UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Problem Set 13
Spring 2022

1. **Sufficient Statistics**

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

(a) Let $X_1, \ldots, X_n$ be drawn 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. Let $\hat{\theta}$ be an estimator for θ with $E[\hat{\theta}^2] < \infty$. Prove that for all θ,

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

Remark: The above result states that $E[\hat{\theta} \mid T]$ is at least as good as $\hat{\theta}$ at estimating θ, in a mean-squared error sense. Since $E[\hat{\theta} \mid T]$ is a function of T, the result implies that we should be looking for estimators of θ that are functions of sufficient statistics.

2. MMSE and Conditional Expectation

Let $X, Y_1, \ldots, Y_n$ be square integrable random variables. The MMSE of X given $(Y_1, \ldots, Y_n)$ is defined as the function $\phi(Y_1, \ldots, Y_n)$ which minimizes the mean square error

$$\mathbb{E}[(X - \phi(Y_1, \ldots, Y_n))^2].$$

(a) For this part, assume $n = 1$. Show that the MMSE is precisely the conditional expectation $\mathbb{E}[X|Y]$. *Hint:* expand the difference as $(X - \mathbb{E}[X|Y] + \mathbb{E}[X|Y] - \phi(Y))$.

(b) Argue that

$$\mathbb{E}\big[(X - \mathbb{E}[X \mid Y_1, \ldots, Y_n])^2\big] \leq \mathbb{E}\left[\left(X - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X \mid Y_i]\right)^2\right].$$

That is, the MMSE does better than the average of the individual estimates given each Y_i.

3. Balls in Bins Estimation

We throw $n \geq 1$ balls into $m \geq 2$ bins. Let X and Y represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbb{E}[Y \mid X]$.

(b) What are $L[Y \mid X]$ and $Q[Y \mid X]$ (where $Q[Y \mid X]$ is the best quadratic estimator of Y given X)?

Hint: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the MMSE.

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(d) Compute $\text{var}(X)$.

(e) Compute $\text{cov}(X, Y)$.

(f) Compute $L[Y \mid X]$ using the formula. Ensure that your answer is the same as your answer to part (b).

4. **Exam Difficulties**

The difficulty of an EECS 126 exam, Θ, is uniformly distributed on $[0, 100]$ (i.e. continuous distribution, not discrete), and Alice gets a score X that is uniformly distributed on $[0, \Theta]$. Alice gets her score back and wants to estimate the difficulty of the exam.

(a) What is the MLE of Θ? What is the MAP of Θ?

(b) What is the LLSE for Θ?

5. Geometric MMSE

Let N be a geometric random variable with parameter $1-p$, and $(X_i)_{i\in\mathbb{N}}$ be i.i.d. exponential random variables with parameter λ. Let $T = X_1 + \cdots + X_N$. Compute the LLSE and MMSE of N given T.

Hint: Compute the MMSE first.

6. Projections

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

(a) Let $\mathcal{H} := \{X : X \text{ is a real-valued random variable with } \mathbb{E}[X^2] < \infty\}$. Prove that $\langle X, Y \rangle := \mathbb{E}[XY]$ makes $\mathcal{H}$ into a real inner product space. [1]

(b) Let U be a subspace of a real inner product space V and let P be the projection map onto U. Prove that P is a linear transformation.

(c) Suppose that U is finite-dimensional, $n := \dim U$, with basis $\{v_i\}_{i=1}^n$. Suppose that the basis is orthonormal. Show that $Py = \sum_{i=1}^n \langle y, v_i \rangle v_i$. (Note: If we take $U = \mathbb{R}^n$ with the standard inner product, then P can be represented as a matrix in the form $P = \sum_{i=1}^n v_i v_i^\mathsf{T}$.)

[1] To be perfectly correct, it is possible for $X \neq 0$ but $\mathbb{E}[X^2] = 0$; this occurs if $X = 0$ with probability 1. To fix this, we need to define two random variables X and Y to be *equal* if $\mathbb{P}(X = Y) = 1$. In other words, we consider *equivalence classes* of random variables, defined by the relation $\overset{a.s.}{=}$. With this definition, then if $X \neq 0$ we do indeed have $\mathbb{E}[X^2] > 0$.