**Statistical Theory** MATH-442
Spring 2025 Ho Yun & Yoav Zemel

---

**Exercise sheet 8**

---

The support of a distribution $P$ on $\mathbb{R}^d$ (or any Polish space) is the set of points $x$ such that $P(B_\epsilon(x)) > 0$ for all $\epsilon > 0$, where $B_\epsilon(x) = \{y : \|y - x\| < \epsilon\}$. You may use without proof that $\mathbb{P}(X \in \text{supp}P_X) = 1$, where $\text{supp}P_X$ is the support of the distribution of $X$. The proof of this is given below, but is **not** examinable.

For each $x \in \mathbb{R}^d$ let $r(x) := \sup\{r \geq 0 : P_X(B_r(x)) = 0\}$, with $r(x) = 0$ for $x \in \text{supp}(P_X)$ and $r(x) > 0$ otherwise. For each $x \notin \text{supp}(P_X)$ there exists $x' \in \mathbb{Q}^d$ with $\|x - x'\| \leq r(x)/4$. This satisfies $P_X(B_{r(x)/2}(x')) \leq P_X(B_{3r(x)/4}(x)) = 0$ and so $r(x') \geq r(x)/2$ and $\|x - x'\| \leq r(x')/2$. Hence

$$\mathbb{P}(X_0 \notin \text{supp}(P_X)) \leq P_X\left(\bigcup_{x' \in \mathbb{Q}^d \setminus \text{supp}(P_X)} B_{r(x')/2}(x')\right) \leq \sum_{x' \in \mathbb{Q}^d \setminus \text{supp}(P_X)} P_X(B_{r(x')/2}(x')) = 0$$

as required.

**Exercise 1** Here we give an alternative proof that $\overline{X}_n$ is admissible in a Gaussian model with squared loss. Let $\delta$ have $R(\theta, \delta) \leq 1/n$ for all $\theta$, with strict inequality for some $\theta_0$. We wish to obtain a contradiction. By continuity of $\theta \mapsto R(\theta, \delta)$ we can find $\epsilon > 0$ and $\theta_1 > \theta_0$ such that $R(\theta, \delta) < 1/n - \epsilon$ for all $\theta \in (\theta_0, \theta_1)$.

For $\tau > 0$ consider the prior $\pi_\tau = N(0, \tau^2)$.

1. Show that for the $\pi_\tau$-Bayes estimator $\delta_\tau$,

$$\frac{\frac{1}{n} - r(\pi_\tau, \delta)}{\frac{1}{n} - r(\pi_\tau, \delta_\tau)} = \frac{\int (\frac{1}{n} - R(\theta, \delta)) \frac{1}{\sqrt{2\pi}\tau} \exp(-\theta^2/2\tau^2) d\theta}{\frac{1}{n} - \frac{1}{n + \tau^{-2}}}$$

2. Show that as $\tau \to \infty$, this fraction converges to $\infty$ and deduce a contraction.

**Exercise 2** This problem considers minimaxity in nonparametric classes of distributions with squared loss.

1. Let $\mathcal{F}$ be the class of distributions with variance bounded by 1. Suppose we are interested in the mean $\mu = \mu(F)$. Show that $\overline{X}_n$ is minimax for the estimation of $\mu$.

2. Let $\mathcal{F}$ be the class of all distributions on $[0, 1]$. Find a minimax estimator for the mean $\mu = \mu(F)$. *Hint: we have a candidate from the previous exercise set. Show that it is indeed minimax. Write .*

**Exercise 3** Let $g^* : \mathbb{R}^d \to \{0, 1\}$ be the Bayes classifier.

1. Prove that
$$\mathbb{P}(g^*(X) \neq Y) = \mathbb{E}\{\min(\eta(X), 1 - \eta(X))\}.$$

2. Show that for any classifier $g : \mathbb{R}^d \to \{0, 1\}$,
$$\mathbb{P}(g^*(X) \neq Y) \leq \mathbb{P}(g(X) \neq Y).$$

3. For $\tilde{\eta}(x)$ and $\tilde{g}(x) = 1$ if $\tilde{\eta}(x) \geq 1/2$, prove that

$$\mathbb{P}(\tilde{g}(X) \neq Y) - \mathbb{P}(g^*(X) \neq Y) \leq 2\mathbb{E}|\eta(X) - \tilde{\eta}(X)|.$$

**Exercise 4** Denote the probability measure for $X$ by $P_X$. Fix $x \in \mathrm{supp}(P_X) \in \mathbb{R}^d$ and reorder the data $(X_1, Y_1), \ldots, (X_n, Y_n)$ according to increasing values of $||X_i - x||$. The reordered data sequence is denoted by

$$(X_{(1)}(x), Y_{(1)}(x)), \ldots, (X_{(n)}(x), Y_{(n)}(x)).$$

If $\lim_{n\to\infty} k/n = 0$, then prove that $||X_{(k)}(x) - x|| \to 0$ with probability one.

Show that if $X_0$ is independent of the data and has probability measure $P_X$, then $||X_{(k)}(X_0) - X_0|| \to 0$ with probability one whenever $k/n \to 0$.

**Exercise 5** Here we give an alternative argument that $\mathbb{P}(\|X_{(k)}(X) - X\| > \delta) \to 0$ for all $\delta > 0$ for the $k$-nearest neighbour classifier when $k/n \to 0$ and $k \to \infty$. Let $U_{(k)}$ be the $k$-th order statistic of independent $U_1, \ldots, U_n \sim [0, 1]$. Using that $U_{(k)}$ has mean $k/(n + 1)$ and variance $k(n - k + 1)/[(n + 1)^2(n + 2)]$, show that

$$\mathbb{P}\left(U_{(k)} > \tfrac{2k}{n}\right) \to 0.$$

For $x \in \mathrm{supp}(P_X)$ define $F_x(t) = \mathbb{P}(\|X_1 - x\| \leq t)$. Let $F_x^{-1}$ denote the corresponding quantile function. Show that $\lim_{s\searrow 0} F_x^{-1}(s) = 0$. Deduce that $\mathbb{P}(\|X_{(k)}(x) - x\| > \delta) \to 0$ for all $\delta > 0$. Deduce further that $\mathbb{P}(\|X_{(k)}(X) - X\| > \delta) \to 0$, where $X$ is independent of the sequence $X_1, \ldots$ and has the same distribution as $X_1$.