COSC1125/1127 Artificial Intelligence

School of Computer Science and IT
RMIT University
Semester 1, 2019

Tutorial Sheet 10
Probability Reasoning

We use $P(\cdot)$ (sometimes also written $\Pr(\cdot)$) to refer to a probability function or probability distribution. When using an upper letter (e.g., X or $Cavity$), we refer to the random variable; when using lowercases we refer to a specific value of the corresponding random variable. So, $P(Cavity)$ is a probability distribution over variable $Cavity$; whereas $P(cavity)$ is a shorthand for $P(Cavity = true)$ and $P(\neg cavity)$ is a shorthand for $P(Cavity = false)$.

1. Prove, formally, that $P(A \mid B \wedge A) = 1$.

Answer

You need to use the definition of conditional probability, $P(X \mid Y) = P(X \wedge Y) \mid P(Y)$, and the definitions of the logical connectives. It is not enough to say that if $B \wedge A$ is "given", then A must be true. From the definition of conditional probability and the fact that $A \wedge A \iff A$ and that conjunction is commutative and associative we have,

$$P(A \mid B \wedge A) = \frac{P(A \wedge (B \wedge A))}{P(B \wedge A)} = \frac{P(B \wedge A)}{P(B \wedge A)} = 1$$

2. Consider the domain of dealing 5-card poker hands from a standard deck of 52 cards, under the assumption that the dealer is fair.

   (a) How many atomic events are there in the joint probability distribution (i.e., how many 5-card hands are there)?

   (b) What is the probability of each atomic event?

   (c) What is the probability of being dealt a royal straight flush (the ace, king, queen, jack and ten of the same suit)?

   (d) What is the probability of being dealt four-of-a-kind (i.e., four cards of different suit but same face value)?

   (e) You are told that the probability drawing two cards from a deck of 52 and them both being the same face value is $\frac{1}{221}$. You take the deck and draw the first card — it is the ace of spades! What is the probability that the second card you draw will also be an ace? (even though this is easy to work out using binomials, use conditional probabilty for this question)

Answer

This is a classic combinatorics question. The point here is to refer to the relevant axioms of probability, principally, the following axioms:

(1) All probabilities are between 0 and 1. $0 \leq P(A) \leq 1$

(2) $P(True) = 1$ and $P(False) = 0$

(3) The probability of a disjuction is given by $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

The question also helps students to grasp the concept of joint probability distribution over all possible states of the world.

(a) It is important to note here that the hand $\{\clubsuit 2, \diamondsuit 3, \heartsuit 4, \diamondsuit 5, \spadesuit 6\}$, is identical to the hand $\{\spadesuit 6, \diamondsuit 5, \heartsuit 4, \diamondsuit 3, \clubsuit 2\}$. If the order of the cards dealt mattered, then the number of hands would simply be $\frac{52!}{47!}$, or $52 \times 51 \times 50 \times 49 \times 48$, because there are 52 choices for the first card, 51 for the second, 50 for the third, and so on.

Generally though, when dealing a hand of cards, it doesn't matter the order in which they are dealt. So, you need to divide this number by the number of possible permutations for a hand of 5 cards, which is $5 \times 4 \times 3 \times 2 \times 1 = 5!$. This means that the total number of 5 card hands that are possible in a 52 card deck is $\frac{52!}{47! \times 5!} = 2,598,960$.

In combinatorics this is written as the binomial coefficient $\binom{52}{5}$, and means: "out of 52 cards, choose 5". In general, $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

(b) By the fair-dealing assumption, each of these is equally likely. Each hand therefore occurs with probability $\frac{1}{2{,}598{,}960}$

(c) There are four hands that are royal straight flushes (one in each suit). By axiom 3, since the events are mutually exclusive, the probability of a royal straight flush is just the sum of the probabilities of the atomic events, i.e., $\frac{4}{2{,}598{,}960} = \frac{1}{649{,}740}$.

(d) Again, we examine the atomic events that are four-of-a-kind events. There are 13 possible kinds and for each, the fifth card can be one of 48 possible other cards. The total probability is therefore $\frac{13 \times 48}{2{,}598{,}960} = \frac{1}{4{,}165}$.

(e) Even though we are given the suit of the first ace, that information is immaterial as the probability you are given does not specify suit, so we can ignore it. If we let $P(A)$ be the probability of drawing the first ace, its value is simply $\frac{4}{52}$. Let $P(A \wedge B)$ be the probability of drawing two aces in a row, which is $\frac{1}{221}$. We can work out the probability of drawing a second ace, *given* we have already drawn an ace, $P(B \mid A)$, using the definition of conditional probability.

$$P(B \mid A) = \frac{P(A \wedge B)}{P(A)} = \frac{\frac{1}{221}}{\frac{4}{52}} = \frac{52}{4 \times 221} = \frac{3}{51}$$

3. Given the full joint distribution shown in the table below, calculate the following:

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	.108	.012	.072	.008
$\neg cavity$	.016	.064	.144	.576

(a) $P(toothache)$

(b) $P(Cavity)$

(c) $P(Toothache \mid cavity)$

(d) $P(Cavity \mid toothache \vee catch)$.

Answer

Note that $P(Cavity)$ denotes a vector of values for the probabilities of each individual state of Cavity. Also note here we use the uppercase (e.g., $Cavity$) to denote a variable, whereas lowercase (e.g., $cavity$) to denote a constant.

(a) $P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

(b) $P(Cavity) = \langle 0.108 + 0.012 + 0.072 + 0.008, 0.016 + 0.064 + 0.144 + 0.576 \rangle = \langle 0.2, 0.8 \rangle$

(c) $P(Toothache \mid cavity) = \langle \frac{0.108+0.012}{0.108+0.012+0.072+0.008}, \frac{0.072+0.008}{0.108+0.012+0.072+0.008} \rangle = \langle 0.6, 0.4 \rangle$

(d) $P(Cavity \mid toothache \vee catch)$
$= \langle \frac{0.108+0.012+0.072}{0.108+0.012+0.072+0.016+0.064+0.144}, \frac{0.016+0.064+0.144}{0.108+0.012+0.072+0.016+0.064+0.144} \rangle$
$\approx \langle 0.462, 0.538 \rangle$

4. After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for serious disease and that the test is 99% accurate (i.e., the probability of testing positive when you do have the disease is 0.99, as is the probability of testing negative when you don't have the disease). The good news is that this is a rare disease, striking only 1 in 10,000 people of your age. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

Answer

We are given the following information:

$$
\begin{aligned}
&P(test \mid disease) = 0.99 \\
&P(\neg test \mid \neg disease) = 0.99 \\
&P(disease) = 0.0001 \\
&test
\end{aligned}
$$

where $test$ means that the test is positive. What the patient is concerned about is $P(disease \mid test)$. Roughly speaking, the reason it is a good thing that the disease is rare is that $P(disease \mid test)$ is proportional to $P(disease)$, so a lower prior probability for $Disease$ will mean a lower value for $P(disease \mid test)$. By and large, if 10,000 people take the test, we expect 1 to actually have the disease, and most likely test positive, while the rest do not have the disease, but 1% of them (about 100 people) will test positive anyway, so $P(disease \mid test)$ will be about 1 in 100. More precisely, using the following:

$$
\begin{aligned}
P(disease \mid test) &= \frac{P(test \mid disease)P(disease)}{P(test)} \\
&= \frac{P(test \mid disease)P(disease)}{P(test \mid disease)P(disease) + P(test \mid \neg disease)P(\neg disease)} \\
&= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.01 \times 0.9999} \\
&= 0.009804
\end{aligned}
$$

Note that in the above, to calculate $P(test)$, we need to sum over all other hidden variables, but in this case there is only one, that is $Disease$:

$$P(test) = P(test \wedge disease) + P(test \wedge \neg disease)$$

by the product rule, we have:

$$
\begin{aligned}
P(test \wedge disease) &= P(test \mid disease)P(disease) \\
P(test \wedge \neg disease) &= P(test \mid \neg disease)P(\neg disease)
\end{aligned}
$$

The moral is that when the disease is much rarer than the test accuracy, a positive result does not mean the disease is likely. A false positive reading remains much more likely.

5. Prove, formally, that $P(A \wedge B \wedge C) = P(A \mid B \wedge C) \times P(B \mid C) \times P(C)$.

Answer

Rearranging the definition of conditional probability,

$$P(X \mid Y) = \frac{P(X \wedge Y)}{P(Y)}$$

gives the product rule:

$$P(X \wedge Y) = P(X \mid Y) \times P(Y)$$

Using the product rule and substituting X/A and $Y/[B \wedge C]$ into $P(A \wedge B \wedge C)$, we have,

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

Finally, applying the product rule again to $P(B \wedge C)$ and substituting X/B and Y/C gives:

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

6. Prove Bayes' Theorem: $P(A \mid B) = \dfrac{P(B \mid A) \times P(A)}{P(B)}$.

Answer

The probability of two events A and B happening, $P(A \wedge B)$, is the probablity of A, $P(A)$, times the probability of B given that B has occurred, $P(B \mid A)$.

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

On the other hand, the probability of A and B is also equal to the probability of B times the probability of A given B.

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

Equating the two yields:

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

and thus,

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