = 11) = P((5,6)) + P((6+5))$$

$$= \frac{1}{36} + \frac{1}{36} = \frac{1}{18}$$

9.3.1 Unconditional Probabilities

Unconditional (or prior) probabilities refer to the degrees of belief in propositions when *there is no any other information*. For example, $P(Cavity = true) = 0.2$ shows the degree of belief that a person has cavity.

```
unconditional  P(Cavity=true) = 0.2
               pick a person at random and before having
               a dentist checks his teeth.

conditional    P(Cavity=true | Toothache=true) = 0.6
```

given

information that we known

```
                pick a person having toothache at random
```

9.3.2 Conditional Probabilities

Conditional (or posterior) probabilities refer to the degrees of belief in propositions when *we know some information.* For example, $P(Cavity = true \mid Toothache = true) = 0.6$ shows the degree of belief that a person, who has toothache, has cavity.

Conditional probabilities are defined in terms of unconditional probabilities:

$$\boxed{P(a \mid b) = \frac{P(a \wedge b)}{P(b)}}$$

This holds whenever $P(b) > 0$. It can be written as

$$\boxed{P(a \wedge b) = P(a \mid b)P(b)} = P(b \mid a)P(a) \quad \text{(product rule)}$$

(when $P(a) > 0$)

9.3.3 Random Variables and Domain

In probability theory, variables are called **random variables**. We use an uppercase letter to begin their names, e.g. *Total*, Die_1, *Cavity*. Each random variable has a **domain** i.e. the set of possible values the variable can take on. For example, the domain of *Cavity* is $\{true, false\}$; the domain of *Total* is $\{2, \ldots, 12\}$.

rolling 2 dice

Conventionally, $A = true$ is abbreviated as a, and $A = false$ is abbreviated as $\neg a$.

```
P(Cavity = true)  =  P(cavity)
P(Cavity = false) =  P(¬cavity)
```

9.3.4 Probability Distribution

A **probability distribution** shows the probabilities of the values in the domain.

a matrix of probabilities

Example 9.4 A random variable *Weather* has four possible values

$$\langle sunny, rain, cloudy, snow \rangle$$ ← domain of Weather

.

We may have the probabilities as follow:

$$
\begin{aligned}
P(Weather = sunny) &= 0.6\\
P(Weather = rain) &= 0.1\\
P(Weather = cloudy) &= 0.29\\
P(Weather = snow) &= 0.01
\end{aligned}
$$

or

probability distribution

sunny rain cloudy snow

$$\boxed{\mathbf{P}(Weather)} = \langle 0.6, 0.1, 0.29, 0.01 \rangle$$

$\mathbf{P}(X)$ gives the values of $P(X = x_i)$ for each possible i, and $\mathbf{P}(X \mid Y)$ gives the values of $\mathbf{P}(X = x_i \mid Y = y_j)$ for each possible i, j pair.

Example 9.5 $\mathbf{P}(Weather, Cavity)$ represents a **joint probability distribution**. It is a 4×2 tables of probabilities.

multiple variables

$$\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather \mid Cavity)\mathbf{P}(Cavity)$$

represents 8 equations ($W = Weather$ and $C = Cavity$):

$$
\begin{aligned}
P(W = sunny \land C = true) &= P(W = sunny \mid C = true)P(C = true)\\
P(W = rain \land C = true) &= P(W = rain \mid C = true)P(C = true)\\
P(W = cloudy \land C = true) &= P(W = cloudy \mid C = true)P(C = true)\\
P(W = snow \land C = true) &= P(W = snow \mid C = true)P(C = true)\\
P(W = sunny \land C = false) &= P(W = sunny \mid C = false)P(C = false)\\
P(W = rain \land C = false) &= P(W = rain \mid C = false)P(C = false)\\
P(W = cloudy \land C = false) &= P(W = cloudy \mid C = false)P(C = false)\\
P(W = snow \land C = false) &= P(W = snow \mid C = false)P(C = false)
\end{aligned}
$$

9.3.5 Probability Formulae

$$P(\neg a) = 1 - P(a) \qquad \text{(probability of negation)}$$

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b) \qquad \text{(inclusion-exclusion principle)}$$

Calculate probability value for an event

9.4 Inference using Full Joint Distribution

Here, we explain a technique called **probabilistic inference**.

Marginalization or **summing out** is a process that we sum up the probabilities for each possible value of the other variables.

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z}) \qquad \text{(marginalization rule)}$$

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y} \mid \mathbf{z}) P(\mathbf{z}) \qquad \text{(conditioning rule)}$$

$\mathbb{P}(Weather, Cavity)$

$P(weather = sunny) = P(Weather = sunny \wedge cavity)$
$+ P(Weather = sunny \wedge \neg cavity)$

Exercise 9.2 Given a full joint distribution for the *Toothache*, *Cavity*, *Catch*, find the following probabilities values:

	toothache		$\neg$*toothache*	
	catch	$\neg$*catch*	*catch*	$\neg$*catch*
cavity	0.108	0.012	0.072	0.008
$\neg$*cavity*	0.016	0.064	0.144	0.576

cavity

catch ∧ ¬ cavity

1. $P(cavity \vee toothache) =$

   P(cavity) + P(toothache) - P(cavity, toothache)

   = (0.108+0.012+0.072+0.008) + (0.108+0.012+0.016+0.064)
   - (0.108+0.012) = 0.28

2. $P(cavity) =$ 0.108 + 0.012 + 0.072 + 0.008
   = 0.2

3. $P(catch \wedge \neg cavity) =$ 0.016 + 0.144 = 0.16

4. $P(cavity \mid toothache) =$ P(cavity, toothache) / P(toothache)

   = (0.108 + 0.012) / (0.108+0.012+0.016+0.064)
   = 0.6

5. $P(catch \mid toothache \wedge cavity) = \dfrac{\text{P(catch, toothache, cavity)}}{\text{P(toothache, cavity)}}$

$$= \frac{0.108}{0.108 + 0.012} = 0.9$$

Exercise 9.3 Two distinct dice are rolled. Calculate the probability that the second die rolled will show a larger number than the first, given that the first die rolled is a 4.

A

B

$\{(4,5), (4,6)\} \rightarrow P(A,B) = \frac{2}{36}$

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{2/36}{6/36} = \frac{1}{3}$$

$\{(4,1), (4,2), (4,3), \ldots, (4,6)\} \rightarrow P(B) = 6/36$

Exercise 9.4 A sequence of two playing cards is drawn at random (without replacement) from a standard deck of 52 cards. What is the probability that the first card is red and the second is black.

A

B

$$P(A \wedge B) = \underbrace{P(B|A)}_{\text{the 2nd card is black given that the 1st card is red}} \underbrace{P(A)}_{\text{the 1st card is red.}} = \frac{26}{51} \times \frac{26}{52} = 0.255$$

Exercise 9.5 If 60% of a department store's customers are female and 75% of the female customers have charge accounts at the store, what is the probability that a customer selected at random is a female and has a charge account?

a = customer is a female $\quad P(a) = 0.60$

b = customer has charge accounts $\quad P(b|a) = 0.75$

$$P(a \wedge b) = P(b|a)P(a) = 0.75 \times 0.60 = 0.45$$

9.5 Independence

If a and b are any events, we say that a and b are **independent** if and only if

$$P(a \wedge b) = P(a)P(b)$$

otherwise, a and b are said to be dependent.
Two events are independent if

$$P(a \mid b) = P(a) \qquad \text{or} \qquad P(b \mid a) = P(b)$$

Exercise 9.6 Suppose that we draw two consecutive cards from a 52-card deck. Let A and B be the events

A = "second card is black"
B = "first card is red."

Check if $P(A \wedge B) = P(A)P(B)$?

Are these events independent?

$P(B) = \frac{26}{52}$ $\quad$ $P(A) = P(A \wedge B) + P(A \wedge \neg B)$
$= \left(\frac{26}{52} \times \frac{26}{51}\right) + \left(\frac{26}{52} \times \frac{25}{51}\right)$

$P(A \wedge B) \neq P(A)P(B)$ $\quad$ A and B are NOT independent.

Exercise 9.7 Suppose that we toss a coin twice and record the sequence of heads and tails. Let A and B be the events

A = "head on the second toss"
B = "head on the first toss"

(H, H) (H, T) (T, H) (T, T)

Are these events independent?

$P(A) = \frac{2}{4}$ $\quad$ $P(B) = \frac{2}{4}$ $\quad$ $P(A, B) = \frac{1}{4}$

$P(A)P(B) = P(A \wedge B)$ $\quad$ A and B are independent.

Exercise 9.8 A vacation resort consists of 500 apartments with sizes and vies as indicated below. Assume that all apartments have the same likelihood of being the next one offered for sale.

	One Bedroom	Two Bedroom
Sea view	240	60
No sea view	160	40

1. What is the probability that the next apartment for sale has a sea view?

$$P(\text{seaview}) = \frac{240+60}{500} = 0.6$$

2. If the next apartment offered for sale has two bedrooms, what is the probability that it has a sea view?

$$P(\text{seaview} \mid \text{twobedrooms}) = \frac{60}{100} = 0.6$$

$$= \frac{P(s \wedge t)}{P(t)} = \frac{60/500}{100/500}$$

3. Are the event (has two bedroom) and (sea view) independent?

$$P(\text{twobedrooms}) = \frac{100}{500} = 0.2$$

$$P(\text{seaview} \wedge \text{two bedrooms}) = \frac{60}{500} = 0.12$$

$$=$$

$$P(\text{seaview})P(\text{two bedrooms}) = (0.6)(0.2) = 0.12$$

These two events are independent.

Exercise 9.9 A hospital uses two tests to classify blood. Every blood sample is subjected to both tests. The first test correctly identifies blood type with probability 0.7, and the second test correctly identifies blood type with probability 0.8. The probability that at least one of the tests correctly identifies the blood type is 0.9.

1. Find the probability that both tests correctly identify the blood type.

2. Determine the probability that the second test is correct given that the first test is correct.

3. Determine the probability that the first test is correct given that the second test is correct.

4. Are the events "test I correctly identifies the blood type" and "test II correctly identifies the blood type" independent?

Independence allows us to dramatically reduce the number of probability values required for probabilistic inference.

Exercise 9.10 How many entries are required to store a full joint distribution
$\mathbf{P}(\mathit{Toothache}, \mathit{Catch}, \mathit{Cavity}, \mathit{Weather})$?

Since the weather should not influence one's dental problems, and the dental problem should not influence the weather, we have

$$\mathbf{P}(\mathit{Toothache}, \mathit{Catch}, \mathit{Cavity}, \mathit{Weather}) = \mathbf{P}(\mathit{Toothache}, \mathit{Catch}, \mathit{Cavity})\mathbf{P}(\mathit{Weather})$$

Exercise 9.11 How many entries are required the decomposed distribution?

Credit

This lecture note is partially adopted from Asst. Prof. Dr. Nirattaya Khamsemanan's GTS111 partial note.