Université Denis Diderot Paris 7

I. Branchements critiques. ($\star$)

Le but de ce projet est de comprendre l'évolution de la probabilité de survie sur n générations d'un processus de branchement critique, puis de simuler n générations d'un tel processus conditionné à survivre.

On considère un processus de branchement de Galton-Watson $(Z_n, n \geq 0)$, issu d'un unique ancêtre, et dont la loi de reproduction est ν.
Précisément, on introduit des variables $(\xi_{n,j})_{n\geq 0, j\geq 1}$ i.i.d suivant la loi ν : $\xi_{n,j}$ représente le nombre de descendants (à la génération $n+1$) du j-ème individu de la génération n ; et Z_n est le nombre total d'individus présents à la génération n.
On va se concentrer sur le cas particulier (*critique*) :

$$\sum_{k\geq 1} k\nu(k) = 1. \tag{1}$$

Pour éviter un cas dégénéré, on suppose par ailleurs que $\nu(1) < 1$.

1. Expliquer pourquoi

$$Z_0 = 1, \quad \text{et} \quad \forall n \geq 0,\ Z_{n+1} = \begin{cases} \sum_{j=1}^{Z_n} \xi_{n,j} \text{ si } Z_n \geq 1 \\ 0 \text{ sinon .} \end{cases}$$

   Montrer alors que $(Z_n)_{n\geq 0}$ est une chaîne de Markov.

2. Interpréter la condition (1)
3. On note $\zeta_n = \mathbb{P}(Z_n = 0)$, de sorte que ζ_n est la probabilité que la descendance de l'unique ancêtre s'éteigne avant la n-ème génération.
   (a) Dans cette question $\nu(0) = \nu(2) = 1/2$. Ecrire un programme qui permette de simuler $(Z_n)_{n\in\{0,N\}}$. Ecrire alors un programme qui permette d'estimer $(\zeta_n)_{n\in\{0,1000\}}$. Que remarque-t-on ?
   (b) Mêmes questions lorsque ν donnée par $\nu(k) = 1/k, \nu(0) = 1 - 1/k$.
   (c) Mêmes questions lorsque ν est la loi de Poisson de paramètre 1.
   (d) Mêmes questions lorsque ν est la loi de $G - 1$, où $G \sim \text{Geom}(1/2)$.
   (e) Que remarque-t-on dans tous les cas ?
4. (a) Montrer qu'on a toujours $\nu(0) > 0$.
   (b) On note

$$G_\nu(x) = \sum_{k\geq 0} \nu(k)x^k, \qquad G_n(x) = \mathbb{E}[x^{Z_n}], \qquad x \in [0,1].$$

   Montrer que

$$G_{n+1}(x) = G_\nu(G_n(x)).$$

(c) Montrer que

$$G_n(0) = \zeta_n, \ \forall n \geq 1.$$

Quelle est la limite de (ζ_n) ?

(d) Ecrire un programme qui permet de *calculer* $(\zeta_n)_{n\geq 0}$. Appliquer ce programme pour les exemples particuliers de mesures ν étudiés dans la partie précédente.

(e) Dans les différents cas, estimer la vitesse de convergence de (ζ_n) vers sa limite.

5. Pour chacun des exemples particuliers de mesures ν étudiés dans les parties précédentes, écrire un programme qui permette de simuler une trajectoire de $(Z_n, n \leq A)$ conditionnée à $Z_A \neq 0$.

   Représenter des réalisations de ces trajectoires pour $A = 10^6$.

   Qu'observe-t-on ?

II La plus longue série de succès consécutifs ($\star$) On considère un joueur qui effectue n parties. Le but de ce projet est d'estimer la plus longue série de succès consécutifs lors de ces n parties.

On notera systématiquement $X_i = 1$ si la i-ème partie est un succès, et $X_i = 0$ si cette i-ème partie est un échec.

On note la plus longue série de succès consécutifs lors des n parties :

$$L_n := \max\{i \geq 0 : \exists j \geq 0 \text{ tel que } j + i \leq n \text{ et } X_{j+1} = X_{j+2} = \cdots = X_{j+i} = 1\}.$$

1. Dans cette question on fait l'hypothèse que les $(X_i)_{i\geq 1}$ sont i.i.d suivant la loi de Bernoulli de paramètre p.
   (a) Que se passe-t-il lorsque $p = 0$? Que se passe-t-il lorsque $p = 1$?
   (b) Ecrire un programme qui simule L_n.
   (c) Simuler des histogrammes de L_n pour $p = 0.1, 0.2, ..., 0.9$ et $n = 100, 1000, 10^4, 10^5, 10^6$.
   Discuter.
   (d) On définit la longueur de la série initiale de succès par
   $$\ell_1 := \begin{cases} 0 \text{ si } X_1 = 0 \\ \max\{i \geq 1 : X_1 = X_2 = \cdots = X_i\} \text{ sinon.} \end{cases}$$
   Quelle est la loi de ℓ_1 ?
   Que vaut X_{ℓ_1+1} ?
   (e) Par récurrence, on définit alors la longueur de la k-ème série de succès. Si on a en effet déterminé $\ell_1, ..., \ell_k$, on définit
   $\tau_k = (\ell_1 + 1) + (\ell_2 + 1) + \cdots + (\ell_k + 1)$, et
   $$\ell_{k+1} := \begin{cases} 0 \text{ si } X_{\tau_k+1} = 0 \\ \max\{i \geq 1 : X_{\tau_k+1} = X_{\tau_k+2} = \cdots = X_{\tau_k+i} = 1\} \text{ sinon.} \end{cases}$$
   Quelle est la loi de $(\ell_k + 1)_{k\geq 1}$?
   (f) Soit $k_n = \max\{k : \tau_k \leq n\}$. Montrer que
   $$\frac{k_n}{n} \xrightarrow[n\to\infty]{\text{p.s.}} (1 - p).$$
   Montrer par ailleurs que
   $$\max\{\ell_k : 1 \leq k \leq k_n\} \leq L_n \leq \max\{\ell_k, 1 \leq k \leq k_n + 1\}.$$
   Montrer alors que
   $$\lim_{n\to\infty} \mathbb{P}\left(\max_{1\leq k\leq n(1-p)} \ell_k \leq -\alpha\frac{\log(n)}{\log(p)}\right) = \begin{cases} 0 \text{ si } \alpha < 1 \\ \exp(-(1-p)) \text{ si } \alpha = 1 \\ 1 \text{ si } \alpha > 1. \end{cases}$$

(g) Déduire des questions précédentes la limite presque sûre, lorsque $n \to \infty$, de $\frac{L_n}{\log(n)}$. Comparer avec les résultats obtenus par simulation.

2. Dans cette partie on suppose que le succès d'une partie donnée dépend du résultat des parties précédentes : on note Y_n la valeur de la série en cours (i.e. $Y_n = \max\{k : X_{n-1} = ... = X_{n-k} = 1\}$) et on suppose que

$$\mathbb{P}(X_n = 1 \mid \mathcal{F}_{n-1}) = \begin{cases} \frac{1}{2} \text{ si } Y_n = 0 \\ 1 - \frac{1}{2Y_n^\alpha} \text{ si } Y_n > 0 \end{cases}.$$

(a) Le processus $(X_n, n \geq 0)$ est-il une chaîne de Markov ? Si oui, quel est son noyau de transition ?

(b) Le processus $(Y_n, n \geq 0)$ est-il une chaîne de Markov ? Si oui, quel est son noyau de transition ?

(c) Montrer qu'une mesure stationnaire ν de $(Y_n, n \geq 0)$ doit vérifier, pour tout $k \geq 0$,

$$\nu(k+1) = \left(1 - \frac{1}{2k^\alpha}\right) \nu(k)$$

Montrer que $(Y_n, n \geq 0)$ possède une probabilité stationnaire ssi $\alpha \leq 1$.

(d) Simuler des histogrammes de la longueur de la plus longue série de succès lors de n parties, pour $\alpha = 1/2, 1, 2$.

Discuter les résultats obtenus.

III. Puissance du choix Le but de ce projet est de comprendre l'influence d'un choix limité sur les placements successifs de n boules dans n urnes.

1. Dans cette partie on suppose que les placements sont i.i.d suivant la loi uniforme. Pour $1 \leq k \leq n$ on note $B_k^{(n)}$ le nombre de boules qui se trouvent dans l'urne k à l'issue des n placements. On note par ailleurs $M_n = \max\{B_k^{(n)}, 1 \leq k \leq n\}$.
   (a) Pour $k \in \{1, ..., n\}$, quelle est la loi de $B_k^{(n)}$?
   (b) Montrer que pour tout $k \geq 1$,
   $$B_k^{(n)} \xrightarrow[n\to\infty]{\text{(loi)}} B_k,$$
   où $B_k \sim \text{Poisson}(1)$.
   (c) Quelle est la loi du n-uplet $(B_1^{(n)}, ..., B_n^{(n)})$?
   (d) Montrer finalement que pour tous entiers $k_1, ..., k_j$,
   $$(B_{k_1}^{(n)}, ..., B_{k_j}^{(n)}) \xrightarrow[n\to\infty]{\text{(loi)}} (B_{k_1}, ..., B_{k_j}),$$
   où les $B_k, k \geq 1$ sont i.i.d suivant la loi de Poisson de paramètre 1.
   (e) Ecrire un programme qui simule les n placements. En déduire un programme qui permet de vérifier le résultat de la question précédente.
   (f) Représenter un histogramme de M_n, pour $n = 10, 100, ..., 10^7$.
   (g) Soient $(B_k, k \geq 1)$ i.i.d suivant la loi de Poisson(1). On note
   $$\mathcal{M}_n = \max\{B_1, ..., B_n\}.$$
   Les variables $\mathcal{M}_n$ et M_n ont-elles la même loi ?

   On admet que
   $$\mathbb{P}(B_1 > n) \sim_{n\to\infty} \frac{c}{n^{3/2}} \exp(-n\log(n)).$$
   En fonction de $\alpha > 0$, trouver
   $$\lim_{n\to\infty} \mathbb{P}\left(\mathcal{M}_n \leq \alpha \frac{\log(n)}{\log\log(n)}\right).$$

   Que vaut la limite presque sûre de $\mathcal{M}_n \frac{\log\log(n)}{\log(n)}$ lorsque $n \to \infty$?
2. Dans cette partie on suppose que pour chaque placement de boule, on pré-sélectionne deux urnes de façon uniforme parmi les n, puis on place la boule dans celle des deux qui est la *plus remplie*.
   A nouveau on note M_n l'urne la plus remplie à l'issue des n placements.

(a) Ecrire un programme qui simule les n placements.

(b) Représenter un histogramme de M_n, pour $n = 10, 100, ..., 10^7$.

3. Dans cette partie on suppose que pour chaque placement de boule, on pré-sélectionne deux urnes de façon uniforme parmi les n, puis on place la boule dans l'urne *la moins remplie* des deux.

   A nouveau on note M_n l'urne la plus remplie à l'issue des n placements.

   (a) Ecrire un programme qui simule les n placements.

   (b) Représenter un histogramme de M_n, pour $n = 10, 100, ..., 10^7$.

4. Comparer et discuter les résultats obtenus pour les différents types de placement.

IV. Ising ($\star$) Le but de ce projet est d'effectuer une simulation du modèle d'Ising à l'aide de l'algorithme de Metropolis-Hastings.

On considère le modèle d'Ising bi-dimensionnel sur le tore $\mathbb{Z}^2/N$, qui est un modèle pour la magnétisation d'un métal.
En chaque site $x \in \mathbb{Z}^2/N$ se trouve une particule dont le spin $\sigma(x) \in \{-1, +1\}$.
Une *configuration de spins* σ correspond à la donnée des N^2 spins. On suppose que l'énergie d'une configuration σ est donnée par

$$H(\sigma) = - \sum_{\{(i,j) \text{ voisins }\}} \sigma_i \sigma_j.$$

A la température T, la *mesure de Gibbs* d'une configuration σ est donnée par

$$\mu_T(\sigma) = \frac{\exp\left(-\frac{H(\sigma)}{T}\right)}{Z_T}, \qquad \text{où } Z_T = \sum_\sigma \exp\left(-\frac{H(\sigma)}{T}\right).$$

1. Considérations préliminaires
   (a) Quel est le nombre total de configurations de spins ?
   (b) Si σ_i et σ_j sont i.i.d, $\sim$ Unif$\{-1, 1\}$, quelle est la loi de $\sigma_i \sigma_j$?
   (c) Si on tire les spins de façon i.i.d, $\sim$ Unif$\{-1, 1\}$, quel est l'ordre de grandeur de $H(\sigma)$?
   (d) Si on tire une configuration de spins suivant la mesure de Gibbs, quelles configurations favorise-t-on ? Quel est l'effet de la température sur ce "favoritisme" ?
   (e) Quelles sont les configurations d'énergie minimale ? Pourquoi jouent-elles un rôle particulier lorsque $T \to 0$?
   (f) Que se passe-t-il lorsque $T \to \infty$?
2. Si σ est une configuration on note σ^x la configuration obtenue depuis σ en retournant le spin en x : $\sigma^x(y) = \begin{cases} \sigma(y) \text{ si } x = y \\ -\sigma(y) \text{ si } x = y \end{cases}$.

   On considère alors la chaîne de Markov $(\sigma_n, n \geq 0)$ à valeurs dans l'espace des configurations, et dont le noyau de transition est,

$$Q(\sigma, \sigma^x) = \frac{1}{N^2} \begin{cases} \exp(-\frac{H(\sigma^x)-H(\sigma)}{T}) \text{ si } H(\sigma^x) - H(\sigma) > 0 \\ 1 \text{ sinon,} \end{cases}$$

   pour tout σ configuration de spins.
   (a) La chaîne est-elle irréductible ? apériodique ? Montrer que la mesure de Gibbs est une probabilité invariante de $(\sigma_n, n \geq 0)$. Y a-t-il d'autres probabilités invariantes ?

(b) Simuler la chaîne lorsque $N = 10, 50, 100, 200, 300, 400, 500$, et différentes valeurs de T. Qu'observe-t-on ?

(c) Estimer la covariance entre 2 spins de sites voisins, pour différentes valeurs de N et T. Estimer la covariance de 2 spins de sites éloignés (e.g. à distance $N/2$), pour différentes valeurs de N et T. Qu'observe-t-on ? Voyez-vous un lien avec les observations de la question précédente ?

V. Mélange de cartes. ($\star$) Le but de ce projet est de comprendre quand un jeu de cartes est proprement mélangé.

On dispose d'un jeu de r cartes placées dans une pile. On a numéroté les cartes de 1 à r : la carte numéro 1 est celle initialement en haut du tas.

Une étape de notre mélange consiste en la démarche suivante :

- On choisit $M \sim \mathrm{Bin}(r, 1/2)$, et on coupe le jeu au niveau de la M-ème carte (i.e. on sépare les M premières cartes des $r - M$ dernières).
- On "re-mélange" nos 2 paquets (on préserve, bien sûr l'ordre relatif dans chacun des 2 sous-paquets) en choisissant uniformément parmi les $\binom{r}{M}$ façons de le faire.

On note $(X_n, n \geq 0)$ la suite de permutations de $\{1, ..., r\}$ obtenue.

Une sous-suite croissante de longueur n d'une permutation $\sigma \in \mathfrak{S}_r$ est un ensemble d'indices $i_1 < ... < i_n$ tels que

$$\sigma(i_1) < \sigma(i_2) < \cdots < \sigma(i_n).$$

Une *plus longue sous-suite croissante* (p.l.s.s.c.) de longueur n est un tel ensemble d'indices maximal au sens de l'inclusion
(autrement dit si $j \in (i_k, i_{k+1})$, $\sigma(j) \notin (\sigma(i_k), \sigma(i_{k+1}))$).

Par exemple la permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 4 & 7 & 1 & 8 & 10 & 2 & 6 & 3 & 9 & 5 \end{pmatrix}$$

possède, par exemple, les trois p.l.s.s.c. suivantes :

$$\begin{aligned} &\sigma(1) = 4 < \sigma(2) = 7 < \sigma(4) = 8 < \sigma(5) = 10, \\ &\sigma(3) = 1 < \sigma(6) = 2 < \sigma(8) = 3 < \sigma(10) = 5 \\ &\sigma(7) = 6 < \sigma(9) = 9. \end{aligned}$$

1. (a) Montrer que $(X_n), n \geq 0$ est une chaine de Markov. Montrer que son noyau de transition vérifie $P(\sigma_1, \sigma_2\sigma_1) = P(Id, \sigma_2)$ pour tous $\sigma_1, \sigma_2 \in \mathfrak{S}_r$, et que

$$P(Id, \sigma) = \begin{cases} \frac{r+1}{2^r} \text{ si } \sigma = Id, \\ \frac{1}{2^n} \text{ si } \sigma \text{ possède exactement 2 p.l.s.s.c.,} \\ 0 \text{ sinon.} \end{cases}$$

   (b) Montrer que (X_n) est irréductible, apériodique et récurrente. Identifier sa loi invariante.

   (c) Ecrire un programme qui simule ce mélange de cartes.

2. Soit la procédure suivante : à chaque carte du jeu on attribue l'étiquette $\xi_i \sim \mathrm{Ber}(1/2)$; on replace alors les cartes étiquetées 0 en haut du paquet (en respectant leur ordre relatif).

   On note $(Y_n, n \geq 0)$ la suite de permutations de $\{1, ..., r\}$ obtenue.

(a) Montrer que $(Y_n), n \geq 0$ est une chaine de Markov. Montrer que son noyau de transition Q vérifie, pour tous $\sigma_1, \sigma_2 \in \mathfrak{S}_r$,

$$Q(\sigma_2, \sigma_1) = P(\sigma_1, \sigma_2).$$

Expliquer le lien entre les chaînes X et Y.

(b) Ecrire un programme qui simule la chaîne Y.

3. (a) Dans cette partie uniquement on ne considérera que $r = 3, 4, 5$. On note P_n la loi de X_n, et π la mesure invariante trouvée en 1.b. Représenter le graphe de

$$n \to \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_r} |P_n(\sigma) - \pi(\sigma)|.$$

Que remarque-t-on ?

(b) Même question avec Q_n, loi de Y_n.

4. On admet que

$$\sum_{\sigma \in \mathfrak{S}_r} |P_n(\sigma) - \pi(\sigma)| = \frac{1}{2} \sum_{k=1}^{r} E(r,k) \left| \frac{\binom{2^n+r-k}{k}}{2^{nr}} - \frac{1}{r!} \right|,$$

avec

$$E(1,1) = 1, \quad E(1,k) = 0 \text{ pour tout } k \neq 1,$$
$$E(r,k) = (r-k+1)E(r-1,k-1) + kE(r-1,k) \text{ pour tous } r \geq 2,\ k \geq 1.$$

(a) Ecrire un programme qui calcule les valeurs des $E(r,k)$, pour $r = 10, 52, 100$, et $k \in \{1, ..., r\}$.

(b) Utiliser la formule admise pour représenter le graphe de

$$n \to \frac{1}{2} \sum_{\sigma \in \mathfrak{S}_{10}} |P_n(\sigma) - \pi(\sigma)|,$$

à nouveau pour $r = 3, 4, 5$ puis pour $r = 10, 52$ et enfin pour $r = 100$. Discuter.

VI. Marche λ-biaisée sur un arbre binaire. ($\star$) Le but de ce projet est de comprendre quelques propriétés, et estimer certaines quantités telles que le temps de couverture, de marches aléatoires sur un arbre binaire fini en fonction d'un paramètre de biais $\lambda \in (0,1)$.

On considère un arbre binaire $\mathcal{T}$ de hauteur $N \in \mathbb{N}^*$. On pourra par exemple considérer qu'un tel arbre est constitué par

- un premier noeud que l'on note $\emptyset$ et qu'on place à hauteur nulle. Ce noeud n'a pas de parent, on l'appelle donc "racine" de l'arbre.
- deux noeuds à hauteur 1, que l'on note 1 et 2, et qui sont chacun directement reliés à la racine.
- A hauteur 2 on compte 4 noeuds. Les deux descendants issus du noeud 1 seront notés $11, 12$, tandis que les deux descendants de 2 seront notés $21, 22$.
- pour tout $k \in \{1, ..., h-1\}$ on compte 2^k noeuds à hauteur k. Les deux descendants du noeud u de hauteur $k-1$ sont notés $u1, u2$.
- A la hauteur N, on compte 2^N noeuds, mais ils n'ont pas de descendants (on parlera donc de "feuilles" pour les noeuds situés à la hauteur h). Comme précédemment les deux noeuds descendantes d'un noeud u de hauteur $h-1$ sont notées $u1, u2$.

On notera que la hauteur d'un noeud u est tout simplement le nombre de chiffres utilisés pour écrire u. On note hauteur$(u) = |u|$. Un tel noeud possède un unique parent noté $p(u)$, que l'on obtient à partir de u en effaçant le dernier chiffre.

On fixe alors $\lambda \in (0,1)$ et on considère alors la marche aléatoire λ-biaisée $(X_n, n \geq 0)$ sur cet arbre, démarée en la racine de l'arbre. Cette marche est une chaîne de Markov à valeurs dans $\mathcal{T}$. Au temps $n+1$, X saute de X_n vers l'un des plus proches voisins de X_n choisis au hasard de la façon suivante :

- si $X_n = \emptyset$, alors

$$X_{n+1} = \begin{cases} \emptyset \text{ avec probabilité } \lambda \\ X_{n+1} = 1 \text{ avec probabilité } (1-\lambda)/2 \\ X_{n+1} = 2 \text{ avec probabilité } (1-\lambda)/2 \end{cases}.$$

- Si $X_n = u$, avec $1 \leq |u| \leq N-1$, alors

$$X_{n+1} = \begin{cases} p(X_n) \text{ avec probabilité } \lambda \\ X_{n+1} = X_n 1 \text{ avec probabilité } (1-\lambda)/2 \\ X_{n+1} = X_n 2 \text{ avec probabilité } (1-\lambda)/2 \end{cases}.$$

- Si $|X_n| = N$ alors $X_{n+1} = p(X_n)$.

1. (a) Montrer que $(X_n, n \geq 0)$ est irréductible, apériodique, et calculer sa distribution invariante en fonction de λ et N.

   (b) Simuler $(X_k, 0 \leq k \leq n)$ et vérifier les calculs de la question précédente, pour des valeurs de $N \in \{1, ..., 10\}$, et pour $\lambda = 1/5, 1/3, 2/5, 1/2, 3/5, 4/5$.

(c) Montrer que $(Y_n := |X_n|, n \geq 0)$ est une chaîne de Markov, calculer son noyau de transition. Montrer que cette chaîne est irréductible, apériodique, et calculer sa distribution invariante en fonction de λ, N. Que vaut $\mathbb{E}_0(\tau_0^+)$, lorsque $\tau_0^+ = \inf\{k \geq 1 : Y_k = 0\}$?

(d) Simuler directement $(Y_n, n \geq 0)$. Estimer $\mathbb{E}_0[\tau_0^+]$.

(e) On note $\tau_j = \inf\{k \geq 0 : Y_k = j\}$, et $p_k := \mathbb{P}_k(\tau_0 < \tau_N)$. Que valent p_0, p_N ? Montrer que pour tout $k \in \{1, ..., N-1\}$, $p_k = \lambda p_{k-1} + (1-\lambda) p_{k+1}$. Ecrire alors un programme qui calcule les différentes valeurs de $p_k, k \in \{0, ..., N\}$, et tracer le graphe correspondant pour $N = 1, 2, ..., 10, 20, ..., 100, 200, ..., 1000$ et pour les valeurs de $\lambda = 1/3, 0.4, 0.5, 0.6$.

(f) Vérifier les calculs de la question précédente à l'aide de simulations de la chaîne $(Y_n, 0 \leq n \leq \min(\tau_0, \tau_n))$.

2. Dans cette partie on fixe $\lambda = 1/3$ (cas de la marche simple sur l'arbre).

(a) Soit $\tau_u = \inf\{k \geq 0 : X_k = u\}$.
Pour u une feuille de l'arbre, simuler la loi de $\tau_u/(2^{N+1}N)$.

(b) On s'intéresse alors au temps nécessaire pour que la marche $(X_n)_{n\geq 0}$ soit passée au moins une fois par tous les noeuds de l'arbre :

$$\tau_{\text{couv}} = \inf\{n \geq 0 : \forall u \in \mathcal{T}\ \exists k \leq n : X_k = u\}.$$

On rappelle que notre marche est issue de la racine de l'arbre.
Simuler pour $N = 3, 4, ..., 10, 15, 20$ la loi de $\tau_{\text{couv}}/(2^{N+1}N^2)$. Que remarque-t-on ?

3. Par simulation, et pour les mêmes valeurs de N, estimer les quantités τ_u, τ_{couv} définies en 2, mais cette fois pour $\lambda = 1/5, 2/5, 1/2, 3/5, 4/5$. Discuter suivant ces valeurs de λ.

VII. Modèle d'opinion sur $\mathbb{Z}^2/N$ Le but de ce projet est l'étude du modèle suivant, introduit initialement pour modéliser l'évolution d'opinions.

On suppose que $N \in \mathbb{N}^*$, et qu'en tout point de $\mathbb{Z}^2/N$ (le tore discret de dimension 2) se trouve une personne, qui possède, en tout temps, une certaine opinion.

Pour faire simple on suppose qu'il n'y a que deux opinions, que l'on note 0 et 1. L'opinion au temps t de la personne située au noeud i est notée $\xi_t(i)$.

Le mécanisme d'évolution est le suivant. A chaque temps discret $t \in \mathbb{N}^*$, on tire, indépendamment des étapes précédentes, une personne uniformément au hasard dans $\mathbb{Z}^2/N$, disons $j(t)$.

La personne $j(t)$ décide alors de consulter l'un de ses 4 voisins dans $\mathbb{Z}^2/N$, et adopte l'opinion de ce voisin, quelle qu'elle soit. On suppose que le choix de voisin est également uniforme, et indépendant des autres étapes. Les opinions des personnes autres que $j(t)$ restent inchangées au temps t.

On dit qu'il y a consensus lorsque les N personnes partagent la même opinion. On note

$$\tau_{\text{consensus}} := \inf\{t \in \mathbb{N} : \prod_{x\in\{1,\dots,N\}} \xi_t(x) = 1 \text{ ou } \prod_{x\in\{1,\dots,N\}} (1-\xi_t(x)) = 1\}.$$

1. Dans cette partie et la suivante, les opinions initiales sont tirées indépendamment, avec
$$\xi_0(i) \sim \mathrm{Ber}(p),$$
pour un certain $p \in (0,1)$ fixé.
   (a) Montrer que $\xi = (\xi_t(x), x \in \mathbb{Z}^2/N, t \in \mathbb{N})$ est une chaîne de Markov à valeurs dans $\{0,1\}^{\{\mathbb{Z}^2/N\}}$. Quel est son noyau de transition ?
   (b) En observant que les deux cas de consensus sont absorbants, donner deux exemples de lois invariantes pour ξ. Quelle est l'hypothèse du théorème sur l'existence d'une unique mesure invariante qui est ici mise en défaut ?

2. (a) Ecrire un programme qui permet, étant donné N et $\xi_0 \in \{0,1\}^{\mathbb{Z}^2/N}$, de simuler ξ_1.
   (b) Simuler alors $(\xi_t)_t$ pour différentes valeurs de N, e.g. $5, 10, 30, 50, 100$. On représentera le film de cette évolution au cours du temps (pour $N = 5, 10, 30$, on considérera une plage de temps assez large pour parvenir au consensus. Pour $N = 50, 100$, on tentera d'effectuer les simulations sur les plages de temps les plus larges possibles).
   (c) Estimer alors, pour ces différentes valeurs de N et en fonction de p, la probabilité que l'opinion 1 "l'emporte" (i.e. le consensus s'établit sur l'opinion 1).
   (d) Estimer, en fonction de N et p, l'espérance du temps de consensus.

3. Dans cette question, on suppose qu'initialement seule la personne placée à l'origine possède l'opinion 1.

   (a) Pour $N = 5, 10, 30, 50, 100$, effectuer une simulation qui permette d'ébaucher le graphe de

   $$t \to \mathbb{P}(\text{l'opinion 1 est présente au temps } t).$$

   (b) Pour $N = 30$, faire un film de l'évolution de $(\xi_t, t \in \mathbb{N})$ conditionnée à l'événement $\{$l'opinion 1 survit éternellement$\}$.

VIII. Modèle d'épidémie sur le graphe complet ($\star$) Le but de ce projet est l'étude d'un modèle d'épidémie sur le graphe complet à N sommets.

Les sommets de notre graphe complet sont notés $\{1, ..., N\}$. On rappelle que dans le graphe complet, tous les noeuds sont voisins de chacun des autres noeuds.

En chaque noeud du graphe se situe une bactérie, elle est dans l'un des 3 états : susceptible (S), infectée (I), rémise (R).

Une bactérie peut passer de l'état S à l'état I, et de l'état I à l'état R, aucune autre transition n'est possible.

Le système à l'instant $t \in \mathbb{N}$ est donc décrit par $(\xi_t(j), j \in \{1, ..., N\}) \in \{S, I, R\}^N$ la donnée des états des N bactéries.

Sauf mention contraire, on suppose qu'initialement toutes les bactéries sont dans l'état S, sauf la bactérie située au noeud 1 qui est dans l'état I.

Le mécanisme d'évolution est le suivant. On fixe un paramètre $q \in [0, 1]$.

A chaque temps discret $t \in \mathbb{N}^*$, on tire trois variables $(j(t), k(t), B(t))$, indépendamment des étapes précédentes. Les variables $j(t)$ et $k(t)$ sont uniformes sur $\{1, ..., N\}$, tandis que $B(t) \sim \mathrm{Ber}(q)$.

Les états des bactéries qui ne sont situées ni en $j(t)$ ni en $k(t)$ sont inchangés (i.e. si $\ell \notin \{j(t), k(t)\}$ on a toujours $\xi_{t+1}(\ell) = \xi_t(\ell)$).

Intéressons nous désormais à l'évolution entre t et $t + 1$ des états des bactéries situées en $j(t)$, $k(t)$.

Si $\xi_t(j(t)) \in \{S, R\}$, il ne se passe rien.

Sinon, i.e si $\xi_t(j(t)) = I$,

— si $B(t) = 1$, la bactérie placée en $j(t)$ risque d'infecter la bactérie placée en $k(t)$. Autrement dit, si on est dans ce cas et si $\xi_t(k(t)) = S$, alors $\xi_{t+1}(k(t)) = I$. Si on est dans ce cas mais que $\xi_t(k(t)) \in \{I, R\}$, il ne se passe rien.

— Si $B(t) = 0$, la bactérie placée en $j(t)$ effectue une rémission. Autrement dit, dans ce cas $\xi_{t+1}(j(t)) = R$.

Enfin on note

$$\tau = \inf \left\{ t \geq 0 : \xi_t(k) \neq I \ \forall k \in \mathbb{Z}^2/N \right\}.$$

1. Que se passe-t-il si $q = 0$? Que se passe-t-il si $q = 1$? Dans la suite on suppose que $q \in (0, 1)$.
2. Que se passe-t-il après le temps τ ?
3. (a) Montrer que $\xi = (\xi_t(x), x \in \{1, ..., N\}, t \in \mathbb{N})$ est une chaîne de Markov à valeurs dans $\{S, I, R\}^{\{1,...,N\}}$. Quel est son noyau de transition ?

   (b) Donner des exemples de lois invariantes pour ξ. Quelle est l'hypothèse du théorème sur l'existence d'une unique mesure invariante qui est ici mise en défaut ?

   (c) On note $(N_S(t), N_I(t), N_R(t))$ les nombres de bactéries susceptibles, infectées et rémises au temps t. Montrer que $(\mathcal{N}(t) := (N_S(t), N_I(t)))_{t\in\mathbb{N}}$ est une chaîne de Markov. Quel est son espace détat ? Quel est son noyau de transition ? Est-elle irréductible ?

(d) Expliquer pourquoi la chaîne $\mathcal{N}$ contient toutes les informations pertinentes concernant l'évolution de la chaîne ξ.

4. (a) Supposons que la chaîne $\mathcal{N}$ est dans l'état (i, j). Rappeler les possibles transitions depuis l'état (i, j).

Montrer que le temps qu'il faut à la chaîne pour changer d'état est

$$G_{i,j} \sim \text{Geom}(\frac{ijq}{N^2} + \frac{(1-q)j}{N}).$$

Au moment où la chaîne quitte effectivement l'état (i, j), donner les probabilités qu'elle atteigne l'état (k, ℓ), en fonction de i, j, k, ℓ.

(b) A l'aide de la question précédente, pour différentes valeurs de N (e.g. $N = 10, 30, 100, 300, 1000$), et de q (e.g. $q = 0.2, 0.4, 0.6, 0.8$) simuler plusieurs trajectoires de $(\mathcal{N}(t), t \geq 0)$.

Dresser, pour ces différentes valeurs de N et q, un histogramme de $N_S(\tau)$.

(c) On considère qu'il y a pandémie dès lors que plus de 20% de la population a été infectée. Estimer la probabilité de pandémie pour $N = 1000$ et différentes valeurs de q. Représenter le graphe de cette probabilité en fonction de q.

(d) Une définition plus restrictive de pandémie consiste à considérer qu'elle n'a lieu que lorsque plus de 20% de la population est touchée *simultanément.* Reprendre la question précédente avec cette définition plus restrictive.

(e) Dresser un histogramme de τ (on se servira de 3 (a)).

IX. Modèle d'Eden-Richardson

On considère une suite E_n de sous ensembles de $\mathbb{Z}^d$. Initialement, $E_1 = \{0\}$, puis à chaque instant n, chaque point ayant au moins un plus proche voisin dans E_n est ajouté à E_n avec probabilité p (indépendamment des autres points).

1. Dans cette question, on considère le cas $d = 1$.
   (a) Ecrire un programme qui simule l'évolution de E_n jusqu'à l'étape 100000.
   (b) Montrer que l'ensemble des sites occupés est un intervalle d'entiers, puis vérifier que
   $$\mathbb{E}(\#E_n) = 1 + 2p(n-1)$$
   (c) On écrit $E_n = \{G_n, \cdots, D_n\}$. Donner la loi de D_n et expliquer pourquoi on a
   $$\lim_{n\to\infty} \frac{D_n}{n} = p.$$
2. On se place désormais dans le cas $d = 2$.
   (a) Réaliser une simulation dans le cas où $p = 1$, pour $n = 100$. Que remarque-t-on ? Expliquer.
   (b) Réaliser des simulations pour différentes valeurs de p. On pourra prendre, par exemple, $p = 0.1, 0.3, 0.5, 0.7, 0.9$. Que remarque-t-on ? Commenter. Pour l'un de ces p, on pourra réaliser un film/gif animé représentant la croissance de E_n lorsque n croît (à l'aide par exemple de la commande xs2gif, voir l'aide de scilab).
   (c) Pour $p = 0.7$, réaliser une simulation qui compte le nombre de points de E_{100}, puis de E_{200}. Que vaut $\frac{\#E_{200}}{\#E_{100}}$? Expliquer pourquoi il est raisonnable de penser que $\#E_n$ croît proportionnellement à n^2.
3. On se place toujours dans le cas $d = 2$, et on s'intéresse à l'ensemble $\mathcal{E}_n = \{(x, y) \in E_n, x + y = n\}$.
   (a) Pour x compris entre 0 et n, on note $Z_n(x) = 1$ si $(x, n - x) \in E_n$, et $Z_n(x) = 0$ sinon. Montrer que Z_n est une chaîne de Markov, dont les transitions sont données par :
   $$\begin{aligned}
   \mathbb{P}(Z_n(x) = 0 | Z_{n-1}(x-1) = Z_{n-1}(x) = 0) &= 1 \\
   \mathbb{P}(Z_n(x) = 0 | Z_{n-1}(x-1) = 1 \text{ ou } Z_{n-1}(x) = 1) &= 1 - p \\
   \mathbb{P}(Z_n(x) = 1 | Z_{n-1}(x-1) = 1 \text{ ou } Z_{n-1}(x) = 1) &= p
   \end{aligned}$$
   (b) Simuler Z_n pour $p = 0.1$, puis pour $p = 0.9$. Commenter.
4. On considère un arbre généalogique $\mathcal{T}$ construit de la façon suivante : à la première génération, l'arbre ne comporte qu'un individu. À chaque génération, chaque individu a un enfant avec probabilité p, deux avec probabilité p^2, et aucun avec probabilité $1 - p - p^2$. On note T_n le nombre d'enfants à la génération n.

(a) Calculer $\mathbb{E}(T_2)$. Montrer que $\mathbb{E}(T_{n+1}|T_n = k) = (p + 2p^2)k$. En déduire que $\mathbb{E}(T_n) = (p + 2p^2)^n$.

(b) En utilisant l'inégalité de Markov, montrer que $\mathbb{P}(T_n > 0) \leq (p + 2p^2)^n$. En déduire que si $p < 1/2$, la probabilité d'extinction de l'arbre vaut 1.

(c) Expliquer pourquoi on peut comparer $\#\mathcal{E}_n$ et T_n, et en déduire que si $p < 1/2$, $\mathcal{E}_n$ est vide à partir d'un certain rang.

(d) Vérifier ce résultat à l'aide de simulations. Pour quelles valeurs de p semble-t-on avoir $\mathcal{E}_n$ non vide pour tout n ?

X. Bootstrap Percolation

Le Baron de Münchausen, embourbé dans un marécage, se serait d'après la légende tiré d'affaire en tirant sur ses propres attaches de bottes (*bootstraps* en anglais), ce qui l'aurait bien entendu projeté en l'air. Depuis, le terme *bootstrap* désigne la façon dont un système peut s'amorcer puis évoluer par lui-même à partir d'un état initial.

On se place sur $\mathbb{Z}^d$ et on considère l'état initial où chaque point est présent (valeur 1) avec probabilité p, et absent (valeur 0) avec probabilité $1-p$. On suppose de plus que tous ces évènements sont indépendants. On appelle une telle configuration une *percolation* de paramètre p.

On considère la règle d'évolution suivante : à l'étape n, tous les points absents qui ont une majorité (au sens large) de plus proches voisins présents deviennent présents pour l'étape $n+1$. On note B_n l'ensemble à l'étape n.

1. Dans cette question, on considère le cas $d = 1$.
   (a) Écrire un programme qui simule le comportement de B_n sur l'intervalle d'entiers $[1, 1000]$.
   (b) Décrire l'évolution depuis l'état initial. Que remarque-t-on ?
   (c) On considère les suites H_i et H'_i définies par récurrence de la façon suivante.
   $$\begin{aligned} H_1 &= \inf\{i \geq 0, i \notin B_0\}, \\ H'_i &= \sup\{k \geq H_i, \forall j \in \{H_i, \cdots, k\}, j \notin B_0\} \text{ pour } i \geq 1, \text{ et} \\ H_i &= \sup\{k > H'_i, \forall j \in \{H'_{i-1}+1, \cdots, k-1\}, j \in B_0\} \text{ pour } i > 1. \end{aligned}$$
   Que peut-on dire de $\{0, \cdots, H_1 - 1\}$? de $\{H_i, \cdots, H'_i\}$? de $\{H'_i + 1, \cdots H_{i+1} - 1\}$?
   (d) On note $G_1 = H_1, G'_i = H'_i - H_i + 1, G_i = H_{i+1} - H'_i$. Vérifier que G_i est le nombre de points absents dans chaque composante connexe de points absents (la première peut être vide), et que G'_i est le nombre de points présents dans chaque composante connexe de points présents.
   (e) Montrer que les $G_i, i > 1$ suivent une loi géometrique de paramètre $1-p$. Montrer de même que les $G'_i, i \geq 0$ suivent une loi géométrique de paramètre p. On remarquera aussi que $G_1 + 1$ suit une loi géométrique de paramètre $1-p$.
   (f) En déduire que pour $i \neq 1$, $\mathbb{E}(G_i + G'_i) = \frac{1}{p} + \frac{1}{1-p}$. Soit N le plus petit entier tel que $H_N \geq n$ ou $H'_N \geq n$. En utilisant la loi des grands nombres, montrer que
   $$N = \frac{n}{\frac{1}{p} + \frac{1}{1-p}} + o(n).$$
   (g) Pour $k \in \mathbb{N}$, donner la fonction de répartition de $\max_{i=1..k} G'_i$. Montrer que $\max_{i=1..N} G'_i$ est d'ordre $\ln(n)$.

(h) On appelle τ_n le temps de remplissage complet de l'intervalle $[1, n]$. Montrer que $\tau_n = \max_{i=1..N} G'_i/2$. Réaliser un histogramme de $\tau_n/ln(n)$ pour $n = 10000$ et commenter.

2. On se place maintenant dans le cas $d = 2$.

   (a) Simuler le comportement de B_n dans le tore $[1, 1000] \times [1, 1000]$, avec différentes valeurs de p. Commenter. Pour l'un de ces p, on pourra réaliser un film/gif animé représentant la croissance de B_n lorsque n croît (à l'aide par exemple de la commande xs2gif, voir l'aide de scilab).

   (b) Y a-t-il des valeurs de p pour lesquelles B_n ne finit pas par recouvrir le tore tout entier ? Expliquer pourquoi on s'attend à ce que la probabilité de l'évènement "le tore entier est recouvert en un temps fini" soit croissante. Simuler les valeurs de cette probabilité pour différentes valeurs de p.

   (c) Réaliser des simulations pour différentes tailles de tore. La taille du tore a-t-elle un effet sur la probabilité de la question précédente ?

3. On considère, pour $d = 2$, la modification suivante du modèle. L'initialisation du modèle reste la même, mais un site (x, y) devient présent si trois au moins des sites $(x - 1, y), (x + 1, y), (x, y - 1), (x, y - 2), (x, y + 1), (x, y + 2)$ sont présents.

   (a) Simuler le comportement de B_n dans le tore, avec différentes valeurs de p.

   (b) Commenter. Quelle est la différence avec la question précedente ?

XI. Marche simple conditionnée

On considère une marche aléatoire simple S sur $\mathbb{Z}$: $S_0 = 0$, la suite $\{X_k = S_{k+1} - S_k\}_{k=0,1,\dots}$ est IID et $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$.

1. (a) Montrer que S est une chaîne de Markov d'espace d'états $\mathbb{Z}$. Quel est son noyau de transition ? La chaîne est-elle irréductible, apériodique ?
   (b) Exprimer $\mathbb{P}(S_{2k} = 0)$ et $\mathbb{P}(S_{2k+1} = 0)$ en fonction de k (on dénombrera l'ensemble des trajectoires qui prennent la valeur 0 au temps $2k$, resp. $2k+1$).
   (c) En utilisant la formule de Stirling, montrer alors que
   $$\lim_{n\to\infty} \sqrt{\pi n}\,\mathbb{P}(S_{2n} = 0) = 1.$$
   (d) Vérifier numériquement le calcul de la question précédente.
2. Pour $n \in \{1, 2, \dots\}$ on appelle $L_n = \sum_{i=0}^{n} \mathbf{1}_{S_i=0}$ la variable aléatoire qui compte le nombre des visites en zéro jusqu'au temps n.
   (a) Calculer $\mathbb{E}[L_n]$, puis trouver ℓ tel que
   $$\lim_{n\to\infty} \mathbb{E}[L_n]/\sqrt{n} = \ell.$$
   Vérifier numériquement.
   (b) Montrer que
   $$\mathbb{E}[L_n^2] = \sum_{i=0}^{n} \mathbb{P}(S_i = 0) + 2 \sum_{0\le i<j\le n} \mathbb{P}(S_i = S_j = 0).$$
   (c) Montrer que pour $0 \le i < j \le n$,
   $$\mathbb{P}(S_i = S_j = 0) = \mathbb{P}(S_i = 0)\mathbb{P}(S_{j-i} = 0),$$
   et en déduire que
   $$\frac{1}{n} \sum_{0\le i<j\le n} \mathbb{P}(S_i = S_j = 0) \underset{n\to\infty}{\longrightarrow} \frac{1}{\pi} \int_0^1 \sqrt{\frac{1-x}{x}}\,dx,$$
   Déterminer alors la limite de $\mathbb{E}\left[\frac{L_n^2}{n}\right]$, puis vérifier numériquement.
   (d) Etudier *numériquement* le comportement de $L_n/\sqrt{n}$ lorsque $n \to \infty$. Essayer de déterminer la loi de la variable limite Y. Sur la base de l'observation numérique, de quel type de convergence s'agit-il ?
3. Dans cette partie on cherche à étudier le comportement de la marche sous la probabilité $\mathbb{P}(\cdot \mid S_{2n} = 0)$.
   On introduit la variable $T_n := \sum_{k=0}^{n-1} \mathbf{1}_{\{S_k + S_{k+1} > 0\}}$ qui donne le temps passé dans le demi-plan positif par la marche lors de ses n premiers pas.

(a) Tracer un histogramme de la loi de T_{2n} sous $\mathbb{P}(\cdot\,|\, S_{2n}=0)$, en utilisant une méthode du rejet, pour $n = 50, 100, 200, ..., 1000$.

(b) Remarquer que lorsque $S_{2n} = 0$, la marche effectue forcément n pas vers le haut et n pas vers le bas lors de ses $2n$ premiers pas. Il suffit donc de choisir les indices des pas effectués vers le haut pour simuler une trajectoire de la marche sous $\mathbb{P}(\cdot\,|\, S_{2n}=0)$. Tracer alors l'histogramme de T_{2n} sous $\mathbb{P}(\cdot\,|\, S_{2n}=0)$ pour $n = 10^4, 10^5, 10^6$.

(c) En utilisant la méthode de la question précédente, représenter grpahiquement quelques dizaines de trajectoires de $\{\frac{S_k}{\sqrt{n}}, k = 1, ..., 2n\}$ sous $\mathbb{P}(\cdot\,|\, S_{2n}=0)$, pour $n = 10^7$. Tenter alors d'expliquer la forme de l'histogramme de T_{2n} sous $\mathbb{P}(\cdot\,|\, S_{2n}=0)$.

4. On prend maintenant en considération une suite IID $\{\eta_j\}_{j=1,2,...}$ de variables à valeurs dans $\{1, 2, \ldots\}$, avec $\mathbb{P}(\eta_1 = k) = c/k^{3/2}$. On pose $\tau_0 = 0$ et, pour $N = 1, 2, \ldots, \tau_N = \sum_{j=1}^{N} \eta_j$. Finalement on introduit $\mathcal{N}_n := \max\{j : \tau_j \leq n\}$. Etudier $\mathcal{N}_n/\sqrt{n}$ pour n grand (voyez-vous un lien avec la partie 2 ?).

XII. Est, Nord-Est Le but de de ce projet est de simuler les modèle "Est" et "Nord-Est".

1. **[Est]**
   Au temps $t \in \mathbb{N}$, et en chaque site de $\{1, ..., N\}$ se trouve, ou non, une particule. Pour décrire le système, on note $\xi_t(i) = 1$ si il y a une particule au temps t au site i, de sorte que $\xi_t \in \{0,1\}^{\{1,...,N\}}$.
   On suppose qu'initialement, $(\xi_0(i), 1 \le i \le N)$ sont i.i.d. suivant la loi de Bernoulli de paramètre p.

   On va également se servir de variables $(B_t)_{t\ge 1}$ i.i.d suivant la loi de Bernoulli de paramètre $\rho \in [0,1]$.

   Au temps $t + 1/2$, on tire uniformément au hasard un site $i(t) \in \{1, ..., N\}$. On regarde alors à l'est, i.e au site $i(t) + 1$, pour constater si ce site héberge, ou non, une particule (on considérera qu'il n'y a jamais de particule à droite du site N).
   — Si il y a effectivement une particule à l'est de $i(t)$ (i.e. si $i(t) \le N - 1$ et $\xi(i(t) + 1) = 1$), il ne se passe rien
   — Si il n'y a pas de particule à l'est de $i(t)$ (i.e. si $i(t) = N$ ou si $i(t) \le N - 1$ et $\xi_t(i(t) + 1) = 0$) alors on décide que $\xi_{t+1}(i(t)) = B_{t+1}$.

   (a) Montrer que $(\xi_t)_{t\in\mathbb{N}}$ est une chaîne de Markov. Quel est son noyau de transition ? Est-elle irréductible ? apériodique ? Possède-t-elle une/plusieurs mesures invariantes ?
   (b) On suppose dans cette question que $p = \rho$. Calculer alors la loi de $(\xi_1(i), 1 \le i \le n)$. Que peut-on en conclure ?
   (c) Pour différentes valeurs de N, ρ, p simuler $(\xi_t, t \ge 0)$. Commenter.
   (d) Pour différentes valeurs de ρ, p, tracer le graphe de $t \to \mathbb{E}[\xi_t(1)\xi_0(1)] - \mathbb{E}[\xi_t(1)]\mathbb{E}[\xi_0[1]]$, et commenter. Que remarque-t-on en particulier lorsque $\rho \to 1$?

2. **[Nord-Est]** Au temps $t \in \mathbb{N}$, et en chaque site de $\{1, ..., N\}^2$ se trouve, ou non, une particule. Pour décrire le système, on note $\xi_t(i, j) = 1$ si il y a une particule au temps t au site (i, j), de sorte que $\xi_t \in \{0,1\}^{\{1,...,N\}^2}$. Comme précédemment les variables $(B_t)_{t\ge 1}$ sont i.i.d et suivent la loi de Bernoulli de paramètre $\rho \in [0,1]$.

   Au temps $t + 1/2$, on tire uniformément au hasard un site $(i(t), j(t)) \in \{1, ..., N\}^2$. On regarde alors au nord et à l'est, i.e aux sites $(i(t) + 1, j(t))$ et $(i(t), j(t) + 1)$, pour constater si ces sites hébergent, ou non, une particule (on considérera qu'il n'y a jamais de particule à droite d'un site du type (N, j), ou au-dessus d'un site du type (i, N)).
   — Si il y a au moins une particule au nord ou à l'est de $(i(t), j(t))$, il ne se passe rien

— Si il n'y a de particule ni au nord, ni à l'est de $(i(t), j(t))$, alors on décide que $\xi_{t+1}(i(t), j(t)) = B_{t+1}$ (en particulier on considère qu'il n'y a jamais de particule à l'est de $x = N$ et au nord de $y = N$).

Reprendre alors exactement les mêmes questions que dans la partie précédente.

XIII. Turbos !

Dans ce sujet, on cherche à “pousser” une marche aléatoire sur $\{0, \cdots N\}$ en l'aidant à avancer quand elle est sur certains sites, appelés “turbos”.
On suppose que 0 et N sont des états absorbants.
Sur les sites où on a placé un “turbo”, la marche aura une probabilité $p > 1/2$ d'aller vers la droite, et une probabilité $1 - p$ d'aller vers la gauche. Sur le reste des sites, elle se déplace équiprobablement vers la gauche et vers la droite.

On dispose de k “turbos”, que l'on veut répartir de façon à changer le comportement de la marche.

1. On considère la stratégie qui consiste à “pousser” la marche au début de sa trajectoire. On place donc les k “turbos” dont on dispose sur $\{1, \cdots, k\}$.
   (a) Montrer que la position de la marche est une chaîne de Markov. Quelles sont ses transitions ? Est-elle irréductible ? apériodique ? Quels sont les sites transients ?
   (b) On appelle p_i la probabilité qu'une marche partant du point i touche le point N avant de toucher le point 0. Estimer toutes les valeurs des p_i à l'aide de simulations dans le cas où $N = 100$, $k = 19$, $p = 0.501, 0.6, 0.7, 0.8, 0.9, 0.999$. Tracer les graphes des résultats obtenus.
   (c) Que valent p_0, p_N ? Trouver une équation exprimant p_i en fonction de p_{i-1} et p_{i+1} lorsque $0 < i < k + 1$, puis une autre qui traite le cas où $k < i < N$.
   (d) Ecrire une équation de la forme $AP = 0$, où A est une matrice $(N + 1) \times (N + 1)$, vérifiée par le vecteur $P = (p_0, \cdots, p_N)^t$.
   (e) En utilisant la commande `linsolve`, trouver les valeurs des p_i. Représenter les graphes de p_i en fonction de i, pour $N = 100$, $k = 19$, $p = 0.501, 0.6, 0.7, 0.8, 0.9, 0.999$. Comparer aux graphes précédents.
   (f) Montrer que l'équation trouvée plus haut peut être vue comme une équation faisant seulement intervenir les quantités $p_i - p_{i-1}$ et $p_{i+1} - p_i$. Ecrire les deux versions de cette équation (selon la valeur de i). En déduire l'expression explicite de p_i.
   (g) On suppose maintenant que le point 0 est réflexif, c'est à dire que si la marche est en 0, elle va en 1 avec probabilité 1. On note t_i le temps d'atteinte du point N en partant du point i. Comme dans les questions précédentes, estimer les t_i, trouver une relation de récurrence vérifiée par t_i, puis la résoudre numériquement et exactement.
2. On considère la stratégie qui consiste à “pousser” la marche à des points répartis uniformément le long de $\{0, \cdots, N\}$. On suppose que $(k + 1)|N$, et l'on placera les “turbos” sur les sites $\{iN/(k + 1), i = 1, \cdots, k\}$.
   (a) Reprendre les questions de la partie précédente (on ne fera pas les résolutions explicites). Comparer les graphes obtenus.

(b) Laquelle de ces deux stratégies faut-il choisir pour avoir la meilleure probabilité de sortir en N lorsque l'on part de 1 ? (critère A)

(c) Laquelle de ces deux stratégies faut-il choisir pour minimiser l'espérance du temps de sortie si 0 est réflexif et que l'on commence en 0 ? (critère B)

3. On cherche maintenant à évaluer ces deux stratégies.

   (a) Ecrire un programme qui, étant données deux configurations, les compare selon chaque critère.

   (b) Simuler un grand nombre de configurations prises uniformément, puis chercher celle qui est la meilleure selon le critère A. La représenter. Est-elle meilleure que la stratégie de la première question ?

   (c) Simuler un grand nombre de configurations prises uniformément, puis chercher celle qui est la meilleure selon le critère B. La représenter. Est-elle meilleure que la stratégie de la deuxième question ?

   (d) Quelle serait la conclusion à tirer de ces simulations ?

XIV. Réseau de concurrence ($\star$)

On cherche à modéliser le scénario suivant : 3 entreprises se disputent 3 marchés, et chaque entreprise est présente sur seulement deux marchés. Ainsi, l'entreprise A vend des ordinateurs et des téléphones, l'entreprise B vend des téléphones et des tablettes, et l'entreprise C vend des tablettes et des ordinateurs.

On associe à chaque entreprise une *taille*, qui correspond à une taille initiale, plus le nombre de produits vendus (on somme les ventes sur les deux marchés de l'entreprise). On suppose que sur chaque marché, un nouveau client choisit l'entreprise où il va acheter le produit avec une probabilité proportionnelle à la taille de l'entreprise. Ainsi, un client qui veut acheter un téléphone l'achète à l'entreprise A avec probabilité $\frac{T_A}{T_A+T_B}$, et à l'entreprise B sinon.

A chaque instant n, un nouveau client arrive simultanément sur chacun des marchés. A l'instant n, on note les tailles des entreprises T_A^n, T_B^n, T_C^n. On suppose que les trois tailles initiales sont strictement positives.

1. (a) Montrer que le triplet (T_A^n, T_B^n, T_C^n) est une chaîne de Markov. Quelles sont ses transitions ? Est-elle irréductible ? Quels sont les sites récurrents ?

   (b) Que vaut $T_A^n + T_B^n + T_C^n$?

   (c) Réaliser une simulation de cette expérience pour $(T_A^0, T_B^0, T_C^0) = (1, 1, 1)$, et un grand nombre de clients. Que peut-on dire du triplet suivant :
$$\left(\frac{T_A^n}{T_A^n + T_B^n + T_C^n}, \frac{T_B^n}{T_A^n + T_B^n + T_C^n}, \frac{T_C^n}{T_A^n + T_B^n + T_C^n}\right)?$$

   (d) Réaliser des simulations avec d'autres conditions initiales. Qu'observe-t-on ?

   (e) On se place à l'instant n, et on suppose que $T_A^n = \alpha n$, $T_B^n = \beta n$, et $T_C^n = \gamma n$. Calculer l'espérance de T_A^{n+1}, puis de T_B^{n+1} et de T_C^{n+1}.

   (f) Trouver une condition sur (α, β, γ) pour que l'on ait $\mathbb{E}[T_A^{n+1}] = \alpha(n+1)$, $\mathbb{E}[T_B^{n+1}] = \beta(n+1)$, et $\mathbb{E}[T_A^{n+1}] = \gamma(n+1)$.

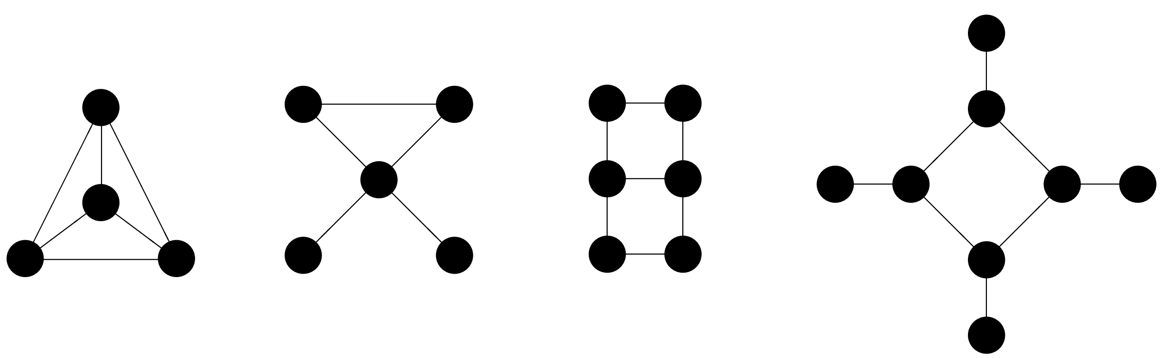
   (g) En raisonnant sur quelques exemples, expliquer pourquoi seule cette valeur de (α, β, γ) semble être une limite possible pour le triplet $\left(\frac{T_A^n}{n}, \frac{T_B^n}{n}, \frac{T_C^n}{n}\right)$.

2. On va maintenant généraliser le marché à un nombre quelconque de concurrents. On se place dans $\mathbb{Z}/N\mathbb{Z}$, et on considère que l'entreprise numéro i est présente sur exactement deux marchés, l'un où elle est en concurrence avec l'entreprise $i-1$, et l'autre où elle est en concurrence avec l'entreprise $i+1$ (ainsi l'entreprise 1 est en concurrence avec l'entreprise N sur un marché, et avec l'entreprise 2 sur l'autre). On considère que l'état initial est une taille 1 pour toutes les entreprises.

   (a) Lorsque $N = 2$, de quel modèle étudié en cours se rapproche-t-on ? Expliquer les similitudes et les différences. Réaliser les histogrammes des répartitions asymptotiques dans les deux cas (modèle du cours, notre modèle), et commenter.

   (b) Lorsque $N = 4$, simuler l'état du marché après un long temps n. Semble-t-il se stabiliser ? Expliquer pourquoi d'après la simulation, on peut décrire l'état final du marché avec une seule valeur r_∞. Que vaut-elle ici ?
   (c) Toujours avec $N = 4$, réaliser un histogramme de r_∞. Comparer avec le cas $N = 2$.
   (d) Réaliser une simulation pour $N = 5$. Commenter.
   (e) Réaliser des simulations pour différentes valeurs de N. Commenter les différences observées.

3. On veut maintenant généraliser le modèle à une structure de marché quelconque. On considère donc un graphe, dont les sommets sont les entreprises et les arêtes sont les marchés. Une entreprise est présente sur un marché si son sommet est connecté à cette arête, et chaque marché ne fait intervenir que deux entreprises, les deux sommets de l'arête.
   (a) Représenter les graphes correspondant aux situations décrites dans les questions précédentes.
   (b) Pour chacun des graphes suivants, réaliser des simulations. Expliquer le comportement asymptotique observé. Quels points communs y a-t-il avec les situations étudiées dans les questions précédentes ? Quelles différences ?



XV. Mangeur de Cookies ($\star$)

On se place sur $\mathbb{Z}^d$, et on considère un marcheur qui se déplace sur le réseau. Initialement il y a un cookie en chaque site. Lorsque le marcheur arrive sur un site qu'il n'avait pas encore visité, il mange le cookie qui s'y trouve.
Sa règle de déplacement dépend ensuite de s'il a ou non mangé un cookie en arrivant sur le site.

— S'il n'a pas mangé de cookie, il se déplace uniformément au hasard vers l'un des plus proches voisins du site où il est (donc chaque voisin est choisi avec probabilité $\frac{1}{2d}$).

— S'il a mangé un cookie, il se déplace vers la droite (dans la direction $+e_1$) avec probabilité $\frac{1}{2d} + p$, et dans chacune des autres directions avec probabilité $\frac{1}{2d} - \frac{p}{2d-1}$.

Le marcheur modifie donc son environnement en faisant disparaître des cookies (il les mange) au cours de sa trajectoire. On cherche à comprendre comment les cookies influencent le comportement du marcheur.

On note X_n la position du marcheur au temps n, et on suppose que $X_0 = 0$. De plus, on note $C_n(x)$ le nombre de cookies à l'instant n au point x (on a $C_n(x) \in \{0,1\}$, et on considère le cas échéant qu'à l'instant n, le marcheur a déjà mangé le cookie en X_n.).

1. On s'intéresse au modèle en tant que chaîne de Markov.
   (a) Montrer que (X_n) n'est pas une chaîne de Markov.
   (b) Montrer que $\left(X_n, (C_n(x))_{x\in\mathbb{Z}^d}\right)$ est une chaîne de Markov. Quelles sont ses transitions ?
2. On se place dans le cas où $d = 1$.
   (a) Réaliser une simulation de la trajectoire de la marche pour différentes valeurs de $p \in]0, 1/2[$. Combien de fois le site 0 a-t-il été visité à l'étape n ?
   (b) Représenter le nombre de total de cookies mangés par le marcheur en fonction du temps.
   (c) On suppose que le marcheur est parti de 0 et visite le site $x > 0$ pour la première fois au temps n. Que vaut $C_n(y)$ pour $y \in \{0, \cdots, x\}$?
   (d) On appelle q_x la probabilité pour que le marcheur retourne alors en 0 avant de visiter le site $x + 1$. Estimer par des simulations q_x pour différentes valeurs de x et de p.
   (e) En utilisant un argument de ruine du joueur, calculer explicitement q_x. Comparer aux résultats de la question précédente.
   (f) Quelle est la probabilité que le marcheur visite les points $\{x + 1, \cdots x + m\}$ avant de revenir en 0 ?
   (g) Pour $p = 0.25$, $x = 1$ et $m = 500$, simuler quelques trajectoires conditionnées à visiter les points $\{x + 1, \cdots x + m\}$ avant de revenir en 0. Commenter.

(h) Quelle est la probabilité pour le marcheur de ne jamais revenir en 0 ? Que peut-on dire de X_n ?

3. On se place maintenant dans le cas $d = 2$.

   (a) Simuler la trajectoire du marcheur pour différentes valeurs de p. Qu'observe-t-on ?

   (b) On suppose qu'à partir d'un certain rang N (éventuellement aléatoire), le marcheur ne mange plus jamais de cookies. On note $X_{N+t} = X_N + Z_t$. Quelle est la loi de Z_t ?

   (c) Donner un argument qui prouve que $||Z_t||$ prend des valeurs arbitrairement grandes. En déduire une contradiction.

   (d) Que peut-on en conclure sur le nombre de cookies mangés par le marcheur ?

   (e) On note $X_n = (X_n^1, X_n^2)$. Simuler le comportement asymptotique de X_n^1/n et de X_n^2/n pour différentes valeurs de p. Expliquer pourquoi le terme "vitesse moyenne" est approprié. Semble-t-elle être strictement positive ? Tracer son graphe en fonction de p.

   (f) Pour $n = 1000$, tracer un histogramme de X_n^2. Commenter.

XVI. Percolation d'invasion ($\star$)

On veut modéliser le territoire envahi par une armée fainéante qui, à chaque instant, conquiert le territoire le plus accessible depuis le territoire qu'elle occupe déjà. Le royaume ainsi créé sera connexe, mais sera-t-il fin et sinueux ? épais et convexe ? Aura-t-il des trous ?
On se place sur $\mathbb{Z}^d$, et à chaque arête on associe un poids i.i.d. qui suit une loi uniforme sur $[0, 1]$.
On crée ensuite un “amas d'invasion” A_n récursivement en partant de l'origine et en ajoutant à l'amas, à chaque étape n, l'arête de plus petit poids reliée à l'amas A_{n-1} (parmi toutes les arêtes connectées à l'amas).

1. Expliquer pourquoi ce modèle est adapté pour modéliser la situation décrite dans l'introduction.
2. On se place dans le cas $d = 1$.
   (a) Montrer que l'amas A_n est un intervalle de longueur n qui contient l'origine.
   (b) Simuler une réalisation de l'agrégat A_n en fonction de n (on présentera plusieurs images ou un film ou un gif).
   (c) On suppose que l'amas trouve sur sa gauche une arête de poids p. Quelle est la loi du nombre d'étapes avant que cette arête soit ajoutée à l'agrégat ? Réaliser des histogrammes pour différentes valeurs de p.
   (d) Toujours dans ce cas, quelle est la loi du poids de l'arête découverte à droite par l'amas lorsque l'arête de poids p est finalement ajoutée à l'agrégat ?
   (e) On considère la suite des arêtes “bloquantes” situées alternativement à gauche et à droite de l'agrégat. On note $(a_1, a_2, \cdots)$ leur position, et on note la suite des poids $(p_1, p_2, \cdots)$. Quelle est la loi de a_1 ? de p_1 ? Vérifier à l'aide d'histogrammes. Quelle est la limite de p_n ?
   (f) Montrer que la suite p_n est une chaîne de Markov.
   (g) En utilisant la question (2c), trouver une relation de récurrence entre $\mathbb{E}[p_{n+1}|p_n]$ et p_n. En déduire une relation entre $\mathbb{E}[p_{n+1}]$ et $\mathbb{E}[p_n]$. Calculer $\mathbb{E}[p_n]$ pour tout n. Vérifier à l'aide de simulations.
   (h) Réaliser un histogramme de $p_n - \mathbb{E}[p_n]$.
   (i) Montrer qu'il existe une constante $A > 0$ telle que p.s., à partir d'un certain rang, $p_n \geq 1 - \frac{An^2}{2^n}$.
   (j) Rappeler la loi de $|a_{n+2}| - |a_n|$ pour une certaine valeur de p_{n+1}. En utilisant la question précédente, montrer que l'on peut majorer la probabilité de l'évènement $\{|a_{n+2}| - |a_n| > M\}$ à partir d'un certain rang.
   (k) A-t-on $|a_{n+2}| - |a_n| \to \infty$ p.s. ? Les simulations confirment-elles ce résultat ?
3. On se place maintenant dans le cas $d = 2$.
   (a) Réaliser une simulation de la croissance de l'amas. On présentera plusieurs images ou un film ou un gif.

(b) On choisit $p > 0.5$, par exemple $p = 0.8$. Représenter l'ensemble des arêtes reliées aux bords d'une grande boite $[-n, n] \times [-n, n]$ par des arêtes de poids inférieur à p. Que remarque-ton ? On appelle cet ensemble C_p.

(c) Montrer que pour $q < p$, on a $C_q \subset C_p$.

(d) On choisit n assez grand pour que l'amas A_n touche C_p. Montrer qu'alors A_n touchera le bord de la boite avant de sortir de C_p. Vérifier sur une simulation, avec $p = 0.8$, puis $p = 0.6$ (On représentera la croissance de l'agrégat depuis l'origine).

(e) En représentant de différentes couleurs deux ou trois différents C_p, simuler la croissance de l'agrégat.

(f) L'agrégat finira-t-il par occuper tout $C_{0.8}$ avant de sortir de la boite ?

(g) Pour différentes valeurs de p, estimer à l'aide de simulations la probabilité que l'agrégat ne contienne aucune arête de de poids supérieur à p lorsqu'il recontre le bord de la boite $[-n, n] \times [-n, n]$.

(h) Représenter le graphe de cette probabilité en fonction de p. Commenter.

XVII. Polymères ($\star$)

On veut modéliser la disposition d'un polymère le long d'une interface (par exemple une membrane cellulaire). Le polymère est représenté par une chaîne d'atomes $\{1, \cdots, n\}$, et sa disposition par la suite $\{X_1, \cdots, X_n\} \in \mathbb{Z}^n$, où X_k indique la position de l'atome k par rapport à l'interface.
On suppose que $X_1 = 0$ et que $X_{k+1} - X_k$ ne peut prendre que les valeurs $\{-1, 1\}$. Ainsi, chaque polymère est représenté par une trajectoire d'entiers distants de 1. Pour une disposition X donnée, on note $C_X = \sum_{i=1}^{n} \mathbf{1}_{X_i=0}$ son nombre de points de contacts avec la membrane.

1. On tire une trajectoire Y uniformément parmi toutes les trajectoires possibles.
   (a) Donner la loi de Y_2.
   (b) A-t-on que Y_k est une chaîne de Markov ?
   (c) Réaliser un histogramme de C_Y pour $n = 1000$.
2. On cherche à classifier les marches suivant la valeur de C.
   (a) Pour de petites valeurs de n, écrire un programme qui renvoie dans un tableau, en fonction de i, le nombre total de marches dont le nombre de passages en 0 vaut i.
   (b) Ecrire un programme qui renvoie une trajectoire uniformément parmi toutes les trajectoires dont le nombre de passages en 0 vaut i.
3. On associe à chaque trajectoire X une énergie $H_\beta(X) = -\beta C_X$, où β est une constante positive. On considère alors une variable aléatoire S dont la loi est donnée par
$$\mathbb{P}(S = X) = \frac{1}{Z_\beta} e^{-H_\beta(X)}.$$
   (a) On note Ω l'ensemble des trajectoires possibles. Montrer que
$$Z_\beta = \sum_{X \in \Omega} e^{-H_\beta(X)}.$$
   (b) Quelles sont les configurations dont l'énergie est la plus haute ? Sont-elles plus ou moins probables que les autres ?
   (c) Quelle est la probabilité d'obtenir une configuration dont l'énergie est minimale ? Comment cette probabilité varie-t-elle avec β ? A quelle condition est-elle proche de 1 ? Que peut-t-on alors dire de S ?
   (d) En utilisant les questions précédentes, simuler une, puis plusieurs, réalisations de S pour quelques valeurs de β. Pour quelles valeurs de n le programme termine-t-il en un temps raisonnable ? Expliquer pourquoi.
4. A une trajectoire donnée X, on associe la trajectoire modifiée en $i \geq 1$, nommée X^i, obtenue de la façon suivante. Pour $k < i$, $X^i_k = X_k$. Si $X_i - X_{i-1} = 1$, on aura $\forall k \geq i, X^i_k = X_k - 2$. Dans le cas contraire, (c'est à dire si $X_i - X_{i-1} = -1$), on aura $\forall k \geq i, X^i_k = X_k + 2$.

(a) Vérifier que X^i est encore une trajectoire acceptable pour représenter un polymère (c'est à dire que $X_0^i = 0$ et que $X_{k+1}^i - X_k^i$ ne peut prendre que les valeurs $\{-1, 1\}$).

(b) On considère la chaîne de Markov dont le noyau est le suivant :

$$Q(X, X^i) = \frac{1}{Cn} \begin{cases} e^{-(H_\beta(X^i) - H_\beta(X))} \text{ si } H_\beta(X^i) - H_\beta(X) > 0 \\ 1 \text{ sinon }, \end{cases}$$

pour tout i et toute trajectoire X, avec C une constante appropriée. Cette chaîne est-elle irréductible ? Apériodique ?

(c) Montrer que la loi de S, telle que définie ci-dessus, est une probabilité invariante de la chaîne. Y a-t'il d'autres probabilités invariantes ?

(d) Simuler la chaîne pour différentes valeurs de n et de β.

(e) Peut-on en déduire une méthode pour simuler une réalisation de S ?

(f) Cette méthode fonctionne-t-elle pour de plus grandes valeurs de n que dans la question (3d) ? Expliquer.

(g) Représenter plusieurs trajectoires de S pour différentes valeurs de β. Qu'observe-t-on ? Commenter.