ENSIIE, Évry, France

Mars 21, 2023

Projet mathématique - Partie I

Introduction aux chaînes de Markov, simulation et estimation par Monte Carlo

---

- Pour le projet mathématique, vous utiliserez Python.
- Vous devez rendre un notebook (format .ipynb) contenant le code, les simulations et les explications mathématiques de vos calculs.
- **Veuillez publier votre notebook sur GitLab avant le 4 avril à 23h59. Les notebooks téléchargés après cette date et heure recevront automatiquement une note de zéro.**

---

1. *(Simulation d'une variable aléatoire discrète, 3/20)* Supposons que $p_1, \ldots, p_n$ sont des nombres réels positifs tels que $\sum_{i=1}^{n} p_i = 1$. Soient $x_1 < x_2 < \cdots < x_n$ des nombres réels arbitraires. Considérons une variable aléatoire discrète X prenant des valeurs dans l'ensemble $\{x_1, x_2, \ldots, x_n\}$ telle que

$$\mathbb{P}(X = x_i) = p_i, \quad i = 1, \ldots, n.$$

Pour simuler une réalisation y de la variable aléatoire X, on peut suivre la procédure suivante :

(a) Simuler u, une réalisation d'une variable aléatoire suivant une loi uniforme U sur $[0, 1]$, c'est-à-dire $U \sim \mathrm{Unif}(0, 1)$.

(b) Trouver le premier indice j^* tel que $u \leq \sum_{i=1}^{j^*} p_i$, c'est-à-dire

$$j^* = \min \left\{ j : u \leq \sum_{i=1}^{j} p_i \right\}.$$

(c) Définir $y = x_{j^*}$.

En utilisant l'algorithme décrit ci-dessus, écrivez une **FONCTION** $y = \mathrm{sim\_dis}(p, x, M)$ qui prend en **ENTRÉES**

- $x = (x_1, \ldots, x_n)$: *le vecteur des valeurs possibles pour* X *;*
- $p = (p_1, \ldots, p_n)$: *le vecteur des probabilités correspondantes, c'est-à-dire la loi de* X *;*
- M : *le nombre désiré de simulations de la variable* X *;*

et renvoie en **SORTIE** un vecteur $y = (y_1, \ldots, y_M)$ dont les composantes sont M simulations indépendantes de la variable aléatoire X.

2. *(Simulation d'une marche aléatoire, 2/20)* Considérons $X_1, X_2, \ldots, X_N$ des variables aléatoires i.i.d. telles que

$$\mathbb{P}(X_i = 1) = p \in (0, 1), \quad \mathbb{P}(X_i = -1) = q = 1 - p, \quad i = 1, \ldots, N.$$

Soit S_n la marche aléatoire (entre 0 et N) définie par

$$S_n = S_0 + X_1 + \ldots + X_n, \quad 1 \leq n \leq N.$$

Supposons que $N = 100$, $S_0 = 5$ et $p = 0.5$ (marche aléatoire symétrique). Simulez $M = 1000$ trajectoires indépendantes de cette marche aléatoire (entre 0 et N). Tracez 10 de ces trajectoires sur un seul graphique. *Conseil :* À chaque étape i, utilisez la fonction $\mathrm{sim\_dis}((q, p), (-1, 1), M)$ de la partie précédente pour simuler M copies indépendantes de X_i.

3. *(Estimation d'une probabilité par la méthode de Monte Carlo pour une marche aléatoire, 3/20)* En utilisant les simulations de la partie précédente, estimez la probabilité $\mathbb{P}(S_N \geq 5)$ en calculant

$$\frac{\#\ (\text{trajectoires telles que } S_N \geq 5)}{M}.$$

Quel est l'intervalle de confiance à 95% de votre estimation ? La vraie valeur de la probabilité se trouve-t-elle dans l'intervalle de confiance ?

4. *(Simulation du mouvement d'un étudiant sur le plateau de jeu, 3/20)* Maintenant, supposons que $(X_i)_{i\geq 0}$ est une chaîne de Markov modélisant les mouvements d'un des étudiants sur le plateau de jeu. La chaîne de Markov prend des valeurs dans $\{1, 2, \ldots, 10\}$ et a une matrice de probabilité de transition $P = (p_{ij})_{1\leq i,j\leq 10} \in [0, 1]^{10\times 10}$. Supposez que $X_0 = 1$ et choisissez aléatoirement une matrice de probabilité de transition avec des coefficients strictement positifs, c'est-à-dire que $p_{ij} > 0$ pour tous les $i, j \in \{1, 2, \ldots, 10\}$. Prenez $N = 3$ et $M = 1000$. En utilisant la fonction sim_dis de la première partie, simulez M réalisations de cette chaîne de Markov pour $i = 0, 1, \ldots, N$ et tracez un histogramme des valeurs X_3 (c'est-à-dire la position après trois itérations).

5. *(Estimation de Monte Carlo de la probabilité d'être dans une zone donnée après un certain nombre d'itérations, 3/20)* En utilisant les simulations de la partie précédente, estimez la probabilité $\mathbb{P}(X_3 = 10)$ en calculant

$$\frac{\#\ (\text{trajectoires telles que } X_3 = 10)}{M}.$$

Quel est l'intervalle de confiance à 95% de votre estimation ? La vraie valeur de la probabilité se trouve-t-elle dans l'intervalle de confiance ?

6. *(3/20)* Répétez les deux parties précédentes avec $N = 100$.

7. *(Simulation du mouvement de plusieurs étudiants et du monstre, 3/20)* Nous modélisons les mouvements de N_e étudiants sur le plateau de jeu avec N_s zones via N_e chaînes de Markov indépendantes $X^{(i)}$, $i = 1, \ldots, N_e$, prenant des valeurs dans $\{1, 2\ldots, N_s\}$, en commençant toutes dans la zone 1, et avec une matrice de transition commune $P \in [0, 1]^{10\times 10}$. Nous modélisons également les mouvements du monstre avec une chaîne de Markov indépendante Y, avec $Y_0 = N_s$ et une matrice de transition Q. Écrivez une fonction death_time(N_e, N_s, P, Q, N) qui calcule le temps de survie de la population d'étudiants sur une simulation des mouvements jusqu'à l'itération N. Choisissez aléatoirement une matrice de transition avec des coefficients strictement positifs P et une matrice de transition arbitraire Q. Prenez $N_e = 5$, $N_s = 10$ et $N = 100$. En utilisant la fonction death_time(N_e, N_s, P, Q, N), tracez un histogramme du temps de survie de la population d'étudiants sur $M = 100$ simulations. Calculez le temps de survie moyen sur les $M = 100$ simulations.