|  |  |
| --- | --- |
| 成绩 |  |

**云南大学2024学年春季学期数学与统计学院2021级**

**《应用多元统计分析实验》**

**期末论文**

**题目**

**我国科创板上市企业财务绩效评价  
——基于因子分析和聚类分析**

学 院：　数学与统计学院

专 业：　　　统计学

姓 名：　　　枫叶

学 号：

教 师：　　　李会琼

**我国科创板上市企业财务绩效评价  
——基于因子分析和聚类分析**

摘要：本文以2023年披露财务信息的566家科创板上市公司为对象，基于因子分析法对样本公司的财务指标进行综合评价，并结合聚类分析结果揭示各类别公司存在的问题，在此基础上，本文提出了一定建议。

关键词：因子分析；聚类分析；科创板；绩效评价

**目录**

[一、 研究背景及意义 3](#_Toc170658085)

[二、 国内外研究现状 3](#_Toc170658086)

[三、 数据说明 5](#_Toc170658087)

[四、 描述性统计 5](#_Toc170658088)

[五、 因子分析 6](#_Toc170658089)

[（一） 可行性验证 6](#_Toc170658090)

[（二） 因子提取 6](#_Toc170658091)

[（三） 综合评价 8](#_Toc170658092)

[六、 聚类分析 8](#_Toc170658093)

[七、 总结 10](#_Toc170658094)

[八、 附录 12](#_Toc170658095)

# 研究背景及意义

近年来，中国资本市场的快速发展和科技创新的不断推进，使得科创板（科创企业板）的设立成为资本市场深化改革的重要举措之一。科创板自2019年设立以来，作为中国资本市场的新兴板块，吸引了众多高科技、高成长企业的上市。这一板块的设立不仅为科技创新企业提供了直接融资的平台，也为投资者提供了更多的投资选择和机会。为促进科创板企业的长期健康发展，有必要对其财务特征进行综合评价，以此评估科创板企业的经营绩效，从中找出存在的问题，并通过聚类分析，把握企业在行业内所处位置及优劣环节，为企业或政府改善企业经营情况提供一定参考。

# 国内外研究现状

企业财务绩效评价指运用财务指标对企业业绩进行科学的评价，由于主要针对财务进行评价，作为企业综合绩效的评估来说具有一定的片面性，但由于财务在企业业绩中的核心地位以及上市公司财务数据的公开性，对上市公司进行财务绩效评价已经是企业绩效评价的重要组成部分。

国内外对财务绩效进行评价的方法主要有主成分法、因子分析法、熵值法、层次分析法、经济增加值法（EVA）和平衡记分卡（SBC）等，如史本山等 (2005)从10个主要财务指标出发, 基于主成分, 以全国上市公司1996-2002年的财务数据为样本数据进行经营绩效综合评价。杜龙波和黄业德 (2010) 使用因子分析法对所选取的33家鲁中地区的上市公司的有关的财务数据计算其因子得分, 从而评价上市公司的财务绩效。朱顺泉等 (2003) 的研究中采取了包括获利能力、成长能力、运营能力、偿债能力、和资本结构等五个指标体系, 并运用熵值法对财务指标进行了客观赋权, 使用模糊评价模型对公司财务状况进行了合理的评价。温素彬等 (2009) 的研究中提出了以企业经济绩效、社会绩效和生态绩效等共生模式和矩阵式评价指标体系, 以权变理论为基础构建出企业三重绩效的层次变权综合评价模型, 从而对三进行均衡化处理, 最后使得绩效评价能够更加具有全面性和科学性以及决策相关性。马方和赵金杰 (2010) 通过对EVA评价方法在推广的分析出其在应用中存在的难点和问 性的收入和股权资本成本的计量问题, 并指出由于以上问题很大程度上阻碍了EVA考核指标的推广应用, 并且 有多个省 (区) 向国资委正式提出申请, 并指出应当暂缓实行该考核评价办法。曹阳 (2009) 通过对比分析了国内众多企业的方法实施失败的原因, 指出了平衡记分卡的实施过程, 即是确定关键绩效指标、评价实施、评价面谈、制定绩和绩效改进指导等过程, 并在此基础上提出了能够解决相关问题的对策。

除此之外，在计算出综合评价得分之后，也有学者尝试使用聚类分析、回归分析等方法得到更加精细化的结论，如叶陈毅和管晓（2021）就以2018年医药制造业74家上市公司为对象，在因子分析法的基础上进一步用聚类分析对企业进行划分。

# 数据说明

本文选取2023年完整披露财务信息的566家科创板上市公司作为样本，所有数据来源于RESET数据库，结合现有研究，本文选取了从盈利能力、偿债能力、现金结构、成长能力四个方面选取了12个代表性财务指标，分别为基本每股收益，净资产收益率，总资产报酬率，营业利润率，流动比率，速动比率，现金负债比，现金流动负债比，利润总额同比增长，总资产同比增长，资产负债率，完整数据见附录。

# 描述性统计

表 1描述统计表

| 变量 | 最小值 | 中位数 | 平均值 | 最大值 |
| --- | --- | --- | --- | --- |
| 基本每股收益 | -8.0800 | 0.7800 | 1.1839 | 41.0000 |
| 净资产收益率 | -9650.398 | 9.811 | -6.615 | 76.868 |
| 总资产报酬率 | -121.771 | 8.269 | 9.349 | 112.481 |
| 营业利润率 | -948649.4 | 17.3 | -2189.6 | 101.1 |
| 流动比率 | 0.4435 | 3.3561 | 5.2645 | 40.1849 |
| 速动比率 | 0.3069 | 2.9591 | 4.6162 | 39.6011 |
| 现金负债比 | -3.19660 | 0.14555 | 0.24992 | 3.42990 |
| 现金流动负债比 | -4.47790 | 0.17990 | 0.30202 | 5.22970 |
| 利润总额同比增长 | -2045.663 | 22.392 | 107.353 | 10755.240 |
| 总资产同比增长 | -68.06 | 29.82 | 56.98 | 899.14 |
| 资产负债率 | 2.615 | 25.484 | 29.595 | 97.419 |

表 1给出了样本数据的描述性统计结果，其中净资产收益率和营业利润率的负值最小值和负值平均值反映了一些公司存在严重亏损现象。然而，中位数和最大值表明许多公司仍有良好的盈利能力；流动比率和速动比率显示公司在流动资产方面的表现良好，中位数和平均值都远高于最低值，表明大部分公司流动性充足；利润总额同比增长和总资产同比增长的最大值极高，显示有些公司在快速增长，但负值最小值和较低的中位数表明整体增长情况参差不齐；资产负债率的中位数和平均值接近，表明大多数公司资产负债情况较为稳定，但最高值接近100%，显示个别公司负债率较高。

# 因子分析

## 可行性验证

首先基于KMO值和Bartlett球形度检验考察因子分析的可行性。其中KMO值度量了相关系数与偏相关系数的相对大小，如果原始数据中确实存在公共因子，则各变量之间的偏相关系数应该很小，那么KMO值将接近1，一般认为KMO值在0.6以上即适合进行因子分析。Bartlett球形度检验的原假设是相关系数矩阵为单位阵，若拒绝原假设则认为变量间存在相关性，适合作因子分析。

表 2 KMO和Bartlett检验

|  |  |  |
| --- | --- | --- |
| KMO值 | | 0.63 |
| Bartlett球形度检验 | 近似卡方 | 5157.695 |
| 自由度 | 55 |
| P值 | 0.000 |

从表 2可以看到，由样本数据计算出的KMO值大于0.6，说明存在公共因子的可能性较大；Bartlett球形度检验的p值小于0.05，在0.05的显著性水平下可以认为变量间存在相关性，故所选取的样本及变量适合做因子分析。

## 因子提取

使用主成分法求解因子载荷矩阵，并基于方差最大化准则进行因子旋转，因子载荷阵如表 3所示，为方便阅读，其中大于0.6（包括接近0.6）的载荷均作了加粗处理。

表 3方差最大化旋转后的因子载荷阵

| 变量/因子 | 现金结构 | 偿债能力 | 成长能力 | 盈利能力 |
| --- | --- | --- | --- | --- |
| 基本每股收益 | -0.496 | -0.085 | 0.454 | 0.260 |
| 净资产收益率 | -0.060 | 0.021 | -0.073 | **0.887** |
| 总资产收益率 | -0.611 | -0.056 | 0.343 | **0.592** |
| 营业利润率 | -0.399 | -0.364 | -0.065 | -0.229 |
| 流动比率 | -0.075 | **0.965** | 0.054 | -0.023 |
| 速动比率 | -0.085 | **0.965** | 0.050 | -0.030 |
| 现金负债比 | **-0.936** | 0.189 | 0.011 | 0.070 |
| 现金流动负债比 | **-0.932** | 0.131 | -0.014 | 0.036 |
| 利润总额同比增长 | -0.036 | -0.037 | **0.722** | -0.144 |
| 总资产同比增长 | 0.025 | 0.316 | **0.644** | 0.123 |
| 资产负债率 | 0.254 | **-0.733** | 0.081 | -0.163 |
| 贡献率 | 0.237 | 0.245 | 0.117 | 0.121 |
| 累积贡献率 | 0.237 | 0.482 | 0.598 | 0.719 |

可以看到四个因子的方差贡献率已经达到了71.9%，可以解释大部分的样本信息，其中第一个因子主要与现金负债比、现金流动负债比相关，可以将其命名为现金结构因子，因子得分越高意味着公司的现金越多，资金利用越不充分；第二个因子主要与流动比率、速动比率和资产负债率有关，可以将其命名为偿债能力因子，因子得分越高意味着公司偿债能力越稳健，具有更强的抗风险能力；第三个因子主要与利润总额同比增长、总资产同比增长有关，可以将其命名为成长能力因子，因子得分越高意味着公司的成长情况越好，公司规模正在不断扩大；最后一个因子主要与净资产收益率、总资产报酬率有关，可以将其命名为盈利能力因子，因子得分越高意味着公司的营收状况越好。

## 综合评价

基于提取出的因子得分对公司绩效进行综合评价，使用各因子方差贡献率除以总贡献率作为因子权重，对各因子得分加权求得综合得分，权重如表 4所示。基于综合得分，表 5给出了排名前十的公司，可以看到亚虹医药居于绝对的领先地位，其后公司的综合得分出现了断崖式下跌，综合各项得分来看，亚虹医药的盈利能力是首屈一指的，但可能由于其规模已经较大，成长能力受限，在现金结构方面，亚虹医药得分较高，说明其现金得到了充分利用，没有过多闲置资金，同时其偿债能力也遥遥领先，说明亚虹医药具有较好的抗风险能力。

表 4因子权重

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子 | 现金结构 | 偿债能力 | 成长能力 | 盈利能力 |
| 权重 | 32.9% | 34.0% | 16.3% | 16.8% |

表 5综合得分前十名公司

| 公司 | 现金结构 | 偿债能力 | 成长能力 | 盈利能力 | 综合得分 |
| --- | --- | --- | --- | --- | --- |
| 亚虹医药 | 16.2256121 | 18.367694 | 0.5787635 | 3.7354773 | 12.305141 |
| 智翔金泰 | 6.0992786 | 10.460305 | 1.5747619 | 0.4305105 | 5.892178 |
| 寒武纪 | 6.6814378 | 11.959232 | -1.0116242 | -1.6189386 | 5.827455 |
| 芯导科技 | -4.3495107 | 15.527874 | 6.7835856 | 1.4467347 | 5.197264 |
| 康希通信 | 9.7406967 | 2.965881 | 1.6186875 | -0.5012810 | 4.392720 |
| 禾迈股份 | -1.2257436 | 7.778322 | 6.4133830 | 1.4390777 | 3.528506 |
| 慧智微 | 9.8260754 | 2.133647 | -1.5865703 | -1.9490863 | 3.372161 |
| 鼎阳科技 | -1.2389861 | 8.417350 | 3.8128518 | 0.8555893 | 3.219506 |
| 东芯股份 | -1.6326521 | 7.942386 | 5.1016371 | 0.3557114 | 3.054595 |
| 西山科技 | -0.3858891 | 7.936586 | 2.5010471 | 0.3682598 | 3.041020 |

# 聚类分析

在因子分析的基础之上，进一步基于提取出的公共因子进行聚类分析。从聚类方法上来说，一般有系统聚类法和动态聚类法两种手段，本文使用动态聚类中的K-means方法进行聚类，首先考察最佳类别数

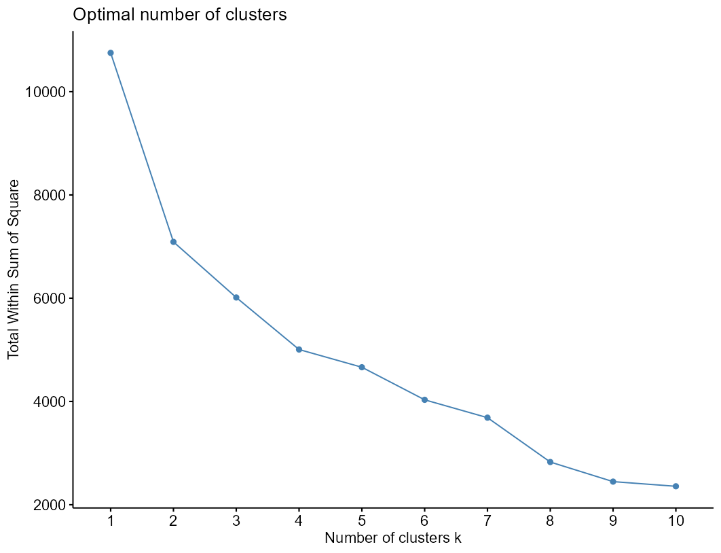


图 1碎石图

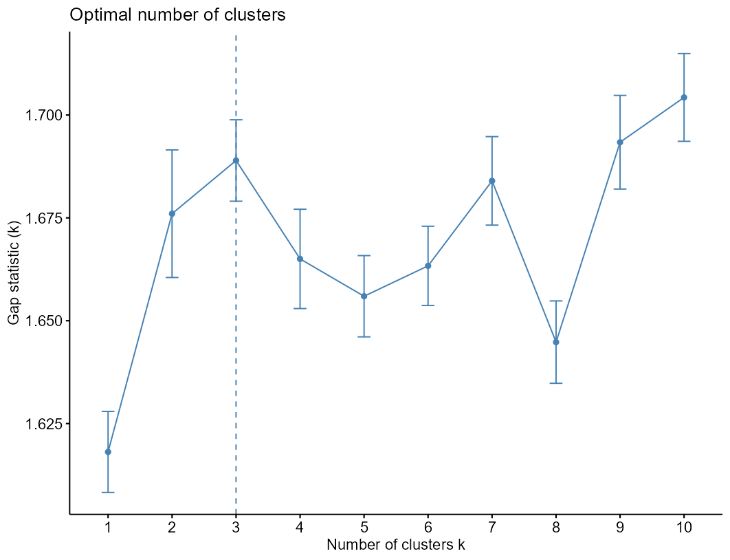


图 2间隔统计量

从图 1和图 2来看，肘形图由陡峭转为平缓的点在9类或10类，而间隔统计量最大的点在10类，考虑到部分极端样本的存在，考虑将样本划分为10类，完整类别情况见附录，各个类别的因子平均得分如表 6所示

表 6聚类后各类别平均因子得分

| 组别 | 现金结构 | 偿债能力 | 成长能力 | 盈利能力 |
| --- | --- | --- | --- | --- |
| 1 | -16.848 | 0.170 | 9.890 | 7.660 |
| 2 | 1.655 | 9.560 | 2.090 | 0.467 |
| 3 | -3.737 | 0.373 | 1.176 | 1.453 |
| 4 | -0.742 | 2.377 | -0.120 | 0.138 |
| 5 | 14.181 | -5.530 | -2.736 | -28.022 |
| 6 | -5.965 | 5.472 | 0.612 | 0.972 |
| 7 | 1.216 | -2.106 | -0.334 | -0.345 |
| 8 | 3.538 | 0.216 | -1.108 | -0.938 |
| 9 | 0.839 | -0.329 | 5.278 | -0.485 |
| 10 | -0.157 | -0.305 | -0.193 | 0.135 |

可以看到其中仅有第二类的企业各项平均得分均为正，且其成长能力居于中间水平，偿债能力居于前列，可以认为该组别或许代表了企业发展已经较为成熟，抗风险能力较强，规模已经较大的企业，检查具体的企业样本发现亚虹医药归属于第二类，这在一定程度上印证了上述推断的合理性。与之相对的是第五类的企业，其除现金结构之外的各项得分均为非常低的水平，事实上该组仅有首药控股一家公司，可以视作样本中的离群点，考虑到该公司其他能力的水平，其较高的现金结构得分或许应该解释为账面流动资金欠缺，负债较高。

基于聚类分析的结果可以对各个组别给出一定建议，如第二组的企业应当继续保持好自己的优势，稳扎稳打，实现长久发展；第五组的企业急需改变自身的经营现状，找到市场定位，优化管理模式，否则极有可能面临退市风险。

# 总结

本报告通过因子分析和聚类分析对2023年披露财务信息的566家科创板上市公司进行了财务绩效评价。研究发现，不同公司在财务指标上表现出显著的异质性，通过因子分析提取了现金结构、偿债能力、成长能力和盈利能力四个公共因子，解释了大部分样本信息，并对各公司进行了综合评分和排名。

基于综合评分，亚虹医药表现出绝对的领先地位，其盈利能力、偿债能力和现金结构均优于其他公司。然而，整体结果显示许多公司在成长能力和盈利能力方面存在显著差异，部分公司面临严重亏损，表现出较大的财务风险。

通过聚类分析，将公司划分为十类，进一步揭示了各类别公司在不同因子上的特征和差异。第二类公司表现出较为成熟的发展态势，具有较强的抗风险能力和良好的成长前景，而第五类公司则表现出较大的财务困境，亟需改进其经营管理模式以避免退市风险。

本研究的结果为政策制定者和企业管理者提供了有价值的参考。政策制定者可以根据不同类别公司存在的问题，制定有针对性的扶持政策，以促进科创板企业的健康发展。企业管理者则应根据自身的财务状况，优化资源配置，提升经营效益，确保企业的可持续发展。

科创板作为中国资本市场的一个重要组成部分，其企业的财务绩效直接影响着市场的稳定与发展。本研究通过科学的方法和严谨的分析，为理解和改善科创板企业的财务绩效提供了新的视角和实证依据。未来的研究可以进一步探讨其他影响企业绩效的因素，如行业特征、市场环境等，以更加全面地评价科创板企业的综合绩效。

# 附录

附录一：原始数据见压缩包中“科创板部分公司财务数据.csv”

附录二： 完整的因子得分及分类结果见压缩包中“因子分析及聚类结果.csv”

附录三：代码文件如下，或可查看压缩包中“期末报告代码.R”

|  |
| --- |
| library(tidyverse)  library(psych)  library(flextable)  library(cluster)  library(factoextra)  #预处理  data <- read.csv("D:/预删除文件夹/大三下/多元统计/科创板部分公司财务数据.csv",encoding = "UTF-8") %>%  `colnames<-`(c("时间","股票代码","简称","基本每股收益","净资产收益率","总资产报酬率",  "营业利润率","流动比率","速动比率","现金负债比","现金流动负债比",  "营业收入同比增长","利润总额同比增长","总资产同比增长","存货周转率",  "应收账款周转率","总资产周转率","资产负债率")) %>%  mutate(across("股票代码",as.numeric)) %>%  na.omit() %>%  group\_by(股票代码) %>%  slice(1) %>%  ungroup() %>%  select(c(1:11,13,14,18))  data\_cal <- data[-c(1:3)]  #描述统计  summary(data\_cal) %>%  as.data.frame() %>%  separate(Freq,into = c("Var","value"),sep = ":") %>%  mutate(across(everything(),str\_remove," ")) %>%  pivot\_wider(names\_from = Var,values\_from = value) %>%  flextable() %>%  save\_as\_docx(path = "临时.docx")  #主成分法函数  fa\_pcm = function(S, k){  p = nrow(S); diag\_S= diag(S); sum\_rank = sum(diag\_S)  rowname = rownames(S); colname = paste('Factor', 1:k, sep='')  A = matrix(0, nrow=p, ncol=k, dimnames=list(rowname, colname))  eig = eigen(S)  for ( i in 1:k)  A[,i] = sqrt(eig$values[i])\*eig$vectors[,i]  h = diag(A%\*%t(A))  rowname = c('SS loadings','Proportion Var', 'Cumulative Var')  B = matrix(0, nrow=3,ncol=k, dimnames=list(rowname, colname))  for (i in 1:k){  B[1,i] = sum(A[,i]^2); B[2,i] = B[1,i]/sum\_rank  B[3,i ] = sum(B[1, 1:i])/sum\_rank  }  A = round(A,3); B = round(B,3); h = round(h,3)  diag\_S = round(diag\_S,3); method = c('Principal Component Method')  list(method=method,loadings=A,var=cbind(common=h,specific=diag\_S-h),B=B)  }  #因子分析  ##可行性检验  KMO(data\_cal)  cortest.bartlett(cor(data\_cal),566)  ##因子得分  a <- fa\_pcm(cor(data\_cal),4)$loadings %>%  varimax(normalize = F) %>%  .$loadings  a[1:11,1:4] %>%  as.data.frame() %>%  round(3) %>%  flextable() %>%  save\_as\_docx(path = "临时.docx")  ##综合评价  scores <- as.matrix(apply(data\_cal,2,scale)) %\*% as.matrix(a[1:11,1:4]) %>%  data.frame(data[3],.) %>%  `colnames<-`(c("公司","现金结构","偿债能力","成长能力","盈利能力")) %>%  mutate(综合得分=现金结构\*0.329+偿债能力\*0.34+成长能力\*0.163+盈利能力\*0.168) %>%  arrange(desc(综合得分))  scores[1:10,] %>%  flextable() %>%  save\_as\_docx(path = "临时.docx")  #聚类  fviz\_nbclust(scores[2:5],kmeans,method = "wss")  fviz\_nbclust(scores[2:5],kmeans,method = "gap")  scores <- data.frame(scores,group=kmeans(scores[c(2:5)],10)$cluster)  ##聚类结果  scores %>%  group\_by(group) %>%  summarise(现金结构=mean(现金结构),  偿债能力=mean(偿债能力),  成长能力=mean(成长能力),  盈利能力=mean(盈利能力),  偿债能力=mean(偿债能力)) %>%  round(3) %>%  flextable() %>%  save\_as\_docx(path="临时.docx")  scores %>%  filter(group==5)  #完整结果  write.csv(scroes,"因子分析及聚类结果.csv",fileEncoding = "GBK",row.names = F) |