Statistical Theory for High and Infinite Dimensions

May 24, 2017

Exercise 5.2: Minimax estimation in a non-regular parametric model.

Consider the parametric model

$$dX_t = \operatorname{sign}(t-\theta)dt + \frac{1}{\sqrt{n}}dW_t$$

where $\operatorname{sign}(x) = 1_{x\geq 0} - 1_{x<0}$ and $\theta \in (0,1)$ is an unknown parameter.

5.2a: *Prove the minimax lower bound*

$$\inf_{\hat{\theta}} \sup_{t\in(0,1)} \mathrm{E}_\theta|\theta - \hat{\theta}| \gtrsim \frac{1}{n}$$

Idea: apply **Proposition 5.2.1**. Define

$$\mathcal{F} := \left\{ f_\theta : f_\theta = \operatorname{sign}(t-\theta),\ \theta \in (0,1) \right\} \subset L_2$$

and let the (semi)-metric d be the L^1-norm. Let $\theta_0 = 1/2$, $\theta_1 = 1/2 + 1/n$ and consider $f_0 = f_{\theta_0}, f_1 = f_{\theta_1}$, for $n > 2$. Then, the first assumption in the proposition is satisfied:

$$d(f_0, f_1) = \int |f_1 - f_0| d\mu = 2 \cdot \frac{1}{n}$$

For the second assumption, use **Proposition 4.2.2**: For all $f_0, f_1 \in L^2[0,1]$,

$$\frac{1}{2}\inf_{\phi}\left(\mathrm{E}_{f_0}\phi + \mathrm{E}_{f_1}(1-\phi)\right) = 1 - \Phi\left(\frac{1}{2}\sqrt{n}\|f_1 - f_0\|_2\right)$$

Hence, there exist a consistent test for $H_0 : f = f_0, H_1 : f = f_1$ if and only if $\sqrt{n}\|f_1 - f_0\|_2 \to \infty$. We have

$$\|f_1 - f_0\|_2^2 = \int |f_1 - f_0|^2 d\mu = \int_\theta^{\theta+1/n} |2|^2 \mathrm{d}\mu = \frac{4}{n}.$$

Hence, $\sqrt{n}\|f_1 - f_0\|_2 = 2$. Therefore, no consistent test exists. By **Proposition 5.2.1**,

$$\inf_{\hat{\theta}} \sup_{f \in \mathcal{F}} \mathrm{E}_f \left[d(\hat{f}, f) \right] \gtrsim \frac{1}{n}$$

The proof is finished by observing that

$$d(f_{\theta_0}, f_{\theta_1}) = \int |f_{\theta_1} - f_{\theta_0}| d\mu = 2|\theta_1 - \theta_0|$$

5.2b: *Describe the MLE* $\hat{\theta}_{MLE}$ *in this model and show that it achieves the rate* $\frac{1}{n}$*, uniformly over* $(0, 1)$.

From equation (2.6), we know that the MLE for f is the maximizer of

$$p_f(X) = \exp\left\{ n \int_0^1 f(t) dX_t - \frac{1}{2} n \int_0^1 f^2(t) dt \right\}$$

over $\mathcal{F}$. For any $f_\theta \in \mathcal{F}$, we have

$$\begin{aligned} p_{f_\theta}(X) &= \exp\left\{ n \int_0^1 \mathrm{sign}(t - \theta) dX_t - \frac{n}{2} \int_0^1 \mathrm{sign}^2(t - \theta) dt \right\} \\ &= \exp\left\{ n \left(\int_0^\theta \mathrm{sign}(t - \theta) dX_t + \int_\theta^1 \mathrm{sign}(t - \theta) dX_t \right) - \frac{n}{2} \right\} \\ &= \exp\left\{ n \left(-(X_\theta - X_0) + X_1 - X_\theta \right) - \frac{n}{2} \right\} \\ &= \exp\left\{ n \left(X_1 - 2X_\theta \right) - \frac{n}{2} \right\} \end{aligned}$$

almost surely. Hence,

$$\begin{aligned} \hat{\theta}_{MLE} &= \arg\max_{\theta \in (0,1)} p_f(X) \\ &= \arg\max_{\theta \in (0,1)} \exp\left\{ n(X_1 - 2X_\theta) - \frac{n}{2} \right\} \\ &= \arg\min_{\theta \in (0,1)} X_\theta \end{aligned}$$

To show the second part, i.e.

$$\sup_{\theta \in (0,1)} \mathrm{E}_\theta |\hat{\theta}_{MLE} - \theta| \lesssim \frac{1}{n},$$

fix $\theta \in (0, 1)$ and n and consider $n(\hat{\theta} - \theta)$.

$$
\begin{aligned}
&\mathrm{P}_\theta\left(n(\hat{\theta}-\theta) \geq s\right) = \mathrm{P}_\theta(\hat{\theta}-\theta \geq s/n) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} X_t \leq \min_{t<\theta+s/n} X_t\right) \\
&\quad\leq \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} X_t \leq X_\theta\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} (X_t - X_\theta) \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} \frac{1}{\sqrt{n}}(W_t - W_\theta) + (t-\theta) \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} \frac{1}{\sqrt{n}}W_{t-\theta} + (t-\theta) \leq 0\right) && \text{(i)} \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} \frac{t-\theta}{\sqrt{n}}W_{1/(t-\theta)} + (t-\theta) \leq 0\right) && \text{(iv)} \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} \frac{1}{\sqrt{n}}W_{1/(t-\theta)} + 1 \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\geq\theta+s/n} W_{1/(n(t-\theta))} + 1 \leq 0\right) && \text{(iii)} \\
&\quad= \mathrm{P}_\theta\left(\min_{t\,:\,n(t-\theta)\geq s} W_{1/(n(t-\theta))} + 1 \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\,:\,1/n(t-\theta)\leq 1/s} W_{1/(n(t-\theta))} + 1 \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\leq 1/s} W_t + 1 \leq 0\right) \\
&\quad= \mathrm{P}_\theta\left(\min_{t\leq 1/s} W_t \leq -1\right).
\end{aligned}
$$

The following is a classical result. For any $x < 0$, $T > 0$, we have

$$
\mathrm{P}\left(\min_{t\leq T} W_t \leq x\right) = 2\mathrm{P}\left(W_T \leq x\right)
$$

So, for Z standard Gaussian,

$$
\begin{aligned}
\mathrm{P}_\theta\left(n(\hat{\theta}-\theta) \geq s\right) &\leq \mathrm{P}_\theta\left(\min_{t \leq 1/s} W_t \leq -1\right) \\
&= 2\mathrm{P}\left(W_{1/s} \leq -1\right) \\
&= 2\mathrm{P}\left(\frac{1}{\sqrt{s}} Z \leq -1\right) \\
&= 2\mathrm{P}\left(Z \leq -\sqrt{s}\right).
\end{aligned}
$$

In a same manner, we obtain

$$
\begin{aligned}
\mathrm{P}_\theta\left(n(\hat{\theta}-\theta) \leq -s\right) &\leq \mathrm{P}\left(\max_{t \leq 1/s} W_t \geq 1\right) \\
&= 2\mathrm{P}\left(Z \geq \sqrt{s}\right)
\end{aligned}
$$

Hence,

$$\mathrm{P}_\theta\left(n|\hat{\theta}-\theta| > s\right) \leq 4\mathrm{P}(Z \geq \sqrt{s}) = 4\mathrm{P}(Z^2 \geq s)$$

Thus, we know that the distribution has tails that are bounded by exponential decreasing tails. Such a distribution has finite moments. Hence,

$$\mathrm{E}_\theta n|\hat{\theta}-\theta| < \infty$$

Since the bound does not depend on θ, we obtain

$$\sup_{\theta \in (0,1)} \mathrm{E}_\theta n|\hat{\theta}-\theta| < \infty$$

5.2c: *Find the limiting distribution of* $n(\hat{\theta}_{MLE}-\theta)$.

(sketch of the proof) In the derivation of the bound of $\mathrm{P}_\theta\left(n(\hat{\theta}-\theta)\right)$, the only inequality was the third line. However, if we prove

$$\min_{t < \theta + s/n} X_t \to X_\theta$$

almost surely as $n \to \infty$, then a similar, more careful argument, can be used to show that in the limit, we have equality. Hence,

$$\lim_{n \to \infty} \mathrm{P}\left(n(\hat{\theta}-\theta) \geq s\right) = 2\mathrm{P}_\theta\left(Z \leq -\sqrt{s}\right)$$

for $s > 0$, and

$$\lim_{n \to \infty} \mathrm{P}\left(n(\hat{\theta}-\theta) \leq s\right) = 2\mathrm{P}_\theta\left(Z \leq -\sqrt{-s}\right)$$

for $s < 0$. It turns out that this distribution is the Laplace distribution.