Sommaire.

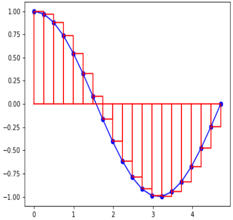
Méthodes de quadrature

En analyse numérique, il existe une vaste famille d'algorithmes dont le but principal est d'estimer la valeur numérique de l'intégrale définie sur un domaine particulier pour une fonction donnée (par exemple l'intégrale d'une fonction d'une variable sur un intervalle).

- Formule des rectangles

Utiliser la méthode des rectangles consiste à approximer la fonction f entre a et b par une somme de surfaces de rectangles de longueur $h = \frac{b-a}{n}$

$$\int_a^b f(x)dx \approx \frac{(b-a)}{n} \sum_{k=0}^{n-1} f(a + k\,h)$$



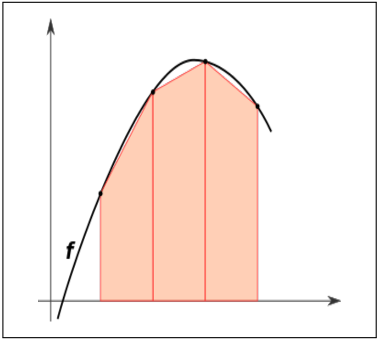
L'erreur faite en prenant cette approximation est majorée par : $\frac{(b-a)^2}{n} \,.\, sup_{x\in[a,b]}\, |f^{'}(x)|$, cela signifie que la méthode des rectangles est exacte pour un polynôme de degré 0.

Méthodes de quadrature

- Formule des trapèze :

Utiliser la méthode des trapèzes consiste à approximer la fonction f entre a et b par une fonction continue affine par morceaux égale à f sur les points de subdivision qui sont les bornes des n intervalles de longueur $h = \frac{b-a}{n}$

$$\int_a^b f(x)dx \approx h \left[\sum_{k=1}^{n-1} f(a+kh) + \frac{(f(a) + f(b)}{2} \right]$$



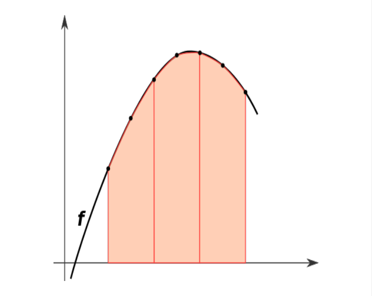
L'erreur faite en prenant cette approximation est majorée par : $\frac{(b-a)^3}{12\,n^2} \,.\, sup_{x\in[a,b]} \quad |f''(x)|$, cela signifie que la méthode des trapèzes est exacte pour un polynôme de degré 1.

Méthodes de quadrature

- Formule de Simpson :

Utiliser la méthode de Simpson consisteà approximer la fonction f prise entre a et b par un polynôme de degré 2 dont nous calculons l'intégrale toujours sur la base des n intervalles de longueur $h = \frac{b-a}{n}$

$$\int_a^b f(x)dx \approx \frac{h}{3}\left[f(a) + f(b) + 2\sum_{k=1}^{n/2-1} f(a+2kh) + 4\sum_{k=1}^{n/2-1} f(a+(2k-1)h)\right]$$



L'erreur faite en prenant cette approximation est majorée par : $\frac{(b-a)^5}{2880\, n^4} \cdot sup_{x\in[a,b]} \ |f^{(4)}(x)|$, cela signifie que la méthode de Simpson est exacte pour un polynôme de degré 3.

Quadrature de Gauss

Technique d'intégration par la méthode de Monte-Carlo

Evaluation d'une intégrale (base de simulations)

le plus simple : Supposons que l'on veuille calculer :

$$I = \int_{[0,1]} f(x)\, dx$$

Plus compliqué : Intégrale de dimension d

$$\int_{[0,1]^d} f(x_1, x_2, \ldots, x_d)\, dx_1,\ dx_2, \ldots,\ dx_d$$

de la fonction f dans l'hypercube unité $[0, 1]^d$.

L'intégrale peut être interprétée comme l'espérance $E[f(X)]$ de la variable aléatoire $f(X)$, où X est une variable aléatoire à valeur dans $\mathcal{R}^d$ avec une distribution uniforme sur $[0, 1]^d$, ce qui signifie que les composantes $X_1, \ldots X_d$ sont indépendantes et identiquement uniformément distribuées sur $[0, 1]$, c'est-à-dire que $X_1, \ldots X_d$ sont des nombres aléatoires.

Espérance Mathématique

- **Espérance Mathématique d'une fonction $f(X)$**

$$E[f(X)] \;=\; \int_{[0,1]} f(x)\, g_X(x)\;\, dx$$

$g_X(x)$ est la fonction densité de probabilité (PDF).

On considère une variable aléatoire continue X uniformément distribuée sur $[0,\ 1]$: $g_X(x) \;=\; \mathbb{I}_{[0,1]}(x)$

$$E[f(X)] \;=\; \int_{[0,1]} f(x)\; dx \quad \Rightarrow \quad I \;=\; E[f(X)].$$

- **Estimateur Monte Carlo**

On considère un n-échantillon $(x_i)_{i\in n}$ de n variables aléatoires indépendantes $(X_i)_{i\in n}$; On calcule la moyenne de $f(x)$ du n-échantillon

$$\widetilde{f_n} \;=\; \frac{1}{n}\sum_{i=1}^{n} f(x_i).$$

$$\widehat{I_n} \;=\widetilde{f_n} \;=\; \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

est appelé estimateur Monte Carlo de $E[f(X)]$.

Loi des grands nombres

- Estimateur MC : estimateur sans biais de l'intégrale

$$E[\widetilde{f_n}] = E\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big] = \frac{1}{n}\sum_{i=1}^{n} E\big[f(x_i)\big] = E\big[f(X)\big] = I \qquad \left(E[\widehat{I_n}] = E\big[f(X)\big] = I\right)$$

- Convergence (loi des grands nombres)

$$\widehat{I_n} = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \xrightarrow[p.s.]{n \to +\infty} E[f(X)] = I.$$

- Erreur quadratique moyenne (MSE) :

$$E\left[\left(\widetilde{f_n} - E[f(X)]\right)^2\right] \left(E\left[(\widehat{I_n} - I)^2\right]\right) = Var(\widetilde{f_n}) \left(Var(\widehat{I_n})\right) = \frac{1}{n}\, Var[f(X)]$$

Exemple

- La vitesse de convergence de la méthode de Monte Carlo classique est en $\mathcal{O}(n^{-1/2})$.
- La méthode Monte-Carlo, avec une vitesse en $1/\sqrt{n}$ est cependant insensible à la dimension et peut donc s'avérer plus avantageuse dès que l'on travaille en dimension grande ou sur une fonction irrégulière
- On peut remarquer que cette vitesse de convergence ne permet pas d'estimer $E[f(X)]$ avec une relativement grande précision, chaque chiffre significatif supplémentaire, nécessitant un coût de simulation 100 fois supérieur.

On veut calculer l'estimateur de l'intégrale suivante par la méthode Monte Carlo :

$$I = \int_0^1 e^{-x}\,dx$$

et la comparer à sa valeur exacte

- Estimateur MC

R > x < − runif(10000)
R > mean(exp(-x))

$\widehat{I_n} = 0.6312546$

- calcul analytique : $I = 1 - e^{-1} \Rightarrow$ I = 0.6321206

Extension à un intervalle quelconque

On veut calculer l'estimateur Monte Carlo de

$$I = \int_a^b f(x)\,dx$$

- I se ré-écrit comme
$$I = (b-a)\int_a^b f(x)\,\frac{1}{b-a}\,dx \;=\; (b-a)\,E[f(X)]$$
où $X \sim \mathcal{U}([a,b])$.
- L'estimateur Monte Carlo (CMC) est alors :
$$\widehat{I_n} = \frac{b-a}{n}\sum_{i=1}^{n} f(X_i)$$
où $X_1, \ldots, X_n$ sont i.i.d. sur $\mathcal{U}([a,b])$.

Variance Monte Carlo

- Calcul de la variance :

$$Var(f(X)) = \sigma^2 = E\left[(f(X) - E[f(X)])^2\right] = \int_{x\in\mathcal{X}} (f(x) - E[f(X)])^2 \; g_X(x)\,dx$$

$$Var(f(X)) = \sigma^2 = E\left[f(X)^2\right] - E[f(X)]^2 = \int_{x\in\mathcal{X}} f(x)^2\, g_X(x)\,dx - E[f(X)]^2$$

- Estimateur non biaisé de la variance

$$\widehat{Var(f(X))} = \widehat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}[f(x_i) - \widetilde{f}_n]^2$$

- L'estimateur non biaisé de $Var\left(\widetilde{f}_n\right)$ est :

$$Var\left(\widetilde{f}_n\right)\left(Var(\widehat{I}_n)\right) = \frac{\widehat{Var(f(X))}}{n} = \frac{1}{n(n-1)}\sum_{i=1}^{n}[f(x_i) - \widetilde{f}_n]^2$$

ou

$$Var\left(\widetilde{f}_n\right)\left(Var(\widehat{I}_n)\right) = \frac{1}{n}\left(\frac{1}{n-1}\sum_{i=1}^{n}[f(x_i)]^2 - \frac{n}{n-1}(\widetilde{f}_n)^2\right)$$

- Théorème Central Limite

Lorsque la variance de $f(X)$ est finie, le théorème Central Limite permet d'établir que $\widetilde{f}_n - E[f(X)]$ suit asymptotiquement une loi normale.

$$P\left(\left|\widetilde{f}_n - E[f(X)]\right| \leq \epsilon \frac{\left(Var(f(X))\right)^{1/2}}{\sqrt{n}}\right) \underset{n\to+\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_{-\epsilon}^{\epsilon} e^{-t^2/2}\, dt$$

- On en déduit l'intervalle de confiance bilatéral symétrique au niveau de confiance asymptotique $1 - \alpha$

$$I.C. = \left[\widetilde{f}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \,;\, \widetilde{f}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] \Rightarrow P\left(E[f(X)] \in \left[\widetilde{f}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \,;\, \widetilde{f}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]\right) = 1 - \alpha$$

- Le niveau de confiance usuel est $1 - \alpha = 0.95; \quad z_{1-\frac{\alpha}{2}} = 1.96$

$$P\left(E[f(X)] \in \left[\widetilde{f}_n - 1.96 \frac{\sigma}{\sqrt{n}} \,;\, \widetilde{f}_n + 1.96 \frac{\sigma}{\sqrt{n}}\right]\right) = 0.95 \quad P\left(I \in \left[\widehat{I}_n - 1.96 \frac{\sigma}{\sqrt{n}} \,;\, \widehat{I}_n + 1.96 \frac{\sigma}{\sqrt{n}}\right]\right) = 0.95$$

- Lorsque la variance est finie, $Var(f(X))$ peut être remplacé par son estimateur empirique $\widehat{\sigma}_n^2$. On obtient une approximation de l'erreur quadratique moyenne mais également de l'intervalle de confiance.

$$P\left(I \in \left[\widehat{I}_n - 1.96 \frac{\widehat{\sigma}_n}{\sqrt{n}} \,;\, \widehat{I}_n + 1.96 \frac{\widehat{\sigma}_n}{\sqrt{n}}\right]\right) = 0.95$$

Hit or Miss Integration

Soit un évènement A de probabilité p. On considère une séquence de variables alátoires de Bernoulli indépendantes et identiquement distribuées : $X_1, \; X_2 \; \ldots \; X_n$

$$X_k = \begin{cases} 1, & \text{if event A occurs} \\ 0, & \text{otherwise} \end{cases} \qquad k = 1, 2, \ldots$$

On pose $\{X = 1\}$ pour hit et $\{X = 0\}$ pour miss.

L'estimation de p par la méthode des moments et par le maximum de vraisemblance basée sur $X_1, \; X_2, \; \ldots \; X_n$ s'exprime $\widehat{p} \; = \; \frac{\sum_{k=1}^{n} X_k}{n}$

Cet estimateur est sans biais, $E[\widehat{p}] \; = \; p$.
Il suit, en utilisant la loi forte des grands nombres, que :

$$P\left(lim_{n \to \infty} \frac{\sum_{k=1}^{n} X_k}{n} \; = \; p \right) = 1$$

Intégration par la méthode du "Hit or Miss"

On veut calculer l'estimateur MC de

$$I = \int_a^b f(x)\,dx$$

- On assume que $0 \le g(x) \le c$, et on écrit I comme

$$I = \int_a^b \int_0^{f(x)} 1\,dxdy = c(b-a) \int_a^b \int_0^c \frac{\mathbf{1}\{y \le f(x)\}}{c(b-a)}\,dydx$$

- Alors, $I = c(b-a)E(\mathbf{1}\{Y \le f(X)\})$, où $Y \sim \mathcal{U}([0,c])$ et $X \sim \mathcal{U}([a,b])$
- En générant indépendamment $X_1, \ldots, X_n$ distribuée uniforément $\mathcal{U}([a,b])$ et $Y_1, \ldots, Y_n$ sur $\mathcal{U}([0,c])$ nous obtenons le "hit or miss" estimateur de I

$$I_n = \frac{c(b-a)}{n} \sum_{i=1}^{n} \mathbf{1}\{Y_i \le f(X_i)\}$$

Application : calcul du nombre π

On considère le 1/4 de disque de rayon 1 centré en 0 inscrit dans un carré de côté 1. Soit p la probabilité de trouver un point généré au hasard dans le cercle : $p = \frac{\pi}{4} = \frac{\mathcal{A}_{cercle}}{\mathcal{A}_{total}}$

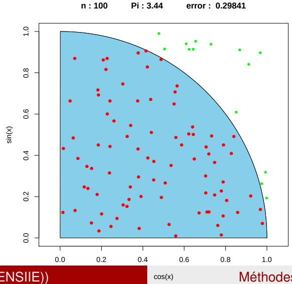
On tire aléatoirement les valeurs de U_1 et U_2 i.i.d. dans $[0,\ 1]^2$ suivant une loi uniforme.
On considère la variable X associée à l'évènement "trouver un point dans le disque" :

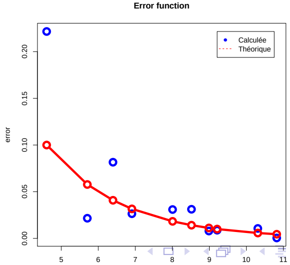
$$X = \begin{cases} 1, & \text{if } U_1^2 + U_2^2 \leq 1 \\ 0, & \text{sinon} \end{cases}$$

$$\text{alors} \qquad E[X] = P(U_1^2 + U_2^2 \leq 1) = \frac{\pi}{4}$$

Pour une séquence de n variables i.i.d., $X_1,\ X_2,\ \ldots\ X_n$:

$$lim_{n\to\infty} \frac{\sum_{k=1}^{n} X_k}{n} = \frac{\pi}{4} \quad \text{avec une probabilité de 1 (loi forte des grands nombres).}$$





Sommaire.

Réduction de variance

Objectif : réduire la variance de l'estimateur MC

$$E\left[(\widehat{I_n} - I)^2\right] = \frac{1}{n} Var[f(X)]$$

augmenter la précision de l'estimateur peut être obtenue pour un nombre donné d'estimation.

Pour cela il y a quelques techniques.

- *Noter qu'une tentative mal réalisée de réduction de la variance, au pire conduit à une variance plus grande, mais généralement rien de pire. Par conséquent, il n'y a pas trop à perdre en utilisant de telles techniques. Et il suffit d'avoir la certitude que la technique employée réduit réellement la variance.*

- *Noter également que les techniques de réduction de la variance portent souvent des noms très fantaisistes, mais que les idées qui les sous-tendent sont toujours relativement simples.*

Méthode de stratification

Consiste à découper le domaine d'intégration en un nombre fini k de parties deux-à-deux disjointes :
Le domaine d'intégration $M = [0,1]^d$ est ainsi partitionné en k régions $M_1, \ldots, M_k$.

$$I = E\Big[f(X)\Big] = \sum_{j=1}^{k} E\Big[f(X)\,\mathbb{1}_{X \in M_j}\Big] = \sum_{j=1}^{k} p_j\, E\Big[f(X) \mid X \in M_j\Big] \qquad \text{où} \quad p_j = P(X \in M_j)$$

Pour chaque région M_j, on estime $I_j := E\Big[f(X) \mid X \in M_j\Big]$ avec n_j tirages indépendants :

$$\widehat{I}^{j}_{n_j} = \frac{1}{n_j} \sum_{i=1}^{n_j} f(x_i^j)$$

L'estimateur de I par échantillonnages stratifiés avec $n = \sum_{j=1}^{k} n_j$ simulations s'obtient en sommant ces k estimations :

$$\widehat{I}^{strat}_n = \sum_{j=1}^{k} p_j\, \widehat{I}^{j}_{n_j} = \sum_{j=1}^{k} \frac{p_j}{n_j} \sum_{i=1}^{n_j} f(x_i^j)$$

Méthode de stratification

On note $\sigma_j^2 = Var\Big[f(X) \mid X \in M_j\Big]$

$$Var(\widehat{I}_n^{strat}) = \sum_{j=1}^{k} p_j^2 \, Var(\widehat{I}_{n_j}^j) = \sum_{j=1}^{k} p_j^2 \, \frac{\sigma_j^2}{n_j}$$

L'estimateur sans biais de la variance de I_j est égale à :

$$\widehat{\sigma}_j^2 = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \Big(f(x_i^j) - \widehat{I}_{n_j}^j\Big)^2$$

La méthode MC a permis d'estimer I : $\widehat{I}_n$ de variance σ^2/n avec $\sigma^2 = Var[f(X)]$

$$\begin{aligned}
\sigma^2 &= E[f(X)^2] - E[f(X)]^2] = \sum_{j=1}^{k} p_j \, E\big[f(X)^2 \mid X \in M_j\big] - \Big(\sum_{j=1}^{k} p_j \, E[f(X) \mid X \in M_j]\Big)^2 \\
&= \sum_{j=1}^{k} p_j \, Var\big(f(X) \mid X \in M_j\big) + \sum_{j=1}^{k} p_j \, E\big[f(X) \mid X \in M_j\big]^2 - \Big(\sum_{j=1}^{k} p_j \, E[f(X) \mid X \in M_j]\Big)^2 \\
&\geq \sum_{j=1}^{k} p_j \, Var\big(f(X) \mid X \in M_j\big) \\
&= \sum_{j=1}^{k} p_j \, \sigma_j^2 \qquad \geq \Big(\sum_{j=1}^{k} p_j \, \sigma_j\Big)^2 \qquad (1)
\end{aligned}$$

Méthode de stratification

- L'objectif de la méthode est de minimiser la variance selon les tailles n_i des différents échantillons, lorsque le nombre total de tirages est fixé égal à n, i.e. $\sum_j^k n_j = n$. Le pb se traite comme un pb de minimisation sous contrainte. On montre que la variance minimale est alors :

$$\sum_{j=1}^{k} p_j^2 \frac{\sigma_j^2}{n_j} = \frac{1}{n}\left(\sum_{j=1}^{k} p_j \, \sigma_j\right)^2$$

- Une façon de choisir les entiers n_j est de les prendre proportionnels à la probabilité p_i, i.e. la partie entière de np_i. Toujours en négligeant la partie entière, on a alors un estimateur de I dont la variance est égale à :

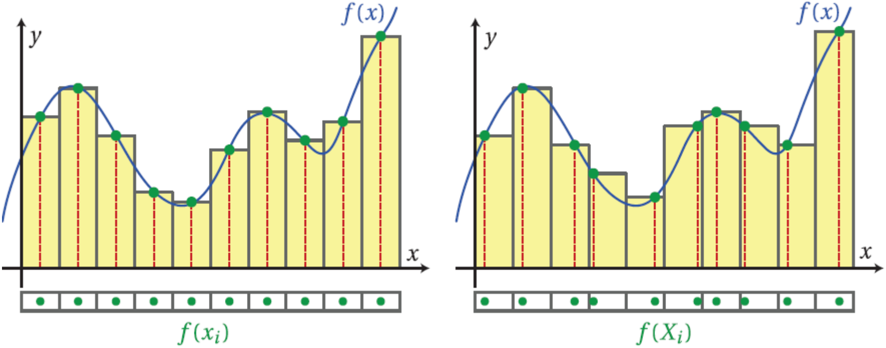
$$\frac{1}{n}\sum_{j=1}^{k} p_j^2 \frac{\sigma_j^2}{n_j} = \frac{1}{n}\sum_{j=1}^{k} p_j \, \sigma_j^2$$

$$\frac{1}{n}\sum_{j=1}^{k} p_j \frac{\sigma_j^2}{n_j} \leq \frac{1}{n}\sigma^2$$

$$Var(\widehat{I_n^{strat}}) \leq Var(\widehat{I_n})$$

on a un estimateur de I dont la variance est inférieure à la variance initiale. Ainsi, on peut toujours construire un estimateur stratifié plus efficace que l'estimateur de Monte Carlo classique, au sens où il est de variance plus faible.

Comparaison à une méthode déterministe de quadrature



Les techniques de quadrature déterministes telles que la sommation de Riemann (à gauche) échantillonnent la fonction intervalles réguliers. Conceptuellement, l'intégration Monte Carlo stratifiée (à droite) effectue une sommation très similaire, mais évalue la fonction à un emplacement aléatoire dans chacune des strates.

Fonction d'importance

Consiste à modifier la loi du vecteur aléatoire X sous laquelle est définie l'espérance $E[f(X)]$

$$E[f(X)] = \int_{\mathbb{R}} f(x)\, g(x)\, dx.$$

On considère $\widetilde{g}$ une fonction densité de probabilité $\left(\widetilde{g}(x) > 0 \text{ et } \int_{\mathbb{R}} \widetilde{g}(x)\, dx = 1\right)$, $I = E_g[f(X)]$ peut s'écrire :

$$I = E_g[f(X)] = \int_{\mathbb{R}} \left(\frac{f(x)g(x)}{\widetilde{g}(x)}\right) \widetilde{g}(x)\, dx = E_{\widetilde{g}}[Y]$$

Y est une v.a.r. $\frac{f(Z)g(Z)}{\widetilde{g}(Z)}$ qui admet $\widetilde{g}$ pour densité pour Z.

$$E_g[f(X)] = E_{\widetilde{g}}\left[f(Z)\, \frac{g(Z)}{\widetilde{g}(Z)}\right]$$

- Estimateur Monte Carlo

$$\widehat{I_n} = \frac{1}{n} \sum_{i=1} f(Z_i)\, \frac{g(Z_i)}{\widetilde{g}(Z_i)}$$

- Estimateur sans biais :

$$E_{\widetilde{g}}[\widehat{I_n}] = I$$

- Convergence

$$\widehat{I_n} = \frac{1}{n} \sum_{i=1} f(Z_i) \frac{g(Z_i)}{\widetilde{g}(Z_i)} \underset{n \to +\infty}{\longrightarrow} E_{\widetilde{g}}\left[f(Z)\, \frac{g(Z)}{\widetilde{g}(Z)}\right] = E_g[f(X)] = I$$

Fonction d'importance

- Variance

$$
\begin{aligned}
Var(\widehat{I_n}) \;=\; \frac{1}{n} Var_{\widetilde{g}}\left[f(Z)\,\frac{g(Z)}{\widetilde{g}(Z)}\right] \;&=\; \frac{1}{n}\left(E_{\widetilde{g}}\left[\left(f(Z)\,\frac{g(Z)}{\widetilde{g}(Z)}\right)^2\right] \;-\; E_{\widetilde{g}}\left[f(Z)\,\frac{g(Z)}{\widetilde{g}(Z)}\right]^2\right) \\
&=\; \frac{1}{n}\left(E_{\widetilde{g}}\left[\left(f(Z)\,\frac{g(Z)}{\widetilde{g}(Z)}\right)^2\right] \;-\; E_g\left[f(X)\right]^2\right) \\
&=\; \frac{1}{n}\left(\int_{\mathbb{R}}\left(\frac{f(z)g(z)}{\widetilde{g}(z)}\right)^2 \widetilde{g}(z)\,dz \;-\; E_g\left[f(X)\right]^2\right) \\
&=\; \frac{1}{n}\left(\int_{\mathbb{R}}\left(\frac{f(z)^2 g(z)}{\widetilde{g}(z)}\right) g(z)\,dz \;-\; E_g\left[f(X)\right]^2\right) \\
&=\; \frac{1}{n}\left(E_g\left[\frac{f(X)^2 g(X)}{\widetilde{g}(X)}\right] \;-\; E_g\left[f(X)\right]^2\right)
\end{aligned}
$$

On a donc une réduction de la variance si et seulement si

$$
E_g\left[\frac{f(X)^2 g(X)}{\widetilde{g}(X)}\right] \;\leq\; E_g\left[f(X)^2\right]
$$

Variables de contrôle

- Consiste à utiliser la solution analytique d'un problème similaire au problème posé mais plus simple et à utiliser l'erreur d'estimation de quantités connues, pour améliorer l'estimation de quantités inconnues.

- Soit $E[X]$ à estimer ($E[X]$ inconnue et $\{X_k,\ k \geq 0\}$ sont des v.a.r. i.i.d. de même distribution que X. Il existe une variable aléatoire Y de moyenne connue $\mu \equiv E[Y]$ et $\{Y_k,\ k \geq 0\}$ sont des v.a.r. i.i.d. de même distribution que Y. X et Y sont corrélées positivement.

- $W \equiv X + \beta(Y - \mu)$ est utilisée comme un estimateur "controllé" de $E[X]$ pour chaque valeur de la constante β .

  β choisi, W reste un estimateur sans biais de $E[X]$

$$E[W] = E[X] + \beta\, E[Y - \mu] = E[X].$$

- De plus $Var[W] = Var[X] + \beta^2\, Var[Y] + 2\,\beta\, Cov[X, Y]$

- $Var(W) < Var(X) \quad \Rightarrow \quad \beta^2\, Var[Y] + 2\beta Cov[X, Y] < 0$

- Le succès de la méthode dépend à la fois du choix de β et du choix de Y.

Choix optimal pour β

- Minimiser la variance $Var[W]$:

$$\beta = -Cov[X,Y]/Var[Y].$$

- En remplaçant par la valeur de β :

$$Var[W] = Var[X] - \frac{Cov[X,Y]^2}{Var[Y]} = \left(1 - \rho_{X,Y}^2\right) Var[X];$$

où $\rho_{X,Y} = \frac{Cov[X,Y]}{Var[X]\ Var[Y]}$ est le coefficient de corrélation des varaibles X and Y.

- Plus X and Y sont corrélées, plus grande est la réduction de variance.

Exemple

On cherche une solution numérique à l'intégrale :

$$I = \int_0^1 e^u \,\mathrm{d}u \quad \left(= e - 1 \right)$$

Soit $U \sim \mathcal{U}([0,\ 1])$, l'intégrale $I = E[f(U)]$, où $f(u) = e^u$.

L'espérance s'estime par la moyenne empirique. La variance associée se calcule alors explicitement :

$$\sigma^2 = E[f(U)^2] - E[f(U)]^2 = \int_0^1 (e^u)^2 \,\mathrm{d}u - (e - 1)^2$$

$$\sigma^2 == \frac{e^2 - 1}{2} - (e-1)^2 = -\frac{e^2}{2} + 2e - \frac{3}{2} \approx 0,2420.$$

En utilisant le développement de Taylor de la fonction f au voisinage de 0, on peut choisir la fonction $g(u) = 1 + u$ pour variable de contrôle de sorte que $E[g(U)] = \int_0^1 (1+u)\,\mathrm{d}u = 3/2.$
On a alors

$$Var(f(U) - g(U)) = Var(e^U - (U+1)) = -\frac{e^2}{2} + 3e - \frac{53}{12} \approx 0,0436$$

d'où une variance inférieure.

- L'utilisation de la variable de contrôle réduit sensiblement la variance dans cet exemple.

Variables antithétiques

Alors que l'intégration Monte Carlo ordinaire utilise des échantillons aléatoires constitués d'observations indépendantes, il peut être avantageux d'utiliser des échantillons comportant des paires d'observations négativement corrélées entre elles. Il s'agit de tirer parti de certaines symétries d'une distribution et de la corrélation négative entre les deux variables aléatoires.

- Exemple : Soit $U \sim \mathcal{U}([0, 1])$, alors $U' = (1 - U) \sim \mathcal{U}([0, 1])$. $F^{-1}(U)$ and $F^{-1}(1 - U)$ ont les deux la même distribution F mais sont antithétiques entre elles car F^{-1} est une fonction monotone.

- Exemple : Soit $Y \sim \mathcal{N}(\mu, \sigma)$, alors $Y' = 2\mu - Y \sim \mathcal{N}(mu, \sigma)$. $F^{-1}(Y)$ and $F^{-1}(2\mu - Y)$ ont les deux la même distribution F mais sont antithétiques entre elles. Y et Y' sont négativement corrélées.

Objectif : estimer $E[Y]$ en utilisant une séquence de paires de variables antithétiques. $(Y_1, \widetilde{Y}_1), (Y_2, \widetilde{Y}_2), \ldots, (Yn, \widetilde{Y}_n)$ où Y_i et $\widetilde{Y}_i$ ont les mêmes distributions quelque soit i et les paires $(Y_1, \widetilde{Y}_1), \ \ldots, \ (Y_n, \widetilde{Y}_n)$ sont i.i.d.

L'estimateur de $E[Y]$ est simplement la moyenne des $2n$ observations

$$\widehat{Y}_{2n} = \frac{1}{2n}\left(\sum_{i=1}^{2n} Y_i\right)$$

$$\widehat{Y}_{2n} = \frac{1}{2n}\left(\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \overline{Y}_i\right) = \frac{1}{n}\left(\sum_{i=1}^{n} \frac{Yi + \overline{Y}_i}{2}\right)$$

Par conséquent, si $Z_i := \frac{Y_i + \overline{Y}_i}{2}$ alors $\widehat{Y}_{n,a}$ est la moyenne des n observations indépendantes : $\Big(Z_1, \ \ldots, \ Z_n\Big)$

$$\widehat{Y}_{2n} = \widehat{Y}_{n,a} = E[Y]$$

Variables antithétiques (2)

On cherche un estimateur de l'espérance de Z avec une variance plus petite que celle de Y :

$$Var(\widehat{Y}_{2n}) = Var\left(\frac{\sum_{i=1}^{2n} Y_i}{2n}\right) = \frac{Var(Y)}{2n}$$

$$\begin{aligned}
Var(\widehat{Y}_{n,a}) &= Var\left(\frac{\sum_{i=1}^{n} Z_i}{n}\right) = \frac{Var(Z)}{n} \\
Var(\widehat{Y}_{n,a}) &= Var\left(\frac{Y_i + \widetilde{Y}_i}{2}\right) \\
Var(\widehat{Y}_{n,a}) &= \frac{1}{n} Var\left(\frac{Y + \widetilde{Y}}{2}\right) \\
&= \frac{1}{4n}\left(Var[Y] + Var[\widetilde{Y}] + 2Cov[Y, \widetilde{Y}]\right) \\
&= \frac{1}{2n}\left(Var[Y] + Cov[Y, \widetilde{Y}]\right)
\end{aligned}$$

$$Var(\widehat{Y}_{n,a}) < Var(\widehat{Y}_{2n}) \quad \Leftrightarrow \quad Cov[Y, \widetilde{Y}] < 0$$

.

Exemple

On cherche une solution numérique à l'intégrale :

$$I = \int_0^1 e^{u^2}\, \mathrm{d}u$$

Soit $U \sim \mathcal{U}([0,\ 1])$, l'intégrale $I = E[f(U)]$, où $f(u) = e^{u^2}$.

Estimation par la brute MC

$$n = 1000$$
$$U = rand(2*n, 1)$$
$$Y = exp(U^2)$$
$$[mean(Y) \quad std(Y)^2]$$
$$ans = \quad 1.4696 \quad 0.2315$$

Estimation par paires de variables antithétiques

$$n = 1000$$
$$U = rand(n, 1)$$
$$V = 1 - U$$
$$Z = (exp(U^2) + exp(V^2))/2$$
$$[mean(Z) \quad std(Z)^2]$$
$$ans = \quad 1.4668 \quad 0.0293$$

Exemple

- $$I = \int_0^\infty ln(1 + x^2)\, e^{-x}\, dx$$

Use $t = 1 - e^{-x}$

Simple MC :

$N = 10000;$

$g = @(t) log(1 + log(1 - t).^2);$

$T = rand(1, N); \quad X = g(T);$

disp([mean(X) std(X) 2*std(X)/sqrt(N)])

0.6901 0.7639 0.0153

Antithetic Variables

$T = rand(1, N/2);$

$X = (g(T) + g(1 - T))/2;$

disp([mean(X) std(X) 2*std(X)/sqrt(N/2)])

0.6868 0.3137 0.0089

Use $x = \frac{t}{1-t}$

Simple MC :

$N = 10000;$

$f = @(t) log(1 + x.^2). * exp(-x);$

$h = @(t) f(t./(1 - t))./(1 - t).^2;$

$T = rand(1, N); X = h(T);$

disp([mean(X) std(X) 2*std(X)/sqrt(N)])

0.6839 0.7142 0.0143

Antithetic Variables

$T = rand(1, N/2);$

$X = (h(T) + h(1 - T))/2;$

disp([mean(X) std(X) 2*std(X)/sqrt(N/2)])

0.6863 0.4574 0.0129

References

- Simulation and the Monte Carlo Method, Rubinstein and Kroese (2008)
- Monte Carlo Statistical Methods, Robert and Casella (2004)