Statistical Theory MATH-442

Spring 2025 Ho Yun & Yoav Zemel

Exercise sheet 7

> You may use without proof that if $X \sim N(\mu, \Sigma)$ then $\mathbb{E}\|X\|^2 = \|\mu\|^2 + \text{trace}(\Sigma)$. You may use that the Beta(a, b) distribution has density $\Gamma(a+b)p^{a-1}(1-p)^{b-1}/[\Gamma(a)\Gamma(b)]$ for $p \in (0, 1)$ and expectation $a/(a+b)$ for $a, b > 0$.

Exercise 1 Let $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, \Sigma_X)$ with a prior $N(\mu, \Sigma_\theta)$ for θ, where μ, Σ_X and Σ_θ are fixed and known and Σ_X, Σ_θ invertible. Find the posterior distribution of μ given $\vec{X} = (X_1, \dots, X_n)$. Show that the Bayes risk of the resulting Bayes rule, with respect to squared loss, is equal to trace(Σ_n), where you should specify what the matrix Σ_n is.

Exercise 2 Let $X_1, \dots, X_n \sim$Bernoulli(p). Find a minimax estimator for p with respect to squared loss. Deduce that $\overline{X}_n$ is not minimax.

Exercise 3 Let δ_m be π_m-Bayes for a prior π_m and suppose that $r(\pi_m, \delta_m) \to r$. Suppose that δ is an arbitrary estimator with risk bounded above by r. Show that δ is minimax. Deduce that $\overline{X}_n$ is minimax for θ in a $\mathcal{N}(\theta, I_d)$ for any dimension d with respect to squared loss.

Exercise 4 Show that the regularity conditions for the Cramér–Rao lower bound hold in the normal model. That is, if $X_1, \dots, X_n \sim N(\theta, 1)$ are independent and δ is such that $var_\theta(\delta(\vec{X})) < \infty$ for all θ, then

$$\tfrac{\partial}{\partial\theta}\mathbb{E}_\theta(\delta(\vec{X})) = \int \tfrac{\partial}{\partial\theta}\vec{f}(\vec{x};\theta)d\vec{x},$$

where $\vec{f}(;\theta)$ is the density of $N((\theta, \dots, \theta)^\top, I_n)$ is the joint (normal) density of $\vec{X}$.

Hint. Assume first that $n = 1$ and $\theta = 0$ — this setup really conveys all the ideas. Then what needs to be shown is that

$$\tfrac{\mathbb{E}_h(\delta) - \mathbb{E}_0(\delta)}{h} - \int \phi'(x)\delta(x)dx = \tfrac{1}{h}\int [\phi(x-h) - \phi(x) - h\phi'(x)]\delta(x)dx$$

vanishes as $h \to 0$, where ϕ is the density of a standard normal. So we want to show that the integral at the right-hand side is finite. Show that this integral is bounded up to constant by $E_0|p(X)\delta(X)] + E_1|p(X)\delta(X)| + E_{-1}|p(X)\delta(X)|$ where p is a polynomial, and deduce that the integral is indeed finite. Then extend to $\theta \neq 0$ and (optional) to $n > 1$.

Exercise 5 Let $X_1, \dots, X_n \sim \mathrm{B}(p)$ independent. Show that $\overline{X}_n$ is a Bayes rule for the loss function $\mathcal{L}(p, a) = \frac{(p-a)^2}{p(1-p)}$. Explain how can it be that $\overline{X}_n$ is unbiased and Bayes at the same time. Explain why this is a sensible loss function. Show that $\overline{X}_n$ is minimax with respect to this loss function.

Exercise 6 Let $X_1, \dots, X_n \sim \mathrm{B}(p)$ independent and consider the loss function $\mathcal{L}(p, a) = 1(|p-a| > \alpha)$ for a fixed $\alpha \in (0, 1/2)$. Let $U \sim [0, 1]$ be independent of $\vec{X} = (X_1, \dots, X_n)$. Show that if α is sufficiently small, the silly randomised estimator $\delta(\vec{X}, U) = U$ (not using the data at all) has better supremum risk than any non-randomised rule $\delta = \delta(\vec{X})$. Thus no non-randomised rule can be minimax.