

基于完全信息的测算信用等级 违约概率的新方法^①

张金宝

(北京第二外国语学院国际商学院)

研究目标: 为充分利用历史数据的信息测算违约概率提供一种新方法。**研究方法:** 从违约强度与违约概率的关系出发, 提出计算违约概率的新思想。以银行贷款为例, 在分析贷款的起止时间与违约概率考察期间的时序关系之后, 通过创新性地构造特征函数的方式, 描述了考察期间贷款数目的时变特征, 提出了利用历史数据测算违约概率的公式。**研究发现:** 传统的测算方法对违约概率存在一定程度的低估。**研究创新:** 采用该方法测算违约概率时, 考察期间的任何一笔贷款都是潜在的测算样本的数据源。相比于传统方法中“只选择在考察期初存在的贷款作为测算样本”的做法, 本文的方法利用历史违约数据的信息更完全。**研究价值:** 为基于历史数据测算违约概率提供了新思路, 研究成果可广泛用于银行贷款的违约概率的测算。

关键词 违约概率 违约强度 信用评级

中图分类号 F830.5 **文献标识码** A

一、问题的提出

违约概率是指债务人在特定的期限内不能履行债务合同义务的可能性大小。它是信用风险管理的重要参数之一。在商业银行内部, 违约概率和违约损失率可以一起为贷款定价提供重要依据。在风险管理方面, 违约概率可以辅助银行进行内部评级、测算贷款的预期损失、确定合理的损失准备金的计提额度, 还可以帮助银行模拟贷款组合的损失分布, 确定需要配置的经济资本, 计算银行需要缴纳的存款保险费等。此外, 违约概率还是定价与银行信用资产相关的金融衍生品的重要变量。

在债券市场上, 违约概率也是信用评级机构关注的重要参数。自 2007 年金融危机爆发以来, 穆迪、标普、惠誉三大评级机构的评级质量受到各方质疑, 被认为存在“评级高估”现象。2016 年我国债券市场违约事件频发。据 Wind 资讯统计, 2016 年共有 79 只债券发生违约。在一系列的违约事件中, 国内评级机构的市场反应令人关注。多数违约债项是在债务人宣布违约的当天甚至是随后的几天, 评级机构才开始下调债项的级别。在此之前, 这些濒临违约的债项多数维持在 AA 到 AAA 的级别。因此, 市场普遍认为评级机构高估了违约债项的级别。一些理论研究甚至认为, 评级公司从自身利益出发, 存在故意夸大评级的现象

^① 本文获得国家社会科学基金“存款保险定价的理论建模与实证研究”(14BJY199) 的资助。作者感谢审稿人的审稿意见, 当然文责自负。

(Mathis 等, 2009; Bolton 等, 2012; 夏凡和姚志勇, 2013)。为回应这些质疑, 评级机构最常见的做法就是公布不同的信用级别所对应的违约概率, 以彰显自身评级的公信力。这表明, 无论是外部评级还是银行内部评级, 准确测算信用级别的违约概率都是非常必要的。

为了测算违约概率, 业界和学术界推出了各种违约风险模型(管七海和冯宗宪, 2004)。比较研究表明, 不同模型针对同一笔债项测算的违约概率结果不尽相同(Annio 等, 2012; Crouhy 等, 2000)。近年来, 仍有关于违约概率的模型被不断提出(Bellalah 等, 2016; Li 等, 2015; Kaiashcova, 2014; Silva 和 Murtcira, 2009)。这种现象也说明, 已有的违约概率模型仍远未达到理想要求。原因在于, 影响债务人违约的因素非常复杂, 这些因素可能来自宏观层面、中观层面(包括行业和地区)以及微观层面, 既包括可以量化的因素, 也包括无法量化很难被引入到定量模型中加以考虑的因素。所以评级机构公布的违约概率多数是根据历史违约数据直接测算的违约概率, 这是因为历史违约数据是债项违约风险最直观的反映。从某种意义上说, 它不存在模型误差, 测算的违约概率可靠性更高。

基于历史数据测算违约概率的方法看似简单。在特定的考察期间内, 测算某一信用级别的违约概率, 似乎只要用发生违约的债项数目除以该级别的债项总数即可, 然而深究起来却没这么简单。以银行贷款业务为例, 在考察期内的任一时点, 银行既可能有新的贷款产生, 也可能有旧的贷款被清偿, 所以考察期间的贷款数目并不是一成不变的。那么在测算违约概率时究竟该怎样确定考察期间的债项总数才更合理? 这些关键的细节问题, 已有的研究文献尚缺乏深入的分析。为了便于读者理解这个问题, 我们借助于一个简单的例子: 表1模拟的是某银行2016年贷款及违约的抽样数据。不失一般性, 假定贷款的信用等级为BBB。为简单起见, 我们假定表1中的BBB级贷款全部是由银行签订贷款协议发放贷款的行为引起, 暂不考虑由于信用等级迁移而形成的贷款。表1给出了贷款的编号、贷款的发放日和到期日以及违约的情况。在表1中, 若某笔贷款违约用“1”来表示, 若没有违约则用“0”来表示。

表1 BBB等级贷款违约统计

编号	发放日	到期日	违约	编号	发放日	到期日	违约
1	2014.07.01	2017.01.01	0	16	2015.10.28	2017.04.28	0
2	2014.07.10	2017.01.10	0	17	2015.10.30	2017.10.30	0
3	2015.04.01	2017.04.01	0	18	2015.10.31	2017.05.01	0
4	2015.04.10	2017.04.10	1	19	2015.11.01	2017.11.01	0
5	2015.07.21	2017.01.21	0	20	2015.11.04	2017.05.04	0
6	2015.07.22	2017.07.22	0	21	2015.11.06	2017.11.06	0
7	2015.07.28	2017.01.28	0	22	2015.11.09	2017.05.09	0
8	2015.08.01	2017.08.01	0	23	2015.11.13	2017.05.13	0
9	2015.08.08	2017.02.08	0	24	2015.11.15	2017.01.15	0
10	2015.08.12	2017.02.12	0	25	2015.11.22	2018.06.21	0
11	2015.08.24	2017.02.24	0	26	2015.12.03	2017.06.03	0
12	2015.08.31	2017.02.28	0	27	2015.12.07	2017.12.07	0
13	2015.09.10	2017.03.10	0	28	2016.01.01	2017.07.01	0
14	2015.10.19	2017.04.19	0	29	2016.01.11	2017.07.11	1
15	2015.10.21	2016.04.21	0	30	2016.02.22	2016.08.22	0

如果将考察期间设定为 2016 年 1 月 1 日～2016 年 12 月 31 日，可以看出在这个期间内既有贷款被清偿（15 号贷款），也有新的贷款产生（29 号、30 号贷款），也就是说考察期间贷款的数目是变化的，那么该如何测算该银行 BBB 级贷款的违约概率？现有的文献中，可供借鉴的方法并不多。标准普尔、穆迪等评级公司的研究报告略有提及（Diane 和 Nick, 2015、2013；Fons 等, 1994），此外彭建刚等（2009）提出了一种类似生命表的测算方法。这些方法通常是选择在考察期的初始时刻已存在的债项作为一个群体，然后观察这个群体在考察期间的违约数目来测算违约概率。换句话说，对于每个债项都是从考察期的初始时刻 $t=0$ 开始观察，而在 $t=0$ 时刻之后产生的债项则被忽略掉了。参照标准普尔评级公司以及穆迪公司的研究报告，本例中违约概率应这样测算：首先统计在 2016 年 1 月 1 日时刻已存在的贷款数，这意味着需要从表中剔除掉 29 号和 30 号这两笔贷款，然后将总计得到的贷款数目 28 作为分母。接着统计 28 笔贷款中发生违约的数目，将其作为分子。两者相除计算出该级别的违约概率。于是，由表 1 的抽样数据得出该银行 BBB 级贷款的违约概率为：

$$\text{违约概率} = \frac{\text{违约贷款的数目}}{2016 \text{ 年初贷款的总数目}} = \frac{1}{28} = 3.57\% \quad (1)$$

这种测算方法实际上是一种静态方法，其本质是从考察期间发生的所有贷款中有选择地抽取只在 2016 年 1 月 1 日存在的贷款作为样本，然后将样本的违约频率作为违约概率。这种取样方法，严格地说是有失偏颇的。

第一，从我们“考察 2016 年 1 月 1 日～2016 年 12 月 31 日整个期间的 BBB 级贷款违约概率”的研究目的看，测算时似乎没有理由舍弃第 29 号、30 号两笔贷款。因为它们确实在考察期内发生的贷款，尽管它们在 2016 年 1 月 1 日之前并不存在。尤其对于 29 号贷款而言，这种发放略迟却发生违约的贷款，恰恰有可能反映该类贷款最新的违约信息，折射出现实经济中违约因素的最新变化，因此忽略它们未必合理。第二，如表 1 中的第 15 号贷款，尽管它在 2016 年 1 月 1 日之前已经存在，但在 2016 年 4 月 21 日就已经清偿完毕，并没有经历整个考察期间，而标普和穆迪的测算方法在计算违约频数时，将其与那些经历整个考察期间的贷款如 5 号、6 号贷款同等对待，记录在测算公式的分母中，也难免让人质疑这样的处理是否妥当。这意味着，标普或穆迪所计算的违约概率并没有充分利用考察期间所有的贷款数据，是在信息不完全的条件下测算的违约概率。

信息的不完全究竟对违约概率的测算会产生什么样的影响？以表 1 中的数据为例，如果考虑 29 号和 30 号贷款，那么测算的违约概率将达到 6.67%，它几乎是标普和穆迪方法测算结果的两倍，这表明不同的测算方式对违约概率的影响有可能是非常大的。那么如何充分利用考察期间内所有的贷款信息来精确地测算违约概率？有没有一种更合理的测算方法？鉴于违约概率变量在信用风险度量以及信用评级中的重要性，我们需要回答这些问题。

本文将从违约强度和违约概率的关系入手，提出测算违约概率的新思想，并给出利用考察期内所有债项信息来测算违约概率的新方法。考虑到无论是银行贷款还是债券，利用历史数据针对不同信用级别测算违约概率本质上是相通的。因此除非特别说明，在本文的理论与实证部分，我们行文时均以银行贷款的评级为例加以论述。

二、文献研究回顾

违约概率的测算方法多见于信用风险量化管理模型中。阐述信用风险量化管理模型不是

本文的任务，我们只关注其中违约概率测算的部分，目的是说明已有的违约概率测算方法与本文提出的测算方法之间的联系，便于读者理解本文方法的思路和意义。

1. 基于统计学原理的测算方法

设某一样本中有 N 笔贷款，在一个特定的时间内（比如 1 年）有 n 笔贷款违约，我们可以把违约的频率 $f=n/N$ 近似地看作这组贷款的违约概率。该方法原理简单，但其测算的结果往往成为一些信用风险管理模型的基础。例如著名的信用风险管理模型 CreditMetrics 中的信用转移矩阵 A ($a_{i,j}$) 就是借助这种方法测算的 (CSFP, 1997; Gordy, 2000)。

另一类基于统计原理测算违约概率的方法可参见彭建刚等 (2009)、Divino 等 (2013) 的研究工作。这些工作所采用的方法本质上是生存分析的研究方法：一个生物个体在考察期间有“生存”和“死亡”两种状态。同样在考察期间，贷款也可以认为存在两种状态——“违约”和“不违约”。基于这种对应关系，可将最初用于生物个体的生存和死亡概率的统计方法运用到贷款的违约概率测算中。但正如本文开篇所述，该方法在抽样上采用的是静态的方法，因而不可避免地带有一定的误差。

2. 结构化模型的测算方法

代表性的研究成果是穆迪公司的 KMV 模型 (Crosbie 和 Bohn, 2003)。该方法有个隐含的假定：就是债务的违约概率等同于企业的破产概率。若企业的资产价值低于其债务价值，企业就会被破产清算。KMV 模型将企业的股权看作是以企业价值 V 为标的资产，以企业的到期负债 X 为执行价格，到期期限为 T (通常为 1 年) 的欧式看涨期权。假定企业资产价值 V 服从几何布朗运动，根据期权定价公式，建立企业股权价值与企业价值之间的数量方程。同时，根据期权定价的原理建立股权价值的波动率与企业价值的波动率之间的数量方程，通过联立这两个方程，可以求出企业期初的资产价值 V_0 和企业的资产价值波动率 σ_V 。从而测算出 1 年后企业的资产小于到期债务的概率，也就是违约概率 PD。当然，这只是理想化的模型。KMV 基于穆迪的历史数据发现，企业通常并不是在其资产价值低于到期负债时就立刻实行破产清算，而是当它低于一个小于到期负债的阈值时才被清算，这个阈值 KMV 模型称为违约点 DV。为此，KMV 模型根据穆迪公司积累的违约数据库对测算的违约概率进行修正：假定资产的预期回报率为 μ ，则定义 DD 为违约距离，即：

$$DD = \frac{\log(V_0/DV) + (\mu - \sigma_V^2/2)T}{\sigma_V \sqrt{T}} \quad (2)$$

它描述了 1 年后企业价值的平均值偏离违约点的程度，违约距离越大，企业违约概率越小。若某企业的 DD 值为 3，KMV 则根据历史数据统计 DD 值为 3 的企业在 1 年内的违约频率，然后用违约频率作为该企业的预期违约概率。这表明：即使建立在现代金融理论基础上的违约概率测算方法，也仍有不足，需要根据经验数据来校准违约概率的测算结果。从研究的趋势看，改善该模型测算精度的工作主要集中在：第一，放松企业价值遵从几何布朗运动的假设，将跳跃过程引入到描述企业价值变化的模型中，用以反映突发事件对企业价值的影响 (Jessen 和 Lando, 2015)；第二，用障碍期权来代替看涨期权从而更真实刻画企业的破产机制 (Tudela 和 Young, 2005)；第三，如何合理确定违约点，使违约距离与违约概率之间的对应关系更稳健 (Jessen 和 Lando, 2015)。

3. 基于计量经济学的测算方法

比较典型的是 Probit 模型和 Logistic 模型 (Tian 等, 2015; Divino 和 Rocha, 2013;

Campbell 等, 2008)。它的出发点是用一组影响企业违约的变量去解释违约概率。这类模型的原理并不复杂, 关键是如何选取恰当的解释变量。较早的研究主要采用企业的财务指标作为解释变量 (Altman, 1968; Shumway, 2001)。Campbell 等 (2008) 提出加入市场化的指标来提高模型的解释力度。另外, 也有研究把 KMV 模型中违约距离 DD 纳入到解释变量之中 (Duffie 等, 2007; Lando 和 Nielsen, 2013)。2007 年的次贷危机以后, Divino 和 Rocha (2013) 开始尝试把宏观经济变量如基准利率等引入到 Logistic 模型中。

4. 基于保险精算的方法

代表性的成果有瑞士银行金融产品部开发 Creditrisk+模型。该模型假设债务人要么违约, 要么不违约。对于众多的债务人而言, 债务人的违约率都非常小, 并在给定的期限内, 债务人的违约彼此独立。于是发生违约的债务人的数目 k 近似服从泊松分布。如果知道 k 个人的风险暴露以及相应的违约损失率, Creditrisk+模型就可以测算银行信贷组合损失的分布。但首要的问题是要估计单个债务人的违约概率。从瑞士银行披露的技术文档和研究文献来看, 单个债务人违约概率多是根据债务人的债项级别采用历史数据进行估计 (Gordy, 2000; 2002)。

除前述 4 种方法外, Berardi 等 (2004) 利用有风险的债券市场价格信息估计债券的违约概率。该方法必须事先估计债券的违约损失率, 由于违约损失率估计比较困难, 类似的研究工作比较少见。此外近年来一些模式识别技术如分类树 (CT)、神经网络、支持向量机等方法被运用于预测债务是否违约。这些技术一般不能直接给出债务的违约概率 (陈诗一, 2008), 过程比较复杂且难以解释, 导致在测算违约概率方面它们的应用受到限制。

5. 研究评述

除基于统计学原理测算违约概率的方法外, 在违约概率的建模过程中不断有新技术和理论被采用, 这或许从另一个角度说明原有模型的测算精度仍满足不了实际需要。学者们试图采用更丰富的数据信息来提高违约概率的测算精度, 例如 KMV 模型在估计违约距离时采用了企业股票的价格信息、Logistic 模型中采用市场化变量以及宏观经济变量等。尤其是市场化的变量, 它能及时反映市场对债务人经济情况或偿债能力的看法, 相对于财务指标而言会更灵活, 信息量更丰富。不过这些模型测算得是否准确, 很大程度上仍需要历史数据进行检验。只要样本的数量足够多, 利用事后的数据依照统计学原理计算的违约概率, 可以帮助检验模型是否有效。从这个意义上说, 基于统计原理测算违约概率仍有不可替代的作用。

不仅如此, 多数违约概率测算模型的参数或结果需要根据历史或经验数据来校准或修正。比如, CreditMetric 模型中信用转移矩阵的计算、KMV 模型中违约距离与预期违约概率的对应关系、Creditrisk+模型中单笔债务的违约概率的测算、Logistic 模型中样本的配比问题, 这些都依赖于借助历史数据测算的违约概率。值得一提的是, 上述模型对违约概率的测算多是基于一定的时间期限 (通常为 1 年), 且依照特定的分组变量 (信用等级、违约距离等) 来进行的, 所面临的问题和本文开篇提出的问题在本质上具有相同或相似的特点。如何充分挖掘和利用历史数据的信息, 使违约概率测算得更准确, 不仅对评估评级质量、进行有效风险管理至关重要, 也会让其他测算违约概率的方法从中受益, 而这正是本文致力解决的问题。

三、基于违约强度的违约概率测算方法

1. 违约强度和违约概率的期限结构

在前文所描述的问题中，当我们考虑如何处理类似 29 号、30 号这样的贷款时，不妨换个角度：类似 29 号、30 号这样的贷款出现，本质上相当于在考察期间，用来测算违约概率的贷款总数在不断发生变化。为了描述贷款总数以及违约贷款数随时间变化的情形，我们需要借助违约强度的概念。

违约强度概念最早由 Jarrow 和 Turnbull (1995)、Jarrow 等 (1997)、Duffie 和 Singleton (1999) 等文献提出，它反映某项债务在单位时间内发生违约的概率。引入违约强度后，我们能够将贷款违约的变化与时间变量联系起来。假定我们在考察期 $[0, T]$ 范围内考察某信用级别贷款的违约概率，如果将考察期内任意时刻 t ($t \in [0, T]$) 贷款的违约强度 λ 、贷款数 N 均看成是随时间 t 变化的量，那么从 t 时刻开始，在很短的时间 Δt 内，发生违约的贷款数 Δn 可以表示成：

$$\Delta n = N(t)\lambda(t)\Delta t \quad (3)$$

式 (3) 右侧中的 $\lambda(t)\Delta t$ 的含义容易理解：因为违约强度 λ 代表单位时间的违约概率，所以乘以 Δt 后， $\lambda(t)\Delta t$ 就是代表在 Δt 的时间内该级别贷款的违约概率。于是，在考察期 $[0, T]$ 内发生违约的贷款数 n 可以写成：

$$n = \int_0^T N(t)\lambda(t)dt \quad (4)$$

在式 (4) 中，当我们将贷款数目 $N(t)$ 看成是随时间变化的量，诸如 29 号或 30 号贷款的情况就可以方便处理。同样，类似 15 号贷款的情况所引起的贷款数的变化也可以方便处理。

简单起见，我们将参照 Cantor 和 Falkenstein (2001) 将违约强度 $\lambda(t)$ 近似看成是一个不随时间而变化的量。当然，这样做是有前提的：第一，贷款具有较好的同质性，这使得对一组需要考察的贷款而言，不同贷款的违约概率之间的差异很小。第二，考察期限较短，这样在比较短的时间内影响贷款违约强度的各种因素变化会比较小。在本文中，当考察的贷款是某一级别的贷款时，贷款违约风险的同质性较好。如果将考察期限设定为 1 年或者更短的时间，相对于一个宏观的经济周期而言，影响违约强度的因素变化会比较小，这样的处理就比较合理。于是有：

$$n = \lambda \int_0^T N(t)dt \quad (5)$$

为了测算违约概率，我们将式 (5) 进行适当的变换，即：

$$\lambda T = \frac{n}{\frac{1}{T} \int_0^T N(t)dt} \quad (6)$$

考虑到违约强度是单位时间内的违约概率，所以式 (6) 的左侧实际上代表的是一组贷款在考察期 $([0, T])$ 内理论上的违约概率，而公式右侧的分母则代表在 $[0, T]$ 内，该组贷款时间意义上的平均数目，于是违约概率的测算公式可以写成：

$$P = \lambda T = \frac{n}{\frac{1}{T} \int_0^T N(t) dt} = n/\bar{N} \quad (7)$$

式(7)的第一个等式清晰地显示出违约概率的期限结构,表明对违约概率的测算与考察期的长短密切相关。换句话说,如果不能事先言明测算的违约概率所对应的期限,那么违约概率数值将没有意义。因为,即使对于同是 BBB 等级的两组贷款而言,如果每组贷款的考察期不同,那么违约概率的测算结果也不一样。

式(7)给我们提供了一个理论上计算违约概率的新方法。需要指出的是,对于一个确定的考察期间而言,这里的违约贷款数 n 是能够观察到的,因而可以看成是已知的。所以,若要根据历史数据计算违约概率,关键是正确计算考察期内平均的贷款数目 \bar{N} 。为此需要建立一个描述贷款数目随时间变化的函数 $N(t)$ 。

2. 描述贷款数目时变特征函数的构造方法

通常,银行每天需要处理许多贷款业务。因此在考察期内的每个时点,贷款数目随时都有可能发生改变,这使式(7)中描述贷款数随时间变化的函数 $N(t)$ 变得非常复杂,用一般的函数形式很难描述。本文从贷款的起止时刻与违约概率考察期间的时序关系入手,创造性地利用特征函数的方法,构造描述贷款数的时变特征函数 $N(t)$ 。

我们仍以考察期内某一信用级别的贷款数目为对象进行讨论。在图 1 中,我们用一个坐标轴表示时间 t 。为描述方便,我们定义考察期的起始时点为 0, 考察期的结束时点为 T , 用 t_s 和 t_e 分别表示某个贷款的起止时间。这里的“起止时间”与通常贷款的自然寿命的起止时间有所不同。确切地说,这里的 t_s 表示贷款出现在该级别的起始时间,这种“出现”所引起的贷款数目的增加通常有两种情形:一是银行向客户发放了该级别的贷款。二是由于贷款的信用质量发生变化,使原已存在的贷款在 t_s 时点迁移至此级别。例如,由原来的 AAA 级别调整为 AA 级别,这会引起 AA 级别的贷款数增加。同样, t_e 则表示贷款在该级别的实际截止时间。显然,这种“截止”会导致该级别贷款数减少。它通常包括以下情形:第一,贷款正常到期。第二,贷款发生违约。比如,企业经营持续恶化,无力偿还债务;或者当企业或有负债的潜在义务转换成现实义务时,导致企业现金流断裂而出现突发性的违约等。第三,贷款人提前偿还贷款。在利率下行的环境中,企业可能会急于借新债还旧债以降低债务成本,通过提前还贷终止旧的贷款合同。第四,贷款的信用质量发生变化,该贷款迁移至其他信用级别。第五,其他情况,比如因贷款企业被其他企业并购而重新做出债务安排等。于是,贷款的起止时刻与考察期两端的时间节点之间的时序关系就可以用 A、B、C、D 四种情况来表示。其中, A 表示在考察期开始时贷款就已存在,且在考察期内某个时点 t_e ,贷款截止; B 表示在考察期开始时贷款已存在,但考察期结束时贷款尚未截止,合同仍在正常履行; C 表示考察期内某时点出现的贷款,并在考察期内截止; D 表示考察期内出现的贷款,考察期结束时合同仍在正常履行,贷款没有截止。

对于 A 情况,我们针对一笔贷款构造一个在 $[0, T]$ 上的二值函数 $\varphi_A(t)$:

$$\varphi_A(t) = \begin{cases} 1 & \text{if } 0 \leq t < t_e \\ 0 & \text{if } t_e \leq t \leq T \end{cases} \quad (8)$$

函数 $\varphi_A(t)$ 准确描述了单笔贷款在时点 t_e 截止前后所引起贷款数目变化。因此,我

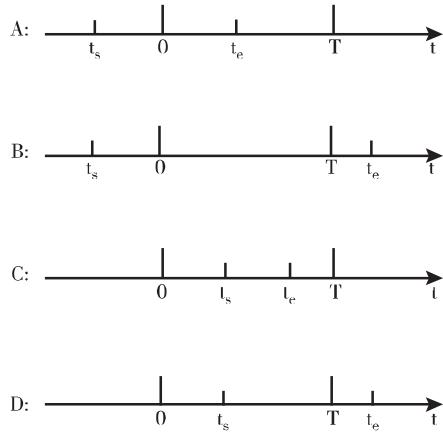


图 1 贷款起止时刻与考察期间的时序关系

们定义 $\varphi_A(t)$ 为 A 情况下描述贷款数目变化的特征函数。类似地，我们可以构造出图 1 中的 B、C、D 三种情况的特征函数（见表 2）。

表 2 A、B、C、D 四种情况特征函数

情况	特征函数	对时间的积分	情况	特征函数	对时间的积分
A	$\varphi_A(t) = \begin{cases} 1 & \text{if } 0 \leq t < t_e \\ 0 & \text{if } t_e \leq t \leq T \end{cases}$	$\int_0^T \varphi_A(t) dt = t_e$	C	$\varphi_C(t) = \begin{cases} 1 & \text{if } t_s \leq t < t_e \\ 0 & \text{else} \end{cases}$	$\int_0^T \varphi_C(t) dt = t_e - t_s$
B	$\varphi_B(t) = 1, (0 \leq t \leq T)$	$\int_0^T \varphi_B(t) dt = T$	D	$\varphi_D(t) = \begin{cases} 1 & \text{if } t_s \leq t \leq T \\ 0 & \text{else} \end{cases}$	$\int_0^T \varphi_D(t) dt = T - t_s$

借助特征函数，我们可以建立描述贷款数随时间变化的函数 $N_A(t)$ 。假定 A 类情况的贷款数总计有 n_A 个，每个 A 类贷款对应的特征函数为 $\varphi_A^{(i)}(t)$ ，这里的 $i \in \{1, \dots, n_A\}$ 表示 A 类贷款中每个贷款的编号。那么描述 A 类贷款总数随时间变化的函数 $N_A(t)$ 可以写成：

$$N_A(t) = \sum_{i=1}^{n_A} \varphi_A^{(i)}(t) \quad (0 \leq t \leq T) \quad (9)$$

类似地，我们可以定义 B、C、D 类贷款总数随时间变化的函数 $N_B(t)$ 、 $N_C(t)$ 、 $N_D(t)$ ：

$$N_B(t) = \sum_{j=1}^{n_B} \varphi_B^{(j)}(t) \quad N_C(t) = \sum_{k=1}^{n_C} \varphi_C^{(k)}(t) \quad N_D(t) = \sum_{q=1}^{n_D} \varphi_D^{(q)}(t) \quad (0 \leq t \leq T) \quad (10)$$

式 (10) 中的 n_B 、 n_C 、 n_D 分别代表 B、C、D 三种情况下的贷款总数。于是，描述贷款数随时间变化的函数 $N(t)$ 可以写成：

$$N(t) = N_A(t) + N_B(t) + N_C(t) + N_D(t) \quad (0 \leq t \leq T) \quad (11)$$

将式 (11) 代入到式 (7) 分母的积分项中，于是有：

$$\int_0^T N(t) dt = \int_0^T N_A(t) dt + \int_0^T N_B(t) dt + \int_0^T N_C(t) dt + \int_0^T N_D(t) dt \quad (12)$$

依据特征函数对时间的积分公式，有下面的数量关系成立：

$$\int_0^T N(t) dt = \sum_{i=1}^{n_A} t_e^{(i)} + \sum_{j=1}^{n_B} T + \sum_{k=1}^{n_C} (t_e^{(k)} - t_s^{(k)}) + \sum_{q=1}^{n_D} (T - t_s^{(q)}) \quad (13)$$

在式 (13) 右侧的计算中，由于考察期的参数 T 为已知，所以需要处理的参数主要是每笔贷款的起止时间 t_s 和 t_e 。站在银行的角度，尽管每笔贷款的起止时间 t_s 和 t_e 具有一定的随机性，但通常会具有一定的统计特征。根据贷款信用等级迁移规律，考虑到对于特定级别的贷款而言，其在后继的时间内发生信用等级迁移的数目只占很小的比例。因此，我们借鉴彭建刚等 (2009) 的研究，主要参照贷款的生命周期规律，假定贷款起始和截止时间参数 t_s 和 t_e 在 $[0, T]$ 区间内服从均匀分布。换句话说，贷款的起始和截止事件在考察期内的每个时点都是等可能发生的。在考察期间比较短，影响贷款供给和需求的因素变化不大的情况下，这种假定是合理的。于是在大样本条件下，式 (13) 可以表示成：

$$\begin{aligned} \int_0^T N(t) dt &= n_A E(t_e^{(i)}) + n_B T + n_C E(t_e^{(k)} - t_s^{(k)}) + n_D E(T - t_s^{(q)}) \\ &= n_A \frac{1}{2} T + n_B T + n_C \frac{1}{6} T + n_D \frac{1}{2} T \end{aligned} \quad (14)$$

式 (14) 中的 $E(\cdot)$ 为求期望值的函数^①。将式 (14) 中的 $\int_0^T N(t) dt$ 代入到式 (7) 中，则在考察期间的违约概率计算公式可以表示成：

$$P = \frac{n}{\frac{1}{T} \int_0^T N(t) dt} = \frac{n}{\frac{1}{2} n_A + n_B + \frac{1}{6} n_C + \frac{1}{2} n_D} \quad (15)$$

式 (15) 即为本文给出的利用历史数据计算违约概率的新方法。该方法在计算考察期间的贷款数目时，并不是像标准普尔或穆迪公司一样只是利用考察期初的贷款数目，也不是简单地将 A、B、C、D 四种情形所对应的贷款数目进行直接加总，而是将 A、C、D 三种情形的贷款数目乘上了一个系数。其根本原因在于，这三类贷款都不是在考察期间 $[0, T]$ 的整个时域范围内全程存在的。因此，可以将 $n_A/2$ 、 $n_C/6$ 、 $n_D/2$ 看作是 A、C、D 类三类贷款在考察期间全时域范围内的等效贷款数目，而 $1/2$ 、 $1/6$ 、 $1/2$ 可看作是 A、C、D 类三类贷款数目 n_A 、 n_C 、 n_D 的“全时域”转换系数。这样计算的违约概率才是真正反映该级别贷款在考察期间 $[0, T]$ 内的违约概率。由于 A、B、C、D 四种情形的划分涵盖了“贷款起止时间与考察期两端时间节点”的全部时序关系，这意味着该方法实际上充分利用了考察期间该级别贷款的所有信息，从而大大减少了信息损失，提高了测算不同信用级别贷款违约概率的准确性。

四、实证研究

实证部分我们用某商业银行的贷款数据来做一个具体的算例，以此说明针对不同信

^① 式 (14) 中的右侧第三项可以用简单的二重积分计算得出。即： $\sum_{k=1}^{n_C} E(t_e^{(k)} - t_s^{(k)}) = \sum_{k=1}^{n_C} \int_0^T \int_0^{t_e} (t_e - t_s) \frac{1}{T^2} dt_s dt_e = n_C \frac{1}{6} T$

用级别的贷款其违约概率的测算方法。我们获得的贷款样本数据的时间跨度为2013年10月1日到2015年3月31日，贷款的客户均为企业，不含面向个人发放的贷款数据。

1. 测算步骤

第一步，界定“违约”的定义。按照巴塞尔协议，如客户出现以下两种情况之一时，即被视为违约。一是除非采取追索措施，如变现抵押品（如果存在的话），否则借款人可能无法全额偿还银行的债务。二是债务人对银行的实质性信贷逾期90天以上。我国商业贷款的5级分类标准关于次级类、可疑类、损失类的贷款定义与巴塞尔协议关于违约的定义基本一致，因此本文参考彭建刚等（2009）将贷款变成次级或次级以下信用等级时，界定为“违约”。

第二步，确定测算违约概率的考察期间 $[0, T]$ 。前已述及，如果将考察期间的违约强度近似看成是常数，那么考察期间就应该尽量取得短些。在兼顾样本数量的前提下，本文以一个季度为违约概率考察期间。

第三步，对于每笔贷款，根据贷款发放的时间，信用等级迁移的时间、违约的时间以及贷款的到期时间，确定贷款的起止时间 t_s 和 t_e 。特别地，如果一个贷款在原信用等级的 t_e^{old} 时点发生信用迁移，那么该时点亦同时被看作它在新的信用等级的起始时间 t_s^{new} 。

第四步，按照贷款的信用级别，统计每个信用级别中，分属A、B、C、D四类情况的贷款数和对应的违约贷款数，然后根据式（15）计算出每个季度该信用级别贷款的违约概率。

2. 初步的测算结果

表3给出了该商业银行在2015年第1季度的4977笔贷款样本按照信用级别进行统计所得到的违约情况的统计结果。

表3 2015年第1季度贷款违约情况及违约概率的测算

客户评级	数量标记	情形				合计	违约概率
		A	B	C	D		
AAA	$N_{default}$	0	0	0	0	0	0.00000
	N_{g_total}	13	40	5	91	149	
AA	$N_{default}$	0	0	0	0	0	0.00000
	N_{g_total}	7	81	55	137	280	
A	$N_{default}$	0	0	0	0	0	0.00000
	N_{g_total}	17	83	27	301	428	
BBB	$N_{default}$	0	1	0	3	4	0.00642
	N_{g_total}	12	106	13	1017	1148	
BB	$N_{default}$	0	3	0	6	9	0.00912
	N_{g_total}	43	287	38	1345	1713	
B	$N_{default}$	0	0	0	0	0	0.00000
	N_{g_total}	32	93	17	1063	1205	

(续)

客户评级	数量标记	情形				合计	违约概率
		A	B	C	D		
CCC	$N_{default}$	0	0	0	0	0	0.00000
	N_{g_total}	0	8	0	0	8	
CC	$N_{default}$	0	0	0	1	1	0.25000
	N_{g_total}	0	0	0	8	8	
C	$N_{default}$	0	4	0	3	7	0.27451
	N_{g_total}	7	13	0	18	38	

注: $N_{default}$ 代表该行统计的是某一级别 A、B、C、D 四种情况下贷款的违约数目 n_a 、 n_b 、 n_c 、 n_d 。 N_{g_total} 代表该行统计的是某一级别 A、B、C、D 四种情况下的贷款总数 n_A 、 n_B 、 n_C 、 n_D (含违约的贷款数)。

以 BBB 级贷款为例, 该级别的贷款总共有 1148 笔贷款。其中 A、B、C、D 四种情形所对应的贷款数目 n_A 、 n_B 、 n_C 、 n_D 分别为 12、106、13、1017 笔, 与之相对应的违约贷款的数目 n_a 、 n_b 、 n_c 、 n_d 分别为 0、1、0、3 笔, 总的违约数为 4 笔。将上述统计结果代入式(15) 即可计算出该银行 BBB 等级的贷款在 2015 年第 1 季度的违约概率为 0.00642。其他信用级别贷款的违约概率测算均与此类似。

类似地, 我们可以计算出 2013 年 10 月~2015 年 3 月连续 6 个季度每个信用级别的违约概率, 若这些季度的违约概率依次表示为 p_q^i ($i=1, 2, 3, 4, 5, 6$), 则随着时间推移, 到 2015 年 3 月 31 日, 该信用级别累积违约概率 P_c 可以表示为:

$$P_c = 1 - \prod_{i=1}^6 (1 - p_q^i) \quad (16)$$

表 4 给出了 2013 年第 4 季度~2015 年第 1 季度不同信用等级的贷款违约率的测算结果, 仍以 BBB 信用等级贷款为例, 表 4 第 2 列至第 7 列依次显示的是 2013 年第 4 季度~2015 年第 1 季度的违约概率测算结果。依此测算出该信用级别的贷款经过连续 6 个季度后, 其累积的违约概率为 0.03667。表 4 中其他信用级别贷款的违约概率、累积违约概率计算与此类似, 不再赘述。需要说明的是, 除 2015 年第 1 季度外, 贷款的样本数据中没有出现 CC 和 C 级别的贷款, 故表 4 中这两个级别的计算结果没有列出。

表 4 2013 年第 4 季度~2015 年第 1 季度不同信用等级的贷款违约概率测算结果

信用等级	概率测算区间的季度时间序列						累计概率 P_c
	2013 年 Q4 (4833)	2014 年 Q1 (3997)	2014 年 Q2 (5472)	2014 年 Q3 (5381)	2014 年 Q4 (7624)	2015 年 Q1 (4977)	
AAA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
AA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
A	0.00000	0.00000	0.00104	0.00000	0.00000	0.00000	0.00104
BBB	0.00820	0.00777	0.00182	0.00684	0.00619	0.00642	0.03667
BB	0.00000	0.01021	0.00960	0.01175	0.01518	0.00912	0.05539
B	0.01447	0.01205	0.01456	0.02377	0.02425	0.00000	0.08604
CCC	0.05797	0.04255	0.05714	0.00000	0.07477	0.00000	0.21318

注: 表头第二行括号中的数字为测算时样本量。

3. 季度违约概率与年度违约概率

根据表4提供的结果，我们可以测算6个季度违约概率的平均值。假定某级别其季度违约概率的平均值为 \bar{p}_q ，那么经过6个季度后，该级别贷款的累积违约概率可以表示为：

$$P_c = 1 - (1 - \bar{p}_q)^6 \quad (17)$$

对比式(16)，于是违约概率的季度平均值可由式(18)给出：

$$\bar{p}_q = 1 - \left(\prod_{i=1}^6 (1 - p_q^i) \right)^{1/6} \quad (18)$$

由此，我们还可以进一步计算出该级别贷款的平均意义上的1年期违约概率为：

$$p_{year} = 1 - (1 - \bar{p}_q)^4 \quad (19)$$

表5给出了不同贷款级别的季度平均违约概率和年度违约概率的测算结果。

表5 季度和年度平均违约概率

信用级别	AAA	AA	A	BBB	BB	B	CCC
\bar{p}_q	0.00000	0.00000	0.00021	0.00626	0.00922	0.01474	0.03438
$\sigma(p_q)$	0.00000	0.00000	0.00052	0.00241	0.00502	0.00876	0.02777
p_{year}	0.00000	0.00000	0.00085	0.02481	0.03638	0.05767	0.13059

4. 两种测算方法的比较

下文我们将本文测算的方法与穆迪和标准普尔的测算方法进行比较。在表6中，我们用标准普尔和穆迪的方法测算了不同信用级别的季度违约概率 p_q^* ，然后将其与本文测算的结果 p_q 进行对比。我们发现，除去表中带有括号的三组数据外，对于有违约记录的几个级别(A级到CCC级)，本文的测算结果均略高于标准普尔和穆迪方法的测算结果。考虑到本文的测算结果是基于银行的贷款抽样，这个比较结果是偶然的呢，还是一种规律使然？为回答这个问题，我们有必要研究本文测算方法在大样本条件下的统计性质。

表6 两种概率测算结果的比较

信用等级	概率标识	概率测算区间的季度时间序列					
		2013年Q3	2014年Q1	2014年Q2	2014年Q3	2014年Q4	2015年Q1
AAA	p_q	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	p_q^*	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
AA	p_q	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
	p_q^*	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
A	p_q	0.00000	0.00000	(0.00104)	0.00000	0.00000	0.00000
	p_q^*	0.00000	0.00000	(0.00127)	0.00000	0.00000	0.00000
BBB	p_q	0.00820	0.00777	0.00182	0.00684	0.00619	(0.00642)
	p_q^*	0.00757	0.00755	0.00171	0.00628	0.00597	(0.00848)
BB	p_q	0.00000	0.01021	0.00960	(0.01175)	0.01518	0.00991
	p_q^*	0.00000	0.00985	0.00938	(0.01197)	0.01497	0.00909

(续)

信用 等级	概率 标识	概率测算区间的季度时间序列					
		2013 年 Q3	2014 年 Q1	2014 年 Q2	2014 年 Q3	2014 年 Q4	2015 年 Q1
B	p_q	0.01447	0.01205	0.01456	0.02377	0.02425	0.00000
	p_q^*	0.01445	0.01195	0.01455	0.02334	0.02396	0.00000
CCC	p_q	0.05797	0.04255	0.05714	0.00000	0.07477	0.00000
	p_q^*	0.04167	0.04167	0.05714	0.00000	0.06383	0.00000

根据式 (15)，本文测算的季度违约概率的公式可以表示成：

$$p_q = \frac{n_a + n_b + n_c + n_d}{\frac{1}{2}n_A + n_B + \frac{1}{6}n_C + \frac{1}{2}n_D} \quad (20)$$

其中， n_a 、 n_b 、 n_c 、 n_d 分别为 A、B、C、D 四种情况下，在考察期内各自产生的违约贷款数目。为了比较两种方法测算的概率，我们对式 (20) 分号上下同除以考察期初的贷款总数目 $n_A + n_B$ ，于是有：

$$p_q = \frac{\frac{n_a + n_b}{n_A + n_B} + \frac{n_c + n_d}{n_C/6 + n_D/2} \times \frac{n_C/6 + n_D/2}{n_A + n_B}}{1 + \frac{n_C/6 + (n_D - n_A)/2}{n_A + n_B}} \quad (21)$$

而 $\frac{n_a + n_b}{n_A + n_B}$ 实际上为标准普尔和穆迪方法测算的违约概率 p_q^* 。如果考虑一个在考察期间 $[0, T]$ 的初始时刻刚刚成立的银行的话，那么本质上式 (21) 分子中的 $\frac{n_c + n_d}{n_C/6 + n_D/2}$ 即为本文测算的考察期间的违约概率。事实上深究而论，无论考察期初银行是否刚刚成立，在大样本条件下，这种数量关系都成立。即：

$$p_q = \lim_{\substack{n_C \rightarrow \infty \\ n_D \rightarrow \infty}} \frac{n_c + n_d}{n_C/6 + n_D/2} \quad (22)$$

于是在式 (21) 中，大样本条件下两种方法测算的概率之间的数量关系可以描述成：

$$p_q = \frac{p_q^* + \delta_1 p_q}{1 + \delta_2} \Rightarrow p_q = \frac{1}{1 - (\delta_1 - \delta_2)} p_q^* \quad (23)$$

这里的 $\delta_1 = \frac{n_C/6 + n_D/2}{n_A + n_B}$ ， $\delta_2 = \frac{n_C/6 + (n_D - n_A)/2}{n_A + n_B}$ 。在经济意义上， δ_1 可以看作是考察期间增加的贷款数与期初的贷款数之比，因此我们可将 δ_1 看成是一个贷款数的增长率，只不过增长的部分是用经过转换的考察期“全时域”范围内等效的贷款数来表示。类似地， δ_2 则可以看作是考察期内贷款数的净增长率。一般情况下，显然有： $\delta_1 > \delta_2$ 成立。于是有：

$$\frac{p_q}{p_q^*} = \frac{1}{1 - (\delta_1 - \delta_2)} > 1 \quad (24)$$

这就证明了大样本条件下，本文方法测算的违约概率要高于标准普尔和穆迪的测算结果。说明表 6 中两种概率的比较结果是受内在的统计规律支配的。由于本文的测算方法充分

利用了考察期间一个信用级别的所有贷款及其违约的信息，这或许表明标准普尔和穆迪测算的违约概率存在一定程度的低估。特别地，如果 $n_A=0$ ，也就是说所有在考察期 $[0, T]$ 的初始时刻就已经存在的贷款，在考察期结束时仍在继续正常履行合同，那么两种方法测算的违约概率是相等的，不过这种情况极为少见。

五、结束语

就某一信用级别的债项而言，历史违约数据是对该级别债项违约风险的最直观反映。因此，充分利用和挖掘历史数据的信息，更加准确测算违约概率是一个值得仔细思考的题目，然而现有的文献对此却缺乏深入细致的讨论。本文从违约强度与违约概率的关系出发，提出计算违约概率的新思想。在分析贷款的起止时间与违约概率考察期间的时序关系之后，通过创新性地构造特征函数的方式，描述了考察期间贷款数目的时变特征，提出了利用历史数据测算违约概率的新方法。

与传统的数据分析关注贷款的“自然寿命”不同，我们对贷款的起止时间进行了严格的界定，这种界定是针对特定的信用级别以及考察期而论的。它与考察期间的时序关系被完整地划分成 A、B、C、D 四种典型类型。换句话说，任何一个属于某信用级别的贷款，只要它在考察期间历史地存在过，不管因为什么原因在什么时点出现，也不管什么原因在什么时点离开，其自身风险经评估后而被赋予的级别符号，都折射出评级机构或银行对该风险级别评估标准的分寸把握以及对贷款风险的体认。因此，在测算该级别违约概率、评价该级别的评级质量时都理应被加以考虑，而本文的测算方法恰好提供了这样的解决方案。于是采用本文的方法测算违约概率时，考察期内的任何一笔贷款都能成为潜在测算样本。相比之下，传统的类似生命表的方法则只选择在考察期的开始时刻存在的贷款作为测算的源数据，这在一定程度上弱化了样本数据的代表性。从这个意义上说，本文的方法利用历史违约数据的信息更完全，因而更有应用价值。

在实证研究中，通过比较两种方法的违约概率测算结果，本文发现：在大样本条件下，传统的测算方法对信用级别的违约概率存在一定程度的低估。对商业银行而言，它为商业银行更加合理地计提损失准备金，测算所需配置的经济资本提供了新的参考，也有助于更合理地测算商业银行的信贷损失分布。就债券市场而言，低估违约概率会造成高估评级，本文的测算结果或许能从另一个视角解释评级机构存在高估评级的现象，也为评级机构基于历史数据更合理地审视自身的评级质量提供了工具。

当然，本文方法运用是以丰富的数据源为前提的。就商业银行而言，贷款是其主要业务之一，在经营过程中能够积累大量的贷款数据。因此，即便是较短的考察期间，客观上也能够获得本方法所需要的数据样本。在证券市场方面，随着市场化改革的深入，发行制度的改革最终会把债券风险识别的责任更多地交给市场，在“产能过剩、经济下行压力加大”的条件下，市场出现少数债券违约的现象将会变成一种“新常态”，当违约数据积累到一定程度后，考察评级机构的评级质量所需的数据条件将更加充分，这也为本文方法的运用提供了更好的前景。

参 考 文 献

- [1] Altman E. I. , 1968, *Financial Ratios, Discriminant Analysis and Prediction of Corporate Bank-*

- ruptcy [J], *Journal of Finance*, 23 (4), 589~610.
- [2] Antnio C., Popova I., Simkins B., 2012, *A Comparative Study of the Probability of Default for Global Financial Firms* [J], *Journal of Banking & Finance*, 36 (3), 717~732.
- [3] Bellalah M., Zouari S., Levyne O., 2016, *The Performance of Hybrid Models in the Assessment of Default Risk* [J], *Economic Modelling*, 52 (1), 259~265.
- [4] Berardi A., Ciraolo S., Trova M., 2004, *Predicting Default Probabilities and Implementing Trading Strategies for Emerging Markets Bond Portfolios* [J], *Emerging Markets Review*, 5 (4), 447~469.
- [5] Bolton P., Freixas X., Shaprio J., 2012, *The Credit Ratings Game* [J], *Journal of Finance*, 67 (1), 85~111.
- [6] Campbell J. Y., Hilscher J., Szilagyi J., 2008, *In Search of Distress Risk* [J], *Journal of Finance*, 63 (6), 2899~2939.
- [7] Cantor R., Falkenstein E., 2001, *Testing for Rating Consistency in Annual Default Rates* [J], *The Journal of Fixed Income*, 11 (2), 36~51.
- [8] Credit Suisse Financial Products, 1997, *Creditrisk+ : A Credit Risk Management Framework* [R], Credit Suisse Financial Products.
- [9] Crosbie P., Bohn J., 2003, *Modeling Credit Risk* [R], Moody's KMV.
- [10] Crouhy M., Galai D., Mark R., 2000, *A Comparative Analysis of Current Credit Risk Models* [J], *Journal of Banking and Finance*, 24 (1~2), 59~117.
- [11] Diane V., Nick W. K., 2015, *Default, Transition, and Recovery: Annual Global Corporate Default Study and Rating Transitions* [R], S&P Global Ratings.
- [12] Diane V., Nick W. K., 2013, *Default, Transition, and Recovery: Annual Global Corporate Default Study and Rating Transitions* [R], S&P Global Ratings.
- [13] Divino J. A., Rocha L. C., 2013, *Probability of Default in Collateralized Credit Operations* [J], *North American Journal of Economic and Finance*, 25 (1), 276~292.
- [14] Duffie D., Singleton K., 1999, *Modeling Term Structure of Defaultable Bonds in a Signaling Jump-Default Model* [J], *Review of Financial Studies*, 12 (4), 687~720.
- [15] Duffie D., Saita L., Wang K., 2007, *Multi-period Corporate Default Prediction with Stochastic Covariates* [J], *Journal of Financial Economics*, 83 (3), 365~384.
- [16] Fons J. S., Carty L., Kaufman J., 1994, *Corporate Bond Default and Default Rates 1970 ~ 1993* [R], Moody's Investors Service, 11~13.
- [17] Gordy M. B., 2000, *A Comparative Anatomy of Credit Risk Models* [J], *Journal of Banking & Finance*, 24 (1~2), 119~149.
- [18] Gordy M. B., 2002, *Saddle Point Approximation of Creditrisk+* [J], *Journal of Banking & Finance*, 26 (7), 1335~1353.
- [19] Jarrow R., Turnbull S., 1995, *Pricing Derivatives on Financial Securities Subject to Credit Risk* [J], *Journal of Finance*, 50 (1), 53~86.
- [20] Jarrow R., Lando D., Turnbull S., 1997, *A Markov Model for the Term Structure of Credit Risk Spreads* [J], *Review of Financial Studies*, 10 (2), 481~523.
- [21] Jessen R., Lando D., 2015, *Robustness of Distance-to-Default* [J], *Journal of Banking & Finance*, 50 (1), 493~505.
- [22] Kaiasheva E. V., 2014, *Modeling Default Probability for Micro and Small Enterprises, Financial Analytics* [J], *Science and Experience*, 17 (2), 44~56.
- [23] Lando D., Nielsen M. S., 2013, *Additive Intensity Regression Models in Corporate Default Analysis* [J], *Journal of Financial Econometric*, 11 (3), 443~485.
- [24] Li X., Shang Y., Su Z., 2015, *Semiparametric Estimation of Default Probability: Evidence from the Prosper Online Credit Market* [J], *Economics Letters*, 127 (2), 54~57.

- [25] Mathis J., McAndrew J., Rochet J. C., 2009, *Rating the Raters: Are Reputation Concerns Powerful Enough to Discipline Rating Agencies?* [J], Journal of Monetary Economics, 56 (5), 657~674.
- [26] Silva S., Murtreira J., 2009, *Estimation of Default Probabilities Using Incomplete Contracts Data* [J], Journal of Empirical Finance, 16 (3), 457~465.
- [27] Shumway T., 2001, *Forecasting Bankruptcy More Accurately: A Simple Hazard Model* [J], Journal of Business, 74 (1), 101~124.
- [28] Tian S., Yan Y., Guo H., 2015, *Variable Selection and Corporate Bankruptcy Forecasts* [J], Journal of Banking & Finance, 52 (3), 89~100.
- [29] Tudela M., Young G., 2005, *A Merton-Model Approach to Assessing the Default Risk of UK Public Companies* [J], International Journal of Theoretical and Applied Finance, 8 (6), 737~761.
- [30] 陈诗一:《德国公司违约概率预测及其对我国信用风险管理的启示》[J],《金融研究》2008年第8期。
- [31] 管七海、冯宗宪:《信用违约概率测度研究:文献综述与比较》[J],《世界经济》2004年第1期。
- [32] 彭建刚、易宇、李樟飞:《测算商业银行贷款违约概率的贷款违约表法探讨》[J],《管理学报》2009年第6期。
- [33] 夏凡、姚志勇:《评级高估与低估:论国际信用评级机构“顺周期”行为》[J],《金融研究》2013年第2期。

A New Method of Measuring Default Probability with Complete Information in Credit Ratings

Zhang Jinbao

(International Business School, Beijing International Studies University)

Research Objectives: Present a new method of calculating default probability by making a full use of historical data. **Research Methods:** After analyzing relationship between default intensity and default probability, this paper provides a new way of estimating default probability. Specifically, this paper discusses the temporal relation between the beginning and ending time of loan and inspection horizon of the measuring PD, then proposes a characteristic function to construct a function for describing the time-varied number the loan, and presents a new formula of estimating default probability. **Research Findings:** By an empirical study, we find there is an underestimation of default probability in traditional methods. **Research Innovations:** Using this method, any loan appeared in the inspection horizon can be sampled from historical data and used for calculating default probability. Compared with the traditional methods, this method can make full use of the information implied in historical data and has more practical application value. **Research Value:** This paper gives a new way of calculating probability of default. It can be widely used in the practice of risk management of commercial banks.

Key Words: Default Probability; Default Intensity; Credit Rating

JEL Classification: G21; G24; G32

(责任编辑:陈星星)