

财务比率截面数据分布特征研究

<http://www.criifs.org.cn> 2009年2月16日 刘国光

[摘 要] 财务比率截面数据分布特征对于财务危机预测、信用评级等具有基础性的意义。通过对沪深全部上市公司2000年至2006年8种主要财务比率年度截面数据分布进行拟合分析后发现,高达90%的财务比率截面数据的最佳拟合分布为非正态分布,原始比率数据中有近一半的比率服从对数逻辑分布;经过数据转换后,财务比率的对数逻辑分布特性更为显著。从不同年份间比率截面数据的记忆特性来看,流动性比率表现出完全的记忆,现金负债比率也表现出较强的记忆。通过数据转换或异常值剔除,可以显著提升财务比率数据的正态特性。

[关键词] 财务比率截面数据; 分布特征; 数据转换; 异常值剔除

目前,财务比率分析已被广泛应用于财务危机预测、债券评级、商业信用评分、证券投资分析以及企业并购研究等,这些研究大多是建立在财务比率截面数据为单变量正态分布或多元正态分布的假设基础上,而财务比率的截面数据服从正态分布这一假设是需要严格检验的。如果财务比率截面数据服从严格的非正态分布,就意味着这些研究存在着严重的问题。忽视分布属性这一问题,即使能够得到不错的实际判别效果,也是缺乏基础理论依据的。事实上,应用财务比率预测财务危机和进行信用评级的作用,取决于应用的概率模型能否准确地描述财务比率分布行为。例如,信用管理人员在运用财务比率分析客户的信用水平时,会假定财务比率正态分布确定的信用水平可能显著异于财务比率非正态假设确定的信用水平。

财务比率正态分布对理论模型的构建也会产生影响。例如,最小二乘法和判别分析法通常被用来预测破产风险和进行债券评级,如果财务比率不是服从正态分布,而是服从稳态分布,稳态分布的不连续性就预示着财务比率会发生极端的突然变化,这时,OLS方法将提供无效预测和不准确的置信区间估计,预测财务危机和进行债券评级将变得更加困难。即使应用Logistic回归进行财务危机预测,财务比率为正态分布时的危机预测准确性也明显高出财务比率为非正态分布的预测准确性。此外,如果某种财务比率不服从正态分布,那么,均值将不再是一个合适的基准,将特定比率与均值进行比较就不再恰当。从投资者的角度来说,识别财务比率分布对估计未来预测的不确定性是非常重要的。然而,我国至今没有对财务比率截面数据分布的相关研究,没有人对这种具有基础性意义的研究进行尝试。本文拟通过对沪深股市2000年至2006年全部上市公司的财务比率截面数据分布进行统计假设检验,探讨沪深上市公司财务比率截面数据的分布特征。为了获得数据的正态分布或近似正态分布,我们将进行数据转换和异常值剔除,同时分析截面数据分布的时间序列特征,以研究我国上市公司财务比率截面数据分布的时间变动性。

一、文献综述

Stone 和Rasp (1991) [1]、Stuart Mcleay (2000) [2]分别在各自的研究中发现,当财务比率截面数据为非正态分布、特别是极端非正态时,Logistic 回归的危机预测准确性明显低于正态比率预测的准确性,而在经过偏度或者峰度降低的转换之后,模型的预测能力得以提升。Frecka 和 Hopwood (1983) [3]研究了1950~1979 年COMPUSTAT 数据库中制造业的11 种财务比率截面数据分布,发现异常值数据对分布拟合有显著的影响,通过异常值剔除,大多数比率显示出正态或近似正态分布。Mike Cudd 和Dudda(2000) [4]在预测兼并公司的6 个收购假设时发现,如果不对财务比率的行业属性做出调整,6 个收购假设只有1 个成立;如果考虑财务比率的分布特征,则有4 个假设得到证实。可见,财务比率的分布特征对研究结论具有显著的影响。

国内也有许多学者运用财务比率进行财务危机预测。吕长江、徐丽莉、周琳(2004) [5]在重新界定财务困境和财务破产概念及其差异的基础上,以我国上市公司三年的财务数据为样本,对公司财务困境和财务破产进行了实证分析。刘红霞(2004) [6]运用主成分分析法构建了企业财务危机预警模型,并得到较好的预测效果。方洪全(2004) [7]运用多元判别法对企业信用风险进行了评估,其所建立的信用风险评估模型对国有商业银行的贷款评价有较强的实用价值。还有一些学者运用多元统计方法进行了诸如收购兼并和信用评级研究。多元统计分析中的许多分析方法,如主成分分析、判别分析等,都是建立在变量正态分布或多元正态分布的基础上,而所有这些研究都没有预先进行财务比率的分布检验。

其实,财务危机公司和正常公司财务比率的协方差阵是不可能完全相同的,因为两类公司各项财务指标的波动情况不会完全一致。与其他金融数据一样,企业的财务比率并不服从正态分布,而是具有明显的高峰厚尾分布特征。2004~2006年沪深上市公司各年的流动比率分布以及2000~2006 年沪深全部流动比率最佳拟合分布,就很好地验证了这一特点。可以看出,不论是2004~2006 年各年的截面数据,还是2000~2006 年沪深全部的截面数据,流动性比率分布明显是有偏的,是非正态分布,其最佳拟合分布为非正态的对数逻辑分布。

从理论上说,只要有一个财务指标不服从正态分布,那么,各种财务指标服从联合正态分布的假定就不能成立。然而,国内许多学者在运用判别分析法时,都直接认定数据服从正态分布,且两类总体协方差阵相等。显然,这样得出的结论可能是有偏误的。因此,对财务比率截面数据的分布特征进行研究,具有基础性意义。

二、数据和研究方法

(一) 数据时期及比率选取

本文涉及到的数据来源于CCER 数据库,包括2000 年至2006 年全部沪深上市公司年度财务比率数据,见表1。

表1 各会计年度的公司数

会计年度 公司数

2000 年 1 054

2001 年 1 136

2002 年 1 192

2003 年 1 255

2004 年 1 377

2005 年 1 381

2006 年 1 434

本文对财务比率的选取原则有三条。(1) 肯定了以前研究所采用比率的合理性和科学性。例如, Altman 的Z 值模型中所选用的5 个财务比率就是通过对22 个财务比率进行大量的测算和实验筛选出来的。(2) 反映企业盈利能力和偿债能力原则。企业的盈利是其偿还债务的重要资金来源, 企业的盈利能力越强, 偿还到期债务就越有保障, 发生财务危机的可能性也就越小。偿债能力是指企业偿还到期债务的能力, 是衡量企业财务实力的重要指标, 它与企业财务危机的发生息息相关。在财务比率应用的主要领域, 如危机预测和债券评级, 应首先考虑这两个比率。(3) 反映企业营运能力的原则。企业的资产组合越合理、资产周转速度越快, 就说明企业营运能力越强、资产运用效率越高, 其收入和盈利越多、盈利能力越强, 企业就越是足够的资金偿还到期债务, 发生财务危机的可能性也就越小。

我们借鉴王今等人(2003) [8]的研究成果, 选取了反映盈利能力、偿债能力、营运能力和资本结构的8 种财务比率: 盈利能力选取主营业务利润率和净资产收益率, 偿债能力选取流动比率、速动比率和现金负债比率, 营运能力选取应收账款周转率和资产周转率, 流动性比率选取流动比率和速动比率, 资本结构选取债务资本比率。

(二) 研究方法

本文选取最大似然估计方法进行分布函数的参数估计。最大似然估计方法是通过在给定数据集合前提下使数据的联合分布概率最大化, 估计概率分布函数的参数, 其公式如下:

$$l = \prod_{i=1}^n f(x_i, \alpha) \quad (1)$$
其中, $f(x_i, \alpha)$ 为相应的概率密度函数, x_i 为已知数据集, α 为待求的概率分布函数参数向量。通过使(1) 式最大化可求得参数 α 值, 即:

$$dl/d\alpha = 0 \quad (2)$$

本文主要根据实际财务比率拟合特定分布函数的卡方检验值判断拟合优度, 同时还考虑分布拟合的K - S 检验值和A - D 检验值, 以判断拟合的优劣程度。

K- S 检验可定义为考虑变量所有可能值的实际分布与理论分布之间距离绝对值的最大值, 即:

$$D_{ks} = \max | f_{Emp}(x) - f_{Theo}(x) | \quad (3)$$

其中, $f_{Emp}(x)$ 为实际累积分布函数, $f_{Theo}(x)$ 为特定理论分布函数。

为了获得具有正态分布特征的数据,本文考虑对原始数据进行转换。转换方法包括自然对数转换、平方根转换以及倒数转换,并对原始数据进行异常值剔除。

三、财务比率截面数据分布拟合实证分析

(一) 原始截面比率数据的分布特性

表2 是原始截面比率数据的分布特性,它描述了2000年至2006年比率截面数据分布的矩特征。表中最左边一列指标中的“年度最高值”,是指其对应的行中数据为2000年至2006年各年相应财务比率数据矩的7个年份中的最大值,或者各年截面数据拟合正态分布 χ^2 检验值的7个年份中的最大值。最左边一列指标中的“年度最低值”,是指其对应行中数据为相应7个年份数据中的最小值。“全部数据”对应行中的数据为将不同比率7年数据总合在一起计算的相应指标值。

表2 2000~2006年财务比率截面数据分布特性

净资产收益率 主营业务利润率 流动比率速动比率 现金负债比率 应收账款周转率

资产周转率 债务资本比率

平均值

全部数据- 0.012 0.225 8 1.848 6 1.651 1 0.510 2 56.231 0.532 1 1.594 4

年度最高值0.099 0.282 1 2.158 2 1.859 3 0.601 5 98.932 0.586 8 3.433 8

年度最低值- 0.275 0.079 0 1.661 3 1.458 1 0.396 0 13.649 0.501 5 0.856 1

最高/最低- 0.360 3.570 8 1.299 1 1.275 1 1.518 9 7.248 1.170 0 4.010 9

方差

全部数据0.087 3 0.057 0 3.165 2 0.985 4 0.123 6 3 584 0.196 3 0.774 38

年度最高值0.101 9 0.135 5 4.760 4 1.251 6 0.273 5 74 445 0.252 8 1.227 0

年度最低值0.060 2 0.032 5 1.385 0 0.771 7 0.001 9 752.3 0.128 1 0.224 8

最高/最低1.692 6 4.169 21 3.437 1 1.621 8 143.94 98.956 1.973 4 5.458 1

偏度

全部数据- 64.72 - 8.567 8 9.150 9 9.720 4 9.262 9 28.236 4.111 2 59.601

年度最高值11.627 2.188 4 11.754 2.183 0 1.055 9 34.888 3.605 7 48.741

年度最低值0.124 5 - 9.782 2.198 9 1.112 5 0.623 5 13.729 2.014 1 0.587 5

峰度

全部数据506. 62 189. 79 149. 31 166. 10 159. 61 693. 71 22. 225 414. 10

年度最高值962 181. 8 208. 10 208. 55 217 1 224. 3 28. 593 732

年度最低值61. 301 9. 814 7 10. 297 11. 982 13. 135 219. 6 8. 944 6 3. 984 5

卡方检验值

全部数据17 934 2 566 7 076 796 13 060 28 325 307 12 670

年度最高值27 600 955. 4 1 529 1 529 2 352 36 923 578. 3 27 479

年度最低值106 142. 2 391 490 1 178 22 773 264. 4 189. 6

各种财务比率不同年度的平均值相对差异较大,差异最大的是应收账款周转率,最大年度平均值与最小年度平均值之比为7. 248 2 ;差异最小的是净资产收益率,最大年度平均值与最小年度平均值之比为- 0. 360。

方差表示各种财务比率年度内各自的波动程度,各种比率在不同年度的方差差异较大,现金负债比率和应收账款周转率相对差异高达143. 94、98. 956 ,远高于其他比率的相对差异;差异最小的是净资产收益率和速动比率,分别为1. 692 6 和1. 621 8。

我们对表2 中的偏度和峰度进行相应的u 检验后发现,所有年度财务比率截面数据的| u | 检验值均大于2. 58。可以认为,在99 %的置信水平下,各年度各种财务比率的截面数据为有偏且不是正态峰,各种财务比率截面数据不服从正态分布。表2 中各种比率的 χ^2 检验值不论从最高、最低值还是从平均值来看,其都大于99. 9 %水平下的 χ^2 检验临界值,即在99. 9 %置信水平下可以拒绝财务比率服从正态分布的原假设,可见,财务比率截面数据不服从正态分布。

财务比率截面数据不服从正态分布可能是由以下几个原因造成的:(1) 财务比率是以商的形式表示的两个或两个以上财务变量的结合,即使每个财务变量都服从正态分布且相互独立,它们的商也很难服从正态分布;(2) 出于外部对企业管理的压力,公司为保住配股资格可能采取财务操纵手法,使财务比率保持在可接受的水平,从而很难保证财务比率的正态分布;(3) 一些财务比率的值域范围不满足 $(-\infty, +\infty)$,可能存在为0 的最低边界,导致有偏的正态分布;(4) 财务比率的截面数据来自于不同企业,由于行业、经营环境的原因,数据在总体上缺乏一致性,从而导致比率的非正态分布。

(二) 最佳拟合分布及其时间变化

国外的研究基本上局限于对财务比率为非正态分布的证明,个别学者曾用伽玛分布或一般帕累托分布进行拟合,而国内尚没有人研究过财务比率截面数据分布。本文认为,为了更深入地探讨比率截面数据分布特征,需要运用多种分布函数进行分布拟合,并从中确定最佳拟合分布。由于危机预测和债券评级等都涉及到动态分析研究,因此,还应当从动态角度探讨比率截面数据分布变化,对比率的面板数据(panel data) 进行分析,这也是本文的一个创新。

表3 最佳拟合分布及其时间变化

净资产收益率 主营业务利润率 流动比率速动比率 现金负债比率 应收账款周转率 资产周转率
 债务资本比率

2000年 贝塔分布 正态分布 对数逻辑 对数逻辑 对数正态分布 PT4分布 对数逻辑 PT4分布

2001年 正态分布逻辑斯谛克分布对数逻辑对数逻辑对数正态分布对数正态分布对数逻辑对数逻辑

2002年逆高斯分布 对数逻辑 对数逻辑 对数逻辑 PT4分布 对数正态分布 对数逻辑 PT4分布

2003年 正态分布 逻辑斯谛克分布 对数逻辑 对数逻辑 PT4分布 对数逻辑 PT4分布 PT4分布

2004年 正态分布 逻辑斯谛克分布 对数逻辑 对数逻辑 PT4分布 PT4分布 对数正态分布对数逻辑

2005 年对数逻辑正态分布对数逻辑对数逻辑PT4 分布对数正态分布PT4 分布对数逻辑

2006 年正态分布对数逻辑对数逻辑对数逻辑PT4 分布对数正态分布对数逻辑对数逻辑

为了分析比率的截面数据分布,本文选取了23种分布函数,使用@risk 软件包对各种比率截面数据进行分布拟合。这23 种分布函数分别是:BetaGeneral 分布、卡方分布、误差函数分布、Erlang 分布、指数分布、极值分布、伽玛分布、几何分布、超几何分布、逆高斯分布、逻辑斯谛克分布、对数逻辑分布、对数正态分布、负二项分布、正态分布、帕累托分布、皮尔逊第四型分布(Pearson Type VI Distribution ,PT4 分布)、普阿松分布、瑞利分布、T 分布、三角分布、均匀分布和韦布尔分布(实际拟合结果见表3)。

根据 χ^2 检验值、K- S 检验值和A - D 检验值可以确定最佳拟合分布,结果见表3。总体来看,在56 组截面数据中,仅有6 组数据的最佳拟合分布为正态分布,占10. 7 %;有12 组数据服从PT4 分布,占21. 5 %;有26 组数据服从对数逻辑(loglogistic) 分布,占46. 4 %。56 组截面数据中有近一半的数据服从对数逻辑(loglogistic) 分布,有近70 %的数据服从对数逻辑(loglogistic) 分布和PT4 分布,这是本文的一个发现。

财务比率截面数据分布特征在不同的年度呈现出明显的差异,其中,流动性比率(流动比率和速动比率) 在各年度的截面数据均服从对数逻辑分布,显示出完全的一致性。最佳拟合分布为正态分布的全部6 种比率都集中在反映盈利能力(净资产收益率、主营业务利润率) 的14 组比率中,占43 % ,明显有别于其他比率。现金负债比率的7 组比率从1999 年起连续5 年为PT4 分布,显示出PT4 分布特性。其他比率,如债务资本比率、应收账款周转率以及资产周转率等,其分布在不同年份间呈现出较大差异。

总之,从不同年份的比率截面数据记忆特性来看,流动性比率表现出完全的记忆性,其各年份的分布均服从对数逻辑分布。现金负债比率也表现出较强的记忆性,1997、1998 年为对数正态分布,其

他年度均为PT4 分布。其他比率没有表现出明显的记忆特性。

(三) 财务比率截面数据转换与拟合正态分布

基于正态分布的优良特性, 研究者希望通过对原始数据进行变换, 或者通过对原始数据异常值的剔除, 使转换的数据服从或渐进服从正态分布。本文以2003 年沪深所有上市公司的8 种财务比率截面数据为代表, 对8 种财务比率原始数据进行自然对数转换 $\ln(x)$ 、倒数转换 $1/x$ 、平方根转换 \sqrt{x} 和异常值剔除, 然后对转换的数据进行正态分布拟合检验, 以观察转换的数据是否服从正态分布或接近正态分布。本文进行异常值剔除的方法是, 如果某一原始数据高出该数据组的一个标准差, 就将其剔除出数据组。

表4 2003 年财务比率截面数据转换后拟合正态分布的 χ^2 检验值

净资产收益率 主营业务利润率 流动比率 速动比率 现金负债比率 应收账款周转率 资产周转率
债务资本比率

X 27 605 2 693 1 529 496. 1 2 068 20 947 536. 8 6 907

$\ln(x)$ 226 65. 71 145. 6 104 61. 99 188. 5 121. 0 65. 05

Sqr (x) 1 187 68. 47 367. 7 163. 4 111. 3 278. 1 119. 5 357. 8

$1/x$ 3 608 36. 37 4 551 781. 4 4 000 2 664 544. 8 2 919

剔除异常值后870 97. 39 161 161 458. 1 7 046 166. 4 499

表4 是2003 年的财务比率截面数据转换后的正态分布拟合检验结果。为了进行对比, 我们同时列出2003 年8 种财务比率原始截面数据正态分布拟合检验值, 表中的数据为 χ^2 检验值。

首先, χ^2 检验值表明, 在99. 9 %的置信水平下拒绝比率分布为正态分布的原假设。即使经过数据转换和异常值剔除, 处理后的数据仍不服从正态分布。

其次, 除了倒数转换外, 其他三种数据处理都能使转换后的数据分布比原始数据更接近于正态分布。其中, 又以自然对数转换后的数据最接近于正态分布, 特别是债务资本比率和现金负债比率, 经过自然对数转换后的截面数据正态分布拟合 χ^2 检验值分别为65. 05 和61. 99 , 非常接近于正态分布。

再次, 剔除异常值后, 数据拟合正态分布 χ^2 检验值与平方根转换后的数据拟合正态分布 χ^2 检验值之间的差异不大。

最后, 倒数转化后的数据拟合正态分布效果最差, 7 个比率中有5 个比率的 χ^2 检验值高于原始数据 χ^2 检验值。所有经过倒数转化后的数据拟合正态分布 χ^2 检验值均显著高于其他三种数据处理方式的 χ^2 检验值。

表5 是以2003 年各种财务比率截面数据为例, 探讨数据经过自然对数转换和平方根转换后的最

佳拟合分布,以综合考虑实际财务比率拟合特定分布函数的卡方 χ^2 检验值、K-S 检验值和A - D 检验值来判断拟合的优劣程度。

表5 数据转换后最佳拟合分布

净资产收益率 主营业务利润率 流动比率 速动比率 现金负债比率 应收账款周转率 资产周转率
债务资本比率

X 正态分布 对数逻辑 对数逻辑 对数逻辑 PT4分布 对数正态分布 对数逻辑 对数逻辑

Ln(x) 对数逻辑 对数逻辑 对数逻辑 对数逻辑 对数逻辑 对数逻辑 对数逻辑 对数逻辑

Sqr (x) 极值分布 对数逻辑 对数逻辑 对数逻辑 对数逻辑 PT4 对数逻辑 极值分布

剔除异常值 对数逻辑 对数逻辑 对数逻辑 对数逻辑 逆高斯分布对数逻辑BetaGeneral 对数逻辑

经过以上三种数据转换之后,没有一个比率的最佳拟合分布为正态分布。数据转换以前的8 个(2003 年)原始比率中有3 个比率不服从对数逻辑分布,经过自然对数转换后,所有比率数据均服从对数逻辑分布。经过平方根转换后,仍然有3 个数据不服从对数逻辑分布。原始比率不服从对数逻辑分布的净资产收益率经过平方根转换后,仍然不服从对数逻辑分布。原始比率服从对数逻辑分布的债务资本比率经过平方根转换后,服从极值分布。原始比率剔除异常值后,有2 个数据不服从对数逻辑分布。总之,经过数据转换后,比率截面数据具有更明显的对数逻辑分布特性。

四、结论

通过对沪深全部上市公司财务比率截面数据分布进行拟合分析,我们发现,绝大部分财务比率截面数据的最佳拟合分布为非正态分布。通过数据转换或异常值剔除,能提高数据的正态特性,这一研究结论对于国内的相关研究具有重要的启示意义。

其一,进行企业危机预测、建立相应的危机预测模型时,要充分考虑模型假设条件和模型对各种变量分布属性的要求,分析危机公司(如ST 公司)财务比率截面数据与正常公司财务比率截面数据的分布特征差异。即使应用Logistic 进行危机预测,截面数据的分布特性对预测结果仍然有明显影响。

其二,在对数据进行处理时,应当考虑对数据进行分布检验,并通过适当的数据转换或异常值剔除,取得更真实、满意的预测效果。本研究的56 组数据中有近一半比率服从对数逻辑分布,经过数据转换后,比率的逻辑分布特性更为显著。因此,在应用财务比率进行相关分析时,不仅要考虑比率数据的非正态分布,而且应当考虑具体的分布形式,使预测或评价研究建立在更为科学、准确的基础上。

其三,在后续研究中,可以进一步探讨财务比率截面数据的行业分布特征。本研究是将沪深上市公司的财务比率结合在一起,但不同行业财务比率截面数据的分布有不同的特点,所以,应增加财务比率数目,利用适当的模型研究比率的时间变动性。

[参考文献]

[1] Stone M,Rasp J . Tradeoffs in the Choice between Logit and OLSfor Accounting Choice Studies[J] . The Accounting Review ,1991 ,7 :170 - 187.

[2] Stuart Mcleay. The Sensitivity of Prediction Models to the Non - normality of Bounded and Unbounded Financial Ratios[J] . BritishAccounting Review ,2000 ,32 :213 - 230.

[3] Frecka TJ ,Hopwood W S. The Effects of Outliers on the Cross - Sectional Distributional Properties of Financial Ratios[J] .Accounting Review ,1983 ,58 (1) .

[4] Mike Cudd ,Rakesh Duggal. Industry Distributional Characteristics of Financial Ratios :An Acquisition Theory Application[J] . The Financial Review ,2000 ,35 (1) .

[5] 吕长江,徐丽莉. 上市公司财务困境与财务破产的比较分析[J] . 经济研究,2004 , (8) .

[6] 刘红霞. 以主成分分析法构建企业财务危机预警模型[J] . 中央财经大学学报,2004 , (4) .

[7] 方洪全. 运用多元判别法评估企业信用风险的实例[J] . 预测,2004 , (4) .

[8] 王 今. 财务危机预警中财务比率的选择研究[J] . 数学的实践与认识,2003 , (8) .

文章来源：山西财经大学学报 (责任编辑： x1)