Méthodes numériques en probabilités et statistique Contrôle continu partie 1

Le but de ce devoir est de simuler des variables aléatoires $\gamma(a,b)$ avec $a,b>0$ quelconques en implémentant l'algorithme d'Ahrens and Dieter 1982. Pour rappel, la densité de la loi $\gamma(a,b)$ est

$$\frac{b^a}{\Gamma(a)}e^{-bx}x^{a-1}\mathbb{1}_{[0,+\infty[}(x) \text{ où } \Gamma \text{ est la fonction gamma.}$$

On va montrer (exercice 1) que ce problème se ramène essentiellement à simuler suivant une loi $\gamma(a,1)$ avec $a<1$ dont on note f_a la densité. Pour simuler selon cette loi (avec $a<1$ fixé), on va appliquer une méthode de rejet (exercice 2) en utilisant comme enveloppe une loi de densité sur $\mathbb{R}$

$$g_a(x)=\frac{ae}{a+e}\left(x^{a-1}\mathbb{1}_{]0,1[}(x)+e^{-x}\mathbb{1}_{[1,+\infty[}(x)\right).$$

Exercice 1

On va montrer des propriétés de la loi γ pour justifier qu'il suffit de savoir simuler suivant une loi $\gamma(a,1)$.

1. Montrer que si $Z\sim\gamma(a,1)$ alors $Z/b\sim\gamma(a,b)$ pour tout $b>0$. Vérifier que la loi $\gamma(1,b)$ correspond à la loi $\mathcal{E}(b)$.
2. Montrer que si $Z_1\sim\gamma(a_1,1)$, $Z_2\sim\gamma(a_2,1)$ et Z_1 et Z_2 sont indépendants alors $Z_1+Z_2\sim\gamma(a_1+a_2,1)$, vous aurez besoin de la définition de la fonction beta https://fr.wikipedia.org/wiki/Fonction_bêta.
3. Puisque pour tout $a>0$, on peut écrire $a=\lfloor a\rfloor+\big(a-\lfloor a\rfloor\big)$ (où $\lfloor\rfloor$ désigne la partie entière, en déduire qu'il suffit de savoir simuler selon la loi $\gamma(a,1)$ pour $a<1$.

Exercice 2

On vérifie que la densité g_a majore à une constante près f_a et qu'on peut utiliser la méthode d'inversion pour simuler suivant g_a.

1. Montrer que, pour $c_a=(a+e)/(ae\Gamma(a))$, on a bien $f_a(x)\le c_a g_a(x)$ pour tout $x\in\mathbb{R}$.
2. On note $G_a(\cdot)$ la fonction de répartition associée à la densité $g_a(\cdot)$. Montrer que la fonction $Ginv_a$ est définie par :

$$Ginv_a(z)=\left(\frac{a+e}{e}z\right)^{1/a}\mathbb{1}_{(z<\frac{e}{a+e})}-\log\left(\frac{a+e}{ae}(1-z)\right)\mathbb{1}_{(z\ge\frac{e}{a+e})}.$$

   est l'inverse de la fonction $G_a(\cdot)$.

3. Tracer la densité f_a et la courbe de la fonction cg_a sur le même graphique, on pourra prendre `x_grid = np.linspace(0,5,100)`.

4. Coder une fonction `simu_g` qui prend n nombre de réalisations et $a \in]0,1[$ en argumant et renvoie n réalisations de v.a. i.i.d. de densité g_a grâce à la méthode d'inversion.

Exercice 3

On applique la méthode de rejet à f_a.

1. Coder une fonction `gamrej`, dépendant de n et de a, qui renvoie n réalisations de v.a. i.i.d. de loi $\gamma(a,1)$. Vérifiez graphiquement (pour $a = 0.5$ et $n = 1000$) que la loi des variables aléatoires obtenues correspond bien à celle de la densité $\gamma(a,1)$. La fonction γ est codée dans `scipy.special.gamma`.

2. Modifier légèrement la fonction `gamrej` pour obtenir de plus en sortie, un vecteur de taille n donnant pour chaque variable aléatoire tirée selon la loi $\gamma(a,1)$, le nombre d'essais qui a été nécessaire (i.e le nombre de passages dans la boucle).

3. Quelle est l'espérance du nombre d'essais nécessaires pour pour simuler une variable selon la loi $\gamma(a,1)$? Vérifier numériquement votre réponse.

Exercice 4

On propose d'utiliser l'algorithme de rejet avec une seconde enveloppe

1. Soit h_a la densité définie par :

$$h_a(x) = \frac{a}{a+1}\left(x^{a-1}\mathbb{1}_{]0,1[}(x) + \frac{1}{x^2}\mathbb{1}_{[1,+\infty[}(x)\right).$$

   Vérifiez que h_a est une densité et trouver une constante d_a telle que $f_a(x) \leq d_a h_a(x)$ pour tout $x \in \mathbb{R}$.

2. Sur un même graphique, tracer la densité f_a, la fonction $c_a g_a$ et la fonction $d_a h_a$. Pour simuler une loi $\gamma(a,1)$ par la méthode du rejet, laquelle des deux enveloppes $d_a h_a$ et $c_a g_a$ est-il préférable de choisir ?

3. Créer une fonction permettant de simuler une variable aléatoire de densité $h_a(\cdot)$ en utilisant la méthode d'inversion.

4. Définir une fonction `gamrej2`, qui utilise la fonction $h_a(\cdot)$ pour simuler un vecteur de n nombres aléatoires tirés suivant la loi $\gamma(a,1)$. Vérifiez graphiquement (pour $a = 0.5$ et $n = 1000$) que la loi des variables aléatoires obtenues correspond bien à celle de la densité $\gamma(a,1)$.

5. Comparer l'efficacité des algorithmes de rejet lorsqu'on utilise les enveloppes g_a et h_a en estimant la moyenne des nombres de tirages nécessaires pour obtenir un réalisation suivant f_a dans les deux cas.

Exercice 5

1. En utilisant la méthode d'inversion, écrire une fonction `expol` d'arguments le paramètre l et un entier n, qui renvoie n variables exponentielle de paramètre l.

2. En utilisant la fonction `expol`, définir une fonction `gamp` d'arguments n et $p \in \mathbb{N}$, qui renvoie n variable gamma de paramètre p.

3. En utilisant les fonctions `gamrej` et `gamp`, définir une fonction `gamaalea` d'arguments n, a et b, qui renvoie un vecteur de nombres aléatoires tirés suivant la loi $\gamma(a, b)$ pour tout $a > 0$ et $b > 0$.

4. Vérifier graphiquement que les variables simulées ont bien la loi souhaitée pour $a = 3.27$, $b = 1$ et $n = 1000$, il faut prendre `x_grid = np.linspace(0,15,100)`.

References

[AD82] Joachim H. Ahrens and Ulrich Dieter. "Generating gamma variates by a modified rejection technique". In: *Communications of the ACM* 25.1 (1982), pp. 47–54.