IN4221 Teoría de Juegos

P1) Considere la siguiente adaptación del algoritmo exponencial al caso de información incompleta. En cada etapa t solamente observamos la componente seleccionada $\ell^t_{i_t}$ y definimos la secuencia $\tilde{\ell}^t$ de vectores de *pseudo-costos* mediante

$$\tilde{\ell}^t_i = \begin{cases} 0 & \text{si } i \neq i_t \\ \ell^t_i / p^t_i & \text{si } i = i_t \end{cases}$$

así como los pseudo-costos acumulados $\tilde{L}^t_i = \sum_{\tau=1}^{t} \tilde{\ell}^\tau_i$. Dados $\epsilon > 0$ y $\beta > 0$ tomamos

$$p^t_i = \epsilon \tfrac{1}{n} + (1-\epsilon)\tfrac{w^t_i}{W^t}$$

con $w^t_i = \exp(-\beta \tilde{L}^{t-1}_i)$ y $W^t = \sum_{i=1}^{n} w^t_i$. Intuitivamente: con probabilidad ϵ escogemos i_t segun una ley uniforme y con probabilidad $(1-\epsilon)$ segun una ley exponencial.

(a) Pruebe que $\tilde{\ell}^t$ es un vector aleatorio cuya esperanza es $\mathbb{E}(\tilde{\ell}^t) = \ell^t$.
(b) Demuestre que este algoritmo tiene regret externo

$$R^T = T[\epsilon + (1-\epsilon)\gamma - 1] + \gamma(1-\epsilon)\tfrac{\ln(n)}{\beta}.$$

donde $\gamma = \frac{n\beta}{\epsilon}/[1 - \exp(-\frac{n\beta}{\epsilon})]$.
(c) Suponin=endo T conocido, determine los parámetros óptimos ϵ, β que minimizan el regret externo R^T.