ENSIIE 2024-2025
2A UE Méthodes de Simulation

Feuille 1 de Travaux pratiques

1. Méthode d'inversion

Rappel de la méthode d'inversion. On suppose que l'on sait simuler des réalisations indépendantes de loi uniforme sur $]0,1[$, c'est-à-dire, une suite $(U_n)_{n\geq 1}$ de variables aléatoires indépendantes de même loi $\mathcal{U}]0,1[$. L'appel à la fonction `rand` génère une réalisation $u = U(\omega)$ de la loi uniforme sur $]0,1[$.

Soit $U \sim \mathcal{U}(]0,1[)$ et soit X une variable aléatoire de fonction de répartition F: $F(x) = \mathbb{P}(X \leq x)$, $\forall\, x \in \mathbb{R}$. Pour simuler des réalisations de variables aléatoires $(X_n)_{n\geq 1}$ indépendantes de même loi que X on utilise le résultat suivant: si

$$F^{-1}(u) := \inf\{t \in \mathbb{R} : \quad F(t) \geq u\}, \qquad \text{pour tout } \ u \in]0,1[$$

alors X et $F^{-1}(U)$ ont même loi: on note $X \overset{d}{=} F^{-1}(U)$.

Exercice 1. (*Loi de Bernoulli*) Soit U une loi uniforme sur $]0,1[$, $p \in]0,1[$ et soit X une variable aléatoire de loi de *Bernoulli* de paramètre p: $X \sim \mathcal{B}ern(p)$. Soit F la fonction de répartition de X.

1. Montrer que $1-U$ et U ont la même loi de probabilité: si $U \sim \mathcal{U}(]0,1[)$ alors $1-U \sim \mathcal{U}(]0,1[)$.
2. Vérifier que X a la même loi de probabilité que $\mathbb{1}_{\{U\leq p\}}$ et $\mathbb{1}_{\{1-U\leq p\}}$.
3. Utiliser la question précédente pour simulez (et afficher) un échantillon $(X_i)_{1\leq i\leq N}$ de taille N de la loi de B*ern*(p), pour $p = 0.8$, pour $N = 10$.
4. Pour $N = 100$, $N = 1000$ et puis 10000, simulez un échantillon$(X_i)_{1\leq i\leq N}$ de taille N de la loi de B*ern*(p), pour $p = 0.6$, et, pour chaque N, calculez et comparez la quantité suivante avec p:
$$\frac{\#\{i \in \{1,\cdots,N\} \text{ tel que } X_i = 1\}}{N}.$$
5. On rappelle le *Théorème Central Limite* (TCL): Pour toute suite de variables aléatoires $(X_n)_{n\geq 1}$ iid de moyenne $\mu = \mathbb{E}(X_1)$ et de variance finie σ^2, on a:
$$Z_n = \sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \quad \overset{\mathcal{L}}{\longrightarrow} \quad Z \sim \mathcal{N}(0,1), \text{ lorsque } n \to \infty, \tag{1}$$
où $\bar{X}_n = (X_1 + \ldots + X_n)/n$ est la moyenne empirique associée à l'échantillon.
   (a) Gérérez (et affichez) un échantillon $(Z_n^i)_{i=1,\ldots,N}$ de taille $N = 10$ de Z_n pour $n \in \{10, 30, 50\}$.
   (b) Pour $n = 10$, gérérez un échantillon $(Z_n^i)_{i=1,\ldots,N}$ de taille $N = 10^5$ de la loi de Z_n et tracez sur le même graphique l'histogramme empirique associé à l'échantillon $(Z_n^i)_{i=1,\ldots,N}$ et la densité de la loi normale standard
$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, \quad x \in \mathbb{R}.$$
   (c) Reprendre la question précédente pour $n = 30$ et pour $n = 50$.
   (d) Commentez les graphiques obtenus dans les deux questions précédentes.

Exercice 2. (*Loi discrète quelconque*) Soit X une variable aléatoire prenant ses valeurs dans $\{a_1, a_2, a_3, a_4\}$ avec $a_1 := 0.5$; $a_2 := -9$ $a_3 := -1.5$; $a_4 := 7$ et soit $(p_i)_{i=1,\cdots,4}$ les poids associés définis par

$$\begin{cases} p_1 = \mathbb{P}(X = a_1) = 1/4 \\ p_2 = \mathbb{P}(X = a_2) = 1/8 \\ p_3 = \mathbb{P}(X = a_3) = 1/8 \\ p_4 = \mathbb{P}(X = a_4) = 1/2. \end{cases}$$

1. Simuler un échantillon indépendant de taille $N = 10000$ de X et comparez les fréquences empiriques avec leurs poids associés.

2. Soit $\mu = \mathbb{E}(X)$, $\sigma^2 = \mathit{Var}(X)$ et soit $(Z_n)_{n\geq 1}$ une suite de v.a. définie pour tout $n \geq 1$ par

$$Z_n := \sqrt{n}\,\frac{\bar{X}_n - \mu}{\sigma} \tag{2}$$

   où $\bar{X}_n = (X_1 + \ldots + X_n)/n$ est la moyenne empirique associée à l'échantillon.

3. Pour $n = 10$, gérérez un échantillon $(Z_n^i)_{i=1,\ldots,N}$ de taille $N = 10^5$ de la loi de Z_n et tracez sur le même graphique l'histogramme empirique associé à l'échantillon $(Z_n^i)_{i=1,\ldots,N}$ et la densité de la loi normale standard

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad x \in \mathbb{R}.$$

4. Reprendre la question précédente pour $n = 30$ et pour $n = 50$.

5. Commentez les graphiques obtenus dans les deux questions précédentes.

Exercice 3. Soit $n \geq 2$, $p < n$ et $E = \{1, \ldots, n\}$. On veut constituer aléatoirement p groupes de la façon suivante. L'ensemble des groupes est noté $G = \{g_1, \ldots, g_p\}$ où g_i désigne le groupe i, pour tout $i \in G$. On suppose initialement que pour tout $i \in G$, $g_i = \emptyset$. Lorsque p ne divise pas n: $n = q \times p + r$, avec $r \neq 0$, un des groupes sera constitué de $q + r$ élément(s) et le reste de q élément(s). On répète les étapes suivantes tant que *card*$(E) > 1$:

1. On choisit un groupe g de façon aléatoire selon la loi uniforme $\mathcal{U}(G)$ sur G: $g \sim U(G)$.

2. On génère un nombre aléatoire $x \sim \mathcal{U}(E)$ et on pose $E = E\backslash\{x\}$ et $g = g \cup \{x\}$.

3. Si le groupe g précédemment choisi est tel que *card*$(g) = q$, alors, on pose $G = G\backslash\{g\}$ et on suppose que le groupe g est complet.

Coder la fonction précédente pour $(p, n) = (3, 9)$ et pour $(p, n) = (10, 101)$.

Exercice 4. (*Loi géométrique*) Soit X une v.a. de loi géométrique $\mathcal{G}(p)$ de paramètre de succès p. Ecrivez une function qui simule une échantillon de taille donnée n de X et implémentez-la pour $n = 10$ et $p = 0.4$.

2. Méthodes de rejet et de transformation

Exercice 5. On veut générer une réalisation de la loi $\mathcal{N}(0, 1)$ par la méthode de rejet.

1. Montrer que si X est de loi $\mathcal{N}(0, 1)$, alors $|X|$ a pour densité

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \mathbb{1}_{\{x>0\}}.$$

2. Montrer que pour tout $x > 0$,

$$f(x) \leq \sqrt{\frac{2e}{\pi}}\, e^{-x}.$$

3. En déduire un algorithme de simulation de $|X|$ par la méthode de rejet.

4. Tracer l'histogramme des fréquences pour un échantillon de taille N assez grand et le comparer avec la densité théorique de $|X|$.

5. Soit Θ une v.a. à valeurs dans $\{-1, +1\}$ avec $\mathbb{P}(\Theta = +1) = \mathbb{P}(\Theta = -1) = 1/2$. Montrer que $X \overset{d}{=} \Theta\,|X|$.

6. Simuler un échantillon indépendant de taille N de loi normale en utilisant les questions précédentes et comparer les densités théoriques et empiriques.

Exercice 6. Soit X une variable aléatoire de loi uniforme sur la sphère unité $A = \{(x, y, z) \in \mathbb{R}^3,\ x^2 + y^2 + z^2 \leq 1\}$ et soit le domaine $S = \{(x, y, z) \in \mathbb{R}^3,\ x \in]-1, +1[,\ y \in]-1, +1[,\ z \in]-1, +1[\}$.

Générer un échantillon de taille 10000 de la v.a. X et représenter les points simulés sur $\mathbb{R}^3$.

Exercice 7. (*Méthode de Box-Muller*). Soit R une v.a. de loi exponentielle de paramètre $1/2$: $R \sim \mathrm{Exp}(1/2)$, et soit $\Theta \sim \mathcal{U}(]0, 2\pi[)$.

1. En utilisant le fait que le vecteur $(X_1, X_2) = (\sqrt{R}\cos(\Theta), \sqrt{R}\sin(\Theta)) \sim \mathcal{N}(0; I_2)$, où I_2 est la matrice identité de $\mathbb{R}^2$, simulez et affichez un échantillon de taille $N = 10$ de la gaussienne bidimensionnelle.

2. Générez un échantillon $(X_1^k, X_2^k)_{\{1 \leq k \leq N\}}$ de (X_1, X_2) de taille $N = 10^5$ et représentez graphiquement le nuage de points $\{(X_1^k, X_2^k),\ 1 \leq k \leq N\}$.

3. Tracez dans deux graphiques différents, la densité de (X_1, X_2) et l'histogramme associé à un échantillon de taille $N = 10^5$ de (X_1, X_2).

4. En utilisant la question 1., simuler un échantillon de taille $N = 10^5$ de $X_1 \sim \mathcal{N}(0; 1)$. Tracez l'histogramme associé et le comparer (dans le même graphe) à la densité d'une loi $\mathcal{N}(0; 1)$.

5. Soit $\mu_1, \mu_2 \in \mathbb{R}$ et $\sigma_{ij} \geq 0$, pour $i, j \in \{1, 2\}$. Soit

$$\begin{cases} Z_1 = \mu_1 + \sigma_{11}X_1 + \sigma_{12}X_2 \\ Z_2 = \mu_2 + \sigma_{21}X_1 + \sigma_{22}X_2. \end{cases}$$

   On montre que $Z = (Z_1, Z_2) \sim \mathcal{N}(\mu, \Sigma)$ où

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

   et σ_1, σ_2, ρ sont définis par:

$$\sigma_1^2 = \sigma_{11}^2 + \sigma_{12}^2, \qquad \sigma_2^2 = \sigma_{21}^2 + \sigma_{22}^2, \qquad \rho = \frac{\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}}{\sigma_1\sigma_2}.$$

   (a) On pose $\mu_1 = \mu_2 = 0$. Choisir des valeurs de σ_{ij} pour que ρ prenne les valeurs $0.1,\ 0.5,\ 0.9$. Pour chaque valeur de $\rho \in \{0.1, 0.5, 0.9\}$, générez et représentez graphiquement un échantillon de taille $N = 10^4$ de (Z_1, Z_2). Commentez les graphes.

   (b) Pour chaque $\rho \in \{0.1, 0.5, 0.9\}$, représentez la densité de (Z_1, Z_2) qui s'écrit:

$$f_{(Z_1,Z_2)}(z_1, z_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Big(-\frac{z}{2(1-\rho^2)}\Big), \quad \text{avec } z = \frac{z_1^2}{\sigma_1^2} - \frac{2\rho z_1 z_2}{\sigma_1\sigma_2} + \frac{z_2^2}{\sigma_2^2}.$$

   Commentez les graphes.

Exercice 8. (*Mélange de lois*) Soient $(p_1, p_2, p_3) = (1/6, 1/3, 1/2)$ et X une v.a. de densité

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + p_3 f_3(x)$$

où $f_1(x) = \mathbb{1}_{]0,1]}(x),\ f_2(x) = \frac{1}{2}(2x-1)\mathbb{1}_{]1,2]}(x),\ f_3(x) = \frac{2}{3}(-3x+9)\mathbb{1}_{]2,3]}(x)$.

Générer un échantillon de taille $N = 10^5$ de X et comparez l'histogramme associé à la densité de X.

Exercice 9. Soit $X_1 \sim \mathcal{N}(-3, 1)$ et $X_2 \sim \mathcal{N}(3, 1)$ deux v.a. indépendantes de densités respectives f_1 et f_2. Soit X la v.a. de densité

$$f(x) = p_1 f_1(x) + p_2 f_2(x), \qquad p_1, p_2 \in [0, 1],\ p_1 + p_2 = 1.$$

Générer un échantillon de taille $N = 10^5$ de X et comparez l'histogramme associé à la densité de X dans chaque cas suivant: $(p_1, p_2) = (1/2, 1/2)$, $(p_1, p_2) = (1/8, 7/8)$, $(p_1, p_2) = (7/8, 1/8)$.