

Solutions to Exercises for Module 4 – Bayesian Networks

Exercise 1: Mr. Dupont is a professional wine taster. When given a French wine, he will identify it 90% of the time correctly as French, and will mistake it for a Californian wine 10% of the time. When given a Californian wine, he will identify it 80% of the time correctly as Californian, and will mistake it for a French wine 20% of the time. Suppose that Mr. Dupont is given ten unlabelled glasses of wine, three French and seven Californian. He randomly picks a glass, tries the wine, and says: "French". Build a Bayesian network to determine the probability that the wine he tasted was Californian?

Solution:

If the wine is French then:

- The probability that it will be identified as French = 90%
- The probability that it will be identified as Californian = 10%

If the wine is Californian then:

- The probability that it will be identified as French = 20%
- The probability that it will be identified as Californian = 80%

The prior probability that the wine is French = 30%

The prior probability that the wine is Californian = 70%

The Bayesian network:

Wine	
French	30.0
Californian	70.0

Taste_Test	
French	41.0
Californian	59.0

NPT for Wine node:

French	Californian
30	70

NPT for Taste test node:

Wine	French	Californian
French	90	10
Californian	20	80

Inserting the evidence into the Bayesian network:

Wine	
French	65.9
Californian	34.1

Taste_Test	
French	100
Californian	0

Exercise 2: As accounts manager in your company, you classify 75% of your customers as "good credit" and the rest as "risky credit" depending on their credit rating. Customers in the "risky" category allow their accounts to go overdue 50% of the time on average, whereas those in the "good" category allow their accounts to go overdue only 10% of the time. Build a Bayesian network to determine the percentage of overdue accounts that are held by customers in the "risky credit" category?

Solution:

For customers in the risky category:

- The probability that accounts will be overdue = 50%
- The probability that accounts will not be overdue = 50%

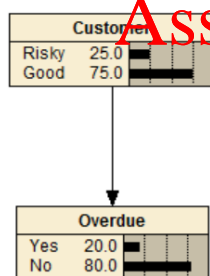
For customers in the good category:

- The probability that accounts will be overdue = 10%
- The probability that accounts will not be overdue = 90%

The prior probability of a risky customer = 25%

The prior probability of a good customer = 75%

The Bayesian network:



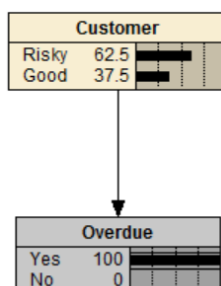
NPT for Customer node:

Risky	Good
25	75

NPT for Overdue node:

Customer	Yes	No
Risky	50	50
Good	10	90

Inserting the evidence into the Bayesian network:



Exercise 3: You are on a jury in a murder trial. After a few days of testimony, you are 80% sure that the defendant is guilty. Then, at the end of the trial, the prosecution presents a new piece of evidence fresh from the lab. The defendant's blood type is found to match that of blood found at the scene of the crime, which could only be the blood of the murderer. The particular blood type occurs in 5% of the population. Build a Bayesian network to determine what your revised probability that the defendant is guilty should be?

Solution:

The null hypothesis is that the defendant is innocent, H = Defendant is innocent

The alternative hypothesis is that the defendant is guilty, not H = Defendant is guilty

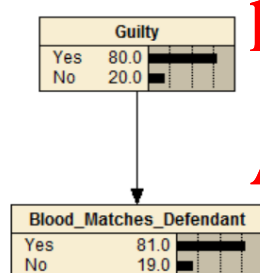
$P(\text{not } H) = 80\%$ because you are 80% sure that the defendant is guilty (i.e. not innocent)

$P(H) = 20\%$ because you are 80% sure that the defendant is guilty, so there is a 20% chance the defendant is innocent

$P(E|H) = 5\%$ because 5% of the population have the same blood type as the defendant. In other words, if the defendant is innocent then there is a 5% chance that someone else with the same blood type as the defendant could have committed the crime.

$P(E|\text{not } H) = 100\%$ because there is 100% chance that the defendant is guilty if the blood found at the scene of the crime is the same type as the defendant.

The Bayesian network:



<https://powcoder.com>

Add WeChat powcoder

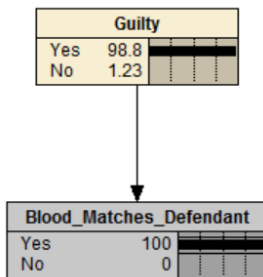
NPT for Guilty node:

Yes	No
80	20

NPT for Blood Matches Defendant node:

Guilty	Yes	No
Yes	100	0
No	5	95

Inserting the evidence into the Bayesian network:



Exercise 4: It is known that there is a single thief in a room of 100 people. A lie detection machine, which is 95% accurate, is used to identify the thief. If a person is selected from the room and the lie detector identifies them as the thief, build a Bayesian network to determine the chance that the correct person has been identified?

Solution:

The prior probability of a person in the room being the thief is 1/100 or 1%

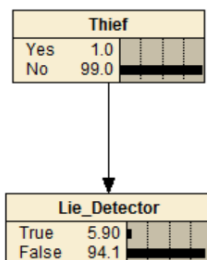
$P(\text{Thief}) = 1\%$

$P(\text{not Thief}) = 99\%$

If a person is the thief:

- The probability that the lie detector will identify them as a thief = 95%
- The probability that the lie detector will not identify them as a thief = 5%

The Bayesian network:



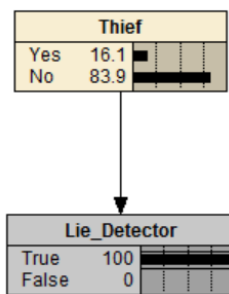
NPT for Thief node:

Yes	No
1	99

NPT for Lie Detector node:

Thief	True	False
Yes	95	5
No	5	95

Inserting the evidence into the Bayesian network:



Exercise 5: Consider the example of a test to determine if a person has a particular disease. First construct the following Bayesian network to predict the probability of a person having the disease if they return a positive test result. 1 in 1000 people within the population have the disease. If a person has the disease then there is a 100% chance that the test will be positive and if the person does not have the disease then there is a 5% chance that the test will be positive.



NPT for Person has disease node:

Yes	No
0.1	99.9

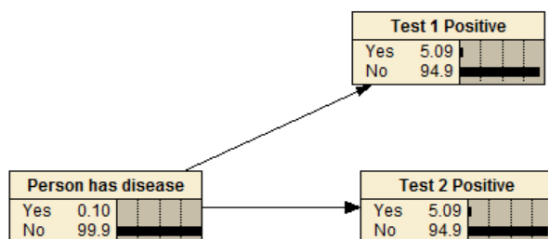
NPT for Test 1 Positive node:

Person has disease	Yes	No
Yes	100	0
No	5	95

Now suppose we run the test on a person twice. Assuming the two tests are independent, construct a new Bayesian network to calculate the probability that a person has the disease if both the first and second test results are positive.

Solution:

The Bayesian network:



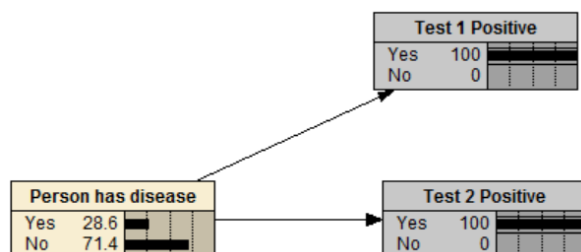
NPT for Person has disease node:

Yes	No
0.1	99.9

NPT for both Test 1 Positive and Test 2 Positive nodes:

Person has disease	Yes	No
Yes	100	0
No	5	95

Inserting the evidence into the Bayesian network:

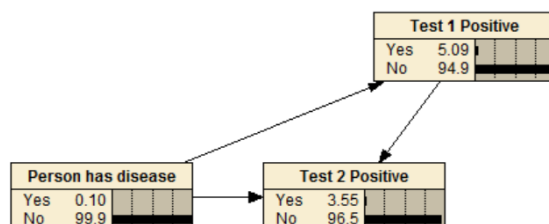


Now assume that the two tests are dependent: If a person has the disease then there is a 100% chance that the second test will be positive. If the person does not have the disease then there is a 50% chance that the second test will be positive if the first test is positive and a 1% chance that the second test will be positive if the first test is negative. Update your Bayesian network to calculate the probability that a person has the disease if both the first and second test results are positive and the tests are dependent.

Solution:

Add WeChat powcoder

The Bayesian network:



NPT for Person has disease node:

Yes	No
0.1	99.9

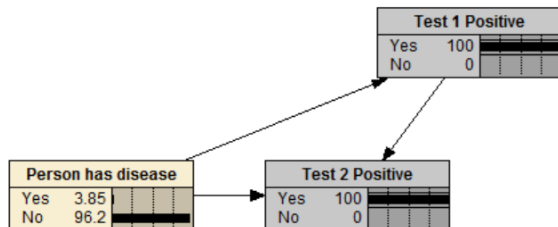
NPT for Test 1 Positive node:

Person has disease	Yes	No
Yes	100	0
No	5	95

NPT for Test 2 Positive node:

Person has disease	Test 1 Positive	Yes	No
Yes	Yes	100	0
Yes	No	100	0
No	Yes	50	50
No	No	1	99

Inserting the evidence into the Bayesian network:

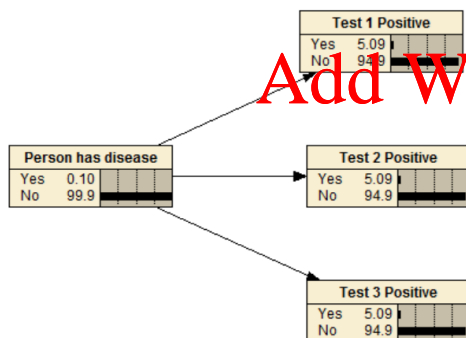


Now extend both of your Bayesian network models (for independent and dependent test) to three tests. Use your models to calculate the probability that a person has the disease if all three test results are positive for the case that the tests are independent and for the case that the tests are dependent. For the case that the tests are dependent, assume that if the person does not have the disease then there is an 80% chance that the third test will be positive if the first and second tests are positive, and a 0.5% chance that the third test will be positive if the first and second tests are negative.

Solution for Independent Tests:

<https://powcoder.com>

The Bayesian network:



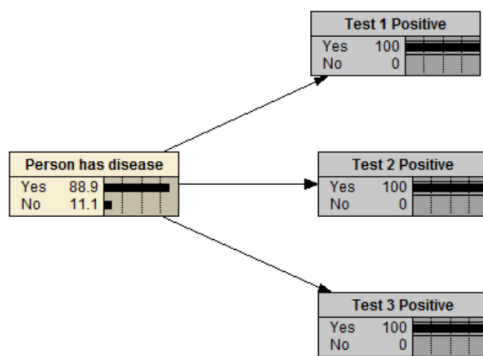
NPT for Person has disease node:

Yes	No
0.1	99.9

NPT for Test 1 Positive, Test 2 Positive and Test 3 Positive nodes:

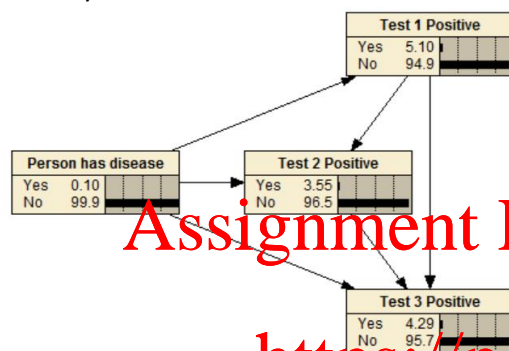
Person has disease	Yes	No
Yes	100	0
No	5	95

Inserting the evidence into the Bayesian network:



Solution for Dependent Tests:

The Bayesian network:



NPT for Person has disease node:

Yes	No
0.1	99.9

NPT for Test 1 Positive node:

Person has disease	Yes	No
Yes	100	0
No	5	95

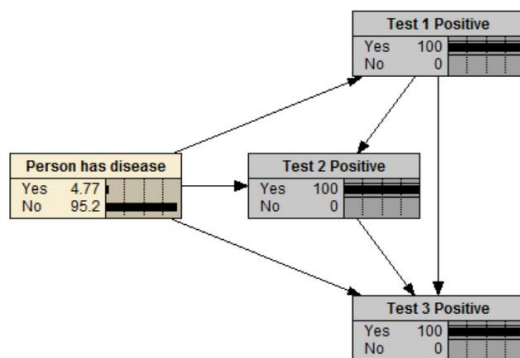
NPT for Test 2 Positive node:

Person has disease	Test 1 Positive	Yes	No
Yes	Yes	100	0
Yes	No	100	0
No	Yes	50	50
No	No	1	99

NPT for Test 3 Positive node:

Person has disease	Test 1 Positive	Test 2 Positive	Yes	No
Yes	Yes	Yes	100	0
Yes	Yes	No	100	0
Yes	No	Yes	100	0
Yes	No	No	100	0
No	Yes	Yes	80	20
No	Yes	No	50	50
No	No	Yes	50	50
No	No	No	0.5	99.5

Inserting the evidence into the Bayesian network:



Exercise 6: In the following “Printer Fault Diagnosis” Bayesian Network, if there is an Application fault, which nodes within the network will have their probabilities updated?

Solution:

The following nodes will have their probabilities updated because they are serially connected to the Application node:

Application Fault
GDI Input
GDI Output
Print Default Out
Printer Output

Assignment Project Exam Help

<https://powcoder.com>

Print to File will also have its probabilities updated because it is serially connected to Print Default Out.

Add WeChat powcoder

Exercise 7: Open the “Chest Clinic” Bayesian network. Note that, with no evidence entered, the probabilities for “Smoker”, “World Travel” and “Lung Cancer” are respectively 50%, 1% and 5.5%

- Enter soft evidence for the node “Smoker”. Specifically, enter the values 0.9 and 0.1 respectively to capture the idea that you are 90% certain the person is a smoker. Now run the model. What are the updated probabilities for: “Smoker”, “World Travel” and “Lung Cancer” respectively?
- Remove the evidence you entered in (i) for ‘Smoker’ and enter soft evidence for the node “World Travel”. Specifically, enter the values 0.9 and 0.1 respectively to capture the idea that you are 90% certain the person has travelled. Run the model. What are the updated probabilities for: “Smoker”, “World Travel” and “Lung Cancer” respectively?
- Explain the apparent discrepancy between the way soft evidence for “Smoker” and “World Travel” was handled.

Solution:

- 90%, 1%, 9.1%
- 50%, 8.33%, 5.5%
- After entering ‘soft evidence’ into a node the posterior probability distribution for that node will not necessarily equal the soft evidence. The posterior probability distribution for the “Smoker” node equals the soft evidence entered into the node because the prior probabilities for the states (yes/no) were equal (50%). However, in the “World Travel” node, the prior probabilities for the

states are not equal; in fact the prior probability of 'yes' is so low that the "90%" soft evidence entered for 'visit' can only shift the posterior probability from 1% to 8.33%. In other words, you have to be very confident to make a major shift away from the prior probabilities. If you enter soft evidence of 99.9% for 'visit' then the posterior probability moves to 99%.

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder