Artificial Intelligence Assignment – 3
Exact Inference using Variable Elimination

1) P(Sick | Lecture = True, Doctor = True)

Exact Inference: We will use variable elimination[1] to compute the exact inference for the above query. The above query can be written as P(S | l, d).

$$P(S \mid l, d) = \alpha P(S) \sum_p P(p) \sum_h P(h|s,p) P(l|h) P(d|h)$$

Now, we factorize L (Lecture) and D (Doctor). After that, we will get the factor product of H, L, and D. Finally we marginalize the factor and eliminate H (Headache).

$$= \alpha P(S) \sum_p P(p) \left[0.6 \times 0.3 \times \begin{pmatrix} 0.95 & 0.90 \\ 0.40 & 0.10 \end{pmatrix} + 0.8 \times 0.01 \times \begin{pmatrix} 0.05 & 0.10 \\ 0.60 & 0.90 \end{pmatrix}\right]$$

$$= \alpha P(S) \sum_p P(p) \begin{pmatrix} 0.1714 & 0.1628 \\ 0.0768 & 0.0252 \end{pmatrix}$$

$$= \alpha P(S) \left[0.1 \times \begin{pmatrix} 0.1714 \\ 0.1628 \end{pmatrix} + 0.9 \times \begin{pmatrix} 0.0768 \\ 0.0252 \end{pmatrix}\right]$$

The same way we eliminate P (Pub).

$$= \alpha \begin{pmatrix} 0.05 \\ 0.95 \end{pmatrix} \times \begin{pmatrix} 0.08626 \\ 0.03896 \end{pmatrix}$$

$$= \alpha \begin{pmatrix} 0.004313 \\ 0.037012 \end{pmatrix}$$

Finally, we normalize and get the exact inference of our query.

$$\approx \mathbf{< 0.104, 0.896 >}$$

2) P(Sick | Doctor = False)

Exact Inference: We will use variable elimination[1] to compute the exact inference for the above query. The above query can be written as P(S = true | D = false).

$$P(S = true \mid D = false) = \alpha \sum_{P,H,L} P(S = true, P, H, L, D = true)$$

$$= \alpha \sum_{P,H,L} [P(S = true) P(P) P(H \mid S = true, P) P(L \mid H)\, P(D = false, H)]$$

Evaluating the above expression from right to left, we notice that $\sum_L P(L \mid H)$ is equal to 1 by definition. Hence, there was no need to include it in the first place; the variable L is irrelevant to this query.

$$= \alpha P(S = true) \sum_H [P(D = false|H) \sum_P P(P) P(H|S = true, P)]$$

$$= \alpha P(S = true) \sum_H P(D = false \mid H) f_P(H)$$

Where $f_P(H = true) = (0.95)(0.1) + (0.4)(0.9) = 0.455$ and $f_P(H = false) = (0.05)(0.1) + (0.6)(0.9) = 0.545$.

Continuing,

$$= \alpha 0.05(0.455 \times 0.4 + 0.545 \times 0.99)$$

$$= 0.03607752\alpha$$

Similarly, we can evaluate

$$P(S = false \mid D = false) = \alpha P(S = false) \sum_H P(D = false \mid H) f'_P(H)$$

Where $f'_P(H = true)$ = (0.9)(0.1) + (0.1)(0.9) = 0.18 and $f'_P(H = false)$ = (0.1)(0.1) + (0.9)(0.9) = 0.82.

Continuing,

= α0.95(0.18 × 0.4 + 0.82 × + 0.99)

= 0.83961α

Therefore, P(S = true | D = false) = $\frac{0.03607752\alpha}{0.03607752\alpha + 0.83961\alpha}$ = 0.0411 and P(S = false | D = false) = 1 – 0.0411 = 0.9589.

Finally, we normalize and get the exact inference of our query.

$$\approx < \mathbf{0.0411, 0.9589} >$$

3) P(Pub | Lecture = False)

Exact Inference: We will use variable elimination[1] to compute the exact inference for the above query. The above query can be written as P(P = true | L = false).

P(P = true | L = false) = $\alpha \sum_{S,H,D} P(P = true, S, H, D, L = true)$

= $\alpha \sum_{S,H,D}[P(P = true)P(S)P(H \mid P = true, P)P(D \mid H)\, P(L = false, H)]$

Evaluating the above expression from right to left, we notice that $\sum_D P(D \mid H)$ is equal to 1 by definition. Hence, there was no need to include it in the first place; the variable D is irrelevant to this query.

= $\alpha P(P = true) \sum_H [P(L = false|H) \sum_S P(S)P(H|P = true, S)]$

= $\alpha P(S = true) \sum_H P(L = false \mid H) f_S(H)$

Where $f_S(H = true)$ = (0.95)(0.05) + (0.4)(0.95) = 0.4275 and $f_S(H = false)$ = (0.05)(0.05) + (0.6)(0.95) = 0.5725.

Continuing,

= α0.1(0.7 × 0.4275 + 0.2 × 0.5725)

= 0.0445252α

Similarly, we can evaluate

P(P = false | D = false) = $\alpha P(P = false) \sum_H P(L = false \mid H) f'_S(H)$

Where $f'_S(H = true)$ = (0.9)(0.05) + (0.1)(0.95) = 0.14 and $f'_S(H = false)$ = (0.1)(0.05) + (0.9)(0.95) = 0.86.

Continuing,

= α0.9(0.14 × 0.7 + 0.86 × + 0.2)

= 0.27α

Therefore, P(P = true | D = false) = $\frac{0.0445252\alpha}{0.0445252\alpha + 0.27\alpha}$ = 0.1415 and P(S = false | D = false) = 1 – 0.1415 = 0.8585.

Finally, we normalize and get the exact inference of our query.

$$\approx < \mathbf{0.1415, 0.8585} >$$

4) P(Pub | Lecture = False, Doctor = True)

Exact Inference: We will use variable elimination[1] to compute the exact inference for the above query. The above query can be written as P(P = true | L = false, D = true).

$$P(P = true \mid L = false, D = true) = \alpha \sum_{S,H} P(P = true, S, H, D = true, L = true)$$

$$= \alpha \sum_{S,H} [P(P = true) P(S) P(H \mid P = true, P) P(D = true \mid H)\, P(L = false, H)]$$

$$= \alpha P(P = true) \sum_{H} P(L = false \mid H) P(D = true \mid H) f_S(H)$$

Where $f_S(H = true) = (0.95)(0.05) + (0.4)(0.95) = 0.4275$ and $f_S(H = false) = (0.05)(0.05) + (0.6)(0.95) = 0.5725$.

Continuing,

$$= \alpha 0.1(0.7 \times 0.6 \times 0.4275 + 0.2 \times 0.01 \times 0.5725)$$

$$= 0.0180695\alpha$$

Similarly, we can evaluate

$$P(P = false \mid L = false, D = true) = \alpha P(P = false) \sum_{H} P(L = false \mid H) P(D = true \mid H) f'_S(H)$$

Where $f'_S(H = true) = (0.9)(0.05) + (0.1)(0.95) = 0.14$ and $f'_S(H = false) = (0.1)(0.05) + (0.9)(0.95) = 0.86$.

Continuing,

$$= \alpha 0.9(0.7 \times 0.6 \times 0.14 + 0.2 \times 0.01 \times 0.86)$$

$$= 0.054468\alpha$$

Therefore, P(P = true | L = false, D = true) = $\frac{0.0180695\alpha}{0.0180695\alpha + 0.054468\alpha}$ = 0.2491 and P(P = false | L = false, D = true) = 1 – 0.2491 = 0.7509.

Finally, we normalize and get the exact inference of our query.

$$\approx < \mathbf{0.2491, 0.7509} >$$

Approximate Inference using Rejection Sampling

For computing approximate inference, we need to generate random samples, therefore using random generator we have generated 45 samples, which are listed below in the table with the random values and their topological order {S, P, H, L, D}.

As noted in the algorithm of rejection sampling [2], as more samples are collected, the estimate will converge to the true answer. The standard deviation of the error in each probability will be proportional to $1/\sqrt{n}$, where n is the number of samples used in the estimate. The only major problem with rejection sampling is that it rejects so many samples.

Random Values	Samples
0.85 0.97 0.14 0.11 0.83	{-S, -P, -H, +L, -D}
0.67 0.38 0.67 0.18 0.58	{-S, -P, -H, +L, -D}
0.90 0.46 0.34 0.70 0.76	{-S, -P, -H, +L, -D}
0.58 0.18 0.08 0.67 0.89	{-S, -P, +H, -L, -D}
0.42 0.35 0.72 0.37 0.56	{-S, -P, -H, +L, -D}
0.41 0.53 0.62 0.43 0.01	{-S, -P, -H, +L, -D}
0.79 0.30 0.31 0.15 0.62	{-S, -P, -H, +L, -D}
0.27 0.10 0.84 0.75 0.17	{-S, -P, -H, +L, -D}
0.49 0.18 0.25 0.43 0.87	{-S, -P, -H, +L, -D}
0.14 0.85 0.06 0.19 0.79	{-S, -P, +H, +L, -D}
0.84 0.47 0.71 0.37 0.45	{-S, -P, -H, +L, -D}
0.69 0.61 0.37 0.94 0.47	{-S, -P, -H, -L, -D}
0.97 0.87 0.85 0.75 0.97	{-S, -P, -H, +L, -D}
0.60 0.27 0.98 0.43 0.92	{-S, -P, -H, +L, -D}
0.67 0.01 0.04 0.91 0.38	{-S, +P, +H, -L, +D}
0.02 0.32 0.76 0.84 0.64	{+S, -P, -H, -L, -D}
0.79 0.87 0.38 0.54 0.83	{-S, -P, -H, +L, -D}
0.23 0.59 0.66 0.83 0.53	{-S, -P, -H, -L, -D}
0.17 0.23 0.05 0.64 0.21	{-S, -P, +H, -L, +D}
0.33 0.19 0.70 0.04 0.81	{-S, -P, -H, +L, -D}
0.89 0.97 0.96 0.39 0.85	{-S, -P, -H, +L, -D}
0.51 0.82 0.49 0.38 0.11	{-S, -P, -H, +L, -D}
0.14 0.05 0.44 0.82 0.89	{-S, +P, +H, -L, -D}
0.64 0.60 0.83 0.46 0.89	{-S, -P, -H, +L, -D}

0.00 0.43 0.37 0.41 0.17	{+S, -P, +H, -L, +D}
0.58 0.90 0.23 0.64 0.12	{-S, -P, -H, -L, -D}
0.62 0.48 0.73 0.56 0.72	{-S, -P, -H, +L, -D}
0.76 0.51 0.62 0.68 0.75	{-S, -P, -H, +L, -D}
0.49 0.01 0.78 0.91 0.70	{-S, +P, +H, -L, -D}
0.05 0.75 0.67 0.67 0.69	{-S, -P, -H, +L, -D}
0.27 0.01 0.23 0.49 0.80	{-S, +P, +H, -L, -D}
0.30 0.89 0.50 0.31 0.83	{-S, -P, -H, +L, -D}
0.88 0.23 0.93 0.01 0.51	{-S, -P, -H, +L, -D}
0.73 0.42 0.35 0.40 0.53	{-S, -P, -H, +L, -D}
0.52 0.33 0.44 0.73 0.67	{-S, -P, -H, +L, -D}
0.96 0.03 0.49 0.02 0.25	{-S, +P, +H, +L, +D}
0.24 0.51 0.38 0.56 0.44	{-S, -P, -H, +L, -D}
0.53 0.93 0.04 0.48 0.24	{-S, -P, +H, -L, +D}
0.22 0.24 0.74 0.05 0.56	{-S, -P, -H, +L, -D}
0.40 0.90 0.13 0.53 0.79	{-S, -P, -H, +L, -D}
0.88 0.08 0.11 0.16 0.24	{-S, +P, +H, +L, +D}
0.94 0.89 0.26 0.19 0.11	{-S, -P, -H, +L, -D}
0.78 0.18 0.56 0.23 0.70	{-S, -P, -H, +L, -D}
0.41 0.58 0.29 0.76 0.99	{-S, -P, -H, +L, -D}
0.01 0.11 0.02 1.00 0.55	{+S, -P, +H, -L, +D}

Code to generate random values in Python between 0 to 1:

```
//import random
//file = open('tmp.txt', 'w')
//for a in range(45):
//	for i in range(0, 5):
//		n = random.random()
//		formatted_n = "{:.2f}".format(n)
//		file.write(formatted_n)
//		print(formatted_n, end = " ")
//	print(end='\n')
//file.close()
```

3) P(Pub | Lecture = False)

Approximate Inference: We will use Rejection Sampling [2] to compute the approximate inference for the above query. The above query can be written as P(P = true | L = false).
In the Rejection Sampling, we reject/remove those samples which do not match the evidence, and for the above query, the evidence is L = false. So After rejecting the samples our Sample list will be as follows.

Random Values	Samples
0.58 0.18 0.08 0.67 0.89	{-S, -P, +H, -L, -D}
0.69 0.61 0.37 0.94 0.47	{-S, -P, -H, -L, -D}
0.67 0.01 0.04 0.91 0.38	{-S, +P, +H, -L, +D}
0.02 0.32 0.76 0.84 0.64	{+S, -P, -H, -L, -D}
0.23 0.59 0.66 0.83 0.53	{-S, -P, -H, -L, -D}
0.17 0.23 0.05 0.64 0.21	{-S, -P, +H, -L, +D}
0.14 0.05 0.44 0.82 0.89	{-S, +P, +H, -L, -D}
0.00 0.43 0.37 0.41 0.17	{+S, -P, +H, -L, +D}
0.58 0.90 0.23 0.64 0.12	{-S, -P, -H, -L, -D}
0.49 0.01 0.78 0.91 0.70	{-S, +P, +H, -L, -D}
0.27 0.01 0.23 0.49 0.80	{-S, +P, +H, -L, -D}
0.53 0.93 0.04 0.48 0.24	{-S, -P, +H, -L, +D}
0.01 0.11 0.02 1.00 0.55	{+S, -P, +H, -L, +D}

$$\text{P(Pub = true | Lecture = false)} \approx \frac{N_{PS}(Pub = true,\ \ Lecture = false)}{N_{PS}(Lecture = false)}$$

$$\text{P(Pub = true | Lecture = false)} = \frac{4}{13}$$

Similarly, P(Pub = false | Lecture = false) = 1 – P(Pub = true | Lecture = false)

$$= 1 - \frac{4}{13} = \frac{9}{13}$$

Finally, we get the approximate inference of our query.

$$\approx < \mathbf{0.3076, 0.6924} >$$

Comparing with the exact inference ($< \mathbf{0.1415, 0.8585} >$), the approximate inference has much larger error. The main reason behind this error is that when we use rejection sampling, as we generate samples at the same time based on the evidence it rejects a lot of samples as well, thus leading to poor inference results.

4) P(Pub | Lecture = False, Doctor = True)

Approximate Inference: We will use Rejection Sampling [2] to compute the approximate inference for the above query. The above query can be written as P(P = true | L = false, Doctor = true).
In the Rejection Sampling, we reject/remove those samples which do not match the evidence, and for the above query, the evidence is L = false, Doctor = true. Considering the topological order of {S, P, H, L, D}, we first reject samples based on the evidence of Lecture = false and after that Doctor = true. So After rejecting the samples our Sample list will be as follows.

Random Values	Samples
0.67 0.01 0.04 0.91 0.38	{-S, +P, +H, -L, +D}
0.17 0.23 0.05 0.64 0.21	{-S, -P, +H, -L, +D}
0.00 0.43 0.37 0.41 0.17	{+S, -P, +H, -L, +D}
0.53 0.93 0.04 0.48 0.24	{-S, -P, +H, -L, +D}
0.01 0.11 0.02 1.00 0.55	{+S, -P, +H, -L, +D}

$$\text{P(Pub = true | Lecture = false, Doctor = true)} \approx \frac{N_{PS}(Pub = true,\ Lecture = false,\ Doctor = true)}{N_{PS}(Lecture = false,\ Doctor = true)}$$

$$\text{P(Pub = true | Lecture = false, Doctor = true)} = \frac{1}{5}$$

Similarly, P(Pub = false | Lecture = false, Doctor = true) = 1 – P(Pub = true | Lecture = false, Doctor = true)

$$= 1 - \frac{1}{5} = \frac{4}{5}$$

Finally, we get the approximate inference of our query.

$$\approx <\mathbf{0.20}, \mathbf{0.80}>$$

Comparing with the exact inference ($<\mathbf{0.2491}, \mathbf{0.7509}>$) the approximate inference has a much lower error. The main reason behind this low error is that based on randomly generated samples, even after rejecting such samples, more samples were not rejected based on the evidence criteria, thus leading to much lower error based inference results.

References:

[1] Russell, Stuart, and Peter Norvig. "Artificial intelligence: a modern approach." (2002). PP: e524 – e527
[2] Russell, Stuart, and Peter Norvig. "Artificial intelligence: a modern approach." (2002). PP: e532 – e533