1st sheet of exercises
due: 25 October 2007

Exercise 1.

(a) Show that the set consisting of all densities is convex and $L_1(\mathbb{R}^d)$-closed.

(b) Consider the set of all densities which additionally lie in $L_2(\mathbb{R}^d)$. Is it $L_1(\mathbb{R}^d)$-closed, too?

Exercise 2. (Proof of Lemma 1.1) Assume two real-valued random variables X and Y defined on the same probability space. Show that the regression function $r(x) := E(Y \mid X = x)$ minimizes the term

$$E|Y - g(X)|^2$$

among all measurable functions $g : \mathbb{R} \to \mathbb{R}$ with $E|g(X)|^2 < \infty$.

Exercise 3. We assume that a density f lies in the class of densities on $[0, 1]$ which are piecewise constant on an equidistant grid with the bin width $1/m$, i.e. $f \in \mathcal{F}_m$ where

$$\mathcal{F}_m := \Big\{ f \text{ density} : f(x) = \sum_{j=0}^{m-1} f_j \cdot \chi_{[j/m,(j+1)/m)}(x) \, , \, x \in [0,1] \Big\} ,$$

where χ_I denotes the indicator function of a set I.

(a) As a reasonable estimator $\hat{f}(x)$ of $f(x)$ based on the i.i.d. data $X_1, \ldots, X_n$ having the density $f \in \mathcal{F}_m$ with known integer $m > 0$, we suggest

$$\hat{f}(x) := \frac{m}{n} \sum_{j=0}^{m-1} \#\big\{ l : X_l \in [j/m, (j+1)/m) \big\} \cdot \chi_{[j/m,(j+1)/m)}(x) \, .$$

Calculate its expectation and variance for some arbitrary but fixed $x \in [0, 1]$.

(b) Now assume that $m = 2^k$ with unknown integer $k > 0$. Modify the estimator to handle this situation.

Exercise 4. We assume that f is an arbitrary univariate density which is supported and continuous on $(0, 1)$. Construct a pointwise consistent estimator $\hat{f}(x)$ of $f(x)$ based on i.i.d. real-valued data $X_1, \ldots, X_n$ having the density f; more precisely, the convergence

$$E\big|\hat{f}(x) - f(x)\big|^2 \stackrel{n\to\infty}{\longrightarrow} 0$$

shall be satisfied for all $x \in (0, 1)$.
tip: It is helpful to consider the estimator from Exercise 3 while $m = m_n$ may now depend on the sample size n.