Statistics 581, Problem Set 8

Wellner; 11/14/2018

Reading: Chapter 3, Sections 3-5; start reading Chapter 4 (to be handed out on Monday, November 20);
Ferguson, ACILST, Chapter 20, pages 133-139; Chapter 22, pages 144-150;
vdV, Asymp. Statist., pages 85 - 107; Sections 6.1 - 6.2; 7.1 - 7.6.

Due: Wednesday, November 21, 2018.

1. (a) Show that if $\theta_n = cn^{-1/2}$ and T_n is the Hodges super-efficient estimator discussed in class, then the sequence $\{\sqrt{n}(T_n - \theta_n)\}$ is uniformly square-integrable.
(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 $n \to \infty$.

2. (Super-efficiency at two parameter values) Suppose that $X_1, \ldots, X_n$ are i.i.d. $N(\theta, 1)$ where $\theta \in \mathbb{R}$) Let $a, b \in [0, 1)$ and define the estimator T_n as follows:
$$T_n = \begin{cases} \overline{X}_n & \text{if } |\overline{X}_n - 1| > n^{-1/4} \ \text{ and } \ |\overline{X}_n + 1| > n^{-1/4}, \\ a\overline{X}_n + (1-a) & \text{if } |\overline{X}_n - 1| \le n^{-1/4}, \\ b\overline{X}_n + (1-b)(-1) & \text{if } |\overline{X}_n + 1| \le n^{-1/4}. \end{cases}$$
(a) Find the limiting distribution of $\sqrt{n}(T_n - \theta)$ when:
(i) $\theta \neq 1$ and $\theta \neq -1$; (ii) $\theta = 1$; (iii) $\theta = -1$.
(b) Find the limiting distribution of $\sqrt{n}(T_n - \theta_n)$ when:
(i) $\theta_n = 1 + cn^{-1/2}$; (ii) $\theta_n = -1 + cn^{-1/2}$.
(c) Could we have super-efficiency at a countable collection of parameter values?

3. Suppose that $X_1, \ldots, X_n$ are i.i.d. with distribution function F having a continuous density function f. Let $\mathbb{F}_n$ be the empirical distribution function of the X_i's, suppose that b_n is a sequence of positive numbers, and let
$$\hat{f}_n(x) = \frac{\mathbb{F}_n(x + b_n) - \mathbb{F}_n(x - b_n)}{2b_n}.$$
(a) Compute $E\{\hat{f}_n(x)\}$ and $Var(\hat{f}_n(x))$.
(b) Show that $E\hat{f}_n(x) \to f(x)$ if $b_n \to 0$.
(c) Show that $Var(\hat{f}_n(x)) \to 0$ if $b_n \to 0$ and $nb_n \to \infty$.
(d) Use some appropriate central limit theorem to show that (perhaps under some suitable further conditions that you might need to specify)
$$\sqrt{2nb_n}(\hat{f}_n(x) - E\hat{f}_n(x)) \to_d N(0, f(x)).$$
Hint: Write $\hat{f}_n(x)$ in terms of some Bernoulli random variables and identify $p = p_n$.

4. Suppose that $(T|Z) \sim \text{Weibull}(\lambda^{-1}e^{-\gamma Z}, \beta)$, 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 re-parametrize 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) Weibull variable T, and $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ where $\theta_4 \equiv \eta$.

(a) Derive the joint density $p_\theta(t, z)$ of (T, 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?

(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.

5. **Optional bonus problem 1:** Lehmann and Casella, Problem 2.13, page 501.
Let $b_n(\theta) = E_\theta(T_n) - \theta$ be the bias of Hodges estimator T_n.

(a) Show that

$$b_n(\theta) = \frac{-(1-a)}{\sqrt{n}} \int_{-n^{1/4}}^{n^{1/4}} x\phi(x - \sqrt{n}\theta) dx.$$

(b) Show that $b_n'(\theta) \to 0$ for any $\theta \neq 0$ and $b_n'(0) \to 1 - \alpha$.

(c) Use (b) to explain how the Hodges estimator T_n can violate $V^2(\theta)$ without violating (Cramér-Rao) information inequality.

6. **Optional bonus problem 2:** Suppose that $X_1, \ldots, X_n$ are i.i.d. F on $\mathbb{R}$, and let $\mathbb{F}_n$ denote the empirical d.f. of the X_i's. Let Φ denote the standard normal distribution function, $\Phi(x) = \int_{-\infty}^x \phi(y)dy$ where $\phi(y) = (2\pi)^{-1/2}\exp(-y^2/2)$ is the standard normal density. Let $0 < a < 1$ and define a new estimator $\widetilde{F}_n$ of F by

$$\widetilde{F}_n(x) = \begin{cases} (1-a)\Phi(x) + a\mathbb{F}_n(x), & \text{if } \|\mathbb{F}_n - \Phi\|_\infty \le n^{-1/4}, \\ \mathbb{F}_n(x), & \text{if } \|\mathbb{F}_n - \Phi\|_\infty > n^{-1/4}. \end{cases}$$

(a) Find the limiting distribution of the process $\{\sqrt{n}(\widetilde{F}_n(x) - F(x)) : x \in \mathbb{R}\}$ when $F = \Phi$.
(b) Find the limiting distribution of the process $\{\sqrt{n}(\widetilde{F}_n(x) - F(x)) : x \in \mathbb{R}\}$ when $F \ne \Phi$.
(c) Show that $\widetilde{F}_n$ is not a regular estimator of F at $F = \Phi$ (in an appropriate sense to be defined), but that F is a regular estimator of F at any $F \ne \Phi$.

7. **Optional bonus problem 3:** This is a continuation of problem 3 above.
(a) Suppose that f is differentiable at x. What further assumptions on f' or f'' and b_n do you need to show that $\sqrt{nb_n}\left\{E(\hat{f}_n(x)) - f(x)\right\} \to 0$? (This says that the bias of $\hat{f}_n$ for estimating x is $o((nb_n)^{-1/2})$.)
(b) Combine the conclusion of (a) with the conclusion of problem 3(d) to find the limiting distribution of $\sqrt{2nb_n}(\hat{f}_n(x) - f(x))$.
(c) Now suppose that $x, y \in \mathbb{R}$ satisfy $x \ne y$, $f(x) > 0$ and $f(y) > 0$. Find the limiting joint distribution of

$$\sqrt{2nb_n}\begin{pmatrix} \hat{f}_n(x) - f(x) \\ \hat{f}_n(y) - f(y) \end{pmatrix}$$

under appropriate further hypotheses on derivatives of f at x and y and the sequence b_n.
(d) Compare the result in (c) with the joint limiting distribution of

$$\sqrt{n}\begin{pmatrix} \mathbb{F}_n(x) - F(x) \\ \mathbb{F}_n(y) - F(y) \end{pmatrix}$$

obtained in Chapter 2 (what is it explicitly?).
(e) Let k be a (bounded) density on $\mathbb{R}$ with mean 0 and finite variance. A *kernel density estimator* $\hat{f}_n$ with kernel k and bandwidth b_n is given by

$$\hat{f}_n(x) = \int_{-\infty}^{\infty} \frac{1}{b_n} k\left(\frac{x-y}{b_n}\right) d\mathbb{F}_n(y)$$

where k is a (bounded) density on $\mathbb{R}$ with mean 0 and finite variance. The estimator $\hat{f}_n$ in (a) - (d) of problem 3 is the special case of this with $k(x) = 2^{-1}1_{[-1,1]}(x)$, the uniform density on $[-1, 1]$. How does the conclusion of 3(d) change for a general kernel k? Explain why.