

基于 SIR 模型的新冠肺炎疫情的研究报告

摘要

新冠肺炎作为 2020 年最热门的话题，收到了全球无数人的热切关注。本文将主要采用 SIR 模型并结合其他若干模型对新冠肺炎的疫情进行分析和预测，并在此基础上提出可行的防控策略或合适的风险防控建议。

本论文基于 SIR 模型，针对新冠疫情的防护与治理问题，通过 Logistics 函数拟合和以及高级机器学习算法，如指数三光滑平滑 (ETS)，预测某一地区内短时间内的疫情状况。同时对不同地区内的不同时间段，不同政府采取的措施进行比较，得出较为合适的防疫方法。

针对问题一，基于病毒传播基本公式，选取武汉未封城之前，基本没有外力介入的情况下的数据，采用 logistic 函数，求解公式并进行多项式拟合，最后拟合得出系数。

针对问题二，国内，我们选用 SIR 模型，对封城后数据列出微分方程组，采用差分递推法进行计算和预测。但是由于微分方程的解十分复杂和难解，我们最终决定通过迭代形求解。

针对国外地区，由于数据过于复杂以及防控措施的不一致，我们选用更适合的高级机器学习算法，如指数三光滑平滑 (ETS) 进行预测，并得出未来一段时间内的疫情预测数据。

针对问题三，通过 logistic 函数预测国内武汉若不采取任何措施，各时间段内的患病人数。从这里面，找出我国政府采取措施前四天和后四天的数据，接着，利用这两天数据分别开始迭代，与实际情况进行相对比，对是否耽误时间做了评判。并且，结合“群防群控”和“群体免疫”两个不同的政策，给出了目前状况下，一些可行的策略。

针对问题四，基于前三个问题的结论，综合考虑现有实际状况，得出适合的流行疾病防控建议。

关键词：SIR 模型 Logistic 函数 差分递推法 迭代 指数三光滑平滑 (ETS)

一、问题重述

1.1 问题背景

传染病在人类历史的长河中一直威胁着人类的生存。14 世纪，“黑死病”（鼠疫）流行于亚洲、欧洲和非洲，仅在欧洲，黑死病就夺去了 2500 万人的生命；1918-1919 的“西班牙大流感”造成了历史上死亡人数达 4000-5000 多万人的一次瘟疫。乃至本世纪 2003 年“严重急性呼吸道综合征”（非典型性肺炎），2009 年“甲型 H1N1 流感”，2014 年和 2018 年“埃博拉病毒”，2015-2016 年“寨卡病毒”，2012 年和 2015 年“中东呼吸综合征”等令人闻风丧胆传染病。

2019 年底，新型冠状病毒肺炎（Corona Virus Disease 2019, COVID-19）疫情突然爆发，世界各国人民的生命健康和经济发展受到极大的威胁。根据最新报道，全球感染新型冠状病毒肺炎人数已超过 660 多万。虽然各国都出台了一系列病毒防控措施，但由于国情和民风的不同，导致防控效果有较大差异。

“群防群控”是控制病毒传播的强而有力的策略，会在短时间内切断病毒的所有传播途径，但这会以经济停滞和人民失业为代价。而欧美某些国家则遵循“群体免疫”的策略，通过适当措施“缓和”病毒传播的速度，而不采用封城这样极端的措施。采用何种策略需要充分考量病毒本身的传播特征。

在流行病学中，病毒传播系数即基本传染数，是衡量一个病毒传染能力的重要指标。所谓基本传染数，是指在没有外力介入，同时所有人都没有免疫力的情况下，一个感染到某种传染病的人，会把疾病传染给其他多少个人的平均数。

1.2 待解决问题

1. 确定新型冠状病毒肺炎的传播系数。由于大部分疫情实时数据都是通过防控后得到的，但病毒传播系数是在没有外力介入的情况下得出来的，因此要从数据中选取部分数据来计算传播系数。请明确指出所选取的数据，如：某国或地区某月某日至某月某日的数据，并说明选取的理由。

2. 根据现有疫情数据，并选择针对不同类型防控策略与措施的几个国家或地区，以及治愈率，死亡率和疫苗的接种等相关因素，建立数学模型来分析、预测未来的疫情数据。

3. 在问题 2 基础上，建立适当的数学模型说明为什么认为某些国家在新型冠状病毒肺炎暴发前期耽误了时间，并结合“群防群控”和“群体免疫”给出当前状况下这些国家真正可行的防控策略及其有效性。

4. 基于上述问题讨论，并充分考虑不同传染病病毒本身的传播特征，建立数学模型分析，给出健全我国传染病类突发公共卫生事件风险防控建议。

二、问题分析

2.1 问题一的分析

针对问题一：确定新型冠状病毒肺炎的传播系数，在流行病学中，病毒传播系数即基本传染数。一般的认为基本传染数的一个可行计算公式：

$$R_0 = 1 + \beta T_{av}^{[1]}$$

其中 β 为某一新冠肺炎的感染者感染未感染人群的平均概率， T_{av} 为平均感染天数，即某一患者从被感染到死亡或康复或隔离(失去感染能力)的平均天数。由于病毒传播系数(基本传染数)属于病毒的基本性质，故本文中我们选取了武汉市在未封城(采取措施)之前的时间段：12月9日——12月31日。针对参数 β ，我们发现 logistic 模型在预测流行疾病初期的感染情况效果较好，所以我们将选取的时间段的数据进行多项式拟合，发现在传染的初期病毒平均感染率较高但会随着时间的推移逐渐较少并趋于平稳。对于平均传染天数 T_{av} ，根据文献[2]给出的 T_{av} 可能取值为 7.5d。

2.2 问题二的分析

在传染病动力学中，SIR 模型由 Kermack 与 McKendrick 在 1927 年研究流行于伦敦的黑死病时提出的仓室模型，且至今仍应用广泛并不断发展。于是我们借鉴经典的 SIR 模型对封城后武汉市新冠肺炎疫情进行分析与预测。根据 SIR 模型，我们将武汉市全市人口分为未感染人群(S)，感染人群(I)，退出人群(R)，具体关系图 1 所示：



图 1. SIR 传播模型

根据三者关系，列出微分方程组。但是经过计算发现，我们无法求出 I 关于 t 的精确解，故我们将采用差分递推法写出 I(t) 的差分形式在进行整理得到 I(t) 的迭代方程，由此便可得到对武汉市新冠肺炎疫情的近似预测结果。

以上是中国地区最有代表性的预测例子，但针对国外未采取防控措施或采取措施时间太晚地区，无法使用 SIR 模型进行预测分析。其主要原因为：国外地区疫情防控力量太弱以及检测力度不够导致的数据波动性很大，致使预测的结果误差较大。故本文中使用的拟合以及高级机器学习算法，如指数三光滑平滑(ETS)预测国外地区短时间内的疫情状况。

2.3 问题三的分析

在问题三中，我们以武汉 1 月 24 日封城为转折点，分别研究武汉市提前一周和延迟一周封城所导致新冠肺炎感染效果。因为在封城之前，在武汉地区新型冠状病毒是以正常传播系数传播的，因此我们延用第一问的 logistic 模型分别求出在没有外力介入的情况下，武汉市在 1 月 20 日和 1 月 28 日的累计感染人数，并以此人数为初始数据带入至问题二中的 SIR 模型中，求得提前采取措施与延迟采取措施的最终疫情传播情况。

2.4 问题四的分析

不同传染病病毒传播特征呈现多样化，常见的传染病模型按照传染病类型分为 SI、SIR、SIRS、SEIR 模型等。本文中我们将对这四类不同模型进行分析，并结合疫情特征图给出合适的防控建议

三、基本假设

- 1.假设从网络上获取的数据贴近疫情的真实情况；
- 2.假设武汉市在封城之前没有采取任何措施；
- 3.只对确诊人群进行分析和预测；
- 4.武汉市地区采取封城措施之后，不再考虑迁移率等因素；
- 5.感染人群康复之后再次复发不会具有传染性；
- 6.不考虑疫情期间人群的自然死亡；

四、符号说明

图 2.符号系统

符号	说明
t	时间
$S(t)$	第 t 天未受感染人群数量
$I(t)$	第 t 天未感染人群数量
$R(t)$	第 t 天退出人群(无传染能力)数量
β	封城前平均传染率
μ	治愈率
η	死亡率
R_0	病毒传播系数

T_{av}	平均传染天数
R^2	相关系数
N	某地区全体人口数量
α	封城后的感染率

五、模型的建立与求解

5.1 针对问题一模型的建立与求解

5.1.1 问题一模型的建立

新型冠状病毒肺炎的传播系数是预测新冠肺炎疫情的一个重要指标，在流行病学中，病毒传播系数即基本传染数。一般的认为基本传染数的一个可行计算公式：

$$R_0 = 1 + \beta T_{av}^{[1]} \quad (1)$$

针对公式中的 β 值，因为 logistic 函数在模拟传染病在初期的传播情况效果较好，在本文中我们将选取的时间段的数据带入至 logistic 函数中进行拟合：

$$I(t) = \frac{N}{1 + (\frac{N}{I_0} - 1)e^{-\beta t}} \quad (2)$$

其中 N 为武汉市常驻人口。

病毒传播系数是在没有外力介入的情况下得出来的，故在本文中我们选取了武汉市从第一例患者出现：12 月 9 日至 12 月 31 日的数据来计算病毒传播系数。其主要原因有以下两点：

1. 在该时间段内武汉市没有采取任何的防控措施，因而此段时间内新型冠状病毒肺炎是在没有任何外力介入的情况下进行自然传播的；
2. 相比较于国外地区在疫情爆发初期统计患者数据(波动性太大)，武汉市的前期数据更加贴适于 logistic 函数模拟出来的传染病在初期的增长曲线；

针对公式中的 T_{av} 值，《新英格兰医学杂志》最近的研究指出：COVID-19 潜伏期中位数和四分位距为 4.0d (2.0d-7.0d)；从出现发热等症状到发展为确诊病例的时间中位数和 IQR 为 3.0d 和 1.0d-6.0d。由于 COVID-19 的潜伏期同样具有传染性，且从 IQR 来看，此两项数据均存在一定程度的右偏，故将 T_{av} 取值为 7.5d 在当前的研究阶段是有依据的。

5.1.2 问题一模型的求解

Logistics 函数的建立：

设武汉市累计感染人数最大值为 N，当 $I(t)=N$ 时， $\beta=0$ 。

$$\frac{dI}{dt} = \beta I \quad (3)$$

$$I(0) = 0$$

$$\beta(I) = \beta - \alpha I \quad (\alpha > 0, \beta > 0) \quad (4)$$

上述两式求解，可得： $\alpha = \frac{\beta}{N}$ ，于是得到：

$$\beta(I) = \beta(1 - \frac{I}{N}) \quad (5)$$

将(5)代入到(3)中又得到：

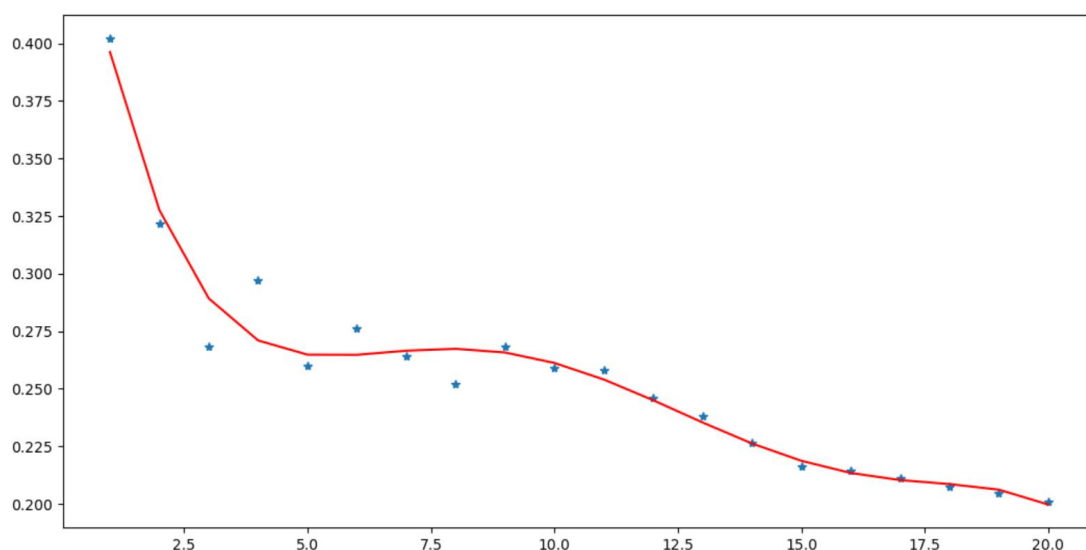
$$\begin{cases} \frac{dI}{dt} = \beta I(1 - \frac{I}{N}) \\ I(0) = I_0 \end{cases} \quad (6)$$

解得(2)：

$$I(t) = \frac{N}{1 + (\frac{N}{I_0} - 1)e^{-\beta t}} \quad (2)$$

将数据代入(见附录 5)求得当疫情传播一定时间后，平均感染率 β 会趋于平稳并呈现逐渐下降的趋势：

图 3. β 值的拟合曲线



解得 β 的拟合多项式：

$$\beta = -1.983 \times 10^{-6} x^5 + 0.0001218 x^4 - 0.002787 x^3 + 0.02899 x^2 - 0.1378 x + 0.5078$$

对 β 值的拟合曲线我们对其进行了显著性检验，求出其相关系数：

$$R^2 = 0.960570284$$

$$R^2 > 0.95$$

以上说明了 β 值的拟合曲线相关性良好，误差小，故我们选取了当 β 趋于平稳时一段时间的平均值：

综上，我们分别将 β 和 T_{av} 代入到公式(1)中解得病毒传播系数：

$$R_0=2.901063$$

5.2 问题二模型的建立与求解

5.2.1 问题二模型的建立

本问题中我们将基于 SIR 模型分别分析和预测武汉市地区(采取措施)和地区(未采取措施)的疫情状况。

其中 S 表示未感染人群在该地区的总人数；I 表示感染者在该地区的总人数；R 表示退出人群即死亡或康复或不再具备传染能力的人在该地区的总人数。故根据**问题分析 2.2**列出的 SIR 传播模型的示意图列出 SIR 微分方程组：

$$\begin{cases} \frac{dS}{dt} = -\alpha SI \\ \frac{dI}{dt} = \alpha SI - (\mu + \eta)I \\ \frac{dR}{dt} = (\mu + \eta)I \\ S + I + R = N \end{cases} \quad (7)$$

平均感染率 β 和推出率 $\mu + \eta$ 是 SIR 模型中两个描述疫情传播和控制重要的参数指标，在本文中我们将通过差分递推法求解这两个参数：

$$\alpha = -\frac{dS}{dt}/(SI) \quad (8)$$

$$\mu + \eta = -(\frac{dS}{dt} + \frac{dI}{dt})/I \quad (9)$$

$\frac{dS}{dt}$ 表示每日未感染人群减少的人数或每日新增的感染人数， $\frac{dI}{dt}$ 每日感染人数变化。将武汉市 1 月 24 日起的疫情数据带入(8) (9)中(数据见附录 2)，进行多项式拟合。

求解出 β 和 $\mu + \eta$ 两个重要参数后，我们使用差分递推法的定义：

$$\frac{dS}{dt} = S_t - S_{t-1} \quad (10)$$

$$\frac{dI}{dt} = I_t - I_{t-1} \quad (11)$$

将(10) (11)又分别代入到(8) (9)中得到：

$$\begin{cases} S_t = S_{t-1}/(\beta I_t + 1) \\ I_t = I_{t-1}/[1 + (\mu + \eta) - \beta S_{t-1}] \end{cases} \quad (12)$$

经过计算，我们发现该方程组无法求出关于 I_t 的精确解，参考文献[3]，得到(12)的迭代形

$$\begin{cases} S_t = \left[\alpha I_{t-1} - 1 + (\mu + \eta) + \alpha S_{t-1} - \sqrt{(\alpha I_{t-1} - 1 + (\mu + \eta) + \alpha S_{t-1})^2 - 4\alpha(1 + \mu + \eta)} \right] / 2\alpha \\ I_t = I_{t-1} / (1 + \mu + \eta - \alpha S_t) \end{cases} \quad (13)$$

由附录 6 中武汉 1 月 24 日起的疫情数据将 $S_0=11000000$ (人)， $I_0=922$ (人)代入至(13)进行迭代，得到武汉市未来的疫情预测图表。

针对英国地区的疫情预测，因上文叙述过的其数据波动性太大，无法使用 SIR 模型进行分析和预测，故在本文中，我们使用 Excel 中高级机器学习算法即指数三光滑平滑 (ETS) 算法预测。ETS 算法的使用“AAA 版本或基于现有值 (历史) 预测文莱之。预测值时指定的目标或日期，应为时间线的延续或者历史值得延续。”

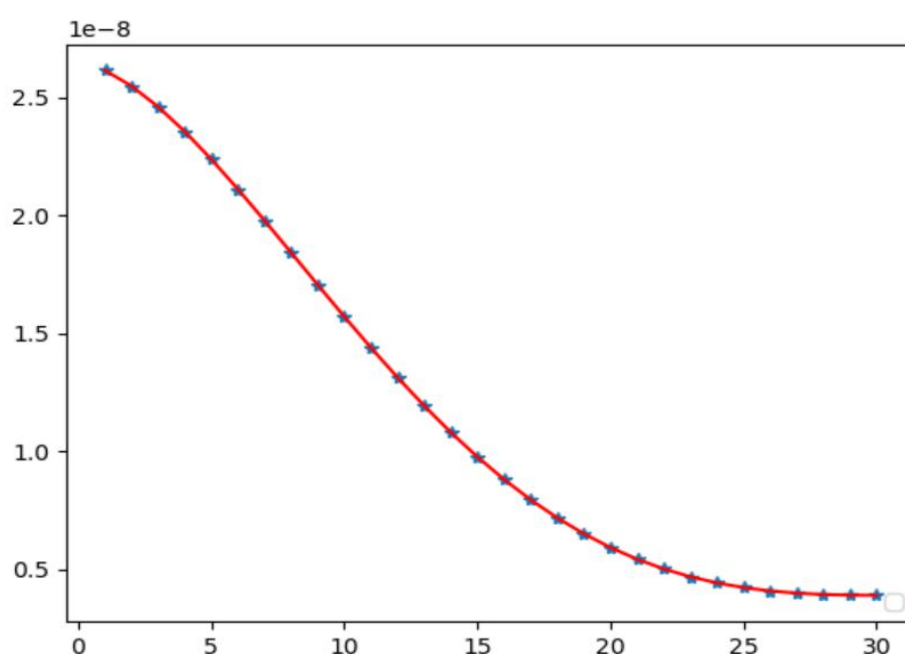
5.2.2 问题二模型的求解

武汉地区的预测结果：

α 的拟合结果：

$$\alpha = 1.387 \times 10^{-15} x^5 - 1.928 \times 10^{-13} x^4 + 9.214 \times 10^{-12} x^3 - 1.569 \times 10^{-10} x^2 - 2.564 \times 10^{-10} x + 2.653 \times 10^{-18}$$

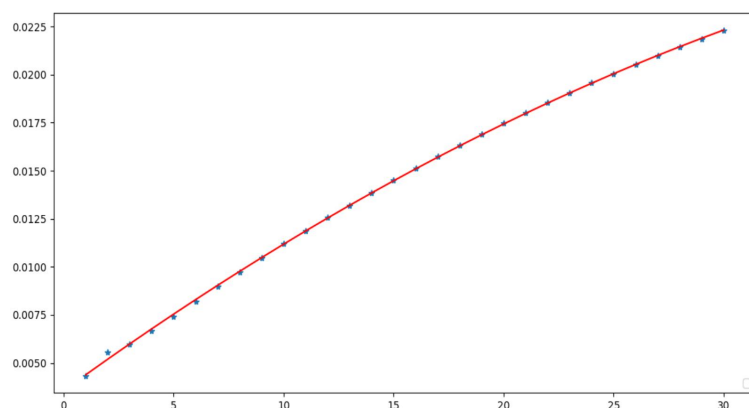
图 4. α 的拟合曲线



$\mu + \eta$ 的拟合结果：

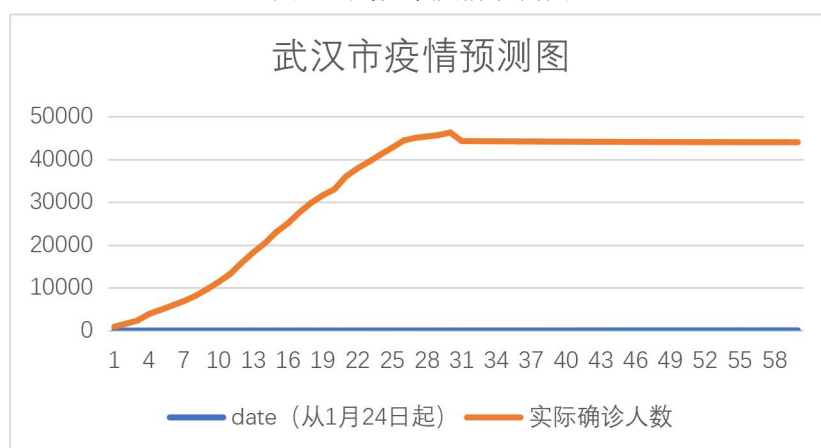
$$\mu + \eta = -6.84 \times 10^{-6} x^2 + 0.0008303x + 0.003564$$

图 5. $\mu + \eta$ 的拟合曲线



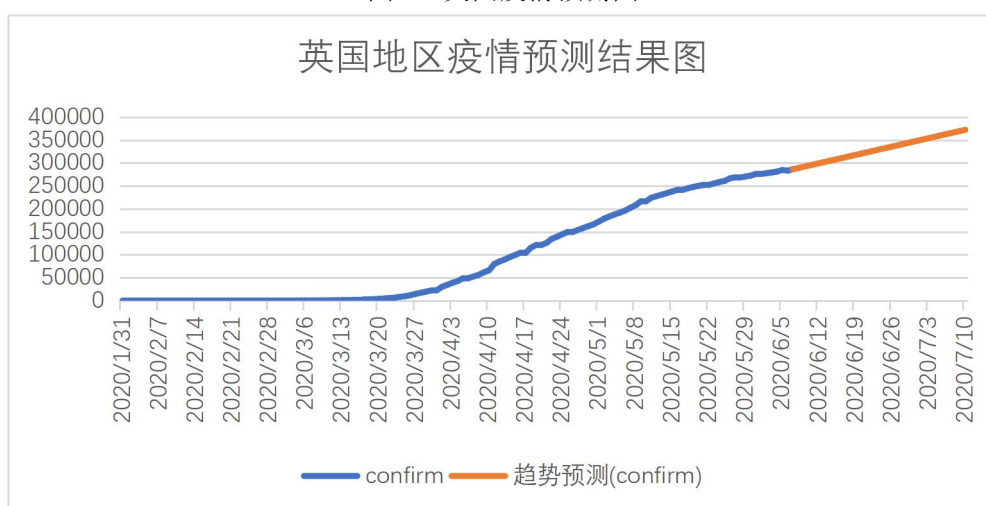
将附录 6 中武汉 1 月 24 日起的疫情数据将 $S_0=11000000$ (人), $I_0=922$ (人)代入至(13)进行迭代, 得武汉市未来的疫情预测图表:

图 6. 武汉市疫情预测图



英国地区预测结果:

图 7. 英国疫情预测图



通过分析数据与观察疫情预测图发现: 针对采取不同措施的不同地区, 采取严格措施的武汉市现在的疫情已经稳定下来, 每日新增的感染患者数量在个位数波动; 而选择群体免疫的英国如今的感染者数量依然以一个较大的速率增加。

5.3 问题三模型的建立与求解

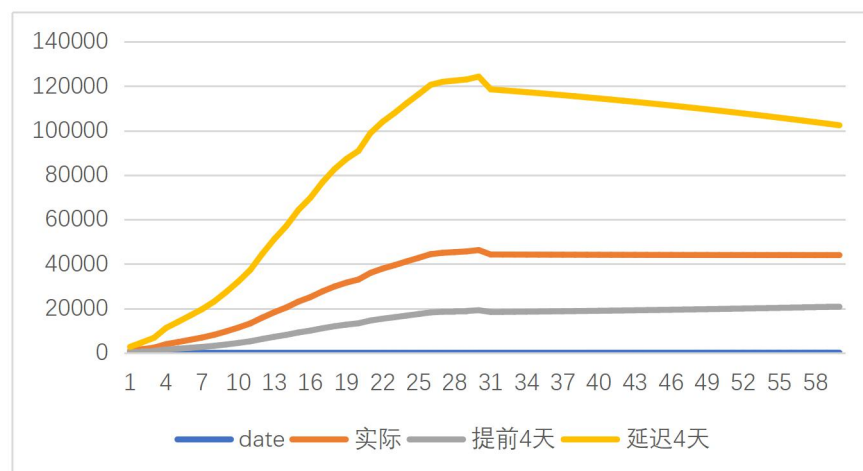
5.3.1 问题三模型的建立

在问题三中，我们将武汉市 1 月 24 日封城作为转折点，利用 logistic 模型求出在没有封城的情况下武汉市新冠肺炎在自然传播状态下传播至 1 月 20 日和 1 月 28 日的武汉市感染患者数量，分别以这两日为新封城日期，以这两日感染患者数量为新输入，在问题二 SIR 模型的基础上进行迭代，预测出在提前四天封城和延迟四天封城的前提，感染患者数量曲线。

5.3.2 问题三模型的求解

经过迭代，得到：

图 8. 武汉市感染患者的实际情况和提前(延迟)4 天封城 预测图



如图所示，在实际情况下武汉市(自 1 月 24 日起)第 31 天疫情趋于稳定，而若提前 4 天封城会使得武汉市疫情第 28 天趋于稳定，且感染人数远少于实际情况；若延迟 4 天封城会导致疫情控制的时间大大增加，并且感染患者的数量会远大于实际情况。

由此可见，在流行疾病的传染性不变的情况下，采取措施的时间越早，总感染人数便越少，且疫情控制的时间会显著减少。由此可见，政府第一时间采取防控措施、民众提升防护意识对疫情控制的重要性。

根据现有的 SIR 模型，模型内有判断标准，当 $S_0 < \frac{1}{\sigma}$ 时，传染病就不会继续向外扩散了，因此达到了制止传染病蔓延的效果，除此之外，在提高医疗和卫生等其他水平以外，使阈值 $\frac{1}{\sigma}$ 变大以外，另一个更加有效的途径就是想办法降低 S_0 。降低 S_0 ，这可以通过比如预防接种使群体免疫的办法做到。

在不考虑初始病人比例 i_0 的情况下，有公式 $S_0 = 1 - r_0$ ，那么，在这种情况下，传染病被阻断蔓延的条件 $S_0 < \frac{1}{\sigma}$ 就可以转换为：

$$r_0 \geq 1 - \frac{1}{\sigma} \quad (14)$$

依次来看，只需要通过群体免疫在初始时刻的移出者比例(即免疫比例)就可以制止传染病的蔓延。

要想让这种方法生效必须要求免疫者在人群中均匀分布，但是在现实中实现群体免疫的难度很大。据估计当时印度等国天花传染病的接触数 $\sigma=5$ ，则至少要有 80%的人接受免疫才行。据世界卫生组织报告，即使花费大量资金提高 γ ，也因很难做到免疫者的均匀分布，使得天花直到 1977 年才在全世界根除。而有些传染病的 σ 更高，根除就更加困难。

5.4 问题四模型的建立与求解

不同传染病病毒传播特征多样化，常见的传染病模型按照传染病类型分为 SI、SIR、SIRS、SEIR 模型等。

一般的白传染病流行范围内的人群分成以下几类：

1. S 类 健康人群(易受感染者)(Susceptible), 指没有受到感染的人群, 缺乏对于流行疾病的抵抗能力, 接触了感染者后易被感染;
2. E 类 暴露者(Exposed), 指接触了感染者, 处于潜伏期, 暂无传染能力;
3. I 类 感染者(Infection), 指染上传染病的人群, 可以传染给 S 类人群, 将其变为 E 类或 I 类;
4. R 类 康复者(Recovered), 指被隔离或康复或死亡而失去感染能力的感染者;

SI 模型

将人群分为 S 类和 I 类, SI 模型里面, 只考虑了易感者和感染者, 并且感染者不能够恢复, 因此可建立如下微分方程:

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI \end{cases} \quad (15)$$

这里 β 为传染率。在疾病传播期内, 所考察地区的总人数 $S(t) + I(t) = K$ 保持不变。利用这一守恒关系得:

$$\begin{cases} \frac{dI}{dt} = rI(1 - \frac{I}{K}) \\ r = \beta K \end{cases} \quad (16)$$

这是一个 **logistics** 模型。其指数增长率 $r = \beta K$ 正比于总人数 K 和传染率 β 。这个模型有两个主要结论:

- (1) 指数增长率 r 正比于总人数。当传染率 β 一定时，一定染病地区内的总人数 K 越多，传染病爆发的速度越快，说明了隔离的重要性，必须要减少感染者和易感者的接触；
- (2) 在 $I=K/2$ 时，患病人数增长最快，此时各医院门诊处于最危险的时刻，医护人员需要做好防护，加大力度，稳定感染者；

SIR 模型

有的时候，感染者在康复了之后，就有了抗体，于是后续就不再会获得此类病症，这种时候，考虑 SIS 模型就不合适了，需要考虑 SIR 模型。此类病症有麻疹，腮腺炎，风疹等。SIR 模型进一步考虑了病人的康复过程。模型的微分方程为：

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \quad (17)$$

总人数 $S(t) + I(t) + R(t) = N$ 为常数。这里假设病人康复后就获得了永久免疫，因而可以移出系统。对于致死性的传染病，死亡的病人也可以归入 R 类。因此 SIR 模型只有两个独立的动力学变量 I 和 S ，它们的相轨迹满足

$$\frac{dI}{dS} = \frac{\gamma}{\beta S} - 1 \quad \longrightarrow \quad I + S - \frac{\gamma}{\beta} \ln S = \text{const} \quad (18)$$

给定 $t=0$ 时刻的条件 $S = S_0$ ，随着 S 从 S_0 开始单调递减，染病人数 I 在 $S = \gamma/\beta$ 时达到峰值，随后一直回落，直到减为零。此时剩余一部分易感人群 S_∞ ，而疾病波及到的总人数为 R_∞ ，二者可由总人数守恒和相轨迹方程解出。

SIRS 模型

有很多小病是可以恢复并且反复感染的，例如日常的感冒，发烧等。在这种情况下，感染者就有一定的几率重新转化成易感者，形如

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \alpha R \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I - \alpha R \end{cases} \quad (19)$$

总人数 $S(t) + I(t) + R(t) = N$ 为常数。参数 α 决定康复者获得免疫的平均保持时间。系统有两个不动点 $S = N (I = R = 0)$ 或 $S = \gamma/\beta$ 。前者表示疾病从研究

地区消除，而后者则是流行状态。消除流行病的参数条件是 $\gamma > \beta N$ 。若做不到，则要尽量减小 α 而增加 γ ，使更多人保持对该疾病的免疫力。

SEIR 模型

如果所研究的传染病有一定的潜伏期，与病人接触过的健康人并不马上患病，而是成为病原体的携带者，归入 E 类。此时有

$$\begin{cases} \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \alpha E - \gamma_2 I \\ \frac{dE}{dt} = \beta IS - (\alpha + \gamma_1) E \\ \frac{dR}{dt} = \gamma_1 E + \gamma_2 I \end{cases} \quad (20)$$

仍有守恒关系 $S(t) + E(t) + I(t) + R(t) = N$ ，病死者可归入 R 类。潜伏期康复率 γ_1 和患者康复率 γ_2 一般不同。潜伏期发展为患者的速率为 α 。与 SIR 模型相比，SEIR 模型进一步考虑了与患者接触过的人中仅一部分具有传染性的因素，使疾病的传播周期更长。疾病最终的未影响人数 S_∞ 和影响人数 R_∞ 可通过数值模拟得到。

综上 健全我国传染病类突发公共卫生事件风险防控建议

应对突发公共卫生事件主要为三部走战略：早发现、早隔离、早治疗。

早发现：

1. 对现有尤其是在一线岗位的医护人员进行培训，加强其辨别能力。
2. 本次新冠肺炎在早期，因为一些工作人员的不作为和抱着“宁求无过，不求有功”的心态，没有及时上报中央，给防卫工作耽误了极大的时间，所以应该建立一个恰当的突发疫情上报机制，在传播早期及时上报，尽量缩短病毒在指数传播过程中的时间。
3. 指派区域医生，实行责任到人制度，在突发公共卫生事件确定后，由区域医生指挥对所负责区域人员的安全检测。

早隔离：

1. 在全国各地建立常规隔离点，并配备一定的医疗卫生工具，方便在突发公共卫生事件后及时启用，避免各大医院拥挤不堪，从而影响到其他正常病患者的及时治疗
2. 对于疫情的源头地区和疑似源头地区，应该防止人员外流，尽量避免造成大规模传播，防止疫情发生二次扩散。同时对于重点地区应该着力做好防控工作

3. 利用新兴高科技设备来为防控提供支持，在现有科技大力发展的背景下，利用智能设备和云服务平台以及大数据监控，减少人与人直接接触的次数，对于信息的收集，填报，分析以及开会研讨等简单的工作可以直接放到互联网上进行，尽一切可能阻止疫情扩散。

早治疗：

1. 及时组织国内外优秀的医学科研人员，用科学的方式去研究病毒，用准确的逻辑链去调查传染源和传播途径，及时监察病毒是否发生变异，加快特效药，有效要和疫苗的研发进度，为这些药品的上市和流通提供紧急特殊渠道，加快面世进度，提高医治率，降低死亡人数。

2. 从长期来看，我们应该科学地做好医学准备和应急防控措施，加大相关学科研究和基础建设，加大人才引进步伐和人才培养计划。

六、模型评价

模型优缺点：

SIR 的基础是微分方程，根据微分方程求解，SIR 模型可以通过现有的数据非常准确地得出拟合曲线，并利用拟合出的曲线，进行科学分析，得出不让病毒继续向外扩散的具体方法

但是，此模型对于人群的分类不够精细，没有考虑到无症状感染人员和疑似，治愈患者，同时也没考虑到环境要素例如有效的隔离措施。

对于及时有效的反馈，该模型也未考虑到这一点，因此，在通过该模型预测的过程中，越往后，预测的准确性和精度越差。采用已有数据无法体现出这些因素对疫情控制的影响，这是模型致命的漏洞。为此必须引入反馈机制达到自我调整的功能。

还有一点，微分方程求解是很麻烦的，并且对初始数据要求较高，这样一来，就增加了数据预处理的步骤和难度，复杂度。

基于上述分析，我们在本次建模过程中加入了一个新的反馈方法。虽然无法做到实时反馈，但我们只需要将连续的时间划分为等距的时间段，然后按照时间段反馈，这和每日统计疫情数据比较相似。于是，连续模型就改为离散模型。

七、参考文献

- [1] Wallinga J ,Lipstch M 2007 How generation intervals shape the relationship between growth rates and reproductive numbers.B 274 599
- [2] LiQ, GuanXH, WuP, WangXY, ZhouL, TongYQ, RenRQ, LeungS, LauE, WongJ, XingXS, XiangNJ, WuY, LiC, ChenQ, LiD, LiuT, ZhaoJ, LiM, TuWX, ChenCD, JinLM, YangR, Wang Q, ZhouSH, WangR, LiuH, LuoYB, LiuY, ShaoG, LiH, TaoZF, YangY, DengZQ, LiuBX, MaZ T, ZhangYP, ShiGQ, LamT, WuJ, GaoG, Cowling B, Yang B, Leung G, Feng Z J 2020 N. Engl. J. Med. 382 1199
- [3] 盛华雄, 吴琳, 肖长量 新冠肺炎疫情传播建模分析与预测[J], 系统仿真学报, 2020, 759-766
- [4] 王聪, 严洁, 王旭, 李敏 新型冠状病毒肺炎早期时空传播特征分析[J], 物理学报, 2020, 69(8), 249-258
- [5] 王志心, 刘治, 刘兆军 基于机器学习的新型冠状病毒(COVID-19)疫情分析及预测[J] 生物医学工程研究 2020, 39(1), 1-5

八、附录

附录一：问题一的求解代码（python3.8）

```
import numpy as np
import pandas as pd
import math

#p 就是上文提到的 $\beta$ 
n=11000000
i0=1
p=0
k=[1,1,2,5,5,5,8,8,12,14,16,25,29,37,40,45,47,49,59,68,78,90,102]
for i in range(2,len(k)):
    cnt=(n*i0/k[i]-i0)/(n-i0)
    p+=(-math.log(cnt))/(i+1)
p/=len(k)-2
print(str(p))
#p=0.2534751872377965
```

附录 2： 校验国内前期 β 的准确性（python3.8）

```
import math

ave=0.2534751872377965
sum=0
n=11000000
```



```

i0=1
p=0
k2=[1,1,2,5,5,5,8,8,12,14,16,25,29,37,40,45,47,49,59,68,78,90,102]
k=list()
for i in range(3,len(k2)):
    cnt=(n*i0/k2[i]-i0)/(n-i0)
    k.append((-math.log(cnt))/(i+1))

for x in range(1,21):
    p=-1.983e-06*x**5 + 0.0001218*x**4 - 0.002787*x**3 + 0.02899*x**2 -
0.1378*x + 0.5078
    sum=sum+(p-k[x-1])**2

sum2=0
for x in range(1,21):
    p=-1.983e-06*x**5 + 0.0001218*x**4 - 0.002787*x**3 + 0.02899*x**2 -
0.1378*x + 0.5078
    sum2+=(p-0.254596492)**2
sum3=1-(sum/sum2)
print(str(sum3))
#sum3=0.9605702849424397

```

附录 3： 武汉封城后 λ 的拟合（python3.8）问题二中的 α 就是 λ

```

# coding: utf-8
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from datetime import datetime

s=11000000
china = pd.read_csv('wuhan_30.csv', encoding='GBK')
data=china.iloc[:,].values
k=list()
for i in range(30):
    k.append(data[i][4]/((s-data[i][1])*data[i][1]))
    print(k[i])

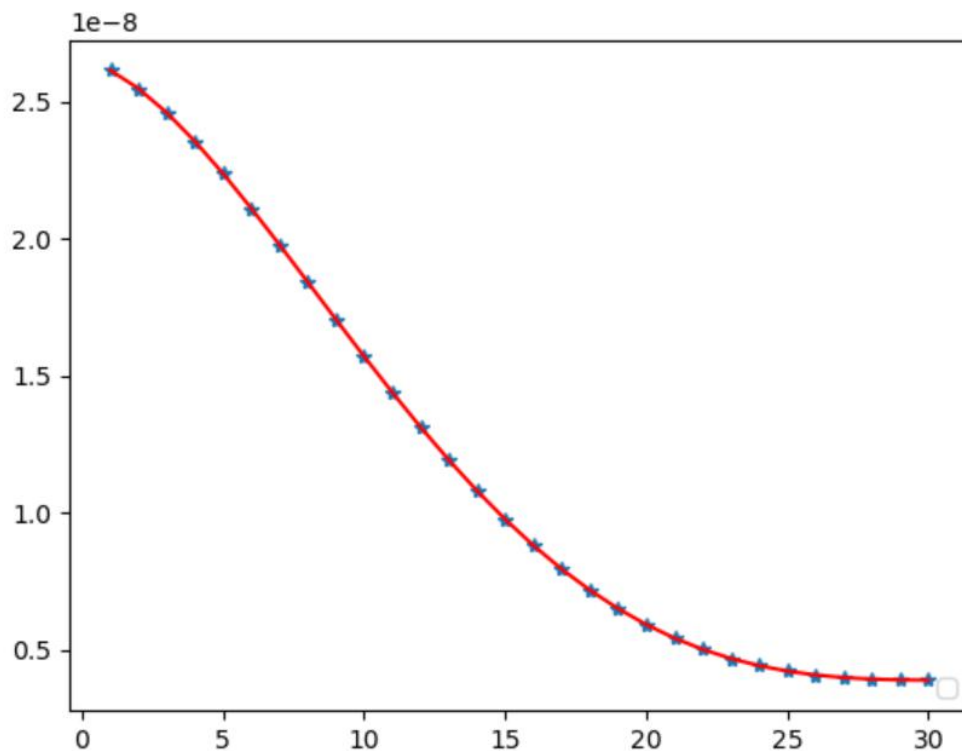
```

```

'''
for i in range(30):
    k.append(data[i][1]/((s-data[i][2])*data[i][2]))
    print(k[i])
'''

x = np.arange(1, len(k)+1, 1)
y = np.array(k)
z1 = np.polyfit(x, y, 5) # 用 5 次多项式拟合
p1 = np.poly1d(z1)
print(p1) # 在屏幕上打印拟合多项式
yvals=p1(x)
plot1=plt.plot(x, y, '*')
plot2=plt.plot(x, yvals, 'r')
plt.legend(loc=4)
plt.show()
#          5          4          3          2
#1.387e-15 x - 1.928e-13 x + 9.214e-12 x - 1.569e-10 x - 2.564e-10 x + 2.653e-08

```



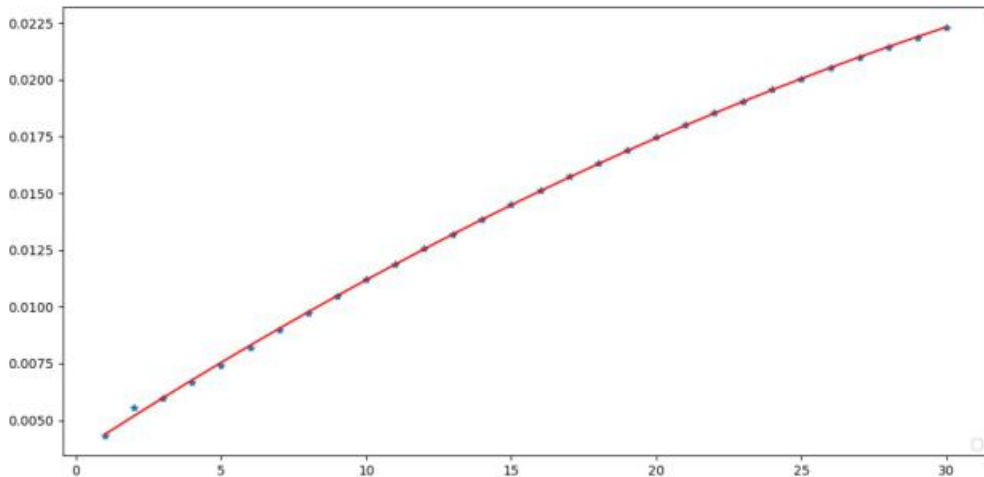
附录 3：武汉封城后 $\mu+\eta$ 的拟合曲线（python3.8）

```
# coding: utf-8
import requests
import json
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from datetime import datetime

s=11000000
china = pd.read_csv('wuhan_30.csv', encoding='GBK')

data=china.iloc[:,].values
k=list()
for i in range(30):
    k.append((data[i][4]*2-data[i][5])/data[i][1])
    print(k[i])

x = np.arange(1, len(k)+1, 1)
y = np.array(k)
z1 = np.polyfit(x, y,2)#用 2 次多项式拟合
p1 = np.poly1d(z1)
print(p1) #在屏幕上打印拟合多项式
yvals=p1(x)
plot1=plt.plot(x, y, '*')
plot2=plt.plot(x, yvals, 'r')
plt.legend(loc=4)
plt.show()
```



```
#          2
#-6.84e-06 x + 0.0008303 x + 0.003564
```

附录 4： 迭代出正常情况下感染人数增长情况（python3.8）

```
# coding: utf-8
import requests
import json
import math
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from datetime import datetime

e = pd.read_csv('wuhan_30.csv', encoding='GBK')
s=11000000
#it=46063
'''
for i in range(1,31):
    x=i
    la=1.387e-15*i**5 - 1.928e-13*x**4 + 9.214e-12*x**3 - 1.569e-10*x**2 -
2.564e-10*x + 2.653e-08
    u=-6.782e-06*x**2 + 0.0008275*x + 0.003428
    st=(la*it+1+u+la*(s-it)-math.sqrt((((la*it+1+u+la*(s-it))**2-4*la*(1+u))))/(2*la)
    it=it/(1+u+la*st)
```

```

print(str(it))
'''
for i in range(31,61):
    st=0.9593030450458995
    it=45256.602628212655
    x=i
    la=1.387e-15*i**5 - 1.928e-13*x**4 + 9.214e-12*x**3 - 1.569e-10*x**2 -
2.564e-10*x + 2.653e-08
    u=-6.782e-06*x**2 + 0.0008275*x + 0.003428
    st=(la*it+1+u+la*(s-it)-math.sqrt((((la*it+1+u+la*(s-it))**2-4*la*(1+u))))/(2*la)
    it=it/(1+u+-la*st)
    print(str(int(it)))

```

附录 5： logic 函数推导出不采取措施的数据（python3.8）

```

import math

for i in range(1,101):
    k=math.e**(-i*0.254596492)
    k=k*(11000000-1)+1
    k=11000000/k
    print(str(k))

```

附录五： 问题一中武汉 12 月 9 日-12 月 31 日数据

date	确诊	死亡	治愈	新增	退出
1	922	38	32	265	526
2	1618	45	40	453	897
3	2348	63	42	635	1256
4	3890	85	42	1007	1988
5	4855	104	75	1194	2352
6	5861	129	82	1359	2670
7	6889	159	103	1497	2932
8	8115	192	139	1642	3205
9	9659	224	171	1810	3519
10	11342	265	224	1957	3787
11	13234	313	303	2092	4027
12	15851	362	368	2286	4373
13	18267	414	431	2393	4545
14	20418	478	534	2423	4563
15	23053	545	698	2471	4608
16	25082	608	877	2426	4473
17	27652	681	1044	2413	4391
18	29854	748	1206	2354	4221
19	31608	820	1377	2259	3984
20	32994	1036	1915	2147	3718
21	35991	1016	2016	2146	3644
22	37914	1123	2502	2090	3477
23	39462	1233	2915	2031	3310
24	41152	1309	3458	2000	3195
25	42752	1381	4219	1985	3113
26	44412	1497	4895	1993	3074
27	45027	1585	5448	1974	3003
28	45346	1684	6214	1959	2946
29	45660	1774	7206	1959	2920
30	46259	1856	8171	1981	2931

附录六：问题二中武汉地区疫情预测数据以及提前四天和延迟四天封城预测数据

date	实际	提前4天	延迟4天	自由传播
1	922	349	2676	1
2	1618	613	4690	1
3	2348	893	6788	2
4	3890	1484	11216	2
5	4855	1860	13956	3
6	5861	2254	16794	4
7	6889	2661	19673	5
8	8115	3147	23097	7
9	9659	3761	27401	9
10	11342	4433	32074	12
11	13234	5192	37310	16
12	15851	6240	44563	21
13	18267	7214	51217	27
14	20418	8089	57098	35
15	23053	9160	64304	45
16	25082	9995	69790	58
17	27652	11050	76755	75
18	29854	11963	82670	97
19	31608	12701	87317	126
20	32994	13295	90923	162
21	35991	14541	98948	209
22	37914	15359	103989	270
23	39462	16030	107977	349
24	41152	16761	112331	450
25	42752	17460	116417	581
26	44412	18187	120643	749
27	45027	18490	122008	966
28	45346	18674	122554	1247
29	45660	18858	123073	1608
30	46259	19162	124345	2074
31	44258	18392	118613	2676
32	44240	18446	118198	3452
33	44223	18502	117775	4452
34	44207	18560	117342	5743
35	44192	18621	116901	7407
36	44177	18683	116447	9552
37	44163	18749	115984	12319
38	44149	18816	115509	15886
39	44136	18885	115025	20483
40	44123	18957	114529	26408
41	44111	19030	114024	34042
42	44100	19107	113510	43873
43	44089	19185	112983	56528
44	44079	19265	112448	72809
45	44069	19348	111900	93740
46	44060	19433	111343	120621
47	44052	19521	110777	155101
48	44044	19610	110199	199257
49	44037	19702	109612	255687
50	44030	19796	109012	327613
51	44024	19892	108403	418983
52	44018	19990	107783	534560
53	44013	20091	107152	679971
54	44009	20194	106513	861678
55	44005	20300	105861	1086830
56	44002	20407	105200	1362903
57	43999	20517	104527	1697099
58	43997	20629	103844	2095424
59	43996	20743	103151	2561497
60	43995	20860	102447	3095203

附录七：问题二中英国地区疫情预测数据

2020/2/2	2
2020/2/3	2
2020/2/4	2
2020/2/5	2
2020/2/6	3
2020/2/7	3
2020/2/8	3
2020/2/9	3
2020/2/10	8
2020/2/11	8
2020/2/12	8
2020/2/13	9
2020/2/14	9
2020/2/15	9
2020/2/16	9
2020/2/17	9
2020/2/18	9
2020/2/19	9
2020/2/20	9
2020/2/21	9
2020/2/22	9
2020/2/23	9
2020/2/24	13
2020/2/25	13
2020/2/26	13
2020/2/27	15
2020/2/28	19
2020/2/29	23
2020/3/1	35
2020/3/2	40
2020/3/3	51
2020/3/4	85
2020/3/5	206
2020/3/6	163
2020/3/7	163
2020/3/8	319
2020/3/9	373
2020/3/10	456
2020/3/11	590
2020/3/12	798
2020/3/13	1140
2020/3/14	1140
2020/3/15	1543
2020/3/16	1543
2020/3/17	2626
2020/3/18	2626
2020/3/19	3269
2020/3/20	3983
2020/3/21	5027
2020/3/22	5683
2020/3/23	6650
2020/3/24	8077
2020/3/25	9529
2020/3/26	11658
2020/3/27	14579
2020/3/28	17089
2020/3/29	19522
2020/3/30	22141
2020/3/31	22472
2020/4/1	29842
2020/4/2	34123
2020/4/3	38681
2020/4/4	42433
2020/4/5	48388
2020/4/6	48451
2020/4/7	52302
2020/4/8	55957
2020/4/9	61497
2020/4/10	65872
2020/4/11	79841
2020/4/12	85173
2020/4/13	89554
2020/4/14	94823
2020/4/15	99455
2020/4/16	104133
2020/4/17	104145
2020/4/18	115299
2020/4/19	121168
2020/4/20	121174