**一．实验名称：新冠疫情预测**

**二．小组成员学号姓名及分工：**

分工：单人独立完成实验报告所有内容

**三．实验过程：**

**1.实验目的**

搜集全球某一个国家的新冠数据，研究数据的函数类型，根据某一国家截止2022年5月底的数据，预计该国家感染的最终规模，判断该国家疫情何时结束。

**2.数据采集**

选择泰国截止至2022年5月30日的新冠感染人数数据并制作成如下表格：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5.9 | 5.10 | 5.11 | 5.12 | 5.13 | 5.14 | 5.15 | 5.16 | 5.17 | 5.18 | 5.19 |
| 4029959 | 4063844 | 4106230 | 4148090 | 4180863 | 4209571 | 4250949 | 4271815 | 4290824 | 4308319 | 4324850 |
|  |  |  |  |  |  |  |  |  |  |  |
| 5.20 | 5.21 | 5.22 | 5.23 | 5.24 | 5.25 | 5.26 | 5.27 | 5.28 | 5.29 | 5.30 |
| 4337568 | 4353237 | 4367752 | 4382977 | 4394915 | 4406755 | 4415593 | 4424750 | 4434511 | 4442648 | 4450457 |

**3.模型选择**

在实际生活中，新冠感染人数达到一定数量时，由于各种社会因素，感染人数增加速度将会减慢，直到最后不再增长。所以，我们选取三个函数模型分别处理上述表格中的数据，对比三种处理结果，选择最优的模型进行预测。选取的三种模型依次为Logistic模型,和

**4.实验步骤**

首先使用Logistic模型进行预测，其matlab代码如下：

x=1:22;

y=[4029959,4063844,4106230,4148090,4180863,4209571,4250949,4271815,4290824,4308319,4324850,4337568,4353237,4367752,4382977,4394915,4406755,4415593,4424750,4434511,4442648,4450457];

a=[4800000,2,0.2];

f=@(a,x)a(1)./(1+a(2)\*exp(-a(3)\*x));

[A,resnorm]=lsqcurvefit(f,a,x,y)

t=27;

while f(A,t+1)-f(A,t)>=1

t=t+1;

end

t

f(A,t)

t=1:60;

plot(x,y,'\*',t,f(A,t))

为了方便记录，下面两种模型代码中的y将简记为y=[…]

运行此代码得到的结果如下（相关图形见图1）：

A =

1.0e+06 \*

4.4996 0.0000 0.0000

resnorm =

4.5608e+08

t =

103

ans =

4.4996e+06

结果表示疫情将在5月9日后的103天后结果，最终感染规模为456080000人。

然后使用模型进行预测，其matlab代码如下：

x=1:22;

y=[…];

a=[4800000,2.5,0.1,0.1];

f=@(a,x)a(1)./(1+a(2)\*exp(-a(3)\*x-a(4)\*x.^2));

[A,resnorm]=lsqcurvefit(f,a,x,y)

t=27;

while f(A,t+1)-f(A,t)>=1

t=t+1;

end

t

f(A,t)

t=1:40;

plot(x,y,'\*',t,f(A,t))

运行此代码得到的结果如下（相关图形见图2）：

A =

1.0e+06 \*

4.5525 0.0000 0.0000 -0.0000

resnorm =

4.1863e+08

t =

51

ans =

4.5068e+06

结果表示疫情将在5月9日后的51天后结果，最终感染规模为418630000人。

最后使用模型进行预测，其matlab代码如下：

x=1:22;

y=[…];

a=[4.5,0.122,0.1];

f=@(a,x)a(1)\*exp(exp(a(2)\*x+a(3)));

[A,resnorm]=lsqcurvefit(f,a,x,y)

t=27;

while abs(f(A,t)-f(A,t+1))>1

t=t+1;

end

t

f(A,t+1)

t=1:40;

plot(x,y,'\*')

hold on;

plot(t,f(A,t));

运行此代码得到的结果如下（相关图形见图3）：

A =

1.0e+06 \*

4.1564 0.0000 -0.0000

resnorm =

1.4526e+11

t =

190

ans =

Inf

结果表示疫情将在5月9日后的190天后结果，最终感染规模为145260000000人。

**5.实验结果及分析**

查阅资料可知：泰国新冠疫情官方宣布结束日期为2022年10月1日（距离5月9日约140天）。比较上面三个模型得出的预测数据，得出结论：使用Logistic模型预测出来的结果与实际最接近，误差最小。而其余两种模型预测出的结果与实际相差较大，可能的原因有以下几点：1.泰国新冠感染人数的趋势更接近于Logistic模型。2.模型预测时选取的初值不同导致预测结果偏离较大3.数据值太大，预测准确度较低。