

Mesures et Probabilités Statistiques

Rapport de TP : Simulation de lois de probabilités

Allan De Clercq¹

¹Université de Lille

14 octobre 2024

Tuteurs : Assi N'Guessan Assi.Nguessan@polytech-lille.fr

Abstract :

Ce document présente le rapport d'un travail pratique dont l'objectif était d'explorer et d'analyser les propriétés statistiques de différentes distributions de probabilité. Ce projet explore la simulation de lois de probabilité à travers une série de travaux pratiques. En suivant les consignes et le code fournis dans le lien du TP [1], nous avons généré diverses distributions probabilistes et analysé leurs propriétés statistiques. Les résultats obtenus montrent une concordance significative avec les théories statistiques attendues, confirmant ainsi la validité des méthodes utilisées.

Ce travail a permis de renforcer la compréhension des concepts statistiques fondamentaux et de leur application pratique à travers des simulations. Les méthodes employées ont été validées par des résultats cohérents avec les attentes théoriques, démontrant l'efficacité des techniques de simulation utilisées.

En conclusion, ce projet a non seulement permis d'approfondir les connaissances en statistiques, mais aussi de développer des compétences pratiques en programmation et en analyse de données. Le code source complet, incluant les scripts pour générer les figures et tableaux, est disponible sur le répertoire GitHub suivant : <https://github.com/AllanDCQ/Simulation-de-lois-de-probabilit-s/>.

Table des matières

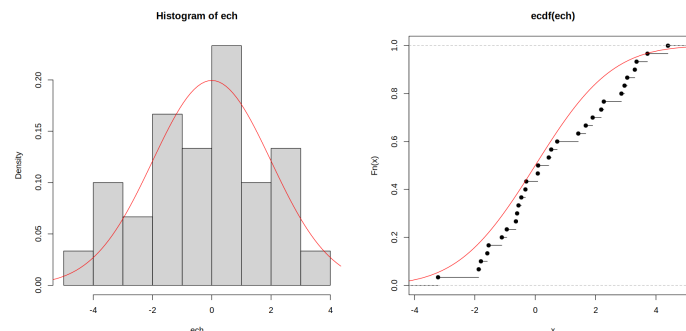
Introduction	1
0.1 Exemple pour la loi normale	1
0.2 Analyse des résultats	1
1 Lois discrètes	1
1.1 Loi uniforme discrète	1
1.1.1 Exemple - Simulation de la loi uniforme discrète	2
1.1.2 Analyse des résultats	2
1.2 Loi de Bernoulli	2
1.2.1 Exemple - Simulation de la loi de Bernoulli	2
1.2.2 Analyse des résultats	3
1.3 Loi binomiale	3
1.3.1 Exemple - Simulation de la loi de binomiale	3
1.3.2 Analyse des résultats	3
1.4 Loi Géométrique	4
1.4.1 Exemple - Simulation de la loi géométrique	4
1.4.2 Analyse des résultats	4
1.5 Loi de Poisson	4
1.5.1 Exemple - Simulation de la loi Poisson	5
1.5.2 Analyse des résultats	5
2 Lois continues	6
2.1 Loi Uniforme sur $[0, 1]$, notée $\mathcal{U}([0, 1])$	6
2.1.1 Exemple - Simulation de la loi uniforme sur $[0, 1]$	6
2.1.2 Analyse des résultats	7
2.2 Loi Uniforme sur $[a, b]$, notée $\mathcal{U}([a, b])$ avec $a < b$	7
2.2.1 Exemple - Simulation de la loi uniforme sur $[a, b]$	7
2.2.2 Analyse des résultats	7
2.3 Loi Normale centrée réduite, notée $\mathcal{N}(0, 1)$	8
2.3.1 Exemple - Simulation de la loi normale centrée réduite $\mathcal{N}(0, 1)$	8
2.3.2 Analyse des résultats	8
2.3.3 Démonstration	8
2.4 Loi Normale, notée $\mathcal{N}(\mu, \sigma^2)$	9
2.4.1 Exemple - Simulation de la loi normale centrée réduite $\mathcal{N}(1, 4)$	9
2.4.2 Analyse des résultats	9
2.5 Théorème	9
2.6 Exercices sur la loi normale	10
2.6.1 Question 1 : Voir Code TP, section : Exercices sur la loi normale, Question 1 [1] . .	10
2.6.2 Question 2 : Voir Code TP, section : Exercices sur la loi normale, Question 2 [1] . .	10
2.6.3 Question 3 : Voir Code TP, section : Exercices sur la loi normale, Question 3 [1] . .	11
2.6.4 Question 4 : Voir Code TP, section : Exercices sur la loi normale, Question 4 [1] . .	11
2.6.5 Question 5 : Voir Code TP, section : Exercices sur la loi normale, Question 5 [1] . .	12
2.6.6 Question 6	12
2.7 Lien entre les lois normales	13
2.8 Sommes de variables aléatoires gaussiennes	14
2.9 Somme de carrés de variables gaussiennes centrées réduites	14
2.10 Exercice	15
2.11 Loi Exponentielle	15
Code annexe	16

Introduction

Dans le logiciel R, on peut simuler très simplement à peu près toutes les lois classiques. Dans ce TP, on se propose, pour chaque loi usuelle, de simuler des échantillons, et de calculer la moyenne et la variance de ces échantillons, ainsi que faire des graphiques représentant ces lois. Il y a quatre commandes à connaître pour chaque loi : **rmaloi** : permet de simuler selon loi, **dmaloi** : permet de calculer la densité de maloi, **pmaloi** : permet de calculer la fonction de répartition de maloi, **qmaloi** : permet de calculer le quantile de maloi.

0.1 Exemple pour la loi normale

Voir Code Github [2] : Section 2.1 Introduction



(a) Histogramme de ech avec courbe de densité (b) Graphique de la fonction de répartition

FIGURE 1 – Graphiques pour la loi normale

	X5.	X10.	X25.	X50.	X75.	X90.	X95.
quantile.echantillon	-3.41902	-2.48670	-1.65578	-0.49161	0.34339	1.22235	2.01371
quantile.normale	-3.28970	-2.56310	-1.34897	0.0	1.34897	2.56310	3.28970

TABLE 1 – Tableau de comparaison des quantiles de la loi normale

0.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille 30 selon une loi normale d'espérance 0 et d'écart-type 2.

L'histogramme (Figure 1a) montre une divergence entre la densité théorique et la simulation, due à la petite taille de l'échantillon. Cette variabilité est également visible sur le graphique de la fonction de répartition (Figure 1b) et dans le tableau de comparaison des quantiles (Table 1).

En conclusion, un échantillon plus grand est nécessaire pour une meilleure concordance avec la loi normale, conformément à la **loi des grands nombres** [3].

1 Lois discrètes

1.1 Loi uniforme discrète

Soit X une variable aléatoire distribuée selon une loi uniforme discrète sur $\{1, \dots, n\}$ (voir [4]). La distribution est définie comme suit :

$$\forall k \in \{1, \dots, n\}, p(X = k) = \frac{1}{n} \Leftrightarrow X \rightarrow \mathcal{U}(\{1, \dots, n\}) \quad (1)$$

1.1.1 Exemple - Simulation de la loi uniforme discrète

Voir Code Github [2] : 2.2.1 Loi uniforme discrète

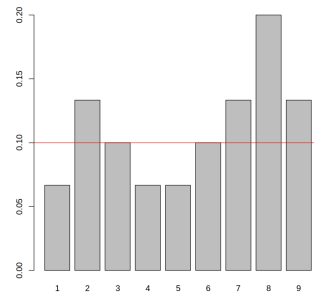


FIGURE 2 – Histogramme de simulation par loi uniforme discrète avec sa droite de probabilité théorique

1.1.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille 30 selon une loi uniforme discrète $\mathcal{U}(\{1, \dots, 10\})$ (1). Théoriquement, chaque événement est équiprobable, donc :

$$\forall k \in \{1, \dots, 10\}, p(X = k) = \frac{1}{10}$$

En analysant les résultats de l'histogramme (Figure 3), nous constatons grâce à la droite horizontale rouge représentant la probabilité théorique $p(X) = \frac{1}{10}$, que les fréquences observées varient et ne sont pas toutes égales. Par exemple,

$$p(X = 2) = 0.03 \quad p(X = 4) = 0.23$$

Cette variation est encore une fois due à la taille réduite de l'échantillon. Les fréquences observées devraient se rapprocher des probabilités théoriques avec un échantillon plus important, conformément à la loi des grands nombres [3]. En conclusion, un échantillon plus important est nécessaire pour démontrer la validité de la loi uniforme discrète de manière plus précise.

1.2 Loi de Bernoulli

Cette loi est notée $\mathcal{B}(p)$. On dit que $X \rightsquigarrow \mathcal{B}(p)$ si et seulement si X est une variable binaire telle que (voir [5]) :

$$p(X = 1) = p \quad \text{et} \quad p(X = 0) = 1 - p \quad (2)$$

C'est un cas particulier de la loi binomiale avec $n = 1$. On peut vérifier facilement par le calcul que $\mathbb{E}[X] = p$ et $\mathbb{V}[X] = p(1 - p)$.

1.2.1 Exemple - Simulation de la loi de Bernoulli

Voir Code Github [2] : 2.2.2 Loi de Bernoulli

0	1
0.73333	0.26667

TABLE 2 – Fréquences observées de l'échantillon simulé par la loi de Bernoulli $\mathcal{B}(0.3)$

1.2.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille 30 selon une loi de Bernoulli $\mathcal{B}(0.3)$. La loi de Bernoulli est une loi binaire qui ne prend que deux valeurs possibles : 0 ou 1. Dans notre cas, les fréquences théoriques sont donc :

$$p(\text{Échec} = 0) = 0.7 \quad p(\text{Succès} = 1) = 0.3$$

En observant les fréquences (Table 2), la fréquence de succès est de 0.27 et celle d'échec est de 0.73. Ces fréquences sont proches des probabilités théoriques, avec une légère variation due à la taille de l'échantillon. Un échantillon plus grand réduirait cette variation, conformément à la loi des grands nombres [3].

1.3 Loi binomiale

On dit que $X \rightsquigarrow \mathcal{B}(n, p)$ si X est la somme de n variables de Bernoulli indépendantes de paramètre p . X représente le nombre de succès parmi n essais indépendants (voir [6]).

Pour résumer, soit $Y \rightsquigarrow \mathcal{B}(p)$ pour $i = 1, \dots, n$ où les Y_i sont indépendantes. On note :

$$X = \sum_{i=1}^n Y_i \rightsquigarrow \mathcal{B}(n, p) \quad (3)$$

Les valeurs possibles pour la variable X sont $\{0, \dots, n\}$ et

$$\forall k \in \{0, \dots, n\}, p_k = p(X = k) = C_n^k p^k (1-p)^{n-k} \quad (4)$$

1.3.1 Exemple - Simulation de la loi de binomiale

Voir Code Github [2] : 2.2.3 Loi de Binomiale

5	6	7	8	9	10
0.02	0.06	0.30	0.36	0.16	0.09

TABLE 3 – Tableau partiel [5, 10] des fréquences observées de la simulation binomiale de $p(\text{Succès}) = 0.8$ et $N = 100$ observations

0	1	2	3	4	6
0.06	0.29	0.37	0.22	0.05	0.01

TABLE 4 – Tableau partiel [0, 6] des fréquences observées de la simulation binomiale de $p(\text{Succès}) = 0.2$ et $N = 100$ observations

5	6	7	8	9	10
0.026	0.088	0.201	0.301	0.268	0.107

TABLE 5 – Tableau partiel [5, 10] des fréquences théoriques de la loi binomiale $\mathcal{B}(10, 0.8)$, arrondies à 3 décimales

0	1	2	3	4	6
0.107	0.268	0.302	0.201	0.088	0.026

TABLE 6 – Tableau partiel [0, 6] des fréquences théoriques de la loi binomiale $\mathcal{B}(10, 0.2)$, arrondies à 3 décimales

1.3.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille 100 selon deux lois binomiales : $\mathcal{B}(10, 0.8)$ et $\mathcal{B}(10, 0.2)$. Pour chaque loi, nous avons observé le nombre de succès sur 10 essais répétés 100 fois de manière indépendante.

On peut s'apercevoir que pour la loi $\mathcal{B}(10, 0.8)$, la fréquence la plus élevée est $p(X = 8) = 0.36$ et pour la loi $\mathcal{B}(10, 0.2)$, la fréquence la plus élevée est $p(X = 2) = 0.37$. Ce résultat s'explique par la propriété d'espérance (moyenne pondérée des valeurs possibles) de la loi binomiale : $\mathbb{E}[X] = np$.

— Pour la loi $\mathcal{B}(10, 0.8)$, on s'attend à avoir $\mathbb{E}[X] = np = 10 * 0.8 = 8$ succès en moyenne.

— Pour la loi $\mathcal{B}(10, 0.2)$, on s'attend à avoir $\mathbb{E}[X] = np = 10 * 0.2 = 2$ succès en moyenne.

En comparant les fréquences observées (Tableaux 3 et 4) avec les fréquences théoriques (Tableaux 5 et 6), nous remarquons une bonne concordance, bien que des variations existent en raison de la taille du nombre de répétitions.

Pour vérifier cela, si nous augmentons la taille de l'échantillon à $N = 10000$, les fréquences observées se rapprochent encore plus des fréquences théoriques :

— Pour la loi $\mathcal{B}(10, 0.8)$: $p(X = 8) = 0.3026$

— Pour la loi $\mathcal{B}(10, 0.2)$: $p(X = 2) = 0.3033$

1.4 Loi Géométrique

On dit que X suit une loi géométrique $\mathcal{G}(p)$ avec $0 < p < 1$. On répète, de façon indépendante, une même expérience (qui a une probabilité p de réussite et une probabilité $q = 1 - p$ d'échec) et on note X le rang du premier succès (voir [7]). Les valeurs prises par X sont donc $X(\Omega) = \mathbb{N}^*$ et les probabilités associés sont :

$$\forall k \in \mathbb{N}^*, p_k = P(X = k) = (1 - p)^{k-1}p = q^{k-1}p \quad (5)$$

1.4.1 Exemple - Simulation de la loi géométrique

Voir Code Github [2] : 2.3 Loi Géométrique

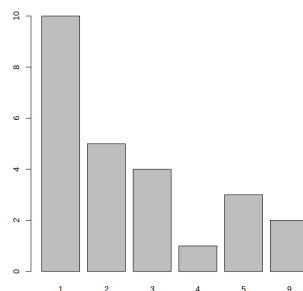


FIGURE 3 – Histogramme de simulation par loi géométrique avec une probabilité $p(X = \text{Succès}) = 0.3$ et une taille d'échantillon de 25

1.4.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille 25 selon une loi géométrique $\mathcal{G}(0.3)$. Comme nous pouvons le remarquer dans l'histogramme (Figure 3), les fréquences observées sont décroissantes, ce qui est conforme à la loi géométrique (5). Par exemple :

$$\begin{aligned} p(X = 1) &= \frac{10}{25} = 0.4 \approx (1 - 0.3)^{1-1}0.3 = 0.3 \\ p(X = 2) &= \frac{5}{25} = 0.2 \approx (1 - 0.3)^{2-1}0.3 = 0.21 \end{aligned}$$

1.5 Loi de Poisson

On dit que X suit une loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda > 0$. On peut voir cette loi comme la loi des événements rares. Cette loi est une approximation de la loi binomiale, $\mathcal{B}(n, p)$ lorsque n est grand et p petit. On pose alors $\lambda = np$. En pratique, on dit que cette approximation est bonne dès que $n > 20$ et $p \leq 0.1$ et $np \leq 5$ (voir [8]).

Les valeurs possibles pour X sont dans N et les probabilités associées :

$$\forall k \in \mathbb{N}, p_k = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (6)$$

1.5.1 Exemple - Simulation de la loi Poisson

Voir Code Github [2] : 2.4 Loi Poisson

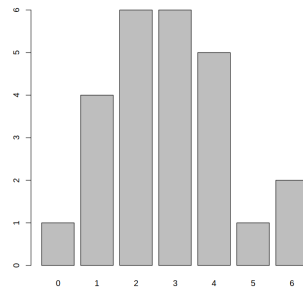


FIGURE 4 – Histogramme de simulation par loi poisson avec un paramètre $\lambda = 3$ et une taille d'échantillon de 25

1.5.2 Analyse des résultats

Dans cet exemple, nous avons simulé un échantillon de taille $n = 25$ selon une loi de Poisson $\mathcal{P}(3)$. Ce qui signifie $\lambda = 3 = np$ donc cet exemple à une probabilité de succès de $p = 0.12$. On se trouve bien dans le cas où $n > 20$ et $p \leq 0.1$ et $np \leq 5$.

En observant l'histogramme (Figure 4), nous constatons que les fréquences observées sont conformes à la loi de Poisson (6). Par exemple :

$$\begin{aligned} p(X = 0) &= \frac{1}{25} = 0.04 \approx e^{-3} \frac{3^0}{0!} = 0.049 \\ p(X = 1) &= \frac{4}{25} = 0.16 \approx e^{-3} \frac{3^1}{1!} = 0.149 \\ p(X = 2) &= \frac{6}{25} = 0.24 \approx e^{-3} \frac{3^2}{2!} = 0.224 \end{aligned}$$

Nous pouvons observer grâce à la Figure 5 ci-dessous que la loi Poisson converge est une approximation de la loi binomiale pour n grand et p petit.

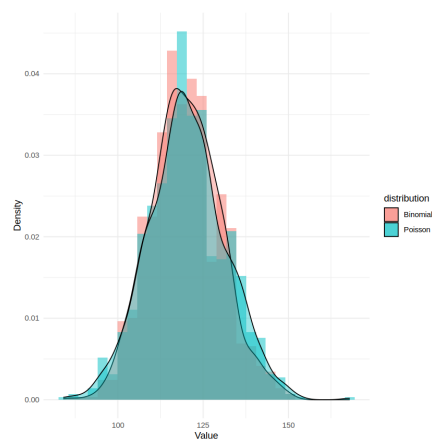
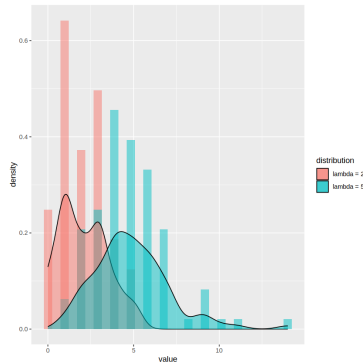


FIGURE 5 – Comparaison de la loi de Poisson et de la loi binomiale pour $n = 1000$ et $p = 0.12$

Lorsque l'on fait varier le paramètre λ de la loi de Poisson, on peut observer que la distribution se déplace vers la droite (Figure 6). Cela s'explique par le fait que λ est l'espérance de la loi de Poisson, dans le tableau ci-dessous Table 7, on peut voir que pour $\lambda = 2$ la médiane et la moyenne sont égales à 2 (ou très proche en fonction de la taille de l'échantillon n) Alors que pour $\lambda = 5$, la moyenne est de 5 et la médiane est de 5 (**Voir code annexe 1 : Code : Variation des paramètres de la loi Poisson**)



λ	n	Moyenne	Médiane
2	100	2.05	2.00
5	100	4.87	5.00

TABLE 7 – Comparaison des paramètres de la loi de Poisson pour $\lambda = 2$ et $\lambda = 5$

FIGURE 6 – Comparaison de la loi de Poisson selon le paramètre λ

2 Lois continues

2.1 Loi Uniforme sur $[0, 1]$, notée $\mathcal{U}([0, 1])$

On dit que X suit une loi uniforme sur $[0, 1]$ ssi sa fonction densité est :

$$f_X(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases} \quad (7)$$

La fonction de répartition d'une loi uniforme sur $[0, 1]$ est :

$$\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0, 1] \\ 1 & \text{si } x > 1 \end{cases} \quad (8)$$

L'espérance de la loi uniforme $\mathcal{U}_{[0,1]}$ est $\mathbb{E}[X] = \frac{1}{2}$. Le moment d'ordre 2 est $\mathbb{E}[X^2] = \frac{1}{3}$ et la variance est $\mathbb{V}[X] = \frac{1}{12}$ (voir [9]).

2.1.1 Exemple - Simulation de la loi uniforme sur $[0, 1]$

Voir code GitHub [2] : Section 3.1 Loi uniforme sur $[0, 1]$

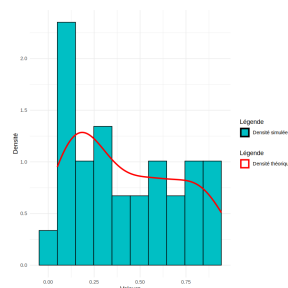


FIGURE 7 – Histogramme de simulation par loi uniforme sur $[0, 1]$ avec sa courbe de densité

2.1.2 Analyse des résultats

La figure 7 montre un histogramme de simulation S de la loi uniforme sur $[0, 1]$ avec sa courbe de densité. L'histogramme montre que les valeurs générées tendent à être réparties uniformément sur l'intervalle $[0, 1]$ et que sa courbe de densité tend à être constante et égale à 1 (voir 7).

La moyenne des valeurs par la probabilité d'apparition de cette simulation est : $\mathbb{E}[S] = 0.426$ et la variance est : $\mathbb{V}[S] = 0.083$. D'après les valeurs théoriques, on a $\mathbb{E}[X] = \frac{1}{2} \approx \mathbb{E}[S]$ et $\mathbb{V}[X] = \frac{1}{12} \approx \mathbb{V}[S]$. On peut donc conclure que les valeurs simulées sont proches des valeurs théoriques.

2.2 Loi Uniforme sur $[a, b]$, notée $\mathbb{U}([a, b])$ avec $a < b$

On dit que X suit une loi uniforme sur $[a, b]$ ssi sa fonction densité est :

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases} \quad (9)$$

La fonction de répartition d'une loi uniforme sur $[a, b]$ est :

$$\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases} \quad (10)$$

L'espérance de la loi uniforme $U_{[a,b]}$ est $\mathbb{E}[X] = \frac{a+b}{2}$. Le moment d'ordre 2 est $\mathbb{E}[X^2] = \frac{a^2+ab+b^2}{3}$ et la variance est $\mathbb{V}[X] = \frac{(b-a)^2}{12}$ (voir [9]).

2.2.1 Exemple - Simulation de la loi uniforme sur $[a, b]$

Voir code GitHub [2] : Section 3.2 Loi uniforme sur $[a, b]$

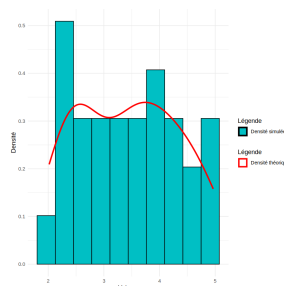


FIGURE 8 – Histogramme de simulation par loi uniforme sur $[2, 5]$ avec sa courbe de densité

2.2.2 Analyse des résultats

La figure 11 montre un histogramme de simulation S de la loi uniforme sur $[2, 5]$ avec sa courbe de densité. L'histogramme montre que les valeurs générées tendent à être réparties uniformément sur l'intervalle $[2, 5]$ et que sa courbe de densité tend à être constante et égale à $\frac{1}{3}$ car $\frac{1}{b-a} = \frac{1}{5-2}$ (voir équation 9).

La moyenne des valeurs par la probabilité d'apparition de cette simulation est : $\mathbb{E}[S] = 3.421$ et la variance est : $\mathbb{V}[S] = 0.805$. D'après les valeurs théoriques, on a

$$\mathbb{E}[X] = \frac{a+b}{2} = \frac{5+2}{2} = 3.5 \approx \mathbb{E}[S]$$

$$\mathbb{V}[X] = \frac{(b-a)^2}{12} = \frac{(5-2)^2}{12} = 0.75 \approx \mathbb{V}[S]$$

On peut donc conclure que les valeurs simulées sont proches des valeurs théoriques.

2.3 Loi Normale centrée réduite, notée $\mathcal{N}(0, 1)$

Si $X \rightsquigarrow \mathcal{N}(0, 1)$ alors la fonction densité de X est :

$$\forall x \in \mathbb{R}, f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (11)$$

L'espérance de la loi normale centrée réduite est $\mathbb{E}[X] = 0$ et la variance est $\mathbb{V}[X] = 1$ (voir [10]).

2.3.1 Exemple - Simulation de la loi normale centrée réduite $\mathcal{N}(0, 1)$

Voir code GitHub [2] : Section 3.3 Loi Normale centrée réduite notée $\mathcal{N}(0, 1)$

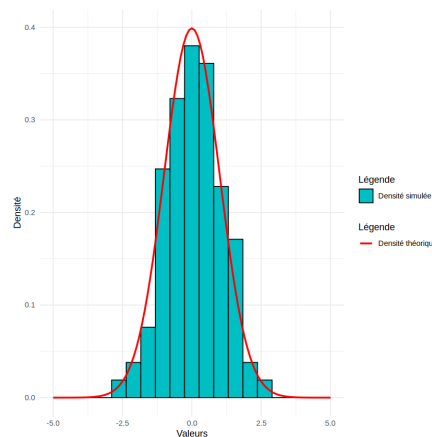


FIGURE 9 – Histogramme de simulation par loi normale centrée réduite $\mathcal{N}(0, 1)$ avec la courbe de densité théorique

2.3.2 Analyse des résultats

La figure 9 illustre la distribution d'une loi normale centrée réduite, laquelle est symétrique par rapport à $x = 0$, son point de densité maximale, et décroît de part et d'autre. En analysant la moyenne et la variance de la simulation S représenté sous forme d'histogramme, nous obtenons :

$$\mathbb{E}[S] = 0.07 \approx \mathbb{E}[X] = 0$$

$$\mathbb{V}[S] = 1.02 \approx \mathbb{V}[X] = 1$$

La symétrie autour de $x = 0$ s'explique par le paramètre $\mu = 0$ (moyenne) de la distribution. L'variance, quant à lui, détermine la dispersion des valeurs autour de la moyenne, influençant ainsi la forme de la courbe de densité.

2.3.3 Démonstration

La proposition de départ est :

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi} \quad (12)$$

1. Nous savons que $\sqrt{2\pi} = 2.506628$. Il est possible de simuler cette equation sur R (voir code annexe 2). En utilisant avec la fonction *interval* nous obtenons : 2.506628 avec une erreur absolue < 0.00023 . Ce qui confirme la proposition de départ.
2. Nous pouvons également comparer graphiquement les 2 equations, voir figure 10 et nous remarquons que les 2 courbes se superposent parfaitement.

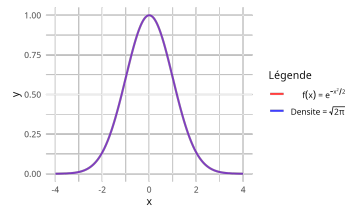


FIGURE 10 – Comparaison graphique des 2 équations (Voir Code Github : section 3.4 [2])

2.4 Loi Normale, notée $\mathcal{N}(\mu, \sigma^2)$

Si $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$ alors la fonction densité de X est :

$$\forall x \in \mathbb{R}, f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (13)$$

2.4.1 Exemple - Simulation de la loi normale centrée réduite $\mathcal{N}(1, 4)$

Voir code TP : Section Loi Normale $\mathcal{N}(\mu, \sigma^2)$ [1]

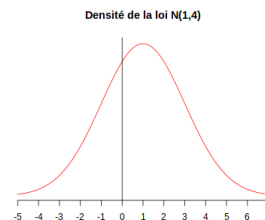


FIGURE 11 – Histogramme de simulation par loi normale centrée réduite $\mathcal{N}(1, 4)$

2.4.2 Analyse des résultats

Le graphique de la figure 11 suit une loi normale centrée réduite, $\mathcal{N}(1, 4)$ laquelle est symétrique par rapport à $x = 1 = \mu$, son point de densité maximale. Et $\sigma = 4$ détermine la dispersion des valeurs autour de la moyenne, influençant ainsi la forme de la courbe de densité. Elle paraît donc plus aplatie.

2.5 Théorème

$$\text{Si } X \rightsquigarrow \mathcal{N}(\mu, \sigma^2) \text{ alors } \frac{X - \mu}{\sigma} \rightsquigarrow \mathcal{N}(0, 1) \quad (14)$$

Pour vérifier le théorème, il faut tracer un graphique (voir figure 12) de la loi normale centrée réduite $\mathcal{N}(0, 1)$ et de la loi normale $\mathcal{N}(1, 4)$ centrée en $\mu = 1$ et d variance $\sigma^2 = 4$. Dans la figure ci dessous, les 2 courbes de densité tendent à se superposer pour un échantillon de taille $n = 2000$. (voir code annexe 3)

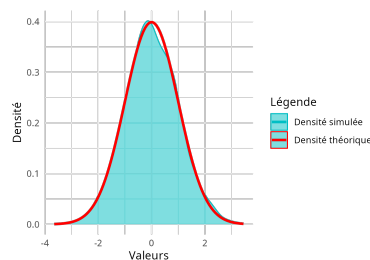


FIGURE 12 – Comparaison des courbes de densité de la loi normale centrée réduite $\mathcal{N}(0, 1)$ et de la loi normale $\mathcal{N}(1, 4)$ centrée en $\mu = 1$ et d variance $\sigma^2 = 4$

2.6 Exercices sur la loi normale

Pour simuler n valeurs aléatoirement suivant la loi normale de moyenne m et d'écart-type s on utilise sous R la commande **rnorm**. Nous avons effectués 5 simulations de la loi normale d'une taille $n = 20$ et de paramètres : $\mathcal{N}(0, 1)$, $\mathcal{N}(1, 1)$, $\mathcal{N}(2, 1)$, $\mathcal{N}(0, 2)$, $\mathcal{N}(0, 4)$. Voici ci-dessous le tableau résultant de ces simulations :

$n = 20$	$\mathcal{N}(0, 1)$	$\mathcal{N}(1, 1)$	$\mathcal{N}(2, 1)$	$\mathcal{N}(0, 2)$	$\mathcal{N}(0, 4)$
nb de valeurs < -2	1	0	0	4	5
nb de valeurs < 0	12	3	0	11	10
nb de valeurs $= 0$	0	0	0	0	0
nb de valeurs > 2	1	1	11	3	5

TABLE 8 – Tableau de simulation des 5 lois normales de taille $n = 20$

2.6.1 Question 1 : Voir Code TP, section : Exercices sur la loi normale, Question 1 [1]

Avec les mêmes valeurs des moyennes et des variances, on simule maintenant des échantillons de taille $n=2000$ individus :

$n = 20$	$\mathcal{N}(0, 1)$	$\mathcal{N}(1, 1)$	$\mathcal{N}(2, 1)$	$\mathcal{N}(0, 2)$	$\mathcal{N}(0, 4)$
nb de valeurs < -2	48	4	0	326	595
nb de valeurs < 0	1006	308	48	997	978
nb de valeurs $= 0$	0	0	0	0	0
nb de valeurs > 2	55	336	985	319	624

TABLE 9 – Tableau de simulation des 5 lois normales de taille $n = 2000$ (voir code annexe 4)

2.6.2 Question 2 : Voir Code TP, section : Exercices sur la loi normale, Question 2 [1]

On considère uniquement le premier échantillon $nd1$. Que trace la fonction `plot` ?

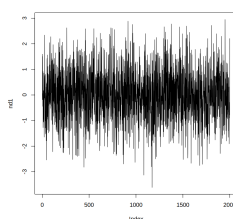


FIGURE 13 – Graphique en courbe de la loi normale $\mathcal{N}(0, 1)$

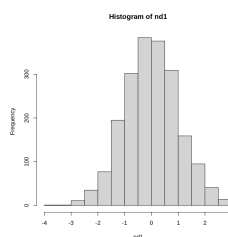


FIGURE 14 – Histogramme de la loi normale $\mathcal{N}(0, 1)$

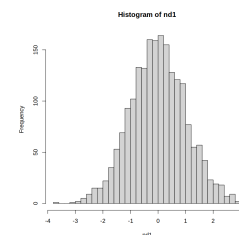


FIGURE 15 – Histogramme de la loi normale $\mathcal{N}(0, 1)$ avec 20 classes

La figure 22 représente un graphique obtenu à l'aide de la fonction **plot**. Ce graphique trace les valeurs de l'échantillon $nd1$ en fonction de leur indice. L'axe des abscisses (x) correspond aux indices des valeurs de l'échantillon, variant de 1 à $n = 2000$, tandis que l'axe des ordonnées (y) représente les valeurs de l'échantillon, lesquelles suivent une distribution normale.

Ensuite, les graphiques des figures 23 et 24 sont tout deux obtenus à l'aide de la fonction **hist**. Mais nous avons choisi manuellement le nombre de classes pour chaque histogramme pour la figure 24 grâce à la variable **breaks** = 20. Alors que pour le graphique de la figure 23, nous avons laissé R choisir automatiquement le nombre de classes.

2.6.3 Question 3 : Voir Code TP, section : Exercices sur la loi normale, Question 3 [1]

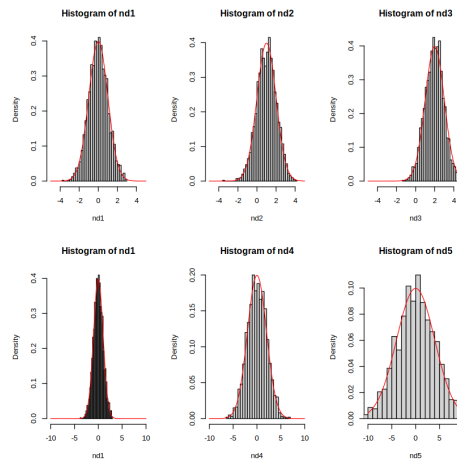


FIGURE 16 – Histogrammes des simulations avec leurs courbes de densité théorique des échantillons : $\mathcal{N}(0, 1)$, $\mathcal{N}(1, 1)$ et $\mathcal{N}(2, 1)$. Puis des échantillons $\mathcal{N}(0, 1)$, $\mathcal{N}(0, 2)$ et $\mathcal{N}(0, 4)$

Comparer les histogrammes des échantillons nd1, nd2 et nd3 : qu'est-ce qui change ?

En examinant les trois premiers histogrammes de la figure 17, on constate qu'ils suivent une distribution normale avec des variances identiques mais des moyennes différentes. Cela se traduit graphiquement par des courbes de densité centrées et symétriques autour de $x = 0$ pour $\mathcal{N}(0, 1)$, $x = 1$ pour $\mathcal{N}(1, 1)$, et $x = 2$ pour $\mathcal{N}(2, 1)$. Les histogrammes de simulation se déplacent donc également vers la droite, en suivant la moyenne de chaque distribution.

Comparer les histogrammes des échantillons nd1, nd4 et nd5 : qu'est-ce qui change ?

On remarque cette fois-ci pour les 3 derniers histogrammes de la figure 17 que les courbes de densité sont centrées autour de $x = 0$ mais avec des variances différents. Graphiquement, cela se traduit par des courbes de densité plus ou moins étalées, plus ou moins aplaties, en fonction de l'variance de chaque distribution. Les histogrammes de simulation suivent donc la dispersion des valeurs autour de la moyenne.

2.6.4 Question 4 : Voir Code TP, section : Exercices sur la loi normale, Question 4 [1]

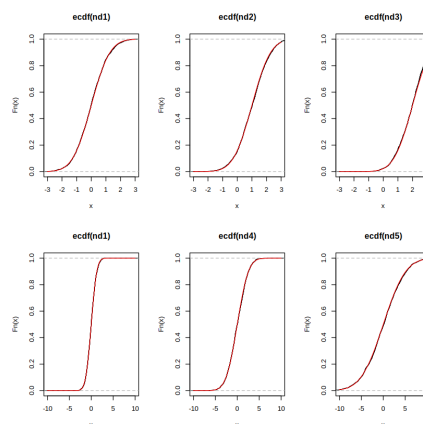


FIGURE 17 – Fonction de répartition simulées avec leur courbe théorique des échantillons : $\mathcal{N}(0, 1)$, $\mathcal{N}(1, 1)$ et $\mathcal{N}(2, 1)$. Puis des échantillons $\mathcal{N}(0, 1)$, $\mathcal{N}(0, 2)$ et $\mathcal{N}(0, 4)$

Comparer les fonctions de répartitions des échantillons nd1, nd2 et nd3 : qu'est-ce qui change ?

En examinant les 3 premiers graphiques, nous observons que la forme de chaque courbe reste identique car l'écart type est identique. Mais leurs positions sur l'axe horizontal x changent en fonction de la moyenne : plus la moyenne est élevée, plus la courbe est décalée vers la droite. De plus, elles sont symétriques par rapport à l'axe des ordonnées de leur moyenne respective.

Comparer les fonctions de répartitions des échantillons nd1, nd4 et nd5 : qu'est-ce qui change ?

Pour des échantillons de même moyenne mais d'écart types différents, les courbes sont toutes symétriques par rapport à l'axe des ordonnées de leur moyenne respective : $x = 0$. Mais les courbes sont plus ou moins étalées, plus ou moins aplaties, en fonction de l'écart type de chaque distribution.

2.6.5 Question 5 : Voir Code TP, section : Exercices sur la loi normale, Question 5 [1]

Donner ci-dessous la densité de la loi normale de moyenne 0 et de variance 1 :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Quel est le lien entre la densité et la fonction de répartition ?

La fonction de répartition $F(x)$ est la primitive de la densité $f(x)$, c'est-à-dire l'intégrale de $f(x)$ de $-\infty$ à t . Elle permet de calculer la probabilité que la variable aléatoire $X \leq t$. (voir [11])

$$F(t) = \int_{-\infty}^t f(x) dx$$

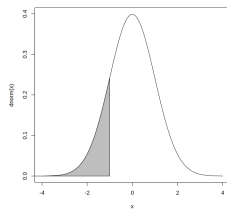


FIGURE 18 – Courbe de densité de $\mathcal{N}(0,1)$ avec la zone de probabilité $P(X \leq -1)$

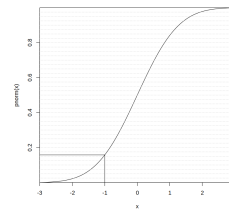


FIGURE 19 – Fonction de répartition de la loi normale $\mathcal{N}(0,1)$ avec intersection en $x = -1$

Les graphiques ci-dessus permettent de visualiser la probabilité $P(X \leq -1) = 0.1587$ pour la loi normale $\mathcal{N}(0,1)$. Dans le graphique de la figure 18, la zone coloriée en gris représente l'aire sous la courbe de densité de la loi normale centrée réduite $\mathcal{N}(0,1)$ pour $x \leq -1$ qui est égal à 0.1587. La figure 19 illustre la fonction de répartition de $\mathcal{N}(0,1)$, laquelle permet en lisant l'intersection de la courbe de répartition en $x = -1$: $P(X \leq -1) = 0.1587$

2.6.6 Question 6

Aire de la figure	Pr représentée par l'aire	Ecriture avec F(t)	Valeur de pnorm
1	$P(X \leq -1)$	$F(-1)$	0.308
2	$P(X \leq 2)$	$F(2)$	0.841
3	$P(X \geq 2)$	$1 - F(2)$	0.159
4	$P(-1 \leq X \leq 2)$	$F(2) - F(-1)$	0.532

TABLE 10 – Tableau de comparaison des valeurs de probabilités entre les fonctions de répartition et la fonction pnorm pour une loi normale $\mathcal{N}(0,4)$, voir code annexe 5

Soit X une variable aléatoire gaussienne de moyenne 15 et de variance 9. Quelle est la $Pr(X \leq 17)$?

La fonction **`pnorm(17,15,3)`** de paramètres 17 (x), 15 (moyenne) et 3 (écart-type) permet de calculer la probabilité $Pr(X \leq 17) = 0.747$

Quelle est $Pr(X > 7)$?

La fonction **`1-pnorm(7,15,3)`** permet de calculer $Pr(X > 7) = 0.996$

Enfin quelle est la $Pr(7 < X < 15)$?

La fonction **`pnorm(15,15,3)-pnorm(7,15,3)`** permet de calculer $Pr(7 < X < 15) = 0.496$

2.7 Lien entre les lois normales

Soit X une variable aléatoire distribuée suivant $\mathcal{N}(0, 1)$. On souhaite montrer que les variables $X_1 = X + 1$, $X_2 = X - 1$, $X_3 = 2X$ et $X_4 = X/2$ suivent aussi une loi normale. Pour cela on simule un grand nombre de valeurs suivant la loi normale $\mathcal{N}(0,1)$ et l'on construit les nouvelles variables.

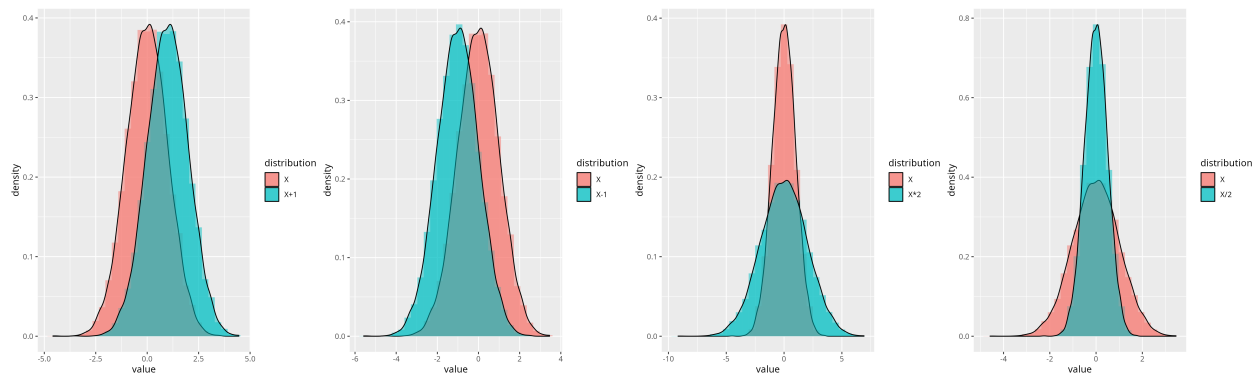


FIGURE 20 – Comparaison des lois normales $\mathcal{N}(0, 1)$ avec les variables $X_1 = X + 1$, $X_2 = X - 1$, $X_3 = 2X$ et $X_4 = X/2$ suivant une loi normale

Le graphique de la figure 21 démontre que les variables $X_1 = X + 1$, $X_2 = X - 1$, $X_3 = 2X$ et $X_4 = X/2$ suivent également une loi normale (**voir code annexe 6**). Plus précisément, nous observons que X_1 suit la loi $\mathcal{N}(1, 1)$, X_2 la loi $\mathcal{N}(-1, 1)$, X_3 la loi $\mathcal{N}(0, 2)$ et X_4 la loi $\mathcal{N}(0, 0.5)$. Ces résultats confirment ainsi que toute combinaison affine d'une variable normalement distribuée demeure normale, tout en modifiant la moyenne et la variance selon les opérations effectuées.

Soit maintenant Y une variable aléatoire de moyenne 4 et de variance 3, quelle est la distribution de la variable $Z = \frac{Y-2}{\sqrt{3}}$?

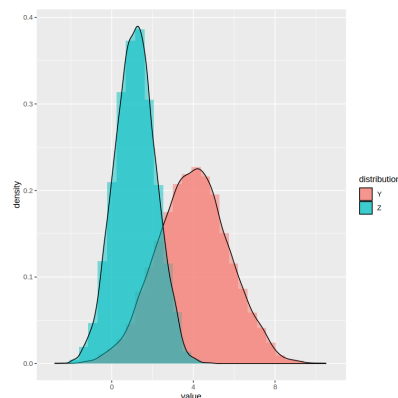


FIGURE 21 – Comparaison des lois normales $\mathcal{N}(4, 3)$ avec la variable $Z = \frac{Y-2}{\sqrt{3}}$ suivant une loi normale

Nous remarquons que Z suit une loi normale $\mathcal{N}(1, 1)$. A noter que si Z respectait le théorème de la loi centrée réduite $Z^* = \frac{Y - \mu}{\sigma} = \frac{Y - 4}{\sqrt{3}}$ alors Z^* suivrait une loi normale $\mathcal{N}(0, 1)$.

2.8 Sommes de variables aléatoires gaussiennes

Donner la loi de $X + Y$ et de $X - Y$ si $X \rightsquigarrow N(m_1, \sigma_1^2)$ et $Y \rightsquigarrow N(m_2, \sigma_2^2)$

$$X + Y \rightsquigarrow N(m_1 + m_2, \sigma_1^2 + \sigma_2^2) \quad \text{et} \quad X - Y \rightsquigarrow N(m_1 - m_2, \sigma_1^2 + \sigma_2^2) \quad (15)$$

Voir code annexe 7 : Code : Sommes de variables aléatoires gaussiennes et Code TP, section : Sommes de variables aléatoires gaussiennes [1]

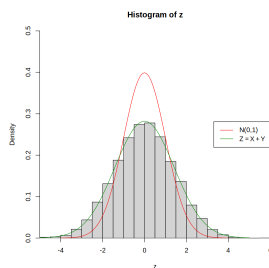


FIGURE 22 – Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$

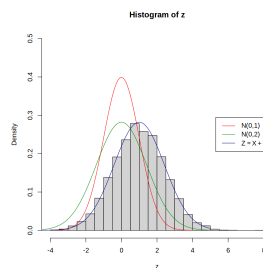


FIGURE 23 – Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$ et $\mathcal{N}(1, 1)$

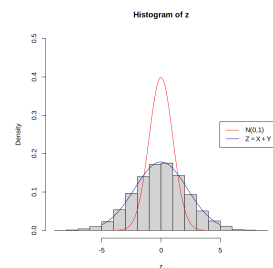


FIGURE 24 – Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$ et $\mathcal{N}(0, 2)$

Déduire de ce qui précède la loi de la variable $\bar{X} = \frac{1}{n}(X_1 + X_2 + X_3 + \dots + X_n)$ où toutes les variables X_i sont gaussiennes centrées réduites.

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n) \rightsquigarrow N\left(\frac{1}{n} \sum_{i=1}^n m_i, \frac{1}{n} \sum_{i=1}^n \sigma_i^2\right) \quad (16)$$

Voir code annexe 8 : Code : Démonstration de la loi d'une moyenne de lois gaussienne \bar{X}

2.9 Somme de carrés de variables gaussiennes centrées réduites

Trouver le nom de la nouvelle loi des carrés et comment on choisit ses paramètres ?

$$X_1^2 + X_2^2 + X_3^2 + \dots + X_n^2 \rightsquigarrow \chi^2(n) \quad (17)$$

Sur R, il faut utiliser la fonction **dchisq(n)** pour avoir la distribution de la loi. Elle prend en paramètre x la valeur à évaluer et n le nombre de degrés de liberté (dans notre cas c'est l'index n de l'équation 17).

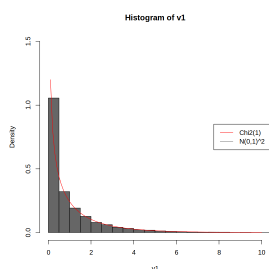


FIGURE 25 – Histogramme de la somme de carrés de 2 lois normales $\mathcal{N}(0, 1)$

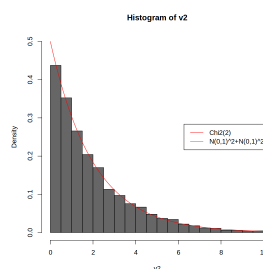


FIGURE 26 – Histogramme de la somme de carrés de 3 lois normales $\mathcal{N}(0, 1)$

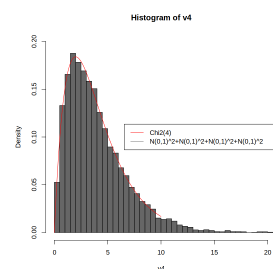


FIGURE 27 – Histogramme de la somme de carrés de 4 lois normales $\mathcal{N}(0, 1)$

2.10 Exercice

Voir code annexe 9 : Code : Exercice sur la loi normale

La taille X d'un homme français tiré au sort dans la population est supposée suivre une loi normale de moyenne 172cm et de variance 196cm^2 .

Question 1 Quelle est la probabilité que X soit inférieure à 160cm ?

Pour calculer la probabilité que X soit inférieure à 160 cm , nous devons calculer $P(X < 160) = F(160)$ où F est la fonction de répartition de la loi normale $\mathcal{N}(172, 196)$.

$$P(X < 160) = F(160) = 0.196$$

Question 2 Quelle est la probabilité qu'un homme tiré au sort mesure plus de deux mètres ?

Pour calculer la probabilité qu'un homme tiré au sort mesure plus de deux mètres, nous devons calculer $P(X > 200) = 1 - F(200)$ où F est la fonction de répartition de la loi normale $\mathcal{N}(172, 196)$.

$$P(X > 200) = 1 - F(200) = 0.023$$

Question 3 Que vaut $P(165 < X < 185)$?

Pour calculer la probabilité que X soit comprise entre 165 et 185 cm , nous devons calculer $P(165 < X < 185) = F(185) - F(165)$ où F est la fonction de répartition de la loi normale $\mathcal{N}(172, 196)$.

$$P(165 < X < 185) = F(185) - F(165) = 0.515$$

Question 4 La taille des femmes suit une loi gaussienne de moyenne 162cm et de variance 144cm^2 . Quelle est la probabilité qu'un homme tiré au sort soit plus grand qu'une femme choisie au hasard

Pour calculer la probabilité qu'un homme tiré au sort soit plus grand qu'une femme choisie au hasard, nous devons calculer on définit Z la variable aléatoire de la différence de taille entre un homme et une femme (voir équation de la somme des lois gaussiennes 15). X variable aléatoire de la taille d'un homme et Y variable aléatoire de la taille d'une femme.

$$Z = X - Y \rightsquigarrow \mathcal{N}(172 - 162, 196 + 144) = \mathcal{N}(10, 340)$$

$$P(X > Y) = P(X - Y > 0) = 1 - F(0) = 0.706$$

2.11 Loi Exponentielle

On dit que X suit une loi exponentielle de paramètre $\lambda > 0$, ce que l'on note $X \rightsquigarrow \mathcal{E}(\lambda)$ ssi sa fonction densité est :

$$\forall x \in \mathbb{R}, f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (18)$$

Vérifier que la fonction f est bien une densité de probabilité

Pour vérifier que la fonction f est bien une densité de probabilité, il faut vérifier que l'intégrale de f sur \mathbb{R} est égale à 1. La fonction f doit également être positive sur \mathbb{R} et continue.

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^{+\infty} \lambda \exp(-\lambda x) dx = 1$$

Voir code annexe 10 : Code : Vérification de la densité de probabilité de la loi exponentielle

Une simulation S avec un échantillon $n = 30$ et $\lambda = 0.001$ donnerait un histogramme avec une distribution exponentielle décroissante avec une espérance de $\mathbb{E}(S) \approx \frac{1}{0.001} \approx 1000$ et une variance de $\mathbb{V}(S) \approx \frac{1}{0.001^2} \approx 10^6$. Sa fonction de répartition est donnée par $F(x) = 1 - \exp(-\lambda x)$ et sera quant à elle une fonction exponentielle croissante. (voir code TP, section : Loi Exponentielle [1])

Code annexe

```

1 n <- 100
2 lamda_1 <- 2
3 lamda_2 <- 5
4 poiss <- rpois(n, lambda = lamda_1)
5 poiss2 <- rpois(n, lambda = lamda_2)
6 df <- data.frame(value = c(poiss, poiss2), distribution = c(rep("lambda = 2", n), rep(
  "lambda = 5", n)))
7 ggplot(df, aes(x=value, fill=distribution)) + # Histogramme
8   geom_histogram(aes(y=..density..), position="identity", alpha=0.5, bins=30) +
9   geom_density(alpha=0.5)
10
11 # Simulez differentes lois de Poisson en faisant varier lambda ainsi que la taille de
  l'echantillon
12 lambda <- 30 # paramètre de la loi de Poisson
13 n <- 1000
14 ech <- rpois(n, lambda) # simulation de la loi de Poisson
15 summary(ech)
16 barplot(table(ech))

```

Listing 1 – Code : Variation des paramètres de la loi Poisson

```

1 equation_1 <- integrate(function(x) exp(-x^2/2), -Inf, Inf)
2 equation_2 <- sqrt(2*pi)
3 ggplot(df, aes(x = x)) + # Tracer la fonction densité
4   geom_line(aes(y = y, color = "f1"), alpha = 0.5, size = 1) + # Courbe rouge avec
  transparence
5   geom_line(aes(y = y2, color = "f2"), alpha = 0.5, size = 1) + # Courbe bleue avec
  transparence
6   labs(x = "x", y = "y", color = "Légende")

```

Listing 2 – Code : Vérification de l'égalité

```

1 n <- 2000
2 mu <- 1
3 sigma <- 4
4 ech <- rnorm(n, mu, sigma)
5 ggplot(df, aes(x = (x - mu) / sigma)) +
6   geom_density(aes(color = "Densité simulée"), fill = "#00BFC4", alpha = 0.5) +
7   stat_function(fun = dnorm, aes(color = "Densité théorique"), size = 1.2) +
8   labs(x = "Valeurs", y = "Densité", color = "Légende")

```

Listing 3 – Code : Probabilités des lois normales

```

1 # A suivre du code TP, section : Exercices sur la loi normale, Question 1
2 print(colSums(nd_en_col < -2) / n * 2000)
3 print(colSums(nd_en_col < 0) / n * 2000)
4 print(colSums(nd_en_col == 0) / n * 2000)
5 print(colSums(nd_en_col > 2) / n * 2000)

```

Listing 4 – Code : Probabilités des lois normales

```

1 pnorm(-1, 0, 2) # F(-1) : P(X <= -1)
2 pnorm(2, 0, 2) # F(2) : P(X <= 2)
3 1-pnorm(2, 0, 2) # 1 - F(2) : P(X > 2)
4 pnorm(2, 0, 2) - pnorm(-1, 0, 2) # F(2) - F(-1) : P(-1 < X <= 2)

```

Listing 5 – Code : Probabilités par fonctions de répartition

```

1 y <- rnorm(10000, 4, sqrt(3)) # On simule des variables aléatoires gaussiennes
2 z <- (y - 2) / sqrt(3) # On applique la transformation linéaire
3 # Graphique de comparaison entre y et z

```

```
4 df <- data.frame(value = c(y, z), distribution = c(rep("Y", length(y)), rep("Z",
  length(z))))
5 ggplot(df, aes(x = value, fill = distribution)) +
6   geom_histogram(aes(y = ..density..), position = "identity", alpha = 0.5, bins =
  30) +
7   geom_density(alpha = 0.5)
```

Listing 6 – Code : Lien entre les lois normales

```
1 # Meme ecart type mais pas de meme esperance
2 lines(x, dnorm(x, 1, sqrt(1*1 + 1*1)), col="blue4") #la loi proposée
3
4 # Meme esperance mais pas de meme ecart type
5 lines(x, dnorm(x, 0, sqrt(1*1 + 2*2)), col="blue4") # la loi proposée
```

Listing 7 – Code : Sommes de variables aléatoires gaussiennes

```
1 n <- 10000 # Nombre de simulations
2 x1<-rnorm(n,0,1) # On simule des variables aléatoires gaussiennes
3 x2<-rnorm(n,0,2)
4 x3<-rnorm(n,3,1)
5 x4<-rnorm(n,2,1)
6
7 x_bar <- (x1 + x2 + x3 + x4) / 4 # On calcule la moyenne de ces variables aléatoires
8
9 hypothese <- rnorm(10000, (0+0+3+2)/4, sqrt(1*1 + 2*2 + 1*1 + 1*1)/4) # On simule l'
  hypothese de la loi d'une moyenne de lois gaussienne (voir rapport section 2.8)
10
11 # Comparaison entre la moyenne et une loi normale
12 df <- data.frame(value = c(moyenne, hypothese),
13   distribution = c(rep("Moyenne", length(moyenne)), rep("Hypothese",
  length(hypothese))))
14
15 ggplot(df, aes(x = value, fill = distribution)) +
16   geom_histogram(aes(y = ..density..), position = "identity", alpha = 0.5) +
17   geom_density(alpha = 0.5)
18 # Conclusion: On perçoit bien que les deux densités se superposent parfaitement.
```

Listing 8 – Code : Démonstration de la loi d'une moyenne de lois gaussienne \bar{X}

```
1 moyenneH <- 172 # moyenne de la taille des hommes
2 varianceH <- 196 # variance de la taille des hommes
3 question1 <- pnorm(160, moyenneH, sqrt(varianceH)) # F(160)
4 question2 <- 1 - pnorm(200, moyenneH, sqrt(varianceH)) # 1 - F(200)
5 question3 <- pnorm(185, moyenneH, sqrt(varianceH)) - pnorm(165, moyenneH, sqrt(
  varianceH)) # F(185) - F(165)
6 moyenneF <- 162 # moyenne de la taille des femmes
7 varianceF <- 144 # variance de la taille des femmes
8 question4 <- 1 - pnorm(0, moyenneH - moyenneF, sqrt(varianceH + varianceF)) # 1 - F(0)
```

Listing 9 – Code : Exercice sur la loi normale

```
1 x <- seq(0, 10, 0.1)
2 lambda <- 0.5
3 f <- lambda * exp(-lambda * x)
4 # Tracer la fonction densité avec ggplot
5 df <- data.frame(x = x, f = f)
6 ggplot(df, aes(x = x, y = f)) + geom_line(color = "blue")
7 # vérifier l'intégration à 1
8 integrate(function(x) lambda * exp(-lambda * x), 0, Inf)
```

Listing 10 – Code : Vérification de la densité de probabilité de la loi exponentielle

Table des figures

1	Graphiques pour la loi normale	1
2	Histogramme de simulation par loi uniforme discrète avec sa droite de probabilité théorique	2
3	Histogramme de simulation par loi géométrique avec une probabilité $p(X = \text{Succès}) = 0.3$ et une taille d'échantillon de 25	4
4	Histogramme de simulation par loi poisson avec un paramètre $\lambda = 3$ et une taille d'échantillon de 25	5
5	Comparaison de la loi de Poisson et de la loi binomiale pour $n = 1000$ et $p = 0.12$	5
6	Comparaison de la loi de Poisson selon le paramètre λ	6
7	Histogramme de simulation par loi uniforme sur $[0, 1]$ avec sa courbe de densité	6
8	Histogramme de simulation par loi uniforme sur $[2, 5]$ avec sa courbe de densité	7
9	Histogramme de simulation par loi normale centrée réduite $\mathcal{N}(0, 1)$ avec la courbe de densité théorique	8
10	Comparaison graphique des 2 équations (Voir Code Github : section 3.4 [2])	9
11	Histogramme de simulation par loi normale centrée réduite $\mathcal{N}(1, 4)$	9
12	Comparaison des courbes de densité de la loi normale centrée réduite $\mathcal{N}(0, 1)$ et de la loi normale $\mathcal{N}(1, 4)$ centrée en $\mu = 1$ et d'variance $\sigma^2 = 4$	9
13	Graphique en courbe de la loi normale $\mathcal{N}(0, 1)$	10
14	Histogramme de la loi normale $\mathcal{N}(0, 1)$	10
15	Histogramme de la loi normale $\mathcal{N}(0, 1)$ avec 20 classes	10
16	Histogrammes des simulations avec leurs courbes de densité théorique des échantillons : $\mathcal{N}(0, 1)$, $\mathcal{N}(1, 1)$ et $\mathcal{N}(2, 1)$. Puis des échantillons $\mathcal{N}(0, 1)$, $\mathcal{N}(0, 2)$ et $\mathcal{N}(0, 4)$	11
17	Fonction de répartition simulées avec leur courbe théoriques des échantillons : $\mathcal{N}(0, 1)$, $\mathcal{N}(1, 1)$ et $\mathcal{N}(2, 1)$. Puis des échantillons $\mathcal{N}(0, 1)$, $\mathcal{N}(0, 2)$ et $\mathcal{N}(0, 4)$	11
18	Courbe de densité de $\mathcal{N}(0, 1)$ avec la zone de probabilité $P(X \leq -1)$	12
19	Fonction de répartition de la loi normale $\mathcal{N}(0, 1)$ avec intersection en $x = -1$	12
20	Comparaison des lois normales $\mathcal{N}(0, 1)$ avec les variables $X_1 = X + 1$, $X_2 = X - 1$, $X_3 = 2X$ et $X_4 = X/2$ suivant une loi normale	13
21	Comparaison des lois normales $\mathcal{N}(4, 3)$ avec la variable $Z = \frac{Y-2}{\sqrt{3}}$ suivant une loi normale	13
22	Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$	14
23	Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$ et $\mathcal{N}(1, 1)$	14
24	Histogramme de la somme de 2 lois normales $\mathcal{N}(0, 1)$ et $\mathcal{N}(0, 2)$	14
25	Histogramme de la somme de carrés de 2 lois normales $\mathcal{N}(0, 1)$	14
26	Histogramme de la somme de carrés de 3 lois normales $\mathcal{N}(0, 1)$	14
27	Histogramme de la somme de carrés de 4 lois normales $\mathcal{N}(0, 1)$	14

Liste des tableaux

1	Tableau de comparaison des quantiles de la loi normale	1
2	Fréquences observées de l'échantillon simulé par la loi de Bernoulli $\mathcal{B}(0.3)$	2
3	Tableau partiel $[5, 10]$ des fréquences observées de la simulation binomiale de $p(\text{Succès}) = 0.8$ et $N = 100$ observations	3
4	Tableau partiel $[0, 6]$ des fréquences observées de la simulation binomiale de $p(\text{Succès}) = 0.2$ et $N = 100$ observations	3
5	Tableau partiel $[5, 10]$ des fréquences théoriques de la loi binomiale $\mathcal{B}(10, 0.8)$, arrondies à 3 décimales	3
6	Tableau partiel $[0, 6]$ des fréquences théoriques de la loi binomiale $\mathcal{B}(10, 0.2)$, arrondies à 3 décimales	3
7	Comparaison des paramètres de la loi de Poisson pour $\lambda = 2$ et $\lambda = 5$	6
8	Tableau de simulation des 5 lois normales de taille $n = 20$	10

9	Tableau de simulation des 5 lois normales de taille $n = 2000$ (voir code annexe 4)	10
10	Tableau de comparaison des valeurs de probabilités entre les fonctions de répartition et la fonction pnorm pour une loi normale $\mathcal{N}(0, 4)$, voir code annexe 5	12

Listings

1	Code : Variation des paramètres de la loi Poisson	16
2	Code : Vérification de l'égalité	16
3	Code : Probabilités des lois normales	16
4	Code : Probabilités des lois normales	16
5	Code : Probabilités par fonctions de répartition	16
6	Code : Lien entre les lois normales	16
7	Code : Sommes de variables aléatoires gaussiennes	17
8	Code : Démonstration de la loi d'une moyenne de lois gaussienne \bar{X}	17
9	Code : Exercice sur la loi normale	17
10	Code : Vérification de la densité de probabilité de la loi exponentielle	17

Références

1. *TP 2 : Simulation de lois de probabilités*. Aussi disponible à l'adresse : <https://fcorset.github.io/TPs/tp2.html>.
2. *Github TP : Simulation de lois de probabilités*. Aussi disponible à l'adresse : <https://github.com/AllanDCQ/Simulation-de-lois-de-probabilit-s/>.
3. P. L. HSU, Robbins Herbert. Complete Convergence and the Law of Large Numbers [<https://doi.org/10.1073/pnas.33.2.25>]. 1947.
4. *Loi uniforme discrète*. Aussi disponible à l'adresse : <https://bibmath.net/dico/index.php?action=affiche&quoi=%2F1%2Floiunifdisc.html>.
5. *Loi de Bernoulli*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/1/loibernoulli.html>.
6. *Loi binomiale*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php/index.php?action=affiche&quoi=/1/loibinomiale.html>.
7. *Loi géométrique*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/1/loigeometrique.html>.
8. *Loi de Poisson*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/1/loipoisson.html>.
9. *Loi Uniforme Continue*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/1/loiunifcont.html>.
10. *Loi Normale centrée réduite*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/1/loinormale.html>.
11. *Fonctions de répartition et de masse*. Aussi disponible à l'adresse : <https://www.bibmath.net/dico/index.php?action=affiche&quoi=/f/foncrepar.html>.