Concours Blanc n°2 - Math 1 - sujet B

Mercredi 8 Mars
Durée : 4 heures

Exercice 1

On considère la suite (u_k) définie, pour $k \geq 2$ par $u_k = \dfrac{1}{k-1} - \dfrac{1}{k}$.

(1) Vérifier la suite (u_k) peut être considérée comme une distribution de probabilité. On note alors Y une variable aléatoire telle que $Y(\Omega) = [\![2, +\infty[\![$ et $P(Y = k) = u_k$.
(2) La variable aléatoire Y admet-elle une espérance ?

Soit $n \in \mathbb{N}^*$. On dispose d'une urne contenant n boules, indiscernables au toucher, numérotées de 1 à n et on effectue des tirages successifs, avec remise, dans cette urne.
On introduit la variable aléatoire X_n qui prend la valeur du rang du tirage pour lequel, pour la première fois, la boule piochée a un numéro supérieur ou égal à celle piochée lors du tout premier tirage.
Pour tout $k \in \mathbb{N}^*$, on note T_k la variable aléatoire qui correspond au numéro de la boule piochée au $k-$ième tirage.

(1) Quel est, pour tout $k \in \mathbb{N}^*$, la loi de T_k? *On observera que la loi de T_k dépend de n.*
Rappeler la commande Python qui permet d'en renvoyer une simulation.

(2) Que vaut $X_n(\Omega)$?

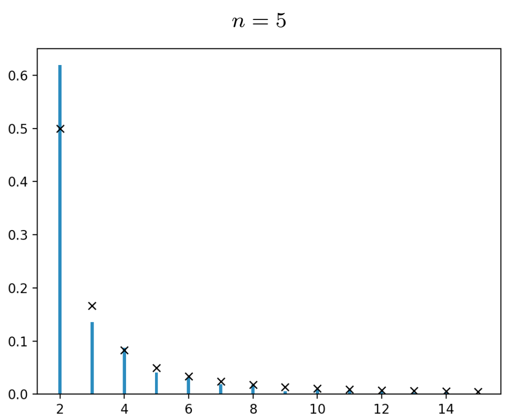
(3) (a) Recopier et compléter la fonction suivante de sorte qu'elle renvoie une simulation de X_n.

```
import numpy as np
import numpy.random as rd

def simul_X(n) :
    t = .........
    k = ........
    while .......... :
        k=k+1
    return k
```

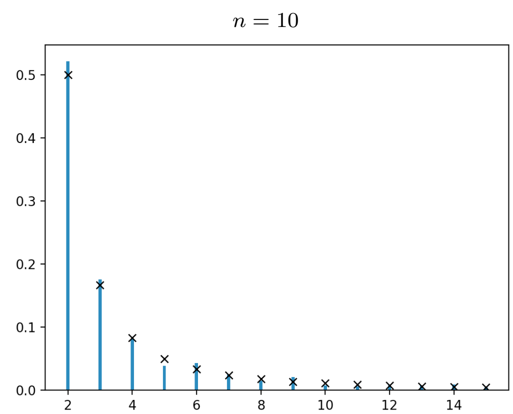
(4) On complète la fonction précédente par le code suivant dont l'exécution permet l'affichage reproduit ci-après (pour différentes valeurs de n). Que peut-on conjecturer? On précisera notamment ce que fait la fonction `mystère`.

```
import matplotlib.pyplot as plt

def mystere(n):
    S=[simul_X(n) for k in range(1000)]
    m=max(S)
    f=np.zeros(m)
    for k in range(1000):
        f[S[k]-2]+=1
    return f/1000

def u(k):
    return 1/(k-1)-1/k

n=5 # puis n=10, puis n=25 puis n=100
obs=mystere(n)
N=min(len(obs), 14)
x=[k for k in range(2, N+2)]
U=[u(k) for k in x]
h=[obs[k] for k in range(N)]
plt.plot(x, U, ’kx’)
plt.bar(x, h, width=0.1)
plt.show()
```

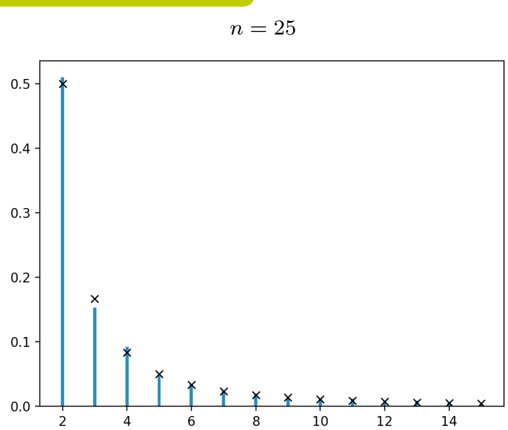
Affichage Python



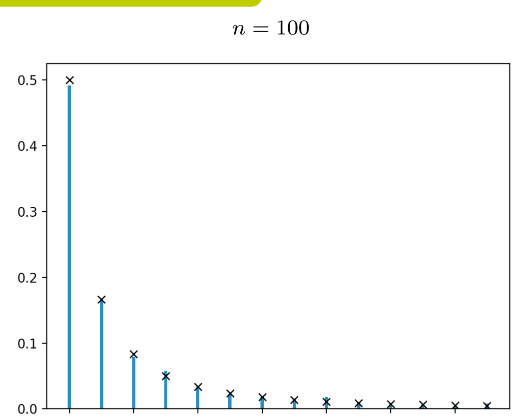
Affichage Python



Affichage Python



Affichage Python



(4) (a) Soient $k \geq 2$ et $1 \leq i \leq k$. Écrire l'évènement $[X_n = k] \cap [T_1 = i]$ à l'aide d'évènements dépendant uniquement des variables aléatoires T_j.

(b) En déduire que, pour tout $k \geq 2$,

$$P(X_n = k) = \frac{1}{n} \sum_{i=1}^{n} \left(\left(\frac{i-1}{n} \right)^{k-2} - \left(\frac{i-1}{n} \right)^{k-1} \right).$$

(5) (a) Montrer que, pour tout $N \geq 2$,

$$\sum_{k=2}^{N} kP(X_n = k) = \frac{1}{n} \sum_{i=1}^{n} \left(1 - \left(N + \frac{n}{n-i+1} \right) \left(\frac{i-1}{n} \right)^{N-1} + \frac{n}{n-i+1} \right)$$

(b) En déduire que X_n admet une espérance et que

$$E(X_n) = 1 + \sum_{j=1}^{n} \frac{1}{j}.$$

(6) Montrer que (X_n) converge en loi vers Y.
On pourra utiliser le résultat suivant: si f est une fonction continue sur $[a;b]$ alors

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f\left(a + \frac{k}{n} \right) = \int_a^b f(t) \mathrm{d}t.$$

Exercice 2

On considère deux variables aléatoires X_1 et X_2 indépendantes de même loi $\mathcal{U}(\{-1; 1\})$ définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ (que l'on ne cherchera pas à expliciter) et on pose, pour tout $\omega \in \Omega$,

$$M(\omega) = \begin{pmatrix} X_1(\omega) & X_2(\omega) \\ X_2(\omega) & X_1(\omega) \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}).$$

(1) Justifier que, pour tout $\omega \in \Omega$, $M(\omega)$ est une matrice diagonalisable. Est-elle inversible ?

(2) Montrer que, pour tout $\omega \in \Omega$,

$$\mathrm{Sp}(M(\omega)) = \{X_1(\omega) + X_2(\omega);\ X_1(\omega) - X_2(\omega)\}.$$

(3) Déterminer une matrice inversible $P \in \mathcal{M}_2(\mathbb{R})$ telle que, pour tout $\omega \in \Omega$,

$$P^{-1} M(\omega) P = D(\omega)$$

où $D(\omega)$ est une matrice diagonale à expliciter.

(4) On considère le système différentiel linéaire à coefficients *aléatoires*

$$(\mathcal{S}) \qquad \left\{ \begin{array}{lcl} y_1' & = & X_1 y_1 + X_2 y_2 \\ y_2' & = & X_2 y_1 + X_1 y_1 \end{array} \right.$$

Remarquons que pour tout $\omega \in \Omega$, on obtient un système différentiel linéaire à coefficients constants. On pose, pour tout $t \in \mathbb{R}$, $Y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$. On admet que $Y'(t) = \begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix}$.

(a) Montrer que, pour tout $\omega \in \Omega$, Y est solution de $(\mathcal{S})$ si et seulement si $Y' = M(\omega)Y$.

(b) Soit $\omega \in \Omega$ fixé.

(i) Quels sont les états d'équilibre du système?

(ii) Déterminer, pour tout $\omega \in \Omega$ l'ensemble des solutions de $(\mathcal{S})$.

(c) Calculer la probabilité que toutes les trajectoires soient convergentes.

Exercice 3

Pour tout réel $a \in \mathbb{R} \setminus \left\{0; -\frac{1}{2}\right\}$, on considère la fonction f_a, définie sur $\mathbb{R}^2$ par

$$f_a(x,y) = \left(1 + y + xy + ax^2\right) e^y.$$

(1) Justifier que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$.

(2) (a) Calculer les dérivées partielles d'ordre 1 de f_a.
(b) En déduire que f_a possède deux points critiques dont on précisera les coordonnées.

(3) Calculer les dérivées partielles d'ordre 2 de f_a.

(4) Montrer que, si λ_1 et λ_2 sont les racines d'un trinôme de la forme $x^2 - sx + p$ alors

$$\begin{cases} \lambda_1 + \lambda_2 &= s \\ \lambda_1 \lambda_2 &= p \end{cases}$$

(5) (a) Examiner, pour chacun des deux points critiques, à quelle condition portant sur a, f_a présente en ces points un extremum local. (*On utilisera la question précédente.*)
(b) Déterminer, en distinguant trois cas, si f_a présente sur $\mathbb{R} \times \mathbb{R}$ un maximum local ou un minimum local et donner sa valeur en fonction de a.

Problème

Partie 1 - Sommes de loi exponentielles indépendantes

Pour $n \in \mathbb{N}$ et $\lambda > 0$, on considère l'intégrale $I_n = \displaystyle\int_0^{+\infty} t^n e^{-\lambda t} \mathrm{d}t$.

(1) (a) Justifier que, pour tout $n \in \mathbb{N}$, I_n est une intégrale convergente.
(b) Montrer, à l'aide d'une intégration par parties que, pour tout $n \in \mathbb{N}$, on a $\lambda I_{n+1} = (n+1) I_n$.
(c) En déduire que, pour tout $n \in \mathbb{N}$, $I_n = \dfrac{n!}{\lambda^{n+1}}$.

Pour tout $n \in \mathbb{N}^*$ et $\lambda > 0$, on considère la fonction

$$g_n(t) = \begin{cases} 0, & \text{si } t < 0 \\ \dfrac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t}, & \text{si } t \geq 0 \end{cases}$$

(2) Justifier que g_n peut être considérée comme une densité de probabilité.
Quelle loi retrouve-t-on pour $n = 1$?

On considère $Y_1, Y_2, ..., Y_n$ des variables aléatoires indépendantes de même loi exponentielle de paramètre $\lambda > 0$ et on pose $S_n = Y_1 + Y_2 + ... + Y_n$.

(3) Déterminer $E(S_n)$ et $V(S_n)$.

Dans toute la suite, on **admet** que S_n est une variable aléatoire à densité, admettant pour densité la fonction g_n.

Partie 2 - Loi de Pareto

Soient a et b des réels strictement positifs. On dit qu'une variable aléatoire X suit la loi de Pareto de paramètres a et b si X admet pour densité la fonction

$$\forall t \in \mathbb{R}, \qquad f(t) = \begin{cases} 0, & \text{si } t < b \\ a\dfrac{b^a}{t^{a+1}}, & \text{si } t \geq b \end{cases}$$

(4) Vérifier que f définit une densité de probabilité. On note X une variable de densité f, suivant donc la loi de Pareto de paramètres a et b.

(5) (a) Déterminer une condition nécessaire et suffisante sur a pour que X admette une espérance et préciser, dans ce cas là, sa valeur.
(b) Même question avec la variance.

(6) Déterminer la fonction de répartition F_X de X.

(7) On pose alors $Y = \ln\left(\dfrac{X}{b}\right)$.
(a) Montrer que Y suit une loi exponentielle dont on précisera le paramètre.
(b) Déduire de la question précédente l'écriture d'une fonction Python d'en-tête `def simul_X(a,b):` qui renvoie une simulation de la loi de Pareto de paramètres a et b.

Partie 3 - Estimation des paramètres d'une loi de Pareto

On considère dans cette partie une loi de Pareto de paramètres a et b dont seule la valeur de b est connue. On dispose d'un $n-$échantillon $(X_1, ..., X_n)$ (avec $n \geq 2$) de X, et on note f_a une densité de X. On va estimer la valeur de a par la méthode du maximum de vraisemblance. Soient $x_1, ...x_n$ des réels strictement supérieurs à b. On introduit la fonction L définie sur $\mathbb{R}_+^*$ par

$$L(a) = \prod_{i=1}^{n} f_a(x_i).$$

(8) Exprimer $L(a)$ puis $\ln(L(a))$.
(9) On considère la fonction φ de $\mathbb{R}_+^*$ dans $\mathbb{R}$ définie par

$$\forall a > 0, \qquad \varphi(a) = n\ln(a) + na\ln(b) - (a+1)\sum_{i=1}^{n}\ln(x_i).$$

(a) Démontrer que φ atteint un maximum, atteint en un unique réel w que l'on exprimera en fonction de $x_1, .., x_n$ et b.
(b) Que peut-on dire de w pour la fonction L?

(10) On pose dorénavant $W_n = \dfrac{n}{\displaystyle\sum_{i=1}^{n}\ln\left(\dfrac{X_i}{b}\right)}$.
(a) Montrer que la variable aléatoire $\displaystyle\sum_{i=1}^{n}\ln\left(\dfrac{X_i}{b}\right)$ admet pour densité la fonction g_n avec $\lambda = a$.
(b) En déduire que $E(W_n) = \dfrac{na}{n-1}$ puis proposer un estimateur W_n', construit à partir de W_n par transformation linéaire, non biaisé pour a.