

在粮食单产预测中应用马尔柯夫链——以吉林省为例

常春水^{1,2}, 张文秀^{2*}, 宿桂红²

(1. 吉林农业科技学院经济管理学院, 吉林吉林 132101; 2. 四川农业大学经济管理学院, 四川雅安 625014)

摘要 以1981~2006年共26年的统计资料为实例, 用加权的马尔柯夫链模型预测和分析了吉林省未来粮食生产的丰欠年景。

关键词 粮食单产; 加权马尔柯夫链; 预测

中图分类号 S11+9 文献标识码 A 文章编号 0517-6611(2008)27-11621-02

Application of Markov Chain in the Prediction of Unit Grain Yield

CHANG Chunshui et al (College of Economy Management, Jilin Agricultural Science and Technology College, Jilin, Jilin 132101)

Abstract Taking the statistical data from 1981 to 2006 as examples, the harvest states of grain production in Jilin Province in the future were predicted and analyzed by using the weighted Markov chain model.

Key words Unit area grain yield; Weighted Markov chain; Prediction

马尔柯夫链模型作为一种比较成熟的预测方法, 已经得到了广泛的应用, 学者们运用这一原理在很多领域开展了大量的研究。农业领域主要是对耕地需求^[1]、农作物年景^[2-3]、粮食单产风险预警^[4-5]、降水丰枯^[6]等方面进行预测。

吉林省是粮食生产大省, 其粮食生产对全国的贡献率很大。而粮食生产受自然因素影响大、生产周期长、不确定因素多, 所以具有很强的不稳定性。笔者以吉林省1981~2006年共26年的粮食单产数据进行序列规范化处理, 然后以各阶自相关系数为权, 利用加权的马尔柯夫链预测未来吉林省的粮食生产年景, 为确定农业风险管理策略提供信息工具, 同时也为进一步加强我国粮食安全提供参考建议。

1 加权马尔柯夫链预测模型^[2-4]

马尔柯夫过程是具有“无后效性”的随机过程, 即在已知某一随机过程“现在”的条件下, 其“将来”与“过去”是独立的。其数学表达如下:

定义在概率空间 (Ω, F, P) 上的随机序列 $\{X(t), t \in T\}$, 其中 $T = \{0, 1, 2, \dots\}$, 状态空间 $I = \{0, 1, 2, \dots\}$ 称为马尔柯夫链, 如果对任意正整数 l, m, k , 及任意非负整数 $j_1 > \dots > j_2 > j_1$ ($m > j_1$) $P\{X(m+k) = i_{m+k} | X(m) = i_m, X(j_1) = i_{j_1}, \dots, X(j_2) = i_{j_2}, X(j_1) = i_{j_1}\} = P\{X(m+k) = i_{m+k} | X(m) = i_m\}$ 有

$$P\{X(m+k) = i_{m+k} | X(m) = i_m, X(j_1) = i_{j_1}, \dots, X(j_2) = i_{j_2}, X(j_1) = i_{j_1}\} = P\{X(m+k) = i_{m+k} | X(m) = i_m\} \quad (1)$$

成立。这里假定:

$$P\{X(m) = i_m, X(j_1) = i_{j_1}, \dots, X(j_2) = i_{j_2}, X(j_1) = i_{j_1}\} > 0$$

马尔柯夫链的性质和特征定义很多, 这里只考虑其“无后效性”。实际应用中, 一般只考虑齐次马尔柯夫链, 即对任意 $m, k \in T$, 有:

$$p_{ij}(m, k) = p_{ij}(k), i, j \in E \quad (2)$$

式中, $p_{ij}(m, k)$ 表示“系统时刻 m 时处于状态为 i , 经 k 步转移至状态 j 的概率”, $p_{ij}(k)$ 表示“从状态 i 经 k 步转移至状态 j 的概率”。此时转移概率与初始时刻无关。

2 实证分析

以吉林省26年间的粮食单产量数据进行序列规范化处

作者简介 常春水(1975-), 男, 吉林吉林人, 在读硕士, 讲师, 从事农业经济与区域经济方面的研究。* 通讯作者, 博士生导师, 教授, E-mail: cndzwx@163.com。

收稿日期 2008-07-28

理, 利用加权马尔柯夫链预测未来的粮食生产年景(表1)。

表1 1981~2006年粮食单产序列及其状态

Table 1 The unit grain yield sequence and its status during 1981-2006

kg/hm ²					
年份	产量	年景	年份	产量	年景
Year	Yield	Harvest patterns	Year	Yield	Harvest patterns
1981	2 625	5	1994	5 655	2
1982	2 820	5	1995	5 565	2
1983	4 125	4	1996	6 420	1
1984	4 665	3	1997	5 040	3
1985	3 735	4	1998	7 020	1
1986	4 035	4	1999	6 555	1
1987	4 815	3	2000	4 875	3
1988	4 950	3	2001	5 820	1
1989	3 945	4	2002	5 490	2
1990	5 805	1	2003	5 625	2
1991	5 355	2	2004	5 820	1
1992	5 205	2	2005	6 015	1
1993	5 385	2	2006	6 285	1

注: 数据来源于《吉林统计年鉴》(2007)及整理。

Note: The data are from Jilin Statistical Yearbook(2007).

经计算, 得各阶自相关系数及各种滞时的马尔柯夫链权重, 见表2。

表2 各阶自相关系数及各种滞时的马尔柯夫链权重

Table 2 All-order autocorrelation coefficient and Markov chain weight of each time delay

步长	自相关系数 r_k	权重 w_k
Step length	Autocorrelation coefficient	Weight
1	0.543	0.270
2	0.393	0.196
3	0.448	0.220
4	0.371	0.185
5	0.250	0.125

结果得出, 该序列的均值为 $\bar{x} = 5 145 \text{ kg/hm}^2$, 无偏估计的标准差:

$$s = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 72$$

根据表2, 将该序列划分5个级别(表3)。按表3的分级标准, 确定年景状态, 如表1中“年景”栏中所列。

表3 粮食生产年景状态分级

Table 3 Classification of grain production harvest state

年景状态 Harvest state	级别 Grade	分级标准 Classification standards	年均单产量区间 kg/ hm ² Interval of annual average unit yield
1	大丰年	X > 柳 + 0.5s	X > 5 685
2	丰年	柳 X < 柳 + 0.5s	5 145 X < 5 685
3	均年	柳 0.5s X < 柳	4 605 X < 5 145
4	灾年	柳 1.5s X < 柳 - 0.5s	3 435 X < 4 605
5	重灾年	X < 柳 - 2.0s	X < 3 435

通过马尔柯夫链模型的计算,可得各种滞时的状态转移概率矩阵P。

由此预测2006 年粮食年景,如表4 所示。

由表4 可知,Max{ P_{i,j} } = 0.54, 所对应的i = 1, 即2006 年的年景状态为1(丰年), 与实际相符。同理,可以预测2007 年的年景状态为1(丰年)。

以相依性最强的步长为1 的马尔柯夫链的特征分析为例。该链的状态是遍历的,存在唯一的平稳分布,且此时的平稳分布即为它的极限分布。

$$P^{(1)} = \begin{pmatrix} \frac{3}{7} & \frac{2}{7} & \frac{2}{7} & 0 & 0 \\ \frac{2}{7} & \frac{5}{7} & 0 & 0 & 0 \\ \frac{2}{5} & 0 & \frac{1}{5} & \frac{2}{5} & 0 \\ \frac{1}{4} & 0 & \frac{2}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, P^{(2)} = \begin{pmatrix} \frac{3}{6} & \frac{2}{6} & \frac{1}{6} & 0 & 0 \\ \frac{3}{7} & \frac{2}{7} & \frac{1}{7} & 0 & 0 \\ \frac{2}{5} & \frac{1}{5} & 0 & \frac{2}{5} & 0 \\ 0 & \frac{1}{4} & \frac{2}{4} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

表4 2006 年粮食年单产预测

Table 4 The forecast of annual unit yield of grain in 2006

初始年 Beginning year	初始状态 Beginning status	滞时 Time delay	权重 Weight	1	2	3	4	5	概率来源 Probability sources
2005	1	1	0.270	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	0	0	P ⁽¹⁾
2004	1	2	0.196	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	0	0	P ⁽²⁾
2003	2	3	0.220	$\frac{4}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	0	0	P ⁽³⁾
2002	2	4	0.185	$\frac{4}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	0	0	P ⁽⁴⁾
2001	1	5	0.125	$\frac{3}{5}$	$\frac{2}{5}$	0	0	0	P ⁽⁵⁾
P _i (权重和)				0.54	0.29	0.17	0	0	

表5 平稳分布、极限分布及各状态重现期情况

Table 5 Stationary distribution, limit distribution and the recurrence period of each status

状态(j) Status	j	T _j (T _j = 1/ j) 年
1	0.354	2.8
2	0.354	2.8
3	0.190	7.7
4	0.100	10.0
5	0	0

$$P^{(3)} = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} & 0 & 0 & 0 \\ \frac{4}{7} & \frac{2}{7} & \frac{1}{7} & 0 & 0 \\ \frac{1}{5} & \frac{2}{5} & \frac{2}{5} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{2}{4} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, P^{(4)} = \begin{pmatrix} \frac{1}{5} & \frac{3}{5} & \frac{1}{5} & 0 & 0 \\ \frac{4}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P^{(5)} = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} & 0 & 0 & 0 \\ \frac{3}{5} & 0 & \frac{2}{5} & 0 & 0 \\ \frac{1}{5} & \frac{3}{5} & 0 & \frac{1}{5} & 0 \\ \frac{1}{4} & \frac{2}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

$$j = \sum_{i=1}^5 j P_{ij} \quad j = 1, 2, 3, 4, 5$$

依照步长为1 的转移矩阵P⁽¹⁾, 可得到各种状态的重现期T_j = 1/ j, 如表5 所示。

依据表5 的信息推断,在长期过程中,吉林省粮食生产出现大丰年与丰年的机会最多,大约平均每隔2.8 年出现1 次,概率为0.354;其次是均年,大约平均每隔7.7 年出现1 次,概率为0.190;灾年出现的概率较小,为0.100,即每10 年出现1 次;大灾年出现的可能性为0。

参考文献

[1] 刘耀林,刘艳芳,张玉梅.基于灰色-马尔柯夫链预测模型的耕地需求量预测研究[J].武汉大学学报:信息科学版,2004(7):575-579,596.
 [2] 夏乐天,彭志行,沈永梅.加权马尔柯夫链在农作物年景预测中的应用[J].数学的实践与认识,2005(12):30-35.
 [3] 赵利民,周西利.马尔可夫链在大白菜年景预报中的应用[J].西北农业学报,2003(4):139-142.
 [4] 聂荣.基于加权马尔柯夫链的粮食单产风险预测——以辽宁省为例[J].农业技术经济,2007(5):88-92.
 [5] 徐学荣,林少伟.福建粮食单产风险预警的马尔柯夫方法[J].福建农林大学学报:哲学社会科学版,2005(1):38-40.
 [6] 孙才志,张戈,林学钰.加权马尔柯夫模型在降水丰枯状况预测中的应用[J].系统工程理论与实践,2003(4):100-105.