Méthodes de Monte-Carlo et algorithmes stochastiques

Séance 4 – Réduction de variance II : stratification et importance sampling

M2 Finance – Université Gustave Eiffel

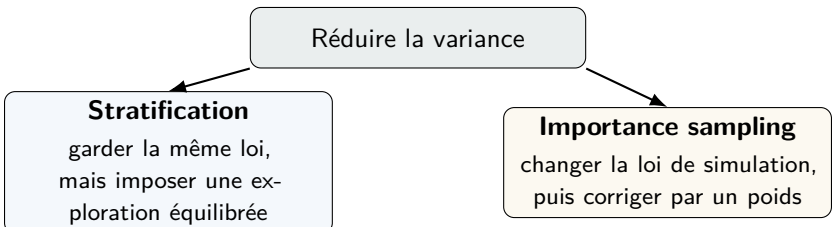
But de la séance : terminer le chapitre 2 en contrôlant où les simulations se placent, puis en changeant directement la loi de simulation lorsque la loi naturelle visite mal la région utile.

Panorama du cours

Chapitre 1	Moyennes de Monte-Carlo, convergence, intervalles de confiance, choix de n, méthode delta, simulation par inversion et acceptation–rejet.
Chapitre 2	Réduction de variance : antithétiques, variables de contrôle, conditionnement, stratification, importance sampling, événements rares.
Chapitre 3	Espérances conditionnelles, projection L^2, régression, programmation dynamique, Longstaff–Schwartz.
Chapitre 4	Quasi-Monte-Carlo, suites à faible discrépance, Robbins–Monro, Kiefer–Wolfowitz.

Séance 4. On termine le chapitre 2 : stratification, allocation optimale, importance sampling, événements rares, erreur relative et diagnostics de poids.

Fil directeur : deux manières de mieux utiliser les simulations



Idée à retenir

Aujourd'hui, le problème n'est plus seulement de transformer une variable donnée : on organise ou on déplace les tirages pour qu'ils visitent mieux les régions importantes.

Pourquoi stratifier ?

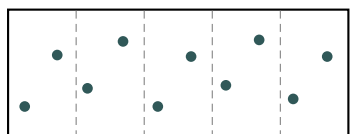
Dans une simulation naïve, certaines régions importantes peuvent être mal représentées par hasard.

tirages naïfs



certaines strates sont vides, d'autres surreprésentées

tirages stratifiés



chaque région imposée est visitée

Message

La stratification retire une source de hasard : le nombre de points qui tombent dans chaque région.

Cadre mathématique de la stratification

On veut estimer

$$I = \mathbb{E}(h(X)).$$

On découpe l'espace en strates disjointes

$$A_1, \dots, A_K, \qquad \bigcup_{k=1}^{K} A_k = E.$$

On note

$$p_k = \mathbb{P}(X \in A_k), \qquad I_k = \mathbb{E}(h(X) \mid X \in A_k).$$

Alors, par la formule des probabilités totales,

$$I = \sum_{k=1}^{K} p_k I_k.$$

Idée à retenir

La stratification remplace une moyenne globale par une somme pondérée de moyennes conditionnelles.

Estimateur stratifié

Dans chaque strate A_k, on simule

$$X_{k,1}, \dots, X_{k,n_k} \quad \text{i.i.d. de loi} \quad \mathcal{L}(X \mid X \in A_k).$$

L'estimateur stratifié est

$$\widehat{I}^{\text{strat}} = \sum_{k=1}^{K} p_k \left(\frac{1}{n_k} \sum_{i=1}^{n_k} h(X_{k,i}) \right).$$

Lecture pratique

- p_k est le poids de la strate dans la loi naturelle ;
- la moyenne interne estime I_k ;
- les n_k sont choisis par l'utilisateur.

Biais et variance de l'estimateur stratifié

On pose

$$\sigma_k^2 = \mathrm{Var}(h(X) \mid X \in A_k).$$

Proposition

L'estimateur stratifié est sans biais et

$$\mathrm{Var}(\widehat{I}^{\mathrm{strat}}) = \sum_{k=1}^{K} \frac{p_k^2 \sigma_k^2}{n_k}.$$

Conséquence

La variance dépend non seulement du découpage, mais aussi de la répartition des simulations entre les strates.

Preuve : biais et variance

Dans chaque strate,

$$\mathbb{E}\left(\frac{1}{n_k}\sum_{i=1}^{n_k} h(X_{k,i})\right) = I_k.$$

Donc

$$\mathbb{E}(\widehat{I}^{\text{strat}}) = \sum_{k=1}^{K} p_k I_k = I.$$

Pour la variance, les simulations entre strates sont indépendantes :

$$\begin{aligned}\operatorname{Var}(\widehat{I}^{\text{strat}}) &= \sum_{k=1}^{K} p_k^2 \operatorname{Var}\left(\frac{1}{n_k}\sum_{i=1}^{n_k} h(X_{k,i})\right) \\ &= \sum_{k=1}^{K} p_k^2 \frac{\sigma_k^2}{n_k}.\end{aligned}$$

Allocation proportionnelle

Le choix le plus simple est

$$n_k \simeq np_k.$$

On parle d'allocation proportionnelle. Dans ce cas,

$$\mathrm{Var}(\widehat{I}^{\mathrm{strat}}) \simeq \frac{1}{n}\sum_{k=1}^{K} p_k \sigma_k^2.$$

Comparaison avec le naïf

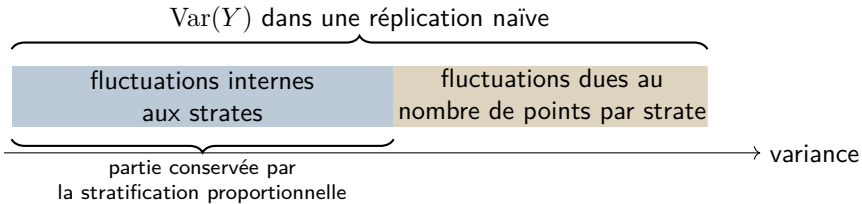
Si $Y = h(X)$, la formule de variance totale donne

$$\mathrm{Var}(Y) = \sum_{k=1}^{K} p_k \sigma_k^2 + \sum_{k=1}^{K} p_k (I_k - I)^2.$$

Donc

$$\mathrm{Var}(\widehat{I}^{\mathrm{strat}}) \leq \frac{\mathrm{Var}(Y)}{n}.$$

Ce que retire la stratification proportionnelle



Idée à retenir

L'allocation proportionnelle enlève la variance due au hasard du nombre de tirages par strate. Le gain est important lorsque les moyennes I_k varient fortement d'une strate à l'autre.

Exemple exact : $\int_0^1 x \, \mathrm{d}x$

Naïvement, $U \sim \mathcal{U}([0,1])$ et

$$\mathrm{Var}(U) = \frac{1}{12}, \qquad \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} U_i\right) = \frac{1}{12n}.$$

Découpons $[0,1]$ en K intervalles de longueur $1/K$, avec n/K tirages par strate. Dans chaque strate,

$$\sigma_k^2 = \frac{1}{12K^2}, \qquad p_k = \frac{1}{K}, \qquad n_k = \frac{n}{K}.$$

Donc

$$\mathrm{Var}(\widehat{I}^{\mathrm{strat}}) = K\frac{(1/K)^2(1/(12K^2))}{n/K} = \frac{1}{12nK^2}.$$

Lecture

La variance est divisée par K^2. Avec $K = 10$, l'erreur standard est divisée par 10.

Allocation optimale de Neyman

Si les variances conditionnelles σ_k^2 sont très différentes, on ne doit pas forcément tirer proportionnellement aux masses p_k.

Proposition

On minimise

$$\mathrm{Var}(\widehat{I}^{\mathrm{strat}}) = \sum_{k=1}^{K} \frac{p_k^2 \sigma_k^2}{n_k}$$

sous la contrainte $n_1 + \cdots + n_K = n$. En négligeant les contraintes d'intégralité, l'allocation optimale est

$$n_k = n \frac{p_k \sigma_k}{\sum_{j=1}^{K} p_j \sigma_j}.$$

La variance minimale vaut

$$\frac{1}{n} \left(\sum_{k=1}^{K} p_k \sigma_k \right)^2.$$

Allocation optimale : preuve par Lagrange

On minimise

$$F(n_1, \ldots, n_K) = \sum_{k=1}^{K} \frac{p_k^2 \sigma_k^2}{n_k}$$

sous la contrainte $\sum_k n_k = n$. Le lagrangien est

$$\mathcal{L}(n_1, \ldots, n_K, \lambda) = \sum_{k=1}^{K} \frac{p_k^2 \sigma_k^2}{n_k} + \lambda \left(\sum_{k=1}^{K} n_k - n \right).$$

La condition d'optimalité donne

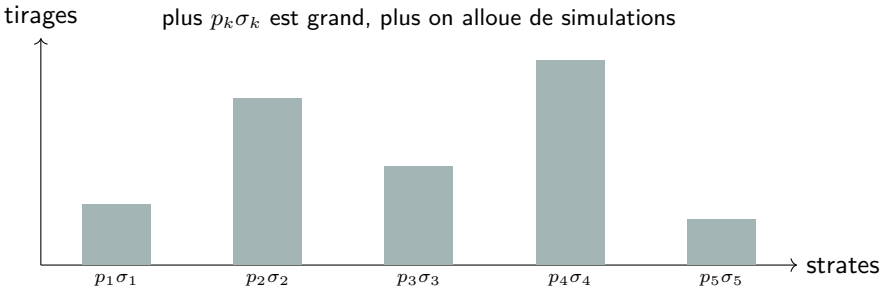
$$-\frac{p_k^2 \sigma_k^2}{n_k^2} + \lambda = 0,$$

donc

$$n_k \propto p_k \sigma_k.$$

La normalisation $\sum_k n_k = n$ donne la formule.

Comment lire l'allocation optimale ?



Point de vigilance

La formule optimale demande les σ_k, souvent inconnus. En pratique, on les estime avec une simulation pilote, puis on répartit le budget principal.

Procédure pilote en pratique

Quand les σ_k sont inconnus, une procédure standard est :

1. faire un petit échantillon pilote dans chaque strate ;
2. calculer des estimations $\widehat{\sigma}_k$;
3. choisir les allocations
$$n_k \propto p_k \widehat{\sigma}_k ;$$
4. lancer la simulation principale avec ces n_k.

Point pratique

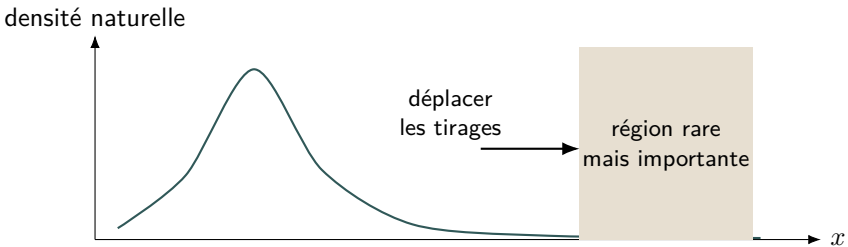
Il faut garder une partie suffisante du budget pour la simulation principale. Le pilote sert à apprendre où l'aléa est fort, pas à épuiser le budget.

Transition : quand stratifier ne suffit plus

La stratification garde la loi naturelle de simulation, mais organise mieux les tirages.

Limite

Si la région qui contribue le plus à l'espérance est très rare sous la loi naturelle, même une stratification simple peut rester insuffisante.



Idée à retenir

L'importance sampling change la loi de simulation pour visiter volontairement la région utile.

Importance sampling : changer la loi de simulation

On veut calculer

$$I = \mathbb{E}_p(h(X)) = \int h(x)p(x)\ \mathrm{d}x,$$

où p est la densité naturelle. On choisit une densité q telle que

$$q(x) > 0 \quad \text{dès que} \quad h(x)p(x) \neq 0.$$

Alors

$$I = \int h(x)\frac{p(x)}{q(x)}q(x)\ \mathrm{d}x = \mathbb{E}_q\left(h(X)\frac{p(X)}{q(X)}\right).$$

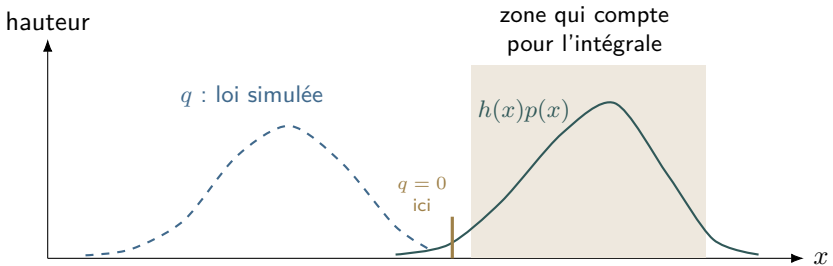
Poids d'importance

$$L(X) = \frac{p(X)}{q(X)}.$$

Le poids corrige le changement de loi : un point tiré sous q est repondéré comme il aurait pesé sous p.

Condition de support : le piège à éviter

L'importance sampling remplace la densité naturelle p par une densité de simulation q, puis corrige par le poids p/q.



Point de vigilance

Si $q(x) = 0$ dans une zone où $h(x)p(x) \neq 0$, on ne peut jamais simuler cette zone. Le poids p/q ne peut pas corriger une région qui n'est jamais visitée.

Estimateur par importance sampling

Soient $X_1, \dots, X_n$ i.i.d. de densité q. On pose

$$\widehat{I}_n^{\mathrm{IS}} = \frac{1}{n} \sum_{i=1}^{n} h(X_i) \frac{p(X_i)}{q(X_i)}.$$

Proposition

Si

$$\mathbb{E}_q \left(h(X)^2 \frac{p(X)^2}{q(X)^2} \right) < +\infty,$$

alors $\widehat{I}_n^{\mathrm{IS}}$ est sans biais et

$$\mathrm{Var}(\widehat{I}_n^{\mathrm{IS}}) = \frac{1}{n} \left(\int h(x)^2 \frac{p(x)^2}{q(x)} \, \mathrm{d}x - I^2 \right).$$

Importance sampling : preuve

L'absence de biais vient du changement de densité :

$$\mathbb{E}_q\left(h(X)\frac{p(X)}{q(X)}\right) = \int h(x)\frac{p(x)}{q(x)}q(x)\,\mathrm{d}x = I.$$

Pour la variance, une réplication a pour second moment

$$\mathbb{E}_q\left(h(X)^2\frac{p(X)^2}{q(X)^2}\right) = \int h(x)^2\frac{p(x)^2}{q(x)}\,\mathrm{d}x.$$

Comme l'estimateur est une moyenne de réplications indépendantes,

$$\mathrm{Var}(\widehat{I}_n^{\mathrm{IS}}) = \frac{1}{n}\left(\int h(x)^2\frac{p(x)^2}{q(x)}\,\mathrm{d}x - I^2\right).$$

Densité optimale théorique

Pour diminuer la variance, on minimise le second moment d'une réplication IS :

$$\mathbb{E}_q\left[\left(h(X)\frac{p(X)}{q(X)}\right)^2\right] = \int h(x)^2\frac{p(x)^2}{q(x)}\,\mathrm{d}x.$$

Proposition

Une densité optimale est

$$q^\star(x) = \frac{|h(x)|p(x)}{\int |h(u)|p(u)\,\mathrm{d}u}.$$

Si $h \geq 0$, alors $q^\star(x) = h(x)p(x)/I$ et l'estimateur a variance nulle.

Point de vigilance

Cette densité dépend généralement de I, donc elle sert surtout de guide pour choisir une bonne loi q.

Densité optimale : preuve par Cauchy–Schwarz

Pour toute densité q admissible,

$$\left(\int |h(x)|p(x)\,\mathrm{d}x\right)^2 = \left(\int \frac{|h(x)|p(x)}{\sqrt{q(x)}}\sqrt{q(x)}\,\mathrm{d}x\right)^2$$
$$\leq \left(\int h(x)^2\frac{p(x)^2}{q(x)}\,\mathrm{d}x\right)\left(\int q(x)\,\mathrm{d}x\right).$$

Comme q est une densité,

$$\int q(x)\,\mathrm{d}x = 1.$$

L'égalité dans Cauchy–Schwarz impose

$$q(x) \propto |h(x)|p(x).$$

Idée à retenir

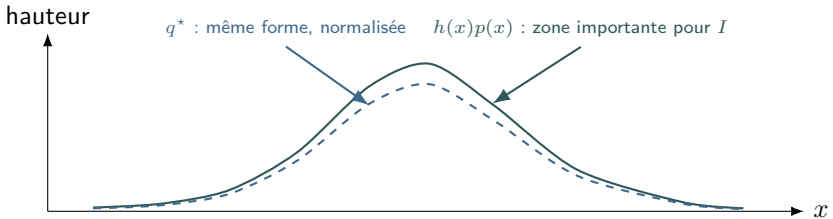
La bonne loi d'importance doit être grande là où la contribution $|h|p$ est grande.

Interprétation : où faut-il tirer ?

On écrit

$$I = \int h(x)p(x)\,\mathrm{d}x.$$

Si $h \geq 0$, les zones où $h(x)p(x)$ est grand sont celles qui apportent le plus à I. La densité optimale $q^\star$ a la même forme, simplement normalisée pour avoir une aire totale égale à 1.



Message

Une bonne loi d'importance tire souvent là où $h(x)p(x)$ est grand, puis corrige le changement de loi par le poids p/q.

Exemple gaussien : le problème du naïf

Soit

$$X \sim \mathcal{N}(0,1), \qquad I_\beta = \mathbb{E}\left(\mathrm{e}^{\beta X}\mathbf{1}_{\{X>0\}}\right).$$

On a

$$I_\beta = \mathrm{e}^{\beta^2/2}\Phi(\beta).$$

L'estimateur naïf moyenne

$$Y = \mathrm{e}^{\beta X}\mathbf{1}_{\{X>0\}}.$$

Son second moment est

$$\mathbb{E}(Y^2) = \mathbb{E}\left(\mathrm{e}^{2\beta X}\mathbf{1}_{\{X>0\}}\right) = \mathrm{e}^{2\beta^2}\Phi(2\beta).$$

Point de vigilance

Pour β grand, l'espérance est dominée par de très grandes valeurs dans la queue droite de la normale, que la simulation naïve visite rarement.

Exemple gaussien : où est la région importante ?

L'intégrande non normalisé est

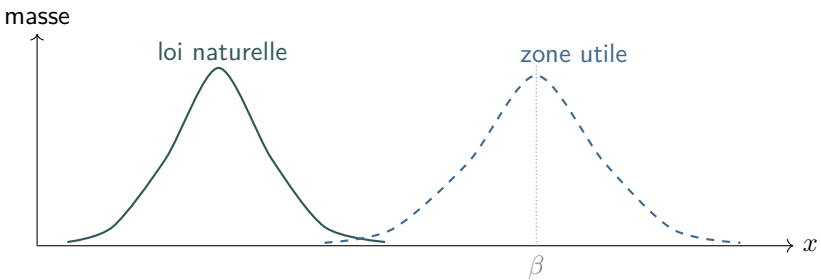
$$\mathrm{e}^{\beta x}\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\beta x - \frac{x^2}{2}\right).$$

La fonction

$$x \mapsto \beta x - \frac{x^2}{2}$$

est maximale en

$$x = \beta.$$



Translation gaussienne

On simule maintenant sous

$$q_\mu = \mathcal{N}(\mu, 1).$$

Si $p = \varphi$ est la densité de $\mathcal{N}(0,1)$, alors

$$\frac{p(x)}{q_\mu(x)} = \exp\left(-\mu x + \frac{\mu^2}{2}\right).$$

L'estimateur d'importance devient

$$\mathrm{e}^{\beta X} \mathbf{1}_{\{X>0\}} \exp\left(-\mu X + \frac{\mu^2}{2}\right), \qquad X \sim \mathcal{N}(\mu, 1).$$

Idée à retenir

Le choix $\mu = \beta$ recentre les simulations près du maximum de la contribution dominante.

Cas $\mu = \beta$: simplification et variance

Pour $\mu = \beta$, l'estimateur devient

$$Z_\beta = \mathrm{e}^{\beta^2/2}\mathbf{1}_{\{X>0\}}, \qquad X \sim \mathcal{N}(\beta, 1).$$

Il est bien sans biais :

$$\mathbb{E}(Z_\beta) = \mathrm{e}^{\beta^2/2}\mathbb{P}_{\mathcal{N}(\beta,1)}(X > 0) = \mathrm{e}^{\beta^2/2}\Phi(\beta) = I_\beta.$$

Sa variance est

$$\mathrm{Var}(Z_\beta) = \mathrm{e}^{\beta^2}\Phi(\beta)(1 - \Phi(\beta)).$$

Exemple numérique

Pour $\beta = 5$, le rapport des variances entre naïf et importance sampling est d'environ 2.5×10^{17}.

Événements rares : l'erreur absolue ne suffit pas

Pour un événement rare, on estime souvent

$$p = \mathbb{P}(A) = \mathbb{E}(\mathbf{1}_A), \qquad p \ll 1.$$

L'estimateur naïf a une réplication de variance

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

La variance relative d'une réplication vaut

$$\frac{\mathrm{Var}(\mathbf{1}_A)}{p^2} = \frac{1-p}{p} \simeq \frac{1}{p}.$$

Point de vigilance

Lorsque p devient très petit, le nombre de simulations nécessaire pour une précision relative donnée explose. C'est un défaut structurel du naïf.

Variance relative et erreur relative

Soit $\widehat{p}_n$ un estimateur sans biais de $p > 0$.

Définition

La variance relative est

$$\frac{\mathrm{Var}(\widehat{p}_n)}{p^2}.$$

Pour une moyenne de n réplications indépendantes Y_i,

$$\frac{\mathrm{Var}(\widehat{p}_n)}{p^2} = \frac{1}{n}\frac{\mathrm{Var}(Y)}{p^2}.$$

Lien avec l'erreur relative

Pour un estimateur sans biais,

$$\mathbb{E}\left[\left(\frac{\widehat{p}_n - p}{p}\right)^2\right] = \frac{\mathrm{Var}(\widehat{p}_n)}{p^2}.$$

Exemple rare : intervalle exponentiel

Soit $X \sim \mathcal{E}(1)$ et

$$p_\ell = \mathbb{P}(X \in [\ell, \ell+1]) = \mathrm{e}^{-\ell}(1 - \mathrm{e}^{-1}).$$

L'estimateur naïf est une Bernoulli :

$$Y = \mathbf{1}_{\{X \in [\ell, \ell+1]\}}, \qquad \mathrm{Var}(Y) = p_\ell(1 - p_\ell).$$

Une méthode d'importance consiste à tirer directement

$$U \sim \mathcal{U}([\ell, \ell+1])$$

et à écrire

$$p_\ell = \int_\ell^{\ell+1} \mathrm{e}^{-x}\, \mathrm{d}x = \mathbb{E}(\mathrm{e}^{-U}).$$

Idée à retenir

Toutes les simulations sont maintenant dans la région utile : l'estimateur ne produit jamais de zéro.

Exemple rare : comparaison des ordres

Pour la méthode uniforme sur $[\ell, \ell+1]$,

$$\mathrm{Var}(\mathrm{e}^{-U}) = \mathrm{e}^{-2\ell}\left(\frac{1-\mathrm{e}^{-2}}{2} - (1-\mathrm{e}^{-1})^2\right).$$

Donc

$$\mathrm{Var}(\mathrm{e}^{-U}) = O(\mathrm{e}^{-2\ell}), \qquad p_\ell^2 = O(\mathrm{e}^{-2\ell}).$$

En revanche, pour le naïf,

$$\mathrm{Var}(\mathbf{1}_{\{X\in[\ell,\ell+1]\}}) = O(\mathrm{e}^{-\ell}).$$

Conclusion

La variance relative de la méthode d'importance reste bornée, alors que celle du naïf diverge comme e^{ℓ}.

Erreur relative bornée

Considérons des probabilités rares $p_\ell \to 0$ et des estimateurs sans biais Y_ℓ.

Définition

La méthode a une erreur relative bornée si

$$\sup_\ell \frac{\mathrm{Var}(Y_\ell)}{p_\ell^2} < +\infty.$$

Interprétation

Le nombre de simulations nécessaire pour obtenir une précision relative donnée ne diverge pas quand l'événement devient rare.

Point de vigilance

Pour les événements rares, il faut contrôler une erreur relative : une petite erreur absolue peut être énorme par rapport à p_ℓ.

Autonormalisation des poids

Si $W(x) = p(x)/q(x)$ et si p est normalisée, alors

$$I = \mathbb{E}_p(h) = \mathbb{E}_q(Wh), \qquad \mathbb{E}_q(W) = 1.$$

Quand la densité cible n'est connue qu'à une constante multiplicative près, on utilise souvent

$$\widehat{I}_n^{\mathrm{SN}} = \frac{\sum_{i=1}^n W_i h(X_i)}{\sum_{i=1}^n W_i}.$$

Asymptotique

Si

$$\mathbb{E}_q\left[W^2(h-I)^2\right] < +\infty,$$

alors

$$\sqrt{n}(\widehat{I}_n^{\mathrm{SN}} - I) \Rightarrow \mathcal{N}\left(0, \mathrm{Var}_q(W(h-I))\right).$$

Autonormalisation : preuve de l'idée

On écrit exactement

$$\widehat{I}_n^{\mathrm{SN}} - I = \frac{\frac{1}{n}\sum_{i=1}^n W_i(h(X_i) - I)}{\frac{1}{n}\sum_{i=1}^n W_i}.$$

Par la loi forte,

$$\frac{1}{n}\sum_{i=1}^n W_i \to \mathbb{E}_q(W) = 1.$$

Le numérateur est une moyenne centrée car

$$\mathbb{E}_q[W(h - I)] = \mathbb{E}_q(Wh) - I\mathbb{E}_q(W) = 0.$$

Le TCL appliqué au numérateur, puis Slutsky, donne le résultat.

Point de vigilance

L'autonormalisation n'est pas une réduction de variance automatique : des poids très déséquilibrés peuvent rendre le rapport instable.

Dégénérescence des poids et ESS

On normalise les poids par

$$\widetilde{w}_i = \frac{w_i}{\sum_{j=1}^{n} w_j}.$$

Le diagnostic de taille effective d'échantillon est

$$\mathrm{ESS} = \frac{1}{\sum_{i=1}^{n} \widetilde{w}_i^2} = \frac{\left(\sum_{i=1}^{n} w_i\right)^2}{\sum_{i=1}^{n} w_i^2}.$$

On a toujours

$$1 \leq \mathrm{ESS} \leq n.$$

Lecture

Si tous les poids sont égaux, $\mathrm{ESS} = n$. Si un seul poids porte presque toute la masse, ESS est proche de 1.

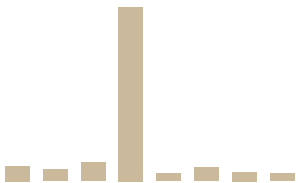
ESS : visualisation

poids équilibrés



ESS proche de n

poids dégénérés



ESS proche de 1

Point de vigilance

L'ESS est un diagnostic, pas un théorème de précision. Une méthode peut avoir des poids dispersés mais rester bonne pour une fonction h donnée, ou inversement.

Limite asymptotique de l'ESS

Si les poids W_i sont i.i.d. positifs, avec $0 < \mathbb{E}(W) < \infty$ et $\mathbb{E}(W^2) < \infty$, alors

$$\frac{\mathrm{ESS}}{n} \to \frac{\mathbb{E}(W)^2}{\mathbb{E}(W^2)} = \frac{1}{1 + \mathrm{CV}(W)^2} \qquad \text{p.s.}$$

Preuve

On écrit

$$\frac{\mathrm{ESS}}{n} = \frac{\left(\frac{1}{n}\sum_{i=1}^n W_i\right)^2}{\frac{1}{n}\sum_{i=1}^n W_i^2}.$$

La loi forte donne la convergence du numérateur vers $\mathbb{E}(W)^2$ et du dénominateur vers $\mathbb{E}(W^2)$.

Synthèse : choisir une méthode de réduction de variance

Méthode	Idée	Situation favorable
Antithétique	coupler deux tirages opposés	symétrie naturelle, covariance négative
Contrôle	soustraire une fluctuation connue	variable auxiliaire très corrélée
Conditionnement	remplacer par une espérance conditionnelle	bruit interne intégrable analytiquement
Stratification	équilibrer les régions de l'espace	moyennes conditionnelles très différentes
Importance sampling	changer la loi et corriger par un poids	événement rare, queue, intégrande concentré

Idée à retenir

Une bonne méthode répond à la cause de la variance : symétrie, corrélation, bruit interne, hétérogénéité spatiale ou rareté.

Procédure numérique raisonnable

Une bonne procédure de réduction de variance suit souvent cette logique :

1. commencer par un estimateur naïf et mesurer sa variance empirique ;
2. identifier la source principale des fluctuations ;
3. choisir une méthode adaptée à cette source ;
4. comparer les demi-largeurs d'IC à temps de calcul comparable ;
5. vérifier la stabilité lorsque les paramètres changent.

Diagnostic minimal

Ne jamais conclure uniquement avec une moyenne. Il faut regarder la variance, l'IC, le coût, et pour l'importance sampling, la stabilité des poids.

Messages à retenir

1. La stratification impose une exploration équilibrée et supprime une partie de la variance inter-strates.
2. L'allocation optimale met plus de tirages dans les strates de grand $p_k\sigma_k$.
3. L'importance sampling change la loi de simulation et corrige par le poids p/q.
4. Une bonne densité d'importance suit la contribution $|h|p$, pas nécessairement la densité naturelle p.
5. Pour les événements rares, la variance relative est souvent le critère pertinent.
6. Les diagnostics de poids, notamment l'ESS, sont indispensables mais ne remplacent pas une comparaison d'IC à budget fixé.