STA3000F: Homework 1

Due: October 12, 2025, 11:59pm on Quercus

Q1: Median-of-means estimator

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with

$$\mathbb{E}[X_i] = \mu, \quad \text{and} \quad \mathrm{var}(X_i) = \sigma^2 < +\infty.$$

The median-of-means estimator is constructed as follows: partition the data into k groups of equal size (assume k divides n), compute the sample mean within each group, and take the median of these k means as the estimator $\widehat{\mu}_{\mathrm{MOM}}$.

Prove that there exists universal constant $c_1, c_2 > 0$, such that for any $\delta \in (0, 1)$, by choosing $k = \lceil c_1 \log(2/\delta) \rceil$, with probability $1 - \delta$, the median-of-means estimator satisfies

$$|\widehat{\mu}_{\mathrm{MOM}} - \mu| \leq c_2 \sigma \sqrt{\frac{\log(2/\delta)}{n}}.$$

Q2: sufficient statistics

Consider the class of distributions

$$\mathcal{P} = \Big\{ \mathbb{P} \ : \ \mathbb{P} \text{ is a distribution on } \mathbb{R}, \ \mathbb{E}_{\mathbb{P}}[|X|] < +\infty \Big\}.$$

Given n i.i.d. samples $X_1, X_2, \ldots, X_n$ from some distribution $\mathbb{P} \in \mathcal{P}$.

1. Let $(X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)})$ be the order statistics of the samples, i.e., rearranging the samples in non-decreasing order. Show that the vector of order statistics $T(X_1, X_2, \cdots, X_n) := (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is a sufficient statistic for the family $\mathcal{P}$.

2. Let $\widehat{\mu}(X_1, X_2, \ldots, X_n) := X_1$ be an estimator of $\mu := \mathbb{E}_{\mathbb{P}}[X]$. Find the Rao–Blackwellized estimator $\widehat{\mu}_{\mathrm{RB}}$ of $\widehat{\mu}$ given the sufficient statistic T.

Q3: local non-asymptotic minimax lower bound

Consider a parametric family of distributions $(\mathbb{P}_\theta : \theta \in \mathbb{R})$ with density p_θ. Given $\theta_0 \in \mathbb{R}$, we assume that the Fisher information $I(\theta_0) > 0$, and that $\sup_{\theta\in\mathbb{R}} |I''(\theta)| \le c$. Show that there exists a constant $c_0 > 0$ (which may depend on c), such that

$$\inf_{\widehat{\theta}} \sup_{|\theta-\theta_0|\le\varepsilon} \mathbb{E}_\theta[(\widehat{\theta}-\theta)^2] \ge \left(nI(\theta_0) + c_0 n\varepsilon^2 + \frac{c_0}{\varepsilon^2}\right)^{-1},$$

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

Q4: asymmetric Le Cam's two-point method

Consider the parameter estimation problem for a class $(\mathbb{P}_\theta : \theta \in \Theta)$. Let $\theta_0, \theta_1 \in \Theta$ be two parameters. For any $q > 0$, show that

1. For any test ϕ, we have

$$q \cdot \mathbb{P}_{\theta_0}(\phi(X) = 1) + \mathbb{P}_{\theta_1}(\phi(X) = 0) \geq 1 - \frac{1}{q}\Big\{1 + \chi^2(\mathbb{P}_{\theta_1} || \mathbb{P}_{\theta_0})\Big\},$$

   where we define the χ^2-divergence as

$$\chi^2(\mathbb{P}||\mathbb{Q}) := \int \Big(\frac{d\mathbb{P}}{d\mathbb{Q}} - 1\Big)^2 \mathrm{d}\mathbb{Q}.$$

2. Let $0 < \psi_0 < \psi_1$ be two fixed positive scalars. Show that

$$\inf_{\widehat{\theta}} \sup_{i \in \{0,1\}} \frac{1}{\psi_i^2} \mathbb{E}_{\theta_i}[|\widehat{\theta} - \theta_i|^2] \geq \frac{|\theta_0 - \theta_1|^2}{8} \Big\{ \frac{1}{\psi_1^2} - \frac{\psi_0^2}{\psi_1^4} \Big\{1 + \chi^2(\mathbb{P}_{\theta_1} || \mathbb{P}_{\theta_0})\Big\}\Big\}.$$