Statistics 581
Problem Set 5
Wellner; 10/24/2018

Reading: Course Notes, Chapter 2, pages 30-40;
Ferguson, ACLST, Chapters 13 and 14, pages 87 - 100;
vdVaart, Asymp. Stat., Sections 21.1-21.2, pages 304-310.

Due: Wednesday, October 31, 2018.

Reminder: Midterm exam, Friday, November 2, 2018

1. van der Vaart, problem 3.8, page 34, modified. Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(p) with $0 < p < 1$.
(a) Find the limit distribution of $\sqrt{n}(\overline{X}_n^{-1} - p^{-1})$.
(b) Show that $E|\overline{X}_n^{-1}| = \infty$ for every n.
(c) Connect the example in (a) to a result in the 581 Course Notes, Section 2.4.

2. van der Vaart, problem 3.6, page 34: Let $X_1, \ldots, X_n$ be i.i.d. with expectation μ and variance 1. Find constants a_n and $b_n$such that $a_n(\overline{X}_n^2 - b_n)$ converges in distribution when $\mu = 0$ or $\mu \neq 0$.

3. van der Vaart, problem 19.4, page 290: Suppose that $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are independent samples from distribution functions F and G respectively. The Kolmogorov-Smirnov statistic for testing the null hypothesis $H: \ F = G$ versus $K: \ F \neq G$ is the supremum distance $K_{m,n} \equiv \|\mathbb{F}_m - \mathbb{G}_n\|_\infty$ between the empirical distributions of the two samples.
(a) Find the limiting distribution of $\sqrt{mn/N} K_{m,n}$ under the null hypothesis. Do this first assuming that $\lambda_N \equiv m/N \equiv m/(m+n) \to \lambda \in [0,1]$ as $m \wedge n \to \infty$. What can you say if the latter hypothesis is dropped?
(b) Show that the Kolmogorov - Smirnov test is asymptotically consistent against every alternative $F \neq G$.
(c) Find the asymptotic power function as a function of (Δ_F, Δ_G) for alternatives (F_m, G_n) where $\{F_m\}$ and $\{G_n\}$ satisfy, much as in our discussion in class on 26 October, $\|F_m - F_0\}_\infty \to 0$, $\|G_n - F_0\|_\infty \to 0$ and, for functions $\Delta_F, \Delta_G : [0,1] \to \mathbb{R}$, $\|\sqrt{m}(F_m - F_0) - \Delta_F(F_0)\|_\infty \to 0$ and $\|\sqrt{n}(G_n - F_0) - \Delta_G(F_0)\|_\infty \to 0$.

4. 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 $M_n \equiv \max_{1 \leq i \leq n} X_i$.
(e) For what values of r is $E|M_n|^r < \infty$?
(f) Find a sequence of constants b_n so that $M_n/b_n \ \to_d$ and find the limiting distribution. [Hint: see Ferguson, ACLST, Theorem 14, page 95.]
(g) Find the densities of M_n/b_n with b_n as in (f). Do these densities converge pointwise to a limit density? If so, what can you conclude from Scheffé's theorem?

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

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

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 $\hat{\beta}_n$?

6. **Optional bonus problem 1:** Suppose that $X_1, \ldots, X_n$ are i.i.d. F on R. Let $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^n 1_{(-\infty,x]}(X_i)$ be the empirical d.f. of the sample. Let $\alpha \in (0,1)$. The goal of the following problem is to find a number c so that

$$P_F\{c\mathbb{F}_n(x) \le F(x) \text{ for all } -\infty < x < \infty\} = 1 - \alpha \,, \tag{1}$$

i.e. so that $c\mathbb{F}_n(x)$ is a lower $1-\alpha$ confidence bound for F if F. Let

$$A_n(c, F) = \{\mathbb{F}_n(x) \le F(x)/c \text{ for all } -\infty < x < \infty\} .$$

(a) Show that $P_F(A_n(c, F)) = P(B_n(c)$ where, for $\mathbb{G}_n$ the empirical distribution function of $\xi_1, \ldots, \xi_n$ i.i.d. Uniform$(0,1)$ random variables,

$$B_n(c) = \{\mathbb{G}_n(x) \le x/c \text{ for all } 0 < x \le 1\} .$$

(b) Re-express the event $B_n(c)$ in terms of the order statistics $0 \le \xi_{(1)} \le \ldots \le \xi_{(n)} \le 1$ of the Uniform$(0,1)$ sample. [Hint: draw a picture first!]
(c) Compute $P(B_n(c))$ using the re-expression of the event $B_n(c)$ you found in (b) and the joint density of the uniform order statistics for $n = 1$, $n = 2$, and $n = 3$.
(d) Extend the calculations in (b) to a general n.
(e) Find c explicitly as a function of α and give the resulting lower confidence bound.

7. **Optional bonus problem 2:** van der Vaart, problem 19.10, page 290, modified. Suppose that $X_1, \ldots, X_n$ are i.i.d. F with $E|X_1|^2 < \infty$ and such that F has a density f satisfying $f(F^{-1}(1/2)) > 0$.
(a) Find the asymptotic joint distribution of $\overline{X}_n$ and $\mathbb{F}_n^{-1}(1/2)$. Hint: use asymptotic linearity.
(b) Find the asymptotic distribution of the mean absolute deviation from the median: i.e. with $m(F) \equiv F^{-1}(1/2)$ and $A(F) \equiv E|X - m(F)|$, find the limiting distribution of $\sqrt{n}(A(\mathbb{F}_n) - A(F))$.