Tutorial 2 - Probability and Inference

COMP9418 – Advanced Topics in Statistical Machine Learning

Lecturer: Gustavo Batista

Lecture: Propositional Logic and Probability Calculus

Topic: Questions from lecture topics

Last revision: Tuesday 22nd September, 2020 at 20:33

Question 1

(From the textbook) For each of the following pairs of sentences, decide whether the first sentence implies the second. If the implication does not hold, identify a world at which the first sentence is true, but the second is not.

a. $(A \implies B) \wedge \neg B$ and A.
b. $(A \vee \neg B) \wedge B$ and A.
c. $(A \vee B) \wedge (A \vee \neg B)$ and A.

Answer

Let us build a table with all possible worlds for each item.

a. $(A \implies B) \wedge \neg B$ and A

A	B	$\neg B$	$A \implies B$	$(A \implies B) \wedge \neg B$	$(A \implies B) \wedge \neg B \models A$?
1	1	0	1	0	Yes
1	0	1	0	0	Yes
0	1	0	1	0	Yes
0	0	1	1	1	No

b. $(A \vee \neg B) \wedge B$ and A

A	B	$\neg B$	$A \vee \neg B$	$(A \vee \neg B) \wedge B$	$(A \vee \neg B) \wedge \neg B \models A$?
1	1	0	1	1	Yes
1	0	1	1	0	Yes
0	1	0	0	0	Yes
0	0	1	1	0	Yes

c. $(A \vee B) \wedge (A \vee \neg B)$ and A

A	B	$\neg B$	$A \vee B$	$A \vee \neg B$	$(A \vee B) \wedge (A \vee \neg B)$	$(A \vee \neg B) \wedge \neg B \models A$?
1	1	0	1	1	1	Yes

A	B	$\neg B$	$A \vee B$	$A \vee \neg B$	$(A \vee B) \wedge (A \vee \neg B)$	$(A \vee \neg B) \wedge \neg B \models A$?
1	0	1	1	1	1	Yes
0	1	0	1	0	0	Yes
0	0	1	0	1	0	Yes

Question 2

Which of the following pairs of sentences are mutually exclusive? Which are exhaustive? If a pair of sentences is not mutually exclusive, identify a world at which they both hold. If a pair of sentences is not exhaustive, identify a world at which neither holds.

a. $A \vee B$ and $\neg A \vee \neg B$.
b. $A \vee B$ and $\neg A \wedge \neg B$.
c. A and $(\neg A \vee B) \wedge (\neg A \vee \neg B)$.

Answer

a. $A \vee B$ and $\neg A \vee \neg B$ are not mutually exclusive. The worlds which both hold are marked with "yes" in column "Both". These sentences are exhaustive.

A	B	$\neg A$	$\neg B$	$A \vee B$	$\neg A \vee \neg B$	Both
1	1	0	0	1	0	No
1	0	0	1	1	1	Yes
0	1	1	0	1	1	Yes
0	0	1	1	0	1	No

b. $A \vee B$ and $\neg A \wedge \neg B$ are mutually exclusive and exhaustive.

A	B	$\neg A$	$\neg B$	$A \vee B$	$\neg A \wedge \neg B$
1	1	0	0	1	0
1	0	0	1	1	0
0	1	1	0	1	0
0	0	1	1	0	1

c. A and $(\neg A \vee B) \wedge (\neg A \vee \neg B)$ are mutually exclusive and exhaustive.

A	B	$\neg A$	$\neg B$	$\neg A \vee B$	$\neg A \vee \neg B$	$(\neg A \vee B) \wedge (\neg A \vee \neg B)$
1	1	0	0	1	0	0
1	0	0	1	0	1	0
0	1	1	0	1	1	1
0	0	1	1	1	1	1

Question 3

Suppose that 24% of a population are smokers and that 5% of the population have cancer. Suppose further that 86% of the population with cancer are also smokers. What is the probability that a smoker will also have cancer?

Answer

Let us take note of the given probabilities. We will use variable S to denote smokers and C for people with cancer.

$P(+s) = .24$

$P(+c) = .05$

$P(+s|+c) = .86$

The exercise asks for $P(+c|+s)$. Therefore, we will need to invert the conditional probabilities given. It is a direct application of the Bayes rule.

$P(+c|+s) = \frac{P(+s|+c)p(+c)}{p(+s)} = \frac{.86 \times .05}{.24} = .1792$

Therefore it is expected that 17.91% of the smokers also have cancer.

Question 4

(After Koller) Suppose that a tuberculosis (TB) skin test is 95 per cent accurate. That is, if the patient is TB-infected, then the test will be positive with probability 0.95, and if the patient is not infected, then the test will be negative with probability 0.95. Now suppose that a person gets a positive test result. What is the probability that he is infected? Suppose that 1 in 1000 of the subjects who get tested is infected.

To answer this question, provide the following intermediate quantities:

$P(TB = +) =$

$P(Test = +|TB = +) =$

$P(Test = +|TB = -) =$

$P(Test = +) =$

Which equation provides a direct answer to the question posed in this problem?

Answer

From the problem statement, we get:

$P(TB = +) = 0.001$

$P(Test = +|TB = +) = 0.95$

$P(Test = +|TB = -) = 0.05$

For the next one, we can first use the sum rule in the following way:

$P(Test = +) = P(Test = +, TB = +) + P(Test = +, TB = -)$

As we do not have the joint probabilities, we need to calculate them from the conditional probabilities using the product rule:

$P(Test = +, TB = +) = P(Test = +|TB = +)P(TB = +) = 0.95 \times 0.001 = 0.00095$

$P(Test = +, TB = -) = P(Test = +|TB = -)P(TB = -) = 0.05 \times 0.999 = 0.04995$

Therefore,

$P(Test = +) = 0.00095 + 0.04995 = 0.0509$

The problem asks for $P(TB = +|Test = +)$. Therefore, we will need the Bayes Rule to invert $P(Test = +|TB = +)$

$P(TB = +|Test = +) = \frac{P(Test=+|TB=+)P(TB=+)}{P(Test=+)} = \frac{0.95 \times 0.001}{0.0509} \approx 0.0187$

Thus, although a subject with a positive test is much more probable to be TB-infected than is a random subject, fewer than 2 per cent of these subjects are TB-infected.

Question 5

Consider the following distribution over three variables:

A	B	C	$P(A, B, C)$
1	1	1	.27
1	1	0	.18
1	0	1	.03
1	0	0	.02
0	1	1	.02
0	1	0	.03
0	0	1	.18
0	0	0	.27

For each pair of variables, state whether they are independent. State also whether they are independent given the third variable. Justify your answer.

Answer

Let us compute the marginals $P(A, B)$, $P(A, C)$ and $P(B, C)$.

A	B	$P(A, B)$
1	1	.45
1	0	.05
0	1	.05
0	0	.45

A	C	$P(A, C)$
1	1	.30
1	0	.20
0	1	.20
0	0	.30

B	C	$P(B, C)$
1	1	.29
1	0	.21
0	1	.21
0	0	.29

Also, let us compute the priors $P(A)$, $P(B)$ and $P(C)$.

A	$P(A)$
1	.50

A	$P(A)$
0	.50

B	$P(B)$
1	.50
0	.50

C	$P(C)$
1	.50
0	.50

From these probability distributions, we observe that no pair of variables is independent since $P(X,Y) \neq P(X)P(Y)$. Now, let us use these tables to compute the conditional probabilities using the equation $P(X,Y|Z) = \frac{P(X,Y,Z)}{P(Z)}$.

C	A	B	$P(A,B|C)$
1	1	1	.54
1	1	0	.06
1	0	1	.04
1	0	0	.36
0	1	1	.36
0	1	0	.04
0	0	1	.06
0	0	0	.54

Also, use the equation $P(X|Y) = \frac{P(X,Y)}{P(Y)}$ to compute the following conditional probabilities.

A	C	$P(A|C)$
1	1	.60
1	0	.40
0	1	.40
0	0	.60

C	B	$P(B|C)$
1	1	.58
1	0	.42
0	1	.42
0	0	.58

From these probability tables, we observe that A is not independent of B given C since $P(A,B|C) \neq P(A|C)P(B|C)$.

We can use the same approach to conditional independence of B and C given A. Let us first compute

$P(B, C|A)$.

A	B	C	$P(B, C\|A)$
1	1	1	.54
1	1	0	.36
1	0	1	.06
1	0	0	.04
0	1	1	.04
0	1	0	.06
0	0	1	.36
0	0	0	.54

Also, we need to compute $P(B|A)$ and $P(C|A)$.

B	A	$P(B\|A)$
1	1	.90
1	0	.10
0	1	.10
0	0	.90

C	A	$P(C\|A)$
1	1	.60
1	0	.40
0	1	.40
0	0	.60

From these probability tables, we observe that $B \perp\!\!\!\perp C|A$ since $P(B, C|A) = P(B|A)P(C|A)$.

Finally, we test if $A \perp\!\!\!\perp B|C$. We start computing $P(A, C|B)$.

B	A	C	$P(A, C\|B)$
1	1	1	.54
1	1	0	.36
1	0	1	.04
1	0	0	.06
0	1	1	.06
0	1	0	.04
0	0	1	.36
0	0	0	.54

Also, we need to compute $P(A|B)$ and $P(C|B)$.

B	A	$P(A\|B)$
1	1	.90
1	0	.10
0	1	.10
0	0	.90

B	C	$P(C\|B)$
1	1	.58
1	0	.42
0	1	.42
0	0	.58

From these probability tables, we observe that A is not independent of B given C since $P(A, C|B) = P(A|B)P(C|B)$.

Question 6

We have three binary random variables: season (S), temperature (T) and weather (W). Let us suppose you are given the following conditional probability distributions (CPDs):

S	$P(S)$
summer	0.5
winter	0.5

S	T	$P(T\|S)$
summer	hot	0.7
summer	cold	0.3
winter	hot	0.3
winter	cold	0.7

S	T	W	$P(W\|T, S)$
summer	hot	sun	0.86
summer	hot	rain	0.14
summer	cold	sun	0.67
summer	cold	rain	0.33
winter	hot	sun	0.67
winter	hot	rain	0.33
winter	cold	sun	0.43
winter	cold	rain	0.57

Calculate the joint probability distribution $P(T, S, W)$ using the chain rule.

S	T	W	$P(T, S, W)$
summer	hot	sun	
summer	hot	rain	
summer	cold	sun	
summer	cold	rain	
winter	hot	sun	
winter	hot	rain	
winter	cold	sun	
winter	cold	rain	

Answer

To solve this one, we need to remember that the chain rule has $n!$ possibilities. We need to choose one that matches the information at hand. In this case, we need to obtain $P(S)$, $P(T|S)$ and $P(T, S, W)$ from $P(T, S, W)$. Therefore, we choose the following one:

$P(T, S, W) = P(S)P(T|S)P(W|T, S)$

Now, filling in the table is similar to a join operation in databases. We multiply the matching rows for values of T, S and W.

S	T	W	$P(T, S, W)$
summer	hot	sun	$0.5 \times 0.7 \times 0.86 = 0.301$
summer	hot	rain	$0.5 \times 0.7 \times 0.14 = 0.049$
summer	cold	sun	$0.5 \times 0.3 \times 0.67 = 0.1005$
summer	cold	rain	$0.5 \times 0.3 \times 0.33 = 0.0495$
winter	hot	sun	$0.5 \times 0.3 \times 0.67 = 0.1005$
winter	hot	rain	$0.5 \times 0.3 \times 0.33 = 0.0495$
winter	cold	sun	$0.5 \times 0.7 \times 0.43 = 0.1505$
winter	cold	rain	$0.5 \times 0.7 \times 0.57 = 0.1995$

Question 7

(From Ben Lambert's book "A Student's Guide to Bayesian Statistics") Suppose that, in an idealised world, the ultimate fate of a thrown coin - head or tails - is deterministically given by the angle at which you throw the coin and its height above the table. Also, in this ideal world, the heights and angles are discrete. However, the system is chaotic[1] (highly sensitive to initial conditions), and the results of throwing a coin at a given angle (in degrees) and height (in meters) are shown in the next table.

Angle (degree)	0.2	0.4	0.6	0.8	1
0	T	H	T	T	H
45	H	T	T	T	T
90	H	H	T	T	H
135	H	H	T	H	T
180	H	H	T	H	H
225	H	T	H	T	T
270	H	T	T	T	H
315	T	H	H	T	T

a. Suppose that all combinations of angles and heights are equally likely to be chosen. What is the probability that the coin lands heads up?

b. Now suppose that some combinations of angles and heights are more likely to be chosen than others, with the probabilities shown in the next table. What are the new probabilities that the coin lands heads up?

Angle (degree)	0.2	0.4	0.6	0.8	1
0	0.05	0.03	0.02	0.04	0.04

[1]The authors of the following paper experimentally tested this and found it to be the case, "The three-dimensional dynamics of the die throw", Chaos, Kapitaniak et al. (2012).

Angle (degree)	0.2	0.4	0.6	0.8	1
45	0.03	0.02	0.01	0.05	0.02
90	0.05	0.03	0.01	0.03	0.02
135	0.02	0.03	0.04	0.00	0.04
180	0.03	0.02	0.02	0.00	0.03
225	0.00	0.01	0.04	0.03	0.02
270	0.03	0.00	0.03	0.01	0.04
315	0.02	0.03	0.03	0.02	0.01

c. We force the coin thrower to throw the coin at an angle of 45 degrees. What is the probability that the coin lands heads up?

d. We force the coin-thrower to throw the coin at a height of 0.2m. What is the probability that the coin lands heads up?

e. If we constrained the angle and height to be fixed, what would happen in repetitions of the same experiment?

Answer

a. A count of heads in the first table results in $P(H) = \frac{19}{40}$.

b. A weighted average where the weights are provided by the second Table results in $P(H) = 0.5$.

c. A weighted average given an angle of 45 degrees results in $P(H|angle = 45) \approx 0.23$.

d. Similar to the previous question, but constrained to a height of 0.2, we have $P(H|height = 0.2) \approx 0.70$.

e. If both height and angle are fixed, the outcome becomes deterministic.

Question 8

(After Koller) An often useful rule in dealing with probabilities is known as reasoning by cases. Let X, Y , and Z be random variables, then $P(X|Y) = \sum_z P(X, z|Y)$.

Prove this equality using the basic properties of (conditional) distributions.

Answer

Let us use the Bayes conditioning to prove this property. We will also need to know that $\sum_z P(z|\mathbf{W}) = 1$ for each instantiation $\mathbf{w} \in \mathbf{W}$.

$$\begin{aligned}
\sum_z P(X, z|Y) &= \sum_z \frac{P(X, z, Y)}{P(Y)} \\
&= \frac{\sum_z P(X, z, Y)}{P(Y)} \\
&= \frac{\sum_z P(z|X, Y)P(X, Y)}{P(Y)} \\
&= \frac{P(X, Y)\sum_z P(z|X, Y)}{P(Y)} \\
&= \frac{P(X, Y)}{P(Y)} \\
&= P(X|Y)
\end{aligned}$$