Stat 581,2004: Homework 5: Outline solutions

1. (a) We find that $\prod_1^n x_i$ is maximized s.t. $\sum_1^n x_i = K$ when all the x_i are equal, so $G_n^n = \prod_1^n x_i \le \overline{x_n}^n$, so $G_n \le A_n$. Repeating with $1/x_i$ in place of x_i we get $G_n^{-1} \le H_n^{-1}$. So for all real positive values of x_i, $H_n \le G_n \le A_n$, and hence for non-negative random variables X_i with probability 1.

(b) If $\mathrm{E}|X_i| < \infty$, $\mathrm{E}(1/|X_i|) < \infty$, and $\mathrm{E}|\log(X_i)| < \infty$, and $\mathrm{E}(X_i) = a$, $\mathrm{E}(1/X_i) = h^{-1}$, and $\mathrm{E}(\log(X_i)) = \log(g)$ then, by SLLN,
$(A_n, \log(G_n), 1/H_n) \to_{a.s.} (a, \log(g), h^{-1})$, Then , by continuous mapping theorem $(A_n, G_n, H_n) \to_{a.s.} (a, g, h)$. (Hence also converges in probability.)

(c) If also $\mathrm{E}X_i^2 < \infty$, $\mathrm{E}((1/X_i)^2) < \infty$, and $\mathrm{E}(\log(X_i)^2 < \infty$, then by the multivariate CLT

$$\sqrt{n}(A_n - a,\ \log(G_n) - \log(g),\ H_n^{-1} - h^{-1}) \to_d Z \sim N_3(0, \Sigma)$$

where $\Sigma_{11} = \mathrm{var}(X_i)$, $\Sigma_{12} = \mathrm{Cov}(X_i, \log(X_i))$, $\Sigma_{13} = \mathrm{Cov}(X_i, 1/X_i)$, etc.
Now use δ-method, with $v(x, y, z) = (x, \exp(y), 1/z)$,
so $\bigtriangledown v = diag(1, \exp(y), -z^{-2}) = diag(1, g, -h^{-2})$ at the mean point, and

$$\sqrt{n}(A_n - a,\ G_n) - g,\ H_n - h) \to_d (\bigtriangledown v) \cdot Z \sim N_3(0, (\bigtriangledown v)\Sigma(\bigtriangledown v)')$$

2. $Z_i = \min(X_i, Y_i)$, $\delta_i = I(X_i \le Y_i)$, $f_{X,Y}(x, y) = f(x;\theta)g(y)$.
(a) On $Z = z$, $\delta = 1$, pdf $= f(z;\theta)g(y)$ on $y > z$) integrates to $f(z;\theta)(1 - G(z))$.
On $Z = z, \delta = 0$, pdf $= f(x;\theta)g(z)$ on $x > z$) integrates to $g(z)(1 - F(z;\theta))$.
Together: $h(z, \delta; \theta) = (f(z;\theta)(1 - G(z)))^\delta (g(z)(1 - F(z;\theta)))^{1-\delta}$.

(b)

$$\begin{aligned}
\log h &= \delta \log f + \delta \log(1 - G) + (1 - \delta) \log g + (1 - \delta) \log(1 - F) \\
&= -\delta(\log\theta + Z/\theta) + \ldots - (1 - \delta)Z/\theta \\
\frac{\partial \log h}{\partial \theta} &= -\delta/\theta + Z/\theta^2 \\
\frac{\partial^2 \log h}{\partial \theta^2} &= \delta/\theta^2 - 2Z/\theta^3
\end{aligned}$$

Now $\mathrm{E}(\delta) = P(X \le Y) = \int F(y)g(y)dy = \int(1 - \exp(y/\theta))g(y)dy$.
Also $\mathrm{E}(\partial \log h/\partial\theta) = 0$ gives $\mathrm{E}(Z) = \theta\mathrm{E}(\delta)$.
So $I_1(\theta) = \theta^{-3}(2\theta\mathrm{E}(\delta) - \theta\mathrm{E}(\delta)) = \mathrm{E}(\delta)/\theta^2$. For an n-sample, $I_n(\theta) = nI_1(\theta) = n\mathrm{E}\delta/\theta^2$.
CRLB $= 1/I_n(\theta) = \theta^2/n(\int(1 - e^{-y/\theta})g(y)dy)$
Note this makes sense: if observe all the X_i, info is n/θ^2 and we expect to observe a proportion $\mathrm{E}(\delta)$ of them.

3. (a) $p_\theta(x) = \theta f_1(x) + (1 - \theta)f_2(x)$.

$$\begin{aligned}
\frac{\partial \log p}{\partial \theta} &= \frac{f_1(x) - f_2(x)}{p_\theta(x)} \text{ so} \\
I_1(\theta) &= \mathrm{E}\left(\left(\frac{\partial \log p}{\partial \theta}\right)^2\right) = \int \frac{(f_1(x) - f_2(x))^2}{p_\theta(x)} dx \\
I_n(\theta) &= nI_1(\theta), \quad \mathrm{var}(T) \ge 1/I_n(\theta), \text{ if } \mathrm{E}(T) = \theta.
\end{aligned}$$

(b) Let S_i be the support of f_i: $S_1 \cap S_2 \;=\; \Phi$.

$$\begin{aligned} I_1(\theta) &= \int_{S_1} \frac{f_1^2}{\theta f_1} dx \;+\; \int_{S_2} \frac{f_2^2}{(1-\theta) f_2} dx \\ &= \theta^{-1} \;+\; (1-\theta)^{-1} \;=\; 1/\theta(1-\theta) \end{aligned}$$

(c) Regarding this as a missing data problem (see JAW 3.10, as per hint)

$$\begin{aligned} Y_i &= (X_i, \delta_i),\ \ \delta_i = 1, 0 \text{ as } X_i \sim f_1, f_2 \\ q_\theta(x,\delta) &= (f_1(x)\theta)^\delta (f_2(x)(1-\theta))^{1-\delta} \\ \log q &= \text{const} \;+\; \delta \log\theta \;+\; (1-\delta)\log(1-\theta) \\ \frac{\partial \log q}{\partial \theta} &= \frac{\delta}{\theta} - \frac{1-\delta}{1-\theta},\quad \frac{\partial^2 \log q}{\partial\theta^2} \;=\; -\frac{\delta}{\theta^2} - \frac{1-\delta}{(1-\theta)^2},\quad I_1^{(q)}(\theta) \;=\; 1/\theta(1-\theta) \\ \text{Now } \frac{\partial \log p}{\partial\theta} &= \frac{f_1(x) - f_2(x)}{p_\theta(x)} \;=\; \mathrm{E}\left(\frac{\partial \log q}{\partial\theta} \;\mid\; X = x\right) \text{ since } \mathrm{E}(\delta | X = x) = \frac{\theta f_1(x)}{p_\theta(x)} \\ \text{So } \mathrm{E}\left(\left(\frac{\partial \log p}{\partial\theta}\right)^2\right) &= \mathrm{E}\left(\mathrm{E}\left(\frac{\partial \log q}{\partial\theta} \;\mid\; X\right)^2\right) \\ &\le \mathrm{E}\left(\mathrm{E}\left(\frac{\partial \log q}{\partial\theta}\right)^2 \;\mid\; X\right) \;=\; \mathrm{E}\left(\left(\frac{\partial \log q}{\partial\theta}\right)^2\right) \end{aligned}$$

4. (a) $K \;\sim\; \mathcal{P}(\mu B)$, so $\ell(\mu) = \text{const} - \mu B + K(\log(\mu) + \log(B))$, so $\widehat{\mu} \;=\; K/B$ which has variance $B^{-2}(\mu B) \;=\; \mu/B$.

(b) Presence and observation of plants are independent, so this is still a (now non-homogeneous) Poisson process. Expected number of plants observed is $2\int_0^\infty \exp(-\lambda x)(\mu L)dx \;=\; 2\mu L/\lambda$. Further, $\Pr(x|\text{observed}) \;\propto\; \Pr(\text{obs }|x)\Pr(x) \;\propto\; \exp(-\lambda x)$, so normalizing the density we have $f_X(x;\lambda) \;=\; \lambda\exp(-\lambda x)$.

(c) The likelihood for (μ, λ) is

$$\begin{aligned} \Pr(K = k; 2\mu L/\lambda) f_\lambda(x_1, \ldots, x_k | K = k) &\propto \left(\frac{mu}{\lambda}\right)^k \exp\left(-\frac{2\mu L}{\lambda}\right) \prod_1^k \lambda\exp(-\lambda x_i) \\ &= \mu^k \exp(-\frac{2\mu L}{\lambda} - \lambda\sum_{i=1}^k x_i) \end{aligned}$$

(d) $\ell \;=\; k\log(\mu) \;-\; 2\mu L/\lambda \;-\; \lambda\sum_i x_i$ and solving the likelihood equations gives $\lambda \;=\; (\overline{x})^{-1}$ and $\mu = k\lambda/2L$. Taking negative of expected second derivatives gives the information matrix (inverse of asymptotic variance-covariance matrix) as $I_{\mu\mu} \;=\; 2L/\lambda\mu$, $I_{\mu\lambda} \;=\; -2L/\lambda^2$, and $I_{\lambda\lambda} \;=\; 4\mu L/\lambda^3$.

(e) If $B = 2L/\lambda$, $I^{\mu\mu} \;=\; B/\mu$, confirming (a).
Then $I_{\mu\mu\cdot\lambda} \;=\; (B/\mu) - (B/\lambda)^2(\lambda^2/2\mu B) \;=\; B/(2\mu)$, so exactly one half the information about μ is lost by the need to estimate λ.