

# Modeling and Forecasting of Online Women's Sales Volume in E-Commerce Platform

Rongxiang Zhao, Yichao Yue, Siqi Liu, Lintao Zheng, Jin Yang

Dalian Minzu University, Dalian Liaoning  
Email: baixd518@126.com

Received: Apr. 4<sup>th</sup>, 2019; accepted: Apr. 19<sup>th</sup>, 2019; published: Apr. 26<sup>th</sup>, 2019

---

## Abstract

This paper thinks that the sales data of Rakumachi, Only, and Taiping Bird in 2014-2017 can be regarded as a time series. First, the type of time series is determined by time series timing diagram. Then, the data is preprocessed according to different types. Since the sequence has a cluster effect, the heteroscedasticity is eliminated by taking the logarithm. Then, a preliminary mathematical model is proposed in turn, and linear fitting and residual fitting are performed to build a residual autoregressive model. By establishing the most appropriate time series model for the model to continuously identify, fit, test and optimize, the predicted value of the logarithm of the sales volume is obtained, and the sales volume of the three brands is predicted in the next six months.

## Keywords

Electronic Business Platform, Time Series Model, Timing Chart Test, Linear Fit, Forecast

---

## 电商平台线上女装销售量的建模与预测

赵戎祥, 岳逸超, 刘思琪, 郑林涛, 杨 金

大连民族大学, 辽宁 大连  
Email: baixd518@126.com

收稿日期: 2019年4月4日; 录用日期: 2019年4月19日; 发布日期: 2019年4月26日

---

## 摘 要

本文认为乐町、Only、太平鸟2014~2017年的销售量数据可以看成一个时间序列, 首先通过时间序列的

时序图判断其时间序列的类型, 然后根据不同的类型对数据进行预处理, 这里由于序列具有集群效应, 进而通过取对数消除异方差性, 然后依次提出初步的数学模型, 进行线性拟合和残差拟合来建立残差自回归模型, 通过对模型不断识别、拟合、检验和优化, 建立最恰当的时间序列模型, 得出销售量对数的预测值, 并以此预测三个品牌的销售量未来半年的走势。

## 关键词

电商平台, 时间序列模型, 时序图检验, 线性拟合, 预测

Copyright © 2019 by authors and Hans Publishers Inc.

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

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



Open Access

## 1. 前言

随着电子商务的迅速普及和个性化消费趋势的发展, 各种女装品牌在网络上悄然兴起[1]。以淘宝网为例, 截至 2017 年 12 月, 正在运营的个人网店达 900 多万家, 每天访问淘宝网的人数也早已过亿, 并且在持续增加。在如此庞大的消费群体中, 女性网民成为主导消费潮流的新生力量, 因此女装成为最具代表性的线上商品。面对愈演愈烈的市场竞争, 服装品牌企业对运营程度的专业化要求越来越高, 企业管理模式必须以科学管理为重中之重[2]。为了给服装产业提供一个更好的发展前景, 企业需要最大程度地满足顾客需求, 把焦点集中在服装销售方式的发展和变革上。线上与线下如何更好联动发展已成为传统专业市场研究的重大理论和实践问题[3], 服装品牌要想在终端销售方面大展身手, 必须时时掌握最准确的销售趋势。

在服装品牌遇到信息化建设、营销模式转型以及新业务拓展等战略转变时期, 女装销售方式的研究已变得极其重要[4]。因此通过对线上女装销售量的分析来得出某种结论是迫在眉睫的。

## 2. 问题分析

电商平台线上女装销售量可以看成是一个时间序列, 首先通过其时间序列图判断其是否为平稳时间序列, 若不是则需要先进行对数据的预处理, 之后初步提出分析模型, 不断通过对时序图检验、线性拟合及残差检验等建立最优时间序列模型, 并用来预测电商平台上女装销售量的走势。

## 3. 名词解释

1) 时序图: 亦称为序列图或循序图, 是一种 UML 行为图。它通过描述对象之间发送消息的时间顺序显示多个对象之间的动态协作。

2) 时间序列模型: 时间序列模型, 是在生产和科学研究中, 对某一个或一组变量  $x_{(t)}$  进行观察测量, 将在一系列时刻  $t_1, t_2, \dots, t_n$  ( $t$  为自变量) 按照时间次序排列, 并用于解释变量和相互关系的数学表达式。所得到的离散数字组成序列集合  $x(t_1), x(t_2), \dots, x(t_n)$ , 我们称之为时间序列, 这种有时间意义的序列也称为动态数据。根据对系统进行观测得到的时间序列数据, 用曲线拟合方法对系统进行客观的描述。

## 4. 时间序列模型的建立

### 4.1. 时序图检验

当我们需要掌握时间序列的一些基本分布特征时, 最直观的方法是利用一个横轴表示时间、纵轴表

示序列取值的二维平面坐标图，即时序图(Timing chart)。所以时序分析的第一步通常是绘制时序图。R 具有非常便捷与强大的绘图功能，利用 R 绘制时序图主要使用 plot 函数，这里横轴表示以每个月为单位 4 年的时间，纵轴表示销售量。如图 1，图 2，图 3。

三种品牌的时序图清晰地显示，销售量序列都呈现明显的递增趋势，尤其是 2017~2018 年间的销售量远远高于前三年，所以它一定不是平稳序列。除此之外，序列还有不明显的集群效应，这可能是由于季节的变动和近几年电商行业的崛起，我们可以通过差分后平方的时序图可以进一步观察特征。

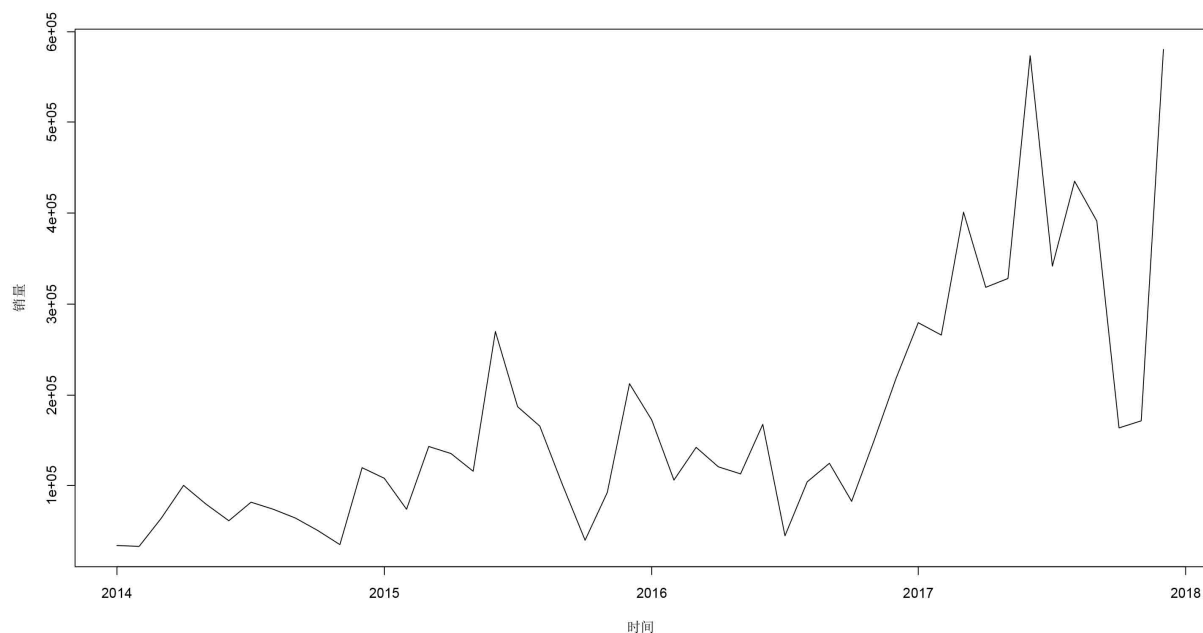


Figure 1. Only timing chart

图 1. Only 时序图

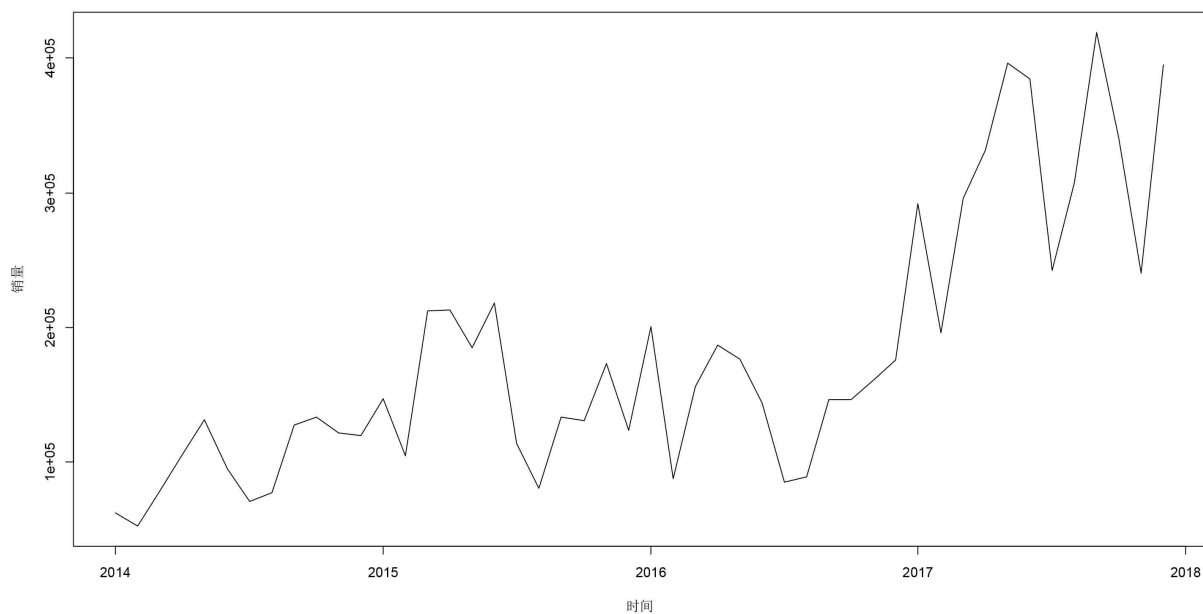


Figure 2. Taiping Bird timing chart

图 2. 太平鸟时序图

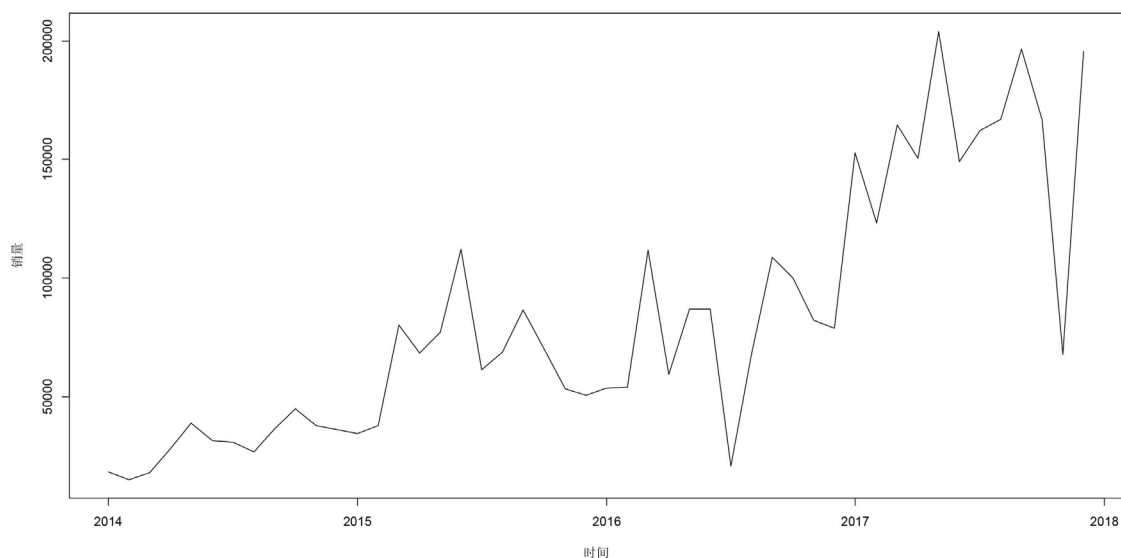


Figure 3. Rakumachi timing chart

图 3. 乐町时序图

## 4.2. 差分

集群效应就意味着在整个序列观察期，序列的方差基本是齐性的，但在某一段或几段时期方差与期望方差有显著差异，这与时序图显示的特征相吻合。确定性时序分析方法和随机时序分析方法是两种常见的时间序列方法，是以无论是采取哪种时间序列分析方法，分析的第一步都是要通过一定的步骤提取序列中所包含的确定性信息。为了可以对序列进行分析，大多数情况下可选择差分法，当序列经过平稳化处理，自然可以进一步观察出序列的特征。差分方法是确定性信息提取方法中比较有效和直接的确定性信息提取方法。在具体的研究中，序列的特点不尽相同，我们也会分别选择不同的差分方式。由于差分方式具有一定的误差，不能轻易使用，但这里为了观察异方差性，我们可以利用差分平方的时序图。差分平方的时序图如图 4、图 5、图 6：

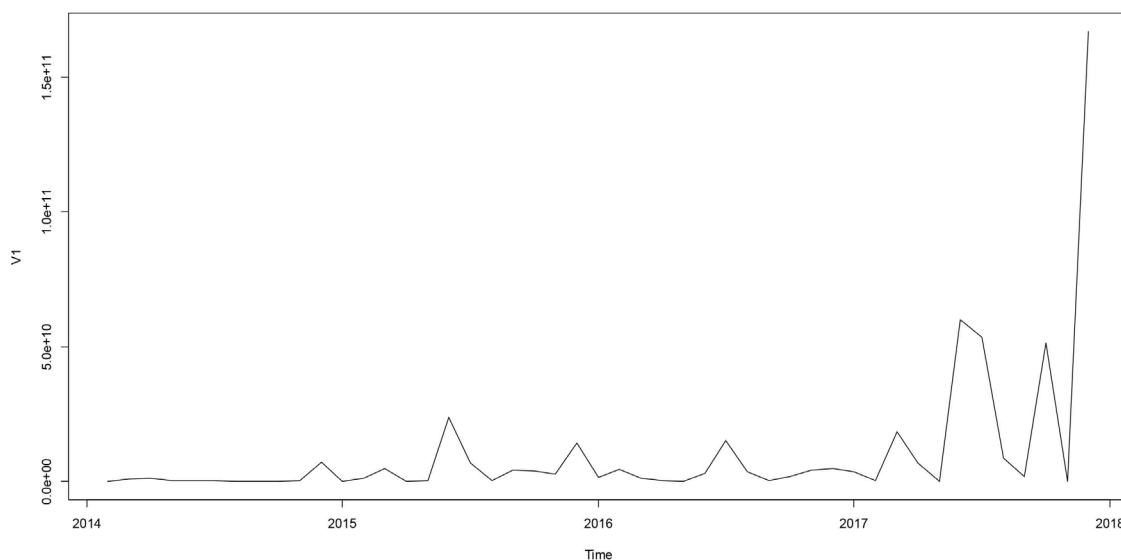


Figure 4. Only timing chart

图 4. Only 时序图

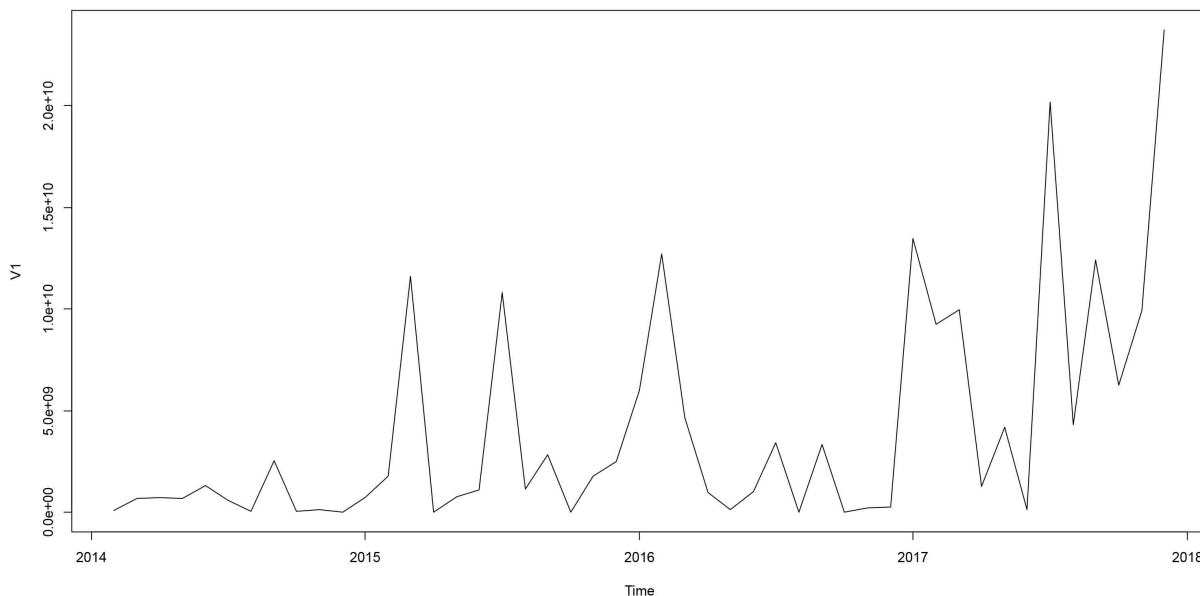


Figure 5. Taiping Bird timing chart  
图 5. 太平鸟时序图

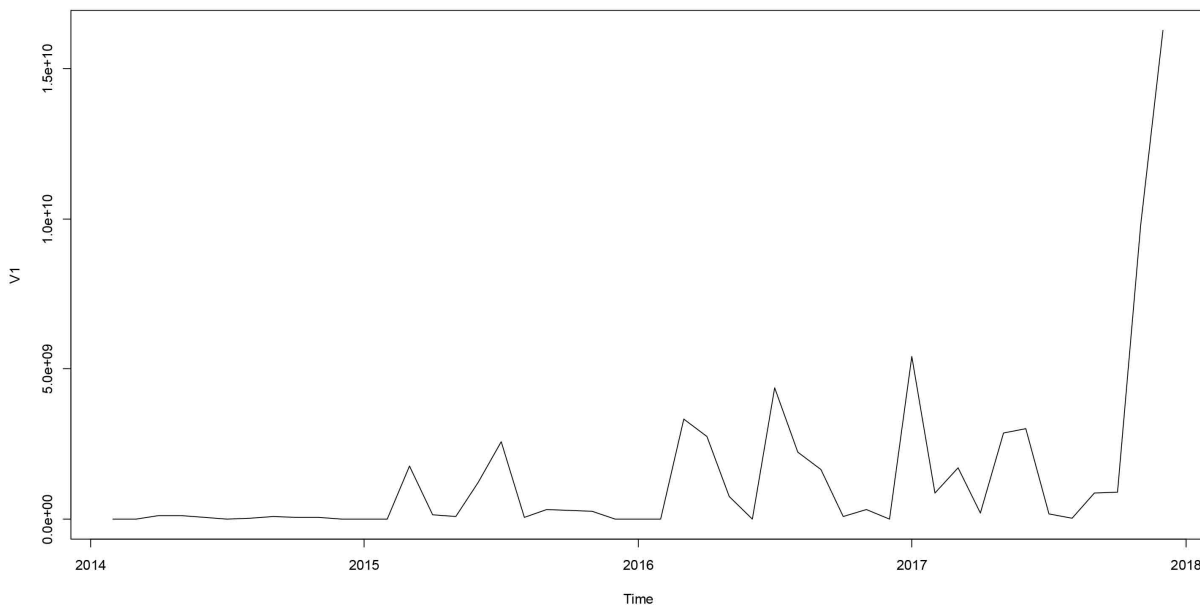


Figure 6. Rakumachi timing chart  
图 6. 乐町时序图

观察三个品牌差分后的序列平方图，序列的波动在大部分时间是平稳的，但在 2017~2018 一年的时间里波动偏大，说明不同时期的方差不等，这是具有异方差性的典型特征。忽视异方差的存在会导致残差的方差被严重低估，继而导致模型的拟合精度受影响，所以要采取有效的手段消除异方差性。

### 4.3. 消除异方差性

通过检查数据的平稳性，发现三组数据的异方差性，根据所学的有限知识，可以通过对数据取对数的方法消除异方差性，从而对销售量的对数进行建模。取对数时序图如下：(图 7~图 9)

通过观察如上的对数时序图，序列有着明显的递增趋势。ARIMA 模型虽然可以对非平稳序列拟合，但很难科学地对模型进行直观的解释。所以，这里可以使用更具现实意义的残差自回归模型，分两步提取序列的确定性信息。

#### 4.4. 拟合时间序列模型

##### 4.4.1. 线性拟合及残差检验

和大部分模型一样，残差自回归模型也是提取序列中主要的确定性信息的一种方式，它通过确定性因素分解方法首先对趋势进行线性拟合。因为序列呈现明显的递增趋势，这里进行线性拟合的基本原理就是首先提取序列的长期递增信息。

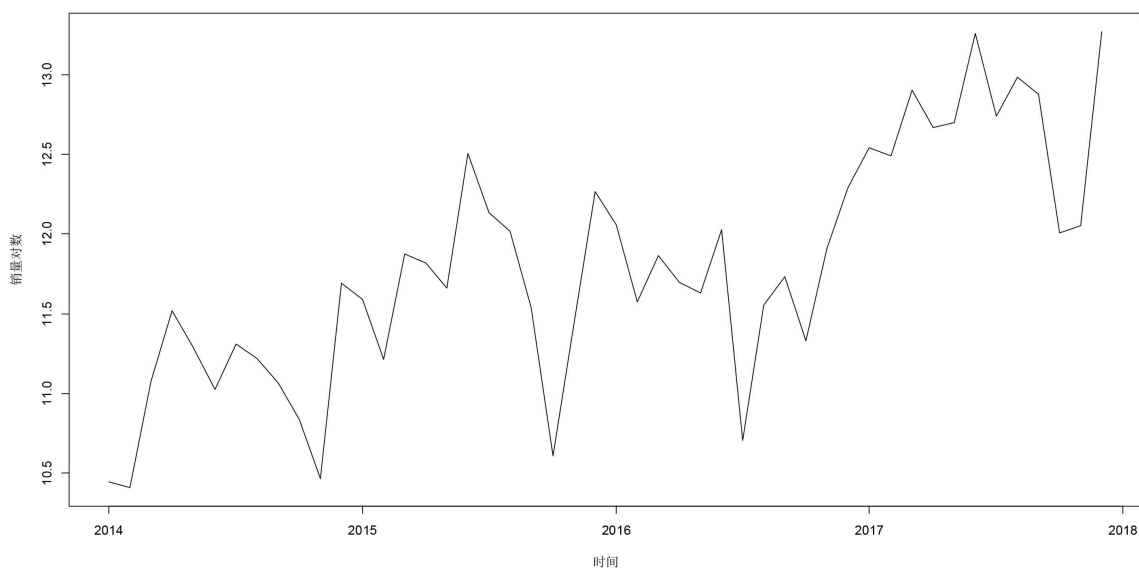


Figure 7. Only timing chart

图 7. Only 时序图

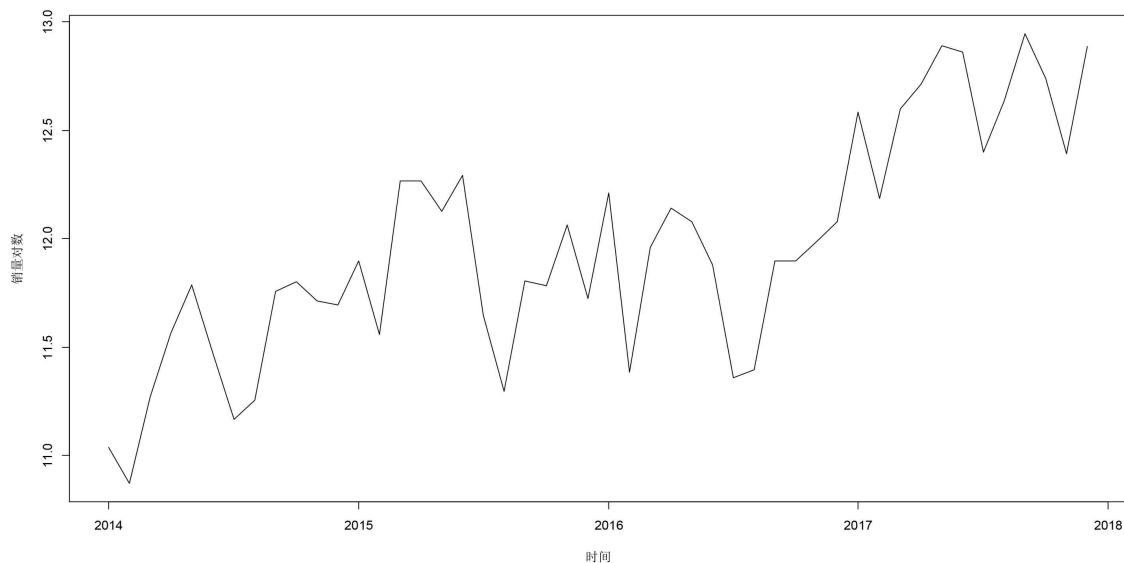
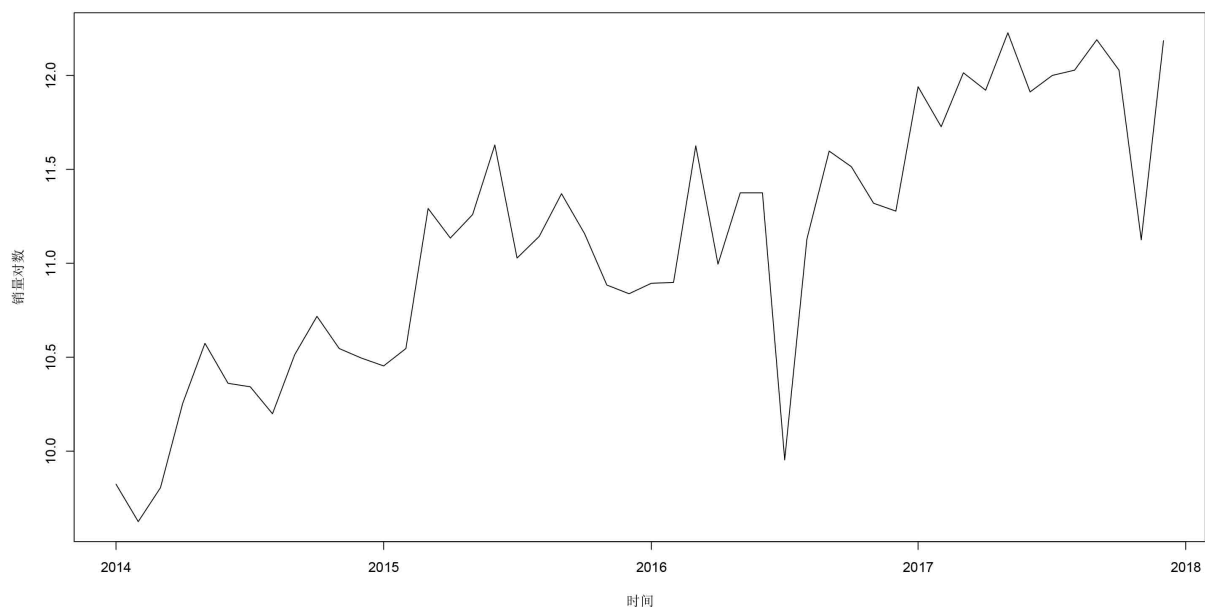


Figure 8. Taiping Bird timing chart

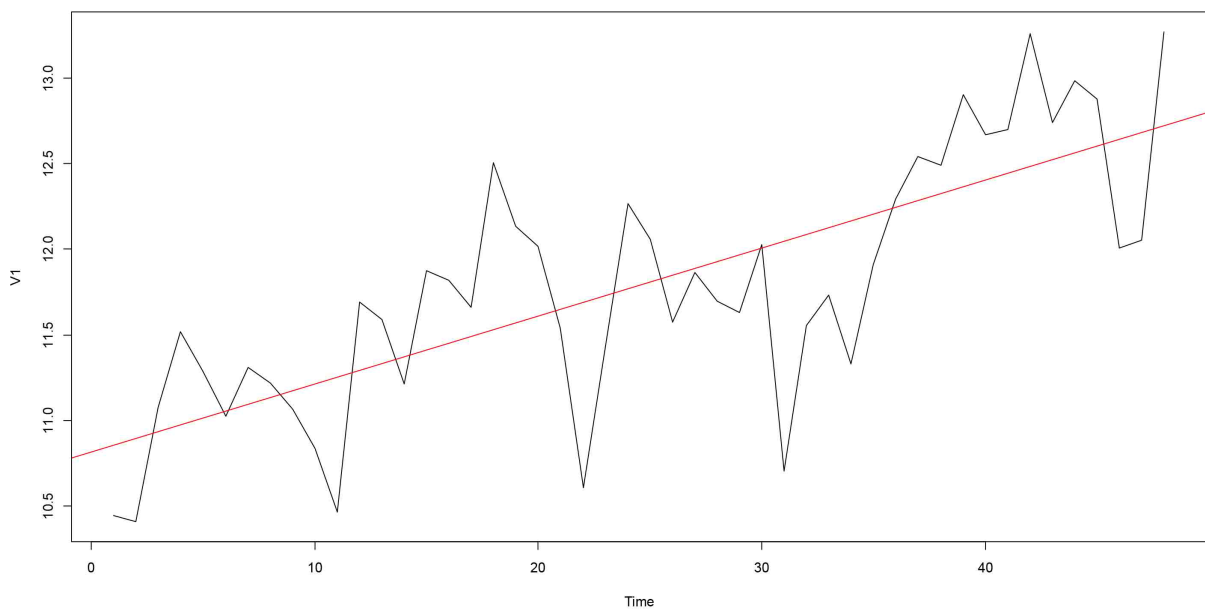
图 8. 太平鸟时序图



**Figure 9.** Rakumachi timing chart  
**图 9.** 乐町时序图

其次，接下来需要对拟合效果进行检验，这是确定模型后的必要步骤。这里采用 DW 检验，可以进行残差自相关性检验。若是残差序列显示出显著的自相关性，就说明序列中的相关信息提取的不够充分，线性拟合的精确度不高，此时需要提取其中残存的相关信息，对残差序列继续拟合，以此提高模型的精度。

拟合效果如图 10。



**Figure 10.** Only timing chart  
**图 10.** Only 时序图

线性回归模型为：

$$x_t = 10.816522 + 0.039712t + \varepsilon_t$$

残差检验结果显示, DW 统计量的值小于 2, 且 p 值小于显著性水平(0.05), 这说明残差序列高度正相关, 为了充分提取相关信息, 我们需要对残差序列继续拟合。

拟合效果如图 11。

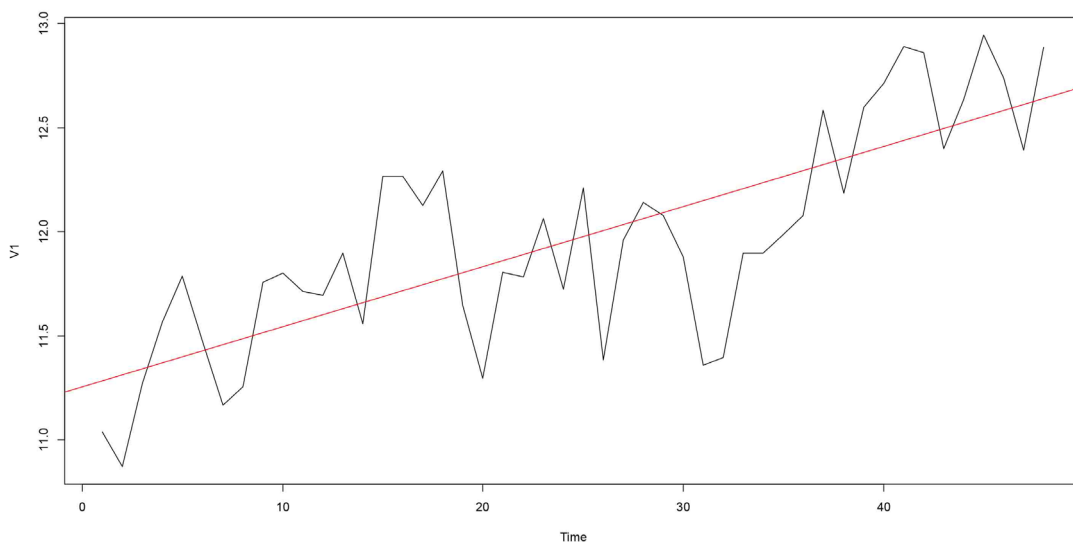


Figure 11. Taiping Bird timing chart

图 11. 太平鸟时序图

线性回归模型为:

$$x_t = 11.256435 + 0.02883t + \varepsilon_t$$

通过残差检验显示, 可以看出结果与上一个品牌类似, DW 统计量的值小于 2, 且 p 值远远小于显著性水平(0.05), 这说明残差序列高度正相关, 为了充分提取相关信息, 我们需要对残差序列继续拟合。

拟合效果如图 12。

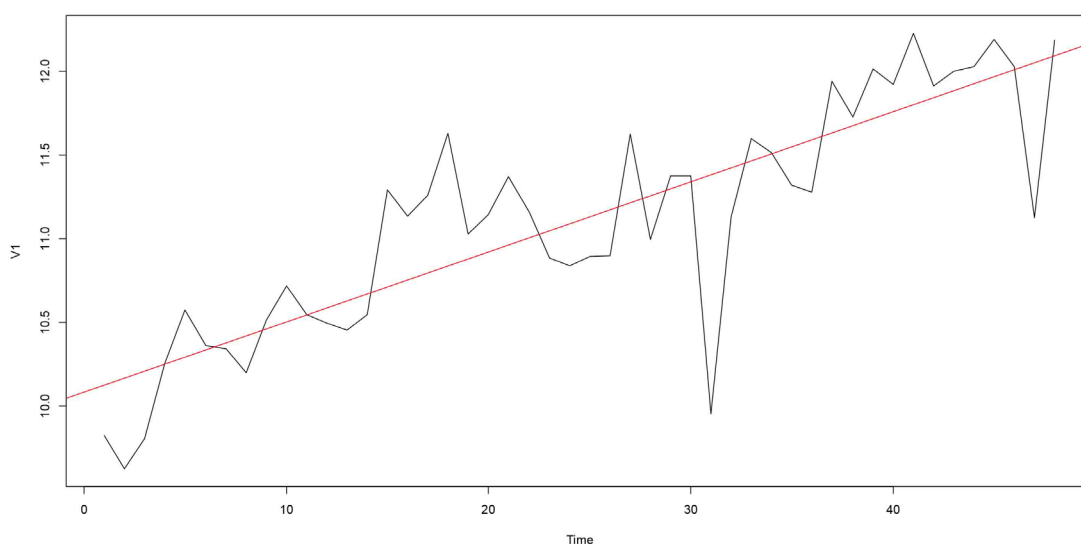


Figure 12. Rakumachi timing chart

图 12. 乐町时序图



线性回归模型为:

$$x_t = 10.084547 + 0.041794t + \varepsilon_t$$

残差检验结果如上, DW 统计量的值小于 2, 且 p 值小于显著性水平(0.05), 这说明残差序列高度正相关, 为了充分提取相关信息, 我们需要对残差序列继续拟合。

#### 4.4.2. 残差拟合

残差时序图如图 13。

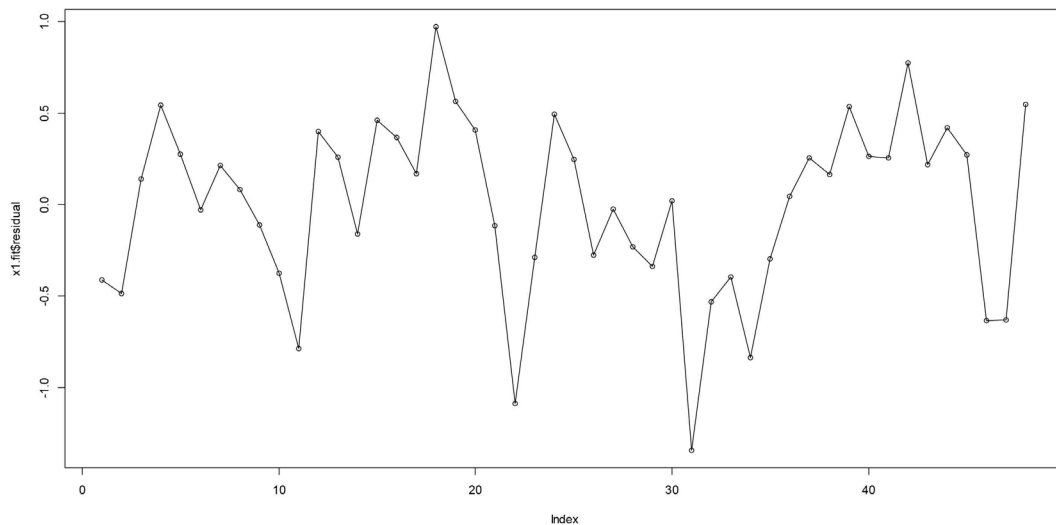


Figure 13. Only residual timing chart

图 13. 残差时序图

观察图 13 的残差序列, 时序图呈现基本平稳的特征, 因此可以进一步观察自相关图和偏自相关图。自相关图 Autocorrelation Chart 如图 14, 偏自相关图 Partial Autocorrelation Chart 如图 15。

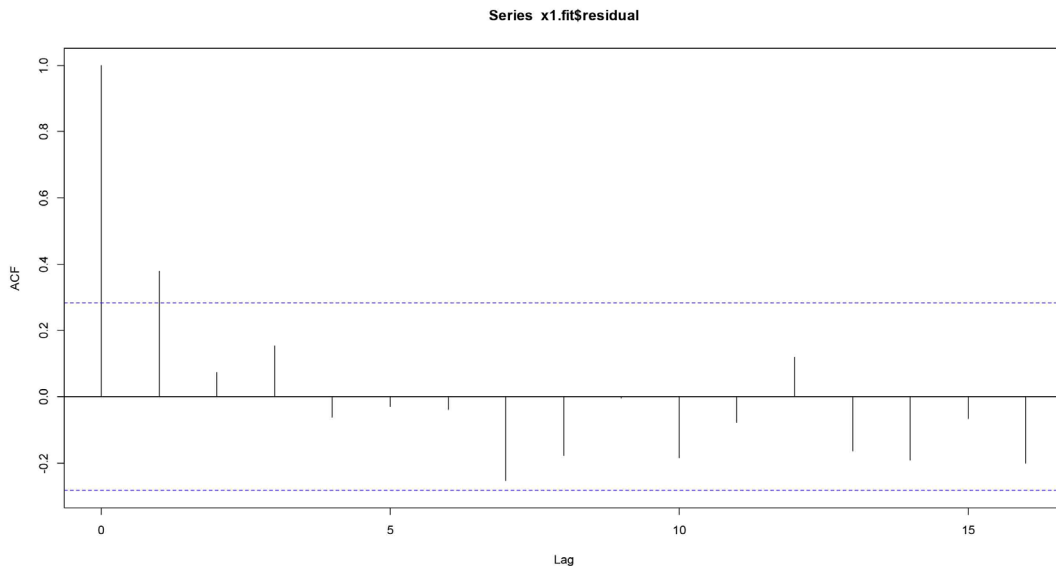


Figure 14. Autocorrelation chart

图 14. 自相关图

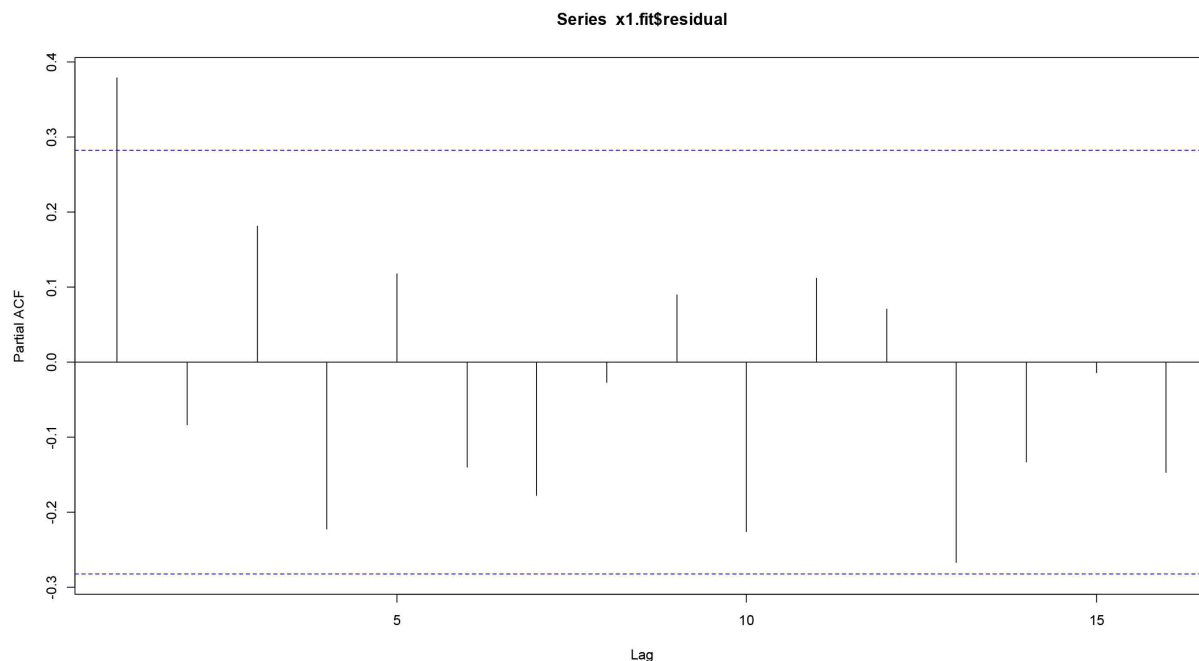


Figure 15. Partial autocorrelation chart  
图 15. 偏自相关图

自相关图和偏自相关图显示出来的特征和该序列时序图平稳的性质是非常吻合的。所以判定此时该序列为平稳时间序列。根据残差序列的相关图和偏自相关图呈现的性质，偏自相关图 1 阶截尾，可以考虑拟合 ARMA(1, 1)模型和 AR(1)模型，从通过检验的模型中选择最优模型。

从检验结果看，两种模型的残差序列延迟 6 阶与 12 阶的  $p$  值都大于显著性水平，接受原假设，认为残差显著自相关，显示两种拟合模型都成立，进而通过 AIC、BIC 准则进行模型优化。为了选择最优模型，其中一种有效的规则就是 AIC 准则，其基本规则是在所有通过检验的模型中选择使得 AIC 函数达到最小的模型，其为相对最优模型[5]。

第二个模型的 AIC 和 BIC 分别为 59.81、64.88，都小于第一个模型，根据 AIC 准则，确定最优模型为 ARMA(1, 1)，拟合结果为：

$$\begin{cases} x_t = 10.816522 + 0.039712t + \varepsilon_t \\ \varepsilon_t = -0.2631\varepsilon_{t-1} + 0.8118e_{t-1} + e_t, e_t \sim N(0, 0.1828) \end{cases}$$

残差时序图如图 16：

观察图 16 的残差序列，时序图呈现基本平稳的特征，因此可以进一步观察自相关图和偏自相关图。自相关图 Autocorrelation Chart 如图 17，偏自相关图 Partial Autocorrelation Chart 如图 18。

自相关图和偏自相关图显示出来的特征和该序列时序图平稳的性质是非常吻合的。所以判定此时该序列为平稳时间序列。由图 18 可以看出偏自相关图 1 阶截尾，同样可以考虑拟合 ARMA(1, 1)模型和 AR(1)模型，从通过检验的模型中选择最优模型。

从检验结果看，两种模型的残差序列延迟 6 阶与 12 阶的  $p$  值都大于显著性水平，接受原假设，认为残差显著自相关，显示两种拟合模型都成立，进而通过 AIC、BIC 准则进行模型优化。

第二个模型的 AIC 和 BIC 分别为 23.5、26.97，都小于第一个模型，根据 AIC 准则，确定最优模型为 AR(1)，拟合结果为：

$$\begin{cases} x_t = 11.256435 + 0.02883t + \varepsilon_t \\ \varepsilon_t = 0.4492\varepsilon_{t-1} + e_t, e_t \sim N(0, 0.8884) \end{cases}$$

残差时序图如图 19。

观察图 19 的残差序列，时序图呈现基本平稳的特征，因此可以进一步观察自相关图和偏自相关图。自相关图 Autocorrelation Chart 如图 20，偏自相关图 Partial Autocorrelation Chart 如图 21。

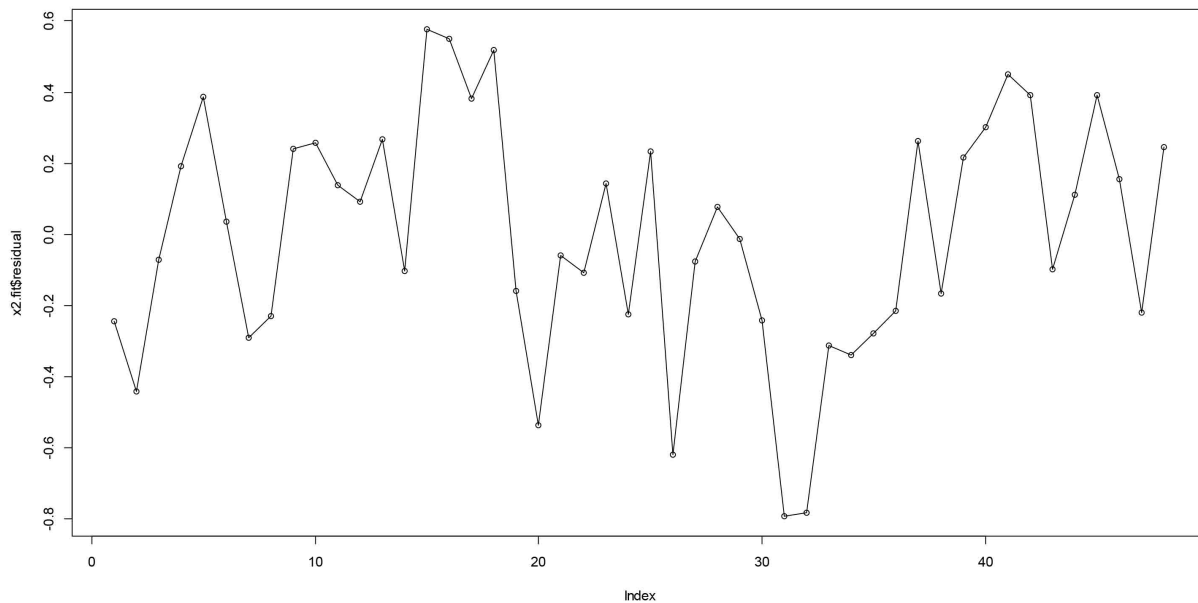


Figure 16. Taiping Bird residual timing chart

图 16. 太平鸟残差时序图

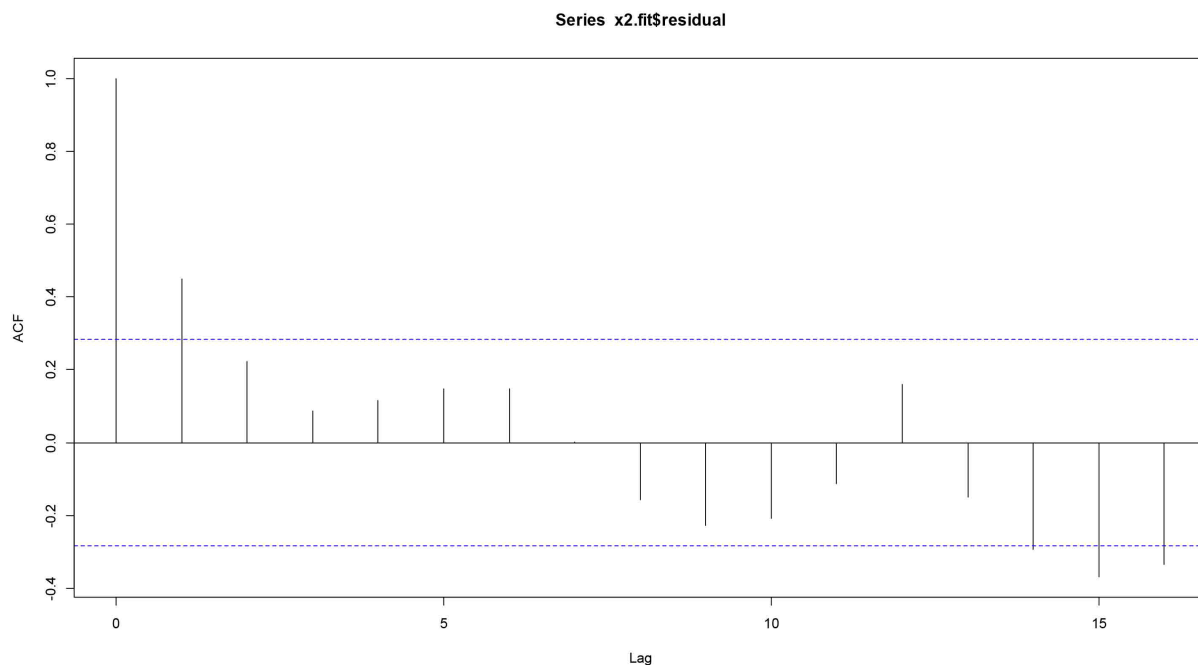


Figure 17. Autocorrelation chart

图 17. 自相关图

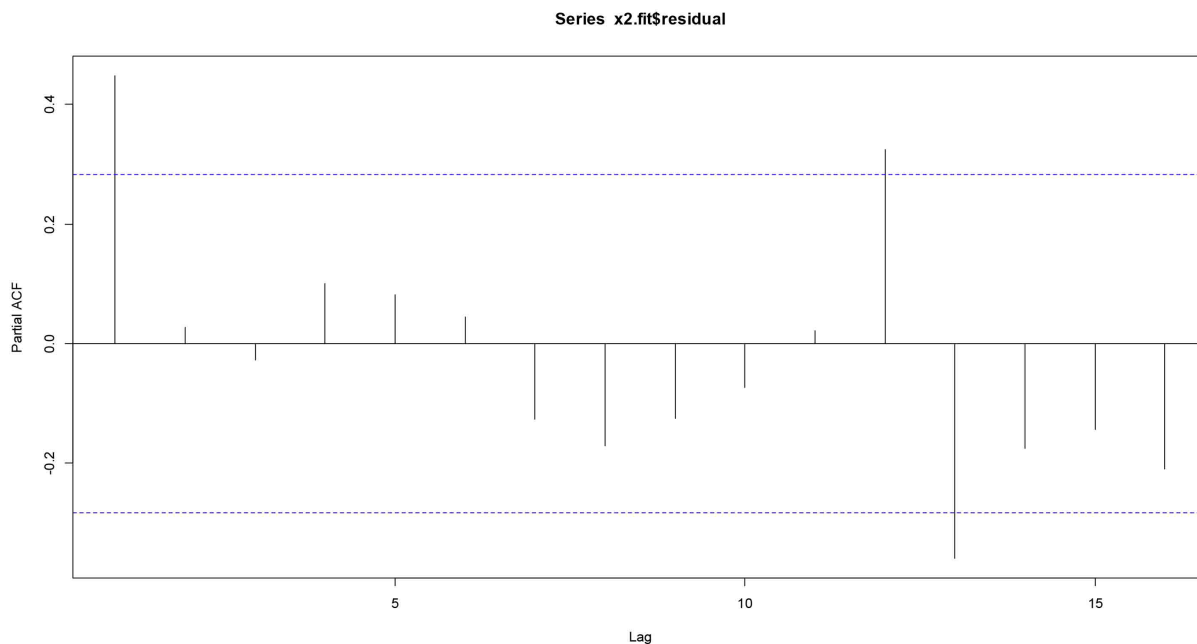


Figure 18. Partial autocorrelation chart  
图 18. 偏自相关图

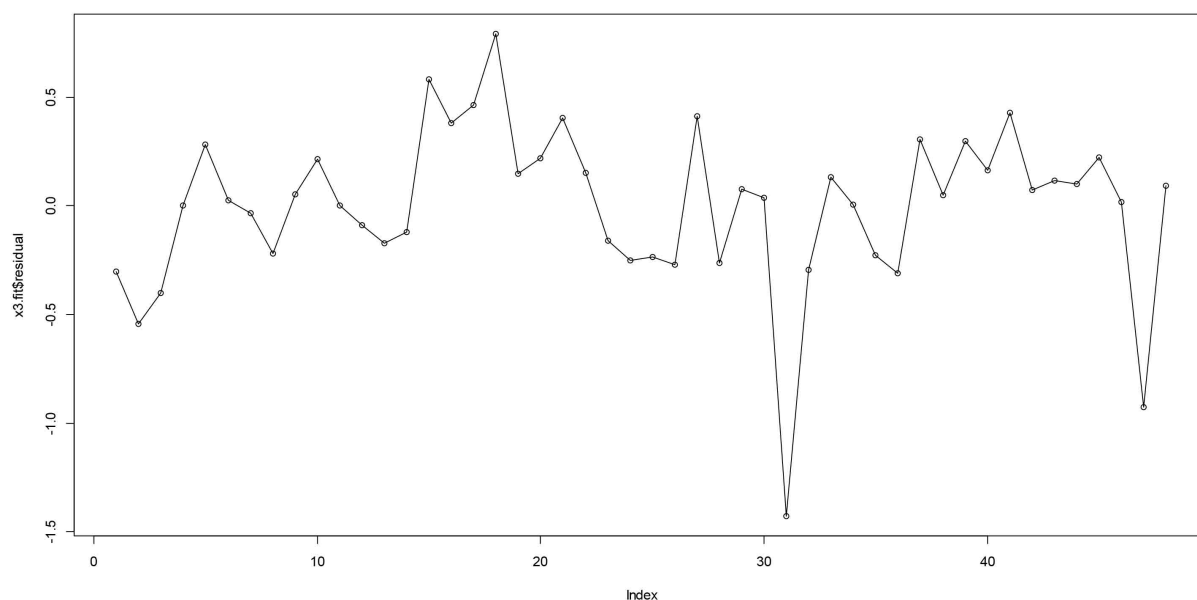


Figure 19. Rakumachi residual timing chart  
图 19. 乐町残差时序图

自相关图和偏自相关图显示出来的特征和该序列时序图平稳的性质是非常吻合的。所以判定此时该序列为平稳时间序列。根据残差序列的相关图和偏自相关图呈现的性质，自相关图 1 阶截尾，也可以考虑拟合 ARMA(1, 1)模型和 AR(1)模型，从通过检验的模型中选择最优模型。

从检验结果看，两种模型的残差序列延迟 6 阶与 12 阶的  $p$  值远远都大于显著性水平，接受原假设，认为残差显著自相关，显示两种拟合模型都成立，进而通过 AIC、BIC 准则进行模型优化。

第二个模型的 AIC 和 BIC 分别为 38.16、43.23，都小于第一个模型，根据 AIC 准则，确定最优模型

为 ARMA(1,1), 拟合结果为:

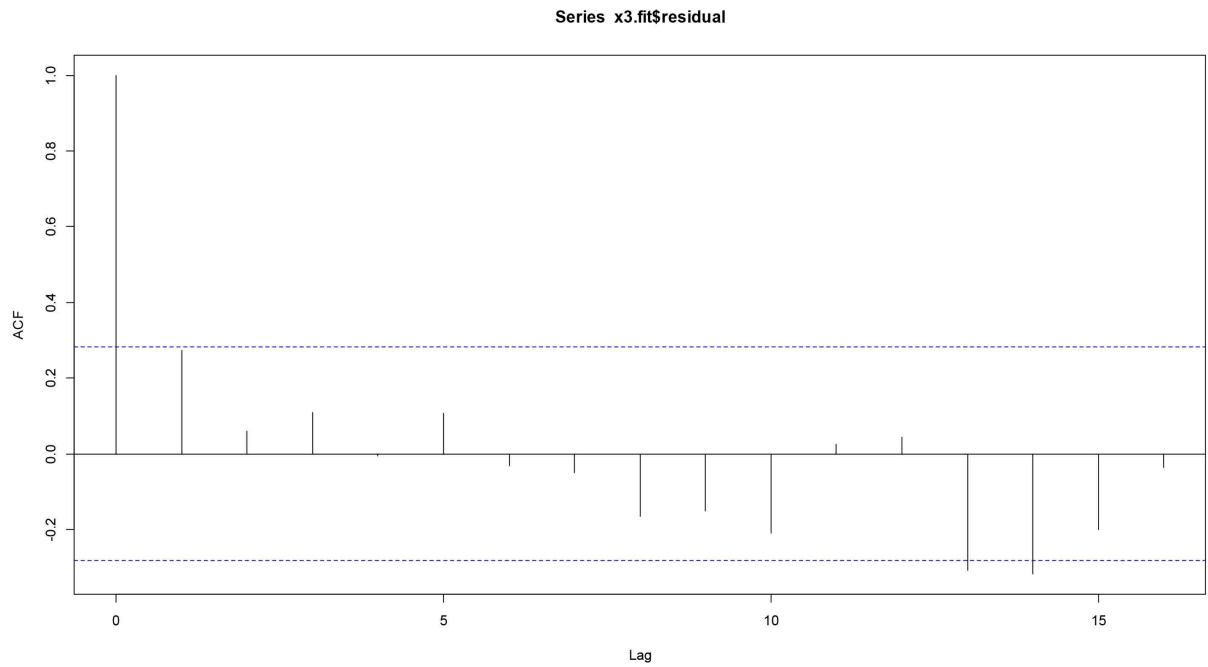


Figure 20. Autocorrelation Chart  
图 20. 自相关图

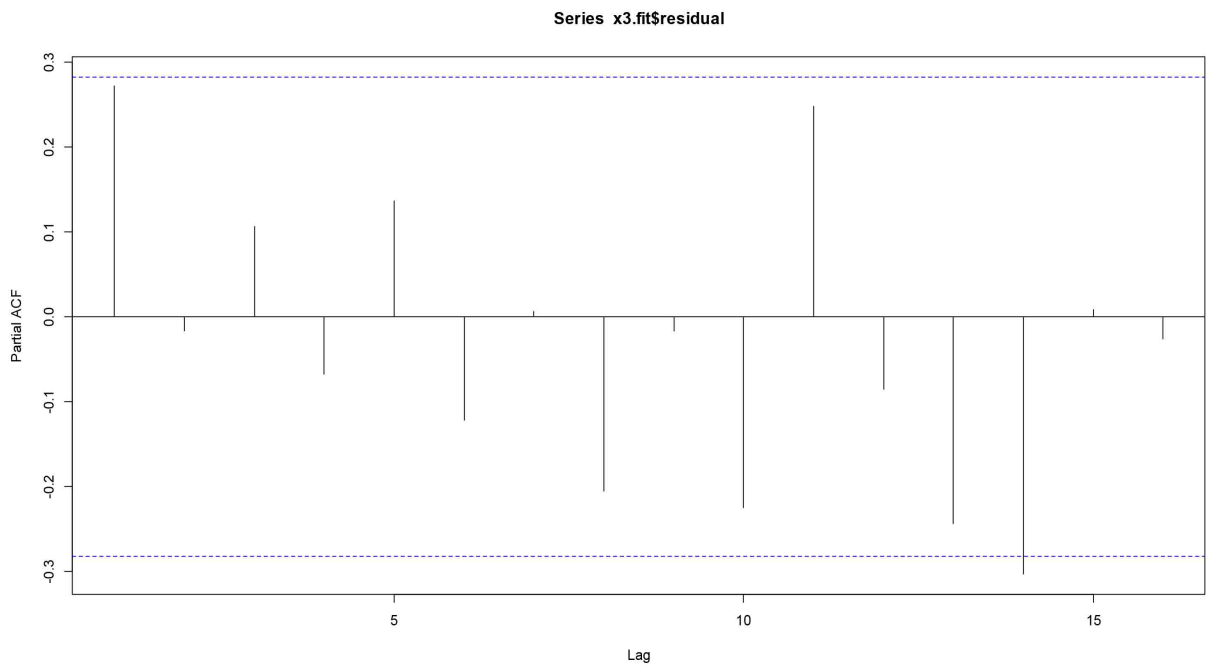


Figure 21. Partial autocorrelation chart  
图 21. 偏自相关图

$$\begin{cases} x_t = 10.084547 + 0.041794t + \varepsilon_t \\ \varepsilon_t = -0.5638\varepsilon_{t-1} + 0.9427e_{t-1} + e_t, e_t \sim N(0, 0.1155) \end{cases}$$

## 5. 预测

包括平稳性检验、取对数消除异方差性、线性拟合、残差检验、残差拟合及模型检验、模型优化在内的所有工作，是我们到目前为止为了预测所做的铺垫。通过观察序列已知的样本值，寻找规律来建立模型，对未来某个时刻或一段时间内的取值进行估计，这个过程称之为预测。由于我们是对销售量的对数进行建模，得出的预测结果也一定是对数，以此得出未来五个月的销售量对数，反应未来的销售量趋势。

#only

预测值如下：我们可以查询到 2018 年 1 月 only 的线上真实销售量数值，经计算得对数为 12.75023，以此作为真实值，预测值为 13.16904，计算出相对误差为 0.033。趋势图如图 22。

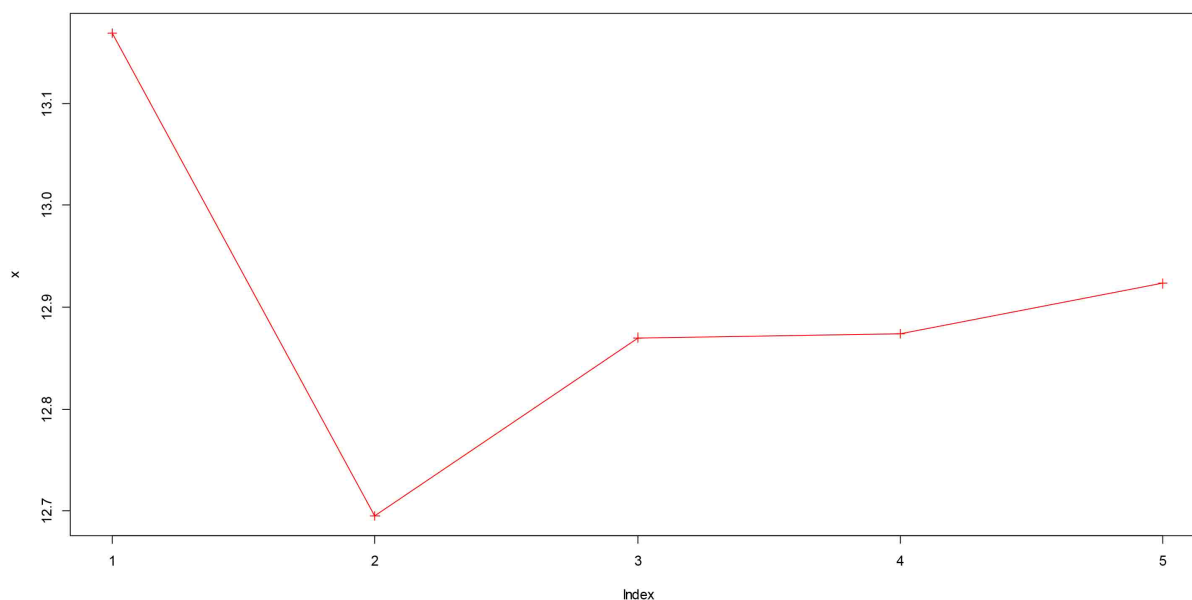


Figure 22. Tendency chart

图 22. 趋势图

#太平鸟 Taiping Bird

预测值如下：我们可以查询到 2018 年 1 月太平鸟的线上真实销售量数值，经计算得对数为 12.63014，以此作为真实值，预测值为 12.88189，计算出相对误差为 0.02。趋势图如图 23。

预测值如下：我们可以查询到 2018 年 1 月的乐町线上真实销售量数值，经计算得对数为 11.82221，以此作为真实值，预测值为 12.28687，计算出相对误差为 0.039。趋势图如图 24。

从三种品牌的趋势图看出未来五个月的销售量总体呈先下降后上升趋势。only 和太平鸟 2、3 月份线上销售量会遭遇低谷，这与双 12 过后网民的消费能力相关[6]，并且 2 月份的春节也会对大家的线上消费造成影响，假期的到来、实体店的折扣让更多女性朋友选择实体店消费[7]。因此，在经历两个月的低迷后，卖家不用过度忧虑，行业将在春季迎来另一个消费风潮，备足春夏装迎接这个时期才是可取之道[8]。尤其是更年轻化的乐町，3 月份将会持续上升。此外，为了迎接网购年龄的日益年轻化，淘宝原创品牌的产生呼之欲出，在此同时，女装企业也应在照顾门店的基础上迎接更多线上销售挑战。

## 6. 总结

本文根据 2014~2017 年三个线上女装品牌 48 个月的销售量统计资料，针对序列的集群效应，通过取

对数消除异方差性，建立残差自回归模型，并进行未来五个月的预测。计算结果表明，该模型能较好地解决线上女装销售量对数的估计和预测问题，预测精度较为精确。服装行业可以以此趋势作为参考，对未来的营销策略做出调整，实现最有效的营销，达到卖家与买家双赢的效果，为实现“行业 + 互联网”新模式打下坚实的基础。然而，由于序列的异方差性，我们只能对销售量对数进行建模预测，因此可以看出销售量的大致趋势，却不能得出销售量的预测值，这其中的误差对企业战略的制定带来不可避免的干扰。

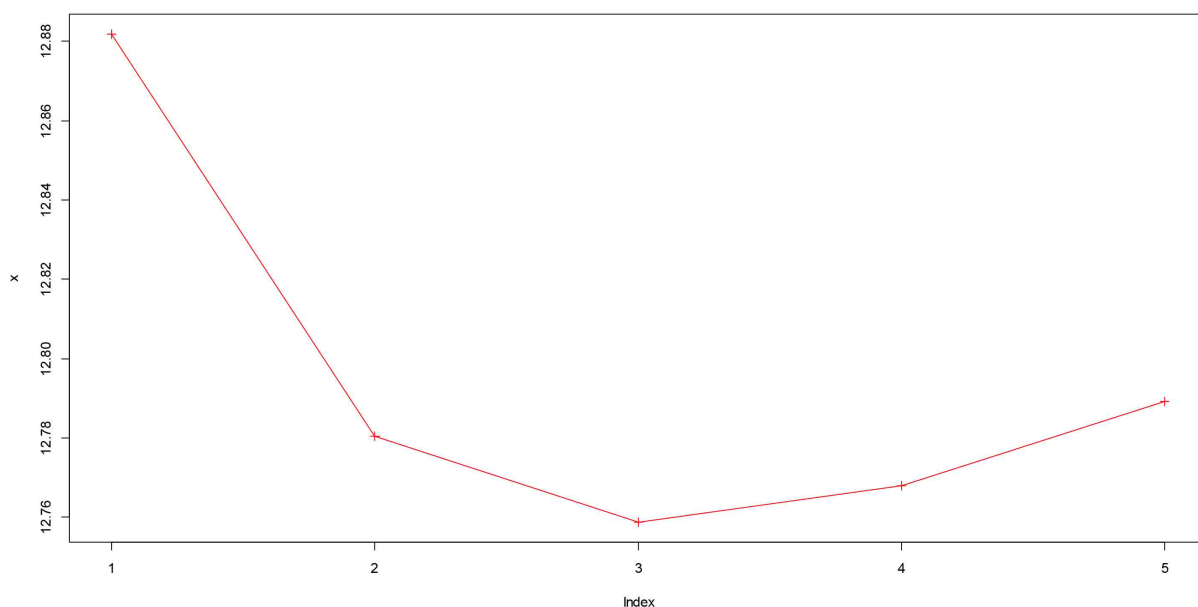


Figure 23. Tendency chart

图 23. 趋势图

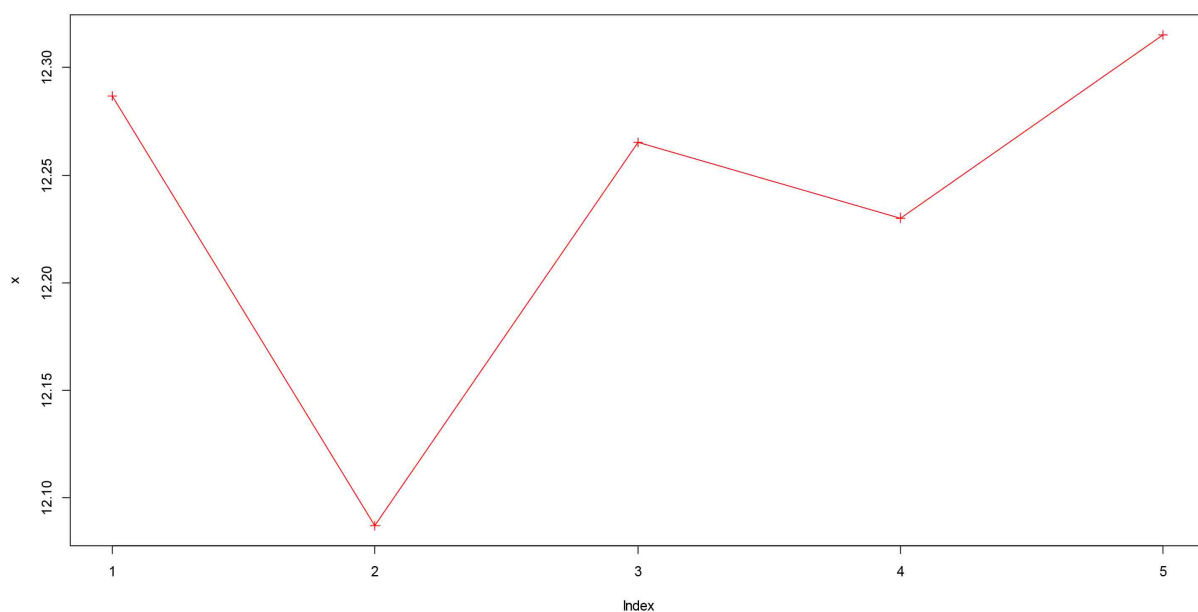


Figure 24. Tendency chart

图 24. 趋势图

在今后的研究中, 我们可以通过学习更多的时间序列模型来完成时间序列的拟合, 比如条件异方差模型就是适用于具有集群效应的序列。由于在实际情况下, 样本大小肯定是有限的, 因此有些情况还需要采用二次加权过程, 以便获得更有效的系数估计, 这是对异方差变系数模型进行统计推断的主要证据。不同的时间序列都有适合自己的建模方式, 因此, 具体问题要具体分析。

## 参考文献

- [1] 蔡建梅, 李易儒. 基于淘宝网络销售平台的原创女装品牌发展策略研究[J]. 浙江理工大学学报, 2012, 29(4): 526-531.
- [2] 万艳敏, 黄珊, 张斯雯. 女装品牌直营终端销售业务流程指标体系研究[J]. 北京服装学院学报, 2013, 33(2): 12-17.
- [3] 文春英, 张淑梅. 网络购物说服策略的运用-基于淘宝网女装销售文本的内容分析[J]. 新媒体, 2013(5): 77-80.
- [4] 李家华, 王睿, 陈佩芳. 浅谈电子商务模式下女装服饰销售策略分析[J]. 西部皮革, 2016, 38(12): 71-71.
- [5] 王燕. 时间序列分析: 基于 R [M]. 北京: 中国人民大学出版社, 2015: 94, 162.
- [6] 中华全国商业信息中心. 女装上位[J]. 商业消费趋势, 2014(7): 92-93.
- [7] 艾维娜, 杨坚争. B2B 电子商务平台发展及存在的问题[J]. 电子商务, 2018(1): 9-10.
- [8] 女装商战跳过春装买夏装[J]. 纺织服装周刊, 2015(15): 77.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [sa@hanspub.org](mailto:sa@hanspub.org)