Méthodes de Monte-Carlo et algorithmes stochastiques
Correction du TD/TP 1

M2 Finance – Université Gustave Eiffel

Exercice 1 – Une même intégrale, plusieurs estimateurs

On considère

$$I = \int_0^1 \cos\left(x^3\right) \mathrm{e}^{-x}\, dx.$$

1) Si $U \sim \mathcal{U}([0,1])$, alors la densité de U vaut 1 sur $[0,1]$. Donc, pour toute fonction intégrable f,

$$\mathbb{E}\left[f(U)\right] = \int_0^1 f(u)\, du.$$

Avec $f(u) = \cos\left(u^3\right) \mathrm{e}^{-u}$, on obtient

$$I = \mathbb{E}\left[\cos\left(U^3\right) \mathrm{e}^{-U}\right].$$

Un estimateur Monte-Carlo naturel est donc

$$\widehat{I}_n^{(1)} = \frac{1}{n} \sum_{k=1}^{n} \cos\left(U_k^3\right) \mathrm{e}^{-U_k},$$

où $U_1, \ldots, U_n$ sont i.i.d. de loi uniforme sur $[0,1]$.

2) Si $X \sim \mathcal{E}(1)$, alors X a pour densité $x \mapsto \mathrm{e}^{-x} \mathbf{1}_{x \geq 0}$. Ainsi

$$\begin{aligned}
\mathbb{E}\left[\cos\left(X^3\right) \mathbf{1}_{\{X \leq 1\}}\right] &= \int_0^{+\infty} \cos\left(x^3\right) \mathbf{1}_{\{x \leq 1\}} \mathrm{e}^{-x}\, dx \\
&= \int_0^1 \cos\left(x^3\right) \mathrm{e}^{-x}\, dx \\
&= I.
\end{aligned}$$

On peut donc prendre

$$\widehat{I}_n^{(2)} = \frac{1}{n} \sum_{k=1}^{n} \cos\left(X_k^3\right) \mathbf{1}_{\{X_k \leq 1\}},$$

où $X_1, \ldots, X_n$ sont i.i.d. exponentielles de paramètre 1.

3) La densité de Y est

$$p(y) = \frac{\mathrm{e}^{-y}}{1 - \mathrm{e}^{-1}} \mathbf{1}_{0 \leq y \leq 1}.$$

Alors

$$\begin{aligned}
\left(1 - \mathrm{e}^{-1}\right) \mathbb{E}\left[\cos\left(Y^3\right)\right] &= \left(1 - \mathrm{e}^{-1}\right) \int_0^1 \cos\left(y^3\right) \frac{\mathrm{e}^{-y}}{1 - \mathrm{e}^{-1}}\, dy \\
&= I.
\end{aligned}$$

Le troisième estimateur est donc

$$\widehat{I}_n^{(3)} = \left(1 - \mathrm{e}^{-1}\right) \frac{1}{n} \sum_{k=1}^{n} \cos\left(Y_k^3\right),$$

où $Y_1, \ldots, Y_n$ sont i.i.d. de densité p.

4) Les trois estimateurs sont sans biais. En effet, dans chaque cas, la variable aléatoire que l'on moyenne a pour espérance I.
Introduisons précisément ces trois variables :

$$\begin{aligned} Z_1 &= \cos(U^3)\mathrm{e}^{-U}, \\ Z_2 &= \cos(X^3)\mathbf{1}_{\{X\leq 1\}}, \\ Z_3 &= (1-\mathrm{e}^{-1})\cos(Y^3). \end{aligned}$$

On a donc, pour $j \in \{1,2,3\}$,

$$\mathbb{E}(Z_j) = I, \qquad \widehat{I}_n^{(j)} = \frac{1}{n}\sum_{k=1}^{n} Z_{j,k},$$

où $Z_{j,1}, \ldots, Z_{j,n}$ sont i.i.d. de même loi que Z_j.

Variance des estimateurs. Par indépendance des $Z_{j,k}$,

$$\begin{aligned} \mathrm{Var}\left(\widehat{I}_n^{(j)}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{k=1}^{n} Z_{j,k}\right) \\ &= \frac{1}{n^2}\,\mathrm{Var}\left(\sum_{k=1}^{n} Z_{j,k}\right) \\ &= \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(Z_{j,k}) \\ &= \frac{n\,\mathrm{Var}(Z_j)}{n^2} \\ &= \boxed{\frac{\mathrm{Var}(Z_j)}{n}}. \end{aligned}$$

À nombre de simulations n fixé, comparer les variances des trois estimateurs revient donc à comparer $\mathrm{Var}(Z_1)$, $\mathrm{Var}(Z_2)$ et $\mathrm{Var}(Z_3)$.

Pourquoi il n'est pas nécessaire de connaître I. Pour chacune des trois variables,

$$\mathrm{Var}(Z_j) = \mathbb{E}(Z_j^2) - \left(\mathbb{E}(Z_j)\right)^2 = \mathbb{E}(Z_j^2) - I^2.$$

Cette formule contient I^2, mais cela ne signifie pas qu'il faille connaître I pour comparer les estimateurs. En effet, le terme $-I^2$ est exactement le même dans les trois variances. Ainsi, pour deux indices i et j,

$$\begin{aligned} \mathrm{Var}(Z_i) - \mathrm{Var}(Z_j) &= \left(\mathbb{E}(Z_i^2) - I^2\right) - \left(\mathbb{E}(Z_j^2) - I^2\right) \\ &= \mathbb{E}(Z_i^2) - \mathbb{E}(Z_j^2). \end{aligned}$$

Par conséquent,

$$\boxed{\mathrm{Var}(Z_i) < \mathrm{Var}(Z_j) \quad \Longleftrightarrow \quad \mathbb{E}(Z_i^2) < \mathbb{E}(Z_j^2).}$$

On peut donc comparer les variances en calculant uniquement les seconds moments, sans calculer I.

Calcul des trois seconds moments.
Pour le premier estimateur,

$$\begin{aligned} \mathbb{E}(Z_1^2) &= \mathbb{E}\left[\cos^2(U^3)\mathrm{e}^{-2U}\right] \\ &= \int_0^1 \cos^2(x^3)\mathrm{e}^{-2x}\,dx, \end{aligned}$$

car U est uniforme sur $[0,1]$.

Pour le deuxième estimateur,

$$\begin{aligned}\mathbb{E}(Z_2^2) &= \mathbb{E}\left[\cos^2(X^3)\mathbf{1}_{\{X\le 1\}}^2\right]\\ &= \mathbb{E}\left[\cos^2(X^3)\mathbf{1}_{\{X\le 1\}}\right]\\ &= \int_0^{+\infty}\cos^2(x^3)\mathbf{1}_{\{x\le 1\}}\mathrm{e}^{-x}\,dx\\ &= \int_0^1\cos^2(x^3)\mathrm{e}^{-x}\,dx.\end{aligned}$$

On a utilisé ici le fait que $\mathbf{1}_{\{X\le 1\}}^2 = \mathbf{1}_{\{X\le 1\}}$.

Enfin, pour le troisième estimateur,

$$\begin{aligned}\mathbb{E}(Z_3^2) &= (1-\mathrm{e}^{-1})^2\mathbb{E}\left[\cos^2(Y^3)\right]\\ &= (1-\mathrm{e}^{-1})^2\int_0^1\cos^2(y^3)\frac{\mathrm{e}^{-y}}{1-\mathrm{e}^{-1}}\,dy\\ &= (1-\mathrm{e}^{-1})\int_0^1\cos^2(y^3)\mathrm{e}^{-y}\,dy.\end{aligned}$$

Si l'on pose

$$M_2 = \int_0^1\cos^2(x^3)\mathrm{e}^{-x}\,dx,$$

alors les deux derniers calculs donnent simplement

$$\mathbb{E}(Z_2^2) = M_2, \qquad \mathbb{E}(Z_3^2) = (1-\mathrm{e}^{-1})M_2.$$

Or

$$0 < 1-\mathrm{e}^{-1} < 1 \qquad \text{et} \qquad M_2 > 0.$$

Donc

$$\mathbb{E}(Z_3^2) < \mathbb{E}(Z_2^2),$$

et, puisque le terme $-I^2$ est commun aux deux variances,

$$\mathrm{Var}(Z_3) < \mathrm{Var}(Z_2).$$

Finalement,

$$\boxed{\mathrm{Var}\left(\widehat{I}_n^{(3)}\right) < \mathrm{Var}\left(\widehat{I}_n^{(2)}\right).}$$

Cette conclusion est obtenue *sans jamais calculer* I. Pour comparer également le premier estimateur aux deux autres, on peut comparer numériquement les trois seconds moments ci-dessus. Dans le TP Python de l'exercice 6, on fera plutôt cette comparaison à partir des variances empiriques observées sur les simulations.

La variance est essentielle en Monte-Carlo car, à n fixé, elle détermine directement l'erreur-type de l'estimateur :

$$\sqrt{\mathrm{Var}\left(\widehat{I}_n^{(j)}\right)} = \frac{\sqrt{\mathrm{Var}(Z_j)}}{\sqrt{n}}.$$

Deux estimateurs peuvent donc être tous les deux sans biais tout en ayant des précisions très différentes.

Idée à retenir

Le deuxième estimateur gaspille une partie des simulations : lorsque $X > 1$, la contribution vaut zéro. Le troisième estimateur force au contraire les tirages à rester dans $[0,1]$, tout en corrigeant la densité par le facteur $1-\mathrm{e}^{-1}$. On vient de montrer directement qu'il est plus précis que le deuxième. La comparaison avec l'estimateur uniforme sera complétée numériquement dans le TP.

Exercice 2 – Estimation de π, variance et taille d'échantillon

1) Le point (U_i, V_i) est uniforme dans le carré $[-1,1]^2$, de surface 4. Le disque unité a une surface π. Donc

$$p = \mathbb{P}\left(U_i^2 + V_i^2 \leq 1\right) = \frac{\pi}{4}.$$

Comme B_i est une Bernoulli de paramètre p, on a $\mathbb{E}(B_i) = p$. Un estimateur de π est alors

$$\widehat{\pi}_n = 4\frac{1}{n}\sum_{i=1}^{n} B_i.$$

2) On part de

$$\widehat{\pi}_n = 4\overline{B}_n = \frac{4}{n}\sum_{i=1}^{n} B_i.$$

Comme les B_i sont indépendantes et ont toutes la même variance,

$$\begin{aligned}\mathrm{Var}(\widehat{\pi}_n) &= \mathrm{Var}\left(\frac{4}{n}\sum_{i=1}^{n} B_i\right) \\ &= \frac{16}{n^2}\,\mathrm{Var}\left(\sum_{i=1}^{n} B_i\right) \\ &= \frac{16}{n^2}\sum_{i=1}^{n}\mathrm{Var}(B_i).\end{aligned}$$

Or $B_i \sim \mathcal{B}(p)$, donc $\mathrm{Var}(B_i) = p(1-p)$. Ainsi

$$\boxed{\mathrm{Var}(\widehat{\pi}_n) = \frac{16}{n}p(1-p).}$$

On retrouve bien le comportement général de Monte-Carlo : la variance de la moyenne est divisée par n, et son écart-type est donc d'ordre $1/\sqrt{n}$.

3) Pour estimer cette variance à partir des données simulées, on peut d'abord regarder les variables individuelles

$$Z_i = 4B_i.$$

Leur moyenne empirique vaut précisément

$$\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} 4B_i = \widehat{\pi}_n.$$

On estime alors $\mathrm{Var}(Z_i)$ par la variance empirique corrigée

$$S_Z^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Z_i - \overline{Z}_n\right)^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(4B_i - \widehat{\pi}_n\right)^2.$$

Puisque $\widehat{\pi}_n$ est la moyenne des Z_i, on a théoriquement

$$\mathrm{Var}(\widehat{\pi}_n) = \frac{\mathrm{Var}(Z_i)}{n}.$$

On estime donc la variance de l'estimateur par

$$\boxed{\widehat{\mathrm{Var}}(\widehat{\pi}_n) = \frac{S_Z^2}{n}.}$$

Il existe ici une seconde façon de faire, spécifique au fait que B_i est une Bernoulli. On estime d'abord

$$p = \mathbb{E}(B_i) \quad \text{par} \quad \widehat{p}_n = \frac{1}{n}\sum_{i=1}^{n} B_i,$$

puis on remplace p par $\widehat{p}_n$ dans la formule théorique :

$$\widehat{\mathrm{Var}}(\widehat{\pi}_n) = \frac{16}{n}\widehat{p}_n(1-\widehat{p}_n).$$

Les deux approches sont très proches pour n grand.

4) L'objectif

$$\mathbb{P}\left(|\widehat{\pi}_n - \pi| \leq 0.01\right) \geq 0.95$$

est équivalent à demander

$$\mathbb{P}\left(|\widehat{\pi}_n - \pi| > 0.01\right) \leq 0.05.$$

L'inégalité de Chebyshev donne, pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(|\widehat{\pi}_n - \pi| > \varepsilon\right) \leq \frac{\mathrm{Var}(\widehat{\pi}_n)}{\varepsilon^2} = \frac{16p(1-p)}{n\varepsilon^2}.$$

On ne veut pas utiliser ici la valeur exacte de p. On utilise donc la borne classique

$$p(1-p) \leq \frac{1}{4},$$

qui est valable pour tout $p \in [0,1]$. Il vient

$$\mathbb{P}\left(|\widehat{\pi}_n - \pi| > \varepsilon\right) \leq \frac{16}{n\varepsilon^2} \times \frac{1}{4} = \frac{4}{n\varepsilon^2}.$$

Avec $\varepsilon = 0.01$, une condition suffisante est donc

$$\frac{4}{n(0.01)^2} \leq 0.05.$$

En isolant n,

$$n \geq \frac{4}{0.05(0.01)^2} = \frac{4}{0.05 \times 10^{-4}} = 800000.$$

Cette condition garantit la précision demandée, mais elle est très pessimiste.

5) Le TCL donne une approximation beaucoup plus précise lorsque n est grand. Puisque

$$\mathbb{E}(4B_i) = \pi, \qquad \mathrm{Var}(4B_i) = 16p(1-p),$$

on a approximativement

$$\frac{\widehat{\pi}_n - \pi}{4\sqrt{p(1-p)/n}} \simeq \mathcal{N}(0,1),$$

ou, de façon équivalente,

$$\widehat{\pi}_n - \pi \simeq \mathcal{N}\left(0, \frac{16p(1-p)}{n}\right).$$

Pour une loi normale centrée réduite, environ 95% de la masse est comprise entre -1.96 et 1.96. La demi-largeur de l'intervalle de confiance est donc approximativement

$$1.96 \times 4\sqrt{\frac{p(1-p)}{n}}.$$

On veut que cette quantité soit au plus 0.01, soit

$$1.96 \times 4\sqrt{\frac{p(1-p)}{n}} \leq 0.01.$$

Après élévation au carré et isolement de n,

$$n \geq \left(\frac{1.96 \times 4\sqrt{p(1-p)}}{0.01} \right)^2 .$$

En utilisant $p = \pi/4 \simeq 0.7854$, donc

$$p(1-p) \simeq 0.16855,$$

on trouve

$$n \gtrsim 103599.$$

On retiendra donc environ 104000 simulations. Cette valeur est beaucoup plus petite que la borne de Chebyshev, car le TCL exploite la forme approximativement normale de l'erreur et pas seulement sa variance.

Exercice 3 – Dimensionnement d'une simulation par le TCL

1) Si $U \sim \mathcal{U}([0,1])$, sa densité vaut 1 sur $[0,1]$. Par conséquent,

$$\mathbb{E}[\sin(\sqrt{U})] = \int_0^1 \sin(\sqrt{u})\, du = I.$$

En prenant $U_1, \ldots, U_n$ i.i.d. uniformes sur $[0,1]$, l'estimateur Monte-Carlo naturel est

$$\boxed{\widehat{I}_n = \frac{1}{n} \sum_{k=1}^{n} \sin(\sqrt{U_k}).}$$

Comme chaque terme a pour espérance I,

$$\mathbb{E}(\widehat{I}_n) = I,$$

et l'estimateur est donc sans biais.

2) Pour calculer exactement I, on pose $x = t^2$. Lorsque x varie de 0 à 1, t varie aussi de 0 à 1, et

$$dx = 2t\, dt.$$

Ainsi

$$I = 2 \int_0^1 t \sin t\, dt.$$

On intègre par parties avec $u = t$ et $dv = \sin t\, dt$, donc $du = dt$ et $v = -\cos t$. On obtient

$$\begin{aligned} \int_0^1 t \sin t\, dt &= [-t\cos t]_0^1 + \int_0^1 \cos t\, dt \\ &= -\cos 1 + \sin 1. \end{aligned}$$

Finalement,

$$\boxed{I = 2(\sin 1 - \cos 1) \simeq 0.602337.}$$

3) Posons

$$Z_k = \sin(\sqrt{U_k}).$$

Les Z_k sont i.i.d., d'espérance I et de variance

$$\sigma^2 = \mathrm{Var}(Z_1) = \mathrm{Var}(\sin(\sqrt{U})).$$

Puisque $\widehat{I}_n = n^{-1}\sum_{k=1}^n Z_k$, l'indépendance donne

$$\begin{aligned}\mathrm{Var}(\widehat{I}_n) &= \frac{1}{n^2}\sum_{k=1}^n \mathrm{Var}(Z_k)\\ &= \frac{1}{n^2}\, n\sigma^2 = \boxed{\frac{\sigma^2}{n}}.\end{aligned}$$

Le TCL affirme alors que, pour n grand,

$$\frac{\sqrt{n}(\widehat{I}_n - I)}{\sigma} \simeq \mathcal{N}(0,1).$$

Pour un niveau de confiance de 99%, on laisse environ 0.5% dans chaque queue de la loi normale, d'où le quantile $z_{0.995} \simeq 2.576$. Si σ était connue, l'intervalle serait approximativement

$$\widehat{I}_n \pm 2.576\frac{\sigma}{\sqrt{n}}.$$

En pratique, σ est généralement inconnue. On l'estime par

$$S_n^2 = \frac{1}{n-1}\sum_{k=1}^n \left(Z_k - \overline{Z}_n\right)^2, \qquad \overline{Z}_n = \widehat{I}_n,$$

et on utilise donc

$$\boxed{\widehat{I}_n \pm 2.576\frac{S_n}{\sqrt{n}}.}$$

Le terme $S_n/\sqrt{n}$ est l'estimation de l'écart-type de l'estimateur $\widehat{I}_n$, souvent appelée erreur-type.

4) La demi-largeur théorique de l'intervalle précédent est

$$2.576\frac{\sigma}{\sqrt{n}}.$$

On veut qu'elle soit au plus 0.01, donc

$$2.576\frac{\sigma}{\sqrt{n}} \le 0.01.$$

Comme toutes les quantités sont positives, on peut isoler $\sqrt{n}$:

$$\sqrt{n} \ge \frac{2.576\sigma}{0.01}.$$

En élevant au carré,

$$\boxed{n \ge \left(\frac{2.576\sigma}{0.01}\right)^2.}$$

Cette formule montre explicitement le coût en $1/\sqrt{n}$: si l'on divise la demi-largeur souhaitée par 10, il faut multiplier n par 100.

5) Il reste à calculer σ^2. Par définition,

$$\sigma^2 = \mathbb{E}[Z_1^2] - \mathbb{E}[Z_1]^2 = \mathbb{E}[\sin^2(\sqrt{U})] - I^2.$$

On commence donc par le second moment. Comme précédemment, avec $u = t^2$,

$$\mathbb{E}[\sin^2(\sqrt{U})] = \int_0^1 \sin^2(\sqrt{u})\, du = 2\int_0^1 t\sin^2 t\, dt.$$

On utilise l'identité

$$\sin^2 t = \frac{1-\cos(2t)}{2}.$$

Ainsi

$$\begin{aligned} 2\int_0^1 t\sin^2 t\,dt &= \int_0^1 t(1-\cos(2t))\,dt \\ &= \frac{1}{2} - \int_0^1 t\cos(2t)\,dt. \end{aligned}$$

Pour la dernière intégrale, une intégration par parties donne

$$\begin{aligned} \int_0^1 t\cos(2t)\,dt &= \left[\frac{t\sin(2t)}{2}\right]_0^1 - \frac{1}{2}\int_0^1 \sin(2t)\,dt \\ &= \frac{\sin 2}{2} - \frac{1-\cos 2}{4}. \end{aligned}$$

Par conséquent,

$$\mathbb{E}[\sin^2(\sqrt{U})] = \frac{3}{4} - \frac{\sin 2}{2} - \frac{\cos 2}{4} \simeq 0.399388.$$

D'autre part,

$$I^2 = 4(\sin 1 - \cos 1)^2.$$

Donc

$$\begin{aligned} \sigma^2 &= \frac{3}{4} - \frac{\sin 2}{2} - \frac{\cos 2}{4} - 4(\sin 1 - \cos 1)^2 \\ &\simeq 0.036578, \end{aligned}$$

et ainsi

$$\sigma \simeq 0.191253.$$

On remplace maintenant σ dans la condition obtenue à la question précédente :

$$\begin{aligned} n &\geq \left(\frac{2.576 \times 0.191253}{0.01}\right)^2 \\ &\simeq 2427.21. \end{aligned}$$

Comme n doit être entier, il faut arrondir au supérieur :

$$\boxed{n \geq 2428.}$$

En pratique, lorsque σ n'est pas connue à l'avance, on peut effectuer une première simulation pilote, estimer σ par S_n, puis utiliser cette estimation pour choisir une taille d'échantillon adaptée.

Exercice 4 – Aiguille de Buffon : une autre estimation de π

1) La demi-longueur de l'aiguille vaut 1. Si l'aiguille fait un angle Θ avec les rainures, sa projection dans la direction perpendiculaire aux rainures a pour longueur

$$1 \times \sin\Theta = \sin\Theta.$$

Le centre de l'aiguille est à distance X de la rainure la plus proche. Une extrémité atteint ou dépasse cette rainure exactement lorsque cette distance est plus petite que la projection précédente. L'événement de croisement est donc

$$\boxed{\{X \leq \sin\Theta\}.}$$

2) On conditionne par rapport à Θ. Pour une valeur fixée $\Theta = \theta$, on a $0 \leq \sin\theta \leq 1$, et comme $X \sim \mathcal{U}([0,1])$,

$$\mathbb{P}(X \leq \sin\theta \mid \Theta = \theta) = \sin\theta.$$

En moyennant par rapport à Θ,

$$\begin{aligned} p &= \mathbb{P}(X \leq \sin\Theta) \\ &= \mathbb{E}[\mathbb{P}(X \leq \sin\Theta \mid \Theta)] \\ &= \mathbb{E}(\sin\Theta). \end{aligned}$$

Or $\Theta \sim \mathcal{U}([0,\pi/2])$, donc sa densité vaut $2/\pi$ sur cet intervalle. Ainsi

$$\begin{aligned} \mathbb{E}(\sin\Theta) &= \frac{2}{\pi}\int_0^{\pi/2} \sin\theta\, d\theta \\ &= \frac{2}{\pi}[-\cos\theta]_0^{\pi/2} = \frac{2}{\pi}. \end{aligned}$$

Donc

$$\boxed{p = \frac{2}{\pi}.}$$

3) On définit

$$B_i = \mathbf{1}_{\{X_i \leq \sin\Theta_i\}}.$$

Chaque B_i est une Bernoulli de paramètre $p = 2/\pi$. La moyenne empirique

$$\widehat{p}_n = \frac{1}{n}\sum_{i=1}^{n} B_i$$

est donc un estimateur naturel de p. Comme $p = 2/\pi$, on inverse la relation pour obtenir

$$\pi = \frac{2}{p}.$$

On propose donc

$$\boxed{\widehat{\pi}_n = \frac{2}{\widehat{p}_n}},$$

dès que $\widehat{p}_n > 0$.

4) On a bien

$$\mathbb{E}(\widehat{p}_n) = p,$$

car $\widehat{p}_n$ est une moyenne de Bernoulli de paramètre p. En revanche, appliquer une fonction non linéaire à un estimateur sans biais ne conserve en général pas le sans-biais. Ici la fonction est $g(x) = 2/x$, et

$$\mathbb{E}[g(\widehat{p}_n)] \neq g(\mathbb{E}(\widehat{p}_n)) = g(p) = \pi$$

en général. L'estimateur de π n'est donc pas exactement sans biais pour n fini.
Il est cependant consistant. En effet, la loi des grands nombres donne

$$\widehat{p}_n \longrightarrow p = \frac{2}{\pi}$$

presque sûrement. Comme la fonction $x \mapsto 2/x$ est continue en $p > 0$, on obtient

$$\widehat{\pi}_n = \frac{2}{\widehat{p}_n} \longrightarrow \frac{2}{p} = \pi.$$

5) Par exemple :

```python
import numpy as np
rng = np.random.default_rng(123)

for n in [1_000, 10_000, 100_000]:
    x = rng.uniform(0.0, 1.0, size=n)
    theta = rng.uniform(0.0, np.pi/2, size=n)
    p_hat = np.mean(x <= np.sin(theta))
    print(n, 2.0 / p_hat)
```

Exercice 5 – Volume d'une boule par Monte-Carlo

1) Le vecteur $U_i = (U_i^{(1)}, U_i^{(2)}, U_i^{(3)})$ est uniforme dans le cube $[-1,1]^3$. Ce cube a pour volume

$$2 \times 2 \times 2 = 8.$$

L'indicatrice B_i vaut 1 exactement lorsque U_i tombe dans la boule unité. Par conséquent,

$$\begin{aligned}\mathbb{E}(B_i) &= 1 \times \mathbb{P}(B_i = 1) + 0 \times \mathbb{P}(B_i = 0) \\ &= \mathbb{P}(U_i \text{ est dans la boule}).\end{aligned}$$

Pour une loi uniforme, cette probabilité est le rapport des volumes :

$$\boxed{p := \mathbb{E}(B_i) = \frac{V_3}{8}.}$$

2) La relation précédente donne $V_3 = 8p$. On estime naturellement p par la proportion de points simulés qui tombent dans la boule,

$$\widehat{p}_n = \overline{B}_n = \frac{1}{n}\sum_{i=1}^{n} B_i.$$

On obtient donc l'estimateur

$$\boxed{\widehat{V}_{3,n} = 8\widehat{p}_n = \frac{8}{n}\sum_{i=1}^{n} B_i.}$$

Comme le volume théorique de la boule unité de $\mathbb{R}^3$ est

$$V_3 = \frac{4\pi}{3},$$

on retrouve

$$p = \frac{V_3}{8} = \frac{4\pi/3}{8} = \frac{\pi}{6}.$$

3) Chaque B_i est une Bernoulli de paramètre p, donc

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

Comme

$$\widehat{V}_{3,n} = \frac{8}{n}\sum_{i=1}^{n} B_i,$$

l'indépendance donne

$$\begin{aligned}\mathrm{Var}(\widehat{V}_{3,n}) &= \frac{64}{n^2}\sum_{i=1}^{n}\mathrm{Var}(B_i) \\ &= \frac{64}{n^2}\, np(1-p) \\ &= \boxed{\frac{64}{n}p(1-p)}.\end{aligned}$$

En remplaçant p par $\pi/6$, on peut aussi écrire

$$\mathrm{Var}(\widehat{V}_{3,n}) = \frac{64}{n}\frac{\pi}{6}\left(1 - \frac{\pi}{6}\right).$$

4) En dimension d, le cube $[-1,1]^d$ a volume 2^d. Si U est uniforme dans ce cube et si V_d désigne le volume de la boule unité de $\mathbb{R}^d$, alors

$$p_d := \mathbb{P}(\|U\| \leq 1) = \frac{V_d}{2^d}.$$

Ainsi

$$V_d = 2^d p_d,$$

et, avec $B_i = \mathbf{1}_{\{\|U_i\| \leq 1\}}$,

$$\widehat{V}_{d,n} = \frac{2^d}{n}\sum_{i=1}^{n} B_i.$$

Le problème est que $p_d = V_d/2^d$ devient très petit lorsque la dimension augmente. Autrement dit, une proportion de plus en plus faible des points du cube tombe dans la boule. La plupart des B_i valent alors zéro, et il faut énormément de simulations pour observer suffisamment de points dans la boule. La méthode devient donc peu efficace en grande dimension.

5) Par exemple :

```
import numpy as np
rng = np.random.default_rng(123)

for n in [1_000, 10_000, 100_000]:
    u = rng.uniform(-1.0, 1.0, size=(n, 3))
    inside = np.sum(u**2, axis=1) <= 1.0
    v_hat = 8.0 * np.mean(inside)
    print(n, v_hat)
```

Exercice 6 – TP Python : comparer plusieurs estimateurs d'une même intégrale

Voici une version complète du code. Les parties qui étaient à compléter dans l'énoncé correspondent exactement à la construction des variables aléatoires individuelles dont on prend ensuite la moyenne. C'est important pour lire correctement la sortie de `mc_summary` : la variance affichée est la variance empirique de ces valeurs individuelles, tandis que la variance de leur moyenne est obtenue en divisant cette quantité par n.

```python
import numpy as np
from scipy.integrate import quad

rng = np.random.default_rng(123)


def mc_summary(samples):
    """Renvoie moyenne, variance empirique des valeurs et IC asymptotique 95%."""
    n = len(samples)
    mean = np.mean(samples)
    sd = np.std(samples, ddof=1)
    half_width = 1.96 * sd / np.sqrt(n)
    return mean, sd**2, mean - half_width, mean + half_width


def estimator_uniform(n):
    u = rng.uniform(0.0, 1.0, size=n)
    return np.cos(u**3) * np.exp(-u)


def estimator_exponential(n):
    x = rng.exponential(scale=1.0, size=n)
    return np.cos(x**3) * (x <= 1.0)


def estimator_truncated_exponential(n):
    u = rng.uniform(0.0, 1.0, size=n)
    y = -np.log(1.0 - u * (1.0 - np.exp(-1.0)))
    return (1.0 - np.exp(-1.0)) * np.cos(y**3)

true_value, _ = quad(lambda x: np.cos(x**3) * np.exp(-x), 0.0, 1.0)
print("valeur de reference :", true_value)

estimators = [
    ("uniforme", estimator_uniform),
    ("exponentielle", estimator_exponential),
    ("exp. tronquee", estimator_truncated_exponential),
]

for n in [1_000, 10_000, 100_000]:
    print("\nn =", n)
    for name, estimator in estimators:
        mean, var, lower, upper = mc_summary(estimator(n))
        print(f"{name:15s} moyenne={mean:.6f} var={var:.6f} "
              f"IC=[{lower:.6f}, {upper:.6f}]")
```

Commentaire sur le code

Pour comprendre la fonction `mc_summary`, notons $Z_1, \dots, Z_n$ les valeurs contenues dans `samples`. Alors :

— `mean = np.mean(samples)` calcule

$$\overline{Z}_n = \frac{1}{n}\sum_{k=1}^{n} Z_k,$$

qui est précisément l'estimation Monte-Carlo $\widehat{I}_n$;

— `sd = np.std(samples, ddof=1)` calcule la racine carrée de

$$S_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(Z_k - \overline{Z}_n)^2,$$

qui estime la variance $\mathrm{Var}(Z_1)$ des valeurs individuelles ;

— la variance de la moyenne $\widehat{I}_n$ est $\mathrm{Var}(Z_1)/n$, donc elle est estimée par S_n^2/n ;

— son écart-type est donc estimé par $S_n/\sqrt{n}$, ce qui explique la ligne `half_width = 1.96 * sd / np.sqrt(n)`.

Ainsi, la variable appelée `var` dans l'affichage du programme correspond à S_n^2, et non directement à la variance de $\widehat{I}_n$.

La simulation de l'exponentielle tronquée utilise par ailleurs la méthode d'inversion. Si Y a pour densité

$$p(y) = \frac{\mathrm{e}^{-y}}{1-\mathrm{e}^{-1}}\mathbf{1}_{0\le y\le 1},$$

alors, pour $0 \le y \le 1$,

$$\begin{aligned} F(y) = \mathbb{P}(Y \le y) &= \int_0^y \frac{\mathrm{e}^{-t}}{1-\mathrm{e}^{-1}}\,dt \\ &= \frac{1-\mathrm{e}^{-y}}{1-\mathrm{e}^{-1}}. \end{aligned}$$

Pour simuler Y, on prend $U \sim \mathcal{U}([0,1])$ et on résout $U = F(Y)$:

$$\begin{aligned} U &= \frac{1-\mathrm{e}^{-Y}}{1-\mathrm{e}^{-1}}, \\ U(1-\mathrm{e}^{-1}) &= 1-\mathrm{e}^{-Y}, \\ \mathrm{e}^{-Y} &= 1-U(1-\mathrm{e}^{-1}), \\ Y &= -\log\left(1-U(1-\mathrm{e}^{-1})\right). \end{aligned}$$

C'est cette dernière formule qui apparaît dans la ligne `y = -np.log(...)`.

La valeur de référence obtenue par quadrature numérique est

$$I \simeq 0.603395.$$

Cette valeur sert uniquement à vérifier numériquement les estimations Monte-Carlo.

Pour interpréter les variances affichées par le programme, reprenons les trois variables individuelles

$$\begin{aligned} Z_1 &= \cos(U^3)\mathrm{e}^{-U}, \\ Z_2 &= \cos(X^3)\mathbf{1}_{\{X\le 1\}}, \\ Z_3 &= (1-\mathrm{e}^{-1})\cos(Y^3). \end{aligned}$$

Dans chaque cas, le programme simule $Z_1, \dots, Z_n$, puis calcule leur moyenne. La quantité affichée sous le nom `var` doit donc être comparée à la variance théorique de la variable individuelle correspondante.

Pour la représentation uniforme,

$$\begin{aligned} \mathbb{E}(Z_1^2) &= \int_0^1 \cos^2(x^3)\mathrm{e}^{-2x}\,dx \simeq 0.410537, \\ \mathrm{Var}(Z_1) &= \mathbb{E}(Z_1^2) - I^2 \simeq 0.410537 - (0.603395)^2 \\ &\simeq \boxed{0.04645}. \end{aligned}$$

La variance de l'estimateur lui-même est donc

$$\mathrm{Var}(\widehat{I}_n^{(1)}) \simeq \frac{0.04645}{n}.$$

Pour l'exponentielle non tronquée,

$$\begin{aligned}\mathbb{E}(Z_2^2) &= \mathbb{E}\Big[\cos^2(X^3)\mathbf{1}_{\{X\le 1\}}\Big] \\ &= \int_0^1 \cos^2(x^3)\mathrm{e}^{-x}\,dx \simeq 0.581483, \\ \mathrm{Var}(Z_2) &= 0.581483 - (0.603395)^2 \simeq \boxed{0.21740}.\end{aligned}$$

Ainsi

$$\mathrm{Var}(\widehat{I}_n^{(2)}) \simeq \frac{0.21740}{n}.$$

Pour l'exponentielle tronquée,

$$\begin{aligned}\mathbb{E}(Z_3^2) &= (1-\mathrm{e}^{-1})^2\mathbb{E}[\cos^2(Y^3)] \\ &= (1-\mathrm{e}^{-1})\int_0^1 \cos^2(y^3)\mathrm{e}^{-y}\,dy \\ &\simeq 0.367603,\end{aligned}$$

d'où

$$\begin{aligned}\mathrm{Var}(Z_3) &= 0.367603 - (0.603395)^2 \\ &\simeq \boxed{0.003481}.\end{aligned}$$

La variance de l'estimateur vaut donc

$$\mathrm{Var}(\widehat{I}_n^{(3)}) \simeq \frac{0.003481}{n}.$$

On doit donc observer dans le programme des variances empiriques proches de

$$0.04645, \qquad 0.21740, \qquad 0.003481,$$

avec des fluctuations dues au caractère aléatoire de la simulation. Lorsque n augmente, ces variances empiriques se stabilisent autour de leurs valeurs théoriques, tandis que les intervalles de confiance se resserrent comme $1/\sqrt{n}$.

Les trois estimateurs sont sans biais, mais le troisième est nettement plus efficace : sa variance est environ 13 fois plus petite que celle de la représentation uniforme et plus de 60 fois plus petite que celle de l'exponentielle non tronquée. L'explication est simple : l'estimateur exponentiel non tronqué gaspille les tirages pour lesquels $X > 1$, puisque leur contribution est alors nulle. L'exponentielle tronquée concentre au contraire tous les tirages dans la région utile $[0, 1]$, tout en corrigeant exactement la normalisation par le facteur $1 - \mathrm{e}^{-1}$.