

赛区评阅编号（由赛区组委会填写）：

2020 高教社杯全国大学生数学建模竞赛

承 诺 书

我们仔细阅读了《全国大学生数学建模竞赛章程》和《全国大学生数学建模竞赛参赛规则》（2019 年修订稿，以下简称为“竞赛章程和参赛规则”，可从全国大学生数学建模竞赛网站下载）。

我们完全清楚，在竞赛开始后参赛队员不能以任何方式，包括电话、电子邮件、“贴吧”、QQ 群、微信群等，与队外的任何人（包括指导教师）交流、讨论与赛题有关的问题；无论主动参与讨论还是被动接收讨论信息都是严重违反竞赛纪律的行为。

我们完全清楚，抄袭别人的成果是违反竞赛章程和参赛规则的行为；如果引用别人的成果或资料（包括网上资料），必须按照规定的参考文献的表述方式列出，并在正文引用处予以标注。

我们以中国大学生名誉和诚信郑重承诺，严格遵守竞赛章程和参赛规则，以保证竞赛的公正、公平性。如有违反竞赛章程和参赛规则的行为，我们将受到严肃处理。

我们授权全国大学生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号（从 A/B/C/D/E 中选择一项填写）： A

我们的报名参赛队号（12 位数字全国统一编号）： 202010038016

参赛学校（完整的学校全称，不含院系名）： 南京大学

参赛队员 (打印并签名)：1. 陈万驰 181840033 数学系

2. 姜玉骅 181870080 工程管理学院

3. 余帅杰 181860077 计算机科学与技术系

指导教师或指导教师组负责人 (打印并签名)： 教练组

（指导教师签名意味着对参赛队的行为和论文的真实性负责）

日期： 2020 年 08 月 1 日

（请勿改动此页内容和格式。此承诺书打印签名后作为纸质论文的封面，注意电子版论文中不得出现此页。以上内容请仔细核对，如填写错误，论文可能被取消评奖资格。）

赛区评阅编号（由赛区组委会填写）：

2020 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅记录（可供赛区评阅时使用）：

评 阅 人						
备 注						

送全国评阅统一编号（由赛区组委会填写）：

全国评阅随机编号（由全国组委会填写）：

（请勿改动此页内容和格式。此编号专用页仅供赛区和全国评阅使用，参赛队打印后装订到纸质论文的第二页上。注意电子版论文中不得出现此页。）

基于修正 SEIR 模型的新冠肺炎疫情预测与评估

摘要

2020 年伊始,随着新型冠状病毒感染肺炎的迅速蔓延,全国笼罩在了疫情的阴影下,一时间几乎整个中国只有一个主题,就是如何共同抗击新型冠状病毒疫情(以下简称新冠疫情)。本文结合基于新冠病毒特征修改的 SEIR 传染病模型,对武汉市的疫情发展进行了模拟和对防控措施进行了评估,以及对印度疫情发展进行了模拟和预测。比较了单次检测和混样检测方法,给出了已知病毒携带率的条件下可以采取的检测次数最少的方案。

针对问题一,基于传染病状况以及武汉市政府采取的控制措施,我们建立了一个修正的 SEIR 模型,将人群分为易感者(S)、潜伏者(E)、感染者(I)、住院者(H)、移出者(R)、被隔离易感者(S_q)、被隔离潜伏者者(E_q)。在通过查阅参考文献以及经过适当调试之后,我们确定了各种人群之间相互转化的比率并建立了微分方程来表示转化过程。对微分方程求解之后我们得到了对武汉疫情发展的预测图像并将之与实际情况比对,最终我们得到了比较符合武汉实际情况的预测数据。

针对问题二,基于问题一所建立的模型,我们通过改变隔离率、接触率、治愈率的取值来模拟不同严格程度的隔离措施对疫情发展的影响。通过观察数据及图像,发现武汉诸多严格的措施对疫情防控产生了巨大的正面影响。这也说明了在我国疫情防控初期,全国医护人员支援武汉的必要性和其积极影响。

针对问题三,我们选取印度共和国作为此题的研究对象。印度是世界第二人口大国,庞大的人口基数和较不均衡的发展使得该国在疫情防控方面面临巨大的压力。通过对问题一建立的模型进行适当的修改和调试,利用官方发布的感染数据,我们绘制出了印度疫情发展的相关图像。通过观察预测数据和图像,我们发现如果不加强防控措施,印度的疫情还将呈指数型增长。

针对问题四,我们利用伯努利分布对混样检测方式进行建模,将单人所需检测次数的期望作为优化的目标函数。在给定病毒携带者比例的条件下,我们利用下降单纯形法(Downhill Simplex)算出了使得目标函数最小的混样检测样本数量。最终,我们绘制了目标函数最小值与人群病毒携带率的关系图像并得出结论:在预期病毒携带者大于 30% 的地区,应该采取单人检测的方法,在小于 30% 的地方,应该采取混样检测的方法。

关键字: 新冠肺炎疫情 SEIR 模型 混样检测

一、问题背景与重述

2019 年 12 月初，武汉爆发新型冠状病毒 (COVID-19) 疫情，很快蔓延到全国各地，并在世界各国大流行。时至当下，病毒仍然在世界许多国家肆虐，世界各国人民的正常生活、生命健康以及社会经济持续受到威胁。由此，对新型冠状病毒的传播建立数学模型就非常有意义。

题目要求解决以下四个问题：

1. 根据国家卫生健康委员会每日疫情通报数据，建立武汉市新型冠状病毒传播数学模型（需给出 2 月 12 日由于确诊标准改变而产生数据突变的合理方案），评价其合理性和实用性。
2. 说明 2020 年 1 月 23 号武汉封城、2020 年 1 月下旬武汉新建雷神山医院和火神山医院、2 月初建立了 20 家方舱医院、武汉市市民居家隔离、全国各地有近 4.2 万医务人员驰援武汉抗击疫情等一系列措施在控制疫情发展中发挥的作用，如果没有采取（或提早、延后采取），对疫情传播所产生的影响作出评估。
3. 根据世界卫生组织报告的疫情通报，选定某个国家和地区疫情数据，建立数学模型，评价其合理性和实用性
4. 给出面对给定人群中预期病毒携带者的比例，采用单次检测还是混样检测（混检分组方式），达到既可以筛选病毒携带者，同时检测次数最少的方案。

二、问题分析

2.1 问题一分析

疫情传播类问题，可将人群分为易感者 (Susceptible, S)、感染者 (Infected, I)、潜伏者 (Exposed, E) 和移除人群 (Removed, R)，即经典的 SEIR 模型。考虑到新冠肺炎的隔离措施，在 SEIR 模型的基础上添加隔离人群以适用实际情况。由于各个时间段疫情防护的严格程度不同，建立模型选取的参数不能一概而论，因此可以考虑划分时间段，不同时间段选取不同的模型参数。如果以 2 月 12 日作为一个时间段分割点，这也能解决 2 月 12 日由于确诊标准改变而产生数据突变的拟合问题。

2.2 问题二分析

各种防护措施影响的是模型参数的选取，如接触率、感染率、治愈率等。通过分析措施对参数的影响，将参数的变化映射到模型的变化，即能评估出该措施对疫情传播产生的影响。

2.3 问题三分析

问题三与问题一基本一致，考虑沿用问题一的模型，用于建立印度新冠肺炎疫情传播模型，对疫情发展进行了模拟和预测。

2.4 问题四分析

欲比较混样检测和单次检测方法，单次检测方法即每人检测次数为一次，因此问题转变为给定人群病毒携带率，混样检测平均每人检测次数是否少于 1。若是则考虑使用混样检测方法，否则使用单次检测方法。

混样检测，顾名思义就是将几个样本混合在一起进行一次核酸检测。某次需要对 N 个人进行检测，病毒的携带率为 p 。为了减少工作量，一次性将 k 个人的样本混合，如果混合样本为阴性，则说明这 k 个人都没有患病。如果混合样本为阳性，说明这 k 个人中至少有 1 个人的样本是阳性，那么接下来对这 k 个人的分别单独做检测。而单次检测简单直观，即对所有人单独做检测。考虑用基本概率论知识，比较混样检测和单次检测方法对应到平均每人的检测次数，选取更合理的方案。

三、模型假设

1. 假设人群中所有个体都有被感染的概率。
2. 假设被感染个体痊愈后，会产生抗体，不会再被感染。
3. 假设所有感染者是同质的，即病情的严重程度、死亡率相同。
4. 鉴于 1 月 23 日武汉市封城，假设武汉市总人口的不变。
5. 假设核酸检测准确率为 100%。
6. 假设混样检测的结果不会因混样而改变，即若至少有一人感染，混样检测呈阳性；若无人感染，混样检测呈阴性。

四、符号说明

表 1 符号说明表

参数	定义	单位
W	负重上限	千克
M	初始资金	元
t	天数	天
T	总天数	天
m_w	每箱水的质量	千克/箱
m_f	每箱食物的质量	千克/箱
p_w	水的基准价格	元/箱
p_f	食物的基准价格	元/箱
n_{sw}	晴朗天气下水的基准消耗量	箱
n_{hw}	高温天气下水的基准消耗量	箱
n_{ow}	沙暴天气下水的基准消耗量	箱
n_{sf}	晴朗天气下食物的基准消耗量	箱
n_{hf}	高温天气下食物的基准消耗量	箱
n_{of}	沙暴天气下食物的基准消耗量	箱
n	玩家数	人
P_{it}	第 t 天开始时玩家 i 所处的位置	/
W_{it}	第 t 天开始时玩家 i 剩余的水	箱
F_{it}	第 t 天开始时玩家 i 剩余的食物	箱
Q_{it}	第 t 天开始时玩家 i 剩余的资金	元
S_{it}	第 t 天玩家 i 所处的地点特征	/
We_t	第 t 天的天气	/

五、模型建立与分析

5.1 游戏模型的建立

我们首先将该游戏利用数学语言加以描述。显然该局游戏的最终目的是使玩家到达终点时的收益最大，即

$$\max Q_{30} + \frac{1}{2}p_w W_{30} + \frac{1}{2}p_f F_{30} \quad (1)$$

其中 Q_t, W_t, F_t 分别表示第 t 天时玩家所剩下的资金、水和食物量。如果玩家在第 30 天前到达终点，则其各个属性将会在未来几天视作不变，所以我们以第三十天为统一结束时间。该目标函数有如下约束：

$$Q_t = Q_{t-1} + 1000Mine_t - Shop_t[10ShopF_t + 20ShopW_t] \quad (2)$$

其中

$$Mine_t = \begin{cases} 0, & \text{如果第 } t \text{ 天不挖矿} \\ 1, & \text{如果第 } t \text{ 天挖矿} \end{cases}$$
$$Shop_t = \begin{cases} 0, & \text{如果第 } t \text{ 天不购物} \\ 1, & \text{如果第 } t \text{ 天购物} \end{cases}$$

即每天结束时的资金等于前一天的资金加上当天挖矿获得的 1000 元（如果挖矿的话），再减去在村庄购买食物和水花费的钱（如果购买的话）。其中 $ShopF_t$ 和 $ShopW_t$ 分别表示玩家在第 t 天购买的食物量和水量（如果购买的话）。

$$F_t = F_{t-1} - \Delta F_t + Shop_tShopF_t \quad (3)$$

即每天结束时的食物量等于前一天的食物量减去当天的食物消耗量再加上在村庄购买的食物量（如果购买的话）。

$$W_t = W_{t-1} - \Delta W_t + Shop_tShopW_t \quad (4)$$

即每天结束时的水量等于前一天的水量减去当天的水消耗量再加上在村庄购买的水量（如果购买的话）。

$$P_t = P_{t-1}(1 - Move_t) + \bar{P}_{t-1}Move_t \quad (5)$$

其中 P_t 表示第 t 天的位置， \bar{P}_t 表示第 $t+1$ 天可以到达的几个相邻区域之一，

$$Move_t = \begin{cases} 1, & \text{如果第 } t \text{ 天移动} \\ 0, & \text{如果第 } t \text{ 天不移动（包括挖矿、停留）} \end{cases}$$

$$Shop_t \leq If_0[(P_t - C_1)(P_t - C_2) \cdots (P_t - C_n)] \quad (6)$$

$$Mine_t \leq If_0[(P_{t-1} - K_1)(P_{t-1} - K_2) \cdots (P_{t-1} - K_m)] \quad (7)$$

其中 C_1, C_2, \dots, C_n 表示 n 个村庄的位置， K_1, K_2, \dots, K_m 表示 m 个矿山的位置，

$$If_0(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases}$$

ΔF_t 和 ΔW_t 分别表示第 t 天所需要消耗的食物量和水量，由 Lagrange 插值算法：

$$L(x) = \sum_{i=1}^n f(x_i) \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}$$

可得

$$\Delta F_t = Move_t(-5Weather_t^2 + 3Weather_t + 14) + Mine_t(-7.5Weather_t^2 + 4.5Weather_t + 21) \quad (8)$$

$$\Delta W_t = Move_t(-11Weather_t^2 + 17Weather_t + 10) + Mine_t(-16.5Weather_t^2 + 25.5Weather_t + 15) \quad (9)$$

其中

$$Weather_t = \begin{cases} 0, & \text{晴朗天气} \\ 1, & \text{高温天气} \\ 2, & \text{沙暴天气} \end{cases}$$

表示第 t 天的天气情况。

可以看出，问题一的核心关键在于玩家是否前往矿山挖矿、在挖矿之后是否需要前往村庄补给物资以及是否可以在矿山和村庄之间往返。我们将问题作如下简化：

选取图中起点、终点、村庄和矿山作为特殊点，利用 Dijkstra 算法计算出每两个特殊点之间的最短路径（即不考虑沙暴天气的影响下，所需最少的天数），如图1：

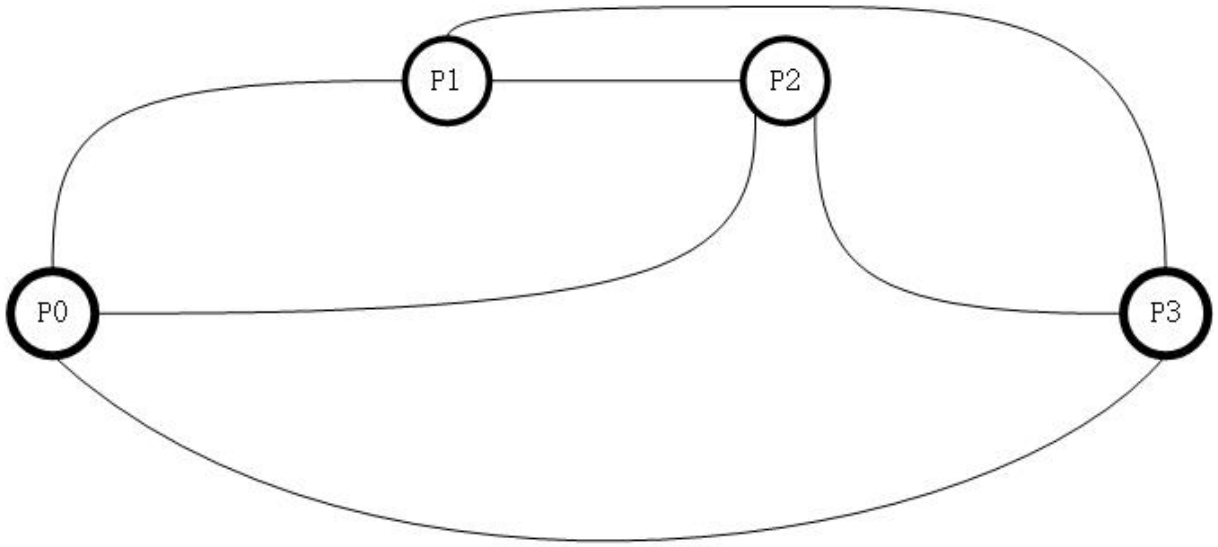


图 1 简化后的 Map1

该图是一个完全图（即图中每两点之间都有一条边），且每条边上的数字代表着两个特殊点之间的最短路径（受天气影响可能会有变化）。其中： P_0 代表起点， P_1 代表村庄， P_2 代表矿山， P_3 代表终点。对于问题一，因为天气是预先给定的，整局游戏没有随机因素，所以我们假定玩家在游戏途中始终向着某个特殊点（终点、村庄、矿山）前进并且选择最合适的路径，而非漫无目的地随机移动，即游戏旅程由几条有向线段叠加而成：

$$\vec{P} = \sum_{k=1}^n \overrightarrow{P_{i_k} P_{i_{k+1}}} \quad (10)$$

其中 $i_k \leq 3$, $k = 1, 2, \dots, n$.

参考文献

- [1] Xia W, Sanyi T, Yong C, et al. When will be the resumption of work in Wuhan and its surrounding areas during COVID-19 epidemic? A data-driven network modeling analysis[J]. SCIENTIA SINICA Mathematica, 2020.
- [2] 曹盛力, 冯沛华, 时朋朋. 修正 SEIR 传染病动力学模型应用于湖北省 2019 冠状病毒病 (COVID-19) 疫情预测和评估 [J]. 浙江大学学报 (医学版), 2020, 49(1): 0-0.

- [3] Lakdawalla D, Keeler E, Goldman D, et al. Getting Americans back to work (and school) with pooled testing[J]. USC Schaeffer Center White Paper, 2020.
- [4] Sarkar K, Khajanchi S, Nieto J J. Modeling and forecasting the COVID-19 pandemic in India[J]. Chaos, Solitons & Fractals, 2020: 110049.