Homework 8
Due Thursday Nov 6 by 3:00

1. (Relationship between tests and confidence sets.) Let $X_1, \ldots, X_n \sim p(x;\theta)$ where $\theta \in \mathbb{R}$.

   (a) Let C_n be a $1-\alpha$ confidence interval for θ. Consider testing
$$H_0 : \theta = \theta_0 \quad \text{and} \quad H_1 : \theta \neq \theta_0.$$
   Define a test as follows: reject H_0 if $\theta_0 \notin C_n$. Show that this defines a test with type I error α.

   (b) For each $\theta_0 \in \mathbb{R}$, let $\phi(\theta_0, X_1, \ldots, X_n)$ be a level α test of
$$H_0 : \theta = \theta_0 \quad \text{and} \quad H_1 : \theta \neq \theta_0.$$
   In other words, $\phi(\theta_0, X_1, \ldots, X_n) \in \{0, 1\}$ and
$$\mathbb{P}_{\theta_0}(\phi(\theta_0, X_1, \ldots, X_n) = 1) = \alpha.$$
   Let
$$C_n = \left\{ \theta_0 : \ \phi(\theta_0, X_1, \ldots, X_n) = 0 \right\}.$$
   Show that C_n is a $1-\alpha$ confidence set. In other words, for every θ,
$$P_\theta(\theta \in C_n) = 1 - \alpha.$$

2. Let $X_1, \ldots, X_n \sim P$ where P has a density p on $[0, 1]$. Assume that the density p satisfies the following condition:
$$|p(x) - p(y)| \leq L \ |x - y|$$
   for all $x, y \in [0, 1]$, where $L > 0$ is a constant. Given an integer m, let $h = 1/m$ and define m bins:
$$B_1 = [0, h], \quad B_2 = (h, 2h], \ldots.$$
   Let n_j denote the number of observations in bin B_j. The histogram density estimator is
$$\widehat{p}(x) = \frac{\widehat{\pi}_j}{h} \quad \text{for } x \in B_j$$
   where $\widehat{\pi}_j = n_j/n$.

   (a) Find $\mathbb{E}(\widehat{p}(x))$ and find $V(\widehat{p}(x))$.

   (b) Find an approximate expression for $R_n(h) = \mathbb{E} \int (\widehat{p}(x) - p(x))^2 dx$.

   (c) Find h_n to minimize your expression and find $R_n(h_n)$. You should get that $R_n(h_n) = O(n^{-2/3})$.

3. Let $T(F)$ be a statistical functional. Suppose that T satisfies the following condition: there exists $L > 0$ such that, for any F and G,

$$|T(F) - T(G)| \leq L \sup_x |F(x) - G(x)|. \tag{1}$$

Let $\theta = T(F)$ and $\widehat{\theta}_n = T(\widehat{F}_n)$.

(a) Show that $\widehat{\theta}_n \xrightarrow{P} \theta$.

(b) Let $\mathcal{F}$ be all distributions on the real line and define

$$T(F) = \inf\{x : \; F(x) = 1\}.$$

Show that T does not satisfy (1).

4. Let $X_1, \ldots, X_n \sim P$ and let $\mu = \mathbb{E}(X_i)$ and $\sigma^2 = V(X_i)$. Let $\overline{X}_n = (1/n)\sum_{i=1}^n X_i$ and $s^2 = (1/n)\sum_{i=1}^n (X_i - \overline{X}_n)^2$. Let $X_1^*, \ldots, X_n^*$ denote a bootstrap sample and let $\overline{X}_n^* = \frac{1}{n}\sum_{i=1}^n X_i^*$.

(a) Show that $\mathbb{E}(\overline{X}_n^* | X_1, \ldots, X_n) = \overline{X}_n$.

(b) Show that $\mathrm{Var}(\overline{X}_n^* | X_1, \ldots, X_n) = s^2$.