Essais cliniques adaptatifs
à l'aide des bandits stochastiques

21 juillet 2022

Le problème

- on dispose de N médicaments
- on souhaite déterminer lequel de ces médicaments est le plus efficace
- pour cela, on va donner ces médicaments à des patients (dans le cadre d'essais cliniques) tout en cherchant à attribuer le meilleur médicament le plus grand nombre de fois possible

Plan

1. Formalisation à l'aide des bandits stochastiques
2. Stratégie ϵ-greedy
3. Algorithme Upper Confidence Bound (UCB)
4. Application à des données réelles

Formalisation du problème

1. N médicaments
2. chacun a une certaine probabilité fixée de faire guérir le patient
3. on dispose d'un certain nombre de patients
4. But : donner le meilleur médicament au plus grand nombre de personnes possible

1. N actions : $0,1,\dots N-1$
2. chaque action a, lorsqu'elle est choisie, donne une récompense 1 avec une probabilité $q(a)$
3. le pas de temps $t = 0,1,\dots$ on note A_t et R_t
4. But : récolter un nombre maximal de récompenses

Stratégie ϵ-greedy

- pour chaque action a, on calcule la moyenne des récompenses obtenues après avoir choisi a. On la note $Q_t(a)$.
- **exploitation** : choisir l'action a qui a la plus grande moyenne $Q_t(a)$
- **exploration** : choisir une action au hasard

Stratégie ϵ-greedy

- pour chaque action a, on calcule la moyenne des récompenses obtenues après avoir choisi a. On la note $Q_t(a)$.
- **exploitation** : choisir l'action a qui a la plus grande moyenne $Q_t(a)$ avec une probabilité $1 - \epsilon$
- **exploration** : choisir une action au hasard avec une probabilité ϵ

Algorithme ϵ-greedy

pour chaque patient :

avec une probabilité $1 - \epsilon$:

donner le médicament A_t qu'on pense être le meilleur

sinon:

donner un médicament A_t aléatoire

observer l'état du patient R_t et mettre à jour Q_t

Résultats

$\texttt{actions\_q} = [0.9, 0.1, 0.8]$

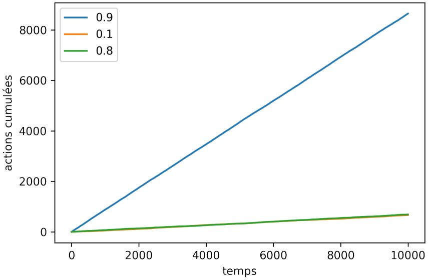


FIG.: ϵ-greedy ($\epsilon = 0.2$)

$\texttt{actions\_q} = [0.9, 0.1, 0.88]$

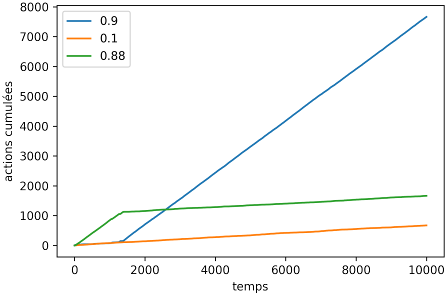
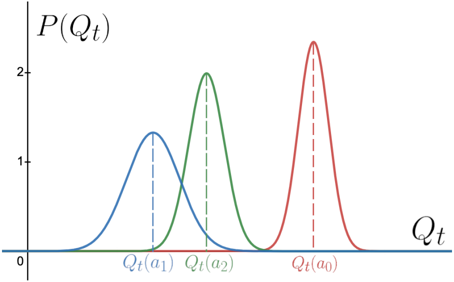
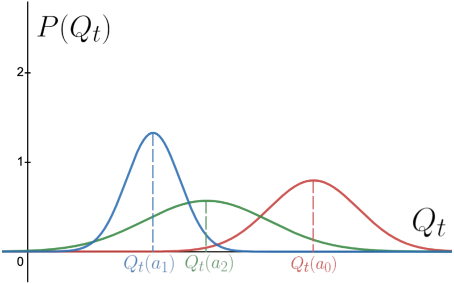


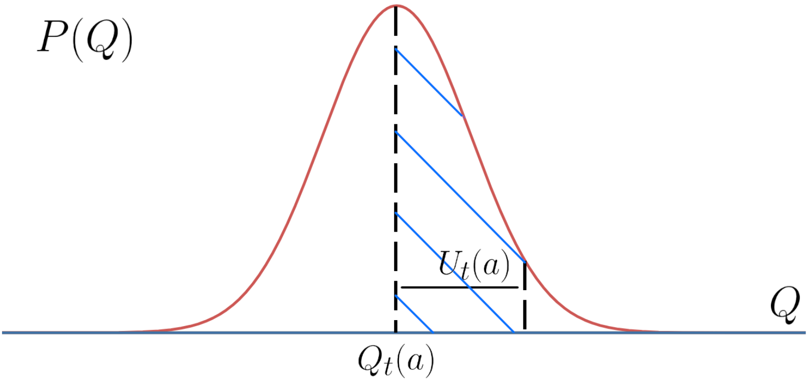
FIG.: ϵ-greedy ($\epsilon = 0.2$)

Incertitude





Upper Confidence Bound



$$P\Big(q(a) \geq Q_t(a) + U_t(a)\Big) \leq p$$

Calcul de $U_t(a)$

Inégalité de Hoeffding

Soit $(X_1,\dots,X_t)$ des variables aléatoires réelles mutuellement indépendantes à valeurs presque sûrement dans $\{0,1\}$, et $S_t = \frac{1}{t}\sum_{k=1}^{t} X_k$. Alors,

$$\forall u \in \mathbb{R}_+^*,\ P\big(\mathbb{E}[S_t] \geq S_t + u\big) \leq \mathrm{e}^{-2tu^2}$$

Calcul de $U_t(a)$

Inégalité de Hoeffding

Soit $(X_1,\ldots,X_t)$ des variables aléatoires réelles mutuellement indépendantes à valeurs presque sûrement dans $\{0,1\}$, et $S_t = \frac{1}{t}\sum_{k=1}^{t} X_k$. Alors,

$$\forall u \in \mathbb{R}_+^*,\ P\big(\mathbb{E}[S_t] \geq S_t + u\big) \leq \mathrm{e}^{-2tu^2}$$

Application aux récompenses données par une action a :

$$P\Big(q(a) \geq Q_t(a) + U_t(a)\Big) \leq \mathrm{e}^{-2N_t(a)U_t(a)^2}$$

Calcul de $U_t(a)$

Application aux récompenses données par une action a :

$$P\Big(q(a) \geq Q_t(a) + U_t(a)\Big) \leq \mathrm{e}^{-2N_t(a)U_t(a)^2}$$

on veut

$$\mathrm{e}^{-2N_t(a)U_t(a)^2} = p$$

d'où

$$U_t(a) = \sqrt{\frac{-\ln(p)}{2N_t(a)}}$$

Algorithme UCB

on choisit $p = t^{-4}$, d'où :

$$U_t(a) = \sqrt{\frac{2\ln(t)}{N_t(a)}}$$

alors :

$$\boxed{A_t \leftarrow \text{argmax}_a \; Q_t(a) + U_t(a)}$$

$\texttt{actions\_q} = [0.9, 0.1, 0.8]$

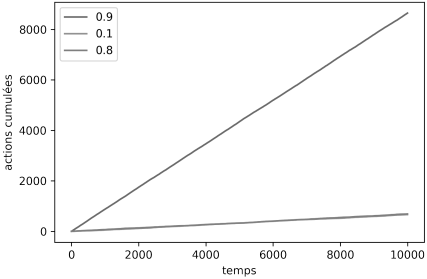


FIG.: ϵ-greedy ($\epsilon = 0.2$)

$\texttt{actions\_q} = [0.9, 0.1, 0.8]$

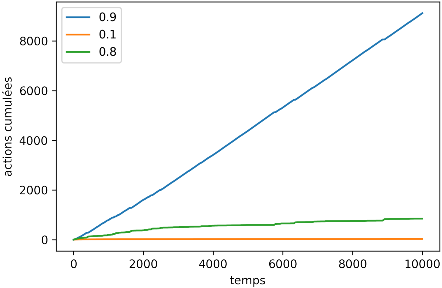


FIG.: UCB

$$\texttt{actions\_q} = [0.9, 0.1, 0.88]$$

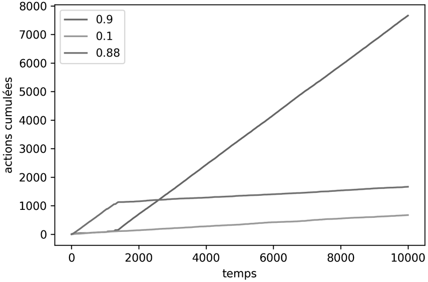


FIG.: ϵ-greedy ($\epsilon = 0.2$)

$$\texttt{actions\_q} = [0.9, 0.1, 0.88]$$

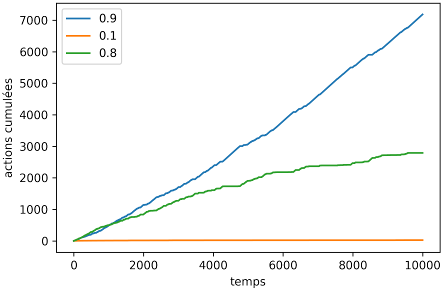


FIG.: UCB

Application : la base de données International Stroke Trials

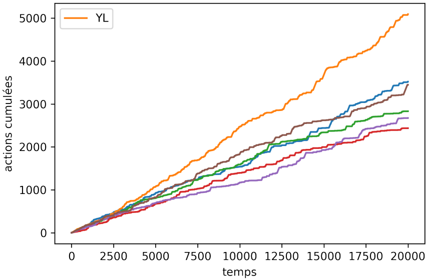
- près de 20 000 patients atteints d'AVC en phase aiguë
- 6 traitements (différentes doses d'aspirine et d'héparine)
- donnée sur la mort des patients à J+14

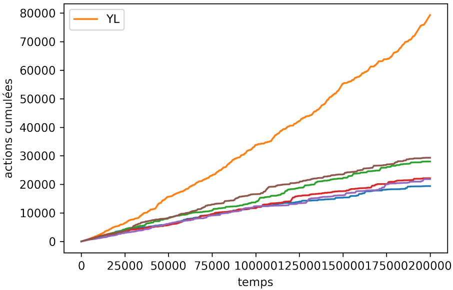
	SEX	AGE	RXASP	RXHEP	DDEAD
0	M	69	Y	N	N
1	M	76	N	L	N
2	F	71	Y	N	N
3	M	81	N	M	N
4	M	78	Y	M	N
...	...	...	...	...	...
19430	M	66	N	L	N
19431	M	75	N	M	N
19432	M	77	N	N	N
19433	F	87	N	N	Y
19434	M	54	Y	M	N

Résultats

- Données expérimentales :
  actions_q $= [0.889, 0.907, 0.896, 0.893, 0.892, 0.893]$
- But :
  - s'assurer de la convergence d'UCB vers le meilleur traitement (YL : aspirine $=$ Y, héparine $=$ L)
  - étudier le nombre d'attributions du meilleur traitement

Résultats





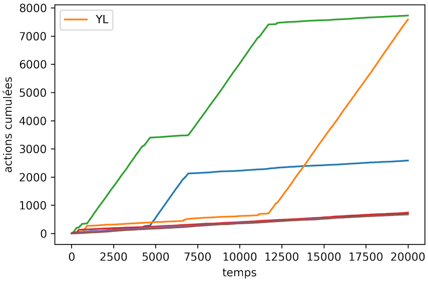
	classique	UCB
nombre de traitements YL attribués	3319	4665

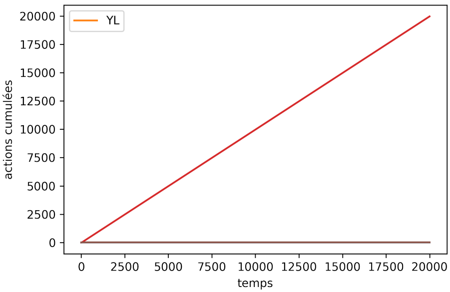
Ouverture

- utilisation concrète encore limitée
- adaptation pratique de la stratégie pour permettre une plus grande vitesse de distribution des médicaments aux patients
- essai clinique `DisCoVeRy` dans le cadre de la Covid-19, mené en France

Annexe : ϵ-greedy sur IST

avec $\epsilon = 0.2$





Annexe : cas concret d'un essai adaptatif

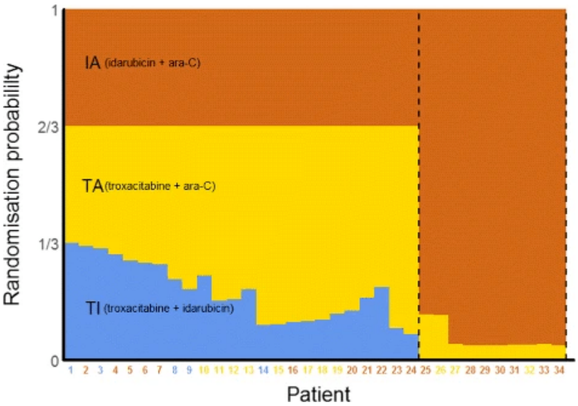
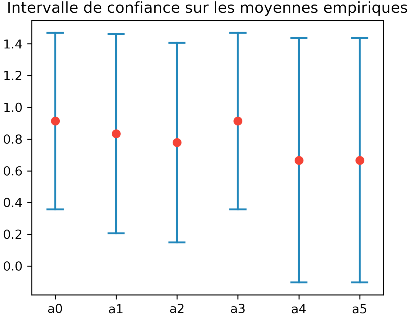


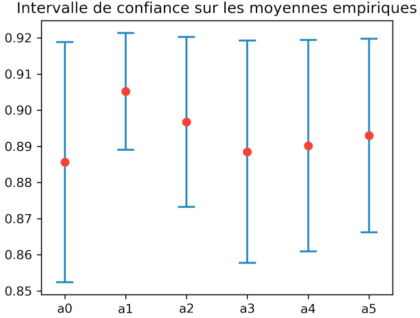
FIG.: PALLMAN ET AL., *BMC Medicine*, [Adaptive designs in clinical trials: why use them, and how to run and report them]

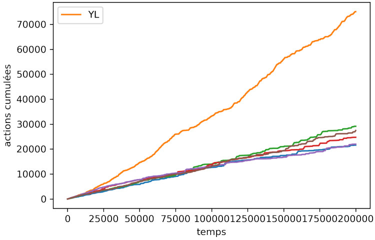
Annexe : incertitudes en fonction du temps

t = 1000



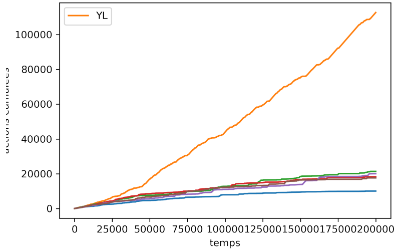
t = 190000

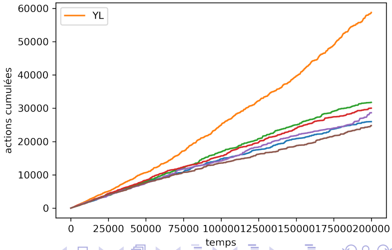




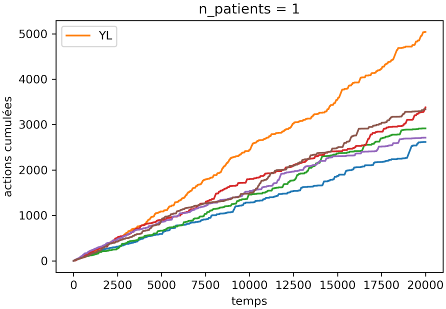
$$p = \frac{1}{t^2}$$

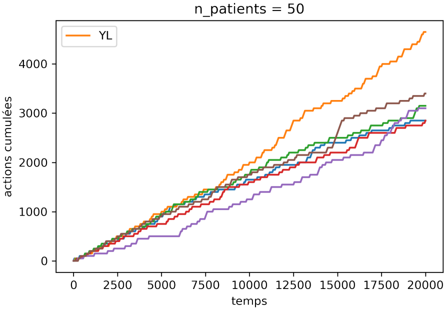
$$p = \frac{1}{t^{10}}$$



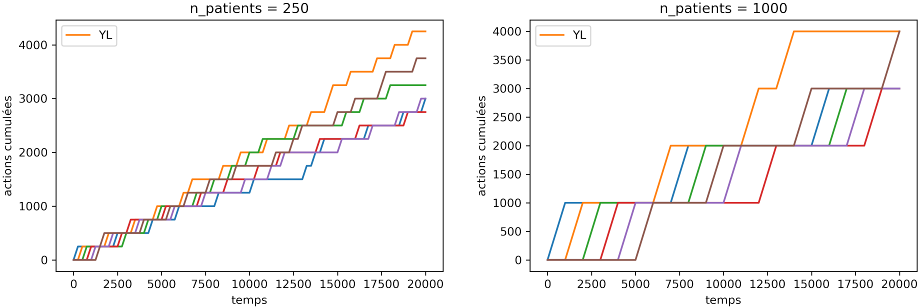


- on traite des groupes de patient et on fait les mises à jour *par paquets*





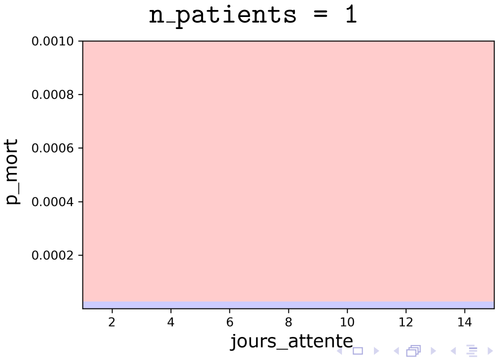
Annexe : UCB avec des groupes de patients



Annexe : comparaison entre les essais cliniques randomisés et adaptatifs

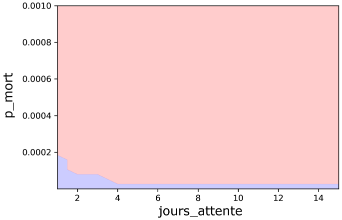
6 traitements, 20 000 patients

- jours_attente : temps nécessaire entre l'administration d'un traitement et l'observation de son résultat.
- p_mort : probabilité quotidienne qu'a un patient atteint de la maladie de mourir

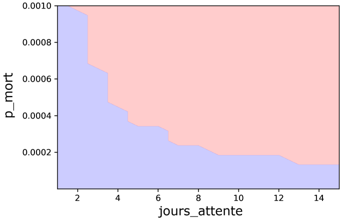


Annexe : comparaison entre les essais cliniques randomisés et adaptatifs

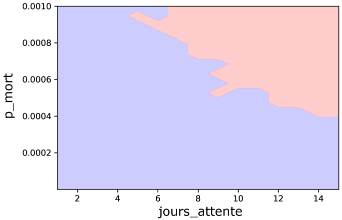
n_patients = 50



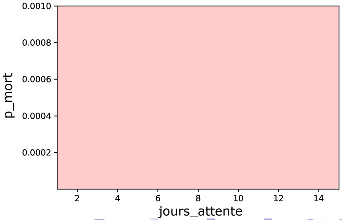
n_patients = 500



n_patients = 2000



n_patients = 10000



Lien avec le RL

- problème de bandit $\iff$ MDP avec un seul état
- chaque épisode est de longueur 1
- model-free

Des bandits ?

bandit manchot (*one-armed bandit*)

bandit à plusieurs bras (*multi-armed bandit*)