

110

68-72

商业银行信用风险评估及其实证研究¹王春峰² 万海暉 张维

(天津大学系统工程研究所)

F832.33

【摘要】旨在将判别分析法应用于商业银行信用风险评估中,并且通过与logit方法相比较,进一步研究了判别分析法的有效性,最后讨论判别分析法作为一种传统的建模工具的优缺点。

关键词:判别分析, logit分析, 信用风险

商业銀行

0 引言

现代商业银行在社会经济发展过程中,发挥着筹集融通资金、引导资产流向、提高资金运用效率和调节社会总需求的作用,是国民经济的“总枢纽”和“调节器”。然而商业银行从她诞生时起,就经受着金融风险的威胁,特别是目前,我国正处于由计划经济向市场经济过渡的时期,银行业作为灵敏度最高的行业,其面临的风险非常突出,而信用风险作为我国商业银行当前面临的主要风险,它的测定、防范与控制更是一个不容忽视的问题。本文主要就商业银行的信用风险进行评估、测度。

所谓信用风险,是指信贷资金安全系数的不确定性,表现为企业由于各种原因,不愿或无力偿还银行贷款本息,使银行贷款无法收回,形成呆帐损失的可能性^[1]。如果商业银行出现较大的信用风险,不能按期收回贷款本息,除影响商业银行贷款业务的正常发展以外,其更严重的后果是一旦现金不足,不能应付提款,激发挤兑,威胁银行的正常经营,甚至使银行面临破产。

在对商业银行信用风险的评估中,继比率分析之后,广泛采用了基于统计判别方法上的预测模型。这些模型被表述为一类分类问题,是定义在基于财务比率集合的多维空间上的模型。其中有著名的Altman的Z-score模型和在此基础上改进的ZETA模型以及评价企业放款的乔瑟模型。这

些预测模型所应用的方法有:回归分析法(Meyer and Pifer, 1970^[2])、多元判别分析法(Altman, 1968^[3])、logit法(Press and Wilson, 1978^[4])、probit法(Barth and Brumbaugh, 1989^[5])、因子-logistic法(West, 1985^[6])等。这些模型主要在两方面有所不同:一是样本分布的假定,二是判别函数的形式。由于建模技术、假设条件和分类标准存在着不同,无法对这些统计预测模型进行直接比较。下面主要采用多元判别分析法来研究商业银行信用风险的评估问题。

1 信用风险评估模型

1.1 判别分析(Discriminant Analysis)

对于这一类问题的研究,都是在Fisher于1936年作出了启发性的研究^[7]之后进行的。在这些预测模型中,最常用的是多元判别分析法(Multivariate Discriminant Analysis)。判别分析法(Discriminant Analysis,简称DA)是根据观察到的一些统计数字特征,对客观事物进行分类,以确定事物的类别。它的特点是已经掌握了历史上每个类别的若干样本,总结出分类的规律性,建立判别公式。当遇到新的事物时,只要根据总结出来的判别公式,就能判别事物所属的类别。

对商业银行信用风险的测度可以转化为衡量企业财务状况的问题,因为信用风险的形成——

¹ 国家自然科学基金重大项目(79790130)和国家教委优秀人才基金资助项目(96-170)。

² 王春峰,博士,教授,通讯地址:天津大学系统工程研究所,邮编:300072。

企业能否如期还本付息,主要取决于企业的财务状况.影响企业财务状况的因素很多,不能仅根据企业的某一项指标,而应根据企业的多项指标来判断企业财务状况的强弱.可把由多个指标构成的综合判断指标叫做判别函数,而 DA 的关键就在于建立判别函数.目前,统计学建立判别函数常用方法有:一是未知总体分布情况下,根据个体到各个总体的距离进行判别的距离判别函数^[8];二是已知总体分布的前提下求得平均误判概率最小的分类判别函数,通常称为贝叶斯(Bayes)判别函数;三是未知总体分布或未知总体分布函数前提下的根据费歇(Fisher)准则得到的最优线性判别函数^[9,10].在这里,选用费歇准则,借助方差分析的思想,通过采用极大化组间差和组内差之比的方法建立判别函数.在满足如下条件时,可以证明,由费歇方法所得出的判别规则最优极小化误判代价.

- a) 每组均服从多元正态分布.
- b) 每组的协方差矩阵相同.
- c) 每组的均值向量、协方差矩阵、先验概率和误判代价是已知的.

由于线性判别函数(LDA)在实际使用中是最方便的,如在距离判别和贝叶斯判别中,在正态总体等协方差时,均导出一个线性判别函数,所以只研究线性判别函数.在满足上述 3 个假设的条件下,该判别函数使误判概率达到最小.

设有 l 个总体 G_1, \dots, G_l , 相应的均值向量和协方差矩阵分别为 $\mu^{(1)}, \dots, \mu^{(l)}; V^{(1)}, \dots, V^{(l)}$. 任给一个体 x , 考虑它的线性函数 $y(x) = b^T x$, 则 x 是来自 G_t 的条件下, $y(x)$ 的均值和方差为

$$\begin{aligned} \bar{y}_t &= E(y(x)/G_t) = b^T \mu^{(t)} & t = 1, \dots, l \\ V_t^2 &= V(y(x)/G_t) = b^T V^{(t)} b & t = 1, \dots, l \end{aligned}$$

令

$$B_0 = \sum_{t=1}^l (\bar{y}_t - \frac{1}{l} \sum_{t=1}^l \bar{y}_t)^2$$

$$E_0 = \sum_{t=1}^l V_t^2 = \sum_{t=1}^l b^T V^{(t)} b$$

B_0 相当于一元方差分析中的组间差, E_0 相当于组内差, 运用方差分析的思想, 选择 b 使

$$\Delta(b) = \frac{B_0}{E_0}$$

达到极大, $\Delta(b)$ 称为判别效率, 如令

$$\mu' = (\mu^{(1)}, \dots, \mu^{(l)})$$

$$B = \mu'^T (I_l - \frac{1}{l}) \mu'$$

$$E = \sum_{t=1}^l V_t$$

则易见

$$\Delta(b) = \frac{B_0}{E_0} = \frac{b^T B b}{b^T E b}$$

显然, B, E 均为非负定阵, 由统计学知识可知, $\Delta(b)$ 的极大值为方程

$$|B - \lambda E| = 0 \quad (*)$$

的最大特征根, 记方程的非零特征根为 $\lambda_1 \geq \dots \geq \lambda_l$. 取 b 对应于 λ_1 的特征向量就能达到要求. 于是有:

定理 费歇准则下的线性判别函数 $y(x) = b^T x$ 的解 b 为方程 (*) 的最大特征根 λ_1 所对应的特征向量 b_1 , 且相应的判别效率 $\Delta(b_1) = \lambda_1$.

若仅用一个线性判别函数不能很好区分各个总体, 可取 λ_2 对应的特征向量 b_2 , 建立第 2 个线性判别函数 $y = b_2^T x$. 如果还不够, 可用 λ_3 建立第 3 个, 依此类推.

当 $l = 2$, 即两个总体的判别时, 可以证明费歇准则下的线性判别函数为

$$DF(x) = x^T \sum^{-1} (\mu^{(1)} - \mu^{(2)}) - 1/2(\mu^{(1)} - \mu^{(2)})^T \sum^{-1} (\mu^{(1)} + \mu^{(2)})$$

其中, μ_1, μ_2 和 \sum 分别是各组均值和共同协方差矩阵, 依据一定判别规则即可对原始样本进行分类. 显然, 判别函数 $DF(x)$ 是各独立变量(或指标)的线性函数. 在满足上述 a, b, c 假设下, 它能使误判达到最小.

此外, 根据 p 个指标所构成的判别函数 $DF(x)$, 其判别有效性是可以进行检验的. 这种显著性检验, 一般采用马哈拉诺比斯(Mahalanobis) D^2 统计量来进行. 这个统计量可直接地设想为总体 G_1 与 G_2 之间的距离, 也叫做“综合距离函数”, 这个统计量如下

$$d^2 = b(\bar{X}^1 - \bar{X}^2)$$

其中, $(\bar{X}^1 - \bar{X}^2)$ 为子样 X^1 和 X^2 指标均值的差异列向量; b 为判别函数 $DF(x)$ 的系数行向量.

可以证明, 下列的统计量

$$F_{l, n_1 - n_2 - p - 1} = \left[\frac{n_1 \cdot n_2}{(n_1 - n_2)(n_1 + n_2 - 2)} \right] \cdot \left[\frac{n_1 + n_2 - p - 1}{p} \right] \cdot d^2$$

是服从自由度为 p 和 $(n_1 + n_2 - p - 1)$ 的 F 分布, 其中 p 为指标的个数, n_1 和 n_2 分别为样本 X^1 和 X^2 的容量. 通过查 F 分布表进行检验, 如果有显著差异, 上面的判别是有效的; 否则, 无效.

1.2 logit 方法

MDA 模型的使用依赖于正态分布的假定, 因此, 当研究一些样本不服从正态分布的问题时, 对应用 MDA 模型得到的结果就会持一种怀疑的态度. 为了消除这种假设的影响, 研究无分布的方法就显得非常必要. 其中一种常用的方法就是 logit 方法.

Logit 分析与判别分析法的本质差异在于前者不要求满足正态分布或等方差, 其模型采用 logistic 函数,

$$Y = \frac{1}{1 + e^{-\eta}}, \eta = c_0 + \sum_{i=1}^p c_i x_i$$

其中 $x_i (1 \leq i \leq p)$ 表示第 i 个指标, c_i 是第 i 个指标的系数, Y 为因变量, 由于 $Y \in (0, 1)$, 所以 Y 又被理解为属于某一类的概率.

由于 logistic 回归不假定任何概率分布, 其判别正确率应高于判别分析结果 (Press and Wilson, 1978^[4]).

2 样本数据和模型构造

2.1 样本数据

以某国有商业银行为例, 选择它的企业客户为对象, 考察它们的短期贷款偿还情况. 所用的财务比率是参照国内财政部考核企业财务状况和国外相关财务比率指标, 对反应企业流动性、盈利性、增长性、速动性及偿债性等方面指标, 利用 STATGRAPHICS 软件包进行主因子分析, 得到以下 5 个比率:

$$X_1 = \text{运营资金} / \text{资产总额}$$

$$X_2 = \text{保留盈余} / \text{资产总额}$$

$$X_3 = \text{息税前利润} / \text{资产总额}$$

$$X_4 = \text{普通股、优先股市场价值总额} / \text{负债帐面价值总额}$$

$$X_5 = \text{销售收入} / \text{资产总额}$$

调查取得的样本总数为 129 家, 其中 65 家企业不能偿还银行信用贷款, 这 129 家企业被划分为两个子样本, 训练子样本由 74 家企业构成, 包

括 36 家能够偿还银行贷款的企业和 38 家不能按时偿还贷款的企业, 检验子样本由 55 家企业构成, 包括 28 家能够偿还贷款的企业和 27 家不能按时偿还贷款的企业.

2.2 模型构造

2.2.1 线性判别模型

应用上述的 5 个比率来构造判别分析模型. 对根据样本数据计算得到的 5 个比率进行 Kolmogorov-Smirnov 检验以考察是否服从正态分布的假设. 在删除异常数据后, X_1, X_2, X_3, X_5 均服从正态分布, 只有 X_4 不服从正态分布, 且明显右偏, 这将会对判别分析的结果产生一定的影响. 其次, 检验样本数据是否满足等协方差矩阵的假设. 这一假设是进行其他分析的前提条件. 放松这个假设不仅会影响各组均值有无显著差异的检验, 而且会对变量显著性的解释和判别函数的形式产生影响. 经检验其协方差矩阵大体上是相等的. 同时, 为了增加模型的有效性, 通过考察相关系数矩阵来检验解释变量的线性相关性, 发现没有变量是其它变量的线性组合. 除了 $\text{corr}(\text{RETA}, \text{EBTA}) = 0.7181$ 外, 其相关系数均不大. 判别分析模型是根据 STATGRAPHICS 软件包计算而得到的.

2.2.2 logit 模型

由于 logit 分析不假定任何概率分布, 因此不需要象判别分析法那样事先进行检验, 可直接运用样本数据, 计算得到 logit 模型.

3 结果分析

应用 STATGRAPHICS 软件包, 最终得到线性判别模型的判别系数和 logit 模型的系数, 分别如表 1 和表 2 所示. 5 个比率中的 4 个在 1% 的显著水平下服从正态分布.

每 1 个模型的预测精度都同时通过训练样本和检验样本表现出来. 通常精度都用两类错误来度量. 把第 1 类错误定义为将不能偿还贷款的企业误判为能偿还贷款的企业错误, 第 2 类错误定义为将能够偿还贷款的企业误判为不能偿还贷款的企业错误. 结果如表 3 所示. 总的看来, 判别分析法在训练样本中的误判要多一些, 而在检验样本中的准确率要比 logit 方法高, 但两种方法

在检验样本中的准确率都比训练时要低得多。就第1类错误而言,判别分析法比logit法总是要高一些;而它的第2类错误率又总是要低一些。在两类错误中,第1类错误显然更为严重,因而它的误判代价也就更高一些。所以从总体看来,在满足判

别分析的假设的前提下,判别分析法与logit法各有优劣,训练时判别分析与logit分析的准确率均达到了90%,就是检验中的准确率也有70%以上,结果还是非常理想的。且经过检验判别分析是有效的。

表1 判别分析结果

变量	WCTA*	RETA*	EBTA*	MVTD	STA*	常量
系数	0.32054	1.78868	5.16077	0.07098	-0.06743	-0.52731

* 在1%的显著水平下服从正态分布

表2 logit分析结果

变量	WCTA	RETA	EBTA	MVTD	STA
系数	-10.96324	18.86221	43.14611	0.59414	-2.26517

表3 训练样本和检验样本的误判

模型	训练样本			检验样本		
	第1类错误	第2类错误	总误判	第1类错误	第2类错误	总误判
LDA	5(13.16%)	0(0.00%)	5(6.76%)	11(40.74%)	3(10.71%)	14(25.45%)
logit	2(5.26%)	1(2.78%)	3(4.05%)	9(33.33%)	7(25.00%)	16(29.09%)

4 结束语

作为多元统计分析中应用性较强的一种方法,判别分析法常运用于生产、科研和日常生活中经常遇到需要判别的问题。它能够比较完善地、迅速地解决问题。与一些比较流行的非参数方法相比,判别分析法建立的模型具有更强的透明度,是一种“白盒”技术,因此这一模型也就比较容易进行解释。模型中变量的系数都具有一定的涵义,代表了指标的重要性程度,可以比较各个变量对模型贡献的大小。

但是,在判别分析模型中也存在着一些因素限制了它的有效应用,因为传统的统计方法总是建立在一定的假设检验的基础之上的,判别分析法也不例外。在商业、经济和金融领域的应用中,统计问题尽管不是我们所遇到的唯一的问题,但也是最常碰到的问题。这些问题大致包括以下几方面:(1)样本分布,(2)各组的离散度,(3)独立变量的显著性解释,(4)减少维数,(5)误判率的估计等^[12]。除此之外还有一些其它的问题,如时间序

列的问题。一些学者将判别分析和因子分析、主成分分析结合起来使用,以解决其中的某些问题。

针对我国商业银行信贷资产质量低下和风险加大的客观因素,建立完善的风险管理机制已成为刻不容缓的任务。而信用风险评估作为风险管理的第1个环节则更需要得到妥善的解决。在本文中,判别分析法作为一种定量化方法,将这一问题转化为根据以往模式对企业进行分类的问题,既避免了人们的主观臆断,又简化了问题的复杂性,为商业银行信贷业务的开展提供了借鉴的手段。

参考文献

- 1 曾国坚,何五星. 银行风险论. 北京:中国计划出版社, 1995,4
- 2 Meyer P A, Pifer H. Prediction of Bank Failures. Journal of Finance, 1970, 25: 853~868
- 3 Altman E I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Journal of Finance, 1968, 23: 589~609

- 4 Press S J, Wilson S. Choosing between Logistic Regression and Discriminant Analysis. J. Amer. Statist. Assoc., 1978, 73: 699~705
- 5 Barth J R, Brumbaugh R D, Sauerhaft D. Thrift Institution Failures: Estimating the Regulator's Closure Rule. G. G. Kaufman (Ed), Research in Financial Services, Greenwich, CT, JAI Press, 1989, 1.
- 6 West R C. A Factor-analysis Approach to Bank Condition. Journal of Banking Finance, 1985, 9: 253~266
- 7 Fisher R A. The Use of Multiple Measurements in Taxonomic Problems. Ann. Eugenics, 1936, 7, 179~188
- 8 方开泰,孙尚拱. 距离判别. 应用数学学报, 5, 2
- 9 张尧庭,方开泰. 多元统计分析引论. 北京: 科学出版社, 1982, 215~224
- 10 方开泰. 实用多元统计分析. 上海: 华东师范大学出版社, 1989, 179~213
- 11 Odom M D, Sharda R. A Neural Network for Bankruptcy Prediction. International Joint Conference on Neural Networks, June 17~20, 1990, II: (San Diego, Calif.), 163~168
- 12 Eisenbeis R A. Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. Journal of Finance, 1977, 32, 875~900

Credit Risk Assessment in Commercial Banks and Its Test

Wang Chunfeng, Wan Haihui, Zhang Wei

Institute of Systems Engineering, Tianjin University

Abstract This article applies discriminant analysis method in credit risk assessment in commercial banks. By comparing to logit method, we study efficiency of discriminant analysis. Issues relating to the potential and limitations of discriminant analysis as a traditional modeling tool are also addressed.

Keywords: discriminant analysis, logit analysis, credit risk