

社科研讨课——人口增长预测问题

杨隽祎，黄小科，郑伟一

May 2, 2022

1 Introduction

近年来人口增长成为一大热点，作为人口大国，对人口的预测影响着我国政策的制定与实施，进行恰当的建模与数据处理，在此切实应用的问题中，熟练对知识的掌握与应用。

2 Question

中国是一个人口大国，人口问题始终是制约我国发展的关键因素之一。根据已有数据，运用数学建模的方法，对中国人口做出分析和预测。

从中国的实际情况和人口增长的上述特点出发，参考相关数据，建立中国人口增长的数学模型，并由此对中国人口增长的中短期和长期趋势做出预测，特别要指出模型中的优点与不足之处。

3 Learning And Discussion

本题来源于实际问题，需要真实有效的数据。确定好方向后，我们三人便从网络寻找数据，例如近些年来人口总数，出生率，死亡率，各年龄段人口总数等。

此次我们做了稍微调整，讨论学习相关数据处理，学习讨论模型的合理性，多角度切入，在结果处给出模拟结果与真实情况的对比，以此评定模型的合理性，进而分析相关结论。

4 Difficulty

尽管有上学期运筹学建模题目的引导，我们组仍然面对诸多问题。

1. 模型建立较为多样，如何选择合适的模型。
2. 较为专业的内容难以涉猎。
3. 自身知识的有所欠缺，编程能力较为薄弱
4. 疫情影响，前期准备阶段组员之间线下联系受阻。

5 Solution

认识到上述问题，我们组讨论后决定，改变上学期分组方式，采取三人讨论给出若干相关模型，再由组员分别切入不同模型，同时编程处理数据，最终讨论各个模型的优劣。

其次我们数据均来自国家统计局（《中国统计年鉴》）。模型知识学习自姜启源老师的《数学模型》及其译制的《数学建模》。

6 Model one

6.1 model

模型：灰色动态模型-GM(1, 1)

将输入的数据记作一系列数据（记为： $x_0 = (x_0(1), \dots, x_0(n))$ ），引入累加，逆累加，均值，级比生成算子：

累加生成：将原序列的数据依次累加的到生成序列（记为： $x_1 = (x_1(1), \dots, x_1(n))$ ）

逆累加生成：将生成序列相邻数据作差生成新序列（记为： $y_1 = (y_1(1), \dots, y_1(n))$ ）

均值生成：生成序列相邻数据作均值生成新序列（记为： $z_1 = (z_1(2), \dots, z_1(n))$ ）

级比生成：原序列相邻数据作比生成新序列（记为： $\sigma = (\sigma(2), \dots, \sigma(n))$ ）

数学表达：

$$x_1(k) = \sum_{m=1}^k x_0(m) \quad y_1(k) = x_1(k) - x_1(k-1)$$

$$z_1(k) = \frac{(x_1(k) + x_1(k-1))}{2} \quad \sigma(k) = \frac{x_0(k-1)}{x_0(k)}$$

则定义序列间基本关系： $x_0(k) + az_1(k) = b$

其中设

$$P = \begin{pmatrix} a \\ b \end{pmatrix} \quad Y = \begin{pmatrix} x_0(2) \\ x_0(3) \\ \dots \\ x_0(n) \end{pmatrix} \quad B = \begin{pmatrix} -z_1(2) & 1 \\ -z_1(3) & 1 \\ \dots & \dots \\ -z_1(n) & 1 \end{pmatrix} \quad (1)$$

最小二乘法估计参数列： $\hat{P} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y$

利用离散数据序列建立近似的微分方程模型： $\frac{dx_1}{dt} + ax_1 = b$

解得时间相应函数： $x_1(t) = (x_0(1) - \frac{b}{a})e^{-at} + \frac{b}{a}$

由此可得原始数据的预测值：

$$\hat{x}_0(k+1) = \hat{x}_1(k+1) - \hat{x}_1(k) = (1 - e^{\hat{a}})(x_0(1) - \frac{\hat{b}}{\hat{a}})e^{-\hat{a}k}$$

精度检验引入残差，相对误差的定义：

残差： $q(k) = x_0(k) - \hat{x}_0(k)$

相对误差： $\epsilon(k) = \frac{q(k)}{x_0(k)} \times 100\%$

精度： $p^0 = (1 - \frac{1}{n-1} \sum_{k=2}^n |\epsilon(k)|) \times 100\%$

6.2 application

7 Model Two

8 Conclusion

Summarize your findings and add your comments here.

9 Reflection

从薄利多销问题到人口预测，两道题目均是运筹学范畴。无论是建模还是对数据的处理，误差分析，都要求我们对曲线的拟合讨论，这恰恰是二者最重要的关系。相比薄利多销问题，人口预测问题对模型处理，编程能力都提出了更高的要求。其中的Leslie模型更是能将现有的线代知识加以应用，深化认知。