DURÉE : 3 heures

INTRODUCTION

On considère une suite $(X_n)_{n \geqslant 1}$ de variables aléatoires réelles qui modélisent les résultats obtenus lors d'une suite d'expériences aléatoires. Les variables aléatoires X_n sont définies sur un même espace de probabilité $(\Omega, \mathscr{A}, P)$ et peuvent être discrètes ou continues. Pour tout $n \geqslant 1$, la n-ième expérience est un succès si l'événement $(X_n \in A)$ est réalisé, A étant dans le problème soit une partie donnée de $\mathbb{N}$, soit un intervalle de $\mathbb{R}$ éventuellement aléatoire, dans un sens que l'on précisera.

On désire étudier dans certains cas la variable N^A qui désigne le nombre d'expériences nécessaires pour obtenir un succès, ou plus généralement la variable N_k^A qui désigne le nombre d'expériences nécessaires pour obtenir une suite ininterrompue de k succès ($k \geqslant 2$). Dans une première partie, $(X_n)_{n \geqslant 1}$ est une suite de variables aléatoires indépendantes et de même loi ; on étudiera, pour certains ensembles A, la variable N^A puis, plus généralement, la variable N_k^A. On étudiera dans une seconde partie un problème concret de ce type modélisé par une suite de variables aléatoires réelles non indépendantes et de lois distinctes ; notre objectif sera de calculer $E(N^A)$ pour un certain ensemble A, espérance qui représentera le nombre moyen d'expériences à effectuer pour résoudre le problème effectivement posé.

Les deux parties du problème peuvent être traitées indépendamment l'une de l'autre.

PREMIÈRE PARTIE

On effectue une suite $(X_n)_{n \geqslant 1}$ d'observations indépendantes d'une même variable aléatoire réelle X. On rappelle que la suite $(X_n)_{n \geqslant 1}$ est une suite de variables aléatoires réelles indépendantes dès que pour tout $k \geqslant 2$, la famille $(X_1, X_2, \ldots, X_k)$ est une famille de variables aléatoires réelles indépendantes.

Si X est une variable discrète, A désigne une partie quelconque de $\mathbb{N}$, et si X est une variable continue, A désigne un intervalle quelconque de $\mathbb{R}$. On note $p = P(X \in A)$, $q = 1 - p$ et l'on suppose $0 < p < 1$.

A. Étude de N^A dans le cas où A est fixé quelconque.

1. Montrez que la variable N^A définie dans l'introduction est une variable aléatoire. Déterminez sa loi. On rappelle que la fonction génératrice d'une variable aléatoire N, à valeurs dans $\mathbb{N}$, est la série entière de terme général $P(N = n)\, z^n$. Calculez le rayon de convergence R de la fonction génératrice de N^A et donnez la forme explicite de cette fonction pour $|z| < R$. En déduire les valeurs de $E(N^A)$ et $\sigma^2(N^A)$ [espérance et variance de N^A].

2. On suppose dans cette question que la variable X suit une loi de Poisson de paramètre $\lambda > 0$ et que A désigne l'ensemble des entiers pairs de $\mathbb{N}$. Calculez alors $p = P(X \in A)$ et déduire de la première question les valeurs de $E(N^A)$ et $\sigma^2(N^A)$.

B. Étude de $E(N^A)$ dans le cas où A est aléatoire.

On suppose que la fonction de répartition F, de la variable X, est strictement croissante et continue. La fonction F admet donc une fonction réciproque notée F^{-1}. Pour tout $y \in \mathbb{R}$ on définit $A^y =]y, +\infty[$. Soit Y une variable aléatoire réelle définie sur $(\Omega, \mathscr{A}, P)$ admettant G pour fonction de répartition.

Soit $T(Y)$ l'application de Ω dans $\mathbb{R}$ définie par $T(Y)(\omega) = E(N^{A^{Y(\omega)}})$ pour tout $\omega \in \Omega$. On admet que $T(Y)$ est une variable aléatoire.

1. *a.* Donnez l'expression de $E(N^{A^y})$ pour tout $y \in \mathbb{R}$, en fonction de F et de y. En déduire que $T(Y)$ prend toutes ses valeurs dans l'intervalle $[1, +\infty[$.

   b. Déterminez, en fonction de F^{-1} et de G, la fonction de répartition H de la variable aléatoire $T(Y)$.

   c. Dans le cas $G = F$, montrez que la loi de $T(Y)$ admet une densité que l'on déterminera.

d. En supposant que F et G soient continûment dérivables, déterminez la densité de la loi de T (Y) au point x, en fonction de x, $\rho_Y \circ u$ et $\rho_X \circ u$ où ρ_Y et ρ_X sont les densités respectives de la loi de Y et de celle de X et u une fonction de x que l'on précisera.

2. Étude d'exemples.

a. On suppose dans cette question que Y suit une loi exponentielle de paramètre $\lambda = 1$ et que X suit une loi de densité $\dfrac{2\,e^{-2x}}{(1+e^{-2x})^2}$ sur $\mathbb{R}$. Montrez que la fonction de répartition F de X est strictement croissante et continue sur $\mathbb{R}$. Calculez sa fonction réciproque et en déduire l'expression de la fonction de répartition H de T (Y) ainsi que la densité de sa loi.

b. On suppose maintenant que la loi de X admet la densité $\rho(x) = \dfrac{1}{2}e^{-|x|}$ sur $\mathbb{R}$. Montrez que sa fonction de répartition F est strictement croissante et continue. Donnez l'expression de la fonction de répartition H de T (Y) dans les deux cas suivants :

- Y suit une loi exponentielle de paramètre 1 ;
- Y suit une loi uniforme sur [0, 1].

C. Étude de N_k^A, A fixé quelconque.

Pour tout entier $k \geqslant 2$, on désigne par N_k^A la variable, à valeurs entières, définie par : N_k^A épreuves ont été nécessaires pour obtenir une suite ininterrompue de k succès. Plus précisément, N_k^A prend la valeur k si les k premières expériences ont été chacune un succès et N_k^A prend la valeur n (où $n \geqslant k+1$) dès que les $(n-k-1)$ premières expériences n'ont pas donné de suite ininterrompue de k succès, la $(n-k)$-ième expérience a été un échec et les k dernières expériences ont été chacune un succès. On admettra que pour tout $k \geqslant 2$, la variable N_k^A est une variable aléatoire.

a. Soit pour tout $k \geqslant 2$, $n \geqslant 1$, le nombre $P_n^k = P(N_k^A = n)$.

Établir les relations de récurrence :

(1) $$P_{n+k+1}^k = q\,p^k\left(1 - \sum_{j=1}^{n} P_j^k\right)$$

(2) $$P_{n+k+1}^k = P_{n+k}^k - qp^k P_n^k.$$

L'entier k étant fixé, déterminez les valeurs de P_n^k pour $n = 1, \ldots, k-1$, puis pour $n = k, k+1, \ldots, 2k$.

En utilisant (2), calculez les valeurs de P_n^k pour $n = 2k+1, \ldots, 3k$. Montrez que l'ensemble de ces résultats et l'équation (2) définissent les P_n^k pour tout n.

À partir toujours de (2) vérifiez que l'on a bien $\displaystyle\sum_{n=1}^{\infty} P_n^k = 1$.

b. Pour tout $k \geqslant 2$ on note $g_k(x)$ la fonction génératrice de la variable N_k^A et R_k le rayon de convergence de la série entière définissant g_k. Déduire de (2) la formule :

$$q\,p^k\,g_k(x) = \sum_{n=1}^{\infty} x^n (P_{n+k}^k - P_{n-k-1}^k),$$

et, à partir de cette formule, donnez l'expression de $g_k(x)$ pour tout $|x| < R_k$. Justifiez l'inégalité $R_k \geqslant 1$.

c. On suppose admis que $R_k > 1$. Donnez alors l'expression de $E(N_k^A)$ pour tout $k \geqslant 2$.

DEUXIÈME PARTIE

Dans cette partie, on désigne par n un entier strictement positif donné quelconque.

On se propose de résoudre le problème suivant : un horticulteur a une plantation de n fleurs exactement à faire dans un jardin. Chaque bulbe utilisé ne donne une fleur qu'avec la probabilité p ($0 < p < 1$ et on pose $q = 1 - p$) et les événements concernant, pour chaque bulbe, la production de la fleur sont supposés indépendants. Après chaque plantation il remplace les bulbes qui n'ont pas fleuri et ceci jusqu'à ce que le jardin soit totalement fleuri, le nombre de fleurs nécessaires pour une floraison totale étant n. Soit alors T la variable aléatoire égale au nombre de fois où il devra effectuer des plantations pour obtenir une floraison totale (autrement dit, si $T = j$, le jardin n'est pas totalement fleuri à la $(j-1)$-ième plantation). On cherche alors la moyenne de T, valeur finie ou non, donnée par la formule :

$$E(T) = \sum_{k=1}^{+\infty} k\,P(T = k)$$

Pour résoudre le problème on considère la suite $(X_k)_{k \geqslant 0}$ de variables aléatoires définie par :

$$X_0 = 0$$

$$X_k = \begin{cases} \text{le nombre de fleurs apparues après } k \text{ plantations si } X_{k-1} < n \\ n \quad \text{si} \quad X_{k-1} = n \end{cases}$$

pour tout $k \geqslant 1$.

On remarque, et c'est admis, que si l'on pose $A = \{ n \}$, la variable T représente la variable N^A définie dans l'introduction.

1. Soient n_1, n_2 deux entiers positifs ou nuls tels que $n_1 + n_2 \leqslant n$; on pose $v_1 = n_1$, $v_2 = n_1 + n_2$.
   Calculez alors $P(X_1 = v_1)$ puis $P(X_1 = v_1, X_2 = v_2)$. En déduire les lois respectives des variables aléatoires X_1 et X_2.

2. Montrez alors par récurrence que la loi de X_k est, pour tout $k \geqslant 2$, une loi binomiale $\mathscr{B}(n, 1 - q^k)$.

3. *a.* Montrez, pour tout j entier strictement positif, l'égalité des événements $(T \leqslant j)$ et $(X_j = n)$.

   b. Soit (a_{ij}) un ensemble de réels positifs ou nuls indexé sur $\mathbb{N}^* \times \mathbb{N}$. On admettra l'égalité, dans le cas où le premier membre est une série convergente :

$$\sum_{j=0}^{+\infty} \left(\sum_{i=j+1}^{+\infty} a_{ij} \right) = \sum_{i=1}^{+\infty} \left(\sum_{j=0}^{i-1} a_{ij} \right).$$

   Montrez alors que :

$$E(T) = \sum_{j=0}^{+\infty} P(T > j) = \sum_{j=0}^{+\infty} (1 - P(X_j = n)).$$

   Indication : on pourra commencer par montrer que $\sum_{j=0}^{+\infty} (1 - (1 - q^j)^n)$ est convergente et qu'elle est égale à :

$$\sum_{j=0}^{+\infty} \sum_{i=j+1}^{+\infty} P(T = i).$$

   En déduire, sous forme d'une somme finie, la valeur de E (T).