Projet MSRO

Yassir ANKOUDY

Simulation d'une loi Gaussienne

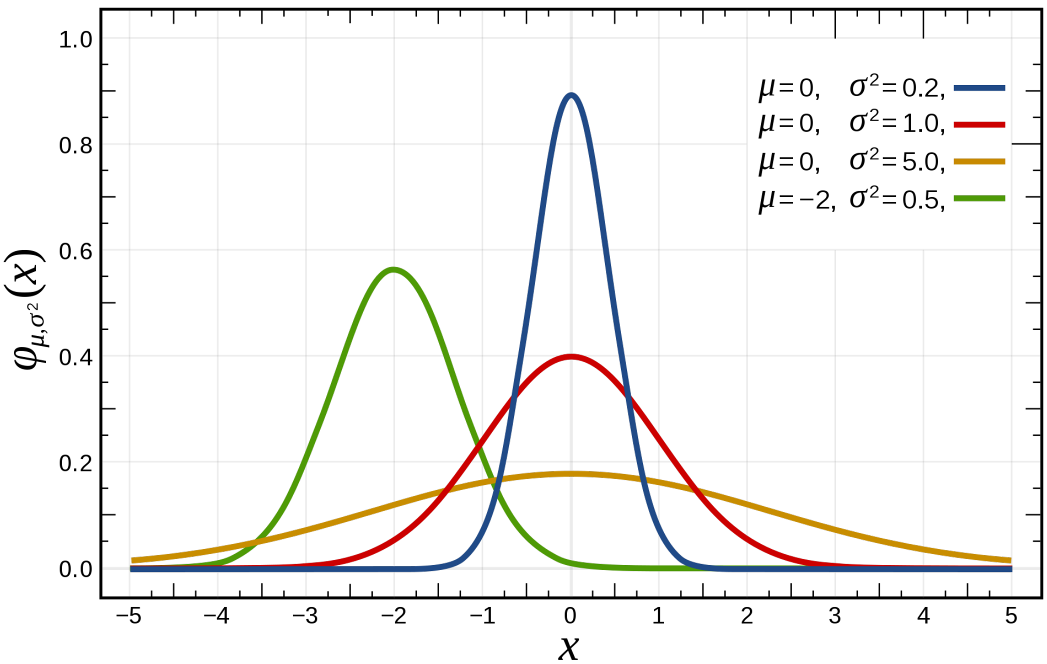


Figure 1: Densité de probabilité de la loi normale

October 23, 2024

Sommaire

1 Introduction

Le but de ce projet est de simuler numériquement une loi gaussienne. Mais il est difficile pour un ordinateur de simuler cette loi, donc pour pouvoir simuler cette loi il faut effectuer une transformation pour permettre a la machine de faciliter la simulation. Pour cela nous allons utiliser une loi uniforme,loi qui est facilemment simulable car elle fait appel a une fonction aleatoire sur un intervalle deja donné.Ce type de fonction est très répandu et simple à mettre en place dans n'importe quel langage de programmation. On peut grâce à ce processus simuler de nombreuses lois discrètes comme la loi de Bernouilli, la loi Binomiale, la loi géométrique ou encore la loi de Poisson et ainsi résoudre numériquement de nombreux problèmes de probabilités.

Ce rapport sera divisé en deux parties, on commencera par définir la loi gaussienne et les propriétés qui lui sont liées, par la suite, nous présenterons une méthode permettant de simuler cette loi, et enfin nous implémenterons la méthode en Python afin de vérifier son efficacité.

2 Partie théorique

2.1 Définition: Loi Gaussienne

On parle de loi Gaussienne lorsque l'on a affaire à une variable aléatoire continue dépendant d'un grand nombre de causes indépendantes dont les effets s'additionnent et dont aucune n'est prépondérante, c'est ce que l'on appelle les conditions de Borel. Cette loi acquiert sa forme définitive avec Gauss en 1809 et Laplace en 1812. C'est pourquoi elle porte également les noms de : loi de Laplace, loi normale et loi de Laplace-Gauss.

Parmi les lois de probabilité, les lois gaussiennes prennent une place particulière grâce au théorème central limite. En effet, elles correspondent au comportement, sous certaines conditions, d'une suite d'expériences aléatoires similaires et indépendantes lorsque le nombre d'expériences est très élevé. Grâce à cette propriété, une loi gaussiennes permet d'approcher d'autres lois et ainsi de modéliser de nombreuses études scientifiques comme des mesures d'erreurs ou des tests statistiques, en utilisant par exemple les tables de la loi normale centrée réduite.

Cas général loi normale

Soit X une variable aléatoire continue qui suit une loi normale de paramètres (μ, σ).

- **Notation** $X \hookrightarrow \mathcal{N}(\mu, \sigma)$.
- **Densité de probabilité:** La densité de probabilité de la loi gaussienne se définit comme suit:

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

$$x \longmapsto f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

avec $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}^+$

2.2 Propriétés de la loi gaussienne

Proposition

Si X suit une loi normale $N(m, \sigma)$, $Z = \frac{X-\mu}{\sigma}$ suit une loi normale centre réduite $N(0, 1)$.

Preuve

Soit deux réels a, b tel que $a < b$

$$P(a \leq Z \leq b) = P\left(a \leq \frac{X-\mu}{\sigma} \leq b\right)$$

$$= P(a\sigma \leq X - \mu \leq \sigma b)$$

$$= P(m + a\sigma \leq X \leq m + \sigma b)$$

$$= \int_{m+a\sigma}^{m+b\sigma} \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right) dx$$

On effectue le changement de variable : $z = \frac{X-\mu}{\sigma}$

$$P(a \leq Z \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} exp\left(-\frac{z^2}{2}\right) dz$$

Cette égalite étant vraie pour tous réels $a < b$, on en déduit que $\frac{1}{\sqrt{2\pi}} exp\left(-\frac{z^2}{2}\right)$ est la densité de Z, autrement dit, Z suit une loi normale centrée réduite $N(0, 1)$.

Proposition

Soit une variable $X \sim N(\mu, \sigma^2)$ on a :

$$E(X) = \mu \textbf{ et } V(X) = \sigma^2$$

Preuve

immédiat avec la proposition précédente et la linéarite de l'esperance. On peut ainsi facilement se ramener a la loi normale centrée réduite. Comme la fonction de répartition φ de la loi normale centrée réduite est tabulée, on se ramenera meme systématiquement a cette loi pour calculer des probabilités. Plus précisement, si X est de loi $N(\mu, \sigma)$, et a et b sont deux reels, on calcule $P(a \leq X \leq b)$ ainsi :

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$$

$$= P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right)$$

ou la variable Z est de loi $N(0, 1)$ .Donc

$$P(a \leq X \leq b) = \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

On peut aussi remarquer que si Z est de loi $N(0, 1)$, alors $-Z$ aussi et donc, pour $u > 0$,

$$P(-u \leq Z \leq u) = P(Z \leq u) - P(u \geq u)$$

$$= P(Z \leq u) - (1 - P(Z \leq u))$$

$$= 2P(Z \leq u) - 1$$

Exemple:

Soit X une variable aléatoire qui suit une loi normale $N(15, 4)$.
Quelle est la probabilité $P(10 \leq X \leq 22)$?

$$P(10 \leq X \leq 22) = P\left(\frac{10-15}{4} \leq \frac{X-15}{4} \leq \frac{22-15}{4}\right)$$

$$= P(-1.25 \leq Y \leq 1.75)$$

où la variable Y suit une loi normale $N(0, 1)$ ainsi on a:

$$P(10 \leq X \leq 22) = P(Y \leq 1.75) + P(Y \leq 1.25) - 1$$

$$= 0.9599 + 0.8944 - 1 = 0.8543$$

2.3 Définition: Loi normale centrée réduite

La loi normale centrée réduite est un cas particulier de la loi gaussienne (ou normale).

Cas centrée réduite

Pour une variable aléatoire X, on dit qu'elle suit une loi normale centrée réduite si sa densité est donnée par:

- **Notation** $X \hookrightarrow \mathscr{N}(\mu, \sigma)$
- **Densité de probabilité:**

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

Pour la loi normale centrée réduite, la moyenne μ est nulle et l'écart type σ est unitaire.

Rappel

Si Xsuit une loi gaussienne $N(0,1)$ alors $\forall a$, tel que$a < b$, on a:

$$P[a \leq X \leq b] = \frac{1}{\sqrt{2\pi}} \int_a^b exp\left(-\frac{x^2}{2}\right) dx = \varphi(b) - \varphi(a)$$

où la fonction φ est la fonction de répartion de X.

Rappelons que φ est une primitive de la densite f. Mais il n'existe pas de forme analytique de la primitive de f. On doit donc lire les valeurs de φ dans le tableau de définition, ou effectuer les calculs avec un logiciel adapte (ex : excel, matlab, R).

Courbe densité

La courbe de la densité de la loi normale $N(0,1)$ porte le nom de "courbe en cloche".

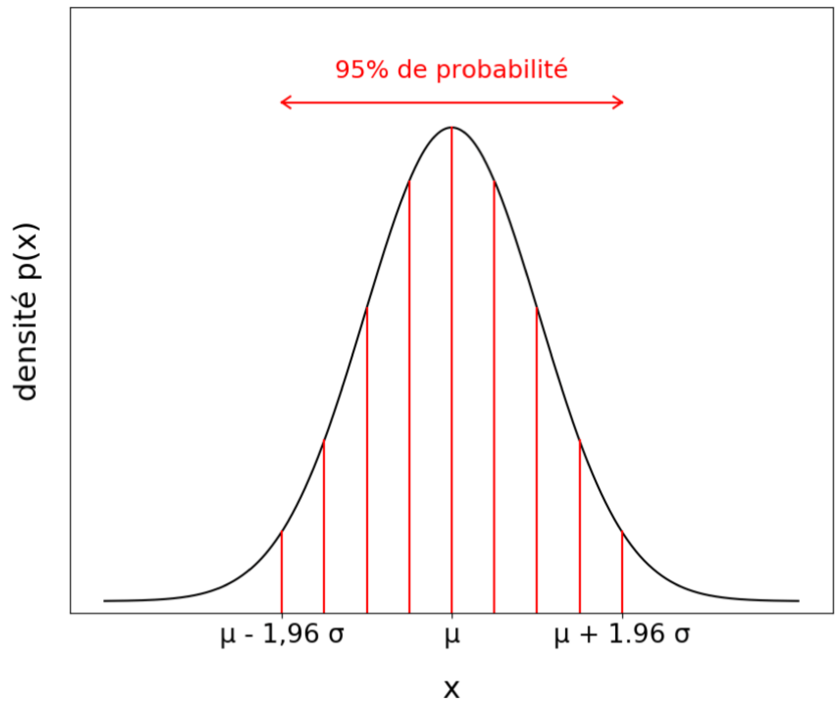


Figure 2: Courbe de la densité de la loi normale centrée réduite

Elle tend vers 0 en l'infini, est croissante sur $\mathbb{R}^-$, puis décroissante. Elle admet donc un maximum en 0. On peut voir aussi qu'elle est symetrique, de centre de symetrie 0.

2.4 Propriétés de la loi normale centrée réduite

Proposition

La variable X est centrée si $E\left(X\right) = 0$ c'est-à-dire de moyenne nulle, et la variable X réduite si $V\left(X\right) = 1$ c'est-à-dire de variance 1.

De plus la variable $Y = -X$ suit encore une loi normale centrée réduite.

Preuve

Le calcul de l'espérance est immédiat quand on a observe que $xf\left(x\right)$ est une fonction impaire. Pour le calcul de la variance, il faut effectuer une intégration par parties. Enfin, montrons que X et $-X$ ont meme loi. Pour tous $a, b \in R$,

$$P\left(a \leq -X \leq b\right) = P\left(-b \leq X \leq -a\right) = \int_{-b}^{-a} f\left(x\right) dx$$

on effectue un changement de variable : $y = -x$

$$P\left(a \leq -X \leq b\right) = \int_{a}^{b} f\left(-y\right) dy = \int_{a}^{b} f\left(y\right) dy$$

par symetrie de la fonction f.

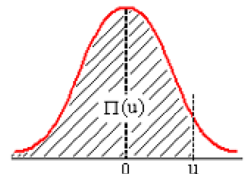
D'où,

$$P\left(a \leq -X \leq b\right) = P\left(a \leq X \leq b\right)$$

2.5 Utilisation des tables

Pour calculer $P(a \leq X \leq b)$ ou $P(X \leq x)$, on a recours au calcul numerique sur ordinateur ou, plus simplement, a une table qui donne $P(X \leq x)$ pour toute décimale x à deux chiffres après la virgule. Voici la table utilisée en général:

Table de Loi Normale
P(x<u)



	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	.0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8254	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	.0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0.9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	6,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	.0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Figure 3: Table de la loi normale centrée réduite

Puis il faut utiliser le fait que :

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a)$$

On lit les probabilités dans la table si a et b sont positifs. Pour trouver $P(X \leq -x)$. quand $x > 0$, on utilise le fait que X et $-X$ suivent la même loi:

$$P(X \leq -x) = P(-X \geq x) = P(X \geq x) = 1 - P(X \leq x)$$

Exemple 1:

On cherche à calculer $P(1 \leq X \leq -1)$.

$$P\left(-1 \leq X \leq 1\right) = P\left(X \leq 1\right) - P\left(X \leq -1\right) = P\left(X \leq 1\right) - \left(1 - P\left(X \leq 1\right)\right)$$

$$= 2P\left(X \leq 1\right) - 1 = 2 \times 0.8413 - 1 = 0.6826$$

Exemple 2:

On cherche $u \in \mathbb{R}$ tel que $P\left(-u \leq X \leq u\right) = 0.90$.

$$P\left(-u \leq X \leq u\right) = P\left(X \leq u\right) - P\left(X \leq -u\right) = P\left(X \leq u\right) - \left(1 - P\left(X \leq u\right)\right)$$

$$= 2P\left(X \leq u\right) - 1$$

D'où $P\left(X \leq u\right) = 0.95$ et $u = 1.6446$

3 Méthode de simulation d'une loi gaussienne

Il existe différentes méthodes qui permettent de simuler une loi gaussienne mais dans le cadre de ce projet nous allons nous focaliser sur la méthode de Box-Muller.

3.1 Méthode de Box-Muller

C'est une méthode pour simuler un couple de variables aléatoires indépendantes de loi normale centrée réduite. Elle est basée sur le théorème suivant :

Théorème Box-Muller

Soient X et Y deux variable aléatoires gaussiennes centré réduite indépendantes. Definissons (R,θ) les coordornnée polaire de (X,Y):

$$X = Rcos(\theta) \textbf{ et } Y = Rsin(\theta)$$

avec $R \geq 0$ et $\theta \in [0, 2\Pi]$.
Alors R^2 et θ sont deux variables aléatoires indépendantes, tel que la variable R^2 suit une loi exponentielle de paramètre $\frac{1}{2}$ c'est-à-dire $\left(R^2 \sim \varepsilon(\frac{1}{2})\right)$ et la variable θ suit une loi uniforme sur $[0, 2\pi]$, soit $(X \sim ([0, 2\pi]))$. De plus, les variables X et Y sont indépendant de loi normale $N(0,1)$.

Démonstration

On pose $S = R^2$, soit f une fonction borélienne bornée. On a:

$$E\left[f\left(X,Y\right)\right] = E\left[\int\left(Rcos\left(\theta\right), Rsin\left(\theta\right)\right)\right]$$

$$= \int_{\mathbb{R}+}\int_{[0,2\pi]} f\left(\sqrt{s}cos\left(\theta\right), \sqrt{s}sin\left(\theta\right)\right) dsd\theta$$

ensuite effectuons un changement de variable en posant :

$$\begin{cases} x = & \sqrt{s}cos\left(\theta\right) \\ y = & \sqrt{s}sin\left(\theta\right) \end{cases}$$

De plus on a :

$$E\left[f\left(X,Y\right)\right] = \int_{\mathbb{R}+}\left(f\left(x,y\right)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\right)dxdy$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}^2} f\left(x,y\right)e^{-\frac{\left(x^2+y^2\right)}{2}}dxdy$$

On en déduit que les variables X et Y sont indépendantes et suivent la loi normale centrée réduite $N\left(0,1\right)$.
Le lemme (résultat) suivant est important pour la suite de notre projet car il va nous permettre de construire les variables nous permettant de simuler une gaussienne numériquement.

Lemme Box-Muller

si Uet V sont indépendantes et de loi uniforme sur $]0,1[$ alors, les variables aléatoires

$$X = \sqrt{-2log(U)}cos(2\pi V)$$

et

$$Y = \sqrt{-2log(U)}sin(2\pi V)$$

sont indépendantes et suivent la loi normale $N(0,1)$.

Afin de démontrer le lemme de Box-Muller, nous allons utiliser la propriété suivante:

Lemme Box-Muller

Soit U une variable aléatoire uniforme sur l'intervalle $[0,1]$, et définissons $X = \alpha U + \beta$, où $\alpha > 0$ et $\beta \in \mathbb{R}$..

Démonstration

Nous voulons montrer que X suit une loi uniforme sur l'intervalle $[\beta, \alpha+\beta]$.

1. **Bornes de** X :

$$\begin{gathered} 0 \leq U \leq 1 \\ \Longleftrightarrow 0 \leq \alpha U \leq \alpha \\ \Longleftrightarrow \beta \leq \alpha U + \beta \leq \alpha + \beta \\ \Longleftrightarrow \beta \leq X \leq \alpha + \beta \end{gathered}$$

Les bornes de X sont vérifiées.

2. **Fonction de répartition de** X :

La fonction de répartition $(F_X(x))$ d'une variable aléatoire continue (X) suivant une loi uniforme sur l'intervalle $[a,b]$ est donnée par :

$$F_X(x) = \frac{x-a}{b-a}, \quad \textbf{pour } a \leq x \leq b.$$

Dans le cas de la loi uniforme sur l'intervalle $[0,1]$, nous avons $a = 0$ et $b = 1$.

Soit U une variable aléatoire qui suit une loi uniforme sur $[0,1]$. La fonction de répartition de U est donnée par :

$$F_U(u) = \begin{cases} 0, & \textbf{si } u < 0 \\ u, & \textbf{si } 0 \leq u < 1 \\ 1, & \textbf{si } u \geq 1 \end{cases}$$

Nous voulons trouver la fonction de répartition de X notée $F_X(x)$.

$$F_X(x) = P(X \leq x) = P(\alpha U + \beta \leq x) = P(U \leq \frac{(x-\beta)}{\alpha}) = F_U(\frac{(x-\beta)}{\alpha})$$

En substituant $F_U(u)$ dans l'équation ci-dessus, nous obtenons :

$$F_X(x) = \begin{cases} 0 & \textbf{si } \frac{(x-\beta)}{\alpha} < 0 \\ (x-\beta)/\alpha & \textbf{si } 0 \leq \frac{(x-\beta)}{\alpha} < 1 \\ 1 & \textbf{si } \frac{(x-\beta)}{\alpha} \geq 1 \end{cases}$$

Cela peut être réécrit comme :

$$F_X(x) = \begin{cases} 0 & \textbf{si } x < \beta \\ \frac{(x-\beta)}{\alpha} & \textbf{si } 0 \leq \beta \leq x < \beta + \alpha \\ 1 & \textbf{si } x \geq \alpha + \beta \end{cases}$$

On retrouve donc la fonction de répartition d'une variable aléatoire qui suit une loi uniforme sur $[\beta, \alpha + \beta]$.

3. **Densité de probabilité de** X : La densité de probabilité de X est la dérivée de la CDF :

$$f_X(x) = \frac{d}{dx} F_X(x)$$

La dérivée est nulle pour $x < \beta$ et $x > \alpha + \beta$.

Dans l'intervalle $\beta \leq x \leq \alpha + \beta$, la dérivée est constante :

$$f_X(x) = \frac{1}{\alpha}$$

Ainsi, la densité de probabilité de X est constante sur l'intervalle $[\beta, \alpha + \beta]$.

Vérification : Vérifions que l'intégrale de la densité de probabilité sur l'intervalle $[\beta, \alpha+\beta]$ est égale à 1 :

$$\int_{\beta}^{\alpha+\beta} \frac{1}{\alpha} \, dx = \frac{x}{\alpha}\Big|_{\beta}^{\alpha+\beta} = \frac{\alpha + \beta - \beta}{\alpha} = 1$$

Ainsi, la densité de probabilité de X satisfait les conditions nécessaires.

En conclusion, nous avons démontré que si U suit une loi uniforme sur $[0, 1]$, alors la variable aléatoire $X = \alpha U + \beta$ suit une loi uniforme sur $[\beta, \alpha + \beta]$.

3.2 Démonstration du Lemme de Box-Muller

Le vecteur aléatoire (U, V) suit la loi uniforme sur $[0, 1]^2$ de densité respective $f_{U,V} = 1_{]0,1[^2}$.

Effectuons un changement de variable, soit $g : (U, V) \mapsto (x, y)$ avec $\begin{cases} x = & \sqrt{-2log\,(U)}cos\,(2\pi V) \\ y = & \sqrt{-2log\,(U)}sin\,(2\pi V) \end{cases}$

Le vecteur aléatoire $(X, Y) = g\,(U, V)$ a une densité $f_{X,Y}$donné par

$$f_{X,Y}\,(x, y) = f_{U,V}(g\,(x, y)) \left| jac\left(g^{-1}\right)(x, y) \right| 1_{D'}\,(x, y)$$

Pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, $g^{-1}\,(x, y) \in [0, 1]^2$ et comme la densité $f_{U,V}$ vaut 1 sur $[0, 1]$ alors le premier facteur dans la formule ci-dessus s'écrit $f_{U,V}\,(g^{-1}\,(x, y)) = 1$.

De plus pour calculer $jac\,(g^{-1})\,(x, y)$ on calacule $jac\,(g)\,(U, V)$ qui est la matrice de Jacobi et pour la calculer on utilise la formule suivante:

$$jac\left(g^{-1}\right)(x, y) = \frac{1}{jac\,(g)\,(g^{-1}\,(x, y))}.$$

en notant $g = (g_1, g_2)$, les derivées partielles de g sont :

$$\frac{\partial g_1}{\partial u} = -u^{-1}\,(-2log\,(u))^{-\frac{1}{2}}\,cos\,(2\pi v)$$

$$\frac{\partial g_1}{\partial v} = -2\pi\,(-2log\,(u))^{\frac{1}{2}}\,sin\,(2\pi v)$$

$$\frac{\partial g_2}{\partial u} = -u^{-1}\left(-2log\left(u\right)\right)^{\frac{-1}{2}} sin\left(2\pi v\right)$$

$$\frac{\partial g_2}{\partial v} = 2\pi\left(-2log\left(u\right)\right)^{\frac{1}{2}} cos\left(2\pi v\right)$$

on en déduis que :

$$jac\left(g\right)\left(u,v\right) = \begin{vmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{vmatrix}$$

$$= \frac{2\pi}{u} cos^2\left(2\pi v\right) + \frac{2\pi}{u} sin^2\left(2\pi v\right) = \frac{2\pi}{u}\left(cos^2\left(2\pi v\right) + sin^2\left(2\pi v\right)\right)$$

$$= \frac{2\pi}{u}$$

car $cos^2\left(2\pi v\right) + sin^2\left(2\pi v\right) = 1$.

pour exprimer u en fonction de (x,y) on remarque que $x^2+y^2 = -2log\left(u\right)$, d'où $u = exp\left(-\left(x^2+y^2\right)/2\right)$ et

$$\left|jac\left(g^{-1}\right)\left(x,y\right)\right| = \frac{1}{2\pi} exp\left(-\frac{x^2+y^2}{2}\right).$$

Finalement, la densité du couple (X,Y) est donnée par

$$f_{X,Y}\left(x,y\right) = \frac{1}{2\pi} exp\left(-\frac{x^2+y^2}{2}\right) 1_{D'}\left(x,y\right).$$

comme $1_{D'}$et $1_{\mathbb{R}^2}$ ne different que du singleton $\{(0,0)\}$ qui est un ensemble de mesure nulle , (X,Y) admet aussi pour densité

$h_{X,Y}\left(x,y\right) = \left(2\pi\right)^{-1} exp\left(-\left(x^2+y^2\right)/2\right)$ qui est la forme classique de la densité dun couple de variables aléatoire de loi normale $N\left(0,1\right)$

Donc les variables X et Y des variables gaussiennes de loi normale $N\left(0,1\right)$ indépendantes .

3.3 Autre démonstration sans Jacobienne

Soient U et V deux variables aléatoires indépendantes qui suivent une loi uniforme sur [0,1]. On pose :

$$Z_0 = \sqrt{-2\ln U}\cos(2\pi V)$$

$$Z_1 = \sqrt{-2\ln U}\sin(2\pi V)$$

On veut montrer que Z_0 et Z_1 sont indépendantes et suivent une loi normale centrée réduite. Pour cela, on utilise le fait que la fonction de répartition conjointe de Z_0 et Z_1 est égale au produit des fonctions de répartition marginales si et seulement si Z_0 et Z_1 sont indépendantes. On a :

$$F_{Z_0,Z_1}(z_0,z_1) = P(Z_0 \leq z_0, Z_1 \leq z_1)$$

$$= P(\sqrt{-2\ln U}\cos(2\pi V) \leq z_0, \sqrt{-2\ln U}\sin(2\pi V) \leq z_1)$$

$$= P(-2\ln U \leq z_0^2 + z_1^2, 2\pi V \in [0, \arctan(z_1/z_0)] \cup [\pi + \arctan(z_1/z_0), 2\pi])$$

$$= P(U \geq e^{-\frac{z_0^2+z_1^2}{2}}, V \in [\frac{1}{2\pi}\arctan(z_1/z_0), \frac{1}{2\pi}(\pi + \arctan(z_1/z_0))])$$

$$= \int_{e^{-\frac{z_0^2+z_1^2}{2}}}^{1} \int_{\frac{1}{2\pi}\arctan(z_1/z_0)}^{\frac{1}{2\pi}(\pi+\arctan(z_1/z_0))} 1 du dv$$

$$= (1 - e^{-\frac{z_0^2+z_1^2}{2}})\frac{1}{2\pi}(\pi + \arctan(z_1/z_0) - \arctan(z_1/z_0))$$

$$= (1 - e^{-\frac{z_0^2+z_1^2}{2}})\frac{1}{2}$$

En dérivant deux fois par rapport à Z_0 et Z_1, on obtient la densité conjointe de Z_0 et Z_1 :

$$f_{Z_0,Z_1}(z_0, z_1) = \frac{\partial^2 F_{Z_0,Z_1}(z_0, z_1)}{\partial z_0 \partial z_1}$$

$$= \frac{1}{2} e^{-\frac{z_0^2+z_1^2}{2}} \frac{1}{2\pi}(2\pi)$$

$$= \frac{1}{2\pi} e^{-\frac{z_0^2+z_1^2}{2}}$$

On reconnaît la densité d'une loi normale bivariée centrée réduite. On en déduit que Z_0 et Z_1 sont indépendentes et suivent une loi normale centrée réduite.

4 Simulation et résultats

Nous allons implémenter en Python deux façons de simuler une gaussienne et les comparer. Tout d'abord, la méthode directe qui est la suivante: Pour simuler une variable aléatoire gaussienne de moyenne μ et de variance σ^2, il suffit de poser $X = \mu + \sigma Y$, où Y suit une distribution normale standard $N(0, 1)$. Et ensuite en utilisant la méthode Box-Muller que nous avons présenté précedemment.

4.1 Codes Python

Voici le code Python pour génerer une gaussienne avec la méthode directe.

```
def simulate_gaussian_direct(mu, sigma, size=1):
    # Générer des échantillons selon une distribution normale standard (N(0, 1))
    standard_normal_samples = np.random.normal(0, 1, size)

    # Appliquer la transformation pour obtenir une gaussienne de moyenne mu et variance sigma^2
    gaussian_samples = mu + sigma * standard_normal_samples

    return gaussian_samples
```

Voici le code Python pour générer une gaussienne avec la méthode de Box-Muller.

```
def simulate_gaussian_box_muller(mu, sigma, size=1):
    # Générer des échantillons selon la méthode de Box-Muller
    u1 = np.random.uniform(0, 1, size)
    u2 = np.random.uniform(0, 1, size)
    z0 = np.sqrt(-2 * np.log(u1)) * np.cos(2 * np.pi * u2)

    # Appliquer la transformation pour obtenir une gaussienne de moyenne mu et variance sigma^2
    gaussian_samples = mu + sigma * z0

    return gaussian_samples
```

Le fichier Python est disponible en PJ du rendu.

4.2 Résultats et comparaison

Test Box-Muller

Notre code Box-Muller fonctionne pour différents paramètres μ et σ.

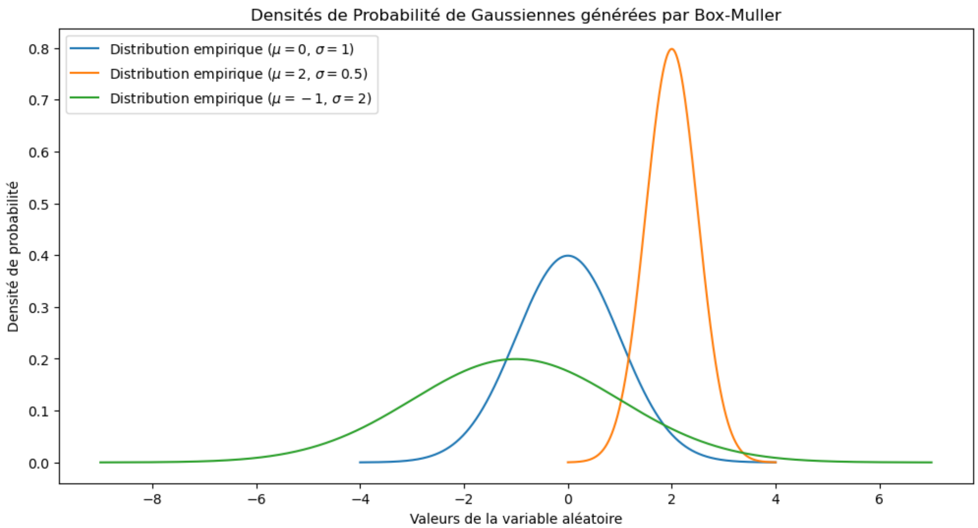


Figure 4: Densité de probabilité de la loi gaussienne pour différents paramètres

Comparaison des méthodes

Ici, on affiche la fonction de répartition, la densité de probabilité et la distribution empirique de chaque loi. Pour ce test, on prendra $\mu = 2.0$ **et** $\sigma = 1.5$. On simule donc deux lois Gaussiennes suivant $N(2.0, 1.5^2)$.

Moyenne empirique (Transformation directe) : 1.9713552971145183

Moyenne empirique (Box-Muller) : 1.95664765753577

Variance empirique (Transformation directe) : 2.2156923805950925

Variance empirique (Box-Muller) : 2.370483386534606

On remarque que les moyennes et variances empiriques sont très proches.

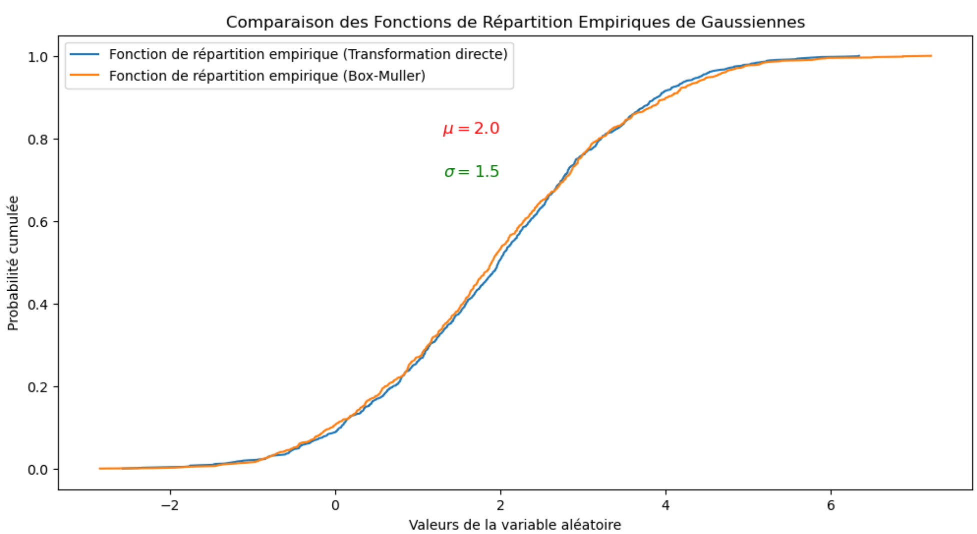


Figure 5: Fonctions de répartition des lois Gaussiennes simulées par 2 méthodes

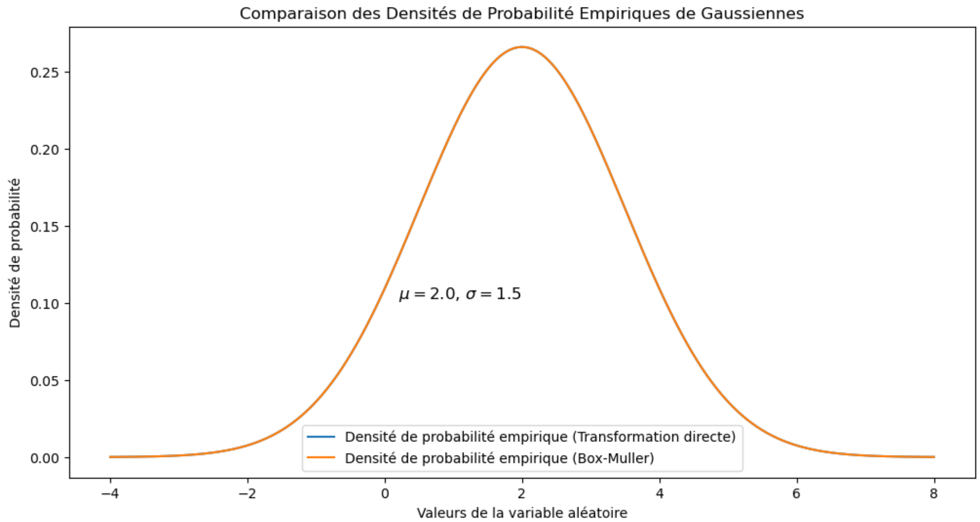


Figure 6: Densités de probabilités des lois Gaussiennes simulées par 2 méthodes

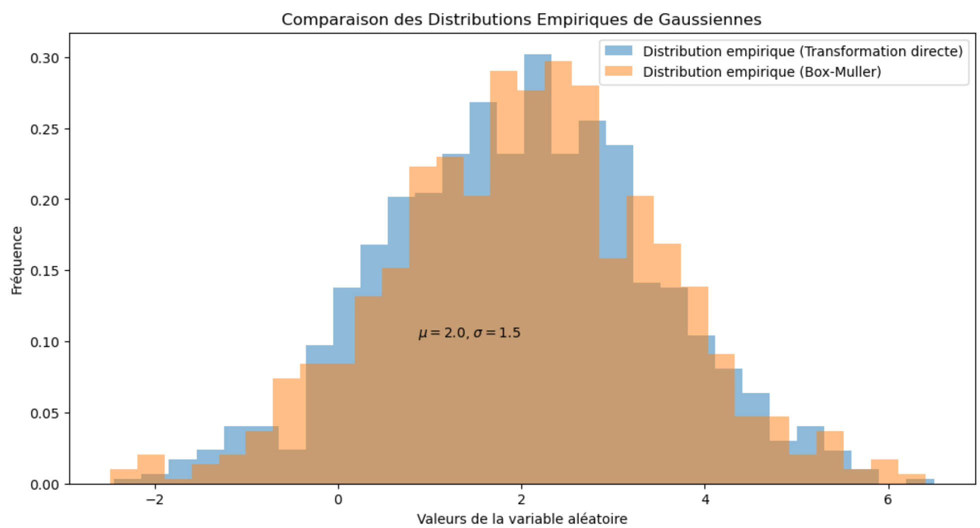


Figure 7: Distributions des lois Gaussiennes simulées par 2 méthodes

On remarque à chaque fois que les courbes se superposent et présentent une allure similaire, ce qui nous permet de dire que les deux méthodes sont identiques et que la méthode de Box-Muller permet de générer une loi Gaussienne tout comme la méthode directe.

5 Conclusion

Pour conclure, nous pouvons dire que ce projet nous a permis de comprendre la loi gaussienne ainsi que toutes ses propriétés, et qu'il nous a permis de découvrir différentes méthodes, dont celle de Box-Muller, qui permettent de simuler une gaussienne et surtout de verifier l'efficacité de cette méthode en la comparant avec la méthode directe.