Cours : Méthodes de Monte Carlo, l'essentiel

Ayoub OULAD ALI

Janvier 2026

Table des matières

1 Méthodes de Monte-Carlo, fondements théoriques **3**
1.1 Motivation : Le problème des événements rares 4
1.2 Formalisation et Propriétés 5

2 Algorithmes **6**
2.1 Monte Carlo Standard (MC) 6
2.2 Monte Carlo par Importance Sampling (MCIS) 6

1 Méthodes de Monte-Carlo, fondements théoriques

Les méthodes de Monte-Carlo, sont des méthodes de simulations basées sur le tirage aléatoire de variables *i.i.d* i.e indépendantes et identiquement distribuées.

Tout le long on considerera $(X_i)_i \overset{\text{i.i.d.}}{\sim} f$ tel que $\forall i, X_i \in \mathrm{L}^2(\mathbb{R})$?

Théorème 1.1 LFGN

Sous les hypothèses précédentes :

$$\lim_{n\to\infty} \bar{X}_n = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}(X_1)$$

Théorème 1.2 TCL

Sous les hypothèses précédentes :

$$\sqrt{n}(\bar{X}_n - \mathbb{E}(X_1)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2) \text{ avec } \sigma^2 = Var(X_1)$$

Définition 1.1 : Estimateur de Monte Carlo

Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable telle que $\mathbb{E}[|g(X_1)|] < \infty$. On souhaite calculer la quantité $I = \mathbb{E}[g(X)]$. L'estimateur de Monte Carlo de I est défini par :

$$\hat{I}_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

Consistance : D'après la **Loi Forte des Grands Nombres** (Théorème 1), l'estimateur $\hat{I}_n$ est un estimateur fortement consistant de I. C'est-à-dire que l'erreur converge vers 0 presque sûrement :

$$\lim_{n\to\infty} \hat{I}_n = \mathbb{E}[g(X)] \quad \text{p.s.}$$

Exemple 1.1 : Calcul de probabilité (Queue de Gaussienne)

Considérons le cas où les tirages suivent une loi normale centrée réduite :

$$X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$$

On cherche à estimer la probabilité p que la variable dépasse un seuil $a \in \mathbb{R}$:

$$p = \mathbb{P}(X > a) = \mathbb{E}[\mathbb{1}_{\{X>a\}}]$$

Ici, la fonction d'intérêt est la fonction indicatrice $g(x) = \mathbb{1}_{\{x>a\}}$. L'estimateur s'écrit :

$$\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i>a\}}$$

Calcul de la variance : Posons $Y_i = \mathbb{1}_{\{X_i>a\}}$. La variable Y_i ne peut prendre que les valeurs 0 ou 1. Elle suit donc une loi de Bernoulli de paramètre p.

$$\mathrm{Var}(Y_i) = p(1-p)$$

Par indépendance des X_i (et donc des Y_i), la variance de l'estimateur est :

$$\begin{aligned}\mathrm{Var}(\hat{p}_n) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(Y_i) \\ &= \frac{1}{n^2}\cdot n\cdot p(1-p) \\ &= \frac{p(1-p)}{n}\end{aligned}$$

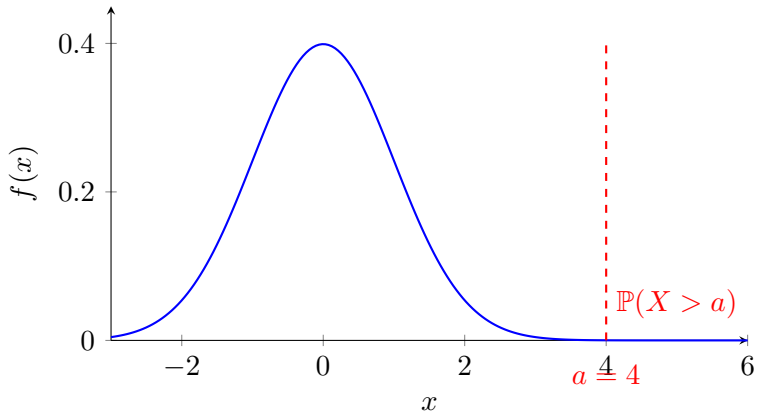
Remarque
On observe que l'erreur quadratique moyenne (liée à la variance) décroît en $1/n$, donc l'écart-type (l'erreur de l'estimation) décroît en $1/\sqrt{n}$.

1.1 Motivation : Le problème des événements rares

Reprenons l'exemple précédent où nous cherchons à estimer $p = \mathbb{P}(X > a)$ avec $X \sim \mathcal{N}(0,1)$.

Remarque
Si a est grand (par exemple $a = 5$), la probabilité p est extrêmement faible (de l'ordre de 10^{-7}). Dans notre estimateur Monte Carlo classique $\hat{p}_n = \frac{1}{n}\sum \mathbb{1}_{\{X_i > a\}}$, la très grande majorité des X_i seront inférieurs à a. L'indicatrice vaudra presque toujours 0.
Pour espérer observer ne serait-ce qu'un seul événement $\{X_i > a\}$, il faudrait en moyenne $n \approx 1/p = 10^7$ tirages ! La variance relative $\frac{\sqrt{\mathrm{Var}(\hat{p}_n)}}{p} = \frac{\sqrt{p(1-p)/n}}{p} \approx \frac{1}{\sqrt{np}}$ explose lorsque $p \to 0$.

Visualisons le problème. On cherche à intégrer la densité f (la courbe bleue) sur la zone rouge, qui est minuscule.



L'idée intuitive : Au lieu de tirer "à l'aveugle" selon la loi $\mathcal{N}(0,1)$ (qui met peu de masse en a), pourquoi ne pas simuler selon une autre loi qui favorise la région d'intérêt ?

Par exemple, si l'on simule des $Y_i \sim \mathcal{N}(a,1)$ (une gaussienne centrée sur a), beaucoup de points tomberont au-delà de a. Cependant, ces points ne sont pas ceux de la loi originale. Il faut donc corriger ce biais en les repondérant. C'est le principe de l'Importance Sampling.

1.2 Formalisation et Propriétés

Soit X une variable aléatoire de densité f sur $\mathbb{R}$. On cherche à estimer $I = \mathbb{E}_f[h(X)] = \int h(x)f(x)dx$.

Introduisons une nouvelle densité de probabilité g (appelée *densité instrumentale* ou *de proposition*), telle que $g(x) > 0$ partout où $h(x)f(x) \neq 0$.

On utilise l'astuce suivante :

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

Si l'on considère une variable aléatoire Y de densité g, cette intégrale n'est autre que l'espérance suivante :

$$I = \mathbb{E}_g\left[h(Y)\frac{f(Y)}{g(Y)}\right]$$

On pose le poids d'importance $w(y) = \frac{f(y)}{g(y)}$. On a $I = \mathbb{E}_g[h(Y)w(Y)]$.

Définition 1.2 : Estimateur d'Importance Sampling

Soit $Y_1, \ldots, Y_n$ une suite de variables aléatoires i.i.d. de densité g. L'estimateur par échantillonnage préférentiel de I est :

$$\hat{I}_n^{IS} = \frac{1}{n}\sum_{i=1}^{n} h(Y_i)\frac{f(Y_i)}{g(Y_i)} = \frac{1}{n}\sum_{i=1}^{n} h(Y_i)w(Y_i)$$

Cet estimateur conserve les bonnes propriétés de l'estimateur MC classique, mais sous la loi g.

Théorème 1.3 : Consistance de l'IS

Si $\mathbb{E}_g[|h(Y)w(Y)|] < \infty$, alors par la Loi Forte des Grands Nombres appliquée à la variable $Z = h(Y)w(Y)$:

$$\lim_{n\to\infty} \hat{I}_n^{IS} = \mathbb{E}_g[h(Y)w(Y)] = I \quad \text{p.s.}$$

L'estimateur est donc fortement consistant (sans biais).

Théorème 1.4 : TCL pour l'IS

Si la variance sous g est finie, c'est-à-dire si :

$$\sigma_g^2 := \mathrm{Var}_g(h(Y)w(Y)) = \mathbb{E}_g\left[\left(h(Y)\frac{f(Y)}{g(Y)}\right)^2\right] - I^2 < \infty$$

Alors, par le Théorème Central Limite :

$$\sqrt{n}(\hat{I}_n^{IS} - I) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_g^2)$$

Remarque

Le but du jeu est de choisir une densité g qui minimise la variance σ_g^2. Un mauvais choix de g (par exemple si le poids f/g peut devenir très grand) peut entraîner une variance infinie et rendre l'estimateur inutilisable.

2 Algorithmes

Voici la mise en œuvre pratique des deux méthodes étudiées pour estimer $I = \mathbb{E}_f[h(X)]$.

2.1 Monte Carlo Standard (MC)

Cet algorithme suppose que l'on sait simuler directement selon la loi f.

Algorithme 1 : Estimateur Monte Carlo Naïf

Données : n : Nombre de simulations
h : Fonction d'intérêt
$Simul_f()$: Générateur de nombres aléatoires selon la loi f
Résultat : Estimation $\hat{I}_n$

$S \leftarrow 0$; `// Initialisation de la somme`
pour $i \leftarrow 1$ **à** n **faire**
  $X \leftarrow Simul_f()$; `// Tirage selon la loi cible`
  $S \leftarrow S + h(X)$; `// Accumulation`
fin
retourner S/n ; `// Moyenne empirique`

2.2 Monte Carlo par Importance Sampling (MCIS)

Cet algorithme est utilisé quand on ne peut pas simuler selon f, ou pour réduire la variance (événements rares). On simule selon une loi instrumentale g.

Algorithme 2 : Importance Sampling

Données : n : Nombre de simulations
h : Fonction d'intérêt
f : Densité de la loi cible (à évaluer)
g : Densité de la loi instrumentale (à évaluer)
$Simul_g()$: Générateur de nombres aléatoires selon la loi g
Résultat : Estimation $\hat{I}_n^{IS}$

$S \leftarrow 0$;
pour $i \leftarrow 1$ **à** n **faire**
  $Y \leftarrow Simul_g()$; `// Tirage selon la loi instrumentale`
  $W \leftarrow f(Y)/g(Y)$; `// Calcul du poids d'importance`
  $S \leftarrow S + h(Y) \cdot W$; `// Accumulation pondérée`
fin
retourner S/n

lien vers exemple de simulation