小龙虾养殖数据预测与捕捞量模型

摘要

本文针对小龙虾池塘养殖系统的数学建模进行了研究，旨在解决三个关键问题：预测未来十年小龙虾的数量、确定最高年收获量以及制定最优捕捞策略。

针对问题一对一个小龙虾池塘养殖系统的未来十年小龙虾数量预测问题进行了研究。通过对过去 20 年小龙虾数量数据的分析，采用线性回归模型对小龙虾数量随时间的变化趋势进行了建模。模型假设小龙虾数量随时间呈线性增长，并通过最小二乘法拟合出了线性回归方程。预测结果显示，从 2021 年到 2030 年，小龙虾数量将稳定增长，每年平均增加约 0.093 千只，到 2030 年小龙虾数量将达到 6.65 千只。

针对问题二最高年收获量分析，本文旨在实现小龙虾养殖场长期稳定运营的目标。为此，研究通过建立数学模型分析如何在每年开始捕捞时各年龄虾数量基本保持不变的情况下获得最高的年收获量（捕获龙虾总重量）。小龙虾被分为四个年龄组，各年龄段具有相同的自然死亡率，并且只有三、四岁的小龙虾可以捕捞和产卵。基于各年龄段小龙虾的重量、自然死亡率、单只产卵量、卵存活率和捕捞强度等信息，本文建立了目标函数和约束条件，以确保每年不同年龄组的龙虾数量变化符合模型预测，控制捕捞强度以确保两个年龄组之间的捕捞行为相对均衡，避免过度捕捞导致虾群数量急剧下降，并确保所有虾群数量和捕捞量都是整数，从而在满足一系列规则的同时达到最大化捕获龙虾总重量的目标。

关键词：最小二乘法；约束条件；长期运营

一、 问题重述

1.1 问题背景

小龙虾是一种非常受欢迎的食物，因其体型较大、肉质丰富且味道鲜美，被广泛用于各种菜肴的制作。根据《中国小龙虾产业发展报告（2021）》中的数据，2021 年中国小龙虾养殖业的总产量达到了 208.96 万吨，产业总产值达到了 4110 亿元人民币，其中小龙虾养殖业的产值约为 710 亿元。

有一个荒废的池塘，里面养殖了小龙虾。随着时间的变化，小龙虾的数量也发生了变化。根据记录，从第一年到第二十年，小龙虾的数量逐年递增。

1.2 问题重述

1.未来十年龙虾数量预测已知池塘中龙虾数量随时间变化，建立数学模型预测未来十年龙虾数量，并列表给出预测结果。

2.最高年收获量分析，为了实现养殖场长期运营，需要每年开始捕捞时各年龄虾数量基本保持不变。建立数学模型分析如何在此前提下得到最高的年收获量（捕获龙虾总重量）。

3.养殖场最优捕捞策略，某龙虾养殖场要求 5 年后虾群的生产能力不能受到太大破坏。已知承包时各年龄组虾群数量，如果采用相同的捕捞强度，建立数学模型分析该养殖场应采取怎样的策略才能使总收获量最高。

二、问题分析

针对问题一，通过对 20 年的小龙虾数量数据的分析，对此采用线性回归对小龙虾进行预测，首先将年份与小龙虾数据进行线性拟合，建立一个简单的线性模型来描述增长趋势，考虑到线性回归模型能够较好地处理此类趋势性问题，所以将其作为首选方法。在模型建立过程中，将年份作为自变量，龙虾数量作为因变量，通过最小二乘法拟合出线性回归方程。分析结果表明，线性回归模型能够捕捉到龙虾数量随时间增长的基本趋势。

针对问题二——最高年收获量分析，目标是在确保小龙虾养殖场长期稳定运营[1]的前提下，通过建立数学模型来分析如何在每年开始捕捞时各年龄虾数量基本保持不变的情况下获得最高的年收获量（捕获龙虾总重量）。考虑到小龙虾分为四个年龄组，各年龄段具有相同的自然死亡率，并且只有三、四岁的小龙虾可以捕捞和产卵，基于各年龄段小龙虾的重量、自然死亡率、单只产卵量、卵存活率和捕捞强度等信息，制定了目标函数和约束条件。这些约束条件包括确保每年不同年龄组的龙虾数量变化符合模型预测、控制捕捞强度以确保两个年龄组之间的捕捞行为相对均衡、避免过度捕捞导致虾群数量急剧下降，以及确保所有虾群数量和捕捞量都是整数。

三、问题假设

1. 小龙虾数量随时间的变化趋势是线性的，即随着时间的推移，小龙虾数量的增长速度保持不变。

2.假设自变量（年份相对于起始年份的时间）与模型中的其他潜在自变量之间不存在高度相关性。

四、符号说明

序号	符号	符号说明
1	y	小龙虾的数量
2	x	年份
3	m	斜率，一年小龙虾增加率
4	b	截距,起始年份小龙虾数量
5	n	卵存活率
6	k	四年虾捕捞强度系数
7	r	产卵总量

五、模型建立与求解

5.1 问题一模型建立与解析

5.1.1 数据处理

通过整理题目中现有数据，可以得知通过题目中的近 20 年内连续数量的表格（表 1）需要预测未来 10 年龙虾的数量，为了更加直观的能看出历年来小龙虾数量的变化，因此把小龙虾数量做成图，如下图所示：

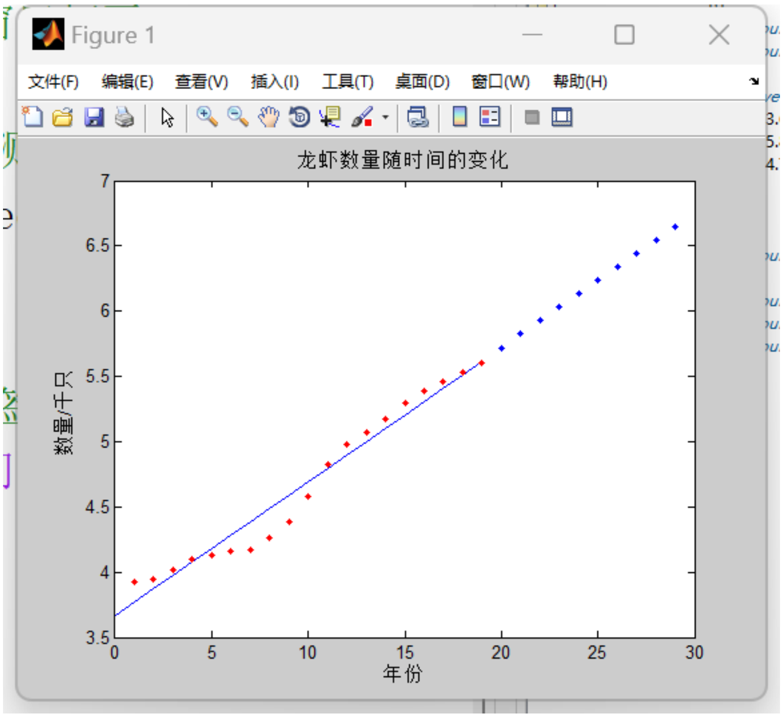

图 1 小龙虾变化趋势图

从这张图中，在图中，红色圆点代指过去二十年（1～20 年）的实际小龙虾数量，而蓝色圆点则代表未来十年（21～30 年）的预测龙虾数量，图中还有一条直线，这条直线就是过年 20 年拟合出来的趋势线，从图中可以看出，随着年份的增加，龙虾的数量也在逐渐增加。具体来说，在 0 到 5 年之间，龙虾的数量相对较少，增长速度较慢；但从第 6 年开始，增长速度明显加快，尤其是在 10 年到 20 年之间，龙虾的数量几乎呈指数级增长。

5.1.2 构建线性回归模型

线性回归模型是一种统计方法，用于分析一个或多个自变量与一个因变量之间的关系。在这个案例中，需要关注的是小龙虾数量随时间的变化趋势。为了预

测未来十年小龙虾的数量，可通过建立线性回归模型[3]来预测数据，线性回归模型的一般形式为：

$$y = mx + b$$

将线性回归模型带入问题一中：

y 代表小龙虾的数量，是模型中要预测的目标变量。

x 代表年份，相对起始年份的时间，即是模型中的自变量，即是通过该自变量来预测。

m 代表斜率，小龙虾数量每一年的增长率，也是一个关键参数，小龙虾数量随着时间的变化速度。

b 代表截距，表示起始年份小龙虾的数量，即是当 $x = 0$ 时 y 的值。

5.1.3 模型参数估计

要估计模型参数 m 和 b，我们使用最小二乘法拟合[2]线性回归模型。在 MATLAB 中，这可以通过 polyfit 函数完成。对于一阶多项式模型，polyfit 函数的数学定义如下:

$$p = polyfit(x, y, 1)$$

其中：

x 表示为年份

y 表示为小龙虾的数量

1 表示为拟合一阶多项式（线性模型）。

该函数返回一个包含两个元素的向量 p:

$$p = [m, b]$$

5.1.4 决定系数 R^2

为了更好的评估模型的拟合，先计算出决定系数 R^2， R^2 的计算公式如下:

$$R^2 = 1 - \frac{\sum (y_i - y_\alpha)^2}{\sum (y_i - y_\beta)^2}$$

其中：

y_α 是由模型得到小龙虾的数量预测值

y_i 时实际观察到小龙虾的数量

y_β 是小龙虾的平均值

5.1.5 问题一求解

为了预测未来十年小龙虾的数量，通过前文给出的线性回归方程，带入

MATLAB 软件中即可计算出模型信息，以下为输出结果：

模型方程： $y = mx + b$

斜率（m）：0.1028
截距（b）：3.6644

决定系数（R^2）：0.9649

参数解释

（1）斜距：如果 $m = 0.1028$，这意味着每过一年，小龙虾的数量预计会增加 0.1028 单位.

（2）截距：如果 $b = 3.6644$，这意味着起始年份小龙虾的数量为 3.6644 单位。

（3）决定系数：模型解释了小龙虾数量变化的 96.49%的变异，较高 R^2 的值表明模型很好地拟合了数据，即模型能够很好地解释小龙虾数量随时间的变化趋势。

显示预测结果

现已知所有基本参数，所以只需要将参数带入到线性回归模型即可求出未来十年每年预测小龙虾数量，如下所示：

表 1 未来十年小龙虾预测数量

年份	数量（千只）
2021	5.72
2022	5.82
2023	5.93
2024	6.03
2025	6.13
2026	6.23
2027	6.34
2028	6.44
2029	6.54
2030	6.65

从表 2 中可以看出从 2021 年到 2030 年，小龙虾数量总共增加了 0.93 千只。并且每年小龙虾数量的增长相对稳定，平均每年增加约 0.093 千只。

5.1.6 总结

根据提供的数据，可以看出小龙虾数量随时间呈现出稳定的增长趋势。从 2021 年到 2030 年，小龙虾数量逐渐增加，每年的增加量相对稳定。这些数据为资源规划提供了重要的参考信息。

5.2 问题二模型建立与求解

5.2.1 题目信息

由题目给定条件，小龙虾分为四个年龄组，各年龄段死亡率相同，且为季节性集中 产卵繁殖。规定捕捞期为前八个月，产卵孵化期为后四个月，且只有三、四岁小龙虾可 以捕捞和产卵。各年龄段小龙虾重量、自然死亡率、单只产卵量、卵存活率和捕捞强度 等信息，具体整理为以下表格。

表 2 虾群各年龄段信息特征

分组	重量（克）	自然死亡率	产卵量	卵存活率	不捕捞强度
一岁	5.17	0.79	0	\	0
两岁	11.58	0.79	0	\	0
三岁	17.66	0.79	5.54	n	0.42k
四岁	22.98	0.79	11.08	n	K

其中：

1.自然死亡率：：单位时间内死亡的虾的数量与虾的总量之比，

2.卵成活率：一岁虾数量与产卵总量 r 的比值，计算式为：

$$n = \frac{1.21\times 10^{11}}{1.21\times 10^{11} + r}$$

3.捕捞强度系数：单位时间捕捞量与各年龄组虾群条数的比例，规定年捕捞强度固定不变。

5.2.2 建立模型

通过题目信息，可以建立以下目标函数和约束条件，分析之间存在的关系。

目标函数

$$G = 17.66\cdot b_{j,3} + 22.98\cdot b_{j,4}$$

目标函数G表示在第 j 年捕获的龙虾总重量。其中$b_{j,3}$和$b_{j,4}$分别代表从年龄组 3 和年龄组 4 中捕获的龙虾数量。系数 17.66 和 22.98 分别是年龄组 3 和年龄组 4 龙虾的平均重量。

最大化G，即在满足所有约束条件下，尽可能地增加捕获龙虾的总重量。

约束条件

1.年龄组龙虾数量基本保持不变

$$\begin{aligned}
n_{j+1,1} &= (n_{j,4} - b_{j,4})\cdot 5.54 + (n_{j,3} - b_{j,3})\cdot 2.27 \\
n_{j+1,2} &= n_{j+1,1}\cdot(1-0.79) \\
n_{j+1,3} &= n_{j+1,2}\cdot(1-0.79) \\
n_{j+1,4} &= n_{j+1,3}\cdot(1-0.79) - b_{j,3}
\end{aligned}$$

$n_{j+1,1}$表示第 $j+1$年年龄组 1 的龙虾数量，它由前一年年龄组 3 和 4 的龙虾数量以及捕捞量决定。

$n_{j+1,2}$到$n_{j+1,4}$表示 $j+1$年龄组 2 到 4 的龙虾数量，由前一年年龄组相应存活率（1-0.79）决定的，年龄组 4 还减去了捕捞量$b_{j,3}$。

目的：确保每年不同年龄组的龙虾数量变化符合模型预测。

2.捕捞量与各年龄组虾数量的比例关系

$$\frac{b_{j,3}}{n_{j,3}} : \frac{b_{j,4}}{n_{j,4}} = 0.42$$

这个比例关系表明捕捞量$b_{j,3}$和$b_{j,4}$与各自年龄组的虾群数量$n_{j,3}$和$n_{j,4}$之比应保持一致，且该比例为 0.42。

目的：控制捕捞强度，确保两个年龄组之间的捕捞行为相对均衡。

3.捕捞量不超过龙虾数量

$$b_{j,3} < n_{j,3}, b_{j,4} < n_{j,4}$$

确保捕捞量不会超过相应年龄组的虾群数量。

目的：避免过度捕捞导致虾群数量急剧下降。

4.**龙虾数量和捕捞量均为整数**

$$n_{j,y} \in Z \geq 0, \quad b_{j,3}, b_{j,4} \in Z \geq 0$$

所有的虾群数量和捕捞量必须是整数。

目的：确保模型结果能够反映实际场景。

总结

这些约束条件共同作用于模型，确保了模型在满足一定规则的同时，达到最大化捕获龙虾总重量的目标。其中，年龄组龙虾数量保持不变的约束条件是整个模型的核心，因为它定义了如何从一年过渡到下一年，而其他约束条件则帮助控制捕捞活动的合理性，并确保模型的稳定性和可行性。

六、模型评价与推广

6.1 模型优点

1.线性回归模型结构简单，易于理解和应用，适合用来描述小龙虾数量随时间的线性增长趋势。

2.模型能够捕捉到小龙虾数量随时间增长的基本趋势，为未来十年小龙虾数量的预测提供了稳定的基础。

3.预测结果为资源规划提供了重要参考信息，有助于养殖场提前做好准备，应对小龙虾数量的变化。

6.2 模型缺点

1.模型假设小龙虾数量随时间的变化趋势是线性的，这一假设可能忽略了非线性因素的影响，例如环境变化、疾病爆发等因素可能会对小龙虾数量产生影响。

6.3 模型推广

1.模型预测结果为资源规划提供了重要参考信息，有助于养殖场提前做好准备，应对小龙虾数量的变化。

2.线性回归模型可以根据不同的数据集和预测目标进行调整，适用于多种养殖环境下的预测需求。

参考文献

[1].丁红霞,宋长太.小龙虾养殖提质增效技术措施[J].农村新技术,2024-07-05(07).

[2].李翔.基于 MCM 的多项式最小二乘法拟合的不确定度评定[J].环境技术,2024-07-25(07).

[3].冷建飞,高旭,朱嘉平.多元线性回归统计预测模型的应用[J].河海大学商学院,复旦大学财务金融系,2016(07).

附 录

附录 1
介绍：支撑材料的文件列表
问题 1 的源代码 Xiaolongxiashuliang.m 注：m 文件是 MATLAB 程序

附录 2
介绍：MATLAB 代码，求小龙虾未来十年数量

```
% 已有数据
years = [0:19]; % 将年份转换为相对于第一年的年数
lobsters = [3.81, 3.92, 3.95, 4.02, 4.10, 4.13, 4.16, 4.17, 4.26, 4.39, ...
            4.58, 4.83, 4.98, 5.07, 5.17, 5.30, 5.39, 5.46, 5.53, 5.60]; % 数量/千只

% 构建线性回归模型
p = polyfit(years, lobsters, 1);
m = p(1); % 斜率
b = p(2); % 截距

% 计算模型的决定系数 R²
```

```
y_mean = mean(lobsters);
SST = sum((lobsters - y_mean).^2); % 总平方和
SSR = sum((polyval(p, years) - y_mean).^2); % 回归平方和
R_squared = SSR / SST;

% 预测未来十年
futureYears = (20:29); % 未来十年的年数
predictedLobsters = polyval(p, futureYears);

% 输出模型信息
disp('模型方程: lobsters = m * year + b');
fprintf('斜率 (m): %.4f\n', m);
fprintf('截距 (b): %.4f\n', b);
fprintf('决定系数 (R²): %.4f\n', R_squared);

% 绘制已有数据
plot(years, lobsters, 'r. '); % 使用红色圆点表示实际数据
hold on; % 保持绘图窗口打开

% 在同一图表上绘制预测数据
plot(futureYears, predictedLobsters, 'b.'); % 使用蓝色圆点表示预测数据
```

```
hold off;

% 设置图表标题和轴标签
title('龙虾数量随时间的变化');
xlabel('年份');
ylabel('数量/千只');

% 显示预测结果
for i = 1:length(futureYears)
    fprintf('%d\t%.2f\n', futureYears(i)+1, predictedLobsters(i));
end
```