

Research on the Risk of Individual Stock Investment -- Take the Brokerage Sector as an Example

Zheng Ma^{1,2,3}, Hongmei Zhang^{1,2,3}, Zhiqiang Chen^{1,2,3}, Ziyang Wang^{1,3}

¹Guizhou Institution for Technology Innovation and Entrepreneurship Investment, Guizhou University of Finance and Economics, Guiyang Guizhou 550025, China

²Guizhou Institution for Urban Economy and Development, Guizhou University of Finance and Economics, Guiyang Guizhou 550025, China

³School of Finance, Guizhou University of Finance and Economics, Guiyang Guizhou 550025, China

个人股票投资的风险研究——以券商板块为例

马征^{1,2,3}, 张红梅^{1,2,3}, 陈志强^{1,2,3}, 王资燕^{1,3}

¹ 贵州财经大学贵州科技创新创业投资研究院, 贵州贵阳 550025, 中国

² 贵州财经大学贵州城镇经济与发展研究院, 贵州贵阳 550025, 中国

³ 贵州财经大学金融学院, 贵州贵阳 550025, 中国

Abstract

Today, Invest in the stock market has increasingly perfect continuous development, individual investors has gradually become the main force of stock investment, but due to various constraints, individual investors for face of massive investment information, they cannot identify the valuable investment intelligence information, choice of the value of the investment way are different degrees of defects and individual investors in the face of abnormal fluctuations in the stock market's risk bearing capacity is low. In order to identify individual investors in securities investment risk, this paper selects Shanghai Stock Exchange traded 15 brokerage stocks as a sample, using SPSS factor analysis and cluster analysis methods sample stock ranking and grouping, so as to identify the stock risk and for individual investors in securities investment provide relevant reference.

Keywords: Stock Investment Risk; Cluster Analysis; Factor Analysis

摘要

在股票投资市场日益完善不断发展的今天,个人投资者逐渐成为了股票投资的主力军,但是个人投资者在面对海量的投资信息时无法甄别出其中有价值的投资情报信息,选择的价值投资方式存在不同程度上的缺陷,并且个人投资者在面对股市异常波动时的风险承受能力偏低。为了识别个人投资者证券投资中的风险,本文选取上海证券交易所挂牌交易的15只券商股票作为样本,运用SPSS软件中的因子分析及聚类分析的方法将样本股票排名、分组,从而识别股票的风险,进而为个人投资者证券投资提供相关参考。

关键词: 股票投资风险; 聚类分析; 因子分析

1. 引言

九十年代以来,总体上我国的股票市场出现了两次规模比较大、持续时间比较长的异常波动状况:第一次是在2005年6月,股指从底部998点一路飙升至2007年10月份的6124.04点,最大涨幅达到了513.6%,随后股票指数在2008年10月急剧下跌至1664.93点;第二次是在2014年11月至2015年6月。也就是说,股票指数在短短半年的时间里,总涨幅达到了110%。然后股票指数于2015年7月一度下探至3373.54点,跌幅达到35%;创业板则下探至2304.76点。作为股票市场中基数较大的个人投资者——个人投资者占到中国股市成交量的80%至90%,在上述两次的异常波动中尤其是在下跌阶段损失惨重,严重套牢。投资股票的较大不确定因素,这使得需要合理利用数据分析的方式来分析数据,确保能够从泛滥的数据中提取有效的投资信息以规避投资风险获得收益。但是,股市具有变化无常、涨跌莫测的特点,加之每个交易日每只股票都会产生庞大的交易数据和财务数据,想要在这些繁琐的数据中获取对个人投资者真实有效的数据其困难是巨大的。所以,本文希望通过以券商股票板块为例,根据需要充分研究和分析上市公司的发展前景、历史业绩等相关财务数据,在聚类分析和因子分析方法的基础上,采用SPSS.13软件分析,找到一种合理判断上市公司股票价值,规避个人股票投资风险的方法。

2. 文献综述

统计方法应用于证券分析的文章已屡见不鲜。毛哲敏:在《房地产上市公司资本结构影响因素实证分析》(2009)一文中,以上市公司为研究对象,运用主成分分析法,对我国民营上市公司的资本结构影响因素进行了实证分析。李艳在《房地产上市公司股价影响因素实证分析》(2012)一文中选取了房地产上市公司股票价格指数为因变量,选择上证综指、财务业绩、CPI、人民币贷款利率、房地产价格指数五个因素指标为自变量进行了多元回归统计建模,检验结果表明财务业绩对上市公司股价的影响显著。吕美艳在《房地产上市公司财务指标的因子分析》(2011)一文中,

选取以万科A、保利地产、招商地产等共计75家沪深两市的上市房地产公司,采用20个财务指标,对上述房地产上市公司股价波动进行因子分析,其结果显示房地产上市公司股价波动与财务业绩具有正相关的关系。技术分析大师艾略特在1938年根据这一发现他提出了一套相关的市场分析理论,精炼出市场的13种形态(Pattem)或波浪(Waves),在市场上这些形态重复出现,但是出现的时间间隔及幅度大小并不一定具有再现性。而后他又发现了这些呈结构性形态之图形可以连接起来形成同样形态的更大图形。张祖国:在《投资分析》(2014)一书中指出基本面分析主要从国家的宏观环境及公司的经营情况几方面来考察公司的股票价值情况。从以上来看,以往的研究主要是对理论方面的研究,本文试图从量化的角度来分析证券的投资价值。

3. 数据和指标选取

3.1. 数据选取

本文选择在上海证券交易所挂牌上市的券商板块的股票进行分析,排除了B股、H股、数据不可得和有特殊性质的股票后,进而得到样本股票15只。所以选择券商板块股票进行研究,是因为券商的主营业务比较单一,利润来源比较集中,影响收入和利润的因素比较少;还有绝大部分的个人投资者进行股票投资时都是通过券商进入股票交易系统进行交易的,之所以券商板块的股票对个人投资者来讲比较熟悉,数据的选取具有可得性,全面性。数据来源东方财富股票交易软件,始数据截止到2016年3月31日的。

3.2. 指标选取

目前在评价上市公司基本面状况时,处于核心地位的财务数据指标是上市公司的盈利能力、经营能力、偿债能力和成长能力。由此,本文在上述15只股票的财务数据中选取7个最能反映这些能力的指标。

(1) 盈利能力:是反映上市公司资金增值的能力,盈利能力的大小决定了股票价值的高低,反映盈利能力的指标有:

①净资产收益率=净利润/平均净资产,反映上市公司资本的盈利能力,指上市公司的所有者

通过投入资本在生产经营过程中所取得利润的能力。该指标是评价上市公司自有资本及其积累所获得报酬水平的综合指标；

②每股收益=净利润/总股数，该比率反映了每股普通股创造税后利润的能力。

(2) 偿债能力：是上市公司偿还到期债务（包括本息）的能力，反映上市公司资产的安全性，反映偿债能力的指标有：

③资产负债率=负债总额/资产总额，反映了上市公司的资本结构，常被用于考察企业的长期偿债能力。

④资产总计（总资产）=期末资产总计余额，反映上市公司实际拥有或者控制的资产。

(3) 成长能力：成长性好的上市公司在总资产扩张能力，股本扩张能力，资产重组方面都比较强，反映成长能力的指标有：

⑤营业收入=本年主营业务的总收入。该指标反映上市公司从本身主要经营的业务中获取资金的能力以及未来的成长能力。

⑥净利润=本年净利润，该比率越高表明上市公司效益越好，成长能力越强。

⑦营业利润=本年营业收入-本年年初营业成本，该比率反映上市公司扩大市场规模的能力

4. 实证分析

4.1. 聚类分析概述

聚类分析就是对抽象对象进行集合分类，其结果是由具有某些相似特征的类似对象构成多个分类。在聚类分析过程中生成的簇是一种数据对象组合，同一簇中的对象相似性很高，而在不同簇中的对象之间差异很大。计算差异程度的依据是被描述对象的某种实际属性。聚类分析方法其实就是把给定的数据集合划分成不同簇的集合，即在某一空间中规定数据库中可以获得有限集合或者获得有限取样点集合。聚类分析的主要目标是把具有相关性的数据聚集成一类，使每一类之间的相似性最小，而类内数据的相似性最大。

聚类分析目前已经在众多领域中得到运用，而且能够在许多应用中结合其他的分析技术。运用聚类分析方法可以准确识别出稀疏区域和密集区域，而且可以在聚类过程中得到整体的布局方式和数据间固有的联系。聚类分析的基本特征：分析方式简单直观。聚类分析在研究性分析上应用较多，能从不同的角度理解分析聚类结果，最后在不同分析结果中找到与实际情况最相符的结果。在进行分析的过程中聚类变量其实就是聚类分析最后得到的解，增减变量个数，都会影响最后聚类的结果。

4.2. 聚类分析

通过以上对聚类分析的阐述，结合聚类分析的特点：事先并不知道根据什么分类，分成几类。本文通过 SPSS 软件对样本数据进行系统聚类分析，在聚类分析中，首先对数据进行无量纲化处理，其次进行标准化处理，其次用 SPSS 软件中的分析中选择分类按钮对数据进行分类。标准化处理结果如表 1 所示：

表 1 凝聚顺序表

编号	簇集 1	簇集 2	系数	簇集 1	簇集 2	下阶段
1	2	10	.211	0	0	3
2	3	9	.418	0	0	3
3	2	3	.575	1	2	6
4	11	12	.847	0	0	6
5	7	15	.952	0	0	9
6	2	11	1.060	3	4	10
7	1	4	1.105	0	0	12
8	5	6	1.279	0	0	11
9	7	8	1.455	5	0	10
10	2	7	1.826	6	9	11
11	2	5	5.203	10	8	13
12	1	13	9.981	7	0	13
13	1	2	21.843	12	11	14
14	1	14	41.727	13	0	0

表 1 输出结果第一行数据的含义是在聚类分析的第一步中，2 号个案与 10 号个案聚成一类，样本离差平方和是 0.211，这个小类将会在聚类分析的第 3 步用到。同理，第 3

Risk Analysis and Crisis Response in Big Data Era (RAC-16)

行数据的含义是在聚类分析的第三步中，3号个案与第一步聚成的小类又聚成另外一小类，他们的离差平方和是 0.575，这个小类将在聚类分析的第六步用到。可见，经过 14 步聚类过程，15 个个案最后聚成了一大类。根据上述聚类分析的方法与过程，通过 SPSS 软件可以将 15 个个案根据离差平方和分为 4 类，具体的分类情况如表 2 所示。

表 2 分类簇集表

1:中信证券	1	9:兴业证券	2
2:国金证券	2	10:东吴证券	2
3:西南证券	2	11:东兴证券	2
4:海通证券	1	12:东方证券	2
5:招商证券	2	13:国泰君安	3
6:华泰证券	2	14:安信信托	4
7:光大证券	2	15:国投安信	2
8:方正证券	2		

由表2可以看出，第一类上市公司包含国金证券、方正证券、东吴证券、西南证券、兴业证券、东兴证券、东方证券、光大证券、国投安信，这一类公司业绩一般，表现平平，是市场中的主体部分，占据市场主体；第二类公司对应的股票有招商证券、华泰证券，这一类公司比第一类表现稍好一些；第三类公司包含海通证券、中信证券、国泰君安，这一类股票是市场上表现较好的股票，风险相对较低；第四类上市公司包含安信信托，这一类股票是在市场上表现较差的一类股票。

5. 因子分析

虽然在以上部分对 15 只股票进行了分类，但是在实际投资过程中，只是进行分类是不够的，比较片面。为了能够进一步了解是什么因素造成了股票之间的不同市场表现，下面运用因子分析，选取 7 个常用财务指标，也是投资者很容易获取的几个指标。从中找出具有主导影响力因素。

5.1. 因子分析概述

因子分析模型的建立步骤大致为：

第一，建立指标体系，构成原始矩阵，保证全部指标同向化对样本数据进行标准化处理得到标准化矩阵，计算变量的简单相关系数矩阵。第二，解特征方程 $|R - \lambda E| = 0$ ，计算相关矩阵的特征值 λ ，若 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ ，以 $\lambda_p \geq 1$ 为标准确定公共因子个数 P。第三，计算初始因子载荷矩阵和公共因子方差，用正交或斜交旋转的方法求得正交或斜交因子载荷矩阵；根据正交或斜交因子载荷矩阵相关系数绝对值，确定并命名公共因子。最后，计算公因子得分和最后得分。

5.2. 因子分析

对原始数据进行 KMO 和 Bartlett 检验，其结果如表 3 所示：

表 3 KMO 和 Bartlett 检验

Kaiser-Meyer-Olkin:取样足够度	.679
Bartlett 的球形检验近似卡方	153.188
df	21
显著性	.000

其中 KMO 值为 0.679，表示比较适合进行因子分析；Bartlett 的球形检验的原假设为相关系数矩阵为单位阵，结果显示显著性水平为 0.000，小于显著性水平 0.05，因此拒绝原假设，说明变量之间存在相关关系，适合做因子分析。

表 4 因子贡献率

成分	初始特征值			提取平方和载入			旋转平方和载入		
	总计	差的%	累计%	总计	差的%	累计%	总计	差的%	累计%
1	3.93	56.23	56.237	3.93	56.23	56.23	3.93	56.19	56.19
2	2.59	37.10	93.338	2.59	37.10	93.33	2.60	37.14	93.33
3	.260	3.712	97.050						
4	.130	1.862	98.912						
5	.055	.788	99.700						
6	.019	.272	99.972						
7	.002	.028	100.00						

Risk Analysis and Crisis Response in Big Data Era (RAC-16)

表 4 给出了因子贡献率的结果。左侧部分为初始特征值，中间为提取主因子结果，右为旋转后的主因子结果。其中前两个因子的特征值大于 1，并且前两个因子的特征值总和占总特征值的 98.912%，因此，提取前两个子作为主因子。

表 5 旋转因子负荷矩阵

	元件	
	1	2
每股收益	.433	.840
净资产收益率 (%)	-.157	.973
营业收入 (亿)	.969	-.036
营业利润 (亿元)	.971	.098
净利润 (亿元)	.991	.062
总资产 (亿元)	.909	-.254
资产负债比率 (%)	.180	-.932

表 5 的二、三列分别是两个征值所对应的特征向量。可以看到表中的七指标：营业收入、营业利润、净利润以及总资产在第一公因子上具有正向负荷，说明可以将第一个主因子命名为“盈利因子”，主要反映上市公司获取收入，赢的利润的状况。每股收益、净资产收益率在第二公因子上具有较大的正向负荷，说明可以将第二个公因子命名为“收益因子”，主要反映上市功能是公司股东获得收益的状况。

表 6 成分得分系数矩阵

	成分	
	1	2
每股收益	.115	.326
净资产收益率 (%)	-.034	.373
营业收入 (亿)	.246	-.008
营业利润 (亿元)	.248	.044
净利润 (亿元)	.252	.030
总资产 (亿元)	.230	-.092
资产负债比率 (%)	.040	-.358

由表 6 可以得到：

$F1=0.115zscore(\text{每股收益})+0.246zscore(\text{营业收入})+0.248zscore(\text{营业利润})+0.252zscore(\text{净利润})+0.230zscore(\text{总资产})+0.040zscore$

$(\text{资产负债率})-0.034zscore(\text{净资产收益率})$
 $F2=0.326zscore(\text{每股收益})+0.373zscore(\text{净资产收益率})+0.044zscore(\text{营业利润})+0.030zscore(\text{净利润})-0.008zscore(\text{营业收入})-0.092zscore(\text{总资产})-0.358(\text{资产负债率})$

标准化之后的数据代入上述方程就可以计算出每一只股票的各个因子的评价得分，然后以各因子方差贡献率占两个公因子总方差贡献率比重作为权重进行加权汇总，得出每只股票的综合得分，根据综合得分可以获得上市公司的排名，如表 7：

表 7 综合排名

排名	名称	综合得分
1	中信证券	4.7100
2	海通证券	4.5321
3	国泰君安	4.0322
4	招商证券	3.9612
5	光大证券	3.5543
6	西南证券	3.2100
7	华泰证券	3.002
8	兴业证券	2.9623
9	东方证券	2.9510
10	方正证券	2.6336
11	东吴证券	2.4897
12	东兴证券	2.0156
13	国投安信	1.9991
14	国金证券	1.8752
15	安信信托	1.0211

通过综合实力得分可以得出的结论：第三类股票中中信证券、海通证券和国泰君安综合实力得分最高，进行股票投资的时候可以买入，风险相对较低；第二类股票中招商证券和华泰证券得分比较高，代表这两只股票在买入时需加以考虑分析其收益的情况，风险相对第三类较高一些；第一类股票的综合实力得分排名占据主体位置，表明国金证券、方正证券、东吴证券、西南证券、兴业证券、东兴证券、东方证券、光大证券和国投安信在市场的表现一般，风险比较高，在选择购买时多加权衡；第四类股票，安信信托得分最低，表明这只股票在市场当中的表现不好，再进行投资的时候尽量规避。

6. 总结

基于 SPSS 软件的因子分析的方法可以有效地减少在个人投资上市公司股票时需要考虑的指标数目,聚类分析可以清晰的将具有不同因子特征的上市公司分类,找出适合个人投资者投资风险承受的股票。这两类方法的优点在于:

(1)可以浓缩上市公司公开的财务指标,用很少的因子代替繁杂的指标种类,使个人投资者对上市公司有一个简明、清晰的认识:

(2)个人投资者可以清晰的看到每个因子下每只股票的得分排名情况以及综合得分排名情况,便于比较上市公司在经营管理能力等各方面的优劣;

(3)投资者可以找到适宜达到自己投资决策的目标类型的股票来权衡自己的投资风险。

7. 参考文献

- [1]何晓群.多元统计分析.北京:中国人民大学出版社,2012.
[2]范坤,冯长焕.因子分析中指标数据如何正

确预处理.财会月刊:2013(003):85-88.

- [3]曹裕,陈晓红,王傅强.我国上市公司生命周期划分方法实证比较研究.系统管理学报,2010(3):313-322.
[4]陈琦.聚类分析和判别分析在股票投资中的应用.中国市场,2011,26:69-72.
[5]邓秀勤.聚类分析在股票市场板块分析中的应用.数理统计与管理,1999,5(18):(1-4).
[6]李德荣,何莉敏,李玉.聚类分析和因子分析在股票投资中的应用.内蒙古统计,2011,01:29-31.
[7]李艳.房地产上市公司股价影响因素实证分析.特区经济,2012,09:84-86.
[8]吕美艳.房地产上市公司财务指标的因子分析.现代商业,2011,36:209.
[9]毛哲敏.房地产上市公司资本结构影响因素实证分析.商场现代化,2009,18:80-81.
[10]J.K Yu, Z.F. Zhou. The errors estimate of the multistage combined investment risk assessment. Journal of Risk Analysis and Crisis Response, 2011 (1): 106-109.