Using 3 Late Days

HW4

Ammar Saldi…

1) $P(K) = \gamma \quad P(C) = 1-\gamma \quad P_K(x;\lambda) = \frac{2^x \lambda^x e^{-2\lambda}}{x!} \quad P_C(x;\mu) = \frac{\mu^x e^{-\mu/2}}{2^x x!}$

(a) $\mathcal{L}(\text{Dataset}) = \prod_{i \in \text{Kroger} \to K} \gamma P_K(x^{(i)} \mid z=K) \prod_{i \in \text{Costco} \to C} (1-\gamma) P_C(x^{(i)} \mid z=C)$

Likelihood

$$\boxed{\mathcal{L}(\text{Dataset}) = \prod_{i \in K} \gamma \frac{2^{x^{(i)}} \lambda^{x^{(i)}} e^{-2\lambda}}{x^{(i)}!} \prod_{i \in C} (1-\gamma) \frac{\mu^{x^{(i)}} e^{-\mu/2}}{2^{x^{(i)}} x^{(i)}!}}$$

(b)

$$\ln(\mathcal{L}(\text{Dataset})) = \ln \prod_{i \in K} \gamma \frac{2^{x^{(i)}} \lambda^{x^{(i)}} e^{-2\lambda}}{x^{(i)}!} \prod_{i \in C} (1-\gamma) \frac{\mu^{x^{(i)}} e^{-\mu/2}}{2^{x^{(i)}} x^{(i)}!}$$

$$= \sum_{i \in K} \ln(\gamma) + x^{(i)} \ln 2\lambda - 2\lambda - \ln(x^{(i)}!) + \sum_{i \in C} \ln(1-\gamma) + x^{(i)} \ln \frac{\mu}{2} - \frac{\mu}{2} - \ln x^{(i)}!$$

$$= \boxed{n_K \ln(\gamma) + m_K \ln(2\lambda) - 2\lambda n_K - S_K + n_C \ln(1-\gamma) + m_C \ln\left(\frac{\mu}{2}\right) - n_C \frac{\mu}{2} - S_C}$$

(c) $\frac{\partial}{\partial \gamma} \ln(\mathcal{L}(\text{Dataset})) = \frac{n_K}{\gamma} + \frac{n_C}{1-\gamma} = 0$

$n_K(1-\gamma) = n_C \gamma \rightarrow \frac{n_K}{n_K + n_C} = \gamma$ prob Kroger

$n_K = 4$

$n_C = 6$

$$\boxed{\gamma = \frac{4}{10} = \frac{2}{5} \qquad 1-\gamma = \frac{3}{5}}$$ → prob Costco

$\frac{\partial}{\partial \lambda} \ln(\mathcal{L}(\text{Dataset})) = \frac{m_K}{\lambda} - 2n_K = 0$

$\frac{m_K}{2n_K} = \lambda$

$\lambda = \frac{\sum_{i \in K} x^{(i)}}{2(4)} = \boxed{\frac{23}{8}}$

$\frac{\partial}{\partial \mu} \ln(\mathcal{L}(\text{Dataset})) = \frac{m_C}{\mu} - \frac{n_C}{2} = 0$

$\frac{2m_C}{n_C} = \mu = \frac{2 \sum_{i \in C} x^{(i)}}{n_C} = \frac{2 \cdot 21}{6} = \frac{21}{3} = 7$

$\boxed{\mu = 7}$

2 Expectation Maximization

(a) Use Bayes Rule $\rightarrow$ $p(z^{(i)}=k \mid x^{(i)}; \theta^{(t)}) = \dfrac{p(\bar{x}^{(i)} \mid z^{(i)}=k; \bar{\theta}^{(t)})\, p(z^{(i)}=k \mid \bar{\theta}^{(t)})}{p(\bar{x}^{(i)} \mid \bar{\theta}^{(t)})}$

$p(\bar{x}^{(i)} \mid \bar{\theta}^{(t)}) = p(\bar{x}^{(i)} \mid z^{(i)}=A; \bar{\theta}^{(t)})\, p(z^{(i)}=A \mid \bar{\theta}^{(t)}) + p(\bar{x}^{(i)} \mid z^{(i)}=B; \bar{\theta}^{(t)})\, p(z^{(i)}=B \mid \bar{\theta}^{(t)})$

Since: $p(z^{(i)}=A \mid \bar{\theta}^{(t)}) = p(z^{(i)}=B \mid \bar{\theta}^{(t)}) = \frac{1}{2}$

$p(z^{(i)}=k \mid \bar{x}^{(i)}; \bar{\theta}^{t}) = \dfrac{p(\bar{x}^{(i)} \mid z^{(i)}=k; \bar{\theta}^{(t)})\,\frac{1}{2}}{\frac{1}{2}p(x^{(i)} \mid z^{(i)}=A; \bar{\theta}^{(t)}) + \frac{1}{2}p(\bar{x}^{(i)} \mid z^{(i)}=B; \bar{\theta}^{(t)})} = \dfrac{p(\bar{x}^{(i)} \mid z^{(i)}=k; \bar{\theta}^{(t)})}{p(\bar{x}^{(i)} \mid z^{(i)}=A; \bar{\theta}^{(t)}) + p(\bar{x}^{(i)} \mid z^{(i)}=B; \bar{\theta}^{(t)})}$

Sub in $p(x^{(i)} \mid z^{(i)}=k; \theta^{(t)}) = \theta_{k_s}^{s_i}\theta_{k_c}^{c_i}\theta_{k_r}^{r_i}$

$$p(z^{(i)}=k \mid \bar{x}^{(i)}; \theta^{(t)}) = \boxed{\dfrac{\theta_{k_s}^{s_i}\theta_{k_c}^{c_i}\theta_{k_r}^{r_i}}{\theta_{A_s}^{s_i}\theta_{A_c}^{c_i}\theta_{A_r}^{r_i} + \theta_{B_s}^{s_i}\theta_{B_c}^{c_i}\theta_{B_r}^{r_i}}}$$

(b) $\theta^{(0)} = [.5, .3, .2, .2, .3, .5]$ $M^1 \rightarrow s_1 = 4,\ c_1 = 2,\ r_1 = 4$

$P(z^{(1)}=A \mid x^{(1)}; \bar{\theta}^{(0)}) = \dfrac{.5^4 . 3^2 . 2^4}{.5^4 . 3^2 . 2^4 + .2^4 . 3^2 . 5^4} = .5$ $\quad P(z^{(1)}=B \mid x^{(1)}; \bar{\theta}^{(0)}) = 1-.5 = .5$

$z_{(1)}^{(1)} = A$ tie pick lexicographical order ($A$)

month $\rightarrow$ $M^2 \rightarrow s_2 = 7,\ c_2 = 7,\ r_2 = 7$ $\quad P(z^{(2)} \mid x^{(2)}; \theta^{(0)}) = \dfrac{.5^7 . 3^1 . 2^2}{.5^7 . 3^1 . 2^2 + .2^7 . 3^1 . 5^2} = .99$

$P(z^{(2)}=B \mid x^{(2)}; \bar{\theta}^{0}) = 1-.99 = .01$

$z_{(1)}^{(2)} = A$

$M^3 \rightarrow s_3 = 1,\ c_3 = 2,\ r_3 = 7$ $\quad P(z^{(3)}=A \mid x^{(3)}; \theta^{(0)}) = \dfrac{.5^1 . 3^2 . 2^7}{.5^1 . 3^2 . 2^7 + .2^1 . 3^2 . 5^7} = .004$

$P(z^{(3)}=B \mid x^{(3)}; \bar{\theta}^{(0)}) = 1-.004 = .996$, $z_{(1)}^{(3)} = B$

$M^4 \rightarrow s_4 = 8,\ c_4 = 1,\ r_4 = 1$ $\quad P(z^{(4)}=A \mid x^{(4)}; \bar{\theta}^{(0)}) = \dfrac{.5^8 . 3^1 . 2^1}{.5^8 . 3^1 . 2^1 + .2^8 . 3^1 . 5^1} = .998$ $\quad P(z^4=B \mid x^4; \bar{\theta}^0) = 1-.998 = .002$

$z_{(1)}^{(1)} = A$

$M^5 \rightarrow s_5 = 3,\ c_5 = 3,\ r_5 = 4$ $\quad P(z^5=A \mid x^5; \theta^0) = \dfrac{.5^3 . 3^3 . 2^4}{.5^3 . 3^3 . 2^4 + .2^3 . 3^3 . 5^4} = .286$ $\quad P(z^5=B \mid x^5; \theta^0) = 1-.286 = .714$

$z_{(1)}^{(1)} = B$

$M^5 \rightarrow s_6 = 3,\ c_6 = 2,\ r_6 = 5$ $\quad P(z^6=A \mid x^6; \theta^0) = \dfrac{.5^3 . 3^2 . 2^5}{.5^3 . 3^2 . 2^5 + .2^3 . 3^2 . 5^5} = .138$ $\quad P(z^6=B \mid x^6; \theta^0) = 1-.138 = .862$

$z_{(1)}^{(6)} = B$

$$\boxed{z_{(1)} = [A, A, B, A, B, B] \quad \bar{\theta}^{(1)} = \frac{1}{30}[19, 4, 7, 7, 7, 16]}$$

(R=)

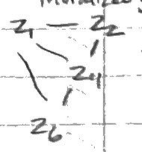
1. Draw Ancestral Graph – only vars mentioned and all ancestors
2. Moralize by marrying the parents
3. Disorient → directed with undirected
4. Delete the givens and their edges
5. 1) Vars disconnected → independent
2) not independent if connected

Bayes Net ass$^n$: Each node cond independent of its non-descendents given parents

$P(A,B|C)$
$= P(A|C)P(B|C)$
$P(A|B,C) = P(A|C)$

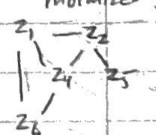
3 (a) $\exists z_x \; P(z_1, z_6 | z_x) = P(z_1|z_x)P(z_6|z_x)$?

moralized graph

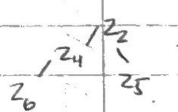



→ No, there exists no node that will make $z_1$ and $z_6$ conditionally independent because there is an edge that directly connects the two, so when you run D-separation they will still be connected and thus not conditionally ind.

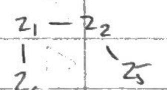
moralized-graph



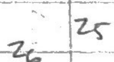

→ No, $z_5$ and $z_6$ are not marginally independent because there exists a path from one to the other in the moralized graph.




→ $z_5, z_6 | z_1$? No, not conditionally ind given $z_1$ as even with $z_1$ removed along with its connections from the graph, there is a path from $z_6$ to $z_5$. All paths from z5 to z6 do not go through z1.




→ $z_5, z_6 | z_4$? No, not conditionally ind given z4 as even with z4 removed along with its connections from the graph, there is a path from $z_6$ to $z_5$. All paths from z5 to z6 do not go through z4.




→ Yes, Given both $z_4$ and $z_1$, $z_5$ and $z_6$ are conditionally independent as there is no path b/w them once conditioned vars and edges removed.

(b) $P(z_1, z_2, z_3, z_4, z_5, z_6) = P(z_1)P(z_2|z_1)P(z_4|z_1,z_2)P(z_5|z_2)P(z_3|z_1)P(z_6|z_1,z_4)$

$\Pi$ all vars given their parents using Bayes net Ass$^n$

(c) $P(z_1, z_2, z_3, z_4, z_5, z_6) = P(z_1)P(z_2|z_1) \cdot P(z_3|z_1)P(z_4|z_1,z_2) \cdot P(z_5|z_2)P(z_6|z_1,z_4)$

$= n-1 + n(n-1) + n(n-1) + n(n-1) + n^2(n-1) + n^2(n-1)$

$= n-1 + 3n(n-1) + 2n^2(n-1) = n-1 + 3n^2 - 3n + 2n^3 - 2n^2 = \boxed{1n^2 + 2n^3 - 2n - 1}$

4) $P(R=T \mid H=T) = \dfrac{P(R=T, H=T)}{P(H=T)} = \dfrac{\sum P(R=T, H=T, W, S)}{P(H=T, S, W, R)} = \boxed{\dfrac{.2781}{.6471} = .42976}$

↳ bottom page

$$P(R=T, H=T, W, S) = \sum_S \sum_W P(S=s)P(R=T \mid S=s)P(W=w \mid S=s)P(H=T \mid R=T, W=w)$$

$$= \sum_w \tfrac{1}{2}\left(P(R=T \mid S=T)P(W=w \mid S=T)P(H=T \mid R=T, W=w)\right) + \tfrac{1}{2}\left(P(R=T \mid S=F)P(W=w \mid S=F)P(H=T \mid R=T, W=w)\right)$$

$$= \tfrac{1}{2}\tfrac{1}{10} P(W=F \mid S=T)P(H=T \mid R=T, W=F) + \tfrac{1}{2}\tfrac{1}{10} P(W=T \mid S=T)P(H=T \mid R=T, W=T)$$
$$+ \tfrac{1}{2}\tfrac{1}{2} P(W=F \mid S=F)P(H=T \mid R=T, W=F) + \tfrac{1}{2}\tfrac{1}{2} P(W=T \mid S=F)P(H=T \mid R=T, W=T)$$

$$= \tfrac{1}{2}\tfrac{1}{10}\tfrac{1}{5}\tfrac{9}{10} + \tfrac{1}{2}\tfrac{1}{10}\tfrac{4}{5}\tfrac{99}{100} + \tfrac{1}{2}\tfrac{1}{2}\tfrac{4}{5}\tfrac{9}{100} + \tfrac{1}{2}\tfrac{1}{2}\tfrac{1}{5}\tfrac{99}{100} = \underline{.2781}$$ ✓ Numerator

$$P(H=T, S, W, R) = \sum \sum_W \sum_R P(S=s)P(W=w \mid S=s)P(R=r \mid S=s)P(H=T \mid W=w, R=r)$$

$$= \tfrac{1}{2}\Big[\sum_W \sum_R P(W=w \mid S=T)P(R=r \mid S=T)P(H=T \mid W=w, R=r) + P(W=w \mid S=F)P(R=r \mid S=F)P(H=T \mid W=w, R=r)\Big]$$

$$= \tfrac{1}{2}\Big[\sum_w P(W=w \mid S=T)P(R=T \mid S=T)P(H=T \mid W=w, R=T) + P(W=w \mid S=T)P(R=F \mid S=T)P(H=T \mid W=w, R=F)$$
$$+ P(W=w \mid S=F)P(R=T \mid S=F)P(H=T \mid W=w, R=T) + P(W=w \mid S=F)P(R=F \mid S=F)P(H=T \mid W=w, R=F)\Big]$$

$$= \tfrac{1}{2}\Big[P(W=T \mid S=T)P(R=T \mid S=T)P(H=T \mid W=T, R=T) + P(W=F \mid S=T)P(R=T \mid S=T)P(H=T \mid W=F, R=T)$$
$$+ P(W=T \mid S=T)P(R=F \mid S=T)P(H=T \mid W=T, R=F) + P(W=F \mid S=T)P(R=F \mid S=T)P(H=T \mid W=F, R=F)$$
$$+ P(W=T \mid S=F)P(R=T \mid S=F)P(H=T \mid W=T, R=T) + P(W=F \mid S=F)P(R=T \mid S=F)P(H=T \mid W=F, R=T)$$
$$+ P(W=F \mid S=F)P(R=F \mid S=F)P(H=T \mid W=T, R=T) + P(W=F \mid S=F)P(R=F \mid S=F)P(H=T \mid W=F, R=F)\Big]$$

$$= \tfrac{1}{2}\Big[.8 \cdot .1 \cdot .99 + .2 \cdot .1 \cdot .9 + .8 \cdot .9 \cdot .99 + .2 \cdot .9 \cdot 0 + .2 \cdot .5 \cdot .99 + .8 \cdot .5 \cdot .9 + .2 \cdot .5 \cdot .9 + 0\Big] = \underline{.6471}$$

Denominator