

# 中小微企业的信贷决策

## 摘要

中小微型企业规模小、资产少,银行在中小微型企业贷款过程中为确保自身利益,需对企业信贷违约风险进行评估,从而建立优化的信贷决策模型。本文针对 123 家有信贷记录企业和 302 家无信贷记录企业的企业信息及发票数据进行分析,通过数据整理与量化处理,选取多项指标作为银行放贷风险的评估依据。使用主成分分析法建立具体风险评估模型,给出各企业风险排序。通过设计信贷策略模型,综合上述企业风险排序序列,求得最佳信贷策略方案。

对于问题一,本文分别建立了风险评估模型和信贷决策模型。风险评估模型以主成分分析法为主导,以量化分析所构建的信誉评级、是否违约、月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和增长稳定率这九组数据为重要指标,给出各企业风险指数及风险排序。信贷决策模型在遵循银行利率优先原则的基础上,采用分级划分贷款额和年利率的方法,对高信誉、低风险企业给与更高贷额和更低年利率,对低信誉、高风险企业给与低额度贷款和较高年利率,通过动态调整为各级分配的名额来求取最大利润。

对于问题二,问题二的求解建立在问题一模型的基础之上。302 家无信贷记录业数据缺失信誉评级与是否违约这两项指标,通过构建分类预测模型,以有信贷记录企业的数据为指导,根据 302 家无信贷记录企业数据预测出相应的信誉评级预测和违约预测。为此本文采用 Logit 模型作为分类预测模型。最后在总放贷额度为一亿元人民币的约束条件下,求得银行从中获得的最大利润金额为 852.7 万元,相应的额度划分策略为:批准 75 万元贷款共 64 家,批准 50 万元贷款共 104 家企业。

对于问题三,在问题二模型的基础上,进一步加入了突发因素对银行贷款策略调整的影响。根据实际情况选取了几项突发因素,分别讨论不同突发因素对不同行业产生的影响,综合考虑突发因素影响的规模和受影响行业的影响程度和发展前景等,对已有放贷策略进行调整,以满足银行稳定经济和自身盈利的双重目标。

**关键词:** 量化分析 主成分分析 Logit 预测 整数规划

## 一、问题重述

对于中小微企业的信贷问题,从一方面来说,其企业规模相对较小、缺少用于抵押的资产,并且他们的经营容易受到外部因素的影响,贷款风险较大。与此同时,他们的流动性较差,负债能力也极其有限<sup>[1]</sup>;从另一方面来说:投资更多的中小微企业有利于分散风险,也方便找到未来的新兴产业,所以银行通常依据信贷的相关政策、企业的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,并可以对信誉高、信贷风险小的企业给予利率优惠。

题目中已知:可以贷款的企业的贷款额度为 10~100 万元、年利率为 4%, 15%、贷款期限为 1 年。该题目要求我们根据实际情况和附件中的数据信息,通过建立恰当的数学模型来研究中小微企业的信贷策。

问题一需要我们对 123 家企业(附件 1)的信贷风险进行量化分析,并给出该银行在固定年度信贷总额的条件下对这些中小微企业的信贷策略。

问题二要求我们在问题 1 的基础上,对 302 家企业(见附件 2)的信贷风险进行量化分析,并给出银行在年度信贷总额为 1 亿元时对这些企业的信贷策略。

问题三已知一些突发因素可能会影响企业的生产经营和经济效益,而且突发因素对不同行业、不同类别的企业往往会有不同的影响。针对附件 2 中各企业,综合考虑的信贷风险和可能出现的突发因素(例如:新冠疫情)对各企业的影响,最后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

针对问题一的主要目标是为 123 家有信贷记录企业的制定基本的信贷策略。为此,本文设计了两套模型来进行信贷策略的最终制定。风险评估模型建立在对附件数据量化分析的基础上。通过分析企业信息与票据信息得到:信誉评级、是否违约、月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和增长率稳定性九项指标,采用主成分分析法对上述指标建立决策模型,并利用该模型求解出各企业的风险指数并输出风险排序序列。

信贷策略模型采取分级划档的原则对贷款额和年利率分级,每级分配变量个数名额,按各企业风险序列分配,低风险、高信誉的企业享有更高的贷款额和更低的年利率。最后给出银行所得利润计算公式,通过穷举各级分配的名额数,得到最大利润情况下的具体信贷策略<sup>[1]</sup>。

针对问题二,研究对象为 302 家无信贷记录企业的信贷策略。问题二的主要目标是制定一个为 302 家无信贷记录的企业提供总计一亿元人民币贷款额的信贷策略。本问的解决建立在第一问所提出模型的基础之上。在套用风险评估模型之前,需要设计

一个预测模型，根据所提供的发票数据，为 302 家无信贷企业做出信誉评级与违约预测，使其指标项目满足风险评估模型的输入需求。本文采用 Logit 预测模型对这两项未知指标进行预测。之后套用风险评估模型取得风险序列，再为信贷决策模型增加一亿元贷款总额这一约束条件，最终由信贷决策模型给出最大利润与具体决策方案。

针对问题三，讨论常见的突发因素对银行贷款策略调整的影响。选取了气象与水旱灾害、突发公共卫生事件和贸易冲突作为常见的影响因素，分别讨论了对不同行业的影响。之后对银行的放贷策略调整进行讨论，使得银行的放贷策略在考虑推动经济稳定发展的同时还能满足自身的盈利目标

### 三、模型假设

1. 假设贷款利率仅与顾客流失率有关，不考虑其他影响因素；
2. 银行针对各企业的信贷评估仅考虑企业实力、企业信誉两方面；
3. 银行通过综合评估企业的信贷风险来确定具体的信贷策略，不考虑其他因素。
4. 企业的信贷风险仅由题目中所反映的企业实力、信誉以及供求关系的稳定性所决定，不考虑经营者情况等其他主观因素。
5. 计算企业总收益时不考虑企业的其他成本以及其他需要缴纳的税额。

### 四、符号说明

符号	符号描述
$A$	主成分数据矩阵
$R$	相关系数
$u_j$	第 $j$ 个主成分特征向量
$\lambda_j$	第 $j$ 个主成分特征值
$b_j$	第 $j$ 个主成分信息贡献率
$\alpha_p$	第 $p$ 个主成分累计贡献率
$Z$	信贷风险综合评价
$Rank$	风险序列
$T$	贷款额度上限

$E_k$	第 $k$ 家企业的信誉评级
$A_k$	第 $k$ 家企业的分配额度
$W$	银行放贷获取的利润

## 五、模型建立与求解

### 5.1 信贷风险评估模型

信贷风险评估模型需要对企业信息和发票信息中的数据进行量化处理。由于作废发票无意义，首先应剔除作废发票。企业信息中有信誉评级和是否违约需要量化，发票中蕴含了金额、税额、退货退款等信息，通过量化并计算能够得到企业的月收益增长率  $k$ 、月收益增长稳定性  $S$ 、月均进项金额  $I$ 、月均销项金额  $O$ 、月收益金额  $G$ 、进项税额退款比率  $U$  以及销项税额退款比率  $V$ 。

#### 5.1.1 企业各指标的量化

企业实力可以由企业的月收益增长率  $k$ 、月收益增长稳定性  $S$ 、月均进项金额  $I$ 、月均销项金额  $O$ 、月收益金额  $G$  来表示。

企业的盈利能力由月收益金额  $G$  来表示，企业的月收益金额越高，说明企业的盈利能力越强。月收益  $G$  为：

$$G = O - I$$

其中  $I$  为月均进项金额、 $O$  为月均销项金额，这两项都通过将总金额除以月份得到。

企业的月收益增长率  $k$  是用来衡量企业的收益前景的一个指标，增长率越大说明企业未来的成长潜力越大。企业的月收益增长率需要用到最小二乘拟合，已知在  $n$  个不同的月份  $x_1, x_2, \dots, x_n$  下分别有测量值  $y_1, y_2, \dots, y_n$ ，先假设月收益与时间的关系如下：

$$f(x_i, k, b) = kx_i + b$$

其中， $k$  即为所求月收益增长率。为了寻找参数  $k$  的最优估计值，要使得  $f(x_i, k, b)$  与  $y_i$  的和方差

$$L(f(x_i, k, b), y_i) = \sum_{i=1}^n [f(x_i, k, b) - y_i]^2,$$

最小，即  $\min L(f(x_i, k, b), y_i)$ ,  $L$  越接近于0，说明模型拟合越好。

月收益增长稳定性  $S$  描述的是企业收益增长的稳定程度，稳定性越好说明资金链和现金流越健康。月收益增长稳定性  $S$  为：

$$S = \frac{L(f(x_i, k, b), y_i)}{n} = \sum_{i=1}^n \frac{[L(f(x_i, k, b), y_i) - y_i]^2}{n}$$

其中， $n$  为月份数， $L$  为计算企业的月收益增长率时用到的和方差方程。当月收益增长稳定性  $S$  越接近于 0，说明了月收益增长越稳定。

### 5.1.2 企业信誉的量化

企业信誉由信誉评级  $Q$ 、是否违约  $P$ 、进项税额退款比率  $U$  和销项税额退款比率  $V$  四部分组成。

企业的信誉评级  $V$  由 A、B、C、D 四级组成，其中信誉评级  $Q$  为 D 的企业在原则上无法从银行取得贷款。由于企业信誉评级  $Q$  满足独立同分布，且 A、B、C、D 四级在统计上来看数量近似，根据辛钦大数定律可知，总体的 A、B、C、D 数量也是近似的，由于每级之间的差值很小，可以直接将 A、B、C、D 分别量化为 4、3、2、1。

企业是否违约  $P$  是一个离散值，可以采用 0-1 规划的思路，当违约时  $P=1$  不违约时为  $P=0$ 。

企业的进项税额退款比率  $U$  和销项税额退款比率  $V$  的计算皆为退货并退款的发票数量除以总的发票数量。

### 5.1.3 主成分分析与决策模型建立

信誉评级、是否违约、月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和月收益增长稳定性，这些都是银行可以用于评价一个企业贷款风险的指标。但是围绕一个企业的不同指标不可避免的有相关性，一个企业月收益金额低，必然容易获得较低的信誉评级，也容易发生还款违约的情况。为了能够得到更为独立的评价指标，首先采用主成分分析<sup>[2]</sup>方法来获得独立的特征指标。

用  $x_1, x_2, \dots, x_9$  分别表示信誉评级、是否违约、月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和月收益增长稳定性。用  $i = 1, 2, \dots, 123$  分别表示企业  $E_1, E_2, \dots, E_{123}$ ，第  $i$  家企业  $x_1, x_2, \dots, x_9$  的取值分别记为  $[a_{i1}, a_{i2}, \dots, a_{i9}]$ ，构建数据矩阵  $A = (a_{ij})_{123 \times 9}$

第一步：由于不同指标的量级存在较大的差异，为了公平的考虑各个因素的影响，需要对原始数据的标准化。将各指标值  $a_{ij}$  转换成标准化指标  $\tilde{a}_{ij}$ ，有：

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j}, i = 1, 2, \dots, 123, j = 1, 2, \dots, 9$$

其中,  $\mu_j = \frac{1}{123} \sum_{i=1}^{123} a_{ij}$  为第  $j$  个指标的样本均值;

$$s_j = \sqrt{\frac{1}{123-1} \sum_{i=1}^{123} (a_{ij} - \mu_j)^2}, j = 1, 2, \dots, 9, \text{为第 } j \text{ 个指标的样本标准差。}$$

标准化处理后得到的标准化数据阵记为  $\tilde{A}$ 。

对应地, 有标准化指标变量:

$$\tilde{x}_i = \frac{(x_i - \mu_i)}{\sigma_i}, i = 1, 2, \dots, 123$$

第二步: 计算相关系数矩阵  $R$ 。令  $x_1, x_2, \dots, x_9$  的相关系数矩阵为  $R = (r_{ij})_{9 \times 9}$ , 有

$$r_{ij} = \frac{\sum_{k=1}^{123} \tilde{a}_{ki} \cdot \tilde{a}_{kj}}{123-1}, i, j = 1, 2, \dots, 9$$

其中,  $r_{ii} = 1$ ,  $r_{ij} = r_{ji}$  称为第  $i$  个指标与第  $j$  个指标的相关系数。

第三步: 计算相关矩阵  $R$  的特征值、特征向量并确定主成分。解特征方程  $|\lambda E - R| = 0$  得到相关系数矩阵  $R$  的特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_9 \geq 0$ , 再计算特征向量  $u_1, u_2, \dots, u_9$ , 其中  $u_j = [u_{1j}, u_{2j}, \dots, u_{9j}]^T$ , 由特征向量组成 9 个新的特征指标

$$\begin{aligned} y_1 &= u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \dots + u_{91}\tilde{x}_9 \\ y_2 &= u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \dots + u_{92}\tilde{x}_9 \\ &\vdots \\ y_9 &= u_{19}\tilde{x}_1 + u_{29}\tilde{x}_2 + \dots + u_{99}\tilde{x}_9 \end{aligned}$$

其中,  $y_j$  为第  $j$  主成分,  $j = 1, 2, \dots, 9$ ,

特征值  $\lambda_j$  ( $j = 1, 2, \dots, 5$ ) 的信息贡献率为

$$\begin{aligned} b_j &= \frac{\lambda_j}{\sum_{k=1}^9 \lambda_k}, j = 1, 2, \dots, 9, \\ \alpha_p &= \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^9 \lambda_k} \end{aligned}$$

其中, 当  $\alpha_p \geq 0.90$  时选择前  $p$  个指标变量  $y_1, y_2, \dots, y_p$  作为  $p$  个主成分, 代替原来的 9 个指标变量, 从而可对  $p$  个主成分进行分析。

第四步: 计算综合评价。综合评价为:

$$Z = \sum_{j=1}^p b_j y_j$$

其中,  $b_j$  为第  $j$  个主成分的信息贡献率, 根据综合评价就可以对企业信贷风险

进行评价。

#### 5.1.4 主成分分析模型的求解

利用 *Matlab* 软件计算的相关系数矩阵的前 9 个特征值及其贡献率如表 1 所示。

表 1 特征值及其贡献率

序号	特征根	贡献率	累计贡献率	序号	特征根	贡献率	累计贡献率
1	2.46	27.37	27.37	6	0.90	10.09	90.78
2	1.44	16.05	43.42	7	0.66	7.35	98.14
3	1.19	13.27	56.69	8	0.17	1.85	100.00
4	1.15	12.80	69.50	9	5.02e-16	5.57e-15	100.00
5	1.01	11.19	80.69				

可以看出前 6 个特征值的累计贡献率超过 90%，所以选取前 6 个主成分进行综合评价。

前 6 个特征值对应的特征向量见表 2。

表 2 特征值对应的特征向量

$u_{ij}$	第 1	第 2	第 3	第 4	第 5	第 6
$\tilde{x}_1$	0.1724	-0.0154	0.7144	-0.1625	-0.0183	-0.0018
$\tilde{x}_2$	-0.0397	0.0243	-0.4054	-0.4333	0.567	0.4321
$\tilde{x}_3$	0.6296	-0.0118	-0.1147	-0.0226	0.0094	-0.0663
$\tilde{x}_4$	0.5807	0.2636	-0.0928	-0.0553	-0.0125	-0.1319
$\tilde{x}_5$	-0.4576	0.5206	0.1077	-0.0479	-0.0462	-0.0837
$\tilde{x}_6$	0.1162	0.171	0.0906	0.4541	-0.219	0.8315
$\tilde{x}_7$	-0.0127	0.1727	-0.3027	0.6873	0.2175	-0.3037
$\tilde{x}_8$	0.0532	0.0839	0.438	0.2313	0.7591	-0.0048
$\tilde{x}_9$	0.0962	0.7696	-0.0157	-0.2189	-0.0655	-0.029

将 6 个主成分对应的特征向量代入主成分方程中得到主成分：

$$y_i = \sum_{j=1}^9 u_{ij} \tilde{x}_j, i = 1, 2, \dots, 6$$

再分别以 6 个主成分的贡献率为权重，构建主成分综合评价模型：

$$Z = 27.37y_1 + 16.05y_2 + 13.27y_3 + 12.80y_4 + 11.19y_5 + 10.09y_6$$

把各企业的 6 个主成分值代入上式，可以得到各企业的信贷风险评估结果(下表仅展示一部分，详见附录 1)其中风险低的企业排名靠前且值大，风险高的企业排名靠后且值小。代码见附录 5。

表 3 各企业的信贷风险评估结果

企业	评分	企业	评分	企业	评分	企业	评分
E1	3.80	E87	0.09	E28	-0.06	E46	-0.22
E42	1.65	E95	0.09	E74	-0.06	E37	-0.22
E38	1.07	E24	0.08	E72	-0.07	E75	-0.22
E71	0.76	E33	0.08	E57	-0.07	E53	-0.22
E16	0.32	E18	0.08	E73	-0.07	E43	-0.23
E60	0.30	E47	0.07	E39	-0.07	E77	-0.24
E55	0.28	E49	0.06	E34	-0.07	E106	-0.24

### 5.1.5 信贷策略模型

在风险评估模型建立的基础上,对企业信贷风险进行评估与分档的方法固定风险因素,构建利润与信贷策略关系方程,通过调节信贷策略来达到银行利润最大化。

#### 5.1.5.1 风险序列的设计

由于原则上对信誉评级为D的企业不予放贷,所以在接下来的讨论中可以直接将信誉评级为D的企业剔除。

令参与贷款的企业信誉评级为  $E_k$ , 通过风险评估模型中的主成分分析过程, 可以将企业的信誉评级按照已得到的各企业信贷风险综合评分  $Z$  从高到低的顺序排序, 即可得到风险序列:

$$Rank = (E_k)_{1 \times n}, k = 1, 2, \dots, n,$$

其中,  $E_k \in \{A, B, C\}$ ,  $n$  为取得信贷的企业数, 且  $Z_k > Z_{k+1} (k = 1, 2, \dots, n-1)$

#### 5.1.5.2 贷款额度的划分

本着对信誉高、信贷风险小的企业给予优惠的信贷优惠原则, 对贷款额度上限  $T$  ( $T=100$  单位: 万元) 实行  $m$  级阶梯划分,  $1 \leq m \leq 10$  且  $m \in N_+$ , 为保证最低贷款额度大于等于10万元, 第  $i$  级的额度为:

$$T \left[ \frac{m-i+1}{m} \right]$$

其中,  $i = 1, 2, \dots, m$ 。将贷款额度划分为4级, 分别为25万元, 50万元, 75万元, 100万元, 针对每一阶梯额度, 设置相应的分配企业数目  $x_i (i = 1, 2, \dots, m)$ , 则各额度分配的企业数之和  $\sum_{i=1}^m x_i$  应与参与贷款的企业总数  $n$  相等, 即

$$\sum_{i=1}^m x_i = n$$

按风险序列的队列优先原则对贷款额度进行分配, 则分配额度  $A_k$  与  $E_k$  的映射关系为:



$$A_k = f(x_i, k) = \begin{cases} 100, k < x_1 \\ 75, k < x_1 + x_2 \\ 50, k < x_1 + x_2 + x_3 \\ 25, else \end{cases}$$

在此种额度划分策略下,低风险、高信誉的企业将能够获得更高贷款额度的批准,切实体现了信贷优惠原则,同时也是银行方面的风控策略的体现。

### 5.1.5.3 贷款年利率的划分

针对利率划分策略,同样遵循对信誉高、信贷风险小的企业给予优惠的信贷优惠原则。

通过对附件3表格数据通过SPSS绘制散点图并拟合后不难发现,随着贷款年利率的增长,客户的流失率也随之增长,二者之间总体趋势呈正相关,如图1所示。

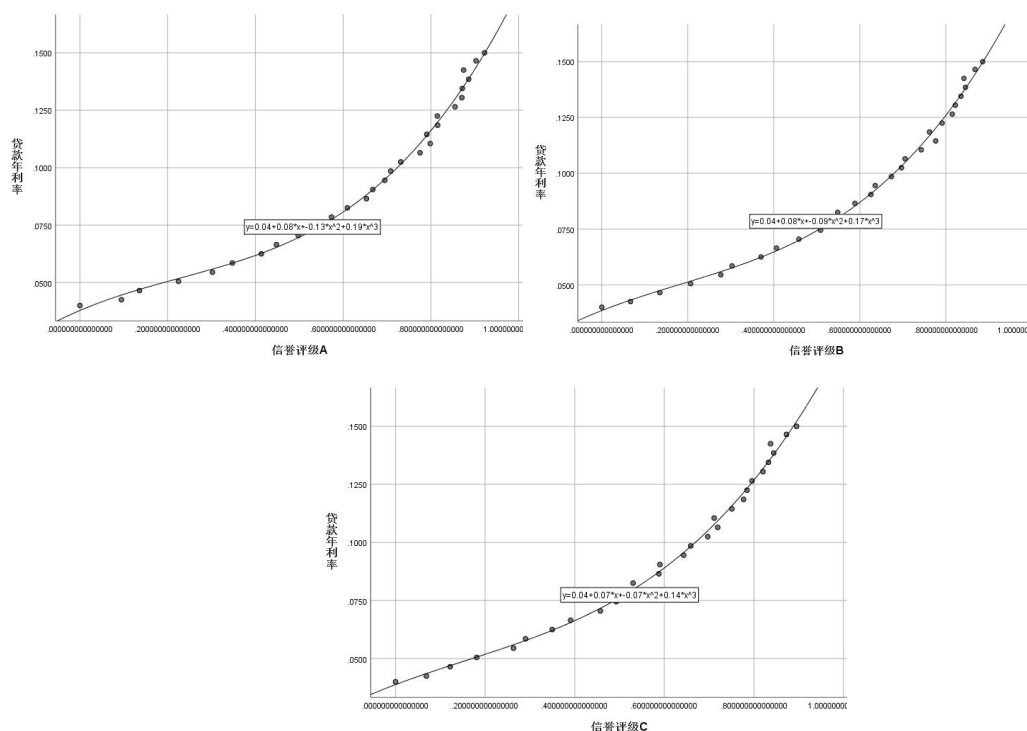


图 1 贷款年利率和客户的流失率

年利率与客户流失率关系图看出更低的贷款利率能够更大程度的保有客户。

在额度划分的分析中明确了:低风险、高信誉的企业拥有更高贷款额度的批准,高风险、低信誉的企业将批准较少的贷款额度。对于前者,设置较低的贷款利率能够让银行最大程度保有与该企业的信贷业务同时保证银行拥有稳定的利润来源且不会面临过高风险;对于后者,由于风险问题和信誉问题的存在,放贷金额低且业务关系不稳定,故设置相对较高利率来最大化银行的收益。

参考额度划分策略,对贷款年利率及客户流失率实行  $m$  级划分。划分后,设贷

款年利率为  $b$ ，客户流失率为  $p$ ，则客户保有率为  $1-p$ ，则平均到所有贷款申请人的年利率为：

$$\bar{R} = b(1-p),$$

$\bar{R}$  值统一了利率与客户流失率这两个变量，且与利润呈正相关。

在分级利润划分的策略下，分别提取  $\bar{R}$  值在各评级下、各分级区域内的极大值作为序列  $E_k$  的  $\bar{R}$  值映射，记为  $R_{km}$ ，映射方式使用额度分配所采用的优先队列方法，并在此基础上加入信誉评级的分支映射模式，可得

$$R_{km} = g(x_i, E_k), i = 1, 2, \dots, m, k = 1, 2, \dots, n$$

选择  $R_{km}$  对应的利率  $b_{km}$  代入计算。

#### 5.1.5.4 利润方程的描述

通过上述分析，已经取得了计算银行利润的全部指标。

对风险序列的每一项信誉评级  $E_k$  都有

$$W_k = b_{km} A_k Y,$$

其中， $W_k$  表示银行从第  $k$  家企业获得的利润， $b_{km}$  表示第  $k$  评级第  $m$  分级  $R_{km}$  对应的利率， $A_k$  表示  $E_k$  对应取得的贷款额度  $Y=1$  表示贷款期（以年为单位）。

对全部的  $n$  家企业，银行取得的利润合计  $W$  为：

$$W = \sum_{k=1}^n W_k$$

将式  $A_k = f(x_i, k)$  带入上式得：

$$W = \sum_{k=1}^n f(x_i, k) b_{km} Y$$

考虑到上式中  $k, b_{km}, Y, n$  均为已知量，故  $W$  是关于  $x_i$  的多元函数，即

$$W = \sum_{k=1}^n f(x_i, k) b_{km} Y$$

$$s.t. \begin{cases} \sum_{i=1}^m x_i = n \\ \sum_{k=1}^n A_k = c \\ x_i > 0 \end{cases}$$

其中， $c$  是信贷总额。

在确定分级数目  $m$  的情况下，最终将利润表达为由  $x_1, x_2$  和  $x_3$  构成的三元一次线性方程， $x_4$  用  $n - x_1 - x_2 - x_3$  表示，在约束条件  $x_i > 0$ 、 $\sum_{i=1}^m x_i = n$  和  $\sum_{k=1}^n A_k = c$  的限制下，存在一组或多组  $(x_1, x_2, x_3)$  使利润  $W$  取得最大。利用穷举算法，计算满足约束的全部  $(x_1, x_2, x_3)$  对应的利润  $W$ ， $W_{\max}$  即为最终求取的最大贷款利润额。

### 5.2 问题二的模型的建立与求解

为了预测附件 2 中 302 家中小微企业的违约情况和信誉评级, 必须找到已经获得的数据与该企业违约情况、信誉评级之间的关系。已知企业的违约情况  $P$  是一个离散值, 令违约为  $P=1$ , 反之为  $P=0$ ; 企业的信誉评级  $Q$  由  $A, B, C, D$  四级组成, 分别量化为 4、3、2、1。由于 Logit 模型是离散选择模型, 相较于线性回归模型能处理因变量是定性变量的情况, 且不要求参加拟合的因子满足正态分布或者等方差, 该模型使解释变量所对应的所有预测值 (概率值) 都落在  $(0, 1)$  间, 所以采用 Logit 模型求解。

#### 5.2.1 基于 Logit 模型的违约情况和信誉评级模型

建立多元 Logit 模型<sup>[3]</sup>如下:

$$P = \frac{\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \beta_{k+1})}{1 + \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \beta_{k+1})}$$

对该模型进行变换, 得到:

$$\ln \frac{P}{1-P} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \beta_{k+1}$$

其中,  $k=1,2,\dots,n$ ,  $x_k$  是参加拟合的因子,  $\beta_{k+1}$  是需要拟合的参数,  $P$  是预测得到的概率值。

#### 5.2.2 违约情况模型的求解

在违约情况模型的建立过程中参与拟合的因子  $x_k(k=1,2,\dots,7)$  分别是月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和月收益增长稳定性。

由于企业的违约情况是用 0, 1 来表示, 0 表示不违约, 1 表示违约, 将问题一中的各项因子和是否违约情况分别作为自变量和因变量代入计算可得到企业违约情况模型的回归参数, 如表 5 所示。代码见附录 7。

表 5 企业违约情况模型的回归参数

$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$
0	-4.08e-16	3.09e-15	5.70e-16	1.94e-16	-1.06e-16	-1.07e-15	-1.26351

将 302 家相应的数据代入所得模型, 可以预计这 302 家皆为 0, 即不存在违约的情况, 如图:

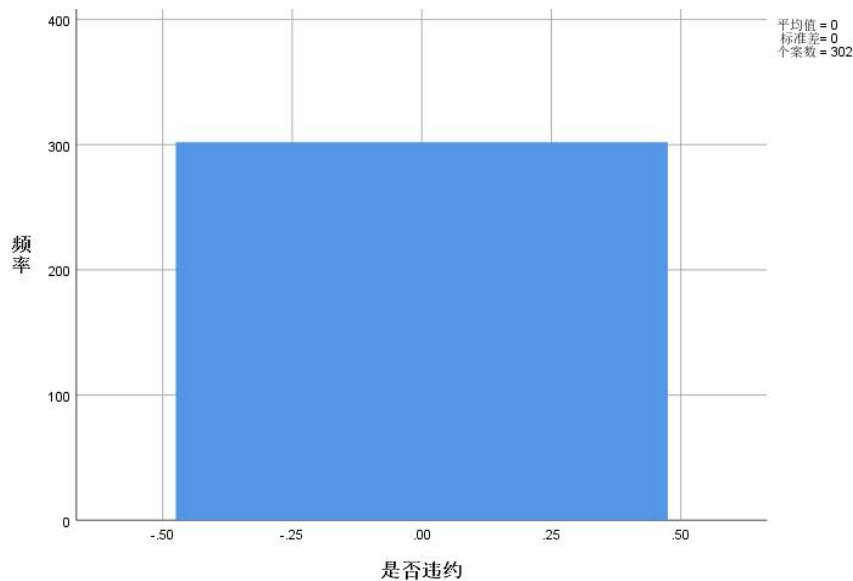


图 2 302 家无信贷公司违约情况

在信誉评级模型的建立过程中参与拟合的因子  $x_k(k=1,2,\dots,8)$  分别是是否违约、月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和月收益增长稳定性。

考虑到有  $ABCD$  四个级别，所以选择分两轮运用 Logit 回归模型对 123 家有信贷记录的企业数据进行回归拟合，得出两次拟合的回归参数值。在第一轮运用 Logit 回归模型时，将 123 家企业按照信誉评级分为  $AB$  和  $CD$  组共两组；在第二轮时，分别对  $AB$  组和  $CD$  组组内运用 Logit 模型，从而将其分为  $A, B$  和  $C, D$  共 4 类，计算可得到企业信誉评级模型的回归参数如表 6 所示。代码见附录 8。

表 6 企业信誉评级模型的回归参数

	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\beta_9$
第一轮	-0.139	0.000	0.008	-0.119	-0.950	-0.170	0.118	0.023	-0.006
第二轮 $AB$ 组	0.240	0.021	-0.578	0.000	0.370	0.120	0.026	-0.134	-0.165
$CD$ 组	-0.650	-0.056	0.000	0.022	-0.072	-0.043	-0.065	-0.09	0.293

在获得上述数据后，对无信贷记录的 302 家企业运用第一轮得到的模型进行第一轮分类，将企业的信誉评级分为  $AB$  组和  $CD$  组，接下来再分别对这两组企业运用第二轮得到的模型进行最后的分类，即可求得这 302 家企业的信誉评级，其中 4、3、2、1 分别表示  $A$ 、 $B$ 、 $C$ 、 $D$ ，具体分类情况如图 3 所示。

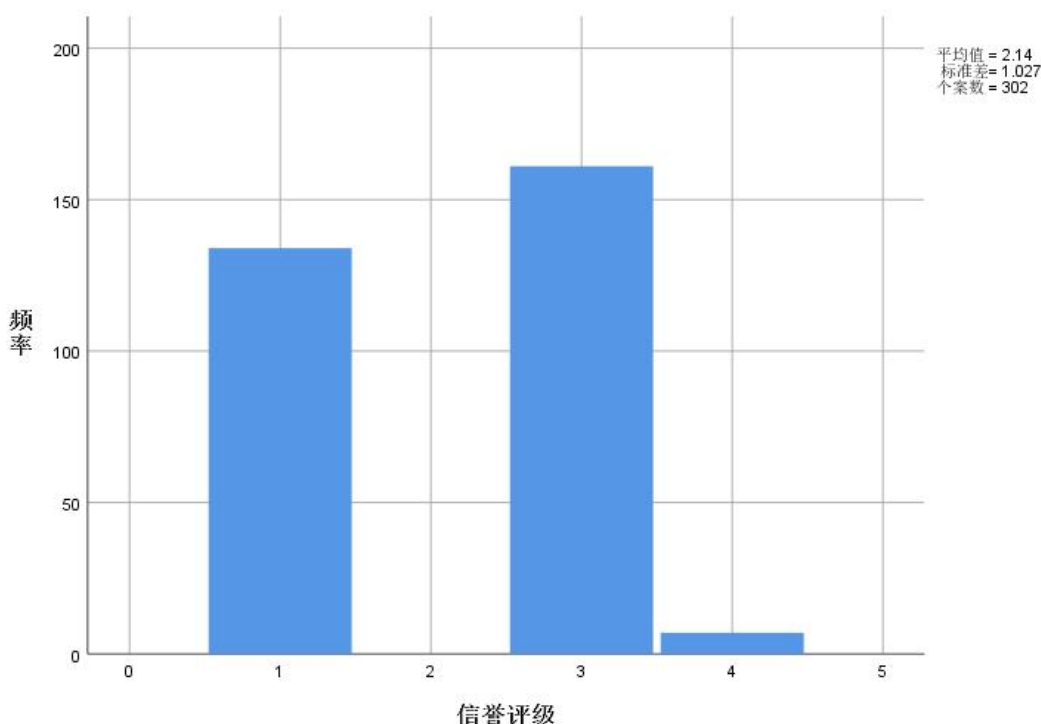


图 3 302 家企业的信誉评级

可以发现，评誉等级为 A 的有 7 家公司；评誉等级为 B 的有 161 家公司；评誉等级为 C 的有 0 家公司评誉等级为 D 的有 134 家公司。

### 5.2.3 基于风险评估模型和信贷策略模型的放贷策略

在问题一中，本文建立了信贷风险评估模型和信贷决策模型。在信贷决策模型中，基于信贷风险评估模型提供的主成分分析方程，设计了企业信贷风险序列作为决策的重要数据指标。

利用上文建立的 Logit 预测模型，可以得到 302 家企业信誉评级和违约概率这两组数据指标。

通过对附 2 中 302 家无信贷记录企业进销票据信息的整理，可以得到该 302 家企业月均进项额、月均销项额、月收益金额、进项税额退款比率、销项税额退款比率、月收益增长率和月收益增长稳定性这七组数据指标。

由于是否违约的值均为否，因此该变量对最后的结果没有影响，因此可以直接剔除。合并上述取得的 302 家企业的 8 组数据指标后，按银行对信誉评级为 D 的企业不予放贷的原则对 302 家企业进行筛选，剔除 34 家信誉评级为 D 的企业后，对剩下符合银行贷款条件的 168 家企业进行信贷风险评估，将 8 项指标带入主成分分析方程中，利用 *Matlab* 软件计算的相关系数矩阵的前 8 个特征值及其贡献率如表 7 所示。

表 7 302 家特征值及其贡献率

序号	特征根	贡献率	累计贡献率	序号	特征根	贡献率	累计贡献率
1	3.47	43.41	43.41	5	0.64	7.96	93.38
2	1.54	19.37	62.78	6	0.34	4.30	97.68
3	0.98	12.25	75.02	7	0.19	2.32	100.00
4	0.83	10.40	85.42	8	1.20e-16	1.50e-15	100.00

可以看出前 5 个特征值的累计贡献率超过 90%，所以选取前 5 个主成分进行综合评价。前 5 个特征值对应的特征向量见表 8。

表 8 特征值对应的特征向量

$u_{ij}$	第 1	第 2	第 3	第 4	第 5
$\tilde{x}_1$	0.4634	-0.0641	-0.0818	0.0293	0.0627
$\tilde{x}_2$	0.1474	0.3731	0.7496	-0.3578	0.2753
$\tilde{x}_3$	0.5197	0.0793	0.1123	-0.0479	0.0423
$\tilde{x}_4$	0.5139	-0.0484	-0.1485	0.0771	-0.0534
$\tilde{x}_5$	-0.0052	0.4941	0.2299	0.7950	-0.2612
$\tilde{x}_6$	-0.1012	0.5363	-0.4693	0.0272	0.6867
$\tilde{x}_7$	0.0387	0.5622	-0.2945	-0.4678	-0.6102
$\tilde{x}_8$	0.4665	-0.0140	-0.1920	0.1059	0.0591

分别以 5 个主成分的贡献率为权重，构建主成分综合评价模型：

$$Z = 43.41y_1 + 19.37y_2 + 12.25y_3 + 10.40y_4 + 7.96y_5$$

把各企业的 5 个主成分值代入上式，计算得到该 168 家企业的风险排序，风险程度随排序递增。将各企业信用评级按风险程度的顺序列出，即可构建风险序列。得到各企业的信贷风险评估结果（下表 9 仅展示一部分，详见附录 2），其中风险低的企业排名靠前且值大，风险高的企业排名靠后且值小。代码见附录 5。

表 9 168 家企业的信贷风险评估结果

企业	评分	企业	评分	企业	评分	企业	评分
E125	6.88	E184	0.07	E216	-0.21	E225	-0.44
E124	3.97	E151	0.07	E177	-0.21	E208	-0.45
E145	3.30	E159	0.06	E204	-0.21	E229	-0.46
E127	2.71	E156	0.06	E209	-0.21	E265	-0.46
E126	2.64	E175	0.06	E271	-0.22	E212	-0.47
E155	2.01	E193	0.06	E243	-0.22	E248	-0.49

#### 5.2.4 分级确定

在信贷决策模型中，本文确立了贷款额度和年利率按风险分级的策略。

问题一中令  $m=4$ , 贷款额度分级情况如表 10 所示。

表 10 贷款额度划分情况

第一级	第二级	第三级	第四级
100 万元	75 万元	50 万元	25 万元

### 5.2.5 最大利润的求解

在问题一的信贷决策模型中, 给出了银行贷款利润的方程组表达, 根据年度信贷总额为 1 亿元的题设条件下, 将该条件转换为约束项加入到方程组中, 即

$$W = \sum_{k=1}^n f(x_i, k) b_{km} Y,$$

$$s, t \left\{ \begin{array}{l} \sum_{i=1}^m x_i = n \\ \sum_{k=1}^n A_k = c \\ x_i > 1 \end{array} \right.$$

其中  $c=10000$  万元,  $m=4$ ,  $n=168$ 。

最终将利润表达为由  $x_1, x_2$  和  $x_3$  构成的三元一次线性方程,  $x_4$  用  $n-x_1-x_2-x_3$  表示, 在约束条件  $x_i > 0$ 、 $\sum_{i=1}^m x_i = n$  和  $\sum_{k=1}^n A_k = c$  的限制下, 存在一组或多组  $(x_1, x_2, x_3)$  使利润  $W$  取得最大。

利用穷举算法, 计算满足约束的全部  $(x_1, x_2, x_3)$  对应的利润  $W$ ,  $W_{\max}$  即为最终求取的最大贷款利润额。

综上所述, 放贷策略和最大贷款利润如表 11 所示。代码见附录 6。

表 11 放贷策略和最大贷款利润

	贷款额度 (单位: 万元)	分配数量
第一级	100	0
第二级	75	104
第三级	50	64
第四级	25	0
利润: 852.7 万元		

### 5.3 问题三的模型的建立与求解

中小微企业作为国民经济的重要基础, 在社会经济的发展中占据了十分重要的地位。中小微企业是我国市场经济的主力军, 数量达 10140 万户, 占据了整个市场主体

数量的 90%，为我国近 80%的城镇居民解决了就业问题，确保了国民经济的活力与繁荣。

但除此之外，中小额企业大多数具有经营规模小、风险承受能力差、管理机制不健全、自有资产不足等特点，处于社会产业链分工的底层，供应链与市场不稳定，风险防范能力差，极易因为突发因素的影响导致经营困难情况的出现。

银行除了要满足自身的盈利目的，还要承担起保就业、稳经济、利民生的社会责任，当中小微企业处于风雨飘摇时，银行会调整信贷策略，引导资金流向更需要的地方。

### 5.3.1 突发因素的选择

根据中国 2007 年起施行的《中华人民共和国突发事件应对法》的规定，突发事件，是指突然发生，造成或者可能造成严重社会危害，需要采取应急处置措施予以应对的自然灾害、事故灾难、公共卫生事件和社会安全事件<sup>[4]</sup>。综合法律、论文和实际情况，在这四大类中选取了常见的气象与水旱灾害、突发公共卫生事件和贸易冲突共 3 项突发因素作为讨论对象。

### 5.3.2 帮扶情况下的策略调整

银行会选择在自身利益不会大规模受损的情况下对需要帮扶的企业提供额外的贷款资助。由于对行业的定向帮扶贷款属于低息贷款，可以确定银行从中的获利将相对减少，因此定向帮扶贷款的规模应可能小，剩余信贷额度遵循问题二的放贷策略，此种情况下获利最丰。

## 1. 气象与水旱灾害

近年来，随着全球变暖等环境问题的出现，导致气候变化带来的气象灾难日趋增多，气象突发情况对我国中小微企业有着诸多影响，其中对农业生产型的企业发展影响最为巨大。2020 年 6 月 2 日至 7 月 2 日，中央气象台连续 31 天发布暴雨预警。洪涝灾害造成农作物受灾面积 861 千公顷，按照同等减产程度推算，今年已因洪涝灾害减产粮食 16.25 亿公斤，相当于 2018 年全国粮食总产量的 0.25%。

我国农业生产具有农业基础设施落后、机械化水平较低、人均耕地面积少等特点，且我国人口基数大，所需口粮多，农业生产一旦受到破坏，将严重威胁到我国的食物安全。因此，银行对受灾农业生产企业的贷款救助起到了至关重要的作用。

在 302 家申请贷款的企业中，涉及到农、林、牧、渔业的企业占比 0.99%，合计 3 家，其中有 1 家不满足贷款条件。可以降低贷款门槛，设立 200 万元利率为 4%的专项资金用作给灾后农业生产重建贷款。

## 2. 突发公共卫生事件



以 2020 年爆发的新冠肺炎为例，在新冠肺炎疫情的持续冲击下，我国经济收到了严重的影响，与疫情发生前我国的经济变动相比，在疫情爆发情况下，我国中小微企业面对的不确定性因素显著增加，相应的企业的信誉评级与企业实力也出现较大变化。

通过数据调研，一季度国内生产总值 206504 亿元，按可比价格计算，同比下降 6.8%。如果分产业看，第一产业增加值 10186 亿元，下降 3.2%；第二产业增加值 73638 亿元，下降 9.6%；第三产业增加值 122680 亿元，下降 5.2%。

总体来看，我国各行各业的企业在疫情的影响情况下均受到了不小的打击，急需大量的现金流来确保资金链的不断裂。

由于疫情的应对措施有国家统一的政策引导，每家银行只需执行与问题二一致的放贷策略，同时在国家的统一指导下，面向制造业、服务业、农业等紧密关系着就业率和国计民生的产业采取降息 10 个基准点的策略。

### 3. 贸易冲突

自 2018 年以来，我国和美国之间爆发了严重的贸易摩擦，考虑到贸易摩擦对各类企业的不同影响，贸易摩擦对我国企业的影响主要体现在制造业和与进出口相关的企业上。

根据对美国出口金额与销售产值比重可以估算对就业的潜在影响，各类行业受影响的潜在规模合计为 330.5 万人，这占各工业行业从业人数的 3.8%，大约会拉高总体失业率 0.4 个百分点。目前美国对我国加收关税约为 12%，对于大部分利润率小于 12% 的企业将难以实现盈利，面临着倒闭的可能性。考虑到制造业本身行业的总规模为 25.17%，同时我国当前的制造业也存在同质化严重、生产效率低下、产能过剩等问题，进行供给侧结构性改革，对落后产能、过剩产能的淘汰也是必不可少的举措，此举还能推动产业升级和转型。

在生物学中，种群在自然条件下呈现逻辑斯蒂增长的特点，逻辑斯蒂方程为：

$$\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right)$$

其中， $r$  是瞬时增长率， $K$  是环境容纳量， $N$  是种群数量，特定种群的  $r$  时定值。此方程是  $\frac{dN}{dt}$  关于  $N$  的二次函数，当  $N = \frac{K}{2}$  时， $\frac{dN}{dt} = \frac{rK}{4}$  种群的最大增长速率，即当前的种群数量为环境容纳量的一半时，整个种群能迸发出最大的活力与生命力。

同理，经过改革开放四十来年的发展，我国的制造业逐渐出现僵化的特征，类比生物学的种群增长特点，我国的制造业已逼近环境容纳量。因此可以选择按照制造业规模的一半进行贷款，即按信贷总额的 12.5% 进行放贷，通过银行从市场中筛选出一半有活力、有技术、有前景的制造业企业进行放贷，让剩余的企业自然淘汰，同时使

得制造业企业占环境容纳量的一半，更好地激发科技创新，推动技术发展。

## 六、模型评价与推广

### 6.1 模型的优点

1. 文章中配有大量的插图与图表，配合文字与算法表述，使模型更加直观易懂；
2. 本题涉及大规模数据分析与处理，模型中提出的各类指标为量化分析确立了明确的方向；
3. 模型中关于信用等级和违约预测所使用的 Logit 模型是信贷风险管理研究领域非常成熟的预测模型，预测结果可靠性强，对输入的特征数据要求低，操作简单便捷；
4. 本文所提出的模型具备高度模块化的特点，风险评估模型与信贷策略模型高度封装，通过风险序列指标相联系，拥有较高的可移植性，可以方便的应用于该领域的其他工作中。

### 6.2 模型的缺点

1. 在建模初期的数据分析与处理阶段，难免存在尚未考虑到的指标；
2. 穷举算法在处理大规模企业数量的贷款业务时会增加贷款利润的计算耗时；
3. 在面临特殊经济环境时，模型的性能表现将会收到一定程度的影响。

### 6.3 模型的推广

1. 通过更加深入细致的数据整理，可以为风险评估模型提供更多的量化因素，以此增强风险评估结果的可靠性，从而增强输出风险序列的可靠性，最终增强银行决策的可靠性；
2. 通过寻找更优的整数规划算法，将较大程度缩减模贷款利润的计算耗时

## 七、参考文献

- [1]周孝华，张保帅. 微型企业信贷风险及化解机制研究[J]. 经济体制改革，2012，000(004):126-130.
- [2]王文博，陈秀芝. 多指标综合评价中主成分分析和因子分析方法的比较[J]. 统计与信息论坛，2006(05):21-24.
- [3]李萌. Logit 模型在商业银行信用风险评估中的应用研究[J]. 管理科学，2005(2):33-38.
- [4]国务院. 国家突发公共事件总体应急预案[J]. 安全与健康:上半月，2006，12(001):77-79.
- [5] 中国会计学会. 金融保险企业财务制度[M]. 中国金融出版社，1993

附录 1 123 家企业的信贷风险评估结果

企业	评分	企业	评分	企业	评分	企业	评分
E1	3.80	E87	0.09	E28	-0.06	E46	-0.22
E42	1.65	E95	0.09	E74	-0.06	E37	-0.22
E38	1.07	E24	0.08	E72	-0.07	E75	-0.22
E71	0.76	E33	0.08	E57	-0.07	E53	-0.22
E16	0.32	E18	0.08	E73	-0.07	E43	-0.23
E60	0.30	E47	0.07	E39	-0.07	E77	-0.24
E55	0.28	E49	0.06	E34	-0.07	E106	-0.24
E89	0.27	E10	0.06	E6	-0.07	E80	-0.25
E13	0.26	E64	0.06	E88	-0.08	E79	-0.26
E27	0.24	E65	0.05	E30	-0.08	E84	-0.29
E83	0.22	E29	0.05	E68	-0.09	E26	-0.32
E19	0.20	E31	0.05	E25	-0.10	E76	-0.32
E85	0.18	E8	0.04	E67	-0.10	E90	-0.33
E93	0.18	E23	0.04	E50	-0.10	E70	-0.39
E21	0.18	E91	0.03	E66	-0.12	E3	-0.44
E110	0.17	E63	0.02	E61	-0.12	E105	-0.49
E54	0.16	E92	0.02	E44	-0.13	E104	-0.52
E11	0.16	E59	0.01	E45	-0.13	E4	-0.55
E14	0.16	E62	-0.01	E81	-0.15	E98	-0.60
E12	0.15	E58	-0.01	E40	-0.16	E7	-0.67
E15	0.15	E97	-0.01	E41	-0.17	E69	-2.16
E32	0.15	E17	-0.03	E56	-0.18	E106	-0.24
E9	0.14	E51	-0.03	E48	-0.18		
E20	0.13	E5	-0.04	E22	-0.18		
E86	0.11	E35	-0.05	E96	-0.19		
E2	0.11	E78	-0.06	E94	-0.21		

附录 2 168 家企业的信贷风险评估结果

企业	评分	企业	评分	企业	评分	企业	评分
E125	6.88	E184	0.07	E216	-0.21	E225	-0.44
E124	3.97	E151	0.07	E177	-0.21	E208	-0.45
E145	3.30	E159	0.06	E204	-0.21	E229	-0.46
E127	2.71	E156	0.06	E209	-0.21	E265	-0.46
E126	2.64	E175	0.06	E271	-0.22	E212	-0.47
E155	2.01	E193	0.06	E243	-0.22	E248	-0.49
E130	1.75	E200	0.05	E215	-0.22	E247	-0.50
E189	1.18	E140	0.04	E211	-0.22	E254	-0.50
E146	1.16	E252	0.03	E162	-0.23	E256	-0.51
E154	1.02	E217	0.02	E185	-0.24	E231	-0.52
E258	0.91	E167	0.00	E201	-0.25	E274	-0.52
E129	0.89	E169	0.00	E171	-0.26	E222	-0.52
E170	0.74	E285	0.00	E244	-0.27	E242	-0.53
E136	0.74	E157	-0.02	E220	-0.29	E276	-0.56
E149	0.69	E148	-0.03	E194	-0.31	E251	-0.58
E163	0.59	E183	-0.03	E223	-0.31	E262	-0.58
E135	0.57	E241	-0.04	E236	-0.31	E275	-0.60
E267	0.53	E178	-0.04	E218	-0.31	E228	-0.61
E164	0.49	E203	-0.04	E190	-0.31	E257	-0.63
E150	0.47	E180	-0.05	E227	-0.31	E250	-0.64
E133	0.44	E186	-0.07	E269	-0.31	E279	-0.65
E139	0.43	E142	-0.09	E207	-0.31	E253	-0.65
E134	0.41	E181	-0.10	E287	-0.32	E273	-0.65
E168	0.38	E166	-0.10	E206	-0.32	E263	-0.65
E153	0.31	E188	-0.10	E239	-0.32	E278	-0.66
E230	0.30	E240	-0.11	E221	-0.33	E272	-0.66
E187	0.29	E238	-0.11	E198	-0.33	E246	-0.66
E147	0.28	E192	-0.12	E172	-0.33	E288	-0.67
E132	0.26	E160	-0.12	E266	-0.35	E282	-0.67
E165	0.24	E249	-0.14	E234	-0.37	E290	-0.68
E176	0.23	E161	-0.14	E191	-0.38	E259	-0.68
E137	0.22	E179	-0.16	E277	-0.38	E289	-0.68
E131	0.19	E237	-0.16	E260	-0.39	E280	-0.69
E152	0.19	E196	-0.16	E202	-0.39	E286	-0.70
E219	0.18	E205	-0.17	E291	-0.39	E268	-0.71
E143	0.16	E281	-0.17	E261	-0.41	E283	-0.74
E128	0.16	E195	-0.18	E235	-0.41		
E199	0.16	E182	-0.19	E226	-0.41		

E138	0.15	E210	-0.19	E264	-0.41
E144	0.13	E197	-0.19	E232	-0.42
E158	0.11	E213	-0.19	E284	-0.43
E141	0.10	E214	-0.19	E245	-0.43
E174	0.09	E270	-0.20	E233	-0.43
E173	0.08	E255	-0.21	E224	-0.44

### 附录 3

%企业发展前景与运营波动

% clear;clc;close all;

%相关mat 文件: P1.mat P2\_4.mat

j = 1;n = 1;

for i=1:length(T11)

if(i==1 || (i>1 && T11(i,1) ~= T11(i-1,1)))

n = 1;

end

if(i==1 || (i>1 && T11(i,3) ~= T11(i-1,3)))

I(j,1) = T11(i,1);%企业编号

I(j,2) = T11(i,2);%年份

I(j,3) = T11(i,3);%月份

I(j,4) = T11(i,7);%进项加价税合计

j = j+1;n = n + 1;

else

I(j-1,4) = I(j-1,4) + T11(i,7);%逐月累加进项加价税合计

end

end

j = 1;n = 1;

for i=1:length(T12)

if(i==1 || (i>1 && T12(i,1) ~= T12(i-1,1)))

n = 1;

end

if(i==1 || (i>1 && T12(i,3) ~= T12(i-1,3)))

S(j,1) = T12(i,1);%企业编号

S(j,2) = T12(i,2);%年份

S(j,3) = T12(i,3);%月份

S(j,4) = T12(i,7);%销项加价税合计

j = j+1;n = n + 1;

else

S(j-1,4) = S(j-1,4) + T12(i,7);%逐月累加销项加价税合计

end

end

%连接 I 与 S

```

for i=1:length(I)
    for j=1:length(S)
        if(I(i,1) == S(j,1) && I(i,2) == S(j,2) && I(i,3) == S(j,3))
            I(i,5) = S(j,4);
            break;
        end
        I(i,5) = 0;%无销项加税总计记为0
    end
end
end

```

%流水计算

```

n = 1;h = 1;
for i=1:length(I)
    if((i>1&&I(i-1,1)~=I(i,1)))
        for j=1:n-1
            %I(i-j,7) = n-1;
            B(h,1) = n-1;%企业有效数据月数
        end
        n = 1;h = h+1;
    end
    I(i,6) = I(i,5) - I(i,4);%月流水计算
    n = n+1;
end
B(h,1) = n-1;

```

%企业发展趋势曲线拟合

```

for i=1:length(B)
    x = 1:B(i,1);
    h = polyfit(x,I(1:B(i,1),6),1);%企业发展趋势曲线拟合
    B(i,2) = h(1);%拟合系数 k
    B(i,3) = h(2);%拟合系数 b
end

```

%企业-月流水矩阵

```

n = 1;
for i=1:length(B)
    for j =1:B(i,1)
        Y(i,j) = I(n,6);
        n = n + 1;
    end
end
end

```

%残差（企业经营稳定程度）

```

for i=1:length(B)

```

```

D = 0;
for j =1:B(i,1)
    D = D + (Y(i,j)-(B(i,2)*j+B(i,3)))^2;
end
B(i,4) = D/B(i,1);
end

```

## 附录 4

```

%-----退款比率计算-----%
%clear;clc;close all;
%相关mat 文件: refund_rate.mat
src = table2array(S2);
% S1:123 家有信贷记录企业进项税额
% S2:123 家有信贷记录企业销项税额
% S3:302 家无信贷记录企业销项税额
% S4:302 家有信贷记录企业进项税额
[m n] = size(src);
rfnum = zeros(1,123);
rfnum_ = zeros(1,302);
total = zeros(1,123);
total_ = zeros(1,302);
for i = 1:m
    total(src(i,1)) = total(src(i,1)) + 1;
    %total_(src(i,1)-123) = total_(src(i,1)-123) + 1;
    if src(i,2) < 0
        rfnum(src(i,1)) = rfnum(src(i,1)) + 1;
        %rfnum_(src(i,1)-123) = rfnum_(src(i,1)-123) + 1;
    end
end
res = rfnum./total
%res_ = rfnum_./total_

```

## 附录 5

```

%-----主成分决策模型-----%
clc,clear
format long
% data=load('sn1.txt');% 将原始数据保存到纯文本文件中
data=load('sn2.txt');% sn1 为问题一, sn2 为问题二
[m,n]=size(data);
[data,mean,sigma]=zscore(data)% 数据标准化
r=corrcoef(data)% 生成相关系数矩阵
% 下面利用相关系数矩阵进行主成分分析, vec1 的列为 r 的特征向量, 即主成分的系数
[vec1,lamda,beta]=pcacov(r);% lamda 为 r 的特征值, beta 为各个主成分的信息贡献率
lamda,beta

```

```

f= repmat(sign(sum(vec1)),size(vec1,1),1);% 构造与 x 同维数的元素为正负 1 的矩阵
vec2=vec1.*f;% 修改特征向量的正负号，每个特征向量乘以所有分量和的符号函数值

alpha=cumsum(beta) %计算累积贡献率 alpha，第 i 个分量表示前 i 个主成分的累积贡献率
% 通过累积贡献率>=0.90 自动选择主成分的个数，num 为选取的主成分的个数
for num=1:n-1
    if(alpha(num)>=90.0)
        break;
    end
end
num

df=data*vec2(:, [1:num]);% 计算各个主成分的得分
tf=df*beta(1:num)/100;% 计算综合得分
[stf,ind]=sort(tf,'descend');% 把得分按照从高到低的次序排列
stf=stf',ind=ind'

```

## 附录 6

```

%-----利润模型-----%
clear,clc,close all;
amt = [100 75 50 25];%分段额度
% cnt = 168;%申请公司数
cnt = 168;%申请公司数
lev_ = [4 2 4 3 2 2 4 4 4 4 4 2 4 4 3 4 4 4 4 4 3 3 4 4 4 3 4 4 3 4 4 3 3 4 3 4 3
4 4 3 3 2 3 3 3 3 3 3 3 2 3 3 3 2 3 2 3 3 3 3 3 3 2 2 2 2 3 2 2 2 4 2 2 2 2
2 2 2 2 2 3 2 3 3 2 2 2 2 2 2 3 2 2 2]; %公司风险排序
% lev = [4 4 4 4 3 3 4 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3 3 3]; %公司风险排序
MAX = 0;%最大利润
TOT = 1e4;%总金额
xx = [];
%遍历求解
for n1 = cnt:-1:0
    for n2 = (cnt-n1):-1:0
        for n3 = (cnt-n2-n1):-1:0
            tot = n1*amt(1)+n2*amt(2)+n3*amt(3)+(cnt-n1-n2-n3)*amt(4);
            if (tot == TOT)
                w = 0;
                for i = 1:cnt

```



```

%
%A 级客户
    if lev(i) == 4
if lev_(i) == 4
    if i <= n1
        r = 0.0465;
    elseif i <= n2+n1 && i > n1
        r = 0.0745;
    elseif i <= n3+n2+n1 && i > n1+n2
        r = 0.0985;
    elseif i <= cnt
        r = 0.1265;
    end
end
%B 级客户
    if lev(i) == 3
if lev_(i) == 3
    if i <= n1
        r = 0.0585;
    elseif i <= n2+n1
        r = 0.0705;
    elseif i <= n3+n2+n1
        r = 0.0985;
    elseif i <= cnt
        r = 0.1265;
    end
end
%C 级客户
    if lev(i) == 2
if lev_(i) == 2
    if i <= n1
        r = 0.0585;
    elseif i <= n2+n1
        r = 0.0825;
    elseif i <= n3+n2+n1
        r = 0.0985;
    elseif i <= cnt
        r = 0.1265;
    end
end
if i <= n1
    w = w+amt(1)*r;
elseif i <= n2+n1
    w = w+amt(2)*r;
elseif i <= n3+n2+n1

```

```

        w = w+amt(3)*r;
    elseif i <= cnt
        w = w+amt(4)*r;
    end
end
end
if w >= MAX
    MAX = w;
    MAX
    N1 = n1
    N2 = n2
    N3 = n3
    N4 = cnt - n1 - n2 - n3
end
end
end
end
end
end

```

## 附录 7

```

% clear,clc,close all;
% 相关mat 文件: P2.mat

PA = 27/123;%评级为 A 的企业的概率
PB = 38/123;%评级为 B 的企业的概率
PC = 34/123;%评级为 C 的企业的概率
PD = 24/123;%评级为 D 的企业的概率

%违约情况系数估计
X=T0(:,3:3+7)';
X = mapminmax(X,0,1);%因子归一化
X = X';

for i=1:length(X)
    X(i,7 + 1) = 1;%构建常数系数
end

for i=1:length(T0)%概率映射
    if (T0(i,1)==0)
        Z(i,1) = 96/123;%没有违约的概率
    else
        Z(i,1) = 27/123;%违约的概率
    end
end
end

```

```

Y = log(Z./(1-Z));
[b,bint,r,rint,stats] = regress(Y,X);%违约情况系数矩阵

%logit 信誉级别系数估计
k = 8;%因子个数

X0=T0(:,2:2+k)';
X0 = mapminmax(X0,0,1);%因子归一化
X0 = X0';

for i=1:length(X0)
    X0(i,k + 1) = 1;%构建常数系数
end

for i=1:length(T0)%概率映射
    if (T0(i,1)==4 || T0(i,1)==3)
        Z0(i,1)=PA + PB;%AB 评级的总概率
    else
        Z0(i,1)=PC + PD;%CD 评级的总概率
    end
end

Y0 = log(Z0./(1-Z0));
[b0,bint,r0,rint0,stats0] = regress(Y0,X0);%第一次决策系数矩阵

%=====
X1=T1(:,2:2+k)';
X1 = mapminmax(X1,0,1);%因子归一化
X1 = X1';
for i=1:length(X1)
    X1(i,k + 1) = 1;%构建常数系数
end

for i=1:length(T1)%概率映射
    if (T1(i,1)==4)
        Z1(i,1)=PA/(PA+PB);%A 在 AB 的概率
    else
        Z1(i,1)=PB/(PA+PB);%B 在 AB 的概率
    end
end

Y1 = log(Z1./(1-Z1));
[b1,bint,r1,rint1,stats1] = regress(Y1,X1);%%AB 级别再决策系数矩阵

```

```

%=====
X2=T2(:,2:2+k)';
X2 = mapminmax(X2,0,1);%因子归一化
X2 = X2';
for i=1:length(X2)
    X2(i,k + 1) = 1;%构建常数
end

for i=1:length(T2)%概率映射
    if (T2(i,1)==2)
        Z2(i,1)=PC/(PC+PD);%C 在 CD 的概率
    else
        Z2(i,1)=PD/(PC+PD);%D 在 CD 的概率
    end
end

Y2 = log(Z2./(1-Z2));
[b2,bint,r2,rint2,stats2] = regress(Y2,X2);%%CD 级别再决策系数矩阵
%=====

```

## 附录 8

```

% clear,clc,close all;
%相关mat 文件: P2_3.mat

X=T3(:,3:9)';
X = mapminmax(X,0,1);%因子归一化
X = X';
for i=1:length(T3)
    pai0 =
    exp(b(1,1)*X(i,1)+b(1,2)*X(i,2)+b(1,3)*X(i,3)+b(1,4)*X(i,4)+b(1,5)*X(i,5)+b(1,6)*
    X(i,6)+b(1,7)*X(i,7)+b(1,8))/(1+exp(b(1,1)*X(i,1)+b(1,2)*X(i,2)+b(1,3)*X(i,3)+b(1,
    4)*X(i,4)+b(1,5)*X(i,5)+b(1,6)*X(i,6)+b(1,7)*X(i,7)+b(1,8)));
    if(pai0 <= 0.5)
        T3(i,2) = 0;
    else
        T3(i,2) = 1;
    end
end

%
X1=T3(:,2:9)';
X1 = mapminmax(X1,0,1);%因子归一化
X1 = X1';

```

```

for i=1:length(T3)
    pai0 =
    exp(b(2,1)*X1(i,1)+b(2,2)*X1(i,2)+b(2,3)*X1(i,3)+b(2,4)*X1(i,4)+b(2,5)*X1(i,5)+b(
    2,6)*X1(i,6)+b(2,7)*X1(i,7)+b(2,8)*X1(i,8)+b(2,9))/(1+exp(b(2,1)*X1(i,1)+b(2,2)*X
    1(i,2)+b(2,3)*X1(i,3)+b(2,4)*X1(i,4)+b(2,5)*X1(i,5)+b(2,6)*X1(i,6)+b(2,7)*X1(i,7)
    +b(2,8)*X1(i,8)+b(2,9)));
    if(pai0 <= 0.5)
        T3(i,1) = 5;
    else
        T3(i,1) = -5;
    end
end

for i=1:length(T3)%编号赋予
    T3(i,11) = i+123;
end

X31=T31(:,2:9)';
X31 = mapminmax(X31,0,1);%因子归一化
X31 = X31';
for i=1:length(T31)
    pai1(i,1) =
    exp(b(3,1)*X31(i,1)+b(3,2)*X31(i,2)+b(3,3)*X31(i,3)+b(3,4)*X31(i,4)+b(3,5)*X31(i,
    5)+b(3,6)*X31(i,6)+b(3,7)*X31(i,7)+b(3,8)*X31(i,8)+b(3,9))/(1+exp(b(3,1)*X31(i,1)
    +b(3,2)*X31(i,2)+b(3,3)*X31(i,3)+b(3,4)*X31(i,4)+b(3,5)*X31(i,5)+b(3,6)*X31(i,6)+
    b(3,7)*X31(i,7)+b(3,8)*X31(i,8)+b(3,9)));
    if(pai1(i,1) <= 0.5)
        T31(i,1) = 4;
    else
        T31(i,1) = 3;
    end
end

X32=T32(:,2:9)';
X32 = mapminmax(X32,0,1);%因子归一化
X32 = X32';
for i=1:length(T32)
    pai0 =
    exp(b(4,1)*X32(i,1)+b(4,2)*X32(i,2)+b(4,3)*X32(i,3)+b(4,4)*X32(i,4)+b(4,5)*X32(i,
    5)+b(4,6)*X32(i,6)+b(4,7)*X32(i,7)+b(4,8)*X32(i,8)+b(4,9))/(1+exp(b(4,1)*X32(i,1)
    +b(4,2)*X32(i,2)+b(4,3)*X32(i,3)+b(4,4)*X32(i,4)+b(4,5)*X32(i,5)+b(4,6)*X32(i,6)+
    b(4,7)*X32(i,7)+b(4,8)*X32(i,8)+b(4,9)));
    if(pai0 <= 0.5)
        T32(i,1) = 2;
    end
end

```

```
    else
        T32(i,1) = 1;
    end
end
```