

The Financial Early Warning Model of Listed Companies Based on Lasso Method and Logistic Regression

Lu Qin^{1,2}, Yujia Jin^{1,2}, Zhuoxi Yu^{1,2*}

¹School of Management Science and Information Engineering, Jilin University of Finance and Economics, Changchun Jilin

²Jilin University of Finance and Economics Key Laboratory of Internet Finance, Changchun Jilin

Email: 1484734011@qq.com, 29984404@qq.com, *yzx8170561@163.com

Received: Jul. 2nd, 2017; accepted: Jul. 18th, 2017; published: Jul. 24th, 2017

Abstract

Generally, the more the financial data indicators of listed companies are, the better the early warning is; However, due to a variety of factors, excessive financial indicators lead to multiple collinearity among variables. This paper presents a financial early-warning model of Logistic regression listed companies based on Lasso method. Firstly, the Lasso method is used to select variables for high-dimensional data, which can reduce the data dimension and eliminate the collinearity between variables. Then, the Logistic regression method is used to predict the financial status of listed companies. Simulation experiment shows that the method proposed in this paper can effectively eliminate the redundancy of data, improve the accuracy of early warning, and provide effective reference for enterprise operators.

Keywords

High Dimensional Data, Lasso Method, Logistic Regression, Early Warning

基于Lasso方法与Logistic回归的上市公司财务预警分析

秦璐^{1,2}, 靳雨佳^{1,2}, 于卓熙^{1,2*}

¹吉林财经大学管理科学与信息工程学院, 吉林 长春

²吉林省互联网金融重点实验室, 吉林 长春

Email: 1484734011@qq.com, 29984404@qq.com, *yzx8170561@163.com

*通讯作者。

收稿日期：2017年7月2日；录用日期：2017年7月18日；发布日期：2017年7月24日

摘要

一般情况下，上市公司财务数据指标越多，预警效果越好，但由于多种因素影响，财务指标过多会导致变量间具有多重共线性。本文提出一种基于Lasso方法的Logistic回归上市公司财务预警模型。首先应用Lasso法对高维数据进行变量选择，达到降低数据维度和消除变量间共线性的目的，再用Logistic回归法实现对上市公司财务状况的预警。仿真实验结果表明，本文提出的方法能够有效地消除数据的冗余性，提高预警的精确性，为企业经营者提供有效的参考意见。

关键词

高维数据，Lasso方法，Logistic回归，财务预警

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近几年，我国 A 股市场中出现的一些上市公司，由于各种财务问题被证券交易所执行特别处理 (Special Treatment, ST)。被特别处理的股票，除了涨跌幅度受限制以外，证监会要求在其股票名称前加上提醒性注释“ST”。此外，这些上市公司的中期报告必须审计。如果一个 ST 企业持续亏损，那么它将面临退市的风险。因此，利用上市公司财务数据，建立有效、稳定的财务危机预警模型具有重要的现实意义，为企业投资者制定合理的财务制度，锁定财务风险提供有价值的参考。

本文根据上市公司财务数据的特征，将 Lasso 方法引入到财务危机预警模型中，与 Logistic 回归进行有机结合，提出一种基于 Lasso 方法的 Logistic 回归上市公司财务预警模型。仿真实验结果证明了本文所提出的方法能够提高预警的准确性，有效地反映上市公司财务危机状况，为上市公司财务预警提出新的方法。

2. Lasso 方法

Tibshirani 于 1996 年提出最小绝对收缩选择法 (Least absolute shrinkage and selection operator, Lasso)，它是一种基于系数压缩方法的新的变量选择方法。Lasso 方法的基本思想是在约束各个回归系数的绝对值之和小于某个特定值的情况下，求解使得回归方程的残差平方和达到最小的回归系数的估计值。Lasso 方法能够有效地降低数据的维度，非常适用于高维数据的变量选择。Lasso 方法的数学表达式如下所示：

$$\arg \min \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

$$\text{s.t. } \sum_{j=1}^p |\beta_j| \leq t, \quad t \in [0, +\infty) \quad (2)$$

$$i = 1, 2, 3, \dots, n, \quad j = 1, 2, 3, \dots, p \quad (3)$$

其中， x_{ij} 表示第 i 个样本对应的输入变量， y_i 表示第 i 个样本对应的输出变量， $\lambda \sum_{j=1}^p |\beta_j|$ 表示模型的惩

罚函数, λ 为惩罚参数, λ 取值越大, 剔除掉的变量越少; 反之, λ 取值越小, 剔除掉的变量越多。

通过控制调和参数 t , 使回归系数总体变小, 即自动压缩回归系数。当 t 取一定值时, 一些相关程度较低的系数减小为 0, 得到一些稀疏解, 从而达到变量选择的目的。

$\hat{\beta}_j$ 定义为回归模型的最小二乘估计, $s = \frac{t}{\sum \hat{\beta}_j}$ 表示正则化参数, 则在区间 $[0,1]$ 上, 用 s 的取值对预测误差

进行估计[1]。预测误差如下所示:

$$CV(s) = \sum (y_i - \sum \beta_j(s)x_{ij})^2 \quad (4)$$

当 s 取适当值时, 使误差达到最小, 此时, 即可确定回归系数。

求解 Lasso 模型的方法也层出不穷, 最初, Tibshirani 提出用二次规划进行求解, 一些学者又提出了新的算法, 例如 shooting 算法, 同伦算法等。目前应用较为广泛的是最小角回归算法(Least Angle Regression, 即 Lars)。

Lars 算法的过程是一个残差拟合的过程, 即对变量进行最小二乘回归, 它的主要思想是: 最初令所有变量系数为 0, 从所有自变量中寻找出与因变量相关性最强的一个, 记为 x_{j_1} , 引入回归方程; 然后在 x_{j_1} 的方向上搜索另一个自变量 x_{j_2} , 使 x_{j_2} 与当前残差的相关系数和 x_{j_1} 与当前残差的相关系数相同, 引入回归方程; 接着, 沿着 x_{j_1} 和 x_{j_2} 的等角线方向继续搜索, 找到变量 x_{j_3} , 使 x_{j_3} 与当前残差的相关系数相同, 引入回归方程; 再沿着平分变量 x_{j_1} 、 x_{j_2} 和 x_{j_3} 的夹角方向搜索到满足相关性最强的第四个变量, 以此类推, 直到所有变量进入方程即可。具体过程详见参考文献[2]。

本文采用 C_p 准则作为变量选择的标准, 它是基于残差平方和的准则, 目前得到广泛应用[3]。假设有 m 个自变量 $\{x_1, x_2, \dots, x_m\}$, 从中筛选出 p 个变量 $\{x_{i_1}, x_{i_2}, \dots, x_{i_p}\}$, 其中 $\{i_1, i_2, \dots, i_p\}$ 为 $\{1, 2, \dots, m\}$ 的一个子集, 则 C_p 准则公式如下所示:

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p - 2) \quad (5)$$

其中, RSS_p 是变量 $x_{i_1}, x_{i_2}, \dots, x_{i_p}$ 与因变量的残差平方和, n 为样本容量, $\hat{\sigma}^2$ 是 Y 与 x_1, x_2, \dots, x_m 方差的无偏估计量, 其公式如下:

$$\hat{\sigma}^2 = \frac{RSS_m}{n - m - 1} \quad (6)$$

3. 基于 Lasso 方法的 Logistic 回归模型

Logistic 回归模型是一种非线性概率模型, 它具有对自变量类型无限制, 不要求数据满足方差齐性和正态性, 系数有较强的可解释性等优势, 因此, Logistic 回归模型广泛应用于医学、经济学、社会学以及地理研究等领域。

Logistic 回归分析是将求解自变量与因变量之间关系问题转化成求解某一事件发生的概率问题。结合本文研究的内容, 上市 ST 公司和非 ST 公司作为因变量是二元离散型变量, 这里将被执行特别处理的 ST 公司视为发生了财务危机, 没有被执行特别处理的非 ST 公司视为没有发生财务危机。定义如下:

$$y_i = \begin{cases} 0 & \text{ST公司(财务危机企业), } y_i^* \leq 0 \\ 1 & \text{非ST公司(正常企业), } y_i^* > 0 \end{cases} \quad (7)$$

其中, y_i^* 为假设的连续变量, 代表事件发生的概率。

由于 Lasso 不仅适用于线性模型, 也适用于广义线性模型和半参数模型等其他一些统计模型, 因此, 本文采用基于 Lasso 方法的 Logistic 回归模型, 假设 y_i^* 与自变量 x_{ki} 之间存在某种线性关系, 具体形式如下:

$$y_i^* = \alpha + \sum_{k=1}^n \beta_k x_{ki} + \varepsilon_i \quad (8)$$

由式(8)得到

$$P(y_i = 1 | x_i) = P\left[\left(\alpha + \sum_{k=1}^n \beta_k x_{ki} + \varepsilon_i\right) > 0\right] = P\left[\varepsilon_i > -\left(\alpha + \sum_{k=1}^n \beta_k x_{ki}\right)\right] \quad (9)$$

其中, y_i 表示公司是否发生财务危机; x_{ki} 表示第 i 家公司的第 k 个财务指标。

通常, 误差项 ε_i 服从 Logistic 分布, 属于对称分布, 因此, 公式(9)等价于下面的公式:

$$P(y_i = 1 | x_i) = P\left[\varepsilon_i \leq \left(\alpha + \sum_{k=1}^n \beta_k x_{ki}\right)\right] = \frac{1}{1 + e^{-\varepsilon_i}} \quad (10)$$

公式(10)称为 Logistic 函数, 它呈现 S 型曲线分布, 且值域为 0 到 1 之间。将公式(10)转换成如下形式, 得到 Logistic 回归模型:

$$p_i = \frac{1}{1 + e^{-\varepsilon_i}} = \frac{1}{1 + e^{-(\alpha + \sum_{k=1}^n \beta_k x_{ki})}} = \frac{e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})}}{1 + e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})}} \quad (11)$$

其中, p_i 表示第 i 家公司发生财务危机的概率, 它是由变量 x_{ki} 组成的非线性函数, 并且可以转换成线性函数。

定义第 i 家公司不发生财务危机的概率为

$$1 - p_i = 1 - \frac{e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})}}{1 + e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})}} = \frac{1}{1 + e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})}} \quad (12)$$

由上面可以得到, 第 i 家公司发生财务危机与不发生财务危机的概率之比为

$$\frac{p_i}{1 - p_i} = e^{(\alpha + \sum_{k=1}^n \beta_k x_{ki})} \quad (13)$$

将公式(13)称为事件的发生比, 其取值范围为 0 到 ∞ , 因此, 通过 Logit 变换, 可以将上式非线性函数转换成线性函数, 如下所示:

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \sum_{k=1}^n \beta_k x_{ki} \quad (14)$$

Logistic 回归模型将上市公司发生财务危机概率的预测问题转化成预测某个公司发生财务危机的发生比问题, 该模型不仅能够预测上市公司是否发生财务危机, 还能够直观地反映出公司发生财务危机的概率, 客观地反映公司经营现状[4]。

4. 实验模拟与结果分析

4.1. 数据指标选取

本文选取了 2014 年~2016 年间我国 A 股市场中因财务状况异常而被特别处理的上市公司(即 ST 公司)作为主要研究对象, 剔除存在缺失值的数据后, 共选取 22 家 ST 公司。因为上市公司出现财务危机需要经历一个过程, 上市公司被特别处理是由于之前两年连续亏损, 因此本文选取第 T-3 年的指标数据建立

预警模型, 预测公司在第 T 年是否因财务状况异常而被特别处理。此外, 按照 ST 公司与非 ST 公司 1:3 的配对比例随机选取 66 家非 ST 公司作为匹配样本。选取的这些配对非 ST 公司要与 ST 公司处于相同行业, 资产规模相近, 并且在同一时间段内。最终确定了由 23 家 ST 公司和与之匹配的 66 家非 ST 公司共同组成的研究样本集[5], 如表 1 所示。

Table 1. 2014-2016 A-share listed company
表 1. 2014 年~2016 年 A 股上市公司

公司性质	最新公司全称	股票代码
ST 公司	安徽古井贡酒股份有限公司	000596
ST 公司	鞍山第一工程机械股份有限公司	400021
ST 公司	厦门海洋实业(集团)股份有限公司	400022
ST 公司	汕头宏业(集团)股份有限公司	400025
ST 公司	深圳市中侨发展股份有限公司	400026
ST 公司	珠海鑫光集团股份有限公司	400028
ST 公司	吉林成城集团股份有限公司	600247
ST 公司	南宁化工股份有限公司	600301
ST 公司	西安宏盛科技发展股份有限公司	600817
ST 公司	振兴生化股份有限公司	000403
ST 公司	仁和药业股份有限公司	000650
ST 公司	恒逸石化股份有限公司	000703
ST 公司	湖北洪湖生态农业股份有限公司	400027
ST 公司	包头明天科技股份有限公司	600091
ST 公司	云南景谷林业股份有限公司	600265
ST 公司	山东金泰集团股份有限公司	600385
ST 公司	上海中技投资控股股份有限公司	600634
...
非 ST 公司	山东民和生物科技股份有限公司	836448
非 ST 公司	北京兆信信息技术股份有限公司	430073
非 ST 公司	长城信息产业股份有限公司	000748
非 ST 公司	青海华鼎实业股份有限公司	600243
非 ST 公司	江苏三房巷实业股份有限公司	600370
非 ST 公司	重庆路桥股份有限公司	600106
非 ST 公司	宁夏新日恒力钢丝绳股份有限公司	600165
非 ST 公司	中山公用事业集团股份有限公司	000685
非 ST 公司	武汉高德红外股份有限公司	002414
非 ST 公司	华润双鹤药业股份有限公司	600062
非 ST 公司	吉林紫鑫药业股份有限公司	002118
非 ST 公司	福建海源自动化机械股份有限公司	002529
非 ST 公司	江阴中达软塑新材料股份有限公司	835905
非 ST 公司	浙江万好万家文化股份有限公司	600576
非 ST 公司	武汉禾元生物科技股份有限公司	833101
非 ST 公司	华联控股股份有限公司	000036
非 ST 公司	广州珠江实业开发股份有限公司	600684
非 ST 公司	北京鸢峰科技开发股份有限公司	400010
非 ST 公司	彩虹显示器件股份有限公司	600707

企业的财务指标反映了公司的财务状况和经营成果, 不同时期的指标数据反映出企业不同时期的经营状况[6]。本文从公司的盈利能力、偿债能力、成长能力、现金流量、运营能力、资本结构等方面考虑, 初步选取了 19 个反映公司财务状况的指标变量进行研究[7], 如表 2 所示。本研究数据主要来源于 RESSET 数据库, 使用 SPSS 和 R 软件进行数据处理以及构建模型。

4.2. 实验结果与分析

使用 R 软件中的 Lars 程序包对数据处理, 根据变量的移动顺序来判断指标的重要性, 运行 Lasso 方法得到变量选择次序表, 如表 3 所示[8]。

Table 2. Financial early-warning index

表2. 财务预警指标

所属体系	指标名称	表示符号
盈利能力	每股收益	x_1
	净资产收益率	x_2
	销售净利率	x_3
	营业利润率	x_4
偿债能力	流动比率	x_5
	速动比率	x_6
	产权比率	x_7
成长能力	营业收入增长率	x_8
	净利润增长率	x_9
	净资产增长率	x_{10}
	总资产增长率	x_{11}
运营能力	存货周转率	x_{12}
	应收账款周转率	x_{13}
	流动资产周转率	x_{14}
	总资产周转率	x_{15}
现金流量	销售现金比率	x_{16}
	经营现金净流量	x_{17}
	营业收入现金含量	x_{18}
资本结构	资产负债率	x_{19}

Table 3. Variable selection order table

表3. 变量选择次序表

	x_{19}	x_4	x_2	x_6	x_{17}	x_{18}	x_{12}	x_9	x_7	x_{14}
变量	19	4	2	6	17	18	12	9	7	14
步骤	1	2	3	4	5	6	7	8	9	10
	x_{11}	x_{15}	x_8	x_1	x_{10}	x_{16}	x_{13}	x_5	x_3	
变量	11	15	8	1	10	16	13	5	3	
步骤	11	12	13	14	15	16	17	18	19	

从上表可以看出, 变量选择的顺序依次是 $x_{19}, x_4, x_2, x_6, x_{17}, x_{18}, \dots, x_{13}, x_5, x_3$ 。

在变量选择过程中, 根据 C_p 准则来判断最优选择效果, 即 C_p 值越小, 变量选择效果越好, 计算过程中的各项参数值如表 4 所示。

从表 4 中可以看出, 在第 5 步时, 当 C_p 达到最小值为 1.1919, 此时变量选择效果达到最优, 因此最终选择的变量是 $x_{19}, x_4, x_2, x_6, x_{17}$ 。

图 1 为回归系数解路径图, 可以直观地判断财务指标的选择过程。

从图 1 可以看出, 当回归系数估计值逐渐增大时, 不同指标系数呈现出不同的离散程度[9], 当回归系数估计值逐渐减小时, 部分变量的标准化系数为 0。根据上面所述, 在第 5 步时变量选择效果达到最优, 则可以较为直观地看出所选择的变量分别为 $x_{19}, x_4, x_2, x_6, x_{17}$ 。

对所选择的这 5 个变量进行共线性检验, 如下表 5 所示。

从上表可以看出, 所有条件指数均小于 10, 则可认为这 5 个变量间不存在共线性, 也进一步证明了使用 Lasso 方法做变量选择达到了消除共线性的目的。

从样本数据中随机选取 68 家上市公司(其中包括 17 家 ST 公司和 51 家非 ST 公司)作为训练样本, 其余的 20 家上市公司(其中包括 5 家 ST 公司和 15 家非 ST 公司)作为测试样本。将 Lasso 方法选择的 5 个

Table 4. Table of parameter values

表4. 各项参数值表

步骤	自由度	残差平方和	C_p 值
0	1	16.500	14.9075
1	2	16.207	15.1189
2	3	15.605	13.4344
3	4	13.946	5.2873
4	5	13.491	4.5058
5	6	12.622	1.1919
6	7	12.464	2.2265
7	8	12.304	3.2483
8	9	11.997	3.3706
9	10	11.964	5.1699
10	11	11.775	6.0101
11	12	6.010	7.4879
12	13	11.661	9.3159
13	14	11.570	10.7560
14	15	11.503	12.3480
15	16	11.401	13.7257
16	17	11.313	15.1847
17	18	11.229	16.6689
18	19	11.211	18.5652
19	20	11.119	20.0000

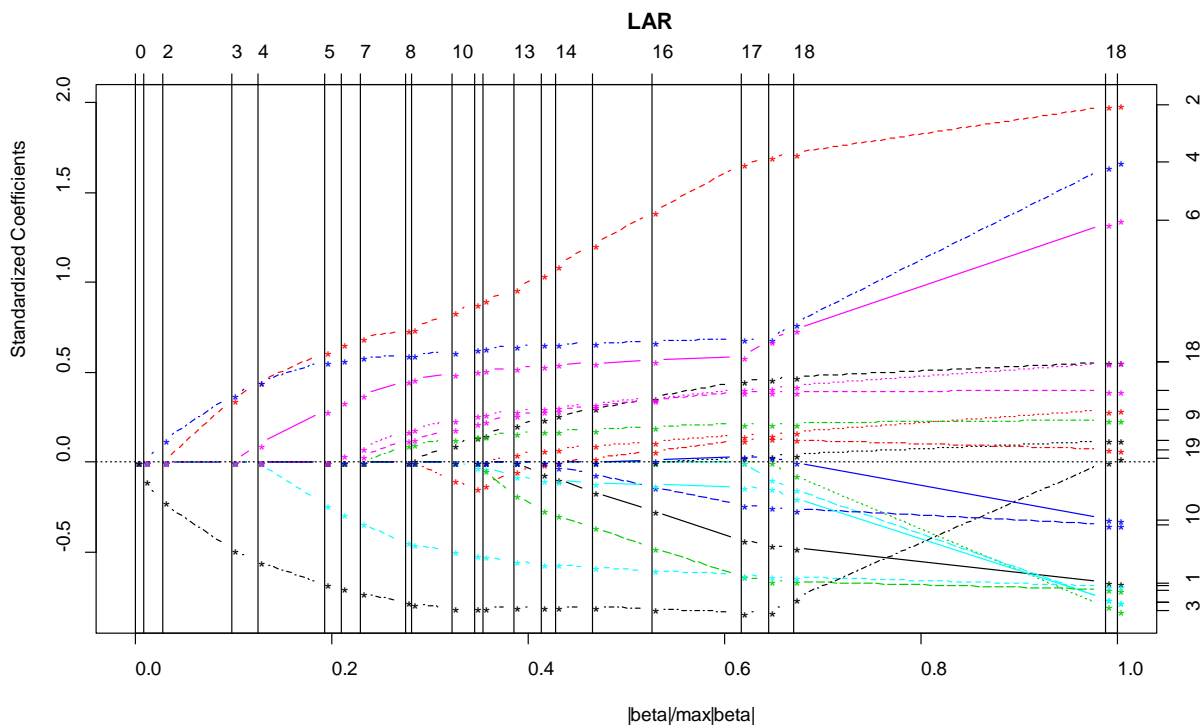


Figure 1. Regression coefficient solution path diagram

图 1. 回归系数解路径图

Table 5. Collinearity diagnosis

表 5. 共线性诊断

维数	特征值	条件指数	方差比例					
			(常量)	x_2	x_4	x_6	x_{17}	x_{19}
1	1.745	1.000	0.12	0.01	0.11	0.07	0.00	0.11
2	1.394	1.119	0.06	0.12	0.10	0.17	0.01	0.09
3	1.008	1.316	0.01	0.04	0.00	0.00	0.93	0.00
4	.901	1.392	0.01	0.84	0.05	0.04	0.03	0.06
5	.541	1.797	0.00	0.00	0.70	0.00	0.00	0.71
6	.411	2.061	0.80	0.00	0.05	0.73	0.04	0.03

变量 $x_2, x_4, x_6, x_{17}, x_{19}$ 作为解释变量, 选择训练样本的数据建立 Logistic 回归模型, 得到的模型统计量, 如表 6 所示。

根据表 6, 将 Logistic 回归分析结果代入公式(15), 得到最终的上市公司财务预警模型为:

$$P_i = \frac{1}{1 + e^{-(0.69 + 2.119x_{2i} - 3.14x_{4i} - 2.266x_{6i} + 0.117x_{17i} - 2.89x_{19i})}} \quad (15)$$

其中, P_i 是第 i 家公司发生财务危机的概率, x_{2i} 表示第 i 家企业的净资产收益率, x_{4i} 表示第 i 家企业的营业利润率, x_{6i} 表示第 i 家企业的速动比率, x_{17i} 表示第 i 家企业的经营现金净流量, x_{19i} 表示第 i 家企业的资产负债率。

Table 6. Model statistic
表6. 模型统计量

	B	S.E.	Wals	df	Sig.	Exp (B)
净资产收益率	2.119	0.844	6.305	1	0.012	8.322
营业利润率	-3.14	1.205	6.797	1	0.009	0.043
速动比率	-2.266	1.055	4.615	1	0.032	0.104
经营现金净流量	0.117	0.556	0.044	1	0.833	1.124
资产负债率	-2.89	1.586	3.318	1	0.069	0.056
常量	0.69	0.452	2.335	1	0.127	1.994

Hosmer 和 Lemeshow 检验财务预警模型的拟合优度, 如表 7 所示。

从表 7 可以看出, 拟合优度检验得到的 p 值为 0.293, 大于 0.05, 说明模型的拟合程度较好。

通常选取概率值为 0.5 作为财务状况正常与财务状况异常的临界点, 即当 $P_i > 0.5$ 时, 则表示该公司财务状况发生异常; 当 $P_i < 0.5$ 时, 则表示该公司财务状况正常。上市公司财务预警模型对训练样本的回判结果如表 8 所示。

回判结果包含两类错误, 第一类错误是将财务危机公司误判为财务正常公司, 第二类错误是将财务正常公司误判为财务危机公司[10]。从表 8 可以看出, 训练样本的回判结果中, 第一类错误率为 23.5%, 17 家 ST 公司中判定有 13 家 ST 公司, 4 家非 ST 公司, 正确率为 76.5%; 第二类错误率为 3.9%, 51 家非 ST 公司中判定有 49 家非 ST 公司, 2 家 ST 公司, 正确率为 96.1%。68 家上市公司的总正确率为 91.2%, 表明该预警模型具有较好的实用性。

同样, 应用上市公司财务预警模型对测试样本进行财务危机预警, 得到的预测结果如表 9 所示。

从表 9 可以看出, 测试样本的检验结果中, 第一类错误率为 20%, 5 家 ST 公司中判定有 4 家 ST 公司, 1 家非 ST 公司, 正确率为 80%; 第二类错误率为 0%, 15 家非 ST 公司全部判定正确, 正确率为 100%。20 家上市公司的总正确率为 95%。

为了检验本文提出的上市公司财务预警模型对财务状况预警的有效性, 应用相同的训练数据和测试数据, 将 Logistic 回归模型与本文提出的模型的预测结果作进一步对比, 图 2 和图 3 分别为训练数据和测试数据的对比结果。

从图 2 和图 3 的对比中可以看出, 与 Logistic 回归模型比, 上市公司财务预警模型对公司财务状况预警具有较高的准确性, 因此, 基于 Lasso 方法的 Logistic 回归上市公司财务预警模型更适用于我国 A 股市场的分析与预测。

综上所述, 从模型的拟和优度检验和预测准确率两方面分析可以看出, 基于 Lasso 方法的 Logistic 回归建立的上市公司财务预警模型不仅能够较好地判定公司的财务状况, 而且具有较强的预测能力, 能够帮助企业管理者及时做好财务风险防控[2]。

5. 结语

由于上市公司财务数据指标较多, 变量之间往往具有多重共线性, 应用 Logistic 回归模型对上市公司进行财务预警时, 当变量之间相关程度较高时, 往往得不到理想的预警结果。针对此, 本文提出一种基于 Lasso 方法的 Logistic 回归上市公司财务预警模型。应用 2014 年~2016 年间我国 A 股上市公司财务数据进行拟合, 仿真实验结果表明, 本文提出的上市公司财务预警模型能够合理地分析和预警公司的财务状况, 提高预警的精确性, 帮助企业经营者有效地应对企业的财务状况。

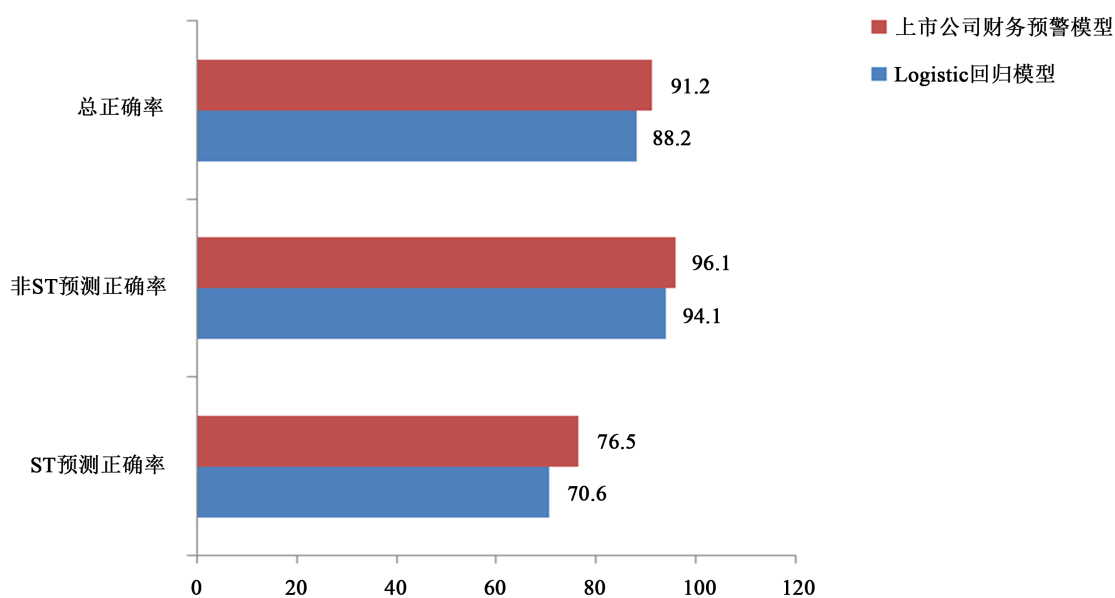


Figure 2. Training data comparison results

图 2. 训练数据对比结果

Table 7. Hosmer and Lemeshow test

表 7. Hosmer和Lemeshow检验

步骤	卡方	df	Sig.
1	9.615	8	0.293

Table 8. Training sample return result

表 8. 训练样本回判结果

已观测		已预测		百分比校正
		企业性质		
		ST	非ST	
企业性质	ST	13	4	76.5
	非ST	2	49	96.1
总计百分比				91.2

Table 9. Testing sample return result

表 9. 测试样本回判结果

已观测		已预测		百分比校正
		企业性质		
		ST	非ST	
企业性质	ST	4	1	80.0
	非ST	0	15	100.0
总计百分比				95.0

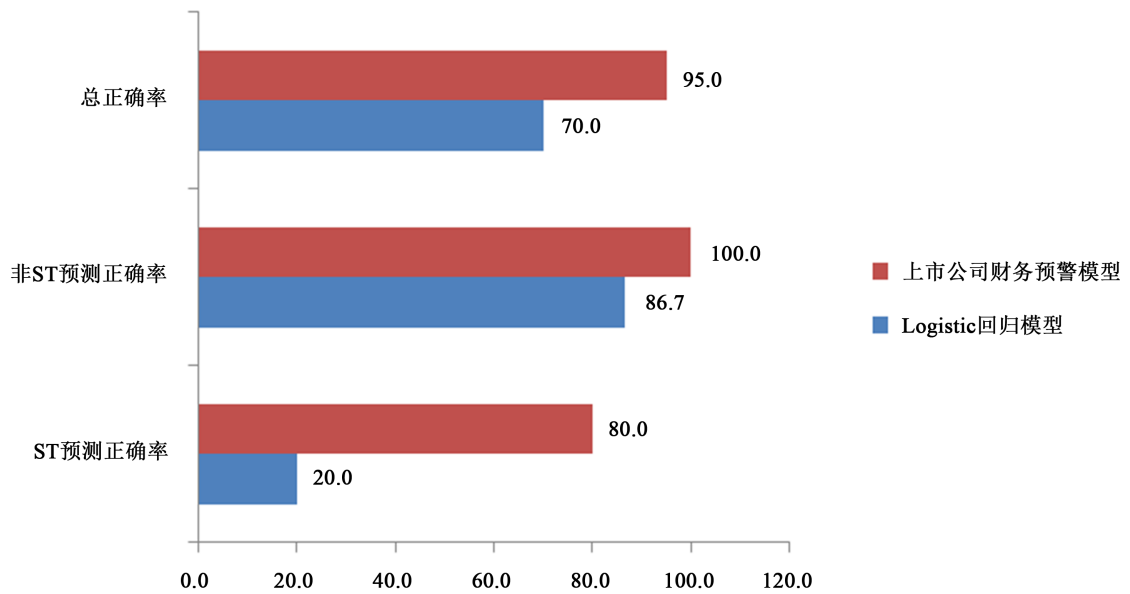


Figure 3. Test data comparison results
图 3. 测试数据对比结果

基金项目

国家社会科学基金(16BTJ020)。

参考文献 (References)

- [1] 张艳丽, 尤晓琳, 强薇, 等. 基于 LASSO 的企业财务危机预警与关键指标选择[J]. 河南师范大学学报: 自然科学版, 2016(3): 160-165.
- [2] Hui, Z. and Trevor, H. (2005) Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society*, **67**, 301-320.
- [3] 孙浩. 基于 Logistic 模型的财务预警研究[D]: [硕士学位论文]. 衡阳: 南华大学, 2013.
- [4] 连晓丽. 我国 A 股上市公司财务危机预警模型实证研究[D]: [硕士学位论文]. 厦门: 厦门大学, 2014.
- [5] 秦志敏. 我国上市公司财务预警变量选择研究[D]: [博士学位论文]. 大连: 东北财经大学, 2012.
- [6] 成鹏伟, 费宇. LASSO 和 A-LASSO 方法在财务预警模型变量选择中的应用[J]. 中国证券期货, 2013(3): 110-111.
- [7] 顾云燕. 基于 Lasso 和 Cox 模型的上市中小企业财务预警分析[D]: [硕士学位论文]. 兰州: 兰州大学, 2016.
- [8] 杨二宝. 上市公司财务危机预警模型的实证研究[D]: [硕士学位论文]. 西安: 西北大学, 2005.
- [9] 李敏捷. Lasso-Logistic 与 Group Lasso-Logistic 模型在出生缺陷研究中的应用[D]: [硕士学位论文]. 太原: 山西医科大学, 2016.
- [10] 罗昊. 基于自适应 LASSO 变量选择的 Logistic 信用评分模型研究[D]: [硕士学位论文]. 南京: 东南大学, 2016.

期刊投稿者将享受如下服务：

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：aam@hanspub.org