

摘要

金融市场受多重因素影响，具有高度复杂性和不确定性，是一个复杂的动态变化系统。对股市的历史数据研究分析遇到了越来越多的挑战。诸如数据噪声，过去的历史数据不能准确反映未来的趋势等问题。

基于主体（Agent）的建模方法，首先对微观个体进行建模，然后通过计算机模拟实验，模拟仿真股票市场交易的过程，进而在宏观上观察系统的动态演变。基于主体的人工股票市场提供了一个新的研究视角，使得我们可以研究股票市场的微观结构和宏观现象。

本文将使用该方法，利用Python语言构建一个人工股票市场，通过该模型模拟真实的股票市场，并在此基础上进行分析。本文的人工股市模型中通过面向对象的思想设计了三类异质的交易者主体。其中父类是BaseInvestor（噪音投资者），再根据预测股价的方式不同实现了两个子类：ValueInvestor（机构投资者）和TechnologyInvestor（技术投资者）。采用连续竞价的形式，通过双向拍卖进的机制而使得买卖双方达成交易，并在每个交易日结束时，更新交易者的股票持有数量和现金余额，记录每个交易日交易者的财富变化。

首先为了证明模型的可靠性、可行性。对模型进行描述性输出验证^[1]。将计算生成的输出与实际金融系统的数据进行比较，并对比其与实际金融数据是否具有相同的统计特征。通过多次重复实验，发现上证指数、深成指数、恒生指数、标普500，这四支股票指数的收益率分布服从尖峰、厚尾型的非正态分布，并且有明显的波动聚集性和长期记忆性的特征。这种统计特征在本文构建的人工股市模型中同样存在，这也说明模型具备反映真实股票市场的某些特征的能力。

实验发现当噪声交易者预测的噪声标准差较大（为5）时，也就是当市场交易者中出现一些投资者盲目买入卖出时，市场会产生剧烈波动，趋势反转者也会取得优势地位。也会使股票价格时常偏离其真实价值。

对交易主体财富分布进行研究发现：随着交易的进行，发现财富逐渐集中，表现为基尼系数逐渐增大，直到达到某一值后，平稳波动。也就是高财富人群的财富占总财富的比例逐渐增大。这种财富逐渐集中的现象，无论股价是下跌还是上涨都会持续发生。这种现象也不难解释：一方面不同投资者的风险偏好和购买的股票数量不一样，在股票上涨阶段，增加的财富也不一样。另一方面，基尼系数只是代表“分蛋糕”平均程度的一个指标，“分蛋糕”的大小并不能反映出来。

实验还发现无论是单一股票市场还是多股票市场，适当增加机构投资者比例能够在一定程度缓解股市财富集中问题，表现为基尼系数和泰尔指数都随着机构投资者数量增加而减小，当机构投资者增加到100以后，基尼系数和泰尔指数表现为数不再下降，而是平稳波动。另外当机构投资者较少时（数量为30），增加技术投资者或噪音投资者会加剧财富分布的不平等。而当机构投资者较多时（数量为100），增加技术投资者或噪音投资者没有明显加剧财富分布不平等。根据中国银行研究院统计，截至2019年末，我国境内机构投资者持有A股流通市值占比仅为11.6%，国际对比来看，我国机构投资者持股市值占比处于低位。这也从事实的角度说明增加机构投资者比例是可取的。

关键词：基于主体的建模，人工股票市场，连续竞价，财富不平等

ABSTRACT

Affected by multiple factors, the financial market is highly complex and uncertain, and is a complex dynamic system. Through the research and analysis of the historical data of the stock market, more and more challenges have been encountered. Issues such as data noise, past data not accurately reflecting future trends, etc.

Through an agent-based modeling approach, First, the micro-individual is modeled, Then through computer simulation experiments, simulate the process of simulating stock market transactions, Then observe the dynamic evolution of the system at a macro level. Agent-based artificial stock market provides a new research perspective, It allows us to study the microstructure of the stock market and the emergence of macroscopic phenomena.

Affected by multiple factors, the financial market is highly complex and uncertain, and is a complex dynamic system. The research and analysis of historical data of the stock market has encountered more and more challenges. Issues such as data noise, past data not accurately reflecting future trends, etc.

Agent-based modeling method, First, the micro-individual is modeled, Then through computer simulation experiments, simulate the process of simulating stock market transactions, Then observe the dynamic evolution of the system at a macro level. The agent-based artificial stock market provides a new research perspective, It allows us to study the micro structure and macroscopic phenomena of the stock market.

This article will use this method to build an artificial stock market using Python language, and simulate the real stock market through this model. and analyze on this basis. In the artificial stock market model in this paper, three types of heterogeneous trader subjects are designed through object-oriented thinking. The parent class is BaseInvestor (noise investor), and two subclasses are realized according to the way of predicting stock prices: ValueInvestor

(institutional investors) and Technology Investor (technology investors). In the form of continuous bidding, the buyer and the seller reach a deal through a two-way auction mechanism. And at the end of each trading day, update traders' stock holdings and cash balances, Record changes in the wealth of traders for each trading day.

Firstly, to prove the reliability and feasibility of the model. Descriptive output validation of the model. Comparing the output generated by the calculation with the data of the actual financial system, And compare whether it has the same statistical characteristics with the actual financial data. Through repeated experiments, it was found that Shanghai Composite Index, Shenzhen Component Index, Hang Seng Index, S&P 500, The return distributions of these four stock indexes obey the non-normal distribution with peaks and thick tails. And it has the characteristics of obvious volatility aggregation and long-term memory. This statistical feature also exists in the artificial stock market model constructed in this paper, which also shows that the model has the ability to reflect some characteristics of the real stock market.

The experiment found that when the noise standard deviation predicted by noise traders is large (5), That is, when some investors blindly buy and sell among market traders, There will be wild swings in the market, and trend reversals will gain the upper hand. It will also cause the stock price to deviate from its true value from time to time.

A study on the wealth distribution of transaction subjects found that: as the transaction proceeds, Find wealth gradually concentrated, The performance is that the Gini coefficient gradually increases until it reaches a certain value and then fluctuates steadily. That is to say, the proportion of the wealth of the high-wealth group to the total wealth is gradually increasing. This gradual concentration of wealth will continue regardless of whether the stock price is falling or rising. This phenomenon is not difficult to explain: on the one hand, the risk preferences of different investors and The number of stocks purchased is not the same, and the increase in wealth is also different in the stage of stock rise. On the other hand, the Gini coefficient is only an indicator representing the average degree of the "partial pie", and the size of the "partial pie" does not reflect it.

The experiment also found that whether it is a single stock market or a multi-

stock market, Appropriately increasing the proportion of institutional investors can alleviate the problem of wealth concentration in the stock market to a certain extent, as shown by the Gini coefficient and Theil index. The number of institutional investors increases but decreases. When the number of institutional investors increases to 100, the Gini coefficient and Theil index no longer decline, but fluctuate steadily. In addition, when there are few institutional investors (the number is 30), adding technology investors or noise investors will exacerbate the inequality of wealth distribution. And when there are many institutional investors (the number is 100), adding technology investors or noise investors does not significantly increase the inequality of wealth distribution. According to the statistics of Bank of China Research Institute, As of the end of 2019, the market capitalization of A shares held by domestic institutional investors was only 11.6%. From an international comparison, the share of market value held by institutional investors in my country is at a low level. This also shows that increasing the proportion of institutional investors is desirable from a factual point of view.

Keywords: Agent-based modeling, artificial stock markets, continuous auctions, wealth inequality

目 录

1.绪 论	1
1.1背景与意义	1
1.1.1研究背景	1
1.1.2研究意义	2
1.2国内外研究历史与现状	4
1.2.1人工股市综述	4
1.2.2综述总结	5
1.2.3基于主体的仿真工具研究综述	6
1.3本文的主要贡献与创新	7
1.4本论文的结构安排	7
1.5本章小结	8
2.基础理论介绍	9
2.1金融股票市场的复杂性	9
2.2AGENT相关理论	10
2.2.1多Agent系统模型概述	11
2.2.2多Agent的适用领域	14
2.2.3小结	14
2.3投资者主体	15
2.3.1基本面投资者	15
2.3.2技术分析投资者	15
2.3.3噪音投资者	16
2.3.4小结	16
2.4人工股市模型的校准	16
2.4.1尖峰厚尾分布的检验	17

2.4.2波动聚集性的检验	18
2.4.3长期记忆性的检验	18
2.4.4小结	19
2.5竞价策略	20
2.5.1集合竞价	20
2.5.2连续竞价	20
2.5.3连续竞价原则	20
2.5.4小结	21
2.6财富的集中度	21
2.6.1洛伦兹曲线	21
2.6.2基尼系数	22
2.6.3泰尔指数	22
2.6.4小结	23
2.7本章小结	23
3.国内外股票指数收益率统计特征校验	24
3.1股票选取	24
3.2正态性检验	24
3.3波动聚集性检验	26
3.4长记忆性检验	27
3.5本章小结	28
4.人工股市的搭建	29
4.1模型总体设计	29
4.2人工股市基本市场框架	29
4.3投资者的预测行为	30
4.4投资者的买卖报价及数量	31
4.5交易撮合	31
4.5.1多股票引入的问题	31
4.5.2预扣款冻结机制	32
4.5.3撮合的数据结构	33
4.5.4撮合流程	35

4.6涨停板设置	39
4.7本章小结	39
5.人工股市实验分析	40
5.1实验准备	40
5.1.1仿真实验的参数设定	40
5.2人工股市的检验	41
5.2.1正态性检验	41
5.2.2波动聚集性检验	41
5.2.3长期记忆性检验	42
5.3财富不平等演变	43
5.3.1市场剧烈波动时	43
5.3.2财富分布演变方向	44
5.3.3机构投资者越多有利于公平	46
5.4本章小结	53
6.全文总结与展望	54
6.1全文总结	54
6.2后续工作展望	56
6.3本章小结	56
参考文献	57
致 谢	62

1.绪 论

1.1 背景与意义

1.1.1 研究背景

传统的经济现象分析方法往往注重整体的总结，简化了经济现实，忽视了人为因素，极少关注到个体和环境相互作用。因此尽管可以获得十分严密的逻辑推理，却无法准确反映实际情况，解决实际问题。一些研究人员尝试借鉴自然科学的科学方法。例如实验、观察和模拟来证明假设。实验经济学通过对经济环境的模拟，以及对人类行为的实验研究^[2]，从而试图理解经济现象和经济决策。借助实验模拟现实经济问题，为经济理论研究提供了新的方向。以实验为导向的研究方法，在一定程度上克服了传统数学模型对经济问题研究的局限性。然而，经济活动常常需要多个经济个体参与，由于传统的实验方法的实验规模受限和数据管理不易等局限性，难以建立符合研究规模的实验来模拟现实从而进行研究分析^[3]。

正是因为计算机具有强大的处理能力和模拟能力，实验经济学才得以发挥其巨大的潜力，并得到了快速的发展。随着复杂系统和人工智能研究的发展，Agent（主体）技术开始应用于计算机仿真，使得计算机不再是只能进行简单计算而是具有智能性。这正是基于此，使得基于主体的实验能够很好地模拟经济活动，应用于经济问题的研究。至此研究经济学问题也如同自然科学问题一样，可以通过实验和模拟来研究真实的经济问题。这种研究方法更注重经济问题的动态特征和经济个体的相互作用。

基于Agent的计算经济学的一个重要方法是应用复杂自适应系统（Complex Adaptive System, CAS）的理论来研究经济系统中的问题。复杂自

适应系统是一种基于Agent的模型，它以人类活动为基础，考虑到社会和客观环境的复杂性，以及系统中参与者的行为。它的目的是模拟经济系统的演化，帮助人们理解经济系统的复杂性。

通过将经济系统看作是由若干相互作用的个体进行复杂交互的复杂系统，自适应系统的方法突破了传统的经济学的框架。它最根本的思想是放弃了完全理性、全知全能的“经济人”的假设。在微观层面，用“规则”代替计算，对每个个体进行建模。这些规则可以是简单的认知法则，也可以是复杂的行为模式，而个体差异在于不同的规则和规则的选择。而在宏观层面上，通过摒弃简单的还原论思想，将经济系统看成是由若干相互作用的个体进行复杂交互的复杂系统。最后，在方法论上，使用计算机模拟技术代替纯数学推理。

复杂适应系统理论认为经济主体的不确定性、有限性和差异性现实社会中经济系统的核心特征。该理论提出，经济主体是一个复杂的自适应系统，由许多相互关联的个体组成，他们的行为是基于信息和观察到的环境的不确定性和变化的结果。因此，分析经济问题的过程可以通过模拟经济主体的行为和关系，从而预测整个经济系统的发展情况。

美国圣达菲研究所开发的人工股票市场ASM(Artificial Stock Market)是最著名的利用CAS思想研究经济现象的实验模型。ASM可以模拟股票交易来仿真股票市场的行为，模拟市场决策者的行为来研究市场机制^{[4][5][6][7]}。人工股票市场可以看作是一个真实股票的数字化克隆，所以通过很低的成本就可以完成一系列在真实世界中代价很大的实验^[8]。比如考察不同类型投资者占比份额对股票市场价格波动和财富分布变化的情况。仅仅利用历史市场数据是很难通过调整不同交易者类型占比情况对财富分配的影响进行实证研究，而基于人工股票市场的研究方法对于制定政策、评估和预测政策市场效果起到了很大作用^[9]。

1.1.2 研究意义

在当今世界，贫富之间的差距正在扩大。财富日益集中的现象对全球经济产生了深远的影响，也是许多问题的根源。根据西南财经大学中国家庭金

融调查与研究中心的2014年《中国家庭财富的分布及高净值家庭财富报告》显示，中国的财富分布出现严重不平等。财富排名前1%的家庭占有社会总体财富的三分之一左右，说明财产集中在少数高财富家庭^{[10][11]}。数十年来，日益扩大的财富差距一直是一个值得讨论的话题，并且仍然是一个尚未解决的问题。富人和弱势群体之间日益扩大的差距已经导致世界各地的贫困、不平等和经济不稳定的情况增加。随着这种差距继续扩大，必须研究财富日益集中的原因和影响，以确定解决这一紧迫问题的最佳方法。

股票市场被视为提高资本配置效率的工具之一。在股票市场上，各种信息可以通过价格反映，股票价格的变动反映公司的经营状况、行业趋势、宏观经济环境等因素，从而引导投资者做出投资决策。这种价格发现机制有助于优化资源配置，促进经济发展。同时，股票市场还可以提供各种融资机会，帮助公司筹集资金，扩大规模，提高效率，进一步促进经济发展。因此，股票市场作为资本市场的重要组成部分，可以说是提高资本配置效率的工具之一。

虽然股票市场在过去确实注重市场效率，但也没有忽略公平原则。虽然诸如内幕交易、操纵市场等不公平行为也时有发生，但是市场监管机构也一直在努力维护市场的公平和透明。同时，现代股票市场也越来越注重公平。例如采取了诸如信息披露、监管规则、公平交易平台等措施来保证市场的公平性。

股票市场作为金融体系的一个关键组成部分，撮合投资者的需求和行业的资源，以满足社会和投资者的利益需求。随着股票市场的发展，市场资源的分配效率日益提高，但是在有关资源配置公平性方面却相对缺乏研究。

对于多主体建模在股票市场的研究，关注资源配置公平性的研究意义重大，因为它可以直接影响到投资者的收益和社会整体财富分配。考虑到股票市场的复杂性和多主体之间的博弈，公平性是一个十分重要的考量因素，值得进行深入研究和探讨。

多主体建模还可以帮助研究者探究股票市场中不同类型的参与者之间是否存在不平等的资源分配，以及在股票市场中如何实现公平的资源配置。总

之，多主体建模为股票市场的研究带来新的视角，有助于更好地理解市场的运行机制。

本论文研究的目的是，通过改进ASM模型，运用多Agent的理论和方法，构建基于多主体的人工股票市场，来模拟研究交易者策略行为^[12]，并通过分析不同类型交易者占比对交易者整体财富演变的影响。通过合理调整投资者的组成，促进财富分配更加公平。对促进财富更加公平地分配提供了理论指导和实验支撑，因此本研究从现实层面和理论层面都有很大的研究价值。

1.2 国内外研究历史与现状

1.2.1 人工股市综述

基于Agent的仿真模型研究始于上世纪90年代，其中最著名就是圣塔菲研究所的人工股票市场模型（SFI-ASM）^[13]。SFI-ASM的建立为研究金融市场的动态过程和模型提供了一个新的研究方向，特别是在行为金融学和复杂系统方面的研究。它向研究人员提供了一种更加真实的方式模拟金融市场，并且由于它是由一群智能体互相作用的系统，因此能够更好地模拟市场中的多样性和复杂性^[14]。

Lebaron（2001）^[15]通过构建人工金融市场，得到了一些真实股票市场中价格时间序列的统计特征。例如，收益率并不是符合正态分布，而是有尖峰、厚尾分布特征，表明股票市场的结构与传统理论的预期不尽相同。

Chiarella（2002）^[16]建立了一个基于指令驱动连续双向拍卖市场。交易者根据固定规则设置出价和要价并发布市场或限价订单。还研究不同的交易策略如何影响价格、买卖差价、交易量和波动性的动态。也分析了市场设计的一些特征。例如刻度大小和订单生命周期是如何影响市场流动性的。

高宝俊（2006）^[17]建立了一个连续竞价市场的人工股市模型，将投资者分成基本面分析交易者和技术分析交易者两类，进而研究了不同类型的交易者对股市价格波动、交易量等的影响等问题。

Lebaron (2007)^[18]在收益率时间序列统计特征基础上,进一步考虑交易者之间的相互学习和模仿,并得出股票市场存在的长期记忆性的特征。

崔宾阁(2012)^[19]提出了一个具有财务信息的新型股票市场概念模型。

Arifovic (2013)^[20]利用人工股市建模,研究了市场透明度和主体行为对市场效率的影响,并证实了如果交易者能够获取和利用其他投资者的行动信息,并进行学习,将会提高市场分配效率。

Ahn (2018)^[21]研究表明初始参数值的微小变化会导致金融市场出现相对较大的波动,可以将其视为复杂或混乱的系统。这也意味着金融市场由于描述市场的模型参数的微小意外变化而变得不稳定。

Ivan Jericevicha (2021)^[22]验证了使用事件驱动、异步撮合引擎的可行性。以消息流的形式连续提供订单确认和更新,以便更接近实时交易环境。

周融天(2022)^[23]通过利用人工股票市场模拟方法对A股融券费率改革方案进行测试,研究发现,费率设计不够灵活是制约A股融券发展的重要原因。进而提出了针对A股融券市场费率机制设计的优化方案。

1.2.2 综述总结

根据市场参与者的行为和其影响,对基于主体建模的人工股市进行分类。第一类是基于投资者行为的多主体建模,主要研究对市场有重大影响的主要投资者。这种多主体建模通过模拟投资者行为,来深入分析投资者的决策过程以及投资者的实际行为如何影响市场的运行。第二类是基于市场构建的多主体模型,它主要是研究市场的构建对股票市场的影响。比如市场的结构、法律制度、监管机构制定的规则等。这类建模可以对于制定政策、评估和预测政策市场效果起到很大作用。

综上所述研究股票市场的多主体建模具有重要的理论价值和实践意义。

1.2.3 基于主体的仿真工具研究综述

Swarm是一个基于Agent的仿真平台，它可以用于模拟人类和其他生物的行为。然而，Swarm也存在一些缺点，包括：1、已停止维护：Swarm的最后一个版本发布于2007年，已经停止了维护和更新。这意味着Swarm可能无法适应新的计算机系统和编程环境，也可能无法修复一些已知的错误和漏洞。2、学习曲线较陡：Swarm需要用户具备一定的编程和数学知识，学习曲线较为陡峭。对于不熟悉编程的用户来说，可能需要更长的时间来学习和掌握Swarm的使用方法。3、可视化功能较弱：Swarm的可视化功能相对较弱，用户需要花费额外的时间和精力来构建和调整可视化界面。4、可扩展性有限：Swarm的扩展性相对较弱，用户可能需要自己编写代码来实现某些特定的功能和模型。5、缺乏社区支持：Swarm的社区支持相对较弱，用户可能难以找到其他用户和开发者来共同解决问题和交流经验。

NetLogo是一个广泛使用的主体仿真工具，它可以用于模拟人类和其他生物的行为。NetLogo具有良好的可视化界面和丰富的模型库，可以帮助用户快速构建仿真模型。此外，NetLogo还提供了许多教育资源和社区支持。但是，NetLogo的学习曲线较为陡峭，需要较长时间的学习和实践。

AnyLogic是一款商业仿真工具，它可以用于多种应用领域，例如交通、制造业、物流等。AnyLogic具有丰富的特性和可视化功能，可以帮助用户构建复杂的仿真模型。此外，AnyLogic还提供了基于Java的API和PythonAPI，可以支持更多的编程语言和工具。但是，AnyLogic的使用价格较高，不太符合普通研究者的使用需求。

总的来说，现有的仿真工具存在一些缺点，比如Swarm已经停止维护，不再更新版本。NetLogo的学习曲线较为陡峭，需要较长时间的学习和实践。商业仿真软件AnyLogic，功能强大但使用价格较高，不太符合普通研究者的使用需求。因此，本文选择了开源语言Python作为开发语言。作为市场份额不断攀升的语言，Python可以满足用户对仿真程序的使用需求并实现很好的仿真效果。

1.3 本文的主要贡献与创新

本文通过Python语言构建了一个符合真实股票市场的基于多主体的多股票的人工股票市场。其中参与股票交易的主体分为三类：通过信息优势掌握股票价值的机构投资者、通过历史数据预测的股票价格技术投资者和盲目的噪音投资者。

然后通过撮合这三类投资者，使得这三类投资者互相交易。通过分析基尼技术和泰尔指数进而发现了一些宏观趋势。比如股票市场的财富有逐渐集中化的趋势。表现为基尼系数逐渐增大，达到某一值后，平稳波动。而当增大机构投资者占比后会缓解财富集中的趋势和减小基尼系数的上限。

本文的创新主要是模型上的创新和观点上的创新，模型上的创新是指以往的人工股票市场都是单一股票市场，本文通过预扣款机制和互斥锁实现多股票市场，较之前单一市场模型更加贴切现实。

观点上的创新是指通过分析泰尔指数和基尼系数发现财富有逐渐集中的趋势。

1.4 本论文的结构安排

第一章首先简单阐述本文研究的背景和意义。从金融系统的复杂性和传统金融方法的缺陷出发，引出基于主体的金融方法。随后，又简要介绍基于Agent的人工股票市场的研究思路 and 基础理论，总结了历年以来国内外不同学者的研究成果，第一类是基于投资者行为的多主体建模，第二类是基于市场构建的多主体模型。也简要说明了选择Python作为建模语言的原因。最后对本文的结构安排做了一个简要描述。

第二章从金融市场的复杂性出发，首先说明传统方法的弊端，然后引入Agent的狭义定义、广义定义和Agent四个基本特征。再者介绍了股票市场中重要的三类投资者：基本面投资者、技术投资者和噪音投资者。也对双向拍卖机制和校验人工股票市场的方法做了简单的阐述。最后介绍了分析财富集中度的基尼系数和泰尔指数。

第三章是对股票收益率时间序列进行校验。选择国内外四只股票指数，计算其收益率。最后对其收益率进行校验，校验是否满足尖峰、厚尾分布，波动聚集性和长期记忆性的特征。

第四章是人工股票交易模型的框架构建。这部分介绍了模型中主体的定义、主体的预测规则和市场清算规则，并通过流程图、示意图、类图和伪代码等形式详细解释了主体是如何通过撮合引擎进行交易。

第五章是人工股票市场模型校验与财富不平等分析。首先对模型进行校验，确认其是否符合真实股票市场统计特征。通过重复多次实验，保证实验的可重复性，避免随机因素的干扰。实验分析，人工股票市场中的投资者财富的分布和投资者成分占比的关系，发现适当增加机构投资者能够让投资者的财富分布更加公平。当机构投资者数量较少时，增加技术投资者和噪音投资者数量都会增加财富不平等。

第六章总结全文的研究内容，同时指出研究过程中的不足，并对基于Agent的人工股票市场研究的未来研究和应用前景进行展望。

1.5 本章小结

本章首先介绍背景与意义：贫富之间的差距正在扩大。这种日益集中的财富对全球经济产生了深远的影响。利用CAS思想，建立基于Agent的人工股市市场。通过相对简单的微观个体活动突显出宏观层面的复杂行为，进而可以对于制定政策、评估和预测政策市场效果起到很大作用。然后将国内外的研究做一个简要梳理，基于主体建模的人工股市研究可以分为两类。同时也对本文的主要贡献和创新做了一个说明。最后将本文的安排做一个简单描述，简述每一章的结构安排。

2.基础理论介绍

2.1 金融股票市场的复杂性

金融市场的复杂性是一个长期以来一直被经济学家和金融从业者所研究的问题。这种复杂性体现在许多方面，其中一些主要特征如下：

1.非线性：金融市场的变化通常不是线性的，即市场变量之间的关系不是简单的比例关系。相反，市场变量之间的关系通常是非线性的，可能存在不稳定性和突发性的变化。

2.分形结构：金融市场的变化通常表现出分形结构，这意味着市场变化在不同的时间尺度上都有相似的形态和规律。

3.对初始条件的敏感：金融市场对初始条件的敏感性非常高，微小的初始变化可能会导致大规模的市场波动和变化。

4.易于出现涌现现象：金融市场通常表现出涌现现象，即由于大量个体主体的相互作用而产生的不可预测的整体行为。

5.自组织行为：金融市场通常表现出自组织行为，即个体主体之间相互作用的结果导致整体市场行为的出现。

经济学家用多种方法研究金融市场的复杂性。一种方法是使用数学模型，如博弈论和计量经济学，分析市场行为并理解市场变量之间的关系。另一种方法是使用主体模型^{[5][6][24][25][26][27][28]}，它允许经济学家模拟并分析市场中单个主体之间的相互作用。这种方法提供了对个体行为是如何相互作用以产生整体市场行为的见解^[29]。通过这些模型，经济学家可以了解市场中的不同参与者之间的相互作用，以及他们如何影响市场的变化和波动^[30]。

使用数学方法对系统建模，系统被看作一个整体，使用数学表达式和方程组等来描述系统的主要或基本特性。但是，当组成系统的各个元素之间关

系非常复杂时，数学模型很难描述这些关系。而基于主体的建模方法则是在计算机科学领域中的一种新的系统建模方法，它将系统的各个部分看作独立的个体，称为Agent，以每个Agent的行为和相互作用为基本元素，从而构建出整个系统的模型。

基于主体的模型强调了系统中各个元素之间的相互影响和相互关系。与传统的系统建模方法不同，基于主体的模型将系统中的各个元素看作是有意识的、有态度的主体，而非单纯的数学变量。这种建模方法更加符合人们对现实世界的理解，因此能够更好地反映现实世界的复杂性。利用计算机的特性，基于多Agent技术建立的模型能够实现更大规模的主体，模型交互性和运行速度都有所提高。

2.2 Agent 相关理论

Agent，一词来源于拉丁语agents，原意是“代理（人）”，也就是指：一个人代表他人或组织去完成一些事情。国内资料中对它还缺乏一个统一的译名，因此一般称为智能体、代理或主体等。Agent技术是一种用于模拟复杂系统中的个体行为的技术。它允许模拟系统中的个体做出决策，并与其他个体相互作用。

Agent是一个具有自治能力的实体，可以在动态环境中运行。它的目标是接受来自系统、计算机、用户等实体的委托，并为之提供服务。Agent会根据其委托任务的驱使进行认识和必要的活动，并对动态环境的变化做出合理的反应，与服务主体之间的关系也相对独立。

Agent是指一种独立的、有自主决策和行动能力的主体。在系统建模、人工智能、复杂系统等领域，Agent常常被作为一种抽象的模型，用来描述和研究复杂系统中的主体。Agent可以是人类、机器人、软件程序、组织、团体等。它们可以通过感知、理解和决策等能力与环境进行交互，实现其目标^[31]。

Agent的研究在近年来已在多个领域得到广泛发展，应用范围也十分广泛。然而，由于不同学科和研究内容的差异，学者们对Agent的理解和定义

也不尽相同，导致目前仍难以找到统一的定义。因此，我们将其分为狭义和广义两个定义。

狭义Agent的概念是：该智能实体具备自治性、社会性、反应性和能动性等基本行为特性^[32]。

广义概念则是指，该智能实体不仅具备了狭义概念中的四种基本特征，还应当具有某些人类才具备的基本特征。例如移动性、协同性、理性、知识、责任、信念等更抽象的概念。对于某些特定领域的研究，还需要赋予Agent一些更高层次的特征，使其更符合特定研究对象的特性^[33]。下面是对Agent的几个基本特征的说明^[34]：

(1) 自治性：其自治性主要表现在：Agent具有自主决策能力和自我管理能力的特征。在基于主体的系统建模中，Agent是一个具有独立性和自主性的主体，它可以自主决策和行动。自治性可以帮助Agent自主地解决问题，更好地适应环境，并且最大限度地实现其目标。在复杂系统中，Agent的自治性是一种重要的技能，使系统更灵活，更容易适应变化。

(2) 社会性：Agent能够感知环境变化，根据这些变化及时做出响应并采取行动。

(3) 反应性：指主体根据输入信息和命令做出反应并完成特定任务的能力。这种能力决定了主体是否能够快速做出决策，并针对不断变化的环境做出相应反应。

(4) 能动性：Agent既可以被指派来完成特定任务，也可以根据规则自主探索环境，自我学习，从而更好地解决目标问题。

总的来说，Agent是一种用于描述和研究复杂系统中主体的抽象模型，它具有自治性、社会性和反应性等特征。

2.2.1 多 Agent 系统模型概述

多Agent系统模型的基本构成包括智能体、环境和交互三个部分。其中，智能体是系统的基本单元，具有自主性和决策性，可以感知环境并根据感知信息做出决策；环境是智能体活动的背景，包括智能体所处的空间、时间和

其他物体；交互则是指智能体之间的信息交流和协作行为，是智能体之间相互影响的核心。

在经济学、社会学、生态学等领域，多Agent系统已经发挥了巨大的作用，主要因为社会系统中的自然人和经济人，或者生物系统中生物有着相似性，采用多Agent观点可以很自然地对这些系统建模^[35]。ABM(Agent-Based Modeling)的根本观点是：通过研究系统中自主主体的行为和相互作用，可以理解复杂系统的行为。在ABM中，主体被建模为具有自身属性、行为和决策规则的自主实体。这些主体相互作用，与环境相互作用，以形成系统级的自发行为。

ABM采取自下而上的方法，建模和模拟个体主体的行为，并研究系统行为的结果。这种方法使研究人员可以研究对系统整体行为有贡献的微观层面的过程，并深入了解主体之间的复杂关系和相互作用^[36]。

目前，ABM已经成为探索复杂系统、提取系统本质规律的一门主要方法^[35]。因为这样的系统中个体数量以及类别较多，个体还可能具有适应、学习能力，个体之间的相互作用往往与时间、空间以及个体类型有关，是非线性的，很难用解析方法求解，一般要借助强大的计算能力进行仿真研究，这就形成了对计算实验技术的需求。多Agent系统通常必须具备二个主要的功能：智能体可以自主决定行动以实现某种设计目标，而且能够与其他智能体交互，这种交互不仅是数据交换，而是参与某种共同行为。

多Agent系统模型是一种用于描述和研究由多个自主决策的智能体组成的复杂系统的模型。在多Agent系统模型中，每个智能体都具有自主性、感知性、决策性和优化机制四个基本特征，可以通过感知环境、思考和执行行动来达到预定的目标^[35]。

Agent的表达式如下：

$$Agent = \langle R_t, E_t, V_t, Y_t, N \rangle$$

(1) 特征集合：

R_t 是Agent的特征集合，包括Agent在计算实验过程中不变的特征（如Agent的标识）和可能变化的特征（如Agent可能扮演的角色种类）。

(2) 感知功能：

E_t 是Agent感知到的、对其状态及行为可产生刺激作用的外在事件的集合。Agent在组织中对自身所处局部环境的感知会直接影响其行为决策。例如，通过了解同类Agent的行为信息，可以判断出自己目前的行为决策是否恰当，进而决定是否需要做出适应性改变。

(3) 行为功能:

V 是Agent的行为集合，包括Agent自发采取的，以及受外在事件刺激采取的所有行为。Agent的行为是根据其属性、可获取的资源 and 局部环境，综合计算出下一阶段的行为反应。Agent的行为主要通过对自身资源的封装和调用来实现，形成抽象的服务功能等，并以此作为实现Agent协同的基础。Agent行为规则如下:

$$\forall v_t \in V, \varepsilon(\alpha, v_t, E_t, Y_t) \Rightarrow \langle S_{t+1}, N, \Lambda \rangle$$

其中 $\varepsilon(\alpha, v_t, E_t, Y_t)$ 表示Agent α 在环境 E_t 和决策 Y_t 的情况下完成行为 v_t 后的结果评估， \Rightarrow 表示“满足”， N 为Agent α 所受的约束条件， Λ 为任务标准。

(4) 决策功能:

Y 是Agent在感受外在事件刺激以及与其他Agent进行交互学习过程中所采取的决策机制。Agent在协作过程中，为了保证自身利益与协作成功，需要根据外部环境的变化和自身的规则库不断做出分析和决策。比如是否调整个体行为，是否选择或更换合作伙伴等。此外，Agent在协作过程中，它将不断从环境中获取信息，学习新的规则，改进自身的规则库，调整自己的决策策略，进一步优化协作行为，更好地与环境交互。

$$Y_t = Y_{t-1} \oplus E_{t-1}$$

$$v_t = f(Y_t) |_{\varepsilon(\alpha, v_t, E_t, Y_t) \Rightarrow \langle S_{t+1}, N, \Lambda \rangle}$$

其中， \oplus 表示决策机制 Y_{t-1} 基于环境状态 E_{t-1} 的更新； $f(Y_t)$ 为决策 Y_t 对行为 v_t 的函数映射。

(5) 感知功能:

N 是Agent所受的约束条件，包括环境、其他Agent以及任务目标对该Agent的限制。

2.2.2 多 Agent 的适用领域

多Agent的应用领域非常广泛，可以用于各种复杂的自动化任务和决策系统。例如：

人工智能：Agent可以应用于语音识别、自然语言处理、图像识别、机器翻译、游戏智能等人工智能领域。

自动控制：Agent可以应用于自动驾驶汽车、自动化生产线、智能家居、机器人控制等自动化控制领域。

信息管理：Agent可以应用于智能搜索、信息过滤、推荐系统等信息管理领域。

经济学：Agent可以应用于市场模拟、竞价策略分析、供应链管理等经济学领域。

生物学：Agent可以应用于生物模拟、生物信息学等生物学领域。

社会科学：Agent可以应用于社会模拟、社会网络分析、群体行为模拟等社会科学领域。

总之，Agent在各种需要自主决策、自主学习、自主适应环境的场景中都有广泛的应用。多Agent系统模型为研究复杂系统提供了一种新的思路和工具，可以通过模拟和仿真来揭示系统内部的运作机理和行为特征，对于解决实际问题 and 提高系统性能具有重要意义。

2.2.3 小结

本节主要概述ABM是研究复杂系统、提取系统本质规律的一种主流技术^[37]。多Agent系统一般具有两种重要的能力：首先，Agent至少可以在某种程度上自治行动。其次，每个Agent都可以与其他的Agent进行交互。Agent结构的主体由四个基本特征构成：感知、决策、行为和优化机制。

此外，多Agent系统提供了一种自下而上的研究模式。通过Agent之间的自主合作，多Agent系统可以完成复杂任务，并在演化中产生出宏观规律。因此，可以利用多Agent系统的思想为复杂适应系统的建模和仿真研究开辟一条更加有效的路径^[38]。

2.3 投资者主体

股票市场有形形色色的投资者，但主要可以分为三类：基本面投资者、技术分析投资者和噪音投资者。

2.3.1 基本面投资者

Chiarella^[16]认为基本面分析是一种证券分析方法，其目的是评估公司的基本财务状况和商业模式，以确定其价值和未来表现。机构投资者是最重要的基本面投资者，有信息优势、专业优势，因此能挑选出管理与业绩兼优的股票，其大量持股使之有动力、实力和能力去参与公司治理，包括参加股东大会、提出议案与建议等等，从而间接地促进了股市的健康成长。基本面分析者通常会研究公司的财务报告、收益、营收、现金流等因素，以了解公司的财务和经营状况。希望通过对公司基本面的评估，预测公司的未来表现，为投资决策提供依据。基本面分析投资者通过分析公司财富和确定经营状况来确定基本价格。这个基准价作为对股票真实价值的估计，根据基准价格是否高于市场价格，再决定买入或者卖出。就是说市场价格较基准价格低时，买入。反之卖出^[39]。

2.3.2 技术分析投资者

技术分析投资者^[16]会通过分析股市历史数据以及股票价格的趋势来预测股市的行情。他们通常会使用股票价格的历史图表，包括股票价格的波动、成交量以及技术指标（例如移动平均线）等数据。通过对这些数据的分析，他们可以预测股票的价格可能会在未来几天、几周、几个月或几年内上涨或下跌。他们希望通过对证券价格走势的评估，预测证券的未来表现，为投资决策提供依据。由于包含于历史价格的特征是多样的，因此技术分析投资者的类型有很多。

移动平均线技术是一种最常用的技术分析工具，它是通过对特定周期内的收盘价格的平均价格来计算出来的，可以帮助交易者发现市场趋势并做出

投资决策。移动平均线技术有两种，长期移动平均线和短期移动平均线，当短期移动平均线向上穿过长期移动平均线时，可以作为买入信号；而当短期移动平均线向下穿过长期移动平均线时，可以作为卖出信号^[40]。

2.3.3 噪音投资者

一般情况下，投资者是基于信息进行交易的，但有些投资者可能由于过于自信或者其他原因，错误地把噪声误认为是信息从而依据噪声进行交易。噪声交易是指投资者依据噪声（如短期内的随机价格波动）所进行的交易^[41]。不知情的交易者没有关于股票真实价值的信息，错误地认为他们拥有正确的信息，并根据该信息进行交易。这种交易行为，也会使得股票价格在一定程度上偏离其真实价格。

2.3.4 小结

作为最常见的三类投资者，基本面投资者、技术分析投资者和噪音投资者代表着不同的投资理念。机构投资者依据基本面分析，通过获取研究市场相关信息来估算股票的价值，从而建立一个基准价格，以此来决定是否买入或卖出。技术分析则是通过观察一段时间内的股价走势，来预测股价可能的走向，从而做出投资决策。噪音投资者误把噪声误认为是信息从而依据噪声进行交易。

2.4 人工股市模型的校准

由于计算金融学数据来源于模拟实验，因此我们无法对比真实股市数据与人工股市模型中的参数和结果的差别，验证其是否能反映真实经济活动。因此对人工股市模型进行有效性校准是必要的。目前该领域的研究中，对计算金融模型校准主要有三种说法^[1]：

第一，描述性输出验证，将计算生成的输出与实际金融系统的数据进行比较，并对比其与实际金融数据是否具有相同的统计特征。以确定其反映真实经济活动情况的准确性。

第二，预测输出验证，将计算生成的数据与真实系统的发展趋势相比较。显然这个方法的问题是需要的的时间周期太长。

第三，输入验证，确保模型中包含的基本结构、行为和制度条件再现了实际系统的主要方面，这就是我们所说的事前验证。

在上述三种方法中，本文选择第一种方法校验，即描述性输出验证。

在股票市场中，许多特征性事实表明：在较长的时期内，股票收益率不遵循正态分布，具有显著的尖峰厚尾特征，收益率波动具有显著的聚集性和长期记忆性^{[42][43][44][42][45]}。

2.4.1 尖峰厚尾分布的检验

在图形上，尖峰厚尾分布与正态分布有很大的差别。正态分布的图形是一个对称的钟形，具有连续性，而尖峰厚尾分布的图形则比较不对称，峰值较高，尾部较厚。从统计学角度来看，这意味着有更大的概率出现极端值。

总的来说，判断是否服从正态分布有很多方法，最直观的方法是绘制分布图。还有其他检验方法，比如偏度和峰度检验、JB检验和KS检验等。

偏度是用来衡量分布形态的对称性的一种度量，特别是相对于正态分布来说的。正态分布是一种完全对称的分布。因此，如果一组数据的偏度为0，则表明它的分布状况与正态分布类似。如果偏度大于0，表明数据分布情况偏向于大值，反之，如果偏度小于0，表明数据分布情况偏向于小值^[46]。如果数据采样自一个正态分布的总体，那么JB检验渐近于 $\chi^{(2)}$ 分布。

$$SK = \frac{m_3}{\frac{m_2^2}{3}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{\frac{3}{2}}} \quad (2-1)$$

薄尾型分布的概率密度函数通常以指数函数的速度衰减至零。这意味着，当随机变量的值离均值越远，其对应的概率密度函数值就越小。因此，薄尾型分布的数据分布在均值附近，极端值很少，常见的薄尾型分布有正态分布和指数分布。长尾分布的概率密度函数通常以幂函数的速度衰减至零。这意

味着，当随机变量的值离均值越远，其对应的概率密度函数值减小的速度越慢^[47]。

一般来说，峰度是一种用于描述数据分布形状的重要指标，峰度定义如下：

$$K = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} - 3 \quad (2-2)$$

峰度（K）可以作为评估数据分布是否符合正态分布的一个快速参考指标。当K=0值时，表明分布的峰度与正态分布相同，没有尖峰问题；而K>0则表明该分布比正态分布具有更高的峰度。

2.4.2 波动聚集性的检验

许多学者通过实证研究发现，我国股票市场的日收益率存在明显的聚集性特征。波动聚集性指的是，在收益率大幅度波动之后，会紧邻着另一次大幅度的波动^[48]。

ARCH检验就是其中一种检验时间序列波动性聚集的方法。具体步骤如下：首先为了检验ARCH效应，先建立均值模型，拟合 μ_t ，计算残差 $a_t = r_t - \mu_t$ 。用残差序列的平方作ARCH效应检验。其次对 a_t^2 作Ljung-Box白噪声检验，检验不显著时没有ARCH效应，检验显著时有ARCH效应。

2.4.3 长期记忆性的检验

股票价格的长期记忆性^{[18] [49] [50] [51]}是指不同时期的股票价格在相隔较远周期时仍然具有一定的相关性，股票价格序列的自相关系数具有持续性，既不像平稳过程呈现指数迅速衰减，也不像单位根过程呈现不衰减，而是双曲线式的衰减，衰减得更慢，持续时间更长。更进一步，当自相关系数为正时，股票维持原来的运行趋势；相反，当自相关系数为负时，股票的运行趋势发生逆转^[52]。

检验收益长期记忆性的经典方法是英国水文学家赫斯特（Harold Edwin Hurst）在于1951年提出的R/S分析（重新标度极差分析）^[53]。1951年，Hurst

在研究尼罗河水位变化时发现了时间序列中存在的长记忆性（long term memory），即时间序列当前（或过去）的取值以远超随机扰动所能达到的程度影响该时间序列在未来的取值^[54]。

R/S分析具体步骤为，将长度为N的时间序列Y等分为n个子集，每个子集的长度为k=N/n，分别计算每个子集的最大值与最小值的差值和标准差。由公式2-3和公式2-4可以求得 (R/S)_n。再根据最小二乘法估算出公式2-5中Hurst指数。

$$R_n = \max(Y_1, \dots, Y_n) - \min(Y_1, \dots, Y_n) \quad (2-3)$$

$$S_n = \sqrt{\frac{(x_r - \bar{x})^2}{n}} \quad (2-4)$$

$$(R/S)_n = c \cdot n^H \quad (2-5)$$

时间序列的Hurst指数居于0-1之间。以0.5为间隔，时间序列在不同的区间会表现不同的特性：

当 $H=0.5$ 时，则说明该时间序列具有随机性，也就是数据没有明显的趋势或规律^[55]。

当 $0.5 < H < 1$ 时，则说明该时间序列具有长期相关性，也就是数据的趋势是稳定的，并且在一段时间内具有相似的波动。

当 $H < 0.5$ 时，则说明该时间序列具有反趋势，也就是数据的趋势是不稳定的，在一段时间内波动的幅度随着时间的推移而变大^[56]。

2.4.4 小结

股票市场中有许多特征性事实表现：在较长周期内，收益率分布服从尖峰、厚尾型的非正态分布，并且价格或收益率序列有明显的波动聚集性和长期记忆性^[42]。是否是正态分布可以通过绘制分布图，偏度，峰度，JB检验等方法检验。波动聚集性检验可以用残差序列的平方作ARCH效应检验。长期记忆性的检验则可以使用Hurst指数进行检验。

2.5 竞价策略

双向拍卖^{[57] [58]}是一种拍卖方式，它允许卖方和买方通过竞价来达成交易。在双向拍卖中，买方和卖方都可以提交自己的报价，卖方提交最低售价，而买方提交最高购买价。然后，通过逐渐逼近对方的价格，买方和卖方在双方都认为合适的价格上达成交易。

在双向拍卖中，卖方和买方可以根据自己的需求来自由调整价格，这使得双向拍卖相对于其他拍卖方式更加灵活和公正。双向拍卖在各种交易中都有广泛应用。例如房地产交易、股票交易、艺术品拍卖等等。根据双向拍卖交易规则中的买卖双方发生交易的时间不同，将双向拍卖划分为连续竞价和集合竞价。

2.5.1 集合竞价

所谓集合竞价^{[59] [60]}就是在当天还没有成交价的时候，欲要入市的投资者可根据前一天的收盘价和对当日股市的预测来输入股票价格，而在这段时间输入的所有价格都是平等的，不需要按照时间优先和价格优先的原则交易，而是按最大成交量的原则来定出股票的价位，这个价位就被称为集合竞价的价位，而这个过程就被称为集合竞价。

2.5.2 连续竞价

所谓连续竞价是指对买卖申报按照逐笔连续撮合的竞价方式。集合竞价结束后，证券交易所开始当天的正式交易，交易系统按照价格优先、时间优先的原则，确定每笔证券交易的具体价格^[60]。

2.5.3 连续竞价原则

连续竞价的两个原则是：首先是价格优先，即买单则价格较高者能优先成交，卖单则是价格较低者能优先成交。其次是时间优先：即买卖方向和价格相同的委托单，先申报的委托单会比后申报的委托单优先成交。

2.5.4 小结

双向拍卖主要包括集合竞价和连续竞价。集合竞价主要是为了达成最大交易量。连续竞价有时间优先和价格优先两个原则。两种竞价方式在交易机制上存在一些差异：

时间不同：连续竞价在交易日开始时就开启，并持续至交易日结束；而集合竞价则只在交易日开盘前和收盘后的短暂时间内进行。

报价方式不同：连续竞价中买卖双方可以根据市场行情实时报价；而集合竞价中买卖双方则需在规定时间内提交限价订单，即指定买入或卖出证券的数量和价格。

交易价格确定方式不同：连续竞价中买卖双方根据市场实时行情进行撮合交易，最终交易价格由市场供需情况决定；而集合竞价中，交易所根据所有限价订单的价格和数量进行匹配，最终交易价格由一定的匹配规则决定，开盘价和收盘价也由此确定。

2.6 财富的集中度

财富的集中是指财富、收入和资源在不同社会阶层之间的不平等分配。这意味着很小一部分人拥有大量的财富，而其他人口拥有的财富要少得多。这种财富的集中通常是收入不平等的结果，即最高收入者和最低收入者之间的差距过大。这种财富集中是造成贫富差距扩大的主要因素，因为富人变得更富有，穷人变得更贫穷。

财富的集中不仅是一个经济问题，也是一个社会问题。它导致贫困和不平等的加剧，因为富人能够积累更多的财富，而穷人则被抛在后面。它还可能导致社会流动性的下降，因为那些财富较少的人无法获得富人获得的资源和机会。

2.6.1 洛伦兹曲线

洛伦兹曲线（Lorenz curve）是用来描述收入分配不平等程度的一种可视化工具。它是一条反映收入分配情况的曲线，横轴表示人口比例，纵轴表

示收入比例。洛伦兹曲线研究的是国民收入在国民之间的分配问题，该曲线得名于美国统计学家M.O.洛伦兹（Max Otto Lorenz），而基尼系数则是在此基础上由意大利经济学家基尼所定义的。

2.6.2 基尼系数

基尼系数^[61] (Gini coefficient) 是一种经济学和统计学的指标，用来衡量一个国家或社会内部的财富分配不均衡程度。它是由意大利经济学家Corrado Gini于1912年首先提出的^[62]。基尼系数是国际上通用的、用以衡量一个国家或地区居民收入差距的常用指标之一。其包括收入基尼系数（Income Gini）和财富基尼系数（Wealth Gini）。两者的算法大致相同，区别在于收入基尼系数的数据是来自某地区的家庭收入统计，财富基尼系数的数据是来自某地区的家庭总资产统计。公式如下：

$$G = \frac{A}{A + B}$$

其中A为实际收入分配曲线和收入分配绝对平等线之间的面积，如果A为零，即基尼系数为0，表示完全均衡，也就是所有人的财富数量相等。B为实际收入分配曲线和收入分配绝对不平等线之间的面积。如果B为零，则系数为1，代表收入完全集中在少数人手中。

2.6.3 泰尔指数

泰尔指数（Theil index）是一种用于度量收入或财富分配不平等程度的统计指标，常用于经济学和社会学等领域。它是由荷兰统计学家Theil在1967年提出的。

$$T = \sum_{k=1}^K w_k \ln\left(\frac{w_k}{e_k}\right)$$

其中， w_k 表示第k组收入占总收入的比重， e_k 表示第k组人口数占总人口数的比重。

泰尔指数^[63]作为收入不平等程度的测度指标具备良好的可分解性质，即将样本分为多个群组时，泰尔指数可以分别衡量组内差距与组间差距对总差距的贡献。假设包含 n 个体的样本被分为 K 个群组，每组分别为 g_k ，第 k 组 g_k 中的个体数目为 n_k ，则有 $\sum_{k=1}^K n_k = n$ 。

记 T_b 与 T_w 分别为组间差距和组内差距，则可将泰尔指数分解如下：

$$T = T_b + T_w = \sum_{k=1}^K y_k \ln\left(\frac{y_k}{n_k/n}\right) + \sum_{k=1}^K y_k \left(\sum_{i \in g_k} \frac{y_i}{y_k} \ln \frac{y_i/y_k}{1/n_k} \right)$$

其中 y_i 表示个体 i 的收入份额（占总收入的比例）， y_k 表示第 k 组的收入份额（占总收入的比例）。

2.6.4 小结

本节主要是介绍通过绘制洛伦兹曲线进而计算基尼系数，可以衡量财富的集中度。基尼系数的范围从0（表示完全平等）到1（表示完全不平等）。而泰尔指数则可以分析不同组内和组间的财富不平等程度。

2.7 本章小结

本章节主要从金融市场的复杂性出发，首先说明传统方法的弊端，再重点阐述Agent狭义定义、广义定义和多Agent系统模型概述。也是建立基于多主体的人工股市所需要的知识进行梳理。其次介绍了股票市场重要的三类投资者：基本面投资者、技术投资者和噪音投资者。然后引出集合竞价和连续竞价。集合竞价主要是为了达成最大交易量。连续竞价有时间优先和价格优先两个原则。最后介绍了衡量财富集中度的基尼系数和衡量组间和组内不平等程度的泰尔指数。

3.国内外股票指数收益率统计特征校验

3.1 股票选取

本文选择四个具有代表性的国内外股票指数指标，分别为标普500、深成指数、上证指数、恒生指数，对其统计特征进行检验。首先利用第三方工具pandas_datareader从雅虎财经上获取这四个指数，2010年至2020年每日收盘数据。然后计算对数收益率，对数收益率计算公式如下：

$$R_t = \ln p_t - \ln p_{t-1}$$

其中 p_t 表示第 t 个交易日的收盘价格。

3.2 正态性检验

本文首先绘制直方图，能够直观地观察收益率不符合正态分布。如图3-1所示。再绘制收益率分位图（quantile-quantile plot, QQ图），如图3-2所示。可以观察这些点没有集中到一条直线上。也表明不符合正态分布。最后计算偏度、峰度和JB统计量结果如下表3-1所示。

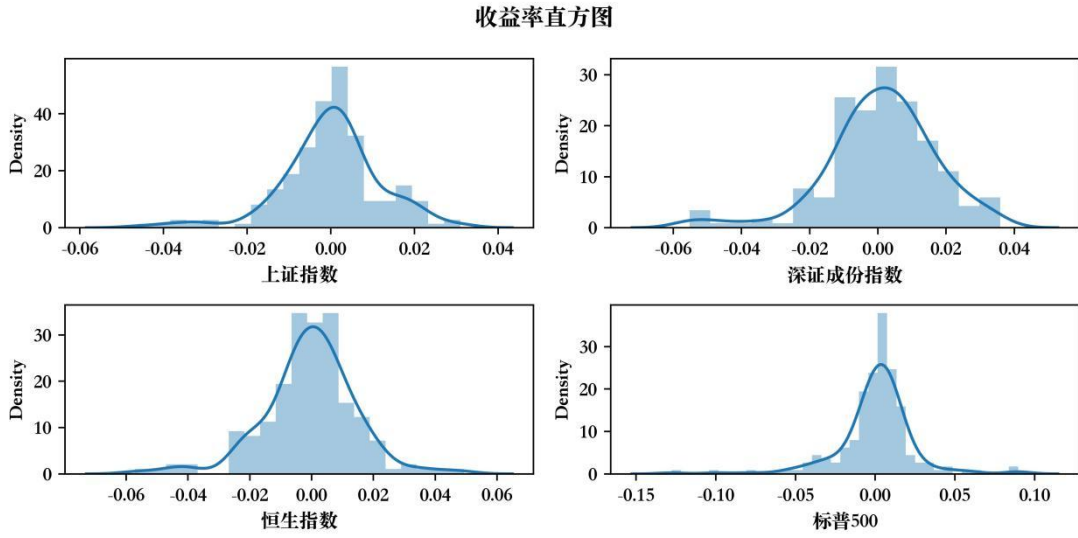


图 3-1 股票收益率直方图

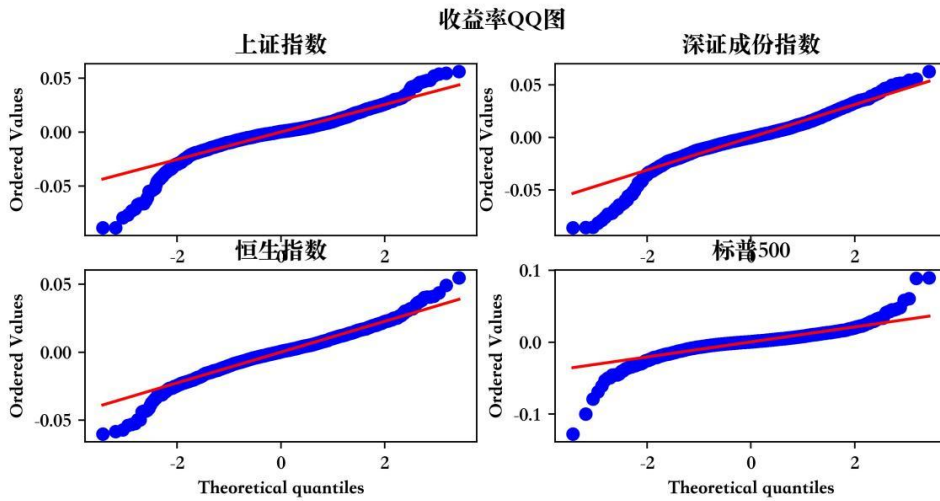


图 3-2 收益率 QQ 图

表 3-1 上证指数、深证成份指数、恒生指数、标普 500 收益率统计量

	均值	方差	偏度	峰度	JB统计量
上证指数	0.00051713	0.0138594	-0.431853	8.60768	679.451
深证成份指数	0.0013433	0.017338	-1.0	3.9391	177.453
恒生指数	-0.000205542	0.0156748	-0.335128	2.33328	52.8172
标普500	0.000621422	0.0227429	-0.857991	7.92566	597.514

综合图3-1、图3-2和表格3-1均表现出尖峰、厚尾的非正态分布的特征。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/025340223202011043>