Homework 6: Hypothesis Testing - II

CR07: Selected Topics Information Theory (Fall 2019). ENS de Lyon

samir.perlaza@inria.fr – Deadline: November 11, 2019 at 11:59 p.m.

Consider two probability spaces $(\mathcal{X}^n, \mathscr{F}, P_0)$ and $(\mathcal{X}^n, \mathscr{F}, P_1)$, with $\mathcal{X}$ a discrete set such that $|\mathcal{X}| = t < \infty$, $n \in \mathbb{N}$ and $\mathscr{F}$ the largest σ-algebra containing $\mathcal{X}^n$. Consider also two probability spaces $(\mathcal{X}, \mathscr{G}, Q_0)$ and $(\mathcal{X}, \mathscr{G}, Q_1)$, where $\mathscr{G}$ is the largest σ-algebra containing $\mathcal{X}$. Within this context, consider the hypotheses

$$H_0: \qquad X \sim P_0 \tag{1a}$$

$$H_1: \qquad X \sim P_1, \tag{1b}$$

where for all $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ and for all $i \in \{0, 1\}$,

$$P_i\left(\{\boldsymbol{x}\}\right) = \prod_{t=1}^{n} Q_i\left(\{x_t\}\right). \tag{2}$$

For all $(i, j) \in \{0, 1\}^2$, let $C_{ij} > 0$ be the cost of accepting hypothesis H_i, when the ground-truth is hypothesis H_j. Let also $\pi_i \in [0, 1]$ be the ***à priori*** probability with which hypothesis H_i is accepted. Hence, given a test $T : \mathcal{X}^n \to \{0, 1\}$ the Bayesian risk is

$$R(T) = \pi_0\left(C_{00}P_0\left(\mathcal{X}_0\right) + C_{10}P_0\left(\mathcal{X}_1\right)\right) + (1-\pi_0)\left(C_{01}P_1\left(\mathcal{X}_0\right) + C_{11}P_1\left(\mathcal{X}_1\right)\right). \tag{3}$$

The test that minimizes the Bayesian risk is such that for all $\boldsymbol{x} \in \mathcal{X}^n$:

$$T_n(\boldsymbol{x}) = \mathbb{1}_{\{\boldsymbol{x} \in \mathcal{X}_1\}}, \tag{4}$$

where

$$\mathcal{X}_1 = \left\{\boldsymbol{x} \in \mathcal{X}^n : \frac{P_1\left(\{\boldsymbol{x}\}\right)}{P_0\left(\{\boldsymbol{x}\}\right)} > \frac{\pi_0\left(C_{10} - C_{00}\right)}{(1-\pi_0)\left(C_{01} - C_{11}\right)}\right\}. \tag{5}$$

1. Prove that the acceptance region of hypothesis H_1, i.e., the set $\mathcal{X}_1$ in (5), can be written as:

$$\mathcal{X}_1 = \left\{\boldsymbol{x} \in \mathcal{X}^n : D\left(P_{\boldsymbol{x}}||Q_0\right) - D\left(P_{\boldsymbol{x}}||Q_1\right) > \frac{1}{n}\log\left(\frac{\pi_0\left(C_{10} - C_{00}\right)}{(1-\pi_0)\left(C_{01} - C_{11}\right)}\right)\right\}. \tag{6}$$

2. Prove that the probability of false alarm $P_0(\mathcal{X}_1)$ satisfies:

$$\lim_{n\to\infty} -\frac{1}{n}\log P_0(\mathcal{X}_1) = D(P^*||Q_0), \tag{7}$$

where $P^* \in \triangle(\mathcal{X})$ satisfies

$$\begin{aligned} D(P^*||Q_0) = \min_{P\in\triangle(\mathcal{X})} \quad & D(P||Q_0) \\ \text{s.t.} \quad & D(P||Q_0) - D(P||Q_1) > \tfrac{1}{n}\log\left(\tfrac{\pi_0(C_{10}-C_{00})}{(1-\pi_0)(C_{01}-C_{11})}\right). \end{aligned} \tag{8}$$

3. What is the exact probability measure P^* in (7) ?

4. Prove that the probability of misdetection $P_1(\mathcal{X}_0)$ satisfies:

$$\lim_{n\to\infty} -\frac{1}{n}\log P_1(\mathcal{X}_0) = D(Q^*||Q_1), \tag{9}$$

where $Q^* \in \triangle(\mathcal{X})$ satisfies

$$\begin{aligned} D(Q^*||Q_1) = \min_{Q\in\triangle(\mathcal{X})} \quad & D(Q||Q_1) \\ \text{s.t.} \quad & D(Q||Q_0) - D(Q||Q_1) < \tfrac{1}{n}\log\left(\tfrac{\pi_0(C_{10}-C_{00})}{(1-\pi_0)(C_{01}-C_{11})}\right). \end{aligned} \tag{10}$$

5. What is the exact probability measure Q^* in (9) ?

6. Provide an upper-bound on the Bayesian risk in (3) of the optimal test in (4) in the case in which the number of observations is finite ($n < \infty$).

7. Provide an upper-bound on the Bayesian risk in (3) of the optimal test in (4) in the case in the asymptotic regime of observations ($n \to \infty$). HINT: provide the largest achievable exponent of the Bayesian risk.

8. What is the impact of the prior $(\pi_0, 1-\pi_0)$ in the upper-bounds on the Bayesian risk in both the asymptotic and non-asymptotic regimes?

9. Prove that under certain conditions, the biggest achievable exponent of the Bayesian risk is identical to the Chernoff information $C(Q_0, Q_1)$, that is,

$$C(Q_0, Q_1) = -\min_{\lambda\in[0,1]} \log\left(\sum_{a\in\mathcal{X}} (Q_0(a))^{\lambda}(Q_1(a))^{1-\lambda}\right). \tag{11}$$

10. Describe a few applications of the Chernoff Information in computer science, e.g., image segmentation, edge detection, etc.