

# 基于时间序列模型的蔬菜类商品的补货和定价策略分析

田 静, 吕 平\*

杭州师范大学, 数学学院, 浙江 杭州

收稿日期: 2024年5月6日; 录用日期: 2024年5月29日; 发布日期: 2024年6月6日

## 摘 要

在生鲜商超中, 蔬菜类商品一般都有着较短的保鲜期, 它们的品相随销售时间的增加而变差。因此, 对于商超而言, 为了使蔬菜能够在当天售出, 制定恰当的补货和定价策略显得尤为重要。基于某商超经销的6个蔬菜品类在某段时间的相关销售数据, 本文首先根据价格需求弹性的计算公式, 对各品类取对数后的日销售量和日销售价格进行线性回归, 得到各品类销售总量与成本加成定价的关系表达式。之后对各品类的日销售量进行时间序列分析, 建立乘积ARIMA模型, 预测得到各品类未来一周的日销售量, 并根据损耗率计算公式得到各品类未来一周的日补货总量。最后以商超每日净收益最大为目标函数, 建立数学规划模型, 得到各品类未来一周的定价策略和日收益。

## 关键词

线性回归, 价格需求弹性曲线, 时间序列, 数学规划

# Replenishment and Pricing Strategy Analysis of Vegetable Commodities Based on Time Series Model

Jing Tian, Ping Lyu\*

School of Mathematics, Hangzhou Normal University, Hangzhou Zhejiang

Received: May 6<sup>th</sup>, 2024; accepted: May 29<sup>th</sup>, 2024; published: Jun. 6<sup>th</sup>, 2024

## Abstract

In the fresh supermarket, vegetable products generally have a short shelf life, and their products

\*通讯作者。

deteriorate with the increase of sales time. Therefore, it is particularly important for supermarkets to develop appropriate replenishment and pricing strategies in order to enable vegetables to be sold on the same day. Based on the relevant sales data of 6 vegetable categories distributed by a supermarket in a certain period of time, this paper first conducts linear regression on the daily sales volume and daily selling price of each category after taking logarithm according to the calculation formula of price demand elasticity, and obtains the relationship expression between the total sales volume of each category and the cost plus pricing. Then, the daily sales volume of each category is analyzed in time series, and the product ARIMA model is established to predict the daily sales volume of each category in the next week, and the total daily replenishment volume of each category in the next week is obtained according to the calculation formula of the depletion rate. Finally, taking the maximum daily net return of the supermarket as the objective function, a mathematical planning model is established to get the pricing strategy and daily return of each category in the next week.

## Keywords

Linear Regression, Price Demand Elasticity Curve, Time Series, Mathematical Programming

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



## 1. 引言

随着生鲜市场进入刚需期, 生鲜商超作为新型零售模式已成为我国生鲜销售的主导力量[1], 越来越多的生鲜商超出现在人们的日常生活中。为了保证商品的新鲜程度, 生鲜商超需要在当日尽可能多地将生鲜售出, 特别是蔬菜类商品。蔬菜类商品的补货和定价策略对商超的收益有着重要的影响, 同时, 综合考虑蔬菜的保鲜期、销售时间、品相、损耗率、需求量、供应量等多种因素, 对于商超做出正确决策至关重要。

本文根据 2023 年全国大学生数学建模竞赛 C 题, 基于题中所给的某商超经销的 6 个蔬菜品类(花叶类、花菜类、水生根茎类、茄类、辣椒类和食用菌类)商品的相关数据, 通过建立相关的数学模型, 来分析和探究各个蔬菜品类的销售总量和成本加成定价的关系、日补货总量、净收益最大化的定价策略等。根据本文的分析, 该商超可以制定未来一周各个蔬菜品类的合理科学的补货和定价策略, 并取得最大化的净日收益。

## 2. 基于线性回归拟合的销售总量与成本加成定价的关系

### 2.1. 价格需求弹性

蔬菜品类的销售总量与成本加成定价, 简单来说就是需求量与价格的关系。根据经济学原理中典型的价格需求弹性曲线  $P = f(Q)$  (单调递减), 类似地有“蔬菜品类的价格需求弹性曲线”。在公式(1)中,  $\varepsilon$  表示价格需求弹性系数,  $\Delta Q$  和  $\Delta P$  分别为某一小段时间内需求量变化量和价格变化量,  $Q$  和  $P$  分别表示该时间段内初始需求量和价格。

$$\varepsilon = \frac{\Delta Q/Q}{\Delta P/P} \quad (1)$$

如果  $0 < \varepsilon < 1$ , 那么需求缺乏弹性; 如果  $\varepsilon > 1$ , 那么需求富有弹性; 如果  $\varepsilon = 1$ , 即单位价格需求弹

性, 那么说明需求量变动幅度和价格变动幅度相同。

## 2.2. 基于价格需求弹性公式建立与求解线性回归模型

### 2.2.1. 模型的建立

在蔬菜类商品的需求量  $Q$  和价格  $P$  所构成的散点图中, 若作一需求曲线  $Q = f(P)$  可以拟合这些点, 则利用拟合的需求曲线函数的理论值可以近似替代  $Q$  和  $P$  的实际值[2], 同时拟合曲线价格需求弹性的理论值可以近似代替该商品价格需求弹性的实际值。

对于  $Q = f(P)$ , 其价格弹性系数  $\tilde{\varepsilon} = -\Delta Q/Q \div \Delta P/P$  可表示为微分形式  $\tilde{\varepsilon} = -dQ/Q \div dP/P$  [3]。对  $dQ/Q = -\tilde{\varepsilon} \cdot dP/P$  两边取积分, 得  $\int dQ/Q = -\tilde{\varepsilon} \cdot \int dP/P$ , 从而  $\ln Q = -\tilde{\varepsilon} \cdot \ln P + \ln K$ , 即

$$Q = K \cdot P^{-\tilde{\varepsilon}} \quad (2)$$

对于公式(2), 令  $Y = \ln Q$ ,  $X = \ln P$ ,  $c = \ln K$ , 建立线性回归方程  $Y = -\tilde{\varepsilon} \cdot X + c$ , 即为一元线性回归, 其中  $-\tilde{\varepsilon}$  为回归系数,  $c$  为常数项。

### 2.2.2. 模型的求解

Step 1: 将原始数据(销售总量和成本加成定价)取对数, 得到对数化后的处理数据。

Step 2: 基于处理后的数据建立线性回归方程, 得到线性回归方程  $\ln Q = -\tilde{\varepsilon} \cdot \ln P + \ln K$ 。

Step 3: 检验线性回归方程是否通过 p 值检验。

Step 4: 将线性回归方程逆推为原始方程  $Q = K \cdot P^{-\tilde{\varepsilon}}$ 。

### 2.2.3. 模型的结果

如表 1 所示, 这 6 类蔬菜数据都能够通过方程的检验, 因此较为可信。

如图 1 所示, 从左到右, 从上到下依次为花叶类、花菜类、水生根茎类、茄类、辣椒类和食用菌类, 可以得到各个品类销售总量与成本加成定价的拟合曲线。由图 1 和表 2 可知, 对于花叶类、花菜类、茄类、辣椒类和食用菌类, 弹性系数满足  $0 < \tilde{\varepsilon} < 1$ ; 对于水生根茎类, 弹性系数满足  $\tilde{\varepsilon} > 1$ 。因此, 结合生活常识可知大部分蔬菜类商品缺乏需求弹性, 本文的结果也恰好印证了这一事实。

Table 1. Parameter test results of linear regression for each category

表 1. 各个品类线性回归的参数检验结果数据

	花叶类	花菜类	水生根茎类	茄类	辣椒类	食用菌类
p 值	$2.64 \times 10^{-16}$	$3.8 \times 10^{-25}$	$2.72 \times 10^{-78}$	$1.66 \times 10^{-12}$	$1.43 \times 10^{-29}$	$4.51 \times 10^{-23}$
F 值	69.2	113	415	51.1	135	102

Table 2. The fitting equation of total sales volume and cost plus pricing for each category

表 2. 各个品类销售总量与成本加成定价的拟合方程

品类	拟合方程
花叶类	$Q_1 = 5.7750 \cdot P_1^{-0.39376}$
花菜类	$Q_2 = 5.0842 \cdot P_2^{-0.72862}$
水生根茎类	$Q_3 = 6.5480 \cdot P_3^{-1.46850}$
茄类	$Q_4 = 3.9506 \cdot P_4^{-0.50431}$
辣椒类	$Q_5 = 5.2224 \cdot P_5^{-0.45044}$
食用菌类	$Q_6 = 5.3736 \cdot P_6^{-0.62093}$

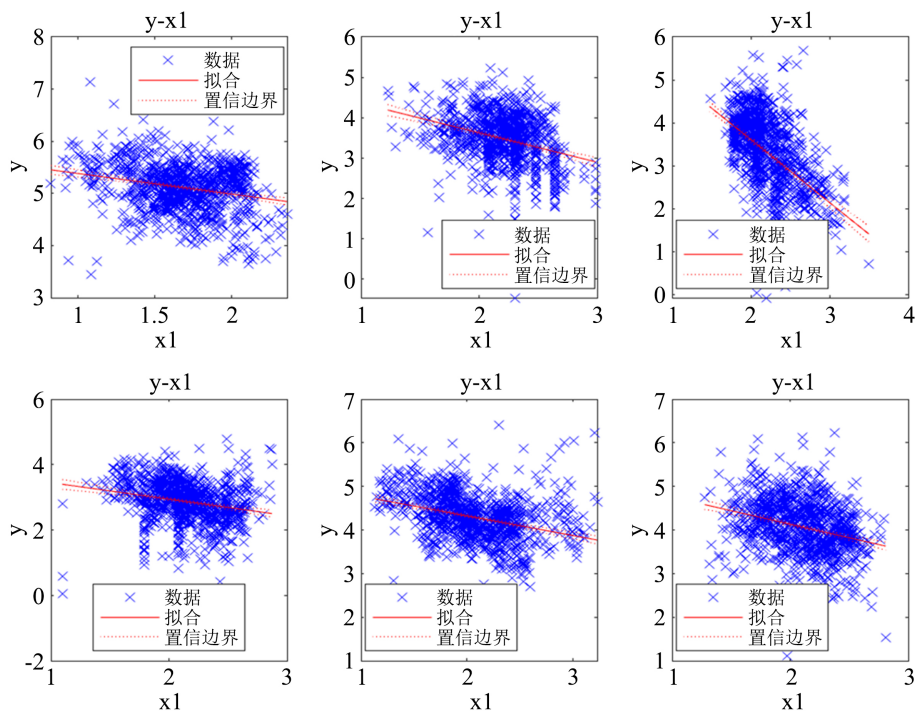


Figure 1. The fitting curve of total sales and cost plus pricing for each category  
图 1. 各个品类销售总量与成本加成定价的拟合曲线

### 3. 基于时间序列预测的销售总量和批发价格

由于需要对 6 个品类的销售总量和批发价格进行预测, 即共有 12 次预测分析。但受篇幅限制, 本文仅呈现其中 1 次的具体预测过程, 其余 11 次预测同理。以下将以花叶类的销售量为例。

#### 3.1. 模型的建立与求解

##### 3.1.1. 数据可视化: 时序图的变化规律

通过 R 软件绘制出的时序图, 如图 2 所示。根据时序图中的折线变化, 可知原始的时间序列是一组单位间隔为天数的数据, 有周期性趋势(周期为 7), 是非平稳的序列。

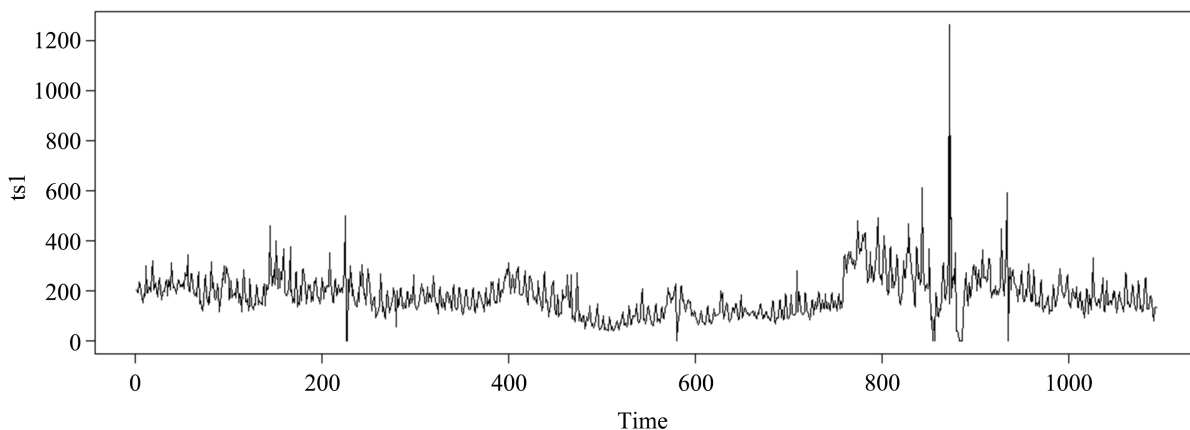
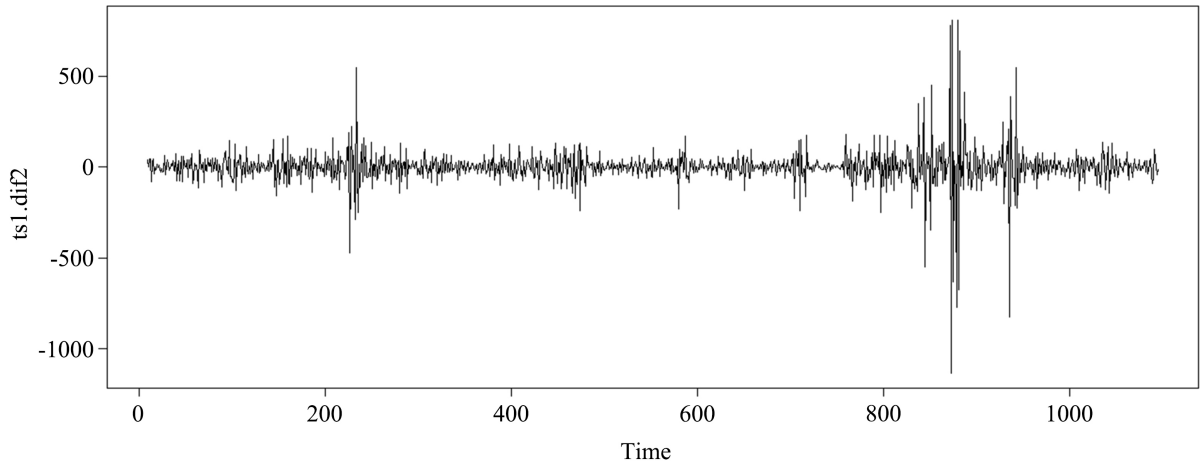


Figure 2. Time series diagram of total sales volume of mosaic  
图 2. 花叶类销售总量的时序图

### 3.1.2. 数据处理与分析

首先, 为了消除原始时间序列的周期性[4], 对原始时间序列进行 7 步差分, 最终可得到  $ts1.tf2$  序列。根据  $ts1.tf2$  序列的时序图(如图 3 所示), 可知差分以后的序列无明显的长期性和周期性。



**Figure 3.** Timing diagram of  $ts1.tf2$  sequence  
**图 3.**  $ts1.tf2$  序列的时序图

**Table 3.** p-value table of unit root test and pure randomness test of  $ts1.tf2$  sequence  
**表 3.**  $ts1.tf2$  序列的单位根检验和纯随机性检验的 p 值表

序列	单位根检验	纯随机性检验	
		Lag = 6	Lag = 12
$ts1.dif2$	p = 0.01	p = 4.23e-106	p = 7.378e-11

对  $ts1.tf2$  序列进行单位根检验(结果如表 3 所示), 得到 p 值为 0.01, 小于 0.05, 故拒绝原假设( $H_0$ : 序列非平稳),  $ts1.tf2$  序列平稳。

对  $ts1.tf2$  序列进行纯随机性检验(结果如表 3 所示), 在延迟 6 阶时, p 值为 4.23e-106, 小于 0.05; 在延迟 12 阶时, p 值为 7.378e-11, 也小于 0.05。故拒绝原假设( $H_0$ : 序列是纯随机序列),  $ts1.tf2$  序列是非纯随机序列。

因此, 可以对  $ts1.tf2$  序列进行时间序列建模。

### 3.1.3. 建立 ARIMA 模型

通过 R 软件绘制  $ts1.tf2$  序列的自相关图和偏自相关图, 如图 4 所示。

长期性定阶: 通过自相关图可知  $\rho(k)$  拖尾, (1,2,6) 在外; 通过偏自相关图可知  $\phi_{kk}$  拖尾, (1,2,4,6) 在外。

周期性定阶: 通过自相关图可知  $\rho(k)$  2 阶截尾; 通过偏自相关图可知  $\phi_{kk}$  4 阶拖尾。

因此, 将模型初步定阶为 “ARIMA((1,2,4,6),1,(1,2,6))\*(0,1,2)<sub>7</sub>”。对该模型进行参数估计, 依次剔除未通过 t 检验的最差的参数, 得到最终模型为

$$(1-B)(1-B^7)X_t = \frac{1-0.5801B^3-0.5466B^5+0.2678B^6}{1-0.7116B-0.3652B^2+0.2170B^3+0.4593B^5+0.1287B^6} \cdot \frac{1-0.9806B^7}{1} \cdot \xi_t$$

其中,  $\xi_t \sim WN(0,3537)$ 。延迟 6 阶时, p 值为 0.9821; 延迟 12 阶时, p 值为 0.7186。

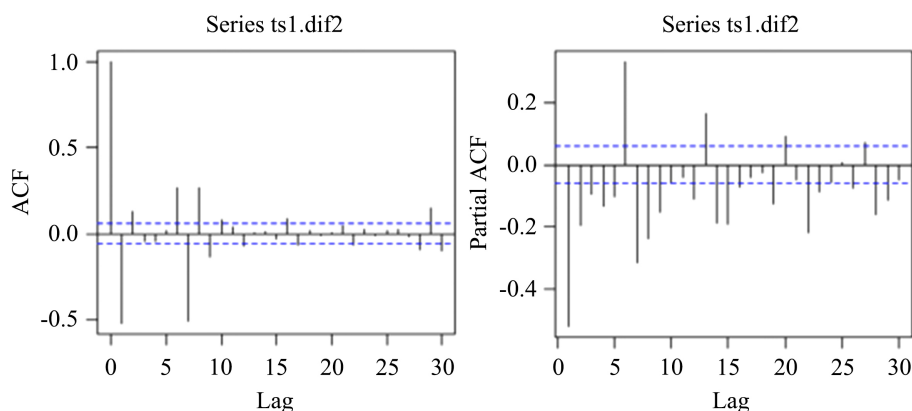


Figure 4. Autocorrelation and partial autocorrelation of *ts1.tf2* sequence

图 4. *ts1.tf2* 序列的自相关图和偏自相关图

### 3.2. 模型的结果

如表 4 所示, 通过时间序列预测花叶类未来 7 天的销售量。同理, 通过以上方法可以得到其他 5 个品类未来 7 天的销售量。由于还需要预测 6 个品类未来 7 天的批发价格, 可以采用同样的方法得出结果。

Table 4. The mosaic time series predicts sales for the next 7 days

表 4. 花叶类时间序列预测未来 7 天的销售量

时间	2023.7.1	2023.7.2	2023.7.3	2023.7.4	2023.7.5	2023.7.6	2023.7.7
花叶类的日销售量	193.06	184.59	137.06	126.24	125.98	124.51	149.03

## 4. 收益最大化的规划模型

### 4.1. 模型的建立

#### 4.1.1. 决策变量

① 预测销量  $d_{ij}$ 、预测批发价  $TP_{ij}$

通过时间序列预测模型, 可以预测得到各个品类第  $j$  天的日销售量和第  $j$  天的批发价格。因此本文引进变量  $d_{ij}$  和  $TP_{ij}$ , 分别表示第  $i$  个品类第  $j$  天的预测销量和第  $i$  个品类第  $j$  天的预测批发价格。

② 损耗率  $L_i$

由于已知的近期损耗率是各个单品的近期损耗率, 为了更好地表示品类的近期损耗率, 可以使用各个品类下单品损耗率的数学期望来表示品类的近期损耗率。因此本文引进变量  $L_i$ , 表示第  $i$  个品类的近期损耗率。

$$\begin{cases} L_i = \sum_k p_{ik} l_{ik} \\ p_{ik} = s_{ik} / \sum_k s_{ik} \end{cases}$$

其中,  $l_{ik}$  表示第  $i$  个品类下第  $k$  个单品的近期损耗率,  $p_{ik}$  表示第  $i$  个品类下第  $k$  个单品近期损耗率为  $l_{ik}$  的概率,  $s_{ik}$  表示第  $i$  个品类下第  $k$  个单品的最近一周的销售总量。

根据损耗率的计算公式“损耗率 = 损耗量 ÷ 净用量”, 可以得到第  $i$  个品类未来第  $j$  ( $j=1, \dots, 7$ ) 天的日补货总量  $rd_{ij}$ , 即

$$rd_{ij} = d_{ij}(1 + L_i)$$

③ 成本加成定价  $P_{ij}$ 、净利润  $W_{ij}$

为了更好地表示商超每日的净收益大小, 引进变量  $P_{ij}$  和  $W_{ij}$ , 分别表示第  $i$  个品类第  $j$  天的成本加成定价和第  $i$  个品类第  $j$  天的净利润。

④ 需求量  $Q_{ij}$

根据前文拟合的各个品类销售总量和成本加成定价之间的关系表达式, 可以计算出第  $j$  天第  $i$  个品类在价格  $P_{ij}$  下的需求量  $Q_{ij}$ , 即

$$Q_{ij} = Q_i \cdot P_{ij}$$

#### 4.1.2. 约束条件

① 需求限制

由于大部分蔬菜品种如果当日未售出, 隔日就无法再售。为了避免这种损耗, 可以将需求量  $Q_{ij}$  的下界约束为预测销量  $d_{ij}$ 。

$$d_{ij} \leq Q_{ij}$$

② 利润率限制

为了使商超每日的净收益达到最大, 根据成本加成定价的计算公式“成本加成定价 = 成本  $\times$  (1 + 利润率)”, 可以将各个品类的定价的下界约束为该品类近 3 年平均利润率下的成本加成定价大小。

$$P_{ij} \geq TP_{ij}(1 + \alpha_i)$$

其中,  $\alpha_i$  表示第  $i$  个品类近三年的平均利润率大小。

#### 4.1.3. 目标函数

根据需求限制, 每天陈列的蔬菜会全部售出, 不存在因当日未售出而产生的损耗, 故可以将商超每日的净收益表示收入和成本之差。即目标函数为:

$$\max \sum_{i=1}^6 W_{ij} = \sum_{i=1}^6 (d_{ij} \cdot (P_{ij} - TP_{ij}) - d_{ij} \cdot L_i \cdot TP_{ij}) \quad j=1, \dots, 7$$

## 4.2. 模型的求解与结果

利用 Matlab 中 lsqcurvefit 函数, 可以求得各个品类未来 7 天的日补货总量(如表 5 所示)和定价策略(如表 6 所示), 同时得到该商超未来一周每天的最大收益情况(如表 7 所示)。

**Table 5.** Daily replenishment total for each category in the next 7 days

**表 5.** 各个品类未来 7 天的日补货总量

	2023.7.1	2023.7.2	2023.7.3	2023.7.4	2023.7.5	2023.7.6	2023.7.7
花叶类	213.661	204.281	151.686	139.708	139.425	137.792	164.933
花菜类	33.216	31.765	17.361	16.783	16.440	17.228	21.701
水生根茎类	37.659	32.875	16.891	16.008	15.350	14.095	21.920
茄类	27.857	27.217	19.654	18.220	18.500	20.324	24.363
辣椒类	115.433	110.325	78.636	77.048	77.651	81.868	100.037
食用菌类	68.256	64.649	42.979	40.828	46.538	43.361	57.214

由表 5 可知, 在未来 7 天里, 每天补货量最多的品类是花叶类, 较少的品类是花菜类、水生根茎类和茄类。从整体上看, 可以发现当前一天的补货量较多时, 后一天的补货量就会下降, 这满足一般规律, 能够使得损耗减少, 收益增大。

**Table 6.** Pricing strategies for the next 7 days for each category

**表 6.** 各个品类未来 7 天的定价策略

	2023.7.1	2023.7.2	2023.7.3	2023.7.4	2023.7.5	2023.7.6	2023.7.7
花叶类	5.579	5.543	8.764	10.800	10.856	11.186	7.085
花菜类	12.109	12.194	24.109	25.256	25.982	24.365	17.749
水生根茎类	18.164	17.979	18.497	18.332	18.820	18.558	18.565
茄类	7.660	7.661	7.743	8.997	8.730	7.808	7.809
辣椒类	6.032	6.028	8.096	8.471	8.325	7.403	6.007
食用菌类	6.589	7.191	13.879	15.076	12.210	13.683	8.755

由表 6 可知, 在未来 7 天里, 各个品类每天的价格与日补货量呈现了一种相反趋势。以花叶类为例: 当日补货总量下降时, 其定价上升; 当日补货总量上升时, 其定价下降。说明在某品类蔬菜日补货量上升时, 商超需要适当降低定价来促进其销售, 减少滞销带来的亏损; 在某品类蔬菜日补货量下降时, 商超需要适当提高定价来确保自身收益。

**Table 7.** The trader's daily earnings for the next 7 days

**表 7.** 该商超未来 7 天的日收益

时间	2023.7.1	2023.7.2	2023.7.3	2023.7.4	2023.7.5	2023.7.6	2023.7.7	总计
日收益	3318.960	3152.395	3167.469	3362.555	3311.530	3290.970	2982.832	22586.712

由表 7 可知, 在未来 7 天里, 该商超将在未来第 4 天取得最大日收益, 在未来第 7 天取得最少日收益。整体而言, 该商超日收益比较稳定, 未来 7 天的总收益比较可观。

## 5. 结论

时间序列模型在处理时间序列数据时具有强大的适应性, 能够通过分析历史数据有效地预测未来的发展趋势, 常被用于股票市场预测、销售预测等。与其他模型相比, 时间序列模型往往能提供更为准确的预测结果。因此, 本文基于某商超经销的 6 个蔬菜品类在某段时间的相关销售数据, 通过依次建立关于销售总量与成本加成定价的线性回归模型、关于销售总量和批发价格的时间序列模型以及收益最大化的规划模型, 帮助该商超得到了各品类未来一周净利润最大化的补货和定价策略。在考虑蔬菜品类的销售总量与成本加成定价关系时, 本文采用了价格需求弹性曲线函数来对数据进行拟合, 并将所得拟合方程运用在后续模型建立中, 提升了模型结果的严谨性和科学性。同时通过时间序列预测未来一周的销售总量和批发价格, 保障了模型数据的准确性。

通过改变 6 个品类的弹性系数对模型进行灵敏度分析, 可以发现 6 个品类弹性系数的变化不容易引起收益的较大变化。说明相互替代品关系的存在, 使模型呈现出较好的稳定性。因此, 本文所建立的模



型可以适当推广至其他生鲜产品的定价与补货决策中, 例如: 海鲜类和肉禽类等。

### 参考文献

- [1] 李晓璐, 周曙光. 我国生鲜商超零售业发展问题研究[J]. 商业经济研究, 2021(23): 35-37.
- [2] 苏娟, 方舒, 邢广进, 等. 考虑价格需求弹性的 CS-SVM 短期负荷预测方法[J]. 江苏大学学报(自然科学版), 2022, 43(3): 319-324.
- [3] 杜远福. 价格需求弹性的计算与应用[J]. 商丘师范学院学报, 2000(4): 60-63.
- [4] 徐任超, 阎威武, 王国良, 等. 基于周期性建模的时间序列预测方法及电价预测研究[J]. 自动化学报, 2020, 46(6): 1136-1144.