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。

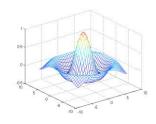
我们授权五一数学建模竞赛组委会,可将我们的论文以任何形式进行公开展示(包括进行网上公示,在书籍、期刊和其他媒体进行正式或非正式发表等)。

参赛题号	(从 A/B/C 中选技	¥一项填写) :	В	
参赛队号:	T	21037582375424		
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:	何依		-
	队员2姓名:	刘冀鑫		<u></u>
	队员3姓名:	周国秦		<u></u>
联系方式:	Email:19′	706648@qq.com	联系电话:	18980697207

日期: 2021年5月1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: 基于消防救援出警情况的优化分析

关键词: Matlab 编程 灰色预测 曲线拟合 最优化模型 相关性分析

摘 要:

针对问题一,依据 2 月、5 月、8 月、11 月中第一天三个时间段的接警次数,通过计算每个时间段接警的概率值来分配这四个月的第一天三个时间段的值班安排人数。结论如下: 2 月 1 日第一时间段值班人数 5 人,第二时间段值班人数 13 人,第三时间段值班人数 12 人;第三时间段值班人数 5 人,第二时间段值班人数 12 人,第三时间段值班人数 13 人;第三时间段值班人数 14 人;第三时间段值班人数 5 人,第二时间段值班人数 11 人,第三时间段值班人数 14 人;11 月 1 日第一时间段值班人数 5 人,第二时间段值班人数 17 人,第三时间段值班人数 8 人。

针对问题二,需要我们建立消防救援出警次数的预测模型方案,通过微分方程模型和灰色预测模型方案的对比发现,灰色预测方案的相对误差较低。所以我们选择最佳预测模型为灰色预测模型。并且获得 2020 年的预测数值见表格 7。

针对问题三,依据7种类别事件的发生时间,需要我们建立各类事件发生次数与月份关系的多种数学模型。通过对事件发生次数和月份之间的统计,使用曲线拟合对7种类别事件进行曲线判优,由此得到每类事件发生次数的最优模型,见文中图2——图8

针对问题四,需要我们分析该地区 2016-2020 年各类事件密度在空间上的相关性,所以我们可以根据求解事件密度,得出各类事件密度-面积的统计表,再判断该数据是否符合皮尔逊相关系数。然后通过公式模型求解得出相关系数,最后通过相关性比较求出相关性最强的事件类别为事件②。

针对问题五,依据附件1的各地区人口数据算出人口密度,再利用问题四求出的事件密度,通过对人口密度和事件密度进行曲线拟合,由此得到人口密度与事件密度之间的关系为各地区事件密度随人口密度的增大而增大。

最后,我们模型的优点就是思路清晰明了,考虑较为全面,具有一定的可操作性, 在现实生活中具有实际意义。

1. 问题重述

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?

2.问题分析

2.1 问题一的分析

针对问题一,依据2月、5月、8月、11月中第一天三个时间段的接警次数,通过计算每个时间段接警的概率值来分配这四个月的第一天三个时间段的值班安排人数。

2.2 问题二的分析

针对问题二,需要我们建立消防救援出警次数的预测模型方案,通过微分方程模型和灰色预测模型方案的对比发现,灰色预测方案的相对误差较低,delta为()。

所以我们选择最佳预测模型为灰色预测模型。并且获得2020年的预测数值。

2.3 问题三的分析

针对问题三,依据7种类别事件的发生时间,需要我们建立各类事件发生次数与月份关系的多种数学模型。通过对事件发生次数和月份之间的统计,使用曲线拟合对7种类别事件进行曲线判优,由此得到每类事件发生次数的最优模型。

2.4 问题四的分析

针对问题四,需要我们分析该地区 2016-2020 年各类事件密度在空间上的相关性,所以我们可以根据求解事件密度,得出各类事件密度-面积的统计表,再判断该数据是否符合皮尔逊相关系数。然后通过公式模型求解得出相关系数,最后通过相关性比较求出相关性最强的事件类别。

2.5 问题五的分析

针对问题五,依据附件1的各地区人口数据算出人口密度,再利用问题四求出的事件密度,通过对人口密度和事件密度进行曲线拟合,由此得到人口密度与事件密度之间的关系。

2.6 问题六的分析

针对问题六,根据问题中的图建立邻接矩阵,通过 Floyd 算法求出各地区间的最短路径,筛选出到各地方最近的地区。再统计出 16 年~20 年各区域的出警人数。以这两个数据计算出除区域 J 和区域 N 以外所有的平均出警距离,通过这几个数据里确定消防站的具体修建地。

3.符号说明

表格 1

序号	符号	意义
1	m	月份
2	tn	事件 n 发生次数
3	Р	人口密度
4	Q	事件密度

4.模型建立与模型求解

4.1 问题一的模型建立与模型求解

首先,我们对附件 2 的数据进行筛选统计。将 2 月、5 月、8 月、11 月的第一天三个时间段的接警次数统计如下表:

表格 2

	接警次数	Ⅱ接警数	Ⅲ接警数	总接警数
2月1日	0	8	7	15
5月1日	2	7	8	17
8月1日	1	3	4	8
11月1日	0	4	2	6

然后计算 2016-2020 年的每个时间段占总接警数的概率如下:

表格 3

	时间段概率	时间段概率	Ⅲ时间段概率
2月1日	0	0.533333	0.466667
5月1日	0.117647	0.411765	0.470588
8月1日	0.125	0.375	0.5
11月1日	0	0.666667	0.333333

由上表可推测人数:由表中数值可知,四个月份在 I 时间段接警的概率较小,按照题意只能在这个时段分配 5 人。2 月 1 日 II、III时间段值班人数总共为 25 人,根据概率值分配得到,II 时间段值班人数为 13 人,III时间段值班人数为 12 人。5 月 1 日 II、III时间段值班人数总共为 25 人,根据概率值分配得到,II 时间段值班人数为 12 人,III时间段值班人数为 13 人。8 月 1 日 II、III时间段值班人数总共为 25 人,根据概率值分配得到,II 时间段值班人数为 14 人。2 月 1 日 II、III时间段值班人数总共为 25 人,根据概率值分配得到,II 时间段值班人数为 14 人。2 月 1 日 II、III时间段值班人数为 8 人。总结如下表:

表格 4

	值班人数	Ⅱ值班人数	Ⅲ值班人数	总值班人数
2月1日	5	13	12	30
5月1日	5	12	13	30
8月1日	5	11	14	30
11月1日	5	17	8	30

5.2 问题二的模型建立与模型求解

(1) 原始数据。

首先,统计各月的接警数总和。如下表:

表格 5

With 9						
	2016 年	2017 年	2018年	2019年	年总和	
1月	79	58	54	67	258	
2月	95	28	129	76	328	
3月	98	44	68	73	283	
4月	51	14	71	44	180	
5月	275	148	107	138	668	
6 月	67	87	119	146	419	
7月	63	23	59	72	217	
8月	40	42	50	30	162	
9月	50	49	46	41	186	
10 月	31	33	44	26	134	
11 月	38	65	40	25	168	
12 月	43	68	63	36	210	

(2) 模型建立。(以1月为例,后续月份以此类推)

1. 问题二要求以月份为单位,我们就以此为一个为变量,通过对灰色预测模型的选择,我们选择 GM(1.1)预测最优模型,再根据题目约束条件,求出最优解。再在 matlab 软件中编程得到预测模型和预测数值(见附录)5.1,2 模型建立极比检验。建立消防救援出警次数的预测模型;序列如下:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(4) = (79,58,54,67))$$

2. 求级比 λ (k),有

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}$$

$$\lambda = (\lambda(2), \lambda(3), \lambda(4) = (1.3621, 1.0741, 0.8060))$$

- 3. 级比判断。由于所有的 λ (k) \in [0.8060,1.3621],k=2,3,4,故可以用 $x^{(0)}$ 作 GM(1,1) 建模。
- 4. GM(1,1)建模

对原始数据 $x^{(0)}$ 作一次累加,得

$$x^{(1)} = (79,137,191,258)$$

构造数据矩阵 B 及数据向量 Y,有

$$\mathbf{B} = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ -\frac{1}{2}(x^{(1)}(3) + x^{(1)}(4)) & 1 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \end{bmatrix}.$$

计算:

$$\hat{u} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y = \begin{bmatrix} -0.0791 \\ 46.5768 \end{bmatrix}$$

于是得到 $\hat{a} = -0.0791$, $\hat{b} = 46.5768$.

模型求解:

$$\frac{dx^{(1)}}{dt} + \hat{a}x^{(1)} = \hat{b}$$

求解得:

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}\right)e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} = 667.8344e^{0.0791k} + 588.8344$$

求生成序列预测值 $\hat{x}^{(1)}(k+1)$ 及模型还原值 $\hat{x}^{(0)}(k+1)$,令 k=1,2,3,由上式的时间响应函数可算得 $\hat{x}^{(1)}$, 其中取 $\hat{x}^{(1)}(1)=\hat{x}^{(0)}(1)=x^{(0)}(1)=79$,由 $\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)$,取 k=1,2,3,得

$$\hat{x}^{(0)} = \left(\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4)\right) = (79,55.0,59.5,64.4)$$

売格 6

年份	原始值	预测值	残差	相对误 差	级比偏 差
2016	79	79.0000	0	0%	
2017	58	54.9704	3.0296	0.05%	-0.4742
2018	54	59.4947	-5.4947	0.10%	-0.1625
2019	67	64.3914	2.6086	0.04%	0.1277

经验证,该模型的精度较高,可进行预测和预报。[1] 通过 matlab 代码编译,可得 2021 年每月的出警次数预测值为:

表格 7

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 问题三的模型建立与模型求解

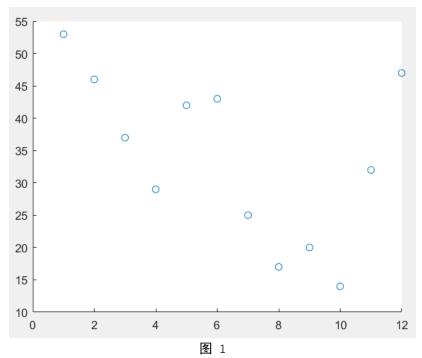
根据问题三,需要我们根据 7 种类别事件的发生时间,建立各类事件发生次数与月份关系的多种数学模型。通过以拟合度最优为评价标准,确定每类事件发生次数的最优模型。所以我们采用最优化拟合模型对数据进行处理。

丰权	0
42 1H	റ

	事件①	事件②	事件③	事件④	事件⑤	事件⑥	事件⑦
1月	53	9	56	32	1	22	113
2月	46	10	86	26	2	19	162
3 月	37	9	60	18	2	16	177
4月	29	10	24	22	6	9	126
5月	42	15	219	20	7	25	403
6月	43	7	60	35	9	37	284
7月	25	11	13	18	2	32	137
8月	17	6	3	19	11	23	112
9月	20	8	7	20	2	39	127
10 月	14	5	7	17	2	15	101
11月	32	12	36	25	2	12	102
12 月	47	10	46	32	2	20	116

以事件①为例 (后续事件以此类推):

设 t 为事件①发生次数,m 为月份,在 m-t 平面上画散点图,由此得知 t 与 m 大致为非线性关系。



通过 matlab cftool 曲线拟合工具,对散点进行曲线拟合后,得到 Degree 中的拟

合优度 R^2 如下表:

表格 9

, ,	771H -				
Degree	拟合优 度				
1	-0.2131				
2	0.4979				
3	0.6108				
4	0.7730				
5	0.7861				
6	0.8830				
7	0.8670				
8	0.8714				
9	0.9998				

由上表可知: 当 Degree 等于 9 时,拟合优度 R^2 最接近于 1,即为最优。 由此可得非线性回归方程为:

 $t1 = -0.0004978m^9 + 0.02898m^8 - 0.7169m^7 + 9.828m^6 - 81.59m^5 + 421.2m^4$ $-1336m^3 + 2482m^2 - 2422m + 980.7$

其拟合图像如下图:

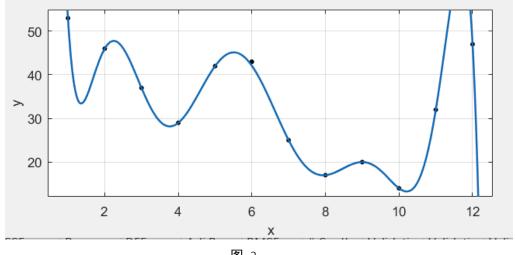
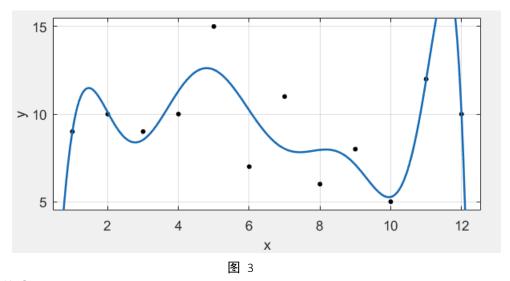


图 2

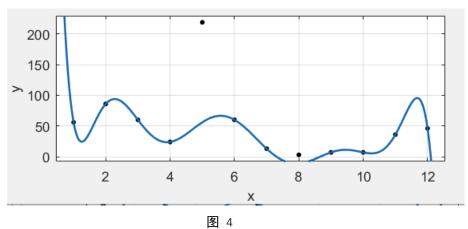
其他事件同上述操作可得: 事件②:

$$t2 = -3.561m^9 - 8.677m^8 + 16.65m^7 + 36.77m^6 - 26.57m^5 - 44.79m^4 + 20.25m^3 + 14.75m^2 - 8.291m + 8.851$$



事件③:

$$t3 = -133.5m^9 + 13.26m^8 + 689.9m^7 - 95.81m^6 - 1186m^5 + 208.5m^4 + 782.9m^3 - 128.2m^2 - 184.1m + 38.89$$



事件4):

 $t4 = -0.0002109m^9 + 0.01229m^8 - 0.3014m^7 + 4.167m^6 - 34.52m^5 + 177.4m^4 - 559.3m^3 + 1032m^2 - 1004m + 417.2$

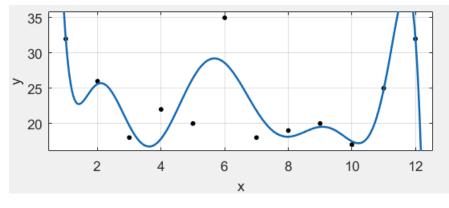
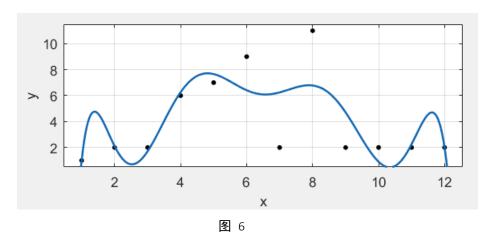


图 5

事件⑤:

$$t5 = 1.056e - 05m^9 - 0.0008846m^8 + 0.02923m^7 - 0.5115m^6 + 5.259m^5 - 32.87m^4 + 123.8m^3 - 266.6m^2 + 293.2m - 121.3$$



事件⑥:

 $t6 = -0.0006588m^9 + 0.03867m^8 - 0.9658m^7 + 13.38m^6 - 112.5m^5 + 589.2m^4 - 1901m^3 + 3595m^2 - 3562m + 1401$

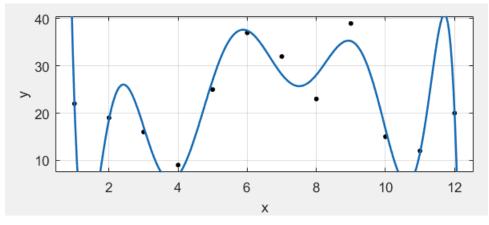


图 7

事件⑦:

$$t7 = -0.004812m^9 + 0.2797m^8 - 6.896m^7 + 94.02m^6 - 774.4m^5 + 3956m^4 - 1.24e + 04m^3 + 2.27e + 04m^2 - 2.127e + 04x + 8261$$

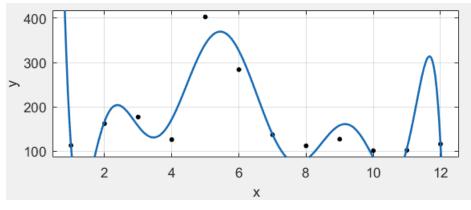


图 8

通过以上7次对每个事件的最优度拟合,可以确定各类事件发生次数的最优模型。

5.4 问题四的模型建立与模型求解

针对问题四,需要我们分析该地区 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,所以我们必须先求出事件事件密度。 在获取事件密度的情况下,建立起事件密度和区域面积的表格,如下表:

区域	区域面积	事件①	事件②	事件③	事件④	事件⑤	事件⑥	事件⑦
А	90	0.087%	0.027%	0.101%	0.087%	0.005%	0.014%	0.398%
В	120	0.065%	0.010%	0.055%	0.024%	0.000%	0.007%	0.240%
С	88	0.056%	0.005%	0.056%	0.042%	0.009%	0.000%	0.253%
D	75	0.104%	0.049%	0.137%	0.071%	0.011%	0.011%	0.455%
Е	111	0.093%	0.015%	0.234%	0.033%	0.007%	0.011%	0.515%
F	85	0.126%	0.005%	0.121%	0.044%	0.010%	0.005%	0.431%
G	113	0.073%	0.007%	0.120%	0.069%	0.011%	0.007%	0.426%
Н	93	0.066%	0.027%	0.049%	0.049%	0.009%	0.000%	0.265%
-	125	0.020%	0.003%	0.033%	0.020%	0.000%	0.003%	0.161%
J	74	0.100%	0.072%	0.156%	0.061%	0.006%	0.044%	0.406%
К	132	0.062%	0.009%	0.100%	0.031%	0.000%	0.006%	0.352%
L	128	0.016%	0.006%	0.074%	0.045%	0.010%	0.019%	0.154%
М	119	0.052%	0.021%	0.173%	0.041%	0.010%	0.021%	0.429%
N	89	0.060%	0.023%	0.139%	0.028%	0.000%	0.005%	0.203%
Р	10	7.202%	2.058%	9.671%	5.267%	0.782%	9.588%	32.963%

表格 10

之后通过对原数据进行分析,该表数据符合皮尔逊相关系数,故我们对此建立相关系数的模型。

假设有两组数据 $X:\{X_1,X_2,\dots,X_n\}$ 和 $Y:\{Y_1,Y_2,\dots,Y_n\}$

样本均值:
$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

样本协方差: Cov(X,Y)=
$$\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{n-1}$$

样本 Pearson 相关系数:
$$r_{XY} = \frac{cov(X,Y)}{s_X s_Y}$$

其中: S_X (sigma X)是 X 的样本标准差, $S_X = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$,同理 $S_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}}$ 相关系数的绝对值越大,相关性越强,相关系数越接近于 1 或-1,相关度越强,相关系数越接近于 0,相关度越弱。

通常情况下通过以下取值范围判断变量的相关强度:

相关系数: 0.8-1.0 极强相关

0.6-0.8 强相关

0.4-0.6 中等程度相关

0.2-0.4 弱相关

0.0-0.2 极弱相关或无相关[2]

模型求解:

Step1: 根据公式我们求出事件①—⑦的相关系数为

表格 11

WIH II		
相关系数		
-0.78414		
-0.79164		
-0.77983		
-0.78187		
-0.78125		
-0.77786		
-0.77994		

Step2: 由相关系数的值得出相关性: 相关性最强的为事件②。[3]

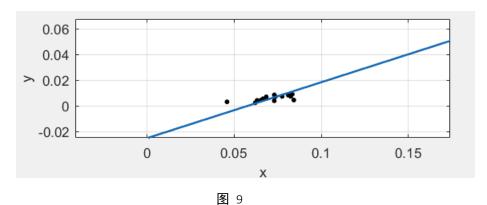
5.5 问题五的模型建立与模型求解

根据题目要求,算出各地区的事件密度和人口密度如下表:

表格 12

区域	人口密	事件密
	度	度
Α	0.0736	0.0072
В	0.0730	0.0040
С	0.0631	0.0043
D	0.0810	0.0084
Е	0.0833	0.0091
F	0.0775	0.0074
G	0.0684	0.0071
Н	0.0654	0.0046
I	0.0620	0.0024
J	0.0730	0.0086
K	0.0666	0.0056
L	0.0458	0.0032
М	0.0824	0.0075
N	0.0842	0.0046
Р	1.6061	0.6753

设人口密度为 P,事件密度为 Q,根据上表使用 matlab cftool 曲线拟合工具进行拟合(如下图),可知人口密度和事件密度之间满足线性关系:



其满足 P=0.4359O-0.02509

由此可推断出,各地区事件密度随人口密度的增大而增大。

5. 参考文献

[1] 司守奎, 孙兆亮, 数学建模算法与应用(第2版), 北京: 国防工业出版社, 2019;

[2] 星码,数学建模----皮尔逊相关系数, https://blog.csdn.net/qq_43779658/article/details/107748177?ops_request_misc=%257B %2522request%255Fid%2522%253A%2522161995577216780366516497%2522%252C %2522scm%2522%253A%252220140713.130102334..%2522%257D&request_id=1619 95577216780366516497&biz_id=0&utm_medium=distribute.pc_search_result.none-task-blog-2~all~sobaiduend~default-1-

107748177.nonecase&utm_term=%E7%9A%AE%E5%B0%94%E9%80%8A%E6%A8%A1%E5%9E%8B, 2020.08.03;

[3] 司守奎, 孙兆亮, 数学建模算法与应用习题解答(第2版), 北京: 国防工业出版社, 2019;

6.附录

```
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);
%预测后续数据
F=[];F(1)=A(1);
for i=2:(n+2)
    F(i)=(A(1)-b/a)/\exp(a*(i-1))+b/a;
end
for i=2:(n+2)
                    %向后预测2年
    G(i)=F(i)-F(i-1);
end
t1=2016:2019;
t2=2016:2021;
G
plot(t1,A,'o',t2,G) %原始数据与预测数据的比较
clc,clear
function coeff = myPearson(X, Y)
问题 4 代码:
% 本函数实现了皮尔逊相关系数的计算操作
%
%输入:
% X: 输入的数值序列
% Y: 输入的数值序列
%
%输出:
% coeff: 两个输入数值序列 X, Y 的相关系数
%
if length(X) \sim = length(Y)
  error('两个数值数列的维数不相等');
  return;
end
fenzi = sum(X .* Y) - (sum(X) * sum(Y)) / length(X);
fenmu = \operatorname{sqrt}((\operatorname{sum}(X .^2) - \operatorname{sum}(X)^2 / \operatorname{length}(X)) * (\operatorname{sum}(Y .^2) - \operatorname{sum}(Y)^2 / \operatorname{length}(X))
);
coeff = fenzi / fenmu;
```

%函数 myPearson 结束 disp('预测数据为: ');6 coeff end