Non-Paramatric Statistics Exercise 3

Osman Ceylan, Jiahui Wang, Zhuoyao Zeng

24. November 2020

Exercise 1.5

Let $(X, \mathcal{A})$ be a measurable space and μ be a probability measure on X. Moreover, let $\mathbf{P}$ be a probability measure on X that has a μ-density $h : X \to [0, \infty)$, and let $L_{\text{dens}} : X \times \mathbb{R} \to [0, \infty)$ be the corresponding loss function. Compute the excess L_{dens}-risks for both $\mathbf{P}$ and μ. Which one is better suited for capturing the intuitive goal of density estimation?

Solution:

h is a μ-density of $\mathbf{P} \;\Rightarrow\; \forall A \in \mathcal{A} : \mathbf{P}(A) = \displaystyle\int_A h \,\mathrm{d}\mu$.

Recall from lecture: $\forall x \in X, t \in \mathbb{R} : L_{\text{dens}}(x, t) = |h(x) - t|$.

$\forall x \in X : L_{\text{dens}}(x, h(x)) = 0 \;\Rightarrow\;$ The Bayes Risk of L_{dens} related to any measure of X is 0. $\quad (*)$

Suppose now we have an estimated density h' of h.

This means that: $h' \geq 0 \;\wedge\; \displaystyle\int_X h' \,\mathrm{d}\mu = 1 \;\wedge\; \forall A \in \mathcal{A} : \mathbf{P}'(A) := \int_A h' \,\mathrm{d}\mu$.

As the excess L_{dens}-risks of μ, we obtain:

$$\mathcal{R}_{L_{\text{dens}},\mu}(h') - \mathcal{R}^*_{L_{\text{dens}},\mu} \overset{(*)}{=} \mathcal{R}_{L_{\text{dens}},\mu}(h') \overset{\text{Def.}}{=} \int_X L_{\text{dens}}(x, h'(x)) \,\mathrm{d}\mu(x) = \int_X |h - h'| \,\mathrm{d}\mu$$

As for the excess L_{dens}-risks of $\mathbf{P}$, we obtain:

$$\mathcal{R}_{L_{\text{dens}},\mathbf{P}}(h') - \mathcal{R}^*_{L_{\text{dens}},\mathbf{P}} \overset{(*)}{=} \mathcal{R}_{L_{\text{dens}},\mathbf{P}}(h') \overset{\text{Def.}}{=} \int_X L_{\text{dens}}(x, h'(x)) \,\mathrm{d}\mathbf{P}(x) = \int_X |h - h'| \,\mathrm{d}\mathbf{P} = \int_X |h - h'| \, h \,\mathrm{d}\mu.$$

The excess risk relating to $\mathbb{P}$is more suitable for capturing the intuitive goal of density estimation because we can see that compared with that of excess risk under $\mathbb{P}$, the integrand of excess risk under μ has an additional function h, and thus, the value of the integrand will be larger for $x \in X : h(x) > 1$, and smaller for $x \in X : h(x) < 1$. This means that the risk penalty for the part, where h is "relatively large", is larger, and the risk penalty for the part, where h is "relatively small", is then smaller.

Exercise 2.1

Let $X \neq \emptyset$ and $\mathfrak{A} = (A_j)_{j\in J}$ be an at most countable partition of X. Describe the space $\mathcal{L}_0(X)$ of measurable functions $X \to \mathbb{R}$ for the corresponding σ-algebra $\mathcal{A} := \sigma(\mathfrak{A})$.

Solution:

We construct $\Sigma := \{S \subseteq X | \exists I \subseteq J : S = \bigcup_{i\in I} A_i\}$.

$\mathfrak{A} = (A_j)_{j\in J}$ is a countable partition of $X \;\Leftrightarrow\; X = \dot{\bigcup}_{j\in J} A_j$.

$\Rightarrow$ A simple verification of the axioms of σ-algebra shows that Σ is a σ-algebra containing $\mathfrak{A}$.

Meanwhile it is also easy to see that $\Sigma \subseteq \sigma(\mathfrak{A}) \;\Rightarrow\; \Sigma = \sigma(\mathfrak{A})$. $\quad (i)$

Moreover, recall the following theorem from measure theory:

> Let $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$ be mesrable spaces and $f : \Omega_1 \to \Omega_2$.
>
> Also let $(M_n)_{n\in\mathbb{N}}$ a sequence from $\mathcal{F}_1$ and at the same time a partition of Ω_1.
>
> For $\forall n \in \mathbb{N}$ denote $\mathcal{F}_1 \mid M_n := \{A \cap M_n | A \in \mathcal{F}_1\}$.
>
> Then it holds that: f is $\mathcal{F}_1$-$\mathcal{F}_2$-measurable $\Leftrightarrow\; \forall n \in \mathbb{N} : f|_{M_n}$ is $\mathcal{F}_1|M_n$-$\mathcal{F}_2$-measurable.

$(A_j)_{j\in J}$ is paarwise disjoint $\;\Rightarrow\; \forall j \in J : \mathcal{A}|A_j = \{\emptyset, A_j\}$.

$\Rightarrow\; \forall j \in J \, \forall f \in \mathcal{L}_0(X) : f|_{A_j}^{-1}(\mathcal{B}) \in \mathcal{A}|A_j = \{\emptyset, A_j\}$. $\quad (ii)$

We also know that $\forall c \in \mathbb{R} : \{c\} \in \mathcal{B}$.

With (ii) we can then deduce that $\forall j \in J \; \forall f \in \mathcal{L}_0(X) : f|_{A_j}^{-1}(\{c\}) \in \{\emptyset, A_j\}$. $\quad (iii)$

Combining all considerations , we conclude that:

$$
\begin{aligned}
\mathcal{L}_0(X) &\overset{\text{Def}}{=} \{f : X \to \mathbb{R} \mid f^{-1}(\mathcal{B}) \subseteq \mathcal{A}\} \\
&\overset{(i)}{=} \{f : X \to \mathbb{R} \mid \forall B \in \mathcal{B} \; \exists I_B \subseteq J : f^{-1}(B) = \textstyle\bigcup_{i\in I_B} A_i\} \\
&\overset{(ii)}{=} \{f : X \to \mathbb{R} \mid \forall j \in J : f|_{A_j}^{-1}(\mathcal{B}) = \emptyset \;\vee\; f|_{A_j}^{-1}(\mathcal{B}) = A_j\} \\
&\overset{(iii)}{=} \{f : X \to \mathbb{R} \mid f = \sum_{i\in I} c_i \mathbb{1}_{A_i}\}.
\end{aligned}
$$