

# 中国农村居民“十一五”期间 收入和消费的灰色预测

相丽驰<sup>1</sup>, 金晓彤<sup>2</sup>

(1. 浙江万里学院, 浙江宁波 315100 2. 吉林大学, 长春 130012)

**摘要:** 本文利用 2000年 - 2005年我国农村居民人均纯收入和人均生活消费数据, 采用灰色预测理论, 分别建立了收入和消费 GM (1, 1) 灰色模型, 依此对“十一五”期间我国农民的纯收入和生活消费进行预测。模型检验的结果表明, 模型精度高, 从而获得的预测结果具有很高的可信度。从预测数据可见, “十一五”期间我国农民的收入、消费呈现攀升的态势, 但城乡收入差别问题仍然凸现。

**关键词:** 灰色系统; 农民; 收入; 消费; 预测

我国农村居民收入的提高和消费的扩大是促进农村经济可持续发展的重要环节。对“十一五”期间农民收入和消费的走势进行预测, 无疑对国家制定宏观经济政策等诸多方面具有重要意义。灰色预测是指对既含有已知信息又含有不确定信息的系统进行预测, 经济系统其系统本身和系统环境、系统内部和系统外部边界若明若暗, 呈现这一特点, 用灰色系统理论进行研究能够收到很好的效果。灰预测的独特之处在于所需数据少、预测剪度高、能够修正。它可以弥补在计量经济学基础上建立模型的一些不足, 如: 线性模型难以把握非线性现象、指数平滑法对上升数据预测偏低、ARMA 模型需要数据较多等。本文将采用灰预测方法对“十一五”期间我国农民收入和消费进行预测。

## 一、灰色预测模型

### (一) 灰色预测模型 GM (1, 1)<sup>[1][2]</sup>

灰色预测模型 GM (Grey Model) 包括一阶单变量的 GM (1, 1) 模型和 n 阶 h 个变量的 GM (n, h) 模型。它兼有微分方程、差分方程和指数方程的特性。一般常用的是 GM (1, 1) 模型。

设  $X^{(0)}$  为 GM (1, 1) 的建模序列,  $X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$ , 令  $X^{(1)}$  为  $X^{(0)}$  的一次累加生成序列,  $X^{(1)}(k) = \sum_{i=1}^k X^{(0)}(i)$  ( $i = 1, 2, \dots, n$ ), 令  $Z^{(1)}(k) = 0.5X^{(1)}(k) + 0.5X^{(1)}(k-1)$ 。

由  $X^{(1)}$  建立白微分方程:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu \quad (1)$$

$a, \mu$  分别称为发展灰数和灰作用量。

利用最小二乘法可得:

$$\alpha = \begin{pmatrix} a \\ \mu \end{pmatrix} = (B^T B)^{-1} B^T Y_n$$

其中:  $B = \begin{pmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \dots & \dots \\ -Z^{(1)}(n) & 1 \end{pmatrix}$   $Y_n = \begin{pmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \dots \\ X^{(0)}(n) \end{pmatrix}$

对白化微分方程求解, 可得响应式为:

$$X^{(1)}(k+1) = [X^{(0)}(1) - \frac{\mu}{a}] e^{-ak} + \frac{\mu}{a} \quad (k=0, 1, 2, \dots, \Delta)$$

累减生成预测值  $X^{(0)}(k+1) = X^{(1)}(k+1) - X^{(1)}(k)$   $X^{(0)}(1) = X^{(1)}(1)$

### (二) GM (1, 1) 模型检验<sup>[3]</sup>

1. 事前检验。事前检验是对原始数据序列  $X^{(0)}$  作 GM (1, 1) 建模的可行性检验。当级比  $\sigma^{(0)}(k) = \frac{X^{(0)}(k)}{X^{(0)}(k-1)}$  值落在可容区为  $(e^{-2}, e^2)$  内, 即  $\sigma^{(0)}(k) \in (0.1353, 7.389)$ , 表明  $X^{(0)}$  是光滑的, 可作灰预测。若  $\sigma^{(0)}(k) \in (e^{-\frac{1}{n}}, e^{\frac{1}{n}})$ , 则可获得精度较高的 GM (1, 1) 模型。当灰发展系数  $-a \leq 0.3$  GM (1, 1) 模型可以用于中长期预测。

2. 事中检验。事中检验是对 GM (1, 1) 模型的精度检验。一般采用残差检验、后验差检验、级比偏差值检验。

残差检验 绝对残差  $\Delta^{(0)} = X^{(0)}(k) - X^{(0)}(k)$ ,

**作者简介:** 相丽驰 (1964年 - ), 女, 吉林省长春市人, 浙江万里学院基础学院, 副教授, 学士。研究方向为经济数学方法。金晓彤 (1964年 - ), 女, 吉林省长春市人, 吉林大学商学院教授, 博士生导师, 经济学博士。研究方向为商业经济学。

\* **基金项目:** 本文为“浙江省社会科学界联合会 2006 研究课题” (项目编号 06N93)、“2004 年度国家社会科学基金项目” (项目编号 04BJY047) 的阶段性研究成果

相对残差  $\epsilon(k) = \frac{\Delta^{(0)}(k)}{X^{(0)}(k)} \times 100\%$ , 平均残差  $\epsilon(\text{avg})$

$$= \frac{1}{n-1} \sum_{k=2}^n |\epsilon(k)| \times 100\%, \text{ 平均精度 } p^0 = (1 - \epsilon(\text{avg})) \times 100\%.$$

一般要求  $\epsilon(k) < 20\%$ , 最好是  $\epsilon(k) < 10\%$ ; 一般要求  $p^0 > 80\%$ , 最好是  $p^0 > 90\%$ 。

后验差检验 原始序列的标准差:

$$S_1 = \sqrt{\frac{\sum [X^{(0)}(k) - \bar{X}^{(0)}]^2}{n-1}}$$

绝对残差序列的标准差:

$$S_2 = \sqrt{\frac{\sum [\Delta^{(0)}(k) - \bar{\Delta}^{(0)}]^2}{n-1}}$$

方差比  $C = \frac{S_2}{S_1}$ , 小误差概率  $P = p\{\Delta^{(0)}(k) - \bar{\Delta}^{(0)} | < S_0\}$  其中  $S_0 = 0.6745S_1$ 。

一般要求  $0.80 < P < 0.95$   $0.35 < C < 0.5$  最好是  $P > 0.95$   $C < 0.35$ 。

级比偏差值检验 级比偏差:

$$\rho(k) = 1 - \left(\frac{1 - 0.5a}{1 + 0.5a}\right) \frac{X^{(0)}(k-1)}{X^{(0)}(k)} 100\% = 1 - u\sigma^{(0)}(k) 100\%$$

一般要求  $\rho(k) < 20\%$ , 最好是  $\rho(k) < 10\%$ 。

3. 事后检验。事后检验是预测检验。一般可采用滚动检验。令  $X^{(0)}$  原始序列,  $X_i^{(0)}$  为建模序列  $X^{(0)}$  的 4 维新陈代谢子列。如  $X_4^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4)\}$ ,  $X_5^{(0)} = \{X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5)\}$  ..... 对  $X_i^{(0)}$  作 GM (1, 1) 建模。

滚动残差:

$$\epsilon_{run}(i+1) = \frac{X^{(0)}(i+1) - X^{(0)}(i+1)}{X^{(0)}(i+1)} \times 100\%$$

$$\text{平均滚动残差: } \epsilon_{run}(\text{avg}) = \frac{1}{n-4} \sum_{i=4}^{n-1} |\epsilon(i+1)| \times 100\%$$

平均滚动精度:  $p_{run}^0 = (1 - \epsilon_{run}(\text{avg})) \times 100\%$

一般要求  $p_{run}^0 > 80\%$ , 最好是  $p_{run}^0 > 90\%$ 。

(三) 新陈代谢 GM (1, 1) 模型<sup>[4]</sup>

为了充分利用新信息的作用, 进一步提高精度, 可以建立新陈代谢 GM (1, 1) 模型, 即由已经建立的 GM (1, 1) 预测一个值, 补充到已知序列中, 同时去掉一个最老的信息, 再继续建模预测一个值, 如此增加新信息去掉旧信息。

根据以上建模思想方法和表一数据资料, 分别建立我国农民收入和消费的灰色预测模型。

表一 2000年 - 2005年中国农村居民家庭平均每人纯收入和生活消费支出 单位: 元

年份	2000	2001	2002	2003	2004	2005
纯收入	2253	2366	2476	2622	2936	3255
生活消费	1670	1741	1834	1943	2185	2555

资料来自中国统计年鉴、2005 统计公报

$$26316.02e^{0.083527k} - 24063.02 \quad (3)$$

## 二、我国农民收入 GM (1, 1) 模型建立及检验

(一) 收入 GM (1, 1) 模型建立

设 2000年 - 2005年纯收入数据为原始序列  $X^{(0)} = \{2253 \ 2366 \ 2476 \ 2622 \ 2936 \ 3255\}$ , 生成新序列  $X^{(1)} = \{2253 \ 4619 \ 7095 \ 9717 \ 12653 \ 15908\}$ , 均值序列  $Z^{(1)} = \{3436 \ 5857 \ 8406 \ 11185 \ 14280.5\}$ 。根据 GM (1, 1) 建模方法得  $\alpha = \begin{bmatrix} a \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} -0.083527 \\ 2009.918 \end{bmatrix}$ , 确定白化模型方程为:

$$\frac{dX^{(1)}}{dt} - 0.083527X^{(1)} = 2009.918$$

依据式 (2) 确定响应式为:

$$X^{(1)}(k+1) = [X^{(0)}(1) - \frac{\mu}{a}] e^{-ak} + \frac{\mu}{a} =$$

(二) 收入 GM (1, 1) 模型检验

依据模型检验的方法和公式, 得到如下系列结果。

1. 事前检验。级比  $\sigma^{(0)} = \{0.952 \ 0.956 \ 0.944 \ 0.893 \ 0.902\}$ ,  $n = 6$   $\sigma^{(0)}(k) \in (e^{-\frac{a}{2}}, e^{\frac{a}{2}})$  即界区为  $(0.7515 \ 1.3307)$ 。级比都落在界区内, 可以采用 GM (1, 1) 模型且可获得精度较高的模型。灰发展系数  $-a = 0.083527 < 0.3$  可以进行中长期预测。

2. 事中检验。残差检验 根据式 (3) 计算纯收入预测序列  $X^{(1)}, X^{(0)}$  进一步得  $\Delta^{(0)}(k)\epsilon(k)$ , 结果见表二。

从而  $\epsilon(k) < 10\%$ ,  $\epsilon(\text{avg}) = 1.803\% < 5\%$ ,  $p^0 = 98.17\% > 90\%$ , 模型精度好, 通过残差检验。

表二 农民纯收入残差检验精度表

年份	$X^{(0)}$	$X^{(1)}$	$X^{(1)}$	$X^{(0)}$	绝对残差 $\Delta^{(0)}(k)$ $X^{(0)}(k)X^{(0)}(k)$	相对残差 $\epsilon(k)$ $\Delta^{(0)}(k) X^{(0)}(k)$
2000	2253	2253	2253.00	2253.00	0.00	0.00000
2001	2366	4619	4545.52	2292.52	73.48	0.03105
2002	2476	7095	7037.76	2492.24	-16.24	-0.00656

2003	2622	9717	9747.11	2709.35	- 87.35	- 0.03331
2004	2936	12653	12692.49	2945.38	- 9.38	- 0.00319
2005	3255	15908	15894.45	3201.96	53.04	0.01629

后验差检验  $S_1 = 378.72$   $\bar{\Delta}^{(0)} = 2.26$   $S_2 = 56.86$   $C = 0.1501 < 0.35$ 。由于  $S_0 = 255.45$  而  $|\Delta^{(0)}(i) - \bar{\Delta}^{(0)}| = \{2.26, 71.22, 18.50, 89.61, 11.63, 50.78\}$  各个值均小于  $S_0$ ，所以  $P = 1 > 0.95$ ，满足  $P > 0.95$   $C < 0.35$  模型精度好，通过后验差检验。

级比偏差值检验 由  $\rho(k) = 1 - u\sigma^{(0)}(k)$  得  $\rho = \{-0.0352, -0.0389, -0.0266, 0.0291, 0.0194\}$ ，有各级比偏差值  $\rho(k) < 10\%$ ，模型精度好，通过级比偏差检验。

3. 事后检验。用  $X_4^{(0)} = \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4)\}$  建模得预测值  $X^{(0)}(5) = 2755.2$  用  $X_5^{(0)} = \{X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5)\}$  建模得预测值  $X^{(0)}(6) = 3175.1$ 。可得  $\epsilon_{min}(5) = 0.0616$   $\epsilon_{min}(6) = 0.0245$   $\epsilon_{min}(avg) = 4.30\%$   $p_{min}^0 = 95.70\% > 90\%$ ，预测精度好，预测模型具有 95.7% 可信程度。

### 三、我国农民消费 GM (1, 1) 模型及检验

表三 农民消费的残差检验精度表

年份	$X^{(0)}$	$X^{(1)}$	$X^{(4)}$	$X^{(0)}$	绝对残差 $\Delta^{(0)}(k)$ $X^{(0)}(k)X^{(0)}(k)$	相对残差 $\epsilon(k)$ $\Delta^{(0)}(k) X^{(0)}(k)$
2000	1670	1670	1670.00	1670.00	0.00	0.00000
2001	1741	3411	3332.17	1662.17	78.83	0.04528
2002	1834	5245	5168.14	1835.97	- 1.97	- 0.00108
2003	1943	7188	7196.09	2027.95	- 84.95	- 0.04372
2004	2185	9373	9436.09	2240.00	- 55.00	- 0.02517
2005	2555	11928	11910.31	2474.22	80.78	0.03162

$\epsilon(avg) = 2.93\% < 5\%$ ， $p^0 = 97.07\% > 90\%$ ，通过残差检验。

后验差检验:  $S_1 = 348.76$   $S_2 = 67.72$   $C = \frac{S_2}{S_1} = 0.1942 < 0.35$  小误差概率  $P = 1 > 0.95$  通过后验差检验。

级比偏差值检验:  $\rho = \{-0.0596, -0.0486, -0.0427, 0.0177, 0.0553\}$ ，有  $\rho(k) < 10\%$  成立，通过级比偏差检验。

3. 事后检验。由 4 维子序列预测  $X^{(0)}(5) = 2050.6$   $X^{(0)}(6) = 2367.3$   $\epsilon_{min}(5) = 0.0615$   $\epsilon_{min}(6) = 0.0730$   $\epsilon_{min}(avg) = 6.72\%$ ， $p_{min}^0 = 93.28\% > 90\%$ ，预测精度好，预测模型具有 93.28% 可信程度。

四、基于 GM (1, 1) 模型的农民收入和消费预测

(一) 消费 GM (1, 1) 模型建立  
根据 2000 年 - 2005 年消费数据和 GM (1, 1) 模型建立方法，求得  $\alpha = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_n =$

$$\begin{bmatrix} -0.099450 \\ 1414.799 \end{bmatrix}$$

从而确定 GM (1, 1) 白化模型方程：  
 $\frac{dX^{(1)}}{dt} - 0.099450X^{(1)} = 1414.799$

白化响应式为：  
 $X^{(4)}(k+1) = [X^{(0)}(1) - \frac{\mu}{a}]e^{-ak} + \frac{\mu}{a} = 15896.20e^{0.099450k} - 14226.20$  (4)

(二) 消费 GM (1, 1) 模型检验  
1. 事前检验。级比  $\sigma^{(0)} = \{0.959, 0.949, 0.944, 0.889, 0.855\}$  在界区中，可以建立 GM (1, 1) 模型。  
2. 事中检验。残差检验 根据式 (4) 计算消费的预测序列  $X^{(4)}, X^{(0)}$  进而得  $\Delta^{(0)}(k)\epsilon(k)$ ，结果见表三。

$\epsilon(avg) = 2.93\% < 5\%$ ， $p^0 = 97.07\% > 90\%$ ，通过残差检验。

后验差检验:  $S_1 = 348.76$   $S_2 = 67.72$   $C = \frac{S_2}{S_1} = 0.1942 < 0.35$  小误差概率  $P = 1 > 0.95$  通过后验差检验。

级比偏差值检验:  $\rho = \{-0.0596, -0.0486, -0.0427, 0.0177, 0.0553\}$ ，有  $\rho(k) < 10\%$  成立，通过级比偏差检验。

3. 事后检验。由 4 维子序列预测  $X^{(0)}(5) = 2050.6$   $X^{(0)}(6) = 2367.3$   $\epsilon_{min}(5) = 0.0615$   $\epsilon_{min}(6) = 0.0730$   $\epsilon_{min}(avg) = 6.72\%$ ， $p_{min}^0 = 93.28\% > 90\%$ ，预测精度好，预测模型具有 93.28% 可信程度。

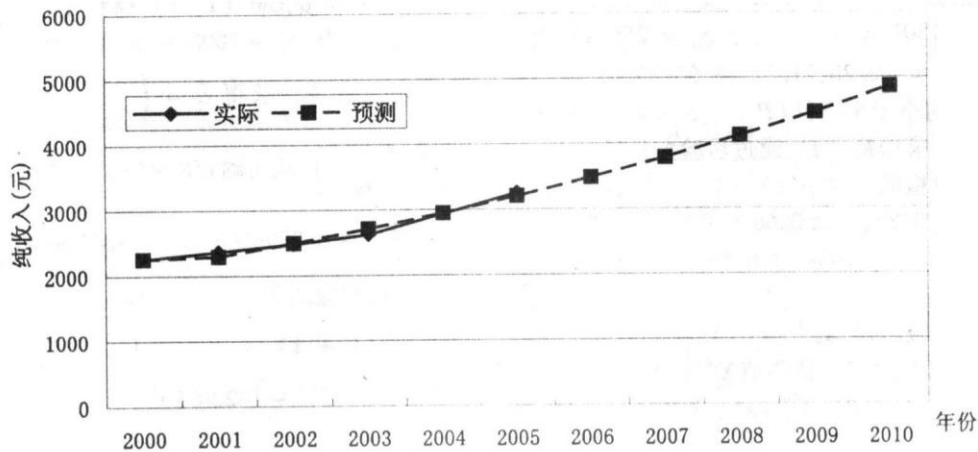
上述检验表明两个 GM (1, 1) 模型具有很好的精度，可以直接根据式 (3)、式 (4) 对“十一五”期间我国农村居民人均纯收入及消费分别进行预测。为了更加充分体现新信息的作用，在上述两个 GM (1, 1) 模型基础上，可分别建立了收入和消费的六维新陈代谢 GM (1, 1) 模型群，并利用该模型群，对农民的收入和消费进行预测。两种预测结果见表四。

表四 2006 年 - 2010 年中国农村居民家庭平均每人纯收入和生活消费支出预测

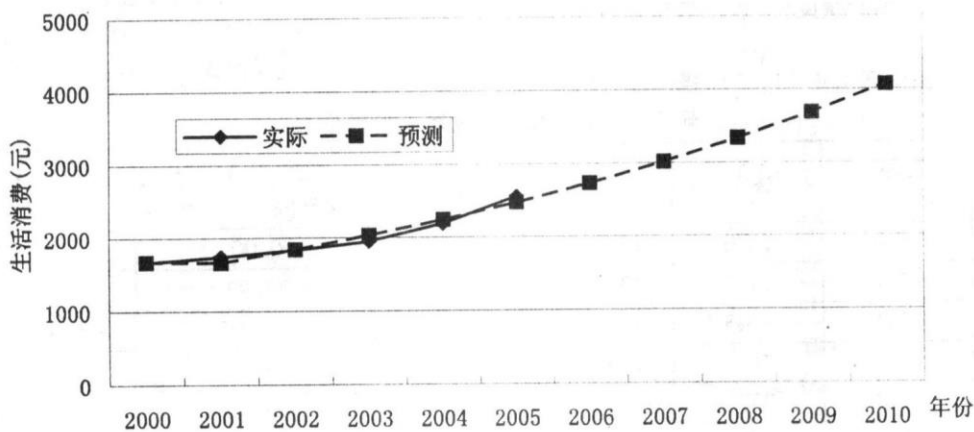
单位: 元

年份	纯收入		生活消费支出	
	基础 GM (1, 1)	新陈代谢 GM (1, 1)	基础 GM (1, 1)	新陈代谢 GM (1, 1)
2006	3480.9	3480.9	2732.9	2732.9
2007	3784.1	3831.9	3018.7	3069.9

2008	4113.8	4207.0	3334.4	3444.6
2009	4472.2	4577.3	3683.0	3826.5
2010	4861.8	4987.6	4068.1	4232.7



图一 2000年-2010年农民纯收入基础GM(1,1)预测值和实际值



图二 2000年-2010年农民生活消费基础GM(1,1)预测值和实际值

## 五、结束语

从上述所建模型的事前、事中、事后检验过程中,我们看到模型具有非常好的精度,从图一、图二也可以看到 2000年 - 2005年模型计算值与实际值的拟合很好,因而通过模型得到的预测结果具有很高的可信程度。灰色预测理论利用有限的信息获取高精度的预测模型,为社会经济研究领域不可缺少的理论。

预测结果表明“十一五”期间我国农民的收入、消费呈现平稳攀升的态势。农民收入的提高,将带来消费市场的日益繁荣,届时我国农民的整体收入水平和消费水平将跃上一个新的台阶。尽管农民收入在不断提高,但是城乡收入差别问题仍然凸现。2005年农民人均纯收入 3255 元,仍然不足城镇居民收入的 1/3 即使 2010 年到达预测值 4900 元左右,也只是相当于 1996 年城镇居民人均可支配收入水平,两者相差近十五年。“十一五”期间除了免收农业税、取消特产税等政策之外,中央还需要推出更重要举措,

寻找新的收入增长点,如:发展优质高效农业、鼓励农民发展养殖、园艺等劳动密集型产品的生产、引导富余劳动力向非农产业和城镇转移、政府要扩大以工代赈规模等。农民的收支两旺,才能使农村真正意义上实现“社会主义新农村”。

## 参考文献:

- [1] 邓聚龙. 灰理论基础 [M]. 武汉: 华中科技大学出版社, 2002. 221 - 225
- [2] 徐国祥. 统计预测和决策 [M]. 上海财经大学出版社, 2005. 208 - 209
- [3] 邓聚龙. 灰预测与灰决策 [M]. 武汉: 华中科技大学出版社, 2002. 99 - 105
- [4] 刘思峰, 郭天榜, 党耀国. 灰色系统理论及应用 [M]. 北京: 科学出版社, 1999. 117 - 118

(编辑校对: 韦群跃 张清述)