

MODÈLE ÉPIDÉMIE SIR EN ÉPIDÉMIOLOGIE : PRÉSENTATION ET ILLUSTRATION

TURPIN Yohann, GEHU Lauric, FOSSET Jaurel

14 janvier 2024

Table des matières

1	Introduction	3
2	Les modèles déterministes	4
2.1	Le modèle SIR	4
2.1.1	Histoire et amélioration	4
3	Les modèles stochastiques	5
3.1	Reed et Frost	5
3.2	Les améliorations ammenées par le modèle de Bartlett	6
3.2.1	Quelques rappels	6
3.2.2	Le modèle de Bartlett	6
3.3	La construction de Selke	7
3.3.1	Définition du modèle	7
3.3.2	Taille finale de l'épidémie	8
3.4	Que fait-on du début de l'épidémie	10
4	Conclusion	11

Remerciements

Nous tenons à exprimer notre profonde gratitude à l'enseignante Karine Serier pour son soutien, ses conseils et surtout sa patience tout au long de la réalisation de ce rapport. Nous sommes reconnaissants de l'avoir eu comme encadrante et nous espérons que sa passion des mathématiques et sa bonne humeur se ressentiront dans le travail fourni.

Introduction

Les modèles déterministes sont plus faciles à étudier, mais ils ne reflètent pas toujours la réalité. Les modèles probabilistes, quant à eux, sont une façon naturelle de modéliser l'évolution d'une épidémie, car chaque individu a une certaine probabilité d'être infecté par la maladie.

L'étude de ces problèmes stochastiques est importante pour déterminer si, quand la taille de la population augmente, ils convergent vers un problème déterministe. Nous nous intéresserons plus particulièrement aux épidémies qui se transmettent d'individu en individu et confèrent une immunité aux contacts suivants en cas de rétablissement. Les maladies infantiles, sexuellement transmissibles ou même la grippe correspondent à ce genre de pathologie.

On appelle ce type de modèle une épidémie SIR (Susceptible, Infecté, Rétabli). La population est ainsi donc séparée en trois catégories. La population est séparée en trois catégories : les Susceptibles qui sont les individus sains qui sont sensibles à la maladie. Les Infectés qui sont atteints par l'épidémie et peuvent donc la transmettre pendant une certaine période de contagion. A la fin de cette période, ils deviennent Rétablis et sont immunisés contre une nouvelle infection.

Les modèles déterministes

2.1 Le modèle SIR

2.1.1 Histoire et amélioration

Le modèle SIR est un modèle mathématique simple qui permet de modéliser la propagation d'une maladie infectieuse dans une population. Le modèle divise la population en trois groupes : les individus susceptibles d'être infectés (S), les individus infectés (I) et les individus guéris ou rétablies (R). Les individus passent d'un groupe à l'autre en fonction de leur état de santé. Le modèle est décrit par les équations différentielles suivantes :

•

$$\frac{dS}{dt} = -\frac{\beta SI}{N}$$

•

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I$$

•

$$\frac{dR}{dt} = \gamma I$$

où S, I et R représentent respectivement le nombre de personnes dans chaque groupe, N est la taille totale de la population, β est le taux de contact et γ est le taux de guérison. Le modèle de Reed et Frost est un autre modèle mathématique simple qui permet de modéliser la propagation d'une maladie infectieuse dans une population. Dans ce modèle, les individus sont divisés en trois groupes : les individus susceptibles d'être infectés (S), les individus infectés (I) et les individus rétablies (R). Le modèle suppose que chaque individu infecté a une probabilité fixe de transmettre la maladie à un individu susceptible chaque semaine.

Les modèles stochastiques

3.1 Reed et Frost

Ainsi comme annoncée lors de l'introduction, le modèle déterministe traité juste avant possède certaines limites et hypothèses simplificatrices faisant de lui, aux premiers abords, un mauvais outil de comparaison avec la réalité. C'est sur ce même constat qu'en 1928, les chercheurs Reed et Frost posent les bases du premier modèle épidémiologique stochastique de type chaîne binomiale, c'est à dire faisant intervenir la loi binomiale.

Dans ce modèle, on trouve des générations successives d'infectieux d'indice $t = 0, 1, 2, \dots$, qui sont juste capables de contaminer une unique génération de susceptibles. Par la suite, ils ne participent plus au processus de l'épidémie. Supposons que la taille de la population soit égale à une constante n .

On pose $X(t)$ et $Y(t)$, les nombres respectifs de susceptibles et d'infectieux de la génération t . On a donc, comme condition initiale : $X(0) + Y(0) = n$; puis $X(t+1) + Y(t+1) = X(t)$, avec $t = 0, 1, 2, \dots$ puisque les infectieux et les susceptibles de la génération $t+1$ sont vus comme étant issus des susceptibles de la génération t .

Ensuite, on suppose que le nombre d'infectieux de la génération $t+1$ est une variable aléatoire qui suit une loi binomiale de paramètres $X(t)$ et $p(Y(t))$, cette dernière étant la probabilité qu'un susceptible soit infecté, lorsque le nombre d'infectieux est $Y(t)$. Alors, pour décrire tout ce phénomène, on dispose de la formule suivante :

$$P(Y_{t+1} = k | X_t = x, Y_t = y) = \binom{x}{k} p(y)^k (1 - p(y))^{x-k}.$$

Or, le modèle de Reed et Frost suppose que la probabilité de ne pas être infectieux lorsque l'on est susceptible est : $1 - p(y) = (1 - p)^y$

Ainsi, on se retrouve avec la formule suivante :

$$P(Y_{t+1} = k | X_t = x, Y_t = y) = \binom{x}{k} (1 - q^y)^k (q^y)^{x-k}$$

Démonstration. Calculons $P(Y_{t+1} = k | X_0 = x_0, Y_0 = y_0, \dots, X_t = x, Y_t = y)$

Du fait que l'état de la génération $(t+1)$ ne dépends seulement que de la génération t , on a :

$$\bullet P(Y_{t+1} = k | X_0 = x_0, Y_0 = y_0, \dots, X_t = x_t, Y_t = y_t) = P(Y_{t+1} = k | X_t = x, Y_t = y)$$

Puis, en appliquant la formule de la probabilité conditionnelle, on aboutit à :

$$\bullet P(Y_{t+1} = k | X_t = x, Y_t = y) = \binom{x}{k} p(y)^k (1 - p(y))^{x-k}$$

On rappelle enfin que $1 - p(y) = (1 - p)^y$, et on injecte cette égalité dans le résultat précédent pour finalement obtenir :

$$\bullet P(Y_{t+1} = k | X_t = x, Y_t = y) = \binom{x}{k} (1 - q^y)^k (q^y)^{x-k}$$

□

3.2 Les améliorations ammenées par le modèle de Bartlett

3.2.1 Quelques rappels

Afin d'introduire et permettre une meilleure compréhension de ce qui va suivre, permettons-nous quelques rappels.

Processus de Poisson

Un processus de Poisson est un type de processus stochastique qui modélise le nombre d'événements se produisant dans un intervalle de temps donné ou sur un intervalle spatial donné. De façon plus consciencieuse, elle se définit comme suit :

Soit (S_1, \dots, S_n) une suite de variables aléatoires exponentielles indépendantes de paramètre λ . Posons $T_n = \sum_1^n S_i (T_0 = 0)$ et $\forall t \in \mathbf{R}$:

$$N_t = \sum_{n \geq 0} \mathbf{1}_{\{T_n \leq t\}}$$

La famille aléatoire (N_t) s'appelle le **processus de Poisson standard issu de 0 d'intensité λ** . Ce processus compte alors le nombre de variables T_i dans $[0, t]$.

Un processus de Poisson a alors les propriétés suivantes :

- **Incrément constant** : $\forall \omega, t \mapsto N_t$ est croissante, continue à droite, constante par morceaux et ne croît que par sauts de 1, avec N_t finie presque sûrement.
- **Événement unique** : si $N_{t-}(\omega)$ désigne la limite à gauche de $N_t(\omega)$ au point t alors $\forall t, N_t = N_{t-}$ presque sûrement.
- **Loi de Poisson** : $\forall t, N_t$ suit une loi de Poisson de paramètre λt .
- **Indépendance** : $\forall (t_1, t_2, \dots, t_n)$ tels que $0 < t_1 < t_2 < \dots < t_n$ les variables $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ sont indépendantes.
- **Invariance** : si $s < t$, la loi de $N_t - N_s$ est la même que celle de N_{t-s} .

Loi de poisson

On rappelle aussi que si une variable aléatoire X suit une loi de poisson alors on a :

$$P(X = k) = \exp(-\lambda) \frac{\lambda^k}{k!}$$

3.2.2 Le modèle de Bartlett

Les rappels étant maintenant terminés, revenons-en au modèle de Bartlett.

C'est en 1949 que le statisticien le statisticien Maurice S. Bartlett propose ce qui sera le modèle épidémiologique standard. Celui-ci sera défini ainsi :

Au temps initial, la population se compose de m infectés et n susceptibles. Chaque infecté a une période de contagion qui suit une loi I (moyenne i variance σ^2) indépendante de celle des autres.

Et pendant cette période il rencontre un autre individu donné suivant les instants de saut d'un processus de Poisson d'intensité $\frac{\lambda}{n}$. Ainsi, pendant un temps dt , la probabilité qu'un infecté rencontre un individu initialement susceptible vaut λ .

On appellera $E_{n,m}(\lambda, I)$ une telle épidémie.

3.3 La construction de Selke

Une construction équivalente, proposée par Selke en 1983, se base sur l'idée d'affecter à chaque individu une capacité de résistance à l'épidémie et de définir, à partir du nombre d'infectés, une pression de l'épidémie et un critère d'infection sur la résistance de l'individu.

Pour y arriver, les chaînes de Markov sont utilisées pour modéliser l'évolution de l'épidémie au fil du temps, en prenant en compte les transitions entre les différents états de l'épidémie (susceptible, infecté, rétabli). De plus, le processus de Poisson précédemment définie est aussi utilisé afin de modéliser le nombre d'infections qui se produisent dans un intervalle de temps donné, en supposant que les infections se produisent de manière aléatoire et indépendante les unes des autres.

De plus cette construction permet aussi d'identifier un paramètre important quand il s'agit de modéliser une épidémie : sa taille finale. Ici, l'expression de la taille finale de l'épidémie dépend du nombre de reproductions de base R_0 , qui est le nombre moyen de personnes qu'un individu infecté infecte au début de l'épidémie, défini par $R_0 = \lambda i$

Plus le nombre de reproductions de base est élevé, plus la taille finale de l'épidémie sera grande.

L'expression de la taille finale de l'épidémie peut être calculée à partir de l'équation $1 - \tau = e - R_0\tau$, où τ est la proportion de la population qui sera infectée à la fin de l'épidémie.

Chaîne de Markov

Rappelons cette notion mathématique.

Une chaîne de Markov est un processus aléatoire à temps discret dont la principale caractéristique est l'absence de mémoire et l'existence de probabilités de transition entre les états de la chaîne. Cela signifie que seul l'état actuel du processus a une influence sur l'état qui va suivre.

En considérant S un ensemble fini (ou dénombrable) et $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires définie sur le même univers Ω munit de la mesure de probabilité \mathbb{P} et à valeurs dans S . (X_n) est chaîne de Markov homogène si :

— **Propriété de Markov** : $\forall n \in \mathbf{N}$ et $(x_i)^n \in S^n (i \in [[0, n + 1]])$

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_{0:n} = x_{0:n}) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n)$$

— **Homogénéité** : $\mathbb{P}(X_{n+1} = y | X_n = x)$ ne dépend pas de n , $\forall (x, y) \in S^2$. On note alors $p(x, y)$ cette probabilité.

La loi de X_0 est appelée loi initiale de la chaîne de Markov, $P = (p(x, y))_{x, y \in S}$ est appelée matrice de transition.

3.3.1 Définition du modèle

Ainsi, après avoir vu tous les outils qui sont utilisés au sein de ce modèle, revenons-en à sa définition en elle même.

Plus précisément, revenons-en à ses particularités qui sont les notions de pression et résistance à l'épidémie.

On va indexer la population de la façon suivante :

-Les infectés initiaux seront notés $-(m - 1), -(m - 2), \dots, -1, 0$

-Les susceptibles de 1 à n .

On associe à chaque individu un temps de contagion $I - (m - 1), \dots, I_n$ indépendantes et identiquement

distribuées de loi I. A chaque susceptible on associe aussi des seuils Q_1, Q_n indépendants et identiquement distribués suivant une loi exponentielle de paramètre 1.

On peut alors définir à chaque temps t , avec $Y(t)$ le nombre d'infectés, la pression de l'épidémie :

$$A(t) = \frac{\lambda}{n} \int_0^t Y(u) du$$

Un susceptible j est alors infecté quand la pression dépasse le seuil Q_j . Il devient contagieux pendant le temps I_j

3.3.2 Taille finale de l'épidémie

On réordonne les individus susceptibles en fonction de leur seuil croissant que l'on noteras Q_1, \dots, Q_n .

On peut noter que l'épidémie s'arrête lorsque la pression exercée par les individus infectés n'est pas suffisante pour atteindre le prochain seuil. Chaque individu infecté participe à la pression ce qui nous donne la pression exercée après la i ème infection : $\frac{\lambda}{n} \sum_{j=-(m-1)}^i I_j$. Si Q_{i+1} est supérieur à la valeur de la dernière expression, l'épidémie s'arrête.

On a donc que la taille finale de l'épidémie Z est :

$$Z = \min \left\{ \frac{\lambda}{n} \sum_{j=-(m-1)}^i I_j \right\}$$

Theorem 1 (Identité de Wald). *Soit une épidémie $E_{n,m}(\lambda, I)$ de pression finale Z . Alors $\forall \theta \geq 0$, on a :*

$$\mathbb{E} \left[\frac{e^{-\theta P(\infty)}}{\Phi\left(\frac{\lambda\theta}{n}\right)^{Z+m}} \right] = 1$$

où $\Phi(\theta) = \mathbb{E}[e^{-\theta T}]$ est la fonction de transformée de Laplace de T .

Démonstration. Calculons $\Phi\left(\frac{\lambda\theta}{n}\right)^{n+m}$

En utilisant la fonction de transformée de Laplace, on obtient :

- $\Phi\left(\frac{\lambda\theta}{n}\right)^{n+m} = \mathbb{E}[e^{-(\frac{\lambda\theta}{n}) \sum_{j=-(m-1)}^n I_j}]$

En réalisant un changement des bornes à l'insertion de la formule de la taille finale d'une épidémie, on se retrouve avec :

- $\mathbb{E}[e^{-(\frac{\lambda\theta}{n}) \sum_{j=-(m-1)}^i I_j}] = \mathbb{E}[e^{-\theta(Z + \frac{\lambda}{n} \sum_{j=Z+1}^n I_j)}]$

Ce qui nous permet de déduire l'égalité suivant grâce à la fonction de transformée de Laplace :

- $\mathbb{E}[e^{-\theta(Z + \frac{\lambda}{n} \sum_{j=Z+1}^n I_j)}] = \mathbb{E}[e^{-\theta Z} \Phi\left(\frac{\lambda\theta}{n}\right)^{n-Z}]$

On se retrouve ainsi avec l'égalité :

- $\Phi\left(\frac{\lambda\theta}{n}\right)^{n+m} = \mathbb{E}[e^{-\theta Z} \Phi\left(\frac{\lambda\theta}{n}\right)^{n-Z}]$

En divisant les deux côtés de l'égalité par $\Phi\left(\frac{\lambda\theta}{n}\right)^{n+m}$, on se retrouve finalement avec l'identité de Wald :

- $\mathbb{E} \left[\frac{e^{-\theta P(\infty)}}{\Phi\left(\frac{\lambda\theta}{n}\right)^{Z+m}} \right] = 1$

□

De ce précédent théorème, on peut obtenir notre prochain théorème :

Theorem 2. *Dans ce modèle d'épidémie, les probabilités P_k^n que k individus initialement susceptibles soit finalement infectés sont décrites par l'expression :*

$$\sum_{k=0}^l \binom{n-k}{l-k} \frac{P_k^n}{\Phi(\frac{n-l}{n})^{k+m}} = \binom{n}{l}$$

Démonstration.

□

3.4 Que fait-on du début de l'épidémie

Conclusion

Theorem 3 (Loi faible des grands nombres). *Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variable aléatoire indépendante et identiquement distribuée de même loi que X (loi mère). Si $\mathbb{E}[|X|] < +\infty$, alors $\overline{X_n} \rightarrow \mathbb{E}[X]$*

On ne rappelle pas la loi faibles des grands nombres pour rien.

En effet, elle peut être observée dans les modèles SIR déterministes et stochastiques.

Plus précisément, on peut montrer le modèle stochastique converge vers le modèle déterministe correspondant lorsque la taille de la population tend vers l'infini. Ainsi, on peut en conclure que la faible loi des grands nombres est un concept qui fournit une justification mathématique à l'utilisation de modèles déterministes pour approcher le comportement des modèles stochastiques parmi une grande population, tel que pendant une épidémie mondiale.

Table des figures

Liste des tableaux