30100: Mathematical Statistics I **Winter 2022**

Homework 5

Lecturer: Chao Gao

1. Consider $X \sim N(\theta, I_p)$. Find the minimax test for the following problem:

$$H_0 : \theta = \mu, \quad H_1 : \|\theta - \mu\| \geq \epsilon.$$

A testing procedure ϕ is minimax if it minimizes $\sup_{\theta \in \Theta_0} \mathbb{P}_\theta \phi + \sup_{\theta \in \Theta_1} \mathbb{P}_\theta (1 - \phi)$.

(a) Let π be the uniform distribution on the sphere $\{\theta : \|\theta - \mu\| = \epsilon\}$. Write down the testing procedure that minimizes the average testing error

$$\mathbb{P}_\mu \phi + \int \mathbb{P}_\theta (1 - \phi) \pi(\theta) d\theta,$$

denoted by ϕ_{opt}.

(b) Define a vector $\Delta \in \mathbb{R}^p$ whose first coordinate is $\|X - \mu\|$ and the remaining coordinates are 0. Explain why

$$\int e^{\delta \langle X - \mu, \theta - \mu \rangle} \pi(\theta) d\theta = \int e^{\delta \langle \Delta, \theta - \mu \rangle} \pi(\theta) d\theta$$

is true.

(c) Show that if $Z \sim N(0, I_p)$, then $\frac{\epsilon Z}{\|Z\|}$ follows a uniform distribution on the sphere $\{\theta : \|\theta\| = \epsilon\}$.

(d) Show that if $Z \sim N(0, I_p)$, then $\frac{Z_1^2}{\|Z\|^2}$ follows a Beta distribution. What is this Beta distribution? Then, find the density of $\frac{Z_1}{\|Z\|}$.

(e) Use the above facts to prove that $\int e^{\delta \langle X - \mu, \theta - \mu \rangle} \pi(\theta) d\theta$ is an increasing function of $\|X - \mu\|$ when $\delta > 0$.

(f) Show that $\phi_{\text{opt}} = \{\|X - \mu\|^2 > t(\epsilon)\}$, and thus ϕ_{opt} is a chi-squared test.

(g) Show the chi-squared test ϕ_{opt} is minimax. Hint: you have learned how to prove minimaxity in Lecture 4. A similar argument can be used in a testing problem.

2. Consider the mixture model $X_1, ..., X_n \sim \frac{1}{2} N(\theta, 1) + \frac{1}{2} N(-\theta, 1)$ with $\theta \in [-1, 1]$. The goal here is to estimate $|\theta|$ (note that one can only identify θ up to a sign change).

(a) Find an estimator $\hat{T}$ of $|\theta|$ that satisfies

$$\sup_{|\theta| \leq 1} \mathbb{E}_\theta (\hat{T} - |\theta|)^2 \leq \frac{C}{n^{1/2}},$$

for some universal constant $C > 0$. Hint: it is easy to see that $\frac{1}{n} \sum_{i=1}^n X_i^2$ estimates $\theta^2 + 1$.

(b) For densities p and q, use Cauchy-Schwarz to prove

$$\int |p-q| \leq \sqrt{\int \frac{(p-q)^2}{p}}.$$

The quantity $\chi^2(P\|Q) = \int \frac{(p-q)^2}{p}$ is known as the χ^2-divergence between p and q.

(c) Show $\chi^2(P^n\|Q^n) = \left(\int \frac{q^2}{p}\right)^n - 1$.

(d) Let p be the density of $N(0,1)$ and q be the density of $\frac{1}{2}N(-cn^{-1/4},1)+\frac{1}{2}N(cn^{-1/4},1)$. Show $\int \frac{q^2}{p} = \frac{1}{2}\left(e^{c^2n^{-1/2}} + e^{-c^2n^{-1/2}}\right)$, and thus $\chi^2(P^n\|Q^n) \leq \frac{1}{4}$ for some sufficiently small c. (Hint: to make your calculation clean, establish $\int \frac{\phi(x-a)\phi(x-b)}{\phi(x)}dx = e^{ab}$ first, where $\phi(x-a)$ denotes the density of $N(a,1)$.)

(e) With Le Cam's two-point argument, show

$$\inf_{\hat{T}} \sup_{|\theta|\leq 1} \mathbb{E}_\theta(\hat{T} - |\theta|)^2 \geq \frac{C_1}{n^{1/2}},$$

for some universal constant $C_1 > 0$.

(f) Typically, a parametric model always has minimax rate n^{-1} with respect to the squared error loss. However, for the Gaussian mixture model above, you have proved that the minimax rate is $n^{-1/2}$. Can you explain what makes the Gaussian mixture model special? Hint: What is the Fisher information at $\theta = 0$? What would be the minimax rate if we exclude a neighborhood of 0 from the parameter space $[-1,1]$?

3. Find a UMVUE that is inadmissible.