Homework 2 – Xavier Gitiaux

March 2019

1 Exercise 1

Joint Distribution Denote C for coherence, D for difficulty, I for intelligent, G for grade, S for SAT, J for job, L for letter and H for happy. Then

$$Pr(C, D, G, H, I, J, L, S) = Pr(C)Pr(I)Pr(D|C)Pr(G|D, I)Pr(S|I)Pr(L|G)Pr(J|L, S)Pr((H|G, J)$$

Rewrite the joint distribution using factors

$$Pr(c, d, g, h, i, j, l, s) = f_C(c)f_I(i)$$

$$Pr(J = j) = \sum_{d,g,h,i,l,s} Pr(g|d, i)Pr(s|i)Pr(l|g)Pr(j|l, s)Pr(h|g, j)m_1(d) \tag{1}$$

where

$$m_1(d) = \sum_c Pr(c)Pr(d|c)$$

$$Pr(J = j) = \sum_{g,h,i,l,s} Pr(s|i)Pr(l|g)Pr(j|l, s)Pr(h|g, j)m_2(g, i) \tag{2}$$

where

$$m_2(g, i) = \sum_d m_1(d)Pr(g|d, i)$$

Then

$$Pr(J = j) = \sum_{g,h,l,s} Pr(l|g)Pr(j|l, s)Pr(h|g, j)m_3(g, s) \tag{3}$$

where

$$m_3(g, s) = \sum_i m_2(g, i)Pr(s|i)$$

Then

$$Pr(J = j) = \sum_{g,h,l} Pr(l|g)Pr(h|g, j)m_4(g, l, j) \tag{4}$$

where

$$m_4(g,l,j) = \sum_s m_3(g,s)Pr(j|l,s)$$

Then

$$Pr(J=j) = \sum_{g,h} Pr(h|g,j)m_5(g,j) \tag{5}$$

where

$$m_5(g,j) = \sum_l m_4(g,l,j)Pr(l|g)$$

We finish by doing

$$Pr(J=j) = \sum_g m_6(g,j)m_5(g,j) \tag{6}$$

$$m_6(g,j) = \sum_s Pr(h|g,j).$$

2 Exercise 2

2.1 Question (a)

$$\begin{aligned}
Pr[B|J=j,M=m] &\propto \sum_{a,e} Pr(a,B,e,j,m) \\
&\propto \sum_{a,e} Pr(e)Pr(B)Pr(A=a|B,E=e)Pr(J=j|A=a)Pr(M=m|A=a) \\
&\propto \sum_{a,e} Pr(e)Pr(B)Pr(A=a|B,E=e)Pr(J=j|A=a)Pr(M=m|A=a) \\
&\propto Pr(B)\sum_e Pr(e)\left(\sum_a Pr(A=a|B,E=e)Pr(J=j|A=a)Pr(M=m|A=a)\right) \\
&\propto Pr(B)\sum_e Pr(e)m(e,B,j,m)
\end{aligned} \tag{7}$$

where

$$m(e,B,j,m) = \sum_a Pr(A=a|B,E=e)Pr(J=j|A=a)Pr(M=m|A=a)$$

Therefore,

$$\begin{aligned}
m(1,1,1,1) &= Pr(A=1|B=1,E=1)Pr(J=1|A=1)Pr(M=1|A=1) \\
&\quad + Pr(A=0|B=1,E=1)Pr(J=1|A=0)Pr(M=1|A=0) \\
&= (0.95)(0.90)(0.70) + (0.05)(0.05)(0.01) \\
&= 0.598525
\end{aligned} \tag{8}$$

and

$$\begin{aligned}m(1,0,1,1) &= Pr(A=1|B=0,E=1)Pr(J=1|A=1)Pr(M=1|A=1) \\ &+ Pr(A=0|B=0,E=1)Pr(J=1|A=0)Pr(M=1|A=0) \\ &= (0.29)(0.90)(0.70) + (0.71)(0.05)(0.01) \\ &= 0.183055\end{aligned} \tag{9}$$

and

$$\begin{aligned}m(0,1,1,1) &= Pr(A=1|B=1,E=0)Pr(J=1|A=1)Pr(M=1|A=1) \\ &+ Pr(A=0|B=1,E=0)Pr(J=1|A=0)Pr(M=1|A=0) \\ &= (0.94)(0.90)(0.70) + (0.06)(0.05)(0.01) \\ &= 0.59223\end{aligned} \tag{10}$$

and

$$\begin{aligned}m(0,0,1,1) &= Pr(A=1|B=0,E=0)Pr(J=1|A=1)Pr(M=1|A=1) \\ &+ Pr(A=0|B=0,E=0)Pr(J=1|A=0)Pr(M=1|A=0) \\ &= (0.001)(0.90)(0.70) + (0.999)(0.05)(0.01) \\ &= 0.0011295\end{aligned} \tag{11}$$

Therefore

$$\begin{aligned}Pr[B=1|J=1,M=1] &\propto Pr(B=1)\left(Pr(E=1)m(1,1,1,1) + Pr(E=0)m(0,1,1,1)\right) \\ &= (0.01)\left((0.02)(0.598525) + (0.98)(0.59223)\right) \\ &= 0.005923559\end{aligned} \tag{12}$$

$$\begin{aligned}Pr[B=0|J=1,M=1] &\propto Pr(B=0)\left(Pr(E=1)m(1,0,1,1) + Pr(E=0)m(0,0,1,1)\right) \\ &= (0.01)\left((0.02)(0.183055) + (0.98)(0.0011295)\right) \\ &= 0.0000476801\end{aligned} \tag{13}$$

After renormalization,

$$Pr[B=1|J=1,M=1] = \frac{0.005923559}{0.005923559 + 0.0000476801} \sim 0.992$$

and

$$Pr[B=0|J=1,M=1] \sim 0.008.$$

2.2 Question b

With the elimination method, to compute each m, we need 4 multiplications and 1 addition so a total of 16 multiplications and 4 additions. To compute $P(B=b|J=1,M=1)$, we need 3 multiplications and 1 addition. Therefore, in total, we need 22 multiplications and 6 additions to get $P(B|J=1,M=1)$.

With the enumeration method, to compute $P(B=b|J=1,M=1)$, since there are two hidden binary variables A and E and since we need to compute 4 multiplications for each (a,e), in total we need 16 multiplications and 4 additions. Therefore to get $P(B|J=1,M=1)$, we need 32 multiplications and 8 additions.

2.3 Question c

We have

$$Pr(X_1, ..., X_n) = P(X_1)\prod_{i=2}^{n} Pr(X_i|X_{i-1}) \tag{14}$$

Therefore,

$$\begin{aligned}
Pr(X_n|X_1 = 1) &= Pr(X_1 = 1)\sum_{x_2,...,x_n}\prod_{i=2}^{n} Pr(x_i|x_{i-1}) \\
&= Pr(X_1 = 1)\sum_{x_4,...,x_n}\prod_{i=2}^{n} Pr(x_i|x_{i-1})\underbrace{\left(\sum_{x_2} Pr(x_3|x_2)Pr(x_2|X_1 = 1)\right)}_{=m_2(x_3)} \\
&= Pr(X_1 = 1)\sum_{x_4,...,x_n}\prod_{i=2}^{n} Pr(x_i|x_{i-1})m_2(x_3) \\
&= Pr(X_1 = 1)\sum_{x_5,...,x_n}\prod_{i=2}^{n} Pr(x_i|x_{i-1})m_3(x_4) \\
&... \\
&= Pr(X_1 = 1)m_n(x_n)
\end{aligned} \tag{15}$$

with

$$m_{k-1}(x_k) = \sum_{x_{k-1}} Pr(x_k|x_{k-1}m_{k-2}(x_{k-1})$$

Each $m_k(x)$, given the previous one, takes 2 multiplications and 1 addition so a total of 4 multiplications and 2 additions to compute m_k. Therefore, if we compute m_k recursively from m_2 to m_n, we need $4(n-1)$ multiplications and $2(n-1)$ additions, so $O(n)$ operations.

Using a brute force enumeration would have cost $n-1$ multiplications for each $\prod_{i=2}^{n} Pr(x_i|x_{i-1})$, so a total of $O(2^n)$ additions and $O(2^n(n-1))$ multiplications.

3 Exercise 3

3.1 Question (a)

The cdf is determined by $CDF(j) = P(X \leq x_j)$ for $j = 1, ..., k$. We can obtain the whole cdf with the following recursion

$$CDF(j) = \begin{cases} p_j + CDF(j-1) & \\ 0 & \text{if } j = 0 \end{cases} \tag{16}$$

By memoization, we can obtain $CDF(1)$, ... $CDF(k)$ in $O(k)$ times. Note that the resulting array CDF is sorted.

To obtain one random sample from the CDF, draw $u \sim U(0, 1)$, find j such that $CDF_j \leq u < CDF(j + 1)$ (assuming $CDF(0) = 0$) and return x_j. Since CDF is sorted, finding j will take $O(log(k))$ time.

3.2 Question (b)

Compute for $j = 1, ..., k$, $N_j = p_j N$, which takes a total of $O(k)$ operations. Then, write a sample with N_j copies of x_j for $j = 1, ..., k$. Writing the sample takes $N_1 + ... + N_k = N$ operations. Therefore the total running time is $O(N + k)$ and the per-sample time is $O\left(\frac{N+k}{N}\right) \approx O(1)$ since $k << N$.

4 Exercise 4

First,

$$P(H, S, R) = P(S)P(R)P(H|S, R) \tag{17}$$

Therefore we can compute the joint distribution as in table **??**

H	R	S	P(H, R, S)
1	1	1	0.007
1	1	0	0.0027
1	0	1	0.4851
1	0	0	0.0297
0	1	1	0
0	1	0	0.0003
0	0	1	0.2079
0	0	0	0.2673

Table 1: Joint Distribution (question 4)

Using the table **??**, we can compute

$$\begin{aligned} Pr(H = 1, S = 0) &= P(H = 1, S = 0, R = 1)P(R = 1) + P(H = 1, S = 0, R = 0)P(R = 0) \\ &= 0.0294 \end{aligned} \tag{18}$$

$$\begin{aligned} Pr(R = 1|H = 1, S = 0) &= \frac{P(H = 1, S = 0, R = 1)}{P(H = 1, S = 0)} \\ &\sim 0.0917 \end{aligned} \tag{19}$$

$$\begin{aligned} Pr(R = 1, H = 1) &= P(H = 1, S = 1, R = 1)P(S = 1) + P(H = 1, S = 0, R = 1)P(S = 0) \\ &= 0.00571 \end{aligned} \tag{20}$$

and

$$\begin{aligned} Pr(R = 1|H = 1) &= \frac{P(H = 1, R = 1)}{P(H = 1)} \\ &= \frac{0.00571}{0.5245} \sim 0.01089 \end{aligned} \tag{21}$$