

统计学在投资组合中的应用

刘怀琛

上海上实剑桥教育中心, 上海 200120

摘要： 本文较为详细地介绍了数理统计在投资组合中的详细应用，其中包括概率统计学理论知识，投资组合的理论分析，系统、科学地用数理统计学方法来确定所选股票和投资组合的收益率和方差，并对收益和风险进行了分析。利用 EXCEL 进行规划求解，确立有效前沿曲线，确定投资组合的有效性。根据分析结果，得出在投资组合中各股票的最优占比，分析出性价比最高和风险最小的投资组合方式。

关键词： 数理统计；投资组合；投资风险；有效前沿

Application of Statistics in Investment Portfolios

Liu Huaichen

Shanghai Shangshi Cambridge Education Center, Shanghai 200120

Abstract： This article introduces the detailed application of mathematical statistics in investment portfolios, including theoretical knowledge of probability and statistics, theoretical analysis of investment portfolios, and the systematic and scientific use of mathematical statistics methods to determine the return rate and variance of selected stocks and portfolios. Additionally, it analyzes returns and risks. EXCEL is used for programming solutions, establishing an efficient frontier curve, and determining the effectiveness of investment portfolios. Based on the analysis results, the optimal proportion of each stock in the investment portfolio is obtained, and the investment portfolio with the highest cost-effectiveness and minimal risk is analyzed.

Keywords： mathematical statistics; investment portfolio; investment risk; efficient frontier

引言

随着互联网和经济格局的变动和日益成熟，金融市场中的一些问题和形势需要逐渐规范处理；这需要统计学中的理论和模型进行分析，并寻求解决方案；其在证券期货交易中的作用也日益增长^[1]。基于以上，本文将详细分析统计学在投资组合，以及投资风险中的具体应用。

一、投资组合和风险的理论原理

投资组合是金融领域中的一个概念，指投资者为实现财务目标而持有的一系列资产，包括股票、债券、基金、现金等，通过将不同的投资资产按照一定比例组合到一起，以达到最佳的风险收益平衡点，其目的是为了分散风险^[2]。在进行投资组合分析时，投资者需要考虑多个方面的因素，包括不同资产类别的收益率，风险水平，相关性等。在进行投资组合分析时，有几种常见的方法，包括均值-方差模型、资本资产定价模型（CAPM）和有效前沿分析^[3]。

美国经济学家马科维茨在1952年首次提出投资组合理论^[4]。该理论包含两个重要内容：均值-方差分析方法和投资组合有效边界模型。了解现代投资组合理论，收集关于证券投资、股票的相关

数据，掌握聚类分析、回归分析等分析方法，建立均值-方差等模型，对已有的投资组合进行分析、预测，最后评价投资组合的优劣，从中选出最优的投资组合。马科维茨投资组合理论提出了一个重要的前提假设，即投资者对于证券投资的主观意愿是恒定值、没有差别，换言之就是投资者的投资行为并不受其主观意愿的影响，同时现代投资组合理论之中资产指标被看作整体^[5]。此外也提出了有效边界的概念，用以区分投资者建立的组合是否具有有效性，因此在这一假设下，我们可以运用数理统计的方法来确定组合证券预期收益率、方差，从而规划投资组合的确立。

基于理论，本文自选了证券投资组合，对投资组合理论展开了深层次地探究。运用数理统计的方法来确定投资组合的收益率和方差，从而规划求解，最终确立有效前沿曲线，确定投资组合的有效性^[6]。

作者简介：刘怀琛（2007.03-），汉族，内蒙古呼和浩特人，研究方向：统计学。

二、实证分析——基于数理投资组合确立投资组合及有效前沿曲线

(一) 数据来源

由于本文所建立的时间序列的时间为2024年8月起，因此股票需选择上市日期在2024年8月之前，同时考虑到金融、房地产、制造业，信息技术作为股票的四大热门板块。因此本文以深圳证券交易所为基础，具体的股票（括号内为股票代码）为：平安银行(000001)、大悦城(000031)、美的集团(000333)、国华人安(000004)，从而建立本文所研究的投资组合。具体而言，选取了从2018年12月28日到2023年12月29日这四支股票五年内的每月收盘价，进行投资组合和有效前沿分析。

(二) 主要模型

本文依据马科维茨资产组合理论^[7]构建所使用的模型，其中一个重要假设即马科维茨模型的基本假定，投资者均为理性的投资者，通过一定收益率水平下的风险衡量来确定最优的投资组合理论^[8]。具体模型如下：

$$\min \sigma^2 = \sum \sum x_i x_j \text{cov}(R_i, R_j)$$

σ^2 代表投资的风险， x_i ， x_j 为股票 i 、 j 的投资比例， R_i ， R_j 为第 i 只股票、第 j 只股票的收益， $\text{cov}(R_i, R_j)$ 为两只股票之间的协方差。

(三) 选取投资单品

选取了2018年12月28日到2023年12月29日平安银行、大悦城、美的集团，以及国华人安这四支股票五年内的每月收盘价。进而我们可以得到这五年内这四支股票每月的历史净值，不妨比较出在同一时间内投资单品的涨跌幅。

在一众投资单品中选择合适的产品进行组合投资之前，需首先计算投资品之间的相关系数^[9]，计算公式如下：

$$\rho_{xy} = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(X) * \text{Var}(Y)}} = \frac{\text{Cov}(x, y)}{\sqrt{\sigma_x * \sigma_y}}$$

其中， $\text{Cov}(x, y)$ 是两个产品的协方差， $\text{Var}(X) * \text{Var}(Y)$ 是两个产品标准差的乘积，当 ρ_{xy} 为正数，且越大时，两个产品之间的正相关性越强，说明这两组产品高度相似，在进行投资组合时无法分散风险和提提高收益率，需要合并。通过计算得到我们所选取的四组产品之间的相关性，如表1所示。

表1：组间相关系数

相关系数				
	平安银行	大悦城	美的集团	国华人安
平安银行	100%			
大悦城	46%	100%		
美的集团	56%	38%	100%	
国华人安	1%	16%	-1%	100%

这四支股票的正相关性不是特别高，可以进行组合投资。

(四) 衡量投资单品的投资风险

在金融投资中，我们通常最关心的是收益率，因此收益率的不确定性就是本研究所认定的投资风险^[10]。月均回报率 r 的计算公式如下：

$$r = E(P) = \frac{(\omega_1 + \omega_2 \dots \omega_n)}{T} = \frac{\sum_{t=0}^n \omega(t)}{T}$$

通过计算求得结果，如表2所示。

表2：投资单品的月均回报率

	平安银行	大悦城	美的集团	国华人安
月均回报率 r	0.55%	-0.32%	1.28%	1.17%

由于在后续和日常分析中我们讨论最多的是年均回报率，所以需要进入如下计算：

$$\text{年均回报率} = (1+r)^{12} - 1$$

通过计算求得结果，如表3所示。

表3：投资单品的年均回报率

	平安银行	大悦城	美的集团	国华人安
年均回报率	6.85%	-3.73%	16.45%	14.93%

用标准差（方差）是衡量投资组合和单一资产风险的指标，反映了资产收益的波动性。当标准差较大时，表明资产收益在短期内有较大的波动，投资者可能会面临较大的风险，相反的，当标准差较低时，则意味着资产的收益相对稳健，投资者所面临的风险较小。投资单品的标准差的计算公式如下：

$$\sigma = \text{Std}(X) = \sqrt{\frac{((x_1 - \mu)(x_2 - \mu) \dots (x_n - \mu))^2}{N}} = \sqrt{\frac{\sum_{i=0}^n (X - \mu)^2}{N}}$$

计算结果如表4所示。

表4：投资单品的标准差

	平安银行	大悦城	美的集团	国华人安
标准差	9.34%	9.55%	8.63%	16.15%

投资单品的方差 σ^2 如表5所示。

表5：投资单品的方差

	平安银行	大悦城	美的集团	国华人安
方差	0.87%	0.91%	0.74%	2.61%

由此我们可以得出，投资国华人安的风险较大，投资美的集团的收益相对稳健。

(五) 投资单品的性价比

夏普比率是基金绩效评价的标准化指标，计算公式如下：

$$\text{SharpeRatio} = \frac{r}{\sigma}$$

其中， r 是月均回报率， σ 是标准差，夏普比率是可以同时对收益与风险综合考虑的指标。在风险相同的情况下，夏普比率越高的投资品具有更高的投资价值，性价比更高。本文所选取的四个投资单品的夏普比率如表6所示。

表6：投资单品的夏普比率

	平安银行	大悦城	美的集团	国华人安
夏普比率	5.93%	-3.31%	14.81%	7.22%

由此可以得出，投资美的集团性价比相对比较。

(六) 投资组合

综上分析，美的集团在这五年的月均回报率最高，为16.45%，标准差（风险）最小，为8.63%，夏普比率最高14.81%，最值得投资。美的集团近几来的发展已经不再是像最初一样局限在家电领域，逐步扩展到电梯、高低压变频器、医学影像产品、机器人及自动化系统业务，其他业务还包括智能供应链、家电原材料销售、批

发及加工业务，以及吸收存款、同业拆借、消费信贷、买方信贷及融资租赁等金融业务。其业务广泛，发展稳健上升。但本着“鸡蛋不放在同一篮子里”的原则，我们也无法预测每家公司未来的发展状况，为了降低风险，确保并提升收益，还需进行分散投资。

投资组合的回报率和标准差

在计算投资组合中各产品的投资比重前，我们要定义数据结构，权重总和为100%。接下来计算投资组合的回报率，公式如下：

$$r_p = w_1 * r_1 + w_2 * r_2 \dots w_n * r_n = \sum_{i=0}^n w_i r_i$$

其中， w_i 为投资单品的权重， r_i 为投资单品的回报率。

计算投资组合的标准差，公式如下：

$$\sigma_p = \sqrt{(w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 \dots w_n^2 \sigma_n^2) + (2w_1 w_2 \rho_{12} \sigma_1 \sigma_2 + 2w_3 w_3 \rho_{23} \sigma_2 \sigma_3 \dots) + 2w_{n-1} w_n \rho_{n-1} \sigma_{n-1} \sigma_n}$$

之后，我们利用 EXCEL 中的“规划求解”工具计算投资组合前沿曲线上的有效集，其本质为如下公式：

$$MIN \sigma_p(r_p) = \sigma_p = \sqrt{\sum_{i=0}^n \sum_{j=0}^n w_i w_j cov(P_i, P_j) \sigma_{ij}}$$

给定一个投资组合的回报率，求出投资组合中各单品的投资比例，并且使波动性，也就是方差，最小。

那么，我们给定的投资组合的回报率的范围，起点要比这四支投资单品中，年回报率最低的大悦城（-3.73%）高一点，终点要比年回报率最高的美的集团（16.45%），范围在 -3% ~ 16%。最终，规划求解得出如下投资组合。具体投资权重，年收益率，月收益率，月标准差，月方差，以及夏普比率如下表7所示。

表7：最优投资组合

	投资权重 P				年收益率	月收益率	月标准差	月方差	夏普比率
	平安银行	大悦城	美的集团	国华网安					
性价比最高	0%	0%	79%	21%	16%	1.25%	7.60%	0.58%	16.51%
风险最小	21%	24%	40%	16%	9%	0.73%	6.81%	0.46%	10.74%

最终，经过数理统计的分析，以深圳证券交易所为基础，选取的四支股票：平安银行(000001)、大悦城(000031)、美的集团(000333)、国华网安(000004)构建的投资组合中，性价比最高的投资权重为：平安银行0%，大悦城0%，美的集团79%，国华网安21%，年收益率为16%；风险最小的投资权重为：平安银行21%，大悦城24%，美的集团40%，国华网安16%，年收益率为9%。

三、总结

投资组合理论的核心是通过投资多种不同类型的资产，构建一个多元化的投资组合，以达到风险和收益之间的最佳平衡。通过实证可以看出，投资组合中的投资单品的占比会对整个组合的效益产生影响。其中，风险和收益成正相关，当收益较大时，其需要承担的风险也在上升。通过计算所选股票和投资组合的收益率，方差，以及夏普比率，对收益和风险进行了分析。利用 Microsoft Excel 中的分析工具进行规划求解，建立有效前沿曲线，并确定投资组合的有效性。根据计算和分析结果，得出所建立的投资组合中四支股票的最优占比，分析出性价比最高和风险最小的投资组合方式。

随着大数据和信息化的不断发展，现如今金融领域的发展离不开数理统计分析，不论是日常投资理财，国际经济形势分析和预测，还是经济政策的建立。将金融体系和大数据、信息化相结合已经是当今的发展趋势。通过数理统计的方法对信息和数据进行收集、获取、判断和分析，这样不仅大大提高工作效率，也减少了不必要的工作损失，让金融工作更加数字化、智能化。同样，各行各业的高需求也促使数据驱动的统计的发展空间得到进一步的提升，形成新的数智发展趋势。

参考文献

[1] 李从珠,李文蕾,江铁军. 统计学在证券期货市场中的应用 [J]. 数理统计与管理, 2000(9): 51 - 64.
 [2] 李婧涵, 彭于校, 阳晓玲. 浅谈统计学在证券期货市场中的应用 [J]. 四川水泥, 2017: 305.
 [3] 裴永苹. 证券投资技术分析的组合作用 [J]. 佳木斯教育学院学报, 2013, (05): 471+474.
 [4] 孙江睿. 统计学在证券期货市场中的应用 [J]. 统计与管理, 2015(5): 54 - 55.
 [5] 杨桂元, 唐小我. 组合证券投资的统计决策方法探讨 [J]. 统计研究, 1999(7): 53 - 55.
 [6] 王海霞, 林伽伊. 证券投资基金持股影响股票价格波动性分析 [J]. 中国产经, 2023, (06): 88-90.
 [7] 邵玉姣. 财经专业虚拟仿真证券投资实训体系改革与创新研究 [J]. 老字号品牌营销, 2024(06): 63-65.
 [8] 徐高. 理解“证券投资研究” [J]. 金融博览, 2022(01): 30-32.
 [9] 谭璐. 证券投资基金风险及应对分析 [J]. 投资与创业, 2021, 32(24): 27-29.
 [10] 郑新宇. 证券投资中随机指标判断法则的运用 [J]. 上海商业, 2022(03): 220-221.