ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 2 (2012) No. 3, pp. 40-49

Markov chain analysis and prediction on the fluctuation cycle of vegetable price

Xiaoxia Zhu[1], Xiuquan Xu[2]

[1] School of Sciences, Hebei University of Science and Technology, Shijiazhuang Hebei 050018, China

[2] Schools of Economics and Management, Beijing Jiaotong University Beijing100044, China

Abstract: In this paper, in order to solve the problems of heavy price fluctuations and unbalanced resource allocation in the vegetable wholesale market caused by huge daily demand and varied supply channels, we establish the predication model of Markov chain (MC), and a case just like the future price fluctuation interval and cycle of some kind of seasonal vegetable in Fushan area, the analysis and forecasting results indicate that the predication model of MC is a feasible technique to analyze the changing laws of something and forecasts its future trend with the basic principles and methods of MC. Meanwhile, the principle of economics that the commodity price fluctuates around the commodity value in the free market economy is confirmed from the mathematical point.

Keywords: Price of Vegetable, Markov Chain, Fluctuation Cycle, Prediction

1. Exordium and propaedeutics

Daily necessities have large influence on our lives, Related to the happiness index of national life, The Chinese government attaches great importance to the commodity price. Particularly grain and vegetables, wheat, soybeans have been incorporated into the national macro-control, some are being incorporated into the national macroeconomic control, however, due to Preservation of seasonal vegetables, Government should not macro-control, This resulted in a seasonal vegetable prices volatility, Such as oversupply, will inevitably result in a large number of vegetables slow-moving, Farmers damaged, Such as The large number of Henan farmers deliver their hard-planting vegetables to the public free of charge in 2011; Such as In short supply, will inevitably result in high prices of vegetables, and Affect the grassroots life. For the above reasons, in this paper, this paper established a mathematical model by using Markova chain, Prediction the price fluctuation and price fluctuation cycle of some area of some vegetables in the future period of time, to provide reference for the local government departments.

Markov chain is a kind of random process, it has" no aftereffect"(Markov), That is, to determine the future status of the process, Know its state at the moment is enough, Does not need to know its previous state, Markov chain definition is given below:

Definition 1.1(Markov chain) Random process $\{X_n, n = 0,1,2,\cdots\}$ called a Markov chain, if it takes a finite or countable value s_0 , s_1 , s_2 , $\cdots$,(using $\{0,1,2,\cdots\}$ to mark s_0 , s_1 , s_2 , $\cdots$,And call it the status of the process, $\{0,1,2,\cdots\}$ or its subset is credited as S , Referred to as the process state space), For any $n \geq 0$ And the state $i, j, i_0, i_1, i_2, \cdots, i_{n-1}$,

$$P\left\{\begin{matrix} X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, \cdots, \\ X_{n-1} = i_{n-1}, X_n = i \end{matrix}\right\} = P\{X_{n+1} = j \mid X_n = i\} \qquad (1)$$

(1) is called Markov chain one-step transfer probability; it describes the characteristics of Markov chain, which is The Markov property. If said chain is time homogeneous, One-step transfer probability only with the state i, j , has nothing to do with the time n , Denoted $p_{ij}\,(i, j \in S)$ Markov chain one-step transfer probability, denote

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} & \cdots \\ p_{10} & p_{11} & p_{12} & p_{13} & \cdots \\ p_{20} & p_{21} & p_{22} & p_{23} & \cdots \\ p_{30} & p_{31} & p_{32} & p_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \qquad (2)$$

Called P as Transition probability matrix, Referred to as the transfer matrix, Easy to see nature of $p_{ij}\,(i, j \in S)$:

$$p_{ij} \geq 0,\ i, j \in S\ , \sum_{j \in S} p_{ij} = 1\ \ \forall i \in S \tag{3}$$

Definition 1.2(n step Transition probability**) call** Conditional probability

$$p_{ij}^{(n)} = P\{X_{m+n} = j \mid X_m = i\},\ i, j \in S\ \ m \geq 0\ n \geq 0 \tag{4}$$

as Markova chain n step transition probability, correspondingly called $P^{(n)} = \left(p_{ij}^{(n)}\right)$ as the n step transition matrix by Chapman-Kolmogorov equation (C-K equation) knowing $P^{(n)} = P^n$.

Definition 1.3 called state i reachable state j （ $i, j \in S$ ）,if existing $n \geq 0$ to make $p_{ij}^{(n)} \geq 0$, Credited as $i \to j$, If at the same time $j \to i$, Said i and j interworking, Credited as $i \leftrightarrow j$.

According to the interworking nature between the state, the two interworking state can be divided into a class, and any one state cannot belong to two different categories.

Definition 1.4 if there is only one class of Markova chain, Says it is irreducible; otherwise known as the covenant.

Definition 1.5 (Periodic) if the set $\{n : n \geq 1, p_{ii}^{(n)} > 0\}$ is not empty, then call its greatest common divisor $d = d(i)$ as the cycle of the state i .if $d > 1$, Call state i is periodic; if $d = 1$, Call state i is not periodic. If the above collection is empty, called cycle of state i is infinite.

Theorem1.1 if the state i and j belong to the same class, then $d(i) = d(j)$.

Definition 1.6 (Recurrent) for state $i, j \in S$, starting from the state i, after n step, first reached the state j probability denoted as $f_{ij}^{(n)}$, then

$$f_{ij}^{(n)} = 0,\ f_{ij}^{(n)} = P\{X_n = j, X_k \neq j, k = 1, 2, \cdots, n-1 \mid X_0 = i\}\ \ n \geq 1 \tag{5}$$

let $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$, if $f_{ii} = 1$, call state i as recurrent State; if $f_{ii} \leq 1$, call state i as non-recurrent State or instantaneous State.

In fact, f_{ij} denotes the probability starting from state i after finite steps can reach the state j .when state i is recurrent State, Means the probability from the state i ,after a finite number of steps to return to the state i is 1;whereas if state

i is non-recurrent State, it means Not to return to the state i with probability $1-f_{ii}$,that is slide by state i .

For recurrent State i , define:

$$\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)} \tag{6}$$

That is μ_i said from the state i starting return to state i a desired average step length (or time).

Definition 1.7 for recurrent State i , if $\mu_i < +\infty$, then call state i as Normal return state; if $\mu_i = +\infty$, then call state i as zero recurrence state. In particular, if State i is normal return state and non periodic, then call state i as ergodic state. If State i is ergodic state, and $f_{ii}^{(1)} = 1$, then call state i as absorbing state.

Theorem1.2 the same category of state i, j in Markova chain, they are of the same as the recurrent state or non-recurrent state, and when they are recurrent state, they are of the same as normal return state or zero recurrent state.

Theorem1.3 irreducible Markov chain with Cycle d , whose state space S can be uniquely decomposed into d mutually disjoint subsets,

$$S = \bigcup_{r=0}^{d-1} s_r \quad s_r \cap s_t = \varnothing, r \neq t \tag{7}$$

And starting from any state in s_r , by 1 step will enter s_{r+1} (where $s_d = s_0$).

Theorem1.4 let Irreducible、Normal return state、Cycle d Markova chain, its state space is S, then for state $i \to j$, $i, j \in S$,

$$\lim_{n\to\infty} p_{ij}^{(nd)} = \begin{cases} \dfrac{d}{\mu_j} & \textit{if i and j are all belong to subset } s_r, \\ 0 & \text{other,} \end{cases} \tag{8}$$

Where $S = \bigcup_{r=0}^{d-1} s_r$ is defined in theorem1.3.In particular, when $d = 1$, then $\forall\, i, j \in S$, $\lim_{n\to\infty} p_{ij}^{(nd)} = \dfrac{1}{\mu_j}$.

Definition 1.8 For Markova chain, Probability distribution $\{\pi_j\ ,\ j \in S\}$ is called Invariant, if $\pi_j = \sum_{i \in S} \pi_i p_{ij}$.

Definition 1.9 if all states of Markova chain are interlinked and are all positive recurrent state with Cycle of 1,then call it ergodic. For ergodic Markova chain

$$\lim_{n \to \infty} p_{ij}^{(n)} = \upsilon_j\ ,\ j \in S \tag{9}$$

Theorem1.5 for Irreducible aperiodic Markov chain:

(1)If it is ergodic, then $\pi_j = \lim_{n \to \infty} p_{ij}^{(n)} > 0$ （ $j \in S$ ） Is invariant distribution and is the only invariant distribution;

(2) If the state is instantaneous or zero recurrence, the invariant distribution does not exist.

By theorem 1.4 and theorem1.5 showing that Ergodic Markovian chain, average step length μ_j that starting from state j then return to state j and Invariant distribution π_j , satisfy the following relations

$$\mu_j = 1/\pi_j\ . \tag{10}$$

2. Model construction

The following hypotheses are made before constructing the model:

(1) The seasonal vegetable is a kind of commodity in a certain area, with large demand and a variety of supply channels, which will not cause monopoly of some merchant;

(2) The growth cycle of the seasonal vegetable is appropriate.

Condition (1) ensures that the price of the seasonal vegetable is affected by supply and demand. If the current price of the seasonal vegetable is comparatively high, more vegetable will be transported into the local area, or there will be more growth of the seasonal vegetable; if the current price of the seasonal vegetable is comparatively low, vegetable grower will reduce the production of the seasonal vegetable, therefore the price of the seasonal vegetable will be affected by the current price, while it may have nothing to do with the past price of the seasonal vegetable.

Condition (2) ensures that it is meaningful to construct the model. If the growth period of the season vegetable is short, the model may be of little value for the local government.

In order to facilitate research, the price can be divided into several intervals according to specific situation. Therefore let

$$0 < t_1 < t_2 < \cdots < t_{m-1} < +\infty , \quad m \in N^+ \tag{11}$$

In particular, let $t_0 = 0, t_m = +\infty$, we get the price interval.

$$s_r = [t_r , t_{r+1}) \qquad 0 \le r < m \quad r \in N \tag{12}$$

Make all the price intervals constitute the sample space.

$$\Omega = \{ s_r \mid 0 \le r < m \quad r \in N \} . \tag{13}$$

To take a certain day in the past as the first day, then let the price of the seasonal vegetable on that day X_1 , and then X_n means the price of the seasonal vegetable in day n . Since X_n will be concluded in a certain interval, to judge whether the events happen or not depends on whether X_n belongs to s_r or not. Let Γ be the algebra σ of the event. The appropriate probability $\hat{P}$ can be selected to construct probability space $(\Omega, \Gamma, \hat{P})$. Then X_n is a random variable in the probability space $(\Omega, \Gamma, \hat{P})$. So we can obtain a group of random process $\{X_n , n \in N^+\}$ in the probability space $(\Omega, \Gamma, \hat{P})$. Moreover hypothesis (1) shows that the current price of the seasonal vegetable has something to do with yesterday's, while it has nothing to do with the price in the earlier days. So the random process meets Markov property, which means that $\{X_n , n \in N^+\}$ is a group of Markov chain, and it is homogeneous. In order to mark easily, let $S = \{0, 1, \cdots, m-1\}$ indicate the state space $\{ s_0, s_1, \cdots, s_{m-1} \}$ of the Markov chain, then make

$$p_{ij} = P\{X_{n+1} = j \mid X_n = i\} \tag{14}$$

Be the transition probability from state i to state j for the Markov chain at the time of n , then the transfer matrix is as follows.

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0(m-1)} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1(m-1)} \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2(m-1)} \\ p_{30} & p_{31} & p_{32} & \ddots & \vdots \\ p_{(m-1)0} & p_{(m-1)1} & p_{(m-1)2} & \cdots & p_{(m-1)(m-1)} \end{bmatrix}. \tag{15}$$

If we know that the price of the seasonal vegetable in day n belongs to the state i (i.e., the price of the seasonal vegetable in the day belongs to the interval i), then the probability of expectation of the price of the seasonal vegetable transferring from the state i after the first day will be as follows:

$$\mu_i^{(1)} = \sum_{j=0}^{m-1} j \times p_{ij} \tag{16}$$

The state expectations of the seasonal vegetable price after day k will be as follows:

$$\mu_i^{(k)} = \sum_{j=0}^{m-1} j \times p_{ij}^{(k)} \tag{17}$$

Among which $P^{(k)} = \left(p_{ij}^{(k)}\right) = P^k$, so we can estimate the price in day k according to the state expectation. If $\mu_i^{(k)}$ is an integer r , then the estimated price in day k will be in the interval $(t_r , t_{r+1}]$; if $\mu_i^{(k)}$ is not an integer, we provide the following two methods for reference:

(1) Changing $\mu_i^{(k)}$ to be an integer by rounding up;

(2) Translation for the state interval, such as $\mu_i^{(k)}$ = 2.7, then we can get the price valuation interval $(2.7 \times t_1 , 2.7 \times t_2]$.

The first method is adopted in the case study of this paper. If the Markov chain is ergodic, theorem 1.4 and theorem 1.5 shows that, for ergodic Markov chain, the average step length μ_j with the beginning state j and returning state j and invariant distributions satisfy the following relations:

$$\mu_j = 1/\pi_j \tag{18}$$

μ_j is the average time of the price beginning with the state j and ending with returning to the state j. While invariant distribution can be solved by the following linear equation.

$$\pi_0 = \sum_{i=0}^{m-1} \pi_i p_{ij} , \pi_1 = \sum_{i=0}^{m-1} \pi_i p_{ij} , \ldots, \pi_{m-1} = \sum_{i=0}^{m-1} \pi_i p_{ij} . \qquad (19)$$

3. Application analysis

We check the model based on Fo Shan vegetable wholesale market as an example. According to the lettuce price of vegetable wholesale market in Fo Shan city in 2011, we define four status as $s_1 = [0, 1)$, $s_2 = [1, 2)$, $s_3 = [2, 3)$, $s_4 = [3, 4)$.

As the price doesn't change much in a short time, we define 5 days as a period, and the price is the average of the 5 days. The transition probability p_{ij} can be replaced by N_{ij} / N_i, where N_i means number of period when lettuce wholesale price of Foshan on 2011 is among the interval i, and N_{ij} means number of period transfer from interval *i* to *j*. Then we got the transition probability matrix based on lettuce wholesale price of Foshan on 2011 as below:

$$P = \begin{bmatrix} 0.7000 & 0.2000 & 0.1000 & 0 \\ 0.1333 & 0.8000 & 0 & 0.0667 \\ 0.0455 & 0.0909 & 0.7273 & 0.1364 \\ 0 & 0 & 0.2667 & 0.7333 \end{bmatrix}$$

The 5-day-average price of lettuce of Foshan city from 25th to 29th December 2011 is 1.82 Yuan/Kg, define this time as X_1, thus the current state is 2, calculating with Matlab7.8 we got the result $\mu_2^{(1)} = 2.0001$, $\mu_2^{(2)} = 2.0357$, $\mu_2^{(3)} = 2.0871$, $\mu_2^{(4)} = 2.1431$, $\mu_2^{(5)} = 2.1978$, $\mu_2^{(6)} = 2.2482$, $\mu_2^{(7)} = 2.2933$. The result is shown Table1.

Then repeat the calculation and take 10 days as a period, we got a similar result. According to the result above we can confirm that it is feasible to predict the price of vegetable in some certain region using Markov chain model, and the key element of the model is proper interval of time and price.

Now we analyze period of waves of price. According to the probability matrix we can define the relationship between each Markov chain as $1 \leftrightarrow 2$, $1 \leftrightarrow 3$, $3 \leftrightarrow 4$ which is irreducible. Here we find the period of waves 1,2,3,4 is 1 which equals to the period of waves of the Markov chain. We get equations with the same distribution according to Theorem 1.5 as below:

$$\pi_1 = 0.7\pi_1 + 0.1333\pi_2 + 0.0455\pi_3, \pi_2 = 0.2\pi_1 + 0.8\pi_2 + 0.0909\pi_3,$$

$$\pi_3 = 0.1\pi_1 + 0.7273\pi_3 + 0.2667\pi_4, \pi_4 = 0.0667\pi_2 + 0.1364\pi_3 + 0.7333\pi_4, 1 = \pi_1 + \pi_2 + \pi_3 + \pi_4$$

Calculate it and we get: $(\pi_1, \pi_2, \pi_3, \pi_4) = (0.1811, 0.3106, 0.2849, 0.2234)$

Also we can get it from formula $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$ by taking n as a large value such as 500, and the distribution is:

$(0.1817, 0.3115, 0.2856, 0.2239)$.

Compare the results we find they are similar, and we get the period of waves of the four statuses from $\mu_j = 1/\pi_j$ as below:

$(5.5, 3.2, 3.5, 4.5)$

Table1. Comparison of the actual price and the prediction result

Number	Date	Actual price	prediction result
1	2011-12-30	1.6	（1，2）
2	2011-12-31	1.6	（1，2）
3	2012-01-05	1.7	（1，2）
4	2012-01-06	1.9	（1，2）
5	2012-01-07	1.5	（1，2）
6	2012-01-09	1.5	（1，2）
7	2012-01-10	1.5	（1，2）
8	2012-01-11	1.5	（1，2）
9	2012-01-12	1.5	（1，2）
10	2012-01-13	1.5	（1，2）
11	2012-01-14	1.4	（1，2）
12	2012-01-16	1.5	（1，2）
13	2012-01-17	1.5	（1，2）
14	2012-01-18	1.5	（1，2）
15	2012-01-19	1.5	（1，2）
16	2012-01-20	1.3	（1，2）
17	2012-01-21	1.3	（1，2）
18	2012-01-30	1.8	（1，2）
19	2012-01-31	1.9	（1，2）
20	2012-02-01	1.9	（1，2）

4. Conclusions

It is feasible to predict the price of vegetable in some certain region using Markov chain model, the result is similar to the real situation. The key element is the definition of the status space, the smaller the price interval is, the more correct and valuable the result is. Also the definition of time period is very important. In this paper we take 5 days and 10 days as a period representatively, and the results are similar. How to choose it correctly? From the result we found that the prediction of the period is quite different from the real situation, this is caused by the definition of status space. In this paper, we consider discrete Markov chain model, moreover, it worth discussing whether continuous Markov chain model is feasible in the future.

Acknowledgements

This paper is supported by the project from National Key Technologies R & D Program of China (2012BAH20B03) and Supported by Foundation of Hebei University of Science and Technology (XL200761).

References

Batabval A. A. (2006). Markov chains Models, algorithms and applications. *Interfaces*, 36(6), 609-610.

Hu, T. B. & Ye, J. K. (2008). The Application of Marco Chain Model in GIS Data Prediction. *Applications of Computer Systems*, 8, 90-93.

Lu, D, J. (2006). *Random Process and its Applications*. Beijing, Tsinghua University press.

Run, Y. Q. (2007). An Improved Prediction Model of Markov Chain. *Science Technology and Engineering*, 7, 5143-5155.

Yan, R. F. (1999). The Random Process Model on Stock Market Prediction. *Journal of Northwest Normal University (Natural Science Edition)*, 3, 44-46.