Homework 4

2021-31645-T1 Machine Learning
Emmanouil Palaiologos

December 12, 2021

Conditional Independence (Graph Models)

Bishop, C. M. (2006). Pattern recognition. Machine learning, 128(9).

Exercise 8.3

In order to prove by direct evaluation that the equality $P(a, b) = P(a)P(b)$ does **NOT** hold, we just need only *one* combination of values that is not satisfying this relation.

By looking on the table, we can acquire easily the marginal probabilities of $P(a)$ and $P(b)$, for the example case of $a = 0$ and $b = 0$. All we need to do is consider all possible combinations of the remaining two variables, when we fix one of them to be 0. That is, we are basically applying the **sum rule** on the *marginals* :

a	b	c	$p(a,b,c)$
0	0	0	0.192
0	0	1	0.144
0	1	0	0.048
0	1	1	0.216
1	0	0	0.192
1	0	1	0.064
1	1	0	0.048
1	1	1	0.096

$$\begin{cases} P(a=0) = \sum_{b,c} P(a=0, b, c) \\ P(b=0) = \sum_{b,c} P(a, b=0, c) \end{cases} \tag{1}$$

$$\implies \begin{cases} P(a=0) = 0.192 + 0.144 + 0.048 + 0.216 \\ P(b=0) = 0.192 + 0.144 + 0.192 + 0.064 \end{cases} \tag{2}$$

$$\boxed{P(a=0) = 0.6} \quad \text{and} \quad \boxed{P(b=0) = 0.592} \tag{3}$$

Similarly for the joint probability :

$$P(a=0,b=0)=\sum_{c} P(a=0,b=0,c)=0.192+0.144=0.336 \tag{4}$$

Comparing (3) and (4), we can deduce that

$$P(a=0)P(b=0)=0.355 \neq 0.336=P(a=0,b=0) \tag{5}$$

$$\Longrightarrow \boxed{P(a,b) \neq P(a)P(b)} \quad \text{q.e.d.} \tag{6}$$

Now, we shall verify the very foundation of *Markov Condition* about *"conditional independence"*. Given c, we prove that :

$$P(a,b|c)=P(a|c)P(b|c) \quad \text{, for both c=0,1} \tag{7}$$

- For $\underline{c=0}$, we apply **Bayes Theorem** for the left-hand side of (7):

$$P(a,b|c=0)=\frac{P(a,b,c=0)}{\sum_{a,b} P(a,b,c=0)} \tag{8}$$

$$\Longrightarrow \begin{cases} P(a=0,b=0|c=0)=\frac{0.192}{0.192+0.048+0.192+0.048}=\frac{0.192}{0.48}=0.4 \\ P(a=0,b=1|c=0)=\frac{0.048}{0.192+0.048+0.192+0.048}=\frac{0.048}{0.48}=0.1 \\ P(a=1,b=0|c=0)=\frac{0.192}{0.192+0.048+0.192+0.048}=\frac{0.192}{0.48}=0.4 \\ P(a=1,b=1|c=0)=\frac{0.048}{0.192+0.048+0.192+0.048}=\frac{0.048}{0.48}=0.1 \end{cases} \tag{9}$$

Next, we compute the right-hand side of (7) :

$$P(a=0|c=0)=\frac{P(a=0,c=0)}{P(c=0)}=\frac{\sum_b P(a=0,b,c=0)}{\sum_{a,b} P(a,b,c=0)}=\frac{0.24}{0.48}=0.5 \tag{10}$$

$$P(a=1|c=0)=\frac{\sum_b P(a=1,b,c=0)}{\sum_{a,b} P(a,b,c=0)}=\frac{0.24}{0.48}=0.5 \tag{11}$$

$$P(b=0|c=0)=\frac{\sum_a P(a,b=0,c=0)}{\sum_{a,b} P(a,b,c=0)}=\frac{0.384}{0.48}=0.8 \tag{12}$$

$$P(b=1|c=0)=\frac{\sum_a P(a,b=1,c=0)}{\sum_{a,b} P(a,b,c=0)}=\frac{0.096}{0.48}=0.2 \tag{13}$$

If we create factors of pairs $P(a_i|c=0)P(b_i|c=0)$ and we compare them with the results of (9), we are witnessing an **identity**! That means that we have proved (7) for $c=0$. When we condition over $c=0$ the probabilities become independent.

- For $\underline{c = 1}$ we perform similar calculations:

$$P(a,b|c=1) = \frac{P(a,b,c=0)}{\sum_{a,b} P(a,b,c=1)} \tag{14}$$

$$\Longrightarrow \begin{cases} P(a=0,b=0|c=1) = \frac{0.144}{0.144+0.216+0.064+0.096} = \frac{0.144}{0.52} = 0.277 \\ P(a=0,b=1|c=1) = \frac{0.216}{0.144+0.216+0.064+0.096} = \frac{0.216}{0.52} = 0.415 \\ P(a=1,b=0|c=1) = \frac{0.064}{0.144+0.216+0.064+0.096} = \frac{0.064}{0.52} = 0.123 \\ P(a=1,b=1|c=1) = \frac{0.096}{0.144+0.216+0.064+0.096} = \frac{0.096}{0.52} = 0.185 \end{cases} \tag{15}$$

and

$$P(a=0|c=1) = 0.692 \qquad P(a=1|c=1) = 0.308 \tag{16}$$
$$P(b=0|c=1) = 0.400 \qquad P(b=1|c=1) = 0.600 \tag{17}$$

Again multiplying the probabilities between any pair of (16), we get exactly the values of $P(a,b|c=1)$ as computed above. Equality (7) proven also for $c = 1$.

When we condition over $c = 0, 1$ the probabilities become independent.

Exercise 8.4

Solved in Python

.

.

mlhw

December 12, 2021

[ ]:
```
#Inserting table values, to a numpy array 2x2x2

import numpy as np
p=np.zeros((2,2,2))
for a in (0,1):
  for b in (0,1):
    for c in (0,1):
      p[a][b][c]=input("Enter P({},{},{})".format(a,b,c))
```

```
Enter P(0,0,0)0.192
Enter P(0,0,1)0.144
Enter P(0,1,0)0.048
Enter P(0,1,1)0.216
Enter P(1,0,0)0.192
Enter P(1,0,1)0.064
Enter P(1,1,0)0.048
Enter P(1,1,1)0.096
```

[ ]:
```
#initializations

p_a1=0 # p(a=0)
p_a2=0 # p(a=1)
p_c1=0 # p(c=0)
p_c2=0 # p(c=1)
p_bc=np.zeros((2,2))   # p(b|c)
p_ca=np.zeros((2,2))   # p(c|a)

for a in (0,1):
  for b in (0,1):
    for c in (0,1):
      if c==0:
        p_c1+=p[a][b][0]
      elif c==1:
        p_c2+=p[a][b][1]

      if a==0:
        p_a1+=p[a][b][c]
```

```
      elif a==1:
        p_a2+=p[a][b][c]


for a in (0,1):

  for b in (0,1):
    for c in (0,1):

      if c==0:
        p_bc[b][c]=(p[0][b][c]+p[1][b][c])/p_c1

      else:
        p_bc[b][c]=(p[0][b][c]+p[1][b][c])/p_c2
        #print("for {},{}, we get p(b|c)={} ".format(b,c,p_bc[b][c]))
 ↪#---debug

      if a==0:

        p_ca[c][a]=(p[a][0][c]+p[a][1][c])/p_a1
        #print("for {},{}, we get p(c|a)={} ".format(c,a,p_ca[c][a]))
 ↪#---debug



      elif a==1:

        p_ca[c][a]=(p[a][0][c]+p[a][1][c])/p_a2
```

```
for 0,1, we get p(c|a)=0.6
for 1,1, we get p(c|a)=0.39999999999999997
for 0,1, we get p(c|a)=0.6
for 1,1, we get p(c|a)=0.39999999999999997
```

We validate that it holds : $ p(a,b,c)= p(a)p(c|a)p(b|c)$

[ ]:
```
bad_count=0
bad_save=[]
cond_ind=True
for a in (0,1):
  for b in (0,1):
    for c in (0,1):
      if a==0:
        if p[a][b][c]!=round(p_a1*p_bc[b][c]*p_ca[c][a],4):

          bad_count+=1
          #bad_save.append((a,b,c)) #--debug
```

```
        elif a==1:
          if p[a][b][c]!=round(p_a2*p_bc[b][c]*p_ca[c][a],4):
            bad_count+=1
            #bad_save.append((a,b,c))  #--debug
if bad_count!=0:
  cond_ind=False

print(cond_ind)
```

```
True
```

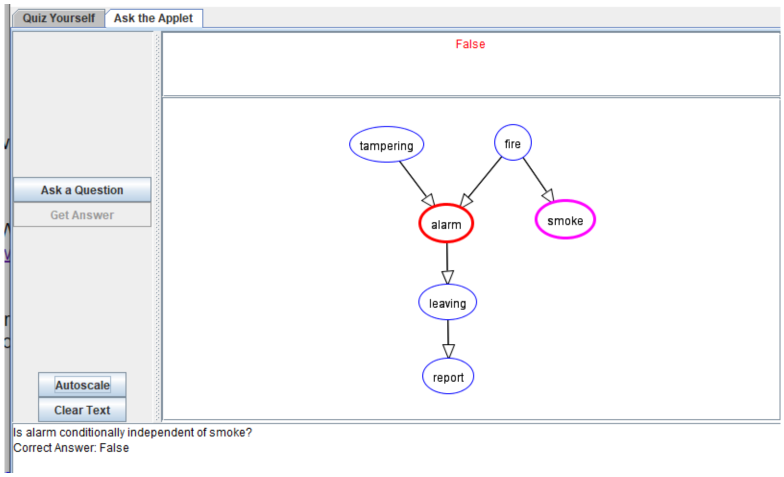
Q.E.D.! Now if we want to check for example a specific (random) combination like \$a=0 , b=1 , c=0 \$ we get:

[ ]:
```
p[0][1][0]==round(p_a1*p_bc[1][0]*p_ca[0][0],4)    #(we've put p_a1<--> p(a=1)␣
 ↪as a=1 here )
```

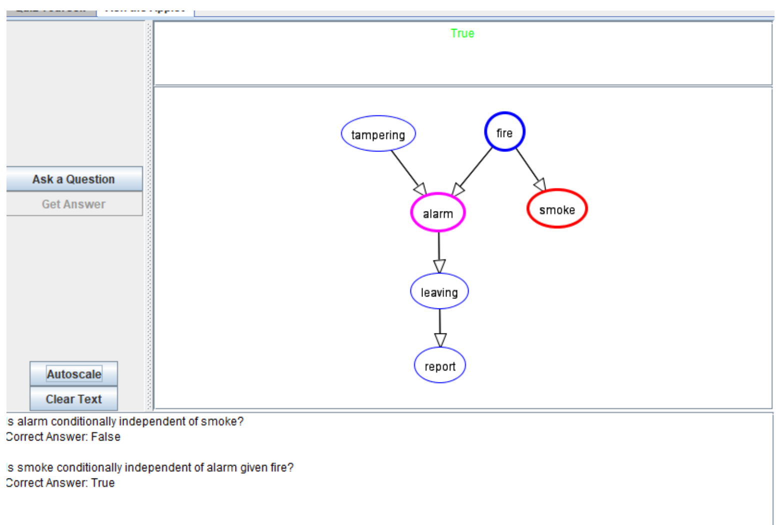
[ ]: True

Independence AIspace Quiz

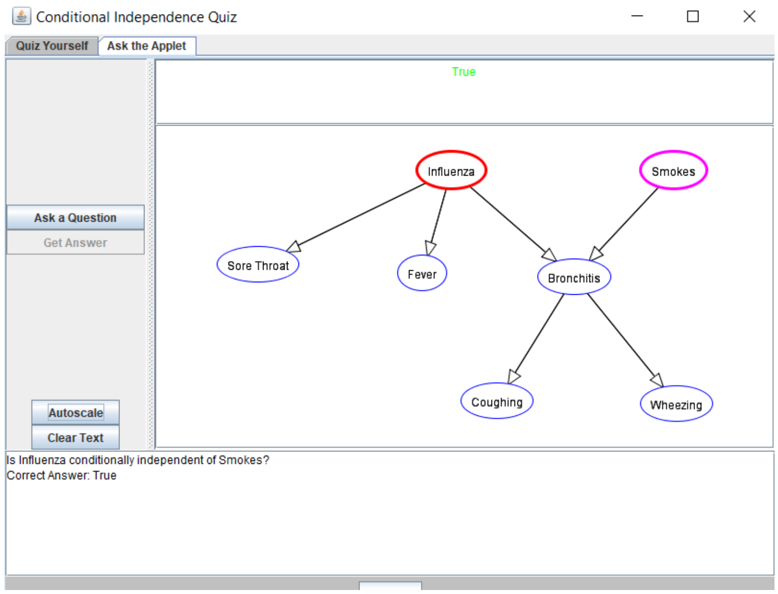
- Tail-to-Tail node : (unblocked) $\Longrightarrow$ Dependent (False)



- Tail-to-Tail node: (blocked) $\Longrightarrow$ Independent (True)



- Head-to-Head node: (*"d-seperated"-no active chain*) $\implies$ Independent (True)



- Head-to-Head node: (*"d-connected"-active chain due to descendant node*) $\implies$ Dependent (False)