

新冠疫情传播模型分析及可视化探究

乔姝婧 段成凤 苏思宁 张楠

2020/04/24

一. 主要内容介绍

主要内容

- 一. 社会背景及模型意义
- 二. SIR 模型介绍
- 三. SEIR 模型以及进阶的 C-SEIR 模型
- 四. 模型可视化
 - 常见的动力学模型的可视化
 - 基于 SEIR 模型对武汉市疫情传播过程的模拟

1.1 社会背景

2019 年 12 月底首次发现新冠病毒感染者。由于疫情初期对于病毒未诊断完全，且不了解其传播途径及机制，对此并没有引起民众足够的重视，此时民众防范心理还不强，加之，当时处于春节前，国内有大量人员流动，为病毒的传播提供了契机。随着时间的推移，感染人数呈爆发式增长，疫情给全国人民带来了一定程度的恐慌。此次疫情对我国的经济、政治等各个方面都产生了极大的影响。目前，虽然我国的疫情得到了遏制，但是，世界各个国家正处于疫情的高峰时期，境外病例的输入同样应引起重视，在严加防护的同时，防止病毒再次传播。

1.2 模型的意义

- 在疫情中，我们想了解会有多少人感染，在何时会达到高峰，何时会出现拐点，何时会平息，它的传播与什么因素有关，相关程度大小如何，我们该采取什么措施抑制它的传播等等相关问题。
- 利用数学模型来分析受感染人数的变化规律，预报传染病的高峰时间，分析这些疾病的传播过程，得出控制传染病蔓延的方法，进而抑制其传播。
- 因此我们对疫情传播模型进行分析并进行可视化探究。

1.3 常用模型

●SIR、SEIR、C-SEIR 模型

(1) SIR 模型

SIR 模型是一种传播模型，多用于流行病研究。该模型将人群分为三类，在一定的假设条件下，得到其传染机制，通过分析提出一些可行性建议。传播过程大致如下：最初，所有的节点都处于易感染状态。然后，部分节点接触到信息后，变成感染状态，这些感染状态的节点试着去感染其他易感染状态的节点，或者进入恢复状态。感染一个节点即传递信息或者对某事的态度。恢复状态，即免疫，处于恢复状态的节点不再参与信息的传播。

延伸应用：在社交网络中，由于信息的传播方式与传染病在人群中的传播方式比较相像，因此 SIR 流行病模型也可用于描述信息在社交网络中的传播。当信息在社交网络中的某一用户产生时，其余不知情的用户都处于易感期。

1.3 常用模型

接下来，信息会依赖网络中的好友关系在网络中传播，在这个时期，已经接收到并且传播过信息的用户将处于移出期，而刚刚接收到但还未传播信息的用户将处于传染期，而还未收到信息的用户依然处于易感期。最后网络中的所有用户都收到了这个消息，已经没有传播的必要，信息的传播过程到此结束。（在这我只介绍该模型的大致机制，具体理论由我的小伙伴为大家介绍）

（2）SEIR 模型

该模型是一个带潜伏期的传染病模型，而 SIR 模型是没有考虑潜伏期的，是 SIR 模型的一种优化。

（3）C-SEIR 模型

C-SEIR 模型是 SEIR 模型的进阶版，该模型相比于 SEIR 模型又有了一定的改进，将 SEIR 模型中忽略的点纳入模型。

二. SIR 模型介绍

2.1 SIR 模型

SIR 模型是一种传播模型，是信息传播过程的抽象描述，是传染病模型中最经典的模型，类似的还有 SI 和 SIS 两种，模型中把传染病流行范围内的人群分成三类：S 类，易感者 (Susceptible)，指未得病者，但缺乏免疫能力，与感染者接触后容易受到感染；I 类，感染者 (Infective)，指染上传染病的人，它可以传播给 S 类成员；R 类，移出者 (Removal)，指被隔离，或因病愈而具有免疫力的人。

三者之间的关系如下图：



2.2 SIR 模型

在病毒最开始的时候 $S=N$ ，其中 N 是总人口，然后 S 以每天 α 的速度变到 I ， I 又以每天 β 的速度变到 R ，不同时刻 t 下这三者的关系为：

$$N(t) = S(t) + I(t) + R(t)$$

$$S(t+1) = S(t) - \alpha S(t)$$

$$I(t+1) = I(t) + \alpha S(t) - \beta I(t)$$

$$R(t+1) = R(t) + \beta I(t)$$

2.3 数学模型

SIR 模型的建立基于以下三个假设：

1. 不考虑人口的出生、死亡、流动等种群动力因素。人口始终保持一个常数，即

$$N(T) \equiv K$$

。

2. 一个病人一旦与易感者接触就必然具有一定的传染力。假设 t 时刻单位时间内，一个病人能传染的易感者数目与此环境内易感者总数 $s(t)$ 成正比，比例系数为 β ，从而在 t 时刻单位时间内被所有病人传染的人数为 $\beta S(t)I(t)$ 。

3. t 时刻，单位时间内从染病者中移出的人数与病人数量成正比，比例系数为 γ ，单位时间内移出者的数量为 $\gamma I(t)$ 。

2.3 数学模型

基于以上三个假设条件，感染机制如下所示：

$$\left\{ \begin{array}{l} S(i) + I(j) \xrightarrow{\beta} I(i) + I(j) \\ I(i) \xrightarrow{\gamma} R(i) \end{array} \right.$$

2.3 数学模型

在以上三个基本假设条件下，可知：当易感个体和感染个体充分混合时，感染个体的增长率为

$$\beta S(t)I(t) - \gamma I(t)$$

，易感个体的下降率为 $\beta S(t)I(t)$ ，恢复个体的增长率为 $\gamma I(t)$ 。易感者从患病到移出的过程可以用微分方程表示如下：

$$\begin{cases} \frac{ds(t)}{dt} = -\beta i(t)s(t) \\ \frac{di(t)}{dt} = \beta i(t)s(t) - \gamma i(t) \\ \frac{dr(t)}{dt} = \gamma i(t) \end{cases}$$

2.3 数学模型

解得微分方程的解为：

$$I = (S_0 + I_0) - S + \frac{1}{\sigma} \ln \frac{S}{S_0}$$

(S_0 和 I_0 表示初始值)，其中 σ 是传染期接触数

$$\sigma = \frac{\beta}{\gamma}$$

为保证传染病不蔓延，需要满足

$$S_0 < \frac{1}{\sigma}$$

2.3 数学模型

为了达到这个目的：

- (1) 可以提高阈值 $\frac{1}{\sigma}$ ，需降低 σ ，即减小日接触率 β ，可通过提高卫生水平的方式；
- (2) 增大日治愈率 γ ，可以通过提高医疗水平的方式。
- (3) 也可以通过群体免疫来提高 R_0 ，从而降低 S_0 ，使病情不蔓延。

2.4 缺点

- SIR 模型对人群的分类不够细致，没有明确考虑隔离的因素。而现实中对疑似病人的隔离是控制疫情传播的有效手段。
- 模型没有引入反馈机制，在预测过程中，单纯依据已有数据预测未来较长一段时间的数据，必然会使准确度降低。
- 微分方程组求解较为困难，且对初值比较敏感，这对模型的稳健性是一个很大的影响。

三. SEIR 模型以及进阶的 C-SEIR 模型

3.1 基础的 SEIR 模型

SEIR 模型，类似于 SIR 模型，但是增加了对潜伏期的定义，将人群分为易感人群 (Susceptible)、已被感染但无症状处于潜伏期的人群 (Exposed)、已表现出症状但未被隔离的患病人群 (Infected)、康复人群 (Recovered) 四类，因此更适用于具有一定潜伏期的传染病。这四类人群加起来的总数为 N ，即 $S(t)+E(t)+I(t)+R(t)=N(t)$ 。

SEIR 模型假设：

- 只存在人传人
- 现阶段无治疗的特效药
- 不考虑外界环境因素对病死率等参数的影响
- 忽略人口出生、死亡与国际间流动

3.1 基础的 SEIR 模型

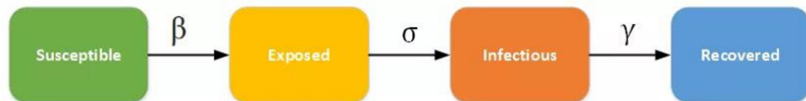
传染病中人群分类：



Figure 1: picture.

3.1 基础的 SEIR 模型

其状态转移图如下所示：



微分方程组：

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

$$\frac{dE}{dt} = \frac{\beta SI}{N} - vE - \sigma E$$

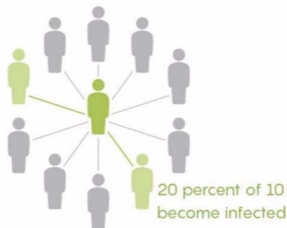
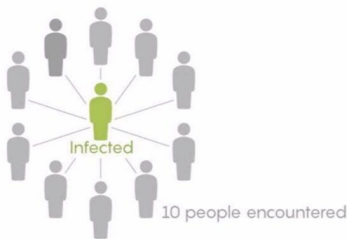
$$\frac{dI}{dt} = \sigma E - \gamma I - vI$$

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

3.1 基础的 SEIR 模型

该模型所涉及的参数主要为：可再生数 R_0 、平均潜伏期时间 DE 和平均收治时间 DI。其中，后两种参数均可直接从官方发布中获得，而 R_0 ，即一名被感染者平均每天传染到的人数，其值较难估计。

基本传染数 (R_0)：一个病人能让几个人得病的能力



3.1 基础的 SEIR 模型

该模型的不足之处：

1. 考虑到政府防控措施的实施与升级， R_0 的取值不应设为定值。
2. 封城作为非常严厉的防控手段执行得非常彻底，必须在参数设置中有效的反映。
3. 社区隔离措施作为后来使用的控疫手段也必须考虑。

而第三点在现有的 SEIR 模型中无法模拟，为此，北京大学前沿计算研究中心陈宝权课题组引入了 C-SEIR 模型。

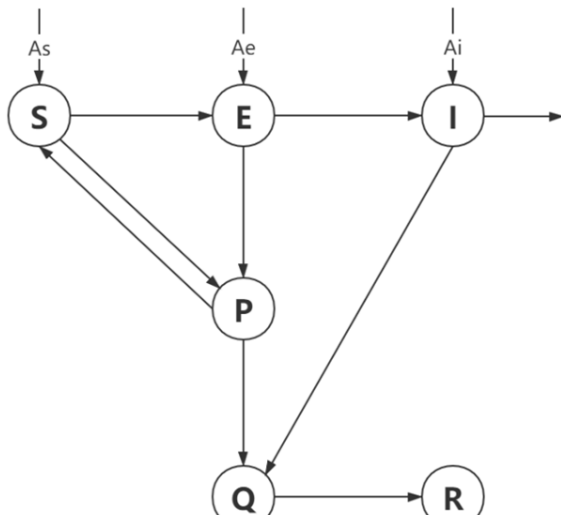
3.2 C-SEIR 模型

C-SEIR 模型相比于 SEIR 模型主要有以下两点改进：

1. 考虑政府的隔离措施，将人群进一步划分出隔离患者和未隔离患者，隔离患者不具备传播能力；
2. 考虑政府措施的加强和群众防护意识上升，病毒的基本可再生系数 (R_0) 应该随时间变化而不是一个固定值，因此通过真实数据拟合出病毒的传染率曲线代替 R_0 。

针对第一点，C-SEIR 在 SEIR 的四类人群基础上增加了两类新的人群：被隔离疑似感染人群 (P)，已确诊并被隔离的患病人群 (Q)。注意在 P 类中的人包括新型冠状病毒的患者，也包括了症状相似但未感染新型肺炎的人群，可以假设这一部分人群不具备向外传染病毒的能力，即病毒的传染能力只与 I 和 E 有关。同时，在考虑湖北省之外的省市时还需要考虑来自武汉的人群输入。

3.2 C-SEIR 模型



3.2 C-SEIR 模型

对于第二点，首先定义病毒的传染率函数 $\beta(t)$ 为每日新增的被传染人数 ΔE 除以未被收治的患病人数 I ，考虑到新冠病毒在潜伏期可能也有一定的传染能力，可以写成公式为：

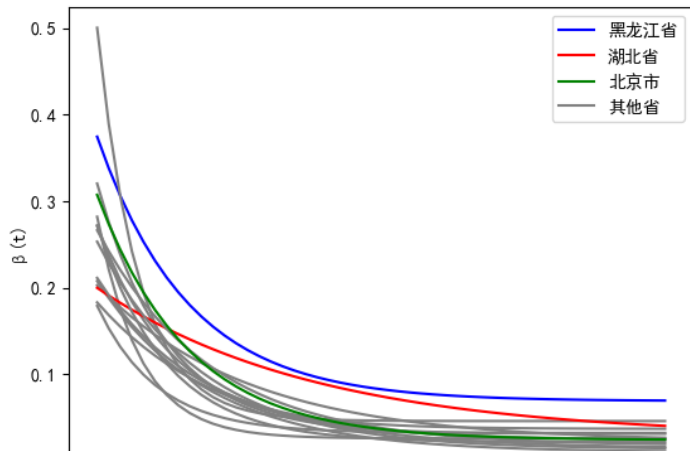
$$\beta(t) = \frac{\Delta E}{I + kE}$$

这里 k 取 0.1，表示潜伏期传染能力是表现症状时的 0.1 倍。为了确定 $\beta(t)$ 的具体形式，首先使用卫健委公布的每日确诊人数以及估算的潜伏期时长反推每日的潜伏期人数和感染人数来计算 $\beta(t)$ 的估计值，再选取函数对数据进行拟合。

3.2 C-SEIR 模型

在大部分传染病的传播过程中，传染率 $\beta(t)$ 会随时间指数衰减。所以使用指数函数，来拟合逐日估计的 $\beta(t)$ 散点值。累计到 2 月 7 日 24 时，累计感染人数超过 200 的各省份的拟合结果如下图所示：

3.2 C-SEIR 模型



四. 疫情可视化

4.1 提要

- 1. 几种常见的动力学模型
- 2. 基本参数
- 3. 武汉市疫情爆发时间点
- 4. 划分阶段适配模型

4.2 传染病动力学模型可视化

1. 动力学模型：

- SI 模型
- SIS 模型
- SIR 模型
- SEIR 模型
- SEITR 模型

2. 基本参数：

- N 为人群总数
- β 为传染率系数, $\beta = ab$, 即每人平均接触的人的数量与单次接触传染概率的乘积
- γ 为恢复率系数, 取决于治疗效率与疾病严重程度, 也包含死亡数量。
- T_e 为疾病潜伏期
- I_0 为感染者的初始人数, E_0 为潜伏者的初始人数, R_0 为治愈者的初始人数或者比率。

4.2 传染病动力学模型可视化

3. 利用 SEITR 模型对武汉市疫情模拟的背景时间节点：

2019.12.1：有记录第一名确诊患者（源于 2019.1.24 柳叶刀文章）

2019.12.31：武汉市通报并开展不明原因病毒性肺炎病例搜索和回顾性调查工作（源于 2020.1.5 武汉卫健委）

2020.1.3：武汉通报称未发现明显人传人证据（源于 1.3 武汉卫健委）

2020.1.5：排除传染性非典型（SARS）可能性（源于 1.5 武汉卫健委）

2020.1.15：武汉方面称病毒尚未发现明确的人传人证据，不能排除有限人传人的可能，但持续人传人的风险较低。（源于 2020.1.15 武汉卫健委）

2020.1.20：钟南山院士表示基本可以确定人传人（源于央视新闻）

2020.1.23：武汉封城（源于肺炎疫情防控指挥部 1 月 23 日发布公告）

截止 2020.1.25 晚：30 个省区市启动重大突发公共卫生事件一级响应。（源于人民日报）

2020.2.3：火神山医院收治病人（源于人民日报）

2020.2.6：雷神山医院收治病人（源于人民日报）

4.2 传染病动力学模型可视化

4. 划分阶段适配模型

第一阶段：2019.12.31 以前。病毒在相对自然条件下进行传播，并未引起大规模的重视与措施，参数相对符合未加限制的自然环境。

第二阶段：2019.12.31——2020.1.23。相关部门已经有了相应的措施，但由于武汉卫健委的“未发现明显人传人证据”使得广大易感人群并未作出相关防护。再加上春节的各项活动使得传染率较高。

第三阶段：2020.1.23——2020.2.2。易感人群的防护意识逐渐加强，武汉封城，也使得传染率相对较低，但医院收治能力有限，大部分患者无法及时就诊，存在患者滞留问题。

第四阶段：2020.2.3 以后。易感人群普遍加强防护，传染率得到有效控制，新建的火神山及之后的雷神山医院与多所方舱医院使得换着收治效率增加，且治愈率有所增加。