DS4 - Lundi 4 décembre 2023

Consignes

Tous les feuillets doivent être identifiables et numérotés par le candidat.
Les candidats sont invités à soigner la présentation de leur copie, à mettre en évidence les principaux résultats, à respecter les notations de l'énoncé et à donner des démonstrations complètes - mais brèves - de leurs affirmations.

Exercice 1 : tirages dans une urne

Soit n un entier supérieur ou égal à 2 . On considère une urne contenant n boules indiscernables numérotées de 1 à n.

On tire au hasard une boule et on la retire de l'urne ainsi que toutes les boules ayant un numéro supérieur à celui de la boule tirée. On réitère l'expérience jusqu'à ce que l'urne soit vide et l'on note X_n la variable aléatoire égale au nombre de tirages réalisés pour vider l'urne.

Pour tout entier i, on pourra noter N_i la variable aléatoire égale au numéro de la i-ème boule tirée s'il y a eu au moins i tirages, et 0 sinon.

1. Trouver la loi de X_2 puis donner son espérance et sa variance.
2. Trouver la loi de X_3 et donner son espérance.
3. Donner l'ensemble des valeurs que peut prendre X_n.
4. Déterminer $P(X_n = 1)$ et $P(X_n = n)$.
5. Simulation informatique :
   On suppose que l'on a déjà importé le package `numpy.random as rd`
   Ecrire une fonction Python, intitulée `def Exercice1(n):` qui simule cette expérience aléatoire et renvoie la valeur de X_n.
6. Prouver que pour tout $k \geq 2$, on a :
$$P(X_n = k) = \frac{1}{n} \sum_{i=2}^{n} P(X_{i-1} = k-1).$$
7. En déduire que $E(X_{n+1}) - E(X_n) = \dfrac{1}{n+1}$.
8. En déduire une expression de $E(X_n)$ sous forme d'une somme.
9. (a) Prouver que pour tout entier $k \geq 2$, on a : $\displaystyle\int_k^{k+1} \frac{1}{t}\, dt \leq \frac{1}{k} \leq \int_{k-1}^{k} \frac{1}{t}\, dt$.
   (b) En déduire que $\displaystyle\sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{\sim} \ln(n)$.
   (c) En déduire un équivalent de $E(X_n)$ quand n tend vers $+\infty$.

Exercice 2 : la loi de Weibull

Soit λ un réel fixé strictement positif. Soit X une variable aléatoire, dont la fonction de répartition $F_\lambda : \mathbb{R} \to \mathbb{R}$ est telle que

$$\forall x \in \mathbb{R}, \quad F_\lambda(x) = \begin{cases} 1 - \exp(-\lambda.\sqrt{x}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

1. Justifier que X est une variable aléatoire à densité. On dit alors que X suit la loi $\mathcal{W}(\lambda)$.
2. Donner une densité de X.
3. Etudier la convexité de F_λ sur $]0, +\infty[$ et la dérivabilité de F_λ en 0 à droite .
4. Représenter l'allure de la courbe représentative de F_λ dans le plan.
5. Soit Y la variable aléatoire $Y = \lambda \cdot \sqrt{X}$.
   (a) Justifier que Y suit une loi exponentielle dont précisera le paramètre et on donnera une densité.
   (b) Justifier que pour tout entier $r \in \mathbb{N}^*, E\left(Y^r\right)$ existe et préciser la valeur de $E\left(Y^r\right)$.
   (c) En déduire que pour tout entier $r \in \mathbb{N}^*, E\left(X^r\right)$ existe et préciser la valeur de $E\left(X^r\right)$. Préciser $E(X)$ et vérifier que $V(X) = \dfrac{20}{\lambda^4}$.
6. Soit Z une variable aléatoire suivant une loi exponentielle de paramètre λ.
   Prouver que la variable aléatoire Z^2 a même loi que X. Retrouver $E(X)$.

Problème : temps d'attente

Les parties 1 et 2 sont indépendantes ; la partie 3 utilise des résultats établis dans les parties 1 et 2. $(\Omega, \mathcal{A}, P)$ est un espace probabilisé et toutes les variables aléatoires considérées sont relatives à cet espace.

Partie 1 : Temps d'attente pour deux guichets

p est un réel fixé où $p \in]0, 1[$ et on note $q = 1 - p$.

Une gare dispose de deux guichets .Trois clients notés C_1, C_2, C_3 arrivent en même temps.
Les clients C_1 et C_2 se font servir tandis que le client C_3 attend puis effectue son opération dès que l'un des deux guichets se libère.
On définit X_1, X_2, X_3 les variables aléatoires égales à la durée de l'opération des clients C_1, C_2, C_3 respectivement. Ces durées sont en minutes et arrondies à l'unité supérieure ou égale.
On suppose que les variables aléatoires X_1, X_2, X_3 sont mutuellement indépendantes et suivent la même loi géométrique de paramètre $p \in]0, 1[$.

On note T la variable aléatoire où $T = max(X_1, X_2)$, Z la variable aléatoire $Z = min(X_1, X_2)$ et Δ la variable aléatoire $\Delta = |X_1 - X_2|$.

1. (a) Justifier pour tout $k \in \mathbb{N}^*$, $P(X_1 > k) = q^k$.
   (b) Calculer $P(Z > k)$ pour tout $k \in \mathbb{N}^*$.
   (c) Justifier que $P(Z = 1) = 1 - q^2$.

(d) Reconnaître la loi de Z. Préciser $E(Z)$, $V(Z)$.

2. (a) Exprimer $X_1 + X_2$ et Δ en fonction de Z et T.
   (b) Déterminer alors $E(T)$ en fonction de p.
   (c) Les variables aléatoires Z et T sont-elles indépendantes ?

3. (a) Préciser l'ensemble $\Delta(\Omega)$ des valeurs prises par $\Delta = |X_1 - X_2|$.
   (b) Déterminer $P(\Delta = 0)$ en fonction de p.
   (c) Soit $n \in \mathbb{N}^*$. Déterminer $P(X_1 - X_2 = n)$ en fonction de p.
   En déduire que
   $$P(\Delta = n) = \frac{2p.q^n}{2-p}$$
   (d) Justifier que la variable aléatoire Δ admet une espérance $E(\Delta)$ et la calculer.

4. Que représente l'événement $A = [X_3 > \Delta]$? Déterminer $P(A)$ en fonction de p.

5. Simulation en Python On suppose que l'on a déjà importé `numpy as np` et `numpy.random as rd`
   Ecrire un programme qui demande $p \in]0,1[$ à l'utilisateur, qui simule la variable Δ et affiche sa valeur et qui indique si l'événement A est réalisé ou non.

Partie 2 : Fonction génératrice associée à une VARD

Pour toute v.a.r. Y discrète définie sur Ω, à valeurs dans $\mathbb{N}$, on note

$$\forall t \in \mathbb{R}, \quad G_Y(t) = \sum_{k=0}^{+\infty} P(Y = k).t^k$$

6. Soit $t \in [-1, 1]$.
   Justifier que la série de terme général $P(Y = k).t^k$ où $k \in \mathbb{N}$ converge absolument.
   En déduire l'existence de l'espérance de la variable t^Y.
   Exprimer $E(t^Y)$ à l'aide de G_Y.

7. Que vaut $G_Y(1)$? Justifier que pour tout $t \in [0, 1]$, $0 \leq G_Y(t) \leq 1$.

8. Deux cas particuliers

   (a) Soit N une variable suivant la loi de Poisson de paramètre $\lambda > 0$.
   Justifier que G_N est une fonction définie sur $\mathbb{R}$ et exprimer, pour tout $t \in \mathbb{R}$, $G_N(t)$ en fonction de t et λ.
   Vérifier que G_N est dérivable sur $\mathbb{R}$ et vérifier que $E(N) = G'_N(1)$.
   (b) Soit X une variable suivant la loi géométrique de paramètre $p \in]0, 1[$.
   Déterminer l'ensemble de définition $\mathcal{D}_X$ de G_X et exprimer, pour $t \in \mathcal{D}_X$, $G_X(t)$ en fonction de t et p.
   Justifier que G_X est dérivable sur $\mathcal{D}_X$ et montrer que $E(X) = G'_X(1)$.

9. Soit une suite $(U_k)_{k\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, admettant toutes une espérance.
   On admet que dans ce cas, pour tout $n \in \mathbb{N}^*$, $E(U_1.U_2 \cdots U_n)$ existe et que
   $$E(U_1. \cdots .U_n) = E(U_1) \cdots E(U_n)$$

On considère une suite $(Y_k)_{k\in\mathbb{N}^*}$ de variables mutuellement indépendantes, de même loi que Y. On note, pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=1}^{n} Y_k$.

Soit $t \in [-1, 1]$. Justifier que pour tout $n \in \mathbb{N}^*$, $G_{S_n}(t) = (G_Y(t))^n$.

Dans la suite du problème, on admet le résultat suivant :

$\boxed{\text{si } G_Y \text{ est dérivable en 1, alors } Y \text{ admet une espérance et } E(Y) = G'_Y(1)}$

Partie 3 : étude d'un seul guichet

On considère un seul guichet qui reçoit et traite successivement d'éventuels clients.
On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables mutuellement indépendantes, où X_k est la durée de traitement en minutes du k-ième client.
Les variables X_k suivent une même loi géométrique de paramètre $p \in]0, 1[$.

On considère aussi la variable N égale au nombre de clients traités par le guichet dans une journée donnée.
Une étude a prouvé que N suit une loi de Poisson de paramètre $\lambda > 0$.
Par ailleurs les variables X_k où $k \in \mathbb{N}^*$ et N sont mutuellement indépendantes.
On note X_0 la variable certaine nulle.
On considère la variable définie par $S = \sum_{k=0}^{N} X_k$, c'est-à-dire que

$$\forall \omega \in \Omega, \quad S(\omega) = \sum_{k=0}^{N(\omega)} X_k(\omega)$$

S est égale à la durée totale de traitement des clients dans la journée considérée.
On note également pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{k=0}^{n} X_k$.

(a) Soit $n \in \mathbb{N}^*$. Justifier que pour tout $i \in \mathbb{N}$, $P_{[N=n]}(S = i) = P(S_n = i)$.
En déduire l'existence et la valeur de l'espérance conditionnelle $E(S \mid [N = n])$.

(b) Justifier que $E(S| \, [N = 0]) = 0$.

(c) A l'aide du théorème de l'espérance totale, montrer que S admet une espérance et déterminer $E(S)$ en fonction de λ et de p.

10. (a) Soit $n \in \mathbb{N}^*$ et $t \in [-1, 1]$. A l'aide du théorème de transfert et de certains résultats de la partie 2, justifier que l'espérance conditionnelle $E(t^S|[N = n])$ existe et l'exprimer en fonction de $G_{S_n}(t)$ puis de $G_{X_1}(t)$.
On admet que $E(t^S| \, [N = 0]) = 1$.

(b) En utilisant le théorème de l'espérance totale, montrer que :

$$\forall t \in [0, 1], \quad G_S(t) = G_N(G_{X_1}(t))$$

(c) En utilisant certains des résultats de la partie 2, justifier alors que S admet une espérance et retrouver $E(S)$.