

战略性新兴产业 企业信用风险评价体系研究

Study on the Credit Risk Evaluation System
of Enterprises in Strategic Emerging Industry

张 目 李 伟 著



科学出版社

科学研究规划基金项目“战略性新兴产业企业信用风险评价
项目编号:11YJA630196)

战略性新兴产业企业信用风险 评价体系研究

张 目 李 伟 著

科 学 出 版 社

北 京

内 容 简 介

本书基于教育部人文社会科学研究规划基金项目,以战略性新兴产业企业信用风险评价体系为研究对象,共分为14章,对战略性新兴产业的发展现状及区域分布、银行与战略性新兴产业企业的信贷行为、战略性新兴产业企业信用风险影响因素、战略性新兴产业企业信用风险评价指标体系及指标赋权、战略性新兴产业企业信用风险的组合评价模型、战略性新兴产业企业信用风险的分行业动态评价、不确定性方法在战略性新兴产业企业信用评价中的应用等问题展开研究,并得出相关研究结论与建议。

本书可供管理学、应用经济学等相关学科的高校师生,产融结合及信用风险管理领域的科研工作者,战略性新兴产业管理者,银行金融机构管理者,以及对科技金融领域感兴趣的读者阅读参考。

图书在版编目(CIP)数据

战略性新兴产业企业信用风险评价体系研究 / 张目, 李伟著. —北京: 科学出版社, 2018.7

ISBN 978-7-03-051044-0

I. ①战… II. ①张… III. ①高技术企业-中小企业-企业信用-贷款风险管理-研究-中国 IV. ①F279.244.4 ②F832.42

中国版本图书馆 CIP 数据核字 (2016) 第 303843 号

责任编辑: 莫永国 / 责任校对: 彭 映 武雯雯

责任印制: 罗 科 / 封面设计: 墨创文化

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

四川煤田地质制图印刷厂印刷

科学出版社发行 各地新华书店经销

*

2018年7月第 一 版 开本: 787×1092 1/16

2018年7月第一次印刷 印张: 10.25

字数: 260千字

定价: 79.00元

(如有印装质量问题, 我社负责调换)



前 言

战略性新兴产业是以重大技术突破和重大发展需求为基础,对经济社会全局和长远发展具有重大引领带动作用,知识技术密集、物质资源消耗少、成长潜力大、综合效益好的产业。加快培育和发展战略性新兴产业对推进我国现代化建设具有重要的战略意义。加快培育和发展战略性新兴产业,必须健全财税金融政策支持体系,加大扶持力度,引导和鼓励社会资金投入。其中,引导金融机构建立适应战略性新兴产业特点的信贷管理和贷款评审制度,是鼓励金融机构加大信贷支持力度的重要举措之一。

战略性新兴产业企业(指符合战略性新兴产业发展方向的企业)具有较强的扩张能力,对资金的需求巨大,但由于受到技术生命周期、技术成熟度、市场容量、经营管理能力、技术创新能力、区域性产业聚集度、区域性产业配套能力、区域投资环境等因素的影响,其未来发展和经营业绩等具有较大的不确定性。因此,战略性新兴产业企业在成长过程中,必然伴随着不同程度的信用风险。如何构建科学的战略性新兴产业企业信用风险评价体系,是金融机构建立适应战略性新兴产业特点的信贷管理和贷款评审制度的基础性和关键性课题。

本书以多属性综合评价为逻辑主线,既吸收了国内外信用风险评价的先进技术与方法,又结合了国内战略性新兴产业的发展现状及特征,以战略性新兴产业企业信用风险评价体系为研究对象,较为系统地研究了银行与战略性新兴产业企业的信贷行为、战略性新兴产业企业信用风险影响因素、战略性新兴产业企业信用风险评价指标体系及指标赋权、战略性新兴产业企业信用风险的组合评价模型、战略性新兴产业企业信用风险的分行业动态评价、不确定性方法在战略性新兴产业企业信用评价中的应用等问题。

本书在战略性新兴产业企业信用风险评价指标体系中引入了技术创新能力、产品市场前景、区域经济发展水平等一级指标,突出了战略性新兴产业特点;在指标赋权中,以熵权法为基础,引入基于相对熵的组合赋权法,较好地兼顾了主观偏好和客观信息;在信用风险评价方法上,采用正态云模型、可变模糊集理论、模糊积分、灰色关联投影法、未确知测度等不确定多属性综合评价方法,体现了信用风险评价的模糊性、随机性、动态可变性、灰色性或未确知性;同时,引入组合评价法,提高了综合评价的全面性、科学性和合理性。本书注重理论联系实际,对各种方法的原理只作简要介绍,对计算步骤的介绍则力求详尽,突出了实用性和可操作性。本书集中体现了不确定多属性综合评价方法在信用风险评价领域的应用,内容新颖、丰富,可为学界后续研究和业界实际操作提供参考。

在本书的撰写过程中,研究生李伟参与完成了第4、5、10章,研究生匡暑炎、闫慧杰、李岩、张珂、陈华丰、董礼做了大量的基础性工作,科学出版社对本书的编写和

出版给予了大力支持，莫永国编辑为本书做了大量的组织工作，在此一并表示感谢！由于作者水平有限，书中难免出现疏漏和错误，希望广大读者提出宝贵意见，以便进一步修改完善。



2016 年 12 月于贵阳

目 录

第 1 章 绪论	1
1.1 研究背景、意义及目的	1
1.2 国内外研究现状	2
1.2.1 信用风险的概念与成因	2
1.2.2 传统的信用风险评价方法	3
1.2.3 多元统计分析方法	4
1.2.4 人工智能方法	11
1.2.5 战略性新兴产业企业信用风险评价相关研究	17
1.3 研究内容与技术路线	18
1.3.1 研究内容	18
1.3.2 技术路线	19
1.4 主要创新点	20
第 2 章 战略性新兴产业的发展现状分析及区域分布研究	22
2.1 引言	22
2.2 战略性新兴产业统计分类目录的重构	23
2.2.1 战略性新兴产业的外延	23
2.2.2 战略性新兴产业统计分类目录	24
2.3 战略性新兴产业的发展现状研究	26
2.3.1 指标选取及数据来源	26
2.3.2 战略性新兴产业的发展现状	26
2.4 战略性新兴产业的区域分布研究	29
2.4.1 区位熵理论	29
2.4.2 战略性新兴产业区位熵的计算	29
2.4.3 战略性新兴产业的区域分布特征	31
2.5 本章小结	31
第 3 章 银行与战略性新兴产业企业信贷行为的演化博弈分析	33
3.1 引言	33
3.2 演化博弈理论的基本概念	34
3.2.1 基本概念的形成	34
3.2.2 演化稳定策略	34
3.2.3 复制者动态	35
3.3 演化博弈模型的构建	36

3.3.1	模型假设	36
3.3.2	支付矩阵	36
3.3.3	演化稳定策略分析	37
3.4	模型的稳定性分析	38
3.5	本章小结	40
第4章	战略性新兴产业企业信用风险影响因素研究	41
4.1	引言	41
4.2	战略性新兴产业的内涵及特征分析	42
4.2.1	战略性新兴产业的内涵	42
4.2.2	战略性新兴产业的特征	42
4.3	战略性新兴产业企业信用风险的定义及违约行为分析	44
4.3.1	战略性新兴产业企业信用风险的定义	44
4.3.2	战略性新兴产业企业的违约行为分析	45
4.4	战略性新兴产业企业信用风险影响因素分析	47
4.4.1	财务因素分析	47
4.4.2	非财务因素分析	48
4.5	实证检验	50
4.5.1	因子分析的数学模型	50
4.5.2	因素变量的选取与数据来源	50
4.5.3	因子分析过程与结果分析	51
4.6	本章小结	53
第5章	战略性新兴产业企业信用风险评价指标体系构建	55
5.1	引言	55
5.2	评价指标的选取原则	56
5.3	战略性新兴产业企业信用风险评价指标体系	57
5.3.1	指标体系的构成	57
5.3.2	指标的具体含义	58
5.4	评价指标的筛选	62
5.4.1	评价指标的相关性分析	63
5.4.2	评价指标的鉴别力分析	64
5.5	本章小结	65
第6章	基于信息熵的战略性新兴产业企业信用风险评价指标赋权研究	66
6.1	引言	66
6.2	信息熵的基本原理	67
6.2.1	信息熵	67
6.2.2	离散型随机变量概率分布的熵	68
6.2.3	连续型随机变量概率分布的熵	68
6.3	熵权法的计算步骤	69
6.3.1	建立决策矩阵	69

6.3.2	指标数据标准化	69
6.3.3	确定指标权重	70
6.4	实证分析	70
6.4.1	样本与原始数据	70
6.4.2	指标熵权的计算	72
6.5	本章小结	73
第 7 章	基于正态云模型的战略性新兴产业企业信用评价研究	74
7.1	引言	74
7.2	正态云理论	75
7.2.1	正态云及其数字特征	75
7.2.2	正向正态云发生器	76
7.3	基于正态云模型的战略性新兴产业企业信用评价模型	76
7.4	实证分析	77
7.4.1	指标体系与样本数据	77
7.4.2	指标权重的确定	78
7.4.3	模糊综合评价	78
7.5	本章小结	81
第 8 章	基于相对熵和可变模糊集理论的战略性新兴产业企业信用评价研究	83
8.1	引言	83
8.2	战略性新兴产业企业信用评价模型	84
8.2.1	基于相对熵的组合权重计算	84
8.2.2	计算待评企业 u 对级别 h 的指标相对隶属度	85
8.2.3	计算待评企业 u 对级别 h 的综合相对隶属度	86
8.2.4	计算待评企业 u 的级别特征值	86
8.3	实证分析	87
8.3.1	指标体系与样本数据	87
8.3.2	基于相对熵的组合权重计算	87
8.3.3	可变模糊评价	88
8.4	本章小结	90
第 9 章	战略性新兴产业企业信用风险的组合评价模型研究	91
9.1	引言	91
9.2	模糊 Borda 法	91
9.3	实证分析	92
9.3.1	指标体系与样本数据	92
9.3.2	指标权重的确定	93
9.3.3	正态云模型评价结果	93
9.3.4	可变模糊评价结果	94
9.3.5	组合评价	94
9.4	本章小结	95

第 10 章 战略性新兴产业企业信用风险的分行业动态评价研究.....96

10.1 引言96

10.2 动态综合评价模型的构建97

10.2.1 模糊测度与 Choquet 模糊积分97

10.2.2 语意变量与正三角模糊数98

10.2.3 动态综合评价模型98

10.3 实证分析99

10.3.1 指标体系99

10.3.2 样本数据选取99

10.3.3 动态综合评价过程101

10.3.4 结果分析与比较102

10.4 本章小结105

第 11 章 灰色系统理论在战略性新兴产业企业信用评价中的应用研究106

11.1 引言106

11.2 战略性新兴产业企业信用评价的灰色关联投影模型107

11.2.1 建立决策矩阵107

11.2.2 确定指标权重108

11.2.3 决策矩阵的标准化处理108

11.2.4 确定加权灰色关联决策矩阵108

11.2.5 确定灰色关联投影值109

11.2.6 战略性新兴产业企业信用综合评分110

11.2.7 战略性新兴产业企业信用评价步骤110

11.3 实证分析111

11.4 本章小结113

第 12 章 Sugeno 模糊积分在战略性新兴产业企业信用评价中的应用研究114

12.1 引言114

12.2 模糊测度与模糊积分115

12.3 基于 Sugeno 模糊积分的企业信用评价模型116

12.4 基于 Sugeno 模糊积分的企业信用评价步骤117

12.5 实例分析117

12.6 本章小结120

第 13 章 未确知测度在战略性新兴产业企业信用评价中的应用研究121

13.1 引言121

13.2 企业信用评价的未确知测度模型122

13.2.1 单指标未确知测度122

13.2.2 指标权重的确定122

13.2.3 多指标综合测度评价矩阵123

13.2.4 置信度识别准则124

13.2.5 排序124

13.3 应用实例 124

13.3.1 指标体系与样本数据 124

13.3.2 指标权重的确定 125

13.3.3 战略性新兴产业企业信用评价 126

13.4 本章小结 127

第 14 章 MLP 神经网络在战略性新兴产业企业信用评价中的应用研究 128

14.1 引言 128

14.2 指标体系与样本数据 129

14.3 判别分析法的分类结果与分析 133

14.3.1 Fisher 判别函数的生成 133

14.3.2 分类结果与分析 133

14.4 MLP 神经网络模型的分类结果与分析 135

14.4.1 MLP 神经网络模型 135

14.4.2 分类结果与分析 136

14.5 本章小结 137

参考文献 139

附录 152

第1章 绪 论

1.1 研究背景、意义及目的

国际金融危机的爆发使世界各国都认识到战略性新兴产业发展的重要性。战略性新兴产业是引导未来经济社会发展的重要力量。发展战略性新兴产业已成为世界主要国家抢占新一轮经济和科技发展制高点的重大战略。我国正处在全面建设小康社会的关键时期，必须按照科学发展观的要求，抓住机遇，明确方向，突出重点，加快培育和发展战略性新兴产业^[1]。战略性新兴产业是以重大技术突破和重大发展需求为基础，对经济社会全局和长远发展具有重大引领带动作用，知识技术密集、物质资源消耗少、成长潜力大、综合效益好的产业。加快培育和发展战略性新兴产业对推进我国现代化建设具有重要的战略意义^[1]。

2009年11月3日，时任国务院总理温家宝在向首都科技界的讲话《让科技引领中国可持续发展》中，完整表述了大力发展战略性新兴产业、争夺经济科技制高点、努力实现经济社会可持续发展的战略构想，并提出了现阶段我国重点发展的战略性新兴产业主要包括节能环保、新一代信息技术、生物、高端装备制造、新能源、新材料和新能源汽车七大产业^[2, 3]。在政策的鼓励和扶持下，我国战略性新兴产业呈现良好的发展态势，产业规模持续扩大，一批关键性技术取得突破，对经济发展的拉动作用不断增强，支柱产业地位日益凸显。2010年末，战略性新兴产业增加值已占国内生产总值(gross domestic product, GDP)的4%。2012年，国家信息中心经济预测部分析预测显示^[4]，我国节能环保产业成长迅速，“十二五”将迎来重要发展机遇期；新一代信息技术产业的国民经济支柱性产业地位日益凸显；生物产业发展规模初具，市场前景广阔；高端装备制造业发展状况参差不齐，但潜力可观；新能源产业总体规模已居世界前列，未来将继续快速增长；新材料产业应用前景广阔，行业将保持平稳增长；新能源汽车产业市场逐步成熟，“十二五”时期将驶入快车道。从发展趋势看，“十二五”期间我国七大战略性新兴产业均面临着良好机遇，发展前景广阔，但不同行业由于政策支持、产业基础、市场应用等方面情况各不相同，发展状况将出现一定程度的分化，表现出各自不同的发展趋势和特征。《“十二五”国家战略性新兴产业发展规划》(2012)提出^[5]，到2015年，战略性新兴产业增加值占国内生产总值比重达到8%左右，对产业结构升级、节能减排、提高人民健康水平、增加就业等的带动作用明显提高。到2020年，力争使战略性新兴产业成为国民经济和社会发展的重要推动力量，增加值占国内生产总值比重达到15%，部分产业和关键技术跻身国际先进水平，节能环保、新一代信息技术、生物、高端装备制造产业成为国民经济支柱产业，

新能源、新材料、新能源汽车产业成为国民经济先导产业。

加快培育和发展战略性新兴产业，必须健全财税金融政策支持体系，加大扶持力度，引导和鼓励社会资金投入。其中，引导金融机构建立适应战略性新兴产业特点的信贷管理和贷款评审制度，是鼓励金融机构加大信贷支持力度的重要举措之一。战略性新兴产业企业具有较强的扩张能力，对资金的需求巨大，但由于受到技术生命周期、技术成熟度、市场容量、经营管理能力、技术创新能力、区域性产业聚集度、区域性产业配套能力、区域投资环境等因素的影响，其未来发展和经营业绩等具有较大的不确定性。因此，战略性新兴产业企业在成长过程中，必然伴随着不同程度的信用风险。如何构建科学的战略性新兴产业企业信用风险评价体系，是金融机构建立适应战略性新兴产业特点的信贷管理和贷款评审制度的基础性和关键性课题。解决好这一课题，首先，战略性新兴产业企业的信用状况将从新生伊始就有据可依，有案可查，并由授信方实施动态监控，这将十分有助于规范战略性新兴产业企业的生产经营行为，并促进其持续健康的发展；其次，有助于降低合约双方交易成本，减小信息非对称性，在一定程度上化解市场风险和金融风险，改善战略性新兴产业企业的投融资渠道；第三，对现代金融信用的分层与结构化以及信贷资产的证券化等也将具有积极的推动作用。

本书将对战略性新兴产业的发展现状及区域分布、银行与战略性新兴产业企业的信贷行为、战略性新兴产业企业信用风险影响因素、战略性新兴产业企业信用风险评价指标体系及指标赋权、战略性新兴产业企业信用风险的组合评价模型、战略性新兴产业企业信用风险的分行业动态评价、灰色系统理论、Sugeno 模糊积分和未确知测度在战略性新兴产业企业信用评价中的应用、MLP 神经网络在战略性新兴产业企业信用评价中的应用等问题展开一系列研究。本书的研究将拓展信用风险评价理论和方法的研究领域，对解决战略性新兴产业企业信用风险的识别与评价问题、促进科技和金融结合、加快培育和发展战略性新兴产业具有重要的理论意义和现实意义。

1.2 国内外研究现状

1.2.1 信用风险的概念与成因

人们对信用风险的认识，存在一系列不同的观点。观点之一认为，信用风险是指交易对手无力履约的风险，或者说，是债务人不能按期偿还其债务而给债权人带来的风险。观点之二认为，信用风险是指所有因客户违约而带来的风险。例如，资产业务中的借款不能按时还本付息、负债业务中的借款人提前挤兑、表外业务因为交易对手违约而转化为表内负债等。观点之三认为，信用风险是指由于借款人信用等级的转移和履约能力的变化，导致其债务的市场价值波动而引起损失的可能性。因此，信用风险的高低主要由交易对手的财务状况和风险状况决定^[6]。观点之四认为，信用风险可以分为广义的信用风险和狭义的信用风险，广义的信用风险是指所有因客户违约而带来的风险，狭义的信用风险则是指信贷风险^[7]。信用风险与信贷风险在概念上既有联系，又有区别^[6]。信贷风险是指在信贷过

程中,借款人受各种不确定性因素影响,无法按时偿还商业银行贷款,造成贷款本金和利息损失的可能性。从商业银行角度来看,信贷风险与信用风险在主体上是一致的,都是由债务人信用状况发生变化给商业银行经营带来的风险。信贷风险与信用风险的主要区别在于:两者所涵盖的金融资产范围不同,信用风险不仅包括信贷风险,而且还包括一系列存在于其他表内或表外业务中的风险。

信用风险的成因主要包括内部原因和外部原因^[7]。其中,内部原因是形成信用风险的最根本原因,借款人由于经营管理不善或决策失误等自身原因,导致无法按期偿还贷款的本金和利息,或导致其信用等级下降,从而产生信用风险。宏观经济环境、法律法规制度、政治制度等外部原因,均可能会造成借款人的信用状态发生变动,从而引发信用风险。一般认为,履行合约的能力和履行合约的意愿是信用风险的直接成因,因此,人们评价信用风险的基本思路就是对借款人的履约能力和履约意愿进行分析。

人们对信用风险评价方法的探索可追溯到20世纪30年代,20世纪60年代以后其逐渐演变为热点。最早被商业银行广泛采用的信用风险评价方法是专家分析法。20世纪80年代以来,随着金融自由化和全球化的发展,国际金融市场波动加剧,各国商业银行和投资者面临着空前的信用风险挑战,国际金融界加强了对信用风险评价方法的研究,各种新型的信用风险评价方法层出不穷,信用风险管理理论与技术日趋完善^[8]。综合来看,信用风险评价方法大体上经历了一个从简单到复杂、从定性分析到定量计算、从单一资产的信用风险评价到资产组合的信用风险评价的渐进发展过程^[9],先后出现了传统的信用风险评价方法、多元统计分析方法、人工智能方法和信用风险计量模型等。

1.2.2 传统的信用风险评价方法

(1)“5C”要素分析法。商业银行通过对借款人的道德品质(character)、还款能力(capacity)、资本实力(capital)、担保(collateral)和经营环境条件(condition)等因素进行定性分析,以判别借款人的偿还意愿和偿还能力,这一方法被业界称为“5C”要素分析法。另外,部分商业银行还归纳出“5W”或“5P”因素。上述方法的共同点在于,对每一要素逐一进行评分,以确定借款人的信用等级,根据借款人的信用等级进行信贷决策^[8]。

(2)LAPP原则。商业银行通过对借款人资产的流动性(liquidity)、活动性(activity)、盈利性(profitability)和潜力(potentialities)等指标进行定性分析,以评价借款人的信用状况^[9]。

(3)财务比率分析法。把企业各项财务指标作为一个整体,对企业财务状况和经营情况进行系统剖析和综合评价。杜邦财务分析体系和沃尔比重评分法是这类方法的主要代表。杜邦财务分析体系以净值报酬率为龙头,以资产净利润率为核心,重点对企业获利能力进行揭示。沃尔比重评分法分别给出七项财务指标的分数比重,通过与行业平均比率进行比较,确定各项财务指标的得分及总体指标的累计分数,由此得出对企业财务状况的综合评价结果,并确定企业的信用等级^[8]。

简评:传统的信用风险评价方法具有直观明了、操作简单等优点,但其缺点也显而易

见^[9]。例如,以定性分析为主,其主观性较强,缺乏客观的评价基础;借款人财务数据的真实性和准确性直接影响着信用风险评价结果的可靠性,而现实中的财务数据常常失真;综合分析能力较差,对财务状况缺乏整体的概括。

1.2.3 多元统计分析方法

1.多元判别分析法

多元判别分析法的基本思路是:从已经掌握的历史上每个类别(如违约类和正常类)的若干样本中,总结出其中的分类规律,建立判别函数,并对新样本进行分类。Altman^[10]率先将多元判别分析法运用于财务危机、公司破产及违约风险评估,提出了包含总资产、营运资本、未分配利润、总债务和销售额等5个变量的Z-score模型。该模型认为,Z值小于1.81为高风险区,Z值大于2.99为低风险区,而Z值介于1.81~2.99为灰色区域,即不能确定贷款的风险程度。Altman^[11]进一步对Z-score模型进行了改进,建立了所谓的“Zeta”判别分析模型。

国内部分学者对多元判别分析法进行了应用研究。例如,王春峰等^[12]运用多元判别分析法对我国商业银行的信用风险进行评估,通过与Logistic模型的比较,验证了多元判别分析法的有效性。施锡铨和邹新月^[13]分析了典型判别分析在企业信用风险评估中的应用。张玲和曾维火^[14]基于Z-score模型对我国上市公司进行了信用评级研究。张明等^[15]利用多元判别分析技术(multivariate discriminant analysis, MDA)对电子商务客户信用进行分类,取得了较好的预测效果。杨彬等^[16]将多元线性判别分析的结果与其他变量一起作为BP(back propagation)神经网络的输入变量,建立了混合两阶段个人信用评分模型,同时满足了预测精度和稳健性的双重要求。此外,胡胜^[17]选取我国上市公司财务数据进行实证分析,证实多元判别分析法的第一类错误率达30%左右。因此,商业银行在运用多元判别分析法对上市公司信用风险进行判断时要十分谨慎。霍雨佳^[18]将财务预警理论与公司债券违约风险的度量相结合,运用多元判别分析方法,建立财务预警指标体系和预警模型对公司债券违约进行预测,并对模型的有效性进行检验。

简评:多元判别分析法的优势在于其具有明显的解释性,而缺陷在于其前提条件过于严格,如要求样本数据服从正态分布和等协方差^[9],而现实中大量的信用数据并不满足这些条件。

2.Logistic 模型

Press 和 Wilson^[19]首次提出了 Logistic 模型。Ohlson^[20]率先将 Logistic 模型应用于商业银行信用风险评价领域,即根据一系列财务指标变量来预测公司破产或违约概率,并基于风险偏好程度设置风险警戒线,从而判断企业的信用类别。其基本形式如下:

$$P = \frac{1}{1 + e^{-y}}$$

其中, $y = C_0 + \sum_{i=1}^n C_i X_i$, 为信用风险评价中的企业财务指标; $P \in [0, 1]$ 为借款人的违约概率。

在国内, 于立勇和詹捷辉^[21]在结合国有商业银行实际数据的基础上, 通过 Logistic 回归模型构建了违约概率的测算模型。李萌^[22]以不良贷款率作为信用风险衡量标准, 构造了商业银行信用风险评估的 Logistic 模型, 并结合 t 检验和主成分分析法对模型进行实证分析。石晓军^[23]对 Logistic 违约率模型的最优样本配比与分界点进行了模拟分析。李关政和彭建刚^[24]针对传统 Logistic 模型在信用风险评估时忽略系统性风险的不足, 从经济周期和经济转型两个方面引入系统性风险因子, 结合财务因子, 构建了新的 Logistic 模型。田秋丽^[25]分析我国中小企业信用风险度量的适用模型, 并用实证方法对 Logistic 模型的度量效果进行检验。结果表明, Logistic 模型在中小企业信用风险度量中具有较好的适用性。李关政^[26]把传统的 Logistic 模型扩展为包含宏观冲击因子的 MF-Logistic 模型, 将其用于信用风险压力测试。实证结果显示: 基于 MF-Logistic 模型的信用风险压力测试能科学地度量宏观冲击因子和微观风险因素对信用风险的影响, 并且能直观地显示银行在不同压力情景下的资本充足水平。邓晶等^[27]以我国沪深两市 A 股上市公司为研究对象, 选取 2010~2012 年首次成为 ST 的 81 家公司和对应的 81 家非 ST 公司为研究样本, 运用 20 个财务指标进行因子分析, 并根据 Logistic 模型对所得因子进行回归分析, 建立预测模型。实证结果显示, 模型具有良好的预测效果, 在上市公司出现信用风险之前的预测准确率在 85% 以上。吴会咏和刘艳春^[28]选取代表中小企业的经营与发展能力、利润构成与盈利能力、资产与负债和现金流量 4 个方面的 13 个指标作为解释变量, 并应用偏最小二乘法提取 PLS 成分, 排除系统中的噪声干扰, 构建一个度量信用风险的二分类因变量的 Logistic 模型。实证结果表明, 该模型不仅具有良好的平稳性和准确性, 而且具有较强的解释过程变化的能力。此外, 胡心瀚等^[29]应用结合变量选择的 Logistic 模型估计了上市公司的违约概率, 采用核密度估计方法和 Archimedean copula 函数分别拟合了上市公司的违约概率变化和市值变化的边缘分布和联合分布, 并在大样本下通过条件 VaR 检验了结构化信用分析模型在中国市场的适用性。孙克和蒋岳祥^[30]构建 Logistic 模型对影响债券信用等级迁移的财务指标进行实证分析, 结果发现, 偿债能力和盈利能力都对企业债的信用等级迁移具有显著影响, 但是信用升级和信用降级分别对不同的财务指标敏感。孙焱林等^[31]采用蒙特卡罗模拟方法, 以 Logistic 回归模型为例, 对无约束数据回归模型和受约束数据回归模型进行了对比分析, 证实了这种有偏性的存在性, 并且可能会对信用风险评价结果产生影响。侯博等^[32]建立基于融资企业、核心企业、供应链运行状况及融资模式的信用风险评价指标体系, 并选取汽车中小企业为样本, 采用主成分分析法和 Logistic 模型构建信用风险评价模型。实证结果表明, 该模型具有较高的准确率和适用性。

简评: Logistic 模型的优点在于不要求样本数据服从正态分布和等协方差, 因而得到了广泛应用。Logistic 模型的不足之处在于对灰色区域的差别敏感性较强; 另外, 如果样本点完全分离, 此时对参数的最大似然估计不一定存在。

3. Probit 模型

Probit 模型是假设事件发生概率服从累积正态分布函数的二分类因变量模型,也称为 Normit 模型,其基本思路是:假设每一个体都面临二选一的选择,且其选择依赖于可分辨的个体特征,旨在寻找描述个体的一组特征与该个体所做某一特定选择的概率之间的关系^[33]。将 Probit 模型应用于预测企业破产概率的基本步骤是:首先确定样本的极大似然函数;然后通过求解极大似然函数的极大值来获取参数;最后利用公式求出企业破产概率,其分类判别规则与 Logistic 模型相同。

Zmijewski^[34]采用 Probit 模型预测了破产概率。高培业和张道奎^[35]把深市上市公司分为制造业和非制造业,采用一年的财务数据,运用线性判别模型、Probit 模型等五种模型进行企业失败预测实证研究。郑昱^[36]以在浙江省随机抽样的 506 个自然人为样本,运用 Probit 模型进行了个人信用风险实证研究。此外,张鹏和曹阳^[37]利用 KMV 模型计算出样本上市公司的违约距离,并将其作为 PROBIT 模型的自变量计算出上市公司的违约概率。实证结果表明,违约距离能较好地识别上市公司的信用风险,将违约距离作为自变量进行 PROBIT 建模分析,明显提高了模型的统计显著性和预测精度。程天笑和闻岳春^[38]使用 Probit、Logistic 与 Extreme Value 模型,讨论融资融券业务个人客户的违约概率计量。实证结果表明,三种模型均具有较强的区分能力,但 Extreme Value 模型准确性最强,拟合效果最好,最适合于计量融资融券业务的违约概率。

简评:Probit 模型和 Logistic 模型的共同之处在于均采用极大化似然函数。两者的不同之处在于^[7]:①假设条件不同。前者要求样本数据服从标准正态分布,后者则没有严格的假设条件。②求解参数的方法不同。前者采用极大似然函数求极值的方法,后者采用线性回归方法。③求企业破产概率的方法不同。前者采用积分方法,后者采用取对数的方法。

4. 非参数方法

k-近邻判别(k-nearest neighbor)属于非参数统计方法。k-近邻判别在一定的距离概念下,根据若干定量指标,从样本中选取与确定向量距离最短的 k 个样本组成一组,从而实现对本样本的分类。该方法适于对本样本初始分布和数据采集范围限制较少的情况,并减少了以函数形式表达内容的要求。另外,k-近邻判别通过划分任意多个决策区间而近似于样本实际分布^[39]。Tam 和 Kiang^[40]对 k-近邻判别在信用风险评价中的应用进行了研究,选取 19 个财务指标,基于马氏距离,对样本企业进行分类。其结果表明,k-近邻判别的分类效果不如多元判别分析法。Henley 和 Hand^[41]、姜明辉等^[42]也分别将 k-近邻判别运用到信用风险评价中。

聚类分析(cluster analysis)是一类非参数统计方法。在信用风险分析中,聚类分析根据样本点(代表借款人)在样本空间中的相互距离来进行分类。该类方法的突出优点在于不要求样本总体服从任何具体分布,既可对变量采用名义尺度,也可对变量采用次序尺度,因而非常适于信用风险分析中按照定量指标(如财务指标等)和定性指标(如管理者素质、发展前景等)对并不服从一定分布特性的数据信息进行分类的要求^[39]。

Lundy^[43]首次运用聚类分析对消费贷款申请者信用风险进行评价,即根据申请者年

龄、职业、婚姻状况和居住条件等指标,通过聚类分析判断消费贷款申请者所属的信用类别。在国内,肖北溟和李金林^[44]利用国有商业银行的贷款历史数据,通过因子分析、聚类分析等方法构建内部信用评级模型,并对模型的有效性进行了检验。刘淑莲等^[45]采用因子分析和聚类分析法对抽样选取的307家上市公司的信用风险进行综合评价,将具有相似信用风险水平的受评上市公司划分为同类,确定出评级标准,从而构造出上市公司信用评级模型。实证结果发现,该模型具有较好的识别和预测能力。奚胜田等^[46]运用因子分析法和聚类分析法,对在沪深交易所上市的33家农业机械行业和建筑机械行业的生产企业进行信用评级,并对照标准普尔(S&P)的信用评级体系确定各个企业的评级级别。此外,郭军华和李帮义^[47]提出了一种将模糊聚类(fuzzy clustering, FC)和变精度粗糙集理论(variable precision rough, VPRS)结合进行信用风险评价的模型。张洪祥和毛志忠^[48]提出基于多维时间序列模糊聚类与模糊规则提取技术相结合的模糊分类系统,将其应用于信用评价研究。张目和周宗放^[49]提出一种基于投影寻踪和最优分割的企业信用评级模型。该模型运用投影寻踪对样本企业进行信用综合评分,将信用综合得分由大到小排序,生成有序样品序列;利用最优分割法对有序样品进行聚类,得出明确的聚类结果;将最优分割点对应的信用综合得分作为划分信用等级的阈值,从而实现对样本企业的信用评级。王化中和强凤娇^[50]将灰色统计理论中的三角白化权函数聚类决策方法应用到中小企业信用评价中,并通过重新设计灰色聚类决策步骤,克服了三角白化权函数聚类系数不规范可能带来的聚类错误或影响类内排序问题。李战江等^[51]通过逼近理想点赋权模型确定评级指标的权重,建立评级总得分的测算模型;通过非参数核密度估计方法与切片取样方法获得反映小样本分布规律的评级大样本;通过ward聚类对评级大样本进行有序分类,建立了可分为9个信用级别的小样本评级模型。吴凤和吴义能^[52]选取2015年我国创业板上市公司的多元财务指标,运用因子分析法,在对每家企业进行K-means聚类分析的基础上,用Z值进行分类,将之作为该年度创业板上市公司信用评级的基准。

简评:非参数方法在实际应用中面临着一定的问题。例如,在样本容量相同的情形下,对某一具体问题,如果确实存在特定的参数模型,且能够找出,此时参数模型的效率高于非参数方法^[39]。k-近邻判别不要求样本数据服从正态分布,但是,当样本数据的维数较高时,存在所谓的“维数祸根”,表现为即便样本量很大,其撒在高维空间中仍然显得非常稀疏,绝大多数样本点附近根本没有样本点,这将导致k-近邻判别难以使用。此外,当数据维数较高时,聚类分析也会遇到很大的困难。

5. 分类树方法

分类树是一种基于统计理论的、由计算机实现的非参数识别技术^[53],是20世纪80年代末利用机器学习发展起来的符号方法,它可以实现对原始样本的最佳分类判别。分类树的基本思路为,样本空间 X 包含A和B两类样本,将其作为根结点,按照一定的规则进行分割,产生两个结点,即子集 X_1 和 X_2 ,它们满足: $X = X_1 \cup X_2$,且 $X_1 \cap X_2 = \emptyset$ 。这一过程递归地对子集 X_1 和 X_2 重复进行,直至按照某种标准,结点无法再分,从而成为最终的叶结点,叶结点所表示的数据子空间的特征决定其属于A类还是B类。

Frydman等^[54]根据现金流/总负债、保留盈余/总资产、总负债/总资产、现金流/销售收