Statistics 581, Problem Set 9 Solutions

Wellner; 12/07/2001

1. (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) Lehamnn 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 $\overline{X}$ when $\overline{X} > 0$ and does not exist when $\overline{X} \le 0$.
(d) Lehmann and Casella, problem 6.3.18, page 502. [**Note:** It seems to me that 3.15(b) should be 3.15(c) since $C(0,a)$ is a *scale family*).] In problem 3.15(c) with f the Cauchy density $C(0,a)$, the likelihood equation has a unique root $\hat{a}$ and $\sqrt{n}(\hat{a} - a) \rightarrow_d N(0, 2a^2)$.

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 for ξ is

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

for $\xi > 0$. When $\overline{X} > 0$, this is maximized by $\hat{\xi} = \overline{X}$, but when $\overline{X} \le 0$, the supremum of $l_n(\xi)$ over $\xi > 0$ is not achieved (in the set $\xi > 0$). If we change the parameter space to $\xi \ge 0$, then the supremum is achieved, and we find that the MLE is given by $\hat{\xi} = \overline{X}1_{[\overline{X}\ge 0]}$.
(d) First consider problem 3.15(c): Here the log-likelihood for the scale family for a density f is

$$l_n(a) = n\log a + \sum_{i=1}^{n}\log f(aX_i), \qquad a > 0\,.$$

Hence the score equation is

$$\begin{aligned} 0 = \dot{l}_n(a) &= \frac{n}{a} + \sum_{i=1}^{n}\frac{f'(aX_i)}{f(aX_i)}X_i \\ &= \frac{1}{a}\left\{n - \sum_{i=1}^{n}g(aX_i)\right\} \equiv \frac{1}{a}\left\{n - h_n(a)\right\} \end{aligned}$$

where $g(x) \equiv -xf'(x)/f(x)$. If $xf'(x)/f(x)$ is strictly decreasing in x, then $g(x)$ is strictly increasing in x, and hence $h_n(a)$ is strictly increasing in a. Hence there is at most one value of a satisfying $h_n(a) = n$. If a solution exists, it is unique.
For the particular case of a Cauchy density f, it is easy to compute

$$g(x) = \frac{2x^2}{1+x^2}$$

which is strictly increasing (from 0 at $x = 0$ to 2 at $x = \infty$). Moreover, in this case the likelihood equation becomes

$$h_n(a) = \sum_{i=1}^{n}\frac{2a^2X_i^2}{1+a^2X_i^2} = n\,.$$

But the left side converges to 0 as $a \downarrow 0$, and converges to $2n$ as $a \uparrow \infty$. Since it is monotone increasing and continuous, there is a unique solution $\hat{a}_n$. All the hypotheses of Theorem 1.5 hold in this case, and

$$I_{scale}(f) = \int_{-\infty}^{\infty}\left\{1 + x\frac{f'(x)}{f(x)}\right\}^2 f(x)dx = \frac{1}{2}\,.$$

Thus it follows that

$$\sqrt{n}(\hat{a}_n - a) \to_d N(0, 2a^2)\,.$$

2. Suppose that $(Y|Z) \sim \text{Poisson}(\lambda e^{\gamma Z})$, and $Z \sim Bernoulli(\eta)$, and $\theta = (\lambda, \gamma, \eta)$. Let $X = (Y, Z)$, and suppose that we observe $X_1, \ldots, X_n$ i.i.d. as X.
   (a) Find the score equations for estimation of θ.
   (b) Give conditions on the data $X_1, \ldots, X_n = (Y_1, Z_1), \ldots, (Y_n, Z_n)$ guaranteeing that the score equations have a unique solution which maximizes the likelihood. Call the resulting estimators $\widehat{\theta}_n = (\widehat{\lambda}_n, \widehat{\gamma}_n, \widehat{\eta}_n)$.
   (c) What does theorem 4.1.5 (Chapter 4, page 4), say about the asymptotic distribution of $\sqrt{n}(\widehat{\theta} - \theta_0)$ when the distribution of the data is given by P_{θ_0}.
   (d) Suppose that $\theta_1 \neq \theta_0$ is the "true" value of the parameter θ, and we consider the likelihood ratio $L_n(\theta_1)/L_n(\theta_0)$ where $L_n(\theta) \equiv \prod_{i=1}^n p_\theta(X_i)$. Show that $n^{-1}\log(L_n(\theta_1)/L_n(\theta_0) \rightarrow_p$ some constant, and identify the constant explicitly in terms of θ_1, θ_0.

   Solution: (a) From the solution of problem 1, problem set #8,

$$\begin{aligned}
\dot{l}_\lambda(y, z) &= \frac{y}{\lambda} - e^{\gamma z} = \frac{1}{\lambda}(y - \lambda e^{\gamma z}), \\
\dot{l}_\gamma(y, z) &= yz - \lambda e^{\gamma z} z = z(y - \lambda e^{\gamma z}), \\
\dot{l}_\eta(y, z) &= \frac{z}{\eta} - \frac{1-z}{1-\eta}.
\end{aligned}$$

   Thus the score equations for $\theta = (\lambda, \gamma, \eta)$ are

$$\begin{aligned}
0 &= \sum_{i=1}^n \dot{l}_\lambda(Y_i, Z_i) = \frac{1}{\lambda}\sum_{i=1}^n (Y_i - \lambda e^{\gamma Z_i}), \\
0 &= \sum_{i=1}^n \dot{l}_\gamma(Y_i, Z_i) = \sum_{i=1}^n Z_i(Y_i - \lambda e^{\gamma Z_i}), \\
0 &= \sum_{i=1}^n \dot{l}_\eta(Y_i, Z_i) = \sum_{i=1}^n \left\{\frac{Z_i}{\eta} - \frac{1-Z_i}{1-\eta}\right\} = (\eta(1-\eta))^{-1}\sum_{i=1}^n (Z_i - \eta).
\end{aligned}$$

   (b) The third equation always has the unique solution $\hat{\eta} = \overline{Z}_n$. It is clear that the score equation for λ has only the solution $\lambda = 0$ if all $Y_i = 0$, and in this case there is clearly no unique solution for γ. So as least one Y_i must be non-zero in order to get a unique solution. If all the Z_i's are equal to 1, then the two equations agree, and it is clear that all we can estimate is the product λe^γ. Similarly, if all the Z_i's are equal to 0, then the score equation for γ is trivially satisfied (for any γ), and the score equation for λ gives just an estimator of λ. For the first and second equations, we compute

$$\begin{aligned}
\ddot{l}_{n,\lambda\lambda} &= -\frac{1}{\lambda^2}\sum_{i=1}^n Y_i, \\
\ddot{l}_{n,\lambda\gamma} &= -\sum_{i=1}^n Z_i e^{\gamma Z_i} = \ddot{l}_{n,\gamma\lambda}, \\
\ddot{l}_{n,\gamma\gamma} &= -\lambda\sum_{i=1}^n Z_i^2 e^{\gamma Z_i}.
\end{aligned}$$

Thus the matrix of second partial derivatives (the Hessian) fails to be negative definite if $\sum_1^n Y_i = 0$ or if

$$\left(\lambda \sum_{i=1}^n Z_i^2 e^{\gamma Z_i}\right)\left(\lambda^{-2}\sum_{i=1}^n Y_i\right) \le \left(\sum_{i=1}^n Z_i e^{\gamma Z_i}\right)^2 .$$

Because the Z_i's are Bernoulli,

$$\sum_{i=1}^n Z_i^2 e^{\gamma Z_i} = \sum_{i=1}^n Z_i e^{\gamma Z_i} = e^{\gamma}\sum_{i=1}^n Z_i .$$

Thus the above inequality holds if and only if

$$\sum_{i=1}^n Y_i \le \lambda e^{\gamma}\sum_{i=1}^n Z_i . \tag{0.1}$$

(Note that upon division by n the right side converges in probability to $\eta_0\lambda e^{\gamma}$ while the left side (divided by n) converges to $\eta_0\lambda_0 e^{\gamma_0} + \lambda_0(1-\eta_0)$.) At the MLE, the inequality (0.1) holds if and only if

$$\sum_{i=1}^n Y_i \le \sum_{i=1}^n Y_i Z_i .$$

This inequality can occur only if $Z_i = 1$ whenever $Y_i > 0$. Thus it becomes clear that the equations will have a unique solution yielding a maximum of the likelihood if at least one $Z_i = 0$ and $Y_i \ge 1$.
(c) Theorem 4.1.5 says that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \to_d N_3(0, I(\theta_0)^{-1})$$

where

$$I(\theta) = \begin{pmatrix} \lambda^{-1}E(e^{\gamma Z}) & E(Ze^{\gamma Z}) & 0 \\ E(Ze^{\gamma Z}) & \lambda E(Z^2 e^{\gamma Z}) & 0 \\ 0 & 0 & (\eta(1-\eta))^{-1} \end{pmatrix} .$$

(d) When θ_1 is true,

$$\begin{aligned} n^{-1}\log\frac{L_n(\theta_1)}{L_n(\theta_0)} &= n^{-1}\sum_{i=1}^n \log\frac{p_{\theta_1}}{p_{\theta_0}}(X_i) \\ &\to_p E_{\theta_1}\log\frac{p_{\theta_1}}{p_{\theta_0}}(X_1) \\ &= K(P_{\theta_1}, P_{\theta_0}) . \end{aligned}$$

Here

$$\begin{aligned} \log\frac{p_{\theta_1}}{p_{\theta_0}}(x) &= y\log\left(\frac{\lambda_1 e^{\gamma_1 z}}{\lambda_0 e^{\gamma_0 z}}\right) - \lambda_1 e^{\gamma_1 z} + \lambda_0 e^{\gamma_0 z} \\ &\quad + z\log\frac{\eta_1}{\eta_0} + (1-z)\log\frac{1-\eta_1}{1-\eta_0} , \end{aligned}$$

so

$$
\begin{aligned}
K(P_{\theta_1}, P_{\theta_0}) &= E_{\theta_1} \log \frac{p_{\theta_1}}{p_{\theta_0}}(X_1) \\
&= E_{\theta_1}\left(Y \log \frac{\lambda_1 e^{\gamma_1 Z}}{\lambda_0 e^{\gamma_0 Z}}\right) + \eta_1(\lambda_0 e^{\gamma_0} - \lambda_1 e^{\gamma_1}) + (1-\eta_1)(\lambda_0 - \lambda_1) \\
&\qquad + \eta_1 \log \frac{\eta_1}{\eta_0} + (1-\eta_1)\log\frac{1-\eta_1}{1-\eta_0} \\
&= E_{\theta_1}\left(\lambda_1 e^{\gamma_1 Z} \log\left(\frac{\lambda_1 e^{\gamma_1 Z}}{\lambda_0 e^{\gamma_0 Z}}\right)\right) + \eta_1(\lambda_0 e^{\gamma_0} - \lambda_1 e^{\gamma_1}) + (1-\eta_1)(\lambda_0 - \lambda_1) \\
&\qquad + \eta_1 \log \frac{\eta_1}{\eta_0} + (1-\eta_1)\log\frac{1-\eta_1}{1-\eta_0} \\
&= \eta_1\left(\lambda_1 e^{\gamma_1}\log\left(\frac{\lambda_1 e^{\gamma_1}}{\lambda_0 e^{\gamma_0}}\right)\right) + (1-\eta_1)\left(\lambda_1 \log\left(\frac{\lambda_1}{\lambda_0}\right)\right) \\
&\qquad + \eta_1(\lambda_0 e^{\gamma_0} - \lambda_1 e^{\gamma_1}) + (1-\eta_1)(\lambda_0 - \lambda_1) \\
&\qquad + \eta_1 \log \frac{\eta_1}{\eta_0} + (1-\eta_1)\log\frac{1-\eta_1}{1-\eta_0} .
\end{aligned}
$$

3. For the same set-up as in problem 1, consider taking a "profile likelihood" approach to the estimation of γ as follows:
(a) Let $l_n(\theta) = l_n(\gamma, \lambda, \eta)$: consider first maximizing this as a function of λ and η for each fixed value of γ to find

$$(\widehat{\lambda}(\gamma), \widehat{\eta}(\gamma)) \equiv \text{argmax}_{(\lambda,\eta)} l_n(\lambda, \gamma, \eta) .$$

Compute the maximizer $(\widehat{\lambda}(\gamma), \widehat{\eta}(\gamma))$ as explicitly as possible, and then form the "profile log-likelihood" $l_n^{profile}(\gamma)$ defined by

$$l_n^{profile}(\gamma) \equiv l_n(\widehat{\lambda}(\gamma), \gamma, \widehat{\eta}(\gamma)) .$$

(b) Now maximize $l_n^{profile}(\gamma)$ with respect to γ. Find the resulting "profile likelihood" score equation for γ.
(c) Does the equation you derived in (b) follow from the original score equations?
(d) Does the "profile score function" which appears in (b) correspond to or relate to the efficient score for γ in any way?

Solution: (a) The value of γ doesn't influence the maximization with respect to η and we find $\hat{\eta}(\gamma) = \hat{\eta} = \overline{Z}_n$ for all γ. Solving the score equation for λ for a fixed γ yields

$$\sum_1^n Y_i = \lambda \sum_1^n e^{\gamma Z_i}, \qquad \text{or} \qquad \hat{\lambda}(\gamma) = \frac{\sum_1^n Y_i}{\sum_1^n e^{\gamma Z_i}} .$$

Substitution of these into the log-likelihood yield the profile log-likelihood

$$
\begin{aligned}
l_n^{profile}(\gamma) &= l_n(\hat{\lambda}(\gamma), \gamma, \hat{\eta}) \\
&= \sum_{i=1}^n \left\{ Y_i \log(\hat{\lambda}(\gamma) e^{\gamma Z_i}) - \hat{\lambda}(\gamma) e^{\gamma Z_i} - \log(Y_i!) \right\}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + n\overline{Z}\log\overline{Z} + (n - n\overline{Z})\log(1-\overline{Z}) \\
&= n\overline{Y}\log\hat{\lambda}(\gamma) + \gamma\sum_{i=1}^n Y_iZ_i - \hat{\lambda}(\gamma)\sum_{i=1}^n e^{\gamma Z_i} + \text{constant in } \gamma\,.
\end{aligned}
$$

(b) Differentiating the profile log-likelihood with respect to γ yields

$$
\dot{l}_{n,\gamma}^{profile}(\gamma) = \frac{d}{d\gamma}\log\hat{\lambda}(\gamma)n\overline{Y}_n + \sum_{i=1}^n Y_iZ_i = \sum_{i=1}^n Y_iZ_i - n\overline{Y}\frac{\sum_1^n Z_ie^{\gamma Z_i}}{\sum_1^n e^{\gamma Z_i}}
$$

since

$$
\frac{d}{d\gamma}\log\hat{\lambda}(\gamma) = -\frac{\sum_1^n Z_ie^{\gamma Z_i}}{\sum_1^n e^{\gamma Z_i}}\,.
$$

Thus the profile score equation for γ becomes: $\hat{\gamma}^{profile} = \hat{\gamma}$ satisfies

$$
\frac{\sum_1^n Y_iZ_i}{\sum_1^n Y_i} = \frac{\sum_1^n Z_ie^{\hat{\gamma}Z_i}}{\sum_1^n e^{\hat{\gamma}Z_i}}\,. \tag{0.2}
$$

(c) If we solve the original score equation for λ for fixed γ, then we obtain

$$
\hat{\lambda}(\gamma) = \frac{\sum_1^n Y_i}{\sum_1^n e^{\gamma Z_i}}
$$

as in (a). Substitution of this into the score equation for γ yields

$$
\begin{aligned}
0 &= \sum_{i=1}^n Z_iY_i - \hat{\lambda}(\gamma)\sum_{i=1}^n Z_ie^{\gamma Z_i} \\
&= \sum_{i=1}^n Z_iY_i - \frac{\sum_1^n Y_i}{\sum_1^n e^{\gamma Z_i}}\sum_{i=1}^n Z_ie^{\gamma Z_i}\,,
\end{aligned}
$$

and this implies that (0.2) holds.

(d) To see the connection between the profile score function for γ and the efficient score function for γ, note that

$$
\dot{l}_n^{profile}(\gamma) = \dot{l}_{n,\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\frac{d}{d\gamma}\hat{\lambda}(\gamma) + \dot{l}_{n,\gamma}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\,, \tag{0.3}
$$

and, since $0 = \dot{l}_{n,\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})$, by differentiating with respect to γ we have

$$
0 = \ddot{l}_{n,\lambda\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\frac{d}{d\gamma}\hat{\lambda}(\gamma) + \ddot{l}_{n,\gamma\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\,,
$$

and hence

$$
\frac{d}{d\gamma}\hat{\lambda}(\gamma) = -\left(\ddot{l}_{n,\lambda\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\right)^{-1}\ddot{l}_{n,\gamma\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\,. \tag{0.4}
$$

Substitution of (0.4) into (0.3) yields

$$
\begin{aligned}
\dot{l}_n^{profile}(\gamma) &= \dot{l}_{n,\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\frac{d}{d\gamma}\hat{\lambda}(\gamma) + \dot{l}_{n,\gamma}(\hat{\lambda}(\gamma),\gamma,\hat{\eta}) \\
&= \dot{l}_{n,\gamma}(\hat{\lambda}(\gamma),\gamma,\hat{\eta}) - \ddot{l}_{n,\gamma\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\left(\ddot{l}_{n,\lambda\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta})\right)^{-1}\dot{l}_{n,\lambda}(\hat{\lambda}(\gamma),\gamma,\hat{\eta}) \\
&= \sum_{i=1}^n\left\{\dot{l}_\gamma(X_i) - \hat{I}_{\gamma\lambda}\hat{I}_{\gamma\gamma}^{-1}\dot{l}_\lambda(X_i)\right\}\Big|_{\theta=(\hat{\lambda}(\gamma),\gamma,\hat{\eta})}\,.
\end{aligned}
$$

4. 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

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

D. Does a maximum likelihood estimate of $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ exist? Is it unique?
E. Compute an approximate (one - step) maximum likelihood estimate $\check{\theta}$ of θ using the method of moment estimators $\overline{\theta}_n$ as the preliminary estimators based on the following data (with $n = 19$):

0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50,
7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89 .

[These are failure times in minutes for an insulating fluid between two electrodes subject to a voltage of 34 kV. – from Nelson, *Applied Life Data Analysis*, page 105.]
F. Compute the maximum likelihood estimator $\hat{\theta}_n$, and compare it with the one step estimator computed in E.

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

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

Then since $(\overline{\alpha}, \overline{\beta}) = g(\overline{T}, S_T^2)$ and $(\alpha, \beta) = g(E_\theta T, Var_\theta(T))$ where $g \equiv (g_1, g_2) : R^2 \to R^2$ 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 $\underline{\tilde{Z}} \equiv (Z_1, \sqrt{2}\sigma_T^2 Z_2)$ that

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

where

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

D. The maximum likelihood estimator exists and is unique in this model if not all the X_i's are equal (which happens with probability 1 if the model holds). The following solution is from Lehmann, TPE, page 536 (with slightly different notation).

We first reparametrize the Weibull model by writing

$$\begin{aligned} p_\theta(x) &= \frac{\beta}{\alpha}(\frac{x}{\alpha})^{\beta-1} \exp(-(\frac{x}{\alpha})^\beta) 1_{(0,\infty)}(x) \\ &= \frac{\beta}{\eta} x^{\beta-1} \exp(-\frac{x^\beta}{\eta}) \\ &\equiv p_\gamma(x) \end{aligned}$$

where $\eta \equiv \alpha^\beta$ and $\gamma \equiv (\beta, \eta)$. Then

$$l(\gamma|\underline{X}) = n\log\beta - n\log\eta + (\beta - 1)\sum_{i=1}^n \log X_i - \frac{1}{\eta}\sum_{i=1}^n X_i^\beta.$$

Thus, with $\gamma_1 \equiv \beta$, $\gamma_2 \equiv \eta$, the likelihood equations become

$$\dot{l}_1(\gamma|\underline{X}) = \frac{n}{\beta} + \sum_{i=1}^n \log X_i - \frac{1}{\eta}\sum_{i=1}^n X_i^\beta \log X_i = 0, \tag{0.5}$$

and

$$\dot{l}_2(\gamma|\underline{X}) = -\frac{n}{\eta} + \frac{1}{\eta^2}\sum_{i=1}^{n} X_i^{\beta} = 0, \tag{0.6}$$

or

$$\hat{\eta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^{\hat{\beta}} \tag{0.7}$$

from 0.6. Substitution of 0.7 into 0.5 yields the equation

$$\frac{\sum_i X_i^{\hat{\beta}} \log X_i}{\sum_i X_i^{\hat{\beta}}} - \frac{1}{\hat{\beta}} = \frac{1}{n}\sum_{i=1}^{n} \log X_i\,, \tag{0.8}$$

or

$$h(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n} \log X_i \tag{0.9}$$

where

$$h(\beta) \equiv \frac{\sum_i X_i^{\beta} \log X_i}{\sum_i X_i^{\beta}} - \frac{1}{\beta} < \frac{\sum_i X_i^{\beta} \log X_i}{\sum_i X_i^{\beta}}$$

since $\beta > 0$. Now

$$\begin{aligned} h'(\beta) &= \frac{\sum_i X_i^{\eta}(\log X_i)^2}{\sum_i X_i^{\beta}} - \Big(\frac{\sum_i X_i^{\beta} \log X_i}{\sum_i X_i^{\beta}}\Big)^2 + \frac{1}{\beta^2} \\ &\equiv I + II \\ &> I, \end{aligned}$$

and furthermore,

$$I = \sum a_i^2 p_i - \big(\sum a_i p_i\big)^2 = \mathrm{Var}_p(a)$$

since, with $a_i \equiv \log X_i$, $p_i \equiv X_i^{\beta}/\sum_j X_j^{\beta} \geq 0$, $\sum_i p_i = 1$. Thus $I > 0$ and hence $h'(\beta) > 0$ from (0.10) while

$$-\infty = \lim_{\beta\to 0} h(\beta) < \frac{1}{n}\sum_{i=1}^{n} \log X_i < \log X_{(n)} = \lim_{\beta\to\infty} h(\beta).$$

[Draw the picture!] (To see this last limit, note that with $p_{(i)} \equiv X_{(i)}^{\beta}/\sum_j X_j^{\beta}$,

$$\begin{aligned} p_{(i)} &= \frac{1}{(\frac{X_{(1)}}{X_{(i)}})^{\beta} + \ldots + (\frac{X_{(n)}}{X_{(i)}})^{\beta}} \\ &\to \begin{cases} 0, & i \leq n \quad\quad (\text{so } X_{(n)}/X_{(i)} > 1) \\ 1, & i = n \quad (\text{so } X_{(j)}/X_{(n)} < 1,\ j < n) \end{cases} \end{aligned}$$

as $\beta \to \infty$.) Thus (0.9) has a unique solution $\hat{\beta}$. By taking this value of $\hat{\beta}$ in (0.7), we see that the MLE $\hat{\gamma}$ of γ exists and is unique. Thus the unique MLE of $\theta = (\alpha, \beta)$ is $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ with $\hat{\alpha} = \hat{\eta}^{1/\hat{\beta}}$.
E. The one step estimator using $\hat{I}(\overline{\theta}_n) = I(\overline{\theta}_n)$ is

$$\check{\theta}_n \equiv \overline{\theta}_n + \hat{I}_n^{-1}(\overline{\theta}_n)\Big(\frac{1}{n}\dot{l}(\overline{\theta}_n)\Big) = (12.27\ldots\,, 0.7421\ldots).$$

The one - step estimator using $\hat{I}_n(\overline{\theta}_n) = (-n^{-1}\ddot{l}_n(\overline{\theta}_n)$ is

$$\check{\theta}_n = (11.778, 0.7633).$$

F. The maximum likelihood estimate $\hat{\theta}_n = (12.222..., 0.77082...)$; see the following pages.