Statistics 581, Problem Set 7 Solutions

Wellner; 11/22/2006

1. Compute and plot the *score for location*, $-(f'/f)(x)$ when:
   A. $f(x) = \phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, (normal or Gaussian);
   B. $f(x) = \exp(-x)/(1+\exp(-x))^2$, (logistic);
   C. $f(x) = \frac{1}{2}\exp(-|x|)$, (double exponential);
   D. $f = t_k$, the $t-$distribution with k degrees of freedom;
   E. $f(x) = \exp(-x)\exp(-\exp(-x))$, Gumbel or extreme value.

   Soluton: A. For $f(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, it follows that $\log f(x) = -x^2/2 +$ *constant* so that $(-f'/f)(x) = x$, $-1 - x(f'/f)(x) = x^2 - 1$.
   B. For $f(x) = e^{-x}/(1+e^{-x})^2$, $\log f(x) = -x - 2\log(1+e^{-x})$ and

$$-\frac{f'}{f}(x) = \frac{1-e^{-x}}{1+e^{-x}},$$

while

$$-1 - x\frac{f'}{f}(x) = x\frac{1-e^{-x}}{1+e^{-x}} - 1 \sim |x| - 1 \qquad \text{as} \qquad |x| \to \infty.$$

C. For $f(x) = 2^{-1}\exp(-|x|)$,

$$\log f(x) = -|x| + \textit{constant},$$

and

$$-\frac{f'}{f}(x) = \begin{cases} -1 & x < 0 \\ \text{undefined} & x = 0 \\ +1 & x > 0 \end{cases},$$

while

$$-1 - x\frac{f'}{f}(x) = |x| - 1, \qquad \text{for} \qquad x \neq 0.$$

D. For the t_k distribution, $f(x) = \frac{\Gamma(\frac{1}{2}(k+1))}{\Gamma(\frac{1}{2}k)}\frac{1}{\sqrt{\pi k}}(1+\frac{x^2}{k})^{-(k+1)/2}$,

$$\log f(x) = -\frac{k+1}{2}\log(1+\frac{x^2}{k}),$$

and

$$-\frac{f'}{f}(x) = \frac{k+1}{k}\frac{x}{1+\frac{x^2}{k}},$$

while

$$-1 - x\frac{f'}{f}(x) = k\frac{x^2-1}{x^2+k}.$$

E. For $f(x) = \exp(-x)\exp(-\exp(-x))$,

$$\log f(x) = -x - \exp(-x),$$

and

$$-\frac{f'}{f}(x) = 1 - \exp(-x),$$

while

$$-1 - x\frac{f'}{f}(x) = -1 + x(1 - \exp(-x)).$$

Plots of these score functions for location are given in Figure 1. Note that they are *odd functions* in cases A-D, which are all symmetric densities about zero. In case E, corresponding to the asymmetric extreme value density, the score for location does not have any obvious symmetry property.

Figure 1: Scores for location.

2. Compute $I_f = \int (f'(x)/f(x))^2 f(x) dx$, the information for location, for each of the densities in problem 1.

   Solution: A. In this case $I_f = \int x^2 \phi(x) dx = Var(Z) = 1$ where $Z \sim N(0,1)$.
   B. For the logistic density the information for location is

$$\begin{aligned} I_f &= \int_{-\infty}^{\infty} (\frac{1-e^{-x}}{1+e^{-x}})^2 dF(x) = \int_{-\infty}^{\infty} (2F(x)-1)^2 dF(x) \\ &= \int_0^1 (2u-1)^2 du = 4Var(U) = 4\frac{1}{12} = \frac{1}{3}. \end{aligned}$$

   C. For the double-exponential density, $[(-f'/f)(x)]^2 = 1$, so $I_f = 1$.
   D. For the $t-$ distribution with k degrees of freedom, by using a change of variables and letting T_r denote a random variable with the $t-$ distribution with

r degrees of freedom,

$$\begin{aligned} I_f &= \int_{-\infty}^{\infty} (\frac{k+1}{k})^2 \frac{x^2}{(1+x^2/k)^2} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{k}{2})\sqrt{\pi k}} \frac{1}{(1+x^2/k)^{(k+1)/2}} dx \\ &= \frac{(k+1)(k+2)}{(k+4)(k+3)} Var(T_{k+4}) \\ &= \frac{(k+1)(k+2)}{(k+4)(k+3)} \frac{k+4}{k+2} = \frac{k+1}{k+3} \end{aligned}$$

since $Var(T_r) = r/(r-2)$ for $r > 2$.

E. For the extreme value distribution $F(x) = \exp(-\exp(-x))$ and therefore if $X \sim F$, the random variable $Y \equiv \exp(-X) \sim$ exponential(1):

$$\begin{aligned} P(Y \geq y) &= P(\exp(-X) \geq y) = P(X \leq -\log(y)) \\ &= \exp(-\exp(\log(y))) = \exp(-y). \end{aligned}$$

Since $-(f'/f)(x) = -1 + e^{-x}$, it is easy to see that

$$I_f = E[-\frac{f'}{f}(X)]^2 = E[\exp(-X) - 1]^2 = E[Y-1]^2 = Var(Y) = 1.$$

3. Consider the two parameter location-scale model

$$\mathcal{P} = \{P_\theta : \frac{dP_\theta}{d\lambda} = p_\theta : \theta \in \Theta\}$$

where $\Theta = \mathbb{R} \times \mathbb{R}^+$,

$$p_\theta(x) = \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right),$$

and the (known) density f has a derivative f' almost everywhere with respect to Lebesgue measure λ.

(a) Calculate the information matrix $I(\theta)$ for θ.

(b) For which of the densities in A-E of problem 1 is $I_{12}(\theta)$ not zero?

Solution: (a) The score for location is

$$\begin{aligned} \dot{l}_1(x) &= \frac{\partial}{\partial\theta_1} \log\left\{\frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right)\right\} \\ &= -\frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\frac{1}{\theta_2} \end{aligned}$$

and the score for scale is

$$\begin{aligned} \dot{l}_2(x) &= \frac{\partial}{\partial\theta_2} \log\left\{\frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right)\right\} \\ &= -\frac{1}{\theta_2} - \frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\frac{(x-\theta_1)}{\theta_2^2} \end{aligned}$$

Thus we compute

$$
\begin{aligned}
I_{11}(\theta) &= E\dot{\mathbf{l}}_1^2(X) = \frac{1}{\theta_2^2}\int \left(\frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\right)^2 \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right) dx \\
&= \frac{1}{\theta_2^2}\int \left(\frac{f'}{f}(y)\right)^2 f(y)dy \equiv \frac{1}{\theta_2^2} I_{f,loc}, \\
I_{22}(\theta) &= E\dot{\mathbf{l}}_2^2(X) = \frac{1}{\theta_2^2}\int \left(-1 - \frac{(x-\theta_1)}{\theta_2}\frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\right)^2 \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right) dx \\
&= \frac{1}{\theta_2^2}\int \left(-1 - y\frac{f'}{f}(y)\right)^2 f(y)dy \equiv \frac{1}{\theta_2^2} I_{f,scale}, \\
I_{12}(\theta) &= E\dot{\mathbf{l}}_1(X)\dot{\mathbf{l}}_2(X) = I_{21}(\theta) \\
&= \frac{1}{\theta_2^2}\int \left(-\frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\right)\left(-1 - \frac{(x-\theta_1)}{\theta_2}\frac{f'}{f}\left(\frac{x-\theta_1}{\theta_2}\right)\right) \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right) dx \\
&= \frac{1}{\theta_2^2}\int y \left(\frac{f'}{f}(y)\right)^2 f(y)dy \equiv \frac{1}{\theta_2^2} I_{12,f}.
\end{aligned}
$$

(b) Note that $I_{12,f} = 0$ for all symmetric densities f since y is odd while $(f'/f)^2 f$ is even. Thus $I_{12}(\theta) = 0$ for cases A-D in problem 1, while $I_{12}(\theta) = \theta_2^{-2} I_{12,f} \neq 0$ in case E. I calculate

$$
\begin{aligned}
I_{12,f} &= \int y \left(\frac{f'}{f}(y)\right)^2 f(y)dy \\
&= \int_{-\infty}^{\infty} y(-1+e^{-y})^2 e^{-y}\exp(-e^{-y})dy \\
&= -\int_0^{\infty} (\log v)(v-1)^2 e^{-v} dv = -(1-\gamma)
\end{aligned}
$$

where γ is Euler's constant. In fact this is related to I_{12} that we calculated already for the Weibull family in Example xx.xx in the notes and in the first display at the top of page 2 of the Handout on Gamma, Digama, and Polygamma.

4. Suppose that $X \sim \text{Gamma}(\alpha, \beta)$; i.e. X has density p_θ given by

$$
p_\theta(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp(-\beta x) 1_{(0,\infty)}(x), \quad \theta = (\alpha, \beta) \in (0,\infty)\times(0,\infty) \equiv \Theta\, .
$$

Consider estimation of : A. $q_A(\theta) \equiv E_\theta X$. B. $q_B(\theta) \equiv F_\theta(x_0)$ for a fixed x_0; here $F_\theta(x) \equiv P_\theta(X \le x)$.
(i) Compute $I(\theta) = I(\alpha, \beta)$; compare Lehmann & Casella page 127, Table 6.1
(ii) Compute $q_A(\theta)$, $q_B(\theta)$, $\dot{q}_A(\theta)$, and $\dot{q}_B(\theta)$.
(iii) Find the efficient influence functions for estimation of q_A and q_B.
(iv) Compare the efficient influence functions you find in (iii) with the influence functions ψ_A and ψ_B of the natural nonparametric estimators $\overline{X}_n$ and $\mathbb{F}_n(x_0)$ respectively; in particular, show that $\psi_A \in \dot{\mathcal{P}}$, while $\psi_B \notin \dot{\mathcal{P}}$.

Solution: For the Gamma(α, β) parametrized *my* way:

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

Thus

$$\log p_\theta(x) = (\alpha-1)\log x + \alpha\log\beta - \log\Gamma(\alpha) - \beta x,$$

and hence

$$\dot{l}_\alpha(x) = \log x + \log\beta - \frac{\Gamma'}{\Gamma}(\alpha) = \log(\beta x) - \psi(\alpha),$$
$$\dot{l}_\beta(x) = \frac{\alpha}{\beta} - x$$

Furthermore,

$$\ddot{l}_{\alpha\alpha}(x) = -\psi'(\alpha),$$
$$\ddot{l}_{\alpha\beta}(x) = \frac{1}{\beta} = \ddot{l}_{\beta\alpha}(x),$$
$$\ddot{l}_{\beta\beta}(x) = -\frac{\alpha}{\beta^2}.$$

Hence

$$I(\theta) = \begin{pmatrix} \psi'(\alpha) & -1/\beta \\ -1/\beta & \alpha/\beta^2 \end{pmatrix}.$$

(ii). Now $q_A(\theta) = \alpha/\beta$, and

$$\begin{aligned} q_B(\theta) = P_\theta(X \le x_0) &= \int_0^{x_0} \frac{1}{\Gamma(\alpha)}\beta^\alpha x^{\alpha-1}e^{-\beta x}\,dx \\ &= \int_0^{\beta x_0} \frac{1}{\Gamma(\alpha)} t^{\alpha-1}e^{-t}\,dt \\ &\equiv \frac{\Gamma(\alpha,\beta x_0)}{\Gamma(\alpha)} \end{aligned}$$

where $\Gamma(\alpha, y)$ is the incomplete gamma function; note that $\Gamma(\alpha,\infty) = \Gamma(\alpha)$. Therefore

$$\begin{aligned} \dot{q}_A^T(\theta) = \left(\frac{\partial}{\partial\alpha}q_A, \frac{\partial}{\partial\beta}q_B\right) &= (\frac{1}{\beta}, -\frac{\alpha}{\beta^2}) = \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \\ &= \mathrm{Cov}_\theta(X - E_\theta(X), \dot{l}_\theta^T(X)), \end{aligned}$$

while, with

$$\psi(\alpha, y) \equiv \frac{\partial}{\partial\alpha}\log\Gamma(\alpha,y) \equiv \Gamma'(\alpha,y)/\Gamma(\alpha,y)\,,$$

$$\begin{aligned} \dot{q}_B^T(\theta) &= \left(\frac{\Gamma'(\alpha,\beta x_0)}{\Gamma(\alpha)} - \frac{\Gamma(\alpha,\beta x_0)\Gamma'(\alpha)}{\Gamma^2(\alpha)}, \frac{(\beta x_0)^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta x_0}\frac{x_0}{\beta}\beta\right) \\ &= \left(\frac{\Gamma(\alpha,\beta x_0)}{\Gamma(\alpha)}\{\psi(\alpha,\beta x_0) - \psi(\alpha)\}, \frac{x_0}{\beta}p_\theta(x_0)\right) \\ &= (q_B(\theta)\{\psi(\alpha,\beta x_0) - \psi(\alpha)\}, \frac{x_0}{\beta}p_\theta(x_0)) \\ &= \mathrm{Cov}_\theta[(1_{[0,x_0]}(X) - F_\theta(x_0)), \dot{l}_\theta^T]\,. \end{aligned}$$

(iii). The scores are given by

$$\dot{l}_\theta(x) = \begin{pmatrix} \dot{l}_\alpha(x) \\ \dot{l}_\beta(x) \end{pmatrix} = \begin{pmatrix} \log(\beta x) - \psi(\alpha) \\ \frac{\alpha}{\beta} - x \end{pmatrix}$$

and the information matrix is

$$I(\theta) = \begin{pmatrix} \psi'(\alpha) & -1/\beta \\ -1/\beta & \alpha/\beta^2 \end{pmatrix}.$$

Thus

$$I^{-1}(\theta) = \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1},$$

and the efficient influence function for estimation of q_A is

$$\begin{aligned} \tilde{l}_A &= \dot{q}_A(\theta)^T I^{-1}(\theta)\dot{l}_\theta \\ &= \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1} \begin{pmatrix} \log(\beta x) - \psi(\alpha) \\ \frac{\alpha}{\beta} - x \end{pmatrix} \\ &= \frac{\beta}{\alpha\psi'(\alpha) - 1}\{0 \cdot (\log(\beta x) - \psi(\alpha)) + (\frac{1}{\beta} - \frac{\alpha}{\beta}\psi'(\alpha))(\frac{\alpha}{\beta} - x)\} \\ &= (x - \frac{\alpha}{\beta}). \end{aligned}$$

Note that $X - E_\theta(X) \in [\dot{l}_\theta] = \dot{\mathcal{P}}$; in fact, $X - E_\theta(X) = -\dot{l}_\beta(X)$.
Similarly, $\tilde{l}_B(x) = \dot{q}_B(\theta)I^{-1}(\theta)\dot{l}_\theta(x)$; unfortunately, this does not simplify much, largely due to the fact that $1_{[0,x_0]}(X) - F_\theta(x_0) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$.
(iii) The information bound for estimation of q_A is

$$\begin{aligned} I^{-1}(P|q_A, \mathcal{P}) &= \dot{q}_A^T I^{-1}(\theta)\dot{q}_A \\ &= \frac{1}{\beta}(1, -\frac{\alpha}{\beta}) \begin{pmatrix} \alpha/\beta^2 & 1/\beta \\ 1/\beta & \psi'(\alpha) \end{pmatrix} \frac{\beta^2}{\alpha\psi'(\alpha) - 1} \begin{pmatrix} 1 \\ -\alpha/\beta \end{pmatrix} \frac{1}{\beta} \\ &= \frac{\alpha}{\beta^2} = Var_\theta(X). \end{aligned}$$

Similarly,

$$I^{-1}(P|q_B, \mathcal{P}) \;\; = \;\; = \dot{q}_B^T I^{-1}(\theta)\dot{q}_B,$$

which does not simplify appreciably because $1_{[0,x_0]}(X) - F_\theta(x_0) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$. However, since we know that $\tilde{l}_B = \Pi(1_{[0,x_0]}(x) - F(x_0)|\dot{\mathcal{P}})$, it follows easily that

$$I^{-1}(P|q_B, \mathcal{P}) < E_\theta(1_{[0,x_0]}(X) - F_\theta(x_0))^2 = F_\theta(x_0)(1 - F_\theta(x_0));$$

i.e. it is possible to improve on the natural nonparametric estimator $\mathbb{F}_n(x_0)$ of $q_B(\theta) = F_\theta(x_0)$ when the model holds.

5. Suppose that $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, $\Theta \subset R^k$ is a parametric model satisfying the hypotheses of the multiparameter Cramér - Rao inequality. Partition θ as $\theta = (\nu, \eta)$ where $\nu \in R^m$ and $\eta \in R^{k-m}$ and $1 \le m < k$. Let $\dot{\mathbf{l}} = \dot{\mathbf{l}}_\theta = (\dot{\mathbf{l}}_1, \dot{\mathbf{l}}_2)$ be the corresponding partition of the (vector of) scores $\dot{\mathbf{l}}$, and, with $\widetilde{\mathbf{l}} \equiv I^{-1}(\theta)\dot{\mathbf{l}}$, the

efficient influence function for θ, let $\widetilde{\mathbf{l}} = (\widetilde{\mathbf{l}}_1, \widetilde{\mathbf{l}}_2)$ be the corresponding partition of $\widetilde{l}$. In both cases, $\dot{\mathbf{l}}_1$, $\widetilde{\mathbf{l}}_1$ are $m-$vectors of functions, and $\dot{\mathbf{l}}_2$, $\widetilde{\mathbf{l}}_2$ are $k-m$ vectors. Partition $I(\theta)$ and $I^{-1}(\theta)$ correspondingly as

$$I(\theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$$

where I_{11} is $m \times m$, I_{12} is $m \times (k-m)$, I_{21} is $(k-m) \times m$, I_{22} is $(k-m) \times (k-m)$. Also write

$$I^{-1}(\theta) = [I^{ij}]_{i,j=1,2}.$$

Verify that:

A. $I^{11} = I_{11\cdot 2}^{-1}$ where $I_{11\cdot 2} \equiv I_{11} - I_{12} I_{22}^{-1} I_{21}$,

$I^{22} = I_{22\cdot 1}^{-1}$ where $I_{22\cdot 1} \equiv I_{22} - I_{21} I_{11}^{-1} I_{12}$,

$I^{12} = -I_{11\cdot 2}^{-1} I_{12} I_{22}^{-1}$,

$I^{21} = -I_{22\cdot 1}^{-1} I_{21} I_{11}^{-1}$.

This amounts to formulas (5) and (6) of section 3.2, page 15.

B. Verify that

$\widetilde{\mathbf{l}}_1 = I^{11}\dot{\mathbf{l}}_1 + I^{12}\dot{\mathbf{l}}_2 = I_{11\cdot 2}^{-1}(\dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1}\dot{\mathbf{l}}_2)$, and

$\widetilde{\mathbf{l}}_2 = I^{21}\dot{\mathbf{l}}_1 + I^{22}\dot{\mathbf{l}}_2 = I_{22\cdot 1}^{-1}(\dot{\mathbf{l}}_2 - I_{21} I_{11}^{-1}\dot{\mathbf{l}}_1)$.

The first of these is (7) on page 15, section 3.2.

Solution: A. This is just block inversion/multiplication of matrices:

$$\begin{aligned}
\begin{pmatrix} I^{11} & I^{12} \\ I^{21} & I^{22} \end{pmatrix} \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} &= \begin{pmatrix} I_{11\cdot}^{-1} & -I_{11\cdot 2}^{-1} I_{21} I_{22}^{-1} \\ -I_{22\cdot 1}^{-1} I_{21} I_{11}^{-1} & I_{22\cdot 1} \end{pmatrix} \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix} \\
&= \begin{pmatrix} I_{11\cdot 2}^{-1}(I_{11} - I_{12} I_{22}^{-1} I_{21}) & I_{11\cdot 2}^{-1}(I_{12} - I_{12}) \\ I_{22\cdot 1}^{-1}(-I_{21} + I_{21}) & I_{22\cdot 1}^{-1}(-I_{21} I_{11} I_{12} + I_{22}) \end{pmatrix} \\
&= \begin{pmatrix} Ident & 0 \\ 0 & Ident \end{pmatrix} = Identity.
\end{aligned}$$

by using the definition of $I_{11\cdot 2}$ and $I_{22\cdot 1}$.

B. This follows immediately from the formulas for I^{11} and I^{12} by just plugging into the formula $\tilde{\mathbf{l}}_1 = I^{11}\dot{\mathbf{l}}_1 + I^{12}\dot{\mathbf{l}}_2$ for $\tilde{\mathbf{l}}_1$:

$$\begin{aligned}
\tilde{\mathbf{l}}_1 &= I_{11\cdot 2}^{-1}\dot{\mathbf{l}}_1 - I_{11\cdot 2}^{-1} I_{12} I_{22}^{-1}\dot{\mathbf{l}}_2 \\
&= I_{11\cdot 2}^{-1}(\dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1}\dot{\mathbf{l}}_2) = I_{11\cdot 2}^{-1}\mathbf{l}_1^* .
\end{aligned}$$