PROJET MATHÉMATIQUE

Juin 2025

Martin Baratte

Nils Villien

1A Informatique et réseaux

SOMMAIRE

1 Introduction

Dans ce rapport, nous allons comparer diverses stratégies de décisions adaptatives dans un contexte médical. L'expérience en question est un essai clinique porté sur 5 traitements A, B, C, D et E sur 1000 patients. La problématique est qu'on ne connaît pas les probabilité théoriques d'efficacité de ces traitements, d'où l'intérêt de cette démarche. Trois stratégies seront étudiés et testés numériquement avec python :

- La stratégie du choix uniforme
- La stratégie du choix de traitement ayant la meilleure probabilité d'efficacité observée
- La stratégie du traitement par inférence bayésienne sur les probabilités d'efficacité

2 Étude des stratégies

2.1 Stratégie du choix uniforme

On cherche à déterminer la moyenne théorique du nombre total de patients ayant un succès (i.e. un traitement efficace). Chaque traitement est choisi uniformément au hasard, avec $N=1000$ patients ainsi que des probabilité d'efficacité inconnues p_k, avec $k \in [\![1;5]\!]$ associés aux traitement A, B, C, D, E. Pour y parvenir, on devra construire un modèle probabiliste dont on devra calculer l'espérance.

2.1.1 Modélisation

Soit $X_n \in \{0;1\}$ le résultat (respectivement échec et succès) pour le patient n. On devra calculer :

$E(\sum_{n=1}^{1000} X_n)$ ou $1000 \cdot E(X_n)$ cherchons X_n

Pour déterminer la loi de X_n deux évènements sont à prendre en compte :

(1) Le tirage uniforme indépendant d'un traitement de probabilité $\frac{1}{5}$

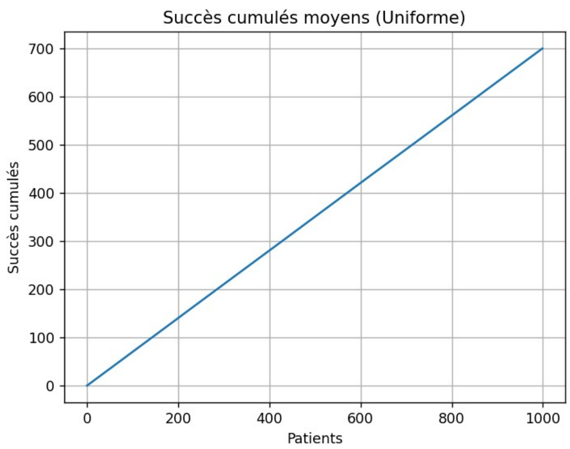
(2) L'efficacité du traitement qui est une variable aléatoire de bernoulli de paramètre p_k dépendant du traitement k administré

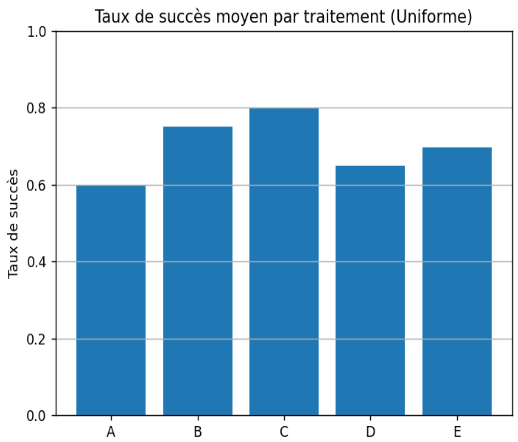
On en déduit que l'on peut exprimer $E(X_n)$ sous la forme d'une moyenne simple (ou arithmétique) :

$$E(X_n) = \frac{1}{5}\sum_{n=1}^{5} p_k$$

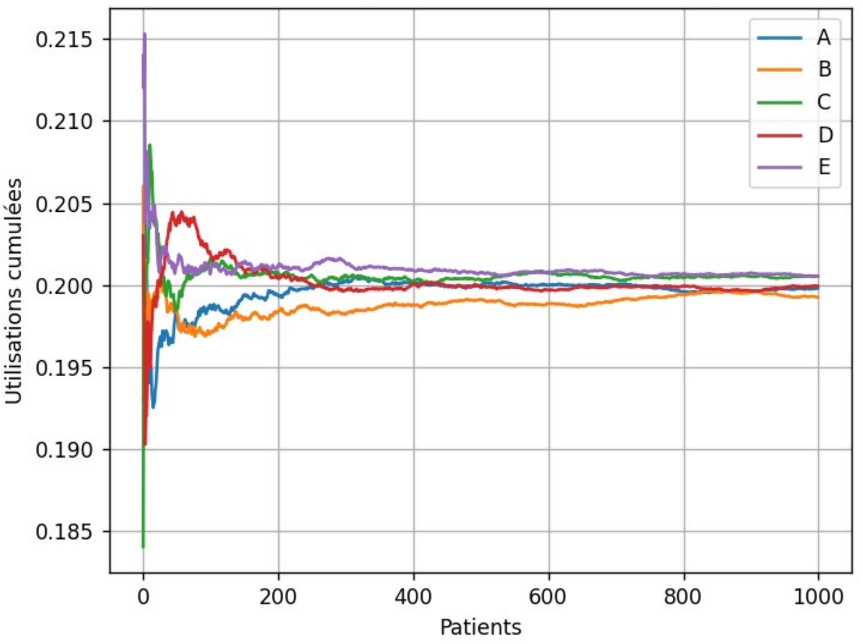
Ainsi $E(\sum_{n=1}^{1000} X_n) = 1000(\frac{1}{5}\sum_{n=1}^{5} p_k) = 200\sum_{k=1}^{5} p_k$

2.1.2 Résultats





Proportion des traitements utilisés au cours du temps (uniforme)



Succès total = 700/1000 = 70%

On remarque que la proportion des traitements tend vers 20% pour chacun des 5 traitements (proportion homogène attendue comme on utilise une stratégie uniforme)

Le taux de succès semble être le suivant : C=80% B=75%, E= 70%, D= 65% , A= 60%

Cela explique pourquoi le succès moyen est de 70%

La courbe des "succès cumulés moyen " reste affine dans les prochaines méthodes, il n'est donc pas pertinent de l'étudier

2.1.3 Bilan de la stratégie

On remarque que la moyenne théorique du nombre de succès (ici 700) est proportionnelle à la somme des probabilités d'efficacité des traitements. Cette stratégie est donc peu pertinente car les résultats dépendent uniquement de la moyenne globale des p_k mais pas de leur variance.

2.2 Stratégie du choix du traitement ayant la meilleure probabilité d'efficacité observée

On suppose que pour chaque $T_k \in \{A;B;C;D;E\}$, l'efficacité suit une loi de Bernoulli de paramètre inconnu $p_k \in (0;1)$. Les issues (succès = 1, échec = 0) sont des variables aléatoires $X_i^{(k)} \sim B(p_k)$, indépendantes et uniforme pour chaque k.

2.2.1 Application de la fonction log-vraisemblance

D'après la fonction de log-vraisemblance pour une loi de Bernoulli on a :

$$\begin{aligned}\log L(p_k|x_1,\dots,x_n) &= \sum_{i=1}^{N_k(n)} \log\left(p_k^{x_i}(1-p_k)^{1-x_i}\right)\\ &= \sum_{i=1}^{N_k(n)} \left(x_i \cdot \log(p_k) + (1-x_i)\log(1-p_k)\right)\\ &= \left(\sum_{i=1}^{N_k(n)} x_i\right)\log(p_k) + \left(\sum_{i=1}^{N_k(n)} (1-x_i)\right)\log(1-p_k)\\ &= S_k(n)\log(p_k) + (N_k(n)-S_k(n))\log(1-p_k)\end{aligned}$$

- Avec $N_k(n)$, le nombre total d'essais pour le traitement k jusqu'à l'instant n.
- Avec $S_k(n)$, le nombre de succès pour le traitement k jusqu'à l'instant n.

On maximise, ce qui nous donne $\widehat{p_{k,n}} = \dfrac{S_k(n)}{N_k(n)}$

On remarque que l'estimateur utilisé est une moyenne empirique :

$$\widehat{p_{k,n}} = \frac{S_k(n)}{N_k(n)} = \frac{1}{N_k(n)} \sum_{i=1}^{N_k(n)} X_i^{(k)}$$

ainsi $E(\widehat{p_{k,n}}) = p_k$ donc $\widehat{p_{k,n}} \underset{n \to +\infty}{\to} p_k$, de ce fait l'estimateur est fortement croissant (ce qui assure que les estimations se stabilisent à long terme).

2.2.2 Implémentation algorithmique

(1) Donner à chaque traitement un nombre minimal d'essais pour éviter d'avoir un estimateur non défini (à cause d'une division par zéro).

(2) À chaque instant n

- Calculer tous les $\widehat{p_{k,n}}$
- Choisir le traitement T_k avec la plus grande $\widehat{p_{k,n}}$
- Observer un nouveau résultat (succès ou échec), mettre à jour S_k et N_k

<u>Remarque :</u> Il est possible qu'un traitement initialement chanceux soit choisi en boucle, ce qui implique un biais de sélection précoce

2.2.3 Stratégie de décision

Au vu du constat précédent il est préférable d'intégrer la stratégie de décision Epsilon-Greedy[1] qui nous permettra d'alterner entre phase d'exploitation et d'exploration :

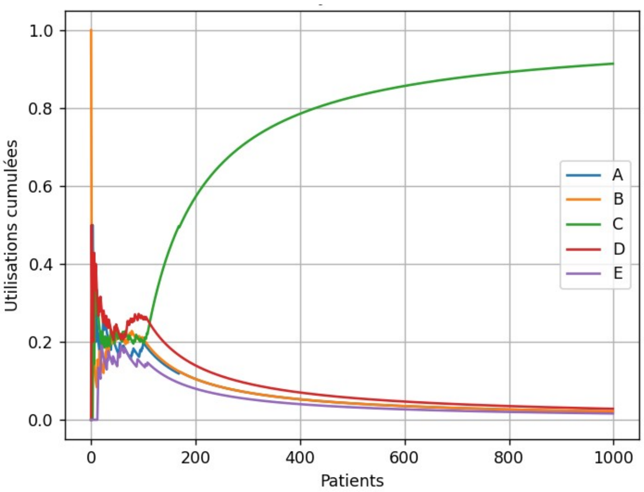
- (exploitation) avec une probabilité $1-\varepsilon$, on choisit le traitement avec le meilleur $\widehat{p_{k,n}}$
- (exploration) avec une probabilité ε, on choisit un traitement uniformément au hasard

On choisira $\varepsilon = 0{,}05$

Ainsi tous les traitements continuent à être testés même si leur estimation est très faible

2.2.4 Résultats

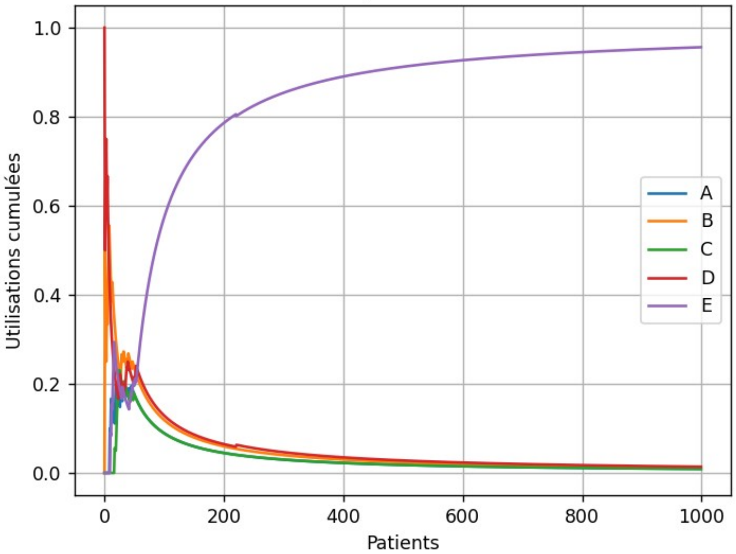
Proportion des traitements utilisés au cours du temps (Choix du meilleur)



Sur une population de 1000 patients, la meilleure distribution que j'ai obtenue est celle-ci avec 794/1000 succès (proche de 80%)

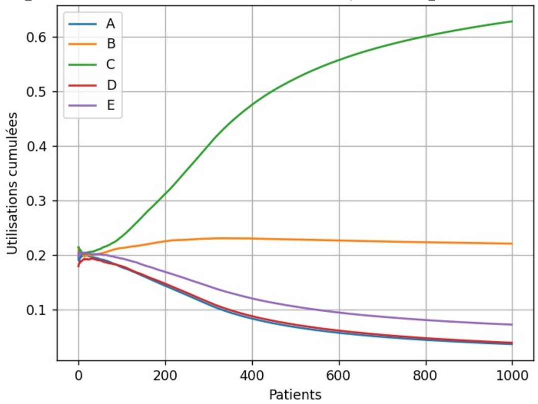
Cela est dû au fait que le traitement C a été choisi comme le meilleur traitement avec un nombre de tests uniformes faibles (environ 10% de la population totale) et les traitements moins bons ne sont plus réutilisé par la suite.

On peut aussi avoir une estimation préalable qui n'est pas bonne et qui choisit E comme le meilleur puis change d'avis et commence à utiliser B quand le % de réussite en essai de B finit par rattraper le % effectif de E et commence à utiliser C quand le % de réussite effectif de B et E passe en dessous du % en essai de C (ici C n'a pas eu de chance malgré presque 200 essais)

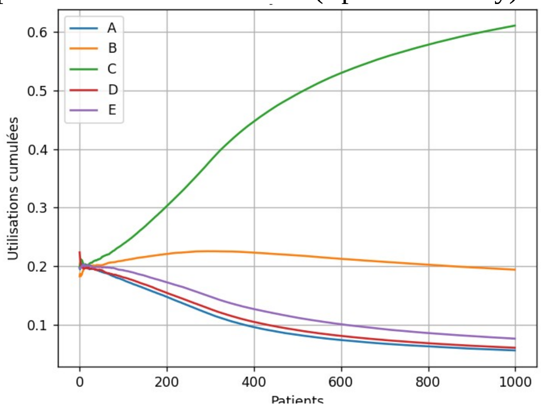


C'est pour lutter contre ça que on met en place stratégie de décision Epsilon-Greedy

Proportion sur 1000 iterations (sans Epsilon-Greedy)



Proportion sur 1000 iterations (Epsilon-Greedy)



On remarque une diminution des extrêmes sur un grand nombre d'itérations, cela s'explique par la diminution d'occurrence de "pas de chance de chance" mais aussi des "beaucoup de chance "

Le taux de réussite est donc similaire avec ou sans Epsilon-Greedy : 766 (sans) contre 767 (avec) sur 1000 itérations. Ce qui reste largement supérieur à la stratégie uniforme.

2.2.5 Bilan de la stratégie

Grâce à la méthode du maximum de vraisemblance on a prouvé que :

- $\widehat{p_{k,n}} = \frac{S_k}{N_k}$ est la moyenne empirique des succès
- $\widehat{p_{k,n}}$ est sans biais et converge vers p_k
- la stratégie s'appuie sur des mises à jour progressives, mais doit comporter de l'exploration pour corriger les biais dus à la variabilité initiale.

2.3 Stratégie du choix du traitement par inférence bayésienne sur les probabilités d'efficacité

Dans cette partie on utilisera une approche bayésienne dans le but de modéliser et actualiser nos connaissances sur les probabilités d'efficacité p_k des traitements, en prenant en compte l'incertitude a posteriori. On choisira le traitement le plus prometteur à partir d'un tirage aléatoire sur la loi de probabilité a posteriori de p_k.

Dans cette approche :

- On considère p_k comme une variable aléatoire
- On commence avec une loi à priori sur p_k, puis on la met à jour avec les observations pour obtenir une loi a posteriori.

2.3.1 Loi Bêta comme loi à priori

Pour un paramètre de Bernoulli p_k, la loi Bêta est un choix naturel comme loi à priori :

$$p_k \sim Beta(\alpha_k, \beta_k)$$

où α_k et β_k sont les pseudo-comptages de succès et d'échecs pour le traitement k et on a aucune connaissance au préalable : $\alpha_k = \beta_k = 1$ (loi uniforme sur $[0;1]$).

2.3.2 Loi Bêta comme loi à posteriori

L'efficacité d'un traitement suit une loi binomiale, de ce fait on obtiendra également une loi Bêta à posteriori.

- On observera : $S_k(n)$ succès et $F_k(n)$ échecs
- Loi à posteriori : $p_k|données \sim Beta(\alpha_k + S_k(n),\ \beta_k + F_k(n))$ car si $X_1, \dots, X_n \sim B(p_k)$ et $p_k \sim Beta(\alpha, \beta)$, alors $p_k|X_1, \dots, X_n \sim Beta(\alpha + S, \beta + (n - S))$

2.3.3 Stratégie de décision

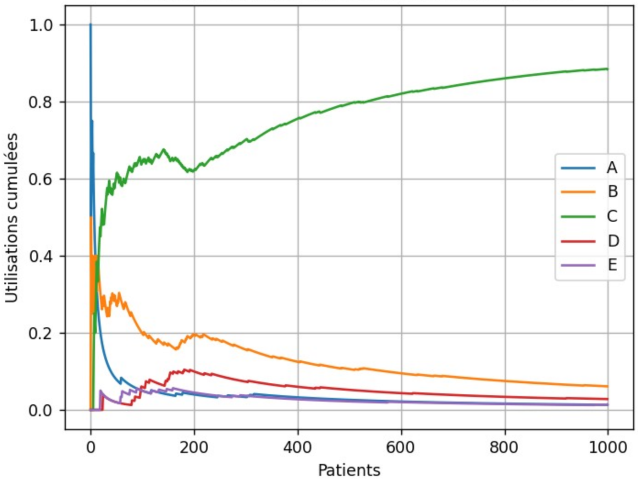
Pour chaque patient :

(1) pour chaque traitement k tirer une valeur $p_k' \sim Beta(\alpha_k + S_k, \beta_k + F_k)$

(2) sélectionner le traitement ayant le plus grand p_k'.

2.3.4 Résultats

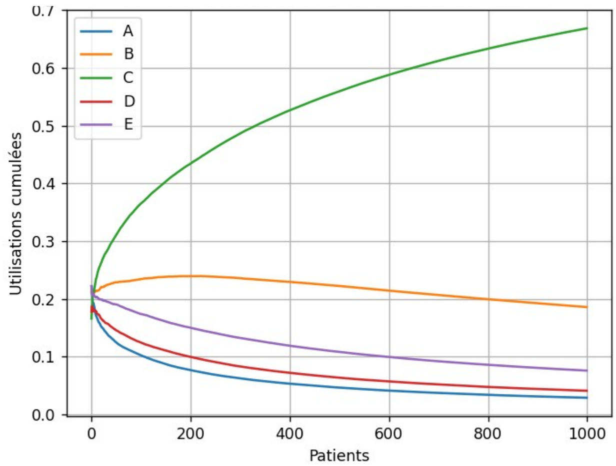
Proportion des traitements utilisés au cours du temps (Bayésienne)



Le traitement C est identifié comme le meilleur traitement très rapidement sans passer par une longue phase d'essai, ce qui permet d'atteindre une proportion de traitements par C très élevée (environ 85%)

Et donc un taux de réussite global important : ici 794/1000 (soit presque 80% qui est le taux maximum)

Proportion des traitements au cours du temps sur 1000 iterations (Bayesien)



Sur un grand nombre d'itérations, on obtient des proportions qui sont légèrement plus intéressantes qu'Epsilon-Greedy sans avoir besoin d'une phase d'essai

Le taux de réussite global est de 770/1000, soit 77 % ce qui est plus élevé que les autres méthodes

2.3.5 Bilan de la stratégie

On remarque que plus un traitement est testé, plus sa loi Bêta se concentre autour de sa véritable valeur d'efficacité, donc le choix gagne en stabilité. On retient que l'approche par inférence bayésienne permet de trouver un équilibre dynamique entre phase d'exploration et d'exploitation.

3 Conclusion

La stratégie la plus pertinente semble être le recours à l'inférence bayésienne, car elle offre les meilleurs résultats tout en évitant une phase d'essai préalable potentiellement discutable sur le plan éthique. En effet, cette approche ne nécessite pas de traiter un groupe de patients à titre expérimental avant de généraliser le traitement à l'ensemble de la population concernée.

4 Sources

[1] https://www.geeksforgeeks.org/epsilon-greedy-algorithm-in-reinforcement-learning/