

Mémoire de modélisation probabiliste en épidémiologie

Romain Lacoste, Mathis Fitoussi

8 juillet 2024

Table des matières

1	Introduction & notations	1
1.1	Objectifs	1
1.2	Modèle épidémiologique	1
1.3	Représentation en graphe	2
2	Préliminaires mathématiques	2
2.1	Procédures d'échantillonnage symétriques	2
2.2	Polynômes de Gontcharoff	5
3	Ensembles de susceptibilité	7
4	Taille totale	10
5	Taille totale et sévérité	12

1 Introduction & notations

1.1 Objectifs

Le but de ce rapport est d'expliquer comment on peut faire une représentation en graphe d'un modèle SIR classique. Notamment, on ne cherche pas à savoir comment une épidémie se propage sur un graphe représentant des interactions sociales pré-existantes, mais simplement à donner une nouvelle interprétation du modèle SIR connu pour essayer d'en tirer de nouveaux résultats, notamment le nombre de gens touchés par l'épidémie et le temps cumulé que des individus ont passé malades.

1.2 Modèle épidémiologique

Chaque individu, une fois infecté, est infectieux pendant une durée aléatoire. Ces durées sont i.i.d. selon une loi connue. On les note $\underbrace{I_{-(a-1)}, \dots, I_0}_{\text{Infectieux}}, \underbrace{I_1, \dots, I_n}_{\text{Susceptibles}}$. Durant ce laps de temps, un infectieux rencontre chaque autre individu de la population avec un taux λ . Les mesures dirigeant ces processus de poisson sont mutuellement indépendants.

On considère que si un infectieux rencontre un susceptible durant sa période infectieuse, il l'infecte nécessairement (en particulier, on ne considère pas de période latente). Ainsi, si on note $W_{i,j}$ le temps entre l'instant d'infection de i et sa rencontre avec j , $W_{i,j} \sim Exp(\lambda)$ et j est infecté ssi $W_{i,j} \leq I_i$, c'est à dire si lors de la rencontre, i est encore infectieux.

On part de la situation suivante :

- n individus susceptibles
- a individus infectieux

On note $N = n + a$ le nombre total d'individus.

On pose $\mathcal{S}_0^{inf} := \{-(a-1), -(a-2), \dots, 0\}$ les infectieux à l'instant initial et $\mathcal{S}_0^{sus} := \{1, 2, \dots, n\}$ les susceptibles à l'instant initial.

On note $E_{n,a}(\lambda)$ ce modèle.

1.3 Représentation en graphe

A partir de ces considérations, on peut construire un graphe de l'épidémie. Notons :

- $V = \{-a+1, \dots, n\}$ l'ensemble des sommets.
- $A \subset V^2$ un ensemble d'arêtes tel que $(i \neq j) \in A \iff W_{i,j} \leq I_i$ et $\forall i \in V, (i, i) \in A$

Le graphe $\mathcal{G}_E = (V, A)$ représente notre épidémie. A noter que $(i, j) \in V$ n'implique pas nécessairement que i infecte j , car ce dernier avait potentiellement déjà été infecté.

On notera de plus $i \rightarrow j$ pour un chemin de i à j .

Que peut-on déjà dire sur cette représentation en graphes ?

- j finit par être infecté ssi il existe un chemin menant de $\mathcal{S}_0 = \{-a+1, \dots, 0\}$ à j
- La taille finale de l'épidémie (i.e. le nombre d'individus infectés tout au cours de l'épidémie une fois celle-ci terminée) est $|\{j \in \{1, \dots, n\} : \exists i \in \mathcal{S}_0 : i \rightarrow j\}|$
- Tel quel, ce modèle ne donne pas d'informations sur l'évolution temporelle de l'épidémie. Il faudrait pour cela ajouter les poids $W_{i,j}$ aux arêtes du graphe pour représenter les durées avant infection.

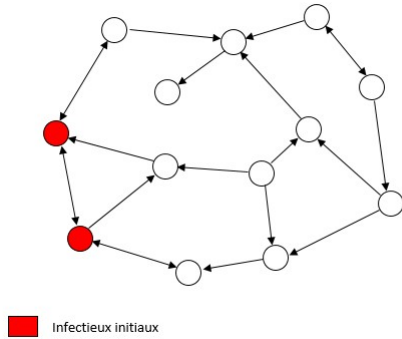


FIGURE 1 – Graphe en début d'épidémie

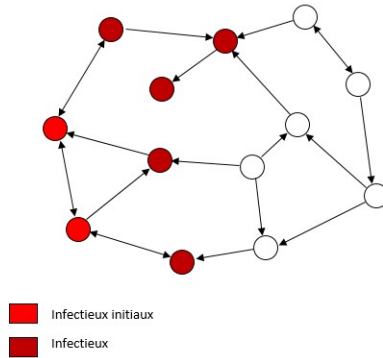


FIGURE 2 – Graphe en fin d'épidémie

2 Préliminaires mathématiques

2.1 Procédures d'échantillonnage symétriques

On considère une population \mathcal{N} finie de taille N .

Par convention, on pose $\mathcal{N} = \{1, \dots, N\}$ et soit $X \subseteq \mathcal{N}$ un sous-ensemble aléatoire.

Pour $A \subseteq \mathcal{N}$, soit $p_A := \mathbb{P}(X = A)$ et $r_A := \mathbb{P}(X \supseteq A)$

Pour chaque individus $i \in \mathcal{N}$, on pose : $\chi_i = \mathbb{1}_{\{i \in X\}}$

On peut remarquer que :

$$\mathbb{E} \left[\prod_{i \in A} \chi_i \right] = \begin{cases} 1 & \text{si } i \in X, \forall i \in A \\ 0 & \text{sinon} \end{cases}$$

Donc puisque $X \supseteq A \Leftrightarrow i \in X, \forall i \in A$, on a alors :

$$r_A = \mathbb{E} \left[\prod_{i \in A} \chi_i \right]$$

Lemme 2.1. Pour tout $n \in \mathbb{N}$, et pour tous réels $(a_i, b_i) \in \mathbb{R}^{2n}$,

$$\prod_{i=1}^n (a_i + b_i) = \sum_{J \subset \{1, \dots, n\}} \prod_{i \in J} a_i \prod_{i \in \{1, \dots, n\} \setminus J} b_i$$

Démonstration. On montre le résultat par récurrence sur n .

$n = 0$ fonctionne car un produit vide vaut 1 et \emptyset est le seul sous-ensemble de \emptyset pour le membre de droite.

$n \geq 0$ et supposons la propriété vraie au rang n .

$$\begin{aligned} \prod_{i=1}^{n+1} (a_i + b_i) &= a_{n+1} \prod_{i=1}^n (a_i + b_i) + b_{n+1} \prod_{i=1}^n (a_i + b_i) \\ &= a_{n+1} \sum_{J \subset \{1, \dots, n\}} \prod_{i \in J} a_i \prod_{i \in \{1, \dots, n\} \setminus J} b_i + b_{n+1} \sum_{J \subset \{1, \dots, n\}} \prod_{i \in J} a_i \prod_{i \in \{1, \dots, n\} \setminus J} b_i \\ &= \sum_{J \subset \{1, \dots, n\}} \prod_{i \in J \cup \{n+1\}} a_i \prod_{i \in \{1, \dots, n\} \setminus J} b_i + \sum_{J \subset \{1, \dots, n\}} \prod_{i \in J} a_i \prod_{i \in \{1, \dots, n\} \setminus J \cup \{n+1\}} b_i \\ &= \sum_{J \subset \{1, \dots, n+1\}} \prod_{i \in J} a_i \prod_{i \in \{1, \dots, n+1\} \setminus J} b_i \end{aligned}$$

□

Proposition 2.1.

$$p_A = \sum_{A \subseteq B \subseteq \mathcal{N}} (-1)^{|B| - |A|} r_B$$

Démonstration. En développant grâce au lemme précédent :

$$\begin{aligned}
p_A = \mathbb{P}(X = A) &= \mathbb{E} \left[\prod_{i \in A} \chi_i \prod_{j \in A^c} (1 - \chi_j) \right] \\
&= \mathbb{E} \left[\prod_{i \in A} \chi_i \sum_{C \in A^c} \prod_{j \in C} (-\chi_j) \right] \\
&= \sum_{C \in A^c} (-1)^{|C|} \mathbb{E} \left[\prod_{i \in A} \chi_i \prod_{j \in C} \chi_j \right] \\
&= \sum_{C \in A^c} (-1)^{|C|} \mathbb{E} \left[\prod_{i \in A} \prod_{j \in C} \chi_i \chi_j \right]
\end{aligned}$$

On remarque ensuite que le terme $\prod_{i \in A} \prod_{j \in C} \chi_i \chi_j$ vaut 1 si et seulement si $A \subset X$ et $C \subset X$, si et seulement si $B = A \cup C \subset X$. Si on change l'indexation de la somme, on a donc, puisque $|C| = |B| - |A|$:

$$p_A = \sum_{A \subset B} (-1)^{|B| - |A|} \mathbb{E} \left(\prod_{i \in B} \chi_i \right) = \sum_{A \subseteq B \subseteq \mathcal{N}} (-1)^{|B| - |A|} r_B$$

□

Dans le cadre d'une procédure d'échantillonnage symétrique, on considère que la dépendance de p_A envers A se limite à la connaissance de sa taille : $|A| = a$.

Ainsi, il suffit de tirer X de cardinal a et il y a $\binom{N}{a}$ choix possible d'un tel ensemble.

Il suit que

$$p_A = \frac{p_a}{\binom{N}{a}}$$

où $p_a = \mathbb{P}(|X| = a)$

Il suit que r_A ne dépend lui aussi que de a et on a :

$$p_a = \binom{N}{a} \sum_{b=a}^N (-1)^{b-a} \binom{N-a}{N-b} r_b$$

Remarque 2.1. *Le choix d'une procédure d'échantillonnage symétrique sous entend des hypothèses de modélisation : chaque individus à la même probabilité d'être choisi. En d'autres termes, cela signifie que les individus sont échangeables. Ainsi, dans des modèles où l'âge ou le sexe des individus a une influence sur le choix cela n'est plus le cas.*

Pour $s, k \in \mathbb{Z}_+$ on pose $s_{[k]} := s(s-1)\dots(s-k+1)$

Lemme 2.2. *Pour une procédure d'échantillonnage symétrique, on a :*

$$\mathbb{E} [|X|_{[k]}] = N_{[k]} r_k, \quad \forall k = 0, 1, \dots$$

Démonstration. Soit k un entier fixé, on a :

$$\begin{aligned}
\mathbb{E} [|X|_{[k]}] &= \sum_{a=k}^N a_{[k]} p_a \\
&= \sum_{a=k}^N \frac{a_{[k]}}{a!} \frac{N!}{(N-a)!} \sum_{b=a}^N (-1)^{b-a} \frac{(N-a)!}{(N-a-N+b)!(N-b)!} \\
&= \sum_{a=k}^N \frac{1}{(a-k)!} \sum_{b=a}^N (-1)^{b-a} \frac{N_{[b]} r_b}{(b-a)!} \\
&= \sum_{b=k}^N N_{[b]} r_b \sum_{a=k}^b (-1)^{b-a} \frac{1}{(a-k)!(b-a)!} \\
&= \sum_{b=k}^N \frac{N_{[b]} r_b}{(b-k)!} \sum_{a=k}^b (-1)^{b-a} \frac{(b-k)!}{(a-k)!(b-k-(a-k))!} \\
&= \sum_{b=k}^N \frac{N_{[b]} r_b}{(b-k)!} \sum_{a=k}^b (-1)^{b-a} \binom{b-k}{a-k} \\
&= \sum_{b=k}^N \frac{N_{[b]} r_b}{(b-k)!} \underbrace{\sum_{i=0}^{b-k} (-1)^{b-k-i} \binom{b-k}{i}}_{C_b}
\end{aligned}$$

D'après la formule du binôme de Newton on a :

$$\begin{aligned}
C_b &= \sum_{i=0}^{b-k} \binom{b-k}{i} 1^i (-1)^{(b-k)-i} \\
&= (1-1)^{b-k} \\
&= \begin{cases} 0 & \text{si } b \neq k \\ 1 & \text{si } b = k \end{cases}
\end{aligned}$$

Finalement, on obtient donc :

$$\begin{aligned}
\mathbb{E} [|X|_{[k]}] &= \sum_{b=k}^N \frac{N_{[b]} r_b}{(b-k)!} \delta_{bk} \\
&= N_{[k]} r_k
\end{aligned}$$

□

On termine cette sous-section en introduisant une dernière notation.

Pour I un infecté, on pose $\phi_I(\theta) := \mathbb{E} [\exp(-\theta I)]$ pour $\theta \geq 0$ sa transformée de Laplace.

Ainsi $\phi_I(-\theta)$ est la fonction génératrice des moments de I .

Pour $k \in \mathbb{N}$, on pose $q_k := \phi_I(k\lambda)$.

On a $q_0 = \phi_I(0) = 1$ et pour $k \geq 1$ q_k est la probabilité que l'infectieux I échoue à contacter les individus d'un ensemble de k susceptibles durant sa période infectieuse.

2.2 Polynômes de Gontcharoff

Définition 2.1. soit $u = (u_n) \in \mathbb{R}^{\mathbb{N}}$ une suite de réels. La suite $(G_i(x|u))_{i \in \mathbb{N}}$ de polynômes de Gontcharoff associée est définie par :

$$\forall k \in \mathbb{N}, \sum_{i=0}^k \frac{u_i^{k-i}}{(k-i)!} G_i(x|u) = \frac{x^k}{k!}$$

On remarque déjà que $G_0(x|u) = 1$ et $\forall k > 0, \deg G_k(x|u) = k$. On en déduit notamment la proposition suivante :

Proposition 2.2. $\forall k \in \mathbb{N}, (G_i)_{0 \leq i \leq k}$ forme une base de $\mathbb{R}_k[X]$

Proposition 2.3. $\forall (i, j) \in \mathbb{N}^2,$

$$G_i^{(j)}(u_j|u) = \delta_{i,j} \quad (1)$$

De plus, cette propriété caractérise les polynômes de Gontcharoff, c'est à dire qu'une famille de polynômes vérifiant $\forall (i, j) \in \mathbb{N}^2, P_i^{(j)}(u_j) = \delta_{i,j}$ est égale à la famille des polynômes de Gontcharoff.

Démonstration. On prouve la première partie par récurrence sur i .

$i = 0$: $G_0(x|u) = 1$ et donc $G_0^{(j)}(u_j|u) = \delta_{0,j}$.

$i \geq 1$ et supposons le résultat vrai jusqu'au rang i .

$$G_i(x|u) = - \sum_{p=0}^{i-1} \frac{u_p^{i-p}}{(i-p)!} G_p(x|u) + \frac{x^i}{i!}$$

Soit $j < i$.

$$G_i^{(j)}(x|u) = - \sum_{p=0}^{i-1} \frac{u_p^{i-p}}{(i-p)!} G_p^{(j)}(x|u) + \frac{x^{i-j}}{(i-j)!}$$

On évalue en u_j et on utilise l'hypothèse de récurrence :

$$G_i^{(j)}(u_j|u) = - \sum_{p=0}^{i-1} \frac{u_p^{i-p}}{(i-p)!} G_p^{(j)}(u_j|u) + \frac{u_j^{i-j}}{(i-j)!} = - \sum_{p=0}^{i-1} \frac{u_p^{i-p}}{(i-p)!} \delta_{p,j} + \frac{u_j^{i-j}}{(i-j)!} = 0$$

On reprend la formule de la dérivée pour constater que pour $j = i, G_i^{(j)}(x, u) = 1$ et on conclut par récurrence.

Pour la seconde partie, soit R un polynôme de degré i . Comme les polynômes de Gontcharoff sont échelonnés en degré, les i premiers forment une base de $\mathbb{R}_i[X]$ et donc $\exists a_0, \dots, a_i \in \mathbb{R}^i$ tels que :

$$\begin{aligned} R(x) &= \sum_{l=0}^i a_l G_l(x|u) \\ \implies R^{(j)}(u_j) &= \sum_{l=0}^i a_l \delta_{l,j} = a_j \end{aligned}$$

Ainsi, si un tel polynôme satisfait (1), $a_j = \delta_{i,j}$, et donc $R(x) = G_i(x|u)$. □

Remarque 2.2. On lit également dans la preuve l'expression de tout polynôme de degré i dans la base des polynômes de Gontcharoff :

$$R(x) = \sum_{l=0}^i R^{(l)}(u_l) G_l(x|u) \quad (2)$$

Proposition 2.4. $\forall 0 \leq j \leq i$,

$$G_i^{(j)}(x|u) = G_{i-j}(x|E^j u)$$

où on a noté $E^j u = (u_{n+j})_{n \in \mathbb{N}}$.

Démonstration. $G_i^{(j)}(x|u)$ est un polynôme de degré $i - j$ donc d'après (2) en utilisant la base $(G_0(x|E^j u), \dots, G_{i-j}(x|E^j u))$,

$$G_i^{(j)}(x|u) = \sum_{l=0}^i G_i^{(j+l)}(u_{j+l}|u) G_l(x|E^j u) = \sum_{l=0}^i \delta_{i,j+l} G_l(x|E^j u) = G_{i-j}(x|E^j u)$$

□

Proposition 2.5. pour $(a, b) \in \mathbb{R}^2$,

$$G_i(ax + b|au + b) = a^i G_i(x|u) \quad (3)$$

Démonstration. $G_i(ax + b|au + b)$ est un polynôme de degré i donc d'après (2) en utilisant la base usuelle,

$$G_i(ax + b|au + b) = \sum_{l=0}^i a^l G_i^{(l)}(au_l + b|au + b) G_l(x|u) = \sum_{l=0}^i a^l \delta_{i,l} G_l(x|u) = a^i G_i(x|u)$$

□

3 Ensembles de susceptibilité

Définition 3.1. Pour $A \subseteq V$, l'ensemble de susceptibilité \mathcal{S}_A de A est défini par :

$$\mathcal{S}_A = \{j \in V \setminus A : j \rightarrow i, \text{ pour un } i \in A\}$$

Avec la convention $\mathcal{S}_\emptyset = \emptyset$

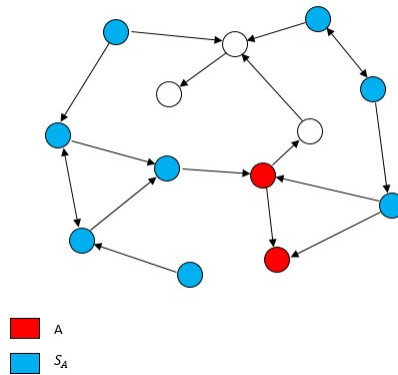


FIGURE 3 – Ensemble de susceptibilité \mathcal{S}_A de A

Autrement dit, c'est l'ensemble des individus tel que s'ils sont infectés, ils vont introduire l'épidémie dans A .

Ainsi on a en particulier $A \subseteq \mathcal{S}_A$.

Remarque 3.1. On rappelle que \mathcal{S}_0^{inf} et \mathcal{S}_0^{sus} désignent respectivement l'ensemble des infectés et susceptibles initiaux. Si $A \subseteq \mathcal{S}_0^{sus}$, c'est-à-dire que A est entièrement composé de susceptibles initiaux, alors A évite l'infection si et seulement si son ensemble de susceptibilité ne contient pas d'infectés initiaux, i.e. si et seulement si $\mathcal{S}_A \cap \mathcal{S}_0^{inf} = \emptyset$.

On pose $S_A := |\mathcal{S}_A|$ la taille de \mathcal{S}_A . Pour $A \subseteq V$, la distribution de S_A ne dépend de A que de par sa taille $|A|$ (cela est justifié par la nature échangeable des individus). Si on suppose que $|V| = N$ et $|A| = j$ pour $0 \leq j \leq N$ on pose $\mathbb{P}_{jN}(S_A = l)$ pour la probabilité que \mathcal{S}_A soit de taille l ($l = j, j+1, \dots, N$). Enfin on rappelle que $q_k = \phi_I(k\lambda)$ ($k = 0, 1, \dots$).

Lemme 3.1. Pour N fixé et $j \in \{0, \dots, N\}$ on a :

$$\mathbb{P}_{jN}(S_A = l) = (N-j)_{[l-j]} G_{l-j}(1|E^j U) q_l^{N-l} \quad (l = j, j+1, \dots, N) \quad (4)$$

où la suite U est donné par $u_k = q_k$ ($k = 0, 1, \dots$).

Démonstration. Soit $j \in \{1, \dots, N\}$ fixé.

On numérote les éléments de V $1, 2, \dots, N$.

On note $\mathbf{0}$ l'ensemble vide, et pour tout $i \in \{1, \dots, N\}$ on note \mathbf{i} l'ensemble $\{1, \dots, i\}$.

Quitte à échantillonner de nouveau, on peut prendre $A = \mathbf{j}$

Puisque l'échantillonnage est symétrique (les individus sont échangeables), il y a $\binom{N-j}{l-j}$ choix possibles d'ensembles de taille $l-j$ pour \mathcal{S}_A .

Ainsi

$$\mathbb{P}_{jN}(\mathcal{S}_A = \mathbf{l}) = \frac{\mathbb{P}_{jN}(S_A = l)}{\binom{N-j}{l-j}} \quad (5)$$

Pour $l \in \{j, j+1, \dots, N\}$, l'ensemble de susceptibilité \mathcal{S}_A peut être construit en deux étapes :

- ① On construit l'ensemble de susceptibilité de A parmi \mathbf{l} , qu'on note $\mathcal{S}_{A,\mathbf{l}}$
 $\mathcal{S}_{A,\mathbf{l}} := \{j \in \mathbf{l} \setminus A : j \rightarrow i, \text{ pour } i \in A\}$
- ② On construit l'ensemble de susceptibilité de $\mathcal{S}_{A,\mathbf{l}}$ parmi V .

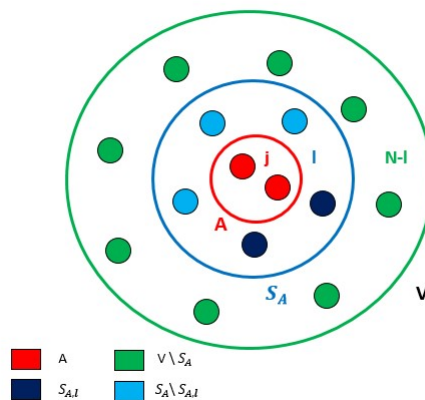


FIGURE 4 – Représentation schématique de la construction

En particulier, de cette construction on tire l'équivalence suivante :

$\mathcal{S}_A = \mathbf{l}$ si et seulement si $\mathcal{S}_{A,\mathbf{l}} = \mathbf{l}$ et aucun des individus de $\mathbf{N} \setminus \mathbf{l}$ ne contactent un individu de \mathbf{l} . On note \mathcal{A} cet évènement. On rappelle que q_k est la probabilité qu'un infectieux échoue à contacter

chaque autre individu dans un ensemble de k susceptibles durant sa période infectieuse. Ainsi puisque les individus de $\mathbf{N} \setminus \mathbf{l}$ ne doivent pas avoir de contact avec un ensemble de l individus et que les contact se font indépendamment, la probabilité de l'évènement \mathcal{A} est q_l^{N-l} . De plus \mathcal{A} et l'évènement $\mathcal{S}_{A,\mathbf{l}} = \mathbf{1}$ sont indépendants puisque les individus entrent en contact de manière indépendante. Ainsi,

$$\begin{aligned} \mathbb{P}_{jN}(\mathcal{S}_A = \mathbf{1}) &= \mathbb{P}_{jN}(\{\mathcal{S}_{A,\mathbf{l}} = \mathbf{1}\} \cap \mathcal{A}) \\ &= \mathbb{P}_{jN}(\{\mathcal{S}_{A,\mathbf{l}} = \mathbf{1}\}) \mathbb{P}_{jN}(\mathcal{A}) \\ &= \mathbb{P}_{jl}(\{\mathcal{S}_A = \mathbf{1}\}) q_l^{N-l} \end{aligned}$$

Or $\mathbb{P}_{jl}(\mathcal{S}_A = \mathbf{1}) = \frac{\mathbb{P}_{jl}(S_A=l)}{\binom{l-j}{l-j}} = \mathbb{P}_{jl}(S_A = l)$

En injectant dans (5) on obtient :

$$\mathbb{P}_{jN}(S_A = l) = \binom{N-j}{l-j} \mathbb{P}_{jl}(S_A = l) q_l^{N-l} \quad (6)$$

Puisque $\sum_{l=j}^N \mathbb{P}_{jl}(S_A = l) = 1$ on a :

$$\begin{aligned} 1 &= \sum_{l=j}^N \binom{N-j}{l-j} \mathbb{P}_{jl}(S_A = l) q_l^{N-l} \\ &= \sum_{i=0}^{N-j} \frac{(N-j)!}{(N-j-i)! i!} \mathbb{P}_{j,j+i}(S_A = j+i) q_{i+j}^{N-j-i} \end{aligned}$$

Soit

$$\sum_{i=0}^{N-j} \frac{q_{i+j}^{N-j-i}}{(N-j-i)!} \frac{\mathbb{P}_{j,j+i}(S_A = j+i)}{i!} = \frac{1}{(N-j)!}$$

En posant $k = N - j$ et $u_i = q_i$ on a :

$$\sum_{i=0}^k \frac{u_{i+j}^{k-i}}{(k-i)!} \frac{\mathbb{P}_{j,j+i}(S_A = j+i)}{i!} = \frac{1}{k!}$$

Et d'après la définition (2.1) on trouve :

$$\frac{\mathbb{P}_{j,j+i}(S_A = j+i)}{i!} = G_i(1|E^j U)$$

Finalement en injectant dans (6) on obtient :

$$\begin{aligned} \mathbb{P}_{jN}(S_A = l) &= \binom{N-j}{l-j} (l-j)! G_{l-j}(1|E^j U) q_l^{N-l} \\ &= (N-j)_{[l-j]} G_{l-j}(1|E^j U) q_l^{N-l} \end{aligned}$$

□

Remarque 3.2. — *Ce lemme donne un lien entre les polynômes de Gontcharoff et les ensembles de susceptibilité. En, effet la distribution de la taille d'un tel ensemble admet une expression sous la forme d'un polynôme de Gontcharoff.*

— Il n'est pas si surprenant de voir apparaître les polynômes de Gontcharoff ici. En effet, on peut reformuler leur définition en :

$$\forall k \in \mathbb{N}, \sum_{i=0}^k \binom{k}{i} (i! u_i^{k-i} G_i(x|u)) = x^k$$

Cette formule fait intervenir des expressions usuelles en dénombrement et donc en probabilités. Ici, on fait notamment intervenir cette expression en $x = 1$ pour pouvoir appliquer sur le membre droit une condition de normalisation. On obtient alors les valeurs des G_i en 1, et, grâce aux propriétés des polynômes de Gontcharoff, cela permet d'en déduire ce que valent les polynômes en un x quelconque. On utilisera cela plus tard pour faire le lien entre les fonctions génératrices de variables d'intérêt de notre problème et polynômes de Gontcharoff.

4 Taille totale

Proposition 4.1. Soit S le nombre de susceptibles restant à la fin de l'épidémie. $\forall j \in \mathbb{N}$,

$$\mathbb{E} [S_{[j]}] = \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i^{(j)}(1|u)$$

On rappelle que n est le nombre initial de susceptibles, a le nombre initial d'infectieux et $\forall k, u_k = q_k$.

Démonstration. Soit $j \in 1, \dots, n$ et soit A un ensemble de susceptibles initiaux de taille j . Remarquons que l'on procède à une procédure d'échantillonnage symétrique sur l'ensemble des susceptibles initiaux, car la variable aléatoire A ne dépend que de $|A|$. D'après 2.2, on a donc (en prenant $S = |X|$) :

$$\mathbb{E} [S_{[j]}] = n_{[j]} r_j = n_{[j]} \mathbb{P}(A \subset S)$$

On doit donc calculer $\mathbb{P}(A \subset S)$, la probabilité que tous les individus d'un ensemble évitent l'infection. Pour cela, on conditionne par rapport au nombre S_A d'individus dans l'ensemble de susceptibilité de A qui étaient initialement infectieux, c'est à dire le nombre d'individus initialement infectieux qui seront amenés à infecter au moins un individu de A . Rappelons également que q_k est la probabilité qu'un infectieux échoue à contacter chaque autre individu dans un ensemble de k susceptibles durant sa période infectieuse.

$$\mathbb{P}(A \subset S | S_A = i) = q_i^a$$

$$\begin{aligned} \Rightarrow \mathbb{E} [S_{[j]}] &= n_{[j]} \sum_{i=j}^n \mathbb{P}_{j,n}(S_A = i) q_i^a && \text{(probas totales)} \\ &= \sum_{i=j}^n \underbrace{n_{[j]}(n-j)_{[i-j]}}_{n_{[i]}} G_{i-j}(1|E^j U) q_i^{n-i+a} && \text{(lemme 3.1)} \\ &= \sum_{i=j}^n n_{[i]} G_i^{(j)}(1|U) q_i^{n-i+a} && \text{(proposition 2.4)} \\ &= \sum_{i=0}^n n_{[i]} G_i^{(j)}(1|U) q_i^{n-i+a} && \text{(premiers termes nuls car } G_i \text{ a degré } i) \end{aligned}$$

□

Remarque 4.1. — On a utilisé que :

$$(n-j)_{[i-j]}n_{[j]} = \prod_{p=0}^{i-j-1} (n-j-p) \times n_{[j]} = \prod_{u=j}^{i-1} (n-u) \prod_{u=0}^{j-1} (n-u) = \prod_{u=0}^{i-1} (n-u) = n_{[i]}$$

— Soit M_k l'évènement : "un infectieux échoue à contacter chaque autre individu dans un ensemble de k susceptibles durant sa période infectieuse". Les auteurs (et donc nous aussi dans la preuve précédente) ont utilisé que $q_k = \mathbb{P}(M_k)$. Cela se justifie presque par le calcul suivant :

$$\begin{aligned} \mathbb{P}(M_k) &= \prod_{j=1}^k \mathbb{P}(W_{i,j} > I) \\ &= \mathbb{P}(W_{i,j} > I)^k \end{aligned} \quad \text{par indépendance des } W$$

Or $\mathbb{P}(W_{i,j} > I) = \mathbb{E}(1_{W_{i,1} > I}) = \mathbb{E}(\mathbb{E}(1_{W_{i,1} > I} | I)) = \mathbb{E}(e^{-\lambda I})$ car les W suivent une loi $\text{Exp}(\lambda)$.

Néanmoins, nous ne comprenons pas le passage de la puissance k à l'intérieur de l'espérance, puisque les variables ne sont pas indépendantes.

Pour nous, le fait que $q_1 = \mathbb{P}(M_1) = \mathbb{E}(e^{-\lambda I})$ est correct, mais, en revanche, on a $q_k = q_1^k \neq \mathbb{E}(e^{-k\lambda I})$.

Néanmoins, si jamais nous avons raison dans cette remarque, il est aisé de corriger le problème en prenant la bonne valeur pour les $q_k = \mathbb{E}(e^{-\lambda I})^k$.

Théorème 4.1. Soit $f_{n,a}(x) = \mathbb{E}(x^S)$ la fonction génératrice de S . On a :

$$f_{n,a}(x) = \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i(x|u)$$

Démonstration. Remarquons d'abord que comme S est à espace d'états fini, $f_{n,a}(x)$ est un polynôme en x , et donc que son développement de Taylor en 1 est exact :

$$f_{n,a}(x) = \sum_{j=0}^{\infty} \frac{(x-1)^j}{j!} f_{n,a}^{(j)}(x)$$

Ensuite, par définition, $f_{n,a}^{(j)}(x) = \mathbb{E}[S_{[j]}]$, et donc :

$$f_{n,a}(x) = \sum_{j=0}^{\infty} \frac{(x-1)^j}{j!} \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i^{(j)}(1|u)$$

On rappelle que la première somme est en réalité finie (car on décrit un polynôme), et donc on peut intervertir :

$$f_{n,a}(x) = \sum_{i=0}^n n_{[i]} q_i^{n+a-i} \sum_{j=0}^{\infty} \frac{(x-1)^j}{j!} G_i^{(j)}(1|u)$$

Enfin, le terme $\sum_{j=0}^{\infty} \frac{(x-1)^j}{j!} G_i^{(j)}(1|u)$ est le développement de Taylor de $G_i(x|u)$ en 1, qui est un polynôme, et vaut donc $G_i(x|u)$:

$$f_{n,a}(x) = \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i(x|u)$$

□

Théorème 4.2. Soit $Z = n - S$ le nombre total d'infectés par les infectés initiaux, et $\mu_{n,a} = \mathbb{E}(Z)$.

$$\mu_{n,a} = n - \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_{i-1}(1|E^1 u)$$

Démonstration.

$$\begin{aligned} \mu_{n,a} &= \mathbb{E}(Z) = n - \mathbb{E}(S) = n - \mathbb{E}(S_{[1]}) = n - f_{n,a}^{(1)}(1) \\ &= n - \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i'(1|u) \\ &= n - \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_{i-1}(1|E^1 u) \end{aligned}$$

□

Objectif accompli, on sait calculer la taille finale de l'épidémie en fonction des données du problème, à savoir le nombre de susceptibles et d'infectieux initiaux et la loi de I !

5 Taille totale et sévérité

La **sévérité** T_A d'une épidémie $E_{n,a}(\lambda, I)$ est la somme de toute les périodes infectieuses de tout ses individus infectieux (infectieux à l'instant initial inclut).

On rappelle que S est le nombre de susceptibles qui restent non-infectieux à la fin de l'épidémie.

Dans cette section on va donner quelques résultats sur la loi jointe (S, T_A)

Tout d'abord on pose : $\phi_{n,a}(x, \theta) = \mathbb{E} [x^S \exp(-\theta T_A)]$ ($x \in \mathbb{R}, \theta \geq 0$) la transformée de Laplace de la loi jointe (S, T_A) . On rappelle que la transformée de Laplace du couple caractérise sa loi.

On rappelle que $q_k = \phi_I(k\lambda)$ est la probabilité qu'un infectieux échoue à contacter chaque autre individu dans un ensemble de k susceptibles. On note I sa période infectieuse et on note A_k l'évènement "il échoue à contacter chaque autre individu dans un ensemble de k susceptibles".

Soit $q_k(\theta) = \mathbb{E} [e^{-\theta I} \mathbb{1}_{A_k}]$ ($k \geq 1, \theta \geq 0$).

Et $q_0(\theta) = \phi_I(\theta)$.

Puisque, conditionnellement à sa période infectieuse I , l'infectieux infecte les susceptibles indépendamment avec probabilité $1 - e^{-\lambda I}$, on a :

$$\begin{aligned} q_k(\theta) &= \mathbb{E} [\mathbb{E} [e^{-\theta I} \mathbb{1}_{A_k} | I]] \\ &= \mathbb{E} [e^{-\theta I} \mathbb{P}(A_k | I)] \\ &= \mathbb{E} [e^{-\theta I} e^{-k\lambda I}] \\ &= \phi_I(\theta + \lambda k) \end{aligned}$$

On note \mathcal{Q} la suite q_0, q_1, \dots et on écrit $f_{n,a}(x)$ comme $f_{n,a}(x; \mathcal{Q})$ pour mettre en évidence explicitement la dépendance en \mathcal{Q} .

Le lemme suivant donne une expression de la transformée de Laplace de la loi jointe (S, T_A) .

Lemme 5.1. *Pour $n, a \in \mathbb{N}$ fixés, on a :*

$$\phi_{n,a}(x, \theta) = (q_0(\theta))^{n+a} f_{n,a} \left(\frac{x}{q_0(\theta)}; \tilde{\mathcal{Q}}(\theta) \right)$$

où $\tilde{\mathcal{Q}}(\theta)$ est donné par $\tilde{q}_i(\theta) = \frac{q_i(\theta)}{q_0(\theta)}$

Démonstration. Si $a = 0$, c'est-à-dire si il n'y a aucun infectieux à l'état initial, on a :

- D'un côté on a : $\phi_{n,0}(x, \theta) = \mathbb{E} [x^S] = x^n$ car $T_A = 0$ (comme il n'y a aucun infectieux la somme des périodes infectieuses est nulle).
- De l'autre :

$$\begin{aligned} (q_0(\theta))^{n+0} f_{n,0} \left(\frac{x}{q_0(\theta)}; \tilde{\mathcal{Q}}(\theta) \right) &= (q_0(\theta))^{n+0} \sum_{i=0}^n n_{[i]} q_i^{n+0-i} G_i \left(\frac{x}{q_0(\theta)} \middle| \tilde{\mathcal{Q}}(\theta) \right) \\ &= (q_0(\theta))^n \sum_{i=0}^n \frac{q_i^{n-i}}{(n-i)!} n! G_i \left(\frac{x}{q_0(\theta)} \middle| \tilde{\mathcal{Q}}(\theta) \right) \\ &= (q_0(\theta))^n \left(\frac{x}{q_0(\theta)} \right)^n && \text{(d'après la définition 2.1)} \\ &= x^n \end{aligned}$$

Donc dans ce cas l'égalité est bien satisfaite.

On suppose à présent $a > 0$ et on considère l'épidémie $E_{n,a}(\lambda, I)$.

On se donne un infectieux à l'instant initial, noté i^* et on note I sa période infectieuse.

La réalisation finale de $E_{n,a}(\lambda, I)$ peut être construit en deux étapes :

- ① En commençant par considérer l'ensemble des susceptibles que i^* a réussi à contacter. On suppose qu'il y a un nombre Z_0 de tels susceptibles.
- ② Puis en considérant l'épidémie parmi les $n - Z_0$ susceptibles restants et avec $a - 1 - Z_0$ infectieux à l'instant initial.

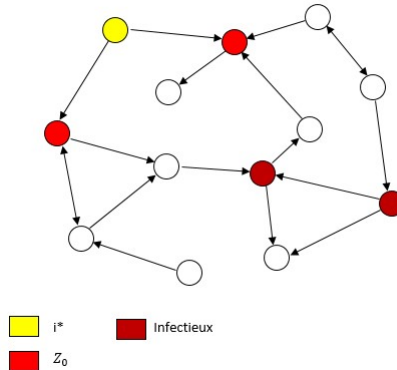


FIGURE 5 – Graphe après l'étape ① de la construction

De cette construction et en sommant sur les valeurs de Z_0 on tire l'égalité suivante :

$$\phi_{n,a}(x, \theta) = \sum_{k=0}^n \mathbb{E} [e^{-\theta I} \mathbb{1}_{\{Z_0=k\}}] \phi_{n-k, a+k-1}(x, \theta) \quad (7)$$

On pose $p_k^{(n)}(\theta) := \frac{\mathbb{E}[e^{-\theta I} \mathbb{1}_{\{Z_0=k\}}]}{q_0(\theta)}$ ($k \in \mathbb{N}$)

On fait remarquer que $p_k^{(n)}(\theta)$ est la fonction de masse du nombre d'éléments qui ne sont pas échantillonnés dans le processus d'échantillonnage symétrique induit par les $\tilde{q}_k(\theta)$ (puisque qu'à l'inverse q_k correspond aux individus qui ont été contactés par l'infectieux).

On pose pour $n, a \in \mathbb{N}$:

$$\psi_{n,a}(x, \theta) := \frac{\phi_{n,a}(x, \theta)}{(q_0(\theta))^{n+a}} \quad (8)$$

Si $n = 0$ alors $S = 0$ (aucun susceptible donc aucun susceptible à la fin de l'épidémie) et T_A est la somme de a copies i.i.d. de I (il n'y a que des infectieux donc l'ensemble des périodes infectieuses correspond à la somme des périodes infectieuses de ces derniers).

De ces deux points on tire :

$$\begin{aligned} \phi_{0,a}(x, \theta) &= \mathbb{E} [x^0 \exp(-\theta(I + \dots + I))] \\ &= \mathbb{E} [\exp(-\theta I)^a] \\ &= \phi_I(\theta)^a \end{aligned} \quad (\text{par indépendance des copies de } I)$$

Ainsi $\psi_{0,a}(x, \theta) = \left(\frac{\phi_I(\theta)}{q_0(\theta)}\right)^a = 1$

Si $a = 0$ alors $S = n$ et $T_A = 0$ et donc (comme on l'a vu plus haut) $\phi_{n,0}(x, \theta) = x^n$.

Ainsi $\psi_{n,0}(x, \theta) = \left(\frac{x}{q_0(\theta)}\right)^n$

De tout cela et en partant de (7), $\psi_{n,a}(x, \theta)$ est déterminé par :

$$\begin{cases} \psi_{n,a}(x, \theta) = \sum_{k=0}^n p_k^{(n)}(\theta) \psi_{n-k, a+k-1}(x, \theta) \\ \psi_{0,a}(x, \theta) = 1 \\ \psi_{n,0}(x, \theta) = \left(\frac{x}{q_0(\theta)}\right)^n \end{cases} \quad (9)$$

En conditionnant sur Z_0 on peut aussi montrer que $f_{n,a}(x) = f_{n,a}(x; \tilde{\mathcal{Q}}(\theta))$ est déterminé par :

$$\begin{cases} f_{n,a}(x) = \sum_{k=0}^n p_k^{(n)}(0) f_{n-k, a+k-1}(x) \\ f_{0,a}(x) = 1 \\ f_{n,0}(x) = x^n \end{cases} \quad (10)$$

Les solutions de (9) et (10) étant unique, on a finalement par unicité de la solution :

$$\psi_{n,a}(x, \theta) = f_{n,a} \left(\frac{x}{q_0(\theta)}; \tilde{\mathcal{Q}}(\theta) \right)$$

et le résultat suit en utilisant (8). \square

Remarque 5.1. Ce lemme donne donc une caractérisation de la loi du couple (S, T_A) et met en évidence un lien avec la fonction génératrice de S . Ce résultat, portant sur la sévérité et la taille totale, permet donc de quantifier la forme finale de l'épidémie. Le théorème suivant donne une expression de la transformée de Laplace du couple faisant intervenir les polynômes de Gontcharoff.

Théorème 5.1. Pour $n, a \in \mathbb{N}$ on a :

$$\phi_{n,a}(x, \theta) = \sum_{i=0}^n n_{[i]} (q_i(\theta))^{n+a-i} G_i(x|U(\theta)) \quad (x \in \mathbb{R}, \theta \in \mathbb{R}_+)$$

où $U(\theta)$ est donné par $u_k(\theta) = q_k(\theta) \quad k \in \mathbb{N}$

Démonstration. D'après le lemme 5.1 on a : $\phi_{n,a}(x, \theta) = (q_0(\theta))^{n+a} f_{n,a}\left(\frac{x}{q_0(\theta)}; \tilde{\mathcal{Q}}(\theta)\right)$

D'après le théorème 4.1 on a : $f_{n,a}(x) = \sum_{i=0}^n n_{[i]} q_i^{n+a-i} G_i(x|u)$

En combinant ces deux résultats on obtient :

$$\begin{aligned} \phi_{n,a}(x, \theta) &= (q_0(\theta))^{n+a} \sum_{i=0}^n n_{[i]} \left(\frac{q_i(\theta)}{q_0(\theta)}\right)^{n+a-i} G_i\left(\frac{x}{q_0(\theta)} \middle| \tilde{\mathcal{Q}}(\theta)\right) \\ &= (q_0(\theta))^{n+a} \sum_{i=0}^n n_{[i]} \left(\frac{q_i(\theta)}{q_0(\theta)}\right)^{n+a-i} G_i\left(\frac{x}{q_0(\theta)} \middle| \frac{1}{q_0(\theta)} U(\theta)\right) \end{aligned}$$

avec $U(\theta)$ est donné par $u_k(\theta) = q_k(\theta)$

Or d'après la proposition 2.5, en prenant $a = \frac{1}{q_0(\theta)}$ on a :

$$G_i\left(\frac{x}{q_0(\theta)} \middle| \frac{1}{q_0(\theta)} U(\theta)\right) = \frac{1}{(q_0(\theta))^i} G_i(x|U(\theta))$$

Finalement on obtient :

$$\begin{aligned} \phi_{n,a}(x, \theta) &= (q_0(\theta))^{n+a} \sum_{i=0}^n n_{[i]} \left(\frac{q_i(\theta)}{q_0(\theta)}\right)^{n+a-i} G_i\left(\frac{x}{q_0(\theta)} \middle| \tilde{\mathcal{Q}}(\theta)\right) \\ &= (q_0(\theta))^{n+a} \sum_{i=0}^n n_{[i]} \left(\frac{q_i(\theta)}{q_0(\theta)}\right)^{n+a-i} \frac{1}{(q_0(\theta))^i} G_i(x|U(\theta)) \\ &= \sum_{i=0}^n n_{[i]} \left(\frac{q_i(\theta)}{q_0(\theta)}\right)^{n+a-i} (q_0(\theta))^{n+a-i} G_i(x|U(\theta)) \\ &= \sum_{i=0}^n n_{[i]} (q_i(\theta))^{n+a-i} G_i(x|U(\theta)) \end{aligned}$$

□