

Durée 3h.

Le document suivant liste une série de compétences qu'il faut maîtriser.

Le travail demandé est donc de répondre aux questions, de créer d'éventuels scripts, de les enregistrer.

On fournira à la fin du tp un compte-rendu sous la forme d'un pdf et un fichier code *Python*.

Dans la suite du tp, on utilisera les modules *Python* : *math*, *random*, *matplotlib* et *numpy*

On utilisera les imports suivants :

```
from math import *
from random import *
from matplotlib import pyplot
import numpy as np
```

On rappelle aussi qu'une k -liste est une liste de k éléments.

Notions de base

- Créer une fonction *Python* , nommée *matrice* qui prend en arguments deux entiers naturels n et p , et donne en sortie :

- une matrice A $n \times p$ dont les coefficients sont des nombres aléatoires uniformément distribués compris entre 0 et 1
- une matrice B qui regroupe les deux premières colonnes de A

Rappel : Une matrice $n \times p$ est codée en *Python* par une p -liste de n -listes.

Le module *numpy* de *Python* permet la génération de nombres aléatoires.

Voici quelques fonctions statistiques qui existent dans Python :

<code>numpy.random.uniform(0, 2, 7)</code>	une liste de 7 valeurs issues d'une loi uniforme entre 0 et 2
<code>numpy.random.binomial(10, 0.3, 7)</code>	une liste de 7 valeurs d'une loi binomiale de 10 tirages avec probabilité de succès de 0.3
<code>numpy.random.poisson(1, 7)</code>	une liste de 7 valeurs issues d'une loi de Poisson de paramètre 1
<code>numpy.random.normal(5, 2, 7)</code>	une liste de 7 valeurs issues d'une loi normale de moyenne 5 et écart-type 2
<code>numpy.mean(A)</code>	Calcul de la moyenne empirique (i.e. arithmétique) de A
<code>numpy.var(A)</code>	Calcul de la variance empirique de A
<code>numpy.std(A)</code>	Calcule l'écart-type empirique de A

Un vecteur comportant n réalisations indépendantes d'une même loi est aussi appelé n -échantillon.

- Créer une fonction *Python* nommée *stat_unif* qui prend en argument une entier naturel n et donne en sortie une liste dont les éléments sont, dans l'ordre, la liste, la moyenne, la variance et l'écart-type d'un n -échantillon d'une loi uniforme.
- (a) Générer un 2000-échantillon X de loi uniforme sur $[0; 1]$. Comparer sa moyenne empirique avec $E[X]$.
 (b) Coder une fonction *Python* , nommée *produit_terme* qui prend un entrée deux listes, et fournit en sortie une liste dont les termes sont les produits deux à deux des termes des listes en entrée.
 Par exemple, si $A=[1,2,3]$ et si $B=[1,2,3]$, alors `produit_terme(A,B)` est `[1,4,9]`.
 (c) Comparer la moyenne empirique de XY avec $E[XY]$, lorsque X et Y suivent toutes les deux une loi uniforme sur $[0; 1]$.

- (d) Générer un 2000-échantillon de loi uniforme sur $[0; 1]$ et déterminer le nombre N d'éléments de votre échantillon qui sont inférieurs à 0,3.
- (e) N est donc aléatoire, ou plus précisément N est une variable aléatoire. Quelle est sa loi ?
4. En théorie des probabilités et en statistique, la loi normale est l'une des lois de probabilité les plus adaptées pour modéliser des phénomènes naturels issus de plusieurs événements aléatoires.
- Par exemple, des mesures faites sur une population de grande taille donnent des valeurs qui sont distribuées selon une loi similaire à la loi normale. On peut penser à titre d'exemple à la taille des femmes adultes d'une population donnée.

Générer un 2000-échantillon d'une loi normale centrée de variance 4. Obtenir sa moyenne empirique et sa variance empirique.

On illustrera cette question par un histogramme qui représentera l'échantillon obtenu.

5. En théorie des probabilités et en statistiques, la loi de Poisson est une loi de probabilité discrète qui décrit le comportement du nombre d'événements se produisant dans un laps de temps fixé, si ces événements se produisent avec une fréquence moyenne connue et indépendamment du temps écoulé depuis l'événement précédent.

Depuis quelques décennies, son champ d'application s'est considérablement élargi. Actuellement, on l'utilise beaucoup dans les télécommunications (pour compter le nombre de communications dans un intervalle de temps donné), le contrôle de qualité statistique (nombre de défauts en SPC), la description de certains phénomènes liés à la désintégration radioactive (la désintégration des noyaux radioactifs suivant, par ailleurs, une loi exponentielle de paramètre noté aussi lambda), la biologie (mutations), la météorologie, la finance pour modéliser la probabilité de défaut d'un crédit

Générer un 2000-échantillon d'une loi de Poisson de paramètre 2. Obtenir sa moyenne empirique et sa variance empirique.

On illustrera cette question par un histogramme qui représentera l'échantillon obtenu.

6. Si x et y sont deux vecteurs de même taille, alors la commande $plot(x,y)$ permet la construction du graphe formé à partir des points $(x[i]; y[i])$.

Utiliser cette commande pour obtenir le graphe de la fonction normale centrée réduite.

On créera avant tout une fonction *gauss* qui prend en entrée trois paramètres : x la variable réelle, la moyenne m de la loi normale et l'écart-type *sigma* de la loi normale utilisée.

Problèmes

7. On décide de se déplacer sur un axe gradué. Le déplacement s'effectue de manière aléatoire, est long d'une unité et peut se faire sur l'axe dans les deux sens. On démarre d'un point O et on effectue 40 déplacements.
- Proposer un algorithme donnant la position du point d'arrivée.
 - Généraliser votre algorithme en listant les positions d'arrivée de n déplacements aléatoires.
 - Tracer les fréquences des différentes positions.
8. En 2013, dans un village perdu du Jura, 25 enfants sont nés, parmi lesquels 21 garçons. Entre 2000 et 2010, dans un autre village du Jura, 128 enfants sont nés dont 45 garçons.
- Estimer la proportion de garçons en France à la naissance en 2012.
 - Faire 100 simulations pour chacun des deux villages. Que peut-on en déduire ?