Grandes Déviations en épidémiologie

Etienne Pardoux
en collaboration avec Peter Kratz

Aix Marseille Université

Ecole CIMPA, ABIDJAN, 22 Mars 2014

Plan

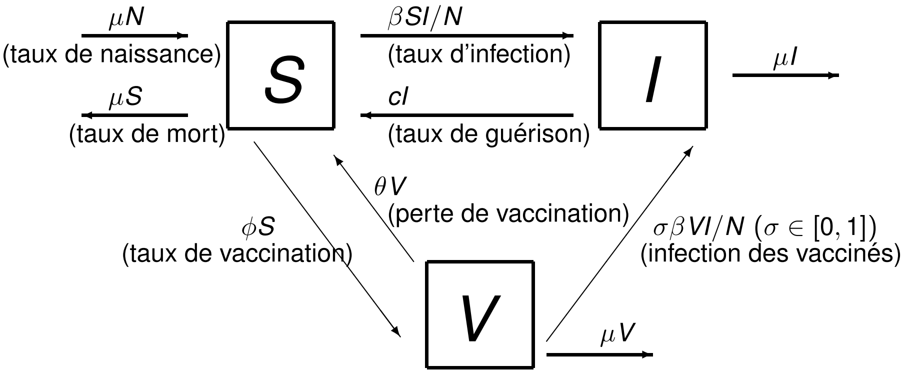
1. Motivation
   - Modèles déterministes compartimentaux
   - Comportement en temps long
   - Modèles stochastiques
2. Un modèle général
   - Modèles poissoniens
   - Loi des Grands Nombres
   - Théorème Central Limite
3. Approximation diffusion
4. Grandes déviations
   - Fonction de taux
   - Principe des Grandes Déviations
   - Sortie d'un domaine

Plan

1. Motivation
   - Modèles déterministes compartimentaux
   - Comportement en temps long
   - Modèles stochastiques
2. Un modèle général
   - Modèles poissoniens
   - Loi des Grands Nombres
   - Théorème Central Limite
3. Approximation diffusion
4. Grandes déviations
   - Fonction de taux
   - Principe des Grandes Déviations
   - Sortie d'un domaine

Un modèle avec vaccination

- Modèle SIV de Kribs–Zaleta et Velasco–Hernandez
- S =# de susceptibles, I = # d'infectieux,
  V = # de vaccinés, $N = S + I + V$ taille totale de la population



Représentation sous forme d'EDO

- $$\begin{aligned} S' &= \mu N - \beta\frac{SI}{N} - (\mu+\phi)S + cI + \theta V \\ I' &= \beta\frac{(S+\sigma V)I}{N} - (\mu+c)I \\ V' &= \phi S - \sigma\beta\frac{VI}{N} - (\mu+\theta)V \end{aligned} \tag{1}$$
- L'équation (1) a une solution unique qui satisfait $0 \leq S, I, V \leq S+I+V = N$

EDO pour les quantité renormalisées

- Posons

$$s(t) = \frac{S(t)}{N},\ i(t) = \frac{I(t)}{N},\ v(t) = \frac{V(t)}{N}.$$

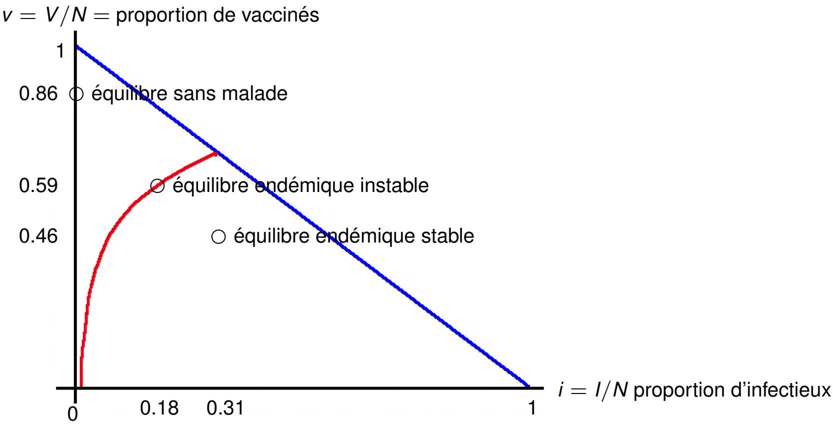
  Les quantités $s(t)$, $i(t)$ et $v(t)$ sont les proportions de susceptibles, infectieux et vaccinés dans la population totale.
- Grâce à l'identité $s = 1 - i - v$, on peut se ramener à une EDO en dimension 2 pour le couple (i, v)
- Cette EDO s'écrit

$$\begin{aligned} i' &= \beta((\sigma - 1)v - i)i + (\beta - \mu - c)i \\ v' &= \phi - \phi i - \sigma\beta v i - (\mu + \theta + \phi)v. \end{aligned}$$

EDO et équilibres

- On veut comprendre le comportement en temps long du modèle
  - Est-ce que la maladie s'éteint ou devient endémique ?
- Trouver les équilibres de l'EDO (1)
  - $R_0 =$ "nombre de reproduction de base"
    = "# moyen d'individus qu'un infectieux infecte, au début de l'épidémie"
    un équilibre sans malade ($I = 0$) de (1) existe
    Si $R_0 < 1 \Rightarrow$ cet équilibre est asymptotiquement stable
  - $\tilde{R}_0 =$ nombre de reproduction de base sans vaccination
    $\tilde{R}_0 > 1 \Rightarrow$ l'équilibre sans malade est instable
  - $R_0 < 1 < \tilde{R}_0$ (vrai avec des choix ad hoc des paramètres)
    $\Rightarrow$ il y a deux équilibres endémiques ($I > 0$)
    L'un est asymptotiquement stable, l'autre est instable

Equilibres de l’EDO



Le modèle SIS

- Dans ce modèle :
  1. Chaque infectieux infecte un susceptible donné au taux β/N.
  2. Chaque infectieux guérit au taux μ, et redevient instantanément susceptible.
- On a donc deux compartiments S et I. On suppose que la taille totale de la population est constante, égale à N. On peut ramener le modèle à une seule EDO, qui s'écrit
- en fonction de I

$$I' = (\beta - \mu)I - \beta\frac{I^2}{N}.$$

- en fonction de $i = I/N$

$$i' = (\beta - \mu)i - \beta i^2.$$

Equilibres du modèle SIS

- $R_0 =$ le nombre moyen qu'un infectieux infecte au début de l'épidémie $= \beta/\mu$
- Si $R_0 \leq 1$, i.e. $\beta \leq \mu$, $i' \leq 0$, 0 est l'unique équilibre, qui est stable. C'est l'équilibre sans malade.
- Si $R_0 > 1$, i.e. $\beta > \mu$, alors 0 est toujours un équilibre, mais qui est instable. Il y a un équilibre stable, l'équilibre endémique, avec comme fraction d'infectés à l'équilibre $i^* = 1 - \mu/\beta$.

Modèles stochastiques

- Comment déduire le modèle stochastique du modèle déterministe ?
- Les taux deviennent ceux de processus de Poisson
  - Un individu de type S devient de type I à l'instant de saur du processus de Poisson correspondant.
  - Les taux sont constants entre deux sauts sucessifs
  - Exemple. Le taux d'infection (à l'instant t) = $\beta \frac{S(t)I(t)}{N}$
- Questions
  - Quelle est la différence entre la solution de l'EDO et celle de l'EDS poissonienne quand N est grand ?
  - La solution de l'EDS peut-elle passer du bassin d'attraction d'un équilibre stable de l'EDF à un autre (avec N grand) ?
  - Combien de temps faut–il attendre pour que cela se produise ?
  - Suivant la taille N de la population est–ce possible/probable ?

Plan

1 Motivation
- Modèles déterministes compartimentaux
- Comportement en temps long
- Modèles stochastiques

2 Un modèle général
- Modèles poissoniens
- Loi des Grands Nombres
- Théorème Central Limite

3 Approximation diffusion

4 Grandes déviations
- Fonction de taux
- Principe des Grandes Déviations
- Sortie d'un domaine

Modèles poissonniens

$$Z^N(t) := x + \frac{1}{N}\sum_{j=1}^{k} h_j P_j\Big(\int_0^t N\beta_j(Z^N(s))ds\Big) \qquad (2)$$

$$= x + \int_0^t b(Z^N(s))ds + \frac{1}{N}\sum_j h_j M_j\Big(\int_0^t N\beta_j(Z^N(s))ds\Big)$$

- d = nombre of compartiments (susceptibles, infectieux, ...)
  N = "taille" de la population. k = nombre de types de sauts.
  $Z_i^N(t)$ = proportion d'individus dans le compartiment i à l'instant t
  A = ensemble (compact) des valeurs prises par le processus
- P_j ($j = 1, \ldots, k$) : processus de Poisson standard mutuellement indépendants
- $M_j(t) = P_j(t) - t$: processus de Poisson compensés
- $h_j \in \mathbb{Z}^d$: valeurs des sauts $\beta_j : A \to \mathbb{R}_+$: intensités des sauts. $b(x) = \sum_j h_j\beta_j(x)$

Le modèle stochastique SIS

- Dans ce cas $d = 1$, $k = 2$, $A = [0,1]$, $h_1 = 1$, $h_2 = -1$, $\beta_1(z) = \beta z(1-z)$, $\beta_2(z) = \mu z$.
- L'EDS poissonnienne s'écrit

$$\begin{aligned} Z^N(t) =& x + \int_0^t [(\beta-\mu)Z_s^N - \beta(Z_s^N)^2] ds \\ &+ \frac{1}{N} M_1 \left(N\beta \int_0^t Z_s^N(1-Z_s^N) ds \right) \\ &- \frac{1}{N} M_2 \left(N\mu \int_0^t Z_s^N ds \right). \end{aligned}$$

Loi des Grands Nombres

- Modèle déterministe

$$\phi(t) := x + \int_0^t b(\phi(s))ds = x + \int_0^t \sum_{j=1}^k h_j \beta_j(\phi(s))ds \qquad (3)$$

Théorème (Kurtz)

Soit $x \in A$, $T > 0$, $\beta_j : \mathbb{R}^d \to \mathbb{R}_+$ bornées et Lipschitz. Il existe des constantes $C_1(\epsilon), C_2 > 0$ ($C_1(\epsilon) = \Theta(1/\epsilon)$ quand $\epsilon \to 0$, C_2 indépendente de ϵ) telles que pour $N \in \mathbb{N}$, $\epsilon > 0$

$$\mathbb{P}\Big[\sup_{t\in[0,T]} |Z^N(t) - \phi(t)| \geq \epsilon\Big] \leq C_1(\epsilon) \exp\Big(-C_2 \frac{N}{\log N}\epsilon^2\Big).$$

En particulier, $Z^N \to \phi$ p.s., uniformément pour $t \in [0, T]$.

Théorème Central Limite

- Posons $U^N(t) := \sqrt{N}[Z^N(t) - \phi(t)]$.
- Alors $U^N(t) \Rightarrow U(t)$, où $U(t)$ est un processus d'Ornstein–Uhlenbeck, solution de l'EDS

$$dU(t) = b'(\phi(t))U(t)dt + \sum_{j=1}^{k} h_j \sqrt{\beta_j(\phi(t))}dW_j(t).$$

Plan

1 Motivation
- Modèles déterministes compartimentaux
- Comportement en temps long
- Modèles stochastiques

2 Un modèle général
- Modèles poissoniens
- Loi des Grands Nombres
- Théorème Central Limite

3 Approximation diffusion

4 Grandes déviations
- Fonction de taux
- Principe des Grandes Déviations
- Sortie d'un domaine

Approximation diffusion

- $$Y^N(t) = x + \int_0^t b(Y^N(s))ds + \frac{1}{N}\sum_j h_j W_j\Big(\int_0^t N\beta_j(Y^N(s))ds\Big),$$
  - W_j ($j = 1, \ldots, k$) : mouvements brownien standard mutuellement indépendants.

Théorème (Kurtz)

Il existe une VA $X = X(N, T)$ dont la loi ne dépend pas de N et qui vérifie $\mathbb{E}[\exp(\lambda X)] < \infty$ pour certains $\lambda > 0$ t.q.

$$\sup_{0\le t\le T} |Z^N(t) - Y^N(t)| \le X \frac{\log N}{N}.$$

Une formulation alternative

- L'EDS de la page précédente se réécrit

$$Y^N(t) = x + \int_0^t b(Y^N(s))ds + \frac{1}{\sqrt{N}} \sum_j h_j \int_0^t \sqrt{\beta_j(Y^N(s))}dW_j(s)$$

  On voit ici une EDS avec "petit bruit", donc on est dans le cadre de la théorie de Wentzell–Freidlin.

- Attention : Le résultat de convergence de Kurtz ne dit rien du comportement en temps long. Il a été montré que les fonctionnelles de Grandes Déviations du systèmes poissonnien et de l'approximation diffusion peuvent être très différentes !

Plan

1. Motivation
   - Modèles déterministes compartimentaux
   - Comportement en temps long
   - Modèles stochastiques
2. Un modèle général
   - Modèles poissoniens
   - Loi des Grands Nombres
   - Théorème Central Limite
3. Approximation diffusion
4. **Grandes déviations**
   - Fonction de taux
   - Principe des Grandes Déviations
   - Sortie d'un domaine

Evénements rares

- D'après la LGN, $Z^N \to \phi$ p.s., uniformément pour $t \in [0, T]$
- Mais : Une (grande) déviation de Z^N par rapport à la solution ϕ de l'EDO est néanmoins possible (même avec N grand !)
- Fixons $T > 0$; $D([0, T]; A) := \{\phi : [0, T] \to A | \phi \text{ càdlàg}\}$;
- On veut quantifier

$$\mathbb{P}[Z^N \in G], \quad \mathbb{P}[Z^N \in F]$$

pour $G \subset D([0, T]; A)$ ouvert, $F \subset D([0, T]; A)$ fermé (N grand), avec typiquement $\phi \notin G$, $\phi \notin F$.

Transformée de Legendre–Fenchel

- Transformée de Legendre–Fenchel
  A la position $x \in A$ et à la direction de mouvement $y \in \mathbb{R}^d$, on associe
  $$L(x,y) := \sup_{\theta \in \mathbb{R}^d} \ell(\theta, x, y)$$
  avec
  $$\ell(\theta, x, y) = \langle \theta, y \rangle - \sum_j \beta_j(x)(e^{\langle \theta, h_j \rangle} - 1)$$
- $L(x,y) \geq L\left(x, \sum_j \beta_j(x) h_j\right) = 0$
- $L(x,y) < \infty$ ssi
  $\exists \mu \in \mathbb{R}_+^k$ t.q. $y = \sum_j \mu_j h_j$ et $\mu_j > 0 \Rightarrow \beta_j(x) > 0$
- $L(x,y)$ peut être considérée comme une “mesure locale” de l’énergie requise pour un mouvement au départ de x dans la direction y.

Fonction de taux

- Fonction de taux. Soit $x \in A$.

$$I_{x,T}(\tilde{\phi}) := \begin{cases} \int_0^T L(\tilde{\phi}(t), \tilde{\phi}'(t))dt & \text{si } \tilde{\phi}(0) = x \text{ and } \tilde{\phi} \text{ est abs. cont.} \\ \infty & \text{sinon} \end{cases}$$

- $I_{x,T}(\phi) = 0$ ssi ϕ est solution de (3) sur $[0, T]$
- $I_{x,T}(\tilde{\phi})$ s'interprète comme l'énergie nécessaire pour que le système suive $\tilde{\phi}$ plutôt que ϕ

Principe des grandes déviations

- Sous des hypothèses à préciser (qui sont vérifiées dans le cas du modèle SIV avec immigration)

Théorème

Si $G \subset D([0,T];A)$ est ouvert et $x \in A$,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}[Z^N \in G] \geq - \inf_{\tilde{\phi}\in G} I_{x,T}(\tilde{\phi}).$$

Si $F \subset D([0,T];A)$ est fermé et $x \in A$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[Z^N \in F] \leq - \inf_{\tilde{\phi}\in F} I_{x,T}(\tilde{\phi}).$$

Remarques bibliographiques

- La théorie des grandes déviations a été bien développée pour les EDS dirigées par un Brownien. Voir en particulier la théorie des “petites perturbations des systèmes dynamiques” de Wentzell et Freidlin 1984. Voir aussi Dembo et Zeitouni 2009.
- La littérature sur les Grandes Déviations pour les processus de Poisson et les EDS dirigées par des processus de Poisson est moins fournie. Signalons cependant Dupuis et Ellis 1997, Feng et Kurtz 2006, Shwartz et Weiss 1995.
- Une difficulté pour nous : certains taux s’annulent quand le processus atteint le bord, donc le logarithme correspondant devient infini ! Motivés par les modèles informatiques, Shwartz et Weiss ont traité en 2005 le cas où certains taux peuvent s’annuler sur certaines frontières. Mais leurs hypothèses ne sont pas satisfaites dans nos modèles, et nous avons dû généraliser encore un peu (Kratz et Pardoux en préparation).

Sortie d'un domaine

- $G =$ domaine d'attraction d'un équilibre x^* ; $x \in G$.
- Quand Z^N sort–il de G (et entre dans le domaine d'attraction d'un autre équilibre) ?
- Par quel point du bord Z^N sort–il de G (et en suivant quelle trajectoire) ?
- L'énergie minimale nécessaire pour aller de y à z dans l'intervalle de temps $[0, T]$ est

$$V(y, z, T) := \inf_{\phi:\, \phi(0)=y, \phi(T)=z} I_{y,T}(\phi)$$

- L'énergie minimale nécessaire pour aller de y à z est

$$V(y, z) := \inf_{T>0} V(y, z, T)$$

- L'énergie minimale pour aller de x^* à la frontière est

$$\bar{V} := \inf_{z \in \partial G} V(x^*, z)$$

Temps de sortie

- Sous des hypothèses adéquates :

Corollaire

$x \in G$, $\delta > 0$.

$$\lim_{N\to\infty} \mathbb{P}[e^{N(\bar{V}+\delta)} > \tau^N] = 1, \quad \lim_{N\to\infty} \mathbb{P}[e^{N(\bar{V}-\delta)} < \tau^N] = 1.$$

- Autrement dit, $\tau^N \simeq e^{N\bar{V}}$.
- C'est une conséquence du principe des Grandes Déviations.

Lieu de sortie

- Sous des hypothèses adéquates :

Théorème (travail en cours avec Brice Samegni)

$x \in G$, $F \subset \partial G$ fermé, $\inf_{z \in F} V(x^*, z) > \bar{V}$.

$$\lim_{N \to \infty} \mathbb{P}[Z^N(\tau^N) \in F] = 0.$$

En particulier, s'il existe $z^* \in \partial G$ tel que pour tout $z \neq z^*$ $V(x^*, z^*) < V(x^*, z)$, alors pour tout $\delta > 0$,

$$\lim_{N \to \infty} \mathbb{P}[|Z^N(\tau^N) - z^*| < \delta] = 1.$$

- Problème : ∂G est la "frontière caractéristique" de G, i.e., pour $x \in G$, $\lim_{t \to \infty} \phi(t) \neq x^*$.

Application au modèle SIS

- Dans ce cas $G = (0, 1]$. Il y a un unique point d'équilibre stable de l'EDO dans $(0, 1]$ si $R_0 > 1$, qui est l'équilibre endémique.
- Nos résultats permettent de répondre à une question ouverte en épidémiologie, à savoir le temps d'extinction d'une situation endémique.
- On a une formule analytique pour la transformée de Legendre–Fenchel $L(x, y)$.
- Supposons par exemple que $\beta = 1.5$ et $\mu = 1$. Alors $R_0 = 1.5$, et par une méthode numérique on peut calculer $\bar{V}$. on trouve $\bar{V} \simeq 0.0702$.
- On en déduit que pour N suffisamment grand, $\tau^N \simeq \exp(0.0702N)$.

Application au modèle SIV

- Dans ce cas, nous n'avons pas encore obtenu les résultats.
- Le problème de contrôle pour le calcul de $\bar{V}$ est plus compliqué, notamment parce que $L(x, y)$ n'est plus connue explicitement. Il faut l'évaluer numériquement de façon approchée.

Une prochaine Ecole CIMPA

- Il y aura du 5 au 16 décembre 2015 une Ecole CIMPA
- à Ziguinchor, Sénégal
- sur le thème

Probabilistic Models in Epidemiology

- Organisateurs : Alassane Diedhiou, Etienne Pardoux
- Principaux orateurs : Frank Ball (Nottingam), Tom Britton (Stockholm), Catherine Larédo (Jouy en Josas), Etienne Pardoux (Marseille), Vier Chi Tran (Lille).
- page web
  http ://www.cmi.univ-mrs.fr/ pardoux/Cimpa2015.html

Une référence importante

- Cette Ecole CIMPA en 2015 traitera de nombreux aspects des modèles probabilistes en épidémiologie, notamment ceux qui étudient la propagation des épidémies sur les graphes de relations sociales, mais aussi les méthodes statistiques adaptées à l'épidémiologie.
- Pour une introduction au monde très riche des modèles mathématiques (déterministes et stochastiques) en épidémiologie, je recommande le livre récent

  Diekmann O., Heesterbeek, H., Britton, T. : *Mathematical tools for understanding infectious disease dynamics*, Princeton University Press, 2013.

MERCI POUR VOTRE ATTENTION !