HEC 2021

Dans ce problème, on s'intéresse à un modèle, inspire du modèle de Cori, de propagation d'un virus au sein d'une population.
La partie 1 introduit des outils théoriques permettant de définir et d'étudier ce modèle.
Les parties 2 et 3 concernent cette étude. Si l'on fait abstraction des définitions, des notations et de la question 17, la partie 3 est indépendante des parties 1 et 2.

On admet dans ce problème les *théorèmes de limite monotone :*

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements (*ie :* $\forall\, n \in \mathbb{N},\ A_n \subset A_{n+1}$), alors

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n)$$

- Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements (*ie :* $\forall\, n \in \mathbb{N},\ A_{n+1} \subset A_n$), alors

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{N\to+\infty} \mathbb{P}(A_n)$$

Partie 1 - Lois composées

On considère :

- un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ et J un sous-ensemble non vide de $\mathbb{R}^+$;
- une variable aléatoire Y sur cet espace à valeurs dans J.
- une famille $(X_t)_{t\in J}$ de variables aléatoires sur cet espace **à valeurs dans** $\mathbb{N}$ et **indépendantes de** Y telles que pour tout $t \in J$,

  $$X_t \text{ suit la loi } \mu(t)$$

  $\mu(t)$ désignant une loi de probabilité de paramètre t.

On définit la variable aléatoire Z sur cet espace par :

$$\forall \omega \in \Omega,\ \text{si } Y(\omega) = t \text{ alors } Z(\omega) = X_t(\omega)$$

et on dit que Z suit la loi $\mu(Y)$.
On considère dans cette partie une telle variable Z qui suit la loi $\mu(Y)$.
Pour tout $k \in \mathbb{N}$, on définit aussi la fonction f_k de J dans $[0,1]$ par :

$$f_k(t) = \mathbb{P}([X_t = k])$$

1. *Un exemple avec Python.*
   On considère le script Python suivant :

```
def X(t):
    r = 1
    while rd.random() > ...
        ...
    return r

Y = rd.random()
Z = .....
print(Z)
```

   En considérant les notations précédentes avec $J =]0,1[$ et en notant Y la variable aléatoire dont Y est une simulation, compléter le script précédent pour que Z soit une simulation d'une variable aléatoire qui suit la loi géometrique $\mathscr{G}(Y)$.

Cas où Y est discrète.
On suppose dans les questions 2 et 3 que Y est discrète.

2. (a) Soit $y \in Y(\Omega)$. Montrer que, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}\left([Z=k] \cap [Y=y]\right) = f_k(y)\mathbb{P}\left([Y=y]\right)$$

et si $\mathbb{P}\left([Y=y]\right) \neq 0$,

$$\mathbb{P}_{[Y=y]}\left([Z=k]\right) = f_k\left(y\right)$$

(b) En déduire que :

$$\mathbb{P}\left([Z=k]\right) = \mathbb{E}\left(f_k\left(Y\right)\right) \tag{1}$$

(c) *Un exemple où* $J = \mathbb{N}^*$.
Soit $p \in]0,1[$. Si pour tout $n \in \mathbb{N}^*$, X_n suit la loi uniforme sur $[\![1,n]\!]$ et si la loi de Y est définie par, pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}\left([Y=n]\right) = np^2\left(1-p\right)^{n-1}$$

montrer que Z suit la loi géométrique de paramètre p.

3. On suppose que pour tout $t \in J$, $\mathbb{E}\left(X_t\right)$ existe. On note $g\left(t\right)$ cette espérance et on suppose que $\mathbb{E}\left(g\left(Y\right)\right)$ existe.

(a) Montrer que :

$$\mathbb{E}\left(g\left(Y\right)\right) = \sum_{y \in Y(\Omega)} \left(\sum_{k=0}^{+\infty} k f_k\left(y\right) \mathbb{P}\left([Y=n]\right)\right)$$

(b) En admettant que l'on peut inverser l'ordre des sommes, montrer que $\mathbb{E}\left(Z\right)$ existe et que :

$$\mathbb{E}\left(Z\right) = \mathbb{E}\left(g\left(Y\right)\right) \tag{2}$$

On admet que les résultats établis dans les questions 2 et 3, en particulier (1) **et** (2)**, sont encore vrais lorsque** Y **n'est plus discrète.**

4. *Un premier exemple.* On suppose que $J =]0,1[$, que la loi de X_t est la loi géométrique de paramètre t et que Y suit la loi uniforme sur $]0,1[$.
Montrer que pour tout $k \in \mathbb{N}^*$, $\mathbb{P}\left([Z=k]\right) = \dfrac{1}{k\left(k+1\right)}$. La variable aléatoire Z admet-elle une espérance ?

5. *Un deuxième exemple.* On suppose que $J = [0,+\infty[$, que la loi de X_t est la loi de Poisson de paramètre t et que Y suit la loi exponentielle de paramètre $\lambda > 0$.
Par suite, Z suit la loi $\mathscr{P}\left(Y\right)$.
Par convention, la loi de Poisson de paramètre 0 est la loi de la variable aléatoire nulle.

(a) Montrer que pour tout $k \in \mathbb{N}$,

$$\mathbb{P}\left([Z=k]\right) = \int_0^{+\infty} \frac{t^k}{k!}\lambda e^{-(\lambda+1)t}\,\mathrm{d}t = \frac{\lambda}{(\lambda+1)^{k+1}} \int_0^{+\infty} \frac{x^k}{k!} e^{-x}\,\mathrm{d}x$$

(b) En raisonnant par récurrence sur $k \in \mathbb{N}$, justifier que pour tout $k \in \mathbb{N}$,

$$\int_0^{+\infty} \frac{x^k}{k!} e^{-x}\,\mathrm{d}x = 1$$

(c) Déterminer la loi de Z. Reconnaître la loi de Z + 1.

(d) En déduire $\mathbb{E}(Z)$. Ce résultat est-il cohérent avec l'égalité (2) ?

Partie 2 - Le modèle de Cori

On considère une population d'effectif infini dans laquelle un individu donné est infecté le jour 0 par un virus contagieux.
Soit $d \in \mathbb{N}^*$. On suppose que :

- tout individu infecté par le virus est immédiatement contagieux et sa contagiosité ne dure que $(d+1)$ jours, du jour n où il est infecté jusqu'au jour $(n+d)$ $(n \in \mathbb{N})$;
- une fois infectés, les individus présentent un même profil de contagiosité donné par un $(d+1)$-uplet $(\alpha_0, \alpha_1, \dots, \alpha_d)$ qui dépend généralement de facteurs biologiques.

Pour tout $k \in [\![0, d]\!]$, on dit que α_k est la contagiosité de tout individu ayant éte infecté k jours plus tôt.
Autrement dit, on peut considérer que α_k, lié à la nature du virus, détermine la proportion d'individus contaminés par un individu infecté, parmi tous ceux avec lesquels il est en contact k jours après sa contamination.

Finalement, les réels $\alpha_0, \alpha_1, \dots, \alpha_d$ sont tels que, pour tout $k \in [\![0, d]\!]$, $\alpha_k \in]0, 1[$ et on note $\alpha = \sum_{k=0}^{d} \alpha_k$, ce qui signifie que α est la contagiosité globale d'un individu infecté sur toute la période où il est infecté.
On utilise les notations et définitions de la partie 1 avec $J = \mathbb{R}^+$.
On suppose que les variables aléatoires qui interviennent par la suite sont définies sur l'espace $(\Omega, \mathscr{A}, \mathbb{P})$.

- Pour tout $n \in \mathbb{N}$, on note R_n la variable aléatoire qui désigne le nombre moyen de contacts réalisés le jour n par un individu contagieux ce jour-là.

  On suppose, pour tout $n \in \mathbb{N}$, l'existence de $\mathbb{E}(R_n)$ et on pose $r_n = \mathbb{E}(R_n)$.
- Pour tout $n \in \mathbb{N}$, on note Z_n la variable aléatoire égale au nombre total d'individus qui sont infectés et donc deviennent contagieux le n-ième jour. Par exemple, $Z_0 = 1$.
- Pour tout $n \in \mathbb{N}$, on note I_n la variable aléatoire égale à la contagiosité globale de la population le n-ième jour, définie par :
$$I_n = \sum_{k=0}^{\min(n,d)} \alpha_k Z_{n-k} \tag{*}$$
- On suppose enfin que, pour tout $n \in \mathbb{N}$, I_n et R_n sont indépendantes et que si l'on pose $Y_n = R_n I_n$, on a :
$$Z_{n+1} \text{ suit la loi } \mathscr{P}(Y_n)$$
  où $\mathscr{P}$ désigne la loi de Poisson. Ainsi la loi de Z_{n+1} ne dépend que des lois de R_n et de I_n.

6. Donner une justification de $(*)$.
7. (a) Soit $n \in \mathbb{N}$. On suppose que $\mathbb{E}(I_n)$ existe. Montrer que $\mathbb{E}(Y_n)$ existe et en utilisant un résultat de la partie 1, montrer que $\mathbb{E}(Z_{n+1})$ existe et vaut $r_n \mathbb{E}(I_n)$.

   (b) Montrer que pour tout $n \in \mathbb{N}$, $z_n = \mathbb{E}(Z_n)$ existe et vérifie la relation de récurrence
$$z_{n+1} = r_n \sum_{k=0}^{\min(n,d)} \alpha_k z_{n-k} \tag{3}$$
8. **Programmation de z_n avec Python.**

   On suppose que la suite $(r_n)_{n \in \mathbb{N}}$ vérifie, pour tout $n \in \mathbb{N}$, $r_n = \dfrac{n+2}{n+1}$.

   On note Δ la matrice ligne $\begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_d \end{pmatrix}$.

   Écrire une fonction Python d'entête `z(Delta,n)` qui calcule z_n si `Delta` représente la matrice ligne Δ.
9. Soit $(U_n)_{n \geqslant 0}$, $(V_n)_{n \geqslant 0}$, deux suites d'événements tels que $\lim_{n \to +\infty} \mathbb{P}(U_n) = \lim_{n \to +\infty} \mathbb{P}(V_n) = 1$.

   Montrer que $\lim_{n \to +\infty} \mathbb{P}(U_n \cap V_n) = 1$.

On rappelle que l'on dit qu'un événement A **est presque sûr lorsque** $\mathbb{P}(A) = 1$.

10. On note pour tout $n \in \mathbb{N}^*$, $A_n = \bigcap_{k=n}^{+\infty} [Z_k = 0]$ et B l'événement « la contamination s'éteint au bout d'un nombre fini de jours ».

(a) Montrer que $\mathbb{P}(\mathrm{B}) = \lim_{n\to+\infty} \mathbb{P}(\mathrm{A}_n)$.

(b) En distinguant les cas où $\mathbb{P}\left(\bigcap_{k=n}^{n+d} [\mathrm{Z}_k = 0]\right)$ est nulle ou pas, établir que, pour tout $p \geqslant d$,

$$\mathbb{P}\left(\bigcap_{k=n}^{n+p} [\mathrm{Z}_k = 0]\right) = \mathbb{P}\left(\bigcap_{k=n}^{n+d} [\mathrm{Z}_k = 0]\right)$$

puis que $\mathbb{P}(\mathrm{A}_n) = \mathbb{P}\left(\bigcap_{k=n}^{n+d} [\mathrm{Z}_k = 0]\right)$.

(c) En déduire que B est presque sûr si et seulement si $\lim_{n\to+\infty} \mathbb{P}([\mathrm{Z}_n = 0]) = 1$.

(d) Montrer que cela équivaut aussi au fait que $(\mathrm{Z}_n)_{n\in\mathbb{N}}$ converge en loi vers 0.

11. (a) Montrer, en utilisant un résultat de la partie 1, que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}([\mathrm{Z}_{n+1} = 0]) = \mathbb{E}\left(e^{-\mathrm{Y}_n}\right)$$

(b) On suppose que $\lim_{n\to+\infty} z_n = 0$. En déduire que B est presque sûr (on pourra montrer que pour tout x reel, $e^{-x} \geqslant 1 - x$).

Partie 3 - Limite du nombre moyen de contaminations journalières

Dans cette partie, on conserve les notations de la partie 2 et on s'intéresse au comportement asymptotique de la suite $(z_n)_{n\in\mathbb{N}}$, definie par la relation (3) et $z_0 = 1$, sous trois hypothèses différentes concernant la suite $(r_n)_{n\in\mathbb{N}}$.
Pour tout réel x, on identifie x et la matrice carrée d'ordre 1 dont l'unique coefficient est x.
Pour tout $k \in [\![0, d]\!]$, on pose $a_k = \dfrac{\alpha_k}{\alpha}$.

12. On suppose, dans cette question, qu'il existe $\mathrm{N} \in \mathbb{N}$ et $\rho \in]0,1[$ tels que, pour tout $n \geqslant \mathrm{N}$, $r_n\alpha \leqslant \rho$.
On note (H_1) cette hypothese.

(a) Que vaut $\lim_{t\to 1} \sum_{k=0}^{d} a_k t^{d-k}$?

En déduire qu'il existe $\theta \in]0,1[$ tel que $\theta^{d+1} \geqslant \rho\left(\sum_{k=0}^{d} a_k \theta^{d-k}\right)$ (on pourra raisonner par l'absurde).

On pose $\mathrm{M} = \max_{k\in[\![\mathrm{N},\mathrm{N}+d]\!]} \dfrac{z_k}{\theta^k}$.

(b) Montrer que pour tout $n \geqslant \mathrm{N}$, $z_n \leqslant \mathrm{M}\theta^n$.

(c) En déduire que $\lim_{n\to+\infty} z_n = 0$.

On montrerait de même que s'il existe $\mathrm{N} \in \mathbb{N}$ *et* $\rho > 1$ *tels que, pour tout* $n \geqslant \mathrm{N}$, $r_n\alpha \geqslant \rho$, *on a* $\lim_{n\to+\infty} z_n = +\infty$.
On note (H_2) *cette hypothèse.*

On suppose, dans les questions 13 à 16 , que **la suite** $(r_n)_{n\in\mathbb{N}}$ **est constante de valeur** $\dfrac{1}{\alpha}$. On note (H_3) cette hypothèse.
On pose pour tout $n \in \mathbb{N}$,

$$\mathrm{U}_n = \begin{pmatrix} z_n \\ z_{n-1} \\ \vdots \\ z_{n-d} \end{pmatrix}$$

avec $z_{-1} = \ldots = z_{-d} = 0$.

13. (a) Montrer qu'il existe une matrice A carrée d'ordre $d+1$, de première ligne $\mathrm{L} = \begin{pmatrix} a_0 & a_1 & \ldots & a_d \end{pmatrix}$, telle que pour tout $n \in \mathbb{N}$, $\mathrm{U}_{n+1} = \mathrm{A}\mathrm{U}_n$.

(b) En déduire que, pour tout $n \geqslant 0$, $\mathrm{U}_n = \mathrm{A}^n\mathrm{U}_0$ puis que $z_{n+1} = \mathrm{L}\mathrm{A}^n\mathrm{U}_0$.

14. Dans cette question, $d = 2$ et $\mathrm{L} = \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}$.

   (a) Montrer que $\mathrm{Sp}(\mathrm{A}) = \left\{1, -\frac{1}{2}, -\frac{1}{3}\right\}$.

   (b) Déterminer une base $(\mathrm{V}_1, \mathrm{V}_2, \mathrm{V}_3)$ de $\mathcal{M}_{3,1}(\mathbb{R})$, où V_1 est un vecteur colonne propre de A pour la valeur propre 1, V_2 pour $-\frac{1}{2}$, V_3 pour $-\frac{1}{3}$, ces colonnes ayant leur premier coefficient égal a 1.

   (c) Déterminer $(s_1, s_2, s_3) \in \mathbb{R}^3$, tel que $\mathrm{U}_0 = s_1\mathrm{V}_1 + s_2\mathrm{V}_2 + s_3\mathrm{V}_3$.

   (d) En déduire que la suite $(z_n)_{n\in\mathbb{N}}$ converge vers s_1.

15. On revient au cas général.

   (a) Montrer que $\lambda \in \mathrm{Sp}(\mathrm{A})$ si et seulement si $\lambda^{d+1} = \sum_{k=0}^{d} a_{d-k}\lambda^k$ et que les sous-espaces propres de A sont de dimension 1.

   (b) Montrer que 1 est valeur propre de A et déterminer le vecteur colonne propre associé V dont la somme des composantes vaut $d+1$.

   (c) Établir que $-1 \notin \mathrm{Sp}(\mathrm{A})$ et que si $|\lambda| > 1$, alors $\lambda \notin \mathrm{Sp}(\mathrm{A})$.

16. On pose pour tout $k \in [\![0, d]\!]$, $b_k = \sum_{i=k}^{d} a_i$. On définit aussi le sous-espace vectoriel H de $\mathcal{M}_{d+1,1}(\mathbb{R})$ formé des matrices $\mathrm{W} = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{pmatrix}$ telles que $\sum_{k=0}^{d} b_k w_k = 0$.

   (a) Montrer que pour tout $\mathrm{W} \in \mathrm{H}$, $\mathrm{AW} \in \mathrm{H}$.

   (b) Déterminer l'unique réel s tel que $\mathrm{U}_0 - s\mathrm{V} \in \mathrm{H}$.

   (c) Nous admettons que, pour tout $\mathrm{W} \in \mathrm{H}$, $\mathrm{LA}^n\mathrm{W} \to 0$ quand $n \to +\infty$.
   En déduire que $\lim_{n\to+\infty} z_n = s$.

17. Sous quelle(s) hypothèse(s), parmi les trois hypothèses (H_1), (H_2) et (H_3) faites dans cette partie, la série $\sum_{n=0}^{+\infty} z_n$ est-elle convergente ? Comment interpréter ce résultat ?