

类别	全日制研究生	√
	教育硕士	
	同等学力	

沈阳师范大学

硕士学位论文

题 目：用大数据方法研究股市问题

培 养 单 位：数学与系统科学学院

专 业 名 称：统计学

指 导 教 师：程丛电

研 究 生：马晶鑫

完 成 时 间：2017 年 5 月 19 日

沈阳师范大学研究生处制

学位论文独创性声明

本人所呈交的学位论文是在导师的指导下取得的研究成果。据我所知，除文中已经注明引用的内容外，本论文不包含其他个人已经发表或撰写过的研究成果。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确说明并表示了谢意。

作者签名：_____ 日期：_____

学位论文使用授权声明

本人授权沈阳师范大学研究生处，将本人硕士学位论文的全部或部分内容编入有关数据库进行检索；有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版，允许论文被查阅和借阅；有权可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。保密的学位论文在解密后适用本规定。

作者签名：_____ 日期：_____

用大数据方法研究股市问题

中文摘要

股票市场是一个非常庞大而复杂的系统，其功能主要是对已经发行的股票进行转让、买卖和流通；其状况如何与国民经济的发展息息相关。在中国，伴随着国民经济的迅猛崛起，股票市场也在蓬勃的发展，越来越多的股民怀着支持国家建设与投资理财的热情投身于股市当中。因此，关于股市的研究具有重要的意义。虽说这方面的研究工作已经很多，但大多结果都不够令人满意，其主要不足之处大多在于以下两点：一、受到信息摄取量和储存水平的限制与计算水平的限制，用于支撑研究的数据量较小；二、没有充分地考虑利率与内部收益率的影响。当代计算机与互联网络的发展，使得我们有能力克服这两个不足之处，做出更加理想的科研成果。本文试从扩大数据量与重视利息率和收益率的视域出发建构三项探讨股市奥秘的方法并进行一定的实证分析。

文章共五章。第一章介绍股票的意义与背景和大数据方法的意义与背景，以及相关的研究动态，并说明笔者所要进行的工作与文章的结构。第二章首先简介行为金融学的内容与意义，以及相关研究动态；然后从大数据、机器学习和行为金融学的角度出发，通过设计一个算法定义出一个表现按照某种初级炒股行为买卖一支股票的收益状况的随机变量 R ，并进行有关研究。基于这支股票的一定的折现历史数据运用所设计的算法求出 R 的一个样本；根据该样本做出一个经验分布；再根据该经验分布建立其近似解析分布的一个序，并进而给出一种求优化近似解析分布的方法。基于 R 的生成方式及相关探讨展示它的意义，并揭示出一些值得进一步研究的问题。第三章先从消除利率与内部收益率影响的视角出发，定义一种现值均线，而后建立一种根据长短现值均线预测股价与股指走势的方法，并基于过去多年的历史数据进行实证分析。第四章先简介马尔科夫链的有关基础知识，定义一种拟内部收益率，而后建立一种基于拟内部收益率折现历史数据与马尔科夫链预测股票与股指趋势的方法，并根据一定的历史数据进行实证分析。第五章是结论和展望，总结性的描述所做工作；指不足，并提出有待进一步研究的问题，及笔者对于未来的展望。

所做工作可为炒股提供一定的启示，可为从大数据、机器学习和行为金融学的角度出发进一步研究股市提供一定的启示。

关键词: 大数据, 股票, 利率, 均线, 马尔科夫链

Research on the stock market by big data method

Abstract

The stock market is a very large and complex system. Its function is mainly to transfer, sell and circulate the issued shares, and its condition is closely related to the development of the national economy. In China, with the rapid rise of the national economy, the stock market is booming, more and more investors are investing in the stock market with the enthusiasm of the support of national construction and investment management. Therefore, the research on the stock market is of great significance. Although there have a lot of research work in this area, but most of the results are not satisfactory, the main shortcoming are mostly in the following two points: first, the level of information is limited and calculation of intake and storage level limit for a smaller amount of data supporting research; two, did not fully consider the impact of interest rate and internal rate of return. With the development of modern computer and Internet, we have the ability to overcome these two shortcomings and make better achievements in scientific research. This paper tries to construct three ways to explore the mysteries of the stock market from the view of enlarging the amount of data and paying attention to the interest rate and the rate of return and carrying out some empirical analysis.

This article consists of five chapters. The first chapter introduces the significance and background of the stock and the big data method, dynamic related research, and explains the structure how the work to be carried out. The second chapter introduces the content and significance of behavioral finance, and the related research trends. Then, from the point of view of large data, machine learning and behavioral finance, we design an algorithm to define a kind of performance a random variable R of the yield status of the stock and conducts the relevant study. The discount to the historical data of the stock by using the designed algorithm to derive a sample based on R . According to the samples to make a empirical distribution, the empirical distribution establish the approximation of a sequence distribution, then gives an approximate analytical method for optimizing the distribution and generation of R . To demonstrate its significance based on, and reveals some problems for further study. The third chapter starts from the elimination of the interest rate and the internal rate of return of the perspective of the definition of a moving average value, establish a method to predict the stock price and stock index trend according to the length of the average

value, makes an empirical analysis of historical data in the past years based on the relevant basic knowledge. The fourth chapter first introduces the definition of Markov chain, an internal rate of return, set up an internal rate of return of discounted based on the historical data and the Markov chain prediction of stock and stock index trend, and to analyze the historical data. The fifth chapter is the conclusion and prospect, summarizes the description of the work, points out the shortcomings, and puts forward the problems to be further studied, and the author's future prospects.

The work can provide some inspiration for the stock market, which can provide some inspiration for the further study of the stock market from the perspective of large data, machine learning and behavioral finance.

Key words: big data, stock, interest rate, average, markov chain

目录

中文摘要.....	I
Abstract.....	II
第一章 引言.....	1
1.1 背景与意义.....	1
1.2 研究动态.....	2
1.3 大数据.....	3
1.4 大数据与股市研究.....	4
1.5 本文工作.....	4
1.6 本文结构.....	4
第二章 用随机变量表现炒股行为.....	5
2.1 行为经济学.....	5
2.2 模型建立.....	5
2.3 分布估计.....	6
2.4 展望.....	9
第三章 一种用均线预测走势的方法.....	11
3.1 股票的平均价格.....	11
3.2 均线.....	13
3.3 利用长短均线预测股票走势.....	13
3.4 实证分析.....	13
第四章 运用马尔科夫链预测股票走势.....	16
4.1 马尔科夫链的相关基础知识.....	16
4.2 研究方法.....	17
4.3 实证分析.....	19
结论与展望.....	25
参考文献.....	26
附录.....	28
致谢.....	40
个人简历及攻读硕士学位期间已发表论文.....	41

第一章 引言

1.1 背景与意义

一种由股份公司以融资为目的而发行的，持有者可以此分得一定的股息和红利的有价证券，我们称之为股票。股票市场是一个非常庞大而且复杂的体系，包括股票的买卖、转让以及股票市场的流通等复杂过程，因此社会上各方面因素的错综运转无时无刻不影响着股市的变化。

在我国，现实意义上的股票市场是在改革开放的计划经济体制下开始萌芽的。股票实际上是一种融资的手段，通过发放股票来筹集社会上的闲置的资金，成为企业和集团的进一步发展的主要资金来源。除此以外，任何企业公司在经营和发展的过程中都会存在着一定的经营风险。在经营过程中，公司资金集中所承受的风险就比较大。通过发行股票的方式来构建一个股份公司可以分散投资风险。即使投资失败，股东们也只需要承担也是非常小的部分的损失。中国的股票市场从最初的试点到完善市场管理，再到现在的蓬勃发展经历了非常曲折的历程，特别是最近的几年内，中国的经济快速发展已几乎覆盖全球，因而所有人都在关注着中国股票市场的发展。在亚太地区来说，中国股票市场在上市公司的数目以及总市值方面都有非常大的优势。中国股市发展时间较美国来说，可谓是在极短的时间内飞速发展，能够取得今天的规模和地位实属不易。由此可见，股票市场也反映着一个国家的经济发展状况，维持股市的平稳发展对于现代和谐社会的建立是必不可少的。

1.2 研究动态

现如今，经济的健康发展，投资者更多地选择将资产投入到股市中，期待获得高额的收益。但是股票价格的波动趋势具有非线性、时变性和高度的不稳定性等特点，因此投资者在期待高额回报的同时也要承担一定的风险。因此，要想做到“保值增值”就必须找到一个有效的预测方法，尽量降低风险，减少损失。与此同时，股市的状况又与国民经济的发展息息相关。因此关于股市的研究具有重要的意义，在过去的十几年里从各种不同的角度出发研究股市的文献层出不穷。

Kim 和 Lee^[1]（2004）从优化变换的神经网络的角度研究了股市预测问题；Fonseka 和 Liyanage^[2]（2008）给出了一个运用滞后相关分析对我国现阶段股市数据的数据挖掘算法；李涛和郭杰^[3]（2009）探究了城市居民的风险态度与其投资行为之间存在的关系；王嘉展和刘丽霞^[4]（2014）探讨了随机利率下股票价格服从几何分数布朗运动的幂期权定价问题；郑薇等^[5]（2015）考察了统计分析方

法在股票最优投资方面的实际应用；朱泓嘉^[6]（2015）验证了马尔科夫预测法应用到短期的股票价格波动趋势预测的有效性；张力^[7]（2016）讨论了一种股票市场投资组合策略构造与其相关模型的检验问题；黄子颖^[8]（2016）讨论了将 Markov 链过程理论应用于股票交易市场对股价综合指数的分析预测模型；黄玉洁^[9]等利用多元统计分析的方法对上市公司与股票价格之间的关系进行分析；张莉莉和许英^[10]基于 Markov 链对股票价格趋势加以预测；王越^[11]主要基于马尔科夫和布朗运动建立股票价格预测模型，通过实例分析检验模型。

1.3 大数据

大数据，也叫做巨量资料，指的是当需要处理的数据量的规模大到无法利用现有的计算技术和软件工具分析处理，从而得到有利用价值的信息。大数据具有五个 V 特点^[12]：容量（Volume）大、多样性（Variety）、高速性（Velocity）、精准性（Veracity）、多变性（Variability）。事实上，大数据不仅仅是数据的数量庞大和处理大数据的技术，而是利用计算机和软件工具，在大规模数据的基础上完成一些在小规模数据的基础上是无法实现的事情。换句话说，大数据的出现使得我们急需探究一种新的方式去处理巨量的存储数据和实时数据，通过整理数据从中提取到隐藏在数字中的有价值的信息，进而调整企业和公司的发展方向，改善股票市场环境，最终形成变革之力。

现代大数据主要有以下几个方面^[13]的特征：

（1）大数据十分碎片化

伴随着现在社会的迅猛发展，我们的消费内容实在是五花八门，种类繁多。我们在日常生活中产生数据的能力远远大于我们的处理数据的能力。迄今为止，已产生巨量数据，数据信息十分碎片化。

（2）大数据是小数据的拓展

多年以来，国内的大型企业、公司一直在寻求能够解决现有问题的技术支持，投入巨额费用在数据技术分析的研發，实施以及应用上。小数据实际上是进一步开拓大数据的基础。小数据行使着重要的描述性职责：企业利用现有的数据可以判断公司内部各项业务的运转情况，可以了解现状，也能够帮助企业发现最新趋势。

（3）大数据既非模糊也非确切

就一定意义而言，大数据不能改变任何事物。虽然大数据可以提高我们判断事物变化趋势的能力，但是不能赋予我们持续的、准确的预测未来的能力。在大数据时代，大数据处于不断变动的状态，尽管真实存在，却也难以准确预测。

事实上，大数据并不是遥不可及的一个单调的词汇，而是一个可以以任何形

式你我身边真实存在。虽然你感觉不到或者是忽视了它的存在，但它却时刻在影响着我们的生活，大数据与我们的日常生活密不可分，例如：我们每天都要关注的微博头条、马云旗下的支付宝蚂蚁借贷、QQ 微信以及百度地图等都是以大数据为基础的，它们的具体意义如下。

（1）微博头条

微博是最早的一款社交软件，微博中是可以进行搜索的，在网友的大量、重复性的搜索下出现的热词，就是现在的微博头条。大数据的应用主要体现在：通过搜索对海量的搜索词汇进行统计处理、汇总，按热度排出网民现在关注的焦点。

（2）支付宝蚂蚁借贷

支付宝是目前国内最火爆的软件之一，在支付宝的软件交易平台上，每天都会存储巨量的数据和交易信息，阿里公司对客户的交易数据进行分析，判断客户是否存在潜在的资金问题，并安排平台与客户进行沟通。由此可见，利用专业的计算技术对大量的实时数据进行分析处理，可以让企业公司的经营情况透明化。

（3）QQ 微信上的大数据

随着 QQ 微信的开发及推广，多数人都选择使用 QQ 微信来交流，在朋友圈中分享图片文字来记录日常的心情和生活。QQ 微信还添加了语音视频聊天等新功能，除此以外，还可以提供充值、预约、网络购票、还款以及生活缴费等服务。QQ 微信的出现，为人们生活提供了很多便利，全国上下有数亿的人每天都在使用着 QQ 微信，这一普通的行为却让 QQ 微信的后台每天都产生巨量的数据信息，而我们利用着大数据完成了网络互动。

（4）百度地图

作为中国最大的搜索引擎，百度提供的一项服务---百度地图，使得我们能够通过使用软件在地图上查看自己现在所处的位置，以及周边的建筑及周边服务。这款方便实用的软件上的地图、界面都是由大数据堆积出来的。

1.4 大数据与股市研究

随着市场经济的迅猛发展，科技的不断创新，互联网时代已经到来。互联网为人们的生活带来便捷，人们已经习惯于通过网络获取各类信息，网络支付购物，信息搜索等活动。互联网的发展和传播，使得人们能够更容易的获取到市场信息，互联网已经成为大众获取信息的主要途径，也正以如此才产生了巨量数据信息，使我们迎来了大数据时代。正因为股市的发展与国民经济的发展息息相关，所以关于股市的研究具有重要的意义，在庞大而复杂的大数据环境下，股票市场的发展机遇与挑战并存。

随着云时代的到来，大数据也受到了广泛的关注，学者们开始从数据挖掘的

角度研究股市预测，例如，唐滔^[14]（2013）研究了云计算在股票数据分析领域的应用与实现问题；王春峰^[15]等（2014）利用开发的金融数据系统对媒体信息进行数据挖掘，并探讨了媒体信息对股票收益的影响，同时预测网络信息将成为影响股票市场的重要风险因素。席文帅^[16]等探讨了大数据的时代背景下应如何正确的进行股票投资。

1.5 本文工作

受到以上几个方面研究动态的启发，本文尝试运用大数据的方法与注重利率和内部收益率的视域出发，研究几个股市的问题。

一、以行为金融学为出发点，建立一个能够在一定程度上表现某种初级炒股行为收益状况的随机变量模型，并根据一定的实际参考数据生成样本以及估计分布，最大程度的模拟出一种炒股人的炒股行为，旨在能够帮助投资人获取更大利益；

二、在掌握均线的基础知识后，参考大数据的特点并考虑利率变化带来的影响，利用折线均线从而构建一种长短均线模型，再利用计算机软件对某支股票的数据进行整理、分析、估计，尝试预测出股票的走势情况；

三、结合马尔科夫链的相关知识，考虑利率与内部收益率的影响，在此基础上，利用两支股票数据进行实例分析，暨给出一些值得探讨的问题。

1.6 本文结构

本文的内容安排如下：

第一章为引言部分，首先介绍了股票市场的背景和对我国经济发展的重要意义，以及简介大数据研究的背景，进而说明了本文的研究意义，阐述了关于股市研究的动态以及本文所要进行的工作。

第二章以行为经济学为基础，利用实际股票数据构建随机变量模型，给出一列随机变量的生成方式算法及其估计分布，其次实际验证模型的有效性和实用性，在此基础上叙述了研究工作中的缺陷，并指出可进一步讨论的问题。

第三章叙述了股票的平均价格以及修正平均价格的涵义，尝试利用均线预测股价走势，建立一种关于当前股票平均价格的判别模型并验证。

第四章通过对马尔科夫链的相关知识的学习中，探究马尔科夫链预测的方法并选取两支实际股票数据进行比较分析，以此检验马尔科夫链预测法对股市研究问题的实用性。

第五章是结论和展望，全面总结本文的工作并指出有待进一步研究的问题。

第二章 用随机变量表现炒股行为

2.1 行为经济学

经济学由理论与经验事实构成，是研究人们的决策如何相互影响的。理论可以帮助我们理解真实世界，并对真实世界做出正确的预测；经验事实则可能支持理论，也可能颠覆理论。正因为如此，“以更贴近经验事实为目的，对具有现实心理特征的当事人如何参与配置稀缺资源的研究”称为行为经济学。

行为经济学是在心理学的基础上提出的行为假设，它将行为分析理论与实际的经济规律结合起来。传统的金融学理论，不重视股民的投资情绪，认为投资是理性的行为。而实际上，金融学者们发现，股民的“跟风”思维、心理定势以及环境影响对投资行为有很大的影响。行为经济学的核心思想是，对经济的研究是建立在现实的心理特征的基础上，现实中投资者的决策偏好是内生的而非外生的，投资者的决策模式和行为偏好可以通过市场反应出来。行为经济学更贴近实际经济，对经济学领域的发展有重要意义。

在新兴的边缘科学——行为金融学的影响下，近几年来越来越多的学者和专家开始从行为金融学的角度来探究股市的相关问题，例如，李静^[17]（2012）运用行为金融学的知识对股票市场投资者的行为进行研究；杜飞^[18]（2014）从行为金融学的角度出发考查了股票市场动量效应和反转效应的形成机制；李涛和张文韬^[19]（2015）考虑了人格特征与股票投资的相互关系。黄润鹏和左文明^[20]等以行为金融理论为基础，验证了微博情绪信息能够影响并预测股票市场的假设；蒋翠清，梁坤^[21]等研究利用社会媒体的信息资源对股票行为预测构建回归预测模型，并已实际验证了此方法的有效性和实用性。

本章主要思想是利用一段时间内的股票数据从统计的角度出发，利用概率论与数理统计和金融方面的知识，以行为金融学为基础对炒股人的炒股行为以及炒股人的心理进行分析探究，以购买的某支股票历史数据为依据，模拟生成一个表现某种初级炒股行为的随机变量的方法，并拟合其对应的分布估计，以此来预测股票的走势，有利于投资者获取更大的收益。

2.2 模型建立

在现实生活中，很多刚刚步入股市的新人选择这样炒股：瞄准一支股票（或板块），观察一段时间后估算出其近期的平均价格，然后见到某刻股票的价格在一定程度低于估算的平均价格时选择买入，若见到其股票的价格在一定程度高于估算的平均价格时即卖出，若是过一定的时间段后仍然没有出现卖点的话，便不管此刻股票价格的高低即时抛售。我们称此行为做初级炒股行为(Junior actions of

speculation on stocks), 并简记为 JASS。为了探究按照 JASS 的方式炒某支股票 X 的收益状况及其与股市其它方面所存在的关系, 兹我们通过如下算法 1 与其输出 R 来抽象地刻画该炒股行为及其收入状况。

算法 1

输入: 正整数 l 和 s , 正实数 ε 和 δ ; 股票 X 的股价:

$$P_i^{(o)}, P_i^{(c)}, P_i^{(h)}, P_i^{(l)}, i = -l, -l+1, 0, 1, \dots, s.$$

输出: 收益值 R 。

1. 令 $k=0, R=0$ 。 $m = \frac{1}{l} \left[\sum_{i=-l}^{-1} \frac{1}{4} (P_i^{(o)} + P_i^{(c)} + P_i^{(h)} + P_i^{(l)}) \right]$ 。
2. 若 $k < s$ 且 $P_k^{(o)} > m - \varepsilon$, 令 $k := k + 1$,
 $m = \frac{1}{l} \left[\sum_{i=-l+k}^{k-1} \frac{1}{4} (P_i^{(o)} + P_i^{(c)} + P_i^{(h)} + P_i^{(l)}) \right]$, 返回第 2 步; 否则进行下一步。
3. 若 $k < s$, 令 $k := k + 1$, 然后进行下一步; 否则转 6。
4. 若 $k < s$ 且 $P_k^{(o)} < m + \delta$, 令 $k := k + 1$, 返回第 4 步; 否则进行下一步。
5. (i) $k < s$, 令 $R = P_k^{(h)} - m + \varepsilon$, 然后进行下一步;
(ii) $k = s$, 令 $R = P_k^{(o)} - m + \varepsilon$, 进行下一步。
6. 输出 R , 然后停机。

注 1: 该算法中, l 表示观察过的一段时间, s 表示实际炒股时间; ε 和 δ 分别用以确定买入点和卖出点; $P_i^{(o)}, P_i^{(c)}, P_i^{(h)}, P_i^{(l)}, i = -l, -l+1, 0, 1, \dots, s$, 分别表示过去第 i 日股票 X 的开盘价, 收盘价, 最高价和最低价; m 表示过去 l 日的平均价格; R 表示按照 JASS 方式炒做股票 X 时购买一支股票的收益, 称为收益随机变量。

由上可知, 算法 1 和输出 R 抽象地刻画出 JASS 的炒股方式与其收入的关系。所以, 我们可以借此来研究 JASS 的炒股方式与其带来的收益, 以及从某一特定的角度对股市分析探究, 在一定程度上揭示出股市这个巨复系统的奥秘。

2.3 分布估计

定义收益随机变量 R 后, 自然要考虑 R 的分布状况。本节专门讨论这一问题。折现:

有效地估计 R 的分布状况, 需要较大的样本, 由于利率关系, 当观察期较长时前后的数据具有某种较大的“异质性”, 即同是 1 个单位的货币前后的实际价值差异较大; 为了消除利率的这种不良影响, 我们应首先统一“单位”, 即进行“折现”。如下的算法 2 可用于完成该项任务。

算法 2

输入: $T (\geq 1)$ (年数), i_t (表示第 t 年的利率), $p_{it}^{(o)}$ (第 t 年第 i 日的开盘价), $p_{it}^{(c)}$ (第 t 年第 i 日的收盘价), $p_{it}^{(h)}$ (第 t 年第 i 日的最高价), $p_{it}^{(l)}$ (第 t 年第 i 日的最低价), $i = 1, 2, \dots, 365; t = 1, 2, \dots, T$.

输出: $\bar{P}_{it}^{(o)}, \bar{P}_{it}^{(c)}, \bar{P}_{it}^{(h)}, \bar{P}_{it}^{(l)}$ (表示不同时点各价格在第一年开始时的现值)。

1. 令 $t := 0, I = 1$.

2. $t := t + 1, I := I \times (1 + i_t)$.

(1) $i := 0$;

(2) $i := i + 1$,

$$P_{it}^{(o)} = \frac{p_{it}^{(o)}}{\left[I / (1 + i_t) \right]^{\frac{365-i}{365}}}, \quad P_{it}^{(c)} = \frac{p_{it}^{(c)}}{\left[I / (1 + i_t) \right]^{\frac{365-i}{365}}},$$

$$P_{it}^{(h)} = \frac{p_{it}^{(h)}}{\left[I / (1 + i_t) \right]^{\frac{365-i}{365}}}, \quad P_{it}^{(l)} = \frac{p_{it}^{(l)}}{\left[I / (1 + i_t) \right]^{\frac{365-i}{365}}},$$

$i < 365$, 返回第 2 步; 否则进行第 3 步。

(3) $t < T$, 转第 2 步; 否则进行第 3 步。

3. 输出 $\bar{P}_{it}^{(o)}, \bar{P}_{it}^{(c)}, \bar{P}_{it}^{(h)}, \bar{P}_{it}^{(l)}$. 停机。

运用算法 2 可将过去 T 年中每日的股价折成初始时刻的现值, 从理论上讲, 这就在一定程度上消除了利率的不良影响。

已知过去 T 年中每日的股价: $p_{it}^{(o)}, p_{it}^{(c)}, p_{it}^{(h)}, p_{it}^{(l)}$, 用算法 2 可得到它们在初始时刻的现值 $\bar{P}_{it}^{(o)}, \bar{P}_{it}^{(c)}, \bar{P}_{it}^{(h)}, \bar{P}_{it}^{(l)}$, 再通过令 $j = 365(t-1) + i$,

$$\bar{P}_{it}^{(o)} = \bar{P}_j^{(o)}, \bar{P}_{it}^{(c)} = \bar{P}_j^{(c)}, \bar{P}_{it}^{(h)} = \bar{P}_j^{(h)}, \bar{P}_{it}^{(l)} = \bar{P}_j^{(l)}$$

将它们重新排列为: $\bar{P}_j^{(o)}, \bar{P}_j^{(c)}, \bar{P}_j^{(h)}, \bar{P}_j^{(l)}, j = 1, 2, \dots, 365T$. 以每一 $j (1 \leq j \leq 365T - s)$ 作为算法 1 中的当前时刻, 运用算法 1 便可得到 R 的容量为 $(365T - s - l)$ 的一个样本 $\{R_j\}$. 根据这个样本就可以估计 R 的分布。

以沪市的大连热电举例如下, 选取该支股票从 2011 年初至 2015 年末每日的开盘价, 收盘价, 最高价和最低价作为参考数据, 并取 $l = 30, s = 60, \varepsilon = 0.2, \delta = 0.1$, 先用算法 2 进行折现处理, 再按上述方式重新排序, 然后用下面的算法 3 便可获得 R 的一个容量为 $(365T - 90)$ 的一个样本 $\{R_j\}$ (由于存在周六、周日, 及其它一些缺失值, 我们按每年 244 天进行统计与计算, 我们所得到的样本的容量实际上为

$(244 \times 5 - 90) = 1130$)。

算法 3

输入：正整数 T, l 和 s ，正实数 ε 和 δ ；股票 X 的股价：

$$\bar{P}_j^{(o)}, \bar{P}_j^{(c)}, \bar{P}_j^{(h)}, \bar{P}_j^{(l)}, j=1, 2, \dots, 365T.$$

输出：收益值 R 的样本 $\{R_j\}$ ， $l < j < 365T - s$ 。

1. 令 $j := l$ 。
2. 置 $j := j + 1$ 。若 $j \geq 365T - s$ ，转 3，否则置 $i := -l$ ，继续往下进行。
 - (1) $P_j^{(o)} := \bar{P}_{j+i}^{(o)}, P_j^{(c)} := \bar{P}_{j+i}^{(c)}, P_j^{(h)} := \bar{P}_{j+i}^{(h)}, P_j^{(l)} := \bar{P}_{j+i}^{(l)}$ 。
 - (2) $i := i + 1$ ，返回 (1)，否则继续往下进行。
 - (3) 调用算法 1。
 - (4) $R_j = R$ 。
3. 输出 $\{R_j\}$ ，然后停机。

注 2: (i) 对于周日、周六未开盘没有股票数据，其它停盘存在的缺失值，我们用他们两端值的均值进行代替。

(ii) 我们的计算是通过在一台 Intel(R) Pentium(R) Dual E2160 @ 1.80GHz 1.80 GHz, 0.99GB RAM 的个人计算机上运行算法 3 的 Matlab 语言程序而完成的。

表 (1) 处理数据分组结果

数据分组	[-6, -5) [-5, -4) [-4, -3) [-3, -2) [-2, -1) [-1, 0) [0, 0] (0, 1]							
频数	26	2	2	2	2	128	995	3
频率	0.0224	0.0018	0.0018	0.0018	0.0018	0.11	0.8578	0.0026
最值	最大值 0.0011		最小值 -5.9087					

由 $\{R_j\}$ 可以得到关于 R 的如表 (1) 所示不完全数据。

根据该组数据又可做出 R 经验密度曲线，见下面的图 (1)。

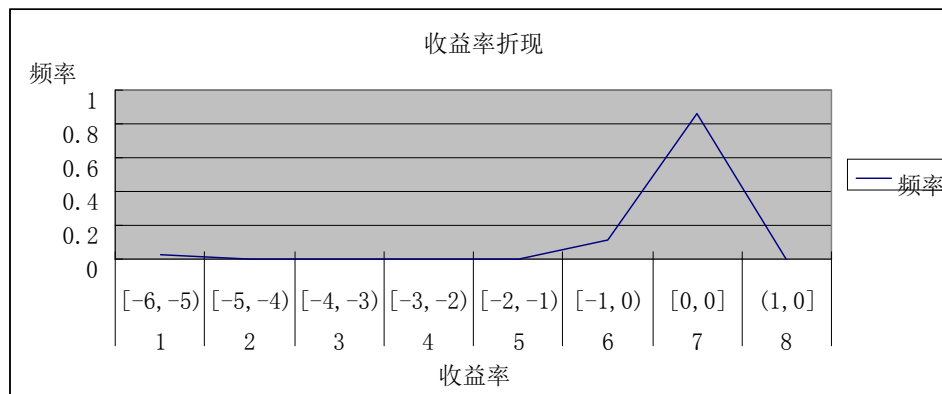


图 (1) 中 $f(x)$ 为关于沪市大连热电的 R 的经验密度曲线

由于上面曲线是选取大连热电股票近 5 年内的数据进行分析得到，根据曲线走势可以发现收益率为 0 的情况高达 85% 以上，收益率为负值的情况为 14.34%。由此可见，对于初级炒股行为，简单的通过对股价的开盘价、收盘价、最高价、最低价进行运算买入，是具有极大风险性的行为，初级炒股者想要获得高额收益几率甚小。本文数据结论，比较符合现实股市行情与股市规律。

虽然 R 的经验分布简单易求且直观，但是不利于我们的理论研究，为了便于进行相关的理论研究，我们可根据样本 $\{R_j\}$ 与所求 R 的经验分布，利用某种分布估计方法，进一步来探索 R 的具有较好拟合度的理论分布。例如我们可按下面 (A)、(B)、(C) 三步获取 R 的近似分布。

(A) 按文献[22]、[23]的方法求出 R 的近似 g-h 分布；或先根据以往的经验及有关思考，拟设 R 服从某类参数分布 $f(x, \theta_1, \theta_2, \dots, \theta_k)$ ，然后用最大似然估计法求出相应最大似然估计；或用其它某种方法求出 R 的某近似解析分布。

(B) 设 g, h 是 R 的两个近似解析分布的密度， f 为图 (1) 中经验密度，若

$$\int_{-\infty}^{+\infty} |g(x) - f(x)| dx \leq \int_{-\infty}^{+\infty} |h(x) - f(x)| dx, \quad (1)$$

认为 g 优于 h 。运用这种方法比较 R 的近似分布。

(C) 运用 (A) 多求出几个 R 的近似分布， f_1, f_2, \dots, f_k ，然后通过 (B) 求出最优的 f^* 。

运用 (A)、(B) 和 (C) 所展示的方法，根据样本 $\{R_j\}$ ，可以求出 R 的优化近似解析分布；为了突出主题与节省篇幅，细节和具体计算在此从略。

注 3: (1) 在实践中，我们可根据实际情况用适当的 a 与 b 分别替换式 (1) 中的 $-\infty$ 与 $+\infty$ 。(2) 因为 f 是分段函数，

$$\begin{aligned} \int_a^b |g(x) - f(x)| dx &= \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |g(x) - f(x)| dx, \quad x_0 = a, x_n = b; \\ \int_a^b |h(x) - f(x)| dx &= \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |h(x) - f(x)| dx, \end{aligned}$$

由此我们可通过计算机进行相关计算。

2.4 展望

本章主要阐述了一种初级炒股行为及买卖股票的收益状况，基于上述已完成的工作可进一步研究的问题。

求出收益的随机变量 R 以及对应的样本 $\{R_j\}$ 与经验分布之后，我们可以进一步讨论如下问题：

(1) 探索 R 的解析分布。

- (2) 计算出 R 的数字特征，如均值，方差，峰度和偏度等。
- (3) 讨论 R 的半方差，VaR 和 TvaR 等风险度量。
- (4) 考虑算法 1 中参数 l 和 s ， ε 和 δ 的最优值。
- (5) 研究同板块的不同支股票的收益变量 R 之间的关系。
- (6) 研究 R 与其它技术指标，如沪市的综合指数，K 线等的关系。
- (7) 通过对于 R 的研究，以此来预测股市的走势。
- (8) 运用机器学习的相关知识，研究开发出能够实时跟踪股市的行情，自动生成 R 的当前值，并经过一定的运算，为股民提供有效的参考建议的应用软件。
- (9) 在实际经济市场的大环境下，从各个企业公司运营角度进行讨论，考虑股票分红以及其他融资行为对股市带来的影响，给股民提出合理化的建议。

第三章 一种用均线预测走势的方法

3.1 股票的平均价格

股票的平均价格^[24]是一种用来反映股票的价格波动的标准。人们通过对不同时点、不同股票间的平均价格进行比较，可以分析预测出股票价格的变动趋势。现在我国的股票价格的平均值主要是来自证券交易所、金融类公司及银行等的统计与整理。从宏观上讲，股票的平均价分为横向与纵向两类。其横向，主要有股价算术平均数、修正股价平均数以及加权股价平均数三类；其纵向，主要是各种日期的均值。

(1) 股价算术平均数^[24]

股价算术平均数 \bar{x} 是利用抽样样本中的单支股票每日的价格（一般用收盘价）计算求和再除以其抽样样本的个数而得出来的，即：

$$\bar{x} = (P_1 + P_2 + P_3 + \cdots + P_n) / n,$$

其中 P_i 为样本 i 当前的股价， n 为样本容量。即通过计算每日股价的算术平均数的方法，是世界上首次用来衡量股票价格的指标，也称为道·琼斯股价平均值。算术平均数的优势在于计算比较简便，而其不足之处是：一旦抽样样本中的股票出现对该股票进行赢利分红、企业融资等情况时，股价的算术平均数会中断，此时不再具有连续性，无法再利用时间序列平均值数据进行对比分析，因为此时的股票价格变化与算术平均数序列所表现出的结果是不一致。

(2) 修正股份平均数^[25]

修正股份平均数是通过除数修正法（道式修正法）而计算出来的，这是由美国的学者道·琼斯在 1928 年提出的一种计算股价平均数的方法，见文献[26]。计算方法为：

新的股数 = 折股后的股票总价格 / 折股前的平均股价，

修正的股价平均数 = 折股前的股价总额 / 新的股数。

例如：选取某股票样本见下表 2

表 2 抽样样本股票价格表

抽样样本股票	A 股票	B 股票	C 股票	D 股票	E 股票
股票价格	10 元/股	15 元/股	20 元/股	25 元/股	30 元/股

可求出此时的股票市场的股价的算术平均数为：

$$\bar{x} = (10 + 15 + 20 + 25 + 30) / 5 = 20;$$

由于受到一些因素的影响，这 5 支股票价格在某时点出现涨跌幅，见下表 3：

表 3 另一时点抽样样本股票价格表

抽样样本股票	A 股票	B 股票	C 股票	D 股票	E 股票
股票价格	12 元/股	18 元/股	25 元/股	30 元/股	35 元/股

求得此刻股票市场的股价的修正算术平均数为：

$$\text{新的股数} = (12 + 18 + 25 + 30 + 35) / 20 = 6,$$

将新的股数代入下列的算式中，可得：

$$\text{修正的股价平均数} = (12 + 18 + 25 + 30 + 35) / 6 = 20.$$

修正股份平均数的意义在于：一旦某支股票出现分红、融资、发行新股等情况时，可通过修正新的股数，使得股价的算术平均数不会发生改变，能够保持算术平均数的连续性和可比性，有助于对股票趋势的研究。

除此之外，加权股价平均数^[27]也是一种股票价值的计算方法。加权股价平均数是根据抽样样本的各支股票间的相对重要性进行加权，然后再计算平均数的一种股价平均数的算法。其权数可以是股票总市值、成交股数以及股票发行量等。

(3) 日股价平均数

日股价平均数可根据计算股价平均值的时间差分为：5 日平均数，周平均数，月平均数等。一般地， n 日平均数记为：

$$P = \frac{p_1 + p_2 + \cdots + p_n}{n},$$

这里， p_1, p_2, \cdots, p_n 为样本过去 n 日每日的平均数。

3.2 均线

设 $\{p_i | i=1, 2, \dots\}$ 为抽样样本的股价平均数序列, P_i 为 n 日的平均数,

$$i = n, n+1, \dots,$$

序列 $\{P_i | i=1, 2, \dots\}$ 的图像称为 n 日均线。当样本容量为 1, 即样本仅一支股票时, n 日均线称为该股票的 n 日均线。

由于选取的股票的历史数据量比较大, 因此考虑利率 α 对股票价格的影响因素, 给出折现均线公式如下:

$$P_i = p_i (1 + \alpha)^{\frac{l}{244T}}, \quad l=1, 2, \dots, n, n=244T.$$

下面文中提到的所有均线皆为折线均线。

3.3 利用长短均线预测股票走势

分别用 P_i 和 P'_i , $i=1, 2, \dots$, 表示某支股票 l (long) 日与 s (short) 日均线, $l > s$ 。已知该股票过去 T 年的每日价格 (某种均价, 或开盘价等) $p_i, i=1, 2, \dots, 365T$, 令

$$D = \left| \left\{ i \mid P_i > P'_i > p_i, p_i > P_{i+l} \right\} \right|,$$

$$I = \left| \left\{ i \mid P_i < P'_i < p_i, p_i < P_{i+l} \right\} \right|,$$

这里 $|A|$ 表示集合 A 中元素的个数。当 I 充分大于 D 时, 我们认为往后满足 $P_i > P'_i > p$ 的时点为买入的好时机, 即往后 l 日内长的可能性较大。

3.4 实证分析

(1) 选取中国石油 (601857) 从 2010 年 1 月 4 日至 2015 年 12 月 31 日的股票的历史数据为参考资料 (参考数据见附表 1), 分别用 P_i 和 $P'_i, i=1, 2, \dots$, 表示该支股票 30 日与 5 日均线。

首先参考这支股票股价过去 6 年内的每日股价 (本文中的股价只看收盘价), 通过计算机对该数据进行计算处理, 可求出股价的 5 日平均数和 30 日平均数 (详

见附表 1)，由以上处理数据可以画出如下折线图 2：

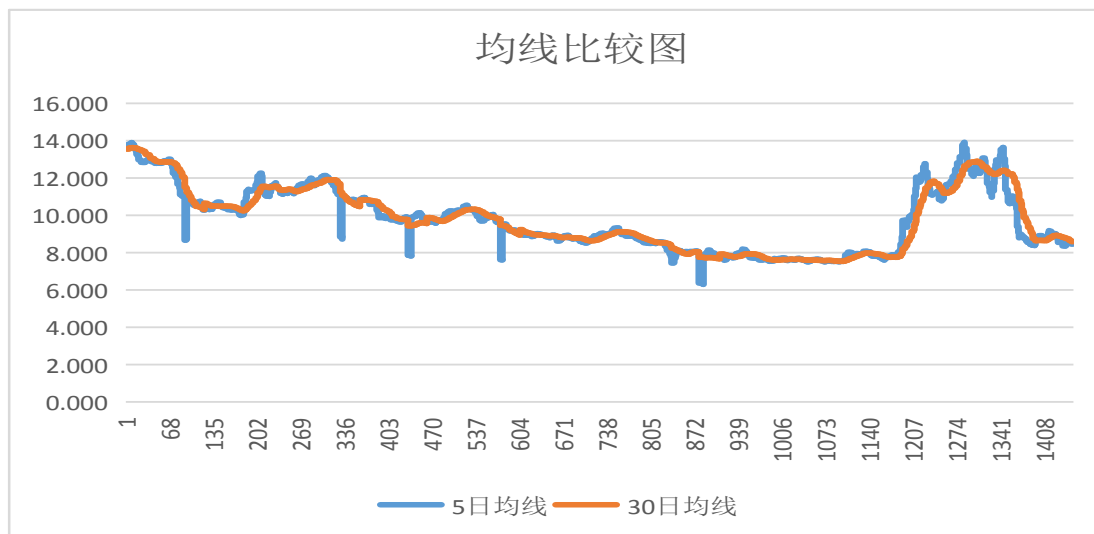


图 2 股票数据均线比较图

从上图可以看出，5 日均线 and 30 日均线的走势大致相同，都可以拟合代表这支股票近期的整体走势。但是，我们发现 5 日均线和 30 日均线数据两者在某些时点存在较大的差异，这是由于股票价格的涨幅趋势受股票市场的各类综合因素的影响。

再从数据差异的角度出发，利用建立的模型对抽样样本进行数据量的统计：

$$D = \left| \left\{ i \mid P_i > P'_i > p_i, p_i > P_{i+l} \right\} \right|,$$

$$I = \left| \left\{ i \mid P_i < P'_i < p_i, p_i < P_{i+l} \right\} \right|,$$

得到这组股票数据的统计结果：D=691，I=434。

分析结果：

(1) 通过对这支股票的处理，我们发现 $D > I$ ，说明这支股票在该时间段内，出现降低趋势的可能性高于出现增长趋势，总体来说这支股票近期呈现降低趋势，与图表走势保持一致。

(2) 通过对单支股票的短时均线和长时均线数据量的对比统计：当 I 充分小于 D 时，我们认为往后满足 $P_i < P'_i < p_i$ 的时点时，表示近期这支股票总体处于增势，此时为抛售的好时机，即往后 l 日内跌的可能性较大；当 I 充分大于 D 时，我们认为往后满足 $P_i > P'_i > p_i$ 的时点时，表示近期这支股票总体处于减势，此时为买入的好时机，即往后 l 日内涨的可能性较大。

(3) 可利用该模型进一步来研究比较某板块内的股票的短时均线和长时均线数据量的多少，或者进行板块间的比较，可以作为比较股票走势的一个标准。

(4) 本文中，这一模型的提出仅限于理论研究。后续研究工作，可在此基础上，考虑其他因素的影响，继续修正该模型。

第四章 运用马尔科夫链预测股票走势

众所周知，马尔科夫链在预测股票走势领域具有重要的应用，不少文献对此都有体现，例如：伍海华的专著^[28]中给出了一种运用马尔科夫链预测股票走势的方法，并进行了一定的实证研究。以往在这方面的工作主要不足之处，有以下两点：其一，没有完全注重利率与内部收益率的影响；其二，受到计算技术的限制，数据量相对不足，如文章^[6]只用了 24 天的股票数据，而专著^[28]中仅用了三年数据。本章主要从改进以上两点不足的角度出发，建立一种基于大数据的运用马尔科夫链预测股票走势的方法。

4.1 马尔科夫链的相关基础知识

1907 年前后，俄国数学家马尔科夫（A.A.Markov, 1856—1922）对一系列有着特定相依性的随机变量进行研究，后人将此列随机变量命名为马尔科夫链（Markov Chain），马尔科夫预测法由此而来。马尔科夫预测法主要是通过统计学中的应用概率统计知识，利用马尔科夫链的理论和性质对一系列随机变量的模型进行研究并分析预测所处状态。

马尔科夫预测法的核心思想是：通过构建马尔科夫链，通过目前的已知条件来确定初始状态的概率向量以及对应的状态的转移矩阵，从而预测模型在未来某一时刻的状态。马尔科夫链的定义及主要性质如下：

定义 1^[29]：设有随机过程 $X_T = \{X_t, t \in T = (0, 1, 2, \dots)\}$ ，其状态空间为 $I = \{0, 1, 2, \dots\}$ ，若对任意正整数 k ，任意 $t_i \in T$ ， $t_i < t_{i+1}$ ， $i = 0, 1, 2, \dots, k+1$ 及任意非负整数 i_0, i_1, \dots, i_{k+1} ，有

$$P\{X_{t_{k+1}} = i_{k+1} | X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_k} = i_k\} = P\{X_{t_{k+1}} = i_{k+1} | X_{t_k} = i_k\},$$

则称 X_T 为 Markov 链或马氏链。

注：如果把时刻 t_k 看成现在，那么 t_{k+1} 是将来时刻， t_0, t_1, \dots, t_{k-1} 则是过去时刻。

Markov 链即在确定现在时刻的状态下，将来的状态与过去无关。

定义 2^[29]：我们称状态的变化为转移，所以条件概率 $P\{X_{n+k} = j | X_n = i\}$ 称为在时刻 n 系统从状态 i 经过 k 步转移到状态 j 的 k 步转移概率，记为 $p_{ij}^{(k)}(n) = P\{X_{n+k} = j | X_n = i\}$ ， $i, j \in I$

注 3：（1）转移概率 $p_{ij}^{(k)}(n)$ 不仅与状态 i, j 有关而且与时刻 n 有关。

（2）当 $k=1$ 时，把 $p_{ij}^{(1)}$ 记为 p_{ij} ，称 p_{ij} 为 Markov 链的一步转移概率。

（3）若用 $P^{(k)}$ 表示 Markov 链的 k 步转移概率所组成的矩阵，则

$$P^{(k)} = (p_{ij}^{(k)}) = \begin{bmatrix} p_{00}^{(k)} & p_{01}^{(k)} & p_{02}^{(k)} & \cdots \\ p_{10}^{(k)} & p_{11}^{(k)} & p_{12}^{(k)} & \cdots \\ \vdots & \vdots & \vdots & \\ p_{i0}^{(k)} & p_{i1}^{(k)} & p_{i2}^{(k)} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \quad \text{称为 } k \text{ 步转移概率矩阵。}$$

(4) 转移概率矩阵性质：① $p_{ij}^{(k)} \geq 0, i, j \in I$ ；② $\sum_{j \in I} p_{ij}^{(k)} = 1, i \in I, k = 1, 2, \dots$ 。

(5) (C-K 方程) 对任意正整数 k, l ，及 $i, j \in I$ ，有 $p_{ij}^{(k+l)} = \sum_{r \in I} p_{ir}^{(k)} p_{rj}^{(l)}$ 。

(6) 设 Markov 链 $\{X_n, n \geq 0\}$ 的状态空间为 I ，若对一切 $i, j \in I$ ，存在不依赖于 i 的极限 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = p_j$ ，则称 Markov 链具有遍历性^[29]。（直观意义：具有遍历性的 Markov 链看做一个系统，系统不论从哪一个状态 i 出发，当转移步数 n 充分大后，转移到状态 j 的概率都接近于 p_j ，即经历足够长的一段时间后，系统会达到某平稳状态。）

4.2 预测方法

内部收益率（ IRR ）是一个相关度量指标，是指项目在整个计算期内的各年净现金流量的现值累计恰好等于零时的折现率（即净现值为零时的折现率）。内部收益率的求解公式如下：

$$NPV = \sum_{t=0}^n (CI - CO)_t (1 + IRR)^{-t} = 0.$$

其中， NPV 表示净现值（即项目的现金流的现值）； IRR 为内部收益率； $(CI - CO)_t$ 表示第 t 年的净现金流量； $(1 + IRR)^{-t}$ 表示第 t 年的折现系数。

关于内部收益率的详细内容，请自行了解。（参见[30---32]）

4.2.1 一种拟内部收益率及其求法

要求出一只股票的内部收益率需要有该公司或企业相应的资金流水明细，就目前状况来说调查收集企业的详细资金明细是难以办到，为了进行有关讨论，兹我们给出一种近似替代方法。

(1) 定义

设 \bar{P}_k 为某支股票的第 k 年的平均价格， $k = 1, 2, \dots, T$ ，则称

$$i = \frac{1}{T-1} \sum_{k=1}^{T-1} \frac{(\bar{P}_{k+1} - \bar{P}_k)}{\bar{P}_{k+1}}.$$

i 为这支股票的 T 年的拟内部收益率。

(2) 定义 4.2.1 具有如下的意义与应用：

- ①该拟内部收益率可在一定程度上近似替代实际内部收益率。
- ②由于在定义中没有进行先“折现”再求内部收益率，实际上将二者进行了结合，因而在实际中为了消除利率与内部收益率的影响，可只考虑拟内部收益率的影响，其可用于对股价进行“折现”，来统一单位。
- ③当研究股票问题时，利用所求拟内部收益率对相关股价、股指折现后再进行研究，可以更加有效地凸显出股价的波动趋势。

4.2.2 分类

为了利用马尔科夫链预测股票走势，要先按照如下做法对股价进行分类。

设 p 为某时刻的股价（本文中股价指每日股票的收盘价）， p_i 为一个价位，满足：

$$0 \leq p_1 < p_2 < \cdots < p_k < \infty, \quad i = 1, 2, \dots, k,$$

若 $p_i \leq p \leq p_{i+1}$ ，则称此时刻的股价处于状态 i 。

进行这样的分类后，一支股票在某一时刻的股票价格只有 $1, 2, \dots, k$ 个状态，于是其各个状态的价格便可以认为是状态空间为 $I = \{1, 2, \dots, k\}$ 的一个时间序列。为了方便研究，我们假设其为一个马尔科夫链。

4.2.3 求状态转移矩阵

当我们将一支股票的每日价格取某种平均 \bar{p} （例如：可取值为每日股价的开盘价、收盘价、最高价、最低价四者的平均值或者其他适当方法均可），按 4.2.2 中方法确定为 I 中的一个状态后，可将每日的价格所成序列视为一个马尔科夫链，利用公式： $X(t+k) = X(t)P(k)$ 进行预测，这里 $X(t)$ 为马尔科夫链在 t 时的取值， k 为 t 时对应的随机变量， $P(k)$ 为 k 步转移概率矩阵。这就需要我们先求出转移概率矩阵 P 。

设 $P_1, P_2, \dots, \{P_i\}_{i=1}^n$ 为一只股票从第 1 日至第 n 日的股价序列，再设 S_i 为处于 i 的状态数量， S_{ij} 为由 i 状态经一步转移到 j 状态的数量，一步转移概率矩阵为：

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

则 $p_{ij} \approx \frac{S_{ij}}{S_i}$ 。

因此，只要我们现分别统计出 S_i, S_{ij} ，则可近似求出 \mathbf{P} 。

4.2.4 预测方法

(1) 输入股价 p_1, p_2, \dots, p_T ，正整数 T ，计算出 $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_n$ ， $n = 365T$ ；

(2) 求出 T 年的拟内部收益率 i ；

(3) 折现：令 $p'_l = \bar{p}_l \frac{1}{(1+i)^{\frac{l}{244T}}}$ ， $l = 1, 2, \dots, n, n = 244T$ ；

(4) 按 4.2.2 的方法将每个 p' 确定为 I 中的一个状态 p''_l ；

(5) 根据 $\{p''\}$ ，按照 4.2.3 的方法求出概率转移矩阵 P ；

(6) 根据当前状况预测第 k 日的状况走势， $X(k) = X(0)P^{(k)}$ ，其中 $p_{ij}^{(k)}$ 是内部收益率从 i 出发经过 n 步转移到状态 j 的概率。

4.2.5 讨论：

在上述工作的基础上还可以进行如下的研究：

(1) 求出所讨论的股票的转移概率状况；

(2) 利用马尔科夫链相关知识，求得所讨论的股票数据的多步转移概率矩阵；

(3) 对比概率转移矩阵，讨论所讨论的股票是否具备马氏链的互通性、平稳性；

(4) 如若具备平稳性，求出其平稳状态的概率；

(5) 通过求出的转移概率估算值与股票实际的历史股价走势的比较，观察此种方法是否与符合实际股票走势保持一致，分析探究本章研究是否具备有效性。

4.3 实证分析

(1) 选取大连热电（600719）2011 年 1 月 4 日至 2015 年 12 月 31 日的股票历史收盘价作为参考数据（由于存在周六、周日，及其它情况会存在一些缺失值，我们按每年 244 天进行统计与计算）：

① 首先通过对股价的数据 p_1, p_2, \dots, p_T 统一折现处理，得到标准化数据 $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_k$ （见附表 1.PC）；

② 然后利用公式： $i = \frac{1}{T-1} \sum_{k=1}^{T-1} \frac{(\bar{p}_{k+1} - \bar{p}_k)}{\bar{p}_k}$ ，求出 2011 年---2015 年这 5 年的拟内部

部收益率为：

计算求得： $\bar{P}_1 = 10.868$ ， $\bar{P}_2 = 7.392$ ， $\bar{P}_3 = 7.112$ ， $\bar{P}_4 = 6.682$ ， $\bar{P}_5 = 9.9$ 。

代入公式：

$$\begin{aligned}
 i &= \frac{1}{T-1} \sum_{k=1}^{T-1} \frac{(\bar{P}_{k-1} - \bar{P}_k)}{\bar{P}_{k-1}} \\
 &= \frac{1}{4} \left(\frac{10.868 - 7.392}{10.868} + \frac{7.392 - 7.112}{7.392} + \frac{7.112 - 6.682}{7.112} + \frac{6.682 - 9.9}{6.682} \right) \\
 &\approx \frac{1}{4} \cdot (-0.06) \\
 &= -0.015
 \end{aligned}$$

③利用求得的拟内部收益率 i 折现：

$$p'_l = \bar{p}_l (1+i)^{-\frac{l}{244T}}, l=1, 2, \dots, n, n=244T.$$

其中， $i = -0.015$ ， $T=1, 2, 3, 4, 5$ 。

④求得概率转移矩阵 p'_l 见附表 2.(1)。根据折现出的矩阵 p'_l ，从低到高划分为 4 个四个状态：

表 4 数据的概率转移状态分组

状态	p''_1	p''_2	p''_3	p''_4
数据区间	(0, 0.25]	(0.25, 0.5]	(0.5, 0.75]	(0.75, 1]
出现次数	395	413	241	171

⑤根据上表处理可得状态转移矩阵如下表所示，表中每个数据表示由所在行的当前状态经过一步转移到达所在列状态的次数：

表 5 数据的概率转移状态

下一步状态 当前状态	p''_1	p''_2	p''_3	p''_4
p''_1	393	1	0	0
p''_2	1	412	0	0
p''_3	0	1	240	0
p''_4	0	0	1	170

并估计出该数据的转移概率矩阵 P ：

$$P = \begin{bmatrix} \frac{393}{394} & \frac{1}{394} & 0 & 0 \\ \frac{1}{413} & \frac{411}{413} & \frac{1}{413} & 0 \\ 0 & \frac{1}{241} & \frac{239}{241} & \frac{1}{241} \\ 0 & 0 & \frac{1}{171} & \frac{170}{171} \end{bmatrix} = \begin{bmatrix} 0.9975 & 0.0025 & 0 & 0 \\ 0.0024 & 0.9952 & 0.0024 & 0 \\ 0 & 0.0041 & 0.9917 & 0.0041 \\ 0 & 0 & 0.0058 & 0.9942 \end{bmatrix},$$

则 2、3、4、5 步转移概率矩阵分别为：

$$P^{(2)} = \begin{bmatrix} 0.9949 & 0.0051 & 0 & 0 \\ 0.0048 & 0.9904 & 0.0048 & 0 \\ 0 & 0.0082 & 0.9835 & 0.0082 \\ 0 & 0 & 0.0116 & 0.9884 \end{bmatrix}, P^{(3)} = \begin{bmatrix} 0.9924 & 0.0076 & 0 & 0 \\ 0 & 0.9856 & 0.0072 & 0 \\ 0 & 0.0123 & 0.9754 & 0.0123 \\ 0 & 0.0001 & 0.0173 & 0.9826 \end{bmatrix}$$

$$P^{(4)} = \begin{bmatrix} 0.9899 & 0.0100 & 0 & 0 \\ 0.0096 & 0.9809 & 0.0095 & 0.0001 \\ 0.0001 & 0.0163 & 0.9674 & 0.0162 \\ 0 & 0.0001 & 0.0229 & 0.9770 \end{bmatrix}, P^{(5)} = \begin{bmatrix} 0.9874 & 0.0125 & 0.0001 & 0 \\ 0.0119 & 0.9762 & 0.0118 & 0.0001 \\ 0.0001 & 0.0202 & 0.9595 & 0.0202 \\ 0 & 0.0002 & 0.0284 & 0.9713 \end{bmatrix},$$

当转移步数趋向于无穷大时，转移概率矩阵为：

$$P^{(\infty)} = \begin{bmatrix} 0.3233 & 0.3388 & 0.1977 & 0.1403 \\ 0.3233 & 0.3388 & 0.1977 & 0.1403 \\ 0.3233 & 0.3388 & 0.1977 & 0.1403 \\ 0.3233 & 0.3388 & 0.1977 & 0.1403 \end{bmatrix}.$$

(2) 同样，选取另为一种股票，中国石油（601857）2012 年 1 月 4 日至 2016 年 12 月 31 日的股票历史收盘价作为参考数据（由于存在周六、周日，及其它情况会存在一些缺失值，我们按每年 244 天进行统计与计算），作为参考对照组进行分析：

①同理，通过对股价的数据 p_1, p_2, \dots, p_T 统一折现处理，得到标准化数据 $\bar{p}_1, \bar{p}_2, \dots, \bar{p}_k$ ；

②然后利用公式： $i = \frac{1}{T-1} \sum_{k=1}^{T-1} \frac{(\bar{p}_{k+1} - \bar{p}_k)}{\bar{p}_k}$ ，求出 2012 年---2016 年这 5 年的拟内

部收益率为：

计算求得： $\bar{p}_1 = 10.866$, $\bar{p}_2 = 9.353$, $\bar{p}_3 = 8.624$, $\bar{p}_4 = 11.526$, $\bar{p}_5 = 7.681$.

代入公式：

$$\begin{aligned}
 i &= \frac{1}{T-1} \sum_{k=1}^{T-1} \frac{(\bar{P}_{k-1} - \bar{P}_k)}{\bar{P}_{k-1}} \\
 &= \frac{1}{4} \left(\frac{10.866 - 9.353}{10.866} + \frac{9.353 - 8.624}{9.353} + \frac{8.624 - 11.526}{8.624} + \frac{11.526 - 7.681}{11.526} \right) \\
 &\approx \frac{1}{4} \square (0.21) \\
 &= 0.0525
 \end{aligned}$$

③利用求得的拟内部收益率 i 折现:

$$p'_l = \bar{p}_l (1+i)^{-\frac{l}{244T}}, l=1, 2, \dots, n, n=244T.$$

其中, $i=0.0525$, $T=1, 2, 3, 4, 5$.

④求得概率转移矩阵 p'_l 见附表 2.(2)。根据折现出的矩阵 p'_l , 从低到高划分为 4 个四个状态:

表 6 数据的概率转移状态分组

状态	p''_1	p''_2	p''_3	p''_4
数据区间	(0, 9]	(9, 11]	(11, 13]	(13, ∞]
出现次数	644	329	206	41

⑤根据上表处理可得状态转移矩阵如下表所示, 表中每个数据表示由所在行的当前状态经过一步转移到达所在列状态的次数, 并估计出转移概率矩阵 P :

表 7 数据的概率转移状态

下一步状态 当前状态	p''_1	p''_2	p''_3	p''_4
p''_1	629	12	0	0
p''_2	12	312	5	0
p''_3	0	6	197	8
p''_4	0	0	8	30

$$P = \begin{bmatrix} \frac{629}{329} & \frac{12}{329} & 0 & 0 \\ \frac{641}{329} & \frac{641}{329} & \frac{5}{329} & 0 \\ 0 & \frac{6}{211} & \frac{197}{211} & \frac{8}{211} \\ 0 & 0 & \frac{8}{38} & \frac{30}{38} \end{bmatrix} = \begin{bmatrix} 0.9813 & 0.0187 & 0 & 0 \\ 0.0365 & 0.9483 & 0.0152 & 0 \\ 0 & 0.0284 & 0.9336 & 0.0379 \\ 0 & 0 & 0.2105 & 0.7895 \end{bmatrix},$$

则 2、3、4、5 步转移概率矩阵分别为：

$$P^{(2)} = \begin{bmatrix} 0.9636 & 0.0361 & 0.0003 & 0 \\ 0.0704 & 0.9004 & 0.0286 & 0.0006 \\ 0.0010 & 0.0535 & 0.8801 & 0.0653 \\ 0 & 0.0060 & 0.3628 & 0.6313 \end{bmatrix}, P^{(3)} = \begin{bmatrix} 0.9469 & 0.0523 & 0.0008 & 0 \\ 0.1019 & 0.8560 & 0.0405 & 0.0015 \\ 0.0030 & 0.0758 & 0.8363 & 0.0849 \\ 0.0002 & 0.0160 & 0.4717 & 0.5121 \end{bmatrix}$$

$$P^{(4)} = \begin{bmatrix} 0.9311 & 0.0674 & 0.0016 & 0 \\ 0.1312 & 0.8149 & 0.0512 & 0.0028 \\ 0.0057 & 0.0957 & 0.7998 & 0.0988 \\ 0.0008 & 0.0286 & 0.5484 & 0.4222 \end{bmatrix}, P^{(5)} = \begin{bmatrix} 0.9161 & 0.0813 & 0.0025 & 0.0001 \\ 0.1585 & 0.7767 & 0.0607 & 0.0041 \\ 0.0091 & 0.1136 & 0.7690 & 0.1083 \\ 0.0018 & 0.0427 & 0.6014 & 0.3541 \end{bmatrix},$$

当转移步数趋向于无穷大时，转移概率矩阵为：

$$P^{(\infty)} = \begin{bmatrix} 0.5444 & 0.2794 & 0.1493 & 0.0269 \\ 0.5444 & 0.2794 & 0.1493 & 0.0269 \\ 0.5444 & 0.2794 & 0.1493 & 0.0269 \\ 0.5444 & 0.2794 & 0.1493 & 0.0269 \end{bmatrix}.$$

通过对比两支股票的五年数据求得的多步转移概率矩阵，发现以下特点：

(1) 互通性：在对上述两只股票数据处理的过程中，通过对我们所划分的各个状态间的相互转移进行观察分析发现：无论股价从是哪个状态出发的，经过一段时间后，都会到达其他状态（概率大于 0），也就是说股票的各个状态之间存在着互通性。

(2) 平稳性：以内具有互通性，所以由 4.1 基础知识可知，该过程是遍历的。即从任意某一状态出发，经过足够长的转移步数后整个系统都会达到一种平稳状态，简称稳态。

(3) 除此之外，我们也可以通过解方程组的方式来得出股价的平稳状态：

$$\begin{cases} p_1'' = (393/394) \cdot p_1'' + (1/413) \cdot p_2'' + 0 \cdot p_3'' + 0 \cdot p_4'' \\ p_2'' = (1/394) \cdot p_1'' + (411/413) \cdot p_2'' + (1/241) \cdot p_3'' + 0 \cdot p_4'' \\ p_3'' = 0 \cdot p_1'' + (1/413) \cdot p_2'' + (239/241) \cdot p_3'' + (1/171) \cdot p_4'' \\ p_4'' = 0 \cdot p_1'' + 0 \cdot p_2'' + (1/241) \cdot p_3'' + (170/171) \cdot p_4'' \end{cases}, \quad (1)$$

$$p_1'' + p_2'' + p_3'' + p_4'' = 1,$$

即有

$$q_1 = 0.3233, q_2 = 0.3388, q_3 = 0.1977, q_4 = 0.1403.$$

$$\begin{cases} p_1'' = (629/641) \cdot p_1'' + (12/329) \cdot p_2'' + 0 \cdot p_3'' + 0 \cdot p_4'' \\ p_2'' = (12/641) \cdot p_1'' + (312/329) \cdot p_2'' + (6/211) \cdot p_3'' + 0 \cdot p_4'' \\ p_3'' = 0 \cdot p_1'' + (5/329) \cdot p_2'' + (197/211) \cdot p_3'' + (8/38) \cdot p_4'' \\ p_4'' = 0 \cdot p_1'' + 0 \cdot p_2'' + (8/211) \cdot p_3'' + (30/38) \cdot p_4'' \end{cases}, \quad (2)$$

$$p_1'' + p_2'' + p_3'' + p_4'' = 1,$$

即有

$$p_1'' = 0.5444, p_2'' = 0.2794, p_3'' = 0.1493, p_4'' = 0.0269.$$

总结上述论证：当转移步数无限大时，股价的概率转移矩阵达到稳态，概率分布保持一致性。

(4) 通过两组股票数据得出的转移概率估算值，可以看出股票涨幅的概率比较小，符合现实中两支股票的历史股价走势的实际情况。因此可以通过马尔科夫链，大致的预测出两只股票价格的走势。

马尔科夫预测法是一种只受上一状态所影响的随机过程的预测方法，它通过对初始状态和状态转移矩阵来预测，短期内具有有效性。但长远角度来看，状态转移矩阵必定会因各种因素的影响发生变化，所以马尔科夫预测法不适合长期预测，仅适合短期预测且需要考虑股票价格的综合影响因素进行进一步的分析探究。

结语与展望

本文受以往对股票市场的各类研究的启发，在此基础之上与时代的特点相结合，符合当前学术领域对大数据与行为经济学的研究需求。本文的主要内容是：首先建立了一个表现某种初级炒股行为收益状况的随机变量 R 并给出了一种基于某支股票的一定的折现历史数据运用所设计的获取其样本的方法，给出一种求优化近似解析分布的方法，基于 R 的生成方式及相关探讨展示它的意义，并揭示出一些值得进一步研究的问题。所做工作可为炒股提供一定的启示，也为从大数据、机器学习和行为金融学的角度出发进一步研究股市提供一定的启示。然后，提出了一种用均线来预测股价走势的方法，利用股票历史大数据的建立长短均线预测走势模型进行分析研究。除此以外，运用马尔科夫预测法的思想对股票大数据分析探讨：通过构建马尔科夫链，运用目前的已知数据来确定初始状态的概率向量并求出其对应的状态转移矩阵，从而预测模型在未来某一时刻的状态。利用马尔科夫链预测法对股票价格进行进一步的研究，旨在能够帮助炒股人更精准更有效的预测股市价格的波动趋势。

通过本文的三个方法研究以及相应的实证分析，可以得出如下结论：三种方法在一定程度上都可以对股市进行有效的预测股市价格的变动趋势，其中初级炒股行为收益状况的随机变量 R 较适合利用股票大数据来进行长期的数据趋势预测汇总；而均线预测股价模型比较适合预估股价的变动幅度大小；马尔科夫预测模型对短期预测的准确性更高。

本文的研究是建立在一定的假设条件下，利用计算机软件与实际的股票数据相结合进行的初步的分析研究，因此具有一定的局限性。学者们可以建立更加符合实际情况的股票模型，充分利用大数据与计算机进行讨论研究。衷心地期望我们的工作能够激发同志们利用大数据和行为经济学方法研究股市的热情，共同推动股市研究的深入发展！

参考文献

- [1] KIM K-J, LEE W B. Stock market prediction using artificial neural networks with optimal feature transformation[J]. NEURAL COMPUT APPL, 2004,13: 255-260.
- [2] FONSEKA C, LIYANAGE L. A data mining algorithm to analyse stock market data using lagged correlation[C]//ICIAFS08,2008:163-166.
- [3] 李涛, 郭杰. 风险态度与股票投资[J].经济研究,2009(2):56-67.
- [4] 王嘉展, 刘丽霞. 随机利率下股票价格服从几何分数布朗运动的幂期权定价[J], 河北师范大学学报(自然科学版), 2014, 38 (2) :113 - 116.
- [5] 郑薇,开璇,谢心庆.统计分析方法在股票最优投资的应用[J].贵州商业高等专科学校学报,2015,28(1): 49-52.
- [6] 朱泓嘉.马尔科夫预测法在股票价格预测中的应用[J].时代金融,2015(610): 294-299.
- [7] 张力.股票市场投资组合策略构造及模型检验[J].海南热带海洋学院学报,2016,23(5):108-113.
- [8] 黄子颖.股票走势预测的随机分析方法研究[J].物流工程与管理,2016, (07): 251-252.
- [9] 黄玉洁, 钱丽丽, 陶凤梅.股票市场数据的统计实证分析[J].鞍山师范学院学报, 2016, (02):19-23.
- [10] 张莉莉, 许英.基于 Markov 链对股票价格的预测应用[J].科技经济市场, 2016,(11):89-91.
- [11] 王越.基于马尔科夫和布朗运动的股票价格预测模型[J].科技致富向导, 2015,(05):163.
- [12] 王宏志.大数据技术丛书 大数据算法[M].北京: 机械工业出版社.2015.
- [13] (美) 西蒙.大数据应用 商业案例实践[M].北京: 人民邮电出版社.2014.
- [14] 唐滔.基于云计算的股票数据分析研究与实现[D],湖南:湖南大学,2013.
- [15] 王春峰,李嘉毅,房振明. 大数据下媒体关注度与股票收益关系研究[J]. 天津大学学报(社会科学版),2016,(02):103-108.
- [16] 席文帅,申瑞新,吕文英,纪楠. 大数据时代股票投资策略的探讨[J]. 现代经济信息, 2016, (11):295.
- [17] 李静.基于行为金融学的股票市场投资者行为研究[D]. 北京: 中国社会科学院研究生院, 2012.
- [18] 杜飞.行为金融学视角下的股票市场动量效应和反转效应形成机制[J].时代金融,2014(8): 114-114.
- [19] 李涛,张文韬.人格特征与股票投资[J].经济研究, 2015,6:103-116.
- [20] 黄润鹏,左文明,毕凌燕.基于微博情绪信息的股票市场预测[J].管理工程学报,2015,(01):47-52.
- [21] 蒋翠清,梁坤,丁勇,刘士喜,刘尧. 基于社会媒体的股票行为预测[J]. 中国管理科学,2015,(01):17-24.
- [22] 陈倩,李金林,邹庆忠.基于 g-h 分布的股票收益率风险价值研究[J].兵工学报,2009,30 (增刊 1) : 175 - 180.
- [23] 钟波,山宇.基于随机模拟与 g-h 分布的 VaR 计算方法[J].统计与决策,2013(15):8.
- [24] 王思全.统计学在股票价格统计数据中的应用[J].中国市场,2016,(05):103.
- [25] 崔红军主编.证券学基础[M].上海: 华东理工大学出版社.2015.
- [26] 江生忠, 张雪冰.保险投资[M].北京: 中国金融出版社.1993.
- [27] 李德荃主编.金融市场学[M].北京: 对外经济贸易大学出版社.2011.
- [28] 伍海华, 杨德平编著.随机过程 金融资产定价之应用[M].北京: 中国金融出版社.2002.
- [29] 宋占杰,王家生,王勇.随机过程基础[M].天津: 天津大学出版社,2011.

- [30] 徐蓉.工程经济学[M].北京：冶金工业出版社.2015.
- [31] 徐景峰.金融数学（精算师）[M].北京：中国财政经济出版社.2010.
- [32] 王力国.图解经济学[M].北京：石油工业出版社.2015.

6.52 6.73 6.80 6.75 6.69 6.69;6.65 6.65 6.65 6.65 6.77 6.80 6.78 6.48 6.54 6.53 6.34 6.38 6.35
6.16 6.24 6.27 6.30 6.24 6.17 6.26 6.08 5.96 6.06 5.98 6.17 6.38 6.64 6.64 6.55 6.52 6.43 6.46
6.62 6.59 6.72 6.86 7.14 6.80 6.55 6.74 6.56 6.71 6.44 6.35 6.39 6.21 6.19 6.27 6.48 6.29 6.27
6.35 6.33 6.43 6.36 6.29 6.28 6.41 6.40 6.51 6.64 6.04 6.01 5.98 5.81 5.94 6.00 5.99 5.92 6.13
6.12 6.13 6.34 6.38 6.43 6.51 6.47 6.54 6.28 6.21 6.14 6.21 6.15 6.05 6.12 6.08 6.05 6.07 6.16
6.03 6.07 5.96 6.04 6.00 6.03 6.14 6.12 6.33 6.24 6.15 6.23 6.15 5.95 5.86 5.95 5.85 6.09 6.20
6.17 6.16 6.28 6.23 6.36 6.32 6.21 6.19 6.05 6.36 6.68 6.98 7.75 7.69 7.82 7.60 7.71 7.63 7.78
7.69 7.55 7.57 7.71 7.98 8.02 7.94 8.19 8.16 8.14 8.38 8.55 8.06 8.30 7.90 7.73 7.60 7.51 7.53
7.84 7.59 7.55 7.53 7.43 7.27 7.52 7.45 7.46 7.39 7.73 7.73 7.65 7.85 7.60 7.61 7.80 7.82 8.00
8.16 7.80 7.75 7.79 7.08 6.87 7.01 6.60 6.81 6.81 6.89 6.69 6.54 6.65 6.53 6.41 6.28 6.18 6.13
6.20 6.12 5.96 6.01 6.16 6.47 6.50 6.60 6.86 6.91 6.73 6.88 6.80 6.67 6.85 7.28 7.25 7.14 6.99
6.92 7.00 7.17 7.13 7.06 7.12 7.04 7.11 7.16 7.21 6.89 6.81 6.78 6.83 7.06 6.77 6.64 6.69 6.63
6.66 6.59 6.48 6.68 6.73 6.79 6.69 6.08 5.97 5.85 5.79 5.79 5.86 5.56 5.77 5.98 5.95 6.03 5.75
5.49 5.75;5.86 5.72 5.73 5.77 5.97 6.09 5.99 6.44 6.82 6.83 6.80 6.44 6.66 6.61 7.07 7.37 7.55
7.61 7.62 7.76 8.25 8.76 8.63 8.49 8.29 8.42 8.74 9.11 9.59 9.57 8.70 8.78 8.46 8.55 8.43 8.12
8.35 8.39 8.33 8.29 8.41 8.39 8.12 8.16 8.19 8.21 8.10 8.24 8.42 7.82 7.64 7.67 8.13 8.15 8.52
8.54 8.46 8.43 8.18 8.15 8.22 8.15 8.60 8.85 8.79 8.94 9.09 9.41 9.18 9.07 9.23 9.26 9.36 9.29
9.20 9.28 9.46 9.28 9.25 9.57 9.83 9.68 9.71 9.74 10.08 10.34 10.38 9.57 9.12 9.26 9.42 9.64 9.89
10.08 9.72 9.55 9.46 9.37 9.68 10.46 10.86 10.92 11.25 11.39 10.77 10.83 10.79 10.54 10.62 11.45
11.40 11.91 11.92 11.69 11.81 11.92 11.88 11.53 11.68 11.43 11.34 11.33 11.38 11.38 10.98 11.06
10.78 10.94 11.08 10.78 10.54 10.28 10.54 10.36 10.30 10.89 11.20 11.16 10.82 10.84 10.94 10.91
10.67 10.61 10.34 10.65 10.36 10.01 11.11 11.58 10.53 11.11 11.05 11.87 11.46 11.73 11.35 11.51
11.23 11.40 11.20 11.19 11.07 10.40 10.23 10.18 10.67 10.58 10.17 10.55 10.92 11.52 12.13 12.10
11.84 11.41 11.59 11.49 11.69 12.15 12.20 11.53 11.00 11.05 10.76 10.85 10.61 10.78 10.65 10.93
10.88 11.04 11.69 11.67 11.75 11.32 11.35 11.71 11.71 11.47 11.94 12.59 12.46 12.65 12.21 11.25
11.34 11.21 11.04 10.04 9.94 10.06 10.28 10.14 10.38 10.28 10.00 9.90 9.71 9.82 9.59 9.68 9.11
9.40 9.42 9.36 9.26 9.10 9.01 9.06 9.56 9.57 10.27 9.89 9.94 9.89 10.24 10.58 10.47 10.73 11.06
11.42 11.40 11.98];

附表 2

(1) $p'_l =$

Columns 1 through 11

0.9983 0.9966 0.9950 0.9933 0.9916 0.9900 0.9883 0.9867 0.9850 0.9834 0.9817

Columns 12 through 22

0.9801 0.9784 0.9768 0.9751 0.9735 0.9719 0.9702 0.9686 0.9670 0.9654 0.9637

Columns 23 through 33

0.9621 0.9605 0.9589 0.9573 0.9557 0.9541 0.9525 0.9509 0.9493 0.9477 0.9461

Columns 34 through 44

0.9445 0.9429 0.9414 0.9398 0.9382 0.9366 0.9351 0.9335 0.9319 0.9304 0.9288

Columns 45 through 55

0.9273 0.9257 0.9241 0.9226 0.9210 0.9195 0.9180 0.9164 0.9149 0.9134 0.9118

Columns 56 through 66

0.9103 0.9088 0.9072 0.9057 0.9042 0.9027 0.9012 0.8997 0.8981 0.8966 0.8951

Columns 67 through 77										
0.8936	0.8921	0.8906	0.8891	0.8877	0.8862	0.8847	0.8832	0.8817	0.8802	0.8788
Columns 78 through 88										
0.8773	0.8758	0.8743	0.8729	0.8714	0.8700	0.8685	0.8670	0.8656	0.8641	0.8627
Columns 89 through 99										
0.8612	0.8598	0.8584	0.8569	0.8555	0.8540	0.8526	0.8512	0.8498	0.8483	0.8469
Columns 100 through 110										
0.8455	0.8441	0.8427	0.8412	0.8398	0.8384	0.8370	0.8356	0.8342	0.8328	0.8314
Columns 111 through 121										
0.8300	0.8286	0.8272	0.8259	0.8245	0.8231	0.8217	0.8203	0.8189	0.8176	0.8162
Columns 122 through 132										
0.8148	0.8135	0.8121	0.8107	0.8094	0.8080	0.8067	0.8053	0.8040	0.8026	0.8013
Columns 133 through 143										
0.7999	0.7986	0.7972	0.7959	0.7946	0.7932	0.7919	0.7906	0.7893	0.7879	0.7866
Columns 144 through 154										
0.7853	0.7840	0.7827	0.7814	0.7800	0.7787	0.7774	0.7761	0.7748	0.7735	0.7722
Columns 155 through 165										
0.7709	0.7696	0.7683	0.7671	0.7658	0.7645	0.7632	0.7619	0.7606	0.7594	0.7581
Columns 166 through 176										
0.7568	0.7556	0.7543	0.7530	0.7518	0.7505	0.7492	0.7480	0.7467	0.7455	0.7442
Columns 177 through 187										
0.7430	0.7417	0.7405	0.7393	0.7380	0.7368	0.7355	0.7343	0.7331	0.7318	0.7306
Columns 188 through 198										
0.7294	0.7282	0.7269	0.7257	0.7245	0.7233	0.7221	0.7209	0.7197	0.7185	0.7173
Columns 199 through 209										
0.7160	0.7148	0.7136	0.7125	0.7113	0.7101	0.7089	0.7077	0.7065	0.7053	0.7041
Columns 210 through 220										
0.7029	0.7018	0.7006	0.6994	0.6982	0.6971	0.6959	0.6947	0.6936	0.6924	0.6912
Columns 221 through 231										
0.6901	0.6889	0.6878	0.6866	0.6855	0.6843	0.6832	0.6820	0.6809	0.6797	0.6786
Columns 232 through 242										
0.6775	0.6763	0.6752	0.6741	0.6729	0.6718	0.6707	0.6696	0.6684	0.6673	0.6662
Columns 243 through 253										
0.6651	0.6640	0.6628	0.6617	0.6606	0.6595	0.6584	0.6573	0.6562	0.6551	0.6540
Columns 254 through 264										
0.6529	0.6518	0.6507	0.6496	0.6485	0.6474	0.6464	0.6453	0.6442	0.6431	0.6420
Columns 265 through 275										
0.6410	0.6399	0.6388	0.6377	0.6367	0.6356	0.6345	0.6335	0.6324	0.6314	0.6303
Columns 276 through 286										
0.6292	0.6282	0.6271	0.6261	0.6250	0.6240	0.6229	0.6219	0.6208	0.6198	0.6188
Columns 287 through 297										
0.6177	0.6167	0.6157	0.6146	0.6136	0.6126	0.6115	0.6105	0.6095	0.6085	0.6074
Columns 298 through 308										
0.6064	0.6054	0.6044	0.6034	0.6024	0.6014	0.6003	0.5993	0.5983	0.5973	0.5963

Columns 309 through 319										
0.5953	0.5943	0.5933	0.5923	0.5913	0.5904	0.5894	0.5884	0.5874	0.5864	0.5854
Columns 320 through 330										
0.5844	0.5835	0.5825	0.5815	0.5805	0.5796	0.5786	0.5776	0.5766	0.5757	0.5747
Columns 331 through 341										
0.5737	0.5728	0.5718	0.5709	0.5699	0.5690	0.5680	0.5670	0.5661	0.5651	0.5642
Columns 342 through 352										
0.5633	0.5623	0.5614	0.5604	0.5595	0.5585	0.5576	0.5567	0.5557	0.5548	0.5539
Columns 353 through 363										
0.5529	0.5520	0.5511	0.5502	0.5492	0.5483	0.5474	0.5465	0.5456	0.5447	0.5437
Columns 364 through 374										
0.5428	0.5419	0.5410	0.5401	0.5392	0.5383	0.5374	0.5365	0.5356	0.5347	0.5338
Columns 375 through 385										
0.5329	0.5320	0.5311	0.5302	0.5293	0.5285	0.5276	0.5267	0.5258	0.5249	0.5240
Columns 386 through 396										
0.5232	0.5223	0.5214	0.5205	0.5197	0.5188	0.5179	0.5170	0.5162	0.5153	0.5144
Columns 397 through 407										
0.5136	0.5127	0.5119	0.5110	0.5101	0.5093	0.5084	0.5076	0.5067	0.5059	0.5050
Columns 408 through 418										
0.5042	0.5033	0.5025	0.5017	0.5008	0.5000	0.4991	0.4983	0.4975	0.4966	0.4958
Columns 419 through 429										
0.4950	0.4941	0.4933	0.4925	0.4917	0.4908	0.4900	0.4892	0.4884	0.4875	0.4867
Columns 430 through 440										
0.4859	0.4851	0.4843	0.4835	0.4827	0.4819	0.4810	0.4802	0.4794	0.4786	0.4778
Columns 441 through 451										
0.4770	0.4762	0.4754	0.4746	0.4738	0.4730	0.4722	0.4715	0.4707	0.4699	0.4691
Columns 452 through 462										
0.4683	0.4675	0.4667	0.4659	0.4652	0.4644	0.4636	0.4628	0.4621	0.4613	0.4605
Columns 463 through 473										
0.4597	0.4590	0.4582	0.4574	0.4567	0.4559	0.4551	0.4544	0.4536	0.4528	0.4521
Columns 474 through 484										
0.4513	0.4506	0.4498	0.4491	0.4483	0.4475	0.4468	0.4460	0.4453	0.4446	0.4438
Columns 485 through 495										
0.4431	0.4423	0.4416	0.4408	0.4401	0.4394	0.4386	0.4379	0.4372	0.4364	0.4357
Columns 496 through 506										
0.4350	0.4342	0.4335	0.4328	0.4320	0.4313	0.4306	0.4299	0.4292	0.4284	0.4277
Columns 507 through 517										
0.4270	0.4263	0.4256	0.4249	0.4241	0.4234	0.4227	0.4220	0.4213	0.4206	0.4199
Columns 518 through 528										
0.4192	0.4185	0.4178	0.4171	0.4164	0.4157	0.4150	0.4143	0.4136	0.4129	0.4122
Columns 529 through 539										
0.4115	0.4108	0.4101	0.4095	0.4088	0.4081	0.4074	0.4067	0.4060	0.4054	0.4047
Columns 540 through 550										
0.4040	0.4033	0.4026	0.4020	0.4013	0.4006	0.3999	0.3993	0.3986	0.3979	0.3973

Columns 551 through 561										
0.3966	0.3959	0.3953	0.3946	0.3939	0.3933	0.3926	0.3920	0.3913	0.3907	0.3900
Columns 562 through 572										
0.3893	0.3887	0.3880	0.3874	0.3867	0.3861	0.3854	0.3848	0.3842	0.3835	0.3829
Columns 573 through 583										
0.3822	0.3816	0.3809	0.3803	0.3797	0.3790	0.3784	0.3778	0.3771	0.3765	0.3759
Columns 584 through 594										
0.3752	0.3746	0.3740	0.3733	0.3727	0.3721	0.3715	0.3709	0.3702	0.3696	0.3690
Columns 595 through 605										
0.3684	0.3678	0.3671	0.3665	0.3659	0.3653	0.3647	0.3641	0.3635	0.3628	0.3622
Columns 606 through 616										
0.3616	0.3610	0.3604	0.3598	0.3592	0.3586	0.3580	0.3574	0.3568	0.3562	0.3556
Columns 617 through 627										
0.3550	0.3544	0.3538	0.3532	0.3526	0.3520	0.3515	0.3509	0.3503	0.3497	0.3491
Columns 628 through 638										
0.3485	0.3479	0.3474	0.3468	0.3462	0.3456	0.3450	0.3444	0.3439	0.3433	0.3427
Columns 639 through 649										
0.3421	0.3416	0.3410	0.3404	0.3399	0.3393	0.3387	0.3381	0.3376	0.3370	0.3365
Columns 650 through 660										
0.3359	0.3353	0.3348	0.3342	0.3336	0.3331	0.3325	0.3320	0.3314	0.3309	0.3303
Columns 661 through 671										
0.3297	0.3292	0.3286	0.3281	0.3275	0.3270	0.3264	0.3259	0.3253	0.3248	0.3243
Columns 672 through 682										
0.3237	0.3232	0.3226	0.3221	0.3215	0.3210	0.3205	0.3199	0.3194	0.3189	0.3183
Columns 683 through 693										
0.3178	0.3173	0.3167	0.3162	0.3157	0.3151	0.3146	0.3141	0.3135	0.3130	0.3125
Columns 694 through 704										
0.3120	0.3115	0.3109	0.3104	0.3099	0.3094	0.3088	0.3083	0.3078	0.3073	0.3068
Columns 705 through 715										
0.3063	0.3058	0.3052	0.3047	0.3042	0.3037	0.3032	0.3027	0.3022	0.3017	0.3012
Columns 716 through 726										
0.3007	0.3002	0.2997	0.2992	0.2987	0.2982	0.2977	0.2972	0.2967	0.2962	0.2957
Columns 727 through 737										
0.2952	0.2947	0.2942	0.2937	0.2932	0.2927	0.2922	0.2917	0.2912	0.2907	0.2903
Columns 738 through 748										
0.2898	0.2893	0.2888	0.2883	0.2878	0.2873	0.2869	0.2864	0.2859	0.2854	0.2849
Columns 749 through 759										
0.2845	0.2840	0.2835	0.2830	0.2826	0.2821	0.2816	0.2811	0.2807	0.2802	0.2797
Columns 760 through 770										
0.2793	0.2788	0.2783	0.2779	0.2774	0.2769	0.2765	0.2760	0.2755	0.2751	0.2746
Columns 771 through 781										
0.2742	0.2737	0.2732	0.2728	0.2723	0.2719	0.2714	0.2709	0.2705	0.2700	0.2696
Columns 782 through 792										
0.2691	0.2687	0.2682	0.2678	0.2673	0.2669	0.2664	0.2660	0.2655	0.2651	0.2647

Columns 793 through 803										
0.2642	0.2638	0.2633	0.2629	0.2624	0.2620	0.2616	0.2611	0.2607	0.2603	0.2598
Columns 804 through 814										
0.2594	0.2589	0.2585	0.2581	0.2576	0.2572	0.2568	0.2563	0.2559	0.2555	0.2551
Columns 815 through 825										
0.2546	0.2542	0.2538	0.2534	0.2529	0.2525	0.2521	0.2517	0.2512	0.2508	0.2504
Columns 826 through 836										
0.2500	0.2496	0.2491	0.2487	0.2483	0.2479	0.2475	0.2471	0.2466	0.2462	0.2458
Columns 837 through 847										
0.2454	0.2450	0.2446	0.2442	0.2438	0.2434	0.2429	0.2425	0.2421	0.2417	0.2413
Columns 848 through 858										
0.2409	0.2405	0.2401	0.2397	0.2393	0.2389	0.2385	0.2381	0.2377	0.2373	0.2369
Columns 859 through 869										
0.2365	0.2361	0.2357	0.2353	0.2349	0.2345	0.2341	0.2337	0.2334	0.2330	0.2326
Columns 870 through 880										
0.2322	0.2318	0.2314	0.2310	0.2306	0.2302	0.2299	0.2295	0.2291	0.2287	0.2283
Columns 881 through 891										
0.2279	0.2276	0.2272	0.2268	0.2264	0.2260	0.2256	0.2253	0.2249	0.2245	0.2241
Columns 892 through 902										
0.2238	0.2234	0.2230	0.2226	0.2223	0.2219	0.2215	0.2211	0.2208	0.2204	0.2200
Columns 903 through 913										
0.2197	0.2193	0.2189	0.2186	0.2182	0.2178	0.2175	0.2171	0.2167	0.2164	0.2160
Columns 914 through 924										
0.2157	0.2153	0.2149	0.2146	0.2142	0.2138	0.2135	0.2131	0.2128	0.2124	0.2121
Columns 925 through 935										
0.2117	0.2114	0.2110	0.2106	0.2103	0.2099	0.2096	0.2092	0.2089	0.2085	0.2082
Columns 936 through 946										
0.2078	0.2075	0.2071	0.2068	0.2064	0.2061	0.2058	0.2054	0.2051	0.2047	0.2044
Columns 947 through 957										
0.2040	0.2037	0.2033	0.2030	0.2027	0.2023	0.2020	0.2016	0.2013	0.2010	0.2006
Columns 958 through 968										
0.2003	0.2000	0.1996	0.1993	0.1990	0.1986	0.1983	0.1980	0.1976	0.1973	0.1970
Columns 969 through 979										
0.1966	0.1963	0.1960	0.1956	0.1953	0.1950	0.1947	0.1943	0.1940	0.1937	0.1934
Columns 980 through 990										
0.1930	0.1927	0.1924	0.1921	0.1917	0.1914	0.1911	0.1908	0.1905	0.1901	0.1898
Columns 991 through 1001										
0.1895	0.1892	0.1889	0.1886	0.1882	0.1879	0.1876	0.1873	0.1870	0.1867	0.1864
Columns 1002 through 1012										
0.1860	0.1857	0.1854	0.1851	0.1848	0.1845	0.1842	0.1839	0.1836	0.1833	0.1829
Columns 1013 through 1023										
0.1826	0.1823	0.1820	0.1817	0.1814	0.1811	0.1808	0.1805	0.1802	0.1799	0.1796
Columns 1024 through 1034										
0.1793	0.1790	0.1787	0.1784	0.1781	0.1778	0.1775	0.1772	0.1769	0.1766	0.1763

Columns 1035 through 1045

0.1760 0.1757 0.1754 0.1751 0.1748 0.1745 0.1743 0.1740 0.1737 0.1734 0.1731

Columns 1046 through 1056

0.1728 0.1725 0.1722 0.1719 0.1716 0.1714 0.1711 0.1708 0.1705 0.1702 0.1699

Columns 1057 through 1067

0.1696 0.1694 0.1691 0.1688 0.1685 0.1682 0.1679 0.1677 0.1674 0.1671 0.1668

Columns 1068 through 1078

0.1665 0.1663 0.1660 0.1657 0.1654 0.1651 0.1649 0.1646 0.1643 0.1640 0.1638

Columns 1079 through 1089

0.1635 0.1632 0.1629 0.1627 0.1624 0.1621 0.1618 0.1616 0.1613 0.1610 0.1608

Columns 1090 through 1100

0.1605 0.1602 0.1600 0.1597 0.1594 0.1592 0.1589 0.1586 0.1584 0.1581 0.1578

Columns 1101 through 1111

0.1576 0.1573 0.1570 0.1568 0.1565 0.1562 0.1560 0.1557 0.1555 0.1552 0.1549

Columns 1112 through 1122

0.1547 0.1544 0.1542 0.1539 0.1536 0.1534 0.1531 0.1529 0.1526 0.1524 0.1521

Columns 1123 through 1133

0.1518 0.1516 0.1513 0.1511 0.1508 0.1506 0.1503 0.1501 0.1498 0.1496 0.1493

Columns 1134 through 1144

0.1491 0.1488 0.1486 0.1483 0.1481 0.1478 0.1476 0.1473 0.1471 0.1468 0.1466

Columns 1145 through 1155

0.1463 0.1461 0.1459 0.1456 0.1454 0.1451 0.1449 0.1446 0.1444 0.1441 0.1439

Columns 1156 through 1166

0.1437 0.1434 0.1432 0.1429 0.1427 0.1425 0.1422 0.1420 0.1417 0.1415 0.1413

Columns 1167 through 1177

0.1410 0.1408 0.1406 0.1403 0.1401 0.1399 0.1396 0.1394 0.1392 0.1389 0.1387

Columns 1178 through 1188

0.1385 0.1382 0.1380 0.1378 0.1375 0.1373 0.1371 0.1368 0.1366 0.1364 0.1362

Columns 1189 through 1199

0.1359 0.1357 0.1355 0.1352 0.1350 0.1348 0.1346 0.1343 0.1341 0.1339 0.1337

Columns 1200 through 1210

0.1334 0.1332 0.1330 0.1328 0.1325 0.1323 0.1321 0.1319 0.1317 0.1314 0.1312

Columns 1211 through 1220

0.1310 0.1308 0.1306 0.1303 0.1301 0.1299 0.1297 0.1295 0.1292 0.1290

(2) $p'_l =$

Columns 1 through 11

11.3595 11.4190 11.5985 11.6780 11.7875 11.6971 11.7166 11.8760 11.7456 12.0150 11.9045

Columns 12 through 22

11.9440 11.9435 11.7931 11.8825 11.8520 11.8815 11.9010 11.9105 11.7401 11.9395 11.9789

Columns 23 through 33

11.9984 11.9380 11.9175 11.9270 11.8366 11.8461 11.9754 11.9350 11.9445 11.9939 12.1332

Columns 34 through 44

12.2725	12.3319	12.2315	12.1411	12.3203	12.1801	11.9999	11.9594	12.0088	12.0083	11.9579
Columns 45 through 55										
12.0572	11.8371	11.8965	11.9759	11.9255	11.8651	11.7449	11.6845	11.6740	11.6835	11.6531
Columns 56 through 66										
11.4730	11.2830	11.2626	11.2920	11.2816	11.1814	11.2208	11.2702	11.4393	11.3989	11.3685
Columns 67 through 77										
11.3680	11.4872	11.4668	11.5560	11.4359	11.4554	11.5247	11.5042	11.5337	11.7724	11.7420
Columns 78 through 88										
11.7017	11.5915	11.5312	11.2916	11.2114	11.2109	11.2005	11.1402	11.0501	11.1492	11.0392
Columns 89 through 99										
11.0487	11.0980	11.0677	11.0373	10.9970	11.0962	11.1156	11.0753	10.9653	10.9748	10.7851
Columns 100 through 110										
10.8245	10.8241	10.5946	10.5643	10.6833	10.6331	10.6625	10.6421	10.6815	10.7109	10.6607
Columns 111 through 121										
10.7399	10.5702	10.5300	10.4798	10.4694	10.4092	10.4884	10.4979	10.5671	10.5169	10.5364
Columns 122 through 132										
10.5558	10.2669	10.3858	10.3954	10.3452	10.3746	10.3244	10.3339	10.3434	10.4226	10.4023
Columns 133 through 143										
10.2825	10.2622	10.2717	10.2116	10.3006	10.2505	10.2600	10.3192	10.3188	10.4476	10.4472
Columns 144 through 154										
10.4368	10.4463	10.4558	10.4156	10.3953	10.3948	10.3149	10.3145	10.3737	10.3136	10.3330
Columns 155 through 165										
10.3227	10.2726	10.2423	10.1525	10.3607	10.1914	10.2605	10.1905	10.1901	10.1599	10.1495
Columns 166 through 176										
10.1590	10.4069	10.3965	10.3067	10.3261	10.2860	10.3550	10.2851	10.1556	10.1552	9.9761
Columns 177 through 187										
10.0253	9.9653	10.0145	9.8553	10.0832	10.1522	10.0823	10.2109	10.2303	10.1902	10.2295
Columns 188 through 198										
10.2389	10.2484	10.3671	10.3865	10.3265	10.2269	10.2165	10.2161	10.1859	10.0863	10.0661
Columns 199 through 209										
10.0954	10.0255	10.0945	10.1437	10.1532	10.1428	10.1821	10.0726	10.0325	10.1015	9.9623
Columns 210 through 220										
10.0015	9.9416	9.8917	9.8813	9.8611	9.9201	9.8206	9.8400	9.8198	9.8788	9.8388
Columns 221 through 231										
9.8087	9.7785	9.7781	9.8966	10.0249	10.0047	10.0538	10.0732	10.0233	10.0724	10.0125
Columns 232 through 242										
10.2696	10.2791	10.2786	10.2683	10.2679	10.1684	10.1482	10.3458	10.3354	10.2855	10.3445
Columns 243 through 253										
10.4232	10.4228	10.1254	10.1250	10.0850	10.0846	10.0643	9.9650	10.0932	10.0928	10.0330
Columns 254 through 264										
9.9633	10.0618	10.1801	10.0709	10.1001	10.0997	10.0597	10.1384	10.1380	10.4442	10.3448
Columns 265 through 275										
10.3839	10.4626	10.4127	10.3530	10.3427	10.4213	10.5000	10.3710	10.3903	10.1131	10.0336
Columns 276 through 286										

10.0628	9.9240	10.0027	10.1110	10.0611	9.8730	9.8923	9.9611	9.8816	9.9108	9.9104
Columns 287 through 297										
9.9989	9.8898	9.8993	9.9581	9.8688	9.9178	9.9865	9.9663	9.9165	9.9655	9.8565
Columns 298 through 308										
9.8462	9.7174	9.6972	9.6968	9.7359	9.7256	9.6067	9.6063	9.6454	9.6253	9.6249
Columns 309 through 319										
9.5060	9.5155	9.5052	9.4555	9.5439	9.5534	9.4444	9.4933	9.5028	9.4531	9.3935
Columns 320 through 330										
9.4720	9.5210	9.5107	9.5399	9.5099	9.5292	9.4696	9.4101	9.4195	9.4882	9.5765
Columns 331 through 341										
9.6156	9.5856	9.5556	9.4862	9.5154	9.5248	9.5836	9.5536	9.5138	9.4937	9.5031
Columns 342 through 352										
9.4633	9.4826	9.3245	9.2255	9.0280	9.0572	9.0568	9.0368	9.0364	8.9177	8.7696
Columns 353 through 363										
8.3849	8.2269	8.1379	8.1967	8.4722	8.5605	8.7374	8.9242	8.9436	8.9727	8.9822
Columns 364 through 374										
9.0114	9.0011	9.0599	9.0299	9.0099	8.8913	8.8122	8.7724	8.9197	8.9883	9.0273
Columns 375 through 385										
8.9777	8.9084	8.8490	8.7600	8.9565	8.9856	8.9853	8.9652	8.9648	8.9546	8.8657
Columns 386 through 396										
8.8653	8.8748	9.0515	9.0118	9.0114	8.9225	8.9319	8.9512	8.8427	8.8718	8.8419
Columns 397 through 407										
8.7924	8.8805	8.8605	8.8306	8.8007	8.7020	8.6918	8.7406	8.7993	8.7596	8.7592
Columns 408 through 418										
8.9456	9.0828	9.1513	9.0133	8.9736	8.9634	8.8844	8.7857	8.8345	8.8145	8.7846
Columns 419 through 429										
8.7057	8.7053	8.7049	8.7733	8.7337	8.6842	8.7428	8.7424	8.7420	8.6631	8.6627
Columns 430 through 440										
8.6624	8.7013	8.6518	8.5532	8.4351	8.3365	8.3754	8.4438	8.6300	8.6984	8.7176
Columns 441 through 451										
8.6878	8.6580	8.7754	8.6867	8.6078	8.5682	8.6169	8.4792	8.4690	8.5373	8.6940
Columns 452 through 462										
8.7231	8.7619	8.8989	8.9770	8.8491	8.7801	8.7405	8.7990	8.7692	9.1611	9.0529
Columns 463 through 473										
9.0917	9.0325	9.0615	9.0415	8.9725	8.8741	8.8737	8.8047	8.6475	8.5883	8.5977
Columns 474 through 484										
8.6072	8.5186	8.5084	8.5179	8.5763	8.5270	8.5462	8.5262	8.5357	8.5255	8.5252
Columns 485 through 495										
8.5248	8.5245	8.5241	8.5237	8.2393	8.1997	8.3463	8.3068	8.1987	8.1494	8.0805
Columns 496 through 506										
8.0997	8.1777	8.1382	8.0889	8.3138	8.1862	8.1858	8.2540	8.1753	8.1750	8.0670
Columns 507 through 517										
8.1156	8.0957	8.0855	8.0558	8.1142	8.1432	8.1135	8.0936	8.1129	8.1517	8.1122
Columns 518 through 528										

8.1803	8.5714	8.2384	8.0521	8.0518	8.0906	8.2957	8.2464	8.3048	8.2555	8.1378
Columns 529 through 539										
8.2255	8.3425	8.3617	8.2831	8.1067	8.1260	8.0767	8.1448	8.1152	8.1246	8.1047
Columns 540 through 550										
8.2217	8.2507	8.2308	8.1718	8.1030	8.1124	8.1414	8.1411	8.1505	8.1208	8.1303
Columns 551 through 561										
8.1983	8.2468	8.2758	8.2364	8.2360	8.1673	8.1670	8.1862	8.1467	8.0976	8.1363
Columns 562 through 572										
8.1359	8.1258	8.0669	8.0080	8.0369	8.0464	8.0363	8.0262	8.0063	8.0157	8.0447
Columns 573 through 583										
8.1224	8.1807	8.1901	8.1214	8.1601	8.1207	8.1594	8.2079	8.1685	8.1779	8.1776
Columns 584 through 594										
8.1480	8.1574	8.1570	8.1274	8.1564	8.1755	8.0191	7.9895	7.9989	8.0278	8.0275
Columns 595 through 605										
8.0174	8.0658	8.1435	8.1042	8.0941	8.0352	8.0836	8.0346	8.0440	8.0436	8.0433
Columns 606 through 616										
8.0430	8.0524	8.0228	8.0419	8.0611	8.0510	8.0507	8.0406	7.9915	7.9912	8.0103
Columns 617 through 627										
8.0295	8.0584	8.0873	8.0480	8.0574	8.0083	8.0372	8.0564	8.1632	8.2213	8.3768
Columns 628 through 638										
8.3277	8.5806	8.5608	8.5020	8.5893	8.5694	8.5204	8.3837	8.4223	8.5193	8.4995
Columns 639 through 649										
8.4699	8.3917	8.4692	8.4786	8.5561	8.4779	8.4094	8.4285	8.4574	8.4181	8.4372
Columns 650 through 660										
8.4563	8.4170	8.4361	8.4747	8.5814	8.6296	8.6585	8.6289	8.5605	8.5504	8.5889
Columns 661 through 671										
8.5983	8.5590	8.5781	8.5583	8.4315	8.2561	8.2655	8.3527	8.3329	8.3034	8.3127
Columns 672 through 682										
8.3027	8.3218	8.3506	8.3016	8.2527	8.2426	8.2520	8.2322	8.1930	8.2023	8.1826
Columns 683 through 693										
8.1725	8.0750	8.0747	8.0549	8.1323	8.1902	8.2287	8.3450	8.3252	8.2860	8.2759
Columns 694 through 704										
8.2950	8.2461	8.3623	8.4299	8.4004	8.4292	8.3318	8.3411	8.2340	8.2045	8.2041
Columns 705 through 715										
8.2426	8.3200	8.3487	8.4940	8.5033	8.6874	8.6870	8.9002	8.8998	9.7923	10.7624
Columns 716 through 726										
10.9754	10.1016	10.1012	9.8776	9.9451	9.8186	10.1674	10.3998	10.4091	10.7191	11.0096
Columns 727 through 737										
10.3981	10.0776	10.4554	10.7556	10.9782	11.1038	12.2573	12.0532	12.7702	12.5854	12.4395
Columns 738 through 748										
11.9736	11.9925	11.6527	12.6604	12.9797	11.7869	12.1838	13.0750	13.0163	13.2096	13.4804
Columns 749 through 759										
12.9856	12.6168	12.3353	11.9859	11.5591	12.0431	11.7519	11.4802	11.2181	11.3533	11.5659
Columns 760 through 770										

11.4879	11.5068	11.5063	11.3993	11.4085	11.5630	11.6690	11.7557	11.6487	11.2512	11.2410
Columns 771 through 781										
11.0179	11.0852	11.2106	11.0843	11.0644	11.3447	11.3829	11.5276	11.8078	11.9816	11.9714
Columns 782 through 792										
12.0193	12.1349	11.9215	11.6501	12.3269	12.0361	12.2194	12.0641	12.1990	12.2469	12.1787
Columns 793 through 803										
12.6134	12.8644	12.7381	12.8053	13.0465	12.8622	12.8230	13.3737	13.7502	13.1598	13.3526
Columns 804 through 814										
13.5841	13.7092	13.6893	15.0518	14.7225	14.0839	13.7547	13.8701	13.3669	12.9508	12.5057
Columns 815 through 825										
12.7468	13.0071	13.0356	12.7645	12.8123	12.4349	12.1059	12.5208	12.5010	12.5487	12.9829
Columns 826 through 836										
13.3301	13.3972	13.7057	12.5268	12.2365	12.7285	12.6990	12.6792	12.9394	13.1609	13.6818
Columns 837 through 847										
13.5267	13.3234	13.2456	13.3030	12.9259	12.4041	12.5967	12.1039	11.4567	11.7844	12.1216
Columns 848 through 858										
11.6097	10.7986	11.2323	11.6178	11.0191	11.9835	12.1277	13.3428	13.9018	12.6471	13.1965
Columns 859 through 869										
13.4178	12.8481	12.5968	13.8598	13.7338	13.5018	14.0702	13.8189	13.9726	14.0009	14.0196
Columns 870 through 880										
12.6788	12.1480	11.9354	12.0024	11.3656	10.8350	11.0852	10.8919	10.7372	10.9295	11.3435
Columns 881 through 891										
11.4201	11.2558	11.2746	11.2452	11.0713	11.3985	11.0511	10.5690	10.2410	9.7396	8.7662
Columns 892 through 902										
8.7851	8.9870	9.3623	9.1500	9.4289	8.9566	8.8311	8.9077	9.2251	9.0514	9.0125
Columns 903 through 913										
8.7040	8.6748	8.9632	8.7703	8.6159	8.6637	8.7500	8.5571	8.5952	8.7296	8.6234
Columns 914 through 924										
8.4498	8.4206	8.6319	8.7182	8.9680	8.8810	8.9287	9.0246	9.1493	9.1008	9.1677
Columns 925 through 935										
9.1193	9.0227	9.0319	9.0893	9.0023	8.9827	9.0208	8.9435	8.7219	8.6543	8.9327
Columns 936 through 946										
9.1824	9.2973	9.4412	9.4600	9.4019	9.3054	9.1320	9.1797	9.2369	9.2462	9.2362
Columns 947 through 957										
9.1877	9.1105	9.1005	9.1770	9.1574	8.6381	8.5801	8.5894	8.8196	9.0113	8.7612
Columns 958 through 968										
8.7320	8.5587	8.5296	8.5004	8.4712	8.6053	8.5473	8.6334	8.6523	8.6519	8.7091
Columns 969 through 979										
8.6992	8.7180	8.6504	8.6501	8.5441	8.5918	8.5818	8.5238	8.0436	7.9856	8.2252
Columns 980 through 990										
7.7450	7.9079	7.6196	7.6769	7.5326	7.5515	7.2633	7.1959	7.3395	7.3775	7.1854
Columns 991 through 1001										
7.2522	7.3478	7.0022	7.2033	7.3565	7.3754	7.1641	7.2405	7.1347	7.1824	7.1437
Columns 1002 through 1012										

7.0955	7.2677	7.4304	7.3438	7.3339	7.4391	7.3908	7.4385	6.9972	7.1790	7.1979
Columns 1013 through 1023										
7.2359	7.4465	7.4941	7.7238	7.6851	7.7231	7.7036	7.5979	7.5976	7.6548	7.6928
Columns 1024 through 1034										
7.7882	7.7783	7.7780	7.7489	7.6911	7.6812	7.5373	7.5369	7.4983	7.3831	7.5743
Columns 1035 through 1045										
7.5165	7.5641	7.5638	7.6017	7.5153	7.4192	7.5434	7.5526	7.6863	7.6668	7.6474
Columns 1046 through 1056										
7.4939	7.5510	7.4933	7.4643	7.4831	7.4541	7.4346	7.4247	7.4244	7.3189	7.4047
Columns 1057 through 1067										
7.3565	7.3562	7.2411	7.0973	7.0875	7.1159	7.0965	7.0866	7.0959	7.1338	7.0953
Columns 1068 through 1078										
7.0758	7.1042	7.0944	7.1132	7.0842	7.1126	7.0932	7.1216	7.2646	7.2165	7.2258
Columns 1079 through 1089										
7.2542	7.2348	7.2249	7.1959	7.0810	7.0807	7.1090	7.0609	7.0893	7.0986	7.0792
Columns 1090 through 1100										
7.1075	7.0690	7.0114	7.0302	7.0777	7.1060	7.1153	7.1054	7.1815	7.1908	7.1618
Columns 1101 through 1111										
7.1806	7.1326	7.1705	7.2561	7.2940	7.2555	7.2648	7.2644	7.2355	7.2066	7.2254
Columns 1112 through 1122										
7.1773	7.1675	7.2149	7.1096	7.0998	7.0995	7.0992	7.1180	7.0986	7.1174	7.1171
Columns 1123 through 1133										
7.1454	7.1642	7.1543	7.1827	7.2205	7.3251	7.2962	7.2959	7.2670	7.2857	7.2950
Columns 1134 through 1144										
7.2947	7.3039	7.2845	7.2747	7.2934	7.3599	7.4072	7.3593	7.3399	7.3682	7.3583
Columns 1145 through 1155										
7.3580	7.3291	7.2907	7.1856	7.1662	7.1564	7.1561	7.1462	7.1364	7.1361	7.1453
Columns 1156 through 1166										
7.0593	7.0781	7.0587	7.0965	7.0867	7.1626	7.1813	7.1715	7.1522	7.1709	7.1325
Columns 1167 through 1177										
7.1989	7.1986	7.2078	7.2265	7.2929	7.2545	7.2447	7.2348	7.2441	7.2057	7.1959
Columns 1178 through 1188										
7.1480	7.1572	7.1283	7.1566	7.1848	7.1465	7.2033	7.2601	7.2978	7.2785	7.2877
Columns 1189 through 1199										
7.2779	7.2585	7.3343	7.5243	7.4479	7.4000	7.4567	7.5230	7.6178	7.4748	7.6076
Columns 1200 through 1210										
7.7024	7.4929	7.4546	7.4733	7.4349	7.5012	7.4153	7.5576	7.5953	7.4524	7.4616
Columns 1211 through 1220										
7.5468	7.4799	7.7743	7.9925	7.9542	7.9348	7.8585	7.8676	7.7913	7.7910	

致谢

时光荏苒，岁月如梭。转眼间，我在沈阳师范大学的三年研究生的学习生涯即将结束。首先衷心感谢我的导师程丛电教授，感谢程老师在我的三年研究生学习生涯中对我的谆谆教诲，不仅教授我各种统计专业的知识，让我从懵懵懂懂到渐渐加深了解，而且努力培养我的自主学习能力和不断探究的精神。作为我的硕士研究生导师程老师，从我的毕业论文选题、开题报告、写作、反复修改和最后定稿过程中都投入了大量的心血和精力，程老师在百忙之中给予我细致而深刻的指导，帮助我理顺写作思路并给出了诸多建议。正是由于程老师自始至终对我无微不至的关心、指导、帮助和鼓励，使得我的毕业论文能够顺利的完成。程老师拥有渊博的知识而且秉承严肃的治学态度，对学生认真负责，又不乏宽容，一直感染和激励着我。在此向程老师至于我最真挚的感谢：祝福程老师工作顺利，身体健康，一帆风顺，万事如意。

其次感谢杨光老师，付志慧老师，刘瑞银老师，温学兵老师，王立柱师，李巍老师以及所有在我的研究生生涯中给予我帮助的老师们。感谢研究生期间的辅导员孟宪涛老师，给予我学习上和生活上的指导。感谢梁晓同，马威等同学对我学习和生活上的帮助。感谢母校沈阳师范大学，感谢数学与系统科学学院统计学系，让我在知识的天空里尽情的翱翔。

最后感谢我的家人，是他们给予我无微不至的关怀与支持，使我获得不断向前的动力，获得面对挫折的勇气。

总之，向各位良师益友，在此衷心的表示：“谢谢”。

马晶鑫

2017年3月于沈阳师范大学

个人简历及攻读硕士学位期间已发表论文

姓名：马晶鑫

性别：女

出生日期：1992 年 8 月 6 日

民族：汉

籍贯：辽宁省沈阳市

学士学位授予单位：沈阳师范大学

获学士学位时间：2014 年 6 月 27 日

学习经历：

2010 月-2014 年 7 月 沈阳师范大学数学与系统科学学院统计学

2014 年 9 月-至今 沈阳师范大学数学与系统科学学院统计学专业研究生

攻读硕士学位期间已发表的论文

1. 程丛电, 马晶鑫. 基于一种炒股行为生成的随机变量与分布估计[J]. 沈阳师范大学 (自然科学版), 2017(2).
2. Cheng Chongdian, Ma Jingxin. Approximation algorithm for extended maximum concurrent flow problem with saturated capacity[C]//LEMCS, 2015, 632-636.
3. Cheng Chongdian, Ma Jingxin. Maximum multicommodity flow problem with local requirement[C]//IHMSC08, 2016:433-437.