Statistics 581, Problem Set 8

Wellner; 11/12/2008

Reading: Chapter 3, Sections 3-5;
Ferguson, ACLST, Chapter 20, pages 133-139; Chapter 22, pages 144-150;
Lehmann and Casella, Chapter 6, especially sections 6.1-6.2, pages 429-443.
Due: Wednesday, November 19, 2008.

1. Suppose that $\theta = (\theta_1, \theta_2) \in \Theta \subset R^k$ where $\theta_1 \in R$ and $\theta_2 \in R^{k-1}$. Show that:
   A. $\mathbf{l}_1^* = \dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathbf{l}}_2$ is orthogonal to $[\dot{\mathbf{l}}_2] \equiv \{a' \dot{\mathbf{l}}_2 : a \in R^{k-1}\}$ in $L_2(P_\theta)$.
   B. $I_{11 \cdot 2} = \inf_{c \in R^{k-1}} E_\theta (\dot{\mathbf{l}}_1 - c' \dot{\mathbf{l}}_2)^2$ and that the minimum is achieved when $c' = I_{12} I_{22}^{-1}$. Thus
   $$I_{11 \cdot 2} = E_\theta (\dot{\mathbf{l}}_1 - I_{12} I_{22}^{-1} \dot{\mathbf{l}}_2)^2 = E_\theta [(\mathbf{l}_\theta^*)^2] \, .$$
   C. Prove the formulas (16) and (17) on page 21 of the Chapter 3 notes and interpret these formulas geometrically.

2. Suppose that $(Y|Z) \sim \text{Weibull}(\lambda^{-1} e^{-\gamma Z}, \beta)$, and $Z \sim G_\eta$ on R with density g_η with respect to some dominating measure μ. Thus the conditional cumulative hazard function $\Lambda(t|z)$ is given by
   $$\Lambda_{\gamma,\lambda,\beta}(t|z) = (\lambda e^{\gamma Z} t)^\beta = \lambda^\beta e^{\beta\gamma Z} t^\beta$$
   and hence
   $$\lambda_{\gamma,\lambda,\beta}(t|z) = \lambda^\beta e^{\beta\gamma Z} \beta t^{\beta - 1} \, .$$
   (Recall that $\lambda(t) = f(t)/(1 - F(t))$ and
   $$\Lambda(t) \equiv \int_0^t \lambda(s) ds = \int_0^t (1 - F(s))^{-1} dF(s) = -\log(1 - F(t))$$
   if F is continuous.) Thus it makes sense to reparametrize by defining $\theta_1 \equiv \beta\gamma$ (this is the parameter of interest since it reflects the effect of the covariate Z), $\theta_2 \equiv \lambda^\beta$, and $\theta_3 \equiv \beta$. This yields
   $$\lambda_\theta(t|z) = \theta_3 \theta_2 \exp(\theta_1 z) t^{\theta_3 - 1}$$
   You may assume that
   $$a(z) \equiv (\partial/\partial\eta) \log g_\eta(z)$$
   exists and $E\{a^2(Z)\} < \infty$. Thus Z is a "covariate" or "predictor variable", θ_1 is a "regression parameter" which affects the intensity of the (conditionally) Exponential variable Y, and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ where $\theta_4 \equiv \eta$.
   (a) Derive the joint density $p_\theta(y, z)$ of (Y, Z) for the re-parametrized model.
   (b) Find the information matrix for θ. What does the structure of this matrix say about the effect of $\eta = \theta_4$ being known or unknown about the estimation of $\theta_1, \theta_2, \theta_3$?
   (c) Find the information and information bound for θ_1 if the parameters θ_2 and θ_3 are known?
   (d) What is the information bound for θ_1 if just θ_3 is known to be equal to 1?

3. This is a continuation of the previous problem:
(e) Find the efficient score function and the efficient influence function for estimation of θ_1 when θ_3 is known.
(f) Find the information $I_{11\cdot(2,3)}$ and information bound for θ_1 if the parameters θ_2 and θ_3 are unknown. (Here both θ_2 and θ_3 are in "the second block".)
(g) Find the efficient score function and the efficient influence function for estimation of θ_1 when θ_2 and θ_3 are unknown.
(h) Specialize the calculations in (d) - (g) to the case when $Z \sim \text{Bernoulli}(\theta_4)$ and compare the information bounds.

4. (a) Lehmann and Casella, Problem 2.13, page 501.
(b) Let $R_n(\theta) \equiv nE_\theta(T_n - \theta)^2$ where T_n is the Hodges superefficient estimator as in Example 3.3.1 (so $T_n = \delta_n$ of Example 2.5, Lehmann and Casella pages 440 - 443). Show that $R_n(n^{-1/4}) \to \infty$ as $n \to \infty$.

5. Suppose that $Z \sim N(0,1)$ and, for $\mu \in R$ and $\sigma > 0$, that $X = \mu + \sigma Z \sim P_{\mu,\sigma} = N(\mu, \sigma^2)$.
(a) Compute the likelihood ratio
$$\frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(x) = \frac{\sigma^{-1}\phi((x-\mu)/\sigma)}{\sigma^{-1}\phi(x/\sigma)} \qquad \text{and} \qquad Y \equiv \log \frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X)\,.$$
What is the distribution of Y under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$?
(b) Plot the function $l(\mu; X) \equiv \log(dP_{\mu,\sigma}/dP_{0,\sigma})(X)$ as a function of μ.
(c) Find the maximum value of the function $l(\mu; X)$ in B (as a function of μ) and the value of $\mu \equiv \hat{\mu}$ which achieves the maximum.
(d) What is the distribution of $\hat{\mu}$ under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$? What is the distribution of $l(\hat{\mu}; X)$ under $P_{0,\sigma}$ and under $P_{\mu,\sigma}$?

6. **Optional bonus problem:** (Generalization of problem 5, problem set #6). 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 $n^{-1}\log(L_n(\theta_1)/L_n(\theta_0) \to_p$ some constant, and identify the constant explicitly in terms of θ_1, θ_0.

7. **Optional bonus problem:** Read Note 8.5, Lehmann and Casella, page 145. Explore the identity in the second display in this note and see if it makes sense as written. If not, rewrite the identity in a way that makes sense to you. [Compare with Efron and Johnstone (1990) and/or Bickel, Klaassen, Ritov, and Wellner (1993), pages 420-424.]