审计理论前沿

SHENJILILUNQIANYAN

◆审计署审计科研所 编

2014

中国时代经济出版社

审计理论前沿

SHENJILILUNQIANYAN

◆审计署审计科研所 编

2014

中国时代经济出版社

图书在版编目（CIP）数据

审计理论前沿．2014／审计署审计科研所编．
—北京：中国时代经济出版社，2015.7
ISBN 978－7－5119－2414－8

Ⅰ.①审… Ⅱ.①审… Ⅲ.①审计理论—文集
Ⅳ.①F239.0－53

中国版本图书馆 CIP 数据核字（2015）第 138404 号

书　　名：审计理论前沿．2014
作　　者：审计署审计科研所

出版发行：中国时代经济出版社
社　　址：北京市丰台区玉林里 25 号楼
邮政编码：100069
发行热线：（010）63508271　63508273
传　　真：（010）63508274　63508284
网　　址：www. cmepub. com. cn
电子邮箱：zgsdjj@ hotmail. com
经　　销：各地新华书店
印　　刷：北京市昌平百善印刷厂
开　　本：787×1092　1/16
字　　数：223 千字
印　　张：17
版　　次：2015 年 7 月第 1 版
印　　次：2015 年 7 月第 1 次印刷
书　　号：ISBN 978－7－5119－2414－8
定　　价：50.00 元

前　言

2014年，审计署审计科研所在审计署党组的正确领导下，牢固树立科学审计理念，深入总结审计实践经验，积极探索审计发展规律，大力推进理论创新和实践创新，着力发挥理论引导、实践指导和决策参考的“智库”作用。按照审计署的统一部署，紧紧围绕审计工作重点开展审计理论研究，就审计的理论和现实问题进行探讨，力求突出针对性、前瞻性、建设性和有效性，着力服务审计实践，形成了一批研究成果，并以《审计研究报告》的形式刊出。为了广泛宣传审计署审计科研所的最新理论研究成果，推进国家审计理论创新，我们把当年编发的《审计研究报告》理论文章，按照刊发的先后顺序进行了汇编，结集出版《审计理论前沿2014》。

该书收录了2014年度审计署审计科研所《审计研究报告》编发的30篇理论文章，它基本反映了审计理论的最新研究，内容涵盖了审计监督与国家治理理论、审计制度完善发展理论、审计组织与管理理论、审计实务方法探讨、审计文化与审计作用理论等，现公开出版，供广大读者参考。

审计署审计科研所

2015年1月20日

目　录

基于数据挖掘的商业银行信用风险评估研究方法概述

世界银行对全球银行业危机的研究表明，导致银行破产的主要原因就是信用风险。随着我国金融改革的不断深入，我国金融机构的退出问题得到越来越多的讨论，对商业银行信用风险客观的量化评估变得日益重要。因此，商业银行信用风险也理应是国家金融审计需要重点关注的金融风险之一。本研究将对运用数据挖掘技术实现商业银行信用风险评估的方法进行梳理和概要性研究，主要分为商业银行信用风险基本概念及评估方法现状、基于数据挖掘的企业客户信用风险评估方法和基于数据挖掘的个人客户信用风险评估方法三个部分。

一、商业银行信用风险基本概念及评估方法现状

（一）风险与信用风险

1. 风险概念的二维性

风险定义可以分为两类：一是广义风险，强调结果的不确定性，是指在一定条件下和一定周期内发生的个中结果的变动程度，结果的变动程度越大则相应的风险也就越大，反之则越小；二是狭义风险，强调确定性带来的不利后果，是指一定条件下和一定周期内由于个中结果发生的不确定性，而导致行为主体遭受损失或损害的可能性。这个概念中突出强调了风险的危害性，是一种“损失的危害”，弱化了风险可能带来的效益的一面。无论何种定义，都可以看出风险是一个二维概念，它既涵盖了损失的大小，又涵盖了损失发生概率的大小。

2. 信用风险

对于信用风险概念的理解，最具代表性的观点有三种。第一种观点认为，信用风险指交易对象无力履约的风险，即债务人未能如期偿还其债务造成违约而给经济主体经营带来的风险。第二种观点认为，信用风险有广义和狭义之分。广义的信用风险指所有因客户违约所引起的风险，如资产业务中的借款人不按时还本付息引起的资产质量恶化；表外业务中的交易对手违约引致或有负债转化为表内负债等。狭义的信用风险通常专指信贷风险。第三种观点认为，信用风险指由于借款人或市场交易对手违约而导致损失的可能性，信用风险还包括由于借款人的信用评级的变动和履约能力的变化导致其债务的市场价值变动而引起损失的可能性。这种观点中信用风险的大小主要取决于交易对手的财务状况和风险状况。

（二）商业银行信用风险评估方法现状

信用风险是最古老也是最重要的风险之一。信用风险直接产生于客户的商业选择，通常它不受金融市场运动的影响，而是依赖于交易规则，包括资本市场交易规则和货币市场交易规则。它们对信用风险都会产生影响，信用风险到目前为止仍然难以测量，但是实际情况又要求对信用风险进行量化，例如数量管理关系的发展和信用风险密切相关，尤其是考虑到资本充足率，更要参考信用风险的大小。

管理信用风险可以通过定性和定量两个方面，通过评级机构的评级来测量信用风险等级属于定性分析，它主要依靠的是经验判断。JP Morgan 公司设计的 Credit-Metrics 模型、KMV 公司的 KMV 模型、瑞士银行金融产品开发部的 Credit-Risk + 模型和麦肯锡的 Credit Portfolio View 模型等，进行信用风险的测定属于定量分析。

1. 专家系统法

专家系统法是一种古老的信用风险评估方法，是商业银行在长期经营信贷业务、承担信用风险过程中逐步发展并完善起来的传统信用风险分析方法。专家系统法有：5C 分析法、5W 分析法、5P 分析法等，最常用的是 5C 分析法。5C 分析法主要是指道德品质（Character）、还款能力（Capacity）、资本实力（Capital）、担保（Collateral）和经营环境条件（Condition）

五个方面。

2. 贷款风险分类

贷款风险分类法又称贷款五级分类法，是对银行的贷款质量进行评价并对银行抵御贷款损失的能力进行评估的一套系统方法。其最核心的评价标准是考察贷款归还的可能性，并根据贷后管理的需要将贷款分为五类。

3. Z 评分模型（Z－score model）

Z 评分模型是美国纽约大学斯特商学院教授爱德华·阿尔特曼在前人研究的基础上，于 1968 年提出的。1977 年又对该模型进行了修正和扩展，建立了第二代模型 ZETA 模型。Z 评分模型一经推出，便引起了各界关注，许多金融机构纷纷采用它来预测信用风险，并取得了一定的成效，目前它已经成为西方国家信用风险度量的重要模型之一。阿尔特曼的 Z 评分模型是一种多变量的分辨模型，他是根据数理统计中的辨别分析技术，对银行过去的贷款案例进行统计分析，选择一部分最能反映借款人的财务状况，对贷款质量影响最大、最具预测或分析价值的比率，然后根据所选择的比率指标，设计出一个能最大程度区分贷款风险度的数学模型；最后，借助于模型，对贷款申请人的信用风险及资信情况进行评估、判别，并把贷款申请人划分为不同的信用类别。

4. 判别分析模型

商业银行可以根据已有的数据集中的数据，把其中的违约样本和非违约样本的数据进行归类，然后确定其违约的临界值。判别分析法是对研究对象所属类别进行判别的一种统计分析方法。它是按照一定准则建立一个或多个判别函数，用研究对象的数据资料确定判别函数中的待定系数，并计算判别临界值，从而确定样本所属的类别。从本质上看，判别分析法是一种线性回归。

5. 现代信用风险度量模型

世界著名的中介机构和金融机构向外公布的比较有影响力的信用风险度量模型主要有以下几种：JP 摩根的信用度量术模型（Credit-Metrics）、KMV 公司开发的 KMV 模型、瑞士银行金融产品开发部的信用风险附加模型（Credit Risk +）和麦肯锡公司的信用组合观点模型（Credit Portfolio

View)。其中信用度量术模型（Credit-Metrics）是基于在一段时间内一种信用等级向另一种信用等级（包括违约）迁移的概率模型。KMV 模型是基于 Merton（1974）的期权定价模型，即假设公司的资产满足一个动态变化的随机过程，并且当资产的价值低于某个违约阈值时违约发生，此模型依赖于公司的资本结构。信用风险附加模型（Credit Risk +）是一个保险精算模型，即假设公司债券的违约是外生的并且满足一个泊松过程（Poisson process)。信用组合观点模型（Credit Portfolio View）是一个离散时间多期计量模型，在这里违约概率是建立在多个宏观经济变量（例如：失业率、利率水平、经济增长率、汇率、政府支出）的条件之上。

（三）基于数据挖掘的商业银行信用风险评估方法的提出

专家系统法和贷款分类法属于定性分析方法，依赖于专家的经验及判断，理论依据不充分。多元判别分析采用企业历史财务指标，影响其信用评估时效性。现代信用风险计量模型具有较为严格的应用前提。如 KMV 模型需要大量的上市公司数据；Credit-Metrics 模型依赖于评级公司提供的信用等级以及国家和行业长期的历史数据。Credit Portfolio View 模型考虑宏观经济因素对信用等级转移的影响，而宏观经济因素的个数、各因素的经济含义及它们与信用级别转移的具体函数关系很难确定和检验。

数据挖掘技术为商业银行信用风险评估提供了一种新的方法，该方法没有现代信用风险度量模型苛刻的前提条件，克服了专家系统法和贷款分类法的主观随意性，也没有多元判别分析对自变量与因变量之间成线性关系的假设前提。如何从海量数据中寻找对信用风险分析有用的信息，解决数据丰富但是信息匮乏的局面，成为商业银行信用风险管理中的关键。而在这样存有大量信息的数据仓库中挖掘隐含的知识正是数据挖掘技术的专长，它可以从中客观地挖掘评估规则，然后将这些规则存入专家系统的知识库，从而为决策提供依据，而且这样的过程是自动的，根据这些规则产生的评估结果会更客观、更准确。

二、基于数据挖掘的企业客户信用风险评估方法

企业信贷是商业银行的主要业务，也是重要利润来源之一，企业客户

信用风险评估是商业银行的一个重要风险管理问题，也是系统性金融风险分析中需要关注的重要问题。商业银行储存有大量的事实数据，这些数据涵盖了企业客户的基本信息、管理机制、重大事项、债务结构等诸多数据信息，这些信息对信用风险评估分析来说十分重要。

（一）用于数据挖掘的企业客户信用风险评估指标体系构建

1. 企业客户信用风险评估指标体系建立原则

总的来说，基于我国商业银行企业客户信用风险管理的现状，综合分析国内外研究结论，参考国外权威指标体系和我国商业银行现有的信用风险评估要素，选取评价指标时应遵循以下原则：

一是全面性原则。评估指标体系必须既能够全面地反映企业客户目前的信用综合水平，又要包括企业发展前景的各方面指标；既能代表过去的业绩，又能预测未来的发展趋势。二是科学性原则。评价指标体系的大小也必须适宜，亦即指标体系的设置应有一定的科学性。如果指标体系过大，指标层次过多、指标过细，势必将评价者的注意力吸引到细小的问题上；而指标体系过小，指标层次过少、指标过粗，势必不能充分反映信用风险状况。三是可行性原则。指标数据的来源应易于采集，表达方式简单易懂，而且数据真实、准确。指标数据不能收集或只能得到部分数据，指标体系就失去了其应有的作用，商业银行信用风险评估本身就变得毫无意义。因此，数据统计上的可行性与准确性是不容忽视的重要标准。四是代表性原则。认真分析商业银行经营实践，力求所选指标能反映出银行的真实运行机制。

2. 企业客户信用风险评估指标体系构建分析

企业的经营状况是影响其信用风险状况评估的根本因素。一般来说，企业的经营状况与其盈利能力、经营能力、发展能力、偿债能力、现金流转能力等因素密切相关，是一个统一的整体。因而银行作为企业的债权人，为了准确测度和评价贷款企业的信用风险状况，应在信用风险分析中对这些内容进行综合分析。银行信用风险评估指标体系的设计将从这几类入手：

（1）盈利能力

公司经营的主要目的在于使投资资金获得较高的利润与维持公司的适

度增长，只有盈利才能使经营与规模不断成长与发展。盈利能力是指上市公司赚取利润的能力，也是上市公司财务结构和经营绩效的综合体现。上市公司的盈利能力大，它在获取贷款、增资扩容等方面就有优势；其盈利能力小，在股市中就难有良好的表现。其主要衡量指标有净资产收益率、总资产报酬率、销售净利率和成本费用利润率。

（2）经营能力

企业的经营能力是指企业利用现有资源创造社会财富的能力，具体指各项营运资产的周转效率，它直接影响到企业的盈利能力和偿债能力，体现了企业的经营绩效。其主要衡量指标有总资产周转率、应收账款周转率、存货周转率、流动资产周转率。

（3）偿债能力

偿债能力是指在一定期间内清偿各种到期债务的能力。对于大多企业样本的选取和指标变量的数据分析而言，资金来源除了股东权益以外，还有一部分来自对外负债。偿债能力的强弱是衡量上市公司经营绩效的重要指标，它直接关系到债权人、投资者的切身利益。偿债能力的大小将直接影响到公司的营运能力和净资产变现能力。其主要衡量指标有流动比率、速动比率、超速动比率、资产负债率、负债与所有者权益比率和负债与有形资产比率。

（4）成长能力

成长能力是指上市公司通过自身的生产经营活动不断发展的能力，成长能力反映了上市公司的发展潜力。其主要衡量指标有主营业务收入增长率、净资产增长率和总资产增长率。

（5）现金流量

英美等国国际会计准则委员会等组织的概念结构公告中，认为现金流量指标能较好地反映上市公司资金的流动性和对债务的偿还能力。其主要衡量指标有现金流动负债比。

（6）其他非财务因素分析

商业银行分析贷款企业的信用风险，除了考虑财务因素外，还要分析企业客户的其他非财务因素。这些因素也对信用风险的形成产生重要影响，如管理者素质、经营管理水平、内部控制等。

根据上述商业银行企业信用风险相关因素构建指标体系，并根据指标体系采集相应数据，作为数据挖掘的原始数据集。

（二）企业客户信用风险评估数据挖掘算法

属性约简是数据挖掘领域的一个难题，属性约简的过程本质上是进行“降维”的过程，降低向量维度可以极大地增加样本的密度。

1. 属性约简的重要意义及典型算法

（1）属性约简在数据挖掘分析中的意义

信息系统数据库的信息增加有两个方向：横向和纵向。横向指的是属性数目的不断增加，纵向指的是记录数的增加。对信息系统横向属性数目的约简称之为属性约简或称属性选择。

在数据分析中有两类属性是不必要的，一类是与目标内容不相关的属性，另一类是对于其他属性来说冗余的属性。在实际应用中，这两类不必要的属性可能同时存在，但是由于属性之间的交互作用冗余属性更难于剔除。为了把这两类不必要的属性减到最少，我们需要属性约简。属性约简的目的在于辨别属性的重要程度，删除与学习任务不相关的或不必要的属性，并建立优化的学习模型。属性约简能显著降低归纳算法运行时间，并提高结果模型的准确度。

（2）前馈神经网络的属性映射算法

前馈神经网络可以用来属性约简，具有一个隐藏层的多层感知器被用于属性特征提取。其基本思想是使用隐藏单元作为新提取的特征。让我们考察它如何处理属性约简的三个主要问题。第一个问题是如何评估新属性的性能。估计预测准确率，并用它作为性能度量。这要求应该标记数据所属的类。我们选择会使预测准确率达到最佳的属性集。第二个问题是将原特征映射到新特征。在这种情况下，它是从输入单元到隐藏单元的非线性映射。第三个问题是如何确定新特征个数。显然，后两个问题与神经网络的拓扑结构密切相关。设计两个算法用于构造具有最少隐藏单元（最少属性），并且输入层和隐藏层之间具有最少连接的网络：以提高预测准确率为目标，网络构造算法极度严格地添加每一个隐藏单元；在不影响预测准确率的情况下，网络剪枝算法会剪去输入层和隐藏层之间的冗余连接。

（3）小波变换

离散小波变换是一种线性信号处理技术，用于数据向量 X 时，将它变换成不同的数值小波系数向量 X'。两个向量具有相同的长度。当这种技术用于数据规约时，每个元组看作一个 n 维数据向量，即 $X=(x_1,x_2,\cdots,x_n)$，描述 n 个数据库属性在元组上的 n 个测量值。

小波变换后的数据可以截短，仅存放小部分最强的小波系数，就能保留近似的压缩数据。这样结果表示数据非常稀疏，使得如果在小波空间进行计算的话，利用数据稀疏特点的操作计算得非常快。该技术也能用于消除噪声，而不会光滑掉数据的主要特征，使得它们也能有效地用于数据清理。给定一组系数，使用所用的离散小波变换的逆，可以构造原始数据的近似。

流行的小波变换包括 Haar－2、Daubechies－4 和 Daubechies－6。离散小波变换的一般过程使用一种层次金字塔算法（Pyramid algorithm），它在每次迭代时将数据减半，导致计算速度很快。该方法如下：

①输入数据向量的长度 L 必须是 2 的整数幂。必要时，通过在数据向量后添加 0，这一条件可以满足（$L\geqslant n$）。

②每个变换涉及应用两个函数。第一个使用某种数据光滑，如求和或加权平均。第二个进行加权差分，提取数据的细节特征。

③两个函数作用于 X 中的数据点对，即作用于所有的测量对（x_{2i}，x_{2i+1}）。这导致两个长度为 $L/2$ 的数据集。一般而言，它们分别代表输入数据的光滑后的版本或低频版本和它的高频内容。

④两个函数递归作用于前面循环得到的数据集，直到得到的结果数据集的长度为 2。

⑤由以上迭代得到的数据集中选择的值被指定为数据变换的小波系数。

2. 企业客户信用风险评估数据挖掘典型分类算法

分类是一种重要的数据分析形式，它提取刻画重要数据类的模型。这种模型称为分类器，预测分类的类标号。例如，我们可以构建一个模型，把银行贷款的等级分为正常和非正常等。这种分析形式可以帮助我们更好地全面理解所要分析的商业银行相关数据。

（1）朴素贝叶斯分类方法

①朴素贝叶斯分类法的提出

贝叶斯分类法是统计学分类方法。它们可以预测类隶属关系的概率，如一个给定元组属于一个特定类的概率。贝叶斯分类基于贝叶斯定理。对分类算法的比较发现，一种称为朴素贝叶斯分类法的简单贝叶斯分类法可以与决策树和经过挑选的神经网络分类器媲美。用于大型数据库，贝叶斯分类法也表现出高准确率和高速度。朴素贝叶斯分类法假定一个属性值在给定类上的影响独立于其他属性的值。这一假定称为类条件独立性。做此假定是为了简化计算，并在此意义下称为“朴素的”。

②朴素贝叶斯分类法的基本原理

a. 设 D 是训练元组和它们相关联的类标号的集合。通常，每个元组用一个 n 维属性向量 $X = \{x_1, x_2, \cdots, x_n\}$ 表示，描述由 n 个属性 $A_1, A_2, \cdots, A_n$ 对元组的 n 个测量。

b. 假定有 m 个类 $C_1, C_2, \cdots, C_m$。给定元组 X，分类法将预测 X 属于具有最高后验概率的类（在条件 X 下）。也就是说，朴素贝叶斯分类预测 X 属于类 C_i，当且仅当

$$P(C_i/X) > P(C_j/X) \qquad 1 \leqslant j \leqslant m, j \neq i$$

这样，最大化 $P(C_i/X)$。$P(C_i/X)$ 最大的类 C_i 称为最大后验假设。根据贝叶斯定理：

$$P(C_i/X) = \frac{P(X/C_i)P(C_i)}{P(X)}$$

c. 由于 $P(X)$ 对所有类为常数，所以只需要 $P(X/C_i)P(C_i)$ 最大即可。如果类的先验概率未知，则通常假定这些类是等概率的，即 $P(C_1) = P(C_2) = \cdots = P(C_m)$，并据此对 $P(C_i/X)$ 最大化。否则，最大化 $P(X/C_i)P(C_i)$。类先验概率可以用 $P(C_i) = |C_{i,D}|/|D|$ 估计，其中 $|C_{i,D}|$ 是 D 中 C_i 类的训练元组数。

d. 给定具有许多属性的数据集，计算 $P(C_i/X)$ 的开销可能非常大。为了降低计算 $P(C_i/X)$ 的开销，可以做类条件独立的朴素假定。给定元组的类标号，假定属性值有条件地互相独立（即属性之间不存在依赖关系）。因此，

$$P(X/C_i) = \prod_{k=1}^{n} P(x_k/C_i) = P(x_1/C_i)P(x_2/C_i)\cdots P(x_n/C_i)$$

可以很容易地由训练元组估计概率 $P(x_1/C_i)$，$P(x_2/C_i)$，…，$P(x_n/C_i)$。注意，x_k 表示元组 X 在属性 A_k 的值。对于每个属性，考察该属性是分类的还是连续值的。

e. 为了预测 X 的类标号，对每个类 C_i，计算 $P(X/C_i)P(C_i)$。该分类法预测输入元组 X 的类为 C_i，当且仅当

$$P(X/C_i)P(C_i) > P(X/C_j)P(C_j) \qquad 1 \leqslant j \leqslant m, j \neq i$$

换言之，被预测的类标号是使 $P(X/C_i)P(C_i)$ 最大的类 C_i。

(2) k - 最近邻分类算法

k - 最近邻分类算法是20世纪50年代早期首次引进的。当给定大量数据集时，该方法是计算密集的，直到20世纪60年代计算能力大大增强之后才流行起来。此后它被广泛用于模式识别领域。

最近邻分类法是基于类比学习，即通过将给定的检验元组与和它相似的训练元组进行比较来学习。训练元组用 n 个属性描述。每个元组代表 n 维空间的一个点。这样，所有的训练元组都存放在 n 维模式空间中。当给定一个未知元组时，k - 最近邻分类算法搜索模式空间，找出最接近未知元组的 k 个训练元组。这 k 个训练元组是未知元组的 k 个“最近邻”。

“邻近性”用距离度量，如欧几里德距离。两个点或元组 $X_1 = (x_{11}, x_{12}, \cdots, x_{1n})$ 和 $X_2 = (x_{21}, x_{22}, \cdots, x_{2n})$ 的欧几里德距离是：

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2}$$

换言之，对于每个数值属性，我们取元组 X_1 和 X_2 该属性对应值的差，取差的平方和，并取其平方根。在使用上述公式之前，我们把每个属性的值规范化。这有助于防止具有较大初始值域的属性比具有较小初始值域的属性的权重过大。例如，可以通过计算下式，使用最小——最大规范化把数值属性 A 的值 v 变换到 [0，1] 区间中的 v'

$$v' = \frac{v - \min_A}{\max_A - \min_A}$$

其中，$\min_A$ 和 $\max_A$ 分别是属性 A 的最小值和最大值。

对于 k - 最近邻分类，未知元组被指派到它的 k 个最近邻中的多数类。当

$k = 1$ 时，未知元组被指派到模式空间中最接近它的训练元组所在的类。最近邻分类也可以用于数值预测，即返回给定未知元组的实数值预测。在这种情况下，分类器返回未知元组的 k 个最近邻的实数标号的平均值。

可以通过实验来确定近邻数 k 的值。从 $k = 1$ 开始，使用检验集估计分离器的错误率。重复该过程，每次 k 增值1，允许增加一个近邻。可以选取产生最小错误率的 k。一般而言，训练元组越多，k 的值越大（使得分类和数值预测决策可以基于存储元组的较大比例）。随着训练元组数趋向于无穷并且 $k = 1$，错误率不会超过贝叶斯错误率的2倍。如果 k 也趋向于无穷，则错误率趋向于贝叶斯错误率。

最近邻分类法使用基于距离的比较，本质上赋予每个属性相等的权重。因此，当数据存在噪声或不相关属性时，它们的准确率可能受到影响。然而，这种方法已经被改进，结合属性加权和噪声数据元组的剪枝。距离度量的选择可能是至关重要的。也可以使用曼哈顿距离或其他距离度量。

三、基于数据挖掘的个人客户信用风险评估方法

（一）基于数据挖掘的个人客户信用风险评估的指标选取

可以根据“典型申请指标”选取以下九个指标进行实验：逾期期数、性别、现有房性质、学历、月均收入、婚姻状况、行业、年龄、职业。其中，逾期期数是用于对贷款客户进行分类的，为决策属性或分类属性，而其余八个属性主要描述了贷款客户的相关特征，为条件属性或特征属性。

根据商业银行的个人住房按揭贷款数据记录表，上述指标的数据类型如下表所示：

字段名称	数据类型
ID	int
逾期期数	int
性别	text