CSE 150 Homework 2

Pedro Sousa Meireles

Fall 2018

1 Probabilistic reasoning

(a)

$$P(D=0|S_1=1,...,S_k=1) = \frac{P(S_1=1,...,S_k=1|D=0)\cdot P(D=0)}{P(S_1=1,...,S_k=1)}$$
$$P(D=1|S_1=1,...,S_k=1) = \frac{P(S_1=1,...,S_k=1|D=1)\cdot P(D=1)}{P(S_1=1,...,S_k=1)}$$
$$r_k = \frac{P(S_1=1,...,S_k=1|D=0)\cdot P(D=0)}{P(S_1=1,...,S_k=1|D=1)\cdot P(D=1)}$$

$$\begin{aligned} P(S_1=1,...,S_k=1|D=0) &= \prod_{i=1}^{k} P(S_i=1|parents(S_i),D=0) \\ &= \prod_{i=1}^{k} P(S_i=1|D=0) \\ &= \frac{1}{2^k} \end{aligned}$$

$$\begin{aligned} P(S_1=1,...,S_k=1|D=1) &= \prod_{i=1}^{k} P(S_i=1|parents(S_i),D=1) \\ &= \prod_{i=1}^{k} P(S_i=1|D=1) \\ &= \frac{f(1)}{f(k)} && \text{(intermediate terms cancel each other in the producto} \\ &= \frac{1}{2^k+(-1)^k} \end{aligned}$$

Replacing values in r_k formula:

$$\begin{aligned} r_k &= \frac{\frac{1}{2^k}}{\frac{1}{2^k+(-1)^k}} \\ &= \frac{2^k}{2^k+(-1)^k} \end{aligned}$$

This means $r_k < 1$ when k is even, which means $D=1$, and $r_k > 1$ when k is odd, which means $D=0$.

(b) The diagnosis becomes less certain as more symptoms are observed, because 2^k grows really fast, and at some moment adding or subtracting 1 in the denominator won't make much difference and the value of r_k will remain close to 1.

2 Noisy-OR

X_1	X_2	X_3	$P(Y=1\mid X_1,X_2,X_3)$
0	0	0	0
1	0	0	
0	1	0	$\frac{1}{3}$
0	0	1	
1	1	0	
1	0	1	
0	1	1	$\frac{4}{5}$
1	1	1	$\frac{5}{6}$

(a)

$$P(Y=1|X_1,X_2,X_3) = 1 - \prod_i (1-p_i)^{X_i}$$

$$P(Y=1|X_1=0,X_2=1,X_3=0) = 1-(1-p_2) = \frac{1}{3}$$

$$p_2 = \frac{1}{3}$$

$$P(Y=1|X_1=0,X_2=1,X_3=1) = 1-(1-\frac{1}{3})\cdot(1-p_3) = \frac{4}{5}$$

$$p_3 = \frac{7}{10}$$

$$P(Y=1|X_1=1,X_2=1,X_3=1) = 1-(1-p_1)\cdot(1-\frac{1}{3})\cdot(1-\frac{7}{10}) = \frac{5}{6}$$

$$p_1 = \frac{1}{6}$$

Knowing p_1, p_2 and p_3 it's easy to calculate the remaining probabilities:

$$P(Y=1|X_1=1,X_2=0,X_3=0) = 1-(1-\frac{1}{6}) = \frac{1}{6}$$

$$P(Y=1|X_1=0,X_2=0,X_3=1) = 1-(1-\frac{7}{10}) = \frac{7}{10}$$

$$P(Y=1|X_1=1,X_2=1,X_3=0) = 1-(1-\frac{1}{6})\cdot(1-\frac{1}{3}) = \frac{4}{9}$$

$$P(Y=1|X_1=1,X_2=0,X_3=1) = 1-(1-\frac{1}{6})\cdot(1-\frac{7}{10}) = \frac{3}{4}$$

So, completing the table:

X_1	X_2	X_3	$P(Y=1\|X_1,X_2,X_3)$
0	0	0	0
1	0	0	$\frac{1}{6}$
0	1	0	$\frac{1}{3}$
0	0	1	$\frac{7}{10}$
1	1	0	$\frac{4}{9}$
1	0	1	$\frac{3}{4}$
0	1	1	$\frac{4}{5}$
1	1	1	$\frac{5}{6}$

(b) Suppose $X_1 = burglary$, $X_2 = earthquake$, $X_3 = fire$ and $Y = alarm$.

$$P(X_2=1|Y=0) < P(X_2=1) < P(X_2=1|Y=1,X_1=1,X_3=1) < P(X_2=1|Y=1) < P(X_2=1|Y=1,X_1=0,X_3$$

$P(X_2=1|Y=0) < P(X_2=1)$: It is more likely to an earthquake to occur naturally than if we know the alarm didn't go off.

$P(X_2=1) < P(X_2=1|Y=1,X_1=1,X_3=1)$: If we know the alarm is off, although it may have other explanations, the probability of an earthquake is bigger than if we know nothing.

$P(X_2=1|Y=1,X_1=1,X_3=1) < P(X_2=1|Y=1)$: If we know the alarm has gone off, it is more likely that an earthquake caused it if we don't know that there was a burglary and a fire.

$P(X_2=1|Y=1) < P(X_2=1|Y=1,X_1=0,X_3=0)$: If we know the alarm has gone off and there's no burglary or fire, it is more likely that it was due to an earthquake than if we know nothing besides the alarm has gone off.

(c) (c).1

$$P(X_2=1) = \frac{1}{3}$$

(c).2

$$P(X_2=1|Y=0) = \frac{P(Y=0|X_2=1)\cdot P(X_2=1)}{P(Y=0)}$$

We already know $P(X_2=1)$, but we have to calculate $P(Y=0)$ and $P(Y=0|X_2=1)$:

$$\begin{aligned} P(Y=0) &= \sum_{i,j,k} P(Y=0, X_1=x_i, X_2=x_j, X_3=x_k) \\ &= \sum_{i,j,k} P(Y=0|X_1=x_i, X_2=x_j, X_3=x_k)\cdot P(X_1=x_i, X_2=x_j, X_3=x_k) \end{aligned}$$

$$P(Y=0|X_1=x_i, X_2=x_j, X_3=x_k) = 1 - P(Y=1|X_1=x_i, X_2=x_j, X_3=x_k)$$

and $P(X_1=x_i, X_2=x_j, X_3=x_k) = P(X_1=x_i)\cdot P(X_2=x_j)\cdot P(X_3=x_k)$

$$P(Y=0) = 0.643621399$$

$$P(Y=1) = 1 - P(Y=0) = 0.356378601$$

$$\begin{aligned} P(Y=0|X_2=1) &= \sum_{i,j} P(Y=0, X_1=x_i, X_3=x_j|X_2=1) \\ &= \sum_{i,j} P(Y=0|X_1=x_i, X_2=1, X_3=x_j)\cdot P(X_1=x_i, X_3=x_j|X_2=1) \end{aligned}$$

$P(Y=0|X_1=x_i,X_2=1,X_3=x_j)=1-P(Y=1|X_1=x_i,X_2=1,X_3=x_j)$

and $P(X_1=x_i,X_3=x_j|X_2=1)=P(X_1=x_i)\cdot P(X_3=x_j)$

$P(Y=0|X_2=1)=0.482716049$

$P(Y=1|X_2=1)=1-P(Y=0|X_2=1)=0.517283951$
Replacing in the original equation:

$$P(X_2=1|Y=0)=\frac{0.482716049\cdot\frac{1}{3}}{0.643621399}=0.25$$

(c).3

$$P(X_2=1|Y=1)=\frac{P(Y=1|X_2=1)\cdot P(X_2=1)}{P(Y=1)}$$

We already know all the needed probabilities. Now we just have to replace in the formula:

$$P(X_2=1|Y=1)=\frac{0.517283951\cdot\frac{1}{3}}{0.356378601}=0.483833718$$

(c).4

$$P(X_2=1|Y=1,X_1=0,X_3=0)=\frac{P(Y=1|X_1=0,X_2=1,X_3=0)\cdot P(X_2=1|X_1=0,X_3=0)}{P(Y=1|X_1=0,X_3=0)}$$

$P(X_2=1|X_1=0,X_3=0)=P(X_2=1)$, because X_1, X_2 and X_3 are conditionally independent.
Now we must calculate the value of $P(Y=1|X_1=0,X_3=0)$:

$$\begin{aligned}P(Y=1|X_1=0,X_3=0)&=\sum_i P(Y=1,X_2=x_i|X_1=0,X_3=0)\\&=\sum_i P(Y=1|X_1=0,X_2=x_i,X_3=0)\cdot P(X_2=x_i|X_1=0,X_3=0)\\&=\frac{1}{3}\cdot\frac{1}{3}\\P(Y=1|X_1=0,X_3=0)&=\frac{1}{9}\end{aligned}$$

Replacing in the original formula:

$$P(X_2=1|Y=1,X_1=0,X_3=0)=\frac{\frac{1}{3}\cdot\frac{1}{3}}{\frac{1}{9}}=1$$

(c).5

$$P(X_2=1|Y=1,X_1=1,X_3=1)=\frac{P(Y=1|X_1=1,X_2=1,X_3=1)\cdot P(X_2=1|X_1=1,X_3=1)}{P(Y=1|X_1=1,X_3=1)}$$

$P(X_2=1|X_1=1,X_3=1)=P(X_2=1)$, because X_1, X_2 and X_3 are conditionally independent.
Now we must calculate the value of $P(Y=1|X_1=1,X_3=1)$:

$$\begin{aligned}P(Y=1|X_1=1,X_3=1)&=\sum_i P(Y=1,X_2=x_i|X_1=1,X_3=1)\\&=\sum_i P(Y=1|X_1=1,X_2=x_i,X_3=1)\cdot P(X_2=x_i|X_1=1,X_3=1)\\&=\frac{3}{4}\cdot\frac{2}{3}+\frac{5}{6}\cdot\frac{1}{3}P(Y=1|X_1=1,X_3=1)&=\frac{7}{9}\end{aligned}$$

Replacing in the original formula:

$$P(X_2 = 1|Y = 1, X_1 = 1, X_3 = 1) = \frac{\frac{5}{6} \cdot \frac{1}{3}}{\frac{7}{9}} = 0.357142857$$

3 Hangman

(a) <u>10 least common words:</u>

Word Count Probability

BOSAK, 6, 7.827934689453437e-07

CAIXA, 6, 7.827934689453437e-07

MAPCO, 6, 7.827934689453437e-07

OTTIS, 6, 7.827934689453437e-07

TROUP, 6, 7.827934689453437e-07

CCAIR, 7, 9.13259047102901e-07

CLEFT, 7, 9.13259047102901e-07

FABRI, 7, 9.13259047102901e-07

FOAMY, 7, 9.13259047102901e-07

NIAID, 7, 9.13259047102901e-07

<u>10 most common words:</u>

Word Count Probability

FIFTY, 106869, 0.013942725872119989

FIRST, 109957, 0.014345603577470525

AFTER, 110102, 0.01436452108630337

WHICH, 142146, 0.018545160072784138

THEIR, 145434, 0.018974130893766185

ABOUT, 157448, 0.020541544349751077

WOULD, 159875, 0.02085818430793947

EIGHT, 165764, 0.021626496097709325

SEVEN, 178842, 0.023332724928853858

THREE, 273077, 0.03562714868653127

(b)

correctly guessed	incorrectly guessed	best next guess l	$P(L_i = l$ for some i e {1,2,3,4,5}$\|E)$
- - - - -	{}	E	0.5394
- - - - -	{E, O}	I	0.6366
Q - - - -	{}	U	0.9867
Q - - - -	{U}	A	1
- - Z E -	{A, D, I, R}	O	0.8803
- - - - -	{E, O}	I	0.6366
D - - I -	{}	A	0.8207
D - - I -	{A}	E	0.7521
- U - - -	{A, E, I, O, S}	Y	0.6270

(c)
```
In [1]: import re
```

```
In [2]: file = open('hw2_word_counts_05.txt','r')

In [3]: fileString = file.read()
        fileArray = re.compile(" |\n").split(fileString)
        fileArray.pop() # remove EOF

In [4]: words = fileArray[0:][::2]
        counts = fileArray[1:][::2]
        counts = [int(n) for n in counts]
        probabilities = [n/sum(counts) for n in counts]

In [5]: words = list(zip(words, counts, probabilities))

In [6]: words = sorted(words, key=lambda x:x[2])
```

10 least common words:

```
In [7]: words[:10]

Out[7]: [('BOSAK', 6, 7.827934689453437e-07),
         ('CAIXA', 6, 7.827934689453437e-07),
         ('MAPCO', 6, 7.827934689453437e-07),
         ('OTTIS', 6, 7.827934689453437e-07),
         ('TROUP', 6, 7.827934689453437e-07),
         ('CCAIR', 7, 9.13259047102901e-07),
         ('CLEFT', 7, 9.13259047102901e-07),
         ('FABRI', 7, 9.13259047102901e-07),
         ('FOAMY', 7, 9.13259047102901e-07),
         ('NIAID', 7, 9.13259047102901e-07)]
```

10 most probable words:

```
In [8]: words[len(words)-10:]

Out[8]: [('FIFTY', 106869, 0.013942725872119989),
         ('FIRST', 109957, 0.014345603577470525),
         ('AFTER', 110102, 0.01436452108630337),
         ('WHICH', 142146, 0.018545160072784138),
         ('THEIR', 145434, 0.018974130893766185),
         ('ABOUT', 157448, 0.020541544349751077),
         ('WOULD', 159875, 0.02085818430793947),
         ('EIGHT', 165764, 0.021626496097709325),
         ('SEVEN', 178842, 0.023332724928853858),
         ('THREE', 273077, 0.03562714868653127)]
```

Hangman

```
In [9]: letters = ["A", "B", "C", "D", "E", "F", "G",
                   "H", "I", "J", "K", "L", "M", "N",
                   "O", "P", "Q", "R", "S", "T", "U",
                   "V", "W", "X", "Y", "Z"]

In [10]: def compareWordGuess(word, guess):
             for i in range(0,5):
                 if (not guess[i] == '-' and not guess[i] == word[i])
                     or (guess[i] == '-' and word[i] in guess):
                     return False
             return True
```

```
In [17]: def guessLetter(words, corrGuess, incorrGuess):
             guessWords = [w for w in words
                           if compareWordGuess(w[0], corrGuess)
                           and not any(s in w[0] for s in incorrGuess)]
             wordProbs = [row[2] for row in guessWords]
             guessWords = [[w[0], w[1], w[2]/sum(wordProbs)] for w in guessWords]

             letterProb = [[letter, 0] for letter in letters
                           if letter not in corrGuess
                           and letter not in incorrGuess]
             for l in letterProb:
                 for word in guessWords:
                     if(l[0] in word[0]):
                         l[1] += word[2]
             letterProb = sorted(letterProb, key=lambda x:x[1])
             return letterProb[-1]
```

```
In [11]: print(guessLetter(words, "-----", ["E", "O"]))
         print(guessLetter(words, "D--I-", []))
         print(guessLetter(words, "D--I-", ["A"]))
         print(guessLetter(words, "-U---", ["A", "E", "I", "O", "S"]))
```

```
['I', 0.6365554141009612]
['A', 0.8206845238095238]
['E', 0.7520746887966805]
['Y', 0.6269651101630529]
```

```
In [12]: print(guessLetter(words, "-----", []))
         print(guessLetter(words, "-----", ["E", "O"]))
         print(guessLetter(words, "Q----", []))
         print(guessLetter(words, "Q----", ["U"]))
         print(guessLetter(words, "--ZE-", ["A", "D", "I", "R"]))
```

```
['E', 0.5394172389647974]
['I', 0.6365554141009612]
['U', 0.9866727159303579]
['A', 0.9999999999999999]
['O', 0.8803418803418803]
```

4 Conditional independence

(a) $P(F|H) = P(F|C,H)$

False. F is conditionally dependent on C, because the path $F \leftarrow C$ is not d-separated by H.

(b) $P(E|A,B) = P(E|A,B,F)$

$E \leftarrow \underline{B} \rightarrow F$, d-separated by rule #2

$E \rightarrow G \leftarrow F$, d-separated by rule #3

True. E is conditionally independent on F given A and B.

(c) $P(E, F|B, G) = P(E|B, G) \cdot P(F|B, G)$

$E \leftarrow \underline{B} \rightarrow F$, d-separated by rule #2

$E \rightarrow \underline{G} \leftarrow F$, not d-separated by rule #3

False. E and F are not conditionally independent given B and G.

(d) $P(A, B|D, E, F) = P(A, B|D, E, F, G, H)$

$A \rightarrow \underline{E} \leftarrow B \rightarrow \underline{F} \leftarrow C \rightarrow H$, not d-separated

False. A and B are conditionally dependent of G and H given D, E and F.

(e) $P(F|B, C, G, H) = P(F|B, C, E, G, H)$

$F \rightarrow \underline{G} \leftarrow E$, not d-separated

False. F is conditionally dependent on E given B, C, G and H.

(f) $P(D, E, F) = P(D) \cdot P(E|D) \cdot P(F|E)$

This question is equivalent to $P(F|D, E) = P(F|E)$ due to the expansion of product rule.

$F \leftarrow B \rightarrow E$, not d-separated

False. F and D are conditionally dependent given D.

(g) $P(A|F) = P(A)$

$A \rightarrow E \leftarrow B \rightarrow F$, not d-separated

False. A is conditionally dependent of F.

(h) $P(E, F) = P(E) \cdot P(F)$

This question is equivalent to $P(E|F) = P(E)$ due to product rule.

$E \leftarrow B \rightarrow F$, not d-separated

False. E and F are conditionally dependent.

(i) $P(D|A) = P(D|A, E)$

$D \leftarrow \underline{A} \rightarrow E$, d-separated by rule #2

True. D and E are conditionally independent given A.

(j) $P(B, C) = P(B) \cdot P(C)$

This question is equivalent to $P(B|C) = P(B)$ due to product rule.

$B \rightarrow E \rightarrow G \leftarrow F \leftarrow C$, d-separated by rule #3 (G)

$B \rightarrow E \rightarrow G \leftarrow F \rightarrow H \leftarrow C$, d-separated by rule #3 (G and H)

$B \rightarrow F \leftarrow C$, d-separated by rule #3 (F)

$B \rightarrow F \rightarrow H \leftarrow C$, d-separated by rule #3 (H)

True. B and C are conditionally independent.

5 Subsets

(a) $P(A) = P(A|S)$

$S = \{B, C, F, H\}$

(b) $P(A|C) = P(A|S)$

$S = \{B, C, F, H\}$

(c) $P(A|B, C) = P(A|S)$

$S = \{B, C, F, H\}$

(d) $P(B) = P(B|S)$
$S = \{A, D, C\}$

(e) $P(B|A, E) = P(B|S)$
$S = \{A, C, D, E\}$

(f) $P(B|A, C, E) = P(B|S)$
$S = \{A, C, D, E\}$

(g) $P(D) = P(D|S)$
$S = \{B, C, F, H\}$

(h) $P(D|A) = P(D|S) \cdot P(F)$
$S = \{A, B, C, E, F, G, H\}$

(i) $P(D|C, E) = P(D|S)$
$S = \{C, E\}$

(j) $P(D|F) = P(D|S)$
$S = \{B, C, F, H\}$