STAT 543

HW 5 Solution

Spring 2016

#2.2.30

sol) From the hint, WLOG, we can suppose $n_1 = n_2 = \cdots = n_q = 0$ and $n_{q+1}, \cdots, n_k > 0$.

Then, $L(\underset{\sim}{\theta}) = p(\underset{\sim}{z}; \underset{\sim}{\theta}) = \prod_{j=q+1}^{k} \theta_j^{n_j}$

$$\Rightarrow \quad l(\underset{\sim}{\theta}) \equiv \log L(\underset{\sim}{\theta}) = \sum_{j=q+1}^{k} n_j \log \theta_j$$

$$= \sum_{j=q+1}^{k-1} n_j \log \theta_j + n_k \log\left(1 - \sum_{j=q+1}^{k-1} \theta_j\right) \quad \text{as} \quad \theta_k = 1 - \sum_{j=q+1}^{k-1} \theta_j .$$

$$\Rightarrow \quad \frac{\partial l(\underset{\sim}{\theta})}{\partial \theta_j} = \frac{n_j}{\theta_j} + n_k \cdot \frac{(-1)}{1 - \sum_{j=q+1}^{k-1} \theta_j} = \frac{n_j}{\theta_j} - \frac{n_k}{\theta_k} \overset{set}{=} 0 , \quad j = q+1, \cdots, k-1 .$$

$$\Rightarrow \quad \begin{cases} \dfrac{n_j}{\theta_j} = \dfrac{n_k}{\theta_k} , & j = q+1, \cdots, k-1 \\ \theta_k = 1 - \sum_{j=q+1}^{k-1} \theta_j \end{cases}$$

Thus, the solutions satisfy
$$\begin{cases} \dfrac{\hat\theta_j}{\hat\theta_k} = \dfrac{n_j}{n_k} , & j = q+1, \cdots, k-1 \\ \hat\theta_k = 1 - \sum_{j=q+1}^{k-1} \dfrac{n_j}{n_k} \hat\theta_k . \end{cases}$$

$$\therefore \quad \hat\theta_k = \frac{n_k}{n} \quad \text{and} \quad \hat\theta_j = \frac{n_j}{n} , \quad j = q+1, \cdots, k-1.$$

Since $\dfrac{\partial^2 l(\underset{\sim}{\theta})}{\partial \theta_j^2} = -\dfrac{n_j}{\theta_j^2} < 0$ and
$$|J| = \begin{vmatrix} -\frac{n_1}{\theta_1^2} & 0 & 0 & \cdots & 0 \\ 0 & -\frac{n_2}{\theta_2^2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ & & & & -\frac{n_{k-1}}{\theta_{k-1}^2} \end{vmatrix} > 0 ,$$

$$\underset{\sim}{\hat\theta}_{MLE} = (\hat\theta_1, \hat\theta_2, \cdots, \hat\theta_k) = \left(\frac{n_1}{n}, \frac{n_2}{n}, \cdots, \frac{n_k}{n}\right) \quad //.$$

#2. 2.35

sol) $L(\theta) = p(\underset{\sim}{x};\theta) = \prod_{i=1}^{2} \dfrac{1}{\pi[1+(x_i-\theta)^2]}$

$\Rightarrow \ell(\theta) \equiv \log L(\theta) = -\sum_{i=1}^{2} \log\left(\pi[1+(x_i-\theta)^2]\right) = -2\log\pi - \sum_{i=1}^{2} \log[1+(x_i-\theta)^2]$

$\Rightarrow \ell'(\theta) = \sum_{i=1}^{2} \dfrac{2(x_i-\theta)}{1+(x_i-\theta)^2} \overset{\text{set}}{=} 0$ implies

$$\frac{2(x_1-\theta)}{1+(x_1-\theta)^2} + \frac{2(x_2-\theta)}{1+(x_2-\theta)^2} = 0$$

$\Rightarrow 2(x_1-\theta)[1+(x_2-\theta)^2] + 2(x_2-\theta)[1+(x_1-\theta)^2] = 0$

$\Rightarrow (x_1+x_2-2\theta) + (x_1-\theta)(x_2-\theta)(x_2-\theta+x_1-\theta) = 0$

$\Rightarrow [1+(x_1-\theta)(x_2-\theta)][x_1+x_2-2\theta] = 0$

$\Rightarrow [\theta^2-(x_1+x_2)\theta+x_1x_2+1][\theta-\frac{x_1+x_2}{2}] = 0$

$\Rightarrow [(\theta-\bar{x})^2-\Delta^2+1][\theta-\bar{x}] = 0$ where $\bar{x} = \dfrac{x_1+x_2}{2}$.

(a) $|\Delta| \le 1$.

i) Case 1. $\Delta = 1$: $\ell'(\theta) = 0$ implies $(\theta-\bar{x})^3 = 0$ $\therefore \hat{\theta} = \bar{x}$.

Case 2. $|\Delta| < 1$: Since $[(\theta-\bar{x})^2-\Delta^2+1] > 0$, $\ell''(\theta) = 0$ implies $(\theta-\bar{x}) = 0$ $\therefore \hat{\theta} = \bar{x}$.

Thus, $\ell'(\theta)$ has only one root $\hat{\theta} = \bar{x} = \dfrac{x_1+x_2}{2}$.

ii) $\ell''(\theta) = \sum_{i=1}^{2} \dfrac{-2[1+(x_i-\theta)^2] + 2(x_i-\theta)\cdot 2(x_i-\theta)}{[1+(x_i-\theta)^2]^2} = \sum_{i=1}^{2} \dfrac{2(x_i-\theta)^2-2}{[1+(x_i-\theta)^2]^2}$

$\Rightarrow \ell''(\hat{\theta}) = 2\cdot\left\{2\cdot\left(\frac{x_1-x_2}{2}\right)^2-2\right\}\left\{\sum_{i=1}^{2}\frac{1}{[1+(x_i-\theta)^2]^2}\right\} = 4(\Delta^2-1)\left\{\sum_{i=1}^{2}\frac{1}{[1+(x_i-\theta)^2]}\right\} \le 0$ //

(b) $|\Delta| > 1$

sol) $\quad l'(\theta) = 0 \implies [(\theta-\bar{x})^2 - \Delta^2 + 1](\theta - \bar{x}) = 0.$

Here, as $|\Delta| > 1$, $\quad (\theta-\bar{x})^2 - \Delta^2+1 = 0$ has 2 roots (quadratic equation of θ).

i.e., $\quad \theta^2 - 2\bar{x}\theta + \bar{x}^2 - \Delta^2 + 1 = 0$ has 2 roots as $\hat{\theta} = \bar{x} \pm \sqrt{\Delta^2-1}$

$\therefore$ $l'(\theta) = 0$ has 3 roots : $\hat{\theta}_1 = \bar{x}$, $\hat{\theta}_2 = \bar{x} + \sqrt{\Delta^2-1}$, $\hat{\theta}_3 = \bar{x} - \sqrt{\Delta^2-1}$

Note that

$$L(\hat{\theta}_1) = \prod_{i=1}^{2} \frac{1}{\pi[1+(x_i-\hat{\theta})^2]} = \frac{1}{\pi^2\left[1+\left(\frac{x_1-x_2}{2}\right)^2\right]^2}$$

$$L(\hat{\theta}_2) = \prod_{i=1}^{2} \frac{1}{\pi[1+(x_i-\hat{\theta}_2)^2]} = \frac{1}{\pi^2\left[1+\left(\frac{x_1-x_2}{2} - \sqrt{\Delta^2-1}\right)^2\right]\left[1+\left(\frac{x_2-x_1}{2} - \sqrt{\Delta^2-1}\right)^2\right]}$$

$$L(\hat{\theta}_3) = \prod_{i=1}^{2} \frac{1}{\pi[1+(x_i-\hat{\theta}_3)^2]} = \frac{1}{\pi^2\left[1+\left(\frac{x_1-x_2}{2} + \sqrt{\Delta^2-1}\right)^2\right]\left[1+\left(\frac{x_2-x_1}{2} + \sqrt{\Delta^2-1}\right)^2\right]}.$$

and

$$d_1 \overset{\text{let}}{=} \left[1+\left(\frac{x_1-x_2}{2}\right)^2\right]^2 = [1+\Delta^2]^2 = \Delta^4 + 2\Delta^2 + 1$$

$$d_2 \overset{\text{let}}{=} \left[1+\left(\frac{x_1-x_2}{2} - \sqrt{\Delta^2-1}\right)^2\right]\left[1+\left(\frac{x_2-x_1}{2} - \sqrt{\Delta^2-1}\right)^2\right] = [1+(\Delta-\sqrt{\Delta^2-1})^2][1+(\Delta-\sqrt{\Delta^2-1})^2]$$

$$= 1 + (\Delta-\sqrt{\Delta^2-1})^2 + (\Delta+\sqrt{\Delta^2-1})^2 + (\Delta^2-\Delta^2+1)^2 = 4\Delta^2.$$

$$d_3 \overset{\text{let}}{=} \left[1+\left(\frac{x_1-x_2}{2} + \sqrt{\Delta^2-1}\right)^2\right]\left[1+\left(\frac{x_2-x_1}{2} + \sqrt{\Delta^2-1}\right)^2\right] = [1+(\Delta+\sqrt{\Delta^2-1})^2][1+(\Delta-\sqrt{\Delta^2-1})^2]$$

$$= 4\Delta^2.$$

Since $\quad d_2 = d_3$ and $d_1 - d_2 = 1 + 2\Delta^2 + \Delta^4 - 4\Delta^2 = (1-\Delta^2)^2 \geq 0,$

$L(\hat{\theta}_2) = L(\hat{\theta}_3) \geq L(\hat{\theta}_1)$. Thus, $\hat{\theta}_{MLE} = \bar{x} \pm \sqrt{\Delta^2-1}$, not unique when $|\Delta| > 1$ //

#2.2.29.

sol) i) Since $S \sim Po(n\lambda)$, $S_1 \sim Po(m\lambda_1)$, $S_2 \sim Po(m\lambda_2)$,

S, S_1, S_2 are indep. and $\lambda = \lambda_1 + \lambda_2$,

$$L(\lambda_1, \lambda_2) = \frac{e^{-n(\lambda_1+\lambda_2)}[n(\lambda_1+\lambda_2)]^{S}}{S!} \cdot \frac{e^{-m\lambda_1}\cdot[m\lambda_1]^{S_1}}{S_1!} \cdot \frac{e^{-m\lambda_2}[m\lambda_2]^{S_2}}{S_2!}$$

$$\Rightarrow \ell(\lambda_1,\lambda_2) \equiv \log L(\lambda_1,\lambda_2) = -n(\lambda_1+\lambda_2) + S\log[n(\lambda_1+\lambda_2)] - \log S! - m\lambda_1 + S_1\log(m\lambda_1) - \log S_1!$$

$$-m\lambda_2 + S_2\log(m\lambda_2) - \log S_2!$$

$$\Rightarrow \begin{cases} \dfrac{\partial \ell(\lambda_1,\lambda_2)}{\partial\lambda_1} = -n + \dfrac{S}{\lambda_1+\lambda_2} - m + \dfrac{S_1}{\lambda_1} \overset{\text{set}}{=} 0 & \cdots \text{①} \\ \dfrac{\partial \ell(\lambda_1,\lambda_2)}{\partial\lambda_2} = -n + \dfrac{S}{\lambda_1+\lambda_2} - m + \dfrac{S_2}{\lambda_2} \overset{\text{set}}{=} 0 & \cdots \text{②} \end{cases}$$

$\Rightarrow$ ① $-$ ② $= 0$ implies $\dfrac{S_1}{\lambda_1} = \dfrac{S_2}{\lambda_2} \overset{\text{let}}{=} k$.

$\Rightarrow$ As $S_1 = k\lambda_1$ and $S_2 = k\lambda_2$, $S_1+S_2 = k(\lambda_1+\lambda_2)$ $\therefore \lambda_1+\lambda_2 = \frac{1}{k}(S_1+S_2)$

① $= -n + \dfrac{kS}{S_1+S_2} - m + k = 0$. $\therefore k = \dfrac{(n+m)(S_1+S_2)}{S+S_1+S_2}$

and $\dfrac{1}{k} = \dfrac{S+S_1+S_2}{(n+m)(S_1+S_2)}$

$$\therefore \hat{\lambda}_1 = \frac{S_1}{k} = \frac{(S+S_1+S_2)S_1}{(n+m)(S_1+S_2)}, \quad \hat{\lambda}_2 = \frac{S_2}{k} = \frac{(S+S_1+S_2)S_2}{(n+m)(S_1+S_2)}$$

ii) (1) $\dfrac{\partial^2 \ell(\lambda_1,\lambda_2)}{\partial\lambda_1^2} = -\dfrac{S}{(\lambda_1+\lambda_2)^2} - \dfrac{S_1}{\lambda_1^2} < 0$; $\dfrac{\partial^2 \ell(\lambda_1,\lambda_2)}{\partial\lambda_2^2} = -\dfrac{S}{(\lambda_1+\lambda_2)^2} - \dfrac{S_2}{\lambda_2^2} < 0$

(2)
$$\begin{vmatrix} \frac{\partial^2 l}{\partial \lambda_1^2} & \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_2} \\ \frac{\partial^2 l}{\partial \lambda_1 \partial \lambda_2} & \frac{\partial^2 l}{\partial \lambda_2^2} \end{vmatrix} = \begin{vmatrix} -\frac{S}{(\lambda_1+\lambda_2)^2} - \frac{S_1}{\lambda_1^2} & -\frac{S}{(\lambda_1+\lambda_2)^2} \\ -\frac{S}{(\lambda_1+\lambda_2)^2} & -\frac{S}{(\lambda_1+\lambda_2)^2} - \frac{S_2}{\lambda_2^2} \end{vmatrix}$$

$$= \left[\frac{S}{(\lambda_1+\lambda_2)^2} + \frac{S_1}{\lambda_1^2}\right]\left[\frac{S}{(\lambda_1+\lambda_2)^2} + \frac{S_2}{\lambda_2^2}\right] - \frac{S^2}{(\lambda_1+\lambda_2)^4}$$

$$= \frac{S}{(\lambda_1+\lambda_2)^2} \cdot \frac{S_2}{\lambda_2^2} + \frac{S}{(\lambda_1+\lambda_2)^2} \cdot \frac{S_1}{\lambda_1^2} + \left(\frac{S_1}{\lambda_1^2}\right)\left(\frac{S_2}{\lambda_2^2}\right) > 0.$$

Thus, $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are MLEs. //

Problem 2.

sol) Consider a model for $\underset{\sim}{x} = (W, Y)$ as

$$Y_i | W_i \overset{ind}{\sim} N(\theta_W, 1),$$
$$W_i \overset{iid}{\sim} Ber(\alpha), \qquad i = 1, \cdots, n.$$

Let $\eta = (\alpha, \theta_0, \theta_1)$.

$$L_x(\eta) = f(x; \eta) = \prod_{i=1}^{n} f(y_i | w_i; \eta) \cdot p(w_i; \eta)$$

$$= \prod_{i=1}^{n} \{\phi(y_i - \theta_1) I_{[w_i = 1]} + \phi(y_i - \theta_0) I_{[w_i = 0]}\} \cdot \alpha^{w_i} (1-\alpha)^{1-w_i}$$

$$= \prod_{i=1}^{n} [\alpha \cdot \phi(y_i - \theta_1)]^{w_i} \cdot [(1-\alpha)\phi(y_i - \theta_0)]^{1-w_i}$$

$$\Rightarrow l_x(\eta) \equiv \log L_x(\eta) = \sum_{i=1}^{n} [w_i \log \alpha + w_i \log \phi(y_i - \theta_1) + (1-w_i) \log(1-\alpha) + (1-w_i) \log \phi(y_i - \theta_0)]$$

(1) Pick a starting value $\eta^{(0)} = (\alpha^{(0)}, \theta_0^{(0)}, \theta_1^{(0)})$.

(2) E-step : For each iterate t, $(t = 0, 1, 2, \cdots, T)$,

$$E_{\eta^{(t)}}[l_X(\eta) \mid S(X) = y] = \sum_{i=1}^{n} \Big\{ E[w_i \mid y_i ; \eta^{(t)}] \cdot [\log\alpha + \log\phi(y_i - \theta_1)] +$$

$$(1 - E[w_i \mid y_i ; \eta^{(t)}]) \cdot [\log(1-\alpha) + \log\phi(y_i - \theta_0)] \Big\}$$

where $E[w_i \mid y_i ; \eta^{(t)}] = P(w_i = 1 \mid y_i ; \eta^{(t)})$

$$\overset{\text{Bayes thm}}{=} \frac{P(y_i \mid w_i = 1 ; \eta^{(t)}) \cdot P(w_i = 1 ; \eta^{(t)})}{P(y_i \mid w_i = 0 ; \eta^{(t)}) \cdot P(w_i = 0 ; \eta^{(t)}) + P(y_i \mid w_i = 1 ; \eta^{(t)}) \cdot P(w_i = 1 ; \eta^{(t)})}$$

$$= \frac{\phi(y_i - \theta_1^{(t)}) \cdot \alpha^{(t)}}{\phi(y_i - \theta_0^{(t)})(1 - \alpha^{(t)}) + \phi(y_i - \theta_1^{(t)}) \cdot \alpha^{(t)}} \equiv r_i^{(t)}$$

(3) M-step : Find a maximizer of $E_{\eta^{(t)}}[l_X(\eta) \mid S(X) = y] \equiv Q(\eta \mid \eta^{(t)})$, $\eta^{(t+1)}$

i) $$\frac{\partial Q}{\partial \alpha} = \frac{\sum_1^n r_i^{(t)}}{\alpha} - \frac{n - \sum_1^n r_i^{(t)}}{1-\alpha} \overset{\text{set}}{=} 0 \qquad \therefore \alpha^{(t+1)} = \frac{\sum_1^n r_i^{(t)}}{n}.$$

ii) $$\frac{\partial Q}{\partial \theta_0} = \sum_1^n (1 - r_i^{(t)}) \left[\frac{\partial}{\partial \theta_0} \left(- \frac{(y_i - \theta_0)^2}{2} \right) \right] = \sum_1^n (1 - r_i^{(t)})(y_i - \theta_0) \overset{\text{set}}{=} 0 \qquad \therefore \theta_0^{(t+1)} = \frac{\sum_1^n (1 - r_i^{(t)}) y_i}{\sum_1^n (1 - r_i^{(t)})}$$

iii) $$\frac{\partial Q}{\partial \theta_1} = \sum_1^n r_i^{(t)} \left[\frac{\partial}{\partial \theta_1} \left(- \frac{(y_i - \theta_1)^2}{2} \right) \right] = \sum_1^n r_i^{(t)} (y_i - \theta_1) \overset{\text{set}}{=} 0 \qquad \therefore \theta_1^{(t+1)} = \frac{\sum_1^n r_i^{(t)} y_i}{\sum_1^n r_i^{(t)}}$$

$$\Rightarrow \eta^{(t+1)} = (\alpha^{(t+1)}, \theta_0^{(t+1)}, \theta_1^{(t+1)}).$$

(4) Iterate step (2) and (3) until convergence. //

Problem 3.

① Study 1 : $\underset{\sim}{z}_{1i} \sim \text{Multinom}(1, \underset{\sim}{p})$, $i=1,\cdots,n_1$ with $n_1 = 100$

② Study 2 : $\underset{\sim}{z}_{2i} \sim \text{Multinom}(1, \underset{\sim}{p})$, $i=1,\cdots,n_2$ with $n_2 = 100$

③ Study 3 : $\underset{\sim}{z}_{3i} \sim \text{Multinom}(1, \underset{\sim}{p})$, $i=1,\cdots,n_3$ with $n_3 = 50$

where $\underset{\sim}{p} = (p_{++}, p_{+-}, p_{-+})$ and $p_{--} = 1-(p_{++}+p_{+-}+p_{-+})$

$$\Rightarrow L_x(\underset{\sim}{p}) = \prod_{l=1}^{3}\prod_{i=1}^{n_l} f(x_{li};\underset{\sim}{p}) = p_{++}^{\sum_1^3 n_{l++}} \times p_{+-}^{\sum_1^3 n_{l+-}} \times p_{-+}^{\sum_1^3 n_{l-+}} \times p_{--}^{\sum_1^3 n_{l--}}$$

where $\underset{\sim}{n}_l = (n_{l++}, n_{l+-}, n_{l-+}, n_{l--}) = \sum_{i=1}^{n_l} \underset{\sim}{z}_{li}$.

$$\Rightarrow l_x(\underset{\sim}{p}) = \left(\sum_1^3 n_{l++}\right)\log p_{++} + \left(\sum_1^3 n_{l+-}\right)\log p_{+-} + \left(\sum_1^3 n_{l-+}\right)\log p_{-+} + \left(\sum_1^3 n_{l--}\right)\log(1-p_{++}-p_{+-}-p_{-+})$$

(1) Pick a starting value $\underset{\sim}{p}^{(0)} = (p_{++}^{(0)}, p_{+-}^{(0)}, p_{-+}^{(0)})$. where $p_{--}^{(0)} = 1-p_{++}^{(0)}-p_{+-}^{(0)}-p_{-+}^{(0)}$.

(2) E-step : For each iterate t, $(t=1,2,\cdots,T)$,

$$\begin{aligned} E_{\underset{\sim}{p}^{(t)}}[l_x(\underset{\sim}{p}) \mid S(\underset{\sim}{z})=\underset{\sim}{y}] = &\{E_{\underset{\sim}{p}^{(t)}}[n_{1++} \mid n_{1+\cdot}=10] + E_{\underset{\sim}{p}^{(t)}}[n_{2++} \mid n_{2+\cdot}=20] + 4\}\cdot\log p_{++} \\ &+ \{E_{\underset{\sim}{p}^{(t)}}[n_{1+-} \mid n_{1+\cdot}=10] + E_{\underset{\sim}{p}^{(t)}}[n_{2+-} \mid n_{2+\cdot}=20] + 2\}\cdot\log p_{+-} \\ &+ \{E_{\underset{\sim}{p}^{(t)}}[n_{1-+} \mid n_{1-\cdot}=90] + E_{\underset{\sim}{p}^{(t)}}[n_{2-+} \mid n_{2-\cdot}=80] + 8\}\cdot\log p_{-+} \\ &+ \{E_{\underset{\sim}{p}^{(t)}}[n_{1--} \mid n_{1-\cdot}=90] + E_{\underset{\sim}{p}^{(t)}}[n_{2--} \mid n_{2-\cdot}=80] + 36\}\log(1-p_{++}-p_{+-}-p_{-+}). \end{aligned}$$

Since $n_{l++} \mid n_{l+\cdot} \sim \text{Bin}(n_{l+\cdot}, \frac{p_{++}}{p_{+\cdot}})$ where $n_{l+\cdot} = n_{l++}+n_{l+-}$ and $p_{+\cdot} = p_{++}+p_{+-}$, $l=1,2$

and $n_{l-+} \mid n_{l-\cdot} \sim \text{Bin}(n_{l-\cdot}, \frac{p_{-+}}{p_{-\cdot}})$ where $n_{l-\cdot} = n_{l-+}+n_{l--}$ and $p_{-\cdot} = p_{-+}+p_{--}$, $l=1,2$,

$$E_{p^{(t)}}[n_{1++} \mid n_{1+\cdot}=10] = 10 \cdot \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}}, \qquad E_{p^{(t)}}[n_{2++} \mid n_{2+\cdot}=20] = 20 \cdot \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}}.$$

$$E_{p^{(t)}}[n_{1+-} \mid n_{1+\cdot}=10] = 10 \cdot \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}}, \qquad E_{p^{(t)}}[n_{2+-} \mid n_{2+\cdot}=20] = 20 \cdot \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}},$$

$$E_{p^{(t)}}[n_{1-+} \mid n_{1-\cdot}=90] = 90 \cdot \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}}, \qquad E_{p^{(t)}}[n_{2-+} \mid n_{2-\cdot}=80] = 80 \cdot \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}}$$

$$E_{p^{(t)}}[n_{1--} \mid n_{1-\cdot}=90] = 90 \cdot \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} \qquad E_{p^{(t)}}[n_{2--} \mid n_{2-\cdot}=80] = 80 \cdot \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}}.$$

(3) M-step : Find a maximizer of $E_{p^{(t)}}[\ell_X(p) \mid s(z)=y] \equiv Q(p \mid p^{(t)})$, $p^{(t+1)}$

i) $$\frac{\partial Q}{\partial p_{++}} = \frac{10 \cdot \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}} + 20 \cdot \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}} + 4}{p_{++}} - \frac{90 \cdot \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 80 \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 36}{1 - p_{++} - p_{+-} - p_{-+}} \overset{set}{=} 0$$

ii) $$\frac{\partial Q}{\partial p_{+-}} = \frac{10 \cdot \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}} + 20 \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}} + 2}{p_{+-}} - \frac{90 \cdot \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 80 \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 36}{1 - p_{++} - p_{+-} - p_{-+}} \overset{set}{=} 0$$

iii) $$\frac{\partial Q}{\partial p_{-+}} = \frac{90 \cdot \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}} + 80 \cdot \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}} + 8}{p_{-+}} - \frac{90 \cdot \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 80 \frac{p_{--}^{(t)}}{p_{-\cdot}^{(t)}} + 36}{1 - p_{++} - p_{+-} - p_{--}} \overset{set}{=} 0$$

The solutions of the system of equations i), ii) and iii) are

$$p_{++}^{(t+1)} = \frac{10 \cdot \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}} + 20 \frac{p_{++}^{(t)}}{p_{+\cdot}^{(t)}} + 4}{250},$$

$$p_{+-}^{(t+1)} = \frac{10 \cdot \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}} + 20 \cdot \frac{p_{+-}^{(t)}}{p_{+\cdot}^{(t)}} + 2}{250},$$

$$p_{-+}^{(t+1)} = \frac{90 \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}} + 80 \cdot \frac{p_{-+}^{(t)}}{p_{-\cdot}^{(t)}} + 8}{250}$$

(4) Iterate step (2) and (3) until convergence. //

7.2.1.

sol) $\pi(\theta|\underset{\sim}{x}) \propto f(\underset{\sim}{x}|\theta)\cdot\pi(\theta) \propto \exp\left[-\frac{1}{2\sigma^2}\sum_1^n (x_i-\theta)^2\right]\cdot 1$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\left(n\theta^2-2\sum_1^n x_i\theta\right)\right]$$

$$\propto \underbrace{\exp\left[-\frac{n}{2\sigma^2}(\theta-\bar{x})^2\right]}_{\text{a kenel of Normal dist. family}}$$

$\Rightarrow \theta|\underset{\sim}{x} \sim N(\bar{x}, \frac{\sigma^2}{n})$

$\therefore$ The Improper Bayes rule for squared error loss is

$$\delta^*(\underset{\sim}{x}) = E[\theta|\underset{\sim}{x}] = \bar{x}.$$

//

7.2.2.

sol) ① $\pi(\theta|\underset{\sim}{x}) \propto f(\underset{\sim}{x}|\theta)\cdot\pi(\theta) \propto \theta^{\sum_1^n x_i}(1-\theta)^{n-\sum_1^n x_i}\cdot\theta^{r-1}(1-\theta)^{s-1}$

$$\propto \underbrace{\theta^{\sum_1^n x_i+r-1}(1-\theta)^{n-\sum_1^n x_i+s-1}}_{\text{a kenel of beta dist. family}}$$

$\Rightarrow \theta|\underset{\sim}{x} \sim \text{beta}(\sum_1^n x_i+r,\ n-\sum_1^n x_i+s)$

$\Rightarrow \hat{\theta}_B = E[\theta|\underset{\sim}{x}] = \dfrac{\sum_1^n x_i + r}{n+r+s} = \dfrac{n\bar{x}+(r+s)\cdot\frac{r}{r+s}}{n+r+s} = \dfrac{n}{n+r+s}\bar{x} + \dfrac{r+s}{n+r+s}\cdot\theta_0$

$\therefore \hat{\theta}_B = w\theta_0 + (1-w)\bar{x}$ where $w = \dfrac{r+s}{n+r+s}$

② Note that $\text{Unif}(0,1) \overset{d}{=} \text{beta}(1,1)$ $\therefore \hat{\theta}_B = \dfrac{\sum_1^n x_i+1}{n+2} = \dfrac{S+1}{n+2}$ where $S = \sum_1^n x_i$

//

#3.2.3.

Sol) ① $L(\theta) = f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_1^n x_i}(1-\theta)^{n-\sum_1^n x_i}$

$\Rightarrow \ell(\theta) \equiv \log L(\theta) = \sum_1^n x_i \log\theta + (n-\sum_1^n x_i)\log(1-\theta)$

$\Rightarrow \ell'(\theta) = \dfrac{\sum_1^n x_i}{\theta} - \dfrac{n-\sum_1^n x_i}{1-\theta} \overset{set}{=} 0 \qquad \therefore \hat\theta = \dfrac{\sum_1^n x_i}{n} = \bar x.$

Since $\ell''(\theta) = -\dfrac{\sum_1^n x_i}{\theta^2} - \dfrac{(n-\sum_1^n x_i)}{(1-\theta)^2} < 0, \ \forall\theta, \qquad \hat\theta_{MLE} = \bar x.$

By invariance property of MLE, $\quad \widehat{g(\theta)}_{MLE} = \hat\theta(1-\hat\theta) = \bar x(1-\bar x).$

② Since $\theta|\underset{\sim}{x} \sim beta(\sum_1^n x_i + r,\ n-\sum_1^n x_i + s)$, $\qquad E[\theta|\underset{\sim}{x}] = \dfrac{\sum_1^n x_i + r}{n+r+s}$

Since $Var(\theta|\underset{\sim}{x}) = \dfrac{(\sum_1^n x_i+r)(n-\sum_1^n x_i+s)}{(n+r+s)^2(n+r+s+1)}$,

$$E[\theta^2|\underset{\sim}{x}] = Var[\theta|\underset{\sim}{x}] + \left(E[\theta|\underset{\sim}{x}]\right)^2 = \frac{(\sum_1^n x_i+r)(n-\sum_1^n x_i+s)}{(n+r+s)^2(n+r+s+1)} + \frac{(\sum_1^n x_i+r)^2}{(n+r+s)^2}$$

$$= \frac{(\sum_1^n x_i+r)\left[n-\sum_1^n x_i+s+(\sum_1^n x_i+r)(n+r+s+1)\right]}{(n+r+s)^2(n+r+s+1)} = \frac{(\sum_1^n x_i+r)(\sum_1^n x_i+r+1)}{(n+r+s)(n+r+s+1)}$$

$$\Rightarrow E[g(\theta)|\underset{\sim}{x}] = E[\theta|\underset{\sim}{x}] - E[\theta^2|\underset{\sim}{x}] = \frac{\sum_1^n x_i+r}{n+r+s} - \frac{(\sum_1^n x_i+r)(\sum_1^n x_i+r+1)}{(n+r+s)(n+r+s+1)} = \frac{(\sum_1^n x_i+r)(n-\sum_1^n x_i+s)}{(n+r+s)(n+r+s+1)}$$

but

$$g(\hat\theta_B) = \hat\theta_B(1-\hat\theta_B) = \left(\frac{\sum_1^n x_i+r}{n+r+s}\right)\left(1-\frac{\sum_1^n x_i+r}{n+r+s}\right) = \frac{(\sum_1^n x_i+r)(n+s-\sum_1^n x_i)}{(n+r+s)^2}$$

$\therefore \ E[g(\theta)|\underset{\sim}{x}] \neq g(\hat\theta_B)$

//

7.2.4.

sol)

$$\pi(x) \propto 1 \quad \text{and} \quad x = \frac{\theta}{1-\theta}, \quad 0 < x < \infty$$

$$\Rightarrow \pi(\theta) \propto \pi(x) \cdot \left| \frac{dx}{d\theta} \right| = \frac{1}{(1-\theta)^2}, \quad 0 < \theta < 1.$$

$$\Rightarrow \pi(\theta \mid \underline{x}) \propto f(\underline{x} \mid \theta) \cdot \pi(\theta) \propto \theta^{\sum_1^n x_i} (1-\theta)^{n - \sum_1^n x_i} \cdot (1-\theta)^{-2}$$

$$= \underbrace{\theta^{\sum_1^n x_i} (1-\theta)^{n - \sum_1^n x_i - 2}}_{} \quad \text{a kernel of beta dist. family}$$

$$\therefore \theta \mid \underline{x} \sim \text{beta}\left(\sum_1^n x_i + 1, \ n - \sum_1^n x_i - 1 \right)$$

$$\hat{\theta}_B = \frac{\sum_1^n x_i + 1}{n} = \frac{S' + 1}{n} \quad \text{exists if} \quad 0 < S' < n-1 \quad \text{where} \quad S' = \sum_1^n x_i .$$

//

Problem 5.

(a) Since $K \le 2.72$, take $M = 2.72$ and $h(x) = \phi(x)$.

Then, $f(x) = K\phi(x)\sin^2 x \le M \cdot h(x)$.

$\Rightarrow$ The rejection algorithm for sampling from $f(x)$.

① generate $x^{**} \sim h(x)$ i.e., $x^{**} \sim N(0,1)$

② Independently, generate $u \sim U(0,1)$

③ If $M \cdot u \cdot h(x^{**}) < K \cdot \phi(x^{**}) \sin^2(x^{**})$, then set $x^* = x^{**}$,

o.w., return to ①.

//

(b) let $N = 10{,}000$.

$$\hat{EX} = \frac{1}{N}\sum_{1}^{N} x_i^* \xrightarrow{P} EX \quad \text{by WLLN}$$

$$\hat{Var}X = \frac{1}{N-1}\sum_{1}^{N} (x_i^* - \hat{Ex})^2 \xrightarrow{P} \sigma^2 \quad \text{by using WLLN.}$$

Since $P[0.3 < X < 1.2] = E I_{[0.3 < X < 1.2]}$,

$$\hat{p} = \frac{1}{N}\sum_{i=1}^{N} I_{[0.3 < x_i^* < 1.2]} \xrightarrow{P} P[0.3 < x < 1.2] \quad \text{by WLLN.}$$

//

Problem 6.

(a) Since $h(\mathbf{x}) \le e^1$, take $M = e^1$ and $g(\mathbf{x}) = 1$, then $h(\mathbf{x}) \le M \cdot g(\mathbf{x})$

① Generate $x_i^{**} \sim g(\mathbf{x})$, i.e., $x_i^{**} \sim U(0,1)$, $i = 1, \cdots, 10$.

② Independently, generate $u \sim U(0,1)$

③ If $M \cdot u \cdot g(\mathbf{x}^{**}) \le h(\mathbf{x}^{**})$, set $\mathbf{x}^* = \mathbf{x}^{**}$ where $\mathbf{x}^{**} = (x_1^{**}, x_2^{**}, \dots, x_{10}^{**})$.

o.w. return to ①

(b) For N iterations where N is large enough,

$$\hat{EX_1X_2} = \frac{1}{N}\sum_{1}^{N} x_{i1}^* x_{i2}^* \quad \text{is an estimator based on simulation.}$$

By LNN, $\hat{EX_1X_2} \xrightarrow{P} EX_1X_2$. //

Problem 7.

(a) We have $x_i \overset{\text{ind}}{\sim} Po(\lambda_i)$, where $\lambda_1 = \lambda_2 = \cdots = \lambda_M \equiv \mu_1$, $\lambda_{M+1} = \cdots = \lambda_N \equiv \mu_2$,

$\mu_1 \sim Exp(1)$, $\mu_2 \sim Exp(1)$, $M \sim \text{Unif}\{1, 2, \cdots, N\}$,

μ_1, μ_2 and M are mutually indep.

$$\Rightarrow G(\mu_1, \mu_2, M \mid \underline{x}) = \frac{f(\underline{x} \mid \mu_1, \mu_2, M) \cdot G(\mu_1, \mu_2, M)}{\sum_{M=1}^{N} \int_0^\infty \int_0^\infty f(\underline{x} \mid \mu_1, \mu_2, M) \cdot G(\mu_1, \mu_2, M) \, d\mu_1 d\mu_2}$$

$$= \frac{\left[\prod_{i=1}^{M} \frac{e^{-\mu_1} \mu_1^{x_i}}{x_i!}\right]\left[\prod_{j=M+1}^{N} \frac{e^{-\mu_2} \mu_2^{x_j}}{x_j!}\right] \cdot \frac{1}{N} \cdot e^{-\mu_1} \cdot e^{-\mu_2}}{\sum_{M=1}^{N} \int_0^\infty \int_0^\infty \left[\prod_{i=1}^{M} \frac{e^{-\mu_1} \mu_1^{x_i}}{x_i!}\right]\left[\prod_{j=M+1}^{N} \frac{e^{-\mu_2} \mu_2^{x_j}}{x_j!}\right] \cdot \frac{1}{N} \cdot e^{-\mu_1} e^{-\mu_2} \, d\mu_1 d\mu_2}$$

$$= \frac{e^{-(M+1)\mu_1} \cdot \mu_1^{S_M} \cdot e^{-(N-M+1)\mu_2} \cdot \mu_2^{(T-S_M)}}{\sum_{M=1}^{N} \int_0^\infty \int_0^\infty \underbrace{e^{-(M+1)\mu_1} \cdot \mu_1^{S_M}} \, \underbrace{e^{-(N-M+1)\mu_2} \cdot \mu_2^{(T-S_M)}} \, d\mu_1 d\mu_2}$$

each is the kernel of gamma dist. family

$\Rightarrow$ The integral of the denominator can be evaluated explicitly so can the summation.

$\Rightarrow$ Also, $G(M \mid \underline{x}) = \int_0^\infty \int_0^\infty G(\mu_1, \mu_2, M \mid \underline{x}) \, d\mu_1 d\mu_2$ is the product of two gamma integrals with multiplication of some scalar.

This can be also evaluated explicitly. //

(b)

sol) The posterior dist. of (μ_1, μ_2, M) is

$$g(\mu_1, \mu_2, M \mid \underset{\sim}{z}) = \frac{e^{-M\mu_1} \cdot \mu_1^{S_M} \, e^{-(N-M)\mu_2} \, \mu_2^{(T-S_M)} \cdot g(\mu_1, \mu_2)}{\sum_{M=1}^{N} \int_0^\infty \int_0^\infty e^{-M\mu_1} \mu_1^{S_M} \cdot e^{-(N-M)\mu_2} \mu_2^{(T-S_M)} \, g(\mu_1, \mu_2) \, d\mu_1 \, d\mu_2}$$

$\Rightarrow$ Gibbs sampling algorithm :

(1) Generate $M^{(0)} \sim \text{Unif}\{1, 2, \cdots, N\}$ and $(\mu_1^{(0)}, \mu_2^{(0)}) \sim g(\mu_1, \mu_2)$.

(2) For each iteration, t $(t = 1, 2, \cdots, T)$, generate $(\mu_1^{(t+1)}, \mu_2^{(t+1)}, M^{(t+1)})$ as follows.

i) Generate $M^{(t+1)} \sim g(M \mid \mu_1^{(t)}, \mu_2^{(t)}, \underset{\sim}{z})$:

Since $g(M \mid \mu_1^{(t)}, \mu_2^{(t)}, \underset{\sim}{z}) \propto e^{-M(\mu_1^{(t)} - \mu_2^{(t)})} \cdot \left[\frac{\mu_1^{(t)}}{\mu_2^{(t)}} \right]^{S_M}$,

take $C = N \cdot \max_{M \in \{1, \cdots, N\}} e^{-M(\mu_1^{(t)} - \mu_2^{(t)})} \cdot \left[\frac{\mu_1^{(t)}}{\mu_2^{(t)}} \right]^{S_M}$ and $h(M) = \frac{1}{N}$.

$\Rightarrow$ $C \cdot h(M) \geq e^{-M(\mu_1^{(t)} - \mu_2^{(t)})} \cdot \left[\frac{\mu_1^{(t)}}{\mu_2^{(t)}} \right]^{S_M}$.

By using the rejection sampling algorithm, generate $M^{(t+1)}$ from $g(M \mid \mu_1^{(t)}, \mu_2^{(t)}, \underset{\sim}{z})$.

ii) Generate $\mu_1^{(t+1)} \sim g(\mu_1 \mid M^{(t+1)}, \mu_2^{(t)}, \underset{\sim}{z})$:

Since $g(\mu_1 \mid M^{(t+1)}, \mu_2^{(t)}, \underset{\sim}{z}) \propto e^{-M^{(t+1)}\mu_1} \cdot \mu_1^{S_{M^{(t+1)}}} \cdot g(\mu_1 \mid \mu_2^{(t)})$,

take $C_1 = \max_{\mu_1} e^{-M^{(t+1)}\mu_1} \cdot \mu_1^{S_{M^{(t+1)}}}$ and $h_1(\mu_1) = g(\mu_1 \mid \mu_2^{(t)})$

$\Rightarrow$ $C_1 \cdot h_1(\mu_1) \geq e^{-M^{(t+1)}\mu_1} \cdot \mu_1^{S_{M^{(t+1)}}} \cdot g(\mu_1 \mid \mu_2^{(t)})$

By using the rejection sampling algorithm, generate $\mu_1^{(t+1)}$ from $g(\mu_1 \mid M^{(t+1)}, \mu_2^{(t)}, \underset{\sim}{z})$

iii) Generate $\mu_2^{(t+1)} \sim g(\mu_2 \mid M^{(t+1)}, \mu_1^{(t+1)}, \underset{\sim}{z})$:

Since $g(\mu_2 \mid M^{(t+1)}, \mu_1^{(t+1)}, \underset{\sim}{z}) \propto e^{-(N-M^{(t+1)})\mu_2} \cdot \mu_2^{(T-S_{M^{(t+1)}})} \cdot g(\mu_2 \mid \mu_1^{(t+1)})$,

take $C_2 = \max_{\mu_2} e^{-(N-M^{(t+1)})\mu_2} \cdot \mu_2^{(T-S_{M^{(t+1)}})}$ and $h_2(\mu_2) = g(\mu_2 \mid \mu_1^{(t+1)})$,

By using the rejection sampling algorithm, generate $\mu_2^{(t+1)}$ from $g(\mu_2 \mid M^{(t+1)}, \mu_1^{(t+1)}, \underset{\sim}{z})$.

(3) Go back to step (2) and iterate to convergence.

Problem 8.

(a)

We can get the conditional (posterior) distributions as follows:

$P(\theta_1 \mid \theta_2)$

θ_1 \ θ_2	1	2	3	4
4	0	0	$\frac{0.2}{0.25}$	$\frac{0.1}{0.15}$
3	0	0	$\frac{0.05}{0.25}$	$\frac{0.05}{0.15}$
2	$\frac{0.2}{0.4}$	$\frac{0.1}{0.2}$	0	0
1	$\frac{0.2}{0.4}$	$\frac{0.1}{0.2}$	0	0

$P(\theta_2 \mid \theta_1)$

θ_1 \ θ_2	1	2	3	4
4	0	0	$\frac{0.2}{0.3}$	$\frac{0.1}{0.3}$
3	0	0	$\frac{0.05}{0.1}$	$\frac{0.05}{0.1}$
2	$\frac{0.2}{0.3}$	$\frac{0.1}{0.3}$	0	0
1	$\frac{0.2}{0.3}$	$\frac{0.1}{0.3}$	0	0

$\Rightarrow$ Suppose that we use Gibbs sampling to produce the simulated values, by generating θ_1^* first and then generating θ_2^* from $P(\theta_2 \mid \theta_1^*)$.

If θ_1^* was in state 1 or 2, θ_2^* can never get to state 3 or 4, vice versa. Then, $P(\theta_1 \le 2) = 1$ or 0.

$\therefore$ The Gibbs sampling does not work in this case. //

(b)

sol) ① Note that $P(\underset{\sim}{\theta}_i^* \mid \underset{\sim}{\theta}_{i-1}^*) \propto J(\underset{\sim}{\theta}_i^* \mid \underset{\sim}{\theta}_{i-1}^*) \cdot a_i$ where $J(\underset{\sim}{\theta}_i^* \mid \underset{\sim}{\theta}_{i-1}^*)$ denotes the proposal dist. of $\underset{\sim}{\theta}_i^*$ given $\underset{\sim}{\theta}_{i-1}^*$ and a_i denotes the acceptance rate, $\min(r_i, 1)$ with $r_i = \dfrac{p(\underset{\sim}{\theta}_i^*)}{p(\underset{\sim}{\theta}_{i-1}^*)}$ (as $J_i(\underset{\sim}{\theta}' \mid \underset{\sim}{\theta})$ is uniform).

i)

$\underset{\sim}{\theta} = (\theta_1, \theta_2)$	(1,1)	(1,2)	(2,1)	(2,2)	(3,3)	(3,4)	(4,3)	(4,4)
$p(\underset{\sim}{\theta})$	0.2	0.1	0.2	0.1	0.05	0.05	0.2	0.1

ii) $J(\underset{\sim}{\theta}' \mid \underset{\sim}{\theta}) = \frac{1}{8}$ for any $\underset{\sim}{\theta}' \in \Theta$ given any fixed $\underset{\sim}{\theta} \in \Theta$

where $\Theta = \{(1,1), (1,2), (2,1), (2,2), (3,3), (3,4), (4,3), (4,4)\}$.

iii) Table : The moving ratio (r_i), acceptance rate (a_i), and conditional prob. $(P(\theta_i^* \mid \theta_{i-1}^*))$, $(r_i, a_i, P(\theta_i^* \mid \theta_{i-1}^*))$, for off-diagonal elements and conditional prob. for diagonal elements.

θ_{i-1}^* \ θ_i^*	(1,1)	(1,2)	(2,1)	(2,2)	(3,3)	(3,4)	(4,3)	(4,4)
(1,1)	$\frac{1}{2}$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$
(1,2)	$(2, 1, \frac{1}{8})$	$\frac{1}{4}$	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$
(2,1)	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$\frac{7}{16}$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$
(2,2)	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$\frac{1}{4}$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$
(3,3)	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$\frac{1}{8}$	$(1, 1, \frac{1}{8})$	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$
(3,4)	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$	$\frac{1}{8}$	$(4, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$
(4,3)	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{32})$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$
(4,4)	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$	$(2, 1, \frac{1}{8})$	$(1, 1, \frac{1}{8})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{16})$	$(2, 1, \frac{1}{8})$	$\frac{1}{4}$

* How can we get the conditional prob. of $\underset{\sim}{\theta}_t^*$ given $\underset{\sim}{\theta}_{t-1}^*$ for diagonal element?

i.e., The conditional probability of having the same state as the previous.

= The conditional probability of staying at the same state.

$$= 1 - \sum_{\underset{\sim}{\theta}' \neq \underset{\sim}{\theta}} P(\underset{\sim}{\theta}' \mid \underset{\sim}{\theta}).$$

$\Rightarrow$ For example, $P(\underset{\sim}{\theta}_t^* = (1,1) \mid \underset{\sim}{\theta}_{t-1}^* = (1,1)) = 1 - \left(\frac{1}{16} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{32} + \frac{1}{8} + \frac{1}{16}\right) = \frac{1}{2}$ //

②

$$P(\underset{\sim}{\theta}_t^* = \underset{\sim}{\theta}') = P(\underset{\sim}{\theta}_t^* = \underset{\sim}{\theta}' \mid \underset{\sim}{\theta}_{t-1}^* = \underset{\sim}{\theta}) \cdot P(\underset{\sim}{\theta}) = T^T \cdot \begin{bmatrix} 0.2 \\ 0.1 \\ 0.2 \\ 0.1 \\ 0.05 \\ 0.05 \\ 0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.1 \\ 0.2 \\ 0.1 \\ 0.05 \\ 0.05 \\ 0.2 \\ 0.1 \end{bmatrix}$$

where

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & \frac{1}{32} & \frac{1}{32} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{16} & \frac{7}{16} & \frac{1}{16} & \frac{1}{32} & \frac{1}{32} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{16} & \frac{1}{16} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{16} & \frac{1}{8} & \frac{1}{16} & \frac{1}{32} & \frac{1}{32} & \frac{1}{2} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} & \frac{1}{8} & \frac{1}{4} \end{bmatrix}$$

with $T = \{T_{jk}\}_{8\times 8}$ and $T_{jk} = P(\underset{\sim}{\theta}_t^* = k\text{th state} \mid \underset{\sim}{\theta}_t^* = j\text{th state})$, $j, k = 1, 2, \cdots, 8$. //