M1 – Statistique – 2020/2021
vincent.brault@univ-grenoble-alpes.fr

TP n°1 : Méthode de rejet et estimation

Dans ce TP, nous allons étudier une méthode pour simuler des lois ***non-classiques*** *appelée la méthode de rejet puis estimer les paramètres grâce aux estimations vues en cours.*

Nouvelle loi

Dans ce TP, nous allons étudier la loi de paramètre $\theta^\star > 0$ de densité définie pour tout x dans $\mathbb{R}$ par :

$$f_{\theta^\star}(x) = C_{\theta^\star} \sin(\theta^\star x) \mathbb{1}_{[0;\pi/\theta^\star]}(x)$$

où $C_{\theta^\star}$ est la constante de normalisation.

1. Expliquer pourquoi l'intervalle de définition ne peut pas aller avant 0 ou se prolonger après $\pi/\theta^\star$.

2. Calculer la constante de normalisation.

3. Représenter les densités pour $\theta^\star \in \{0.5, 1, 2\}$.

4. Calculer la valeur maximale $M_{\theta^\star}$ que peut valoir la densité.

Méthode de rejet

La loi n'étant pas classique, il n'existe pas de générateur pré-programmé. Nous allons contourner ce problème en utilisant la méthode de rejet. Pour simuler une variable aléatoire X suivant la loi de densité $f_{\theta^\star}$, nous prenons une constante $\lambda \geq M_{\theta^\star}$ et nous appliquons la méthode suivante :

1. Tirer une variable aléatoire $U_x \sim \mathcal{U}([0, \pi/\theta^\star])$.
2. Tirer une variable aléatoire $U_y \sim \mathcal{U}([0, \lambda])$.
3. Si $U_y > f_{\theta^\star}(U_x)$, recommencer à 1.
4. Renvoyer $X = U_x$.

5. Calculer la probabilité que U_y soit strictement supérieure à $f_{\theta^\star}(U_x)$. Montrer que l'algorithme s'arrête presque sûrement.

6. Déterminer la loi de X. *Remarque : Penser que, pour sortir de la boucle, nous ne devons pas vérifier la condition de la troisième étape de l'algorithme.*

7. Donner la loi de la variable aléatoire N représentant le nombre d'itérations. Quelle(s) valeur(s) de λ préconisez-vous pour optimiser la procédure ?

8. Implémenter la méthode pour simuler une variable.

9. Pour $n \in \{10, 100, 1000\}$, simuler un n-échantillon suivant la loi de densité $f_{\theta^\star}$ pour $\theta^\star \in \{0.5, 1, 2\}$. Représenter les n-échantillons sous forme d'histogrammes et comparer avec les densités.

Estimation

Dans cette partie, nous étudions le comportement empirique des estimateurs du paramètre $\theta^\star$ vus en cours.

Méthode des moments

10. Calculer l'estimateur $\widehat{\theta}_n$ obtenu par la méthode des moments.

11. Implémenter une fonction qui prend en entrée un n-échantillon et renvoie l'estimateur $\widehat{\theta}_n$.

12. Quelle est la loi asymptotique de l'estimateur ?

13. Pour $n \in \{10, 100, 1000\}$ et $\theta^\star \in \{0.5, 1, 2\}$, simuler 100 estimateurs $\widehat{\theta}_n$ et mettre en évidence la loi asymptotique.

Méthode du maximum de vraisemblance

14. Calculer la vraisemblance du modèle.

15. Pour $n \in \{10, 100, 1000\}$ et $\theta^\star \in \{0.5, 1, 2\}$, représenter la fonction de vraisemblance associée à n-échantillon. Quel semble être l'estimateur du maximum de vraisemblance ?

16. Implémenter une fonction qui prend en entrée un n-échantillon et renvoie une estimation l'estimateur $\widetilde{\theta}_n$.

17. Comparer avec l'estimateur que vous avez proposé.

18. Pour $n \in \{10, 100, 1000\}$ et $\theta^\star \in \{0.5, 1, 2\}$, simuler 100 estimateurs $\widetilde{\theta}_n$ et représenter la distribution.

Pour aller plus loin

Une autre technique de simulation consiste à utiliser l'inverse de la fonction de répartition.

19. Montrer que si F est une fonction de répartition continue et bijective d'un intervalle de $\mathbb{R}$ vers $[0, 1]$ et U une variable uniforme sur $[0, 1]$ alors $X = F^{-1}(U)$ est une variable de densité F'.

20. En déduire une autre façon de simuler des variables suivant la loi de densité $f_{\theta^\star}$