

TD 2 : INTÉGRATION MONTE-CARLO

EXERCICE 1 (Échantillonnage préférentiel)

Soit X une variable aléatoire suivant une loi de Pareto de paramètres $(1, 3)$ (dont une densité est $f : x \in \mathbb{R} \mapsto 3x^{-4}\mathbf{1}_{x \geq 1}$). On s'intéresse à l'estimation de $p = \mathbb{P}(X \geq s)$ pour $s > 1$ fixé.

1. Déterminer la valeur de p .

Correction On obtient par intégration : $p = s^{-3}$.

2. Proposer un estimateur Monte-Carlo et déterminer sa loi asymptotique.

Correction Si (X_n) est une suite i.i.d. de loi de Pareto de paramètres $(1, 3)$, un estimateur Monte-Carlo de p est : $\hat{p}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i \geq s}$. On a alors, par le TCL : $\sqrt{n}(\hat{p}_n - p) \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, p(1-p))$.

3. Proposer un estimateur Monte-Carlo par échantillonnage préférentiel construit sur une loi exponentielle translatée. Est-il satisfaisant ?

Correction On considère une loi exponentielle de paramètre 1 translatée de s , dont une densité est : $g : x \in \mathbb{R} \mapsto e^{x-s}\mathbf{1}_{x \geq s}$. En appelant $\text{Leb}_{\mathbb{R}}$ la mesure de Lebesgue sur $\mathbb{R}$, il est clair que l'hypothèse de domination $f\mathbf{1}_{\geq s} \cdot \text{Leb}_{\mathbb{R}} \ll g \cdot \text{Leb}_{\mathbb{R}}$ est vérifiée. Ainsi $\tilde{p}_n = \frac{1}{n} \sum_{i=1}^n f(Y_i)/g(Y_i)\mathbf{1}_{Y_i \geq s}$, avec (Y_n) est une suite i.i.d. de loi de $\text{Exp}(1) + s$ est licite et consistant.

Cependant

$$\mathbb{E} \left[\left(\frac{f(Y_1)}{g(Y_1)} \mathbf{1}_{Y_1 \geq s} \right)^2 \right] = \int_s^\infty 9y^{-8}e^{y-s}dy = \infty.$$

Ainsi, $\text{Var} \left(\frac{f(Y_1)}{g(Y_1)} \mathbf{1}_{Y_1 \geq s} \right)$ n'existe pas et on ne peut n'y appliquer le TCL à $\tilde{p}_n$, ni obtenir d'intervalles de confiance asymptotiques.

EXERCICE 2 (Somme d'exponentielles et conditionnement)

Soient X et Y deux variables aléatoires indépendantes de lois respectives $\text{Exp}(\lambda)$ et $\text{Exp}(\mu)$, pour $\lambda > 0$ et $\mu > 0$. On cherche à estimer $p = \mathbb{P}(S > s)$, où $S = X + Y$ et $s > 0$ est fixé.

1. Proposer un estimateur Monte-Carlo $\hat{p}_n$ de p .

Correction Si (X_n) et (Y_n) sont deux suites i.i.d de loi respectives $\text{Exp}(\lambda)$ et $\text{Exp}(\mu)$, alors on peut choisir comme estimateur $\hat{p}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i + Y_i > s}$.

2. Déterminer $\mathbb{P}(S > s | Y = y)$ pour tout $y > 0$. En déduire un estimateur Monte-Carlo par conditionnement $\tilde{p}_n$.

Correction Puisque X et Y sont indépendantes, on a pour tout $y > 0$

$$\mathbb{P}(S > s | Y = y) = \mathbb{P}(X > s - Y | Y = y) = \mathbb{P}(X > s - y) = e^{\lambda(y-s)}\mathbf{1}_{y \leq s} + \mathbf{1}_{y > s}.$$

Ensuite, puisque

$$p = \mathbb{E}[\mathbb{P}(S > s | Y)] = \mathbb{E}[e^{\lambda(Y-s)}\mathbf{1}_{Y \leq s}] + \mathbb{E}[\mathbf{1}_{Y > s}],$$

un estimateur par conditionnement de p est $\tilde{p}_n = \frac{1}{n} \sum_{i=1}^n (e^{\lambda(Y_i-s)}\mathbf{1}_{Y_i \leq s} + \mathbf{1}_{Y_i > s})$.

EXERCICE 3 (Estimation d'une p -valeur)

On considère un vecteur gaussien à deux dimensions $X = (U, V) \sim \mathcal{N}(\theta, I_2)$, où $\theta \in \mathbb{R}^2$ et I_2 est la matrice identité de dimension 2. On souhaite tester $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$. Pour ce faire, on considère la statistique de test $S(X) = \max(|U|, |V|)$.

1. On note Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$ et F celle de la variable aléatoire $S(X)$ sous H_0 . Exprimer F en fonction de Φ puis proposer un test $T(X)$ de taille $\alpha \in]0, 1[$ pour décider entre les hypothèses H_0 et H_1 .

Correction On se place sous H_0 , i.e. X est un vecteur gaussien standard en dimension 2. Puisque $S(X)$ est une variable aléatoire positive, on a $F(t) = 0$ pour $t \leq 0$. Pour $t \geq 0$, on voit facilement que

$$F(t) = (2\Phi(t) - 1)^2.$$

Soit $\alpha \in]0, 1[$. On considère un test de la forme $T(X) = \mathbf{1}_{S(X) > q}$, avec $q \in \mathbb{R}$. Pour qu'il soit de taille α , on vérifie :

$$\mathbb{P}(S(X) > q) = \alpha \iff 1 - F(q) = \alpha \iff q = \Phi^{-1} \left(\frac{1 + \sqrt{1 - \alpha}}{2} \right).$$

D'où le test $T(X) = \mathbf{1}_{S(X) > \Phi^{-1}((1 + \sqrt{1 - \alpha})/2)}$.

2. Soit $x \in \mathbb{R}^2$ une réalisation de X . Déterminer la p -valeur $\alpha_0(x)$ du test T pour la réalisation x .

Correction D'après ce qui précède,

$$T(X) = 1 \iff \alpha > 1 - [2\Phi(S(x)) - 1]^2.$$

Ainsi la p -valeur vaut $\alpha_0(x) = 1 - [2\Phi(S(x)) - 1]^2$.

3. On rappelle que, puisqu'ici le test est de la forme $T(X) = \mathbf{1}_{S(X) > F^{-1}(1 - \alpha)}$, avec F continue et strictement croissante sur son support, la p -valeur s'écrit aussi $\alpha_0(x) = \mathbb{P}_{\theta_0}(S(X) \geq S(x))$. En déduire un estimateur Monte-Carlo de la p -valeur.

Correction On peut obtenir un estimateur Monte-Carlo en simulant (Y_n) i.i.d selon la loi $\mathcal{N}(0, 1)$ et en définissant la n^e observation de la façon suivante : $X_n = (Y_{2n-1} + \theta_1, Y_{2n} + \theta_2)$. Un estimateur Monte-Carlo de $\alpha_0(x)$ est alors $\sum_{i=1}^n \mathbf{1}_{S(X_i) \geq S(x)} / n$.

4. On suppose maintenant que $X \sim \mathcal{N}(\theta, \Sigma)$, où Σ est une matrice symétrique et définie positive, et on souhaite toujours tester $H_0 : \theta = 0$ contre $H_1 : \theta \neq 0$, avec la même statistique de test $S(X) = \max(|U|, |V|)$. La formule théorique de la p -valeur n'est plus aussi simple et on ne cherchera pas à l'explicitier. Proposer cependant une manière de l'approcher (on admettra que F est dans ce cas continue et strictement croissante).

Correction Puisque F est dans ce cas continue et strictement croissante, on a toujours la formule $\alpha_0(x) = \mathbb{P}_{\theta_0}(S(X) \geq S(x))$ et il suffit donc de simuler (X_n) i.i.d selon la loi $\mathcal{N}(0, \Sigma)$ et de calculer $\sum_{i=1}^n \mathbf{1}_{S(X_i) \geq S(x)} / n$.

EXERCICE 4 ((*) Test du rapport de vraisemblance)

Soient $N \in \mathbb{N}^*$, $\theta \in]0, 1[$ et $X \sim \text{Bin}(N, \theta)$. Pour $\theta_0 \in]0, 1[$, on souhaite tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ au niveau $\alpha \in]0, 1[$. Pour ce faire, on considère la statistique

$$S(X) = -2 \log \frac{\theta_0^X (1 - \theta_0)^{N-X}}{\hat{\theta}_N^X (1 - \hat{\theta}_N)^{N-X}},$$

où $\hat{\theta}_N$ est l'estimateur du maximum de vraisemblance (on prend comme d'habitude la convention usuelle $0^0 = 1$), et le test du rapport de vraisemblance : $T(X) = \mathbf{1}_{S(X) > F^{-1}(1-\alpha)}$, où F est la fonction de répartition de $S(X)$ sous H_0 .

1. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_N$.

Correction On a ici un modèle de lois discrètes $\{\text{Bin}(N, t), t \in]0, 1[\}$, dominé par la mesure de comptage sur $\mathbb{N}$, et dont une densité d'une loi paramétrée par $t \in]0, 1[$ est $x \in \mathbb{N} \mapsto \binom{N}{x} t^x (1-t)^{N-x} \mathbf{1}_{x \leq N}$. Un calcul classique montre ensuite que $\hat{\theta}_N = X/N$ (il n'y a qu'une seule observation!).

2. De manière générale, considérons un modèle statistique paramétrique $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$, un échantillon observé $\mathbf{X}$ de loi P_θ (pour un certain $\theta \in \Theta$) et une alternative de la forme $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$, pour $\theta_0 \in \Theta$ fixé. Soient une statistique de test $S(\mathbf{X})$, de fonction de répartition F_θ et, pour $\alpha \in]0, 1[$, le test : $T(\mathbf{X}) = \mathbf{1}_{S(\mathbf{X}) > F_{\theta_0}^{-1}(1-\alpha)}$. Montrer que $T(\mathbf{X})$ est un test de taille $1 - F_{\theta_0}(F_{\theta_0}^{-1}(1-\alpha))$ et donc de niveau α .

Correction Le niveau est donné par la majoration :

$$\sup_{\theta \in \{\theta_0\}} \mathbb{P}(T(\mathbf{X}) = 1) = \mathbb{P}_{\theta_0}(S(\mathbf{X}) > F_{\theta_0}^{-1}(1-\alpha)) = 1 - F_{\theta_0}(F_{\theta_0}^{-1}(1-\alpha)).$$

Or par définition de l'inverse généralisée (on peut aussi faire un dessin!) : $F_{\theta_0}(F_{\theta_0}^{-1}(1-\alpha)) \geq 1-\alpha$, donc $\sup_{\theta \in \{\theta_0\}} \mathbb{P}(T(\mathbf{X}) = 1) \leq \alpha$.

3. Supposons que la statistique $S(X)$ n'ait qu'un nombre fini d'issues, notées $\{v_1, \dots, v_m\}$. Soit alors $\delta = \min_{1 \leq i, j \leq m} |v_j - v_i|$. En remarquant que $F_{\theta_0}^{-1}(1-\alpha) \in \{v_1, \dots, v_m\}$ et que $S(\mathbf{X}) > F_{\theta_0}^{-1}(1-\alpha) \iff S(\mathbf{X}) - \delta \geq F_{\theta_0}^{-1}(1-\alpha)$, montrer que, pour une réalisation x de $\mathbf{X}$, la p -valeur $\alpha_0(x)$ vaut :

$$\alpha_0(x) = \mathbb{P}_{\theta_0}(S(X) \geq S(x)).$$

Correction Pour une réalisation x de $\mathbf{X}$, la p -valeur $\alpha_0(x)$ vaut :

$$\begin{aligned} \alpha_0(x) &= \inf \{ \alpha \in [0, 1], S(x) = 1 \} \\ &= \inf \{ \alpha \in [0, 1], S(x) > F_{\theta_0}^{-1}(1-\alpha) \} \\ &= \inf \{ \alpha \in [0, 1], S(x) - \delta \geq F_{\theta_0}^{-1}(1-\alpha) \} \\ &= \inf \{ \alpha \in [0, 1], F_{\theta_0}(S(x) - \delta) \geq 1-\alpha \} \quad (\text{propriété fondamentale de l'inverse généralisée}) \\ &= 1 - F_{\theta_0}(S(x) - \delta) \\ &= \mathbb{P}_{\theta_0}(S(\mathbf{X}) > S(x) - \delta) \\ &= \mathbb{P}_{\theta_0}(S(\mathbf{X}) \geq S(x)). \end{aligned}$$

4. Revenons au cadre particulier qui nous intéresse et considérons une réalisation $x \in \{0, \dots, N\}$ de X . Proposer un estimateur Monte-Carlo visant à approcher la p -valeur $\alpha_0(x)$.

Correction Le test qui nous intéresse rentre dans le cadre des deux questions précédentes. On peut donc affirmer que la p -valeur s'écrit $\alpha_0(x) = \mathbb{P}_{\theta_0}(S(X) \geq S(x))$, dont on peut obtenir un estimateur Monte-Carlo en simulant (X_n) i.i.d selon la loi $\text{Bin}(N, \theta_0)$ et en calculant $(\sum_{i=1}^n \mathbf{1}_{S(X_i) \geq S(x)})/n$.

5. Soit la fonction

$$\varphi : x \in]0, 1[\mapsto -2 \log \frac{\theta_0^x (1 - \theta_0)^{1-x}}{x^x (1-x)^{1-x}},$$

de sorte que $S(X) = N\varphi(\hat{\theta}_N)$. Montrer qu'un développement limité de φ en θ_0 donne :

$$\forall x \in]0, 1[, \quad \varphi(x) = \frac{1}{\theta_0(1-\theta_0)}(x - \theta_0)^2 + (x - \theta_0)^2 r(x),$$

où $r :]0, 1[\rightarrow \mathbb{R}$ est une fonction telle que $\lim_{x \rightarrow \theta_0} r(x) = 0$.

Correction On cherche à réaliser un développement limité à l'ordre 2 de φ . Pour ce faire, on remarque que φ est bien deux fois dérivable et, pour tout $x \in]0, 1[$:

$$\varphi'(x) = -2 \left[\log \frac{\theta_0}{1-\theta_0} - \log \frac{x}{1-x} \right], \quad \text{et} \quad \varphi''(x) = \frac{2}{x(1-x)}.$$

Il vient donc, pour tout $x \in]0, 1[$:

$$\begin{aligned} \varphi(x) &= \varphi(\theta_0) + \varphi'(\theta_0)(x - \theta_0) + \frac{\varphi''(\theta_0)}{2}(x - \theta_0)^2 + o((x - \theta_0)^2) \\ &= \frac{1}{\theta_0(1-\theta_0)}(x - \theta_0)^2 + o((x - \theta_0)^2). \end{aligned}$$

6. En remarquant que l'on peut prolonger r par continuité de sorte que $r(\theta_0) = 0$, conclure que, sous H_0 , $S(X) \xrightarrow[N \rightarrow \infty]{\text{Loi}} \chi_1^2$.

Correction De la question précédente, il vient :

$$S(X) = N\varphi(\hat{\theta}_N) = \left[\frac{\sqrt{N}(\hat{\theta}_N - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}} \right]^2 + \left[\sqrt{N}(\hat{\theta}_N - \theta_0) \right]^2 r(\hat{\theta}_N).$$

Or, sous H_0 , X ayant la même loi qu'une somme de N variables aléatoires indépendantes de loi $\text{Ber}(\theta_0)$, le TCL assure que

$$\frac{\sqrt{N}(\hat{\theta}_N - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}} \xrightarrow[N \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, 1).$$

Par suite, le théorème de continuité permet d'affirmer que

$$\left[\frac{\sqrt{N}(\hat{\theta}_N - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}} \right]^2 \xrightarrow[N \rightarrow \infty]{\text{Loi}} \chi_1^2,$$

et que $[\sqrt{N}(\hat{\theta}_N - \theta_0)]^2 r(\hat{\theta}_N) \xrightarrow[N \rightarrow \infty]{\mathbb{P}} 0$. Le lemme de Slutsky permet de conclure.

7. Donner une autre expression de la p -valeur $\alpha_0(x)$ et en déduire une manière de l'approcher lorsque N est grand.

Correction D'après ce qui a été écrit dans le cas général,

$$T(X) = 1 \iff F(S(X) - \delta) \geq 1 - \alpha \iff \alpha \geq 1 - F(S(X) - \delta).$$

Ainsi p -valeur vaut $\alpha_0(x) = 1 - F(S(x)^-) \approx 1 - F_{\chi_1^2}(S(x))$, où on a noté $F_{\chi_1^2}$ la fonction de répartition de la loi χ_1^2 .

EXERCICE 5 (Variables antithétiques)

Nous ne présentons pas cette méthode dans le cadre le plus général, mais plutôt par l'intermédiaire d'une application classique. On souhaite estimer $I = \mathbb{E}[\varphi(X)]$ avec X une variable aléatoire réelle telle que $\mathbb{E}[\varphi(X)^2] < \infty$, F sa fonction de répartition et on suppose disposer d'une formule pour la fonction quantile F^{-1} . Dans ce cas, on sait qu'une façon de simuler X est de considérer U uniforme sur $[0, 1]$ et de poser $X := F^{-1}(U)$. L'estimateur Monte-Carlo standard s'écrit alors

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \varphi(X_i) = \frac{1}{n} \sum_{i=1}^n \varphi(F^{-1}(U_i)),$$

où (U_n) est une suite i.i.d de loi $\text{Unif}([0, 1])$ et pour tout $n \geq 1$, $X_n = F^{-1}(U_n)$. L'estimateur de I par variables antithétiques est défini par

$$\tilde{I}_n := \frac{1}{n} \sum_{i=1}^n \frac{\varphi(F^{-1}(U_i)) + \varphi(F^{-1}(1 - U_i))}{2}.$$

1. Montrer que $\tilde{I}_n$ est sans biais, fortement consistant et asymptotiquement normal, c'est-à-dire

$$\sqrt{n} (\tilde{I}_n - I) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, s^2),$$

avec

$$s^2 := \frac{1}{2} [\text{Var}(\varphi(F^{-1}(U))) + \text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1 - U)))].$$

Correction Notons

$$Y_i := \frac{\varphi(F^{-1}(U_i)) + \varphi(F^{-1}(1 - U_i))}{2},$$

si bien que

$$\tilde{I}_n = \frac{1}{n} \sum_{i=1}^n Y_i.$$

Les variables Y_i sont clairement i.i.d., de même loi qu'une variable générique

$$Y := \frac{\varphi(F^{-1}(U)) + \varphi(F^{-1}(1 - U))}{2}.$$

Puisque les variables U et $(1 - U)$ ont la même loi (i.e. la loi uniforme sur $[0, 1]$), les variables $\varphi(F^{-1}(U))$ et $\varphi(F^{-1}(1 - U))$ ont aussi même loi, donc même espérance et celle-ci vaut

$$\mathbb{E}[\varphi(F^{-1}(U))] = \mathbb{E}[\varphi(X)] = I,$$

ce qui prouve que $\tilde{I}_n$ est sans biais. La consistance forte découle de la loi forte des grands nombres appliquée aux variables Y_i et la normalité asymptotique du TCL. La variance asymptotique vaut

$$s^2 = \text{Var}(Y) = \frac{1}{4} \text{Var}(\varphi(F^{-1}(U)) + \varphi(F^{-1}(1-U))),$$

donc, comme les variables $\varphi(F^{-1}(U))$ et $\varphi(F^{-1}(1-U))$ ont la même variance,

$$s^2 = \frac{1}{2} [\text{Var}(\varphi(F^{-1}(U))) + \text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1-U)))] .$$

2. Montrer que $s^2 \leq \sigma^2 := \text{Var}(\varphi(X))$.

Correction Commençons par noter que

$$\text{Var}(\varphi(F^{-1}(U))) = \text{Var}(\varphi(X)) = \sigma^2 \quad \text{et} \quad \text{Var}(\varphi(F^{-1}(1-U))) = \sigma^2.$$

Par ailleurs, l'inégalité de Cauchy-Schwarz donne

$$\text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1-U))) \leq \sqrt{\text{Var}(\varphi(F^{-1}(U)))} \times \sqrt{\text{Var}(\varphi(F^{-1}(1-U)))} = \sigma^2,$$

d'où le fait que $s^2 \leq \sigma^2$.

3. Soit $\alpha \in]0, 1[$. Déterminer un estimateur fortement consistant $\tilde{s}_n^2$ de s^2 et en déduire un intervalle de confiance de niveau asymptotique $(1 - \alpha)$ pour I .

Correction Si l'on définit

$$\tilde{s}_n^2 := \frac{1}{4n} \sum_{i=1}^n [\varphi(F^{-1}(U_i)) + \varphi(F^{-1}(1-U_i))]^2 - \tilde{I}_n^2,$$

la loi des grands nombres et le Théorème de continuité assurent que

$$\tilde{s}_n^2 \xrightarrow[n \rightarrow \infty]{\text{p.s.}} \frac{1}{4} \mathbb{E} [(\varphi(F^{-1}(U)) + \varphi(F^{-1}(1-U)))^2] - I^2,$$

ce qui s'écrit encore

$$\frac{1}{2} \{ (\mathbb{E} [\varphi(F^{-1}(U))^2] - I^2) + (\mathbb{E} [\varphi(F^{-1}(U))\varphi(F^{-1}(1-U))] - I^2) \} = s^2.$$

Un intervalle de confiance est finalement $[\tilde{I}_n \pm q\tilde{s}_n/\sqrt{n}]$, où q est le quantile de niveau $(1 - \alpha/2)$ de la loi $\mathcal{N}(0, 1)$.

4. Application. Considérons l'exemple jouet suivant : soit $q > \log 2$ fixé connu, $X \sim \text{Exp}(1)$ et $p := \mathbb{P}(X > q) = e^{-q}$. Montrer que $s^2 = (\sigma^2 - p^2)/2$ où σ^2 est la variance asymptotique du Monte-Carlo standard et s^2 celle obtenue par variables antithétiques.

Correction Avec les notations ci-dessus, on a donc $F^{-1}(U) = -\log U$ et $\varphi(x) = \mathbf{1}_{x>q}$. L'estimateur Monte-Carlo naïf est

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i > q} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{-\log U_i > q},$$

tandis que l'estimateur de I par variables antithétiques est défini par

$$\tilde{I}_n := \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{1}_{-\log U_i > q} + \mathbf{1}_{-\log(1-U_i) > q}}{2}.$$

La variance asymptotique du Monte-Carlo naïf est

$$\sigma^2 = \text{Var}(\mathbf{1}_{X > q}) = p(1-p),$$

tandis que celle obtenue par variables antithétiques est

$$s^2 = \frac{1}{2} [\text{Var}(\varphi(F^{-1}(U))) + \text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1-U)))],$$

c'est-à-dire

$$s^2 = \frac{1}{2} [p(1-p) + \text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1-U)))].$$

Or

$$\text{Cov}(\varphi(F^{-1}(U)), \varphi(F^{-1}(1-U))) = \mathbb{E}[\mathbf{1}_{-\log U > q} \mathbf{1}_{-\log(1-U) > q}] - p^2 = -p^2,$$

car, en tenant compte du fait que $q > \log 2$,

$$\mathbb{E}[\mathbf{1}_{-\log U > q} \mathbf{1}_{-\log(1-U) > q}] = \mathbb{P}(U < e^{-q}, U > 1 - e^{-q}) = 0.$$

Bilan : l'estimation par variables antithétiques a permis de gagner un facteur supérieur à 2 en terme de réduction de la variance asymptotique.

EXERCICE 6 ((*) Preuve par couplage de l'inégalité de covariance de Tchebychev)

Soit X une variable aléatoire, φ une fonction croissante et ψ une fonction décroissante telles que $\varphi(X)$ et $\psi(X)$ soient de carré intégrable.

1. Soit X' une variable aléatoire de même loi que X et indépendante de celle-ci. Montrer que

$$\text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) = \mathbb{E}[(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X'))] \leq 0.$$

Correction Puisque X et X' ont même loi, il en va de même pour les variables $\varphi(X)$ et $\varphi(X')$ d'une part, et $\psi(X)$ et $\psi(X')$ d'autre part, lesquelles ont donc la même moyenne, d'où

$$\begin{aligned} & \text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) \\ &= \mathbb{E}[(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X'))] - \mathbb{E}[\varphi(X) - \varphi(X')] \mathbb{E}[\psi(X) - \psi(X')] \\ &= \mathbb{E}[(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X'))]. \end{aligned}$$

Ensuite, sur l'événement $\{X \geq X'\}$, par croissance de φ et décroissance de ψ , $\varphi(X) - \varphi(X') \geq 0$ et $\psi(X) - \psi(X') \leq 0$, donc $(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X')) \leq 0$. De même, sur l'événement $\{X \leq X'\}$, $(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X')) \leq 0$. Ainsi $(\varphi(X) - \varphi(X'))(\psi(X) - \psi(X')) \leq 0$ p.s. et

$$\text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) \leq 0.$$

2. Montrer par ailleurs que : $\text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) = 2 \text{Cov}(\varphi(X), \psi(X))$, et en déduire l'inégalité de covariance de Tchebychev, à savoir que : $\text{Cov}(\varphi(X), \psi(X)) \leq 0$.

Correction Puisque X et X' sont indépendantes, toute fonction de X est indépendante de toute fonction de X' , donc

$$\text{Cov}(\varphi(X), \psi(X')) = \text{Cov}(\varphi(X'), \psi(X)) = 0.$$

La bilinéarité de la covariance donne

$$\text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) = \text{Cov}(\varphi(X), \psi(X)) + \text{Cov}(\varphi(X'), \psi(X')),$$

et le fait que X et X' ont même loi assure donc que

$$\text{Cov}(\varphi(X) - \varphi(X'), \psi(X) - \psi(X')) = 2 \text{Cov}(\varphi(X), \psi(X)).$$

De cette égalité et de l'inégalité de la question précédente, on déduit bien l'inégalité de covariance de Tchebychev :

$$\text{Cov}(\varphi(X), \psi(X)) \leq 0.$$

3. En déduire le principe des variables antithétiques en Monte-Carlo, à savoir que si φ est monotone et h décroissante telle que $h(X)$ ait même loi que X alors, sous réserve d'existence des variances,

$$\text{Var}\left(\frac{\varphi(X) + \varphi(h(X))}{2}\right) \leq \frac{1}{2} \text{Var}(\varphi(X)).$$

Correction Si φ est croissante (resp. décroissante) et h décroissante, alors $\varphi \circ h$ est décroissante (resp. croissante). Dans tous les cas, φ et $\varphi \circ h$ ont des monotonies opposées donc, d'après le résultat précédent,

$$\text{Cov}(\varphi(X), \varphi(h(X))) \leq 0.$$

Or, par bilinéarité,

$$\text{Var}\left(\frac{\varphi(X) + \varphi(h(X))}{2}\right) = \frac{1}{4} (\text{Var}(\varphi(X)) + 2 \text{Cov}(\varphi(X), \varphi(h(X))) + \text{Var}(\varphi(h(X)))),$$

d'où

$$\text{Var}\left(\frac{\varphi(X) + \varphi(h(X))}{2}\right) \leq \frac{1}{4} (\text{Var}(\varphi(X)) + \text{Var}(\varphi(h(X)))) = \frac{1}{2} \text{Var}(\varphi(X)),$$

la dernière égalité étant due au fait que $\varphi(X)$ et $\varphi(h(X))$ ont même loi par hypothèse.

4. Exemple jouet : proposer une méthode de variables antithétiques pour estimer $\mathbb{E}[\exp X]$ lorsque $X \sim \mathcal{N}(0, 1)$.

Correction Si $X \sim \mathcal{N}(0, 1)$, alors $h(X) = -X \sim \mathcal{N}(0, 1)$. La fonction h étant décroissante, la fonction exponentielle croissante, on peut appliquer ce qui précède. L'estimateur antithétique s'écrit

$$\tilde{I}_n = \frac{1}{n} \sum_{i=1}^n \frac{e^{X_i} + e^{-X_i}}{2}, \quad (1)$$

où les X_i sont i.i.d. gaussiennes centrées réduites. Notons que la variance asymptotique

$$s^2 := \frac{1}{2} (\text{Var}(\exp X) + \text{Cov}(\exp X, \exp(-X))) = \frac{1}{2} (\mathbb{E}[e^{2X}] + 1 - 2\mathbb{E}[e^{-X}]^2)$$

est facile à calculer ($\mathbb{E}[e^{2X}] = e^2$ et $\mathbb{E}[e^X] = e^{1/2}$) et donne, en notant $\sigma^2 = \text{Var}(e^X) = e(e-1)$ la variance asymptotique du Monte-Carlo standard,

$$s^2 = \frac{1}{2}(e-1)^2 \implies \frac{s^2}{\sigma^2} = \frac{e-1}{2e} \approx 0.32,$$

soit une réduction de la variance par un facteur 3. Même en tenant compte du fait que cette méthode antithétique est environ deux fois plus coûteuse en calculs (évaluation de $\exp X_i$ et de $\exp(-X_i)$), elle reste avantageuse.

EXERCICE 7 ((†) Marche aléatoire avec dérive et échantillonnage préférentiel)

Soit $(X_n)_{n \geq 1}$ une suite de variables gaussiennes indépendantes $\mathcal{N}(-m, 1)$ avec $m > 0$, et $S_n = X_1 + \dots + X_n$. Les réels $a < 0$ et $b > 0$ étant fixés, on considère le temps d'arrêt $N = \min\{n : S_n < a \text{ ou } S_n > b\}$ et on s'intéresse à la probabilité $p = \mathbb{P}(S_N > b)$ que la marche aléatoire (S_n) sorte de l'intervalle $[a, b]$ par le haut.

1. Montrer que le temps d'arrêt N est presque sûrement fini.

Correction Par continuité monotone décroissante, on a

$$\mathbb{P}(N = \infty) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \{N > n\}\right) = \lim_{n \rightarrow \infty} \mathbb{P}(N > n),$$

or $S_n \sim \mathcal{N}(-nm, n)$ et

$$\mathbb{P}(N > n) \leq \mathbb{P}(S_n > a) = \mathbb{P}(\mathcal{N}(-nm, n) > a) = \mathbb{P}(\mathcal{N}(0, 1) > (a + nm)/\sqrt{n}) \xrightarrow[n \rightarrow \infty]{} 0.$$

2. Proposer un estimateur $\hat{p}_k$ de p basé sur k simulations successives de trajectoires. Intuitivement, pourquoi cet estimateur n'est-il pas satisfaisant ?

Correction On simule k trajectoires $S(1), \dots, S(k)$, avec $S(j) = (S_n(j))_{1 \leq n \leq N(j)}$ la trajectoire obtenue à la j -ème simulation. Pour celle-ci, notons

$$I(j) = \begin{cases} 0 & \text{si } S_{N(j)} < a \\ 1 & \text{si } S_{N(j)} > b \end{cases}$$

L'estimateur Monte-Carlo standard de p s'écrit alors

$$\hat{p}_k = \frac{1}{k} \sum_{j=1}^k I(j).$$

Puisqu'une trajectoire typique a tendance à descendre, la probabilité p est très faible et cette méthode Monte-Carlo standard ne fonctionne pas : à moins de prendre k très grand, on ne verra pas une seule trajectoire sortir par le haut. En terme de variance asymptotique, c'est le problème classique d'estimation d'événement rare par Monte-Carlo standard. L'écart-type asymptotique $\sigma = \sqrt{p(1-p)}$ est beaucoup plus grand que la quantité p à estimer.

3. Plutôt que de simuler X suivant la densité d'une loi $\mathcal{N}(-m, 1)$, on simule X' suivant la loi $\mathcal{N}(m, 1)$. On note (S'_n) la marche aléatoire obtenue et N' le temps de sortie de l'intervalle $[a, b]$. Notons le rapport de vraisemblance $w(s) := f(s)/g(s)$, où f et g sont les densités respectives des variables aléatoires S_N et $S'_{N'}$. Le but est de montrer que $w(s) = \exp(-2ms)$.

- (a) Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$ une fonction test, montrer que

$$\mathbb{E}[\varphi(S_N)] = \sum_{n=1}^{\infty} \mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}].$$

Correction Par Fubini-Tonelli, on a

$$\mathbb{E}[\varphi(S_N)] = \mathbb{E}\left[\sum_{n=1}^{\infty} \varphi(S_n) \mathbf{1}_{N=n}\right] = \mathbb{E}\left[\sum_{n=1}^{\infty} \varphi(S_n) \mathbf{1}_{N=n}\right] = \sum_{n=1}^{\infty} \mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}].$$

- (b) Notant $I = [a, b]$ et I^c son complémentaire, montrer que

$$\mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}] = \int_{\mathbb{R}^n} \varphi(s_n) \mathbf{1}_{I^{n-1}}(s_1, \dots, s_{n-1}) \mathbf{1}_{I^c}(s_n) f(s_1, \dots, s_n) ds_1 \dots ds_n,$$

où, avec la convention $s_0 = 0$,

$$f(s_1, \dots, s_n) = \exp(-2ms_n) \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{((s_i - s_{i-1}) - m)^2}{2}\right\}.$$

Correction On a

$$\mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}] = \mathbb{E}[\varphi(S_n) \mathbf{1}_{I^{n-1}}(S_1, \dots, S_{n-1}) \mathbf{1}_{I^c}(S_n)].$$

Si $f(s_1, \dots, s_n)$ désigne la densité du vecteur aléatoire gaussien $[S_1, \dots, S_n]^t$, ceci donne

$$\mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}] = \int_{\mathbb{R}^n} \varphi(s_n) \mathbf{1}_{I^{n-1}}(s_1, \dots, s_{n-1}) \mathbf{1}_{I^c}(s_n) f(s_1, \dots, s_n) ds_1 \dots ds_n.$$

Puisque $[X_1, \dots, X_n]^t = A[S_1, \dots, S_n]^t$, avec les X_i i.i.d. de loi $\mathcal{N}(-m, 1)$ et A la matrice dont la diagonale vaut 1 et la sous-diagonale vaut -1, la formule de changement de variable donne

$$f(s_1, \dots, s_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{((s_i - s_{i-1}) + m)^2}{2}\right\},$$

c'est-à-dire

$$f(s_1, \dots, s_n) = \exp(-2ms_n) \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{((s_i - s_{i-1}) - m)^2}{2}\right\}.$$

- (c) En déduire que

$$\mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}] = \mathbb{E}[\varphi(S'_n) \exp(-2mS'_n) \mathbf{1}_{N'=n}],$$

et conclure.

Correction Le terme produit dans l'expression précédente est la densité $g(s_1, \dots, s_n)$ du vecteur $[S'_1, \dots, S'_n]^t$, si bien qu'en rembobinant le fil précédent, on aboutit à

$$\mathbb{E}[\varphi(S_n) \mathbf{1}_{N=n}] = \int_{\mathbb{R}^n} \varphi(s_n) \mathbf{1}_{I^{n-1}}(s_1, \dots, s_{n-1}) \mathbf{1}_{I^c}(s_n) \exp(-2ms_n) g(s_1, \dots, s_n) ds_1 \dots ds_n,$$

donc

$$\mathbb{E}[\varphi(S_n)\mathbf{1}_{N=n}] = \mathbb{E}[\varphi(S'_n)\exp(-2mS'_n)\mathbf{1}_{N'=n}].$$

En remarquant que, pour les mêmes raisons que pour (S_n) , le temps d'arrêt N' pour la marche (S'_n) est presque sûrement fini, il ne reste plus qu'à sommer sur n pour obtenir

$$\mathbb{E}[\varphi(S_N)] = \mathbb{E}[\varphi(S'_{N'})\exp(-2mS'_{N'})],$$

ce qui prouve que le rapport de vraisemblance entre les deux densités vaut $f(s)/g(s) = \exp(-2ms)$.

Remarque : La densité $f(s)$ de la variable aléatoire S_N n'est pas triviale, les calculs précédents montrent en effet qu'elle s'écrit

$$f(s) = \sum_{n=1}^{\infty} \int_{\mathbb{R}^{n-1}} \mathbf{1}_{I^{n-1}}(s_1, \dots, s_{n-1}) \mathbf{1}_{I^c}(s) f(s_1, \dots, s_{n-1}, s) ds_1 \dots ds_{n-1}.$$

En revanche, le rapport des densités f et g est, lui, élémentaire.

4. Dédurre de la question précédente un nouvel estimateur $\tilde{p}_k$ de p basé sur k simulations successives.

Correction Notons $S'(j) = (S'_n(j))_{1 \leq n \leq N'(j)}$ la trajectoire ainsi obtenue à la j -ème simulation et

$$I'(j) = \begin{cases} 0 & \text{si } s'_{N'(j)} < a \\ 1 & \text{si } s'_{N'(j)} > b \end{cases}$$

Pour un total de k simulations, on en déduit l'estimateur par échantillonnage préférentiel

$$\tilde{p}_k = \frac{1}{k} \sum_{j=1}^k I'(j) \exp(-2ms'_{N'(j)}).$$

Remarque : par symétrie, il est clair que cette fois presque toutes trajectoires sortiront par le haut, c'est-à-dire que $I'(j) = 1$ avec très grande probabilité.

5. Dédurre de la forme de $\tilde{p}_k$ que $p \leq \exp(-2mb)$.

Correction Puisque $I'(j) = 1$ lorsque $s'_{N'(j)} > b$, il vient que pour tout k ,

$$\tilde{p}_k \leq \frac{1}{k} \sum_{j=1}^k e^{-2mb} = e^{-2mb}$$

donc

$$p = \mathbb{E}[\tilde{p}_k] \leq e^{-2mb}.$$

Par exemple, pour $m = 1$ et $b = 5$, ceci donne $p \leq 4.5 \times 10^{-5}$.