CPE Lyon 3ETI 2023-2024

Probabilités continues : TP5

Exercice 1 : Simulation de la loi normale (méthode de Box-Muller)

Nous avons vu dans le TP précédent la simulation de lois continues à l'aide de la méthode d'inversion qui consiste à calculer la variable aléatoire X associée à la loi par la formule $X = F^{-1}(Y)$ où F est la fonction de répartition de X et Y est une variable aléatoire uniformément distribuée sur $[0, 1]$.

La fonction e^{-x^2} n'admet pas de primitive pouvant s'exprimer simplement à l'aide des fonctions usuelles. On ne peut donc donner une expression littérale de la fonction de répartition d'une loi normale, et par conséquent on ne peut appliquer la méthode d'inversion pour simuler une loi normale.

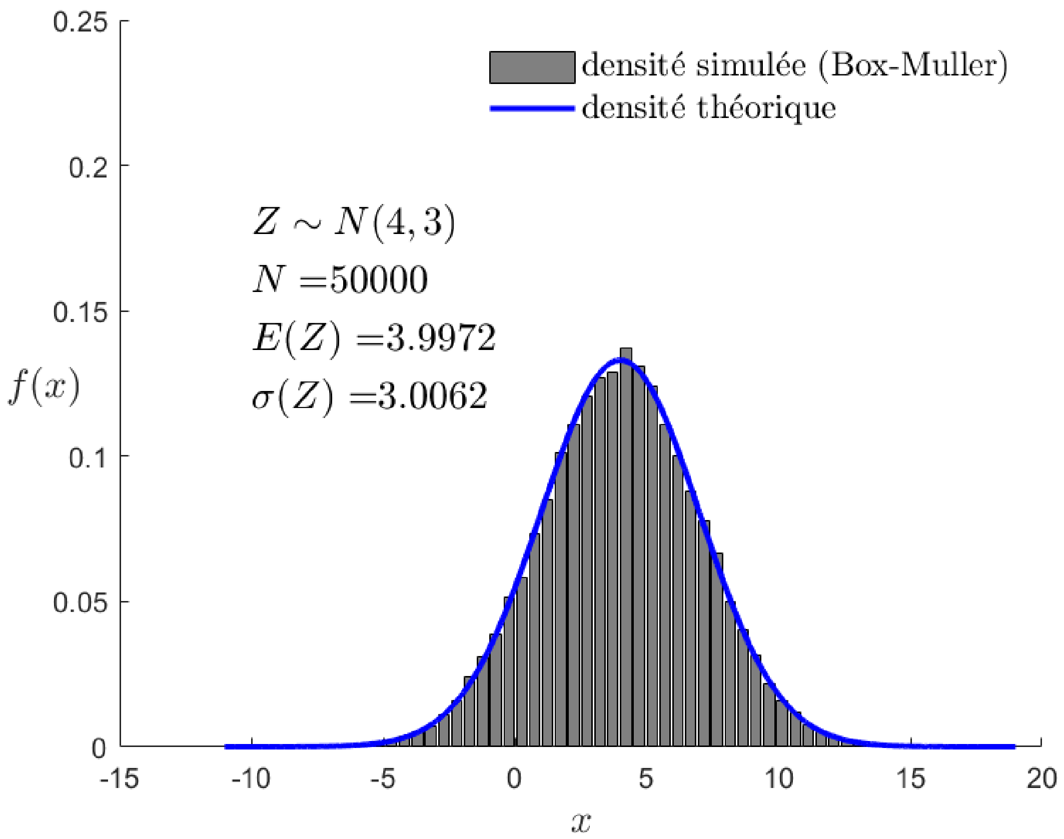
L'objectif de cet exercice est d'introduire la méthode dite de Box-Muller qui permet la simulation simultanée de 2 variables aléatoires normales centrées réduites.

Soit (X, Y) un couple de variables aléatoires représentant les coordonnées d'un point dans le plan. On suppose que X et Y sont indépendantes et qu'elles sont de distribution normale centrée réduite. On note $f_{X,Y}$ leur densité conjointe.

On note R et Θ les variables aléatoires correspondantes aux coordonnées polaires du point :

$$\begin{cases} X = R\cos\Theta \\ Y = R\sin\Theta \end{cases}$$

1) En vous aidant des résultats obtenus à l'exercice 10 du chapitre III du cours de probabilités continues, vérifier à l'aide d'un programme Matlab, que les variables X et Y suivent une loi normale centrée réduite.
2) Comment peut-on alors simuler une variable aléatoire obéissant à une loi normale de paramètres quelconques ? Illustrer cela à l'aide de Matlab (on affichera l'histogramme de la variable aléatoire que l'on comparera à la densité théorique, voir figure ci-dessous).



Exercice 2 : Théorème Central Limite

Attention ! cet exercice a pour objectif d'illustrer (et de mieux faire comprendre) le théorème central limite, il n'y a donc aucune simulation de loi de probabilité ici, et par conséquent, <u>aucune utilisation</u> des commandes `rand`, `histogram`, `mean` *et* `std`.

Rappel : La densité de probabilité d'une somme de deux variables aléatoires indépendantes X et Y, de densités f_X et f_Y respectivement, s'obtient à l'aide du produit de convolution :

$$f_{X+Y} = f_X * f_Y$$

Ce résultat se généralise à un nombre quelconque de variables aléatoires indépendantes, c'est-à-dire si $(X_k)_{1\leq k\leq n}$ est un ensemble de variables aléatoires indépendantes, et si on note $Z = X_1 + X_2 + \cdots + X_n$ alors :

$$f_Z = f_{X_1} * f_{X_2} * \dots * f_{X_n}$$

Dans le cas particulier où toutes les variables aléatoires suivent la même loi, alors la densité de Z est proche de celle d'une loi normale : c'est le **théorème central limite**.

1) On considère une variable aléatoire X dont la loi de probabilité est donnée par le tableau suivant :

k	3	4	5	6	7	8
$P(X=k)$	$\frac{4}{20}$	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{7}{20}$	$\frac{2}{20}$	$\frac{3}{20}$

a. Afficher l'histogramme de cette loi (voir figure page suivante) et calculer, à l'aide de Matlab, l'espérance mathématique et l'écart type de X (attention, il s'agit ici d'une loi théorique, les commandes `mean` et `std` ne sont d'aucune utilité).

b. Soient X_1 et X_2 deux variables aléatoires suivant la même loi que X. Calculer « à la main » le produit de convolution :

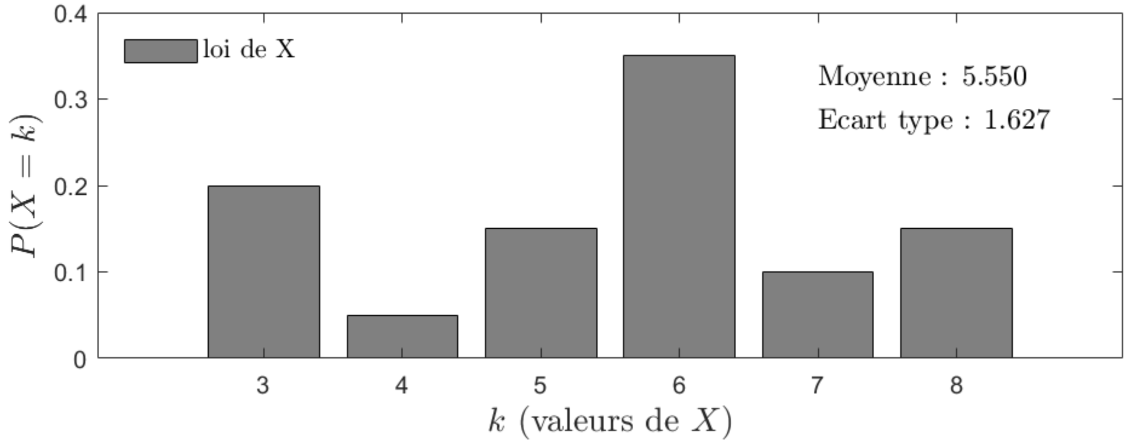
$$\frac{1}{20}[4,1,3,7,2,3] * \frac{1}{20}[4,1,3,7,2,3]$$

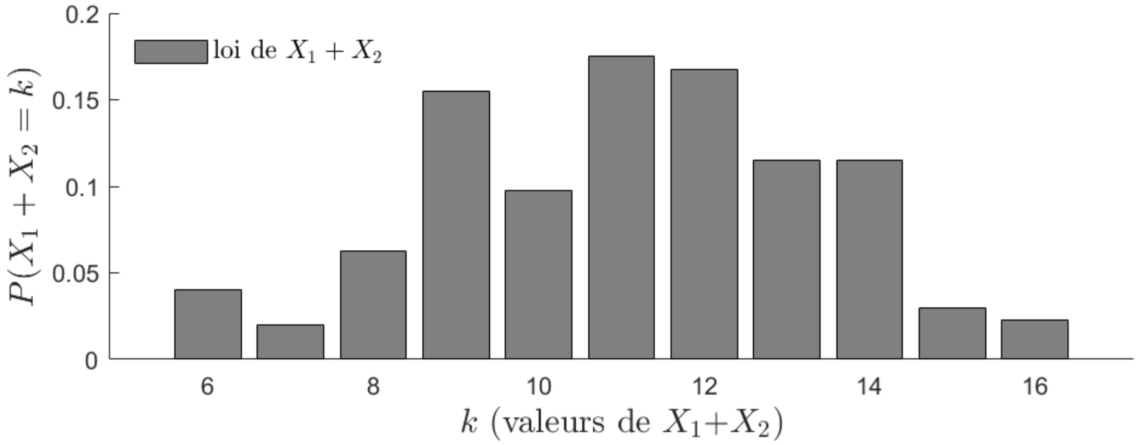
- vérifier votre résultat à l'aide de la commande `conv` de Matlab, puis afficher, dans la même fenêtre mais sur un graphique différent (utiliser la commande `subplot`), l'histogramme du résultat obtenu (voir figure page suivante).
- de quelle nouvelle variable aléatoire cet histogramme représente-t-il la loi de probabilité ?
- quelles sont les valeurs prises par cette variable ? (rectifier l'affichage de l'histogramme si nécessaire)

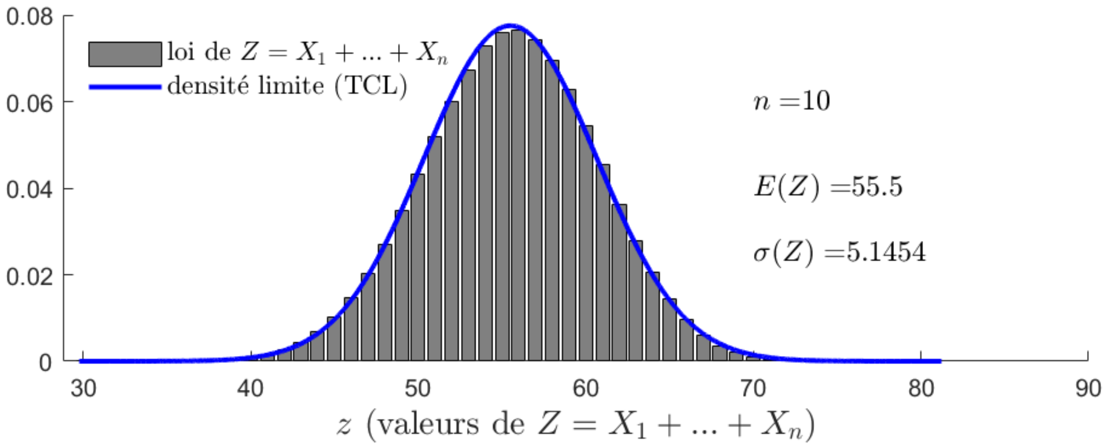
c. Toujours à l'aide de la commande `conv` de Matlab calculer et afficher (sous forme d'histogramme, toujours dans la même fenêtre mais sur un 3ème graphique) la « densité » de la variable aléatoire égale à la somme de $n = 10$ variables aléatoires suivant la même loi que X. Faire varier la valeur de n et commenter.

d. Calculer les valeurs théoriques de l'espérance mathématique et de l'écart type de cette variable aléatoire « somme ».

e. Afficher, en superposition de l'histogramme précédent (voir figure page suivante), la densité de la loi normale sous forme d'une courbe continue donnée par le théorème central limite (on pourra la tracer sur le support $[\mu - 5\sigma ; \mu + 5\sigma]$ où μ et σ sont la moyenne et l'écart type théoriques de la variable « somme » calculés en *d.*).


loi de X
Moyenne : 5.550
Ecart type : 1.627
$P(X = k)$
k (valeurs de X)


loi de $X_1 + X_2$
$P(X_1 + X_2 = k)$
k (valeurs de X_1+X_2)


loi de $Z = X_1 + ... + X_n$
densité limite (TCL)
n =10
$E(Z)$ =55.5
$\sigma(Z)$ =5.1454
z (valeurs de $Z = X_1 + ... + X_n$)