Statistics 581, Problem Set 6

Wellner; 11/1/2006

Reminder: Midterm exam: Monday, November 6.
Reading: Lecture Notes Chapter 3, sections 1-2;
Ferguson, ACILST, chapters 19-20, pages 126 - 139;
Lehmann and Casella, TPE, Sections 2.5 and 2.6, pages 113 - 129;
and Section 6.2, pages 437 - 443.
Due: Wednesday, November 15, 2006.

1. Chapter 2, Exercise 5.3, page 25. [Hint: use the fact that $\mathbb{S}_n(t_j) - \mathbb{S}_n(t_{j-1}) = n^{-1/2}\sum_{i=[nt_{j-1}]+1}^{[nt_j]} X_i$, $j = 1, \ldots, t_k$ with $t_0 \equiv 0$ are independent random variables.]

2. Consider a function $T : \mathcal{F} \to \mathbb{R}$ where $\mathcal{F}$ is some (sub) class of distribution functions F (examples include the mean, $T(F) = \mu(F) = \int x dF(x)$, the variance $T(F) = \sigma^2(F) = \int (x - \int y dF(y))^2 dF(x)$, the median $T(F) = F^{-1}(1/2)$, linear combinations of order statistics $T(F) = \int_0^1 F^{-1}(u)w(u)du$, the *mean residual life function* at $x > 0$ $T(F) \equiv e(x, F) \equiv \int_{(x,\infty)}(1 - F(u))du/(1 - F(x)) = E(X - x|X > x)$, and so forth). [The mean residual life function gives the mean life conditional on surviving beyond x.] The "principle of substitution" says that $T(F)$ can be estimated by $T(\widehat{F}_n)$. for some estimator $\widehat{F}_n$ of F. If T is sufficiently "smooth", then frequently the empirical distribution function $\mathbb{F}_n$ can be taken as the estimator $\hat{F}_n$ of F.

   Give a treatment of consistency and asymptotic normality of the estimator $e(x, \mathbb{F}_n)$ of $e(x, F)$ based on our results from sections 2.4 and 2.6. You may assume that with $X \sim F$ on $(0, \infty)$ we have $E_F X < \infty$, $E_F X^2 < \infty$, and $1 - F(x) > 0$ (as well as any other additional assumptions you need).

3. 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
$$l(\mu; X) \equiv \log\frac{dP_{\mu,\sigma}}{dP_{0,\sigma}}(X)$$
   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 $P_{\mu,\sigma}$?

4. Suppose that $X, X_1, X_2, \ldots, X_n$ are independent Exponential(λ) random variables:
$$P(X \geq x) = \exp(-\lambda x), \qquad x > 0.$$

(a) Show that the r-th moment of X, $\mu_r \equiv \mu_r(\lambda)$ is given by

$$\mu_r(\lambda) = EX^r = \frac{\Gamma(r+1)}{\lambda^r}.$$

(b) Use the moment calculation in (a) to show that

$$\frac{\mu_r(\lambda)}{\mu_{r+1}(\lambda)} = \frac{\lambda}{r+1}$$

and hence that the family of estimators $\{\hat{\lambda}_n^{(k)}\}_{k \geq 0}$ given by

$$\hat{\lambda}_n^{(k)} \equiv (k+1)\frac{\overline{X^k}_n}{\overline{X^{k+1}}_n} \equiv (k+1)\frac{n^{-1}\sum_1^n X_i^k}{n^{-1}\sum_1^n X_i^{k+1}}$$

are all consistent estimators of λ: $\hat{\lambda}_n^{(k)} \rightarrow_p \lambda$ for each $k = 0, 1, 2, \ldots$.
(c) Show that

$$\sqrt{n}(\hat{\lambda}_n^{(k)} - \lambda) \rightarrow_d N(0, \sigma_k^2(\lambda)) \quad \text{as} \quad n \rightarrow \infty$$

and compute $\sigma_k^2(\lambda)$ explicitly as a function of k and λ.
(c) What is the asymptotic relative efficiency of $\hat{\lambda}_n^{(k)}$ to $\widehat{\lambda}_n \equiv \hat{\lambda}_n^{(0)} = 1/\overline{X}_n$ for $k > 1$?

5. Problem 4, page 132, Ferguson, AFCILST, modified slightly.
Suppose that $X_1, \ldots, X_n$ are independent random variables with $X_i \sim$ Poisson$(\exp(\theta z_i))$ for $i = 1, \ldots, n$ where $z_1, \ldots, z_n$ are known real numbers.
(a) Find the Cramér - Rao bound for the variance of an unbiased estimator of θ based on $X_1, \ldots, X_n$.
(b) Find the Cramér - Rao bound of an unbiased estimator of $q(\theta) = P_\theta(X < 3)$.
(c) Repeat (a) and (b) for the case in which we observe (X_i, Z_i) for $i = 1, \ldots, n$ where $(X_i|Z_i) \sim$ Poiss$(\exp(\theta Z_i))$, and $Z_i \sim G$ on R with density g where g is unknown. Compare the bounds for this model with those computed in (a) and (b).