ENEE 621: Solutions Problem Set 3

Spring 2007
Narayan

1. (i)

$$g_{LMSE}(Y) \triangleq \hat{E}[\theta|Y] = \mu_\theta + \Sigma_{\theta Y}\Sigma_Y^{-1}(Y - \mu_Y).$$

$$\mu_\theta = m = \mu_Y$$

$$\Sigma_Y = \sigma^2 + var[Z] = \sigma^2 + \frac{q^2}{12}.$$

$$\begin{aligned}\Sigma_{\theta Y} &= E[(\theta - m)(Y - m)] = E[(\theta - m)(\theta + Z - m)] \\ &= E[(\theta - m)^2] + E[(\theta - m)(Z - 0)]\end{aligned}$$

Since $\theta \perp Z, E[(\theta - m)(Z - 0)] = 0$. Then

$$\Sigma_{\theta Y} = \sigma^2$$

Hence,

$$\hat{E}[\theta|Y] = m + \left(\frac{12\sigma^2}{12\sigma^2 + q^2}\right)(Y - m).$$

(ii) Assume now that θ is an unknown $\mathbb{R}$-valued constant. Then

$$f_\theta(y) = \begin{cases} \frac{1}{q}, & \theta - \frac{q}{2} \leq y \leq \theta + \frac{q}{2} \\ 0, & \text{otherwise.} \end{cases}$$

$\Rightarrow$ $\arg\max_{\theta \in \mathbb{R}} f_\theta(y) =$ any θ in $[y - \frac{q}{2}, y + \frac{q}{2}]$.

$\Rightarrow$ for a given $y \in \mathbb{R}, \exists$ a family of ML estimates given by

$$\{g_{ML}^\alpha(y) = \alpha(y - \frac{q}{2}) + (1 - \alpha)(y + \frac{q}{2}), 0 \leq \alpha \leq 1\}.$$

2. (i)

$$f_\theta(y) = \begin{cases} \frac{1}{\theta}, & 0 < y \leq \theta(< 1) \\ 0, & \text{otherwise.} \end{cases} \qquad (*)$$

Hence,

$$\begin{aligned}g_{ML}(y) &= arg \max_{0<t<1} f_t(y) \\ &= arg \max_{y \leq t < 1} f_t(y), \text{ by } (*) \\ &= y\end{aligned}$$

$\Rightarrow g_{ML}(y) = y, \qquad 0 < y < 1.$

Further, $E_\theta[g_{ML}(Y)] = E_\theta[Y] = \frac{\theta}{2} \Rightarrow g_{ML}$ is biased.

(ii)

$$\nu_y(t) = \frac{f_t(y)\nu(t)}{f(y)}, \quad 0 < y, t < 1.$$

$$= \begin{cases} \frac{1}{t} \cdot \frac{2t}{f(y)}, & 0 < y \leq t < 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{1-y}, & 0 < y \leq t < 1 \text{ (why?)} \\ 0, & \text{otherwise.} \end{cases}$$

i.e., $\nu_y(t)$ is uniform for t in $[y, 1)$.

(a) Thus,

$$g_{MAP}(y) = \text{any } t \text{ in } [y, 1)$$

Therefore, for a given y in $(0, 1)$, there exists a family of MAP estimates given by:

$$\{g^\alpha_{MAP}(y) = \alpha y + (1-\alpha), \quad 0 < \alpha \leq 1\}.$$

Further,

$$\begin{aligned} E[g^\alpha_{MAP}(Y)] &= \alpha E[Y] + (1-\alpha) \\ &= \alpha \left\{ \int_0^1 \left(\frac{1}{t} \int_0^t y dy \right) 2t dt \right\} + (1-\alpha) \\ &= \frac{\alpha}{3} + 1 - \alpha = 1 - \frac{2\alpha}{3}. \end{aligned}$$

Also, $E[\theta] = \int_0^1 t \cdot 2t \cdot dt = \frac{2}{3}$. Thus, $E[g^\alpha_{MAP}(Y)] = E[\theta]$ only when $1 - \frac{2\alpha}{3} = \frac{2}{3}$, i.e., if $\alpha = \frac{1}{2}$, so that g^α_{MAP} is unbiased only for $\alpha = \frac{1}{2}$; else, it is biased.

(b)

$$\begin{aligned} g_{MSE}(y) = E[\theta | Y = y] &= \int_y^1 t\nu_y(t) dt = \frac{1}{1-y} \int_y^1 t dt \\ &= \frac{1+y}{2}, \quad 0 < y < 1. \end{aligned}$$

Since $E[E[\theta|Y]] = E[\theta], g_{MSE}$ is (always) unbiased.

(c) From (b), $g_{LMSE}(y) \triangleq \hat{E}[\theta | Y = y] = \frac{1+y}{2}$; Therefore, g_{LMSE} is unbiased.

3.

$$f_\theta(y) \triangleq P_\theta(Y = y) = \frac{e^{-\theta}\theta^y}{y!}, \quad y = 0, 1, 2, \ldots, \quad \theta \in (0, \infty).$$

$$\ln f_\theta(y) = -\theta + y\ln\theta - \ln(y!).$$
$$\frac{\partial}{\partial\theta}\ln f_\theta(y) = -1 + \frac{y}{\theta}$$
$$\Rightarrow \theta = g_{ML}(y) \text{ satisfied } -1 + \frac{y}{g_{ML}(y)} = 0,$$

whence $g_{ML}(y) = y, y = 0, 1, 2, \ldots$

$$E_\theta[g_{ML}(Y)] = E_\theta[Y] = \theta \Rightarrow g_{ML} \text{ is unbiased.}$$

$$\Sigma_\theta(g_{ML}) = E_\theta[(g_{ML}(Y) - \theta)^2] = E_\theta[(Y-\theta)^2] = \theta.$$
$$M(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\ln f_\theta(Y)\right)^2\right] = \frac{1}{\theta^2}E_\theta[(Y-\theta)^2] = \frac{1}{\theta}.$$

Using $\theta > 0$, we have $M^{-1}(\theta) = \theta = \sum_\theta(g_{ML})$, so that g_{ML} is efficient.

4. For $-1 < \theta < 1$:

$$\begin{aligned} f_\theta(y_1, y_2) &= \frac{1}{2\pi\sqrt{1-\theta^2}}\exp\left[\frac{-1}{2(1-\theta^2)}[y_1 y_2]\begin{bmatrix} 1 & -\theta \\ -\theta & 1\end{bmatrix}\begin{bmatrix} y_1 \\ y_2\end{bmatrix}\right], \quad -\infty < y_1, y_2 < \infty \\ &= \frac{1}{2\pi\sqrt{1-\theta^2}}\exp\left[-\frac{1}{2(1-\theta^2)}(y_1^2 - 2y_1y_2\theta + y_2^2)\right] \end{aligned}$$

$\Rightarrow \ln f_\theta(y_1, y_2) = -\ln 2\pi - \frac{1}{2}\ln(1-\theta^2) - \frac{1}{2(1-\theta^2)}(y_1^2 - 2y_1y_2\theta + y_2^2)$.
Let us use the notation $g_{ML}^\theta(y_1, y_2) \triangleq \theta^*$ for convenience. If $\theta^* \in (-1, 1)$, then

$$\frac{\partial}{\partial\theta}\ln f_\theta(y_1, y_2)|_{\theta=\theta^*} = 0$$

Then,

$$\Rightarrow \theta^*(1-\theta^{*2}) + y_1y_2(1-\theta^{*2}) - \theta^*(y_1^2 - 2y_1y_2\theta^* + y_2^*) = 0$$

i.e.,

$$\theta^{*3} - y_1y_2\theta^{*2} + (y_1^2 + y_2^2 - 1)\theta^* - y_1y_2 = 0. \qquad (\$)$$

Let $f(\theta) = \theta^3 - y_1y_2\theta^2 + (y_1^2 + y_2^2 - 1)\theta - y_1y_2$. Note that $f(-1) = -(y_1+y_2)^2 \leq 0$ and $f(1) = (y_1 - y_2)^2 \geq 0$. Thus, $f(\theta)$ has at least one root in $[-1, 1]$. Next, it can be easily verified that $\theta = 1$ is a root only if $y_1 = y_2$, and that $\theta = -1$ is a root only when $y_1 = -y_2$. Hence for all (y_1, y_2) such that $|y_1| \neq |y_2|$, MLE exists (i.e., in $(-1, 1)$). The actual determination of the MLE (for $|y_1| \neq |y_2|$) is tedious. For certain values of $(y_1, y_2) \quad (|y_1| \neq |y_2|)$, the equation $f(\theta) = 0$ has a single root in $(-1, 1)$; for other (y_1, y_2), it has 3 roots in $(-1, 1)$.

5. For $y^n \triangleq (y, \ldots, y_n)$, we have

$$f_\theta(y^n) = f_{A,\Phi}(y_1, \ldots, y_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - A\sin(\frac{k\pi}{2} + \Phi))^2\right]$$

so that

$$\ln f_{A,\Phi}(y_1, \ldots, y_n) = -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - A\sin(\frac{k\pi}{2} + \Phi))^2.$$

For notational convenience, let $g_{ML}^A(y^n) \triangleq A^*, g_{ML}^\Phi(y^n) = \Phi^*$. Then

$$\frac{\partial}{\partial A} \ln f_{A,\Phi}(y^n)|_{\substack{A=A^* \\ \Phi=\Phi^*}} = 0$$

$$\Rightarrow \sum_{k=1}^{n} \left(y_k - A^* \sin\left(\frac{k\pi}{2} + \phi^*\right)\right) \sin\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

$$\Rightarrow \sum_{k=1}^{n} y_k \sin\left(\frac{k\pi}{2} + \Phi^*\right) - A^* \sum_{k=1}^{n} \sin^2\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

$$\Rightarrow \sum_{k=1}^{n} y_k \sin\left(\frac{k\pi}{2} + \Phi^*\right) = A^* \sum_{k=1}^{n} \sin^2\left(\frac{k\pi}{2} + \Phi^*\right)$$

Using the fact that n is even,

$$\sum_{k=1}^{n} \sin^2\left(\frac{k\pi}{2} + \Phi^*\right) = \frac{n}{2}$$

$$\Rightarrow \cos\Phi^* \left[y_1 - y_3 + y_5 \ldots\right] - \sin\Phi^* \cdot \left[y_2 - y_4 + y_6 \ldots\right] = \frac{A^* n}{2} \qquad (i)$$

Next,

$$\frac{\partial}{\partial \Phi} \ln f_{A,\Phi}(y^n)|_{\substack{A=A^* \\ \Phi=\Phi^*}} = 0$$

$$\Rightarrow \sum_{k=1}^{n} \left(y_k - A^* \sin\left(\frac{k\pi}{2} + \Phi^*\right)\right) (-A^*) \cos\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

$$\Rightarrow \sum_{k=1}^{n} y_k \cos\left(\frac{k\pi}{2} + \Phi^*\right) - A^* \sum_{k=1}^{n} \sin\left(\frac{k\pi}{2} + \Phi^*\right) \cos\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

Since n is even,

$$\sum_{k=1}^{n} \sin\left(\frac{k\pi}{2} + \Phi^*\right) \cos\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

So that,

$$\sum_{k=1}^{n} y_k \cos\left(\frac{k\pi}{2} + \Phi^*\right) = 0$$

$$\Rightarrow \cos\Phi^* \left[-y_2 + y_4 - y_6 + \ldots\right] - \sin\Phi^* \left[y_1 - y_3 + y_5 \ldots\right] = 0$$

$$\Rightarrow \tan\Phi^* = \frac{-y_2 + y_4 - y_6 + \ldots}{y_1 - y_3 + y_5 \ldots}$$

i.e.,

$$g_{ML}^{\Phi}(y_1, \ldots, y_n) = arc\tan\left[\frac{-y_2 + y_4 - y_6 + \ldots}{y_1 - y_3 + y_5 \ldots}\right].$$

Finally, from (i)

$$g_{ML}^{A}(y_1, \ldots, y_n) = \frac{2}{n}\left[(y_1 - y_3 + y_5 \ldots)\cos g_{ML}^{\Phi}(y_1, \ldots, y_n)\right.$$

$$\left. -(y_2 - y_4 + y_6 \ldots)\sin g_{ML}^{\Phi}(y_1, \ldots, y_n)\right].$$

6. $f_\theta(y_1, \ldots, y_n) = \prod_{k=1}^{n}\left(e^{-(y_k - \theta)} u(y_k - \theta)\right)$. Note that given $y^n \triangleq (y_1, \ldots, y_n)$, for $f_\theta(y_1, \ldots, y_n) \neq 0$, we need that $\theta \leq y_k$ for $k = 1, \ldots, n$, i.e., that $\min_{1\leq k\leq n} y_k \geq \theta$. Under this condition, i.e., for $\theta \leq \min_{1\leq k\leq n} y_k$, we have that

$$f_\theta(y_1, \ldots, y_n) = \prod_{k=1}^{n} e^{-(y_k - \theta)} = e^{n\theta} \prod_{k=1}^{n} e^{-y_k} \quad (> 0),$$

which (subject to $\theta \leq \min_{1\leq k\leq n} y_k$) is maximized by $\theta^* = \min_{1\leq k\leq n} y_k$. Thus,

$$g_{ML}(y_1, \ldots, y_n) = \min_{1\leq k\leq n} y_k.$$

7. Let $y^n \triangleq (y_1, \ldots, y_n), \quad Y^n \triangleq (Y_1, \ldots, Y_n)$.

$$f_\theta(y^n) = \frac{1}{(\sqrt{2\pi\sigma^2})^n (\prod_{k=1}^{n} y_k)} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left\{\ln\left(\frac{y_k}{\theta}\right)\right\}^2\right]$$

$$\ln f_\theta(y^n) = -\frac{n}{2}\ln(2\pi\sigma^2) - \sum_{k=1}^{n}\ln y_k - \frac{1}{2\sigma^2}\sum_{k=1}^{n}\left(\ln\frac{y_k}{\theta}\right)^2.$$

Let $\theta^* = g^\theta_{ML,n}$. Then,

$$\frac{\partial}{\partial \theta} \ln f_\theta(y^n)|_{\theta=\theta^*} = 0 = -\frac{1}{2\sigma^2} \sum_{k=1}^{n} \left[2 \left(\ln \frac{y_k}{\theta^*} \right) \frac{\theta^*}{y_k} \left(-\frac{y_k}{\theta^{*2}} \right) \right]$$

so that,

$$\sum_{k=1}^{n} \ln \left(\frac{y_k}{\theta^*} \right) = 0 \quad \Rightarrow \ln \theta^* = \frac{1}{n} \sum_{k=1}^{n} \ln y_k$$

$$\Rightarrow \theta^* = \left(\prod_{k=1}^{n} y_k \right)^{1/n}$$

i.e.,

$$g^\theta_{ML,n}(y^n) = \left(\prod_{k=1}^{n} y_k \right)^{1/n}.$$

Now, use the fact that $Y \sim$ lognormal with parameters $(\theta, \sigma^2) \Leftrightarrow Y = \exp Z$, where $Z \sim \mathcal{N}(\ln \theta, \sigma^2)$, and we get

$$E[Y] = E[\exp Z], \quad E[Y^2] = E[\exp 2Z]$$

The moments of Y can then be obtained from the moment-generating function $M_Z(\cdot)$ of Z. Specifically,

$$M_Z(u) \triangleq E[\exp uZ] = \exp(u \ln \theta + \frac{\sigma^2 u^2}{2}) \quad (\mathit{check!})$$

$$\Rightarrow E[Y] = M_Z(1) = \theta e^{\sigma^2/2}.$$

$$E[Y^2] = M_Z(2) = \theta^2 e^{2\sigma^2}.$$

Returning to our estimation problem:

$$\ln \left(g^\theta_{ML,n} (Y^n) \right) = \frac{1}{n} \sum_{k=1}^{n} \ln Y_k$$

Since $\ln Y_k \sim \mathcal{N}(\ln \theta, \sigma^2)$, $\ln(g^\theta_{ML,n}(Y^n)) \sim \mathcal{N}(\ln \theta, \frac{\sigma^2}{n})$, so that $g^\theta_{ML,n}(Y^n)$ is lognormal with parameters $(\theta, \frac{\sigma^2}{n})$. Hence,

$$E_\theta \left[g^\theta_{ML,n}(Y^n) \right] = \theta e^{\sigma^2/2n}$$

so that $g^{\theta}_{ML,n}$ is **biased**. However, since $\lim_n \theta e^{\sigma^2/2n} = \theta$, we see that the sequence of estimators $\{g^{\theta}_{ML,n}\}_{n=1}^{\infty}$ is **asymptotically unbiased**. Next,

$$\begin{aligned}
\Sigma_\theta\left(g^{\theta}_{ML,n}\right) &= E_\theta\left[\left(g^{\theta}_{ML,n}(Y^n) - \theta\right)^2\right] \\
&= E_\theta\left[\left(g^{\theta}_{ML,n}(Y^n)\right)^2\right] - 2\theta E_\theta\left[g^{\theta}_{ML,n}(Y^n)\right] + \theta^2 \\
&= \theta^2 e^{\frac{2\sigma^2}{n}} - 2\theta\cdot\theta e^{\frac{\sigma^2}{2n}} + \theta^2 \\
&= \theta^2\left[e^{\frac{2\sigma^2}{n}} - 2e^{\frac{\sigma^2}{2n}} + 1\right]
\end{aligned}$$

Thus,

$$\lim_n \Sigma_\theta\left(g^{\theta}_{ML,n}\right) = 0$$

i.e.,

$$\lim g^{\theta}_{ML,n}(Y^n) = \theta \text{ in q.m. under } P_\theta$$

$$\Rightarrow \lim_n g^{\theta}_{ML,n}(Y^n) = \theta \text{ in probability } P_\theta.$$

Hence, $g_{ML,n}$ is a (weakly) consistent estimator. Turning finally to the notion of efficiency, we see that the notion does not apply as $g^{\theta}_{ML,n}$ is **biased**. (We have defined "efficiency" in class only for unbiased estimators.) However, it is of interest to see how $\sum_\theta(g^{\theta}_{ML,n})$ differs from the appropriate Cramér-Rao lower bound (CRLB). For the problem at hand, the CRLB $= b^2_\theta(g^{\theta}_{ML,n}) + \left[1 + \frac{d}{d\theta} b_\theta(g^{\theta}_{ML,n})\right]^2 M^{(n)}(\theta)^{-1}$. To compute CRLB, first note that $M^{(n)}(\theta) = nM(\theta)$, where $M(\theta) = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln f_\theta(Y_1)\right]$. Since

$$\frac{\partial^2}{\partial\theta^2}\ln f_\theta(y) = \frac{\partial}{\partial\theta}\left[\frac{1}{\sigma^2}(\ln\frac{y}{\theta})\cdot\frac{1}{\theta}\right] = \frac{1}{\theta^2\sigma^2}\left[-1 - (\ln y - \ln\theta)\right],$$

we have

$$\begin{aligned}
M(\theta) &= -E_\theta\left[-\frac{1}{\sigma^2\theta^2} - (\ln Y - \ln\theta)\right] \\
&= \frac{1}{\sigma^2\theta^2} + E_\theta\left[\ln Y - \ln\theta\right] \\
&= \frac{1}{\sigma^2\theta^2}, \text{ because } E_\theta[\ln Y - \ln\theta] = 0
\end{aligned}$$

Next,

$$\begin{aligned}
b_\theta(g^{\theta}_{ML,n}) &= F_\theta\left[g^{\theta}_{ML,n}(Y^n)\right] - \theta = \theta e^{\sigma^2/2n} - \theta \\
&= \theta\left[e^{\sigma^2/2n} - 1\right]
\end{aligned}$$

$$\Rightarrow \frac{d}{d\theta} b_\theta\left(g^\theta_{ML,n}\right) = e^{\sigma^2/2n} - 1$$

$$\Rightarrow CRLB = \theta^2\left[\left(e^{\sigma^2/2n}-1\right)^2\right] + \left(1+e^{\sigma^2/2n}-1\right)^2 \frac{\sigma^2\theta^2}{n}$$
$$= \theta^2\left[e^{\sigma^2/n}\left(1+\sigma^{2/n}\right)+1-2e^{\sigma^2/2n}\right]$$

Hence,

$$\Sigma_\theta\left(g^\theta_{ML,n}\right) - CRLB$$
$$= \theta^2\left[e^{2\sigma^2/n} - e^{\sigma^2/n}\left(1+\frac{\sigma^2}{n}\right)\right]$$
$$= \theta^2 e^{\sigma^2/n}\left[e^{\sigma^2/n} - \left(1+\frac{\sigma^2}{n}\right)\right]$$
$$> 0,\ \text{since } e^{\sigma^2/n} > 1 + \frac{\sigma^2}{n}(\text{why?})$$

However,

$$\lim_n\left[\sum_\theta (g^\theta_{ML,n}) - CRLB\right] = \lim_n\left[\theta^2 e^{\sigma^2/n}\left[e^{\sigma^2/n} - (1+\sigma^2/n)\right]\right]$$
$$= 0.$$

8. Unknown parameters: $\theta = \{\theta_{ij}\}_{i,j=1}^M$ where $\theta_{ij} = P(Y_{k+1}=j|Y_k=i), k = 0,1,2,\ldots$. Observations:

$$(y_1,\ldots,y_n),\ \text{with } P(Y_0=y_0)=1. \qquad (*)$$

Then,

$$f_\theta(y_0,\ldots,y_n) = P(Y_0=y_0)\cdot\theta_{y_0y_1}\theta_{y_1y_2}\ldots\theta_{y_{n-1}y_n}(\text{ using the Markov property})$$
$$= \prod_{i=1}^M\prod_{j=1}^M \theta_{ij}^{n(i,j)},\ \text{by}(*)$$

where $n(i,j) = \#$ of different values of k in $\{0,\ldots,n=1\}$ such that $y_k = i$ and $y_{k+1} = j$ (i.e., the number of times the process goes directly from state i to state j). Since $\sum_{j=1}^M \theta_{ij} = 1$ for $i = 1,\ldots,M$, we have $\theta_{iM} = 1 - \sum_{j=1}^{M-1}\theta_{ij}, i = 1,\ldots,M$.

$$\Rightarrow f_\theta(y_0,y_1,\ldots,y_n) = \prod_{i=1}^M\left(1-\sum_{\ell=1}^{M-1}\theta_{i\ell}\right)^{n(i,M)}\prod_{j=1}^{M-1}\theta_{ij}^{n(i,j)}$$

$$\Rightarrow \ln f_\theta(y_0, y_1, \ldots, y_n) = \sum_{i=1}^{M}\left[n(i,M)\ln\left(1 - \sum_{\ell=1}^{M-1}\theta_{i\ell}\right) + \sum_{j=1}^{M-1} n(i,j)\ln\theta_{ij}\right]$$

For notational convenience, let $g_{ML}^{\theta_{ab}}(y_0, y_1, \ldots, y_n) \triangleq \theta^*_{ab}$. Then, for $1 \leq a \leq M, 1 \leq b \leq M-1$,

$$\frac{\partial}{\partial\theta_{ab}} \ln f_\theta(y_0, y_1, \ldots, y_n)|_{\theta_{ab}=\theta^*_{ab}} = 0 = \frac{-n(a,M)}{1 - \sum_{\ell=1}^{M-1}\theta^*_{al}} + \frac{n(a,b)}{\theta^*_{ab}}$$

whence $\theta^*_{ab} = \frac{n(a,b)}{n(a,M)}\theta^*_{aM}$. Summing both sides over b yields:

$$1 = \sum_{b=1}^{M}\frac{n(a,b)}{n(a,M)}\theta^*_{aM} = \frac{\theta^*_{aM}}{n(a,M)}\sum_{b=1}^{M} n(a,b).$$

$$\Rightarrow \theta^*_{ab} = \frac{n(a,b)}{\sum_{\ell=1}^{M} n(a,\ell)} = \frac{n(a,b)}{n(a)}, \text{ where } n(a) \triangleq \sum_{\ell=1}^{M} n(a,\ell).$$

Thus,

$$g_{ML}^{\theta_{ij}}(y_0, y_1, \ldots, y_n) = \frac{n(i,j)}{n(i)}, \quad 1 \leq i,j \leq M.$$

"Intuitively,"

$$g_{ML}^{\theta_{ij}}(y_0, \ldots, y_n) = \frac{(\#\text{ of times process leaves state } i \text{ directly for state } j)}{(\#\text{ of times it leaves state } i \text{ for state } i \text{ or any other state.})}$$

Notice that:

\# of times it leave state i for state i or any other state

= # of times it resides in state i in times $0, \ldots, n-1$

Observations: 1112122112221212

$$n(1,1) = 3; n(1,2) = 5; n(2,1) = 4; n(2,2) = 3, n(1) = 8; n(2) = 7.$$

$$\Rightarrow \theta^*_{11} = \frac{3}{8}, \theta^*_{12} = \frac{5}{8}, \theta^*_{21} = \frac{4}{7}, \theta^*_{22} = \frac{3}{7} \text{ are the ML estimates.}$$

9. (i) $\hat{E}[\theta|Y] = \mu_0 + \sum_{\theta Y}\sum_Y^{-1}(Y - \mu_Y)$. Note that $\sum_{\theta Y} = E[(Y^2 - E[Y^2])Y] = E[Y^3] - E[Y]E[Y^2]$. Since $E[Y^3] = E[Y] = 0, \sum_{\theta Y} = 0$. Hence, $\hat{E}[\theta|Y] = \mu_\theta = E[Y^2] = \frac{1}{3}$, a constant, and clearly a poor estimator.

(ii) $g_{MSE}(Y) = E[\theta|Y] = E[Y^2|Y] = Y^2$.

10. Here $\theta = (\theta_1, \theta_2), 0 \leq \theta_1 < \theta_2$. Let $y^k \triangleq (y_1, \ldots, y_k)$. Then

$$f_\theta(y^k) = \begin{cases} \left(\frac{1}{\theta_2 - \theta_1}\right)^k & \text{if } \min\{y_1, \ldots, y_k\} \geq \theta_1,\ \max\{y_1, \ldots, y_k\} \leq \theta_2 \\ 0 & \text{otherwise.} \end{cases}$$

(i)

$$\Rightarrow g^\theta_{ML,k}(Y^k) = \begin{bmatrix} g^{\theta_1}_{ML,k}(Y^k) \\ g^{\theta_2}_{ML,k}(Y^k) \end{bmatrix} = \begin{bmatrix} \min\{Y_1, \ldots, Y_k\} \\ \max\{Y_1, \ldots, Y_k\} \end{bmatrix}.$$

(ii)

$$\begin{aligned} E_\theta\left[g^{\theta_1}_{ML,k}\left(Y^k\right)\right] &= E_\theta\left[\min\{Y_1, \ldots, Y_k\}\right] \\ &= \int_0^\infty P(\min\{Y_1, \ldots, Y_k\} \geq x)dx, \\ &\qquad \text{using the fact that} \min\{Y_1, \ldots, Y_k\} \geq 0 \\ &= \int_0^\infty [P(Y_1 \geq x)]^k dx. \end{aligned}$$

Now,

$$P(Y_1 \geq x) = \begin{cases} 1, & 0 \leq x \leq \theta_1 \\ \frac{\theta_2 - x}{\theta_2 - \theta_1}, & \theta_1 \leq x \leq \theta_2 \\ 0, & x \geq \theta_2. \end{cases}$$

$$\begin{aligned} \Rightarrow E_\theta\left[g^{\theta_1}_{ML,k}\left(Y^k\right)\right] &= \int_0^{\theta_1} dx + \int_{\theta_1}^{\theta_2} \left(\frac{\theta_2 - x}{\theta_2 - \theta_1}\right)^k dx \\ &= \theta_1 + \int_{\theta_1}^{\theta_2} \left(\frac{\theta_2 - x}{\theta_2 - \theta_1}\right)^k dx. \end{aligned}$$

Since $\int_{\theta_1}^{\theta_2} \left(\frac{\theta_2 - x}{\theta_2 - \theta_1}\right)^k dx > 0$, we see that $g^{\theta_1}_{ML,k}$ is biased for $k = 1, 2, \ldots$.

(iii) $\lim_k E_\theta\left[g^{\theta_1}_{ML,k}(Y^k)\right] = \theta_1 + 0 = \theta_1$ i.e., $\{g^{\theta_1}_{ML,k}(Y^k)\}_{k=1}^\infty$ is asymptotically unbiased.

(iv) Given $\epsilon > 0$,

$$\begin{aligned} &P_\theta\left(|g^{\theta_1}_{ML,k}(Y^k) - \theta_1| > \epsilon\right) \\ &= P_\theta\left(g^{\theta_1(Y^k)}_{ML,k} > \theta_1 + \epsilon\right) + P_\theta\left(g^{\theta_1}_{ML,k}(Y^k) < \theta_1 - \epsilon\right) \end{aligned}$$

Notice that,

$$P_\theta(g^{\theta_1}_{ML,k}(Y^k) < \theta_1 - \epsilon) = 0(\text{why?})$$

So that

$$\begin{aligned}
&P\theta(|g_{ML,k}^{\theta_1}(Y^k) - \theta_1 > \epsilon) \\
&= P_\theta(\min\{Y_1, \ldots, Y_k\} > \theta_1 + \epsilon) = (P_\theta(Y_1 > \theta_1 + \epsilon))^k \\
&= \left(\frac{\theta_2 - \theta_1 - \epsilon}{\theta_2 - \theta_1}\right)^k \\
&= \left(1 - \frac{\epsilon}{\theta_2 - \theta_1}\right)^k \to 0 \text{ as } k \to \infty, \text{ as long as } \epsilon \leq \theta_2 - \theta_1
\end{aligned}$$

If $\epsilon > \theta_2 - \theta_1$, then clearly $P_\theta(Y_1 > \theta_1 + \epsilon) = 0$. Thus, for every $\epsilon > 0$, $\lim_k P_\theta \left(|g_{ML,k}^{\theta_1}(Y^k) - \theta_1| >\right.$ 0. Therefore, $\{g_{ML,k}^{\theta_1}\}_{k=1}^{\infty}$ is (weakly) consistent.

11. $\nu_y(t) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(t - y - \frac{y^2}{2}\right]$, $\infty < t < \infty$, $\Rightarrow$ conditioned on $Y = y$, θ is Gaussian with mean $y + \frac{y^2}{2}$, and variance 1, i.e.,

$$E[\theta|Y] = Y + \frac{Y^2}{2}, \quad E[(\theta - E[\theta|Y])^2] = 1. \tag{\$}$$

(i) $\hat{E}[\theta|Y] = \mu_0 + \sum_{\theta Y} \sum_Y^{-1}(Y - \mu_Y)$

$\mu_\theta = E[\theta] = E[E[\theta|Y]] = E[Y + \frac{Y^2}{2}] = 1 + \frac{2}{2} = 2$. (Verify: $E[Y] = 1, E[Y^2] = 2$)

$$\begin{aligned}
\Sigma_{\theta Y} &= E[\theta Y] - E[\theta]E[Y] = E[\theta Y] - 2 \\
&= E[E[\theta Y|Y]] - 2 = E[YE[\theta|Y]] - 2 = E[Y(Y + \frac{Y^2}{2})] - 2 \\
&= E[Y^2] + \frac{1}{2}E[Y^3] - 2 = 2 + \frac{1}{2}6 - 2 = 3 \text{ (verify } E[Y^3] = 6) \\
&\Rightarrow \hat{E}[\theta|Y] = 2 + \frac{3}{1}(Y - 1) = 3Y - 1.
\end{aligned}$$

(ii)

$$\begin{aligned}
cov[\theta - \hat{E}[\theta|Y]] &= E[(\theta - \hat{E}[\theta|Y])^2] \\
&= E[(\theta - \hat{E}[\theta|Y])\theta], \text{ by orthogonality principle} \\
&= E[\theta^2] - E[\theta(3Y - 1)] = E[\theta^2] - 3.5 + 2 = E[\theta^2] - 13 \qquad (*)
\end{aligned}$$

From (\$) above:

$$\begin{aligned}
1 &= E[(\theta - E[\theta|Y])^2] \\
&= E[(\theta - E[\theta|Y])\theta], \text{ by orthogonality principle} \\
&= E[\theta^2] - E[\theta E[\theta|Y]]
\end{aligned}$$

So that,

$$\begin{aligned}
E[\theta^2] &= 1 + E[\theta E[\theta|Y]] = 1 + E[\theta(Y + \frac{Y^2}{2})] \\
&= 1 + E[\theta Y] + \frac{1}{2}E[\theta Y^2], \text{ notice that } E[\theta Y] = 5 \\
&= 6 + \frac{1}{2}E[E[\theta Y^2|Y]] \\
&= 6 + \frac{1}{2}E[Y^2 E[Y^2 E[\theta|Y]] = 6 + \frac{1}{2}E[Y^2(Y + \frac{Y^2}{2})]
\end{aligned}$$

i.e.,

$$\begin{aligned}
E[\theta^2] &= 6 + \frac{1}{2}E[Y^3] + \frac{1}{4}E[Y^4] \\
&= 6 + \frac{1}{2}\cdot 6 + \frac{1}{4}\cdot 24 \text{ (verify : } E[Y^4] = 24.) \\
&= 15
\end{aligned}$$

$\Rightarrow$ from (*) above, $\text{cov}[\theta - \hat{E}[\theta|Y]] = 15 - 13 = 2$.

12. (i) Letting $y^n \triangleq (y_1, \ldots, y_n)$, we have

$$\begin{aligned}
f_\theta(y^n) &= \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^n y_i} \\
\Rightarrow \ln f_\theta(y^n) &= -n \ln\theta - \frac{1}{\theta}\sum_{i=1}^n y_i \\
\Rightarrow \frac{\partial}{\partial\theta}\ln f_\theta(y^n) &= -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n y_i.
\end{aligned}$$

Setting $\frac{\partial}{\partial\theta}\ln f_\theta(y^n) = 0$ at $\theta = g_{ML,n}(y^n)$, we get

$$g_{ML,n}(y^n) = \frac{1}{n}\sum_{i=1}^n y_i$$

(ii)

$$\begin{aligned}
\Sigma_\theta(g_{ML,n}) &= E_\theta[|\theta - g_{ML,n}(Y^n)|^2] \\
&= E_\theta[(\frac{1}{n}\sum_{i=1}^n Y_i - \theta)^2]
\end{aligned}$$

i.e.,

$$\begin{aligned}
\Sigma_\theta(g_{ML,n}) &= \frac{1}{n^2}E_\theta\left[\left(\sum_{i=1}^n Y_i - n\theta\right)^2\right] \\
&= \frac{1}{n^2}\text{var}_\theta\left[\sum_{i=1}^n Y_i\right] = \frac{1}{n^2}\cdot n\ \text{var}_\theta[Y_1] \\
&= \frac{\theta^2}{n}
\end{aligned}$$

Next,

$$E_\theta[g_{ML,n}(Y^n)] = E_\theta\left[\frac{1}{n}\sum_{i=1}^{n} Y_i\right] = \theta$$

$$\Rightarrow g_{ML,n} \text{ is unbiased.}$$

Compare $\sum_\theta(g_{ML,n})$ with CRLB. To this end, observe that

$$\frac{\partial^2}{\partial\theta^2}\ln f_\theta(y) = \frac{1}{\theta^2} - \frac{2y}{\theta^3}$$

so that

$$M^{(1)}(\theta) = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln f_\theta(Y)\right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \frac{1}{\theta^2}.$$

$$\Rightarrow M^{(n)}(\theta) = \frac{n}{\theta^2}, n = 1, 2, \ldots$$

Since $\sum_\theta(g_{ML,n}) = (M^{(n)}(\theta))^{-1}$, $g_{ML,n}$ is a MVUE.

13. (i) Fist note that

$$P(Y = 1) = \int_0^1 P(Y = 1|\theta = t)dt$$
$$= \int_0^1 tdt = \frac{1}{2} = P(Y = 0).$$

Next, for $0 \le t \le 1$:

$$P(\theta \le t, Y = 1) = \int_0^t P(Y = 1|\theta = \alpha)\nu(\alpha)d\alpha, \text{ where } \nu(\alpha) \text{ is the density of } \theta$$
$$= \int_0^t \alpha d\alpha = \frac{t^2}{2},$$

so that $P(\theta \le t|Y = 1) = G_1(t) = t^2$. Hence,

$$\nu_1(t) = 2t, \quad 0 \le t \le 1$$

where $\nu_1(t)$ is the conditional density of θ at t, when $Y = 1$. Similarly, for $0 \le t \le 1$:

$$P(\theta \le t, Y = 0) = \int_0^t P(Y = 0|\theta = \alpha)\nu(\alpha)d\alpha$$
$$= \int_0^t (1-\alpha)d\alpha = t - \frac{t^2}{2},$$

so that

$$G_0(t) = P(\theta \le t|Y = 0) = 2t - t^2,$$

whence

$$\nu_0(t) = 2(1-t), \quad 0 \le t \le 1.$$

where $\nu_0(t)$ is the conditional density of θ at t, given $Y = 0$. Finally,

$$g_{MSE}(y=1) = E[\theta|Y=1] = \int_0^1 t \cdot 2t \cdot dt = 2/3$$
$$g_{MSE}(y=0) = E[\theta|Y=0] = \int_0^1 t \cdot 2(1-t)dt = 1/3.$$

(ii) Let $y^n \triangleq (y_1, \ldots, y_n)$. Then

$$P(Y^n = y^n|\theta = t) = t^{\sum_{i=1}^n y_i} \cdot (1-t)^{n-\sum_{i=1}^n y_i}$$

$$\Rightarrow P(Y^n = y^n) = \int_0^1 P(Y^n = y^n|\theta = t)\nu(t)dt$$
$$= \int_0^1 t^{\sum_{i=1}^n y_i} \cdot (1-t)^{n-\sum_{i=1}^n y_i} dt$$

Also, for $0 \le t \le 1$:

$$P(\theta \le t, Y^n = y^n) = \int_0^t P(Y^n = y^n|\theta = \alpha)\nu(\alpha)d\alpha$$
$$= \int_0^t \alpha^{\sum_{i=1}^n y_i} \cdot (1-\alpha)^{n-\sum_{i=1}^n y_i} d\alpha$$

$$\Rightarrow P(\theta \le t|Y^n = y^n) = G_{y^n}(t) = \frac{\int_0^t \alpha^{\sum_{i=1}^n y_i} \cdot (1-\alpha)^{n-\sum_{i=1}^n y_i} d\alpha}{\int_0^1 \alpha^{\sum_{i=1}^n y_i} \cdot (1-\alpha)^{n-\sum_{i=1}^n y_i} d\alpha}$$

Hence,

$$\nu_{y^n}(t) = \frac{t^{\sum_{i=1}^n y_i} \cdot (1-t)^{n-\sum_{i=1} y_i}}{\int_0^1 \alpha^{\sum_{i=1}^n y_i} \cdot (1-\alpha)^{n-\sum_{i=1}^n y_i} d\alpha}, \quad 0 \le t \le 1.$$

$$\Rightarrow g_{MAP}(y^n) = \arg \max_{0 \le t \le 1} t^{\sum_{i=1}^n y_i} \cdot (1-t)^{n-\sum_{i=1}^n y_i}$$

Take log of $t^{\sum_{i=1}^n y_i} \cdot (1-t)^{n-\sum_{i=1}^n y_i}$ to get:

$$g_{MAP}(y^n) = \frac{1}{n} \sum_{i=1}^n y_i$$

(This result also follows from the fact that θ being uniform on [0,1] $\Rightarrow g_{ML} = g_{MAP}$.)

(iii) $E[\theta] = \frac{1}{2}; E[g_{MAP}(Y^n)] = E[\frac{1}{n}\sum_{i=1}^n Y_i] = \frac{1}{n}\sum_{i=1}^n E[E[Y_i|\theta]]$. Notice that

$$E[Y_i|\theta] = \theta$$

so that

$$E[g_{MAP}(Y^n)] = E[\theta]$$

Hence, $g_{MAP}(Y^n)$ is unbiased.

14. (i) Since X_1 and X_2 are i.i.d., the conditional distribution of X_1 given $(X_1 + X_2)$ is the same as that of X_2 given $(X_1 + X_2)$

$$\Rightarrow E[X_1|X_1 + X_2] = E[X_2|X_1 + X_2]$$

By adding: $2E[X_1|X_1 + X_2] = E[X_1 + X_2|X_1 + X_2] = X_1 + X_2 = Y$

$$\Rightarrow E[X_1|X_1 + X_2] = \frac{Y}{2}, \text{ using the fact that } \theta = X \text{ and } Y = X_1 + X_2$$

i.e., $g_{MSE}(Y) = \frac{Y}{2}$.

(ii) $E[|\theta - g_{MSE}(Y)|^2] = E[(X_1 - \frac{X_1+X_2}{2})^2] = E[(\frac{X_1-X_2}{2})^2] = \frac{1}{4}\cdot 2 = \frac{1}{2}$.

Problem 15

(a)

$$\begin{aligned}
\mu_\theta &= \frac{1}{\alpha};\ \mu_Y = E[\theta X + N] = E[\theta]E[X] + E[N] = \frac{1}{\alpha}. \\
\Sigma_Y &= E[Y^2] - \mu_Y^2 = E[(\theta X + N)^2] - \frac{1}{a^2} = \dots \frac{3}{\alpha^2} + 1 > 0. \\
\Sigma_{\theta Y} &= E[(\theta - \mu_\theta)(Y - \mu_y)] = E[\theta Y] - \frac{1}{\alpha^2} = E[\theta(\theta X + N)] - \frac{1}{\alpha^2}] = \frac{1}{\alpha^2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\hat{E}[\theta|Y] &= \mu_\theta + \Sigma_{\theta Y}\Sigma_Y^{-1}(Y - \mu_Y) \\
&= \frac{1}{\alpha} + \frac{1}{(3+\alpha^2)}(Y - \frac{1}{\alpha}) \\
&= \frac{1}{(3+\alpha^2)}[Y + \frac{2+\alpha^2}{\alpha}].
\end{aligned}$$

(b)

$$\begin{aligned}
E[(\theta - \hat{E}[\theta|Y])^2] &= E[(\theta - \hat{E}[\theta|Y])\theta], \quad \textit{by OP} \\
&= E[\theta^2] - E\left[\theta\left(\frac{1}{(3+\alpha^2)}\{Y + \frac{2+\alpha^2}{\alpha}\}\right)\right] = \dots \\
&= \frac{2+\alpha^2}{\alpha^2(3+\alpha^3)}.
\end{aligned}$$

Problem 16

(a) For each $\theta > 0$,

$$f_\theta(y_1, \ldots, y_m, m) = \begin{cases} \theta e^{-\theta y_1} \ldots \theta e^{-\theta y_m} e^{-\theta(T - \sum_{i=1}^m y_i)}, & \sum_{i=1}^m y_i \leq T \\ 0, & \text{else} \end{cases}$$

i.e.,

$$f_\theta(y_1, \ldots, y_m, m) = \theta^m e^{-\theta T} \mathbf{1}\left(\sum_{\mathbf{i=1}}^{\mathbf{m}} \mathbf{y_i} \leq \mathbf{T}\right),$$

So that by the Factorization Theorem, $T(y_1, \ldots, y_m, m) = m$ is a nontrival sufficient statistic.

(b)

$$\ln f_\theta(y_1, \ldots, y_m, m) = m \ln \theta - \theta T + \ln \mathbf{1}\left(\sum_{\mathbf{i=1}}^{\mathbf{m}} \mathbf{y_i} \leq \mathbf{T}\right)$$

and

$$\frac{d}{d\theta} \ln f_\theta(y_1, \ldots, y_m, m) = \frac{m}{\theta} - T,$$

whence

$$g_{ML,T}(y_1, \ldots, y_m, m) = \frac{m}{T}.$$

(c) For each $\theta > 0$, M is a Poisson rv with mean θT, so that $E_\theta[g_{ML,T}(Y_1, \ldots, Y_M)] = \frac{1}{T} E_\theta[M] = \frac{\theta T}{T} = \theta$, and so $g_{ML,T}$ is unbiased. Furthermore, by the completeness of the Poisson family of distributions of mean θT as θ ranges over $(0,\infty)$, we obtain that $T(Y_1, \ldots, Y_M, M) = M$ is a complete sufficient statistic. Hence, the required MVUE is:

$$g(Y_1, \ldots, Y_M, M) = E_\theta\left[\frac{M}{T}|M\right] = \frac{M}{T},$$

i.e., the MLE is also a MVUE.

(d) For each $\theta > 0$,

$$\begin{aligned} &E_\theta\left[\left(g_{ML,T}\left(Y_1, \ldots, Y_M, M\right) - \theta\right)^2\right] \\ &= E_\theta\left[\left(\frac{M}{T} - \theta\right)^2\right] \\ &= \frac{1}{T^2} E_\theta\left[(M - \theta T)^2\right] = \frac{1}{T^2} \text{var}_\theta[M] = \frac{\theta T}{T^2} = \frac{\theta}{T}. \end{aligned}$$

Thus, for each $\theta > 0$, it holds that $\lim_{T\uparrow\infty} E_\theta\left[(g_{ML,T}(Y_1,\ldots,Y_M,M)-\theta)^2\right] = 0$, which implies $\lim_{T\uparrow\infty} g_{ML,T}(Y_1,\ldots,Y_M,M) = \theta$ in probability P_θ, i.e., (weak) consistency of the MLE.

(e) For $T = 1$, $f_\theta(y_1,\ldots,y_m,m) = \theta^m e^{-\theta} 1\left(\sum_{i=1}^m y_i < 1\right)$, and $g(t) = \alpha e^{-\alpha t}, t \geq 0$. Hence,

$$\begin{aligned} g_{\theta|Y_1,\ldots,Y_M,M}(t|y_1,\ldots,y_m,m) &= \frac{t^m e^{-t}\mathbf{1}\left(\sum_{\mathbf{i=1}}^{\mathbf{m}} \mathbf{y_i} < \mathbf{1}\right)\alpha \mathbf{e}^{-\alpha \mathbf{t}}}{\int_{\mathbf{0}}^\infty \tau^{\mathbf{m}}\mathbf{e}^{-\tau}\mathbf{1}\left(\sum_{\mathbf{i=1}}^{\mathbf{m}} \mathbf{y_i} < \mathbf{1}\right)\alpha\mathbf{e}^{-\alpha\tau}\mathbf{d}\tau} \\ &= \frac{t^m e^{-(1+\alpha)t}}{\int_0^\infty \tau^m e^{-(1+\alpha)\tau}d\tau}, \quad \sum_{i=1}^m y_i < 1. \end{aligned}$$

Finally,

$$\begin{aligned} g_{MSE}(y_1,\ldots,y_m,m) &= E\left[\theta|Y_1 = y_1,\ldots,Y_m = y_m, M = m\right] \\ &= \frac{\int_0^\infty t^{m+1}e^{-(1+\alpha)t}dt}{\int_0^\infty \tau^m e^{-(1+\alpha)\tau}d\tau} = E[U^{m+1}]/E[U^m], \end{aligned}$$

where U is exponential with mean $\frac{1}{1+\alpha}$. Continuing,

$$g_{MSE}(y_1,\ldots,y_m,m) = \left(\frac{(m+1)!}{(1+\alpha)^{m+1}}\right)/\left(\frac{m!}{(1+\alpha)^m}\right) = \frac{m+1}{1+\alpha}.$$

For $m = 1, g_{MSE}(y_1,1) = \frac{2}{1+\alpha}$.

Problem 17

(a)

$$f_\theta(y_1,y_2) = \frac{1}{4}\exp\left[-\sum_{i=1}^2 |y_i-\theta|\right]$$

whence

$$\ln f_\theta(y_1,y_2) = -ln4 - \left[\sum_{i=1}^2 |y_i-\theta|\right].$$

Hence,

$$g_{ML}(y_1,y_2) = \arg\min_{-\infty<\theta<\infty}\left[\sum_{i=1}^2 |y_i-\theta|\right]. \tag{1}$$

In (1) observe that for $\theta \notin [\tilde{y}_1,\tilde{y}_2]$, where $(\tilde{y}_1,\tilde{y}_2)$ is a rearrangement of (y_1,y_2) such that $\tilde{y}_1 \leq \tilde{y}_2$, we have

$$\sum_{i=1}^2 |y_i-\theta| > \tilde{y}_2 - \tilde{y}_1 \ (\geq 0)$$

whereas for $\theta \in [\tilde{y}_1, \tilde{y}_2]$, we have

$$\sum_{i=1}^{2} |y_i - \theta| = \tilde{y}_2 - \tilde{y}_1.$$

Hence, $g_{ML}(y_1, y_2) =$ any value in $[\tilde{y}_1, \tilde{y}_2]$, so that **all** the maximum-likelihood estimates can be represented as:

$$g_{ML}^{(\alpha)}(y_1, y_2) = \alpha \tilde{y}_1 + (1-\alpha)\tilde{y}_2, \quad 0 \leq \alpha \leq 1. \tag{2}$$

In order to identify the unbiased estimator(s) among these, we proceed as follows. Note that

$$E_\theta \left[g_{ML}^{(\alpha)} (Y_1, Y_2) \right] = \alpha E_\theta \left[\tilde{Y}_1 \right] + (1-\alpha)\, E_\theta \left[\tilde{Y}_2 \right], \quad 0 \leq \alpha \leq 1. \tag{3}$$

Next,

$$\begin{aligned} P_\theta \left[\tilde{Y}_1 \leq y \right] &= P_\theta \left[\{Y_1 \leq y\} \bigcup \{Y_2 \leq y\} \right], \quad y \epsilon \mathbb{R} \\ &= 2F_\theta(y) - F_\theta^2(y), \end{aligned}$$

so that

$$f_\theta^{\tilde{Y}_1}(y) = 2f_\theta(y) - 2F_\theta(y) f_\theta(y), \quad y \in \mathbb{R}.$$

Also,

$$\begin{aligned} P_\theta \left[\tilde{Y}_2 \leq y \right] &= P\theta \left[\{Y_1 \leq y\} \bigcap \{Y_2 \leq y\} \right] \\ &= P_\theta \left[Y_1 \leq y \right] P_\theta \left[Y_2 \leq y \right] = F_\theta^2(y), \end{aligned}$$

so that

$$f_\theta^{\tilde{Y}_2}(y) = 2F_\theta(y) f_\theta(y), \quad y \in \mathbb{R}$$

Hence, in (3),

$$\begin{aligned} E_\theta \left[g_{ML}^{(\alpha)}(Y_1, Y_2) \right] &= \alpha \left[\int_{-\infty}^{\infty} y\{2f_\theta(y) - 2F_\theta(y) f_\theta(y)\} dy \right] \\ &\quad + (1-\alpha) \left[\int_{-\infty}^{\infty} y \cdot 2F_\theta(y) f_\theta(y) dy \right], \end{aligned}$$

from which, for $\alpha = \frac{1}{2}$, we get

$$E_\theta \left[g_{ML}^{(1/2)}(Y_1, Y_2) \right] \int_{-\infty}^{\infty} y f_\theta(y) dy = E_\theta[Y] = \theta, \quad \theta \in \mathbb{R}.$$

Thus, $g_{ML}^{(1/2)}(Y_1, Y_2) = \frac{1}{2}\left(\tilde{Y}_1 + \tilde{Y}_2\right) = \frac{1}{2}(Y_1 + Y_2)$ is the desired unbiased ML estimate.

(b) Since

$$f_t(y) = \begin{cases} \frac{1}{2}e^{t-y}, & y \geq t \\ \frac{1}{2}e^{y-t}, & y \leq t \end{cases}$$

and

$$g(t) = \begin{cases} \frac{1}{2}, & -1 \leq t \leq 1 \\ 0, & \text{else} \end{cases}$$

we get

$$f_{\theta,Y}(t,y) = \begin{cases} \frac{1}{4}e^{t-y}, & -1 \leq t \leq 1, y \geq t \\ \frac{1}{4}e^{y-t}, & -1 \leq t \leq 1, y \leq t \\ 0, & \text{else} \end{cases}$$

from which we get that

$$g_{MAP}(y) = \begin{cases} -1, & y \leq -1 \\ y, & -1 \leq y \leq 1 \\ 1, & y \geq 1. \end{cases}$$

Problem 18

(a) For each $\theta > 0$,

$$\begin{aligned} f_\theta(y^n) &= \left(\frac{1}{2\theta}\right)^n 1\left(\min_i y_i \geq -\theta\right) 1\left(\max_i y_i \leq \theta\right) \\ &= \left(\frac{1}{2\theta}\right)^n 1\left(\max_i |y_i| \leq \theta\right) \end{aligned}$$

so that

$$g_{ML,n}(y^n) = \max_{1\leq i\leq n} |y_i|, \quad y^n \in \mathbb{R}^n .$$

(b) For each $\theta > 0$,

$$g_{ML,n}(Y^n) - \theta = \max_{1\leq i\leq n} |Y_i| - \theta \leq 0 \quad P_\theta - a.s., n = 1, 2, \ldots,$$

so that

$$\sqrt{n}\left(g_{ML,n}(Y^n) - \theta\right) \leq 0 \quad P_\theta - a.s. \text{ for } n = 1, 2, \ldots,$$

and, hence, $\sqrt{n}\left(g_{ML,n}(Y^n) - \theta\right)$ cannot converge to a Gaussian rv as $n \to \infty$.

(c) From $f_1(y) = \frac{1}{2} \cdot 1\left(|y| \leq 1\right)$ and $f_2(y) = \frac{1}{4} \cdot 1(|y| \leq 2)$, together with $P[\theta = 1] = P[\theta = 2] = 1/2$. we obtain

$$f(y) = \begin{cases} \frac{3}{8}, & |y| \leq 1 \\ \frac{1}{8}, & 1 < |y| \leq 2 \\ 0, & \text{else.} \end{cases}$$

Hence,

$$P[\theta = 1|Y = 1] = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{3}{8}} = \frac{2}{3}$$

and $P[\theta = 2|Y = 1] = \frac{1}{3}$. This means that

$$P[\theta \le t|Y = 1] = \begin{cases} 0, & t < 1 \\ \frac{2}{3}, & 1 \le t < 2 \\ 1, & t \ge 2, \end{cases}$$

from which it follows that $g_{MEM}(1) = 1$.

Problem 19

(a) If $\overline{\theta}_t(Y^t) = \hat{\theta}_t(Y^t)$ P-a.s., then $\overline{\theta}_t$ must satisfy the orthogonality principle too. However, considering the affine estimator $g : \mathbb{R}^t \to \mathbb{R}$ given by $g(y^t) = y_1$, $y^t \in \mathbb{R}^t$, we see that

$$\begin{aligned} E\left[\left(\theta - \overline{\theta}_t(Y^t)\right) g(Y^t)\right] &= E\left[\left(\theta - \frac{1}{t}\sum_{\ell=1}^{t} Y_\ell\right) Y_1\right) \\ &= E\left[\left(\theta - \frac{1}{t}\sum_{\ell=1}^{t}(\theta + N_\ell)\right) Y_1\right] \\ &= E\left[\left(\frac{1}{t}\sum_{\ell=1}^{t} N_\ell\right)(\theta + N_1)\right] = \frac{1}{t}E[N_1^2] \\ &= \frac{\sigma^2}{t} > 0, \end{aligned}$$

i.e., $\overline{\theta}_t$ does not satisfy the orthogonality principle.

(b) Since $\overline{\theta}_t$ is a linear estimator of θ on the basis of Y^t, we have

$$\begin{aligned} E\left[\left(\theta - \hat{\theta}_t(Y^t)\right)^2\right] &\le E\left[\left(\theta - \overline{\theta}(Y^t)\right)^2\right] \\ &= E\left[\left(\theta - \frac{1}{t}\sum_{\ell=1}^{t}(\theta + N_\ell)\right)^2\right] \\ &= E\left[\left(\frac{1}{t}\sum_{\ell=1}^{t} N_\ell\right)^2\right] = \frac{\sigma^2}{t}. \end{aligned}$$

Hence, $\lim_t E\left[\left(\theta - \hat{\theta}_t(Y^t)\right)^2\right] = 0$, i.e., $\hat{\theta}_t\left(Y^t\right) \to \theta$ in q.m., which implies that

$$\hat{\theta}_t(Y^t) \to \theta \text{ in probability},$$

so that the sequence of estimators $\{\hat{\theta}_t\}_{t=1}^{\infty}$ is (weakly) consistent.