

# 基于 TOPSIS 和多目标优化的风险量化与信贷策略模型

## 摘要

银行对中小微企业贷款的审核，一般以信贷政策、企业的实力、信誉和经营情况为主要依据，因此需建立评价体系进行风险量化、制定信贷策略。本文借助 TOPSIS 方法，提取代表企业信誉和实力的主要特征，建立风险量化方案计算企业风险评估值，并据此建立多目标优化模型，为银行信贷的合理分配提供了可行策略。之后将模型的应用场景进行推广，讨论了在市场突发变动的条件下，银行信贷策略的调整与优化。

针对问题一**有信贷记录的 123 家企业的风险量化和信贷策略确定**，我们首先选取了资金实力、营业能力、发展潜力、供销关系稳定性、交易过程稳定性五个观测指标，建立风险量化评估体系，并利用 TOPSIS 方法计算了各企业的风险量化评估值。在利用专业信誉评价结果进行准确性校验后，我们依据风险量化评估值，以各企业的信贷金额为决策变量，以银行收益、风险损失为两个目标函数，建立了受到贷款金额总和、利率和贷款金额范围以及银行风险管控政策约束的双目标优化信贷策略模型，设计算法求解后将各企业放贷与否、贷款利率和贷款金额的结果保存在“附件一”文件中。

针对问题二**无信贷记录的 302 家企业的风险量化和信贷策略确定**，我们延续问题一的求解思路，利用在问题一中建立的风险量化评估体系进行各企业风险量化评估值的计算，以此为标准对所有企业进行的信誉评级，将问题一所建立的信贷策略模型进行推广并应用在本问中，设计算法求解后将各企业放贷与否、贷款利率和贷款金额的结果保存在“附件二”文件中。

针对问题三**在突发因素影响下对 302 家企业风险量化方案和信贷策略的调整**，我们考虑到突发因素对银行贷款政策和企业发展情况的影响，引入风险管控因子，通过上调贷款利率和进行贷款限额修正银行在非常时期的贷款政策；引入发展情况映射关系，将突发因素对企业的影响分为市场整体影响和企业个体影响，建立起突发事件前后企业行为的映射关系，修正企业在非常时期的还贷能力。利用上述指标对问题一的模型进行修正，得到在突发因素影响下的风险量化方案和信贷策略模型，设计算法求解后将各企业放贷与否、贷款利率和贷款金额的结果保存在“附件三”文件中，并对问题二、问题三的结果进行了比较，分析在不同时期的信贷策略差别，说明模型的合理性。

最后我们分析了模型的优、缺点，并提出了模型的改进方向。

**关键词：**风险量化，信贷策略，TOPSIS 方法，多目标优化模型

## 一、问题重述

由于中小微企业规模相对较小，也缺少抵押资产，因此在申请贷款时银行通以**信贷政策、企业的交易票据信息**和**上下游企业的影响力**为主要依据，向**实力强、供求关系稳定**的企业提供贷款，并可以对**信誉高、风险小**的企业给予利率优惠。具体操作过程一般为首先根据中小微企业的实力、信誉对其信贷风险做出评估，然后依据信贷风险等因素来确定是否放贷及贷款额度、利率和期限等信贷策略。

某银行对确定要放贷企业的贷款额度为 10~100 万元；年利率为 4%~15%；贷款期限为 1 年。现需根据现实和给出的企业信息，通过建立数学模型研究对中小微企业的信贷策略，并解决以下问题：

- ◆ **问题一：**对有信贷记录的 123 家企业的信贷风险进行量化分析，给出该银行在年度信贷总额固定时对这些企业的信贷策略；
- ◆ **问题二：**对无信贷记录的 302 家企业的信贷风险进行量化分析，给出该银行在年度信贷总额为 1 亿元时对这些企业的信贷策略；
- ◆ **问题三：**综合考虑各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，给出该银行在年度信贷总额为 1 亿元时的信贷调整策略。

## 二、问题分析

从目标的角度看，问题一到问题三分别要求对有信贷记录的企业进行风险量化并建立策略，将策略在有信贷记录企业的数据集中验证后推广使用在无信贷记录的企业，最后要考虑突发因素对风险评估和信贷策略的影响。总的来看，问题的解决鲜明地分成两个步骤——建立风险量化评估方案、建立信贷分配策略。

在建立风险量化评估方案时，需要首先找到影响信贷风险的主要因素并制定指标进行刻画，将各项指标进行统筹形成一套风险量化评估体系。在建立信贷分配策略时，应站在银行的角度，统筹考虑风险和收益，以带有风险量化标签的企业为对象建立优化模型，进行在可控风险范畴下的利益最大化。本问的整体建模和求解思路如图 1 所示。

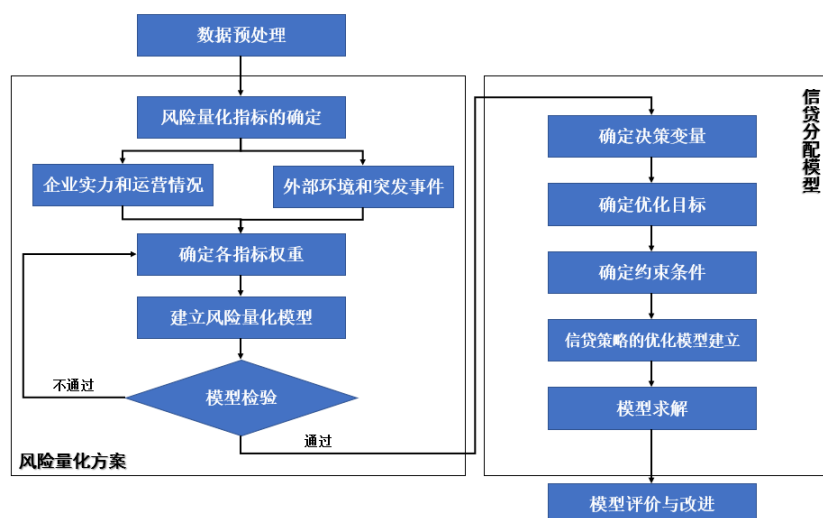


图 1 文章整体建模和求解思路

## 2.1 对问题一的分析

问题一要求对有信贷记录的 123 家企业的信贷风险量化分析，根据分析结果建立信贷总额一定时各企业的贷款金额、贷款利率的制定策略。量化分析的基础是观测指标的确定，首先需根据题目给出的历年进项发票、销项发票数据，构造出能够衡量企业实力、经营状态、信誉情况的指标，建立起风险量化体系。之后可以借助熵值法确定各指标在体系中的权重，通过 TOPSIS 方法计算出衡量企业风险的量化指标。通过计算所得的风险指标和题目给出的历史信贷信息、银行的专业信誉评级进行比较，可以验证评价体系的可行性。

风险量化完成后，建立信贷金额的分配策略。由于银行在放贷过程中主要关注风险控制条件下的利益最大化，因此可以各企业信贷金额为决策变量，以银行一年内总的期望收益为优化目标，建立受到贷款总额、企业贷款限额等条件约束的多目标优化模型。编写算法求解即可完成在本问题下银行信贷策略的策略。

## 2.2 对问题二的分析

问题二要求对无信贷记录的 302 家企业的信贷风险进行量化分析，给出该银行在年度信贷总额为 1 亿元时对这些企业的信贷策略。与问题一相比，区别在于未给出专业的信誉评级和历史违约情况。由于在问题一中已经建立起不依赖历史评价的模型，因此直接按照问题一的数据处理和模型求解方案对问题二数据进行计算，即可完成本问要求。

## 2.3 对问题三的分析

问题三要求综合考虑各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，对模型进行调整并对问题二进行重新计算和比较。我们经过研究发现，突发的大型公共事件会对模型的两方面因素产生影响。

一方面是银行的贷款政策，非常阶段银行将会收紧贷款发放和优惠力度，加大风险管控的力度以减小基准利率的浮动水平，保证国民经济的稳定性。体现在模型中表现为根据企业的分类和性质，降低单一企业的贷款金额上限和上调总体的贷款利率。另一方面是企业的发展情况，突发时间使得大部分企业的违约率升高，经营稳定状况下降。当然，根据利与弊的矛盾共生关系，非常时期也定会有企业异军突起，出现正向增长的情形，需要根据企业类型和性质具体分析。

根据上述分析，我们可以引入新的风险管控因子、发展影响因子，对模型中的相关参数进行修正。因子的具体取值根据企业类型、承压能力等因素来确定，利用问题二中的数据集进行重新求解，比较突发事件产生时贷款金额的变化，结合实际情况可以进一步解释模型的合理性。

# 三、模型假设与变量说明

## 3.1 模型假设

- 1.题目提供的数据真实、完整、有效，选取的企业具有一般性；
- 2.银行对于企业的信用评级准确可靠，能够作为参考的标准；
- 3.银行发放贷款的利率仅由申请贷款公司的风险信誉决定，不考虑其他因素影响；
- 4.假设银行相比利润更加重视风险，需制定严格风险管控机制。

### 3.2 变量说明

本文主要用到的变量名称和解释如表 1 所示，未列出或重复的符号以出现处的定义和解释为准。

表 1 主要变量说明表

变量名称	变量解释
$ASR$	年销售额
$TPTP$	销项发票价税合计值
$TPTS$	进项发票价税合计值
$ASP$	年销售利润
$NEI$	有效发票数量
$TI$	发票总量
$SS$	供方企业集合
$SP$	销方企业集合
$\eta$	年销售利润率
$g$	年销售额增长率
$\beta$	供销稳定指数
$\gamma$	年发票有效率
$C_i$	风险量化评价指标
$e_i$	第 $i$ 家企业的贷款金额
$r_i$	第 $i$ 家企业的贷款利率
$l_i$	第 $i$ 家企业的流失率
$p_i$	第 $i$ 家企业的违约率
$r_f$	无风险利率
$\zeta_1$	利率上调变动因子
$\zeta_2$	贷款限额变动因子
$\lambda$	企业变动因子
$\varepsilon$	市场变动因子

## 四、模型的建立与求解

### 4.1 问题一：基于 TOPSIS 的风险量化估计和信贷策略模型

#### 4.1.1 风险量化观测指标的选取

风险量化指标的选取是建立风险量化模型并进行风险预测的基础，本节的重点在于如何从企业的交易发票信息、信贷记录中提取可以衡量企业信贷风险水平的观测指标。根据题目要求和查阅相关文献<sup>[1]</sup>，对于面向小微企业的贷款银行通常是依据信贷政策、企业交易票据信息和上下游企业的影响力，向实力强、供求关系稳定的企业提供贷款，企业的实力、信誉、盈利能力、发展潜力、供销关系、经营稳定程度等是银行着重关注的指标，因此我们从以下五个方面构造本问进行风险量化的观测指标。

### (1) 资金实力

企业实力是指企业在生产、技术、销售、管理和资金等方面力量的总和。一般来讲，企业资金实力是结合企业资金体量、运营实力、行业年景等因素进行综合考虑的评估结果，是风险大小的重要影响因素。

考虑到题目数据条件有限，仅给出进项发票、销项发票价税合计值，同时为排除 2020 年初突发疫情的偶然因素，我们采用最近一个完整自然年(2019 年)内所有有效销项发票价税合计值(Total of Price and Tax of Purchase, TPTP),粗算 2019 年年销售额(Annual Sales Revenue, ASR)，将其作为企业资金实力的评价指标，即

$$ASR = \sum TPTS_i \quad (1)$$

### (2) 盈利能力

盈利能力是指企业获取利润的能力，也称为企业的资金或资本增值能力，通常表现为一定时期内企业收益数额的多少及其水平的高低<sup>[2]</sup>。企业在销售过程中的利润情况，一定情况下可以反映企业的盈利能力，体现了企业偿还贷款的能力。

我们采用最近一个完整自然年(2019 年)内所有有效销项发票价税合计值与进项发票价税合计值(Total of Price and Tax of Sales, TPTS)之差，粗算 2019 年年销售利润(Annual Sales Profit, ASP)，取其与当年年销售额的比值，即当年年销售利润率  $\eta$ ，作为企业盈利能力的评价指标，即

$$\eta = \frac{ASP}{ASR} = \frac{\sum TPTS_i - \sum TPTP_j}{\sum TPTS_i} \quad (2)$$

### (3) 发展潜力

企业发展潜力是指企业未来的发展趋势，是否能保持高活力、高效益，是否能紧跟市场和行业的发展趋势，预测了未来一段时间企业是否会出现动荡而产生风险。影响企业发展潜力的主要因素包括行业发展空间，企业核心竞争力，公司财务状况等。

本问中，通过盈利水平随年限的统计特征，刻画企业盈利能力的变动，在一定程度上反映企业未来的发展潜力。采用最近一个自然年(2019 年)内年销售额和前一年(2018 年)年销售额的差与当年年销售额之比，粗算年销售额增长率  $g$ ，作为企业发展潜力的评价指标，即

$$g = \frac{ASR' - ASR}{ASR} = \frac{\sum TPTS'_i - \sum TPTS_i}{\sum TPTS_i} \quad (3)$$

#### (4) 供销关系稳定性

供销关系的稳定性表示企业是否长期与固定的供货方、销货方进行交易，一定程度上能体现上下游企业的实力。稳定的供销关系是企业运营的重要条件，因此供销关系稳定性评估是风险评估的重要一环。

本问中，给出了过去数年内与企业产生交易的供、求方企业的数目和交易时间。我们统计最近一个自然年(2019 年)以及前一年(2018 年)内，与企业产生交易的**供方企业集合**(Set of Supplier, SS)、**销方企业集合**(Set of Purchaser, SP)，定义**供销稳定指数**  $\beta$  评价一个企业的供销关系稳定性，即

$$\beta = \alpha_1 \frac{|SS' \cap SS|}{|SS' \cup SS|} + \alpha_2 \frac{|SP' \cap SP|}{|SP' \cup SP|} \quad (4)$$

其中， $\alpha_1$ 、 $\alpha_2$  为经验性系数，一般认为供方比销方对稳定性影响更大，这里  $\alpha_1$  取 0.7， $\alpha_2$  取 0.3<sup>[3]</sup>。

#### (5) 交易过程稳定性

交易是企业运营的重要过程，无论购进过程还是销售过程，一旦出现中断、违约等异常情况都可能对企业运营策略产生影响。因此，评价交易过程的稳定性是风险评估的重要组成部分。

本问中，只给出了票据的相关信息，因此采用**有效发票数量**(Number of Effective Invoice, NEI)占**发票总量**(Total of Invoice, TI)的比例即**年发票有效率**  $\gamma$ ，作为交易过程稳定性的评价指标，即

$$\gamma = \frac{NEI}{TI} \quad (5)$$

其中，定义无效发票为交易未正常进行所产生的异常发票，包括作废发票和负数发票。有效发票为交易正常进行所产生的发票，有效发票和无效发票之和为发票总量。

### 4.1.2 基于 TOPSIS 方法的风险量化方案

#### 4.1.2.1 风险量化体系的建立

经过对企业风险产生的客观原因的分析，结合题目所给出的数据，我们建立了考察企业资金实力、盈利能力、发展潜力、供销关系稳定性和交易过程稳定性五方面，包含年销售额 ASR、年销售利润率  $\eta$ 、年利润增长率  $g$ 、供销稳定指数  $\beta$  和年发票有效率  $\gamma$  五项观测指标的风险量化体系，如图 2 所示。

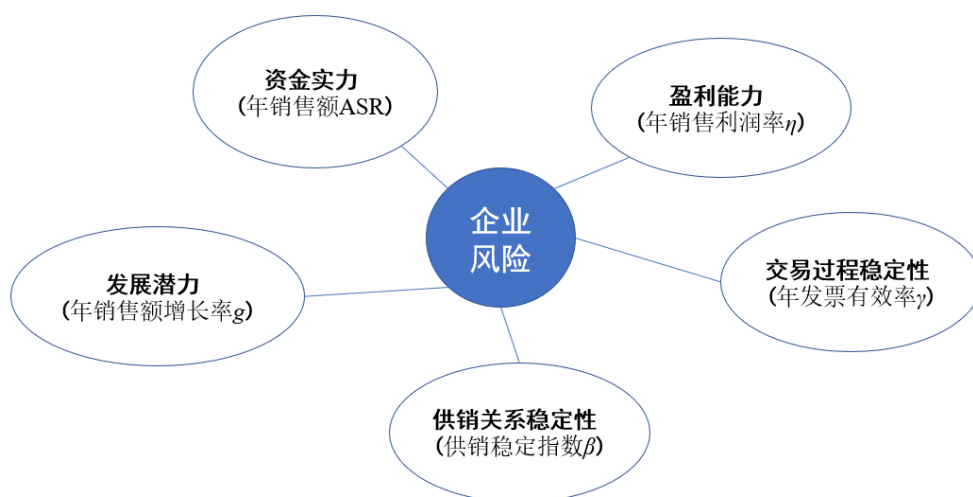


图 2 企业风险量化体系

#### 4.1.2.2 基于熵值法的指标权重确定

基于信息论的熵值法是根据各指标所含信息有序程度的差异性来确定指标权重的客观赋权方法，仅依赖于数据本身的离散程度。熵用于度量不确定性，指标的离散程度越大，则熵值越大，表明指标值提供的信息量越多，则该指标的权重也应越大。具体计算流程如下。

##### ◆ 原始数据矩阵的归一化

对于  $n$  个待评价对象，每个对象都有  $m$  个指标，则初始数据矩阵构造为

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \quad (6)$$

将其中每一项进行归一化，这里采用比值归一化的方法，即

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (7)$$

##### ◆ 各指标熵值的计算

根据熵的定义，计算各指标的熵值为

$$e_j = -k \sum_{i=1}^n p_{ij} \ln p_{ij}, j = 1, 2, \dots, m \quad (8)$$

其中  $k$  与样本数量相关，这里取  $k = 1/\ln n$ 。此外当  $p_{ij}=0$ ，则令  $p_{ij} \ln p_{ij} = 0$ 。

##### ◆ 各指标权系数的计算

根据熵权系数的定义，计算各指标权重为

$$h_j = \frac{1 - e_j}{\sum_{k=1}^m 1 - e_k} \quad (9)$$

熵权系数  $h_j$  越大，则该指标代表的信息量越大，表示其对综合评价的作用越大。

需要指出的是，从理论上讲权重的赋予应是由专业的金融分析人士基于经济学原理进行考量，不同的学者对不同因素的重视不同就会导致权重的不同，从根本上具有不可避免的主观性。这里采取的熵值法推算权重无法保证绝对的准确和有效，我们将利用专家给出的信誉评级对熵值法算出的结果进行校验，来说明我们量化方法在一定程度上的可行性。

#### 4.1.2.3 基于 TOPSIS 风险评估方案的建立

**TOPSIS 法** (Technique for Order Preference by Similarity to an Ideal Solution) 是 C.L.Hwang 和 K.Yoon 于 1981 年首次提出，TOPSIS 法根据有限个评价对象与理想化目标的接近程度进行排序的方法，是在现有的对象中进行相对优劣的评价<sup>[4]</sup>。其基本原理，是通过检测评价对象与最优解、最劣解的距离来进行排序，若评价对象最靠近最优解同时又最远离最劣解，则为最好；否则不为最优。其中最优解的各指标值都达到各评价指标的最优值。最劣解的各指标值都达到各评价指标的最差值。其基本流程如图 3 所示。

TOPSIS Algorithm	
<b>Input</b>	原始数据集 $X = \{x_1, x_2, \dots, x_n\}$ 各指标权重 $\omega = \{\omega_1, \omega_2, \dots, \omega_m\}$
<b>Output</b>	各数据样本 TOPSIS 评价结果
<b>Process</b>	<ol style="list-style-type: none"> <li>1 对原始数据集中的指标属性同向化 <math>X'</math></li> <li>2 构造向量归一化后的标准化矩阵 <math>Z = \{z_1, z_2, \dots, z_n\}</math></li> <li>3 <b>for</b> <math>Z</math> 的每一列 <math>Z_i</math> <b>do</b></li> <li>4     最劣方案 <math>Z</math> 的第 <math>i</math> 维度 <math>\leftarrow Z_i</math> 元素最小值</li> <li>5     最优方案 <math>Z^*</math> 的第 <math>i</math> 维度 <math>\leftarrow Z_i</math> 元素最大值</li> <li>6 <b>end for</b></li> <li>7 <b>for</b> <math>z_i \in Z</math> <b>do</b></li> <li>8     <math>z_i</math> 与最优方案的接近程度 <math>D^+_i \leftarrow</math> 式 15-1</li> <li>9     <math>z_i</math> 与最劣方案的接近程度 <math>D^-_i \leftarrow</math> 式 15-2</li> <li>10    <math>z_i</math> 与最优方案的贴近程度 <math>C_i \leftarrow</math> 式 16</li> <li>11 <b>end for</b></li> <li>12 根据 <math>C_i</math> 大小进行排序</li> </ol>

图 3 TOPSIS 算法流程

下面对各步骤具体操作进行详细讨论<sup>[5]</sup>。

#### ◆ 指标属性正向化



TOPSIS 法使用距离尺度来度量样本差距，使用距离尺度就需要对指标属性进行同向化处理。若一个维度的数据越大越好，另一个维度的数据越小越好，会造成尺度混乱。这里统一采用成本型指标向效益型指标转化。对于任一成本型指标  $x$ ，其转化映射关系为

$$x' = \begin{cases} 2 \frac{x-m}{M-m}, & m \leq x \leq \frac{1}{2} (M+m) \\ 2 \frac{M-x}{M-m}, & \frac{1}{2} (M+m) \leq x \leq M \end{cases} \quad (10)$$

其中  $M$  为指标  $x$  可能取值的最大值， $m$  为指标  $x$  可能取值的最小值。

#### ◆ 归一化初始矩阵构造

设有  $n$  个待评价对象，每个对象都有  $m$  个指标，则初始数据矩阵构造为

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \quad (11)$$

对属性进行归一、规范化处理，每一列除以当前列向量范数(使用余弦距离度量)，即

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad (12)$$

据此得到标准化矩阵  $Z$ ，表示为

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nm} \end{bmatrix} \quad (13)$$

#### ◆ 最优最劣方案的确定

根据构造原理，最优方案  $Z^+$  由  $Z$  中每列元素的最大值构成，最劣方案  $Z^-$  由  $Z$  中每列元素的最小值构成，即

$$\begin{aligned} Z^+ &= \max z_{11}, z_{21}, \dots, z_{n1}, \max z_{12}, z_{22}, \dots, z_{n2}, \dots, \max z_{1m}, z_{2m}, \dots, z_{nm} \\ &= Z_1^+, Z_2^+, \dots, Z_m^+ \\ Z^- &= \min z_{11}, z_{21}, \dots, z_{n1}, \min z_{12}, z_{22}, \dots, z_{n2}, \dots, \min z_{1m}, z_{2m}, \dots, z_{nm} \\ &= Z_1^-, Z_2^-, \dots, Z_m^- \end{aligned} \quad (14)$$

#### ◆ 各评价对象贴近程度的计算

首先用距离刻画各评价对象与最优方案、最劣方案的接近程度，即

$$D_i^+ = \sqrt{\sum_{j=1}^m \omega_j (Z_j^+ - z_{ij})^2} \quad D_i^- = \sqrt{\sum_{j=1}^m \omega_j (Z_j^- - z_{ij})^2} \quad (15)$$

其中  $\omega_j$  为第  $j$  各属性的权重，由 5.1.2.2 中的熵值法确定。据此可以计算各评价对象与最优方案的贴近程度为

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (16)$$

其中  $0 \leq C_i \leq 1$ ， $C_i$  越接近 1 表明评价对象越优。

至此，通过 TOPSIS 方法计算得到的  $C_i$ ，可以作为不同企业的风险评估值并据此排序，完成风险量化的任务。

### 4.1.3 各企业风险量化分析和检验

#### 4.1.3.1 各企业风险量化分析结果

将预处理得到的企业数据代入 5.1.2 模型中，可以得出各企业的风险评估值  $C_i$ 。根据问题一给出的银行对各企业的专业信誉评级，A 级企业 27 家，B 级企业 38 家，C 级别企业 34 家，D 级别企业 24 家，分布比例如图 4 所示。

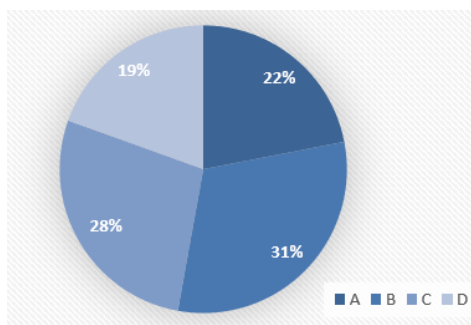


图 4 企业评级分布图

按照相同比例对计算结果按照风险评估值进行分级和排序，得到部分企业的风险量化评估结果，如表 2 所示。完整结果详见附件。

表 2 问题一企业风险量化评估结果

企业编号	风险评估值	风险评级
E1	0.439208	A
E2	0.734273	A
E3	0.731189	A
E4	0.980558	A
E5	0.404883	A
E6	0.520681	A
E7	0.707499	A
...	...	...

#### 4.1.3.2 基于相关性分析的结果校验

现在我们得到了银行给出的专业风险评级，以及经过数据分析得到的计算风险评级，需要利用专业风险评级对计算结果的准确性、可靠性进行校验。这里采用 Spearman 秩

相关性检验法从统计学角度进行校验。

### (1)Spearman 秩相关性检验的基本原理

Spearman 相关系数被定义成等级变量之间的皮尔逊相关系数。对于样本容量为  $n$  的样本， $n$  个原始数据被转换成等级数据，相关系数  $\rho$  为

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (17)$$

实际应用中，变量间的连结是无关紧要的，于是可以通过简单的步骤计算  $\rho$ 。被观测的两个变量的等级的差值，则  $\rho$  为

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (18)$$

其中  $d_i = (x_i - y_i)$ ， $x_i, y_i$  分别表示两个变量按降序排列的秩， $n$  为样本容量。

进行显著性检验时，令原假设为“ $H_0$ : X 和 Y 相互独立， $H_1$ : X 和 Y 正相关”，则假设检验的拒绝域分别为： $W = \{\rho_s \geq c_\alpha\}$ ， $W = \{\rho_s \leq -c_\alpha\}$ ， $c_\alpha$ ， $d_\alpha$  为临界值，在零假设成立时， $t = r_s \sqrt{(n-2)/(1-r^2)}$ ，服从自由度为  $n-2$  的 t 分布。

在显著水平为  $\alpha$  时，统计量的值落在否定域  $\{t | |t| > t_\alpha / (2n-4)\}$  中，拒绝零假设，Spearman 等级相关系数显著；否则，接受零假设，则 Spearman 等级相关系数不显著<sup>[6]</sup>。

### (2)结果和分析

利用 SPSS 进行 Spearman 秩相关性分析，得到的结果如表 3 所示。画出热力图如图 5 所示

表 3 相关性分析结果图

		专业风险评级				
计算 风险 评级		A	B	C	D	Total
	A	15(56)	8(30)	4(15)	0	27
	B	7(18)	15(39)	14(37)	2(5)	38
	C	4(12)	13(38)	13(38)	4(12)	34
	D	1(4)	2(8)	3(13)	18(75)	24
	Total	27	38	34	24	123



图 5 Spearman 秩相关性分析结果热力图

计算风险评级与专业风险评级的 Spearman 相关系数为  $0.5823(P = 1.62 \times 10^{-12})$ ，二者存在相关性，结论具有统计学意义。据此，我们验证了我们风险量化评估方案在结果上具有一定的可靠性，可以用于后续的信贷策略建立以及其他的相关问题。

#### 4.1.4 基于多目标优化的信贷策略模型

##### 4.1.4.1 违约率的确定

在风险量化过程中，我们得到了每个企业的风险评估值  $C_i$ ，并对其可靠性进行了检验，因此风险评估值可以作为风险的量化指标，在信贷策略制定时发挥作用。

我们认为风险评估值的大小和企业违约率存在着相关关系，评估值越大，风险越低，违约率越小；且在评估值较大时，违约率变化速率极低，而评估值在接近 0 的很小区间内，违约率迅速向 0 趋近。基于此，我们利用函数拟合违约率  $p$  与风险评估值  $C_i$  之间的关系，认为二者的关系近似服从函数

$$p = \begin{cases} 1, & 0 \leq x \leq 10^{-4} \\ -1000x + 1.1, & 10^{-4} \leq x \leq 10^{-3} \\ -0.1x + 0.01, & 10^{-3} \leq x \leq 10^{-1} \\ 0, & 10^{-1} \leq x \leq 1 \end{cases} \quad (19)$$

基于此可以建立起风险评估值与违约率之间的关系。

##### 4.1.4.2 放贷与否、贷款金额与贷款利率

###### ◆ 放贷与否

根据题目要求，信誉等级为 D 的不予发放贷款。信誉等级为 A 的，违约出现的概率极低，因此在非特殊情况下应予以发放贷款。而在信誉等级为 B 和 C 的企业中，仅有少

数企业出现了违约情况。通过计算，有违约记录的企业占有所有企业的 21.95%，因此应不予这一部分企业发放贷款。

根据风险评估值的计算结果，排名后 21.95% 的企业的评价值阈值为 0.00015，因此对于小于该值的企业，不应予以发放贷款。

#### ◆ 贷款金额

对于确定放贷的  $n$  个企业，设给予他们的贷款金额序列为  $E = (e_1, e_2, \dots, e_n)$ ，其中  $e_n$  表示银行给予第  $n$  个企业的贷款金额。由于资金在不同的企业将收取不同的利息，承担的风险也有所不同，则不同的分配方式直接决定了银行的可能收益和风险，因此我们将贷款金额作为策略优化模型的决策变量。

#### ◆ 贷款利率

根据题目要求，风险越低的企业，应该在利率上受到更多的优惠。我们采取聚类的方法，根据各企业风险评估值的统计特征将所有确定放贷的企业聚类为 29 类，以对于 29 种不同的利率。得到的各企业贷款利率。表 4 展示了前 6 家企业的利率结果，完整结果详见附件。

表 4 利率计算结果示意图

企业编号	贷款利率
E1	0.075
E2	0.0585
E3	0.0585
E4	0.04
E5	0.1225
E6	0.0505
...	...

#### 4.1.4.3 目标函数的确定

对于信贷，银行的主要目标有二——利息收益最高、风险损失最小。基于此，我们选取银行总的利润和所有可能风险作为优化目标。

##### (1) 利益最大

对于第  $i$  家确定放贷的企业，在 5.1.4.2 中我们给出了给予它的贷款金额为  $e_i$ ，收取的利率为  $r_i$ ，根据附件三可得在该利率下该企业的流失率为  $l_i$ ，则银行在这笔贷款中所收获的利润为

$$e_i \cdot r_i \cdot 1 - l_i \quad (20)$$

同时，考虑到银行如果不发放这一笔贷款，这些钱也将以无风险利率  $r_f$  进行增值，因此应减去无风险收益对利润进行修正，修正后的利润  $w_i$  为

$$w_i = e_i \cdot r_i \cdot 1 - l_i - r_f \quad (21)$$

因此，对于所有确定放贷的  $n$  家企业，总的销售利润为各企业利润之和，我们期望这一值最大，即第一个优化目标为

$$\max \sum_{i=1}^n e_i \cdot r_i \cdot 1 - l_i - r_f \quad (22)$$

根据当前银行政策式中  $r_f$  取 3.8%。

## (2) 风险最小

对于第  $i$  家确定放贷的企业，贷款金额为  $e_i$ ，收取的利率为  $r_i$ ，根据 5.1.4.1 可以计算该企业的潜在违约概率为  $p_i$ 。一旦该企业违约，银行将损失本金、利润和这些资金以无风险利率  $r_f$  进行增值的价值，故损失金额的期望  $d_i$  为

$$d_i = e_i \cdot 1 + r_i \cdot p_i \quad (23)$$

对于所有确定放贷的  $n$  家企业，总的可能损失为各企业损失期望之和，我们期望这一值最小，即第二个优化目标为

$$\min \sum_{i=1}^n e_i \cdot 1 + r_i \cdot p_i \quad (24)$$

### 4.1.4.4 约束条件的确定

根据题目要求和银行政策，模型共受到贷款总额、贷款限额和风险管控三方面约束。

#### (1) 贷款总额

根据题目要求，所有企业贷款金额之和不应超过固定的上限值  $S$ ，即

$$\sum_{i=1}^n e_i \leq S \quad (25)$$

#### (2) 贷款限额

根据题目要求，单一企业贷款金额不应小于 10 万元而不应大于 100 万元，即

$$10 \leq e_i \leq 100, i = 1, 2, \dots, n \quad (26)$$

#### (3) 风险管控

我们经过查阅资料，出于对风险的限制，对一个企业的贷款金额上限与其违约率负相关，违约率越高，贷款金额上限越低，这里我们取

$$e_i \leq 100 \cdot 1 - p_i, i = 1, 2, \dots, n \quad (27)$$

### 4.1.4.5 信贷策略模型的最终确定

综合 4.1.4.3 和 4.1.4.4 的计算和推导，我们建立了以各企业贷款金额为决策变量，以利益和风险为目标，受到贷款总额、贷款限额和风险管控约束的双目标优化模型如下

$$\begin{aligned}
& \max \sum_{i=1}^n e_i \cdot r_i \cdot 1 - l_i - r_f \\
& \min \sum_{i=1}^n e_i \cdot 1 + r_i \cdot p_i \\
& s.t. \begin{cases} \sum_{i=1}^n e_i \leq S \\ 10 \leq e_i \leq 100, i = 1, 2, \dots, n \\ e_i \leq 100 \cdot 1 - p_i, i = 1, 2, \dots, n \end{cases}
\end{aligned} \tag{28}$$

## 4.1.5 模型的求解与结果

### 4.1.5.1 求解算法

由于模型中两种优化目标决策变量相同，且具有明显线性性，因此可将二者相减构造单目标优化的求解。观察求解目标的结构发现，最终的优化模型转化为定系数的  $n$  元线性函数，形如

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \tag{29}$$

因此求解此模型只需将系数降序排列，按照从大到小的顺序进行贷款金额分配即可。具体求解算法如下。

**Step1** 为所有确定放贷企业预分配下限金额 10 万元；

**Step2** 计算各企业贷款金额系数向量  $(\alpha_1, \alpha_1, \dots, \alpha_n)$ ；

**Step3** 将系数向量降序排列，得到新的序列  $(\beta_1, \beta_1, \dots, \beta_n)$ ；

**Step4** 按照从大到小的顺序依序进行贷款分配，将每家企业贷款金额补充至受约束贷款金额最大值，直至贷款总额全部被分配或所有企业被遍历；

**Step5** 按照企业编号顺序输出贷款金额向量  $(e_1, e_1, \dots, e_n)$ 。

### 4.1.5.2 求解结果

利用上述算法对模型求解，得到第一问向各企业发放的贷款金额。表 5 展示了其中若干企业的结果，所有结果详见附件。

表 5 问题一求解结果示意图

企业编号	计算信用评级	是否放贷	贷款金额	贷款利率
E1	A	是	100	0.0705
E2	A	是	100	0.0585
E21	B	是	99.737	0.1185
E64	C	是	99.737	0.1025
E65	C	是	94.118	0.1305
E111	D	否	—	—
...	...	...	...	...

## 4.2 问题二：模型在无信贷记录企业中的推广应用

### 4.2.1 问题分析

问题二要求给出对无信贷记录企业的信贷策略，包括是否放贷、贷款金额和贷款利率。在问题一中，我们已经建立了对于一批企业的风险量化分级体系，给出了每一家企业的风险评估值  $C_i$ ，并建立了信贷分配策略模型

$$\begin{aligned}
 & \max \sum_{i=1}^n e_i \cdot r_i \cdot 1 - l_i - r_f \\
 & \min \sum_{i=1}^n e_i \cdot 1 + r_i \cdot p_i \\
 & s.t. \begin{cases} \sum_{i=1}^n e_i \leq S \\ 10 \leq e_i \leq 100, i = 1, 2, \dots, n \\ e_i \leq 100 \cdot 1 - p_i, i = 1, 2, \dots, n \end{cases}
 \end{aligned} \tag{30}$$

需要强调的是，模型(30)的建立虽然根据有信贷记录的企业数据而来，但是在计算和求解的过程全程未用到历史信贷记录或银行专业评级，问题一中的附加数据仅被用作模型可靠性的检验，即模型(30)是仅依赖于企业票据、经营数据的策略模型，与历史信贷记录无关，具有较强泛用性，在本问中依旧适用。

### 4.2.2 模型的求解与结果

利用 4.1.5.1 中算法对模型求解，得到第二问向各企业发放的贷款金额。表 6 展示了其中若干企业的结果，所有结果详见附件。

表 6 问题二求解结果示意图

企业编号	计算信用评级	是否放贷	贷款金额	贷款利率
E1	A	是	100	0.0665
E2	A	是	100	0.0665
E3	A	是	100	0.0745
E119	B	是	99.737	0.0945
E301	D	否	—	—
E302	D	否	—	—
...	...	...	...	...

## 4.3 问题三：突发因素下信贷策略模型的调整



#### 4.3.1 问题分析

根据题目要求和查阅资料，突发因素对突发的大型公共事件会对银行的贷款政策、企业的发展情况两方面因素产生影响。

针对银行贷款政策的调整，我们认为银行主要通过上调基准利率水平和下调贷款额度上限来进行风险管控。对这两点我们分别引入风险管控因子，作为乘数对原模型中的利率和约束进行控制性修正。针对企业经营状况的影响，我们将附加风险分为固定风险和经营风险变动两方面，通过线性函数对原违约率和新违约率建立映射关系，从而对模型进行修正。

#### 4.3.2 风险管控因子的引入

根据分析，在突发状况下银行贷款政策发生变化，体现在利率和贷款额度上限的变化。因此我们引入乘子对

$$\xi_1, \xi_2 \quad (31)$$

前者作为利率的修正指标，后者作为贷款额度上限的修正指标。则修正后，各企业银行贷款利率变动为

$$r_i' = \xi_1 r_i \quad (32)$$

随之发生变动的还包括新贷款利率条件下的客户流失率  $l_i'$ ，可通过查阅附件三获取。银行的贷款额度上限变动为

$$\xi_2 \cdot 100 \cdot 1 - p_i' \quad (33)$$

这一变化将改变模型的第三条约束条件，式中  $p_i'$  指引入发展影响映射关系进行修正后的违约率，将在下一节中给出。

#### 4.3.3 发展影响映射关系的引入

根据分析，我们认为突发因素对于企业的影响可以分为两个方面，一方面是对原有违约概率的变动，在负面事件发生时对于大多数企业而言体现为降低，也会有少数领域出现提高，另一方面是市场大形势下任何企业难以避免的客观运营风险，体现为固定的违约概率提升。故我们引入发展影响映射关系

$$p_i' = \lambda p_i + \varepsilon \quad (34)$$

式中  $\varepsilon$  为常数，对于所有企业一致，由市场整体状况所决定； $\lambda$  为可变参数，根据企业的分类和性质决定。

#### 4.3.4 模型的修正结果

综上，我们额外考虑突发因素对银行贷款策略和企业运营状况的影响，引入了参数组 $(\xi_1, \xi_2, \lambda, \varepsilon)$ 对模型进行修正，修正后的模型为

$$\begin{aligned} & \max \sum_{i=1}^n e_i \cdot r_i' \cdot 1 - l_i' - r_f \\ & \min \sum_{i=1}^n e_i \cdot 1 + r_i' \cdot p_i' \\ & s.t. \begin{cases} \sum_{i=1}^n e_i \leq S \\ 10 \leq e_i \leq 100, i = 1, 2, \dots, n \\ e_i \leq \xi_2 \cdot 100 \cdot 1 - p_i', i = 1, 2, \dots, n \\ r_i' = \xi_1 r, i = 1, 2, \dots, n \\ p_i' = \lambda p_i + \varepsilon, i = 1, 2, \dots, n \end{cases} \end{aligned} \quad (35)$$

#### 4.3.5 模型求解和结果分析

##### 4.3.5.1 模型求解算法

本问中模型与问题一模型没有本质性差异，因此沿用问题一的解法。首先利用新的模型参数重新进行风险量化，再将双目标优化转化为单目标，计算各决策变量权重，并依降序进行贷款金额分配，得到本问的求解结果。

##### 4.3.5.2 模型求解结果

在新冠肺炎发生的背景下，重新对问题二数据集进行求解，得到银行对各企业的贷款分配策略。表 7 展示了其中的部分结果，所有企业结果详见附件。

表 7 问题三求解结果示意图

企业编号	计算信用评级	是否放贷	贷款金额	贷款利率
E1	A	是	98.737	0.0798
E2	A	是	98.737	0.0798
E3	A	是	98.737	0.0894
E299	D	否	—	—
E301	D	否	—	—
E302	D	否	—	—
...	...	...	...	...

##### 4.3.5.3 结果的比较和分析

以贷款企业编号为横坐标，分别以该企业贷款金额、贷款利率为纵坐标，绘制在问

题二、问题三背景下贷款策略的折线图和条形图如图 6 所示，直观反映在不同社会、市场背景下银行贷款策略的不同。

从图上看，在新冠肺炎疫情背景下，银行向绝大部分企业的贷款金额有了明显下降，而贷款利率也作了上调，这与银行的风险管控策略相吻合；而有少部分企业在疫情背景下可申请贷款金额比正常时期有所提高，通过对这些企业详细信息的查询，他们共同的特征为都是新兴的科技企业。贷款限额的提高反映了银行对这些企业发展前景的看好，在过去的一段时间里科技企业的快速发展也从历史的角度印证了模型求解结果在一些方面的可靠性。

## 五、模型评价与改进方向

### 5.1 模型优点

- ◆ **适用性强，适用范围广。**模型直接建立了票据信息与风险评估值之间的联系，使得无论企业是否有信贷记录都可以由历史交易信息快速进行评估和贷款策略分析；
- ◆ **准确率较好，分类预测结果准确。**模型计算信誉评级与专业人士给出的信誉评级符合程度高，信誉水平、风险水平评价较准确；
- ◆ **泛化能力强。**模型参数与实际意义联系密切，容易进行修正、推广和改进，在不同使用条件下由较好的泛化能力。

### 5.2 模型缺点

- ◆ **风险量化的指标需进一步改进。**在计算过程中发现模型的评价倾向和专业认识的评价倾向并不能完全重合，需要了解金融学知识进行改进；
- ◆ **参数的选取主观性较强。**部分经验性系数、权重等的确定带有主观性，不能完全准确地反映出客观的原理和情形。

### 5.3 改进方向

- ◆ **更为完整的风险评估体系。**可以跳脱出题目数据，从企业和市场规律的角度对风险量化观测指标进行重新制定，使其更符合显示情况，与银行的考察角度更契合；
- ◆ **风险评估算法的更新。**在更完整的数据条件下，运用随机森林的相关算法，对风险量化模型进行更深层的学习和调整，以提高准确率；
- ◆ **更丰富的原理性分析。**可以结合实际和经济学原理，对模型的计算结果进行更多现实意义下的解释，提高可读性。

## 六、参考文献

- [1] 张婕. 信用风险量化管理在我国商业银行应用分析[D]. 西南财经大学, 2007.
- [2] 金磊, 李钢. 中国企业盈利能力与竞争力[J]. 中国工业经济, 2007(11):7-16.
- [3] 张青. 家电行业供销合作关系稳定性影响因素研究[D]. 南京工业大学, 2015.
- [4] 李明. 基于 AHP 的 TOPSIS 法在供应商选择和评价中的应用研究[J]. 物流工程与管理, 2012,

34(008):67-69.

- [5] Deng H , Yeh C H , Willis R J . Inter-company comparison using modified TOPSIS with objective weights[J]. Computers & Operations Research, 2000, 27(10):963-973.
- [6] 贾俊平,何晓群,金勇进. 统计学(第四版)[M]. 北京:中国人民大学出版社,2009.
- [7] 周志华. 机器学习[M]. 北京:清华大学出版社,2016.

## 七、附件清单

编号	文件名	文件类型	文件内容
1	问题一求解结果	xlsx	第一问中各企业信贷策略的求解结果
2	问题二求解结果	xlsx	第二问中各企业信贷策略的求解结果
3	问题三求解结果	xlsx	第三问中各企业信贷策略的求解结果
4	风险量化与评级 1	xlsx	第一问中各企业风险量化评估值和计算信誉评级
5	风险量化与评级 2	xlsx	第二问中各企业风险量化评估值和计算信誉评级
6	风险量化与评级 3	xlsx	第三问中各企业风险量化评估值和计算信誉评级
7	贷款利率	xlsx	贷款利率与评估值的对应关系
8	preProcess	py	数据预处理 python 程序
9	TOPSIS	py	TOPSIS 风险量化 python 程序
10	cluster	py	贷款利率计算 python 程序
11	loan	py	模型求解 python 程序
12	info	py	相关性校验 python 程序

\*所有附件打包为压缩文件“支撑材料.zip”另行上交