# Problem 1

## Intuition

Goal: use $Y^* (Y_1^*, \dots, Y_{1000}^*)$ to learn $\sigma^2$ and $q$

$$p(x,y) = \frac{1}{2}\left(\prod_{i=1}^{1000} f_i(y_i|x_i)\right)\left(\prod_{k=1}^{999} a(x_{k+1}, x_k)\right)$$

$\hookrightarrow$ depends on $\sigma^2$ $\quad\quad$ $\llcorner$ depends on $q$

Without using EM:

$\boxed{\text{We have}}$ $Y^*$, $P_{xy}$

$\boxed{\text{So}}$, we could find the marginal dist of $y$ and pick $\sigma^2$ and $q$ to maximize the likelihood of observe data

$\boxed{\text{Essentially}}$, $P_{xy}(x,y) \Rightarrow \sum_x P_{xy}(x,y) = P_y(Y^*)$

$$\max_{\sigma^2, q} \log\left(\sum_x P_{xy}^{q\sigma^2}(x, Y^*)\right)$$

$\boxed{\text{Observe}}$, this is really really computationally heavy!!! So we EM.

EM

① how to define $q_{k+1}$ $\sigma^2_{k+1}$?

$\to$ Note: we know this

$\to$ Note: we don't know, our goal

$$(q_{k+1}, \sigma^2_{k+1}) = \underset{(q,\sigma^2)}{\arg\max}\ \mathbb{E}_{X|Y^*}^{q_k,\sigma_k^2}\left[\log\left(P_{xy}^{q\sigma^2}(x, y^*)\right)\right]$$

② Let's dissect what's in the expectation.

$$\log\left(P^{q\sigma^2}(x,y)\right) = \log\left(\tfrac{1}{2}\right) - 1000\log(2\pi\sigma^2) + \sum_{i=1}^{1000}\left(\frac{-(x_i - y_i)^2}{2\sigma^2}\right) + \sum_{i=1}^{999}\log\left(a(x_i, x_{i+1})\right)$$

③ Let's take expectation of this ②

$$E_{X|Y^*}^{q_k\sigma_k^2}\left[-1000\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{1000}(x_i - y_i)^2 + \sum\log\left(q\mathbb{1}_{x_i = x_{i+1}} + (1-q)\mathbb{1}_{x_i \neq x_{i+1}}\right)\right]$$

$$\Updownarrow$$

$$\underbrace{-500\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{1000} E_{X|Y^*}^{q_k\sigma_k^2}\left[(x_i - Y_i^*)^2\right]}_{\text{depends only on } \sigma^2} + \underbrace{\sum_{i=1}^{999} E_{X|Y=y^*}^{q_k\sigma_k^2}\left[\log\left(q\mathbb{1}_{x_i = x_{i+1}} + (1-q)\mathbb{1}_{x_i \neq x_{i+1}}\right)\right]}_{\text{depends only on } q}$$

$\boxed{\text{Note}}$

④ Find expression that minimize $E[\,]$:

For $\sigma^2$:

① $A := \sum_{i=1}^{1000} E_{X|Y=y^*}^{q_k\sigma_k^2}\left[(x_i - Y_i^*)^2\right]$

② $\max_{\sigma^2} -1000\log(\sigma^2) - \frac{1}{2\sigma^2}A$

③ take the derivative of ② in respect to $\sigma^2$ and set to 0

$$-\frac{1000}{\sigma^2} + \frac{A}{2\sigma^4} = 0$$

$$\frac{A}{2\sigma^4} = \frac{1000}{\sigma^2}$$

$$\sigma^2 = \frac{A}{2000}$$

For q

① $\sum_{i=1}^{999} E_{X_i|Y=y^*}^{q_n \, \sigma_n} [\underbrace{\log(q \mathbb{I}_{X_i = X_{i+1}} + (1-q)\mathbb{I}_{X_i \neq X_{i+1}})}_{\Downarrow \text{ simplify}}]$

$$\log(q)\mathbb{I}_{X_i = X_{i+1}} + \log(1-q)\mathbb{I}_{X_i \neq X_{i+1}}$$

$$= \log(q) \underbrace{\sum_{i=1}^{999} E_{X|y^*}^{\sigma_n^2 q^k} [\mathbb{I}_{X_i = X_{i+1}}]}_{\text{denote as } B} + \log(1-q)(999 - B)$$

② $\max_q \log(q) B + \log(1-q)(999 - B)$

③ take the derivative of ② in respect to q and set to 0

$$\frac{B}{q} - \frac{999 - B}{1-q} = 0$$

$$\frac{1-q}{q} = \frac{999-B}{B} \Rightarrow \frac{1}{q} - 1 = \frac{999}{B} - 1$$

$$q = \frac{B}{999}$$

⑤ Let's compute A and B,

For A

$$E[(X_i - Y_i^*)^2 \mid Y_i = Y_i^*] = (1-y_i^*)^2 P(X_i = 1 \mid Y_i = y_i^*) + (0 - Y_i^*)^2 P(X_i = 0 \mid Y_i = y_i^*)$$

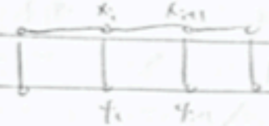
$\boxed{\text{note}}$ $P_{X|Y} = \frac{P_{Y|X} P_X}{P_Y}$ In hw5 we showed $P_X$ is invariant meas~ (...) s.t. $P_X = \frac{1}{2}$

$$P_{X|Y} = \frac{P_{Y|X=0}}{P_{Y|X=0} + P_{Y|X=1}}$$

For B

$E^{6,9}_{x|y}[\mathbb{I}_{x_i = x_{i+1}}]$

Note




① $E^{6,9}_{x|y}[\mathbb{I}_{x_i = x_{i+1}}] = \mathbb{I}_{(x_i=0)=(x_{i+1}=0)} P(x_i, x_{i+1} \mid y_i^*, y_{i+1}^*)$
$+ \mathbb{I}_{(x_i=1)=(x_{i+1}=1)} P(x_i, x_{i+1} \mid y_i^* \; y_{i+1})$

↳ expression we want to solve

$\underline{\hspace{6cm}}$

$\mathbb{I}_{0=0} P + \mathbb{I}_{1=1} P + \mathbb{I}_{1\neq 0} P + \mathbb{I}_{0\neq 1} P$

② lets break $P(x_i, x_{i+1} \mid y_i^*, y_{i+1}^*)$ into subproblem

please omit

$P(x_i \mid y_i^* \; y_{i+1}^*) \; P(x_{i+1} \mid x_i \; y_i^* \; y_{i+1}^*)$ ← deduce using bayes rule

$\boxed{\text{For } P(x_i \mid y_i^* \; y_{i+1}^*)}$, we can observe from the graph that it could be rewritten as $P(x_i \mid y_i^*)$.

This is similar to $\boxed{\text{For A}}$

$$P_{x|y} = \frac{P_{y|x}}{\sum_x P_{y|x}}$$

For $P(x_{i+1} \mid y_i^*, y_{i+1}^*, x_i^*)$

$$P_{x_{i+1} \mid y_i^*, y_{i+1}^*, x_i} = \frac{P_{y_i^*, y_{i+1}^*, x_i \mid x_{i+1}} \; P_{x_{i+1}}}{\underbrace{P_{y_i^* \, y_{i+1} \, x_i}}_{\text{constant}}}$$

$$P_{x_{i+1}=0 \mid y_i^* y_{i+1}^* x_i=0} = \frac{P_{y_i^* y_{i+1}^* x_i=0 \mid x_{i+1}=0}}{P_{y_i^* y_{i+1} x_i \mid x_{i+1}=0} + P_{y_i y_{i+1} x_i \mid x_{i+1}=1}} = a(x_i=0)=(x_{i+1}=0) f(y_i \mid$$

$$P_{x_{i+1}=1 \mid y_i \, y_{i+1} \, x_i=1} = \frac{P_{y^* \, y_{i+1} \, x=1 \mid x_{i+1}=1}}{P_{y_i \, y_{i+1} \, x_i \mid x_{i+1}=0} + P_{y_i \, y_{i+1} \, x_i \mid x_{i+1}=1}}$$

$P(x_i, x_{i+1} \mid y_i \, y_{i+1}) = P(x_i \mid y_i \, y_{i+1}) \, P(x_{i+1} \mid x_i \, y_i \, y_{i+1}) \propto a(x_i = x_{i+1}) f(x_i \mid y_i) f(x_{i+1} \mid y_{i+1}$

$$P(x_i, x_{i+1} \mid y_i \, y_{i+1}) = \frac{a(x_i = x_{i+1}) f(x_i \mid y_i) f(x_{i+1} \mid y_{i+1})}{\sum a() f() f()}$$

## Computation

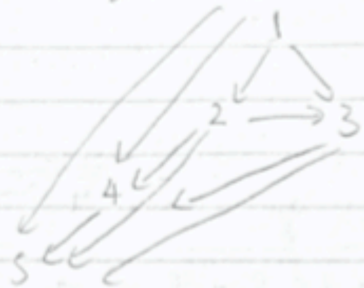
```r
library("readxl")
y <- read_excel("yvalues.xlsx",col_names=FALSE)
y<-as.vector(y$...1)

f_y_x<-function(x,y,s){
  (1/sqrt(2*pi*s))*exp(-(x-y)^2/(2*s))
}

A<-function(y,sigma_k){
  a=((1-y)^2*f_y_x(1,y,sigma_k))/(f_y_x(1,y,sigma_k)+f_y_x(0,y,sigma_k))
  b=((0-y)^2*f_y_x(0,y,sigma_k))/(f_y_x(1,y,sigma_k)+f_y_x(0,y,sigma_k))
  sum(a+b)
}

a<-function(q,x_i,x_k){
  q*ifelse(x_i==x_k,1,0)+(1-q)*ifelse(x_i!=x_k,1,0)
}
sigma_2<-1
max_q<-0.5
for (i in 1:1000) {
  sigma_2<-sum(A(y,sigma_2))/1000
  upper<-a(max_q,1,1)*f_y_x(1,y,sigma_2)[1:999]*f_y_x(1,y,sigma_2)[2:1000]+
    a(max_q,0,0)*f_y_x(0,y,sigma_2)[1:999]*f_y_x(0,y,sigma_2)[2:1000]
  down<-a(max_q,1,1)*f_y_x(1,y,sigma_2)[1:999]*f_y_x(1,y,sigma_2)[2:1000]+
    a(max_q,0,0)*f_y_x(0,y,sigma_2)[1:999]*f_y_x(0,y,sigma_2)[2:1000] +
    a(max_q,0,1)*f_y_x(0,y,sigma_2)[1:999]*f_y_x(1,y,sigma_2)[2:1000]+
    a(max_q,1,0)*f_y_x(1,y,sigma_2)[1:999]*f_y_x(0,y,sigma_2)[2:1000]
  p<-upper/down
  max_q<-sum(p)/999
}
max_q
sigma_2
```

```
> max_q
[1] 0.950075
> sigma_2
[1] 0.9949241
```

# Problem 2

Q2 Yes

$P_1 P_{2|1} P_{3|1,2} P_{4|1,2,3} P_{5|1,2,3,4} \cdots = \prod P_{i|1:i}$




This is a dag because it will never form a cycle as child-n will never be the condition for parent.

This also forms joint distribution

$P_{2|1} = \frac{P_{21}}{P_1}$

$\frac{P_{21}}{P_1} \cdot P_1 = P_{21}$ $\Big\}$ $P_{12} = P_1 P_{2|1}$

$P_{3|21} = \frac{P_{321}}{P_{21}}$

$\frac{P_{321}}{P_{21}} \cdot P_{21} = P_{321}$ $\Big\}$ $P_{321} = P_{3|21} \cdot P_{2|1} P_1$

$\prod_{i=1}^{n} P_{i|1:i} = P(X_1, \ldots X_n)$

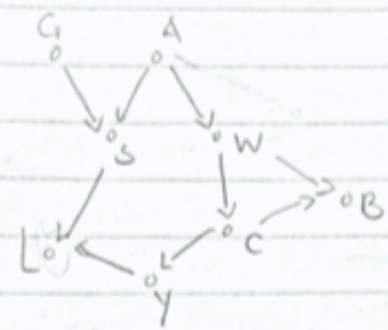
# Problem 3

Q3, Gender (G), smoking habit (S), severity of Angina (A), age (Y), Weight (W), Blood pressure (B), Cholesterol level (C), and Lung (L)

$P = P_G \cdot P_A \cdot P_{S|A,G} \cdot P_{W|A} \cdot P_{C|W} P_{B|C,W} P_{Y|C} P_{L|S,Y}$

↳ joint dist

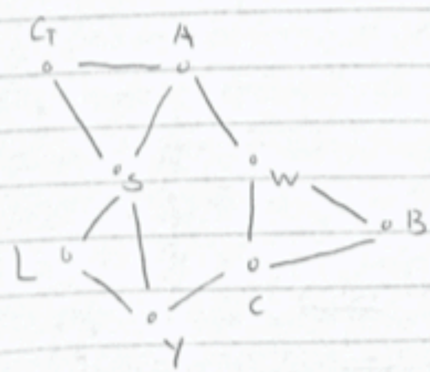
a,




b, moral graph

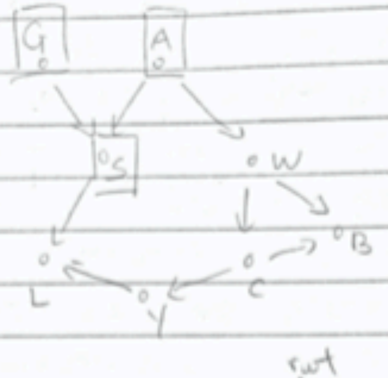
Definition: Let $G = (V, E)$ be a DAG. The "moral graph" of $G$ is the graph $G' = (V, E')$ (undirected Graph)

- $\{u, v\} \in E'$ if $u \to v$
- $\{u, v\} \in E'$ if $u$ ; $v$ have common child



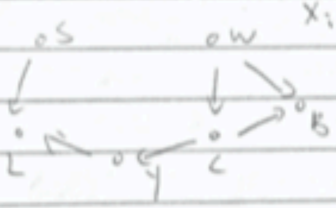
C, i conditional dist of (L, Y, C, B, W) given G=g, A=a,
S=s

$P(L,Y,C,B,W) \mid G=g, A=a, S=s$




Note for BN's: conditional rule: can only condition on root

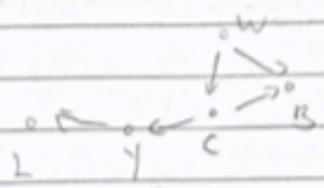
$$\text{step 1} \quad P(L,Y,C,B,W) \mid \overbrace{\underline{\underline{G=g}}, \underline{\underline{A=a}}, \underline{S=s}}^{\text{root}}$$




$X_i \mid G=g, A=a$

$$\text{step 2} \quad P_{LYCBW \mid G=g, A=a, S=s} = P_{W|A} P_{B|CW} P_{Y|C} P_{L|SY}$$

random

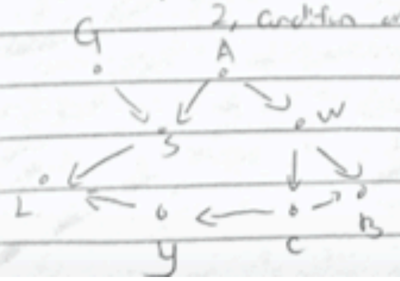



:: $P_{CWB \mid GAS}$

Strategy 1. Marginalize over leaves

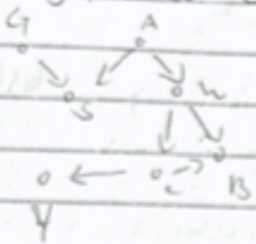
2. Condition on roots

Note for BN: can only marginalize over leaves

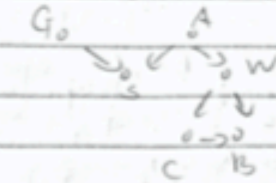


step1 marginalize over L and y

① marginalize over L b/c L is a leave




② now L is a leaf so marginalize over y

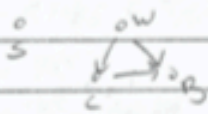



$X_{GAWCB}$

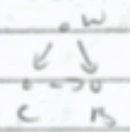
step2 apply conditioning rule

① conditioning over wrt G, A

$P_{CWB|GAS}$




② conditioning over S

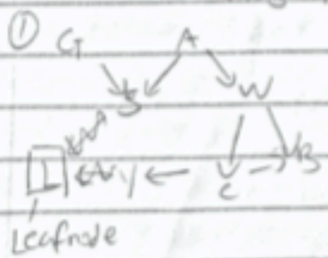


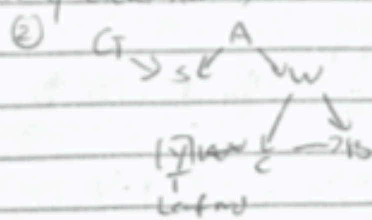

iii,

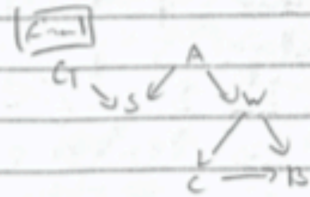
$P_{GASWCB}$

strategies ① marginalize over L (leaf node)

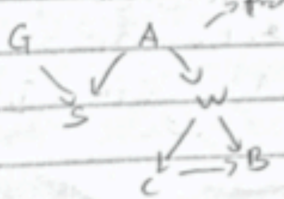
② marginalize over y (leaf node)



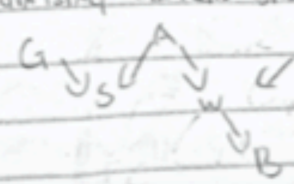

iv $P_{GASWB}$ → from previous, we can marginalize over C




$\sum_C P_G P_A P_{S|AG} P_{W|A} P_{C|W} P_{B|WC} = P_G P_A P_{S|AG} P_{W|A} \sum_C P_{C|W} P_{B|WC}$

$= P_G P_A P_{S|AG} P_{W|A} \sum_C P_{BC|W}$

$= P_G P_A P_{S|AG} P_{W|A} P_{B|W}$



d, $P(L=2 \mid S \geq 30) = \underbrace{\frac{P_{LS}}{P_S}}_{\text{Bayes rule}}$

$$\sum_L \sum_Y \sum_C \sum_W \sum_G \sum_A P_A P_G P_{S|AG} P_{W|A} P_{C|W} P_{Y|C} P_{L=2|SY} = P_{LS}$$

$30 \cdot 94 \cdot 270 \cdot 340 \cdot 2 \cdot 5 \approx 10^7$ sums

$$\sum_S \sum_A \sum_G P_{S|AG} P_A P_G = P_S$$

$34 \cdot 5 \cdot 2 = 340$ sums

$$\frac{P_{LS}}{P_S} = P(L=2 \mid S \geq 30)$$

e, $P(C \geq 120 \mid Y=50, A \geq 3, S=[20,30]) = \frac{P_{CYAS}}{P_{YAS}}$

$$\sum_C \sum_S \sum_A \sum_W \sum_G P_A P_{S|AG} P_G P_{W|A} P_{W|C} P_{C|Y=50} = P_{CYAS}$$

$230 \cdot 10 \cdot 2 \cdot 340 \cdot 2 = 7820$ sums

$$\sum_C \sum_S \sum_A \sum_W \sum_G P_A P_G P_{S|AG} P_{W|A} P_{W|C} P_{C|Y=50} = P_{YAS}$$

$270 \cdot 10 \cdot 2 \cdot 340 \cdot 2 = 3{,}670{,}000$ sums

Combine these two $\frac{P(C \geq 120, Y=50, A \geq 3, S=[20,30])}{P(Y=50, A \geq 3, S=[20,30])} =$

f $P(L \geq 1 \mid Y=50, A \geq 3, S=[20,30]) = \frac{P_{LYAS}}{P_{YAS}}$

$$\sum_L \sum_S \sum_C \sum_W \sum_A \sum_G P_A P_G P_{S|AG} P_{W|A} P_{W|C} P_{C|Y=50} P_{L|SY} = P_{LYAS}$$

$2 \cdot 10 \cdot 270 \cdot 340 \cdot 2 \cdot 2 \approx 10^6$ sums

we could get $P_{YAS}$ from $\boxed{e}$ so

$$\frac{P_{LYAS}}{P_{YAS}} = P(L \geq 1 \mid Y=50, A \geq 3, S=[20,30])$$