Statistics 581, Problem Set 7
Wellner; 11/10/2010

Reminder: Second make-up lecture: Monday, November 15, 11:30 - 12:20, Savery 132.
Reading: Chapter 3, Section 2;
Ferguson, ACILST, Chapter 19, pages 126-132, Chapter 20, pages 133-134;
Lehmann and Casella, pages 113-129, and 439- 443.
Due: Wednesday, November 17, 2010.

1. Suppose that $X_1, \ldots, X_n$ are i.i.d. with the Weibull distribution F_θ given by
$$1 - F_\theta(x) = \exp(-(x/\alpha)^\beta), \qquad x \ge 0$$
where $\theta = (\alpha, \beta) \in (0,\infty) \times (0,\infty)$.
(a) Find the inverse (or quantile function) $F_\theta^{-1}(u)$ corresponding to F_θ in terms of α, β, 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 α and β. Prove that your estimators are consistent.
(c) Prove that your estimators $\hat\alpha_n$ and $\hat\beta_n$ satisfy
$$\sqrt{n} \begin{pmatrix} \hat\alpha_n - \alpha \\ \hat\beta_n - \beta \end{pmatrix} \to_d N_2(0, \Sigma)$$
and identify Σ as a function of α, β, and t.
(d) How would you choose r and s to minimize the asymptotic variance of $\hat\beta_n$?

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

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?

4. Suppose that $X \sim \text{Beta}(\alpha, \beta)$; i.e. X has density p_θ 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 $F_\theta(x) \equiv P_\theta(X \le x)$.
(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. Do we have either $\psi_A \in \dot{\mathcal{P}}$, while $\psi_B \notin \dot{\mathcal{P}}$?

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 $\dot{l} = \dot{l}_\theta = (\dot{l}_1, \dot{l}_2)$ 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 θ, 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. $I^{11} = I_{11\cdot 2}^{-1}$ 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 (3) and (4) of section 3.2, page 14.
B. Verify that
$\tilde{l}_1 = I^{11}\dot{l}_1 + I^{12}\dot{l}_2 = I_{11\cdot 2}^{-1}(\dot{l}_1 - I_{12}I_{22}^{-1}\dot{l}_2)$, and
$\tilde{l}_2 = I^{21}\dot{l}_1 + I^{22}\dot{l}_2 = I_{22\cdot 1}^{-1}(\dot{l}_2 - I_{21}I_{11}^{-1}\dot{l}_1)$.

6. **Optional bonus problem 1:** Suppose that Y has the Beta(α, β) density as in problem 4 above. Now let $W \equiv \log(Y/(1-Y))$.
(a) Find the density $f = f_{\alpha,\beta}$ of W and relate it to the standard logistic density in problem 2. Plot several of these densities for $\alpha = \beta > 1$ and $\alpha \neq \beta > 1$. (b) Now suppose that $X \equiv \sigma W + \mu$ where W has density $f_{\alpha,\beta}$ as in (a). Find the density $g_{\alpha,\beta,\mu,\sigma}$ of X. Is there any natural reparametrization of the family $\mathcal{P} \equiv \{g_{\alpha,\beta,\mu,\sigma} : \alpha > 0, \beta > 0, \mu \in \mathbb{R}, \sigma > 0\}$ in terms of fewer parameters?
(c) What is the information matrix for the model $\mathcal{P}$ in (b)?

7. **Optional bonus problem 2:** Suppose that we want to model the survival of twins with a common genetic defect, but with one of the two twins receiving some treatment. Let X represent the survival time of the untreated twin and let Y represent the survival time of the treated twin. One (overly simple) preliminary model might be to assume that X and Y are independent with Exponential(η) and Exponential($\theta\eta$) distributions, respectively:

$$f_{\theta,\eta}(x,y) = \eta e^{-\eta x} \eta\theta e^{-\eta\theta y} 1_{(0,\infty)}(x) 1_{(0,\infty)}(y)$$

Compute the Cramér-Rao lower bound for unbiased estimates of θ based on $Z = X/Y$, the maximal invariant for the group of scale changes $g(x,y) = (cx, cy)$ with $c > 0$. Compared this bound to the information bounds for estimation of θ based on observation of (X,Y) when η is known and unknown.

8. **Optional bonus problem 3:** A more realistic model involves assuming that the common parameter η for the two twins varies across sets of twins. There are several different ways of modeling this: one approach involves supposing that each pair of twins observed (X_i, Y_i) has its own fixed parameter η_i, $i = 1, \ldots, n$. In this model we observe (X_i, Y_i) with density f_{θ,η_i} for $i = 1, \ldots, n$; i.e.

$$f_{\theta,\eta_i}(x_i, y_i) = \eta_i e^{-\eta_i x_i} \eta_i \theta e^{-\eta_i \theta y_i} 1_{(0,\infty)}(x_i) 1_{(0,\infty)}(y_i) \,. \tag{0.1}$$

This is sometimes called a "functional model" (or model with incidental nuisance parameters).

Another approach is to assume that $\eta \equiv Z$ has a distribution, and that our observations are from the mixture distribution. Assuming (for simplicity) that $Z = \eta \sim \text{Gamma}(a,b)$ with density $g_{a,b}(\eta)$, it follows that the (marginal) distribution of (X,Y) is

$$\begin{aligned} p_{\theta,a,b}(x,y) &= \int_0^\infty f_{\theta,z}(x,y) g_{a,b}(z) dz \\ &= \frac{\theta}{b^2} \left(\frac{b}{b + x + \theta y} \right)^{a+2} \frac{\Gamma(a+2)}{\Gamma(a)} \,. \end{aligned} \tag{0.2}$$

This is sometimes called a "structural model" (or mixture model).
(a) Find the information for θ in the functional model.
(b) Find the information for θ in the structural model.
(c) Compare the information bounds you computed in (a) and (b). When is the information for θ in the functional model larger than the information for θ in the structural model?