Problem 1:

a) E-step: $Q(\theta;\theta^0) = \underset{D_b}{E}\left[\log P(X_g, X_b;\theta) \mid (X_g, \theta^0)\right]$

$$D = \left\{\binom{1}{1}, \binom{3}{3}, \binom{2}{*}\right\} \rightarrow D_g = \{x_{11}, x_{12}, x_{21}, x_{22}, x_{31}\} \quad D_b = \{x_{32}\}$$

$$\rightarrow Q(\theta,\theta^0) \underset{\text{امیر یافتی}}{\overset{\text{طبق تعریف}}{=}} \int_{-\infty}^{\infty} \log\left(\prod_{k=1}^{3} P(x_k|\theta)\right) P(x_{32}|X_g, \theta^0)\, dx_{32}$$

$$= \int_{-\infty}^{\infty} \sum_{k=1}^{3} \log P(x_k|\theta) \frac{P(x_{32}, x_{31}=2 \mid \theta^0)}{P(x_{31}=2 \mid \theta^0)} dx_{32}$$

$$= \int_{-\infty}^{\infty} \underbrace{\sum_{k=1}^{2} \log P(x_k|\theta)}_{\text{ثابت}} \underbrace{\frac{P(x_{32}, x_{31}=2|\theta^0)}{P(x_{31}=2|\theta^0)} dx_{32}}_{\text{انتگرال}=1} + \int_{-\infty}^{\infty} \log P(x_3|\theta) \frac{P(x_{32}, x_{31}=2|\theta^0)}{P(x_{31}=2|\theta^0)}$$

$$= \sum_{k=1}^{2} \log P(x_k|\theta) + \int_{-\infty}^{\infty} \log P\left(\binom{2}{x_{32}}\Big|\theta\right) \times \frac{P\left(\binom{2}{x_{31}}\Big|\theta^0\right)}{\int_{-\infty}^{\infty} P\left(\binom{2}{x'_{32}}\Big|\theta^0\right) dx'_{32}} dx_{32}$$

$$\int_{-\infty}^{\infty} P\left(\binom{2}{x'_{32}}\Big|\theta^0\right) dx'_{32} = \int_0^4 \frac{1}{\theta_1^0} e^{-2\theta_1^0} \times \frac{1}{\theta_2^0} dx'_{32} \overset{\theta(2,4)}{=} \frac{1}{8} \int_0^4 e^{-4} dx'_{32} = \frac{1}{2e^4}$$

$$\rightarrow Q(\theta,\theta^0) = \sum_{k=1}^{2} \log P(x_k|\theta) + 2e^4 \begin{cases} \int_0^{\theta_2} \log\left(\frac{1}{\theta_1} e^{-2\theta_1} \times \frac{1}{\theta_2}\right) \times \frac{1}{2} e^{-4} \times \frac{1}{4} & 3 \leq \theta_2 \leq 4 \\ \int_0^4 \quad '' \qquad\qquad '' & \theta_2 \geq 4 \\ 0 & \text{otherwise} \end{cases}$$

C

$$\rightarrow Q(\theta,\theta^0) = \log\left(\frac{e^{-\theta_1}}{\theta_1} \times \frac{1}{\theta_2}\right) + \log\left(\frac{e^{-3\theta_1}}{\theta_1} \times \frac{1}{\theta_2}\right)$$

$$= -\theta_1 - \log(\theta_1\theta_2) - 3\theta_1 - \log(\theta_1\theta_2) + C = -4\theta_1 - 2\log(\theta_1\theta_2) + C$$

نرمالیزیشن $\rightarrow \int_{-\infty}^{\infty} p(x_1)\, dx_1 = 1 \rightarrow \int_{0}^{\infty} \frac{1}{\theta_1} e^{-\theta_1 x_1} dx_1 = 1 \rightarrow \frac{-1}{\theta_1^2} e^{-\theta_1 x_1}\Big|_0^{\infty} = 1 \rightarrow \boxed{\theta_1 = 1}$

$$C = \begin{cases} \frac{1}{4}\int_0^{\theta_2} \log\left(\frac{e^{-2\theta_1}}{\theta_1\theta_2}\right) dx_{32} & 3 \leq \theta_2\ 4 \\ \frac{1}{4}\int_0^{4} \log\left(\frac{e^{-2\theta_1}}{\theta_1\theta_2}\right) dx_{32} = \ln\left(\frac{e^{-2\theta_1}}{\theta_1\theta_2}\right) & \theta_2 \geq 4 \\ 0 & \text{Otherwise} \end{cases}$$

برابراست با:

b) M-step: $\underline{3 \leq \theta_2 \leq 4} \rightarrow Q(\theta;\theta^0) = -4\theta_1 - 2\log(\theta_1\theta_2) + \frac{1}{4}\int_0^{\theta_2} \log\left(\frac{e^{-2\theta_1}}{\theta_1\theta_2}\right) dx_{32}$

$$\overset{\theta_1=1}{=} -4 - 2\log(\theta_2) + \frac{1}{4} \times \theta_2 \times (-2 - \log(\theta_2))$$

$$= -4 - 2\log\theta_2 - \frac{(\log\theta_2)\theta_2}{2} - \frac{\theta_2}{2} \rightarrow$$ نزولی برحسب θ_2

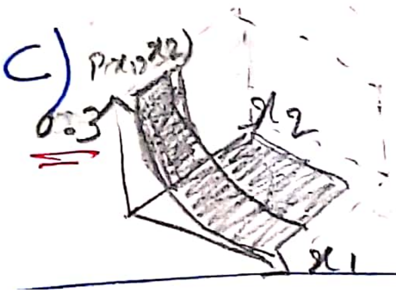
$$\rightarrow \underset{\theta_2}{\operatorname{argmax}}\, Q(\theta,\theta^0) = 3 \rightarrow Q \simeq -8$$

$\underline{\theta_2 \geq 4} \rightarrow Q(\theta,\theta^0) = -4 - 2\log(\theta_2) + \ln\left(\frac{e^{-2}}{\theta_2}\right) = -6 - 3\log\theta_2 \rightarrow$ نزولی برحسب θ_2

$$\underset{\theta_2}{\operatorname{argmax}}\, Q = 4 \rightarrow Q = 4$$

$$\rightarrow Q\Big|_{\theta_2=3} > Q\Big|_{\theta=4} \rightarrow \boxed{\theta_2 = 3}$$

c)



$\theta = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ ☆

$\theta^0 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

Problem 2:

a) $P(D|\theta) = \prod_{i=1}^{n} p(x_i|\theta) \longrightarrow MLE(\theta) = \underset{\theta}{\operatorname{argmax}} \prod_{i=1}^{n} p(x_i|\theta)$

$$p(x_i|\theta) = \begin{cases} \frac{1}{\theta} & 0 \leq x_i \leq \theta \\ 0 & x_i > \theta \end{cases} \longrightarrow MLE(\theta) = \underset{\theta}{\operatorname{argmax}} \begin{cases} \frac{1}{\theta^n} & \forall_i \ 0 \leq x_i \leq \theta \\ 0 & o.w \end{cases}$$

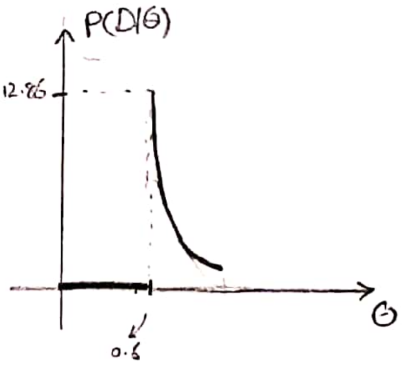
$$0 \leq \frac{1}{\theta^n} \longrightarrow MLE(\theta) = \underset{\theta}{\operatorname{argmax}} \frac{1}{\theta^n} = \underset{\theta}{\operatorname{argmax}} \log \frac{1}{\theta^n} = \underset{\theta}{\operatorname{argmax}} -n\log\theta \quad 0 \leq x_i \leq \theta \ \text{I}$$

$$\theta \downarrow \ -n\log\theta \uparrow \longrightarrow \underset{\theta}{\operatorname{argmax}} -n\log n = \text{the smallest possible } \theta \ \text{II}$$

$$\begin{cases} \text{I} & 0 \leq x_i \leq \theta \to \theta \geq \max_K x_K \to \theta \geq \max[D] \\ \text{II} & \end{cases} \Longrightarrow \theta = \max[D]$$

b) همانطور که در بخش قبل ذکر شد:

$$P(D|\theta) = \begin{cases} \frac{1}{\theta^n} & \forall_i \ 0 \leq x_i \leq \theta \to \theta \geq \max[D] \geq 0 \\ 0 & o.w \end{cases}$$



$$n=5 \to P(D|\theta) = \begin{cases} \theta^{-5} & \theta \geq 0.6 \\ 0 & o.w \end{cases}$$

'roblem 3:

$$p(x,\theta) = \prod_{i=1}^{d} \theta_i^{x_i}(1-\theta_i)^{1-x_i}$$

observations : $D = (x_1, \dots, x_N)$ where we assume x_i are i.i.d

$$\rightarrow L(\theta, D) = p(D|\theta) = \prod_{K=1}^{N} p(x_K|\theta) = \prod_{k=1}^{N}\prod_{i=1}^{d} \theta_i^{x_i^{(K)}}(1-\theta_i)^{1-x_i^{(K)}}$$

$$\ell(\theta) = \log L(\theta|D) = \circ \rightarrow \sum_{K=1}^{N}\sum_{i=1}^{d} x_i^{(K)}\log(\theta_i) + (1-x_i^{(K)})\log(1-\theta_i)$$

$$\nabla_\theta \ell(\theta) = 0 \rightarrow \forall(1 \leq j \leq d) : \frac{\partial \ell(\theta)}{\partial \theta_j} = 0 \rightarrow \sum_{K=1}^{N} x_j^{(K)} \frac{\partial(\log(\theta_j))}{\partial \theta_j} + \sum_{K=1}^{N}(1-x_j^{(K)})\frac{\partial(1-\theta_j)}{\partial \theta_j} = 0$$

$$\rightarrow \sum_{K=1}^{N} \frac{x_j^{(K)}}{\theta_j} - \sum_{K=1}^{N}\frac{1-x_j^{(K)}}{1-\theta_j} = 0$$

$$\rightarrow (1-\theta_j)\sum_{K=1}^{N} x_j^{(K)} = \theta_j\left(N - \sum_{K=1}^{N} x_j^{(K)}\right) \rightarrow \sum_{K=1}^{N} x_j^{(K)} = \theta_j N$$

$$\rightarrow \forall(1 \leq j \leq d) : \theta_j = \frac{\sum_{K=1}^{N} x_j^{(K)}}{N}$$

$$\rightarrow \hat{\theta} = \frac{\sum_{K=1}^{N} x^{(K)}}{N}$$

Problem 4:

a.

می دانیم در یادگیری میانگین توزیع نرمال داریم:

$$P(\mu|D) = \frac{1}{\sqrt{2\pi}\,\sigma_n} \exp\left(\frac{-1}{2}\left(\frac{\mu-\mu_n}{\sigma_n}\right)^2\right) \longrightarrow \begin{cases} \mu_n = \dfrac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\,\overline{x}_n + \dfrac{\sigma^2}{n\sigma_0^2+\sigma^2}\,\mu \\[2ex] \sigma_n^2 = \dfrac{\sigma_0^2\,\sigma^2}{n\sigma_0^2+\sigma^2} \end{cases}$$

$$N(\mu_n, \sigma_n^2)$$

$$n_0 = \frac{\sigma^2}{\sigma_0^2} \rightarrow \sigma^2 = n_0\sigma_0^2$$

$$\rightarrow \mu_n = \frac{n\sigma_0^2}{n\sigma_0^2+n_0\sigma_0^2} \times \frac{\sum_{k=1}^{n} x_k}{n} + \frac{n_0\sigma_0^2}{n\sigma_0^2+n_0\sigma_0^2}\,\frac{\sum_{k=-n_0+1}^{0} x_k}{n}$$

$$= \frac{\sigma_0^2(n)}{\sigma_0^2(n+n_0)} \times \frac{\sum_{k=1}^{n} x_k}{n} + \frac{\sigma_0^2(n_0)}{\sigma_0^2(n+n_0)} \times \frac{\sum_{k=-n_0+1}^{n} x_k}{n}$$

$$= \frac{\sum_{k=1}^{n} x_k}{n+n_0} + \frac{\sum_{k=-n_0+1}^{0} x_k}{n+n_0} = \frac{1}{n+n_0}\sum_{k=-n_0+1}^{n} x_k$$

Problem 4. continued

$$\sigma_n^2 = \frac{\sigma_0^2 \sigma^2}{n\sigma_0^2 + \sigma^2} = \frac{\sigma_0^2 n_0 \sigma_0^2}{n\sigma_0^2 + n_0\sigma_0^2} = \frac{n\sigma_0^4}{\sigma_0^2(n+n_0)} = \frac{n\sigma_0^2}{n+n_0} = \frac{\sigma^2}{n+n_0}$$

b.

نتیجه قسمت قبل این است که با توجه به اینکه بدست آوردیم،

$$\mu_n = \frac{1}{n+n_0} \sum_{k=-n_0+1}^{n} x_k \quad , \quad \sigma_n^2 = \frac{\sigma^2}{n+n_0}$$

در نتیجه تخمین maximum likelihood روی همه‌ی $n+n_0$ مشاهده همان نتیجه‌ی تخمین بیزی روی n نمونه دارد و درنتیجه با انتخاب μ_0 به عنوان میانگین n_0 نمونه و نیز $n_0 = \frac{\sigma^2}{\sigma_0^2}$ ، تخمین بیزی مانند MLE عمل می‌کند.

سوال ۵)

$$f(\mu) = \frac{1}{\sigma_\mu^2}\,\mu\, e^{-\frac{\mu^2}{2\sigma_\mu^2}}$$

$$\hat{\mu}_{MP} = \underset{\mu}{\arg\max}\; p(\mu|D) = \underset{\mu}{\arg\max}\; P(\mu)\,P(D|\mu)$$

$$= \underset{\mu}{\arg\max}\; \log\mu - \frac{\mu^2}{2\sigma_\mu^2} + \sum_{k=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\frac{\partial}{\partial\mu} \longrightarrow \frac{1}{\mu} - \frac{\mu}{\sigma_\mu^2} + \sum_{k=1}^{n}\left(\log\frac{1}{\sqrt{2\pi}\,\sigma} - \frac{(x-\mu)^2}{2\sigma^2}\right) = 0$$

$$\rightarrow \frac{1}{\mu} - \frac{\mu}{\sigma_\mu^2} + \frac{n\bar{x}}{\sigma^2} - n\mu = 0$$

$$\rightarrow \left(n + \frac{1}{\sigma_\mu^2}\right)\mu^2 - \frac{n\bar{x}}{\sigma^2}\mu - 1 = 0$$

$$\rightarrow \mu = \frac{\frac{n\bar{x}}{\sigma^2} \pm \sqrt{\frac{n^2\bar{x}^2}{\sigma^2} + 4\left(n + \frac{1}{\sigma_\mu^2}\right)}}{2\left(n + \frac{1}{\sigma_\mu^2}\right)}$$