

FEUILLE DE TRAVAUX PRATIQUES - PYTHON #3 MÉTHODE DE MONTE CARLO

Emeline LUIRARD

1 De la Loi des Grands Nombres à la méthode de Monte-Carlo

Soit E un sous-ensemble de \mathbb{R}^d , $d \geq 1$.

Théorème 1 (Loi des Grands Nombres). *Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. i.i.d., à valeurs dans E , ayant un moment d'ordre 1. Alors*

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[X_1].$$

La LGN est dite faible lorsque la convergence a lieu en probabilité. La LGN est dite forte lorsque la convergence a lieu ps ou dans L^p .

Autrement dit, la moyenne arithmétique converge vers la moyenne probabiliste. Le résultat de convergence peut également se réécrire, dans le cas où les (X_n) suivent une loi de densité f ,

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow \infty]{} \int_E x f(x) dx.$$

ou encore, avec une fonction mesurable $g : E \rightarrow \mathbb{R}$ telle que $g(X_1) \in L^1$.

$$\frac{1}{n} \sum_{k=1}^n g(X_k) \xrightarrow[n \rightarrow \infty]{} \int_E g(x) f(x) dx.$$



Il est important d'avoir la condition d'intégrabilité !

Cela va nous permettre d'approcher la valeur d'une intégrale. C'est ce qu'on appelle **la méthode de Monte Carlo**. En terme d'estimation statistique, la méthode de Monte Carlo nous fournit un estimateur $\frac{1}{n} \sum_{k=1}^n g(X_k)$ de l'intégrale $\int_E g(x) f(x) dx$, qui est **fortement consistant et sans biais**.

Prenons un premier exemple : on se donne $-\infty < a < b < +\infty$ et une fonction $g : [a, b]^d \rightarrow \mathbb{R}$ intégrable. On souhaite approcher (estimer) l'intégrale

$$I(g) := \int_{[a,b]^d} g(x) dx.$$

Remarquons que si (X_n) est une suite de v.a. i.i.d. de loi uniforme sur $[a, b]^d$, alors l'estimateur

$$I_n(g) := \frac{1}{n} (g(X_1) + \dots + g(X_n))$$

converge presque sûrement et dans \mathbb{L}^1 vers la limite $\mathbb{E}[g(X_1)] = \int_{[a,b]^d} g(x) \frac{1}{(b-a)^d} dx = \frac{1}{(b-a)^d} I(g)$.

Ainsi, $(b-a)^d I_n(g)$ est un estimateur fortement consistant de $I(g)$.



Quand on intègre sur un intervalle de mesure différente de 1, on n'oublie pas de renormaliser :

$$(b-a)^d \mathbb{E}[g(X_1)] = \int_{[a,b]^d} g(x) dx.$$

D'où l'importance de se placer sur un compact ici...

💡 On peut considérer d'autres lois que la loi uniforme sur un pavé !

Le premier exemple connu d'application de cette méthode remonte au dix-huitième siècle et au fameux problème de l'aiguille de Buffon qui permet un calcul approché de π . Le nom "méthode de Monte-Carlo" est un nom de code qu'a utilisé Ulam alors qu'il développait ces méthodes avec von Neuman, Fermi et Metropolis, dans le laboratoire de Los Alamos, où se préparait la première bombe à hydrogène.

Exercice 1. Premiers exemples

À l'aide de la méthode Monte-Carlo, calculer des approximations des intégrales suivantes :

$$\int_0^1 4\sqrt{1-x^2} dx, \quad \int_{[-1,1]^3} \mathbb{1}_{\{x^2+2y^2+3z^2 \leq 1\}} dx dy dz.$$

Exercice 2. Volume de la boule unité

On considère la boule unité \mathbb{B}_d dans \mathbb{R}^d , i.e. $\mathbb{B}_d = \{(x_1, \dots, x_d) \in \mathbb{R}^d, x_1^2 + x_2^2 + \dots + x_d^2 \leq 1\}$.

1. Écrire une fonction qui, en entrée, prend deux entiers n et $d \geq 2$ et en sortie, donne une approximation du volume de la boule euclidienne basée sur un n -échantillon de variables uniformes dans le cube $[-1, 1]^d$.
2. Comparer l'écart à la valeur théorique en fonction de d et n .

Aide : Volume de la boule unité de dimension d : $V(d) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2} + 1)}$

2 La méthode de Monte carlo VS des méthodes déterministes

Nous allons voir que contrairement aux méthodes déterministes, la méthode de Monte Carlo est certes plus lente mais ne dépend pas de la régularité de l'intégrande et peu de la dimension d .

2.1 Étude de la méthode de Monte Carlo

Théorème 2 (TCL). Soit $(X_n)_n$ une suite de v.a. i.i.d. à valeurs dans E , admettant un moment d'ordre 2. Alors

$$\frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}[X_1] \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(0, 1),$$

où $\sigma^2 = \text{Var}(X_1)$.

D'après la définition de vitesse de convergence d'un estimateur, le TCL nous indique que la **vitesse de convergence de la méthode Monte Carlo est $\frac{\sqrt{n}}{\sigma}$** . Ce théorème donne également un moyen d'obtenir des **intervalles de confiance asymptotiques afin de contrôler l'erreur de la méthode**. D'après le théorème central limite, si (x_1, \dots, x_n) est une observation de (X_1, \dots, X_n) , l'intervalle ci-dessous est un intervalle de confiance asymptotique de niveau 95% pour $\mathbb{E}[g(X)]$:

$$\left[\frac{1}{n} \sum_{k=1}^n g(x_k) - \frac{1.96 \sigma}{\sqrt{n}}, \frac{1}{n} \sum_{k=1}^n g(x_k) + \frac{1.96 \sigma}{\sqrt{n}} \right].$$

Remarquez que ces intervalles sont spécifiques à l'échantillon observé. La variance σ^2 est donnée par $\sigma^2 = \mathbb{E}[g^2(X)] - \mathbb{E}[g(X)]^2$. Il est fort possible que cette quantité soit inconnue. La stratégie est alors de remplacer la variance théorique par son estimateur empirique

$$\sigma_n^2 := \frac{1}{n} \sum_{k=1}^n \left(g(X_k) - \frac{1}{n} \sum_{i=1}^n g(X_i) \right)^2.$$

D'après le lemme de Slutsky, si (x_1, \dots, x_n) est une réalisation de (X_1, \dots, X_n) , alors

$$\left[\frac{1}{n} \sum_{k=1}^n g(x_k) - \frac{1.96 \sigma_n}{\sqrt{n}}, \frac{1}{n} \sum_{k=1}^n g(x_k) + \frac{1.96 \sigma_n}{\sqrt{n}} \right].$$

est encore un intervalle asymptotique au niveau 95% de $\mathbb{E}[g(X)]$.

Exercice 3. À l'aide de la méthode de Monte-Carlo, écrire un programme qui

1. calcule une approximation de l'intégrale $I = \int_{\mathbb{R}} \cosh(x) e^{-x^2} dx$,
2. calcule l'estimateur empirique de la variance associé,
3. trace sur le même graphique l'estimateur de l'intégrale et l'intervalle de confiance associé en fonction du nombre de données.

En notant $\epsilon_n := I_n(g) - I(g)$,

$$\mathbb{P}(|\epsilon_n| < 0.01) = \mathbb{P}\left(-0.01 \frac{\sqrt{n}}{\sigma} \leq \frac{\sqrt{n}}{\sigma} (I_n(g) - I(g)) \leq 0.01 \frac{\sqrt{n}}{\sigma}\right) \rightarrow \mathbb{P}\left(-0.01 \frac{\sqrt{n}}{\sigma} \leq \mathcal{N}(0, 1) \leq 0.01 \frac{\sqrt{n}}{\sigma}\right).$$

Afin d'avoir un niveau d'erreur inférieur à 0.01, avec un niveau de confiance de 95 %, on trouve, à l'aide des tables, qu'il suffit d'avoir $0.01\sqrt{n}/\sigma \geq 1.96$ ou encore

$$n \geq \frac{(1.96\sigma)^2}{0.01^2}.$$

Remarquons que la taille de l'échantillon n est linéaire en σ^2 . Il est donc judicieux de chercher à **minimiser la variance σ^2** .

2.2 Comparaison à la méthode des rectangles

On a vu que la méthode de Monte Carlo a une vitesse de convergence de l'ordre de \sqrt{n} , c'est plus lent que des méthodes déterministes, comme la méthode des rectangles qui a une vitesse de convergence de l'ordre de n sous certaines conditions.

Cependant, cette vitesse ne dépend pas de la régularité de l'intégrande et peu de la dimension d , contrairement aux méthodes déterministes : La méthode des rectangles nous dit qu'on peut approcher l'intégrale $\int_{[0,1]^d} g(x) \, dx$ par $\frac{1}{n^d} \sum_{1 \leq k_1, \dots, k_d \leq n} g\left(\frac{k_1}{n}, \dots, \frac{k_d}{n}\right)$. Sous l'hypothèse que g soit dérivable de dérivée bornée par M , on obtient comme borne de l'erreur $|\epsilon_n| \leq M \frac{\sqrt{d}}{n}$.

Le **coût de calcul** avec la méthode des rectangles est en $O(n^d)$ alors que celui de la méthode de Monte Carlo est de $O(nd)$.

Pour un temps de calcul N , la précision de la méthode des rectangles est donc en $N^{-1/d}$, tandis que celle de la méthode de Monte-Carlo est en $N^{-1/2}$. Ainsi, la méthode de Monte Carlo est plus précise à partir de la dimension 3, dans le cas de fonctions suffisamment régulières. Les exercices de ce TP ont été choisis par souci de simplicité, mais gardez ce fait en tête.

3 Comment minimiser la variance ?

Comme on l'a vu plus haut, plus la variance σ^2 est faible, meilleure est l'estimateur obtenu par la méthode de Monte-Carlo. Diverses méthodes ont été proposées pour réduire cette variance. Nous allons en mettre en pratique quelques unes dans les exercices sur l'exemple suivant, qui sera notre fil conducteur, et dont les motivations sont données à la section 4. On se donne une variable X de loi $\mathcal{N}(0, 1)$ et l'on souhaite donner une approximation de l'intégrale/espérance $C := \mathbb{E}[(e^X - 1)_+]$, où $x_+ = \max(0, x)$.

3.1 Échantillonnage préférentiel

On souhaite calculer une intégrale du type

$$I = \int g(x) f(x) \, dx,$$

où f est une densité de probabilité. La méthode de Monte-Carlo "naïve" consiste à approcher I par

$$I_n := \frac{1}{n} \sum_{k=1}^n f(X_k)$$

où les variables X_i sont i.i.d. de densité f . Soit h une autre densité de probabilité supposée strictement positive. On peut alors écrire

$$I = \int \frac{f(x)g(x)}{h(x)} h(x) \, dx,$$

ce qui suggère l'approximation

$$J_n = \frac{1}{n} \sum_{k=1}^n \frac{f(Y_k)g(Y_k)}{h(Y_k)},$$

où cette fois, les variables Y_i sont de densité h . Il y a gain de variance si

$$\text{var} \left(\frac{f(Y_1)g(Y_1)}{h(Y_1)} \right) \leq \text{var} (g(X_1)),$$

ou encore

$$\int \left(\frac{f(x)g(x)}{h(x)} \right)^2 h(x) dx \leq \int g(x)^2 f g(x) dx.$$

On doit donc choisir h de sorte que, d'une part, les variables Y_i soient faciles à simuler, et d'autre part, que l'inégalité ci-dessus soit satisfaite. Une méthode possible pour cela consiste à choisir h "proche" de la fonction fg de sorte que la variance (sous h) soit faible, puis de normaliser h pour en faire une densité.

Exercice 4. On souhaite donner une approximation de

$$C := \mathbb{E}[(e^X - 1)_+] = \frac{1}{\sqrt{2\pi}} \int (e^x - 1)_+ e^{-x^2/2} dx.$$

On note que, pour x proche de zéro, on a $e^x - 1 \approx x$ ce qui motive le calcul suivant

$$C = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{e^x - 1}{x} x e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{e^{\sqrt{2y}} - 1}{\sqrt{2y}} e^{-y} dy = \frac{1}{\sqrt{2\pi}} \mathbb{E} \left[\frac{e^{\sqrt{2Y}} - 1}{\sqrt{2Y}} \right]$$

où Y suit une loi exponentielle $\mathcal{E}(1)$. Estimer C par la méthode de Monte-Carlo naïve, puis par la méthode de Monte-Carlo basée sur la représentation en terme de variable exponentielle. Comparer les variances empiriques.

3.2 Variable de contrôle

On souhaite toujours approcher une intégrale du type $I = \mathbb{E}[g(X)]$. Supposons que l'on sache calculer explicitement une intégrale du même type, disons $\mathbb{E}[h(X)]$ pour une certaine fonction h . On peut alors écrire $I = \mathbb{E}[g(X) - h(X)] + \mathbb{E}[h(X)]$ et on aura un gain de variance dès que $\text{var}(g(X) - h(X)) \leq \text{var}(g(X))$.

Exercice 5. On revient sur le calcul approché de $C = \mathbb{E}[(e^X - 1)_+]$ où $X \sim \mathcal{N}(0, 1)$. On introduit la quantité $P := \mathbb{E}[(1 - e^X)_+]$ et on remarque que $C - P = \mathbb{E}[e^X - 1] = e^{1/2} - 1$, de sorte que $C = P + e^{1/2} - 1$. Estimer C via P par la méthode de Monte-Carlo. Comparer les variances par rapport aux approximations précédentes.

3.3 Symétrisation

On souhaite encore et toujours donner une valeur approchée d'une intégrale/espérance du type $I = \mathbb{E}[g(X)]$. Supposons que pour une certaine transformation T , les variables X et $T(X)$ aient même loi. On peut alors écrire

$$I = \frac{\mathbb{E}[g(X)] + \mathbb{E}[g(T(X))]}{2},$$

et l'approcher, via la méthode de Monte-Carlo par

$$I_n = \sum_{i=1}^n \frac{g(X_i) + g(T(X_i))}{2}.$$

Le calcul de la variance donne alors

$$\begin{aligned}
\text{Var}\left(\frac{g(X_1) + g(T(X_1))}{2}\right) &= \mathbb{E}\left[\left(\frac{g(X_1) + g(T(X_1))}{2}\right)^2\right] - I^2 \\
&= \frac{1}{2} (\mathbb{E}[g(X_1)^2] + \mathbb{E}[g(X_1)g(T(X_1))]) - I^2 \\
&\stackrel{CS}{\leq} \frac{1}{2} (\mathbb{E}[g(X_1)^2] + \mathbb{E}[g(X_1)^2]) - I^2 \\
&= \text{var}(g(X_1)).
\end{aligned}$$

Il y donc un gain de variance.

Exercice 6. Si $X \sim \mathcal{N}(0, 1)$, alors $-X \sim \mathcal{N}(0, 1)$ et l'on peut donc écrire

$$C = \mathbb{E}[(e^X - 1)_+] = \frac{\mathbb{E}[(e^X - 1)_+] + \mathbb{E}[(e^{-X} - 1)_+]}{2},$$

Estimer C par la méthode de Monte-Carlo basée sur cette dernière écriture et comparer avec les approximations précédentes.

4 Option d'achat et de vente

Dans cette dernière section, nous revenons sur l'exemple qui nous a servi de fil conducteur dans la section précédente, à savoir le calcul de l'intégrale $C = \mathbb{E}[(e^X - 1)_+]$. Nous tâchons d'expliquer pourquoi le calcul d'une telle quantité est naturel et important dans la pratique.

Supposons que vous êtes un fabricant de biscuits à l'épeautre. Vous achetez vos matières premières, en particulier la farine d'épeautre tous les mois. Les prix de ces matières premières varient quotidiennement, du fait de l'offre et de la demande et des spéculateurs. Dans six mois, vous savez que vous aurez besoin de dix tonnes de farine. Le cours actuel est de 2500 euros la tonne. Dans six mois, selon la demande, la météo etc., ce cours pourra être encore de 2500 euros la tonne, il pourra avoir baissé à 2000 euros ou au contraire il pourra avoir flambé jusqu'à 3000 euros. Ces variations auront naturellement un impact fort sur votre trésorerie au moment de l'achat.

Pour se prémunir d'une éventuelle flambée des prix, vous pouvez émettre une option d'achat, aussi appelé un "call", auprès d'un vendeur de céréales. Cela consiste à payer un montant C (convenu à l'avance entre vendeur et acheteur), pour qu'à une date fixée T (ici dans six mois), vous puissiez exercer votre droit d'acheter ou non la marchandise au vendeur avec qui vous souscrivez le contrat, à un prix K lui aussi fixé à l'avance, et ce quelque soit le cours de la tonne de farine à l'instant T . Comment fixer la valeur d'une telle option d'achat ?

Observons votre gain/perte selon le cours de la farine au temps final T . Soit (X_t) le cours (aléatoire) de la farine d'épeautre à l'instant $0 \leq t \leq T$. Si $X_T \geq K$, vous exercez votre droit i.e. vous achetez l'action au vendeur au prix K et vous gagnez ainsi $X_T - K - C = (X_T - K)_+ - C$. En revanche, si $X_T < K$, vous n'exercez pas votre droit et n'achetez pas l'action à ce vendeur et votre gain/perte est $-C = (X_T - K)_+ - C$. En moyenne (selon les aléas), votre gain au cours de la transaction avec ce vendeur sera donc de

$$\mathbb{E}[(X_T - K)_+] - C.$$

Pour que le jeu soit équitable entre acheteur et vendeur, il faut donc que le prix C de l'option d'achat soit tel que

$$C = \mathbb{E}[(X_T - K)_+].$$

On peut naturellement jouer au même jeu avec le point de vue du vendeur qui veut se prémunir d'une baisse importante du cours d'une action, auquel cas l'option de vente aussi appelée "put" est donnée par $P = \mathbb{E}[(K - X_T)_+]$.

Dans la réalité, on ne connaît bien sûr pas la loi de la variable X_T , i.e. dans notre exemple, le cours de l'action de farine d'épeautre dans six mois. L'un des objets principaux des mathématiques financières consiste précisément à modéliser l'évolution $(X_t)_{0 \leq t \leq T}$ du cours d'une action au cours du temps. Pour des modèles simplistes, on peut calculer explicitement le prix C de l'option d'achat. En revanche, dans des modèles un tant soit peu réalistes d'évolution des cours, la loi de X_T reste inconnue, et l'on recourt alors à la méthode de Monte-Carlo pour estimer le coût de cette option.

Références

- [CBCC16] Alexandre Casamayou-Boucau, Pascal Chauvin, and Guillaume Connan. *Programming en Python pour les mathématiques - 2e éd.* Dunod, Paris, 2e édition edition, January 2016.
- [Rub] Sylvain Rubenthaler. Méthodes de monte carlo. <https://math.unice.fr/~rubentha/enseignement/poly-cours-monte-carlo-m1-im.pdf>.
- [Vig18] Vincent Vigon. *python proba stat*. Independently published, October 2018.