**Exercice 1** (**"Anytime policy"**)
Nous nous plaçons dans le cadre de la prédiction séquentielle avec avis d'experts: les observations sont notées $y_1, y_2, \ldots$ A la date $t$, l'agent doit prendre une décision $d_t$ connaissant

- les décisions $f_{1,t}, \ldots, f_{i,t}, \ldots, f_{K,t}$ proposées par $K$ experts
- tout le passé (les $y_s$ pour $1 \le s < t$ et les décisions passées des experts $f_{i,s}$ pour $1 \le s < t$ et $1 \le i \le K$).

La perte instantanée encourue au temps $t$ est $\ell(y_t, d_t) \in [0,1]$. On considère la stratégie $S$ à poids exponentiels définie par: avec probabilité

$$p_{i,t} = \frac{\exp(-\eta_t L_{i,t-1})}{\sum_{k=1}^{K} \exp(-\eta_t L_{k,t-1})},$$

la décision choisie est $d_t = f_{i,t}$, où $\eta_t = \sqrt{\frac{4 \log K}{t}}$.
Le regret de $S$ après $n$ étapes est

$$R_n = \sum_{t=1}^{n} \ell(y_t, d_t) - \min_{1 \le i \le K} \sum_{t=1}^{n} \ell(y_t, f_{i,t}).$$

1) Soit $q = (q_1, \ldots, q_K)$ un vecteur de probabilités (i.e. $\forall i,\ q_i \ge 0$ et $\sum_i q_i = 1$). Soient $a_1, \ldots, a_K$ des nombres réels. *Montrer que la fonction $\phi : \eta \mapsto \frac{1}{\eta} \log\big(\sum_i q_i \exp(\eta a_i)\big)$ définie sur $\mathbb{R}_+^*$ est croissante.* Indication: on pourra introduire $\chi : u \mapsto u \ln(u) + 1 - u$, et écrire la dérivée de $\phi$ en fonction de $\sum_i q_i \chi\big(\exp(\eta a_i)/\sum_j q_j \exp(\eta a_j)\big)$, et étudier le signe de $\chi$.
2) En utilisant le résultat précédent et en adaptant la démonstration vue en cours, *montrer que pour tout $n$, on a $\mathbb{E}R_n \le \frac{\log K}{\eta_n} + \frac{\sum_{t=1}^{n} \eta_t}{8}$ et donc que $\mathbb{E}R_n \le \sqrt{n \log K}$.*
3) *Quel est l'avantage de cette stratégie de décision par rapport à celle vue en cours?*

**Exercice 2** (**Regret par rapport à la meilleure alternance d'experts**)
On se place dans le cadre de la prédiction séquentielle avec $K$ experts sur $n$ pas de temps. Soit $(y_t)_{t \ge 1}$ une suite de nombres de l'intervalle $[0;1]$. L'expert $k$ fournit au temps $t$ la prédiction $f_{k,t}$ dans l'intervalle $[0;1]$. Soit $\ell$ une fonction (de perte) de $[0;1] \times [0;1]$ dans $[0;1]$ telle que pour tout $y$, la fonction $\big[y' \mapsto \ell(y, y')\big]$ est convexe.

Soit $s \in \mathbb{N}$ tel que $1 \le s < n$. Pour $\vec{k} = (k_1, \ldots, k_s)$ un vecteur d'entiers de $\{1, \ldots, K\}$, et pour $\vec{t} = (t_0, \ldots, t_s)$ un vecteur d'entiers tels que $t_0 = 1 < t_1 < t_2 < \cdots < t_{s-1} < n < n+1 = t_s$, on définit l'expert $g_{(\vec{t},\vec{k})}$ par: pour tout $i \in \{1, \ldots, s\}$, pour tout $t \in [t_{i-1}, t_i[$, sa recommandation est $f_{k_i,t}$. L'expert $g_{(\vec{t},\vec{k})}$ alterne donc entre les différents experts (initiaux) suivant la période où il se trouve.

Soit $\mathcal{G} = \big\{g_{(\vec{t},\vec{k})} : \vec{k} \in \{1, \ldots, K\}^s$ et $\vec{t} \in \mathbb{N}^{s+1}$ tel que $t_0 = 1 < t_1 < t_2 < \cdots < t_{s-1} < n < n+1 = t_s\big\}$. Soit $R_n$ le regret par rapport au meilleur expert dans $\mathcal{G}$: $R_n$ est la différence entre la perte cumulée de l'agent et celle du meilleur expert dans $\mathcal{G}$.

1) *Proposer une stratégie pour laquelle on démontrera que le regret $R_n$ est borné par $\sqrt{\frac{n \ln(K^a n^b)}{2}}$ où on précisera l'expression de $a$ et $b$ en fonction de $s$.*

2) *Que retrouvons-nous pour $s = 1$?*