

基于 ARIMA 模型的卷烟销量预测研究

易治国¹, 向立梨²

(1.中国烟草总公司深圳市公司, 广东 深圳 518048; 2.深圳烟草工业有限责任公司, 广东 深圳 518109)

摘要:近几年全国的卷烟销量总体平稳,符合烟草行业持续稳健和高质量发展的指导要求。本文以深圳市 2009 年至 2019 年的卷烟季度销量数据作为时间序列,采用 ARIMA 模型,对深圳卷烟市场未来一年的销量进行短期预测,为卷烟市场的合理投放,提供一定的参考价值 and 指导意义。

关键词: 卷烟销量; ARIMA 模型; 销量预测

【DOI】10.12231/j.issn.1000-8772.2021.11.149

1 引言

深圳卷烟市场经历了 2015 年的烟草行业“提税顺价”和 2016 年的“去库存”后,卷烟销量趋于平稳且略有回升。合理有效的把握消费者需求,提升按需求组织货源和投放市场的能力,加快推动卷烟市场化取向改革,离不开卷烟市场的准确判断和预测。因此,提高卷烟销量预测的准确度,对于改善卷烟供求关系、保持卷烟市场良好的发展状态和提升卷烟营销水平具有重要意义。

2 卷烟销量预测模型构建

2.1 ARIMA 模型

时间序列简单的说就是各时间点上形成的数值序列。卷烟销售受季节性和节日因素影响较大,时间序列的季节周期性是指每经过一定时间间隔后,这组数据都呈现出某种相似性。从卷烟销售的实际情况来看,当期或后期的卷烟销量不仅与前各期的销量值有关,还与随机干扰项有关;ARIMA 模型就是一种适合数据具有稳定的季节性特征的短期预测方法,具有较强的灵活性,为卷烟季度或月度的销量预测,提供了较好的应用价值,预测结果较为可靠。

2.2 数据处理

根据深圳烟草 2009 年至 2019 年的卷烟季度销量数据,制作销量时序图,如图 1 所示。

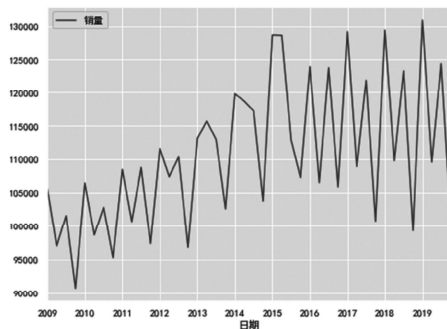


图 1 2009 年-2019 年季度销量时序图

从图 1 来看,卷烟销量具有较强的递增趋势,初步判断时间序列非平稳。对卷烟季度销量时间序列进行平稳性检验,从原始序列的单位根检验结果来看,单位根统计量对应的 P 值显著大于 0.05,最终判断该序列为非平稳序列。对卷烟销量进行一阶差分处理,不难发现,一阶差分后序列的时序图在均值附近比较平稳的波动,接近一个平稳序列。卷烟季度销量一阶差分序列的单位根检验结果和白噪声检验结果所示。从单位根检验结果来看,单位根检验输出的 P 值远小于 0.05,所以一阶差分后的序列是平稳序列;再根据差分序列的白噪声检验结果来看,输出 P 值也远小于 0.05,所以一阶差分之后的序列是平稳非白噪声序列。

2.3 模型识别与定阶

对一阶差分之后的平稳非白噪声序列拟合 ARIMA 模型,建立 ARIMA(p,d,q)模型,其中 d=1,为确认模型中 p 和 q 的值,以 BIC 准则为评价该时间序列数据的准则,取其中 BIC 信息量达到最小的模型阶数。利用 Python 语言的 statsmodels 模块进行测试,来确定 p、q 的阶数,进行相对最优模型识别。根据计算出数值。然后输出一份模型报告如下图 2 所示。从图 2 可以发现,MA1 和 MA2 的 P 值远低于显著性水平 0.05,构建的 ARIMA(0,1,2)模型符合要求。

表 1

| adf | cValue | | | p 值 |
|----------|---------|---------|---------|------------|
| | 1% | 5% | 10% | |
| -14.7167 | -3.6056 | -2.9371 | -2.6070 | 2.7925E-27 |

表 2

| stat | p 值 |
|---------|----------|
| 28.1233 | 1.14E-07 |

Results: ARIMA

| | | | |
|---------------------|------------------|----------------------|------------|
| Model: | ARIMA | BIC: | 921.6632 |
| Dependent Variable: | D. 销量 | Log-Likelihood: | -453.31 |
| Date: | 2020-12-02 16:28 | Scale: | 1.0000 |
| No. Observations: | 43 | Method: | css-mle |
| Df Model: | 3 | Sample: | 06-30-2009 |
| Df Residuals: | 40 | | 12-31-2019 |
| Converged: | 1.0000 | S.D. of innovations: | 8889.450 |
| No. Iterations: | 28.0000 | HQIC: | 917.216 |
| AIC: | 914.6184 | | |

| | Coef. | Std.Err. | t | P> t | [0.025 | 0.975] |
|-------------|----------|----------|----------|--------|-----------|-----------|
| const | 438.3978 | 343.1731 | 1.2775 | 0.2088 | -234.2092 | 1111.0048 |
| ma.L1.D. 销量 | -1.4463 | 0.1006 | -14.3808 | 0.0000 | -1.6434 | -1.2492 |
| ma.L2.D. 销量 | 0.6962 | 0.1054 | 6.6046 | 0.0000 | 0.4896 | 0.9027 |

| | Real | Imaginary | Modulus | Frequency |
|------|--------|-----------|---------|-----------|
| MA.1 | 1.0388 | -0.5978 | 1.1985 | -0.0831 |
| MA.2 | 1.0388 | 0.5978 | 1.1985 | 0.0831 |

图 2 模型报告图

2.4 模型预测检验

将预测长度设定为 3 个季度,利用该模型对 2020 年前三个季度的卷烟销量进行预测,预测数据如下表 3 所示。

表 3 2020 年前三季度销量预测数据表

| 项目 | 预测结果 | 标准误差 | 置信区间 |
|-----|------------|----------|-----------------------|
| 一季度 | 123,874.72 | 7,556.03 | 109,065.17-138,684.27 |
| 二季度 | 115,657.05 | 8,274.38 | 99,439.55-131,874.54 |
| 三季度 | 116,029.68 | 8,487.04 | 99,395.40-132,663.98 |

根据预测数据计算得到 2020 年前三个季度的卷烟销量为 355,561 箱,与 2020 年前三个季度的卷烟实际销量相差 8,794 箱,误差率为 2.41%,误差在可接受的范围内。另外,通过比较 2020 年四个季度的销量预测数据和实际销售数据来看,总体误差率低于 1%,表明 ARIMA 模型对于预测卷烟短期销量具有较强的指导意义。

3 结束语

本文通过对深圳市 2009 年至 2019 年卷烟销售的季度销量时间序列进行纯随机性和平稳性检验,最后选择 ARIMA 模型进行分析,采用 BIC 信息量进行相对最优模型的识别,从而确定相对最优模型的阶数。根据销量预测结果与实际销量的比较来看,预测的误差较小,应用效果明显。结合当前卷烟市场的销售状态,依据卷烟销量的预测结果,为制定卷烟销售计划,科学安排销售进度提供了较好的参考价值。

参考文献

- [1]程幸福,陈厚铭.季节 ARIMA 模型在企业销量预测中的应用-以销售卷烟为例[J].中国商论,2016,8(23).
- [2]吴明山,王冰,等.卷烟销量组合预测模型研究[J].中国烟草学报,2019,(7).

作者简介:易治国(1977,3-),男,湖南人,汉,本科,注册会计师,单位:中国烟草总公司深圳市公司,研究方向:财务管理,数据分析。