Homework 3 10/36-702: Due March 22

1. Let $X_1, \ldots, X_n \sim P$ where P has a density p on $[0,1]$. Suppose that $p \in \mathcal{P}$ where

$$\mathcal{P} = \Big\{ p : \ p \geq 0, \ \int p = 1, \ |p(y) - p(x)| \leq L\,|x - y|, \ \text{for all } x, y \in [0,1] \Big\}.$$

   We want to estimate the density p. Let $d(p,q) = \sqrt{\int_0^1 (p(x) - q(x))^2 dx}$. Let R_n be the minimax risk:

$$R_n = \inf_{\widehat{p}} \sup_{p \in \mathcal{P}} \mathbb{E}_P(d(p, \widehat{p})).$$

   Find the minimax rate, i.e., find an upper and lower bound on R_n that converge to 0 at the same rate.

2. Let $\mathcal{P} = \{p_\theta : \ \theta \in \Theta\}$ be a parametric family where $\Theta \subset \mathbb{R}$. You may assume regularity conditions on the model as needed. Show that

$$K(p_\theta, p_{\theta+\epsilon}) = \frac{\epsilon^2}{2} I(\theta) + o(\epsilon^2)$$

   where K is the Kullback-Leibler distance and I is the Fisher information. Explain why this implies that the minimax risk for estimating θ with loss $|\theta - \widehat{\theta}|$ is bounded from below by $C/\sqrt{n}$ as long as $I(\theta) > 0$.

3. Consider the normal means problem:

$$Y_j = \theta_j + \frac{1}{\sqrt{n}} \epsilon_j, \quad j = 1, 2, \ldots.$$

   Let $\theta = (\theta_1, \theta_2, \ldots)$. Assume that $\theta \in \Theta$ where

$$\Theta = \Big\{ \theta : \sum_{j=1}^{\infty} \theta_j^2 j^{2p} \leq C \Big\}.$$

   Define the prior π such that $\theta_1, \theta_2, \ldots$ are independent and

$$\theta_j \sim N(0, j^{-2q}), \quad j = 1, 2, \ldots.$$

   (a) Find the posterior for θ.

   (b) Find the Bayes estimator under squared error loss (i.e. find the posterior mean).

   (c) Find the maximum risk over Θ of the Bayes estimator.

   (d) Show that if $q = p + (1/2)$ the Bayes estimator has risk of order $n^{-2p/(2p+1)}$. (This is the minimax rate.)

   (e) Show that $\pi(\Theta) > 0$ requires $q > p + (1/2)$.