Statistics 581, Problem Set 9 Solutions

Wellner; 12/30/2005

1. (a) Lehmann and Casella, Problem 2.13, page 501.
(b) Let $R_n(\theta) \equiv nE_\theta(T_n - \theta)^2$ where T_n is the Hodges superefficient estimator as in Example 3.3.1 (so $T_n = \delta_n$ of Example 2.5, Lehmann and Casella pages 440 - 443). Show that $R_n(n^{-1/4}) \to \infty$ as $n \to \infty$.

Solution: (a) (a') First recall that (with $\delta_n = T_n$) since $\sqrt{n}(\overline{X} - \theta) \stackrel{d}{=} Z \sim N(0,1)$ we can write

$$\begin{aligned}
\sqrt{n}(T_n - \theta) &= \sqrt{n}(\overline{X}_n 1_{[|\overline{X}_n| > n^{-1/4}]} + a\overline{X}_n 1_{[|\overline{X}_n| \le n^{-1/4}]} - \theta) \\
&\stackrel{d}{=} Z 1_{[|Z+\theta\sqrt{n}| > n^{1/4}]} + [aZ + \sqrt{n}\theta(a-1)] 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\
&= Z + [(a-1)Z + (a-1)\sqrt{n}\theta] 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\
&= Z - (1-a)[Z + \sqrt{n}\theta] 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]}.
\end{aligned}$$

Thus

$$\begin{aligned}
b_n(\theta) &= E_\theta(T_n) - \theta \\
&= n^{-1/2}\left\{EZ - (1-a)E[Z + \sqrt{n}\theta] 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]}\right\} \\
&= -\frac{1-a}{\sqrt{n}} E[Z + \sqrt{n}\theta] 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\
&= -\frac{1-a}{\sqrt{n}} \int_{-n^{1/4}}^{n^{1/4}} x\phi(x - \sqrt{n}\theta)dx
\end{aligned}$$

since $Z + \theta\sqrt{n} \sim N(\theta\sqrt{n}, 1)$.
(b') Differentiating the result in (a') gives

$$\begin{aligned}
b_n'(\theta) &= -\frac{1-a}{\sqrt{n}} \int_{-n^{1/4}}^{n^{1/4}} x\phi'(x - \sqrt{n}\theta)(-\sqrt{n})\ dx \\
&= -(1-a) \int_{-n^{1/4}}^{n^{1/4}} x(x - \sqrt{n}\theta)\phi(x - \sqrt{n}\theta)\ dx \qquad \text{since} \quad \phi'(x) = -x\phi(x) \\
&\to 0 \qquad \text{if} \quad \theta \ne 0
\end{aligned}$$

by the dominated convergence theorem since $x(x - \sqrt{n}\theta)\phi(x - \sqrt{n}\theta)1_{[-n^{1/4}, n^{1/4}]}(x) \to 0$ for each fixed x and is dominated by the integrable function $x^2\phi(x)$ (for $n \ge (2/|\theta|)^{1/4}$). When $\theta = 0$

$$b_n'(0) = -(1-a)\int_{-n^{1/4}}^{n^{1/4}} x^2\phi(x)\ dx \to -(1-a)\int_{-\infty}^{\infty} x^2\phi(x)dx = -(1-a).$$

(c') The information inequality implies that

$$Var_\theta(\sqrt{n}(T_n - \theta)) \ge \frac{(b_n'(\theta)+1)^2}{I(\theta)} = (b_n'(\theta)+1)^2$$

since $I(\theta) = 1$. At the point $\theta = 0$ the right side converges to a^2, while the limit inferior of the left side is the variance of the limiting distribution at $\theta = 0$, namely a^2. Thus there is no contradiction with the information inequality.

(b) Using the distributional identity in (a) yields

$$\begin{aligned} R_n(\theta) &= 1 + (1-a)^2 E(Z + \sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}| \le n^{1/4}]} \\ &\quad - 2(1-a)E\{Z(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}\} \\ &= 1 + \{(1-a)^2 - 2(1-a)\}E(Z+\sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad + 2(1-a)\sqrt{n}\theta E\{(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}\} \\ &= 1 - (1-a^2)E(Z+\sqrt{n}\theta)^2 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad + 2(1-a)\sqrt{n}\theta E\{(Z+\sqrt{n}\theta)1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}\} \end{aligned}$$

(This confirms the first identity in Lehmann's example 4.7, page 442.) Squaring out the expectation in the second term and writing the third term as the sum of two terms yields, with $\alpha_n \equiv n^{1/4} - \sqrt{n}\theta$, $\beta_n \equiv -n^{1/4} - \sqrt{n}\theta$,

$$\begin{aligned} R_n(\theta) &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad - 2(1-a^2)\sqrt{n}\theta E Z 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad - (1-a^2)n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2(1-a)n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2(1-a)\sqrt{n}\theta E\{Z1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}\} \\ &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad +(1-a)^2 n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad -2a(1-a)\sqrt{n}\theta E(Z1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}) \end{aligned}$$

where

$$\begin{aligned} E(Z1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]}) &= \int_{\alpha_n}^{\beta_n} z\phi(z)dz \\ &= -\int_{\alpha_n}^{\beta_n} \phi'(z)dz \qquad \text{since} \quad \phi'(z) = -z\phi(z) \\ &= -(\phi(\beta_n) - \phi(\alpha_n)). \end{aligned}$$

Thus it follows that

$$\begin{aligned} R_n(\theta) &= 1 - (1-a^2)EZ^2 1_{[|Z+\theta\sqrt{n}|\le n^{1/4}]} \\ &\quad + (1-a)^2 n\theta^2(\Phi(\beta_n) - \Phi(\alpha_n)) \\ &\quad + 2a(1-a)\sqrt{n}\theta(\phi(\beta_n) - \phi(\alpha_n)). \end{aligned}$$

(This confirms the second identity in Lehmann's problem 4.7, page 442.) Now we take $\theta = \theta_n = n^{-1/4}$, and note that $\alpha_n = -2n^{1/4}$, $\beta_n = 0$. Since the expectation of in the second term in the last display is bounded below by zero and above by 1 we find that

$$\begin{aligned} R_n(n^{-1/4}) &\geq a^2 + (1-a)^2 n^{1/2}(1/2 - \Phi(-2n^{1/4}) \\ &\qquad +2a(1-a)n^{1/4}(\phi(0) - \phi(-2n^{1/4})) \\ &\to a^2 + \infty + \infty = \infty \end{aligned}$$

since $\Phi(-2n^{1/4}) \to 0$ and $\phi(-2n^{1/4}) \to 0$.

2. (a) Lehmann and Casella, Problem 6.6, page 142.
   (b) Consider the somewhat more general family

   $$\mathcal{P} = \{(1-\epsilon)\phi(x-\xi) + (\epsilon/\tau)\phi((x-\mu)/\tau)) : \ \theta = (\epsilon, \xi, \mu, \tau) \in [0,1] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+\}$$

   where ϕ is the standard normal density. Is the information matrix for $(\epsilon, \xi, \mu, \tau)$ in this family always nonsingular? [Hint: what is the relationship of the model $\mathcal{P}$ in (b) to the model in Lehmann and Casella's problem 5.21, page 139?]
   (c) Suppose that $(X, \Delta) \in \mathbb{R} \times \{0, 1\}$ has distribution given by $p_\theta(x|\Delta = 0) = \phi(x)$, $p_\theta(x|\Delta = 1) = \phi((x-\mu)/\tau)/\tau$, and $\Delta \sim$ Bernoulli(ϵ). Show that the marginal distributions of X are in the family $\mathcal{P}$ of part (a). What is the relationship of the information for ϵ based on observation of (X, δ) to the information for ϵ based on observation of X alone?

   Solution: (a) I would prefer to write $\theta = (\xi, \epsilon, \tau)$. The first thing to note about the model

   $$\mathcal{P} = \{P_\theta : \ (dP_\theta/d\lambda)(x) = p(x;\theta) = p(x)\}$$

   is that it is a location family: $p(x; \xi, \epsilon, \tau) = p_0(x - \xi; \epsilon, \tau)$ where

   $$p_0(x; \epsilon, \tau) = (1-\epsilon)\phi(x) + (\epsilon/\tau)\phi(x/\tau).$$

   Moreover, all the distributions in the family $\mathcal{P}_0 = \{P_{0,\epsilon,\tau} : \epsilon \in [0,1], \tau > 0\}$ are *symmetric about* 0. Here are plots of these densities for $\tau = 3$ and $\epsilon = 0, .2, .4, .6, .8, 1$; the second plot is of the densities p_0 for $\epsilon = .2$ and $\tau = 1.5, 4, 8, 16, 32, 64$.

Figure 1: Densities $p(x)$ for $\xi = 0$, $\tau = 3$, $\epsilon = 0, .2, .4, .6, .8, 1.0$

Figure 2: Densities $p(x)$ for $\xi = 0$, $\epsilon = .2$, $\tau = 1.5, 4, 8, 16, 32, 64$

Since $\phi'(x) = -x\phi(x)$, the score functions for ϵ, ξ, and τ are given by

$$
\begin{aligned}
\dot{l}_\xi(x) &= \frac{1}{p(x)}\left\{(x-\xi)(1-\epsilon)\phi(x-\xi) + \frac{x-\xi}{\tau^2}\frac{\epsilon}{\tau}\phi\left(\frac{x-\xi}{\tau}\right)\right\}, \\
\dot{l}_\epsilon(x) &= \frac{1}{p(x)}\left\{\frac{1}{\tau}\phi\left(\frac{x-\xi}{\tau}\right) - \phi(x-\xi)\right\}, \\
\dot{l}_\tau(x) &= \frac{1}{p(x)}\frac{\epsilon}{\tau^2}\phi\left(\frac{x-\xi}{\tau}\right)\left\{\left(\frac{x-\xi}{\tau}\right)^2 - 1\right\}.
\end{aligned}
$$

Thus with $\theta \equiv (\xi, \epsilon, \tau)$ and $\underline{\dot{l}}_\theta$, the information matrix $I(\xi, \epsilon, \tau) = I(\theta)$ is given by

$$
\begin{aligned}
I(\theta) &= E_\theta\{\underline{\dot{l}}_\theta(X)\underline{\dot{l}}_\theta(X)^T\} \\
&= \left(E_\theta(\dot{l}_i(X)\dot{l}_j(X))\right).
\end{aligned}
$$

Hence it becomes clear that all of the elements of $I(\epsilon, \xi, \tau)$ are constant functions of ξ; hence it suffices to compute the information matrix for $\xi = 0$. Thus we take $\xi = 0$ in the rest of the argument. Now note that the scores for ϵ and τ are even functions of x: $\dot{l}_\epsilon(-x) = \dot{l}_\epsilon(x)$ and similarly for $\dot{l}_\tau$. On the other hand, the score function for ξ is an odd function of x: $\dot{l}_\xi(-x) = -\dot{l}_\xi(x)$. It follows that

$$
E_\theta\{\dot{l}_\xi(X)\dot{l}_\epsilon(X)\} = 0, \qquad \text{and} \qquad E_\theta\{\dot{l}_\xi(X)\dot{l}_\tau(X)\} = 0.
$$

Thus the only non-zero entry off the diagonal is $I_{23}(\theta) = I_{\epsilon,\tau}(\theta)$, and the information matrix is given by

$$
I(\theta) = \begin{pmatrix} I_{11}(\theta) & 0 & 0 \\ 0 & I_{22}(\theta) & I_{23}(\theta) \\ 0 & I_{32}(\theta) & I_{33}(\theta) \end{pmatrix}
$$

where

$$I_{11}(\theta) = \int \frac{[x(1-\epsilon)\phi(x) + (x/\tau^2)(\epsilon/\tau)\phi(x/\tau)]^2}{p_{(0,\epsilon,\tau)}(x)}\, dx = I_{11}(0,\epsilon,\tau)$$

$$I_{22}(\theta) = \int \frac{[(1/\tau)\phi(x/\tau) - \phi(x)]^2}{p_{(0,\epsilon,\tau)}(x)}\, dx = I_{22}(0,\epsilon,\tau)$$

$$I_{33}(\theta) = \int \frac{(\epsilon/\tau^2)^2\phi(x/\tau)^2\{(x/\tau)^2-1\}^2}{p_{(0,\epsilon,\tau)}(x)}\, dx = I_{33}(0,\epsilon,\tau)$$

$$I_{23}(\theta) = I_{32}(\theta) = \int \frac{[(x/\tau)\phi(x\tau) - \phi(x)][(\epsilon/\tau^2)\phi(x/\tau)\{(x/\tau)^2-1\}}{p_{(0,\epsilon,\tau)}(x)}\, dx = I_{23}(0,\epsilon,\tau).$$

I do not know any "closed form" results for $I_{11}(\theta)$, $I_{22}(\theta)$, $I_{33}(\theta)$, or $I_{23}(\theta)$, but it is not hard to compute them as functions of $\theta = (\epsilon, \xi, \tau)$ especially since they depend only on (ϵ, τ). See Figures 3-5 below.

Figure 3: Information for ϵ as a function of τ for $\epsilon = .1, .2, .3, \ldots, .9$

Figure 4: Information for ξ as a function of τ for $\epsilon = .1, .2, .3, \ldots, .9$

Figure 5: Information for τ as a function of τ for $\epsilon = .1, .2, .3, \ldots, .9$

(b) Note that the model of problem 5.21, page 139, is the submodel of $\mathcal{P}$ with $\epsilon = \epsilon_0 = 1/2$, $\tau = 1$, $\xi = \theta$, $\mu = -\theta$:

$$\mathcal{P}_0 = \{p_{(\epsilon_0, \theta, -\theta, \tau=\tau_0)} \in \mathcal{P} : \theta \in \mathbb{R}\}$$

with $\epsilon_0 = 1/2$, $\tau_0 = 1$. For this submodel we calculate

$$\dot{\mathbf{l}}_\theta(x) = \frac{(1/2)\phi(x-\theta)(x-\theta) + (1/2)\phi(x+\theta)(-x-\theta)^2}{p_{(}x;\theta)},$$

and therefore,

$$
\begin{aligned}
I(\theta) &= \int \frac{[(1/2)\phi(x-\theta)(x-\theta)+(1/2)\phi(x+\theta)(-x-\theta)^2]^2}{p(x;\theta)}dx \\
&= \int \frac{[(1/2)x\phi(x)-(1/2)x\phi(x)]^2}{\phi(x)}dx \qquad \text{at} \;\; \theta = 0 \\
&= 0
\end{aligned}
$$

since the numerator inside the integral vanishes.

Figure 6: Densities in problem 5.21

Figure 7: Information in problem 5.21 as a function of θ

3. (a) Exercise 2.1.6, page 10, chapter 2 notes; i.e. show that $d_{TV}(P,Q) = 1 - \int p \wedge q \, d\mu$.
(b) Exercise 2.1.7, page 10, chapter 2 notes; i.e. show that

$$H^2(P,Q) \le d_{TV}(P,Q) \le H(P,Q)\{1 + \rho(P,Q)\}^{1/2} \le \sqrt{2}H(P,Q).$$

Solution: (a) From the proof of proposition 1.13, chapter 2 notes, page 9, we see that

$$\begin{aligned} d_{TV}(P,Q) &= \frac{1}{2}\int |p-q| d\mu = \int_{[p\ge q]} (p-q)\ d\mu = \int_{[p\ge q]} p d\mu - \int_{[p\ge q]} p \wedge q\ d\mu \\ &= \int_{[p\ge q]} p d\mu + \int_{[p<q]} p d\mu - \int_{[p\ge q]} p \wedge q\ d\mu - \int_{[p<q]} p\ d\mu \\ &= \int p d\mu - \int_{[p\ge q]} p \wedge q\ d\mu - \int_{[p<q]} p \wedge q\ d\mu \\ &= 1 - \int p \wedge q\ d\mu \equiv 1 - \eta(P,Q). \end{aligned}$$

Alternatively, use the identity $|a-b| = a + b - 2(a \wedge b)$ for all $a, b \in \mathbb{R}$ to deduce that

$$|p(x) - q(x)| = p(x) + q(x) - 2p(x) \wedge q(x)$$

for each fixed x, and hence

$$\begin{aligned} d_{TV}(P,Q) &= \frac{1}{2}\int |p-q| d\mu = \frac{1}{2}\left(\int p d\mu + \int q d\mu - 2\int p \wedge q d\mu\right) \\ &= 1 - \int p \wedge q d\mu \equiv 1 - \eta(P,Q). \end{aligned}$$

(b) To see the first inequality, note that $H^2(P,Q) = 1 - \rho(P,Q)$ where

$$\rho(P,Q) = \int \sqrt{pq}\ d\mu \ge \int p \wedge q\ d\mu \equiv \eta(P,Q)$$

since $\sqrt{p(x)q(x)} \ge p(x) \wedge q(x)$ for all x. Thus we have

$$H^2(P,Q) \;=\; 1 - \rho(P,Q) \le 1 - \eta(P,Q) = d_T V(P,Q).$$

For the second inequality, write $|p-q| = |(\sqrt{p} - \sqrt{q})(\sqrt{p} + \sqrt{q})|$ and then apply the Cauchy-Schwarz inequality: thus

$$\begin{aligned} 2d_{TV}(P,Q) &= \int |p-q| d\mu = \int |(\sqrt{p} - \sqrt{q})(\sqrt{p} + \sqrt{q})| d\mu \\ &\le \left(\int |\sqrt{p} - \sqrt{q}|^2 d\mu\right)^{1/2} \left(\int |\sqrt{p} + \sqrt{q}|^2 d\mu\right)^{1/2} \\ &= \sqrt{2}H(P,Q)\left\{\int (p + 2\sqrt{pq} + q) d\mu\right\}^{1/2} \\ &= \sqrt{2}H(P,Q)\left\{2 + 2\rho(P,Q)\right\}^{1/2} \\ &= 2H(P,Q)\left\{1 + \rho(P,Q)\right\}^{1/2}, \end{aligned}$$

and this yields the claimed inequality. The third inequality is easy since $\rho(P,Q) \le 1$ by Cauchy-Schwarz again.

4. (a) Lehmann and Casella, problem 6.3.1, page 501.
   (b) Lehmann and Casella, problem 6.3.2, page 501.
   (c) Lehmann and Casella, problem 6.3.4, page 501.

5. (a) Lehmann and Casella, problem 6.3.1, page 501: Let X have the binomial distribution $Bin(n,p)$, $0 \le p \le 1$. Determine the MLE of p:
   (i) by the usual calculus method determining the maximum of a function.
   (ii) by showing that $p^x q^{n-x} \le (x/n)^x[(n-x)/n]^{n-x}$.
   (b) Lehmann and Casella, problem 6.3.2, page 501: In the preceding problem, show that the MLE does not exist when p is restricted to $0 < p < 1$ and when $X = 0$ or $X = 1$.
   (c) Lehmann and Casella, problem 6.3.4, page 501: suppose that $X_1, \ldots, X_n$ are i.i.d. as $N(\xi, 1)$ with $\xi > 0$. Show that the MLE is $\overline{X}_n$ when $\overline{X}_n > 0$ and does not exist when $\overline{X}_n \le 0$.

   Solution: (a)(i) Since $\log P_p(X = x) = x \log p + (n-x)\log(1-p)$, we have $l(p|X) = X \log p + (n-X)\log(1-p)$; differentiating this with respect to p yields

$$\dot{l}(p|X) = \frac{X}{p} - \frac{n-X}{1-p} = \frac{X(1-p)-(n-X)p}{p(1-p)}$$

   and this equals 0 if $p = \hat{p} \equiv X/n$. Since the second derivative is

$$\ddot{l}(p|X) = -\frac{X}{p^2} - \frac{n-X}{(1-p)^2} < 0$$

   it follows that $\hat{p} = X/n$ is the MLE of $p \in [0,1]$.
   (a)(ii) Since $(\prod_{i=1}^n y_i)^{1/n} \le n^{-1}(y_1 + \cdots + y_n)$ for any numbers $y_i \ge 0$, it follows, with $y_i \equiv np/X$ for $i = 1, \ldots, X$, and $y_i \equiv nq/(n-X)$, $i = X+1, \ldots, n$, that

$$\left\{ \left(\frac{np}{X}\right)^X \left(\frac{nq}{n-X}\right)^{n-X} \right\}^{1/n} \le n^{-1}\left\{ X\frac{np}{X} + (n-X)\frac{nq}{n-X} \right\} = 1\,,$$

   or, equivalently,

$$p^X(1-p)^{n-X} \le \left(\frac{X}{n}\right)^X \left(\frac{n-X}{n}\right)^{n-X},$$

   with equality if and only if $p = X/n \equiv \hat{p}$. Thus $\hat{p} = X/n$ is the MLE of $p \in [0,1]$.
   (b) When the closed interval $[0,1]$ is replaced by the open interval $(0,1)$, then the MLE exists if $0 < X < n$ and is $\hat{p} = X/n \in (0,1)$ in this case. If $X = 0$, then the log-likelihood equals $n\log(1-p)$, so $\sup_{p\in(0,1)} l(p) = 0$, but this supremum is not achieved (in the set $(0,1)$). Thus the MLE does not exist in this case. Similarly, if $X = n$, the the log-likelihood equals $n \log p$, so $\sup_{p\in(0,1)} l(p) = 0$, but this supremum

is not achieved (in the set $(0,1)$).
(c) The log-likelihood is

$$\begin{aligned} l_n(\theta) &= -\frac{1}{2}\sum_{i=1}^{n}(X_i-\xi)^2-(n/2)\log(2\pi) \\ &= n\xi\overline{X}_n - n\xi^2/2-(n/2)\log(2\pi) \\ &= -(n/2)(\xi-\overline{X}_n)^2+(n/2)\overline{X}_n^2-(n/2)\log(2\pi). \end{aligned}$$

When $\overline{X}_n > 0$ this is maximized by $\xi = \hat{\xi} = \overline{X}_n$.
When $\overline{X}_n \le 0$ this is maximized over $\xi \ge 0$ by $\xi = \hat{\xi} = 0$. Since the maximization in the problem as stated is over the open set $\xi > 0$, the supremum is not attained.

6. Consider the Weibull family of example 3.2.5: $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ with $\Theta \subset R^{+2}$ given by the (Lebesgue) densities

$$p_\theta(x) = \frac{\beta}{\alpha}(\frac{x}{\alpha})^{\beta-1}\exp(-(\frac{x}{\alpha})^\beta)1_{[0,\infty)}(x)$$

where $\theta \equiv (\alpha,\beta) \in (0,\infty)\times(0,\infty) \subset R^2$. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with density function p_θ.
A. If $X \sim P_\theta \in \mathcal{P}$, show that the distributions of $\log X$ form a location and scale family from a Gumbel (extreme value) density on R.
B. Use the result of A to construct method of moments estimators or quantile based estimators $\overline{\theta}_n$ of $\theta = (\alpha,\beta)$.
C. Show that the method of moments or quantile estimators $\overline{\theta}_n$ of θ are asymptotically normal, and find the asymptotic distribution; i.e. show that

$$\sqrt{n}(\overline{\theta}_n - \theta) \to_d N_2(0,\Sigma) \qquad \text{for some} \qquad \Sigma.$$

[We will use these estimators as "starting points" approximate (or one-step) maximum likelihood estimators in problem set 10.]

Solution: A. Recall that $Y \equiv (X/\alpha)^\beta \sim \exp(1)$, and that $W \equiv -\log(Y) \sim$ Gumbel:

$$P(W \le w) = P(-\log(Y) \le w) = P(Y \ge e^{-w}) = \exp(-e^{-w}).$$

Thus it follows that

$$W = -\log(Y) = \beta\{-\log(X) + \log(\alpha)\},$$

or equivalently that

$$T \equiv -\log(X) = \frac{1}{\beta}W - \log(\alpha).$$

Thus the distributions of $T \equiv -\log(X)$ form a location - scale family of the Gumbel (extreme value) distribution with d.f. $\exp(-\exp(-x))$.

B. Now $T = -\log X$ has

$$E(T) = \frac{\gamma}{\beta} - \log\alpha, \qquad Var(T) = \frac{1}{\beta^2}\frac{\pi^2}{6}$$

where $\gamma = .577...$ is Euler's constant (don't confuse this with the γ above!). Since $\overline{T} = -1.7864...$ and $S_T = 1.4853...$, moment estimators of (α, β) based on (8) are given by

$$\overline{\beta}_n \equiv \frac{\pi}{\sqrt{6}}\frac{1}{S_T} = .8639,$$

$$\overline{\alpha} = \exp(-\overline{T} + \frac{\gamma}{\overline{\beta}}) = 11.6407$$

for the given data.

C. Asymptotic normality of $(\overline{\alpha}_n, \overline{\beta}_n)$ follows from joint asymptotic normality of $(\overline{T}_n, S_T^2)$ and the delta method: by the multivariate CLT and Slutsky's theorem

$$\begin{pmatrix} \sqrt{n}(\overline{T} - ET)/\sigma \\ \sqrt{n}(S_T^2 - \sigma_T^2)/(\sqrt{2}\sigma_T^2) \end{pmatrix} \rightarrow_d \underline{Z} \sim N_2(0, \Sigma).$$

Then since $(\overline{\alpha}, \overline{\beta}) = g(\overline{T}, S_T^2)$ and $(\alpha, \beta) = g(E_\theta T, Var_\theta(T))$ where $g \equiv (g_1, g_2) : R^2 \rightarrow R^2$ is defined by

$$g_1(x, y) = \exp(\frac{\gamma\sqrt{6}}{\pi}\sqrt{y} - x),$$

$$g_2(x, y) = \frac{\pi/\sqrt{6}}{\sqrt{y}},$$

it follows by the delta method with $\tilde{\underline{Z}} \equiv (Z_1, \sqrt{2}\sigma_T^2 Z_2)$ that

$$\sqrt{n}\left((\overline{\alpha}_n, \overline{\beta}_n)^T - (\alpha, \beta)^T\right) \rightarrow_d \nabla g \tilde{\underline{Z}}$$

where

$$\nabla g \equiv \nabla g(E_\theta T, \mathrm{Var}_\theta T) = \begin{pmatrix} -\alpha & (3\gamma/\pi^2)\alpha\beta \\ 0 & -3\beta^3/\pi^2 \end{pmatrix}.$$