E2–201 Information Theory Discussion: Saturday 22 February 2025

Problem Set 5

Instructor: Rajesh Sundaresan *TAs: Deepthi, Prachi, Narasimha, Tejashree*

Remarks:

- Collaboration, discussion, and working in teams to solve problems is strongly encouraged.
- To test your understanding, write the solution to each problem in your own words without referring to a friend, text, or class notes.

Problems:

1. Let $P^{(0)}$ and $P^{(1)}$ be two PMFs on A such that for every $a \in A$, both $P^{(0)}(a) > 0$ and $P^{(1)}(a) > 0$. $X \sim P^{(0)}$ under H_0 and $X \sim P^{(1)}$ under H_1. Consider the likelihood ratio test with threshold 0, i.e.,

$$D_1 = \left\{ a \in A \ \Big| \ \frac{P^{(1)}(a)}{P^{(0)}(a)} \geq 0 \right\}.$$

Consider a decision region C_1 such that $P^{(0)}(C_1) \leq P^{(0)}(D_1)$. Argue that

$$P^{(1)}(C_1^c) \geq P^{(1)}(D_1^c).$$

2. Consider the following two hypotheses with alphabet $\mathbb{R}$:

$$\begin{aligned} H_0 &: X = \mu_0 + Z \\ H_1 &: X = \mu_1 + Z \end{aligned}$$

where Z has the Normal distribution with zero mean and unit variance, and $\mu_1 > \mu_0$. Find the likelihood ratio (of H_1 w.r.t. H_0), the log-likelihood ratio, and the relative entropy. Discuss some interesting properties of this relative entropy.

3. Consider two zero-means Gaussians $N(0, \sigma_0^2)$ and $N(0, \sigma_1^2)$ of different variances. Compute the relative entropy: $D(N(0, \sigma_1^2) \parallel N(0, \sigma_0^2))$.

4. Let X and Y be Poisson random variables with means λ and μ respectively. Compute the relative entropy $D(P_X \parallel P_Y)$.

5. In 100 throws of an unbiased dice, the observed average number of dots per throw was 5. Find the PMF P^* in whose neighbourhood the empirical distribution will lie with high probability. (E.g., 80 sixers, 20 singles, and none of the others? 33 sixers, 33 fours, 34 fives, and none of the others?) You may need the help of a computer to find the exact P^*. Those who have played bridge may have some intuition on questions of this nature.

6. Chernoff bound: Suppose that $X_1, X_2, \cdots, X_n$ are iid with mean 0. Consider $\Pr\{n^{-1}\sum_{i=1}^{n} X_i > a\}$ for $a > 0$. In order to get the Chebyshev inequality, you squared both sides and applied Markov inequality. Instead, assume X_1 has exponential moments, apply the function $e^{t(\cdot)}$ to both sides with $t \geq 0$, and apply Markov inequality. Optimising the upper bound over $t \geq 0$, can you show that the probability decays exponentially fast to zero?

7. Suppose that $\{P_{X|\theta}, \theta \in \Theta\}$ is a family of PMFs on A. Let $Y : A \to B$, and let there be functions $g_\theta : B \to \mathbb{R}_+$ and $h : B \to \mathbb{R}_+$ such that

$$P_{X|\theta}(a|\theta) = g_\theta(T(a))h(a), \qquad \text{for all } a \in A \text{ and } \theta \in \Theta.$$

Show that T is a sufficient statistic for θ.

8. Suppose that there are two coins with biases p and r respectively. One of these coins was picked and tossed n times. Let $X_1, X_2, \cdots, X_n$ be the outcomes of these tosses with $X_i = 1$ if the ith toss was a "Head" and 0 otherwise. To save space, your lab mate stored $Y = \sum_{i=1}^{n} X_i$ and deleted the exact sequence of toss outcomes. Identify the loss in relative entropy and interpret your answer in terms of being able to decide on the bias.