1.1 Probabilistic reasoning

Variable assumption:

- M represents my children's situation, $M = 1$ means meltdown, $M = 0$ means not;
- L represents if I will be late, $L = 1$ means late, $L = 0$ means not;

According to the question:

$$\begin{aligned} P(M=1) &= 0.01 \\ P(L=1|M=1) &= 0.98 \\ P(L=1|M=0) &= 0.03 \end{aligned}$$

And we have to get $P(M = 1|L = 1)$:

$$\begin{aligned} P(M=1|L=1) &= \frac{P(L=1|M=1)P(M=1)}{P(L=1|M=1)P(M=1)+P(L=1|M=0)P(M=0)} \\ &= \frac{0.98 \times 0.01}{0.98 \times 0.01 + 0.03 \times (1-0.01)} \\ &\approx 0.2481 \end{aligned}$$

1.2 Conditioning on background evidence

(a)

$$\begin{aligned} P(X|Y,E) &= \frac{P(X,Y,E)}{P(Y,E)} \\ &= \frac{P(Y|X,E)P(X|E)P(E)}{P(Y|E)P(E)} \\ &= \frac{P(Y|X,E)P(X|E)}{P(Y|E)} \end{aligned}$$

(b)

$$\begin{aligned} P(X|E) &= \frac{P(X,E)}{P(E)} \\ &= \frac{\sum_y P(X,Y=y,E)}{P(E)} \\ &= \frac{\sum_y P(X,Y=y|E)P(E)}{P(E)} \\ &= \sum_y P(X,Y=y|E) \end{aligned}$$

1.3 Conditional independence

$(i) \Rightarrow (ii)$:

$$\begin{aligned} P(X|Y,E) &= \frac{P(X,Y|E)}{P(Y|E)} \\ &\overset{(i)}{=} \frac{P(X|E)P(Y|E)}{P(Y|E)} \\ &= P(X|E) \end{aligned}$$

$(ii) \Rightarrow (iii)$:

$$\begin{aligned} P(Y|X,E) &= \frac{P(X|Y,E)P(Y|E)}{P(X|E)} \\ &\overset{(ii)}{=} \frac{P(X|Y,E)P(Y|E)}{P(X|Y,E)} \\ &= P(Y|E) \end{aligned}$$

$(iii) \Rightarrow (i)$:

$$\begin{aligned} P(X,Y|E) &= P(Y|X,E)P(X|E) \\ &\overset{(iii)}{=} P(X|E)P(Y|E) \end{aligned}$$

In conclusion, these three statements are equivalent.

1.4 Creative writing

(a)

- X represents if a student can pass the test, $X = 1$ means pass, $X = 0$ means not;
- Y represents if the test is difficult, $Y = 1$ means difficult, $Y = 0$ means not;
- Z represents if the student is smart, $Z = 1$ means smart, $Z = 0$ means not;

If the student is smart, it will raise the probability of passing the test. Therefore:

$$P(X = 1|Z = 1) > P(X = 1)$$

If the test is difficult, it will decrease the probability that a smart student can pass the test. Therefore:

$$P(X = 1|Y = 1, Z = 1) < P(X = 1|Z = 1)$$

(b)

- X represents if a student can pass the test, $X = 1$ means pass, $X = 0$ means not;
- Y represents if the test is easy, $Y = 1$ means easy, $Y = 0$ means not;
- Z represents if the student is smart, $Z = 1$ means smart, $Z = 0$ means not;

If the test is easy, it will raise the probability of passing the test. Therefore:

$$P(X=1) < P(X=1|Y=1)$$

If the student is smart, it will raise the probability passing the easy test. Therefore:

$$P(X=1|Y=1) < P(X=1|Z=1, Y=1)$$

In conclusion:

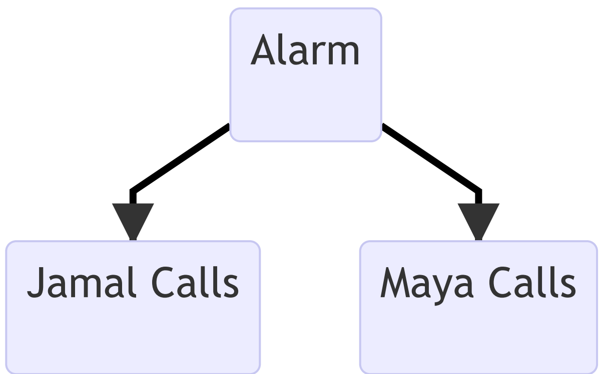
$$P(X=1) < P(X=1|Y=1) < P(X=1|Z=1, Y=1)$$

(c)

Actually it's hard for me to create proper example fitting the problem. We can use the example in class and set proper numerical value to fit it.

- Z represents if the alarm is ringing, $Z=1$ means ringing, $Z=0$ means not;
- X represents if Jamal will call, $X=1$ means calling, $X=0$ means not;
- Y represents if Maya will call, $Y=1$ means calling, $Y=0$ means not;

And the relevance of X, Y and Z are as follows:



Suppose that:

$$\begin{aligned}
&P(Z=1) = 0.9 \\
&P(X=1|Z=0) = 0.8 \\
&P(X=1|Z=1) = 0.1 \\
&P(Y=1|Z=0) = 0.1 \\
&P(Y=0|Z=1) = 0.2
\end{aligned}$$

In the class we have know that X and Y are conditionally independent given Z. Therefore:

$$P(X, Y|Z) = P(X|Z)P(Y|Z)$$

Then:

$$
\begin{aligned}
P(X,Y) &= \sum_z P(X,Y,Z=z) \\
&= \sum_z P(X|Y,Z=z)P(Y|Z=z)P(Z=z) \\
&= \sum_z P(X|Z=z)P(Y|Z=z)P(Z=z)
\end{aligned}
$$

And:

$$
\begin{aligned}
P(X) &= \sum_z P(X|Z=z)P(Z=z) \\
P(Y) &= \sum_z P(Y|Z=z)P(Z=z)
\end{aligned}
$$

Using the concluded formula, we can get:

$$
\begin{aligned}
P(X=1,Y=1) &= \sum_{i=0}^{1} P(X=1|Z=i)P(Y=1,Z=i)P(Z=i) \\
&= 0.8 \times 0.1 \times (1-0.9) + 0.1 \times 0.2 \times 0.9 \\
&= 0.026
\end{aligned}
$$

And:

$$
\begin{aligned}
P(X=1) &= P(X=1|Z=0)P(Z=0) + P(X=1|Z=1)P(Z=1) \\
&= 0.8 \times (1-0.9) + 0.1 \times 0.9 \\
&= 0.17 \\
P(Y=1) &= P(Y=1|Z=0)P(Z=0) + P(Y=1|Z=1)P(Z=1) \\
&= 0.1 \times (1-0.9) + 0.2 \times 0.9 \\
&= 0.19
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
P(X=1)P(Y=1) &= 0.17 \times 0.19 \\
&= 0.0323
\end{aligned}
$$

In conclusion:

$$
P(X=1,Y=1) < P(X=1)P(Y=1)
$$

1.5 Probabilistic inference

(a)

$$
\begin{aligned}
P(B=1|A=1) &= P(B=1,E=0|A=1) + P(B=1,E=1|A=1) \\
&= \frac{P(A=1|B=1,E=0)P(B=1,E=0)}{P(A=1)} + \\
&\frac{P(A=1|B=1,E=1)P(B=1,E=1)}{P(A=1)} \\
&= \frac{P(A=1|B=1,E=0)P(B=1)P(E=0)}{P(A=1)} + \\
&\frac{P(A=1|B=1,E=1)P(B=1)P(E=1)}{P(A=1)}
\end{aligned}
$$

Then $P(A=1)$ need to be calculated:

$$
\begin{aligned}
P(A=1) &= \sum_{b,e} P(A=1,B=b,E=e) \\
&= \sum_{b,e} P(A=1|B=b,E=e)P(B=b,E=e) \\
&= 0.002 \times 0.995 \times 0.999 + 0.96 \times 0.995 \times 0.001 + \\
&0.35 \times 0.005 \times 0.999 + 0.98 \times 0.005 \times 0.001 \\
&= 0.00469636
\end{aligned}
$$

In conclusion:

$$
\begin{aligned}
P(B=1|A=1) &= \frac{0.96 \times 0.001 \times 0.995}{0.00469636} + \frac{0.98 \times 0.001 \times 0.005}{0.00469636} \\
&\approx 0.204435
\end{aligned}
$$

(b)

$$
\begin{aligned}
P(B=1|A=1,E=0) &= \frac{P(A=1|B=1,E=0)P(B=1|E=0)}{P(A=1|E=0)} \\
&= \frac{P(A=1|B=1,E=0)P(B=1)}{\sum_b P(A=1|B=b,E=0)P(B=b|E=0)} \\
&= \frac{P(A=1|B=1,E=0)P(B=1)}{\sum_b P(A=1|B=b,E=0)P(B=b)} \\
&= \frac{0.96 \times 0.001}{0.002 \times (1-0.001) + 0.96 \times 0.001} \\
&\approx 0.32454
\end{aligned}
$$

(c)

$$
\begin{aligned}
P(A=1|J=1) &= \frac{P(J=1|A=1)P(A=1)}{P(J=1)} \\
&= \frac{P(J=1|A=1)P(A=1)}{P(J=1,A=0)+P(J=1,A=1)} \\
&= \frac{P(J=1|A=1)P(A=1)}{P(J=1|A=0)P(A=0)+P(J=1|A=1)P(A=1)} \\
&= \frac{0.93 \times 0.00469636}{0.1 \times (1-0.00469636) + 0.93 \times 0.00469636} \\
&\approx 0.04204
\end{aligned}
$$

(d)

$$
\begin{aligned}
&P(A=1|J=1,M=1) \\
&= \frac{P(J=1|M=1,A=1)P(M=1|A=1)P(A=1)}{P(J=1,M=1)} \\
&= \frac{P(J=1|A=1)P(M=1|A=1)P(A=1)}{P(J=1,M=1,A=0)+P(J=1,M=1,A=1)} \\
&= \frac{P(J=1|A=1)P(M=1|A=1)P(A=1)}{P(J=1|A=0)P(M=1|A=0)P(A=0)+P(J=1|A=1)P(M=1|A=1)P(A=1)} \\
&= \frac{0.93 \times 0.65 \times 0.00469636}{0.1 \times 0.01 \times (1-0.00469636) + 0.93 \times 0.65 \times 0.00469636} \\
&\approx 0.740418
\end{aligned}
$$

(e)

$$
\begin{aligned}
P(A=1|M=0) &= \frac{P(M=0|A=1)P(A=1)}{P(M=0)} \\
&= \frac{P(M=0|A=1)P(A=1)}{P(M=0|A=0)P(A=0)+P(M=0|A=1)P(A=1)} \\
&= \frac{(1-0.65) \times 0.00469636}{(1-0.01) \times (1-0.00469636) + (1-0.65) \times 0.00469636} \\
&\approx 0.0016654
\end{aligned}
$$

(f)

$$
\begin{aligned}
P(A=1|M=0,E=1) &= \frac{P(M=0|A=1,E=1)P(A=1|E=1)}{P(M=0|E=1)} \\
&= \frac{P(M=0|A=1)P(A=1|E=1)}{P(M=0|E=1)}
\end{aligned}
$$

Calculate $P(A=1|E=1)$:

$$
\begin{aligned}
P(A=1|E=1) = {} & P(A=1,B=0|E=1)+P(A=1,B=1|E=1) \\
& = P(A=1|B=0,E=1)P(B=0|E=1)+ \\
& P(A=1|B=1,E=1)P(B=1|E=1) \\
& = P(A=1|B=0,E=1)P(B=0)+ \\
& P(A=1|B=1,E=1)P(B=1) \\
& = 0.35\times(1-0.001)+0.98\times 0.001 \\
& = 0.35063
\end{aligned}
$$

Calculate $P(M=0|E=1)$:

$$
\begin{aligned}
P(M=0|E=1) &= \sum_a P(M=0,A=a|E=1) \\
&= \sum_a P(M=0|A=a,E=1)P(A=a|E=1) \\
&= \sum_a [P(M=0|A=a)\sum_b P(A=a,B=b|E=1)] \\
&= \sum_a [P(M=0|A=a)\sum_b P(A=a|B=b,E=1)P(B=b|E=1)] \\
&= \sum_a [P(M=0|A=a)\sum_b P(A=a|B=b,E=1)P(B=b)] \\
&= (1-0.01)\times[(1-0.35)\times(1-0.001)+(1-0.98)\times 0.001]+ \\
&(1-0.65)\times[0.35\times(1-0.001)+0.98\times 0.001] \\
&= 0.7655968
\end{aligned}
$$

In conclusion:

$$
\begin{aligned}
P(A=1|M=0,E=1) &= \frac{P(M=0|A=1)P(A=1|E=1)}{P(M=0|E=1)} \\
&= \frac{(1-0.65)\times 0.35063}{0.7655968} \\
&= 0.16029
\end{aligned}
$$

The results of (a) and (b), (c) and (d), (e) and (f) are all different. They are consistent with commonsense patterns of reasoning.

1.6 Kullback-Leibler distance

(a)

Assume that:

$$f(z)=\log z-(z-1)$$

Then:

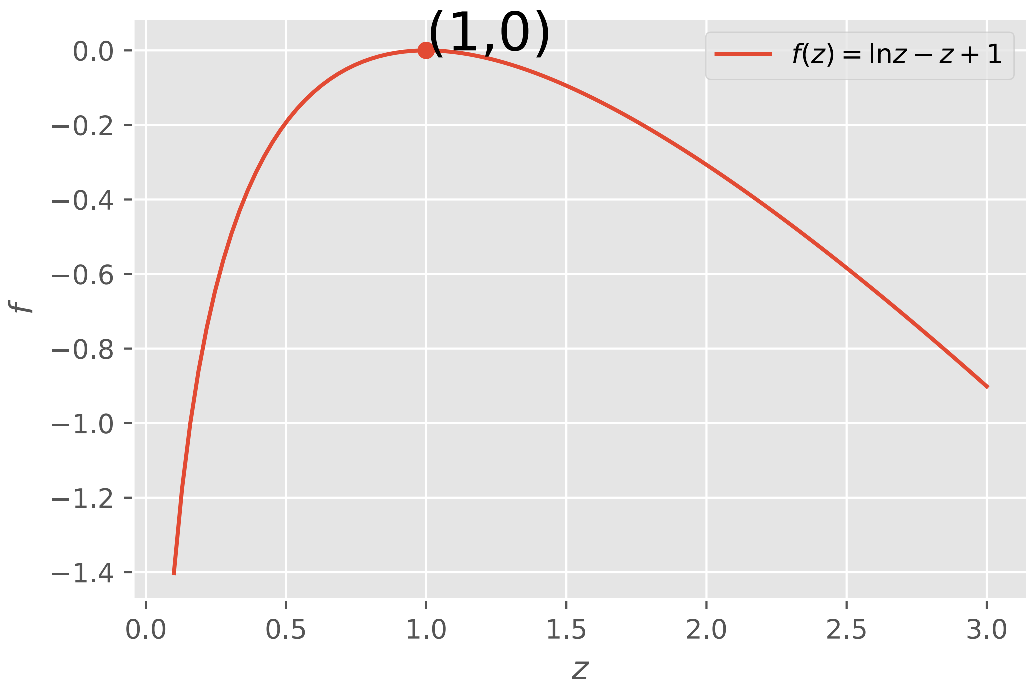
$$\frac{\mathrm{d}}{\mathrm{d}z}f(z)=\frac{1}{z}-1=\frac{1-z}{z}$$

The maximum of $f(z)$ is $f(1)=0$. Therefore:

$$\begin{aligned}&f(z) \le 0\\ \Leftrightarrow\ &\log z \le z-1\end{aligned}$$

Only if $z=1$, then $\log z = z-1$.

The figure of $f(z)$ is as follows:



(b)

Using the conclusion of (a):

$$\begin{aligned}\mathrm{KL}(p,q) &= \sum_i p_i \log(\frac{p_i}{q_i})\\ &= -\sum_i p_i \log(\frac{q_i}{p_i})\\ &\ge -\sum_i p_i(\frac{q_i}{p_i}-1)\\ &= \sum_i p_i - \sum_i q_i\\ &= 0\end{aligned}$$

In conclusion:

$$\mathrm{KL}(p,q) \ge 0$$

Only if $\frac{p_i}{q_i} = 1$ for all i, then$\mathrm{KL}(p,q)=0$.

(c)

Assumption:

- X represents if a student pass the exam, $X=1$ means pass, $X=0$ means not;
- E represents the test is easy;

- E' represents the student is smart;

Then assume that:

$$p: P(X=1|E)=0.8, P(X=0|E)=0.2;$$
$$q: P(X=1|E')=0.6, P(X=0|E')=0.4;$$

Using the formula:

$$\begin{aligned}
\mathrm{KL}(p,q) &= \sum_i p_i \log(\frac{p_i}{q_i}) \\
&= 0.8\log\frac{0.8}{0.6} + 0.2\log\frac{0.2}{0.4} \\
&\approx 0.091516 \\
\mathrm{KL}(q,p) &= \sum_i q_i \log(\frac{q_i}{p_i}) \\
&= 0.6\log\frac{0.6}{0.8} + 0.4\log\frac{0.4}{0.2} \\
&\approx 0.104650
\end{aligned}$$

Therefore, KL distance is not a symmetric function, which means that $\mathrm{KL}(p,q) \neq \mathrm{KL}(q,p)$.