

Statistics 581, Problem Set 7 Solutions - Corrected

Wellner; 11/17/2010

1. Suppose that  $X_1, \dots, X_n$  are i.i.d. with the Weibull distribution  $F_\theta$  given by

$$1 - F_\theta(x) = \exp(-(x/\alpha)^\beta), \quad x \geq 0$$

where  $\theta = (\alpha, \beta) \in (0, \infty) \times (0, \infty)$ .

(a) Find the inverse (or quantile function)  $F_\theta^{-1}(u)$  corresponding to  $F_\theta$  in terms of  $\alpha$ ,  $\beta$ , and  $u \in (0, 1)$ , and show that

$$\log F_\theta^{-1}(u) = \log \alpha + \frac{1}{\beta} \log \log \left( \frac{1}{1-u} \right).$$

(b) Fix  $r \in (0, 1/2)$  and  $s \in (1/2, 1)$ . Use the  $r$ -th and  $s$ -th quantiles of the  $X_i$ 's, namely  $\mathbb{F}_n^{-1}(r)$  and  $\mathbb{F}_n^{-1}(s)$ , to obtain simple consistent estimators  $\hat{\alpha}_n$  and  $\hat{\beta}_n$  of  $\alpha$  and  $\beta$ . Prove that your estimators are consistent.

(c) Prove that your estimators  $\hat{\alpha}_n$  and  $\hat{\beta}_n$  satisfy

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}_n - \alpha \\ \hat{\beta}_n - \beta \end{pmatrix} \rightarrow_d N_2(0, \Sigma)$$

and identify  $\Sigma$  as a function of  $\alpha$ ,  $\beta$ , and  $t$ .

(d) How would you choose  $r$  and  $s$  to minimize the asymptotic variance of  $\hat{\beta}_n$ ?

**Solution:** (a) Since  $1 - F_\theta(x) = \exp(-(x/\alpha)^\beta)$ , it follows we can solve  $F_\theta(x) = u$  for  $x = F_\theta^{-1}(u)$ . This yields

$$F_\theta^{-1}(u) = \alpha(-\log(1-u))^{1/\beta},$$

or

$$\log F_\theta^{-1}(u) = \log \alpha + \frac{1}{\beta} \log \log \left( \frac{1}{1-u} \right). \quad (0.1)$$

(b) Since we can estimate  $F_\theta^{-1}(r)$  and  $F_\theta^{-1}(s)$  respectively by  $\mathbb{F}_n^{-1}(r)$  and  $\mathbb{F}_n^{-1}(s)$  respectively, the relationship in (0.1) suggests that we estimate  $\alpha$  and  $\beta$  as the solutions  $\hat{\alpha}$  and  $\hat{\beta}$  of the pair of equations

$$\log \mathbb{F}_n^{-1}(r) = \log \hat{\alpha} + \frac{1}{\hat{\beta}} \log \log 1/(1-r), \quad (0.2)$$

$$\log \mathbb{F}_n^{-1}(s) = \log \hat{\alpha} + \frac{1}{\hat{\beta}} \log \log 1/(1-s). \quad (0.3)$$

Letting  $A_r \equiv \log \log 1/(1-r)$ , and  $B_s \equiv \log \log 1/(1-s)$ , we find that

$$\begin{aligned} 1/\hat{\beta} &= \frac{1}{B_s - A_r} (\log \mathbb{F}_n^{-1}(s) - \log \mathbb{F}_n^{-1}(r)) \\ &\equiv a_{r,s} \log \mathbb{F}_n^{-1}(s) - a_{r,s} \log \mathbb{F}_n^{-1}(r) \end{aligned}$$

and

$$\begin{aligned}\log \hat{\alpha} &= \frac{-A_r}{B_s - A_r} \log \mathbb{F}_n^{-1}(s) + \frac{B_s}{B_s - A_r} \log \mathbb{F}_n^{-1}(r) \\ &\equiv c_{r,s} \log \mathbb{F}_n^{-1}(s) + d_{r,s} \log \mathbb{F}_n^{-1}(r)\end{aligned}$$

where

$$a_{r,s} \equiv \frac{1}{B_s - A_r}, \quad c_{r,s} \equiv -A_r a_{r,s} \quad d_{r,s} \equiv B_s a_{r,s}.$$

Since  $(\mathbb{F}_n^{-1}(r), \mathbb{F}_n^{-1}(s)) \rightarrow_{a.s.} (F_\theta^{-1}(r), F_\theta^{-1}(s))$ , It follows easily by the continuous mapping theorem that

$$\frac{1}{\hat{\beta}} \rightarrow_{a.s.} a_{r,s} \log F_\theta^{-1}(s) - a_{r,s} \log F_\theta^{-1}(r) = \frac{1}{\beta},$$

and

$$\log \hat{\alpha} \rightarrow_{a.s.} c_{r,s} \log F_\theta^{-1}(s) + d_{r,s} \log F_\theta^{-1}(r) = \log \alpha,$$

and hence by the continuous mapping theorem,  $(\hat{\alpha}, \hat{\beta}) \rightarrow_{a.s.} (\alpha, \beta)$ .

(c) First, we know that

$$\sqrt{n} \begin{pmatrix} \mathbb{F}_n^{-1}(r) - F^{-1}(r) \\ \mathbb{F}_n^{-1}(s) - F^{-1}(s) \end{pmatrix} \rightarrow_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \frac{r(1-r)}{f^2(F^{-1}(r))} & \frac{r(1-s)}{f(F^{-1}(r))f(F^{-1}(s))} \\ \frac{r(1-s)}{f(F^{-1}(r))f(F^{-1}(s))} & \frac{s(1-s)}{f^2(F^{-1}(s))} \end{pmatrix}.$$

This implies that

$$\sqrt{n} \begin{pmatrix} \log \mathbb{F}_n^{-1}(r) - \log F^{-1}(r) \\ \log \mathbb{F}_n^{-1}(s) - \log F^{-1}(s) \end{pmatrix} \rightarrow_d D\underline{Z} \sim N_2(0, D\Sigma D^T)$$

where

$$D = \begin{pmatrix} 1/F^{-1}(r) & 0 \\ 0 & 1/F^{-1}(s) \end{pmatrix}.$$

Hence it follows that

$$\begin{aligned}&\sqrt{n} \begin{pmatrix} 1/\hat{\beta} - 1/\beta \\ \log \hat{\alpha} - \log \alpha \end{pmatrix} \\ &= M\sqrt{n} \begin{pmatrix} \log \mathbb{F}_n^{-1}(r) - \log F^{-1}(r) \\ \log \mathbb{F}_n^{-1}(s) - \log F^{-1}(s) \end{pmatrix} \\ &\rightarrow_d MD\underline{Z} \sim N_2(0, MD\Sigma D^T M^T).\end{aligned}$$

where

$$M = \begin{pmatrix} -a_{r,s} & a_{r,s} \\ d_{r,s} & c_{r,s} \end{pmatrix} = a_{r,s} \begin{pmatrix} -1 & 1 \\ B_s & -A_r \end{pmatrix}.$$

Finally, with  $g(x, y) = (g_1(x), g_2(y))$ ,  $g_1(x) = 1/x$ ,  $g_2(y) = \exp y$ , we find, by the delta-method, that

$$\begin{aligned} \sqrt{n} \begin{pmatrix} \hat{\beta} - \beta \\ \hat{\alpha} - \alpha \end{pmatrix} \\ \rightarrow_d \nabla g M D \underline{Z} \sim N_2(0, \nabla g M D \Sigma D^T M^T \nabla g^T) \end{aligned}$$

where

$$\nabla g = \begin{pmatrix} -\beta^2 & 0 \\ 0 & \alpha \end{pmatrix}.$$

We begin combining all this by noting that  $D\Sigma D^T$  involves the function

$$\begin{aligned} F^{-1}(u)f(F^{-1}(u)) &= \alpha \left( \log \left( \frac{1}{1-u} \right) \right)^{1/\beta} \frac{\beta}{\alpha} \left( \log \left( \frac{1}{1-u} \right) \right)^{(\beta-1)/\beta} (1-u) \\ &= \beta(1-u) \log \left( \frac{1}{1-u} \right) \equiv \beta g(u) \end{aligned}$$

at the points  $u = r$  and  $u = s$ . Computing  $D\Sigma D^T$  yields

$$D\Sigma D^T = \beta^{-2} \begin{pmatrix} \frac{r(1-r)}{g(r)^2} & \frac{r(1-s)}{g(r)g(s)} \\ \frac{r(1-s)}{g(r)g(s)} & \frac{s(1-s)}{g(s)^2} \end{pmatrix} \equiv \beta^{-2} \begin{pmatrix} c_{11}(r, s) & c_{12}(r, s) \\ c_{12}(r, s) & c_{22}(r, s) \end{pmatrix}.$$

Since the matrix  $M$  just depends on  $r, s$ , we find that the matrix

$$M D \Sigma D^T M^T = \beta^{-2} a_{r,s}^2 \begin{pmatrix} d_{11}(r, s) & d_{12}(r, s) \\ d_{12}(r, s) & d_{22}(r, s) \end{pmatrix},$$

where

$$\begin{aligned} d_{11}(r, s) &= c_{11}(r, s) - 2c_{12}(r, s) + c_{22}(r, s) \\ d_{12}(r, s) &= B_s(c_{12}(r, s) - c_{11}(r, s)) - A_r(c_{22}(r, s) - c_{12}(r, s)) \\ d_{22}(r, s) &= A_r^2 c_{22}(r, s) - 2A_r B_s c_{12}(r, s) + B_s^2 c_{11}(r, s). \end{aligned}$$

Thus we conclude that the asymptotic covariance matrix of  $(\hat{\beta}, \hat{\alpha})$  is given by

$$\nabla g M D \Sigma D^T M^T \nabla g^T = a_{r,s}^2 \begin{pmatrix} \beta^2 d_{11}(r, s) & -\alpha d_{12}(r, s) \\ -\alpha d_{12}(r, s) & (\alpha/\beta)^2 d_{22}(r, s) \end{pmatrix}.$$

(d) The asymptotic variance of  $\hat{\beta}$  is

$$\beta^2 a_{r,s}^2 d_{11}(r, s) = \beta^2 (c_{11}(r, s) - 2c_{12}(r, s) + c_{22}(r, s)) a_{r,s}^2.$$

This is minimized by  $r = r_0 \approx .1704$ ,  $s = s_0 \approx .97$ , and the minimum value is  $\beta^2(.917) > \beta^2(6/\pi^2)$ . This ad-hoc estimator  $\hat{\beta}$  based on quantiles is *inefficient*; its asymptotic variance (for any value of  $r, s$ , including the minimizing  $r_0, s_0$ ) is larger than the best possible asymptotic variance, which is  $\beta^2(6/\pi^2)$  as we will see in Chapter 3. In fact the ARE when  $r = r_0$  and  $s = s_0$  is  $(6/\pi^2)/.917 = .663$

The asymptotic variance of  $\hat{\alpha}$  is

$$(\alpha/\beta)^2 a_{r,s}^2 d_{22}(r, s) = (\alpha/\beta)^2 (B_s^2 c_{11}(r, s) - 2A_r B_s c_{12}(r, s) + B_s^2 c_{22}(r, s)).$$

This is minimized by  $r = r_0 \approx .398$ ,  $s = s_0 \approx .82$ , and the minimum value is  $(\alpha/\beta)^2(1.359) > (\alpha/\beta)^2(1.11)$ . This ad-hoc estimator  $\hat{\alpha}$  based on quantiles is also *inefficient*; its asymptotic variance (for any value of  $r, s$ , including the minimizing  $r_0, s_0$ ) is larger than the best possible asymptotic variance, which is about  $(\alpha/\beta)^2(1.11)$  as we will see in Chapter 3. For the estimator based on the optimal  $r_0, s_0$  (for  $\alpha!$ ), the ARE is  $\approx 1.109/1.359 = .816$

2. (a) Compute and plot the *score for location*,  $-(f'/f)(x)$  when:
- $f(x) = \phi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ , (normal or Gaussian);
  - $f(x) = \exp(-x)/(1 + \exp(-x))^2$ , (logistic);
  - $f(x) = \frac{1}{2} \exp(-|x|)$ , (double exponential);
  - $f = t_k$ , the  $t$ -distribution with  $k$  degrees of freedom;
  - $f(x) = \exp(-x) \exp(-\exp(-x))$ , Gumbel or extreme value.
- (b) Compute  $I_f = \int (f'(x)/f(x))^2 f(x) dx$ , the information for location, for each of the densities in (a).

**Soluton:** (a)

A. For  $f(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ , it follows that  $\log f(x) = -x^2/2 + \text{constant}$  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

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

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

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

and

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

while

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

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.

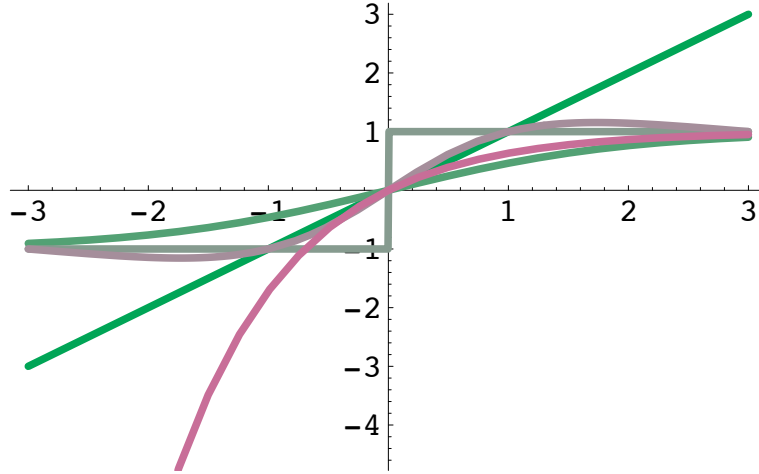


Figure 1: Scores for location.

(b) A. In this case  $I_f = \int x^2 \phi(x) dx = \text{Var}(Z) = 1$  where  $Z \sim N(0, 1)$ .

B. For the logistic density the information for location is

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

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,

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

since  $\text{Var}(T_r) = r/(r-2)$  for  $r > 2$ .

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

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

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

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

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  $\lambda$ .

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

(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

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

and the score for scale is

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

Thus we compute

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

(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  $\gamma$  is Euler's constant. In fact this is related to  $I_{12}$  that we calculated already for the Weibull family in Example 3.2.5 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  $X \sim \text{Beta}(\alpha, \beta)$ ; i.e.  $X$  has density  $p_\theta$  given by

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

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  $F_\theta(x) \equiv P_\theta(X \leq x)$ .

- (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  $\psi_A$  and  $\psi_B$  of the natural nonparametric estimators  $\bar{X}_n$  and  $\mathbb{F}_n(x_0)$  respectively. Do we have either  $\psi_A \in \dot{\mathcal{P}}$ , while  $\psi_B \notin \dot{\mathcal{P}}$ ?

**Solution:** For the Beta( $\alpha, \beta$ ) density:

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

Thus

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

and hence

$$\begin{aligned} \dot{l}_\alpha(x) &= \log x + \psi(\alpha + \beta) - \psi(\alpha), \\ \dot{l}_\beta(x) &= \log(1 - x) + \psi(\alpha + \beta) - \psi(\beta). \end{aligned}$$

Furthermore,

$$\begin{aligned} \ddot{l}_{\alpha\alpha}(x) &= \psi'(\alpha + \beta) - \psi'(\alpha), \\ \ddot{l}_{\alpha\beta}(x) &= \psi'(\alpha + \beta), \\ \ddot{l}_{\beta\beta}(x) &= \psi'(\alpha + \beta) - \psi'(\beta). \end{aligned}$$

Hence

$$I(\theta) = \begin{pmatrix} \psi'(\alpha) - \psi'(\alpha + \beta) & -\psi'(\alpha + \beta) \\ -\psi'(\alpha + \beta) & \psi'(\beta) - \psi'(\alpha + \beta) \end{pmatrix}. \quad (0.4)$$

This is positive definite for all  $\alpha > 0$ ,  $\beta > 0$ .

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

$$q_B(\theta) = P_\theta(X \leq x_0) = \int_0^{x_0} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} dx,$$

Therefore

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

while, with

$$\begin{aligned} \dot{q}_B(\theta) &= \begin{pmatrix} E_\theta(1_{(0,x_0]}(X) \log X) + (\psi(\alpha + \beta) - \psi(\alpha))F_\theta(x_0) \\ E_\theta(1_{(0,x_0]}(X) \log(1 - X)) + (\psi(\alpha + \beta) - \psi(\beta))F_\theta(x_0) \end{pmatrix} \\ &= \text{Cov}_\theta[(1_{[0,x_0]}(X) - F_\theta(x_0)), \dot{l}_\theta^T]. \end{aligned}$$

(iii). The scores are given by

$$\dot{l}_\theta(x) = \begin{pmatrix} \dot{l}_\alpha(x) \\ \dot{l}_\beta(x) \end{pmatrix} = \begin{pmatrix} \log(x) - (\psi(\alpha) - \psi(\alpha + \beta)) \\ \log(1 - x) - (\psi(\beta) - \psi(\alpha + \beta)) \end{pmatrix}$$

and the information matrix is as given in (0.4) Thus

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

where

$$\det(I(\theta)) = (\psi'(\alpha) - \psi'(\alpha + \beta))(\psi'(\beta) - \psi'(\alpha + \beta)) - \psi'(\alpha + \beta)^2,$$

and the efficient influence function for estimation of  $q_A$  is

$$\tilde{l}_A(x) = \dot{q}_A(\theta)^T I^{-1}(\theta) \dot{l}_\theta(x) \in \dot{\mathcal{P}}$$

and hence is a (centered) linear combination of  $\log x$  and  $\log(1-x)$ . Note that  $X - E_\theta(X) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$ , and hence the sample mean is inefficient for estimation of  $E_\theta(X)$  in this model.

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  $1_{[0,x_0]}(X) - F_\theta(x_0) \notin [\dot{l}_\theta] = \dot{\mathcal{P}}$ .

(iv) The information bound for estimation of  $q_A$  is

$$\begin{aligned} I^{-1}(P|q_A, \mathcal{P}) &= \dot{q}_A^T I^{-1}(\theta) \dot{q}_A \\ &= (\alpha + \beta)^{-4} (\beta, -\alpha) \frac{1}{\det I(\theta)} \begin{pmatrix} \psi'(\beta) - \psi'(\alpha + \beta) & \psi'(\alpha + \beta) \\ \psi'(\alpha + \beta) & \psi'(\alpha) - \psi'(\alpha + \beta) \end{pmatrix} \begin{pmatrix} \beta \\ -\alpha \end{pmatrix} \\ &< \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta - 1)} = \text{Var}_\theta(X) \end{aligned}$$

where the inequality holds since  $\psi_A(X) = (X - E_\theta(X)) \notin \dot{\mathcal{P}}$ . 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  $\tilde{l}_B = \Pi(1_{[0,x_0]}(x) - F_\theta(x_0) | \dot{\mathcal{P}})$ , 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 estimators  $\bar{X}_n$  and  $\mathbb{F}_n(x_0)$  of  $q_A(\theta) = E_\theta(X)$  and  $q_B(\theta) = F_\theta(x_0)$  when the model holds. (If we had considered  $q_C(\theta) = E_\theta \log(X/(1-X))$  or  $q_D(\theta) = E_\theta \log X$ , this story would change! It is also an instructive exercise to consider the sub-model consisting of the beta densities with  $\alpha = \beta$ .)

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  $\theta$  as  $\theta = (\nu, \eta)$  where  $\nu \in R^m$  and  $\eta \in R^{k-m}$  and  $1 \leq m < k$ . Let  $\dot{l} = \dot{l}_\theta = (\dot{l}_1, \dot{l}_2)$  be the corresponding partition of the (vector of) scores  $\dot{l}$ , and, with  $\tilde{l} \equiv I^{-1}(\theta) \dot{l}$ , the *efficient influence function* for  $\theta$ , let  $\tilde{l} = (\tilde{l}_1, \tilde{l}_2)$  be the corresponding partition of  $\tilde{l}$ . In both cases,  $\dot{l}_1, \tilde{l}_1$  are  $m$ -vectors of functions, and  $\dot{l}_2, \tilde{l}_2$  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.  $I^{11} = I_{11.2}^{-1}$  where  $I_{11.2} \equiv I_{11} - I_{12} I_{22}^{-1} I_{21}$ ,

$$\begin{aligned}
I^{22} &= I_{22 \cdot 1}^{-1} \text{ where } I_{22 \cdot 1} \equiv I_{22} - I_{21} I_{11}^{-1} I_{12}, \\
I^{12} &= -I_{11 \cdot 2}^{-1} I_{12} I_{22}^{-1}, \\
I^{21} &= -I_{22 \cdot 1}^{-1} I_{21} I_{11}^{-1}.
\end{aligned}$$

This amounts to formulas (3) and (4) of section 3.2, page 14.

B. Verify that

$$\begin{aligned}
\tilde{l}_1 &= I^{11} \dot{l}_1 + I^{12} \dot{l}_2 = I_{11 \cdot 2}^{-1} (\dot{l}_1 - I_{12} I_{22}^{-1} \dot{l}_2), \text{ and} \\
\tilde{l}_2 &= I^{21} \dot{l}_1 + I^{22} \dot{l}_2 = I_{22 \cdot 1}^{-1} (\dot{l}_2 - I_{21} I_{11}^{-1} \dot{l}_1).
\end{aligned}$$

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

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

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  $\tilde{\mathbf{l}}_1$ :

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

Similarly,

$$\begin{aligned}
\tilde{\mathbf{l}}_2 &= -I_{22 \cdot 1}^{-1} I_{21} I_{11}^{-1} \dot{\mathbf{l}}_1 + I_{22 \cdot 1}^{-1} \dot{\mathbf{l}}_2 \\
&= I_{22 \cdot 1}^{-1} (\dot{\mathbf{l}}_2 - I_{21} I_{11}^{-1} \dot{\mathbf{l}}_1) = I_{22 \cdot 1}^{-1} \mathbf{l}_2^*.
\end{aligned}$$