

# 基于多元线性回归的股价分析及预测

王培冬

(河南财经政法大学, 河南 郑州 450046)

**摘要:** 中国是世界上最大的发展中国家, 中国股票市场的股票价格是序列相关的, 这意味着股票的历史信息可以用来预测未来股价。本文以沪深300指数为实例, 以其成交金额、最低价、成交量、次日开盘、开盘价、收盘价、最高价为解释变量, 以该股票的次日收盘价作为被解释变量建立多元线性回归模型, 之后对模型进行异方差性检验和自相关性检验, 最后运用模型对股价进行预测。结果表明, 股票的历史信息可以用来预测未来短时间内的股票价格。

**关键词:** 多元线性回归; 股价分析; 自相关性检验

## 1 现状

沪深300指数是根据科学客观的原理, 挑选出深圳和上海资本市场中最具代表性股票组成样本股, 用以综合反映证券市场最具市场影响力的一批优质大盘企业的整体状况。沪深300指数是能够反映深圳股市状态的优质股票。

自中国进入新时代以来, 人民生活水平日益提高, 个人可支配收入日益增多, 股票成为人们日常理财方式, 股票价格的变动也成了劳动人民关心的问题, 也是国家关心和照顾的重点问题。中国是世界上最大的发展中国家, 股票市场也正在完善发展中, 中国股票市场的股票价格是序列相关的, 这意味着股票的历史信息可以用来预测未来股价。

## 2 多元线性回归模型在股价预测中的应用

在实际生活中, 股票的价格间存在一定的线性关系, 本文将采用回归分析的方法, 以沪深300指数和上证指数的次日开盘价、收盘价格、最高价格、最低价格、成交金额、开盘价格、成交量为自变量建立模型。

### 2.1 多元线性回归模型的建立

本文使用沪深300指数2018年的历史数据, 即以

次日开盘价格、收盘价格、最高价格、最低价格、开盘价格、成交量、成交金额为自变量, 以该股票的次日收盘价为因变量建立多元线性方程并进行求解预测。本文所用数据来源于网易财经股票行情历史交易数据, 选取沪深300指数2018年一整年的数据进行建模分析预测。

### 2.2 多元线性回归模型在沪深300指数中的预测研究

将股票的历史数据输入SPSS软件中, 运用SPSS进行数据处理得到模型 $R^2$ 值为0.989, 调整 $R^2$ 值为0.988, Durbin-Watson值为2.011。其中,  $R$ 表示拟合优度, 是用来衡量模型的拟合程度, 其数值越接近1越好。 $R^2$ 表示决定系数, 其数值用于反映模型能够解释的方差中占因变量方差的百分比, 数值越接近1越好。调整 $R^2$ 是对决定系数 $R^2$ 的校正, 比 $R^2$ 更加严谨, 其数值越接近1越好。其次日开盘价格、收盘价格、最高价格、最低价格、开盘价格、成交量、成交金额变量系数分别为0.978、-0.226、0.188、0.205、-0.153、-8.408E-10、1.297E-10, 常数项为21.768。

设次日收盘价为 $Y$ , 次日开盘价为 $x_1$ , 收盘价为 $x_2$ , 最高价为 $x_3$ , 最低价为 $x_4$ , 开盘价为 $x_5$ , 成交

金, 从而使河北省民营企业在创新能力提升中受到了很大的资金限制。

### 4.4 民营企业与高校及研究机构的创新合作不紧密

就目前来看, 河北省民营企业在创新能力提升中, 并未与高校及研究机构等建立紧密的合作关系, 这也导致高校及研究机构的研究成果及数据难以发挥出较大作用, 而且高校和研究机构在科研过程中也并未以满足民营企业的需求为目的, 这也造成许多的科技成果难以被有效应用到民营企业之中。

## 5 结语

总而言之, 在京津冀协同发展环境下, 河北省民营企业在提升创新能力中既有其优势所在, 也有其不足

之处, 在发展环境上也同样是有利有弊。因此, 对于河北省民营企业来说, 必须要对京津冀协同发展所带来的变化进行科学分析, 以此制定和调整创新战略, 最大限度的消除不利影响, 这样才能使民营企业的核心竞争力得到根本性的提高, 进而确保京津冀协同发展战略的顺利实施。

参考文献:

- [1]高寿华. 基于SWOT模型的浙江民营企业国际化经营对策分析[J]. 中国市场, 2015(34):103-105+107.
- [2]何艳娜. 新常态下河南省民营经济发展SWOT分析[J]. 经济界, 2017(04):49-55.
- [3]耿丽萍. 民营企业助推乡村振兴的SWOT分析及策略研究[J]. 晋城职业技术学院学报, 2019, 12(03):48-51.

量为 $x_6$ ,成交金额为 $x_7$ ,可以得出多元回归方程:  
 $Y=21.768+0.978x_1-0.226x_2+0.188x_3+0.205x_4-0.153x_5-8.408 \times 10^{-10}x_6+1.297 \times 10^{-10}x_7$

其中 $R^2$ 为0.989。

下图2-1是输出的残差对预测值的散点图,残差图上的 $n$ 个点在图中的位置的散布是随机的,没有任何规律可言,可以说明沪深300指数模型满足模型的基本假设,不存在异方差性。

由输出数据可知,沪深300指数模型的 $R^2$ 值为0.989,接近于1,DW值为2.011,接近于2,调整 $R^2$ 值为0.988,也接近于1。DW=2.011,值在2左右。综上,由定义可知,可认为该模型不存在序列的相关性。

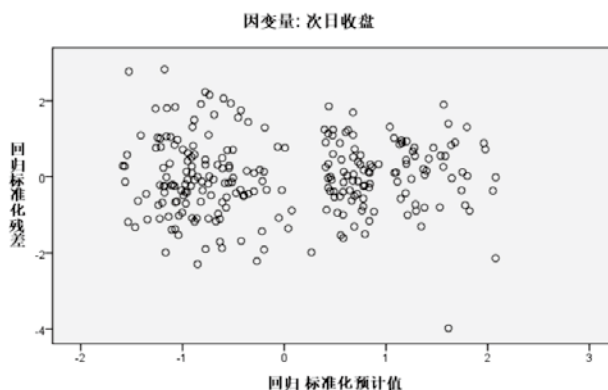


图2-1 沪深300指数残差散点图

### 3 预测分析

表3-1 沪深300预测结果

次日收盘	2969.535	3010.654	2990.506	3002.033	3017.282	3038.198	3029.395	3067.417	3091.126	3128.427
预测值	3018.476	2998.94	3045.005	3007.675	3006.048	3014.715	3055.585	3086.095	3136.594	3137.443
误差率	0.016481	0.00389	0.018224	0.00188	0.00372	0.00773	0.008645	0.006089	0.014709	0.002882

其中误差率=(预测值-实际值)/实际值。此预测结果是将历史数据带入下边方程:

$Y=49.575-0.355x_1+0.143x_2+0.402x_3+0.795x_4+0.718x_5-1.692 \times 10^{-9}x_6+2.064 \times 10^{-10}x_7$ 中所得到的,即上表3-1。计算得出预测误差均值为22.93015545,误差率均值为0.00739866,由此可知,模型的预测误差率在0.003左右,误差数值较实际值来说较小,由此得出模型的预测效果较好的结论。

## 4 建议

### 4.1 对上市公司的建议

对于上市公司来说,合理安排使用公司上市募集的资金,在盈利时及时给股东分红,通过公布财务报表等让投资者了解本公司的现金流等资产的状态都是健康的,可以给投资者信心,在其他公开信息中也可以展现出本公司的优点。

### 4.2 适当对股市进行调控

股票市场的平稳运行,需要国家政府等部门的自主调控,例如18年实施的去杠杆政策,有助于将房地产行业中的泡沫挤出,将在房地产产业内部空转的资金挤出,以便降低杠杆率,降低由于房地产市场波动导致的股票市场的波动,进而调控股市的整体状况。

### 4.3 合理安排外汇储备

2018年内,美联储连续加息,对中国的资本市场影响也特别巨大。美联储加息会使美元升值,相对之下,人民币币值会降低,中国庞大的外汇储备中有

三万亿左右美元,当外部资金流入时,我国强大的外汇储备可以应付因为外汇风险导致的股市崩盘情况,但外汇储备数额巨大,要对其进行合理的安排。18年第四季度,美国又与中国打贸易战,在我看来,贸易摩擦就像是两个经济体的战争,输赢的代价很大,输的一方的经济增长可能会停滞不前甚至是倒退。

### 4.4 对投资者的建议

股票市场是一个十分复杂的金融体系,股票的价格也会受多种状况的影响,对于投资人来说洞察买卖的良机是重中之重,因此投资者可以利用本模型进行预测分析,最大程度地实现股票获益。但对于长期投资的投资者而言需要考虑各方面的因素,如政府政策因素、企业经营状况等。因此投资者投资需要十分谨慎,无论哪个因素的改变,都可能会导致股票市场的价格波动。所以在选择一支优秀的股票的时候,我们可以用此模型进行分析,来确定这只股票的投资必要性,但面对变化莫测的股票市场,仅仅靠这一个技术分析是不能够确保万无一失的,我们应当关注上市公司的各种公开信息,在深思熟虑、权衡利弊之后,做出合理的决策。

参考文献:

- [1]姜富伟,涂俊,David E.Rapach,Jack K.Strauss,周国富.中国股票市场可预测性的实证研究[J].金融研究,2011(09):107-121.
- [2]李潇宁.多元线性回归与时间序列模型在股票预测中的应用[J].科技创业月刊,2019,32(02):153-155.
- [3]周鹏飞,卢泽雨.基于SPSS多元线性回归模型在城市用水量的预测[J].水利科技与经济,2018,24(05):6-10.