Machine Learning Recitation 5

Sumeet Maan

12 February, 2019

1 Problem

To show $P(a, b, c|d, e) = P(a, b|c, d, e)\ P(c|d, e)$
Solution: We can write the LHS of the equation by product rule as:
$P(a, b, c|d, e) = \frac{P(a,b,c,d,e)}{P(d,e)}$ — equation 1 by product rule
Further, the joint probability distribution can be expressed by chain rule
$P(a, b, c, d, e) = P(a, b|c, d, e)P(c, d, e)$ —equation 2
Put equation 2 in 1.
$P(a, b, c|d, e) = \frac{P(a,b|c,d,e)P(c,d,e)}{P(d,e)}$
$P(c, d, e)$= $P(c|d, e)$P(d,e)
$P(a, b, c|d, e) = \frac{P(a,b|c,d,e)P(c|d,e)P(d,e)}{P(d,e)}$
On cancelling $P(d, e)$ common terms we get,
$P(a, b, c|d, e) = P(a, b|c, d, e)\ P(c|d, e)$ Hence, proved.

2 Problem

To show, $P(a|b, c)$=$\frac{P(b|a,c)P(a|c)}{P(b|c)}$
Solution: LHS of the equation can be written as
$P(a|b, c)$= $\frac{P(a,b,c)}{P(b,c)}$ — by product rule
$P(a|b, c)$= $\frac{P(b|a,c)P(a,c)}{P(b,c)}$— —equation 1 by chain rule
P(a,c) = $P(a|c)P(c)$
Similarly, P(b,c) = $P(b|c)P(c)$
Put both the above equations in equation 1, we get
$P(a|b, c) = \frac{P(b|a,c)P(a|c)P(c)}{P(b|c)P(c)}$
On cancelling $P(c)$ common term from numerator and denominator, we get,
$P(a|b, c)$=$\frac{P(b|a,c)P(a|c)}{P(b|c)}$

3 Problem

Solution:
$P(cavity|\neg catch) = \frac{P(cavity \wedge \neg catch)}{P(\neg catch)}$

$P(cavity \wedge \neg catch) = \frac{0.012+0.008}{0.012+0.008+0.064+0.576}$ – from the table given
$P(cavity \wedge \neg catch) = \frac{0.02}{0.66}$
$P(cavity \wedge \neg catch) = 0.03$ is the answer

4 Problem

To show, $P(Y|X, Z)$= $P(Y|X)$
Solution:
Applying product rule on the LHS of the equation.
To show, $P(Y|X, Z)$= $\frac{P(X,Y,Z)}{P(X,Z)}$—equation 1
P(X,Y,Z) can be written as :
P(X,Y,Z)= $P(X)P(Y|X)P(Z|X)$ —equation 2 by chain rule and looking at tail to tail connection in Bayesian network.
P(X,Z)=$P(Z|X)P(X)$— equation 3 by chain rule Substitute equation 2 and 3 in 1.
$P(Y|X, Z)$=$\frac{P(X)P(Y|X)P(Z|X)}{P(Z|X)P(X)}$
Cancelling the common terms P(X) and $P(Z|X)$
We get, $P(Y|X, Z)$= $P(Y|X)$
Hence, proved

5 Problem

To find $P(W|R)$ given the bayesian network
Solution:
We need to marginalize over S given R.
$P(W|R) = \sum_s (P(W, S|R)$—equation 1
$(P(W, S|R) = \frac{P(W,S,R)}{P(R)}$—equation 2 by product rule
P(W,R,S) = $P(W|R, S)P(R)P(S)$—equation 3 by chain rule
Substitute equation 3 in 2 and 2 in 1, we get
$P(W|R) = \sum_s \frac{P(W|R,S)P(R)P(S)}{P(R)}$
Cancelling common terms, P(R) we get,

$P(W|R) = \sum_s P(W|S,R)P(S)$
$P(W|R) = P(W|S,R)P(S) + P(W|\neg S,R)P(\neg S)$
$P(W|R)$= 0.95 x 0.2 + 0.90 x 0.8
$P(W|R)$= 0.19 + 0.72
$P(W|R)$= 0.91 is the answer

6 Problem

Solution a.
$P(W|\neg C,R)$= $\sum_s(P(W,S|\neg C,R)$— equation 1
Applying product rule to the RHS of equation 1
$P(W|\neg C,R)$= $\sum_s \frac{(P(W,S,\neg C,R)}{P(\neg C,R)}$
Applying chain rule on the RHS.
$P(W|\neg C,R)$= $\sum_s \frac{(P(W|S,\neg C,R)P(S|R,\neg C)P(R|\neg C)P(\neg C)}{P(R|\neg C)P(\neg C)}$
Cancelling $P(R|\neg C)P(\neg C)$, we get
$P(W|\neg C,R)$= $\sum_s P(W|S,\neg C,R)P(S|R,\neg C)$—equation 2
Considering the first term in the RHS
The below equation can be written without $\neg C$ considering the bayesian network as S and R block the path from W to C,
$P(W|S,\neg C,R) = P(W|S,R)$
And the second term can be written as,
$P(S|R,\neg C) = P(S|\neg C)$ as we don't need R to compute the probability of S given R and $\neg C$as S and R are in tail to tail connection
Re-writing equation 2, we have
$P(W|\neg C,R)$= $\sum_s P(W|S,R)P(S|\neg C)$
$P(W|\neg C,R)$=$P(W|S,R)P(S|\neg C)$+$P(W|\neg S,R)P(\neg S|\neg C)$
$P(W|\neg C,R)$=0.95 x 0.5 + 0.9 x 0.5
$P(W|\neg C,R)$= 0.475 + 0.45
$P(W|\neg C,R)$= 0.925 is the answer

Solution b.
$P(W|C,R)$= $\sum_s(P(W,S|C,R)$— equation 1
Applying product rule to the RHS of equation 1
$P(W|C,R)$= $\sum_s \frac{(P(W,S,C,R)}{P(C,R)}$
Applying chain rule on the RHS.
$P(W|C,R)$= $\sum_s \frac{(P(W|S,C,R)P(S|R,C)P(R|C)P(C)}{P(R|C)P(C)}$
Cancelling $P(R|C)P(C)$, weget
$P(W|C,R)$= $\sum_s P(W|S,C,R)P(S|R,C)$—equation 2

Considering the first term in the RHS
The below equation can be written without C considering the bayesian network as S and R block the path from W to C,
$P(W|S,C,R) = P(W|S,R)$
And the second term can be written as,
$P(S|R,C) = P(S|C)$ as we don't need R to compute the probability of S given R and C as S and R are in tail to tail connection
Re-writing equation 2, we have
$P(W|C,R)= \sum_s P(W|S,R)P(S|C)$
$P(W|C,R)=P(W|S,R)P(S|C)+P(W|\neg S,R)P(\neg S|C)$
$P(W|C,R)$=0.95 x 0.1 + 0.9 x 0.9
$P(W|C,R)$= 0.095 + 0.91
$P(W|C,R)$= 0.905 is the answer

7 Problem

Using d-separation algorithm to answer the questions asked

1. P and Q are conditionally dependant given S because P and Q are in head to head connection. S does not block the path between P and Q. They are connected by R.
2. P and Q are conditionally dependant given V because P and Q are in head to head connection by R. V does not block the path between P and Q. They are connected by R.
3. P and W are conditionally independant given R because P and W are in chain of head to tail connection by RSV. And R blocks the path between P and W.
4. S and T are conditionally dependant given P because S and T are in tail to tail connection. P does not block the path between S and T. They are connected by R.
5. U and W are conditionally independant given V because U and W are in chain of tail to tail connection by SV. And V blocks the path between U and W.