Topics in Statistical Theory — Example Sheet 1

Lucas Riedstra

8 November 2020

Question 4. *Let $X \sim \mathrm{Bin}(n, p)$. Compare the Hoeffding, Bennett, and Bernstein upper bounds on $\mathbb{P}(X/n \geq \frac{1}{2})$ as $p \to 0$.*

Solution. Note that X/n is the average of n i.i.d. random variables $Y_i \sim \mathrm{Bern}(p)$, where $Y_i \in [0, 1]$ for all i.

1. We start with Hoeffding's inequality. In this case, we have

$$\mathbb{P}\left(X/n - p \geq \frac{1}{2}\right) \leq \exp\left(-\frac{2n^2(1/2)^2}{n}\right) = \exp\left(-\frac{n}{2}\right),$$

   and this bound also holds as $p \to 0$.

2. We continue with Bennett's inequality. We consider the mean-zero random variables $Y_i - p$, which are bounded from above by $b = 1 - p$. Now Bennett's inequality tells us that

$$\mathbb{P}\left(X/n \geq \frac{1}{2}\right) \leq \exp\left(-\frac{np(1-p)}{(1-p)^2} h\left(\frac{1-p}{2p(1-p)}\right)\right) = \exp\left(-\frac{np}{1-p} \cdot h\left(\frac{1}{2p}\right)\right).$$

   We compute

$$h\left(\frac{1}{2p}\right) = \left(1 + \frac{1}{2p}\right)\log\left(1 + \frac{1}{2p}\right) - \frac{1}{2p} = \log\left(1 + \frac{1}{2p}\right) + \frac{1}{2p}\left(\log\left(1 + \frac{1}{2p}\right) - 1\right),$$

   and therefore

$$\frac{np}{1-p} h\left(\frac{1}{2p}\right) \geq \frac{n}{2(1-p)}\left(\log\left(1 + \frac{1}{2p}\right) - 1\right) \overset{p \to 0}{\to} \infty.$$

   It follows that the Bennett upper bound converges to 0 as $p \to 0$.

3. We finish with Bernstein's inequality. We have for $q \geq 3$ that

$$\frac{2(1-p)}{q!} \leq \frac{2}{q!} \leq 3^{2-q},$$

   and therefore we have

$$\begin{aligned} \mathbb{E}[(Y_i - p)_+^q] &= p(1-p)^q = \sigma_p^2(1-p)^{q-1} = (q!\sigma_p^2(1-p)^{q-2}/2) \cdot (2(1-p)/q!) \\ &\leq q!\sigma_p^2((1-p)/3)^{q-2}/2, \end{aligned}$$

   so $Y_i - p$ satisfies Bernstein's condition with $c = \frac{1-p}{3}$. Now Bernstein's inequality tells us that

$$\mathbb{P}(X/n \geq \frac{1}{2} \leq \exp\left(-\frac{n(1/2)^2}{2(\sigma_p^2 + (1-p)/6)}\right) = \exp\left(-\frac{n}{8\sigma_p^2 + 4(1-p)/3}\right) \overset{p \to 0}{\to} \exp\left(-\frac{3n}{4}\right).$$

Of course, the true limit is 0 for any n, which is only given by Bennett's inequality. We also see that Hoeffding's inequality gives the most loose bound.

Question 10. *(a) Verify the algebraic identity*

$$\varphi_\sigma(x-\mu)\varphi_{\sigma'}(x-\mu') = \varphi_{\sigma\sigma'/(\sigma^2+\sigma'^2)^{1/2}}(x-\mu^*)\varphi_{(\sigma^2+\sigma'^2)^{1/2}}(\mu-\mu')$$

where $\mu^* := (\sigma'^2\mu + \sigma^2\mu')/(\sigma^2+\sigma'^2)$, and φ_σ is the $N(0,\sigma^2)$ density.

(b) Let $X_1,\dots,X_n \overset{\text{iid}}{\sim} N(0,\sigma^2)$. Taking K to be the $N(0,1)$ density, show that the MISE of the kernel density estimate $\hat{f}_n$ with kernel K and bandwidth h can be expressed exactly as

$$\mathrm{MISE}(\hat{f}_n) = \frac{1}{2\pi^{1/2}}\left\{\frac{1}{nh} + (1-\frac{1}{n})\frac{1}{(h^2+\sigma^2)^{1/2}} - \frac{2^{3/2}}{(h^2+2\sigma^2)^{1/2}} + \frac{1}{\sigma}\right\}.$$

Proof. (a) We have

$$\begin{aligned}
&\frac{(x-\mu)^2}{\sigma^2} + \frac{(x-\mu')^2}{\sigma'^2}\\
&= \frac{\sigma'^2(x-\mu)^2 + \sigma^2(x-\mu')^2}{\sigma^2\sigma'^2}\\
&= \frac{(\sigma^2+\sigma'^2)x^2 - 2(\sigma'^2\mu+\sigma^2\mu')x + \sigma'^2\mu^2 + \sigma^2\mu'^2}{\sigma^2\sigma'^2}\\
&= \frac{(\sigma^2+\sigma'^2)(x^2-2\mu^*x) + \sigma'^2\mu^2 + \sigma^2\mu'^2}{\sigma^2\sigma'^2}\\
&= \frac{(\sigma^2+\sigma'^2)(x^2-2\mu^* + \mu^{*2}) - (\sigma^2+\sigma'^2)\mu^{*2} + \sigma'^2\mu^2 + \sigma^2\mu'^2}{\sigma^2\sigma'^2}\\
&= \frac{(x-\mu^*)^2}{(\sigma\sigma'/(\sigma^2+\sigma'^2)^{1/2})^2} + \frac{\sigma'^2\mu + \sigma^2\mu'^2 - (\sigma'^2\mu+\sigma^2\mu')^2/(\sigma^2+\sigma'^2)}{\sigma^2\sigma'^2}\\
&= \frac{(x-\mu^*)^2}{(\sigma\sigma'/(\sigma^2+\sigma'^2)^{1/2})^2} + \frac{(\sigma^2+\sigma'^2)(\sigma'^2\mu+\sigma^2\mu'^2) - (\sigma'^2\mu+\sigma^2\mu')^2}{\sigma^2\sigma'^2(\sigma^2+\sigma'^2)}\\
&= \frac{(x-\mu^*)^2}{(\sigma\sigma'/(\sigma^2+\sigma'^2)^{1/2})^2} + \frac{(\mu-\mu')^2}{\sigma^2+\sigma'^2}\\
&= \frac{(x-\mu^*)^2}{(\sigma\sigma'/(\sigma^2+\sigma'^2)^{1/2})^2} + \frac{(\mu-\mu')^2}{((\sigma^2+\sigma'^2)^{1/2})^2},
\end{aligned}$$

which proves the claim.

(b) Let $K=\varphi_1$ and define $K_h(x) := h^{-1}K(x/h)$ so $K_h = \varphi_h$. Then recall from the lectures that

$$\mathrm{MISE}(\hat{f}_n) = \frac{1}{n}\int_{\mathbb{R}}\left[(\varphi_h^2 * \varphi_\sigma)(x) - (\varphi_h*\varphi_\sigma)^2(x)\right]\mathrm{d}x + \int_{-\infty}^{\infty}\left[(\varphi_h*\varphi_\sigma)(x) - \varphi_\sigma(x)\right]^2\mathrm{d}x$$

We will use the previous exercise to compute all these terms. We have in general

$$\begin{aligned}
(\varphi_h*\varphi_\sigma)(x) &= \int_{\mathbb{R}}\varphi_h(x-z)\varphi_\sigma(z)\,\mathrm{d}z\\
&= \int_{\mathbb{R}}\varphi_\sigma(z)\varphi_h(z-x)\,\mathrm{d}z\\
&= \varphi_{(\sigma^2+h^2)^{1/2}}(x)\int_{\mathbb{R}}\varphi_\xi(z-\mu^*)\,\mathrm{d}z\\
&= \varphi_{(\sigma^2+h^2)^{1/2}}(x). \qquad (1)
\end{aligned}$$

We also have

$$\varphi_\sigma^2(x-\mu) = \varphi_{\sigma/\sqrt{2}}(x-\mu)\varphi_{\sqrt{2}\sigma}(0) = \frac{1}{2\sigma\sqrt{\pi}}\varphi_{\sigma/\sqrt{2}}(x-\mu). \tag{2}$$

Combining eqs. (1) and (2) we get

$$\begin{aligned}(\varphi_h^2 * \varphi_\sigma)(x) &= \int_{\mathbb{R}} \varphi_h^2(x-z)\varphi_\sigma(z)\,\mathrm{d}z \\ &= \frac{1}{2h\sqrt{\pi}}\int_{\mathbb{R}} \varphi_{h/\sqrt{2}}(x-z)\varphi_\sigma(z)\,\mathrm{d}z \\ &= \frac{1}{2h\sqrt{\pi}}(\varphi_{h/\sqrt{2}} * \varphi_\sigma)(x) \\ &= \frac{1}{2h\sqrt{\pi}}\varphi_{(\sigma^2+h^2/2)^{1/2}}(x)\end{aligned}$$

We also get

$$(\varphi_h * \varphi_\sigma)^2(x) = \varphi^2_{(\sigma^2+h^2)^{1/2}}(x) = \frac{1}{2(\sigma^2+h^2)^{1/2}\sqrt{\pi}}\varphi_{(\sigma^2+h^2)^{1/2}/\sqrt{2}}(x).$$

Finally, we have

$$((\varphi_h * \varphi_\sigma)(x) - \varphi_\sigma(x))^2 = (\varphi_h * \varphi_\sigma)^2(x) - 2(\varphi_h * \varphi_\sigma)(x)\varphi_\sigma(x) + \varphi_\sigma^2(x).$$

The first term we already computed, the third term is $\frac{1}{2\sigma\sqrt{\pi}}\varphi_{\sigma/\sqrt{2}}(x)$, so we only need to compute

$$(\varphi_h * \varphi_\sigma)(x)\varphi_\sigma(x) = \varphi_{(\sigma^2+h^2)^{1/2}}(x)\varphi_\sigma(x) = \varphi_\xi(x)\varphi_{(2\sigma^2+h^2)^{1/2}}(0) = \frac{1}{\sqrt{2\pi}(2\sigma^2+h^2)^{1/2}}\varphi_\xi(x),$$

where ξ is an irrelevant constant.

Combining all these terms and using that $\varphi_\sigma(x-\mu)$ integrates to 1 for any μ, σ, we get

$$\begin{aligned}\operatorname{MISE}\hat{f}_n &= \frac{1}{n}\left(\frac{1}{2h\sqrt{\pi}} - \frac{1}{2(\sigma^2+h^2)^{1/2}\sqrt{\pi}}\right) + \frac{1}{2(\sigma^2+h^2)^{1/2}\sqrt{\pi}} - \frac{\sqrt{2}}{\sqrt{\pi}(2\sigma^2+h^2)^{1/2}} + \frac{1}{2\sigma\sqrt{\pi}} \\ &= \frac{1}{2\sqrt{\pi}}\left(\frac{1}{nh} + (1-\frac{1}{n})\frac{1}{(\sigma^2+h^2)^{1/2}} - \frac{2^{3/2}}{(2\sigma^2+h^2)^{1/2}} + \frac{1}{\sigma}\right).\end{aligned}$$

□

Question 11. *Use the expression from 10(b) to derive an upper bound of the form* $\operatorname{MISE}\hat{f}_n \leq C_1/(nh) + C_2^2h^4$ *(where you should specify* C_1, C_2*). Show that* $\varphi_\sigma \in \mathcal{F}_{\mathcal{N}}(2, L)$ *with* $L^2 = 3/(8\pi^{1/2}\sigma^5)$*, and hence compare the bound from the first part of this question with that obtained from the general theory from lectures.*

Solution. We have

$$\left(1-\frac{1}{n}\right)\frac{1}{(\sigma^2+h^2)^{1/2}} - \frac{2^{3/2}}{(2\sigma^2+h^2)^{1/2}} \leq \frac{1}{(\sigma^2+h^2)^{1/2}} - \frac{2}{(\sigma^2+h^2/2)^{1/2}} < 0,$$

I do not know how to obtain an upper bound of the form $C_1/(nh) + C_2^2 h_4$ from this expression.

To show that $\varphi_\sigma \in \mathcal{F}_{\mathcal{N}}(2, L)$, we must show that $\varphi_\sigma \in \mathcal{F}_{\mathcal{N}}(2, L)$. By question 9, it suffices to show that $\|\varphi''_\sigma\|^2_{L^2} \leq L^2$. A simple computation gives

$$\varphi''_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma^5}}(x^2-\sigma^2)\exp\left(-\frac{x^2}{2\sigma^2}\right) \leq \frac{1}{\sqrt{2\pi\sigma^5}}x^2\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

so

$$\|\varphi''_\sigma\|_{L^2}^2 \le \frac{1}{2\pi\sigma^{10}} \int_{\mathbb{R}} x^4 \exp\left(-\frac{x^2}{\sigma^2}\right) \mathrm{d}x \overset{\star}{=} \frac{1}{2\pi\sigma^{10}} \cdot \frac{3}{4}\sqrt{\pi}\sigma^5 = \frac{3}{8\sqrt{\pi}\sigma^5} = L^2,$$

where $\star$ can be computed using the fact that the integral is, up to scaling, the fourth moment of $N(0, \sqrt{2}\sigma)$ distribution.

Note that for $K = \varphi_1$, we have

$$R(K) = \int_{-\infty}^{\infty} \varphi_1^2(x)\,\mathrm{d}x = \frac{1}{2\sqrt{\pi}},$$

while

$$\mu_2^2(K) = \int_{-\infty}^{\infty} x^2 \varphi_1(x)\,\mathrm{d}x = 1.$$

Plugging all the above into theorem 27 shows that

$$\mathrm{MISE}(\hat{f}_n) \le \frac{1}{2\sqrt{\pi}} \frac{1}{nh} + \frac{3}{8\sqrt{\pi}\sigma^5} h^4.$$