



西南财经大学

SOUTHWESTERN UNIVERSITY OF FINANCE AND ECONOMICS

硕士学位论文

MASTER'S DISSERTATION

全球典型国家新冠肺炎时间序列的变点检测及研究

Change Point Detection for COVID-19 Time Series in Typical
Countries around the World

学位申请人 陈牟瑶

指导教师 余睿

学科专业 应用统计

学位类别 专业学位

分类号 _____ 密级 _____

U.D.C _____

全球典型国家新冠肺炎时间序列的变点检测及研究

Change Point Detection for COVID-19 Time Series in Typical
Countries around the World

学位申请人: _____ 陈牟瑶

学 号: _____ 220025200062

学 科 专 业: _____ 应用统计

研 究 方 向: _____ 大数据分析

指 导 教 师: _____ 余睿

定 稿 时 间: _____ 2022 年 5 月 16 日

西南财经大学

学位论文原创性及知识产权声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明，因本学位论文引起的法律结果完全由本人承担。

本人同意在校攻读学位期间论文工作的知识产权单位属西南财经大学。本人完全了解西南财经大学有关保留、使用学位论文的规定，即学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权西南财经大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印、数字化或其他复制手段保存和汇编本学位论文。

本学位论文属于

1、☐ 保密，在____年解密后适用本授权书。

2、☒ 不保密

特此声明。

学位申请人：陈宇瑶
2022年5月16日

摘要

新型冠状病毒肺炎自 2019 年 12 月爆发以来备受关注，全球各个国家及地区的疫情发展状况差异明显，疫情时间序列呈现出不同的发展特征，防控效果也各不相同。本文在已有新冠肺炎研究成果的基础上，将“滑动窗口”的理念与变点统计分析方法相结合并应用到公共卫生领域当中，选取美国、巴西、英国、俄罗斯、印度、印度尼西亚、菲律宾、中国共 8 个新冠肺炎典型国家作为研究对象，对 2020 年 1 月 3 日至 2021 年 11 月 1 日的累计确诊病例日增长率时间序列进行均值变点检测，定量刻画每个国家的疫情发展态势，结合相关疫情防控政策与措施分析变点存在的原因，并进行多个维度的拓展性研究。主要研究结果如下：第一，模拟实验得出 SN-based 检验效果比 KS 检验更加稳定，因此选择 SN-based 检验用于本文新冠肺炎时间序列的研究；第二，变点所在位置与政策措施的实施时间紧密相关，且普遍存在一定的滞后效应，即在政策措施实施时间之后的一段时间确诊病例增长率才会出现明显变化，从而出现变点；第三，根据累计确诊病例曲线走向可将 8 个国家划分为三种类型：长期平稳型、阶段波动型和持续扩张型；第四，经国家之间不同维度的均值变点检测结果及疫情发展形势对比发现，自变点检测起始位置起 180 天内的疫情走向与国家各自首次发布居家隔离、社交距离限制等严格防疫措施的时间关联性较强；第五，为进一步验证均值变点检测结果，选定适当的 ARMA 模型对相邻变点间的平稳子序列进行拟合，拟合效果因国家而异，个别国家拟合优度较高，可达到 0.85 以上。

关键词：新冠肺炎；均值变点；滑动窗口；疫情防控

Abstract

The COVID-19 has attracted much attention since its outbreak in December 2019 and the development in various countries and regions around the world is obviously different. Based on the existing research results, this paper combines the concept of ‘sliding window’ with the statistical analysis method of change points and applies it to the field of public health. It selects the United States, Brazil, the United Kingdom, Russia, India, Indonesia, Philippines and China as the research objects, and conducts mean change point test on the time series of daily growth rates of cumulative confirmed cases from January 3, 2020 to November 1, 2021. It aims to quantitatively describe the development trend of each country’s epidemic situation, analyze the reasons for the change points and conduct extended researches on multiple dimensions. The main results obtained are listed as follows: First, the simulation experiment shows that the SN-based test is more stable than the KS test, so the SN-based test is selected for the study of COVID-19 time series in this paper. Second, the location of the change point and the implementation time of the policies and measures is closely related. There is generally a certain lag effect. Third, according to the trend of the cumulative confirmed case curve, 8 countries can be divided into three types: long-term stable, phased fluctuation and continuous expansion. Fourth, the comparison of the detection results of the mean change point and the development situation of COVID-19 between countries shows that, the trend within 180 days from the starting position of the change point detection has a strong time correlation with the country's first release of strict epidemic prevention measures such as home isolation and social distancing restrictions. Fifth, in order to further verify the mean change point detection results, ARMA models are selected to fit the stationary subsequence between adjacent change points. The effect of the fitting model varies from country to country, and individual countries have a high goodness of fit, which can reach more than 0.85.

Keywords: COVID-19; Mean change point; Sliding window; Epidemic prevention

目录

摘要.....I

Abstract.....II

目录.....III

1. 引言.....1

 1.1 研究背景.....1

 1.2 研究意义.....1

2. 文献综述.....3

 2.1 变点问题简述.....3

 2.1.1 变点理论.....3

 2.1.2 变点分类及常用方法.....4

 2.2 变点问题研究进程.....4

 2.2.1 国外研究现状.....4

 2.2.2 国内研究现状.....5

 2.3 新冠肺炎研究进程.....6

 2.3.1 国外研究现状.....6

 2.3.2 国内研究现状.....7

 2.4 变点问题在新冠肺炎研究中的应用.....8

3. 时序均值变点检验设计.....10

 3.1 变点检测方法.....10

 3.1.1 CUSUM 算法.....10

 3.1.2 KS 检验.....10

 3.1.3 SN-based 检验.....11

 3.2 KS 检验与 SN-based 检验方法比较.....11

 3.2.1 检验方法模拟.....12

 3.2.2 KS 检验与 SN-based 检验比较结果.....13

4. 累计确诊病例日增长率的均值变点检测.....15

 4.1 研究数据.....15

 4.1.1 数据来源.....15

 4.1.2 数据处理.....16

4.2 基于滑动窗口的序列均值变点检测.....	16
4.3 日增长率序列的变点检测结果与原因分析.....	17
4.3.1 美洲区域	17
4.3.2 欧洲区域	22
4.3.3 东南亚区域	26
4.3.4 西太平洋区域	30
5. 基于变点检测结果的扩展研究.....	36
5.1 国家新冠肺炎时间序列的变点对比.....	36
5.1.1 各区域内的国家对比.....	36
5.1.2 各区域确诊病例最高的国家对比.....	39
5.1.3 疫情发生最早与最晚国家对比.....	40
5.1.4 人口相近的国家对比.....	41
5.2 国家新冠肺炎平稳时间子序列的拟合	42
6. 总结与展望.....	44
6.1 研究结论.....	44
6.2 研究展望.....	45
参考文献.....	46
附录.....	50
致谢.....	51

1. 引言

1.1 研究背景

新型冠状病毒肺炎（Corona Virus Disease 2019, COVID-19），简称“新冠肺炎”，自 2019 年 12 月爆发以来，加剧蔓延扩散至全球各个国家和地区，其传播路径呈现多中心化的特征。

日内瓦时间 2020 年 1 月 30 日，世界卫生组织（WHO）召开记者会宣布：新型冠状病毒疫情被列为“国际关注的突发公共卫生事件”（Public Health Emergency of International Concern，即 PHEIC）。在之后的一段时间内，全球范围内的 COVID-19 确诊病例数、死亡人数和受疫情影响国家的数量不断攀升。3 月 11 日，WHO 正式宣布可以用“Pandemic”形容新型冠状病毒的流行形势，即新型冠状病毒进入全球大流行状态。

即使在提倡全面接种新冠疫苗的背景下，新冠肺炎疫情目前仍然在全球多个国家内持续蔓延。截至北京时间 11 月 8 日，全球新冠肺炎累计确诊 250,584,988 例，累计死亡 5,064,202 例。全球单日新增确诊 352,844 例，新增死亡 4,826 例。所有新增确诊病例中有一半以上来自欧洲国家，美国的严重疫情未有好转，而日本、德国、印度等地区的疫情不断反复，无法有效得到全面控制，之所以存在这样的现象，其中一个重要原因就是不断出现的新冠病毒变体，大大增加了疫情防控的难度和力度。

此次肆虐全球的公共卫生事件对各个国家来说，都是一场持久战，同时也彰显了国际社会之间互相合作的必要性，吸取经验、完善措施，才能与病毒进行有效抗争。根据钟南山院士的预计，如果按照目前的形势，新冠病毒对我们人类身体健康的影响可能会逐渐降低，以后可能就跟流感一样，每年定期打疫苗，人类和新冠病毒逐渐发展为长期共存的一个状况。

1.2 研究意义

新型冠状病毒肺炎爆发以来，全球各个国家及地区呈现出不同的发展态势，疫情时间序列呈现出不同的发展特征，防控效果也各不相同。全球不同国家及地区的疫情发展到底呈现什么特征，不同特征背后的主要原因是什么，对于这一问题的揭示将有助于研究疫情时间序列的发展规律以及不同防控效果的有效性。

目前针对新冠肺炎时间序列特征及其影响因素的研究虽然已有不少，但大部分对国家疫情爆发的时间序列特征研究比较单一，国家及地区之间疫情防控措施与效力也缺乏系统的对比。本文将统计学中的变点问题思路应用到公共卫生领域当中，进行不同维度的探索发现，这同时也是一个相对全新维度的研究方向。变点问题是统计学中的一个经典分支，“模型中的某个或某些量起突然变化之点”是对“变点”的简单定义，它通常反映事物或过程的某种质的变化，在各个领域内都很常见。本文将对全球新冠肺炎典型国家的累计确诊病例日增长率时间序列进行变点检测与分析，基于已检测到的变点可将序列划分为多个平稳子序列，进而从变点分析的角度定量刻画每个国家的疫情发展态势，结合相关政策法规、新闻事件资料分析变点存在的原因，选择适当的模型对变点检测结果进行验证，从多个维度对比各国的防控措施与防控效果，从而有针对性地提出建议。

2. 文献综述

2.1 变点问题简述

2.1.1 变点理论

变点问题是统计学中的一个经典分支，陈希孺(1991)在《变点统计分析简介》中将变点定义为“模型中的某个或某些量起突然变化之点”^[1]，通常反映事物或过程的某种质的变化，在各个领域内都很常见。也就是在一个序列或过程中，当某个统计特性（分布类型、分布参数）在某时间点受系统性因素而非偶然性因素影响发生变化，我们就称该时间点为变点。具体定义：假设存在一个数据集，每个数据观测值相互独立，如果在某一时刻，模型中的某个或某些变量突然发生了变化，即存在一个时间点，在该点之前，数据集符合一个分布，在该点之后，数据集符合另外一个分布，则该点为该数据集的变点。

变点检测即利用一定的统计指标或统计方法，对时间序列的状态进行观测，以便准确有效地估计出变点的位置。假设存在时间序列 $\{X_t\}(t = 1, 2, 3, \dots, n)$ ，则均值变点检验原假设和备择假设如下：

$$H_0: E(X_1) = \dots = E(X_n) = \mu,$$

$$H_1: E(X_1) = \dots = E(X_{k^*}) \neq E(X_{k^*+1}) = \dots = E(X_n),$$

其中 $1 \leq k^* < n$ 且 k^* 未知。

统计学中对变点的研究历史不长，主要集中在变点理论的研究方面，涉及较多的数学理论知识与公式证明。近年来，逐渐有学者将变点的理论研究与实际数据相结合进行一些简单的变点检测实证研究。

2.1.2 变点分类及常用方法

陈希孺(1991)在《变点统计分析简介》第一篇“变点问题提法”中论及变点问题模型主要分为三个类别：均值变点、回归变点和概率变点，各自对应的突然发生变化的“某个统计特性”分别是均值、回归系数和事件的概率。

变点统计分析中常用方法有最小二乘法、极大似然法、累计次数法、贝叶斯法、局部比较法等。最小二乘法是在处理变点问题时使用频率较高的一种方法，且模型中随机误差项的分布不需要有特定的假设。其主要原理就是将观察值与理论值之差的平方和设为目标函数，得到其达到极小值时的点，作为有关参数即变点的点估计。极大似然法、累计次数法、贝叶斯法是在处理概率变点问题时比较常用的方法。极大似然法就是将变点看作参数，通过求取似然函数的极大值来估计该参数，其中的依赖关系较复杂，因此涉及更多的计算。累计次数法是基于最小二乘法中事先肯定变点个数为1时衍生的思想，计算简单，优点在于可以求出其在原假设（不存在变点）之下的渐近分布，因此主要用于检验变点是否存在。贝叶斯法的原理是，将包括变点在内的模型中的参数作为随机变量，确定其先验分布，通过先验分布和样本分布来推断变点这个参数的后验分布及其他相关信息，优点在于可以对变点进行点估计、区间估计并检验存在与否。局部比较法适用于各类变点模型，其重点在于考察“某个统计特性”在各个“局部”内的变化并进行比较，从而发现差别，取“局部”内变化最为显著之处作为变点估计，该方法涉及复杂的分布问题和随机误差。

2.2 变点问题研究进程

2.2.1 国外研究现状

对于变点问题的研究最早可追溯到 Page(1954)在国际顶级期刊 *Biometrika* 上发表的一篇关于连续抽样检验的文章^[2]，文章主要阐述了变点在质量控制方面的应用，通过抽检生产线上的产品来检测其质量是否存在显著波动，

从而能够控制总体质量在一个稳定范围内。在这以后,变点问题被广泛应用于多个领域,比如:经济学、金融学、气象科学、医学、流行病学等。

Yao, Yi-Ching 和 Richard A. Davis(1986)对满足正态性且独立同分布的时间序列进行均值变点检验,发现当变点处于序列中不早不晚的位置时,贝叶斯(Bayes)检验的效果要优于似然比(LR)检验^[3]。D.L. Hawkins(1988)基于非参数分布应用 Kolmogorov-Smirnov 型统计量,并评估该统计量的检验效果^[4]。Ian B. MacNeill(1974)在独立同分布序列中使用 CUSUM 统计量,并得出在原假设下其收敛于布朗桥的结论^[5]。Brown, Dubin 和 Evans(1975)提出累加平方和算法(CUSUM-SQ),适用于对时间序列中存在的多个结构性方差变点进行检测^[6]。Inclan, Carla 和 George C. Tiao(1994)基于 CUSUM-SQ 算法进行优化,提出新的算法即迭代累加平方和算法(Iterative Cumulative Sums of Squares, ICSS)^[7]。ICSS 算法一经提出便被众多研究者广泛应用于金融时间序列的分析当中,具有重大意义。Vera R. Eastwood(1993)介绍了变点检测问题中不同的非参数方法并给出相应示例^[8]。Csorg 和 Horvath(1997)对变点统计分析问题中的极限理论进行深入研究,总结了近二十年来这一领域的研究进展与成果^[9]。

2.2.2 国内研究现状

变点问题是一个比较新的统计研究方向,在国内从事这方面研究的人员虽然不是很多,但已发表了部分成果并取得显著成效。我国统计学者在这一领域的研究开始于上个世纪 90 年代。陈希孺教授(1991)在《数理统计与管理》一书的《变点统计分析简介》篇章中详细介绍了变点统计模型及针对不同变点模型采用的具体方法^[1]。在当时,变点统计分析这一课题的发展还不够成熟,但这一篇章的刊载为之后的研究奠定了重要的理论基石。

杨喜寿和杨洪昌(1996)提出 U-似然比检验统计量,用于检验济南市 1919-1988 年平均气温序列的变点,有助于划分气候阶段^[10]。谭智平和缪柏其(2000)基于分布变点问题进行非参数推断,采用 Kolmogorov-Smirnov 型检验统计量对香港股票市场的隐含波动率进行变点分析^[11]。谭智平和缪柏其(2003)提出改进的 Kolmogorov-Smirnov 型检验统计量,模拟检测变点的程序并证

明所得结果具备强相合性^[12]。熊立华、周芬等(2003)将贝叶斯方法与 MCMC 抽样理论相结合,对水文时间序列进行变点分析^[13]。谭常春、赵林城等(2007)利用滑窗方法对时间序列参数变点的假设检验和估计问题进行讨论,得出变点估计的强弱收敛速度与强相合性,并对中国股市收益率进行了实证分析^[14]。李绍刚和杨少华(2010)研究了基于 GARCH 变点模型的上证收益率伪波动持续性,通过 ICSS 算法检测变点,并且结合历史背景检验变点,通过分段建模来消除其伪波动性^[15]。Xiaofeng Shao 和 Xianyang Zhang(2010)基于传统的 Kolmogorov-Smirnov 型检验统计量,提出 self-normalization(SN)的思想并将其运用至时间序列的变点检测问题^[16]。

2.3 新冠肺炎研究进程

新型冠状病毒肺炎爆发以来,全球各个国家及地区呈现出不同的发展态势,防控效果也各不相同。根据《数据资讯:全球新冠肺炎及其疫苗研究》^[17],截至 2021 年 1 月 31 日,各国学者针对新冠肺炎进行了大量研究,其中研究论文发文量前 3 位的国家分别是美国、中国、英国,研究角度多数与冠状病毒等医疗卫生主题相关,除此之外,主要可以分为两个维度:疫情时间序列的研究和疫情影响因素的研究。

2.3.1 国外研究现状

自疫情爆发以来,国外学者陆续展开对于新冠肺炎的研究,主要包含:对所在国家疫情进行时间序列预测、对比世界部分国家及地区疫情发展趋势、探索政策措施与疫情防控效果的关联性等。

Talkhi Nasrin、Fatemi Narges Akhavan 等(2021)基于 2020 年 2 月 20 日至 8 月 15 日伊朗确诊及死亡病例数据,应用了 NNETAR、ARIMA、Hybrid 等 9 个模型来预测未来 30 天的数据,其中 MLP 和 Holt-Winter 分别是预测确诊人数与死亡人数效果最佳的模型,预测误差相对较低^[18]。Saswat Singh、Chandreyee Chowdhury 等(2021)基于截至 2020 年 6 月印度的确诊病例数据,采取多种方法对 2020 年 8 月数据进行预测,比如 TBAT、ARIMA、

N-BEATS 等, 最终将预测效果最好的 ARIMA 模型用于研究印度封锁时期与解封时期疫情的发展形势, 在病毒传播模式不会发生异常变化的前提下, 这种方法对于预测是非常具有价值的^[19]。Palmer William R、Davis Richard A 等(2021)为更好地观察 COVID-19 每日死亡人数的动态变化, 提出了一个广义非线性状态空间模型即一个涉及二阶差分和 AR(1)-ARCH(1)模型的潜在状态过程, 该模型拟合结果展示了与多个动态阶段相关的动态变化, 且适用于纽约市及德克萨斯州其他地区的死亡人数时间序列^[20]。Gyeongseung Han、Jeewon Han 等(2021)对全球、韩国及美国的疫情发展序列进行探索性数据分析(Exploratory Data Analysis, EDA), 得出此次全球大流行对经济产生一定的负面影响, 且受 COVID-19 影响最大的工作类型是与休闲或酒店相关的工作^[21]。Sri Rahayu、Intan Zainafree 等(2020)采用横断面研究, 随机选取 350 名印度尼西亚三宝垄地区市民作为样本, 设定自变量为年龄、职业等社区特征, 因变量为戴口罩、洗手等针对 COVID-19 的预防措施, 对以上变量中的分类数据和数值数据分别进行 Logistic 检验和 Mann Whitney 检验, 旨在研究社区特征与 COVID-19 预防措施之间的关联性^[22]。Zhang Junyi、Zhang Runsen 等(2021)从政策维度出发, 评估澳大利亚、加拿大、日本、新西兰、英国和美国 6 个国家在 2020 年不同月份的共 418 项政策措施(其中 244 条为交通措施)的有效性并进行对比, 每项政策的有效性由 9 个感染和流动性指标来衡量, 对应于政策实施前后的三个时期。此外还构建了动态贝叶斯多级广义结构方程模型用于研究政策之间的一系列因果关系, 为交通和公共卫生部门制定更有效的 COVID-19 政策提供了科学合理的依据^[23]。

2.3.2 国内研究现状

新冠肺炎疫情爆发并报告于中国湖北省武汉市, 而国内学者也陆续针对本次“大流行”进行大量研究, 分析角度各不相同, 对于整个公共卫生领域的研究具有一定意义。

白璐、郭佩汶等(2020)选取 ARIMA(1,1,1)模型拟合湖北省 2020 年 1 月 20 日至 3 月 4 日的新冠肺炎确诊病例时间序列, 并进行 6 期预测, 得出疫情仍存在上涨趋势但有所减缓的结论, 并提出相应的政策建议^[24]。邵年、钟敏

等(2020)在已有的时滞动力学模型(TDD-NCP)基础上,结合中国疾病预防控制中心(CCDC)的经验分布和随即动力行为数据,衍生出新的模型即 FUDAN-CCDC,用于预测上海市、武汉市的疫情发展趋势^[25]。胡海文(2020)采用随机森林重要性评分法,进行特征筛选,确定对美国新冠肺炎确诊人数与死亡人数分别影响最大的 10 个因素,建立长短时记忆网络(LSTM)模型对美国疫情发展进行预测^[26]。谢聪慧、吴世新等(2021)采取基于特征的时间序列聚类方法,提取标准差、Hurst 指数等 6 个主要特征,对全球疫情相对严重的 60 个国家的每日新增确诊病例序列进行谱系聚类,最终将国家分为三个模式:C 类、S 类、I 类,并从经济、医疗、人文冲突角度对该结果进行成因分析^[27]。付吉丽、黄立强(2020)选取中国、美国等 6 个新冠肺炎代表性国家,分别对确诊病例、死亡比例序列进行数学特征相关研究,并建立指标对各个国家的疫情防控效果进行综合评价^[28]。武蓓蓓(2021)以江西省 2020 年 1 月 23 日至 2 月 19 日每日新增病例数据为基础,对新冠肺炎发病率进行空间自相关分析,并绘制相应的空间分布专题图,从可视化的角度直观展现了疫情分布状况^[29]。郭文璇和房昊(2021)从人口迁徙、流动的角度来分析中国疫情防控措施的有效性,研究除港澳台之外 31 个省市自治区在 2020 年 1 月 1 日至 2 月 29 日的日度数据,证明人口流动强度更高的地区,疫情传染的风险会更大,这也是对党中央“内防扩散、外防输出”政策的进一步支持^[30]。

2.4 变点问题在新冠肺炎研究中的应用

2003 年于我国爆发的 SARS 疫情,使国内外学者开始重视“公共卫生”这一领域的问题研究,尤其是对于突发事件的分析和预测,显得更为重要。李仲来、崔恒建等(2003)对北京市 2003 年 4 月-6 月的 SARS 疫情数据进行分析,分别建立 SI 模型和分段 SI 模型用于预测,其中分段 SI 模型引入了“变点”的理念,即在某一时刻增长率发生了转折性变化。研究结果也表明,在 4 月 20 日左右,由于政府对于疫情采取了一定措施,使得 4 月 24 日之后 SARS 的确诊病例增长率发生明显变化^[31]。被称为“全球大流行”的新冠肺炎病毒疫情自 2020 年 12 月开始不断蔓延,是世界各个国家及地区公共卫生

体系面临的共同挑战，部分国内外学者将变点问题的思路应用到新冠肺炎研究当中，获得不同的研究成果。

Vokó Zoltán 和 Pitter János György(2020)对 28 个欧洲国家的新冠肺炎发病数据进行中断时间序列分析，采取线性门限回归模型的方法估计序列中的变点位置，运用泊松回归模型观察变点处的具体变化程度，并基于 Google Community Mobility Reports 计算社会距离指数(Social Distance Index, SDI)，旨在刻画社会距离与新冠时间序列之间的联系。结果发现，SDI 四分位数最高的国家的病例数在统计上显著下降，而 SDI 四分位数不是很高的国家的疫情相对于最初也有明显改善，由此得出“居家”这一国家政策对遏制欧洲的 COVID-19 大流行做出了有意义的贡献^[32]。Han Lin Shang 和 Ruofan Xu(2021)运用一种分层分裂的估计方法来检测比利时 2019 年第 27 周至 2020 年第 20 周的周度超额死亡率序列的多个变点，并按性别及年龄分组进行进一步检测。研究发现变点为 2020 年第 13 周和第 18 周，这与新冠肺炎疫情的扩散密切相关，且呈现明显的年龄差别，即老年组死亡人数的激增比其他年龄组出现得更早，持续时间也更长^[33]。Coughlin Steven S.、Yiğiter Ayten 等(2020)基于全球 20 个国家以及包含 27 个国家在内的欧盟的 2020 年 1 月至 5 月 COVID-19 每日新增确诊病例数据，使用 B-spline 模型进行拟合，并采取统计方法去估计序列中的变点位置，并结合各个国家应对疫情实施的政策措施进行研究，发现有效的措施与疫情得到控制之间存在一定的滞后效应^[34]。Jiang F、Zhao Z 等(2020)将新生的 SN-based 变点检验方法与 NOT 算法相结合，应用于全球 30 个国家的 COVID-19 累计确诊病例与死亡序列当中，进行多变点估计，对比发现不同国家之间疫情防控的关联性，此外还设计了一个两阶段方案用于预测美国未来的累计死亡病例，取得良好结果^[35]。刘君和罗晓媛(2021)基于我国某省份 2019 年 12 月 1 日至 2021 年 1 月 6 日新冠肺炎确诊病例间隔时间的数据，补全左截断和右删失部分数据，求得完整数据服从指数分布下的似然函数，利用 MCMC 方法估计间隔时间的突变时刻即变点，该方法是将 Gibbs 抽样与 Metropolis-Hastings 算法相结合得到联合样本，对后验分布的统计量进行估计的方法，最终得出结论：确诊病例间隔时间为 2 天或 5 天，说明疫情得到有效控制^[36]。

3. 时序均值变点检验设计

3.1 变点检测方法

3.1.1 CUSUM 算法

本文应用 Page 提出的 CUSUM(Cumulative Sum)算法对变点进行识别,这也是近几十年来识别变点的经典算法。CUSUM 算法主要通过研究序列均值和方差的信息,从而判断过程是否发生变化,其基本思想是:给定平均水平条件 θ ,将每个监测量与平均水平差的累加和作为观察对象,当监测量的 CUSUM 明显比正常平稳运行条件下的平均水平高或者低时,意味着系统内发生了变化。一般通过监测样本的均值、方差等统计量作为平均水平条件。由于使用方便、判断准则简单、易于操作等原因, CUSUM 算法在工业质量控制、自动故障监测、经济、金融等方面应用广泛。

CUSUM 过程^[37]的计算公式如下:

$$T_n(\lfloor nr \rfloor) = n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} (X_t - \bar{X}_n), \quad r \in [0,1], \quad (3.1)$$

其中 $\bar{X}_n = n^{-1} \sum_{t=1}^n X_t$, 是序列 $\{X_t\}$ 的均值。

3.1.2 KS 检验

Kolmogorov Smirnov 检验,简称 KS 检验,是统计学中的一种非参数假设检验,用来检测单样本是否服从某一分布,或者两样本是否服从相同分布。基于 CUSUM 计算公式 3.1,可以得到:

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor nr \rfloor} \{X_t - E(X_t)\} \Rightarrow \sigma B(r), \quad (3.2)$$

其中 $\sigma^2 = \lim_{n \rightarrow \infty} n \text{var}(\bar{X}_n)$, $B(r)$ 表示一维布朗运动。在原假设的前提下,进一步得到 $T_n(\lfloor nr \rfloor) \Rightarrow \sigma\{B(r) - rB(1)\}$,由此可以计算 KS 检验统计量^[38],公式如下:

$$KS_n = \sup_{r \in [0,1]} |T_n([nr])/ \hat{\sigma}_n| = \sup_{k=1,2,\dots,n} |T_n(k)/ \hat{\sigma}_n|, \quad (3.3)$$

这里的 $\hat{\sigma}_n^2$ 是 σ^2 的一致估计量, $\hat{\sigma}_n^2 = \sum_{k=-l_n}^{l_n} \hat{\gamma}(k)K(k/l_n)$, 其中 $\hat{\gamma}(k)$ 是样本滞后 k 阶的自协方差估计, $K(\cdot)$ 是核函数, l_n 是窗宽参数, 本文设定 $l_n = \lfloor n^{1/3} \rfloor$ 。

设定显著性水平 $\alpha = 0.05$, 若 KS 检验统计量大于临界值 1.36 (KS 检验临界值见附表 1), 则拒绝原假设, 认为该时间序列存在变点。为获得变点所在具体位置, 可通过以下公式求得:

$$\hat{k} = \arg \max_{k=1,2,\dots,n} |T_n(k)/ \hat{\sigma}_n|. \quad (3.4)$$

3.1.3 SN-based 检验

参考 Xiaofeng Shao 和 Xianyang Zhang(2010)^[16]将 self-normalization 思想进行延伸, 提出了新的检验方法, 就是 SN-based 检验, 简单来说就是将 KS 检验统计量计算中的 $\hat{\sigma}_n$ 替换为 D_n , $D_n^2 = n^{-2} \sum_{t=1}^n \{ \sum_{j=1}^t (X_j - \bar{X}_n) \}^2$ 。为方便之后的计算过程, 对于 $k = 1, 2, \dots, n-1$, 定义 V_n 如下:

$$V_n(k) = n^{-2} \left[\sum_{t=1}^k \{ S_{1,t} - (t/k) S_{1,k} \}^2 + \sum_{t=k+1}^n \{ S_{t,n} - (n-t+1)/(n-k) S_{k+1,n} \}^2 \right], \quad (3.5)$$

其中, 对于 $t_1 \leq t_2$, $S_{t_1,t_2} = \sum_{j=t_1}^{t_2} X_j$ 。所以 SN-based 检验统计量计算公式如下:

$$SN_n = \sup_{k=1,2,\dots,n-1} T_n(k)' V_n^{-1}(k) T_n(k). \quad (3.6)$$

设定显著性水平 $\alpha = 0.05$, 若 SN-based 检验统计量大于临界值 40.1 (SN-based 检验临界值见附表 2), 则拒绝原假设, 认为该时间序列存在变点。为获得变点所在具体位置, 可通过以下公式求得:

$$\hat{k} = \arg \max_{k=1,2,\dots,n-1} T_n(k)' V_n^{-1}(k) T_n(k). \quad (3.7)$$

3.2 KS 检验与 SN-based 检验方法比较

3.2.1 检验方法模拟

在统计学中，假设检验的关键是确定检验统计量和临界区域，而好的检验统计量必然是能够很好地区分原假设和备择假设。参考《Econometric Methods with Applications in Business and Economics》第 56 页“Size and power”部分的解读^[39]，可以根据以下两种概率判断检验方法的效果：

1) 第 I 类错误（弃真）：原假设为真，但是检验的结果落在拒绝域，因而拒绝原假设。犯第 I 类错误的概率称为检验水平，用 size 表示；

2) 第 II 类错误（存伪）：原假设为假，但是检验的结果未落在拒绝域，因而未拒绝原假设。正确拒绝第 II 类错误的概率称为检验功效，用 power 表示。

在对新冠肺炎时间序列进行变点检测之前，本文首先进行初步模拟，依据 size 和 power 的大小及稳定性，比较 KS 检验和 SN-based 检验的效果，选择检验效果更优的一种方法用于之后的实证分析。

构建时间序列模型 $\{Y_t\}$ ($t = 1, 2, \dots, n; Y_0 = 0$):

$$Y_t = \varphi_t + \varepsilon_t, \quad (3.8)$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t, \quad u_t \sim i.i.d N(0,1), \quad (3.9)$$

其中 φ_t 是时间序列中的常数变化量 c ， ε_t 是一个一阶自回归模型，即 AR(1)， ρ 是自相关系数($-1 < \rho < 1$)。建立变点检验的原假设 $H_0: E(Y_1) = \dots = E(Y_n) = \mu$ ，备择假设 $H_1: E(Y_1) = \dots = E(Y_{k^*}) \neq E(Y_{k^*+1}) = \dots = E(Y_n)$ ， $1 \leq k^* < n$ 且 k^* 未知。给定显著性水平 $\alpha = 0.05$ ，计算检验统计量，如统计量大于相应临界值，则拒绝原假设 H_0 ，认为时间序列 $\{Y_t\}$ 存在变点。

本文将通过调整模型参数来计算不同参数取值下的 size 和 power。其中自相关系数 ρ 取值 $\{-0.8, -0.4, 0, 0.4, 0.8\}$ ；常数变化量 c 取值 $\{-1, -0.5, 0, 0.5, 1\}$ ，这里 $c=0$ 为特殊情况，即时间序列不存在变点；时间序列 $\{Y_t\}$ 的样本量 n 取值 $\{200, 500\}$ ；重复次数 m 取值 $\{100, 200\}$ 。最终判定检验效果的标准是检验水平 size 值稳定且接近显著性水平 $\alpha = 0.05$ ，检验功效 power 值稳定且尽可能大。

3.2.2 KS 检验与 SN-based 检验比较结果

基于以上思路,利用 Python 对已构建的时间序列 $\{Y_t\}$ 分别进行 KS 检验和 SN-based 检验,详细的检验结果见表 3.1、表 3.2。

表 3.1 KS 检验模拟结果

replications: m = 100						
n	ρ	size	power			
		($c=0$)	($c=-1$)	($c=-0.5$)	($c=0.5$)	($c=1$)
200	-0.8	0	1	0.42	0.46	1
	-0.4	0.02	1	0.8	0.83	1
	0	0.03	1	0.9	0.87	1
	0.4	0.09	1	0.9	0.93	1
	0.8	0.34	1	0.99	0.99	1
500	-0.8	0	1	0.99	0.99	1
	-0.4	0.03	1	1	1	1
	0	0.03	1	1	1	1
	0.4	0.08	1	1	1	1
	0.8	0.29	1	1	1	1
replications: m = 200						
n	ρ	size	power			
		($c=0$)	($c=-1$)	($c=-0.5$)	($c=0.5$)	($c=1$)
200	-0.8	0.005	0.995	0.47	0.36	1
	-0.4	0.02	1	0.765	0.795	1
	0	0.05	1	0.855	0.88	1
	0.4	0.07	1	0.935	0.86	1
	0.8	0.255	1	0.965	0.985	1
500	-0.8	0.01	1	0.975	0.98	1
	-0.4	0.02	1	0.99	0.995	1
	0	0.03	1	1	1	1
	0.4	0.07	1	1	1	1
	0.8	0.275	1	1	1	1

表 3.2 SN-based 检验模拟结果

replications: m = 100						
n	ρ	size	power			
		($c=0$)	($c=-1$)	($c=-0.5$)	($c=0.5$)	($c=1$)
200	-0.8	0.02	0.99	0.69	0.78	1
	-0.4	0.02	1	0.81	0.77	0.99
	0	0.04	0.99	0.77	0.81	0.99
	0.4	0.03	1	0.79	0.79	1
	0.8	0.07	1	0.82	0.84	1
500	-0.8	0.04	1	0.99	0.97	1
	-0.4	0.03	1	1	0.99	1
	0	0.05	1	0.98	0.97	1
	0.4	0.05	1	0.97	0.97	1
	0.8	0.04	1	1	0.99	1

replications: $m = 200$

n	ρ	size	power			
		($c=0$)	($c=-1$)	($c=-0.5$)	($c=0.5$)	($c=1$)
200	-0.8	0.035	0.995	0.675	0.7	0.985
	-0.4	0.025	0.995	0.755	0.795	0.995
	0	0.05	1	0.81	0.77	0.99
	0.4	0.06	1	0.755	0.75	1
	0.8	0.06	1	0.83	0.8	1
500	-0.8	0.02	1	0.99	0.98	1
	-0.4	0.045	1	0.995	0.985	1
	0	0.05	1	0.98	0.975	1
	0.4	0.055	1	0.985	0.98	1
	0.8	0.06	1	0.985	0.985	1

根据以上检验结果，可以分析得到 KS 检验与 SN-based 检验之间的一些异同点如下：第一，随着自相关系数 ρ 的变化，KS 检验在 $\rho=0.8$ 时的 size 上上升至 0.3 左右，扭曲严重，远超过显著性水平 $\alpha=0.05$ ，因此过度拒绝原假设，表现不稳定；而 SN-based 检验的 size 表现更加稳定，基于显著性水平 $\alpha=0.05$ 上下浮动，幅度很小；第二， $n=200$ ， $c=-0.5$ 和 0.5 时，随着自相关系数 ρ 的变化，KS 检验的 power 由 0.4 左右上升至 0.9 以上，表现不稳定；而 SN-based 检验的 power 虽然不是很高，却始终稳定在 0.65-0.85 范围内；第三，除第二点中提到的情况外，其余情况下 KS 检验与 SN-based 检验的 power 值相近，均在 0.9 以上，最高可达到 1。

通过梳理总结 KS 检验与 SN-based 检验的异同点，发现二者在大部分参数设定的情况下检验结果相近，但是 SN-based 检验具备一项显著优势，即随自相关系数 ρ 的变化，size 表现更佳，始终基于显著性水平 $\alpha=0.05$ 上下浮动，且浮动幅度很小，比 KS 检验更加稳定。因此，本文选择将 SN-based 检验方法应用于新冠肺炎累计确诊病例日增长率的均值变点检测。

4. 累计确诊病例日增长率的均值变点检测

4.1 研究数据

4.1.1 数据来源

本文研究采用的数据是 8 个新冠肺炎典型国家的日度累计确诊病例，时间范围是 2020 年 1 月 3 日至 2021 年 11 月 1 日，来源于世界卫生组织官方网站（<https://www.who.int/>）。

本文所研究的 8 个国家分别是美国、巴西、英国、俄罗斯、印度、印度尼西亚、菲律宾和中国。世界卫生组织（World Health Organization, WHO）会员国分成 6 个区域，每个区域设立一个区域办事处。截至 2021 年 11 月 1 日，这 6 个区域的累计确诊病例排序如表 4.1 所示，本文选取美洲区域、欧洲区域、东南亚区域、西太平洋区域各 2 个国家，其中美国、英国、印度、菲律宾分别是相应区域内累计确诊病例最多的国家，而巴西、俄罗斯、印度尼西亚、中国分别是相应区域内比较有代表性的国家。截至 2021 年 11 月 1 日，8 个国家的累计确诊病例如表 4.2 所示。

表 4.1

区域	累计确诊病例
美洲区域 Americas	93,652,790
欧洲区域 Europe	76,995,605
东南亚区域 South-East Asia	43,984,655
东地中海区域 Eastern Mediterranean	16,365,479
西太平洋区域 Western Pacific	9,442,686
非洲区域 Africa	6,152,212

注：数据来源：<https://covid19.who.int/>

表 4.2

国家	累计确诊病例	国家	累计确诊病例
----	--------	----	--------

美国	45,635,708	巴西	21,804,094
英国	9,057,633	俄罗斯	8,554,192
印度	34,285,814	印度尼西亚	4,244,761
菲律宾	2,787,276	中国	126,078

注：数据来源：<https://covid19.who.int/table>

4.1.2 数据处理

由于累计确诊病例数据过于庞大，本文采用累计确诊病例的日增长率作为序列均值变点的检测对象。累计确诊病例日增长率 $rate_t$ 的计算公式如下：

$$rate_t = \frac{cases_t - cases_{t-1}}{cases_{t-1}}, \quad (4.1)$$

其中 $cases_t$ 表示日期为 t 的累计确诊病例数量。

4.2 基于滑动窗口的序列均值变点检测

本文研究的阶段性目的是检测出 8 个国家新冠肺炎累计确诊病例日增长率序列的全部均值变点，由于已有的变点检测方法只能检测出某时间序列内的单一变点，所以参考田腾、石茂林等(2021) 提出的基于滑动窗口的时间序列异常检测方法^[40]，将“滑动窗口”的理念引入至变点检测当中。

原始累计确诊病例数据时间范围是 2020 年 1 月 3 日至 2021 年 11 月 1 日，共 669 天。由于各个国家疫情开始的时间即第一个新冠肺炎确诊病例出现的时间不一致；且在疫情发生的第一个月，全球各个国家及地区的疫情均处于急速扩张的状态，这一阶段的序列一定是非平稳的，即存在变点，所以本文均值变点检测起始位置是国家首例确诊日期之后的第 30 天(1 个月左右)，受检测数据是检测起始位置至 2021 年 11 月 1 日的累计确诊病例日增长率序列 $\{rate_t\} (t = 1, 2, \dots, n)$ 。8 个国家的首例确诊日期、序列均值变点检测起始位置、序列样本量的具体说明见表 4.3。

表 4.3

国家	首例确诊	检测起始位置	样本量
美国	01-20	02-19	622
巴西	02-26	03-27	585
英国	02-01	03-02	610

俄罗斯	01-31	03-01	611
印度	01-30	02-29	612
印度尼西亚	03-02	04-01	580
菲律宾	01-31	03-01	611
中国	01-04	02-03	638

注：表中日期所在年份均为 2020 年

基于滑动窗口检测时间序列变点的具体思想：设置滑动窗口长度 $l = 30$ ，则初始滑窗子序列 $window_1 = \{rate_1, rate_2, \dots, rate_{30}\}$ ，利用 SN-based 检验该子序列是否存在变点。(1)如果存在变点，变点位置记为 k_1 ，滑动窗口继续检测下一个子序列 $window_2 = \{rate_{k_1+1}, rate_{k_1+2}, \dots, rate_{k_1+30}\}$ ，即变点位置 k_1 之后的 30 天；(2)如果不存在变点，滑动窗口继续检测下一个子序列 $window_2 = \{rate_1, rate_2, \dots, rate_{30}, rate_{31}, \dots, rate_{60}\}$ ，即在原子序列基础上增加往后滑动的 30 天，形成新的受检测子序列。持续滑动窗口进行检测，直至滑动至新冠肺炎累计确诊病例日增长率序列结束为止，最终得到该序列中的全部变点： $k_1, k_2, k_3, \dots, k_s$ ，共计 s 个变点。

本文研究所用到的工具是 Python。

4.3 日增长率序列的变点检测结果与原因分析

8 个国家的均值变点检测具体结果见相应的累计确诊病例时间序列图（图中橙色竖线是变点检测起始位置）和表格（表中“间隔范围”指相邻变点之间的时间范围，“平均增长率”即所在间隔范围的累计确诊病例日增长率的均值）。

4.3.1 美洲区域

(1) 美国

基于美国 2020 年 2 月 19 日至 2021 年 11 月 1 日（622 天）累计确诊病例日增长率序列，共检测到 22 个变点。具体结果见图 4.1 和表 4.4。美国疫情整体呈现波浪形发展态势，先后出现四个高峰。

2020年2月27、28日,因出现感染路径不明的第15号病例,美国疾控中心(CDC)修改核酸检测标准,纳入对感染路径不明、无接触史患者的排查,并开始将检测权限下放至各州实验室。在这之后新增确诊病例一路攀升,平均增长率由5.32%上升至20.78%,变化幅度高达15.46%,在4月初达到第一个高峰,出现第2个变点4月3日。2020年3月13日,特朗普宣告美国进入国家紧急状态,下令实施“保持社交距离”措施;3天后,白宫正式发布《总统的美国冠状病毒指导方针:15天减缓传播》(The President's Coronavirus Guidelines for America: 15 days to slow the spread)的疫情防控指南。3月19、20日,加利福尼亚州、纽约州率先宣布全州范围的“居家隔离令”。4月4日,CDC在个人防护建议中新增“佩戴布质口罩”。这些政策措施也是在4月初之后的一段时间内疫情开始缓慢回落的主要原因,变点4月3日之后平均增长率不断下降。

因疫情有所缓和,特朗普公布“再次开放美国”的联邦指南,分三个阶段全面重启经济,5月起,美国50个州陆续不同程度地重启经济活动,且部分州未实行“口罩令”。疫情形势在6月中旬之后开始恶化,在7月下旬达到第二个高峰,出现第8个变点7月27日,6月25日至7月27日的平均增长率为1.81%,相较上一阶段上升了0.6个百分点。此次高峰的出现使得美国再次加强疫情防控,截至7月23日,“口罩令”已在超过35个州推行。8月-10月疫情再度缓慢回落,可以看到第10个与第11个变点即8月31日与10月14日之间间隔长达45天。

第三个高峰出现在2020年11月至2021年1月。由于天气变冷、美国大选、圣诞节等原因导致室内活动增多,且早在8月6日美国国务院就已宣布解除“第四级全球旅行警告”,使得年底国际旅行人数增多,美国疫情形势不容乐观,新增确诊人数一路激增。10月30日、12月26日相继出现两个变点,时间跨度较长,平均增长率变化状态均为上升,10月31日至12月26日的平均增长率相较上一阶段上升了0.52%。2021年1月20日拜登上台之后,实施全面推进疫苗接种计划,承诺要在执政100天内完成1亿剂疫苗接种,并严格执行“口罩令”。新的抗议举措使得美国疫情在往后长达半年的时间内有明显缓和,2021年1月25日至7月9日,共出现6个变点,变点之间间隔的平均增长率均呈现下降态势,其中最大幅度为下降了0.44%。3月,各州陆续取

消封锁措施并重启大型社交活动；5月13日，CDC发布新的口罩指导方针，不再强制要求已完全接种新冠疫苗的人们在室外以及大部分室内情况下佩戴口罩。

随着新冠病毒“德尔塔”变异毒株的传播，美国疫情出现反弹，在2021年7-9月迎来第四个高峰。8月4日、9月19日相继出现两个变点，平均增长率变化状态均为上升，最高上升至0.38%，与2月上半月的平均增长率相近，形势一度回转。由于该病毒在易感人群中广泛流行，美国再次加强疫情防控，比如：7月30日，CDC强调不论疫苗接种状态，在公共室内、交通工具等相对密闭空间内均需佩戴口罩或面罩；纽约市长宣布于8月16日起要求人们在室内商业场所出示接种新冠疫苗证明……

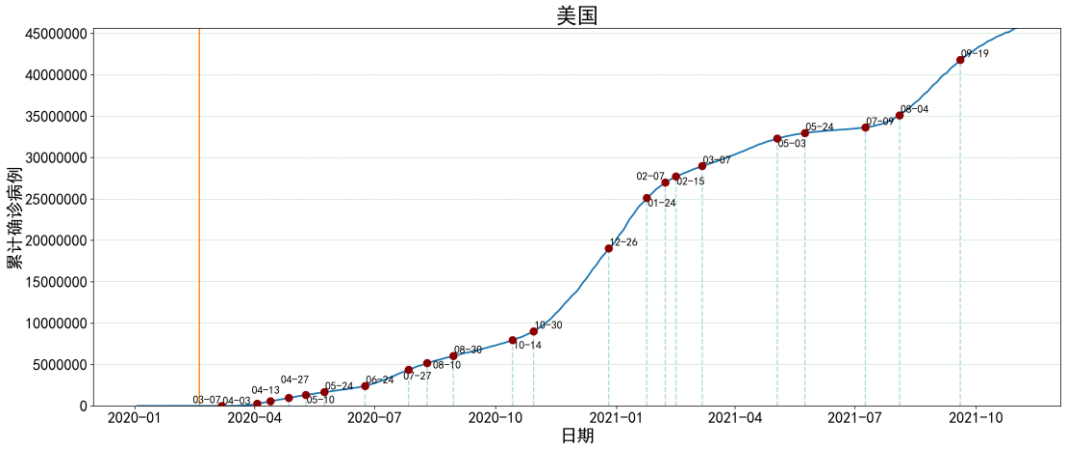


图 4.1 美国累计确诊病例时间序列

表 4.4

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-03-07	2020-02-19 ~ 2020-03-07	17	5.32	
2	2020-04-03	2020-03-08 ~ 2020-04-03	27	20.78	+15.46
3	2020-04-13	2020-04-04 ~ 2020-04-13	10	8.34	-12.44
4	2020-04-27	2020-04-14 ~ 2020-04-27	14	3.91	-4.43
5	2020-05-10	2020-04-28 ~ 2020-05-10	13	2.41	-1.50
6	2020-05-24	2020-05-11 ~ 2020-05-24	14	1.66	-0.75
7	2020-06-24	2020-05-25 ~ 2020-06-24	31	1.21	-0.45
8	2020-07-27	2020-06-25 ~ 2020-07-27	33	1.81	+0.60
9	2020-08-10	2020-07-28 ~ 2020-08-10	14	1.19	-0.62

10	2020-08-30	2020-08-11 ~ 2020-08-30	20	0.79	-0.40
11	2020-10-14	2020-08-31 ~ 2020-10-14	45	0.61	-0.18
12	2020-10-30	2020-10-15 ~ 2020-10-30	16	0.80	+0.19
13	2020-12-26	2020-10-31 ~ 2020-12-26	57	1.32	+0.52
14	2021-01-24	2020-12-27 ~ 2021-01-24	29	0.96	-0.36
15	2021-02-07	2021-01-25 ~ 2021-02-07	14	0.52	-0.44
16	2021-02-15	2021-02-08 ~ 2021-02-15	8	0.33	-0.19
17	2021-03-07	2021-02-16 ~ 2021-03-07	20	0.22	-0.11
18	2021-05-03	2021-03-08 ~ 2021-05-03	57	0.19	-0.03
19	2021-05-24	2021-05-04 ~ 2021-05-24	21	0.10	-0.09
20	2021-07-09	2021-05-25 ~ 2021-07-09	46	0.04	-0.06
21	2021-08-04	2021-07-10 ~ 2021-08-04	26	0.16	+0.12
22	2021-09-19	2021-08-05 ~ 2021-09-19	46	0.38	+0.22
		2021-09-20 ~ 2021-11-01	44	0.20	-0.18

(2) 巴西

基于巴西 2020 年 3 月 27 日至 2021 年 11 月 1 日（585 天）累计确诊病例日增长率序列，共检测到 12 个变点。具体结果见图 4.2 和表 4.5。巴西疫情整体呈现持续蔓延扩张的态势，变点仅有 12 个，说明累计确诊病例日增长率序列相对平稳，疫情没有明显的好转。

2020 年 10 月（包含 10 月）之前共有 8 个变点。虽然变点之间间隔范围的平均增长率变化幅度均为下降，但实际上疫情形势是非常不容乐观的，背后存在诸多原因。2020 年 3 月以来，巴西国内围绕“居家隔离”这一防疫措施的分歧十分严重。卫生部长曼代塔坚持必须严格执行 WHO 建议的“保持社交距离”措施，但总统博索纳罗多次公开带头反对该建议，强调恢复经济运行，禁止限制居民活动。联邦最高法院表示，地方州市有权决定各自地区的防疫政策，总统无权干涉。3 月 24 日，圣保罗州开始采取居家隔离、限制商业活动等防疫措施，但并未严格落实；5 月 27 日，又宣布逐步放宽社会隔离措施，在这之后，各州陆续放松，计划恢复经济活动。之所以出现平均增长率逐渐下降，很大程度是因为巴西医疗卫生体系不发达，核酸检测未能全面覆盖。7 月，一项由巴西卫生部与佩洛塔斯联邦大学联合进行的调查显示，巴西实际的新冠肺炎感染人数已超过 1050 万，是官方数据的 7 倍左右。

截至论文研究时间结束，2020 年 10 月之后仅有 4 个变点。第 9、10 个变点分别是 2020 年 11 月 18 日、2021 年 4 月 12 日，间隔长达 145 天，且平均

增长率相较上一间隔上升 0.16%，说明在这将近 5 个月（145 天）内，新增确诊病例一路攀升，保持较高的日增长率，疫情形势十分严峻。这一阶段的疫情反弹使得巴西政府再次加强防控措施。2020 年 12 月 17 日，巴西最高联邦法院(STF)认为有必要对未接种疫苗的人采取限制性措施。疫情最为严重的圣保罗州宣布，从 1 月 25 日起，全州进入防疫最为严格的红色阶段。2021 年 1 月中旬之后陆续开始疫苗接种工作，截至 9 月 5 日，已有 64.79%的人口接种至少一剂疫苗，图 4.2 可以看到在第 11 个变点 6 月 28 日之后，巴西疫情曲线明显趋缓。

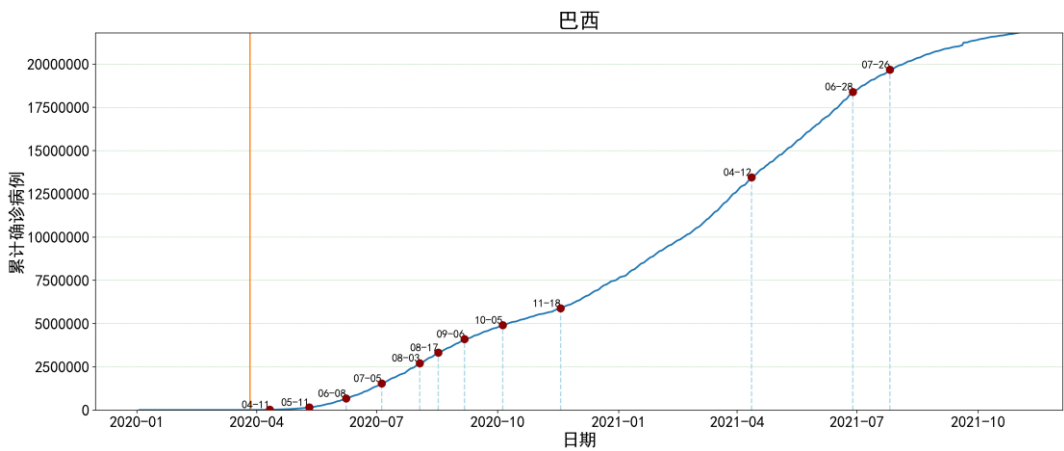


图 4.2 巴西累计确诊病例时间序列

表 4.5

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-04-11	2020-03-27 ~ 2020-04-11	15	13.41	
2	2020-05-11	2020-04-12 ~ 2020-05-11	30	7.51	-5.90
3	2020-06-08	2020-05-12 ~ 2020-06-08	28	5.37	-2.14
4	2020-07-05	2020-06-09 ~ 2020-07-05	27	3.12	-2.25
5	2020-08-03	2020-07-06 ~ 2020-08-03	29	1.97	-1.15
6	2020-08-17	2020-08-04 ~ 2020-08-17	14	1.46	-0.51
7	2020-09-06	2020-08-18 ~ 2020-09-06	20	1.06	-0.40
8	2020-10-05	2020-09-07 ~ 2020-10-05	29	0.63	-0.43
9	2020-11-18	2020-10-06 ~ 2020-11-18	44	0.41	-0.22
10	2021-04-12	2020-11-19 ~ 2021-04-12	145	0.57	+0.16
11	2021-06-28	2021-04-13 ~ 2021-06-28	77	0.41	-0.16
12	2021-07-26	2021-06-29 ~ 2021-07-26	28	0.24	-0.17

		2021-07-27 ~ 2021-11-01	99	0.11	-0.13
--	--	-------------------------	----	------	-------

4.3.2 欧洲区域

(1) 英国

基于英国 2020 年 3 月 2 日至 2021 年 11 月 1 日（610 天）累计确诊病例日增长率序列，共检测到 22 个变点。具体结果见图 4.3 和表 4.6。英国疫情整体呈现波浪形发展态势，先后历经四轮爆发过程。

第一轮爆发是疫情发现后的 2020 年 3 月至 4 月，这一阶段共检测到 4 个变点，间隔时间都相对较短，平均增长率较高。这一阶段重点关注的是英国政府提出的“群体免疫”政策。3 月 12 日，首相宣布英国正式从“遏制”阶段转移到“延缓”阶段，即之前采取的外防输入措施失败，疫情在本土开始全面蔓延。“群体免疫”的不积极态度收到民众的反对，于是在 3 月 20 日，英国政府正式下令关闭学校、餐厅、咖啡厅等场所，并从 3 月 23 日起实施“全民居家隔离”。这一措施使得疫情出现明显缓和，变点 4 月 10 日、5 月 3 日各自对应的平均增长率下降 8.25%、6.34%。5 月 10 日，英国政府宣布分阶段实施解封计划，允许一定条件下的自由外出；从 7 月 4 日起正式开始英格兰大规模解封计划。

过早的解封计划使得英国从 2020 年 9 月起迎来疫情的第二轮爆发。连续出现四个平均增长率变化均为上升的变点：8 月 26 日、9 月 25 日、10 月 1 日、11 月 7 日，其中变点 9 月 25 日与 10 月 1 日之间仅相隔 6 天，这一阶段的平均增长率变化幅度最大，增加了 0.93%，说明疫情反弹十分迅速。10 月 12 日，针对疫情风险最高的英格兰地区，英国政府出台更为严格的“三级封锁”措施，即按照疫情严重程度将英格兰各地区划分为三个等级：中等风险、高风险和非常高风险，分别采取不同的防控措施；并从 10 月 30 日起启动“封城”。之后的 11、12 月初疫情明显缓和，变点 11 月 20 日、12 月 11 日对应的平均增长率分别下降了 0.73%、0.79%，幅度较大。

由于新冠病毒变种，英国疫情在 2020 年 12 月中旬开始第三轮爆发。可以看到 2020 年 12 月 12 日至 2021 年 1 月 17 日这一阶段的平均增长率跳跃式上升至 1.72%，图 4.3 中对应的曲线斜率明显增大。2021 年 1 月 4 日，英国

再次启动第三轮全国范围的疫情封锁措施。图 4.3 中自 2021 年 2 月以后，疫情明显得到一定控制，曲线斜率即日增长率逐渐接近于 0，尤其是第 20、21 个变点之间间隔长达 67 天，且平均增长率下降，说明疫情十分稳定。主要原因是英国政府呼吁全民接种疫苗，2 月 15 日前已有超过 1500 万人接种了第一剂疫苗，进展顺利。

2021 年 6 月英国疫情再度出现第四轮爆发，此次是受德尔塔及其他变异病毒的影响。最后一个变点出现在 6 月 26 日，在往后的四个月左右（129 天）的时间内再未检测到任何变点，且对应曲线斜率十分陡峭，说明这一阶段的日增长率持续维持在较高的水平，十分稳定。根据 BBC 报道，英国首相在 7 月 5 日的发布会上宣布英国将于 7 月 19 日解除现行所有封锁措施，取消保持一米社交距离、强制戴口罩等措施，自此之后将真正意义上回归正常生活。这也是英国疫情自 6 月之后形势变得更为严峻，且不再得到有效控制的主要原因。

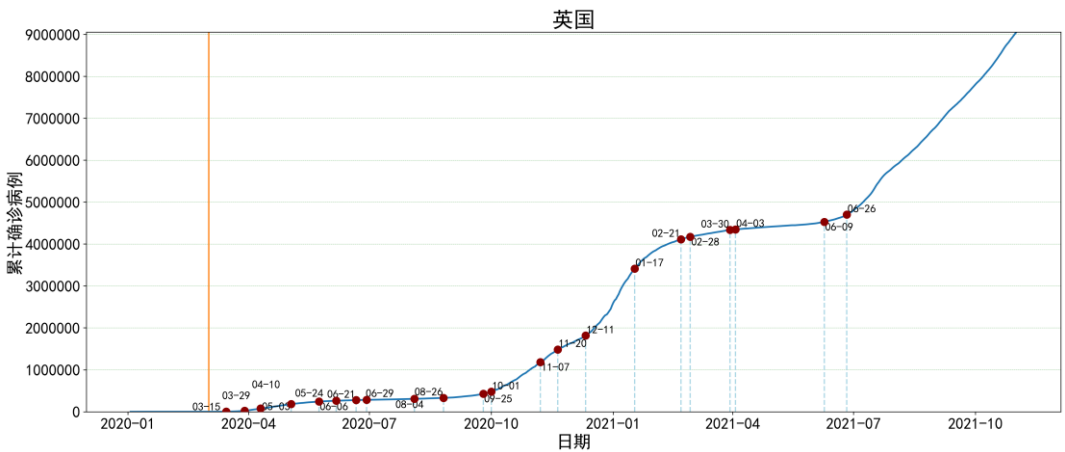


图 4.3 英国累计确诊病例时间序列

表 4.6

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-03-15	2020-03-02 ~ 2020-03-15	13	33.58	
2	2020-03-29	2020-03-16 ~ 2020-03-29	14	18.43	-15.15
3	2020-04-10	2020-03-30 ~ 2020-04-10	12	10.18	-8.25
4	2020-05-03	2020-04-11 ~ 2020-05-03	23	3.84	-6.34
5	2020-05-24	2020-05-04 ~ 2020-05-24	21	1.38	-2.46

6	2020-06-06	2020-05-25 ~ 2020-06-06	13	0.58	-0.80
7	2020-06-21	2020-06-07 ~ 2020-06-21	15	0.37	-0.21
8	2020-06-29	2020-06-22 ~ 2020-06-29	8	0.27	-0.10
9	2020-08-04	2020-06-30 ~ 2020-08-04	36	0.22	-0.05
10	2020-08-26	2020-08-05 ~ 2020-08-26	22	0.33	+0.11
11	2020-09-25	2020-08-27 ~ 2020-09-25	30	0.88	+0.55
12	2020-10-01	2020-09-26 ~ 2020-10-01	6	1.81	+0.93
13	2020-11-07	2020-10-02 ~ 2020-11-07	37	2.49	+0.68
14	2020-11-20	2020-11-08 ~ 2020-11-20	13	1.76	-0.73
15	2020-12-11	2020-11-21 ~ 2020-12-11	21	0.97	-0.79
16	2021-01-17	2020-12-12 ~ 2021-01-17	37	1.72	+0.75
17	2021-02-21	2021-01-18 ~ 2021-02-21	35	0.54	-1.18
18	2021-02-28	2021-02-22 ~ 2021-02-28	7	0.20	-0.34
19	2021-03-30	2021-03-01 ~ 2021-03-30	30	0.13	-0.07
20	2021-04-03	2021-03-31 ~ 2021-04-03	4	0.09	-0.04
21	2021-06-09	2021-04-04 ~ 2021-06-09	67	0.06	-0.03
22	2021-06-26	2021-06-10 ~ 2021-06-26	17	0.22	+0.16
		2021-06-27 ~ 2021-11-01	129	0.51	+0.29

(2) 俄罗斯

基于俄罗斯 2020 年 3 月 1 日至 2021 年 11 月 1 日（611 天）累计确诊病例日增长率序列，共检测到 21 个变点。具体结果见图 4.4 和表 4.7。俄罗斯疫情整体呈现较为温和的波浪形发展态势，先后历经三轮爆发过程。

第一轮爆发是 2020 年 3 月下旬至 5 月。第一个变点位于变点检测起始位置之后的第 49 天，即 4 月 19 日，出现较晚，原因是早期俄罗斯采取的硬核管控措施。针对疫情爆发最早的中国，俄罗斯政府决定自 1 月 31 日起关闭中俄陆路客运口岸；2 月起，暂停向中方发放工作、旅游、学习、私人访问签证，临时限制上述人员入境。直到 3 月 2 日才发现首例本土确诊。根据俄罗斯海关的数据，3 月下旬的 10 天内约有 100 万人从国外回到俄罗斯，多数为从意大利、法国、西班牙等欧洲疫情严重国家的回国公民；另外随着核酸检测试剂下发至各地，俄罗斯确诊病例快速上升。3 月 27 日，政府宣布采取全国带薪休假到 4 月 30 日、全民居家隔离等严格防疫措施，进入“封城”状态。疫情在一定程度上得到控制，5 月 12 日起开始分三阶段逐步解除疫情限制。6 月至 9 月共检测到 5 个变点，平均增长率均呈现小幅度的下降。

解除封锁使得俄罗斯与欧洲其他疫情严重国家一样在 2020 年 10 月至 12 月迎来第二轮爆发。连续检测到 3 个变点：10 月 9 日、10 月 29 日、12 月 8 日，各自对应的间隔范围平均增长率均上升，幅度分别为 0.30%、0.26%、0.08%，逐渐减小，说明采取的防疫措施有一定效果。10 月 28 日起正式生效法令：启动新一轮新冠肺炎限制措施，要求俄罗斯人在公共场所、交通工具、停车场和电梯里佩戴口罩；与此同时，也在最大限度地扩大疫苗接种力度。图 4.4 中可以看到 2021 年 2 月至 6 月的曲线变得平缓，第 15、16、17 个变点所对应的平均增长率仅下降 0.03%、0.00%、0.01%，说明疫情防控不够严厉，容易出现反弹。

由于德尔塔变异病毒的快速传播，再加上疫苗接种滞后、防空松懈等原因，俄罗斯在 2021 年 6 月中旬迎来疫情的第三轮爆发。变点 6 月 23 日和 7 月 27 日对应的平均增长率变化幅度均为上升。莫斯科市长宣布重启封锁，且下令人们留在家中，并呼吁人们积极接种疫苗。这一措施似乎并未使疫情得到明显控制，其中一个重要原因就是俄罗斯人民不愿意接种疫苗，且对于政府的防疫措施并不完全配合，导致直到 2021 年 10 月俄罗斯疫苗接种率仅有 33%，疫情或会呈现长期化上升趋势。

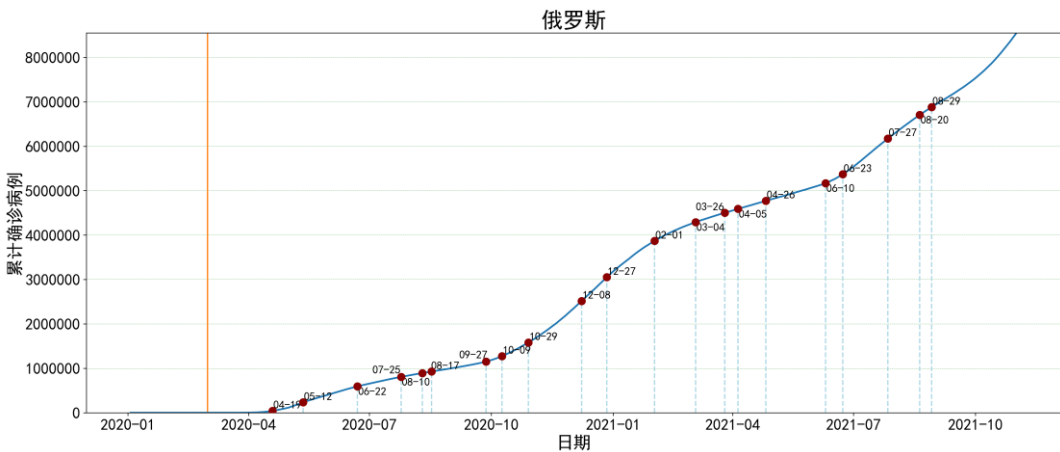


图 4.4 俄罗斯累计确诊病例时间序列

表 4.7

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-04-19	2020-03-01 ~ 2020-04-19	49	25.54	

2	2020-05-12	2020-04-20 ~ 2020-05-12	23	7.64	-17.90
3	2020-06-22	2020-05-13 ~ 2020-06-22	41	2.31	-5.33
4	2020-07-25	2020-06-23 ~ 2020-07-25	33	0.94	-1.37
5	2020-08-10	2020-07-26 ~ 2020-08-10	16	0.63	-0.31
6	2020-08-17	2020-08-11 ~ 2020-08-17	7	0.55	-0.08
7	2020-09-27	2020-08-18 ~ 2020-09-27	41	0.53	-0.02
8	2020-10-09	2020-09-28 ~ 2020-10-09	12	0.83	+0.30
9	2020-10-29	2020-10-10 ~ 2020-10-29	20	1.09	+0.26
10	2020-12-08	2020-10-30 ~ 2020-12-08	40	1.17	+0.08
11	2020-12-27	2020-12-09 ~ 2020-12-27	19	1.02	-0.15
12	2021-02-01	2020-12-28 ~ 2021-02-01	36	0.66	-0.36
13	2021-03-04	2021-02-02 ~ 2021-03-04	31	0.33	-0.33
14	2021-03-26	2021-03-05 ~ 2021-03-26	22	0.22	-0.11
15	2021-04-05	2021-03-27 ~ 2021-04-05	10	0.19	-0.03
16	2021-04-26	2021-04-06 ~ 2021-04-26	21	0.19	-0.00
17	2021-06-10	2021-04-27 ~ 2021-06-10	45	0.18	-0.01
18	2021-06-23	2021-06-11 ~ 2021-06-23	13	0.29	+0.11
19	2021-07-27	2021-06-24 ~ 2021-07-27	34	0.41	+0.12
20	2021-08-20	2021-07-28 ~ 2021-08-20	24	0.35	-0.06
21	2021-08-29	2021-08-21 ~ 2021-08-29	9	0.29	-0.06
		2021-08-30 ~ 2021-11-01	65	0.34	+0.05

4.3.3 东南亚区域

(1) 印度

基于印度 2020 年 2 月 29 日至 2021 年 11 月 1 日（612 天）累计确诊病例日增长率序列，共检测到 22 个变点。具体结果见图 4.5 和表 4.8。印度疫情曲线起伏较大，明显看出存在两个高峰。

第一个变点位于检测位置之后的 43 天，即 2020 年 4 月 12 日，时间跨度超过一个月，说明前期处于相对平稳的状态。3 月 24 日，莫迪总理宣布从 3 月 25 日起全国进入全面封锁状态，持续 21 天，要求居民居家隔离；该封锁令之后被延长至 5 月 3 日。5 月 30 日，印度政府公布三阶段解封策略，高风险地区延长封锁至 6 月 30 日，其他地区的公共场所从 6 月 8 日起开始分阶段陆续开放。直到第 6 个变点 8 月 3 日出现，疫情出现小幅度反弹，在 8 月-9 月达到第一个高峰。虽然平均增长率一直下降，但从图 4.5 可以看到该阶段确诊病例明显上升，单日新增确诊病例数最高超过 9 万。由此可见，全面解封

措施在一定程度上使得疫情出现转折，需进一步加强管控。进入 10 月，新增确诊病例开始缓慢减少。2021 年 1 月 16 日，印度开始全民接种疫苗工作，选择优先为 60 岁以上老年人接种。疫苗工作的有序进行使得疫情得到控制，第 15、16 个变点即 1 月 10 日与 3 月 18 日之间间隔长达 67 天，且平均增长率仅有 0.14%。

2021 年 3 月，莫迪政府宣布印度疫情进入尾声，过早解除一系列严格的居家隔离与封锁措施，拆除此前修建的临时隔离中心，居民防疫心理出现疲态。印度疫情在 4 月迎来第二个高峰，由于新出现的变种病毒“B.1.617（德尔塔）”广泛传播，确诊病例一路暴增，出现 3 个变点：4 月 10 日、4 月 17 日、5 月 9 日，对应平均增长率连续上升，从图 4.5 也可以看到 4-5 月的曲线明显变得陡峭，疫情反弹迅速。印度卫生部 4 月 29 日公布的数据显示，已连续 8 天单日新增病例超过 30 万。第二波感染风暴来临，4 月 21 日，总理再次敦促各邦实行封锁政策，呼吁居民尽可能待在室内；此外，政府积极与其他国家领导人接触，尝试获得医疗援助，并宣布从 5 月 1 日开始，所有 18 岁以上的居民都将接种疫苗，私人疫苗供应商也能销售和管理疫苗。进入 6 月，疫情明显缓和，恢复至 2020 年年底的水平。最后一个变点位于 7 月 2 日，往后的 123 天内再无变点，平均增长率逐渐稳定。

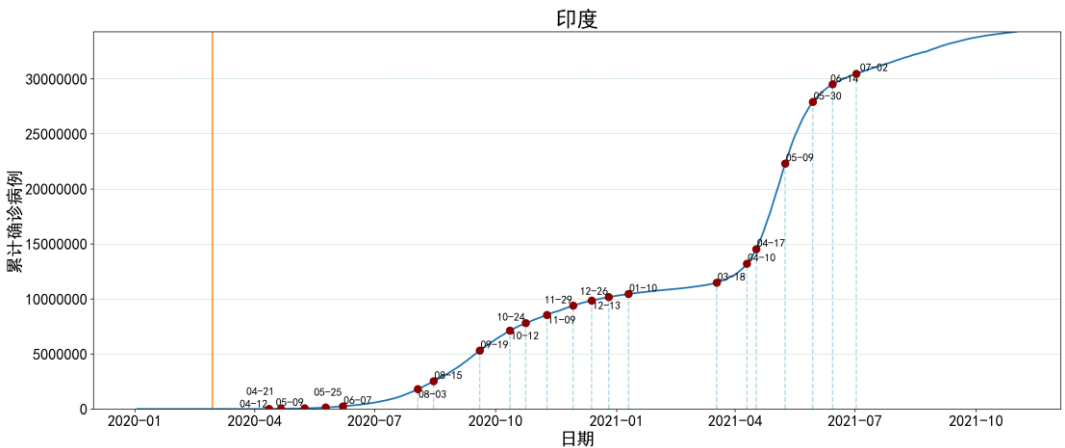


图 4.5 印度累计确诊病例时间序列

表 4.8

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
----	------	------	------	---------	-------------

1	2020-04-12	2020-02-29 ~ 2020-04-12	43	20.11	
2	2020-04-21	2020-04-13 ~ 2020-04-21	9	9.31	-10.80
3	2020-05-09	2020-04-22 ~ 2020-05-09	18	6.69	-2.62
4	2020-05-25	2020-05-10 ~ 2020-05-25	16	5.42	-1.27
5	2020-06-07	2020-05-26 ~ 2020-06-07	13	4.52	-0.90
6	2020-08-03	2020-06-08 ~ 2020-08-03	57	3.55	-0.97
7	2020-08-15	2020-08-04 ~ 2020-08-15	12	2.85	-0.70
8	2020-09-19	2020-08-16 ~ 2020-09-19	35	2.14	-0.71
9	2020-10-12	2020-09-20 ~ 2020-10-12	23	1.29	-0.85
10	2020-10-24	2020-10-13 ~ 2020-10-24	12	0.78	-0.51
11	2020-11-09	2020-10-25 ~ 2020-11-09	16	0.57	-0.21
12	2020-11-29	2020-11-10 ~ 2020-11-29	20	0.47	-0.10
13	2020-12-13	2020-11-30 ~ 2020-12-13	14	0.35	-0.12
14	2020-12-26	2020-12-14 ~ 2020-12-26	13	0.24	-0.11
15	2021-01-10	2020-12-27 ~ 2021-01-10	15	0.18	-0.06
16	2021-03-18	2021-01-11 ~ 2021-03-18	67	0.14	-0.04
17	2021-04-10	2021-03-19 ~ 2021-04-10	23	0.61	+0.47
18	2021-04-17	2021-04-11 ~ 2021-04-17	7	1.37	+0.76
19	2021-05-09	2021-04-18 ~ 2021-05-09	22	1.97	+0.60
20	2021-05-30	2021-05-10 ~ 2021-05-30	21	1.07	-0.90
21	2021-06-14	2021-05-31 ~ 2021-06-14	15	0.38	-0.69
22	2021-07-02	2021-06-15 ~ 2021-07-02	18	0.18	-0.20
		2021-07-03 ~ 2021-11-01	123	0.10	-0.08

(2) 印度尼西亚

基于印度尼西亚 2020 年 4 月 1 日至 2021 年 11 月 1 日 (580 天) 累计确诊病例日增长率序列, 共检测到 15 个变点。具体结果见图 4.6 和表 4.9。印度尼西亚疫情起始较晚, 有两轮波动比较明显。

2020 年 4 月至 5 月, 印度尼西亚疫情前期处于相对温和的阶段, 确诊病例较少, 变点所对应的平均增长率持续下降, 主要原因有两个。其一, 雅加达特区省长巴斯维丹下令, 自 3 月中旬起关闭印尼的学校及娱乐场所, 要求市民居家隔离及办公, 尽量不要出门; 且在 3 月 31 日, 总统佐科制订大规模社区限制政策(Pembatasan Sosial Berskala Besar, PSBB), 采取严格的防疫措施; 4 月 2 日, 政府正式宣布禁止外国人入境, 印尼人回国必须强制隔离和提供健康证明。其二, 印度尼西亚医疗体系落后, 初期每千人的新冠病毒核酸检测次数仅 0.05 次, 很可能存在大量感染患者未发现的情况。及时的抗议政策使得疫情有所缓和, 内政部发布命令, 6 月 1 日至 30 日开启新一阶段的防疫封锁,

划分“严控区”和“非严控区”，针对后者开始陆续放松限制措施，重启经济。变点 7 月 31 日、10 月 4 日、2 月 17 日之间的间隔分别是 65 天、136 天，跨度很长，说明在这近 5 个月的时间内，确诊病例处于稳定上升的状态，既无明显好转也未受到控制。其中因雅加达医疗系统逐渐崩溃，市长宣布于 9 月 14 日起重新执行 PSBB，到 10 月以后再次放松。2021 年 1 月，印尼启动全民接种疫苗计划，疫情得到小幅度的控制，图 4.6 中曲线斜率变小。

2021 年 6 月下旬，印度尼西亚疫情出现明显转折，连续出现 2 个变点：7 月 4 日、8 月 1 日，对应间隔范围 6 月 21 日至 7 月 4 日、7 月 5 日至 8 月 1 日的平均增长率较前一时间间隔分别上升了 0.65%、0.48%，图 4.6 中该阶段曲线变得陡峭，反弹迅速。出现这一现象的原因有两方面：其一，4-5 月印度发现“德尔塔”变种病毒且迅速传播，但印度尼西亚并不重视，在 5 月开始放宽限制措施，且恰逢开斋节结束，据印尼内阁估计，这一时间陆续从全国各地返回首都雅加达的各类车辆高达 100 万，庞大的人员流动促使了德尔塔病毒的超级传播；其二，截至 8 月 2 日，只有不到 10% 的目标人群完成了疫苗接种，而印尼的疫苗接种工作早在 1 月就已开始。疫情的快速反弹使得政府再次加强防控，于 7 月 3 日起实施更严格的社会活动限制，即紧急 PPKM。进入 8 月以后，疫情得到改善，变点对应的平均增长率一路下降。

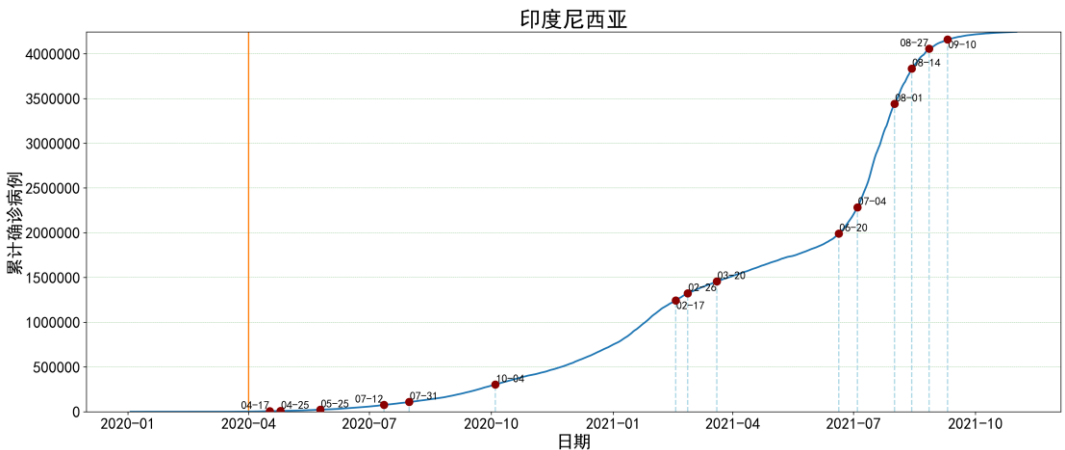


图 4.6 印度尼西亚累计确诊病例时间序列

表 4.9

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%

1	2020-04-17	2020-04-01 ~ 2020-04-17	16	8.31	
2	2020-04-25	2020-04-18 ~ 2020-04-25	8	4.79	-3.52
3	2020-05-25	2020-04-26 ~ 2020-05-25	30	3.30	-1.49
4	2020-07-12	2020-05-26 ~ 2020-07-12	48	2.54	-0.76
5	2020-07-31	2020-07-13 ~ 2020-07-31	19	1.91	-0.63
6	2020-10-04	2020-08-01 ~ 2020-10-04	65	1.60	-0.31
7	2021-02-17	2020-10-05 ~ 2021-02-17	136	1.04	-0.56
8	2021-02-26	2021-02-18 ~ 2021-02-26	9	0.69	-0.35
9	2021-03-20	2021-02-27 ~ 2021-03-20	22	0.44	-0.25
10	2021-06-20	2021-03-21 ~ 2021-06-20	92	0.34	-0.10
11	2021-07-04	2021-06-21 ~ 2021-07-04	14	0.99	+0.65
12	2021-08-01	2021-07-05 ~ 2021-08-01	28	1.47	+0.48
13	2021-08-14	2021-08-02 ~ 2021-08-14	13	0.84	-0.63
14	2021-08-27	2021-08-15 ~ 2021-08-27	13	0.44	-0.40
15	2021-09-10	2021-08-28 ~ 2021-09-10	14	0.18	-0.26
		2021-09-11 ~ 2021-11-01	53	0.04	-0.14

4.3.4 西太平洋区域

(1) 菲律宾

基于菲律宾 2020 年 3 月 1 日至 2021 年 11 月 1 日（611 天）累计确诊病例日增长率序列，共检测到 16 个变点。具体结果见图 4.7 和表 4.10。菲律宾整体疫情发展先慢后快，先后出现三个较为明显的高峰。

2020 年 1 月至 3 月初，菲律宾疫情防控重点在防止境外输入。2 月 2 日开始禁止所有来自包括港澳的中国地区的游客入境，且菲律宾公民和持有菲律宾居留证的外国人入境后必须接受 14 天的强制隔离。这一阶段管控较为成功，仅有 3 例输入型病例，直到 3 月 6 日出现首例本土确诊病例，自此菲律宾防疫工作重点转向控制本土传播。政府制定的社区隔离政策分为 4 个等级，从严格至宽松依次是强化社区隔离、改进版强化社区隔离、一般社区隔离、改进版一般社区隔离。3 月 15 日，政府下令在包括首都大马尼拉在内的吕宋岛实施“强化社区隔离”措施。3 月 20 日，对所有发生疫情本土传播国家的旅客实施旅行禁令，国际航班停止，全菲进入“隔离抗疫”。严格的防控措施使得菲律宾的初期疫情控制得很好，直到 4 月 4 日才出现第 1 个变点，且平均增长率持续下降。自 5 月中旬起开始逐步解封，社区隔离措施等级不断调低，

6月1日进入“一般社区隔离”。疫情开始恶化，呈现多点爆发的态势，到达第一个高峰，菲律宾成为西太平洋地区疫情增长最快的国家。第5与第6个变点之间间隔长达85天，这一阶段即6月6日至8月29日的平均增长率是2.78%，出现小幅上涨。8月2日，政府同意将包括首都大马尼拉地区在内的五个省份的防疫措施升级至“改进版强化社区隔离”。经历这次反弹之后，疫情逐渐转好，可以看到第10与第11个变点之间间隔为121天，说明2020年11月至2021年3月初疫情十分稳定，图4.7中这一阶段曲线趋于水平。

从2021年3月开始，疫情再次出现反复，到达第二个高峰，连续出现两个变点3月20日和4月20日，对应平均增长率分别上涨0.46%和0.43%。政府决定3月29日至4月11日期间，首都大马尼拉地区及周边四个省份实行最严格的“强化社区隔离”，4月12日至4月30日下调至“改进版强化社区隔离”。疫情自4月下旬以后得到一定控制，增长速度减缓。

受德尔塔变异毒株的扩散影响，菲律宾疫情自8月以后再次爆发，在9月到达第三个高峰。最后一个变点出现在8月1日，之后的三个月内的平均增长率再度上涨，确诊病例持续增长，据卫生部报告，8月23日新增确诊病例18332例，为疫情爆发以来单日确诊病例数新高。8月6日起，大马尼拉地区再度实施为期两周的“强化社区隔离”，从图4.7中可以看到直到10月下旬左右，疫情才出现明显缓和。

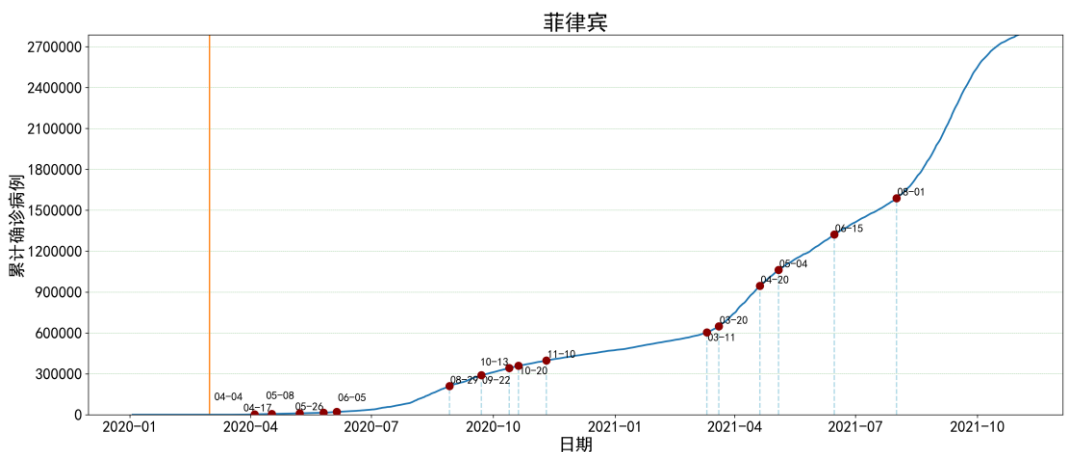


图 4.7 菲律宾累计确诊病例时间序列

表 4.10

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-04-04	2020-03-01 ~ 2020-04-04	34	23.93	
2	2020-04-17	2020-04-05 ~ 2020-04-17	13	4.99	-18.94
3	2020-05-08	2020-04-18 ~ 2020-05-08	21	2.92	-2.07
4	2020-05-26	2020-05-09 ~ 2020-05-26	18	1.82	-1.10
5	2020-06-05	2020-05-27 ~ 2020-06-05	10	3.60	+1.78
6	2020-08-29	2020-06-06 ~ 2020-08-29	85	2.78	-0.82
7	2020-09-22	2020-08-30 ~ 2020-09-22	24	1.37	-1.41
8	2020-10-13	2020-09-23 ~ 2020-10-13	21	0.80	-0.57
9	2020-10-20	2020-10-14 ~ 2020-10-20	7	0.67	-0.13
10	2020-11-10	2020-10-21 ~ 2020-11-10	21	0.50	-0.17
11	2021-03-11	2020-11-11 ~ 2021-03-11	121	0.34	-0.16
12	2021-03-20	2021-03-12 ~ 2021-03-20	9	0.80	+0.46
13	2021-04-20	2021-03-21 ~ 2021-04-20	31	1.23	+0.43
14	2021-05-04	2021-04-21 ~ 2021-05-04	14	0.83	-0.40
15	2021-06-15	2021-05-05 ~ 2021-06-15	42	0.52	-0.31
16	2021-08-01	2021-06-16 ~ 2021-08-01	47	0.39	-0.13
		2021-08-02 ~ 2021-11-01	93	0.61	+0.22

(2) 中国

基于中国 2020 年 2 月 3 日至 2021 年 11 月 1 日（638 天）累计确诊病例日增长率序列，共检测到 19 个变点。具体结果见图 4.8 和表 4.11。中国整体疫情发展先快后慢，初期爆发态势猛烈，在得到控制之后保持稳定，但仍历经多轮小幅反弹。

新冠肺炎疫情爆发于中国湖北武汉，2020 年 1 至 2 月处于疫情爆发期，确诊病例增长速度非常之快。第一个变点出现在 2 月 14 日，此时累计确诊病例为 63932 例，前一日即 2 月 13 日新增确诊多达 15152 例，图 4.8 中可以看到曲线此时明显变得陡峭，原因是数据口径进行调整，将临床诊断病例纳入确诊病例统计。2020 年 1 月 23 日，习近平要求立即对湖北省、武汉市人员流动和对外通道实行严格封闭的交通管控，武汉正式进入“封城”状态；1 月 25 日，国家卫生健康委发布通用、旅游、家庭、公共场所、公共交通工具、居家观察等 6 个公众预防指南；1 月 26 日，国务院办公厅印发通知，决定延长 2020 年春节假期，各地大专院校、中小学、幼儿园推迟开学。一系列严格

的防控措施使得中国疫情在进入 3 月以后很快得到控制, 3 月中旬每日新增病例控制在个位数以内, 本轮流行高峰已经过去。3 月 25 日起, 湖北省有序解除离鄂通道管控措施, 武汉市以外地区逐步恢复正常生产生活秩序; 4 月 8 日, 武汉市解除持续 76 天的“封城”。往后的三个月内全国疫情发展平稳, 明显好转。

在经历初期的流行高峰之后, 全国各地疫情陆续出现小幅波动, 呈现多地局部爆发和零星散发的常态化状态, 主要原因是全球多个国家出现新冠病毒变种毒株, 疫情传播速度加快, 多条传播链源于物流、冷链等相关渠道, 我国目前面临的主要风险是“境外输入”。从变点检测结果可以看出, 自 2020 年 3 月流行高峰过去, 直到 2021 年 11 月 1 日为止, 共出现 4 轮疫情的小幅波动。

第一轮: 2020 年 6 月中旬起, 北京、新疆乌鲁木齐、辽宁大连等地接连出现本土确诊病例, 因此出现变点 7 月 17 日和 8 月 14 日, 对应平均增长率分别上涨 0.03% 和 0.12%。乌鲁木齐市、大连市分别在 7 月 18 日、7 月 25 日进入疫情防控“战时状态”, 停办各类聚集性活动, 对社区实施封闭式管理, 全市范围内开展核酸检测等。本轮疫情在 8 月中旬进入尾声, 地方疫情得到控制。

第二轮: 受国外变种病毒出现和冬季天气寒冷的双重影响, 我国北方地区的疫情自 2020 年 10 月下旬之后开始反复, 涉及辽宁、黑龙江、河北、北京等地, 变异毒株的传播速度更快, 加上 2021 年 1 月进入春运高峰期, 感染风险提高, 因此出现变点 11 月 22 日和 2 月 7 日, 对应平均增长率分别上涨 0.01% 和 0.08%, 且二者之间间隔长达 77 天, 证实这一阶段疫情处于持续小幅增长的状态。各地及时采取与之前相同的防控措施, 另外增加“建议居民非必要不返乡, 尽量留在当地过春节”这一条。本轮疫情在 2 月初进入尾声, 之后长达 98 天内再无变点出现, 疫情发展态势转好。

第三轮: 受德尔塔病毒传播影响, 我国疫情于 2021 年 5 月中旬出现明显恶化, 变点 5 月 16 日与 6 月 13 日之间这一段的平均增长率上涨 0.36%, 是自疫情流行高峰结束之后的最大的上升幅度。在得到短暂控制之后, 7 月 20 日起由南京禄口国际机场引发的疫情一度蔓延至数十余个省份, 尤以湖南、

江苏最为严峻，因此出现变点 8 月 13 日，对应平均增长率上涨 0.04%。本轮疫情在各地采取严格管控措施后开始收尾，到 8 月 22 日实现本土零新增。

第四轮：最后一个变点 9 月 12 日至 11 月 1 日之间的平均增长率较前一阶段上涨 0.01%，出现这一现象的主要原因是 10 月 17 日由“西北地区旅行团”引发的新一轮本土疫情。截至 11 月 16 日，本轮疫情已波及 16 省 31 市，病毒均为德尔塔变异株，由多个不关联的境外输入源头引起。

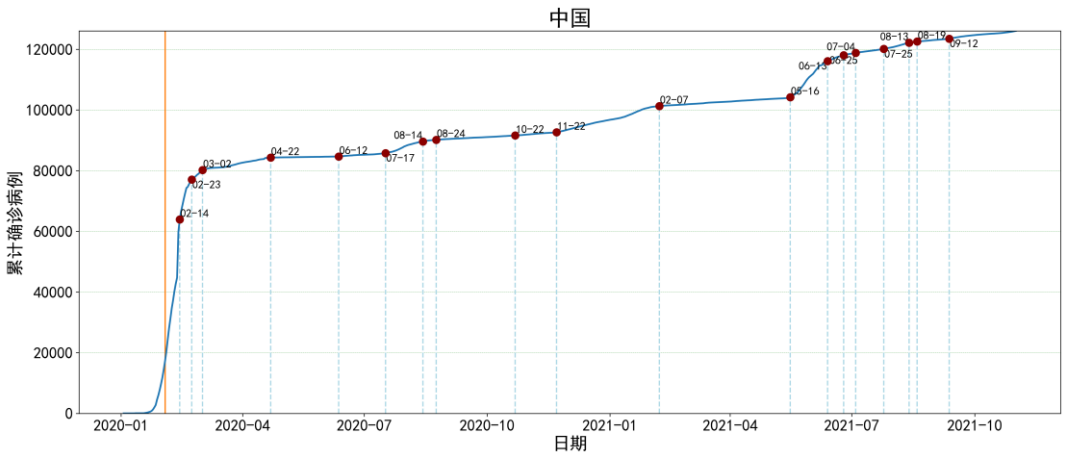


图 4.8 中国累计确诊病例时间序列

表 4.11

序号	变点位置	间隔范围	间隔天数	平均增长率/%	平均增长率变化幅度/%
1	2020-02-14	2020-02-03 ~ 2020-02-14	11	13.49	
2	2020-02-23	2020-02-15 ~ 2020-02-23	9	2.10	-11.39
3	2020-03-02	2020-02-24 ~ 2020-03-02	8	0.50	-1.60
4	2020-04-22	2020-03-03 ~ 2020-04-22	51	0.10	-0.40
5	2020-06-12	2020-04-23 ~ 2020-06-12	51	0.01	-0.09
6	2020-07-17	2020-06-13 ~ 2020-07-17	35	0.04	+0.03
7	2020-08-14	2020-07-18 ~ 2020-08-14	28	0.16	+0.12
8	2020-08-24	2020-08-15 ~ 2020-08-24	10	0.06	-0.10
9	2020-10-22	2020-08-25 ~ 2020-10-22	59	0.03	-0.03
10	2020-11-22	2020-10-23 ~ 2020-11-22	31	0.04	+0.01
11	2021-02-07	2020-11-23 ~ 2021-02-07	77	0.12	+0.08
12	2021-05-16	2021-02-08 ~ 2021-05-16	98	0.03	-0.09
13	2021-06-13	2021-05-17 ~ 2021-06-13	28	0.39	+0.36
14	2021-06-25	2021-06-14 ~ 2021-06-25	12	0.14	-0.25
15	2021-07-04	2021-06-26 ~ 2021-07-04	9	0.07	-0.07

16	2021-07-25	2021-07-05 ~ 2021-07-25	21	0.05	-0.02
17	2021-08-13	2021-07-26 ~ 2021-08-13	19	0.09	+0.04
18	2021-08-19	2021-08-14 ~ 2021-08-19	6	0.05	-0.04
19	2021-09-12	2021-08-20 ~ 2021-09-12	24	0.03	-0.02
		2021-09-13 ~ 2021-11-01	51	0.04	+0.01

5. 基于变点检测结果的扩展研究

5.1 国家新冠肺炎时间序列的变点对比

基于 4.3 变点检测结果与原因的陈述, 本文将从四个维度入手, 进一步对国家之间的新冠肺炎时间序列变点检测结果以及所采取的抗疫政策进行比较, 从而针对疫情防控提出相应的建议。四个维度分别是: 1. 各区域(美洲区域、欧洲区域、东南亚区域、西太平洋区域)内的 2 个国家; 2. 各区域确诊病例最高的国家(4 个); 3. 疫情发生最早与最晚的国家; 4. 人口相近的国家。由于疫情初期的防控政策与措施对于疫情之后的发展走向具有重要意义, 所以本文选定变点检测起始往后的半年(180 天)作为国家之间对比的时间范围。

5.1.1 各区域内的国家对比

(1) 美国与巴西

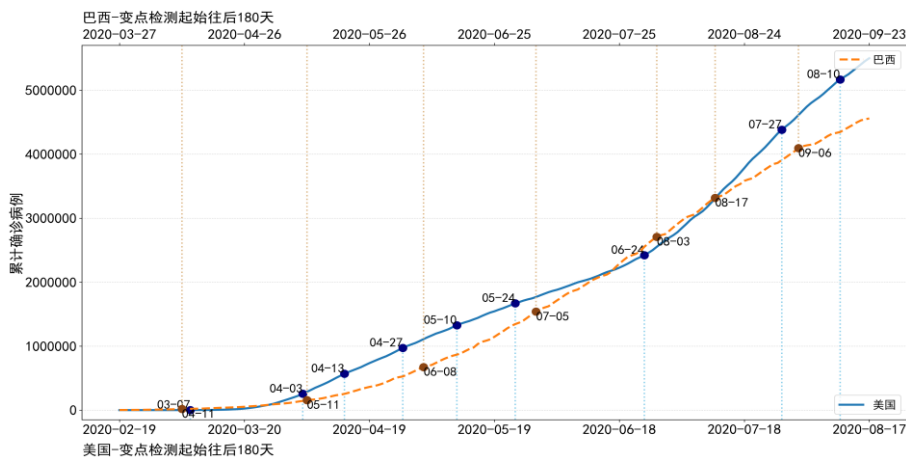


图 5.1 美国与巴西变点检测结果对比

由图 5.1 可以看到, 自变点检测起始往后的 180 天内, 美国和巴西的疫情发展趋势相似, 均呈现不断扩张的特点。二者检测到的变点分别是 9 个和 7 个。巴西变点分布相对更加均匀, 大约每隔 15 天就会出现一个变点; 美国在 4-5 月变点出现频率较高。两个国家开始采取居家隔离等限制措施的时间分别

为 3 月 13 日和 3 月 24 日，也是各自疫情出现后的第 53 天和第 27 天，由于美国采取有效措施的时间落后，导致前期疫情扩张速度比巴西更快。直到历经 3 个半月左右，出现转折，即巴西的累计确诊病例数超过美国且持续约 1 个月，出现该现象的主要原因是巴西从 6 月起开始逐步解除限制，居民防疫心态放松，导致疫情快速反弹。自此之后，美国累计确诊病例持续高于巴西，二者的疫情发展状况均未见良好转变，分别是全球累计确诊病例第一和第三的国家。

(2) 英国与俄罗斯

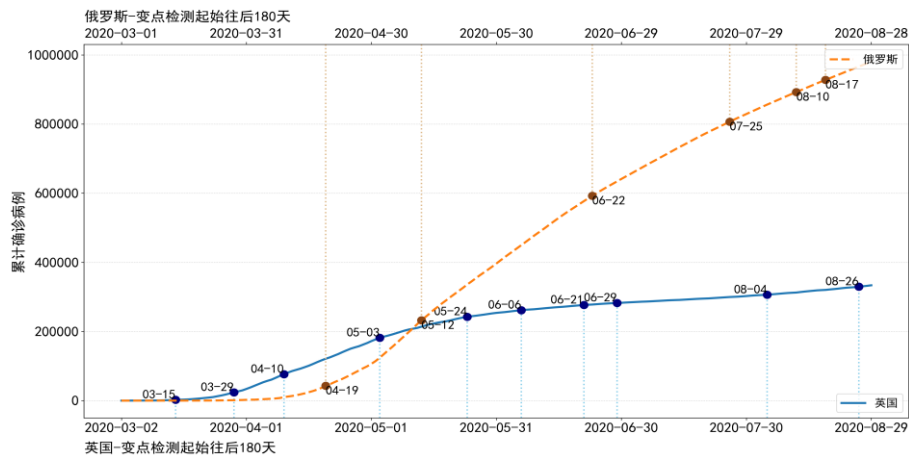


图 5.2 英国与俄罗斯变点检测结果对比

由图 5.2 可以看到，自变点检测起始往后的 180 天内，英国疫情发展较稳定，俄罗斯疫情初期状态良好，后期急速扩张。二者检测到的变点分别是 10 个和 6 个。英国变点分布较为均匀；俄罗斯首个变点出现较晚，且变点之间间隔相对更长。两个国家疫情出现时间仅相差一天，开始采取居家隔离等限制措施的时间分别为 3 月 23 日和 3 月 27 日，也是各自疫情出现后的第 51 天和第 56 天，但疫情发展趋势大相径庭，主要原因是俄罗斯在初期对欧洲意大利等严重疫情国家人员输入的防控疏漏，导致确诊病例在 4 月之后一路攀升，在 5 月上旬超过英国，而且于 5 月 12 日开始逐步解除限制，使得之后二者的差距越来越大。

(3) 印度与印度尼西亚

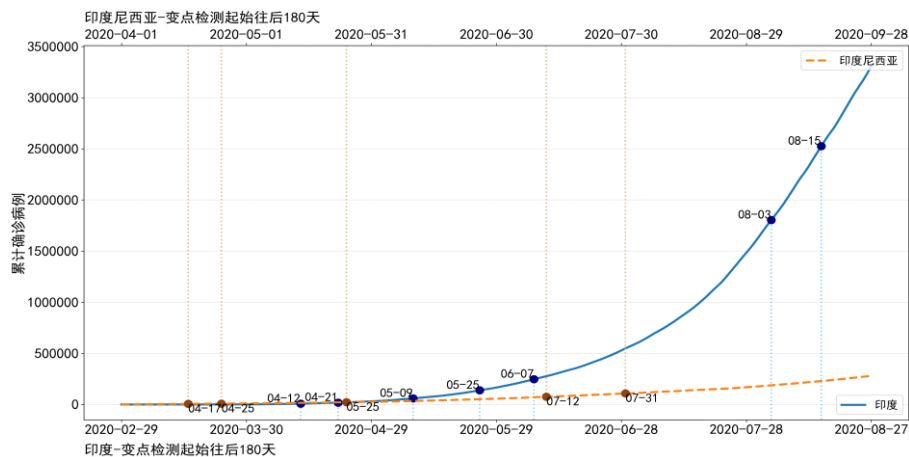


图 5.3 印度与印度尼西亚变点检测结果对比

由图 5.3 可以看到,自变点检测起始往后的 180 天内,印度和印度尼西亚在前期疫情都控制得比较好,但在约 3 个月之后,印度确诊病例开始呈现持续增长的发展特点,而印度尼西亚仍保持稳定。二者检测到的变点分别是 7 个和 5 个。印度变点出现较晚,且在 4-6 月出现频率较高;印度尼西亚变点分布相对分散。两个国家开始采取居家隔离等限制措施的时间分别为 3 月 25 日和 3 月 31 日,也是各自疫情出现后的第 55 天和第 29 天。虽然二者在前期疫情发展都比较温和,但由于印度尼西亚采取措施要远早于印度,且印度人口远高于印度尼西亚,是世界人口第二大国,导致印度在后期疫情开始猛烈爆发;此外,印度尼西亚卫生体系相对落后,存在感染人数统计疏漏的可能性。

(4) 菲律宾与中国

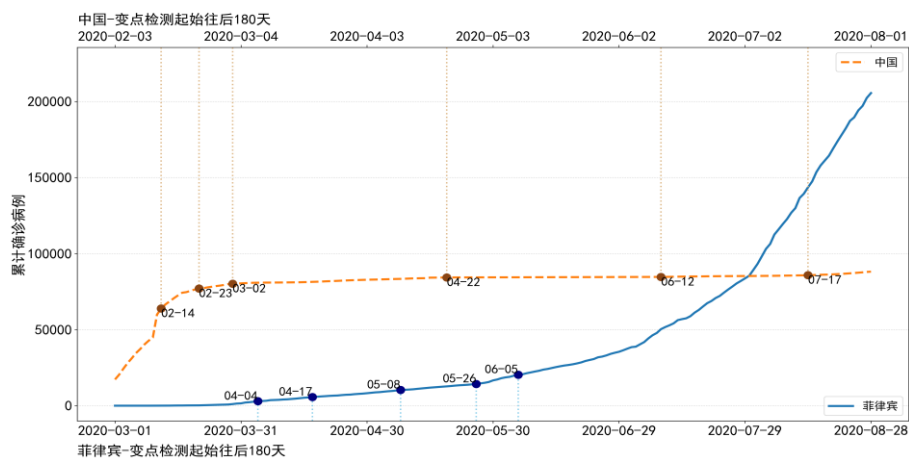


图 5.4 菲律宾与中国变点检测结果对比

由图 5.4 可以看到,自变点检测起始往后的 180 天内,菲律宾和中国的整体确诊病例都较少,二者疫情发展趋势恰好相反,菲律宾是先慢后快,中国是先快后慢。二者检测到的变点分别是 5 个和 6 个。菲律宾变点出现较晚,且集中分布在 4-6 月;中国变点集中分布在疫情开始的爆发期。两个国家开始采取居家隔离等限制措施的时间分别为 3 月 15 日和 1 月 23 日,也是各自疫情出现后的第 44 天和第 19 天。菲律宾在疫情初期严格的防控措施使得其发展稳定,但在 5 月中旬开始下调隔离等级之后疫情出现明显反弹,扩张迅速,之后超过中国的累计确诊病例数,由于中国采取防疫措施的时效性之高,使得疫情在两个月左右的时间内得到控制,此后一路维持稳定。

5.1.2 各区域确诊病例最高的国家对比

四大区域确诊病例最高的国家分别是美国、英国、印度和菲律宾,图 5.5 是变点检测起始往后 180 天内四个国家的疫情发展及变点检测结果对比。

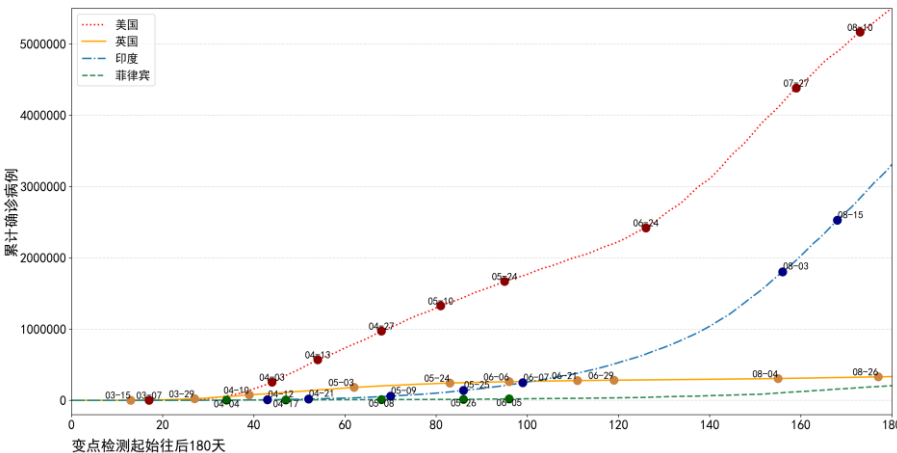


图 5.5 各区域确诊病例最高的国家变点检测结果对比

从疫情发展的角度对比来看,英国、菲律宾的疫情控制效果良好且发展较稳定,印度在前 80 天内发展平稳,之后以较快的速度持续扩张,而美国则是在 30 天之后持续以相对平均的速度在扩张,且该扩张规模远超过印度。

从变点的角度对比来看,四个国家的共同点在于变点相对更集中分布在第 30 至 120 天内,其中菲律宾更是全部落于该时间段内,说明在这一阶段受政策、措施等因素影响,疫情变化转折较多;不同点则主要在于国家之间不同

的防疫政策直接关系到疫情走向的不同。四个国家开始采取居家隔离、社交距离限制等有力措施的时间分别是变点检测起始位置之后的第 23 天、第 21 天、第 25 天和第 14 天, 经历疫情首轮爆发之后, 开始逐步解封的时间分别是第 72 天、第 69 天、第 100 天、第 75 天, 其中美国与英国时间相近, 但疫情发展截然相反, 重要原因在于美国急于重启经济, 选择在疫情并未得到真正缓和之时就开始解封, 确诊病例仍持续快速增长; 印度是在这一阶段采取限制措施时间最长的国家, 防控效果良好, 但在第 100 至 180 天之间的曲线趋势与美国相似, 且变点仅有 2 个, 说明反弹的势头没有得到及时控制; 菲律宾采取限制措施时间最早, 且疫情发展最佳, 但其在第 100 至 180 天之间未检测到变点, 且曲线开始缓慢倾斜, 接近英国, 说明该阶段疫情出现反弹, 但并未得到有效控制, 确诊病例稳定缓慢增长。

5.1.3 疫情发生最早与最晚国家对比

在本文研究的 8 个国家中, 新冠肺炎疫情发生最早与最晚的国家分别是中国与印度尼西亚, 时间分别是 2020 年 1 月 4 日和 3 月 2 日, 相差 58 天, 图 5.6 是变点检测起始往后 180 天内二者的变点检测结果对比。

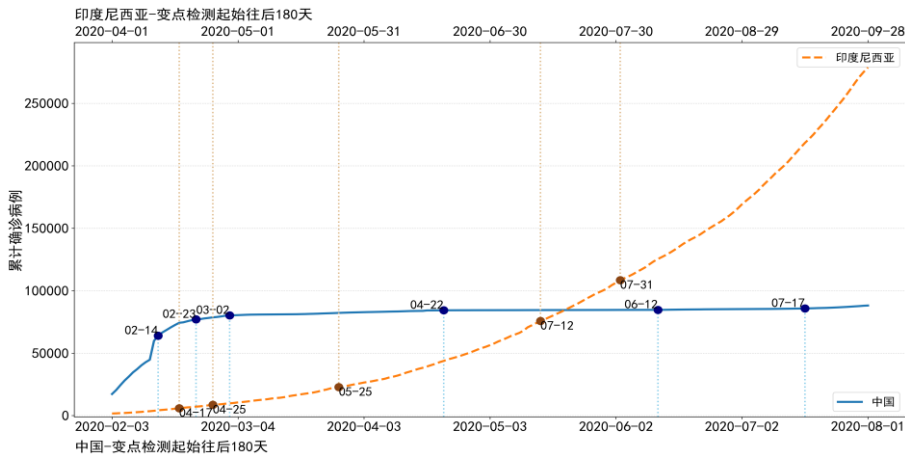


图 5.6 中国与印度尼西亚变点检测结果对比

由图 5.6 可以看到, 自变点检测起始往后的 180 天内, 中国疫情发展特点是先快后慢, 印度尼西亚则是先慢后快。因中国采取居家隔离等限制措施的时间很早, 是疫情出现后的第 19 天, 所以之后才能得到有效的稳定和控制;

而印度尼西亚借鉴相关经验，在疫情出现后的第 29 天就及时采取严格防疫措施，才会呈现前期的稳定态势。但在之后印度尼西亚确诊病例一路上升，超过中国，说明受医疗卫生体系落后等诸多因素影响，印度尼西亚的疫情发展逐渐恶化。二者检测到的变点分别是 6 个和 5 个，分布相似，均集中落在前 30 天内，之后分布稀疏，变点之间时间间隔均在 15 天以上。

5.1.4 人口相近的国家对比

根据世界银行(<https://data.worldbank.org/country>)截至 2019 年末的世界各国人口数据统计，中国(13.98 亿)、印度(13.66 亿)分别位列第一、第二。人口是衡量一个国家新冠肺炎严重程度的重要因素，本文将对变点检测起始往后的半年（180 天）内中国和印度的疫情发展和变点，对比结果见图 5.7。

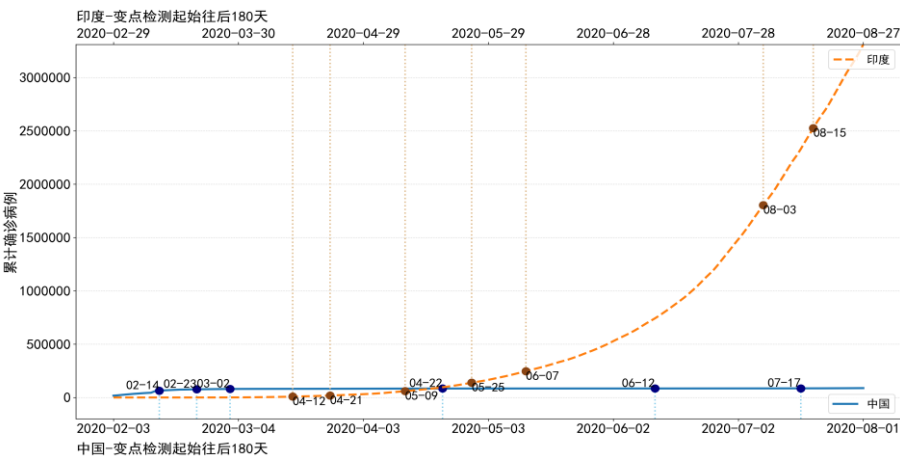


图 5.7 中国与印度变点检测结果对比

由图 5.7 可以看到，自变点检测起始往后的 180 天内，中国疫情发展特点是先快后慢，印度则是先慢后快，且后期的“快”远超过中国初期爆发的速度。二者检测到的变点分别是 6 个和 7 个。总体来看，二者的变点分布特点均是先密集后稀疏，但其中又存在巨大的差异。其一，中国的变点集中分布在疫情开始的爆发期，印度的变点集中分布在 4-6 月，也正是在这一阶段印度累计确诊病例超过中国，且增长速度持续加快，曲线逐渐趋向倾斜；其二，二者在第 120-180 天内的变点均仅有 2 个，但疫情曲线截然相反，中国疫情发展稳定，而印度则是加速扩张，在第 180 天累计确诊病例达到 3310234 例，是中国此

时 88122 例的约 37.56 倍。由此可见,人口相近的两个国家疫情发展趋势大相径庭,背后关联诸多因素,比如国家自身的医疗卫生体系发达程度、疫情防控政策措施的效力、公众的防疫意愿与态度等等。

5.2 国家新冠肺炎平稳时间子序列的拟合

由于序列中相邻变点之间的该段子序列是平稳的,所以为进一步验证 4.3 中的变点检测结果,本文将选取某段累计确诊病例日增长率平稳子序列进行模型拟合,时间范围是最后一个变点位置至 2021 年 11 月 1 日,选定的模型是 $ARMA(p, q)$, 利用 Python 绘制相应的 ACF、PACF 图以确定将要拟合模型的阶数 p 、 q , 对该段序列进行拟合并度量拟合效果。

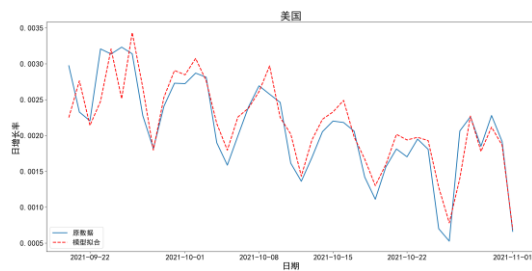


图 5.8 美国日增长率拟合



图 5.9 巴西日增长率拟合

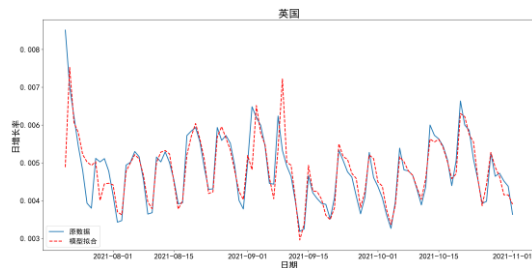


图 5.10 英国日增长率拟合



图 5.11 俄罗斯日增长率拟合

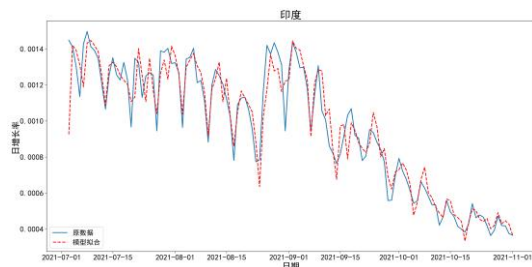


图 5.12 印度日增长率拟合



图 5.13 印度尼西亚日增长率拟合

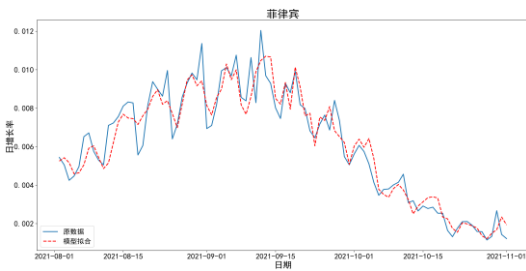


图 5.14 菲律宾日增长率拟合



图 5.15 中国日增长率拟合

图 5.8 至 5.15 为 8 个国家累计确诊病例日增长率平稳子序列的模型拟合结果，蓝色实线为原数据，红色虚线为相应模型拟合结果。依据自相关函数 ACF 和偏自相关函数 PACF 图像确定最优阶数，所拟合的模型及拟合优度具体见表 5.1。其中俄罗斯、印度、印度尼西亚、菲律宾和中国日增长率序列的模型拟合结果 ACF 具备共同特征，即 ACF 明显拖尾且 PACF 截尾，拟合优度相对更高，尤以俄罗斯、印度、菲律宾效果最佳，拟合优度均超过 0.9；美国、巴西和英国日增长率序列的 ACF、PACF 均存在截尾现象，拟合效果相对较差。

表 5.1 模型拟合结果

国家	模型	拟合优度
美国	ARMA(6,1)	0.7910
巴西	ARMA(5,1)	0.3991
英国	ARMA(9,4)	0.6533
俄罗斯	ARMA(1,0)	0.9481
印度	ARMA(8,0)	0.9157
印度尼西亚	ARMA(10,0)	0.8723
菲律宾	ARMA(11,0)	0.9206
中国	ARMA(4,0)	0.6786

6. 总结与展望

6.1 研究结论

本文在已有新冠肺炎研究成果的基础上,将“滑动窗口”的理念与变点统计分析方法相结合,对全球 8 个新冠肺炎疫情典型国家的累计确诊病例日增长率序列进行变点检验,并查阅相关疫情防控政策与措施,进行变点原因分析、多维度国家间对比以及对变点检测得到的平稳子序列进行拟合,最终的结论主要包含以下几个方面:

(1) 基于本文构建的简单时间序列模型,分别采取 KS 检验和 SN-based 检验进行变点检测模拟实验,最终发现二者的变点检测结果相近,但 SN-based 检验效果比 KS 检验更加稳定。因此选择 SN-based 检验用于新冠肺炎累计确诊病例日增长率序列的变点统计分析。

(2) 美国、巴西、英国、俄罗斯、印度、印度尼西亚、菲律宾、中国的累计确诊病例日增长率序列检测到的变点个数分别是 22、12、22、21、22、15、16、19。经国家相关防疫政策和措施的比对发现,变点所在位置与政策措施的实施时间紧密相关,且普遍存在一定的滞后效应,即在政策措施实施时间之后的一段时间确诊病例增长率才会出现明显变化,从而出现变点。

(3) 从累计确诊病例曲线整体来看,8 个国家疫情发展状况各自呈现不同特点,可划分为三种类型:长期平稳型(中国),特点是初期爆发猛烈,但很快得到控制且长期处于稳定发展的状态;阶段波动型(英国、印度、印度尼西亚、菲律宾),特点是存在 2 至 4 个明显的波动起伏,疫情反弹恶化与得到控制的阶段容易区分;持续扩张型(美国、巴西、俄罗斯),特点是疫情曲线持续向上,几乎不存在明显的平稳时间段。

(4) 经国家之间不同维度的均值变点检测结果及疫情发展形势对比发现,自变点检测起始位置起 180 天内的疫情走向与国家各自首次发布居家隔离、社交距离限制等严格防疫措施的时间关联性较强。比如:美国、巴西采取措施

较晚,疫情发展形势不容乐观,而中国、印度尼西亚采取措施较早,疫情相对来说更加稳定。

(5) 通过对各个国家的累计确诊病例日增长率序列中最后一个变点位置至 2021 年 11 月 1 日该段平稳子序列的 ARMA 模型拟合,得到俄罗斯、印度、印度尼西亚、菲律宾所拟合模型效果较好,拟合优度均超过 0.85,美国、巴西、英国、中国所拟合模型效果不佳,拟合优度均低于 0.8。

6.2 研究展望

本文对全球新冠肺炎疫情典型国家的累计确诊病例日增长率序列进行的变点统计分析是将变点理念应用至公共卫生领域的一项拓展性研究,存在创新与不足之处。

创新点其一是将滑动窗口的序列监测思想与变点检验方法相结合,以达到检测出日增长率序列中全部均值变点的目的;其二是基于已检测到的变点位置来探寻其存在的原因,途径是查询国家各自的疫情防控政策和措施,结合变点进行疫情发展的整体分析与原因浅析;其三是从四个不同维度出发来对比国家之间的变点检测结果和疫情发展趋势,可明显得出国家防控力度和效果的优劣性。不足之处其一是变点是否存在受到诸多因素的影响,本文仅考虑到最为关键的因素即防疫政策与措施,未能考虑到其他细小因素比如国家经济状况、人口流动等;其二是采用模型 1: $y_t = \mu_t + \varepsilon_t$ (μ_t 为累计确诊病例日增长率均值) 将日增长率作为主要研究对象进行分析,未考虑到生物种群增长特征,之后会采用更广泛的模型 2: $y_t = \beta + g\left(\frac{t}{n}\right) + \varepsilon_t$ 对确诊病例数进行深入探索,模型 2 是模型 1 的延伸形式,当模型 2 中的时间 t 变化幅度很小时,则函数 $g\left(\frac{t}{n}\right)$ 趋于 0,即为模型 1;其三是仅选取了 8 个典型国家进行分析,之后也可以尝试对更多国家进行变点检测,依据检测结果对所有国家进行聚类分析,以此将国家分为不同类型,更具有普适性。

变点分析仍然是统计学领域比较新的研究方向,近年来多个学者提出更加灵活崭新的检验方法,为推动该方向的研究做出重要贡献,而将变点与公共卫生相结合更是一个全新维度的理念,相信在未来定会有更大的突破。

参考文献

- [1] 陈希孺.变点统计分析简介[J].数理统计管理,1991,10(2):52-53.
- [2] Page,E.S.. Continuous Inspection Schemes[J].Biometrika,1954, (42):100-114.
- [3] Yao,Yi-Ching and Richard A.Davis. The Asymptotic Behavior of the Likelihood Ratio Statistic for Testing a Shift in Mean in a Sequence of Independent Normal Variates[J].The Indian Journal of Statistics, Series A,1986,48(3): 339-353.
- [4] D.L. Hawkins. Retrospective and sequential tests for a change in distribution based on Kolmogorov Smirnov-type statistics[J].Sequential Anal,1988, (7):23-51.
- [5] Ian B. MacNeill. Tests for Change of Parameter at Unknown Times and Distributions of Some Related Functionals on Brownian Motion[J].Ann. Statist,1974, 2(5):950-962.
- [6] Brown,R.L.,J.Durbin and J.M. Evans. Techniques for Testing the Constancy of Regression Relationships over Time[J].Journal of the Royal Statistics Society,1975,37(2):149-163.
- [7] Inclan,Carla and George C. Tiao. Use of Cumulative Sums of Squares for Retrospective Detection of Changes in Variance[J].Journal of the American Statistical Association,1994,89:913-923.
- [8] Vera R.Eastwood. Some Nonparametric Methods for Changepoint Problems[J].The Canadian Journal of Statistics/La Revue Canadienne De Statistique,1993,21(2):209-222.
- [9] Csorgo,M. and Horvath,L.. Limit Theorems in Change-Point Analysis[M]. New York:Wiley & Sons,1997.
- [10] 杨喜寿,杨洪昌.气候时间序列变点的推断[J].大气科学,1996,(1):47-53.
- [11] 谭智平,缪柏其.分布变点模型的非参数检验和区间估计[J].数学年刊,2000,(22):617-626.
- [12] 谭智平,缪柏其.关于分布变点问题的非参数统计推断[J].中国科学技术大学学报,2000,30(3):270-277.

- [13] 熊立华,周芬,肖义,郭生练.水文时间序列变点分析的贝叶斯方法[J].水电能科学,2003,21(4):39-41.
- [14] 谭常春,赵林城,缪柏其.至多一个变点的 Γ 分布的统计推断及在金融中的应用[J].系统科学与数学,2007(1):2-10.
- [15] 李绍刚,杨少华.基于 GARCH 变点模型的上证收益率伪波动持续性研究[J].价格月刊,2010,(7):91-94.
- [16] Xiaofeng Shao,Xianyang Zhang.Testing for Change Points in Time Series[J]. Journal of the American Statistical Association,2010,105:1228-1240.
- [17] 数据资讯: 全球新冠肺炎及其疫苗研究[J].中国科学院院刊,2021,36(02): 233-237.
- [18] Talkhi Nasrin,Fatemi Narges Akhavan,Ataei Zahra,Nooghabi Mehdi Jabbari. Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A Comparison of Time Series Forecasting Methods[J]. Bio-medical Signal Processing and Control,2021,66:102494.
- [19] Saswat Singh,Chandreyee Chowdhury,Ayan Kumar Panja,Sarmistha Neogy. Time Series Analysis of COVID-19 Data to Study the Effect of Lockdown and Unlock in India[J]. Journal of The Institution of Engineers (India): Series B,2021.
- [20] Palmer William R,Davis Richard A,Zheng Tian. Count-Valued Time Series Models for COVID-19 Daily Death Dynamics[J]. Stat (International Statistical Institute),2021:e369.
- [21] Gyeongseung Han,Jeewon Han,Seungmin Han,Hyunkyung Jeong. Analyzing the Impact of COVID 19 on Global Trends and Predicting Future Cases[J]. American Journal of Theoretical and Applied Statistics,2021,10(1):1-8.
- [22] Sri Rahayu,Intan Zainafree,Aufiena Nur Merzistya,Tika Cahyani. Community Characteristics in COVID-19 Preventive Precautions[P]. Proceedings of the 5th International Seminar of Public Health and Education, ISPHE 2020, 22 July 2020, Universitas Negeri Semarang, Semarang, Indonesia,2020.
- [23] Zhang Junyi,Zhang Runsen,Ding Hongxiang,Li Shuangjin,Liu Rui,Ma Shuang,Zhai Baoxin,Kashima Saori,Hayashi Yoshitsugu. Effects of transport-related COVID-19 policy measures: A case study of six developed countries[J]. Transport Policy,2021,110:37-57.

- [24] 白璐,郭佩汶,范晋蓉.湖北省新冠肺炎确诊人数的建模与预测分析[J].检验检疫学刊,2020,30(02):10-12.
- [25] 邵年,钟敏,程晋,陈文斌.基于 FUDAN-CCDC 模型对新冠肺炎的建模和确诊人数的预测[J].数学建模及其应用,2020,9(01):29-32.
- [26] 胡海文.基于 LSTM 模型的新冠病情预测和影响因素分析[J].现代信息技术,2021,5(07):91-94.
- [27] 谢聪慧,吴世新,张晨,孙文涛,何海芳,裴韬,罗格平.基于谱系聚类的全球各国新冠疫情时间序列特征分析[J].地球信息科学学报,2021,23(02):236-245.
- [28] 付吉丽,黄立强,王玉文.基于新型冠状病毒传播的数学模型研究[J].哈尔滨师范大学自然科学学报,2020,36(05):18-23.
- [29] 武蓓蓓.新冠肺炎疫情信息的空间分析及可视化研究[J].测绘与空间地理信息,2021,44(10):20-23+28.
- [30] 郭文璇,房昊.人口流动、传染病蔓延应急管控与中国新冠疫情防控经验[J].中国经贸导刊(中),2021(10):47-52.
- [31] 李仲来,崔恒建,杨华,李小文.SARS 预测的 SI 模型和分段 SI 模型[J].遥感学报,2003(05):345-349.
- [32] Vokó Zoltán,Pitter János György. The effect of social distance measures on COVID-19 epidemics in Europe: an interrupted time series analysis[J]. GeroScience,2020,42(4):1075-1082.
- [33] Han Lin Shang,Ruofan Xu. Change point detection for COVID-19 excess deaths in Belgium[J]. Journal of Population Research,2021.
- [34] Coughlin Steven S.,Yiğiter Ayten,Xu Honyan,Berman Adam E.,Chen Ji-e. Early detection of change patterns in COVID-19 incidence and the implementation of public health policies: a multi-national study[J]. Public Health in Practice,2020,2.
- [35] Jiang F, Zhao Z, Shao X. Time series analysis of COVID-19 infection curve: A change-point perspective[J]. Journal of Econometrics. 2020.
- [36] 刘君,罗晓媛.删失截断下新冠疫情初期数据多变点 Bayes 估计[J].黑河学院学报,2021,12(09):181-182.
- [37] CsörgőM,Horvath L.Nonparametric tests for the changepoint problem[J]. J Statist.Plann.Inference,1987,(17):1-9.

- [38] D.L. Hawkins. Retrospective and sequential tests for a change in distribution based on Kolmogorov Smirnov-type statistics[J]. Sequential Anal, 1988, (7):23-51.
- [39] Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, Herman K. van Dijk. Econometric Methods with Applications in Business and Economics [M]. New York: Oxford University Press, 2004:56.
- [40] 田腾, 石茂林, 宋学官, 马跃, 冯翔宇. 基于滑动窗口的时间序列异常检测方法[J]. 仪表技术与传感器, 2021(07):112-116.

附录

附表 1 KS 检验临界值

α	
0.40	0.87
0.20	1.07
0.10	1.22
0.05	1.36
0.04	1.37
0.01	1.63

附表 2 SN-based 检验临界值

$1-\alpha\%$	
90%	29.6
95%	40.1
97.5%	52.2
99%	68.6
99.5%	84.6
99.9%	121.9

致谢

两年的硕士学习与研究生活即将在这个夏天划上句号，但对于我的人生却只是一个逗号，我将再次踏上新的征程。值此论文付梓之际，回顾过去生活点点滴滴，感触颇深。人生难得是欢聚，惟有别离多，所有故事都会有结局，无论如何都要感谢在这两年中我所遇到的可爱的老师和同学，所经历的一切美好事物。

在这次毕业论文的撰写过程中，非常感谢余睿老师对我的耐心指导与建议。从选择论文研究方向、查阅文献资料、撰写开题报告、中期检查答辩，一直到现在正式确定论文终稿，老师总能从繁忙的工作中抽出时间给予我多方面的建议，结合专业所学知识与实际生活，才能将理论更好地付诸于实践。正是因为有老师的悉心教导与同学的积极鼓励，才使我的论文完成的比较顺利，在此谨向我的导师以及所有给予我关心的老师与同学表达诚挚的感谢和祝福。

最后要感谢我的父母和挚友，两年来正是因为有他们一如既往的支持与鼓励，我才能以梦为马，砥砺前行，始终不忘期许认真完成学业，生活中的失意与难过也因为我拥有父母的关心这样宝贵的财富才得以一一化解。

大鹏一日同风起，扶摇直上九万里，我终将要迎接这毕业季的匆匆而来，未来的日子定会更加熠熠生辉。

陈牟瑶

2022 年 5 月



经世济民 孜孜以求

西南财经大学

SOUTHWESTERN UNIVERSITY OF FINANCE AND ECONOMICS

校址：四川成都温江柳台大道555号

电话：028-87092032 传真：028-87092632

邮编：611130

网址：<http://www.swufe.edu.cn>

Address: Liulin Campus (Main Campus): 555, Liutai Avenue,

Wenjiang District, Chengdu, Sichuan, P. R. China

Tel: 86-28-87092032 Fax: 86-28-87092632 Postcode: 611130