CS 188
Spring 2021

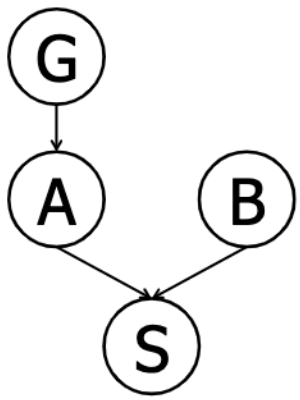
Regular Discussion 5 Solutions

1 Probability

Suppose that a patient can have a symptom (S) that can be caused by two different, independent diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. A model and some conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

$P(G)$	
$+g$	0.1
$-g$	0.9

$P(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$P(B)$	
$+b$	0.4
$-b$	0.6

$P(S \mid A, B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

(a) Compute the following entry from the joint distribution:

$P(+g, +a, +b, +s) =$

$P(+g)P(+a|+g)P(+b)P(+s|+b,+a) = (0.1)(1.0)(0.4)(1.0) = 0.04$

(b) What is the probability that a patient has disease A?

$P(+a) =$ $P(+a|+g)P(+g) + P(+a|-g)P(-g) = (1.0)(0.1) + (0.1)(0.9) = 0.19$

(c) What is the probability that a patient has disease A given that they have disease B?

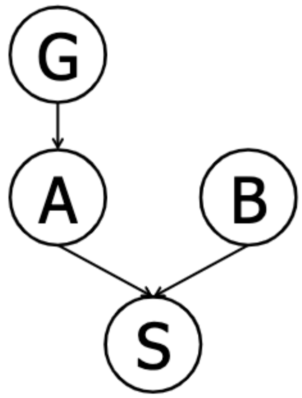
$P(+a|+b) =$ $P(+a) = 0.19$

This equality holds true as we have $A \perp\!\!\!\perp B$ which is given in the problem. You can also infer this from the graph of the Bayes' net via d-separation, which is covered in future lectures.

The figures and table below are identical to the ones on the previous page and are repeated here for your convenience.

$P(G)$	
$+g$	0.1
$-g$	0.9

$P(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$P(B)$	
$+b$	0.4
$-b$	0.6

$P(S \mid A, B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

(d) What is the probability that a patient has disease A given that they have symptom S and disease B?

$$P(+a|+s,+b) = \frac{P(+a,+b,+s)}{P(+a,+b,+s)+P(-a,+b,+s)} = \frac{P(+a)P(+b)P(+s|+a,+b)}{P(+a)P(+b)P(+s|+a,+b)+P(-a)P(+b)P(+s|-a,+b)}$$
$$= \frac{(0.19)(0.4)(1.0)}{(0.19)(0.4)(1.0)+(0.81)(0.4)(0.8)} = \frac{0.076}{0.076+0.2592} \approx 0.2267$$

(e) What is the probability that a patient has the disease carrying gene variation G given that they have disease A?

$$P(+g|+a) = \frac{P(+g)P(+a|+g)}{P(+g)P(+a|+g)+P(-g)P(+a|-g)} = \frac{(0.1)(1.0)}{(0.1)(1.0)+(0.9)(0.1)} = \frac{0.1}{0.1+0.09} = 0.5263$$

2 Independence

1. Suppose you have two random variables, C and N. C is the result of flipping a biased coin that lands on heads (h) with probability 0.8 and tails (t) with probability 0.2. N is the number of heads that result from two independent coin flips of a fair coin.

   (a) Fill in the probability tables for $P(C)$, $P(N)$, and $P(C, N)$.

C	$P(C)$
h	0.8
t	0.2

N	$P(N)$
0	0.25
1	0.5
2	0.25

C	N	$P(C,N)$
h	0	0.2
h	1	0.4
h	2	0.2
t	0	0.05
t	1	0.1
t	2	0.05

   Since the flips of the fair coin and the biased coin are independent from each other, we can compute joint probabilities as the product of the probabilities of the two independent events.

   (b) Using the probability tables above, what is $P(N = 1|C = t)$?

   Because the events N and C are independent, this is just $P(N = 1) = 0.5$, reading directly off of the table.

2. Simplify each of the following into a single probability expression using the given independence assumption.

   (a) Given that $A \perp\!\!\!\perp B$, simplify $\sum_a P(a|B)P(C|a)$.

   If $A \perp\!\!\!\perp B$, then $P(A|B) = P(A)$. Therefore, the expression is just $\sum_a P(a)P(C|a) = P(C)$ by applying Product rule and the Law of Total Probability.

   (b) Given that $B \perp\!\!\!\perp C|A$, simplify $\frac{P(A)P(B|A)P(C|A)}{P(B|C)P(C)}$.

   From our independence assumption, we know that $P(C|A, B) = P(C|A)$, since when conditioned on A, knowing B gives no more information about C. Applying this, we can see that both the numerator and denominator are applications of the Chain Rule, and simplify to get $\frac{P(A,B,C)}{P(B,C)}$, which by the definition of conditional probability is just $P(A|B, C)$.

   (c) Given that $A \perp\!\!\!\perp B|C$, simplify $\frac{P(C,A|B)P(B)}{P(C)}$.

   By Product Rule,the numerator is simplified to get $\frac{P(A,B,C)}{P(C)}$, which by the definition of conditional probability is just $P(A, B|C)$. Our independence assumption allows us to observe that $P(A, B|C) = P(A|C)P(B|C)$, but this doesn't actually let us simplify any further.

3. Mark **all** expressions that are equal to $P(R, S, T)$, given no independence assumptions:

   - ■ $P(R \mid S, T)\ P(S \mid T)\ P(T)$
   - ☐ $P(T \mid R, S)\ P(R)\ P(S)$
   - ☐ $P(R \mid S)\ P(S \mid T)\ P(T)$
   - ☐ $P(R \mid S, T)\ P(S \mid R, T)\ P(T \mid R, S)$
   - ■ $P(T, S \mid R)\ P(R)$
   - ■ $P(T \mid R, S)\ P(R, S)$
   - ☐ None of the above