**云南大学数学与统计学院**

**实验报告**

**实验课名称： 应用多元统计分析实验**

**指导教师： 李会琼**

**专业（年级）： 统计学2021级**

**学生姓名： 枫叶 学号:**

**实验名称： 实验13**

**实验成绩：**

**《应用多元统计分析实验》实验报告 13**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验名称 | 多元正态分布均值向量的检验 | | 实验成绩 |  | |
| 学号 |  | | 姓名 | 枫叶 | |
| 实验时间 | 2024年6月21日 | 实验地点 | 格物楼3508 | 指导教师 | 李会琼 |
| 1. **实验目的**   学习使用R软件进行因子分析   1. **实验要求**   1. 对所使用的方法与所得到的结果进行适当的文字描述。  2. 在实验结果的相应部分附上完整的代码与适当的注释。  3. 采用一定的可视化方法体现出对应计算结果。   1. **实验内容**   企业绩效评价   1. **实验软件**   R语言   1. **实验结果**  一、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 二、因子分析 （1）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 1 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 1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。  （2）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 2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 2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。  （3）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 3所示。基于综合得分，表 4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 3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 4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 |   代码：  library(tidyverse) library(psych) library(flextable) #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)] #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") | | | | | |