



学校代码: 10184

分 类 号:

应用统计 硕士专业学位 论文

基于SIR模型对新冠肺炎疫情基本再生数 的研究

RESEARCH ON THE BASIC REPRODUCTION NUMBER
OF COVID-19 BASED ON THE SIR MODEL

薛 依 琳

应用统计

延 边 大 学

分类号

密级 公开

UDC

学号 2020050003

延边大学硕士学位论文

基于 SIR 模型对新冠肺炎疫情基本 再生数的研究

研 究 生 姓 名

薛依琳

培 养 单 位

延边大学

指导教师姓名、职称

尹哲 教授

学 科 专 业

应用统计


研 究 方 向

大数据和数据挖掘

论 文 提 交 日 期

2022 年 6 月 5 日

本论文已达到应用统计硕士专业学位论文要求

答辩委员会主席	尹哲	
答辩委员会委员	文香丹	
答辩委员会委员	洪成	
答辩委员会委员	金哲植	
答辩委员会委员	尹永学	

延 边 大 学

2022 年 6 月 5 日

学位论文独创性声明

本人郑重声明：所呈交的学位论文系本人在导师指导下独立完成的研究成果。尽我所知，除了文中特别加以标记和致谢的部分外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含本人为获得任何教育机构的学位或学历而使用过的材料。与我一同工作的同事对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

本人如违反上述声明，愿意承担由此引发的一切责任和后果。

研究生签名：薛恒林 日期：2022年6月4日

学位论文使用授权声明

本人在导师指导下所完成的学位论文，学校有权保存其电子和纸制文档，可以借阅或上网公布本学位论文的全部或部分内容，可以向有关部门或机构送交并授权其保存、借阅或上网公布本学位论文的全部或部分内容。对于保密论文，按保密的有关规定和程序处理。

本学位论文属于：

1. 保密 ☐，在 年解密后适用于本声明；2. 不保密 ☒.

研究生签名：薛恒林 导师签名：尹哲 日期：2022年6月4日

摘 要

传染病问题从古至今一直是威胁人类生存的重大问题。自 2019 年末新型冠状病毒肺炎疫情首先在中国湖北省爆发,三年来病毒肆虐全球各地,造成了难以估计的损失,是人类共同面对的浩劫。我国一直以来积极采取各项医疗防控措施,为世界人民做出表率。因此,研究新型冠状病毒肺炎的传播机制、传播途径、感染周期、潜伏期等各种问题都是至关重要的,通过对病毒的研究能够让人们增加了解与认知,以便于采取各项防控措施,打赢这场抗疫之战。

本文基于 SIR 经典仓室模型,选取湖北和长春两地数据,对新型冠状病毒肺炎疫情展开研究。首先介绍了当前我们所面临的疫情现状,对此引入 SIR 模型并对其相关的理论基础与数学知识进行简要说明;接下来利用湖北与长春两地疫情相关数据进行实证分析,采用最小二乘法对传染率系数 β 和移出率系数 γ 进行参数估计,计算出衡量疾病情况的重要指标基本再生数 R_0 ,并利用 R_0 分析了采取不同的防控措施对病毒传播情况的影响。

关键词: 新型冠状病毒, SIR 模型, 最小二乘法, 基本再生数

Abstract

The problem of infectious diseases has been a major problem threatening human existence since ancient times. Since the novel coronavirus pneumonia first broke out in Hubei Province, China at the end of 2019, the virus has ravaged all over the world in the past three years, causing incalculable losses, and it is a catastrophe faced by all mankind. My country has been actively taking various medical prevention and control measures to set an example for the people of the world. Therefore, it is very important to study the transmission mechanism, transmission route, infection cycle, incubation period and other issues of the new coronavirus pneumonia. Through the study of the virus, people can increase their understanding and cognition so as to facilitate various prevention and control measures and to win this battle against the epidemic.

Based on the SIR classic warehouse model, this paper selects data from Hubei and Changchun to conduct research on the outbreak of novel coronavirus pneumonia. First, the current situation of the epidemic situation we are facing is introduced, and the SIR model is introduced and its related theoretical basis and mathematical knowledge are briefly explained. The infection rate coefficient β and the emigration rate coefficient γ were used for parameter estimation, and the basic reproduction number, an important indicator to measure the disease situation, was calculated, and the impact of different prevention and control measures on the virus transmission was analyzed.

Key words: COVID-19, SIR model, least squares method, basic reproduction number

目 录

摘要	I
Abstract	II
第一章 绪论	1
1.1 研究背景	1
1.1.1 人类与传染病	1
1.1.2 COVID-19 病毒以及发展情况	2
1.2 文献综述	5
1.2.1 传染病模型研究综述	5
1.2.2 新冠疫情研究综述	6
1.3 研究内容	7
第二章 SIR 传染病模型	8
2.1 传染病模型基础	8
2.1.1 传染病动力学基本知识	8
2.1.2 几种传染病动力学模型	9
2.2 数学知识	13
2.3 SIR 仓室模型	14
2.3.1 SIR 经典模型	14
2.3.2 重要参数估计	19
第三章 实证分析	21
3.1 数据选取	21
3.2 数据分析	23
3.2.1 湖北疫情数据分析	23
3.2.2 长春疫情数据分析	27
3.3 参数估计	28
3.3.1 针对湖北疫情的 SIR 模型参数估计	29
3.3.2 针对长春疫情的 SIR 模型参数估计	30
3.4 模型应用与分析	33

第四章 总结与反思	38
参考文献	40
致 谢	43
附录 A 湖北省 COVID-19 疫情数据	44
附录 B 长春市 COVID-19 疫情数据	46
附录 C SIR 模型 python 代码	47

第一章 绪论

1.1 研究背景

1.1.1 人类与传染病

传染病是一种由各种病原体引起的疾病，它最大的特点在于能够在人与人或人与动物之间互相传染进而达到广泛流行，这也是对人类来讲最致命的一点。病毒既可以直接传播，也可以借由各种媒介进行传播，常见的如气体传播、水源传播、食物传播、接触传播、土壤传播、垂直传播（母婴传播）、体液传播、粪口传播等。按照疾病传播的范围以及严重程度可以分为：散发、流行、大流行、暴发。一旦达到“大流行”及以上标准即可称为“瘟疫”，危害难以预计，除可造成死亡、摧毁城市、政治、国家、瓦解文明，甚至可以歼灭族群、物种。因此在人类历史上，传染病的预防与治疗一直是关乎人类福祉的重要问题。

有史以来，人类从未停止受到各种传染病的肆虐侵扰，可以说人类发展史伴随着人类与传染病不屈不挠的斗争。早在公元 2 世纪，Antonine 瘟疫^[1]在罗马帝国流行，引起了社会急剧动荡，人心惶惶，外敌借此趁虚而入，导致了罗马帝国最终毁灭。14 世纪，亚欧大陆鼠疫大流行，当时这种瘟疫被人们称为“黑死病”。可怕的黑死病在欧洲猖獗了 3 个世纪，2500 万余人丧生于这场浩劫，当时约 30% 的欧洲人口死亡，恐怖程度不亚于任何一场战争。1507 年前后，一个携带病毒的黑人奴隶将天花从欧洲带到拉美地区，一片净土就此沦陷。残暴不仁的殖民者无视印第安人的死活，故意将各种患病病人接触过的传染源如枕头、被子等各种生活用品送给印第安人。由欧洲传来的腮腺炎、麻疹、霍乱、淋病和黄热病等恶性传染疾病也随之而来。仅 100 年的时间，美洲大陆上 3000 万的原住民人口只剩下不到 100 万人，传染病的威力强大到足以令印第安人灭绝。而在第一次世界大战时期，美军军营当中部分人突然出现了普通的感冒症状，并没有引起各方的密切关注以及及时救治，这直接导致疾病迅速在战场上爆发，而且传播到战场以外的各个国家，美洲、欧洲、亚洲等国家均未能幸免。据不完全统计，当时处在战场上 20 万英军、40 万法军以及 50 万德军，均无一例外地感染了此病，百分之十的致死率让军队上下人心惶惶，一场流感提前终止了战争并造成全球至少有 2000 万人丧命。到了 20 世纪下半叶以后，另外一个严重影响人类健康的传染病

是艾滋病（AIDS）。艾滋病是慢性传染病，自 1980 年出现第一例患者至今都没有很好的治疗方法，全世界约有 7570 万人感染艾滋，其中，3270 万人（约 43%）已经死亡。截至 2020 年，世界卫生组织宣告艾滋病是全球第十九大死因。到 2101 年如果不采取有效措施，艾滋病将导致非洲撒哈拉沙漠以南地区一半人口死亡^[2]。2002 年 11 月，非典型肺炎的爆发让我国人民承受了深重的灾难。它爆发于我国广东省，历时仅仅半年左右，我国患病人数达 8422 例，传播速度令人咂舌，最终扩散到东南亚甚至全球范围。随着时间的推移，甲型 H1N1 流感、脑膜炎、埃博拉、海地霍乱、刚果麻疹、中东呼吸综合征以及离我们最近的新型冠状病毒肺炎等等各种传染病肆虐着人类的精神与躯体，影响着我们生活的方方面面，给人类社会带来不可估量的打击与损失。

无论历史还是今天，传染病问题都是全人类甚至全物种所面临的重大公共威胁。因此，传染病的防治问题迫在眉睫，应当引起全人类的重视。研究传染病的原理、传播机制、患病周期、治疗方法、干预措施等等问题都是十分必要并且至关重要的。

1.1.2 COVID-19 病毒以及发展情况

新型冠状病毒肺炎，简称“新冠肺炎”。2019 年 12 月以来，湖北省武汉市陆续收到多名肺炎患者，发病原因起初并不明确，这些患者均与华南海鲜市场产生联系，后证实这是一种急性呼吸道传染病，由新型冠状病毒感染产生故简称新冠肺炎，世界卫生组织将其命名为“2019 冠状病毒病”，英文名称为“COVID-19”，CO 代表冠状（Corona），VI 代表病毒（Virus），D 代表疾病（Disease）。国际病毒分类委员会则将新型冠状病毒命名为“SARS-CoV-2”（Severe Acute Respiratory Syndrome Coronavirus 2），并认定这种病毒是 SARS 冠状病毒的姊妹病毒。

根据现有的病例资料，新冠肺炎主要症状为发烧、咳嗽、浑身无力，少数患者可能会存在流鼻涕、拉肚子、鼻塞等常见的类似于胃肠感冒的症状。重症病例大多数在 7 天左右出现难以呼吸的症状，更为严重的是逐渐发展为窒息、休克甚至是发展为全身性的代谢性问题、凝血问题以及器官衰竭等。目前的研究能够确定的是新冠肺炎传播途径主要有以下几种：直接传播、接触传播和气溶胶传播。直接传播是指被感染者说话、咳嗽、打喷嚏、呼吸等与外界发生气体交换而同时

他人近距离直接吸入导致的感染；气溶胶传播是指上述各种飞沫携带病毒混合在空气中，形成气溶胶，吸入后导致感染；接触传播是指各种飞沫堆积在我们能够接触到的物品上，病毒由物品经过我们的手，再接触口腔、鼻腔、眼睛等粘膜，导致感染。

2019 年末武汉疫情迅速爆发，在短短 30 天内，确诊人数超过了 2003 年“非典”确诊病例数。2020 年 1 月 30 日，世界卫生组织宣布，将新型冠状病毒疫情列为国际关注的突发公共卫生事件。面对严峻态势，中国政府积极防疫，封锁城市并建立火神山、雷神山医院对患者进行集中隔离与治疗，无数抗疫人员前赴后继。截至 2020 年 3 月 1 日，全国确诊 80026 例，死亡 2912 例，与此同时中国以外共 61 个国家和地区报告新冠肺炎累计确诊病例 7600 余例，已有七国的累计确诊病例逾一百例。香港大学医学院院长梁卓伟分析：“从技术角度来看，新冠疫情目前已经算是全球大流行。疫情在国内的第一波已经大致受控，但是在国外的第一波才刚刚开始，大爆发已经不可避免。”就此疫情席卷全球。2020 年 4 月 3 日，全球新冠肺炎确诊病例突破 100 万例，死亡病例达 51485 例。这是本轮疫情的一个重要时间节点，它意味着本轮新冠肺炎疫情已经成为当代全球最大规模的公共卫生危机。在我国政府与人民坚持不懈的努力下，4 月 8 日零时，武汉市解除离汉离鄂通道管控措施，标志着中国阶段性战胜了疫情。随着时间的推移，北京、新疆、大连、青岛等地陆续成为高风险地区。2020 年底，全球累计确诊新冠肺炎病例超 8000 万例。

2021 年我国研制出新冠疫苗并倡导广大人民群众积极注射第一针、第二针以及加强针，对疫情的控制起到了较好的作用，累计接种新冠疫苗 282341.8 万剂次。国内疫情基本稳定，全年新增确诊病例 231 例，其中境外输入病例 56 例。而世界人民依然处于水深火热当中，全球累计确诊新冠病例超 2.5 亿例。

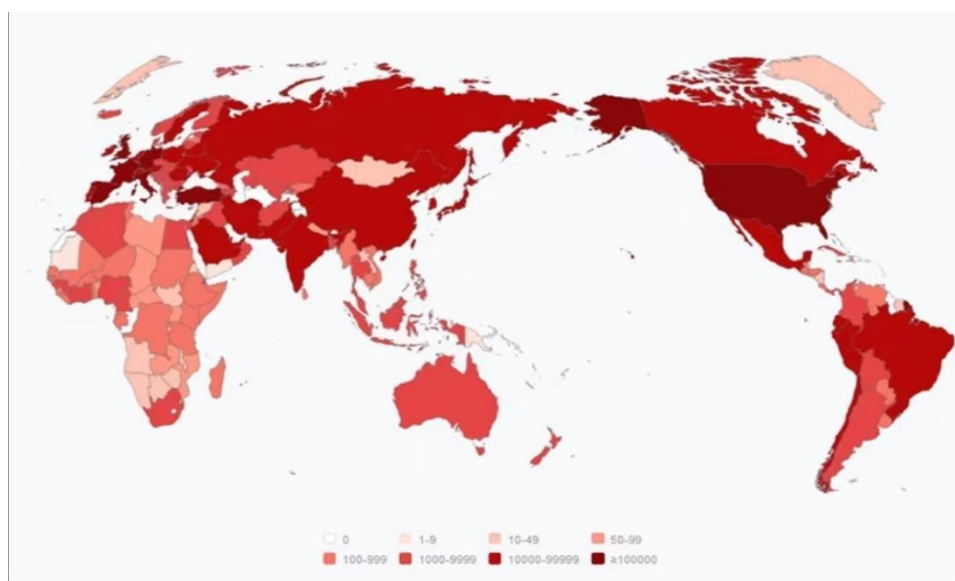


图 1.1 全球新冠肺炎累计确诊分布

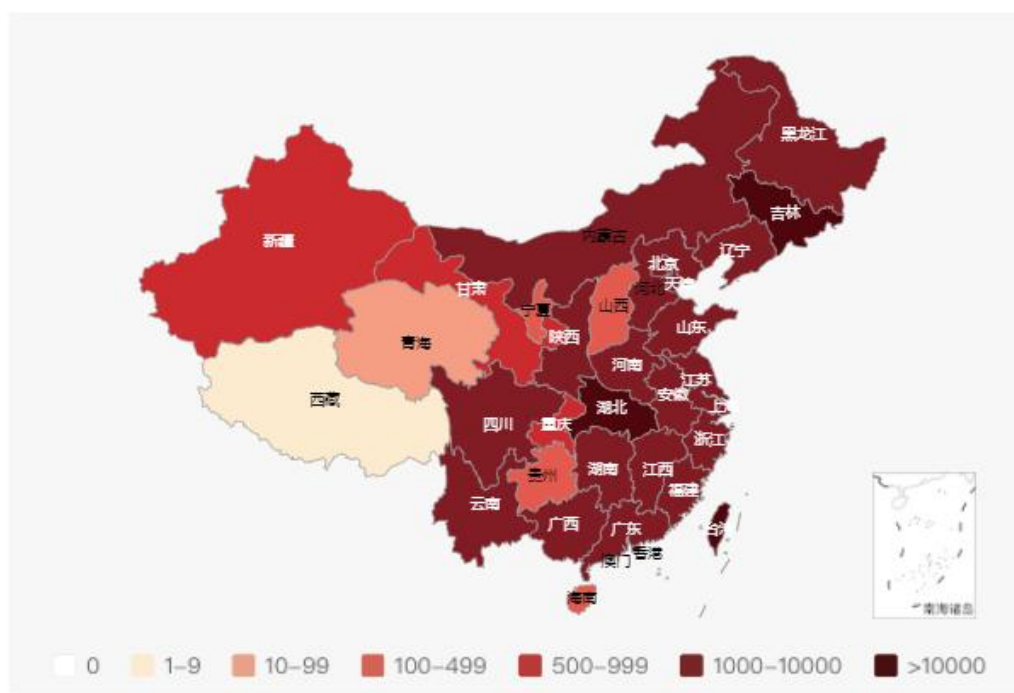


图 1.2 中国新冠肺炎累计确诊分布

随着“奥密克戎”变异株的出现，2022 年我国多地出现无症状感染者。吉林省成为当前疫情的重灾区。病毒至今肆虐全球，严重影响着人民的生活，社会的运转，给我们的生存带来不容忽视的威胁。有关于病毒、防控等依然是当前所有问题的重中之重，应当继续加大力度深入研究。中国政府在此次事件当中积极抗疫、科学防控、迅速采取有效措施，人民群众积极配合、主动隔离、众志成城守卫家园，是全人类抗击疫情的先锋，为世界人民做出了表率。

1.2 文献综述

1.2.1 传染病模型研究综述

从古至今，传染病问题一直伴随着人类的发展史。古代，人类对传染病知识了解太少，医疗水平低下，瘟疫造成了“千村薜荔人遗矢，万户萧疏鬼唱歌”的凄凉景象；而在现代，随着医疗水平的提升，人类对传染病知识的了解越来越多，各种新型传染病也层出不穷，某种程度上的损失依然是我们所避免不了的。历史和现实都在告诫我们：人类长期遭受着传染病的侵害与威胁，解决传染病问题迫在眉睫，是全世界需要引起特殊关注的重大问题。

目前，对传染病的研究方法主要有四种：描述性研究、分析性研究、实验性研究和理论性研究^[3]。传染病动力学是对传染病进行理论性定量研究的一种重要方法。它是根据种群生长的特性，疾病的发生及在种群内部的传播、发展规律，以及与之有关的社会等因素，建立能反映传染病动力学特征的数学模型，通过对模型动力学性态的定性、定量分析和数值模拟，来显示疾病的发展过程，揭示其流行规律，预测其变化发展趋势，分析疾病流行的原因及关键因素，寻求对其预防和控制的最优策略，为人们防治决策提供理论基础和数量依据。

早在 1760 年，法国数学家 Daniel Bernoulli^[4]曾尝试对天花的传播给出定量描述。在他的论文《天花死亡率新分析以及对预防性接种疫苗的优势研究》里，伯努利提出了一个数学模型，将感染者人数随时间的变化曲线视作指数增长模型，并在模型的基础上证明了抑制增长的有效方式是对人群接种疫苗。此项研究非常具有前瞻性，由此之后确定性的传染病模型逐渐产生。1906 年，Harmer^[5]面对一波又一波的麻疹冲击为此构造了一个离散的时间模型。1911 年 Ross 博士^[6]利用微分方程模型对疟疾在蚊子与人群之间的传播动态行为进行了深入的研究并借此获得了诺贝尔医学奖，结果表明当蚊子的数量降低到某一数值一下那么疟疾将不再流行。1927 年苏格兰科学家 William O. Kermack 和 Anderson G. McKendrick^[7]为研究伦敦黑死病提出了著名的 SIR 仓室模型。随后两人又在 1932 年提出了 SIS 仓室模型并且在模型的基础上提出了判断疾病是否流行的“阈值理论”^[8]，为传染病动力学的研究做出了史无前例的伟大贡献。SIR 模型也成为传染病模型中最经典、最基本的模型。随后以此为依托的 SEIR 模型、SIRS 模型、SIQR 模型等陆续出现，传染病数学模型研究得到了飞速地发展。20 世纪 50 年

代,现代计算机的创始人之一冯诺依曼提出了元细胞自动机的概念,Ahmed 等^[9]定义了一类基于元细胞自动机的传染病模型^[10],讨论了传播强度对最终结果的影响及传播结果具体细化的分类。1983 年 Fred Cohen 博士证实了计算机病毒^{[11][12]}的存在,1993 年传染病模型被应用到计算机病毒上且由此建立了第一个完整的计算机病毒传播模型^{[13][14]}。2003 年 SARS 流行期间,石耀霖^[15] 利用越南数据建立了 SARS 病毒传播的系统动力学模型并进行蒙特卡罗实验,结果显示传染率系数是所有因子当中最活跃的一个,关注系数随时间的变化趋势有助于解决问题。2010 年,刘彦麟^[16]将非自治的 SIRS 传染病模型应用到甲型 H1N1 病毒的传播,利用模型分析了系统产生各种性质如稳定性等问题所需要的条件。2015 年,韩新杰等人^[17]为了准确描述埃博拉病毒具体详细的传播运转过程,建立了改进的 SI 模型。当疫情已被控制的情况下,将病人纳入考虑潜伏期,建立了完善的 SEIR 模型,使用迭代方法来获得其非线性的数值解,进而计算治愈人群的数量和药物的需要量。一直到今天,各种传染病模型越来越趋于完善,研究者所考虑的因子更加向实际靠拢,同时伴随着人工智能、机器学习等领域的逐渐兴起,层出不穷的新方法也逐渐被应用到传染病模型的研究当中。

1.2.2 新冠疫情研究综述

2019 年末新冠疫情于武汉迅速暴发,给人民的生产生活带来了无可估量的影响,至今蔓延全球造成数以百万计的死亡,对人类社会的打击是近年来空前巨大的。各国科研工作者对此引起高度重视,对 COVID-19 病毒的传播途径、潜伏期、传播速度、隔离影响、患病周期等各种问题进行了建模研究。无数研究人员基于最经典的 SIR 模型以及 SEIR 模型对疫情的发展进行可视化的分析及预测并由此提出行之有效的防控措施。喻孜等^[18] 采用 SIR 模型同时结合时变参数,针对武汉地区的新冠肺炎疫情进行了评估和预测; Biswas K^{[19][20][21]}等利用经典 SEIR 模型分析和预测疾病的传播趋势,并利用结果来指导预防和控制措施;张艳霞等^[22]采用改进的 SIR 模型方程和龙格库塔法进行模拟计算,进而推导出新冠疫情不蔓延的必要条件,得出最行之有效的防控措施是降低日接触率; Alenezi 等^[23]利用 SIR 模型分析并预测了科威特的新冠疫情,计算出了该国的感染峰值; Lympelopoulou^[24]做出了基于神经网络的经典 SIR 模型由此得到的预测效果良好;于振华等人^[25]在 SLEIR 模型当中加入对低危群体的分析; Jia^[26]则是利用动态扩

展的 eSIR 模型来模拟意大利疫情走势；Cooper^[27]将目光转向了美国得克萨斯州的不同社区；张钰倩等^[28]建立了一类具有复发效应的 SEAIR 模型并给出了模型的基本再生数，研究结果建议采取严格的隔离措施和佩戴口罩来降低疾病的传染率和二次复发；Liu 等人^[29]考虑了无症状感染者，死亡患者和转阴复阳等更多实际问题，提出了融合这些因素的 SEAIRD 模型，其对疫情数据的模拟效果更加贴合现实情况。

无数学者的建模分析使我们对新冠疫情的发展状况有了更明确的认知，我们能够从这些分析当中更加清楚地看到疫情扩散的速度、传播途径、患病周期等各种问题，并且这些结论指导我们对其使用科学的控制措施，逐步战胜病毒，战胜疫情。

1.3 研究内容

本文按照内容共分为以下四章：

第一章是绪论。主要分为两个部分：第一部分介绍了本文的研究背景，描述了传染病的概念、发展历史，对现今新型冠状病毒肺炎疫情的发展情况进行了说明；第二部分是文献综述，介绍了传染病动力学模型发展史上的重要突破以及国内外学者利用传染病模型在新冠疫情上的研究与成果。

第二章是 SIR 传染病模型。主要分为以下三个部分：第一部分介绍几种不同的传染病动力学模型，对模型内容以及相关的动力学基础做简要描述；第二部分介绍本文模型当中所应用到的数学知识如最小二乘法等内容；第三部分展开说明本文重点研究的 SIR 经典模型的概念、方程、性质等并对其重要参数进行估计。

第三章是实证分析。选取具有代表性的湖北省和长春市两地数据，对模型参数进行估计并提出防控建议。

第四章是总结。对本文所做的工作进行总结与反思，提出不足与改进。

第二章 SIR 传染病模型

2.1 传染病模型基础

2.1.1 传染病动力学基本知识

1927 年，苏格兰科学家 Kermack 与 McKendrick 提出了著名的 SIR 仓室（compartment）模型，奠定了传染病动力学的发展，被后世无数学者沿用至今并在其基础上逐渐完善。接下来将对 SIR 仓室模型中涉及到的理论基础与基本概念做出介绍。

SIR 仓室模型，是针对某类传染病将其对应的传播地区人群划分为三个类别，即三个仓室：

易感者（susceptibles），表示 t 时刻没有被感染但有可能被该类疾病传染的人群，记为 $S(t)$ 。

感染者（infectives），表示 t 时刻已经被感染成为病人并出现相应症状具有传染力的人群，记为 $I(t)$ 。

移出者（removed），表示 t 时刻从感染者中移出的人群，这里既包含康复者也包含因该病死亡者，记为 $R(t)$ 。

定义 2.1.1 接触率 假定疾病是通过与他人接触传播的。我们将单位时间内某个感染者与其他人接触的次数定义为**接触率**，这明显极大可能依赖于该区域内的总人口数 N ，故将其记为 $U(N)$ 。

定义 2.1.2 有效接触率 并不是病人的每次接触都会造成疾病传染，当感染者接触人群当中的易感者，并且以 β_0 的概率将疾病传染成功，就将此时的接触率定义为**有效接触率**，记为 $\beta_0 U(N)$ 。

想要达到有效接触需要两个条件：第一感染者需要与易感者接触才会发生传染，当某感染者接触其他感染者或者免疫者等人群时无法发生传染。因而通过有效接触率能够反映病人的生存环境、活动范围、行动能力等问题。第二，经过接触后病毒能否成功地传染给下一个人是有概率的，并不是百分之百。所以有效接触率也能反映病毒的致病能力以及易感者自身的免疫程度。

定义 2.1.3 传染率与发生率 易感者 S 在总人口 N 中所占的比例为 S/N ，则每个感染者对易感者的有效接触率为 $\beta_0 U(N)S/N$ ，即每个感染者平均对易感者的传染

率，简称为**传染率**。那么在 t 时刻单位时间内由易感者转为感染者的人数就是

$\beta_0 U(N) \frac{S(t)}{N(t)} I(t)$ ，称为该病的**发生率**。

传染率当中又存在**双线性发生率**和**标准发生率**：当接触率与区域内总人口成正比时， $U=kN$ ，令 $\beta_0 kN = \beta N$ ，此时该病的发生率就可化简为 $\beta S(t)I(t)$ ，将其定义为**双线性发生率**，也称**简单质量作用律发生率**。但在实际情况中接触率与总人口的关系不一定成正比，我们假定接触率为常数 m ，令有效接触率 $\beta = \beta_0 m$ ，此时

该病的发生率为 $\beta \frac{S(t)}{N(t)} I(t)$ ，将其定义为**标准发生率**。

我们可以看到这两种发生率的主要差别在于有效接触率的定量方式，而在历史研究中还采用过其他种类如介于二者之间的饱和接触率等。在解决实际问题时还需要根据具体的情况进行选择和分析以期效果最优化。

2.1.2 几种传染病动力学模型

自 Kermack-McKendrick 的 SIR 仓室模型建立以来，各种传染病模型层出不穷，如 SI 模型、SIS 模型、SIRS 模型、SEIR 模型等。对于分析研究各种传染病问题起到了不可忽视的作用，帮助我们更好地制定防治措施。不考虑出生与自然死亡等种群动力学因素，最基本的传染病动力学模型大致有以下几类：

● 无疾病潜伏期

我们将 t 时刻易感者的数量记为 $S(t)$ ，感染者的数量记为 $I(t)$ ，移出者的数量记为 $R(t)$ ，分别简记为 S 、 I 、 R 。 t 时刻单位时间内，一个病人能够传染的人数为 β ，移出者在病人中所占的比例为 γ 。

(1) SI 模型

对人群只考虑了易感者和感染者，并且感染者不能够被治愈。此类病症有 HIV 等。该模型的传播机制如下：

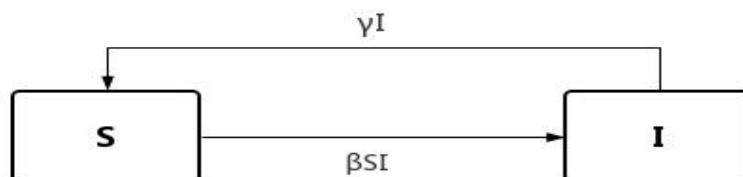


该模型对应的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI \end{cases}$$

(2) SIS 模型

对人群依然只考虑易感者和感染者两类人群。但感染者能够被治愈，并且可能重新被感染，进而再次循环。此类病症有季节性流感、脑炎等。该模型的传播机制如下：

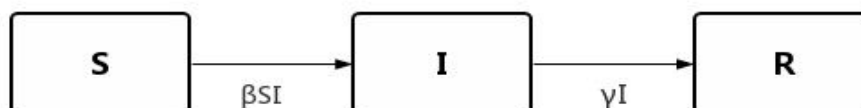


该模型对应的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \gamma I \\ \frac{dI}{dt} = \beta SI - \gamma I \end{cases}$$

(3) SIR 模型

对人群考虑易感者、感染者和移出者三类人群。感染者能够被治愈，并且产生抗体获得终生免疫不再被感染。此类病症有麻疹、腮腺炎、风疹等。该模型的传播机制如下：

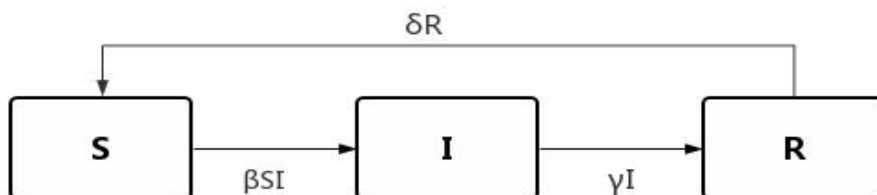


该模型对应的微分方程组如下：

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

(4) SIRS 模型

对人群考虑易感者、感染者和移出者三类人群。感染者能够被治愈并获得短暂免疫力，但有可能被再次感染进行新一轮传播。它与 SIS 模型的区别在于，SIS 模型当中患者被治愈后没有免疫期，可以立即被二次感染；而 SIRS 模型有暂时的免疫期，患者先康复变为移出者后再以比例 δ 丧失免疫力被感染。该模型的传播机制如下：



该模型对应的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \delta R \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I - \delta R \end{cases}$$

● 有疾病潜伏期

在患病之前有一段病菌潜伏期，假定在潜伏期内感染者不具备传染力。我们将 t 时刻潜伏期的人数记为 $E(t)$ ，简记为 E 。疾病的平均潜伏期记为 $\frac{1}{\omega}$ 。

(1) SEIR 模型

对人群考虑易感者、潜伏者、感染者和移出者四类人群。易感者被感染后没有马上发病而是进入潜伏期，再以 ω 的概率成为真正的感染者，被治愈后获得终

生免疫。该模型的传播机制如下：

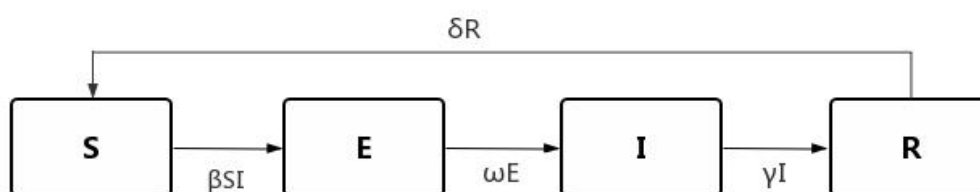


该模型对应的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dE}{dt} = \beta SI - \omega E \\ \frac{dI}{dt} = \omega E - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

(2) SEIRS 模型

对人群考虑易感者、潜伏者、感染者和移出者四类人群。易感者被感染后先进入潜伏期，再以 ω 的概率成为真正的感染者，被治愈后仅有暂时免疫力，过一段时间以 δ 的概率重新被感染。该模型的传播机制如下：



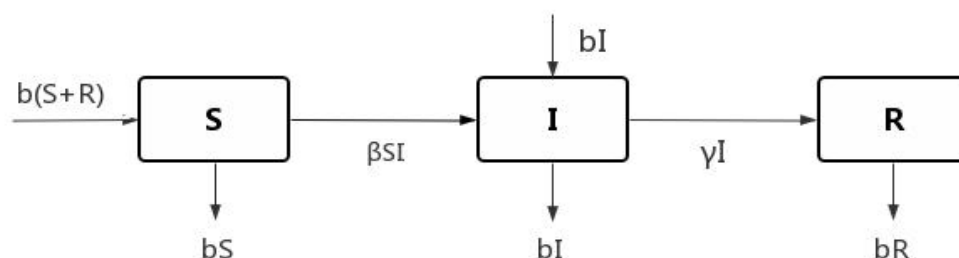
该模型对应的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \delta R \\ \frac{dE}{dt} = \beta SI - \omega E \\ \frac{dI}{dt} = \omega E - \gamma I \\ \frac{dR}{dt} = \gamma I - \delta R \end{cases}$$

以上早期模型均未考虑种群动力学因素。近期的模型会更加贴合实际，有些考虑时滞因素、年龄结构、隔离影响、变动人口等；有些增加了模型的维数，考虑疾病在多个群体中的传播和交叉感染；有些结合具体的疾病特征展开更为深入的研究。模型由于实际情况的改变呈现出更为多变且具有针对性的形式。所以我们在本文中对此类模型仅以一种形式做简要举例说明。

● 总人口恒定且有垂直传染的 SIR 模型

假定出生率系数与自然死亡率系数相等，记为 b ，即环境内总人口是一个常数，记为 K 。母亲的疾病会直接传染给新生儿并且康复者的新生儿也不具备免疫力。该模型的传播机制如下：



2.2 数学知识

● 最小二乘估计

设 Y 为因变量， X 为自变量，满足线性关系：

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{p-1} X_{p-1} + e,$$

其中 e 为误差项，表示除了 X 之外其他因素对 Y 的影响； $\beta_0, \beta_1, \dots, \beta_{p-1}$ 为待估参数。且有 Y 和 X 的 n 组观测值

$$(x_{i1}, \dots, x_{i,p-1}, y_i), \quad i = 1, \dots, n$$

满足

$$y_i = \beta_0 + x_{i1}\beta_1 + \dots + x_{i,p-1}\beta_{p-1} + e_i, \quad i = 1, \dots, n \quad (2.1)$$

其中误差项 e_i 满足：

- (1) $E(e_i) = 0$,
- (2) $Var(e_i) = \sigma^2$,
- (3) $Cov(e_i, e_j) = 0, \quad i \neq j$.

即 Gauss-Markov 假设。且式 (2.1) 的矩阵形式为

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1,p-1} \\ 1 & x_{21} & \cdots & x_{2,p-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{n,p-1} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{p-1} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix},$$

即

$$y = X\beta + e,$$

其中, y 为 $n \times 1$ 的变量观测向量, X 为 $n \times p$ 的已知设计矩阵, β 为 $p \times 1$ 的未知参数向量, e 为随机误差向量。进而上述假设的矩阵形式可写为

$$E(e) = 0, Cov(e) = \sigma^2 I_n.$$

这样我们就得到线性模型:

$$y = X\beta + e, E(e) = 0, Cov(e) = \sigma^2 I_n. \quad (2.2)$$

对于模型 (2.2) 中的未知参数 β 采用最小二乘法进行估计:

令

$$Q(\beta) = \|y - X\beta\|^2 = (y - X\beta)'(y - X\beta), \quad (2.3)$$

要使上式 (2.3) 取得最小值, 对 β 求偏导得

$$\frac{\partial Q}{\partial \beta} = 2X'(X\beta - y), \quad (2.4)$$

令上式 (2.4) 等于零, 当矩阵 $X'X$ 正定时, 方程组有唯一解

$$\hat{\beta} = (X'X)^{-1}X'y$$

即参数 β 的最小二乘估计。

2.3 SIR 仓室模型

2.3.1 SIR 经典模型

SIR 仓室模型基于以下三个假设:

(1) 不考虑人口出生与自然死亡、迁入迁出等种群动力学因素。将疾病传播区域视作一个封闭环境,环境内的总人口 $N(t)$ 始终保持不变为一个常数 K ,即,

$$N(t) \equiv S(t) + I(t) + R(t) \equiv K. \quad (2.5)$$

(2) t 时刻单位时间内一个感染者能够传染的易感者数量与易感者总数 $S(t)$ 成正比,比例系数为 β 。此时一个感染者所传染的易感者数量为 $\beta S(t)$,那么所有感染者所传染的易感者数量为 $\beta S(t)I(t)$,人群从易感者向感染者转移,即感染者数量新增 $\beta S(t)I(t)$,同时易感者数量减少 $\beta S(t)I(t)$ 。

(3) t 时刻单位时间内从感染者移出的数量与感染者总数 $I(t)$ 成正比,比例系数为 γ 。此时移出人数为 $\gamma I(t)$,人群从感染者向移出者转移,即移出者数量新增 $\gamma I(t)$,同时感染者数量减少 $\gamma I(t)$ 。

定义 2.3.1 移出率(系数)与恢复率(系数) γ 是单位时间内移出者在感染者中的比例,定义为**移出率(系数)**。而移出者这个群体不仅包含了康复者还包含因病死亡者、隔离者等,理想情况下,所有移出者均为康复者,此时 γ 称为**恢复率(系数)**。

基于以上三个假设,我们将 $S(t)$ 、 $I(t)$ 、 $R(t)$ 分别简记为 S 、 I 、 R , SIR 仓室模型所体现的传播机制如下:

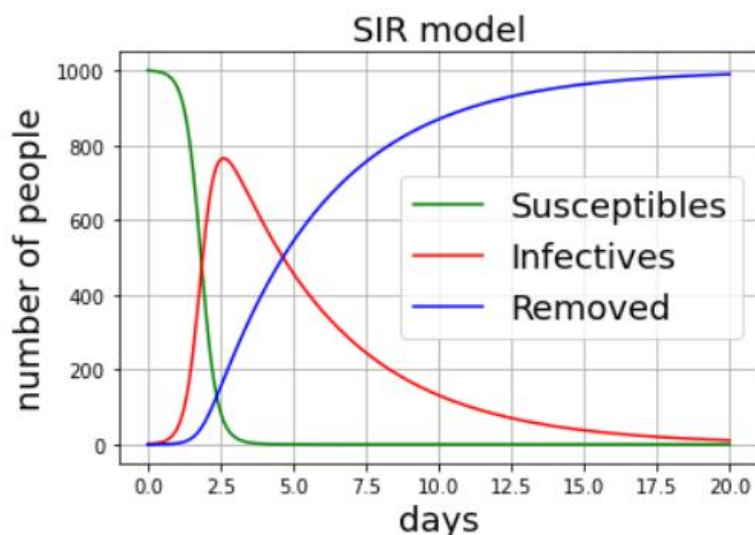


该模型对应的微分方程组如下:

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases} \quad (2.6)$$

式中 S 、 I 、 R 三者的初始值分别为 $S(0)$, $I(0)$, $R(0)$, 分别代表疾病开始时的易感者、感染者、移出者的数量,均为非负整数。 β 为有效接触率, γ 为移出率系数,均为非负数。

假定某地区疾病初始时 $S(0)=1000$, $I(0)=1$, $R(0)=0$, $\beta=0.004$, $\gamma=0.25$, 利用 python 画出 S , I , R 三者随时间变化的图像,代码见附录 C,



图（2.1） SIR 经典模型 S,I,R 三者的变化趋势

结合方程组（2.6）与图像（2.1），易得

（1）对易感者 $S(t)$: $S'(t) < 0$, $S(t)$ 单调递减且有下界，图像位于 x 轴上方并渐近于 x 轴，因此 $S(t)$ 存在极限，且 $\lim_{t \rightarrow \infty} S(t) = S_{\infty}$ ；

（2）对移出者 $R(t)$: $R'(t) > 0$, $R(t)$ 单调递增且有上界，图像位于 x 轴上方并逐渐趋于平稳，因此 $R(t)$ 存在极限，且 $\lim_{t \rightarrow \infty} R(t) = R_{\infty}$ ；

（3）对感染者 $I(t)$: 由式（2.5）， $S(t)$ 、 $R(t)$ 的极限都存在，因此 $I(t)$ 的极限也存在，且 $\lim_{t \rightarrow \infty} I(t) = I_{\infty} = 0$ ；

图像位于 x 轴上方，先单调递增后单调递减，令 $I'(t) = 0$ ，即 $\beta SI - \gamma I = 0$ ，解得，

$$S(t) = \frac{\gamma}{\beta},$$

$I(t)$ 取得极大值同时也是最大值，也就是说当 $S(t) = \frac{\gamma}{\beta}$ 时就是病毒传染的最高峰。

（4）为了更直观地体现 S 与 I 的关系，我们将方程组（2.6）中前两个方程作比，得到

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta S}, \quad \text{令 } \rho = \frac{\gamma}{\beta}$$

即，

$$\frac{dI}{dS} = -1 + \frac{\rho}{S}$$

解得,

$$I(S) = I(0) + S(0) - S + \rho \ln \frac{S}{S(0)}$$

当 $S(0) > \rho$ 时, 随着时间的推移, 感染者 I 的数量随着易感者 S 的数量减少先呈现增大趋势即疾病流行, 取得最大值 $I(\rho)$ 后逐渐减小; 当 $S(0) = \rho$ 时, 感染者 I 的数量起初就是最大值后逐渐减少; 当 $S(0) < \rho$ 时, 感染者 I 的数量直接递减最终趋于 0。也就是说当 $S(0) \leq \rho$ 时, 疾病不会流行。而这就是著名的阈值定理当中所阐述的内容。

定理 2.3.1 阈值定理 设 $S(t)$, $I(t)$ 是初值问题的解, 如果 $\frac{S(0)}{\rho} < 1$, 那么, 当 $t \rightarrow \infty$ 时, $I(t)$ 单调减少趋于零。如果 $\frac{S(0)}{\rho} > 1$, 当 $t \rightarrow \infty$ 时, $I(t)$ 先增加达到最大值

$1 - \rho - \rho \ln \frac{S(0)}{\rho}$, 此时 $S = \rho$, 而后单调减少趋于零, $S(t)$ 是一个单调减少函数,

并且其极限 $\lim_{t \rightarrow \infty} S(t) = S_{\infty}$ 是方程 $1 - S + \rho \ln \frac{S}{S(0)} = 0$ 在 $(0, \rho)$ 内的根。

定义 $\rho = \frac{\gamma}{\beta}$ 为阈值, 若 $S(0) - \rho = \delta$ ($\delta > 0$), 当 δ 与 ρ 相比是一个小量且 $I(0)$

与 $S(0)$ 相比是一个小量, 传染病被消灭时 $S(t)$ 恰好比阈值 ρ 低 δ , 即这场流行病中, 人群中共有 2δ 的人被染上过病。

由于阈值定理体现的是 $S(0)$ 与 ρ 的动态关系, 我们令

$$R_0 = \frac{S(0)}{\rho} = \frac{\beta S(0)}{\gamma}, \quad (2.7)$$

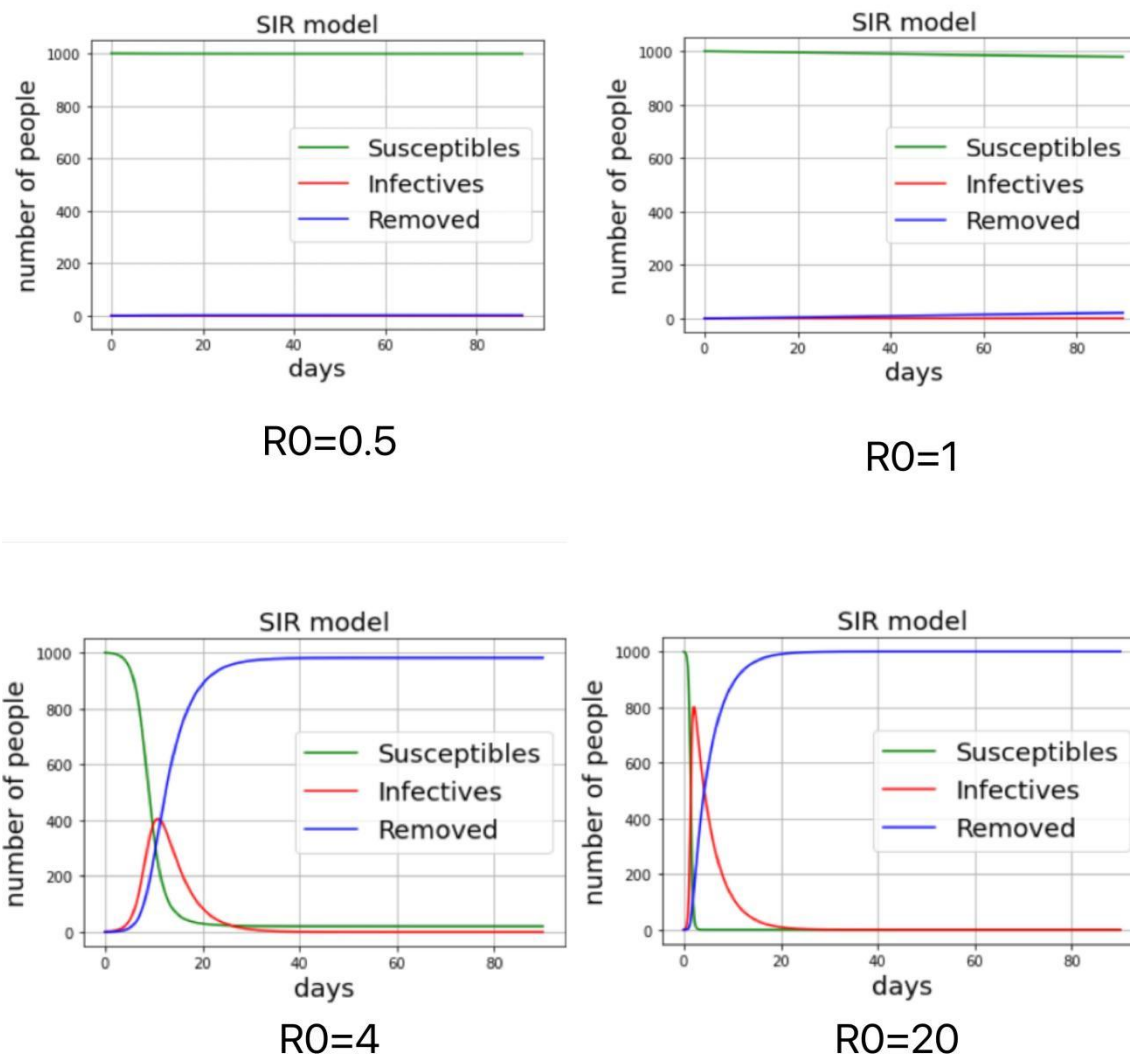
式中 $\frac{1}{\gamma}$ 表示平均移出时间, 即平均患病期。

当 $R_0 > 1$ 时, 疾病始终保持而形成地方病; 当 $R_0 < 1$ 时, 疾病逐渐消亡。因此 $R_0 = 1$ 是区分疾病是否流行的阈值, 并有,

定义 2.3.1 基本再生数 在发病初期而人均均为易感者时, 一个病人在平均患病期

内所传染的人数，称为**基本再生数**。记为 R_0 ，表达式为 $R_0 = \frac{S(0)}{\rho} = \frac{\beta S(0)}{\gamma}$ 。

也就是说， $R_0 > 1$ 所对应的实际含义是一个病人在平均患病期所传染的人数超过一个，进而导致疾病的流行；反之 $R_0 < 1$ 所对应的实际含义是一个病人在平均患病期所传染的人数少于一个，进而疾病逐渐衰退。通过图像能够更清楚地看到 R_0 的作用，以下是分别给基本再生数 R_0 取不同值时 S、I、R 三者变化趋势的图像：



图（2.2）SIR 经典模型 $R_0 = 0.5, 1, 4, 20$ 对应变化趋势

那么我们要想控制疾病的流行，就要尽可能地减小基本再生数 R_0 ，通过 R_0 的表达式（2.7）可以看出：

- （1）减小 $S(0)$ 。 $S(0)$ 代表的是疾病初始时刻的易感人者数。一个人是否易

感取决于个体的免疫力,要想增强免疫力可以给人群注射疫苗使易感者免疫直接成为移出者,进而达到易感者人数减少;

(2) 减小 β 。 β 代表有效接触率。有效接触率一方面在于感染者是否与易感者进行接触,对此可以将感染者进行隔离来减少接触人数与次数;另一方面在于病毒自身的传染能力,可以对环境进行消杀灭菌以减少病毒的存活几率、阻断病毒可能传播的各种途径;

(3) 增大 γ 。 γ 代表移出率系数,理想情况下称为恢复率。提高医疗水平,对病毒进行更为深入的研究,对病人加强治疗才是所有问题的核心关键。

与基本再生数相对应的一个概念是有效再生数,

定义 2.3.2 有效再生数 在疾病传播过程中而人均均为易感者时,一个病人在平均患病期内所传染的人数,称为有效再生数。记为 R_t , 表达式为 $R_t = \frac{S(t)}{\rho} = \frac{\beta S(t)}{\gamma}$ 。

也就是说 t 时刻内一个患者所传染的易感者人数就是有效再生数。相当于动态的基本再生数,利用它我们能更方便地随时了解疾病发展的情况。

2.3.2 重要参数估计

由上文的叙述,我们可以明确地看到感染率系数 β 和移出率系数 γ 是整个模型当中最重要的两个参数,只有对这两个参数进行相对较为准确的估计才会使模型更加完整并且发挥真正的作用,来帮助我们更好地研究问题解决问题。

在本文当中我们采用最小二乘法对感染率系数 β 和移出率系数 γ 进行估计。最小二乘法,又被称为最小平方法,是一种数学优化的方法。它通过判断误差平方和取得最小值的位置来确定参量具体的取值。利用最小二乘法可以便利地求得模型预估数据,并令模型数据与实际数据之间误差的平方和为最小。普通最小二乘估计量具有线性特性、无偏性和最小方差性。

在本文的 SIR 经典仓室模型当中,我们假设待估参数为 $\theta = (\beta, \gamma)$, 实际样本数据为 $Y = (y_1, y_2, \dots, y_n)^T$, 模型拟合所得数据为 $\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T$, 则每个个体的残差为

$$e_i = y_i - \hat{y}_i \text{ 即 } e = Y - \hat{Y},$$

那么残差平方和为

$$SSE(\theta) = \|e\|^2 = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

对 β , γ 分别求偏导得,

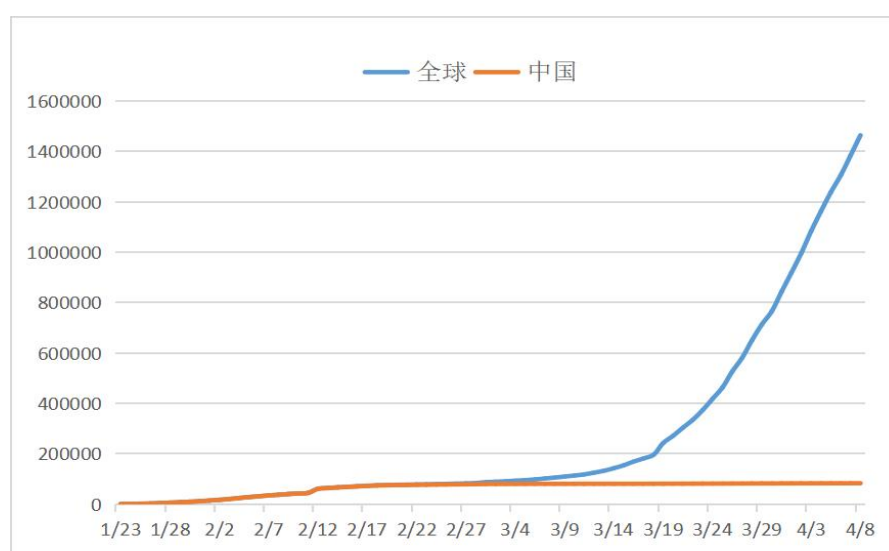
$$\begin{aligned}\frac{\partial SSE(\theta)}{\partial \beta} &= -2 \sum_{i=1}^n (y_i - \hat{y}_i) \frac{\partial \hat{y}_i}{\partial \beta}, \\ \frac{\partial SSE(\theta)}{\partial \gamma} &= -2 \sum_{i=1}^n (y_i - \hat{y}_i) \frac{\partial \hat{y}_i}{\partial \gamma}.\end{aligned}$$

令两个偏导分别等于零, 则能得到对 β , γ 的最小二乘估计量。

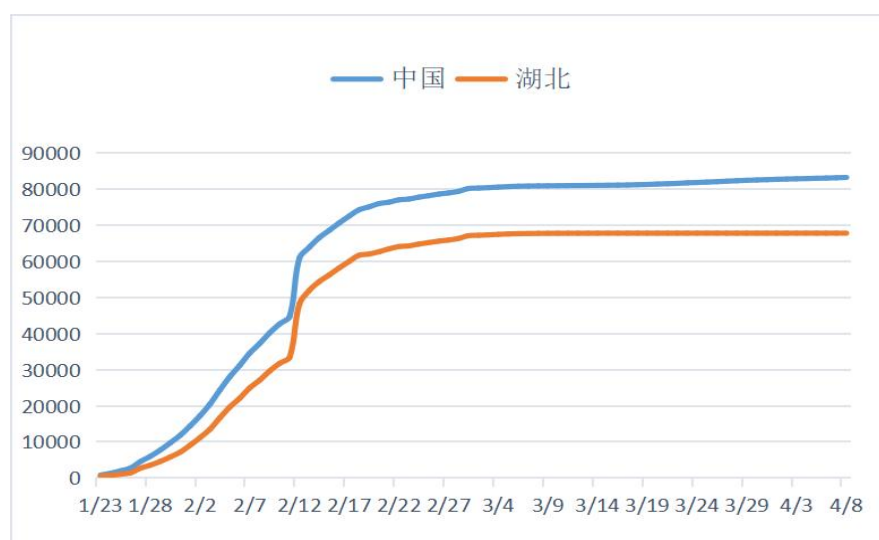
第三章 实证分析

3.1 数据选取

2019 年末新型冠状病毒肺炎在中国湖北武汉爆发，北京、上海、广东等多地陆续产生确诊病例，病毒迅速蔓延至全国范围，34 个省份无一不被卷入此次疫情当中。武汉作为本次国内疫情首发地同时也是重灾区，在 2020 年 1 月 23 日正式实施封城措施，经过各方人员的顽强抗争，疫情逐渐获得基本稳定，确诊人数不再上升，于 2020 年 4 月 8 日解除封禁，国内疫情防控取得了阶段性的胜利。与此同时国外各地疫情出现爆发形态。如图为世界与中国累计确诊病例对比走势、中国与湖北省累计确诊病例走势：



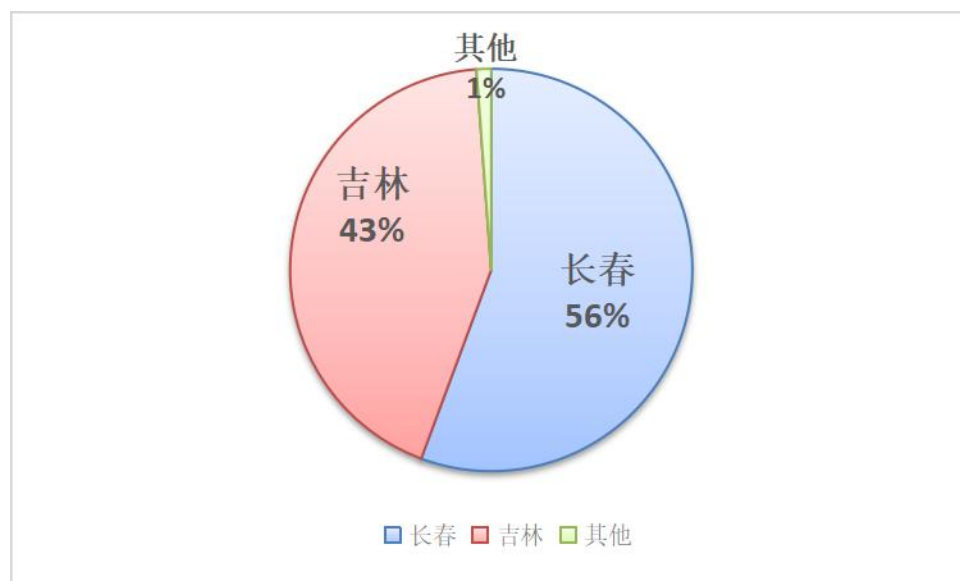
图（3.1）世界与中国累计确诊病例走势



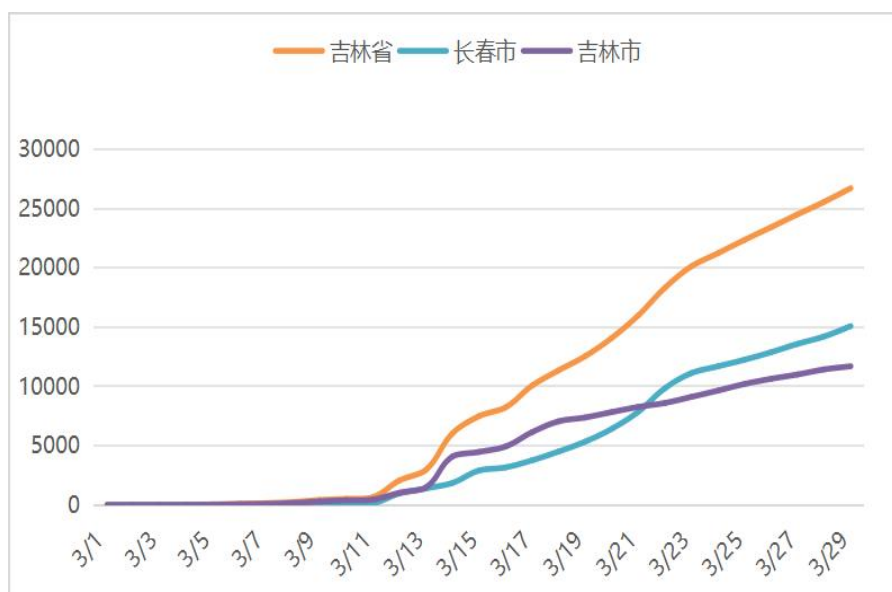
图（3.2）中国与湖北累计确诊病例走势

由图（3.1）（3.2）可知，2020 年 3 月以前疫情集中在中国爆发，世界其他国家并未受到显著影响，而中国疫情则集中在湖北。截至 2020 年 4 月 8 日，中国湖北省累计确诊人数达 67803 例，累计死亡人数达 3215 例，全国累计确诊人数达 83249 例，累计死亡人数达 3344 例。湖北省累计确诊与累计死亡人数分别占全国的 81.4%和 96.1%，在全国范围内受到的影响最为巨大，因此针对疫情发展第一阶段本文选择湖北省疫情数据用于实证分析，数据来源于丁香园见附录 A，选择最具代表性的 2020 年 1 月 23 日至 2020 年 4 月 8 日这一时段。

2020 年第一阶段结束后，疫情相继又在北京、大连、青岛等多地爆发，新冠病毒一直没有停止对人类的侵害，2022 年疫情已经发展到了第三阶段。2022 年 3 月 1 日吉林省延边朝鲜族自治州新增确诊病例 4 例，先后扩散至吉林市、长春市、四平市等多地。仅仅一个月的时间，吉林省累计本土确诊病例达 26710 例，累计无症状感染者达 13568 例。全国累计确诊病例仅 28163 例，显然，吉林省确诊数据占全国确诊数据约 94%以上，成为新一轮疫情的重灾区。而在吉林省内，截至 2022 年 3 月 29 日，长春市累计确诊 15081 例；吉林市累计确诊 11697 例；四平市累计确诊 188 例；延边地区累计确诊 117 例.....各地情况如下所示：



图（3.3）吉林省内累计本土确诊人数分布占比



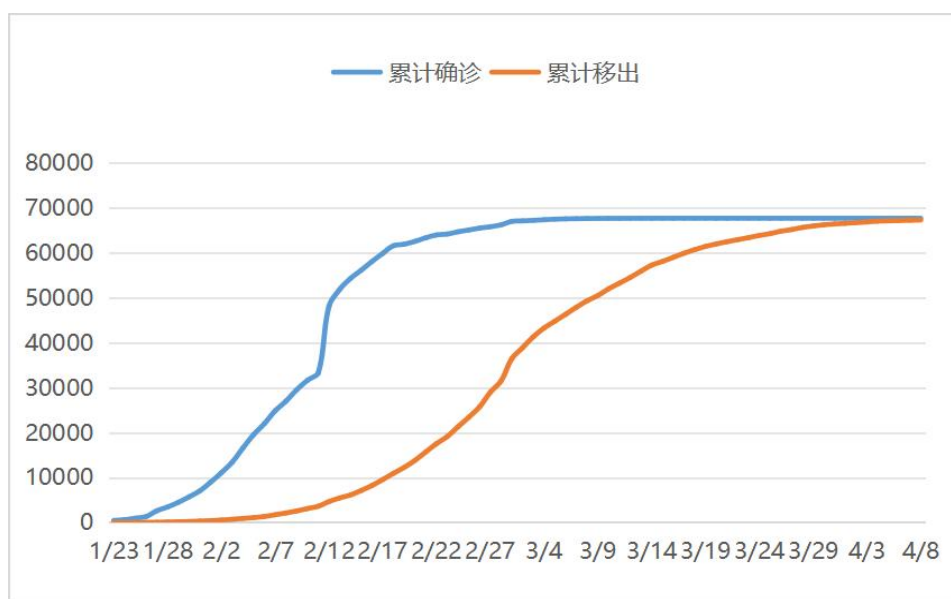
图（3.4）吉林省、长春市、吉林市累计本土确诊人数走势

由图（3.3）（3.4）可知，吉林省内疫情主要集中在长春市与吉林市两地。因此针对疫情发展第三阶段本文选择吉林省长春市疫情数据用于实证分析，数据来源于吉林省卫生健康委员会官方网站^[30]见附录 B，选择 2022 年 3 月 1 日至 2022 年 4 月 3 日这一时段。

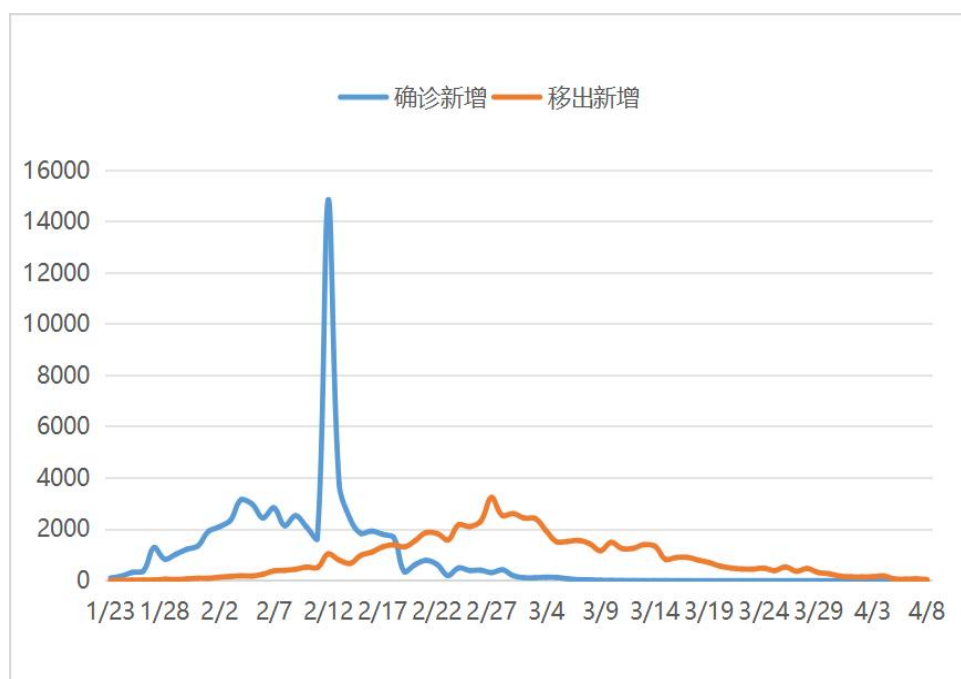
3.2 数据分析

3.2.1 湖北疫情数据分析

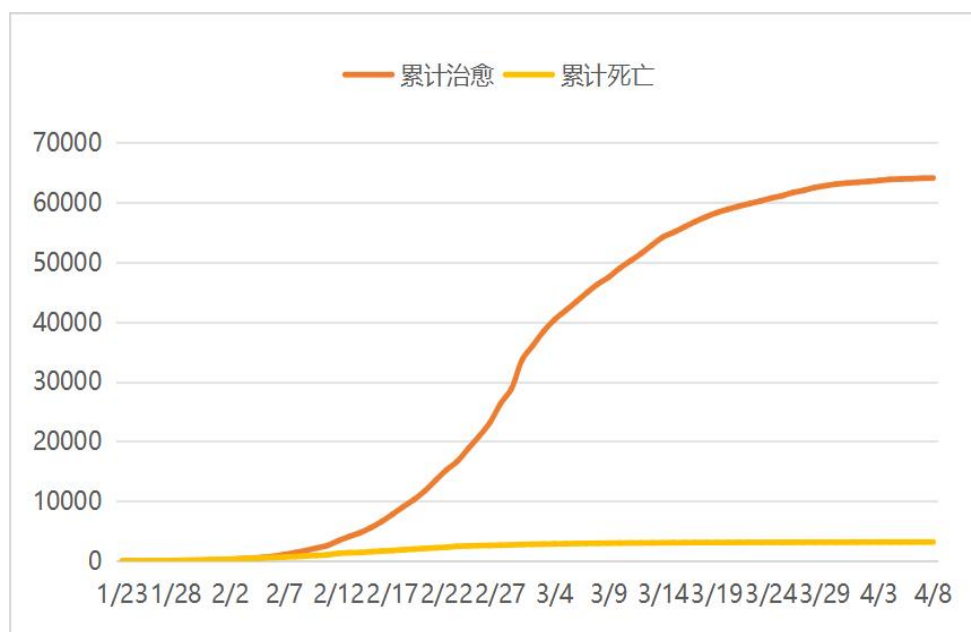
由丁香园获取的原始数据显示：截至 2020 年 4 月 8 日，湖北省累计确诊病例人数共 67803 例，累计死亡病例人数共 3215 例，累计治愈病例人数共 64187 例。做出各项指标随时间变化的图像，如下图所示：



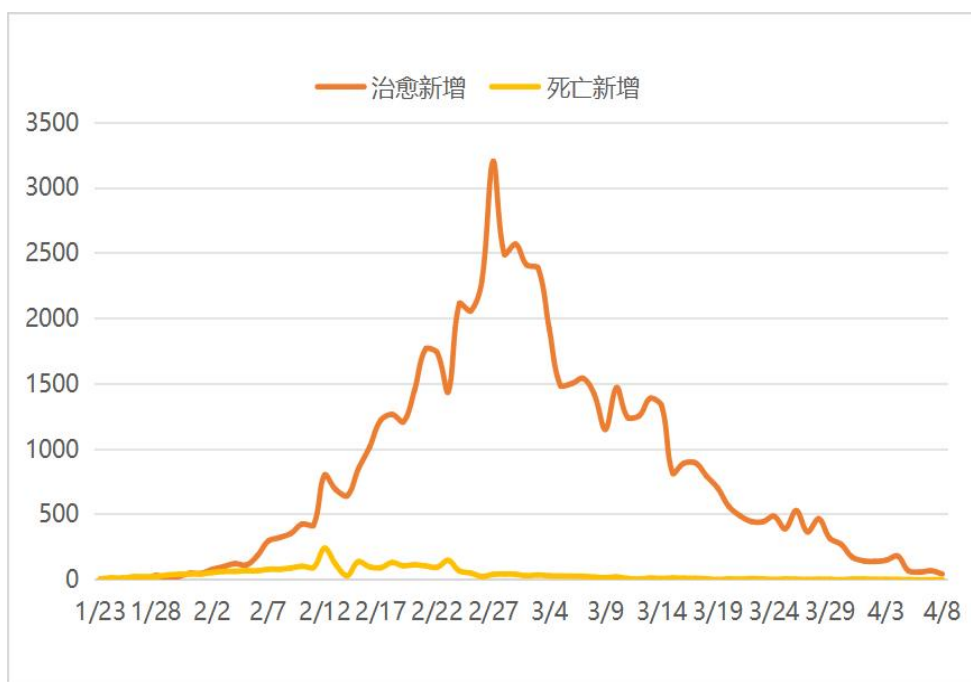
图（3.5）湖北省累计确诊与累计移出人数走势



图（3.6）湖北省确诊新增与移出新增人数走势



图（3.7）湖北省累计治愈与累计死亡人数走势

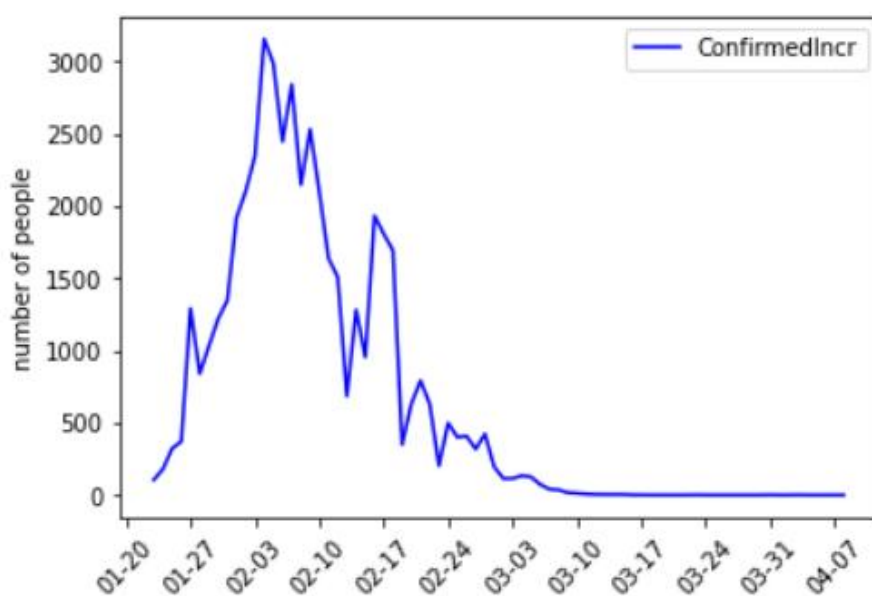


图（3.8）湖北省治愈新增与死亡新增人数走势

由数据及图像可知，疫情发展阶段，累计确诊病例、累计移出病例人数均呈现上升趋势。1月至2月，确诊病例迅猛增加并一度呈现指数爆发式增长，由2月下旬开始增长速度逐渐减缓直至3月份疫情初步得到基本控制，单日确诊增量降低至两位数及以下，4月份已无新增确诊。

结合图（3.5）（3.6），由2020年1月23日至2020年2月2日，确诊病例以

平均每日 1000 以内的速度增加，移出速度较为缓慢基本低于 100 每日。2 月 2 日起确诊病例单日增量突破 2000，确认病例迅速增加，而移出速度依然没有显著提升。2 月 12 日，确诊病例由 33366 激增至 48206，单日增量达到 14840，而由图（3.6）显示 12 日数据异常跳跃，根据湖北卫健委官网^[31]，国家卫生健康委办公厅、国家中医药管理局办公室印发的《新型冠状病毒感染的肺炎诊疗方案(试行第五版)》在湖北省的病例诊断分类中增加了“临床诊断”，以便患者能及早按照确诊病例接受规范治疗，进一步提高救治成功率。根据该方案，湖北省对既往的疑似病例开展了排查并对诊断结果进行了订正，对新就诊患者按照新的诊断分类进行诊断。而临床诊断的判别依据是疑似病例只要 CT 影像有新冠肺炎肺部病变的特征，在没有核酸检测的情况下，就成为临床诊断病例。当日湖北省新增临床诊断病例就有 13332 例，故而产生数据异常点。这对于湖北的疫情防控无疑是更好的举措，但为了便于研究问题，我们对 2 月 12 日附近的确诊病例数据恢复至没有“临床诊断”这一项。重新绘制确诊增量随时间变化的图像如下：



图（3.9）修正后的确诊人数增量走势

经修正，确诊病例增量走势大体上依然是先增大后减小，但峰值减小到 3500 以下。从 2 月 19 日开始，新增病例显著减少，基本达到 500 以下，防控与治疗措施发挥了应有的作用。一直到 3 月 10 日新增病例被控制在个位数并减少为 0，疫情开始走向好转。

结合图（3.7）与图（3.8），治愈病例实现从无到有并且稳步增加，2 月末治

愈增量出现峰值；死亡病例始终以及其缓慢的速度增加并且要低于治愈速度，一直到4月8日，治愈病例新增45例，死亡病例新增2例。我们对于重症病例以及病毒研究的重视都起到巨大的作用，这场抗疫之战取得了阶段性的胜利。

3.2.2 长春疫情数据分析

由吉林省卫生健康委员会官方网站获取的原始数据显示：2022年3月4日，长春市出现首例本土阳性确诊患者，不到一个月的时间病毒迅速蔓延，统计每日累计确诊与新增确诊病例人数走势如下图：

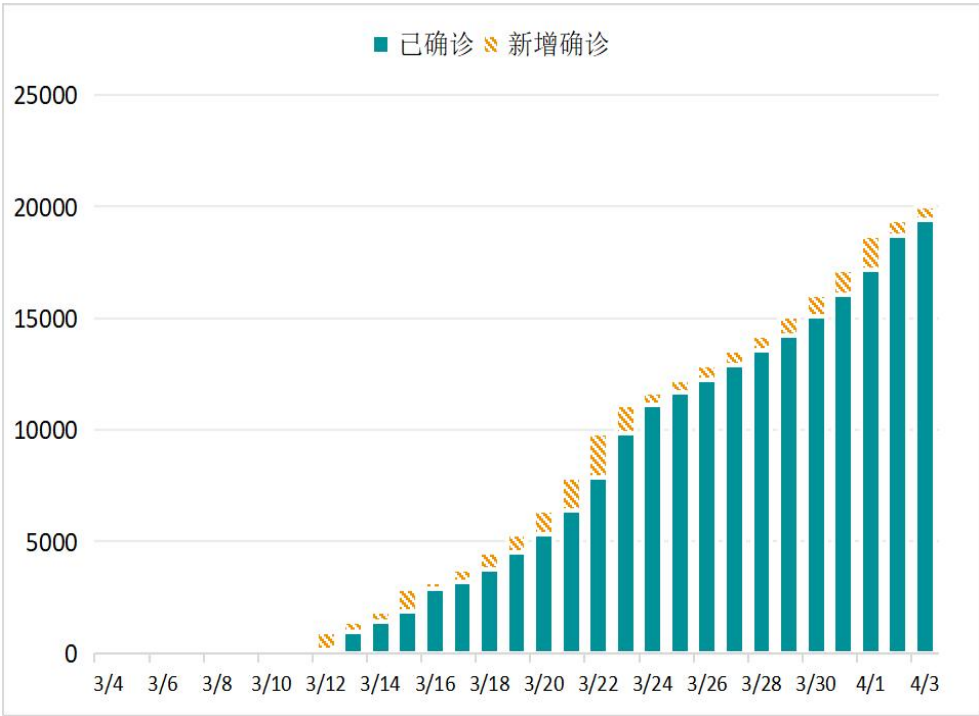
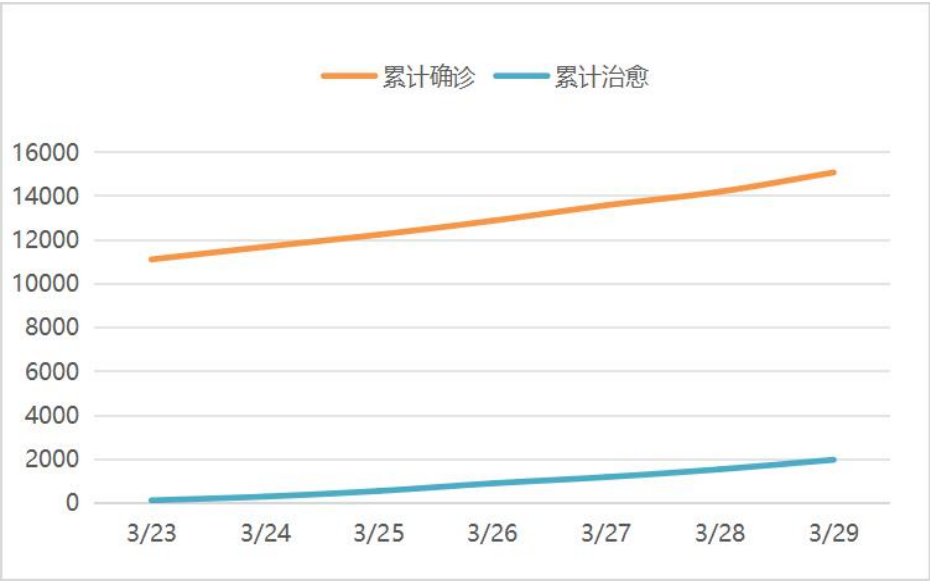


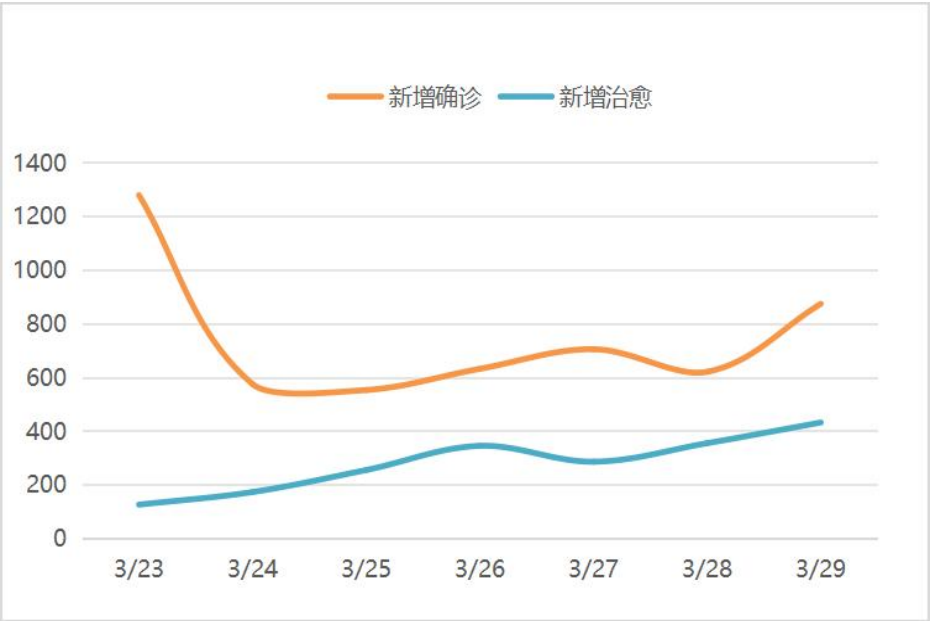
图 (3.10) 每日确诊构成

3月4日至3月11日，确诊人数以每日100例以下的速度增长，较为缓慢，处在可控范围内；由于3月9日首轮“全民核酸”开启，3月12日新增确诊病例人数有了明显的突跃，达831例，此后以日均400~500例的速度继续增长；3月15日新增确诊病例人数突破1000例；自3月16日起增长速度明显逐渐增大，3月19日累计确诊病例人数突破5000例；3月22日新增确诊病例人数达1979例，是这一段时间以来的峰值；由于医疗工作者的不懈努力以及政府管控措施的一再加强，3月23日开始前期被方舱医院收治的患者中核酸阴性患者开始陆续出院，每日新增确诊患者数量显著降低至1000例以下。

3月23日开始，官方数据当中开始出现“治愈病例新增”一项。做出一周内确诊与治愈病例相关指标对比如下：



图（3.11） 累计确诊与累计治愈人数走势



图（3.12） 新增确诊与新增治愈人数走势

由图（3.11）（3.12）可知，3月23日后一周内传染速度与治愈速度基本一致，传染速度略有震动且稍高于治愈速度。

3.3 参数估计

3.3.1 针对湖北疫情的 SIR 模型参数估计

由前文第二章式 (2.6)，SIR 经典模型的微分方程组如下：

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases},$$

式中 $S(t)$ ， $I(t)$ ， $R(t)$ 的值均可由湖北省实际统计数据得到，初值 $I(0) = 494$ ， $R(0) = 55$ 。而初始易感者人数 $S(0)$ ，传染率系数 β 以及移出率系数 γ 是我们需要估计的参数。

由原始数据我们重新定义：

$I(t)$ = 累计确诊人数 - 累计治愈人数 - 累计死亡人数；

$R(t)$ = 累计治愈人数 + 累计死亡人数；

$S(t)$ = 总人口数量 - 感染者数量 - 移出者数量。

我们利用 python 对未知参数进行估计。通过 3.2 的分析我们能够观察到新冠肺炎患者的平均患病期 $\frac{1}{\gamma}$ 大概在 15 天左右，而这与现有研究结论也是相符的，

因此我们首先给定移出率系数 $\gamma = \frac{1}{15} \approx 6.67 \times 10^{-2}$ ，由于我们选择的时间节点在封城管控之后，参考文献^[32]给出了初值 $S(0) = 110810$ 百人，我们对传染率系数 β 进行最小二乘估计。将 $I(0) = 494$ ， $R(0) = 55$ ， $\gamma = 6.67 \times 10^{-2}$ ， $S(0) = 110810$ 带入模型，得到参数 β 的待估区间为 $[3.7 \times 10^{-4}, 8.0 \times 10^{-7}]$ ，对该区间进行遍历，遍历次数为 300，得到当 $\beta = 3.52 \times 10^{-6}$ ，对应的误差平方和最小，此时 $SSE = 1.882 \times 10^8$ 。

这样，参数的估计我们就完成了。将 $\beta = 3.52 \times 10^{-6}$ ， $\gamma = 6.67 \times 10^{-2}$ 带回 SIR 经典模型模拟感染者 $I(t)$ 的变化趋势，得到下图：

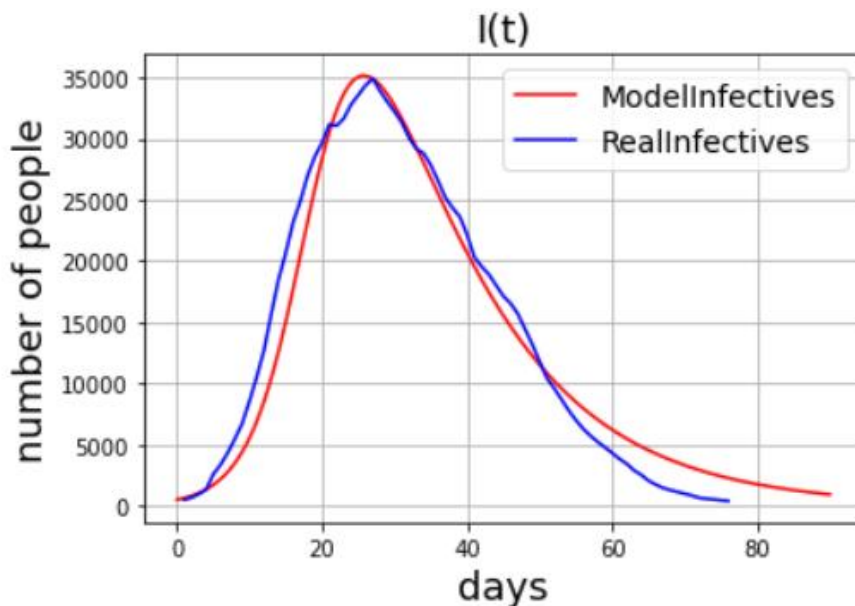


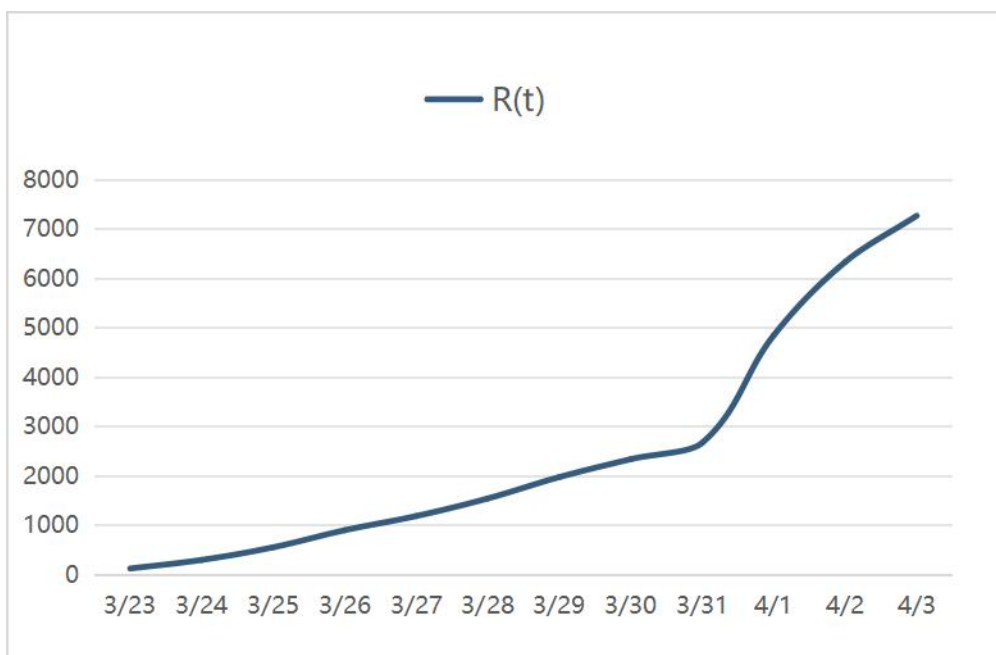
图 (3.13) 湖北省感染者人数 $I(t)$ 模型估计值 (红) 与实际值 (蓝) 对比

由图 (3.13) 可知, 初期 20 天实际值高于估计值, 低估了疫情发展的速度, 模拟效果不佳; 快达到峰值的一周左右, 实际值略低于估计值, 由峰值及其后十天左右的实际值与估计值基本吻合, 这一阶段的模拟效果是比较好的; 后期 38 天左右之后实际值与估计值忽上忽下, 略有差距, 模拟效果一般。

3.3.2 针对长春疫情的 SIR 模型参数估计

疫情由 2020 年发展至今, 新冠病毒已经由最初的原始毒株逐渐进化出德尔塔变异株、奥密克戎变异株等多种形态。与此同时我们的防控措施也从无到有, 小到勤洗手、佩戴口罩, 大到居家隔离、静态管控。各方面因素都在随着时间发展产生改变。因此对于本阶段疫情不能再使用湖北省的相关参数, 针对长春市近期发生的病毒传播情况我们重新调整参数。

首先, 对于初始易感者人数 $S(0)$, 我们依然是参照文献^[32]的方式给出初值, $S(0) = 90669$ 百人。其次, 从 2022 年 3 月 23 日开始有患者陆续出院, 并且病毒目前的致死率极低, 全市死亡人数仅为个位数, 因此我们忽略死亡人数, 将移出者 $R(t)$ 全部看做治愈患者的数量, 做出 $R(t)$ 随时间变化的图像:

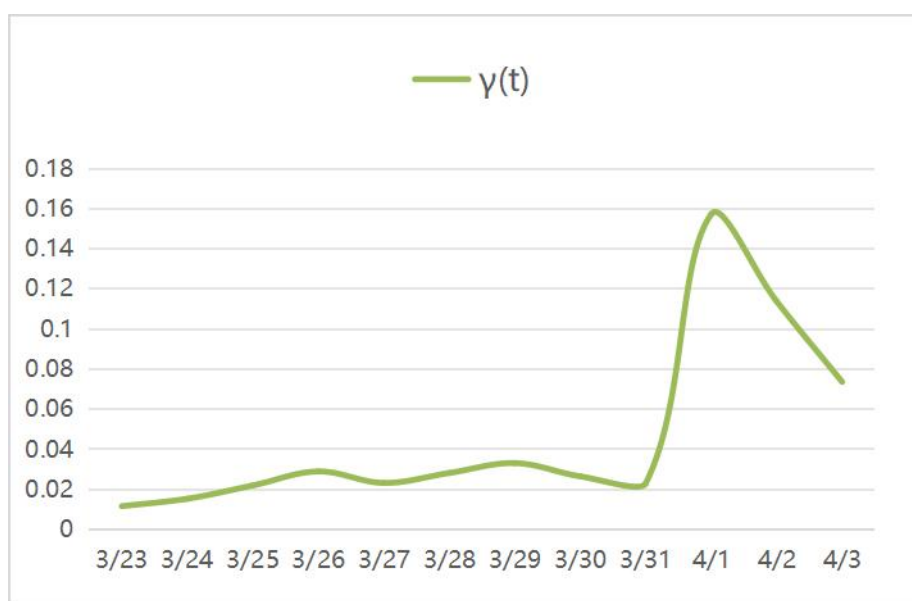


图（3.14） $R(t)$ 随时间变化图像

由 SIR 模型式（2.6）中最后一个方程，参考文献中^[33]的方法，我们可以近似地定义日移出率即移出率系数 γ 随时间变化的函数

$$\gamma(t) = \frac{\Delta R}{I(t)} = \frac{R(t) - R(t-1)}{I(t)}, \quad (3.1)$$

做出 $\gamma(t)$ 的变化如图所示：

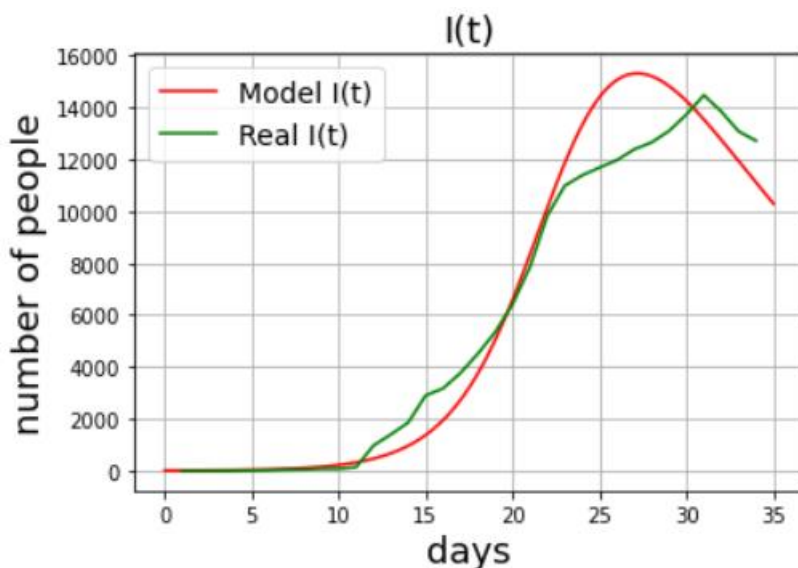


图（3.15） $\gamma(t)$ 随时间变化图像

取日移出率 $\gamma(t)$ 的平均值作为 γ 的估计值，则 $\hat{\gamma} = 4.63 \times 10^{-2}$ 。

接下来对传染率系数 β 进行最小二乘估计，将 $S(0)=90669$, $I(0)=5$, $R(0)=0$, $\gamma = 4.63 \times 10^{-2}$ 带入 SIR 经典模型，得到参数 β 的待估区间为 $[2.7 \times 10^{-3}, 8.0 \times 10^{-6}]$ ，对该区间进行遍历，遍历次数为 300，得到当 $\beta = 5.93 \times 10^{-4}$ ，对应的误差平方和最小，此时 $SSE = 2.37 \times 10^7$ 。

最终将所有的参数值 $S(0)=90669$, $I(0)=5$, $R(0)=0$, $\gamma = 4.63 \times 10^{-2}$, $\beta = 5.93 \times 10^{-4}$ 代入模型，得到感染者 $I(t)$ 随时间的变化的趋势如下：



图（3.16） 长春市感染者人数 $I(t)$ 模型估计值（红）与实际值（绿）对比

由图可知，模型在病毒爆发初期 10 日以内拟合效果较好，此时采取防控措施最为稳妥；11-18 日，实际值高于拟合值，这是由于长春市从这一时段开始密集实行“全民核酸”政策造成了实际数据的突跃；19-22 日，模型基本完全拟合实际；23 日实际值出现了较为明显的下降，原始数据显示患者从这一天开始出院也就是移出者这一项才开始出现，而对应这一时段的模型值要高于实际值，并提前 3 天左右出现极值点，面对当前的疫情严重程度按照模型的走势采取防控措施是有利而无害的，但综合考虑可能会对社会经济等方面造成损失。因此在疾病

发展中后期不适合继续利用原模型以及相应参数再对情况进行预测分析。而疾病发展前期得到的效果是比较符合实际情况的。

3.4 模型应用与分析

基本再生数 R_0 是衡量疾病发展程度的重要指标，它代表一个病人在平均患病期能够传染的人数最大值。由 3.3 实证分析可以得到湖北省基本再生数：

$$R_{01} = \frac{\beta S(0)}{\gamma} = \frac{3.52 \times 10^{-6} \times 110810}{6.67 \times 10^{-2}} \approx 5.84$$

显然 $R_{01} > 1$ ，单个患者最大可传染 5 个易感者，疾病大爆发。

而长春市的基本再生数

$$R_{02} = \frac{5.93 \times 10^{-4} \times 90669}{4.63 \times 10^{-2}} \approx 11.62$$

$R_{02} > 1$ ，单个患者最大可传染 11 个易感者，而且 $R_{02} > R_{01}$ ，病毒的传染力显著增强，防疫措施必须加大力度，才能有效遏制疾病的传播。

根据现有的研究结果^[34]，新冠病毒原始毒株的基本再生数 R_0 值在 2~5 之间，德尔塔变异株的 R_0 值在 7 左右，而奥密克戎变异株的 R_0 值在 10 左右。将本文的研究结果与之进行分时段对照：

基本再生数	原始毒株	奥密克戎毒株
资料	2~5	10
本文	5.84	11.62

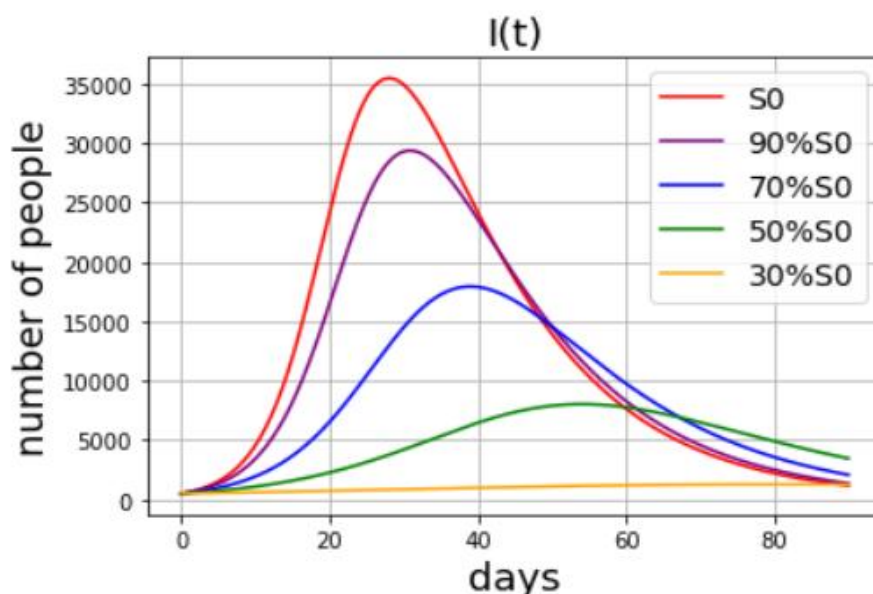
本文结果与资料结果基本相似，但两个结果均偏大，其根本原因在于本文选择的原始样本区域内部已经形成了社区传播状态，超级传播者在小范围内聚集，在传染前期对于本区域内部反而是更大的风险；文献当中的数据则是更加针对于病毒本身的具有平均意义的基本再生数。

由 2.3.1 的分析，控制基本再生数 R_0 是分别从初始易感者人数 $S(0)$ ，传染率系数 β ，移出率系数 γ 三个角度入手的：

(1) 注射疫苗。注射新冠疫苗作为国家大力提倡的防控手段，其作用在于提升人群免疫力，进而达到减少初始易感者人数 $S(0)$ 的目的。我们假设疫苗接种后易感者获得长期免疫成为移出者，在现有情况的基础上，控制 β 与 γ 的值不变，改变疫苗的接种率，同时与之对应的 $S(0)$ 如下表所示：

疫苗接种率	0%	10%	30%	50%	70%
$S(0)$	$S(0)$	$90\%S(0)$	$70\%S(0)$	$50\%S(0)$	$30\%S(0)$

感染者人数 $I(t)$ 的变化趋势如图所示：

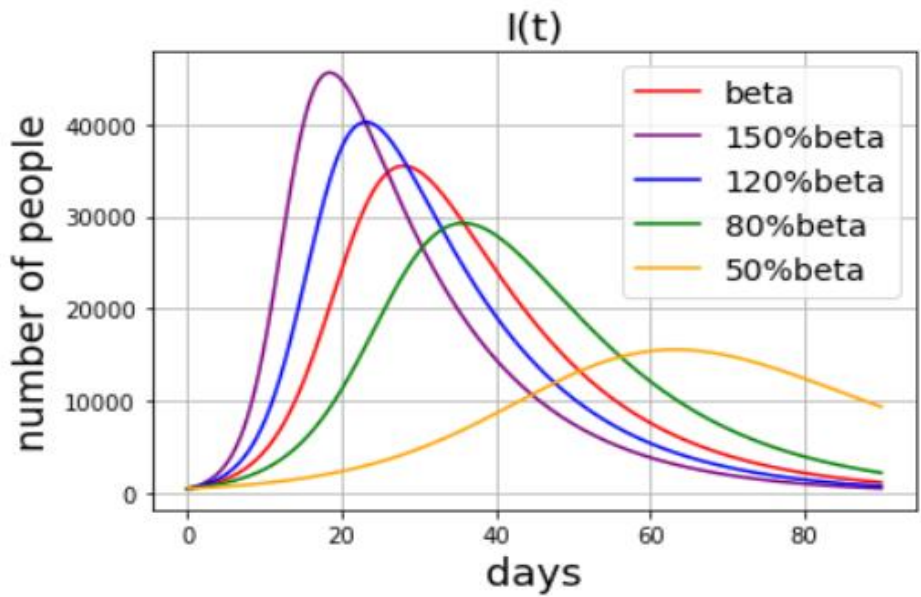


图（3.17） 改变 $S(0)$ 感染者人数走势对比

由上至下，疫苗接种率逐渐提高，感染者峰值逐渐减小，并且峰值对应的时间逐渐后移。也就是说疫苗接种率越高，病毒传播的速度越慢，范围越小，甚至达到 70% 以上时，不会造成指数型大规模爆发。这一点可以由 2022 年的疫情现状得到验证，由国家卫生健康委员会数据显示我国当前疫苗接种率达 87%，因此长春市的疫情并没有达到超级传播状态。因此，注射疫苗，提高免疫力应作为全民最重要的预防手段。同时，提升疫苗的有效率，响应政策及时注射疫苗加强针也是防疫过程当中不容忽视的一环。

(2) 隔离人群。隔离人群是最直接的降低有效接触率的方式，将有效接触尽可能地降到最低也就意味着病毒缺乏传播对象，从而降低传染率系数 β 。我们假设隔离强度与传染率系数成反比，在现有情况的基础上隔离强度分别降低与升

高 20%和 50%，即对应 $150\%\beta, 120\%\beta, \beta, 80\%\beta, 50\%\beta$ 时，控制 $S(0)$ 与 γ 的值不变，感染者人数 $I(t)$ 的变化趋势如图所示：



图（3.18） 改变 β 感染者人数走势对比

由上至下，隔离强度逐渐增大，传染率系数 β 逐渐减小，感染者峰值逐渐减小，整体周期拉长。但形成大规模传播的时间还是要远远小于患者移出时间，因此实施隔离措施最好一定是在病毒传播的早期，这就要求民众具有自觉意识，早发现，早上报，避免形成大规模传播。

结合长春市本次疫情，为了市民满足基本的生存需求以及响应政府“全民核酸”的号召，实际的隔离强度并没有达到理想状态。再加上目前奥密克戎变异株虽然致死率大大降低，但基本再生数可以达到原始毒株的三倍以上，同时平均代际间隔缩小至原始毒株的 $1/3$ ，平均潜伏期变为 $1/2$ ，传染力度大大增强。因此隔离强度必须加大。不必要的人员流动一定要避免。首先，市民所必须的生活物资应当由相关部门指定人员统一采买发放，过程严格把关，对一切流经的人员、场所都进行严格的消毒杀菌，并且按照一定的时间周期如 10 天、两周等进行无接触式配送，限制配送的人数与次数，严格控制由物资运输产生的病毒传播链形成。其次，“全民核酸”政策能够有效地统计患病者数量并且防患于未然。但其中的隐患在于短期内的“全民核酸”一定是要求速度的，这会导致某一小范围内的人员聚集，反而增加了交叉感染的风险。因此，市民出门做核酸检测时需要提高自

觉意识,人与人之间主动间隔一米以上,尽量避免人多的时候去核酸点进行检测,最好是相关人员执行“错峰检测”的形式,合理规定市民做核酸的时间段与人流流量。除此之外,在统计情况基本明确后,对某一大区域中的各个小区域可以采取不同的管控措施。对于相当一段时间内没有阳性患者的低风险区域,隔离期间减少核酸检测的频率;而对于出现阳性患者的区域实施更为严格的管控,禁止除防疫人员以外的一切流动,对阳性患者最邻近的周边区域加大排查,并且增强消杀力度。当然,无接触式核酸检测也是一种很好的手段。在保证医疗物资充裕的前提下,对各家各户发放核酸测试剂盒,市民进行自测,这样能够良好的避免交叉感染的风险。但同时对市民朋友的自觉性要求较高。

(3) 消杀灭菌。消杀灭菌能够对外部环境造成影响,阻断病毒一切可能的传播途径与媒介,降低病毒的生存能力。本质上依然是降低传染率系数 β 。

(4) 提高医疗水平。对病人加强治疗,增大移出率系数 γ 同时可理解为降低平均患病周期。因此要早发现早治疗,在发病初期及时服用药物遏制病情,缩短患病时间。我们依然观察当移出率系数取不同值时对应的感染者人数 $I(t)$ 的变化趋势:

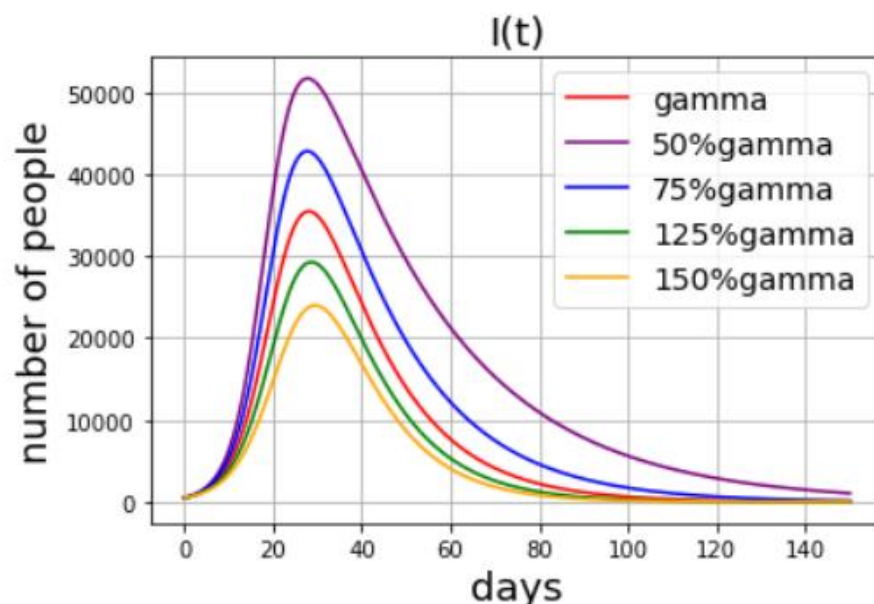


图 (3.19) 改变 γ 感染者人数走势对比

由上至下,移出率系数 γ 逐渐增大,平均患病周期的差距有约两个月之久。感染者峰值随之较为均匀地下降,整体规模显著缩小。因此,大力发展医疗,提升医疗研究水平、提高医护人员素质、备足各种医疗资源等都是现阶段亟待解决

的问题。

第四章 总结与反思

本文基于传染病学经典的 SIR 仓室模型,对新型冠状病毒肺炎疫情展开了建模研究,分别选择疫情发展第一阶段的湖北省数据以及第三阶段的长春市数据作为实证分析。研究了新型冠状病毒肺炎疫情在不同的发展阶段所对应的不同情况。对模型当中重要的参数作出估计,并对防控措施提出了分析与改进。

首先,第一章与第二章对当前的疫情发展状况以及各项研究进行简要的介绍,并对此提出结合 SIR 传染病模型进行更为细致的研究。接下来详细地说明了 SIR 经典模型的原理、重要概念如传染率系数、基本再生数等问题,并且对参数估计的方法——最小二乘法进行简要介绍。为后面的实证分析打好铺垫。

第三章进入实证分析。针对新型冠状病毒肺炎疫情发展的不同阶段分别选取了具有代表性的湖北省与长春市两地数据进行研究,数据分别来源于丁香园与吉林省卫生健康委员会官方网站。首先分别对两地数据进行预处理,对原始数据中各项指标进行基本的可视化分析,找到数据的异常点,如湖北省数据当中 2020 年 2 月 12 日的确诊突增是由于判断指标的改变,对此将前后的数据进行调试。处理过后再将各项指标与模型当中的 $S(t), I(t), R(t)$ 三个量进行对应以求模型拟合的准确性。接下来对重要的参量传染率系数 β 、移出率系数 γ 等进行最小二乘估计,先基本确定待估参数区间,再对每个区间进行数次遍历以找到误差平方和 SSE 取得最小值的位置。将估计参数带回原始模型并与实际情况进行对比。综合来看,湖北省的模型得到了更好的拟合效果,模拟值与实际值基本一致;而对于长春市疫情的模拟仅在前期得到的效果较好,后期参考意义不大。原因在于随着疫情的发展,人们采取的防控手段逐渐起到了良好的效果。通过两地参数的估计我们计算出衡量疾病发展状况的重要指标基本再生数 R_0 ,该值显著地反馈了病毒发展到今天传播能力已经远远高于初期,对此做出分析并提出具有实际意义的防控措施建议。

本文的不足之处有以下几个方面:

SIR 经典模型当中只考虑了易感者、感染者和移出者三种人群,而实际情况当中新型冠状病毒肺炎是具有潜伏期的。随着病毒一步一步发展到今天,潜伏期的数据越来越难以观察,在患者就诊之前难以界定病毒存在了多久,这会造成较大

的个体差异，因此本文未引入潜伏者这一人群类别；此外，长春市目前的发展状况仅能获取到有限数据，这是硕士阶段目前为止的研究成果。本文仅对确诊病例做了模型的拟合，没有考虑无症状感染者，将来如能获取更多数据将更加关注二者之间的转化以及提出更具现实意义防控措施。

参考文献

- [1] R Sharp, A J Lotka. A problem in age distribution. *Philos. Mag.*, 1911, 21: 435~438.
- [2] G F Webb. *Theory of Nonlinear Age-dependent Population*. Addison-Wesley, Reading, 1968.
- [3] 马知恩, 周义仓等. 传染病动力学的数学建模与研究. 第一版. 北京: 科学出版社, 2004.
- [4] Daniel Bernoulli. Essai d' une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l' inoculation pour al prévenir, in *Mémoires de Mathématiques et de physique*. Paris : Académie Royale des Sciences, 1760, 1-45.
- [5] Hamer, W. H., Epidemic disease in England-the evidence of variability and of persistency of type. *Lancet*, 1906, 1: 733-739.
- [6] Ross. R., *The Prevention of Malaria*[M]. 2nd ed. John Murray, London, 1911.
- [7] Kermack, W.O., and McKendrick, A.G., A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 1927, 115: 700-721.
- [8] Kermack, W.O., and McKendrick, A.G., A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 1932, 138: 55-83.
- [9] Ahmed E, Agiza HN, Hassan SZ. On modeling hepatitis B transmission using cellular automata[J]. *Journal of Statistical Physics*, 1998, 92, Nos 3/4.
- [10] Ahmed E, Agiza HN. On modeling epidemics Including latency, incubation and variable susceptibility[J]. *Physica A*, May 1, 1998, 253(1-4): 347-352.
- [11] 韦而泊·施拉姆. 大众传播媒介与社会发展[M], 华夏出版社, 1990; 85-87.
- [12] 尼葛洛庞帝. 数字化生存[M]. 海口: 海南出版社, 1997: 26-29.
- [13] 金耀基. 大学之理念[M]. 北京: 生活 读书 新知三联书店, 2001: 3-5.
- [14] 汪凯. 转型中国: 媒体, 民意与公共政策[M]. 上海: 复旦大学出版社, 2005: 172, 181.
- [15] 石耀霖. SARS 传染扩散的动力学随机模型. *科学通报*, 2003, 48: 1373-1377.
- [16] 刘彦麟. H1N1 流行病传播动力模型研究[D]. 天津大学, 2010.
- [17] 韩新杰, 李姗姗. 博拉病毒传播问题的数学模型研究[J]. 2015.
- [18] 喻 孜, 张贵清. 基于时变参数-SIR 模型的 COVID-19 疫情评估和预测 [J]. *电子科技大学学报*, 2020, 49(3): 357-361.
- [19] Biswas K, Khaleque A. Covid-19 spread: Reproduction of data and prediction using a SIR model on Euclidean network[J]. *arXiv preprint arXiv:2003.07063*,

2020.

- [20] Lei S, Jiang F et al. Clinical characteristics and outcomes of patients undergoing surgeries during the incubation period of COVID-19 infection[J]. E Clinical Medicine, 2020, 21: 100331.
- [21] Achterberg M A, Prasse B, et al. Comparing the accuracy of several network-based COVID-19 prediction algorithms[J]. International Journal of Forecasting, 2020.
- [22] 张艳霞, 李进. 基于 SIR 模型的新冠肺炎疫情传播预测分析, doi: 10.3969/j.issn.1671-7872.2020.01.015
- [23] ALENEZI M N, ALANZI F S. Building a sensible SIR estimation model for COVID-19 outbreak in Kuwait [J]. Alexandria Engineering Journal, 2021, 60(3): 3161-3175.
- [24] LYMPEROPOULOS I N. Stay home to contain Covid-19: Neuro-SIR – Neurodynamical epidemic modeling of infection patterns in social networks [J]. Expert Systems With Applications, 2021, 165: 113970.
- [25] 于振华, 黄山阁. SLEIR 新冠肺炎传播动力学模型构建与预测[J/OL]. 西安交通大学学报. <https://kns.cnki.net/kcms/detail/61.1069.T.20220104.1208.002.html>
- [26] JIA Wangping, HAN Ke. Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China [J]. Front. Med. 2020, 7:169-186.
- [27] COOPER L, MONDAL A, G C. A SIR model assumption for the spread of COVID-19 in different communities [J]. Chaos, Solitons and Fractals, 2020, doi: <https://doi.org/10.1016/j.chaos.2020.110057>.
- [28] 张钰倩, 张太雷. 具有复发效应的 SEAIR 模型及在新冠肺炎传染病中的应用 [J]. 山东大学学报(理学版), 2022, 57(1): 56-68.
- [29] Liu X X, Fong S. Towards a Realistic Model for Simulating Spread of Infectious COVID-19 Disease[C]//Proceedings of the 2020 the 4th International Conference on Big Data and Internet of Things. 2020: 96-101.
- [30] http://wsjkw.jl.gov.cn/xwzx/xwzx/202204/t20220401_8425363.html
- [31] http://wjw.hubei.gov.cn/bmdt/ztl/fkxxgzbdgrfyyq/xxfb/202002/t20200213_2025580.shtml
- [32] 陈兴志, 田宝单. 基于 SEIR 模型的 COVID-19 疫情防控效果评估和预测 [J]. 应用数学和力学, 2021, 42(2): 199-211.
- [33] 陈创. 基于 SIR 模型在新冠肺炎中的应用与研究[D]. 兰州大学, 2021.

[34]<https://new.qq.com/omn/20220515/20220515A08GMH00.html>

致 谢

值此论文完成之际，我的校园生活也将画上句号。两年的时间不长不短，但 是在我的人生当中却留下了浓墨重彩的一笔。在身体与精神相对更加趋于完整的 年纪，还能留在校园这片净土，充实大脑，丰富内心，耐心雕琢自己，这段时光 将永远成为我宝贵的财富。在此要对一路陪伴我成长成熟的各位老师同学家长说 一声衷心的感谢！

首先，感谢我的导师尹哲老师。对科学，老师坚持钻研，秉持着一丝不苟、 孜孜不倦的态度，对重要问题绝不放过，以自身的行动督促我、激励我，在我感 到迷茫和困惑的时候为我指明方向，阐明问题的根源为我解答疑惑；对生活，老 师积极乐观，对生活当中出现的大事小情保持着自己的热忱，独立思考独立探索 的精神一直是我的榜样。

通过这两年老师的教诲，培养了我终身学习的意识、严于律己的品格、奋发 向上的态度和独立解决问题的能力。无论是在哪个方面都使我成为更好的自己， 能够成为尹哲老师的学生是我的幸运。

其次，感谢我在延边大学遇见的所有老师同学。老师们不辞辛苦的传授给我 各项知识与技能，同学们陪伴着我度过求学的日子，你们是我生活当中的有机组 成成分，是你们的存在让我的研究生生涯更加丰富多彩！

最后，感谢我的父母家人。从咿呀学语到如今长大成人，二十多年来的悉心 培育背后是你们浓浓的爱。希望我也能够逐渐成长为你们可以依靠的人，不再为 我操心费神。希望我能成为你们幸福快乐的源泉。

谢谢你们！我爱的人！

附录 A 湖北省 COVID-19 疫情数据

日期	确诊 累计	确诊 新增	治愈 累计	治愈 新增	死亡 累计	死亡 新增
1/23	549	105	31	3	24	7
1/24	729	180	32	1	39	15
1/25	1052	323	42	10	52	13
1/26	1423	371	44	2	76	24
1/27	2714	1291	47	3	100	24
1/28	3554	840	80	33	125	25
1/29	4586	1032	90	10	162	37
1/30	5806	1220	116	26	204	42
1/31	7153	1347	166	50	249	45
2/1	9074	1921	215	49	294	45
2/2	11177	2103	295	80	350	56
2/3	13522	2345	396	101	414	64
2/4	16678	3156	520	124	479	65
2/5	19665	2987	633	113	549	70
2/6	22112	2447	817	184	618	69
2/7	24953	2841	1115	298	699	81
2/8	27100	2147	1439	324	780	81
2/9	29631	2531	1795	356	871	91
2/10	31728	2097	2222	427	974	103
2/11	33366	1638	2639	417	1068	94
2/12	48206	14840	3441	802	1310	242
2/13	51986	3780	4131	690	1426	116
2/14	54406	2420	4774	643	1457	31
2/15	56249	1843	5623	849	1596	139
2/16	58182	1933	6639	1016	1696	100
2/17	59989	1807	7862	1223	1789	93
2/18	61682	1693	9128	1266	1921	132
2/19	62031	349	10337	1209	2029	108
2/20	62662	631	11788	1451	2144	115
2/21	63454	792	13557	1769	2250	106
2/22	64084	630	15299	1742	2346	96
2/23	64287	203	16738	1439	2495	149
2/24	64786	499	18854	2116	2563	68
2/25	65187	401	20912	2058	2615	52
2/26	65596	409	23200	2288	2641	26
2/27	65914	318	26403	3203	2682	41
2/28	66337	423	28895	2492	2727	45

3/1	67103	196	33757	2570	2803	42
3/2	67217	114	36167	2410	2834	31
3/3	67332	115	38556	2389	2871	37
3/4	67466	134	40479	1923	2902	31
3/5	67592	126	41966	1487	2931	29
3/6	67666	74	43468	1502	2959	28
3/7	67707	41	45011	1543	2986	27
3/8	67743	36	46433	1422	3007	21
3/9	67760	17	47585	1152	3024	17
3/10	67773	13	49056	1471	3046	22
3/11	67781	8	50298	1242	3056	10
3/12	67786	5	51553	1255	3062	6
3/13	67790	4	52943	1390	3075	13
3/14	67794	4	54278	1335	3085	10
3/15	67798	4	55094	816	3099	14
3/16	67799	1	55987	893	3111	12
3/17	67800	1	56883	896	3122	11
3/18	67800	0	57678	795	3130	8
3/19	67800	0	58381	703	3132	2
3/20	67800	0	58942	561	3139	7
3/21	67800	0	59432	490	3144	5
3/22	67800	0	59879	447	3153	9
3/23	67801	1	60323	444	3160	7
3/24	67801	0	60810	487	3163	3
3/25	67801	0	61201	391	3169	6
3/26	67801	0	61731	530	3174	5
3/27	67801	0	62098	367	3177	3
3/28	67801	0	62565	467	3182	5
3/29	67801	0	62882	317	3186	4
3/30	67801	0	63153	271	3187	1
3/31	67802	1	63326	173	3193	6
4/1	67802	0	63471	145	3199	6
4/2	67802	0	63612	141	3203	4
4/3	67803	1	63762	150	3207	4
4/4	67803	0	63945	183	3210	3
4/5	67803	0	64014	69	3212	2
4/6	67803	0	64073	59	3212	0
4/7	67803	0	64142	69	3213	1
4/8	67803	0	64187	45	3215	2

附录 B 长春市 COVID-19 疫情数据

日期	确诊新增	无症状转确诊新增	无症状新增
3/4	5	0	0
3/5	7	0	0
3/6	7	0	5
3/7	17	0	6
3/8	12	0	10
3/9	23	0	25
3/10	2	0	21
3/11	63	0	97
3/12	831	0	42
3/13	430	22	3
3/14	460	1	3
3/15	1044	3	3
3/16	268	0	5
3/17	595	0	0
3/18	743	1	1
3/19	833	0	3
3/20	1079	2	1
3/21	1437	2	0
3/22	1979	57	0
3/23	1280	0	253
3/24	576	68	290
3/25	554	7	403
3/26	633	56	495
3/27	706	120	330
3/28	622	11	310
3/29	875	19	348
3/30	997	14	422
3/31	1078	21	444
4/1	1544	19	894
4/2	723	25	3100
4/3	574	64	2346

附录 C SIR 模型 python 代码

```
import numpy as np

import matplotlib.pyplot as plt

from scipy.integrate import odeint

def SIR(sir, t, beta, gamma):
    S, I, R = sir

    dsdt = - beta * S * I
    didt = beta * S * I - gamma * I
    drdt = gamma * I

    return [dsdt, didt, drdt]

t = np.linspace(0, 20, 5000)

S0, I0, R0 = 1000, 1, 0

beta, gamma = 0.004, 0.25

result = odeint(SIR, [S0, I0, R0], t, args=(beta, gamma))

St, It, Rt = result[:, 0], result[:, 1], result[:, 2]

plt.plot(t, St, c="g", label="Susceptibles")
plt.plot(t, It, c="r", label="Infectives")
plt.plot(t, Rt, c="b", label="Removed")

plt.title("SIR model", fontsize=18)
plt.xlabel("days", fontsize=18)
plt.ylabel("number of people", fontsize=18)
plt.legend(fontsize=18)
plt.grid(True)
plt.show()
```