

## Statistics 581, Problem Set 7

Wellner; 11/8/2000

**Reading:** Chapter 3, Section 2;

Ferguson, ACLST, Chapter 19, pages 126-132, Chapter 20, pages 133-134;

Lehmann and Casella, pages 113-129, and 439- 443.

**Due:** Wednesday, November 15, 2000.

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.
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.
3. Lehmann and Casella, TPE, Problem 6.5, part (a), page 142.
4. 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  $I_{11.2} \equiv I_{11} - I_{12}I_{22}^{-1}I_{21}$ ,

$I^{22} = I_{22.1}^{-1}$  where  $I_{22.1} \equiv I_{22} - I_{21}I_{11}^{-1}I_{12}$ ,

$I^{12} = -I_{11.2}^{-1}I_{12}I_{22}^{-1}$ ,

$I^{21} = -I_{22.1}^{-1}I_{21}I_{11}^{-1}$ .

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

B. Verify that

$\tilde{l}_1 = I^{11}\dot{l}_1 + I^{12}\dot{l}_2 = I_{11.2}^{-1}(\dot{l}_1 - I_{12}I_{22}^{-1}\dot{l}_2)$ , and

$\tilde{l}_2 = I^{21}\dot{l}_1 + I^{22}\dot{l}_2 = I_{22.1}^{-1}(\dot{l}_2 - I_{21}I_{11}^{-1}\dot{l}_1)$ .

5. **Optional bonus problem:** [This is example 7.2.5 and 7.2.7 in Lehmann and Casella, TPE, section 6.2; also see problems 6.2.12 - 6.2.14, Lehmann and Casella, TPE, page 501.] Suppose that  $X_1, \dots, X_n$  are i.i.d.  $N(\theta, 1)$  so  $I(\theta) = 1$ . Let

$0 < a < 1$  and define  $T_n \equiv \bar{X}_n 1_{[|\bar{X}_n| \geq n^{-1/4}]} + a\bar{X}_n 1_{[|\bar{X}_n| < n^{-1/4}]}$ . This is Hodges superefficient estimator of  $\theta$ .

(a) Show that  $\sqrt{n}(T_n - \theta) \rightarrow_d N(0, V(\theta))$  where

$$V(\theta) = \begin{cases} 1, & \theta \neq 0 \\ a^2, & \theta = 0 \end{cases}$$

(b) Show that  $T_n$  is *not a regular estimator* of  $\theta$  at  $\theta = 0$ , but that it is regular at every  $\theta \neq 0$ . If  $\theta_n = t/\sqrt{n}$ , find the limiting distribution of  $\sqrt{n}(T_n - \theta_n)$  under  $P_{\theta_n}$ .

C. For  $\theta_n = t/\sqrt{n}$  show that

$$R_n(\theta_n) = nE_{\theta_n}(T_n - \theta_n)^2 \rightarrow E(aZ + t(a - 1))^2 = a^2 + t^2(1 - a)^2$$

where  $Z \sim N(0, 1)$ . This is *larger* than 1 if  $t^2 > (1 + a)/(1 - a)$ , and hence superefficiency also entails worse risks in a local neighborhood of the point(s) where the asymptotic variance is smaller.