Exercise 4: Stochastic bandits with a continuum of arms (can be solved after Course #4)

Consider the setting of stochastic bandits with a continuum of arms indexed by $\mathcal{A} = [0,1]$. A bandit problem is given by a mean-payoff function f, which we assume to be continuous, thus bounded; for simplicity we consider $f : [0,1] \to [0,1]$. When the player picks arm $X_t \in [0,1]$ at round t, she gets a payoff $Y_t \in [0,1]$ drawn at random according to a distribution with expectation $f(X_t)$, conditionally to X_t. The question is to upper bound the regret defined as

$$R_T = T \max_{[0,1]} f - \mathbb{E}\left[\sum_{t=1}^{T} Y_t\right].$$

Consider the following two-stage methodology. We divide $[0,1]$ into the $K \geqslant 2$ regular intervals $[(i-1)/K,\, i/K]$, for $i \in \{1, \ldots, K\}$. An auxiliary algorithm picks an interval index $I_t \in \{1, \ldots, K\}$. An arm X_t is then drawn at random within bin I_t and a payoff Y_t is obtained.

We already considered this setting and strategy in an exercise in the lecture notes, with Lipschitz mean-payoff functions and the UCB algorithm.

In the present exercise, we rather assume that the mean-payoff function f is α–Hölder, for some $\alpha > 0$: there exists $L > 0$ such that for all $x, x' \in [0,1]$,

$$\bigl|f(x) - f(x')\bigr| \leqslant L|x - x'|^{\alpha}.$$

Also, we rather consider MOSS as the auxiliary algorithm; we recall that its distribution-free regret bound is $K + 45\sqrt{KT}$ against K–tuples of probability distributions over $[0,1]$.

1. Show that for a fixed number of bins $K \geqslant 2$ and a fixed horizon $T \geqslant 2$, the regret of the two-stage strategy above is upper bounded by $K + 45\sqrt{KT} + TL/K^{\alpha}$.
2. Explain how to pick K and which regret bound is obtained when T and α are known.
3. What could we do when α is known but T is unknown? Provide the (order of magnitude of the) corresponding regret bound.

Dealing with an unknown α is much more challenging, but was optimally solved by Hédi Hadiji in his PhD thesis, which he defended in December 2020.