STAT 543

HW 3 Solution

Spring 2016

Problem 1.

1.6.2 (a)

$$f(\underset{\sim}{x};\theta) = \prod_{i=1}^{n} \theta \cdot x_i^{\theta-1} \cdot I_{[0<x_i<1]} = \theta^n \left[\prod_{i=1}^{n} x_i\right]^{\theta-1} \cdot \prod_{i=1}^{n} I_{[0<x_i<1]}$$

$$= \prod_{i=1}^{n} I_{[0<x_i<1]} \cdot \exp\left[(\theta-1)\sum_{i=1}^{n} \ln x_i + n\ln\theta\right] \overset{\text{let}}{=} h(\underset{\sim}{x}) \exp\left[\eta(\theta) T(\underset{\sim}{x}) - B(\theta)\right]$$

where $h(\underset{\sim}{x}) = \prod_{i=1}^{n} I_{[0<x_i<1]}$, $\eta(\theta) = \theta - 1$, $T(\underset{\sim}{x}) = \sum_{i=1}^{n} \ln x_i$ and $B(\theta) = -n\ln\theta$. //

(b) For any fixed $a>0$,

$$f(\underset{\sim}{x};\theta) = \prod_{i=1}^{n} \theta a x_i^{a-1} \exp(-\theta x_i^a) \cdot I_{[x_i>1]} = \theta^n a^n \left[\prod_{i=1}^{n} x_i\right]^{a-1} \cdot \prod_{i=1}^{n} I_{[x_i>1]} \cdot \exp(-\theta x_i^a)$$

$$= a^n \left[\prod_{i=1}^{n} x_i\right]^{a-1} \cdot \prod_{i=1}^{n} I_{[x_i>1]} \exp\left[-\theta \sum_{i=1}^{n} x_i^a + n\ln\theta\right] \overset{\text{let}}{=} h(\underset{\sim}{x}) \exp\left[\eta(\theta) T(\underset{\sim}{x}) - B(\theta)\right]$$

where $h(\underset{\sim}{x}) = a^n \left[\prod_{i=1}^{n} x_i\right]^{a-1} \prod_{i=1}^{n} I_{[x_i>1]}$, $\eta(\theta) = -\theta$, $T(\underset{\sim}{x}) = \sum_{i=1}^{n} x_i^a$ and $B(\theta) = -n\ln\theta$ //

(c) For any fixed $a>0$,

$$f(\underset{\sim}{x};\theta) = \prod_{i=1}^{n} \frac{\theta a^\theta}{x_i^{\theta+1}} \cdot I_{[x_i>a]} = \theta^n a^{n\theta} \cdot \left[\prod_{i=1}^{n} x_i\right]^{-(\theta+1)} \cdot \prod_{i=1}^{n} I_{[x_i>a]}$$

$$= \prod_{i=1}^{n} I_{[x_i>a]} \cdot \exp\left[-(\theta+1)\sum_{i=1}^{n} \ln x_i + n\ln\theta + n\theta \ln a\right] \overset{\text{let}}{=} h(\underset{\sim}{x}) \exp\left[\eta(\theta) T(\underset{\sim}{x}) - B(\theta)\right]$$

where $h(\underset{\sim}{x}) = \prod_{i=1}^{n} I_{[x_i>a]}$, $\eta(\theta) = -(\theta+1)$, $T(\underset{\sim}{x}) = \sum_{i=1}^{n} \ln x_i$ and $B(\theta) = -n\ln\theta - n\theta\ln a$ //

#1.6.5 (a)

$$f(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} I_{[0<x<1]}$$

$$= \exp\left[(\alpha-1)\ln x + (\beta-1)\ln(1-x) - \ln B(\alpha,\beta)\right]\cdot I_{[0<x<1]}$$

Let $\eta = (\eta_1, \eta_2) \equiv (\alpha-1, \beta-1)$, $T(x) = (\ln x, \ln(1-x))$, $B(\eta) = \ln B(\alpha,\beta)$, and

$h(x) = I_{[0<x<1]}$.

Suppose that $\exists$ a, b, c s.t. $a\ln x + b\ln(1-x) = c$ for $\forall x \in (0,1)$.

$\Rightarrow$ $a=b=c=0$. ☺ 1) $-a\ln 2 - b\ln 2 = c$ for $x=\frac{1}{2}$

2) $-2a\ln 2 - 2b\ln 2 = c$ for $x=\frac{1}{4}$

1) − 2); $\underbrace{a\ln 2}_{>0} + b\underbrace{\ln 2}_{>0} = 0$ $\Leftrightarrow$ $a=b=0$ $\therefore$ $c=0$.

Thus, $\ln x$ and $\ln(1-x)$ are linearly indep.

$\therefore$ Beta family is two-parameter exponential family. //

(b)

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} I_{[x>0]} = \exp\left[(\alpha-1)\ln x - \frac{x}{\beta} - \ln\Gamma(\alpha) - \alpha\ln\beta\right]\cdot I_{[x>0]}$$

Let $\eta = (\eta_1, \eta_2) \equiv (\alpha-1, -\frac{1}{\beta})$, $T(x) = (\ln x, x)$, $B(\eta) = \ln\Gamma(\alpha) + \alpha\ln\beta$ and $h(x) = I_{[x>0]}$

Suppose that $\exists$ a, b, c s.t. $a\ln x + bx = c$ for all $x>0$.

$\Rightarrow$ $a=b=c=0$ ☺ 1) $0 + b = c$ for $x=1$ $\therefore$ $b=c$

2) $-a\ln 2 + \frac{b}{2} = b$ for $x=\frac{1}{2}$

3) $a\ln 2 + 2b = b$ for $x=2$

2)+3) $\frac{5}{2}b = 2b$ $\therefore b=c=0$ $\Rightarrow$ $a\ln x = 0$, $\forall x$ $\therefore a=0$.

Thus,

#1.6.6

$$f(\underset{\sim}{x};\underset{\sim}{\alpha}) = \frac{\Gamma\left(\sum_{j=1}^{r}\alpha_j\right)}{\prod_{j=1}^{r}\Gamma(\alpha_j)} \cdot \prod_{j=1}^{r} x_j^{\alpha_j - 1} \cdot I_{[0<\forall x_j<1,\ \sum_{j=1}^{r} x_j = 1]}$$

$$= \exp\left[\sum_{j=1}^{r}(\alpha_j - 1)\ln x_j + \ln\Gamma\left(\sum_{j=1}^{r}\alpha_j\right) - \sum_{j=1}^{r}\ln\Gamma(\alpha_j)\right] \cdot I_{[0<\forall x_j<1,\ \sum_{j=1}^{r} x_j = 1]}$$

Let $\underset{\sim}{\eta} = (\eta_1, \cdots, \eta_r) \equiv (\alpha_1 - 1, \cdots, \alpha_r - 1)$, $T(\underset{\sim}{x}) = (\ln x_1, \cdots, \ln x_r)$,

$B(\underset{\sim}{\alpha}) = -\ln\Gamma\left(\sum_{j=1}^{r}\alpha_j\right) + \sum_{j=1}^{r}\ln\Gamma(\alpha_j)$ and $h(\underset{\sim}{x}) = I_{[0<\forall x_j<1,\ \sum_{j=1}^{r} x_j = 1]}$

Suppose that $\exists\ c_0, c_1, c_2, \cdots, c_r$ s.t. $\sum_{j=1}^{r} c_j \ln x_j = c_0$, $\forall 0<x_j<1$, $\sum_{j=1}^{r} x_j = 1$.

$\Rightarrow c_1 = c_2 = \cdots = c_r = c_0 = 0$.

Thus, all $\ln x_j$ s $(j=1,\cdots,r)$ are linearly indep.

$\therefore$ Dirichlet family is r-parameter exponential family. //

#1.6.34 (a)

p.f) $f_p(\underline{x}) = f_p(x_1, \cdots, x_n) = f_p(x_1) \cdot f_p(x_2|x_1) \cdot f_p(x_3|x_1, x_2) \cdots f_p(x_n|x_1, x_2, \cdots, x_{n-1})$

$= f_p(x_1) \cdot f_p(x_2|x_1) \cdot f_p(x_3|x_2) \cdots f_p(x_n|x_{n-1})$ by Markov chain property

$= \frac{1}{2} \cdot p_{00}^{N_{00}} p_{01}^{N_{01}} p_{10}^{N_{10}} \cdot p_{11}^{N_{11}} = \frac{1}{2} p^{N_{00}+N_{11}} (1-p)^{N_{01}+N_{10}}$

$= \frac{1}{2} p^T (1-p)^{n-1-T}$, $x_i \in \{0,1\}$. $\forall i$

$\Rightarrow f_p(\underline{x}) = \frac{1}{2} p^T (1-p)^{n-1-T} \cdot \prod_{i=1}^{n} I_{[x_i \in \{0,1\}]}$

$= \exp\{T \ln p + (n-1-T) \ln (1-p)\} \cdot \frac{1}{2} \prod_{i=1}^{n} I_{[x_i \in \{0,1\}]}$

$= \exp\{T \ln \frac{p}{1-p} + (n-1) \ln (1-p)\} \cdot \frac{1}{2} \prod_{i=1}^{n} I_{[x_i \in \{0,1\}]}$

Let $\eta = \ln \frac{p}{1-p}$, $T(\underline{x}) = T$, $B(p) = -(n-1) \ln (1-p)$ and $h(\underline{x}) = \frac{1}{2} \prod_{i=1}^{n} I_{[x_i \in \{0,1\}]}$.

∴ This is a one-parameter exponential family. //

(b)

Note $T = N_{00} + N_{11} = \#$ of same states.

Define $I_k = \begin{cases} 1 & \text{if } x_{k+1} = x_k \\ 0 & \text{o.w.} \end{cases}$, $k = 1, 2, \cdots, n-1$.

Then $T = \sum_{k=1}^{n-1} I_k$. Since $I_k \overset{ind}{\sim} Ber(p)$, $T \sim Bin(n-1, p)$

∴ $ET = (n-1)p$. //

Problem 2. (a)

$$L(\eta) = f(x \mid \eta) = h(x) \exp[\eta T(x) - B(\eta)]$$

$$\Rightarrow \ell(\eta) \overset{\text{let}}{=} \ln L(\eta) = \ln h(x) + \eta T(x) - B(\eta)$$

$$\Rightarrow \ell'(\eta) = T(x) - B'(\eta)$$

$$\Rightarrow \ell''(\eta) = -B''(\eta) \qquad \therefore \quad I_x(\eta) = -E[\ell''(\eta)] = B''(\eta). \quad //$$

(b)

$$I_x(\eta'; \eta) = E_{\eta'}\left[\ln \frac{f(x \mid \eta')}{f(x \mid \eta)}\right] = E_{\eta'}[\ell(\eta') - \ell(\eta)]$$

$$= E_{\eta'}[\eta' T(x) - B(\eta') - \eta T(x) + B(\eta)] = (\eta' - \eta) E_{\eta'} T(x) - [B(\eta') - B(\eta)]$$

Since $E_{\eta'} \ell'(\eta') \int \frac{\partial \log f(x \mid \eta')}{\partial \eta'} f(x \mid \eta')\, dx = \int \frac{\frac{\partial f(x \mid \eta')}{\partial \eta'}}{f(x \mid \eta')} \cdot f(x \mid \eta')\, dx = \int \frac{\partial f(x \mid \eta')}{\partial \eta'} dx$

$$= \frac{\partial}{\partial \eta'} \int f(x \mid \eta')\, dx = 0,$$

$$E_{\eta'} \ell'(\eta') = E_{\eta'}[T(x) - B'(\eta')] = 0 \quad \Leftrightarrow \quad E_{\eta'} T(x) = B'(\eta').$$

$$\therefore \quad I_x(\eta'; \eta) = (\eta' - \eta) B'(\eta') - [B(\eta') - B(\eta)]. \quad //$$

Problem 3. (See the handout settings)

$\Leftarrow$: Suppose that $T(X)$ is sufficient.

By the factorization thm, $f_\theta(x) = p(T(x), \theta) \cdot h(x)$. and

$$g_\theta(t) \equiv \sum_{x \mid T(x)=t} f_\theta(x) = p(t,\theta) \sum_{x \mid T(x)=t} h(x).$$

Then,

$$I_X(\theta) = E_\theta\left[\left(\frac{f_\theta'(x)}{f_\theta(x)}\right)^2\right] = \sum_x \left[\frac{f_\theta'(x)}{f_\theta(x)}\right]^2 \cdot f_\theta(x)$$

$$= \sum_t \sum_{x \mid T(x)=t} \left[\frac{f_\theta'(x)}{f_\theta(x)}\right]^2 f_\theta(x)$$

$$= \sum_t g_\theta(t) \sum_{x \mid T(x)=t} \left[\frac{f_\theta'(x)}{f_\theta(x)}\right]^2 \cdot \frac{f_\theta(x)}{g_\theta(t)}$$

$$= \sum_t g_\theta(t) \sum_{x \mid T(x)=t} \left[\frac{p'(t,\theta)}{p(t,\theta)}\right]^2 \cdot \frac{p(t,\theta)\cdot h(x)}{g_\theta(t)}$$

$$= \sum_t g_\theta(t) \cdot \left[\frac{p'(t,\theta)}{p(t,\theta)}\right]^2 \cdot \frac{p(t,\theta)\sum_{x \mid T(x)=t} h(x)}{g_\theta(t)}$$

$$= \sum_t g_\theta(t) \cdot \left[\frac{p'(t,\theta)}{p(t,\theta)}\right]^2 \underset{\uparrow}{=} \sum_t \left[\frac{g_\theta'(t)}{g_\theta(t)}\right]^2 g_\theta(t) = I_{T(X)}(\theta)$$

//

$$\because \frac{g_\theta'(t)}{g_\theta(t)} = \frac{p'(t,\theta)}{p(t,\theta)}$$

$\Rightarrow$: Suppose that $I_X(\theta) = I_{T(X)}(\theta)$.

From the handout, we have that

$$I_X(\theta) = \sum_x \left[\frac{f_\theta'(x)}{f_\theta(x)}\right]^2 f_\theta(x) = \sum_t g_\theta(t) \cdot \sum_{x|T(x)=t} \left[\frac{f_\theta'(x)}{f_\theta(x)}\right]^2 \cdot \frac{f_\theta(x)}{g_\theta(t)}$$

by Jensen's ≠

$$\geq \sum_t g_\theta(t) \left[\sum_{x|T(x)=t} \frac{f_\theta'(x)}{f_\theta(x)} \cdot \frac{f_\theta(x)}{g_\theta(t)}\right]^2$$

$$= \sum_t \left[\frac{\sum_{x|T(x)=t} f_\theta'(x)}{g_\theta(t)}\right]^2 g_\theta(t)$$

$$= \sum_t \left[\frac{g_\theta'(t)}{g_\theta(t)}\right]^2 g_\theta(t) = I_{T(X)}(\theta).$$

Since the equality holds, $\frac{f_\theta'(x)}{f_\theta(x)}$ is a constant when $T(x)=t$, by Jensen's inequality.

$$\Rightarrow \quad \frac{f_\theta'(x)}{f_\theta(x)} = \frac{\partial \log f_\theta(x)}{\partial \theta} \overset{\text{let}}{=} K(t,\theta) \qquad \text{for any given } T(x)=t.$$

$$\Rightarrow \quad \log f_\theta(x) = \int K(t,\theta)\,d\theta + c(x) \qquad \text{for any given } T(x)=t$$

$$\therefore \quad f_\theta(x) = e^{\int K(T(x),\theta)\,d\theta} \cdot e^{c(x)} \overset{\text{let}}{=} g(T(x),\theta)\cdot h(x) \qquad \text{where } g(T(x),\theta) = e^{\int K(T(x),\theta)\,d\theta} \text{ and } h(x) = e^{c(x)}$$

By the factorization thm,

$T(X)$ is sufficient for θ. //

#3.3.14

pf) $K(q,\pi) - K(p,\pi) = \int_{\Theta} \left[E_\theta \log \frac{p_\theta(x)}{q(x)} \right] \pi(\theta)\, d\theta - \int_{\Theta} \left[E_\theta \log \frac{p_\theta(x)}{p(x)} \right] \pi(\theta)\, d\theta$

$$= \int_{\Theta} E_\theta \left[\log p_\theta(x) - \log q(x) - \log p_\theta(x) + \log p(x) \right] \pi(\theta)\, d\theta$$

$$= \int_{\Theta} E_\theta \left[\log \frac{p(x)}{q(x)} \right] \pi(\theta)\, d\theta = \int_{\Theta} \left[\int_{\mathcal{X}} \log \frac{p(x)}{q(x)} \cdot p_\theta(x) \cdot dx \right] \pi(\theta)\, d\theta$$

$$= \int_{\mathcal{X}} \int_{\Theta} \log \frac{p(x)}{q(x)}\, p_\theta(x)\, \pi(\theta)\, d\theta\, dx$$

$$= \int_{\mathcal{X}} \log \frac{p(x)}{q(x)} \left[\int_{\Theta} p_\theta(x) \pi(\theta) d\theta \right] dx = \int_{\mathcal{X}} \log \frac{p(x)}{q(x)}\, p(x)\, dx$$

by Jensen's ≠

$$= E\left[\log \frac{p(x)}{q(x)} \right] = E\left[-\log \frac{q(x)}{p(x)} \right] \geq -\log E\left[\frac{q(x)}{p(x)} \right]$$

$$= -\log \int q(x)\, dx = 0$$

where the equality holds iff $\frac{q(x)}{p(x)}$ is constant (or degenerate r.v.)

$\therefore$ $K(q,\pi) \geq K(p,\pi)$ $\Rightarrow$ $p(x)$ minimizes $K(q,\pi)$, and

the minimum is $K(p,\pi) = \int K(p_\theta, p)\, \pi(\theta)\, d\theta = \int E_\theta \log \frac{p_\theta(x)}{p(x)}\, \pi(\theta)\, d\theta.$ //

#7.7.15

(1) $N(\theta, \sigma_0^2)$.

i) $L(\theta) = f(x;\theta) = (2\pi\sigma_0^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma_0^2}(x-\theta)^2\}$

$\Rightarrow \ell(\theta) \overset{let}{=} \log L(\theta) = -\frac{1}{2}\log 2\pi\sigma_0^2 - \frac{1}{2\sigma_0^2}(x-\theta)^2$

$\Rightarrow \ell'(\theta) = \frac{1}{\sigma_0^2}(x-\theta)$

$\Rightarrow I_X(\theta) = E_\theta\left[\ell'(\theta)^2\right] = E_\theta\left[\frac{1}{\sigma_0^4}(X-\theta)^2\right] = \frac{1}{\sigma_0^4}E_\theta(X-\theta)^2 = \frac{1}{\sigma_0^2}$.

∴ Jeffrey's prior density, $\pi(\theta) \propto \sqrt{\frac{1}{\sigma_0^2}}$ ∴ $\pi(\theta) \propto 1$.

ii) Suppose that $X_1, \cdots, X_n \overset{iid}{\sim} N(\theta, \sigma_0^2)$, $\pi(\theta) \propto 1$.

$\Rightarrow \pi(\theta|\underset{\sim}{x}) \propto f(\underset{\sim}{x};\theta)\cdot\pi(\theta) \propto \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\theta)^2\right] = \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\theta)^2\right]$

$= \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \frac{1}{2\sigma_0^2}n(\bar{x}-\theta)^2\right]$

$\propto \exp\left[-\frac{n}{2\sigma_0^2}(\theta-\bar{x})^2\right]$ ← a kernel of Normal dist. family

Thus, $\theta|\underset{\sim}{x} \sim N(\bar{x}, \frac{\sigma_0^2}{n})$

∴ Bayes rule for squared loss is $E[\theta|\underset{\sim}{x}] = \bar{x}$. //

(2) $N(\mu_0, \theta)$

i) $L(\theta) = f(x;\theta) = (2\pi\theta)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\theta}(x-\mu_0)^2\right]$

$\Rightarrow l(\theta) \overset{let}{=} \log L(\theta) = -\frac{1}{2}\log 2\pi\theta - \frac{1}{2\theta}(x-\mu_0)^2$

$\Rightarrow l'(\theta) = -\frac{1}{2\theta} + \frac{1}{2}(x-\mu_0)^2 \cdot \frac{1}{\theta^2}$

$\Rightarrow l''(\theta) = \frac{1}{2\theta^2} + \frac{1}{2}(x-\mu_0)^2 \cdot \left[\frac{-2\theta}{\theta^4}\right] = \frac{1}{2\theta^2} - \frac{(x-\mu_0)^2}{\theta^3}$

$\therefore I_X(\theta) = -E_\theta[l''(\theta)] = -\frac{1}{2\theta^2} + \frac{1}{\theta^3}E_\theta(X-\mu_0)^2 = -\frac{1}{2\theta^2} + \frac{1}{\theta^2} = \frac{1}{\theta^2}$.

$\therefore$ The Jeffrey's prior density of θ, $\pi(\theta) \propto \sqrt{\frac{1}{\theta^2}} = \frac{1}{\theta}$.

ii) Suppose that $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu_0, \theta)$, $\pi(\theta) \propto \theta^{-1}$.

$\Rightarrow \pi(\theta|x) \propto f(x;\theta)\cdot\pi(\theta) \propto \theta^{-\frac{n}{2}} \exp\left[-\frac{1}{2\theta}\sum_{i=1}^n (x_i-\mu_0)^2\right] \times \theta^{-1}$

$= \theta^{-\frac{n}{2}-1} \exp\left[-\frac{1}{\theta}\left\{\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2\theta}\right\}\right] \Leftarrow$ a kernel of Inverse gamma dist. family.

Thus, $\theta|x \sim IG\left(\frac{n}{2}, \frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2}\right)$.

$\therefore$ Bayes rule for squared loss is $E[\theta|x] = \frac{\sum_{i=1}^n (x_i-\mu_0)^2}{n-2}$. //

(3) $B(n,\theta)$

i) $L(\theta) = f(x;\theta) = \binom{n}{x}\theta^{x}(1-\theta)^{n-x}$

$\Rightarrow \ l(\theta) \overset{let}{=} \log L(\theta) = \log\binom{n}{x} + x\log\theta + (n-x)\log(1-\theta)$

$\Rightarrow \ l'(\theta) = \frac{x}{\theta} - \frac{n-x}{1-\theta}$

$\Rightarrow \ l''(\theta) = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$

$\Rightarrow \ I_X(\theta) = -E_\theta[l''(\theta)] = \frac{1}{\theta^2}E_\theta X + \frac{1}{(1-\theta)^2}[n - E_\theta X] = \frac{n}{\theta} + \frac{n}{1-\theta} = \frac{n}{\theta(1-\theta)}$.

$\therefore$ Jeffrey's prior, $\pi(\theta) \propto \sqrt{\frac{1}{\theta(1-\theta)}}$

ii) Suppose that $X_1, \cdots, X_m \overset{iid}{\sim} B(n,\theta)$, $\pi(\theta) \propto \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$

$\Rightarrow \ \pi(\theta|\underline{x}) \propto f(\underline{x};\theta)\cdot\pi(\theta) \propto \theta^{\sum_{i=1}^{m} x_i}(1-\theta)^{\sum_{i=1}^{m}(n-x_i)} \cdot \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$

$\propto \theta^{\sum_1^m x_i - \frac{1}{2}}(1-\theta)^{nm - \sum_1^m x_i - \frac{1}{2}}$ $\Leftarrow$ a kernel of Beta dist. family

Thus, $\theta|\underset{\sim}{x} \sim Beta\left(\sum_{i=1}^{m} x_i + \frac{1}{2},\ nm - \sum_{i=1}^{m} x_i + \frac{1}{2}\right)$

$\therefore$ Bayes rule for squared loss is $E[\theta|\underset{\sim}{x}] = \dfrac{\sum_{i=1}^{m} x_i + \frac{1}{2}}{nm+1}$ //

Problem 5.

$$L(\lambda) = f(x;\lambda) = \frac{e^{-\lambda}\cdot\lambda^{x}}{x!}$$

$$\Rightarrow \ell(\lambda) \overset{let}{=} \log L(\lambda) = -\lambda + x\log\lambda - \log x!$$

$$\Rightarrow \ell'(\lambda) = -1 + \frac{x}{\lambda}$$

$$\Rightarrow \ell''(\lambda) = -\frac{x}{\lambda^2}$$

$$\Rightarrow I_X(\lambda) = -E_\lambda[\ell''(\lambda)] = \frac{1}{\lambda^2}\cdot E_\lambda(X) = \frac{1}{\lambda}$$

$\therefore$ The Jeffrey's prior, $\pi(\lambda) \propto \lambda^{-\frac{1}{2}}$.

Then, $\pi(\lambda|x) \propto f(x;\lambda)\cdot\pi(\lambda) \propto e^{-\lambda}\cdot\lambda^{x}\cdot\lambda^{-\frac{1}{2}} = e^{-\lambda}\cdot\lambda^{x-\frac{1}{2}}$ ← a kernel of Gamma dist. family

$\therefore \lambda|X \sim$ Gamma $(X+\frac{1}{2}, 1)$.

//

Problem 6. (a)

$$L(\underset{\sim}{\theta}) = f(x; \mu, \sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$

$$\Rightarrow \ l(\underset{\sim}{\theta}) \overset{let}{=} \log L(\underset{\sim}{\theta}) = -\frac{1}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2 .$$

$$\Rightarrow \ \frac{\partial l(\underset{\sim}{\theta})}{\partial \mu} = \frac{1}{\sigma^2}(x-\mu) \ ; \quad \frac{\partial l(\underset{\sim}{\theta})}{\partial \sigma} = -\frac{1}{2}\cdot\frac{2\sigma}{\sigma^2} + \frac{1}{2}(x-\mu)^2 \cdot \frac{2\sigma}{\sigma^4} = -\frac{1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}$$

$$\Rightarrow \ \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \mu^2} = -\frac{1}{\sigma^2} \ ; \quad \frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \mu \, \partial \sigma} = -(x-\mu)\cdot\frac{2\sigma}{\sigma^4} = -\frac{2(x-\mu)}{\sigma^3} \ ;$$

$$\frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \sigma^2} = \frac{1}{\sigma^2} - (x-\mu)^2 \cdot \frac{3\sigma^2}{\sigma^6} = \frac{1}{\sigma^2} - \frac{3(x-\mu)^2}{\sigma^4}$$

$$\Rightarrow \ I_{11}(\underset{\sim}{\theta}_0) = -E\left[\left.\frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \mu^2}\right|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}_0}\right] = \frac{1}{\sigma_0^2} ,$$

$$I_{12}(\underset{\sim}{\theta}_0) = -E\left[\left.\frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \mu \partial \sigma}\right|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}_0}\right] = \frac{2}{\sigma_0^3} E_{\underset{\sim}{\theta}_0}(x-\mu_0) = 0$$

$$I_{22}(\underset{\sim}{\theta}_0) = -E\left[\left.\frac{\partial^2 l(\underset{\sim}{\theta})}{\partial \sigma^2}\right|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}_0}\right] = \frac{1}{\sigma_0^2} - \frac{3}{\sigma_0^4}\cdot E_{\underset{\sim}{\theta}_0}(x-\mu_0)^2 = \frac{2}{\sigma_0^2} .$$

$$\therefore \ I_X(\underset{\sim}{\theta}_0) = \begin{bmatrix} \frac{1}{\sigma_0^2} & 0 \\ 0 & \frac{2}{\sigma_0^2} \end{bmatrix} \quad //$$

(b) $f(x;\mu,\sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right]$

Let $\eta_1 = \frac{\mu}{\sigma^2}$ and $\eta_2 = -\frac{1}{2\sigma^2}$ $\Rightarrow$ $\sigma^2 = -\frac{1}{2\eta_2}$ and $\mu = -\frac{\eta_1}{2\eta_2}$

$\Rightarrow L(\eta) = f(x;\eta) = \left[-\frac{\pi}{\eta_2}\right]^{-\frac{1}{2}} \exp\left[\eta_2\left(x + \frac{\eta_1}{2\eta_2}\right)^2\right]$

$\Rightarrow \ell(\eta) \stackrel{\text{let}}{=} \log L(\eta) = -\frac{1}{2}\log\left(-\frac{\pi}{\eta_2}\right) + \eta_2\left(x + \frac{\eta_1}{2\eta_2}\right)^2$

$= -\frac{1}{2}\log\pi + \frac{1}{2}\log(-\eta_2) + \eta_2 x^2 + \eta_1 x + \frac{\eta_1^2}{4\eta_2}$

$\Rightarrow \frac{\partial \ell(\eta)}{\partial \eta_1} = x + \frac{\eta_1}{2\eta_2}$; $\frac{\partial \ell(\eta)}{\partial \eta_2} = \frac{1}{2\eta_2} + x^2 - \frac{\eta_1^2}{4}\cdot\frac{1}{\eta_2^2}$

$\Rightarrow \frac{\partial^2 \ell(\eta)}{\partial \eta_1^2} = \frac{1}{2\eta_2}$; $\frac{\partial^2 \ell(\eta)}{\partial \eta_1 \partial \eta_2} = -\frac{\eta_1}{2}\cdot\frac{1}{\eta_2^2}$; $\frac{\partial^2 \ell(\eta)}{\partial \eta_2^2} = -\frac{1}{2\eta_2^2} + \frac{\eta_1^2}{4}\cdot\frac{2\eta_2}{\eta_2^4}$

$= -\frac{1}{2\eta_2^2} + \frac{\eta_1^2}{2\eta_2^3}$

$\Rightarrow I_{11}(\eta^0) = -E\left[\frac{\partial^2 \ell(\eta)}{\partial \eta_1^2}\bigg|_{\eta=\eta^0}\right] = -\frac{1}{2\eta_2^0}$

$I_{12}(\eta^0) = -E\left[\frac{\partial^2 \ell(\eta)}{\partial \eta_1 \partial \eta_2}\bigg|_{\eta=\eta^0}\right] = \frac{\eta_1^0}{2\eta_2^{0\,2}}$

$I_{22}(\eta^0) = -E\left[\frac{\partial^2 \ell(\eta)}{\partial \eta_2^2}\bigg|_{\eta=\eta^0}\right] = \frac{1}{2\eta_2^{0\,2}} - \frac{\eta_1^{0\,2}}{2\eta_2^{0\,3}}$

$$\therefore I_X(\eta^0) = \begin{bmatrix} -\frac{1}{2\eta_2^0} & \frac{\eta_1^0}{2\eta_2^{0\,2}} \\ \frac{\eta_1^0}{2\eta_2^{0\,2}} & \frac{1}{2\eta_2^{0\,2}} - \frac{\eta_1^{0\,2}}{2\eta_2^{0\,3}} \end{bmatrix}$$ //

(c) Based on the results from (a) and (b), we cannot simply put θ_0 into the matrix from a). (Note : $I_{12}(\underset{\sim}{\theta_0}) = 0$ and $I_{12}(\eta^0) = \frac{\eta_1^0}{2\eta_2^{0\,2}}$) //

Problem 7. (a)

i) $I_X(f(x|1), f(x|2)) = E_{\theta=1}\left[\log \frac{f(x|1)}{f(x|2)}\right] = \sum_x \log \frac{f(x|1)}{f(x|2)} \cdot f(x|1) = 0.2 \log 2 = 0.1386$

$I_X(f(x|2), f(x|1)) = E_{\theta=2}\left[\log \frac{f(x|2)}{f(x|1)}\right] = \sum_x \log \frac{f(x|2)}{f(x|1)} \cdot f(x|2) = \infty$. $\therefore$ Not same.

↳ In general, the measure of K-L information does not satisfy the property of symmetry.

↳ Note that $f(x|1) = 0$ but $f(x|2) = 0.1 > 0$ when $x = 1$. This definitely tells us that it came from $f(x|2)$ and not $f(x|1)$. This is "infinite" information.

(b)

$I_X(f(x|1), f(x|3)) = \sum_x \log \frac{f(x|1)}{f(x|3)} \cdot f(x|1) = 0.3348 > I_X(f(x|1), f(x|2)) = 0.1386$

$\therefore$ X is more informative for discriminating $\theta = 1$ from $\theta = 3$. //

(c)

X	1	2	3	4	5	6
$T(x) \equiv \frac{f(x|3)}{f(x|1)}$	$\frac{0}{0}$	$\frac{1}{2}$	$\frac{1}{2}$	3	$\frac{1}{2}$	$\frac{1}{2}$

where $T(1) \equiv 1$,
(arbitrary except $\{\frac{1}{2}, 3\}$)

$T(X) = t$	$\frac{1}{2}$	1	3
$\theta = 3$	0.4	0	0.6
$\theta = 1$	0.8	0	0.2

$\Rightarrow I_{T(X)}[f_T(t|1), f_T(t|3)] = \sum_t \frac{\log f_T(t|1)}{\log f_T(t|3)} \cdot f_T(t|1)$

$= 0.3348$

//