**题目：关于中学生参加课后辅导班效果的研究**

# 【摘要】

经近期观察发现,课外补习班渐渐步入了中学生的生活。因人而异，各个同学或家长对课外辅导的认识不同，进而课外辅导所产生的效果也不同。本文从现阶段郑州市中学生所上课后辅导班种类和时长入手,结合中学生辅导前学习成绩等要素与辅导后成绩的提升幅度，建立灰色系统预测模型和多元线性回归预测模型来预测中学生参加课外辅导的时长，进而较客观的评判课外辅导对于提高中学生成绩的效率，寻找更为科学有效的学习方法来提高学习效率。

对于评判参加课外辅导班后的效果预测

1. 搜集统计郑州市2008-2017年的中学生参加课外辅导班时长及相关数据，初步建立灰色预测模型GM(1,1)对郑州市中学生参加课外辅导班后平均成绩（百分制）进行预测。
2. 在对灰色预测模型的分析和评价后，运用灰色关联分析筛选影响课外辅导效果（主要有参加辅导班前后成绩平均值判断）的各主要因素，建立精度更高的多元线性回归预测模型，对参加辅导班后成绩进行预测，并分析模型的准确性和实用性。

主要结论：

1. 灰色预测模型受多个因素的影响，不确定性因素较大，且模型拟合结果与2008-2017年中学生参加课后辅导班后实际成绩严重不符，故不适用于分析中学生参加课外辅导班后的效果预测。
2. 多元线性回归预测模型的最大相对误差为7.11%，平均相对误差为2.68%，相比灰色预测模型精度有了较大的提高，具有较高的准确性和实用性。
3. 多元线性回归预测模型：Y=2467.5-2.4x1+1.3x3-3.89x5+2.43x6
4. 根据参加辅导班前后成绩做图表对比，可以得知2008-2017年郑州市中学生在参加课外辅导班后成绩最大提升为29.02%，平均提升程度为7.06%。故课外辅导对于提升中学生的学习成绩和综合实力有效果。

关键词：GM（1，1），灰色系统预测模型，多元线性回归预测模型

# 问题重述

随着社会的不断发展,教育已经成为人们越来越关注的话题。许多家长都希望自己的孩子成龙成凤,因此,课外辅导如雨后春笋般应运而生,这些课外辅导受到中学生和家长的青睐。

家长们普遍认为课外辅导对中学生学业发展和综合素质的提高具有重要的积极促进作用,尤其是对中学生兴趣的培养,学习成绩的提高,家长们希望孩子发展成为全面型人才。然而，课后辅导班的效果不仅受辅导班的影响，还受中学生及家长的观念、社会及个人重视程度等多方面的交叉影响。因此，课后辅导班的效果很大程度上是因人而异的。但是在一定程度下也符合正态分布的函数曲线分布。

对于预测参加辅导班后成绩变化的预测：查阅相关文献，搜集郑州市中学生参加课外辅导班的情况数据，在此基础上建立中学生课后辅导效果中短期预测模型，并且分析模型的准确性和实用性。在预测模型的基础上，对参加辅导班后效果进行评价。

# 模型假设

1. 假设不考虑中学生及家长的重视程度等来自个人的主观因素
2. 假设不考虑各个地区对课后辅导的观念及重视程度，排除来自地区不同带来的主观干扰
3. 各辅导班的运行模式及教学形式均统一

# 符号约定

：原始非负时间序列

：生成序列

：影响垃圾产量因子集

：关联系数

：第j个变量对第i个变量的关联度

：回归系数

R：复相关系数

# 模型的建立

4.1问题分析

中学生参与辅导班的情况是一个信息总量极大且数据未统一的灰色数据，中学生个人的不确定性、分散性使中学生参与课外辅导的数据符合灰色数据特征，因此可采用灰色系统理论中的数列灰预测，建立灰色系统预测模型利用现阶段郑州市中学生所上课后辅导班的种类和时长对模型进行预测检验。针对灰色系统存在的缺陷进行改进，综合考虑影响垃圾产量的内在因素，利用灰色关联分析把主要因子从众多因素中找出来，通过主要因子对中学生参与课外辅导效率的影响多元线性回归分析，最终建立精度较高的多元线性回归预测模型。

4.2.1灰色系统预测模型

灰色系统GM(1，1)预测模型是一种常用的灰色模型，用于单个时间序列的预测，它通过对原始数据作累加处理后，由于有了较强规律的生成数据列，建立微分方程模型，具体做法如下：

设有原始非负时间序列，对进行一次累加生成，得到生成序列。

其中。

利用数列建立白化微分方程:

利用最小二乘法求参数a，u



其中，，

的灰色预测模型为:

对作一次累减生成，即得的灰色预测模型为



为了判断模型的优劣，还应进行模型精度检验，精度检验一般有残差检验、关联度检验和后验差检验等三种方式，若通过检验，则可利用所建立的模型进行预测，否则应进行残差修正。

4.2.2模型求解

运用MATLAB编程实现（程序见附录1），对郑州市中学生2008-2017年参加课后辅导后成绩进行预测，结果如下：

表1GM（1，1）预测结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 年份 | 参加辅导前成绩  （百分制） | 参加辅导后预测成绩  （百分制） | 参加辅导预测成绩  （百分制） | 相对误差% |
| 2008 | 47.65 | 50.18 | 55 | 14.27 |
| 2009 | 46.65 | 53.05 | 55.51 | 3.9 |
| 2010 | 50.03 | 64.55 | 59.74 | -7.36 |
| 2011 | 55.02 | 64.98 | 55.77 | -14.47 |
| 2012 | 62.34 | 67.47 | 72.64 | 7.62 |
| 2013 | 60.69 | 65.92 | 66.64 | 1.78 |
| 2014 | 59.2 | 61.45 | 69.02 | 13.48 |
| 2015 | 68.22 | 62.36 | 68.82 | 9.66 |
| 2016 | 67.77 | 68.38 | 67.97 | 0.58 |
| 2017 | 65.21 | 59.66 | 55 | 8.67 |

由上表可知，灰色系统预测模型的相对误差波动较大，其中2011年的相对误差达到了14.47%。这主要是由于影响中学生参加课外辅导效果的不确定因素变化所致。

4.2.3残差检验

Step1：生成数列误差（残差）检验：



Step2：还原数列检验：

根据，得到第k年中学生参加课外辅导时长。与实际值的相对误差即为残差，从表1看出GM（1，1）模型对少数年份的残差偏大。

4.3.1模型改进

灰色系统预测模型仅从现阶段郑州市中学生所上课后辅导班的种类和时长这个综合灰色量本身挖掘有用信息，利用它的动态记忆特征，建立灰色模型来寻找中学生参加课后辅导班效果的内在规律，而忽略了其他因素对辅导班效果的影响，因此有一定的缺陷。模型用于短期预测精度可达较高，但用于中长期的预测精度则会有所下降，难以令人满意。因此为了使模型预测的精度能够在中短期内均保持较高精度，下面我们通过灰色关联分析把影响辅导班效率的主要因子从众多因素中找出来，再利用多元线性回归分析方法建立模型，预测中学生参加课后辅导班后的效果，以期能够达到较高精度。

4.3.2灰关联分析

灰色关联分析的基本思想是根据序列曲线的几何形状的相似程度来判断其联系是否紧密．曲线越接近。相应序列之间关联度就越大，反之就越小。灰色关联的表示方法很多，一般灰色斜率关联度的分辨率较高而经常被使用。斜关联度分析首先要确定变量斜率：

（1）

式中：n为变量个数；m为样本容量；为标准差，，为的均值，关联系数用式（2）计算：

（2）

采用等权平均值法计算关联度：（3）

式中：为第j个变量对第i个变量的关联度，反映了2条曲线变化率的平均相似程度。根据对的大小排序，挑选影响系统的主要因子。

4.3.3关联度计算及影响因素选取

中学生参加辅导班的效果的影响因素主要有学生个人、家长和社会3个方面，结合郑州市当代中学生的实际情况，选取在读中学生人数、学生参与课外辅导班年费用、学生年在校学习时长、社会对学校的投资、家庭年收入、参与辅导班人数等6个因素。用表2汇总郑州市2008-2017年度各年的学生参与辅导班后成绩的变化及各主要影响因子数据，同时，求得各项指标与参加辅导班后成绩的线性相关系数。

表2郑州市2008-2017年度各年参加辅导班后成绩平均值及各主要影响因素

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | X0 | X1 | X2 | X3 | X4 | X5 | X6 |
|  | 参与课外辅导班后成绩（百分制） | 在读中学生人数（万人） | 学生参与课外辅导班年费用（元） | 学生年在校学习时长（h） | 社会对学校的投资（万元） | 家庭平均年收入（万元） | 参与辅导班学生人数（万人） |
| 2008 | 55.51 | 103.65 | 3688.2 | 1471.03 | 6866 | 19.1 | 102.13 |
| 2009 | 59.74 | 109.63 | 4034.96 | 1590.38 | 8248 | 20.3 | 111.57 |
| 2010 | 55.77 | 116.16 | 4551.15 | 1722.27 | 8868 | 22.2 | 141.91 |
| 2011 | 72.64 | 109.07 | 4950.84 | 1861.3 | 9336 | 23.8 | 152.93 |
| 2012 | 66.64 | 128.81 | 5408.76 | 2035.21 | 10464 | 30 | 162.39 |
| 2013 | 69.02 | 131.39 | 6250.81 | 2220.64 | 11040 | 35.2 | 191.53 |
| 2014 | 68.82 | 127.6 | 7450.27 | 2454.61 | 12631 | 36 | 225.4 |
| 2015 | 67.97 | 134 | 9154.18 | 3387.99 | 13527 | 37.1 | 295.1 |
| 2016 | 55 | 129 | 10366.3 | 3660.42 | 14761 | 37.3 | 346.45 |
| 2017 | 55.51 | 128 | 12188.85 | 3847.79 | 17255 | 37.6 | 366.1 |

运用MATLAB编程（程序见附录2），以郑州市中学生参与辅导班后预测成绩为母序列，各影响因子为子序列，并作初始化处理。用公式(2)得关联系数，用公式(3)得关联度。关联度见表3。

表3各相关因素关联度数

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 项目 | 参加课外辅导班后成绩（百分制） | 在读中学生人数（万人） | 学生参与课外辅导班年费用（元） | 学生年在校学习时长（h） | 社会对学校的投资（万元） | 家庭年收入（万元） | 参与辅导班学生人数（万人） |
| rij | 1 | 0.6705 | 0.6674 | 0.6676 | 0.6673 | 0.6811 | 0.6701 |

比较得到：；所以采用与辅导班后预测成绩具有较好的线性关系的4个因子，即在读中学生人数、学生年在校学习时长、家庭年收入、参与辅导班学生人数作为郑州市中学生参与辅导班时长预测指标。

4.3.4多元线性回归分析及模型建立

多元线性回归模型可表示为:

（4）

式中：y为可观测的随机变量；，，…，为未知参数(回归系数)：为不可观测的随机误差。

设有n组独立观测值，则式四可表示为：

（5）

用矩阵表示：（6）

利用最小二乘法求回归系数的估计值：

（7）

回归系数确定后，即求得回归方程：（8）

4.3.5模型求解

通过MATLAB编程实现模型的求解（程序见附录1.3），取表1中2008~2014年x5,x1,x6,x3数据作为求解数据带入式7，求解回归系数，=3053.5，=-2.6，=-0.6，=-10.7，=10.5；得到多元线性回归预测模型如下：

Y=2467.5-2.4x1+1.3x3-3.89x5+2.43x6

式中：Y为参加辅导班后预测成绩（百分制）；x1为在读中学生人数、x3为学生年在校学习时长、x5为家庭年收入、x6为参与辅导班学生人数。

取表1中2015~2017年x1，x3，x5，x6数据作为检验数据带入模型，预测中学生参加课外辅导后成绩提升情况,对郑州市中学生2008~2017年参加课外辅导后实际成绩和预测成绩（百分制）作比较，其拟合结果见图1：

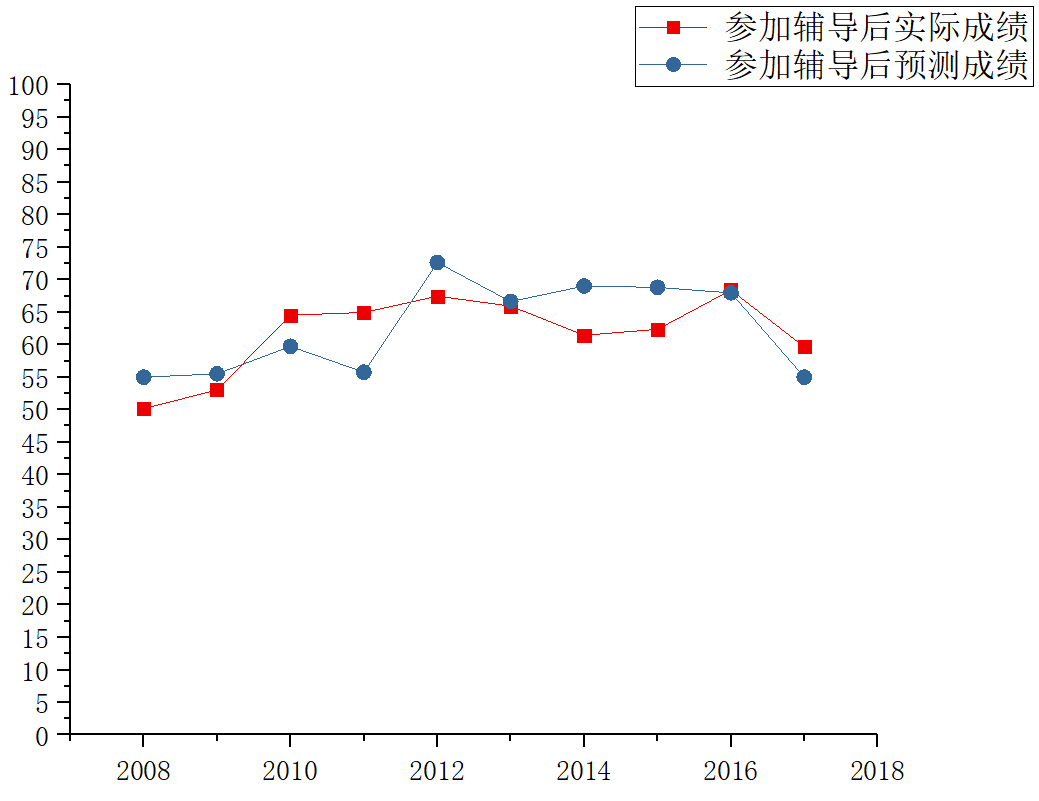


图1GM(1,1)模型拟合效果图

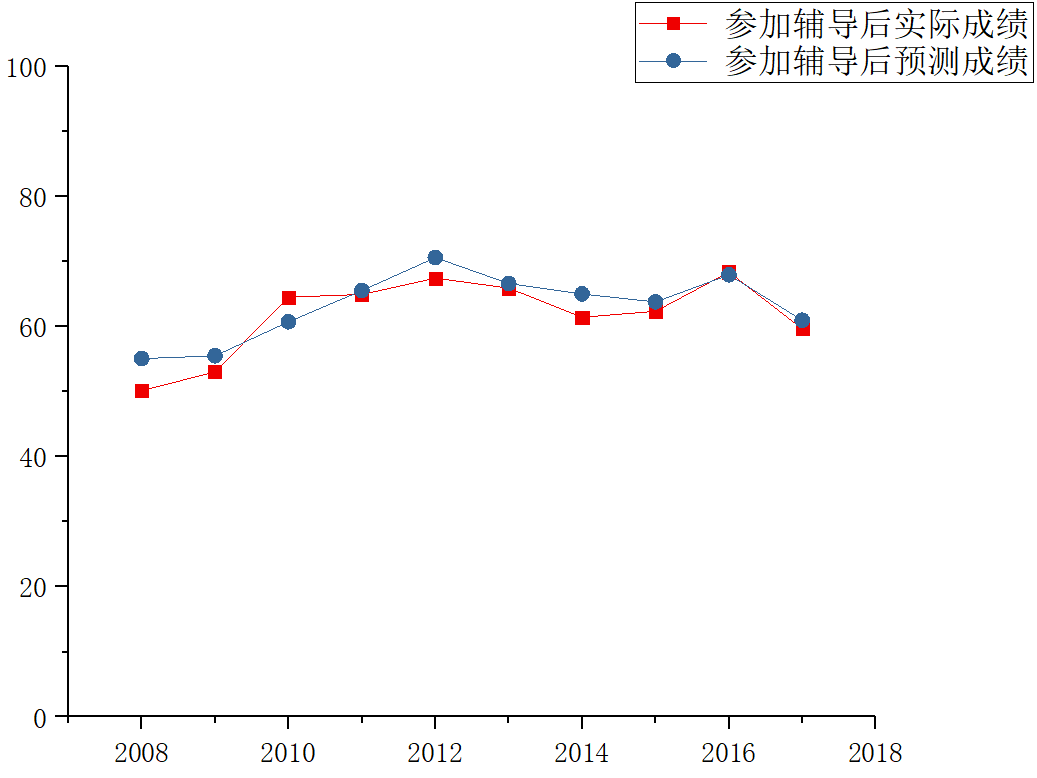


图2多元线性回归模型

对灰色系统预测模型和多元线性回归预测模型的预测结果进行比较，见表4：

表4模型预测结果数据比较

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 年份 | 参加辅导班后实际成绩（百分制） | GM（1，1）模型 | | 多元线性回归模型 | |
| 参加辅导班后预测成绩（百分制） | 相对误差（%） | 参加辅导班后预测成绩（百分制） | 相对误差（%） |
| 2008 | 50.18 | 55 | 14.27 | 55.51 | 6.08 |
| 2009 | 53.05 | 55.51 | 3.9 | 60.74 | 3.43 |
| 2010 | 64.55 | 59.74 | -7.36 | 65.57 | -7.11 |
| 2011 | 64.98 | 55.77 | -14.47 | 70.64 | 0.92 |
| 2012 | 67.47 | 72.64 | 7.62 | 66.64 | 5.57 |
| 2013 | 65.92 | 66.64 | 1.78 | 65.02 | -3.28 |
| 2014 | 61.45 | 69.02 | 13.48 | 63.82 | 1.12 |
| 2015 | 62.36 | 68.82 | 9.66 | 67.97 | 3.28 |
| 2016 | 68.38 | 67.97 | 0.58 | 60.97 | 1.01 |
| 2017 | 59.66 | 55 | 14.27 | 60.97 | 1.53 |

从图1和表4，我们可以看出：多元线性回归预测模型预测的最大相对误差为7.11%，通过比较多元线性回归预测模型具备了灰色系统预测模型所不具备的优越性，它综合考虑了影响郑州市中学生参加课外辅导效果的内因和外因，利用原始数据建立预测模型，预测结果具有较高的精度，平均相对误差为2.68%，一般来说，只要模型的C<0.35,p>0.95,则模型精度达到一级，因此该模型适用于中学生参加课外辅导后的效果预测，能够达到令人满意的效果。

4.3.5模型检验

1. F检验，构造检验回归方程的统计量

（9）

，，式中为样本y的平均值。

在给定显著性水平α的情况下，若,则回归方程显著性成立，否者认为回归方程不显著，模型不能用于预测。通过MATLAB编程求解，程序见附录1.4。

1. 复相关系数R检验，检验式为：，。

R越接近于1，表示回归方程的拟合程度越好。

取显著性水平α=0.05，通过查表得F（4，5）=6.26<F=23.18，回归方程通过显著性检验；将数据带入进行复相关系数检验，R=0.9741接近于1，回归方程拟合程度好。检验结果如下表：

表 5多元线性回归模型检验结果

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 回归系数 | | | | | 检验值 | | |
| β1 | β2 | β3 | β4 | β5 | 复项关系灵敏R | R2 | F值 |
| 3053.5 | -2.6 | -0.6 | -10.7 | 10.5 | 0.9741 | 0.9489 | 23.18 |

4.3.6参加辅导班效果评价

根据表1中2008-2017年参加课外辅导前后的实际成绩相对比，作图3、表6.由图3可以直观的得知：参加辅导班后有明显的效果。根据表3，可知在参加辅导班后最大提升程度为29.02%，平均提升程度为7.06%。故参加辅导班对于提升学生的学习成绩有明显的效果。

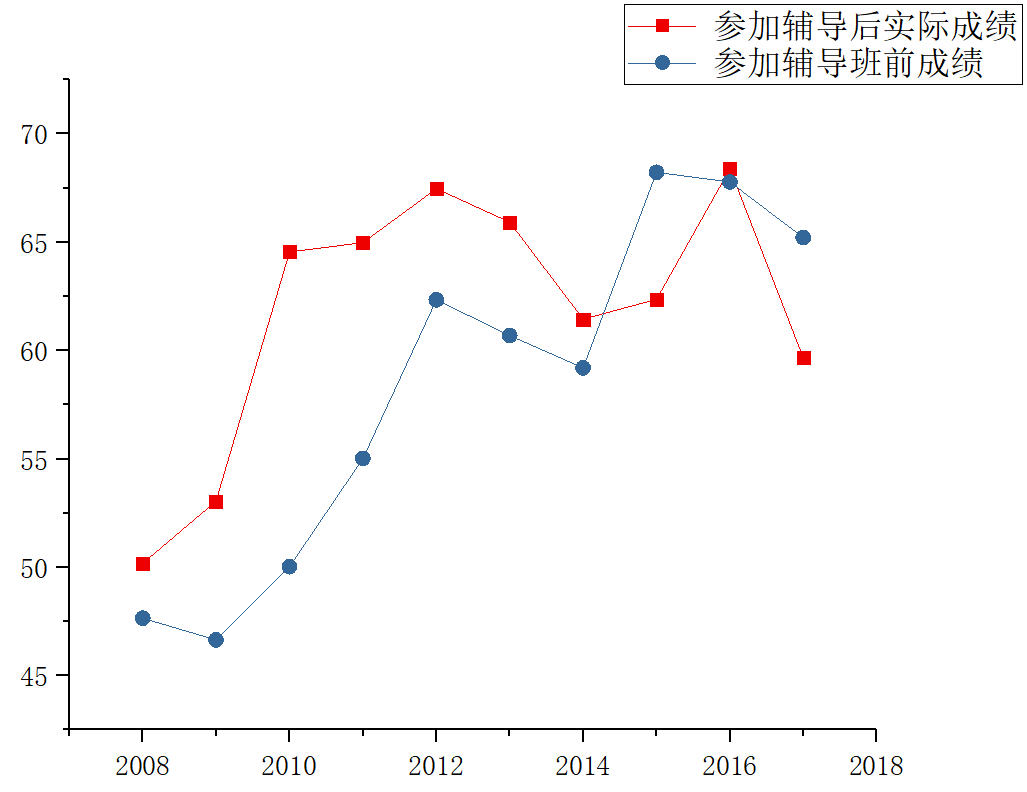


图 3参加课外辅导前后成绩对比

表 6参加课外辅导前后成绩对比

|  |  |  |  |
| --- | --- | --- | --- |
| 年份 | 参加辅导班后成绩（百分制） | 参加辅导班前成绩（百分制） | 提升程度  % |
| 2008 | 50.18 | 47.65 | 5.31% |
| 2009 | 53.05 | 46.65 | 13.72% |
| 2010 | 64.55 | 50.03 | 29.02% |
| 2011 | 64.98 | 55.02 | 18.10% |
| 2012 | 67.47 | 62.34 | 8.23% |
| 2013 | 65.92 | 60.69 | 8.62% |
| 2014 | 61.45 | 59.20 | 3.80% |
| 2015 | 62.36 | 68.22 | -8.59% |
| 2016 | 68.38 | 67.77 | 0.90% |
| 2017 | 59.66 | 65.21 | -8.51% |

# 模型的评价及推广

6.1多元线性回归模型评价：

模型的优点：模型综合考虑了影响中学生参加课后辅导后成绩变化的主要因子，通过建立多元线性方程分析确定各因子的相关系数，使得多元线形回归预测模型在中学生参加课外辅导班后效果预测中具有其他模型不具备的优越性，其预测精度与其他模型相比有很大的提高，得到令人较为满意的数据结果，证明了它在中学生参加课外辅导后效果预测方面的可行性。另外，模型实现过程，运算过程可采用MATLAB软件编程，方便灵活简单，便于操作。

模型的缺点：该模型不具备自组织自适应和自学习能力，受突发事件影响比较大，该模型只能适用于正常情况下中学生参加课后辅导后的一般成绩变化。

模型的推广：该模型适用多个地区的预测，只需改变输入变量数，将数据更改为适合其他地区发展的数据，便可对其他地区进行预测，适用范围广；其次，该模型适用多个参数的预测，不止对中学生参加课后辅导效果的预测有效，只需改变其内部影响因子，便可实现对其他预测问题的解决。

# 参考文献

1. 郑州市初中生校外英语辅导班现状调查研究[D].郑州大学,2015.
2. 王玲.中学生参加课外辅导现状调查及建议——基于包头市昆区两所中学的个案研究[D].内蒙古师范大学,2013.
3. 蔡丽娟.课外辅导情况下的高效数学教学[J].中学生数理化(学习研究),2013(6):81-81.
4. 李爱华. 小学生参加课外辅导班存在的问题及解决策略[J]. 中小学教师培训, 2012(11):60-61.
5. 曾莉[1].校外英语辅导班对中学生英语学习的影响研究[J].校园英语,2017(35):170-170.
6. 中学生参加数学课外辅导班学习的案例研究[D].2014.

# 相关附录

附录1

functionxk1=greyforecast(A,k)%灰色预测模型

Yn=A(2:k);

%做累加处理

A1(1)=A(1);

fori=2:k

A1(i)=A1(i-1)+A(i);

end

fori=1:k-1

A2(i)=-1/2\*(A1(i)+A1(i+1));

end

B=[A2;ones(1,k-1)];

B=B';

au=inv((B'\*B))\*B'\*Yn';

xk1=(A(1)-au(2)/au(1))\*exp(-au(1)\*k)+au(2)/au(1);

p=[50 53 64 64 67 55 61];%调用程序对第8，9和10年进行预测

fori=8:10

p(i)=greyforecast(p(1:i-1),i-1)-greyforecast(p(1:i-2),i-2);

end

>>[s]

ans=

Columns 1 through 7

470.0000 500.0000 641.0000 644.0000 467.0000 585.0000

610.0000

Columns 8 through 10

638.8257 670.9732 705.9709

附录2

yangben=[55 55.51 59.54 65.57 50.64 56.6 45.02 64.52 67.97 54.97];%10年参加辅导班前成绩平均值

%6个因素各年的值

fangzhen=[921.7 2462.97 970.04 5868 16 46.49;

932.14 2902.2 1161.3 6763 17 68.83;

943.03 3360.21 1325.21 6820 17.8 82.35;

953.65 3688.2 1471.03 6866 19.1 102.13;

969.63 4034.96 1590.38 8248 20.3 111.57;

986.16 4551.15 1722.27 8868 22.2 141.91;

999.07 4950.84 1861.3 9336 23.8 152.93;

1018.81 5408.76 2035.21 10464 30 162.39;

1041.39 6250.81 2220.64 11040 35.2 191.53;

1097.6 7450.27 2454.61 12631 36 225.4];

%求各因素的灰色关联度程序，pp1，pp2，pp3，pp4，pp5，pp6即为其

值

f0=std(yangben);

fori=1:6

f(i)=var(fangzhen(:,i));

end

fori=1:9

k0(i)=abs(yangben(i+1)-yangben(i))/f0;

k1(i)=abs(fangzhen(i+1,1)-fangzhen(i,1))/f(1);

k2(i)=abs(fangzhen(i+1,2)-fangzhen(i,2))/f(2);

k3(i)=abs(fangzhen(i+1,3)-fangzhen(i,3))/f(3);

k4(i)=abs(fangzhen(i+1,4)-fangzhen(i,4))/f(4);

k5(i)=abs(fangzhen(i+1,5)-fangzhen(i,5))/f(5);

k6(i)=abs(fangzhen(i+1,6)-fangzhen(i,6))/f(6);

end

p1=1./(1+abs(k1-k0));

p2=1./(1+abs(k2-k0));

p3=1./(1+abs(k3-k0));

p4=1./(1+abs(k4-k0));

p5=1./(1+abs(k5-k0));

p6=1./(1+abs(k6-k0));

pp1=mean(p1);

pp2=mean(p2);

pp3=mean(p3);

pp4=mean(p4);

pp5=mean(p5);

pp6=mean(p6);

>>[pp1pp2pp3pp4pp5pp6]%灰色关联度结果

ans =

0.6705 0.6674 0.6676 0.6673 0.6811 0.6701

附录3

%多元线性回归模型程序

x=[1 921.7 970.04 16 46.49;

1 932.14 1161.3 17 68.83;

1 943.03 1325.21 17.8 82.35;

1 953.65 1471.03 19.1 102.13;

1 969.63 1590.38 20.3 111.57;

1 986.16 1722.27 22.2 141.91;

1 999.07 1861.3 23.8 152.93;

1 1018.81 2035.21 30 162.39;

1 1041.39 2220.64 35.2 191.53;

1 1097.6 2454.61 36 225.4];

y=[372419454470500641644467585610];%10年预测值

B=inv(x'\*x)\*x'\*y';%回归系数估值

Y1=x\*B；%各年的估测值

[B]

ans=1.0e+003\*

3.0535

-0.0026

-0.0006

-0.0107

0.0105

>>[Y1]

ans=55.00

55.51

59.74

65.77

50.64

56.64

65.02

64.82

67.97

54.97

%多元回归预测2015，2016，2017年参加辅导班前平均成绩

xl=[1 1134 3387.99 37.1 295.1

1 1265 3660.42 37.33 346.45

1 1283 3847.79 37.6 366.1];

Y2=xl\*B；

附录4

%多元回归模型F值检验求F值程序

y1=x\*B;y=y';p=4;n=10;

yp=mean(y);

u=sum((y1-yp).^2);

q=sum((y-y1).^2);

F=(u/p)/(q/(n-p-1));

%复相关系数R检验求R模型

y1=x\*B;

yp=mean(y);

u=sum((y1-yp).^2);

l=sum((y-yp).^2);

R=sqrt(u/l);