TP1 : Simulation de variables aléatoires

Li-Thiao-Té Sébastien

1 Lancers de pièces

On souhaite simuler une loi de Bernoulli $B(p), p \in [0,1]$ à l'aide de tirages d'une pièce de monnaie équilibrée (exercice 1 de la feuille de TD). Pour cela, on commence par simuler une variable aléatoire de loi uniforme d'après la formule :

$$U = \sum_{k=1}^{+\infty} \frac{B_k}{2^k}$$

où les B_k sont des lois de Bernoulli de paramètre 1/2. On obtient un tirage de $B(p) = \mathbb{1}_{U<p}$ en comparant le développement dyadique de p avec la suite des B_k. On obtient une procédure rapide en arrêtant les tirages dès que la comparaison avec p est possible. Soit i le premier indice différent, on vérifie que le simulateur est efficace en calculant $\mathbb{E}[i] = \sum_i i/2^i = 2$.

Pièce de monnaie Implémenter une fonction `piece` qui simule une variable aléatoire de loi de Bernoulli de paramètre 1/2. Dans la suite de la section, le seul générateur aléatoire autorisé est la fonction `piece`. Vérifier le code en tirant un grand échantillon (compter le nombre de 0 obtenus).

Simulation d'une variable uniforme On simule une loi uniforme sur $[0,1]$ en reconstruisant la somme $\sum_{k=1}^{+\infty} \frac{B_k}{2^k}$. Implémenter une fonction `runif(e)` qui simule la loi uniforme. On reconstruira la somme petit à petit jusqu'à la précision `e` (avec une boucle `while`). Vérifier en tracant la densité empirique.

Simulation d'une variable de Bernoulli Reprendre le code de la fonction `runif(e)` et le modifier en une fonction `rbern(p)` afin de faire la comparaison avec p à chaque itération. On s'arrête dès que la comparaison est possible, et non plus dès que la précision est atteinte. Vérifier à l'aide d'un grand échantillon.

Propriétés du simulateur

- Modifier la fonction `rbern` pour rendre le nombre de tirages de `piece` utilisés *en plus* du tirage. Vérifier que le simulateur est efficace.
- On sait que $\sum_i i/2^i = 2$. On souhaite vérifier la convergence des sommes partielles vers 2. Comment le faire en deux lignes de code ? (Utiliser la fonction `cumsum`).

2 Inversion de la fonction de répartition

2.1 Lois discrètes à support fini

Représentation d'une loi de probabilité On se donne une loi de probabilité portée par les entiers de $[0,k]$. Pour représenter cette loi, il suffit de fournir les valeurs $\mathbb{P}[X = i] = p_i$, $i \in [0,k]$. On les représente sous la forme d'un vecteur de longueur $k+1$ contenant $(p_0, p_1, \ldots, p_k)$.

Fonction de répartition discrète On représente la fonction de répartition d'une loi discrète en donnant le vecteur $(a_0, a_1, \ldots, a_k)$ tel que $a_i = \mathbb{P}[X \leq i]$. Comment calculer la fonction de répartition à partir de la loi ? (une seule ligne de code)

Tirage d'une loi discrète Il s'agit de calculer $F^-(U)$ où U est une variable aléatoire uniforme. Pour cela, on cherche le plus grand indice i tel que $F(i) < U$. Implémenter une fonction `rproba(p)` qui réalise le tirage. Vérifier en tracant un histogramme.

Bonus Comment tirer une loi de support arbitraire ?

2.2 Lois continues

Lois gaussiennes Implémenter une fonction `rgauss(m,s)` qui renvoit un tirage d'une loi gaussienne de moyenne `m` et d'écart-type `s`. Pour le calcul de la fonction réciproque, utiliser la fonction `erf`. Vérifier que la moyenne empirique et la variance empirique correspondent.

Quantile-Quantile plot Vérifier la moyenne et la variance n'est pas tout à fait suffisant. On peut vérifier de manière simple tous les quantiles en comparant les quantiles théoriques $q(x)$ et empiriques $q_n(x)$. Pour cela on affiche les couples $(q(x), q_n(x))$ pour plusieurs valeurs de x avec la fonction `qqnorm` de la stixbox. Vérifier que les points affichés se répartissent autour de la diagonale. Savez-vous décider (test statistique) si les deux lois correspondent à partir du graphique ?

Loi de Cauchy Implémenter une fonction `rcauchy` qui renvoit un tirage d'une loi de Cauchy. Dans ce cas, on ne peut pas vérifier la moyenne et la variance. Illustrer le fait que $1/n \sum_k C_k$ n'est pas convergente en tracant plusieurs trajectoires sur le même graphique. Peut-on encore vérifier à l'aide d'un tracé quantile-quantile ?

3 Méthode de rejet

Loi uniforme sur une ellipse Pour tirer une loi uniforme à l'intérieur d'une ellipse, il suffit de simuler une loi uniforme sur un domaine rectangulaire contenant l'ellipse, et de rejeter les points hors de l'ellipse. Implémenter un tel tirage dans une fonction `rellipse`. On rappelle qu'une ellipse peut être paramétrée par l'équation $(x/a)^2 + (y/b)^2 = 1$.

Vérification visuelle

- Affichez sur une même figure l'ensemble des points tirés et le tracé de l'ellipse.
- Sélectionnez quelques rectangles inclus dans l'ellipse et vérifiez la loi des grands nombres.

Question : comment tirer des rectangles uniformément à l'intérieur de l'ellipse ?

Bonus Comment tirer des points uniformément sur une surface ? (cf exercice 4 de la feuille de TD)

4 Méthode de Monte Carlo

Calcul d'une espérance par la méthode de Monte Carlo Soit X une variable aléatoire de loi normale. On propose d'évaluer $\mathbb{P}[X > t] = \mathbb{E}[\mathbb{1}_{X>t}]$ par la méthode de Monte Carlo avec une fonction `montecarlo(n,t)` où `n` est le nombre de tirages. Comparez le résultat avec la valeur donnée par la stixbox.

Réduction de la variance On diminue la variance de l'estimateur par translation en considérant : $\mathbb{P}[X > t] = e^{-c^2/2}\mathbb{E}\left[\mathbb{1}_{X>t}\, e^{c(c-X)}\right]$ où $X \sim \mathcal{N}(c, 1)$. Implémenter le nouvel estimateur dans une fonction `montecarlo(n,t,c)`.

Comparaison entre les deux méthodes On fixe un nombre de tirages autorisés à $n = 100$. Faites un grand nombre d'estimations. Illustrer le TCL. Reproduire pour l'estimateur avec variance réduite. Comparer les deux estimateurs en superposant les deux histogrammes correspondants.