Département Statistique
$2^{\text{éme}}$ année

2022-2023
Méthodes de simulation

TP5

Objectif

On se propose d'implémenter des techniques de simulation telles que l'échantillonnage préférentiel et les variables antithétiques qui visent à réduire la variance de l'estimateur de Monte-Carlo.

Échantillonnage préférentiel (Importance Sampling)

Il s'agit d'une méthode d'estimation de l'espérance d'une fonction d'une variable aléatoire X, lorsque l'évaluation directe de cette espérance est difficile ou impossible.

$$\mu = E(X) = E\left(h(X)\frac{\pi(X)}{g(X)}\right)$$

La méthode consiste à tirer des échantillons d'une distribution instrumentale (aussi appelée loi d'importance ou de proposition) notée $g(X)$ qui est plus facile à échantillonner que la distribution d'intérêt (ou la distribution cible) notée $\pi(X)$. Les échantillons sont pondérés par le rapport entre la densité de probabilité de la distribution cible et celle de la loi instrumentale, appelé poids d'importance.

$$W(i) = \frac{\pi(X(i))}{g(X(i)}$$

La moyenne pondérée des échantillons fournit alors une estimation de l'espérance de la fonction sous la distribution cible.

$$\hat{\mu} = \frac{W(1)h(X(1)) + ... + W(m)h(X(m))}{m}$$

Cette méthode peut être utilisée pour évaluer des intégrales, des probabilités, ou pour effectuer une optimisation stochastique et réduire la variance.

On cherche à estimer $E(e^{-X+\cos(X)})$.

1. Estimer μ avec Monte-Carlo naive.

2. En considérant $X_1, ..., X_n$ des réalisations de variables aléatoires i.i.d. suivant la loi de $EXP(1)$, donner une proposition de distribution instrumentale.

3. Calculer les poids d'importance.

4. Donner une estimation de μ avec la méthode d'échantillonnage préférentiel. Comparer le résultat avec celui de Monte Carlo naive.

Variables antithétiques

Cette méthode consiste à utiliser des variables aléatoires corrélées négativement afin de réduire la variance de l'estimateur. Plus précisément, au lieu de générer une seule suite de variables aléatoires, on en génère deux qui sont corrélées négativement, c'est-à-dire que lorsque l'une prend une valeur élevée, l'autre prend une valeur faible et vice versa.

L'estimateur d'intérêt est ensuite calculé en utilisant la moyenne des observations des deux suites de variables aléatoires. Ainsi, l'estimateur obtenu est plus précis que celui obtenu avec la méthode Monte-Carlo classique ; Si $X_1, X_2, ..., X_n$ sont des variables aléatoires indépendantes et identiquement distribuées avec une moyenne μ et une variance σ^2, alors l'estimateur de Monte-Carlo classique est donné par :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Avec la méthode des variables antithétiques, on génère également une suite de variables aléatoires $Y_1, Y_2, ..., Y_n$ telles que $Y_i = g(X_i)$, où g est une fonction qui établit une corrélation négative entre X et Y. L'estimateur d'intérêt est alors donné par :

$$\hat{\mu} = \frac{1}{2n} \sum_{i=1}^{n} (X_i + Y_i)$$

Avec cette méthode, la variance de l'estimateur est réduite de manière significative par rapport à l'estimateur classique de Monte-Carlo, car la covariance négative entre X et Y permet de compenser les erreurs de chaque suite de variables aléatoires.

On souhaite calculer :

$$I = \int_0^2 e^{-x^2} dx$$

1. Utiliser R pour calculer I directement.

2. (a) Estimer I par la méthode de Monte-Carlo classique de deux façons différentes.

   (b) Donner les intervalles de confiance au niveau 0.95 correspondant aux estimateurs précédents.

3. (a) Pour chacun des estimateurs précédents, proposer une méthode alternative utilisant une variable antithétique.

   (b) Donner les intervalles de confiance au niveau 0.95 correspondant aux estimateurs précédents.