1 Solution for HW1 / CSE254 / 2020

1.1 Setup

We observe a bit sequence $Y_1, Y_2, \ldots \in \{0, 1\}$. We generate a sequence of predictions $p_1, p_2, \ldots \in [0, 1]$ and would like to minimize the cumulative loss $L_T = \sum_{t=1}^{T} |p_t - Y_t|$.

We consider two scenarios:

1. We assume that the sequence is generated by IID coin flips where $P(Y_t = 1) = q$. q is unknown.
2. We don't assume anything about how the sequence is generated. We want the cumulative loss to not be much larger than the cumulative loss of the best (in hindsight) fixed prediction $h \in [0, 1]$

1.2 Analysis for IID scenario

Suppose that we are in the IID scenario, and that we know the bias q. It is not hard to convince yourself that if $q > 1/2$ the best prediction is $p = 1$ and if $q < 1/2$ the best prediction is $p = 0$. If $q = 1/2$ the expected loss is $1/2$ independent of the value of p

As we don't know the value of q, we need to estimate it, the minimal variance estimator is

$$\hat{q}_t = \frac{\sum_{i=1}^{t} Y_i}{t}$$

The best prediction, given this estimate, is

$$p_t = \begin{cases} 0 & \text{if } \hat{q}_t \le 1/2 \\ 1 & \text{if } \hat{q}_t > 1/2 \end{cases}$$