Partiel Mag 372 – Modélisation en probabilités 20-02-2025

1. Le sujet comporte 2 exercices indépendants.
2. Vous ne pouvez pas vous connecter à votre session personnelle, vous vous connectez à la session `mathens` avec le mot de passe `mathens`. Tous vos fichiers devront être enregistrés dans le dossier "`remise_copie`" présent sur le bureau. Il faut impérativement inclure votre nom dans le nom des fichiers.
3. L'idéal est de transmettre l'exportation en pdf de votre feuille Jupyter.
4. La réponse aux questions théoriques doit être sur les copies papier.
5. Les graphes doivent avoir un titre et être légendés (les graphes doivent être présents, le code qui les génère ne suffit pas).
6. On rappelle les commandes d'import des bibliothèques adéquates :

```python
import math
import numpy as np
import scipy.stats as scs
import matplotlib.pyplot as plt
import numpy.random as random
```

Exercice 1. Un lave-vaisselle est utilisé tous les jours, et chaque jour, il a une petite probabilité $p = \frac{1}{1000}$ de tomber en panne. L'objectif de l'exercice est d'évaluer le nombre de pannes X qu'aura le lave-vaisselle pendant sa durée de vie, qu'on estime à $N = 3000$ jours.

1. Pour $i \in [\![1, N]\!]$, on note X_i la variable aléatoire qui vaut 1 si le lave-vaisselle tombe en panne le i-ième jour, et 0 sinon. On suppose que ces variables de Bernoulli sont indépendantes, avec $\mathbb{P}[X_i = 1] = p$ pour tout i. Rappeler la valeur de

$$\mathbb{P}[X = k], \quad \text{avec } k \in [\![0, N]\!] \text{ et } X = X_1 + X_2 + \cdots + X_N.$$

2. Écrire un programme `fonction_repartition(L)` qui calcule la probabilité $\mathbb{P}[X \leq L]$ pour un entier L compris entre 0 et N. On pourra importer la fonction de deux variables `binom` pour calculer les coefficients binomiaux, à l'aide de la commande :

```python
from scipy.special import binom
```

3. Donner les valeurs de la fonction $\mathbb{P}[X \leq L]$ pour $L = 3$, $L = 5$ et $L = 7$. Vérifier que $\mathbb{P}[X \leq 7] \geq 98\%$. Dans ce qui suit, tous les histogrammes demandés pourront donc être dessinés sur l'intervalle entier $[\![0, 7]\!]$.
4. Écrire un programme `nombre_pannes()` qui simule la variable X. Utiliser ce programme pour obtenir un vecteur de 10 simulations indépendantes de la variable X.
5. Dessiner l'histogramme des fréquences empiriques

$$F_n(k) = \frac{\text{nombres de simulations } X^{(j)} \text{ égales à } k}{n},$$

où $X^{(1)}, X^{(2)}, \ldots, X^{(n)}$ sont des réalisations indépendantes de la variable X. On pourra prendre $n = 1000$ et, comme indiqué dans la question précédente, $k \in \{0, 1, \ldots, 7\}$.

6. Comme $N \gg 1$ et $p \ll 1$, il est bien connu que la loi de la question 1. est approximée par la loi Poisson(λ) avec $\lambda = Np$; cette loi de Poisson est la loi discrète sur les entiers donnée par
$$\text{Poisson}(\lambda)[k] = \mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!}.$$
Superposer à l'histogramme précédent l'histogramme de la loi de Poisson de paramètre $\lambda = Np$ (on pourra comme précédemment importer de `scipy.special` la fonction `factorial` pour les factorielles). Commenter ce dessin.

Exercice 2. On considère l'espace des états $\mathfrak{X} = \{0, 1, 2\}$ et la matrice P dont les lignes et colonnes sont indicées par $\mathfrak{X}$, et dont les entrées sont :
$$P = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}.$$

1. Écrire un programme qui vérifie qu'une matrice carrée, donnée sous la forme d'un array *numpy*, est une matrice stochastique. Utiliser ce programme avec la matrice P pour vérifier qu'elle est stochastique.
2. Trouver l'unique vecteur de probabilité $\pi = (\pi(0), \pi(1), \pi(2))$ avec $\pi(0)+\pi(1)+\pi(2) = 1$ et $\pi\,P = \pi$ (loi stationnaire). On pourra s'aider de la fonction `np.linalg.solve` (on rappelle par ailleurs qu'un produit matriciel de deux array *numpy* est obtenu avec `np.matmul`).
3. En utilisant la fonction `np.linalp.eigvals`, déterminer les valeurs propres de la matrice P (on pourra mettre de côté les minuscules termes imaginaires complexes s'ils apparaissent). Vérifier que le polynôme caractéristique de P s'écrit
$$\det(XI_3 - P) = (X - 1)(X - \alpha)^2,$$
où α est un nombre rationnel à determiner.
4. Trouver deux vecteurs $\nu = (\nu(0), \nu(1), \nu(2))$ et $\mu = (\mu(0), \mu(1), \mu(2))$ tels que $\nu\,P = \alpha\,\nu$ et $\mu\,P = \alpha\,(\mu + \nu)$ (on pourra imposer $\nu(0) = \mu(0) = 1$). Que peut-on dire de la famille (π, ν, μ) dans $\mathbb{R}^3$?
5. On note $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov sur $\mathfrak{X}$ de matrice de transition P. Pour toute loi initiale π_0, rappeler le lien entre π_0, P et la loi marginale
$$\pi_n = (\mathbb{P}[X_n = 0], \mathbb{P}[X_n = 1], \mathbb{P}[X_n = 2]).$$
Montrer que pour toute loi initiale π_0, il existe des coefficients réels a et b tels que pour tout $n \geq 0$,
$$\pi_n = \pi + \alpha^n\left((a + nb)\nu + b\mu\right).$$
En déduire que $\lim_{n\to\infty} \pi_n = \pi$. Commenter ce résultat.
6. Implémenter une fonction qui engendre une trajectoire de la chaîne de Markov sur n étapes, à partir d'un état initial donné $i \in \{0, 1, 2\}$.
7. Simuler une trajectoire de la chaîne avec $X_0 = 0$ et $n = 10\,000$ étapes. On pourra rendre le résultat sous la forme d'une liste contenant $X_1, X_2, \ldots, X_n$.

8. Dessiner sur un même graphe l'histogramme de la loi stationnaire de la question 2., et l'histogramme des fréquences de visites $(f_n(0), f_n(1), f_n(2))$ avec

$$f_n(k) = \frac{1}{n} \sum_{m=1}^{n} 1_{(X_m = k)}$$

et avec la trajectoire calculée précédemment. Commenter ces histogrammes.