Statistical Inference and Learning- Ex-4

Yiqing Wang and Boaz Nadler

due by Sun. Jan. 10, 2016

Q1. Let X be a Gaussian distributed random variable with mean μ and variance σ^2. We would like to estimate its mean μ from a single observed realisation. We are interested in a *linear* estimator of the form $aX + b$ and define its risk function under the square loss as

$$R(aX + b, \mu) = \mathbb{E}[(aX + b - \mu)^2].$$

Calculate the risk function under the conditions below and show that they are inadmissible.

1. when $a > 1$.
2. when $a = 1, b \neq 0$.
3. when $a < 0$.

Q2. Consider the same setup as in Question 1 of Exercise 2: let $(x_i)_{1\le i\le n}$ be n i.i.d. samples from a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ with known variance σ^2 but unknown population mean $\mu \in \mathbb{R}$. We want to estimate μ under square error loss.

Recall that the MLE for μ is the sample average, $\hat{\mu}_{\mathrm{ML}} = \frac{1}{n}\sum_i X_i$. The goal of this exercise is to show that in one dimension (X is a real-valued random variable), the MLE is an admissible estimator.

1. Assume a Gaussian prior $\mathcal{N}(0, \Delta^2)$ on μ. Derive the associated Bayesian estimator $\hat{\mu}_\Delta$ for μ. Show that as Δ^2 tends to infinity, $\hat{\mu}_\Delta$ converges to the MLE. Can you explain this phenomenon ?

2. Calculate the risk functions $R(\hat{\mu}_{\mathrm{ML}}, \mu)$ and $R(\hat{\mu}_\Delta, \mu)$.

3. Let $\hat{\mu}$ be another estimator of μ. Denote

$$\mathbb{E}_\Delta R(\hat{\mu}, \mu) := \frac{1}{\sqrt{2\pi\Delta^2}} \int_{\mathbb{R}} R(\hat{\mu}, \mu) e^{-\frac{\mu^2}{2\Delta^2}} d\mu.$$

Show that $\mathbb{E}_\Delta R(\hat{\mu}, \mu) \geq \mathbb{E}_\Delta R(\hat{\mu}_\Delta, \mu)$.

4. Assume that the risk function $R(\hat{\mu}, \mu)$ is continuous in μ. Show that if the estimator $\hat{\mu}$ dominates the MLE, that is, $R(\hat{\mu}_{\mathrm{ML}}, \mu) \geq R(\hat{\mu}, \mu)$ for all $\mu \in \mathbb{R}$, and

$$\sup_{\mu\in\mathbb{R}} [R(\hat{\mu}_{\mathrm{ML}}, \mu) - R(\hat{\mu}, \mu)] > 0,$$

we can find a strictly positive constant ϵ and a finite (but non-empty) interval I such that

$$\mathbb{E}_\Delta R(\hat{\mu}_{\mathrm{ML}}, \mu) - \mathbb{E}_\Delta R(\hat{\mu}_\Delta, \mu) \geq \frac{\epsilon}{\sqrt{2\pi\Delta^2}} \int_I e^{-\frac{\mu^2}{2\Delta^2}} d\mu.$$

5. Show that such a strictly positive ϵ cannot exist (Hint: what happens to the quantities on both sides of the inequality as Δ tends to infinity). Conclude that $\hat{\mu}_{\mathrm{ML}}$ is admissible.

Q3. Consider a random vector $(X_1, \cdots, X_d) \in \mathbb{R}^d$ where $d \geq 3$. Assume that its entries satisfy

$$\forall i \in \{1, \cdots, d\}, \quad X_i = \mu_i + N_i$$

with unknown but fixed $(\mu_1, \cdots, \mu_d) \in \mathbb{R}^d$ and Gaussian distributed i.i.d. $(N_i)_{1 \leq i \leq d}$ of zero mean and known variance σ^2. We want to show that, rather surprisingly and in contrast to the one dimensional case, the MLE here is *inadmissible* under the square error loss.

1. Suppose that only one observation $\boldsymbol{x} = (x_1, \cdots, x_d)$ is available. Derive the MLE and its risk function under the square error loss.

2. Calculate the risk function for the following estimator (known as the James-Stein estimator),

$$\hat{\mu}_{\text{JS}} = \left(1 - \frac{(d-2)\sigma^2}{\|\boldsymbol{x}\|^2}\right) \boldsymbol{x}$$

where $\|\boldsymbol{x}\|^2 = \sum_{i=1}^d x_i^2$. Conclude that this estimator dominates the MLE.

Q4. Let $(x_i)_{1 \leq i \leq n}$ be n i.i.d. real-valued samples from a Gaussian distribution with known variance σ^2 and unknown expectation μ. We want to test $H_0 : \mu = 0$ against $H_1 : \mu > 0$.

1. Construct a size α test (type I error rate α). Is your test optimal?

2. Describe how, under the alternative hypothesis H_1, your test's power varies with n and μ.