5.6 (Minimax lower bound for nonparametric regression).Consider the fixed design regression model in which we observe $Y_1, \cdots, Y_n$ satisfying the relation

$$Y_i = f(i/n) + \varepsilon_i, \qquad \text{for } i = 1, \dots, n.$$
$$f \in \mathcal{C}_1^\beta[0,1]$$

where $\varepsilon_1, ..., \varepsilon_n$ are independent standard normal variables and f is the unknown function of interest. In this statistical setting, derive a minimax lower bound for the pointwise risk $\mathbb{E}_f \left(\hat{f}(t) - f(t)\right)^2$ over $\mathcal{C}_1^\beta[0,1]$ at a fixed point $t \in [0,1]$.

Solution. We can derive a minimax lower bound by studying a collection of a finite collection of indistinguishable functions $\mathcal{G}$ in the Holder ball since

$$\inf_{\hat{f}} \sup_{f \in \mathcal{C}_1^\beta[0,1]} \mathbb{E}_f\, |\hat{f}(t) - f(t)|^2 \geq \inf_{\hat{f}} \sup_{f \in \mathcal{G}} \mathbb{E}_f\, |\hat{f}(t) - f(t)|^2$$

Due to the distributional assumption on the error term we can find an expression for the likelihood ratio:

$Y_i \sim (f(i/n), 1)$ and consequently the likelihood is given by: $\ell(f|Y_i) \propto \exp\left[-\frac{1}{2}\left(Y_i - f(i/n)\right)^2\right]$. By independence the likeliood given the full sample is given by:

$$\ell(f|Y) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^n \left(Y_i - f(i/n)\right)^2\right]$$

The likelihood ratio is given by te quotient of the likelihoods:

$$\begin{aligned}
\frac{p_{f_0}}{p_{f_1}}(X) &= \exp\left[\sum_{i=1}^n y_i(f_0(i/n) - f_1(i/n)) - \frac{1}{2}\sum_{i=1}^n (f_0(i/n)^2 - f_1(i/n)^2)\right] \\
&= \exp\left[\langle f_0(i/n) - f_1(i/n), Y_j - f_1(i/n)\rangle - \frac{1}{2}\left\|f_0(i/n) - f_1(i/n)\right\|^2\right] \\
&\stackrel{d}{=} \exp\left[\left\|f_0(i/n) - f_1(i/n)\right\| Z - \frac{1}{2}\left\|f_0(i/n) - f_1(i/n)\right\|^2\right]
\end{aligned}$$

Where all inner products and norms are Euclidian and $f(i/n)$ denotes the n-dimensional vector $(f(1/n), \cdots f(n/n))'$.

In this setting the Neyman-Pearson lemma also holds: The most powerfull test of level α (for two simple hypotheses)is the Likelihood Ratio test:

$$\varphi(X) = 1_{\left\{\frac{p_{f_1}(X)}{p_{f_0}(X)} > c_\alpha\right\}}$$

where $c_\alpha = \exp\left[\|f_1(i/n) - f_0(i/n)\| \xi_{1-\alpha} - \frac{1}{2}\|f_1(i/n) - f_0(i/n)\|^2\right]$ and

$$\frac{p_{f_1}}{p_{f_0}}(X) \stackrel{d}{=} \exp\left[\|f_0(i/n) - f_1(i/n)\| Z - \frac{1}{2}\|f_0(i/n) - f_1(i/n)\|^2\right]$$

Similarly to theorem 4.2.2 it can be shown a consistent test for two simple hypotheses excicts iff

$$\|f_0(i/n) - f_1(i/n)\|^2 \to \infty$$

Let $\mathcal{G}$ our proposed collection of indistinguisable functions be given by $\{f_0, f_1\}$ with $f_0 \equiv 0$ and $f_1(x) = \sigma^\beta K(\frac{x-t}{\sigma})$.

$$\|f_0(i/n) - f_1(i/n)\|^2 = \sum_{i=1}^{n} f_1(i/n)^2 \sim n\|f_1\|_{L^2}^2$$

This follows since:

$$\begin{aligned}
\left|\frac{1}{n}\sum_{i=1}^{n} f_1(i/n)^2 - \int_0^1 f_1^2(t)dt\right| &\leq \sum_{i=1}^{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} |f_1^2(t) - f_1^2(i/n)|^\beta dt \\
&\leq \sum_{i=1}^{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} C|i/n - t|^\beta dt \ \text{ Since } f_1^2 \in \mathcal{C}_1^\beta[0,1] \\
&\leq C\sum_{i=1}^{n}\int_{\frac{i-1}{n}}^{\frac{i}{n}} \left(\frac{1}{n}\right)^\beta dt \ \text{ Since } t \in \left(\frac{i-1}{n}, \frac{i}{n}\right) \\
&\leq Cn^{-\beta} \to 0 \text{ as } n \text{ tends to } \infty
\end{aligned}$$

Thus

$$\|f_0(i/n) - f_1(i/n)\|^2 \sim n\|f_1\|_{L^2}^2 \sim n\sigma^{2\beta+1}$$

If we choose $\sigma \sim n^{-\frac{1}{2\beta+1}}$ then the norm does not go to infinity and therefore there excicts no consistent test. Note that their seperation rate is given by:
$d(f_0, f_1) = |f_0(t) - f_1(t)|^2 = \sigma^{2\beta}K(0)^2 \sim 2cn^{-\frac{2\beta}{2\beta+1}}$, if we choose $\sigma \sim n^{-\frac{1}{2\beta+1}}$
f_0 and f_1 are $2cr_n$-seperated, where $r_n = n^{-\frac{2\beta}{2\beta+1}}$

In this statistical setting a theoreom similar to 5.2.1 can be formulated (I think the proof will be the same in this setting) and thus the minimax lowerbound is equal to the seperation rate of the indistinguisable functions:

$$\inf_{\hat{f}} \sup_{f \in \mathcal{C}_1^\beta[0,1]} \mathbb{E}_f \, |\hat{f}(t) - f(t)|^2 \geq cn^{-\frac{2\beta}{1+2\beta}}$$

□