**Exercice (Bandit en milieu adverse)**
Dans le cadre du bandit en milieu adverse à $K \geq 2$ bras, on considère la stratégie vue en cours : la probabilité de tirer le bras $i$ au temps $t$ est

$$p_{i,t} = \frac{e^{\eta \sum_{s=1}^{t-1} \tilde{g}_{i,s}}}{\sum_{k=1}^{K} e^{\eta \sum_{s=1}^{t-1} \tilde{g}_{k,s}}}$$

avec $\tilde{g}_{i,s} = 1 - \frac{1-g_{I_s,s}}{p_{i,s}} \mathbb{1}_{I_s=i}$, $\eta > 0$, et $g_{i,s} \in [0,1]$ la récompense que donne le bras $i$ s'il est tiré au temps $s$. Introduisons $\ell_{i,t} = 1 - g_{i,t}$, $\tilde{\ell}_{i,t} = 1 - \tilde{g}_{i,t}$ et $\tilde{L}_{i,t} = \sum_{s=1}^{t} \tilde{\ell}_{i,s}$ pour tout $i \in \{1, \ldots, K\}$ et $t \in \{1, \ldots, n\}$.

1) *Pour $t \in \{1, \ldots, n\}$, montrer l'égalité*

$$\ell_{I_t,t} = \frac{1}{\eta} \ln \left( \sum_{i=1}^{K} p_{i,t} e^{-\eta(\tilde{\ell}_{i,t} - \sum_{k=1}^{K} p_{k,t} \tilde{\ell}_{k,t})} \right) - \frac{1}{\eta} \ln \left( \sum_{i=1}^{K} p_{i,t} e^{-\eta \tilde{\ell}_{i,t}} \right).$$

2) *Pour $i \in \{1, \ldots, K\}$ et $t \in \{1, \ldots, n\}$, donner une expression de $p_{i,t}$ faisant uniquement intervenir $\eta$ et les $\tilde{L}_{k,t-1}$, $k \in \{1, \ldots, K\}$.*

3) *Soit $X$ une variable aléatoire à valeurs positives et de carré intégrable. Montrer que*

$$\ln \left( \mathbb{E} e^{-X} \right) \leq -\mathbb{E} X + \frac{\mathbb{E}(X^2)}{2},$$

*et en déduire*

$$\ln \left( \mathbb{E} e^{-\eta(X - \mathbb{E} X)} \right) \leq \frac{\eta^2}{2} \mathbb{E}(X^2).$$

4) *Conclure des questions précédentes l'inégalité*

$$\ell_{I_t,t} \leq -\frac{1}{\eta} \ln \left( \frac{\sum_{i=1}^{K} e^{-\eta \tilde{L}_{i,t}}}{\sum_{i=1}^{K} e^{-\eta \tilde{L}_{i,t-1}}} \right) + \frac{\eta}{2} \sum_{i=1}^{K} p_{i,t} \tilde{\ell}_{i,t}^2.$$

5) Pour tout $i \in \{1, \ldots, K\}$, introduisons $L_{i,n} = \sum_{t=1}^{n} \ell_{i,t}$.

(a) Rappeler la loi de $I_t$ conditionnellement à tout le passé (i.e. pour $I_1$, …, $I_{t-1}$ fixés), et calculer l'espérance par rapport à cette loi de $\tilde{\ell}_{i,t}$ pour $i \in \{1, \ldots, K\}$.

(b) En déduire l'égalité $\mathbb{E} \tilde{L}_{i,n} = \mathbb{E} L_{i,n}$ pour $i \in \{1, \ldots, K\}$.

6) Ecrire $\sum_{i=1}^{K} p_{i,t} \tilde{\ell}_{i,t}^2$ uniquement en fonction de $\ell_{I_t,t}$ et $p_{I_t,t}$, puis en déduire l'inégalité $\mathbb{E} \sum_{i=1}^{K} p_{i,t} \tilde{\ell}_{i,t}^2 \leq K$ pour tout $t \in \{1, \ldots, n\}$.

7) Conclure des questions précédentes l'inégalité

$$\mathbb{E} \sum_{t=1}^{n} \ell_{I_t,t} - \min_{i \in \{1,\ldots,K\}} \mathbb{E} L_{i,n} \leq \frac{\eta n K}{2} + \frac{\ln K}{\eta}.$$

8) Quelle valeur de $\eta$ conseillez-vous de prendre ?