STA3000F: Homework 1

Due: October 11, 2024, before class on Quercus

Q1: Application of concentration inequalities

For any integer $n > 0$, define the Hamming distance on the hypercube $\{0,1\}^n$

$$d_H(x, y) := \sum_{i=1}^{n} \mathbf{1}_{x_i \neq y_i}, \quad \text{for any } x, y \in \{0,1\}^n.$$

Show that there exists a universal constant $c > 0$, such that for any $n > 0$, there exists a subset $A \subseteq \{0,1\}^n$ with $|A| \geq e^{cn}$, satisfying

$$d_H(x, y) \geq \frac{n}{4}, \quad \text{for any pair } x, y \in A.$$

[Hint: consider a subset formed by i.i.d. uniform random samples from the hypercube.]

Q2: weighted loss function

Consider a class of probability models $(\mathbb{P}_\theta : \theta \in \mathbb{R})$ with density functions p_θ for $\theta \in \mathbb{R}$. We want to estimate a functional $g(\theta)$ under the loss function

$$L(\theta; a) = (a - g(\theta))^2 w(\theta),$$

for a known non-negative weight function w.

1. Let π be a prior distribution, find the Bayes estimator under the loss function L. (Express it using the posterior distribution; you can assume integrability of relevant functions).

2. Let $\mathbb{P}_\theta := \text{Ber}(\theta)$ for $\theta \in (0, 1)$, take $g(\theta) = \theta$, and let the weight function be $w(\theta) = \frac{1}{\theta(1-\theta)}$. Find a minimax estimator under this loss function

Q3: (Bayes) Crámer–Rao lower bounds

Given $\theta \in \mathbb{R}^d$, we observe the pair $(X, Y) \in \mathbb{R}^d \times \{0, 1\}$ as follows

$$X \sim \mathbb{P}, \quad \text{and} \quad Y|X \sim \text{Ber}\Big(\frac{1}{1 + e^{-\theta^\top X}}\Big).$$

We assume that $\mathbb{P}$ has a density with respect to Lebesgue measure and that $\mathbb{E}\big[\,\|X\|_2^2\,\big] < +\infty$.

1. Derive the Fisher information matrix $I(\theta)$ for estimating θ (express it as an expectation under the distribution of X).

2. Let us consider i.i.d. samples $(X_i, Y_i)_{i=1}^n$ from the distribution above. Consider the special case of $d = 1$ and $X \sim \mathcal{N}(0, 1)$ for simplicity. Show that there exists a universal constant $c > 0$, such that

$$\inf_{\widehat{\theta}} \sup_{\theta \in [\theta_0 - \varepsilon, \theta_0 + \varepsilon]} \mathbb{E}\big[|\widehat{\theta} - \theta|^2\big] \geq \Big(nI(\theta_0) + cn\varepsilon + \frac{c}{\varepsilon^2}\Big)^{-1},$$

   valid for any $n \geq 1$ and $\varepsilon \in (0, 1)$.

Q4: Le Cam's two-point method

Consider the parameter estimation problem for a class $(\mathbb{P}_\theta : \theta \in \Theta)$. Let $g(\theta) \in \mathbb{R}$ be the functional of interest, and consider a mean-squared loss function.

1. Show that

$$\inf_{\delta} \sup_{\theta \in \Theta} \mathbb{E}\big[|\delta(X) - g(\theta)|^2\big] \geq \frac{1}{8} \sup_{\theta_0, \theta_1 \in \Theta} \Big\{ |g(\theta_0) - g(\theta_1)|^2 \cdot \big(1 - d_{\mathrm{TV}}(\mathbb{P}_{\theta_1}, \mathbb{P}_{\theta_0})\big) \Big\}.$$

   [Hint: use the testing lower bound.]

2. Consider the following special case: $\mathbb{P}_\theta = \mathrm{Unif}([0, \theta])$, and $\theta \in [1, 2]$; we are interested in estimating $g(\theta) = \theta$. Given samples $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathbb{P}_\theta$, show a lower bound on the minimax risk

$$\inf_{\widehat{\theta}_n} \sup_{\theta \in [1,2]} \mathbb{E}\big[|\widehat{\theta}_n(X_1, X_2, \cdots, X_n) - g(\theta)|^2\big].$$

   It suffices to give a tight rate of convergence as n grows. The constant pre-factor in the lower bound does not matter.