Statistics 581
Problem Set 5
Wellner; 10/25/00

Reading: Ferguson, ACLST, Chapters 13 and 14, pages 87 - 100; Chapter 2, sections 3 and 4.
Due: Wednesday, November 1, 2000.

1. Suppose that $Y_i = \alpha + \theta'(x_i - \overline{x}) + \epsilon_i$, $i = 1, \ldots, n$, where $\epsilon_i \sim (0, \sigma^2)$ are i.i.d. and the x_i's are known vectors in R^k. Equivalently, $\underline{Y} = X\underline{\beta} + \underline{\epsilon}$ where
$$X^T = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 - \overline{x} & x_2 - \overline{x} & \cdots & x_n - \overline{x} \end{pmatrix}$$
so that X is an $n \times (k+1)$ matrix. Let $\hat{\underline{\beta}}$ be the least squares estimator of $\underline{\beta} = (\alpha, \theta')'$; i.e. $\hat{\underline{\beta}} = (X^T X)^{-1} X^T \underline{Y}$. Suppose that $n^{-1}(X^T X) \rightarrow D$ where D is positive definite.
(a) What additional condition(s) do you need to impose to prove that
$$\sqrt{n}(\hat{\beta}_n - \beta) \rightarrow_d N_{k+1}(0, \text{“}something\text{”})\,?$$
(b) Find “something” in part (a).

2. Suppose that $X_1, \ldots, X_n$ are i.i.d. Cauchy$(0,1)$; so the density of each X_i with respect to Lebesgue measure on R is $f(x) = \pi^{-1}(1 + x^2)^{-1}$, $x \in R$.
(a) Compute the distribution function F of the X_i's.
(b) Compute and plot the inverse distribution function F^{-1} corresponding to F.
(c) For what values of $r > 0$ is $E|X_1|^r < \infty$?
(d) Find the distribution function of $M_n \equiv \max_{1 \le i \le n} X_i$.
(e) For what values of r is $E|M_n|^r < \infty$?
(f) Find a sequence of constants b_n so that $M_n / b_n \rightarrow_d$ and find the limiting distribution. [Hint: see Ferguson, ACLST, Theorem 14, page 95.]

3. Suppose that $X_1, \ldots, X_n$ are i.i.d. with continuous distribution function F. Let F_0 be a fixed, specified distribution function. Suppose we

want to test $H: \; F = F_0$ versus $K: \; F \neq F_0$. Consider the *Cramér - von Mises statistic* given by

$$C_n^2 \equiv \int_{-\infty}^{\infty} n(\mathbb{F}_n(x) - F_0(x))^2 dF_0(x) \, .$$

(a) Show that if $F = F_0$ is true, then

$$C_n^2 =_d \int_0^1 n(\mathbb{G}_n(t) - t)^2 dt \, ,$$

where $\mathbb{G}_n$ is the empirical d.f. of n i.i.d. Uniform$(0,1)$ rv's.
(b) Show that when the null hypothesis is true,

$$C_n^2 \rightarrow_d \int_0^1 \mathbb{U}(t)^2 \, dt$$

where $\mathbb{U}$ is a standard Brownian bridge process.
[Hint: Use the fact that $\mathbb{U}_n \Rightarrow \mathbb{U}$ in $(D[0,1], \| \cdot \|_\infty)$ and the continuous mapping theorem.]
(c) Suppose that the null hypothesis fails. Thus $F \neq F_0$. Show that in this case

$$n^{-1} C_n^2 \rightarrow_{a.s.} \int_{-\infty}^{\infty} (F(x) - F_0(x))^2 dF_0(x) > 0 \, ,$$

and hence the test based on C_n^2 is consistent for all $F \neq F_0$.
(d) Suppose that $F = F_n$ satisfies $\|F_n - F_0\|_\infty \rightarrow 0$ and $\sqrt{n}(F_n(x) - F_0(x)) \rightarrow g(x)$ in $L_2(F_0)$; i.e.

$$\int [\sqrt{n}(F_n(x) - F_0(x)) - g(x)]^2 dF_0(x) \rightarrow 0 \, .$$

Describe the limiting distribution of C_n^2 under the local alternatives F_n in terms of a Brownian bridge process $\mathbb{U}$ and g.

4. Suppose that $X_1, \ldots, X_n$ are i.i.d. with the Weibull distribution F_θ given by

$$1 - F_\theta(x) = \exp(-(x/\alpha)^\beta), \qquad x \geq 0$$

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 $t \in (0, 1/2)$. Use the $t-$th and $(1-t)-$th quantiles of the X_i's, namely $\mathbb{F}_n^{-1}(t)$ and $\mathbb{F}_n^{-1}(1-t)$, 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 t to minimize the asymptotic variance of $\hat{\beta}_n$?

5. **Optional bonus problem:** Let c^2 denote the constant on the right side in Problem 3(c) above. In the set-up of that problem, show that when $F \neq F_0$ it follows that

$$\sqrt{n}(n^{-1} C_n^2 - c^2) \to_d N(0, V^2)$$

and find V^2.
[Hint: Use $\sqrt{n}(\mathbb{F}_n - F) =_d \mathbb{U}_n(F)$, $\mathbb{U}_n \Rightarrow \mathbb{U}$, and the continuous mapping theorem.]