运用非参数 Dirichlet 生存模型进行财务困境企业预警

易莹莹(博士)

(南京邮电大学经济与管理学院 南京 210046)

【摘要】企业陷入财务困境往往是一个动态持续的过程,且具有经常性特点。本文从生存时间的研究视角,采用生存模型对我国上市制造公司中财务困境企业的生存时间进行预警,并针对模型中残差分布不易确定的问题,提出用非参数Dirichlet过程放松分布假设,并作为其的先验分布,利用贝叶斯估计法和MCMC算法进行模型参数和生存曲线的后验推断,以获取财务困境企业生存时间的稳健估计值和更为精确的生存概率预警分析值。

【关键词】生存分析 加速失效时间模型 区间删失 Dirichlet过程 MCMC算法

一、关于生存回归模型的运用

企业陷入财务困境往往是一个动态持续的过程,且具有经常性的特点,一旦公司发生财务困境,会给管理者、债权人、投资者等等造成重大的影响。如果这种状况持续下去,就会对公司生存状况造成极大威胁,甚至导致企业破产。因此,对陷入财务困境的企业进行预警研究具有重要的理论意义和应用价值。

虽自20世纪30年代开始,就有学者提出对企业财务困境预测的分析方法,但传统研究中大多是把它当作影响公司管理者、债权人、投资者决策的要素,很少涉及企业的生存时间与生存率问题。而且以往研究中多采用判别分析模型和离散选择模型,这些方法往往需要配对样本分析,而配对样本的选择不当会严重影响研究与分析的结果。

笔者认为在当今全球经济环境快速变化的情况下,需要以生存时间为新的研究视角,采用生存分析模型探讨公司财务困境和企业生存的相关问题。

生存分析早期主要用于生物和医学,而后逐渐用于社会学、市场学、保险精算学、经济金融等领域。与传统模型相比,生存分析模型不需要配对样本,避免了样本选择的难题。国外研究表明,生存分析不仅在理论上能适用于对组织的生存困境和财务风险领域的分析,在实证研究中也发现生存模型的分析效果要好于传统的静态模型。比如,William(1985)运用比例风险模型对银行经营失败问题进行研究,发现其实证结果优于传统的判别分析模型。Lee和Urrutia(1996)采用基于威尔布分布的生存模型对1980~1991年间82家保险公司的比较与分析结果说明,利用生存模型预测保险公司破产的准确性更高。Shumway(2001)从数学上证明了生存分析的风险模型避免了选择偏差问题,比Logit等静态模型更具优越性。Saretto(2004)采用带时序协变量对生存时间分布不加任何限定的最大似然估计法,估计信用风险模型提高了预测的准确度。

而国内近几年才开始利用生存模型对企业信用风险主要影响因素进行研究与分析。宋雪枫(2006)与陈磊、任若恩(2007)等应用比例风险模型分析了信贷企业发生财务危机的概率。李文娟(2008)通过生存分析的寿命表方法以及Cox比例风险模型研究特别处理企业摘帽与退市的影响因素。

生存回归模型通常包括加速失效时间模型和比例风险模型,为更有效地对陷入财务困境企业生存时间进行预警,我们采用加速失效时间的生存模型进行研究,但由于它需要事先假设基准生存函数的具体分布形式,如果这些假设不成立,会使不同的模型假设产生截然相反的结论。因此,虽然它较比例风险模型更易解释,但可能会产生一些误差。为解决这些问题,提高预测精度,我们提出运用非参数Dirichlet过程放松对加速失效时间模型中基准生存函数的分布形式设定的限制,并利用贝叶斯估计和MCMC算法进行估计和推断,对我国深沪两市A股上市制造公司中的财务困境公司的生存时间和生存概率做定量分析。

二、加速失效时间模型与非参数贝叶斯估计法

对于第i个观察个体,加速失效时间模型的基本表达式如下:

$$T_i=\exp(-x_i'\beta)v_i,\ i=1,L,n \tag{1}$$

$$v_1,L,v_n|G\overset{iid}{\sim}G$$

其中,$x=(x_1,L,x_n)'$是协变量,β是对应的未知系数,它反映了对数生存时间与协变量之间的关系,也是我们主要推断的对象。

残差项v服从某种独立同分布G(·),其对应的密度函数为f(·),它也表示在没有协变量影响(x=0)情况下的基准生存时间,如果用S_0表示,有:

$$S_0(t)=P(T>t|x=0)=P(v>t) \tag{2}$$

这样可以推导出该模型的生存函数:

$$S(t|x)=P(T>t|x)=P\{\ln(T)>\ln(t)|x\}$$
$$=P\{\ln(T)+x'\beta>\ln(t)+x'\beta|x\}$$

$$=P\{\exp[\ln(T)+x'\beta]>t\exp(x'\beta)|x\}$$
$$=P\{v>t\exp(x'\beta)|x\}$$
$$=S_0(t\exp(x'\beta)) \qquad (3)$$

这里注意到，如果没有协变量的影响，个体的实际生存时间为v，但是，正因为有了协变量的影响，它通过因子 exp(x′β)改变了原来的时间尺度。时间可以以常数因子加速，也可以以常数因子减速，这取决于 β 的符号。因此，因子 exp(x′β)常被称为加速因子，反映了协变量取值的变化对生存时间的影响，我们称此模型为加速失效时间模型。

由于实际分析中对残差项 v 进行的参数假设太过于严格，常常很难正确地判别其服从的参数分布。因此，我们讨论其为非参数分布 Dirichlet 过程 $DP(\alpha, G_0)$的情形。其中，G_0 代表对残差项 v 所服从分布的先验信念，α 表示残差项 v 所服从的分布集中于 G_0 的集中度。参数 β 解释了对数生存时间与协变量之间的关系，在模型中令它服从均值为 β_0、协方差为 S_{β_0} 的正态分布。

对于区间删失数据，加速失效时间模型的形式可以描述为：

$$T_i=\exp(-x_i'\beta)v_i, i=1,L,n, T\in[a,b)$$
$$\beta|\beta_0, S_{\beta_0}\sim N(\beta_0, S_{\beta_0})$$
$$v_i|G\sim G$$
$$G|\alpha, G_\theta\sim DP(\alpha, G_\theta), \theta\sim f(\theta) \qquad (4)$$

根据 Hanson 和 Johnson(2004)的研究成果，假设 $\beta\perp\theta$，$\beta\perp v$，$\beta\perp G|\theta$ 以及 $v\perp(\theta,\beta)|G$，且对于所有的观测 i 和所有的 $\theta\in\Theta$ 而言，$G_\theta\{[a_i,b_i)\}>0$。因此，有$[G,\theta|v,\beta,T\in[a,b)]=[\theta|v][G|\theta,v]$，$[\beta,v|G,\theta,T\in[a,b)]=[\beta|G,T\in[a,b)][v|\beta,G,T\in[a,b)]$。

给定 G、β，模型(4)的似然函数设定为：

$$P(T\in[a,b)|\beta,G)=\prod_{i=1}^{n}G(J_i(\beta)) \qquad (5)$$

其中，$J_i(\beta)\overset{def}{=}[a_i\exp(x_i'\beta), b_i\exp(x_i'\beta))$。因此，参数 β 的后验分布为：

$$f(\beta|G,T\in[a,b))\propto f(\beta)\prod_{i=1}^{n}G(J_i(\beta)) \qquad (6)$$

此时，v 的完全条件分布可以描述为：

$$v_i|v^{(i)},\beta,\theta,T\in[a,b)\begin{cases}\sim G_0 & \text{在 } J_i(\beta)\text{截断} \quad Prob=w/ap\\ =v_k & v_k\in J_i(\beta), k\neq i \quad Prob=w/p\end{cases} \qquad (7)$$

其中，$1/p=\alpha G_\theta(J_i(\beta))+\sum_{j\neq i}\delta_{J_i(\beta)}(v_j)$。

β 的完全条件分布为：

$$f(\beta_j|\beta^{(j)},v,T\in[a,b))\propto f(\beta_j|\beta^{(j)})I_{R(v,\beta^{(j)})}(\beta_j) \qquad (8)$$

其中，$R(v,\beta^{(j)})=\bigcap_{i=1}^{n}\{x_{ij}\beta_j\in[\ln(\frac{v_i}{b_i})-x_i^{(j)}\beta^{(j)}, \ln(\frac{v_i}{a_i})-x_i^{(j)}\beta^{(j)}\}$。

θ 的完全条件分布为：

$$[\theta|v]\sim f(\theta|v)\propto f(\theta)\prod_{i=1}^{d(v)}g_\theta(v_i^*) \qquad (9)$$

其中，$\{v_i^*\}_{i=1}^{d(v)}$是 v 的离散值集合，$g_\theta(v)$是 G_θ 关于勒贝格测度(Lebesgue measure)的密度函数。

Hanson & Johnson 认为当删失区间很小时，根据式(8)直接对参数 β 的完全条件分布抽样会容易在迭代过程中产生自相关性，因此，往往需要迭代足够多次数以获取有代表性的样本来推断后验分布。他认为除非删失区间足够宽，否则该直接方法既无效率又不实用。因此提出使用 MCMC 算法中的 Metropolis-Hastings 算法对参数 β 进行抽样，并定义接受概率为：

$$\rho(\beta,\beta^*)=\min\{\frac{f(\beta^*)\prod_{i=1}^{n}G(J_i(\beta^*))}{f(\beta)\prod_{i=1}^{n}G(J_i(\beta))},1\} \qquad (10)$$

其中，$\rho(\beta^*,\beta)=\rho(\beta,\beta^*)$。在马尔可夫链进行第 g 步迭代时，取：

$$\beta^{g+1}=\begin{cases}\beta^* & prob=\rho(\beta^g,\beta^*)\\ \beta^g & prob=1-\rho(\beta^g,\beta^*)\end{cases} \qquad (11)$$

为了对式(10)进行计算，G 必须在区间$[0,\infty)$上的有限划分中加以具体说明，因为它是 Dirichlet 过程，如果划分设定为$\{Ij\}_{j=1}^{M}$，则有：

$$\{G(I_j)\}|V,\theta\sim DP(\{\alpha G_\theta(I_j)+\sum_{i=1}^{n}\delta_{v_i}(I_j)\}) \qquad (12)$$

由于 v_i 处于$J_i(\beta)$中，定义：

$$J_i(\beta)=\bigcup_{j=n_{1,i}}^{n_{2,i}}I_j, i=1,L,n+1 \qquad (13)$$

其中，$J_{n+1}(\beta)=[0,\infty)$。

并且定义随机变量 $S_i|\{G(I_j)\}$为：

$$P(S_i=s)=\begin{cases}\frac{G(I_s)}{\sum_{j=n_{1,i}}^{n_{2,i}}G(I_j)} & s\in\{n_{1,i},L,n_{2,i}\}\\ 0 & s\notin\{n_{1,i},L,n_{2,i}\}\end{cases} \qquad (14)$$

其中，$n_{1,n+1}=1$，$n_{2,n+1}=M$。S_i 表示 v_i 在集合$\{Ij\}_{j=1}^{M}$中所处的区间。于是有：

$$P(v_i\in Ij|G,v_i\in J_i(\beta))=P(v_i\in Ij|\{G(I_k)\},v_i\in J_i(\beta))$$
$$=P(S_i=j) \qquad (15)$$

对生存数据进行分析时，预测其生存曲线是很重要的。式(4)的生存曲线为：

$$\hat{S}(t|x,T\in[a,b))=\frac{a}{a+n}G_\theta(t\exp(x'\beta),\infty)+\frac{1}{a+n}\delta_{v_i}(t\exp(x'\beta),\infty) \qquad (16)$$

Hanson 和 Johnson 指出式(16)可以近似为：

$$\hat{S}(t|x,T\in[a,b))=\int\hat{S}(t|x,\beta,v,\theta)dP(\beta,v,\theta|T\in[a,b))$$
$$=\frac{1}{N}\sum_{g=1}^{N}\{aG_{\theta^g}((t\exp(x'\beta^g),\infty))+\sum_{i=1}^{n}\delta_{v_i^g}((t\exp(x'\beta^g),\infty))\}\times\{a+n\}^{-1} \qquad (17)$$

其中，N 指迭代的总次数。

为了对模型参数进行合理估计，通常利用马尔可夫蒙特卡罗(MCMC)算法计算。根据 Hanson 和 Johnson 的研究成果，令初始值为 $X^0=(\beta^0,V^0,\theta^0)$，第 g 步迭代步骤为：

(1)首先根据式(9)对 $\theta^g|v^{g-1}$ 进行抽样。

(2)从区间端点$\{J_i(\beta^*)\}_{i=1}^{n}$和$\{J_i(\beta^{g-1})\}_{i=1}^{n}$获取划分$\{I_j\}_{j=1}^{M}$,对建议分布$\beta^* \sim q(\beta^* | \beta^{g-1})$进行抽样。

(3)根据式(12)对$[G(I_j)]|v^{g-1}, \theta^g$进行抽样。

(4)根据式(10)计算$\rho(\beta^{g-1}, \beta^*)$。取$U \sim U(0,1)$:如果$U>\rho(\beta^{g-1}, \beta^*)$,令$\beta^g=\beta^{g-1}$,$v^g=v^{g-1}$,以及$v_{n+1}^{g}=v_{n+1}^{g-1}$,并停止进行抽样;如果$U \leqslant \rho(\beta^{g-1}, \beta^*)$,令$\beta^g=\beta^*$,并继续进行步骤(5)。

(5)根据式(14)对S_1, L, S_{n+1}进行抽样。令$m_j=\sum_{i=1}^{n+1}\delta_{jS_i}$,$\varepsilon_j$是包含元素$\{i \in \{1, L, n+1\}: S_i=j\}$的向量,那么$m_j$是$\varepsilon_j$的长度。

(6)为了对生存曲线(17)进行抽样,令$J=\{j: m_j>0\}$,对于每个$j \in J$。

第一步,对$U_1, L, U_{m_j} \overset{iid}{\sim} U(0,1)$进行抽样。

第二步,对$W_1, L, W_{M_j} \overset{iid}{\sim} beta(1, \alpha G_{\theta^g}(I_j)+\sum_{i=1}^{n}\delta v_i^{g-1}(I_j))$进行抽样,并构成变量$W_i^*=W_i[\prod_{s=1}^{i-1}(1-W_s)]$,直至$\sum_{i=1}^{M_j}W_i^*>\max\{U_1, L, U_{m_j}\}$。其中,$M_j$是随机变量。

第三步,在划分I_j中对$v_1^*, L, v_{M_j}^* \sim \frac{\alpha G_{\theta^g}+\sum_{i=1}^{n}\delta v_i^{g-1}}{\alpha G_{\theta^g}(I_j)+\sum_{i=1}^{n}\delta v_i^{g-1}(I_j)}$,进行抽样,然后更新向量v,使得$W_{\varepsilon_{js}}^{\%}=v_k^*$,且对于$s=1, L, m_j$,满足$\sum_{i=1}^{k-1}W_i^* \leqslant U_s < \sum_{i=1}^{k}W_i^*$。

(7)令$v^g=W^{\%}$。

(8)对精度参数α进行抽样。

将上述步骤进行多次迭代,直至马尔可夫链达到收敛。

三、实证研究

1. 变量选择与模型设计。根据我国上市公司的有关监管政策,对财务状况或其他状况出现异常的上市公司的股票交易要进行特别处理,即对公司ST(Special Treatment)。这些异常状况往往会影响投资者对其发展未来难以进行合理的判断,因此,在本文中我们把公司被ST作为其发生信用风险的标志。

上市公司被ST的原因有:①最近两个会计年度净利润均为负值;②最近一个会计年度股东权益低于注册资本;③最近两个会计年度净利润均为负值,且最近一个会计年度每股净资产低于股票面值;④被注册会计师出具无法表示意见或否定意见的审计报告;⑤交易所或中国证监会认定财务状况异常;⑥追溯调整导致最近两年连续亏损;⑦出现其它异常状况,实施特别处理;⑧追溯调整导致最近两年连续亏损,最近一个会计年度股东权益低于注册资本;⑨在法定期限内未依法批露定期报告;⑩最近三年连续亏损。本文主要考虑的ST公司是由于财务状况出现异常原因导致的。

我国自1998年4月开始在沪深两市实施ST政策,因此,本文将样本观测期设定为1998年4月至2009年12月,即把1998年4月后上市的公司作为研究对象(本研究所采用的时间尺度为月)。从生存分析的角度说,就是把公司上市时间作为生存时间的起点,公司被ST作为发生的"失效"事件,以此作为生存时间的终点。

因而本文样本数据由两类组成,一类是公司在观测期结束前被ST,该类ST公司的生存时间是完整的;另一类是公司在观测期结束时仍未被ST,该类非ST公司确切的生存时间是未知的。对于ST公司而言,它从上市到被ST是其生存时间;对于非ST公司而言,它从上市到2009年12月是其生存时间。

参照国内外的研究成果,本文选取反映公司财务状况的主要指标有:盈利能力、偿债能力、营运能力、成长能力、现金流量能力和股票市场价值等方面的20个指标(见表1)。由于各家公司的规模、总资产等都不同,为了排除这些因素的影响,我们选取的都是比率指标。又由于陷入特别处理的ST公司并不是暂时的状况,基本都是经过一个较长过程而形成的,而且具有持续性和经常性的特点,因此,本文在选取影响因素时,对所有的财务指标以历史整体数值计算出的平均结果作为各指标的数据值。对于ST公司,用其上市当期至被特别处理当期的财务数据的平均结果作为计算财务指标;对于非ST公司,用其上市当期至2009年12月31日的财务数据的平均结果作为计算的财务指标值。

我们进行模型分析变量与指标的选择过程是,首先运用非参数Wilcoxon检验(令显著性水平=0.01)寻找在ST公司和非ST公司之间具有显著差异的指标,剔除不能区分ST公司和非ST公司的指标。由于剩下的指标之间可能存在多重共线性,于是利用相关性检验(令显著性水平=0.01)对剩余指标进行再次筛选。最终入选模型的分析指标有:资产负债率、已获利息倍数、应收账款周转率和平均收益率(表1中带*的指标)。本文数据主要来源于天软上市公司财务数据库。

表1　　影响公司生存状况的分析指标

盈利能力	资产报酬率、销售净利率、营运报酬率
偿债能力	流动比率、速动比率、超速动比率、 长期债务/营运资金、资产负债率*、已获利息倍数*
营运能力	应收账款周转率*、存货周转率、总资产周转率
成长能力	营业利润(同比增长率)、营业收入(同比增长率)
现金流量能力	经营活动产生的现金流量净额/营业收入、 经营活动产生的现金流量净额(同比增长率)
股票市场价值	净资产收益率、平均收益率*、 市盈率、市净率

因此,最终建立企业生存时间的预测模型如下:

$$\ln T_i=-(\beta_1x_1+\beta_2x_2+\beta_3x_3+\beta_4x_4)+V_i,$$
$$\ln T_i \in [\ln(a_i), \ln(b_i)] \qquad (18)$$

其中,协变量$x=(x_1, x_2, x_3, x_4)$分别是指影响公司生存时间的资产负债率、已获利息倍数、应收账款周转率和平均收益率因素;a_i为第i个公司的上市时间;对于ST公司,b_i为第i个公司被ST的时间,对于非ST公司,b_i指2009年12月。

2. 数据描述。由于运用生存分析时需要确保样本中企业失败过程的同质性(邓晓岚等,2007),本文选取深沪两市A股上市制造业公司为研究对象。

为提高样本数据质量,我们还剔除了以下样本:①上市不足3个完整的会计年度或上市不足3个会计年度即被ST的公司。由于公司在刚上市的三年内,相应的盈利能力比

较强，几乎不存在被特别处理的风险。因此，本文所研究的ST公司对象至少是在2001年4月后被ST的公司。②数据无法收集完整的公司。这样，最终得到的有效样本数是397家，其中ST公司有40家，非ST公司有357家。

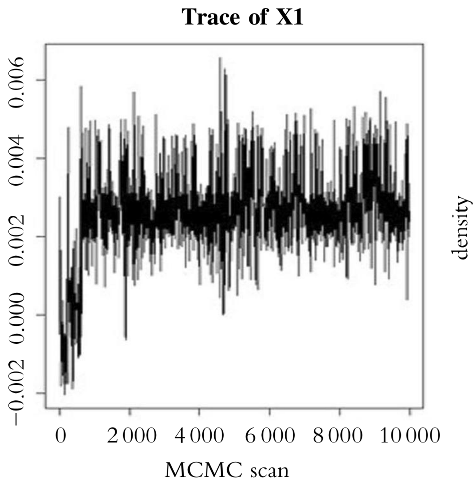
为了预测该模型的结果，我们还进行样本外预测检验。把检验的样本观测期设定为2010年，即在2010年符合条件的ST公司选定为我们要进行检验的样本。这样，最终得到的有效样本数是12家，见表2。

表2　两市制造业公司(A股)被ST的时间分布(2001~2009年；2010年)

年＼月	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
2002年	0	0	0	1	0	0	0	0	0	0	0	0
2003年	0	0	1	3	0	0	0	0	0	0	0	0
2004年	0	0	1	2	0	0	0	0	0	0	0	0
2005年	0	1	0	7	0	0	0	0	0	0	0	0
2006年	0	0	3	2	0	0	0	0	0	0	0	0
2007年	0	0	0	6	6	0	0	0	0	0	0	0
2008年	0	0	0	5	0	0	0	0	0	0	0	0
2009年	0	0	0	2	0	0	0	0	0	0	0	0
总计	0	1	5	28	6	0	0	0	0	0	0	0
2010年	0	1	0	10	1	0	0	0	0	0	0	0
总计	0	1	0	10	1	0	0	0	0	0	0	0

3. 实证与结果分析。对模型进行计算时，我们首先设置了平行运行2条链，每条链先退火(burn in)20 000次，当这2条链达到收敛后，再迭代10 000次，然后根据最后的10 000个后验抽取进行分析。另外，为了克服因为连续抽取而导致的自相关性，我们设计thin=100，即所得的随机数每隔100次抽取用于后验量的计算。

图1为模型的轨迹图，其中X_1、X_2、X_3和X_4分别代表资产负债率、已获利息倍数、应收账款周转率和平均收益率。从图1可以看出，各个参数的抽取值都达到了稳定状态。并且从表3的左半部分也可以看出，各个参数的G-R统计量接近于1。因此，可以认为各个参数在迭代过程中已经达到收敛。



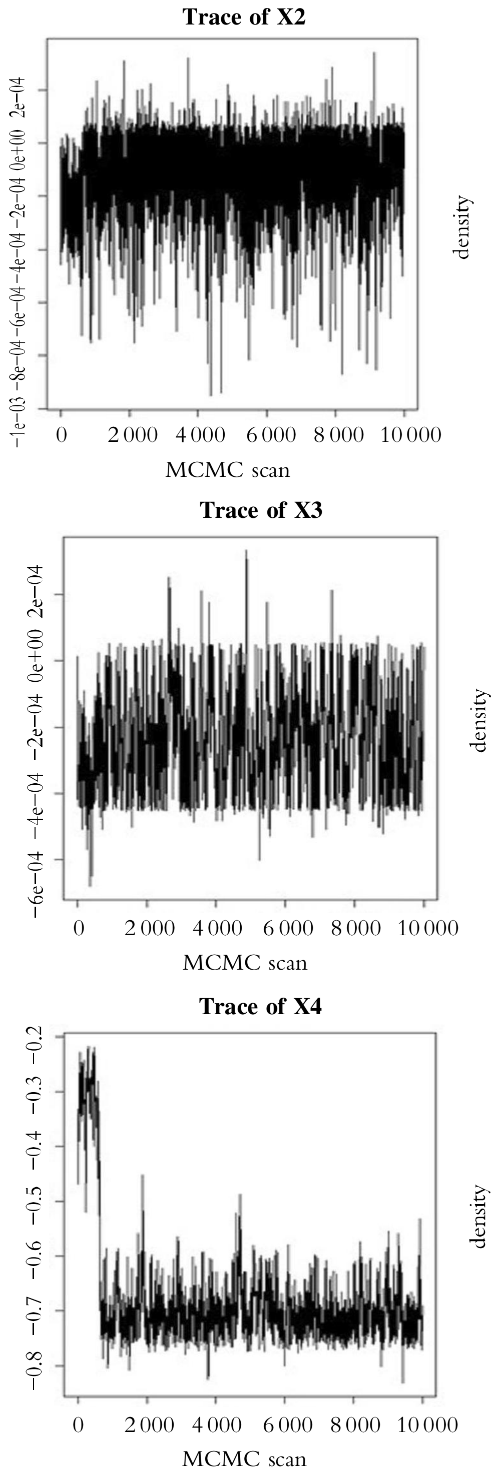


图1　各个参数的MCMC估计

表3右半部分为模型中资产负债率、已获利息倍数、应收账款周转率、平均收益率的后验预测结果。可以看出，参数估计系数分别为0.002 7、-0.000 1、-0.000 2、-0.679 6。说明资产负债率较低，已获利息倍数、应收账款周转率和平均收益率较高的股票，财务风险越低，其公司的生存时间长，投资价值高；相反，资产负债率较高，已获利息倍数、应收账款周转率和平均收益率较低的股票，其公司的生存时间短，投资价值低。

表 3　各参数的 G-R 统计量和后验估计值

	G-R统计量		后验估计	
	点估计	97.5%估计	均值	标准差
资产负债率	1.03	1.11	0.002 7	0.001 1
已获利息倍数	1.01	1.04	-0.000 1	0.000 2
应收账款周转率	1.00	1.01	-0.000 2	0.000 1
平均收益率	1.02	1.06	-0.679 6	0.104 6
基准分布:均值	1.00	1.00	0.494 3	0.780 9
方差	1.00	1.00	70.129 4	83.283 8
精度参数:α	1.00	1.00	1.000 0	1.748 0

为了对该模型进行验证，我们把 2010 年 12 家 ST 公司的相关数据代入模型。出于文章篇幅考虑，这里只给出代码为 600562 的公司的验证曲线。根据估计结果，其生存曲线如图 2 所示。

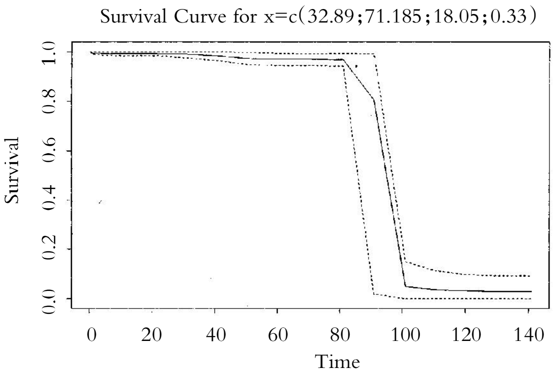


图 2　生存曲线图

图 2 生存曲线图中横轴代表生存时间，纵轴代表对应的生存概率，实线部分描绘了该公司的生存曲线，虚线部分是对应的其生存曲线的区间估计。一般而言，生存曲线越是靠上，说明对于一个给定的时间点，生存时间超过该点的概率越大、生存状况越好。

实证分析结果表明，该公司的实际生存时间为86 个月，从图中我们也可以看出，该公司在 86 个月生存时间的概率仅为 0.4。通过上述分析可知，运用该模型能较好地对我们所研究的上市公司的生存时间进行分析，从而能够帮助投资者作出正确的投资决策。

四、结论

由于我国证券市场发展的历史短，市场处于需要不断完善的过程，而且公司经营过程中面临内部状况发生变动和环境动态变化及竞争压力，使其随时都有可能陷入财务困境。而陷入财务困境通常是一个动态过程，一旦发生而且持续，就会导致企业破产。

因此，建立财务困境企业的监测机制与危机预警，利用公开的财务报表信息预测上市 ST 公司的状态，对影响公司生存时间的因素进行分析具有重要意义，也是股东、投资者以及债权人共同关心的问题。

以往对企业财务困境问题进行研究多采用判别分析模型和离散选择模型，这些方法往往需要配对样本分析，而配对样本的选择不当会严重影响研究与分析的结果。虽然也有学者使用比例风险模型进行研究，但也存在模型假设条件过于苛刻的问题。

为克服这些缺陷，本文利用生存分析中的加速失效时间模型分析上市公司的生存时间，针对模型中残差的分布难以确定的问题，利用非参数 Dirichlet 过程放松其分布假设，并利用贝叶斯估计法和 MCMC 算法以获取参数和生存曲线的稳健估计值。以深沪两市 A 股上市制造公司为研究对象，选取一些能够反映公司财务状况的指标，并利用非参数 Wilcoxon 检验和相关性检验剔除了不具有区分 ST 公司和非 ST 公司的指标，最后，研究发现：对于资产负债较低、已获利息倍数、应收账款周转率和平均收益率较高的股票，其公司的生存时间越长，投资价值越高；反之则相反。

公司陷入财务困境往往是经历了一个动态的演变过程，在这个过程中，存在着大量的预警信息。在我们的研究中可以看出，对于我国的上市 A 股制造公司，资产负债率、已获利息倍数、应收账款周转率和平均收益率是关系其生存时间的关键要素。这些实证分析与研究结果的发现，不仅有利于监管部门对其进行有效的监督和指导，而且能够有效地帮助股东、投资者以及债权人加强风险意识，以便能够作出正确的决策。

【注】本文系南京邮电大学引进人才基金(项目编号：NYS211025)和江苏省教育厅高校哲学社会科学研究项目(项目编号：09SJB790032)的阶段性研究成果。

主要参考文献

1. William，R.L. Stephen，W.L.，James.An application of the Cox model to bank failure. The Financial Review，1985；20

2. Lee，S.H.，Urrutia，J.L. Analysis and prediction of Insolvency in the property-liability insurance industry:a comparison of Logit and hazard models. Journal of Risk and Insurance，1996；6

3. Shumway，T. Forecasting bankruptcy more accurately:a simple hazard rate model. Journal of Business， 2001；7

4. Saretto. Predicting and pricing the probability of default. Working paper， 2004；5

5. 宋雪枫，杨朝军，徐仁重. 商业银行信用风险评估的生存分析模型及实证研究.金融论坛，2006；11

6. 陈磊，任若恩.基于比例危险和主成分模型的公司财务困境预测.财经问题研究，2007；7

7. 李文娟.中国 ST 企业的生存分析.厦门大学硕士论文，2008

8. Hanson，T.， Johnson，W.O.. A Bayesian semiparametric AFT model for interval-censored data. Journal of Computational and Graphical Statistics，2004；13

9. 邓晓岚，陈朝晖，王宗军.公司财务困境非参数生存分析模型评价.武汉理工大学学报，2007；6

10. Knox，K.J. Blankmeyer，E.C. Trinidad，J.A.，Stutzman，J.R.. Predicting bankruptcy in the Texas nursing facility industry. The Quarterly Review of Economics and Finance，2009；3