SOLUTION TO BIOS260 FINAL EXAM

1. (a) $X_n \to_{a.s.} X \Rightarrow X_n \to_p X \Rightarrow X_n \to_d X$.

   (b) Let U be a r.v. and $U \sim Unif(0,1)$. Define $X_{2^k+m} = I(\frac{m}{2^k} \le U < \frac{m+1}{2^k})$ for $k = 0, 1, 2, ...$ and $m = 0, 1, ..., 2^k - 1$. Then $X_n \to_p 0$ but X_n does not converge to 0 almost surely.

   (c) Let X and Y be independent random variable and follow $N(0,1)$. Define $X_n = X$. Then $X_n \to_d Y$ but X_n does not converge to Y in probability.

2. (a) The likelihood function is given by

$$\lambda^n \prod_{i=1}^n \omega_i \exp\{-\lambda \sum_{i=1}^n \omega_i X_i\} I(X_{(1)} > 0).$$

   This is a one-parameter exponential family with complete sufficient statistics

$$T_n = \sum_{i=1}^n \omega_i X_i.$$

   (b) Since $E[T_n] = n/\lambda$, the UMVUE for $1/\lambda$ is T_n/n.

   (c) The variance the UMVUE is $\sum_{i=1}^n \omega_i^2 Var(X_i)/n^2 = 1/(n\lambda^2)$. On the other hand,

$$E[\ddot{l}_\lambda] = -n/\lambda^2.$$

   Then the C-R bound for $1/\lambda$ is given by $1/(n\lambda^2)$. The UMVUE attains this bound.

   (d) The score equation is

$$\frac{n}{\lambda} - \sum_{i=1}^n \omega_i X_i = 0.$$

   Thus, the MLE for $1/\lambda$ is equal to T_n.

   (e) We derive the asymptotic distribution of $\sqrt{n}(T_n - E[T_n])/\sqrt{Var(T_n)}$, which is equal to

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\omega_i X_i - 1/\lambda)\lambda.$$

   Note $\omega_i X_i$ are i.i.d with distribution $\lambda \exp\{-\lambda x\}$. From the CLT,

$$\sqrt{n}\lambda(T_n - 1/\lambda) \to_d N(0,1).$$

3. (a) (Y_{1i}, Y_{2i}) follows a bivariate-normal distribution with mean (μ, μ) and covariance Σ equal to

$$\begin{pmatrix} 1+\sigma^2 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}.$$

   The joint log-likelihood function is

$$-n \log\left\{2\pi\sqrt{|\Sigma|}\right\} - \frac{1}{2}\sum_{i=1}^n (Y_{1i}-\mu, Y_{2i}-\mu)\Sigma^{-1}\begin{pmatrix} Y_{1i}-\mu \\ Y_{2i}-\mu \end{pmatrix} = -n\log\left\{2\pi\sqrt{\sigma^4+2\sigma^2}\right\}$$

$$-\frac{1}{2}\sum_{i=1}^n \left\{\frac{1+\sigma^2}{\sigma^4+2\sigma^2}(Y_{1i}-\mu)^2 - \frac{2}{\sigma^4+2\sigma^2}(Y_{1i}-\mu)(Y_{2i}-\mu) + \frac{1+\sigma^2}{\sigma^4+2\sigma^2}(Y_{2i}-\mu)^2\right\}.$$

(b) By direct calculation,

$$-E[\ddot{l}_\theta] = -\begin{pmatrix} -\frac{2}{\sigma^2+2} & 0 \\ 0 & * \end{pmatrix}.$$

Thus, the efficient information bound for μ is $(\sigma^2+2)/2$. The efficient influence function is given by

$$(1,0)I(\theta)^{-1}\begin{pmatrix} \frac{1}{\sigma^2+2}(Y_1-\mu)+\frac{1}{\sigma^2+2}(Y_2-\mu) \\ * \end{pmatrix} = \frac{1}{2}\left\{(Y_1-\mu)+(Y_2-\mu)\right\}.$$

(c) The joint density of (Y_{1i}, Y_{2i}, X_i) is

$$\frac{1}{2\pi\sigma^2\sqrt{2\pi}}\exp\{-(Y_{1i}-X_i)^2/2\sigma^2-(Y_{2i}-X_i)^2/2\sigma^2-(X_i-\mu)^2/2\}.$$

Thus, the complete score equation is

$$\sum_{i=1}^n (X_i-\mu)=0,$$

$$-\frac{n}{\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^n\left\{(Y_{i1}-X_i)^2+(Y_{i2}-X_i)^2\right\}=0.$$

(d) Since (X_i, Y_{i1}, Y_{i2}) follows a multivariate normal distribution with mean (μ,μ,μ) and covariance matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1+\sigma^2 & 1 \\ 1 & 1 & 1+\sigma^2 \end{pmatrix}.$$

The conditional distribution of X_i given Y_{i1} and Y_{i2} is a normal distribution with mean

$$\mu+(1,1)\begin{pmatrix} 1+\sigma^2 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}^{-1}\begin{pmatrix} Y_{i1}-\mu \\ Y_{i2}-\mu \end{pmatrix}$$

and variance

$$1-(1,1)\begin{pmatrix} 1+\sigma^2 & 1 \\ 1 & 1+\sigma^2 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(e) The conditional expectation of the complete score equations given the observed data is

$$\sum_{i=1}^n (E[X_i|Y_{i1},Y_{i2}]-\mu)=0,$$

$$-\frac{n}{\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^n\left\{Y_{i1}^2+Y_{i2}^2-2(Y_{i1}+Y_{i1})E[X_i|Y_{i1},Y_{i2}]+2E[X_i^2|Y_{i1},Y_{i2}]\right\}=0.$$

Then, in the M-step, we compute

$$\mu=\frac{1}{n}\sum_{i=1}^n E[X_i|Y_{i1},Y_{i2}]$$

and

$$\sigma^2=\frac{1}{2n}\sum_{i=1}^n\left\{Y_{i1}^2+Y_{i2}^2-2(Y_{i1}+Y_{i1})E[X_i|Y_{i1},Y_{i2}]+2E[X_i^2|Y_{i1},Y_{i2}]\right\}.$$

The two expectations $E[X_i|Y_{i1}, Y_{i2}]$ and $E[X_i^2|Y_{i1}, Y_{i2}]$ are calculated in the E-step. Specifically,

$$E[X_i|Y_{i1}, Y_{i2}] = \mu + \frac{Y_{i1} - \mu + Y_{2i} - \mu}{2 + \sigma^2},$$

$$E[X_i^2|Y_{i1}, Y_{i2}] = \left\{\mu + \frac{Y_{i1} - \mu + Y_{2i} - \mu}{2 + \sigma^2}\right\}^2 + \frac{\sigma^2}{2 + \sigma^2}.$$

As a result, at the kth iteration, the EM gives

$$\mu^{(k+1)} = \mu^{(k)} \frac{\sigma^{2(k)}}{2 + \sigma^{2(k)}} + \frac{\bar{Y}_{n1} + \bar{Y}_{n2}}{2 + \sigma^{2(k)}},$$

$$\begin{aligned} \sigma^{2(k+1)} &= \frac{1}{2n} \sum_{i=1}^{n} \Bigg[Y_{i1}^2 + Y_{i2}^2 - 2(Y_{i1} + Y_{i1})(\mu^{(k)} + \frac{Y_{i1} + Y_{2i} - 2\mu^{(k)}}{2 + \sigma^{2(k)}}) \\ &\quad +2\left\{\mu^{(k)} - \frac{Y_{i1} + Y_{2i} - 2\mu^{(k)}}{2 + \sigma^{2(k)}}\right\}^2 + \frac{2\sigma^{2(k)}}{2 + \sigma^{2(k)}} \Bigg]. \end{aligned}$$