q1/

$$p(x_i=1)=q$$
$$p(x_i=0)=1-q$$

1/

$$p(x_2,x_3|z_5)=\frac{p(x_2,x_3,z_5)}{p(z_5)}=\frac{p(x_2,x_3)}{p(z_5)}$$

$$p(z_5=1)=p(x_2=0,x_3=1)+p(x_2=1,x_3=0)$$
$$p(z_5=0)=p(x_2=0,x_3=0)+p(x_2=1,x_3=1)$$

$$p(x_2=0,x_3=1|z_5=1)=\frac{(1-q)(q)}{2(1-q)(q)}=\frac{1}{2}$$

$$p(x_2,x_3|z_5=1)=\frac{\overset{\curvearrowright p(x_2)p(x_3)}{p(x_2,x_3)}}{p(x_2=0,x_3=1)+p(x_2=1,x_3=0)}$$

$$p(x_2=1,x_3=0|z_5=1)=1/2$$
$$p(x_2=1,x_3=1|z_5=1)=0$$
$$p(x_2=0,x_3=0|z_5=1)=0$$

$$p(x_2,x_3|z_5=0)=\frac{p(x_2,x_3)=p(x_2)p(x_3)}{p(x_2=0,x_3=0)+p(x_2=1,x_3=1)}$$
↳for $x_2=0,x_3=0$
$x_2=1,x_3=1$

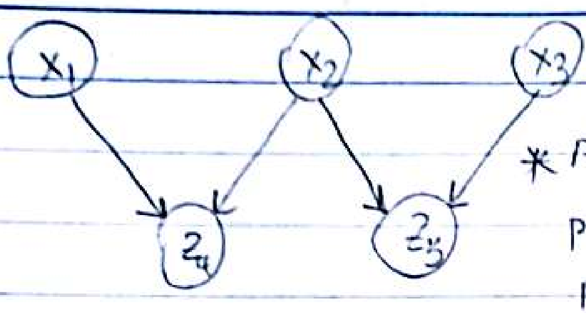
$$p(x_2=0,x_3=0|z_5=0)=\frac{(1-q)(1-q)}{(1-q)^2+q^2}=\frac{(1-q)^2}{(1-q)^2+q^2}$$

$$p(x_2=1,x_3=1|z_5=0)=\frac{q^2}{(1-q)^2+q^2}$$

$$p(x_2=1,x_3=0|z_5=0)=0 \leftarrow p(z=0|x_3=0,x_2=1)=0$$

$$p(x_2=0,x_3=1|z_5=0)=0$$

2/




$* \; p(x_1, x_2, x_3, z_4, z_5) =$
$p(x_1)\, p(x_2)\, p(x_3)$
$p(z_5 \mid x_2, x_3)\, p(z_4 \mid x_1, x_2)$

$p(x_i = 0) = 1 - q \quad , \quad p(x_i = 1) = q$

| $X_1$ | $X_3$ | $z_5=0$ | $z_5=1$ |
|---|---|---|---|
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 |

| $X_1$ | $X_2$ | $z_4=0$ | $z_4=1$ |
|---|---|---|---|
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 |
| 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 |

$X_2 \perp X_3 \mid (X_1, z_4)$
$X_2 \perp X_3 \mid X_1$
$z_4 \perp (X_3, z_5) \mid (X_1, X_2)$
$z_4 \perp z_5 \mid (X_1, X_2, X_3)$
$z_5 \perp (X_1, z_4) \mid (X_2, X_3)$

$X_1 \perp (X_2, X_3, z_5)$
$X_2 \perp (X_1, X_3)$
$X_2 \perp X_1 \mid (X_3, z_5)$
$X_1 \perp z_5 \mid (X_2, X_3, z_4)$
$(X_1, z_4) \perp (X_3, z_5) \mid X_2$
$X_1 \perp (X_2, X_3) \mid z_5$
$X_1 \perp (X_2, z_5) \mid X_3$
$X_1 \perp (X_3, z_5) \mid (X_2, z_4)$

$X_3 \perp (X_1, X_2, z_4)$
$X_3 \perp (X_1, X_2) \mid z_4$
$X_3 \perp (X_2, z_4) \mid X_1$
$X_3 \perp (X_1, z_4) \mid (X_2, z_5)$
$X_3 \perp z_4 \mid (X_1, X_2)$
$X_3 \perp z_4 \mid (X_1, X_2, z_5)$
$X_3 \perp (X_1, z_4) \mid X_2$

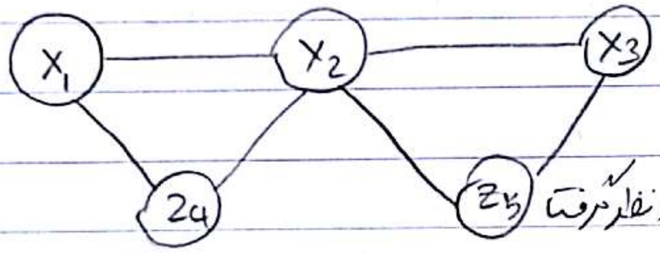
potential

توابع وزن را به 2 شکل زیر می توان تعریف کرد:

3/

$$\begin{cases} \phi_1(x_2, z_5, x_3) = p(x_3)\,p(x_2)\,p(z_5 \mid x_2, x_3) \\ \phi_2(x_1, z_4, x_2) = p(x_1)\,p(z_4 \mid x_1, x_2) \end{cases}$$

$$\begin{cases} \phi_3(x_2, z_5, x_3) = p(x_3)\,p(z_5 \mid x_2, x_3) \\ \phi_4(x_1, z_4, x_2) = p(x_1)\,p(x_2)\,p(z_4 \mid x_1, x_2) \end{cases}$$




می توان 2 حالت بالا را به صورت روبه‌رو نمایش داد:

در واقع تمام مثلث ها را در نظر گرفتیم

هر کلیک یک مثلث ← روی 3 یال تعریف شده است

| | |
|---|---|
| $x_1 \perp x_3 \mid (z_4, x_2, z_5)$ | $(x_1, z_4) \perp (x_3, z_5) \mid x_2$ |
| $x_1 \perp z_5 \mid (z_4, x_2, x_3)$ | $z_4 \perp (x_3, z_5) \mid (x_1, x_2)$ |
| $z_4 \perp x_3 \mid (x_1, x_2, z_5)$ | $(x_1, z_4) \perp z_5 \mid (x_2, x_3)$ |
| $z_4 \perp z_5 \mid (x_1, x_2, x_3)$ | $x_1 \perp (x_3, z_5) \mid (z_4, x_2)$ |
| $(x_1, z_4) \perp x_3 \mid (x_2, z_5)$ | $x_1 \perp x_3 \mid x_2$ |
| $x_1 \perp (x_3, z_5) \mid (z_4, x_2)$ | $z_4 \perp x_3 \mid x_2$ |

تمامی استقلال هایی که میتوان از روابط بالا نتیجه گرفت نیز برای این گراف صادق است.

$$p(x_1, x_2, x_3, z_4, z_5) = \frac{1}{Z}\phi_1(\cdot\cdot)\,\phi_2(\cdot\cdot) = 1/Z\,\phi_3(\cdots)\,\phi_4(\ )$$

$\phi_1$ و $\phi_2$ و $\phi_3$ و $\phi_4$ از معادله‌ی توزیع joint شکل 2 حاصل شده‌اند.

*

4/ $z_5 \perp x_3 \ , \ z_4 \perp x_1$

$$p(z_5=1) = p(z_5=1 \mid x_3=0) = p(z_5=1 \mid x_3=1)$$

$$p(z_5=1 \mid x_3=0) = p(x_2=1) = q \qquad 0<q<1$$

$$p(z_5=1 \mid x_3=1) = p(x_2=0) = 1-q$$

$$p(z_5=1) = p(z_5=1 \mid x_3=1)\, p(x_3=1) + p(z_5=1 \mid x_3=0)\, p(x_3=0)$$

$$= 2q(1-q) = 1$$

پس اگر $0<q<1$ باشد ما استقلال را فقط به شرط $q=1/2$ خواهیم داشت که این مورد توسط گراف مشخص نفی شود.

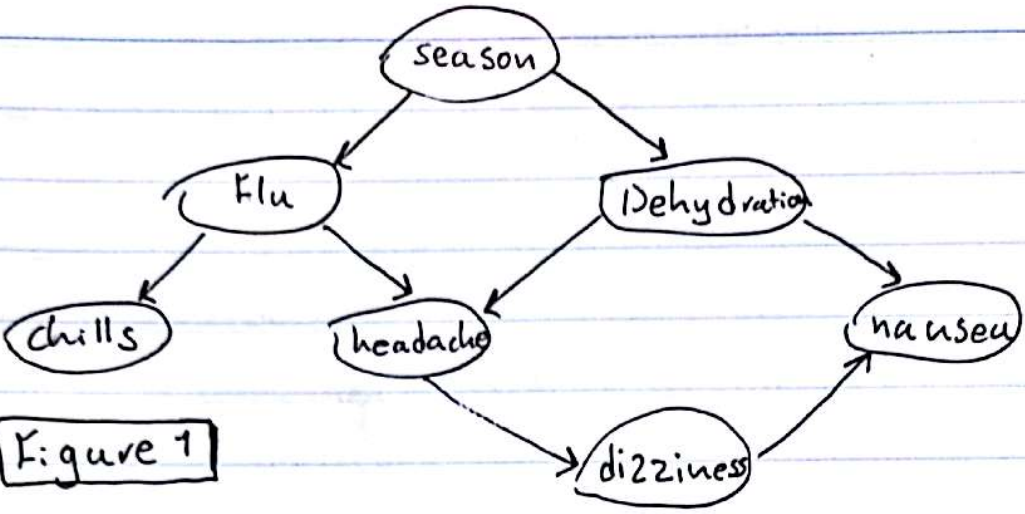



Figure 1

1 season ⊥ chills

False: season → Flu → chills : Flu should be observed

2. season ⊥ chills | (Flu)

True season → Flu → chills → d-sepration

3. season ⊥ headache | Flue

False season → dehydration → headache

مسیر بالا قطع نشده، با dehydration given باید شود

4. season ⊥ headache | Flue, dehydration

True تمام مسیرهای موازی قطع شده اند

5. Season ⊥ Nausea | Dehydration

False season → Flu → headache → dizziness → nausea باید این مسیر قطع شود

6. season ⊥ Nausea | Dehydration, Headache

True تمامی مسیرهای موازی بین این 2 متغیر قطع شده

7. Flu ⊥ Dehydration False

Flu ← seaso → dehydration با مسیر روبه رو وابستگی داریم

8. Flu ⊥ Dehydration | season, headache False

Flue → headache ← dehygdration

در v-stru بالا با given کردن headache ، وابستگی ایجاد می شود

9. Flu ⊥ Dehydration | season

True تمام مسیرها برای ایجاد وابستگی قطع شده اند

10. Flu ⊥ Dehydration | season, Nausea

False Flu → Headache → Dizziness → Nausea ← Dehydration

با given کردن Nausea یک V-struct فعال شده ، وابستگی ایجاد می‌کند

11. chills ⊥ Nausea

False chills ← Flu ← season → Dehydration → Nausea

یکی از 3 متغیر مسیر بالا بین chills , Nausea باید given شود تا استقلال ایجاد شود.

12. chills ⊥ Nausea | headache

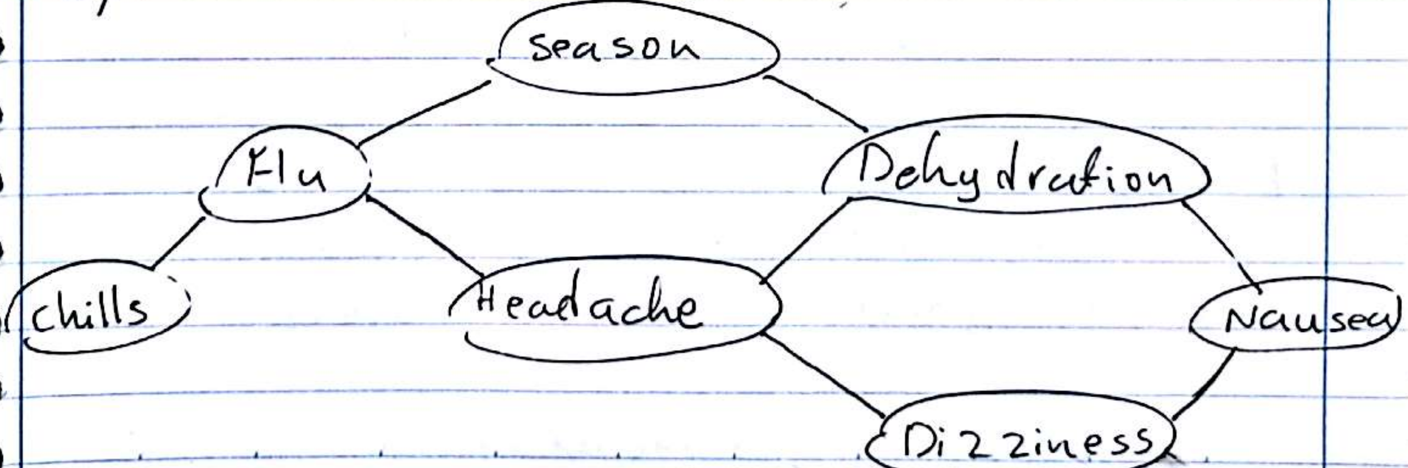
False chills ← Flu → headache ← Dehydration → Nausea

با given کردن headache در مسیر بالا یک V-struct. فعال می‌شود و وابستگی ایجاد می‌کند

1/ $p(S,F,D,C,H,N,Z) = p(S)\,p(F|S)\,p(D|S)$

$\times\, p(C|F)\,p(H|F,D)\,p(N|D,Z)\,p(D|H)$

2/



$$p(S, F, D, C, H, Z, N) = 1/Z \; \phi_1(S)\, \phi_2(F)\, \phi_3(D)$$

$$\phi_4(C)\, \phi_5(H)\, \phi_6(Z)\, \phi_7(N)\, \phi_8(S, F)\, \phi_9(S, D)$$

$$\phi_{10}(F, C)\, \phi_{11}(F, H)\, \phi_{12}(D, H)\, \phi_{13}(D, N)\, \phi_{14}(N, Z)$$

$$\phi_{15}(H, Z)$$

ضربی از توابع وزن روی هر نود و یال

Evaluate probabilities

1 $p(F = true) = \sum_{s} p(F = true, S = s) = \sum_{s} p(F = true \mid S = s)\, p(s)$

$= P(F = true \mid S = w)\, p(S = w) + p(F = true \mid S = s)\, p(S = s)$

$= 0.4 \times 0.5 + 0.1 \times 0.5 = 0.25$

2 $p(F = true \mid S = w) = 0.4$

3 $p(F = true \mid S = w, H = true) = \dfrac{p(F = true, S = w, H = true)}{p(S = w, H = true)}$

$p(F = true, S = w, H = true) = \sum_{d} p(F = true, S = w, H = true, D = d)$

$= \sum_{d} \Big[ p(H = true \mid F = true, D = d)\, p(F = true \mid S = w)$
$\quad p(D = d \mid S = w)\, p(S = w) \Big]$

$= 0.9 \times 0.4 \times 0.1 \times 0.5 + 0.8 \times 0.4 \times 0.9 \times 0.5 = 0.162$

$P(S=w, H=true) = \sum_{f,d} P(F=f, S=w, H=true, D=d)$

$= \sum_{f,d} \Big[ P(D=true \mid F=f, D=d)\, P(F=f \mid S=w)$

$P(D=d \mid S=w)\, P(S=w) \Big]$

$= 0.9 \times 0.4 \cdot 0.7 \times 0.5 + 0.8 \times 0.4 \times 0.9 \times 0.5 +$

$0.8 \times 0.6 \times 0.1 \times 0.5 + 0.3 \times 0.6 \times 0.9 \times 0.5 = 0.267$

$\Rightarrow P(F=true \mid S=w, H=true) = \frac{0.162}{0.267} = 0.61$

4/ $P(F=true \mid S=w, H=true, D=true)$

$= \frac{P(F=true, S=w, H=true, D=true)}{P(S=w, H=true, D=true)}$ ①

$\hookrightarrow = \sum_{f} P(F=f, S=w, H=true, D=true)$

$= \sum_{f} \Big[ P(H=true \mid F=f, D=true)$

$P(F=f \mid S=w)\, P(D=true \mid S=w)\, P(S=w) \Big]$

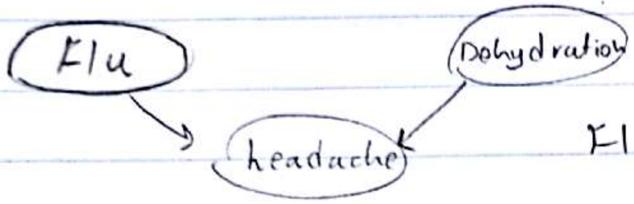
$= 0.9 \times 0.4 \times 0.1 \times 0.5 + 0.8 \times 0.6 \times 0.1 \times 0.5$

① $= P(H=true \mid F=true, D=true)\, P(F=true \mid S=w)$

$P(D=true \mid S=w)\, P(S=w) = 0.9 \times 0.4 \times 0.1 \times 0.5$

$\Rightarrow P(F=true \mid S=w, H=true, D=true) = 0.43$

5/ دانستن اینکه dehydrated هستیم likelihood اینکه Flu داریم را کاهش
می دهد این تاثیر فهم است به وسیله اینکه سردرد (headache)
به وسیله dehydration توضیح داده می شود




یعنی اگر سردرد داشته باشیم
در مورد Flu و dehydrate
به یکدیگر وابسته می شود.
یعنی اگر dehydrate باشیم و سردرد داشته باشیم احتمال را likelihood
اینکه آنفولانزا داشته باشیم کاهش می یابد

q3/

① $\ln q_m^*(H_m) = E_{H_{-m}}[\ln p(H_m | H_{-m}, O_{1:N})] + \text{const}$

$q(H) = \underbrace{q_1(H_1)}\,\underbrace{q_2(H_2)}\cdots\underbrace{q_M(H_M)}$

$p(H|O) = p(H_1|O)\,p(H_2|O)\cdots p(H_M|O)$

$$q^*(z) = \underset{q(H)\in Q}{\arg\min}\ KL(q(H)\,\|\,p(H|O))$$

$$= KL(q(H)\,\|\,p(O,H)) + \log p(O)$$

$$\log p(O) = -KL(q(H)\,\|\,p(O,H)) + \underbrace{KL(q(H)\,\|\,p(H|O))}_{\geq 0 \ ①}$$

$$\underbrace{-KL(q(H)\,\|\,p(O,H))}_{ELBO} \leq \log p(O) \quad ②$$

طبق 1 و 2 تقریب می‌زنیم:

$$q^*(H) = \arg\max \underbrace{-KL(q(H)\,\|\,p(O,H))}_{ELBO \text{ یا } L}$$

chain rule: $p(O_{1:N}, H_{1:M}) = p(O_{1N}) \prod_{i=1}^{M} p(H_i | H_{1(i-1)}, O_{1N})$

$$ELBO = -\int_H q(H) \log \frac{q(H)}{p(O,H)}\, dH =$$

$$E_{q(H)}\left[-\log \frac{q(H)}{p(O,H)}\right] = E_{q(H)}[\log p(O,H)] - E_{q(H)}[\log q(H)]$$

$$E_q[\log p(O,H)] \doteq E_q[\underbrace{\log p(O_{1:N})}_{\text{نسبت به } q \text{ ثابت}}] + \underbrace{\sum_{i=1}^{M} E_q[\log p(H_i | H_{1(i-1)}, O_{1:N})]}$$

$$= \text{const} + E_q[\log p(H_m | H_{1:m-1}, O_{1:N})]$$ ما فقط برای $H_m$ می خواهیم

$$E_q[\log p(H_m \mid H_{1:m-1}, O_{1:N})] =$$

$$\int_{H_m} q_m(H_m) \int_{H_1}\int_{H_2}\cdots\int_{H_M} q_1(H_1)\cdots q_M(H_M) \overbrace{\phantom{x}}^{\log p(H_m \mid H_{1:m-1}, O_{1:N})} dH_1 \ldots dH_M \, dH_m$$

$$= E_{q_m}[E_{H_{-m}}[\log p(H_m \mid H_{-m}, O_{1:N})]] + \text{const} \quad (1)$$

$$E_q[\log q(H)] = \sum_{i=1}^{M} \log E_{q_i(H_i)}[\log q_i(H_i)]$$

$$= E_{q_m}[\log q_m(H_m)] + \text{const} \quad (2)$$

$$(1), (2) \Rightarrow q_m^*(H_m) = E_{q_m}[E_{H_{-m}}[\log p(H_m \mid H_{-m}, O_{1:N})]]$$

$$- E_{q_m}[\log q_m(H_m)] + \text{const}$$

$$\Rightarrow q_m^*(H_m) = E_{q_m}[E_{H_{-m}}[\log p(H_m \mid H_{-m}, O_{1:N})] - \log q_m(H_m)] + C$$

$$= \underbrace{\int_{H_m} q_m(H_m)(E_{H_{-m}}[\log p(H_m \mid H_{-m}, O_{1:N})] - \log q_m(H_m)]}_{G(q_m(H_m))} dH_m + C$$

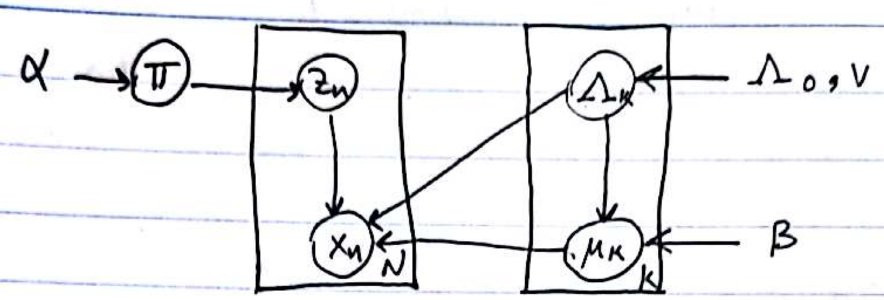
عبارت G باید نسبت به $q_m$ ، max شود پس طبق Euler-lagrange:

$$\frac{dG}{dq_m} = E_{-H_m}[\log p(H_m \mid H_{-m}, O_{1:N})] - \log q_m^*(H_m)$$

$$- \frac{1}{q_m^*(H_m)} q_m^*(H_m) = 0$$

$$\Rightarrow \log q_m^*(H_m) = E_{-H_m}[\log p(H_m \mid H_{-m}, O_{1:N})] + C$$

(3)




$p(\pi) = \text{Dirichlet}(\alpha)$

$p(\Lambda_k) = \text{wishart}(\Lambda_0, \nu)$

$p(\mu_k \mid \Lambda_k) = \text{Normal}(0, (\beta \Lambda_k)^{-1})$

$p(z_n \mid \pi) = \text{categorical}(\pi)$

$p(x_n \mid z_{nk} = 1, \Lambda_k, \mu_k) = \text{Normal}(\mu_k, \Lambda_k^{-1})$

a) $p(\pi \mid \text{the other random variables})$ معادله (2)

$$= \text{Dirichlet}\left(\alpha_1 + \sum_{n=1}^{N} z_{n1}, \ldots, \alpha_K + \sum_{n=1}^{N} z_{nK}\right)$$

$$p(z) = \prod_{k=1}^{K} \pi_k^{z_k} \rightarrow p(z_n \mid \pi) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{nk}}$$

$$p(x \mid z) = \prod_{k=1}^{K} N(x \mid \mu_k, \Sigma_k)^{z_k}$$

$$p(x \mid z, \mu, \Lambda) = \prod \prod N(x_n \mid \mu_k, \Lambda_k^{-1})^{z_{nk}}$$

$$p(\pi) = \frac{\Gamma(k\alpha)}{\Gamma(\alpha)^k} \prod_{k=1}^{K} \pi_k^{\alpha - 1} = \alpha \prod_{k=1}^{K} \pi_k^{\alpha - 1}$$

$$p(\mu \mid \Lambda) = \prod_{k=1}^{K} N(\mu_k \mid 0, (\beta \Lambda_k)^{-1})$$

$$p(\Lambda) = \prod_{k=1}^{k} W(\Lambda_k \mid \Lambda_0, \nu)$$

$$\Rightarrow p(\mu, k) = p(\mu \mid \Lambda) \, p(\Lambda)$$

(1) $p(X, z, \pi, \mu, \Lambda) = p(X \mid z, \mu, \Lambda)\, p(z \mid \pi)\, p(\mu \mid \Lambda)\, p(\Lambda)$
$\times\, p(\pi)$

$$\mathcal{N}(x \mid \mu, \Lambda) = \frac{|\Lambda|^{1/2}}{(2\pi)^{D/2}} \exp\left[-\frac{1}{2}(x-\mu)^T \Lambda (x-\mu)\right]$$

$$\mathcal{W}(\Lambda \mid \Lambda_0, \nu) = \underbrace{B(\Lambda_0, \nu)}\, |\Lambda|^{\frac{(\nu-D-1)}{2}} \exp\left[-\tfrac{1}{2}\, tr(\Lambda_0^{-1} \cdot \Lambda)\right]$$

$$\left[|\Lambda_0|^{\frac{\nu}{2}}\, 2^{\frac{\nu D}{2}}\, \pi^{\frac{D(D-1)}{4}} \prod_{j=1}^{D} \Gamma\left(\frac{\nu}{2} - \frac{j-1}{2}\right)\right]^{-1}$$

(2) $p(\pi \mid \text{other})$

$$E_{-\pi}\left[\underbrace{\ln p(x, z, \pi, \mu, \Lambda)}\right] = E\left[\ln p(\pi \mid z, x, \mu, \Lambda)\right]$$

$$\ln\left[\underbrace{p(x \mid z, \mu, \Lambda)}_{\text{ثابت}}\, p(z \mid \pi)\, \underbrace{p(\mu \mid \Lambda)\, p(\Lambda)}_{\text{ثابت}}\, p(\pi)\right)$$

$$= E_{-\pi}\left[\underbrace{\ln p(z \mid \pi)}_{Cat(\pi)} + \underbrace{\ln p(\pi)}_{Dir(\alpha)}\right]$$

$$\ln p(\pi) = (\alpha - 1)\sum_{k=1}^{K} \ln(\pi_k) + \ln \alpha$$

$$\ln p(z \mid \pi) = \sum_{n=1}^{N}\sum_{k=1}^{K} z_{nk} \ln(\pi_k)$$

$$E\left[\ln p(\pi \mid \text{other})\right] = E_{-\pi}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} z_{nk} \ln(\pi_k) + (\alpha - 1)\sum_{k=1}^{K} \ln(\pi_k) + \ln \alpha\right]$$

$$\xrightarrow{\ln \text{ حذف}} E_{-\pi}\left(\underbrace{\alpha \prod_{k=1}^{K} \pi_k^{\sum_{n=1}^{N} z_{nk}} \prod_{k=1}^{K} \pi_k^{\alpha-1}}_{\text{فرم درکله}}\right) = E_{-\pi}\left(\alpha \prod_{k=1}^{K} \pi_k^{\alpha - 1 + \sum_{n=1}^{N} z_{nk}}\right)$$

$$Dir\left(\alpha_1 + \sum_{n=1}^{N} z_{n1}, \ldots, \alpha_K + \sum_{n=1}^{N} z_{nk}\right)$$

(3) معادله

$$p(\Lambda|other) = Wishart\left(\left(\Lambda_0^{-1} + \beta \mu_K \mu_K^T + \sum_{n=1}^{N} (x_n - \mu_K)(x_n - \mu_K)^T z_{nK}\right)^{-1}, v + 1 + \sum_{n=1}^{N} z_{nK}\right)$$

$$E[\ln p(\Lambda|other)] = E_{\cdot\Lambda_j}[\ln(all)] + const$$

$$\ln(all) = \underbrace{\sum_{i=1}^{n} \ln p(x_i|z_i, \mu_{z_i}, \Lambda_{z_i})}_{(1)} + \underbrace{\sum_{i=1}^{n} \ln(p(z_i|\pi)) +}_{constant}$$

$$\underbrace{\sum_{i=1}^{k} \ln p(\mu_i|\Lambda_i)}_{\ln p(\mu_j|\Lambda_j)} + \underbrace{\sum_{i=1}^{k} \ln p(\Lambda_i)}_{\ln p(\mu_j)} + \underbrace{\ln p(\pi)}_{constant}$$

$$(1)\ p(x_i|z_i, \Lambda, \mu) = \prod_{k=1}^{k} p(x_i|\mu_k, \Lambda_k)^{z_{ik}}$$

$$(1) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \ln p(x_i|\mu_k, \Lambda_k)$$

$$= \sum_{i=1}^{N} z_{ij} \ln p(x_i|\mu_j, \Lambda_j)$$

$$\ln p(\Lambda_j) = \ln B(\Lambda_0, v) + \left(\frac{v - D - 1}{2}\right) \log|\Lambda_j|$$

$$-1/2\ trace(\Lambda_0^{-1} \Lambda_j)$$

$$\ln p(\mu_j|\Lambda_j) = \ln \mathcal{N}(\mu_j|0, (\beta\Lambda_j)^{-1})$$

$$= \ln \frac{1}{(\beta\Lambda_j)^{-1/2}\sqrt{2\pi}} \exp\left[-1/2\left(\frac{x - \mu_i|0}{(\beta\Lambda_j)^{-1}}\right)^2\right]$$

$$= \ln \frac{(\beta\Lambda_j)^{1/2}}{\sqrt{2\pi}} - \underbrace{\frac{1}{2}(x - \mu_j)^2 \cdot (\beta\Lambda_j)^2}_{-1/2\ trace(\mu_K \mu_K^T \beta\Lambda_j)}$$

$$\ln \underbrace{p(\Lambda_j)}_{\text{wishart}} + E_{\mu_j}\left[\ln p(\mu_j \mid \Lambda_j)\right] + \sum_{i=1}^{n} \underbrace{E_{z_j}[z_j]}_{z_{nk}} E_{\mu_j}\left[\ln p(x_i \mid \mu_j, \Lambda_j)\right]$$

$$E_{\mu_j}\left[\underbrace{\ln p(x_i \mid \mu_j \Lambda_j)}\right]$$

$$\ln \frac{(\Lambda_j)^{1/2}}{(2\pi)^{d/2}} + \text{trace}\left[-1/2\,(x_i-\mu_j)^T \Lambda (x_i-\mu_j)\right]$$

$\Rightarrow$ $\ln$ حذف

$$p(\Lambda \mid \text{other}) = \text{wishart}(\Lambda_0, \nu) \times$$

$$\left(\frac{\beta \Lambda_j}{2\pi}\right)^{1/2} \times -\frac{1}{2}\,\text{trace}(\mu_k \mu_k^T \beta \Lambda_j)$$

$$\times \sum_{i=1}^{n} z_{nk} \cdot \frac{(\Lambda_j)^{1/2}}{(2\pi)^{d/2}} \cdot -\frac{1}{2}\,\text{trace}\left((x_i-\mu_j)^T \Lambda (x_i-\mu_j)\right]$$

$$= \text{wishart}\left((\Lambda_0^{-1} + \beta \mu_k \mu_k^T + \sum_{n=1}^{N} (x_n-\mu_k)(x_n-\mu_k)^T z_{nk})^{-1}\right.$$

$$\left., \nu + 1 + \sum_{n=1}^{N} z_{nk}\right)$$

ادامه(4)

$$p(\mu_k | other) \propto \exp\left[-\frac{1}{2}\mu_k^T(\beta\Lambda_k + \Lambda_k\sum_{n=1}^{N} z_{nk})\mu_k + \mu_k^T(\Lambda_k \sum_{n=1}^{N} x_n z_{nk})\right]$$

$$E_{-\mu_k}[\ln p(x, z, \pi, \mu, \Lambda)] =$$

$$E[\ln p(\mu_k | other)]$$

$$\Rightarrow \underbrace{\sum_{i=1}^{n} \ln p(x_i | z_k, \mu_k, \Lambda_k)} + \underbrace{\ln p(z|\pi)}_{constant} + E_{\Lambda_k}[\ln p(\mu_k|\Lambda_k)] + \underbrace{\ln p(\Lambda) + \ln p(\pi)}_{const}$$

$$\underbrace{E_{z_k}[z_k]}_{z_{nk}} \times E_{\Lambda_k}[\downarrow] \qquad \ln p(x_i | \mu_k, \Lambda_k)$$

$$p(x_i | \mu_k, \Lambda_k) = \frac{|\Lambda_k|^{1/2}}{(2\pi)^{d/2}} \exp\left[-\frac{1}{2}(x_i - \mu_k)^T \Lambda_k (x_i - \mu_k)\right]$$

$$p(\mu_k | \Lambda_k) = N(\mu_k | 0, (\beta\Lambda_k)^{-1})$$

$$= \left(\frac{\beta\Lambda_k}{2\pi}\right)^{1/2} \cdot \exp\left[-\frac{1}{2}(x - \mu_k)^2 (\beta\Lambda_k)^2\right]$$

$$\Rightarrow p(\mu_k | other) = \left(\frac{\beta\Lambda_k}{2\pi} \cdot \frac{|\Lambda_k|}{(2\pi)^d}\right)^{1/2} \cdot$$

$$\exp\left[\sum_{i=1}^{m} -\frac{1}{2}(x_i - \mu_k)^T \Lambda_k (x - \mu_k) z_{nk} \cdot - \frac{1}{2}(x - \mu_k)^2 (\beta\Lambda_k)^2\right]$$

$$\propto \exp\left[-\frac{1}{2}\mu_k^T(\beta\Lambda_k + \Lambda_k \sum_{i=1}^{N} z_{ik})\mu_k + \mu_k^T(\Lambda_k \sum_{i=1}^{N} x_i z_{ik})\right]$$

$$p(z_{nk}=1 \mid \text{other}) = \text{categorical}\left(\frac{\varepsilon_1}{\sum_{k=1}^{K}\varepsilon_k}, \dots, \frac{\varepsilon_k}{\sum_{k=1}^{K}\varepsilon_k}\right)$$

$$E_{-z_{nk}}[\ln p(\text{all})] = E[\ln p(z_{nk} \mid \text{other})]$$

$$\ln p(\text{all}) = \text{constant} + \sum_{i=1}^{N} \ln p(z_i \mid \pi) + \sum_{i=1}^{N} \ln p(x_i \mid z_i, \Lambda_{z_i}, \mu_{z_i})$$

$$\Rightarrow E_{-z_{nk}}\left[\sum_{n=1}^{N}\sum_{k=1}^{K} z_{nk} \ln \pi_k + \ln \mathcal{N}(x_n \mid \mu_k, \Lambda_k^{-1})\right]$$

$$= \sum_{n=1}^{N}\sum_{k=1}^{K}\left[z_{nk}\left[E_{\pi_k}(\ln \pi_k) + \frac{1}{2} E_{\Lambda_k}(\ln|\Lambda_k|)\right.\right.$$

$$\left.\left. - \frac{D}{2}\ln(2\pi) - \frac{1}{2} E_{\Lambda_k,\mu_k}[(x_n-\mu_k)^T \Lambda_k (x_n - \Lambda_k)]\right]\right]$$

$$\ln P_{nk} = E[\ln \pi_k \cdot \mathcal{N}(x_n \mid \mu_k, \Lambda_k^{-1})] \xrightarrow{\text{بتوان } e}$$ ،

$$p(z_{nk}=1 \mid \text{other}) = \prod_{n=1}^{N}\prod_{k=1}^{K}\left(\frac{P_{nk}}{\sum_{j=1}^{K} P_{nj}}\right)^{z_{nk}}$$

$$\Rightarrow \text{cat}\left(\frac{P_{n1}}{\sum_{j=1}^{K} P_{nj}}, \dots, \frac{P_{nk}}{\sum_{j=1}^{K} P_{nj}}\right)$$

$$\Rightarrow P_{nk} \text{ or } \varepsilon_k = \pi_k \mathcal{N}(x_n \mid \mu_k, \Lambda_k^{-1})$$

94

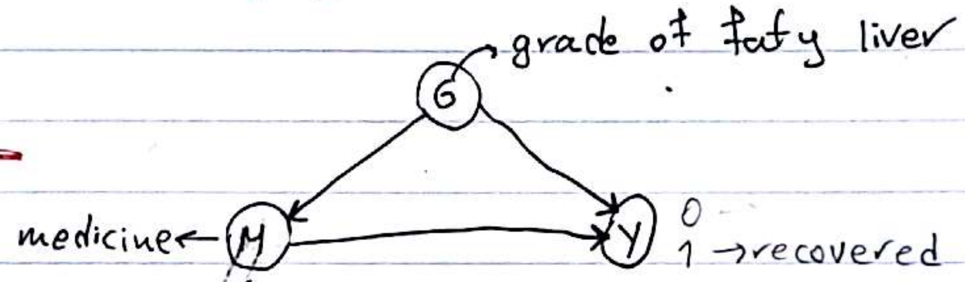
| | A (heated) | | B | |
|---|---|---|---|---|
| fatty liver grade 1 | 81/87 | =0.93 | 234/270 | 0.86 |
| fatty liver grade 4 | 192/263 | 0.73 | 55/80 | 0.68 |
| total | 273/350 | 0.78 | 289/330 | 0.82 |

$$\frac{87}{350}\times 0.93 + \frac{263}{350}\times 0.73 = 0.78$$
$$0.2311 \qquad 0.55$$

$$\frac{270}{350}\times 0.86 + \frac{80}{350}\times 0.68$$
$$0.66 \qquad 0.15 = 0.81$$

=> simpson's paradox

senario 1/




اگر grade بیماری در انتخاب نوع دارو تاثیرگذار باشد یعنی به این صورت تقسیم بگیریم که اگر بیمار دارای grade 4 بود داروی A و اگر دارای grade 1 بود داروی B تجویز می شود حال علت شفای بیمار در حالت کلی به grade بیماری نیز وابسته است چراکه در داروی A برای بیماران grade4 تجویز شده که حال وخیمی داشته اند پس احتمال شفا در این بیماران کم بوده پس درصد کمتری از خوب شدن در total داروی A دیده می شود. این در حالی است که داروی B برای بیمارانی تجویز شده که دارای grade1 بوده اند و احتمال خوب شدن آنها بیشتر بوده است پس در total دارو B تاثیرگذاشته با این سناریو در نهایت داروی A بهتر است.

$$\boxed{0.832} = P(Y=1 \mid do(M=A)) = P(Y=1 \mid M=A, G=1)\,P(G=1)$$
$$+ P(Y=1 \mid M=A, G=4)\,P(G=4) = 0.93\times\frac{357}{700} + 0.73\times\frac{343}{700}$$

$$\boxed{0.78} = P(Y=1 \mid do(M=B)) = P(Y=1 \mid M=B, G=1)\,P(G=1)$$
$$+ P(Y=1 \mid M=B, G=4)\,P(G=4) = 0.86\times\frac{357}{700} + 0.68\times\frac{343}{700}$$

طبق فرمول ها داروی A بهتر است