

# Échantillonnage préférentiel adaptatif multiple

**Pierre PUDLO**, Aix-Marseille Université, I2M

**Jean-Michel MARIN**, Université de Montpellier, IMAG

**Mohammed SEDKI**, Université de Paris-Sud, Inserm

**Mots-clés** : Monte Carlo, algorithmes adaptatifs, échantillonnage préférentiel, génétique des populations

On peut voir les méthodes de Monte Carlo comme une méthode d'approximation d'une distribution, dite loi cible ou cible, par une loi discrète en échantillonnant aléatoirement l'espace à parcourir. L'échantillonnage préférentiel est une méthode qui utilise une distribution auxiliaire pour faire des propositions qui sont ensuite pondérées par le rapport des densités de la loi cible et de la loi de proposition pour corriger l'écart entre ces lois et entre sans biais. Pour être efficace, l'échantillonnage préférentiel requière un ajustement assez fin de la distribution de proposition, ce d'autant plus que la dimension de l'espace à parcourir est grande. La calibration de la loi de proposition requière des informations sur la distribution cible dont on ne dispose pas en règle générale. La loi de proposition est donc réglée avec des algorithmes itératifs, qui permettent d'apprendre l'information manquante pour la calibration en améliorant progressivement une approximation courante de la loi cible.

Parmi les propositions de la littérature, nous nous intéressons à l'algorithme d'échantillonnage adaptatif préférentiel multiple, *adaptive multiple importance sampling*, ou AMIS [1]. Cette méthode a produit une amélioration significative en stabilité en proposant une technique de recyclage efficace des différentes approximations construites au cours des itérations. Dans cet exposé, nous donnerons une idée des résultats asymptotiques que nous avons obtenu [3]. Contrairement à Douc et al. [2], nos résultats sont démontrés dans un régime asymptotique où le nombre d'itérations tend vers l'infini, mais où le nombre de tirages par itérations est une suite de nombres entiers croissante, mais fixée. Nous montrerons également le comportement numérique de l'algorithme sur des exemples de statistique bayésienne où la loi cible est la loi *a posteriori*. Nous pensons en particulier aux cas issus de la génétique des populations, où la densité de la loi cible est difficile à calculer, et où, donc, la stratégie de recyclage fait grandement gagner en efficacité.

## Références

- [1] J-M. Cornuet, J-M. Marin, Antonietta Mira, and C. P. Robert. Adaptive Multiple Importance Sampling. *Scandinavian Journal of Statistics*, 39:798–812, 2012.
- [2] R. Douc, A. Guillin, J. M. Marin, and C. P. Robert. Convergences of adaptive mixtures of importance sampling schemes. *The Annals of Statistics*, 35(1):420–448, 2007.
- [3] J. M. Marin, P. Pudlo and M. Sedki. Consistency of Adaptive Importance Sampling and Recycling Schemes. To appear in *Bernoulli*.

**Pierre PUDLO**, Aix Marseille Université, CNRS, Centrale Marseille, I2M UMR 7373, 13453, Marseille CEDEX 13.

`pierre.pudlo@univ-amu.fr`

**Jean-Michel MARIN**, Université de Montpellier, Institut Montpellierain Alexander Grothendieck IMAG, Place E. Bataillon, 34095 Montpellier CEDEX

`jean-michel.marin@umontpellier.fr`

**Mohammed SEDKI**, Université Paris-Sud, U1181, Hôpital Paul Brousse, Bât. Inserm 15/16, 16 Avenue Paul Vaillant-Couturier 94807 Villejuif CEDEX

`mohammed.sedki@u-psud.fr`