

# 基于 ARMA 模型对股票 “青岛海尔”成交量的分析预测<sup>\*</sup>

林 杰 朱家明<sup>\*\*</sup> 陈富媛

( 安徽财经大学 )

**【摘 要】**针对股票“青岛海尔”的日成交量,选取 2016 年下半年数据,利用 B-J 法建立 ARMA 模型,借助 Eviews 8 软件的求解,得到与实际数据高度吻合的 ARMA(3,2) 模型,并对未来 4 天日成交量进行预测,然后与 2017 年初的实际数据对比,发现预测结果比较好.结论表明:模型的精确度比较高,适合股票短线也即日成交量的分析与预测.

**【关键词】**青岛海尔;成交量;ARMA 模型;B-J 法;预测;Eviews8

中图分类号:O211.67 文献标识码:A 文章编号:1000-5617(2016)06-0006-05

## 0 引言

2017 年是我国股票改革的第 13 个年头,目前已经形成了与我国经济发展相适应的特色道路.股票投资已经成为我国投资者的主要投资途径.然而,应该分析哪些指标,如何选择合适的股票实现最大收益,仍然是令许多股民困惑的问题.在股票市场众多的指标中,成交量与收益率是最重要的两个统计指标.宋红宇<sup>[1]</sup>通过建立回归方程得出我国股票市场中交易量与价格变化正相关的结论并根据股票成交量给出了组合投资优化方案;李超<sup>[2]</sup>通过建立成交量与收益率的回归模型,得出股票收益率与股票成交量的变化率呈正相关关系的结论.因此,股票成交量可以作为最优投资方案选择的一个重要参考指标.股票成交量的变化是一个时间序列,它是随机过程的一次实现,与存在重复抽样的横截面数据不同.后者可以通过均值、方差等进行描述性统计分析,前者就需要通过特定的方法对其进行分析

研究. ARMA 模型在进行数据分析预测的过程中,能够同时考虑时间序列的相关性和随机波动的干扰性,具有可拓展性、易学易用性、高效性等优点.在金融领域,姜乐<sup>[3]</sup>、冯盼<sup>[4]</sup>结合包钢股份和招商银行数据利用该模型对股票价格进行了分析预测,数据拟合精度高,预测效果良好.

青岛海尔股份有限公司于 1993 年 11 月在上交所上市交易,股票代码 600690. 青岛海尔作为一家早期成立的大型企业,上市时间也较早,筹资方式多样.公司目前经营状况良好,研究青岛海尔较为典型<sup>[5]</sup>. 该文将结合青岛海尔的股票成交量历史数据建立 ARMA 模型,并利用该模型对接下来一周的成交量进行预测.

## 1 ARMA 模型介绍与建模步骤

### 1.1 ARMA 模型介绍

ARMA 模型的全称是自回归滑动平均模型,它是目前最常用的拟合线性平稳时间序列的模型<sup>[6]</sup>. 其可细分为 AR 模型、MA 模型和 ARMA

收稿日期:2016-11-01

<sup>\*</sup> 国家自然科学基金项目(11601001)

<sup>\*\*</sup> 通讯作者: zhujm1973@163.com.

模型三大类。

一般的 ARMA(  $p, q$  ) 模型表示为:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (1)$$

其中  $\{\varepsilon_t\}$  是白噪声序列  $p$  和  $q$  都是非负整数。

对于随机扰动项不同,则回归方程不一样,以下有两种情况:

①如果随机扰动项是一个白噪声( $\mu_t = \varepsilon_t$ ), 则称(1)式为一  $p$  阶自回归过程,记为 AR(  $p$  ):

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t$$

②如果  $\mu_t$  不是一个白噪声,通常认为它是一个  $q$  阶的移动平均过程记为 MA(  $q$  ):

$$\mu_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

## 1.2 建模步骤

ARMA 模型是由美国学者 George Box 和英国统计学家 Gwily Jenkins 共同建立的<sup>[5]</sup>,他们给出了对平稳时间序列数据使用 ARMA 模型建模、估计、检验和预测的方法,该方法也称为 B-J 法。使用该方法的前提是这个时间序列是平稳的,对于给定的时间序列进行 ARMA 模型建模,通常步骤如下:

### (1) 平稳性检验

检验时间序列的平稳性可以先通过时间序列整体图形进行大致判断,平稳序列的时序图应该显示出序列始终围绕一个常数值波动,且波动的范围不大。如果观察序列的时序图显示出该序列有明显的趋势或周期,那它通常不是平稳序列。再运用 ADF 单位根检验方法<sup>[7]</sup>,通过统计量来进一步判定该序列的平稳性。如果时间序列是不平稳的,可以通过低阶差分、季节差分、对数差分等方法,使其变为平稳序列。对原始或者处理后的平稳序列还需要进行白噪声检验,白噪声序列没有可以提取信息的序列,白噪声检验一般采用检验统计量  $Q$  的方法。

(2) ARMA 模型定阶定阶方法主要有相关函数定阶法和 AIC 准则<sup>[7]</sup>。首先,一般通过观察平稳时间序列的偏自相关和自相关系数图的性质(见表1)初步估计  $p, q$  的值,选取合适的 ARMA(  $p, q$  ) 模型进行拟合。

表1 模型定阶表

	AR( $p$ )	MA( $q$ )	ARMA( $p, q$ )
自相关系数	拖尾	$q$ 阶截尾	拖尾
偏自相关系数	$p$ 阶截尾	拖尾	拖尾

由于样本的随机性,相关系数一般不会呈现出理论截尾的情形。为了进一步精确判断  $p, q$  的值,还需要通过 AIC 最小准则进行判断。对模型中的  $p$  和  $q$  两个参数的组合选择,可以根据自相关系数和偏自相关系数的特征进行多次尝试,最终选择 AIC 值最小的模型作为相对最优模型<sup>[8]</sup>。

### (3) 参数估计及检验

进行参数估计时,常选用最小二乘法,借助相关统计软件即可实现。估计参数后,同时可以得到相关参数的统计量等信息,根据其显著性,选择保留或者剔除。

### (4) 模型的有效性检验

模型拟合的效果主要通过检验残差项是否为白噪声序列进行评价。当残差项为白噪声序列时,说明残差项中不包含相关信息,时间序列中的信息被完全提取,此时模型拟合效果显著。

### (5) 模型的预测

通过以上几步建立了拟合效果显著的合理的模型,可以根据原始序列对未来值进行估计预测。预测方法通常选用线性最小均方预测方法,得到原始序列未来取值及趋势<sup>[9]</sup>。

## 2 实证分析及预测

在网易-股票-青岛海尔板块([http://quotes.money.163.com/trade/ljsysj\\_600690.html#01b07](http://quotes.money.163.com/trade/ljsysj_600690.html#01b07))选取了2016年8月1日至12月30日(不包含节假日及双休日),共计110天股票成交量的数据。下面借助 Eviews 8 软件建立 ARMA 模型,对数据进行分析建模。

### 2.1 股票成交量序列平稳化检验

将数据导入 Eviews 8 作出其时序图(如图1所示),先通过时序图观察其平稳性,考虑是否进一步进行 ADF 检验。

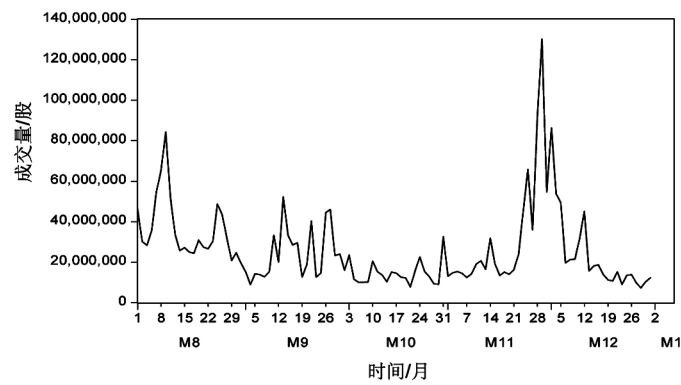


图 1 股票日成交量时序图

由图 1 可以看出,在这段时间内股票的成交量波动较大,是不平稳的. 对其进行 ADF 检验(见表 2),考虑是否对原始数据进行差分平稳化.

表 2 原始序列 ADF 检验图

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.418105	0.0124
Test critical values:		
1% level	-3.491928	
5% level	-2.888411	
10% level	-2.581176	

从表 2 中可以看出,在显著性水平 0.05 下,原始序列拒绝存在一个单位根的原假设,即有 95% 的把握认为该序列是平稳的,可以为后续建模使用.

## 2.2 股票成交量序列模型定阶

作出股票成交量数据的自相关图和偏自相关图(见表 3).

从表 3 中看到,自相关系数滞后 4 阶滞后落入 2 倍标准差范围之内,且没有明显的截尾性;偏自相关系数滞后 1 阶滞后落入 2 倍标准差范围之内,且没有明显的截尾性. 考虑使用 ARMA 模型,此时  $p$  的可能取值为 2 或 3,  $q$  的可能取值为 2 或 4. 因此,可以考虑建立 ARMA(2, 2) 模型、ARMA(2, 3) 模型、ARMA(2, 4) 模型、ARMA(3, 2) 模型、ARMA(3, 3) 模型或 ARMA(3, 4) 模型. 分别求解其 AIC 值(见表 4),根据最小 AIC 值原则,最终选择 ARMA(3, 2) 模型.

表 3 原始序列相关表



























Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.679	0.679	52.045	0.000
		2	0.549	0.164	86.402	0.000
		3	0.466	0.083	111.41	0.000
		4	0.300	-0.151	121.87	0.000
		5	0.180	-0.082	125.66	0.000
		6	0.099	-0.030	126.82	0.000
		7	0.050	0.034	127.12	0.000
		8	0.020	0.025	127.16	0.000
		9	0.031	0.063	127.28	0.000
		10	0.028	-0.006	127.38	0.000
		11	0.004	-0.057	127.38	0.000
		12	-0.009	-0.041	127.39	0.000
		13	-0.047	-0.066	127.68	0.000

表 4 各模型 AIC 值

模型	ARMA(2, 2)	ARMA(2, 3)	ARMA(2, 4)	ARMA(3, 2)	ARMA(3, 3)	ARMA(3, 4)
AIC	35.85823	35.7737	35.7862	35.6640	35.6679	35.6843

## 2.3 股票成交量模型参数求解及检验

通过上文对模型的选择,将 ARMA(3, 2) 模型作为最佳拟合预测模型. 对该模型进行参数求解并进行检验(见表 5).

从表 5 中可以看到,模型中的每项系数  $p$  值均小于 0.001,具有明显的统计意义,且其对应的特征根均大于 1,满足平稳性要求. 因此,该模

型对应的表达式为:

$$X_t = 23938312 + 1.522X_{t-1} - 1.217X_{t-2} + 0.446X_{t-3} + \varepsilon_t - 1.101\varepsilon_{t-1} + 0.978\varepsilon_{t-2}$$

## 2.4 股票成交量模型残差检验

估计参数后,对模型的残差序列进行白噪声

表 5 ARMA(3,2) 参数求解及检验结果

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23938312	4890221.	4.895139	0.0000
AR(1)	1.522002	0.087666	17.36131	0.0000
AR(2)	-1.217008	0.123356	-9.865858	0.0000
AR(3)	0.445712	0.084616	5.267457	0.0000
MA(1)	-1.013844	0.029179	-34.74557	0.0000
MA(2)	0.978227	0.026781	36.52738	0.0000
R-squared	0.590873	Mean dependent var	25927950	
Adjusted R-squared	0.570620	S.D. dependent var	19945979	
S.E. of regression	13070025	Akaike info criterion	35.66398	
Sum squared resid	1.73E+16	Schwarz criterion	35.81386	
Log likelihood	-1902.023	Hannan-Quinn criter.	35.72474	
F-statistic	29.17348	Durbin-Watson stat	2.141720	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.69	.41+.69i	.41-.69i	
Inverted MA Roots	.51+.85i	.51-.85i		

检验(见表 6)。如果残差序列不是白噪声序列,说明原始序列信息没有被完全提取,模型需要改进;如果残差序列的样本自相关系数均落入随机

区间内,说明残差序列是纯随机的,为白噪声序列。

从表 6 中可以看到,每个残差项的概率值均大于 0.05,说明 Q 统计量均小于检验水平为 0.05 的卡方分布临界值。说明此时的残差项为白噪声序列。

2.5 股票成交量模型的预测

由于 ARMA 模型短期预测精确度较高,利用上文中建立的 ARMA(2,3) 模型对青岛海尔股票未来四天(2017 年 1 月 3 日至 1 月 6 日)成交量进行预测<sup>[10]</sup>(见表 7)。

表 6 残差序列相关表



























Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.021	-0.021	0.0472	0.828
		2 -0.078	-0.079	0.7289	0.695
		3 0.123	0.120	2.4144	0.491
		4 -0.025	-0.027	2.4844	0.647
		5 -0.038	-0.020	2.6464	0.754
		6 -0.066	-0.087	3.1475	0.790
		7 -0.032	-0.033	3.2686	0.859
		8 -0.063	-0.072	3.7388	0.880
		9 0.019	0.030	3.7798	0.925
		10 0.035	0.029	3.9231	0.951
		11 -0.015	0.001	3.9488	0.971
		12 0.026	0.013	4.0307	0.983
		13 0.030	0.015	4.1423	0.990

表 7 2017 年 1 月 3 日至 6 日股票成交量预测值与实际值比较(单位:股)

预测期	预测值	真实值	误差比
1 月 3 日	13471711	12842863	0.04897
1 月 4 日	13066676	17610333	0.25801
1 月 5 日	14910101	12801074	0.16475
1 月 6 日	14763134	11529182	0.28050

由表 7 可得,2017 年 1 月 3 日至 6 日股票成交量预测值与实际值的误差比均小于 0.3,说明预测值较为准确,可以对未来短期内的股票成交进行预测<sup>[11]</sup>。进一步验证了该文建立模型的可靠性,说明该模型对未来成交量预测具有实际意义。

3 结束语

该文利用时间序列分析的 Box-Jenkins 思想<sup>[5]</sup>对青岛海尔股票日成交量这一时间序列建立 ARMA 模型进行实证分析和预测,分析研究了股票日成交量的基本变化特征<sup>[12]</sup>。利用 B-J 法,首先,针对原始数列进行平稳性检验得到原始数据基本平稳;其次,运用原始序列的自相关

图和偏自相关图进行模型识别,确定可建立的模型;然后,对两个模型进行参数估计与检验,根据最小 AIC 准则,确定合适的模型为 ARMA(3,2),并根据参数估计结果给出模型的表达式;随后,通过残差检验验证了所设计模型的合理性<sup>[13]</sup>;最后,运用已经确立好的模型,对青岛海尔股票未来四天日成交量进行预测,通过与实际值进行比较,再一次验证了模型的合理性,模型具有较好的预测效果。综上所述,ARMA 模型能够较好地拟合股票日成交量的时间分布问题,借助 EViews 软件能够十分简便地将 ARMA 模型应用于股票日成交量的分析与预测,为股民进行股票投资提供参考,从而降低风险,提高收益<sup>[14]</sup>。

## 参 考 文 献

- [1] 宋红宇. 股票成交量对投资组合优化影响的研究[D]. 广州大学 2010.
- [2] 李超. 成交量与股票收益率关系的实证分析[J]. 新乡学院学报 2016 33(4): 15-19.
- [3] 姜乐. 基于时间序列的股票价格分析研究与应用[D]. 大连理工大学 2015.
- [4] 冯盼, 曹显兵. 基于 ARMA 模型的股价分析与预测的实证研究[J]. 数学的实践与认识 2011 41(22): 84-90.
- [5] 宝音朝古拉, 白静. 上市公司证券发行及资本成本与资本结构研究——以青岛海尔为例[J]. 财经理论研究 2016 (3): 104-112.
- [6] 朱家明, 刘红杉, 朱运良, 程瑶瑶, 于静. 基于收益率的股票相关性计量分析[J]. 菏泽学院学报 2015(02): 10-14.
- [7] 陈瑶. 基于时间序列分析的我国 GDP 预测模型[D]. 苏州科技学院 2015.
- [8] 魏宁. 时间序列分析方法研究及其在陕西省 GDP 预测中的应用[D]. 西北农林科技大学 2010.
- [9] 吕婕, 朱家明, 胡学峰, 等. 基于对数收益率的股票网络结构研究[J]. 时代金融 2016(8): 146-147, 153.
- [10] 徐孝琳, 周小伟, 朱家明. 基于时间序列模型的上海证券交易所融券余额分析与预测[J]. 河北北方学院学报: 自然科学版 2016(1): 53-60.
- [11] 安潇潇. ARMA 相关模型及其应用[D]. 燕山大学 2008.
- [12] 周立敏, 曹喆, 朱家明, 孙文康. 基于关系数和最佳阈值的股票网络构建及应用[J]. 浙江科技学院学报 2015(4): 256-262.
- [13] 席爽, 朱家明. 沪深 300 股票指数期权定价实证研究[J]. 怀化学院学报 2015(7): 35-38.
- [14] 侍冰雪, 魏慧茹, 朱韶东, 朱家明. 基于网络结构的股票相关性研究[J]. 浙江科技学院学报 2015(1): 62-67.

## Analysis and Forecast of Stock "Qingdao Haier" Volume Based on ARMA Model

Lin Jie, Zhu Jiaming, Chen Fuyuan  
(Anhui Finance and Economics University)

**Abstract:** Based on the trading volume on the stock "Qingdao Haier", selected data of the first half in 2016, ARMA model is established by using the B-J method. The solution of Eviews 8 software was highly consistent with the actual data of ARMA (3, 2) model. And carried on the forecast to the future 4 day turnover, the actual data was compared with that at the beginning of 2017. The results show the prediction results are better, the accuracy of the model is relatively high, and it is suitable for the analysis and prediction of the short-term and current stock volume.

**Keywords:** Qing dao Haier; Stook; ARMA mode; B-J method; Forecast; Eviews 8

(责任编辑: 季春阳)