Stat 581,2004: Homework 7: Outline solutions

1. (a) $L(\theta,\psi) = (2\psi)^{-n} I(\theta-\psi \leq X_{(1)} \leq X_{(n)} \leq \theta+\psi)$
The min.sufft statistic is $(X_{(1)}, X_{(n)})$.
Likelihood is maximized by $\theta-\psi = X_{(1)},\ \theta+\psi = X_{(n)}$. So $\widehat{\psi} = \frac{1}{2}(X_{(n)} - X_{(1)})$, $\widehat{(\theta)} = \frac{1}{2}(X_{(n)} + X_{(1)})$.

(b) $\ell(\theta) \ = \ \sum_{i=1}^{n} |X_{(i)} - \theta|$
The min.sufft statistic is $(X_{(1)}, \ldots, X_{(n)})$.
The log-likelihood function is piecewise linear, with gradient $n - 2r$ on $X_{(r)} < \theta < X_{(r+1)}$.
Thus $\widehat{\theta}$ is the median of the X_i.

2. (a) $F_n^{-1}(p) \ \rightarrow_{a.s.} \ F^{-1}(p)$, so
$T_n \ \rightarrow_{a.s.} \ \frac{1}{2}(F^{-1}(p) + F^{-1}(1-p)) \ = \ F^{-1}(1/2) = \theta$ by symmetry.

(b) Let $f_1 = f(F^{-1}(p))$, $f_2 = f(F^{-1}(1-p))$. By symmetry, $f_2 = f_1 = f$ say. Now

$$\sqrt{n}\left(\left(\begin{array}{c} F_n^{-1}(p) \\ F_n^{-1}(1-p) \end{array}\right) - \left(\begin{array}{c} F^{-1}(p) \\ F^{-1}(1-p) \end{array}\right)\right) \ \rightarrow_d \ N_2\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} p(1-p)/f_1^2, & p^2/f_1 f_2 \\ p^2/f_1 f_2, & p(1-p)/f_2^2 \end{array}\right)\right)$$

$$\text{so } \sqrt{n}(\frac{1}{2}(F_n^{-1}(p) + F_n^{-1}(1-p)) \ - \ \theta) \ \ \rightarrow_d \ \ N(0, 2p/(4f^2))$$

so to minimize asymptotic variance we maximize $f(F^{-1}(p))/\sqrt{p}$.

(c) For $X_i \sim U(\theta-\psi, \theta+\psi)$, $f = 1/2\psi$ on its range. So choose p as small as possible, i.e. $p = 1/n$. Then $T_n^{(1/n)} \ = \ \frac{1}{2}(X_{(1)} + X_{(n)}) \ = \ \widehat{\theta}$.

For $X_i \sim DE(\theta)$: for $p < \frac{1}{2}$, $F^{-1}(p) = \theta - \log(2p)$, and $f(F^{-1}(p)) \ = \ \frac{1}{2}\exp(\log(2p)) \ = \ p$. So maximize $p/\sqrt{p} = \sqrt{p}$, i.e. $p = \frac{1}{2}$. Then $T_n^{(1/2)} \ = \ \text{median} \ = \ \widehat{\theta}$.

3. (a) Let $W_i \equiv (X_i, \delta_i)$ where $\delta_i \equiv I(X_i = 1)$. Then the W_i are i.i.d, $\mathrm{E}(W_i) = (\theta, \theta e^{-\theta})$, and

$$\mathrm{var}(W) \ = \ V(\theta) \ = \ \left(\begin{array}{cc} \theta, & \theta e^{-\theta} - \theta^2 e^{-\theta} \\ \theta e^{-\theta} - \theta^2 e^{-\theta} & \theta e^{-\theta}(1 - \theta e^{-\theta}) \end{array}\right).$$

Hence, by the multivariate CLT $n^{\frac{1}{2}}(\overline{W_n} - (\theta, \theta e^{-\theta})) \ \ \rightarrow_d \ \ N_2(0, V(\theta))$

(b) Now $(\widehat{P_{1,n}}, Z_n) \ = \ g(\overline{W_n})$ where $g(u,v) = (ue^{-u}, v)$, so

$$\bigtriangledown g(u,v) \ = \ \left(\begin{array}{cc} (1-u)e^{-u}, & 0 \\ 0 & 1 \end{array}\right).$$

Hence, using the delta-method:
$n^{\frac{1}{2}}((\widehat{P_{1,n}}, Z_n) - (\theta e^{-\theta}, \theta e^{-\theta})) \ \ \rightarrow_d N_2(0, (\bigtriangledown g)V(\theta)(\bigtriangledown g)^t)$ where $\bigtriangledown g$ is evaulated at $(\theta, \theta e^{-\theta})$ and hence

$$(\bigtriangledown g)V(\theta)(\bigtriangledown g)^t \ = \ \left(\begin{array}{cc} \theta(1-\theta)^2 e^{-2\theta} & \theta(1-\theta)^2 e^{-2\theta} \\ \theta(1-\theta)^2 e^{-2\theta} & \theta e^{-\theta}(1-\theta e^{-\theta}) \end{array}\right)$$

(c) Since these are both regular asymptotically Gaussian estimators, the A.R.E. is the ratio of the asymptotic variances, and

$$ARE(Z_n, \widehat{P_{1,n}}) \ = \ \frac{\theta(1-\theta)^2 e^{-2\theta}}{\theta e^{-\theta}(1-\theta e^{-\theta})} \ = \ \frac{(1-\theta)^2 e^{-\theta}}{(1-\theta e^{-\theta})} \ < \ 1$$

Thus the MLE (not surprisingly) is more efficient, for all $\theta > 0$.

4. Likelihood equation for ρ reduces to

$$g(\rho) \;=\; \rho^3 \;-\; S_{xy}\rho^2 \;+\; (S_{xx}+S_{yy}-1)\rho \;-\; \S_{xy} \;=\; 0$$

where $S_{xy} = \sum_i x_i y_i$, $S_{xx} = \sum_i x_i^2$, $S_{yy} \;=\; \sum y_i^2$.
Now $g(-1) < 0$, $g(1) > 0$ (justify these), so there is at least one real root in $-1 < \rho < 1$.

Now

$$\begin{aligned} g(\rho) - g(\rho_0) &= (\rho-\rho_0)H(\rho,\rho_0) \text{ where} \\ H(\rho,\rho_0) &= (\rho^2 + \rho\rho_0 + \rho_0^2 \;-\; S_{xy}(\rho+\rho_0) + S_{xx} + S_{yy} - 1) \\ &\to_p \rho^2 + 1 \;>\; 0 \text{ for all } \rho \end{aligned}$$

So for n large enough $\Pr(H(\rho,\rho_0) \;>\; 0) \;\to\; 1$, and then $\Pr(g(\rho) < 0,\ \forall\rho < \rho_0) \;\to\; 1$ and $\Pr(g(\rho) > 0,\ \forall\rho > \rho_0) \;\to\; 1$, and the root is unique.
(Cite approp. theorems, and unif. cgce on compact interval.)

5.(a)

$$\begin{aligned} -2\ell(\theta,\psi,r) &= (x-\theta)^t(x-\theta) \;+\; (y-\psi)^t)(y-\psi) \;+\; (z - r\theta - (1-r)\psi)^t(z - r\theta - (1-r)\psi) \\ \ell_\theta = 0 &\Rightarrow (\theta - x) \;=\; r(z - r\theta - (1-r)\psi) \\ \ell_\psi = 0 &\Rightarrow (\psi - y) \;=\; (1-r)(z - r\theta - (1-r)\psi) \\ \ell_r = 0 &\Rightarrow (z - r\theta - (1-r)\psi)^t(\psi - \theta) \;=\; 0 \end{aligned}$$

Solving the first two of these for θ and ψ in terms of r gives $\theta - X \;=\; rW(r)/h(r)$ and $\psi - Y \;=\; (1-r)W(r)/h(r)$ as required.

(b) Substituting for θ and ψ into $(z - r\theta - (1-r)\psi)$ we get $(z - r\theta - (1-r)\psi) \;=\; W(r)/h(r)$ as required, and then

$$-2\ell(\widehat{\theta},\widehat{\psi},r) \;=\; (h(r))^{-2}(r^2W^tW \;+\; (1-r)^2W^tW \;+\; W^tW) \;=\; W^tW/h(r)$$

Hence we need to maximize $S(r) = \; -W(r)^tW(r)/h(r)$.

(c) Now

$$\frac{W(r)^tW(r)}{h(r)} \;=\; \frac{r^2d_{xy}^2 - r(d_{yz}^2 + d_{xy}^2 - d_{xz}^2) + d_{yz}^2}{1 + r^2 + (1-r)^2}$$

and differentiating, the result follows after slightly messy arithmetic.

(d)

$$g(r) \;=\; r^2(d_{xz}^2 - d_{yz}^2) \;+\; 2r(d_{yz}^2 - d_{xy}^2) \;+\; (d_{xy}^2 - d_{xz}^2)$$

The quadratic equation $g(r) = 0$ has two real roots (which we could write down). If $d_{xz} > d_{yz}$, $g(r)$ is positive for large positive and large negative r, so the smaller root of $g(r) = 0$ is a maximum, the other a minimum. The log-likelihood increases, then decreases, and finally increases again as $r \to \infty$.
(i) $g(0) < 0,\ g(1) < 0$, maximum is in $r < 0$ and MLE within $0 \le r \le 1$ is $\widehat{r} = 0$.
(ii) $g(0) > 0,\ g(1) < 0$; MLE is the root in $0 < r < 1$.
(iii) $g(0) < 0,\ g(1) > 0$. The root in $0 < r < 1$ is a MINIMUM of the likelihood. The maximum is in $r < 0$ and the MLE is $\widehat{r} = 0$.

(This rather lengthy example is to show that in some cases you have to be careful. You cannot assume the root within the parameter space $0 \le r \le 1$ is necessarily going to be a maximum. Note the answers do make sense: for a geometric interpretation, see the cited paper.)