

第三届“ScienceWord 杯”数学中国

数学建模网络挑战赛

承 诺 书

我们仔细阅读了第三届“ScienceWord 杯”数学中国数学建模网络挑战赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们允许数学中国网站(www.madio.net)公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛报名号为：

参赛队员（签名）：

队员 1：刘硕

队员 2：袁学玲

队员 3：王学霞

参赛队教练员（签名）：

参赛队伍组别：大学生组

报名号 # 1110

第三届“ScienceWord 杯”数学中国

数学建模网络挑战赛

编 号 专 用 页

参赛队伍的参赛号码：（请各个参赛队提前填写好）：
1110

竞赛统一编号（由竞赛组委会送至评委团前编号）：

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

报名号 # 1110

2010 年第三届“ScienceWord 杯”数学中国 数学建模网络挑战赛

题 目 高校图书馆智能服务

关 键 词 聚类分析 图形统析 图书超期

摘 要：

本文将问题一中的问题转化为合理分类及数据处理的问题。问题二是评价和设计借阅规则的问题。

问题一：

首先我们分别应用聚类分析法中的 K-means 算法以及图形统计分析两种方法将超期天数分为三类，经对比，两种方法得到的结果基本吻合，分类如下：第 I 类——60 天以内；第 II 类——60 天至四年；第 III 类——大于四年。建立图书超期统析模型，利用 EXCEL 软件和 Visio 软件，将各类数据处理，从连续图、离散图、和综合图，对各类数据进行横纵向全面的比较，得出各类数据的特点以及彼此间的联系。并对超期书籍分类，根据数据处理分析，结论如下：第 III 类书籍需求特殊，以自然科学类为主。对于第 I、II 类，自然科学类和社会科学类丛书比例为 1.23: 1，可见自然科学类书籍比较容易超期，另外，经随机统计，社科类的小说及艺术类休闲系列丛书比例为 35%。

超期原因主要有如下四种：1、过失超期；2、意外事故；3、故意超期；4、图书馆管理制度不严或个人意识不高。

问题二：

评价：

通过模型中得到的结论，对现有借阅规则进行评价。

借阅期限：30 天对于理工科的专业用书比较合理，但针对那些科研用书及热门图书不是很合理。

续借期限：超期在 40 天内的概率占第一类数据的 80%左右，所以 15 天期限过短。

罚款：综合经济因素与超期频率考虑，可知罚款设置如此简单，0.1 元不合理。

设计：

借阅期限：教师等另行考虑；学生按对图书的需求设置不同期限，另加人性化管理。

续借期限：设置一定的条件，续借只可在第一次借期前 7 天办理，期限 40 天。

罚款：金额先高后低，提高大家的意识，并缓解严重超期带给读者的经济压力。

参赛队号 1110

所选题目 C 题

参赛密码 _____
(由组委会填写)

高校图书馆智能服务

1 问题的背景与重述

1.1 问题背景

随着知识更新速度的加快，用户对信息资源的需求量越来越大，图书馆作为象牙塔中的一座必不可少的知识宝库，高校师生用户对图书馆中各类信息资源的利用率和需求值越来越高，超期借阅现象也随之而来。如今，读者借阅超期现象十分普遍，尤其是图书馆实现自动化管理以来，逾期涉及自动计算罚金，很多读者不能严格遵守图书借阅规章制度，借阅图书馆的书刊后，不及时归还，甚至长期占有，严重影响了图书的流通，降低了文献资源的周转利用率，同时也大大影响了图书馆整体的服务形象和工作效率。为加强书刊文献的有序流动，保障图书馆的正常运行，保护读者的合法权益，就必须解决读者借书逾期不还这一因素。

如今，各图书馆都制定了图书超期罚款制度，制约读者，可是实行超期罚款后，带来了各种各样的矛盾和问题。这些矛盾和问题如果处理不当，会影响图书馆对外服务形象，增加了馆员与读者之间的矛盾。在构建和谐图书馆中产生消极影响，特别是影响读者对图书馆的信任度和利用率。

因此，图书馆必须深刻自省发现超期现象的缘由，减少超期罚款这个倍受关注的话题给其所带来带来的压力，并采取一定措施，督促读者按期还书，从而充分发挥馆藏文献的作用^[1]。

1.2 问题重述

图书馆源于保存记事的习惯。图书馆是为读者在馆内使用文献而提供的专门场所。而高校的图书馆为教学和科研服务，具有服务性和学术性强的特点。

现在的高校图书馆存在着许多不良的现象，比如对书籍的损坏、借阅超期等。在搜索引擎下输入“图书超期”等关键字，成千上万的网页被检索出，大多都是超期读者的信息列表。超期现象是读者借阅图书资料时超过规定借阅期限后不归还的一种普遍现象。处理超期罚款是流通服务中常遇到的事，特殊情况则需酌情处理。这一现象虽是小事，但处理不当会影响图书馆与读者之间的关系，从而影响图书馆的声誉，以及馆藏资源的利用率，最终影响图书馆功能的正常发挥。随着信息化的发展，图书的预约服务也逐渐普及。图书的预约与借阅超期的现象也有很大的内在联系。试通过数学模型来研究以下的问题：

(1) 对于借阅图书超期，原因众多。考虑到高校的特点，大体上可以把超期原因分为三类：1、无意错过还书期限；2、由于教学科研需要，长期使用；3、由于罚款数目较少，所以故意拖延。请通过给出的数据或自行查找相关数据，建立合理的模型，分析哪些书容易发生超期现象，以及当发生超期时，分析其可能的原因，以便图书馆采取相应的措施改进服务。

(2) 假设某高校的借书规则为：借书时限为 1 个月，可以续借一次，续借时限为 15 天，超期一天罚款 0.1 元。试根据问题(1)中的模型评价这种借书规则的优劣。请设计

一种更加合理的借阅规则。

2 问题分析

2.1 问题分析

问题一分析：

根据题目中的提示，图书超期现象发生的原因主要有三个，建立模型，用合适的算法，将超期天数分为三类，然后通过作图对数据进行统计，根据数据图形特征，对数据进行合理的分析。在对数据进行分析之后，可以得出图书借阅超期的原因。而图书馆借阅的图书进行分类，总体上可分为两大类：自然科学类和社会科学类，根据国家标准，图书又被分为22类，根据对所有书籍分类的统计，依据数据特征可以分析出超期天数出现最多的一类或几类书。然后从自己的经验，以及对图书馆的调查或了解，分析出该高校此几类书超期现象严重的原因。

问题二分析：

在问题一所建立的模型和数据统计分析的基础上，依据合适的评价方法对该高校的借阅规则进行分析与评价，并针对该借阅规则中不合理的地方，提出相应的改进措施。而后根据自己所了解的知识，以及对该高校图书馆借阅规则的分析，提出一种更合理、更智能化、更人性化的借阅规则。在对该高校的分析基础之上，还可以给高校图书馆领导写一封信，向他表达一些经分析得到的关于图书馆借阅规则的好建议。

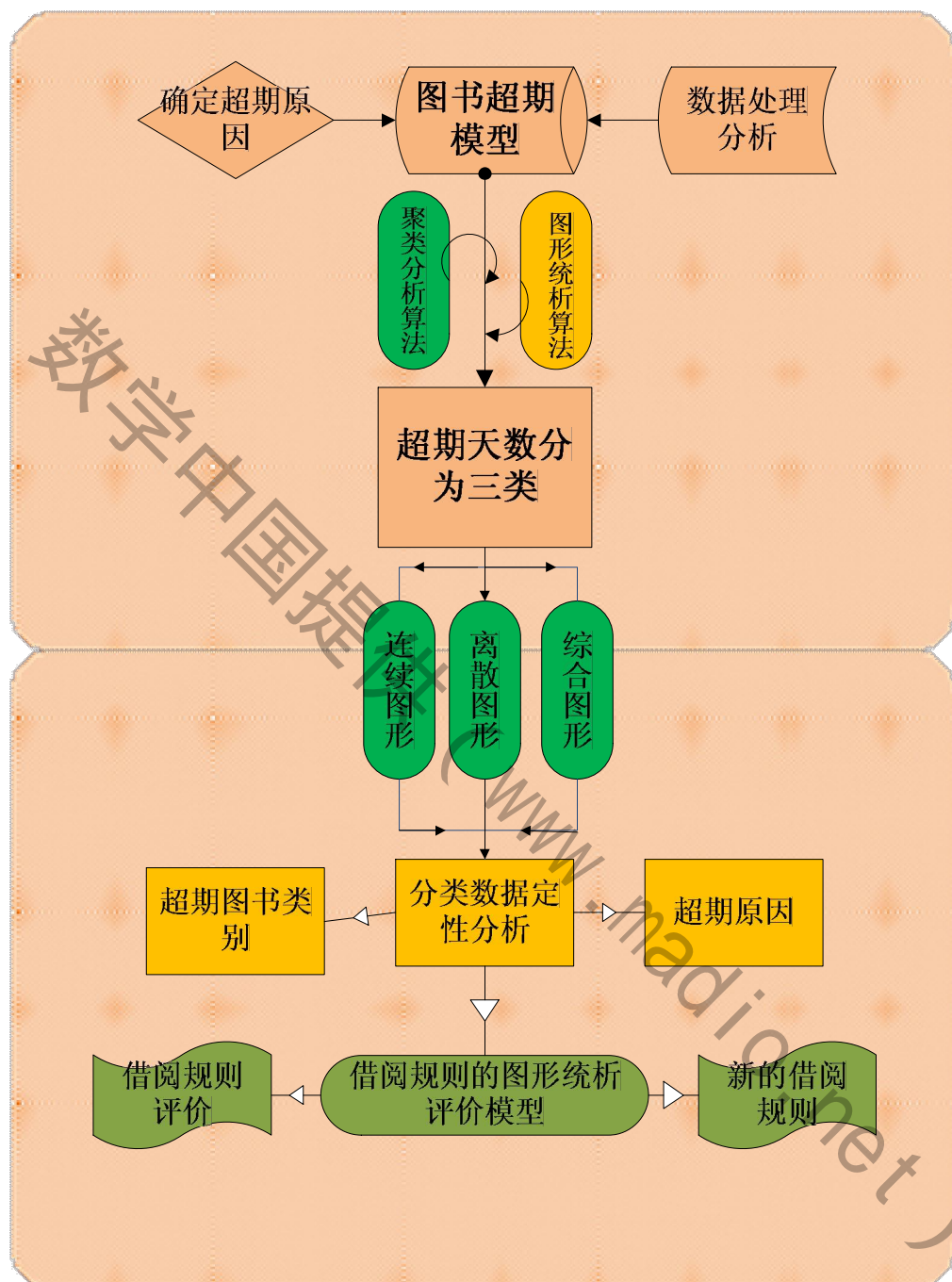
3 模型假设与建模思路图

3.1 模型假设

- 1、假设图书馆的计算机工作运转状况良好，超期天数统计无任何差错。
- 2、假设各个学校图书馆针对即将离校毕业生均设立“超期图书结账”制度，毕业生结帐后超期天数不再继续统计。
- 3、假设此数据来源于真实学校，且此学校的图书管理制度具有代表性。
- 4、假设此图书馆严格按照借书条例运行，对因超期天数少或因意外原因不能及时还书者给与的宽限制度不予考虑。
- 5、假设超期频率在前两个月内基本相同，用第二个月内的数据可以基本反映前两个月内总体的超期分布密度。
- 6、假设各学校借书卡形式统一，均为一卡通。
- 7、忽略图书馆中计算机智能化管理所要求的技术限制。

3.2 建模思路图

下面的思路流程图是我们文章结构的一个缩影,它将完整而形象的反应我们文章的建模思路。



4 符号系统

为了便于描述问题,我们用一些符号来代替问题中涉及的一些基本变量,如表一:符号说明一览表所示,其它一些变量将在文中陆续说明。

表一: 主要符号说明一览表

符号意义	字母表示	单位
借阅超期天数	T	天
借阅人次	V	无量纲
超期书次数	M	无量纲
读者号	W	无量纲
超期书的数目	S	无量纲

5 图书超期统析模型的建立与分析

5.1 确定图书超期的基本原因

1) 过失超期

过失超期是读者不知道图书借阅到期，因而未能按时还书所造成的，即主观上不希望超期。过失超期的情形主要有：（1）过于自信的过失，即读者实际不知道或不记得借阅到期的准确日期，因过于自信自以为是而导致超期；（2）疏忽大意的过失，即读者因忙于学习或工作等种种原因完全忘记借阅到期，甚至忘记自己曾借过图书而导致超期。

[2]

2) 故意超期

故意超期，就是读者明知图书借阅到期并且能够及时办理还书手续，而不予办理相关手续的行为，即明知故犯。故意超期的情形主要有：（1）借期比较短，学生读者只有一个月，对有些专业书来说一个月时间根本来不及看完，图书借阅到期时，图书馆无库存复本，且按规定不能续借，而读者因某种急需，如考研、考级等，还需继续借阅热门藏书，于是通过违章继续占用到期图书，即“明知故犯”；（2）有的读者认为，每天0.1元的超期罚款算不了什么，超期了无非是交点钱，因而未把按时主动还书当回事，导致借阅超期；（3）有的图书很紧俏，市场上不易买到，有的读者出于种种原因，为达到长期独占目的，在故意超期后谎报丢失，即使交超期罚款和赔书款也自认为值得。

3) 意外事件

意外事件是指既非故意，又非过失而导致超期的情形，主要有：（1）所借的图书被盗、遗失或借给别人而无法追还等。（2）由于教学科研需要，长期使用，有的老师借了一本或几本书，一直不归还，直至退休或者离校才到图书馆办理还书手续或者赔偿手续；有的老师借阅的专业书，他因为需要经常查阅，一直到不再需要该书才到图书馆办理还书手续。^[3]

5.2 模型的建立

为了确定超期天数与借阅的书的种类、超期原因等之间的关系，需要从所给的已知数据中找出超期天数分布的规律，从而对其进行分析。从所给数据，可以看出数据量很

大，规律不是很明显，因而需要将数据统计、分析。下面我们就超期天数和超期原因、超期的书的种类因素建立模型，分析超期天数与各个影响因素之间的关系，得出经常出现超期的原因。

5.3 模型的求解方法

图书超期现象发生的原因主要有三个，在建立模型的基础上，我们要首先对所给的大量数据进行分类，再依据分类进行分析。对于上述模型，我们用聚类分析算法和图形统析两种算法来求解。

5.3.1 聚类分析算法分类

5.3.1.1 算法描述

聚类分析（Cluster Analysis）又称群分析，是根据“物以类聚”的道理，对样品或指标进行分类的一种多元统计分析方法，它们讨论的对象是大量的样品，要求能合理地按各自的特性来进行合理的分类，没有任何模式可供参考或依循，即是在没有先验知识的情况下进行的。聚类分析起源于分类学，在古老的分类学中，人们主要依靠经验和专业知识来实现分类，很少利用数学工具进行定量的分类。随着人类科学技术的发展，对分类的要求越来越高，以致有时仅凭经验和专业知识难以确切地进行分类，于是人们逐渐地把数学工具引用到了分类学中，形成了数值分类学，之后又将多元分析的技术引入到数值分类学形成了聚类分析。

聚类分析是根据事物本身的特性研究个体的一种方法，目的在于将相似的事物归类。它的原则是同一类中的个体有较大的相似性，不同类的个体差异性很大。这种方法有三个特征：

（1）适用于没有先验知识的分类。如果没有这些事先的经验或一些国际标准、国内标准、行业标准，分类便会显得随意和主观。这时只要设定比较完善的分类变量，就可以通过聚类分析法得到较为科学合理的类别；

可以处理多个变量决定的分类。例如，要根据消费者购买量的大小进行分类比较容易，但如果在进行数据挖掘时，要求根据消费者的购买量、家庭收入、家庭支出、年龄等多个指标进行分类通常比较复杂，而聚类分析法可以解决这类问题；

（3）聚类分析法是一种探索性分析方法，能够分析事物的内在特点和规律，并根据相似性原则对事物进行分组，是数据挖掘中常用的一种技术。

5.3.1.2 算法计算方法过程

系统聚类法是目前聚类分析诸方法中用得最多、最为广泛接受的聚类方法，故采用系统聚类方法。但是由于用系统聚类法对 n 个样品： $(x_1, x_2, x_3, \dots, x_n)$ 进行聚类时，若 n 较大，距离矩阵 D 的计算量就大。另外，用系统聚类法聚类时，样品一旦归到某个类以后就不变了，但是如果换一种计算距离的方法，聚类结果可能发生变化。为了减少计算量采用了快速聚类法，其基本思想是：在待聚类的样品比较多时，先给出一个大致初始分类，然后用某种原则进行修改，直到分类结果比较合理为止。

其聚类过程大致可以由图 1 表示^[4]：

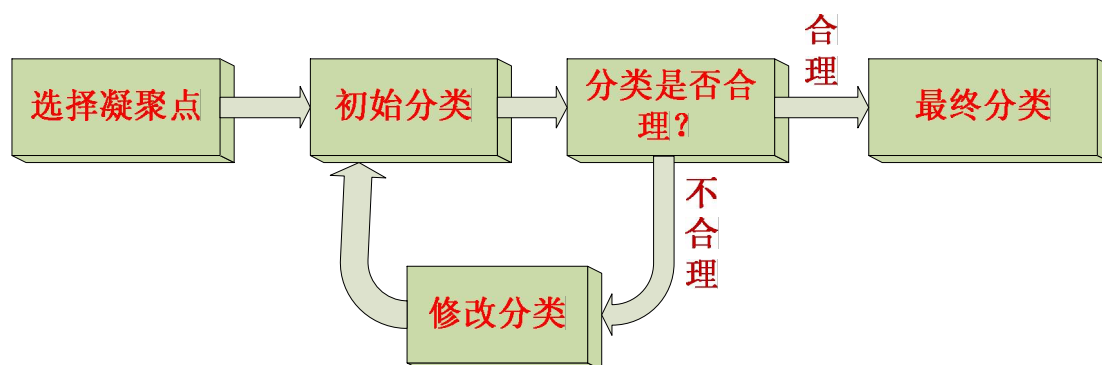


图 1：快速聚类法过程

为了得到初始的分类, 先设一些凝聚点, 再让样品按某种原则向凝聚点凝聚, 根据相关原理选择凝聚点^[5]。修改分类的原则选择最简单的逐个修改法, 即为下面用的 K-means 方法, 其具体实现过程如下^[6]:

(1) 首先, 随机地选择 k 个对象, 每个对象初始地代表了一个簇的平均值或中心。对剩余的每个对象, 根据其与其各个簇中心的距离将它赋给最近的簇。即初始化中心 C_i 然后

将输入样本 $X(p)$ 按最近邻规则分组, 即将 $\{X_{(1)}, X_{(2)}, \dots, X_{(p)}\}$ 分配给中心为 C_i

$(i=1, 2, 3 \dots k)$ 的输入样本聚类集合 $S_i (i=1, 2, 3 \dots k)$ 要求满足:

$$d_j = \min P_{X^{(p)} - C_i} (i=1, 2, 3 \dots k; j=1, 2, 3 \dots n)$$

式中: d_j 为剩余对象与簇中心的距离。

(2) 然后重新计算每个簇的平均值。这个过程不断重复, 直到准则函数收敛。根据样本 S_i 的平均值调整中心 C_i 。重复以上步骤, 直到中心的分布不再变化。

5.3.1.3 计算结果

采用目前国际上应用广泛的多元统计分析软件 Matlab^[7], 应用软件中的聚类分析功能, 找到聚类的中心, 并做出最后结果, 最终结果如表二所示:

表二：聚类中心

借阅超期分类	聚类中心
I类	45.2
II类	803.6
III类	3031.7

5.3.2 图形统析算法

5.3.2.1 算法介绍

杜威曾说过: 大学图书馆是一个大学的心脏, 由于其普及性不难分析出本题在所给

数据提供解题切入点的同时，经验和常识也为本题的分析提供了一定的方向^[8]。鉴于此，我们在理论分析的基础上结合实际，通过对大量数据的处理、统计和分析，再通过经验以及自身对自己所在学校图书馆工作人员和超期读者的调查，对借阅超期天数分为三类。

5.3.2.2 数据分类

首先，我们对超期天数的概况，通过直观的图形进行粗略估算；然后结合实际和常理，对借阅情况分为三类。其主要依据有两个——图形直观分析和实际经验。

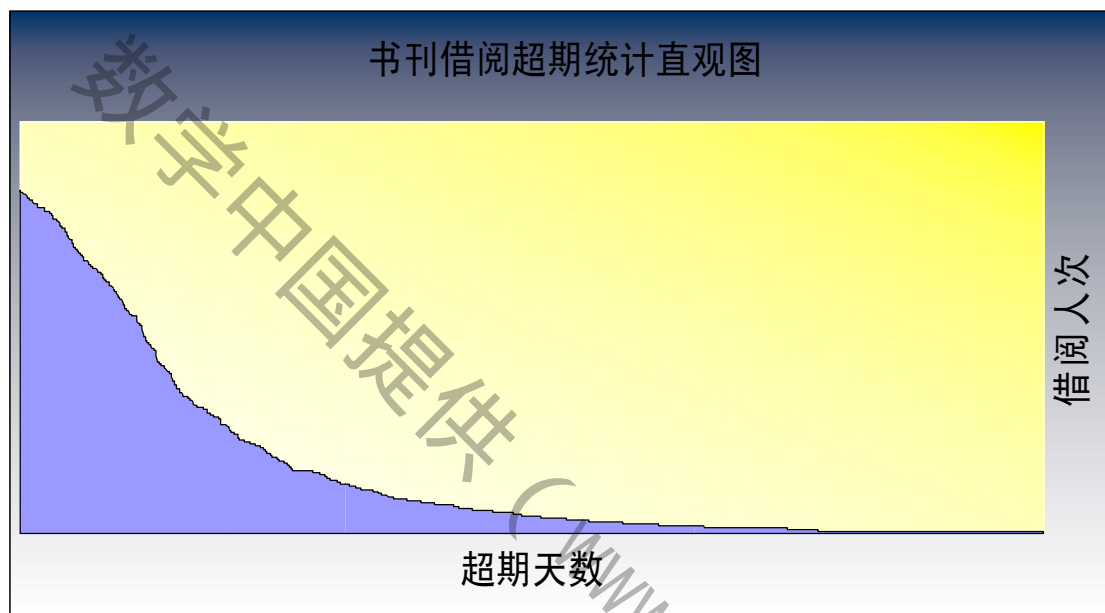


图 2：书刊借阅超期统计直观图

图2为书刊借阅超期统计直观图，横轴表示按升序排列的超期天数 T ，纵轴表示与超期天数对应的借阅人次 V 。从图中蓝色阴影面积中可以看出，随着超期天数的增加，借阅人次累计越少。由此可粗略推断：短期超期人数占的比例较大，长期超期的人数，尤其是超过统计数据总数的 $2/3$ 时，变得更小。此图依然可以通过积分面积的方法得出不同超期阶段的人数密度。由此，能更加精确的表现出两者的对应关系。考虑到非线性的影响，可以避免因为代数平均值所带来的盲目分类。另外，各类之间不能存在过大差异，否则会导致在比较分析中所带来的较大误差。由此图的曲线走向大致可以对分类的间距作出几乎合理的判断。

联系校园生活，读者大体分为学生和教员两大类。学生根据学校不同也可以分为不同的几类，而且所占比例也有很大的差异。但综合而言，绝大多数学校以本科生为主（专门针对本科学校而言）。由于本科学校的普及性，对图书馆的制度更能准确的给与表达，所以我们针对对本科学校的借书情况的分析将超期现象分为三大类。

第I类：在现有的数据下，将超期天数 T 界定在图书到期之日开始至60天。

在对本校图书馆老师的专访中获知的图书馆书刊借阅超期的大致状况，证实了经验带给我们的判断：违规者的超期天数大多处在两个月内，而45天前尤为普遍。由于此次缺少一定的数据量，我们只在现有数据上作出定性判断。最终分析结果与曲线图的积分

面积可以得到大致的吻合，这进一步验证了我们判断的合理性。

第II类：超期天数 T 界定在60天至1460天（即4年）。

此类判断主要从读者来考虑，由于学生在读者中所占的比例相当，因此以其作为主要参考依据。对于学生而言，借阅超期两个月，单单针对图书馆的借阅管理条例制度而言，已属违规比较严重的一类，这与意外事件（如忘记及时归还或因事情而意外耽搁等）等非恶意行为关联的比例不是很大。这时往往会与个人因素有关。因此，在第I类的基础上从此处界定初值。另外，考虑到本科生在校期以四年为主，而离校期间一般都会因相关制度的制约而结束在图书馆的这笔“旧账”。当然，也不排除例外，如五年制的学生或在校续读生等。在这里我们只对一般情况进行分析。从考虑上述原因，我们将第II类的下限定在四年。

与第I类相比，此类时间间隔与其相差悬殊，但从上述面积积分图上可以得出这类分离的依据的科学性与合理性。

第III类：超期天数 T 界定在1460天（4年）以上。

与第II类分类依据类似，从读者角度考虑，超期四年以上的多数为教学人员。虽然这些人员所占比例不大，在图中也能得以体现，但从图中横坐标可以看出，其区间长度高于前两类很多，这就缩小了与前两者的差异，基本符合基于此曲线上的分类原则。

根据以上分析，超期天数的三个分类如表格二：

表二：分类情况

分类	超期天数 T /天
第I类	$0 \leq T \leq 60$
第II类	$60 \leq T \leq 1460$
第III类	$T \geq 1460$

6 分类数据的定性分析

6.1 各类数据统计分析

通过以上分类，用EXCEL软件将各类数据进行统计，得出三组数据结果如图3所示。超期书的数目 S 的数据如下：第一类：1088；第二类：2391；第三类：911

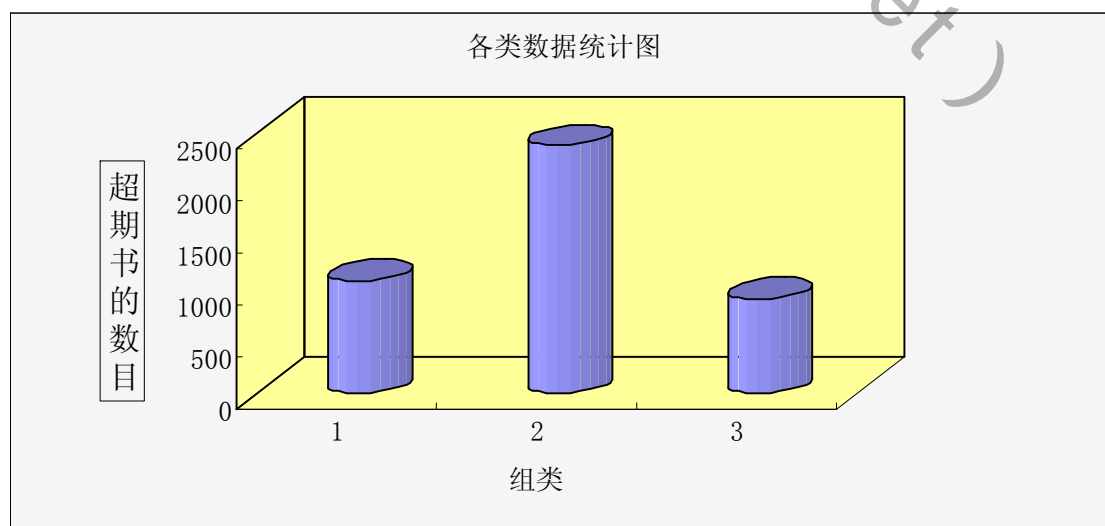


图 3：各类数据统计图

分析上图知第二组数据最高，但因第 I 组数据不足，少前 30 天的数据，所以与其他两类无法做定性比较，因此，在此对第一组不给与讨论。

与第 II 类相比，第三类所占比例不高，所以也不对其作讨论。

6.2 第 II 类中书籍分类定性分析

第 II 类数据占得比例较大，所以对其进行单独分析。应用 EXCEL 将所给数据进行处理，根据借阅书刊的名称对本类数据中所包含的书刊进行分类。分类时可参照国标的索书码查阅不容易区分的类型。由于书比较繁多，所以大体上将书分为两类——自然科学和社会科学。结果统计如下图 4：

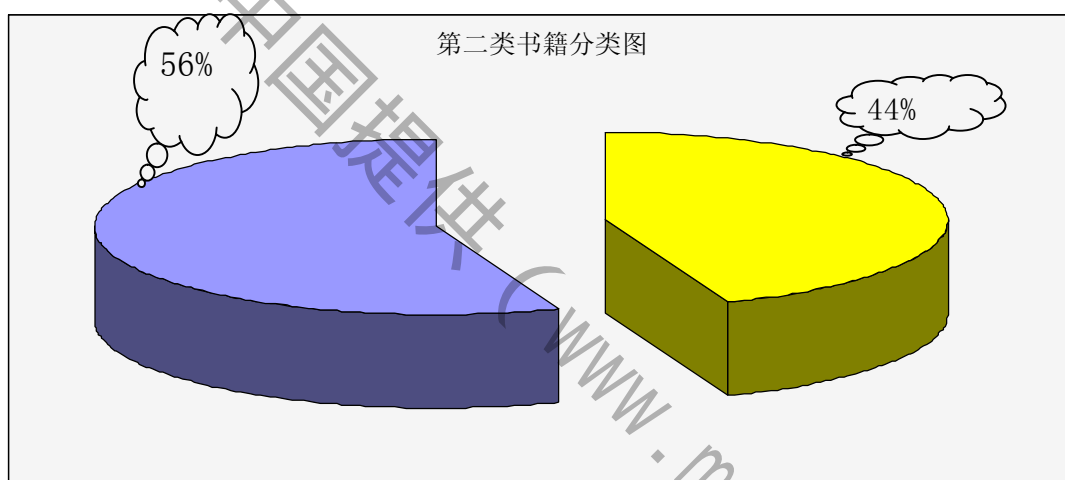


图 4：
各类
书籍
分类
图

图 中
蓝 色
饼 块
代 表
自 然
科 学

类，黄色饼块代表社会科学类。由此可知，在该段区域中自然科学相对而言使用率比较频繁。

此现象受多种因素的制约，首先从学校角度考虑，偏重理工科的学校需求知识的来源更倾向于自然科学类的书籍；而以文科为特色的学校，如财经大学更欢迎社会科学类的丛书。

从题中所给数据可以推断：该校以理工科为主的倾向较大。从社科类的数据中不难发现小说和艺术类等休闲系列丛书很受读者们的青睐。

6.3 读者因素定性分析

为了分析读者这个因素对三类数据中超期天数的影响，我们对三类数据中的超期天数与对应的读者的关系经过排序处理统计制图，结果如下（见图 5）

报名号 # 1110

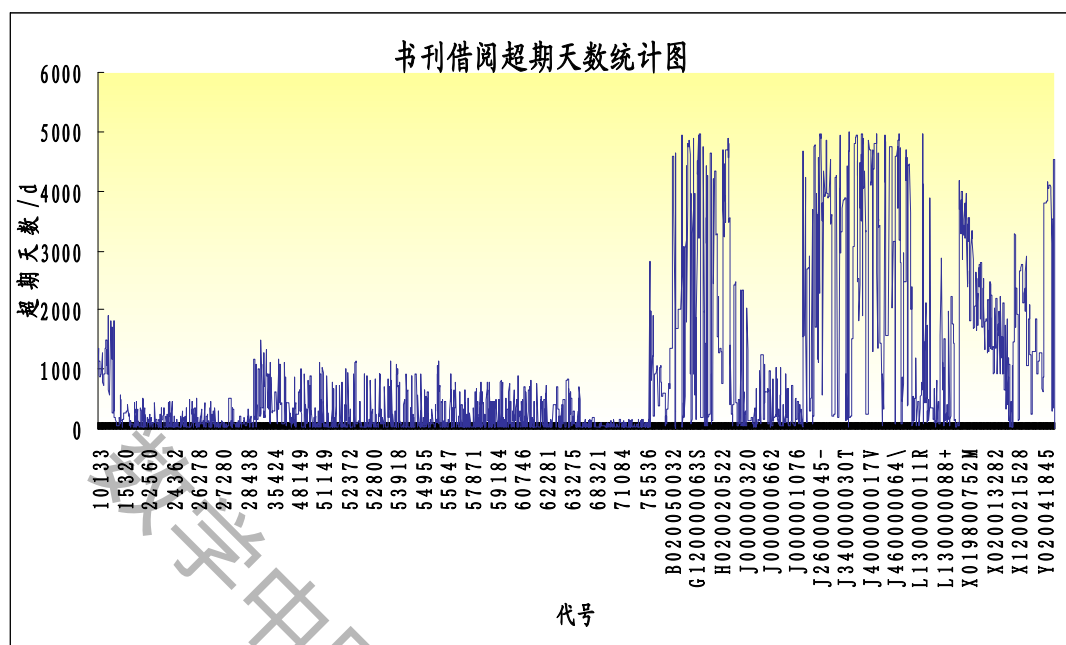


图 5：书刊借阅超期天数统计图

由图中曲线可知：该曲线具有一定的规律性，这些特征可以反映出如下信息：

1. 峰高大致符合成区分布，而且规律明显。因为横坐标按读者号升序排列，而读者号一般在群体分配中会采取连续排列，鉴于此，可以认为在一定的区间内的读者号是同一类，读者属于同一类型，如一个专业甚至一个学院或年级。由此可以判断，同类型读者超期现象基本相似持平，可见环境氛围和一个集体所制定的相关制度对遵守借阅制度的影响之大。
2. 曲线的纵轴高度可以很明显的分为两类，且峰高差距显著。综合来看，读者编号全为数字的基本处于低峰区；而以字母为首的读者号大多处于高峰区。纵轴代表超期天数，根据上述分类可以知道低峰段代表第一二类，而高峰段代表第三类。从横轴分析，横坐标中带字母的读者号基本位于高峰段，也就是第三类的读者号与前两类相比具有很明显的字母标志，不是普通学生的编号，可见这类读者以教师为主，这与分类时的经验依据相吻合，进一步证明了分类的合理性。
3. 进一步仔细分析此图曲线走向，峰高基本符合分区原则，但不符合递增规则。这在带字母区读者号的区间内体现得更为突出。这个区间内有一部分相对低峰区，但与同等高度的低峰区相比，所占比例不高。而在分类中他应属于第二类，这就说明第二类中有一部分教师，但还是以学生为主，这个结论在分类中的依据我们已考虑到，进一步验证了经验性的判断。

6.4 离散图定性分析

通过以上分析，除了得出新的结论外，还验证了第一步分类依据的合理性与科学性。

从上述中的分析，从连续曲线为整体分析提供了依据，整个分析过程陈述了全局的变化与联系，但各类点的分布在图（5）中不能体现，自然针对个体的表现规律就无法分析。对此，我们又在原有坐标基础上绘制了散点分布图，如图（6）所示。

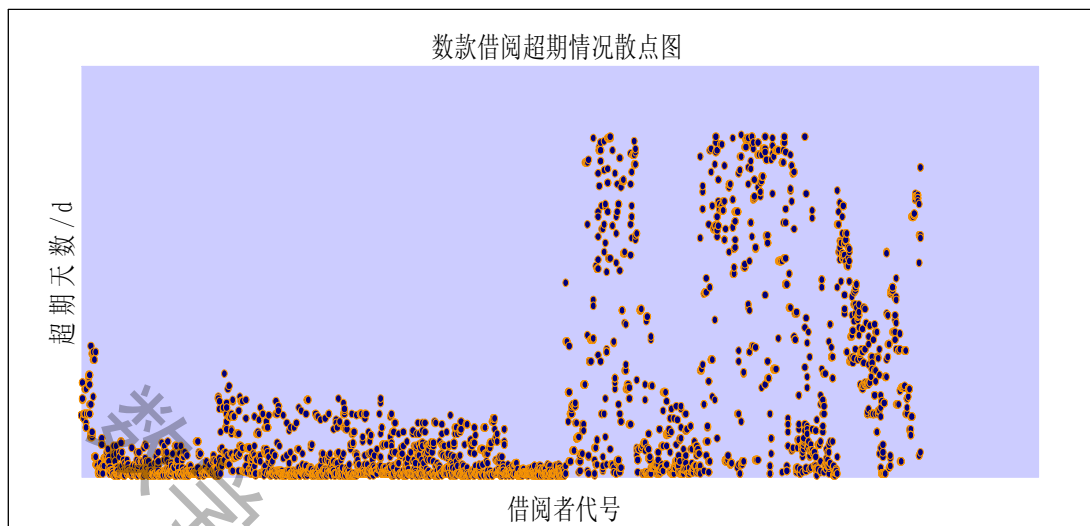


图 6：书籍借阅超期情况散点图

此图中的散点代表着题中所给的每一个数据，因此从此图中可以得出个体的分布统计以及在一定区间内个体堆积的疏密程度，在这里我们称它为整体的分布密度。

因图（5）与图（6）拥有相同的坐标轴，所以结合图（5），在图（6）中可以看出第一类数据总体分布密度最大，第二类分布密度也较大。尽管没有前三十天的数据，但通过对分散密度的分析，可以看出第一二类数据中所对应的原因为主要朝其原因，其中第一类所占比重更大些。

综上可以推断，借阅超期现象以短期超期为主，多为过失超期和意外事件有主要关系。可见图书馆的借阅制度应在此方面给与高度重视，建议图书馆为那些无意的读者提供一定的帮助或提醒，防止这种状况的高频率发生。也许在管理制度和借阅制度方面同时改善，会为解决图书馆资源的有效利用的问题提供一定的帮助。

另外，第三类数据虽然总体分布密度较小，但其所处高度之高也反映出了以教师为主的读者们与图书管理制度之间的矛盾有待于解决。

6.5 离散与综合定性分析

以上对整体和个体两方面分别进行讨论得出上述结论。整体从连续模型的角度出发，通过连续曲线对总体进行统计；而个体通离散模型角度出发，通过散点分布图对个体及其之间彼此的联系进行分析。下面我们将连续与离散综合的思路用于此模型中，得出两种矛盾的解题方法在综合应用中的联系与综合统析。基于此思路，我们将连续图与离散图合并，得出以下综合图形（图 7）。

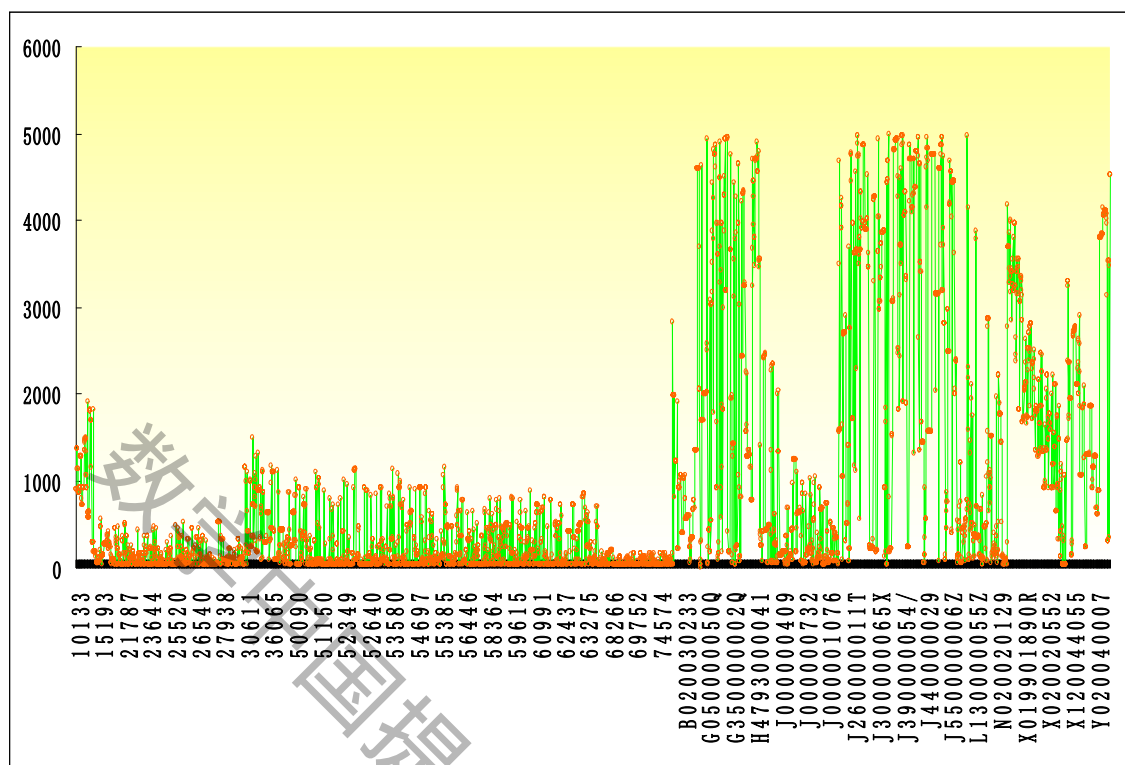


图 7：离散和连续综合图

图（7）是连续与离散的综合，图中橘黄色的散点反映了离散，绿色的连续曲线反映了连续。

连续曲线主要特征在于它的峰高明显，比离散的更加直观；它的最大优势在于它可以将整体的表现规律勾画出来，通过曲线的走向，可以对各类数据进行比较分析，方法直观，而离散对整体的轮廓不能给与良好的放映。

离散点主要特征在于它的堆积密度一目了然，比连续的更具有说服力；在于他可以将个体的分布表现形式放映出来，通过个体之间的疏密堆积程度，可以对各类数据的分布密度给与清晰地描述，而连续对个体之间的联系则体现的不明显。

连续与离散各自有各自的优势和特点，两者放映的是同一对象，但所解决的问题却各不相同。将两者结合，可以将整体与局部很好的过渡，为两者的联系提供了方法。从本体着手，连续曲线主要突出纵轴上（超期天数 T ）可以说明的问题，通过超期时间的不同，将不同区间的超期时间给与统计，在峰高上体现不同区间上书籍超期数目所占整体的比例。而离散点主要强调横轴（读者号 W ）所带来的信息，通过读者号可以反映同类读者，即在一段区域内超期的概率。可见两者分别从横向与纵向两个方面对同类以及不同区间的超期现象做了一个总结。

该方法从整体到局部，从连续到离散，从纵向到横向，全面的系统的解释了超期天数与读者的联系。

7 借阅规则的图形统析评价体系

7.1 借阅规则

某高校的借书规则为：借书时限为 1 个月，可以续借一次，续借时限为 15 天，超期一天罚款 0.1 元。

7.2 借阅规则分析

7.2.1 借阅时限

借阅期限是图书馆为满足读者阅读需求，提高图书的利用率，对读者借阅图书资料规定的期限。借阅期限因馆而异，一般多定为一至二个月。图书馆确定借阅期限主要作用是为了最大限度地发挥本馆馆藏书刊的作用，满足大部分服务对象的阅读需求^[9]。合理的借阅期限可以使读者自由安排读书研究时间，充分利用图书文献，在一定程度上满足读者的阅读需求。通过问题一中聚类分析模型和图形统析模型的分析，可以知道该学校的图书超期四年以上的对象主要是教师，短期内超期的对象主要是学生，该学校的理工科专业图书超期现象比较严重，而一些考研、考级需要用的辅导书超期现象更严重。

由以上分析，可以得出，该图书馆一个月的借书时限，对本科生借阅专业性的自然科学类图书和社会科学类图书而言，是比较合理的。而对于教学和科研较多，需要长期使用学术著作的教师和研究生而言，一个月的借书时限是远远不够的。此外，一些关于考研、考英语四六级、考计算机二三级、考注册会计师书的图书，在特定的时间段里特别热，最容易出现超期现象，图书馆一个月的借书时限会限制该书在这一最需要的时间段里的流通范围，影响该类书的使用价值，也不能满足该类图书服务对象的需求，对该类书一个月的借阅期限也是不合理的。

7.2.2 续借时限

一般图书馆的续借次数为一次，续借时限的天数根据各个图书馆的不同情况而定，在前一问的模型中，我们已经根据超期原因的分析将超期天数分为三类，先对第一类短期内超期天数的书本量进行统计，统计图如图 3 所示。

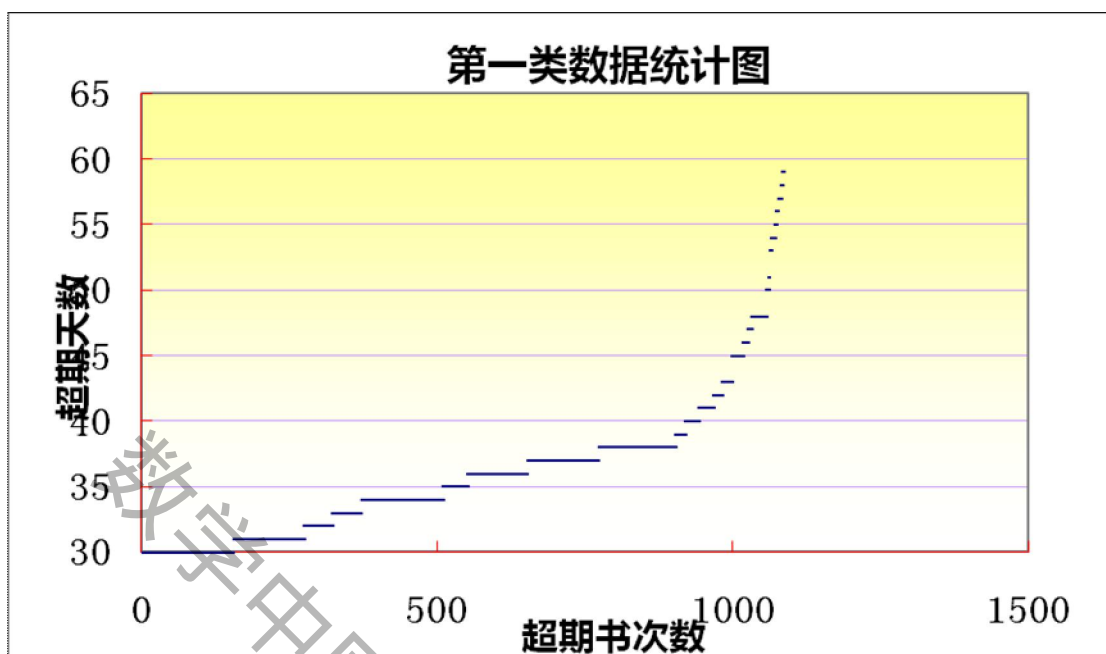


图 3: 第一类数据统计图

对数据统计图进行分析，超期天数在五天之内，即超期天数从30天到35天，书本量达到第一类超期书次数的将近一半。超期天数在十天之内，即超期天数从30天到40天，书本量达到第一类超期书次数M的百分之八十以上。由此可见，该学校超期40天之内的现象很普遍，而该学校的借阅规则中的续借时限仅为15天，这15天显然远远不能满足学生的需求，因而该学校借阅规则中的续借时限的规定是不合理的，而应该根据学生的需求，适当增减续借期限。从第一问中，对超期的书的类别分析中，可以知道，自然科学类的图书超期现象是社会科学类图书的1.5倍左右，同时考虑到图书的进馆的时间长短，针对该学校图书的流通率和学生使用图书类别的多少，可以具体地根据图书的类别，分别对各种类别的图书制定出不同时间长度的续借时限。此外，从前面的分析中，我们已经知道，老师借阅的图书超期时间最长，而老师们由于教学需要和科研需要，借阅的图书大部分都是学术著作和教科书，同时为了保证教科书也能够在学生中很好的流通，对老师的续借时限针对图书的类别也应该制定不同的续借时限。

7.2.3 罚款制度

随着信息需求量的增大，超期借阅在高校图书馆中不可避免，久而久之，对文献资源利用会有不同程度的损害，图书馆有义务保护本馆的图书资源和用户对文献资料的公平使用权利，因此，高校图书馆规定一定的借阅期限是合法合理的，同时，为了保障所有用户利用馆藏图书资源的权利，也应当对超期借阅用户给予一定的处理措施。于是该校的借阅规则中的超期罚款制度是合理的，而该校规定超期一天罚款 0.1 元的规则，是需要讨论的。

从上一问中对超期天数统计的分析中，可以得知，超期天数在 60 天以内的短期之内的人相对要比较多，而在一年与四年之间的人数要相对来说比较少。经过对图书馆工作人员以及超期读者的调查得知，短期内超期的原因主要是，有些学生认为罚款很少，

不在乎那些罚款的钱，即出现最初分析的故意超期现象。此外，超期时间超过一年之后，罚款显然已经很多，已经是所借阅书籍的十几倍，超过了学生的经济承受能力，因而很多超期较长的学生，怕交过多的图书罚款，就一直拖着罚款，不再还书，导致逾期时间越来越长^[10]。在第一问模型中对老师的分析，老师超期时间要远远长于学生的超期时间，为了督促老师不再自己占用图书资源的坏习惯，应该对老师的罚款制定特别的规定。

根据以上的分析，针对超期现象的数量规律，应该是最初的罚款较重，以督促学生以经济的考虑而早日归还所占有的图书资源。而有些学生因为特殊原因而无法早日归还，为避免学生担心罚款较重而不再还书的现象，在一定时间之后的罚款应该减少到学生的经济承受能力之内，由此鼓励由于某种特殊原因而短期内没有及时归还图书的学生归还所借阅的图书。

7.3 制定合理的借阅规则

随着新时期我国高等教育的快速发展，高校图书馆在硬件方面的建设得到了很大的改善。如今，宽敞的馆舍和先进的设备已经难以成为高校图书馆竞争力的主要优势，因为只要资金投入到位，图书馆的硬件建设都能够达到很高的、大致相当的水平，高效优质的服务质量才是衡量各图书馆办馆水平的关键所在。完善的服务规范是提高服务质量的重要途径和根本保证。作为与读者有最直接关联的图书借阅规则无疑是检验图书馆服务水平的高低、服务是否规范的一把标尺。其中，借书期限及逾期还书日数的计算标准既是图书借阅规则所必备，又是图书馆在办理读者借书超期罚款手续的重要依据，同时又由于它涉及到借阅工作中许多细致的操作问题而容易成为引发读者对图书馆服务不满甚至造成矛盾冲突的重要因素，因而这些问题使得很多高校图书馆的工作人员普遍感到头疼和棘手。因此，就借阅规则，我们提出以下可行性方案：

首先，图书馆要在借阅规则中明确、具体地提出借书期限的规定。我们认为，在借书期限的规定中，应对借期、借书期限、还书日期、逾期罚款、续借等等一系列容易歧义的内容作出具体的解释，以减少很多读者对这些概念发生不同的理解。例如，很多读者对借书期限理解为“如果在借书期限30天内遇到节假日，有几天节假日借书期限就应当向后顺延几天”^[11]。这样的理解显然是不正确的^[12]。在上述统计的超期罚款的教师读者当中，有部分也是由于这样的错误理解而导致图书超期。因此制定规则如下：

(一) 严格界定什么是借期、借书期限以及还书日期。

为了让读者的理解与图书馆“立法”宗旨相一致，避免在实施过程中发生争议，必须作如下界定：（1）借期：每周以七日计算，包括星期六、星期日及法定节假日在内；（2）借书期限：自办理借书之次日起算至还书日止；（3）还书日期：依据借阅规则规定借书期限为30天，则设定借期为26天（电脑系统应设定显示），并给予读者4天的催还期，但还期适逢星期六、星期日、寒暑假、法定节假日或本馆全日不开放借还业务时，分别规定如下：A、还期逢星期六、星期日、寒暑假、法定节假日：以假期届满之次日为还期；B、还期逢本馆全日不开放借还业务时，以下一开放借还业务之首日为还期。

(二) 明确规定逾期日数计算标准。

逾期日数：自第1点（1）、（2）、（3）条规定的还期后之次日起算，在逾期中不管遇什么假日、本馆不开放日，或因不可抗力不能还书者也一并算入；超过催还期的图书（逾期未还图书），将按日、按册数罚款，一周之内，每天每册罚款人民币0.5元；逾期一个月之内，罚款0.2元，如果逾期一年可适当减少罚款，如每周每册罚款人民币0.1元；有图书逾期未还的读者，本馆暂停其借书的权利^[13-14]。

(三) 续借应设置一定条件。

图书馆规定学生读者可以续借 1 次、老师 2 次, 续借只可于第一次借期届满前 7 天内办理, 借期自次日起算, 可续借 40 天, 提前及逾期均不再办理续借。

其次, 图书馆要建立规范的催还制度, 配套服务使读者感到满意。图书馆的服务理念是“善待每一位读者”, 就是要多站在读者的立场上考虑问题, 尊重读者, 关心读者。因此, 图书馆有义务为借书即将当期或已经到期的读者提供及时的、善意的提醒, 体现出为读者着想的诚心。

1) 催还分类:

借期届满的催还和续借期间的催还。

2) 催还规定:

A、借期届满的催还。(1) 催还期: 设定为 4 天, 参照“还书日期”规定的内容执行。在催还期内归还的图书不罚款, 超过催还期的图书电脑系统将按日、按册计算罚款。

(2) 在开放借还业务工作日里, 每天应给当天新进入催还期的图书和其名下的读者汇总, 便于发布催还通知。B、续借期间的催还: 随着图书馆为读者的服务项目越来越细, 将来借书预约业务开展起来后, 续借期间若有人办理预约, 需于本馆催还通知发出日起算 7 天内归还。这样规定体现读者利用图书资料的公平原则^[15]。

3) 催还方式:

(1) 以“通知的形式”进行催还

在学生密集的地方, 张贴姓名通知单(催还的读者可只写书证号, 不写姓名), 例如教学区、餐厅门口, 宿舍区必经通道处, 图书馆必经处等等学生、教师经常到的地方, 进行张榜通知提示图书到期的读者。

(2) 以“电话形式”进行催还

在读者借阅图书的时候, 图书馆老师负责让读者把其有效的联系方式(如手机号、固定电话)留下, 当读者所借阅的书籍即将到期或者已到期的, 老师们及时地给读者发短信或打电话做友情提醒, 进行催还服务。至于电话费用, 各图书馆可根据自己的实际情况来定, 可以由读者自己来承担, 也可以由图书馆来承担, 老师记录好通知情况, 如果某一读者多次要图书馆来通知还书, 那么老师可以和他们要取电话费用等。

(3) 利用“网络资源优势”进行催还

在如今这个网络大环境下, 高校成员几乎人人都要利用电脑进行工作、学习。在校园网上或图书馆主页上, 利用网上通知的方式, 提醒有到期图书的读者引起注意。腾讯 QQ 现在是世界也是中国上使用人数最多的免费聊天软件, 2006 年 6 月 3 日, 腾讯公司宣布, 腾讯 QQ 最高同时在线用户突破 2000 万, 是一个妇孺皆知的聊天软件, 也是当今大学生获取信息不可缺少的聊天软件。因此, 图书馆在建立读者档案的时候, 可以把读者的电子邮箱和 QQ 号码留下存档, 工作人员便可以利用网络, 与读者沟通, 进行网上催还服务。我们认为, 这是一种高效、经济、快捷的书刊催还方式。

针对虽收到图书馆友情提醒, 但在逾期一个月以后仍未还书的读者, 图书馆应采取严重的催还方式。如今, 大多数高等学校都是采用“一卡通”模式, 因此, 我们建议, 此时图书馆可以采取的催还方式, 将借阅逾期过长的读者的校园卡冻结, 以此来提醒其应该将图书归还到图书馆, 便于图书的流通。

对于老师借阅图书的情况, 图书馆可以采取别类方法。由于老师们要做科研活动、教学工作等, 需要长期借阅一些图书, 但为了方便更多的读者借阅, 因此当有读者需要该类书籍时, 图书馆老师可以给借阅的老师打电话, 让他将图书暂时借给其他读者借用一下。

再者, 图书馆要主动宣传图书馆制度, 同时建立读者诚信制度。

1、要想让读者了解图书馆制度，首先得让他们感兴趣，因此学校和图书馆一定要做好宣传工作，如制作并派发图书馆制度手册；鼓励工作人员到读者中间去宣传，鼓励学生班干部以身作则，在同学中间宣传；重视新生入馆教育，把它当作展示、宣传图书馆的最佳时机，为新生正确利用图书馆打下基础。

2、建立读者诚信制度，冻结有过期图书的读者的使用权限，“逼”其还书读者依法享有使用文献资源的权力，同时也有不得侵害其他读者合法使用的义务，不能按时归还所借图书，随之就丧失了继续使用图书馆的其他权力，直至归还超期图书之后。读者诚信制度是管理文献的一种辅助措施，具体来说就是全程跟踪并记录读者的借阅行为，并依据所设定的分值及遵章守纪的程度，对读者行为打分。如对损坏图书、超期等读者减分，而对借阅频繁、按期归还图书的读者加分等，然后在图书馆的权限范围内，对诚信分值高的读者给予一定的特别服务，而对于诚信分值过分低者，则要限制其某方面的权利。以此来引导读者讲究诚信、遵守图书馆的规章制度^[16]。

最后，图书馆要适当增加还书渠道和方式

对于还书有困难的读者，可以提供特别服务，如异地实习的学生，我们可以在实习地设立还书站或者对要进行实习的高年级读者的借期放宽到回校的时间。年龄大的或退休教师读者数量相对较少，在工作能力范围内，本着爱护老人、尊敬老人的原则，我们可以提供上门取书还书的特殊服务^[17]。

以上就是我们提出的借阅规则，希望此规则可以给大多高校图书馆提供便利，即方便图书馆老师，又方便借阅图书的老师同学和即将借阅图书的老师和同学。

7.4 针对图书超期现象给图书馆的建议

给高校图书馆领导的一封信

尊敬的图书馆领导：

随着学校的发展，图书馆的硬件建设都能够达到很高的、大致相当的水平，高效优质的服务质量才是衡量各图书馆办馆水平的关键所在。完善的服务规范是提高服务质量的重要途径和根本保证。作为与读者有最直接关系的图书借阅规则无疑是检验图书馆服务水平的高低、服务是否规范的一把标尺。图书超期现象是各个图书馆都会发生的现象，为了减小此现象的发生，图书馆应该在图书借阅规则上做一些人性化和合理化的改动。

其中，借书期限及逾期还书日数的计算标准既是图书借阅规则所必备，也是图书馆在办理读者借书超期罚款手续的重要依据，同时又由于它涉及到借阅工作中许多细致的操作问题而容易成为引发读者对图书馆服务不满甚至造成矛盾冲突的重要因素。通过我们对一所具有代表性高校图书馆图书超期现象和图书借阅规则的分析，我们对高校图书馆借阅规则提出一些改进意见。

一、 根据实际需要，合理制定图书借阅期限和续借时限。

针对借阅对象和服务方式的不同而实际分析，高校图书馆的借阅对象是教师、研究生和学生，教师因教学和科研的需要，借阅期限可适当延长。其次根据服务区域的大小改动图书借阅期限，服务区域大，则借期应相应长一些。另外根据图书类型，制定不同借阅期限，利用率高和新到馆的图书期限须短些。

二、 利用现代信息技术，制定智能化、人性化的催还提醒。

利用手机短信、校园卡系统等现代信息技术，在读者借阅图书到期前几天之内，自动发出催还提醒。

三、 制定专门针对教师的借阅规则。

教师是学校一个特殊的借阅图书群体，对老师的借阅图书数量以及借阅期限要特殊考虑，但要确保教师有还书的意识，能够在不在使用之后，及时把图书换到图书馆，以方便其他人使用。

只要科学分析读者的阅读需求，充分吸取读者意见，根据各馆实际情况，制定出的借阅期限一定会符合读者的要求。当然，科学制定借书期限只是作好服务的一个方面，另一方面也要加强宣传，使制定借阅期限的目的和要求深入人心，读者自觉遵守借阅期限，积极配合工作人员把借阅工作搞好。

8 模型的科学性和现实意义

8.1 模型的科学性分析

本文我们从题目中所给的实际数据开始入手，先将图书超期这一宽广的问题划分为几个小问题，使文章具有很强的连贯性与层次感，从建立图书超期聚类分析数学模型的建立，再到提出及建立图形统析模型，极强的逻辑性是数学建模论文思维紧密所必不可少的前提，我们在处理好数据的基础上做好每一步的推导工作，通过对两个模型分别建立评价体系对该高校的借阅规则进行了详细的评价，因此，准确的数据、清晰的图形以及准确的分析是本文科学性的强有力的说明。

我们思路、方法及数学模型的合理性主要体现在以下几个方面：

1) 假设的合理性

无论是从题目中归纳出的假设条件，还是根据查找文献材料所确定的假定内容，我们都做到了有依可寻，避免了主观的臆断而出现模型过于简化的情况，合适的假设是我们模型求解的关键，从求解结果的分析来看假设也是合情合理的。

2) 方法的科学性

本题的问题虽然只有一问，但是我们却把它划分为许多小问进行处理，化整为散，逐渐深入，使得问题解决得更加完善。同时建立两个不同的模型，针对不同的数学模型相应的使用聚类分析算法、图形及数据统计等数学工具来求解，使问题解决得更为科学，更有说服力。

3) 数据处理的合理性与清晰性

该题中涉及了许多数据，这些数据的处理利用强大的数学分析软件Matlab和数据分析软件Excel，进行合理的处理，并作出清晰、关系明了的图形。合理的数据处理，为数据分析进而对图书超期原因的分析及借阅规则的提出，提供了有力的凭据。

4) 求解方法的可靠性。

我们用软件 **Matlab** 及Excel分别对模型进行了求解，并且都在模型求解之后

对模型进行了评价，保证了模型求解的可靠性。

5) 题目与实际情况的结合

题目中所提供的该所大学图书馆的借阅情况，与我们所在大学图书馆借阅情况非常相似，依据我们对自己学校情况的了解对自己学校图书馆工作人员和图书超期读者的调查，我们对该校图书超期现象进行了合理的分析。

8.2 模型的现实性意义

通过建立合理的数学模型和详细的数据处理，以及对数据进行图形处理，对某一所高校的图书超期现象出现的原因进行了分析，并根据所建立的模型建立了合理的评价体系，对该高校的借阅规则进行了评价，并根据该校的具体情况提出了相应比较合理的借阅规则。模型具有很强的实用性，可以根据任意一所高校图书馆的图书超期记录，对该校的超期现象进行分析，并对该校提出相应合理的借阅规则。综上所述，我们所建立的模型，对于解决现在各高校以及各个单位的图书馆图书超期和图书借阅规则的制定问题有着极其重要的作用和意义。

9 模型的评价与推广

9.1 模型的优点

- 1) 本文所用数据为高校图书馆中心提供，具有很强的真实性与代表性；
- 2) 本文运用在数据分析中有很重要作用的聚类分析算法，对数据进行合理的分类处理，并建立评价体系，对借阅规则合理的评价，由此文章具有很强的逻辑性与合理性；
- 3) 本文将该高校的借阅规则分为若干部分，并对各个部分分别进行讨论分析，然后针对该小部分提出合理的建议，文章的思路清晰且具有很好的连贯性；
- 4) 本文在根据数据进行图书超期原因分析时，结合自己学校实际情况以及对图书馆工作人员和超期读者的调查，总结出的原因较为全面，使模型及体系具有很强的使用性与代表性；
- 5) 本文中，两个模型相对比，更能相互验证各个模型所得到结果的正确性与合理性。

9.2 模型的缺点

- 1) 模型中只是对超期现象出现的主要原因进行了分析，并未对其中的某些特殊原因进行分析。
- 2) 模型的假设中，对图书超期第一个月与第二个月书的超期频率相同的假设，与实际中的情况不相符。
- 3) 对该校借阅规则评价之时，考虑的因素不全面，未从图书馆工作人员的工作方便角度比如：工作人员处理藏书周转的能力、流通部的人员数量及工作能力等方面来考虑，也并未考虑该图书馆的藏书量。

9.3 模型的推广

图书超期现象原因的分析问题属于一类对常见实际问题进行原因分析的问题，而对借阅规则的评价属于一类评价问题。

本文所建的模型具有很大的通用性,例如在研究生录取问题,医疗体系评价等方面都可以应用。

参考文献

- [1]安园园,浅析高校读者借阅超期现象的原因及应对策略[J].现代服务,2008(23):162-163.
- [2]胡福文,高校图书馆读者借阅超期有关问题的思考[J].学理论,2009(17):158-159.
- [3]陈竞,高校图书馆图书超期借阅超期罚款中存在的矛盾及对策——以上海城市管理职业技术学院图书馆为例[J].2009,8:89-91.
- [4]向东进.实用多元统计方法[M].武汉:中国地质大学出版社,2005.
- [5]张尧庭,方开泰.多元统计分析引论[M].北京:科学出版社,1982.
- [6]于泓漪.道路交通事故原因的聚类分析[M].长春:吉林大学硕士学位论文,2005.
- [7]陈平雁.SPSS13.0统计软件应用教程[M].北京:人民卫生出版社,2005.
- [8]鲍婷婷,涂娟,刍议高校图书馆超期借阅处罚制度[J].农业图书情报学刊,2008,(02):172-173.
- [9]范洁,浅谈高校图书馆借阅期限的确定问题[J].高校图书馆工作,2001(5).
- [10]牛晓芳,马红月,霍宏,邵红,图书逾期不还的原因分析与对策[J].中华医学图书情报杂志,2005(05):40.
- [11]马波,关于高校图书馆罚款制度法律依据的探讨,山东图书馆季刊,2008(3):
- [12]李莉,医院图书馆借阅期限的合理确定[J].医学信息,2003,10.
- [13]李立新,图书馆罚款论,图书馆理论与实践,2007(7):73-77
- [14]国立东华大学图书馆,读者违规处理规定,2006,09,18.
- [15]张冬梅,解决流通管理自动化中的若干问题,情报探索,2006,(12)
- [16]李润秀,图书馆开架借阅出现的问题及解决措施[J].贵州社会主义学院学报,2005,1.

附录

Matlab 求解聚类分析的源程序代码：

```
%std1.m, 用极差标准化法标准化矩阵
function std=std1(vector)
max=max(vector); % 对列求最大值
min=min(vector);
[a,b]=size(vector); % 矩阵大小 ,a 为行数 ,b 为列数
for i=1:a
for j=1:b
std(i,j)=(vector(i,j)-min(j))/(max(j)-min(j));
end
end

%ds1.m, 用绝对值法求距离
function d=ds1(vector);
[a,b]=size(vector);
d=zeros(a);
for i=1:a
for j=1:a
or k=1:b
(i,j)=d(i,j)+abs(vector(i,k)-vector(j,k));
end
end
end

fprintf(' 绝对值距离矩阵如下： \n');
disp(d)

%min1.m, 求矩阵中最小值，并返回行列数及其值
function [v1,v2,v3]=min1(vector);%v1 为行数， v2 为列数， v3 为其值
[v,v2]=min(min(vector'));
[v,v1]=min(min(vector));
v3=min(min(vector));

%min2.m , 比较两数大小，返回较小的值
function v1=min(v2,v3);
if v2>v3
```



```

        v1=v3;
    else
        v1=v2;
    end

%cluster.m, 最短距离聚类法
function result=cluster(vector);
[a,b]=size(vector);
max=max(max(vector));

    for i=1:a
        for j=i:b
            vector(i,j)=max;
        end
    end

end;
for k=1:(b-1)
[c,d]=size(vector);
fprintf(' 第 %g 次聚类 : \n',k);
[e,f,g]=min1(vector);
fprintf(' 最小值 =%g, 将第 %g 区和第 %g 区并为一类, 记作 G%g\n\n',g,e,f,c+1);
for l=1:c
    if l<=min2(e,f)
        vector(c+1,l)=min2(vector(e,l),vector(f,l));
    else
        vector(c+1,l)=min2(vector(l,e),vector(l,f));
    end
end;
vector(1:c+1,c+1)=max;
vector(1:c+1,e)=max;
vector(1:c+1,f)=max;
vector(e,1:c+1)=max;
vector(f,1:c+1)=max;
end

%print1, 调用各子函数
function print=print1(filename,a,b); %a 为地区个数, b 为指标数
fid=fopen(filename,'r')
vector=fscanf(fid,'%g',[a b]);
fprintf(' 标准化结果如下: \n')
v1=std1(vector)
v2=ds1(v1);
cluster(v2);

```

% 输出结果

```
print1('fname',9,7)
```

数学中国提供 (www.madio.net)