Statistics 210B, Spring 1998
Class Notes

P.B. Stark

stark@stat.berkeley.edu

www.stat.berkeley.edu/∼stark/index.html

January 27, 1999

Eigth Set of Notes

1 Application of linear programming

Density estimation with shape restrictions. (See Hengartner and Stark, 1995. Finite-sample confidence envelopes for shape-restricted densities, *Ann. Stat., 23*, 525-550.) Suppose we observe $\mathbf{X} = \{X_j\}_{j=1}^n$ i.i.d. F, where F is a distribution with a density f w.r.t. Lebesgue measure, $\mathrm{supp}\{F\} \subset \mathbf{R}^+$, and f is monotone decreasing on $\mathbf{R}^+$. We seek a confidence interval for $f(x_0)$ for some $x_0 \geq 0$. Let F denote not only the distribution (measure), but also the cdf of the measure: $F(x) = F((-\infty, x])$. Let $\mathcal{P}$ be the set of probability measures on $\mathbf{R}^+$, and let $\mathcal{Q}$ be the set of subprobability measures on $\mathbf{R}^+$. For any measure $G \in \mathcal{Q}$, define the Kolmogorov-Smirnov (K-S) norm

$$\|G\| \equiv \sup_{x \in \mathbf{R}} |G((-\infty, x])| = \sup_{x \in \mathbf{R}} |G(x)|. \tag{1}$$

Let $\hat{F}_n$ be the empirical measure correspoonding to the cdf

$$\hat{F}_n(x) \equiv \frac{1}{n} \#\{X_j \leq x\}, \tag{2}$$

and let

$$\chi = \chi_n(\alpha) = \sqrt{\frac{\ln \frac{2}{\alpha}}{2n}}. \tag{3}$$

Massart (The tight constant in the Dvoretsky-Kiefer-Wolfowitz inequality, *Ann. Prob., 18*, 1269-1283, 1990.) shows that for all n,

$$\mathbf{P}_F \left\{ \|F - \hat{F}_n\| > \chi \right\} \le \alpha. \tag{4}$$

Define

$$D = D_\chi^\zeta \equiv \{G \in \mathcal{Q} : -\chi \le \hat{F}_n(x) - G(x) \le \zeta, \ \forall x \in \mathbf{R}\}, \tag{5}$$

and $D_\chi = D_\chi^\chi$. Because of Massart's result,

$$\mathbf{P}_F\{D_\chi \ni F\} \ge 1 - \alpha. \tag{6}$$

Let C be those measures in $\mathcal{X}$ whose densities are monotone decreasing on $\mathbf{R}^+$. Then

$$\mathbf{P}_F\{C \cap D_\chi \ni F\} \ge 1 - \alpha. \tag{7}$$

Consider a fixed functional $T : \mathcal{X} \to \mathbf{R}$. We have

$$\mathbf{P}_F\{\inf_{G \in C \cap D_\chi} T(G) \le T(F) \le \sup_{G \in C \cap D_\chi} T(G)\} \ge 1 - \alpha; \tag{8}$$

that is, if we set

$$T^-(\mathbf{X}) = \inf_{G \in C \cap D_\chi} T(G) \tag{9}$$

and

$$T^+(\mathbf{X}) = \sup_{G \in C \cap D_\chi} T(G), \tag{10}$$

the interval $[T^-, T^+]$ is a $1 - \alpha$ confidence interval for $T(F)$. For that matter, let A be an arbitrary index set, and let $\{T_\alpha\}_{\alpha \in A}$ be an arbitrary collection of functionals on $\mathcal{X}$. Then

$$\mathbf{P}_F\{[T_\alpha^-(\mathbf{X}), T_\alpha^+(\mathbf{X})] \ni T_\alpha(F) \forall \alpha \in A\} \ge 1 - \alpha; \tag{11}$$

that is, the *simultaneous coverage probability* for any collection of confidence intervals derived from the set $C \cap D_\chi$ is at least $1 - \alpha$. (We shall discuss simultaneous confidence intervals and multiplicity in hypothesis tests in greater detail presently.)

Take $T(G) = g(y)$, the value of the density g of the measure G at the point y. Let $T^+ = T^+(y) = \sup_{G \in C \cap D_\chi} g(y)$ and $T^- = T^-(y) \inf_{G \in C \cap D_\chi} g(y)$. Finding T^+ and T^- are infinite-dimensional linear programming problems.

In terms of the densities g, the problem of finding T^- is

$$\inf\{g(y) : g \text{ is monotone, and } \forall x \in \mathbf{R}^+, g(x) \geq 0, \quad \text{and } -\chi \leq \int_0^x g(u)du - \hat{F}_n(x) \leq \chi\}. \tag{12}$$

The constraints are linear inequalities in g, and the objective functional is linear in g. The unknown is infinite-dimensional, and there are an infinite number of constraints.

It happens that these problems can be reduced exactly to finite-dimensional linear programs. Notice that the maximum vertical distance between $G(x)$ and $\hat{F}_n(x)$ must occur at one of the data X_j. With probability one, the data $\{X_j\}$ are distinct, and the smallest datum is greater than zero. Wlog assume that the data are ordered such that $0 \leq X_1 < X_2 < \cdots < X_n < \infty$. Let N denote the number of elements in the set $\{0, y\} \cup \{X_j\}_{j=1}^n$. With probability one, N is either $n+1$ or $n+2$. For $j = 1, \ldots, N$, let y_j be the jth smallest element in the set $\{0, y\} \cup \{X_j\}_{j=1}^n$. For any density $g(x)$, define $\tilde{g}^+(x)$ to be the left-continuous piecewise average of g on the intervals determined by $\{y_j\}$:

$$\tilde{g}^+(x) = \sum_{j=1}^{N-1} 1_{x \in [y_j, y_{j+1})} \frac{1}{y_{j+1} - y_j} \int_{y_j}^{y_{j+1}} g(u)du. \tag{13}$$

Then $\tilde{g}(x)$ is the density of a subprobability measure. Let $\tilde{G}$ be the corresponding measure. If $g(x)$ is monotone decreasing, so is $\tilde{g}(x)$, and $\|\tilde{G} - \hat{F}_n\|_{KS} = \|G - \hat{F}_n\|_{KS}$.