

PROBABILISTIC DERIVATION OF MULTI-POPULATION EPIDEMIC MODEL

WITH $\beta_{ijk} = -\kappa_j \log(1 - c_{ik})$

GIANG, YANN, MARC ET JD

Definition 1. Durant le petit intervalle de temps δt , chaque individu natif de la ville i visite **une seule** ville j (avec probabilité ρ_{ij}) et rencontrera **en moyenne** κ_j individus. Ces individus proviennent de toutes les villes.

1. NOTATION

Notation : Here, we present list of sets and events describing the state of the system at time t :

- C_i is the set of all individuals born in subpopulation i .
- $V_{i,t}$ is the set of all individuals physically located in subpopulation i from time t to time $t + \delta t$. This includes foreigners traveling in subpopulation i at time t , and all natives from subpopulation i which are not traveling abroad at time t .
- S_t, E_t, I_t, R_t are the sets of all individuals respectively susceptible, exposed, infected and recovered at time t . Note that these set include individuals from all subpopulations.
- $S_{i,t}, E_{i,t}, I_{i,t}, R_{i,t}$ are the same sets, restricted to natives of subpopulation i . So formally, $S_{i,t} = S_t \cap C_i$, $E_{i,t} = E_t \cap C_i$, $I_{i,t} = I_t \cap C_i$, and $R_{i,t} = R_t \cap C_i$.
- $Transmit(y, x)$ is an event indicating that individual x gets infected by individual y which was already infected
- $c_{i,k}$ is the probability that a susceptible individual native from i being in contact with another infected individual native from k gets infected.
- κ_j is the average number of contacts per unit of time a susceptible will have when visiting city j .
- ξ_{jk} refers to the probability that an individual y meeting x in C_j comes from C_k .
- $\rho_{i,j}$, the probability that an individual from subpopulation i visits subpopulation j . Of course, $\sum_{j=1}^M \rho_{ij} = 1$.

Proposition 2. *The coefficient κ should also depend on i , because an individual native from city i meets more people in his own city than abroad ($\kappa_{i,i} > \kappa_{i,j}$).*

2. THE BASICS (NEARLY SAME AS IN EARLIER VERSIONS)

Let us write a probabilistic formulation of $\frac{dE}{dt}$:

One general question is always posed “how does the population of exposed individuals of subpopulation i evolve ?”. For the sake of simplicity, in the process of transmission of the SEIR model, we focus on the incidence and we assume for now that the latent period and the recovery rate, respectively $\mu = \sigma = 0$. Thus, we

write a probabilistic formulation of $\frac{dE_i}{dt}$. Assuming the time is discrete, we have $\frac{dE_i}{dt} \approx \mathbb{E}[E_{i,t+1} \setminus E_{i,t}]$. Then,

$$\begin{aligned}
\mathbb{E}[E_{i,t+1} \setminus E_{i,t}] &= \mathbb{E}[E_{i,t+1} \cap S_{i,t}] \\
&= \sum_{x \in C_i} \Pr[x \in E_{t+1} \wedge x \in S_t] \\
&= \sum_{x \in C_i} \Pr[x \in S_t] * \Pr[x \in E_{t+1} \mid x \in S_t] \\
&= \Pr_{x \sim \mathcal{X}_i}[x \in E_{t+1} \mid x \in S_t] * \sum_{x \in C_i} \Pr[x \in S_t] \\
&= |S_{i,t}| \times \Pr_{x \sim \mathcal{X}_i}[x \in E_{t+1} \mid x \in S_t]
\end{aligned}$$

Assume there are M cities. An individual x of the subpopulation i may be visiting another subpopulation, or staying in its own subpopulation. Applying the law of total probabilities, we get:

$$\begin{aligned}
\Pr_{x \sim \mathcal{X}_i}[x \in E_{t+dt} \mid x \in S_t] &= \sum_{j=1}^M \Pr_{x \sim \mathcal{X}_i}[x \in E_{t+dt} \wedge x \in V_{j,t} \mid x \in S_t] \\
&= \sum_{j=1}^M \Pr_{x \sim \mathcal{X}_i}[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}] \cdot \Pr_{x \sim \mathcal{X}_i}[x \in V_{j,t}] \\
&\quad \sum_{j=1}^M \Pr_{x \sim \mathcal{X}_i}[x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}] \times \rho_{ij}
\end{aligned}$$

Where $\rho_{i,j} = \Pr_{x \sim \mathcal{X}_i}[x \in V_{j,t}]$, the probability that an individual from subpopulation i visits subpopulation j . Of course, $\sum_{j=1}^M \rho_{ij} = 1$.

3. STUDY OF CASE WHERE AGENT x NATIVE FROM CITY i VISITS CITY j

Here, we look at the probability that a susceptible $x \sim \mathcal{X}_i$ visiting j gets infected or not after δt time steps. Let \mathcal{Y} be the uniform distribution over $V_{j,t}$. **(IMPORTANT)** The correct mathematical approach for this would be to assume that for each city k , the number of people native from k that we meet during δt follows a Poisson process. So both the number of people we meet and the number of infected people we meet during δt should be random variables.

In the approach described in [keeling2011], the authors did not do this. They assumed that both the number of people we meet and the number of infected people we meet *are fixed* (otherwise the maths they write would have been different). We will call this the “Keeling & Rohani” interpretation

We introduce an alternative approximation, where we assume that the number κ of people we meet during δt is *fixed*, but each of these people has *some probability* to be infected. This is an *in-between interpretation*, easier than the Poisson process maths, but better than Keeling&Rohani’s one. We will call this the “Yann-Giang” interpretation.

3.1. The “Yann-Giang” interpretation.

Proposition 3. *Agent x meets exactly κ_j other individuals, and each of these individuals has a probability $\frac{|I_{k,t}|}{N_k}$ of being infected, where k is its native city. Let $y_1 \dots y_{\kappa_j}$ be the individuals that x meets. We get:*

$$\begin{aligned} & Pr_{x \sim \mathcal{X}_i} [x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] \\ = & Pr_{x \sim \mathcal{X}_i, y_1, \dots, y_{\kappa_j} \sim \mathcal{Y}} \left[\bigwedge_{p=1}^{\kappa_j} \neg (y_p \in I_t \wedge \text{Transmit}(y_p, x)) \mid x \in S_t \wedge x \in V_{j,t} \right] \end{aligned}$$

So we have:

$$\begin{aligned} & Pr_{x \sim \mathcal{X}_i} [x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] \\ = & Pr_{x \sim \mathcal{X}_i, y \sim \mathcal{Y}} [\neg (y \in I_t \wedge \text{Transmit}(y, x)) \mid x \in S_t \wedge x \in V_{j,t}]^{\kappa_j \delta t} \end{aligned}$$

As in the earlier versions of the document, we have:

- The probability so that a susceptible individual x is infected by an infected individual y :

$$\begin{aligned} & Pr_{x \sim \mathcal{X}_i, y \sim \mathcal{Y}} [y \in I_t \wedge \text{Transmit}(y, x) \mid x \in S_t \wedge x \in V_{j,t}] \\ = & \sum_{k=1}^M Pr_{x \sim \mathcal{X}_i, y \sim \mathcal{Y}} [y \in I_t \wedge \text{Transmit}(y, x) \mid x \in S_t \wedge x \in V_{j,t} \wedge y \in C_k] \cdot Pr_{y \sim \mathcal{Y}} (y \in C_k) \\ = & \sum_{k=1}^M \{ Pr_{x \sim \mathcal{X}_i, y \sim \mathcal{X}_k} [y \in I_t \mid x \in S_t \wedge x \in V_{j,t}] \\ & \times Pr_{x \sim \mathcal{X}_i, y \sim \mathcal{X}_k} [\text{Transmit}(y, x) \mid y \in I_t \wedge x \in S_t \wedge x \in V_{j,t} \wedge y \in C_k] \times Pr_{y \sim \mathcal{Y}} (y \in C_k) \} \\ = & \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \end{aligned}$$

$\xi_{jk} = \frac{N_k \rho_{kj}}{\sum_{v=1}^M N_v \rho_{vj}}$ refers to the probability that an individual y meeting x in C_j comes from C_k .

- Therefore, the probability so that a susceptible individual x is not infected by an infected individual y :

$$1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right)$$

- The probability so that a susceptible individual x is not infected after κ_j contacts per unit time δt .

$$\left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]^{\kappa_j \delta t}$$

- Thus, the probability so that a susceptible individual x becomes infected after κ_j contacts per unit time δt .

$$Pr_{x \sim \mathcal{X}_i} [x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] = \left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]^{\kappa_j \delta t}$$

We now apply the *log* approximation which consists in approximating $1 - (1-u)^v$ by $v \log(1-u)$:

$$Pr_{x \sim \mathcal{X}_i} [x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] = -\kappa_j \delta t \log \left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]$$

So...

$$\frac{dPr_{x \sim \mathcal{X}_i} [x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}]}{dt} \simeq -\kappa_j \log \left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]$$

Overall, we have the formule of the infection force as follows:

$$\lambda_i = \sum_j \rho_{ij} \kappa_j \log \left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]$$

Thus, in the forcing case, the contact rate is in the sinusoidal form as follows :

$$\kappa_i = \kappa_{i0} (1 + \kappa_{i1} \cos(\frac{2\pi t}{T} + \phi_i))$$

3.2. “Keeling & Rohani” Interpretation.

Proposition 4. *Agent x meets exactly $\kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}$ other infected individuals native from city k . (UGLY!!!)*

Let $l_k = \kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}$. Let $y_1^k \dots y_{l_k}^k$ be the infected individuals native from k that our individual x meets between t and $t + \delta t$.

$$\begin{aligned} & Pr_{x \sim \mathcal{X}_i} [x \in S_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] \\ &= Pr_{x \sim \mathcal{X}_i} \left[\bigwedge_{\substack{k=1 \dots M \\ p=1 \dots l_k}} \neg (Transmit(y_p^k, x)) \mid x \in S_t \wedge x \in V_{j,t} \right] \\ &= \prod_{k=1}^M Pr_{x \sim \mathcal{X}_i} \left[\bigwedge_{p=1 \dots l_k} \neg (Transmit(y_p^k, x)) \mid x \in S_t \wedge x \in V_{j,t} \right] \\ &= \prod_{k=1}^M (1 - c_{ik})^{\kappa_j \delta t \xi_{jk} \frac{|I_{k,t}|}{N_k}} \end{aligned}$$

We plug this back into the previous formula, and we get:

$$Pr_{x \sim \mathcal{X}_i} [x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] = 1 - \prod_{k=1}^M (1 - c_{ik})^{\kappa_j \xi_{jk} \frac{|I_{k,t}|}{N_k} \delta t}$$

The first order approximation of $1 - \prod_{k=1}^M (1 - c_{ik})^{v_k}$ is $\sum_{k=1}^M -v_k \log(1 - c_{ik})$. Applying this approximation here, we get:

$$Pr_{x \sim \mathcal{X}_i} [x \in E_{t+\delta t} \mid x \in S_t \wedge x \in V_{j,t}] \simeq \delta t \sum_{k=1}^M \left(-\kappa_j \xi_{jk} \frac{|I_{k,t}|}{N_k} \log(1 - c_{ik}) \right)$$

Define $\beta_{ijk} = -\kappa_j \log(1 - c_{ik})$, let δt converge to zero, and we get:

$$\frac{dPr_{x \sim \mathcal{X}_i} [x \in E_{t+dt} \mid x \in S_t \wedge x \in V_{j,t}]}{dt} \simeq \sum_{k=1}^M \left(\xi_{jk} \frac{|I_{k,t}|}{N_k} \beta_{ijk} \right)$$

If there is only one city i , then we fall back to the formula of [keeling2011]
We have :

$$\beta_i = -\kappa_i \log(1 - c_i)$$

$$\frac{d}{dt} \mathbb{E} [|E_{i,t+dt} - E_{i,t}|] \simeq -|S_{i,t}| \left(\frac{|I_i|}{N_i} \beta_i \right)$$

4. FINAL FORMULA

We simply have to plug in the probability ρ_{ij} that i visits j .
We get, for the “Yann-Giang” interpretation :

$$\frac{d}{dt} \mathbb{E} [|E_{i,t+dt} - E_{i,t}|] \simeq -|S_{i,t}| \sum_j \rho_{ij} \kappa_j \log \left[1 - \sum_{k=1}^M \left(\frac{|I_{k,t}|}{N_k} \times c_{ik} \times \xi_{jk} \right) \right]$$

And for the “Keeling & Rohani” Interpretation :

$$\frac{d}{dt} \mathbb{E} [|E_{i,t+dt} - E_{i,t}|] \simeq -|S_{i,t}| \sum_j \rho_{ij} \sum_{k=1}^M \left(\xi_{jk} \frac{|I_{k,t}|}{N_k} \beta_{ijk} \right)$$

REFERENCES

- [keeling2011] Matt J. Keeling, Pejman Rohani (2011) Modeling Infectious Diseases in Humans and Animals