You can find the handwriting portions with larger image at page 10 and onward. This is to help to have a clearer image to read my handwriting

Problem 1

Intuition

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

$$p(x,y) = \frac{1}{2}\Big(\underbrace{\prod_{i=1}^{1000} f_i(y_i|x_i)}_{\text{depends on } \sigma^2}\Big)\Big(\underbrace{\prod_{k=1}^{999} a(x_{k+1}, x_k)}_{\text{depends on } q}\Big)$$

Without using EM:

We have: Y^*, P_{xy}

So, we could find the marginal dist of Y and pick σ^2 and q to maximize the likelihood of observe data

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

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

Observe, this is really really computationally heavy!!! So we EM.

EM

① how to define q_{k+1} σ^2_{k+1}?

$$(q_{k+1}, \sigma^2_{k+1}) = \operatorname*{argmax}_{(q,\sigma^2)} \mathbb{E}_{\underbrace{x|y^*}}^{\overbrace{q_k, \sigma_k^2}^{\text{Note: we know this}}}\Big[\log\big(P_{xy}^{\overbrace{q\sigma^2}^{\text{Note: we don't know, our goal}}}(x, y^*)\big)\Big]$$

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

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

③ Let's take expectation of this ②

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

$$\Updownarrow$$

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

Note

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

For σ^2:

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

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

③ take the derivative of ② in respect to 6^2 and set to 0

$$-\frac{1000}{6^2} + \frac{A}{26^4} = 0$$

$$\frac{A}{26^4} = \frac{1000}{6^2}$$

$$6^2 = \frac{A}{2000}$$

For q

① $\sum_{i=1}^{999} E_{X_1 Y = y^*}^{q_n 6_n} \left[\underbrace{\log(q \mathbb{1}_{x_i = x_{i+1}} + (1-q)\mathbb{1}_{x_i \neq x_{i+1}})}\right]$

⇓ simplify

$\log(q)\mathbb{1}_{x_i = x_{i+1}} + \log(1-q)\mathbb{1}_{x_i \neq x_{i+1}}$

$$= \log(q) \underbrace{\sum_{i=1}^{999} E_{X_1 Y = y^*}^{6_n^2 q^k} [\mathbb{1}_{x_i = x_{i+1}}]}_{\text{denote as } B} + \log(1-q)(999 - B)$$

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

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

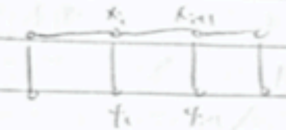
note $P_{X|Y} = \frac{P_{Y|X} P_X}{P_Y}$ In hw5 we showed P_X is invariant meas~ ... s.t $P_X = \frac{1}{2}$

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

For B

$$E^{\hat{\sigma},q}_{X|Y}[\mathbb{1}_{x_i = x_{i+1}}]$$

$\boxed{\text{Note}}$



① $E^{\hat{\sigma},q}_{X|Y}[\mathbb{1}_{x_i = x_{i+1}}] = \mathbb{1}_{(x_i=0)=(x_{i+1}=0)} P(x_i, x_{i+1} \mid Y_i^*, Y_{i+1}^*)$
$+ \mathbb{1}_{(x_i=1)=(x_{i+1}=1)} P(x_i, x_{i+1} \mid Y_i^*\ Y_{i+1})$

↳ expression we want to solve

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

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 $P(X_{i+1} \mid Y_i^*, Y_{i+1}^*, X_i^*)$

$$P_{X_{i+1} \mid Y_i^*, Y_{i+1}^*, X_i} = \frac{P_{Y_i^*, Y_{i+1}^*, X_i \mid X_{i+1}}\ P_{X_{i+1}}^{\to \frac{1}{2}}}{\underbrace{P_{Y_i^*\ Y_{i+1}\ X_i}}_{\text{constant}}}$$

$$P_{X_{i+1}=0 \mid Y_i^*\ Y_{i+1}^*\ X_i=0} = \frac{P_{Y_i^*\ Y_{i+1}^*\ X_i=0 \mid X_{i+1}=0}}{P_{Y_i^*\ Y_{i+1}\ X_i \mid X_{i+1}=0} + P_{Y_i\ Y_{i+1}\ X_i \mid X_{i+1}=1}} = \quad a(x_i=0)=(x_{i+1}=0) f(Y_i \mid$$

$$P_{X_{i+1}=1 \mid Y_i\ Y_{i+1}\ X_i=1} = \frac{P_{Y^*\ Y_{i+1}\ X=1 \mid X_{i+1}=1}}{P_{Y_i\ Y_{i+1}\ X_i \mid X_{i+1}=0} + P_{Y\ Y_i\ X_i \mid X_{i+1}=1}}$$

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

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

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
```

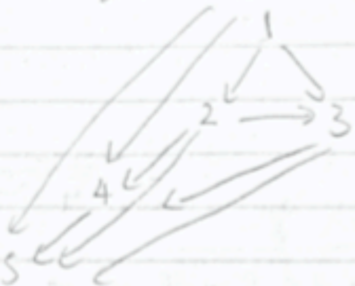
```
> max_q
[1] 0.950075
> sigma_2
[1] 0.9949241
```

My estimated q is 0.95 and estimated sigma^2 is 1

Problem 2

Q2 Yes

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



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

This also forms joint distribution

$$\left.\begin{aligned} &- P_{2|1} = \frac{P_{21}}{P_1} \\ &- \frac{P_{21}}{P_1} \cdot P_1 = P_{21} \end{aligned}\right\} P_{12} = P_1 P_{41}$$

$$\left.\begin{aligned} &- P_{3|21} = \frac{P_{321}}{P_{21}} \\ &- \frac{P_{321}}{P_{21}} \cdot P_{21} = P_{321} \end{aligned}\right\} P_{321} = P_{3|21} \cdot P_{2|1} P_1$$

$$\prod_{i=1}^{n} P_{i|1:i} = P(x_1, \ldots x_n)$$

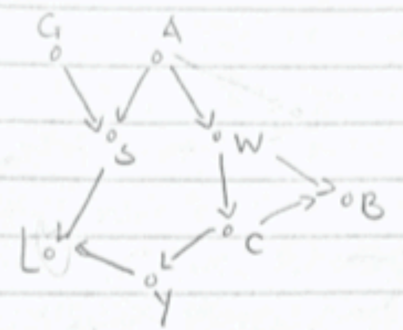
Problem 3

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

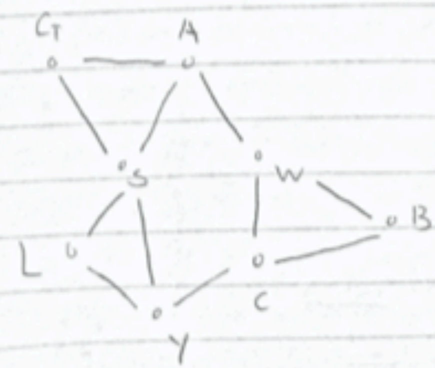
a,



b, moral graph

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

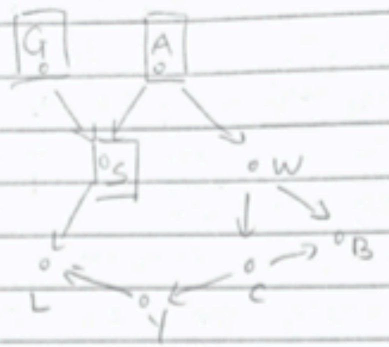
- $\{u, v\} \in E'$ if $u \to v$
- $\{u, v\} \in E'$ if u ; v have common child



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

S=s

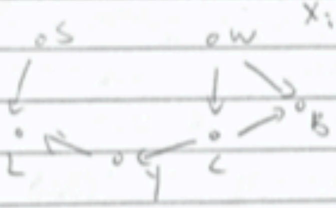
$P(L,Y,C,B,W) \mid G=g, A=a, S=s$



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

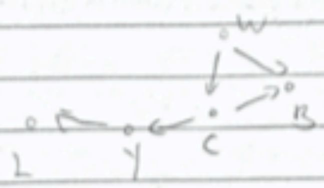
$$\text{step 1} \quad P(L,Y,C,B,W) \mid \underbrace{G=g, A=a, S=s}_{}^{\text{root}}$$

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



random

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

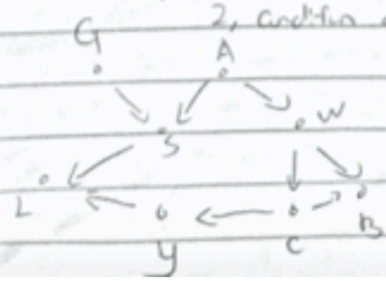


ii, $P_{CWB \mid GAS}$

Strategy 1, marginalize over leaves

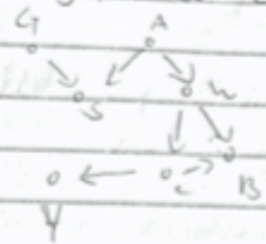
2, condition on roots

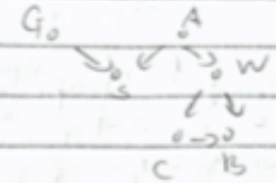
Note for BN: can only marginalize over leaves



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



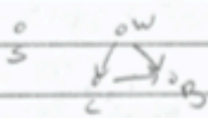


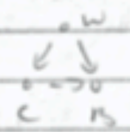
X_{GAWCB}

Step2 apply conditioning rule

① conditioning over wrt G, A   ② conditioning over S

$P_{CWB|\underline{GA}S}$



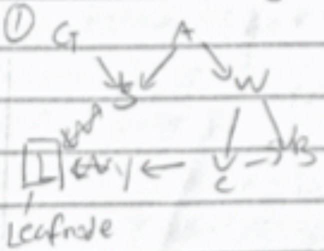


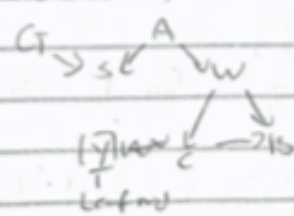
iii,

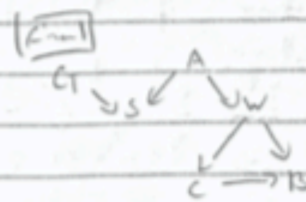
P_{GASWCB}

strategies ① marginalize over L (leaf node)

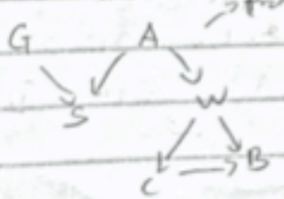
② marginalize over Y (leaf node)







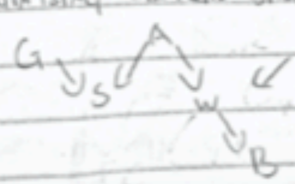
iv P_{GASWB} → from previous we can marginalize over C



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

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

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



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

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

$30 \cdot 94 \cdot 270 \cdot 340 \cdot 2 \cdot 5 \approx 10^7$ 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, $P(C \geq 120 \mid Y=50, A \geq 3, S=[20,30]) = \dfrac{P_{CYAS}}{P_{YAS}}$

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

$230 \cdot 10 \cdot 2 \cdot 340 \cdot 2 = 7820$ 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 $\dfrac{P(C \geq 120, Y=50, A \geq 3, S=[20,30])}{P(Y=50, A \geq 3, S=[20,30])} =$

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

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

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

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

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

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

$$p(x,y) = \frac{1}{2}\Big(\underbrace{\prod_{i=1}^{1000} f_i(y_i|x_i)}_{\text{depends on } \sigma^2}\Big)\Big(\underbrace{\prod_{k=1}^{999} a(x_{k+1}, x_k)}_{\text{depends on } q}\Big)$$

Without using EM:

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

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

$\boxed{\text{Essentially}}$, $P_{XY}(X,Y) \Rightarrow \sum_X P_{XY}(X,Y) = P_Y(Y^*)$

$$\max_{\sigma^2 q} \log\Big(\sum_x P_{XY}^{q\sigma^2}(x, y^*)\Big)$$

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

EM

① how to define q_{k+1} σ^2_{k+1}?

$$(q_{k+1}, \sigma^2_{k+1}) = \underset{(q,\sigma^2)}{\arg\max}\ \mathbb{E}_{X|Y^*}^{\overbrace{q_k,\sigma_k^2}^{\text{Note: we know this}}}\Big[\log\big(P_{XY}^{\overbrace{q\sigma^2}^{\text{Note: we don't know, our goal}}}(x, y^*)\big)\Big]$$

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

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

③ Let's take expection of this ②

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

$$\Updownarrow$$

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

$\boxed{\text{Note}}$

④ Find expression that maximize $E[\,]$

For σ^2:

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

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

③ take the derivative of ② in respect to σ^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 $\displaystyle\sum_{i=1}^{999} E_{X|Y=y^*}^{q_k, \sigma_k^2} [\underbrace{\log(q \mathbb{I}_{x_i = x_{i+1}} + (1-q)\mathbb{I}_{x_i \neq x_{i+1}})}_{\Downarrow \text{ simply}}]$

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

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

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

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

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

$\boxed{\text{note}}$ $\displaystyle P_{X|Y} = \frac{P_{Y|X} P_X}{P_Y}$ In hw5 we showed P_X is invariant measure

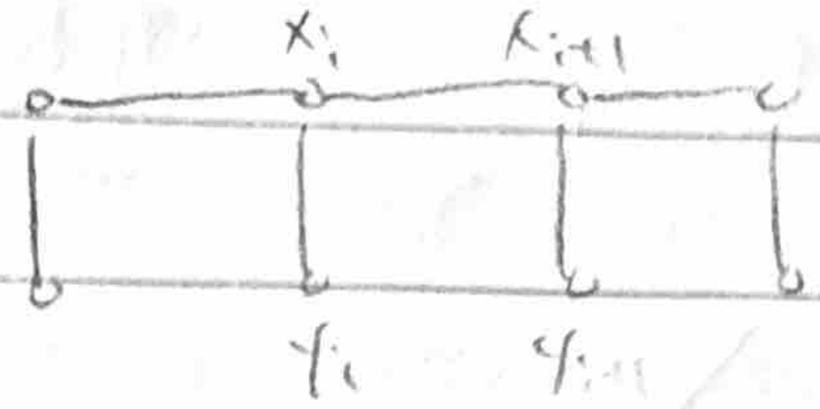
s.t. $P_X = \frac{1}{2}$

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

For B

$E_{X|Y}^{\delta^2 q}[\mathbb{I}_{x_i = x_{i+1}}]$

Note



(1) $E_{X|Y}^{\delta^2 q}[\mathbb{I}_{x_i = x_{i+1}}] = \mathbb{I}_{(x_i=0)=(x_{i+1}=0)} P(x_i, x_{i+1} \mid Y_i^*, Y_{i+1}^*)$
$+ \mathbb{I}_{(x_i=1)=(x_{i+1}=1)} P(x_i, x_{i+1} \mid Y_i^* Y_{i+1})$

↳ expression we want to solve

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

(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

For $P(X_i \mid Y_i^* Y_{i+1}^*)$, we can observe from the graph that It could be rewritten as $P(X_i \mid Y_i^*)$, s

This is similar to For A

$$P_{X|Y} = \frac{P_{Y|X}}{\sum_X P_{Y|X}}$$

For $P(X_{i+1} \mid Y_i^*, Y_{i+1}^*, X_i^*)$

$$P_{X_{i+1} \mid y_i^*, y_{i+1}^*, X_i} = \frac{P_{y_i^*, y_{i+1}^*, X_i \mid X_{i+1}} \; P_{X_{i+1}}^{\to \frac{1}{2}}}{\underbrace{P_{y_i^* y_{i+1} X_i}}_{\text{constant}}}$$

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

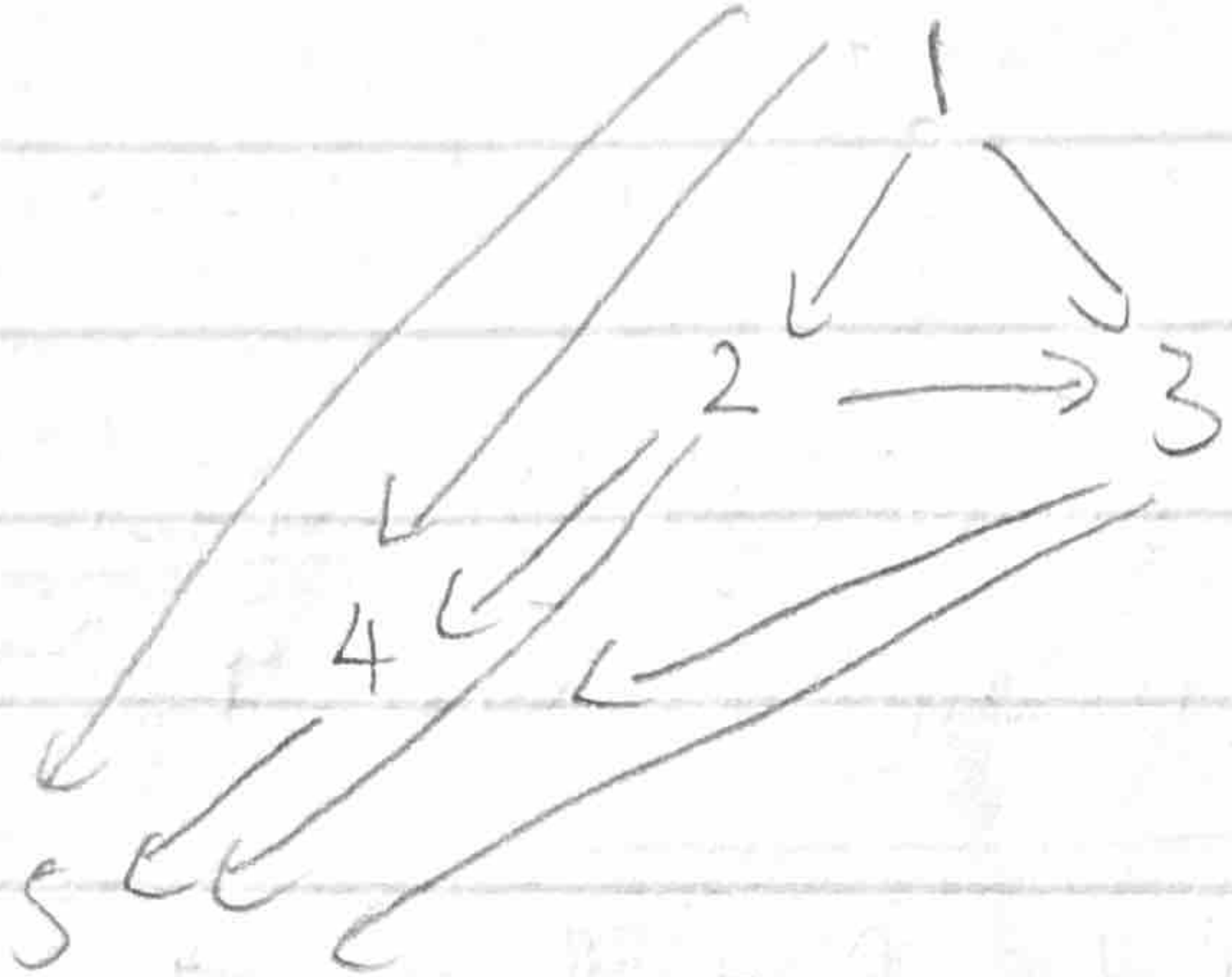
$$P_{X_{i+1}=1 \mid Y_i Y_{i+1} X_i=1} = \frac{P_{y^* y_{i+1} X=1 \mid X_{i+1}=1}}{P_{y_i y_{i+1} X_i \mid X_{i+1}=0} + P_{y y_{i+1} X_i \mid X_{i+1}=1}}$$

$P(X_i, X_{i+1} \mid Y_i Y_{i+1}) = P(X_i \mid Y_i Y_{i+1}) \, P(X_{i+1} \mid X_i Y_i Y_{i+1}) \propto a(X_i = X_{i+1}) f(X_i \mid Y_i) f(X_{i+1} \mid Y_{i+1}$

$$P(X_i, X_{i+1} \mid Y_i Y_{i+1}) = \frac{a(X_i = X_{i+1}) f(Y_i \mid Y_i) f(X_{i+1} \mid Y_{i+1})}{\sum \cancel{a() f() f()}}$$

Q2 Yes

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



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

This also forms joint distribution

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

$-\ \frac{P_{21}}{P_1} \cdot P_1 = P_{21}$

$P_{12} = P_1 P_{2|1}$

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

$-\ \frac{P_{321}}{P_{21}} \cdot P_{21} = P_{321}$

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

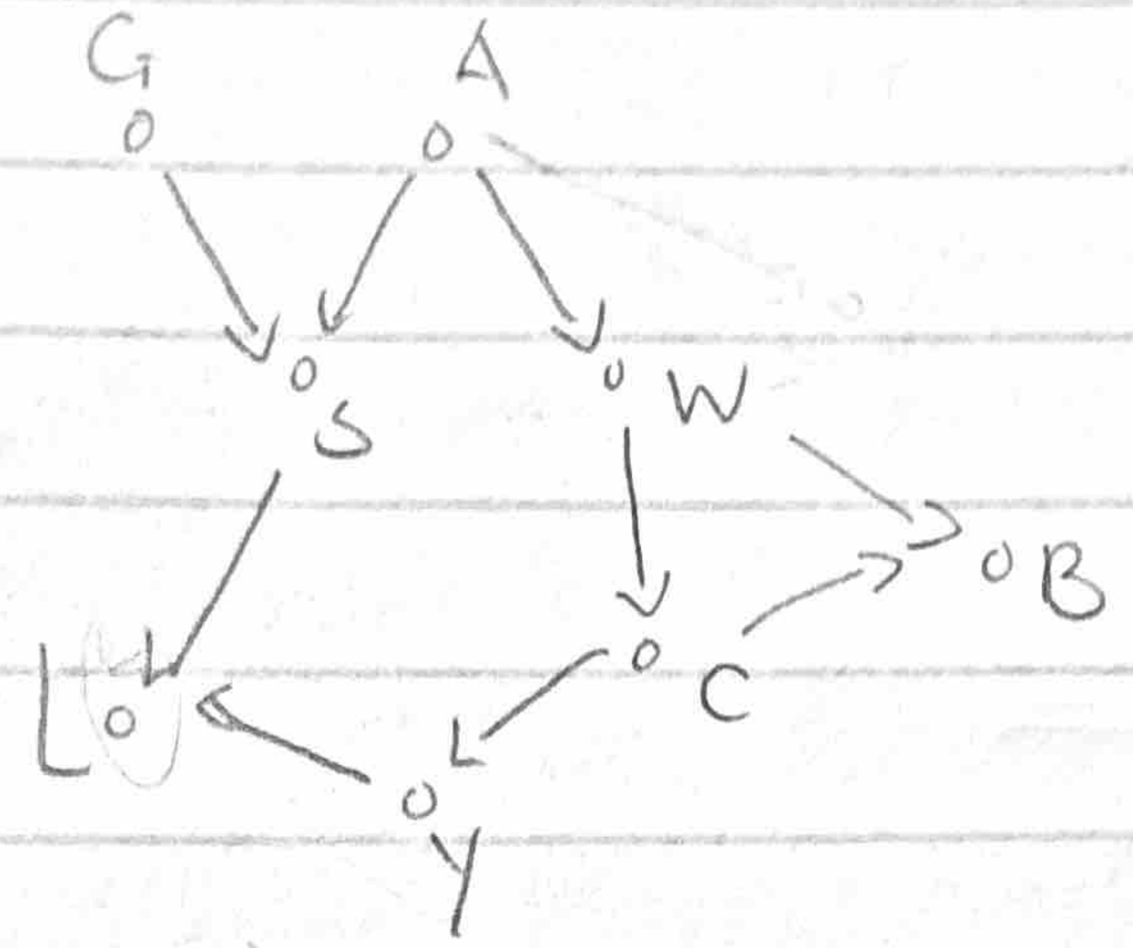
$\prod_{i=1}^{n} P_{i|1:i} = P(X_1, \ldots, X_n)$

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

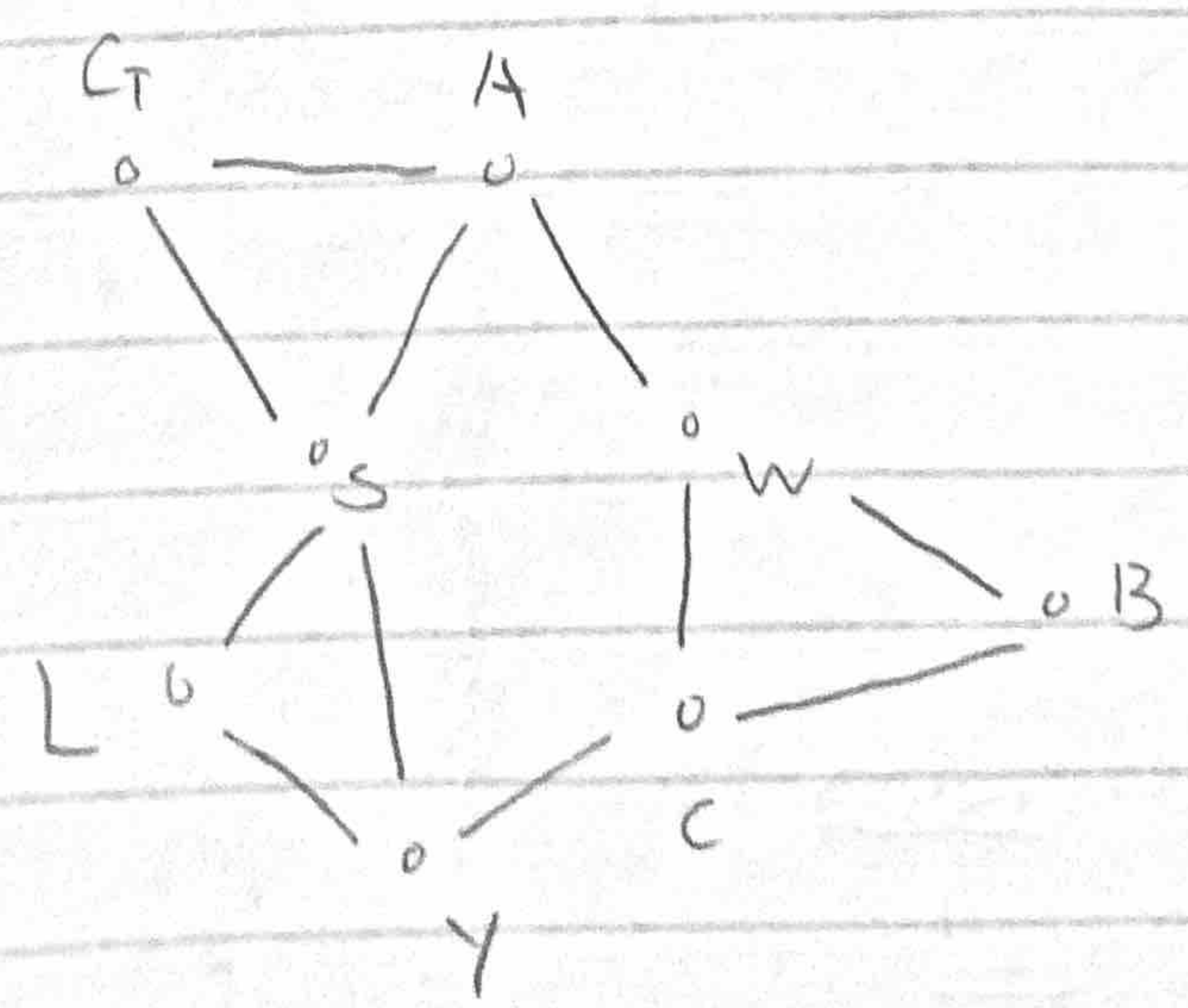
a,



b, moral graph

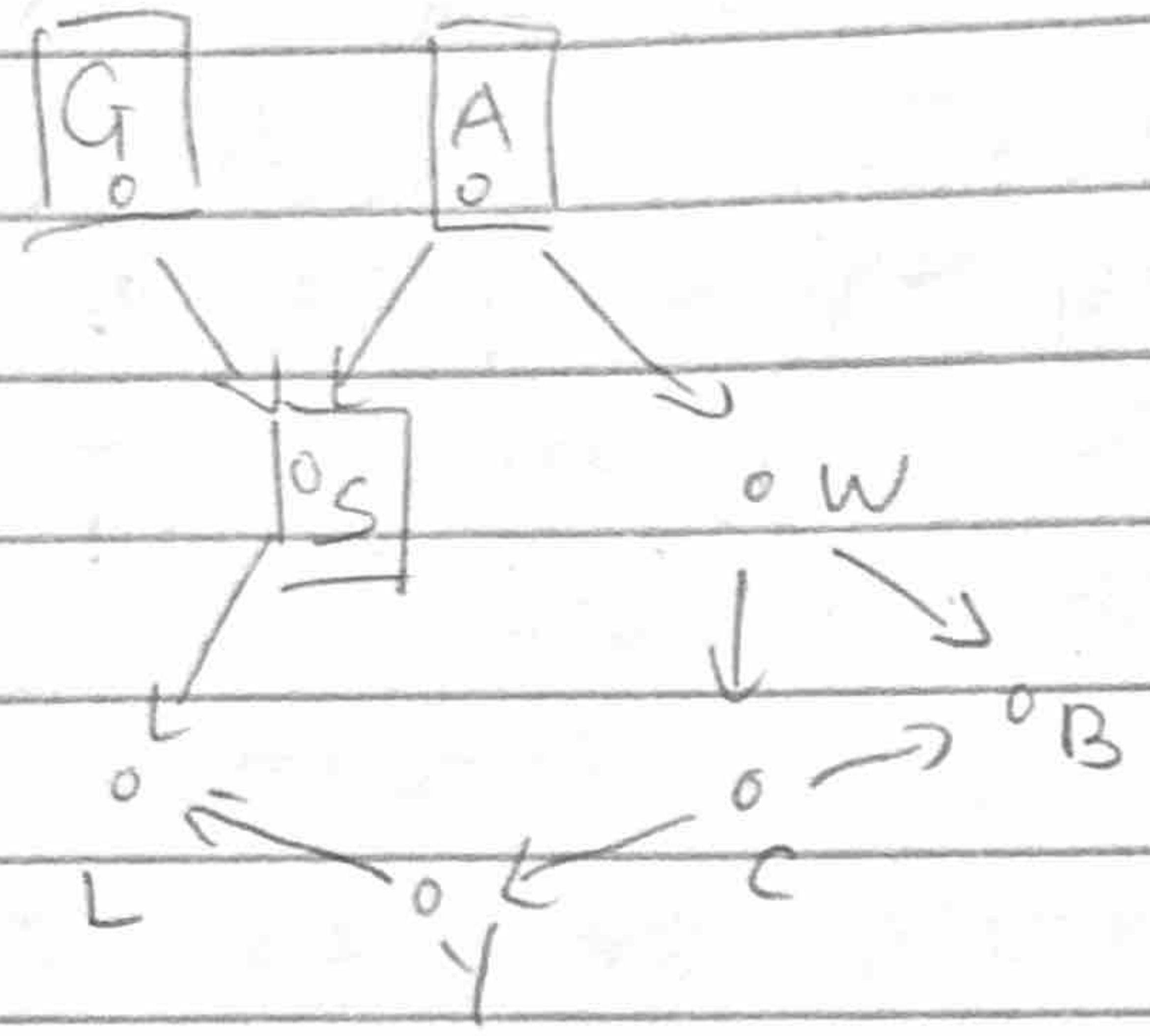
Definition: Let $G = (V, E)$ be a DAG. The "moral graph" of G is the graph $G' = (V, E')$ (undirected graph)

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



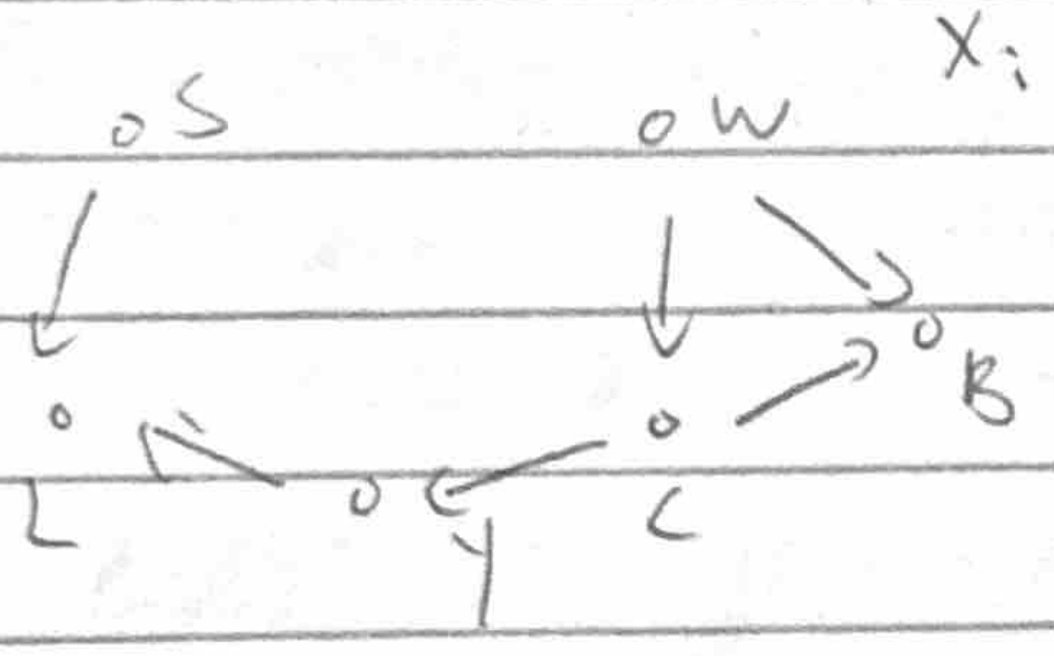
C, ; conditional dist of (L, Y, C, B, W) given G=g, A=a, S=s

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



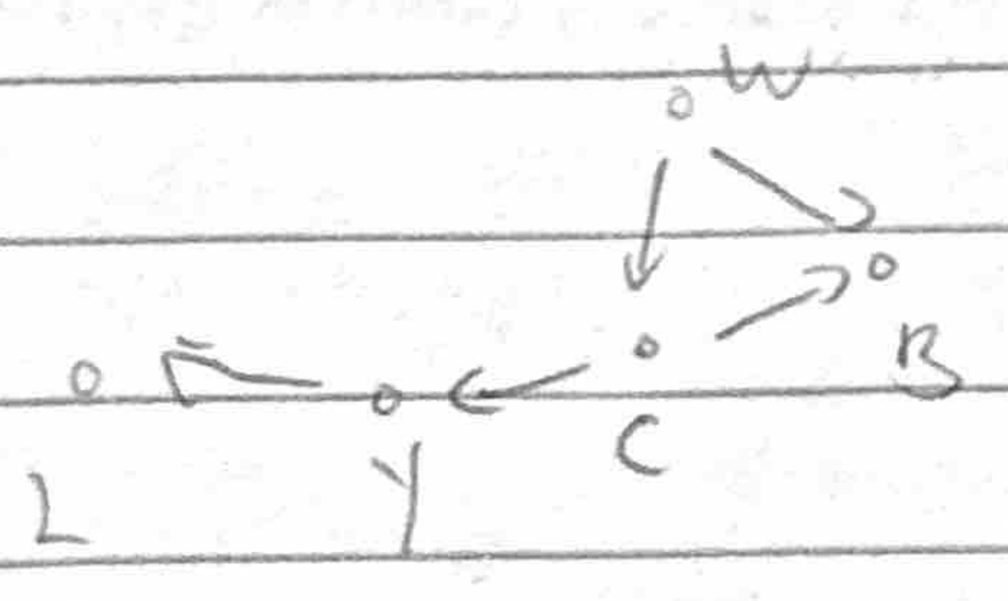
Note for BN's conditional rule: can only condition on root

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



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

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

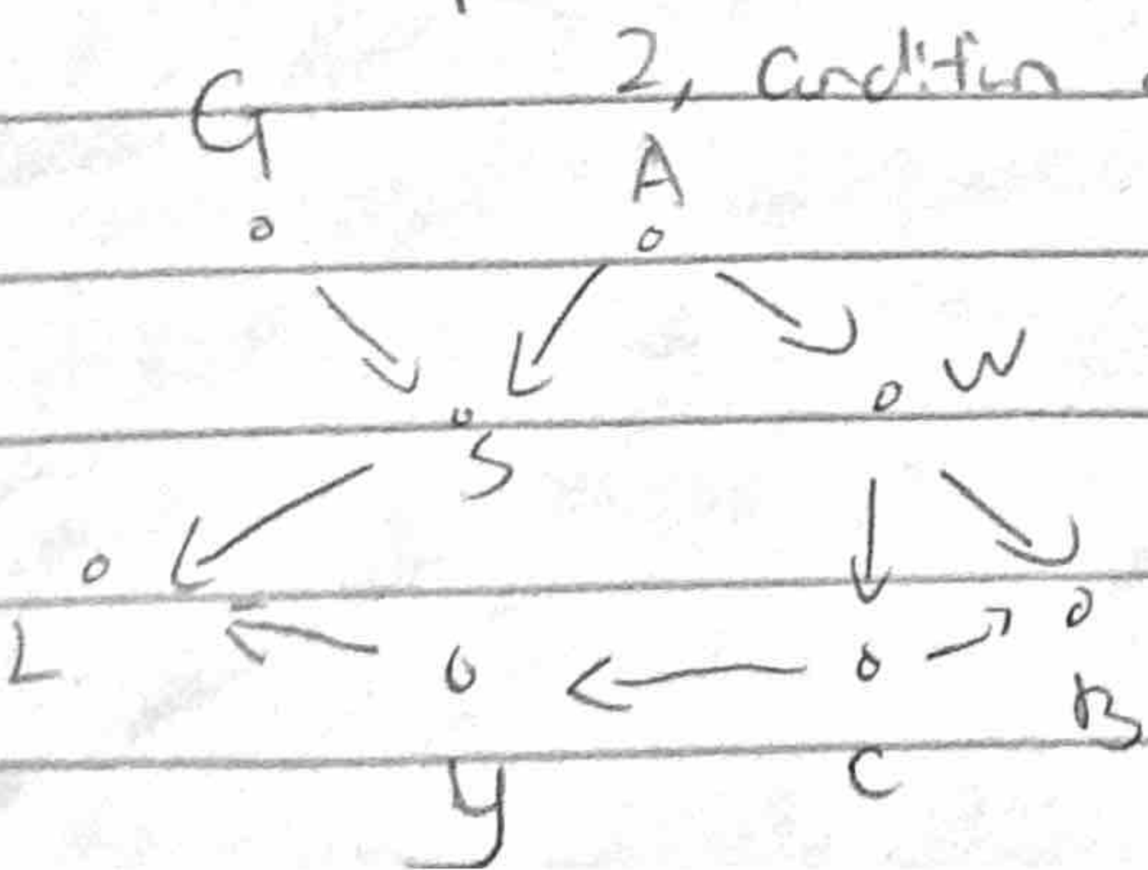


ii; $P_{CWB|GAS}$

Strategy 1, marginalize over leaves

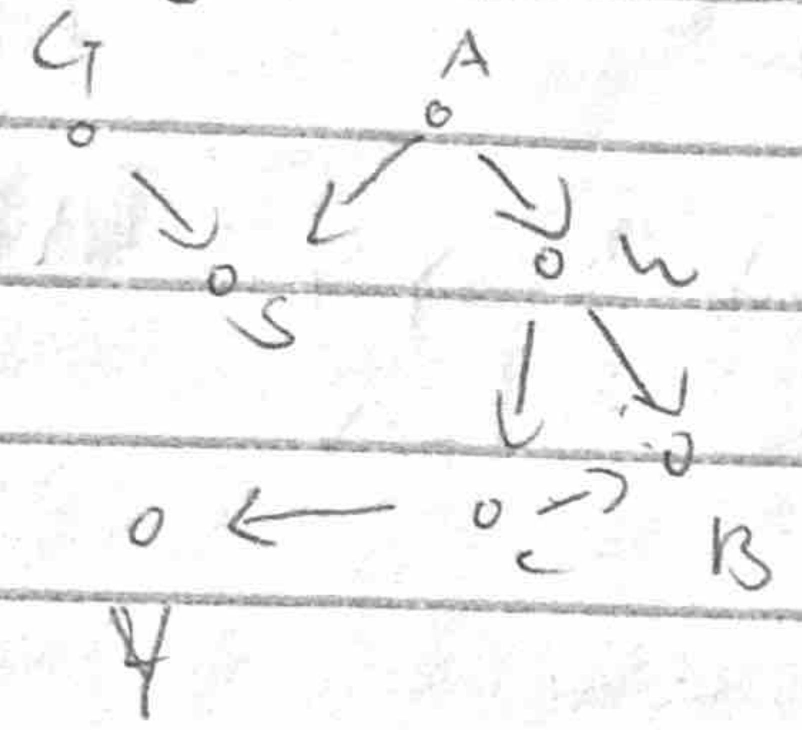
2, condition on roots

Note for BN: can only marginalize over leaves

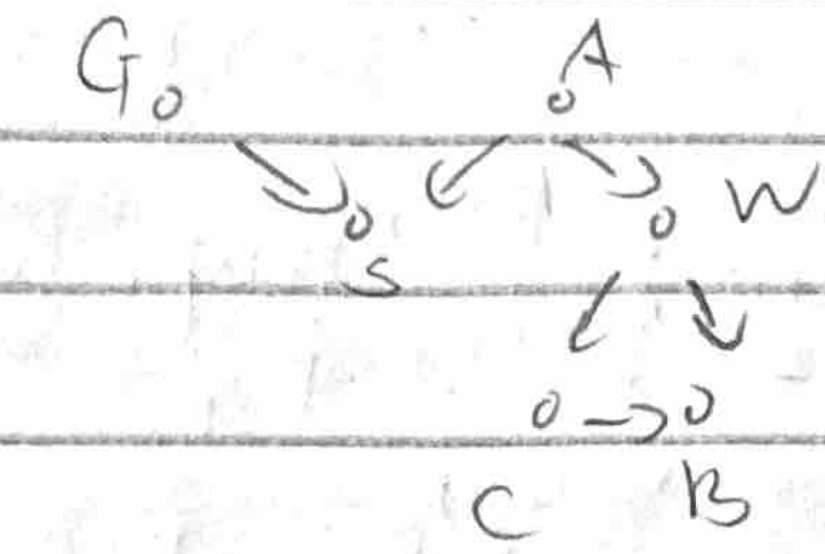


step1 marginalize over L and Y

① marginalize over L b/c L is a leave



② now L is a leaf so marginalize over Y

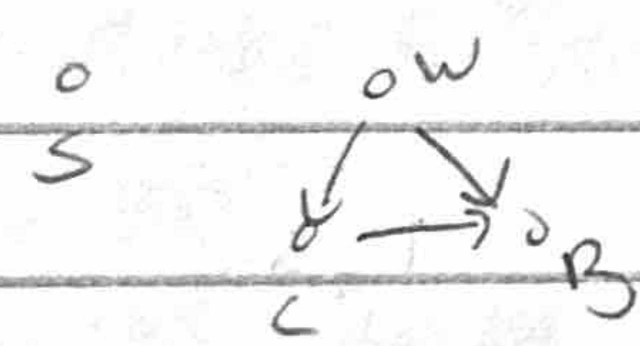


X_{GAWCB}

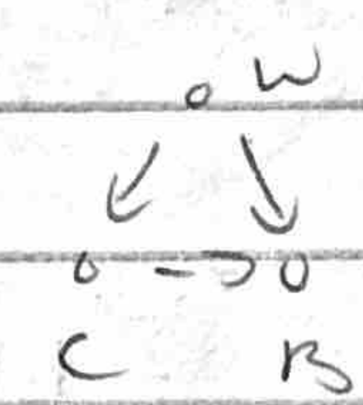
step2 apply conditioning rule

① Conditioning over root G, A

$P_{CWB|\underline{GA}S}$



② conditioning over S



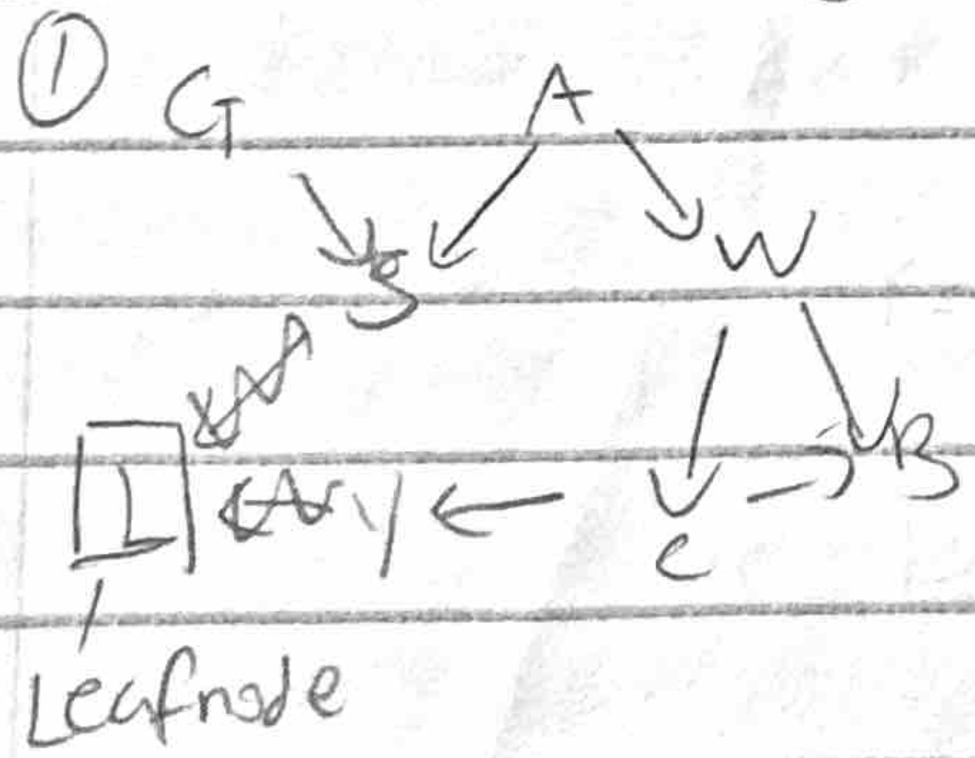
iii

P_{GASWCB}

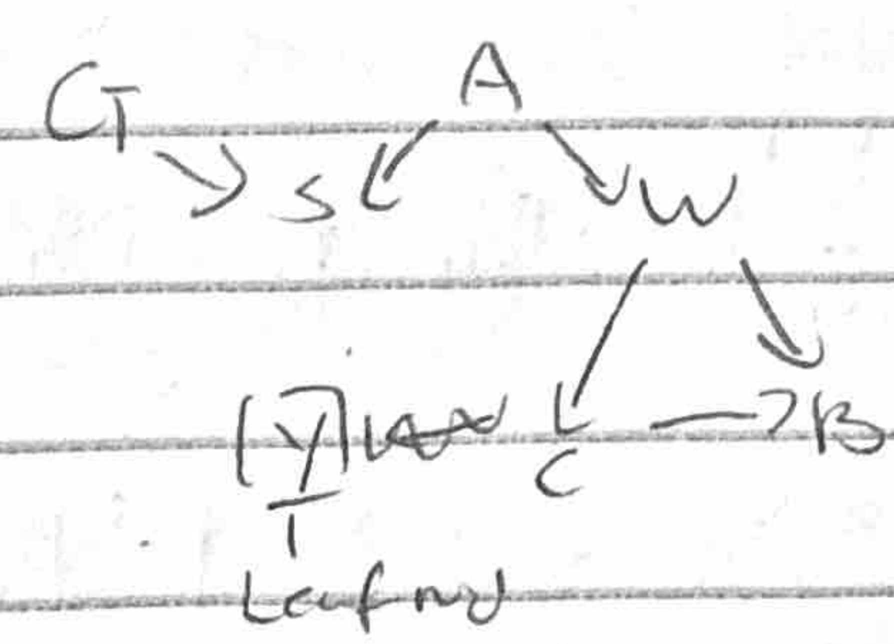
strategies ① marginalize over L (leaf node)

② marginalize over Y (leaf node)

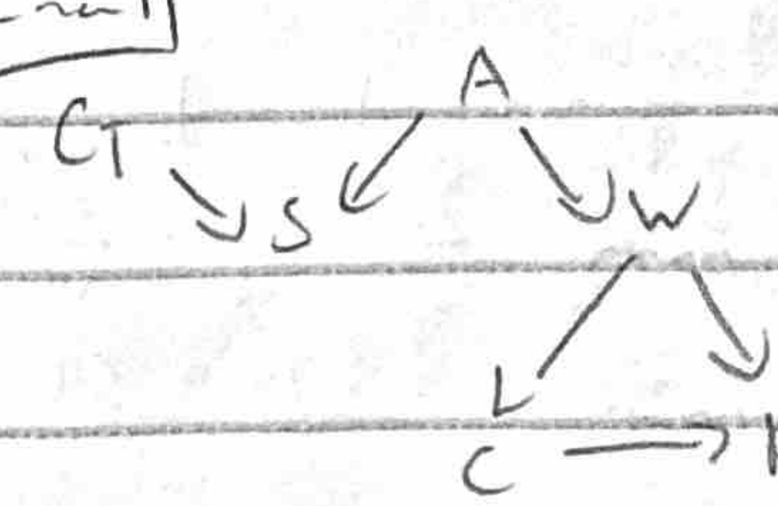
①



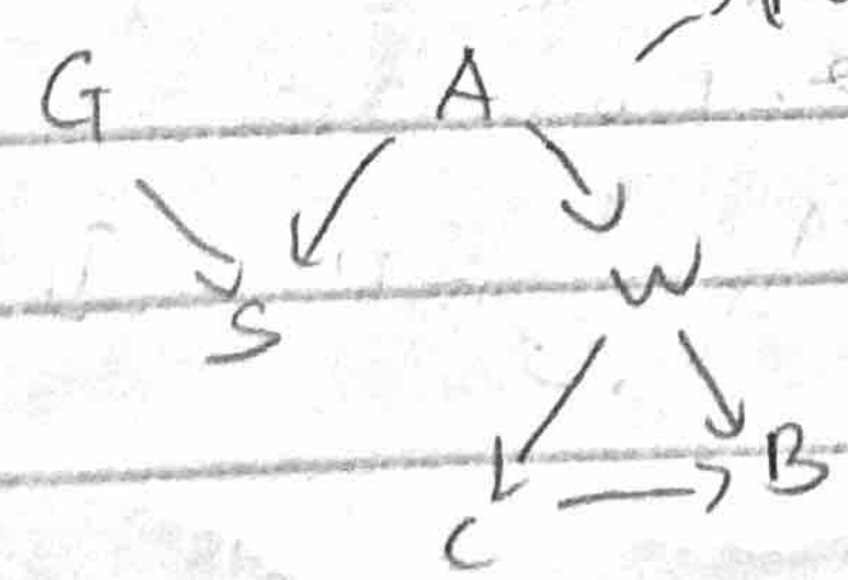
②



$\boxed{\text{Final}}$



iv P_{GASWB} → from previous

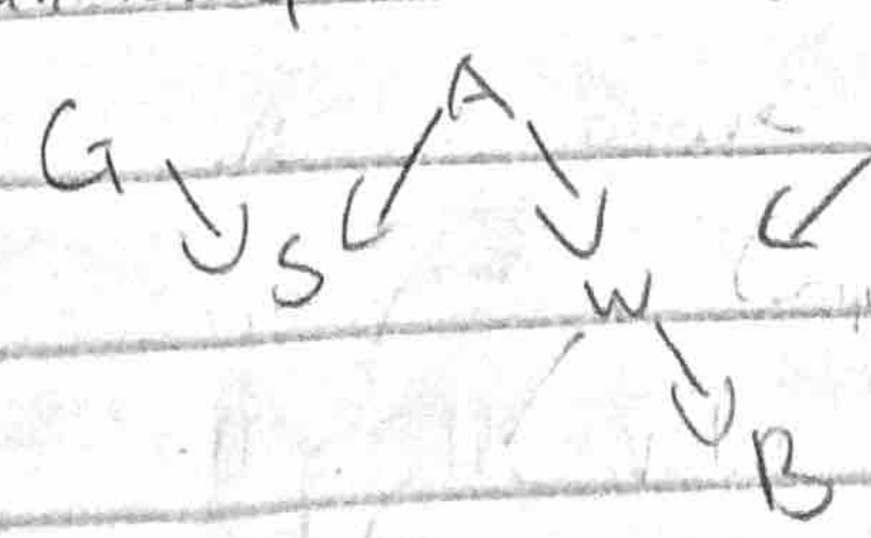


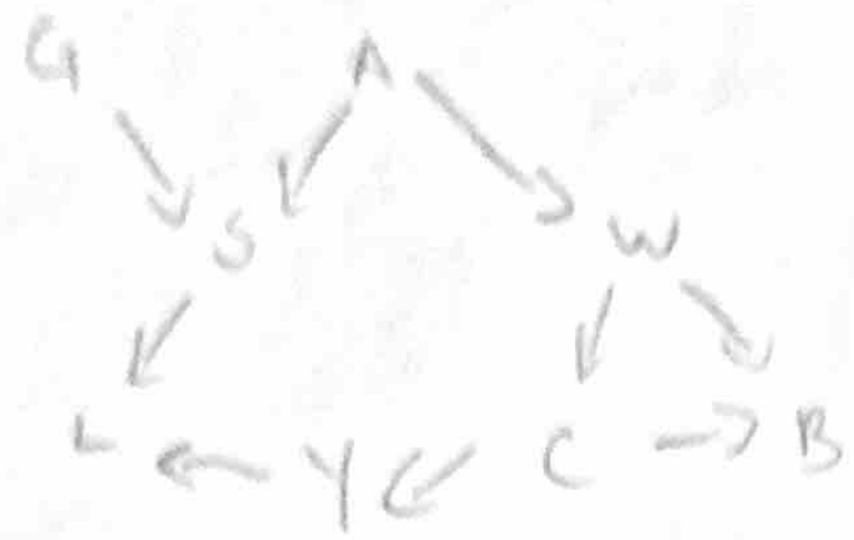
we can marginalize over C

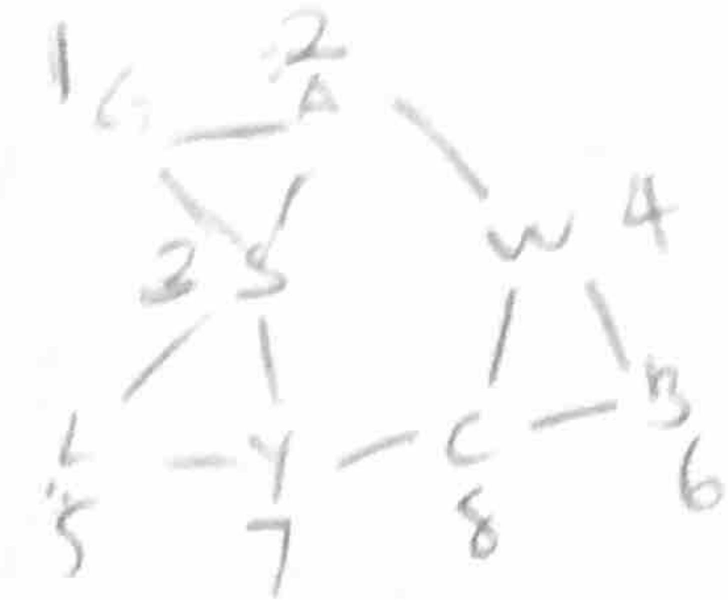
$$\sum_C P_G P_A P_{S|AG} P_{W|A} P_{C|W} P_{B|WC} = P_G P_A P_{S|AG} P_{W|A} \sum_C P_{C|W} P_{B|WC}$$

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

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







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

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

$30 \cdot 5 \cdot 2 = 300$ 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 culd get P_{YAS} from $\boxed{e}$ so

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