Math 6070-1: Spring 2013
Solutions for problem set 3

1. *Let $X_1, X_2, \ldots$ be an i.i.d. sample from a density function f. We assume that f is differentiable in an open neighborhood V of a fixed point x, and $B := \max_{z\in V} |f'(z)| < \infty$.*

   (a) *Prove that for all $\lambda > 0$, $m \geq 1$, and all $x \in \mathbf{R}$,*

   $$\mathrm{P}\left\{\min_{1\leq j\leq m} |X_j - x| \geq \lambda\right\} = \left[1 - \int_{x-\lambda}^{x+\lambda} f(z)\,\mathrm{d}z\right]^m.$$

   Solution: Let $\mathcal{M}_j$ denote the event that $|X_j - x| \geq \lambda$. Then,

   $$\begin{aligned}\mathrm{P}\left\{\min_{1\leq j\leq m} |X_j - x| \geq \lambda\right\} &= \mathrm{P}(\mathcal{M}_1 \cap \cdots \cap \mathcal{M}_m)\\ &= \mathrm{P}(\mathcal{M}_1) \times \cdots \times \mathrm{P}(\mathcal{M}_m)\\ &= [\mathrm{P}(\mathcal{M}_1)]^m,\end{aligned}$$

   since $X_1, \ldots, X_m$ are i.i.d. The claim follows, since

   $$\mathrm{P}(\mathcal{M}_1) = 1 - \mathrm{P}\{|X_1 - x| < \lambda\} = 1 - \int_{x-\lambda}^{x+\lambda} f(z)\,\mathrm{d}z.$$

   (b) *Prove that for all $\epsilon > 0$ small enough,*

   $$\max_{z\in[x-\epsilon\,,x+\epsilon]} \big|f(x) - f(z)\big| \leq 2B\epsilon.$$

   Use this to estimate $|\int_{x-\epsilon}^{x+\epsilon} f(z)\,\mathrm{d}z - 2\epsilon f(x)|$.
   Solution: The first bound is from freshman calculus. Namely, we use the fundamental theorem of calculus to see that

   $$f(x) - f(z) = \int_z^x f'(w)\,\mathrm{d}w.$$

   Therefore, by the triangle inequality for integrals,

   $$\begin{aligned}|f(x) - f(z)| &\leq \int_z^x |f'(w)|\,\mathrm{d}w\\ &\leq B|x - z|.\end{aligned}$$

This yields the bound

$$\max_{z\in[x-\epsilon\,,x+\epsilon]} \left|f(x)-f(z)\right| \le B\epsilon,$$

which is slightly better than the one we are asked to derive.
Next we use a trick that we used in the context of density estimation [for convolutions]. Namely, we write

$$\int_{x-\epsilon}^{x+\epsilon} f(z)\,\mathrm{d}z - 2\epsilon f(x) = \int_{x-\epsilon}^{x+\epsilon} [f(z)-f(x)]\,\mathrm{d}z.$$

Then, apply the triangle inequality for integrals in order to obtain

$$\begin{aligned}\left|\int_{x-\epsilon}^{x+\epsilon} f(z)\,\mathrm{d}z - 2\epsilon f(x)\right| &= \left|\int_{x-\epsilon}^{x+\epsilon} [f(z)-f(x)]\,\mathrm{d}z\right| \\ &\le \int_{x-\epsilon}^{x+\epsilon} |f(z)-f(x)|\,\mathrm{d}z \\ &\le 2B\epsilon^2.\end{aligned}$$

(c) *Suppose that as $m\to\infty$, $\lambda_m\to\infty$ and $\lambda_m^2/m\to 0$. Then, prove that*

$$\lim_{m\to\infty} \frac{-1}{2\lambda_m} \ln \mathrm{P}\left\{\min_{1\le j\le m} |X_j - x| \ge \frac{\lambda_m}{m}\right\} = f(x).$$

Solution: Throughout, let us write

$$\delta_m := \int_{x-(\lambda_m/m)}^{x+(\lambda_m/m)} f(z)\,\mathrm{d}z.$$

We know that $\lim_{m\to\infty}\delta_m = 0$, thanks to freshman calculus. In fact,

$$\delta_m = \frac{2\lambda_m}{m} f(x) \pm \frac{2B\lambda_m^2}{m^2},$$

thanks to part (b).
Part (a) shows that

$$\frac{-1}{2\lambda_m} \ln \mathrm{P}\left\{\min_{1\le j\le m} |X_j - x| \ge \frac{\lambda_m}{m}\right\} = \frac{-m}{2\lambda_m}\ln(1-\delta_m).$$

According to Taylor's expansion, from freshman calculus,

$$\begin{aligned}\ln(1-\delta_m) &= -\delta_m + \frac{\delta_m^2}{2} - \cdots \\ &= -\frac{2\lambda_m}{m} f(x) \mp \frac{2B\lambda_m^2}{m^2} + \frac{1}{2}\left(\frac{2\lambda_m}{m} f(x) \pm \frac{2B\lambda_m^2}{m^2}\right)^2 \\ &= -\frac{2\lambda_m}{m} f(x) \mp \frac{(2B + [2f(x)]^2)\lambda_m^2}{m^2} + \cdots.\end{aligned}$$

The point is that the error of approximation is bounded above by some finite constant times λ_m^2/m^2. In particular,

$$\lim_{m\to\infty} \frac{-m}{2\lambda_m} \ln(1-\delta_m) = f(x),$$

which is another way to rewrite the statement.

(d) *Devise an estimator of $f(x)$ based on the previous steps.*

Solution: Suppose $X_1, \ldots, X_n$ is i.i.d. data from pdf f, and suppose that we can choose and fix an integer $1 \ll m \ll n$ such that n divides m. Block your data into $(n/m) - 1$ sub-blocks of length m; that is,

$$\underbrace{X_1\,, \ldots, X_m}_{\text{first block}}, \underbrace{X_{m+1}\,, \ldots, X_{2m}}_{\text{second block}}, \ldots, \underbrace{X_{n-m+1}\,, \ldots, X_n}_{\text{last block}}$$

These blocks are i.i.d. Moreover, we know from part (c) that

$$\mathrm{P}\left\{\text{Every term in the } j\text{th block is at least } \frac{\lambda_m}{m} \text{ away from } x\right\} \approx \mathrm{e}^{-\lambda_m f(x)},$$

provided only that m is large. By the law of large numbers, since $n \gg m$, and since there are $(n/m) - 1 \approx (n/m)$ blocks,

$$\frac{m}{n} \sum_{j=1}^{(n/m)-1} I\left\{\text{Every term in the } j\text{th block is at least } \frac{\lambda_m}{m} \text{ away from } x\right\} \approx \mathrm{e}^{-\lambda_m f(x)}.$$

Therefore, a reasonable estimator is

$$\widehat{f}(x) := \frac{-1}{\lambda_m} \ln\left(\frac{m}{n} \sum_{j=1}^{(n/m)-1} I\left\{\text{Every term in the } j\text{th block is at least } \frac{\lambda_m}{m} \text{ away from } x\right\}\right).$$

If you cannot find such an $(n\,, m)$ pair exactly, then divide the data into m blocks that are equally-sized, inasmuch as possible. Any such division works just as well as any other, and works as in the preceding.