

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的建模

关 键 词 高校图书馆 数据挖掘 衰减模型 借阅罚款规则 EXCEL

摘 要

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。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。超期罚款政策作为的一项管理手段得到了世界图书馆界的普遍采用，并且在长期的实践中被证明是行之有效的。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，因此我们根据提供的数据对高校超期问题进行如下研究：

第一问，首先我们对提供的数据进行信息挖掘，根据读者证号编码方式的区别以及两种不同读者证号的平均超期借阅天数的差异，将读者分为两类：学生和教师。鉴于学生和教师对图书的需求可能会存在差异，故将学生和教师分别考虑。我们根据中国图书分类法将图书大体分成五类：计算机类，经管类，人文类，理工类和其他类。接着我们利用随机抽样方法对数据进行抽样分析，得到抽样比例模型，根据比例大小，确定学生和教师的容易借阅超期书籍的种类都是理工类、人文类和计算机类。由于还需要知道超期的原因，于是我们分别对学生和教师借阅超期书籍的数量随超期时间作图。接着分别对学生和教师借阅超期书籍的数量随超期时间的关系图进行信息挖掘。根据假设[4]我们可以得到，超期借阅书籍的归还数量与天数的关系和催还清单上超期借阅书籍的数量与天数的关系应该大致相同。通过图形分析，我们可以得到定性分析结论，对于学生和教师而言，较短时间内归还超期书籍的数量都较多，为无意超期的原因，较长时间学生则是故意拖延超期，而教师则是科研或教学需要而超期。

第二问涉及到对现行借书超期惩罚规则进行优劣评价，是第一问的延伸和继续，我们利用第一问分析得到的结论和建立起来的模型对图书馆借书超期惩罚规则进行优劣评判，然后我们结合该图书馆借书超期惩罚规则中的缺点和不足以及学生和教师的借书情况，分别建立相应的图书借阅规则，从而使得新的图书馆借书超期惩罚规则更加符合图书馆的实际情况，从而对症下药，体现图书馆借书超期惩罚的真正作用。

参赛队号 1869

所选题目 C 题

参赛密码 _____
(由组委会填写)

英文摘要（选填）

(此摘要非论文必须部分，选填可加分，加分不超过论文总分的 5%)

报名号 #1869

目录

1	问题的背景与重述.....	2
1.1	问题的背景.....	2
1.2	问题的重述.....	3
2	问题的分析.....	3
2.1	第一问的分析.....	3
2.2	第二问的分析.....	4
3	模型假设.....	4
4	符号系统说明.....	4
5	问题一模型的建立与求解.....	5
5.1	图书的大体分类及抽样比例模型的建立.....	5
5.1.1	抽样方法介绍.....	6
5.1.2	具体抽样应用.....	7
5.1.2.1	总体样本的统计结果.....	7
5.1.2.2	教师样本的统计结果.....	8
5.1.2.3	学生样本的统计结果.....	8
5.1.3	这类书籍容易发生超期现象的原因.....	9
5.2	模型建立.....	10
5.2.1	基于总体数据模型的建立.....	10
5.2.2	基于学生的衰减模型的建立.....	12
5.2.3	基于教师的衰减平稳混合模型的建立.....	14
6	问题二借书规则的优劣评价与新的借书规则的建立.....	16
6.1	该图书馆现行借书规则的优劣评价.....	16
6.2	新的借书规则的设计.....	17
6.2.1	基于学生的借书规则的设计.....	17
6.2.2	基于教师的借书规则的设计.....	18
7	给图书馆馆长的一封信.....	19
8	模型的进一步讨论.....	21
	参考文献.....	22

报名号 #1869

1 问题的背景与重述

1.1 问题的背景

当今世界，随着科学技术与人类文明的飞速发展，信息的需求量不断增大，图书馆作为信息的流通媒介，也面临着越来越多的挑战。

图书馆是搜集、整理、收藏图书资料供人阅览、参考的机构，早在公元前3000年就出现了最早的图书馆，图书馆有保存人类文化遗产、开发信息资源、参与社会教育等职能。图书馆的功能之一，就是要收集、加工、整理、科学管理这些珍贵的文献资源，以便广大的读者借阅使用。图书馆是作为保存各民族文化财富的机构而存在的，它担负着保存人类文化典籍任务。但近年来由于计算机网络化的实现以及科学技术的突飞猛进，图书馆不但保存手写和印刷的文献，还保存其他载体形式的资源，而且保存的目的是为了更好地使用。同时图书馆收藏着大量的文献信息资源，积极地开发，广泛地利用这些文献资源是图书馆的重要职能之一，它也是图书馆承担各种职能的基础。图书馆通过对文献信息资源进行加工整理、科学分析综合、指引，形成有秩序、有规律、源源不断的信息流，进行更加广泛的交流与传递，使读者更好地利用它们。图书馆的文献资源开发包括下面几项内容：第一，对到馆的文献进行验收、登记、分类、编目、加工，最后调配到各借阅室，以便科学排架，合理的流通；第二，对馆外文献信息资源进行搜索、过滤，成为虚拟馆藏，形成更加宽广、快捷的信息通道；第三，通过最现代化的手段——计算机网络操作技术使馆藏文献走向数字化。

高校的图书馆在所有图书馆之中占有很重要的地位，它为教学和科研服务，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，它不仅影响图书馆工作的正常开展，而且对读者的权益也带来不利影响。处理超期罚款是图书馆管理图书流通服务中经常要处理的事务。处理超期罚款虽是小事，但处理不当便会影响图书馆与读者之间的关系，从而影响图书馆的声誉和形象，或者是馆藏资源的利用率，最终影响图书馆功能的正常有效发挥。因而正确地解决超期罚款问题对于图书馆的声誉和形象以及读者的切身利益相当重要。



图 1-1 图书馆藏书

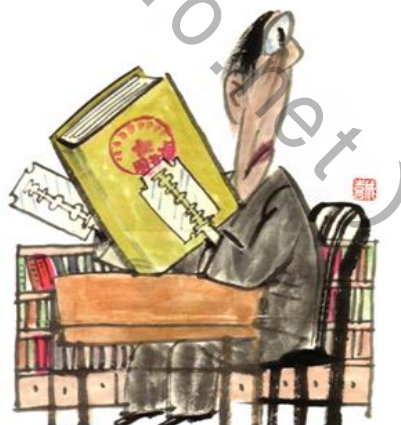


图 1-2 图书超期借阅

报名号 #1869

1.2 问题的重述

C 题：高校图书馆的智能服务

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：

第一阶段问题：

(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 问题的分析

2.1 第一问的分析

题目提供的数据包括读者证号，条码号，还期，书名和过期天数。读者证号有两类，一类由以0开头的一串数字组成，一类由以字母开头的字符串组成。通过模型的假设我们认为这两类应该分别属于学生和教师，于是我们把这两类分开，分别计算超期天数的平均值，以0开头的读者证号的超期天数的平均值（230天）比以字母开头的读者证号的超期天数的平均值（1829天）少很多，因此我们可以认为读者证号以0开头的这类读者为学生，而读者证号以字母开头的这类读者为老师。之所以这样判断是因为学生借阅图书的时间较短，绝大多数都不从事科研，而且学生平均呆在学校的时间最多为5年，但是以字母开头的读者证号平均超期天数为5.1年（1829天），如果是学生的行为，那图书馆的藏书功能已经被破坏，同时学校教育学生方面也将是失职的；而考虑到教师长期呆在高校，并且借阅图书的时间较长，大部分都从事科研或者教学，因此允许借阅的时间会很长。如果允许借阅的时间按年来计量的话，那么造成长时间超期现象也很正常。我们借鉴中国图书分类法将清单中的图书进行分类，分别为：计算机类，经管类，人文类，理工类和其他类。将读者按读者证号分类之后分别对学生，教师和总体超期借阅的书分别进行随机抽样，统计超期的各类书的数量，并建立抽样比例模型，分析哪些书容易发生超期现象。

下面就学生和教师借书超期的原因进行定性分析。对于借阅图书超期，原因众多，但考虑到高校的特点，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。对于学生来说，参与科研的人数毕竟较少，和学生总数相比微不足道，因而可以将这个原因忽略，从而我们将学生借阅书籍超期的原因分为两种：1、无意错过还书期限；2、由于罚款数目较少，所以故意拖延。考虑到大部分的学生的素质还是很高的，素质差、故意拖延的学生

报名号 #1869

的数量还是很少的，客观上讲第一个原因是主要的，同时故意拖延的现象还是存在的，其严重与否因学校学生的总体素质而异。不可能存在大量学生无意超期的时间很长，因而大部分无意超期图书的超期天数应该较短，借阅超期天数较长的应该是故意拖延的情况。对于教师来说，教师为人师表，故意拖延会伤害个人形象，公众影响较为恶劣，同时故意拖延和由于科研需要这两个原因对于教师来说难于区分，因而我们可以不需要考虑教师存在故意拖延而造成超期的原因或者将教师故意拖延归到由于科研需要来处理，那么我们将教师超期原因归结为两点：1、无意错过还书期限；2、由于教学科研需要，长期使用。接着我们分别对学生超期借阅书的数量随超期天数的关系作图，以及对教师超期借阅的书籍的数量随超期时间作图，分别分析每段时间超期的主要原因，以及曲线所反映的情况，然后可以分别对两图的曲线做多项式拟合。

2.2 第二问的分析

首先，图书馆现行借书超期惩罚规则没有轻重之分，这点就存在很大的问题，同时第二问涉及到对现行借书超期惩罚规则优劣评价，是第一问的延伸和继续，我们可以利用第一问分析得到的结论对图书馆借书超期惩罚规则进行评判，找出其缺点和不足之处，然后在新的图书馆借书超期惩罚规则中对缺点和不足之处进行改进，从而使得新的图书馆借书超期惩罚规则更加符合图书馆超期借阅的实际情况，从而对症下药，体现图书馆图书馆借书超期惩罚的真正作用。

3 模型假设

- [1] 假设高校图书馆图书借阅的人仅有学生和教师两类，不考虑外来人员借阅图书所带来的影响；
- [2] 假设学生不参与科学研究，即使有人参与科学研究，其数量与学生的总人数相比较非常小，可以忽略不计，即可以认为学生不参与科学研究；
- [3] 假设学生与教师的读者证号的编码方式不相同，也就是学生与教师的读者证号的编码方式各为一种；
- [4] 假设超期借阅书本的归还情况与超期借阅天数的关系和催还清单上超期借阅书本的数量与超期借阅天数的关系相同，即催还清单上所表示的关系即为超期借阅书本归还的数量与超期借阅天数的关系。
- [5] 假设该高校的学生参与科学研究情况和教师为人素质方面和其他高校相似，也就是学生几乎不参与科学研究，并且教师的素养比较高，不故意拖延借阅图书。

4 符号系统说明

为了便于描述问题，我们用一些符号来替代问题中涉及的一些基本变量，表示如下：

符号说明

- y 超期借阅图书的数量。
- x 超期天数。
- x_1 超期借阅的书籍的数量与超期时间的关系中的时间段（50 天为一段，用于对总体数据和学生数据的分析）。
- x_2 教师超期借阅的书籍的数量与超期时间的关系中时间段（100 天为一段，用于对教师数据的分析）。
- R^2 回归系数，是相关系数的平方，表示所得到的结果的可信程度，值越接近 1 则信度越高。

报名号 #1869

5 问题一模型的建立与求解

5.1 图书的大体分类及抽样比例模型的建立

根据中国图书分类法，我们将所有催还清单中记录的书籍大体分为计算机类，经管类，人文类，理工类和其他类。

通过对搜集到的数据的分析和探讨。我们发现过期天数从30天到5000天不等，过期书的书名五花八门，书的种类繁多，有的还不好判断，难以归类。看到条码号，再想到以往我们借的书上好像也有条码号，我们就跑去学校图书馆去查看，在翻过大量书以及有关条码号编码的书籍后，我们只能判断每一个条码号对应一个书名，但是在图书馆的时候我们还发现我们学校图书馆中同一类的书籍的条码号编码方式也有多种，大部分以C开头，有的以A开头，而有的还以0开头，这样的话我们也没办法通过条码哈号来将超期的图书分类。我们也问了图书馆的相关老师有关条码号和书的分类问题，也同样的不到满意的答案，于是我们猜测条码号的编码可能与图书的进馆时间有关，而且图书馆不同时期使用的编码方式有所不同才导致同一类书的条码号有不同的方式，题中数据按条码号排序后发现相近的条码号对应的书的种类完全不一样，如下表一，而且编码方式也有多种如表二，这也验证了我们的猜测。如此看来我们只能通过手动来将超期的图书进行分类。由于数据有4397个，我们只能进行抽样处理。再看所有读者的读者证号，将所有数据按照读者证号进行排序，发现有大概两类号码，一类是以0开头的一串数字，一类是以字母开头的字符串，如表二所示。于是我们把这两类分开，分别计算超期天数的平均值，以0开头的超期天数的平均值为200多天，而以字母开头的超期天数的平均值为1000多天且接近2000天，于是我们可以认为读者证号以0开头的一类读者为学生，而读者证号以字母开头的一类读者为老师，那么学生超期借阅的书总共有2536本，而老师超期借阅的书总共有1861本。之所以这样判断是因为学生借阅图书的时间较短，绝大多数都不从事科研，而且本科生呆着学校最多4年，大多数学生的经济状况也一般；而老师借阅图书的时间可以较长，大部分都要从事科研或者教学，因此借的书的超期时间会很长，而且老师可以呆着学校很长时间，老师的经济状况一般都不错。

分类之后分别对老师，学生和总体超期借阅的书分别抽样，统计超期的各类书的数目，分析哪些书容易发生超期现象。

表一

读者证号	条码号	还期	书名	超期天数
0000063486	00831717	Wednesday, October 01, 2008	空气污染气象学 =Air pollution m	216
0000061924	00831618	Tuesday, May 06, 2008	基础工程	364
0000010294	00831590	Tuesday, August 09, 2005	举一反三——3ds Max 三维动画 制?/font>	1365
0000030319	00831442	Tuesday, March 07, 2006	谈判就这么简单! =Negotiate thi	1155
0000062746	00831299	Sunday, March 16, 2008	当代语用学	415

报名号 #1869

表二

读者证号	条码号	还期	书名	超期天数
Z360000076E	Z0440159F	Tuesday, December 17, 1996	铁路设计	4522
Y029606397X	X0356590I	Tuesday, February 03, 1998	Oxford Advanced Learner`s Dict	4109
X020022939	W00051588	Tuesday, July 12, 2005	艺术哲学 (上)=:	1393
0000015001	RW032041	Tuesday, March 17, 2009	政治经济学前史	48
0000059441	01268381	Friday, March 27, 2009	BIOS 与 Windows 注册表设置一点通	38

5.1.1 抽样方法介绍

抽样是一种非全面的调查方法。在统计调查和统计分析中都有广泛的应用。抽样调查是按随机原则从全部研究对象中抽取一部分观察，根据样本资料计算样本的特征值，然后以样本的特征值，对总体的特征值做出具有一定可靠性的估计和判断，以反映总体的数量特征和数量表现的一种统计方法。所谓随机原则，即是在抽取样本时，排除人们主观意图的作用，使得总体中的各对象均以相等的机会被抽中。

一. 简单随机抽样

(抽签法、随机样数表法)常用于总体个数较少时，它的主要特征是从总体中逐个抽取；

优点：操作简便易行

缺点：总体过大不易实行

1. 定义：

一般地，设一个总体含有 N 个个体，从中逐个不放回地抽取 n 个个体作为样本 ($n \leq N$)，如果每次抽取时总体内的各个个体被抽到的机会都相等，就把这种抽样方法叫做简单随机抽样。

2. 简单随机抽样方法

(1) 抽签法

一般地，抽签法就是把总体中的 N 个个体编号，把号码写在号签上，将号签放在一个容器中，搅拌均匀后，每次从中抽取一个号签，连续抽取 n 次，就得到一个容量为 n 的样本。

抽签法简单易行，适用于总体中的个数不多时。当总体中的个体数较多时，将总体“搅拌均匀”就比较困难，用抽签法产生的样本代表性差的可能性很大。

(2) 随机数法

随机抽样中，另一个经常被采用的方法是随机数法，即利用随机数表、随机数骰子或计算机产生的随机数进行抽样。

二. 分层抽样

主要特征分层按比例抽样，主要使用于总体中的个体有明显差异。共同点：每个个体被抽到的概率都相等 N/M 。

报名号 #1869

定义：

一般地，在抽样时，将总体分成互不交叉的层，然后按照一定的比例，从各层独立地抽取一定数量的个体，将各层取出的个体合在一起作为样本，这种抽样方法是一种分层抽样。

三. 系统抽样

当总体中的个体数较多时，采用简单随机抽样显得较为费事。这时，可将总体分成均衡的几个部分，然后按照预先定出的规则，从每一部分抽取一个个体，得到所需要的样本，这种抽样叫做系统抽样。

步骤：

一般地，假设要从容量为 N 的总体中抽取容量为 n 的样本，我们可以按下列步骤进行系统抽样：

(1) 先将总体的 N 个个体编号。有时可直接利用个体自身所带的号码，如学号、准考证号、门牌号等；

(2) 确定分段间隔 k ，对编号进行分段。当 N/n (n 是样本容量) 是整数时，取 $k=N/n$ ；

(3) 在第一段用简单随机抽样确定第一个个体编号 l ($1 \leq l \leq k$)；

(4) 按照一定的规则抽取样本。通常是将 l 加上间隔 k 得到第 2 个个体编号 $(l+k)$ ，再加 k 得到第 3 个个体编号 $(l+2k)$ ，依次进行下去，直到获取整个样本。

5.1.2 具体抽样应用

对所有数据一共抽取三个容量分别为 300 的样本。应用的抽样方法有随机抽样和分层抽样。

第一个样本的抽取方法为：所有数据按照条码号进行升序排列，抽取前 300 个数据组成一个样本，称为总体样本。

第二个样本的抽取方法为：将所有老师超期的数据按照条码号进行升序排列，抽取前 300 个数据组成一个样本，称为老师样本。

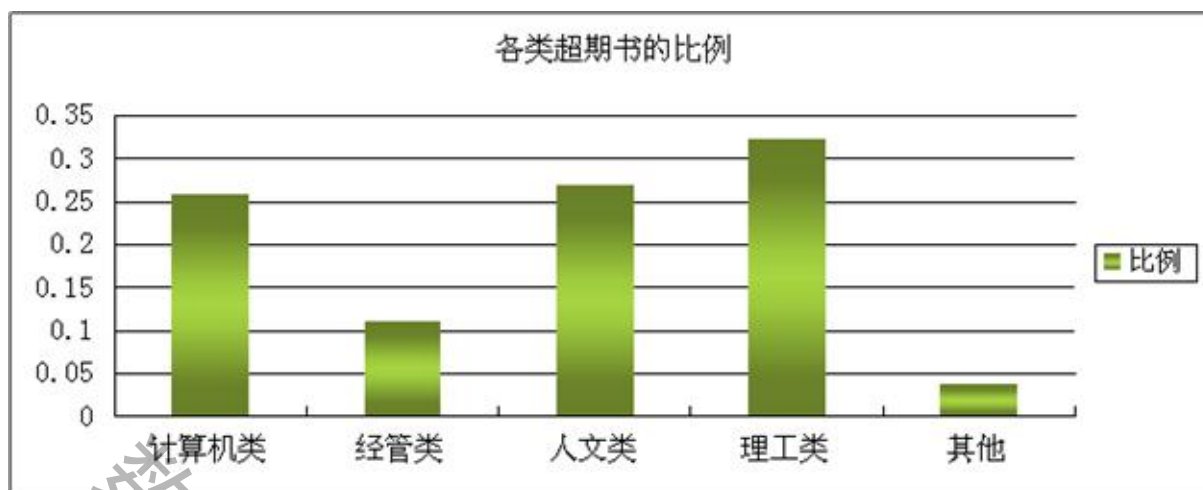
第三个样本的抽取方法为：将所有学生超期的数据按照条码号进行升序排列，抽取前 300 个数据组成一个样本，称为学生样本。

5.1.2.1 总体样本的统计结果

从所有借的书中抽取300个样本的统计结果					
	计算机类	经管类	人文类	理工类	其他
数目	78	33	81	97	11
比例	0.26	0.11	0.27	0.323	0.037

从表中的统计结果来看所有超期的书中计算机类，人文类和理工类的书都有七八十本，三者相加后占样本总体的比例为 85.3%。说明这几类书容易发生超期现象。

报名号 #1869

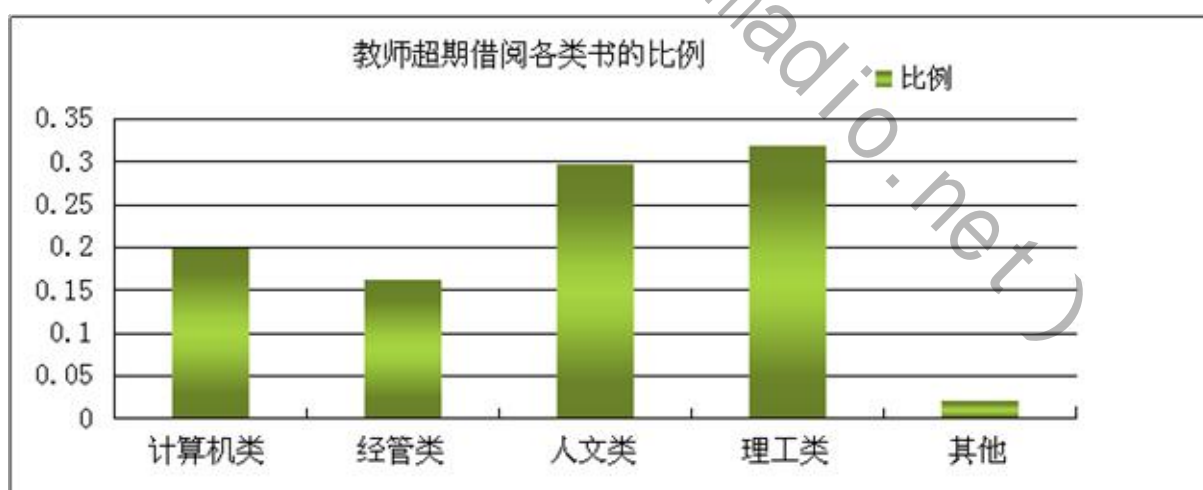


从直方图中也可以直观地看出计算机类，人文类和理工类的书容易发生超期现象。

5.1.2.2 教师样本的统计结果

从教师借的书中抽取300个样本的统计结果					
	计算机类	经管类	人文类	理工类	其他
数目	60	49	89	96	6
比例	0.2	0.163	0.297	0.32	0.02

从表中的统计结果来看被教师借的超期的书中人文类和理工类的书都有八九十本，二者相加后占样本总体的比例为 61.7%。说明这两类书容易发生超期现象。



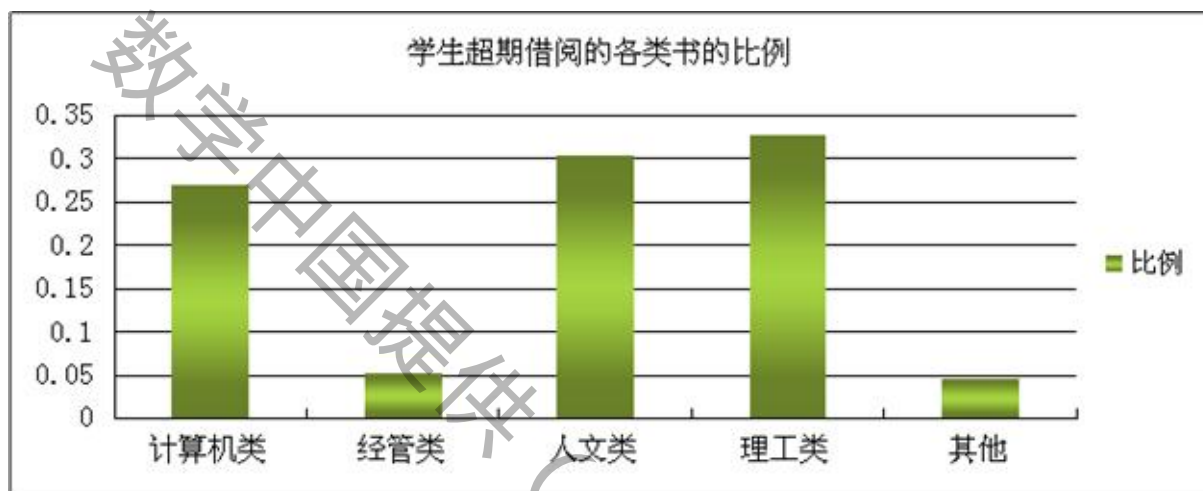
从直方图中也可以直观地看出人文类和理工类的书容易发生超期现象。

5.1.2.3 学生样本的统计结果

报名号 #1869

从学生借的书中抽取300个样本的统计结果					
	计算机类	经管类	人文类	理工类	其他
数目	81	16	91	98	14
比例	0.27	0.053	0.303	0.327	0.047

从表中的统计结果来看被老师借的超期的书中人文类和理工类的书都在九十本以上，二者相加后占样本总体的比例为 63%。说明这两类书容易发生超期现象



从直方图中也可以直观地看出人文类和理工类的书容易发生超期现象

综上所述可知超期的书中人文类和理工类的书占了 60%以上。也就是说人文类和理工类的书容易发生超期现象。原因可能是：大部分学生比较喜欢看人文类的书来消遣，而借理工类的书是课程学习的需要；另外，大部分的老师都要从事科研或教学工作，所以借理工类的书也叫多。

5.1.3 这类书籍容易发生超期现象的原因

造成读者借阅超期的原因主要有读者故意超期、过失超期和其他原因。

(一)故意超期

所谓故意超期,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)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,这类书主要是理工类,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,如有些人文类的,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。

(二)过失超期

与故意超期不同,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

报名号 #1869

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。

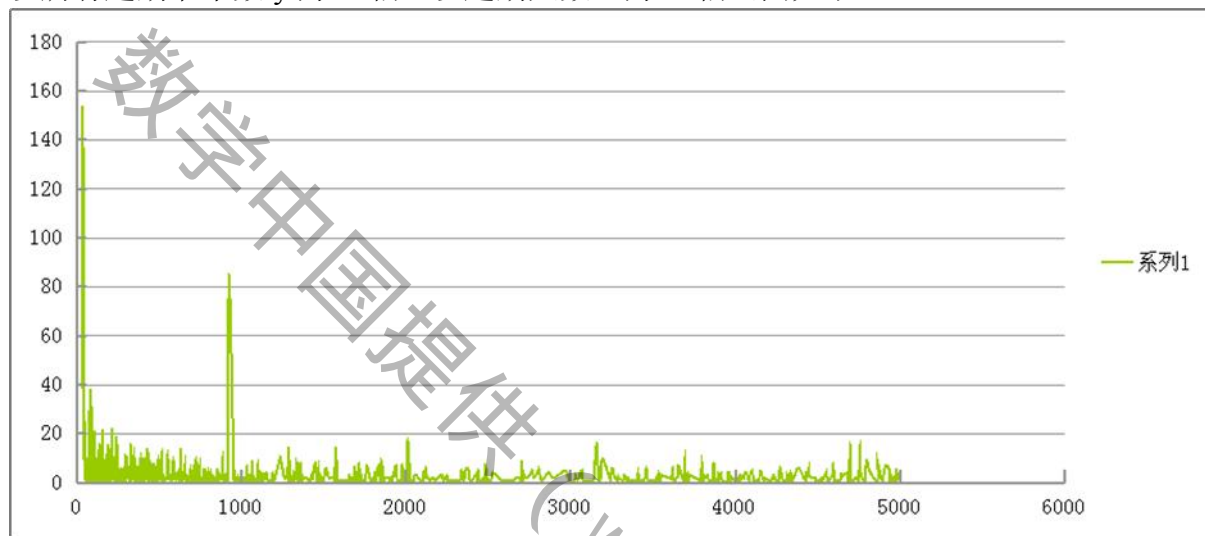
(三)其他原因

其他原因是指读者既非故意,又非过失而导致超期的情形,主要指出现意外事件(如所借的图书被盗、遗失等)后读者未及时到图书馆进行处理而导致超期,造成赔书又超期的双重损失。出现这种情况,只要处理得当,超期罚款原本是可以避免的。

5.2 模型建立

5.2.1 基于总体数据模型的建立

以所有超期书本数 y 为 Y 轴, 以超期天数 x 为 X 轴画图如下。



图中在开始的一段时间内超期书本数 y 在超期天数 x 的增加下降剧烈,这段时间的变化主要是无意造成的;之后下降的趋势较缓直到波动不大一直到最后,这段时间变化主要是故意造成的。

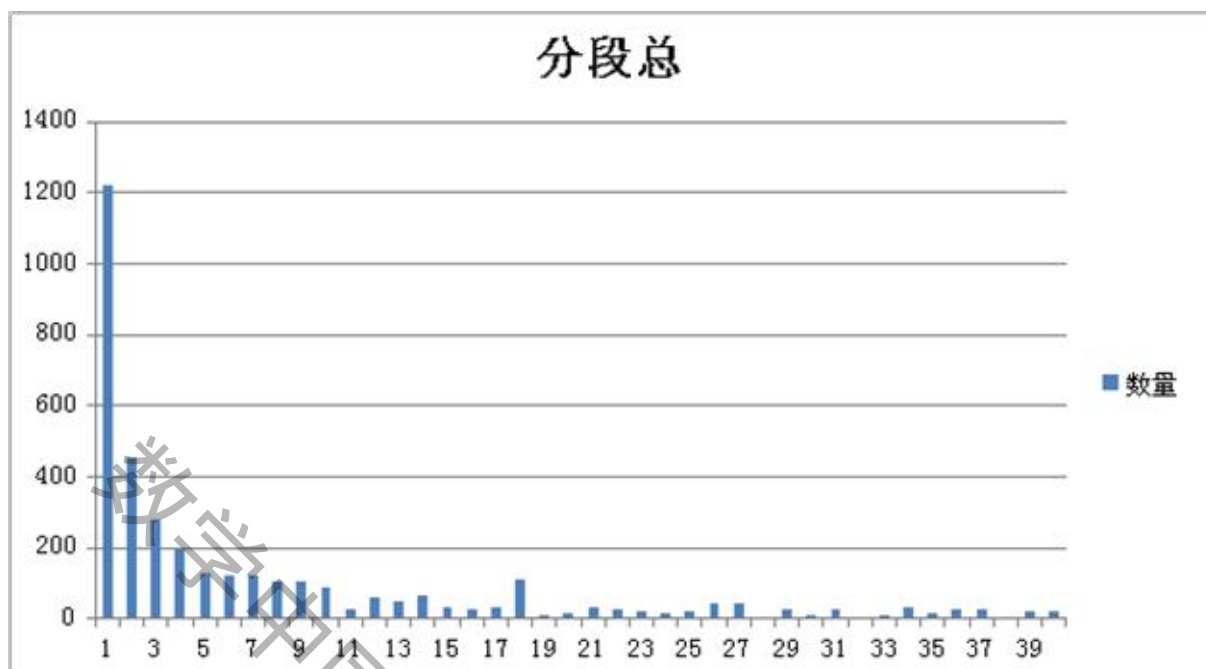
图中在第一个高峰和第二个高峰之间 4 有一个较平缓的段可以看做线段,两端点 x_1, x_2 之间的超期天数相差 28 天,在 $x_3=516$ 和 $x_4=551$ 之间也有一个类似的线段, $x_4-x_3=35$ 。之后还有很多段这样的线段。所以可以认为这段时间内图书馆因为特殊原因而闭馆,可能是寒假,暑假或者其他原因。所以可以将这些线段去掉而将后面的曲线往前平移,从而使曲线更平滑,从而能从中找出规律。

考虑到图书馆具体某天借出图书超期的数量随机性较大,规律性不明显,所以将超期的时间分为若干时间段。在各时间段内,超期图书的数量规律较明显。通过试差法,得到当如下分段方法:

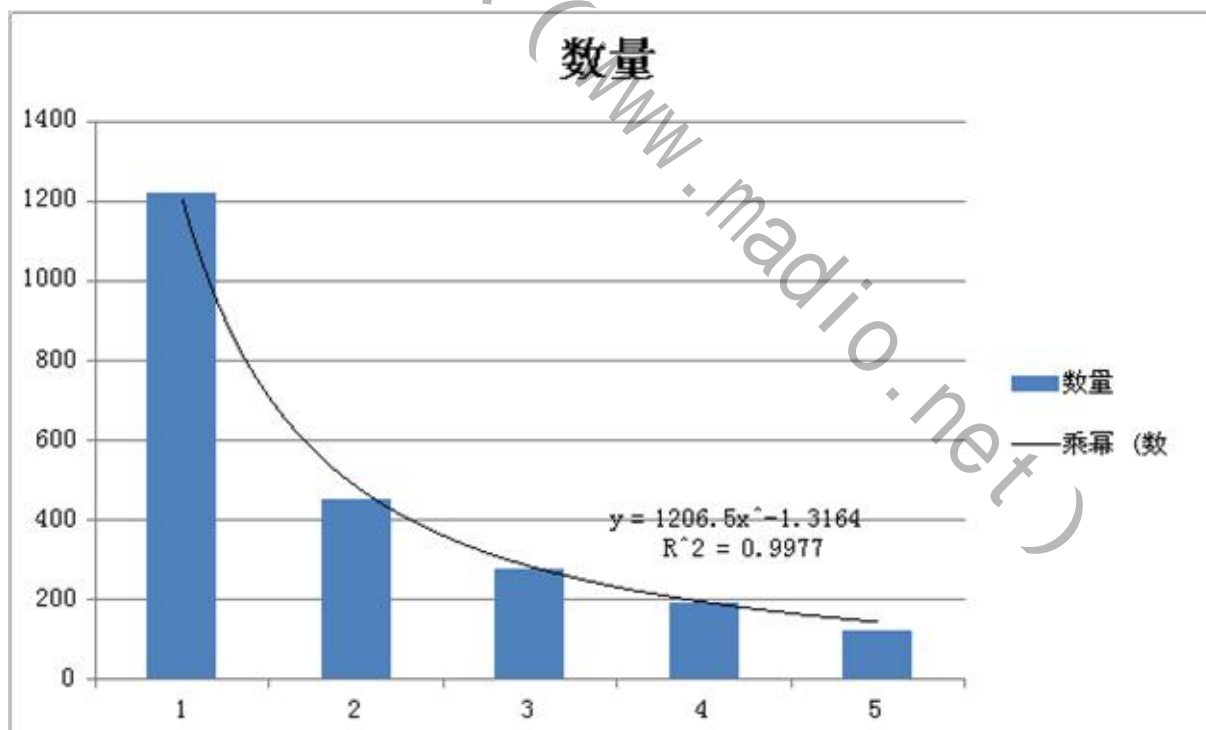
- 1 对总数据及学生数据的分析中,取 50 天为一段最合理。
- 2 对教师数据的分析中,取 100 天为一段最合理

下面是简化后做的图。简化方法为采用第一种简化方法。

报名号 #1869



以 200 天为界限对上图进行分段,即将 $x_1=4$ 作为分界点, $x_1=4$ 之前的一段的图线如下



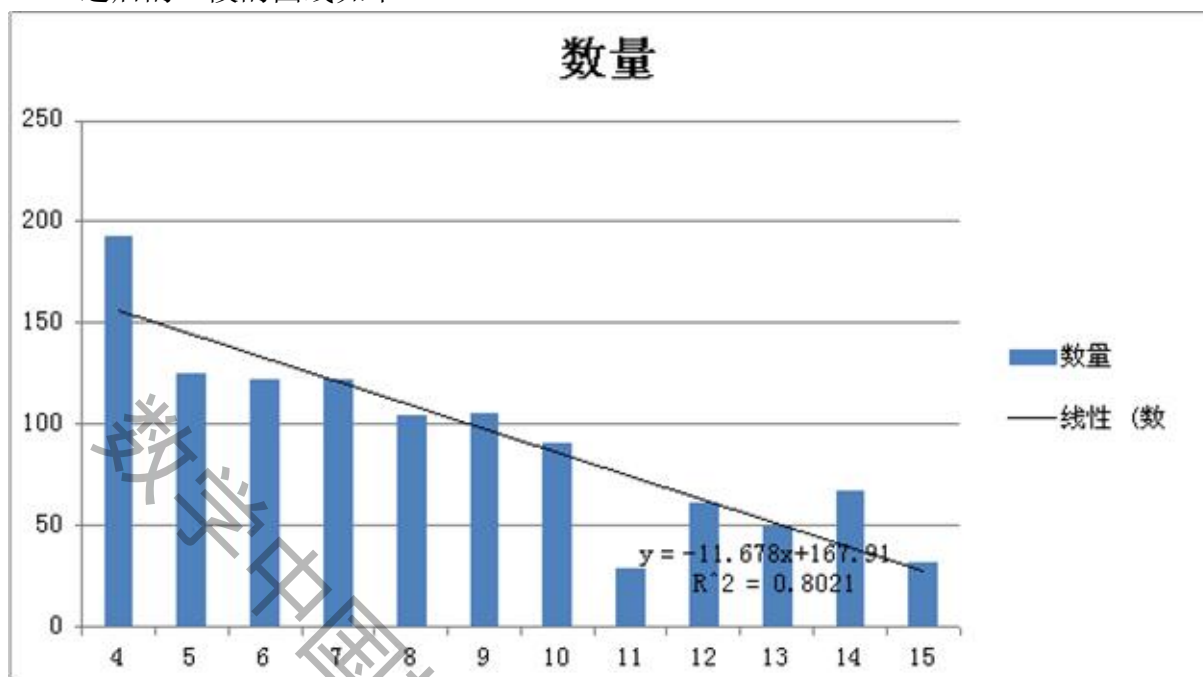
图中在超期书本数 y 在超期天数 x 的增加下降剧烈,这段时间的变化主要是无意造成的。

对简化后的直方图进行拟合后就得到这段时间内所有超期书本数 y 与超期天数 x_1 的数学模型为

$$\text{当 } 1 \leq x_1 \leq 4 \text{ 时, } y = 1219.2x_1^{-1.377}, R^2 = 0.9965$$

报名号 #1869

x1=4 之后的一段图线如下



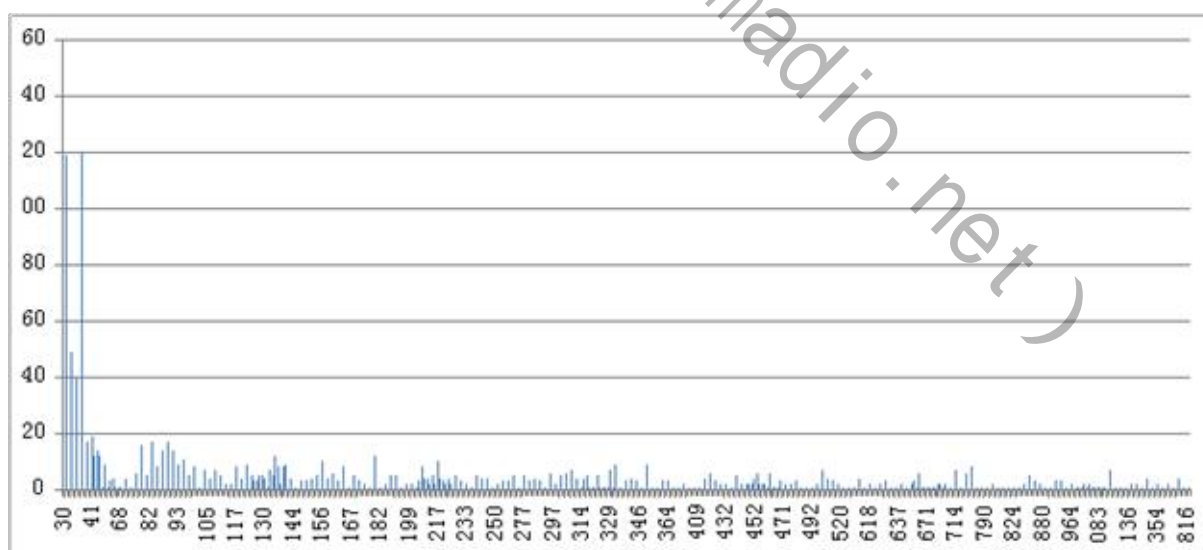
图中在超期书本数 y 随着超期天数 x 的增加而下降的趋势较缓而且斜率较小，这段时间变化主要是故意造成的

拟合后就得到这段时间内所有超期书本数 y 与超期天数 x_1 的数学模型为

$$\text{当 } 4 < x_1 \leq 15 \text{ 时, } y = -11.678x_1 + 167.91, R^2 = 0.8021$$

5.2.2 基于学生的衰减模型的建立

数据简化前超期书本数 y 与超期天数 x 的关系用下图表示，横轴为 x , 纵轴为 y



图中在开始的一段时间内超期书本数 y 随着超期天数 x 的增加下降剧烈，这段时间的变化主要是无意造成的；之后下降的趋势较缓直到波动不大一直到最后，这段时间变化主要是故意造成的。

报名号 #1869

图中：在 $x_1=48$ 和 $x_2=76$ 之间有一段水平的线段， $x_2-x_1=28$

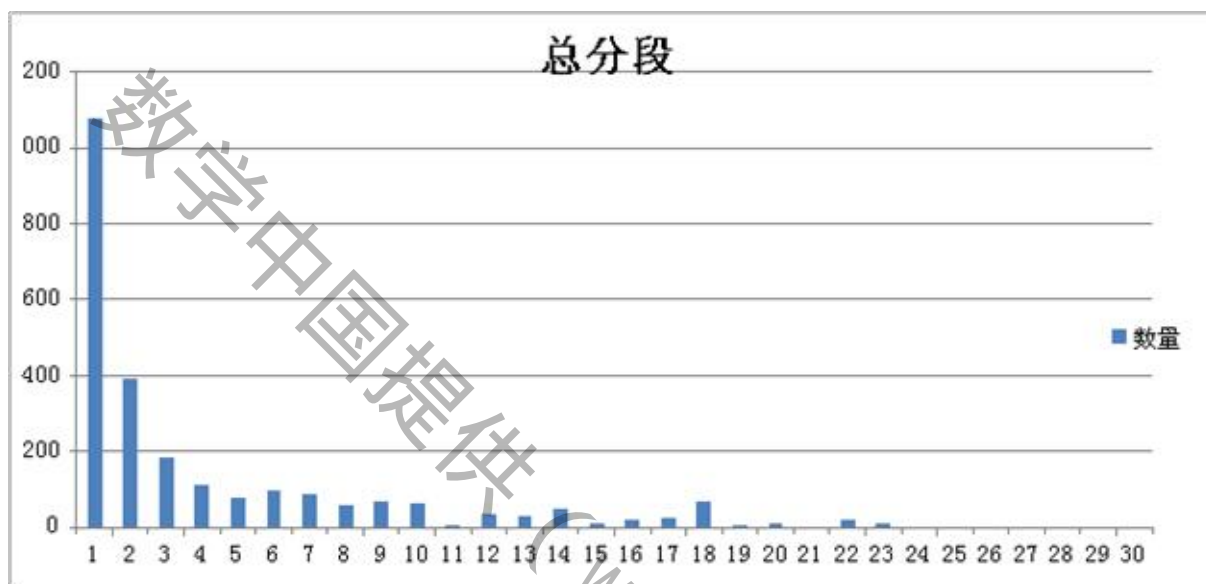
在 $x_2=76$ 和 $x_3=369$ 之间， $x_4-x_3=293$

在 $x_3=369$ 和 $x_4=408$ 之间有一段水平的线段， $x_4-x_3=29$

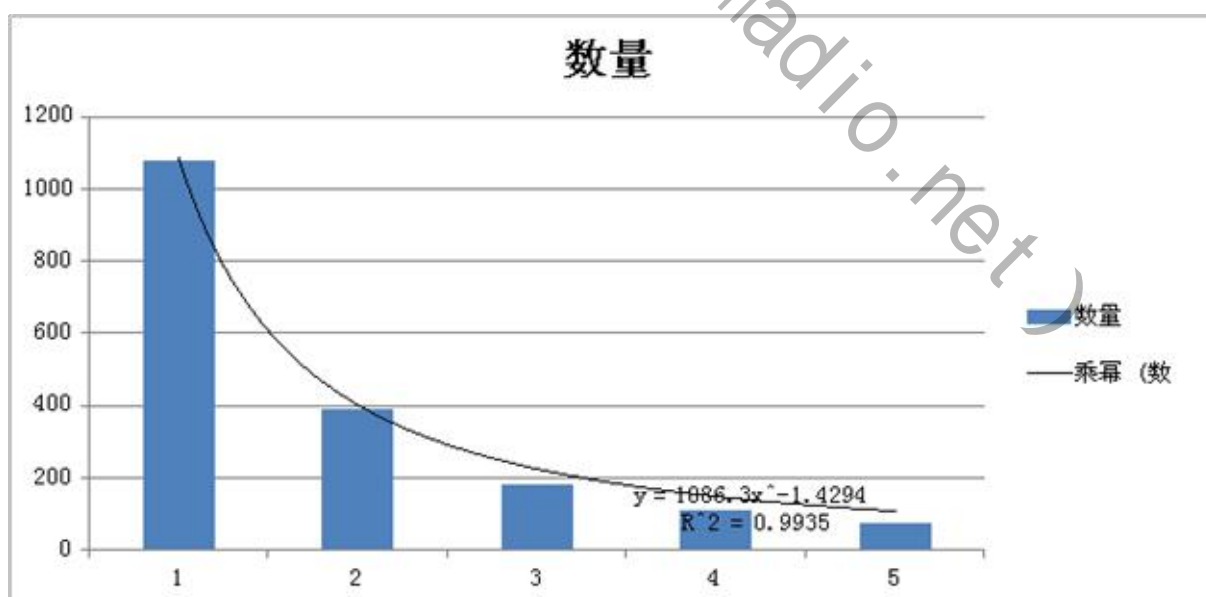
在 $x_4=408$ 和 $x_5=521$ 之间， $x_5-x_4=29$

在 $x_5=521$ 和 $x_6=577$ 之间有一段水平的线段， $x_6-x_5=56$

所以可以认为这线段时间内图书馆因为特殊原因而闭馆，可能是寒假，暑假或者其他原因。所以可以将这些线段去掉而将后面的曲线往前平移，从而使曲线更平滑，从而能从中找出规律。下面是简化后做的直方图。期间应用第一种简化方法



以 200 天为界限对上图进行分段,即将 $x_1=4$ 作为分界点, $x_1=4$ 之前的一段的图线如下



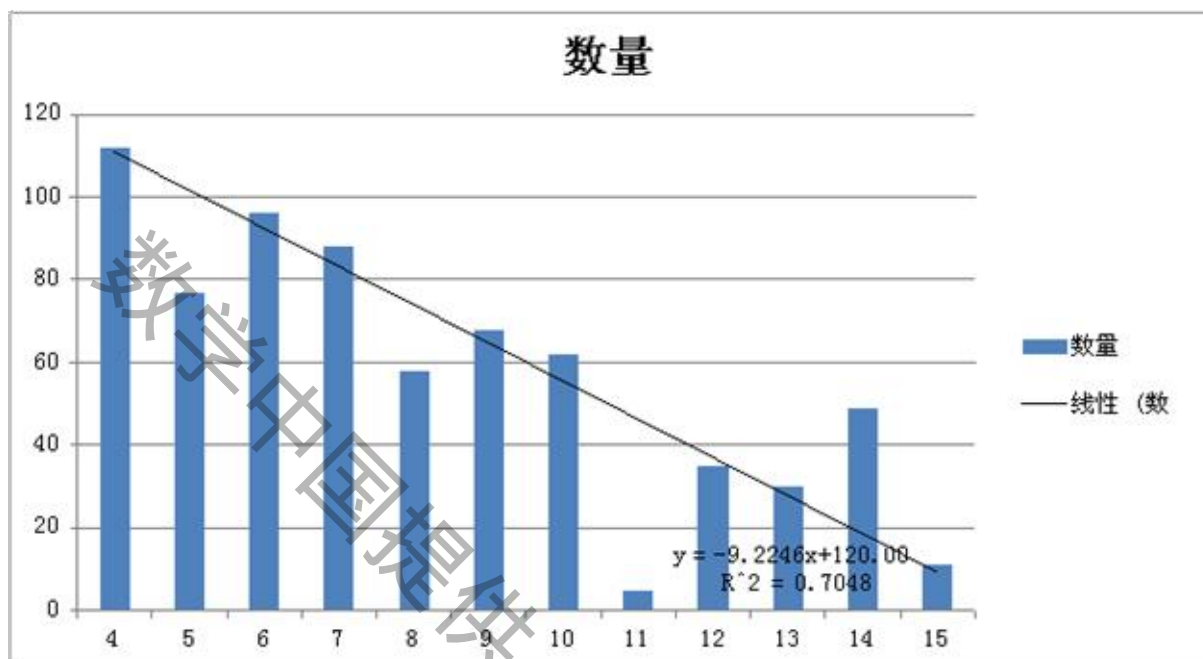
图中超期书本数 y 随着超期天数 x 的增加下降剧烈,这段时间内的变化主要是无意造成的。无意超期的书随超期天数的变化, y 值开始很大,然后随着超期天数 x_1 的增大, y 值急剧下降,就是无意超期的书很快被还,直到基本还清。

报名号 #1869

拟合后就得到这段时间内所有超期书本数 y 与超期天数 x_1 的数学模型为

当 $1 \leq x_1 \leq 4$ 时, $y = 1129.7x_1^{-1.655}$, $R^2 = 0.9974$

$x_1=4$ 之后的一段图线如下



图中超期书本数 y 随着超期天数 x_1 的增加下降较缓, 曲线接近于一条直线, 其斜率较小。这段时间内的变化主要是故意造成的。故意超期的书随超期天数的变化, y 值开始不大, 然后随着超期天数 x 的增大, y 值缓慢下降。图中还包括一段突变段, 可以理解为当 x 等于某个值时, 大量故意超期的书返还。

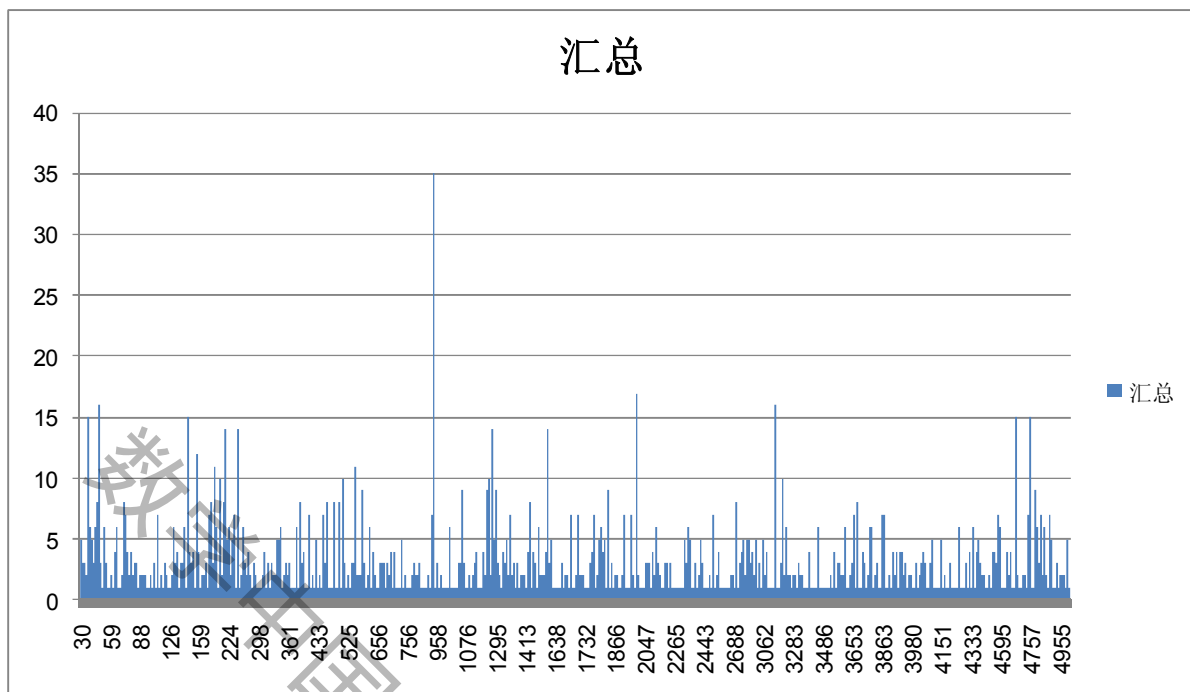
拟合后就得到这段时间内超期书本数 y 与超期天数 x_1 的数学模型为

当 $4 < x_1 \leq 15$ 时, $y = -9.2246x_1 + 120$, $R^2 = 0.7048$

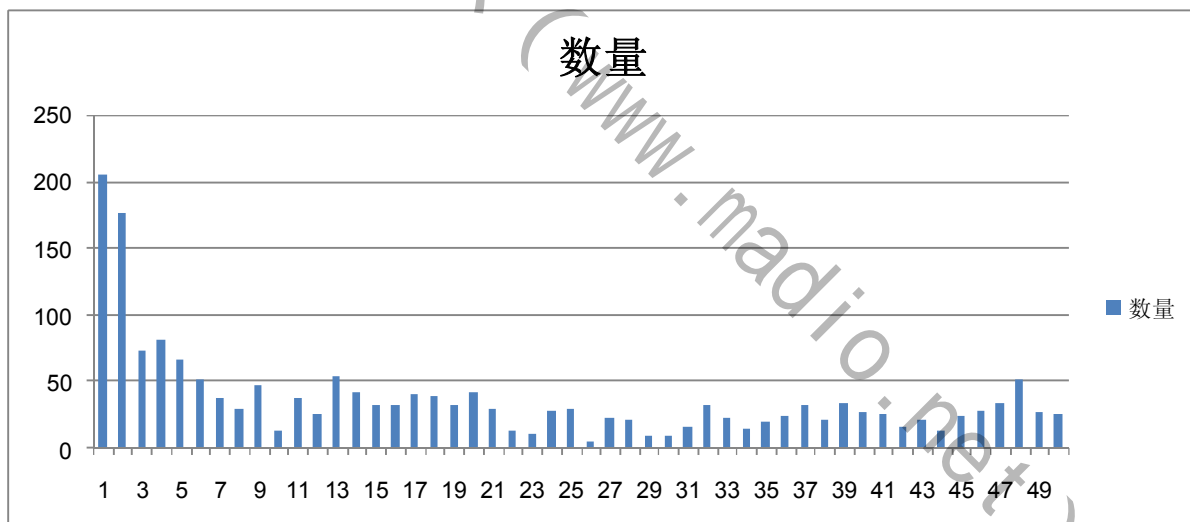
5.2.3 基于教师的衰减平稳混合模型的建立

数据简化前超期书本数 y 与超期天数 x 的关系用下图表示, 横轴为 x , 纵轴为 y

报名号 #1869

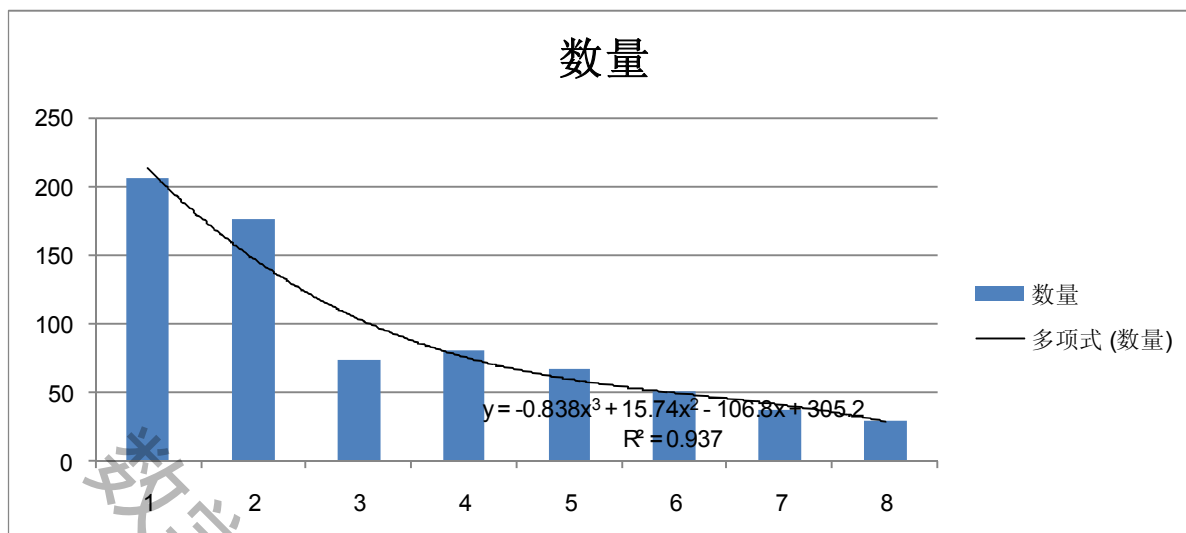


图中在开始的一段时间内超期书本数 y 随超期天数 x 的增加下降较缓, 之后都在某一个定值附近波动而且持续的时间较长。按照第二种简化方法简化后总分段图线如下



以 800 天为界限对上图进行分段,即将 $x_2=8$ 作为分界点, $x_1=8$ 之前的一段的图线如下

报名号 #1869

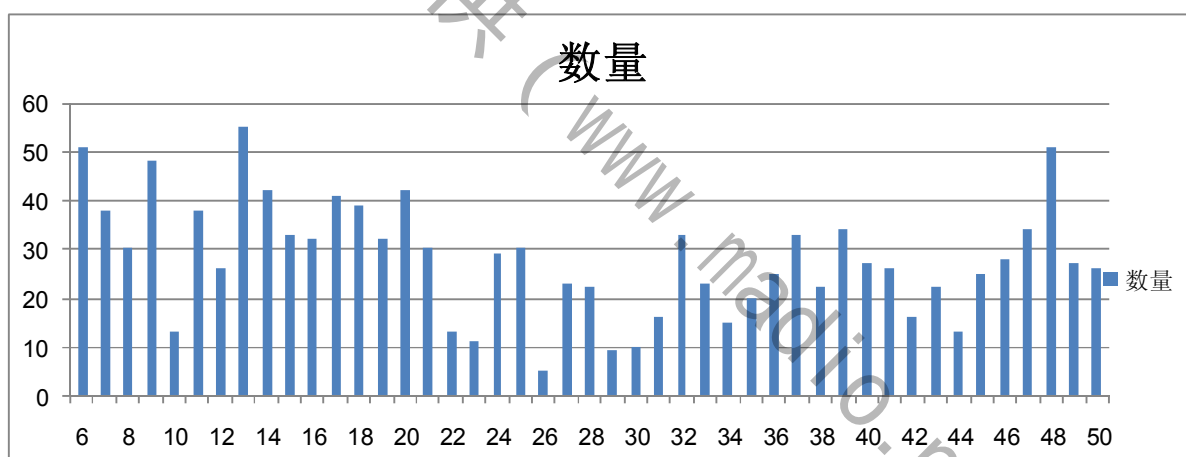


图中超期书本数 y 随超期天数 x 的增加下降较缓直到平缓。引起这段时间的变化的原因有两种：无意和科研或教学。

拟合后就得到这段时间内超期书本数 y 与超期天数 x_2 的数学模型为

$$\text{当 } 1 \leq x_2 \leq 8 \text{ 时, } y = -0.8384x_2^3 + 15.747x_2^2 - 106.82x_2 + 305.21, R^2 = 0.9373$$

$x_1=8$ 之后的一段图线如下



图中超期书本数 y 随超期天数 x 的增加在某一个定值附近波动而且持续的时间较长。引起这段时间的变化的原因只有一个：科研或教学。

6 问题二借书规则的优劣评价与新的借书规则的建立

6.1 该图书馆现行借书规则的优劣评价

题目所给出高校的借书规则为：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。

我们根据第一问建立的抽样比例模型、学生的衰减模型以及教师的衰减平稳混合模型对该规则进行分析，其优缺点主要如下：

优点：

1. 借书时限为 1 个月，对于学生来说比较合理，由于大多数的学生借书期限较短，在

报名号 #1869

短期能就可以将借的书还了，所以借书时限定为一个月还可以；

2. 可以续借图书，那么将会在一定程度上减少部分因为阅读时间不够并且不顾忌较少罚款的故意拖延的人数；

3. 超期一天罚款 0.1 元，考虑到学生的身份来说也比较合理，同时也照顾了大量因为无意借阅超期的人。

缺点：

1. 借书时限都为 1 个月，对于教师来说不是很合理，大部分教师有科学研究和教学的需要，同时科研和教学又都是长期的，因而借书时限太短对教师的科研和教学的影响会比较大，所以对于教师来说借书时限应该加长；

2. 超期一天罚款 0.1 元，没有针对性，对于那些故意长时间拖延不还的人没有起到任何警示作用，所以该图书馆借书规则还是不合理的，应该对图书借阅超期较长的进行较重的罚款；

3. 罚款金额没有上限，那也就是当很大程度上超过读者的财力接受能力时，绝大部分读者也就不会归还图书，那么该图书借阅规则也就变成鼓励读者不归还图书的规则，因此必须存在罚款的上限值，也就是要考虑到大家的财力平均接受能力，对于性质较为恶劣的，应该停止其以后的借阅，直到其缴清上限罚款以及将超期严重的书籍归还后方可允许其借阅。

4. 没有考虑教师和学生财力平均接受能力的差异，罚款还应该考虑到教师和学生财力平均接受能力的差异，从而分别对教师和学生设计相应的图书借阅规则；

5. 该规则本馆数量较少的书和本馆数量较多的书没有区别对待，馆藏较少的书可以不外借，或者相对少的书应该加重超期后每天的罚款金额。

6.2 新的借书规则的设计

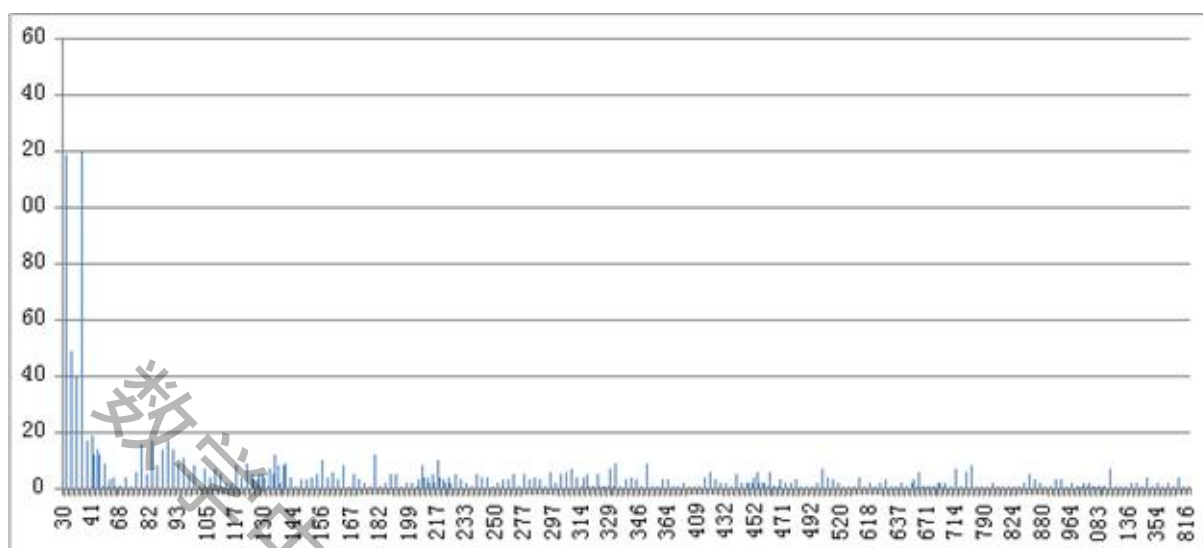
学生和教师的借书用途和财力接受能力的差异较大，所以我们在设计借书规则的时候应该将两者区分对待

6.2.1 基于学生的借书规则的设计

对学生来说，应用学生超期借阅书的数量与超期时间的衰减模型。大量学生借阅的书超期是因为无意，少量是因为故意。衰减模型就是：大量的学生借书之后开始陆续地还书，在开始的一段时间内，图书馆将催还单贴出去后，那些因为无意而借书超期的学生会知道自己还有书没换，于是赶紧去还书，进而导致在这段时间内还的书迅速增加，模型的曲线很陡，斜率很大。在某个时刻上，因为无意而被借超期的书基本还清，只有少量的没换，可能是这些学生没注意到催还单。但这些少量的书相对于因为故意而被借超期的书来说又太少，故可以忽略不计。也就是说，这一点是区分有意和无意的分界点，在这点之前超期借阅的书是因为无意造成的。在这点之后是因为故意造成的，可以将这点成为原因点，也是模型曲线上的第一个转折点。之后，随着时间的推移，又有因为无意而被借超期的书陆续归还，模型曲线慢慢下降直到变平缓。又在某一时刻，超期天数过大，导致需要交付的罚款数目很大，超过了学生的承受极限，这天有大量因为故意而被借超期的书陆续归还了，不过还剩极少数故意借书不还的人，存在着侥幸心理，以为可以一直占有。但是，学生在学校最多只能呆 4 年（不包括研究生），在毕业的时候还是得把书还了，在模型曲线上就表现为所有的书都还清。

报名号 #1869

学生衰减模型超期书本数 y 与超期天数 x 的关系用下图表示，横轴为 x ，纵轴为 y



图中在开始的一段时间内超期书本数 y 随着超期天数 x 的增加下降剧烈，这段时间的变化主要是无意造成的；之后下降的趋势较缓直到波动不大一直到最后，这段时间变化主要是故意造成的。

通过数据模型可以大致确定图中第一个转折点在超期 90 天左右出现，所以可以确定因为无意导致图书借阅超期的大致最大天数为超期 90 天，而学生能够承受的罚款金额的上限大致是在超期 920 天左右呈现，该点周围的图书借阅超期数量都很小，所以可以取定超期 920 天对应的罚款金额为图书借阅超期罚款上限值，如果假设超期一天罚款 0.1 元（经查阅大部分图书馆的超期一天罚款 0.1 元），那么超期 920 天所要的罚款金额上限值为 92 元，于是我们根据上面得到的数据以及分析出的优缺点，设计如下基于学生借书规则：

借书时限为 1 个月，可以续借一次，续借时限为 30 天，超期一天罚款 0.1 元，当其超过超期 60 天之后超期一天罚款为 0.2 元，罚款上限值为 92 元，当该读者所要罚款的金额超过 92 元时，停止其借书的权利，直至其付清罚款并归还借阅超期的图书。

6.2.2 基于教师的借书规则的设计

对老师来说，应用老师的衰减平稳混合模型。大部分老师都要从事教学或科研，而教学和科研持续的时间都很长。衰减平稳混合模型就是：大量的书被老师借走之后开始陆续返还，在开始的一段时间类图书馆将催还单贴出去后，那些因为无意而借书超期的老师会知道自己还有书没还，于是去还书，进而使得在这段时间内还的书较多，反应在模型上就是曲线较陡，斜率较大。在某个时刻上，因为无意而被借超期的书基本还清，只有少量的没换，可能是这些老师没注意到催还单。但这些少量的书相对于因为科研或教学而被借超期的书来说又太少，故可以忽略不计。也就是说，这一点是区分无意和科研原因的分界点，在这点之前超期借阅的书是因为无意和科研或教学造成的。在这点之后是因为科研或教学造成的，同样可以将这点称为原因点，也就是模型曲线上的转折点。之后这些被超期借阅的用于科研或教学的书一直没归还，并且持续很长时间，3 年乃至 10 年，反映在模型曲线上就是曲线就是几乎趋近于一条平行于 x 轴的直线。

根据衰减平稳混合模型模型，将对借书超期老师的处罚分为两类：对于因为无意而超期的老师可以罚款较轻一点，对于因为科研或教学而超期的老师可以罚款较重一点。可以以模型曲线中转折点所对应的超期天数作为标准。由于大多数的老师借书期限很

报名号 #1869

长，在短期不可能能就可以将大部分借的书还了，所以借书时限定为一个月就不合理了，故对于老师的借书时限可以定的长一些，而罚款也要分段操作。不过要减少这类情况还是要学校对老师的借阅进行更具体的约束才行。我们得到的教师数据反映的情况不太明显，但我们根据学生借阅超期平均天数（230 天）和教师借阅超期平均天数（1829 天）、上面学生的借阅规则以及分析出原规则的优缺点，设计出设计如下基于教师的借书规则：

借书时限为 8 个月，超期一天罚款 0.4 元，罚款上限值为 400 元，当该读者所要罚款的金额超过 400 元时，停止其借书的权利，直至其付清罚款并归还借阅超期的图书。

7 给图书馆馆长的一封信

图书馆是搜集、整理、收藏图书资料供人阅览、参考的机构，有保存人类文化遗产、开发信息资源、参与社会教育等职能。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。

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。超期罚款政策作为的一项管理手段得到了世界图书馆界的普遍采用，并且在长期的实践中被证明是行之有效的。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。

根据实际情况以及对模型进行分析，为了解决高校图书馆存在的借阅超期现象，我们提出以下建议：

一，加大宣传，开展读者教育，让读者了解图书馆的规章制度。

要尽量减少超期现象的发生，首要的是加强读者的教育。针对高校特点，要充分利用新生入学教育的机会扎实开展入馆教育，并将超期罚款制度的学习作为读者入馆教育的重点内容。不仅强调超期后如何处罚，更要说明为什么要进行处罚，讲清道理，提高读者的认识，使读者在走进图书馆之前就对超期罚款制度有较深的认识和印象，为读者较好地遵守该项制度打下坚实的基础。

其次，还要抓好日常宣传教育。图书馆可以在醒目位置张贴规章制度，超期处罚措施等。

另外，对超期的读者按规定处以罚款也是必不可少的教育方式。只有将这些教育方式有机结合起来，系统展开，才能受到较好的效果。

二，制订更加合理的借阅规定，设置弹性的借阅期限。

一方面，细分图书类型。根据图书类型及流通方式的不同制定不同的借阅规定。比如可将文献按类型分为图书，期刊，杂志，工具书等，并规定不同的借阅期限。再比如，就图书新旧来说，新书(出版年限 1 年内)的借阅率最高，可以规定为 15 天，稍旧一点的可以延长到 30 天，再旧一些的(出版年限 5 年以上)可以设档延长到 60 天或更长。

另一方面，细分读者类型。如对教师和学生采取不同的超期罚款政策，目前国内绝大多数高校图书馆并无类似规定。在本文中，可以明显看出，学生与教师的借阅情况及超期现象有较大差别，不能以一个标准对待。对教师，可以合理延长借阅期限，从上文中可以明显看出教师借阅平均时间明显较大。

目前国内大学图书馆超期罚款政策虽然简单易记，但对文献的催还效果较差，不能充分体现超期罚款政策促进文献流通，保证资源共享的目的。复杂细致的借阅制度，的确给管理带来了一些不便，但无疑提高了文献利用率，为读者使用图书馆带来了方便。

三，采取复合的图书催还方式。

报名号 #1869

复合的图书催还方式就是从本馆和读者的实际出发，根据读者特点、本馆特点、图书、资料性质等采取不同的图书催还方式。

(1)根据读者特点采取不同的催还方式。对大学生读者，可以采用电子邮件、QQ群、手机短信、网页发布等催还方式。在超期一定期限后（如本题所给数据为30天），可以通过张贴“黑名单”的方式催还。对于教师，可采用电子邮件，手机短信电话等催还方式，但最好不要登上“黑名单”。故意拖延的毕竟为极少数，所以，通过这些方式，绝大多数超期图书能够及时归还。

(2)根据本馆的特点来选取不同的图书催还方式。如果本馆藏书较多，读者群大，并且图书超期的读者人数多。这种情况下，可以使用网页发布、QQ群、张贴“黑名单”等方式催还，这些方式成本比较低廉。如果是藏书较少，读者人数不多，书刊过期的读者也不多的话，可以采用电话，短信的催还方式，这样最直接、最具针对性。

(3)根据图书资料的性质，选取不同的催还方式。如果读者借阅的是本馆比较珍贵的文献的话，也最好采用电话催还这种最为直接的催还方式，保证尽可能追回本馆珍贵的书刊资料。

当然，图书馆员在实际工作中，可以根据实际情况，选用不同的图书催还方法，也可以使用这些方法的组合，尽可能催回本馆的图书，也尽可能使读者没有抵触情绪。

四、制订更加细致的罚款政策。

超期罚款政策制定的出发点是为了促进文献流通，更好的为读者服务。丰富政策内容，具有策略性的政策设计往往更能得到读者的理解，执行的效果也更好。针对目前国内高校图书馆的一些情况，我们提出以下建议。

(1)适当提高现有超期罚款的标准。目前各图书馆所采用的超期罚款标准大多已执行了几十年，一直未作改变，相对于现在人们的收入水平来说，该标准偏低，不能有效抑制故意借阅超期，特别是短期故意超期现象的发生，对此，适当提高超期罚款的标准，增加处罚的力度是非常必要的。

(2)罚款应规定上限，并限制其借阅权利。既然图书馆按规定收取超期罚款只是督促读者按期还书的手段而不是目的，是一种警示而不是为了创收，超期罚款就不应无限制地累计计算。高校图书馆应该对超期罚款的上限作出明确规定，同时相应地对这方面违章的读者的外借权利进行限制，以更有效地敦促读者按期还书。

(3)超期一定程度转为丢失赔偿。在本题所给数据中，超期时间大于一年（365天）的占43.12%，甚至存在超期4990（约13.67年）的。这部分图书回收的概率很小，可以认定为丢失。图书馆可以指定合理的丢失赔偿方案。

(4)公布超期罚款的收支。超期罚款的收入应专款专用，并且每隔一段时间应将罚款的收支情况公布出来，这是消除读者对图书馆误解的最好方法。罚款收入可用于购买新书和用于对读者的奖励。

五、建立健全诚信机制。

图书馆对读者的各种违规行为设定相应的分值，通过对读者的借阅行为进行记录，一旦发现读者有违规违纪行为就减去相应的分数，将读者的诚信划分不同的等级。当读者的诚信值低于某一等级时，图书馆就要对读者借阅权限进行限制；当读者的诚信值高于某一等级时，图书馆可适当放宽借阅权限。另外，图书馆应经常对读者进行道德修养教育，让读者树立起以珍爱图书为荣，以偷书污损图书为耻的观念。图书馆也可以通过宣传栏展出上述借阅中的不良行为，以及在网上谈论以上话题等方式，这样能取得比简单地罚款更好的效果。

报名号 #1869

8 模型的进一步讨论

本文主要解决了如下三个问题：

(1)图书借阅超期中容易超期的图书的种类问题，即理工类，计算机类和人文类比较容易超期。分析其产生这种现象的原因可能是理工类和人文类的分类比较模糊或者是涵盖的面较广，所包含的种类过多，使得图书的数量也较多，从而造成这两类图书被借阅的数量就较多，如果按概率计算，这两类书容易超期也不足为怪。对于计算机类，虽然看似分类较窄，但如果要细分，其包含的种类也就多，图书的数量也较大，并且计算机是一个新兴的行业，对其感兴趣的人数不在少数，因而使得计算机类被借阅的数量也较多，从而产生计算机类容易超期。根据上面的分析得到的启示，我们可以将图书分类进行细分，从而得到更为理想的结果，从而正确指导图书馆的管理。但无论怎么细分，都会存在划分不清的问题，不过可以得到图书馆管理中更需要的结果，那就是更准确的区分容易超期的书，这样也更有利于图书馆的图书管理。

(2)图书借阅超期时，可能存在的原因，也就是学生中发生图书短期借阅超期的可能原因主要是无意借阅超期，较长期借阅超期的可能原因主要是故意拖延超期；而对于教师则是短期借阅超期的可能是无意借阅超期，长期则是因为科研教学需要。实际上原因的分类不够详细，但大体涵盖了较为普遍的原因。给学生借阅图书超期的现象进行划分可能的两个原因时，不考虑学生参与科研或教学这种假设也应该因学校而已，也就是我们只是考虑了大部分高校的实际情况，情况具有普遍意义，对于高校进行整体分析有一定价值，但对于某一个学生参与科研或教学较多的学校来说，其分析作用已经不太明显。所以如果对于某个高校的情况进行分析，那需要学生参与科研或教学的大致数据，这样才能全面分析学生借阅图书超期的原因。同样地对于教师亦是如此，科研与教学需求的原因和故意拖延超期的原因两者难以界定，所以也较难处理。

(3)评判该高校借书规则的优劣以及设计更为合理的借阅规则。该高校的借书规则为：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。可以续借图书，那么将会在一定程度上减少部分因为阅读时间不够并且不顾忌较少罚款的故意拖延的人数；超期一天罚款0.1元，考虑到学生的身份来说也比较合理，同时也照顾了大量因为无意借阅超期的人，但是对于那些故意长时间拖延不还的人没有起到任何警示作用，所以还是不合理的，应该对图书借阅超期较长的进行较重的罚款，但是罚款金额存在一个上限值，也就是要考虑到大家的财力平均接受能力。对于性质较为恶劣的，应该停止其以后的借阅，直到其缴清上限罚款以及将超期严重的书籍归还后方可允许其借阅。同样的，15天的续借时间是否合理，和需要进一步进行分析。所以综上，在借阅规则设计方面要考虑较多的因素，并且如大家的财力平均接受能力是很难确定的值，所以借阅规则的好坏最重要的还是与人相关，如果人人借阅图书都不超期，那也就没有设计图书借阅罚款规则的必要。

报名号 #1869

参考文献

- [1] 赵静等, 数学建模与数学实验, 北京: 高等教育出版社, 2000。
- [2] 胡福文, 高校图书馆读者借阅超期有关问题的思考, 管理研究, 2009 (2): 158-159。
- [3] 戴亚能, 谈高校图书馆大学生不良阅读行为产生的原因及对策, 中国科教创新导刊, 2008 (10): 252-253。
- [4] 曾伟清. 以人为本——高校图书馆的服务与管理理念, 图书馆论坛, 2004 (5): 54-56。
- [5] 唐云富, 高校图书馆读者不良阅读行为的心理分析及教育对策, 河南图书馆学刊, 2005, 25 (2): 61-63。
- [6] 韩宇、朱伟丽, 中美著名大学图书馆“超期罚款”政策比较及思考, 图书情报工作, 2008, 52 (2): 129-131。
- [7] 周迪、宋登汉, 关于公共图书馆“制度性收费”的调查分析, 图书馆杂志, 2006 (2): 41-43。