.

CEFET/RJ
Inteligência Artificial (2018.1)
Prof. Eduardo Bezerra (ebezerra@cefet-rj.br)
Lista de exercícios 05

Créditos: essa lista de exercícios contém a tradução dos exercícios retirados do livro texto da disciplina (AIMA).

1. **(Probabilidades marginais e condicionais)** Considere a distribuição de probabilidades conjunta apresentada abaixo, que envolve as variáveis aleatórias binárias *Toothache*, *Catch* e *Cavity*:

	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

Compute o seguinte:

(a) $\Pr(toothache)$,

(b) $\Pr(Cavity)$,

(c) $\Pr(Toothache \mid cavity)$,

(d) $\Pr(Cavity \mid toothache \wedge catch)$.

Solução

- $\Pr(toothache) = 0{,}108 + 0{,}012 + 0{,}016 + 0{,}064$
- $\Pr(Cavity)$

$$\Pr(cavity) = 0{,}108 + 0{,}012 + 0{,}072 + 0{,}008$$

$$\Pr(\neg cavity) = 0{,}016 + 0{,}064 + 0{,}144 + 0{,}576$$

- $\Pr(Toothache \,|\, cavity)$

$$\Pr(toothache \,|\, cavity) = \frac{0{,}108 + 0{,}012}{0{,}108 + 0{,}012 + 0{,}072 + 0{,}008}$$

$$\Pr(\neg toothache \,|\, cavity) = \frac{0{,}072 + 0{,}008}{0{,}108 + 0{,}012 + 0{,}072 + 0{,}008}$$

- $\Pr(Cavity \,|\, toothache \wedge catch)$

$$\Pr(cavity \,|\, toothache \wedge catch) = \frac{0{,}108}{0 + 0{,}108} = 1$$

$$\Pr(\neg cavity \,|\, toothache \wedge catch) = \frac{0}{0 + 0{,}108} = 0$$

2. **(Regra de Bayes)** Suponha que um exame médico é 99% preciso. Suponha também que aplicação desse exame em um paciente dá resultado positivo. A doença detectada pelo exame é rara: apenas 1 em cada 10.000 indivíduos da população tem a doença. Quais são as chances de o paciente ter realmente a doença, dado que o resultado do exame foi positivo?

Solução

Considere os seguintes eventos:

- P = teste dá positivo
- D = indívíduo tem efetivamente a doença

São fornecidas as seguintes informações no enunciado:

- $\Pr(P|D) = 0.99$
- $\Pr(D) = 0.0001$

Das probabilidades acima, podemos concluir que $\Pr(\neg P|D) = 0.01$ e que $\Pr(\neg D) = 0.9999$. É também informado que o teste resultou positivo para o paciente (essa é a evidência).

A preocupação do paciente é acerca do valor $\Pr(D|P)$. Porque a doença é rara, $\Pr(D|P)$ é proporcional a $\Pr(D)$. Portanto, uma probabilidade *a priori* menor de ter a doença significa um valor menor para $\Pr(D)$. De forma aproximada, se 10.000 pessoas fizerem o teste, esperamos que 1 realmente tenha a doença, e provavelmente tenha exame resultando positivo, enquanto que o restante não tenha a doença. Entretanto, 1% deles (cerca de 100 pessoas) testarão positivo de qualquer maneira, então $\Pr(D|P)$ será cerca de 1 em 100.

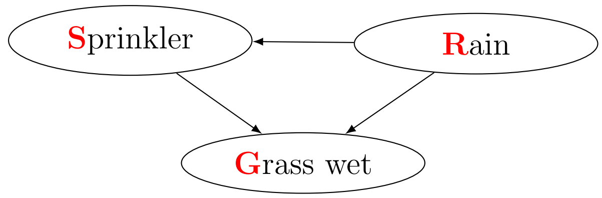
$$\begin{aligned}
\Pr(D|P) &= \frac{\Pr(P|D) \times \Pr(D)}{\Pr(P)} && (1)\\
&= \frac{\Pr(P|D) \times \Pr(D)}{\Pr(P|D) \times \Pr(D) + \Pr(P|\neg D) \times \Pr(\neg D)} && (2)\\
&= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + 0.01 \times 0.9999} && (3)\\
&= 0.009804 && (4)
\end{aligned}$$

A conclusão a que podemos chegar é que, quando a doença é muito mais rara que a exatidão do teste, um resultado positivo no teste não significa que a doença seja provável. Uma leitura falsa positiva continua sendo muito mais provável.

Aqui está um exercício alternativo nas mesmas linhas: um médico diz que crianças que mantêm a cabeça predominantemente para a direita enquanto deita de costas em sua cama será destra, e que crianças que mantêm a cabeça para a esquerda serão canhota. Isabella predominantemente mantém a cabeça para a esquerda. Dado que 90% da população é destra, qual é a probabilidade de Isabella ser destra, se o teste é 90% preciso? E se o teste é 80% preciso? O raciocínio aqui é o mesmo, e a resposta é de 50% de destra se o teste for 90% preciso, e 69% de canhota se o teste for 80% exato.

3. **(Redes Bayesianas)** Considere a rede bayesiana fornecida abaixo. Forneça a distribuição conjunta sobre as três variáveis.

	S	
R	T	F
F	0.4	0.6
T	0.01	0.99



S	
T	F
0.2	0.8

		G	
S		T	F
F	F	0.4	0.6
F	T	0.01	0.99
T	F	0.01	0.99
T	T	0.01	0.99

4. (Redes Bayesianas - inferência) Considerando ainda a rede bayesiana do item anterior. Considere também que a tabela a seguir apresenta uma sequência de números aleatórios, que você pode usar para realizar inferência aproximada sobre a rede e responder a alguma consultas probabilísticas. Considere esses números fornecidos na seguinte ordem: de cima para baixo, e da esquerda para a direita.

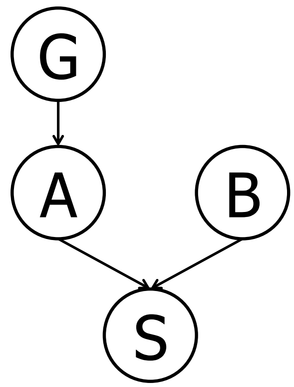
0.02	0.06	0.02	0.84	0.02	0.59
0.33	0.23	0.87	0.49	0.40	0.22
0.33	0.48	0.78	0.42	0.82	0.35

   (a) Quais são as ordenações lineares possíveis na rede bayesiana?
   (b) Usando Amostragem a Priori (*Prior Sampling*), compute aproximações para $\Pr(R)$, $\Pr(S)$, $\Pr(G)$ e $\Pr(G|S = True)$
   (c) Usando Amostragem por Rejeição (*Rejection Sampling*), compute aproximações para $\Pr(R)$, $\Pr(G|S = True)$ e $\Pr(G)$
   (d) Usando Ponderação por Verossimilhança (*Likelihood Weighting*), compute aproximações para $\Pr(R)$, $\Pr(G|S = False)$ e $\Pr(G)$

5. Suponha que um paciente possa ter um sintoma (S) que pode ser causado por duas doenças diferentes (A e B). Sabe-se que a variação do gene G desempenha um grande papel na manifestação da doença A. A rede bayesiana e tabelas de probabilidade condicionais correspondentes para esta situação são mostradas abaixo.
   (a) Compute a seguinte entrada para a distribuição oconjunta: $\Pr(+g, +a, +b, +s) =$
   (b) Qual é a probabilidade de que um paciente tem a doença A? $\Pr(+a) =$
   (c) Qual é a probabilidade de que um paciente tem a doença A dao que el tem a doença B? $\Pr(+a| + b) =$

$\Pr(G)$	
$+g$	0.1
$-g$	0.9

$\Pr(A\|G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9

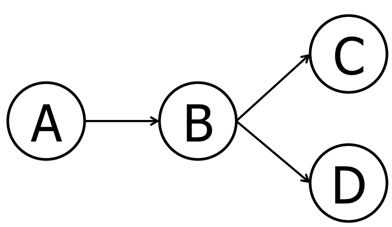


$\Pr(B)$	
$+b$	0.4
$-b$	0.6

$\Pr(S\|A,B)$			
$+a$	$+b$	$+s$	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

6. Considere novamente a rede bayesiana da questão 5.
   (a) Qual é a probabilidade de que um paciente tem a doença A dado que ele tem o sintoma S e a doença B? $\Pr(+a| +s, +b) =$
   (b) Qual é a probabilidade de que um paciente tem a doença de carregar a variante do gene G dado que ele tem a doença A? $\Pr(+g| +a) =$
   (c) Qual é a probabilidade de que um paciente tem a doença de carregar a variante do gene G dado que ele tem a doença B? $\Pr(+g| +b) =$

7. Considere a rede bayesiana abaixo, e as CPTs correspondentes:



A	$\Pr(A)$
$+a$	1/5
$-a$	4/5

A	B	$\Pr(B\|A)$
$+a$	$+b$	1/5
$+a$	$-b$	4/5
$-a$	$+b$	1/2
$-a$	$-b$	1/2

B	C	$\Pr(C\|B)$
$+b$	$+c$	1/4
$+b$	$-c$	3/4
$-b$	$+c$	2/5
$-b$	$-c$	3/5

B	D	$\Pr(D\|B)$
$+b$	$+d$	1/2
$+b$	$-d$	1/2
$-b$	$+d$	4/5
$-b$	$-d$	1/5

   (a) Sua tarefa é estimar $\Pr(+b| -a, -c, -d)$ usando amostragem. Abaixo são fornecidas algumas amostras que foram produzidas por meio de *prior sampling* (i.e., o estágio de rejeição na amostragem por rejeição não aconteceu ainda). Circule as amostras que seriam rejeitadas se fosse aplicada a amotragem por rejeição (*rejection sampling*):

$-a$	$-b$	$+c$	$+d$	$-a$	$-b$	$-c$	$-d$
$+a$	$-b$	$-c$	$+d$	$-a$	$+b$	$+c$	$+d$
$-a$	$-b$	$+c$	$-d$	$+a$	$-b$	$-c$	$-d$

(b) Usando as amostras acima, qual valor você estima para $\Pr(+b|-a,-c,-d)$ usando amostragem por rejeição?

(c) Usando as amostras abaixo (que foram geradas usando ponderação por verossimilhança (*likelihood weighting*)), produza uma estimativa para $\Pr(+b \mid -a,-c,-d)$ ponderação por verossimilhança, ou declare que é a computação é impossível, se for o caso.

$-a$	$-b$	$-c$	$-d$
$-a$	$+b$	$-c$	$-d$
$-a$	$-b$	$-c$	$-d$