21/03/2013

GdT Séquentiel
A. Garivier

Borne inférieure de Lai et Robbins

Récompenses binaires, K bras.

Preuve prise dans [Bubeck et Cesa-Bianchi 2012], origine : [Lai et Robbins '85]
extension [Burnetas & Katehakis '96].

Cadre : • $\theta = (p_a)_{1\leq a\leq K} \in \,]0,1[^K$
$p^* = \max_{1\leq a\leq K} p_a$

• $(X_{a,t})_{\substack{1\leq a\leq K \\ t\geq 1}}$ v.a. indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et t.q.

$\forall 1\leq a\leq K, \quad \forall t\geq 1, \quad X_{a,t} \sim \mathcal{B}(p_a)$.

• Stratégie π (règle dynamique d'allocation)
$\pi = (\pi_t)_{t\geq 1}$ où $\pi_t : (\{1,\dots,K\}\times\{0,1\})^{t-1} \longrightarrow \{1,\dots,K\}$.

On définit par récurrence : A_t = t-ième bras choisi

$N_a(t)$ = nb de fois que le bras a été tiré jusqu'à t

Y_t = t-ième récompense reçue

de la façon suivante : $\begin{cases} A_t = \pi_t(A_1, Y_1, \dots, A_{t-1}, Y_{t-1}) \\ N_a(t) = N_a(t-1) + \mathbb{1}_{\{A_t = a\}} \\ Y_t = X_{A_t, N_a(t)} \end{cases}$

Récompense cumulée : $S(t) = \sum_{s=1}^{t} Y_s$

$\mathcal{F}_t = \sigma(Y_1, \dots, Y_t)$

Regret : pour $T\in\mathbb{N}^*$,
$$R_T \triangleq T p^* - \mathbb{E}[S(t)]$$
$$= Tp^* - \sum_{t=1}^{T} \mathbb{E}\Big[\underbrace{\mathbb{E}[Y_t \mid \mathcal{F}_{t-1}]}_{= p_{A_t} = \sum_{a=1}^{K} p_a \mathbb{1}_{\{A_t=a\}}}\Big]$$
$$= Tp^* - \sum_{a=1}^{K} p_a \, \mathbb{E}\Big[\sum_{t=1}^{T} \mathbb{1}_{\{A_t=a\}}\Big]$$
$$= Tp^* - \sum_{a=1}^{K} p_a \, \mathbb{E}[N_a(T)]$$

D'où $R_T = \sum_{a : p_a < p^*} (p^* - p_a)\, \mathbb{E}[N_a(T)]$ car $T = \sum_{a=1}^{K} N_a(T)$

__Déf__ : la stratégie π est dite

- ⊛ consistante si $\forall \theta$, $R_T = o(T)$, i.e. $\frac{R_T}{T} \to 0 \iff \frac{\mathbb{E}[S_T]}{T} \to p^*$
- ⊛ efficace si $\forall \theta$, $\forall \eta > 0$, $R_T = o(T^\eta)$.

__Théorème__ [Lai & Robbins '85]

analogie avec la "super-efficacité" en statistique

Si π est une stratégie efficace, alors, $\forall \theta \in]0,1[^K$,

$$\liminf_{T\to+\infty} \frac{R_T}{\log T} \geq \sum_{a : p_a < p^*} \frac{p^* - p_a}{kl(p_a, p^*)}$$

où $kl(p,q) = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q} = KL(\mathcal{B}(p), \mathcal{B}(q))$.

C'est une csq de : $\forall a : p_a < p^*$,

$$\liminf_{T\to+\infty} \frac{\mathbb{E}[N_a(T)]}{\log T} \geq \frac{1}{kl(p_a, p^*)} \quad (*)$$

__Preuve de (*)__ : $\theta = (p_a)_{1 \leq a \leq K}$ fixé. Soit a tq $p_a < p^*$.

Soit $\varepsilon > 0$.

On va m.q. $\frac{\mathbb{E}[N_a(T)]}{\log T} \geq \frac{1-\varepsilon}{(1+\varepsilon)\, kl(p_a, p^*)} (1 - o(1))$

Pr cela, il suffit de mq

$$\mathbb{P}\left(N_a(T) < \frac{(1-\varepsilon)\log T}{(1+\varepsilon)\, kl(p_a, p^*)}\right) \xrightarrow[T\to+\infty]{} 0$$

car $\mathbb{E}[N_a(T)] \geq x\, \mathbb{P}(N_a(T) \geq x)$ pour $x = \frac{(1-\varepsilon)\log T}{(1+\varepsilon)\, kl(p_a, p^*)}$

On définit $\theta' = (p'_b)_{1 \leq b \leq K}$ t.q. $p'_b = p_b$ si $b \neq a$

$p'_a = p^* + \delta$ avec $\delta > 0$ tq.

$kl(p_a, p'_a) \leq kl(p_a, p^*)(1+\varepsilon)$

On pose

$$\widehat{kl}_s = \sum_{t=1}^{s} \log \frac{p_a X_{a,t} + (1-p_a)(1 - X_{a,t})}{p'_a X_{a,t} + (1-p'_a)(1-X_{a,t})}$$

$$= \sum_{t=1}^{s} \left(\mathbb{1}_{\{X_{a,t}=1\}} \log \frac{p_a}{p'_a} + \mathbb{1}_{\{X_{a,t}=0\}} \log \frac{1-p_a}{1-p'_a} \right)$$

__Prop__ : si A événement $\in \sigma(X_{a,1}, \dots, X_{a,s}, (X_{b,t})_{b \neq a, t \geq 1})$,

alors $\mathbb{P}'(A) = \mathbb{E}\left[\mathbb{1}_A \exp(-\widehat{kl}_s)\right]$.

Déf : $C_T = \left\{ N_a(T) < \frac{1-\varepsilon}{kl(p_a, p'_a)} \log T \right\} \cap \left\{ \widehat{kl}_{N_a(T)} \leq (1-\frac{\varepsilon}{2}) \log T \right\}$

$$\mathbb{P}'(C_T) = \mathbb{E}\left[\mathbb{1}_{C_T} \exp(-\widehat{kl}_{N_a(T)}) \right]$$

$$\geq \exp\left(-(1-\varepsilon/2)\log T\right) \mathbb{P}(C_T)$$

d'où $\mathbb{P}(C_T) \leq T^{1-\varepsilon/2} \, \mathbb{P}'(C_T)$

$$\leq T^{1-\varepsilon/2} \, \mathbb{P}'\left(N_a(T) < \frac{1-\varepsilon}{kl(p_a, p'_a)} \log T \right)$$

$$= T^{1-\varepsilon/2} \, \mathbb{P}'\left(T - N_a(T) > T - \frac{1-\varepsilon}{kl(p_a, p'_a)} \log T \right)$$

$$\overset{\text{Markov}}{\leq} T^{1-\varepsilon/2} \frac{\mathbb{E}'[T - N_a(T)]}{T - \frac{(1-\varepsilon)\log T}{kl(p_a, p'_a)}}$$

$\mathbb{E}'[T - N_a(T)] = o(T^{\varepsilon/4})$ car π est efficace.

$$\xrightarrow[T \to +\infty]{} 0$$

Il reste à mq. $\mathbb{P}\left(N_a(T) < \frac{1-\varepsilon}{kl(p_a, p'_a)} \log T \right) \xrightarrow[T \to +\infty]{} 0$

Soit $f_T = \frac{1-\varepsilon}{kl(p_a, p'_a)} \log T$

$$\mathbb{P}(C_T) \geq \mathbb{P}\left(N_a(T) < f_T \text{ et } \max_{s < f_T} \widehat{kl}_s \leq (1-\frac{\varepsilon}{2}) \log T \right)$$

$$= \mathbb{P}\left(N_a(T) < f_T \text{ et } \frac{kl(p_a, p'_a)}{(1-\varepsilon)\log T} \max_{s < f_T} \widehat{kl}_s \leq \frac{1-\varepsilon/2}{1-\varepsilon} kl(p_a, p'_a) \right)$$

Lemme : Z_n iid ≥ 0 d'espérance μ.

Si $\frac{1}{n} \sum_{i=1}^{n} Z_i \xrightarrow[n\infty]{p.s.} \mu$, alors $\frac{1}{n} \max_{s \leq n} \sum_{i=1}^{s} Z_i \xrightarrow[n\infty]{p.s.} \mu$.

D'après le lemme :

$$\mathbb{P}\left(\frac{1}{f_T} \max_{s \leq f_T} \widehat{kl}_s \leq \frac{1-\varepsilon/2}{1-\varepsilon} kl(p_a, p'_a) \right) \xrightarrow[T \to +\infty]{} 1$$

car $\frac{1}{s} \widehat{kl}_s = \frac{1}{s} \sum_{t=1}^{s} \left(\mathbb{1}_{\{X_{a,t}=1\}} \ln \frac{p_a}{p'_a} + \mathbb{1}_{\{X_{a,t}=0\}} \ln \frac{1-p_a}{1-p'_a} \right) \xrightarrow[s \to +\infty]{p.s.} p_a \ln \frac{p_a}{p'_a} + (1-p_a) \ln \frac{1-p_a}{1-p'_a}$

$$= kl(p_a, p'_a) \quad \blacksquare$$

Heuristique : H_0 : "$p = p_a$" versus H_1 : "$p = p^*$"

On veut $\alpha \leq 1/T$

Lemme de Stein : $\alpha_n \leq \alpha_0$

$$\frac{1}{n} \log \beta_n \geq -kl(p_a, p^*) \rightsquigarrow \beta_n \geq e^{-n\, kl(p_a, p^*)}$$

On ne peut avoir $\beta_n \approx \frac{1}{T}$

que si $n \approx \frac{\log T}{kl(p_a, p^*)}$