五一数学建模竞赛

承 诺 书

我们仔细阅读了五一数学建模竞赛的竞赛规则。

我们完全明白，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、网上咨询等）与本队以外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的, 如果引用别人的成果或其它公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们愿意承担由此引起的一切后果。

我们授权五一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）。

参赛题号（从 A/B/C 中选择一项填写）： B

参赛队号： T21257513630720

参赛组别（研究生、本科、专科、高中）： 本科

所属学校（学校全称）： 四川大学锦城学院

参赛队员： 队员 1 姓名： 卓楠

队员 2 姓名： 李柯

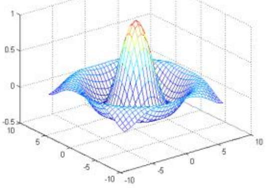
队员 3 姓名：

联系方式： Email： crow_9@163.com 联系电话： 13730620193

日期： 2021 年 5 月 3 日

（除本页外不允许出现学校及个人信息）

五 一 数 学 建 模 竞 赛



题 目：消防救援问题

关键词：伪线性规划模型；灰色预测；平均移动法；概率分布

摘　要：

我国的城市消防救援任务随着城市空间复杂性的上升也越来越多样复杂。现有某地共 15 个区域及相关位置距离信息、各个地区人口数量、占地面积，以及各个地区自 2016 年到 2020 年的消防出警记录等信息。需要分别从消防值班预警的人数安排、消防出警次数预测、各类事件发生次数与月份关系、各类事件发生密度与空间相关性、各类事件发生密度与人口密度关系以及建立消防站这几个问题进行建立模型并结合实际情况进行分析。

问题 1、从 2016 年到 2020 年每年的 2 月 1 日、5 月 1 日、8 月 1 日、11 月 1 日的消防救援出警数据分别导出，对不同时间段的消防救援出警的次数进行统计累加统计，建立伪线性规划模型求解。

问题 2、从 2016 年到 2019 年 4 年间各个月份的接警次数统计作为预测模型的训练集，把 2020 年各个月份的接警次数作为预测模型的测试集，采用灰色预测的方法进行建模，基于 GM（1,1）模型的消防救援接警次数预测(每次预测都只对于同一月份的数据进行处理)。

问题 3、5 年间所有的事件数据按照月份进行分类和统计，创建一个矩阵，利用循环判断数据并使其用于保存事件的类型和月份对应的出警次数总和，然后使用平均移动法对其进行建模。

问题 4、直接用各个区域各个事件的周发生率与该地的占地面积之间的比率结果，即各个区域各个事件的事件密度建立图表模型直观的对数据进行分析求解其相关性。

问题 5、采用概率分布来分析事件密度与人口密度之间的相关性。

附件中对该地消防救援出警的数据只有 2016 年到 2020 年期间一共五年的数据，收集数据量较少，使得结果具的准确性和可信度不高。

在解题过程中，可以很明显发现该地 2020 年有几个月的消防救援出警次数较前四年的次数偏少，结合其数据与实际现实情况分析，应该是该地受到新冠病毒疫情等不确定因素影响导致部分消防事件减少，最终影响结果具有一定偏差。

一、问题的提出

随着我国经济的高速发展，城市空间环境复杂性急剧上升，各种事故灾害频发，安全风险不断增大，消防救援队承担的任务也呈现多样化、复杂化的趋势。对于每一起出警事件，消防救援队都会对其进行详细的记录。

某地有 15 个区域，分别用 A、B、C…表示，各区域位置关系及距离如图 1 所示，各区域的人口及面积见附件 1，该地消防救援队出警数据见附件 2。

请依据该地的消防出警数据，建立数学模型，完成以下问题：

问题 1：将每天分为三个时间段（0:00-8:00 为时段Ⅰ，8:00-16:00 为时段Ⅱ，16:00-24:00 为时段Ⅲ），每个时间段安排不少于 5 人值班。假设消防队每天有 30 人可安排值班，请根据附件数据，建立数学模型确定消防队在每年 2 月、5 月、8 月、11 月中第一天的三个时间段各应安排多少人值班。

问题 2：以该地 2016 年 1 月 1 日至 2019 年 12 月 31 日的数据为基础，以月份为单位，建立消防救援出警次数的预测模型；以 2020 年 1 月 1 日至 2020 年 12 月 31 日的数据作为模型的验证数据集，评价模型的准确性和稳定性，并对 2021 年各月份的消防救援出警次数进行预测，完成表 1。

问题 3：依据 7 种类别事件的发生时间，建立各类事件发生次数与月份关系的多种数学模型，以拟合度最优为评价标准，确定每类事件发生次数的最优模型。

问题 4：根据图 1，请建立数学模型，分析该地区 2016-2020 年各类事件密度在空间上的相关性，并且给出不同区域相关性最强的事件类别（事件密度指每周每平方公里内的事件发生次数）。

问题 5：依据附件 2，请建立数学模型，分析该地各类事件密度与人口密度之间的关系（人口密度指每平方公里内的人口数量）。

问题 6：目前该地有两个消防站，分别位于区域 J 和区域 N，请依据附件 1 和附件 2，综合考虑各种因素，建立数学模型，确定如果新建 1 个消防站，应该建在哪个区域？如果在 2021-2029 年每隔 3 年新建 1 个消防站，则应依次建在哪些区域？

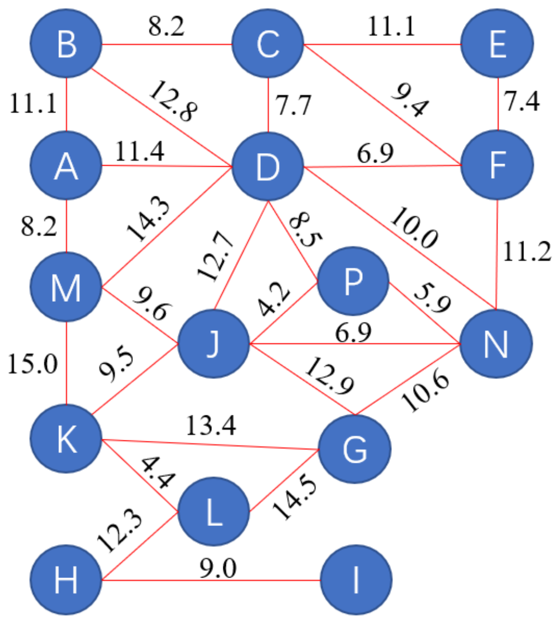


图 1　各区域之间的邻接关系及距离图（图中距离单位为 km）

表 1　问题 2 的结果

月份	预测值（次）
2021 年 1 月	75
2021 年 2 月	152
2021 年 3 月	118
2021 年 4 月	92
2021 年 5 月	115
2021 年 6 月	241
2021 年 7 月	178
2021 年 8 月	27
2021 年 9 月	35
2021 年 10 月	26
2021 年 11 月	10
2021 年 12 月	24

二、问题的分析

2.1 问题背景分析

我国的城市消防救援任务随着城市空间复杂性的上升也越来越多样复杂。现有某地共 15 个区域及相关位置距离信息、各个地区人口数量、占地面积，以及各个地区自 2016 年到 2020 年的消防出警记录等信息。需要分别从消防值班预警的人数安排、消防出警次数预测、各类事件发生次数与月份关系、各类事件发生密度与空间相关性、各类事件发生密度与人口密度关系以及建立消防站这几个问题进行建立模型并结合实际情况进行分析。

2.2 解决问题的分析及思路

问题 1：

想要确定消防队在每年在某一天里的三个时间段各自应该安排多少人值班，应该先将该地所有区域每年在该天的消防救援出警数据导出，根据题意将其按照三个时间段进行分组，确定出该地所有区域每年在该天的三个时间段里发生消防事件的次数，将三个时间段发生的消防事件次数比得出比例的比值，通过不同年份里同一天的比值结果结合题意中值班人数条件建立理论模型，进行值班人数的分配。

问题 2：

根据题意我们可以将 2016 年到 2020 年的消防救援出警的数据分为两个部分，分别是 2016 年到 2019 年的消防救援出警数据（建立预测数据模型）和 2020 年的消防救援出警数据（作为验证数据）两部分。首先我们可以将 2016 年到 2019 年的消防救援出警数据使用灰色预测的方法建立预测模型，通过这个模型对 2020 年的消防救援出警次数进行预测，然后将预测结果与附件 2 中实际记录的 2020 年消防救援出警次数进行比较，并结合一定的实际情况对该预测模型进行准确性和稳定性的分析与评价。

问题 3：

将 5 年间所有的事件数据按照月份进行分类和统计，然后利用工具绘制相关的图形，分析各类事件的发生次数与月份之间的关系找出最佳答案。

问题 4：

这道问题主要抓住各个区域的不同事件发生的次数、各个区域的面积以及通过计算求得各个区域的各个事件的事件密度，分析得出结果即可。分析各类事件在空

间上的相关性主要思路：首先可以先将各类事件在每个区域的事件密度求出，并利用事件密度求出不同区域的事件密度关于区域面积的比率，分析得出不同区域相关性最强的事件类别

问题 5：

这道题与第 4 题有些许相似的地方，不过本题主要的解题突破口就是各个区域不同类型事件的发生密度、各个区域的人口密度，通过比对分析这不同地区、相同事件种类间事件发生密度与人口密度间的关系得出结果。

问题 6：

新建一个消防站需要考虑的因素主要有三个，分别是各个区域发生事件接警的概率、各个已经有两个消防站应避免冗余、消防站到其他各个区域接警的距离。

三、模型假设

3.1 模型假设 1：

假设附件 1 中各个区域人口数量、区域面积在该题解涉及时间线内均保持不变且数据记录准确。

3.2 模型假设 2：

假设附件 2 中接警日期、接警时间点准确以及事件所在区域、事件类别填写记录无误。

四、符号约定

符号	符号说明
i	时段符号Ⅰ、Ⅱ、Ⅲ
A	接警次数
X	值班人数
Y	向量
B	矩阵
α	区域符号
C	次数
T	时间
S	面积
Z	事件密度
L	所有事件密度总和与面积比值比率
P	人口数量
D	人口密度

五、模型建立与求解

5.1 问题 1 的求解

5.1.1 有效数据的提取

将附件 2 中从 2016 年到 2020 年每年的 2 月 1 日、5 月 1 日、8 月 1 日、11 月 1 日的消防救援出警数据分别导出，然后进行对不同时间段的消防救援出警的次数进行统计累加统计，可得出下表：

表 2　2016 年至 2020 年相应日期各个时段共接警次数

接警日期	时段Ⅰ	时段Ⅱ	时段Ⅲ	合计
2 月 1 日	0	8	7	15

5 月 1 日	2	8	7	17
8 月 1 日	1	3	4	8
11 月 1 日	0	3	1	4

5.1.2 模型的建立

由题意假设 i 为对应的时段符号Ⅰ、Ⅱ、Ⅲ，相应日期各个时段对应的接警次数为A_i(单位：次)，相应日期各个时段对应的值班人数为X_i(单位：人),则可利用统计出的接警次数的比例和题意求解：

建立伪线性规划模型求解：

$$\text{s.t.}\begin{cases} X_{\mathrm{I}} + X_{\mathrm{II}} + X_{\mathrm{III}} = 30 \\ X_{\mathrm{I}}, X_{\mathrm{II}}, X_{\mathrm{III}} \geq 5 \\ X_i,\ A_i \in N\ (i = \text{Ⅰ、Ⅱ、Ⅲ}) \end{cases} \quad (1)$$

值班人数与接警次数关系须满足：

$$X_{\mathrm{I}}:X_{\mathrm{II}}:X_{\mathrm{III}} = A_{\mathrm{I}}:A_{\mathrm{II}}:A_{\mathrm{III}} \quad (2)$$

若在比值结果中有任意X_i的值小于 5，即就将其值定为 5，再将剩余未安排值班的人数X_i与其对应的接警次数A_i按比例分配剩余的未安排的值班人数。

5.1.3 答案求解

（1）2 月 1 日 3 个时段安排值班人数：

根据表 2 求解得出的 2 月 1 日近 5 年三个时间段接警次数比为：

$$A_{\mathrm{I}}:A_{\mathrm{II}}:A_{\mathrm{III}} = 0:8:7 \quad (3)$$

根据该比例将总安排值班人数 30 人进行分配，但是结果得出时段Ⅰ安排的值班人数为 0，这显然不符合题意要求每个时段必需安排值班 5 人及以上，所以当遇到这种值班人数结果时应提前考虑到这项题意规定，正确合理的解答应先将值班人数小于 5 的时段分出，保证其时段值班在 5 人的条件下再将剩下的时段和未安排的值班人数进行比例分配，即：

$$\begin{cases} X_{\mathrm{I}} = 5 \\ X_{\mathrm{II}}:X_{\mathrm{III}} = A_{\mathrm{II}}:A_{\mathrm{III}} = 8:7 \\ X_{\mathrm{I}} + X_{\mathrm{II}} + X_{\mathrm{III}} = 30 \end{cases} \quad (4)$$

结果计算得出$X_{\mathrm{II}} = 13$、$X_{\mathrm{III}} = 12$，即在 2 月 1 日当天，时段Ⅰ应该安排 5 人值班，时段Ⅱ应该安排 13 人值班，时段Ⅲ应该安排 12 人值班。

同理按照该方法计算得出：

在 5 月 1 日当天，时段Ⅰ应该安排 5 人值班，时段Ⅱ应该安排 13 人值班，时段Ⅲ应该安排 12 人值班。

在 8 月 1 日当天，时段Ⅰ应该安排 5 人值班，时段Ⅱ应该安排 11 人值班，时段Ⅲ应该安排 14 人值班。

在 11 月 1 日当天，时段Ⅰ应该安排 5 人值班，时段Ⅱ应该安排 17 人值班，时段Ⅲ应该安排 8 人值班。

5.2 问题 2 的求解

5.2.1 数据的预处理

将附件 2 中从 2016 年到 2019 年 4 年间各个月份的接警次数统计作为预测模型

的训练集，把 2020 年各个月份的接警次数作为预测模型的测试集，可得到下表：

表 3 2016 年到 2020 年份各个月份的接警次数

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
2016 年	79	95	98	51	275	67	63	40	50	31	38	43
2017 年	58	28	44	14	148	87	23	42	49	33	65	68
2018 年	54	129	68	71	107	119	59	50	46	44	40	63
2019 年	67	76	73	44	138	146	72	30	41	26	25	36
2020 年	28	23	28	46	62	55	25	29	36	25	51	62

5.2.2 模型的建立

对于这道题的预测方法，我们选择采用灰色预测的方法进行建模，基于 GM(1,1)模型的消防救援接警次数预测(每次预测都只对于同一月份的数据进行处理，下列采用 2016 年到 2019 年 1 月的数据进行建立模型并预测 2020 年 1 月的消防救援接警次数作为示例)。

①原始数据列：

$$x^{(0)} = (x^{(0)}(2016)，x^{(0)}(2017)，\ \ldots，x^{(0)}(2019)) \qquad (5)$$

②选取常数 c 对原始数据列进行下平移交换：

$$y^{(0)}(k) = x^{(0)}(k) + c, k = 2016,2017,\ldots,2019 \qquad (6)$$

序列$y^{(0)}$级比：

$$\lambda_y(k) = \frac{y^{(0)}(k-1)}{y^{(0)}(k)} \in \left(e^{-\frac{2}{n+1}}，\ e^{\frac{2}{n+2}}\right)，k = 2016,2017,\ldots,2019，n = k - 2015 \qquad (7)$$

③对原始数据列做一次累加：

$$x^{(1)} = \left(x^{(0)}(2016), x^{(0)}(2016) + x^{(0)}(2017), \ldots, x^{(0)}(2016) + \cdots + x^{(0)}(2019)\right)$$
$$= \left(x^{(1)}(2016), x^{(1)}(2017), \ldots, x^{(1)}(2019)\right) \qquad (8)$$

④数列$x^{(1)}$的灰导数并建立模型：

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \qquad (9)$$

⑤构造数据向量 Y 和数据矩阵并就得估计值：

$$Y = \left[x^{(0)}(2017), x^{(0)}(2018), x^{(0)}(2019)\right] \qquad (10)$$

$$B = \begin{bmatrix} -Z^{(1)}(2017) & 1 \\ -Z^{(1)}(2018) & 1 \\ -Z^{(1)}(2019) & 1 \end{bmatrix} \qquad (11)$$

(其中:$Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k), k = 2017,2018,2019$)

由最小二乘法可分别求出 a、b 值：

$$\tilde{u}=\left(\tilde{a},\tilde{b}\right)^T=(B^TB)^{-1}B^TY \qquad (12)$$

（此处$\tilde{u}=\left(\tilde{a},\tilde{b}\right)^T$用记号“~”表示所求得的值是近似值，可与 u=$(a,b)^T$相区分。）

⑥由上述微分方程(9)可得出序列的预测值：

$$\tilde{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ab}+\frac{b}{a}, k=2015,2016,2017 \qquad (13)$$

累减还原：

$$\tilde{x}^{(0)}(k+1)=\tilde{x}^{(1)}(k+1)-\tilde{x}^{(1)}(k), k=2016,2017,\dots,2019 \qquad (14)$$

⑦模型检验：

残差序列：

$$\varepsilon(k)=x^{(0)}(k)-\tilde{x}^{(0)}(k), k=2016,2017,\dots,2019 \qquad (15)$$

$$\varepsilon^{(0)}=(\varepsilon(2016),\varepsilon(2017),\dots,\varepsilon(2019)) \qquad (16)$$

相对误差：

$$\Delta(k)=\frac{|\varepsilon(k)|}{x^{(0)}(k)} \qquad (17)$$

$$\Delta=(\Delta(2016),\Delta(2017),\dots,\Delta(2019)) \qquad (17)$$

运用上述式子(7)求级比偏差：

$$\rho(k)=1-\left(\frac{1-0.5a}{1+0.5a}\right)\lambda \qquad (18)$$

将各年份 1 月的消防救援接警次数分别对应并代入该模型，利用 MATLAB 求出模型的预测结果：2020 年 1 月该地区的消防救援接警次数约为 70 次。

5.2.2 模型的准确性和稳定性评价：

利用模型和 MATLAB 预测 2020 年其他月份可得出下表：

表 4　2020 年各个月份预测结果及检验

月份	预测值（次）	实际值（次）	残差	相对误差
2020 年 1 月	70	28	42	1.5
2020 年 2 月	122	23	99	4.304347826
2020 年 3 月	94	28	66	2.357142857
2020 年 4 月	71	46	25	0.543478261
2020 年 5 月	120	62	58	0.935483871
2020 年 6 月	188	55	133	2.418181818
2020 年 7 月	115	25	90	3.6
2020 年 8 月	31	29	2	0.068965517
2020 年 9 月	38	36	2	0.055555556
2020 年 10 月	29	25	4	0.16
2020 年 11 月	16	51	35	0.68627451
2020 年 12 月	32	62	30	0.483870968

结果分析：

从上表数据可得出预测值与实际值之间误差较大，尤其在 2020 年上半年中，数据的相对误差值普遍很大，下半年才有所好转，但是整体的准确性并不可观。结合

实际情况与所给出的数据分析，影响预测结果误差较大的因素有：1、本题所给的只有 2016 年到 2020 年 5 年的数据，数据基数过少导致预测的结果误差较大准确性不高。2、结合现实实际情况在 2020 期间各地均有受到新冠疫情影响，疫情的发展影响到社会一些人为的火灾发生概率，所以导致实际值与预测结果偏差较大。

综上该模型的准确性和稳定性不会太理想。

5.2.3 对 2021 年各月份预测：

利用表 3 的数据以及模型可分别预测求出 2021 年各个月份的消防救援接警次数并填入表 1，同下：

表 1 问题 2 的结果

月份	预测值（次）
2021 年 1 月	75
2021 年 2 月	152
2021 年 3 月	118
2021 年 4 月	92
2021 年 5 月	115
2021 年 6 月	241
2021 年 7 月	178
2021 年 8 月	27
2021 年 9 月	35
2021 年 10 月	26
2021 年 11 月	10
2021 年 12 月	24

5.3 问题 3 的求解

5.3.1 数据的预先处理

将附件 2 中的事件按照类别与月份时间的不同进行分类分别载入模型中，利用 MATLAB 进行基本的数据，依据该数据统计绘制出以下图片：

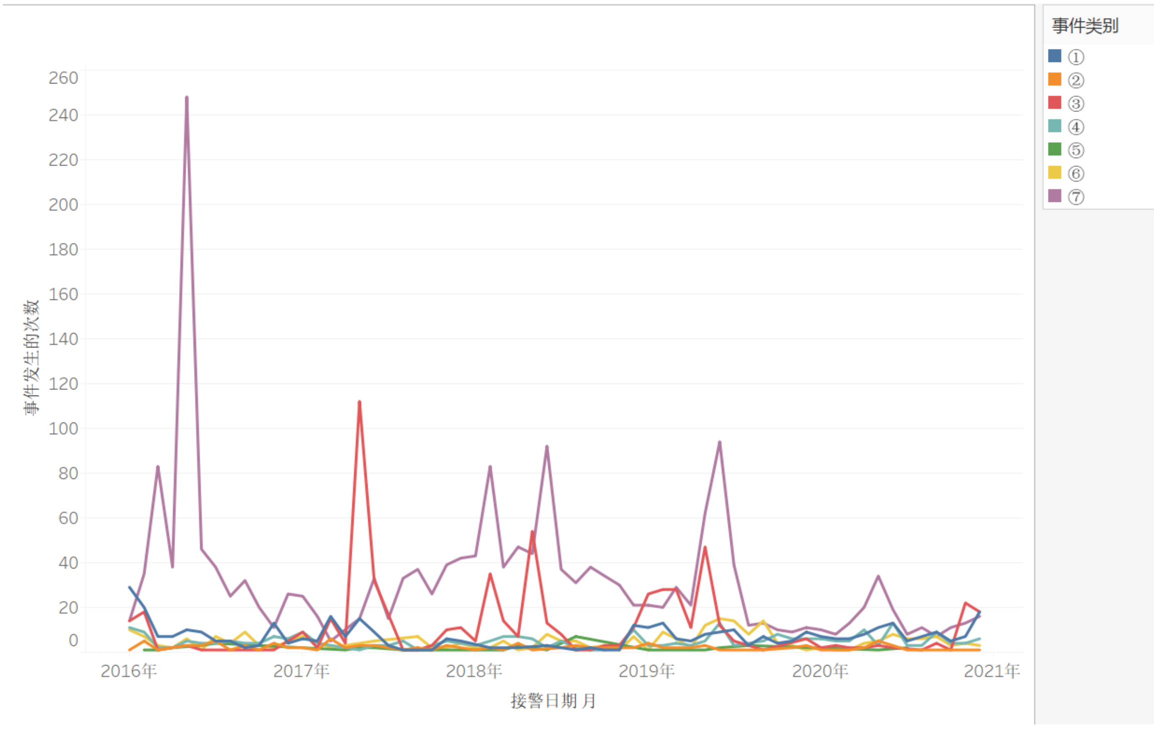


图 2 该地 5 年间每个月各类事件次数(单位：次)

5.3.2 模型的建立

创建一个矩阵，利用循环判断数据并使其用于保存事件的类型和月份对应的出警次数总和，然后使用平均移动法对其进行建模。

假设 Ft 为对下一期的预测值，n 为移动平均的时期个数，At-1,At-2,…,At-n 为前期实验值：

$$Ft = {(\sum_{i=1}^{n} At - i)}/{n} \qquad (19)$$

5.3.3 答案的求解

利用 MATLAB 计算求得平均移动法的平均标准差：86.870517

5.4 问题 4 的求解

5.4.1 解题方法：

直接用各个区域各个事件的周发生率与该地的占地面积之间的比率结果，即各个区域各个事件的事件密度建立图表模型直观的对数据进行分析求解其相关性。

5.4.2 模型的建立与求解

由题意给出的事件密度定义是每周每平方公里内的时间发生次数。假设 C(α)为一段时间某区域发生事件的次数（单位:次），α表示相对应的区域；T 表示发生该事件这些次数所在的这段时间(单位：周)；S(α)表示相对应区域的占地面积(单位；平方公里)；事件密度用 Z(α)表示：

$$Z(\alpha) = \frac{C(\alpha)}{TS(\alpha)} \qquad (19)$$

假设 b 表示①~⑦七种事件种类，且各类事件密度在空间上的相关性用其所有事件密度总和与面积比值比率表示：

$$L(\alpha) = \sum_{b=①}^{⑦} Z(\alpha) : S(\alpha) \qquad (20)$$

利用 MATLAB 进行计算并绘图表示如下：

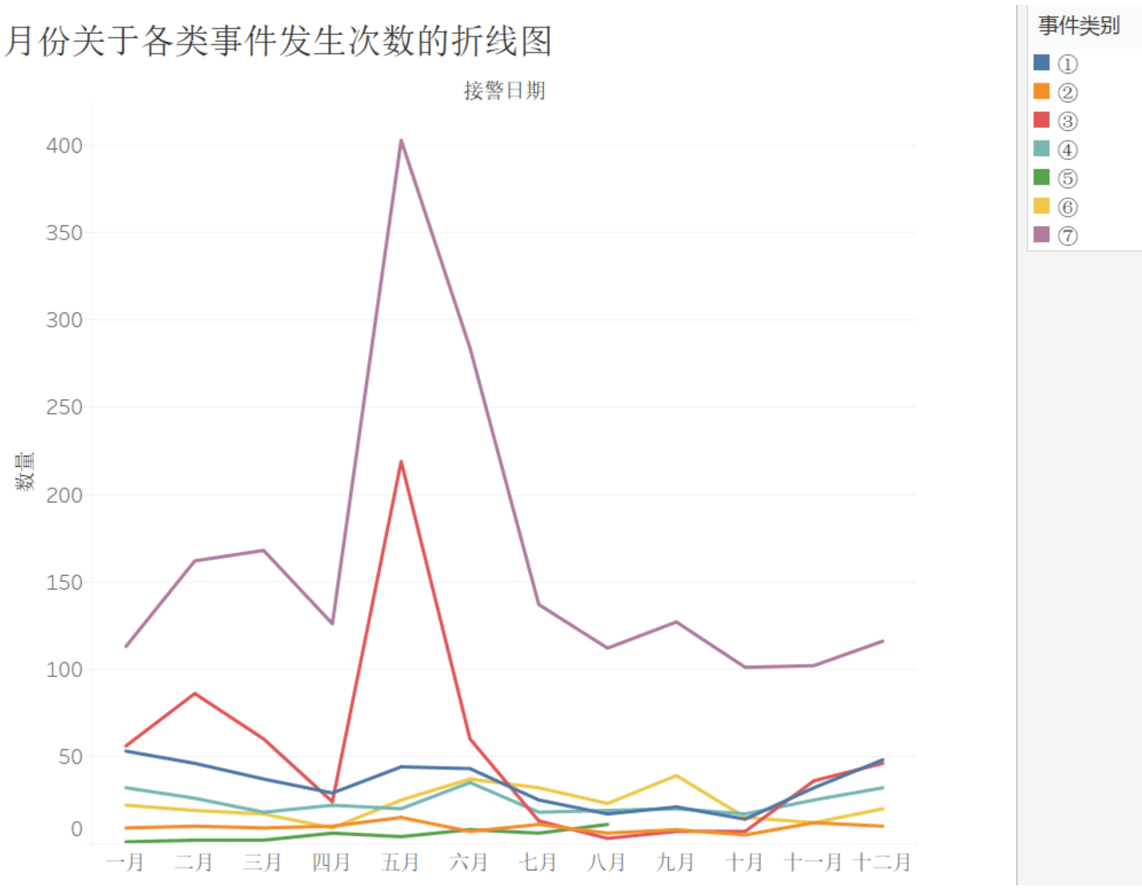


图 3　该地所有区域 5 年各个月份各类事件的平均发生次数(单位：次)

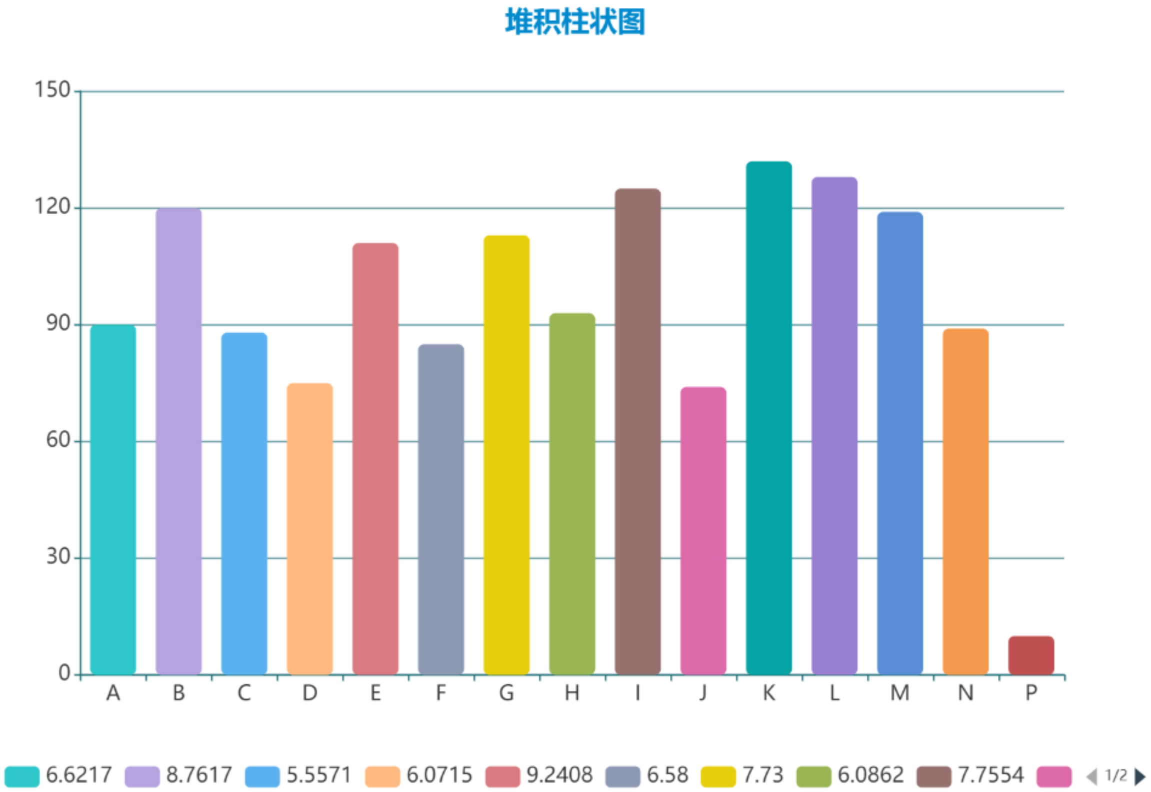


图 4　该地各个区域的占地面积(单位：平方公里)

5.4.3 结果分析

将 MATLAB 计算出的结果汇总制表可得出下表：

表 5 各区域的 L(α)值与最强相关事件类别

区域	所有事件密度总和与面积比值 L(α)	相关性最强事件类别
A	1.7444	⑦
B	0.975	⑦
C	1.0455	⑦
D	2.04	⑦
E	2.2072	⑦
F	1.8021	⑦
G	1.7345	⑦
U	1.129	⑦
I	0.584	⑦
J	2.0946	⑦
K	1.3636	⑦
L	0.78906	⑦
M	1.8151	⑦
N	1.1111	⑦
P	165.7576	⑦

因为此处各类事件密度在空间上的相关性用其所有事件密度总和与面积比值比率表示，即用 L(α)值表示；L(α)的值越大表示在该区域的单位面积内、单位时间范围内，发生事件的几率越大，事件密度在空间上的相关性越强。其中所有区域的最强事件类别是⑦类事件。

5.5 问题五的求解

5.5.1 数据的预处理及解答思路：

这里可以直接借用第 4 题已近计算得出的时间密度数据，将 5.4 中用到的事件类别与区域编号与区域信息表载入模型，然后求出人口密度解答此题。

5.5.2 模型的建立

本题采用概率分布来分析事件密度与人口密度之间的相关性，先计算出人口密度。假设 P(α)为某区域人数(单位：万人);D(α)为某区域人口密度(单位：万人/平方公里)：

$$D(\alpha) = \frac{P(\alpha)}{S(\alpha)} \qquad ()$$

得到下表：

表 6　各个区域的人口密度

区域	人口密度
A	0.073574444
B	0.073014167
C	0.063148864
D	0.080953333
E	0.08325045
F	0.077502945
G	0.06840708
U	0.065443011

I	0.0620432
J	0.072972973
K	0.066559091
L	0.045797656
M	0.082361345
N	0.084175084
P	1.606060606

转置事件类别与区域编号并将每一个元素通过循环除某一区域内发生事件的总数，得到在某地发生某种类别的事件的概率。

5.5.3 答案的求解

表 7　在当前人口密度条件下各区域各事件的发生概率

区域	事件①	事件②	事件③	事件④	事件⑤	事件⑥	事件⑦
A	0.121019	0.038217	0.140127	0.121019	0.006369	0.019108	0.55414
B	0.162393	0.025641	0.136752	0.059829	0	0.017094	0.598291
C	0.130435	0.01087	0.152174	0.097826	0.021739	0	0.586957
D	0.124183	0.058824	0.163399	0.084967	0.013072	0.013072	0.542484
E	0.102041	0.016327	0.257143	0.036735	0.008163	0.012245	0.567347
F	0.169935	0.006536	0.163399	0.058824	0.013072	0.006536	0.581699
G	0.102041	0.010204	0.168367	0.096939	0.015306	0.010204	0.596939
U	0.142857	0.057143	0.104762	0.104762	0.019048	0	0.571429
I	0.082192	0.013699	0.136986	0.082192	0	0.013699	0.671233
J	0.129032	0.083871	0.180645	0.077419	0.006452	0.051613	0.470968
K	0.111111	0.016667	0.177778	0.055556	0	0.011111	0.627778
L	0.049505	0.019802	0.227723	0.138614	0.029703	0.059406	0.475248
M	0.069444	0.027778	0.231481	0.055556	0.013889	0.027778	0.574074
N	0.131313	0.050505	0.30303	0.060606	0	0.010101	0.444444
P	0.106642	0.030469	0.143205	0.078001	0.011578	0.141987	0.488117

六、误差分析

6.1 附件中对该地消防救援出警的数据只有 2016 年到 2020 年期间一共五年的数据，收集数据量较少，使得结果具的准确性和可信度不高。

6.2 在解题过程中，可以很明显发现该地 2020 年有几个月的消防救援出警次数较前四年的次数偏少，结合其数据与实际现实情况分析，应该是该地受到新冠病毒疫情等不确定因素影响导致部分消防事件减少，最终影响结果具有一定偏差。

七、模型的分析与改进

模型的优点：

模型使用了伪线性规划模型、灰色预测、平均移动法、概率分布等，将本题所给出的有限数据中将有效数据利用,从建模与算法上尽量提高结果的可靠性与稳定性。对于问题 1、问题 4 的模型具有可行性。

模型的缺点：

问题 2 题目给出的数据较少，不能更大限度的检测出该模型的结果是否可信，且检验数据受外界影响因素较大，导致该模型的可靠性和稳定性在本题中并不是很

理想。

八、参考文献

[1]卓金武，李必文.MATLAB 在数学建模中的应用（第 3 版）.北京：北京航空航天大学出版社，2020.

九、附录

各题的 MATLAB 运行代码如下：

```
%第一题 MATLAB 代码求解：
load AreaInfo.mat
load PliceData.mat
date = PliceData.Date; %存储事件的发生日期
time = PliceData.Time; %存储事件发生的具体时间
fev = [0,0,0]; %存储所有年份内2月1日三种时间段的数据
may = [0,0,0]; %存储所有年份内5月1日三种时间段的数据
aug = [0,0,0]; %存储所有年份内8月1日三种时间段的数据
nov = [0,0,0]; %存储所有年份内11月1日三种时间段的数据
for i=1:length(date) %在循环内遍历每一个元素
    m = date(i).Month; %将日期时间数据转化为可运算的基本数据类型
    d = date(i).Day; %同上
    if(m==02 && d == 1) %限定时间为2月1日
        fev(bool(time,i)) = fev(bool(time,i))+1;
    elseif(m==05 && d == 1) %限定时间为5月1日
        may(bool(time,i)) = may(bool(time,i))+1;
    elseif(m==08 && d == 1) %限定时间为8月1日
        aug(bool(time,i)) = aug(bool(time,i))+1;
    elseif(m==11 && d == 1) %限定时间为11月1日
        nov(bool(time,i)) = nov(bool(time,i))+1;
    end
end
%输出最优值
disp(linp(fev))
disp(linp(may))
disp(linp(aug))
disp(linp(nov))
function x = linp(a) %定义伪线性规划对传入的数列进行求解
sum = 0;
x = [0,0,0];
for i=1:length(a)
    sum = sum + a(i);
end
for i=1:length(a)
    if((a(i)/sum)*30<=5)
        x(i) = 5;
```

```
        sum = sum-a(i);
    else
        x(i) = round((a(i)/sum)*(30-5));
    end
end
end
function f = bool(time,i) %定义此函数用于划分时间段
if(time(i) < 0.333)
    f = 1;
elseif(time(i) < 0.667)
    f = 2;
else
    f = 3;
end
end
%第二题MATLAB求解:
clc;
clear;
load PliceData.mat
date = PliceData.Date;
time = PliceData.Time;
Pop = zeros(5,12); %定义矩阵Pop用于存储年份（行）月份（列）出警次数的数
据
for i=1:length(date) %遍历所有元素
    y = date(i).Year; %存储事件发生的年份
    m = date(i).Month; %存储事件发生的月份
    %对每个数据进行判断并存入Pop变量中
    if(y ~= 2020)
        if(y == 2016)
            Pop(1,m) = Pop(1,m) + 1;
        elseif(y == 2017)
            Pop(2,m) = Pop(2,m) + 1;
        elseif(y == 2018)
            Pop(3,m) = Pop(3,m) + 1;
        else
            Pop(4,m) = Pop(4,m) + 1;
        end
    else
         Pop(5,m) = Pop(5,m) + 1;
    end
end
% 灰色预测
g = [];
for j=1:12
```

```
    syms a b;
    c = [a b]';
    A = Pop(1:4,j);
    A = A';
    B = cumsum(A);
    n = length(A);
    for i = 1 : n-1
        C(i) = (B(i)+B(i+1))/2;
    end
    D = A;D(1) = [];
    D = D';
    E = [-C;ones(1,n-1)];
    c = inv(E*E') *E *D;
    c = c';
    a = c(1);b = c(2);
    F = [];F(1) = A(1);
    for i = 2:(n + 2)
        F(i) = (A(1)-b/a)/exp(a*(i-1))+b/a;
    end
    G = [];G(1) = A(1);
    for i = 2:(n + 2)
        G(i) = F(i)-F(i-1);
    end
     t1 = 2016:2019;
    t2 = 2016:2021;
    g = [g G(6)];
end
%Pre数据预处理MATLAB代码：
load PliceData.mat
load AreaInfo.mat
category = PliceData.Category;
area = PliceData.AreaID;
a = zeros(1,length(area)); %用于存储处理后的地区ID
c = zeros(1,length(category)); %用于存储处理后的类别
for i=1:length(category)
    %将类别①...替换为1...
    if(category(i)=='①')
        c(i) = 1;
    elseif(category(i)=='②')
        c(i) = 2;
    elseif(category(i)=='③')
        c(i) = 3;
    elseif(category(i)=='④')
        c(i) = 4;
```

```
    elseif(category(i)=='⑤')
        c(i) = 5;
    elseif(category(i)=='⑥')
        c(i) = 6;
    else
        c(i) = 7;
    end
    %将地区名ABC...替换为123...
    if(area(i)=='A')
        a(i) = 1;
    elseif(area(i)=='B')
        a(i) = 2;
    elseif(area(i)=='C')
        a(i) = 3;
    elseif(area(i)=='D')
        a(i) = 4;
    elseif(area(i)=='E')
        a(i) = 5;
    elseif(area(i)=='F')
        a(i) = 6;
    elseif(area(i)=='G')
        a(i) = 7;
    elseif(area(i)=='H')
        a(i) = 8;
    elseif(area(i)=='I')
        a(i) = 9;
    elseif(area(i)=='J')
        a(i) = 10;
    elseif(area(i)=='K')
        a(i) = 11;
    elseif(area(i)=='L')
        a(i) = 12;
    elseif(area(i)=='M')
        a(i) = 13;
    elseif(area(i)=='N')
        a(i) = 14;
    elseif(area(i)=='P')
        a(i) = 15;
    end
end
PliceData = PliceData(:,1:3);
save PliceData c a PliceData;
%第三题MATLAB求解:
clear;
```

```
clc;
load PliceData.mat
date = PliceData.Date;
Pop = zeros(7,12); %定义变量Pop存储类别（行）月份（列）出警次数的数据
for i=1:length(date) %遍历所有元素
    m = date(i).Month;
    Pop(c(i),m) = Pop(c(i),m) + 1; %将每个元素对应放进变量Pop中
end
%平均移动法
for h=1:7
y=Pop(h,:);
m=length(y);
n=1:11;  %当N为11时S最小
for i=1:length(n)
    for j=1:m-n(i)+1
        yhat{i}(j)=sum(y(j:j+n(i)-1))/n(i);
    end
    s(i)=sqrt(mean((y(n(i)+1:m)-yhat{i}(1:end-1)).^2));
end
end
fprintf("平均移动法的平均标准差：%f",sum(s)/length(s))
%非线性规划
%第四题MATLAB求解：
clear;
clc;
load AreaInfo.mat;
load PliceData.mat
id = AreaInfo.AreaID; %存储区域编号
s = AreaInfo.km2; %存储区域面积
E = zeros(7,15); %定义矩阵E用于存储事件类别（行）区域编号（列）
for i=1:length(c) %遍历所有元素
    E(c(i),a(i)) = E(c(i),a(i))+1; %将元素分别存入符合条件的向量中
end
%分析该地区2016-2020年各类事件密度在空间上的相关性
l = zeros(1,15); %用于存储不同区域事件密度密度关于区域面积的比率
for i=1:length(l)
    l(i)=sum(E(:,i))/s(i);
    disp(id(i)+" : "+l(i));
end
%不同区域相关性最强的事件类别
for i=1:length(l)
    [x,y] = max(E(:,i));
    disp(id(i)+" 区域内与 "+y+" 类别相关性最强");
end
```

```
%第五题MATLAB求解：
clear;
clc;
load AreaInfo.mat
load E %事件类别（行）区域编号（列）
%计算人口密度  (万人/平方公里)
len = length(AreaInfo.AreaID);
P_density = zeros(1,len); %人口密度
for i=1:len
   P_density(i) = AreaInfo.Population(i)/AreaInfo.km2(i);  %人口密度=人
口总数/面积
end
E; %为各地区各类事件发生的数量
%该地各类事件密度  (每一个元组代表区域发生不同类型事件的概率)
E_density = E'; %先对E进行转置
for i=1:15
    Sum = sum(E_density(i,:)); %第一个循环记录各类别对应所有区域的出警次
数之和
    for j=1:7
        E_density(i,j) = E_density(i,j)/Sum; %第二个循环记录比率（单个区
域数量/总区域数量）
    end
end
```