Statistics 581, Problem Set 10

Wellner; 12/05/2001

Reading: Chapter 4, Sections 1-4;
Ferguson, ACLST, Chapter 20, pages 133-139; Chapter 22, pages 144-150;
Lehmann and Casella, Chapter 6, especially section 6.5, pages 461-468.
Due: Wednesday, December 12, 2001.

1. Ferguson, ACLST, page 139, problem 3.

2. Ferguson, ACLST, page 118, problem 3. [Also see Lehmann and Casella, Example 7.9, page 482.]

3. Ferguson, ACLST, page 149, problem 2 modified as follows:
   (a) Find the LR test statistic of the null hypothesis $H_0 : \mu = c\theta$ for any fixed number $c > 0$, and find the asymptotic distribution of the LR statistic under H_0.
   (b) Does the theory of our chapter 4 (or Ferguson's chapter 22 apply directly?
   (c) Does the local asymptotic power of your test depend on c?

4. (Right censoring – again.) Consider nonparametric maximum likelihood estimation of F in the censored data problem considered in class, but extend the argument to include ties as follows:
   A. When there are ties, let the distinct Z's be denoted by $T_1 < \ldots < T_k$. Let $m_1, \ldots, m_k$ and $n_1, \ldots, n_k$ be defined by $m_j \equiv \#$ of $Z_i\delta_i = T_j$, $n_j \equiv \#$ of $Z_i(1-\delta_i) = T_j$, and let $p_j \equiv \Delta F(T_j) \equiv F(T_j) - F(T_j-)$, $= 1, \ldots, k$, $p_{k+1} = 1 - F(T_k)$. Show that the likelihood (for F) is

$$L(F|\underline{Z}, \underline{\delta}) = \prod_{i=1}^{k} p_i^{m_i} (\sum_{j=i+1}^{k+1} p_j)^{n_i} .$$

   B. By defining $a_i \equiv p_i / \sum_{j=i}^{k+1} p_j$ for $i = 1, \ldots, k$ and $a_{k+1} = 1$, and rewriting the likelihood in terms of the a_i's, show that the likelihood is maximized by

$$\hat{a}_i = m_i / \sum_{j=i}^{k} (m_j + n_j) = n\Delta\mathbb{H}_n^{uc}(T_i)/n(1 - \mathbb{H}_n(T_i-)),$$

   and hence that the nonparametric MLE of F is (again) the Kaplan - Meier estimator

$$1 - \hat{\mathbb{F}}_n(t) = \prod_{0 \le s \le t} (1 - \Delta\hat{\Lambda}_n(s)).$$

   C. Compute $1 - \hat{\mathbb{F}}_n$ for the following data (length of time until complete remission in weeks for the "maintained group") from a study of the efficacy of chemotherapy for acute Myelogenous leukemia (AML):

$$9, 13, 13+, 18, 23, 28+, 31, 31, 34, 45+, 48, 161+;$$

   here "$+$" indicates censoring ($\delta = 0$).

5. **Bonus problem 1.**
   Ferguson, ACLST, page 150, problem 3. Does the theory in our chapter 4 (or Ferguson's chapter 22) apply directly?

6. **Bonus Problem 2.**
   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.
   Consider testing the hypothesis $H : \gamma = 0$ versus $K : \gamma \neq 0$. (Note that the null hypothesis is *not simple*, but *composite*; the values of λ and η are not specified by the hypothesis H.)
   (a) Propose three different test statistics for testing H versus K, and briefly discuss how you would compute them.
   (b) Do our results in Chapter 4 apply to the (asymptotic) distribution under H of the test statistics you proposed in (a)?
   (c) Consider local alternatives of the form $\gamma_n = tn^{-1/2}$ for $t \in R$ fixed. Give an expression for the local asymptotic power of the tests you proposed in (a) for these alternatives.
   (d) Suppose that $\gamma_1 \neq 0$ is the "true" value of the parameter γ. Show that the test statistics you proposed in (a), when appropriately normalized, converge in probability to positive constants, and identify these constants as explicitly as possible.