D.S. 6 *(4 heures)*

Exercice

Un individu joue avec une pièce non nécessairement symétrique. On note p la probabilité d'obtenir pile et on suppose seulement $p \in]0,1[$.
Dans un premier temps, il lance la pièce jusqu'à obtenir pour la première fois pile. On note N le nombre de lancers nécessaires.
Dans un deuxième temps, il lance N fois cette même pièce et on note X le nombre de piles obtenus au cours de cette seconde série de lancers.

1) Préciser la loi de N, et la loi conditionnelle de X sachant $(N = n)$.

2) Déterminer la loi du couple (N, X).

3) On considère la fonction f définie sur $]-1,1[$ par : $\forall x \in]-1,1[\quad f(x) = \dfrac{1}{1-x}$.
Donner l'expression de la dérivée k-ième de f pour tout $k \geq 0$.
En déduire le développement en série entière de la fonction $x \mapsto \dfrac{1}{(1-x)^{k+1}}$ au voisinage de 0 pour k entier positif.

4) En déduire que la loi de X est donnée par

$$\forall k \geq 1, \quad P(X = k) = \frac{(1-p)^{k-1}}{(2-p)^{k+1}} \quad \text{et} \quad P(X = 0) = \frac{(1-p)}{(2-p)}.$$

5) Soit $\lambda \in]0,1[$, U une variable aléatoire de Bernoulli de paramètre λ et V une variable aléatoire géométrique de paramètre λ indépendante de U. On note $Y = UV$.

a) Sans calculer sa loi, calculer l'espérance de Y.

b) Pour $k \in \mathbb{N}$, calculer $P(Y = k)$ (on pourra traiter séparément le cas $k = 0$).

c) Calculer la variance de Y.

6) En déduire que X a même loi qu'un produit de deux variables aléatoires indépendantes, l'une étant une variable de Bernoulli et l'autre une variable géométrique de même paramètre.

Problème

Dans ce problème, nous étudions le processus de Galton-Watson qui permet entre autres de modéliser le développement d'une population. Ce processus est par exemple utilisé en biologie ou en physique nucléaire.
Dans tout le problème, on se place dans un espace probabilisé $(\Omega, \mathcal{A}, P)$.
Si X est une variable aléatoire entière et positive sur cet espace, on notera G_X, série entière de rayon de convergence au moins 1, la fonction génératrice de X. On rappelle que la fonction génératrice de X est la somme de la série entière :

$$\forall t \in [-1,1],\ G_X(t) = \sum_{n=0}^{\infty} P(X = n)t^n$$

La fonction génératrice d'une variable aléatoire caractérise sa loi. Plus précisément, si X est une variable aléatoire à valeurs dans $\mathbb{N}$ et si (a_n) est une suite de réels positifs tels que, pour tout $t \in [0,1[$, $G_X(t) = \sum_{n=0}^{\infty} a_n t^n$, alors, pour tout $n \in \mathbb{N}$, $a_n = P(X = n)$.
On admet le résultat suivant (lemme de Cesàro) : si $(a_n)_{n\in\mathbb{N}}$ est une suite de nombres réels convergeant vers ℓ et si on pose, pour $n \in \mathbb{N}^*$, $b_n = \dfrac{1}{n}(a_1 + \cdots + a_n)$ alors la suite $(b_n)_{n\geq 1}$ converge vers ℓ.

I – Étude d'une suite récurrente

On considère une fonction f de classe $\mathcal{C}^2$ sur $[0,1]$ à valeurs dans $[0,1]$ telles que f' et f'' soient à valeurs positives. On suppose $f(1)=1$, $f'(0)<1$ et $f''(1)>0$.

On considère de plus la suite récurrente $(u_n)_{n\in\mathbb{N}}$ définie par $u_0=0$ et, pour tout $n\in\mathbb{N}$, $u_{n+1}=f(u_n)$. On pose $m=f'(1)$.

I.A –

I.A.1) Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ est croissante, puis qu'elle est convergente. On note ℓ sa limite.

I.A.2) Montrer que l'équation $f(x)=x$ admet une plus petite solution. Dans toute la suite, on la notera x_f.

I.A.3) Montrer que $\ell=x_f$.

I.B – On suppose $m>1$. Montrer que $x_f\in[0,1[$.

I.C – On suppose maintenant $m\leq 1$. Montrer que $x_f=1$ et que pour tout $n\in\mathbb{N}$, $u_n\neq 1$.

I.D – Dans cette question, on suppose $m=1$.

I.D.1) On pose, pour $n\in\mathbb{N}$, $\varepsilon_n=1-u_n$. Montrer que

$$\lim_{n\to+\infty}\left(\frac{1}{\varepsilon_{n+1}}-\frac{1}{\varepsilon_n}\right)=\frac{f''(1)}{2}.$$

I.D.2) En déduire que, quand n tend vers l'infini, $1-u_n=\varepsilon_n\sim\dfrac{2}{nf''(1)}$. *On pourra utiliser le lemme de Cesàro admis en préambule.*

I.E – On suppose maintenant $m<1$ et on pose encore, pour $n\in\mathbb{N}$, $\varepsilon_n=1-u_n$.

I.E.1) Montrer que la série de terme général ε_n est absolument convergente et en déduire la convergence de celle de terme général $\ln\left(\dfrac{m^{-(n+1)}\varepsilon_{n+1}}{m^{-n}\varepsilon_n}\right)$.

I.E.2) En déduire qu'il existe $c>0$ tel que, quand n tend vers l'infini, $1-u_n\sim c.m^n$.

II – Formule de Wald

Soient $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, mutuellement indépendantes, de même loi à valeurs dans $\mathbb{N}$, et T une variable aléatoire à valeurs dans $\mathbb{N}$ indépendante des précédentes. $(T,X_n)_{n\in\mathbb{N}^*}$ est une famille de variables aléatoires mutuellement indépendantes.

On note G_X la fonction génératrice commune à toutes les X_n.

Pour $n\in\mathbb{N}$ et $\omega\in\Omega$, on pose $S_n(\omega)=\sum\limits_{k=1}^{n}X_k(\omega)$ et $S_0(\omega)=0$, puis $S(\omega)=S_{T(\omega)}(\omega)$.

II.A – On souhaite démontrer l'égalité $G_S=G_T\circ G_X$.

II.A.1) Montrer que, si X et Y sont deux variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes, alors

$$G_{X+Y}=G_XG_Y.$$

II.A.2) En admettant que, pour tout $k\in\mathbb{N}$, S_k est indépendante de X_{k+1}, prouver que

$$\forall k\in\mathbb{N},\quad G_{S_k}=(G_X)^k.$$

II.A.3) En admettant que, pour tout $n\in\mathbb{N}$, T et S_n sont indépendantes, montrer que

$$\forall t\in[0,1[,\ \forall K\in\mathbb{N},\ G_S(t)=\sum_{k=0}^{K}P(T=k)G_X(t)^k+\sum_{n=0}^{\infty}\left(\sum_{k=K+1}^{\infty}P(T=k)P(S_k=n)t^n\right)$$

II.A.4) Pour $K\in\mathbb{N}$ et $t\in[0,1[$, on pose $R_K=\sum\limits_{n=0}^{\infty}\left(\sum\limits_{k=K+1}^{\infty}P(T=k)P(S_k=n)t^n\right)$. Montrer que

$$0\leq R_K\leq\frac{1}{1-t}\sum_{k=K+1}^{\infty}P(T=k).$$

II.A.5) Conclure.

II.B – En déduire que, si T et les X_n sont d'espérance finie, alors S aussi et $E(S) = E(T)E(X_1)$.

II.C – Lors d'une ponte, un insecte pond un nombre aléatoire d'œufs suivant la loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$.
Ensuite, la probabilité qu'un œuf donné devienne un nouvel insecte est $\alpha \in]0, 1[$.

II.C.1) Donner la fonction génératrice d'une variable aléatoire suivant la loi $\mathcal{P}(\lambda)$.

II.C.2) En utilisant la relation de composition ci-dessus, déterminer la loi du nombre d'insectes issus de la ponte.

III – Processus de Galton-Watson

Soit μ une loi de probabilité caractérisée par la suite $(p_k)_{k\in\mathbb{N}}$ de nombre réels entre 0 et 1 telle que $\sum\limits_{k=0}^{\infty} p_k = 1$. Dire qu'une variable aléatoire X sur $(\Omega, \mathcal{A}, P)$ suit la loi μ signie que $X(\Omega) \subset \mathbb{N}$ et, pour tout $k \in \mathbb{N}$, $P(X = k) = p_k$.

On suppose que $p_0 + p_1 < 1$ (ce qui signifie qu'il existe au moins un entier k supérieur ou égal à 2 tel que $p_k \neq 0$).

On étudie un individu qui a un certain nombre de fils. Ces fils ont également chacun (indépendamment les uns des autres) un certain nombre de fils et ainsi de suite. Afin de modéliser la situation, on se donne des variables aléatoires $(X_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}^*}$, indépendantes, qui suivent toutes la loi μ ; on pose Y_0 la variable certaine égale à 1 et, pour $n \in \mathbb{N}$ et $\omega \in \Omega$,

$$\begin{cases} Y_{n+1}(\omega) = 0 & \text{si } Y_n(\omega) = 0 \\ Y_{n+1}(\omega) = \sum\limits_{i=1}^{Y_n(\omega)} X_{n,i}(\omega) & \text{si } Y_n(\omega) \neq 0 \end{cases}.$$

Y_n représente le nombre d'individus à la génération n.
S'il n'y a pas d'individu à la génération n, il n'y en a pas plus à la génération suivante et sinon, le nombre de fils du i-ième élément de la génération n est égal à $X_{n,i}$.

On dit qu'il y a *extinction* lorsqu'il existe un entier n tel que $Y_n = 0$.

On note f la fonction génératrice de la loi μ (et donc de chacune des variables $X_{n,i}$) et, pour $n \in \mathbb{N}$, φ_n la fonction génératrice de la variable aléatoire Y_n.
On a donc en particulier, pour $t \in [0, 1]$, $\varphi_0(t) = t$.

On suppose que toute variable aléatoire suivant la loi μ possède une espérance égale à m et une variance.

III.A – Probabilité d'extinction

III.A.1) Montrer que pour tout $n \in \mathbb{N}$, $\varphi_{n+1} = \varphi_n \circ f$.

III.A.2) Exprimer, pour $n \in \mathbb{N}$, l'espérance de Y_n en fonction de m et de n.

III.A.3a) Vérifier que la probabilité d'extinction est égale à la limite de la suite $\big(\varphi_n(0)\big)_{n\geq 0}$.

III.A.3b) Vérifier qu'on peut appliquer les résultats de la partie **I** à la suite $\big(\varphi_n(0)\big)_{n\geq 0}$.

III.A.4) Si $m \leq 1$, montrer que la probabilité d'extinction est égale à 1.

On définit alors le temps T d'extinction par

$$\forall \omega \in \Omega,\ \begin{cases} T(\omega) = \min\{n \in \mathbb{N} \ / \ Y_n(\omega) = 0\} & \text{s'il existe } n \in \mathbb{N} \text{ tel que } Y_n(\omega) = 0 \\ T(\omega) = -1 & \text{sinon} \end{cases}.$$

On admettra que T est une variable aléatoire.

III.B – Cas sous-critique $m < 1$

On suppose dans cette question que $m < 1$.

III.B.1) Vérifier que T admet une espérance.

III.B.2a) Montrer que, pour tout entier n, $P(Y_n \geq 1) \leq m^n$.

III.B.2b) Montrer que $E(T) = \sum\limits_{k=0}^{\infty} P(T > k)$.

III.B.2c) En déduire une majoration de $E(T)$.

III.C – Étude de la lignée

Dans cette question, on suppose $m \leq 1$.

On note, pour $n \in \mathbb{N}^*$, $Z_n = 1 + \sum\limits_{i=1}^{n} Y_i$ et $Z = 1 + \sum\limits_{n=1}^{\infty} Y_n$.

On admettra que Z est une variable aléatoire définie sur $\bigcup\limits_{k \in \mathbb{N}} (Y_k = 0)$.

III.C.1) Montrer que Z est définie sur un ensemble de probabilité 1.

III.C.2a) Montrer que, pour tout $k \in \mathbb{N}$, $\big(P(Z_n \leq k)\big)_{n \in \mathbb{N}^*}$ est une suite convergente. Déterminer sa limite.

III.C.2b) En déduire que pour tout $k \in \mathbb{N}$, $\big(P(Z_n = k)\big)_{n \in \mathbb{N}^*}$ converge vers $P(Z = k)$.

III.C.2c) Montrer que pour tout $s \in [0, 1[$, tout $n \in \mathbb{N}^*$ et $K \in \mathbb{N}$,

$$|G_{Z_n}(s) - G_Z(s)| \leq \sum_{k=0}^{K} |P(Z_n = k) - P(Z = k)| + \frac{s^{K+1}}{1 - s}.$$

III.C.2d) En déduire que la suite de fonctions (G_{Z_n}) converge simplement vers G_Z sur $[0, 1]$.

III.C.3a) Exprimer G_{Z_1} en fonction de f.

III.C.3b) On admet que, pour tout n entier naturel supérieur ou égal à 2 et pour tout $s \in [0, 1]$,

$$G_{Z_n}(s) = s.f\left(G_{Z_{n-1}}(s)\right).$$

En déduire que, pour tout $s \in [0, 1]$,

$$G_Z(s) = s.f\left(G_Z(s)\right).$$

III.C.3c) Montrer que Z est d'espérance finie si et seulement si $m < 1$. Calculer l'espérance lorsque c'est le cas.

IV – Un exemple

On suppose dans cette partie que, pour tout $k \in \mathbb{N}$, $p_k = \dfrac{1}{2^{k+1}}$.

IV.A – Exprimer, pour $t \in [0, 1]$, $f(t)$ et calculer m.

IV.B – Vérifier que pour tout $t \in [0, 1[$, $\varphi_n(t) \neq 1$. On peut donc poser

$$a_n(t) = \frac{1}{\varphi_n(t) - 1}.$$

IV.C – Montrer que, pour $t \in [0, 1[$, la suite $(a_n(t))_{n \in \mathbb{N}}$ est arithmétique.

IV.D – En déduire que, pour $t \in [0, 1[$ et $n \in \mathbb{N}$,

$$\varphi_n(t) = \frac{n + (1 - n)t}{1 + n - nt}.$$

IV.E – Exprimer, pour $n, k \in \mathbb{N}$, $P(Y_n = k)$ en fonction de n et k.

IV.F – Exprimer, en fonction de $n \in \mathbb{N}^*$, la probabilité de l'événement $(T > n)$. La variable T admet-elle une espérance ?

IV.G – Exprimer, pour $s \in [0, 1[$, $G_Z(s)$ en fonction de s. En déduire la loi de Z.

La théorie, c'est quand on sait tout et que rien ne fonctionne.
La pratique, c'est quand tout fonctionne et que personne ne sait pourquoi.

Mais ici, nous avons réuni théorie et pratique :
rien ne fonctionne et personne ne sait pourquoi.

Albert EINSTEIN