UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Homework 12
Spring 2023

1. **Flipping Coins and Hypothesizing**

   You flip a coin until you see heads. Let $0 < p < q < 1$, and suppose that your hypotheses are

$$X = \begin{cases} 0 & \text{if the bias of the coin is } p \\ 1 & \text{if the bias of the coin is } q. \end{cases}$$

   You observe Y, the number of flips until you see heads. Find a decision rule $\hat{X}$ that maximizes $\mathbb{P}(\hat{X} = 1 \mid X = 1)$ subject to $\mathbb{P}(\hat{X} = 1 \mid X = 0) \le \beta$ for some $\beta \in [0, 1]$.

   Hint: Remember to calculate the randomization constant γ.

2. **One Flip**

   You flip a single coin and observe its result $Y \sim \text{Bernoulli}(p)$. Suppose the hypotheses are

$$X = \begin{cases} 0 & \text{if } p = \frac{1}{3} \\ 1 & \text{if } p = \frac{2}{3}. \end{cases}$$

   a. Find the MLE of X and its associated type I and type II error rates.
   b. Plot the error curve.
   c. Derive the randomized decision rule that minimizes type II error subject to the constraint of $\beta = 0.5$ on the type I error.

      Hint: You should only need to look at the plot from part b.

3. **Exam Difficulty**

   The difficulty of an EECS 126 exam, Θ, is uniformly distributed on $[0, 100]$ (continuously). Alice gets a score X that is uniformly distributed on $[0, \Theta]$, and she wants to estimate the difficulty of the exam given her score.

   a. What is the MLE of Θ? What is the MAP of Θ?

   b. What is the LLSE for Θ?

4. **Gaussian LLSE**

   Let X, Y, Z be i.i.d. $\mathcal{N}(0, 1)$.

   a. Find $\mathbb{L}(X^2 + Y^2 \mid X + Y)$.
   b. Find $\mathbb{L}(X + 2Y \mid X + 3Y + 4Z)$.
   c. Find $\mathbb{L}((X + Y)^2 \mid X - Y)$.

5. **Projections**

The following exercises are from the note on the Hilbert space of random variables. See the notes for some hints.

a. Let $\mathcal{H} \coloneqq \{X : X \text{ is a real-valued random variable with } \mathbb{E}(X^2) < \infty\}$. Prove that $\mathcal{H}$ is closed under addition and scalar multiplication over the real numbers $\mathbb{R}$, and prove that the function $\langle X, Y \rangle := \mathbb{E}(XY)$ is an inner product on $\mathcal{H}$. [1]

b. Let U be a subspace of a real inner product space V. We define the *projection* map P onto U as follows: for each $v \in V$, let Pv be the unique vector in U such that $v - Pv \in U^{\perp}$. Prove that P is a linear transformation.

c. Using part b, prove that $\mathbb{L}(X + Y \mid Z) = \mathbb{L}(X \mid Z) + \mathbb{L}(Y \mid Z)$ for all $X, Y, Z \in \mathcal{H}$.

d. Now, suppose that U is a finite-dimensional subspace, $\dim U \coloneqq n$, with an orthonormal basis $\{u_i\}_{i=1}^{n}$. Prove that $Px = \sum_{i=1}^{n} \langle x, u_i \rangle u_i$ for all $x \in V$.

[1] *Remark.* It is possible for $X \neq 0$ to have $\mathbb{E}(X^2) = 0$, e.g. if $X = 0$ with probability 1. To fix this, we can take almost-sure equivalence classes of random variables, where X and Y are equivalent if $\mathbb{P}(X = Y) = 1$. You may cite this construction when checking that $X \neq 0$ implies $\mathbb{E}(X^2) > 0$.

6. **Sufficient Statistics**

Suppose $X_1, \dots, X_n$ are i.i.d. samples drawn from a probability distribution parameterized by θ. (We are in the non-Bayesian setting, so θ is deterministic but unknown).

A statistic $T(X_1, \dots, X_n)$ is a *sufficient statistic* for θ if for all t, the conditional distribution of $(X_1, \dots, X_n)$ given $T = t$ does not depend on θ. Intuitively, $T(X_1, \dots, X_n)$ "captures all there is to know about θ from the sample $X_1, \dots, X_n$."

a. Let $X_1, \dots, X_n$ be drawn i.i.d. from a Poisson distribution with mean μ. Show that $T = \sum_{i=1}^{n} X_i$ is a sufficient statistic for μ.

b. Let T be a sufficient statistic for θ, and let $\hat{\theta}$ be an estimator for θ with $\mathrm{var}(\hat{\theta}) < \infty$. Show that in mean-squared error sense, $\mathbb{E}[\hat{\theta} \mid T]$ is at least as good as $\hat{\theta}$ at estimating θ:

$$\mathbb{E}[(\mathbb{E}[\hat{\theta} \mid T] - \theta)^2] \leq \mathbb{E}[(\hat{\theta} - \theta)^2].$$

Hint: Consider expanding the decomposition $\mathbb{E}[(\hat{\theta} - \theta)^2] = \mathbb{E}[((\hat{\theta} - \mathbb{E}[\hat{\theta}]) + (\mathbb{E}[\hat{\theta}] - \theta))^2]$.

Remark. Since $\mathbb{E}[\hat{\theta} \mid T]$ is a function of T, the result above suggests we should be looking for estimators of θ that are functions of sufficient statistics.